

K2SZD – Statistika a statistické zpracování dat

(Podklady k teoretické části zkoušky)

1 Výběrové zjišťování

1. Úvod do tématu

- **Definice výběrového zjišťování:** Metoda statistického sběru dat, kdy je zkoumán pouze výběrový soubor, nikoli celá populace.
 - **Význam výběrového zjišťování:** Šetří čas, náklady a umožňuje provádět analýzy i u rozsáhlých souborů dat.
 - **Hlavní úkoly výběrového zjišťování:** Odhad parametrů populace a kontrola spolehlivosti výsledků.
-

2. Základní pojmy výběrového zjišťování

- **Náhodný výběr:** Každý prvek populace má stejnou pravděpodobnost být zahrnut do výběru.
 - **Výběrový soubor:** Podmnožina základního souboru, vybraná pro analýzu.
 - **Statistický model:** Pravděpodobnostní rozdělení náhodné veličiny ve výběrovém souboru.
 - **Výběrový průměr:** Průměr hodnot ve výběru, odhad střední hodnoty populace.
 - **Výběrový rozptyl:** Míra variability hodnot ve výběrovém souboru.
 - **Výběrová směrodatná odchylka:** Odmocnina z výběrového rozptylu, měří průměrnou odchylku hodnot od výběrového průměru.
 - **Výběrová kovariance:** Míra vzájemné závislosti dvou znaků ve výběrovém souboru.
 - **Výběrový lineární korelační koeficient:** Míra síly a směru lineárního vztahu mezi dvěma znaky ve výběru.
 - **Výběrový poměr:** Poměr rozsahu výběru k velikosti populace.
 - **Výběrová chyba:** Rozdíl mezi výsledkem výběrového šetření a skutečnou hodnotou v populaci.
-

3. Metody výběru vzorků

- **Prostý náhodný výběr:** Každý prvek má stejnou šanci být vybrán. Používá se generátor náhodných čísel.
 - **Stratifikovaný výběr:** Populace je rozdělena do homogenních podskupin (strat) a z každé je proveden výběr.
 - **Systematický výběr:** Výběr každého k-tého prvku z uspořádaného seznamu.
 - **Vícestupňový shlukový výběr:** Výběr prováděný ve více stupních (např. okres → město → domácnost).
-

4. Vlastnosti odhadů

- **Nevychýlený odhad:** Průměr hodnot odhadu se rovná skutečné hodnotě parametru populace.
 - **Konzistentní odhad:** S rostoucím rozsahem výběru se hodnota odhadu blíží skutečné hodnotě parametru.
 - **Reprezentativnost výběru:** Schopnost výběru přesně odrážet vlastnosti základního souboru.
-

5. Výběrové charakteristiky a rozdělení

- **Výběrový obecný moment:** Funkce náhodných veličin ve výběrovém souboru.
 - **Centrální limitní věta:** Rozdělení průměrů výběrových souborů se blíží normálnímu rozdělení.
 - **Standardní chyba průměru:** Měří, jak moc se průměr výběru liší od skutečného průměru populace.
-

6. Praktické příklady výběrového zjišťování

- **Průzkum spokojenosti zákazníků:** Výběrový soubor zákazníků reprezentuje celkovou populaci.
 - **Politické průzkumy:** Stratifikovaný výběr podle regionů a demografických charakteristik.
 - **Kontrola kvality ve výrobě:** Systematický výběr výrobků z výrobní linky.
-

7. Kontrolní otázky pro studenta:

1. Co je pravděpodobnostní výběr a proč je důležitý pro statistickou analýzu?

- **Odpověď:** Zajišťuje náhodnost a reprezentativnost výběru, což umožňuje zobecnění výsledků na celou populaci.
2. Jaký je rozdíl mezi prostým náhodným výběrem a stratifikovaným výběrem?
- **Odpověď:** Prostý výběr zahrnuje náhodné prvky, stratifikovaný zajišťuje reprezentaci každé podskupiny.
3. Popište postup při prostém náhodném výběru.
- **Odpověď:** Očíslování prvků, použití generátoru náhodných čísel, výběr prvků, analýza dat.
4. Co je výběrový poměr a jaký je jeho význam?
- **Odpověď:** Poměr počtu prvků ve výběru k celkové populaci, ovlivňuje přesnost odhadů.
5. Jak se vypočítá výběrový průměr?
- **Odpověď:** Součet hodnot ve výběru dělený jejich počtem.
6. Co je výběrový rozptyl a směrodatná odchylka?
- **Odpověď:** Měří variabilitu hodnot ve výběru, směrodatná odchylka je odmocnina rozptylu.
7. Co je výběrová chyba a jak ji lze minimalizovat?
- **Odpověď:** Rozdíl mezi hodnotou z výběru a skutečnou hodnotou populace; zvyšováním rozsahu výběru.
8. Vysvětlete rozdíl mezi odhady konzistentními a nevychýlenými.
- **Odpověď:** Nevychýlený odhad není zatížen systematickou chybou, konzistentní se zpřesňuje s rostoucím vzorkem.
9. Jaký je praktický význam stratifikovaného výběru?
- **Odpověď:** Zajišťuje lepší reprezentaci podskupin populace.
10. Jaké alternativní metody výběru mohou být použity místo prostého náhodného výběru?
- **Odpověď:** Stratifikovaný výběr, systematický výběr, víceúrovňový shlukový výběr.
-

2 Pravděpodobnost

1. Úvod do tématu

- **Definice pravděpodobnosti:** Pravděpodobnost je míra, s jakou lze očekávat, že nastane určitý jev, na základě předem známých podmínek.
 - **Význam pravděpodobnosti:** Používá se v ekonomii, řízení rizik, pojišťovnictví a analýze dat k odhadu nejistých událostí.
 - **Historický vývoj:** Pravděpodobnost vznikla jako nástroj pro analýzu hazardních her a později se rozšířila do vědy a statistiky.
-

2. Základní pojmy pravděpodobnosti

- **Náhodný jev:** Událost, jejíž výsledek není předem jistý (např. hod kostkou).
 - **Náhodná veličina:** Proměnná, jejíž hodnoty závisí na náhodném pokusu.
 - **Klasická pravděpodobnost:** Poměr příznivých výsledků k celkovému počtu možných výsledků.
 - **Geometrická pravděpodobnost:** Pravděpodobnost založená na poměrech geometrických objektů (např. délky, plochy).
 - **Statistická pravděpodobnost:** Pravděpodobnost určená na základě relativní četnosti jevu v dlouhodobém pokusu.
 - **Podmíněná pravděpodobnost:** Pravděpodobnost, že nastane jev A za podmínky, že nastal jev B.
 - **Úplná pravděpodobnost:** Součet pravděpodobností všech možných cest k danému výsledku.
 - **Bayesova věta:** Umožňuje aktualizovat pravděpodobnost na základě nových informací.
 - **Nezávislé jevy:** Jevy, kde výskyt jednoho neovlivňuje pravděpodobnost druhého.
-

3. Typy pravděpodobnosti a výpočty

- **Klasická pravděpodobnost:** Příklady s kostkami, kartami, osudí.
 - **Statistická pravděpodobnost:** Relativní četnosti.
 - **Geometrická pravděpodobnost:** Vztah mezi dvěma geometrickými objekty (např. pravděpodobnost zasažení cíle na ploše).
-

4. Podmíněná pravděpodobnost a Bayesova věta

- **Podmíněná pravděpodobnost:** $P(A|B)=P(A \cap B)/P(B)$.
 - **Bayesova věta:** Umožňuje určit pravděpodobnost příčiny na základě pozorovaného výsledku.
 - **Praktický příklad:** Diagnostika nemoci – pravděpodobnost onemocnění po pozitivním testu.
-

5. Nezávislé a závislé jevy

- **Nezávislé jevy:** $P(A \cap B)=P(A) \cdot P(B)$.
 - **Závislé jevy:** Pravděpodobnost jednoho jevu je ovlivněna výskytem druhého.
 - **Příklad:** Pravděpodobnost dvou hodů kostkou je nezávislá, zatímco pravděpodobnost vytažení karty bez vrácení je závislá.
-

6. Bernoulliho schéma

- **Definice:** Model pro sérii nezávislých pokusů s dvěma možnými výsledky (úspěch/neúspěch).
 - **Binomické rozdělení:** Pravděpodobnost k úspěchů v n pokusech:
$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$
 - **Příklad:** Pravděpodobnost, že při 10 hodech mincí padne panna přesně 6krát.
-

7. Praktické příklady výpočtu pravděpodobnosti

- **Pravděpodobnost výhry v loterii:** Kombinatorické výpočty pro pravděpodobnost výhry.
 - **Pravděpodobnost chyby v účtování:** Pravděpodobnost výskytu chyby při opakovaných transakcích.
 - **Rizikové analýzy:** Aplikace pravděpodobnosti v rozhodování pod nejistotou.
-

8. Kontrolní otázky pro studenta:

1. Jaký je rozdíl mezi klasickou, geometrickou a statistickou pravděpodobností?
 - **Odpověď:** Klasická je založena na počítání možností, geometrická na poměrech a statistická na relativní četnosti.
2. Vysvětlete podmíněnou pravděpodobnost a její použití v praxi.

- **Odpověď:** Pravděpodobnost jevu za předpokladu, že jiný jev již nastal. Používá se v diagnostice a analýzách.
3. Co vyjadřuje Bayesova věta a kde se používá?
- **Odpověď:** Umožňuje aktualizovat pravděpodobnost na základě nových dat, např. v lékařských testech.
4. Jaký je rozdíl mezi závislými a nezávislými jevy?
- **Odpověď:** Nezávislé jevy se neovlivňují, zatímco u závislých výskyt jednoho jevu ovlivňuje pravděpodobnost druhého.
5. Jak se používá Bernoulliho schéma při výpočtu pravděpodobnosti?
- **Odpověď:** Používá se pro odhad počtu úspěchů v sérii nezávislých pokusů.
6. Co je princip úplné pravděpodobnosti?
- **Odpověď:** Pravděpodobnost jevu je součet pravděpodobností všech jeho možných cest.
7. Jak lze vypočítat pravděpodobnost pomocí kombinatoriky?
- **Odpověď:** Pomocí permutací, kombinací a variací.
8. Jaký je praktický význam pravděpodobnosti v oblasti ekonomie?
- **Odpověď:** Analýza rizik, predikce trendů a rozhodování za nejistoty.
9. Co je statistická pravděpodobnost a jak souvisí s relativní četností?
- **Odpověď:** Je založena na podílu výskytu jevu v dlouhodobém pokusu.
10. Jaké jsou základní předpoklady Bernoulliho schématu?
- **Odpověď:** Nezávislost pokusů, dva možné výsledky (úspěch/neúspěch), stálá pravděpodobnost úspěchu (a tedy i neúspěchu).
-

3 Náhodná veličina

1. Úvod do tématu

- **Definice náhodné veličiny:** (Reálná) proměnná, která nabývá hodnot na základě výsledku náhodného pokusu.
 - **Dělení náhodných veličin:** Diskrétní (konečný/spočetný počet hodnot) a spojité (celé intervaly).
 - **Praktické využití náhodných veličin:** Počet zákazníků v obchodě, doba čekání na autobus, výška člověka.
-

2. Diskrétní náhodná veličina

- **Definice:** Nabývá konečného nebo spočetně nekonečného množství hodnot.
 - **Pravděpodobnostní funkce:** Přiřazuje každé hodnotě náhodné veličiny pravděpodobnost.
 - **Distribuční funkce:** Funkce udávající pravděpodobnost, že náhodná veličina nabude hodnoty menší nebo rovné určité hodnotě.
 - **Střední hodnota:** Vážený průměr všech hodnot.
 - **Rozptyl a směrodatná odchylka:** Měří variabilitu hodnot kolem střední hodnoty.
-

3. Spojitá náhodná veličina

- **Definice:** Nabývá všech hodnot v daném intervalu.
 - **Hustota pravděpodobnosti:** Funkce popisující rozložení pravděpodobnosti hodnot.
 - **Distribuční funkce:** Integrál z hustoty pravděpodobnosti od počátečního bodu (obvykle mínus nekonečno) do bodu x .
 - **Střední hodnota a rozptyl:** Integrály vážené hustotou pravděpodobnosti.
-

4. Pravděpodobnostní rozdělení náhodné veličiny

- **Definice:** Popisuje, jak jsou pravděpodobnosti jednotlivých hodnot rozloženy.
 - **Příklady diskrétních rozdělení:** Binomické, Poissonovo.
 - **Příklady spojitých rozdělení:** Normální, exponenciální.
-

5. Číselné charakteristiky náhodných veličin

- **Střední hodnota (očekávaná hodnota):** Průměrná očekávaná hodnota náhodné veličiny.
 - **Rozptyl:** Míra variability hodnot kolem střední hodnoty.
 - **Směrodatná odchylka:** Druhá odmocnina rozptylu, vyjadřuje variabilitu ve stejných jednotkách jako náhodná veličina.
-

6. Praktické aplikace náhodných veličin

- **Ekonomické modely:** Analýza rizik a predikce vývoje trhů.
 - **Logistika:** Modelování doby dodání zboží.
 - **Výrobní procesy:** Počet vadných výrobků na výrobní lince.
-

7. Kontrolní otázky pro studenta:

1. Vysvětlete rozdíl mezi diskrétní a spojitou náhodnou veličinou.
 - **Odpověď:** Diskrétní má konečný/spočetný počet hodnot, spojitá interval hodnot.
2. Co je to distribuční funkce a jaké má vlastnosti?
 - **Odpověď:** Udává pravděpodobnost, že náhodná veličina nabude hodnoty menší nebo rovné určité hodnotě.
3. Jak se počítá střední hodnota diskrétní náhodné veličiny?
 - **Odpověď:** Jako vážený průměr všech hodnot náhodné veličiny.
4. Co je hustota pravděpodobnosti u spojitě náhodné veličiny?
 - **Odpověď:** Funkce popisující pravděpodobnostní rozložení hodnot.
5. Jaký je vztah mezi hustotou pravděpodobnosti a distribuční funkcí?
 - **Odpověď:** Distribuční funkce je integrál hustoty pravděpodobnosti.
6. Jaký je význam rozptylu a směrodatné odchylky?
 - **Odpověď:** Rozptyl měří variabilitu hodnot, směrodatná odchylka vyjadřuje tuto variabilitu ve stejných jednotkách jako veličina.
7. Jaký praktický příklad diskrétní náhodné veličiny můžete uvést?
 - **Odpověď:** Počet zákazníků navštěvujících obchod během jednoho dne.
8. Jaký praktický příklad spojitě náhodné veličiny můžete uvést?
 - **Odpověď:** Doba čekání na autobus.

9. Jaký je význam střední hodnoty náhodné veličiny?

- **Odpověď:** Udává očekávaný průměrný výsledek náhodného pokusu.

10. Jak se liší pravděpodobnostní funkce a hustota pravděpodobnosti?

- **Odpověď:** Pravděpodobnostní funkce se používá u diskrétních veličin, hustota pravděpodobnosti u spojitých veličin.
-

4 Rozdělení pravděpodobnosti

1. Úvod do tématu

- **Definice rozdělení pravděpodobnosti:** Matematický popis rozložení pravděpodobností možných hodnot náhodné veličiny.
 - **Typy rozdělení pravděpodobnosti:** Diskrétní a spojitá rozdělení.
 - **Praktický význam rozdělení:** Modelování reálných jevů, jako jsou počty vadných výrobků, doba čekání nebo výška osob.
-

2. Diskrétní rozdělení pravděpodobnosti

- **Binomické rozdělení:** Pravděpodobnost určitého počtu úspěchů v pevném počtu nezávislých pokusů.
 - **Hypergeometrické rozdělení:** Pravděpodobnost počtu úspěchů při výběru bez vracení.
 - **Poissonovo rozdělení:** Pravděpodobnost výskytu určitého počtu událostí v daném časovém intervalu.
 - **Distribuční funkce diskrétních rozdělení:** Udává pravděpodobnost, že veličina nabude hodnoty menší nebo rovné určité hodnotě.
-

3. Spojitá rozdělení pravděpodobnosti

- **Normální rozdělení:** Symetrické rozdělení s charakteristikou zvonovité křivky.
 - **Studentovo t-rozdělení:** Používá se při malých rozsazích vzorků.
 - **F-rozdělení:** Využívá se v analýze rozptylu.
 - **Chi-kvadrát rozdělení:** Používá se pro testy shody a nezávislosti.
 - **Distribuční funkce spojitých rozdělení:** Integrál z hustoty pravděpodobnosti.
-

4. Charakteristiky pravděpodobnostních rozdělení

- **Střední hodnota:** Očekávaná průměrná hodnota náhodné veličiny.
 - **Rozptyl:** Měří míru variability hodnot kolem střední hodnoty.
 - **Směrodatná odchylka:** Druhá odmocnina rozptylu.
 - **Kvantily spojitých rozdělení:** Hodnota, která rozděluje pravděpodobnost na určité části.
-

5. Využití Excelu pro pravděpodobnostní rozdělení

- **Funkce pro diskrétní rozdělení:** BINOM.DIST, POISSON.DIST.
 - **Funkce pro spojitá rozdělení:** NORM.DIST, T.DIST, F.DIST.
 - **Výpočet kvantilů:** NORM.INV, T.INV, F.INV.
-

6. Praktické aplikace rozdělení pravděpodobnosti

- **Binomické rozdělení:** Pravděpodobnost výhry v loterii nebo počet úspěšných hovorů v call centru.
 - **Poissonovo rozdělení:** Počet zákazníků za hodinu v obchodě.
 - **Normální rozdělení:** Výšky lidí v populaci.
 - **Chi-kvadrát rozdělení:** Test nezávislosti při kontingenčních tabulkách.
-

7. Kontrolní otázky pro studenta:

1. Jaký je rozdíl mezi diskrétním a spojitým rozdělením pravděpodobnosti?
 - **Odpověď:** Diskrétní rozdělení popisuje konečné nebo spočetné hodnoty, spojité interval hodnot.
2. Co popisuje binomické rozdělení a kde se používá?
 - **Odpověď:** Počet úspěchů v pevném počtu nezávislých pokusů, např. testování kvality výrobků.
3. Jaký je význam normálního rozdělení?
 - **Odpověď:** Popisuje mnoho přirozených jevů, např. výšku populace nebo výsledky testů.
4. K čemu slouží Studentovo t-rozdělení?
 - **Odpověď:** Pro testování hypotéz u malých vzorků.
5. Co je kvantil spojitého rozdělení?
 - **Odpověď:** Hodnota, která dělí pravděpodobnost na určité části.
6. Jak se používá rozdělení chi-kvadrát?
 - **Odpověď:** Testy shody, nezávislosti v kontingenčních tabulkách.
7. Jak se v Excelu vypočítá hodnota pravděpodobnosti pro binomické rozdělení?
 - **Odpověď:** Pomocí funkce BINOM.DIST.
8. Jaký je rozdíl mezi hustotou pravděpodobnosti a distribuční funkcí?

- **Odpověď:** Hustota popisuje rozložení hodnot, distribuční funkce pravděpodobnost nižší nebo stejné hodnoty.

9. Co vyjadřuje směrodatná odchylka?

- **Odpověď:** Průměrnou odchylku hodnot od střední hodnoty.

10. Jaké praktické příklady použití rozdělení pravděpodobnosti znáte?

- **Odpověď:** Analýza kvality výroby, odhad doby čekání, testy hypotéz.
-

5 Bodový a intervalový odhad

1. Úvod do tématu

- **Definice odhadu parametrů:** Postup, jak z výběrových dat určit neznámé parametry základního souboru.
 - **Typy odhadů:** Bodový odhad, intervalový odhad.
 - **Praktický význam odhadů:** Odhady parametrů v ekonomii, kvalitě výroby, finančních analýzách.
-

2. Statistické odhady a jejich vlastnosti

- **Statistický odhad:** Funkce výběrových dat, která aproximuje parametr základního souboru.
 - **Nevychýlený odhad:** Střední hodnota odhadu odpovídá skutečné hodnotě parametru.
 - **Konzistentní odhad:** S rostoucím rozsahem výběru se hodnota odhadu blíží skutečné hodnotě parametru.
-

3. Bodový odhad

- **Definice bodového odhadu:** Statistika určená k odhadu parametru základního souboru.
 - **Metoda momentů:** Odhad parametrů pomocí momentů výběru.
 - **(Metoda maximální věrohodnosti:** Hledá odhady parametrů maximalizující funkci věrohodnosti.)
 - **Vlastnosti bodového odhadu:** Nevychýlenost, efektivita, konzistence.
-

4. Intervalový odhad

- **Definice intervalového odhadu:** Interval, ve kterém s určitou pravděpodobností leží skutečná hodnota parametru.
- **Intervalový odhad střední hodnoty:** Výpočet intervalu pro střední hodnotu s danou hladinou spolehlivosti.
- **Intervalový odhad rozptylu:** Výpočet intervalu pro rozptyl pomocí chi-kvadrát rozdělení.
- **Kritické hodnoty:** Hodnoty vypočtené pomocí softwaru nebo z tabulek statistických rozdělení (t-rozdělení, chi-kvadrát, F-rozdělení).

5. Vztah mezi bodovým a intervalovým odhadem

- **Přesnost bodového odhadu:** Ovlivněna velikostí výběru a variabilitou dat.
- **Výhody intervalového odhadu:** Poskytuje rozmezí hodnot s určitou pravděpodobností.
- **Kombinace bodového a intervalového odhadu:** Používá se pro komplexní interpretaci výsledků.

6. Praktické aplikace bodových a intervalových odhadů

- **Odhad průměrného platu zaměstnanců**
- **Odhad průměrné životnosti výrobků**
- **Ekonomické modely pro predikci parametrů trhu**

7. Kontrolní otázky pro studenta (s odpověďmi):

1. Co je to bodový odhad a jaké jsou jeho hlavní vlastnosti?

- **Odpověď:** Bodový odhad je statistika používaná pro odhad parametru základního souboru. Hlavní vlastnosti jsou nevychýlenost, efektivita a konzistence.

2. Jaký je rozdíl mezi bodovým a intervalovým odhadem?

- **Odpověď:** Bodový odhad poskytuje jednu hodnotu pro parametr, intervalový poskytuje rozmezí hodnot s určitou pravděpodobností.

3. Jak se vypočítá intervalový odhad pro střední hodnotu?

- **Odpověď:** Pomocí vzorce pro interval spolehlivosti, který zahrnuje výběrový průměr, kritickou hodnotu a směrodatnou odchylku.

4. Jaká je metoda momentů a kde se používá?

- **Odpověď:** Parametry se odhadují pomocí momentů výběrového souboru. Používá se, když je rozdělení dobře známo.

5. (Jaký je princip metody maximální věrohodnosti?)

- **Odpověď:** Hledá hodnoty parametrů, které maximalizují pravděpodobnost pozorovaných dat.)

6. Co je hladina spolehlivosti a jak ovlivňuje intervalový odhad?

- **Odpověď:** Hladina spolehlivosti určuje pravděpodobnost, s jakou intervalový odhad obsahuje skutečný parametr. Vyšší hladina zvyšuje šířku intervalu.

7. Jak interpretujete kritické hodnoty v intervalovém odhadu?

- **Odpověď:** Kritické hodnoty určují hranice intervalu spolehlivosti na základě zvoleného rozdělení.

8. Jaký je vztah mezi velikostí výběru a přesností odhadu?

- **Odpověď:** S větší velikostí výběru se zvyšuje přesnost odhadu.

9. Co vyjadřuje interval spolehlivosti pro rozptyl?

- **Odpověď:** Intervalový odhad poskytuje pravděpodobné rozmezí hodnot pro rozptyl základního souboru.

10. Jaké praktické příklady odhadů znáte z ekonomie nebo výroby?

- **Odpověď:** Např. odhad průměrného platu zaměstnanců nebo odhad podílu vadných výrobků ve výrobním procesu.
-

6 Testování statistických hypotéz

1. Úvod do tématu

- **Definice statistické hypotézy:** Předpoklad o vlastnostech základního souboru, který se ověřuje pomocí statistických metod.
 - **Nulová a alternativní hypotéza:** Nulová hypotéza (H_0) předpokládá, že mezi zkoumanými jevy není rozdíl (žádný vztah). Alternativní hypotéza (H_1) je protikladem H_0 .
 - **Význam testování hypotéz:** Používá se k ověření hypotéz v ekonomii, vědě a průzkumech trhu.
-

2. Základní pojmy testování hypotéz

- **Nulová hypotéza (H_0):** Předpoklad, který se testuje a potenciálně zamítá.
 - **Alternativní hypotéza (H_1):** Předpoklad, který je platný, pokud je nulová hypotéza zamítnuta.
 - **Hladina významnosti (α):** Pravděpodobnost, s jakou může být H_0 nesprávně zamítnuta.
 - **Kritický obor:** Hodnoty, které vedou k zamítnutí H_0 .
 - **Chyba I. druhu:** Nesprávné zamítnutí pravdivé nulové hypotézy.
 - **Chyba II. druhu:** Nezamítnutí nepravdivé nulové hypotézy.
-

3. Typy statistických testů

- **Jednostranný test:** Testuje odchylku pouze jedním směrem.
 - **Oboustranný test:** Testuje odchylky oběma směry.
 - **Parametrické testy:** Vyžadují předpoklady o rozdělení dat (např. t-test, z-test, F-test).
 - **(Neparametrické testy:** Nevyžadují předpoklady o rozdělení dat.)
-

4. Testovací statistiky

- **t-statistika:** Používá se pro malé výběry a neznámý rozptyl základního souboru.
- **z-statistika:** Používá se pro velké výběry nebo známý rozptyl.
- **F-statistika:** Používá se pro analýzu rozptylu mezi více skupinami.

- **p-hodnota:** Pravděpodobnost získání výsledku, pokud je nulová hypotéza pravdivá.
-

5. Postup při testování hypotéz

1. Formulace nulové a alternativní hypotézy.
 2. Stanovení hladiny významnosti (α).
 3. Výběr vhodného statistického testu.
 4. Výpočet testovací statistiky.
 5. Porovnání s kritickou hodnotou nebo interpretace p-hodnoty.
 6. Rozhodnutí o zamítnutí nebo nezamítnutí hypotézy.
-

6. Interpretace výsledků testů

- **Kritický obor:** Hodnoty testovací statistiky, které vedou k zamítnutí nulové hypotézy.
 - **p-hodnota:** Pokud je p-hodnota menší než α , nulová hypotéza se zamítá.
 - **Praktická interpretace:** Výsledek testování by měl být vždy analyzován v kontextu zkoumaného problému.
-

7. Použití Excelu pro testování hypotéz

- **T.TEST:** Funkce pro t-test v Excelu.
 - **NORM.S.INV:** Výpočet kritických hodnot normálního rozdělení.
 - **Výpočet p-hodnoty:** Pomocí funkcí a nástrojů analýzy dat v Excelu.
-

8. Praktické aplikace testování hypotéz

- Testování účinnosti nového léku
 - Kontrola kvality výroby
 - Analýza výkonnosti zaměstnanců
 - Srovnání výsledků dvou marketingových kampaní
-

9. Kontrolní otázky pro studenta:

- Co je to nulová a alternativní hypotéza?

- **Odpověď:** Nulová hypotéza předpokládá, že mezi jevy není rozdíl (vztah), alternativní předpokládá opak.
 - Jaký je rozdíl mezi jednostranným a oboustranným testem?
 - **Odpověď:** Jednostranný testuje odchylku jedním směrem, oboustranný oběma směry.
 - Co je hladina významnosti a jak ovlivňuje rozhodování při testování hypotéz?
 - **Odpověď:** Je to pravděpodobnost chyby I. druhu.
 - Co vyjadřuje p-hodnota a jak ji interpretujeme?
 - **Odpověď:** Vyjadřuje pravděpodobnost získání výsledku nebo extrémnějšího výsledku při platnosti H_0 .
 - Jaký je rozdíl mezi chybou I. druhu a chybou II. druhu?
 - **Odpověď:** Chyba I. druhu je nesprávné zamítnutí H_0 , chyba II. druhu její nesprávné přijetí.
 - Jaký je postup při testování hypotéz?
 - **Odpověď:** Definice hypotéz (H_0 a H_1), volba testovací statistiky, stanovení hladiny významnosti, výpočet testovací statistiky, určení p-hodnoty nebo kritického oboru, rozhodnutí o zamítnutí či nezamítnutí H_0 .
 - Kdy použijeme t-test, z-test a F-test?
 - **Odpověď:** T-test se používá pro malé vzorky a neznámou směrodatnou odchylku, z-test pro velké vzorky s známou směrodatnou odchylkou, F-test pro porovnání rozptylů dvou skupin.
 - Jaké Excel funkce se používají pro testování hypotéz?
 - **Odpověď:** T.TEST, Z.TEST, F.TEST, CHISQ.TEST a funkce v nástroji Analýza dat.
 - Jak se určuje kritický obor a akceptační obor?
 - **Odpověď:** Kritický obor je oblast hodnot testovací statistiky, při kterých zamítáme H_0 , zatímco akceptační obor zahrnuje hodnoty, při kterých H_0 nezamítáme.
 - Jaké jsou praktické příklady použití testování hypotéz?
 - **Odpověď:** Testování účinnosti léku, kontrola kvality výrobků, srovnání průměrných příjmů mezi dvěma skupinami, analýza zákaznické spokojenosti.
-

7 Parametrické testy

1. Úvod do tématu

- **Definice parametrických testů:** Statistické testy používané k ověření hypotéz o parametrech základního souboru, jako je střední hodnota nebo rozptyl.
 - **Hlavní účel parametrických testů:** Ověřit hypotézy na základě dat, která splňují například předpoklady o normálním rozdělení a homogenitě rozptylů.
 - **Praktické využití parametrických testů:** Ekonomie, medicína, kvalita výroby.
-

2. Základní pojmy parametrických testů

- **Nulová hypotéza (H_0):** Předpoklad, že mezi parametry není rozdíl.
 - **Alternativní hypotéza (H_1):** Předpoklad, že mezi parametry existuje rozdíl.
 - **Hladina významnosti (α):** Pravděpodobnost chyby I. druhu při testování hypotézy.
 - **P-hodnota:** Pravděpodobnost získání výsledku alespoň tak extrémního, pokud platí nulová hypotéza.
 - **Kritická hodnota:** Hodnota, která určuje hraniční bod pro zamítnutí H_0 .
-

3. Typy parametrických testů

- **Jednovýběrový t-test:** Testuje hypotézu o střední hodnotě jedné populace.
 - **Dvouvýběrový t-test:** Porovnává střední hodnoty dvou nezávislých výběrů.
 - **Párový t-test:** Porovnává střední hodnoty dvou závislých výběrů.
 - **F-test:** Testuje hypotézu o rovnosti rozptylů dvou souborů.
-

4. Předpoklady parametrických testů

- Normalita dat
 - Homogenita rozptylů
 - Nezávislost pozorování
-

5. Postup při provádění parametrických testů

1. Formulace nulové a alternativní hypotézy.
2. Stanovení hladiny významnosti (α).

3. Výběr vhodného testu (t-test, F-test).
 4. Výpočet testovací statistiky.
 5. Porovnání s kritickou hodnotou nebo interpretace p-hodnoty.
 6. Rozhodnutí o zamítnutí nebo nezamítnutí hypotézy.
-

6. Interpretace výsledků parametrických testů

- **Význam kritického oboru a p-hodnoty:** Indikace, zda zamítnout H_0 .
 - **Praktická interpretace výsledků:** Závěry musí zohledňovat kontext problému .
-

7. Použití Excelu pro parametrické testy

- **Funkce T.TEST:** Pro jednovýběrový, dvouvýběrový nebo párový t-test.
 - **Funkce F.TEST:** Pro testování rozdílu mezi rozptyly.
 - **Nástroj „Analýza dat“:** Automatizované testování hypotéz.
-

8. Praktické aplikace parametrických testů

- **Ekonomie:** Porovnávání průměrných mezd v různých odvětvích.
 - **Medicína:** Testování účinnosti nového léku.
 - **Výroba:** Kontrola kvality produktů.
-

9. Kontrolní otázky pro studenta (s odpověďmi):

- **Co jsou parametrické testy a jaký je jejich účel?**
 - **Odpověď:** Statistické testy sloužící k ověřování hypotéz o parametrech základního souboru.
- **Jaké jsou základní předpoklady pro použití parametrických testů?**
 - **Odpověď:** Normalita dat, homogenita rozptylů, nezávislost pozorování.
- **Jaký je rozdíl mezi jednovýběrovým a dvouvýběrovým t-testem?**
 - **Odpověď:** Jednovýběrový testuje střední hodnotu jedné populace, dvouvýběrový porovnává dvě nezávislé populace.
- **Kdy použijeme párový t-test?**
 - **Odpověď:** K porovnání středních hodnot dvou závislých výběrů.

- **Jaký je účel F-testu?**
 - **Odpověď:** Testuje hypotézu o rovnosti rozptylů dvou souborů.
 - **Jak interpretujeme p-hodnotu?**
 - **Odpověď:** Pokud je p-hodnota menší než α , zamítáme nulovou hypotézu.
 - **Jaké jsou praktické aplikace parametrických testů?**
 - **Odpověď:** Ekonomické analýzy, medicínské studie, kontrola kvality.
 - **Jak lze v Excelu provést t-test?**
 - **Odpověď:** Pomocí funkce T.TEST nebo nástroje „Analýza dat“.
 - **Co znamená hladina významnosti?**
 - **Odpověď:** Pravděpodobnost chyby I. druhu při testování hypotézy.
 - **Jaký je postup při provádění parametrických testů?**
 - **Odpověď:** Formulace hypotéz, stanovení α , výběr testu, výpočet statistiky, interpretace výsledků.
-

8 Analýza rozptylu (ANOVA)

1. Úvod do tématu

- **Definice analýzy rozptylu (ANOVA):** Statistická metoda používaná k porovnání průměrů mezi více než dvěma skupinami.
 - **Hlavní účel ANOVA:** Zjistit, zda existují statisticky významné rozdíly mezi průměry více skupin.
 - **Praktické využití ANOVA:** Používá se v marketingu, personalistice, výrobě a financích.
-

2. Základní principy analýzy rozptylu

- **Variabilita mezi skupinami (meziskupinová):** Variabilita způsobená rozdíly mezi průměry skupin.
 - **Variabilita uvnitř skupin (vnitroskupinová):** Variabilita způsobená rozdíly uvnitř jednotlivých skupin.
 - **Celková variabilita:** Součet meziskupinové a vnitroskupinové variability.
 - **F-statistika:** Poměr mezi meziskupinovou a vnitroskupinovou variabilitou, který určuje statistickou významnost výsledku.
-

3. Typy ANOVA

- **Jednofaktorová ANOVA:** Porovnává průměry mezi skupinami v rámci jednoho faktoru.
 - **(Dvoufaktorová ANOVA:** Zkoumá vliv dvou faktorů na sledovanou proměnnou a jejich interakci.
 - **ANOVA s opakovanými měřeními (Repeated Measures ANOVA):** Používá se, když jsou hodnoty měřeny opakovaně na stejných subjektech.)
-

4. Předpoklady ANOVA

- **Normalita dat:** Data by měla být normálně rozložena.
 - **Homogenita rozptylů:** Rozptyly skupin by měly být přibližně stejné.
 - **Nezávislost pozorování:** Pozorování musí být nezávislá.
-

5. Postup při provádění ANOVA

1. Formulace nulové a alternativní hypotézy.
 2. Výpočet meziskupinové a vnitroskupinové variability.
 3. Výpočet stupňů volnosti.
 4. Výpočet středních čtverců (MS).
 5. Výpočet F-statistiky.
 6. Porovnání F-statistiky s kritickou hodnotou.
 7. Rozhodnutí o zamítnutí nebo nezamítnutí nulové hypotézy.
-

6. Interpretace výsledků ANOVA

- **F-statistika a kritická hodnota:** Pokud je vypočtená F-statistika větší než kritická hodnota, zamítá se nulová hypotéza.
 - **p-hodnota:** Pokud je p-hodnota menší než hladina významnosti α , rozdíly mezi skupinami jsou statisticky významné.
 - **(Post-hoc testy:** Např. Tukeyho test se používá k identifikaci konkrétních rozdílů mezi skupinami.)
-

7. Použití Excelu pro ANOVA

- **Jednofaktorová ANOVA:** Analýza dat pomocí nástroje „Analýza dat“.
 - **Interpretace výsledků v Excelu:** Zahrnuje hodnoty F, p-hodnotu a tabulku ANOVA.
-

8. Praktické aplikace ANOVA

- **Marketing:** Porovnání účinnosti různých reklamních kampaní.
 - **Personalistika:** Srovnání výkonnosti zaměstnanců v různých týmech.
 - **Výroba:** Testování vlivu výrobních procesů na kvalitu produktu.
 - **Finance:** Analýza výnosů z investičních portfolií.
-

9. Kontrolní otázky pro studenta (s odpověďmi):

1. **Co je to analýza rozptylu a k čemu slouží?**
 - **Odpověď:** ANOVA je statistická metoda pro porovnání průměrů více než dvou skupin za účelem zjištění statisticky významných rozdílů.
2. **Jaké jsou hlavní předpoklady pro použití ANOVA?**

- **Odpověď:** Normalita dat, homogenita rozptylů a nezávislost pozorování.
3. **Vysvětlete rozdíl mezi variabilitou mezi skupinami a variabilitou uvnitř skupin.**
- **Odpověď:** Meziskupinová variabilita zkoumá rozdíly mezi průměry skupin, vnitroskupinová zkoumá rozdíly uvnitř jednotlivých skupin.
4. **Jaký je postup při provádění jednofaktorové ANOVA?**
- **Odpověď:** Formulace hypotéz, výpočet variabilit, stupňů volnosti, středních čtverců, F-statistiky a interpretace výsledků.
5. **Co znamená, pokud je vypočtená hodnota F větší než kritická hodnota?**
- **Odpověď:** Znamená to, že existuje statisticky významný rozdíl mezi průměry skupin.
6. **(Jaký je význam post-hoc testů?)**
- **Odpověď:** Identifikují konkrétní dvojice skupin s významnými rozdíly.)
7. **Jak se interpretují výsledky ANOVA v Excelu?**
- **Odpověď:** Analýzou hodnot F, p-hodnoty a dalších výstupů z tabulky ANOVA.
8. **Jaké praktické aplikace ANOVA znáte?**
- **Odpověď:** Marketing, personalistika, výroba, finance.
9. **(Kdy je vhodné použít dvoufaktorovou ANOVA?)**
- **Odpověď:** Když zkoumáme vliv dvou faktorů a jejich interakce.)
10. **Co vyjadřuje p-hodnota v ANOVA?**
- **Odpověď:** Pravděpodobnost, že získané rozdíly jsou náhodné.
-

9 Korelační analýza

1. Úvod do tématu

- **Definice korelační analýzy:** Statistická metoda zkoumající sílu a směr lineárního vztahu mezi dvěma proměnnými.
 - **Korelační koeficient:** Míra síly a směru lineárního vztahu, pohybuje se mezi -1 a 1.
 - **Praktické využití korelační analýzy:** Ekonomie, medicína, marketing, sociální vědy.
-

2. Typy korelačních koeficientů

- **Pearsonův korelační koeficient:** Měří lineární vztah mezi dvěma spojitými proměnnými.
 - **(Spearmanův korelační koeficient:** Vhodný pro ordinální data nebo nelineární monotónní vztahy.
 - **Kendallův tau:** Používá se pro malé soubory dat.
 - **Point-biserial korelace:** Pro vztah mezi spojitou a binární proměnnou.)
-

3. Výpočet korelačního koeficientu

- Vzorec pro Pearsonův korelační koeficient:

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$$

Význam hodnot korelačního koeficientu:

- **r = 1:** Perfektní pozitivní korelace
 - **r = -1:** Perfektní negativní korelace
 - **r = 0:** Žádná lineární korelace
-

4. Testování významnosti korelačního koeficientu

- **Nulová hypotéza (H₀):** Mezi proměnnými není lineární vztah.
- **Testovací statistika:**

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

- **Interpretace výsledků:** Porovnání hodnoty t s kritickou hodnotou z t -rozdělení.
-

5. Předpoklady pro korelační analýzu

- Lineární vztah mezi proměnnými
 - Normální rozložení obou proměnných
 - Absence výrazně odlehlých hodnot
-

6. Praktické použití Excelu pro korelační analýzu

- **Funkce CORREL:** Výpočet Pearsonova korelačního koeficientu.
 - **Analýza dat v Excelu:** Nástroj pro statistickou analýzu korelace.
 - **Interpretace výstupu:** Hodnota r a p -hodnota pro statistickou významnost.
-

7. Praktické aplikace korelační analýzy

- **Ekonomie:** Závislost mezi výdaji na reklamu a tržbami.
 - **Medicína:** Vztah mezi věkem a krevním tlakem.
 - **Marketing:** Vliv ceny na poptávku.
-

8. Omezení korelační analýzy

- Korelace \neq Kauzalita
 - Vliv odlehlých hodnot
 - Nevhodnost pro nelineární vztahy
-

9. Kontrolní otázky pro studenta (s odpověďmi):

1. **Co je to korelační koeficient a k čemu slouží?**
 - **Odpověď:** Míra síly a směru lineárního vztahu mezi dvěma proměnnými.
2. **Jaké jsou hlavní typy korelačních koeficientů?**
 - **Odpověď:** Pearsonův, Spearmanův, Kendallův tau, Point-biserial.
3. **Jak interpretujeme hodnoty Pearsonova korelačního koeficientu?**
 - **Odpověď:** $r = 1$ (perfektní pozitivní korelace), $r = -1$ (perfektní negativní korelace), $r = 0$ (žádná korelace).

4. **Jaký je vzorec pro výpočet korelačního koeficientu?**

- **Odpověď:** Viz vzorec výše.

5. **Jak testujeme významnost korelačního koeficientu?**

- **Odpověď:** Pomocí t-statistiky a porovnáním s kritickou hodnotou z t-rozdělení.

6. **Jaké jsou předpoklady pro korelační analýzu?**

- **Odpověď:** Lineární vztah, normální rozložení dat, absence odlehlých hodnot.

7. **Jaký nástroj v Excelu používáme pro výpočet korelačního koeficientu?**

- **Odpověď:** Funkce CORREL nebo Analýza dat.

8. **Jaké jsou praktické aplikace korelační analýzy?**

- **Odpověď:** Ekonomie, medicína, marketing.

9. **Co znamená, že korelace neimplikuje kauzalitu?**

- **Odpověď:** Korelace neznamena, že jedna proměnná způsobuje změnu druhé.

10. **Jaké jsou limity použití Pearsonova korelačního koeficientu?**

- **Odpověď:** Nevhodnost pro nelineární vztahy, citlivost na odlehlé hodnoty.
-

10 Lineární regrese

1. Úvod do tématu

- **Definice lineární regrese:** Statistická metoda pro modelování vztahu mezi závislou a nezávislou proměnnou pomocí přímky.
 - **Regresní rovnice:** $Y = \beta_0 + \beta_1 X + \varepsilon Y$.
 - **Praktické využití lineární regrese:** Ekonomie, marketing, predikce prodeje.
-

2. Základní pojmy lineární regrese

- **Závislá proměnná (Y):** Proměnná, kterou chceme předpovědět.
 - **Nezávislá proměnná (X):** Proměnná, která ovlivňuje hodnotu Y.
 - **Směrnice přímky ($\widehat{\beta}_1$):** Udává změnu v Y při jednotkové změně X.
 - **Absolutní člen ($\widehat{\beta}_0$):** Hodnota Y, když X je nulová.
 - **Koeficient determinace (R^2):** Procento variability Y vysvětlené modelem.
-

3. Odhad parametrů regresního modelu

- **Metoda nejmenších čtverců:** Minimalizuje součet čtverců odchylek mezi skutečnými a predikovanými hodnotami.

- **Vzorce pro odhady:**

- $$\widehat{\beta}_1 = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2}$$

- $$\widehat{\beta}_0 = \bar{Y} - \widehat{\beta}_1 \bar{X}$$

4. Předpoklady lineární regrese

- **Linearita:** Vztah mezi X a Y je lineární.
 - **Homoskedasticita:** Konstantní rozptyl reziduí.
 - **Nezávislost:** Hodnoty reziduí jsou nezávislé.
 - **Normalita:** Rezidua mají normální rozdělení.
-

5. Testování významnosti regresních koeficientů

- **Nulová hypotéza (H_0):** Koeficient je roven nule (nemá vliv na Y).
- **Testovací statistika:**

$$t = \frac{\widehat{\beta}_1}{SE(\widehat{\beta}_1)}$$

- **p-hodnota:** Pravděpodobnost získání výsledku za předpokladu platnosti H_0 .
-

6. Interpretace výsledků lineární regrese

- **Regresní koeficienty:** Směrnice a absolutní člen.
 - **Koeficient determinace (R^2):** Podíl vysvětlené variability.
 - **Reziduální analýza:** Hodnocení kvality modelu.
-

7. Použití Excelu pro regresní analýzu

- **Nástroj Analýza dat – Regrese:** Automatizované výpočty.
 - **Interpretace výstupu:** Hodnoty koeficientů, p-hodnoty, R^2 .
 - **Praktické příklady v Excelu:** Predikce hodnot a validace modelu.
-

8. Praktické aplikace lineární regrese

- **Ekonomie:** Predikce tržeb na základě výdajů na reklamu.
 - **Marketing:** Analýza vlivu reklamních kampaní na prodeje.
 - **Finance:** Odhad cen akcií na základě makroekonomických ukazatelů.
-

9. Kontrolní otázky pro studenta (s odpověďmi):

1. Co je to lineární regrese a k čemu slouží?

- **Odpověď:** Metoda pro modelování vztahu mezi závislou a nezávislou proměnnou.

2. Jaké jsou předpoklady lineární regrese?

- **Odpověď:** Linearita, homoskedasticita, nezávislost, normalita.

3. Jak se odhadují parametry regresního modelu?

- **Odpověď:** Pomocí metody nejmenších čtverců.

4. Co vyjadřuje směrnice přímky ($\widehat{\beta}_1$) a absolutní člen ($\widehat{\beta}_0$)?

- **Odpověď:** $\widehat{\beta}_1$ ukazuje změnu Y při jednotkové změně X, $\widehat{\beta}_0$ je hodnota Y při X = 0.

5. Co je to koeficient determinace R^2 a jak se interpretuje?

- **Odpověď:** Udává, jaká část variability Y je vysvětlena modelem.

6. Jaký je postup při testování významnosti regresních koeficientů?

- **Odpověď:** Stanovíme nulovou hypotézu (H_0 : koeficient = 0), vypočítáme testovací statistiku (t-test), porovnáme s kritickou hodnotou nebo p-hodnotou a rozhodneme o významnosti koeficientu.

7. Proč je důležité ověřit předpoklady regresního modelu?

- **Odpověď:** Nesplnění předpokladů (např. normalita, homoskedasticita, nezávislost chyb) může vést k nesprávným závěrům a zkresleným výsledkům analýzy.

8. Jak lze použít Excel pro regresní analýzu?

- **Odpověď:** Pomocí nástroje „Analýza dat“ → „Regrese“ lze odhadnout koeficienty, získat statistiky modelu a ověřit předpoklady.

9. Uvedte příklad aplikace lineární regrese v marketingu.

- **Odpověď:** Analýza vztahu mezi reklamními výdaji a prodejem produktu pro optimalizaci marketingového rozpočtu.

10. Kdy by bylo vhodné použít nelineární regresi?

- **Odpověď:** Když vztah mezi proměnnými není lineární, např. při modelování růstu populace nebo při analýze efektu snižujících se výnosů.
-