VYBRANÉ DVOUVÝBĚROVÉ TESTY

Martina Litschmannová

Obsah přednášky – Vybrané dvouvýběrové testy par. hypotéz

- test o shodě rozptylů (F-test),
- testy o shodě středních hodnot (t-test, Aspinové-Welchův test),
- test o shodě mediánů (Mannův-Whitneyův test),
- test o shodě parametrů dvou binomických rozdělení (test homogenity dvou binomických rozdělení),

- párové testy (párový t-test, párový znaménkový test)

Test o shodě rozptylů (F-test, test homoskedasticity)

H_0: $\sigma_X^2 = \sigma_Y^2$, H_A: $\sigma_X^2 \neq \sigma_Y^2$ (resp. $\sigma_X^2 < \sigma_Y^2$, $\sigma_X^2 > \sigma_Y^2$)

Předpoklady testu:

Mějme dva nezávislé výběry $X_1, X_2, \dots, X_{n_1}$ a $Y_1, Y_2, \dots, Y_{n_2}$, které pocházejí z populací, které mají rozdělení $N(\mu_X; \sigma_X^2)$, resp. $N(\mu_Y; \sigma_Y^2)$.

Testová statistika: $T(\boldsymbol{X}, \boldsymbol{Y}) = \dfrac{\frac{S_X^2}{\sigma_X^2}}{\frac{S_Y^2}{\sigma_Y^2}}$

Nulové rozdělení: Fisherovo-Snedecorovo rozdělení s $n_1 - 1$ stupni volnosti pro čitatele a $n_2 - 1$ stupni volnosti pro jmenovatele

Testy o shodě středních hodnot - dvouvýběrový *t*-test

H_0: $\mu_X = \mu_Y$, H_A: $\mu_X \neq \mu_Y$ (resp. $\mu_X < \mu_Y$, $\mu_X > \mu_Y$)

Předpoklady testu:

Mějme dva nezávislé výběry $X_1, X_2, \dots, X_{n_1}$ a $Y_1, Y_2, \dots, Y_{n_2}$, které pochází z populace mající opět rozdělení $N(\mu_X; \sigma_X^2)$, resp. $N(\mu_Y; \sigma_Y^2)$, kde $\sigma_X^2 = \sigma_Y^2$.

Testová statistika: $T(\boldsymbol{X}, \boldsymbol{Y}) = \dfrac{(\bar{X}-\bar{Y})-(\mu_X-\mu_Y)}{\sqrt{\frac{(n_1-1)s_X^2+(n_2-1)s_Y^2}{n_1+n_2-2}} \cdot \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$

Nulové rozdělení: Studentovo rozdělení s $\nu = n_1 + n_2 - 2$ stupni volnosti

Testy o shodě středních hodnot - Aspinové-Welchův test

$\boldsymbol{H_0}$: $\mu_X = \mu_Y$, $\boldsymbol{H_A}$: $\mu_X \neq \mu_Y$ (resp. $\mu_X < \mu_Y, \mu_X > \mu_Y$)

Předpoklady testu:

Mějme dva nezávislé výběry $X_1, X_2, \dots, X_{n_1}$ a $Y_1, Y_2, \dots, Y_{n_2}$, které pochází z populace mající opět rozdělení $N(\mu_X; \sigma_X^2)$, resp. $N(\mu_Y; \sigma_Y^2)$, kde $\sigma_X^2 \neq \sigma_Y^2$.

Testová statistika: $T(\boldsymbol{X}, \boldsymbol{Y}) = \frac{(\bar{X}-\bar{Y})-(\mu_X-\mu_Y)}{\sqrt{\frac{s_X^2}{n_1}+\frac{s_Y^2}{n_2}}}$

Nulové rozdělení: Studentovo rozdělení s ν stupni volnosti, kde $\nu \cong \frac{\left(\frac{s_X^2}{n_1}+\frac{s_Y^2}{n_2}\right)^2}{\frac{1}{n_1-1}\left(\frac{s_X^2}{n_1}\right)^2+\frac{1}{n_2-1}\left(\frac{s_Y^2}{n_2}\right)^2}$.

Neparametrický test o shodě stř. hodnot – test shody mediánů Mannův-Whitneyův test

H_0: $x_{0,5_X} = x_{0,5_Y}$, H_A: $x_{0,5_X} \neq x_{0,5_Y}$(resp. $x_{0,5_X} < x_{0,5_Y}$, $x_{0,5_X} > x_{0,5_Y}$)

Předpoklady testu:

Nechť $X_1, X_2, \dots, X_{n_1}$ a $Y_1, Y_2, \dots, Y_{n_2}$ jsou dva nezávislé výběry ze spojitých rozdělení se stejným rozptylem a tvarem.

Postup testování: viz Úvod do statistiky, str. 193-194

Test o shodě parametrů dvou binomických rozdělení

H_0: $\pi_1 = \pi_2$, H_A: $\pi_1 \neq \pi_2$ (resp. $\pi_1 < \pi_2, \pi_1 > \pi_2$)

Předpoklady testu:

$\boldsymbol{X}$ a $\boldsymbol{Y}$ jsou náhodné výběry z alternativního rozdělení. Pro provedení tohoto testu musíme mít k dispozici výběry o dostatečném rozsahu n_1, resp. n_2. Rozsahy jednotlivých výběrů lze považovat za dostatečné, pokud jsou splněny podmínky:

$n_1 > \frac{9}{p_1(1-p_1)}$ a $n_2 > \frac{9}{p_2(1-p_2)}$.

Testová statistika: $T(\boldsymbol{X},\boldsymbol{Y}) = \frac{(p_1-p_2)-(\pi_1-\pi_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1}+\frac{p_2(1-p_2)}{n_2}}}$

Nulové rozdělení: normované normální

Párové testy

Jsou-li výsledkem zjišťování dvojice náhodných veličin $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$, které tvoří páry závislých pozorování (jde o veličiny zjišťované na stejné statistické jednotce), musíme při ověřování shody polohy přistoupit k párovým testům.

Definujme soubor rozdílů (diferencí)

$$\boldsymbol{D} = (D_1, D_2, \ldots, D_n), \quad \text{kde } D_i = X_i - Y_i.$$

Lze předpokládat, že náhodné veličiny $(D_1, D_2, \ldots, D_n)$ jsou nezávislé a že mají stejné rozdělení se střední hodnotou $\mu = \mu_1 - \mu_2$. Test o shodě dvou středních hodnot prováděný na základě dvou závislých výběrů můžeme převést na jednovýběrový test o střední hodnotě aplikovaný na soubor diferencí (rozdílů) $\boldsymbol{D}$.

<table>
<tr><th>Typ proměnné</th><th>Požadovaný typ analýzy</th><th colspan="2">Předpoklady</th><th>Testy, resp. intervalové odhady</th></tr>
<tr><td rowspan="9">Dvě nezávislé spojité proměnné</td><td rowspan="3">Ověření shody rozptylů (homoskedasticity)</td><td colspan="2" rowspan="2">Normalita</td><td>F-test
(test shody rozptylů)</td></tr>
<tr><td>Intervalový odhad poměru rozptylů, resp. směr. odchylek</td></tr>
<tr><td colspan="2">—</td><td>Leveneův test</td></tr>
<tr><td rowspan="6">Ověření shody měr polohy (středních hodnot, resp. mediánů)</td><td rowspan="4">Normalita</td><td rowspan="2">Shoda rozptylů (homoskedasticita)</td><td>Dvouvýběrový Studentův t-test
(test shody stř. hodnot)</td></tr>
<tr><td>Intervalový odhad rozdílu stř.hodnot</td></tr>
<tr><td rowspan="2">Různé rozptyly (heteroskedasticita)</td><td>Aspinové-Welchův test
(test shody stř. hodnot)</td></tr>
<tr><td>Intervalový odhad rozdílu stř.hodnot</td></tr>
<tr><td colspan="2">—</td><td>Mannův-Whitneyův test
test shody mediánů</td></tr>
<tr><td colspan="3" style="display:none"></td></tr>
<tr><td rowspan="4">Párová (spojitá) data</td><td rowspan="4">Ověření shody úrovně párových dat</td><td colspan="2" rowspan="2">Normalita</td><td>Párový studentův t-test</td></tr>
<tr><td>Intervalový odhad střední hodnoty rozdílů</td></tr>
<tr><td colspan="2">Výběry většího rozsahu</td><td>Párový znaménkový test</td></tr>
<tr><td colspan="2">Symetrické rozdělení</td><td>Wilcoxnův párový test</td></tr>
<tr><td rowspan="2">Dvě dichotomické proměnné</td><td rowspan="2">Ověření shody pravděpodobností</td><td colspan="2" rowspan="2">$n_i > \frac{9}{p_i(1-p_i)}, i = 1,2$</td><td>Test homogenity dvou binomických rozdělení</td></tr>
<tr><td>Intervalový odhad rozdílu parametru binomických rozdělení</td></tr>
</table>

DĚKUJI ZA POZORNOST!