Kapitola 1.: Základní pojmy matematické statistiky

Cíl kapitoly

Po prostudování této kapitoly budete
- rozumět pojmu „náhodný výběr“
- znát vlastnosti důležitých statistik odvozených z náhodného výběru
- znát vlastnosti bodových a intervalových odhadů parametrů a parametrických funkcí
- umět formulovat nulovou a alternativní hypotézu o parametru či parametrické funkci
- znát tři způsoby, jak testovat nulovou hypotézu proti alternativní hypotéze na dané hladině významnosti
- schopni správně naplánovat pokus
- rozeznávat jednoduché, dvojné a mnohonásobné pozorování
- v rámci dvojného pozorování rozlišovat dvouvýběrové a párové porovnávání
- v rámci mnohonásobného pozorování rozlišovat mnohovýběrové a blokové porovnávání

Časová zátěž

Na prostudování této kapitoly a splnění úkolů s ní spojených budete potřebovat asi 10 hodin studia.

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.

Abychom mohli správně vyhodnotit výsledky pokusu, musí být pokus dobře naplánován. V závislosti na záměrech experimentátora rozeznáváme několik typů uspořádání pokusů: jednoduché pozorování (zkoumají se hodnoty náhodné veličiny pozorované za týchž podmínek), dvojné pozorování (zkoumá se rozdílnost hodnot náhodné veličiny pozorované za dvojích různých podmínek) a mnohonásobné pozorování (zkoumá se rozdílnost hodnot náhodné veličiny pozorované za $r \geq 3$ různých podmínek). Podle typu uspořádání pokusu pak volíme vhodnou statistickou metodu. V tomto textu probereme pouze ty nejjednodušší typy uspořádání pokusů.

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 $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - M)^2$ 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 $X_{11},\dots,X_{1n_1}$, ..., $X_{r1},\dots,X_{rn_r}$ je r stochasticky nezávislých náhodných výběrů o rozsazích $n_1 \geq 2,\dots,n_r \geq 2$. Celkový rozsah je $n = \sum_{j=1}^{r} n_j$. Označme $M_1,\dots,M_r$ výběrové průměry a $S_1^2,\dots,S_r^2$ výběrové rozptyly jednotlivých výběrů. Nechť $c_1,\dots,c_r$ jsou reálné konstanty, aspoň jedna nenulová

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

Statistika $S_*^2 = \frac{\sum_{j=1}^{r}(n_j - 1)S_j^2}{n-r}$ 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 $R_{12} = \begin{cases} \frac{1}{n-1}\sum_{i=1}^{n}\frac{X_i - M_1}{S_1}\cdot\frac{Y_i - M_2}{S_2} & \text{pro } S_1S_2 \neq 0 \\ 0 \text{ jinak} \end{cases}$ výběrový koeficient korelace.

(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(ϑ), 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(ϑ)).

Bodovým odhadem parametrické funkce h(ϑ) je statistika $T_n = T(X_1, ..., X_n)$, která nabývá hodnot blízkých h(ϑ), 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(ϑ) 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(ϑ), ať je hodnota parametru ϑ jakákoliv.

1.3.1. Typy bodových odhadů

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

a) Řekneme, že statistika T je nestranným odhadem parametrické funkce h(ϑ), jestliže
$\forall \vartheta \in \Xi:\ E(T) = h(\vartheta).$
(Význam nestrannosti spočívá v tom, že odhad T nesmí parametrickou funkci h(ϑ) 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(ϑ), 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(ϑ), 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(ϑ), jestliže
$\forall \vartheta \in \Xi \forall \varepsilon > 0: \lim_{n\to\infty} P(|T_n - h(\vartheta)| > \varepsilon) = 0.$
(Význam konzistence spočívá v tom, že s rostoucím rozsahem výběru klesá pravděpodobnost, že odhad se bude realizovat daleko od parametrické funkce h(ϑ).)

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 $X_{11},\dots,X_{1n_1}, ..., X_{r1},\dots,X_{rn_r}$ je r stochasticky nezávislých náhodných výběrů o rozsazích $n_1 \geq 2, ..., n_r \geq 2$ z rozložení se středními hodnotami $\mu_1, ..., \mu_r$ a rozptylem σ^2. Celkový rozsah je $n = \sum_{j=1}^{r} n_j$. Nechť $c_1, ..., c_r$ jsou reálné konstanty, aspoň jedna nenulová. Pak pro libovolné hodnoty parametrů $\mu_1, \dots, \mu_r$ a σ^2 platí:

$$E\left(\sum_{j=1}^{r} c_j M_j\right) = \sum_{j=1}^{r} c_j \mu_j ,$$

$$E(S_*^{\,2}) = \sigma^2.$$

Znamená to, že lineární kombinace výběrových průměrů $\sum_{j=1}^{r} c_j M_j$ je nestranným odhadem lineární kombinace středních hodnot $\sum_{j=1}^{r} c_j \mu_j$ a vážený průměr výběrových rozptylů

$S_*^{\,2} = \dfrac{\sum_{j=1}^{r} (n_j - 1) S_j^{\,2}}{n - r}$ 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: $\forall \vartheta \in \Xi : P(D < h(\upsilon) < H) \geq 1-\alpha$.

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

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

d) Číslo α se nazývá riziko (zpravidla $\alpha = 0{,}05$, méně často 0,1 či 0,01), číslo $1 - \alpha$ 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 $1 - \alpha$. (Tvrzení, že (d,h) pokrývá $h(\vartheta)$ s pravděpodobností aspoň $1 - \alpha$ 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 $1 - \alpha$.)

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 $E(M) = \mu$ 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

$$\left(-\infty, M + \frac{\sigma}{\sqrt{n}}u_{1-\alpha}\right).$$

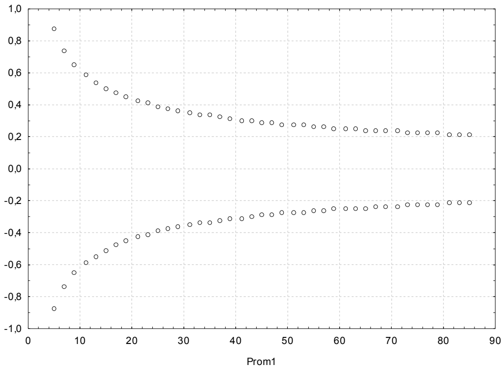
Dosadíme-li do vzorců pro dolní a horní mez číselnou realizaci m výběrového průměru M, dostaneme 100(1-α)% empirický interval spolehlivosti.

1.3.6. Šířka intervalu spolehlivosti

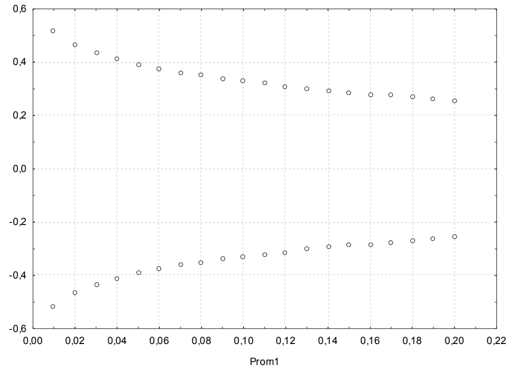
Nechť (d, h) je 100(1-α)% empirický interval spolehlivosti pro $h(\vartheta)$ zkonstruovaný pomocí číselných realizací $x_1, ..., x_n$ náhodného výběru $X_1, ..., X_n$ z rozložení $L(\vartheta)$.

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.

Závislost dolní a horní meze na rozsahu výběru (při konst. riziku)



Závislost dolní a horní meze na riziku (při konst. rozsahu výběru)



1.3.7. Příklad: 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 $\Delta \geq h - d = m + \frac{\sigma}{\sqrt{n}} u_{1-\alpha/2} - \left(m - \frac{\sigma}{\sqrt{n}} u_{1-\alpha/2} \right) = \frac{2\sigma}{\sqrt{n}} u_{1-\alpha/2}$. 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: $h(\vartheta) \leq c$ 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

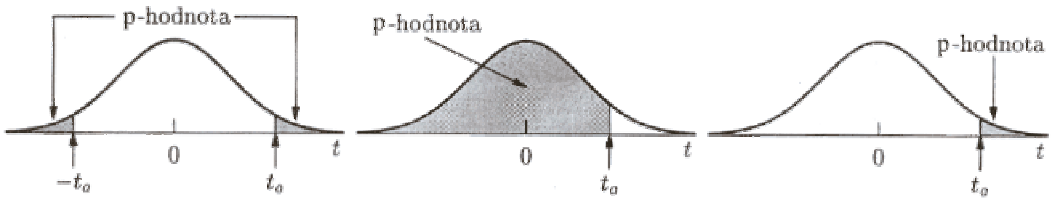
Sestrojíme $100(1-\alpha)\%$ empirický interval spolehlivosti pro parametrickou funkci $h(\vartheta)$. Pokryje-li tento interval hodnotu c, pak H_0 nezamítáme na hladině významnosti α, v opačném případě H_0 zamítáme na hladině významnosti α.
Pro test H_0 proti oboustranné alternativě sestrojíme oboustranný 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 $\leq \alpha$, 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 $p = 2\min\{P(T_0 \leq t_0), P(T_0 \geq t_0)\}$.
Pro levostrannou alternativu $p = P(T_0 \leq t_0)$.
Pro pravostrannou alternativu $p = P(T_0 \geq t_0)$.

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.

Ilustrace významu p-hodnoty pro test nulové hypotézy proti oboustranné, levostranné a pravostranné alternativě:



(Zvonovitá křivka reprezentuje hustotu rozložení, kterým se řídí testové kritérium, je-li nulová hypotéza pravdivá.)

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 $100(1-\alpha)\%$ empirického intervalu spolehlivosti pro střední hodnotu μ při známém rozptylu σ^2 jsou (viz 1.3.5.): $(d, h) = (m - \frac{\sigma}{\sqrt{n}} u_{1-\alpha/2}, m + \frac{\sigma}{\sqrt{n}} u_{1-\alpha/2})$.

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.

1.5. Uspořádání pokusů

Zaměříme se na situaci, kdy zkoumáme hmotnostní přírůstky stejně starých selat téhož plemene při různých výkrmných dietách. Určitou výkrmnou dietu aplikujeme např. po dobu půl roku. Každý den zjišťujeme hmotností přírůstky každého selete a po uplynutí půl roku vypočteme pro každé sele průměrný hmotnostní přírůstek.

1.5.1. Jednoduché pozorování

Náhodná veličina je pozorována za týchž podmínek. Situace je charakterizována jedním náhodným výběrem $X_1, ..., X_n$. (Náhodně vybereme n stejně starých selat téhož plemene, podrobíme je jediné výkrmné dietě a zjistíme hmotnostní přírůstky. Tak dostaneme realizaci jednoho náhodného výběru.)

Pokud lze očekávat, že náhodný výběr pochází z normálního rozložení, můžeme např. konstruovat interval spolehlivosti pro neznámou střední hodnotu, neznámý rozptyl či směrodatnou odchylku průměrných denních hmotnostních přírůstků nebo testovat hypotézu, že střední hodnota průměrných denních hmotnostních přírůstků neklesne pod určitou hranici. (Tyto úkoly budeme řešit ve 3. kapitole.)

1.5.2. Dvojné pozorování

Zkoumá se rozdílnost hodnot náhodné veličiny pozorované za dvojích různých podmínek. Existují dvě odlišná uspořádání tohoto pokusu.

a) Dvouvýběrové porovnávání

Situace je charakterizována dvěma nezávislými náhodnými výběry $X_{11},\ldots,X_{1n_1}$ a $X_{21},\ldots,X_{2n_2}$. (Z populace všech dostupných stejně starých selat téhož plemene náhodně vybereme $n_1 + n_2$ jedinců. Náhodně je rozdělíme na dva soubory o rozsazích n_1 a n_2, první podrobíme výkrmné dietě č. 1 a druhý výkrmné dietě č. 2. Tak dostaneme realizace dvou nezávislých náhodných výběrů.)

Za předpokladu, že dané náhodné výběry pocházejí z normálních rozložení, lze např. konstruovat interval spolehlivosti pro rozdíl středních hodnot či podíl rozptylů průměrných denních hmotnostních přírůstků nebo testovat hypotézu o stejné účinnosti obou výkrmných diet. (Tyto úkoly budeme řešit ve 4. kapitole.)

b) Párové porovnávání

Situace je charakterizována jedním náhodným výběrem (X_{11},X_{12}), ..., (X_{n1},X_{n2}) z dvourozměrného rozložení. Párem se rozumí dvojice (X_{i1},X_{i2}), $i = 1, ..., n$. Úloha se zpravidla převádí na jednoduché pozorování náhodného výběru rozdílů $X_{i1} - X_{i2}$, kde $i = 1, ..., n$. (Náhodně vybereme n vrhů stejně starých selat téhož plemene a z nich vždy dva sourozence a náhodně jim přiřadíme 1. a 2. výkrmnou dietu. Tak dostaneme realizaci náhodného výběru z dvourozměrného rozložení.)

Lze-li dvourozměrný náhodný výběr považovat za výběr z dvourozměrného normálního rozložení, budeme se zabývat konstrukcí intervalu spolehlivosti pro rozdíl středních hodnot průměrných denních hmotnostních přírůstků nebo testovat hypotézu o stejné účinnosti obou výkrmných diet. (Řešení úkolů tohoto typu je popsáno ve 3. kapitole.)

1.5.3. Mnohonásobné pozorování

Zkoumá se rozdílnost hodnot náhodné veličiny pozorované za $r \geq 3$ různých podmínek. Existují dvě odlišná uspořádání tohoto pokusu.

a) Mnohovýběrové porovnávání

Situace je charakterizována r nezávislými náhodnými výběry $X_{11},\ldots,X_{1n_1}$, ..., $X_{r1},\ldots,X_{rn_r}$. (Z populace všech dostupných stejně starých selat téhož plemene náhodně vybereme $n_1 + n_2 + ... + n_r$ jedinců. Náhodně je rozdělíme na r souborů o rozsazích n_1, n_2, ..., n_r. Selata z prvního souboru podrobíme výkrmné dietě č. 1, ..., selata z r-tého souboru podrobíme výkrmné dietě č. r. Tak dostaneme realizace r nezávislých náhodných výběrů.)

Za předpokladu, že všechny náhodné výběry se řídí normálním rozložením s týmž rozptylem, můžeme testovat hypotézu o stejné účinnosti všech r výkrmných diet. (Tomuto problému je věnována 5. kapitola.)

b) Blokové porovnávání

Situace je charakterizována jedním náhodným výběrem $(X_{11},X_{12}, ..., X_{1r})$, ..., $(X_{n1},X_{n2}, ..., X_{nr})$ z r-rozměrného rozložení. Blokem se rozumí r-tice $(X_{i1},X_{i2}, ..., X_{ir})$, $i = 1, ..., n$. (Náhodně vybereme n vrhů starých selat téhož plemene a z nich vždy r sourozenců a náhodně jim přiřadíme 1. až r-tou výkrmnou dietu. Tak dostaneme realizaci náhodného výběru z r-rozměrného rozložení.)

Vyhodnocení výsledků při blokovém porovnávání se provádí např. pomocí Friedmanova testu. Jeho popis se již vymyká náplni předmětu Statistika II. Poučení lze nalézt v doporučené literatuře [Hendl] na str. 360.

Shrnutí

Ústředním pojmem matematické statistiky je pojem *náhodného výběru*, a to jednorozměrného i vícerozměrného. Transformací jednoho nebo více náhodných výběrů vzniká náhodná veličina zvaná *(výběrová) statistika*. K nejdůležitějším statistikám patří *výběrový průměr, výběrový rozptyl, výběrová směrodatná odchylka, hodnota výběrová distribuční funkce, výběrová kovariance, výběrový koeficient korelace*.

Jelikož statistika je náhodná veličina, má smysl počítat její střední hodnotu a rozptyl. Ukázali jsme si vlastnosti střední hodnoty a rozptylu výběrového průměru či hodnoty výběrové distribuční funkce a střední hodnoty výběrového rozptylu, výběrové kovariance a výběrového koeficientu korelace.

Na základě znalosti náhodného výběru aproximujeme neznámou hodnotu parametru či parametrické funkce *bodovým odhadem parametrické funkce*. Zpravidla požadujeme, aby tento odhad měl jisté žádoucí vlastnosti. K těm patří *nestrannost*, resp. *asymptotická nestrannost* či *konzistence*, pokud pracujeme s posloupností bodových odhadů téže parametrické funkce.

Bodové odhady však mají jednu značnou nevýhodu – nevíme, s jakou pravděpodobností odhadují hodnotu neznámé parametrické funkce. Tuto nevýhodu odstraňují *intervalové odhady parametrické funkce*: jsou to intervaly, jejichž meze jsou statistiky a které s předem danou dostatečně velkou pravděpodobností pokrývají hodnotu neznámé parametrické funkce. Pokud do vzorců pro meze 100(1-α)% intervalu spolehlivosti pro danou parametrickou funkci dosadíme číselné realizace náhodného výběru, dostaneme *100(1-α)% empirický interval spolehlivosti*.

Tvrzení o parametrech rozložení, z něhož pochází daný náhodný výběr, nazýváme *nulovou hypotézou*. Proti nulové hypotéze stavíme *alternativní hypotézu*, která říká, co platí, když neplatí nulová hypotéza. Při testování nulové hypotézy proti alternativní hypotéze se můžeme dopustit buď *chyby 1. druhu* (nulovou hypotézu zamítneme, ač ve skutečnosti platí) nebo *chyby 2. druhu* (nulovou hypotézu nezamítneme, ač ve skutečnosti neplatí). Pravděpodobnost chyby 1. druhu se značí α a nazývá se hladina významnosti testu.

Klasický přístup k testování hypotéz spočívá v nalezení vhodného *testového kritéria*. Množina hodnot, jichž může testové kritérium nabýt, se rozpadá *na obor nezamítnutí nulové hypotézy* a na *kritický obor*. Tyto dva neslučitelné obory jsou odděleny *kritickými hodnotami*. Pokud se testové kritérium realizuje v kritickém oboru, nulovou hypotézu zamítáme na hladině významnosti α a přijímáme alternativní hypotézu. V opačném případě nulovou hypotézu nezamítáme na hladině významnosti α. Tím jsme ovšem neprokázali její pravdivost, můžeme pouze říci, že naše data nejsou natolik průkazná, abychom mohli nulovou hypotézu zamítnout.

Test nulové hypotézy proti alternativní hypotéze lze též provést pomocí intervalu spolehlivosti.

Máme-li k dispozici statistický software, můžeme vypočítat *p-hodnotu* jako nejmenší možnou hladinu významnosti pro zamítnutí nulové hypotézy.

Existují tři základní způsoby uspořádání pokusů:

- *jednoduché pozorování* (náhodná veličina je pozorována za týchž podmínek),
- *dvojné pozorování* (náhodná veličina je pozorována za dvojích různých podmínek, přičemž lze použít buď *dvouvýběrové porovnávání* – výsledkem jsou dva nezávislé náhodné výběry nebo *párové porovnávání* – výsledkem je jeden náhodný výběr z dvourozměrného rozložení)
- *mnohonásobné pozorování* (náhodná veličina je pozorována za $r \geq 3$ různých podmínek, přičemž lze použít buď *mnohovýběrové porovnávání* – výsledkem je $r \geq 3$ nezávislých náhodných výběrů nebo *blokové porovnávání* – výsledkem je jeden náhodný výběr z r-rozměrného rozložení).

Správnému uspořádání pokusů je zapotřebí věnovat patřičnou pozornost, neboť při nevhodném uspořádání nelze efektivně využít informace obsažené v datech a prostředky vynaložené na jejich získání jsou znehodnoceny.

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.
9. Popište tři způsoby plánování pokusů.
10. Jak se liší dvouvýběrové a párové porovnávání?
11. Jak se liší mnohovýběrové a blokové porovnávání?

Autokorekční test

1. Která z následujících tvrzení jsou pravdivá?
a) Náhodným výběrem rozumíme objekty základního souboru, které byly vybrány do výběrového souboru náhodně, např. losováním.
b) Náhodným výběrem rozumíme posloupnost stochasticky nezávislých a stejně rozložených náhodných veličin či vektorů.
c) Číselné realizace náhodného výběru uspořádané do vektoru či matice tvoří datový soubor.

2. Která z následujících tvrzení jsou pravdivá?
a) Výběrový rozptyl je aritmetickým průměrem kvadrátů centrovaných složek náhodného výběru.
b) Číselné realizace výběrového průměru se mohou výběr od výběru lišit.
c) V definici váženého průměru výběrových rozptylů hrají roli vah rozsahy jednotlivých náhodných výběrů.

3. Která z následujících tvrzení jsou pravdivá?
a) Statistika je nestranným odhadem parametrické funkce, pokud její střední hodnota je rovna této parametrické funkci, ať je hodnota parametru jakákoliv.
b) Posloupnost statistik je posloupností konzistentních odhadů parametrické funkce, pokud s rostoucím rozsahem náhodného výběru roste pravděpodobnost, že odhady se budou realizovat daleko od parametrické funkce, ať je hodnota parametru jakákoliv.
c) Máme-li dva nestranné odhady téže parametrické funkce, tak za lepší považujeme ten, který má větší rozptyl, ať je hodnota parametru jakákoliv.

4. Která z následujících tvrzení jsou pravdivá?
a) Výběrový průměr je nestranným odhadem střední hodnoty.
b) Výběrová směrodatná odchylka je nestranným odhadem směrodatné odchylky.
c) Výběrový koeficient korelace je nestranným odhadem koeficientu korelace.

5. Která z následujících tvrzení jsou pravdivá?

a) Při konstrukci intervalu spolehlivosti pro parametrickou funkci musíme znát statistiku, která je nestranným bodovým odhadem této parametrické funkce.
b) Empirický 100(1-α)% interval spolehlivosti slouží jako odhad neznámého parametrické funkce v tomto smyslu: pravděpodobnost, že tento interval pokrývá skutečnou hodnotu parametrické funkce, je aspoň 1-α.
c) Při konstantním riziku α klesá šířka empirického 100(1-α)% intervalu spolehlivosti s rostoucím rozsahem náhodného výběru.

6. Která z následujících tvrzení jsou pravdivá?
a) Kritický obor a obor nezamítnutí nulové hypotézy jsou vždy disjunktní.
b) Pravděpodobnost chyby 2. druhu lze určit na základě znalosti rizika α..
c) Pokud byla nulová hypotéza zamítnuta na hladině významnosti 0,01, byla by zamítnuta i na hladině významnosti 0,05.

7. Z následujících tří možností vyberte správnou:
Pokud u několika osob měříme krevní tlak před zátěží a po zátěži, jedná se o
a) jednoduché pozorování
b) dvouvýběrové porovnávání
c) párové porovnávání.

8. Z následujících tří možností vyberte správnou:
Náhodně vybereme dostatečný počet rodin s dětmi a zkoumáme, zda počet dětí ovlivňuje průměrné roční výdaje rodiny na průmyslové zboží. V tomto případě se jedná o
a) párové porovnávání
b) mnohovýběrové porovnávání
c) blokové porovnávání.

9. Z následujících tří možností vyberte správnou:
Náhodně vybereme dostatečný počet mužů a žen se stejným pracovním zařazením. Zkoumáme, zda pohlaví má vliv na výši průměrného ročního platu. Pro tuto situaci využijeme
a) blokové porovnávání
b) párové porovnávání
c) dvouvýběrové porovnávání.

Správné odpovědi: 1b), c) 2b) 3a) 4a) 5a) ,b), c) 6a), c) 7c) 8b) 9c)

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 $E(M) = \mu$, $E(L) = \mu$, 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 $n \geq 3$ 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(μ;0,01). Realizace výběrového průměru je m = 3. Sestrojte 100(1-α)% empirický interval spolehlivosti pro neznámou střední hodnotu μ, je-li a) α = 0,01, b) α = 0,05, c) α = 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(μ;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: μ ≤ 142 proti H_1: μ > 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;∞). 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.