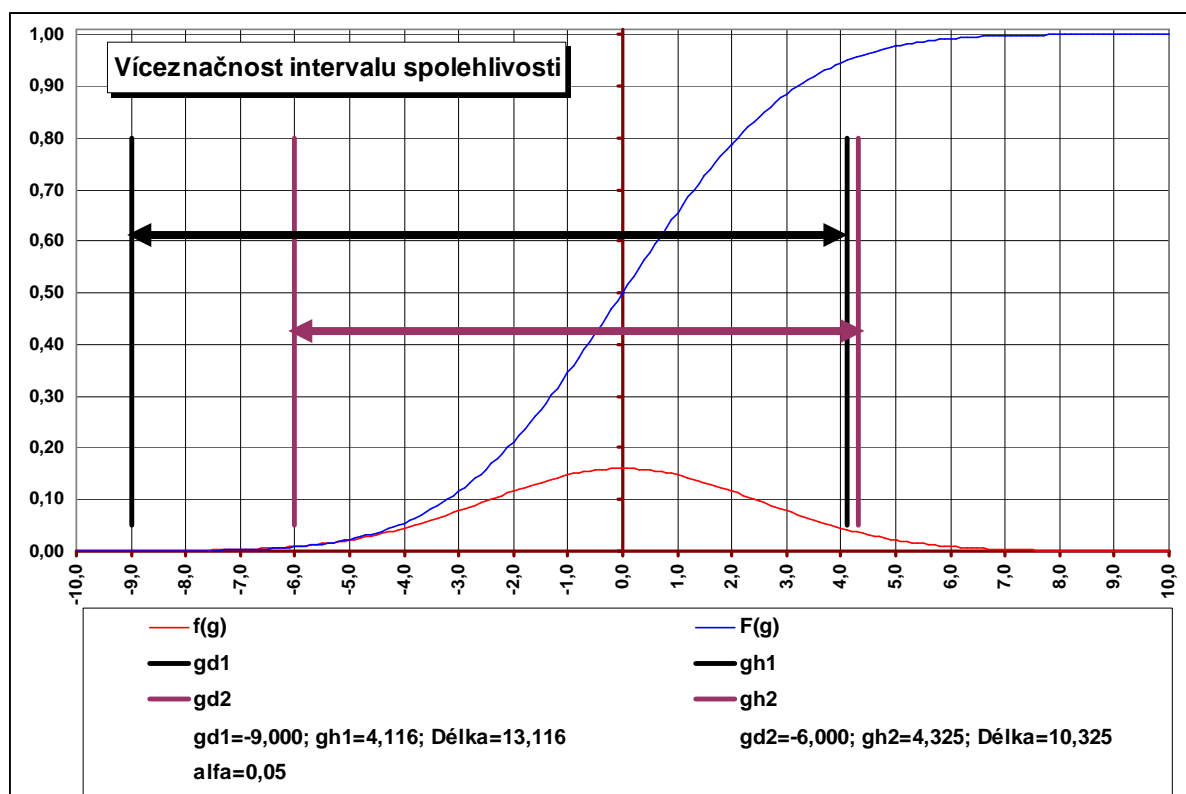


# Intervalové odhady

## Intervaly spolehlivosti

Mějme (iid) náhodný výběr  $\{x_1, x_2, \dots, x_n\}$  rozsahu  $n$  náhodné proměnné  $\xi$  s hustotou  $f(x; g)$ , kde  $g$  je hledaný (odhadovaný) parametr. Dále předpokládáme, že je dáno číslo  $1-\alpha$  nazývané koeficient spolehlivosti. Pod  $100(1-\alpha)\%$  intervalem spolehlivosti pro  $g$  (Neymanovým, konfidenčním) budeme rozumět dvojici statistik  $g_d(x_1, x_2, \dots, x_n) < g_h(x_1, x_2, \dots, x_n)$  takovou, že:  $P(g_d(x_1, x_2, \dots, x_n) < g < g_h(x_1, x_2, \dots, x_n)) \geq 1-\alpha$ .

Je zřejmé, že tímto vztahem nejsou obě meze  $g_d(x_1, x_2, \dots, x_n), g_h(x_1, x_2, \dots, x_n)$  určeny jednoznačně a to ani za případu rovnosti  $P(g_d(x_1, x_2, \dots, x_n) < g < g_h(x_1, x_2, \dots, x_n)) = 1-\alpha$ .



Dále je vhodné si uvědomit, že při tomto pojetí jsou náhodné proměnné obě meze, nikoliv hledaný parametr. Ten je v této úloze považován za neznámou konstantu.

## První metoda konstrukce intervalu spolehlivosti

1. Určíme nějakou statistiku  $T(x_1, x_2, \dots, x_n)$  odhadující parametr  $g$ .
2. Na množině  $T(X^n) \times G$ , kde  $T(X^n)$  je množina možných hodnot zvolené statistiky nad daným výběrovým prostorem a  $G$  je množina možných hodnot parametru  $g$ , zvolíme funkci  $h(T, g)$ . Funkce je zvolena tak, že rozdělení jejích hodnot  $h(T, g)$  nezávisí na

(neznámém) hledaném parametru  $g$ . Na volbu takové funkce je pak dáno i následující omezení:  $\forall t \in T(X^n)$  existuje její inverze vůči  $g$ .

3. K dané funkci  $h(T, g)$  existují čísla  $h_d, h_h$  taková, že  $P(h_d < h(T, g) < h_h) = 1 - \alpha$ . Podotýkáme, že čísla  $h_d, h_h$  nejsou dána jednoznačně (viz výše).
4. Z existence inverze vůči  $g$  pak plyne  $P(h^{-1}(T, h_d) < g < h^{-1}(T, h_h)) = 1 - \alpha$  (to za předpokladu, že funkce  $h(T, g)$  je při daném  $T$  rostoucí; **námět**: zformulujte totéž pro případ, že je klesající).
5. V takovém případě  $g_d(x_1, x_2, \dots, x_n) = h^{-1}(T(x_1, x_2, \dots, x_n), h_d)$  a  $g_h(x_1, x_2, \dots, x_n) = h^{-1}(T(x_1, x_2, \dots, x_n), h_h)$ , jsou hledané konfidenční meze.

*Tato metodika je uvedena podle Machek J.: Teorie odhadu, SPN Praha 1974, skripta MFFUK, str. 114-115.*

**Námět:** Ověřte, za jakých podmínek nepůjde splnit požadovaný předpoklad:  $\forall t \in T(X^n)$  existuje inverze  $h(T, g)$  vůči  $g$ . {Jedna z možností: uvažujte výběr  $\{x_1, x_2, \dots, x_n\}$  z alternativního rozdělení}

**Příklad:** Máme výběr z rovnoměrného rozdělení na intervalu  $\langle 0, g \rangle$ . Možnou statistikou pro odhad neznámého parametru  $g$  je:  $T(x_1, x_2, \dots, x_n) = \max(x_1, x_2, \dots, x_n)$ . Taková statistika má

pak rozdělení  $f_{(n)}(t) = nf(t)F^{n-1}(t) = n \frac{1}{g} \left(\frac{t}{g}\right)^{n-1}$ ,  $F_{(n)}(t) = (F(t))^n = \left(\frac{t}{g}\right)^n$   $0 \Leftrightarrow t \leq 0$   
 $1 \Leftrightarrow t > g$

Funkci  $h(T, g)$  můžeme zvolit např.  $h(t, g) = \frac{t}{g}$ . Potom

$$F_h(x) = P(h < x) = P\left(\frac{t}{g} < x\right) = P(t < gx) = F_n(gx) = x^n \Leftrightarrow 0 \leq x \leq 1, \\ 0 \Leftrightarrow x \leq 0 \\ 1 \Leftrightarrow x > 1$$

Pokud je dáno  $1 - \alpha$  = koeficient spolehlivosti, **můžeme**  $\alpha$  rozdělit do dvou shodných částí  $\alpha = \frac{\alpha}{2} + \frac{\alpha}{2}$  (to pro jednoznačnost volby intervalu spolehlivosti, viz dále). Inverze funkce

$h(t, g) = \frac{t}{g}$  bude  $g = \frac{t}{h}$ . Rovnost  $P(h_d < h(T, g) < h_h) = 1 - \alpha$  bude v tomto konkrétním

případě vypadat  $P\left(h_d < \frac{t}{g} < h_h\right) = \left(1 - \frac{\alpha}{2}\right) - \frac{\alpha}{2}$ . Toho lze dosáhnout **např.** volbou

$$P\left(\frac{t}{g} < h_d\right) = \frac{\alpha}{2} \text{ a } P\left(h_h < \frac{t}{g}\right) = 1 - \frac{\alpha}{2}. \text{ Po dosazení do vztahu } F_h(x) = x^n \Leftrightarrow 0 \leq x \leq 1, \\ 0 \Leftrightarrow x \leq 0 \\ 1 \Leftrightarrow x > 1$$

dostaneme:  $(h_d)^n = \frac{\alpha}{2}$  a  $(h_h)^n = 1 - \frac{\alpha}{2}$ . Řešení těchto vztahů bude  $h_d = \sqrt[n]{\frac{\alpha}{2}}$  a  $h_h = \sqrt[n]{1 - \frac{\alpha}{2}}$ .

$$\text{Dále je } P\left(h_d < \frac{t}{g} < h_h\right) = P\left(\frac{t}{h_h} < g < \frac{t}{h_d}\right) = P\left(\frac{\max(x_1, x_2, \dots, x_n)}{\sqrt[n]{1 - \frac{\alpha}{2}}} < g < \frac{\max(x_1, x_2, \dots, x_n)}{\sqrt[n]{\frac{\alpha}{2}}}\right) = 1 - \alpha.$$

Poznámka: odhad  $T(x_1, x_2, \dots, x_n) = \max(x_1, x_2, \dots, x_n)$  není nestranným, je vychýlený.

Např. pro  $\alpha = 0,05$  a  $n = 10$  bude

$$P(1,002535 \max(x_1, x_2, \dots, x_n) < g < 1,4461256 \max(x_1, x_2, \dots, x_n)) = 0,95$$

a pro  $\alpha = 0,05$  a  $n = 100$  bude

$$P(1,0002532 \max(x_1, x_2, \dots, x_n) < g < 1,0375776 \max(x_1, x_2, \dots, x_n)) = 0,95.$$

**Námět:** Pokuste se aplikovat daný postup pro nestranný odhad  $g$ ,  $\hat{g} = x_{(n)} + \frac{x_{(n)} - x_{(1)}}{n-1}$ .

**Příklad:** Máme výběr z normálního rozdělení se střední hodnotou  $\mu$  a směrodatnou odchylkou  $\sigma$  (známou). Možnou statistikou pro odhad neznámého parametru  $\mu$  je průměr:

$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ . Taková statistika má pak opět normální rozdělení se střední hodnotou  $\mu$  a

rozptylem  $\frac{\sigma^2}{n}$  a tedy s hustotou  $f_{\xi}(x) = \frac{1}{\sqrt{2\pi \frac{\sigma^2}{n}}} e^{-\frac{(x-\mu)^2}{2 \frac{\sigma^2}{n}}}$ . Funkci  $h(\bar{x}, \mu)$  můžeme zvolit

např.  $h(\bar{x}, \mu) = \bar{x} - \mu$ . Funkce  $\bar{x} - \mu$  má hustotu  $f_{\xi}(h) = \frac{1}{\sqrt{2\pi \frac{\sigma^2}{n}}} e^{-\frac{h^2}{2 \frac{\sigma^2}{n}}}$  je tedy nezávislá

(funkčně) na neznámém parametru  $\mu$ . Pro koeficient spolehlivosti  $\alpha = \frac{\alpha}{2} + \frac{\alpha}{2}$  použijeme

obdobný rozklad jako v předchozím případě. Proto

$$1 - \alpha = P(h_d < h(\bar{x}, \mu) < h_h) = P\left(\frac{\sqrt{n}h_d}{\sigma} < \frac{\sqrt{n}(\bar{x} - \mu)}{\sigma} < \frac{\sqrt{n}h_h}{\sigma}\right) = \Phi\left(\frac{\sqrt{n}h_h}{\sigma}\right) - \Phi\left(\frac{\sqrt{n}h_d}{\sigma}\right), \text{ kde } \Phi(x) \text{ je}$$

distribuční funkce centrovaného a normovaného rozdělení  $N(0,1)$ . Rovnost

$$1 - \alpha = \Phi\left(\frac{\sqrt{n}h_h}{\sigma}\right) - \Phi\left(\frac{\sqrt{n}h_d}{\sigma}\right) \text{ lze splnit více způsoby, běžný je } 1 - \frac{\alpha}{2} = \Phi\left(\frac{\sqrt{n}h_h}{\sigma}\right) \text{ a } \frac{\alpha}{2} = \Phi\left(\frac{\sqrt{n}h_d}{\sigma}\right).$$

Její řešení je  $h_h = \frac{\sigma}{\sqrt{n}} \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$  a  $h_d = \frac{\sigma}{\sqrt{n}} \Phi^{-1}\left(\frac{\alpha}{2}\right)$ . Např. pro  $\alpha = 0,05$  je

$\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) = 1,95996$  a  $\Phi^{-1}\left(\frac{\alpha}{2}\right) = -1,95996$ . Řešením rovnice  $1 - \alpha = P(h_d < h(\bar{x}, \mu) < h_h)$  je

pak, včetně inverze funkce  $h(\bar{x}, \mu)$ ,  $1 - \alpha = P\left(\bar{x} + \Phi^{-1}\left(\frac{\alpha}{2}\right) \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \frac{\sigma}{\sqrt{n}}\right)$ .

Konkrétně pro  $\alpha = 0,05$  je  $0,95 = P\left(\bar{x} - 1,95996 \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + 1,95996 \frac{\sigma}{\sqrt{n}}\right)$ .

**Námět:** stanovte takové intervalové odhady za podmínek tohoto příkladu pro  $\alpha = 0,1, 0,05, 0,025, 0,01, 0,005, 0,001$ .

**Příklad:** Máme opět výběr z normálního rozdělení se střední hodnotou  $\mu$  a směrodatnou odchylkou  $\sigma$  (tentokrát neznámou). Možnou statistikou (opět) pro odhad neznámého parametru  $\mu$  je průměr:  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ . Možnou statistikou pro odhad neznámého parametru  $\sigma$

je výběrová směrodatná odchylka:  $s = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$ . Tentokrát se zvolí funkce

$h(\bar{x}, \mu) = \frac{\bar{x} - \mu}{s} \sqrt{n}$ . Její hodnoty nezávisí na neznámých parametrech  $\mu$ ,  $\sigma$  a

$h(\bar{x}, \mu)$  má Studentovo **t-rozdělení** s **n-1** stupni volnosti (dokažte) s hustotou  $f(t, n-1)$  (námět: pokuste se odvodit její explicitní tvar {podíl normálně rozdělené náhodné proměnné a

$\chi^2(n-1)$  náhodné proměnné}). Dále nalezneme řešení rovnice  $\int_{-t_\alpha}^{t_\alpha} f(t, n-1) dt = 1 - \alpha$  (námět: jak vypadá souvislost takového řešení s kvantily Studentova t-rozdělení s n-1 stupni volnosti).

Potom:  $1 - \alpha = P\left(-t_\alpha < \frac{\bar{x} - \mu}{s} \sqrt{n} < t_\alpha\right) = P\left(\bar{x} - t_\alpha \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_\alpha \frac{s}{\sqrt{n}}\right)$  a tím dostaneme intervalový odhad střední hodnoty.

**Srovnání obou příkladů:**

	Dolní mez pro střední hodnotu	Horní mez pro střední hodnotu
$\sigma$ známé	$\bar{x} + \Phi^{-1}\left(\frac{\alpha}{2}\right) \frac{\sigma}{\sqrt{n}}$	$\bar{x} + \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \frac{\sigma}{\sqrt{n}}$
$\sigma$ neznámé	$\bar{x} - t_\alpha \frac{s}{\sqrt{n}}$	$\bar{x} + t_\alpha \frac{s}{\sqrt{n}}$

Protože platí  $u_{1-\alpha/2} \stackrel{\text{def}}{=} \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) = -\Phi^{-1}\left(\frac{\alpha}{2}\right) \stackrel{\text{def}}{=} -u_{\alpha/2}$  (ověřte!), můžeme výše uvedenou tabulku přepsat:

	Dolní mez pro střední hodnotu	Horní mez pro střední hodnotu
$\sigma$ známé	$\bar{x} - u_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}$	$\bar{x} + u_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}$
$\sigma$ neznámé	$\bar{x} - t_\alpha \frac{s}{\sqrt{n}}$	$\bar{x} + t_\alpha \frac{s}{\sqrt{n}}$

Oba odhady se pak liší jen v tom, že při známém  $\sigma$  použijeme toto  $\sigma$  s „korekčním“ faktorem  $u_{1-\alpha/2}$  a při neznámém  $\sigma$  použijeme jeho odhad  $s = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$  s „korekčním“ faktorem  $t_\alpha$  (námět: vyjádřete  $t_\alpha$  pomocí kvantilů Studentova t-rozdělení s  $n-1$  stupni volnosti).

**Námět:** srovnajte tabulku  $u_{1-\alpha/2}$  a  $t_\alpha$  pro různá  $n$  a různá  $\alpha$ . Od jakého  $n$  se oba korekční faktory nebudou prakticky lišit (tj. od jakého  $n$  lze předpokládat znalost  $\sigma$  jeho náhradou za  $s$ ). Pozor, taková aproximace je danou úvahou přijatelná jen a jen pro intervalový odhad střední hodnoty normálního rozdělení).

**Příklad:** Máme opět výběr z normálního rozdělení se střední hodnotou  $\mu$  (neznámou) a směrodatnou odchylkou  $\sigma$  (také neznámou). Možnou statistikou (opět) pro odhad neznámého parametru  $\mu$  je průměr:  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ . Možnou statistikou pro odhad neznámého parametru  $\sigma^2$  je výběrový rozptyl:  $s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$  (ten je nejlepším nestranným odhadem  $\sigma^2$ ). Úkolem je sestavit intervalový odhad pro parametr  $\sigma^2$ .

Veličina  $(n-1) \frac{s^2}{\sigma^2} = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \bar{x})^2$  má  $\chi^2(n-1)$  rozdělení (dokažte: jde vlastně o násobek součtu kvadrátů stejně a normálně rozdělených nezávislých náhodných proměnných). Proto platí:  $1-\alpha = P(h_d < h(\sigma^2, s^2) < h_h) = P\left(h_d < (n-1) \frac{s^2}{\sigma^2} < h_h\right) = P\left(\frac{(n-1)s^2}{\chi_{1-\alpha/2}^2(n-1)} < \sigma^2 < \frac{(n-1)s^2}{\chi_{\alpha/2}^2(n-1)}\right)$ , kde  $\chi_{1-\alpha/2}^2(n-1)$  je  $1-\alpha/2$  kvantil  $\chi^2(n-1)$  rozdělení (s  $n-1$  stupni volnosti) a  $\chi_{\alpha/2}^2(n-1)$  je  $\alpha/2$  kvantil téhož rozdělení.

Z toho, co bylo v tomto příkladu odvozeno, přímo plyne pro odhad parametru  $\sigma$   $1-\alpha = P\left(\sqrt{\frac{(n-1)s^2}{\chi_{1-\alpha/2}^2(n-1)}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi_{\alpha/2}^2(n-1)}}\right)$  a to protože odmocnina z nezáporného čísla je rostoucí funkcí a tedy zachovává nerovnosti.

**Námět:** Odvoďte interval spolehlivosti pro neznámý parametr  $\sigma^2$  u výběru z normálního rozdělení se známou střední hodnotou  $\mu$ .

**Námět:** Odvoďte všechny výše uvedené vztahy pro obecnou dekompozici  $\alpha = \alpha_1 + \alpha_2$  a zvláště pro volby  $(\alpha_1 = 0) \vee (\alpha_2 = 0)$  jednostranné (v pravděpodobnosti) intervaly spolehlivosti). Zde se bude pracovat s kvantily  $p_{1-\alpha_2}$ ,  $p_{\alpha_1}$ .

**Námět:** Odvoďte všechny výše uvedené vztahy pro obecnou dekompozici  $\alpha = \alpha_1 + \alpha_2$  a pro volbu  $1-\alpha = P(h-\varepsilon < h(\bar{x}, \mu) < h+\varepsilon)$ .

**Námět:** Odvoďte všechny výše uvedené vztahy pro obecnou dekompozici  $\alpha = \alpha_1 + \alpha_2$  a pro volbu  $P(h^{-1}(T, h_d) < g < h^{-1}(T, h_h)) = 1 - \alpha$ , kde  $h^{-1}(T, h_d) = k - \varepsilon$  a  $h^{-1}(T, h_h) = k + \varepsilon$  (polohově symetrické intervaly spolehlivosti).

**Námět:** Odvoďte všechny výše uvedené vztahy pro obecnou dekompozici  $\alpha = \alpha_1 + \alpha_2$  a pro volbu  $P(h^{-1}(T, h_d) < g < h^{-1}(T, h_h)) = 1 - \alpha$ , kde  $h_d, h_h$  jsou vybrány tak, aby  $|h^{-1}(T, h_d) - h^{-1}(T, h_h)| \rightarrow \min$  (nejkratší intervaly spolehlivosti).

## Další (obecnější) metoda konstrukce intervalu spolehlivosti

Taková konstrukce je založena na následujícím tvrzení:

Mějme náhodný výběr z rozdělení  $f(x; g); g \in \langle g_{\min}, g_{\max} \rangle$  o rozsahu  $n$ . Dále mějme statistiku  $T(x_1, x_2, \dots, x_n)$ . Pro ni zavedeme  $F(t, g) = P(T \leq t; g); G(t, g) = P(T \geq t; g)$ . Pro funkci  $F(t, g)$  předpokládejme splnění následujících podmínek:

1.  $\forall t \in T(X^n)$  je  $F(t, g)$  spojitá klesající funkce  $g$ .
2.  $\lim_{g \rightarrow g_{\min}} F(t, g) = 1$  a  $\lim_{g \rightarrow g_{\max}} F(t, g) = 0$ .

Pro daná čísla  $0 < \alpha_1, \alpha_2 < 1$  označme funkce  $g_d(t), g_h(t)$  definované vztahy  $F(t, g_h(t)) = \alpha_1$  a  $G(t, g_d(t)) = \alpha_2$ .

Pak platí  $P(g \geq g_h(T); g) \leq \alpha_1$  a  $P(g \leq g_d(T); g) \leq \alpha_2$ . Pokud je takové rozdělení statistiky  $T(x_1, x_2, \dots, x_n)$  spojitě, pak v obou nerovnostech platí rovnost. Z uvedených formulací je zřejmé, že  $g_h(T)$  je  $100(1 - \alpha_1)\%$  horní hranice spolehlivosti a obdobně  $g_d(T)$  je  $100(1 - \alpha_2)\%$  dolní hranice spolehlivosti pro „neznámý“ parametr  $g$ . Dvojice  $g_d(T), g_h(T)$  je tedy  $100(1 - \alpha_1 - \alpha_2)\%$  intervalem spolehlivosti.

*Tvrzení je převzato z: Machek J., Teorie odhadu, SPN Praha 1974, skripta MFFUK, str. 122-123, kde je i uveden důkaz, založený na následující skutečnosti:*

Protože dle uvedeného předpokladu je  $F(t, g)$  spojitou klesající funkcí v  $g$  je nerovnost  $g \geq g_h(T)$  ekvivalentní nerovnosti  $F(t, g) \leq F(t, g_h(T))$ . Ale pro  $F(t, g_h(t))$  platí  $F(t, g_h(t)) = \alpha_1$ , proto je nerovnost  $g \geq g_h(T)$  ekvivalentní i nerovnosti  $F(t, g) \leq \alpha_1$ . Odtud  $P(g \geq g_h(T); g) = P(F(t, g) \leq \alpha_1; g)$ . Protože  $F(t, g) = P(T \leq t; g)$  je i  $P(g \geq g_h(T); g) = P(P(T \leq t; g) \leq \alpha_1; g) = \alpha_1$ , neboť „pravděpodobnost pravděpodobnosti je pravděpodobnost“. Shrnutí:  $P(g \geq g_h(T); g) = \alpha_1$  (toto platí pro situaci kdy je rozdělení statistiky  $T$  spojitě). Pro případ, kdy rozdělení statistiky  $T$  není spojitě, dostaneme:

$$P(g \geq g_h(T); g) = P(F(t, g) \leq \alpha_1; g) = P(T \leq \sup \{t : F(t, g) \leq \alpha_1\}; g) \leq \alpha_1$$

**Námět:** dokončete důkaz i pro druhou nerovnost.

**Námět:** modifikujte uvedené tvrzení pro situaci, kdy bude platit  $\forall t \in T(X^n)$  je  $F(t, g)$  spojitá rostoucí funkce  $g$ , včetně důkazu (modifikujte pro tento případ i druhý předpoklad).

**Námět:** k čemu je třeba v původním tvrzení předpoklad  $\lim_{g \rightarrow g_{\min}} F(t, g) = 1$  a  $\lim_{g \rightarrow g_{\max}} F(t, g) = 0$ , jak bude vypadat tento předpoklad pro modifikaci z předcházejícího námětu?

**Příklad:** Mějme náhodný výběr z alternativního rozdělení  $x_1, x_2, \dots, x_n$ , kde  $P(x_i = 1) = p = P(A); P(x_i = 0) = 1 - p = P(\bar{A})$ , kde  $A$  je sledovaný jev a pozorovaná alternativní náhodná proměnná je indikátorem jeho realizace. Potom statistika  $T(x_1, x_2, \dots, x_n) = \sum_{i=1}^n x_i$  = počet realizací pozorovaného jevu během  $n$  nezávislých experimentů, má binomické rozdělení:  $P(T = t) = \binom{n}{t} p^t (1-p)^{n-t}$ , kde hledáme pro  $p$  jeho interval spolehlivosti.

Funkce  $F(t, p)$  má pak tvar  $F(t, p) = P(T \leq t; p) = \sum_{k=0}^t \binom{n}{k} p^k (1-p)^{n-k}$  a funkce  $G(t, p)$  má pak tvar  $G(t, p) = P(T \geq t; p) = \sum_{k=t}^n \binom{n}{k} p^k (1-p)^{n-k}$ .

Dolní a horní mez pro zvolená čísla  $0 < \alpha_1, \alpha_2, \alpha_1 + \alpha_2 < 1$  nalezneme řešením rovnic:

$$\sum_{i=0}^T \binom{n}{i} p_h^i (1-p_h)^{n-i} = \alpha_1 \text{ a } \sum_{i=T}^n \binom{n}{i} p_d^i (1-p_d)^{n-i} = \alpha_2$$

Tyto dva vztahy je možné, v současnosti, řešit velmi jednoduše numericky (námět: navrhnete konkrétní metodu). Ve starší a klasické statistické literatuře jsou jejich řešení nalézána pomocí některých transformací na kvantily Fischer-Snedecorova F-rozdělení např.:

Jaroslav Hátle, Jiří Likeš: Základy počtu pravděpodobnosti a matematické statistiky. SNTL Praha 1974. Str. 236.  
Machek J.: Teorie odhadu, SPN Praha 1974, skriptá MFFUK, str. 124-126.

**Příklad:** Mějme náhodný výběr z Poissonova rozdělení  $x_1, x_2, \dots, x_n$ , kde

$P(x_i = k) = e^{-\lambda} \frac{\lambda^k}{k!}$ . Potom statistika  $T(x_1, x_2, \dots, x_n) = \sum_{i=1}^n x_i$  = počet realizací pozorovaných Poissonovských událostí v  $n$  nezávislých experimentech, má rozdělení  $P(T = k) = e^{-n\lambda} \frac{(n\lambda)^k}{k!}$ . Hledáme interval spolehlivosti pro neznámý parametr  $\lambda$ .

Dolní a horní mez pro zvolená čísla  $0 < \alpha_1, \alpha_2 < 1$  nalezneme řešením rovnic:

$$\sum_{i=0}^T e^{-n\lambda_h} \frac{(n\lambda_h)^i}{i!} = \alpha_1 \text{ a } \sum_{i=T}^{+\infty} e^{-n\lambda_d} \frac{(n\lambda_d)^i}{i!} = \alpha_2$$

Opět první z rovnic lze bez problémů pomocí současných prostředků řešit numericky, druhá zdánlivě ne. Ovšem jen zdánlivě, pokud si uvědomíme, že:

$$\sum_{i=T}^{+\infty} e^{-n\lambda_d} \frac{(n\lambda_d)^i}{i!} = 1 - \sum_{i=0}^{T-1} e^{-n\lambda_d} \frac{(n\lambda_d)^i}{i!} = \alpha_2.$$

Znova, ve starší a klasické statistické literatuře jsou jejich řešení nalézána pomocí některých transformací (zde pomocí tvrzení o vyjádření binomických a Poissonovských pravděpodobností) na kvantily  $\chi^2$  rozdělení např.:

Jaroslav Hátle, Jiří Likeš: Základy počtu pravděpodobnosti a matematické statistiky. SNTL Praha 1974. Str. 241.  
Machek J.: Teorie odhadu, SPN Praha 1974, skriptá MFFUK, str. 126-128.

Doporučená a zdrojová literatura:

Jiří Reif	Metody matematické statistiky, ZČU v Plzni 2004
Jaroslav Hátle, Jiří Likeš	Základy počtu pravděpodobnosti a matematické statistiky. SNTL Praha 1974.
C. Radhakrishna Rao	Lineární metody statistické indukce a jejich aplikace, ACADEMIA, Praha 1978
Machek J.	Teorie odhadu, SPN Praha 1974, skripta MFFUK