

Systematizace statistických metod

0. Typ observovaných znaků = úroveň měření

A) Kvalitativní = nominální znaky - bez nároku na kvantifikaci

A1) alternativní = dichotomické (2 kategorie)

{muž, žena}, {terapie A, terapie B}, {správně, chybně}, {ano, ne}

A2) multinomické (více než 2 kategorie)

{dgnA, dgnB, dgnC}, {FF, PřF, MFF, LF, FSV}, {ODS, SD, SZ, KDU-ČSL, KSČM}

B) Kvantitativní znaky

B1) ordinální = pořadové

B11) posuzovací škály, rating (úzké = malé rozpětí, 3 – 10 bodové škály)

{školní známka}, {ratingová škála: ++, +, 0, -, --}, {sémantický diferenciál}

B12) posuzovací škály (širší)

{grafická škála}, {subjektivní procentová škála}

B13) lineární kombinace (většinou suma) položkových odpovědí dotazníku

B2) kardinální - desky sametěříš, nejen nějaké dotazníky

B21) čítání

{počet chyb}, {tepová frekvence}, {hrubý skór testu}

B22) pseudointervalová škála

{IQ}, {plošně standardizované skóry testů resp. dotazníků → $N(\mu, \sigma^2)$ }

B23) intervalová škála – fyzikální měření

{čas měření [sec]}, {systolický tlak [bar]}, {kožní odpor [$m\Omega$]}

1. Testy pro porovnávání dvou a více výběrů

Jeden výběr (test shody s modelem)

t- test (jednovýběrový) - populace s diax a IQ, μ norm. rozdělení B22 nebo když nevíme 2 výběry

$H_0: \mu = \text{konstanta}$ (B22), (B21, B23 / $N(\mu, \sigma^2)$ nebo $N > 30$)

Kolmogorov-Smirnovův test = distribuční funkce odpovídá modelové funkci

$H_0: F(x) = F_{\text{model}}(x)$ (B11 – B23)

χ^2 – test dobré shody - p. výběr barev

$H_0: \pi_i = \text{dané konstanty}(i)$ Pro všechny typy dat: (A1 – B23)

Nezávislé výběry

2 nezávislé výběry

rozhodnu se
tedy přijmu
nebo ne H_0
t-test (při homogenní varianci)

$H_0: \mu_1 = \mu_2$ (B22), (B21, B23 / $N(\mu, \sigma^2)$ nebo $N > 30$)

t-test (při heterogenní varianci)

$H_0: \mu_1 = \mu_2$ (B22), (B21, B23 / $N(\mu, \sigma^2)$ nebo $N > 30$)

• Wilcoxon-Whiteův test (= Mann-Whitney test) H_0 : shoda distri. funkce, rozdělení neskončené

$H_0: F_1(x) = F_2(x)$ (B12, B13), (B21, B23 / non $N(\mu, \sigma^2)$ a $N < 30$) *pro nenormální když norm. t-test*

• Mediánový test - rozdělení do 2 skup - nad a pod MED

$H_0: \text{MED}_{ZS1} = \text{MED}_{ZS2}$ (B11, B12)

• χ^2 – test homogenity (= nezávislosti)

H_0 : shodné rozložení $\{\pi_i\}$ v obou ZS (A1, A2, B1)

• Fisherův test - 2 nes. výběry, alternativní znak (M/Z, ano/ne)

$H_0: \pi_1 = \pi_2$ (A1)

H_0 : pravděpodobnost ano je stejná jako u 1. souboru i u 2. souboru
ano stejná u M i Z

k>2 nezávislých výběrů

Jednofaktorový model analýzy variance (ANOVA) *(nic než 2, se k-test)*

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k \quad (B22), (B21, B23 / N(\mu, \sigma^2) \text{ nebo } N > 30)$$

× Kruskal-Wallisův test

$$H_0: F_1(x) = \dots = F_k(x) \quad (B12, B13), (B21, B23 / \text{non } N(\mu, \sigma^2) \text{ a } N < 30)$$

(Tzn. shoda distrib. funkcí ve všech ZS)

Mediánový test

$$H_0: \text{MED}_{ZS1} = \dots = \text{MED}_{ZSk} \quad (B11, B12, B13)$$

χ^2 – test homogenity (= nezávislosti)

$$H_0: \text{shodné rozložení } \{\pi_i\} \text{ ve všech ZS} \quad (A1, A2, B1)$$

Závislé výběry

2 závislé výběry

t-test - *norm, N > 30*

$$H_0: \mu_{DIF} = 0 \quad (B22), (B21, B23 / N(\mu, \sigma^2) \text{ nebo } N > 30)$$

Wilcoxonův test - *norm normalita nebo málo lidí*

$$H_0: \text{MED}_{DIF} = 0 \quad (B12, B13), (B21, B23 / \text{non } N(\mu, \sigma^2) \text{ a } N < 30)$$

analogie Znaménkový (=binomický) test

$$H_0: \pi(+) = \pi(-) \quad (B11, B12)$$

Mc-Nemarův test - *oblibnost položek*

$$H_0: \pi(+, -) = \pi(-, +) \quad (A1)$$

k>2 závislých výběrů

Jednofaktorový model analýzy variance pro opakované observace

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k \quad (B22), (B21, B23 / N(\mu, \sigma^2) \text{ nebo } N > 30)$$

× Friedmanův test - *(rozobrem) Wilcoxonova paroveho testu*

$$H_0: F_1(x) = \dots = F_k(x) \quad (B12, B13), (B21, B23 / \text{non } N(\mu, \sigma^2) \text{ a } N < 30)$$

(Tzn. shoda distribuce ZS ve všech situacích)

2. Korelační metody

typ proměnných X, Y

| | | |
|---|--------|---|
| Pearsonův koeficient korelace | X a Y: | (B22), (B21, B23 a $N(\mu, \sigma^2)$ a lin. regrese) |
| <i>dobře měřitelné, analýza lineárního vztahu</i> Spearmanův koeficient korelace | X a Y: | (B21, B23) (B22 a nelineární regrese) |
| <i>nelineární, nesoraditelné</i> | X: | (B12, B13) |
| | Y: | (B12, B13, B21, B22, B23) |
| Čtyřpolní koeficient korelace | X a Y: | (A1) |
| Biseriální koeficient korelace | X: | (A1) |
| <i>1 aplikace Pearsona</i> | Y: | (B13, B21, B22, B23) |
| <i>1 znak binární, 2. alternativní</i> | | |