Kapitola 1.: Základní pojmy matematické statistiky

1.1. Motivace

Při aplikaci metod popisné statistiky dospíváme pomocí zjištěných dat k závěrům, které se týkají pouze výběrového souboru. Naproti tomu matematická statistika nám umožňuje na základě znalosti náhodného výběru a statistik z něj odvozených (tj. např. výběrového průměru, výběrového rozptylu, výběrového koeficientu korelace, hodnoty výběrové distribuční funkce apod.) učinit závěry o parametrech nebo tvaru rozložení, z něhož daný náhodný výběr pochází. Často se jedná o bodové či intervalové odhady parametrů a parametrických funkcí a testování hypotéz o nich.

1.2. Náhodný výběr a statistiky odvozené z náhodného výběru

1.2.1. Pojem náhodného výběru

Nechť $X_1, ..., X_n$ jsou stochasticky nezávislé náhodné veličiny, které mají všechny stejné rozložení $L(\vartheta)$. Řekneme, že $X_1, ..., X_n$ je náhodný výběr rozsahu n z rozložení $L(\vartheta)$. (Číselné realizace $x_1, ..., x_n$ náhodného výběru $X_1, ..., X_n$ uspořádané do sloupcového vektoru představují datový soubor.)

Nechť $(X_1,Y_1), ..., (X_n,Y_n)$ jsou stochasticky nezávislé dvourozměrné náhodné vektory, které mají všechny stejné dvourozměrné rozložení $L_2(\vartheta)$. Řekneme, že $(X_1,Y_1), ..., (X_n,Y_n)$ je dvourozměrný náhodný výběr rozsahu n z dvourozměrného rozložení $L_2(\vartheta)$. (Číselné realizace $(x_1,y_1), ..., (x_n,y_n)$ náhodného výběru $(X_1,Y_1), ..., (X_n,Y_n)$ uspořádané do matice typu $2\times n$ představují dvourozměrný datový soubor.)

Analogicky lze definovat p-rozměrný náhodný výběr rozsahu n z p-rozměrného rozložení $L_p(\vartheta)$.

1.2.2. Pojem statistiky, příklady důležitých statistik

Libovolná funkce $T = T(X_1, ..., X_n)$ náhodného výběru $X_1, ..., X_n$ (resp. p-rozměrného náhodného výběru) se nazývá statistika.

a) Nechť $X_1, ..., X_n$ je náhodný výběr, $n \geq 2$.

Statistika $M = \frac{1}{n}\sum_{i=1}^{n} X_i$ se nazývá výběrový průměr,

Statistika $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - M\right)^2$ výběrový rozptyl,

Statistika $S = \sqrt{S^2}$ výběrová směrodatná odchylka.

Pro libovolné, ale pevně zvolené reálné číslo x je statistikou též hodnota výběrové distribuční funkce

$$F_n(x) = \frac{1}{n}\operatorname{card}\{i; X_i \leq x\}.$$

b) Nechť $X_{11},\ldots,X_{1n_1}$, ..., $X_{p1},\ldots,X_{pn_p}$ je p stochasticky nezávislých náhodných výběrů o rozsazích $n_1 \geq 2$, ..., $n_p \geq 2$. Celkový rozsah je $n = \sum_{j=1}^{p} n_j$. Označme $M_1, ..., M_p$ výběrové průměry a $S_1^2, ..., S_p^2$ výběrové rozptyly jednotlivých výběrů. Nechť $c_1, ..., c_p$ jsou reálné konstanty, aspoň jedna nenulová

Statistika $\sum_{j=1}^{p} c_j M_j$ se nazývá lineární kombinace výběrových průměrů.

Statistika $S_*^2 = \frac{\sum_{j=1}^{p}(n_j - 1)S_j^2}{n-p}$ se nazývá vážený průměr výběrových rozptylů.

c) Nechť (X_1,Y_1), ..., (X_n,Y_n) je náhodný výběr z dvourozměrného rozložení. Označme $M_1 = \frac{1}{n}\sum_{i=1}^{n} X_i$, $M_2 = \frac{1}{n}\sum_{i=1}^{n} Y_i$

Statistika $S_{12} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - M_1)(Y_i - M_2)$ je výběrová kovariance,

statistika $R_{12} = \begin{cases} \frac{1}{n-1}\sum_{i=1}^{n}\frac{X_i - M_1}{S_1}\cdot\frac{Y_i - M_2}{S_2} & \text{pro } S_1 S_2 \neq 0 \\ 0 \text{ jinak} \end{cases}$ výběrový koeficient korelace.

Pro libovolnou, ale pevně zvolenou dvojici reálných čísel x,y je statistikou též hodnota výběrové simultánní distribuční funkce

$$F_n(x,y) = \frac{1}{n}\text{card}\{i; X_i \le x \wedge Y_i \le y\}.$$

(Číselné realizace m, s^2, s, s_{12}, r_{12} statistik M, S^2, S, S_{12}, R_{12} odpovídají číselným charakteristikám znaků v popisné statistice, ale u rozptylu, směrodatné odchylky, kovariance a koeficientu korelace je multiplikativní konstanta $\frac{1}{n-1}$, nikoli $\frac{1}{n}$, jak tomu bylo v popisné statistice.)

1.3. Bodové a intervalové odhady parametrů a parametrických funkcí

Vycházíme z náhodného výběru X_1, ..., X_n z rozložení $L(\vartheta)$, které závisí na parametru ϑ. Množinu všech přípustných hodnot tohoto parametru označíme Ξ. Parametr ϑ neznáme a chceme ho odhadnout pomocí daného náhodného výběru (případně chceme odhadnout nějakou parametrickou funkci $h(\vartheta)$).

Bodovým odhadem parametrické funkce $h(\vartheta)$ je statistika $T_n = T(X_1, ..., X_n)$, která nabývá hodnot blízkých $h(\vartheta)$, ať je hodnota parametru ϑ jakákoliv. Existují různé metody, jak konstruovat bodové odhady (např. metoda momentů či metoda maximální věrohodnosti, ale těmi se zde zabývat nebudeme) a také různé typy bodových odhadů. Omezíme se na odhady nestranné, asymptoticky nestranné a konzistentní.

Intervalovým odhadem parametrické funkce $h(\vartheta)$ rozumíme interval (D, H), jehož meze jsou statistiky $D = D(X_1, ..., X_n)$, $H = H(X_1, ..., X_n)$ a který s dostatečně velkou pravděpodobností pokrývá $h(\vartheta)$, ať je hodnota parametru ϑ jakákoliv.

1.3.1. Typy bodových odhadů

Nechť X_1, ..., X_n je náhodný výběr z rozložení $L(\vartheta)$, $h(\vartheta)$ je parametrická funkce, T, T_1, T_2, ... jsou statistiky.

a) Řekneme, že statistika T je nestranným odhadem parametrické funkce $h(\vartheta)$, jestliže $\forall\vartheta \in \Xi: E(T) = h(\vartheta)$.
(Význam nestrannosti spočívá v tom, že odhad T nesmí parametrickou funkci $h(\vartheta)$ systematicky nadhodnocovat ani podhodnocovat. Není-li tato podmínka splněna, jde o vychýlený odhad.)

b) Jsou-li T_1, T_2 nestranné odhady téže parametrické funkce $h(\vartheta)$, pak řekneme, že T_1 je lepší odhad než T_2, jestliže
$\forall\vartheta \in \Xi: D(T_1) < D(T_2)$.

c) Posloupnost $\{T_n\}_{n=1}^{\infty}$ se nazývá posloupnost asymptoticky nestranných odhadů parametrické funkce $h(\vartheta)$, jestliže

$$\forall \vartheta \in \Xi : \lim_{n\to\infty} E(T_n) = h(\vartheta).$$

(Význam asymptotické nestrannosti spočívá v tom, že s rostoucím rozsahem výběru klesá vychýlení odhadu.)

d) Posloupnost $\{T_n\}_{n=1}^{\infty}$ se nazývá posloupnost konzistentních odhadů parametrické funkce $h(\vartheta)$, jestliže

$$\forall \vartheta \in \Xi \forall \varepsilon > 0 : \lim_{n\to\infty} P\left(\left|T_n - h(\vartheta)\right| > \varepsilon\right) = 0.$$

(Význam konzistence spočívá v tom, že s rostoucím rozsahem výběru klesá pravděpodobnost, že odhad se bude realizovat daleko od parametrické funkce $h(\vartheta)$.)

Lze dokázat, že z nestrannosti odhadu vyplývá jeho asymptotická nestrannost a z asymptotické nestrannosti vyplývá konzistence, pokud posloupnost rozptylů odhadu konverguje k nule.

1.3.2. Vlastnosti důležitých statistik

a) Nechť $X_1, ..., X_n$ je náhodný výběr z rozložení se střední hodnotou μ, rozptylem σ^2 a distribuční funkcí $\Phi(x)$. Nechť $n \geq 2$ Označme M_n výběrový průměr, S_n^2 výběrový rozptyl a pro libovolné, ale pevně dané $x \in R$ $F_n(x)$ hodnotu výběrové distribuční funkce.

Pak M_n je nestranným odhadem μ (tj. $E(M_n) = \mu$) s rozptylem $D(M) = \frac{\sigma^2}{n}$, S_n^2 je nestranným odhadem σ^2 (tj. $E(S_n^2) = \sigma^2$), ať jsou hodnoty parametrů μ, σ^2 jakékoli. Dále platí, že pro libovolné, ale pevně dané $x \in R$ je výběrová distribuční funkce $F_n(x)$ nestranným odhadem $\Phi(x)$ (tj. $E(F_n(x)) = \Phi(x)$) s rozptylem $D(F_n(x)) = \Phi(x)(1- \Phi(x))/n$, ať je hodnota distribuční funkce $\Phi(x)$ jakákoliv.

Posloupnost $\{M_n\}_{n=1}^{\infty}$ je posloupnost konzistentních odhadů μ. $\{S_n^2\}_{n=1}^{\infty}$ je posloupnost konzistentních odhadů σ^2. Pro libovolné, ale pevně dané $x \in R$ je $\{F_n(x)\}_{n=1}^{\infty}$ posloupnost konzistentních odhadů $\Phi(x)$.

b) Nechť $X_{11},\dots,X_{1n_1}$, ..., $X_{p1},\dots,X_{pn_p}$ je p stochasticky nezávislých náhodných výběrů o rozsazích $n_1 \geq 2$, ..., $n_p \geq 2$ z rozložení se středními hodnotami μ_1, ..., μ_p a rozptylem σ^2. Celkový rozsah je $n = \sum_{j=1}^{p} n_j$. Nechť c_1, ..., c_p jsou reálné konstanty, aspoň jedna nenulová. Pak lineární kombinace výběrových průměrů $\sum_{j=1}^{p} c_j M_j$ je nestranným odhadem lineární kombinace středních hodnot $\sum_{j=1}^{p} c_j \mu_j$, ať jsou střední hodnoty μ_1, ..., μ_p jakékoli a vážený průměr výběrových rozptylů $S_*^2 = \frac{\sum_{j=1}^{p}(n_j - 1)S_j^2}{n-p}$ je nestranným odhadem rozptylu σ^2, ať je rozptyl σ^2 jakýkoliv.

c) Nechť (X_1,Y_1), ..., (X_n,Y_n) je náhodný výběr z dvourozměrného rozložení s kovariancí σ_{12} a koeficientem korelace ρ. Pak výběrová kovariance S_{12} je nestranným odhadem kovariance σ_{12}, ať je kovariance σ_{12} jakákoli, avšak $E(R_{12})$ je rovno ρ pouze přibližně (shoda je vyhovující pro $n > 30$), ať je korelační koeficient ρ jakýkoli.

1.3.3. Pojem intervalu spolehlivosti

Nechť $X_1, ..., X_n$ je náhodný výběr z rozložení $L(\vartheta)$, $h(\vartheta)$ je parametrická funkce, $\alpha \in (0,1)$, $D = D(X_1, ..., X_n)$, $H = H(X_1, ..., X_n)$ jsou statistiky.

a) Interval (D, H) se nazývá 100(1-α)% (oboustranný) interval spolehlivosti pro parametrickou funkci $h(\vartheta)$, jestliže: $\forall \vartheta \in \Xi : P(D < h(\upsilon) < H) \geq 1\text{-}\alpha$.

b) Interval (D, ∞) se nazývá 100(1-α)% levostranný interval spolehlivosti pro parametrickou funkci $h(\vartheta)$, jestliže: $\forall \vartheta \in \Xi : P(D < h(\vartheta)) \geq 1\text{-}\alpha$.

c) Interval (-∞, H) se nazývá 100(1-α)% pravostranný interval spolehlivosti pro parametrickou funkci $h(\vartheta)$, jestliže: $\forall \vartheta \in \Xi : P(h(\vartheta) < H) \geq 1\text{-}\alpha$.

d) Číslo α se nazývá riziko (zpravidla α = 0,05, méně často 0,1 či 0,01), číslo 1 – α se nazývá spolehlivost.

1.3.4. Postup při konstrukci intervalu spolehlivosti

a) Vyjdeme ze statistiky V, která je nestranným bodovým odhadem parametrické funkce $h(\vartheta)$.

b) Najdeme tzv. pivotovou statistiku W, která vznikne transformací statistiky V, je monotónní funkcí $h(\vartheta)$ a přitom její rozložení je známé a na $h(\vartheta)$ nezávisí. Pomocí známého rozložení tzv. pivotové statistiky W najdeme kvantily $w_{\alpha/2}$, $w_{1-\alpha/2}$, takže platí:
$\forall \vartheta \in \Xi : P(w_{\alpha/2} < W < w_{1-\alpha/2}) \geq 1 - \alpha$.

c) Nerovnost $w_{\alpha/2} < W < w_{1-\alpha/2}$ převedeme ekvivalentními úpravami na nerovnost $D < h(\vartheta) < H$.

d) Statistiky D, H nahradíme jejich číselnými realizacemi d, h a získáme tak 100(1-α)% empirický interval spolehlivosti, o němž prohlásíme, že pokrývá $h(\vartheta)$ s pravděpodobností aspoň 1 – α. (Tvrzení, že (d,h) pokrývá $h(\vartheta)$ s pravděpodobností aspoň 1 – α je třeba chápat takto: jestliže mnohonásobně nezávisle získáme realizace $x_1, ..., x_n$ náhodného výběru $X_1, ..., X_n$ z rozložení $L(\vartheta)$ a pomocí každé této realizace sestrojíme 100(1-α)% empirický interval spolehlivosti pro $h(\vartheta)$, pak podíl počtu těch intervalů, které pokrývají $h(\vartheta)$ k počtu všech sestrojených intervalů bude přibližně 1 – α.)

1.3.5. Příklad

Nechť $X_1, ..., X_n$ je náhodný výběr z $N(\mu,\sigma^2)$, kde $n \geq 2$ a rozptyl σ^2 známe. Sestrojte 100(1-α)% interval spolehlivosti pro neznámou střední hodnotu μ.

Řešení: V tomto případě parametrická funkce $h(\vartheta) = \mu$. Nestranným odhadem střední hodnoty je výběrový průměr (viz 1.3.(a)) $M = \frac{1}{n}\sum_{i=1}^{n} X_i$. Protože M je lineární kombinací normálně rozložených náhodných veličin, bude mít také normální rozložení se střední hodnotou E(M) = μ a rozptylem $D(M) = \frac{\sigma^2}{n}$. Pivotovou statistikou W bude standardizovaná náhodná veličina

$U = \frac{M-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$. Kvantil $w_{\alpha/2} = u_{\alpha/2} = -u_{1-\alpha/2}$, $w_{1-\alpha/2} = u_{1-\alpha/2}$.

$\forall \vartheta \in \Xi : 1 - \alpha \leq P(-u_{1-\alpha/2} < U < u_{1-\alpha/2}) =$

$$P\left(-u_{1-\alpha/2} < \frac{M-\mu}{\frac{\sigma}{\sqrt{n}}} < u_{1-\alpha/2}\right) = P\left(M - \frac{\sigma}{\sqrt{n}}u_{1-\alpha/2} < \mu < M + \frac{\sigma}{\sqrt{n}}u_{1-\alpha/2}\right).$$

Meze 100(1-α)% intervalu spolehlivosti pro střední hodnotu μ při známém rozptylu σ^2 tedy jsou:

$D = M - \frac{\sigma}{\sqrt{n}} u_{1-\alpha/2}$, $H = M + \frac{\sigma}{\sqrt{n}} u_{1-\alpha/2}$.

Při konstrukci jednostranných intervalů spolehlivosti se riziko nepůlí, tedy 100(1-α)% levostranný interval spolehlivosti pro μ je $\left(M - \frac{\sigma}{\sqrt{n}} u_{1-\alpha}, \infty \right)$ a pravostranný je $\left(-\infty, M + \frac{\sigma}{\sqrt{n}} u_{1-\alpha} \right)$.

Dosadíme-li do vzorců pro dolní a horní mez číselnou realizaci m výběrového průměru M, dostaneme 100(1-α)% empirický interval spolehlivosti.

1.3.5. Šířka intervalu spolehlivosti

Nechť (d, h) je 100(1-α)% empirický interval spolehlivosti pro $h(\vartheta)$ zkonstruovaný pomocí číselných realizací $x_1, ..., x_n$ náhodného výběru $X_1, ..., X_n$ z rozložení $L(\vartheta)$.

a) Při konstantním riziku klesá šířka h-d s rostoucím rozsahem náhodného výběru.
b) Při konstantním rozsahu náhodného výběru klesá šířka h-d s rostoucím rizikem.

Využití bodu (a) při stanovení minimálního rozsahu výběru z normálního rozložení: Nechť $X_1, ..., X_n$ je náhodný výběr z $N(\mu, \sigma^2)$, kde σ^2 známe. Jaký musí být minimální rozsah výběru n, aby šířka 100(1-α)% empirického intervalu spolehlivosti pro střední hodnotu μ nepřesáhla číslo Δ?

Řešení: Požadujeme, aby $\Delta \geq h - d = m + \frac{\sigma}{\sqrt{n}} u_{1-\alpha/2} - (m - \frac{\sigma}{\sqrt{n}} u_{1-\alpha/2}) = \frac{2\sigma}{\sqrt{n}} u_{1-\alpha/2}$. Z této podmínky dostaneme, že $n \geq \frac{4\sigma^2 u_{1-\alpha/2}{}^2}{\Delta^2}$. Za rozsah výběru zvolíme nejmenší přirozené číslo vyhovující této podmínce.

1.4. Úvod do testování hypotéz

Nulovou hypotézou rozumíme nějaké tvrzení o parametrech nebo typu rozložení, z něhož pochází náhodný výběr. Nulová hypotéza vyjadřuje nějaký teoretický předpoklad, často skeptického rázu a uživatel ji musí stanovit předem, bez přihlédnutí k datovému souboru. Proti nulové hypotéze stavíme alternativní hypotézu, která říká, co platí, když neplatí nulová hypotéza. Např. nulová hypotéza tvrdí, že střední hodnota hmotnosti balíčků cukru balených na automatické lince se nezměnila seřízením automatu, zatímco alternativní hypotéza tvrdí opak. Postup, který je založen na daném náhodném výběru a s jehož pomocí rozhodneme o zamítnutí či nezamítnutí nulové hypotézy, se nazývá testování hypotéz.

1.4.1. Nulová a alternativní hypotéza

Nechť $X_1, ..., X_n$ je náhodný výběr z rozložení $L(\vartheta)$, kde parametr $\vartheta \in \Xi$ neznáme. Nechť $h(\vartheta)$ je parametrická funkce a c daná reálná konstanta.

a) Oboustranná alternativa: Tvrzení H_0: $h(\vartheta) = c$ se nazývá jednoduchá nulová hypotéza. Proti nulové hypotéze postavíme složenou alternativní hypotézu H_1: $h(\vartheta) \neq c$.

b) Levostranná alternativa: Tvrzení H_0: $h(\vartheta) \geq c$ se nazývá složená pravostranná nulová hypotéza. Proti jednoduché nebo složené pravostranné nulové hypotéze postavíme složenou levostrannou alternativní hypotézu H_1: $h(\vartheta) < c$.

c) Pravostranná alternativa: Tvrzení H_0: $h(\vartheta) \le c$ se nazývá složená levostranná nulová hypotéza. Proti jednoduché nebo složené levostranné nulové hypotéze postavíme složenou pravostrannou alternativní hypotézu H_1: $h(\vartheta) > c$.

Testováním H_0 proti H_1 rozumíme rozhodovací postup založený na náhodném výběru $X_1, ..., X_n$, s jehož pomocí zamítneme či nezamítneme platnost nulové hypotézy.

1.4.2. Chyba 1. a 2. druhu

Při testování H_0 proti H_1 se můžeme dopustit jedné ze dvou chyb: chyba 1. druhu spočívá v tom, že H_0 zamítneme, ač ve skutečnosti platí a chyba 2. druhu spočívá v tom, že H_0 nezamítneme, ač ve skutečnosti neplatí. Situaci přehledně znázorňuje tabulka:

skutečnost	rozhodnutí	
	H_0 nezamítáme	H_0 zamítáme
H_0 platí	správné rozhodnutí	chyba 1. druhu
H_0 neplatí	chyba 2. druhu	správné rozhodnutí

Pravděpodobnost chyby 1. druhu se značí α a nazývá se hladina významnosti testu (většinou bývá $\alpha = 0,05$, méně často 0,1 či 0,01). Pravděpodobnost chyby 2. druhu se značí β. Číslo $1-\beta$ se nazývá síla testu a vyjadřuje pravděpodobnost, s jakou test vypoví, že H_0 neplatí.

1.4.3. Testování pomocí kritického oboru

Najdeme statistiku $T_0 = T_0(X_1, ..., X_n)$, kterou nazveme testovým kritériem. Množina všech hodnot, jichž může testové kritérium nabýt, se rozpadá na obor nezamítnutí nulové hypotézy (značí se V) a obor zamítnutí nulové hypotézy (značí se W a nazývá se též kritický obor). Tyto dva obory jsou odděleny kritickými hodnotami (pro danou hladinu významnosti α je lze najít ve statistických tabulkách).

Jestliže číselná realizace t_0 testového kritéria T_0 padne do kritického oboru W, pak nulovou hypotézu zamítáme na hladině významnosti α a znamená to skutečné vyvrácení testované hypotézy. Jestliže t_0 padne do oboru nezamítnutí V, pak jde o pouhé mlčení, které platnost nulové hypotézy jenom připouští.

Pravděpodobnosti chyb 1. a 2. druhu nyní zapíšeme takto:

$P(T_0 \in W/H_0 \text{ platí}) = \alpha$, $P(T_0 \in V/H_1 \text{ platí}) = \beta$.

Stanovení kritického oboru pro danou hladinu významnosti α:

Označme t_{min} (resp. t_{max}) nejmenší (resp. největší) hodnotu testového kritéria.

Kritický obor v případě oboustranné alternativy má tvar

$W = \left(t_{min}, K_{\alpha/2}(T)\right\rangle \cup \left\langle K_{1-\alpha/2}(T), t_{max}\right)$, kde $K_{\alpha/2}(T)$ a $K_{1-\alpha/2}(T)$ jsou kvantily rozložení, jímž se řídí testové kritérium T_0, je-li nulová hypotéza pravdivá.

Kritický obor v případě levostranné alternativy má tvar:

$W = \left(t_{min}, K_{\alpha}(T)\right\rangle$.

Kritický obor v případě pravostranné alternativy má tvar:

$W = \left\langle K_{1-\alpha}(T), t_{max}\right)$.

Doporučuje se dodržovat následující postup:

- Stanovíme nulovou hypotézu a alternativní hypotézu. Přitom je vhodné zvolit jako alternativní hypotézu ten předpoklad, jehož přijetí znamená závažné opatření a mělo by k němu dojít jen s malým rizikem omylu.
- Zvolíme hladinu významnosti α. Zpravidla volíme $\alpha = 0,05$, méně často 0,1 nebo 0,01.
- Najdeme vhodné testové kritérium a na základě zjištěných dat vypočítáme jeho realizaci.
- Stanovíme kritický obor.

- Jestliže realizace testového kritéria padla do kritického oboru, nulovou hypotézu zamítáme na hladině významnosti α. V opačném případě nulovou hypotézu nezamítáme na hladině významnosti α.

1.4.4. Testování pomocí intervalu spolehlivosti

Sestrojíme $100(1-\alpha)\%$ empirický interval spolehlivosti pro parametrickou funkci $h(\vartheta)$. Pokryje-li tento interval hodnotu c, pak H_0 nezamítáme na hladině významnosti α, v opačném případě H_0 zamítáme na hladině významnosti α.
Pro test H_0 proti oboustranné alternativě sestrojíme oboustranný interval spolehlivosti.
Pro test H_0 proti levostranné alternativě sestrojíme pravostranný interval spolehlivosti.
Pro test H_0 proti pravostranné alternativě sestrojíme levostranný interval spolehlivosti.

1.4.5. Testování pomocí p-hodnoty

p-hodnota udává nejnižší možnou hladinu významnosti pro zamítnutí nulové hypotézy. Je-li p-hodnota $\le \alpha$, pak H_0 zamítáme na hladině významnosti α, je-li p-hodnota $> \alpha$, pak H_0 nezamítáme na hladině významnosti α.
Způsob výpočtu p-hodnoty:
Pro oboustrannou alternativu $p = 2 \min\{P(T_0 \le t_0), P(T_0 \ge t_0)\}$.
Pro levostrannou alternativu $p = P(T_0 \le t_0)$.
Pro pravostrannou alternativu $p = P(T_0 \ge t_0)$.

p-hodnota vyjadřuje pravděpodobnost, s jakou číselné realizace $x_1, ..., x_n$ náhodného výběru $X_1, ..., X_n$ podporují H_0, je-li pravdivá. Statistické programové systémy poskytují ve svých výstupech p-hodnotu. Její výpočet vyžaduje znalost distribuční funkce rozložení, kterým se řídí testové kritérium T_0, je-li H_0 pravdivá. Vzhledem k tomu, že v běžných statistických tabulkách jsou uvedeny pouze hodnoty distribuční funkce standardizovaného normálního rozložení, bez použití speciálního software jsme schopni vypočítat p-hodnotu pouze pro test hypotézy o střední hodnotě normálního rozložení při známém rozptylu.

1.4.6. Příklad

10 x nezávisle na sobě byla změřena jistá konstanta μ. Výsledky měření byly: 2 1,8 2,1 2,4 1,9 2,1 2 1,8 2,3 2,2. Tyto výsledky považujeme za číselné realizace náhodného výběru $X_1, ..., X_{10}$ z rozložení $N(\mu, 0{,}04)$. Nějaká teorie tvrdí, že $\mu = 1{,}95$. Proti nulové hypotéze H_0: $\mu = 1{,}95$ postavíme oboustrannou alternativu H_1: $\mu \neq 1{,}95$. Na hladině významnosti 0,05 testujte H_0 proti H_1 všemi třemi popsanými způsoby.

Řešení:

$m = \frac{1}{10}(2 + ... + 2{,}2) = 2{,}06$, $\sigma^2 = 0{,}04$, $n = 10$, $\alpha = 0{,}05$, $c = 1{,}95$

a) Test provedeme pomocí kritického oboru.
Pro úlohy o střední hodnotě normálního rozložení při známém rozptylu používáme pivotovou statistiku $U = \frac{M-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$ (viz 1.3.5.). Testové kritérium tedy bude $T_0 = \frac{M-c}{\frac{\sigma}{\sqrt{n}}}$ a bude mít rozložení N(0, 1), pokud je nulová hypotéza pravdivá. Vypočítáme realizaci testového kritéria: $t_0 = \frac{2{,}06-1{,}95}{\frac{0{,}2}{\sqrt{10}}} = 1{,}74$. Stanovíme kritický obor:

$W = \left(t_{min}, K_{\alpha/2}(T)\right\rangle \cup \left\langle K_{1-\alpha/2}(T), t_{max}\right) = \left(-\infty, u_{\alpha/2}\right\rangle \cup \left\langle u_{1-\alpha/2}, \infty\right) =$
$\left(-\infty, -u_{1-\alpha/2}\right\rangle \cup \left\langle u_{1-\alpha/2}, \infty\right) = \left(-\infty, -u_{0.975}\right\rangle \cup \left\langle u_{0,975}, \infty\right) = \left(-\infty, -1{,}96\right\rangle \cup \left\langle 1{,}96, \infty\right)$. Protože 1,74 $\notin$ W, H_0 nezamítáme na hladině významnosti 0,05.

b) Test provedeme pomocí intervalu spolehlivosti.
Meze 100(1-α)% empirického intervalu spolehlivosti pro střední hodnotu μ při známém rozptylu σ^2 jsou (viz 1.3.5.): (d, h) = $\left(m - \frac{\sigma}{\sqrt{n}} u_{1-\alpha/2}, m + \frac{\sigma}{\sqrt{n}} u_{1-\alpha/2}\right)$.
V našem případě dostáváme: $d = 2{,}06 - \frac{0{,}2}{\sqrt{10}} u_{0{,}975} = 2{,}06 - \frac{0{,}2}{\sqrt{10}} .1{,}96 = 1{,}936$, h = 2,184.
Protože 1,95 $\in$ (1,936; 2,184), H_0 nezamítáme na hladině významnosti 0,05.

c) Test provedeme pomocí p-hodnoty.
Protože proti nulové hypotéze stavíme oboustrannou alternativu, použijeme vzorec
$p = 2 \min\{P(T_0 \le t_0), P(T_0 \ge t_0)\} = 2 \min\{P(T_0 \le 1{,}74), P(T_0 \ge 1{,}74)\} =$
$= 2 \min\{\Phi(1{,}74), 1 - \Phi(1{,}74)\} = 2 \min\{0{,}95907, 1 - 0{,}95907\} = 0{,}08186$. Jelikož 0,08186 > 0,05, nulovou hypotézu nezamítáme na hladině významnosti 0,05.

Kontrolní otázky

1. Vysvětlete pojem „náhodný výběr“ a „statistika“ a uveďte příklady důležitých statistik.
2. K čemu slouží bodový odhad parametrické funkce a jaké typy bodových odhadů znáte?
3. Definujte interval spolehlivosti a popište způsob jeho konstrukce.
4. Jaký vliv na šířku intervalu spolehlivosti má riziko a jaký vliv má rozsah výběru?
5. Co rozumíme pojmem „testování hypotéz“?
6. Popište nulovou a alternativní hypotézu.
7. Vysvětlete rozdíl mezi chybou 1. a 2. druhu.
8. Popište tři způsoby testování hypotéz.

Příklady

1. Nezávisle opakovaná laboratorní měření určité konstanty jsou charakterizována náhodným výběrem $X_1, ..., X_n$ z rozložení se střední hodnotou μ a rozptylem σ^2. Uvažme statistiky $M = \frac{1}{n}\sum_{i=1}^{n} X_i, L = \frac{X_1 + X_n}{2}$. Dokažte, že M a L jsou nestranné odhady konstanty μ a zjistěte, který z nich je lepší.
Výsledek:
Výpočtem zjistíme, že E(M) = μ, E(L) = μ, tudíž statistiky M a L jsou nestranné odhady konstatny μ. Pro posouzení kvality vypočteme D(M) = $\frac{\sigma^2}{n}$, D(L) = $\frac{\sigma^2}{2}$. Vidíme tedy, že pro $n \ge 3$ je lepším odhadem výběrový průměr M.

2. Nechť $X_1, ..., X_n$ je náhodný výběr z rozložení N(μ;0,04). Jaký musí být nejmenší rozsah náhodného výběru, aby šířka 95% empirického intervalu spolehlivosti pro neznámou střední hodnotu μ nepřesáhla číslo 0,16?
Výsledek: 25

3. Nechť $X_1, ..., X_9$ je náhodný výběr z rozložení $N(\mu;0,01)$. Realizace výběrového průměru je $m = 3$. Sestrojte $100(1-\alpha)\%$ empirický interval spolehlivosti pro neznámou střední hodnotu μ, je-li a) $\alpha = 0,01$, b) $\alpha = 0,05$, c) $\alpha = 0,1$.
Výsledek:
ad a) 2,914 mm < μ < 3,086 mm s pravděpodobností aspoň 0,99.
ad b) 2,935 mm < μ < 3,065 mm s pravděpodobností aspoň 0,95.
ad c) 2,945 mm < μ < 3,055 mm s pravděpodobností aspoň 0,90.
Vidíme, že s rostoucím rizikem klesá šířka intervalu spolehlivosti.
4. Nechť $X_1, ..., X_n$ je náhodný výběr z rozložení $N(\mu;0,01)$. Realizace výběrového průměru je $m = 3$. Sestrojte 95% empirický interval spolehlivosti pro neznámou střední hodnotu μ, je-li a) $n = 4$, b) $n = 9$, c) $n = 16$.
Výsledek:
ad a) 2,902 mm < μ < 3,098 mm s pravděpodobností aspoň 0,95.
ad b) 2,935 mm < μ < 3,065 mm s pravděpodobností aspoň 0,95.
ad c) 2,951 mm < μ < 3,049 mm s pravděpodobností aspoň 0,95.
Vidíme, že s rostoucím rozsahem výběru klesá šířka intervalu spolehlivosti.

5. Je známo, že výška hochů ve věku 9,5 až 10 let má normální rozložení s neznámou střední hodnotou μ a známým rozptylem $\sigma^2 = 39,112$ cm^2. Dětský lékař náhodně vybral 15 hochů uvedeného věku, změřil je a vypočítal realizaci výběrového průměru $m = 139,13$ cm. Podle jeho názoru by výška hochů v tomto věku neměla přesáhnout 142 cm s pravděpodobností aspoň 0,95. Lze tvrzení lékaře akceptovat?
Výsledek:
Testujeme H_0: $\mu \leq 142$ proti H_1: $\mu > 142$ na hladině významnosti 0,05.
Testování pomocí kritického oboru: $W = \langle 1,6449, \infty)$, realizace testového kritéria je -1,7773.
Protože testové kritérium se nerealizuje v kritickém oboru, nulovou hypotézu nezamítáme na hladině významnosti 0,05.
Testování pomocí intervalu spolehlivosti: 95% empirický levostranný interval spolehlivosti pro střední hodnotu μ je $(136,47;\infty)$. Protože číslo 142 patří do tohoto intervalu, nulovou hypotézu nezamítáme na hladině významnosti 0,05.
Testování pomocí p-hodnoty: $p = 0,9622$. Protože p-hodnota je větší než hladina významnosti 0,05, nulovou hypotézu nezamítáme na hladině významnosti 0,05.