

8. TESTOVANIE HYPOTÉZ

8.1 Základné pojmy

Často sa vyskytuje úloha overiť, či náhodný výber (ľubovoľný prvok, výberová charakteristika, výberová funkcia alebo celý výber) patrí do základného súboru so známym parametrom. Podobne môžeme získať charakteristiky alebo výberové funkcie z niekoľkých výberov a chceme overiť, či sa tieto výberové charakteristiky alebo funkcie navzájom významne líšia. Tieto úlohy riešime pomocou tzv. **testov významnosti** štatistických hypotéz. Pri riešení problémov testovania štatistických hypotéz je potrebné rozlíšiť:

a) **jednoduché hypotézy**, ktoré vymedzujú rozdelenie vyšetrovanej veličiny jednoznačne tým, že špecifikujú všetky parametre,

b) **zložené hypotézy**, pre ktoré nie je niektorý parameter špecifikovaný jednoznačne a je najčastejšie vymedzený intervalom.

Pri testovaní štatistických hypotéz stojí proti sebe vždy testovaná nulová hypotéza H_0 a alternatívna hypotéza, ktorú označujeme H_1 . **Nulová hypotéza** H_0 je predpoklad existencie základného súboru s istým parametrom θ . Závažná je v tejto súvislosti formulácia **alternatívnej hypotézy** H_1 , t.j. hypotézy, ktorú prijmeme, keď neplatí nulová hypotéza. Táto formulácia bude rozhodujúca pre použitie tzv. jednostrannej alebo dvojstrannej alternatívnej hypotézy.

Testovanie štatistických hypotéz je jednoduchý rozhodovací postup, ktorý na základe náhodného výberu z daného základného súboru vedie buď k **zamietnutiu** testovanej nulovej hypotézy H_0 alebo k **nezamietnutiu** testovanej nulovej hypotézy H_0 .

Testy hypotéz v štatistickej analýze sú určené na odvodenie teoretického rozdelenia na základe vzorku dát (výsledkov merania). Testy hypotéz porovnávajú štatistické veličiny vypočítané zo vzorky dát s kritickými hodnotami vybraného rozdelenia na určitej úrovni hladiny významnosti α , s cieľom rozhodnúť, či vzorka dát patrí do vybraného rozdelenia. Kľúčom testov hypotéz je skonštruovať zo vzorky dát náhodného výberu (meraní) výberovú funkciu a túto porovnať so štatistickou funkciou vybraného rozdelenia v rámci testu nulovej hypotézy.

Keďže každá hodnota náhodného výberu je náhodná veličina, bude i výberová funkcia náhodnou veličinou.

Ďalej sú uvedené bez odvodu niektoré výberové funkcie, ktoré sa používajú v oblasti geodetických aplikácií. Rozdelenie výberových funkcií je konštruované s predpokladom, že náhodný výber je vzorkou základného súboru s normálnym rozdelením, ktorý je charakterizovaný strednou hodnotou a varianciou

$$x \in N(\mu = E(x), \sigma^2). \quad (8.1)$$

V ďalšom texte označenie štatistických veličín má význam:

$\mu = E(x)$ = X - (teoretická) stredná hodnota (skutočná, pravdepodobná hodnota meranej veličiny; $n \rightarrow \infty$),

\bar{x} - výberový priemer (priemerná hodnota, priemer vzorky meraní; $i = 1, 2, n$),

$\sigma^2 = \overline{m}^2$ - variancia (variancia základného súboru),

$\hat{\sigma}^2 = m^2$ - výberová variancia ,

$\sigma = \bar{m}$ - smerodajná odchýlka, základná stredná chyba,

$\hat{\sigma} = m$ - empirická stredná chyba (výberová stredná chyba, smerodajná odchýlka).

Vybrané výberové funkcie:

1. Výberový priemer

$$\bar{x} = \frac{1}{n}(x_1 + x_2 + \dots + x_n), \quad (8.2)$$

má normálne rozdelenie

$$N\left(\mu = E(x), \sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}\right), \quad (8.3)$$

kde x_i sú dielčie priemerné hodnoty vzorky meraní.

Výberový priemer \bar{x} sa počíta z hodnôt x_i s normálnym rozdelením, má normálne rozdelenie s rovnakou strednou hodnotou a menšiu varianciu. Strednú hodnotu výberového priemeru určíme

$$E(\bar{x}) = \frac{1}{n} \sum_{i=1}^n (x_i) = \frac{1}{n} n E(x) = E(x) = \mu \quad (8.4)$$

a varianciu výberového priemeru

$$V(\bar{x}) = \frac{1}{n^2} \sum_{i=1}^n V(x_i) = \frac{1}{n^2} n V(x) = \frac{1}{n} \sigma^2. \quad (8.5)$$

2. Výberová funkcia

$$\frac{\bar{x} - \mu}{\sigma} \sqrt{n}, \quad (8.6)$$

je normovanou hodnotou \bar{x} a má normálne rozdelenie so strednou hodnotou a varianciou $N(0,1)$.

3. Výberová funkcia

$$\frac{\bar{x} - \mu}{\hat{\sigma}} \sqrt{n}, \quad \text{kde } m = \hat{\sigma} = \sqrt{\frac{\sum_{i=1}^n (\bar{x} - x_i)^2}{n'}} \quad (8.7)$$

má Studentovo t rozdelenie s $t(n')$ a n' stupňami voľnosti (nadbytočných meraní), t.j. $n' = n - 1$, obecné $n' = n - k$.

4. Výberová funkcia

$$\frac{n' \hat{\sigma}^2}{\sigma^2}, \text{ kde } \hat{\sigma} = \sqrt{\frac{\sum_{i=1}^n (\bar{x} - x_i)^2}{n'}} \text{ a } n' = n - k, \quad (8.8)$$

je počet stupňov voľnosti, má χ^2 – rozdelenie (chí kvadrát) s n' stupňom voľnosti $\chi^2(n')$.

5. Výberová funkcia: $\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}$, kde $\hat{\sigma}_1$ a $\hat{\sigma}_2$ sú výberové stredné chyby dvoch výberov zo základného súboru s rovnakou varianciou

$$\hat{\sigma}_1 = \sqrt{\frac{\sum_{i=1}^{n_1} (\bar{x}_1 - x_i)^2}{n_1'}}, \quad \hat{\sigma}_2 = \sqrt{\frac{\sum_{j=1}^{n_2} (\bar{x}_2 - x_j)^2}{n_2'}}, \quad (8.9)$$

kde \bar{x}_1 a \bar{x}_2 sú výberové priemery a n_1' , n_2' zodpovedajúce stupne voľnosti F- rozdelenia (Fisherového rozdelenie) $F(n_1', n_2')$.

8.2 Test pravdepodobnosti pri známej variancii základného súboru

Predpokladajme, že x_1, x_2, \dots, x_n sú nezávislé merania zo základného súboru s normálnym rozdelením $N(E(x) = \mu, \sigma^2 = \bar{m}^2)$, kde μ je teoretická stredná hodnota (skutočná, pravdepodobná hodnota) meranej veličiny. Známa je základná (štandardná) stredná chyba σ . Ak x_i patrí do súboru s normálnym rozdelením $N(\mu, \sigma^2)$, pre $1 \leq i \leq n$ platí

$$\bar{x} \approx N\left(\mu, \sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}\right) \text{ a funkcia } \frac{\bar{x} - \mu}{\sigma} \sqrt{n} \left(= \frac{\varepsilon_{\bar{x}}}{m_{\bar{x}}} = \frac{\varepsilon_{\bar{x}}}{\frac{\sigma}{\sqrt{n}}} \right) \text{ ma normálne rozdelenie so strednou}$$

hodnotou a varianciou $N(0,1)$.

Pre hladinu významnosti α (riziko chybného rozhodnutia) dostaneme

$$P\left\{\left|\frac{\bar{x} - \mu}{\sigma} \sqrt{n}\right| \leq x_{\alpha}\right\} = 1 - \alpha \text{ alebo } P\{|\bar{x} - \mu| \leq c_1\} = 1 - \alpha, \quad (8.10)$$

kde x_{α} je kritická hodnota štandardného (základného) normálneho rozdelenia pre hladinu významnosti α a c_1 je

$$c_1 = x_{\alpha} \frac{\sigma}{\sqrt{n}}. \quad (8.11)$$

Pri testovaní výberovej strednej hodnoty merania \bar{x} , či je štatisticky rovná teoreticky strednej hodnote $E(x) = \mu$, hodnotíme dve možné hypotézy:

nulová hypotéza $\bar{x} = \mu$,

alternatívna hypotéza $H_1 \quad \bar{x} \neq \mu$.

Po zvolení hladiny významnosti α nájdeme kritickú hodnotu x_α z tabuľky Studentovho rozdelenia pre $n' = \infty$ a vypočítame c_1 . Ak $|\bar{x} - \mu| < c_1$ prijmem nulovú hypotézu H_0 na hladinu významnosti α . V opačnom prípade odmietneme H_0 .

Poznámka: Od hodnoty $n' > 25$ kritické hodnoty normálneho rozdelenia a Studentovho rozdelenia sú prakticky rovnaké.

Príklad 8.1: Náhodný výber má rozsah $n = 25$ prvkov. Posúdime, či náhodný výber so strednou hodnotou $\bar{x} = 0,267$ zodpovedá výberu zo základného súboru $\mu = 0,269$. Základná stredná chyba $\bar{m} = \sigma = 0,003$. Zvolíme hladiny významnosti $\alpha = 0,05$ a $0,01$. V tabuľke Studentovho rozdelenia na strane 25 pre $n' = \infty$ nájdeme kritické hodnoty x_α

$$\alpha = 0,05 \quad x_\alpha = 1,96 \quad (\approx 2,0),$$

$$\alpha = 0,01 \quad x_\alpha = 2,58 \quad (\approx 2,5).$$

$$\text{Vypočítame } c_1 = x_\alpha \frac{\sigma}{\sqrt{n}} \text{ pre } \alpha = 0,05 \quad c_1 = 1,96 \frac{0,003}{\sqrt{25}} = 0,001_2,$$

$$\alpha = 0,01 \quad c_1 = 2,58 \frac{0,003}{\sqrt{25}} = 0,001_5.$$

$$\text{Z porovnania vzťahov (8.10) } |\bar{x} - \mu| > c_1 \Rightarrow |-0,002| > 0,001_2 \text{ pre } \alpha = 0,05,$$

$$|-0,002| > 0,001_5 \text{ pre } \alpha = 0,01,$$

zamietame nulovú hypotézu H_0 a tvrdíme, že náhodný výber nepatrí do základného súboru s teoretickou strednou hodnotou $E(x) = \mu = 0,269$.

8.3 Test pravdepodobnosti pri neznámej variancii základného súboru

Predpokladajme, že máme rovnakú úlohu ako v predchádzajúcom teste. Vyhodnocujeme, či výberová stredná hodnota meraní je rovná teoretickej strednej hodnote μ za predpokladu, že základná stredná chyba $\bar{m} = \sigma$ je neznáma. Pre tento účel štatistického hodnotenia použijeme Studentovo t -rozdelenie vyjadrené rovnicou

$$t' = \frac{\frac{\bar{x} - \mu}{\sigma} \sqrt{n}}{\sqrt{\frac{(n-1)\hat{\sigma}^2}{\sigma^2} \frac{1}{(n-1)}}} = \frac{\bar{x} - \mu}{\hat{\sigma}} \sqrt{n} \approx t(n-1), \quad (8.12)$$

kde strednú hodnotu \bar{x} a empirickú strednú chybu $\hat{\sigma} = m$ vypočítame z rovníc

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2. \quad (8.13)$$

Pri zvolenej hladine významnosti α dostaneme

$$P\left\{\frac{\bar{x} - \mu}{\hat{\sigma}} \sqrt{n} \leq t_\alpha(n-1) = 1 - \alpha\right\} \text{ alebo } P\{|\bar{x} - \mu| \leq c_2\} = 1 - \alpha, \quad (8.14)$$

kde $t_{\alpha}(n-1)$ je kritická hodnota t rozdelenia s $n - 1$ stupňami voľnosti s hladinou významnosti α , c_2 je

$$c_2 = t_{\alpha}(n-1) \frac{\hat{\sigma}}{\sqrt{n}}. \quad (8.15)$$

Pre zvolenú hladinu významnosti nájdeme kritickú hodnotu t_{α} . Ak $|\bar{x} - \mu| < c_2$ prijímame nulovú hypotézu H_0 . V opačnom prípade platí alternatívna hypotéza H_1 .

Príklad 8.2. Vzďialenosť s bola meraní 10 krát s výsledkami v mm

i	1	2	3	4	5	6	7	8	9	10
s_i	1322	1324	1318	1319	1318	1323	1325	1328	1326	1317

Predpokladáme, že všetky merania sa riadia normálnym rozdelením. Testom zisťujeme, či vyrovnaná vzdialenosť je štatisticky rovná hodnote 1323 mm na hladine významnosti $\alpha = 0,05$ (= 5 %), (t.j. s 95 % spoľahlivosťou). Nulová hypotéza a alternatívna hypotéza sú:

$$H_0 : \bar{x} = \mu = 1323 \text{ mm}; \quad \text{alternatívna hypotéza } H_1 : \bar{x} \neq \mu = 1323 \text{ mm}.$$

Priemerná hodnota a empirická stredná chyba sú ($n = 10$)

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = 1322 \text{ mm}, \quad \hat{\sigma} = \sqrt{\frac{(x_i - \bar{x})^2}{n-1}} = \sqrt{14,67} = 3,830 \text{ mm}.$$

$$\text{Z tabuľky 3.1 Studentovho rozdelenia str. 25 nájdeme } t_{\alpha}(n-1) = t_{\alpha=5\%}(n-1=9) = 2,262. \\ c_2 = t_{\alpha}(n-1) \frac{\hat{\sigma}}{\sqrt{n}} = 2,262 \cdot \frac{3,830}{\sqrt{10}} = 2,74, \quad \bar{x} - \mu = 1322 - 1323 = -1 \text{ mm}.$$

Keďže $|\bar{x} - \mu| < c_2$ prijímame záver, že vyrovnaná hodnota odmeranej vzdialenosti je štatisticky rovná hodnote 1323 mm, pri 5 % riziku chybného rozhodnutia.

8.4 Testovanie hypotézy o rovnosti variancií dvoch normálne rozdelených základných súborov

Testujeme hypotézu, že dve výberové variancie $\hat{\sigma}_1^2$ a $\hat{\sigma}_2^2$ z dvoch výberov o rozsahu n_1 a n_2 odpovedajú výberom z dvoch základných súborov, pre ktoré platí rovnosť variancií, teda $\sigma_1^2 = \sigma_2^2$.

Testovacím kritériom bude veličina

$$F = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}, \quad (8.16)$$

kde $\hat{\sigma}_1$ a $\hat{\sigma}_2$ sú vzťahy (8.9), ktoré majú F -rozdelenie (Fisherové rozdelenie) s $n'_1 = n_1 - 1$ a $n'_2 = n_2 - 1$ stupňami voľnosti. Vo vzťahu (8.16) volíme vždy $\hat{\sigma}_1^2 > \hat{\sigma}_2^2$. Z tabuliek F -rozdelenia nájdeme pre zvolenú hladinu významnosti α kritickú hodnotu F_α na pravej strane rozdelenia (obojsmerný test). Nulovú hypotézu budeme zamietť, ak $F > F_\alpha$.

Príklad 8.3. Vzdialenosť s bola odmeraná dvoma elektronickými teodolitmi (ET) s nasledovnými odmeranými hodnotami a strednými chybami.

$$1. \text{ ET } n_1 = 41, \quad \hat{\sigma}_1 = 4,2 \text{ mm},$$

$$2. \text{ ET } n_2 = 61, \quad \hat{\sigma}_2 = 3,0 \text{ mm}.$$

Testujeme, či obidva teodolity majú rovnakú presnosť pri hladine významnosti $\alpha = 5\%$.

Hypotézy sú: $H_0 : \sigma_1^2 = \sigma_2^2$; alternatívna hypotéza $H_1 : \sigma_1^2 \neq \sigma_2^2$.

Pre F_α z tab. VIIb rozdelenia F nájdeme

$$F_{\alpha=5\%(n_1-1, n_2-2)} = 1,64$$

$$F = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} = \left(\frac{4,2}{3,0} \right)^2 = 1,96 > 1,64.$$

Na základe testu zamietneme nulovú hypotézu o rovnosti variancií oboch súborov meraní pri 5 % riziku chybného rozhodnutia. Inými slovami obidva prístroje štatisticky nemajú rovnakú presnosť.

8.5 Testovanie hypotézy o rovnosti stredných hodnôt dvoch normálne rozdelených základných súborov

Testujeme hypotézu, že dva výbery s výberovými priermi \bar{x}_1 a \bar{x}_2 a výberovými strednými chybami $\hat{\sigma}_1$ a $\hat{\sigma}_2$ sú výbery z dvoch základných súborov, pre ktoré platí rovnosť ich stredných hodnôt $E(x)_1 = E(x)_2$. Testovacie kritérium volíme podľa toho, či variancie základných súborov sú alebo nie sú rovnaké. Musíme najskôr urobiť test rozdielu medzi dvoma výberovými varianciami podľa kap. 8.4, ktorý určí, že:

a) obe variancie sa významne nelíšia, potom použijeme testovacie kritérium

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{(n_1 - 1)\hat{\sigma}_1^2 + (n_2 - 1)\hat{\sigma}_2^2}} \sqrt{n_1 + n_2 - 2} \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \quad (8.17)$$

alebo jednoduchší tvar

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}}, \quad (8.18)$$

kde t má Studentovo rozdelenie s $(n_1 + n_2 - 2)$ stupňami voľnosti.

Z tabuliek Studentovho rozdelenia nájdeme pre zvolenú hladinu významnosti α hodnotu t_α . Nulovú hypotézu budeme zamietť pri $t > t_\alpha$.

b) obe variancie sa významne líšia, potom použijeme testovacie kritérium (8.18) a jeho vypočítanú hodnotu porovnáme s hodnotou

$$t_\alpha^* = \frac{t_{n'_1, \alpha} \frac{\hat{\sigma}_1^2}{n_1} + t_{n'_2, \alpha} \frac{\hat{\sigma}_2^2}{n_2}}{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}, \quad (8.19)$$

kde $t_{n'_1, \alpha}$ a $t_{n'_2, \alpha}$ sú kritické hodnoty z tabuliek Studentovho rozdelenia a t_α^* je ich zväžený priemer pre hladinu významnosti α a stupne voľnosti $n'_1 = n_1 - 1$, $n'_2 = n_2 - 1$. (8.20)

Príklad 8.4. Posúdime významnosť rozdielu medzi dvoma výberovými priermi na hladine významnosti $\alpha = 0,05$. Prvý výber o rozsahu $n_1 = 16$ má priemer $\bar{x} = 225$ a strednú chybu $\sigma_1 = m_1 = 31$; druhý výber o rozsahu $n_2 = 11$ má priemer $\bar{x}_2 = 190$ a empirickú strednú chybu $\sigma_2 = m_2 = 21$.

Nulová hypotéza bude $H_0 : E(x)_1 = E(x)_2$, alternatívna hypotéza $H_1 : E(x)_1 \neq E(x)_2$.

Najprv posúdime významnosť rozdielov medzi dvoma výberovými varianciami testom F – rozdelenia pri hladine významnosti $\alpha = 5\%$. Pre $n'_1 = 15$ a $n'_2 = 10$ je $F_{\alpha=5\%} = 2,84$.

Hodnota testovacieho kritéria bude

$$F = \left(\frac{31}{21} \right)^2 = 2,18.$$

Pretože $F < F_{\alpha=5\%} = 2,18 < 2,84$ nemôžeme zamietnuť nulovú hypotézu. Nezamietneme predpoklad rovnosti variancií oboch základných súborov $\sigma_1^2 = \sigma_2^2$.

Keďže rozdiel medzi obidvoma výberovými varianciami pokladáme za štatisticky nevýznamný, bude hodnota testovacieho kritéria vykonané podľa vzťahu (8.17)

$$t = \frac{|225 - 190|}{\sqrt{15 \cdot 31^2 + 10 \cdot 21^2}} \sqrt{16 + 11 - 2} \cdot \sqrt{\frac{16 \cdot 11}{16 + 11}} = 3,25.$$

Z tabuliek Studentovho rozdelenia nájdeme pre $(n_1 + n_2 - 2) = 25$ stupňov a hladinu významnosti $\alpha = 0,05$ kritickú hodnotu $t_{0,05} = 2,06$. Z porovnania vychádza $t > t_\alpha$, číselne $3,25 > 2,06$, teda zamietame nulovú hypotézu a tvrdíme, že hodnotené výbery nie sú zo základných súborov s rovnakou strednou hodnotou.

Príklad 8.5. Posúdime významnosť rozdielu medzi dvoma výberovými priermi na hladine významnosti $\alpha = 0,05$. Prvý výber o rozsahu $n_1 = 16$ má priemer $\bar{x}_1 = 215$ a výberovú strednú chybu $\hat{\sigma}_1 = 41$; druhý výber o rozsahu $n_2 = 11$ má priemer $\bar{x}_2 = 190$ a strednú chybu $\hat{\sigma}_2 = 21$.

Najprv posúdime obidve výberové variancie. Testovacím kritériom F rozdelenia

$$F = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} = \frac{1681}{441} = 3,81$$

a vypočítanú hodnotu porovnáme s kritickou hodnotou z tabuliek F – rozdelenia pre $\alpha = 0,05$ a $n'_1 = 15$, $n'_2 = 10$, $F_{\alpha = 5\%} = 2,84$. Pretože $F > F$ pokladáme rozdiel variancií za štatisticky významný a pri testovaní rozdielu medzi priemermi postupujeme podľa kap. 8.5b. Formulácia hypotéz bude rovnaká ako v príklade 8.3. Vypočítame hodnoty:

testovacieho kritéria (8.18)

a kritické hodnoty (8.19)

$$t = \frac{215 - 190}{\sqrt{\frac{1681}{16} + \frac{441}{11}}} = \frac{25}{12,05} = 2,07, \quad t_{\alpha}^* = \frac{2,13 \frac{1681}{16} + 2,23 \frac{441}{11}}{\frac{1681}{16} + \frac{441}{11}} = 2,16,$$

kde sme dosadili kritické hodnoty Studentovho rozdelenia $t_{\alpha = 5\%}(n'_1 = 15) = 2,13$ a $t_{\alpha = 5\%}(n'_2 = 10) = 2,23$.

Pretože $t < t_{\alpha}^* = 2,07 < 2,16$ nemôžeme hypotézu o rovnosti stredných hodnôt zamietnuť.

8.6 Rozdelenie χ^2

Ak x_1, x_2, \dots, x_n sú nezávislé náhodné premenné, ktoré majú normálne rozdelenie s parametrami $N(0,1)$, potom nasledujúci súčet

$$\chi^2 = x_1^2 + x_2^2 + \dots + x_n^2 \quad (8.21)$$

má χ^2 (chí kvadrant) rozdelenie s n stupňami voľnosti.

Nech náhodné premenné x_i ($x_i = 1, 2, \dots, n$) majú normálne rozdelenie s parametrami $x_i \sim N(E(x), \sigma^2)$. Potom n nezávislé premenných s normálnym rozdelením má rovnakú strednú hodnotu \bar{x} a varianciu $\hat{\sigma}^2$. Ak $n \rightarrow \infty$ $\bar{x} \rightarrow \mu = X$ stredná hodnota sa rovná skutočnej hodnote meranej veličiny. Nech \bar{x} a $\hat{\sigma}^2 = m^2$ (empirická výberová variancia) sú vzorky strednej hodnoty a variancie vypočítané podľa vzťahov

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad m^2 = \hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2. \quad (8.22)$$

Stredná hodnota vzoriek \bar{x} má normálne rozdelenie s parametrami

$$\bar{x} \sim N\left(E(\bar{x}), \frac{\sigma^2}{n}\right). \quad (8.23)$$

Rozdelenie χ^2 sa konštruuje štatisticky zo vzťahu

$$\chi^2 = \frac{\hat{\sigma}^2}{\sigma^2} = \frac{1}{n-1} \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sigma^2}, \quad \chi^2(n-1) = \frac{\sum \varepsilon \varepsilon}{\sigma^2} = \frac{\sum vv}{\sigma^2}. \quad (8.24)$$

Potom platí

$$(n-1)\chi^2 \sim \frac{\hat{\sigma}^2(n-1)}{\sigma^2}, \quad (8.25)$$

kde $(n-1)$ je počet stupňov voľnosti (nadbytočných meraní).

8.7 Analýza variácií

Štatistický experiment - akým je aj geodetické meranie, môže byť ovplyvnený rôznymi faktormi. Sú to: merací prístroj, merač, atmosferické podmienky, prostredie merania atď. Podľa vplyvu jednotlivých faktorov merania môžeme sériu meraní rozdeliť do skupín (napr. zmena merača, meranie iným prístrojom, meranie za iných atmosferických podmienok, iná etapa merania atď.). Analýzou variácií sa pokúšame vydedukovať, či rozdiely (napr. zmena merača, meranie iným prístrojom, meranie za iných atmosferických podmienok, iná etapa merania atď.). Analýzou variácií sa pokúšame vydedukovať, či rozdiely v skupinách meraní sú štatisticky významné, t.j. či skupiny meraní patria do spoločného základného súboru meraní. Je to možné vykonať internou analýzou variácií každej skupiny meraní a externou analýzou variácií medzi jednotlivými skupinami meraní. Ďalej je uvedená analýza variácií s účinkom vplyvu len jedného faktora.

Predpokladajme, že v každom výsledku meraní F_i ($1 \leq i \leq m$) sa prejavil vplyv faktora na skupinu n_i meraní ℓ_{ij} ($i \leq j \leq n_i$) a všetky merania v každej skupine majú normálne rozdelenie s parametrami $N(\mu = E(l), \sigma^2)$.

Skupina	Meranie	Rozdelenie
F_1	$\ell_{11} \ell_{12} \dots \ell_{1,n_1}$	$N(\mu_1, \sigma^2),$
F_2	$\ell_{21} \ell_{22} \dots \ell_{2,n_2}$	$N(\mu_2, \sigma^2),$
...
F_i	$\ell_{i1} \ell_{i2} \dots \ell_{i,n_i}$	$N(\mu_i, \sigma^2),$
...
F_m	$\ell_{m1} \ell_{m2} \dots \ell_{m,n_m}$	$N(\mu_m, \sigma^2).$

(8.26)

Úlohou je testovať hypotézu teoretických stredných hodnôt μ :

$$H_0: \mu_1 = \mu_2 = \dots = \mu_m. \quad (8.27)$$

Najskôr preskúmame súčet štvorcov jednotlivých variácií S_I^2 všetkých meraní s ohľadom na odpovedajúcu hodnotu pravdepodobnej hodnoty výsledku meraní $\mu_1, \mu_2, \dots, \mu_m$

$$S_I^2(\mu_1, \mu_2, \dots, \mu_m) = \sum_{i=1}^m \sum_{j=1}^{n_i} (\ell_{ij} - \mu_i)^2 = \min. \quad (8.28)$$

Derivujeme funkciu (8.27) a určíme vyrovnanú hodnotu (aritmetický priemer)