

STATISTIKA 2

Statistika a statistické zpracování dat

Blok 3-2

Jiří Fišer

29. listopadu 2024

Kapitola 8

Parametrické testy

Parametrické hypotézy jsou tvrzení o parametrech rozdělení v populaci (např. střední hodnota, rozptyl). Tyto hypotézy můžeme formulovat různými způsoby, například jako rovnost určitého parametru s konkrétní hodnotou (např. „průměrný objem piva ve všech sudech je 50 l“) nebo jako rovnost mezi parametry dvou různých populací (např. „rozptyly objemu piva ve dvou různých výrobních šaržích jsou stejné“).

Parametrické testy jsou statistické testy, které se používají k ověření těchto parametrických hypotéz. Abychom mohli použít parametrický test, musíme předpokládat určité vlastnosti rozdělení dat, například že data pocházejí z normálního rozdělení. Parametrické testy jsou tedy úzce spojeny s parametrickými hypotézami, protože slouží k jejich testování na základě vzorků dat.

Druhy parametrických hypotéz

1. **Hypotézy o parametru jedné populace** (např. střední hodnota, medián, rozptyl, relativní četnost, ...).
2. **Hypotézy o parametrech dvou populací** (např. srovnávání středních hodnot nebo rozptylů mezi dvěma skupinami).
3. **Hypotézy o parametrech více než dvou populací** (např. analýza rozptylu – ANOVA, ...).

Hypotézy o rozptylu

Test významnosti rozdílu dvou rozptylů (F-test)

Předpoklady

Jsou dány dva výběry o rozsazích n_1 a n_2 s výběrovými rozptyly s_1^2 a s_2^2 , vybrané ze dvou základních souborů s rozděleními $N(\mu_1, \sigma_1^2)$ a $N(\mu_2, \sigma_2^2)$, jejichž parametry neznáme.

- Nulová hypotéza: $H_0 : \sigma_1^2 = \sigma_2^2$.
- Alternativní hypotéza: $H_1 : \sigma_1^2 \neq \sigma_2^2$.

Hypotézy o rozptylu

Test významnosti rozdílu dvou rozptylů (F-test)

Testovací statistika

Testovací statistika

$$F = \frac{s_1^2}{s_2^2},$$

kde

- s_1^2 a s_2^2 jsou výběrové rozptyly pro tyto soubory,

má Fisherovo-Snedecorovo rozdělení $F(n_1 - 1, n_2 - 1)$, kde n_1 a n_2 jsou velikosti dvou výběrových souborů.

Hypotézy o rozptylu

Test významnosti rozdílu dvou rozptylů (F-test)

Závěr pro oboustranný test

- $H_0: \sigma_1^2 = \sigma_2^2, \quad H_1: \sigma_1^2 \neq \sigma_2^2$

Jestliže

$$F < F_{\frac{\alpha}{2}}(n_1 - 1, n_2 - 1) \quad \text{nebo} \quad F > F_{1-\frac{\alpha}{2}}(n_1 - 1, n_2 - 1),$$

potom zamítáme hypotézu H_0 (ve prospěch H_1).

Hypotézy o rozptylu

Test významnosti rozdílu dvou rozptylů (F-test)

Závěr pro jednostranné testy

- $H_0: \sigma_1^2 \leq \sigma_2^2, \quad H_1: \sigma_1^2 > \sigma_2^2$

Jestliže

$$F > F_{1-\alpha}(n_1 - 1, n_2 - 1),$$

potom zamítáme hypotézu H_0 (ve prospěch H_1).

- $H_0: \sigma_1^2 \geq \sigma_2^2, \quad H_1: \sigma_1^2 < \sigma_2^2$

Jestliže

$$F < F_{\alpha}(n_1 - 1, n_2 - 1),$$

potom zamítáme hypotézu H_0 (ve prospěch H_1).

Příklad 8.1. Byly sledovány výsledky běhu na 50 m (ve vteřinách) u skupiny desetiletých chlapců (tabulka 6) a dívek (tabulka 5). Posudte získané výsledky z hlediska vyrovnanosti výkonů v jednotlivých skupinách.

Tab. 5: Výsledky běhu na 50 m (ve vteřinách) u skupiny dívek

1	2	3	4	5	6	7	8	9	10	11	12	13	14
10,70	10,80	10,00	10,60	9,20	10,20	9,90	10,00	9,30	10,20	9,80	10,00	10,00	11,00
15	16	17	18	19	20	21	22	23	24	25	26	27	28
12,00	10,00	10,00	11,20	9,40	10,70	9,30	10,10	9,10	10,20	9,30	10,00	9,40	10,90

Tab. 6: Výsledky běhu na 50 m (ve vteřinách) u skupiny chlapců

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
10,80	9,30	9,40	9,90	10,20	9,30	9,40	8,90	9,60	9,70	10,60	9,40	9,50	9,60	10,00	9,30	9,40
18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	
9,40	8,40	9,80	8,80	9,20	9,50	9,80	9,00	10,50	9,40	9,30	9,90	9,10	9,60	8,70	8,10	

Řešení: Hladinu významnosti zvolíme $\alpha = 0,05$.

Určíme potřebné charakteristiky u obou skupin:

$$n_1 = 28, \quad s_1^2 = 0,4689 \quad (\text{dívky}), \quad n_2 = 33, \quad s_2^2 = 0,3405 \quad (\text{chlapci}).$$

$$F_{0,025}(27, 32) = \text{F.INV}(0,025; 27; 32) \approx 0,4722,$$

$$F_{0,975}(27, 32) = \text{F.INV}(0,975; 27; 32) \approx 2,0689.$$

Hodnota testovací statistiky leží mezi kritickými hodnotami, tudíž leží v oboru akceptace nulové hypotézy. Takto tedy nemůžeme zamítnout nulovou hypotézu H_0 . To znamená, že mezi rozptyly není statisticky významný rozdíl. $\square$

Hypotézy o střední hodnotě

Jednovýběrový t-test

Jednovýběrový t-test se používá k ověření, zda průměrná hodnota v základním souboru (populaci) μ se rovná konkrétní hypotetické hodnotě μ_0 , a to na základě údajů získaných z výběrového souboru.

Jednovýběrový t-test

Předpoklady

Předpokládáme, že máme výběr ze základního souboru, který má normální rozdělení $N(\mu, \sigma^2)$.

Výběr má rozsah n , výběrový průměr $\bar{x}$ a výběrový rozptyl s^2 .

Je třeba odlišit, že μ a σ^2 jsou neznámé parametry základního souboru, zatímco $\bar{x}$ a s^2 jsou výběrové charakteristiky.

- Nulová hypotéza: $H_0: \mu = \mu_0$,
- Alternativní hypotéza: $H_1: \mu \neq \mu_0$.

Jednovýběrový t-test

Testovací statistika

Testovací statistika T

$$T = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}},$$

kde

- $\bar{x}$ je výběrový průměr,
- μ_0 je hypotetická hodnota průměru podle nulové hypotézy,
- s je výběrová směrodatná odchylka,
- n je velikost výběrového souboru,

má Studentovo t-rozdělení s $n - 1$ stupni volnosti.

Jednovýběrový t-test

Závěr pro oboustranný test

- $H_0: \mu = \mu_0, \quad H_1: \mu \neq \mu_0$

Jestliže

$$T < t_{\frac{\alpha}{2}}(n-1) \quad \text{nebo} \quad T > t_{1-\frac{\alpha}{2}}(n-1),$$

potom zamítáme hypotézu H_0 (ve prospěch H_1).

$$t_{\frac{\alpha}{2}}(n-1) = -t_{1-\frac{\alpha}{2}}(n-1)$$

Jednovýběrový t-test

Závěr pro jednostranné testy

- $H_0: \mu \leq \mu_0, \quad H_1: \mu > \mu_0$

Jestliže

$$T > t_{1-\alpha}(n-1),$$

potom zamítáme hypotézu H_0 (ve prospěch H_1).

- $H_0: \mu \geq \mu_0, \quad H_1: \mu < \mu_0$

Jestliže

$$T < t_{\alpha}(n-1),$$

potom zamítáme hypotézu H_0 (ve prospěch H_1).

Jednovýběrový t-test

Příklad 8.3. V pivovaru došlo k opravě plnicí linky. Na hladině významnosti $\alpha = 0,05$ ověřte, zda se oprava zdařila, tj., zda linka plní do lahví pivo o objemu 500 ml. Výsledky u vybraných vzorků (v mililitrech) jsou uvedeny v tabulce 7.

Tab. 7: Výsledky u vybraných vzorků objemu piva (v mililitrech)

495,2	496,8	502,1	498,5	501	503	500,7
501,5	501,8	499,1	500,9	502,2	501,7	500,4
500,2	501	499,9	500,2	501,1	500,8	499,3

Řešení: Formulace hypotéz:

Na základě hypotetické hodnoty $\mu_0 = 500$ ml formulujeme následující hypotézy:

- $H_0: \mu = 500$,
- $H_1: \mu \neq 500$.

Výpočet základních charakteristik:

Z poskytnutých dat vypočteme následující výběrové charakteristiky:

$$n = 21, \quad \bar{x} = 500,3571 \text{ ml}, \quad s = 1,77806 \text{ ml}.$$

Testovací statistika:

$$T = \frac{500,3571 - 500}{\frac{1,77806}{\sqrt{21}}} \approx 0,898.$$

Kritická hodnota (vypočteme např. v Excelu pomocí funkce T.INV):

$$t_{0,025}(20) = \text{T.INV}(0,025; 20) \approx -2,08596, \quad t_{0,975}(20) = \text{T.INV}(0,975; 20) \approx 2,08596.$$

Závěr:

Protože hodnota testovacího kritéria je mezi kritickými hodnotami, tak nemůžeme zamítnout nulovou hypotézu. Můžeme tedy konstatovat, že možná odchylka není statisticky významná. Šéfovi oznámíme, že oprava plnicí linky byla úspěšná a linka plní lahve správně. $\square$

Dvouvýběrový t-test

Dvouvýběrový t-test se používá k porovnání středních hodnot dvou základních souborů, na základě dvou nezávislých výběrů z těchto souborů. Budeme rozlišovat dva případy, při rovnosti rozptylů ($\sigma_1^2 = \sigma_2^2$) klasický dvouvýběrový t-test a při nerovnosti rozptylů ($\sigma_1^2 \neq \sigma_2^2$) tzv. Welchův t-test.

Předpoklady:

Předpokládáme, že oba základní soubory mají normální rozdělení $N(\mu_1, \sigma_1^2)$ a $N(\mu_2, \sigma_2^2)$.

Máme k dispozici dva nezávislé výběry o rozsazích n_1 a n_2 , jejichž výběrové průměry označíme $\bar{x}_1$ a $\bar{x}_2$, a výběrové rozptyly s_1^2 a s_2^2 .

- Nulová hypotéza: $H_0: \mu_1 = \mu_2,$
- Alternativní hypotéza: $H_1: \mu_1 \neq \mu_2.$

Dvouvýběrový t-test

Testová statistika ($\sigma_1^2 = \sigma_2^2$)

Testová statistika pro případ **stejných rozptylů** ($\sigma_1^2 = \sigma_2^2$) (což lze prověřit pomocí F-testu):

$$T = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

kde

- $\bar{x}_1$ a $\bar{x}_2$ jsou výběrové průměry,
- s_1^2 a s_2^2 jsou výběrové rozptyly,
- n_1 a n_2 jsou velikosti výběrových souborů,

má Studentovo t-rozdělení s $n_1 + n_2 - 2$ stupni volnosti.

Dvouvýběrový t-test

Testová statistika ($\sigma_1^2 = \sigma_2^2$)

Testová statistika pro případ **stejných rozptylů** ($\sigma_1^2 = \sigma_2^2$) (což lze prověřit pomocí F-testu):

$$T = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

kde

- $\bar{x}_1$ a $\bar{x}_2$ jsou výběrové průměry,
- s_1^2 a s_2^2 jsou výběrové rozptyly,
- n_1 a n_2 jsou velikosti výběrových souborů,

má Studentovo t-rozdělení s $n_1 + n_2 - 2$ stupni volnosti.

Závěr ($\sigma_1^2 = \sigma_2^2$)

Pokud absolutní hodnota testovacího kritéria $|T|$ překročí kritickou hodnotu $t_{1-\frac{\alpha}{2}}(n_1 + n_2 - 2)$, zamítáme nulovou hypotézu H_0 .

Dvouvýběrový t-test

Testová statistika ($\sigma_1^2 \neq \sigma_2^2$)

Testová statistika pro případ, že **rozptyly** obou základních souborů jsou **rozdílné** (což lze prověřit pomocí F-testu),

$$T = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}},$$

kde

- $\bar{x}_1$ a $\bar{x}_2$ jsou výběrové průměry,
- s_1^2 a s_2^2 jsou výběrové rozptyly,
- n_1 a n_2 jsou velikosti výběrových souborů,

má přibližně Studentovo t-rozdělení, jehož počet stupňů volnosti se odhaduje pomocí

$$df \approx \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2-1}}.$$

Dvouvýběrový t-test

Závěr ($\sigma_1^2 \neq \sigma_2^2$)

Jestliže $|T|$ překročí kritickou hodnotu $t_{1-\frac{\alpha}{2}}(df)$, zamítáme nulovou hypotézu H_0 .

Dvouvýběrový t-test

Použití dvouvýběrového t-testu

Dvouvýběrový t-test používáme například k ověřování následujících hypotéz:

- Pocházejí dva vzorky z téhož základního souboru?
- Nedopustili jsme se při dvou měřeních, jejichž výsledkem bylo určení dvou středních hodnot $\bar{x}_1, \bar{x}_2$, systematických chyb?
- Má určitý faktor vliv na zkoumaný argument? Zde zkoumáme dva vzorky - jeden při působení daného faktoru, druhý bez jeho působení.

Dvouvýběrový t-test

Příklad 8.4. Odběratel dostává zářivky od dvou dodavatelů. Při hodnocení kvality zářivek se sleduje také počet zapojení, který snesou zářivky bez poškození. Zkoušky výrobků vedly k těmto výsledkům:

- Dodavatel A: 2139, 2041, 1968, 1903, 1952, 1980, 2089, 1915, 2389, 2163, 2072, 1712, 2018, 1792, 1849
- Dodavatel B: 1947, 1602, 1906, 2031, 2072, 1812, 1942, 2074, 2132

Ověřte hypotézu, že kvalita obou dodávek je stejná. Hladinu významnosti volte $\alpha = 0,05$.

Řešení: V Excelu vypočteme charakteristiky obou souborů:

- $n_1 = 15, \quad \bar{x}_1 = 1998,8, \quad s_1^2 = 27262,17,$
- $n_2 = 9, \quad \bar{x}_2 = 1946,4, \quad s_2^2 = 26498,52.$

Nejdříve provedeme F-test. Testovací statistika:

$$F = \frac{s_1^2}{s_2^2} = \frac{27262,17}{26498,52} \approx 1,0288.$$

Kritické hodnoty:

$$F_{0,025}(14; 8) = \text{F.INV}(0,025; 14; 8) \approx 0,3044, \quad F_{0,975}(14; 8) = \text{F.INV}(0,975; 14; 8) \approx 4,1297$$

Hypotézu o shodě rozptylů $\sigma_1^2 = \sigma_2^2$ nemůžeme zamítnout. Použijeme tedy dvouvýběrový t-test pro rovnost rozptylů:

Testovací statistika:

$$T = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \approx \frac{1998,8 - 1946,4}{\sqrt{\frac{(15-1)27262,17 + (9-1)26498,52}{15+9-2}} \cdot \sqrt{\frac{1}{15} + \frac{1}{9}}} \approx 0,7559.$$

Kritická hodnota:

$$t_{0,975}(22) = \text{T.INV}(0,975; 22) \approx 2,074.$$

Závěr:

Testovací kritérium nepřekročilo kritickou hodnotu, $H_0 : \mu_1 = \mu_2$ nemůžeme zamítnout. Kvalita obou dodávek nevykazuje významné statistické rozdíly. $\square$

Párový t-test

Úvod

Párový t-test se používá k porovnání dvou středních hodnot ze závislých výběrů, které jsou spárovány. To znamená, že každý prvek v prvním výběru x_{1i} má odpovídající prvek v druhém výběru x_{2i} . Tyto páry (x_{1i}, x_{2i}) jsou často výsledkem opakovaných měření na stejném subjektu nebo měření za různých podmínek.

Párový t-test

Předpoklady

Předpokládáme, že oba základní soubory mají normální rozdělení s parametry μ_1, σ_1^2 a μ_2, σ_2^2 . Rozsah obou výběrů je stejný, označíme ho n .

Když přejdeme z párů hodnot na jejich rozdíl $d_i = x_{1i} - x_{2i}$, kde d označíme střední hodnotu této nové veličiny, tak můžeme nulovou hypotézu o rovnosti středních hodnot μ_1 a μ_2 přeformulovat na $d = 0$:

- Nulová hypotéza: $H_0: d = 0$,
- Alternativní hypotéza: $H_1: d \neq 0$.

Výběrový průměr rozdílů $\bar{d}$ se vypočítá jako: $\bar{d} = \frac{\sum_i d_i}{n} = \frac{\sum_i (x_{1i} - x_{2i})}{n}$.

Párový t-test

Testovací statistika

Testovací statistika se vypočte jako

$$T = \frac{\bar{d}\sqrt{n}}{s_d},$$

kde s_d je výběrová směrodatná odchylka rozdílů d_i , vypočtená jako

$$s_d = \sqrt{\frac{\sum_{i=1}^n (d_i - \bar{d})^2}{n - 1}}.$$

Statistika T má při platnosti nulové hypotézy Studentovo t-rozdělení s $n - 1$ stupni volnosti.

Párový t-test

Závěr

Jestliže absolutní hodnota testovacího kritéria $|T|$ překročí kritickou hodnotu $t_{1-\frac{\alpha}{2}}(n-1)$, zamítáme nulovou hypotézu H_0 ve prospěch alternativní hypotézy H_1 .

Párový t-test

Příklad 8.5. Stanovení thiokyanového iontu (SCN^-) bylo paralelně provedeno dvěma metodami (Aldridge a Barker) na 12 vzorcích (tabulka 8). Srovnajte obě metodiky otestováním výsledků. Hladina významnosti $\alpha = 0,05$.

Tab. 8: Výsledky stanovení thiokyanového iontu

	1	2	3	4	5	6	7	8	9	10	11	12
Aldridge	0,38	0,56	0,45	0,49	0,38	0,41	0,6	0,36	0,26	0,41	0,43	0,4
Barker	0,39	0,58	0,44	0,52	0,41	0,45	0,59	0,37	0,28	0,42	0,42	0,38

Řešení: Nejprve vypočteme rozdíly $d_i = x_{1i} - x_{2i}$ pro každý pár měření (tabulka 9):

Tab. 9: Rozdíly d_i hodnoty thiokyanového iontu

	1	2	3	4	5	6	7	8	9	10	11	12
Aldridge	0,38	0,56	0,45	0,49	0,38	0,41	0,6	0,36	0,26	0,41	0,43	0,4
Barker	0,39	0,58	0,44	0,52	0,41	0,45	0,59	0,37	0,28	0,42	0,42	0,38
d_i	-0,01	-0,02	0,01	-0,03	-0,03	0,04	0,01	-0,01	-0,02	-0,01	0,01	0,02

Z těchto rozdílů vypočítáme výběrový průměr $\bar{d}$ a směrodatnou odchylku s_d :

$$\bar{d} = \frac{\sum_i d_i}{n} = \frac{-0,12}{12} = -0,01.$$

Směrodatná odchylka: $s_d = 0,018257$.

Testovací statistika:

$$t = \frac{\bar{d}\sqrt{n}}{s_d} = \frac{-0,01 \cdot \sqrt{12}}{0,018257} \approx 1,8166.$$

Kritická hodnota:

$$t_{0,975}(11) = \text{T.INV}(0,975; 11) \approx 2,201.$$

Závěr:

Testovací kritérium nepřekročilo kritickou hodnotu, tudíž nemůžeme zamítnout nulovou hypotézu H_0 . Obě metodiky dávají přibližně stejné výsledky, bez statisticky významných rozdílů. $\square$

Vyřešte:

8. V tabulce jsou uvedeny výsledky měření tlaku v pneumatikách před a po opravě. Použijte párový t-test k ověření, zda došlo ke statisticky významné změně tlaku po opravě na hladině významnosti $\alpha = 0,05$.

Před opravou	32,5	31,8	33,2	32,0	31,5	33,0	32,3	31,7
Po opravě	32,2	31,9	33,0	31,8	31,6	32,9	32,4	31,8

[p-hodnota = 0,4304, změna není statisticky významná]

Vyřešte:

9. Ověřte pomocí dvouvýběrového t-testu, zda existuje statisticky významný rozdíl mezi průměrnými platy dvou skupin zaměstnanců ve firmě. Data pro obě skupiny jsou uvedena v tabulce. Hladinu významnosti zvolte $\alpha = 0,05$.

Skupina A	45,000	48,000	46,500	47,200	44,800	45,900
Skupina B	50,200	49,500	51,000	48,800	50,100	49,900

[p-hodnota = 0,0001, statisticky významný rozdíl existuje]

Vyřešte:

10. Použijte F-test k porovnání rozptylů výsledků dvou testovacích metod, které byly použity na stejném souboru vzorků. Data jsou uvedena v tabulce. Ověřte hypotézu o shodě rozptylů na hladině významnosti $\alpha = 0,05$.

Metoda 1	8,5	8,3	8,6	8,4	8,7	8,5
Metoda 2	8,1	8,4	8,3	8,2	8,5	8,3

[p-hodnota = 1, hypotéza o shodě rozptylů nezamítnuta]