KGG/STG Statistika pro geografy

7. Testování statistických hypotéz

Mgr. David Fiedor
30. března 2015

Osnova

Dělení testů

- **parametrické** - o parametrech rozdělení základního souboru (průměr, rozptyl, shoda průměrů, . . .)
    - výběr musí pocházet z normálního rozdělení
    - data musí být intervalového nebo poměrového typu
- **neparametrické** - obdoba parametrických
    - používáme v případech, kdy nelze použít test parametrický
    - i pro data ordinálního typu
    - slabší než parametrické testy

Základní charakteristika

- úkolem je řešit úlohy o parametrech normálního rozdělení $N(\mu, \sigma^2)$
- stejně jako u intervalového odhadu (metoda testování) záleží na znalosti, resp. neznalosti těchto parametrů
- z-test, t-test, F-test

z-test

- jednovýběrový z-test
- dvouvýběrový z-test

- z důvodu neexistence speciální funkce pro tento test v programu STATISTICA - nebudeme uvádět pro tento test testování pomocí *p*-hodnoty

Jednovýběrový z-test

Nechť $X_1, \dots, X_n$ je náhodný výběr z rozdělení $N(\mu, \sigma^2)$, kde σ^2 **známe**. Nechť $n \geq 2$ a c je konstanta. Test H_0: $\mu = c$ proti H_1: $\mu \neq c$ se nazývá *jednovýběrový z-test*.

- oboustranná varianta testu

Jednovýběrový z-test

Testování pomocí intervalů spolehlivosti
Vzorce pro meze $100(1-\alpha)\%$ intervalů spolehlivosti pro střední hodnotu μ, jestliže známe rozptyl σ^2:

- oboustranný:
$$(d,h) = \left(\bar{x} - \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2}, \bar{x} + \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2}\right)$$
- levostranný: $(d,\infty) = \left(\bar{x} - \frac{\sigma}{\sqrt{n}} z_{1-\alpha}, \infty\right)$
- pravostranný: $(-\infty,h) = \left(-\infty, \bar{x} + \frac{\sigma}{\sqrt{n}} z_{1-\alpha}\right)$

Stačí rozhodnout, zda daná hodnota c leží v daném intervalu (pokud ne, zamítáme H_0).

Příklad

Úřad pro ochranu životního prostředí vydal nařízení, kterým průmyslovým podnikům určuje maximální přípustnou koncentraci vinylchloridu ve vzduchu, 50 *mg* na kubický kilometr (a to ve vzdálenosti až dvou kilometrů od podniku). Pracovníci úřadu provedli kontrolu na různých místech vzdálených do dvou kilometrů od podniku a v různých dnech, takže bylo zazamenáno celkem 100 měření. Průměrná koncentrace vinylchloridu vypočítaná ze všech měření byla 54 *mg*. Můžeme tvrdit na hladině významnosti $\alpha = 0{,}05$, že podnik dodržuje stanovené limity, jestliže víme, že směrodatná odchylka obsahu emitovaných látek v ovzduší je obecně při měřeních 20 *mg*? Předpokládejme, že rozdělení, z něhož náhodný výběr pochází, je normální.

Řešení

Levostranný interval spolehlivosti:

$$(d, \infty) = \left(54 - \frac{20}{\sqrt{100}} z_{1-0,05}, \infty\right) = (54 - 3{,}29, \infty) = (50{,}71, \infty)$$

$\Rightarrow$ 50 neleží v intervalu spolehlivosti, proto zamítáme nulovou hypotézu na hladině významnosti $\alpha = 0{,}05$

Jednovýběrový z-test

Testování pomocí kritického oboru

- vypočteme realizaci testovacího kritéria $t_0 = \frac{\bar{x} - c}{\frac{\sigma}{\sqrt{n}}}$, kde výraz $\frac{\sigma}{\sqrt{n}}$ označuje *směrodatnou chybu*
- stanovíme kritický obor W
- $t_0 \in W$, H_0 zamítáme na hladině významnosti α a přijímáme H_1
- v opačném případě nezamítáme nulovou hypotézu H_0 na dané hladině významnosti α

Jednovýběrový z-test

Kritické obory jednotlivých variant mají následující tvary:

- oboustranný test (testujeme H_0: $\mu = c$ proti H_1: $\mu \neq c$): kritický obor má tvar $W = (-\infty, -z_{1-\alpha/2}\rangle \cup \langle z_{1-\alpha/2}, \infty)$
- levostranný test (testujeme H_0: $\mu = c$ (resp. H_0: $\mu \geq c$) proti H_1: $\mu < c$): kritický obor má tvar $W = (-\infty, -z_{1-\alpha}\rangle$
- pravostranný test (testujeme H_0: $\mu = c$ (resp. H_0: $\mu \leq c$) proti H_1: $\mu > c$): kritický obor má tvar $W = \langle z_{1-\alpha}, \infty)$

Příklad

Úřad pro ochranu životního prostředí vydal nařízení, kterým průmyslovým podnikům určuje maximální přípustnou koncentraci vinylchloridu ve vzduchu, 50 *mg* na kubický kilometr (a to ve vzdálenosti až dvou kilometrů od podniku). Pracovníci úřadu provedli kontrolu na různých místech vzdálených do dvou kilometrů od podniku a v různých dnech, takže bylo zazamenáno celkem 100 měření. Průměrná koncentrace vinylchloridu vypočítaná ze všech měření byla 54 *mg*. Můžeme tvrdit na hladině významnosti $\alpha = 0{,}05$, že podnik dodržuje stanovené limity, jestliže víme, že směrodatná odchylka obsahu emitovaných látek v ovzduší je obecně při měřeních 20 *mg*? Předpokládejme, že rozdělení, z něhož náhodný výběr pochází, je normální.

Řešení

Testování pomocí kritického oboru

- vypočteme realizaci testovacího kritéria
  $$t_0 = \frac{\bar{x} - c}{\frac{\sigma}{\sqrt{n}}} = \frac{54 - 50}{\frac{20}{\sqrt{100}}} = 2$$
- stanovíme kritický obor
  $W = \langle z_{1-\alpha}, \infty) = \langle 1{,}644854, \infty)$
- $t_0 \in W$, H_0 zamítáme na hladině významnosti α a přijímáme H_1

$\Rightarrow$ lze konstatovat na této hladině významnosti, že podnik porušuje nařízení úřadu pro ochranu životního prostředí

Dvouvýběrový z-test

Nechť $X_{11}, \ldots, X_{1n_1}$ je náhodný výběr z rozdělení $N(\mu_1, \sigma_1^2)$ a $X_{21}, \ldots, X_{2n_2}$ je na něm nezávislý náhodný výběr z rozdělení $N(\mu_2, \sigma_2^2)$, přičemž $n_1 \geq 2, n_2 \geq 2$ a σ_1^2, σ_2^2 **známe**. Nechť c je konstanta. Test H_0: $\mu_1 - \mu_2 = c$ proti H_1: $\mu_1 - \mu_2 \neq c$ se nazývá *dvouvýběrový z-test*.

Dvouvýběrový z-test

Testování pomocí intervalů spolehlivosti
Vzorce pro meze $100(1-\alpha)\%$ intervalů spolehlivosti pro parametrické funkce $\mu_1 - \mu_2$, jestliže známe rozptyly σ_1^2, σ_2^2:

- oboustranný: $(d, h) =$
$$\left(\bar{x}_1 - \bar{x}_2 - \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} z_{1-\alpha/2}, \bar{x}_1 - \bar{x}_2 + \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} z_{1-\alpha/2}\right)$$
- levostranný:
$$(d, \infty) = \left(\bar{x}_1 - \bar{x}_2 - \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} z_{1-\alpha}, \infty\right)$$

Dvouvýběrový z-test

- pravostranný:

$$(-\infty, h) = \left(-\infty, \bar{x}_1 - \bar{x}_2 + \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} z_{1-\alpha}\right)$$

Stačí rozhodnout, zda daná hodnota c leží v daném intervalu (pokud ne, zamítáme H_0).

Dvouvýběrový z-test

Testování pomocí kritického oboru

- vypočteme realizaci testovacího kritéria
$$t_0 = \frac{(\bar{x}_1 - \bar{x}_2) - c}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$
- stanovíme kritický obor W
- $t_0 \in W$, H_0 zamítáme na hladině významnosti α a přijímáme H_1
- v opačném případě nezamítáme nulovou hypotézu H_0 na dané hladině významnosti α

Dvouvýběrový z-test

Kritické obory jednotlivých variant mají následující tvary:

- oboustranný test (testujeme H_0: $\mu_1 - \mu_2 = c$ proti H_1: $\mu_1 - \mu_2 \neq c$): kritický obor má tvar $W = (-\infty, -z_{1-\alpha/2}\rangle \cup \langle z_{1-\alpha/2}, \infty)$
- levostranný test (testujeme H_0: $\mu_1 - \mu_2 = c$ proti H_1: $\mu_1 - \mu_2 < c$): kritický obor má tvar $W = (-\infty, -z_{1-\alpha}\rangle$
- pravostranný test (testujeme H_0: $\mu_1 - \mu_2 = c$ proti H_1: $\mu_1 - \mu_2 > c$): kritický obor má tvar $W = \langle z_{1-\alpha}, \infty)$

t-test

- jednovýběrový t-test
- párový t-test
- dvouvýběrový t-test

Jednovýběrový t-test

Nechť $X_1, \ldots, X_n$ je náhodný výběr z rozdělení $N(\mu, \sigma^2)$, kde σ^2 **neznáme**. Nechť $n \geq 2$ a c je konstanta. Test H$_0$: $\mu = c$ proti H$_1$: $\mu \neq c$ se nazývá *jednovýběrový t-test*.

$\Rightarrow$ rozptyl neznáme - proto využití *Studentova t-rozdělení* pro $\nu = n - 1$ stupňů volnosti

Jednovýběrový t-test

Testování pomocí intervalů spolehlivosti
Vzorce pro meze $100(1-\alpha)\%$ intervalů spolehlivosti pro střední hodnotu μ, jestliže neznáme rozptyl σ^2:

- oboustranný: $(d, h) = \left(\bar{x} - \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1), \bar{x} + \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1)\right)$
- levostranný: $(d, \infty) = \left(\bar{x} - \frac{s}{\sqrt{n}} t_{1-\alpha}(n-1), \infty\right)$
- pravostranný: $(-\infty, h) = \left(-\infty, \bar{x} + \frac{s}{\sqrt{n}} t_{1-\alpha}(n-1)\right)$

Stačí rozhodnout, zda daná hodnota c leží v daném intervalu (pokud ne, zamítáme H_0).

Jednovýběrový t-test

Testování pomocí kritického oboru

- vypočteme realizaci testovacího kritéria $t_0 = \dfrac{\bar{x} - c}{\frac{s}{\sqrt{n}}}$
- stanovíme kritický obor W
- $t_0 \in W$, H_0 zamítáme na hladině významnosti α a přijímáme H_1
- v opačném případě nezamítáme nulovou hypotézu H_0 na dané hladině významnosti α

Jednovýběrový t-test

Kritické obory jednotlivých variant mají následující tvary:

- oboustranný test (testujeme H_0: $\mu = c$ proti H_1: $\mu \neq c$): kritický obor má tvar $W = (-\infty, -t_{1-\alpha/2}(n-1)\rangle \cup \langle t_{1-\alpha/2}(n-1), \infty)$
- levostranný test (testujeme H_0: $\mu = c$ proti H_1: $\mu < c$): kritický obor má tvar $W = (-\infty, -t_{1-\alpha}(n-1)\rangle$
- pravostranný test (testujeme H_0: $\mu = c$ proti H_1: $\mu > c$): kritický obor má tvar $W = \langle t_{1-\alpha}(n-1), \infty)$

Jednovýběrový t-test

Testování pomocí *p*-hodnoty

Význam *p*-hodnoty spočívá v tom, že nám určí nejnižší možnou hladinu významnosti, při které ještě zamítáme nulovou hypotézu H_0, tedy:

a) Je-li $p \leq \alpha$, pak zamítáme nulovou hypotézu H_0 na hladině významnosti α.
b) Je-li $p > \alpha$, pak nezamítáme nulovou hypotézu H_0 na hladině významnosti α.

	Test průměrů vůči referenční konstantě (hodnotě) (Volební účast ORP Jablunkov)							
Proměnná	Průměr	Sm.odch.	N	Sm.chyba	Referenční konstanta	t	SV	p
Účast u voleb 2010	55,73364	2,755733	11	0,830885	62,60000	-8,26392	10	0,000009

Obrázek: Tabulka výsledků jednovýběrového t-testu s vyznačenou *p*-hodnotou

Příklad

Data obsahují procentuální hodnoty účastí voličů u parlamentních voleb v roce 2010 v jednotlivých obcích ORP Jablunkov. Testujte hypotézu na hladině významnosti $\alpha = 0,05$, že voliči v ORP Jablunkov jsou méně zodpovědnými občany než voliči v celé České republice (za úroveň zodpovědnosti voličů vezměte procentuální účast u voleb v jednotlivých obcích a srovnejte tuto hodnotu s celorepublikovým průměrem, který činil $62,6\,\%$).

Příklad

Název obce	Účast u parlamentních voleb 2010 (%)
Bocanovice	55, 26
Bukovec	60, 43
Dolní Lomná	55, 68
Horní Lomná	52, 32
Hrádek	56, 66
Hrčava	51, 23
Jablunkov	55, 62
Milíkov	59, 26
Mosty u Jablunkova	53, 76
Návsí	57, 76
Písek	55, 09

Řešení

- hypotéza H_0, kterou budeme testovat, bude tvaru H_0: $\mu = 62,6$ proti alternativní hypotéze H_1: $\mu < 62,6$
- levostranná varianta t-testu
- první možnost je vypočítat si výběrový průměr a směrodatnou odchylku a hodnoty zadat do testu rozdílů, který získáme takto: *Statistiky – Základní statistiky/tabulky – Testy rozdílů: r, %, průměry – OK*; uprostřed máme volbu *Rozdíly mezi dvěma průměry (normální rozdělení)*, ve které vypíšeme hodnoty *Pr1: x̄, SmOd1: s, N1: n, Pr2: c* a zaškrtneme *Jednostr.* a *Výběrový průměr vs. střední hodnota*
- druhá cesta vedoucí k výsledku je následující: *Statistiky – Základní statistiky/tabulky – t-test, samost. vzorek – OK*

Párový t-test

- párový test se obecně vyznačuje vzájemnou spojitostí obou výběrů
- oba výběry spolu nejenom souvisí, ale mají také stejný rozsah
- náhodné veličiny proto můžeme zapsat takto: $(Y_1, Z_1), \ldots, (Y_n, Z_n)$, kde $n \geq 2$
- předmětem našeho zájmu je zjistit, zda rozdíl středních hodnot je roven nějaké konstantě, resp. porovnat tento rozdíl s nějakou hodnotou

Párový t-test

- lze vytvořit tzv. *rozdílový náhodný výběr* $X_1 = Y_1 - Z_1, \ldots, X_n = Y_n - Z_n$
- střední hodnota μ tohoto rozdílového náhodného výběru je rovna rozdílu středních hodnot veličin Y a Z, tj. $\mu = \mu_1 - \mu_2$

$\Rightarrow$ redukce na jeden výběr - použití vzorce pro jednovýběrový t-test

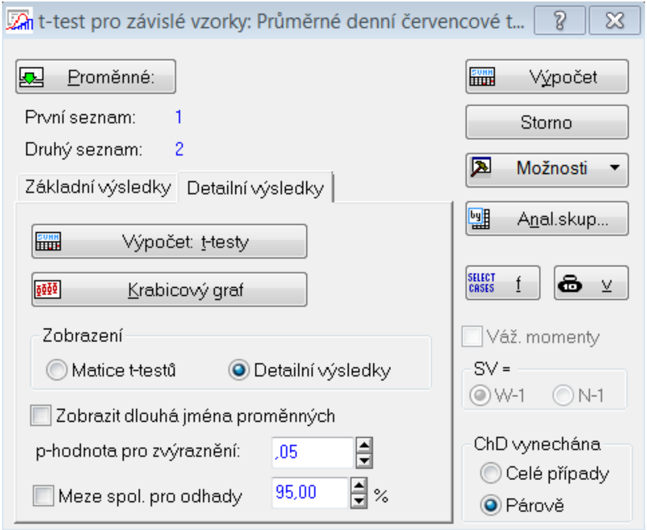
Příklad

Máme k dispozici data obsahující průměrné denní teploty v červenci 2010 a 2011 na stanici Botanická zahrada (Brno). Pokuste se zjistit, zda existuje rozdíl mezi teplotami naměřenými v červenci roku 2010 a v červenci roku 2011. Nulovou hypotézu o shodě středních hodnot průměrných denních teplot testujte na hladině významnosti $\alpha = 0,05$.

Řešení

- teploty měřeny na stejné stanici a rozsah obou výběrů je stejný ($n = 31$) $\Rightarrow$ párový test
- rozdílový náhodný výběr - otestujeme normalitu - použijeme t-test
- *Statistiky – Základní statistiky/tabulky – t-test, závislé vzorky – OK*
- na kartě *Detailní výsledky* máme možnost upravit možné zvýraznění *p*-hodnoty i určit meze intervalu spolehlivosti

Řešení


t-test pro závislé vzorky: Průměrné denní červencové t...
Proměnné:
První seznam: 1
Druhý seznam: 2
Základní výsledky
Detailní výsledky
Výpočet: t-testy
Krabicový graf
Zobrazení
Matice t-testů
Detailní výsledky
Zobrazit dlouhá jména proměnných
p-hodnota pro zvýraznění: ,05
Meze spol. pro odhady 95,00 %
Výpočet
Storno
Možnosti
Anal.skup...
SELECT CASES
Váž. momenty
SV =
W-1
N-1
ChD vynechána
Celé případy
Párově

Dvouvýběrový t-test

Nechť $X_{11}, \ldots, X_{1n_1}$ je náhodný výběr z rozdělení $N(\mu_1, \sigma^2)$ a $X_{21}, \ldots, X_{2n_2}$ je na něm nezávislý náhodný výběr z rozdělení $N(\mu_2, \sigma^2)$, přičemž $n_1 \geq 2, n_2 \geq 2$ a σ^2 **neznáme**. Nechť c je konstanta. Test H_0: $\mu_1 - \mu_2 = c$ proti H_1: $\mu_1 - \mu_2 \neq c$ se nazývá *dvouvýběrový t-test*.

Dvouvýběrový t-test

- nejpoužívanější parametrický test
- již z definice ovšem plynou omezení použití tohoto testu - výběry pochází z normálních rozdělení, jsou nezávislé a ačkoliv neznáme σ^2, víme, že je shodný

$\Rightarrow$ shodu rozptylů ověřujeme **F-testem**

Dvouvýběrový t-test

Testování pomocí intervalů spolehlivosti

Vzorce pro meze $100(1-\alpha)\%$ intervalů spolehlivosti pro parametrické funkce $\mu_1 - \mu_2$, jestliže neznáme rozptyly σ_1^2, σ_2^2, ale víme, že jsou shodné:

- oboustranný:

$$(d, h) = \left(\bar{x}_1 - \bar{x}_2 - s_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{1-\alpha/2}(n_1 + n_2 - 2),\right.$$

$$\left.\bar{x}_1 - \bar{x}_2 + s_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{1-\alpha/2}(n_1 + n_2 - 2)\right)$$

- levostranný: $(d, \infty) =$

$$\left(\bar{x}_1 - \bar{x}_2 - s_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{1-\alpha}(n_1 + n_2 - 2), \infty\right)$$

Dvouvýběrový t-test

- pravostranný: $(-\infty, h) = \left(-\infty, \bar{x}_1 - \bar{x}_2 + s_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{1-\alpha}(n_1 + n_2 - 2)\right)$

Stačí rozhodnout, zda daná hodnota c leží v daném intervalu (pokud ne, zamítáme H_0).

Dvouvýběrový t-test

V intervalech spolehlivosti se vyskytl neznámý symbol s_*, který vychází z již dříve zavedeného váženého průměru výběrových rozptylů a který tudíž vypočítáme takto:

$$s_* = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

Dvouvýběrový t-test

Testování pomocí kritického oboru

- vypočteme realizaci testovacího kritéria
$$t_0 = \frac{(\bar{x}_1 - \bar{x}_2) - c}{s_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$
- stanovíme kritický obor W
- $t_0 \in W$, H_0 zamítáme na hladině významnosti α a přijímáme H_1
- v opačném případě nezamítáme nulovou hypotézu H_0 na dané hladině významnosti α

Dvouvýběrový t-test

Kritické obory jednotlivých variant mají následující tvary:

- oboustranný test (testujeme H$_0$: $\mu_1 - \mu_2 = c$ proti H$_1$: $\mu_1 - \mu_2 \neq c$): kritický obor má tvar $W = (-\infty, -t_{1-\alpha/2}(n_1 + n_2 - 2)\rangle \cup \langle t_{1-\alpha/2}(n_1 + n_2 - 2), \infty)$
- levostranný test (testujeme H$_0$: $\mu_1 - \mu_2 = c$ proti H$_1$: $\mu_1 - \mu_2 < c$): kritický obor má tvar $W = (-\infty, -t_{1-\alpha}(n_1 + n_2 - 2)\rangle$
- pravostranný test (testujeme H$_0$: $\mu_1 - \mu_2 = c$ proti H$_1$: $\mu_1 - \mu_2 > c$): kritický obor má tvar $W = \langle t_{1-\alpha}(n_1 + n_2 - 2), \infty)$

Dvouvýběrový t-test

Testování pomocí *p*-hodnoty

Význam *p*-hodnoty spočívá v tom, že nám určí nejnižší možnou hladinu významnosti, při které ještě zamítáme nulovou hypotézu H_0, tedy:

a) Je-li $p \leq \alpha$, pak zamítáme nulovou hypotézu H_0 na hladině významnosti α.

b) Je-li $p > \alpha$, pak nezamítáme nulovou hypotézu H_0 na hladině významnosti α.

F-test

Nechť $X_{11}, \ldots, X_{1n_1}$ je náhodný výběr z rozdělení $N(\mu_1, \sigma_1^2)$ a $X_{21}, \ldots, X_{2n_2}$ je na něm nezávislý náhodný výběr z rozdělení $N(\mu_2, \sigma_2^2)$, přičemž $n_1 \geq 2, n_2 \geq 2$ a σ_1^2, σ_2^2 **neznáme**. Test H_0: $\frac{\sigma_1^2}{\sigma_2^2} = 1$ proti H_1: $\frac{\sigma_1^2}{\sigma_2^2} \neq 1$ se nazývá *F-test*.

$\Rightarrow$ využívat budeme Fisher-Snedecorova rozdělení pro $(\nu_1, \nu_2) = (n_1 - 1, n_2 - 1)$ stupňů volnosti

F-test

Testování pomocí intervalů spolehlivosti

Vzorce pro meze $100(1-\alpha)\%$ intervalů spolehlivosti pro podíl rozptylů $\dfrac{\sigma_1^2}{\sigma_2^2}$:

- oboustranný: $(d, h) = \left(\dfrac{\frac{s_1^2}{s_2^2}}{F_{1-\alpha/2}(n_1-1, n_2-1)}, \dfrac{\frac{s_1^2}{s_2^2}}{F_{\alpha/2}(n_1-1, n_2-1)} \right)$
- levostranný: $(d, \infty) = \left(\dfrac{\frac{s_1^2}{s_2^2}}{F_{1-\alpha}(n_1-1, n_2-1)}, \infty \right)$

F-test

- pravostranný: $(0, h) = \left(0, \dfrac{\frac{s_1^2}{s_2^2}}{F_\alpha(n_1 - 1, n_2 - 1)}\right)$

Stačí rozhodnout, zda číslo 1 leží v daném intervalu (pokud ne, zamítáme H_0).

F-test

Testování pomocí kritického oboru

- vypočteme realizaci testovacího kritéria $t_0 = \frac{s_1^2}{s_2^2}$
- stanovíme kritický obor W
- $t_0 \in W$, H_0 zamítáme na hladině významnosti α a přijímáme H_1
- v opačném případě nezamítáme nulovou hypotézu H_0 na dané hladině významnosti α

F-test

Kritické obory jednotlivých variant mají následující tvary:

- oboustranný test (testujeme H$_0$: $\frac{\sigma_1^2}{\sigma_2^2} = 1$ proti H$_1$: $\frac{\sigma_1^2}{\sigma_2^2} \neq 1$): kritický obor má tvar $W = \langle 0, F_{\alpha/2}(n_1 - 1, n_2 - 1)\rangle \cup \langle F_{1-\alpha/2}(n_1 - 1, n_2 - 1), \infty)$
- levostranný test (testujeme H$_0$: $\frac{\sigma_1^2}{\sigma_2^2} = 1$ proti H$_1$: $\frac{\sigma_1^2}{\sigma_2^2} < 1$): kritický obor má tvar $W = \langle 0, F_{\alpha}(n_1 - 1, n_2 - 1)\rangle$
- pravostranný test (testujeme H$_0$: $\frac{\sigma_1^2}{\sigma_2^2} = 1$ proti H$_1$:

F-test

Testování pomocí *p*-hodnoty

Význam *p*-hodnoty spočívá v tom, že nám určí nejnižší možnou hladinu významnosti, při které ještě zamítáme nulovou hypotézu H_0, tedy:

a) Je-li $p \leq \alpha$, pak zamítáme nulovou hypotézu H_0 na hladině významnosti α.

b) Je-li $p > \alpha$, pak nezamítáme nulovou hypotézu H_0 na hladině významnosti α.

	t-testy; grupováno:Grupovací proměnná - okres (Podíl věřících) Skup. 1: 1 Skup. 2: 2										
Proměnná	Průměr 1	Průměr 2	t	sv	p	Poč.plat 1	Poč.plat. 2	Sm.odch. 1	Sm.odch. 2	F-poměr Rozptyly	p Rozptyly
Podíl katolíků na celkovém počtu obyvatel (%)	46,10702	67,19388	-7,45369	162	0,000000	77	87	16,54307	19,33868	1,366537	0,165620

Obrázek: Tabulka výsledků dvouvýběrového t-testu společně s F-testem s vyznačenými *p*-hodnotami

Příklad

Zajímavá data nám poskytuje také sčítání lidu z roku 2001. Z dat věnujících se náboženskému vyznání se zaměříme na počty římských katolíků v jednotlivých obcích okresů Frýdek-Místek a Zlín. Na hladině významnosti $0,05$ testujte hypotézu, že se střední hodnoty podílů počtu katolíků na celkovém počtu obyvatel jednotlivých obcí za okresy Frýdek-Místek a Zlín neliší. (Při ověřování normality dat zjistíte mírné porušení, přesto použijte k ověření parametrický test.)

Řešení

- nový datový soubor, do kterého zkopírujeme hodnoty celkového počtu obyvatel a počtu římských katolíků za jednotlivé obce obou okresů
- hodnoty vložíme pod sebe a vytvořením nové proměnné určíme pomocí čísel 0 a 1 příslušnost k okresům Frýdek-Místek (0) a Zlín (1) - *grupovací proměnná*
- do dlouhého jména nové proměnné vložíme vzorec k určení procentuálních hodnot výsledků
- *Statistiky – Základní statistiky/tabulky – t-test, nezávislé, dle skupin*[1] *– OK*

[1] Kdybychom neměli data vložená pod sebou odlišená grupovací proměnnou, ale vedle sebe, použili bychom *t-test, nezávislé, dle proměn.*

Řešení

- zadáme proměnnou (podíl katolíků na celkovém počtu obyvatel) a grupovací proměnnou (hodnoty 1 a 2)
- na kartě *Možnosti* příležitost k volbě konkrétní hladiny významnosti (označeno symbolem *p*)
- varianta dvouvýběrového t-testu se separovanými proměnnými, kterou užíváme, pokud se rozptyly obou výběrů liší a nechceme použít neparametrickou variantu tohoto testu
- na kartě *Detailní výsledky* si můžeme nechat vykreslit různé grafy, včetně N-P plotu a box plotu

Řešení - interpretace testu

	t-testy; grupováno:Grupovací proměnná - okres (Podíl věřících) Skup. 1: 1 Skup. 2: 2										
Proměnná	Průměr 1	Průměr 2	t	sv	p	Poč.plat 1	Poč.plat. 2	Sm.odch. 1	Sm.odch. 2	F-poměr Rozptyly	p Rozptyly
Podíl katolíků na celkovém počtu obyvatel (%)	46,10702	67,19388	-7,45369	162	0,000000	77	87	16,54307	19,33868	1,366537	0,165620

- zkontrolovat *p*-hodnotu F-testu, kterou nám systém automaticky poskytnul pod názvem *p Rozptyly* - musí být větší než zvolená hladina významnosti
- *p*-hodnota je rovna $0,000000$, proto zamítáme hypotézu o shodě středních hodnot a můžeme konstatovat, že na hladině významnosti $\alpha = 0,05$ se střední hodnoty podílů římských katolíků jednotlivých obcí v těchto okresech liší

Neparametrické testy - vlastnosti

- využíváme v případech, kdy neznáme rozdělení základního souboru, z něhož daný náhodný výběr pochází
- v případech, kdy náhodný výběr nepochází z normálního rozdělení
- nebo v případě, kdy zkoumáme data ordinálního typu

- mají obecně menší sílu, tj. nepravdivou hypotézu zamítají s menší pravděpodobností než testy parametrické

Neparametrické testy - vlastnosti

Většina zde uvedených testů patří do skupiny *pořadových*, tzn. že při provádění testů budeme určovat pořadí hodnot, které bude sloužit k výpočtu testovacího kritéria namísto původních hodnot.

Jestliže máme dána reálná čísla $x_1, x_2, \ldots, x_n$, pak *pořadím* čísla x_i nazveme počet těch čísel $x_1, x_2, \ldots, x_n$, která jsou menší nebo rovna číslu x_i. V případě rovnosti hodnot několika čísel zavádíme *průměrné pořadí*, které je rovno aritmetickému průměru pořadí čísel, jejichž hodnoty jsou si rovny.

Neparametrické testy

Vybrané neparametrické testy, jimiž se budeme zabývat:

- Wilcoxonův test
- K-S test
- Kruskalův-Wallisův test (K-W test) a mediánový test

Wilcoxonův test

- **neparametrický ekvivalent t-testu**
- budeme předpokládat spojitost rozdělení, z něhož daný náhodný výběr pochází
- data aspoň ordinálního typu, tj. ordinální, intervalové nebo poměrové
- uvedeme pouze variantu testování pomocí kritického oboru pro všechny typy tohoto testu (jednovýběrový, párový i dvouvýběrový)

Jednovýběrový Wilcoxonův test

Nechť $X_1, X_2, \ldots, X_n$ je náhodný výběr ze spojitého rozdělení o rozsahu n. Budeme testovat hypotézu H_0: $Med(x) = c$, kde c je reálná konstanta, proti jedné z těchto variant:

- oboustranná alternativa H_1: $Med(x) \neq c$
- levostranná alternativa H_1: $Med(x) < c$
- pravostranná alternativa H_1: $Med(x) > c$

Jednovýběrový Wilcoxonův test

Postup provedení testu

a) Nejprve utvoříme rozdíly $Y_i = X_i - c$ (pro $i = 1, 2, \dots, n$). Nastane-li případ, že jsou některé rozdíly rovny nule, pak je vyloučíme a za n již bereme pouze počet nenulových hodnot.

b) Absolutní hodnoty rozdílů, tedy $|Y_i|$, uspořádáme vzestupně podle velikosti a určíme jejich pořadí, resp. průměrné pořadí, pokud jsou si některé hodnoty $|Y_i|$ rovny.

c) Zavedeme důležité označení statistik:
- S_W^+ – označuje součet pořadí přes kladné hodnoty Y_i
- S_W^- – označovat součet pořadí přes záporné hodnoty Y_i

Přitom musí platí vztah: $S_W^+ + S_W^- = \dfrac{n(n+1)}{2}$.

Jednovýběrový Wilcoxonův test

d) Až po tuto chvíli byl postup testování oboustranné hypotézy i jednostranných hypotéz stejný. Nyní je potřeba rozlišit chování testových statistik pro jednotlivé případy, proto:
   - pro oboustrannou hypotézu se testová statistika rovná menší z hodnot S_W^+, S_W^-
   - pro levostrannou hypotézu se testová statistika rovná S_W^+
   - pro pravostrannou hypotézu se testová statistika rovná S_W^-

e) Nulovou hypotézu H_0 zamítáme na zvolené hladině významnosti α, když je testová statistika menší nebo rovna kritické hodnotě, kterou nalezneme v tabulkách (tato hodnota závisí na rozsahu n a hladině významnosti α).

Párový Wilcoxonův test

- postup jako u párového t-testu (převod na jednovýběrový test)
- vytvoříme *rozdílový náhodný výběr* $X_1 = Y_1 - Z_1, \ldots, X_n = Y_n - Z_n$
- testujeme hypotézu H_0: $Med(y) - Med(z) = c$, proti alternativní hypotéze H_1: $Med(y) - Med(z) \neq c$, resp. proti jednostranným alternativám

Příklad

Hodnota indexu stáří[a] pro celou Českou republiku na konci roku 2010 byla rovna $107,8$. Na hladině významnosti $\alpha = 0,1$ testujte hypotézu, že medián hodnot indexů stáří pro okresy Středočeského kraje se neliší od hodnoty $107,8$. Zjištěná data jsou uvedena v tabulce.

[a] Počet 65 letých a starších na 100 obyvatel ve věku 0 – 14.

Příklad

Název okresu	Index stáří (%)
Benešov	105,7
Beroun	98,6
Kladno	101,4
Kolín	107,6
Kutná Hora	119,2
Mělník	96,8
Mladá Boleslav	97,9
Nymburk	96,1
Praha-východ	71,5
Praha-západ	67,2
Příbram	108,1
Rakovník	108,6

Řešení

- po ověření normality zjistíme, že nelze použít parametrický test
- jednovýběrový Wilcoxonův test - není přímo implementován v systému STATISTICA
- obejdeme situaci tím, že použijeme párový Wilcoxonův test
- *Statistiky – Neparametrická statistika – Porovnání dvou závislých vzorků (proměnné) – OK*

	Wilcoxonův párový test (Index stáří Středočeský kraj) Označené testy jsou významné na hladině p <,10000			
Dvojice proměnných	**Počet platných**	T	Z	p-hodn.
Index stáří & Index stáří ČR	12	14,00000	1,961161	0,049861

Řešení

Nedopustili jsme se chyby?

- systém STATISTICA pracuje s asymptotickou variantou testu - my jsme však měli pouze dvanáct pozorování
- je zapotřebí ověřit naše počínání - můžeme však použít vypočtenou hodnotu testovacího kritéria (označeno písmenem $T = 14$)
- z tabulek zjistíme kritickou hodnotu: 13
- na hladině významnosti α zamítáme nulovou hypotézu o shodě mediánu hodnot indexů stáří a konstanty $107,8$

Dvouvýběrový Wilcoxonův test

- v programu STATISTICA pod názvem *test Mann-Whitney*
- obdobně jako dvouvýběrový t-test je nejpoužívanějším neparametrickým testem

Nechť $X_1, \ldots, X_n$ a $Y_1, \ldots, Y_m$ (přičemž rozsahy n a m jsou obecně různé) jsou dva nezávislé náhodné výběry ze dvou spojitých rozdělení, jejichž distribuční funkce se mohou lišit pouze posunutím = musí „vypadat stejně“. Označme $Med(x)$, resp. $Med(y)$ medián prvního, resp. druhého rozdělení. Testujeme hypotézu, že mediány těchto dvou rozdělení jsou shodné proti alternativě, že jsou rozdílné, tj.: H$_0$: $Med(x) - Med(y) = 0$ proti H$_1$: $Med(x) - Med(y) \neq 0$.

Test Mann-Whitney

Postup provedení testu

a) Všechny hodnoty uspořádáme vzestupně podle velikosti ($n + m$ hodnot).

b) Sečteme pořadí všech hodnot $X_1, \ldots, X_n$ a označíme $\sum R_1$. Obdobně sečteme pořadí všech hodnot $Y_1, \ldots, Y_m$ a označíme $\sum R_2$.

c) Vypočítáme testová statistiky U_1 a U_2 podle následujících vzorců, přičemž platí vztah $U_1 + U_2 = mn$:

$$U_1 = nm + \frac{n(n+1)}{2} - \sum R_1$$

$$U_2 = nm + \frac{m(m+1)}{2} - \sum R_2$$

Test Mann-Whitney

d) Pokud je menší z hodnot U_1, U_2 menší nebo rovna než kritická hodnota, kterou nalezneme v tabulkách (v závislosti na rozsazích výběrů a hladině významnosti α), pak nulovou hypotézu o shodě mediánů obou rozdělení zamítáme na dané hladině významnosti α a přijímáme alternativní hypotézu.

Příklad

Ze souboru *Průmysl 1987* vyberte podniky v okresu Karviná a Most (jak víte, oba okresy jsou těžebně zaměřeny) a vytvořte ukazatel spočívající v podílu celkového obratu podniku (podle klasifikace CZ-NACE spadajících do kategorie 10) a počtu všech zaměstnanců. Existuje rozdíl mezi mediány „výnosností“ podniků zaměřených na těžbu v obou regionech? Testujte tuto hypotézu na hladině významnosti $\alpha = 0,05$.

Řešení

- grafické i početní metody ověřování normality dat svědčí v tomto případě o porušení tohoto předpokladu pro použití parametrických metod
- otevřeme si připravený soubor, kde proměnná *Rozdělení podniků podle okresů* je tzv. grupovací proměnnou
- *Statistiky – Neparametrická statistika – Porovnání dvou nezávislých vzorků (skupiny) – OK*
- zvolíme proměnné (proměnná *Obrat přepočtený na jednoho pracovníka* bude závisle proměnnou; grupovací proměnnou jsme již zmínili výše) a hladinu významnosti testu
- zvolíme ikonu *Mann-Whitneyův U test* (použít lze i dvouvýběrový K-S test)

Řešení - interpretace testu

Mann-Whitneyův U test (Průmyslové podniky)
Dle proměn. Rozdělení podniků podle okresu
Označené testy jsou významné na hladině p <,05000

Proměnná	Sčt poř. skup. 1	Sčt poř. skup. 2	U	Z	p-hodn.	Z upravené	p-hodn.	N platn. skup. 1	N platn. skup. 2	2*1str. přesné p
Obrat přepočtený na jednoho pracovníka	169,0000	131,0000	49,00000	-1,07331	0,283132	-1,07331	0,283132	15	9	0,289503

- vystupují zde dvě *p*-hodnoty
- pod symbolem U vystupuje hodnota testové statistiky, což je menší z hodnot U_1, U_2
- symbol Z označuje asymptotickou hodnotu testové statistiky, pro kterou vystupuje *p*-hodnota pod názvem *Úroveň p*
- pro rozsahy výběrů pod 30 porovnáváme s hladinou významnosti α přesnou *p*-hodnotu (2*1 str. přesné p), která je určena pro testovou statistiku U
- *p*-hodnota $= 0,289503$, proto na hladině významnosti $\alpha = 0,05$ nezamítáme nulovou hypotézu o shodě mediánů obou výběrů

K-S test

- použití nejenom pro testování normality dat
- v případě, kdy nelze použít test Mann-Whitney (distribuční funkce se neliší pouze posunutím)
- $X_1, \ldots, X_n$ a $Y_1, \ldots, Y_m$ jsou dva nezávislé náhodné výběry ze dvou spojitých rozdělení - testujeme hypotézu, že distribuční funkce rozdělení, z nichž náhodné výběry pocházejí, jsou shodné
- hodnota testové statistiky udává největší absolutní rozdíl mezi hodnotami výběrových distribučních funkcí pro jakékoliv reálné x
- čím větší hodnota, tím větší pravděpodobnost zamítnutí H_0

K-W test a mediánový test

- oba testy patří mezi vícevýběrové (aspoň dva nezávislé náhodné výběry o obecně různých rozsazích) neparametrické testy
- neparametrická obdoba parametrických testů založených na analýze rozptylu jednoduchého, resp. dvojného třídění
- K-W test je nepatrně silnější

K-W test a mediánový test

Nechť je dáno $m \geq 2$ nezávislých náhodných výběrů o rozsazích $n_1, n_2, \ldots, n_m$. Předpokládejme, že tyto výběry pocházejí ze spojitých rozdělení a označme jejich celkový rozsah $n = n_1 + n_2 + \cdots + n_m$. Na dané hladině významnosti α testujeme hypotézu, že všechny tyto náhodné výběry pocházejí z téhož rozdělení.

- K-W test patří mezi pořadové testy
- mediánový test pracuje s mediánem určeným ze všech n hodnot a testová statistika je pak založena na počtu hodnot jednotlivých výběrů, které jsou větší nebo rovny mediánu
- testové statistiky těchto testů se řídí rozdělením $\chi^2(m-1)$, když H_0 platí

Děkuji za pozornost...