PEDAGOGIKA

MIROSLAV CHRÁSKA

Metody pedagogického výzkumu

ZÁKLADY KVANTITATIVNÍHO VÝZKUMU
2., AKTUALIZOVANÉ VYDÁNÍ

GRADA

PEDAGOGIKA

GRADA®

PEDAGOGIKA

MIROSLAV CHRÁSKA

Metody pedagogického výzkumu

ZÁKLADY KVANTITATIVNÍHO VÝZKUMU

2., AKTUALIZOVANÉ VYDÁNÍ

prof. PhDr. Miroslav Chráska, CSc.

METODY PEDAGOGICKÉHO VÝZKUMU
Základy kvantitativního výzkumu
2., aktualizované vydání

Vydala Grada Publishing, a.s.
U Průhonu 22, 170 00 Praha 7
tel.: +420 234 264 401, fax: +420 234 264 400
www.grada.cz
jako svou 6237. publikaci

Odpovědná redaktorka PhDr. Alena Palčová
Sazba a zlom Milan Vokál
Návrh a zpracování obálky Antonín Plicka
Počet stran 256
Vydání 2., 2016

Vytiskly Tiskárny Havlíčkův Brod, a.s.

ISBN 978-80-271-9225-0 (pdf)
ISBN 978-80-247-5326-3 (print)

OBSAH

ÚVODEM

V pedagogice se v současnosti vedle sebe uplatňují dvě základní paradigmata: paradigma pozitivistické a paradigma postpozitivistické. Těmto paradigmatům odpovídají také dva poměrně rozdílné typy pedagogických výzkumů.

Klasické pedagogické výzkumy vycházejí většinou z pozitivistického paradigmatu. Bývají často označovány jako „vědecké výzkumy" či jako „výzkumy kvantitativně orientované". Vycházejí z přesvědčení, že existuje jedna objektivní realita, která není závislá na našich citech nebo přesvědčení. Předkládaná publikace se zabývá metodologickými problémy zejména tohoto typu klasického (kvantitativně orientovaného) pedagogického výzkumu. Chceme zdůraznit, že pokud v této souvislosti hovoříme o vědeckém výzkumu, nenaznačujeme tím, že ostatní používané metodologie jsou metodami nevědeckými.

V posledních desetiletích se ve světě, ale i u nás, stále více uplatňují výzkumy vycházející z paradigmatu postpozitivistického, které bývají označovány jako „výzkumy kvalitativně orientované". Oba uvedené typy výzkumů mají své silné i slabé stránky a nelze jeden z nich považovat za univerzálně použitelný (ve všech situacích a ke všem účelům). Podle našeho názoru je nejen možné, ale i výhodné ve výzkumné činnosti oba přístupy kombinovat.

Poměrně značná pozornost je v textu věnována statistickým metodám, které se v pedagogických výzkumech využívají. Vzhledem k tomu, že je učební text určen pro studenty pedagogických oborů a pro pedagogy, kteří pedagogický výzkum realizují, bylo snahou autora dosáhnout srozumitelnosti textu i pro ty uživatele, kteří mají jen základní matematické vědomosti a dovednosti.

1. VĚDECKÝ VÝZKUM V PEDAGOGICE

1.1 METODY LIDSKÉHO POZNÁVÁNÍ

Lidské poznávání je velmi složitý proces, který může probíhat nejrozmanitějšími metodami a postupy. Pokud uvažujeme nad tím, co je zdrojem poznání, můžeme mezi bezpočtem individuálních metod a postupů poznávání rozlišit několik základních kategorií. Významný americký filozof Charles Peirce například uvádí (Kerlinger, 1972) čtyři základní metody poznávání:

- **Metoda tradice:** Mnoho z toho, co považujeme za pravdivé, vděčí za svoji existenci právě tradici. Pokud používáme metodu tradice, držíme se svých názorů a „pravd" jen proto, že je jako pravdy znali lidé před námi. Časté opakování takovýchto „pravd" většinou způsobuje zvyšování jejich platnosti v očích lidí. Je zajímavé, že lidé často lpí na tradičních poznatcích, i když mají k dispozici fakta svědčící proti nim.
- **Metoda autority:** U této metody člověk přijímá určité poznatky jako pravdivé jen proto, že je vyslovuje osobnost, která je pro něho autoritou. Na základě této metody je tedy pravdivá ta myšlenka, která má za sebou váhu uznávané osobnosti či veřejného mínění. Metoda autority je ve srovnání s metodou tradice pro lidské poznání důležitější, protože i pomocí ní lze určitého pokroku v poznání dosáhnout. V žádném případě by se však neměla stát ve vědě metodou rozhodující.
- **Metoda a priori:** Kritériem pravdivosti poznání je u této metody „shoda s rozumem" (nikoli tedy shoda se skutečností). Vychází se z názoru, že přirozené sklony táhnou lidi k pravdě. Problematičnost této metody spočívá v tom, že není možno přesně vymezit, co znamená „shodovat se s rozumem". Typickým argumentem při použití této metody je: „vždyť to dá rozum".
- **Metoda vědy:** Pokud lidé při poznávání používají vědeckého přístupu, potom se jejich poznávání od předchozích metod liší zejména tím, že dospívají k novým poznatkům nezávisle na názorech, přáních či postojích badatele. Při správném vědeckém poznávání je činnost vědce natolik kontrolována, že je téměř vyloučeno, aby se uplatnily jeho osobní názory, postoje, emoce apod. Tato velmi důležitá vlastnost vědeckého poznání bývá nejčastěji označována termínem „objektivita".

Poznámka: *Jiní autoři metody lidského poznávání diferencují podrobněji. Například G. Anderson a N. Arsenault (1998) uvádějí pět základních metod poznávání. Tyto metody v podstatě odpovídají metodám shora uvedeným (kromě metody a priori, která je zde podle míry zobecňování členěna na dvě samostatné kategorie).*

1.2 KVANTITATIVNĚ ORIENTOVANÝ PEDAGOGICKÝ VÝZKUM A JEHO HLAVNÍ FÁZE

Filozofickým základem klasických (kvantitativně orientovaných) pedagogických výzkumů je **pozitivismus**, respektive **novopozitivismus**. Tento druh výzkumu má na mysli F. N. Kerlinger (1972), když uvádí: „Vědecký výzkum je systematické, kontrolované, empirické a kritické zkoumání hypotetických výroků o předpokládaných vztazích mezi přirozenými jevy.“ Tato definice je univerzálně použitelná pro vědecký výzkum ve kterékoli oblasti vědy. Pokud hovoříme o kvantitativně orientovaném výzkumu v pedagogice, můžeme jej vymezit jako **záměrnou a systematickou činnost, při které se empirickými metodami zkoumají (ověřují, verifikují, testují) hypotézy o vztazích mezi pedagogickými jevy**.

Jestliže hovoříme o tom, že kvantitativně orientované výzkumy vycházejí z pozitivistické filozofie, považujeme za důležité poznamenat, že část pedagogické veřejnosti chápe pojmy „pozitivismus“ nebo „pozitivistický“ stále ještě (a nezdůvodněně) s jistým pejorativním významovým zabarvením.

Fáze klasického pedagogického výzkumu

Ve výzkumu se řeší buď jeden, nebo více (zpravidla spolu souvisejících) problémů. Řešení vědeckého problému potom představuje řadu navzájem propojených a na sobě závislých kroků a činností. Jednotlivé výzkumy se mohou navzájem lišit co do posloupností jednotlivých realizovaných činností, ale základní schéma postupu bývá následující:

- stanovení problému;
- formulace hypotézy;
- testování (verifikace, ověřování) hypotézy;
- vyvození závěrů a jejich prezentace.

1.2.1 STANOVENÍ PROBLÉMU A JEHO FORMULACE

1.2.1.1 Teoretická analýza poznatků ve zkoumané oblasti

Práce při stanovení problému obvykle začíná tzv. **předběžnou teoretickou analýzou poznatků** v oblasti, kterou hodláme zkoumat. Cílem této analýzy je seznámení se současným stavem a úrovní poznání v dané oblasti a získání co největšího množství dostupných informací o výzkumech, které již byly v této oblasti realizovány.

Základním a nejdůležitějším zdrojem informací je studium příslušné odborné literatury. Kromě studia odborné literatury v tištěné podobě (knižní publikace, časopisy, sborníky, encyklopedie, odborné slovníky apod.) můžeme využívat mnoho dalších zdrojů.

Obrovské informační možnosti skýtá například internet a jeho informační databáze (např. ERIC, EBSCO, PBD a mnohé další). Důležité informace získáváme také na základě konzultací a rozhovorů s odborníky, studiem výzkumných zpráv nebo i na základě přímého empirického sledování pedagogické reality (vlastní pozorování, rozhovory s učiteli, žáky, rodiči apod.). Tuto etapu práce v přípravě výzkumu není radno podceňovat. V současné době je v pedagogice jen velmi málo oblastí, které dosud nebyly nějakým způsobem podrobeny zkoumání. Jestliže se důkladně seznámíme se stavem poznání v dané oblasti, vyvarujeme se tím jednak zbytečného řešení problémů již vyřešených, jednak se vyhneme chybám a omylům, kterých se dopustili autoři před námi.

Začátečníkům lze doporučit, aby seznamování s odbornou literaturou začínali od odborných slovníků a encyklopedií (např. Průcha et al., 2009). Ve slovnících a encyklopediích získáváme základní informace o problematice, kterou máme ve výzkumu řešit, a často také odkazy na další odbornou literaturu.

O každém použitém informačním zdroji (např. tiskem vydané knize, časopise nebo sborníku, on-line publikaci apod.) si pořizujeme **bibliografický záznam**. Tento záznam musí obsahovat údaje, které jsou nutné k identifikaci použitého informačního zdroje. Pořizování bibliografických záznamů a citací se řídí pravidly a zvyklostmi, které zpravidla platí pro určitou zemi, ale mohou se lišit i u různých vydavatelů. V České republice platí v současné době (od 1. dubna 2011) norma ČSN ISO 690, která stanoví, které údaje je nutné v bibliografickém záznamu uvádět. Pro běžnou praxi většinou postačí seznámit se s následujícími příklady bibliografických citací:

Monografické publikace

Jeden autor:

PELIKÁN, Jiří. *Základy empirického výzkumu pedagogických jevů*. Praha: Karolinum, 1998. ISBN 80-7184-569-8.

Dva autoři:

KRÁLÍK, Oldřich a Jiří HARTMANN. *Základy statistiky pro pedagogy*. Brno: Akademické nakladatelství Cerm, 2000. ISBN 80-7204-152-5.

Tři autoři:

PRŮCHA, Jan, WALTEROVÁ, Eliška a Jiří MAREŠ. *Pedagogický slovník*. 3. rozšířené a aktualizované vydání. Praha: Portál, 2001. ISBN 80-7178-579-2.

Více než tři autoři:

PRŮCHA, Jan et al. *Pedagogická encyklopedie*. Praha: Grada Publishing, 2009. ISBN 978-80-7367-546-2.

Kolektiv autorů:

Akademický slovník cizích slov. 2. díl, L–Ž. Praha: Academia, 1995. ISBN 80-200-0524-2.

Kapitola v monografii, článek v časopise, článek ve sborníku

Kapitola v monografii:

OBST, Otto. Učitel ve výuce. In: KALHOUS, Zdeněk et al. *Školní didaktika*. Praha: Portál, 2002, s. 92–120. ISBN 80-7178-253-X.

Článek v tištěném sborníku:

KRAUS, Blahoslav. Volný čas dětí a mládeže v pedagogickém výzkumu. In: *Nové možnosti vzdělávání a pedagogický výzkum.* Sborník příspěvků z 9. celostátní konference ČAPV s mezinárodní účastí. Ostrava: Pedagogická fakulta Ostravské univerzity, 2001, s. 428–433. ISBN 80-7042-181-9.

Článek ve sborníku na CD-ROM:
PRŮCHA, Jan. Deset let České asociace pedagogického výzkumu: Bilance a výhledy. In: *Výzkum školy a učitele.* Sborník referátů z 10. výroční mezinárodní konference ČAPV [CD-ROM]. Praha: Univerzita Karlova, Pedagogická fakulta, 2002.

Článek v odborném časopise

Článek v tištěném časopise:
ŠVEC, Vlastimil. Sebereflexe studentů v pregraduální didaktické přípravě. *Pedagogika*, 1996, roč. 46, č. 3, s. 266–276. ISSN 3330-3815.

Článek v elektronickém časopise:
NEZVALOVÁ, Dana. Akčním výzkumem k zlepšení kvality školy. *e-Pedagogium* [online], 2002, roč. 2, č. 4. [cit. 12. 2. 2003]. Dostupné na: http://epedagog.upol.cz/eped4.2002/index.htm. ISSN 1213-7499.

1.2.1.2 Formulace problému

Dalším krokem, který lze při přípravě výzkumu doporučit, je formulování tzv. **operacionalizovaných definic** pojmů, se kterými budeme ve výzkumu pracovat. Jedná se o definice, které umožní jednotlivé pojmy (konstrukty) jednoznačně „uchopit" (zachytit, změřit). Například při zkoumáni „agresivity u dětí předškolního věku" bude třeba jednoznačně vymezit projevy agresivity, přesně definovat věk dětí, které hodláme zkoumat, apod.

Při formulaci operacionalizovaných definic zpravidla jednotlivé pojmy definujeme poněkud zjednodušeně (vhledem k zaměření výzkumu). Toto zjednodušování definic pojmů má dva důvody. První důvod spočívá v nemožnosti postihnout pedagogické jevy v celé jejich složitosti, vzájemné souvislosti a podmíněnosti. Druhý spočívá v požadavku, aby sledované jevy byly nějakým způsobem zachytitelné (měřitelné). Při formulaci operacionalizovaných definic si musíme být vědomi toho, že určité jevy zjednodušit nelze, nechceme-li zkreslit výsledky výzkumu. Jedině důkladný teoretický rozbor může určit hranice, kam až zjednodušení může sahat, aniž by hrozilo nebezpečí zkreslení (simplifikace).

Jevy nebo vlastnosti, které ve výzkumu vystupují a mezi nimiž hledáme (ověřujeme) existenci vztahů, označujeme jako **proměnné**. Proměnnou je pedagogický jev nebo vlastnost, která se ve výzkumu může měnit (nabývat různých hodnot). Příkladem proměnných je například pohlaví dětí (nabývá dvě možné hodnoty), věk dětí, mentální úroveň dětí, klasifikace žáků v určitém předmětu, chování dětí v určité situaci atd. Proměnné lze rozdělit na tzv. **nezávisle proměnné** a **závisle proměnné**. Nezávisle proměnná je vlastnost (jev), která je příčinou nebo podmínkou vzniku jiné vlastnosti (jevu). Závisle proměnná je vlastnost (jev), která je výsledkem (následkem, důsledkem) působení nezávisle proměnné. Například negativní chování dítěte ve škole (závisle proměnná) může být způsobeno například konfliktními vztahy mezi jeho rodiči (nezávisle proměnná). Správně formulovaný výzkumný problém je otázka, která by měla vyjadřovat vztah mezi proměnnými (měla by se tázat, zda mezi proměnnými existuje vztah).

Při vlastní formulaci problému lze doporučit respektování následujících doporučení:

- Problém by měl být formulován zcela konkrétně, jednoznačně a pokud možno v tázací formě.

- Problém musí implikovat možnost empirického ověřování. Problémy, které nejsou empiricky ověřitelné, nelze ve vědeckém výzkumu zkoumat.
- Problém by měl vyjadřovat vztah mezi dvěma nebo více proměnnými.

Poznámka: *Pokud otázka, na niž hledáme ve výzkumu odpověď, nevyjadřuje vztah mezi proměnnými, nemusí to ještě znamenat, že je bezcenná a že nemá smysl ji řešit. Otázka, která nevyjadřuje vztah mezi proměnnými, ale neumožňuje vyslovit hypotézu, a při jejím řešení se tedy nejedná o výzkum v tom smyslu, ve kterém byl definován výše. Šetření tohoto typu bývají označována jako pedagogické průzkumy.*

1.2.2 HYPOTÉZY A JEJICH MÍSTO V PEDAGOGICKÉM VÝZKUMU

Hypotézy tvoří jádro kvantitativně orientovaných výzkumů. K současnému chápání významu a role hypotéz ve výzkumu významně přispěl **kritický racionalismus**, filozofický směr, jehož zakladatelem je významný filozof vědy Karl R. Popper (1902–1994). Tento autor dospěl k závěru (Popper, 1995), že obecně formulovaná tvrzení (hypotézy) není možno empiricky prokázat (verifikovat). Pro verifikaci hypotéz navrhl tzv. **metodu falzifikace**. Termínem „falzifikace" se v tomto případě rozumí hledání empirických faktů, které **hovoří proti** ověřované hypotéze (v běžném životě má slovo „falzifikace" význam jiný, znamená „padělání" nebo „falšování" něčeho).

Podle K. R. Poppera by vědec ve výzkumu neměl usilovat o dokazování hypotéz, ale pouze o jejich falzifikaci, tj. hledání faktů, svědčících o jejich neplatnosti. Pokud se nepodaří hypotézu ve výzkumu falzifikovat, můžeme ji přijmout, ne ji však považovat za jednou provždy dokázanou. Vždy existuje možnost, že při opakovaném ověřování hypotézy budou nalezena fakta, která s ní nejsou slučitelná. Správně formulovaná vědecká hypotéza musí možnost empirického ověřování (falzifikace) skýtat, tj. musí být falzifikovatelná.

Žádný empirický důkaz nemůže hypotézu nikdy jednoznačně a definitivně dokázat. Je možné říci, že empirický výzkum v podstatě hypotézu nedokazuje, ale pouze zdůvodňuje její přijatelnost. Je-li hypotéza na základě důkladného empirického ověřování přijata, je možné ji zobecnit a doporučit k praktickému využití.

1.2.2.1 Pravidla pro formulaci hypotéz

Při formulaci hypotéz je nutné dodržovat tři základní požadavky, které bývají někdy označovány jako **zlatá pravidla hypotézy** (Gavora, 2000):

- Hypotéza je tvrzení, které je vyjádřeno oznamovací větou (výzkumný problém je naopak vhodné vyjádřit větou tázací).
- Hypotéza musí vyjadřovat vztah mezi dvěma proměnnými (pokud se nejedná o vyjádření vztahů, není možno hovořit o vědecké hypotéze). Proto musí být hypotéza vždy formulována jako **tvrzení o rozdílech, vztazích nebo následcích**.
- Hypotézu musí být možno empiricky ověřovat. Proměnné, které v hypotéze vystupují, musí být měřitelné (byť například jen na základě kategorizace).

Zatímco problém je otázka, která se táže, zda existuje vztah mezi pedagogickými jevy, hypotéza je podmíněným výrokem o vztahu mezi dvěma nebo více proměnnými. Lze také říci, že hypotézy jsou predikcemi (předpověďmi) o vztazích mezi proměnnými. Hypotéza říká, že „nastane-li jev A, nastane také jev B" (jev B je předpovídán na základě existence jevu A). Hypotéza vyjadřující vztah mezi dvěma proměnnými se někdy formálně zapisuje pomocí vztahu

$$Y = f(X).$$

Tento zápis vyjadřuje skutečnost, že jistá proměnná (vlastnost, jev) Y je „funkcí" jiné proměnné X. Jestliže v těchto souvislostech hovoříme o „funkční závislosti" proměnných, je nutné zdůraznit, že v pedagogických výzkumech se ve skutečnosti jedná o závislosti statistické. Ty jsou jiné povahy než skutečné funkční závislosti, jak je známe například z fyziky apod.

Existence jednoduchého hypotetického vztahu $Y = f(X)$ je v oblasti pedagogického zkoumáni málo častá. Daleko častěji se uplatňuje vztah

$$Y = f(X, W, Z, \ldots),$$

tj. jistý účinek je zpravidla vyvoláván celou řadou faktorů (X, W, Z…). Často je však oprávněné a vhodné předpokládat, že z možných faktorů je nejdůležitější jen jeden a ostatní je možno zanedbat. Přípustnost takového zjednodušení je nutné vždy pečlivě zvažovat, aby nedošlo ke zkreslení složité pedagogické reality.

Po zformulování hypotézy se doporučuje provést ještě další myšlenkový krok, který označujeme jako **dedukci důsledků hypotézy**. Při tomto způsobu usuzování vycházíme z platnosti formulované hypotézy a pokoušíme se zpětně dedukovat, který problém z toho vyplývá (*bez* ohledu na problém již dříve formulovaný). Může se stát, že při tomto postupu dospějeme k částečně (*nebo* zcela) jiné formulaci problému, než z jaké jsme původně vyšli. Můžeme také zjistit, že původně stanovený problém není současnými prostředky vědy řešitelný (Kerlinger, 1972).

1.2.2.2 Nejčastější chyby při formulaci hypotéz

Nedostatky při formulaci hypotéz výrazně snižují věrohodnost realizovaného výzkumu a znehodnocují nebo přinejmenším zpochybňují dosažené výsledky. Velmi často se při formulaci hypotéz objevují například následující nedostatky:

- Formulované hypotézy nevyjadřují vztah mezi proměnnými, což znamená, že nevypovídají o rozdílech, vztazích nebo následcích. Příklady nesprávných formulací: „Žáci na prvním stupni základní školy mají rádi matematiku." „Chlapci většinou rádi sportují." „Městské školy jsou dobře vybaveny výpočetní technikou." V uvedených příkladech jsou vztahy jen implicitně naznačeny, jejich obsah však není jednoznačně vymezen.
- Hypotézy nemají formu oznamovací věty. Někdy jsou vyjádřeny pomocí složitých souvětí, z nichž žádné jednoznačné tvrzení nevyplývá.
- Často se objevují neurčité formulace typu „jev A někdy vyvolává jev B". Také při interpretaci výsledků ověřování se někdy vyskytují nejednoznačné formulace typu „hypotéza byla částečně potvrzena" apod. Hypotézy musí být formulovány vždy zcela jednoznačně („natvrdo") a také výsledek ověřování musí být zcela jednoznačný (hypotézu buď přijímáme, nebo odmítáme).

- Jestliže formulujeme hypotézy na začátku výzkumu, potom hovoříme vždy o tzv. **věcných hypotézách**, nikoli o hypotézách statistických (srov. kap. 3). Statistické hypotézy (nulová, alternativní) se formulují a používají až v souvislosti s jejich statistickým ověřováním.

1.2.3 TESTOVÁNÍ HYPOTÉZ VE VĚDECKÉM VÝZKUMU

Při testování (ověřování, verifikaci) hypotézy jde o rozhodnutí, zda můžeme vyslovenou hypotézu přijmout (zda není v rozporu s empirickými fakty). Rozhodnout o přijatelnosti hypotézy lze u kvantitativně orientovaných výzkumů pouze na základě rozsáhlého shromažďování (sběru) dat, jejich tříděním, zpracováním a vyhodnocováním. Data ve výzkumu získáváme metodami, které bývají souborně označovány jako **empirické metody** (např. pedagogické pozorování, dotazník, škály, rozhovor, různé typy testů, sociometrie, Q-metodologie, sémantický diferenciál apod.).

Významné místo při zpracování dat v pedagogických výzkumech i při interpretaci získaných výsledků má **matematická statistika**. Statistika je věda, která se zabývá metodami sběru, zpracování a vyhodnocování **hromadných dat**. Hromadná data získáváme sledováním hromadných jevů, což jsou jevy, které lze sledovat opakovaně (mnohokrát). Jedinečné jevy statistické analýze podrobovat nelze.

Výsledků, ke kterým statistika za několik posledních desetiletí dospěla, není dosud v našich pedagogických výzkumech vždy náležitě využíváno. Mnohé výzkumy při zpracování a interpretaci výsledků využívají jen elementární postupy (jako je výpočet průměrů, procent apod.) a naprosto ignorují možnosti, které moderní statistika nabízí.

Při analýze dat získaných ve výzkumu plní statistika zejména dva základní úkoly. Prvním úkolem, kterým se zabývá tzv. **popisná (deskriptivní) statistika**, je shromážděná data popsat tak, aby poskytovala co možná nejpřesnější, přehlednou a názornou informaci o měřených hromadných jevech. Druhým základním úkolem statistiky je pomáhat při rozhodování, zda mezi sledovanými jevy (proměnnými) je, či není vztah. Tento druhý úkol plní **tzv. induktivní statistika**. Základním myšlenkovým principem induktivní statistiky je usuzování na vlastnosti celku na základě vlastností jeho části.

Na základě výsledků ověřování hypotéz vyslovujeme závěry, ke kterým výzkum dospěl. Konstatujeme přijetí či odmítnutí hypotéz, interpretujeme dosažené výsledky, srovnáváme je s dosavadními výsledky vědy, zdůvodňujeme případné rozdíly. Někdy na základě zjištěných výsledků dedukujeme další podmíněné výroky o vztazích mezi proměnnými. Tyto výroky se mohou stát hypotézami pro případné další výzkumy.

1.2.3.1 Výběr prvků do výzkumných vzorků

Jestliže v běžném životě vyslovujeme soudy o jiných lidech či skupinách lidí, činíme tak většinou na základě znalosti určitého (někdy jen zcela malého) počtu osob. Předpokládáme, že vlastnosti lidí, o kterých se vyslovujeme, jsou stejné (nebo podobné) jako vlastnosti těch, které známe. Podobně jako v běžném životě, ani v pedagogickém výzkumu není zpravidla myslitelné, abychom prozkoumali všechny jedince (nebo situace), kteří nás zajímají. Svoje zjištění opíráme většinou jen o znalost určitého vzorku (výběru). Jde o to, aby vlastnosti námi vybraného vzorku byly pokud možno stejné jako vlastnosti celé skupiny (lidí nebo situací), kterou zkoumáme. Požaduje se, aby vzorek vybraných jedinců (situací) byl co možná nejvíce reprezentativní. V běžném životě se otázkou reprezentativnosti vzorku příliš nezabýváme. Jinak je tomu ve vědeckých výzkumech, kde otázka reprezentativnosti výběru je otázkou klíčového významu.

V dalším výkladu budeme používat dva důležité pojmy: **základní soubor (populace)** a **výběrový soubor (výběr)**. Pojmem „základní soubor" rozumíme všechny prvky (osoby, situace) patřící do skupiny, kterou zkoumáme. Výběrovým souborem (výběrem, vzorkem) rozumíme určitou část prvků vybranou ze základního souboru, která základní soubor zastupuje (reprezentuje).

V některých případech výzkumů (většinou jen v případě malých základních souborů) je zkoumán celý základní soubor. V těchto situacích hovoříme o **vyčerpávajícím (exhaustivním) výběru**. Šetření, ve kterém získáváme data ode všech prvků (osob, situací) v populaci, označujeme také jako **cenzus**.

Druhy výběrů

Existuje více způsobů, jak vybírat jedince (nebo situace) tak, aby danou skupinu osob (nebo situací) dobře reprezentovali. Společným rysem všech těchto postupů je, že rozhodnutí o tom, který prvek (osobu, situaci) vybereme, bude objektivní. Objektivita se zabezpečuje nejčastěji uplatněním náhody. Ve vědeckém výzkumu musí být zaručeno, že při výběru prvků se neuplatní jakýkoli subjektivní zřetel (byť sebelépe míněný), a to ať skrytý či zdánlivě bezvýznamný.

Prostý náhodný výběr (náhodný výběr jednotlivých prvků)

Charakteristickým rysem tohoto výběru je, že všechny prvky souboru mají stejnou pravděpodobnost, že budou vybrány. Každý prvek musí být přitom vybírán nezávisle na ostatních. Tuto podmínku výběr přesně splňuje jen v případě, že se jedná o tzv. **výběr s vracením**. U tohoto výběru se po každém výběru vybraný prvek vrací zpět do základního souboru. Tím, že se vybírá stále ze stejného počtu prvků, je dána stejná pravděpodobnost výběru pro všechny prvky. V praxi se častěji provádí tzv. **výběr bez vracení**, kdy vybrané prvky již zůstávají mimo základní soubor. U početnějších základních souborů nemá smysl mezi výběrem s vracením a bez vracení rozlišovat.

Prakticky se prostý náhodný výběr provádí buď mechanickým losováním (v osudí musí být všechny prvky základního souboru), nebo se používá techniky náhodných čísel. U této techniky se prvkům základního souboru nejdříve přiřadí pořadová čísla a z nich se potom vybírá pomocí náhodných čísel. Náhodná čísla možno získat například z tabulek náhodných čísel, ale v současné době je pohodlnější používat k tomuto účelu počítač nebo i některé typy kalkulátorů. Na některých typech kalkulátorů bývá například funkce označena RND (*random* – náhoda). Stisknutím příslušného tlačítka nám kalkulátor zobrazí náhodné číslo, které je možno použít při výběru.

Prostý náhodný výběr je v mnoha případech v praxi poměrně obtížně uskutečnitelný. Představme si například, že bychom hodlali uskutečnit prostý náhodný výběr žáků 2. ročníku základní školy v České republice o rozsahu řekněme 500 dětí. Vybraný vzorek by byl zřejmě v tomto případě značně rozptýlen a bylo by prakticky nemožné s ním ve výzkumu pracovat. V podobných případech se často místo náhodného výběru jednotlivých prvků provádí výběr skupin (např. výběr školních tříd).

Skupinový výběr

Tato výběrová technika se používá v případě, že základní soubor je uspořádán do určitých skupin (např. základní soubor žáků 1. ročníku základní školy v určitém kraji tvoří jistý počet tříd). Jsou-li skupiny v základním souboru přibližně stejně početné, vybíráme (např. pomocí losování) tak, že skupiny mají stejnou pravděpodobnost, že budou vybrány. Nejsou-li skupiny stejně početné, je možné vybírat s pravděpodobností úměrnou rozsahu těchto skupin.

Skupinový výběr je v pedagogických výzkumech pro svoji snadnou proveditelnost často užívaný. Je ovšem potřeba upozornit na to, že rozsah (velikost) skupinového výběru musí být dostatečný. U tohoto výběru totiž o potřebném rozsahu nerozhoduje jen velikost skupin, ale také jejich počet.

Stratifikovaný výběr

Provádí se u těch základních souborů, které jsou složeny z několika charakteristických podskupin. Chceme-li ze základního souboru, který je složen z podskupin, získat dostatečně reprezentativní výběr, vybíráme z jednotlivých charakteristických podskupin pomocí náhodného výběru vždy určitý počet prvků. Počet vybíraných prvků z podskupin nebývá přesně proporcionální vzhledem ke složení základního souboru.

Například při výzkumu postojů učitelů základní školy k učitelské profesi byl pořízen stratifikovaný výběr učitelů podle délky jejich pedagogické praxe. Všichni učitelé byli rozděleni do následujících pěti podskupin:

1. podskupina – učitelé s délkou pedagogické praxe do 5 let;
2. podskupina – učitelé s délkou pedagogické praxe 6–10 let;
3. podskupina – učitelé s délkou pedagogické praxe 11–15 let;
4. podskupina – učitelé s délkou pedagogické praxe 16–20 let;
5. podskupina – učitelé s délkou pedagogické praxe nad 20 let.

Z každé této podskupiny byl potom náhodně vybrán určitý počet učitelů, čímž bylo zaručeno, že ve výběru se určitým způsobem uplatnily vlastnosti učitelů všech kategorií podle délky pedagogické praxe.

Kontrolovaný výběr (proporcionální stratifikovaný výběr)

Jde o takový stratifikovaný výběr, u něhož počet prvků vybíraných z podskupin je proporcionální k počtu prvků v základním souboru. Při realizaci kontrolovaného výběru učitelů základní školy (podle délky pedagogické praxe) bychom museli respektovat počty učitelů v jednotlivých kategoriích délky praxe. Tímto způsobem bychom dostali výběr, který by byl zmenšeným modelem základního souboru vzhledem k danému rozlišovacímu znaku (délka pedagogické praxe).

Kontrolovaný výběr může být prospěšný v mnoha konkrétních případech výzkumu. Například při řešení problémů, které nějakým způsobem mohou souviset s pohlavím žáků (např. zkoumání volnočasových aktivit, tělesné zdatnosti žáků atd.), bývá nutné zajistit, aby výběr obsahoval stejný počet chlapců a dívek. V tomto případě provádíme kontrolovaný výběr žáků podle pohlaví tak, že náhodně vybereme dvě stejně velké skupiny chlapců a dívek. Výběr je v tomto případě zmenšeným modelem základního souboru vzhledem k pohlaví.

Kontrolované výběry bývají někdy označovány jako **reprezentativní výběry**. Výběr je možno kontrolovat i podle několika důležitých znaků současně. Kontrolované výběry umožňují, i při poměrně malém rozsahu, získávání značně věrohodných výsledků.

Vícenásobný výběr

U tohoto výběru se nezačíná výběrem jednotek (např. žáků), ale výběrem skupin vyššího řádu. Nejdříve se provede výběr skupin nejvyššího řádu a potom se pokračuje ve dvou až třech stupních, až dospějeme k základním jednotkám (žákům). Chceme-li například pořídit vícestupňový výběr žáků pátých ročníků základní školy v České republice, můžeme postupovat tak, že nejdříve vybereme (náhodně) několik krajů (první stupeň výběru), dále ve vybraných krajích vylosujeme několik škol (druhý stupeň výběru) a ve školách náhodně vybereme několik žáků (třetí stupeň výběru).

Výhodou vícestupňového výběru je, že vybrané prvky jsou zpravidla více koncentrovány než u ostatních druhů výběrů. K dosažení věrohodných výsledků je však u tohoto výběru potřeba vybírat vždy poněkud větší počet prvků než u prostého náhodného výběru.

Záměrný výběr

Záměrný výběr se liší od předcházejících druhů výběrů v tom, že zde o výběru jistého prvku nerozhoduje náhoda, ale buď úsudek výzkumníka, nebo úsudek zkoumané osoby. Záměrný výběr může vzniknout v podstatě třemi způsoby:

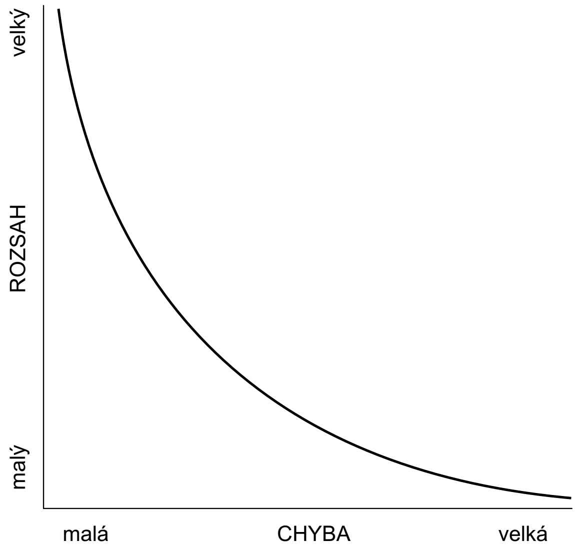
- **anketní výběr** (jedinci se dostávají do výběru sami na základě svého rozhodnutí);
- **výběr „průměrných jednotek"**, při kterém se vybírá určitý objekt (např. škola, třída, žáci), který výzkumník považuje za typický (průměrný) případ. Tato metoda předpokládá vysokou kvalifikaci a erudici výzkumného pracovníka, který musí dobře rozlišovat mezi jevy jedinečnými, zvláštními a obecnými. Je to postup v jistém směru jednodušší, rychlejší a lacinější než ostatní uvedené postupy, zato však přináší nepoměrně méně věrohodné výsledky. Je totiž značně obtížné (ne-li nemožné) dokázat na skutečně vědecké úrovni, že vybraný objekt je typickým reprezentantem základního souboru;
- **kvótní výběr** je jediný záměrný výběr, který z teoretického hlediska vyhovuje. Nejdříve se zvolí určité kontrolní znaky, podle nichž se výběr orientuje. Chceme-li například provést kvótní výběr obyvatelstva určité oblasti, můžeme zjistit, že v tomto základním souboru je určitý počet mužů a žen, že obyvatelstvo má jisté věkové složení, že tyto osoby vykonávají různá povolání, mají různý stupeň vzdělání, že bydlí v různě velkých obcích atd. Podle těchto kontrolních znaků lze vytvořit určité kvóty pro výběr. Kvóty by například v uvedeném příkladu předepisovaly, kolik mužů a žen se má vybrat a v jakém věku, jaké vzdělání mají vybraní jedinci mít, jaké povolání mají vykonávat atd. Kvótní výběr se užívá často v sociologických výzkumech a průzkumech (např. výzkumy veřejného mínění, průzkumy trhu apod.).

  Zvláštní formou kvótního výběru je tzv. **panel**. O panelu hovoříme tehdy, jestliže vytvořenou reprezentativní skupinu osob používáme opakovaně k řadě různých výzkumů (je to možné pouze v případě, že složení základního souboru se příliš rychle nemění).

Mechanický (systémový) výběr
Tento druh výběru je výhodný v případě, že v šetření hodláme zkoumat určité procento prvků ze základního souboru (např. 2 %). Při pořizování mechanického výběru postupujeme tak, že nejdříve všem prvkům základního souboru přiřadíme pořadová čísla. Potom prostým náhodným výběrem určíme počáteční prvek základního souboru a k jeho pořadovému číslu postupně přičítáme konstantu, která odpovídá zvolenému procentu vybíraných prvků. Například v případě, že hodláme vybrat 2 % všech prvků (2 % = 2 / 100 = 1 / 50), postupně přičítáme konstantu 50. (Počáteční prvek základního souboru se v případě výběru 2 % prvků určí z intervalu 1–50.)

Spárované (vyrovnané) výběry
Jde o zvláštní druh kontrolovaných výběrů, kdy získáváme ze základního souboru dva nebo více podobně kontrolovaných výběrů. Pro některé pedagogické výzkumy je například vhodné získat dva výběry se stejným nebo podobným rozdělením určité schopnosti (např. mentální úroveň, tělesná zdatnost). Za tím účelem můžeme vybrané jedince rozdělit do určitých „výkonnostních pásem“ podle výsledků určitého provedeného měření (např. podle výsledků testu rozumových schopností nebo podle výsledků testu tělesné zdatnosti). Každé takto vytvořené výkonnostní pásmo (osoby, které dosáhly určitého výkonu) lze potom náhodně rozdělit do dvou nebo více spárovaných skupin. Při rozdělování jedinců do skupin lze užít velmi jednoduché metody „házení mincí“. Postupujeme tak, že z osob vytvoříme zcela libovolně dvojice a pomocí hodu mincí určíme, do které skupiny jedince zařadíme. K rozdělování jedinců do spárovaných skupin lze také užít techniky náhodných čísel, která generuje počítač nebo kalkulátor. Při použití náhodných čísel lze postupovat tak, že nejdříve jedincům, které máme rozdělit, přiřadíme náhodná čísla. Jedinci, kterým uvedeným způsobem přiřadíme lichá čísla, potom tvoří jednu skupinu a jedinci, kterým byla přiřazena sudá čísla, tvoří druhou spárovanou skupinu (nulu považujeme v tomto případě za sudé číslo).

1.2.3.2 Rozsah výběru
Pokud je výběr v pedagogickém výzkumu proveden adekvátně, měli by vybraní jedinci (situace) mít zhruba stejné vlastnosti jako základní soubor. Veličiny (míry), které jsou založeny na vlastnostech základního souboru, se nazývají **parametry**. Veličiny (míry) odvozené z výběru se nazývají **výběrové charakteristiky**. Parametr má vždy stálou hodnotu, ale většinou není přesně známá. Většinou lze parametr s uspokojivou přesností odhadnout z výběrové charakteristiky. U dobře provedených výběrů není rozdíl mezi parametrem a výběrovou charakteristikou velký. Velikost tohoto rozdílu závisí jednak na druhu (kvalitě) výběru, jednak na rozsahu (velikosti) výběru. Obecně platí, že čím větší je rozsah výběru, tím menší je rozdíl mezi výběrovou charakteristikou a parametrem. Závislost mezi velikostí výběru a chybou odhadu parametru pomocí výběrové charakteristiky názorně ukazuje obrázek 1.



Obr. 1 *Závislost mezi chybou odhadu parametru a rozsahem výběru*

Z uvedeného vyplývá, že čím větší výběr pořídíme, tím více se přiblížíme ke skutečným vlastnostem základního souboru.

Potřebný rozsah výběru lze v některých případech odhadnout pomocí výpočtu. K pochopení dále uváděných postupů je však nejdříve nutné se seznámit se základními pojmy teorie měření a se základními metodami statistického zpracování dat (srov. kap. 2).

Při výpočtech potřebného rozsahu výběru je nejdříve nutné nějakým způsobem odhadnout, jak dalece se zkoumaný znak (měřená vlastnost) v základním souboru rozptyluje, tj. jakou má **variabilitu**. Tento odhad je možno učinit buď na základě již známých (dříve uskutečněných, stejných nebo podobných) výzkumů, anebo je možno uskutečnit předvýzkum a k odhadu variability použít jeho výsledky.

Odhad rozsahu výběru v případě metrických dat

Odhad potřebného rozsahu výběru lze v tomto případě vypočítat pomocí vzorce

$$n = \frac{t_\alpha^2 \cdot s^2}{\Delta^2} \tag{1}$$

kde t_α je koeficient spolehlivosti pro zvolenou spolehlivost α (při běžně požadované spolehlivosti 95 % dosazujeme hodnotu 1,96, při požadované spolehlivosti 99 % dosazujeme hodnotu 2,58), s je směrodatná odchylka a Δ je tzv. požadovaná přesnost (přípustná absolutní chyba). Při výpočtu požadovaného rozsahu výběru tedy vycházíme z odhadu směrodatné odchylky, ostatní dvě potřebné hodnoty víceméně volně volíme na základě zvážení dané situace.

Poučení o tom, co jsou to metrická data, nalezne čtenář v oddíle 2.1 a poučení o výpočtu směrodatné odchylky lze najít v oddíle 2.3.4.

Příklad 1

Máme odhadnout, kolika žákům musíme zadat didaktický test, chceme-li dostatečně spolehlivě a přesně zachytit úroveň jejich vědomostí. V předvýzkumu byla určena směrodatná odchylka $s = 2{,}1$.

Budeme-li požadovat 95% spolehlivost výsledků a spokojíme-li se s přesností ±0,2 bodu, dostáváme ze vzorce (1)

$$n = \frac{1{,}96^2 \cdot 2{,}1^2}{0{,}2^2} \doteq 423$$

Chceme-li tedy didaktickým testem zachytit úroveň vědomostí žáků se zvolenou spolehlivostí a přesností, měli bychom testem vyzkoušet 423 žáků.

Pokud bychom požadovali spolehlivost 99 %, dostali bychom při stejné požadované přesnosti (±0,2 bodu) výsledek $n = 734$ žáků. Pokud bychom požadovali přesnost ±0,1 bodu, potom by při původně požadované spolehlivosti (95 %) byl nutný rozsah $n = 1\,694$ žáků.

Z uvedeného příkladu je zřejmé, že odhad rozsahu výběru je značně závislý na požadavcích na spolehlivost a přesnost.

Odhad rozsahu výběru v případě nominálních nebo ordinálních dat

V případě nominálních nebo ordinálních dat lze potřebný rozsah výběru odhadnout (Nowak, 1965) podle vzorce

$$n = \frac{t_\alpha^2 \cdot p \cdot (1-p)}{d^2} \tag{2}$$

kde n je požadovaný rozsah výběru, t_α je koeficient spolehlivosti pro zvolenou spolehlivost α, p je odhad relativní četnosti zkoumaného znaku v základním souboru a d je požadovaná relativní přesnost (většinou se požaduje přesnost 3–4 %, tj. 0,03–0,04). Uvedený vzorec je velmi podobný vzorci (1). Rozdíl je v tom, že u nominálních nebo ordinálních dat odhadujeme variabilitu pomocí hodnoty $p \cdot (1 - p)$.

Jestliže je obtížné odhadnout relativní četnost zkoumaného znaku v základním souboru, doporučuje se počítat pro jistotu s hodnotou $p = 0{,}5$, poněvadž pro tuto hodnotu vychází rozsah výběru největší.

Příklad 2

Máme rozhodnout, jaký rozsah výběru je nutný k tomu, abychom dotazníkovým šetřením spolehlivě zjistili, kolik procent studentů gymnázia hodlá v určitém školním roce dál studovat na vysoké škole.

Na základě zkušeností z minulých let lze očekávat, že dál studovat na vysoké škole hodlá asi 50 % studentů gymnázia ($p = 0{,}5$). Jestliže budeme požadovat obvyklou spolehlivost 95 % a spokojíme-li se s obvyklou přesností 4 %, tj. $d = 0{,}04$, potom vychází

$$n = \frac{1{,}96^2 \cdot 0{,}5 \cdot (1 - 0{,}5)}{0{,}04^2} \doteq 600$$

K provedení dotazníkového šetření, které by mělo zvolenou spolehlivost a přesnost, by tedy bylo třeba použít výběru o rozsahu 600 studentů.

Je zřejmé, že požadovaný rozsah výběru závisí především na variabilitě (rozptýlení) dat v základním souboru. Jinak řečeno, rozsah výběru závisí na homogenitě základního souboru. Jestliže by téměř všechny prvky základního souboru měly zjišťovaný znak stejný (směrodatná odchylka by byla velmi malá, respektive relativní četnost p by se blížila k 1), potom by stačilo vybrat jen velmi málo prvků (v krajním případě jen jeden prvek), abychom získali správnou představu o základním souboru. Čím více se prvky základního souboru ve sledovaném znaku liší, tím větší výběr je třeba realizovat.

V některých případech lze potřebný rozsah výběru empiricky odhadnout i na základě určení jeho minimální a maximální hodnoty. Často se uvádí, že minimální rozsah výběru n_{min} z populace, která má n prvků, by měl být $n_{min} = 0{,}1 \cdot \sqrt{n}$ a maximální rozsah $n_{max} = \sqrt{n}$.

Úrovně pedagogického výzkumu

Pečlivě připravované kvantitativně orientované výzkumy mohou probíhat (podle důkladnosti přípravy a podle rozsahu a významu zkoumání) postupně na třech úrovních:

- pilotáž;
- předvýzkum;
- vlastní výzkum.

Cílem **pilotáže** je získání předběžných informací o dané problematice. Může se například jednat o volný rozhovor či pozorování, kterým provádíme první sondu do zákonitostí, které hodláme zkoumat. Údajů získaných pilotáží se zpravidla neužívá při vlastním výzkumu. Vhodně provedená pilotáž často umožňuje zpřesnit formulaci problému i hypotézy, může přinést cenné informace o verifikovatelnosti jednotlivých hypotéz atd.

Na pilotáž obyčejně navazuje **předvýzkum**. Předvýzkum by měl být jakýmsi zmenšeným modelem **vlastního výzkumu** (všech jeho hlavních fází). Provádí se většinou na poměrně malém vzorku osob, takže získané výsledky neumožňují činit obecnější závěry. Přesto je důležité, aby předvýzkum proběhl ve všech fázích, včetně testování hypotézy, statistického zpracování výsledků a vyvození závěrů. V předvýzkumu by se měly ověřit všechny metody a techniky, se kterými se počítá při vlastním výzkumu. Pečlivě provedený předvýzkum zmenšuje riziko použití nevhodné metody či techniky a často také přispívá ke zpřesnění formulace problému a hypotéz výzkumu.

1.3 VÝZKUMY EX-POST-FACTO A EXPERIMENTY

Ve vědeckém výzkumu jde o ověřování hypotéz o vztazích mezi přirozenými jevy, tj. o ověřování vztahů mezi proměnnými. Podle toho, zda ve výzkumu nějakým způsobem ovlivňujeme (manipulujeme) působení nezávisle proměnné, rozlišujeme výzkumy ex-post--facto a výzkumy experimentální (experimenty).

Výzkumy ex-post facto jsou takové výzkumy, u nichž se manipulace s nezávisle proměnnou neprovádí, a to buď proto, že není možná, anebo není žádoucí. U tohoto typu výzkumu se postupuje tak, že nejdříve shromáždíme údaje o závisle proměnné a teprve potom se retrospektivně hledá v množině možných nezávisle proměnných pravděpodobná příčina (podmínka) zjištěného stavu. Nevýhodou tohoto postupu je, že nezávisle proměnné lze jen velmi obtížně kontrolovat, a proto výsledky těchto výzkumů bývají méně hodnověrné než výsledky výzkumů experimentálních. Na druhé straně však zpětné hledání příčin (podmínek) vzniku pedagogických jevů je v některých případech jedinou použitelnou možností. Příkladem výzkumu ex-post-facto může být například výzkum příčin agresivního chování dětí ve škole. U tohoto výzkumu bychom zřejmě nejdříve identifikovali skupinu dětí s projevy agresivního chování. Potom bychom na základě zkušeností a teoretických úvah navrhli nejpravděpodobnější příčiny agresivního chování a ověřovali, zda se jednotlivé hypotetické příčiny na vzniku agresivity daných dětí podílejí.

U **experimentů** manipulujeme alespoň jednu nezávisle proměnnou. Manipulovaná nezávisle proměnná je pod kontrolou výzkumného pracovníka, a proto experimentální výzkumy poskytují (za jinak srovnatelných podmínek) věrohodnější výsledky než výzkumy ex-post-facto. Nevýhodou experimentu však je, že není použitelný ve všech situacích. Existují oblasti, kde nelze (nebo by z různých důvodů nebylo vhodné) experimentovat. Experiment totiž nesmí žádným způsobem škodit zkoumaným jedincům, a to ani objektivně, ani subjektivně (např. možnost manipulace proměnných, které vyvolávají negativní změny v osobnostech dětí). Další omezení pro experimentální práci pramení ze skutečnosti, že osoby, s nimiž experimentujeme, se chovají vždy více méně nepřirozeně.

Zkušenosti ukazují, že experiment je vhodnější pro zkoumání procesu vzdělávání než pro zkoumání procesu výchovy (výchova v užším slova smyslu). Výchovný proces je totiž nesrovnatelně složitější než proces vzdělávání, jeho logika je jemnější a hůře postižitelná ve srovnání s vyučováním. Je problematické, zda experiment je vůbec použitelný například na poli mravní výchovy, kde v případě nesprávných předpokladů může dojít k „nácviku" negativního chování apod. Pedagogická realita je nepoměrně složitější než realita fyzikální či biologická, a proto i experimentování v pedagogice je mnohem složitější než experimentování v technice či přírodních vědách. V přírodních vědách lze vyloučit (nebo přidat) jednu z podmínek, a tudíž lze jednoznačně určit, co vyvolalo daný účinek. Pro pedagogiku je typické, že nezávisle proměnné nepůsobí izolovaně, ale ve vzájemně podmíněných interakcích. Odtud pramení menší čistota výsledků při pedagogickém experimentování ve srovnání s experimentováním přírodovědným.

1.3.1 PEDAGOGICKÉ EXPERIMENTY

Význam experimentu pro rozvoj vědy nelze absolutizovat. Je známo, že například řada vědců dospěla k významným objevům, aniž by prováděli rozsáhlejší experimentální výzkum (např. A. Einstein, Ch. Darwin, S. Freud apod.). Lze konstatovat, že experimenty do značné míry upevňují pokrok vědy. Lze říci, že vědecké myšlení může být úspěšné i bez experimentování, experimentování bez myšlení je však bezcenné.

Podle toho, v jakých podmínkách experiment probíhá (a tudíž podle stupně kontrolovatelnosti nezávisle proměnných) lze rozlišit **experiment laboratorní** (*in vitro*) a **experiment přirozený** (*in vivo*).

Podle počtu působících nezávisle proměnných lze rozlišit **experiment jednofaktorový** a **experiment vícefaktorový** (dvou-, tří- … *n*-faktorový).

Podle toho, jakým způsobem je zabezpečována kontrola nad působením nezávisle proměnných, lze rozlišit tři základní techniky experimentu:

- techniku jedné skupiny;
- techniku paralelních skupin;
- techniku rotace faktorů.

U **techniky jedné skupiny** se experimentuje jen v rámci jedné skupiny. V této skupině se manipuluje nezávisle proměnná a také se měří závisle proměnná. Například učitelka mateřské školy může v určitém školním roce ověřovat, zda rychlost vytváření dovedností sebeobsluhy u dětí je závislá na úrovni kladné motivace. Po uplynutí školního roku, ve kterém byly děti k činnostem sebeobsluhy důsledně kladně motivovány, učitelka posuzuje, zda dosažená úroveň dovedností sebeobsluhy je „lepší", nebo „horší" ve srovnání s minulými léty. Výsledky tohoto experimentu jsou velmi málo věrohodné a často dokonce mohou být zavádějící. U jednoskupinového experimentu totiž není k dispozici žádné srovnání, a nelze tudíž tvrdit, že například lepší výsledky dětí jsou způsobeny právě kladnou motivací. Lepší výsledky mohou být docela dobře způsobeny vyšší vstupní úrovní dětí a řadou jiných neexperimentálních vlivů, které nejsou pod naší kontrolou.

Poněkud zdokonalenou formou techniky jedné skupiny je experiment, který F. N. Kerlinger (Kerlinger, 1972) označuje jako „jedna skupina před – po". U této modifikace se před experimentální manipulací nezávisle proměnné měří ve skupině úroveň vlastnosti (proměnné), která bude experimentálním zásahem ovlivňována. Na změnu vyvolanou experimentálním zásahem se usuzuje z rozdílu měření „před" a „po" experimentálním zásahu. Ani toto uspořádání experimentu neposkytuje většinou příliš věrohodné výsledky. Problematičnost postupu spočívá v tom, že určitou vlastnost zpravidla nelze měřit opakovaně za stejných podmínek (např. vliv zapamatování), ale také v tom, že v časovém intervalu mezi prvním a druhým měřením může působit řada nekontrolovaných faktorů. Určitý vliv může mít i zrání organismu (zvláště tehdy, jestliže časový interval mezi prvním a druhým měřením je dlouhý). Jednoskupinový experiment typu „jedna skupina před – po" však může dobře vyhovovat při zkoumání takových proměnných, kde opakované měření není na závadu (např. rychlost čtení, tělesná zdatnost žáků, fyziologické charakteristiky žáků apod.).

Experiment „jedna skupina před – po" může být realizován i v tzv. **simulované podobě**, která spočívá v tom, že první měření se uskuteční v nějaké jiné, podobné skupině. Změna vyvolaná experimentálním zásahem se potom posuzuje na základě rozdílu mezi tímto prvním měřením a druhým měřením, které probíhá ve skupině, v níž byl realizován experimentální zásah (Kerlinger, 1972). Tímto opatřením se sice vyhneme nežádoucímu zkreslení výsledků v důsledku zapamatování prvního měření, ale na druhé straně není nijak zaručeno, že skupina, v níž se provádí měření nezávisle proměnné, je skutečně (v daném směru) podobná té, se kterou experimentujeme.

U **techniky paralelních skupin** se pracuje současně se dvěma nebo více skupinami. Tím, že u této techniky existuje možnost srovnání, se dosahuje (ve srovnání s technikou jedné skupiny) celkově věrohodnějších výsledků. Skupiny, ve kterých se manipuluje nezávisle proměnná (provádí se experimentální zásah), označujeme jako **experimentální**. Skupiny, u nichž se manipulace nezávisle proměnné neprovádí (neprovádí se experimentální zásah), označujeme jako skupiny **kontrolní**.

Často se v souvislosti s experimentální technikou paralelních skupin hovoří o tzv. **experimentálních plánech** (projektech). Spolehlivost experimentů závisí ve velké míře na racionálním uspořádání celého výzkumu. Problémy, na které je třeba při navrhování experimentů myslet, budeme ilustrovat na experimentálních plánech podle E. F. Lindquista (Lindquist, 1967). Tento autor uvádí šest základních experimentálních plánů, které řadí podle dokonalosti (a v důsledku toho také spolehlivosti) od nejméně dokonalého k nejdokonalejšímu. Budeme předpokládat, že pedagogickým experimentem máme posoudit relativní účinnost dvou vyučovacích metod (metoda A a metoda B) v rámci výuky určitého předmětu na základní škole. Měřítkem účinnosti obou vyučovacích metod jsou dosažené vědomosti žáků na konci pokusu (měřené výstupním didaktickým testem):

- **Plán 1.** Experiment probíhá na jedné škole. Žáci daného ročníku jsou náhodně rozděleni do dvou stejně početných tříd. Jedna třída je vyučována metodou A, druhá metodou B, ve třídách vyučují různí učitelé.

  Spolehlivost tohoto experimentu je velmi malá. Náhodně vytvořené skupiny budou (vzhledem k malému počtu žáků, kteří se experimentu účastní) sotva vyrovnané (stejné úrovně). Zřejmě se obě vytvořené skupiny (třídy) budou lišit co do schopnosti učit se. Zjistíme-li na konci experimentu, že mezi výsledky tříd jsou rozdíly, lze je docela dobře zdůvodnit rozdíly ve schopnostech žáků ve skupinách, a nemusí tedy nutně být způsobeny aplikací různých vyučovacích metod. Navíc zde přistupuje ještě další zdroj chyb, a to možnost rozdílné úrovně práce učitele.
- **Plán 2.** Tento experimentální plán je jen nepatrně zdokonaleným plánem předchozím. Zdokonalení spočívá v tom, že obě zkoumané třídy vyučuje týž učitel. Můžeme předpokládat, že úroveň práce učitele v obou třídách bude stejná (i když ani zde nelze vyloučit například větší zaujetí učitele pro jednu z metod apod.), čímž se zčásti vyloučí vliv nestejných vnějších podmínek k učení.

  Spolehlivost tohoto experimentálního plánu je rovněž velmi malá.
- **Plán 3.** Experiment probíhá například na deseti školách (předpokládáme, že na každé z nich se experimentu zúčastní dvě třídy). Na pěti náhodně vybraných školách se vyučuje metodou A, na dalších pěti náhodně vybraných školách metodou B. Výsledky všech žáků vyučovaných jednou metodou jsou srovnávány s výsledky všech žáků vyučovaných metodou druhou.

  Spolehlivost a přesnost tohoto plánu není příliš vysoká, i když je zkoumáno mnohem více žáků než v předchozích případech. Příčinou malé spolehlivosti je hlavně to, že případné rozdíly mezi výsledky obou skupin žáků mohou být způsobeny rozdíly mezi školami (pořizované výběry nejsou prosté náhodné výběry žáků, ale jde o skupinové výběry, přičemž počet skupin je velmi malý). Tento experimentální plán podceňuje chyby způsobené rozdíly mezi skupinami. Z tohoto důvodu se nám mohou jevit rozdíly ve výsledcích obou skupin žáků jako „významné“, ač ve skutečnosti významné nejsou.

- **Plán 4.** Experiment probíhá opět na deseti školách. Na každé škole probíhá vyučování metodou A i metodou B a obě třídy vyučuje také stejný učitel. Experiment je vyhodnocován na základě srovnávání výsledků, jichž dosáhli všichni žáci vyučovaní metodou A a všichni žáci vyučovaní metodou B.

  Spolehlivost tohoto experimentálního plánu je mnohem větší než u plánu 3, i když výzkum zahrnuje stejný počet žáků. Tím, že se na každé škole aplikují obě metody, se dosahuje vyrovnání různých podmínek škol. Nedostatek popisovaného experimentálního plánu spočívá v tom, že pracujeme se dvěma výběry, které jsou si daleko více podobny než náhodné výběry téhož rozsahu (v obou výběrech jsou obsaženy stejné školy). Z uvedeného vyplývá, že tento experimentální plán poněkud přeceňuje chyby způsobené rozdíly mezi skupinami a rozdíly mezi dosaženými výsledky se tudíž mohou jevit jako méně významné, než ve skutečnosti jsou.
- **Plán 5 (experiment spárovaných skupin).** Tento experimentální plán probíhá na jedné škole, na které jsou k dispozici dvě třídy žáků. Před zahájením experimentu je zadán žákům didaktický test, který měří úroveň jejich vědomostí v oblasti, ve které bude probíhat experiment. Na základě výsledků tohoto měření se potom vytvoří dvě třídy (spárované výběry), v nichž bude stejné rozdělení skórů (počtu bodů) – srov. oddíl 1.2.3.1. Experiment se vyhodnocuje na základě srovnávání výsledků obou tříd na konci experimentu.

  Tento experimentální plán poskytuje mnohem přesnější výsledky než plán 1 nebo 2. Jeho nevýhodou je však nutnost vytváření skupin (nových tříd) podle výsledků vstupního testu, což v praxi naráží na nemalé obtíže. Proto se experimentální plán v této podobě realizuje jen výjimečně. Moderní statistické postupy (analýza rozptylu, analýza kovariance) umožňují uskutečnit stejně spolehlivý experiment i bez nutnosti přeskupování žáků do nových (spárovaných) skupin (Lindquist, 1967; Komenda, 1981; Mittenecker, 1968).
- **Plán 6.** V tomto případě se plán 5 opakuje například na deseti školách. Výsledky všech žáků vyučovaných jednou metodou se srovnávají s výsledky všech žáků vyučovaných druhou metodou.

  Tento plán, který je nejspolehlivější ze všech uvedených experimentálních plánů, má význam spíše teoretický, protože v popisované podobě je v praxi jen velmi obtížně realizovatelný.

Technika rotace faktorů je v podstatě kombinací techniky jedné skupiny a techniky paralelních skupin, přičemž do jisté míry zachovává výhody obou postupů. Pracuje se se dvěma nevyrovnanými skupinami jedinců, přičemž experiment probíhá ve dvou fázích. V první fázi je prováděn experimentální zásah v jedné skupině, druhá skupina slouží v této fázi jako kontrolní skupina. Ve druhé fázi experimentu se funkce skupin mění tak, že skupina, ve které byl realizován experimentální zásah, se stává skupinou kontrolní a skupina, která byla v první fázi skupinou kontrolní, se stává experimentální.

Použití techniky rotace faktorů budeme ilustrovat na příkladu výzkumu, ve kterém se ověřovala hypotéza, že provádění frontálních žákovských pokusů ovlivňuje v kladném smyslu úroveň vědomostí žáků z fyziky. Experiment se uskutečnil na škole se dvěma paralelními třídami, přičemž žáci nebyli do tříd nijak speciálně vybíráni (úroveň vědomostí žáků z fyziky nebyla ve třídách stejná). V obou třídách bylo nejdříve provedeno počáteční měření úrovně vědomostí žáků z fyziky (m_{A1} a m_{B1}). Potom byl ve třídě A proveden experimentální zásah (vyučování bylo organizováno tak, že žáci frontálně prováděli pokusy), třída B byla vyučována tradičním způsobem (učitel prováděl jen demonstrační pokusy). Tato fáze experimentu, která trvala čtyři týdny, byla ukončena měřením vědomostí žáků (m_{A2} a m_{B2}). Po určité době (až bylo probíráno učivo, které umožňovalo provádění frontálních žákovských pokusů) byla zahájena druhá fáze experimentu, která trvala opět čtyři týdny a začala počátečním měřením úrovně vědomostí žáků v obou třídách (m_{A3} a m_{B3}). Poté byl realizován experimentální zásah ve třídě B, přičemž třída A plnila funkci kontrolní skupiny. Nakonec bylo provedeno konečné měření úrovně vědomostí žáků v obou skupinách (m_{A4} a m_{B4}). Celý průběh experimentu s využitím techniky rotace faktorů zachycuje tabulka 1.

Ze schématu je možno určit, o kolik vzrostly vědomosti žáků vlivem experimentálního zásahu. Tento nárůst činí

$$(m_{A2} - m_{A1}) + (m_{B4} - m_{B3})$$

Podobně můžeme určit, o kolik vzrostly vědomosti žáků bez experimentálního zásahu

$$(m_{B2} - m_{B1}) + (m_{A4} - m_{A3})$$

Změnu, kterou způsobil experimentální zásah, můžeme vyjádřit jako rozdíl

$$\left[(m_{A2} - m_{A1}) + (m_{B4} - m_{B3})\right] - \left[(m_{B2} - m_{B1}) + (m_{A4} - m_{A3})\right] \tag{3}$$

Na velikosti tohoto rozdílu bude záviset, zda přijmeme, či odmítneme příslušnou hypotézu. Bude-li tento rozdíl skutečně významný (ve prospěch výsledků při experimentování), můžeme hypotézu přijmout.

***Tab. 1** Schéma experimentální techniky rotace faktorů*

	Skupina A	**Skupina B**
první fáze	počáteční měření m_{A1} experimentální zásah měření m_{A2}	počáteční měření m_{B1} – měření m_{B2}
druhá fáze	počáteční měření m_{A3} – konečné měření m_{A4}	počáteční měření m_{B3} experimentální zásah konečné měření m_{B4}

Technika rotace faktorů má výhodu v tom, že nemusí pracovat s vyrovnanými (spárovanými) skupinami, a také, že není třeba kontrolovat neexperimentální faktory. Nedostatkem je však například to, že zpravidla záleží na pořadí, ve kterém je ve skupinách proveden experimentální zásah. Rotace faktorů se tak stává další proměnnou, která v experimentu spolupůsobí. Další podmínkou smysluplného použití techniky rotace faktorů je, že obě experimentální manipulace musí probíhat ve srovnatelných situacích. Dodržení této podmínky nebývá v reálných výzkumech jednoduché.

1.4 KVALITATIVNĚ ORIENTOVANÉ VÝZKUMY

Vedle kvantitativně orientovaných výzkumů se ve světě i u nás v posledních letech velmi silně rozvíjejí tzv. **kvalitativně orientované výzkumy**. Mezi výzkumy kvalitativně a kvantitativně orientovanými jsou některé důležité rozdíly.

Základním rozdílem je, že obě orientace vycházejí z odlišných filozofických základů. Kvantitativně orientované (pozitivistické, klasické) pedagogické výzkumy vycházejí, jak již bylo uvedeno, z **pozitivismu**, respektive **novopozitivismu**. Z této filozofie vyplývá přesvědčení o existenci jedné objektivní reality, která není závislá na našich názorech, citech, postojích nebo přesvědčení. Kvalitativně orientované výzkumy vycházejí naproti tomu zejména z **fenomenologie**, která zdůrazňuje subjektivní aspekty jednání lidí, a tudíž kvalitativně orientované výzkumy připouštějí existenci více realit. Z rozdílných filozofických východisek vyplývají i další rozdíly (cíle výzkumu, přístup badatele apod.).

Nejdůležitější rozdíly mezi kvalitativně orientovaným a kvantitativně orientovaným pedagogickým výzkumem uvádí tabulka 2.

Tyto dvě orientace současných pedagogických výzkumů jsou natolik rozdílné, že jsou mnohdy stavěny proti sobě jako naprosto protikladné a navzájem neslučitelné. Zastáváme názor, že každý z těchto přístupů má své přednosti i nedostatky a že je nejen možné (ale i výhodné) oba v konkrétních výzkumech kombinovat.

***Tab. 2** Srovnání kvalitativně a kvantitativně orientovaného výzkumu (podle P. Gavory, 2000)*

Kvantitativně orientovaný výzkum	Hledisko	Kvalitativně orientovaný výzkum
pozitivismus	filozofická východiska	fenomenologie
jedna realita	existence reality	více realit
vysvětlení jevu	cíle výzkumu	porozumění smyslu
číslo velké skupiny osob zobecnění odstup	přístup	slovo, význam malé skupiny osob jedinečnost vcítění se

2. MĚŘENÍ V PEDAGOGICKÉM VÝZKUMU

2.1 MĚŘENÍ A JEHO DRUHY

Jestliže chceme při studiu pedagogické reality uplatňovat kvantitativní přístup, je třeba, abychom u každého studovaného jevu dokázali vedle postižení jeho kvality zachytit i jeho kvantitu, tj. zachytit jeho velikost nebo množství měřením.

Měření v nejširším smyslu slova je „přiřazování čísel předmětům nebo jevům podle pravidel“ (Kerlinger, 1972). Při tomto vymezení měření je podstatné to, že přiřazování se děje podle jistých pravidel. Pravidla mohou být různě dokonalá a na tom, zda jsou „dobrá“ nebo „špatná“, pochopitelně záleží, zda výsledky měření budou dobré (věrohodné) či špatné (nevěrohodné, chudé).

Stanovení pravidla pro přiřazování je pro kvalitu měření nejdůležitější. Formálně vyjádřeno, jde o to nalézt funkci (pravidlo) pro přiřazování prvků množiny měřených objektů k prvkům množiny čísel. Tuto funkci můžeme obecně vyjádřit zápisem

$$f = \{(x, y)\}.$$

Při zkoumání pedagogické reality se často ocitáme v situaci, kdy proměnná, kterou chceme zachytit, není přímo měřitelná (např. charakteristiky, jako tvořivost, morálka, hostilita apod.). V těchto případech se často uchylujeme k měření tzv. **indikátorů** (ukazatelů), tj. jiných jevů, které s velkou pravděpodobností s danou proměnnou souvisejí. V této souvislosti se také hovoří o tzv. **operacionalizovaných definicích proměnných** (srov. oddíl 1.2.1.2). Určitá proměnná může být operacionálně definována pomocí jednoho nebo i několika indikátorů.

Pokud má měření věrohodně zachycovat vlastnosti měřených objektů, je třeba, aby byly splněny tři základní **postuláty měření**:

- **1. postulát:** Při měření musíme být schopni rozhodnout, zda určitý objekt v daném směru je, nebo není stejný jako jiný objekt. Tuto podmínku můžeme psát:

buď ($a = b$), nebo $a \neq b$, avšak ne oboje

- **2. postulát:** Jestliže objekt a je v daném smyslu roven objektu b a objekt b je roven objektu c, potom objekt a je roven objektu c. Tuto podmínku lze zachytit zápisem:

jestliže [($a = b$) a současně (b = c)], pak ($a = c$)

- **3. postulát:** Jestliže objekt a je větší než objekt b a objekt b je větší než objekt c, potom objekt a je větší než objekt c. Tuto podmínku zachycuje zápis:

jestliže [($a > b$) a současně ($b > c$)], potom ($a > c$)

Při měření v přírodních vědách nebo v technice většinou automaticky očekáváme, že uvedené postuláty platí. V psychologických nebo pedagogických měřeních však platnost těchto podmínek bývá často sporná, a proto by se měla ověřovat. Například při zkoumání postavení jednotlivých členů rodiny můžeme zjistit, že žena dominuje nad manželem ($a > b$), že manžel současně dominuje nad dítětem ($b > c$), ale dítě dominuje nad svojí matkou ($c > a$).

Podle charakteru prováděného přiřazování čísel lze rozlišit čtyři úrovně měření: nominální (klasifikace), ordinální (pořadové), intervalové a poměrové.

Měření nominální

Při tomto měření se užívá čísel pouze jako označení („nálepek") pro určité charakteristiky. Někteří autoři jej proto za měření nepovažují a používají označení klasifikace. Příkladem nominálního měření je například postup, kdy zaznamenáváme pohlaví dětí tím způsobem, že chlapcům přiřazujeme číslo 1 a dívkám číslo 2. U nominálního měření čísla nemají kvantitativní význam a nelze s nimi jako s čísly počítat. Počítat ale lze s četnostmi jednotlivých číselných symbolů. S nominálním měřením se často setkáváme například v pedagogických výzkumech, které využívají data získaná pomocí dotazníků.

Použitelné numerické operace a statistika

Je možné sčítat a odčítat počty případů (četnosti) v každé kategorii, lze určovat modus a některé míry variability (např. nominální varianci), je použitelná frekvenční statistika typu chí-kvadrát, Fisherův test, výpočet procent, stanovení koeficientů kontingence apod.

Měření ordinální (pořadové)

U tohoto měření se objektům přiřazují čísla tak, že vyjadřují pořadí podle určitého kritéria. Například můžeme dětem ve skupině přiřadit čísla podle toho, v jakém pořadí splnily určitý úkol. Tato čísla potom poskytují informaci pouze o pořadí měřených objektů (dětí, situací), nikoli o velikostech rozdílů mezi nimi.

Použitelné numerické operace a statistika

Je možné počítat medián a některé míry variability (např. kvartilovou odchylku), Spearmanův koeficient pořadové korelace, Kendallův koeficient shody, Wilcoxonův test, U-test, Kolmogorovův-Smirnovův test, Kruskalův-Wallisův test apod.

Měření intervalové

Měříme-li objekty na úrovni intervalového měření, přiřazujeme čísla tak, že vyjadřují rozdíly mezi měřenými objekty. Tento druh měření však nemá přirozený nulový bod, nula je na intervalové stupnici stanovena pouze arbitrárně. Čísla získaná intervalovým měřením je možno sčítat a odečítat, nelze je však násobit nebo dělit. Příkladem intervalového měření je například měření úrovně vědomostí didaktickým testem (platí přibližně).

Použitelné numerické operace a statistika

Je možné určovat aritmetický průměr a směrodatnou odchylku, je možné používat Studentův t-test, párový t-test, F-test, analýzu rozptylu, Pearsonův koeficient korelace apod.

Měření poměrové

U poměrového měření přiřazené hodnoty vyjadřují množství vlastnosti, kterou měří. Poměrové měření má (na rozdíl od intervalového) přirozenou nulu. U měření poměrového můžeme využívat všech vlastností reálných čísel, získané hodnoty můžeme sčítat, odčítat, násobit i dělit. Jednotlivé výsledky poměrového měření lze srovnávat na základě otázek „o kolik", ale i „kolikrát". Příkladem poměrového měření je například měření hmotnosti dětí (vážení), určování věku atd. Měření intervalová a poměrová bývají označována souborně jako **měření metrická**. Při měření v pedagogických výzkumech se na úroveň poměrového měření dostáváme jen zřídka (většinou jen při měření proměnných, jako věk dítěte, charakteristiky tělesného vývoje apod.).

Použitelné numerické operace a statistika

Po vhodné kategorizaci lze používat všechny statistické postupy, které byly uvedeny v souvislosti s měřením ordinálním i měřením nominálním. Použijeme-li však pro poměrová data postupů určených pro data ordinální nebo nominální, dochází vždy k určité ztrátě informace.

Někdy se používá statistických postupů určených pro metrická data i v případech, kdy není plně zaručeno, že zpracovávaná data této úrovni měření odpovídají (např. některé škály, školní klasifikace apod.). V těchto případech podstupujeme určité riziko zkreslení výsledků, které při analýze získáme.

2.2 VLASTNOSTI DOBRÉHO MĚŘENÍ

Jestliže realizujeme určité pedagogické měření, nikdy si nemůžeme být dopředu jisti jeho kvalitou. Skutečnou kvalitu měření lze zpravidla dostatečně posoudit až na základě vyhodnocení výsledků již uskutečněného měření. Při posuzování vlastností měření nás obvykle nejvíce zajímá jeho validita, reliabilita a praktičnost.

Validita měření

Českým ekvivalentem pojmu „validita" je „platnost". Měření má dobrou validitu tehdy, jestliže měří skutečně to, co podle předpokladu měřit má. Pro exaktní posouzení validity měření je třeba mít k dispozici nějaké jiné vnější kritérium (např. jiné měření, u něhož je validita nesporná), se kterým se dané měření srovnává. Podle toho, k čemu se validita vztahuje, lze rozlišit validitu:

- **obsahovou** (posuzujeme, do jaké míry se měří stanovený obsah);
- **souběžnou** (posuzujeme, do jaké míry se měření shoduje s jiným měřením týchž objektů);
- **predikční** (posuzujeme, do jaké míry provedené měření vypovídá o budoucím vývoji objektů);

- **konstruktovou** (pojmovou, teoretickou), u které posuzujeme, do jaké míry ovlivňuje výsledky provedeného měření nějaký faktor – konstrukt).

Reliabilita měření

Pojem „reliabilita“ se často nahrazuje termíny „spolehlivost“, „stabilita“, „homogenita“, „přesnost“, „konzistence“ nebo „stálost“, avšak žádný z nich pojem „reliabilita“ plně nevystihuje. Aby měření bylo reliabilní, je třeba, aby při opakování za stejných podmínek poskytovalo stejné (zhruba stejné) výsledky. Tento aspekt reliability je možno označit jako **spolehlivost měření**. V některých případech je reliabilita chápána jen v tomto zúženém smyslu. Jestliže však chápeme reliabilitu širším způsobem, potom požadujeme, aby měření vedle spolehlivosti bylo ještě navíc **přesné**, tj. minimálně zatíženo chybami měření. Za přesné považujeme takové měření, při kterém se dopouštíme jen malého počtu chyb, a tyto chyby nejsou příliš velké. Oba uvedené aspekty, tj. spolehlivost a přesnost, zahrnujeme pod společný pojem „reliabilita měření“. Dostatečně vysoká reliabilita je nutnou podmínkou dobré validity měření, vysoká reliabilita však ještě nezaručuje dobrou validitu.

Stupeň reliability měření se vyjadřuje **koeficientem reliability**. Je to číslo, které může nabývat hodnot od 0 do +1, přičemž platí, že 0 vyjadřuje nulový stupeň reliability a 1 vyjadřuje maximální (ideální) stupeň reliability.

Koeficient reliability je možno určovat mnoha způsoby. Uvedeme jen některé, v praxi často používané postupy:

- **Metoda opakovaného měření.** Měření se provádí opakovaně (stejným měrným nástrojem) za stejných podmínek a koeficient reliability se určuje jako koeficient korelace pro obě provedená měření. Tento způsob stanovení reliability, který postihuje aspekt spolehlivosti měření, není v praxi příliš častý, protože je velmi obtížné (ne-li nemožné) zajistit dvakrát po sobě stejné podmínky pro měření.
- **Metoda paralelního měření.** Měření se provádí opakovaně, za použití různých (ale ekvivalentních) měrných nástrojů (např. určitý problém se opakovaně zkoumá pomocí dvou dotazníků, které se různými způsoby dotazují na tutéž problematiku). Stupeň reliability se v tomto případě vyjadřuje pomocí koeficientu korelace pro obě realizovaná měření. Tato metoda určování reliability postihuje aspekt spolehlivosti měření a je pro svoji náročnost v praxi spíše výjimkou.
- **Metoda půlení** (*half-split method*). U této metody se výsledky provedeného měření (např. výsledky didaktického testu) rozdělují na dvě poloviny (např. výsledky v lichých a sudých položkách testu). Výsledky měření dosažené v obou polovinách měrného nástroje se potom korelují a ze stupně korelace se vychází při stanovení koeficientu reliability. Podrobnosti k výpočtu lze nalézt například v práci Chrásky (1999).
- **Výpočet koeficientu reliability pomocí Kuderova-Richardsonova vzorce.** U této metody výpočtu koeficientu reliability (často používané při stanovení reliability didaktických testů) se vychází z celkového počtu úloh v testu, z obtížnosti jednotlivých testových úloh a z variability provedeného měření (směrodatné odchylky). Podrobnosti výpočtu lze opět najít například v práci Chrásky (1999).
- **Stanovení reliability pomocí Cronbachova koeficientu alfa.** Tato metoda vychází z tzv. dvojnásobné analýzy rozptylu a bývá dostupná při zpracovávání výsledků mě-

ření pomocí počítačových statistických programů (např. Statistica.cz: vícerozměrné průzkumné techniky → analýza spolehlivosti).

Určování stupně reliability měření nemá v našich pedagogických výzkumech příliš dlouhou tradici. Pojem „reliabilita" je zatím většinou spojován jen s didaktickými testy, zatímco ostatní druhy měření (např. měření realizované při pozorování, měření dotazníkem apod.) zpravidla nejsou tomuto kritériu podrobovány.

Praktičnost měření

Pro praxi měření mají velký význam i takové vlastnosti, jako jednoduchost, hospodárnost, úspornost, snadná proveditelnost, malá časová náročnost, malé nároky na kvalifikaci osoby, která měření realizuje atd. Tyto vlastnosti měření označujeme společným názvem praktičnost měření.

Někdy se v literatuře uvádějí i další vlastnosti měření, jako například citlivost (senzibilita), objektivita atd. Lze však prokázat, že tyto vlastnosti jsou součástí vlastností výše uvedených.

2.3 METODY ZPRACOVÁNÍ DAT V PEDAGOGICKÝCH VÝZKUMECH

V kvantitativně orientovaných výzkumech získáváme o studovaných jevech zpravidla velké množství číselných údajů (dat). Abychom z naměřených dat mohli vyčíst potřebné informace, je nutné je nejdříve zpracovat. Při zpracování výsledků pedagogických výzkumů se zpravidla realizují následující kroky:

- uspořádání dat a sestavení tabulek četností;
- grafické znázornění naměřených dat;
- výpočet charakteristik polohy (měr ústřední tendence);
- výpočet charakteristik rozptýlení (měr variability).

2.3.1 USPOŘÁDÁNÍ DAT A SESTAVOVÁNÍ TABULEK ČETNOSTÍ

Základní utřídění dat lze provést například pomocí tzv. **čárkovací metody**. Postup budeme ilustrovat na příkladu, ve kterém se jedná o zpracování metrických dat (výsledky žáků v didaktickém testu).

Příklad 3

Při měření vědomostí žáků didaktickým testem byly získány následující výsledky (počet bodů):

11, 8, 7, 10, 10, 6, 10, 12, 6, 9, 8, 8, 9, 10, 11, 10, 9, 9, 7, 11, 7, 8, 9, 10, 9, 8, 9, 12, 9, 5.

Pomocí čárkovací metody sestavíme tabulku četností.

Při použití čárkovací metody nejdříve zapíšeme do sloupce (vlevo) všechny hodnoty, jichž bylo při měření dosaženo. Hodnoty přitom uvádíme seřazené podle velikosti (od nejmenší po největší). Následně procházíme jednotlivé hodnoty a pomocí čárek zaznamenáváme jejich výskyt. Výsledky čárkovací metody potom převedeme do tabulky četností (tab. 3). Počet žáků, kteří dosáhli určitého výsledku v didaktickém testu, označujeme jako **četnost**. Pokud není uvedeno jinak, rozumí se pod označením „četnost" tzv. **absolutní četnost**.

Tab. 3 *Sestavení tabulky četností čárkovací metodou*

Výsledek v testu (počet bodů)	Žáci, kteří daného výsledku dosáhli	Četnost n_i	Relativní četnost f_i	Kumulativní četnost
5	/	1	0,033	1
6	//	2	0,067	3
7	///	3	0,100	6
8	/////	5	0,166	11
9	////////	8	0,267	19
10	//////	6	0,200	25
11	///	3	0,100	28
12	//	2	0,067	30
		Σ 30	Σ 1,000	

Někdy bývá účelné doplnit tabulku četností ještě o tzv. **relativní četnosti**. Relativní četnost f_i je podíl četnosti (absolutní) n_i a celkové četnosti n

$$f_i = \frac{n_i}{n} \tag{4}$$

Relativní četnosti je možno také vyjádřit v procentech. V tom případě se vypočítaná hodnota f_i vynásobí 100 %. Relativní četnost poskytuje informaci o tom, jak velká část z celkového počtu hodnot připadá na danou hodnotu (kategorii hodnot).

Pro některé statistické analýzy je možné tabulku četností doplnit ještě o tzv. **kumulativní četnosti** n_k. Kumulativní četnost vyjadřuje četnost v určitém řádku tabulky a četnosti ve všech předchozích řádcích dohromady.

Jestliže byl při měření získán velký počet různých hodnot (nebo v případě měření tzv. spojitých náhodných veličin), tabulka četností by obsahovala příliš velký počet řádků a stávala by se nepřehlednou. V těchto případech se většinou získaná data seskupují do tzv. **intervalů**. Většinou se uvádí, že počet intervalů by neměl být větší než 20 a ne menší než 6.

Nejvýhodnější hloubku (šířku) intervalu lze přibližně stanovit podle celé řady empirických vzorců, např.

$$h \approx 0{,}08 \cdot R \tag{5}$$

nebo

$$\frac{R}{24} < h < \frac{R}{12} \tag{6}$$

kde h je hloubka intervalu a R je tzv. **variační šíře** (rozdíl mezi největší a nejmenší naměřenou hodnotou). Vypočítanou hodnotu h je zpravidla nutno zaokrouhlit na vhodné celé číslo. Při stanovení intervalů je třeba dbát na to, aby krajní intervaly (první a poslední) nebyly prázdné, a na to, aby všechny intervaly měly stejnou hloubku.

Sestavení tabulky četností s použitím intervalů a výpočet relativních četností budeme ilustrovat na příkladu.

Příklad 4

Na základní škole se uskutečnil výzkum, ve kterém se měřila výška žáků. Při měření skupiny 31 žáků byly získány následující hodnoty (cm):

144, 149, 145, 142, 146, 147, 141, 150, 143, 146, 150, 141, 148, 148, 144, 141, 145, 148, 144, 143, 155, 133, 158, 154, 151, 140, 136, 137, 153, 139, 138

Vzhledem k poměrně velké variační šíři výsledků (R = 158 – 133 = 25) bude v daném případě vhodné výsledky zapisovat do tabulky četností s intervaly. Podle vzorce (5) vychází optimální hloubka intervalu

$$h \approx 25 \cdot 0{,}08 = 2$$

a podle vzorce (6) velmi podobný výsledek

$$\frac{25}{24} = 1{,}04 < h < \frac{25}{12} = 2{,}08$$

Tab. 4 *Tabulka četností s intervaly*

Výška žáků (cm)	Četnost	Střed intervalu x_i	Kumulativní četnost n_i
133–134	1	133,5	1
135–136	1	135,5	2
137–138	2	137,5	4
139– 140	2	139,5	6
141–142	4	141,5	10
143–144	5	143,5	15
145–146	4	145,5	19
147–148	4	147,5	23
149–150	3	149,5	26
151–152	1	151,5	27
153–154	2	153,5	29
155–156	1	155,5	30
157–158	1	157,5	31
	∑ 31		

Abychom mohli s výsledky (které jsou v tabulce četností uvedeny) počítat, doplňujeme tabulku ještě o další sloupec, který označujeme jako **střed intervalu**. Vzhledem k tomu, že z tabulky četností, která používá intervaly, nelze přesně určit jednotlivé hodnoty, předpokládáme, že hodnoty jsou v intervalech rozloženy „rovnoměrně" se středem v polovině intervalu.

Poznámka: *Při výpočtu aritmetického průměru pomocí středů intervalů se dopouštíme určité chyby, která však většinou není příliš velká (zvláště u souborů dat velkého rozsahu). Pokud zpracováváme soubory dat malého rozsahu a pokud jednotlivé naměřené hodnoty známe, můžeme místo středů intervalů vypočítat pro každý interval průměr hodnot v něm obsažených* ($\overline{x_i}$) *a aritmetický průměr celého souboru dat potom počítat jako průměr z průměrů v jednotlivých intervalech.*

Statistické charakteristiky se zpravidla vypočítávají (při počítačovém zpracování také tisknou) na více desetinných míst, než kolik jich obsahují vstupní údaje. Vypočítané hodnoty proto zpravidla zaokrouhlujeme na dvě až tři platné číslice.

V tabulkách četností by každé pole mělo být vyplněno údajem (číslicí nebo symbolem). Většinou se dodržuje zásada, že vodorovná čárka (–) vyjadřuje skutečnost, že daná hodnota se nevyskytla. Nula (s příslušným počtem desetinných míst) informuje o tom, že naměřená hodnota je tak malá, že ji můžeme považovat za nulovou. Tečka (.) vyjadřuje skutečnost, že hodnotu neznáme.

Sestavování tabulek četností, ale i další statistické operace, nám může usnadnit výpočetní technika. Možnosti využití některých statistických výpočetních systémů při zpracování výzkumných dat budeme uvádět v poznámkách, vždy po vysvětlení příslušných procedur.

Možnosti analýzy na PC
Excel: vložit funkci → statistické → četnosti
Statistica.cz: základní statistiky a tabulky → tabulky četností
SPSS: *descriptive statistics* → *frequencies*

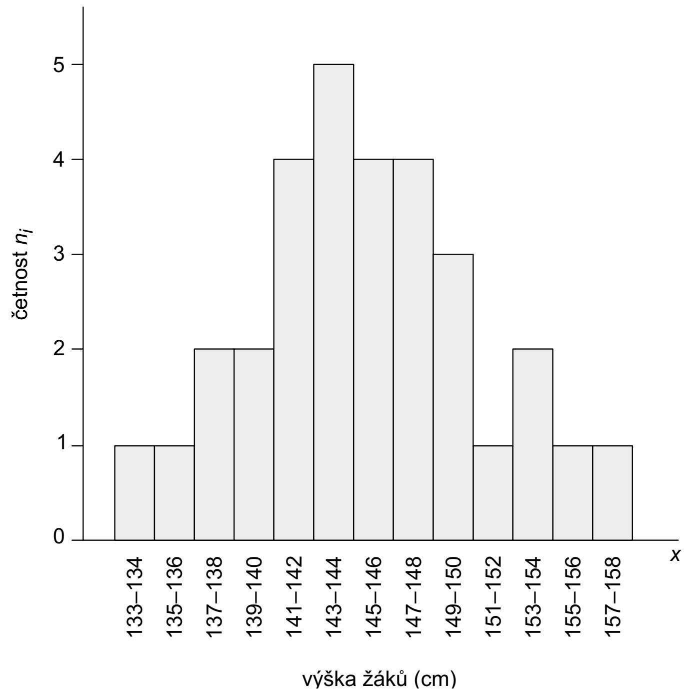
2.3.2 GRAFICKÉ METODY ZOBRAZOVÁNÍ DAT

Data obsažená v tabulce četností je možno prezentovat také v názorné podobě. K tomuto účelu se používají například **histogramy četností, polygony četností, součtové křivky, výsečové grafy** apod.

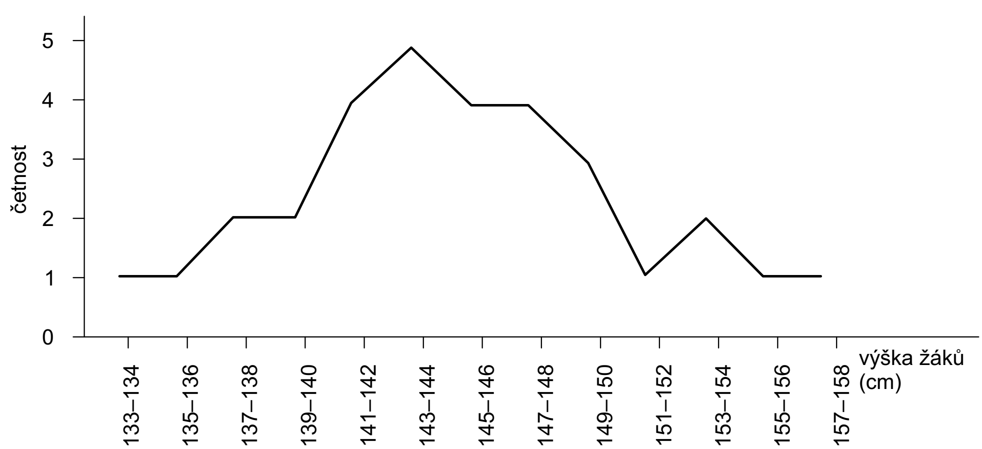
Histogram četností je sloupcový diagram, u kterého na vodorovné ose (x) zobrazujeme jednotlivé naměřené hodnoty (případně intervaly hodnot) a na svislé ose (y) četnosti hodnot n_i. **Polygon četností** se liší od histogramu tím, že se jedná o diagram spojnicový. Obrázky 2 a 3 uvádějí histogram a polygon četností pro výsledky z tabulky 4.

Někdy bývá výhodné graficky znázornit také závislost mezi dosaženými výsledky a kumulativními četnostmi. Graf kumulativních četností bývá označován jako **součtová křivka** nebo **Galtonova ogiva** (obr. 4).

Výsečového diagramu můžeme využít například k názornému zobrazení struktury složení výběrového souboru (obr. 5).

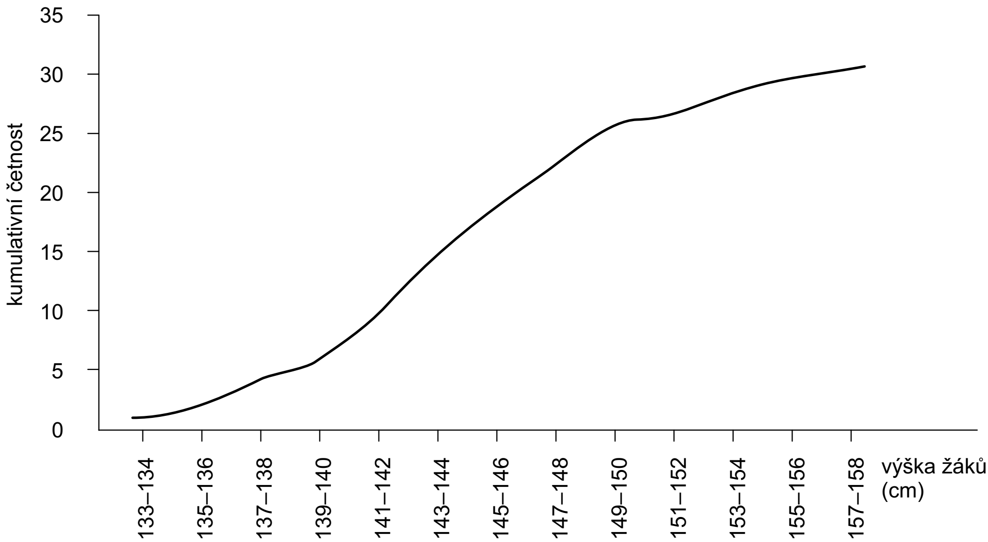


***Obr. 2** Histogram četností*

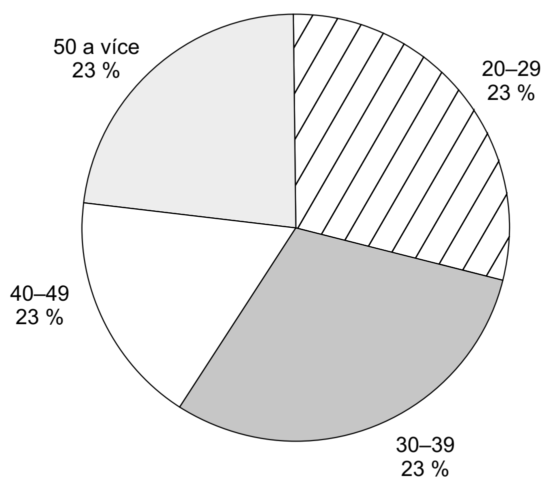


***Obr. 3** Polygon četností*

Možnosti analýzy na PC
Excel: průvodce grafem → sloupcový, spojnicový, výsečový
Statistica.cz: grafy → 2D grafy → histogramy, spojnicové grafy, výsečové grafy



Obr. 4 *Graf kumulativních četností (Galtonova ogiva)*



Obr. 5 *Věková struktura výběrového souboru (počet roků)*

2.3.3 CHARAKTERISTIKY POLOHY (MÍRY ÚSTŘEDNÍ TENDENCE)

Při zpracování hromadných dat zpravidla potřebujeme všechna naměřená data nějakým způsobem výstižně a stručně charakterizovat (jinak řečeno, potřebujeme určit hodnotu, která by všechny naměřené hodnoty dobře „reprezentovala“). V pedagogických výzkumech se k tomuto účelu nejčastěji užívá **aritmetický průměr** (u dat metrických), **medián** (u dat ordinálních) nebo **modus** (u dat nominálních).

Poznámka: *Pomocí aritmetického průměru (respektive mediánu nebo modu) odhadujeme střední hodnotu základního souboru, jejíž skutečná hodnota zpravidla není známá (srov. oddíl Rozsah výběru). Střední hodnota se zpravidla označuje řeckým písmenem μ. V dalším textu nebudeme (pro jednoduchost) rozlišovat mezi střední hodnotou μ a jejími odhady pomocí výběrových charakteristik (aritmetický průměr $\overline{x}$, medián $\widetilde{x}$, modus $\widehat{x}$).*

2.3.3.1 Aritmetický průměr

Aritmetický průměr $\overline{x}$ z číselných hodnot $x_1, x_2, x_3, \ldots$ lze vypočítat podle vzorce

$$\overline{x} = \frac{x_1 + x_2 + x_3 + \ldots + x_n}{n} \tag{7}$$

kde n je celková četnost všech hodnot.

Pro součet všech hodnot x_i ($i = 1, 2, 3, \ldots n$) užíváme často znaku $\sum_{i=1}^{n} x_i$.

Vzorec pro výpočet aritmetického průměru můžeme potom psát ve tvaru

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \tag{8}$$

Ve výzkumech často počítáme aritmetický průměr z dat, která jsou obsažena v tabulce četností. V těchto případech lze aritmetický průměr vypočítat ze vzorce

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{k} n_i \cdot x_i \tag{9}$$

kde n je celková četnost všech hodnot, x_i je určitá hodnota, n_i je četnost hodnoty x_i a k je počet řádků v tabulce četností.

Příklad 5

Výpočet aritmetického průměru z tabulky četností, ve které nebyly použity intervaly, uvádí tabulka 5.

Pro hodnoty uvedené v tabulce 5 vychází aritmetický průměr

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{k} n_i \cdot x_i = \frac{267}{30} = 8{,}9$$

Průměrný počet bodů v didaktickém testu je 8,9 bodu.

Tab. 5 *Výpočet aritmetického průměru z tabulky četností (bez užití intervalů)*

Výsledek v testu (počet bodů)	**Četnost n_i**	**$n_i \cdot x_i$**
5	1	5
6	2	12
7	3	21
8	5	40
9	8	72
10	6	60
11	3	33
12	2	24
	Σ 30	Σ 267

Jestliže jsou v tabulce četností data seskupena do intervalů, potom postupujeme tak, že nejdříve určíme střed každého intervalu a ten potom dosadíme jako hodnotu x_i do příslušného vzorce.

Příklad 6

Výpočet aritmetického průměr z tabulky četností, ve které byly použity intervaly, uvádí tabulka 6. Nejdříve vypočítáme středy intervalů x_i . Pro první řádek tabulky bude například střed intervalu

$x_i = \frac{133+134}{2} = 133{,}5$, pro druhý řádek $x_i = \frac{135+136}{2} = 135{,}5$ atd.

Pro hodnoty uvedené v tabulce 6 potom vychází aritmetický průměr

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} n_i \cdot x_i = \frac{4498{,}5}{31} = 145{,}11$$

Průměrná výška žáků je asi 145 cm.

Tab. 6 *Výpočet aritmetického průměru z tabulky četností (s použitím intervalů)*

Výška žáků (interval)	**Četnost n_i**	**Střed intervalu x_i**	**$n_i \cdot x_i$**
133–134	1	133,5	133,5
135–136	1	135,5	135,5
137–138	2	137,5	275,0

▶

Výška žáků (interval)	**Četnost** n_i	**Střed intervalu** x_i	$n_i \cdot x_i$
139–140	2	139,5	279,0
141–142	4	141,5	566,0
143–144	5	143,5	717,5
145–146	4	145,5	582,0
147–148	4	147,5	590,0
149–150	3	149,5	448,5
151–152	1	151,5	151,5
153–154	2	153,5	307, 0
155 –156	1	155,5	155,5
157–158	1	157,5	157,5
	Σ 31		Σ 4 498,5

Výhodou aritmetického průměru je především to, že jeho matematické vyjádření je jednoduché, a také, že je použitelný při odvozování dalších důležitých vztahů. Mezi výhody lze počítat i to, že jeho hodnota závisí na všech prvcích souboru dat. Nevýhodou aritmetického průměru je však to, že je značně citlivý k tzv. **extrémním hodnotám**, tj. hodnotám, které se od ostatních značně odchylují. Aritmetický průměr je možno počítat z intervalových nebo poměrových (tj. metrických) dat.

V některých situacích (např. tehdy, jestliže chceme postihnout tempo růstu v určité oblasti), se místo aritmetického průměru počítá tzv. **geometrický průměr** $\overline{x}_G$

$$\overline{x}_G = \sqrt[n]{x_1 \cdot x_2 \ldots\ldots x_n} \tag{10}$$

> **Možnosti analýzy na PC**
> Excel: vložit funkci → statistické → průměr
> Statistica.cz: základní statistiky → popisné statistiky

2.3.3.2 Medián

Medián ($\widetilde{x}$) je prostřední hodnota z řady hodnot seřazených podle velikosti. Je to ta hodnota, která rozděluje soubor dat na dvě stejné části (počet hodnot menších nebo stejně velkých jako medián je stejný jako počet hodnot větších nebo stejně velkých jako medián).

Příklad 7

Při měření vědomostí žáků didaktickým testem byly získány následující hodnoty (počty bodů):

14 3 18 4 8 18 4 6 8 10 8.

Pro tyto hodnoty určíme medián.
Nejdříve hodnoty seřadíme podle velikosti: 3 4 4 6 8 8 8 10 14 18 18. Celá řada hodnot má 11 prvků, prostředním prvkem je prvek šestý, tj. číslo 8. Medián proto bude $\tilde{x} = 8$.
Pokud by počet hodnot, u nichž máme určovat medián, byl sudý, určí se medián jako průměr ze dvou prostředních hodnot.
V případě intervalového rozdělení četností (u tabulky četností s intervaly) je možno medián určit na základě interpolace podle vzorce

$$\tilde{x} = L + h \cdot \frac{\frac{n}{2} - n_k}{n_m} \tag{11}$$

kde $\tilde{x}$ je medián, L je dolní hranice tzv. **kritického intervalu**, h je hloubka intervalů, n celková četnost všech hodnot, n_k je kumulativní četnost před dolní hranicí kritického intervalu a n_k je četnost kritického intervalu. Kritický interval je ten interval, ve kterém se nachází medián.

Příklad 8

Výpočet mediánu z tabulky četností s intervaly budeme ilustrovat na stejné situaci, pro niž byl počítán v předcházejícím oddílu aritmetický průměr.

Tabulka 7 obsahuje stejné výchozí údaje jako tabulka 6, je však navíc doplněna sloupcem kumulativních četností. Jak již bylo uvedeno, kumulativní četnost je četnost v určitém řádku tabulky a ve všech předchozích řádcích dohromady. V prvním řádku je tedy kumulativní četnost 1 (před tímto řádkem nepředcházela žádná hodnota), ve druhém řádku tabulky je kumulativní četnost 1 + 1 = 2, ve třetím řádku 2 + 2 = 4 atd.

Při výpočtu mediánu je třeba nejdříve rozhodnout, který interval je kritický. Jak již bylo uvedeno, je to interval, do něhož připadne medián, tj. prostřední z řady hodnot seřazených podle velikosti. V daném případě je v tabulce zachyceno celkem 31 hodnot. Ve sloupci kumulativních četností zjistíme, že prostřední hodnota (tj. hodnota s pořadovým číslem 16) připadá do intervalu 145–146 cm. Tento interval proto označíme jako kritický (v tabulce 7 je označen šedým podtiskem).

Dolní hranice kritického intervalu je v našem případě 144,5, kumulativní četnost před dolní hranicí kritického intervalu je 15, celkový počet hodnot 31 a četnost kritického intervalu 4. Dosadíme-li tyto hodnoty do vzorce (11), dostáváme

$$\tilde{x} = 144{,}5 + 2 \cdot \frac{\frac{31}{2} - 15}{4} = 144{,}75$$

Medián vypočítaný z hodnot udávajících výšku žáků je asi 145 cm.

Tab. 7 *Výpočet mediánu a modu z tabulky četností (s použitím intervalů)*

Výška žáků (interval)	**Četnost n_i**	**Kumulativní četnost**
133–134	1	1
135–136	1	2
137–138	2	4
139–140	2	6
141–142	4	10
143–144	5	15
145–146	4	19
147–148	4	23
149–150	3	26
151–152	1	27
153–154	2	29
155–156	1	30
157–158	1	31
	Σ 31	

Výhodou mediánu je, že není citlivý k extrémním hodnotám, ale také to, že jeho výpočet je někdy možný i v případech, kdy o prvcích souboru dat nemáme úplné informace. K určení mediánu totiž, na rozdíl od aritmetického průměru, není nutné znát všechny hodnoty souboru. Medián lze počítat (určovat) u dat, která mají charakter alespoň dat ordinálních (pořadových).

Možnosti analýzy na PC
Excel: vložit funkci → statistické → medián
Statistica.cz: základní statistiky → popisné statistiky

2.3.3.3 Modus

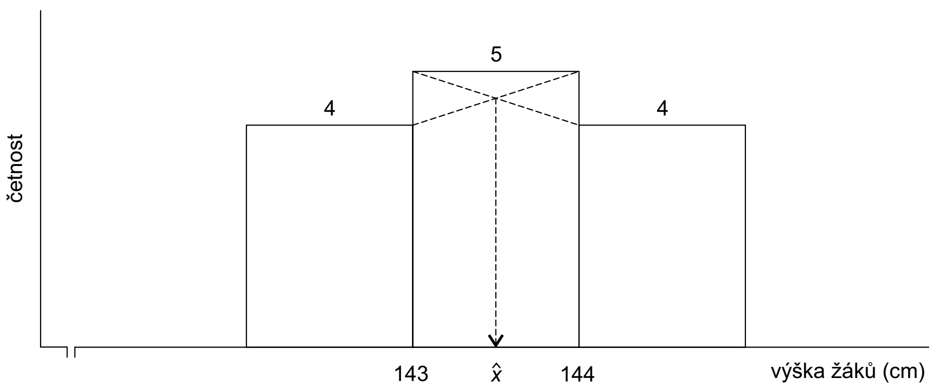
Někdy potřebujeme rychle stanovit alespoň přibližně charakteristiku polohy. K tomuto účelu se dobře hodí **modus**. Modus ($\hat{x}$) je ta hodnota, která se v daném souboru dat vyskytuje nejčastěji (která tedy má největší četnost).

Jestliže známe všechny naměřené hodnoty, lze modus velmi snadno stanovit tak, že zjistíme, která hodnota se v daném souboru dat vyskytuje nejčastěji.

Například v souboru dat, která byla uvedena v souvislosti s určováním mediánu (14 3 18 4 8 18 4 6 8 10 8), je modem hodnota $\hat{x} = 8$ (hodnota 8 je v souboru dat nejčastější).

V případě tabulky četností s intervaly lze modus určit přibližně jako **střed modálního intervalu**. Přesněji lze modus stanovit na základě grafické nebo početní interpolace.

Při grafické interpolaci se postupuje tak, jak je naznačeno na obrázku 6. Vycházíme ze stejné situace, která byla uvažována při výpočtu mediánu. Nejdříve určíme tzv. **modální interval**, což je interval, ve kterém je největší četnost. V tabulce 7 je proto modálním intervalem interval 143–144 cm. Četnost modálního intervalu vyznačíme v obrázku výškou přiměřeně vysokého sloupce. Vedle sloupce vyjadřujícího četnost modálního intervalu zakreslíme ještě sloupce vyjadřující četnosti intervalů sousedních (předcházejícího před kritickým intervalem a následujícího za kritickým intervalem). Další postup interpolace je patrný z obrázku 6.



Obr. 6 *Stanovení modu grafickou interpolací*

Z obrázku je patrné, že modus je asi 143,5 cm.

Početně lze modus určit podle vzorce

$$\hat{x} = L + h \cdot \frac{d_1}{d_1 + d_2} \tag{12}$$

kde $\hat{x}$ je modus, L je dolní hranice modálního intervalu, d_1 je rozdíl mezi četností modálního intervalu a intervalu předcházejícího, d_2 je rozdíl mezi četností modálního intervalu a četností intervalu následujícího a h je hloubka intervalu

Příklad 9

Výpočet modu budeme opět ilustrovat na příkladu měření výšky žáků základní školy (tab. 7).

Z tabulky je patrné, že modálním intervalem je interval 143–144 cm, dolní hranicí modálního intervalu je 142,5, hloubka intervalu je $h = 2$, $d_1 = 1$ a $d_2 = 1$.

Dosadíme-li tyto hodnoty do vzorce (12), dostáváme

$$\hat{x} = 142{,}5 + 2 \cdot \frac{1}{1+1} = 143{,}5$$

Modus vypočítaný z dané tabulky četností je 143,5 cm.

Podobně jako medián je i modus nezávislý na extrémních hodnotách měřené veličiny. Slouží většinou jen jako provizorní charakteristika polohy a neumožňuje další statistickou analýzu. Modus je možno počítat u dat nominálních, ale je použitelný i v případě dat ordinálních nebo metrických.

Určování modu má smysl pouze v případě tzv. jednovrcholového rozdělení (tj. v případě, kdy pouze jedna hodnota má největší četnost). Pokud data získaná ve výzkumu mají dvojvrcholové (bimodální) nebo vícevrcholové rozdělení, potom popsaný způsob určování modu pozbývá smyslu.

Možnosti analýzy na PC
Excel: vložit funkci → statistické → modus
Statistica.cz: základní statistiky → popisné statistiky

2.3.3.4 Vztah mezi charakteristikami polohy

Jestliže srovnáme vypočítané hodnoty aritmetického průměru, mediánu a modu vypočítané ze stejných dat (aritmetický průměr 145, medián 145, modus asi 144), zjišťujeme, že se poněkud liší. Stejných hodnot by tyto střední hodnoty dosáhly v případě přesně symetrického rozdělení četností. Pokud rozdělení není příliš asymetrické, nejsou ani rozdíly mezi charakteristikami polohy příliš velké.

U velkých souborů dat většinou platí přibližný vztah

$$\hat{x} \approx 3\tilde{x} - 2\bar{x} \tag{13}$$

Tento vztah, který byl zjištěn na základě zkušeností, lze použít v případě, že potřebujeme přibližně určit jednu charakteristiku polohy při znalosti zbývajících dvou.

2.3.4 MÍRY VARIABILITY (CHARAKTERISTIKY ROZPTÝLENÍ)

Pomocí charakteristik polohy (měr ústřední tendence) je možno si učinit základní představu o datech, která zpracováváme, ale tato představa není zdaleka úplná. Charakteristika polohy neříká nic o skladbě hodnot, z nichž byla vypočítána. Informaci o tom, jak dalece jsou jednotlivé hodnoty kolem střední hodnoty nakupeny (či naopak rozptýleny), vyjadřují tzv. **míry variability** (charakteristiky rozptýlení).

2.3.4.1 Variační šíře

Pro přibližné posouzení rozptýlení hodnot (posouzení variability) lze užít například **variační šíři** R. Jak již bylo uvedeno v souvislosti s určováním hloubky intervalu v tabulce četností, je to rozdíl mezi největší a nejmenší naměřenou hodnotou

$$R = x_{max} - x_{min} \tag{14}$$

2.3.4.2 Směrodatná (standardní) odchylka

Nejčastěji používanou mírou variability pro data, která byla získána měřením intervalovým nebo poměrovým (metrickým), je směrodatná (standardní) odchylka, respektive rozptyl.

Rozptyl (variance) je **aritmetický průměr čtverců odchylek od aritmetického průměru**. Rozptyl označujeme buď s^2 (v případě, že jej počítáme z hodnot získaných výběrem), nebo σ^2 (v případě, že se vztahuje na celý základní soubor). Výpočet rozptylu pro základní soubor lze provést podle vzorce

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 \tag{15}$$

respektive v případě, že rozptyl počítáme z tabulky četností

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{k}n_i(x_i - \overline{x})^2 \tag{16}$$

kde σ^2 je rozptyl (základního souboru), je celková četnost všech hodnot, x_i je určitá naměřená hodnota, n_i je četnost hodnoty x_i, $\overline{x}$, je aritmetický průměr všech hodnot a k je počet řádků (počet intervalů) v tabulce četností.

Výhodnější však bývá užití upraveného tvaru vzorce

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{k}x_i^2 \cdot n_i - \overline{x}^2 \tag{17}$$

Výpočet podle tohoto (upraveného) vzorce je podstatně snazší, navíc bývá i přesnější (vzhledem k možným zaokrouhlovacím chybám).

Směrodatnou (standardní) odchylku σ vypočítáme jako druhou odmocninu z rozptylu, tj. podle vztahu

$$\sigma = \sqrt{\sigma^2} \tag{18}$$

Pokud máme odhadnout rozptyl základního souboru z dat, která byla zjištěna ve výběru, potom je přesnější použít pro výpočet rozptylu vzorec

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{k}n_i \cdot (x_i - \overline{x})^2 = \frac{1}{n-1}\left[\sum_{i=1}^{k}x_i^2 \cdot n_i - \overline{x}^2 \cdot n\right] \tag{19}$$

Směrodatnou odchylku potom vypočítáme opět jako druhou odmocninu z rozptylu, tj. ze vztahu

$$s = \sqrt{s^2} \tag{20}$$

Příklad 10

Výpočet rozptylu a standardní odchylky (σ i s) budeme ilustrovat na příkladu dat, která byla uvedena v souvislosti s výpočtem aritmetického průměru.

Před výpočtem rozptylu je nejdříve třeba vypočítat aritmetický průměr. Aritmetický průměr pro data, která budeme analyzovat, byl již vypočítán $(\bar{x} = 145{,}1)$.

V tabulce 8 jsou uvedeny všechny potřebné hodnoty pro výpočty podle vzorců (17) i (19).

Tab. 8 *Výpočet rozptylu*

Výška žáků (interval)	Četnost n_i	Střed intervalu x_i	x_i^2	$n_i \cdot x_i^2$
133–134	1	133,5	17 822	17 822
135–136	1	135,5	18 360	18 360
137–138	2	137,5	18 906	37 812
139–140	2	139,5	19 460	38 920
141–142	4	141,5	20 022	80 088
143–144	5	143,5	20 592	102 960
145–146	4	145,5	21 170	84 680
147–148	4	147,5	21 756	87 024
149–150	3	149,5	22 350	67 050
151–152	1	151,5	22 952	22 952
153–154	2	153,5	23 562	47 124
155–156	1	155,5	24 180	24 180
157–158	1	157,5	24 806	24 806
	Σ 31			Σ 653 778

Po dosazení do vzorců dostáváme

$$\sigma^2 = \frac{653\,778}{31} - 145{,}1^2 = 35{,}6 \qquad \sigma = 5{,}97$$

$$s^2 = \frac{1}{31-1} \cdot \left[653\,778 - 145{,}1^2 \cdot 31 \right] = 36{,}8 \qquad s = 6{,}07$$

Srovnáme-li oba výsledky výpočtů směrodatné odchylky, zjišťujeme, že směrodatná odchylka pro výběr (odhad směrodatné odchylky na základě dat z výběru) je poněkud vyšší. Rozlišování vzorců (17) a (19) má smysl zejména při menších četnostech. U velkých četností jsou výsledky u obou postupů prakticky stejné.

Rozptyl a standardní odchylka charakterizují kolísání jednotlivých hodnot kolem aritmetického průměru. Čím více a čím častěji se jednotlivé hodnoty odchylují od aritmetického průměru, tím je rozptyl i standardní odchylka větší. Výpočet rozptylu a standardní odchylky je přípustný pouze v případech, kdy zpracováváme metrická data (intervalová nebo poměrová data).

Možnosti analýzy na PC
Excel: nabídka vložit funkci → statistické → STDEVA; SMODCH
Statistica.cz: základní statistiky → popisné statistiky
SPSS: *descriptive statistics* → *descriptives*

2.3.4.3 Variační koeficient

Jestliže chceme srovnat variabilitu dvou anebo více souborů dat s rozdílnými průměry, můžeme k tomu použít tzv. **variační koeficient** V. Variační koeficient vyjadřuje, kolik procent z průměrné hodnoty směrodatná odchylka činí. Lze jej snadno vypočítat podle vztahu

$$V = \frac{\sigma}{\bar{x}} \cdot 100\ \% \tag{21}$$

kde σ je směrodatná odchylka a $\bar{x}$ aritmetický průměr.

Hodnoty variačního koeficientu V jsou pro různá měření srovnatelné, protože je vyloučen vliv různých jednotek měření i vliv rozdílných průměrů.

2.3.4.4 Kvartilová odchylka

Pokud jsme střední hodnotu souboru dat charakterizovali pomocí mediánu, můžeme jako míru variability použít tzv. **kvartilovou odchylku**.

Při stanovení kvartilové odchylky se (podobně jako při určování mediánu) jednotlivé hodnoty nejdříve seřadí podle velikosti. Ve vytvořené řadě hodnot se potom určí tzv. dolní kvartil Q_1 a horní kvartil Q_3. **Dolní kvartil** Q_1 odděluje dolní čtvrtinu naměřených hodnot, **horní kvartil** Q_3 odděluje horní čtvrtinu naměřených hodnot (medián je vlastně druhý kvartil). Ze známých kvartilů Q_1 a Q_3 se potom vypočítá kvartilová odchylka Q ze vzorce

$$Q = \frac{Q_3 - Q_1}{2} \tag{22}$$

Příklad 11

Při měření vědomostí žáků didaktickým testem byly získány následující hodnoty (počty bodů):

14 3 18 4 8 18 4 6 8 10 8.

Pro tyto hodnoty máme určit kvartilovou odchylku.
Hodnoty nejdříve seřadíme podle velikosti: 3 4 $\boxed{4}$ 6 8 8 8 10 $\boxed{14}$ 18 18. V této řadě hodnot je dolním kvartilem hodnota $Q_1 = 4$ a horním kvartilem $Q_3 = 14$ (medián $\tilde{x} = 8$). Podle vzorce (22) potom vychází kvartilová odchylka

$$Q = \frac{14 - 4}{2} = 5$$

Jestliže máme počítat kvartilovou odchylku z tabulky četností, ve které byly použity intervaly, můžeme oba kvartily vypočítat interpolací (podobně jako u mediánu) podle následujících vzorců

$$Q_1 = L + h \cdot \frac{\frac{1}{4} \cdot n - n_k}{n_m} \tag{23}$$

$$Q_3 = L + h \cdot \frac{\frac{3}{4} \cdot n - n_k}{n_m} \tag{24}$$

Srovnáme-li uvedené vzorce se vzorcem pro výpočet mediánu, zjišťujeme, že jediné rozdíly jsou v čitatelích zlomků. Zatímco při výpočtu mediánu se uvažuje $1/2 \cdot n$, při výpočtech kvartilů se dosazuje buď $1/4 \cdot$ n, nebo $3/4 \cdot n$. Smysl ostatních symbolů ve vzorcích je analogický jako u mediánu. Například ve vzorci pro výpočet Q_1 znamená L dolní hranici intervalu (ve kterém leží dolní kvartil), h je hloubka intervalu, n_k je kumulativní četnost před dolní hranicí intervalu (ve kterém leží dolní kvartil) a n_m je četnost intervalu (ve kterém leží dolní kvartil).

Příklad 12

Výpočet kvartilové odchylky budeme ilustrovat opět na datech z měření výšky žáků. Potřebné hodnoty uvádí tabulka 9.

Bod, který v řadě 31 hodnot odděluje čtvrtinu nejmenších (tj. 31/4 = 7,75 hodnot), leží v intervalu 141–142 (zjistíme ze sloupce kumulativních četností). Dolní kvartil proto bude v tomto intervalu a k němu se také budou vztahovat další hodnoty (L, n_k , n_m) potřebné k výpočtu Q_1 podle vzorce (23). V tabulce 9 jsou intervaly, ve kterých se nacházejí dolní a horní kvartil, označeny podtiskem

$$Q_1 = 140{,}5 + 2 \cdot \frac{\frac{31}{4} - 6}{4} = 141{,}38$$

Obdobně postupujeme i při výpočtu horního kvartilu Q_3 podle vzorce (24)

$$Q_3 = 146{,}5 + 2 \cdot \frac{\frac{3}{4} \cdot 31 - 19}{4} = 148{,}63$$

Kvartilová odchylka je potom

$$Q = \frac{148{,}63 - 141{,}38}{2} = 3{,}63$$

Tab. 9 *Výpočet kvartilové odchylky*

Výška žáků (interval)	Četnost n_i	Kumulativní četnost n_k
133–134	1	1
135–136	1	2
137–138	2	4
139–140	2	6
141–142	4	10
143–144	5	15
145–146	4	19
147–148	4	23
149–150	3	26
151–152	1	27
153–154	2	29
155–156	1	30
157–158	1	31
	Σ 31	

Kvartilová odchylka je mírou rozptýlenosti výsledků kolem mediánu. V případě přibližně symetrického rozdělení četností a při dostatečně velkém souboru dat platí, že v intervalu od $-Q$ do $+Q$ od mediánu se nachází přibližně 50 % všech hodnot. Tato informace může mít význam při popisu daného souboru dat.

2.3.4.5 Míry variability pro nominální data

S výpočtem měr variability u nominálních dat se setkáváme například při zpracování výsledků dotazníkových šetření. Výpočet některých často používaných měr variability objasníme na příkladu.

Příklad 13

V dotazníkovém šetření respondenti odpovídali na otázku:
Jaké máte nejvyšší ukončené vzdělání?

A základní
B středoškolské
C vysokoškolské

V šetření byly získány výsledky, které uvádí tabulka 10.
Máme posoudit, jakou variabilitu mají dosažené výsledky.

Tab. 10 *Stupeň vzdělání respondentů*

Stupeň vzdělání	**Četnost n_i**
A základní	2 121
B středoškolské	1 569
C vysokoškolské	125
	Σ 3 815

Nejjednodušším způsobem lze variabilitu dosažených výsledků posoudit pomocí tzv. **variačního poměru**

$$variační\ poměr = 1 - \frac{n_m}{n} = 1 - \frac{2121}{3815} = 0,44 \tag{25}$$

kde n_m je četnost odpovědí v modální kategorii a n je celková četnost všech odpovědí.

Variační poměr se používá jen k rychlému a orientačnímu posouzení variability. Jeho hodnota závisí na počtu kategorií odpovědí, a proto na základě této míry nelze srovnávat dotazníkové položky s různým počtem kategorií odpovědí. Například pro položku se dvěma kategoriemi odpovědí může nabývat maximální hodnoty 0,5, pro položku s pěti kategoriemi odpovědí může být jeho maximální hodnota 0,8 atd.

Přesnější mírou variability pro nominální data je tzv. **nominální variance**, která se vypočítává podle vzorce

$$nominální\ variance = \frac{n^2 - \sum n_k^2}{n^2} = \frac{3815^2 - (2121^2 + 1569^2 + 125^2)}{3815^2} = 0,52 \tag{26}$$

kde n je celková četnost všech odpovědí a n_k jsou četnosti odpovědí v jednotlivých kategoriích.

Nominální variance je ve srovnání s variačním poměrem přesnější mírou variability. Její nevýhodou však je, podobně jako u variačního poměru, že závisí na počtu kategorií odpovědí.

Chceme-li srovnávat variabilitu u položek s různým počtem kategorií, je vhodné použít tzv. **normované nominální variance**. Normovaná nominální variance je dána poměrem mezi nominální variancí dosaženou a nominální variancí maximálně dosažitelnou a lze ji vypočítat ze vztahu

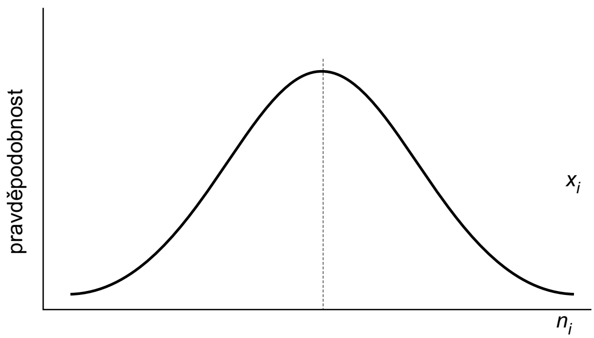
$$\begin{aligned} normovaná\ nominální\ variance &= nominální\ variance \cdot \frac{k}{k-1} \\ normovaná\ nominální\ variance &= 0,52 \cdot \frac{3}{3-1} = 0,78 \end{aligned} \tag{27}$$

kde k je počet kategorií v dotazníkové položce.

Normovaná nominální variance může nabývat hodnot v intervalu od 0 do +1, přičemž tyto hodnoty nezávisí na počtu kategorií odpovědí.

2.4 NORMÁLNÍ ROZDĚLENÍ

Ve výzkumech se často setkáváme se situací, kdy proměnnou, kterou měříme, ovlivňuje současně větší počet poměrně slabých náhodných vlivů. Toto současné působení mnoha vlivů se velmi často projevuje tím, že značná část výsledků se soustřeďuje kolem průměrné hodnoty a na obě strany od ní jsou výsledky stále méně časté, přičemž extrémní hodnoty se vyskytují jen ojediněle. Tuto empirickou zákonitost lze vyjádřit graficky pomocí křivky zvonovitého tvaru, zvané **Gaussova křivka**.



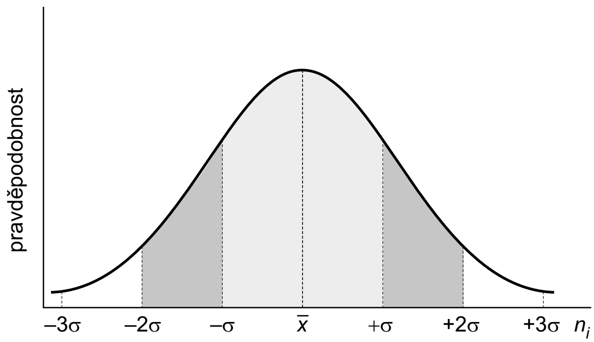
Obr. *7 Gaussova křivka*

Gaussova křivka je souměrná podle osy procházející jejím vrcholem. Vrchol křivky odpovídá aritmetickému průměru všech naměřených hodnot. Křivka nikde neprotíná vodorovnou osu, nýbrž se k ní stále blíží (tato veličina může teoreticky nabývat hodnot od $-\infty$ do $+\infty$). Tvar křivky je závislý na velikosti standardní odchylky. Při zvětšování standardní odchylky se křivka stává plošší a protáhlejší podél vodorovné osy, při zmenšování standardní odchylky se křivka protahuje nahoru a nabývá jehlovitého tvaru.

K získání základní představy o normálním rozdělení je dobré si zapamatovat následující údaje:

- v intervalu od $-\sigma$ do $+\sigma$ kolem aritmetického průměru se nachází přibližně 2/3 (68,27 %) všech hodnot;
- v intervalu od -2σ do $+2\sigma$ leží přibližně 19/20 (95,4 %) všech hodnot;
- v intervalu od -3σ do $+3\sigma$ leží již prakticky všechny hodnoty (99,73 %). V této souvislosti se někdy hovoří o **pravidlu šesti sigma**.

Uvedené skutečnosti jsou dobře patrné i z obrázku 8, kde jsou pravděpodobnosti výskytu hodnot ve zmiňovaných intervalech vyznačeny různě šedou barvou.



Obr. 8 *Pravděpodobnost výskytu hodnot u normálního rozdělení*

Pravděpodobnost výskytu určité hodnoty *x* lze u normálního rozdělení také vypočítat. Byl odvozen vztah, pomocí něhož lze vypočítat tzv. **hustotu pravděpodobnosti** normálního rozdělení. Hustota pravděpodobnosti normálního rozdělení *f (x)* je pravděpodobnost, že náhodná veličina nabude právě hodnoty *x*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\bar{x})^2}{2\sigma^2}} \tag{28}$$

kde σ je směrodatná odchylka základního souboru, *π* je Ludolfovo číslo (tj. číslo 3,14159), *e* je základ přirozených logaritmů (tj. číslo 2,71282), $\bar{x}$ je aritmetický průměr základního souboru a *x* je hodnota, jejíž pravděpodobnost výskytu se vypočítává.

Z uvedeného vztahu je patrné, že pravděpodobnost výskytu určité hodnoty *x* (pravděpodobnost určitého výsledku) závisí, kromě konstant *π* a *e,* pouze na parametrech $\bar{x}$ a σ.

Při řešení praktických problémů s normálním rozdělením používáme často tzv. **distribuční funkci normálního rozdělení** *F(x).* Distribuční funkce normálního rozdělení představuje pravděpodobnost, že normální veličina nabude hodnoty *x* anebo hodnot menších. Distribuční funkci je možno vyjádřit pomocí vztahu

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \int_{-\infty}^{x} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}} dx \tag{29}$$

Protože výpočty hustoty pravděpodobnosti i distribuční funkce normálního rozdělení jsou poměrně velmi pracné, využívají se často místo výpočtů statistické tabulky. Aby nebylo nutné sestavovat tabulky distribuční funkce normálního rozdělení zvlášť pro každou hodnotu aritmetického průměru a pro každou hodnotu směrodatné odchylky, zavádí se tzv. **normovaná normální veličina** *u*, jejíž hodnoty se vypočítávají ze vztahu

$$u = \frac{x - \bar{x}}{s} \tag{30}$$

kde *x* je určitá hodnota normální veličiny, $\bar{x}$ je aritmetický průměr a *s* je směrodatná odchylka. Veličina *u* má tedy průměr 0 a směrodatnou odchylku 1. Veličina *u* vlastně informuje o tom, jak daleko od aritmetického průměru daná hodnota je, přičemž tato vzdálenost se udává ve směrodatných odchylkách. Jestliže například pro určitou hodnotu x vypočítáme, že $u = -2$, znamená to, že tato hodnota je o dvě směrodatné odchylky menší než aritmetický průměr.

V některých statistických aplikacích bývá hodnota normované normální veličiny označována symbolem z (z-skóre).

Ve statistických tabulkách se nejčastěji uvádějí hodnoty **distribuční funkce normovaného normálního rozdělení** $\Phi\,(u)$ (příloha I).

Příklad 14

Ve výzkumu se měřil čas, který žáci potřebují k vyřešení úloh jistého didaktického testu. Zjistilo se, že průměrný čas žáků byl $\bar{x} = 20$ minut a směrodatná odchylka $s = 5$ minut. Máme určit, kolik procent žáků dokáže didaktický test vyřešit za 14 minut, anebo za dobu ještě kratší. (Předpokládáme, že výsledky měření vykazují normální rozdělení a že odhady průměru a směrodatné odchylky jsou dostatečně spolehlivé.)
Nejdříve je třeba hodnotu transformovat pomocí vztahu (30) na veličinu u

$$u = \frac{14 - 20}{5} = -1{,}20$$

Tento výsledek znamená, že hodnota x je o 1,2 směrodatné odchylky menší než aritmetický průměr.
Ve statistických tabulkách (příloha I) zjišťujeme, že $\Phi\,(-1{,}20) = 0{,}1151$.
Pravděpodobnost výskytu času $x = 14$ minut, anebo času ještě kratšího, je tedy asi 0,115, tj. 11,5 %.

Jak již bylo uvedeno, tendenci k normálnímu rozdělení projevují zejména ty náhodné veličiny, u kterých působí více náhodných vlivů současně. Lze to například často pozorovat při měření vědomostí žáků pomocí didaktických testů, při měření rozumových schopností lidí, při měření výkonů žáků v tělesné výchově apod.

U mnoha statistických procedur se předpokládá, že výsledky, s nimiž pracujeme, mají normální rozdělení. Tento předpoklad by se měl vždy alespoň zhruba ověřovat, protože jinak lze snadno dospět ke zkresleným závěrům.

2.5 METODY PRŮZKUMOVÉ ANALÝZY DAT

V poslední době lze v používání statistických kvantitativních metod pozorovat trend, který se snaží umožnit a usnadnit používání kvantitativních metod i výzkumníkům, kteří svojí přípravou a zaměřením k těmto metodám příliš neinklinují. Projevuje se snaha navrhovat a vyvíjet takové postupy a procedury, které by při minimálních nárocích na statisticko-metodologickou přípravu umožňovaly efektivní využívání moderní statistiky. Vzniká tak vlastně nový obor aplikované statistiky, který je znám pod označením **průzkumová analýza dat** (*exploratory data analysis*).

V průzkumové analýze dat se využívají většinou statistické metody vyvinuté pro ordinální (pořadová) data. Typické pro metody průzkumové analýzy dat je to, že se vyznačují velkou **robustností** (to znamená, že nejsou příliš choulostivé na nedodržení podmínek pro jejich oprávněné použití, jak je známe například u tzv. parametrických statistických testů). Metody průzkumové analýzy dat poskytují velmi **názorné grafické výstupy**, které zpřístupňují výsledky výzkumu širokému okruhu zájemců. S metodami průzkumové analýzy se v posledních letech často setkáváme například v mezinárodních srovnávacích výzkumech. Mezi nejznámější metody průzkumové analýzy dat patří tzv. **S-L grafy**, **krabicové grafy**, popřípadě **krabicové grafy s vruby**.

2.5.1 S-L GRAFY

S-L grafy jsou v podstatě určitým druhem histogramů četností. Označení grafů vychází z anglického názvu *stem and leaf display* (graf stonek a list).

Příklad 15

Užití S-L grafu popíšeme na příkladu výzkumu, ve kterém byly měřeny vědomosti žáků didaktickým testem.
Skupina padesáti žáků dosáhla v didaktickém testu ze zeměpisu následující bodové výsledky:

28, 4, 18, 7, 21, 10, 2, 31, 26, 25, 5, 23, 5, 12, 14, 16, 19, 19, 11, 24, 11, 24, 15, 7, 22, 13, 15, 17, 18, 16, 6, 23, 19, 21, 9, 21, 15, 10, 19, 20, 8, 14, 18, 17, 13, 16, 16, 12, 12, 14.

Uvedený soubor dat je zachycen na obrázku 9 pomocí S-L grafu.
Každý dosažený výsledek v testu je v S-L grafu zapsán pomocí dvou kódů: tzv. návěští (číslice vlevo od svislice) a tzv. doplňku (číslice vpravo od svislice). Návěští vyjadřuje u daného výsledku desítky, doplněk vyjadřuje jednotky. Například výsledek 25 bodů má návěští 2 (desítky) a doplněk 5 (jednotky). Každé návěští označuje jeden „stonek", který je tvořen z několika „listů" (doplňků). Na obrázku 9 je celý soubor dat rozdělen do 7 „stonků" (intervalů).

0	24
0	5567789
1	001122233444
1	5556666778889999
2	011123344
2	568
3	1

Obr. 9 *S-L graf pro výsledky didaktického testu*

Protože délka stonku odpovídá četnostem hodnot v jednotlivých intervalech, poskytuje S-L graf informaci o rozložení hodnot v daném souboru. Výhoda S-L grafu ve srovnání s klasickým histogramem spočívá zejména v tom, že zachovává původní čísla, která se dají velmi snadno rekonstruovat. Umožňuje také sledovat, které hodnoty v daném souboru dat chybí (tzv. díry). V uvedeném příkladě S-L grafu například zjišťujeme ve stoncích 2 absenci výsledků 27 a 29 bodů.

Jestliže sestrojujeme S-L graf ručně, postupujeme tak, že nejdříve rozdělíme daný soubor dat do vhodného počtu stonků (intervalů). U souborů s celočíselnými hodnotami a s větší variační šíří je například možné použít pro každou desítku dva stonky (první pro jednotky 0 až 4, druhý pro jednotky 5 až 9). U souborů celočíselných hodnot s menší variační šíří je možno pro každou desítku použít jeden stonek. Do jednotlivých stonků potom zapisujeme příslušné hodnoty, které nakonec vzestupně uspořádáme.

S-L grafy lze použít i k velmi názornému porovnávání dvou souborů dat. V tomto případě je výhodné umístit návěští uprostřed a stonky pro oba soubory kreslit napravo a nalevo od něho. Obrázek 10 uvádí porovnání souboru hodnot z předcházejícího příkladu s výsledky jiné skupiny žáků, která byla testována stejným didaktickým testem.

4431	0	24
8887775	0	5567789
433322110	1	001122233444
8877665555	1	5556666778889999
44332211110000	2	011123344
776555	2	568
21	3	1

Obr. 10 *S-L grafy pro dvě skupiny žáků*

Možnosti analýzy na PC
SPSS: d*escriptive statistics → explore → plots → stem-and-leaf*

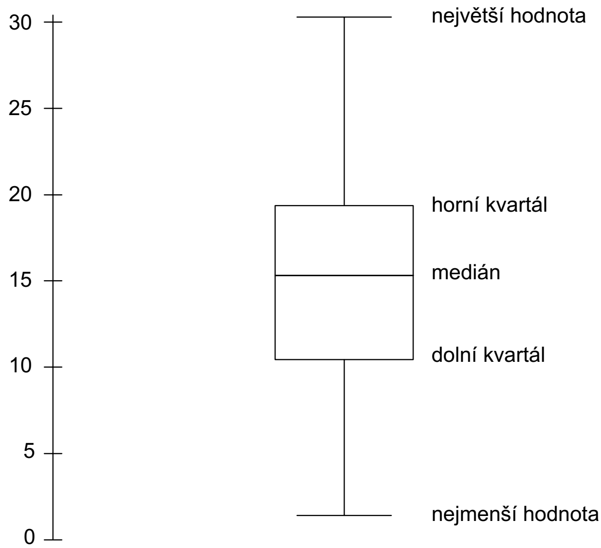
2.5.2 KRABICOVÉ GRAFY

Soubor dat je možné přehledně znázornit také pomocí tzv. **krabicových grafů** (b*ox-and--whiskers-plots* – „krabice s kníry“) (Byčkovský, 1992). Tyto grafy bývají také často označovány jako **kvartilové grafy**.

U krabicových grafů se data znázorňují pomocí zvolené charakteristiky polohy (většinou medián) a odpovídající míry variability (většinou kvartily). Jak již bylo uvedeno v oddíle 2.3, medián je hodnota, která v řadě hodnot (seřazených podle velikosti) odděluje polovinu větších hodnot od poloviny menších hodnot. Dolní kvartil Q_1 je hodnota, která odděluje čtvrtinu nejmenších hodnot, a horní kvartil Q_3 je hodnota, která odděluje čtvrtinu největších hodnot (srov. oddíl 2.3.4.4).

Krabicový graf pro výsledky žáků v didaktickém testu ze zeměpisu (příklad 15) je uveden na obrázku 11. Údaje, které krabicový graf poskytuje, se odčítají na škále umístěné na obrázku vlevo. Z uvedeného kvartilového grafu lze vyčíst, že největší naměřená hodnota

(x_{max}) je 31 bodů a nejmenší naměřená hodnota (x_{min}) je 2 body. Horní hrana „krabice" vyznačuje horní kvartil a spodní hrana určuje dolní kvartil. Medián je vyznačen vodorovnou čarou uvnitř „krabice". Šířka „krabice" většinou odpovídá celkové četnosti v daném souboru (rozsahu výběru). K sestrojení krabicového grafu potřebujeme pět základních údajů: největší a nejmenší naměřenou hodnotu, dolní a horní kvartil a medián.



Obr. 11 *Krabicový graf*

Užití krabicových grafů je výhodné například v případech, kdy potřebujeme porovnat výsledky měření v několika souborech dat.

Příklad 16

Ve výzkumu se ověřovalo, zda existují rozdíly mezi vědomostmi chlapců a dívek ve fyzice na základní škole. Vědomosti chlapců i dívek byly měřeny stejným didaktickým testem a byly získány výsledky, které uvádí tabulka 11. Z hodnot uvedených v tabulce byly vypočítány údaje potřebné ke konstrukci krabicových grafů (obr. 12).

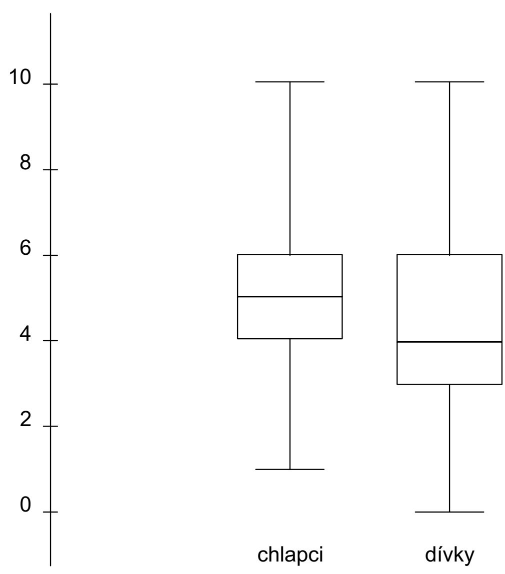
Tab. 11 *Výsledky chlapců a dívek v didaktickém testu z fyziky*

Počet bodů	Četnost chlapců	Četnost dívek
0	0	1
1	3	6
2	8	18
3	18	23

▶

Počet bodů	Četnost chlapců	Četnost dívek
4	28	21
5	23	28
6	20	15
7	9	8
8	9	9
9	8	7
10	7	1
	Σ 133	Σ 137

	Chlapci	Dívky
největší hodnota	10	10
nejmenší hodnota	1	0
horní kvartil	6	6
dolní kvartil	4	3
medián	5	4



***Obr. 12** Krabicové grafy pro výsledky chlapců a dívek*

***Poznámka:** Při vytváření a při interpretaci krabicových grafů je třeba si ověřit, který typ grafu počítačový program vytváří. Některé programy mohou při konstrukci krabicových grafů vycházet z aritmetického průměru a standardní odchylky, a poskytují proto jinou informaci, kterou je také třeba odlišným způsobem interpretovat (horní a dolní hrana „krabice“ v tomto případě nevyjadřují kvartily, nýbrž rozptylování hodnot kolem průměru, které je vyjádřeno pomocí standardní odchylky).*

Možnosti analýzy na PC
Statistica.cz: grafy → 2D grafy → krabicové grafy

2.5.3 KRABICOVÉ GRAFY S VRUBY

Některé statistické programy (např. Statgraphics – nabídka *plotting and descriptive statistics → exploratory data analysis → notched box plots*) nabízí uživateli kromě obvyklých krabicových grafů také tzv. **krabicové grafy s vruby**. Tento druh krabicových grafů je vhodný například tehdy, jestliže srovnáváme dva nebo více souborů dat a zajímá nás, zda mezi nimi jsou **statisticky významné rozdíly**. Tyto grafy poskytují navíc ještě informaci o tzv. **intervalu spolehlivosti**.

Intervalem spolehlivosti rozumíme interval, ve kterém s danou pravděpodobností (např. 95 %) leží všechny naměřené hodnoty. Krabicový graf s vruby pro výsledky chlapců a dívek z předchozího příkladu uvádí obrázek 12. V tomto grafu je interval spolehlivosti 95 % vyznačen pomocí vrubů (zářezů).

Prezentované krabicové grafy s vruby poskytují informaci, kterou je možno vyjádřit následujícím způsobem:

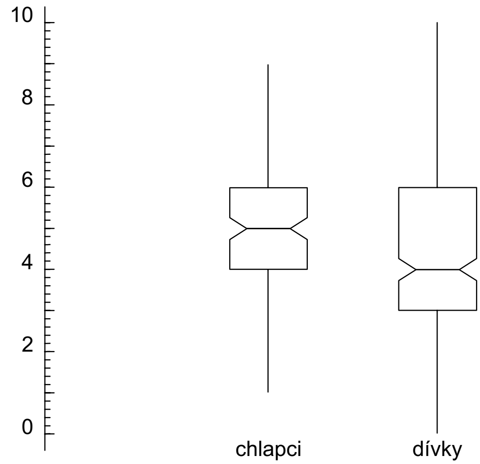
Chlapci dosáhli v didaktickém testu výkonu, který má medián 5 bodů, zatímco dívky dosáhly výkonu s mediánem 4 body. 95% interval spolehlivosti je v tomto případě u chlapců vymezen zhruba rozmezím od 4,7 bodu do 5,3 bodu, zatímco u dívek je vymezen hodnotami od 3,6 bodu do 4,4 bodu (přečteme na škále vlevo). Z uvedeného plyne, že s velkou pravděpodobností (95 %) se výkony všech chlapců budou pohybovat od 4,7 bodu do 5,3 bodu a výkony všech dívek od 3,6 bodu do 4,4 bodu.

Vzhledem k tomu, že v uvedeném příkladě se vruby obou krabicových grafů navzájem nepřekrývají, můžeme usuzovat, že je velmi nepravděpodobné, že by chlapci a dívky dosáhli stejného výsledku (tato pravděpodobnost je menší než 5 %). Rozdíl mezi výkony obou skupin můžeme proto charakterizovat jako **statisticky významný** (na hladině významnosti 0,05 – srov. oddíl 3.3.1). Pokud by se vruby obou krabicových diagramů překrývaly, znamenalo by to, že mezi oběma skupinami žáků nejsou statisticky významné rozdíly.

Krabicové grafy s vruby se většinou kreslí pomocí statistických počítačových programů. Pokud bychom chtěli krabicový graf s vruby sestrojit ručně, potom se velikost vrubu kolem mediánu určuje (pro hladinu významnosti 0,05) ze vztahu

$$velikost\ vrubu = C \cdot 1{,}852 \cdot \frac{Q_3 - Q_1}{\sqrt{n}} \tag{31}$$

kde C je konstanta (v rozmezí od 1,4 do 1,96), $Q_3 - Q_1$ je tzv. kvartilové rozpětí, které je reprezentováno zvolenou výškou „krabice“ a n je rozsah souboru dat. Hodnota konstanty $C = 1{,}96$ se doporučuje v případě, že variabilita výsledků v porovnávaných souborech je výrazně rozdílná, hodnota $C = 1{,}40$ zase tehdy, jestliže se variabilita obou souborů příliš neliší. Počítačové systémy většinou pracují s hodnotou $C = 1{,}7$ (Byčkovský, 1992).



***Obr. 13** Krabicové grafy s vruby pro výsledky chlapců a dívek v didaktickém testu*

> ***Poznámka:** Krabicové diagramy s vruby poskytují většinou také informaci o rozsahu zobrazovaných výběrů (šířky „krabic“ se kreslí v poměru odmocnin rozsahů srovnávaných výběrů).*

Krabicové grafy s vruby jsou zvlášť vhodné pro porovnávání menších a středně velkých souborů (např. od 25 do 400 hodnot). Pro porovnávání velkých souborů (např. pro větší počet hodnot než 1 000) již tyto grafy nevyhovují, protože v těchto případech vychází příliš malé velikosti vrubů.

Metody průzkumové analýzy dat mohou nalézt uplatnění v pedagogickém výzkumu, ale i v pedagogické praxi, například při zpracování výsledků diagnostických šetření. Jejich výhodou je to, že nevyžadují náročné výpočty a nepředpokládají u uživatelů rozsáhlejší znalosti statistiky.

3. STATISTICKÉ METODY POUŽÍVANÉ PŘI TESTOVÁNÍ HYPOTÉZ

3.1 VĚCNÉ A STATISTICKÉ HYPOTÉZY VE VÝZKUMU

V kvantitativně orientovaných výzkumech ověřujeme hypotézy o vztazích mezi jevy (mezi proměnnými). Tyto hypotézy jsou obvykle nejdříve formulovány jako tzv. **věcné hypotézy**, v nichž se k vyjádření jednotlivých proměnných používá věcných termínů. Věcnými hypotézami jsou například tvrzení:

Agresivita dětí předškolního věku se vyskytuje častěji u dětí vyrůstajících v neúplných rodinách než u dětí, které vyrůstají v rodinách úplných.
Chlapci na základní škole dosahují lepších výsledků ve fyzice než dívky.

Proměnné, které ve věcné hypotéze vystupují, se zpravidla následně **operacionalizují**, tj. vyjadřují tak, aby je bylo možno přesně zachytit (změřit). V uvedených příkladech můžeme například vědomosti žáků z fyziky vyjádřit (operacionalizovat) jako výsledky určitého didaktického testu, agresivitu dětí můžeme operacionalizovat jako výskyt určitých projevů v chování dětí za určité období apod.

Aby bylo možné věcné hypotézy ověřovat (testovat) pomocí statistických metod, převádějí se na tzv**. statistické hypotézy**. Statistické hypotézy jsou hypotetická tvrzení o vztazích mezi jevy vyjádřená ve statistických termínech. Například shora uvedeným věcným hypotézám odpovídají následující statistické hypotézy:

Četnost projevů agresivity je u dětí vyrůstajících v neúplných rodinách vyšší než u dětí, které vyrůstají v rodinách úplných.
Průměrný počet bodů v didaktickém testu z fyziky je u chlapců vyšší než průměrný počet bodů u dívek.

Statistickou hypotézu neověřujeme přímo (samu o sobě), nýbrž vždy proti nějakému jinému tvrzení, obyčejně proti tzv. **nulové hypotéze**. Nulová hypotéza je domněnka, která prostřednictvím statistických termínů tvrdí, že mezi proměnnými, které zkoumáme, není vztah. Jestliže například ověřujeme statistickou hypotézu, že průměrný počet bodů v testu u chlapců $\overline{x}_{ch}$ je větší než průměrný počet bodů v testu u dívek $\overline{x}_{d}$ (tj. $\overline{x}_{ch} > \overline{x}_{d}$), potom ji testujeme proti nulové hypotéze, která tvrdí, že $\overline{x}_{ch} = \overline{x}_{d}$.

Pokud se při statistické analýze ukáže, že nulovou hypotézu je možno odmítnout, přijímáme tzv. **alternativní hypotézu**.

Při ověřování výzkumných hypotéz obvykle řešíme dva spolu úzce spjaté problémy. Prvním problémem je otázka, zda vůbec dané proměnné (jevy) spolu souvisí (většinou chceme také vědět, jaké riziko omylu při rozhodování o přijetí hypotézy podstupujeme).

Poté, co existenci vztahu mezi proměnnými prokážeme, zpravidla také usilujeme o postižení těsnosti tohoto vztahu (míry závislosti mezi jevy). O existenci vztahu mezi proměnnými rozhodujeme většinou pomocí tzv. statistických testů významnosti. Těsnost vztahu mezi proměnnými se většinou posuzuje pomocí různých koeficientů (např. koeficienty korelace, regrese, kontingence atd.).

3.1.1 STATISTICKÉ TESTY VÝZNAMNOSTI JAKO PROSTŘEDEK PRO VERIFIKACI HYPOTÉZ

Statistické testy významnosti jsou postupy (procedury), pomocí nichž ověřujeme, zda mezi proměnnými existuje vztah (závislost, souvislost, rozdíl). Výsledkem testů významnosti je rozhodnutí, zda mezi jevy existuje tzv. **statisticky významný vztah**. Jestliže konstatujeme, že určitý výsledek výzkumu je statisticky významný (signifikantní), znamená to, že je velmi nepravděpodobné, že by byl způsoben pouhou náhodou.

Rozhodování ve statistických testech významnosti má vždy pravděpodobnostní charakter (nikdy si nejsme svým rozhodnutím beze zbytku jisti). Pravděpodobnost (riziko), že neoprávněně odmítneme nulovou hypotézu (a tudíž nesprávně přijmeme alternativní hypotézu) se nazývá **signifikance (významnost)**. Při realizaci testů významnosti můžeme rozhodnout, jak velké riziko chyby je v dané výzkumné situaci ještě přijatelné. V této souvislosti hovoříme o zvolené **hladině významnosti**.

Badatel se může při realizaci testů významnosti dopustit dvou druhů chyb. **Chyba prvního druhu** (je zpravidla označována α) spočívá v tom, že neoprávněně odmítneme nulovou hypotézu, ač je správná. Velikost této chyby je většinou dána zvolenou hladinou významnosti. **Chyba druhého druhu** (většinou se označuje β) spočívá v tom, že neoprávněně (nesprávně) přijmeme nulovou hypotézu, ač není správná. Velikost obou chyb spolu souvisí. Jestliže např. při daném rozsahu výběru snížíme chybu prvního druhu (např. místo hladiny významnosti 0,05 použijeme hladinu významnosti 0,01), zvýšíme tím současně chybu druhého druhu. Zmenšování obou chyb současně lze dosáhnout jedině zvětšováním rozsahu výběru.

3.1.2 DRUHY STATISTICKÝCH TESTŮ VÝZNAMNOSTI

Statistické testy významnosti, u nichž se nulová hypotéza týká některého parametru rozdělení náhodné veličiny (např. aritmetického průměru, směrodatné odchylky apod.), se označují jako **testy parametrické**. Parametrické testy vyžadují splnění řady předběžných podmínek a požadavků, má-li být jejich použití oprávněné. Vyžaduje se např., aby rozdělení náhodné veličiny bylo určitého typu (nejčastěji se požaduje, aby rozdělení bylo normální). Splnění těchto podmínek může v praxi znamenat určité komplikace.

Statistické testy významnosti, u nichž se nulová hypotéza netýká parametrů rozdělení (nýbrž nějaké jiné, obecné vlastnosti rozdělení), označujeme jako **testy neparametrické**. Použití neparametrických testů není vázáno na splnění tak velkého počtu požadavků a na tak přísné předběžné podmínky jako u testů parametrických. Použití neparametrických testů významnosti je možné i v případech, kdy není znám typ rozdělení náhodné veličiny.

Větší univerzálnost neparametrických testů významnosti je však vykoupena jejich menší statistickou účinností. Účinností statistického testu významnosti se rozumí schopnost testu rozpoznat i malé odchylky od nulové hypotézy (a odmítnout tedy nulovou hypotézu jako nesprávnou). Parametrické testy významnosti jsou naopak účinnější (ale ne tak univerzálně použitelné) jako testy neparametrické. Neparametrické testy významnosti vyžadují ve srovnání s testy parametrickými větší počet případů (pozorování).

Podle toho, jakým způsobem formulujeme v testu významnosti alternativní hypotézu, lze rozlišit statistické **testy jednostranné** a statistické **testy oboustranné**. Jestliže například ve statistickém testu významnosti byla formulována nulová hypotéza typu $a = b$ (např. aritmetický průměr skupiny A = aritmetický průměr skupiny B), lze alternativní hypotézu formulovat dvojím způsobem: buď $a \neq b$, nebo $a < b$ (respektive $a > b$). Pokud statistický test významnosti pracuje s alternativní hypotézou typu $a \neq b$, jedná se o tzv. **oboustranný test**. Pokud je alternativní hypotéza formulována ve formě $a > b$ (respektive $a < b$), hovoříme o **testech jednostranných**.

Testy jednostranné by se měly používat pouze v případech, kdy je prakticky jisté, že může platit pouze jedna z alternativ (buď $a > b$, nebo $a < b$). V případě, že mohou platit obě alternativy, používáme testy oboustranné. Příklady všech testů významnosti uváděné v této práci jsou testy oboustrannými.

Bližší a podrobnější vysvětlení k těmto dvěma typům statistických testů lze nalézt například v práci Komendy a Klementy (1981).

3.2 STATISTICKÉ METODY PRO ANALÝZU NOMINÁLNÍCH DAT

3.2.1 TEST DOBRÉ SHODY CHÍ-KVADRÁT

U této kategorie testů významnosti se ověřuje, zda četnosti, které byly získány měřením v pedagogické realitě, se odlišují od teoretických četností, které odpovídají dané nulové hypotéze. Princip testu dobré shody chí-kvadrát (ale zároveň i základní principy ostatních testů významnosti) budeme ilustrovat na příkladech fiktivních výzkumných situací (příklady 17–19).

Příklad 17

Skupina 120 žáků základní školy odpovídala v dotazníku na otázku:

Co nejraději děláš ve svém volném čase?

A sport
B četba knih
C poslech hudby
D hry na počítači

Na základě testu dobré shody chí-kvadrát rozhodneme, zda mezi preferencemi volnočasových aktivit žáků jsou statisticky významné rozdíly.

Odpovědi žáků uvádí tabulka 12 ve sloupci „Pozorovaná četnost *P*". Vidíme, že žáci nejčastěji volili odpověď D „hry na počítači" (40 žáků), dále potom (v pořadí podle preference) následuje A „sport" (36 žáků), B „četba knih" (23 žáků) a C „poslech hudby" (21 žáků). Smyslem testu dobré shody chí-kvadrát je rozhodnout, zda zjištěné (pozorované) rozdíly mezi preferencemi jednotlivých četností jsou statisticky významné.

Tab. 12 *Princip testu dobré shody chí-kvadrát*

Preferovaná činnost	**Pozorovaná četnost *P***	**Očekávaná četnost *O***	$P - O$	$(P - O)^2$	$\frac{(P - O)^2}{O}$
sport	36	30	6	36	1,200
četba knih	23	30	−7	49	1,633
poslech hudby	21	30	−9	81	2,700
hry na počítači	40	30	10	100	3,333
	∑ 120	∑ 120			∑ 8,866

Test dobré shody chí-kvadrát (ale i ostatní testy významnosti) začíná formulováním **nulové** a **alternativní hypotézy**.

Obecně platí, že nulová hypotéza (označovaná H_0) je předpoklad, že mezi sledovanými jevy není vztah (souvislost, rozdíl).

Alternativní hypotéza (označovaná H_1 nebo H_A) je naopak předpokladem, že mezi sledovanými jevy vztah (souvislost, rozdíl) existuje.

V uvedeném příkladě budou mít obě hypotézy následující podobu:

H_0: *Četnosti žáků, kteří preferují uvedené volnočasové aktivity, jsou stejně velké.*
H_A: *Četnosti žáků, kteří preferují dané volnočasové aktivity, jsou rozdílné.*

O přijetí nebo odmítnutí uvedených hypotéz rozhodneme na základě testování nulové hypotézy. K tomuto účelu se u testů významnosti vypočítává tzv. **testové kritérium**, což je určitá číselná charakteristika odvozená ze zjištěných dat. U testu dobré shody chí-kvadrát je testovým kritériem hodnota

$$\chi^2 = \sum \frac{(P-O)^2}{O} \tag{32}$$

kde χ^2 je testové kritérium chí-kvadrát, *P* je tzv. **pozorovaná četnost** (skutečná, empiricky zjištěná četnost) a *O* očekávaná četnost. **Očekávaná četnost** je četnost, která odpovídá formulované nulové hypotéze. Pokud by žáci uvedené volnočasové aktivity preferovali zcela stejně, „očekávali" bychom, že jich čtvrtina (tj. 30) zvolí sport, četbu knih, poslech hudby i hry na počítači.

Z tabulky 12 je patrné, že vypočítaná hodnota $\chi^2 = 8{,}866$. Tato hodnota je ukazatelem rozdílu mezi pozorovanou a očekávanou četností. Při rozhodování o platnosti nulové hypotézy zpravidla postupujeme tak, že vypočítanou hodnotu testového kritéria srovnáváme s tzv. **kritickou hodnotou**, kterou lze nalézt ve statistických tabulkách (příloha II). Příslušnou kritickou hodnotu hledáme vždy pro určitou (zvolenou) hladinu významnosti a tzv. **počet stupňů volnosti**.

Hladina významnosti je (jak již bylo uvedeno) pravděpodobnost, že neoprávněně (nesprávně) odmítneme nulovou hypotézu. Tuto pravděpodobnost lze volit podle situace (její závažnosti), ve většině pedagogických výzkumů se však pracuje na hladině významnosti 0,05 (5 %) nebo 0,01 (1 %).

Počet stupňů volnosti závisí u testu dobré shody chí-kvadrát na počtu řádků v tabulce, z níž bylo kritérium chí-kvadrát vypočítáno. Počet stupňů volnosti je počet řádků tabulky, kterým by bylo možno teoreticky přiřknout libovolnou hodnotu a přitom dodržet stanovený sloupcový součet. V našem případě lze součet 120 vytvořit tak, že tři hodnoty zvolíme zcela libovolně (volně), tabulka má proto tři stupně volnosti (tab. 13).

Tab. 13 *Určení počtu stupňů volnosti u testu dobré shody chí-kvadrát*

Preferovaná aktivita	Pozorovaná četnost *P*
sport	libovolná hodnota
četba knih	libovolná hodnota
poslech hudby	libovolná hodnota
hry na počítači	?
	Σ 120

Ve statistických tabulkách (příloha II) zjišťujeme, že pro hladinu významnosti 0,05 a počet stupňů volnosti *f* = 3 je kritická hodnota testového kritéria $\chi^2_{0{,}05}(3) = 7{,}815$.

Vypočítaná hodnota testového kritéria $\chi^2 = 8{,}866$ je větší než hodnota kritická $\chi^2_{0{,}05}(3) = 7{,}815$, proto odmítáme nulovou hypotézu a přijímáme hypotézu alternativní. Mezi preferencemi daných volnočasových aktivit byly prokázány statisticky významné rozdíly (zjištěné rozdíly není možno připsat na vrub náhody). Pravděpodobnost chybného rozhodnutí je v tomto případě menší než 0,05 (5 %).

Příklad 18

V určitém roce zemřelo v České republice na následky dopravních nehod 114 dětí. Tato úmrtí byla rozdělena do jednotlivých měsíců roku tak, jak uvádí tabulka 14 ve sloupci „pozorovaná četnost".

Rozhodneme, zda počet úmrtí je ve všech měsících roku stejný, anebo zda v některých měsících je počet úmrtí vyšší.

Nejdříve formulujeme nulovou hypotézu a alternativní hypotézu:

H_0: Četnosti úmrtí dětí jsou v jednotlivých měsících roku stejné.
H_A: Četnosti úmrtí se v jednotlivých měsících roku liší.

Nulovou hypotézu budeme testovat na hladině významnosti α = 0,05.

Tab. 14 *Smrtelné úrazy dětí v jednotlivých měsících roku*

Měsíc	**Pozorovaná četnost *P***	**Očekávaná četnost *O***	***P – O***	***(P – O)²***	$\frac{(P-O)^2}{O}$
leden	7	9,5	–2,5	6,25	0,658
únor	7	9,5	–2,5	6,25	0,658
březen	8	9,5	–1,5	2,25	0,237
duben	10	9,5	0,5	0,25	0,026
květen	10	9,5	0,5	0,25	0,026
červen	9	9,5	–0,5	0,25	0,026
červenec	14	9,5	4,5	20,25	2,132
srpen	12	9,5	2,5	6,25	0,658
září	10	9,5	0,5	0,25	0,026
říjen	12	9,5	2,5	6,25	0,658
listopad	7	9,5	–2,5	6,25	0,658
prosinec	8	9,5	–1,5	2,25	0,237
	Σ 114	Σ 114,0			Σ 6,000

Pokud by platila nulová hypotéza, potom bychom očekávali, že četnosti úmrtí v jednotlivých měsících roku budou stejně velké a budou činit jednu dvanáctinu z celkového počtu úmrtí, tj. 114 / 12 = 9,5.

Tyto četnosti označíme jako očekávané četnosti *O*. (Skutečnost, že očekávané četnosti nejsou celá čísla, nás nemusí znepokojovat, protože se jedná o *teoretické* četnosti.)

Abychom mohli rozhodnout o platnosti nulové hypotézy, vypočítáme podle vztahu (32) testové kritérium $\chi^2 = 6{,}000$.

Vypočítanou hodnotu srovnáme s kritickou hodnotou tohoto testového kritéria pro zvolenou hladinu významnosti a příslušný počet stupňů volnosti.

Jak již bylo uvedeno, počet stupňů volnosti je dán počtem těch četností v tabulce, které by bylo možné zvolit libovolně při dodržení daného sloupcového součtu (tj. v uvedeném příkladě 114). V naší tabulce četností je celkem 12 řádků, a proto je zde 12 – 1 = 11 stupňů volnosti. Pro hladinu významnosti 0,05 a pro 11 stupňů volnosti lze ve statistických tabulkách (příloha II) nalézt kritickou hodnotu $\chi^2_{0,05}(11) = 19{,}675$.

Vypočítaná hodnota testového kritéria je menší než hodnota kritická, a proto nelze odmítnout nulovou hypotézu. Nebylo prokázáno, že by výskyt úrazů byl v jednotlivých měsících roku rozdílný.

Příklad 19

Jestliže si pozorně prohlédneme četnosti úmrtí během roku v předcházejícím příkladě, můžeme postihnout, že větší četnosti se objevují především v letních měsících, zatímco v zimních měsících jsou četnosti menší. Může nás zajímat odpověď na otázku, zda výskyt úmrtí v období od dubna do října je skutečně významně vyšší než v období od listopadu do března.

Abychom našli odpověď na položenou otázku, musíme nejdříve sdružit četnosti úmrtí do dvou skupin podle uvažovaných období. V období od dubna do října došlo celkem k 77 úmrtím, za období od listopadu do prosince a od ledna do března došlo k 37 úmrtím (tab. 15).

Můžeme vyslovit následující statistické hypotézy:

H_0: Četnosti úmrtí dětí v období od dubna do října jsou stejně velké jako četnosti úmrtí ve zbývajících měsících roku.
H_A: Četnosti úmrtí jsou v období od dubna do října větší než ve zbývajících měsících roku.

Test významnosti provedeme opět na hladině významnosti α = 0,05.

Při výpočtu očekávaných četností vycházíme z úvahy, že za období od dubna do října (tj. za 7 měsíců) bychom (v případě, že platí nulová hypotéza) očekávali 7/12 z celkového počtu úmrtí a ve zbývajících měsících (tj. za 5 měsíců) 5/12 z celkového počtu úmrtí. Na základě této úvahy jsou vypočítány očekávané četnosti,

tj. $\frac{7}{12}\cdot 114 = 66{,}5$ a $\frac{5}{12}\cdot 114 = 47{,}5.$

Tab. 15 *Test dobré shody chí-kvadrát (různě velké očekávané četnosti)*

Období	Pozorovaná četnost *P*	Očekávaná četnost *O*	*P* – *O*	$(P-O)^2$	$\frac{(P-O)^2}{O}$
duben–říjen	77	66,5	10,5	110,25	1,658
listopad–březen	37	47,5	–10,5	110,25	2,321
	Σ 114	Σ 114,0			Σ 3,979

H_0: Smrtelné úrazy dětí se vyskytují stejně často v obou sledovaných obdobích.
H_A: Výskyt smrtelných úrazů dětí je ve sledovaných obdobích roku rozdílný.

Vypočítaná hodnota testového kritéria chí-kvadrát je $\chi^2 = 3{,}979$.

Tabulka četností má v tomto případě 2 – 1 = 1 stupeň volnosti, kritická hodnota testového kritéria pro hladinu významnosti 0,05 je $\chi^2_{0,05}(1) = 3{,}841$. Je zřejmé, že vypočítaná hodnota testového kritéria je větší než hodnota kritická, a proto musíme odmítnout nulovou hypotézu a přijmout hypotézu alternativní. Za období od dubna do října tedy skutečně dochází častěji k úmrtím dětí v důsledku dopravních nehod než ve zbývajících měsících roku.

Na uvedeném příkladu jsme ukázali, že vhodným seskupením četností v tabulce můžeme dosáhnout různých výsledků statistického testu významnosti. Takovéto sdružování četností je oprávněné pouze v případě, že vychází z logiky řešeného problému a nemělo by být pouze spekulací podnícenou daty.

Podmínkou oprávněného použití testu dobré shody chí-kvadrát je, že očekávané četnosti musí být dostatečně velké (většinou se požaduje, aby byly alespoň 5).

Možnosti analýzy na PC
Excel: vložit funkci → statistické → CHITEST (očekávané četnosti je nutno vypočítat)
Statistica.cz: neparametrická statistika → pozorované versus očekávané
SPSS: *nonparametric tests* → *chi-square*

3.2.2 TEST NEZÁVISLOSTI CHÍ-KVADRÁT PRO KONTINGENČNÍ TABULKU

Tento test významnosti je možné použít například v situacích, kdy rozhodujeme, zda existuje souvislost (závislost) mezi dvěma pedagogickými jevy (proměnnými), které byly zachyceny pomocí nominálního (popřípadě ordinálního) měření. Tato situace je častá například při zpracovávání výsledků dotazníkových šetření.

Příklad 20

Vzorku 400 náhodně vybraných studentů pedagogické fakulty byl předložen dotazník, který obsahoval tzv. uzavřené otázky (otázky s nabídnutými odpověďmi). Jedna z otázek zjišťovala, zda studenti byli v uplynulém studijním roce ubytováni na kolejích:

Byl(a) jste v loňském studijním roce ubytován(a) na kolejích?
A byl(a)
B nebyl(a)

Další z otázek zjišťovala, jakého průměrného prospěchu studenti v uplynulém studijním roce dosáhli:

Jaký byl váš průměrný prospěch v loňském studijním roce?
A lepší než 1,6
B 1,6–2,1
C horší než 2,1

Na základě testu nezávislosti chí-kvadrát rozhodneme, zda je vztah (souvislost) mezi tím, zda studenti bydlí na kolejích, a tím, jakých studijních výsledků dosahují.

Výsledky získané dotazníkovým šetřením je nejdříve nutno zapsat do tzv. **kontingenční tabulky** (tab. 16). (Kontingenční tabulka bývá také někdy označována jako „tabulka se dvěma vstupy“.) Čísla v kontingenční tabulce (bez závorek) vyjadřují četnosti studentů, kteří odpověděli určitým způsobem na první otázku a současně určitým způsobem na druhou otázku. Například číslo 93 v tabulce znamená, že 93 studentů odpovědělo, že „bydleli na kolejích“ a současně že měli průměrný prospěch „horší než 2,1“.

Tab. 16 *Kontingenční tabulka*

		průměrný prospěch			
		lepší než 1,6	1,6-2,1	horší než 2,1	Σ
bydlení na kolejích	ano	39 (47,8)	107 (107,55)	93 (83,65)	239
	ne	41 (32,2)	73 (72,45)	47 (56,35)	161
	Σ	80	180	140	400

Čísla uváděná vpravo od tabulky a pod tabulkou jsou tzv. **marginální („okrajové“) četnosti**, tj. součty četností v řádcích a sloupcích tabulky.

Test nezávislosti chí-kvadrát začíná opět formulováním nulové a alternativní hypotézy:

H_0: Mezi četnostmi odpovědí na obě uvedené otázky není závislost (souvislost).
H_A: Mezi odpověďmi respondentů na uvedené otázky je závislost (souvislost).

Testování významnosti provedeme na hladině významnosti 0,05.

Dalším krokem je výpočet očekávaných četností *O* pro každé pole kontingenční tabulky. Jak již bylo uvedeno, očekávané četnosti jsou „teoretické“ četnosti, které odpovídají platnosti nulové hypotézy. Očekávané četnosti jsou v kontingenční tabulce uvedeny v závorkách. Lze odvodit, že očekávanou četnost pro příslušné pole kontingenční tabulky lze vypočítat tak, že násobíme vždy odpovídající marginální četnosti v tabulce a tento součin potom dělíme celkovou četností. Například očekávanou četnost „32,2“ v uvedené kontingenční tabulce vypočítáme

$$O = \frac{80 \cdot 161}{400} = 32{,}2$$

Testové kritérium χ^2 je možné vypočítat podle vztahu (32) z hodnot *P* a *O* pro jednotlivá pole kontingenční tabulky (tab. 17).

Vypočítaná hodnota χ^2 v tabulce 17 je ukazatelem velikosti rozdílu mezi skutečností a vyslovenou nulovou hypotézou. Pro posouzení vypočítané hodnoty χ^2 je dále třeba určit počet stupňů volnosti ve výchozí kontingenční tabulce (ne v tabulce, která sloužila k výpočtu hodnoty χ^2).

***Tab. 17** Výpočet testového kritéria chí-kvadrát pro kontingenční tabulku*

Pozorovaná četnost *P*	**Očekávaná četnost *O***	***P* – *O***	**(*P*–*O*)²**	$\frac{(P-O)^2}{O}$
39	47,80	–8,80	77,44	1,620
107	107,55	–0,55	0,303	0,002
93	83,65	9,35	87,42	1,045
41	32,20	8,80	77,44	2,404
73	72,45	0,55	0,303	0,004
47	56,35	–9,35	87,42	1,551
∑ 400	∑ 400,00			∑ 6,628

Pro kontingenční tabulku o *r* řádcích a *s* sloupcích se určí počet stupňů volnosti podle vztahu

$$f = (r - 1) \cdot (s - 1) \tag{33}$$

kde *r* je počet řádků v kontingenční tabulce a *s* je počet sloupců v kontingenční tabulce. V našem případě vychází pro tabulku o dvou řádcích a třech sloupcích

$$f = (2 - 1) \cdot (3 - 1) = 2$$

Pro vypočítaný počet stupňů volnosti a pro zvolenou hladinu významnosti 0,05 nalezneme ve statistických tabulkách (příloha I) kritickou hodnotu testového kritéria $\chi^2_{0,05}(2) = 5,991$. Srovnáme-li vypočítanou hodnotu testového kritéria s hodnotou kritickou, zjišťujeme, že vypočítaná hodnota je vyšší, a proto můžeme odmítnout nulovou hypotézu. Mezi odpověďmi na uvedené otázky byla prokázána statisticky významná souvislost. Jinak řečeno, bydlení (respektive nebydlení) na kolejích významným způsobem souvisí se studijními výsledky studentů.

Test nezávislosti chí-kvadrát by se neměl provádět v případech, kdy ve více než 20 % polí kontingenční tabulky jsou očekávané četnosti menší než 5 a v případech, kdy v některém poli je očekávaná četnost menší než 1.

***Poznámka:** Počítačové programy analyzují vztahy mezi proměnnými většinou tak, že uživatele informují pouze o výsledku analýzy (uvádí se hodnota testového kritéria, počet stupňů volnosti a hodnota signifikace). Pokud je vypočítaná hodnota signifikace menší než zvolená hladina významnosti (např. 0,05), odmítáme nulovou hypotézu.*

3.2.2.1 Redukce počtu polí v kontingenční tabulce

Ve výzkumech, které pracují s malými výběrovými soubory, bývají požadavky kladené na velikost očekávaných četností v kontingenční tabulce často obtížně splnitelné. Tato situace je častá například ve výzkumech, které jsou realizovány v rámci závěrečných prací vysokoškolského studia (bakalářské práce, diplomové práce). V některých případech lze splnění podmínek pro oprávněné použití testu chí-kvadrát dosáhnout zmenšením počtu řádků nebo sloupců v kontingenční tabulce. Bohužel toto řešení je nutno „zaplatit“ zmenšením citlivosti testu a ztrátou části informace, kterou data obsahují.

Příklad 21

Vysokoškolský student uskutečnil v rámci své diplomové práce výzkum, ve kterém na vzorku 190 náhodně vybraných učitelů základní školy ověřoval hypotézu „Efekt vyhoření se u starších učitelů vyskytuje častěji než u učitelů mladších".

Ve výzkumu byly získány výsledky, které zachycuje následující kontingenční tabulka (tab. 18).

Tab. 18 *Výskyt efektu vyhoření u učitelů základní školy*

věk učitelů (roky)	výskyt efektu vyhoření: ano	výskyt efektu vyhoření: ne	Σ
do 25	6 (4,8)	3 (4,2)	9
26–30	7 (13,3)	18 (11,7)	25
31–35	6 (13,8)	20 (12,2)	26
36–40	6 (5,3)	4 (4,7)	10
41–45	6 (4,3)	2 (3,7)	8
46–50	18 (20,2)	20 (17,8)	38
51–55	18 (10,6)	2 (9,4)	20
56–60	16 (12,8)	8 (11,2)	24
61–65	18 (15,9)	12 (14,1)	30
Σ	101	89	190

Pro všechny pozorované četnosti v kontingenční tabulce (čísla bez závorek) byly vypočítány očekávané četnosti (čísla v závorkách). Zjišťujeme, že pět očekávaných četností (tj. 28 %) je menších než 5 a že nejsou splněny podmínky pro použití testu. (Pole kontingenční tabulky, ve kterých jsou očekávané četnosti menší než 5, jsou označena šedě.)

Danou situaci je možno řešit tak, že v kontingenční tabulce zredukujeme počet polí sloučením věkových kategorií učitelů. Jedno z možných řešení je prezentováno v tabulce 19. Učitele můžeme podle dosaženého věku rozdělit do tří kategorií:

- začínající učitelé (do 25 let);
- učitelé středního věku (26–40);
- učitelé s delší pedagogickou praxí (nad 40 let).

Na základě uvedené kategorizace můžeme vytvořit novou kontingenční tabulku, která má dva řádky a tři sloupce (tab. 19). Úpravou kontingenční tabulky jsme dosáhli toho, že všechny očekávané četnosti mají dostatečně velké hodnoty, a proto můžeme korektním způsobem realizovat test nezávislosti chí-kvadrát (očekávané četnosti jsou v tabulce opět uvedeny v závorkách).

Tab. 19 *Kontingenční tabulka vytvořená pomocí redukce počtu kategorií učitelů*

		věk učitelů (roky)			
		do 25	26–40	nad 40	Σ
výskyt efektu vyhoření	ano	19 (31,9)	30 (29,8)	52 (39,3)	101
	ne	41 (28,1)	26 (26,2)	22 (34,7)	89
	Σ	60	56	74	190

Z pozorovaných a očekávaných četnosti byla vypočítána hodnota testového kritéria $\chi^2 = 19{,}894$. Tabulka má v tomto případě $f = (2 - 1) \cdot (3 - 1) = 2$ stupně volnosti a kritická hodnota pro hladinu významnosti 0,01 činí $\chi^2_{0{,}01}(2) = 9{,}210$.

Vypočítaná hodnota testového kritéria je v uvedeném příkladě výrazně větší než nalezená kritická hodnota, a proto odmítáme nulovou hypotézu a přijímáme hypotézu alternativní. Bylo prokázáno (na hladině významnosti 0,01), že efekt vyhoření se vyskytuje u starších učitelů častěji než u učitelů mladších.

3.2.2.2 Znaménkové schéma pro kontingenční tabulku

Při používání testu nezávislosti chí-kvadrát se zpravidla nespokojujeme s konstatováním, že mezi proměnnými (vlastnostmi, jevy) je nebo není statisticky významný vztah (souvislost). Zajímá nás, kde (ve kterém poli kontingenční tabulky) se vztah projevuje a jak jej můžeme interpretovat. Dobrou pomůckou pro interpretaci výsledků obsažených v kontingenční tabulce je sestavení tzv. **znaménkového schématu kontingenční tabulky**.

Konstrukci a použití znaménkového schématu kontingenční tabulky ukážeme na příkladu výzkumu, ve kterém se ověřoval vztah mezi vzděláním a příjmy obyvatelstva.

Příklad 22

Ve výzkumu se ověřovala hypotéza, že příjmy obyvatelstva určitého regionu závisí na dosaženém stupni vzdělání. Byla získána kontingenční tabulka (tab. 20), která uvedený vztah zachycuje. Kromě pozorovaných četností jsou v závorkách uvedeny také vypočítané očekávané četnosti.

Testem nezávislosti chí-kvadrát bylo prokázáno, že mezi stupněm vzdělání a průměrným měsíčním příjmem obyvatel je statisticky významný vztah (vypočítaná hodnota $\chi^2 = 74{,}27$, kritická hodnota pro hladinu významnosti 0,01 a osm stupňů volnosti je $\chi^2_{0{,}01}(8) = 20{,}090$).

Souvislosti uvnitř uvedené tabulky budeme interpretovat pomocí znaménkového schématu kontingenční tabulky.

Při konstrukci znaménkového schématu pro kontingenční tabulku se postupuje tak, že postupně testujeme významnost rozdílů mezi pozorovanou a očekávanou četností v jednotlivých polích tabulky. K tomuto testování se nejčastěji používá testové kritérium z (z-skóre).

Poznámka: *Veličina z je tzv. normovaná normální veličina, která se také označuje u (srov. oddíl 2.4). Udává, jak daleko je náhodná veličina od aritmetického průměru (vzdálenost se udává ve směrodatných odchylkách).*

Testové kritérium z je možno vypočítat například ze vztahu

$$z = \frac{P_{\%} - O_{\%}}{\sqrt{O_{\%} \cdot (100 - O_{\%})}} \cdot \sqrt{n} \tag{34}$$

kde $P_{\%}$ je pozorovaná četnost v určitém poli vyjádřená v procentech z celkové četnosti, $O_{\%}$ je očekávaná četnost v tomto poli vyjádřená v procentech z celkové četnosti a n je celková četnost v kontingenční tabulce.

Někteří autoři (Řehák a Řeháková, 1986) doporučují počítat hodnoty z-skóre podle vzorce

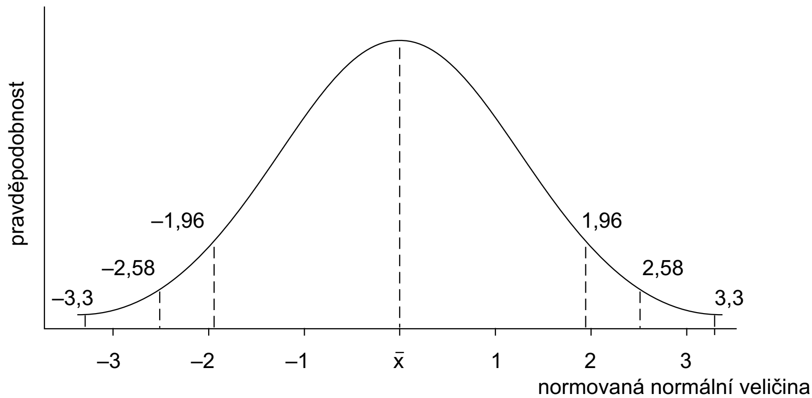
$$z = \frac{n \cdot n_p - n_r \cdot n_s}{\sqrt{n_r \cdot n_s (n - n_r) \cdot (n - n_s)}} \cdot \sqrt{n} \tag{35}$$

kde n je celková četnost v kontingenční tabulce, n_p je pozorovaná četnost v určitém poli kontingenční tabulky, n_r je marginální (okrajová) četnost v řádku tabulky a n_s je marginální četnost ve sloupci kontingenční tabulky.

Tab. 20 *Stupeň vzdělání a průměrné měsíční příjmy respondentů*

průměrný měsíční příjem	**stupeň vzdělání** základní	středoškolské	vysokoškolské	Σ
nižší než 7 000	293 (251)	156 (185)	3 (15)	452
7 000–12 000	876 (845)	609 (625)	34 (50)	1519
13 000–18 000	712 (805)	667 (595)	68 (47)	1447
19 000–24 000	173 (168)	112 (125)	18 (10)	303
nad 24 000	67 (52)	25 (39)	2 (3)	94
Σ	2 121	1 569	125	3 815

Veličinu z je nutno vypočítat pro všechna pole kontingenční tabulky. Pro výpočet hodnot z byl v našem případě použit početně snažší postup podle vztahu (34). Vypočítané hodnoty z se potom simultánně testují, obvykle na hladině významnosti 0,05, 0,01 a 0,001. Kritické hodnoty normované normální veličiny pro uvedené hladiny významnosti jsou patrné z obrázku 14.



Obr. 14 Kritické hodnoty normované normální veličiny (oboustranný test)

Podle výsledků testování přiřazujeme jednotlivým vypočítaným hodnotám *z* znaménka, která se zapisují do znaménkového schématu. Jedno znaménko (buď +, nebo –) se přiřazuje tehdy, pokud je rozdíl mezi pozorovanou a očekávanou četností statisticky významný na hladině významnosti 0,05, tj. pokud platí, že $1{,}96 \leq z < 2{,}58$. Dvě znaménka se přiřazují, pokud rozdíl mezi oběma četnostmi je významný na hladině významnosti 0,01, tj. v případě, že $2{,}58 \leq z < 3{,}30$. Tři znaménka se přiřazují, jestliže rozdíl je významný na hladině významnosti 0,001, tj. tehdy, pokud vypočítaná hodnota odpovídá podmínce $3{,}30 \leq z$. Tabulka 21 uvádí vypočítané hodnoty *z* pro kontingenční tabulku a tabulka 22 uvádí odpovídající znaménkové schéma.

Tab. 21 *Hodnoty z-skóre pro kontingenční tabulku*

Průměrný měsíční příjem	Stupeň vzdělání		
	základní	středoškolské	vysokoškolské
nižší než 7 000	2,74	–2,19	–3,10
7 000–12 000	1,21	–0,61	–2,28
13 000–18 000	–3,73	4,65	3,08
19 000–24 000	0,39	–1,18	2,54
nad 24 000	2,09	–2,24	–0,59

Tab. 22 *Znaménkové schéma kontingenční tabulky*

Průměrný měsíční příjem	Stupeň vzdělání			
	základní	středoškolské	vysokoškolské	Σ
nižší než 7 000	++	–	––	452
7 000–12 000	0	0	-	1 519
13 000–18 000	–––	++	++	1 447
19 000–24 000	0	0	+	303
nad 24 000	+	–	0	94
Σ	2 121	1 569	125	3 815

Význam znamének ve schématu lze shrnout následujícím způsobem:

+++	Pozorovaná četnost je významně větší než četnost očekávaná na hladině významnosti 0,001 (tzn. pozorovaná četnost je značně vyšší než četnost očekávaná).
– – –	Pozorovaná četnost je významně menší než četnost očekávaná na hladině významnosti 0,001 (tzn. pozorovaná četnost je značně menší než četnost očekávaná).
++	Pozorovaná četnost je významně větší než četnost očekávaná na hladině významnosti 0,01.
– –	Pozorovaná četnost je významně menší než četnost očekávaná na hladině významnosti 0,01.
+	Pozorovaná četnost je významně větší než četnost očekávaná na hladině významnosti 0,05.
–	Pozorovaná četnost je významně menší než četnost očekávaná na hladině významnosti 0,05.
0	Mezi pozorovanou četností a četností očekávanou není statisticky významný rozdíl.

Znaménkové schéma umožňuje interpretovat výsledky obsažené v kontingenční tabulce pomocí několika interpretačních výroků. V uvedeném příkladě bychom například mohli vyslovit následující výroky:

- Respondentů s vysokoškolským vzděláním je nejvíce ve vyšších příjmových kategoriích.
- Respondenti se základním vzděláním mají většinou nízké příjmy, ale na druhé straně jich poměrně značná část má příjmy v nejvyšší kategorii.
- Respondenti se středoškolským vzděláním mají nejčastěji příjmy ve střední kategorii apod.

Poměrně pracnou konstrukci znaménkového schématu kontingenční tabulky může usnadnit výpočetní technika (např. program Excel).

3.2.3 TEST NEZÁVISLOSTI CHÍ-KVADRÁT PRO ČTYŘPOLNÍ TABULKU

Zvláštním případem kontingenční tabulky je čtyřpolní tabulka se dvěma řádky a dvěma sloupci. Se čtyřpolní tabulkou se setkáváme v případech, kdy jevy, mezi nimiž ověřujeme vztah, mohou nabývat pouze dvou alternativních kvalit (např. chlapec–dívka, plavec–neplavec, kuřák–nekuřák atd.).

Použití testu nezávislosti chí-kvadrát pro čtyřpolní tabulku budeme ilustrovat na příkladu výzkumu, který se zabýval problémem kouření u mužů a žen.

Příklad 23

Náhodně vybraným vysokoškolským studentům (12 mužů a 36 žen) byla v dotazníku položena otázka, zda kouří; 26 studentů (z toho 15 žen) odpovědělo na otázku kladně.

Na základě dat získaných ve výzkumu máme rozhodnout, zda studenti-muži kouří častěji než studentky-ženy. Výsledky šetření zachycuje tabulka 23.

Tab. 23 *Čtyřpolní tabulka pro test nezávislosti chí-kvadrát*

	kouří	nekouří	Σ
muži	11	1	12
ženy	15	21	36
Σ	26	22	48

Nejdříve formulujeme nulovou a alternativní hypotézu:

H_0: Frekvence kouření je u mužů a žen stejně velká (podíl kuřáků v sociální skupině je u mužů a žen stejný).
H_A: Frekvence kouření je u mužů a žen rozdílná.

Zvolená hladina významnosti: 0,01.

Při testování nulové hypotézy můžeme postupovat tak, jak bylo popsáno u kontingenční tabulky v předchozím příkladě. Mohli bychom nejdříve vypočítat očekávané četnosti pro jednotlivá pole a potom hodnotu testového kritéria χ^2 podle vztahu (32). V případě čtyřpolní tabulky lze však výpočet χ^2 usnadnit použitím vztahu

$$\chi^2 = n \cdot \frac{(ad - bc)^2}{(a+b)\cdot(a+c)\cdot(b+d)\cdot(c+d)} \tag{36}$$

Význam písmen ve vzorci je patrný ze schématu čtyřpolní tabulky (tab. 24).

Čtyřpolní tabulka má jeden stupeň volnosti, vypočítanou hodnotu χ^2 proto srovnáváme s kritickou hodnotou pro jeden stupeň volnosti a zvolenou hladinu významnosti (0,01).

Tab. 24 *Schéma čtyřpolní tabulky*

	α	*non* α	
β	a	b	$a+b$
non β	c	d	$c+d$
	$a+c$	$b+d$	

Po dosazení číselných hodnot do vzorce (36) dostáváme

$$\chi^2 = 48 \cdot \frac{(11 \cdot 21 - 1 \cdot 15)^2}{12 \cdot 36 \cdot 22 \cdot 26} = 9{,}063$$

Zjišťujeme, že vypočítaná hodnota χ^2 je větší než hodnota kritická $\chi^2_{0{,}01}(1) = 6{,}635$, proto odmítáme nulovou hypotézu a přijímáme hypotézu alternativní. Ve sledované sociální skupině se frekvence kouření u mužů a žen liší. Muži kouří častěji než ženy (rozdíl mezi podílem kuřáků – mužů ve skupině a podílem kuřaček – žen ve skupině je statisticky významný). Riziko omylu (pravděpodobnost chybného rozhodnutí) je v uvedeném případě menší než 1 %.

Test nezávislosti chí-kvadrát pro čtyřpolní tabulku lze použít v případech, kdy celková četnost n > 40. Jestliže pro celkovou četnost *n* platí, že $20 < n \leq 40$, potom lze testu použít pouze tehdy, jestliže žádná očekávaná četnost není menší než 5. Pokud jsou četnosti ve čtyřpolní tabulce příliš malé, lze k analýze použít tzv. **Fisherův kombinatorický test**.

Možnosti analýzy na PC
Statistica.cz: neparametrická statistika → 2 × 2 tabulky; základní statistiky → popisné statistiky → kontingenční tabulky
Excel: vložit funkci → statistické → Chi-test
SPSS: *nonparametric tests* → *crosstabs*

3.2.4 FISHERŮV KOMBINATORICKÝ TEST

Tento statistický test je možno použít i v těch případech, kdy četnosti ve čtyřpolní tabulce jsou tak malé, že užití testu chí-kvadrát je vyloučeno.

Příklad 24

V dotazníkovém šetření, kterého se zúčastnilo celkem 32 respondentů, uvedlo 8 dotázaných, že pravidelně chodí do divadla. Dále se zjistilo, že z těch respondentů, kteří pravidelně chodí do divadla, jich má 5 ukončené vysokoškolské vzdělání. Celkový počet respondentů s vysokoškolským vzděláním byl 10. Výsledky uvedeného dotazníkového šetření uvádí tabulka 25.

Máme rozhodnout, zda stupeň dosaženého vzdělání významným způsobem ovlivňuje navštěvování divadla.

Tab. 25 *Čtyřpolní tabulka s malými četnostmi*

		chodí pravidelně do divadla		
		ano	ne	Σ
vysokoškolské vzdělání	ano	5	5	10
	ne	3	19	22
	Σ	8	24	32

Fisherův kombinatorický test začínáme opět formulováním nulové a alternativní hypotézy:

H_0: Podíl respondentů, kteří uvádějí pravidelnou návštěvu divadla, je stejný u osob s vysokoškolským vzděláním i u osob bez vysokoškolského vzdělání.
H_A: Respondenti s vysokoškolským vzděláním uvádějí častěji než respondenti se základním vzděláním, že pravidelně chodí do divadla.

Testování významnosti budeme provádět na hladině významnosti 0,05.

U Fisherova testu postupujeme poněkud odlišným způsobem než u ostatních testů významnosti uváděných v této práci. Zatímco u ostatních testů vypočítáváme testové kritérium a to potom srovnáváme s kritickou hodnotou, u tohoto testu vypočítáváme přímo pravděpodobnost, se kterou bychom mohli nesprávně odmítnout nulovou hypotézu, ač by byla správná. Počítáme tedy přímo významnost (signifikanci), kterou potom srovnáváme se zvolenou hladinou významnosti. Pokud je vypočítaná pravděpodobnost menší než zvolená hladina významnosti, odmítáme nulovou hypotézu a přijímáme hypotézu alternativní. Pokud vypočítaná pravděpodobnost je větší než zvolená hladina významnosti, platnost nulové hypotézy připouštíme.

Abychom mohli příslušnou pravděpodobnost vypočítat, musíme kromě základní čtyřpolní tabulky sestavit ještě další (pomocné tabulky). Postupujeme přitom tak, že nejdříve vyhledáme v základní tabulce nejmenší četnost. Pomocné tabulky potom vytvoříme tak, že tuto nejmenší četnost postupně zmenšujeme o jedničku, až dospějeme k tabulce, v níž minimální četnost je nulová. Marginální (okrajové) četnosti přitom zůstávají u pomocných tabulek nezměněné. Pro základní čtyřpolní tabulku (tab. 25) dostáváme popsaným způsobem další tři pomocné tabulky.

Tab. 26 *Pomocné tabulky pro Fisherův test*

A

6	4	10
2	20	22
Σ 8	24	32

B

7	3	10
1	21	22
Σ 8	24	32

C

8	2	10
0	22	22
Σ 8	24	32

Pro všechny vytvořené tabulky (základní i pomocné) vypočítáme pravděpodobnost p_i, se kterou může tato konfigurace výsledků nastat za předpokladu platnosti nulové hypotézy. Lze odvodit, že tato pravděpodobnost je

$$p_i = \frac{(a+b)!\,(c+d)!\,(a+c)!\,(b+d)!}{n!\,a!\,b!\,c!\,d!} \tag{37}$$

kde n je celková četnost ve čtyřpolní tabulce a hodnoty a, b, c, d jsou četnosti v jednotlivých polích čtyřpolní tabulky podle schématu (tab. 24).

Výsledná pravděpodobnost (signifikance) p se potom určí jako součet všech hodnot p_i, tj.

$$p = \sum p_i \tag{38}$$

Pro hodnoty uvedené v příkladu vychází

$$p_i = \frac{10!\,22!\,24!\,8!}{32!\,5!\,5!\,3!\,19!} = 0{,}036896$$

$$p_{iA} = \frac{10!\,22!\,24!\,8!}{32!\,6!\,4!\,2!\,20!} = 0{,}004611$$

$$p_{iB} = \frac{10!\,22!\,24!\,8!}{32!\,7!\,3!\,1!\,21!} = 0{,}000251$$

$$p_{iC} = \frac{10!\,22!\,24!\,8!}{32!\,8!\,2!\,0!\,22!} = 0{,}000004$$

$$p = 0{,}036896 + 0{,}004611 + 0{,}000251 + 0{,}000004 = 0{,}041$$

Vypočítaná hodnota pravděpodobnosti p vypovídá o tom, že daný výsledek a výsledky ještě nepříznivější mohou nastat (za předpokladu platnosti nulové hypotézy) s pravděpodobností asi 0,041 = 4,1 %. Riziko neoprávněného odmítnutí nulové hypotézy je tedy 0,041, což je hodnota menší než zvolená hladina významnosti 0,05. Proto můžeme na hladině významnosti 0,05 odmítnout nulovou hypotézu a přijmout hypotézu alternativní. Osoby s vysokoškolským vzděláním tedy častěji uvádějí, že pravidelně chodí do divadla.

Možnosti analýzy na PC
Statistica.cz: neparametrická statistika → 2 × 2 tabulky

3.2.5 STUPEŇ ZÁVISLOSTI MEZI JEVY PŘI NOMINÁLNÍM MĚŘENÍ

Pomocí testu chí-kvadrát můžeme objektivně a s předem stanoveným rizikem nesprávného závěru rozhodnout o existenci závislosti mezi dvěma pedagogickými jevy. Výsledek testu významnosti však nevypovídá o stupni této závislosti. K posouzení stupně závislosti mezi jevy v kontingenční tabulce byla navržena řada koeficientů.

3.2.5.1 Koeficienty kontingence

Jako míry závislosti mezi jevy v kontingenční tabulce lze použít **koeficient kontingence** C, který je možno vypočítat podle vzorce

$$C = \sqrt{\frac{\chi^2}{n + \chi^2}} \tag{39}$$

kde n je celková četnost v kontingenční tabulce a χ^2 *je* vypočítaná hodnota testového kritéria chí-kvadrát pro kontingenční tabulku.

Koeficient kontingence C nabývá hodnot v intervalu od 0 do +1, přičemž platí, že čím vyšší je jeho hodnota, tím vyšší je stupeň závislosti. Nevýhodou tohoto koeficientu je, že jeho hodnota závisí na počtu řádků a na počtu sloupců v kontingenční tabulce, a že tudíž dva koeficienty kontingence nelze srovnávat, pokud byly vypočítány z tabulek, které nemají stejné rozměry (tj. stejný počet řádků a sloupců).

Nevýhody koeficientu C do jisté míry odstraňuje normovaný koeficient kontingence C_{norm}. Normovaný koeficient kontingence je definován jako poměr mezi dosaženou hodnotou koeficientu kontingence a největší možnou hodnotou C_{max}, které může koeficient kontingence dosáhnout

$$C_{\max} = \sqrt{\frac{r-1}{r}} \tag{40}$$

$$C_{norm} = \frac{C}{C_{\max}} = \sqrt{\frac{\frac{\chi^2}{n+\chi^2}}{\frac{r-1}{r}}} \tag{41}$$

kde C_{max} je největší možná hodnota koeficientu kontingence, C je dosažená hodnota koeficientu kontingence a r je menší z počtu řádků nebo sloupců v kontingenční tabulce.

Příklad 25

Výpočet normovaného koeficientu kontingence C_{norm} budeme ilustrovat na příkladu kontingenční tabulky, která byla analyzována v oddíle 3.2.2 (tab. 16).

V této kontingenční tabulce je celková četnost $n = 400$ a vypočítaná hodnota $\chi^2 = 6{,}628$. Tabulka má dva řádky a tři sloupce, tudíž $r = 2$. Dosadíme-li příslušné hodnoty do vzorců (39), (40) a (41), dostáváme

$$C = \sqrt{\frac{6{,}628}{400 + 6{,}628}} = 0{,}128$$

$$C_{\max} = \sqrt{\frac{2-1}{2}} = 0{,}707$$

$$C_{norm} = \frac{0{,}128}{0{,}707} = 0{,}181$$

Vypočítaná hodnota $C_{norm} = 0{,}181$ ukazuje, že vztah mezi oběma proměnnými není příliš těsný.

Hodnota C_{norm} nebývá příliš spolehlivá, zvláště u tabulek, kde je velký rozdíl mezi počtem řádků a sloupců. Při srovnávání hodnot C_{norm} pro dvě kontingenční tabulky je dále třeba, aby mezi celkovými četnostmi v obou tabulkách nebyl příliš velký rozdíl.

Čuprovův koeficient K

V případě kontingenční tabulky s rozdílným počtem řádků a sloupců bývá také často doporučován **Čuprovův koeficient *K***, který je možno vypočítat podle vzorce

$$K = \sqrt{\frac{\sqrt{\chi^2}}{\sqrt{n(r-1)\cdot(s-1)}}} \tag{42}$$

kde χ^2 je vypočítaná hodnota testového kritéria chí-kvadrát, n je celková četnost v kontingenční tabulce, r je počet řádků v tabulce a s je počet sloupců v tabulce.

3.2.5.2 Těsnost vztahu mezi jevy ve čtyřpolní tabulce

Fí-koeficient

Jestliže potřebujeme vyjádřit stupeň závislosti mezi oběma alternativními znaky ve čtyřpolní tabulce, potom můžeme vypočítat tzv. **fí-koeficient** r_ϕ. Tento koeficient vychází ze čtyřpolní tabulky a lze ho určit pomocí vzorce

$$r_\Phi = \frac{ad - bc}{\sqrt{(a+b)\cdot(c+d)\cdot(a+c)\cdot(b+d)}} \tag{43}$$

kde význam písmen je stejný jako při výpočtu testového kritéria chí-kvadrát pro čtyřpolní tabulku (srov. oddíl 3.2.3).

Pokud máme pro čtyřpolní tabulku již vypočítáno testové kritérium chí-kvadrát, můžeme fí-koeficient vypočítat snadno ze vztahu

$$r_\Phi = \sqrt{\frac{\chi^2}{n}} \tag{44}$$

Hodnota fí-koeficientu se může pohybovat od –1 do +1, přičemž platí, že čím vyšší je vypočítaná absolutní hodnota, tím vyšší je stupeň závislosti mezi znaky. Znaménko vypočítaného koeficientu je dáno znaménkem výrazu *ad* – *bc* a nemá pro interpretaci výsledku význam. Fí-koeficient nebývá vždy spolehlivým ukazatelem stupně závislosti mezi sledovanými vlastnostmi (jevy). Jeho použití není vhodné například v některých extrémních případech závislostí.

Yulův koeficient asociace

Někdy bývá k posouzení stupně závislosti mezi jevy ve čtyřpolní tabulce doporučován tzv. Yulův koeficient asociace *Q*

$$Q = \frac{ad - bc}{ad + bc} \tag{45}$$

kde význam písmen *a*, *b*, *c*, *d* je stejný jako u fí-koeficientu.

Yulův koeficient asociace může opět nabývat hodnot od 0 do ±1, ale na rozdíl od fí-koeficientu zachycuje jen **jednostrannou závislost mezi jevy**, a poskytuje tudíž také jiné výsledky než fí-koeficient.

Příklad 26

Pro ilustraci vypočítáme hodnoty koeficientů r_ϕ a *Q* pro data uvedená ve čtyřpolní tabulce (tab. 23)

$$r_\Phi = \frac{11\cdot 21 - 1\cdot 15}{\sqrt{12\cdot 36\cdot 22\cdot 26}} = 0,43$$

$$Q = \frac{11 \cdot 21 - 1 \cdot 15}{11 \cdot 21 + 1 \cdot 15} = 0,88$$

Výsledky výpočtů jsou rozdílné, protože každý z koeficientů je odvozen z jiných předpokladů.

Možnosti analýzy na PC
Statistica.cz: neparametrická statistika → tabulky 2 × 2

3.3 STATISTICKÉ METODY PRO ANALÝZU ORDINÁLNÍCH DAT

Jestliže ověřujeme vztah mezi proměnnými, které byly změřeny na úrovni ordinálního (pořadového) měření, můžeme k tomu použít (po vhodné kategorizaci) všech postupů uvedených pro nominální data. Můžeme ale využít také postupů, které byly vyvinuty speciálně pro data ordinální.

3.3.1 ZNAMÉNKOVÝ TEST

Tento statistický test významnosti se užívá v případě dvou opakovaných měření na týchž objektech. Hodnoty naměřené u jednoho objektu ve výběru tvoří vždy pár. Pomocí znaménkového testu lze rozhodnout, zda mezi oběma opakovanými měřeními týchž objektů je významný rozdíl, či nikoli. Jedná se o velmi jednoduchý test, který lze použít všude tam, kde je možno proměnné zachytit (změřit) alespoň na ordinální úrovni (tzn. v případech, kdy je možno alespoň rozhodnout, která z opakovaně naměřených hodnot je vyšší).

Příklad 27

U skupiny 15 dětí byla měřena frekvence mrkání oka v klidové situaci (při volné hře) a při sledování napínavého televizního programu. Výsledky obou měření uvádí tabulka 27.

Máme rozhodnout, zda při sledování napínavého televizního programu je frekvence mrkání oka vyšší než v klidové situaci.

Pro danou situaci nejdříve formulujeme nulovou a alternativní hypotézu:

H_0: Mezi frekvencí mrkání oka v klidové situaci a frekvencí mrkání oka při sledování napínavého programu není rozdíl (případné rozdíly je možno připsat na vrub náhody).
H_A: Frekvence mrkání oka je při sledování napínavého televizního programu vyšší než v klidové situaci.

U znaménkového testu se nejdříve pomocí znamének + a – vyjádří, zda u jednotlivých dětí došlo ke zvětšení nebo zmenšení frekvence mrkání oka (tab. 24, sloupec „změna“). Zjišťujeme, že v 11 případech došlo ke zvýšení frekvence mrkání, zatímco ve čtyřech případech frekvence mrkání poklesla. Pro další úvahy je u znaménkového testu rozhodující počet těch znamének, která se vyskytují méně často (znaménka řidčeji se vyskytujícího druhu). V našem případě se vyskytuje méně znamének –, jejich počet je 4.

Znaménkový test je založen na úvaze, že v případě, že by nebyl mezi oběma měřeními rozdíl, měla by se obě znaménka vyskytovat se stejnou pravděpodobností, tj. měl by jich být stejný počet. Rozdíl mezi oběma opakovanými měřeními se projeví tím, že znaménka jednoho druhu začnou převažovat nad znaménky druhého druhu, a to tím více, čím tento rozdíl bude výraznější. Existují statistické tabulky (příloha III), které umožňují určit, kolikrát se může znaménko, jehož výskyt je méně častý, objevit, máme-li považovat rozdíl mezi měřeními ještě za statisticky nevýznamný.

Nahlédneme-li do tabulky v příloze III, zjišťujeme, že při 15 dvojicích naměřených dat znamená výskyt čtyř znamének řidčeji se vyskytujícího druhu ještě statisticky *nevýznamný* výsledek. Nemůžeme proto odmítnout nulovou hypotézu a konstatujeme, že z naměřených hodnot nelze usuzovat na významné zvýšení frekvence mrkání oka při sledování daného televizního programu.

Tab. 27 *Znaménkový test*

Dítě č.	Frekvence mrkání oka		Změna
	v klidu	při sledování TV	
1	10	11	+
2	8	10	+
3	9	8	–
4	15	14	–
5	12	13	+
6	13	15	+
7	11	13	+
8	14	12	–
9	10	11	+
10	11	13	+
11	12	14	+
12	13	14	+
13	17	16	–
14	16	19	+
15	12	15	+

Znaménkový test je velmi jednoduchý a neklade velké nároky na prováděné měření. Je použitelný i tehdy, jestliže můžeme pouze rozhodnout, jaký směr měla změna, k níž došlo. Jeho nevýhodou však je to, že je málo citlivý při odhalování malých rozdílů mezi opakovanými měřeními. V těchto případech bývá výhodnější použít například Wilcoxonův test, který je ve srovnání se znaménkovým testem citlivější.

Možnosti analýzy na PC
SPSS: *advanced procedures → nonparametric metods → comparison of two samples (volba sign test)*
Statistica.cz: neparametrické statistiky → porovnávání dvou závislých výběrů → znaménkový test

3.3.2 WILCOXONŮV TEST

Tento statistický test významnosti se používá v podobných situacích jako znaménkový test, tedy v případě opakovaných měření týchž objektů. Podmínkou pro jeho použití je, že data, se kterými se pracuje, musí být alespoň ordinální (pořadová). Výhodou Wilcoxonova testu (ve srovnání se znaménkovým testem) je jeho větší účinnost, tj. spíše jím odhalíme malé rozdíly mezi oběma měřeními.

Příklad 28

Použití Wilcoxonova testu vysvětlíme na stejné situaci, která byla analyzována u znaménkového testu. Budeme tedy znovu rozhodovat o tom, zda při sledování napínavého televizního programu je frekvence mrkání oka dětí vyšší než v klidové situaci.
Nulová a alternativní hypotéza budou mít stejnou podobu jako u znaménkového testu.

H_0: Mezi frekvencí mrkání oka v klidové situaci a frekvencí mrkání oka při sledování napínavého programu není rozdíl (případné rozdíly je možno připsat na vrub náhody).
H_A: Frekvence mrkání oka je při sledování napínavého televizního programu jiná než v klidové situaci.

U Wilcoxonova testu postupujeme tak, že nejdříve u každé dvojice hodnot (u každého dítěte) určíme diferenci *d* mezi oběma zjištěnými frekvencemi mrkání oka. Jednotlivým diferencím potom přiřadíme pořadí podle jejich absolutních hodnot.

Tab. 28 *Wilcoxonův test*

Dítě č.	Frekvence mrkání oka		d	Pořadí	+	–
	v klidu	při sledování televize				
1	10	11	–1	4		4
2	8	10	–2	10,5		10,5
3	9	8	1	4	4	
4	15	14	1	4	4	
5	12	13	–1	4		4
6	13	15	–2	10,5		10,5
7	11	13	–2	10,5		10,5
8	14	12	2	10,5	10,5	

▶

Dítě č.	Frekvence mrkání oka		d	Pořadí	+	–
	v klidu	při sledování televize				
9	10	11	–1	4		4
10	11	13	–2	10,5		10,5
11	12	14	–2	10,5		10,5
12	13	14	–1	4		4
13	17	16	1	4	4	
14	16	19	–3	14,5		14,5
15	12	15	–3	14,5		14,5
				∑	22,5	97,5

Nejmenší diference mezi naměřenými hodnotami $|d| = 1$ jsou u sedmi dětí. Těmto diferencím přiřadíme průměrné pořadí (1 + 2 + 3 + 4 + 5 + 6 + 7) / 7 = 4.

Diference $|d| = 2$ jsou u šesti dětí, a proto jim přiřadíme pořadí (8 + 9 + 10 + 11 + 12 + 13) / 6 = 10,5. Diference $|d| = 3$ je u dvou dětí, a přísluší jim proto pořadí 14,5.

Stanovená pořadí diferencí dále rozdělíme podle znaménka do dvou sloupců a každý sloupec sečteme. Menší z obou součtů označíme *T* (v našem případě *T* = *22,5*). Hodnota *T* je testovým kritériem pro Wilcoxonův test. Vypočítanou hodnotu *T* srovnáváme s kritickou hodnotou tohoto kritéria (příloha IV). Pro 15 párů hodnot a hladinu významnosti 0,05 je tabelována kritická hodnota $T_{0,05}(15) = 25$.

Nulovou hypotézu u Wilcoxonova testu zamítáme, jestliže vypočítaná hodnota *T* je *menší nebo rovna* hodnotě kritické. Vzhledem k tomu, že v uvedeném příkladě je vypočítaná hodnota *T* menší než hodnota kritická, odmítáme nulovou hypotézu a přijímáme hypotézu alternativní. Pomocí Wilcoxonova testu jsme tedy prokázali, že mezi oběma provedenými měřeními (frekvence mrkání oka při sledování napínavého televizního programu a frekvence mrkání oka v klidové situaci) jsou na hladině významnosti 0,05 statisticky významné rozdíly.

Poznámka: *Při určování diferencí u Wilcoxonova testu se berou v úvahu jen nenulové diference a také při vyhledávání kritické hodnoty T se počítá jen s počtem dvojic hodnot, kde byly nenulové diference.*

Možnosti analýzy na PC
SPSS: *nonparametric tests → related samples*
Statistica.cz: neparametrické statistiky → porovnávání dvou závislých výběrů → Wilcoxonův test

3.3.3 U-TEST MANNA A WHITNEYHO

Je to velmi vydatný neparametrický test, který lze použít v případech, kdy máme rozhodnout, zda dva výběry mohou pocházet ze stejného základního souboru, tj. zda mají stejné rozdělení četností.

3.3.3.1 U-test pro velmi malé výběry

U tohoto statistického testu se naměřené hodnoty z obou výběrů uspořádávají do jedné řady podle velikosti. U každé hodnoty z prvního výběru se potom zjišťuje, kolik hodnot z druhého výběru jí předchází. Výsledky všech těchto zjištění se sečtou a označí jako *U*. Podle toho, který z obou výběrů zvolíme jako první a který jako druhý, můžeme dostat dva výsledky, které označujeme jako *U* a *U'*. Menší z těchto dvou hodnot je testovým kritériem pro U-test.

U velmi malých výběrů lze testové kritérium *U* určit celkem snadno pouhým seřazením čísel a přímým spočítáním.

Příklad 29

Ve dvou skupinách žáků byly získány následující výsledky (počty bodů v didaktickém testu):

skupina A:	5	7	12	13	15		($n_1 = 5$)
skupina B:	6	9	11	14	16	18	($n_2 = 6$)

Máme rozhodnout, zda mezi výsledky obou skupin žáků je statisticky významný rozdíl.

H_0: *Mezi výsledky žáků v obou skupinách nejsou rozdíly.*
H_A: *Mezi výsledky žáků v obou skupinách jsou rozdíly.*

Zvolená hladina významnosti: 0, 05.

Hodnoty z obou skupin uspořádáme do jedné řady podle velikosti a u každé hodnoty poznamenáme, ze které skupiny pochází:

5	6	7	9	11	12	13	14	15	16	18
(A)	B	(A)	B	B	(A)	(A)	B	(A)	B	B

Hodnotě 5 z výběru A nepředchází žádná hodnota z výběru B, hodnotě 7 z výběru A předchází 1 hodnota z výběru B, hodnotě 12 z výběru A předcházejí 3 hodnoty z výběru B, hodnotě 13 z výběru A předcházejí 3 hodnoty z výběru B a hodnotě 15 z výběru A předcházejí 4 hodnoty z výběru B:

$U = 0 + 1 + 3 + 3 + 4 = 11.$

Podobně můžeme vyjádřit, kolik hodnot z výběru A předchází před jednotlivými hodnotami z výběru B:

$U' = 1 + 2 + 2 + 4 + 5 + 5 = 19.$

Testovým kritériem je menší z obou vypočítaných hodnot, tj. $U = 11$.

Poznámka: *Při výpočtech není třeba určovat obě hodnoty (U i U'), protože mezi veličinami U a U' existuje vztah*

$$U + U' = n_1 \cdot n_2 \qquad (46)$$

Vypočítanou hodnotu *U* srovnáváme s kritickou hodnotou tohoto kritéria pro zvolenou hladinu významnosti a pro dané rozsahy výběrů (příloha V).

Jestliže vypočítaná hodnota *U je menší nebo rovna* hodnotě kritické, odmítáme na zvolené hladině významnosti nulovou hypotézu a přijímáme hypotézu alternativní.

V uvedeném příkladě je vypočítaná hodnota $U = 11$ a kritická hodnota pro zvolenou hladinu významnosti a dané četnosti ve skupinách ($n_1 = 5$ a $n_2 = 6$) $U_{0,05}(5,6) = 3$. Nemůžeme proto odmítnout nulovou hypotézu. Mezi výsledky žáků v obou skupinách nebyly prokázány statisticky významné rozdíly.

Popsaný způsob provedení U-testu je možno doporučit, pokud četnosti ve srovnávaných skupinách jsou menší než 8.

3.3.3.2 U-test pro větší skupiny

U větších skupin by přímé stanovení testového kritéria *U* shora popsaným způsobem již bylo značně nepohodlné, a proto se postupuje jinak. Naměřeným hodnotám (dohromady v obou skupinách) se přiřazují pořadí podle velikosti, a to tak, že pořadí 1 přiřadíme hodnotě nejmenší. Testové kritérium *U* (respektive *U*') je potom možno vypočítat ze vztahů

$$U = n_1 \cdot n_2 + \frac{n_1 \cdot (n_1 + 1)}{2} - R_1 \tag{47}$$

$$U' = n_1 \cdot n_2 + \frac{n_2 \cdot (n_2 + 1)}{2} - R_2 \tag{48}$$

kde n_1 je četnost hodnot v prvním výběru, n_2 četnost hodnot v druhém výběru, R_1 je součet pořadí v první skupině a R_2 je součet pořadí ve druhé skupině.

Pro testování statistické významnosti volíme opět z hodnot *U* a *U*′ hodnotu menší.

Příklad 30

Skupině náhodně vybraných žáků byl zadán didaktický test z fyziky, ve kterém bylo možno získat od 0 do 10 bodů. Ve skupině bylo 20 chlapců a 15 dívek.

Výsledky chlapců:	7, 10, 9, 8, 6, 7, 4, 6, 3, 7, 8, 5, 4, 9, 5, 6, 10, 8, 9, 7
Výsledky dívek:	4, 3, 5, 4, 6, 4, 1, 5, 9, 2, 10, 8, 3, 6, 5

Máme rozhodnout, zda mezi výsledky chlapců a dívek jsou rozdíly

H_0: Mezi dosaženými výsledky obou skupin žáků nejsou rozdíly.
H_A: Výsledky chlapců a dívek v testu z fyziky jsou rozdílné.

Zvolená hladina významnosti: 0,05.

Nejdříve se seřadí dosažené výsledky chlapců i dívek podle velikosti (tab. 29).

Jednotlivým dosaženým výsledkům (u chlapců i dívek) přiřadíme pořadí podle velikosti. Nejmenší dosažená hodnota v testu je 1 bod, tomuto výsledku přiřadíme pořadí 1. Druhá nejmenší hodnota je 2 body – přiřadíme pořadí 2. Výsledku 3 body dosáhli tři žáci, kterým přiřadíme průměrné pořadí, tj. (3 + 4 + 5) / 3 = 4 atd.

Jestliže dosadíme příslušné hodnoty do vzorců (47) a (48), dostáváme

$$U = 20 \cdot 15 + \frac{20 \cdot (20+1)}{2} - 429,0 = 81$$

$$U' = 20 \cdot 15 + \frac{15 \cdot (15+1)}{2} - 201,0 = 219$$

Testovým kritériem je menší z obou vypočítaných hodnot, tj. U = 81. Tuto vypočítanou hodnotu srovnáváme s kritickou hodnotou $U_{0,05}$ (20,15) = 90 pro zvolenou hladinu významnosti a četnosti v obou skupinách (příloha V). Protože vypočítaná hodnota U je menší než hodnota kritická, odmítáme nulovou hypotézu a přijímáme hypotézu alternativní. Mezi dosaženými výsledky chlapců a dívek v testu z fyziky je (na hladině významnosti 0,05) *statisticky významný* rozdíl.

Popsaný postup provedení U-testu lze doporučit u skupin, které mají četnosti přibližně od 9 do 20.

Tab. 29 *U-test Manna a Whitneyho*

Chlapci	
Počet bodů	**Pořadí**
3	4
4	8
4	8
5	13
5	13
6	18
6	18
6	18
7	22,5
7	22,5
7	22,5
7	22,5
8	26,5
8	26,5
8	26,5
9	30,5
9	30,5
9	30,5
10	34
10	34
n_1 = 20	R_1 = 429,0

Dívky	
Počet bodů	**Pořadí**
1	1
2	2
3	4
3	4
4	8
4	8
4	8
5	13
5	13
5	13
6	18
6	18
8	26,5
9	30,5
10	34
n_2 = 15	R_2 = 201,0

3.3.3.3 U-test při velkých četnostech

U-test je vhodný zejména pro testování menších skupin (zhruba do $n = 20$). Mann a Whitney prokázali, že pro velké skupiny (větší než 20) má testové kritérium přibližně normální rozdělení. Nulovou hypotézu lze potom testovat pomocí **normované normální veličiny** *u* (srov. oddíl 2.4). Postupujeme tak, že nejdříve podle vztahu (47), respektive (48), určíme hodnotu testového kritéria U a pomocí této hodnoty potom vypočítáme normovanou náhodnou veličinu u

$$|u| = \frac{U - \frac{n_1 \cdot n_2}{2}}{\sqrt{\frac{n_1 \cdot n_2 \cdot (n_1 + n_2 + 1)}{12}}} \tag{49}$$

kde U je vypočítaná hodnota testového kritéria, n_1 je četnost jedné skupiny a n_2 je četnost druhé skupiny.

Příklad 31

Pomocí U-testu máme ověřit, zda mezi výsledky dvou skupin žáků v didaktickém testu jsou rozdíly. První skupina ($n_1 = 16$) dosáhla výsledky: 8, 7, 2, 1, 3, 3, 2, 7, 4, 2, 5, 5, 5, 2, 2, 5 bodů. Druhá skupina ($n_2 = 23$) dosáhla výsledky: 7, 7, 7, 3, 3, 1, 10, 10, 10, 7, 8, 8, 9, 9, 9, 5, 5, 5, 8, 19, 17, 6, 6 bodů.

Pomocí normované normální veličiny *u* máme rozhodnout, zda rozdíly mezi testovými výsledky obou skupin žáků jsou (na hladině významnosti 0,05) statisticky významné.

Nejdříve byly formulovány statistické hypotézy:

H_0: Mezi výsledky žáků v obou skupinách nejsou rozdíly.
H_A: Mezi výsledky obou skupin žáků jsou rozdíly.

Zvolená hladina významnosti: 0,01.

Výsledky dosažené v obou skupinách byly seřazeny podle velikosti a zapsány do tabulky (tab. 30).

Způsobem, který byl popsán v předchozím příkladu, byly vypočítány hodnoty

$$U = 16 \cdot 23 + \frac{16 \cdot (16+1)}{2} - 200 = 304$$

$$U' = 16 \cdot 23 + \frac{23 \cdot (23+1)}{2} - 580 = 64$$

Testovým kritériem je menší z obou vypočítaných hodnot, tj. $U = 64$.

Dosadíme-li příslušné hodnoty do vztahu (49), dostáváme

$$|u| = \frac{64 - \frac{16 \cdot 23}{2}}{\sqrt{\frac{16 \cdot 23 \cdot (16+23+1)}{12}}} = 3{,}43$$

Vypočítanou hodnotu srovnáváme s kritickou hodnotou $U_{0,05}$ = 1,96 pro hladinu významnosti 0,05 (oboustranný test) nebo s hodnotou $U_{0,05}$ = 2,58 pro hladinu významnosti 0,01 (oboustranný test). Protože vypočítaná hodnota *u* je větší než hodnota kritická pro hladinu významnosti 0,01, odmítáme nulovou hypotézu a přijímáme hypotézu alternativní. Mezi výsledky obou skupin osob tedy jsou na hladině významnosti 0,01 **statisticky významné rozdíly**.

Tab. 30 *U-test při velkých četnostech*

1. skupina		2. skupina	
Počet bodů	**Pořadí**	**Počet bodů**	**Pořadí**
1	1,5	1	1,5
2	5	3	9,5
2	5	3	9,5
2	5	5	16
2	5	5	16
2	5	5	16
3	9,5	6	20,5
3	9,5	6	20,5
4	12	7	24,5
5	16	7	24,5
5	16	7	24,5
5	16	7	24,5
5	16	8	29,5
7	24,5	8	29,5
7	24,5	8	29,5
8	29,5	9	33
$n_1 = 16$	$R_1 = 200$	9	33
		9	33
		10	36
		10	36
		10	36
		17	38
		19	39
		$n_2 = 23$	$R_2 = 580,0$

Jestliže se ve srovnávaných skupinách některé hodnoty opakují, může být vypočítaná hodnota *u* poněkud zkreslená. V těchto případech se doporučuje počítat normovanou normální veličinu podle upraveného vzorce

$$|u'| = \frac{U - \frac{n_1 \cdot n_2}{2}}{\sqrt{\frac{n_1 \cdot n_2}{n \cdot (n-1)} \cdot \frac{n^3 - n}{12} - \sum \frac{r^3 - r}{12}}} \tag{50}$$

kde $|u'|$ je korigovaná absolutní hodnota normované náhodné veličiny, U je vypočítané testové kritérium U, n_1 je četnost první skupiny, n_2 je četnost druhé skupiny, $n = n_1 + n_2$ a r je počet hodnot, které se opakují.

Hodnota 1 se opakuje (v obou srovnávaných skupinách) dvakrát, hodnota 2 se opakuje pětkrát, hodnota 3 se opakuje čtyřikrát, hodnota 5 se opakuje sedmkrát atd.
Potom

$$\sum \frac{r^3 - r}{12} = \frac{2^3-2}{12} + \frac{5^3-5}{12} + \frac{4^3-4}{12} + \frac{7^3-7}{12} + \frac{2^3-2}{12} + \frac{6^3-6}{12} + \frac{4^3-4}{12} + \frac{3^3-3}{12} + \frac{3^3-3}{12} = 70{,}5$$

Podle vzorce (50) byla vypočítána absolutní hodnota normované normální veličiny $|u'|$

$$|u'| = \frac{64 - \frac{16 \cdot 23}{2}}{\sqrt{\frac{16 \cdot 23}{39 \cdot (39-1)} \cdot \frac{39^3 - 39}{12}} - 70{,}5} = 3{,}45$$

Výsledky výpočtů podle vztahů (49) a (50) se v našem případě neliší a vedou ke stejnému závěru. V některých situacích však provádění korekce na opakování hodnot může mít smysl.

Možnosti analýzy na PC
Statistica.cz: neparametrické statistiky → porovnávání dvou nezávislých výběrů → Mann – Whitneyho test
SPSS: *nonparametric tests → related samples*

3.3.4 KOLMOGOROVŮV-SMIRNOVŮV TEST

Tento test je vhodný pro posuzování rozdílů ve složení (struktuře) dvou skupin. Je použitelný v případě tzv. spojitých náhodných veličin. Při použití tohoto testu u diskrétních náhodných veličin klesá výrazně jeho účinnost.

> ***Poznámka:*** *Spojitá náhodná veličina (na rozdíl od tzv. diskrétní náhodné veličiny) je taková veličina, která může nabývat libovolných hodnot (např. čas v běhu žáků na určitou vzdálenost). Příkladem diskrétní náhodné veličiny je např. počet bodů v didaktickém testu, ve kterém lze získat jen celočíselné výsledky (např. 0, 1, 2, 3 … body).*

Kolmogorovův-Smirnovův test pro dva výběry je založen na srovnávání distribučních funkcí ve dvou výběrech. Distribuční funkce náhodné veličiny *F(x)* je pravděpodobnost, že tato veličina dosáhne určité a nebo menší hodnoty. Testové kritérium D je založeno na zkoumání absolutní hodnoty největšího rozdílu distribuční funkce prvního výběru $F_1(x)$ a distribuční funkce druhého výběru $F_2(x)$

$$D = \max \left| F_1(x) - F_2(x) \right| \tag{51}$$

Příklad 32

Použití Kolmogorovova-Smirnovova testu vyložíme na příkladu výzkumu, ve kterém se ověřovaly rozdíly ve výkonech žáků v běhu na určitou vzdálenost, a to u skupiny žáků sportovních tříd a u skupiny žáků běžných tříd (Gajda a Zvolská, 1982).

Výsledky obou skupin žáků jsou patrné z tabulky 31.

H_0: Složení obou skupin žáků je stejné (obě skupiny žáků pocházejí ze stejného základního souboru).
H_A: Mezi složením obou skupin žáků jsou rozdíly (skupiny nepocházejí ze stejného základního souboru).

Zvolená hladina významnosti: 0,05

U obou skupin žáků nejdříve vypočítáme relativní četnosti jednotlivých hodnot v obou skupinách. V první skupině je relativní četnost všech jednotlivých hodnot 1 / 9 = 0,111, ve druhé skupině 1/ 8 = 1,125 (tab. 31). Součet relativních četností v každé skupině je 1.

Tab. 31 *Výsledky žáků sportovních tříd a žáků běžných tříd v běhu*

Žáci sportovních tříd		
číslo žáka	**čas (s)**	**relativní četnost**
1	202	0,111
2	227	0,111
3	237	0,111
4	240	0,111
5	249	0,111
6	251	0,111
7	273	0,111
8	284	0,111
9	213	0,111
$n_1 = 9$		Σ 1,000

Žáci běžných tříd		
číslo žáka	**čas (s)**	**relativní četnost**
1	239	0,125
2	245	0,125
3	252	0,125
4	261	0,125
5	272	0,125
6	281	0,125
7	288	0,125
8	293	0,125
$n_2 = 8$		Σ 1,000

Pro následnou analýzu sestavíme další tabulku, ve které jednotlivé naměřené hodnoty seřadíme vzestupně podle velikosti (tab. 32). Žáci sportovních tříd jsou v tabulce označeni S a žáci běžných tříd B. Tabulka dále obsahuje sloupce kumulativních relativních četností (distribučních funkcí) pro skupinu S a pro skupinu B. Poslední sloupec tabulky (označený D diference) uvádí rozdíly mezi kumulativními relativními četnostmi (distribučními funkcemi) v obou skupinách žáků (největší z těchto diferencí je v tabulce označena šedým podtiskem).

***Tab. 32** Kolmogorovův-Smirnovův test pro dva výběry*

Skupina žáků	**Čas x_i (s)**	**Relativní četnost**		**Kumulativní relativní četnost**		**Diference D**
		skupina S	**skupina B**	**skupina S**	**skupina B**	
S	202	0,111		0,111	0	0,111
S	213	0,111		0,222	0	0,222
S	227	0,111		0,333	0	0,333
S	237	0,111		0,444	0	0,444
B	239		0,125	0,444	0,125	0,319
S	240	0,111		0,555	0,125	0,430
B	245		0,125	0,555	0,250	0,305
S	249	0,111		0,666	0,250	0,416
S	251	0,111		0,777	0,250	**0,527**
B	252		0,125	0,777	0,375	0,402
B	261		0,125	0,777	0,500	0,277
B	272		0,125	0,777	0,625	0,152
S	273	0,111		0,888	0,625	0,263
B	281		0,125	0,888	0,750	0,138
S	284	0,111		1,000	0,750	0,250
B	288		0,125	1,000	0,875	0,125
B	293		0,125	1,000	1,000	0
		Σ 1,000	Σ 1,000			

K rozhodování o platnosti nulové hypotézy se v praxi používá hodnota C

$$C = D_{max} \cdot M \tag{52}$$

kde D_{max} je největší rozdíl mezi kumulativními relativními četnostmi v obou srovnávaných skupinách, M nejmenší společný násobek četností n_1 (četnost první skupiny) a n_2 (četnost druhé skupiny). Nejmenší společné násobky četností n_1 a n_2 jsou uvedeny společně s kritickými hodnotami C ve statistických tabulkách (příloha VI).

V tomto příkladu je $M = 72$, $D_{max} = 0{,}527$, a proto C = 0,527 · 72 = 37,9. V tabulce v příloze VI zjišťujeme, že pro četnost $n_1 = 9$ a $n_2 = 8$ je tabelována kritická hodnota C (8,9) = 46. Protože vypočítaná hodnota je menší než hodnota kritická, přijímáme na hladině významnosti α = 0,05 nulovou hypotézu. Obě skupiny žáků tedy mají zhruba stejné složení a **mohou pocházet z jednoho základního souboru**. K odmítnutí nulové hypotézy by došlo v případě, že vypočítaná hodnota C by byla stejně velká, nebo větší než hodnota kritická.

Možnosti analýzy na PC
Statistica.cz: neparametrické statistiky → porovnávání dvou nezávislých vzorků → Kolmogorovův-Smirnovův test

3.3.5 KRUSKALŮV-WALLISŮV TEST

Pomocí U-testu je možné rozhodnout, zda mezi dvěma skupinami naměřených hodnot jsou statisticky významné rozdíly. Kruskalův-Wallisův test (který je v podstatě zobecněním U-testu) můžeme použít v situaci, kdy máme rozhodnout, zda ve více než dvou skupinách je stejný medián.

Příklad 33

Princip testu objasníme na příkladu fiktivního výzkumu, ve kterém se ověřovalo, jak dalece ovlivňuje kvalitu vědomostí žáků použitá vyučovací metoda. Výzkum probíhal ve třech stejně velkých skupinách (po čtyřech žácích). V jedné skupině byli žáci vyučováni metodou A, ve druhé skupině metodou B a ve třetí skupině metodou C. Po určité době proběhlo ve všech skupinách měření vědomostí didaktickým testem. Výsledky dosažené v testu uvádí tabulka 33a.

Tab. 33a *Výsledky v didaktickém testu*

Metody výuky		
A	**B**	**C**
19	27	*13*
25	30	*22*
18	22	*24*
18	*29*	*20*
∑ 80	∑ 108	∑ 79

Nejdříve formulujeme statistické hypotézy:

H_0: Mediány jsou ve všech skupinách stejné.
H_A: Alespoň pro jednu dvojici mediánů platí, že $\tilde{x}_i \neq \tilde{x}_j$.

Testování provedeme na hladině významnosti 0,05.

Další postup spočívá v tom, že všem hodnotám v tabulce přiřadíme pořadí podle velikosti a těmito pořadími nahradíme původní hodnoty v tabulce 33b.

Tab. 33b *Kruskalův-Wallisův test*

Metody výuky		
A	**B**	**C**
4	10	1
9	12	6,5
2,5	6,5	8
2,5	11	5
∑ 18	∑ 39,5	∑ 20,5

Testovým kritériem pro Kruskalův-Wallisův test je hodnota H, kterou je možno vypočítat ze vzorce

$$H = \left[\frac{12}{n \cdot (n+1)} \sum \frac{R_i^2}{n_i} \right] - 3(n+1) \tag{53}$$

kde n je celková četnost všech hodnot, R_i je součet pořadí v jednotlivých skupinách a n_i jsou četnosti hodnot v jednotlivých skupinách.

Při velkých četnostech hodnot ve skupinách lze pro testování nulové hypotézy použít statistiku chí-kvadrát. Nulovou hypotézu odmítáme, jestliže vypočítané testové kritérium H je větší než kritická hodnota χ^2. Kritickou hodnotu χ^2 vyhledáváme pro $k - 1$ stupňů volnosti (kde k je počet skupin, jejichž mediány srovnáváme).

V uvedeném příkladu vychází testové kritérium

$$H = \left[\frac{12}{12 \cdot 13} \left(\frac{18^2}{4} + \frac{39{,}5^2}{4} + \frac{20{,}5^2}{4} \right) \right] - 3 \cdot 13 = 5{,}317$$

Kritická hodnota testového kritéria chí-kvadrát pro $k - 1 = 3 - 1 = 2$ stupně volnosti a hladinu významnosti 0,05 je $\chi^2_{0,05}(2) = 5{,}991$. Proto nelze odmítnout nulovou hypotézu o rovnosti mediánů ve skupinách.

Poznámka: *V uvedeném fiktivním příkladu byly četnosti hodnot ve skupinách malé, proto by se k testování nulové hypotézy nemělo testové kritérium chí-kvadrát používat.*

Možnosti analýzy na PC
Statistica.cz: neparametrické statistiky → porovnávání nezávislých vzorků → Kruskalova-Wallisova ANOVA a mediánový test

3.3.6 STUPEŇ ZÁVISLOSTI MEZI JEVY PŘI ORDINÁLNÍM MĚŘENÍ

Pomocí výše uvedených testů významnosti bylo možno rozhodnout, zda mezi jevy, které byly zachyceny pomocí ordinálního (pořadového) měření, existuje statisticky významná závislost či nikoli. Zpravidla nás také zajímá, jak těsná tato závislost je (jak velký je stupeň závislosti mezi sledovanými jevy).

3.3.6.1 Spearmanův koeficient pořadové korelace

Tento statistický postup lze využít v případech, kdy máme rozhodnout, jak těsně spolu souvisí dvě proměnné, které byly zachyceny (změřeny) pomocí ordinálního měření. Často například můžeme u skupiny žáků pozorovat vlastnosti, které nám umožňují seřadit je podle míry této vlastnosti. Žáky můžeme například seřazovat podle rychlosti v běhu, podle zručnosti při práci, podle věku, podle hmotnosti (váhy), podle schopnosti zapamatovat si, podle výsledku v určitém didaktickém testu atd. Spearmanův koeficient pořadové korelace umožňuje kvantitativně stanovit, jak dalece jsou si dvě vytvořená pořadí podobná, a tím vlastně určit, jak těsná je souvislost mezi jevy, na základě nichž byla tato pořadí vytvořena.

Příklad 34

Výpočet Spearmanova koeficientu pořadové korelace objasníme na příkladu výzkumného šetření, ve kterém se zjišťovalo, jak těsný je vztah mezi hmotností dětí a jejich rychlostí v běhu. U vybrané skupiny 10 dětí bylo zaznamenáno, v jakém pořadí doběhly do cíle při závodění v běhu (tab. 34). V tabulce je také u všech dětí uvedena jejich hmotnost.

Tab. 34 *Výpočet Spearmanova koeficientu pořadové korelace*

Dítě	Hmotnost (kg)	Pořadí podle hmotnosti	Pořadí podle rychlosti v běhu	d	d^2
Kristina	16,70	9	1	8	64
Vladimír	18,45	7	2	5	25
Milan	16,70	9	3	6	36
Lenka	19,50	6	4	2	4
Robert	16,70	9	5	4	16
Ivan	23,30	2	6	–4	16
Jiří	23,40	1	7	–6	36
Zuzana	19,80	4,5	8	–3,5	12,25
Blanka	19,80	4,5	9	–4,5	20,25
Věra	22,20	3	10	–7	49
					Σ 278,50

Uvedené hmotnosti dětí představují poměrová data. Abychom z těchto dat mohli vypočítat Spearmanův koeficient pořadové korelace, musíme je nejdříve převést na pořadí. Pořadí podle hmotnosti vytvoříme tak, že dítěti, u něhož byla zjištěna největší hmotnost, přiřadíme pořadí 1 a dětem s menší hmotností potom postupně pořadí 2, 3, … atd. Pokud dvě nebo více dětí mají stejnou hmotnost, přiřadíme jim **průměrné pořadí**. Stejnou hmotnost mají například Zuzana a Blanka, které se dělí o 4. a 5. místo ve skupině. Přiřadíme jim proto pořadí (4 + 5) / 2 = 4,5. Kristina, Milan a Robert mají také stejnou hmotnost a dělí se o 8. až 10. místo ve skupině. Proto jim přiřadíme průměrné pořadí (8 + 9 + 10) / 3 = 9.

U každého dítěte v tabulce vypočítáme rozdíl mezi oběma vytvořenými pořadími d a hodnotu d^2. Výpočet Spearmanova koeficientu pořadové korelace se provádí podle vzorce

$$r_s = 1 - \frac{6 \cdot \sum d^2}{n \cdot (n^2 - 1)} \tag{54}$$

kde r_s je Spearmanův koeficient pořadové korelace, n je počet srovnávaných dvojic hodnot (v našem případě počet dětí) a d je rozdíl (diference) pořadí pro jednu dvojici hodnot.

Dosadíme-li do uvedeného vzorce příslušné hodnoty z tabulky, dostáváme

$$r_s = 1 - \frac{6 \cdot 278,5}{10(10^2 - 1)} = -0,69$$

Koeficient r_s může nabývat hodnot od 0 do ±1. Hodnota 0 vypovídá o tom, že mezi srovnávanými jevy není žádný vztah. Čím více se vypočítaná hodnota koeficientu korelace blíží hodnotě 1 (nebo –1), tím těsnější je vztah mezi jevy, které srovnáváme. Kladný výsledek vypovídá o tom, že vyšším hodnotám u jednoho měřeného jevu odpovídají také spíše vyšší hodnoty u druhého jevu a zároveň nižším hodnotám u prvního jevu odpovídají také nižší hodnoty u jevu druhého.

Jestliže je koeficient korelace záporný, znamená to, že mezi jevy, které srovnáváme, je negativní (opačný) vztah, tj. vysokým hodnotám jedné proměnné odpovídají spíše nižší hodnoty druhé proměnné a naopak. Vypočítaný koeficient r_s = –0,69 tedy vypovídá o tom, že mezi hmotností dětí a jejich rychlostí v běhu je negativní vztah, tj. čím větší je hmotnost dítěte, tím menší je jeho rychlost v běhu.

Pro **přibližnou** interpretaci vypočítaného koeficientu korelace je možno použít tabulku 35.

Tab. 35 *Přibližná interpretace hodnot korelačního koeficientu*

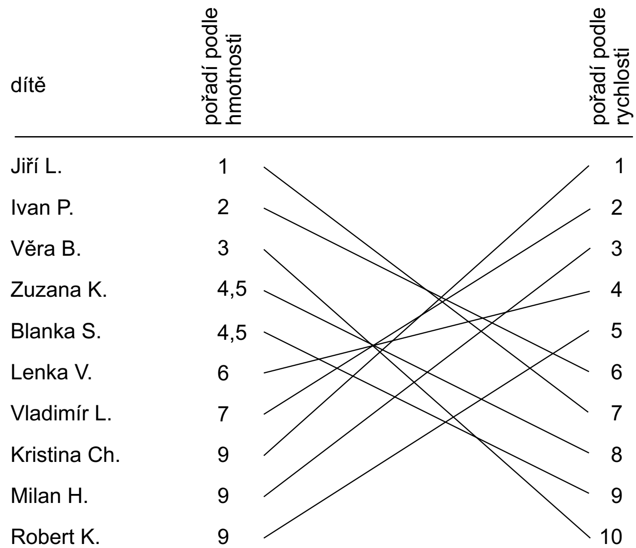
Koeficient korelace	Interpretace
$r = 1$	naprostá závislost (funkční závislost)
$1,00 > r \geq 0,90$	velmi vysoká závislost
$0,90 > r \geq 0,70$	vysoká závislost
$0,70 > r \geq 0,40$	střední (značná) závislost
$0,40 > r \geq 0,20$	nízká závislost
$0,20 > r \geq 0,00$	velmi slabá závislost
$r = 0$	naprostá nezávislost

Záporné hodnoty koeficientu korelace vyjadřují negativní vztah mezi proměnnými a je možné je interpretovat obdobně jako výše uvedené kladné hodnoty.

Ve výzkumech se většinou pracuje s koeficienty korelace, jejichž absolutní hodnota je minimálně 0,40. Uvedená tabulka slouží jen k orientačnímu posouzení vypočítané hodnoty koeficientu korelace. Určitá hodnota korelačního koeficientu může být hodnocena v různých výzkumných situacích různě.

Těsnost vztahu mezi oběma proměnnými je možno v daném případě hodnotit jako „střední" (bližší poučení o interpretaci vypočítaného koeficientu korelace bude uvedeno v oddílu 3.4.3).

Vztah mezi hmotností dětí a jejich rychlostí v běhu můžeme zobrazit graficky metodou „křížem krážem" (obr. 15).



***Obr. 15** Srovnání pořadí dětí podle hmotnosti a rychlosti v běhu*

Příklad 35

Pomocí Spearmanova koeficientu pořadové korelace můžeme například posoudit, jak těsný je vztah mezi klasifikací žáků v určitém vyučovacím předmětu a mezi jejich skutečnými vědomostmi. Toto posuzování můžeme opírat například o klasifikaci žáků na vysvědčení a o výsledky žáků ve spolehlivém a dostatečně validním didaktickém testu.

Tabulka 36 uvádí výsledky jedné školní třídy v matematice. U každého žáka je uvedeno, jakou klasifikaci z matematiky měl na vysvědčení a počet bodů, který dosáhl v didaktickém testu. (Předpokládáme, že didaktický test spolehlivě zkoušel právě ty vědomosti a dovednosti z matematiky, které mají žáci mít osvojené.)

Dosadíme-li příslušná data do vzorce pro výpočet Spearmanova koeficientu pořadové korelace, dostáváme

$$r_s = 1 - \frac{6 \cdot 620{,}5}{26 \cdot (26^2 - 1)} = 0{,}79$$

Vypočítaná hodnota vypovídá o poměrně těsném vztahu (vysoké závislosti) mezi klasifikací žáků z matematiky a jejich skutečnými vědomostmi z matematiky.

Tab. 36 *Spearmanův koeficient pro klasifikaci a výsledky žáků v didaktickém testu*

Žák č.	**Test T**	**Klasifikace K**	**Pořadí T**	**Pořadí K**	**Rozdíl pořadí *d***	d^2
1	28	1	1	1,5	–0,5	0,25
2	27	1	2	1,5	0,5	0,25
3	26	2	3	5,5	–2,5	6,25
4	25	2	4	5,5	–1,5	2,25
5	24	3	5	13,5	–8,5	72,25
6	22	2	6	5,5	0,5	0,25
7	21	2	7	5,5	1,5	2,25
8	20	3	8,5	13,5	–5	25,00
9	20	3	8,5	13,5	–5	25,00
10	19	3	10,5	13,5	–3	9,00
11	19	2	10,5	5,5	5	25,00
12	18	2	12	5,5	6,5	42,25
13	17	4	13,5	22,5	–9	81,00
14	17	3	13,5	13,5	0	0,00
15	16	3	15	13,5	1,5	2,25
16	15	4	16	22,5	–6,5	42,25
17	13	3	18	13,5	4,5	20,25
18	13	4	18	22,5	–4,5	20,25
19	13	4	18	22,5	–4,5	20,25
20	12	3	20,5	13,5	7	49,00
21	12	3	20,5	13,5	7	48,00
22	11	4	22	22,5	–0,5	0,25
23	9	3	24	13,5	10,5	110,25
24	9	4	24	22,5	1,5	2,25
25	9	4	24	22,5	1,5	2,25
26	8	4	26	22,5	3,5	12,25
						Σ 620,50

Spearmanův koeficient pořadové korelace se doporučuje používat pouze tehdy, jestliže počet korelovaných dvojic hodnot není příliš vysoký (maximálně kolem třiceti) a jestliže více než čtyři srovnávané prvky nemají stejné pořadí. Tyto podmínky nebyly v uvedeném příkladě zcela respektovány (stejné pořadí mají více než čtyři žáci).

Možnosti analýzy na PC
SPSS: *correlate* → *bivariate correlations* → Spearman
Statistica.cz: neparametrické statistiky → korelace (Spearman)

3.3.6.2 Kendallův koeficient shody

U Spearmanova koeficientu pořadové korelace se jednalo o určování těsnosti vztahu mezi dvěma pořadími, která byla vytvořena na základě dvou kritérií. Někdy však potřebujeme posoudit, jak těsný je vztah mezi více než dvěma pořadími. V těchto případech lze jako míry těsnosti vztahu použít Kendallova koeficientu shody.

Příklad 36

Výpočet Kendallova koeficientu shody vyložíme na příkladu šetření, ve kterém se srovnávala obliba vyučovacích předmětů u jednotlivých žáků.

Žáci 7. ročníku základní školy se měli vyjádřit k oblibě jednotlivých vyučovacích předmětů tím způsobem, že číslo 1 měli přiřadit předmětu nejoblíbenějšímu, číslo 2 předmětu, který stojí v oblibě na druhém místě, atd. Celkem se žáci vyjadřovali k deseti vyučovacím předmětům. Výsledky hodnocení tří žáků (A, B, C) uvádí tabulka 37.

Tab. 37 *Výpočet Kendallova koeficientu shody*

Vyučovací předmět	Hodnocení žáků			Součet pořadí	X^2
	A	B	C	X	
tělesná výchova	1	2	3	6	36
přírodopis	2	1	1	4	16
výtvarná výchova	3	3	2	8	64
rodinná výchova	4	4	7	15	225
hudební výchova	5	6	5	16	256
dějepis	6	5	9	20	400
zeměpis	7	8	10	25	625
fyzika	8	10	6	24	576
matematika	9	7	4	20	400
český jazyk	10	9	8	27	729
				Σ 165	Σ 3 327

Těsnost vztahu mezi hodnocením obliby vyučovacích předmětů u vybraných žáků posoudíme na základě výpočtu Kendallova koeficientu shody, který lze vypočítat podle vzorce

$$W = \frac{\sum X^2 - \frac{(\sum X)^2}{n}}{\frac{1}{12}k^2 \cdot (n^3 - n)} \tag{55}$$

kde *W* je Kendallův koeficient shody, *X* je součet pořadí, která byla přisouzena jednotlivým posuzovaným objektům (vyučovacím předmětům) všemi posuzovateli, *n* je počet posuzovaných objektů (vyučovacích předmětů) a *k* je počet srovnávaných pořadí (počet hodnotících žáků).

Hodnoty potřebné pro výpočet Kendallova koeficientu shody zachycuje tabulka 34. Pro hodnoty z tabulky vychází

$$W = \frac{3\,327 - \dfrac{165^2}{10}}{\dfrac{1}{12} \cdot 3^2 (10^3 - 10)} = 0,814$$

Kendallův koeficient shody může nabývat hodnot v intervalu od 0 do +1. Čím vyšší je jeho hodnota, tím těsnější je vztah mezi srovnávanými pořadími (tím větší shoda je mezi pořadími).

K posouzení statistické významnosti Kendallova koeficientu shody lze použít testové kritérium chí-kvadrát (Mittenecker, 1968), které se v tomto případě vypočítává ze vztahu

$$\chi^2 = W \cdot k\,(n-1) \tag{56}$$

kde *W* je vypočítaný Kendallův koeficient shody, *k* je počet srovnávaných pořadí a *n* je počet posuzovaných objektů (počet posuzovaných vyučovacích předmětů).

Vypočítaná hodnota testového kritéria χ^2 se srovnává s kritickou hodnotou pro $f = n - 1$ stupňů volnosti (*n* je počet posuzovaných objektů).

Při testování statistické významnosti vypočítaného koeficientu *W* se nejdříve formuluje nulová a alternativní hypotéza:

H_0: Vypočítaný Kendallův koeficient nevypovídá o shodě mezi hodnocením jednotlivých žáků (W = 0).
H_A: Vypočítaný koeficient vypovídá o shodě mezi srovnávanými pořadími (W ≠ 0).

Testování provedeme na hladině významnosti 0,05.

V uvedeném příkladě vychází $\chi^2 = 0,814 \cdot 3\,(10-1) = 21,978$. Tuto vypočítanou hodnotu srovnáváme s kritickou hodnotou $\chi^2_{0,05}(9) = 16,919$. Protože je vypočítaná hodnota ($\chi^2 = 21,978$) větší než hodnota kritická, odmítáme nulovou hypotézu a přijímáme hypotézu alternativní. Mezi pořadími obliby vyučovacích předmětů, které vybraní žáci vytvořili, je statisticky významná shoda.

Možnosti analýzy na PC
Statistica.cz: neparametrické statistiky → korelace (Kendall)
SPSS: *correlate* → *bivariate correlations*→ Kendall

3.4 STATISTICKÉ METODY PRO ANALÝZU METRICKÝCH DAT

Při analýze intervalových nebo poměrových (metrických) dat je možno využívat všech postupů, které byly uvedeny v případě dat nominálních nebo ordinálních.

Data „vyššího“ typu lze totiž vždy (po provedení vhodné kategorizace) převést na data „nižšího“ typu. Takový postup je však zpravidla doprovázen určitou ztrátou informace. V dalším textu uvedeme statistické testy významnosti, které se nejčastěji používají při analýze metrických dat.

3.4.1 FUNKČNÍ A STATISTICKÁ ZÁVISLOST MEZI JEVY

Jestliže chceme ověřovat závislost mezi dvěma veličinami, potom na každém objektu (prvku) statistického souboru provádíme vždy dvě různá měření (pozorování). Data získaná tímto způsobem tvoří tzv. **dvojrozměrný statistický soubor**. U dvojrozměrného statistického souboru tedy máme u každého prvku (jedince, situace) vždy dva údaje o jeho vlastnostech (např. u každého žáka máme změřenu jeho výšku a hmotnost).

Jestliže platí, že určité hodnotě jedné proměnné odpovídá jen jedna určitá hodnota druhé proměnné, potom hovoříme o **funkční závislosti**. Mezi proměnnými je tedy funkční závislost tehdy, jestliže každé hodnotě jedné proměnné odpovídá právě jedna hodnota y druhé proměnné. Funkční závislost lze obecně vyjádřit zápisem

$$y = f(x)$$

kde x je hodnota nezávisle proměnné a y je hodnota závisle proměnné. S funkčními závislostmi se v sociální, psychologické nebo pedagogické oblasti takřka nesetkáváme. Pomocí funkční závislosti jsou popisovány zejména přírodní zákony. Tak např. volný pád tělesa se řídí zákonitostí, která je dána funkčním vztahem

$$s = \frac{1}{2} g \cdot t^2$$

kde s je dráha volného pádu, g konstanta (tzv. tíhové zrychlení) a t je čas volného pádu tělesa. V tomto případě je tedy dráha volného pádu s (závisle proměnná) funkcí času t (nezávisle proměnná). Platí, že určitému času odpovídá pouze jedna určitá dráha s.

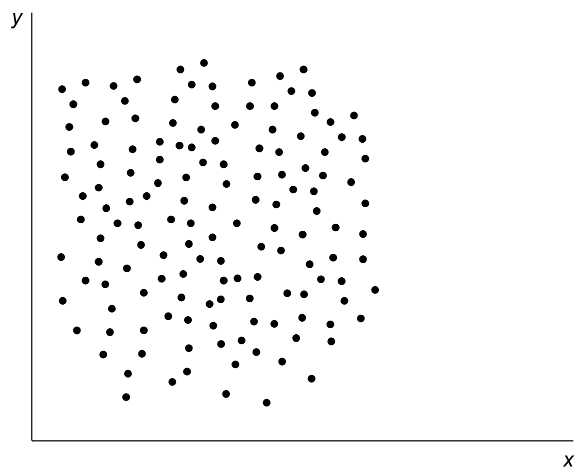
Jestliže jedné hodnotě dané veličiny (proměnné) neodpovídá jen jedna hodnota druhé veličiny (proměnné), nýbrž celý obor hodnot druhé veličiny, potom hovoříme o **statistické závislosti (stochastické závislosti)**.

Příkladem statistické závislosti je například závislost mezi výškou a věkem žáků základní školy. Na základě zkušeností lze říci, že výška žáků se s přibývajícím věkem zvětšuje, ale nelze například tvrdit, že určitému věku odpovídá jen jedna určitá výška žáka.

Při studiu závislostí mezi dvěma náhodnými veličinami označujeme jednu z nich jako nezávisle proměnnou a druhou jako závisle proměnnou. Která z náhodných veličin je závisle a která nezávisle proměnná, je otázka z hlediska statistické analýzy irelevantní, může mít však podstatný význam pro vědní obor, který statistiku využívá.

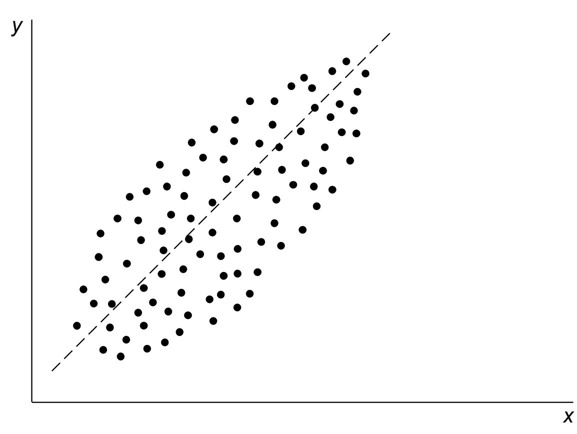
Pro přiblížení povahy statistických závislostí je možné využít **bodových diagramů**. U bodových diagramů se naměřené (pozorované) hodnoty dvou proměnných na týchž objektech znázorňují jako body v pravoúhlé souřadnicové soustavě. Na vodorovnou osu se nanášejí hodnoty proměnné a na svislou osu odpovídající hodnoty proměnné y. Obrazem každé dvojice hodnot je potom jeden bod v diagramu. Jednotlivé body se mohou v bodových diagramech při studiu statistické závislosti seskupovat několika způsoby:

- Jednotlivé body (obrazy dvojic naměřených hodnot) jsou zcela neuspořádané a vyplňují v podstatě celou plochu diagramu. V tomto případě jednotlivým hodnotám nezávisle proměnné *x* odpovídají libovolné hodnoty závisle proměnné *y*. Jde o případ, kdy obě náhodné veličiny jsou **statisticky nezávislé** (obr. 16).



***Obr. 16** Statistická nezávislost dvou náhodných veličin*

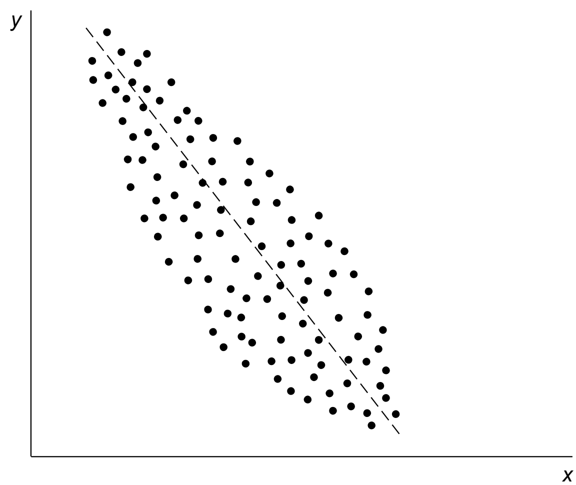
- Body vyplňují přibližně plochu elipsy. Rozmístění bodů v diagramu lze v tomto případě přibližně vystihnout pomocí přímky (hlavní osa elipsy), a proto hovoříme o **lineární statistické závislosti**. Lineární statistická závislost může být dvojího druhu. Jestliže vyšším hodnotám jedné proměnné odpovídají spíše vyšší hodnoty druhé proměnné, potom hovoříme o **přímé závislosti** (obr. 17).



***Obr. 17** Lineární statistická (přímá) závislost*

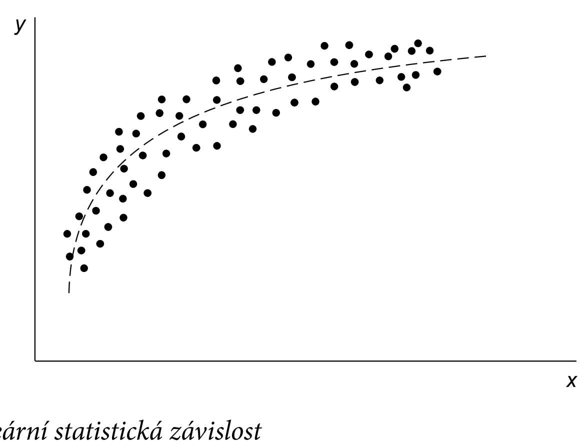
- Jestliže vyšším hodnotám jedné proměnné odpovídají spíše nižší hodnoty druhé proměnné, potom jde o **nepřímou závislost** (obr. 18).

Elipsa, jejíž plochu jednotlivé body vyplňují, určuje charakter statistické závislosti mezi danými proměnnými. Hlavní osa elipsy udává nejpravděpodobnější hodnoty jedné proměnné v závislosti na hodnotách druhé proměnné. Délka vedlejší poloosy elipsy je ukazatelem těsnosti vztahu mezi oběma proměnnými. Jestliže se u elipsy délka vedlejší poloosy zvětšuje, je těsnost vztahu mezi proměnnými stále menší a menší, až nakonec přechází ve statistickou nezávislost obou proměnných (v případě, že délka vedlejší poloosy se rovná délce hlavní poloosy elipsy, elipsa přechází v kružnici). Jestliže se délka vedlejší poloosy elipsy zmenšuje, potom se naopak vztah mezi oběma proměnnými stává těsnější. Mezním případem je situace, kdy délka vedlejší poloosy je rovna nule. V tomto případě přechází elipsa v úsečku a závislost statistická se mění v závislost funkční. Funkční závislost a nezávislost lze tedy považovat za krajní případy statistické závislosti.



Obr. 18 *Lineární statistická (nepřímá) závislost*

- Body se kupí tak, že jejich seskupení lze vystihnout určitou obecnou křivkou. V tomto případě hovoříme o nelineární statistické závislosti (obr. 19).



Obr. 19 *Nelineární statistická závislost*

Čáry, které vystihují průběh závislosti mezi dvěma proměnnými, se nazývají **regresní čáry**. Podle povahy závislosti mluvíme buď o regresních přímkách (obr. 17 a 18) nebo o regresních křivkách (obr. 19).

3.4.2 REGRESNÍ A KORELAČNÍ ANALÝZA

Analýza závislosti mezi dvěma proměnnými má dva základní aspekty. Prvním z nich je nalezení příslušné regresní čáry (nalezení regresní funkce), druhým aspektem je posouzení těsnosti vztahu mezi danými proměnnými.

Úkolem **regresní analýzy** je nalézt regresní funkci, pomocí níž lze ze známých hodnot nezávisle proměnné určit příslušné hodnoty závisle proměnné. Nejjednodušší situace nastává v případě, že mezi proměnnými je lineární statistická závislost. Příkladem takové závislosti je například vztah mezi věkem učitelů na středních školách (x) a jejich průměrnou mzdou (y). V tomto případě je **mzda** učitelů funkcí **věku** učitelů. Čím vyšší je věk učitelů, tím vyšší je jejich průměrná mzda. Platí, že

$$y = f(x) = a + bx \tag{57}$$

kde výrazy a, b jsou konstanty, jejichž skutečnou velikost neznáme, ale můžeme ji odhadnout ze zjištěných empirických dat. Konstanta b se nazývá **koeficient regrese** a udává, o kolik se změní (průměrně) hodnota závisle proměnné y, jestliže hodnota nezávisle proměnné x se zvětší o jednotku. Známe-li obě konstanty (a, b), můžeme pro kteroukoli hodnotu nezávisle proměnné x (např. pro věk učitelů) vypočítat hodnotu závisle proměnné y (např. průměrnou mzdu učitelů). Podrobné vysvětlení regresní analýzy lze nalézt například v práci Králíka a Hartmana (2000).

V pedagogických výzkumech je regresní analýza spíše výjimkou. Daleko častější je posuzování těsnosti vztahu mezi proměnnými, tj. **korelační analýza**. Nejjednodušší případ korelační analýzy nastává v případě, že mezi proměnnými se projevuje lineární statistická závislost (přímá nebo nepřímá). Těsnost vztahu mezi proměnnými je možné vyjádřit například pomocí **koeficientu korelace**.

3.4.3 PEARSONŮV KOEFICIENT KORELACE

Variabilitu hodnot v jednorozměrném statistickém souboru lze charakterizovat pomocí rozptylu neboli variance (srov. oddíl 2.3.4)

$$variance = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \frac{1}{2}\sum_{i=1}^{n}(x_i - \overline{x})\cdot(x_i - \overline{x}) \tag{58}$$

kde n je četnost všech hodnot ve statistickém souboru, x_i jsou jednotlivé hodnoty a $\overline{x}$ je aritmetický průměr všech hodnot.

Variabilitu hodnot v dvojrozměrném statistickém souboru (tj. v případě, že u každého objektu máme dvě naměřené hodnoty) lze charakterizovat pomocí tzv. **kovariance**. Výpočet kovariance je analogický s výpočtem variance

$$kovariance = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})\cdot(y_i - \overline{y}) \tag{59}$$

kde n je četnost dvojic hodnot, x_i, y_i jsou hodnoty obou proměnných naměřené na jednom objektu (dvojice hodnot), $\overline{x}$ je průměrná hodnota jedné proměnné a $\overline{y}$ je průměrná hodnota druhé proměnné.

Pearsonův koeficient korelace r_p je možno definovat jako poměr kovariance a součinu směrodatných odchylek obou proměnných, tj.

$$r_p = \frac{kovariance}{s_x \cdot s_y} \tag{60}$$

kde s_x je směrodatná odchylka jedné proměnné a s_y je směrodatná odchylka druhé proměnné.

Jestliže do vztahu (60) dosadíme za kovarianci výraz ze vztahu (59), dostaneme po úpravě vzorec vhodný pro výpočet **Pearsonova koeficientu korelace** (ve vzorci je pro jednoduchost vynecháno indexování hodnot x a y)

$$r_p = \frac{n\cdot\sum xy - \sum x \cdot \sum y}{\sqrt{\left[n\sum x^2 - \left(\sum x\right)^2\right]\cdot\left[n\sum y^2 - \left(\sum y\right)^2\right]}} \tag{61}$$

kde x, y jsou jednotlivé dvojice hodnot obou proměnných a n počet srovnávaných dvojic hodnot.

Příklad 37

Skupině 10 dětí předškolního věku byly postupně zadány dva úkoly: sestavování obrázkové mozaiky a hledání cesty labyrintem. Byl přitom měřen čas, který jednotlivé děti ke splnění úkolů potřebovaly. Získané výsledky uvádí tabulka 38.

Máme určit, jak těsný je vztah mezi časy potřebnými ke splnění obou úkolů.

Tabulka obsahuje vedle obou naměřených hodnot x, y také hodnoty potřebné k výpočtu koeficientu korelace xy, x^2, y^2). Dosadíme-li příslušné hodnoty do výše uvedeného vzorce, dostáváme

$$r_p = \frac{10\cdot 687305 - 2975\cdot 2220}{\sqrt{\left[10\cdot 912053 - 2975^2\right]\cdot\left[10\cdot 540330 - 2220^2\right]}} = 0{,}750$$

Vypočítaný koeficient $r_p = 0{,}750$ vypovídá o tom, že mezi výsledky dětí, dosahovanými v obou úkolech, je poměrně značný pozitivní vztah. Čím delší je čas, který děti potřebovaly k projití labyrintem, tím delší je také čas, který potřebovaly ke složení obrázkové mozaiky.

***Tab. 38** Výpočet Pearsonova koeficientu korelace*

Dítě č.	x labyrint (s)	y mozaika (s)	xy	x^2	y^2
1	220	164	36 080	48 400	26 896
2	318	246	78 228	101 124	60 516
3	360	306	110 160	129 600	93 636
4	310	248	76 880	96 100	61 504
5	280	114	31 920	78 400	12 996
6	248	102	25 296	61 504	10 404
7	340	295	100 300	115 600	87 025
8	293	253	74 129	85 849	64 009
9	226	212	47 912	51 076	44 944
10	380	280	106 400	144 400	78 400
	Σ 2 975	Σ 2 220	Σ 687 305	Σ 912 053	Σ 540 330

Pearsonův koeficient korelace může nabývat hodnot z intervalu od –1 do +1. Hodnota 0 vypovídá o statistické nezávislosti obou proměnných, hodnota +1 (respektive –1) vypovídá o naprosté (funkční) závislosti proměnných. Čím více se vypočítaná hodnota koeficientu korelace blíží hodnotě 1 (nebo –1), tím těsnější je vztah mezi proměnnými (jevy), které srovnáváme.

Kladný výsledek vypovídá, že vyšším hodnotám jedné proměnné odpovídají také spíše vyšší hodnoty druhé proměnné a zároveň nižším hodnotám první proměnné odpovídají také nižší hodnoty druhé proměnné.

Jestliže je koeficient korelace záporný, znamená to, že mezi proměnnými, které srovnáváme, je negativní (opačný) vztah. V tomto případě vysokým hodnotám jedné proměnné odpovídají spíše nižší hodnoty druhé proměnné a naopak.

Pro přibližnou interpretaci vypočítaného koeficientu korelace je možno použít tabulku 35 (byla uvedena v souvislosti se Spearmanovým koeficientem pořadové korelace). Vypočítanou hodnotu $r_p = 0{,}750$ bychom mohli tímto způsobem interpretovat jako poměrně vysokou.

Vypočítanou hodnotu koeficientu korelace je nutno posuzovat vždy vzhledem k situaci, pro kterou byla vypočítána. Určitá hodnota korelačního koeficientu může být totiž hodnocena v různých výzkumných situacích různě.

V souvislosti s interpretací vypočítané hodnoty korelačního koeficientu je třeba důrazně upozornit na skutečnost, že pouhá existence vysoké korelace mezi dvěma jevy ještě nemusí nutně znamenat existenci skutečného a smysluplného vztahu mezi těmito jevy. V některých případech může být příčinou vysoké korelace působení nějaké jiné, nekontrolované proměnné; někdy může jít i o tzv. **nonsense correlation** (nesmyslnou korelaci). Skutečné vztahy, příčiny a účinky je třeba určit vždy na základě logické analýzy, ve které zvážíme všechny okolnosti zkoumaného vztahu.

Důležitou informaci pro hodnocení vypočítaného koeficientu korelace získáme na základě testování jeho statistické významnosti. Jde o to rozhodnout, zda vypočítaná hodnota korelačního koeficientu je natolik vysoká, abychom mohli hovořit o statisticky významném vztahu. K ověřování statistické významnosti korelačního koeficientu se nejčastěji používá testového kritéria t.

Příklad 38

Máme rozhodnout, zda koeficient korelace $r_p = 0{,}750$ (vypočítaný v předchozím příkladě) vypovídá o statisticky významném vztahu mezi oběma proměnnými (čas potřebný na sestavení obrázkové mozaiky a čas potřebný k projití labyrintem u dětí předškolního věku).

Postup při testování statistické významnosti koeficientu korelace je stejný jako u ostatních testů významnosti. Nejdříve formulujeme nulovou a alternativní hypotézu:

H_0: Vypočítaná hodnota koeficientu korelace nevypovídá o závislosti mezi oběma proměnnými (tvrdíme tedy vlastně, že obě proměnné jsou nezávislé a že $r_p = 0$).
H_A: Vypočítaná hodnota koeficientu korelace vypovídá o vztahu mezi oběma proměnnými ($r_p \neq 0$).

Testování statistické významnosti provedeme na hladině významnosti $\alpha = 0{,}05$.

O přijetí (respektive odmítnutí) nulové hypotézy rozhodneme na základě výpočtu testového kritéria t, které lze v tomto případě vypočítat ze vztahu

$$t = \frac{r_p}{\sqrt{1 - r_p^{\,2}}} \cdot \sqrt{n-2} \tag{62}$$

kde r_p je koeficient korelace a n je počet srovnávaných dvojic hodnot.

V našem případě vychází

$$t = \frac{0{,}750}{\sqrt{1 - 0{,}750^2}} \cdot \sqrt{10-2} = 3{,}207$$

Vypočítanou hodnotu testového kritéria srovnáváme s kritickou hodnotou tohoto kritéria pro zvolenou hladinu významnosti a počet stupňů volnosti, který se v tomto případě určí ze vztahu

$$f = n - 2 \tag{63}$$

kde f je počet stupňů volnosti a n je počet srovnávaných dvojic hodnot. Pro hladinu významnosti $\alpha = 0{,}05$ a $f = 10 - 2 = 8$ stupňů volnosti lze ve statistických tabulkách (příloha X) nalézt kritickou hodnotu $t_{0{,}05}(8) = 2{,}306$. Protože vypočítaná hodnota $t = 3{,}207$ je větší než hodnota kritická, odmítáme nulovou hypotézu a přijímáme hypotézu alternativní. Vypočítaný koeficient korelace tedy vypovídá o statisticky významné závislosti mezi výsledky obou provedených měření.

Uvedený postup testování statistické významnosti koeficientu korelace je použitelný tehdy, jestliže počet srovnávaných dvojic hodnot $n > 8$. Při testování významnosti koeficientu korelace, který byl vypočítán z menšího počtu dvojic hodnot, lze použít tabulek kritických hodnot koeficientu korelace pro malé n (příloha VII). Za statisticky významné lze považovat ty koeficienty, které jsou (v absolutní hodnotě) větší nebo rovny tabelované hodnotě.

Pearsonův koeficient korelace jsme oprávněni počítat pouze v případě splnění následujících podmínek:

- data, z nichž Pearsonův koeficient vypočítáváme, jsou metrická (intervalová nebo poměrová);
- regresní čára musí být přímka;
- základní soubor (z něhož pocházejí data, pro něž počítáme Pearsonův korelační koeficient) má tzv. **dvojrozměrné normální rozdělení** (u dvojrozměrného normálního rozdělení platí, že každé hodnotě první proměnné odpovídá normální rozdělení druhé proměnné a naopak).

Splnění uvedených podmínek by se mělo před každým použitím Pearsonova koeficientu alespoň přibližně kontrolovat.

Pro lepší představu o výpovědní schopnosti daného koeficientu korelace bývá výhodné doplnit statistickou analýzu ještě o výpočet tzv. **koeficientu determinace**. Koeficient determinace je druhá mocnina koeficientu korelace, tj. r_p^2. V našem případě tedy $r_p^2 = 0,750^2 = 0,563$. Tento výsledek lze interpretovat tak, že přibližně z 56 % je výkon dětí při procházení labyrintem ovlivňován výkonem ve skládání obrázkové mozaiky. Zbytek, tj. asi 44 %, připadá na jiné, nezjištěné činitele. Vzhledem k tomu, že u korelační analýzy můžeme považovat za závisle proměnnou jak výkon při průchodu labyrintem, tak výkon při skládání obrazové mozaiky, lze vztah interpretovat i opačným způsobem.

Možnosti analýzy na PC
Excel: vložit funkci → statistické → Pearson
Statistica.cz: základní statistiky → korelační matice
SPSS: *correlate* → *bivariate* → Pearson

3.4.4 BODOVÁ BISERIÁLNÍ KORELACE r_{bb}

Jestliže máme určit těsnost vztahu mezi dvěma proměnnými, z nichž jedna je zachycena na úrovni intervalového nebo poměrového měření a druhá na úrovni nominálního měření s alternativními znaky (nabývajícími dvou hodnot), potom je možno k výpočtu použít tzv. bodové biseriální korelace. **Koeficient bodové biseriální korelace** r_{bb} lze vypočítat podle vzorce

$$r_{bb} = \frac{\overline{x}_p - \overline{x}_q}{s} \cdot \sqrt{p \cdot q} \qquad (64)$$

kde $\overline{x}_p$ je průměr ve skupině hodnot s prvním alternativním znakem, $\overline{x}_q$ je průměr hodnot ve skupině s druhým alternativním znakem, s je směrodatná odchylka vypočítaná z hodnot obou skupin, p je relativní četnost hodnot ve skupině s prvním alternativním znakem a q je relativní četnost hodnot ve skupině s druhým alternativním znakem. Platí tedy, že $q = 1 - p$.

Příklad 39

Výpočet koeficientu bodové biseriální korelace budeme ilustrovat na příkladu výzkumu, ve kterém se ověřovalo, jak dalece ovlivňuje stupeň vzdělání rodičů úroveň všeobecných znalostí u žáků základní školy.

Náhodnému výběru 60 žáků základní školy byl zadán test všeobecných znalostí, ve kterém bylo možno získat maximálně 50 bodů. Ze všech dosažených výsledků byl vypočítán průměr $\bar{x}$ = 30 bodů a směrodatná odchylka $s = 9$.

Bylo zjištěno, že rodiče 27 žáků z daného výběru mají středoškolské nebo vysokoškolské vzdělání a že rodiče zbývajících 33 žáků mají základní vzdělání. Žáci, jejichž rodiče mají středoškolské nebo vysokoškolské vzdělání, dosáhli v testu průměrného výsledku 36 bodů, zatímco žáci, jejichž rodiče mají základní vzdělání, dosáhli průměru 25 bodů.

Máme rozhodnout, jak těsný je vztah mezi stupněm vzdělání rodičů a úrovní všeobecných znalostí žáků základní školy.

Relativní četnost žáků, jejichž rodiče mají středoškolské nebo vysokoškolské vzdělání, je

$$p = \frac{27}{60} = 0{,}45$$

Relativní četnost žáků, jejichž rodiče mají základní vzdělání, je

$$q = \frac{33}{60} = 1 - 0{,}45 = 0{,}55$$

Jestliže dosadíme příslušné hodnoty do vzorce (64), dostáváme

$$r_{bb} = \frac{36 - 25}{9} \cdot \sqrt{0{,}45 \cdot 0{,}55} = 0{,}608$$

Poznámka: *Výpočet bodové biseriální korelace si lze usnadnit pomocí tabulky (příloha VIII), kde jsou tabelovány hodnoty* $\sqrt{p \cdot q}$

Těsnost vztahu mezi úrovní vzdělání rodičů a úrovní všeobecných znalostí jejich dětí je dána koeficientem $r_{bb} = 0{,}608$.

3.4.5 BISERIÁLNÍ KORELACE r_{bis}

V případech, kdy můžeme předpokládat, že i alternativní znak má „ve skutečnosti" normální rozdělení (vznikl na základě kategorizace metrických dat, která měla normální rozdělení), je výhodnější počítat **koeficient biseriální korelace** r_{bis}

$$r_{bis} = \frac{\bar{x}_p - \bar{x}_q}{s} \cdot \frac{p \cdot q}{y} \qquad (65)$$

kde význam písmen je stejný jako ve vzorci (64) a výraz $\frac{p \cdot q}{y}$ je tabelován ve statistických tabulkách (příloha VIII).

Příklad 40

Při statistickém ověřování vlastností úloh didaktického testu byla určitá testová úloha zadána vzorku 200 žáků. Určitou úlohu v testu řešilo správně 136 žáků a nesprávně 64 žáků. Žáci, kteří úlohu řešili správně, dosáhli v celém testu průměru 66,2 bodu, a žáci, kteří úlohu řešili nesprávně, dosáhli v testu průměru 54,0 bodu.

Máme posoudit, jak těsný je vztah mezi úspěšností žáků v dané úloze a celkovými výsledky v testu.

V uvedeném případě máme posuzovat těsnost vztahu mezi dvěma proměnnými, z nichž jedna je výkon žáků v didaktickém testu a druhá příslušnost žáků ke skupině (podle toho, zda danou úlohu řešili, nebo neřešili správně). Můžeme předpokládat, že rozdělení žáků na ty, kteří řešili správně, a ty, kteří úlohu neřešili správně, bylo vytvořeno na základě jejich vědomostí, které mají normální rozdělení.

Relativní četnost žáků, kteří byli při řešení testové úlohy úspěšní je

$$p = \frac{136}{200} = 0{,}68$$

Relativní četnost žáků, kteří byli při řešení testové úlohy neúspěšní, je

$$q = \frac{64}{200} = 1 - 0{,}68 = 0{,}32$$

Ve statistických tabulkách (příloha VIII) vyhledáme pomocnou hodnotu pro vypočítanou hodnotu relativní četnosti p. Zjišťujeme, že pro $p = 0{,}68$ je tabelována hodnota $\frac{p \cdot q}{y} = 0{,}609$.

Z tabulky v příloze VIII je zřejmé, že rozlišování koeficientů r_{bb} a r_{bis} má smysl zejména tehdy, když relativní četnost p je velmi malá (respektive velmi vysoká), protože v těchto případech jsou mezi hodnotami $\sqrt{p \cdot q}$ a $\frac{p \cdot q}{y}$ velké rozdíly. Dosadíme-li příslušné hodnoty do vzorce (65), dostáváme

$$r_{bis} = \frac{66{,}2 - 54{,}0}{12{,}6} \cdot 0{,}609 = 0{,}59$$

Vypočítaný koeficient $r_{bis} = 0{,}59$ vyjadřuje těsnost vztahu mezi úspěšností v dané testové úloze a celkovými výsledky v didaktickém testu. Zjištěná těsnost vztahu je značná, což vypovídá o tom, že daná testová úloha dobře rozlišuje mezi „lepšími" a „horšími" žáky.

Koeficientu biseriální korelace r_{bis} se často užívá při posuzování tzv. **citlivosti (diskriminační hodnoty)** úloh v didaktických testech.

3.4.6 TETRACHORICKÝ KOEFICIENT KORELACE

Při posuzování těsnosti vztahu mezi dvěma proměnnými, které jsou vyjádřeny jako alternativní znaky (můžeme však u obou proměnných předpokládat, že mají „ve skutečnosti" normální rozdělení), můžeme použít tzv. tetrachorického koeficientu korelace (Mráz, 1977). **Tetrachorický koeficient korelace** lze vypočítat podle vztahu

$$r_{tet} = \cos\left(180 \cdot \frac{\sqrt{b \cdot c}}{\sqrt{b \cdot c} + \sqrt{a \cdot d}}\right) \tag{66}$$

kde cos je goniometrická funkce kosinus, čísla a, b, c, d mají význam, který vyplývá ze schématu čtyřpolní tabulky (srov. oddíl 3.2.3).

Také tento koeficient se často používá k posuzování citlivosti úloh v didaktických testech (Chráska, 1999).

3.4.7 TESTOVÁNÍ VÝZNAMNOSTI ROZDÍLU MEZI DVĚMA KOEFICIENTY KORELACE

Ve výzkumech někdy řešíme problém, zda dva vypočítané koeficienty korelace jsou přibližně stejné, anebo zda mezi nimi je významný rozdíl. Abychom mohli dva koeficienty korelace srovnávat, musí být pochopitelně vypočítány pro stejné dvojice proměnných.

Testování významnosti rozdílu mezi dvěma koeficienty korelace vyložíme na příkladu 41.

Příklad 41

U studentů gymnázia byla počítána korelace mezi klasifikací z matematiky a klasifikací z fyziky. U skupiny dívek (n_1 = 64) byl vypočítán koeficient korelace 0,60, u skupiny chlapců (n_2 = 82) koeficient korelace 0,78.

Hledáme odpověď na otázku, zda u chlapců je vztah mezi klasifikací v matematice a klasifikací ve fyzice těsnější než u dívek:

H_0: Vypočítané koeficienty korelace jsou stejné
H_A: Mezi oběma koeficienty korelace je rozdíl.

Zvolená hladina významnosti α = 0,05.

Při testování statistické významnosti rozdílů mezi korelačními koeficienty se nejdříve oba koeficienty korelace transformují (převádějí) na hodnoty **normované normální veličiny** *z* (Fisherova z-transformace) podle vztahu

$$z = \frac{1}{2} \ln \frac{1+r}{1-r} \tag{67}$$

kde r je příslušný koeficient korelace a ln je přirozený logaritmus.

Uvedená transformace se v praxi provádí pomocí tabulek (příloha IX). Pro koeficient korelace r_1 = 0,60 nalézáme v tabulce hodnotu z_1 = 0,6931 a pro koeficient korelace r_2 = 0,78 hodnotu z_2 = 1,0454.

Testovým kritériem pro testování statistické významnosti rozdílů mezi dvěma koeficienty je **normovaná normální veličina** u, kterou lze vypočítat ze vztahu

$$u = \frac{|z_1 - z_2|}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}} \tag{68}$$

kde z_1 a z_2 jsou transformované hodnoty odpovídající oběma koeficientům korelace, n_1 je počet dvojic hodnot (počet žáků), pro něž byl vypočítán koeficient korelace v první skupině, a n_2 je počet dvojic hodnot (počet žáků) ve druhé skupině.

Dosadíme-li do uvedeného vzorce, dostáváme

$$u = \frac{|0,6931 - 1,0454|}{\sqrt{\frac{1}{64 - 3} + \frac{1}{82 - 3}}} = 2,07$$

Vypočítanou hodnotu srovnáváme s kritickou hodnotou normované normální veličiny $u_{0,05} = 1,96$ (pro hladinu významnosti 0,05). Protože vypočítaná hodnota testového kritéria u je větší než hodnota kritická pro zvolenou hladinu významnosti, odmítáme nulovou hypotézu a přijímáme hypotézu alternativní. Mezi koeficienty korelace je tedy statisticky významný rozdíl. Pokud by vypočítaná hodnota testového kritéria u byla menší než hodnota kritická, nulovou hypotézu bychom přijali.

Pokud bychom pracovali na hladině významnosti 0,01, potom by kritická hodnota byla $u_{0,01} = 2,58$ (uváděné kritické hodnoty platí pro oboustranný test – srov. oddíl 3.1.2).

3.4.8 STUDENTŮV t-TEST

Studentův t-test je jedním z nejznámějších statistických testů významnosti pro metrická data. Pomocí Studentova t-testu můžeme rozhodnout, zda dva soubory dat, získané měřením ve dvou různých skupinách objektů (například žáků), mají stejný aritmetický průměr.

Příklad 42

Použití Studentova t-testu budeme ilustrovat na výzkumu, ve kterém se ověřovalo, zda dvě skupiny dětí předškolního věku (dívky a chlapci) mají stejnou průměrnou výšku.

Ve výběrovém souboru pětiletých dětí bylo 20 dívek a 15 chlapců. Při měření jejich výšky byly získány výsledky, které uvádějí tabulky 39 a 40. Z uvedených dat byla vypočítána průměrná výška dívek $\bar{x}_1 = 113,85$ cm a průměrná výška chlapců $\bar{x}_2 = 111,73$ cm.

Máme rozhodnout, zda průměrná výška dívek je větší než průměrná výška chlapců (zda zjištěné rozdíly ve výšce je možno připsat na vrub náhody, či nikoli):

H_0: Průměrná výška chlapců a dívek je stejná.
H_A: Průměrná výška dívek je větší než průměrná výška chlapců.

Zvolená hladina významnosti α = 0,05.

Nulovou hypotézu u Studentova t-testu testujeme pomocí kritéria t, které se vypočítává ze vztahu

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s} \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}} \tag{69}$$

kde $\bar{x}_1$ je průměr jedné skupiny (chlapci), $\bar{x}_2$ je průměr druhé skupiny (děvčata), n_1, n_2 jsou četnosti obou skupin a s je směrodatná odchylka.

Směrodatná odchylka s se vypočítává z hodnot získaných v obou skupinách, z tzv. **nestranného odhadu rozptylu** s^2 podle vzorců

$$s^2 = \frac{1}{n_1 + n_2 - 2}\left[\sum (x_{1i} - \bar{x}_1)^2 + \sum (x_{2j} - \bar{x}_2)^2\right] \tag{70}$$

$$s = \sqrt{s^2} \tag{71}$$

kde x_{1i} a x_{2i} jsou jednotlivé naměřené hodnoty v obou skupinách a význam ostatních symbolů je stejný jako ve vzorci (69).

Pro hodnoty uvedené v tabulkách 39 a 40 vychází

$$s^2 = \frac{1}{20 + 15 - 2}[374{,}557 + 238{,}935] = 18{,}59$$

$$s = 4{,}31$$

$$t = \frac{113{,}85 - 111{,}73}{4{,}31}\sqrt{\frac{20 \cdot 15}{20 + 15}} = 1{,}44$$

Tab. 39 *Výšky u skupiny pětiletých dívek*

Č.	**Výška x_{1i}**	$\boldsymbol{x_{1i} - \bar{x}_1}$	$\boldsymbol{(x_{1i} - \bar{x}_1)^2}$
1	117	3,15	9,923
2	113	–0,85	0,723
3	114	0,15	0,023
4	120	6,15	37,82
5	111	–2,85	8,123
6	112	–1,85	3,423
7	115	1,15	1,323
8	122	8,15	66,423
9	108	–5,85	34,223
10	117	3,15	9,923
11	113	–0,85	0,723
12	114	0,15	0,023
13	114	0,15	0,023
14	103	–10,85	117,723

▶

Č.	**Výška x_{1i}**	$x_{1i} - \bar{x}_1$	$(x_{1i} - \bar{x}_1)^2$
15	113	–0,85	0,723
16	117	3,15	9,923
17	114	0,15	0,023
18	107	–6,85	46,923
19	114	0,15	0,023
20	119	–5,15	26,523
	Σ 2 277		Σ 374,557

Tab. 40 *Výšky u skupiny pětiletých chlapců*

Č.	**Výška x_{2j}**	$x_{2j} - \bar{x}_2$	$(x_{2j} - \bar{x}_2)^2$
1	103	–8,73	76,213
2	106	–5,73	32,833
3	109	–2,73	7,453
4	109	–2,73	7,453
5	110	–1,73	2,993
6	111	–0,73	0,533
7	111	–0,73	0,533
8	111	–0,73	0,533
9	112	0,27	0,073
10	113	1,27	1,613
11	114	2,27	5,153
12	116	4,27	18,233
13	116	4,27	18,233
14	117	5,27	27,773
15	118	6,27	39,313
	Σ 1 676		Σ 238,935

Vypočítanou hodnotu *t* srovnáváme s kritickou hodnotou testového kritéria pro zvolenou hladinu významnosti a příslušný počet stupňů volnosti.

Počet stupňů volnosti se u Studentova t-testu určí podle vztahu

$$f = n_1 + n_2 - 2 \qquad (72)$$

kde *f* je počet stupňů volnosti, n_1 je četnost jedné skupiny a n_2 je četnost druhé skupiny.

V našem případě je $f = 15 + 20 - 2 = 33$. Kritická hodnota Studentova *t* pro 35 stupňů volnosti (nejblíže tabelovaná hodnota v příloze X) a hladinu významnosti 0,05 je $t_{0,05}(35) = 2,030$.

Protože vypočítaná hodnota je menší než hodnota kritická, přijímáme nulovou hypotézu. Mezi průměrnou výškou chlapců a průměrnou výškou děvčat **nejsou statisticky významné rozdíly** (zjištěné rozdíly je možno připsat na vrub náhody). Výsledek testu významnosti lze v tomto případě interpretovat také tak, že výsledky měření v obou skupinách pocházejí ze stejného základního souboru dat.

Nulovou hypotézu bychom mohli odmítnout v případě, že vypočítaná hodnota testového kritéria *t* je větší nebo rovna hodnotě kritické.

Při výpočtech testového kritéria *t* shora uvedeným způsobem je výpočet tzv. **součtů čtverců**, tj. výrazů $\sum(x_{1i}-\overline{x}_1)^2$ a $\sum(x_{2j}-x_2)^2$, poměrně pracný. Navíc je u tohoto postupu větší pravděpodobnost chyby (vzhledem k možným zaokrouhlovacím chybám). Výpočty součtů čtverců lze podstatně zjednodušit použitím vztahu

$$\sum(x_i-\overline{x})^2=\sum x_i^2-\overline{x}\cdot\sum x_i \tag{73}$$

Tento početně výhodnější postup výpočtu testového kritéria *t* budeme ilustrovat příkladem 43.

Příklad 43

Ve výzkumu se ověřovalo, zda dvě skupiny žáků mají stejnou úroveň vědomostí z fyziky. Vědomosti žáků se měřily pomocí didaktického testu, ve kterém bylo možno získat maximálně 10 bodů. Výsledky obou skupin žáků (skupina A a skupina B) uvádějí tabulky 41 a 42.

Pomocí Studentova *t*-testu máme ověřit, zda mezi průměrnými výsledky obou skupin jsou statisticky významné rozdíly.

H_0: Mezi průměrným počtem bodů dosaženým ve skupině A a průměrným počtem bodů dosaženým ve skupině B není rozdíl.
H_A: Mezi dosaženými průměry v obou skupinách jsou rozdíly.

Zvolená hladina významnosti: α = 0,05.

Tab. 41 *Výsledky didaktického testu (skupina A)*

Žák č.	Počet bodů x_i	x_i^2
1	10	100
2	9	81
3	8	64
4	6	36
5	7	49
6	4	16
7	6	36

▶

Žák č.	Počet bodů x_i	x_i^2
8	3	9
9	7	49
10	8	64
11	5	25
12	4	16
13	9	81
14	5	25
15	6	36
16	10	100
17	9	81
18	7	49
$n_A = 18$	Σ 123	Σ 917

Tab. 42 *Výsledky didaktického testu (skupina B)*

Žák č.	Počet bodů x_i	x_i^2
1	4	16
2	3	9
3	5	25
4	4	16
5	6	36
6	4	16
7	1	1
8	5	25
9	9	81
10	5	25
11	2	4
12	10	100
13	8	64
14	3	9
15	6	36
16	5	25
$n_B = 16$	Σ 80	Σ 488

Nejdříve vypočítáme aritmetické průměry v obou skupinách žáků

$$\bar{x}_A = \frac{123}{18} = 6{,}83 \qquad \bar{x}_B = \frac{80}{16} = 5{,}00$$

S použitím vztahů (73), (70) a (71) dostáváme

$$\sum(x_{Ai} - \overline{x}_A)^2 = 917 - 6{,}83 \cdot 123 = 76{,}91$$

$$\sum(x_{Bj} - \overline{x}_B)^2 = 488 - 5{,}0 \cdot 80 = 88{,}0$$

$$s^2 = \frac{1}{18 + 16 - 2} \cdot [76{,}91 + 88{,}0] = 5{,}153 \quad \text{a} \quad s = \sqrt{5{,}153} = 2{,}270$$

Studentovo testové kritérium potom vychází podle vzorce (69)

$$t = \frac{6{,}83 - 5{,}0}{2{,}270} \cdot \sqrt{\frac{18 \cdot 16}{18 + 16}} = 2{,}346$$

Vypočítanou hodnotu srovnáme s kritickou hodnotou Studentova *t* pro zvolenou hladinu významnosti a počet stupňů volnosti $f = 18 + 16 - 2 = 32$. V tabulce (příloha X) je nejblíže tabelovaná hodnota $t_{0{,}05}(30) = 2{,}042$. Vypočítaná hodnota *t* je větší než hodnota kritická, proto odmítáme nulovou hypotézu a přijímáme hypotézu alternativní. Mezi průměrným počtem bodů v testu z fyziky ve skupině A a průměrným počtem bodů v testu z fyziky ve skupině B je **statisticky významný** rozdíl. Skupina A dosáhla v didaktickém testu vyššího průměrného počtu bodů než skupina B.

Studentův t-test patří mezi parametrické testy významnosti. U parametrických testů (jak již bylo uvedeno) se požaduje splnění některých přesně vymezených podmínek, má-li být jejich použití oprávněné. U Studentova t-testu se požaduje, aby:

- základní soubor splňoval požadavek normálního rozdělení;
- byl dodržen požadavek homogenity rozptylu v obou srovnávaných skupinách (požaduje se, aby rozptyl hodnot v obou skupinách byl přibližně stejný);
- měření byla navzájem nezávislá;
- data byla metrická (intervalová nebo poměrová).

Splnění uvedených podmínek by se mělo vždy alespoň přibližně ověřovat. Pokud podmínky nejsou splněny, je lépe použít některého neparametrického testu. Neparametrickou alternativou ke Studentovu t-testu je například U-test Manna a Whitneyho.

Poznámka: *V některých případech lze nedodržení podmínek pro použití Studentova t-testu kompenzovat určitou korekcí výpočtu testového kritéria. Není-li například splněn požadavek homogenity rozptylu, lze testové kritérium vypočítat podle vztahu*

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\dfrac{s_1^2}{n_1 - 1} + \dfrac{s_2^2}{n_2 - 1}}} \tag{74}$$

kde s_1^2 *je rozptyl v jedné skupině a* s_2^2 *je rozptyl v druhé skupině.*

Vypočítané testové kritérium se ovšem v tomto případě srovnává s kritickou hodnotou, která je rovněž korigována na základě rozdílných rozptylů v obou skupinách. Toto korigované testové kritérium t_α^* *se vypočítá podle vzorce*

$$t_\alpha^* = \frac{t'_\alpha \cdot \frac{s_1^2}{n_1 - 1} + t''_\alpha \cdot \frac{s_2^2}{n_2 - 1}}{\frac{s_1^2}{n_1 - 1} + \frac{s_2^2}{n_2 - 1}} \tag{75}$$

kde t'_α *je kritická hodnota testového kritéria, t je vyhledaná ve statistických tabulkách pro* $n_1 - 1$ *stupňů volnosti a* t''_α *je kritická hodnota testového kritéria t vyhledaná pro* $n_2 - 1$ *stupňů volnosti.*

Možnosti analýzy na PC
Excel: vložit funkci → statistické → t-test (volba dva nezávislé výběry)
Statistica.cz: základní statistiky → t-test nezávislé skupiny
SPSS: *compare means → independent samples*

3.4.9 FISHERŮV-SNEDECORŮV F-TEST

Při mnoha statistických analýzách (např. při použití Studentova t-testu, při analýze rozptylu atd.) potřebujeme znát odpověď na otázku, zda ve dvou souborech dat je přibližně stejně velký rozptyl. V těchto případech nám může poskytnout odpověď Fisherův-Snedecorův F-test. U tohoto testu významnosti se rozptyly posuzují pomocí testového kritéria *F*, které se vypočítá ze vztahu

$$F = \frac{s_1^2}{s_2^2} = \frac{\sum (x_{1i} - \bar{x}_1)^2}{\sum (x_{2j} - \bar{x}_2)^2} \cdot \frac{n_2 - 1}{n_1 - 1} \tag{76}$$

kde s_1^2 je rozptyl v první skupině, s_2^2 je rozptyl ve druhé skupině, x_{1i} jsou jednotlivé hodnoty v první skupině, x_{2j} jsou jednotlivé hodnoty ve druhé skupině, $\bar{x}_1$, $\bar{x}_2$ jsou aritmetické průměry hodnot v obou skupinách a n_1, n_2 jsou četnosti v obou skupinách.

Pomocí testového kritéria *F* se potom testuje nulová hypotéza o rovnosti rozptylu v obou skupinách.

Příklad 44

Použití F-testu budeme ilustrovat na situaci, kdy máme rozhodnout, zda byl oprávněně použit Studentův t-test. Jak již bylo uvedeno, jednou z podmínek pro oprávněné použití Studentova t-testu (bez užití korekce) je homogenita rozptylu, tj. požadavek, aby rozptyly v obou srovnávaných skupinách byly zhruba stejné.

Budeme ověřovat situaci, kterou zachycují tabulky 41 a 42 (rozhodování o statistické významnosti rozdílu mezi průměrem skupiny A a průměrem skupiny B):

H_0: Rozptyl výsledků ve skupině A a rozptyl výsledků ve skupině B je stejně velký.
H_A: Rozptyly výsledků ve skupinách jsou rozdílné.

Zvolená hladina významnosti α = 0,05.

Při výpočtu testového kritéria *F* se vždy do čitatele zlomku dosazuje větší rozptyl (větší rozptyl se vždy dělí rozptylem menším), takže hodnota testového kritéria vychází vždy větší než 1.

Rozptyl výsledků ve skupině A vychází

$$s_A^2 = \frac{\sum (x_{Ai} - \bar{x}_A)^2}{n_A - 1} = \frac{76,91}{18-1} = 4,52$$

rozptyl výsledků ve skupině B je

$$s_B^2 = \frac{\sum (x_{Bj} - \bar{x}_B)^2}{n_B - 1} = \frac{88,0}{16-1} = 5,87$$

Větší rozptyl je ve skupině B, testové kritérium *F* vychází

$$F = \frac{5,87}{4,52} = 1,30$$

Vypočítanou hodnotu *F* srovnáváme s kritickou hodnotou tohoto kritéria pro zvolenou hladinu významnosti a počet stupňů volnosti, který se určí zvlášť pro každou skupinu podle vztahů

$$f_1 = n_1 - 1 \tag{77}$$

$$f_2 = n_2 - 1 \tag{78}$$

kde f_1, f_2 jsou počty stupňů volnosti v obou skupinách a n_1, n_2 jsou četnosti hodnot v obou skupinách.

V našem případě je ve skupině s větším rozptylem (tj. skupině B) 15 stupňů volnosti ($f_B = 16 - 1 = 15$) a ve skupině s menším rozptylem (skupině A) 17 stupňů volnosti ($f_A = 18 - 1 = 17$).

Ve statistických tabulkách (příloha XI) kritická hodnota pro 17 a 15 stupňů volnosti uvedena není, a proto použijeme hodnotu nejblíže tabelovanou, tj. hodnotu $F_{0,05}(15;15) = 2,40$. Zjišťujeme, že vypočítaná hodnota je menší než hodnota kritická, a proto přijímáme nulovou hypotézu. Mezi rozptyly v obou skupinách tedy nejsou statisticky významné rozdíly a použití Studentova t-testu proto bylo (z tohoto hlediska) oprávněné.

***Poznámka:** Tabulka v příloze XI platí pro hladinu významnosti 0,05. Pokud bychom chtěli testování významnosti provést na jiné hladině významnosti (např. 0,01), potom bychom museli použít rozsáhlejších statistických tabulek, např. Likeše a Lagy (1978) apod.*

Možnosti analýzy na PC
Excel: vložit funkci → statistické → F-test

3.4.10 PÁROVÝ t-TEST

Tento statistický test významnosti je možno použít například v případech, kdy jsme opakovaně (dvakrát) měřili u téže skupiny osob určitou vlastnost (proměnnou) a chceme rozhodnout, zda mezi výsledky těchto dvou měření jsou statisticky významné rozdíly. V těchto situacích není použitelný Studentův t-test, protože ten předpokládá, že oba výběry jsou nezávislé.

Při opakovaných měřeních téže vlastnosti je třeba mít jistotu, že měříme opravdu za stejných podmínek. V opačném případě by se mohlo stát, že eventuální zjištěná změna bude výsledkem působení faktorů, které nemáme pod svojí kontrolou (a nemusela by tudíž být výsledkem vlivu té proměnné, jejíž vliv analyzujeme).

S opakovaným měřením se setkáváme často při měření fyziologických nebo anatomických charakteristik u žáků. U pedagogických jevů má opakované měření smysl pouze tehdy, pokud druhé měření není ovlivněno zapamatováním měření prvního.

Příklad 45

Pomocí párového t-testu máme rozhodnout, zda u žáků 4. ročníku základní školy došlo během druhého pololetí školního roku ke zvýšení rychlosti hlasitého čtení.

U skupiny 25 žáků 4. ročníku byla změřena rychlost hlasitého čtení na začátku a na konci druhého pololetí. Výsledky obou měření jsou uvedeny v tabulce 43.

H_0: Mezi rychlostí hlasitého čtení žáků na začátku a na konci druhého pololetí není rozdíl.
H_A: Rychlost hlasitého čtení žáků je na konci pololetí jiná než na začátku pololetí.

Zvolená hladina významnosti: α = 0,05.
Nulovou hypotézu budeme testovat pomocí testového kritéria *t*, které se u párového t-testu vypočítává ze vzorce

$$t = \frac{\bar{d} \cdot \sqrt{n \cdot (n-1)}}{\sqrt{\sum (d - \bar{d})^2}} \tag{79}$$

kde *n* je počet párů hodnot, *d* je diference mezi hodnotami u jednoho páru a $\bar{d}$ je průměrná diference.

V uvedené tabulce byla u každého páru hodnot vypočítána diference *d* a průměrná diference $\bar{d} = \frac{\sum d}{n} = \frac{47}{25} = 1{,}88$ (při výpočtu hodnoty $\sum d$ je třeba respektovat znaménka jednotlivých diferencí).

***Tab. 43** Výpočet testového kritéria t pro párový t-test*

Číslo žáka	Rychlost čtení (slov/min)		*d*	d^2
	začátek	konec		
1	154	160	+6	36
2	82	89	+7	49
3	90	88	-2	4
4	111	114	+3	9
5	106	105	-1	1
6	115	118	+3	9
7	94	90	-4	16
8	78	81	+3	9
9	123	126	+3	9
10	125	130	+5	25
11	103	110	+7	49
12	144	138	–6	36
13	203	207	+4	16
14	116	112	–4	16
15	95	94	–1	1
16	73	76	+3	9
17	100	104	+4	16
18	182	178	–4	16
19	159	162	+3	9
20	115	112	–3	9
21	118	123	+5	25
22	53	59	+6	36
23	94	96	+2	4
24	106	110	+4	16
25	125	129	+4	16
			Σ 47	Σ 441

Při výpočtu testového kritéria bylo využito vztahu

$$\sum (d-\bar{d})^2 = \sum d^2 - \bar{d}\sum d \tag{80}$$

Testové kritérium potom vychází

$$t = \frac{1{,}88 \cdot \sqrt{25(25-1)}}{\sqrt{\sum 441 - 1{,}88 \cdot 47}} = 2{,}453$$

Vypočítanou hodnotu $t = 2{,}453$ srovnáváme s kritickou hodnotou testového kritéria pro zvolenou hladinu významnosti a počet stupňů volnosti f, který se určí podle vztahu

$$f = n - 1 \tag{81}$$

kde n je počet párů měření.

Protože vypočítaná hodnota testového kritéria je větší než nalezená kritická hodnota $t_{0,05}(24) = 2{,}064$ (příloha X), **odmítáme nulovou hypotézu** a přijímáme **hypotézu alternativní**. Rychlost hlasitého čtení žáků je na konci druhého pololetí jiná (větší) než na začátku pololetí.

Možnosti analýzy na PC
Excel: vložit funkci → statistické → t-test (volba závislé výběry)
Statistica.cz: základní statistiky → t-test, závislé vzorky
SPSS: *compare means → paired samples*

3.4.10.1 Princip analýzy rozptylu

Analýza rozptylu, kterou odvodil R. A. Fisher před několika desítkami let, je moderní a slibnou statistickou metodou, jež zatím v našich pedagogických výzkumech příliš nezdomácněla. Aplikace analýzy rozptylu může přitom v pedagogickém výzkumu přinášet pozoruhodně přesné a spolehlivé výsledky.

Základní myšlenku analýzy rozptylu je možno vyjádřit následujícím způsobem: Jestliže máme určitý soubor metrických dat (celkem n hodnot), který je rozdělen do několika (k) skupin, potom můžeme vypočítat dva na sobě nezávislé odhady rozptylu:

- první z těchto odhadů vychází z rozptylu mezi průměry skupin;
- druhý vychází z rozptylu uvnitř skupin.

Rozptyl uvnitř skupin můžeme vypočítat tak, že nejdříve vypočítáme rozptyly pro všechny skupiny a z nich potom určíme průměr. Rozptyl uvnitř skupin představuje lepší a spolehlivější odhad, protože není ovlivněn možnými rozdíly mezi skupinami a je dán pouze náhodným kolísáním výsledků. Naproti tomu rozptyl mezi skupinami je na rozdílech mezi skupinami závislý. Jestliže rozptyl „mezi“ skupinami je významně větší než rozptyl „uvnitř“ skupin, znamená to, že skupiny nejsou náhodnými výběry z téhož základního souboru (tzn., že výsledky ve skupinách se statisticky významně liší). K objektivnímu posouzení poměru obou rozptylů se používá F-testu

$$F = \frac{\textit{rozptyl mezi skupinami}}{\textit{rozptyl uvnitř skupin}} \tag{82}$$

3.4.10.2 Jednoduchá analýza rozptylu

Otázka, kterou pomocí analýzy rozptylu často řešíme, je, zda mezi několika zjištěnými průměry naměřených dat jsou, či nejsou významné rozdíly.

Příklad 46

Princip analýzy rozptylu budeme ilustrovat na příkladu jednoduchého fiktivního experimentu, ve kterém se jedná o posuzování vlivu tří metod učení (metoda A, B, C) na výsledky učení žáků.

Celkem bylo zkoumáno 30 žáků, přičemž 12 žáků se učilo metodou A, 10 žáků metodou B a 8 žáků používalo metodu C. Žáci byli rozděleni do skupin náhodně a předpokládáme, že měli k učení jinak stejné předpoklady.

Po uplynutí vhodné doby byl všem 30 žákům zadán didaktický test, který spolehlivě změřil úroveň jejich vědomostí. Získané výsledky uvádí tabulka 44.

Tab. 44 *Výsledky žáků v jednoduchém experimentu*

Skupina (metoda)	Výsledky žáků	Počet žáků	Celkový počet bodů	Σx^2	Aritmetický průměr
A	0 1 2 2 2 2 3 3 4 5 6 7 9	12	44	238	3,66
B	3 4 5 6 7 8 8 9 9 10	10	69	525	6,90
C	1 3 5 5 6 7 7 8	8	42	258	5,25
		Σ 30	Σ 155	Σ 1 021	

Abychom si výpočet odhadů rozptylu usnadnili, vypočítáváme nejdříve tzv. součty čtverců (*SČ*). **Součtem čtverců** se rozumí součet čtverců odchylek od aritmetického průměru, tj. $\sum(x-\bar{x})^2$. Abychom ze součtu čtverců obdrželi rozptyl, je třeba jej dělit počtem stupňů volnosti, tj. $n - 1$ (nikoli n) vzhledem k tomu, že odhadujeme rozptyl základního souboru z výběrových dat.

V uvedeném příkladě lze vypočítat součty čtverců trojího druhu: celkový součet čtverců, součet čtverců mezi skupinami a součet čtverců uvnitř skupin. Lze dokázat, že mezi těmito součty čtverců platí vztah

$$\textit{celkový SČ} = \textit{SČ uvnitř skupin} + \textit{SČ mezi skupinami} \qquad (83)$$

Kromě konečného stadia se při analýze rozptylu pracuje jen se součty čtverců. Má to výhodu v tom, že součty čtverců jsou vždy aditivní (jednotlivé *SČ* je možno sčítat a odčítat). Rozptyly naopak aditivní nejsou, takže je nelze ani sčítat, ani odčítat.

Celkový součet čtverců je možno vypočítat podle vzorce

$$\textit{celkový SČ} = \sum(x-\bar{x})^2 = \sum x^2 - \bar{x}\sum x = \sum x^2 - \frac{\left(\sum x\right)^2}{n} \qquad (84)$$

kde x jsou jednotlivé naměřené hodnoty (ve všech skupinách), $\bar{x}$ je aritmetický průměr všech hodnot, $\sum x$ je součet všech naměřených hodnot.

Pro data uvedená v tabulce 44 dostáváme

$$\textit{celkový SČ} = 1\,021 - \frac{155^2}{30} = 220{,}17$$

Součet čtverců mezi skupinami je mírou variability pro průměry skupin, tj. mírou kolísání průměrů skupin kolem celkového průměru. Lze odvodit, že součet čtverců mezi skupinami lze vypočítat podle vztahu

$$SČ\ mezi\ skupinami = \frac{\left(\sum x_a\right)^2}{n_a} + \frac{\left(\sum x_b\right)^2}{n_b} + \frac{\left(\sum x_c\right)^2}{n_c} - \frac{\left(\sum x\right)^2}{n} \tag{85}$$

kde x_a, x_b, x_c jsou naměřené hodnoty v jednotlivých skupinách, $\sum x$ je součet všech naměřených hodnot (ve všech skupinách), n_a, n_b, n_c jsou četnosti hodnot ve skupinách a n celková četnost všech hodnot.

Dosadíme-li do vztahu (85) hodnoty z tabulky 44, dostáváme

$$SČ\ mezi\ skupinami = \frac{44^2}{12} + \frac{69^2}{10} + \frac{42^2}{8} - \frac{155^2}{30} = 57{,}10$$

Součet čtverců uvnitř skupin lze vypočítat způsobem naznačeným shora, ale snadnější je použití vztahu (83), ze kterého vyplývá, že

$$SČ\ uvnitř\ skupin = celkový\ SČ - SČ\ mezi\ skupinami \tag{86}$$

Pro hodnoty z tabulky 44 potom vychází

$$SČ\ uvnitř\ skupin = 220{,}17 - 57{,}10 = 163{,}07$$

Výsledky analýzy rozptylu je zvykem zapisovat do přehledné tabulky (tab. 45).

Tab. 45 *Výsledky jednofaktorové analýzy rozptylu*

Zdroj rozptylu	SČ	Stupně volnosti	Rozptyl	*F*
mezi skupinami	57,10	2	28,55	4,73
uvnitř skupin	163,07	27	6,04	
celkem	220,17	29		

Rozptyl mezi skupinami je vypočítán ze tří skupinových průměrů, má proto pouze dva stupně volnosti. Rozptyl uvnitř skupin (označovaný také jako rozptyl reziduální) je vypočítán ze tří skupin hodnot, a má proto 11 + 9 + 7 = 30 – 3 = 27 stupňů volnosti. Celkový rozptyl byl vypočítán ze všech hodnot, a má proto 30 – 1 = 29 stupňů volnosti.

Dalším krokem při jednofaktorové analýze rozptylu je rozhodnutí, zda rozptyl mezi skupinami je signifikantně větší než rozptyl uvnitř skupin. Toto rozhodnutí učiníme na základě F-testu. Nejdříve formulujeme nulovou a alternativní hypotézu:

H_0: Mezi rozptylem mezi skupinami a rozptylem uvnitř skupin není rozdíl.
H_A: Rozptyl mezi skupinami je větší než rozptyl uvnitř skupin.

Testování provedeme na hladině významnosti α = 0,05.
Testové kritérium *F* vypočítáme podle vztahu (82)

$$F = \frac{28{,}55}{6{,}04} = 4{,}73$$

Vypočítanou hodnotu $F = 4,73$ srovnáme s kritickou hodnotou F pro hladinu významnosti 0,05 a $f_1 = 2$ a $f_2 = 27$ stupňů volnosti, tj. s hodnotou $F_{0,05}(2,27) = 3,35$ (příloha XI). Protože vypočítaná hodnota je větší než hodnota kritická, odmítáme nulovou hypotézu a přijímáme hypotézu alternativní. Rozptyl mezi skupinami je tedy významně větší než rozptyl uvnitř skupin. Z tohoto zjištění vyplývá, že skupiny nelze považovat za náhodné výběry z téhož základního souboru (jinak řečeno: mezi výsledky žáků v jednotlivých skupinách jsou statisticky významné rozdíly).

Poznámka: *Statisticky významné rozdíly mezi výsledky skupin je samozřejmě možno připsat na vrub používaným metodám učení pouze v případě, že máme záruku, že na žáky nepůsobily žádné jiné (nekontrolované) vlivy. V popsaném uspořádání experimentu by tento předpoklad byl asi obtížně splnitelný a uvedený fiktivní příklad slouží pouze k ilustraci postupu.*

Možnosti analýzy na PC
Statistica.cz: analýza rozptylu ANOVA/MANOVA → jednofaktorová ANOVA
SPSS: *compare means* → *one-way* ANOVA

Duncanův test

Pokud jednofaktorová analýza vede k závěru, že mezi průměry skupin jsou statisticky významné rozdíly, zpravidla nás také zajímá, mezi kterými průměry se signifikantní rozdíly projevují.

Studentův t-test lze použít jen pro srovnávání dvou průměrů, které leží v řadě průměrů seřazených podle velikosti vedle sebe. Pokud bychom použili Studentova t-testu na testování významnosti rozdílů mezi vzdálenějšími průměry (např. mezi prvním a třetím, prvním a čtvrtým atd.), došlo by k podstatnému zvýšení hladiny významnosti a tím ke snížení spolehlivosti testu. Chceme-li posoudit rozdíly mezi všemi průměry na nezměněné hladině významnosti, lze k tomu použít Duncanův test.

Příklad 47

Použití Duncanova testu budeme demonstrovat na situaci, která byla uvažována při výkladu jednofaktorové analýzy rozptylu. Máme rozhodnout, mezi kterými průměry jsou statisticky významné rozdíly.

Nejdříve seřadíme průměry skupin vzestupně podle velikosti:

A	C	B
3,66	5,25	6,90

Rozdíl mezi dvěma průměry je statisticky významný (Komenda a Klementa, 1881), jestliže

$$(\bar{x}_i - \bar{x}_j) \cdot \sqrt{\frac{2n_i \cdot n_j}{n_i + n_j}} \geq s \cdot R_\alpha \tag{87}$$

kde $\overline{x}_i$ je průměr jedné skupiny, $\overline{x}_j$ je průměr druhé skupiny, n_i je počet hodnot v jedné skupině, n_j je počet hodnot v druhé skupině, s je směrodatná odchylka určená z rozptylu uvnitř skupin (reziduálního rozptylu) a R_a je hodnota, která se určí ze statistických tabulek (příloha XII).

Hodnota R_a se v tabulkách vyhledává pro zvolenou hladinu významnosti, pro počet stupňů volnosti f rozptylu uvnitř skupin a pro počet průměrů p, které leží v uspořádané řadě mezi $\overline{x}_i$ a $\overline{x}_j$ (včetně krajních hodnot).

V daném případě vyhledáváme R_a pro hladinu významnosti 0,05, pro počet stupňů volnosti $f = 30 - 3 = 27$ a pro $p = 3$ a $p = 2$ (při srovnávání tří průměrů může být p buď 3, nebo 2).

Nalezené hodnoty R_a zapíšeme do tabulky (tab. 46). Dále určíme z rozptylu uvnitř skupin směrodatnou odchylku $s = \sqrt{s^2} = \sqrt{6{,}04} = 2{,}458$ a vypočítáme hodnoty $s \cdot R_\alpha$. (Hodnoty R_a v tabulce jsou vyhledány pro nejblíže tabelovaných 30 stupňů volnosti.)

Tab. 46 *Duncanův test*

p	2	3
R_a	2,888	3,035
$s \cdot R_\alpha$	7,099	7,460

Průměry seřazené podle velikosti navzájem srovnáváme dosazením do vztahu (87)

$\boxed{\overline{x}_b - \overline{x}_a}$ p = 3

$$(6{,}90 - 3{,}66) \cdot \sqrt{\frac{2 \cdot 10 \cdot 12}{10 + 12}} = 10{,}701$$

10,701 > 7,460 (významný rozdíl)

$\boxed{\overline{x}_c - \overline{x}_a}$ p = 2

$$(5{,}25 - 3{,}66) \cdot \sqrt{\frac{2 \cdot 8 \cdot 12}{8 + 12}} = 4{,}926$$

4,926 < 7,099

$\boxed{\overline{x}_b - \overline{x}_c}$ p = 2

$$(5{,}25 - 3{,}66) \cdot \sqrt{\frac{2 \cdot 8 \cdot 12}{8 + 12}} = 4{,}926$$

4,919 < 7,099

Výsledky Duncanova testu uvádí tabulka 47. Hvězdičkou je označen statisticky významný rozdíl mezi průměrem skupiny A a průměrem skupiny B. Ze tří posuzovaných rozdílů mezi průměry je tedy pouze jeden statisticky významný, zatímco další dva jsou nevýznamné.

Tab. 47 *Rozdíly mezi průměry skupin*

	A	B	C
A	–		
B	3,24*	–	
C	1,59	1,65	–

Výhodou jednofaktorové analýzy rozptylu oproti Studentovu t-testu je, že jedním výpočtem testujeme současně několik rozdílů mezi průměry. Pokud se ukáže, že rozdíly mezi skupinami nejsou statisticky významné, nemá již smysl počítat t (popřípadě provádět Duncanův test).

Poznámka: *V případě dvou skupin se stává test F algebraicky ekvivalentní Studentovu t-testu (v tomto případě platí, že $F = t^2$). Studentův t-test lze tedy považovat za zvláštní případ analýzy rozptylu, kdy počet skupin $k = 2$.*

Předpokladem oprávněného použití analýzy rozptylu je, že ve všech srovnávaných skupinách je přibližně stejně velký rozptyl. Tento předpoklad by se měl před každým použitím analýzy rozptylu ověřovat tzv. **Bartlettovým testem**, který (vzhledem k rozsahu tohoto textu) neuvádíme a odkazujeme na literaturu, například Komenda a Klementa (1981) a další.

V popsaném příkladě analýzy rozptylu byl sledován vliv jen jednoho faktoru (jedné nezávisle proměnné) na výsledek pokusu. Vlastní doménou analýzy rozptylu však jsou výzkumy, ve kterých se současně mění více podmínek (nezávisle proměnných).

3.4.10.3 Dvoufaktorová analýza rozptylu

Tento druh analýzy rozptylu umožňuje postižení vlivu dvou současně působících faktorů (nezávisle proměnných) na určitou závisle proměnnou.

Příklad 48

Postup dvoufaktorové analýzy rozptylu objasníme na příkladu fiktivního výzkumu, ve kterém se sledoval vliv rozumových schopností a pohlaví na úroveň vědomostí žáků ve fyzice.

Skupina 24 žáků základní školy (12 dívek a 12 chlapců) byla vyšetřena testem rozumových schopností a podle výsledků vyšetření byla rozdělena do tří kategorií: „nižší rozumové schopnosti“, „průměrné rozumové schopnosti“ a „vyšší rozumové schopnosti“. Poté byla celá skupina vyzkoušena didaktickým testem z fyziky, který měřil úroveň jejich vědomostí. Výsledky testování uvádí tabulka 48.

Máme rozhodnout, jak se na úrovni vědomostí žáků z fyziky podílí pohlaví žáků (faktor A) a rozumové schopnosti (faktor B).

Tab. 48 *Dvoufaktorová analýza rozptylu*

		Rozumové schopnosti (faktor B)			
		nižší	průměrné	vyšší	Σ
Pohlaví (faktor A)	dívky	22 23 23 24	23 23 23 24	24 25 27 28	289
	chlapci	22 23 23 24	27 28 29 29	24 25 25 26	305
	Σ	184	206	204	594

Z tabulky 48 lze vyčíst, jakých výsledků žáci dosáhli při určitých kombinacích působení obou studovaných faktorů. Například dívky s nižšími rozumovými schopnostmi dosáhly v testu z fyziky 22, 23, 23 a 24 bodů, chlapci s vyššími rozumovými schopnostmi dosáhli v testu z fyziky výsledků:

24, 25, 25 a 26 bodů atd. Celkově tedy můžeme skupinu žáků rozdělit na 6 menších skupin podle kombinace působících faktorů.

Prvním krokem při dvoufaktorové analýze rozptylu bude rozhodnutí, zda mezi výsledky takto vytvořených šesti skupin žáků jsou vůbec statisticky významné rozdíly. Toto rozhodnutí provedeme způsobem, který byl popsán u jednofaktorové analýzy rozptylu. Pokud by se při této analýze ukázalo, že mezi skupinami statisticky významné rozdíly nejsou, potom by další analýza neměla smysl.

Nejprve vypočítáme celkový součet čtverců podle vztahu (84)

$$celkový\ SČ = 22^2 + 23^2 + ... + 26^2 - \frac{594^2}{24} = 108{,}5$$

Dále určíme podle vztahu (85) součet čtverců mezi skupinami

$$SČ\ mezi\ skupinami = \frac{92^2}{4} + \frac{93^2}{4} + \frac{104^2}{4} + \frac{92^2}{4} + \frac{113^2}{4} + \frac{100^2}{4} - \frac{594^2}{24} = 89{,}0$$

Podle vztahu (86) vypočítáme součet čtverců uvnitř skupin

$$SČ\ uvnitř\ skupin = 108{,}5 - 89{,}0 = 19{,}5$$

Výsledky provedené jednofaktorové analýzy rozptylu shrneme do tabulky (tab. 49).

Tab. 49 *Výsledky jednofaktorové analýzy rozptylu*

Zdroj rozptylu	SČ	Stupně volnosti	Rozptyl	F
mezi skupinami	89,0	6 – 1 = 5	17,80	16,48
uvnitř skupin	19,5	24 – 6 = 18	1,08	
celkem	108,5	24 – 1 = 23		

Podle vztahu (82) vypočítáme testové kritérium *F*

$$F = \frac{17{,}80}{1{,}08} = 16{,}48$$

Vypočítanou hodnotu $F = 16{,}48$ srovnáváme s kritickou hodnotou pro hladinu významnosti 0,05 a příslušné počty stupňů volnosti ($f_1 = 6 - 1 = 5$ a $f_2 = 24 - 6 = 18$)). Zjišťujeme, že vypočítaná hodnota je větší než hodnota kritická $F_{0,05}(5,18) = 2{,}773$. Proto odmítáme nulovou hypotézu o rovnosti obou rozptylů a konstatujeme, že rozptyl mezi skupinami je významně větší než rozptyl uvnitř skupin. Z toho plyne, že mezi výsledky skupin jsou statisticky významné rozdíly.

Značně velký rozptyl mezi skupinami má zřejmě tři zdroje: vliv faktoru A (pohlaví žáků), vliv faktoru B (rozumové schopnosti žáků) a vliv vzájemného působení obou faktorů (interakce A x B). Bude platit, že

$$SČ\ mezi\ skupinami\ =\ SČ\ mezi\ A\ +\ SČ\ mezi\ B\ +\ SČ\ A \times B \qquad (88)$$

SČ mezi A a *SČ* mezi B lze vypočítat obdobně jako *SČ* mezi skupinami

$$SČ\ mezi\ A = \frac{289^2}{12} + \frac{305^2}{12} - \frac{594^2}{24} = 10{,}66$$

$$SČ\ mezi\ B = \frac{184^2}{8} + \frac{206^2}{8} + \frac{204^2}{8} - \frac{594^2}{24} = 37{,}00$$

SČ pro interakci A × B určíme na základě vztahu (88)

$$SČ\ A \times B\ =\ SČ\ mezi\ skupinami\ -\ SČ\ mezi\ A\ -\ SČ\ mezi\ B$$

$$SČ\ A \times B\ =\ 89{,}00 - 10{,}66 - 37{,}00 = 41{,}34$$

Výsledky, které jsme v této části analýzy rozptylu obdrželi, shrneme opět do tabulky (tab. 50).

Tab. 50 *Výsledky dvoufaktorové analýzy rozptylu*

Zdroj rozptylu	*SČ*	Stupně volnosti	Rozptyl	*F*
mezi A	10,66	2 – 1 = 1	10,66	9,870
mezi B	37, 00	3 – 1 = 2	18,50	17,130
interakce A × B	41,34	(2 – 1) · (3 – 1) = 2	20,67	19,139
uvnitř skupin	19,50	24 – 6 = 18	1,08	
celkem	108,50	24 – 1 = 23		

Významnost působení jednotlivých faktorů posuzujeme opět pomocí testového kritéria *F*, které vypočítáváme i v tomto případě jako podíl mezi příslušným rozptylem a rozptylem uvnitř skupin. Zjišťujeme, že významný vliv mají oba faktory (A i B), ale i faktor vzájemného působení (interakce) A × B. Vypočítané hodnoty testového kritéria totiž výrazně překračují kritické hodnoty:

$F_{0,05}(1,18) = 4{,}414$ (mezi A)
$F_{0,05}(2,18) = 3{,}555$ (mezi B a A × B)

Jestliže vyjdeme ze zjištěných proporcí mezi součty čtverců, můžeme konstatovat, že sledované faktory (A, B, interakce A × B) vysvětlují celkem asi 82 % variability výsledků, zbytek (tj. asi 18 %) variability zůstává neobjasněn. Faktor A (pohlaví) přitom objasňuje 9,8 % zjištěné variability, faktor B (rozumové schopnosti) vysvětluje 34,1 % variability a faktor interakce A × B (vzájemné působení obou faktorů) vysvětluje 38,1 % zjištěné variability. Uvedené výsledky představují velmi konkrétní odpověď na otázku, do jaké míry sledované faktory ovlivňují úroveň vědomostí žáků ve fyzice.

Poznámka: *Uvedený postup analýzy rozptylu by se měl používat jen v případě, že v každém poli tabulky (v každé dílčí skupině) je stejný počet hodnot (v popisovaném příkladě byly v každém poli tabulky čtyři hodnoty). Pokud by tento předpoklad nebyl splněn, je nutno každý jednotlivý součet pole $\sum x$ a každý součet umocněných naměřených hodnot $\sum x^2$ korigovat podle vztahů*

$$\left(\sum x\right)_{kor} = \frac{\bar{n}}{n_n} \cdot \sum x \tag{89}$$

$$\left(\sum x^2\right)_{kor} = \frac{\bar{n}}{n_n} \cdot \sum x^2 \tag{90}$$

kde $\left(\sum x\right)_{kor}$ *a* $\left(\sum x^2\right)_{kor}$ *jsou opravené součty příslušných hodnot,*
kde je průměrný počet hodnot v polích tabulky, x jsou neopravené hodnoty a n_n *je neopravený (skutečný) počet hodnot v poli tabulky. Tuto korekci je nutno provést vždy ve všech polích tabulky.*

Pokud data, která máme k dispozici, nejsou metrická, můžeme k jejich analýze použít neparametrickou analogii analýzy rozptylu – **Kruskalův-Wallisův H-test** (Komenda a Klementa, 1981).

Možnosti analýzy na PC
Statistica.cz: analýza rozptylu ANOVA/MANOVA → vícenásobná ANOVA

3.5 PRINCIP FAKTOROVÉ ANALÝZY

Faktorová analýza je metoda, která umožňuje určit základní proměnné (faktory), které ovlivňují provedená měření určitého objektu. Jestliže u určitého objektu provedeme několik různých měření, může se stát, že některá měření se budou navzájem tu více, tu méně podobat. Ta měření, která dávají podobné výsledky, je dobré studovat společně, protože vyjadřují něco společného – **společný faktor**. Faktorová analýza umožňuje rozhodnout, které základní faktory (základní proměnné) se v provedených měřeních projevují, a tak umožňuje nahradit velký počet provedených měření několika základními faktory.

Zakladatelem faktorové analýzy je anglický psycholog Charles Spearman, který se pomocí ní pokusil objasnit strukturu lidské inteligence. Zjistil, že základ inteligence tvoří několik faktorů, které bývají označovány například jako verbální schopnosti, matematické schopnosti, abstraktní usuzování, prostorová představivost, paměť atd.

Příklad 49

Podstatu faktorové analýzy vyložíme na příkladu fiktivního výzkumu, ve kterém bylo skupině žáků základní školy postupně zadáno celkem šest různých didaktických testů:

- test z matematiky, který zkoušel řešení slovních úloh (M);
- test z fyziky, který zkoušel řešení fyzikálních úloh o pohybu (F);
- test z chemie, který zkoušel stechiometrické výpočty (Ch);
- test porozumění čtenému textu (Tx);
- test znalostí českého pravopisu (Č);
- test slovní zásoby (z angličtiny) (A).

Pomocí faktorové analýzy budeme hledat odpověď na otázku, zda provedená měření (testování) jsou ovlivněna určitými faktory, které za nimi stojí.

Při faktorové analýze je nejprve nutno všechna měření vyhodnotit a potom vypočítat pro výsledky každého měření (testu) s každým měřením (testem) koeficienty korelace. Tyto koeficienty korelace se zapisují do tzv. **korelační R-matice**. Předpokládejme, že v našem fiktivním případě bychom získali korelační matici, kterou uvádí tabulka 51.

Tab. 51 *Korelační R-matice*

	M	F	Ch	Tx	Č	A
M	–	0,72	0,63	0,09	0,09	0,00
F	0,72	–	0,57	0,15	0,16	0,09
Ch	0,63	0,57	–	0,14	0,15	0,09
Tx	0,09	0,15	0,14	–	0,57	0,63
Č	0,09	0,16	0,15	0,57	–	0,72
A	0,00	0,09	0,09	0,63	0,72	–

Z uvedené R-matice lze snadno vyčíst, které dvojice měření (testů) navzájem vysoce korelují a které korelují velmi málo. Měření (testy), mezi nimiž je vysoká korelace, měří zřejmě něco velmi podobného. Prostudujeme-li hodnoty v R-matici, zjišťujeme, že vysoké korelační koeficienty se objevují mezi testy M, F, Ch a také mezi testy Tx, Č a A. Mezi testy z obou těchto skupin navzájem jsou naopak korelační koeficienty velmi nízké. Na základě studia R-matice můžeme vyslovit závěr, že mezi provedenými měřeními (testy) se vyskytují dvě seskupení – působí zde tedy dva faktory (v R-matici jsou vyznačeny šedým podtiskem).

Zjištěné faktory se při faktorové analýze obvykle označují názvy, které vystihují podstatu toho „společného", co daný faktor vyjadřuje. Při volbě názvu pro zjištěný faktor se jedná o vytvoření určitého konstruktu (konstrukt = pojem uměle vytvořený k určitým vědeckým účelům), který by vyjadřoval a charakterizoval jednotu příslušných měření (provedených testů). V našem fiktivním příkladě se například můžeme ptát, co společného mají testy z matematiky, fyziky a chemie nebo co společného mají testy porozumění čtenému textu, test znalostí českého pravopisu a test slovní zásoby v angličtině. Společným prvkem pro první skupinu testů budou zřejmě výpočty, a proto bychom mohli společný faktor označit například jako **faktor početní**. Obdobně bychom mohli společný faktor vystihující shodu ve druhé skupině testů označit jako **faktor verbální**. Hledání vystihujících označení pro nalezené faktory je ve skutečných výzkumech podstatně složitější než v uvedeném případě. Vzhledem k tomu, že při opakovaném šetření často dostáváme poněkud rozdílné výsledky, musíme i označení faktorů upřesňovat. Názvy faktorů jsou tedy u faktorové analýzy jen předběžné (tentativní).

V popsaném příkladě jednotlivé testy měřily prakticky vždy jen jeden faktor. Jestliže test měří jen jeden faktor, říkáme, že je **faktorově čistý**. Při reálných měřeních nedostáváme téměř nikdy tak jednoznačné a jednoduché výsledky. Obyčejně se zjišťuje, že jednotlivé testy neměří pouze jeden faktor, nýbrž několik faktorů současně. Měří-li test současně několik faktorů, říkáme, že je **sycen několika faktory**. Základním úkolem faktorové analýzy je zjistit, jak dalece jsou jednotlivá měření (testy) určitými faktory sycena.

Prvním výsledkem faktorové analýzy je zpravidla získání tzv. **nerotované faktorové matice** (tab. 52). Čísla uvedená v matici jsou tzv. **faktorové náboje**. Faktorové náboje (mohou nabývat hodnot od –1 přes 0 do +1) vyjadřují stupeň korelace (korelační koeficient) mezi jednotlivými měřeními (testy) a příslušnými faktory. Faktorový náboj +1 by znamenal, že příslušný test je zcela nasycen daným faktorem (šlo by o faktorově čistý test). Faktorový náboj 0 by znamenal, že výsledek testu není daným faktorem vůbec dotčen. Záporné hodnoty faktorových nábojů vypovídají o tom, že daný test je faktorem sycen v negativním smyslu.

Nerotovaná faktorová matice zpravidla není vhodná pro interpretaci výsledků faktorové analýzy. Z tohoto důvodu se nerotovaná faktorová matice zpravidla převádí do „čitelnější podoby“ na základě tzv. **rotace**. Tímto postupem získáváme tzv. **rotovanou faktorovou matici**, která je vhodnější pro interpretaci výsledků (tab. 53). Při rotaci jde v podstatě o nalezení „správného úhlu pohledu“ na data ve faktorové matici. Nalezená (nerotovaná) faktorová matice totiž představuje jen jedno z mnoha možných řešení, která z dané R-matice vyplývají (podrobnější vysvětlení lze nalézt v literatuře, například Kerlinger (1972), Blahuš (1984) apod.).

Tab. 52 *Nerotovaná faktorová matice*

Testy	**1**	**2**	**3**
M	0,61	–0,60	0,03
F	0,65	–0,55	0,11
Ch	0,60	–0,46	–0,09
Tx	0,60	0,46	–0,09
Č	0,65	0,52	0,11
A	0,61	0,60	0,03

Tab. 53 *Rotovaná faktorová matice*

Testy	**1**	**2**	**3**	h^2
M	0,86	0,00	0,03	0,74
F	0,83	0,09	0,11	0,71
Ch	0,75	0,10	–0,09	0,57
Tx	0,10	0,75	–0,09	0,57
Č	0,09	0,83	0,11	0,71
A	0,00	0,86	0,03	0,74

Uvedená rotovaná matice byla získána na základě tzv. **pravoúhlé rotace**. U pravoúhlé rotace platí, že nalezené faktory jsou navzájem nezávislé (korelace mezi faktory je nulová).

Faktor, který jsme v našem fiktivním příkladě pojmenovali jako „faktor početní", je v uvedené faktorové matici označen číslem 1. Číslem 2 je označen faktor, který jsme pojmenovali jako „faktor verbální". Kromě prvních dvou „silných" faktorů byl při faktorové analýze nalezen ještě další, velmi slabý faktor, který je provedenými testy sycen jen nepatrně (označen číslem 3).

V posledním sloupci faktorové matice zprava je ještě uvedena tzv. **komunalita *h*²**. Komunalita vyjadřuje společný faktorový rozptyl příslušného testu (měření) a poskytuje informaci o tom, jaká část variability u daného testu (měření) je vysvětlena nalezenými faktory.

Komunalitu lze vypočítat jako součet čtverců (součet druhých mocnin) faktorových nábojů. Například pro test z matematiky (M) je komunalita

$$h^2 = 0{,}86 + 0{,}00^2 + 0{,}03^2 = 0{,}74$$

Komunalita má význam pro interpretaci výsledků faktorové analýzy.

Faktorová analýza může pomoci při řešení řady obtížných metodologických problémů pedagogiky i psychologie. Umožňuje odhalit strukturu mnoha komplexních proměnných, a tím umožňuje přesnější a hlubší poznání reality. Protože však faktorová analýza představuje vysoce formalizovaný popis reality, klade vysoké nároky na odbornou úroveň výzkumníka.

Existuje mnoho metod, kterými lze faktorovou analýzu provádět. Jedna z nejjednodušších metod je například tzv. **centroidní metoda**, kterou lze použít i v případě, že není k dispozici počítač s příslušným programovým vybavením. Vzhledem k tomu, že tato metoda je v detailech velmi komplikovaná, nebudeme ji na tomto místě popisovat a omezíme se pouze na doporučení, aby ten, kdo hodlá faktorovou analýzu použít, vyhledal podrobnější poučení v odborné literatuře (například Blahuš, 1971, 1985).

Příklad 50

Reálné použití faktorové analýzy popíšeme na příkladu výzkumu (Vala, 2013), ve kterém se měřily prožitky, které mají žáci středního odborného učiliště při četbě poezie.

Žáci středního odborného učiliště (n = 1 088) posuzovali ve výzkumu předložené ukázky textů básní pomocí osmnácti sedmibodových škál (srov. oddíl 4.6):

1.	srozumitelná	1	2	3	4	5	6	7	nesrozumitelná
2.	komplikovaná	1	2	3	4	5	6	7	jednoduchá
3.	nechápu	1	2	3	4	5	6	7	chápu
4.	nejasná	1	2	3	4	5	6	7	jasná
5.	odkrytá	1	2	3	4	5	6	7	skrytá
6.	průhledná	1	2	3	4	5	6	7	neprůhledná
7.	líbí	1	2	3	4	5	6	7	nelíbí

8.	nečetl bych ji znovu	1	2	3	4	5	6	7	četl bych ji znovu
9.	špatná	1	2	3	4	5	6	7	dobrá
10.	ošklivá	1	2	3	4	5	6	7	krásná
11.	přitahuje mne	1	2	3	4	5	6	7	nepřitahuje mne
12.	příjemná	1	2	3	4	5	6	7	nepříjemná
13.	slabá	1	2	3	4	5	6	7	silná
14.	hluboká	1	2	3	4	5	6	7	mělká
15.	ostrá	1	2	3	4	5	6	7	tupá
16.	malá	1	2	3	4	5	6	7	velká
17.	účinná	1	2	3	4	5	6	7	neúčinná
18.	mocná	1	2	3	4	5	6	7	bezmocná

Získané výsledky byly podrobeny faktorové analýze pomocí statistického programu Statistica.cz, při které byla tzv. **metodou hlavních komponent** nejdříve vytvořena nerotovaná faktorová matice pro tři faktory (tab. 54).

***Tab. 54** Nerotovaná matice faktorových nábojů pro tři extrahované faktory*

Škála	Faktor 1	Faktor 2	Faktor 3
1	–0,724036	0,365668	0,216777
2	–0,565605	0,530528	0,227368
3	–0,742935	0,315637	0,234694
4	–0,739326	0,369210	0,221376
5	–0,494346	0,526306	0,315066
6	–0,532097	0,471539	0,261417
7	–0,804258	–0,038682	–0,239445
8	–0,762970	–0,121155	–0,275046
9	–0,755664	–0,070832	–0,238734
10	–0,749495	–0,056165	–0,332569
11	–0,733898	–0,183716	–0,325317
12	–0,708149	–0,012663	–0,372411
13	–0,435988	–0,556145	0,191667
14	–0,258576	–0,634811	0,315655
15	–0,262266	–0,541771	0,450498
16	–0,249042	–0,596249	0,338826
17	–0,635608	–0,438314	0,011964
18	–0,585244	–0,417971	0,069717

Protože výsledky faktorové analýzy v této podobě prakticky není možné interpretovat, byla následně provedena pravoúhlá rotace – typ **varimax normalizovaný**. Výslednou faktorovou matici prezentuje tabulka 55.

Tab. 55 *Rotovaná faktorová matice pro tři faktory*

Škála	Faktor SROZUMITELNOSTI	Faktor PŮSOBIVOSTI	Faktor HODNOCENÍ
1	**0,766609**	0,075452	0,333994
2	**0,781639**	-0,094947	0,181908
3	**0,755617**	0,129709	0,344790
4	**0,780194**	0,080241	0,341738
5	**0,781722**	-0,065192	0,072816
6	**0,742357**	-0,041575	0,144894
7	0,328135	0,152365	**0,758138**
8	0,234015	0,181924	**0,764586**
9	0,279717	0,161689	**0,727086**
10	0,237344	0,095711	**0,781029**
11	0,152068	0,191970	**0,786264**
12	0,219814	0,026892	**0,768950**
13	0,007284	**0,671293**	0,292288
14	-0,083304	**0,744180**	0,093496
15	0,045905	**0,750393**	-0,007097
16	-0,052982	**0,724792**	0,064869
17	0,107026	**0,544319**	0,537143
18	0,119516	**0,545196**	0,458855

Z matice je patrné, že pomocí faktorové analýzy je možné rozdělit použité posuzovací škály do tří kategorií podle toho, který aspekt (faktor) posuzované básně zachycují. U prvních šesti škál se objevují poměrně velké faktorové náboje u faktoru č. 1 (označeno tučným tiskem). Můžeme z toho vyvodit, že tyto škály zachycují (měří) „něco podobného". Pokud se podíváme, co společného tyto škály mají, zjišťujeme, že se zaměřují na srozumitelnost posuzovaných básní, proto byl faktor č. 1 označen jako **faktor srozumitelnosti**.

Podobně jsme postupovali při analýze dalších dvou faktorů. Škály č. 7–12 se zaměřují na to, jak báseň na žáky „působí" – faktor byl označen jako **faktor působivosti**.

Škály č. 13–18 spojuje to, že zachycují „hodnocení" básně, faktor byl pojmenován jako **faktor hodnocení**.

Použití faktorové analýzy v uvedeném případě umožnilo vytvořit nástroj, kterým lze poměrně spolehlivě měřit prožitky žáků při čtení poezie (Vala a Chráska, 2014).

Možnosti analýzy na PC
Statistica.cz: vícerozměrné průzkumné techniky → faktorová analýza

3.6 SHLUKOVÁ ANALÝZA

Shluková analýza je metoda, jejímž cílem je přiřadit jednotky analýzy (např. osoby, případy, události apod.) na základě podobnosti ke skupinám (shlukům, trsům). Přitom charakteristiky shluků ani jejich počet nejsou předem známy – musí být odvozeny z výzkumných dat. Mírou podobnosti (respektive nepodobnosti) jednotek analýzy bývá často korelace nebo tzv. **euklidovská distance *D***.

Poznámka: *Euklidovská distance D je definována jako vzdálenost dvou bodů (A, B) v n-dimenzionálním prostoru. Jestliže bod A má souřadnice $a_1, a_2, \ldots, a_n$ a bod B souřadnice $b_1, b_2, \ldots, b_n$, potom je jejich vzdálenost dána vztahem*

$$D = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2 + \ldots, + (b_n - a_n)^2} = \sqrt{\sum_{i=1}^{n} (b_1 - a_1)^2} \quad (91)$$

Ve shlukové analýze potom vystupují jednotlivé jednotky analýzy (např. osoby, případy apod.) jako body v n-dimenzionálním prostoru. Souřadnice těchto bodů jsou dány vlastnostmi, které jednotky analýzy mají (které tyto jednotky charakterizují).

Shluky jednotek analýzy (např. shluky osob, případů atd.) se mohou vytvářet na základě dvou, tří nebo i více (n) vlastností.

Shluková analýza se dnes prakticky provádí jen pomocí počítačů. Statistické programové systémy nabízejí řadu metod shlukové analýzy. Mezi často používané metody patří například **metoda hierarchického shlukování** nebo **metoda k-průměrů**. Princip těchto dvou metod shlukové analýzy nastíníme v následujících příkladech.

Příklad 51

Hierarchické shlukování

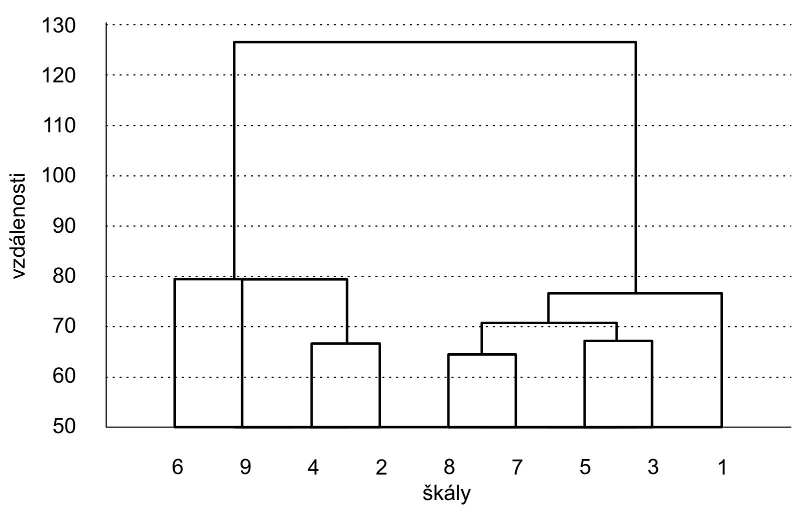
Ve výzkumu se měřily postoje studentů pedagogické fakulty k jejich budoucí učitelské profesi (Chráska, 2003). Postoje se měřily pomocí devíti škál sémantického diferenciálu. Ve výzkumu se (na vzorku 371 studentů) ověřovalo, jakou strukturu mají použité škály (tj. ověřovalo se, zda lze škály sémantického diferenciálu rozdělit do několika typů podle toho, jakou dimenzi postojů zachycují).

V uvedeném případě jsou jednotkami analýzy jednotlivé škály sémantického diferenciálu. Výsledkem provedené shlukové analýzy je tzv. D-matice (tab. 56), která prezentuje euklidovské vzdálenosti mezi jednotlivými škálami. Grafickým výstupem metody hierarchického shlukování je obrázek 20.

Z obrázku je patrné, že škály vytvářejí dva shluky, přičemž do prvního shluku patří škály 8, 7, 5, 3, 1 a do druhého shluku škály 6, 9, 4, a 2. Mezi škálami, které tvoří určitý shluk, jsou malé euklidovské vzdálenosti D, což znamená, že mají podobné vlastnosti (měří něco velmi podobného). Ve zmiňovaném výzkumu škály patřící do 1. shluku vyjadřují tzv. faktor hodnocení a škály patřící do 2. shluku vyjadřují tzv. faktor energie (Chráska, 2003)

Tab. 56 *D-matice*

Škály	1	2	3	4	5	6	7	8	9
1	0	154	80	150	77	139	90	90	155
2	154	0	148	67	140	79	140	142	85
3	80	148	0	145	67	134	73	76	149
4	150	67	145	0	137	80	135	138	79
5	77	140	67	137	0	127	70	75	142
6	139	79	134	80	127	0	132	131	92
7	90	140	73	135	70	132	0	65	139
8	90	142	76	138	75	131	65	0	143
9	155	85	149	79	142	92	139	143	0



Obr. 20 *Shluková analýza – metoda hierarchického shlukování*

Příklad 52

Shlukování metodou k-průměrů

Ve výzkumu, který byl zmíněn v předcházejícím příkladě, byla shluková analýza použita také k identifikaci dvou základních typů studentů podle postojů k učitelské profesi. V tomto případě byli jednotkami pro analýzu jednotliví studenti. Shluková analýza byla provedena metodou k-průměrů (podrobnosti o této metodě lze nalézt v literatuře – např. Osecká, 2001). U každého studenta byly postoje k učitelské profesi vyjádřeny pomocí faktoru hodnocení a faktoru energie. Výsledkem shlukové analýzy bylo nalezení tří shluků (skupin) studentů, které se vyznačují podobným faktorem hodnocení a faktorem energie (obr. 21).

1. shluk ■

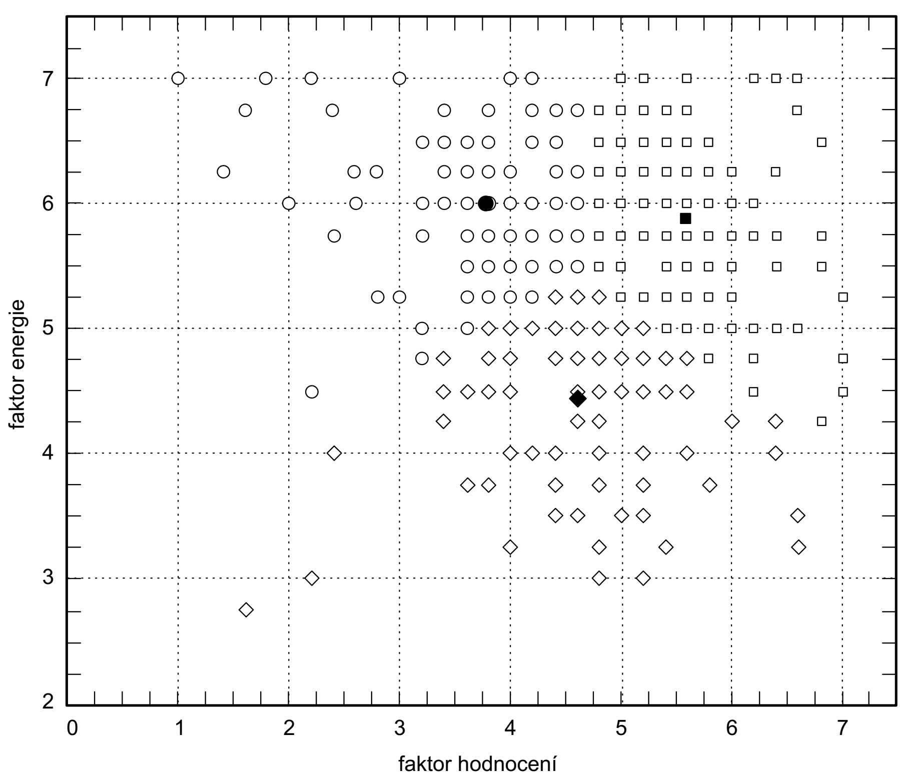
Tento typ studentů se vyznačuje průměrným hodnocením svého budoucího učitelského povolání a přesvědčením, že učitelské povolání je spojeno se značnou námahou, překonáváním překážek apod., tj. s velkým vydáváním energie.

2. shluk ●

Studenti patřící k tomuto typu jsou také přesvědčeni, že učitelské povolání je spojeno se značným vydáváním energie, ale vyznačují se poměrně vysokým hodnocením svého budoucího učitelského povolání.

3. shluk ◆

Tento typ studentů je charakteristický mírně nadprůměrným hodnocením budoucího učitelského povolání a tím, že učitelské povolání nepovažují za příliš namáhavé, „energeticky" náročné.



***Obr. 21** Shluková analýza (bodový diagram pro tři shluky)*

***Poznámka:** Studenti, patřící do jednotlivých shluků jsou v obrázku 21 označeni různými symboly. Střed každého shluku je označen příslušným symbolem s černou výplní.*

Možnosti analýzy na PC
Statistica.cz: vícerozměrné průzkumné techniky → shluková analýza

3.7 METAANALÝZA V PEDAGOGICKÉM VÝZKUMU

Jedním ze základních fenoménů dnešní doby je stále se zrychlující tempo narůstání nových vědeckých poznatků (často se v této souvislosti hovoří o informační explozi). Problém získávání syntetických informací z mnoha dílčích poznatků pomáhá řešit jedna z moderních oblastí statistiky, tzv. **metaanalýza**. Cílem metaanalýzy je **shrnout výsledky dvou nebo více empirických výzkumů, které se zabývají stejným nebo podobným problémem** (Hendl, 2004).

V českých pedagogických výzkumech se metaanalýza zatím prakticky neprovádí. Byla sice publikována řada prací, které uvádějí přehledy a výsledky realizovaných výzkumů jednotlivých problémů, ale většinou se jedná jen o tradiční **„narativní review“**. Tyto přehledy neužívají kvantitativní metody analýzy a mají charakter jen kvalitativního popisu a diskuse.

3.7.1 PŘÍKLADY JEDNODUCHÝCH STATISTICKÝCH METOD METAANALÝZY

Při metaanalýze je možno využít mnoha statistických metod a postupů, které se liší svojí náročností, ale také citlivostí a spolehlivostí. Vzhledem ke složitosti problematiky a rozsahu tohoto textu nebudeme jednotlivé metody popisovat a odkážeme čtenáře na velmi fundované pojednání o metaanalýze v práci J. Hendla (Hendl, 2004, s. 491–528).

Pro naznačení a přiblížení postupu při metaanalýze uvedeme jen dvě nejjednodušší metody.

3.7.1.1 Metoda „počítání hlasů“

U této metody postupujeme tak, že výzkumy, které hodláme podrobit metaanalýze, rozdělíme do tří kategorií:

- výsledek výzkumu pozitivní (statisticky významný);
- výsledek výzkumu negativní (statisticky významný);
- výsledek výzkumu statisticky nevýznamný.

Ta kategorie, ve které je největší četnost, je považována za kategorii nejlépe charakterizující všechny analyzované výzkumy.

Správnější by bylo použít (ve stejné situaci) znaménkový test. V tomto případě bychom označili pozitivní (statisticky významné) výsledky znaménkem + a negativní (statisticky významné) výsledky znaménkem – a dále bychom postupovali způsobem, který jsme popsali v oddíle 3.3.1.

3.7.1.2 Metoda sčítání z-skórů

Metodu sčítání z-skórů můžeme použít například v případech, kdy máme u jednotlivých výzkumů (jež hodláme v rámci metaanalýzy shrnout) informace o zjištěných hodnotách signifikace a o směrech odchylek dvou proměnných (Hendl, 2004).

Příklad 53

V osmi pedagogických výzkumech se ověřovala existence rozdílů mezi dvěma proměnnými (proměnná 1, proměnná 2). V šesti výzkumech se zjistilo, že rozdíl mezi proměnnou 1a proměnnou 2 je kladný, ve dvou výzkumech naopak rozdíl mezi proměnnou 1 a proměnnou 2 byl záporný. Ve všech dílčích výzkumech byla známá hodnota signifikance rozdílu mezi proměnnými. Signifikance (významnost) je pravděpodobnost nesprávného odmítnutí nulové hypotézy. Signifikance zjištěné v jednotlivých výzkumech jsou patrné z tabulky 57.

Tab. 57 *Signifikance rozdílů mezi proměnnými v jednotlivých výzkumech*

Číslo výzkumu	Znaménko rozdílu	Signifikance (oboustranný test)	Odpovídající z-skóre
1	+	0,0324	2,14
2	+	0,0414	2,04
3	–	0,0226	*–2,28*
4	+	0,0500	1,96
5	+	0,1052	1,62
6	+	0,2542	1,14
7	+	0,0026	3,00
8	–	0,0002	*–3,90*
			Σ 5,72

Při integraci výsledků v rámci metaanalýzy postupujeme tak, že nejdříve převedeme hodnoty signifikance u jednotlivých výzkumů na **normovanou normální veličinu** (příloha I), kterou v této statistické aplikaci označujeme písmenem *z* (z-skóre). Protože uvažujeme oboustranný test významnosti, musíme signifikance zjištěné u jednotlivých výzkumů vždy rozdělit na dvě poloviny a výsledek odečíst od 1. Tím dostaneme pravděpodobnost, pro kterou v tabulkách distribuční funkce normovaného normálního rozdělení nalezneme příslušnou hodnotu normované normální veličiny. Z dílčích hodnot z-skóre pro jednotlivé výzkumy potom vypočítáme hodnotu *z*, která vypovídá o celkové významnosti pro všechny integrované výzkumy. Výpočet provádíme podle vztahu

$$z = \frac{\sum z_i}{\sqrt{k}} \tag{92}$$

kde z je celkové z-skóre pro všechny integrované výzkumy, z_i jsou hodnoty z-skóre pro jednotlivé výzkumy a k je počet analyzovaných výzkumů.

Při výpočtu výrazu $\sum z_i$ je nutné respektovat znaménka jednotlivých dílčích z (u výzkumů č. 3 a 8 jsou záporné hodnoty).

Po dosazení hodnot do vzorce (92) dostáváme

$$z = \frac{\sum z_i}{\sqrt{k}} = \frac{(2{,}14 + 2{,}04 - 2{,}28 + 1{,}96 + 1{,}62 + 1{,}14 + 3{,}00 - 3{,}90)}{\sqrt{8}} = 2{,}02$$

Vypočítanou hodnotu srovnáváme s kritickou hodnotou $u_{0,05} = 1{,}96$ (pro hladinu významnosti 0,05, uvažujeme oboustranný test). Výsledek můžeme interpretovat tak, že metaanalýza výsledků v analyzovaných výzkumech prokázala, že mezi danými proměnnými existují statisticky významné rozdíly.

3.8 ZPRACOVÁNÍ A ANALÝZA DAT S VYUŽITÍM POČÍTAČE

Všechny statistické metody a postupy popisované v této publikaci jsou prezentovány tak, aby je bylo možné realizovat i bez použití specializovaného počítačového softwaru.

V současné době je samozřejmě zcela běžné, že výzkumná data se vyhodnocují pomocí počítačových programů. Analýza dat s využitím počítače je jednak mnohonásobně rychlejší než při ručním zpracování, ale také je v tomto případě podstatně menší riziko chyb (např. chyb, které vznikají při zaokrouhlování hodnot). Zpracování dat bez použití počítače je nemyslitelné například v případech, kdy máme zpracovat velké množství dat, nebo v případech, kdy chceme ke statistické analýze použít velmi náročných statistických procedur (např. analýza rozptylu, faktorová analýza apod.).

Na tomto místě je třeba zdůraznit, že snadná dostupnost a zdánlivá jednoduchost provedení statistických analýz pomocí počítače může v některých případech svádět k aplikaci procedur bez náležitého pochopení jejich smyslu, podstaty a podmínek, za nichž jsou použitelné. Mechanická aplikace statistických postupů bez odpovídajících vědomostí může tak vést ke zkresleným nebo chybným závěrům.

Dobrou cestou k pochopení principů statistických procedur je sledovat a promýšlet všechny dílčí kroky s „kalkulačkou v ruce“. Jsme přesvědčeni, že smysluplné využití statistických programů má smysl pouze za předpokladu pochopení alespoň základních principů jednotlivých procedur a jejich logiky.

K běžně dostupným programům obsahujícím základní statistické funkce patří Excel, který je součástí aplikací Microsoft Office. Nabízí jednoduché uživatelské prostředí, které umožňuje zaznamenávat výzkumná data, analyzovat je na popisné i vztahové úrovni a vytvářet grafy. V pedagogických výzkumech bývá nejčastěji využíván jako prostředí pro zápis získaných dat. Tabulka 58 prezentuje obvyklý způsob zápisu dat do tabulky v Excelu.

V řádcích tabulky jsou většinou uvedeni respondenti (případně zkoumané situace), ve sloupcích jednotlivé proměnné (údaje o respondentech).

Tab. 58 *Forma záznamu dat do tabulkového procesoru*

	A	B	C	D	E	G	H	I
1	Číslo respondenta	Pohlaví	Věk	Délka praxe	Typ školy	P1	P2	P3
2	1	muž	56	30	ZŠ	3	2	3
3	2	muž	51	27	ZŠ	3	4	4
4	3	žena	26	2	ZŠ	1	5	1
5	4	muž	29	5	SOŠ	2	1	2
6	5	žena	59	33	G	1	2	2
7	6	žena	38	11	L	2	4	5
8	7	žena	55	30	ZŠ	1		
9	8	žena	52	30	G	4	2	2
10	9	muž	50	27	SOŠ	2	2	2
11	10	muž	56	31	G	5		1
12	11	žena	43	21	L	2	2	1
13	12	muž	59	35	ZŠ	2	3	5
14	13	žena	37	13	SOŠ	2		2
15	14	žena	49	25	G	2	2	1

Každý řádek tabulky (2–15) obsahuje údaje o jednom respondentovi (případně o jedné situaci nebo události). V prvním řádku tabulky jsou uvedeny názvy proměnných (v uvedeném příkladě je to pohlaví respondentů, věk respondentů, délka praxe, typ školy a informace o odpovědích ve třech položkách dotazníku (respondenti vybírali odpovědi na škále od 1 do 5). Data je možno do tabulky zapisovat pomocí čísel nebo textu. Prázdné buňky znamenají, že údaj není znám (např. proto, že respondent danou položku v dotazníku vynechal). Excel se výborně hodí na základní práci s daty (zaznamenání, překódování, doplnění součtů, průměrů atd.). Přestože program Excel není primárně určen pro provádění statistických funkcí, základní z nich realizovat umožňuje. Statistické funkce jsou dostupné v nabídce „Vložit funkci → Statistické → Název funkce". Program Excel je velmi vhodný také pro tvorbu grafů („Průvodce grafem → Typ grafu → Název grafu").

Na trhu je dostupná celá řada počítačových programů zaměřených přímo na statistiku. K nejpoužívanějším patří v současnosti Statistica. Tento program nabízí prostředí, které je primárně určeno pro statistickou analýzu dat. V programu je možné otevřít data, která byla původně zaznamenána v Excelu. To je výhodné například v případě, kdy nejsme vlastníky licence programu a potřebujeme získaná data analyzovat na jiném (např. univerzitním) počítači, kde je tento program k dispozici. V nabídce lze najít velmi široké spektrum operací, které se může mírně lišit podle verze programu. V zásadě se však jedná o stejný systém. V nabídce „Statistiky → Základní statistiky" nalezneme výběr všech popisných statistik, přičemž stačí jen vybrat jednu či více proměnných, které chceme do analýzy zahrnout. Veškeré statistické procedury jsou dále rozděleny v nabídce „Statistiky → Základní statistiky“, „Vícenásobná regrese“, „Anova“, „Neparametrické statistiky“ atd. Program vytváří jako výstup z analýzy přehledné tabulky, které po přidání do protokolu („Domů → Přidat do protokolu“) můžeme upravovat, kopírovat do Excelu, do Wordu nebo ukládat přímo v protokolu. Program ke každé analýze přímo nabízí graf, který je možné si upravit podle vlastních požadavků. Je také možné vytvářet grafy přímo bez předchozí analýzy, a to v nabídce „Grafy“.

Dalším často používaným programem, který byl vytvořen speciálně pro potřeby analýzy dat ve společenských vědách, je SPSS (*Statistical Package for the Social Sciences)*. Nabízí podobné uživatelské prostředí jako Statistica.

4. METODY SBĚRU DAT V KVANTITATIVNĚ ORIENTOVANÝCH PEDAGOGICKÝCH VÝZKUMECH

Předmětem této kapitoly bude objasnění základních postupů, jimiž se realizuje měření a získávání dat v pedagogických výzkumech. Tyto metody označujeme také jako metody sběru dat.

4.1 PEDAGOGICKÉ POZOROVÁNÍ

Pedagogické pozorování je nejstarší a nejrozšířenější metodou získávání dat o pedagogické realitě. Pedagogické pozorování bývá definováno jako „sledování smyslově vnímatelných jevů, zejména chování osob, průběhu dějů aj." (Průcha, Walterová a Mareš, 2001). Někteří autoři považují pozorování za nejdůležitější a nezastupitelný způsob shromažďování materiálu při studiu pedagogické reality. Jiní autoři ale správně upozorňují na to, že běžné pedagogické pozorování se mnohdy nedostává nad jevovou stránku, což poněkud snižuje jeho heuristický význam (Tollingrová, 1971; Travers, 1969).

V literatuře lze nalézt různé přístupy ke klasifikaci **pozorování**. Vzhledem k nárokům na čas se často rozlišuje **pozorování krátkodobé** a **dlouhodobé**. Při vědeckých pozorováních jde většinou o důkladné a dlouhodobé sledování určitého jevu, i když hranice mezi krátkodobým a dlouhodobým pozorováním není nijak ostře vymezena. Jako krátkodobá se zpravidla označují ta pozorování, která netrvají déle než jednu vyučovací jednotku. Krátkodobých pozorování se využívá většinou k praktickým účelům v každodenní praxi. Dlouhodobá, důkladná pozorování, která mohou probíhat na týchž subjektech například po řadu let, jsou označována jako pozorování longitudinální.

Podle toho, kdo pozorování provádí, lze rozlišit sebepozorování – **introspekci** a pozorování jiných – **extrospekci**.

Podle toho, zda se při pozorování pozorovatel setkává přímo s předmětem pozorování či nikoli, hovoříme o pozorování ve vlastním smyslu slova – **vlastním (přímém) pozorování** a o pozorování v nevlastním smyslu slova – **nevlastním pozorování**. Při vlastním pozorování se pozorovatel setkává přímo s předmětem pozorování. Při nevlastním pozorování pracuje s různými výpověďmi o předmětu zkoumání, ať již v podobě řeči psané nebo mluvené. Pro označení stejných druhů pozorování užívají jiní autoři termínů **pozorování přímé** (neboli bezprostřední) a **pozorování nepřímé** (neboli zprostředkované).

Často se také rozlišuje **pozorování standardizované** a **nestandardizované**. Jako standardizované pozorování bývá označována činnost spočívající v záměrném, cílevědomém, systematickém a relativně objektivním sledování smyslově vnímatelných jevů, které nebyly vyvolány zásahem pozorovatele (Mareš, 1983). Běžná školní pozorování, která jsou vždy ve větší či menší míře poznamenána intuitivním přístupem a subjektivitou, bývají označována jako pozorování nestandardizovaná. Pro standardizovaná pozorování je typické používání speciálních observačních technik, které umožňují snížit podíl intuice a subjektivity na únosnou míru.

4.1.1 VLASTNOSTI DOBRÉHO PEDAGOGICKÉHO POZOROVÁNÍ

Na dobré pedagogické pozorování jsou kladeny určité požadavky. V literatuře se často uvádějí například následující čtyři:

- specifikace objektu pozorování (odpovídá na otázku „co se má pozorovat?“);
- zaměřenost pozorování na cíl (odpovídá na otázku „co je třeba zjistit?“);
- organizovanost pozorování (odpovídá na otázku „jak toho dosáhnout?“);
- přesný záznam pozorování (odpovídá na otázku „jak to zachytit?“).

O pozorování můžeme uvažovat také jako o metodě měření pedagogické reality a požadovat, aby splňovalo ty požadavky, které se kladou na každé dobré měření. V tomto případě se požaduje, aby pozorování bylo především dostatečně validní a reliabilní.

Pozorování má dobrou **validitu** tehdy, jestliže se pozoruje skutečně to, co se pozorovat má. U pozorování pedagogických jevů bývá validita často problémem. Jestliže totiž chceme jevy pedagogické reality měřit, jsme vždy nuceni k určitému zjednodušování (simplifikaci). Při tomto zjednodušování se snadno ocitneme v situaci, kdy to, co pozorujeme, nemusí být pro daný jev podstatné, a že pak vlastně pozorujeme něco jiného, než bylo záměrem.

Dobrou **reliabilitu** má pozorování tehdy, jestliže není ve větší míře zatíženo chybami pozorování, tzn. tehdy, jestliže spolehlivě a přesně zachycuje pozorované jevy. Reliabilita pozorování souvisí s validitou, ale není s ní totožná. Zatímco při posuzování validity se uplatňuje celkové chápání pozorovaného jevu, jde u reliability o otázku do značné míry technickou. Má-li mít pozorování dobrou validitu, musí mít vysokou reliabilitu. Má-li však pozorování vysokou reliabilitu, nemusí mít ještě vysokou validitu.

Přesto, že pozorování je základem mnoha pedagogických výzkumů, náslechů, hospitací a inspekcí na školách, až donedávna se v pedagogice nezkoumalo, nakolik jsou výsledky pozorování spolehlivé a přesné (reliabilní), a tudíž nakolik jsme oprávněni činit z nich jakékoli závěry. Teprve v posledních letech se projevuje v naší pedagogice snaha exaktně určovat stupeň reliability standardizovaného pozorování (např. Mareš, 1983; Chráska, 1996).

Reliabilita pozorování závisí (Mareš, 1983) na třech hlavních činitelích:

- na samotném pozorovateli;
- na použité pozorovací technice;
- na okolnostech pozorování.

K exaktnímu vyjádření stupně reliability pozorování se užívá koeficientu reliability. Byla navržena řada postupů, kterými lze tento koeficient stanovit – přehled o použitelných metodách výpočtu podává například studie J. Mareše (1983) nebo studie M. Chrásky (1996, 2003).

4.1.2 SUBJEKTIVNÍ FAKTORY PŮSOBÍCÍ PŘI POZOROVÁNÍ

Má-li být pozorování spolehlivé a přesné, tj. reliabilní, musí být v maximální možné míře objektivní. Malá objektivita je největší slabinou pedagogického pozorování. Subjektivita pozorování má řadu příčin, z nichž některé důležité uvedeme.

Haló efekt

Haló efekt je tendence vnímat jedince pod vlivem celkového (příznivého nebo nepříznivého) dojmu, který si obvykle vytváříme hned na začátku seznámení s hodnoceným člověkem díky nějaké jeho nápadné a často i nedůležité vlastnosti. Jestliže náš celkový dojem je příznivý, zvětšuje to intenzitu kladných vlastností a snižuje intenzitu vlastností záporných a naopak.

Logická chyba

Podstatou této chyby je sklon pozorovatelů hodnotit například povahové vlastnosti lidí tak, jak se to jim samým zdá logické. Pozorovatelé tedy „logicky" předpokládají u pozorovaných osob jisté vlastnosti.

Předsudky

Předsudky jsou například nekritické přejímání názorů na žáka od jiných učitelů nebo nepodložená představa o tom, že děti z určitých společenských vrstev automaticky musí mít určité vlastnosti apod.

Stereotypizace a analogie

Jedná se o sklon vytvářet si určitá schémata obrazů o osobnosti, která mechanicky aplikujeme. O dítěti s určitými projevy se například domníváme, že má stejné vlastnosti jako dítě, které má projevy analogické, aniž přihlížíme k víceznačnosti projevů.

Tradice

Na hodnocení lidí má značný vliv i tradice, zakotvená například v různých rčeních a příslovích. Tak toho, kdo hodně mluví, považujeme někdy neprávem za povrchního, lehkomyslného, nespolehlivého. Podobně hodnotíme žáky s brýlemi a vážnou tváří jako nadané atd.

Figura a pozadí

Tento obecně psychologický mechanismus odvozený z celostních zákonitostí vnímání způsobuje, že například na pozadí hlučné, nepěkně zařízené místnosti jsou naše dojmy o zkoumané osobě nebo věci vcelku nepříznivější. Spojujeme-li určitou osobu s osobou nebo osobami nám sympatickými, vnímáme ji též jako sympatickou.

Aktuální psychický stav

Na posuzování pozorované skutečnosti má nemalý vliv aktuální psychický stav pozorovatele. Jsme-li rozladěni, rozčileni, smutní nebo veselí, jsou naše posudky buď nepříznivější, nebo příznivější.

Tendence k průměru

Jedná se o sklon přisuzovat pozorovaným jevům spíše střední intenzitu než intenzitu vysokou nebo nízkou. Příčinou může být například nejistota pozorovatele, jeho přílišná opatrnost nebo nedostatečná erudice, vliv má i celková zralost osobnosti a bohatost životních zkušeností.

Kontrast

Jde o tendenci podhodnocovat ty vlastnosti žáka, o nichž se pozorovatel domnívá, že v nich on sám vyniká, a naopak.

Shovívavost pozorovatele

Tento faktor způsobuje, že pozorované skutečnosti jsou hodnoceny celkově mírněji, než si zaslouží. Někteří autoři hovoří v této souvislosti o **chybě lenienční**.

4.1.3 TECHNIKY STANDARDIZOVANÉHO POZOROVÁNÍ

Cílem pozorování v pedagogickém výzkumu je nejenom určitý pedagogický jev zachytit a popsat, ale i odhalit a vysvětlit příčiny jeho vzniku a vývoje. Tyto dva spolu úzce související cíle lze charakterizovat otázkami „co" (zachycení a deskripce jevu) a „proč" (vysvětlení příčin, souvislostí). Je pochopitelné, že konečným cílem je zodpovězení otázky „proč", ovšem bez náležitého zodpovězení otázky „co" by naše závěry byly založeny pouze na dohadech a intuici. Je proto nezbytné, aby každé pozorování, které si činí nárok na vědeckost, začínalo vždy přesnou deskripcí pedagogické reality. Pedagogické zkoumání, které uspokojivě neodpovědělo na otázku „co", a přesto se pokouší o zodpovězení otázky druhé (hledá souvislosti, příčiny, vyslovuje nejrůznější soudy atd.), je zavádějící a odsouzeno k nezdaru. Byla vyvinuta řada observačních technik, které umožňují relativně objektivní popis a měření jevů pedagogické reality. Některé z těchto technik jsou pro svoji náročnost vhodné pouze pro pedagogický výzkum, jiné jsou vhodné i pro použití v pedagogické praxi.

4.1.3.1 Technika A. A. Bellacka

Zajímavým a podnětným způsobem se pokusil snížit podíl intuice na výsledcích pedagogického pozorování profesor kolumbijské univerzity A. A. Bellack. Vytvořil originální znakový jazyk, kterým lze „relativně objektivně zachytit dění ve třídě při vyučování a zapsat je simultánně, skoro jako orchestrální partituru" (Tollingerová, 1971).

Vyučovací proces považuje Bellack za komplex interakcí mezi učitelem a žáky, probíhajících podle předem zadaných pravidel, která jsou složitě a vzájemně podmíněna. Výuková interakce je u Bellackovy techniky rozebírána až na elementární části – pedagogické aktivity. Každá pedagogická aktivita je pak popisována *z* hlediska osmi základních kategorií:

1. mluvčí nebo subjekt jednání;
2. typ pedagogické aktivity;
3. učivo a jeho charakteristika;
4. logické operace s tímto učivem;
5. počet řádků záznamu v bodě 3 a 4 (1 řádek = 50 úhozů záznamu);
6. způsob výuky;
7. logické operace, které se na tom podílejí;
8. počet řádků záznamu v bodě 6 a 7.

Každá *z* těchto základních kategorií se dále člení, výsledkem čehož je symbolický deskriptivní jazyk celkem s 54 znaky. Pro ilustraci uvádíme příklad formálního zápisu jedné pedagogické aktivity (Tollingerová, 1971).

Učitel: Které země se spojily k vytvoření společného trhu? Jano!

Nositelem této aktivity je učitel (znak T), z hlediska typu pedagogické aktivity jde o kladení požadavku (znak SOL). Z hlediska 4. kategorie jde o vyjmenování – konstatování (znak FAC). Délka této otázky nepřesahuje 50 úhozů (1 řádek). Z hlediska 6. kategorie slůvko „Jano“ dává pokyn (znak PRC), smyslem tohoto pokynu je navození činnosti (znak PRF), délka nepřesahuje opět 1 řádek. Formální zápis analyzované aktivity vypadá potom takto:

M # 1: /T/SOL/PFT/FAC/1/PRC/PRF/1/

Symbol M # 1 je označení aktivity (její pořadové číslo), následující znaky, které jsou odděleny šikmými čarami, označují jednotlivé základní kategorie (v pořadí od kategorie č. 1 po kategorii č. 8). Podrobnější a přesnější popis Bellackovy techniky uvádí například práce Tollingerové (1971).

Význam Bellackovy techniky spočívá v tom, že umožňuje pomocí vytvořeného symbolického jazyka zachytit to, co se děje ve třídě, když probíhá vyučovací proces. Technika umožňuje zodpovědět řadu závažných otázek, které se týkají průběhu reálného vyučovacího procesu (např. jaká je role učitele a žáka ve vyučovacím procesu, kdo je nositelem pedagogického dění ve třídě, jakým způsobem jsou nejčastěji formulovány dotazy apod.).

Bellackova metoda má význam především v tom, že ukazuje jeden z možných přístupů k minimalizaci chyb pedagogického pozorování. V pedagogických výzkumech je možné principu této metody využít pro vytvoření nástrojů pro pozorování i jiných stránek edukační reality.

4.1.3.2 Technika frekvenční a sekvenční analýzy

U této pozorovací techniky se dění ve třídě při vyučování popisuje pomocí určitého počtu předem stanovených kategorií. Často se používá například deskriptivně analytického systému N. A. Flanderse, který umožňuje veškeré dění ve třídě přiřadit k jedné z deseti následujících kategorií (Travers, 1969):

1. konstruktivní projev sympatií, akceptování žákových citů;
2. souhlas, pochvala, povzbuzení;

3. akceptování nebo rozvíjení myšlenek obsažených v žákově odpovědi;
4. dotazování;
5. výklad; poskytování informací;
6. pokyny, příkazy;
7. nesouhlas, kritika;
8. odpověď nikoli spontánní;
9. spontánní odpovědi;
10. ticho, neverbální činnost, zmatek.

Při použití této pozorovací techniky se nejdříve pořídí přesný záznam vyučovací hodiny (např. videozáznam nebo magnetofonová nahrávka). Záznam se potom rozdělí na krátké časové úseky (většinou třísekundové intervaly). Vycvičený pozorovatel sleduje průběh vyučovací hodiny a podle stanovených pravidel simultánně kóduje činnost učitele i žáků pomocí uvedených deseti kategorií. Jde v podstatě o převedení dění ve třídě do jakéhosi umělého jazyka (soustavy numerických kódů), který popisuje interakci mezi učitelem a žáky.

Příklad 54

Uvedeme příklad písemného záznamu části vyučovací hodiny, která byla analyzována pomocí Flandersova systému kategorií (Hrabal, 1988). V ukázce jsou vyznačeny (•) třísekundové intervaly a čísly v závorkách příslušné kategorie Flandersova systému.

Učitel: • *Kdo z vás mi dovede odpovědět, (4)* • *který byl první státní útvar na (4)* • *našem území? Kdy se poprvé sjednotili (4)* • *pod vládou někoho lidé (4)* • *žijící na našem území? (4)* • *Přemýšlejme, vraťme se trošku zpět. (4)* • *Tak co myslíš, Bílku? (4)* •

Žák: • *Byla to Velká Morava, když barbaři proti Velké Moravě a... (8)* •

Učitel: • *Tak si sedni, to není v (6)* • *pořádku! Já myslím ještě jiný státní útvar, (7)* • *který to byl před Velkou Moravou! Jirko! (4)* •

Žák: • *Bylo to roku 623 (8) a Sámo spojil Slovany. (8)* •

Učitel: • *Ano ... (2)* •

Žák: • *… když bojovali proti Avarům. (8)* •

Učitel: • *Výborně. (2)* • *Tedy jak jsme nazývali tu první říši, byla to... (4)* •

Žák: • *Sámova říše. (8)* •

Učitel: • *Ještě jednou, v kterém roce vznikla? (4)* • *Pěkně nahlas odpovídáme! Vláďo! (6)* •

Žák: • *Bylo to v roce 623. (8)* •

Učitel: • *Dobře. (2)* • *Řekněte mi, co bylo důvodem nebo příčinou toho, že se sjednotili Slované (4)* • *tenkrát dohromady? Proč se potřebovali (4)* • *sjednotit? Magdo!* (4) •

Výsledkem činnosti pozorovatele je posloupnost numerických kódů, které popisují výskyt jednotlivých kategorií činností. Uvedená ukázka, která odpovídá 75 sekundám výuky, je zachycena zápisem:

4 4 4 4 4 4 4 8 6 7 4 8 8 2 8 2 4 8 4 6 8 2 4 4 4

Získané údaje se tabelují do tzv. **interakční matice** ($n = m = 10$). Jednotlivá číselná označení se zaznamenávají do matice tak, jak za sebou následují. Následoval-li například po odpovědi žáka (8) souhlas učitele (2), zaznamenáme to čárkou v 8. řádku a ve 2. sloupci. Jestliže po otázce učitele (4) následuje odpověď žáka (8), zapíšeme tuto skutečnost čárkou ve 4. řádku a 8. sloupci atd. Počet čárek v každém poli matice potom vyjádříme číslem. Tabulka 59 uvádí interakční matici pro uvedenou ukázku části vyučovací hodiny.

Tab. 59 *Interakční matice*

Kategorie	1	2	3	4	5	6	7	8	9	10	Četnost	% času
1											0	0
2				2				1			3	12
3											0	0
4				8		1		3		1	13	52
5											0	0
6							1	1			2	8
7				1							1	4
8		3		1		1		1			6	24
9											0	0
10											0	0
∑	0	3	0	12	0	2	1	6	0	1	25	100

Z údajů obsažených v interakční matici lze získat mnoho zajímavých a důležitých informací nejen o výskytu jednotlivých kategorií činností ve vyučovací hodině, ale i o jejich proporcích, posloupnostech a dalších vlastnostech.

Jestliže sečteme čísla v řádcích interakční matice, získáme informaci o četnostech a proporcích výskytu jednotlivých kategorií ve vyučovací hodině. Pro názornost je možno údaje o výskytu kategorií převést na **procenta času**. Uvedený rozbor označujeme jako **frekvenční analýzu**.

Při sekvenční analýze se zajímáme o to, které typy přechodů mezi kategoriemi činností se ve vyučovací hodině (nebo její části) vyskytovaly. V uvedeném příkladě interakční matice například zjišťujeme, že největší četnost má kategorie „dotazování" (4). Po této kategorii následuje ve většině případů stejná kategorie (pole 4–4), z čehož lze usuzovat na to, že otázky učitele jsou spíše delší (trvají nejméně 2 časové intervaly). Druhou nejfrekventovanější kategorií, která následuje po „dotazování", je „odpověď žáka" (8) atd. Pro každý typ přechodu mezi kategoriemi lze vypočítat pravděpodobnost, se kterou tento přechod ve vyučovací hodině (nebo její části) nastává. Například pravděpodobnost přechodu 4–4 v uvedené ukázce by byla

$$P(4-4)=\frac{8}{13}=0{,}62$$

Pravděpodobnost přechodu 8–2

$$P(8-2)=\frac{3}{6}=0{,}50 \text{ atd.}$$

Výzkum J. Mareše (Mareš, 1977) se například zabýval (kromě jiného) otázkou, jaký je poměr mezi verbální činností učitele, verbální činností žáka a ticha (neverbálními činnostmi) u začínajících učitelů fyziky. Bylo zjištěno, že začínající učitel mluví v hodinách průměrně 39 % času, žáci mluví 10 % času a zbytek, tj. 51 % času zabírá ticho (respektive neverbální činnosti). V uvedeném výzkumu byly zjišťovány také pravděpodobnosti, se kterými po sobě následují jednotlivé kategorie činností. Například objevilo-li se u začínajícího učitele ticho (pauza), následovalo s pravděpodobností 0,78 opět ticho. Z tohoto zjištění lze vyvodit závěr, že neverbální činnosti se u začínajících učitelů vyskytují ve větších celcích.

4.1.3.3 Rating v pedagogickém pozorování

Hovoříme-li o ratingu, máme obyčejně na mysli řadu různých technik a postupů, kterým je společné to, že určité kvalitě jevu je přiřazována kvantitativní hodnota na škále. Toto přiřazování provádí přitom kvalifikovaný odborník – expert.

Termínu „rating" se používá většinou pro označení metody, ale někdy také pro označení výsledků posuzování.

Pojem škála definuje P. N. Kerlinger (1972) takto: „Škála je souborem symbolů nebo čísel, a to tak konstruovaných, že lze symboly nebo čísla přiřadit podle pravidla jedincům (nebo jejich aktům chování), na které se škála aplikuje." Posuzovací škála tedy představuje systém, který je nástrojem kvantifikace pozorovaných jevů. F. N. Kerlinger (1972) uvádí tři hlavní druhy posuzovacích škál:

- kategoriální posuzovací škály;
- numerické posuzovací škály;
- grafické posuzovací škály.

Kategoriální posuzovací škály

Posuzovateli se předkládá několik uspořádaných kategorií, z nichž má vybrat tu, která nejlépe vystihuje pozorovanou skutečnost. Příklad:

Byl učitel pohotový?
- velmi pohotový;
- pohotový;
- nepohotový;
- velmi nepohotový.

Určitým problémem při užívání kategoriálních posuzovacích škál je, že různí posuzovatelé mohou předkládané slovní charakteristiky kategorií chápat odlišným způsobem. Chceme-li u tohoto typu škál získat věrohodné výsledky, je třeba, aby všichni posuzovatelé chápali sémantiku kategorií pokud možno stejným (nebo velmi podobným) způsobem.

Numerické posuzovací škály

U numerické posuzovací škály má posuzovatel před sebou řadu čísel, jež odpovídají různým mírám posuzované vlastnosti s tím, že jednu má označit. Vedle čísel se často uvádějí i popisy jednotlivých měr posuzované vlastnosti. Numerickou posuzovací škálu získáme z kategoriální připojením čísel před kategorie, například:

1	2	3	4
velmi pohotový	pohotový	nepohotový	velmi nepohotový

Někdy se u škály uvádějí jen krajní body, například:

velmi pohotový 1 2 3 4 velmi nepohotový

U posuzovacích škál se doporučuje používat extrémního označení krajních bodů – tzv. **zakotvení škály**, a to i v případech, že se ve skutečnosti zřídka vyskytují. Tyto extrémní body podněcují diferencovaný přístup k ostatním bodům škály.

Numerické posuzovací škály mohou být buď jednostranné (řada čísel tvořících škálu má na jedné straně maximum a na druhé minimum), anebo mohou mít uspořádání bipolární (s nulovým bodem uprostřed).

Optimální počet stupňů škály závisí na konkrétních podmínkách, ale většinou se doporučuje škála čtyř- až devítistupňová. Při měření frekvence nebo kvality pedagogických jevů většinou vystačíme se škálami čtyř- až pětistupňovými, při měření postojů je lépe použít škál šesti- až sedmistupňových. Někdy se diskutuje i problém, zda použitá škála má obsahovat sudý, nebo lichý počet stupňů. Škály s lichým počtem stupňů umožňují neutrální (někdy také alibistická) rozhodnutí, škály se sudým počtem stupňů tuto „sémantickou nulu" nemají, a posuzovatel je proto nucen k rozhodnutí, které konverguje k jednomu z krajních bodů škály. Používání škál s lichým počtem stupňů je však častější.

Grafické posuzovací škály

První dva druhy posuzovacích škál (kategoriální a numerická) byly škály alfanumerické. Kromě těchto škál existují ještě škály grafické, u nichž se míry hodnocení vyjadřují v názorné formě. U těchto škál se většinou kombinují čáry nebo proužky s popisy vlastností. Mají celou řadu forem a podob. Mohou se předkládat v poloze vodorovné (častější), ale také v poloze svislé („teploměr"). Svislá škála se lépe popisuje, ale někdy může její horní bod sugerovat dojem, že jde o „lepší pól".

Doporučuje se, aby úsečky grafických škál byly dlouhé asi 12–15 mm a aby byly nakresleny bez přerušení, čímž se podporuje představa kontinuity.

Příklad grafické posuzovací škály:

Kladl učitel žákům během vyučovací hodiny otázky?

vůbec ne *velmi často*

Kromě již uvedených základních typů škál se používají i mnohé další druhy. Například V. Břicháček (1978) hovoří o tzv. **standardních posuzovacích škálách**, **kumulativních posuzovacích škálách** a **posuzovacích škálách s vynucenou volbou mezi variantami**.

Standardní posuzovací škály

Standardní posuzovací škály pracují s tzv. standardy, což jsou vlastně přesně definované příklady na jednotlivé škálové stupně. Pozorovatelé musí být se všemi standardy nejdříve náležitě seznámeni a teprve potom ke standardům přiřazují ty vlastnosti, které pozorují.

Kumulativní posuzovací škály

U těchto posuzovacích škál se daný jev zachycuje pomocí součtu bodů, který je získán z několika položek. Nejčastější formou těchto škál jsou různé **zaškrtávací listy**, na nichž je například seznam možných (kladných i záporných) vlastností posuzované osoby (např. učitele). Úkolem posuzovatele je označit ty vlastnosti na seznamu, které dobře charakterizují posuzovanou osobu. Pozitivní odpovědi se obvykle skórují +1, negativní –1, neutrální 0. Kumulativní posuzovací škály mají řadu předností. Dobře se předkládají, práce s nimi je pro posuzovatele snadná a pochopitelná. Hlavním problémem u kumulativních posuzovacích škál bývá to, že nemusí vždy tvořit jedinou posuzovací dimenzi. Pokud by tomu tak nebylo, dochází ke kumulování dat, která nejsou slučitelná, což značně zkresluje výsledky i jejich interpretaci.

Posuzovací škály s vynucenou volbou mezi variantami

V případě těchto posuzovacích škál jde o sofistikovanou a poměrně složitou posuzovací techniku, u níž na posuzovateli žádáme, aby označil, kterou vlastnost má posuzovaný objekt v největší míře, kterou v menší a kterou například v nejmenší míře. Ptáme se tedy vlastně na pořadí typických vlastností objektu. V. Břicháček (1978) uvádí na tento druh škály následující příklad:

Jistý výzkum sledoval, jak se v práci učitele uplatňuji čtyři faktory:

- znalost a utřídění látky;
- přiměřený poměr k žákům;
- přiměřený plán a průběh vyučování;
- radost z práce s dětmi.

Byla vytvořena škála s deseti položkami, ve kterých byla uvedena vždy čtyři tvrzení odpovídající výše uvedeným faktorům. Úkolem posuzovatele bylo sestavit v každé položce pořadí tvrzení podle toho, jak dobře popisují určitého učitele. Výsledky škálování umožnily v uvedeném případě sestavit vlastně čtyřrozměrný profil daného učitele. Posuzovací škály s vynucenou volbou mezi variantami jsou (ve srovnání s jinými ratingy) značně spolehlivé, vyžadují však důkladnou a pečlivou přípravu.

Rating představuje velmi slibnou pozorovací techniku, kterou lze užít k relativně objektivnímu pozorování chování (ať již aktuálního nebo zapamatovaného). Velkou předností ratingu je, že umožňuje exaktní stanovení stupně reliability výsledků (Chráska, 1986).

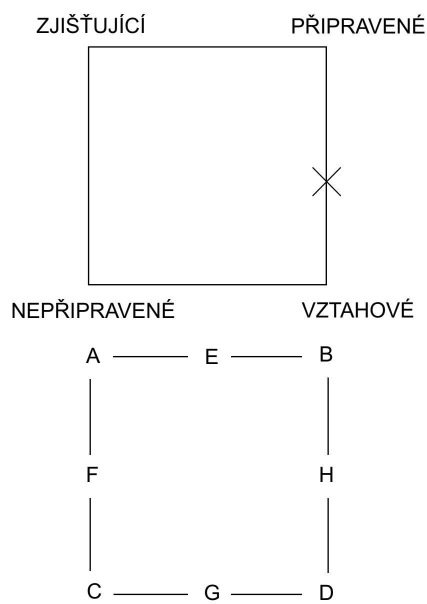
4.1.3.4 Pozorovací technika S. Rysa

U této techniky (Rys, 1975) jde v podstatě o modifikaci techniky škálování, u níž je posuzováno celkem devatenáct oblastí činnosti učitele ve vyučovací hodině (tab. 60).

***Tab. 60** Oblasti posuzování činnosti učitele ve vyučovací hodině*

1.	pohyby učitele	11	užití učebních pomůcek
2.	hlas učitele	12	aktivita žáků
3.	písmo učitele	13	kontrola
4.	řeč učitele	14	učitelovy otázky
5.	styl učitele	15	časový plán
6.	prostředky řízení	16	plán práce
7.	vnímavost situace	17	organizace podmínek
8.	motivace	18	modernizace
9.	zpracování obsahu učiva	19	závěr o práci
10.	způsob podání učiva		

Každá oblast činnosti učitele je hodnocena na devítistupňové škále, která má podobu čtverce. V rozích čtverců jsou slovně vyjádřeny charakteristiky určitého jevu tak, aby v protilehlých rozích byly charakteristiky polární nebo protikladné (pokud to je možné). Strany čtverců jsou pak vyjádřením vztahů nebo kombinací mezi těmito charakteristikami. Obrázek 22 uvádí příklad škály pro oblast č. 14 „učitelovy otázky".



***Obr. 22** Čtvercová škála pro oblast „učitelovy otázky" a její schéma*

Pozorovatel zaznamenává během pozorování do příslušných čtverců zjištěnou kvalitu jevů (např. křížkem). V příkladu, který zachycuje obrázek 22, jsou otázky učitele charakterizovány jako „vztahové – připravené".

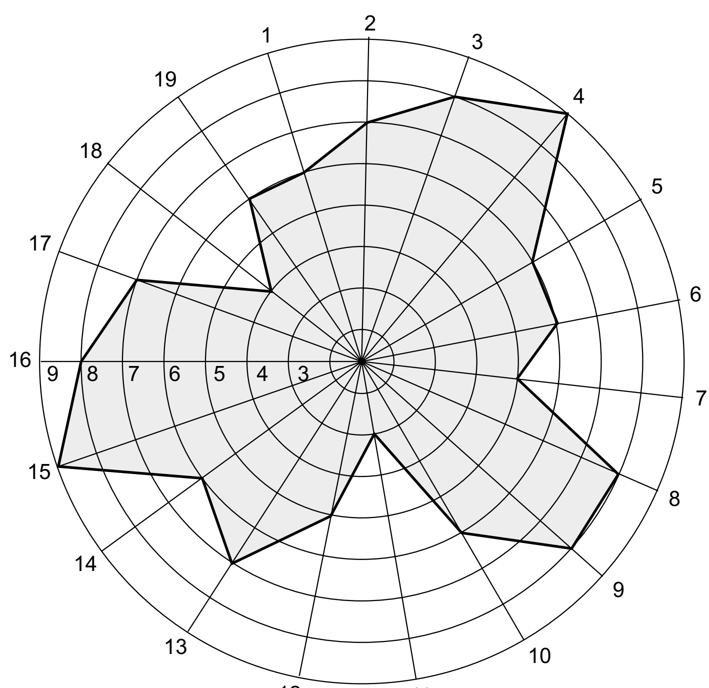
Pro každou škálu je vypracována určitá posuzovací stupnice, pomocí níž lze slovní charakteristiku pedagogického jevu kvantifikovat promítnutím do hodnot použité devítistupňové škály. Tabulka 61 uvádí posuzovací stupnici pro čtvercovou škálu „učitelovy otázky".

Tab. 61 *Posuzovací stupnice pro oblast „učitelovy otázky"*

Počet bodů	Poloha záznamu	Slovní charakteristika
1	–	nevyskytovalo se, nezjištěno
2	F	zjišťující, nahodilé
3	A	zjišťující, nahodilé, připravené
4	E	zjišťující, připravené
5	G	zjišťující nahodilé
6	D	vztahové, nahodilé, připravené
7	C	zjišťující, vztahové, nahodilé
8	H	vztahové, připravené
9	B	zjišťující, vztahové, připravené

Podle tabulky 61 charakteristika „vztahové – připravené" (H) odpovídá hodnocení osmi body (nejvyšší bodové hodnocení přísluší charakteristice B „zjišťující, vztahové, připravené").

Celkově je tedy u této pozorovací techniky činnost učitele charakterizována pomocí 19 číselných hodnot, které se mohou převést na názorný **kruhový diagram** (obr. 23).



Obr. 23 *Kruhový diagram*

Paprsky v kruhovém diagramu odpovídají devatenácti oblastem posuzování, soustředné kružnice vyjadřují hodnoty, kterých bylo v těchto oblastech dosaženo. Spojením jednotlivých bodů na kruhovém diagramu vznikne plošný obraz, který charakterizuje pozorovanou činnost jako celek.

Pozorování pomocí uvedené techniky předpokládá určitou pedagogicko-psychologickou kvalifikaci pozorovatelů.

4.2 DOTAZNÍK V PEDAGOGICKÉM VÝZKUMU

Velmi frekventovanou metodou získávání dat v pedagogickém výzkumu je dotazník. P. Gavora (2000) vymezuje dotazník jako „způsob písemného kladení otázek a získávání písemných odpovědí“. Kladené otázky se mohou vztahovat buď k jevům vnějším (např. názory učitelů na zaváděná organizační opatření), nebo k jevům vnitřním (např. postoje, motivy, citové stavy apod.). Samotný dotazník je soustava předem připravených a pečlivě formulovaných otázek, které jsou promyšleně seřazeny a na které dotazovaná osoba (respondent) odpovídá písemně. Někdy se ve stejném významu jako dotazník užívá termínu **anketa**. Většinou se však oba pojmy rozlišují a za anketu se považuje takové šetření, při kterém se účastníci sami, spontánně do šetření zapojují (např. ankety vyhlašované různými časopisy, rozhlasem, televizí apod.). Někdy se používá i širšího pojmu, tj. **dotazování**, které může mít dvě základní formy: interview a dotazník.

Dotazníkové metodě bývá často oprávněně vytýkáno, že nezjišťuje to, jací respondenti (pedagogická realita) skutečně jsou, ale jen to, jak sami sebe (pedagogickou realitu) vidí nebo chtějí, aby byli viděni. R. M. W. Travers (1969) uvádí, že zatímco v jiných oblastech společenských věd lze pozorovat určitý odklon od používání dotazníku směrem k interview, v pedagogice je frekvence jeho využívání dosud velmi (přesněji řečeno až příliš) vysoká. Vysoká frekvence používání dotazníků v pedagogickém výzkumu je zřejmě dána především jeho zdánlivě snadnou konstrukcí.

Zvlášť velké problémy jsou spojeny s používáním dotazníků rozesílaných poštou. D. Wallace (podle Travers, 1969) vyslovuje dokonce poněkud provokativní názor, že „nejjistější zásadou při rozhodování, zdali použít dotazníků rozesílaných poštou či nikoli, je: nepoužít“.

Velmi malou výpovědní hodnotu mají dotazníky neodborně sestavené a nevhodně použité. Data získaná dotazníkem mají vždy jen podmíněnou platnost a vyžadují vždy velmi obezřetnou interpretaci, abychom odlišili objektivní zjištění od subjektivních soudů. Nespornou výhodou dotazníku na druhé straně je, že umožňuje poměrně rychlé a ekonomické shromažďování dat od velkého počtu respondentů.

4.2.1 DRUHY POLOŽEK V DOTAZNÍKU

Místo termínu položka se často používá termín „otázka“. Označení položka je vhodnější, protože některé položky nemusí mít formu otázky, nýbrž například formu pokynu (např. „Vyberte tvrzení, se kterým souhlasíte“). Položky dotazníku lze třídit podle různých kritérií, z nichž nejčastěji se uvádějí: cíl, pro který je položka určena, forma požadované odpovědi a obsah, který položka zjišťuje.

4.2.1.1 Cíl, pro který je položka určena

Z tohoto hlediska lze rozlišovat položky obsahové (výsledkové) a položky funkcionální. Obsahové položky zjišťují údaje, které jsou nutné pro splnění výzkumného záměru, funkcionální položky mají optimalizovat průběh dotazování. Mezi funkcionální položky řadíme tzv. **kontaktní položky**, **položky funkcionálně psychologické, filtrační** a **kontrolní**.

Kontaktní položky

Kontaktní položky slouží k vytvoření náležitého kontaktu mezi respondentem a výzkumníkem. Bývají snadné a nenáročné, plní funkci úvodu k dotazování a uvádějí respondenta do zkoumané problematiky. Jako kontaktní položky nejsou zpravidla příliš vhodné dotazy na demografické údaje o respondentovi, protože mohou vzniknout pochybnosti o anonymitě dotazování a nedůvěra k výzkumníkovi. Demografické údaje bývá výhodnější požadovat až v závěru dotazníku. Jako kontaktní položky také nejsou vhodné „choulostivé" otázky nebo otázky respondenta nějakým způsobem znepokojující.

Funkcionálně psychologické položky

Funkcionálně psychologické položky slouží k odstranění nežádoucího napětí u respondenta (např. po zneklidňujících otázkách), někdy se používají k „přeladění" respondenta při přechodu od jednoho tématu k druhému nebo pro odstranění stereotypních postojů respondenta ke zkoumané problematice. Při kladení více otázek, které se týkají jednoho problému, se může totiž vytvořit u respondenta určitý stereotyp, na jehož základě potom odpovídá. V těchto případech bývá výhodné přerušit dotazování funkcionálně psychologickou položkou, která odvede pozornost respondenta jiným směrem, a teprve potom se k vlastnímu tématu vrátíme.

Kontrolní položky

Kontrolní položky mají za úkol prověřit věrohodnost zjišťovaných údajů. Je možno použít několika variant kontrolních položek. Jedna z možností spočívá v tom, že na jednu skutečnost se ptáme respondenta více položkami dotazníku. Například položíme otázku „Jste spokojen se svou prací?" a v jiné části dotazníku umístíme otázku „Chtěl byste změnit zaměstnání?". Při zjištění rozporu mezi odpověďmi můžeme položku buď vyloučit jako málo věrohodnou, anebo můžeme provést nějaké jiné doplňující šetření.

Jiná varianta kontrolních položek spočívá v tom, že se do dotazníku zařadí otázka, na niž s naprostou jistotou známe odpověď. Rozpor mezi skutečností a odpovědí respondenta indikuje opět malou věrohodnost jeho odpovědí. Další varianta kontrolních položek užívá otázek, které se ptají na neexistující skutečnosti (např. události, osoby atd.). Odpoví-li respondent určitým způsobem v těchto otázkách, lze z toho rovněž usuzovat na malou serióznost odpovědí. Otázkou zůstává, zda můžeme malou věrohodnost či neserióznost odpovědí (odhalenou kontrolní otázkou) vztahovat na všechny položky dotazníku, nebo jen na některou jeho část. Zřejmě bude záležet na druhu kontrolní otázky a na povaze zkoumaného problému. Důležitou zásadou, kterou musíme při používaní kontrolních položek respektovat, je, že kontrolní otázka nesmí být v dotazníku nikdy umístěna bezprostředně vedle položky, kterou kontrolujeme.

Filtrační položky

Filtrační položky se používají při zkoumání problémů, které se netýkají celého souboru zkoumaných jedinců. Filtrační položky se zpravidla zařazují před položky základní a mají za úkol eliminovat ty jedince, kteří pro šetření nemají význam. Například týká-li se jisté šetření žáků, kteří jsou členy některého sportovního oddílu, ptá se jedna z prvních položek na členství ve sportovních oddílech. Pokud respondent neodpoví žádoucím způsobem, nejsou jeho další odpovědi již brány v úvahu.

4.2.1.2 Forma požadované odpovědi

Podle toho, jakým způsobem má respondent v určité položce dotazníku odpovědět, lze rozdělit položky na otevřené a uzavřené (nestrukturované a strukturované). U otevřených položek respondent odpověď sám vytváří, u položek uzavřených určitým způsobem manipuluje s odpověďmi již navrženými (např. vybírá, seřazuje apod.).

Otevřené (nestrukturované) položky

Otevřené (nestrukturované) položky nenavrhují respondentovi žádné hotové odpovědi. Je u nich určen jen předmět, ke kterému se mají vyjádřit, jinak není respondent zpravidla nijak usměrňován (např. „Co si myslíte o ...?“). Nevýhodou těchto položek je právě jejich volnost, která působí obtíže při vyhodnocování. Po shromáždění všech odpovědí je zpravidla nutné provést dodatečnou kategorizaci, která umožní nepřehledně velký počet individuálních odpovědí převést na menší počet zvolených kategorií, čímž se vždy jistá část informace ztrácí. Provádění dodatečné kategorizace otevřených odpovědí vyžaduje rovněž poměrně kvalifikovaného pracovníka a je časově náročné. Proto při zpracování velkých dotazníkových šetření je používání této formy položek málo reálné. Otevřené položky je výhodné používat v předvýzkumu, kde můžeme nejfrekventovanějších typů odpovědí využít pro konstrukci nabídek pro položky uzavřeného typu. Kladem otevřených položek je, že umožňují často hlubší proniknutí ke sledovaným jevům a lépe postihují skutečné mínění respondentů než položky uzavřené. Výpovědní hodnota otevřených položek také značně závisí na dovednosti nebo ochotě respondentů se vyjadřovat. Otevřené položky jsou vhodné jako položky kontaktní anebo jako položky funkcionálně psychologické. Při grafické úpravě otevřených položek je vždy třeba pamatovat na přiměřeně velké místo pro uvedení odpovědí.

Uzavřené (strukturované) položky

Uzavřené (strukturované) položky se vyznačují tím, že se u nich respondentům předkládá vždy určitý počet předem připravených odpovědí. Hlavní výhodou těchto položek je to, že se podstatně zjednodušuje vyhodnocování odpovědí. Často také respondenti ochotněji vyplňují dotazník s již připravenými odpověďmi. Nevýhodou této formy položek na druhé straně zůstává fakt, že všechny možné kvality odpovědí jsou násilně vtěsnávány do schématu připravených odpovědí. Podle počtu předkládaných odpovědí lze uzavřené položky rozdělit na dichotomické a polytomické. Pokud na položku lze dát jen dvě vzájemně se vylučující odpovědi (např. ano – ne), hovoříme o **položkách dichotomických** (např. „Předplácíte si některý pedagogický časopis? ano – ne"). Pokud na položku existují dvě odpovědi, které se však vzájemně nevylučují, jde o tzv. **nepravou dichotomii**. U **polytomických položek** se předkládá více odpovědí než dvě. Tyto položky je možno dále rozdělit na výběrové, výčtové a stupnicové. Ve výběrových položkách se respondentům předkládá několik odpovědí, *z* nichž jednu mají vybrat. Je důležité, aby kategorie odpovědí (nabídky) byly vyčerpávající, ale ne příliš početné.

Příklad polytomické položky s výběrem odpovědí:

Býváte spokojen(a) se svými výsledky u zkoušek?
A vždy ano
B většinou ano
C někdy ano, někdy ne
D většinou ne
E nikdy

Abychom se vyhnuli nebezpečí, že neuvedeme některou možnou odpověď, můžeme použít i nabídky „jiná odpověď". Tuto nabídku volí respondent v případě, že mu nevyhovuje žádná z nabízených možností. Položky tohoto typu bývají označovány jako **položky polouzavřené**.

Příklad polouzavřené položky:

Na které střední škole jste maturoval(a)?
A na gymnáziu
B na střední odborné škole
C na středním odborném učilišti
D na jiné střední škole (uveďte které) ……………………………………………………

Při řazení nabídek odpovědí dbáme, aby navrhované odpovědi byly vždy (pokud to lze) seřazeny podle určitého kritéria (např. od odpovědi „vždy" k odpovědi „nikdy", podle velikosti, významu, frekvence atd.). Výsledky získané dotazníkovým šetřením lze potom snáze interpretovat.

Zvláštním druhem výběrových položek jsou tzv. **škálové položky**. U škálových položek respondent odpovídá tak, že vybírá určitý bod na předložené škále. Škálových položek je mnoho typů a mívají různou podobu (srov. oddíl 4.1.3.3).

Příklad škálové položky:

Sledujete pravidelně aktuálně vycházející pedagogickou literaturu?
vůbec nesleduji 1 2 3 4 naprosto pravidelně sleduji

Často se v dotaznících užívají tzv. **škály Likertova typu**. U těchto škál se prezentuje určité tvrzení a po respondentovi se požaduje, aby vyjádřil stupeň svého souhlasu, respektive nesouhlasu na hodnotící škále (obyčejně sedmibodové, ale někdy i vícebodové).

Příklad škály Likertova typu:

Efektivitu vyučovacího procesu ovlivňuje používání názorných učebních pomůcek.
naprosto nesouhlasím 1 2 3 4 5 6 7 naprosto souhlasím

Stupeň souhlasu (respektive nesouhlasu) s prezentovaným tvrzením vyjadřuje respondent například zakroužkováním, označením křížkem apod. příslušného čísla na škále.

Výčtové položky se vyznačují tím, že u nich respondent vybírá současně několik odpovědí. Počet odpovědí, které se mají vybrat, je buď neomezený, anebo je určen instrukcí (např. „Vyberte dvě z uvedených odpovědí, se kterými souhlasíte").

Příklad výčtové položky:

Které noviny pravidelně čtete?
A Lidové noviny
B MF Dnes
C Hospodářské noviny
D Právo
E jiné noviny (uveďte které) ……………………………..………………………………

Ve **stupnicových položkách** se respondentům předkládá určitý počet odpovědí s tím, že je mají seřadit podle určitého kritéria (např. podle oblíbenosti, významu apod.).

Příklad stupnicové položky:

Seřaďte následující předměty podle oblíbenosti tak, že předmětu, který máte nejraději, přiřadíte číslo 1 a předmětu nejméně oblíbenému číslo 6:
matematika ……
český jazyk ……
tělesná výchova ……
fyzika ……
dějepis ……
chemie ……

Při používání výčtových a stupnicových položek mohou vznikat obtíže při vyhodnocování odpovědí vzhledem k tomu, že různé kombinace odpovědí (respektive různá pořadí odpovědí) mohou mít různou výpovědní hodnotu. Proto se výčtové a stupnicové položky užívají jen v nezbytných případech.

Odpovědi v uzavřených položkách je možno zaznamenávat buď přímo do formulářů dotazníků (zakroužkováním, podtržením, označením křížkem apod.), nebo lze použít zvláštních **záznamových listů**, které mají výhodu v tom, že samotný formulář dotazníku můžeme používat opakovaně.

4.2.1.3 Obsah, který položka dotazníku zjišťuje

Podle tohoto kritéria můžeme položky v dotazníku rozdělit na položky zjišťující fakta, položky zjišťující znalosti a vědomosti a na položky zjišťující mínění, postoje a motivy respondentů.

Položky zjišťující fakta

Položky zjišťující fakta zpravidla nevyžadují velkou námahu při odpovídání, a proto se často používají jako úvodní položky dotazníku. Užívají se však i v průběhu dotazování, aby si respondent odpočinul od náročnějších otázek. Položky zjišťující fakta bývají velmi často dichotomické (typ ano – ne). Mezi položky zjišťující fakta patří i otázky na demografické údaje (věk, pohlaví, zaměstnání, složení rodiny, sociální postavení apod.). Z psychologického hlediska je nejvhodnější umísťovat otázky na demografické údaje až na konec dotazníku.

Položky zjišťující znalosti nebo vědomosti

Položky zjišťující znalosti nebo vědomosti je nutné v dotazníku formulovat velmi opatrně, aby se respondent necítil kompromitován při neznalosti. Dá se toho docílit například tím, že z formulace položky vyplývá, že eventuální neznalost je zcela běžným jevem. Je možno použít např. formulaci: „Nevzpomínáte si, kdo byl posledním předsedou Federálního shromáždění ČSFR?“ (méně vhodná formulace: „Kdo byl posledním předsedou Federálního shromáždění ČSFR?“) apod.

Položky zjišťující mínění, postoje a motivy

Položky zjišťující mínění, postoje a motivy jsou velmi citlivé na formulaci a na zařazení v dotazníku. Důležitou zásadou je, že v položkách se nesmí projevovat postoje, názory a hodnocení autora dotazníku. Řada otázek tohoto typu může přivést respondenta do rozpaků, vyvolat u něho negativní reakci apod. V těchto případech se doporučuje dát ve formulaci položek najevo, že různost názorů je zcela přirozená a normální.

V položkách zjišťujících mínění, postoje a motivy se užívá často tzv. **nepřímých (projektivních) otázek**. Užívají se zvláště při zkoumání tzv. „choulostivých“ problémů, o nichž respondenti neradi hovoří. V těchto případech se například neptáme přímo na názory dotazovaného, ale na mínění celé skupiny, ke které dotazovaný patří, na mínění „lidí vůbec“ atd.

Příklad nepřímé otázky:

Co si myslí žáci vaší třídy o třídní učitelce? ……………………………………………

U nepřímých otázek předpokládáme, že se respondent ztotožní s příslušnou skupinou a do odpovědi promítne svůj názor. Formulace a používání nepřímých otázek předpokládá určité zkušenosti a schopnost empatie vůči respondentům, jimž je dotazník určen.

Velkým problémem při používání položek, které zjišťují mínění, postoje a motivy, je skutečnost, že respondent může vědomě zkreslovat své odpovědi. Plně to platí u položek, kde je zřejmá souvislost mezi otázkou a tím, co se má zjišťovat. Lze očekávat, že respondenti nebudou ochotni vypovídat pravdivě například tehdy, jestliže ze zaměření otázek dotazníku vyplývá, že cílem šetření je odkrytí jejich negativních vlastností. Jestliže chceme získat věrohodné informace i z těch oblastí, kde lze očekávat vědomé zkreslování odpovědí, můžeme se pokusit použít tzv. **maskované otázky**. U těchto otázek nesmí být na první pohled patrné, co se otázkou zjišťuje. Vytvoření maskovaných otázek je velmi náročné a vyžaduje vedle hlubšího poučení také velké zkušenosti. Navíc, validita maskovaných otázek nebývá příliš vysoká.

4.2.2 NEJDŮLEŽITĚJŠÍ POŽADAVKY NA KONSTRUKCI DOTAZNÍKU

V následujícím přehledu se pokusíme shrnout nejdůležitější pravidla, zásady a požadavky, které bychom měli dodržovat při návrhu jednotlivých položek dotazníku a při sestavování dotazníku jako celku:

- Položky v dotazníku musí být všem respondentům **jasné a srozumitelné**. To znamená, že například musíme respektovat to, jakým respondentům je dotazník určen (věk, vzdělání, motivace). Položky dotazníku by měly být formulovány také co možná nejstručněji.
- Formulace položek v dotazníku musí být naprosto **jednoznačná** a nesmí připouštět chápání více způsoby. Například zkoumáme-li volný čas žáků, je třeba jednoznačně vymezit, co si pod pojmem „volný čas" představujeme.
- Velké opatrnosti je třeba při formulaci položek typu **„proč"**. Zpravidla není možné se přímo ptát například na příčiny určitého chování, vzhledem k tomu, že je respondenti buď neznají, nebo si je plně neuvědomují. Na otázky o příčinách existujících jevů musí dát odpověď většinou výzkumník na základě analýzy výsledků šetření. Jinak je tomu ovšem v případech, kdy zjištění mínění respondenta o příčinách určitých jevů je výzkumným záměrem.
- Položky dotazníku by měly zjišťovat jen **nezbytné údaje**, které nelze získat jiným způsobem. Dotazník by také neměl být příliš rozsáhlý.
- Položky v dotazníku nesmějí být **sugestivní**, tj. takové, že již svou formulací napovídají, jak mají být zodpovězeny.
  Kolik knih jste přečetl za poslední měsíc? ....
  Takto formulovaná otázka respondentovi vlastně naznačuje předpoklad, že alespoň jednu knihu za poslední měsíc přečetl. V takové situaci respondent často odpoví nepravdivě z obavy před kompromitací.
- Pro úspěch každého dotazníkového šetření je nezbytným předpokladem **ochota respondentů spolupracovat**. Ochotu spolupracovat může zvýšit přiměřená motivace v úvodu dotazníku, kde stručně vysvětlíme smysl a potřebnost prováděného šetření. Ochota spolupracovat do značné míry závisí také na tom, jak je vyplňování dotazníku zajímavé a náročné. Nevhodné jsou položky, jejichž zodpovězení je příliš pracné (např. nutnost vyhledávat dokumenty, dlouhé písemné odpovědi apod.). V tomto směru často lépe vyhovují uzavřené, respektive polouzavřené položky.
- Dotazník musí vždy obsahovat **jasné pokyny** k vyplňování. Je to zvlášť důležité u dotazníků rozesílaných poštou.
- Při konstrukci dotazníku je třeba dbát na to, aby získané údaje bylo možno **snadno třídit, tabelovat a zpracovávat**. Přehnaná snaha usnadnit si zpracování výsledků dotazníkového šetření však někdy může vést k dezorientaci respondentů při vyplňování. Považujeme za nepřiměřené a zbytečně matoucí, jestliže se ve formulářích dotazníků objevují označení, jako například znak, kód a jiné symboly nebo výrazy, kterým nemusí respondent rozumět a které mohou odrazovat od vyplňování.
- Při **řazení položek** v dotazníku dáváme vždy přednost pořadí, které vyhovuje z psychologického hlediska, před pořadím logickým. Nejdůležitější položky se doporučuje umísťovat ve střední části dotazníku. Dotazník obyčejně začíná zcela jednoduchými,

konkrétními otázkami, dále následují obsahové položky, které se podle potřeby prokládají položkami filtračními, kontrolními a funkcionálně psychologickými. Někdy bývá pro řazení položek, vztahujících se k jednomu tématu, doporučována tzv. **technika nálevky**. Tato technika spočívá v tom, že soubor položek začíná položkou nejobecnější a další položky se potom postupně zužují. Jiná modifikace této techniky spočívá v tom, že jako první formulujeme obecnou otevřenou položku a na ni potom navazují užší položky uzavřeného typu.
Příklad řazení položek (Skalková et al., 1983):

Charakterizujte úroveň ukázněnosti žáků ve vaší škole. ....
Kontrolujete pravidelně včasný příchod žáků do školy? ano – ne
Projednávají třídní učitelé projevy nekázně některých žáků s vyučujícími? ano – ne
atd.

Někdy se mohou položky vztahující se k určitému tématu řadit i opačným způsobem, tj. od nejužších (nejkonkrétnějších) k nejširším (nejobecnějším). V tomto případě se hovoří o **technice převrácené nálevky**.

4.2.3 VLASTNOSTI DOBRÉHO DOTAZNÍKU

Tak jako každý jiný prostředek měření, měl by i dotazník splňovat základní požadavky kladené na dobré měření. Jsou to zejména validita, reliabilita a praktičnost.

Validita dotazníku spočívá v tom, že dotazník zjišťuje skutečně to, co má zjišťovat, tj. to, co je výzkumným záměrem. Konstrukce dotazníku v klasických pedagogických výzkumech by měla vždy vycházet ze zdůvodněné vědecké hypotézy a jednotlivé položky musí přinášet data pro verifikaci této hypotézy. Posouzení stupně validity dotazníku je vždy do určité míry subjektivní a záleží především na fundovanosti a kompetentnosti autora dotazníku. Lze jen doporučit, aby při posuzování validity dotazníku nevycházel autor jen z vlastních názorů, ale aby nechal vždy posoudit navrhovaný dotazník dalšími odborníky.

Reliabilitou dotazníku se rozumí schopnost dotazníku zachycovat spolehlivě a přesně zkoumané jevy. Dostatečně vysoká reliabilita je nezbytným předpokladem dobré validity dotazníku, i když sama o sobě ještě validitu nezaručuje. Je bohužel skutečností, že uživatelé dotazníků se většinou o stupeň spolehlivosti a přesnosti získávaných výsledků příliš nezajímají. Přitom v mnoha pedagogických výzkumech je dotazník jediným zdrojem informací, o který se šetření opírá. V zásadě je možné stupeň reliability výsledků dotazníkových šetření vždy určitým způsobem odhadovat nebo kontrolovat.

4.2.3.1 Reliabilita měření prováděného dotazníkem

Pojem „reliabilita“ se začal v pedagogice používat až v souvislosti s didaktickými testy. U ostatních metod sběru dat (metod měření) se zatím většinou žádné posuzování reliability neprovádí. Přitom lze pro každý systém měření nalézt způsob, jak stupeň jeho spolehlivosti a přesnosti, tj. stupeň realiability, posoudit, změřit, anebo alespoň přibližně odhadnout.

Při volbě příslušného statisticko-empirického postupu určování reliability dotazníkového šetření musíme v první řadě zvážit, jaký druh dat při šetření získáváme. Existují postupy využitelné pro analýzu nominálních dat, postupy pro stanovení reliability u ordinálních dat, ale také procedury určené pro metrická data.

Nejširší výběr metod pro stanovení reliability se nabízí v případě metrických dat. Zkušenosti ukazují, že například mnohé škály používané v dotaznících poskytují data, která můžeme docela dobře považovat za data metrická, a využívat tudíž při jejich analýze postupů vyvinutých pro metrická data (např. jednofaktorová či dvoufaktorová analýza rozptylu apod.).

4.2.3.2 Metody určování reliability

Pro posouzení reliability výsledků dotazníkového šetření se v sociologicky orientované literatuře někdy doporučuje tzv. **metoda štěpení**. U této metody se srovnávají výsledky, jichž bylo u týchž respondentů dosaženo pomocí dvou různých, ale ekvivalentních forem, dotazníku. Tato metoda je v běžných výzkumech málo reálná, protože vyžaduje velmi náročnou přípravu celého šetření a je navíc příliš náročná časově i ekonomicky.

V některých případech lze reliabilitu dotazníkového šetření posuzovat také tak, že dotazník předložíme týmž respondentům po uplynutí určitého optimálního období znovu. Na stupeň reliability výsledků lze usuzovat ze stupně shody mezi prvním a opakovaným šetřením. V tomto případě záleží na tom, aby doba pro opakované šetření byla určena skutečně optimálně. Je-li příliš krátká, hrozí nebezpečí zapamatování, a tím zkreslení výsledků, je-li naopak příliš dlouhá, může dojít ke změně měřené vlastnosti, takže potom vlastně při opakovaném šetření měříme něco jiného než při šetření prvním. Většinou se uvádí, že nejvhodnější čas pro opakované měření je asi dva až tři týdny po měření prvním. Tato metoda ověřování reliability dotazníkového šetření není pro běžné výzkumy příliš vhodná, protože opakované šetření se stejnými respondenty je jen málokdy uskutečnitelné.

Jako další způsob určení reliability dotazníkového šetření se v sociologické literatuře uvádí postup, kdy se dotazník zadá dvěma reprezentativním výběrům téhož základního souboru. Získané výsledky se potom navzájem srovnávají. Ze stupně shody mezi výsledky v obou výběrech se usuzuje na stupeň reliability provedeného dotazníkového šetření. Modifikací uvedené metody je postup, při kterém přiměřeně velký reprezentativní výběr z jistého základního souboru rozdělíme náhodně na dva stejně velké výběrové soubory. Výsledky získané v těchto dvou výběrových souborech potom srovnáváme (Chráska, 1996).

Poslední uvedená metoda má, ve srovnání s ostatními postupy, výhodu v tom, že nevyžaduje žádné opakované měření ani zadávání jiného dotazníku, nýbrž vychází pouze z dat, která byla získána při běžném dotazníkovém šetření. Metodu lze poměrně snadno aplikovat zejména při ověřování reliability u klasických dotazníků s výběrovými položkami.

Reliabilita pomocí Cohenova koeficientu kappa

Příklad 55

Stanovení reliability dotazníkového šetření budeme ilustrovat na příkladu dotazníkového šetření, ve kterém byly měřeny postoje vysokoškolských studentů ke studiu. Souboru 200 studentů byl předložen dotazník s třiceti položkami, u nichž byl požadován vždy výběr jedné odpovědi z několika nabídnutých alternativ.

Znění jedné z položek dotazníku:

14. Býváte spokojen(a) se svými výsledky u zkoušek?
A vždy ano
B většinou ano
C někdy ano, někdy ne
D většinou ne
E nikdy ne

K posouzení stupně reliability provedeného šetření jsme rozdělili celý výběrový soubor náhodně na dva stejně velké soubory po 100 respondentech. Odpovědi respondentů v jednotlivých položkách dotazníku jsme zapisovali do matic, které zachycují, jak dalece se v jednotlivých odpovědích projevují shody mezi oběma náhodně vytvořenými skupinami. Tabulka 62 uvádí odpovědi respondentů v položce, která byla uvedena jako příklad.

Čísla uvedená vpravo od matice a pod maticí (marginální četnosti) uvádějí počty jednotlivých odpovědí v obou skupinách respondentů. Z matice lze vyčíst, jak dalece se respondenti z obou náhodně vytvořených skupin v odpovědích v dané položce shodli. Čísla na diagonále matice uvádějí vždy počty shodných odpovědí ve skupině I a II. Čím větší četnosti se objevují na diagonále matice, tím větší shoda mezi oběma skupinami nastala, a tím je také větší spolehlivost provedeného dotazníkového šetření.

Tab. 62 *Matice zachycující shody odpovědí v jedné položce dotazníku*

		Skupina II						
		A	B	C	D	E	neodp.	Σ
Skupina I	A	(12)	5					17
	B		(15)					15
	C			(28)				28
	D				(21)			21
	E			8		(8)		16
	neodp.			2			(1)	3
	Σ	12	20	38	21	8	1	100

Míru shody mezi odpověďmi v obou náhodně vytvořených skupinách lze exaktně vyjádřit například pomocí některého z koeficientů, které jsou doporučovány pro posuzování reliability kategoriálních pozorovacích technik. Vhodný je například **Cohenův koeficient kappa** (Mareš, 1983), který je možno vypočítat ze vztahu

$$\kappa = \frac{p_p - p_o}{1 - p_o} \tag{93}$$

$$\text{kde } p_p = \frac{1}{n}\sum n_s \quad \text{a} \quad p_o = \frac{1}{n^2}\sum n_I \cdot n_{II} \tag{94}$$

Ve vzorcích je κ Cohenův koeficient kappa, p_p je tzv. zjištěná (pozorovaná) proporce shody, p_o je očekávaná proporce shody, n_s jsou četnosti shodných odpovědí v dané položce dotazníku (tj. četnosti na diagonále matice, jinak také vždy menší z hodnot n_I a n_{II}), n je celková četnost v matici, n_I jsou četnosti jednotlivých odpovědí v první skupině (marginální četnosti v matici pro první skupinu) a n_{II} jsou četnosti jednotlivých odpovědí v druhé skupině (marginální četnosti v matici pro druhou skupinu). Pro hodnoty uvedené v tabulce 62 dostáváme

$$p_p = \frac{1}{100}(12+15+28+21+8+1) = 0{,}850$$

$$p_o = \frac{1}{100^2}(12\cdot 17+20\cdot 15+38\cdot 28+21\cdot 21+8\cdot 16+1\cdot 3) = 0{,}214$$

$$\kappa = \frac{0{,}850-0{,}214}{1-0{,}214} = 0{,}809$$

U kategoriálních pozorovacích technik se většinou požaduje, aby hodnota Cohenova koeficientu kappa byla minimálně 0,80. Toto kritérium je vhodné i v případě posuzování shody odpovědí u popisované metody. Vypočítanou hodnotu κ = 0,809 lze tedy považovat za vyhovující z hlediska shody mezi náhodně vybranými respondenty, a tudíž i za vyhovující z hlediska reliability.

Používání Cohenova koeficientu kappa je výhodné i tím, že umožňuje testovat statistickou významnost vypočítaného koeficientu pomocí normované normální veličiny u (srov. oddíl 2.4 Normální rozdělení). Toto kritérium lze pro daný koeficient κ vypočítat ze vztahu

$$u = \frac{\kappa}{\sqrt{\dfrac{p_o}{n(1-p_p)}}} \tag{95}$$

kde u je hodnota **normované normální veličiny** a význam ostatních symbolů zůstává stejný jako ve vztazích (93) a (94).
Pro výše uvedený příklad dostáváme

$$u = \frac{0{,}809}{\sqrt{\dfrac{0{,}214}{100\cdot(1-0{,}850)}}} = 6{,}77$$

Vypočítanou hodnotu srovnáváme s kritickou hodnotou pro zvolenou hladinu významnosti (pro hladinu významnosti 0,01 činí kritická hodnota $u_{0,01}$ = 2,58; platí pro oboustranný test – srov. oddíl 3.1.2). Je patrné, že vypočítaná hodnota je v našem případě větší než hodnota kritická, a proto můžeme konstatovat, že vypočítaný koeficient vypovídá o statisticky významné shodě mezi odpověďmi respondentů.

Z koeficientů shody pro jednotlivé položky dotazníku je možno vypočítat průměrný Cohenův koeficient pro celý dotazník, který charakterizuje reliabilitu dotazníku jako celku.

4.2.4 PROVEDENÍ DOTAZNÍKOVÉHO ŠETŘENÍ

Dotazník lze předat respondentům v podstatě třemi způsoby: rozesíláním poštou, osobně nebo prostřednictvím dalších osob.

Snad nejvýhodnější (ale ne vždy proveditelné) je osobní předávání dotazníků, po kterém bezprostředně následuje vyplnění dotazníků respondenty a vybrání dotazníků zpět. Tento postup je dobře proveditelný například ve výzkumech, jimiž se zkoumají mínění vysokoškolských studentů nebo žáků středních či základních škol. Výhodou tohoto způsobu zadávání dotazníků je prakticky stoprocentní návratnost.

Pokud rozesíláme dotazníky poštou, musíme počítat s poměrně malou návratností, zvláště u dotazníků anonymních. Údaje o průměrné návratnosti se v literatuře rozcházejí, ale jsou zhruba v intervalu od 30 % do 60 %. Prakticky to znamená, že u dotazníků rozesílaných poštou je třeba rozesílat alespoň dvojnásobek dotazníků ve srovnání s požadovaným rozsahem výběru. Další stinnou stránkou dotazníků rozesílaných poštou je skutečnost, že vzorek respondentů, kteří dotazník vyplnili a vrátili, nemusí být nutně reprezentativní. Výzkumy totiž ukazují, že vracení dotazníků není jen věcí náhody, nýbrž že se na něm podílejí různé další vlivy. Bylo například prokázáno, že dotazníky vracejí spíše lidé s vyšším vzděláním, lidé s větší odpovědností a kladným postojem ke zkoumané problematice.

Pro úspěch dotazníkového šetření je důležité, aby respondenti měli záruku, že dotazníkem zjištěné skutečnosti nebudou zneužity proti nim. V tomto směru bývá prospěšné užívání **anonymních dotazníků**. Anonymním dotazníkem většinou získáme pravdivější údaje, na druhé straně však anonymní dotazník může svádět k neodpovědnému vyplňování či dokonce k recesi.

Před provedením vlastního dotazníkového šetření je vhodné provést předvýzkum, při kterém se doporučuje navržený dotazník vyzkoušet na vzorku alespoň třiceti respondentů. Pečlivé provedení předvýzkumu zmenší riziko neúspěchu při vlastním dotazníkovém šetření. Na základě výsledků a zkušeností máme možnost navržený dotazník korigovat (upravit formulace položek, vypustit některé položky apod.).

4.2.5 KATEGORIZACE A TŘÍDĚNÍ MATERIÁLU ZÍSKANÉHO DOTAZNÍKEM

Po shromáždění vyplněných dotazníků od respondentů je potřeba získaný materiál nejdříve zkontrolovat z hlediska jeho korektnosti. Doporučuje se vyloučit z dalšího zpracování dotazníky, které jsou vyplněny zjevně nesprávně nebo neúplně (např. dotazníky, ve kterých respondent nedodržel instrukci a vybral více odpovědí apod.).

V dotazníkových šetřeních se často vyskytuje situace, že jednu vlastnost zjišťuje (měří) současně více položek. V těchto případech je možné (někdy dokonce výhodné) vyjádřit úroveň měřené vlastnosti syntetickým způsobem, pomocí tzv. **indexů**. Indexy, které vyjadřují celkovou míru zjišťované vlastnosti, je možno stanovit v podstatě dvojím způsobem. Buď se sečtou dohromady všechny údaje u položek měřících jednu vlastnost, anebo se index vyjádří jako poměr mezi součtem údajů ze všech položek a maximálně možným součtem ve všech položkách. Při dalším zpracování se místo s jednotlivými položkami pracujeme jen s indexy, čímž se celé zpracování i interpretace výsledků podstatně zjednodušuje.

Jednotlivé položky dotazníku vyjadřují různé **znaky** (proměnné) zkoumaného souboru respondentů. Znakem je například údaj o věku určité osoby, údaj o postojích respondenta k určité události, údaj o průměrném prospěchu žáka apod.

Znaky, se kterými se v dotaznících setkáváme, lze rozdělit na čtyři následující druhy (srov. oddíl 2.1 Měření a jeho druhy):

- **Znaky nominální** (kvalitativní), které vypovídají jen o příslušnosti respondenta k určité kategorii odpovědí. Například v položce, která zjišťuje povolání rodičů žáka (měřený znak), zjistíme, kolik rodičů pracuje jako dělníci, zemědělci, státní zaměstnanci, soukromí podnikatelé atd.
- **Znaky pořadové** (ordinální) vypovídají o vzájemném pořadí respondentů podle určitého hlediska. Příkladem položky dotazníku, která měří ordinální znak, je například:

  Jaké máte nejvyšší ukončené vzdělání?
  A základní
  B středoškolské
  C vysokoškolské

- **Znaky intervalové** vypovídají o tom, jak velké jsou rozdíly mezi vlastnostmi respondenta. Příkladem položky, která měří intervalový znak, je například:

  Kolik bodů jste získal(a) v testu při zkouškách v autoškole? ....

- **Znaky poměrové** podávají úplnou informaci o kvantitě měřeného jevu. Poměrový znak informuje nejen o rozdílech mezi respondenty v určité vlastnosti, ale také o tom, kolikrát je určitá vlastnost jednoho jedince větší nebo menší než jiného. Příkladem položky, která měří poměrový znak, je například:

  Kolik měříte? ......
  Kolik vážíte? ......

Intervalové a poměrové znaky bývají také často označovány jako **znaky metrické** nebo **kardinální**. V dotaznících se nejčastěji setkáváme se znaky nominálními nebo ordinálními. Znaky metrické (kardinální) se často pomocí vhodné kategorizace převádějí na znaky ordinální.

Před statistickým zpracováním výsledků dotazníkového šetření je třeba provést (popřípadě jen doplnit) **kategorizaci odpovědí**. Znamená to u každé položky dotazníku jednoznačně určit, které kategorie (druhy) odpovědí přicházejí v úvahu. U uzavřených položek bývá kategorizace již naznačena stavbou položky (nabízené odpovědi tvoří kategorie), ale často je i zde nutno určitou dodatečnou kategorizaci provádět. U výběrových položek například přidáváme kategorii „neodpověděl“, u položek výčtových vytváříme kategorie podle různých kombinací vybraných odpovědí. Komplikovanější situace je u položek stupnicových, kde je nutno vytvořit řadu kategorií podle uváděných pořadí odpovědí.

Výsledky v položkách, v nichž získáváme metrické znaky, se většinou kategorizují do čtyř až šesti kategorií, čímž tyto znaky převádíme, jak již bylo uvedeno, na znaky ordinální. Například odpovědi na otázku „Kolik je vám let?“ můžeme kategorizovat takto:

A méně než 26 let
B 26–35 let
C 36–45 let
D více než 45 let

U položek otevřených je třeba provést **úplnou kategorizaci** odpovědí, tj. musíme všechny individuální odpovědi přiřadit k určitému počtu zvolených kategorií. V některých případech se nevyhneme zavedení kategorie **„jiná odpověď“**, která pro získání nových poznatků většinou mnoho nepřináší. Kategorií by neměl být velký počet (většinou ne více než čtyři až šest), protože větší počet kategorií znemožňuje provedení statistické vztahové analýzy. Kategorie odpovědí bývají někdy také označovány jako třídy znaků.

V případě, že zpracování výsledků dotazníkového šetření se uskuteční na počítači, může po kategorizaci odpovědí následovat ještě tzv. **kódování**, kterým rozumíme přiřazení určitého číselného kódu (číselného označení) každé položce dotazníku a každé kategorii (třídě) odpovědí. Jednotlivé číselné kódy se potom vkládají do počítače. Dalším krokem při zpracování výsledků dotazníkového šetření je tzv. **třídění**. Třídění je postup, pomocí něhož zjišťujeme, kolik respondentů má společný buď jeden, nebo dva, popřípadě více společných znaků.

Pokud zjišťujeme, kolik respondentů má společný jeden znak, hovoříme o **třídění prvního stupně**. Například respondenty můžeme roztřídit podle položky, která zjišťuje jejich pohlaví, na 354 mužů a 297 žen apod. Výsledkem třídění prvního stupně bývá zpravidla tolik tabulek četností, kolik je v dotazníku položek.

Při **třídění druhého stupně** vyhledáváme ty respondenty, kteří mají shodné dva sledované znaky. Vyhledáváme tedy ty respondenty, kteří uvádějí určitou odpověď v jedné položce a současně uvádějí jistou odpověď v druhé položce. Má-li dotazník n položek (znaků), potom výsledkem úplného třídění druhého stupně je $n \cdot (n - 1) / 2$ kontingenčních tabulek. Například pro dotazník o třiceti položkách to představuje 435 kontingenčních tabulek.

Při třídění **třetího stupně** vyhledáváme osoby, které mají společné tři znaky, přičemž počet kontingenčních tabulek, které tyto souvislosti vyjadřují, je u dotazníku s n položkami

$$\frac{n!}{3! \cdot (n-3)!}$$

Pro dotazník o třiceti položkách to představuje již 4 060 kontingenčních tabulek.

Při běžných dotazníkových šetřeních zpravidla vystačíme s tříděním prvního a druhého stupně, přičemž u třídění druhého stupně se obvykle vyhledávají souvislosti jen mezi některými položkami. Volba techniky třídění závisí především na velikosti výběrového souboru. U malých výběrů (přibližně do 100 jedinců) můžeme použít **ruční třídění**. Tato technika spočívá v tom, že formuláře dotazníků rozdělujeme na „hromádky“ podle třídicího znaku, respektive třídicích znaků. Hledané četnosti potom zjistíme spočítáním formulářů dotazníků v jednotlivých hromádkách. Techniku ručního třídění lze u menších výběrových souborů doporučit zejména těm pracovníkům, kteří nemají s vyhodnocováním a interpretací dotazníkových šetření zkušenosti. Tento způsob třídění totiž umožňuje plněji pochopit podstatu této procedury a usnadňuje interpretaci získaných výsledků. U větších výběrových souborů je ruční třídění již příliš pracné, a proto se provádí pomocí počítačových programů (např. Excel, Statistica.cz, SPSS apod.).

4.2.6 POSTUP PŘI ANALÝZE DAT ZÍSKANÝCH DOTAZNÍKEM

Příklad 56

Obvyklý postup při zpracování dat získaných dotazníkem budeme ilustrovat na příkladu výzkumu, v němž se sledovaly postoje vysokoškolských studentů ke studiu.

Výběrovému souboru vysokoškolských studentů (n = 532) byl předložen dotazník, který zjišťoval různé aspekty jejich postojů ke studiu. V dotazníku byly obsaženy také dvě položky, které uvádí tabulka 63.

Tab. 63 *Znění dvou položek v dotazníku (spokojenost studentů s výsledky u zkoušek vs. pohlaví studentů)*

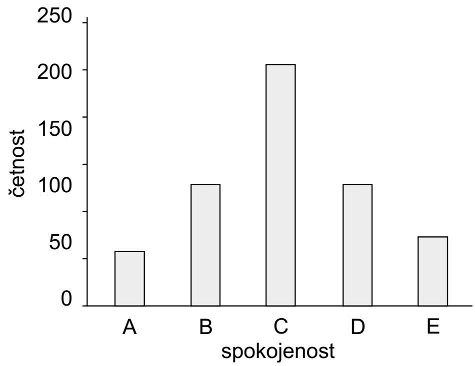
14. Býváte spokojen(a) se svými výsledky u zkoušek? A vždy ano B většinou ano C někdy ano, někdy ne D většinou ne E nikdy ne	30. Jaké je vaše pohlaví? A muž B žena

4.2.6.1 Interpretace výsledků třídění prvního stupně

V rámci třídění prvního stupně byly získány dvě tabulky četností (tab. 64 a 65) a data byla také prezentována graficky (obr. 24 a 25).

Tab. 64 *Spokojenost studentů s výsledky u zkoušek*

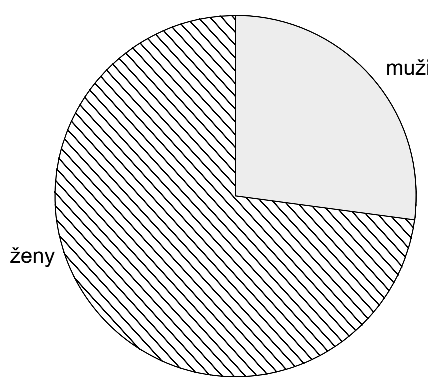
Spokojenost s výsledky	Absolutní četnost	Relativní četnost (%)
A vždy ano	47	8,8
B většinou ano	105	19,7
C někdy ano, někdy ne	213	40,0
D většinou ne	106	19,9
E nikdy ne	61	11,6
	∑ 532	∑ 100,0



***Obr. 24** Histogram četností – spokojenost s výsledky u zkoušek*

***Tab. 65** Pohlaví respondentů*

Pohlaví	Absolutní četnost	Relativní četnost (%)
A muž	144	27,1
B žena	388	72,9
	∑ 532	∑ 100,0



***Obr. 25** Pohlaví respondentů*

Z uvedených tabulek a grafu si lze učinit základní představu o složení výběrového souboru.

Při analýze výsledků získaných v jednotlivých položkách dotazníku nás zpravidla zajímá, jaká je charakteristická (typická) hodnota měřené proměnné (vlastnosti). K této charakteristice se často používá tzv. **modální kategorie**. V položce, která v našem případě zjišťovala spokojenost s výsledky u zkoušek, je modální kategorií kategorie C („někdy ano, někdy ne“).

Při analýze výsledků dosažených v jednotlivých položkách dotazníku můžeme také posuzovat, jak dalece se zjištěná data rozptylují – jakou mají **variabilitu**. Jako míra variability se u nominálních dat může použít například **variační poměr**, **nominální variance** nebo **normovaná nominální variance** (srov. oddíl 2.3.4.5).

K přibližnému posouzení variability stačí použít **variační poměr**. Pro data z tabulky 64 dostáváme podle vzorce (25) výsledek

$$\text{variační poměr} = 1-\frac{n_m}{n}=1-\frac{213}{532}=0,60$$

Přesnější mírou variability je tzv. **nominální variance** nebo **normovaná nominální variance** vzorce. Při použití vzorců (26) a (27) dostáváme

$$\text{nominální variance} = \frac{n^2-\sum n_k^2}{n^2}=\frac{532^2-(47^2+105^2+213^2+106^2+61^2)}{532^2} = 0,74$$

$$\text{normovaná nominální variance} = 0,74\cdot\frac{k}{k-1}=0,74\cdot\frac{5}{5-1}=0,925$$

4.2.6.2 Interpretace výsledků třídění druhého stupně

Ve výzkumu, jehož výsledky uvádějí tabulky 64 a 65 byla ověřována hypotéza, že „studenti (muži) jsou více spokojeni s výsledky u zkoušek než studentky (ženy)".

Ze získaných dat byla sestavena kontingenční tabulka (tab. 66) a formulována nulová a alternativní hypotéza.

Tab. 66 *Kontingenční tabulka (spokojenost versus pohlaví)*

	spokojenost s výsledky u zkoušek					
	A	B	C	D	E	Σ
muži	21 (12,7)	53 (28,4)	54 (57,7)	11 (28,7)	5 (16,5)	144
ženy	26 (34,3)	52 (76,6)	159 (155,3)	95 (77,3)	56 (44,5)	388
Σ	47	105	213	106	61	532

H_0: Spokojenost s výsledky u zkoušek je u mužů a žen stejná.
H_A: Spokojenost s výsledky u zkoušek je u mužů větší než u žen.

Test významnosti provedeme na hladině významnosti 0,01.

Čísla v tabulce (bez závorek) jsou **pozorované četnosti**, čísla v závorkách uvádějí **očekávané četnosti**. Z pozorovaných a očekávaných četností byla vypočítána hodnota testového kritéria chí-kvadrát (postup výpočtu je podrobně vysvětlen v oddíle 3.2.2)

$$\chi^2=\sum\frac{(P-O)^2}{O}=62,921$$

Protože kontingenční tabulka má $f = (2 - 1) \cdot (5 - 1) = 4$ stupňů volnosti, srovnáváme vypočítanou hodnotu χ^2 s kritickou hodnotou $\chi^2_{0,01}(4)=13,277$.

Vzhledem k tomu, že hodnota χ^2 je výrazně větší než hodnota kritická, můžeme odmítnout nulovou hypotézu a přijmout hypotézu alternativní. Mezi spokojeností s výsledky zkoušek u mužů a žen jsou **statisticky významné rozdíly**. Riziko (pravděpodobnost) nesprávného rozhodnutí je v tomto případě menší než 0,01.

Vypočítaná hodnota testového kritéria chí-kvadrát vypovídá o existenci vztahu (souvislosti, závislosti) mezi oběma znaky, ale nevypovídá o těsnosti tohoto vztahu. Těsnost vztahu (stupeň závislosti) mezi dvěma nominálními znaky lze posoudit pomocí některého z koeficientů, které vycházejí z hodnoty testového kritéria chí-kvadrát.

Často bývá doporučován **normovaný koeficient kontingence** C_{norm} (srov. oddíl 3.2.5). Těsnost vztahu mezi spokojeností s výsledky u zkoušek a pohlavím studentů je v tomto případě možno vypočítat podle vzorce (41)

$$C_{norm} = \sqrt{\frac{\dfrac{\chi^2}{n+\chi^2}}{\dfrac{r-1}{r}}} = \sqrt{\frac{\dfrac{62,921}{532+62,921}}{\dfrac{2-1}{2}}} = 0,460$$

Vypočítanou hodnotu je možno považovat za poměrně vysokou, což znamená, že mezi spokojeností studentů s výsledky u zkoušek a jejich pohlavím je poměrně těsný vztah.

Vypočítané hodnoty testového kritéria chí-kvadrát i normovaného koeficientu kontingence platí pro kontingenční tabulku jako celek; nevypovídají o souvislostech mezi jednotlivými kategoriemi odpovědí uvnitř tabulky. Pro interpretaci vztahů uvnitř kontingenční tabulky je (jak již bylo uvedeno) užitečnou pomůckou **znaménkové schéma kontingenční tabulky** (srov. oddíl 3.2.2.2). Pomocí znaménkového schématu dostáváme názornou informaci o tom, ve kterém poli kontingenční tabulky je porušen předpoklad nezávislosti obou srovnávaných znaků. Vypočítané hodnoty **normované normální veličiny** (z-skóre) uvádí tabulka 67.

Tab. 67 *Z-skóre pro jednotlivá pole kontingenční tabulky*

	spokojenost s výsledky u zkoušek				
	A	B	C	D	E
muži	2,36	4,74	–0,52	–3,40	–2,88
ženy	–1,47	–3,04	0,35	2,18	1,80

Vypočítaným hodnotám z-skóre odpovídá následující znaménkové schéma pro kontingenční tabulku (tab. 68).

Tab. 68 *Znaménkové schéma pro kontingenční tabulku*

	spokojenost s výsledky u zkoušek				
	A	B	C	D	E
muži	+	+++	0	--	--
ženy	0	--	0	+	0

Každé pole znaménkového schématu (ve kterém je alespoň jedno znaménko) umožňuje vyslovit určitý interpretační výrok o vztazích mezi sledovanými proměnnými. V uvedeném příkladě se může například jednat o následující výroky:

- Muži uvádějí mnohem častěji než ženy, že jsou se svými výsledky u zkoušek spokojeni (pole muži versus B).
- Ženy uvádějí častěji než muži, že jsou se svými výsledky u zkoušek nespokojeny (pole ženy versus D).
- Muži uvádějí častěji než ženy, že jsou se svými výsledky u zkoušek vždy spokojeni (pole muži versus A) atd.

Možnosti dotazníku poskytovat věrohodné informace o pedagogické realitě jsou často přeceňovány. Dotazník by se měl v pedagogickém výzkumu používat jedině tehdy, pokud není možné zkoumaný problém objasnit pomocí žádné jiné metody. Údaje získané dotazníkem by se měly pokud možno vždy objektivizovat pomocí dalších metod (např. pomocí pedagogického pozorování, rozhovorem, na základě studia dokumentů apod.). Přes všechny oprávněné výhrady k dotazníku však zůstává skutečností, že mnohé oblasti pedagogické reality nedokážeme zachytit lepším a efektivnějším způsobem.

4.3 INTERVIEW V PEDAGOGICKÉM VÝZKUMU

Interview je metoda shromažďování dat o pedagogické realitě, která spočívá v bezprostřední verbální komunikaci výzkumného pracovníka a respondenta. Někdy se v podobném významu používá také obsahově širšího českého termínu „rozhovor“. Protože však ne každý rozhovor je interview, je používání pojmu „interview“ přesnější a výstižnější. Anglický výraz „interview“ je složen ze dvou částí, kde *inter* znamená „mezi“ a *view* znamená „názor“ nebo „pohled“.

Velkou výhodou interview oproti jiným výzkumným metodám je navázání osobního kontaktu, který umožňuje hlubší proniknutí do motivů a postojů respondentů. U interview můžeme sledovat reakce respondenta na kladené otázky a podle nich usměrňovat jeho další průběh.

Úspěšnost interview z velké části závisí na schopnosti výzkumníka navázat přátelský vztah k respondentovi a na vytvoření otevřené atmosféry. Vytvoření vzájemně příjemného, uvolněného vztahu mezi výzkumníkem a respondentem se označuje termínem **raport**.

Podle toho, jak dalece je interview výzkumníkem řízeno, je možno rozlišit interview strukturované, polostrukturované a nestrukturované.

Strukturované interview

Strukturované interview se vyznačuje tím, že při něm tazatel postupuje podle přesně připraveného textu, jsou přesně určeny formulace otázek i jejich pořadí. Tazatel k otázkám nepřidává vlastní komentář, pouze čte otázky a zaznamenává odpovědi respondenta. Důsledně strukturované interview se přibližuje dotazníku, od kterého se liší jen tím, že záznam údajů provádí tazatel. Výhodou strukturovaného interview je, že poskytuje všem respondentům stejné podmínky k odpovědím a také to, že získané výsledky se dají většinou dobře statisticky zpracovávat. Nevýhodou je naopak obtížnější navazování kontaktu mezi tazatelem a respondentem a to, že strukturované interview působí vždy více méně strojeně.

Nestrukturované interview

Nestrukturované interview se více přibližuje přirozené komunikaci mezi lidmi. Tazateli musí být pochopitelně i u tohoto typu rozhovoru jasné, které informace má od respondenta získat. Konkrétní formulace otázek a jejich sled však je ponechán na tazateli. Tazatel se může volně vracet k nejasným nebo zajímavým bodům ve výpovědi respondenta. Výhodou nestandardizovaného interview je především to, že umožňuje snadnější navázání kontaktu mezi tazatelem a respondentem, což může znamenat jeho bezprostřednější a upřímnější projev. Nevýhodou naopak je to, že ne všichni respondenti odpovídají za naprosto stejných podmínek a také to, že nestandardizované interview většinou neposkytuje přímo kvantitativně zpracovatelný materiál.

Polostrukturované interview

Polostrukturované interview je určitým kompromisem mezi výše uvedenými typy interview. Respondentům se v tomto případě nabízí k jednotlivým otázkám vždy několik alternativ odpovědí, ale navíc se od nich požaduje vysvětlení nebo zdůvodnění.

Skupinové interview

Podle počtu zúčastněných osob lze rozlišit rozhovor individuální a skupinový. **Skupinový rozhovor** se užívá např. při zkoumání „choulostivých" otázek. Účastníci rozhovoru v těchto případech nepociťují otázky tak osobně jako při rozhovoru individuálním. Většinou se doporučuje, aby skupina, s níž je prováděn skupinový rozhovor, měla šest až deset členů. Rozhovor by měl probíhat v přirozeném prostředí (např. třída, pracoviště, hřiště apod.). Určitá výhoda skupinového rozhovoru spočívá také v tom, že umožňuje okamžitou kontrolu odpovědí jednotlivců. Ostatní účastníci rozhovoru totiž zpravidla opravují chyby, kterých se vypovídající dopustil. Mezi nevýhody skupinového rozhovoru lze na prvním místě počítat nestejnou intenzitu účasti dotazovaných osob. Tato skutečnost vyplývá jednak z různé míry aktivity či agresivity účastníků (názory aktivních a agresivních osob převládají), jednak z různé úrovně vědomostí či různé úrovně orientovanosti účastníků rozhovoru (stydlivost před osobami s většími vědomostmi či větší orientovaností).

Nejdůležitější pravidla pro realizaci interview:

- Interview by mělo probíhat vždy za **vhodné situace**. Pro interview by měl být vytvořen dostatečný časový prostor. Interview by neměly být svědky osoby, jichž se netýká. Interview by mělo vždy probíhat v přirozeném prostředí.
- Doporučuje se interview **začínat nejobecnějšími otázkami**, které respondenta uvedou do problematiky.
- Při interview je třeba čelit **působení psychologických faktorů**, které mohou negativně ovlivnit výsledky rozhovoru (srov. oddíl 4.1.2).
- Je třeba vytvářet podmínky pro náležité **navázání kontaktu** s respondentem (raport) a pro jeho motivaci ke spolupráci. Tazatel by měl projevovat přiměřený zájem o výpovědi respondenta, měl by být taktní a nevtíravý. Na výsledky interview má vliv i úprava

zevnějšku tazatele, jeho chování během rozhovoru (např. upjatost, familiárnost atd.) a řada dalších jeho osobnostních charakteristik.

- Velký význam má **přesný záznam** průběhu interview. Písemný záznam můžeme provádět přímo během interview, nebo až po jeho ukončení. Provádíme-li zápis během interview, snížíme tím riziko zkreslení výsledků vlivem zapomínání, avšak zapisování odpovědí během interview může znamenat obtížnější navazování kontaktu a vytvoření celkově nepříznivé atmosféry rozhovoru. Zaznamenávání odpovědí respondentů až po skončení interview většinou z psychologického hlediska vyhovuje lépe, klade však větší nároky na přesné zapamatování průběhu interview. Pro záznam průběhu interview je možno použít také technických prostředků (např. diktafonu). Vzhledem k tomu, že používání těchto zařízení je pro většinu respondentů zcela běžnou záležitostí, rušivý vliv na průběh interview nebývá většinou příliš velký.

Pro formulaci otázek interview, pro vedení interview, ale i pro zpracování, vyhodnocování a interpretaci výsledků interview platí v určité míře vše, co bylo řečeno v souvislosti s dotazníkem (srov. oddíl 4.2).

4.4 TESTY V PEDAGOGICKÉM VÝZKUMU

Pojem „test" lze definovat jako „zkoušku, úkol, identický pro všechny zkoumané osoby s přesně vymezenými způsoby hodnocení výsledků a jejich číselného vyjadřování" (Michalička, 1969). Test tedy není jakákoli zkouška, nýbrž zkouška, na kterou jsou kladeny určitě nároky. Testy lze třídit podle různých kritérií. Obecně je přijímáno například dělení na testy schopností, testy osobnosti a testy výkonu.

Testy schopností jsou konstruovány tak, aby se pomocí nich poznalo, jaké schopnosti (předpoklady, dispozice) jedinec má pro řešení určitých úloh nebo situací určitého typu. Snad nejznámější jsou tzv. **testy inteligence**, které zjišťují obecné předpoklady jedince orientovat se v problémových situacích.

Testy osobnosti neměří výkony nebo schopnosti jedinců, nýbrž takové stránky osobnosti, jako jsou vlastnosti temperamentu, zaměření motivace, charakterové vlastnosti, úzkostnost či neuroticismus apod.

V pedagogických výzkumech se velmi často používají **testy výkonu**, které měří výkonnost jedince v určitých oblastech. Nejznámější a nejdůležitější z testů výkonu jsou **testy didaktické**.

4.4.1 DIDAKTICKÉ TESTY A JEJICH DRUHY

Pojem „didaktický test" (angl. *achievement test*) je sice u různých autorů definován různě, ale tato různá vymezení se shodují v tom, že se jedná o zkoušku, která se orientuje na objektivní zjišťování úrovně zvládnutí učiva u určité skupiny osob. Od běžné zkoušky se didaktický test ovšem liší zejména tím, že je navrhován, ověřován, hodnocen a interpretován podle určitých (předem stanovených) pravidel. Stručná a výstižná je definice didaktického testu, jak ji uvádí P. Byčkovský (1982): didaktický test je „nástroj systematického zjišťování (měření) výsledků výuky“. V této definici jsou pod pojmem **výsledky výuky** míněny změny v osobnostech žáků způsobené výukou. Systematičnost postupu je zajišťována tím, že didaktický test je navrhován, ověřován, skórován (bodován) a interpretován podle určitých (předem stanovených) pravidel.

V pedagogických výzkumech se setkáváme s různými druhy didaktických testů, které se liší tím, jaké informace pomocí nich získáváme. Při popisu základních typů didaktických testů budeme postupovat podle klasifikace, kterou navrhl P. Byčkovský (1982).

Testy rychlosti

U těchto testů se zjišťuje, jakou rychlostí je žák schopen řešit určitý typ testových úloh. Testy rychlosti mají pevně stanovený časový limit pro řešení a obsahují většinou velmi snadné úlohy. Předpokládá se, že všichni zkoušení žáci tyto úlohy zvládají a že se liší pouze v rychlosti řešení. Příkladem testu rychlosti je test rychlosti čtení, ve kterém měříme, kolik slov za minutu je žák schopen správně přečíst (aniž přitom přihlížíme ke kvalitě čtení), nebo test přepisu textu na psacím stroji, ve kterém měříme počet správných úhozů za minutu apod.

Testy úrovně

Většina testů používaných v současné době na našich školách jsou testy, které se svým charakterem blíží testům úrovně. Čisté testy úrovně nepoužívají žádné časové omezení (časový limit) a výkon v nich je dán pouze úrovní vědomostí nebo dovedností zkoušeného. Z praktických důvodů však bývá nutné s určitým (i když velmi volným) limitem pracovat zpravidla vždy.

Pokud testy úrovně používají časový limit, pak je volen tak, aby znamenal přerušení práce jen pro ty nejpomalejší žáky. Výzkumy ukazují, že tito nejpomalejší žáci mají ve většině případů také nejslabší vědomosti a ani při dalším prodloužení času nedosahují lepších výsledků. Úlohy jsou totiž v testu zpravidla řazeny se vzrůstající obtížností, takže velmi pomalý žák v okamžiku přerušení práce na testu řeší ty nejobtížnější úlohy, které by byl již sotva schopen vyřešit.

Někdy i testy úrovně používají rychlosti jako vedlejší kritérium pro hodnocení výkonu v testu. Je například možno použít systému kombinovaného hodnocení, při kterém žákovi, který vyřeší správně většinu testových úloh (např. 80 % a více), připíšeme za každou „ušetřenou“ minutu jeden bod navíc.

Testy standardizované

Didaktické testy, které jsou připravovány důkladněji a které také mají úplnější vybavení, se označují jako testy standardizované. Standardizovaný didaktický test je připravován profesionálně, je důkladně ověřen, takže jsou známy jeho základní vlastnosti. Tyto testy vydávají většinou specializované instituce (např. Psychodiagnostika Bratislava). Součástí příslušenství standardizovaného didaktického testu je testová příručka (manuál), ze které se uživatel dozví o vlastnostech testu, o jeho správném použití atd. Většinou je také k dispozici standard (testová norma) pro hodnocení dosažených výkonů.

Nestandardizované didaktické testy

Didaktické testy, u nichž nebyly realizovány všechny kroky obvyklé při přípravě a ověřování testů standardizovaných, označujeme jako testy nestandardizované (učitelské, neformální). Neproběhlo u nich ověřování na větším vzorku žáků, a nejsou tudíž známy všechny jejich vlastnosti. Tyto testy si připravují učitelé sami pro svoji vlastní potřebu. U testů nestandardizovaných není také k dispozici testová příručka ani objektivně stanovený testový standard (testová norma). I při konstrukci těchto testů by však učitelé měli dbát všech základních pravidel a zásad, které se doporučují u testů standardizovaných.

Někdy se užívá i termínu **testy kvazistandardizované**, čímž se rozumí testy připravované dokonaleji než testy učitelské, u nichž ale standardizace nebyla provedena beze zbytku. Kvazistandardizovaným testem je například didaktický test, zjišťující úroveň vědomostí žáků v daném předmětu na určité škole (několik paralelních tříd) nebo na několika školách. Konstrukci těchto testů bývá většinou věnována větší pozornost než u testů nestandardizovaných, bývají známy některé jejich vlastnosti a někdy bývají k dispozici i standardy pro hodnocení testových výsledků.

Testy kognitivní a testy psychomotorické

Dělení didaktických testů na kognitivní a psychomotorické vychází z dělení lidského učení do tří oblastí podle B. S. Blooma (učení kognitivní, afektivní a psychomotorické). Výsledky učení afektivního se didaktickými testy nezjišťují (k tomuto účelu se používají například dotazníky, různé škály apod.). Pokud didaktický test měří úroveň (kvalitu) poznání u žáků, jedná se o test kognitivní. Pokud testem zjišťujeme výsledky psychomotorického učení, jde o test psychomotorický. Příkladem kognitivních testů jsou například testy, ve kterých má žák řešit úlohy z matematiky, překládat text do cizího jazyka atd., příkladem testu psychomotorického je například test psaní na stroji. V současné pedagogické praxi se daleko častěji používají testy kognitivní, užití testu psychomotorického je spíše výjimkou.

Testy výsledků výuky a testy studijních předpokladů

V běžné pedagogické praxi se doposud téměř výlučně používají didaktické **testy výsledků výuky**, které měří to, co se žáci v dané oblasti naučili. **Testy studijních předpokladů** (*aptitude tests*) měří úroveň obecnějších charakteristik jedince, které jsou potřebné k dalšímu studiu. Testy studijních předpokladů by se měly používat zejména při přijímání žáků ke studiu na vyšší typ školy. Zatím však je praxe taková, že testy, které se k tomu účelu na našich školách používají, se příliš neliší od běžných testů, tj. od testů výsledků výuky. Konstrukce testů studijních předpokladů je podstatně náročnější a vyžaduje vedle pedagogické kvalifikace autora také dobrou kvalifikaci psychologickou.

Testy rozlišující (testy relativního výkonu)

Podle toho, jakým způsobem interpretujeme (vysvětlujeme a hodnotíme) výkon žáka v testu, můžeme rozlišit tzv. didaktické testy rozlišující (testy relativního výkonu) a didaktické testy ověřující (testy absolutního výkonu). Testy rozlišující se také označují jako testy statisticko-normativní anebo jako NR testy (*norm-referenced tests*).

Hlavní rozdíl mezi těmito dvěma druhy testů spočívá v tom, že u testů rozlišujících se výkon žáka určuje vzhledem k populaci testovaných, zatímco u testů ověřujících se výkon určuje vzhledem ke všem možným úlohám, které určité učivo reprezentují.

V naší pedagogické praxi se zatím používají téměř výlučně testy rozlišující. Základní ideou, o kterou se opírá koncepce rozlišujících didaktických testů, je snaha dosáhnout maximální možné objektivity a diferencovanosti hodnocení testových výkonů.

Výkon žáka v testu je srovnáván s výkony ostatních žáků, v případě standardizovaných rozlišujících testů s výkony celé žákovské populace. Rozlišující didaktické testy jsou tedy konstruovány tak, že umožňují rozhodnout, jakého výkonu v testu žák dosáhl vzhledem k celé populaci, k níž patří. Umožňují posoudit, zda určitý konkrétní žák je ve srovnání s ostatními žáky například „velmi slabý", „podprůměrný", „průměrný" atd.

Testy ověřující (testy absolutního výkonu)

Didaktické testy ověřující jsou často v literatuře označovány také jako kriteriální testy nebo CR testy (*criterion-referenced tests*). Úkolem ověřujících testů je prověřit úroveň vědomostí a dovedností žáka v přesně vymezené oblasti (části učiva). Výkon testovaného se přitom nesrovnává s výkonem jiných žáků (populace), nýbrž se vyjadřuje vůči všem úlohám, které reprezentují dané učivo.

U ověřujících testů je kritériem úspěchu předem stanovený stupeň zvládnutí učiva. Ověřující testy požadují u vybraných základních poznatků téměř úplné zvládnutí (četnost chyb musí zhruba odpovídat náhodě). Ověřující testy neusilují o diferencované hodnocení žáků, jako je tomu u testů rozlišujících. Cílem je v podstatě rozhodnout, zda žák zvládl učivo, nebo nikoli. Při konstrukci ověřujících testů je základním problémem výběr učiva, které musí žák bezpečně zvládnout. Toto učivo se potom transformuje do testových úloh. Aby se bezpečně ověřilo zvládnutí určitého učiva, požaduje se, aby každý testovaný jev byl pokryt větším počtem testových úloh.

V naší pedagogické praxi se ověřující testy zatím téměř neužívají. Jejich teorie se však rychle rozvíjí a zdá se, že rozšíření tohoto druhu testů bude pro testování vědomostí žáků přínosem.

Testy vstupní, průběžné a výstupní

Vstupní didaktické testy se zadávají na začátku výuky určitého celku učební látky a jejich cílem je postihnout úroveň vědomostí a dovedností, které jsou pro úspěšné zvládnutí daného celku učiva důležité. Zařazení vstupního testu na začátku výuky může posloužit jako zdroj cenných informací například v případě, že hodláme realizovat diferencovanou výuku.

Průběžné didaktické testy se zadávají v průběhu výuky a jejich posláním je poskytovat učiteli zpětnovazební informace potřebné k optimálnímu řízení výuky. Obvykle zkouší jen poměrně malou část učiva a jejich úkolem je sledovat, jak žáci probírané učivo přijímají, chápou a jak si je osvojují. V této souvislosti se často hovoří o tzv. **formativních testech**, které slouží ke sledování procesu formování vědomostí a dovedností u žáků. Tyto testy neslouží většinou k hodnocení žáků, nýbrž k hodnocení výuky.

Výstupní didaktické testy se zadávají buď na konci výukového období, nebo na konci určitého celku a většinou poskytují informace potřebné pro hodnocení žáků. Bývají také označovány jako **testy sumativní**.

Testy monotematické a polytematické

Monotematické testy zkouší jediné téma učební látky, testy polytematické zkouší učivo několika tematických celků. Testy polytematické jsou proto náročnější z hlediska přípravy i konstrukce.

Testy objektivně skórovatelné

Testy objektivně skórovatelné obsahují úlohy, u nichž lze objektivně rozhodnout, zda byly řešeny správně či nikoli. Výhodou těchto testů je, že skórování může provádět jakákoli osoba (někdy i stroj).

Vzhledem k tomu, že velká většina používaných didaktických testů se vyznačuje možností objektivního skórování, vznikla u části pedagogické veřejnosti nesprávná představa, že test je zkouška, která vždy obsahuje pouze objektivně hodnotitelné úlohy (např. úlohy, kde žák vybírá správnou odpověď, nebo úlohy, kde žák formuluje vlastní, ale velmi stručnou, a tudíž objektivně hodnotitelnou odpověď).

Testy subjektivně skórovatelné

Subjektivně skórovatelné testy (označované často jako esej testy) obsahují úlohy, u nichž není možno stanovit jednoznačná pravidla pro skórování. Mezi subjektivně skórovatelné testové úlohy patří například tzv. otevřené široké úlohy, ve kterých žák volně odpovídá na položenou otázku uvedením rozsáhlejší odpovědi. Ukazuje se, že není rozumné vyhýbat se používání takových úloh jen proto, že neumožňují objektivní skórování. Otevřené široké úlohy totiž mohou zkoušet daleko komplexnější vědomosti a dovednosti než úlohy objektivně skórovatelné.

4.4.2 TESTOVÉ ÚLOHY

Didaktický test je sestaven z jednotlivých **testových úloh**. Testovou úlohou rozumíme otázku, úkol nebo problém obsažený v testu. Kromě termínu „testová úloha" se často v literatuře používá výrazů **testová položka** nebo **testový úkol**. Na kvalitě testových úloh závisí v podstatné míře kvalita celého testování. Navrhování a konstrukce testových úloh v pedagogickém výzkumu je velmi náročná činnost, k jejímuž úspěšnému zvládnutí je potřeba kromě zkušeností také náležitého teoretického poučení. Autor didaktického testu by měl být odborníkem v oblasti, pro kterou test připravuje, dobrým pedagogem, ale měl by mít také základní vědomosti ze statistiky.

Podle způsobu, kterým testovaná osoba úlohu řeší, se testové úlohy rozdělují na úlohy otevřené (s tvořenou odpovědí, s volnou odpovědí) a na úlohy uzavřené (s nabízenou odpovědí, s nucenou volbou odpovědí). Úlohy otevřené lze dále podle rozsahu požadované odpovědi rozdělit na široké a na úlohy se stručnou odpovědí. Uzavřené úlohy dělíme na dichotomické, s výběrem odpovědí, přiřazovací a uspořádací.

4.4.2.1 Otevřené široké úlohy

V otevřených širokých úlohách se požaduje od žáka rozsáhlejší odpověď (např. polovina strany nebo i delší). Může se například požadovat pojednání na určité téma (např. „Význam díla K. H. Máchy pro českou poezii", „Které byly hlavní příčiny vzniku první světové války?"), vyřešení určitého problému (např. „Navrhněte postup, kterým je možno určit hustotu neznámé kapaliny"), popis určitého procesu (např. „Popište činnost čtyřdobého zážehového motoru") apod. Požadovaný rozsah odpovědi se žákovi naznačuje velikostí vynechaného místa v testovém zadání.

U otevřených širokých úloh bývá vhodné vymezit strukturu požadované odpovědi, např.: „Výroba surového železa (uveďte hlavní používané suroviny, nakreslete schéma vysoké pece a popište hlavní probíhající chemické reakce").

Otevřené široké úlohy lze doporučit zejména při zkoušení komplexních vědomostí nebo dovedností, osvojovaných v delším časovém období. Jsou vhodné pro zkoušení vyšších úrovní osvojení učiva (např. řešení problémových situací apod.), zatímco pro zkoušení nižších úrovní osvojení učiva (např. zapamatování atd.) jsou výhodnější testy s úlohami objektivně skórovatelnými. Skutečnost, že široké testové úlohy nelze objektivně skórovat, by neměla vést k jejich nepoužívání. Komplexní vědomosti a dovednosti lze sice rozložit na dovednosti dílčí (které se dají zkoušet mnohem snadněji), avšak zvládnutí dílčích dovedností ještě nemusí znamenat zvládnutí dovedností komplexních.

Široké testové úlohy se navrhují poměrně snadno, ale jejich hlavní nevýhodou je nemožnost objektivního skórování. Při skórování širokých úloh se často postupuje tak, že za správné a úplné zodpovězení úlohy se přisuzuje určitý počet bodů (např. deset). Za každou chybnou nebo chybějící část odpovědi se potom strhává určitý počet bodů. V některých případech lze pro skórování otevřené testové úlohy vypracovat detailní předpis, který umožňuje téměř objektivní skórování.

Testy vytvořené z otevřených širokých úloh se někdy označují jako **esej testy**. Tyto testy obsahují zpravidla jen několik úloh. Od běžných písemných zkoušek se esej testy liší tím, že se při jejich konstrukci, hodnocení i při interpretaci výsledků využívají všechna základní pravidla a postupy obvyklé u ostatních didaktických testů.

4.4.2.2 Otevřené úlohy se stručnou odpovědí

Úlohy se stručnou odpovědí požadují od žáka vytvoření a uvedení vlastní krátké odpovědi. Může se požadovat např. uvedení čísla, značky, symbolu, vzorce, jednoduchého grafu, určitého slova, případně několika slov či krátké věty. Úlohy se stručnou odpovědí mohou být dvojího typu: produkční a doplňovací.

Příklady produkčních testových úloh:

Co je jednotkou elektrického napětí? ....
Napište Archimédův zákon ....

Příklady doplňovacích testových úloh:

Hlavním městem Norska je ....
Ptáci zpíval…, včely bzučel… a medvíďata se batolil… po lese.

Mezi výhody úloh se stručnou odpovědí lze počítat zejména to, že se snadno navrhují. Další jejich výhodou je, že neumožňují testovaným osobám tak snadno uhodnout správnou odpověď bez příslušných vědomostí, jako je tomu u úloh s výběrem odpovědí. Většinou se předpokládá, že vytvoření odpovědi je pro testované náročnější než pouhé rozpoznání správné odpovědi mezi nabídnutými alternativami.

4.4.2.3 Dichotomické úlohy

U dichotomických testových úloh jsou testovaným osobám předkládány dvě alternativy odpovědi s tím, že jedna je správná a tu mají označit (např. podtržením, zakroužkováním apod.). Tyto úlohy bývají také někdy označovány jako úlohy s dvoučlennou volbou nebo jako alternativní úlohy (*true-false item*).

Příklady dichotomických testových úloh:

Mistr Jan Hus byl upálen roku 1515: ano – ne
Při vypařování kapaliny se teplo: spotřebovává – uvolňuje

Výhodou dichotomických úloh je, že se velmi snadno navrhují. Tato konstrukční jednoduchost však může svádět k testování jednotlivých detailů, pouhých faktů. Nedostatkem dichotomických úloh je velká pravděpodobnost uhodnutí správné odpovědi i bez příslušných vědomostí. Aby se věrohodnost výsledků získaných testem s dichotomickými úlohami zvýšila, je třeba, aby test obsahoval těchto úloh dostatečný počet.

4.4.2.4 Úlohy s výběrem odpovědí

Úlohy s výběrem odpovědí (v literatuře označované také jako úlohy s vícečlennou či vícenásobnou odpovědí, úlohy polynomické, *multiple-choice*) vděčí za svoji teoretickou rozpracovanost především rozvoji programovaného učení (hlavně tzv. větvených programů). Úloha s výběrem odpovědí se skládá ze dvou částí: problému nebo otázky (tzv. kmenu úlohy) a nabídnutých odpovědí.

Úlohy s výběrem odpovědí se v didaktických testech vyskytují v několika formách.

Úlohy typu „jedna správná odpověď"

U základní formy těchto úloh testované osoby vybírají z několika nabídnutých alternativ **jednu správnou odpověď**.

Příklad úlohy s výběrem odpovědí (jedna odpověď správná):

Cesta <u>vlakem</u> nám velmi rychle uběhla.
Podtržený větný člen je:
A podmět
B předmět
C přívlastek
D příslovečné určení

Úlohy typu „jedna nejpřesnější odpověď"

U těchto testových úloh se požaduje označení nejlepší nebo nejsprávnější odpovědi. Takové úlohy jsou často pro testované osoby velmi obtížné (obtížnější než jim odpovídající úlohy produkční).

Příklad úlohy s výběrem odpovědí (jedna nejpřesnější odpověď):

Které z následujících tvrzení nejlépe odpovídá na otázku „Co je chemický prvek?“
A Prvek je látka, která se skládá z atomů stejného druhu.
B Prvek je látka, kterou již dále nelze dělit.
C Prvek je látka složená z atomů, které mají stejné protonové číslo.
D Žádné z předchozích tvrzení není správné.

Úlohy typu „jedna nesprávná odpověď“

V některých případech lze také požadovat uvedení **nesprávné odpovědi**. V tomto případě je ovšem nutné zápor ve kmenu úlohy patřičně zdůraznit, protože jinak může snadno dojít k přehlédnutí a nesprávné odpovědi přesto, že testovaná osoba má příslušné vědomosti.

Příklad úlohy s výběrem odpovědí (jedna nesprávná odpověď):

Který z následujících dějů ***není*** *formou oxidačního procesu?*
A dýchání
B hnití
C destilace
D rezivění

Úlohy s vícenásobnou odpovědí

Jestliže má testovaná osoba v úloze vybrat několik správných odpovědí, hovoříme o tzv. **vícenásobné odpovědi**. Jestliže se rozhodneme pro použití tohoto typu úlohy, je nutné na to testované předem upozornit. Výběr většího počtu odpovědí může totiž být považován v těchto případech za chybu.

Příklad úlohy s vícenásobnou volbou odpovědi:

Kterými státy protéká (nebo kterých se alespoň dotýká) řeka Odra?
*A Německo **
B Rusko
*C Česká republika **
D Slovensko
*E Polsko **

U tohoto typu úloh mohou někdy vznikat problémy při skórování. Neexistuje totiž jen jedna naprosto správná a jedna naprosto nesprávná odpověď, nýbrž několik částečně správných odpovědí. Lze doporučit dva přístupy, z nichž první lze stručně vyjádřit slovy **„všechno, anebo nic“**. Podle tohoto přístupu přidělíme 1 bod v případě, kdy žák označí všechny správné odpovědi a 0 bodů tehdy, jestliže bude (třeba jen jedna) odpověď nesprávná. Druhý přístup (diferencovanější) spočívá v tom, že přidělíme 1 **„pomocný“** bod za každou označenou správnou odpověď a 1 pomocný bod za každou neoznačenou nesprávnou odpověď. Výsledný součet pomocných bodů potom dělíme počtem nabídek v úloze (maximální bodový zisk v úloze je 1).

Situační úlohy

Zvláštní modifikací testových úloh s výběrem odpovědí jsou úlohy označované jako **úlohy situační** či **interpretační**. Jsou to úlohy, u nichž testovaná osoba vybírá z podstatně většího počtu nabídek, než je obvyklé, přičemž nabídky nejsou předkládány ve formě dlouhého a nepřehledného seznamu, nýbrž vyplynou přímo z dané situace. Pravděpodobnost uhodnutí správné odpovědi bez příslušných vědomostí je u tohoto typu úloh zpravidla velmi malá.

Příklad situační úlohy:

> *Na místo označené hvězdičkou napište takovou číslici, aby výsledné šesticiferné číslo bylo dělitelné sedmi: 823*43*

V této testové úloze nejsou na první pohled nabízeny žádné alternativy odpovědí, ale ve skutečnosti testovaná osoba (která správnou odpověď nezná) vybírá z deseti číslic. Pravděpodobnost náhodné volby správné odpovědi je v tomto případě jen 10 %.

Problém „hádání" správných odpovědí (v úlohách s výběrem odpovědí)

U úloh s výběrem odpovědí existuje vždy určitá pravděpodobnost, že testovaná osoba zvolí správnou odpověď zcela náhodně. Toto nebezpečí se zmenšuje s rostoucím počtem nabízených odpovědí. Jako optimální počet předkládaných odpovědí se uvádí čtyři až pět, praxe se však ustálila na čtyřech odpovědích. Menší počet odpovědí než čtyři se pro velkou pravděpodobnost uhodnutí správné odpovědi nedoporučuje, více než pět odpovědí činí zase úlohu nepřehlednou a také sestavování většího počtu přijatelných odpovědí je značně obtížné.

Pro dvě a tři nabízené odpovědi se někdy doporučuje tzv. **korekce na hádání**. Při používání korekce na hádání se přisoudí žákovi počet bodů podle toho, kolika chyb se v testu dopustil. Vychází se z toho, že žák, který odpověď hádá, se dopouští častěji chyb než ten, který úlohy skutečně řeší a odpovídá jedině tehdy, když odpověď zná. Korekci dosažených bodových výsledků lze provést podle vzorce

$$s_o = s_n - \frac{n}{y-1} \tag{96}$$

kde s_o je tzv. opravené skóre (opravený počet bodů), s_n je neopravené skóre, n je počet nesprávných odpovědí určitého žáka v testu a y je počet nabízených odpovědí v jedné úloze.

Příklad 57

V didaktickém testu, který byl sestaven z 28 dichotomických úloh, určitý žák uvedl 16 správných odpovědí a 6 nesprávných odpovědí (v 6 úlohách neodpověděl). Dosadíme-li tyto hodnoty do vztahu (96) dostáváme

$$s_o = 16 - \frac{6}{2-1} = 10$$

Žákovi proto přisoudíme (přestože odpověděl v 16 úlohách správně) pouze 10 bodů.

Jestliže se provádí korekce na hádání, je nutno na tuto skutečnost testované osoby upozornit. Je třeba jim vysvětlit, že ve sporných případech je pro ně výhodnější neodpovídat vůbec. Naopak v testech, kde se korekce na hádání neprovádí, je pro zkoušeného výhodnější zodpovědět všechny úlohy. Zkušenosti ukazují, že provádění korekce na hádání je sice „statisticky spravedlivé", ale nemusí být spravedlivé vzhledem k jednotlivým žákům. Při tomto postupu dochází například k poškozování žáků, kteří jsou ke své práci velmi kritičtí. Užití korekce na hádání ve výzkumech má však své opodstatnění, protože umožňuje objektivně posoudit úroveň vědomostí celé skupiny testovaných osob.

V souvislosti s problémem hádání správných odpovědí v testu je třeba zvážit, kolik úloh musí test obsahovat, aby informace jím získaná byla dostatečně spolehlivá a přesná. Obsahuje-li test úlohy dichotomické, je pravděpodobnost uhodnutí správně odpovědi v jedné úloze 50 %, což znamená, že v průměru každý druhý testovaný může správnou odpověď uhodnout. Zvětšujeme-li počet úloh v testu, roste spolehlivost celkových výsledků, protože uhodnutí správné odpovědi ve všech úlohách (při dostatečném počtu úloh) je velmi nepravděpodobné. Obecně platí, že test musí obsahovat tím více úloh, čím menší je počet nabízených odpovědí.

Základním konstrukčním požadavkem na úlohy s výběrem odpovědí je, aby všechny předkládané nabídky odpovědí (distraktory) byly pro testované (kteří neznají správnou odpověď) stejně přijatelné (plauzibilní).

4.4.2.5 Přiřazovací úlohy

Přiřazovací úlohy obsahují dvě množiny pojmů a instrukci. Úkolem testovaného je správně přiřadit pojmy jedné množiny k pojmům množiny druhé.

Příklad přiřazovací úlohy:

K názvům států v levém sloupci přiřaďte názvy jejich hlavních měst z pravého sloupce:

Švýcarsko …	*A Dublin*
Norsko …	*B Oslo*
Island …	*C Bonn*
Nizozemsko …	*D Bern*
Finsko …	*E Amsterdam*
	F Reykjavík
	G Helsinky

V uvedeném příkladu je počet nabízených hlavních měst úmyslně větší než počet uvedených států. Tuto zásadu bychom měli dodržovat při navrhování všech úloh tohoto typu. Vyhneme se tím situaci, že ze znalosti některých přiřazení již vyplývají přiřazení další. Výhodou přiřazovacích testových úloh je, že omezují možnost náhodného uhodnutí správné odpovědi na minimální míru. Jejich použití je však možně jen v poměrně omezeném okruhu učiva.

4.4.2.6 Uspořádací úlohy

V uspořádacích úlohách se od testovaného požaduje, aby uspořádal prvky množiny pojmů jedné třídy do řady. Prvky se seřazují podle jistého hlediska, například chronologicky, podle velikosti, podle stupně obecnosti atd.

Příklad uspořádací úlohy:

> *Seřaďte uvedená města podle počtu obyvatel. Seřazení proveďte pomocí pořadových čísel 1–6 tak, aby město s největším počtem obyvatel mělo číslo 1:*
> *Londýn …*
> *Tokio …*
> *Káhira …*
> *New York …*
> *Sao Paulo …*
> *Sydney … …*

Nevýhodou této formy testových úloh je omezená oblast použití, ale také obtížné skórování (vyplývá ze skutečnosti, že nesprávné přiřazení prvků může být provedeno mnoha způsoby, přičemž jde o různě velké chyby).

4.4.3 KONSTRUKCE DIDAKTICKÉHO TESTU

Konstrukci didaktického testu bychom neměli začínat navrhováním testových úloh. Tento postup nejčastěji vede k tomu, že vytvoříme testové úlohy, které se sice snadno navrhují, ale ve svém celku vytvářejí nevyvážený didaktický test. Ten zpravidla nepokrývá celý obsah, který má být zkoušen a zaměřuje se většinou na pouhou reprodukci zapamatovaných poznatků. Nejsnadněji se totiž navrhují testové úlohy, které zkoušejí pouhé zapamatování faktů. Těchto úloh bývá v nekvalifikovaně sestavených testech velká převaha. Chceme-li vytvořit didaktický test skutečně kvalitní, musíme věnovat dostatečnou pozornost jeho plánování.

Prvním krokem při konstrukci didaktického testu je rozhodnutí, k jakému **účelu** má didaktický test sloužit. Testem lze zjišťovat například výsledky výuky na konci určitého tematického celku nebo na konci pololetí či roku, testem lze ale také zjišťovat, jak žáci přijímají a chápou probírané učivo atd.

Dalším krokem je stanovení obsahu, který má test zkoušet. Zpravidla se postupuje tak, že učivo, které má být předmětem testování, se nejdříve rozčlení na určité prvky (fakta, pojmy, vztahy, definice apod.). Každému prvku učiva se potom přidělí určitý počet testových úloh. U jednotlivých úloh se zároveň stanoví, jakou úroveň osvojení poznatků mají zkoušet, tj. zda mají postihovat jen pouhé zapamatování poznatků, nebo také porozumění poznatkům, používání vědomostí k řešení problémových situaci apod. V tomto směru jsou dobrou pomocí různé osvědčené taxonomie výukových cílů (např. Bloomova nebo

Niemierkova taxonomie výukových cílů). Pro upřesnění (specifikaci) obsahu testu lze použít některou z propracovaných technik (např. techniku specifikační tabulky). Potřebné informace lze nalézt například v práci M. Chrásky (1999).

Po stanovení obsahu testu je možné přistoupit k návrhu testových úloh. Podle charakteru testovaného učiva a podle cíle, který má test plnit, se můžeme rozhodnout pro úlohy otevřené, nebo uzavřené. Chceme-li vytvořit kvalitní test, lze doporučit, aby navržené testové úlohy byly po určité době (alespoň několika dnů) znovu autorem posouzeny. Při posuzování úloh si vedle technické kvality úloh všímáme zejména toho, jakou úroveň osvojení poznatků úlohy zkouší. Jako další krok při konstrukci didaktického testu se doporučuje posouzení navržených úloh dalšími odborníky – kompetenty. Jako kompetentů lze využít například učitelů, kteří realizují výuku oboru, v němž testování provádíme.

Na základě vlastního posouzení a na základě doporučení kompetentů provede autor závěrečnou úpravu prototypu testu. Ta spočívá v první řadě ve vyřazení těch úloh, které zjevně vykazují nevhodné vlastnosti. Úlohy se většinou v testu řadí podle vzrůstající obtížnosti, a to tak, že první úlohy jsou nejsnadnější a poslední nejobtížnější (podle subjektivního posouzení autora, respektive kompetentů).

Součástí přípravy testu pro první použití je i předběžné určení času, který bude k vypracování testu třeba, a vypracování pokynů pro práci testovaných.

4.4.4 OVĚŘOVÁNÍ VLASTNOSTÍ DIDAKTICKÉHO TESTU

Má-li vytvořený didaktický test spolehlivě měřit vědomosti, je třeba jej ověřit na přiměřeně velkém vzorku testovaných osob. Teprve při tomto ověřování lze relativně definitivně rozhodnout o vlastnostech testu. Při ověřování didaktického testu se posuzují vlastnosti jednotlivých testových úloh, ale i vlastnosti didaktického testu jako celku.

4.4.4.1 Analýza vlastností testových úloh

Obtížnost testových úloh

Při analýze obtížnosti se vypočítává buď **hodnota obtížnosti *Q*,** nebo **index obtížnosti *P*.** Hodnota obtížnosti udává procento testovaných ve vzorku, kteří danou úlohu zodpověděli nesprávně, anebo ji vynechali

$$Q = 100\frac{n_n}{n} \tag{97}$$

kde Q je hodnota obtížnosti, n_n je počet testovaných ve skupině, kteří odpověděli nesprávně nebo neodpověděli, a n je celkový počet testovaných ve vzorku.

Index obtížnosti je procento testovaných ve skupině, kteří danou úlohu zodpověděli správně

$$P = 100\frac{n_s}{n} \tag{98}$$

kde *P* je index obtížnosti, n_s je počet testovaných ve skupině, kteří odpověděli v dané úloze správně, a *n* je celkový počet testovaných ve skupině.

O vysoké obtížnosti testové úlohy vypovídají vysoké hodnoty obtížnosti *Q* a naopak nízké hodnoty indexu obtížnosti *P*. Za velmi obtížné lze pokládat testové úlohy, u nichž hodnota obtížnosti *Q* je vyšší než 80. Velmi snadné jsou naopak ty úlohy, které vykazují hodnotu obtížnosti *Q* nižší než 20. Velmi obtížných (ale ani velmi snadných úloh) by nemělo být v testu příliš mnoho. Úlohy extrémně obtížné, u nichž se hodnota obtížnosti *Q* blíží ke 100, jsou nevyhovující a je nutno je z testu vyloučit. Úlohu extrémně snadnou, u níž se hodnota obtížnosti *Q* blíží k nule, je možno z psychologických důvodů doporučit jako úvodní úlohu v testu. Může totiž přispět k uklidnění testovaných osob a k vytvoření potřebného pocitu jistoty. Zkušenosti ukazují, že nejvhodnější vlastnosti mají testové úlohy s hodnotou obtížnosti kolem $Q = 50$ (platí pro testy rozlišující).

Citlivost testových úloh

Citlivost úloh bývá často označována také jako rozlišovací hodnota, diskriminační hodnota, rozlišovací ostrost nebo jako rozlišovací schopnost úloh. Vysokou citlivost má úloha, kterou řeší s úspěchem testované osoby, které mají celkově lepší vědomosti, zatímco osoby, které mají celkově horší vědomosti, v této úloze dosahují výsledků špatných. Citlivost úlohy tedy vyjadřuje, jak dalece daná úloha zvýhodňuje testované osoby s lepšími vědomostmi před osobami, které mají vědomosti horší. K rozlišení testovaných osob na osoby **„s lepšími vědomostmi“** a na osoby **„s horšími vědomostmi“** se většinou používá celkových výsledků v ověřovaném didaktickém testu.

Při posuzování citlivosti úloh se většinou nejdříve vzorek testovaných osob rozdělí podle celkového počtu dosažených bodů (hrubého skóre) na dvě části: skupinu **„lepších“** (s vyšším počtem dosažených bodů) a skupinu **„horších“** (s nižším počtem dosažených bodů). Testované osoby se seřadí podle dosaženého celkového počtu bodů v testu, přičemž horní polovinu označíme jako **„lepší“** *(L)* a spodní polovinu jako **„horší“** *(H)*. Někdy je možné obě skupiny žáků vytvořit i z menšího počtu osob, například z 33 % nejlepších a 33 % nejhorších testovaných osob apod.

Citlivost úlohy se dá exaktně posoudit pomocí výpočtu některého z koeficientů citlivosti, kterých byla navržena celá řada. Všechny tyto koeficienty mohou nabývat hodnot od –1 přes nulu do +1. Platí, že čím vyšší hodnotu koeficient má, tím lépe úloha rozlišuje mezi osobami s lepšími vědomostmi a mezi osobami s horšími vědomostmi. Pokud koeficient citlivosti dosahuje hodnoty 0, znamená to, že úloha vůbec nerozlišuje mezi oběma skupinami osob (osoby s lepšími i osoby s horšími vědomostmi jsou v této úloze stejně úspěšné). Záporné hodnoty koeficientu citlivosti vypovídají o tom, že úloha zvýhodňuje osoby, které mají v testu celkově horší výsledky. Kladné hodnoty koeficientu citlivosti naopak vypovídají o tom, že v úloze dosahují lepších výsledků osoby, které mají v testu lepší celkové výsledky (úloha zvýhodňuje osoby, které mají lepší vědomosti).

Velice nízkých (případně záporných) hodnot nabývá koeficient citlivosti například v úlohách, které jsou příliš komplikovaně formulovány a kdy obě skupiny žáků (skupina L a skupina H) mohou volit rozdílné strategie řešení. Zatímco žáci **„lepší“** se snaží k řešení komplikované úlohy dojít složitými úvahami, při kterých se snadno dopouštějí chyb, žáci **„horší“** se složitými úvahami příliš neznepokojují a odpověď s větším či menším štěstím hádají. Nízké hodnoty koeficientu citlivosti se objevují také u některých velmi obtížných úloh s výběrovými odpověďmi nebo u úloh, kde se zkouší formálně zvládnuté učivo.

Výpočet koeficientu citlivosti ULI

Nejjednodušším ukazatelem citlivosti testové úlohy je koeficient ULI (*upper-lower-index*). Tento koeficient vychází z rozdílu mezi obtížností úlohy ve skupině lepších a ve skupině horších testovaných osob

$$d = \frac{n_L - n_H}{0,5N} \tag{99}$$

kde d je koeficient citlivosti ULI, n_L je počet osob z lepší skupiny, které danou úlohu zodpověděly správně, n_H je počet osob ze skupiny horších, které úlohu řešily správně, a N je celkový počet testovaných osob. Uvedený vztah platí pro případ, že obě skupiny byly vytvořeny na základě rozdělení všech testovaných osob podle celkového dosaženého počtu bodů na polovinu.

U koeficientu ULI se požaduje, aby v případě úloh s hodnotou obtížnosti 30–70 bylo d alespoň 0,25 a u úloh s hodnotou obtížnosti 20–30 a 70–80 alespoň 0,15.

Kromě koeficientu ULI se často používají i další koeficienty citlivosti, například **tetrachorický koeficient citlivosti** r_{tet}, **biseriální korelace** r_{bis} apod. (srov. oddíly 3.4.4 a 3.4.6).

Analýza nenormovaných odpovědí

Vedle posuzování obtížnosti a citlivosti testových úloh se v rámci analýzy vlastností úloh provádí také analýza tzv. **nenormovaných odpovědí** (Byčkovský, 1982), tj. rozbor odpovědí vynechaných nebo nesprávných.

Jestliže zjistíme, že v některých úlohách jsou často vynechávány odpovědi, může to znamenat vedle neznalosti učiva také nepochopení formulace úlohy, nedostatek času k vypracování odpovědí atd. V literatuře se uvádí, že je třeba věnovat zvýšenou pozornost zejména těm otevřeným úlohám, ve kterých odpověď vynechalo více než 30–40 % žáků. U uzavřených úloh je však třeba věnovat zvýšenou pozornost i úlohám, kde neodpovědělo více než 20 % žáků.

Rozbor nesprávných odpovědí je velmi jednoduchý u úloh s výběrem odpovědí. V tomto případě postačí zkontrolovat, zda všechny nabídnuté **distraktory** (nesprávné nabídky) jsou pro testované osoby dostatečně atraktivní. To se pozná podle toho, že osoby vybírají skutečně ze všech nabídnutých možností. Ten distraktor, který nikdo (nebo téměř nikdo) nevolí, neplní svoji funkci a měl by být (pokud to je z obsahového hlediska možné) nahrazen jiným, atraktivnějším distraktorem, anebo z úlohy odstraněn.

V případě otevřených testových úloh je možno postupovat tak, že všechny chyby rozdělíme do dvou kategorií, tzv. **základní chyby** a **vedlejší chyby**. Za základní chyby považujeme ty, které jsou způsobené skutečnou neznalostí učiva (jeho nepochopením nebo nezvládnutím). Vedlejší chyby jsou způsobené různými náhodnými vlivy (např. přehlédnutím, numerickou chybou ve výpočtu, nepřesností, špatnou čitelností textu atd.). Jestliže v určité testové úloze převažují vedlejší chyby nad hlavními, může to znamenat, že v této úloze je úspěch řešení věcí náhody. Úlohy, u kterých převažují vedlejší chyby nad základními, by se měly z testu odstranit.

4.4.4.2 Analýza vlastností testu jako celku

Při ověřování didaktického testu věnujeme pozornost také vlastnostem testu jako celku. Základními vlastnostmi didaktického testu jsou validita, reliabilita a praktičnost.

Validita didaktického testu

Validita je základní a nejdůležitější vlastností didaktického testu. Test je dostatečně validní tehdy, pokud se jím zkouší skutečně to, co má být zkoušeno.

U testů studijních výsledků posuzujeme, jak dalece se shoduje obsah testu s cíli a obsahem vyučování. V těchto případech nám jde především o tzv. **obsahovou validitu** testu. Obsah úloh didaktického testu by měl být v těchto případech reprezentativním vzorkem zkoušeného učiva. V praxi se neužívají žádné kvantitativní metody určování validity a při jejím posuzování jsme odkázáni jen na posudek odborníků.

Při posuzování validity testů studijních předpokladů je situace podstatně složitější. Je totiž nutno zvážit, ze kterých vědomostí nebo dovedností můžeme usuzovat na budoucí úspěšnost ve studiu. Testy studijních předpokladů by měly mít především dobrou tzv. **predikční validitu**, tj. schopnost předpovídat budoucí úspěšnost v učení.

Reliabilita didaktického testu

Didaktický test má dobrou reliabilitu tehdy, pokud poskytuje spolehlivé a přesné výsledky. **Spolehlivost** testu spočívá v tom, že při opakování testování za týchž podmínek získáváme stejné nebo velmi podobné výsledky. **Přesnost** testu souvisí s velikostí a četností chyb při testování.

K exaktnímu hodnocení stupně reliability didaktického testu slouží **koeficient reliability**. Koeficient reliability může nabývat hodnot od 0 (pro případ naprosté nespolehlivosti a nepřesnosti didaktického testu) do +1 (pro případ maximální spolehlivosti a přesnosti didaktického testu). Pro individuální testování vědomostí se většinou u didaktických testů požaduje koeficient reliability minimálně 0,80. Čím nižší je koeficient reliability, tím skeptičtěji nutno posuzovat naměřené výsledky.

Stupeň reliability didaktického testu závisí v podstatné míře na počtu úloh v testu. Obecně platí, že čím více úloh test obsahuje, tím vyšší má reliabilitu. U testů s malým počtem úloh (např. deset nebo méně) koeficient reliability zpravidla dosahuje maximálně hodnoty kolem 0,60. Aby byl didaktický test dostatečně validní, musí mít dobrou reliabilitu. Vysoká reliabilita však ještě není zárukou toho, že didaktický test bude validní.

Koeficient reliability je možno vypočítat mnoha způsoby. Často se užívá například tzv. **Kuderův-Richardsonův vzorec** nebo **metoda půlení testu**.

Kuderův-Richardsonův vzorec

Tento model výpočtu koeficientu reliability je vhodný pro didaktické testy úrovně, které jsou složeny z obsahově homogenních úloh. U testů vytvořených z obsahově nehomogenních úloh vychází touto metodou koeficient reliability příliš nízký. Výpočet koeficientu reliability se provádí pomocí Kuderova-Richardsonova vzorce

$$r_{kr} = \frac{k}{k-1}\left(1 - \frac{\sum pq}{s^2}\right) \tag{100}$$

kde k je počet úloh v testu, p je podíl žáků ve vzorku, kteří řešili určitou úlohu v testu správně, $q = 1 - p$ a s je směrodatná odchylka pro celkové výsledky žáků v testu.

Příklad 58

Vzorku 339 žáků základní školy byl zadán didaktický test z fyziky, který obsahoval deset úloh s výběrem odpovědí. Při analýze výsledků testování byly získány údaje, které uvádí tabulka 69. Z údajů uvedených v tabulce byl vypočítán aritmetický průměr a směrodatná odchylka s (srov. oddíly 2.3.3.1 a 2.3.4.2).

Při analýze vlastností testových úloh bylo také zjištěno, kolik žáků řešilo správně jednotlivé úlohy (tab. 70).

Z uvedených údajů vypočítáme koeficient reliability.

Tab. 69 *Výpočet aritmetického průměru a směrodatné odchylky pro výsledky testování*

Počet bodů x_i	**Četnost** n_i	$n_i \cdot x_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$n_i(x_i - \bar{x})^2$
0	8	0	–4,773	22,782	182,256
1	18	18	–3,773	14,236	256,248
2	33	66	–2,773	7,690	253,770
3	44	132	–1,773	3,144	138,336
4	54	216	–0,773	0,598	32,292
5	57	285	0,227	0,052	2,964
6	50	300	1,227	1,506	75,300
7	31	217	2,227	4,960	153,760
8	20	160	3,227	10,414	208,280
9	16	144	4,227	17,868	285,888
10	8	80	5,227	27,322	218,576
Σ	339	1 618			1 807,670

$$\bar{x} = \frac{1\,618}{339} = 4{,}773$$

$$s^2 = \frac{1\,807{,}670}{339 - 1} = 5{,}33 \quad s = 2{,}31$$

Tab. 70 *Hodnoty p a q pro Kuderův-Richardsonův vzorec*

Úloha č.	Počet správných odpovědí	*p*	*q*	*pq*
1	186	0,55	0,45	0,248
2	163	0,48	0,52	0,250
3	207	0,61	0,39	0,238
4	237	0,70	0,30	0,210
5	146	0,43	0,57	0,245
6	170	0,50	0,50	0,250
7	149	0,44	0,56	0,246
8	159	0,47	0,53	0,249
9	139	0,41	0,59	0,242
10	54	0,16	0,84	0,134
				∑ 2,312

Podíl žáků, kteří určitou úlohu odpověděli správně, se určí ze vztahu

$$p = \frac{n_s}{n} \tag{101}$$

kde n_s je počet žáků, kteří určitou úlohu řešili správně a n je celkový počet žáků.

Dosadíme-li všechny potřebné hodnoty do Kuderova-Richardsonova vzorce, dostáváme koeficient reliability

$$r_{kr} = \frac{10}{10-1}\left(1 - \frac{2,312}{2,31^2}\right) = 0,63$$

Vypočítaný koeficient reliability svědčí (s přihlédnutím k tomu, že test obsahoval jen deset úloh) o přijatelném stupni reliability provedeného měření.

Reliabilita didaktického testu metodou půlení

Výhodou této metody je, že se dá použít jak pro testy úrovně, tak pro testy rychlosti. Podmínkou pro použití tohoto modelu výpočtu je, že didaktický test obsahuje sudý počet úloh a jednotlivé úlohy jsou řazeny podle vzrůstající obtížnosti.

Při výpočtu se postupuje většinou tak, že celý test se rozdělí na dvě poloviny, a to tak, že jednu polovinu tvoří úlohy s lichým pořadovým číslem a druhou polovinu úlohy se sudým pořadovým číslem. Výsledky dosažené jednotlivými žáky v obou polovinách testu se potom navzájem korelují. Z hodnoty vypočítaného korelačního koeficientu se vychází při stanovení koeficientu reliability. Samotný výpočet koeficientu reliability metodou půlení se provádí pomocí Spearmanova-Brownova vzorce

$$r_{sb} = \frac{2 \cdot r_p}{1 + r_p} \tag{102}$$

kde r_{sb} je koeficient reliability a r_p je koeficient korelace mezi výsledky žáků v obou polovinách didaktického testu.

Příklad 59

Didaktický test, jehož výsledky byly analyzovány v předcházejícím příkladě, byl rozdělen na dvě poloviny tak, že úlohy č. 1, 3, 5, 7 a 9 vytvořily polovinu L (liché) a úlohy č. 2, 4, 6, 8 a 10 polovinu S (sudé). U každého žáka bylo stanoveno, jakého počtu bodů (hrubého skóre) dosáhl v obou polovinách testu (tab. 71). Výsledky žáků v polovině L jsou označeny x_L a výsledky žáků v polovině S jsou označeny x_S.

Tab. 71 *Výsledky žáků v polovinách didaktického testu*

Jméno žáka	Celkový počet bodů	x_L	x_S	$x_L \cdot x_S$	x_L^2	x_S^2
1. …………	10	5	5	25	25	25
2. …………	10	5	5	25	25	25
3. …………	10	4	6	24	16	36
4. …………	10	3	7	21	9	49
5. …………..	10	6	4	24	36	16
6. …………	10	7	3	21	49	9
7. …………	10	5	5	25	25	25
8. …………	9	5	4	20	25	16
9. …………	9	6	3	18	36	9
10. …………	9	7	2	14	49	4
11. …………	9	4	5	20	16	25
12. …………	8	3	5	15	9	25
13. …………	8	4	4	16	16	16
14. …………	8	4	4	16	16	16
15. …………	7	4	3	12	16	9
338. …………	3	1	2	2	1	4
339. …………	2	1	1	1	1	1
Σ	1 618	716	902	3 197	4 387	5 143

Z hodnot x_L a x_S lze vypočítat Pearsonův koeficient korelace (srov. vzorec (61) v oddíle 3.4.3)

$$r_p = \frac{n\sum x_L \cdot x_S - \sum x_L \cdot \sum x_S}{\sqrt{\left[n\sum x_L^2 - \left(\sum x_L\right)^2\right] \cdot \left[n\sum x_S^2 - \left(\sum x_S\right)^2\right]}} \qquad (103)$$

$$r_p = \frac{339 \cdot 3197 - 716 \cdot 902}{\sqrt{\left[339 \cdot 4387 - 716^2\right] \cdot \left[339 \cdot 5143 - 902^2\right]}} = 0,460$$

Z hodnoty korelačního koeficientu r_p se už snadno vypočítá podle vzorce (102) koeficient reliability r_{sb}

$$r_{sb} = \frac{2 \cdot 0,46}{1 + 0,46} = 0,630$$

4.4.5 STANDARDIZACE DIDAKTICKÉHO TESTU

Dosažený počet bodů v testu (hrubé skóre) sám o sobě nic neříká o tom, zda výkon žáka je průměrný, dobrý či slabý. Získá-li například žák ve dvou testech stejný počet bodů, může to znamenat v jednom testu vynikající výkon, zatímco ve druhém pouze výkon průměrný. Teprve na základě srovnání dosaženého výkonu žáka s výkony ostatních žáků můžeme daného jedince adekvátně posoudit.

Smyslem standardizace je vytvoření testového standardu (testové normy), který umožní zařadit žáka podle dosaženého počtu bodů do určitého „žebříčku“ (stupnice, škály).

U standardizovaných didaktických testů se výkon jednotlivých žáků porovnává s reprezentativním vzorkem žáků (zpravidla se jedná o stovky žáků). Postup, kterým se toto srovnávání realizuje, se nazývá **standardizace testu**. Standardizovat výsledky určitého testu znamená vyjádřit je vzhledem k výsledkům standardizačního vzorku studentů (žáků).

Nejjednodušší metody standardizace jsou založeny na zjišťování procent žáků, kteří v reprezentativním vzorku dosáhli určitého výsledku. Složitější metody standardizace většinou předpokládají, že výsledky testování odpovídají předpokladu **normálního rozdělení** a vycházejí z určování vzdálenosti jednotlivého výsledku od aritmetického průměru (jednotkou této vzdálenosti je směrodatná odchylka).

Percentilová škála

Nejjednodušší metodou standardizace je standardizace pomocí **percentilů**. U této metody se ke každému dosaženému počtu bodů (hrubému skóre) přiřadí tzv. **percentilové pořadí**, které udává, kolik procent testovaných osob ve vzorku dosáhlo horšího výkonu. To umožňuje posoudit, jaké je relativní pořadí určitého jedince ve skupině (např. v populaci žáků).

Percentilové pořadí pro určitý výsledek v testu lze vypočítat podle vzorce

$$PR = 100 \cdot \frac{n_k - \frac{n_i}{2}}{n} \tag{104}$$

kde *PR* je percentilové pořadí testované osoby pro určitý výsledek v testu, n_k je kumulativní četnost daného výsledku, n_i je četnost daného výsledku a n je celková četnost testovaných osob.

Jestliže se určují percentilová pořadí z tabulky četností (ve které nejsou zachyceny všechny jednotlivé výsledky testování, nýbrž jen výsledky v určitých bodových intervalech), potom se výpočet provádí podle vzorce

$$PR = 100 \cdot \frac{n_k - \frac{d_L \cdot n_i}{h}}{n} \tag{105}$$

kde PR a n mají stejný význam jako ve vzorci (104), d_L je rozdíl mezi daným výsledkem a dolní hranicí intervalu, v němž se tento výsledek nachází, n_k je kumulativní četnost v intervalu, v němž se daný výsledek nachází, n_i je četnost v intervalu, v němž se výsledek nachází, a h je hloubka intervalu.

Poznámka: *V pedagogické praxi se někdy místo popsaných odhadů percentilového pořadí vypočítává procento žáků, kteří dosáhli daného a horšího výkonu. V těchto případech se často hovoří o tzv. percentilových hodnotách. Při posuzování výsledků testování je proto vždy potřeba vědět, jakým způsobem byly výsledky získány (zda vyjadřují „procento žáků, kteří dosáhli výkonu horšího" nebo „procento žáků, kteří dosáhli určitého výkonu anebo výkonu horšího").*

Příklad 60

Konstrukci a používání percentilové normy budeme ilustrovat na příkladu standardizace didaktického testu z předcházejícího příkladu. Tento test byl zadán vzorku 339 žáků (výsledky uvádí tabulka 72). Jestliže chceme stanovit pro jednotlivé výsledky v testu percentilové pořadí, musíme nejdříve vypočítat tzv. kumulativní četnosti. Kumulativní četnost je četnost určitého výsledku v tabulce a četnost všech horších výsledků dohromady.

Tab. 72 *Percentilová norma pro didaktický test*

Počet bodů	Četnost	Kumulativní četnost	Percentilové pořadí
0	8	8	1
1	18	26	5
2	33	59	13
3	44	103	24
4	54	157	38
5	57	214	55
6	50	264	71
7	31	295	82
8	20	315	90
9	16	331	95
10	8	339	99
	∑ 339		

Vypočítané kumulativní četnosti použijeme při výpočtu percentilového pořadí pro jednotlivé výsledky podle vzorce (104). Například pro výsledek 4 body dostáváme

$$PP_4 = 100 \cdot \frac{157 - \frac{54}{2}}{339} = 38$$

Z hodnot percentilového pořadí si lze učinit představu o tom, co znamená určitý výkon v porovnání s celým standardizačním vzorkem, a tudíž (pokud se jedná o reprezentativní výběr) s celou populací. Z takto získaného standardu lze vyčíst, že například výkon 10 bodů je výkonem mimořádně dobrým, protože plných 99 % žáků dosahuje horších výsledků. Naproti tomu výsledek 1 bod představuje výkon velmi slabý, protože pouze asi 5 % žáků dosahuje výkonů horších.

C-škála

Výkon žáka v didaktickém testu je možné vyjádřit také pomocí tzv. **C-škály**. Při konstrukci této škály se postupuje tak, že celý standardizační vzorek testovaných osob (uspořádaný podle počtu dosažených bodů) se rozdělí do 11 skupin (stupňů škály) tak, že do první skupiny (bod škály 0) se umístí 1,2 % nejhorších žáků, do druhé skupiny (bod škály 1) se umístí 2,8 % žáků atd. Pro každý bod škály je stanoveno určité procento žáků (tab. 73). Procenta žáků, která odpovídají jednotlivých bodům škály, jsou volena tak, že jsou symetrická vzhledem ke střednímu, tj. 5. bodu škály. Tento střední bod škály obsahuje také největší procento žáků (19,8 %).

Tab. 73 *Procenta případů u C-škály*

Body C-škály	Procenta případů	Kumulativní procenta
0	1,2	1,2
1	2,8	4,0
2	6,6	10,6
3	12,1	22,7
4	17,4	40,1
5	19,8	59,9
6	17,4	77,3
7	12,1	89,4
8	6,6	96,0
9	2,8	98,8
10	1,2	100,0

Příklad 61

Vytvoření C-normy budeme ilustrovat na stejných výsledcích testování jako v předcházejícím příkladě. Výsledky testování jsou uvedeny v tabulce 74 (první dva sloupce zleva).

Při konstrukci C-normy se nejdříve vypočítají **kumulativní četnosti**. Vypočítané kumulativní četnosti se dále převedou na relativní kumulativní četnosti, které se vyjádří v procentech. Tento výpočet se provádí podle vzorce

$$\text{relativní kumulativní četnost} = 100 \cdot \frac{n_k}{n} \tag{106}$$

kde n_k je kumulativní četnost (absolutní) a n je počet testovaných osob. K vypočítaným relativním kumulativním četnostem se potom vyhledávají stejné nebo nejblíže nižší hodnoty relativní kumulativní četnosti z tabulky 73. Tak získáme intervaly pro jednotlivé body (stupně) C-škály.

***Tab. 74** C-norma pro didaktický test*

Počet bodů	Četnost	Kumulativní četnost	Relativní kumulativní četnost (%)	Body C-škály
0	8	8	2,4	0
1	18	26	7,7	1
2	33	59	17,4	2
3	44	103	30,4	3
4	54	157	46,3	4
5	57	214	63,1	5
6	50	264	77,9	6
7	31	295	87,0	6
8	20	315	92,9	7
9	16	331	97,6	8
10	8	339	100,0	10
	∑ 339			

Škála STANIN

Spojením prvních dvou a posledních dvou stupňů C-škály vznikne devítibodová škála STANIN (*standard nine* – standardní devítistupňová škála). První a poslední bod této škály tedy obsahují 1,2 % + 2,8 % = 4 % případů (žáků). Tato škála je zvláště výhodná při počítačovém zpracování výsledků, protože výsledky testování vyjádřené pomocí škály STANIN zabírají jen jeden znak v počítači. Tabulka 75 uvádí porovnání škály STANIN a C-škály.

***Tab. 75** Škála STANIN a C-škála*

<table>
<tr><td>STANIN</td><td colspan="2">1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td colspan="2">9</td></tr>
<tr><td>C-škála</td><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td></tr>
</table>

z-škála

Konstrukce této škály vychází z předpokladu, že výsledky testování mají **normální rozdělení**. Nejdříve by se proto mělo ověřit, zda tento předpoklad je splněn.

Hodnota z-škály (hodnota normované normální veličiny) vyjadřuje, jak daleko je určitý dosažený výsledek od aritmetického průměru, přičemž jednotkou této vzdálenosti je směrodatná odchylka. Platí, že

$$z = \frac{x - \overline{x}}{s} \tag{107}$$

kde z je hodnota z-škály, x je určitý testový výsledek, $\overline{x}$ je aritmetický průměr výsledků v testu, s je směrodatná odchylka pro všechny testové výsledky.

Hodnoty z-škály se pohybují zpravidla v intervalu od –3 do +3, průměrný výsledek je dán hodnotou $z = 0$.

Samotná z-škála se při standardizaci didaktických testů užívá jen zřídka (problémy působí zejména to, že může nabývat záporných hodnot). Užívá se však při konstrukci mnohých dalších standardních škál.

Z-škála

Tato škála vychází ze z-škály a je definována vztahem

$$Z = 100 + 10z = 100 + 10 \cdot \frac{x - \overline{x}}{s} \tag{108}$$

Pokud se výsledky didaktického testu vyjadřují na této škále, potom velká většina (99,7 %) jich je v intervalu od 70 do 130 a průměrný výsledek je dán hodnotou 100 (směrodatná odchylka je 10).

T-škála

Často se k vyjadřování výsledků dosažených v didaktických testech používá tzv. T-škála. T-škála vychází (stejně jako Z-škála) ze z-škály a je definována vztahem

$$T = 50 + 10z = 50 + 10 \cdot \frac{x - \overline{x}}{s} \tag{109}$$

Pro T-škálu platí, že její hodnoty jsou zpravidla v intervalu od 20 do 80 bodů, průměrná hodnota je vyjádřena stupněm 50 (směrodatná odchylka výsledků je 10).

Příklad 62

Při standardizaci didaktického testu bylo nejdříve (testem dobré shody chí-kvadrát) ověřeno, že výsledky testování dobře odpovídají normálnímu rozdělení. Je to názorně patrné i z obrázku 26, kde histogram četností pro výsledky testu je proložen křivkou normálního rozdělení (Gaussova křivka).

Byl vypočítán aritmetický průměr $\overline{x} = 4,773$ a směrodatná odchylka $s = 2,313$. Pomocí vzorců (108) a (109) byly vypočítány hodnoty Z-škály a hodnoty T-škály pro jednotlivé výsledky testování. Například pro výsledek 5 bodů po dosazení a zaokrouhlení výsledků obdržíme

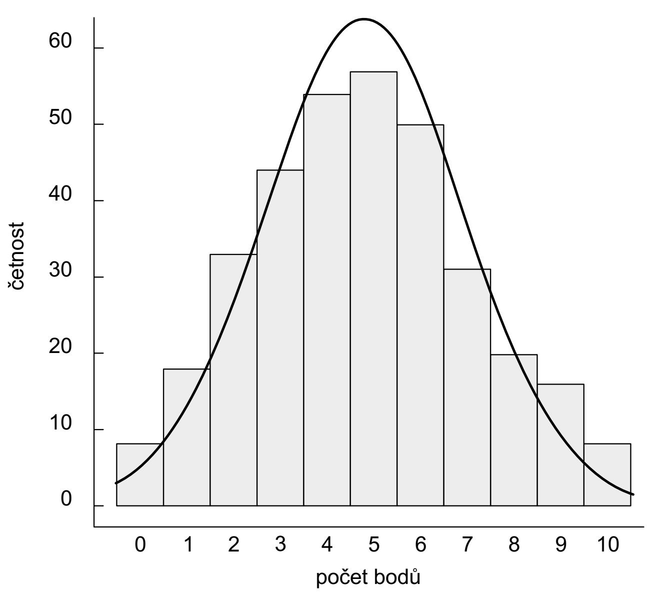
$$Z_5 = 100 + 10 \cdot \frac{5 - 4,773}{2,313} = 101$$

$$T_5 = 50 + 10 \cdot \frac{5 - 4{,}773}{2{,}313} = 51$$

Z výsledků je patrné, že výkon 5 bodů v testu znamená téměř přesně průměrný výkon.

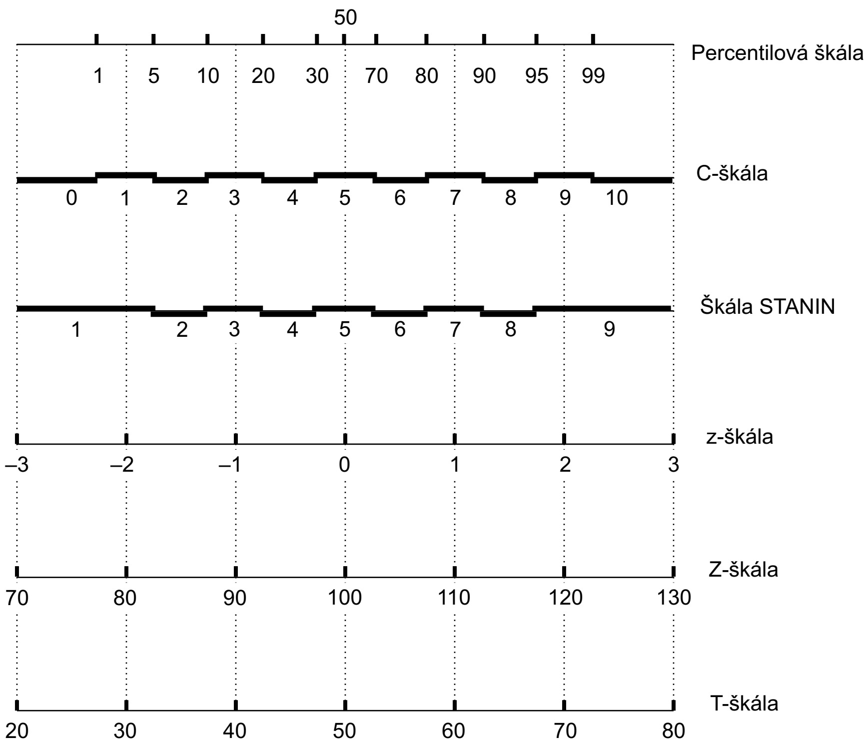
Tab. 76 *Z-škála a T-škála pro didaktický test*

Počet bodů	Četnost	Body Z-škály	Body T-škály
0	8	79	29
1	18	84	34
2	33	88	38
3	44	92	42
4	54	97	47
5	57	101	51
6	50	105	55
7	31	110	60
8	20	114	64
9	16	118	68
10	8	123	73
	∑ 339		



Obr. 26 *Histogram četností pro výsledky testu*

Srovnání nejpoužívanějších druhů škál uvádí obrázek 27.



Obr. 27 *Srovnání nejpoužívanějších druhů škál*

4.5 SOCIOMETRIE V PEDAGOGICKÉM VÝZKUMU

Sociometrií rozumíme soubor specifických výzkumných postupů, které slouží ke zjišťování, popisu a analýze směru a intenzity mezilidských vztahů, jak se projevují v malých sociálních skupinách.

Malá sociální skupina je seskupení lidí, které se vyznačuje některými základními znaky. Mezi ně patří například to, že její členové jsou spojeni určitými společnými cíli, vzájemně se znají a komunikují spolu a vytvářejí společné normy, které jsou pro chování členů skupiny závazné.

Každá malá sociální skupina má určitou vnitřní strukturu. Ve skupině jsou jisté **formální vztahy**, které jsou určeny oficiálně (například vztahy nadřízenosti a podřízenosti), ale také určité **neformální vztahy**, které vznikají spontánně a dobrovolně na základě společné práce, společných zájmů a zálib a z toho pramenících sympatií či antipatií mezi členy skupiny. Sociometrie věnuje pozornost zejména zjišťování neformálních vztahů ve skupinách. Informací získaných pomocí sociometrie je však možno využít i k ovlivnění formálních vztahů ve skupině.

4.5.1 SOCIOMETRICKÝ TEST

Nejdůležitější a základní (i když ne jedinou) technikou, pomocí níž sociometrie získává informace, je **sociometrický test**. Sociometrickým testem zjišťujeme jednak pozitivní volby ve skupině (sympatie, preference, atrakce), jednak negativní volby (odmítání, repulze). V sociometrické praxi je častější zjišťování kladných voleb. Sociometrický test obsahuje jednu nebo více otázek, které umožňují všem členům sociální skupiny volit partnery pro určité situace nebo určité společné činnosti. Sociometrický test se zadává zkoumaným osobám většinou písemně, přičemž tyto osoby odpovídají uvedením příslušných jmen. U malých dětí nebo osob negramotných lze místo písemné formy zadání sociometrického testu použít například třídění fotografií či značek dětí, zanechávání dárků (např. na lavici při odchodu ze třídy) apod. Obsah otázek je závislý vždy na konkrétním cíli sociometrického šetření.

Existenci kladných vztahů ve skupině mohou odhalovat například následující otázky:

S kým ze třídy bys nejraději pracoval na …?
Kteří žáci ze třídy jsou tvoji nejlepší kamarádi?
Kdo ze žáků vaší třídy má největší autoritu?
Koho ze třídy bys vybral, aby zastupoval třídu při jednání s ředitelem školy?

Příklady otázek, které zjišťují existenci záporných vztahů ve skupině:

Kteří žáci ve třídě patří k nejméně oblíbeným?
S kým ze třídy bys rozhodně nechtěl jít na společný výlet?
S kým ze třídy bys na školním výletě nechtěl být ubytován na jednom pokoji?

Zakladatel sociometrie J. L. Moreno zformuloval šest základních pravidel pro konstrukci sociometrického testu (Skalková et al., 1983):

1. Je třeba jasně stanovit hranice sociální skupiny, do které zkoumané osoby přináležejí a v níž mají provádět své volby. Skupinou s jasně stanovenými hranicemi je například školní třída, sportovní oddíl apod.
2. Každý člen dané skupiny má mít možnost provést neomezený počet výběrů. Volí-li například určitý jedinec jenom jednoho z členů skupiny za svého partnera v daném kritériu, je to významný údaj charakterizující ho v porovnání s jiným jedincem, který volí například tři, čtyři, pět nebo více členů.
3. Členové sociální skupiny mají mít kritérium výběru jednoznačně určeno, tj. má být mimo jakoukoli pochybnost, vzhledem k jaké činnosti, události apod. si mají partnery vybírat.
4. Výsledky sociometrického testu mají být vždy spojovány s určitými opatřeními týkajícími se restrukturace zkoumané skupiny. Zkoumané osoby mají získat zkušenost, že sociometrický test jim slouží, že volby, které provedly, mají praktické konsekvence. Má se tak čelit tomu, aby se sociometrické šetření nestalo jen fantazií o možném a nemož-

ném, ale aby bylo šetřením faktického rozhodování zkoumaných osob o mezilidských vztazích se všemi důsledky, které může takové rozhodování mít.

5. Jednotliví členové skupiny nemají vědět o výběrech, které provedli členové ostatní. Díky této zásadě se čelí vzniku konfliktů, které by mohlo zveřejnění voleb navodit.
6. Má být předem prověřeno, zda všichni členové zkoumané skupiny rozumějí otázkám zařazeným do sociometrického testu tak, jak zamýšlel autor šetření.

V praxi se při sociometrických šetřeních ne vždy všechna uvedená pravidla striktně dodržují. Velmi často se nedodržuje například druhá zásada, tj. možnost provedení neomezeného počtu výběrů. Počet voleb bývá redukován na 1–3, což má své výhody (hlavně snadnější zpracovávání), ale také nevýhody (zkoumané osoby jsou pod tlakem provádět určitý počet voleb, i když to neodpovídá jejich potřebě).

Sociometrickým testem získáme údaje o tom, kdo koho ve skupině volí pro určitou situaci či činnost, respektive kdo koho pro tuto situaci odmítá. Tyto údaje jsou ve svém souhrnu zpravidla málo přehledné a je nutné je dále zpracovat. Nejčastěji se provádějí se sociometrickými daty tři druhy analýzy:

- sestavení sociometrické matice;
- konstrukce sociogramů;
- výpočet sociometrických indexů.

4.5.2 SOCIOMETRICKÉ MATICE

V tzv. **neuspořádané sociometrické matici** jsou všichni členové skupiny uvedeni (v libovolném pořadí) v prvním sloupci zleva a ve stejném pořadí také v záhlaví matice. Do řádků matice se postupně zaznamenávají znaménky (+) pozitivní volby, které členové skupiny provedli a znaménky (–) záporné volby. Bývá výhodné označit (např. zakroužkováním) také provedené vzájemné kladné volby. Po okrajích sociometrické matice se obvykle zapisují součty obdržených a součty odevzdaných voleb ve skupině. Diagonála matice se proškrtává, protože nepředpokládáme, že by jedinec volil sám sebe. Příklad neuspořádané sociometrické matice uvádí tabulka 77. V této matici jsou písmeny A až H označeni jednotliví členové sociální skupiny, znaménka (+) a (–) udávají provedené kladné a záporné volby a zakroužkováním jsou označeny vzájemné kladné volby.

***Tab. 77** Neuspořádaná sociometrická matice*

odevzdané volby

		A	B	C	D	E	F	G	H	+	–	Σ
	A	X	+	–	–		⊕			2	2	4
	B	⊕	X	–		⊕	+			3	1	4
	C	–	+	X			+	⊕	+	4	1	5
	D	+	–	–	X		–	–		1	4	5
	E	+	⊕			X	⊕		⊕	4	0	4
	F	⊕	–			+	X	+		3	1	4
	G	+	+	⊕				X		3	0	3
	H	–	+			⊕			X	2	1	3
	+	5	5	1	0	3	4	2	2	22		
obdržené volby	–	2	2	3	1	0	1	1	0		10	
	Σ	7	7	4	1	3	5	3	2			32

Uvedený příklad sociometrické matice byl získán zpracováním odpovědí na dvě následující otázky sociometrického testu:

A *Se kterým spolužákem ze třídy bys rád jel na dvoudenní turistický výlet do Krkonoš?*
B *Se kterým spolužákem ze třídy bys na tento turistický výlet rozhodně nechtěl jet?*

Popsaná neuspořádaná sociometrická matice zpravidla neposkytuje dostatečně rychlou a přehlednou informaci o struktuře sociální skupiny. Neuspořádanou matici lze vhodnou úpravou převést na **matici uspořádanou**, která poskytuje přehlednější informaci o struktuře sociální skupiny. Uspořádání jedinců v uspořádané sociometrické matici by mělo odpovídat následujícím požadavkům (Petrusek, 1969):

- vzájemné výběry se soustřeďují kolem diagonály matice;
- nejvíce odmítnutí se má objevovat v rozích matice vzdálených od diagonály;
- jedinci, kteří provedli výběr do podskupiny a jsou touto podskupinou ignorováni, se soustřeďují v pravém horním okraji;
- jedinci, kteří provedli výběr do podskupiny a jsou zjevně s podskupinou spřízněni, se soustřeďuji při levém dolním okraji;
- izolovaní jedinci mají být umístěni poblíž středu sociometrické matice.

Při úpravě matice se většinou zkusmo hledá takové uspořádání osob, které by v maximální možné míře vyhovovalo výše uvedeným požadavkům. Příklad uspořádané matice, která vznikla úpravou shora uvedené matice neuspořádané, uvádí tabulka 78.

***Tab. 78** Uspořádaná sociometrická matice*

										odevzdané volby		
		C	G	D	H	E	F	B	A	+	−	Σ
	C	X	+		+		+	+	−	4	1	5
	G	+	X					+	+	3	0	3
	D	−	−	X			−	−	+	1	4	5
	H				X	+		+	−	2	1	3
	E				+	X	+	+	+	4	0	4
	F		+			+	X	−	+	3	1	4
	B	−				+	+	X	+	3	1	4
	A	−		−			+	+	X	2	2	4
	+	1	2	0	2	3	4	5	5	22		
obdržené volby	−	3	1	1	0	0	1	2	2		10	
	Σ	4	3	1	2	3	5	7	7			32

Z uvedené matice snadno vyčteme, že v sociální skupině existuje jedna větší podskupina (A, B, E, F, H) a jedna dvojice vzájemně se volících jedinců (C, G). Tato dvojice je však poměrně pevně zprostředkovaně spjata s větší podskupinou (přes jedince A, B, F, H). Jedinec D je ostatními členy skupiny opomíjen.

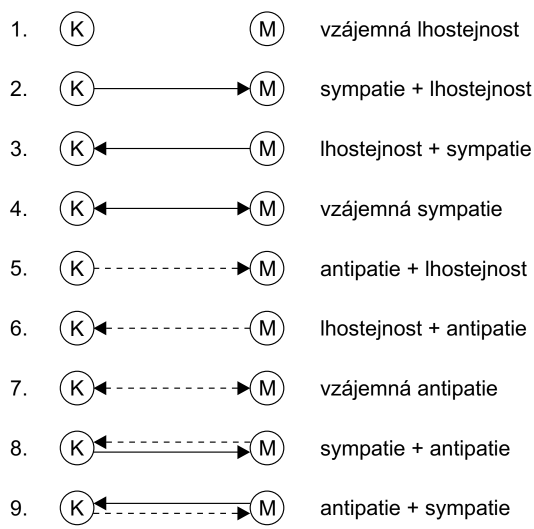
Kromě uspořádané matice lze konstruovat také tzv. **kvadratickou** nebo **kubickou** matici. Kvadratická matice, kterou dostaneme z původní matice umocněním, umožňuje identifikovat zprostředkované vztahy ve skupině („přátele přátel"). Podobně kubická matice umožňuje postihnout trojstupňové vztahy v sociální skupině (vztahy zprostředkované dvěma jedinci). Podrobnější vysvětlení ke konstrukci a čtení matic lze nalézt například v práci M. Petruska (1969).

Ze sociometrických matic lze vyčíst všechny potřebné údaje o provedených volbách ve skupině, a tím i o struktuře osobně výběrových vztahů ve skupině. Informace, které sociometrické matice přinášejí, však zpravidla nebývají příliš přehledné a názorné. Názornější představu o struktuře neformálních vztahů v sociální skupině poskytují tzv. **sociogramy**.

4.5.3 SOCIOGRAMY

Sociogramy jsou vlastně jakýmisi „mapami“ neformálních vztahů v sociálních skupinách. Existuje celá řada různých variant konstrukce sociogramů. Sociogramy má smysl konstruovat pouze pro sociální skupiny s počtem maximálně deset až patnáct členů, protože při větších počtech osob jsou sociogramy již značně nepřehledné a velmi špatně čitelné. Vzhledem k tomu, že běžné školní třídy mají počet žáků větší, nebývá vhodné konstruovat sociogram pro třídu jako celek. Sociogram může dobře posloužit například ke zmapování situace uvnitř přirozeně vznikajících podskupin ve třídě (např. skupina chlapců nebo dívek v rámci smíšené třídy). Při kreslení sociogramů se zpravidla jednostranný kladný výběr znázorňuje plnou čarou se šipkou, přičemž šipka označuje směr výběru. Jednostranný záporný výběr (odmítání) se většinou označuje čárkovanou čarou se šipkou. Vzájemný kladný výběr se zpravidla označuje plnou čarou se šipkami na obou koncích, vzájemný záporný výběr čárkovanou čarou se šipkami na obou koncích.

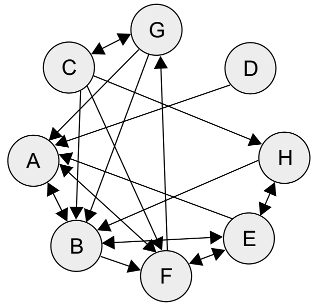
Mezi dvěma jedinci sociální skupiny existuje celkem devět teoretických možností vzájemných vztahů (obr. 28).



Obr. 28 *Vzájemné vztahy mezi dvěma jedinci*

Kruhový sociogram

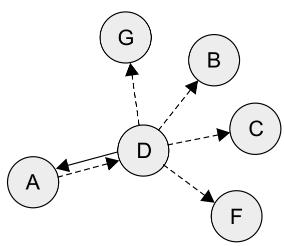
Mezi nejjednodušší sociogramy patří **kruhový sociogram**. Při jeho konstrukci znázorňujeme jednotlivé členy skupiny pomocí kroužků po obvodu kružnice a volby mezi nimi šipkami uvnitř kruhu. Obrázek 29 uvádí kruhový sociogram pro situaci zachycenou v sociometrických maticích (tab. 77 a 78). Kruhový sociogram je velmi jednoduchý, ale při početnějších skupinách je značně nepřehledný a obtížně se čte.



Obr. 29 *Kruhový sociogram (kladné volby)*

Individuální sociogram

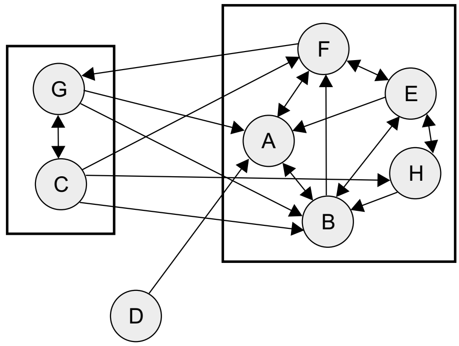
Jiným typem sociogramu je **individuální sociogram**. Tento druh sociogramu nezachycuje strukturu skupiny jako celku, ale pouze vztah určitého jedince k jedincům, které volil, nebo jimiž byl volen. Při jeho konstrukci znázorníme jedince (k němuž se sociogram vztahuje) do středu a kolem něho volené či volící jedince. Volby mezi jedinci se znázorňují opět pomocí plných, respektive čárkovaných čar se šipkami. Individuální sociogram je výhodné použít například tehdy, když potřebujeme analyzovat vztahy jistého jedince k sociální skupině (a vztahy sociální skupiny k tomuto jedinci). V příkladě, který zachycuje analyzovaná sociometrická matice (tab. 77*)* například jedinec D učinil jednu kladnou volbu, čtyři volby záporné a současně jednu zápornou volbu obdržel. Vztahy tohoto jedince k ostatním členům sociální skupiny jsou názorně patrné z obrázku 30.



Obr. 30 *Individuální sociogram*

Strukturální sociogram

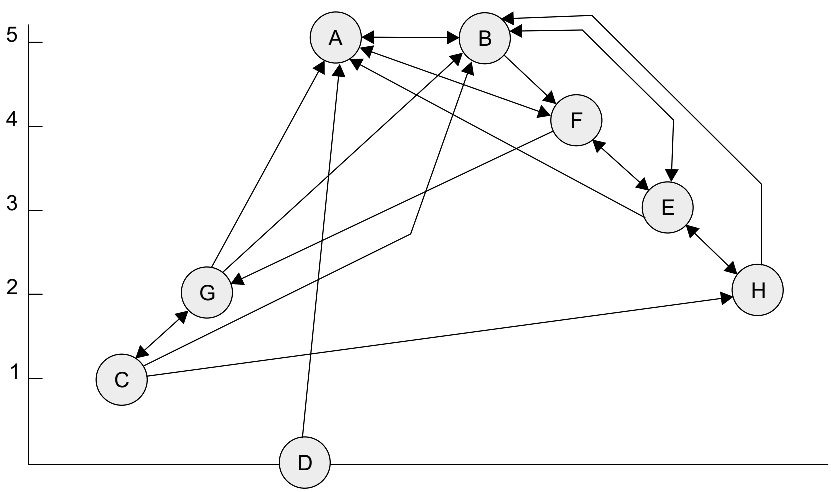
Pokud se ve studované sociální skupině vytvoří relativně samostatné podskupiny vzájemně se volících jedinců, bývá výhodné konstruovat **strukturální sociogram**. U tohoto sociogramu se do středu obrazce znázorňují jedinci s nejvyšším počtem obdržených kladných voleb. Kolem těchto osob zakreslíme jedince, které k nim váže oboustranná kladná volba. Vytvořené podskupiny se většinou v sociogramu graficky zvýrazňují (např. zarámováním nebo vyšrafováním). Strukturální sociogram pro analyzovanou matici uvádí obrázek 31.



Obr. 31 *Strukturální sociogram*

Hierarchický osový sociogram

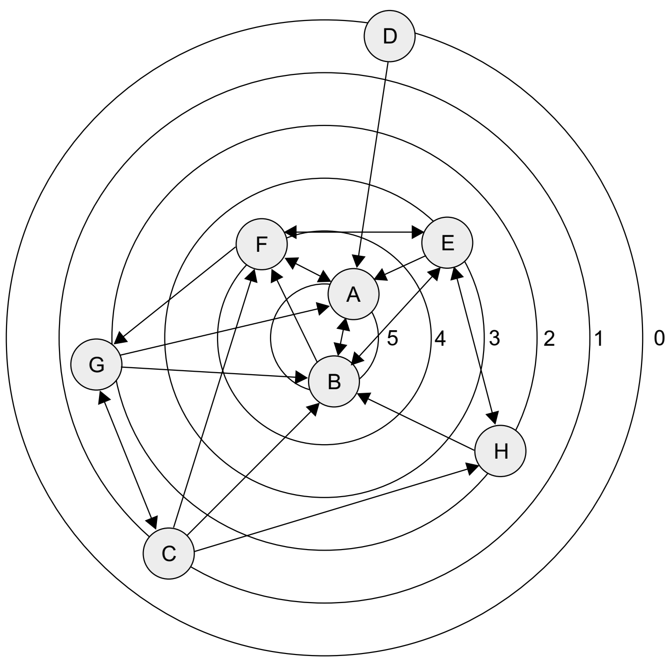
Dalším typem sociogramu je **hierarchický osový sociogram**. Tento sociogram je konstruován v soustavě dvou os, přičemž na svislé ose se zobrazuje počet obdržených voleb a na vodorovné ose se zakreslují jednotliví členové sociální skupiny. Hierarchický osový sociogram je možno zkonstruovat buď pro oblast kladných voleb, nebo pro oblast voleb záporných (odmítání). K dosažení větší přehlednosti tohoto sociogramu se doporučuje umísťovat jedince na vodorovnou osu podle toho, jak jsou si navzájem blízcí. Při konstrukci dbáme, aby se čáry znázorňující jednotlivé volby příliš nekřížily. Obrázek 32 uvádí hierarchický osový sociogram pro data z matice (tab. 77).



Obr. 32 *Hierarchický osový sociogram (kladné volby)*

Terčový sociogram

Jiným, často používaným sociogramem je **terčový sociogram** (hierarchický kruhový sociogram). U tohoto sociogramu se vyjadřuje stupeň neformální autority jedinců tím, že se umisťují do různých vzdáleností od středu sociogramu. Při konstrukci se postupuje tak, že nejdříve sestrojíme $n + 1$ soustředných kružnic, kde n je nejvyšší počet obdržených voleb u jednotlivých členů skupiny. Na největší kružnici potom umísťujeme jedince, kteří neobdrželi žádnou volbu, na další kružnice potom postupně jedince s 1, 2 až n obdrženými volbami. Čím blíže je jedinec ke středu soustředných kružnic, tím více voleb obdržel, tj. tím větší má ve skupině **neformální autoritu**. Symboly jednotlivců je nutné na příslušných kružnicích rozmísťovat tak, aby jedinci, kteří se vzájemně volí, byli zakresleni blízko u sebe a aby čáry znázorňující volby se co nejméně křížily. Terčový sociogram v mnoha případech velmi názorně zobrazuje strukturu neformálních vztahů v sociální skupině. Nevyhovuje však například v případě, kdy vysoce oblíbený jedinec má velmi těsné vztahy s jedincem, kterého volil jen on sám. V tomto případě jsou oba jedinci zakresleni v sociogramu daleko od sebe, což nevystihuje adekvátně skutečnost. Terčový sociogram je možno konstruovat pro kladné volby (obr. 33), ale i pro záporné volby.

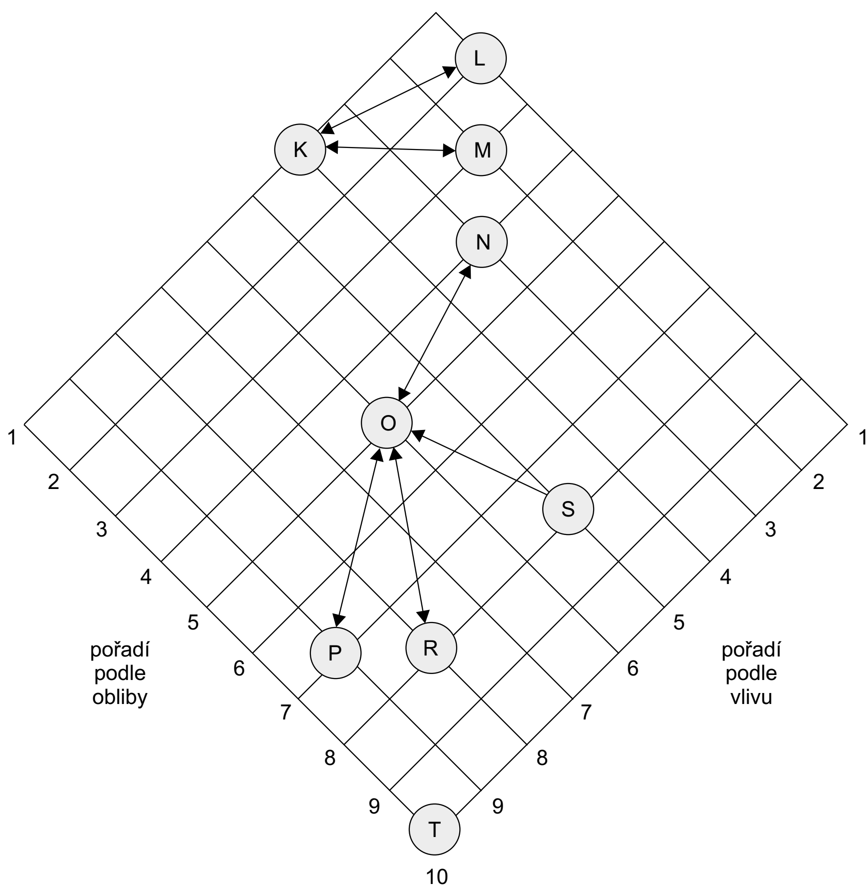


***Obr. 33** Terčový sociogram (kladné volby)*

Uvedené sociogramy můžeme označit jako sociogramy jednorozměrné, protože znázorňují pozici jednotlivce ve skupině podle jednoho kritéria. Pomocí sociometrie však můžeme zkoumat pozici jednotlivců i ze dvou hledisek současně (např. z hlediska oblíbenosti ve skupině a z hlediska vlivu na ostatní členy skupiny). V těchto případech je možno ke znázornění pozic jedince ve skupině použít dvojrozměrného sociogramu (Hrabal, 1989).

Dvojrozměrný (kovariační) sociogram

Dvojrozměrný (kovariační) sociogram je vhodný v těch případech, kdy u skupiny osob sledujeme dvě důležitá a na sobě relativně nezávislá hlediska (např. obliba a vliv). Při konstrukci tohoto typu sociogramu vycházíme z údajů dvou sociometrických matic (např. jedna matice pro oblibu a druhá matice pro vliv). Podle počtu obdržených voleb v obou maticích přiřadíme každému jedinci pořadí (pořadí podle obliby a pořadí podle vlivu). Údaje o pořadích se potom nanášejí do čtvercové sítě, která má podobu **čtverce stojícího na vrcholu** (obr. 34). Strany čtverce jsou rozděleny na tolik dílů, kolik je členů ve skupině. Jedna strana čtvercové sítě slouží k zachycení pořadí jedinců podle jednoho hlediska (např. podle obliby), druhá strana k zachycení pořadí podle druhého hlediska (např. podle vlivu). Každý jedinec je ve čtvercové síti umístěn podle toho, jaké pořadí má podle obou hledisek. Například jedinec K je ve skupině nejoblíbenější (má pořadí 1 podle obliby), ale jeho vliv na ostatní členy skupiny je poněkud menší (podle vlivu má pořadí 4).



Obr. 34 *Dvojrozměrný (kovariační) sociogram*

Uvedený sociogram poskytuje velmi přehlednou informaci o postavení jedinců podle zvažovaných hledisek ve skupině. Umístění ve vertikálním směru informuje o kombinované pozici ve skupině. Umístění ve vodorovném směru poskytuje informaci o poměru obliby a vlivu u jednotlivých členů skupiny. U jedinců umístěných vlevo od diagonály převažuje obliba nad vlivem. Naopak u jedinců, kteří se nacházejí vpravo od diagonály, převažuje vliv nad oblibou. Jedinci umístění na diagonále se vyznačují vyrovnanou pozicí podle obou hledisek.

4.5.4 SOCIOMETRICKÉ INDEXY

Ze sociometrické matice se vychází také při výpočtu sociometrických indexů, které představují kvantitativní vyjádření výsledků sociometrického šetření. Sociometrické indexy lze rozdělit do tří skupin:

- individuální sociometrické indexy (charakterizující neformální postavení jedince ve skupině);
- skupinové sociometrické indexy (charakterizující sociální skupinu jako celek);
- indexy charakterizující strukturu podskupin.

V literatuře (např. Papica, 1974; Petrusek, 1969; Schneider, 1980 a další) lze nalézt sociometrických indexů velký počet. Uvedeme z nich jen ty nejčastěji používané.

4.5.4.1 Individuální sociometrické indexy

Postavení jedince ve skupině je dáno jeho **sociálním statusem**. Status jedince je dán počtem voleb, které jedinec získal, vztaženým k možnému počtu voleb, tj. $N - 1$, kde N je počet členů ve skupině.

Pozitivní sociometrický status (status volby) S_V

$$S_v = \frac{p}{N-1} \tag{110}$$

kde p je počet obdržených pozitivních voleb a N je počet jedinců v sociální skupině. Pro jedince A v tabulce 77 (nejoblíbenější ve skupině) například vychází $S_V = 5 / (8 - 1) = 0{,}71$.

Negativní sociometrický status (status odmítnutí) S_o

$$S_o = \frac{n}{N-1} \tag{111}$$

kde n je počet odmítnutí, kterých se dostalo určitému jedinci, a N je počet jedinců ve skupině. Pro jedince C (jedinec, který obdržel nejvíce záporných voleb) dostáváme

$$S_o = 3 / (8 - 1) = 0{,}43$$

Výsledný status V

$$V = \frac{p-n}{N-1} = S_v - S_0 \tag{112}$$

kde p je počet obdržených pozitivních voleb, n je počet obdržených záporných voleb (odmítnutí) a N je počet jedinců v sociální skupině.

Pro jedince A vychází výsledný status $V = (5 - 2) / (8 - 1) = 0{,}71 - 0{,}29 = +0{,}43$, pro jedince C dostáváme $V = (1 - 3) / (8 - 1) = -0{,}29$.

Index expanzivity E

Tento index je použitelný pouze v případě, že sociometrický test připouštěl neomezený počet voleb

$$E = \frac{v}{N-1} \tag{113}$$

kde v je počet voleb (kladných i záporných), které jedinec učinil, a N je počet osob v sociální skupině.

Pro jedince A vychází index expanzivity $E = (5 + 2) / (8 - 1) = 1$, pro jedince C vychází $E = (1 + 3) / (8 - 1) = 0{,}57$.

Uvedené individuální sociometrické indexy mohou mít velkou poznávací hodnotu nejen v pedagogickém výzkumu, ale i při řešení praktických výchovných problémů.

4.5.4.2 Skupinové sociometrické indexy

Tyto sociometrické indexy slouží k charakteristice sociální skupiny jako celku.

Index pozitivní skupinové expanzivity E_p

Index vyjadřuje průměrný počet kladných voleb, který připadá na jednoho člena skupiny

$$E_p = \frac{p_s}{N} \tag{114}$$

kde p_s je počet kladných voleb provedených ve skupině a N je počet členů ve skupině. Pro celou sociální skupinu (tab. 77) dostáváme $E_p = 22 / 8 = 2{,}75$.

Index negativní skupinové expanzivity E_n

Index vyjadřuje průměrný počet záporných voleb (odmítnutí), který připadá na jednoho člena skupiny

$$E_n = \frac{n_s}{N} \tag{115}$$

kde n_s je počet záporných voleb provedených ve skupině a N počet členů ve skupině. Pro situaci, kterou zachycuje tabulka 77, vychází $E_n = 10 / 8 = 1{,}25$.

Index skupinové koheze (soudržnosti) K

Tento index je dán poměrem provedených vzájemných výběrů (párů) ve skupině k maximálně možnému počtu vzájemných výběrů

$$K = \frac{d}{\frac{N \cdot (N-1)}{2}} \tag{116}$$

kde d je počet vzájemných výběrů (párů, dvojic) provedených ve skupině a N je celkový počet osob ve skupině. Pro data z tabulky 77 dostáváme $K = 6 / 28 = 0{,}21$.

Uvedený vztah pro výpočet indexu skupinové koheze platí v případě, že počet voleb v sociometrickém testu není omezen. V případě **limitovaného počtu voleb** lze index koheze K_L vypočítat ze vztahu

$$K_L = \frac{d}{\frac{a \cdot N}{2}} \tag{117}$$

kde d je opět počet vzájemných výběrů, N je počet jedinců ve skupině a a je počet povolených voleb.

Index skupinové koherence K_o

Hodnota tohoto indexu indikuje velikost vzájemné vázanosti členů skupiny

$$K_o = \frac{d}{p_s} \tag{118}$$

kde d je počet vzájemných voleb (párů) ve skupině a p_s je počet kladných voleb provedených ve skupině. Pro hodnoty z tabulky 77 dostáváme $K_o = 6 / 22 = 0{,}27$.

Index skupinové integrace I_g

Tento sociometrický index slouží k posouzení stupně zapojení členů do sociální skupiny. Výpočet indexu vychází z počtu osob izolovaných (tj. takových, které neprovedly ani neobdržely žádnou kladnou volbu) a počtu osob opomenutých (tj. takových, které sice samy kladnou volbu provedly, ale samy žádnou neobdržely). Tento index lze počítat pouze v případě, že jmenovatel zlomku není roven nule

$$I_g = \frac{1}{N_i + N_o} \tag{119}$$

kde N_i je počet izolovaných jedinců a N_o je počet opomenutých jedinců. Pro analyzovanou sociometrickou matici vychází $I_g = 1 / (0 + 1) = 1$.

Index skupinové sociopreferenční izolace I_z

Také tento index slouží k charakteristice počtu jedinců izolovaných a opomenutých

$$I_z = \frac{N_i + N_o}{N} \tag{120}$$

kde N_i je počet jedinců izolovaných, N_o je počet jedinců opomenutých a N je celkový počet jedinců v sociální skupině. Pro data z tabulky 77 je $I_z = (0 + 1) / 8 = 0{,}13$.

4.5.4.3 Sociometrické indexy charakterizující strukturu podskupin

Strukturu podskupin lze charakterizovat také pomocí indexů, které byly uvedeny v souvislosti s popisem celé sociální skupiny. Z indexů vyvinutých speciálně pro charakteristiku situace uvnitř podskupin si zaslouží pozornost zejména index vnitroskupinové preference.

Index vnitroskupinové preference I_p

Hodnota tohoto indexu je ukazatelem toho, jak členové podskupiny preferují sebe před ostatními členy sociální skupiny

$$I_p = \frac{x_1 \cdot (N - N_1)}{x_2 \cdot (N_1 - 1)} \tag{121}$$

kde x_1 je počet kladných výběrů odevzdaných členy podskupiny členům vlastní podskupiny, N je počet členů v celé skupině, N_1 je počet členů dané podskupiny, x_2 je počet kladných výběrů odevzdaných členy dané podskupiny mimo vlastní podskupinu.

Jestliže hodnota I_p je větší než 1, znamená to, že členové podskupiny preferují sebe. Jestliže je zjištěná hodnota menší než 1, preferují ostatní členy sociální skupiny. Pro pětičlennou podskupinu (tab. 77) vychází

$$I_p = 13 \cdot (8-5) \big/ 1 \cdot (5-1) = 9{,}75$$

4.6 SÉMANTICKÝ DIFERENCIÁL

Tato zajímavá metoda umožňuje měřit individuální, psychologické významy určitých objektů (obyčejně pojmů) u jednotlivých osob. Je známou skutečností, že posuzuje-li jeden objekt více posuzovatelů, každý z nich jej „vidí" poněkud (někdy dokonce velmi) odlišně. Vedle společného kulturního významu má totiž každý pojem ještě další, vedlejší (**konotativní**) významy, které charakterizují jednotlivé posuzovatele. Americký profesor C. Osgood (Osgood et al., 1957) vytvořil v roce 1957 metodu, kterou lze tyto vedlejší, psychologické významy pojmů měřit. Tato metoda je známá pod názvem **sémantický diferenciál**.

4.6.1 KLASICKÝ SÉMANTICKÝ DIFERENCIÁL C. OSGOODA

U sémantického diferenciálu se individuální významy pojmů měří pomocí určitého počtu posuzovacích škál (nejčastěji sedmibodových). Respondenti zaznamenávají svoje mínění o posuzovaných objektech výběrem určitého bodu na těchto škálách. U škál jsou krajní body tvořeny vždy dvojicí adjektiv protikladného významu (antonym). Volbou bodu na škále vyjadřují respondenti míru vlastnosti, jež je vyjádřena příslušnou dvojicí adjektiv. Jednotlivým bodům na škále se potom přiřazují číselné hodnoty 1–7. C. Osgood doporučuje posuzovat každý pojem z hlediska tří faktorů, které označuje jako **faktor hodnocení**, **faktor potence** a **faktor aktivity**. Faktor hodnocení lze interpretovat jako **dobro** či **zlo** pojmu, faktor potence jako **sílu** pojmu a faktor aktivity jako vztah pojmu k **pohybu** *a* **změnám**.

Určíme-li u jistého pojmu tyto tři faktory, je tím podle C. Osgooda určen jeho individuální psychologický význam. Významy jednotlivých pojmů lze potom graficky znázornit jako body v trojdimenzionálním prostoru, jenž se označuje jako **sémantický prostor**.

Použití Osgoodova sémantického diferenciálu budeme ilustrovat na příkladu výzkumu (Chráska, 1997), ve kterém se měřily individuální významy vybraných pojmů u žáků základní školy. Žáci posuzovali ve výzkumu celkem deset pojmů (jsou uvedeny v tabulce 80) pomocí dvanácti škál sémantického diferenciálu. Každý posuzovaný pojem byl předložen na samostatném záznamovém listu. Tabulka 79 uvádí příklad záznamového listu pro pojem NAŠE ŠKOLA.

Písmena *h, p, a* označují faktorovou identifikaci použitých škál (faktor hodnocení, potence (síly) a aktivity).

Aby se snížilo nebezpečí stereotypního posuzování ve škálách, jsou některé škály prezentovány v tzv. **reverzní podobě** (škály mají převrácené krajní body). Reverzní škály jsou v záznamovém listu označeny hvězdičkou (v záznamovém listu, se kterým pracují respondenti, se označení reverzních škál samozřejmě neuvádí).

Tab. 79 *Záznamový list se škálami sémantického diferenciálu*

NAŠE ŠKOLA

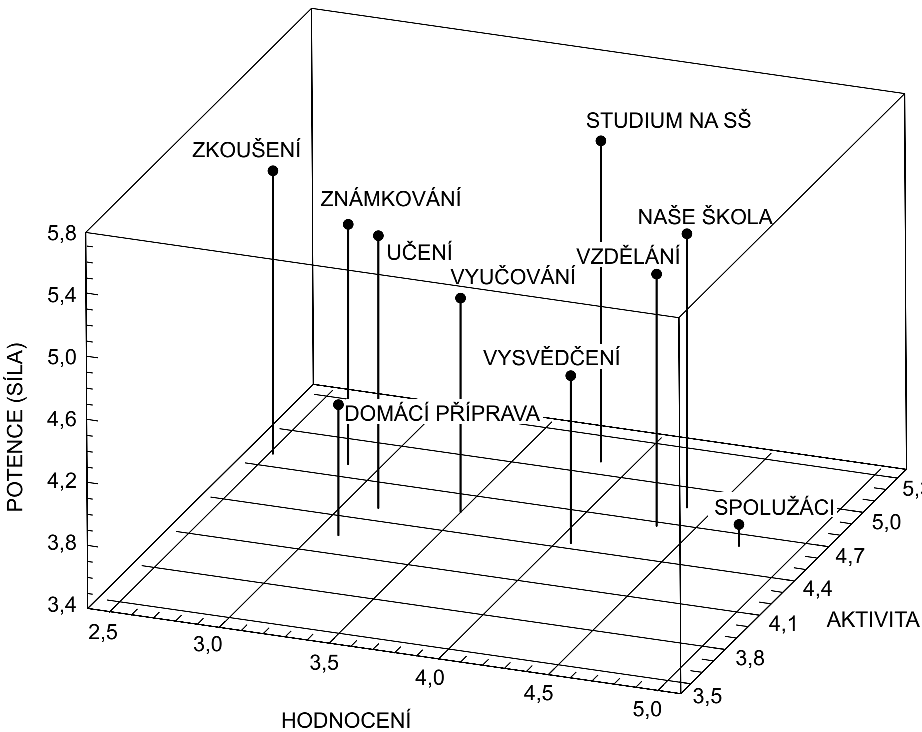
1.	příjemná		X						nepříjemná	h
2.	kulatá*					X			hranatá	a
3.	aktivní					X			pasivní	a
4.	krásná			X					ošklivá	h
5.	hrubá				X				jemná	p
6.	pomalá*				X				rychlá	a
7.	dobrá		X						špatná	h
8.	slabá*					X			silná	p
9.	tupá*			X					ostrá	a
10.	hluboká		X				X		mělká	p
11.	těžká		X						lehká	p
12.	tmavá*				X				světlá	h

Vybraným bodům na škálách (polohám křížků) byla potom přiřazena (zleva doprava) čísla 7–1. Při analýze výsledků byly nejdříve vypočítány průměry ve všech třech faktorech (tab. 80).

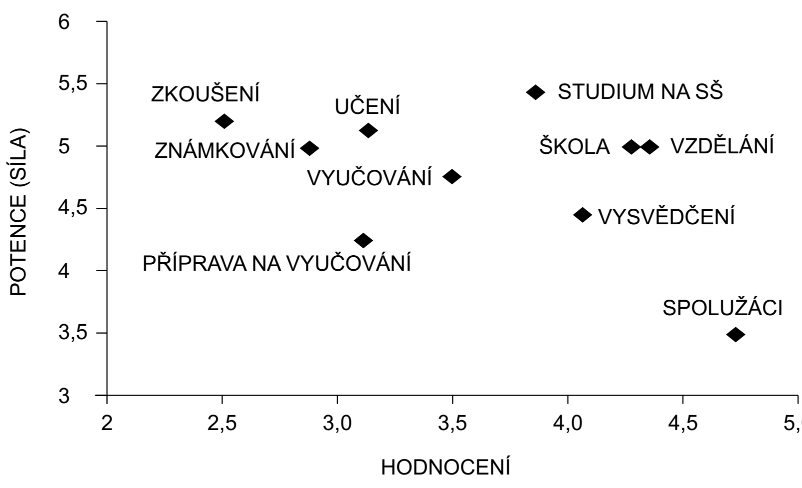
Tab. 80 *Posuzované pojmy (výzkum postojů u žáků základní školy)*

Pojem	Faktor hodnocení	Faktor potence	Faktor aktivity
domácí příprava na vyučování	3,12	4,25	4,25
naše škola	4,29	5,00	5,01
spolužáci	4,73	3,50	4,64
studium na střední škole	3,86	5,43	5,11
učení	3,14	5,12	4,53
vysvědčení	4,06	4,45	4,46
vyučování	3,50	4,76	4,57
vzdělání	4,35	5,01	4,98
zkoušení	2,51	5,20	4,85
známkování	2,88	4,98	4,75

Názornou představu o chápání psychologického významu vybraných pojmů u žáků základní školy si lze učinit na základě grafického zobrazení jejich sémantického prostoru (obr. 35). Někdy se ke grafickému znázornění výsledků používá jen dvojrozměrné zobrazení, do něhož se nepromítají výsledky ve faktoru aktivity (obr. 36).



***Obr. 35** Sémantický prostor žáků základní školy*



***Obr. 36** Dvojrozměrné sémantické pole žáků základní školy*

Pro přesnější popis sémantického prostoru je výhodné určit, jak daleko jsou jednotlivé pojmy v sémantickém prostoru od sebe vzdáleny. Jestliže se dva pojmy v sémantickém prostoru nalézají blízko u sebe, jsou si podobné významem. Naopak, jsou-li v sémantickém prostoru separovány, liší se psychologickým významem, který pro jedince mají. Jako míra vzdálenosti pojmů v sémantickém prostoru se nejčastěji používá tzv. **lineární distance *D*** (Kerlinger, 1972), kterou lze vypočítat ze vztahu

$$D_{ij} = \sqrt{\sum (x_i - x_j)^2} \tag{122}$$

kde D_{ij} je lineární distance mezi pojmem i a pojmem j, x_i je škálová hodnota pojmu i a x_j je škálová hodnota pojmu j.

Při výpočtu lineární distance D_{ij} mezi dvěma pojmy je možno vycházet buď jen z jednoho faktoru (nejčastěji faktoru hodnocení), nebo ze dvou faktorů (nejčastěji faktoru hodnocení a potence), anebo ze všech tří faktorů.

Vypočítané hodnoty lineární distance D se obvykle zapisují do symetrické D-matice (tab. 81). V uvedené matici byly hodnoty lineární distance vypočítány z faktorů hodnocení a potence (síly).

Tab. 81 *D-matice*

	domácí příprava	naše škola	spolužáci	studium na SŠ	učení	vysvědčení	vyučování	vzdělání	zkoušení	známkování
domácí příprava	x	1,39	1,77	1,39	0,87	0,96	0,64	1,45	1,13	0,77
naše škola		x	1,56	0,61	1,16	0,60	0,83	0,06	1,79	1,41
spolužáci			x	2,12	2,27	1,16	1,76	1,56	2,80	2,37
studium na SŠ				x	0,78	1,00	0,76	0,65	1,37	1,08
učení					x	0,67	0,51	1,22	0,64	0,30
vysvědčení						x	0,64	0,63	1,72	1,29
vyučování							x	0,89	1,08	0,66
vzdělání								x	1,85	1,47
zkoušení									x	0,43
známkování										x

Z D-matice je možno vyčíst, které pojmy jsou si významem blízké a které jsou významem vzdálené.

Při analýze dat v D-matici si všímáme zejména malých a velkých hodnot lineární distance D. Z uvedené matice můžeme například vyčíst, že velmi podobně pociťují žáci základní školy pojmy „vzdělání" a „naše škola". Poměrně dosti blízké jsou ale například i pojmy „učení" a „známkování". Největší významová vzdálenost je mezi pojmy „zkoušení" (pojem je „silný", ale ne příliš „dobrý") a „spolužáci" (pojem pociťovaný jako „dobrý", ale ne příliš „silný").

Výběr škál pro měření sémantickým diferenciálem

Škály sémantického diferenciálu je nutno volit podle povahy posuzovaných objektů. Škály uvedené v tabulce 79 se ukázaly jako vhodné pro posuzování pojmů, které mají vztah ke školní realitě žáků základní školy. V jiných případech (např. při posuzování uměleckých děl, literárních útvarů, osob apod.) mohou lépe vyhovovat škály jiné. Určitým vodítkem při volbě škál může být původní seznam padesáti posuzovacích škál, který uvádí C. Osgood (tab. 82).

Tab. 82 *Původní seznam padesáti škál sémantického diferenciálu podle C. Osgooda (1958)*

Č.	Škála	I	II	III	IV	h^2
1	dobrý – špatný	**0,88**	0,05	0,09	0,09	0,79
2	velký – malý	0,06	**0,62**	0,34	0,04	0,51
3	krásný – ošklivý	**0,86**	0,09	0,01	0,26	0,82
4	žlutý – modrý	**–0,33**	–0,14	0,12	0,17	0,17
5	tvrdý – měkký	–0,48	**0,55**	0,16	0,21	0,60
6	sladký – kyselý	**0,83**	–0,14	–0,09	0,02	0,72
7	silný – slabý	0,19	**0,62**	0,20	–0,03	0,46
8	čistý – špinavý	**0,82**	–0,05	0,03	0,02	0,68
9	vysoký – nízký	**0,59**	0,21	0,08	0,04	0,40
10	klidný – vzrušený	**0,61**	0,00	–0,36	–0,05	0,50
11	chutný – odporný	**0,77**	0,05	–0,11	0,00	0,61
12	hodnotný – bezcenný	**0,79**	0,04	0,13	0,00	0,64
13	červený – zelený	–0,33	–0,08	**0,35**	0,22	0,28
14	mladý – starý	0,31	–0,30	**0,32**	0,01	0,29
15	hodný – krutý	**0,82**	–0,10	–0,18	0,13	0,73
16	hlasitý – tichý	–0,39	**0,44**	0,23	0,22	0,45
17	hluboký – mělký	0,27	**0,46**	0,14	–0,25	0,37
18	příjemný – nepříjemný	**0,82**	–0,05	0,28	–0,12	0,77
19	černý – bílý	**–0,64**	0,31	0,01	–0,03	0,51
20	hořký – sladký	**–0,80**	0,11	0,20	0,03	0,69
21	veselý – smutný	0,76	–0,11	0,00	0,03	0,59
22	ostrý – tupý	0,23	0,07	**0,52**	–0,10	0,34
23	prázdný – plný	**–0,57**	–0,26	–0,03	0,18	0,43
24	divoký – klidný	**–0,69**	0,17	0,41	0,02	0,67
25	těžký – lehký	–0,36	**0,62**	–0,11	0,06	0,53

▶

Č.	Škála	I	II	III	IV	h^2
26	vlhký – suchý	0,08	0,07	–0,03	**–0,14**	0,03
27	posvátný – neuctivý	**0,81**	0,02	–0,10	0,01	0,67
28	uvolněný – napjatý	**0,55**	0,12	–0,37	–0,11	0,47
29	odvážný – zbabělý	**0,66**	0,44	0,12	0,03	0,64
30	dlouhý – krátký	0,20	**0,34**	0,13	–0,23	0,23
31	bohatý – chudý	**0,60**	0,10	0,00	–0,18	0,40
32	jasný – zamlžený	**0,59**	0,03	0,10	–0,16	0,38
33	horký – studený	–0,04	-0,06	**0,46**	0,07	0,22
34	tlustý – tenký	–0,06	**0,44**	–0,06	–0,11	0,21
35	pěkný – škaredý	**0,87**	–0,08	0,19	0,15	0,82
36	světlý – tmavý	**0,69**	–0,13	0,26	0,00	0,56
37	basový – pronikavý	–0,33	**0,47**	–0,06	0,02	0,33
38	hranatý – kulatý	–0,17	0,08	**0,43**	0,12	0,23
39	voňavý – páchnoucí	**0,84**	–0,04	–0,11	0,05	0,72
40	poctivý – nečestný	**0,85**	0,07	–0,02	0,16	0,75
41	aktivní – pasivní	0,14	0,04	**0,59**	–0,02	0,37
42	drsný – hladký	**–0,46**	0,36	0,29	0,10	0,44
43	čerstvý – okoralý	**0,68**	0,01	0,22	–0,11	0,52
44	rychlý – pomalý	0,01	0,00	**0,70**	-0,12	0,50
45	poctivý – nepoctivý	**0,83**	0,08	–0,07	0,11	0,71
46	hrbolatý – hladký	–0,42	**0,60**	0,26	0,27	0,68
47	blízký – vzdálený	**0,41**	0,13	0,11	–0,05	0,20
48	štiplavý – nedráždivý	**–0,30**	0,12	0,26	0,05	0,17
49	zdravý – nemocný	**0,69**	0,17	0,09	0,02	0,59
50	široký – úzký	0,26	**0,41**	–0,07	–0,11	0,25
	% z celkové variance	33,78	7,62	6,24	1,52	0,4916

C. Osgood uskutečnil faktorovou analýzu výsledků měření pro padesát škál sémantického diferenciálu a při této analýze extrahoval čtyři společné faktory (I, II, III, IV). Získanou matici faktorových nábojů podrobil pravoúhlé rotaci (srov. oddíl 3.5). V tabulce 82 jsou u jednotlivých škál tučným tiskem označeny dominantní faktorové náboje. Z uvedených výsledků analýzy je patrné, že praktický význam mají pouze první tři faktory. Poslední, čtvrtý faktor (IV) je již prakticky bezvýznamný, protože škály jsou jím jen velmi málo syceny (tento faktor objasňuje již jen velmi malou část celkové variability). Jak již bylo uvedeno, první nejsilnější faktor označuje C. Osgood jako **faktor hodnocení**, druhý faktor jako **faktor potence** a třetí faktor jako **faktor aktivity**.

Při ověřování vlastností škál uvedených v Osgoodově seznamu se ukázalo, že mezi vlastnostmi, které seznam uvádí, a vlastnostmi, které škály v našich sociálně kulturních podmínkách mají, existují určité rozdíly. Rozdíly byly zjištěny zejména ve faktorových nábojích u druhého a třetího faktoru (u faktoru potence a faktoru aktivity). Například u škál **velký** – **malý**, **tvrdý** – **měkký**, **hlasitý** – **tichý** se neprokázalo dominantní sycení faktorem potence, u škály *vysoký – nízký,* se nepotvrdilo dominantní sycení faktorem hodnocení apod. (Chráska, 1996, 1997).

Při měření škálami, jejichž vlastnosti neznáme, je na místě velká opatrnost. Pro ověření faktorové struktury škál je možné použít faktorovou analýzu. Podrobnosti o této proceduře lze nalézt například v pracích P. Blahuše (1985), M. Chrásky (1997) apod.

4.6.2 DVOUFAKTOROVÝ SÉMANTICKÝ DIFERENCIÁL ATER

Při ověřování vlastností a výpovědní hodnoty dat získaných sémantickým diferenciálem (úprava podle C. Osgooda) jsme dospěli k závěru, že ve většině situací je posuzování subjektivního významu pojmů pomocí tří faktorů příliš detailní a nepřináší o mnoho více informací než posuzování pomocí dvou faktorů. Navíc, třetí faktor (faktor aktivity) je natolik subtilní konstrukt, že u něj vždy hrozí značné nebezpečí nesprávné interpretace.

Ve výzkumech jsme ověřili, že při popisu pedagogické (edukační) reality je možné s dostatečnou mírou rozlišovací schopnosti používat sémantický diferenciál, který využívá pouze dva faktory. Tyto dva faktory byly nalezeny na základě faktorové analýzy výsledků posuzování pojmů, které sloužily jako indikátory při měření postojů žáků a studentů k edukační realitě. Nástroj, který jsme pro měření postojů k edukační realitě vytvořili (tab. 84), označujeme jako dotazník ATER (*Attitudes Towards Educational Reality*).

První, nejsilnější faktor označujeme (ve shodě s C. Osgoodem) jako **faktor hodnocení**. Vyjadřuje tedy, jak dalece je posuzovaný objekt osobami vnímán jako **„dobrý“** nebo **„špatný“**.

Druhý faktor, který označujeme jako **faktor energie**, vypovídá o tom, co Osgoodův faktor potence a aktivity dohromady (analogicky jako mechanická energie = potenciální energie + pohybová energie). Faktor energie tedy vyjadřuje, do jaké míry respondenti chápou posuzované pojmy jako **„něco“** spojeného s vydáváním energie, námahou, obtížemi, změnami nebo aktivitou. Tabulka 83 prezentuje jednotlivé škály nástroje ATER, jejich faktorové náboje, komunality a faktorovou identifikaci (h = hodnocení/e = energie). Je patrné, že v prezentovaném dotazníku je pět škál, které měří faktor hodnocení, a pět škál, které měří faktor energie.

Tab. 83 *Faktorové náboje škál dvoufaktorového sémantického diferenciálu ATER*

Číslo	Škála	Faktor hodnocení	Faktor energie	Komunalita h^2	Faktorová identifikace
1	dobrý – špatný	**0,74131**	0,37296	0,68864	h
2	náročný – nenáročný	0,14148	**0,83483**	0,71696	e
3	příjemný – nepříjemný	**0,74483**	0,36803	0,69022	h
4	světlý – tmavý	**0,79144**	0,18782	0,66165	h
5	přísný – mírný	0,27109	**0,77844**	0,67947	e

▶

Číslo	Škála	Faktor hodnocení	Faktor energie	Komunalita h^2	Faktorová identifikace
6	obtížný – snadný	0,30029	**0,83082**	0,78043	e
7	krásný – ošklivý	**0,83103**	0,25093	0,75358	h
8	problémový – bezproblémový	0,37107	**0,64433**	0,55285	e
9	sladký – kyselý	**0,76247**	0,21974	0,62964	h
10	těžký – lehký	0,32978	**0,78108**	0,71884	e

Poznámka: *Uvedená faktorová matice byla získána pomocí tzv. pravoúhlé rotace (typu Varimax), tzn. že oba faktory jsou považovány za navzájem nezávislé. Dominantní faktorové náboje jsou v tabulce 83 zvýrazněny tučným tiskem (srov. oddíl 3.5).*

Tabulka 84 uvádí záznamový list a škály dvoufaktorového sémantického diferenciálu ATER. Aby se snížilo nebezpečí stereotypního posuzování ve škálách, je polovina škál prezentována v **reverzní podobě**. Reverzní škály jsou v záznamovém listu označeny hvězdičkou.

Tab. 84 *Záznamový list a škály dvoufaktorového sémantického diferenciálu ATER*

DOMÁCÍ PŘÍPRAVA

1. dobrá								špatná	*h*
2. nenáročná*								náročná	*e*
3. nepříjemná*								příjemná	*h*
4. světlá								tmavá	*h*
5. přísná								mírná	*e*
6. snadná*								obtížná	*e*
7. krásná								ošklivá	*h*
8. problémová								bezproblémová	*e*
9. kyselá*								sladká	*h*
10. lehká*								těžká	*e*

Metoda sémantického diferenciálu je použitelná pro měření mnoha proměnných, které jsou jinými metodami jen velmi obtížně přístupné. Například J. Vala a M. Chráska (2014) se úspěšně pokusili pomocí metody sémantického diferenciálu měřit **recepci poezie** u žáků základních a středních škol (škály použitého sémantického diferenciálu jsou uvedeny v souvislosti s vysvětlením principu faktorové analýzy (viz příklad 50).

4.7 MĚŘENÍ OBTÍŽNOSTI UČEBNÍHO TEXTU

V pedagogických výzkumech orientovaných na oblast vzdělávání má velký význam měření obtížnosti učebního textu. Obtížnost učebního textu lze posuzovat v podstatě dvojím způsobem.

První způsob spočívá v subjektivním odhadu, který realizují buď experti, nebo uživatelé učebního textu (např. žáci nebo studenti). Často se přitom používá ratingu, kdy posuzovatelé pracují s posuzovacími škálami typu (Gavora, 2000):

jednoduchý jazyk	1	2	3	4	5	N	složitý jazyk
těžké pojmy	1	2	3	4	5	N	přiměřené pojmy
podrobný výklad	1	2	3	4	5	N	všeobecný výklad
zajímavý text	1	2	3	4	5	N	nudný text

Druhý, dokonalejší a spolehlivější způsob měření obtížnosti učebního textu je založen na objektivních parametrech textu, jako je například délka věty, složitost věty, množství a obtížnost pojmů, stupeň opakování pojmů atd.

Komplexní míra obtížnosti textu

Byla vyvinuta celá řada metod, jak z těchto objektivně zjistitelných údajů vypočítat určitý kvantitativní ukazatel (míru) obtížnosti textu. U nás je nejznámější metodou určování tzv. **komplexní míry obtížnosti textu** (autoři Nestlerová, Průcha, Pluskal).

Komplexní míra obtížnosti textu *T* je určena zejména pro měření obtížnosti výkladového textu učebnic. Data potřebná pro výpočet *T* se získávají ze vzorků textu, které jsou získávány podle předem stanovených, standardních pravidel. Většinou se vybírá pět nebo deset vzorků z posuzovaného textu, každý v délce 200 slov; podrobnosti lze nalézt v práci J. Průchy (1998).

Komplexní míru obtížnosti učebního textu je možno vypočítat ze vztahu

$$T = T_s + T_p \tag{123}$$

kde T je komplexní míra obtížnosti textu, T_s je syntaktická obtížnost textu a T_p je sémantická obtížnost textu.

Syntaktická obtížnost T_s je dána průměrnou délkou vět a průměrnou délkou větných úseků a lze ji vypočítat podle vzorce

$$T_s = 0{,}1 \cdot V \cdot U \tag{124}$$

kde V je průměrná délka vět (ve všech vybraných vzorcích textu) a U je průměrná délka větných úseků. Větou se v tomto případě rozumí jakákoli posloupnost slov začínající velkým písmenem a končící tečkou nebo jiným grafickým znakem (dvojtečka, otazník apod.) Průměrnou délkou větného úseku se rozumí počet slov v textu, který připadá na jedno sloveso v určitém tvaru (nepočítají se tedy slovesa v infinitivu).

Sémantická obtížnost T_p je dána četnostmi výskytu pěti následujících kategorií pojmů:

P_1… běžné pojmy;
P_2… odborné pojmy;
P_3… faktografické pojmy;
P_4… číselné údaje;
P_5… opakované pojmy.

Za pojmy se v tomto případě považují všechna podstatná jména včetně podstatných jmen abstraktních a dějových, zpodstatnělých přídavných jmen, osobních jmen a příjmení a zkratek označujících různé pojmy.

Sémantická obtížnost učebního textu se vypočítá podle vzorce

$$T_p = 100 \cdot \frac{P}{N} \cdot \frac{P_1 + 3P_2 + 2P_3 + 2P_4 + P_5}{N} \qquad (125)$$

kde T_p je sémantická obtížnost textu, P je celkový počet pojmů v textu, N je celkový počet slov v textu a P_1, P_2, P_3, P_4 a P_5 jsou četnosti výskytu jednotlivých kategorií pojmů v textu.

Příklad 63

Z posuzované učebnice pro žáky na druhém stupni základní školy bylo vybráno pět vzorků textu. V těchto vzorcích bylo celkem 1 048 slov (N = 10 048) a 88 vět (V = 88). Dále bylo zjištěno, že v posuzovaném textu je 104 sloves a byly určeny četnosti výskytu jednotlivých kategorií pojmů $P_1 = 80$, $P_2 = 157$, $P_3 = 30$, $P_4 = 23$, $P_5 = 66$. Z uvedených informací vypočítáme komplexní míru obtížnosti učebního textu T.

Nejdříve určíme syntaktickou obtížnost textu. Protože ve vzorku textu (celkem 1 048 slov) je 88 vět, činí průměrná délka věty 1 048 : 88 = 11,9 slova. Průměrná délka větného úseku je 1 048 : 104 = 10,1 slova. Syntaktická obtížnost textu potom podle vzorce (124) vychází

$$T_s = 0{,}1 \cdot V \cdot U = 0{,}1 \cdot 11{,}9 \cdot 10{,}1 = 12{,}0$$

Ve vzorku textu bylo celkem 356 pojmů $P = 80 + 157 + 30 + 23 + 66 = 356$
Sémantickou obtížnost textu vypočítáme dosazením příslušných hodnot do vzorce (125)

$$T_p = 100 \cdot \frac{P}{N} \cdot \frac{P_1 + 3P_2 + 2P_3 + 2P_4 + P_5}{N} = 100 \cdot \frac{356}{1\,048} \cdot \frac{80 + 3 \cdot 157 + 2 \cdot 30 + 2 \cdot 23 + 66}{1\,048} = 23{,}4$$

Komplexní míru obtížnosti učebního textu T potom dostáváme ze vzorce (123)

$$T = T_s + T_p = 12{,}0 + 23{,}4 = 35{,}4$$

Hodnota T může nabývat hodnot od 1 (minimální obtížnost) do 100 (maximální obtížnost). J. Průcha (1998) uvádí, že v učebnicích pro základní školu se hodnoty T pohybují v pásmu od $T = 27$ do $T = 63$.

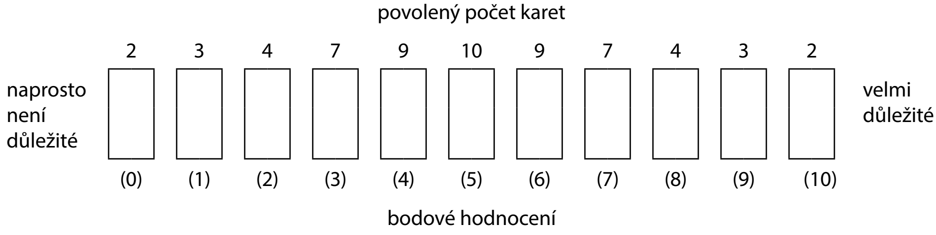
4.8 Q-METODOLOGIE

Názvem Q-metodologie se označuje skupina psychometrických a statistických procedur, které vyvinul v padesátých letech 20. století William Stephenson. Metoda je výhodná ve výzkumech, v nichž máme zjistit, jak určitá skupina respondentů hodnotí určitou množinu objektů, přičemž těchto objektů je velký počet. U Q-metodologie se zkoumaným osobám předkládá soubor (balíček) karet, na nichž jsou uvedeny objekty, které se mají hodnotit (např. výpovědi, názory, estetické objekty, životní hodnoty apod.) s tím, že je mají roztřídit podle určitého kritéria. Jednotlivé karty se označují jako Q-typy. Q-typy je možno třídit podle různých kritérií, například podle významu nebo důležitosti pro zkoumanou osobu, podle vlivu na něco, podle toho, jaký vztah osoba k objektu má apod. Počet karet, které se předkládají ke třídění, bývá obvykle dosti vysoký (jako optimální se doporučuje 60–120).

Příklady Q-typů (posuzování významu pedagogických vědomostí a dovedností pro učitelskou práci):

- umět pracovat s materiálními didaktickými prostředky;
- umět vést pedagogickou dokumentaci;
- znát základní metody výuky a umět v dané situaci zvolit metodu adekvátní;
- umět správně formulovat výukové cíle;
- znát teoretické základy vyučovacího procesu;
- ovládat techniky relaxačních cvičení a umět je vhodně zařazovat do výuky;
- umět adekvátně hodnotit učební výkon žáka;
- umět používat alternativní didaktické postupy;
- umět správně provádět individuální ústní zkoušení;
- umět formulovat učební úlohy v závislosti na výukových cílech;
- umět projektovat vzdělávací programy a umět projektovat výuku;
- umět správně provádět frontální ústní zkoušení;
- umět komunikovat se žáky, učiteli a rodiči;
- umět se pohotově rozhodovat ve standardních i neobvyklých situacích.

Po zkoumané osobě se požaduje, aby karty rozdělila do několika hromádek podle stanoveného kritéria (např. podle důležitosti, významu atd.). Nejčastěji se požaduje takové rozdělení, které přibližně odpovídá normálnímu rozdělení – tzv. **kvazinormální distribuce**. Například při počtu šedesáti karet je možno požadovat třídění do jedenácti hromádek podle schématu, který uvádí obrázek 37.



Obr. 37 *Schéma Q-třídění karet (kvazinormální distribuce)*

Čísla 2, 3, 4 ... 4, 3, 2 v obrázku uvádějí povolené počty karet pro jednotlivé hromádky. Kvazinormální distribuce má výhodu v tom, že umožňuje a usnadňuje statistické zpracování výsledků. Někdy se u Q-metodologie používá i tzv. **pravoúhlá distribuce** Q-typů, při níž se požaduje umístění stejného počtu karet do všech hromádek.

Příklad 64

Použití Q-metodologie budeme ilustrovat na příkladu výzkumu, který zjišťoval význam jednotlivých vědomostí a dovedností z pedagogiky pro učitelskou praxi (Chráska, 1994). K šetření byl vybrán vzorek zkušených učitelů základních škol (expertů) o rozsahu $n = 25$. Této skupině expertů byl předložen soubor šedesáti Q-typů (karet), na nichž byly prezentovány jednotlivé vědomosti nebo dovednosti z pedagogiky. Q-typy navrhli vyučující příslušných pedagogických disciplín na pedagogické fakultě. Použité Q-typy byly strukturovány do pěti skupin:

- A obecné vědomosti z pedagogiky;
- B vědomosti z dějin pedagogiky;
- C vědomosti a dovednosti z pedagogické diagnostiky;
- D didaktické vědomosti a dovednosti;
- E vědomosti (dovednosti) o výchově.

Expertům byl předložen balíček karet, na nichž byly uvedeny jednotlivé vědomosti nebo dovednosti z pedagogiky s tím, že je mají roztřídit podle důležitosti pro učitelskou práci. Výsledky třídění (čísla karet) potom hodnotitelé zapisovali do připraveného formuláře (obr. 38).

					7					
					11					
				12	3	20				
				58	33	27				
			1	40	37	35	2			
			21	44	50	48	43			
			38	52	60	54	46			
			45	57	56	59	29			
		23	24	42	49	39	6	5		
		51	8	18	34	36	47	17		
	31	22	9	26	32	28	13	55	53	
nejméně důležité	25	41	10	19	15	16	14	30	4	nejdůležitější

***Obr. 38** Formulář pro výsledky Q-třídění (výsledky hodnocení jednoho z expertů)*

Umístěním karet do určitých hromádek experti vyjádřili svůj názor na důležitost příslušných vědomostí nebo dovedností z pedagogiky. Jednotlivým polohám umístění Q-typů zleva doprava bylo potom při zpracování výsledků připisováno bodové hodnocení od nuly do deseti (čísla v závorkách). Umístění karty v první hromádce zleva na obrázku tedy znamená, že hodnotitel se domnívá, že vědomost nebo dovednost na kartě uvedená je pro učitelskou práci zcela nedůležitá a tato skutečnost je hodnocena 0 body. Jestliže například zkoumaná osoba umístí jistou kartu v prostřední hromádce (tam je povoleno položit deset karet), znamená to, že připouští, že daná vědomost nebo dovednost má pro učitelskou práci „středně" velký význam a tomuto hodnocení odpovídá pět bodů apod. S uvedeným bodovým hodnocením zkoumané osoby nepracují, ale opírá se o ně celá další statistická analýza.

Naznačeným postupem získáváme informace o tom, jak zkoumané osoby vidí a hodnotí objekty předkládané na jednotlivých Q-typech. V popisovaném příkladu výzkumu byly například zajímavé informace získány rozborem Q-typů, které považovali experti za nejdůležitější a naopak za nejméně důležité pro učitelskou práci (tab. 85 a 86).

Z tabulky 85 je zřejmé, že experti jednoznačně preferují ty pedagogické vědomosti a dovednosti, které mají bezprostřední vztah k jejich didaktické činnosti. Vedle vědomostí a dovedností z didaktiky se však objevují i vědomosti a dovednosti související s výchovou žáků (Q-typy č. 47, 6). Zkušení učitelé připisují velký význam i dovednostem obecnějšího rázu (č. 53, 55), jejichž rozvoj lze jen obtížně přisoudit některé z pedagogických disciplín.

Jestliže studujeme tabulku 86, která uvádí osm (podle expertů nejméně důležitých) vědomostí nebo dovedností z pedagogiky, zjišťujeme, že jsou to vesměs vědomosti získávané studiem dějin pedagogiky.

Při rozboru údajů (uvedených v tabulkách 85 a 86) věnujeme (vedle průměrného hodnocení) pozornost také hodnotám směrodatné odchylky. Směrodatná odchylka vypovídá o tom, jak dalece jsou experti při hodnocení daného Q-typu zajedno, respektive jak dalece se v názorech na význam daného Q-typu liší.

Tab. 85 *Vědomosti a dovednosti z pedagogiky, které považují zkušení učitelé za nejdůležitější*

Pořadí	Q-typ		Průměrné hodnocení	Směrodatná odchylka
1	4.	znát základní metody výuky a umět v dané situaci zvolit metodu adekvátní	9,28	0,79
2	53.	umět komunikovat se žáky, učiteli a rodiči	8,56	1,56
3	17.	umět adekvátně hodnotit učební výkon žáka	8,28	1,28
4	47.	znát metodiku povzbuzování a trestání a umět ji aplikovat ve třídě	7,68	1,31
5–6	5.	umět formulovat výukové cíle	7,64	1,78
5–6	30.	umět formulovat učební úlohy v závislosti na výukových cílech	7,64	1,44
7	55.	umět se pohotově rozhodovat ve standardních i neobvyklých situacích	7,52	1,94
8	6.	znát a umět uplatnit metody vysvětlování, přesvědčování a příkladu ve výchově žáků	7,32	1,22
9	29.	umět správně provádět individuální ústní zkoušení	7,08	1,04
10	13.	znát teoretické základy vyučovacího procesu	6,96	1,81

***Tab. 86** Vědomosti a dovednosti z pedagogiky, které považují zkušení učitelé za nejméně důležité*

Pořadí	Q-typ		Průměrné hodnocení	Směrodatná odchylka
1	25.	umět zhodnotit přínos hlavních představitelů ruské a sovětské pedagogiky	0,40	0,76
2	31.	umět charakterizovat systém středověké výchovy	0,52	0,87
3	22.	umět analyzovat vývoj cílů a obsahu výchovy v antickém Řecku	1,32	0,80
4	51.	chápat příčiny a teoretická východiska reformního hnutí první poloviny 20. století	1,40	0,82
5	23.	umět charakterizovat pedagogický systém J. F. Herbarta	1,44	0,92
6	41.	umět srovnat pedagogické názory O. Chlupa a V. Příhody	1,56	1,04
7	24.	umět charakterizovat pedagogické názory v české pedagogice na přelomu 19. a 20. století	1,80	0,71
8	8.	umět charakterizovat základní mezníky ve vývoji českého školství od poloviny 18. století do současnosti	1,92	0,81

Další možností interpretace výsledků Q-třídění je zjištění, zda mezi hodnocením karet u jednotlivých expertů jsou podobnosti a jak jsou tyto podobnosti velké. K exaktnímu posouzení těsnosti vztahu mezi tříděním dvou osob lze užít výpočtu Pearsonova korelačního koeficientu. Tento výpočet je v případě kvazinormální distribuce výsledků usnadněn tím, že u všech osob je stejné průměrné hodnocení a stejný je i celkový součet všech hodnot. Koeficient korelace lze v tomto případě vypočítat podle vzorce

$$r_p = \frac{\sum xy - n\bar{x}^2}{\sum x^2 - n\bar{x}^2} \tag{126}$$

kde r_p je Pearsonův koeficient korelace, x je hodnocení určitého Q-typu u jedné osoby, y je hodnocení tohoto Q-typu u druhé osoby, n je počet Q-typů a $\bar{x}$ je průměrné hodnocení jedné osoby. Při výpočtech koeficientu korelace stačí pro každou dvojici osob vypočítat jen hodnotu $\sum xy$, protože ostatní hodnoty uváděné ve vzorci jsou pro všechny dvojice osob stejné.

***Tab. 87** Korelační matice z hodnot Q-třídění*

	A	B	C	D
A	x	0,18	0,68	0,13
B	0,18	x	0,06	0,58
C	0,68	0,06	x	0,12
D	0,13	0,58	0,12	x

Koeficienty korelace se vypočítávají pro všechny možné dvojice zkoumaných osob a výsledky se zapisují do korelační matice, jejíž příklad uvádí tabulka 89. Matice obsahuje korelační koeficienty pro hodnocení čtyř učitelů (A, B, C, D) ve výzkumu, který se zabýval hodnocením vědomostí a dovedností z pedagogiky.

Z korelační matice je patrné, že značná podobnost se projevuje mezi tříděním Q-typů u učitelů A a C a také u učitelů B a D. Naopak mezi tříděním učitelů A a B, A a D a učitelů B a C je podobnost velmi malá.

Přesnější informace o **„typech“** třídění u zkoumaných osob může přinést **faktorová**, respektive **shluková analýza** (srov. oddíly 3.5 a 3.6). Pomocí faktorové analýzy lze získat informace například o tom, zda se mezi hodnotícími osobami vyskytovaly určité typy osob, které mají na hodnocení předložených objektů (Q-typů) podobné názory.

Při použití Q-metodologie získáváme poměrně snadno velký počet relativně velice spolehlivých dat. Pro exaktní kontrolu míry spolehlivosti a přesnosti provedeného měření lze doporučit **výpočet koeficientu reliability**. Z doporučovaných metod statistického odhadování reliability u Q-metodologie (Chráska, 1996) lze doporučit **Kendallův koeficient shody W** nebo některý z něho vycházejících koeficientů, například **Taylorův koeficient shody**. Byla také vyvinuta řada koeficientů, které umožňují odhadovat reliabilitu na základě analýzy rozptylu. Z jednoduché (jednofaktorové) analýzy rozptylu vychází například tzv. **koeficient vnitrotřídní korelace**. O podrobnostech těchto postupů se lze informovat v odborné literatuře (např. Bintig, 1980).

Proti používání Q-metodologie bývají někdy vznášeny určité statistické námitky. Statistické operace totiž většinou vyžadují nezávislost odpovědí, což u nucené volby (jak se u Q-metodologie provádí) není přesně splněno. Při velkém počtu karet (Q-typů) však již tato námitka nehraje příliš velkou roli. Navíc lze spolehlivost závěrů zvýšit tím, že zvýšíme požadavky na statistickou významnost (např. místo hladiny významnosti 0,05 použijeme hladinu významnosti 0,01). Někdy se proti Q-metodologii namítá i to, že nucený výběr představuje nepřiměřený nátlak na zkoumanou osobu, která je nucena přizpůsobit se „nerozumnému požadavku“ (Kerlinger, 1972).

Q-metodologie je vhodná zejména k intenzivnímu zkoumání malých skupin osob. Výhodou je také to, že třídění je možné libovolně a mnohokrát opakovat. Q-metodologii lze doporučit zvláště pro objevování nových oblastí výzkumu, kdy na základě práce s malým (ale pečlivě vybraným) vzorkem můžeme získat první informace, které potom dalšími metodami (a na základě přiměřeně velkého výběru) verifikujeme. F. N. Kerlinger (1972) hovoří v této souvislosti o používání Q-metodologie k „heuristickým účelům“.

ZÁVĚREM

Jednotlivé věci, osoby nebo jevy, které nás obklopují, se navzájem odlišují nejen určitými vlastnostmi, ale také mírou těchto vlastností. Chceme-li co nejlépe postihnout kvalitativní rozdíly mezi jevy, musíme se u jednotlivých jevů snažit zachytit i jejich množství, tj. kvantitu. Poznání reality tedy vyžaduje postižení jak kvalitativní, tak kvantitativní stránky zkoumaného jevu.

Při kvalitativním přístupu k realitě získáváme konkrétní, názorný a plastický obraz skutečnosti, který ovšem vychází ze zkušeností a názorů badatele, a je proto často subjektivní. Kvantitativní přístup umožňuje hlubší poznání skutečnosti v její racionální obecnosti. Lze také říci, že při kvalitativním přístupu jde spíše o charakteristiku jedinečnosti různorodých prvků, zatímco při kvantitativním přístupu se postihují četnosti stejnorodých prvků.

K výhodám kvantitativního přístupu patří zejména přehlednost, stručnost a syntetičnost výsledku. Při kvantitativním přístupu je eliminována mnohoznačnost slov (např. při stupňování vlastností posuzovaných objektů). Vysoká míra abstrakce, která je pro číselné nebo funkční vyjadřování závislostí typická, může však v některých případech vést k simplifikaci skutečnosti tím, že se například ztrácí typická pestrost nebo proměnlivost zkoumaných objektů. Toto nebezpečí je možné minimalizovat tím, že každý jev se snažíme postihovat ne pomocí jedné číselné charakteristiky, nýbrž pomocí několika (mnoha) číselných charakteristik současně. Tyto číselné charakteristiky potom představují dílčí pohledy na daný jev a umožňují většinou ve svém souhrnu postižení jevu s uspokojivou přesností a spolehlivostí.

Účelně uplatňovaný kvantitativní přístup vyžaduje, aby čísla, se kterými pracujeme, měla náležitou výpovědní hodnotu. U každého čísla, které při kvantitativní analýze používáme, je třeba důkladně zvažovat, zda (a jak dalece) vypovídá právě o tom aspektu zkoumané reality, který máme na mysli (problém validity měření).

Úsilí o kvantitativní uchopení reality má pochopitelně smysl jen tehdy, jestliže směřuje k hlubšímu proniknutí k podstatě pedagogických jevů. Není-li tomu tak, hrozí nebezpečí, že se úsilí o kvantitativní postižení pedagogické reality stane jen nebezpečnou hrou s čísly.

Při měření jevů v pedagogických výzkumech je třeba mít na paměti, že některé stránky pedagogické reality jsou tak složité, že je velmi obtížné (ne-li nemožné) je kvantitativně zachytit. Tato obtížnost by však neměla být důvodem k tomu, aby pedagogická věda na řešení těchto problémů předem rezignovala. Zkušenosti ukazují, že i ty jevy, které jsou velmi subtilní a obtížně přístupné, se často podaří změřit s až překvapující přesností.

Je ovšem třeba připustit, že existují stránky pedagogické reality (nebo takové vlastnosti), které není možno sebedokonaleji promyšleným systémem měření kvantitativně zachytit. V těchto případech nezbývá pedagogické vědě než s pokorou uznat své ohraničené možnosti a přenechat řešení problému filozofii, náboženství či umění. Vědě přísluší, podle našeho názoru, jen zkoumání empiricky ověřitelných jevů. Pokud by věda předstírala, že její možnosti jsou širší, jde buď o nedorozumění, anebo o nebezpečný sebeklam.

PŘÍLOHA – STATISTICKÉ TABULKY

I DISTRIBUČNÍ FUNKCE NORMOVANÉHO NORMÁLNÍHO ROZDĚLENÍ

u	***Φ*(*u*)**	***u***	***Φ*(*u*)**	***u***	***Φ*(*u*)**
0,00	0,5000	0,50	0,6915	1,00	0,8413
0,02	0,5080	0,52	0,6985	1,02	0,8461
0,04	0,5160	0,54	0,7054	1,04	0,8508
0,06	0,5239	0,56	0,7123	1,06	0,8554
0,08	0,5319	0,58	0,7190	1,08	0,8599
0,10	0,5398	0,60	0,7257	1,10	0,8643
0,12	0,5473	0,62	0,7324	1,12	0,8686
0,14	0,5557	0,64	0,7389	1,14	0,8729
0,16	0,5636	0,66	0,7454	1,16	0,8770
0,18	0,5714	0,68	0,7517	1,18	0,8810
0,20	0,5793	0,70	0,7580	1,20	0,8849
0,22	0,5871	0,72	0,7642	1,22	0,8888
0,24	0,5948	0,74	0,7703	1,24	0,8925
0,26	0,6026	0,76	0,7764	1,26	0,8962
0,28	0,6103	0,78	0,7823	1,28	0,8997
0,30	0,6179	0,80	0,7881	1,30	0,9032
0,32	0,6255	0,82	0,7939	1,32	0,9066
0,34	0,6331	0,84	0,7995	1,34	0,9099
0,36	0,6406	0,86	0,8051	1,36	0,9131
0,38	0,6480	0,88	0,8106	1,38	0,9162
0,40	0,6554	0,90	0,8159	1,40	0,9192
0,42	0,6628	0,92	0,8212	1,42	0,9222
0,44	0,6700	0,94	0,8264	1,44	0,9251
0,46	0,6772	0,96	0,8315	1,46	0,9279
0,48	0,6844	0,98	0,8365	1,48	0,9306
0,50	0,6915	1,00	0,8413	1,50	0,9332

Pro záporné hodnoty veličiny určíme hodnotu distribuční funkce podle vzorce

$$\Phi(-u) = 1 - \Phi(u)$$

Příklad

$$\Phi(-0{,}92) = 1 - \Phi(0{,}92) = 1 - 0{,}8212 = 0{,}1788$$

I DISTRIBUČNÍ FUNKCE NORMOVANÉHO NORMÁLNÍHO ROZDĚLENÍ (POKRAČOVÁNÍ)

u	*Φ*(*u*)	*u*	*Φ*(*u*)	*u*	*Φ*(*u*)
1,50	0,9332	2,00	0,9772	2,50	0,99380
1,52	0,9357	2,02	0,9783	2,55	0,99460
1,54	0,9382	2,04	0,9793	2,60	0,99530
1,56	0,9406	2,06	0,9803	2,65	0,99600
1,58	0,9429	2,08	0,9812	2,70	0,99650
1,60	0,9452	2,10	0,9821	2,75	0,99700
1,62	0,9474	2,12	0,9830	2,80	0,99740
1,64	0,9495	2,14	0,9838	2,85	0,99780
1,66	0,9515	2,16	0,9846	2,90	0,99810
1,68	0,9535	2,18	0,9854	2,95	0,99840
1,70	0,9554	2,20	0,9861	3,00	0,99865
1,72	0,9573	2,22	0,9868	3,05	0,99886
1,74	0,9591	2,24	0,9875	3,10	0,99903
1,76	0,9608	2,26	0,9881	3,15	0,99918
1,78	0,9625	2,28	0,9887	3,20	0,99931
1,80	0,9641	2,30	0,9893	3,25	0,99942
1,82	0,9656	2,32	0,9898	3,30	0,99952
1,84	0,9671	2,34	0,9904	3,35	0,99960
1,86	0,9686	2,36	0,9909	3,40	0,99966
1,88	0,9699	2,38	0,9913	3,45	0,99972
1,90	0,9713	2,40	0,9918	3,50	0,99977
1,92	0,9726	2,42	0,9922	3,60	0,99984
1,94	0,9738	2,44	0,9927	3,70	0,99989
1,96	0,9750	2,46	0,9931	3,80	0,99993
1,98	0,9761	2,48	0,9934	3,90	0,99995
2,00	0,9772	2,50	0,9938	4,00	0,99997

II KRITICKÉ HODNOTY TESTOVÉHO KRITÉRIA CHÍ-KVADRÁT

Stupně volnosti	Hladina významnosti	
	0,050	**0,010**
1	3,841	6,635
2	5,991	9,210
3	7,815	11,341
4	9,488	13,277
5	11,070	15,086
6	12,592	16,812
7	14,067	18,475
8	15,507	20,090
9	16,919	21,666
10	18,307	23,209
11	19,675	24,725
12	21,026	26,217
13	22,362	27,688
14	23,685	29,141
15	24,996	30,578
16	26,296	32,000
17	27,587	33,409
18	28,868	34,805
19	30,144	36,191
20	31,410	37,576
21	32,671	38,932
22	33,924	40,289
23	35,172	41,638
24	36,415	42,980
25	37,652	44,314
26	38,885	45,642
27	40,113	46,963
28	41,337	48,278
29	42,557	49,588
30	43,773	50,892

III ZNAMÉNKOVÝ TEST

Počty znamének řidčeji se vyskytujícího druhu pro znaménkový test
(oboustranný test)

Nulovou hypotézu *odmítáme* na hladině významnosti 0,05, jestliže zjištěný počet znamének řidčeji se vyskytujícího druhu je **menší nebo roven tabelované hodnotě**.

Počet dvojic hodnot	Počet znamének	Počet dvojic hodnot	Počet znamének
5	–	48	16
6	0	49	17
7	0	50	17
8	0		
9	1	51	18
10	1	52	18
		53	18
11	1	54	19
12	2	55	19
13	2	56	20
14	2	57	20
15	3	58	21
16	3	59	21
17	4	60	21
18	4		
19	4	61	22
20	5	62	22
		63	23
21	5	64	23
22	5	65	24
23	6	66	24
24	6	67	25
25	7	68	25
26	7	69	25
27	7	70	26
28	8		
29	8	71	26
30	9	72	27
		73	27

▶

Počet dvojic hodnot	Počet znamének	Počet dvojic hodnot	Počet znamének
31	9	74	28
32	9	75	28
33	10	76	28
34	10	77	29
35	11	78	29
36	11	79	30
37	12	80	30
38	12		
39	12	81	31
40	13	82	31
		83	32
41	13	84	32
42	14	85	32
43	14	86	33
44	15	87	33
45	15	88	34
46	15	89	34
47	16	90	35

IV KRITICKÉ HODNOTY T_α PRO WILCOXONŮV TEST

(oboustranný test)

n	**Hladina významnosti**	
	0,05	**0,01**
6	0	–
7	2	–
8	3	0
9	5	2
10	8	3
11	10	5
12	13	7
13	17	10
14	21	13
15	25	16
16	29	20
17	34	23
18	40	28
19	46	32
20	52	38
21	58	43
22	65	49
23	73	55
24	81	61
25	89	68

V KRITICKÉ HODNOTY TESTOVÉHO KRITÉRIA U_α pro hladinu významnosti 0,05

(oboustranný test)

n_2 \ n_1	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3																		
4		0																
5	0	1	2															
6	1	2	3	5														
7	1	3	5	6	8													
8	2	4	6	8	10	13												
9	2	4	7	10	12	15	17											
10	3	5	8	11	14	17	20	23										
11	3	6	9	13	16	19	23	26	30									
12	4	7	11	14	18	22	26	29	33	37								
13	4	8	12	16	20	24	28	33	37	41	45							
14	5	9	13	17	22	26	31	36	40	45	50	55						
15	5	10	14	19	24	29	34	39	44	49	54	59	64					
16	6	11	16	21	26	31	37	42	48	53	59	64	70	75				
17	6	11	17	22	28	34	39	45	51	57	63	69	75	81	87			
18	7	12	18	24	30	36	42	48	55	61	67	74	80	86	93	99		
19	7	13	19	25	32	38	45	52	58	65	72	78	85	92	99	106	113	
20	8	14	20	27	34	41	48	55	62	69	76	83	90	98	105	112	120	127

VI KRITICKÉ HODNOTY C (n_1, n_2) PRO KOLMOGOROVŮV--SMIRNOVŮV TEST
pro hladinu významnosti α = 0,05

(oboustranný test)

n_1	n_2	C	M	n_1	n_2	C	M	n_1	n_2	C	M	n_1	n_2	C	M
3	5	15	15	6	7	30	42	10	11	60	110	16	17	124	272
	6	6	6		8	17	24		12	33	60		18	64	144
	7	21	21		9	13	18		13	70	130		19	133	304
	8	21	24		10	20	30		14	37	70		20	35	80
	9	8	9		11	43	66		15	16	30				
	10	27	30		12	8	12		16	42	80	17	18	133	306
	11	30	33		13	52	78		17	89	170		19	141	323
	12	10	12		14	27	42		18	46	90		20	146	340
	13	33	39		15	19	30		19	94	190				
	14	36	42		16	30	48		20	11	20	18	19	142	342
	15	12	15		17	62	102						20	76	180
	16	39	48		18	12	18	11	12	72	132				
	17	42	51		19	70	114		13	75	143	19	20	160	380
	18	15	18		20	36	60		14	82	154				
	19	45	57						15	84	165	20	21	173	420
	20	48	60	7	8	40	56		16	89	176				
					9	42	63		17	93	187				
4	5	20	20		10	46	70		18	97	198				
	6	10	12		11	48	77		19	102	209				
	7	24	28		12	53	84		20	107	220				
	8	7	8		13	56	91								
	9	28	36		14	9	14	12	13	81	156				
	10	15	20		15	62	105		14	43	84				
	11	33	44		16	64	112		15	31	60				
	12	9	12		17	68	119		16	24	48				
	13	39	52		18	72	126		17	100	204				
	14	21	28		19	76	133		18	18	36				
	15	44	60		20	79	140		19	108	228				
	16	12	16						20	29	60				
	17	48	68	8	9	46	72								
	18	25	36		10	24	40	13	14	89	182				
	19	53	76		11	53	88		15	96	195				
	20	15	50		12	15	24		16	101	208				
					13	62	104		17	105	221				
5	6	24	30		14	32	56		18	110	234				
	7	28	35		15	67	120		19	114	247				
	8	30	40		16	10	16		20	120	260				
	9	35	45		17	77	136								
	10	8	10		18	40	72	14	15	98	210				
	11	39	55		19	82	152		16	53	112				
	12	43	60		20	22	40		17	111	238				
	13	45	65						18	58	126				
	14	46	70	9	10	53	90		19	121	266				
	15	11	15		11	59	99		20	63	140				
	16	54	80		12	21	36								
	17	55	85		13	65	117	15	16	114	240				
	18	60	90		14	70	126		17	116	255				
	19	61	95		15	25	45		18	41	90				
	20	13	20		16	78	144		19	127	285				
					17	82	153		20	27	60				
					18	10	18								
					19	89	171								
					20	93	180								

VII KRITICKÉ HODNOTY PEARSONOVA A SPEARMANOVA KOEFICIENTU KORELACE PRO POČET DVOJIC HODNOT $n \leq 8$

Koeficient korelace je statisticky významný, jestliže /r/ ≥ tabelované hodnotě.

Počet dvojic hodnot	**Pearsonův koeficient**		**Spearmanův koeficient**	
	hladina významnosti		**hladina významnosti**	
	0,050	**0,010**	**0,050**	**0,010**
3	0,997	1,000	–	–
4	0,950	0,990	1,000	–
5	0,878	0,959	0,900	1,000
6	0,811	0,917	0,829	0,943
7	0,755	0,875	0,745	0,893
8	0,707	0,834	0,691	0,833

VIII POMOCNÉ HODNOTY PRO DVOJŘÁDKOVOU KORELACI

p			$\sqrt{pq}$	*pq* / *y*	*p*			$\sqrt{pq}$	*pq* / *y*
0,01	nebo	0,99	0,099	0,372	0,26	nebo	0,74	0,439	0,593
0,02		0,98	0,140	0,405	0,27		0,73	0,444	0,596
0,03		0,97	0,171	0,428	0,28		0,72	0,449	0,599
0,04		0,96	0,196	0,446	0,29		0,71	0,454	0,602
0,05		0,95	0,218	0,461	0,30		0,70	0,458	0,604
0,06		0,94	0,237	0,474	0,31		0,69	0,462	0,606
0,07		0,93	0,255	0,485	0,32		0,68	0,465	0,609
0,08		0,92	0,271	0,495	0,33		0,67	0,470	0,611
0,09		0,91	0,286	0,504	0,34		0,66	0,474	0,612
0,10		0,90	0,300	0,513	0,35		0,65	0,477	0,614
0,11		0,89	0,313	0,521	0,36		0,64	0,480	0,616
0,12		0,88	0,325	0,528	0,37		0,63	0,483	0,617
0,13		0,87	0,336	0,535	0,38		0,62	0,485	0,619
0,14		0,86	0,347	0,541	0,39		0,61	0,488	0,620
0,15		0,85	0,357	0,547	0,40		0,60	0,490	0,621
0,16		0,84	0,367	0,552	0,41		0,59	0,492	0,622
0,17		0,83	0,376	0,558	0,42		0,58	0,494	0,623
0,18		0,82	0,384	0,563	0,43		0,57	0,495	0,624
0,19		0,81	0,392	0,567	0,44		0,56	0,496	0,625
0,20		0,80	0,400	0,572	0,45		0,55	0,497	0,625
0,21		0,79	0,407	0,576	0,46		0,54	0,498	0,626
0,22		0,78	0,414	0,580	047		0,53	0,499	0,626
0,23		0,77	0,421	0,583	0,48		0,52	0,500	0,626
0,24		0,76	0,427	0,587	0,49		0,51	0,500	0,627
0,25		0,75	0,433	0,590	0,50		0,50	0,500	0,627

IX FISHEROVA z-TRANSFORMACE

$$z = \frac{1}{2} \ln \frac{1+r}{1-r}$$

Pro záporné hodnoty korelačního koeficientu se hodnoty z nalezené v tabulce opatří znaménkem minus.

Pro $r \leq 0{,}2$ platí přibližně $z = r$.

r	**0,01**	**0,01**	**0,02**	**0,03**	**0,04**	**0,05**	**0,06**	**0,07**	**0,08**	**0,09**
0,0	0,0000	0,0100	0,0200	0,0300	0,0400	0,0500	0,0601	0,0701	0,0802	0,0902
0,1	0,1003	0,1104	0,1206	0,1307	0,1409	0,1511	0,1614	0,1717	0,1820	0,1923
0,2	0,2027	0,2132	0,2237	0,2342	0,2448	0,2554	0,2661	0,2769	0,2877	0,2986
0,3	0,3095	0,3205	0,3316	0,3428	0,3541	0,3654	0,3769	0,3884	0,4001	0,4118
0,4	0,4236	0,4356	0,4477	0,4599	0,4722	0,4847	0,4973	0,5101	0,5230	0,5361
0,5	0,5493	0,5627	0,5763	0,5901	0,6042	0,6184	0,6328	0,6475	0,6625	0,6777
0,6	0,6931	0,7089	0,7250	0,7414	0,7582	0,7753	0,7928	0,8107	0,8291	0,8480
0,7	0,8673	0,8872	0,9076	0,9287	0,9505	0,9730	0,9962	1,0203	1,0454	1,0714
0,8	1,0986	1,1270	1,1568	1,1881	1,2212	1,2562	1,2933	1,3331	1,3758	1,4219
0,9	1,4722	1,5275	1,5890	1,6584	1,7380	1,8318	1,9459	2,0923	2,2976	2,6466
0,99	2,6466	2,6996	2,7587	2,8257	2,9031	2,9945	3,1063	3,2504	3,4534	3,8002

X KRITICKÉ HODNOTY TESTOVÉHO KRITÉRIA *t*

(oboustranný test)

Stupně volnosti	Hladina významnosti		Stupně volnosti	Hladina významnosti	
	0,05	0,01		0,05	0,01
1	12,706	63,657	26	2,056	2,779
2	4,303	9,925	27	2,052	2,771
3	3,182	5,841	28	2,048	2,763
4	2,776	4,604	29	2,045	2,756
5	2,571	4,032	30	2,042	2,750
6	2,447	3,707	35	2,030	2,724
7	2,365	3,499	40	2,021	2,705
8	2,306	3,355	45	2,014	2,690
9	2,262	3,250	50	2,009	2,678
10	2,228	3,169	55	2,004	2,668
11	2,201	3,106	60	2,000	2,660
12	2,179	3,055	70	1,994	2,648
13	2,160	3,012	80	1,990	2,639
14	2,145	2,977	90	1,987	2,632
15	2,131	2,947	100	1,984	2,626
16	2,120	2,921	140	1,977	2,611
17	2,110	2,898	200	1,972	2,601
18	2,101	2,878	400	1,966	2,588
19	2,093	2,861	1000	1,962	2,581
20	2,086	2,845		1,960	2,576
21	2,080	2,831			
22	2,074	2,819			
23	2,069	2,807			
24	2,064	2,797			
25	2,060	2,787			

XI KRITICKÉ HODNOTY FISHEROVA-SNEDECOROVA *F* pro hladinu významnosti 0,05

(oboustranný test)

	f_1 (větší rozptyl)								
	1	2	3	4	5	6	7	8	9
1	161,45	199,50	215,71	224,58	230,16	233,99	236,77	238,88	240,54
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,36	19,39
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00
5	6,61	5,79	5,41	5,19	5,04	4,95	4,88	4,82	4,77
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10
7	5,59	5,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18
10	4,97	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21
40	4,09	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04
120	3,92	3,07	2,68	2,45	2,29	2,18	2,09	2,02	1,96
∞	3,84	3,00	2,61	2,37	2,21	2,10	2,01	1,94	1,88

	f_1 (větší rozptyl)							
	10	15	20	30	40	60	120	∞
1	241,88	245,95	248,01	250,09	251,14	252,20	253,25	254,32
2	19,40	19,43	19,45	19,46	19,47	19,48	19,49	19,50
3	8,79	8,70	8,66	8,62	8,59	8,57	8,55	8,53
4	5,96	5,86	5,80	5,75	5,72	5,69	5,66	5,63
5	4,74	4,62	4,56	4,50	4,46	4,43	4,40	4,37
6	4,06	3,94	3,87	3,81	3,77	3,74	3,71	3,67
7	3,64	3,51	3,45	3,38	3,34	3,30	3,27	3,23
8	3,35	3,22	3,15	3,08	3,04	3,01	2,97	2,93
9	3,14	3,01	2,94	2,86	2,83	2,79	2,75	2,71
10	2,98	2,85	2,77	2,70	2,66	2,62	2,58	2,54
15	2,54	2,40	2,33	2,25	2,20	2,16	2,11	2,07
20	2,35	2,20	2,12	2,04	1,99	1,95	1,90	1,84
30	2,17	2,02	1,93	1,84	1,79	1,74	1,68	1,62
40	2,08	1,93	1,84	1,74	1,69	1,64	1,58	1,51
60	1,99	1,84	1,75	1,65	1,59	1,53	1,47	1,39
120	1,91	1,75	1,66	1,55	1,50	1,43	1,35	1,25
∞	1,83	1,67	1,57	1,46	1,39	1,32	1,22	1,00

XII HODNOTY R_α PRO DUNCANŮV TEST pro hladinu významnosti 0,05

f	p								
	2	3	4	5	6	7	8	9	10
1	17,970	17,970	17,970	17,970	17,970	17,970	17,970	17,970	17,970
2	6,085	6,085	6,085	6,085	6,085	6,085	6,085	6,085	6,085
3	4,501	4,516	4,516	4,516	4,516	4,516	4,516	4,516	4,516
4	3,927	4,013	4,033	4,033	4,033	4,033	4,033	4,033	4,033
5	3,635	3,749	3,797	3,814	3,814	3,814	3,814	3,814	3,814
6	3,461	3,587	3,649	3,680	3,694	3,697	3,697	3,697	3,697
7	3,344	3,477	3,548	3,588	3,611	3,622	3,626	3,626	3,626
8	3,261	3,399	3,475	3,521	3,549	3,566	3,575	3,579	3,579
9	3,199	3,339	3,420	3,470	3,502	3,523	3,536	3,544	3,547
10	3,151	3,293	3,376	3,430	3,465	3,489	3,505	3,516	3,522
11	3,113	3,256	3,342	3,397	3,435	3,462	3,480	3,493	3,501
12	3,082	3,225	3,313	3,370	3,410	3,439	3,459	3,474	3,484
13	3,055	3,200	3,289	3,348	3,389	3,419	3,442	3,458	3,470
14	3,033	3,178	3,268	3,329	3,372	3,403	3,426	3,444	3,457
15	3,014	3,160	3,250	3,312	3,356	3,389	3,413	3,432	3,446
16	2,998	3,144	3,235	3,298	3,343	3,376	3,402	3,422	3,437
17	2,984	3,130	3,222	3,285	3,331	3,366	3,392	3,412	3,429
18	2,971	3,118	3,210	3,274	3,321	3,356	3,383	3,405	3,421
19	2,960	3,107	3,199	3,264	3,311	3,347	3,375	3,397	3,415
20	2,950	3,097	3,190	3,255	3,303	3,339	3,368	3,391	3,409
24	2,919	3,066	3,160	3,226	3,276	3,315	3,345	3,370	3,390
30	2,888	3,035	3,131	3,199	3,250	3,290	3,322	3,349	3,371
40	2,858	3,006	3,102	3,171	3,224	3,266	3,300	3,328	3,352
60	2,829	2,976	3,073	3,143	3,198	3,241	3,277	3,307	3,333
120	2,800	2,947	3,045	3,116	3,172	3,271	3,254	3,287	3,314
∞	2,772	2,918	3,017	3,089	3,146	3,193	3,232	3,265	3,294

LITERATURA

ANDERSON, G. a N. ARSENAULT. *Fundamentals of Educational Research*. London: Falmer Press, 1998.

BLAHUŠ, P. *Faktorová analýza a její zobecnění*. Praha: SNTL, 1985.

BLAHUŠ, P. *Základní metody faktorové analýzy v antropometrice*. Praha: Univerzita Karlova, 1971.

BREZINKA, W. K problému vymezení vědy o výchově. *Pedagogika.* 1967, 17 (2), 160–170.

BRICHÁČEK, V. *Úvod do psychologického škálování*. Bratislava: Psychodiagnostické a didaktické testy, 1978.

BROWN, A. a P. DOWLING. *Doing Research / Reading Research.* A mode of Interrogation for Education. London: Falmer Press, 1998.

BYČKOVSKÝ, P. *Didaktické testy a měření výsledků výuky : základní pojmy*. Praha: ČVUT, 1980.

BYČKOVSKÝ, P. Dvě netradiční metody analýzy dat a jejich využití v pedagogickém výzkumu. *Pedagogika.* 1992, 42 (2), 237–249.

BYČKOVSKÝ, P. *Základy měření výsledků výuky*. Praha: Výzkumný ústav inženýrského studia, 1982.

CRESWELL, J. W. *Qualitative inquiry & research design: choosing among five approaches*. 2nd ed. Thousand Oaks: Sage Publications, 2007.

CYHELSKÝ, L. *Úvod do teorie popisné statistiky*. Praha: SNTL, 1974.

DISMAN, M. *Jak se vyrábí sociologická znalost*. Praha: Karolinum,1993.

DITRICH, P. *Pedagogicko-psychologická diagnostika*. Jinočany: Nakladatelství H & H, 1993.

FIELD, A. P. *Discovering statistics using IBM SPSS statistics*. 4th ed. London: SAGE, 2013.

GAJDA, V. a ZVOLSKÁ, J. *Úvod do statistických metod*. Ostrava: Pedagogická fakulta v Ostravě, 1982.

GAVORA, P. *Akí sú moji žiaci? Pedagogická diagnostika žiaka*. Bratislava: Práca, 1999.

GAVORA, P. Kvalitatívny výskum v pedagogike. *Výchova a vzdělávání*. 1990/1991, 1 (8), 170–172.

GAVORA, P. *Tvorba výskumného nástroja pre pedagogické bádanie*. Bratislava: SPN, 2012.

GAVORA, P. *Úvod do pedagogického výzkumu*. Brno: Paido, 2000.

GAVORA, P. *Výzkumné metody v pedagogice.* Brno: Paido, 1996.

GNITECKI, J. *Zarys metodologii badan w pedagogice empirycznej*. Zielona Góra: Wyzsza szkola pedagogiczna, 1993.

HELUS, Z. aj. *Psychologie školní úspěšnosti žáků*. Praha: Státní pedagogické nakladatelství, 1979.

HENDL, J. *Přehled statistických metod zpracování dat. Analýza a metaanalýza dat*. Praha: Portál, 2004.

HENDL, J. *Úvod do kvalitativního výzkumu*. Praha: Karolinum, 1999.

HENDL, J. *Kvalitativní výzkum: základní metody a aplikace*. Praha: Portál, 2005.

HILLEBRANDT, F. *Elementárna štatistika pre psychológov, sociológov a pedagógov*. Bratislava: Slovenské pedagogické nakladatelství, 1968.

HNILIČKOVÁ, J. et al. *Didaktické testy a jejich statistické zpracování*. Praha: Státní pedagogické nakladatelství, 1972.

HORÁK, F. a M.CHRÁSKA. *Metodologie pedagogiky*. Olomouc: Pedagogická fakulta UP, 1986.

HORÁK, F. a M. CHRÁSKA. *Úvod do metodologie pedagogického výzkumu*. Olomouc: Univerzita Palackého, 1989.

HRABAL, V. *Jaký jsem učitel?* Praha: Státní pedagogické nakladatelství, 1988.

HRABAL, V. *Pedagogicko-psychologická diagnostika žáka*. Praha: Státní pedagogické nakladatelství, 1989.

CHRÁSKA, M. a I. KOČVAROVÁ. *Kvantitativní design v pedagogických výzkumech začínajících akademických pracovníků*. 1. vyd. Zlín: Univerzita Tomáše Bati, 2014.

CHRÁSKA, M. a I. KOČVAROVÁ. *Kvantitativní metody sběru dat v pedagogických výzkumech.* 1. vyd. Zlín: Univerzita Tomáše Bati, 2015.

CHRÁSKA, M. a V. JANÁK. *Statistika pro pedagogy*. Olomouc: Pedagogická fakulta UP, 1990.

CHRÁSKA, M. a V. KLAPAL. Typologie učitelů základní školy podle preferencí učitelských dovedností. In: *Profese učitele a současná společnost.* [CD-ROM]. Brno: Pedagogická fakulta MU, 2003.

CHRÁSKA, M. *Didaktické testy*. Brno: Paido, 1999.

CHRÁSKA, M. *Empirická pedagogická šetření a jejich statistické vyhodnocování*. Olomouc: Pedagogická fakulta UP, 1988.

CHRÁSKA, M. *Hypotézy a jejich ověřování v klasických pedagogických výzkumech*. Olomouc: Votobia, Pedagogická fakulta UP, 2005.

CHRÁSKA, M. Jaké jsou postoje žáků a studentů ke škole a edukační realitě? *Pedagogika*. 1998, 48 (1), 54–66.

CHRÁSKA, M. *K současným trendům pedagogického výzkumu ve světě*. Olomouc: Pedagogická fakulta UP, 1995.

CHRÁSKA, M. Které vědomosti a dovednosti z pedagogiky považují učitelé za důležité? *Pedagogika*. 1996, 46 (3), 256–265.

CHRÁSKA, M. Měření postojů k učitelské profesi pomocí sémantického diferenciálu. In: Acta *Fakultatis Paedagogicae Universitatis Presoviensis*, Annus III, Prešov – Olomouc. Prešov: Pedagogická fakulta Prešovskej univerzity, 2004. s. 25–33.

CHRÁSKA, M. *Metodologie řešení vybraných problémů v pedagogickém výzkumu*. Olomouc: Pedagogická fakulta UP, 1992.

CHRÁSKA, M. *Metody pedagogické diagnostiky*. Olomouc: Pedagogická fakulta UP, 1988.

CHRÁSKA, M. *Metody sběru a statistického vyhodnocování dat v evaluačních pedagogických výzkumech*. Olomouc: Pedagogická fakulta, Votobia, 2003.

CHRÁSKA, M. Odhady reliability měření sémantickým diferenciálem. In: *Sborník příspěvků z 6. konference České asociace pedagogického výzku*mu. České Budějovice: Pedagogická fakulta Jihočeské univerzity, 1998, s. 46–53.

CHRÁSKA, M. Spolehlivost a přesnost měření v evaluačních pedagogických výzkumech. In: *Sborník referátů ze 4. konference České asociace pedagogického výzkumu.* Olomouc: Pedagogická fakulta UP, 1996, s. 27–33.

CHRÁSKA, M. *Úvod do výzkumu v pedagogice. Základy kvantitativně orientovaného výzkumu.* Olomouc: Vydavatelství Univerzity Palackého, 2003.

CHRÁSKA, M. *Základy výzkumu v pedagogice*. Olomouc: Pedagogická fakulta UP, 1998.

CHRÁSKA, M. Zkoumání individuální koncepce výuky učitelů pomocí Q-metodologie. In: *Sborník AUPO.* Olomouc: Pedagogická fakulta UP, 1992, s. 122–145.

CHRÁSKA, M. Změny v sémantickém prostoru studentů pedagogické fakulty. *Pedagogika*. 1995, 45 (1), 64–76.

JEDLIČKA, J. *Úvod do statistických metod v pedagogice*. Ostrava: Pedagogická fakulta, 1976.

KAHUDA, F. *Výzkumné metody v sociologii*. Praha: SPN, 1965.

KAPR, J. a Z. ŠAFÁŘ. *Sociologie nebo zdravý rozum?* Praha: Mladá fronta, 1969.

KERLINGER, F. N. *Základy výzkumu chování*. Praha: Academia, 1972.

KOMENDA, S. Pedagogika a edukometrie. *Pedagogika*. 1993, 43 (4), 391–404.

KOMENDA, S. *Politometrie.* Olomouc: Vydavatelství Univerzity Palackého, 1995.

KOMENDA, S. *Základní biometrické metody*. Olomouc: Univerzita Palackého, 1968.

KOMENDA, S. *Základy pravděpodobnostních a statistických metod v psychologic-kém a pedagogickém výzkumu.* Praha: Státní pedagogické nakladatelství, 1967.

KOMENDA, S. a J. KLEMENTA. *Analýza náhodného v pedagogickém experimentu a praxi.* Praha: SPN 1981.

KOSCHIN, F. et al. *Statgraphics aneb statistika pro každého.* Praha: Grada, 1992.

KOVÁŘ, R. a P. BLAHUŠ. *Aplikace vybraných statistických metod v antropomotorice.* Praha: Univerzita Karlova, 1989.

KRÁLÍK, O. a J. HARTMANN. *Základy statistiky pro pedagogy*. Brno: Akademické nakladatelství Cerm, 2000.

LAMSER, V. a L. RŮŽIČKA. *Základy statistiky pro sociology*. Praha: Svoboda, 1970.

LAPITKA, M. Meranie a škálovanie v edukačnom výskume. In: ŠVEC, Š. et al. *Metodológia vied o výchove: Kvantitatívno-scientistické a kvalitatívno-humanistické prístupy.* Bratislava: IRIS, 1998, s. 85–93.

LIKEŠ, J. a J. LAGA. *Základní statistické tabulky*. Praha: SNTL, 1978.

LINDQUIST, E. F. *Statistická analýza v pedagogickém výzkumu*. Praha: SPN, 1967.

LINHART, J. a D. HOLDA. Faktorová analýza. *Sociologický časopis*. 1970, 6(6), 575–592.

LOUČKOVÁ, Ivana. *Integrovaný přístup v sociálně vědním výzkumu*. Praha: Sociologické nakladatelství, 2010.

MAŇÁK, Josef et al. *Kapitoly z metodologie pedagogiky*. Brno: Masarykova univerzita, 1996.

MAREŠ, J. Jak zjišťovat reliabilitu pozorování? *Pedagogika*. 1983, č. 2.

MAREŠ, J. Komunikace učitel – žáci v hodinách fyziky. In: BEDNAŘÍK, M. et al. *Sborník referátů a sdělení přednesených na konferenci Aplikace kybernetické pedagogiky ve vyučování fyzice*. Praha: MFF UK, 1977, s. 13-31.

MAREŠ, P., RABUŠIC, L. a P. SOUKUP. *Analýza sociálněvědních dat (nejen) v SPSS.* 1. vyd. Brno: Masarykova univerzita, 2015.

McKENZIE, G. et al. *Understanding Social esearch: Perspectives on Metodology and Practice*. London: Falmer Press, 1997.

MEŠKO, D., KATUŠČÁK, D. et al. *Akademická príručka*. Martin: Nakladateľstvo Osveta, 2004.

MICHALIČKA, M. Pedagogické testy a problémy jejich použití v pedagogické praxi. *Pedagogika*. 1969, č. 1.

MITTENECKER, E. *Plánování a statistické hodnocení experimentů*. Praha: Státní pedagogické nakladatelství, 1968.

MRÁZ, V. *Analýza a standardizace testů pedagogické a psychologické diagnostiky*. Praha: Pedagogická fakulta UK, 1977.

MUŽIČ, V. *Testy vědomostí*. Praha: Státní pedagogické nakladatelství, 1971.

NIČKOVIČ, R. *Metodológia pedagogického výskumu*. Bratislava: Státní pedagogické nakladatelství, 1968.

NOWAK, S. *Metody badán socjologicznych*. Warszawa: 1965.

ONDREJKOVIČ, Peter. *Úvod do metodologie sociálných vied: Základy metodológie kvantitativného výskumu.* Nitra: Regent, 2005.

OSGOOD, C. et al. *The Measurement of Meaning*. Urbana, III, University of Illinois Press, 1957.

PAPICA, J. *Metody sociálně psychologického výzkumu*. Praha: Univerzita Karlova, 1974.

PAPICA, J. *Základy psychometrie*. Olomouc: Filozofická fakulta UP, 1984.

PEERS, I. *Statistical Analysis for Education & Psychology Researchers*. London: Falmer Press, 1996.

PELIKÁN, J. *Metodologie výzkumu osobnosti středoškolského profesora a jeho pedagogického působení*. Praha: Univerzita Karlova, 1991.

PELIKÁN, J. *Základy empirického výzkumu pedagogických jevů*. Praha: Karolinum, 1998.

PETRUSEK, M. *Sociometrie*. Praha: Svoboda, 1969.

PIETER, J. *Zarys metodologii pracy naukowej*. Warszawa: PWN, 1975.

PICHOT, P. *Mentální testy*. Praha: Státní pedagogické nakladatelství, 1970.

POPPER, K. R. *Logika vědeckého výzkumu.* Praha: Oikúmené, 1997.

POPPER, K. R. *Objective Knowledge*. Oxford: 1972.

POSPÍŠIL, L. *Empirická kritika marxistických teorií práva a řízení společnosti*. Olomouc: Filozofická fakulta Univerzity Palackého, 1990.

PRŮCHA, J. *Moderní pedagogika.Věda o edukačních procesech*. Praha: Portál, 1997.

PRŮCHA, J. *Pedagogická evaluace*. Brno: Centrum pro další vzdělávání učitelů, Masarykova univerzita, 1996.

PRŮCHA, J. *Pedagogické teorie a výzkumy na Západě.* Praha: Univerzita Karlova, 1992.

PRŮCHA, J. *Pedagogický výzkum: Uvedení do teorie a praxe*. Praha: Karolinum, 1995.

PRŮCHA, J. *Učebnice: Teorie a analýzy edukačního média*. Brno: Paido, 1998.

PRŮCHA, J. *Učitelství: Obraz profese*. Praha: Portál, 2002.

PRŮCHA, J. et al. *Pedagogická encyklopedie*. Praha: Grada, 2009.

PRŮCHA, J., WALTEROVÁ, E. a J. MAREŠ, *Pedagogický slovník*. Praha: Portál, 1991.

PŘÍHODA, V. *Praxe školského měření*. Praha: Dědictví Komenského, 1936.

PŘÍHODA, V. *Teorie školského měření*. Praha: Bakulův ústav, 1932.

Příručka pro sociology. Praha: Svoboda, 1980.

REISENAUER, R. *Metody matematické statistiky a jejich aplikace*. Praha: SNTL, 1970.

REITEROVÁ, E. *Základy statistiky pro studenty psychologie*. Olomouc: Univerzita Palackého, 2000.

RYS, S. *Hospitace v pedagogické praxi*. Praha: Státní pedagogické nakladatelství, 1975.

ŘEHÁK, J. a B. ŘEHÁKOVÁ. *Analýza kategorizovaných dat v sociologii*. Praha: Academia, 1986.

ŘÍČAN, P. *Úvod do psychometrie*. Bratislava: Psychodiagnostické a didaktické testy, 1978.

SCHMIDT, H. D. *Posudzovanie ľudského správania rating-škálami*. Bratislava: Slovenské pedagogické nakladatelstvo, 1970.

SCHNEIDER, M. *Úvod do základů sociologického výzkumu*. Olomouc: Univerzita Palackého, 1980.

SINGULE, F. *Současné pedagogické směry a jejich psychologické souvislosti*. Praha: Státní pedagogické nakladatelství, 1992.

SKALKOVÁ, J. et al. *Úvod do metodologie a metod pedagogického výzkumu*. Praha: Státní pedagogické nakladatelství, 1983.

SMÉKAL, V. *Poznávání a posuzování osobností žáků*. Praha: Státní pedagogické nakladatelství, 1970.

STRAUSS, A. aj. CORBIN. *Základy kvalitativního výzkumu: postupy a techniky metody zakotvené teorie*. Boskovice: Albert, 1999.

SWOBODA, H. *Moderní statistika*. Praha: Svoboda, 1977.

ŠVAŘÍČEK, R a K. ŠEĎOVÁ. *Kvalitativní výzkum v pedagogických vědách*. Praha: Portál, 2007.

ŠVEC, Š. et al. *Metodológia vied o výchove: Kvantitativno - scientické a kvalitativno - humanitné prístupy v edukačnom výskume*. Bratislava: IRIS, 1998.

ŠVEC, V. et al. *Slovník pedagogické metodologie*. Brno: Paido 2005.

TOLLINGEROVÁ, D. Bellackova metoda mikroanalýzy a její formální zápis. *Psychológia a patopsychológia dieťaťa*. 1971, č. 3.

TRAVERS, R. M. W. *Úvod do pedagogického výzkumu*. Praha: SPN, 1969.

VALA, J. a M. CHRÁSKA. Recepce poezie u žáků středního odborného učiliště a gymnázia. *Studia Paedagogica*. 2014, 19 (1), 43–63.

VOGT, W. P. a B. JOHNSON. *Dictionary of statistics & methodology: a nontechnical guide for the social sciences*. 4th ed. Thousand Oaks, Calif.: SAGE, 2011.

ZAPLETAL, I. *K obecné metodologii pedagogických věd*. Praha: Státní pedagogické nakladatelství, 1977.

ZICH, F. *Sociologický výzkum*. Praha: Svoboda, 1976.

REJSTŘÍK

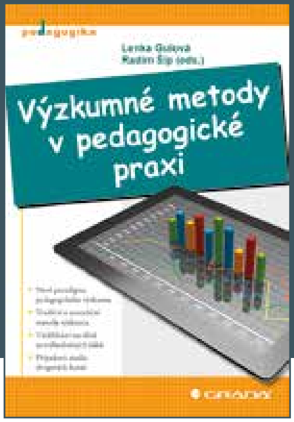
PEDAGOGIKA

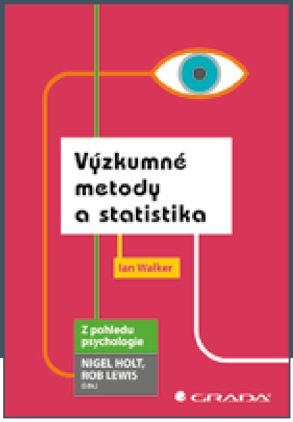
Aktualizovaná publikace renomovaného českého autora objasňuje podstatu klasického (kvantitativního) pedagogického výzkumu. Analyzuje jeho základní fáze, informační zdroje a také jeho výhody i nevýhody. Kniha je určena zejména studentům pedagogických oborů a pedagogům.

Publikace rozvíjí zejména následující témata:

- otázky měření v pedagogickém výzkumu včetně základních metod, s jejichž pomocí se zpracovávají výsledky měření
- statistické metody používané při testování hypotéz
- nejčastěji používané metody sběru dat

Zaměřuje se nejen na metody, které jsou v pedagogickém výzkumu běžně používané, ale také na ty méně časté (sémantický diferenciál, Q-metodologie, faktorová analýza, shluková analýza). Velkou péči autor věnoval srozumitelnosti textu, aby byl přístupný i těm, kdo mají jen základní matematické znalosti a dovednosti.





GRADA

Grada Publishing, a.s.
U Průhonu 22, 170 00 Praha 7
tel.: +420 234 264 401, fax: +420 234 264 400
e-mail: obchod@grada.cz

www.grada.cz

ISBN 978-80-247-5326-3