Josef Machek

Poznámka k teorii intervalového odhadu

Acta Universitatis Carolinae. Mathematica et Physica, Vol. 3 (1962), No. 1, 25--38

Persistent URL: http://dml.cz/dmlcz/142145

1962 ACTA UNIVERSITATIS CAROLINAE MATHEMATICA ET PHYSICA NO 1. PAG. 25—38

POZNÁMKA K TEORII INTERVALOVÉHO ODHADU

JOSEF MACHEK

Katedra matematické statistiky na matematicko-fysikální fakultě.

Poznámka pojednává o vlastnostech intervalových odhadů, odvozených z testů statistických hypotéz s určitými optimálními vlastnostmi.

Intervalem spolehlivosti se ve statistické teorii odhadu rozumí interval, jehož koncové body, $\underline{\Theta}$, $\overline{\Theta}$, jsou funkcemi náhodného výběru (a tedy náhodnými veličinami), který má tu vlastnost, že skutečná hodnota odhadovaného parametru Θ v příslušném rozdělení je jím pokryta s předem zvolenou pravděpodobností $1 - \alpha$,

$$P\{\underline{\Theta} \leqq \Theta \leqq \overline{\Theta}\} = 1 - \alpha$$

nezávisle na skutečné hodnotě Θ. Číslo $1 - \alpha$ se nazývá koeficientem spolehlivosti. Obvykle se interval spolehlivosti konstruuje některým z následujících dvou postupů:

1. zvolí se vhodná výběrová charakteristika pro odhadovaný parametr, $T = T(X_1, \ldots, X_n)$, kde $X_1, \ldots, X_n$ jsou prvky náhodného výběru (tj. navzájem nezávislé náhodné veličiny se stejným rozdělením); tato charakteristika se standardisuje tak, aby její rozdělení nezáviselo na odhadovaném parametru, tzn., zavede se $T' = T'(T, \Theta)$, funkce charakteristiky T a odhadovaného parametru Θ, jejíž rozdělení nezávisí na Θ. Pak lze najít hodnoty T_1' a T_2' s vlastností

$$P\{T_1' \leqq T'(T, \Theta) \leqq T_2'\} = 1 - \alpha .$$

Do $T'(T,\Theta)$ se dosadí pozorovaná hodnota charakteristiky T), tj. hodnota, vypočtená z výběru), a je-li možné, invertují se nerovnosti

$$T'_1 \leqq T'(T, \Theta) \leqq T'_2$$

vzhledem k Θ.

2) Zvolí se opět vhodná charakteristika T pro odhad Θ, najde se její distribuční funkce

$$F_T(t, \Theta) = P\{T \leqq t\} .$$

Tato distribuční funkce závisí na Θ, což je v zápisu vyjádřeno. Je-li $F_T(t, \Theta)$ při každém t klesající spojitou funkcí Θ, pak kořeny $\underline{\Theta}$, $\overline{\Theta}$ resp. rovnic

$$F_T(T, \overline{\Theta}) = \alpha_1, \quad F_T(T - 0, \underline{\Theta}) = 1 - \alpha_2$$

jsou hranicemi intervalu spolehlivosti pro Θ s koeficientem spolehlivosti $1 - \alpha_1 - \alpha_2$; α_1 je pravděpodobnost, že celý interval bude vlevo od skutečné hodnoty, α_2 je pravděpodobnost, že celý interval bude vpravo od skutečné hodnoty.

Další způsob konstrukce intervalu spolehlivosti souvisí těsně s teorií ověřování (testování) statistických hypotéz. Tvrdí-li hypotéza, že parametr Θ má hodnotu rovnou danému číslu Θ_0, je snadné se přesvědčit, že rozhodovací pravidlo „přijmout hypotézu $\Theta = \Theta_0$, když interval spolehlivosti $\underline{\Theta}, \overline{\Theta}$ obsahuje bod Θ_0, a zamítnout $\Theta = \Theta_0$, když bod Θ_0 leží mimo interval spolehlivosti $\underline{\Theta}, \overline{\Theta}$" je testem hypotézy $\Theta = \Theta_0$ s hladinou významnosti rovnou doplňku koeficientu spolehlivosti do 1. Značíme-li tedy koeficient spolehlivosti $1 - \alpha$, je hladina významnosti testu, sdruženého s danou metodou konstrukce intervalu spolehlivosti, rovna α. Naopak se z této skutečnosti nabízí možnost sdružit s každým testem statistické hypotézy o hodnotě parametru interval spolehlivosti, totiž interval, jehož dolní hranici tvoří infimum množiny všech možných hodnot Θ_0, pro které test při daném pozorování nevede k zamítnutí hypotézy $\Theta = \Theta_0$, a horní hranici supremum téže množiny. Je třeba jen se postarat o to, aby množina všech možných hodnot parametru „přijatelných" podle zvoleného testu byla intervalem; to prakticky znamená omezit se na určitou skupinu testů s jistými vlastnostmi. (Jinak by koeficient spolehlivosti nebyl doplňkem příslušné hladiny významnosti).

Tento třetí způsob konstrukce intervalu spolehlivosti byl již bezpochyby aplikován; je nejspíše mlčky znám a uznáván. Na souvislost odhadů a testů ostatně poukázali NEYMAN a PEARSON ve svých pracích o intervalech spolehlivosti.

Přesto se jeví účelným věnovat této konstrukci poznámku, explicitně popsat konstrukci intervalu spolehlivosti a explicitně vyslovit jejich vlastnosti. I z jiných důvodů zaslouží tento postup pozornosti.

Při výkladu teorie intervalového odhadu se totiž souvislosti intervalů spolehlivosti s testy hypotéz často nevyužívá (jak ukazují učebnice např. [1]), nebo se jí užívá spíše opačným směrem (SCHMETTERER [2]), ačkoliv teorie ověřování hypotéz je zpravidla dosti podrobně známa. Teorie ověřování hypotéz dává pro řadu praktických situací návod k volbě testu s dobře odůvodněnými optimálními vlastnostmi, naproti tomu interval spolehlivosti konstruuje aplikující statistik často intuitivně, aniž si uvědomuje přednosti nebo nedostatky svého postupu. Proto se jeví prospěšným přímo vyslovit vlastnosti intervalů spolehlivosti, odvozených z různých testů, podat jednoduché důkazy těchto vlastností a uvést podmínku pro to, aby z testu vyplynul skutečně *interval* spolehlivosti (a ne množina jiného tvaru).

Podrobněji vyslovena zní úloha intervalového odhadu takto: je dán systém $\mathbf{F}$ distribučních funkcí $\mathrm{F}_\omega(\mathrm{x})$ definovaných na vícerozměrném euklidovském prostoru W, který budeme nazývat výběrovým prostorem, $\mathbf{F} = \{\mathrm{F}_\omega(\mathbf{x}) : \omega \in \Omega\}$. Distribuční funkce systému $\mathbf{F}$ se od sebe liší indexy ω; množinu Ω všech možných indexů nazveme parametrickým prostorem. Index ω nemusí být číslo, může to být též dvojice čísel apod. Systémem $\mathbf{F}$ může například být množina všech n-rozměrných normálních rozdělení s momentovou maticí tvaru $\sigma^2 . \mathrm{E}$ (kde E je jednotková matice typu $\mathrm{n} \times \mathrm{n}$ a σ^2 konstanta), s vektorem středních hodnot $\mathrm{\mu} = (\mu, \ldots, \mu)$. Parametrickým prostorem pak bude množina dvojic (μ, σ^2). S tímto systémem máme co dělat při odhadu parametrů jednorozměrného normálního rozdělení z n navzájem nezávislých pozorování. Budeme značit prvky W malými tučnými písmeny, náhodnou veličinu (vektorovou) nabývající hodnot z W velkým tučným písmenem.

Na parametrickém prostoru je definována reálná funkce Θ (ω), zobrazující prostor Ω na interval T, který může být i neohraničený. V právě uvedeném příkladě může být třebas $\Theta(\mu, \sigma^2) = \mu$, nebo $\Theta(\mu, \sigma^2) = \sigma$, nebo $\Theta(\mu, \sigma^2) = \mu + k\sigma$ (kde k je daná konstanta), nebo $\Theta(\mu, \sigma^2) = \Phi\left(\frac{H - \mu}{\sigma}\right)$ kde H je dané číslo a Φ značí distribuční funkci normálního rozdělení s nulovou střední hodnotou a s jednotkovým rozptylem).

Je žádán interval spolehlivosti pro $\Theta(\omega)$, tzn. dvě funkce $\underline{\Theta}(\mathbf{x})$, $\overline{\Theta}(\mathbf{x})$ na výběrovém prostoru, s vlastností

$$\inf_{\omega \in \Theta^{-1}(\Theta_0)} P\{\underline{\Theta}(\mathbf{X}) \leqq \Theta_0 \leqq \overline{\Theta}(\mathbf{X})\} = 1 - \alpha$$

kde Θ_0 je skutečná hodnota parametru.

Interval spolehlivosti pro $\Theta(\omega)$ má být odvozen z testu hypotézy o Θ. Aby byla k dispozici soustava testů, umožňující ověřit hypotézu $\Theta(\omega) = t$ pro jakékoliv $t \in T$, je třeba vybrat ve výběrovém prostoru vhodné podmnožiny $w(t) \subset W$, $t \in T$, takové, že hypotéza $\Theta(\omega) = t$ se zamítne, když $\mathbf{X} \in w(t)$. Tato podmnožina $w(t)$ je tzv. *kritický obor pro test hypotézy* $\Theta(\omega) = t$ a volí se tak, aby byly vhodným způsobem omezeny pravděpodobnosti zamítnutí hypotézy $\Theta(\omega) = t$, když ta je správná a jejího nezamítnutí, když ve skutečnosti $\Theta(\omega) \neq t$.

Pro konstrukci intervalů spolehlivosti bude výhodné užívat kritických oborů s některými zvláštními vlastnostmi. Především budeme žádat, aby

1) $\sup_{\omega \in \Theta^{-1}(t)} P\{\mathbf{X} \in w(t)\} = \sup_{\omega \in \Theta^{-1}(t)} \int_{w(t)} dF_\omega(\mathbf{x}) = \alpha$ pro všechna $t \in T$,

kde $\Theta^{-1}(t)$ značí $\{\omega : \Theta(\omega) = t\}$ a α je zvolená konstanta. Kritický obor $w(t)$, splňující (1) nazveme *kritickým oborem velikosti* α. Je jasné, že nebude pro všechny úlohy existovat systém kritických oborů $w(t)$ velikosti α, zpravidla bude možno podmínku (1) splnit pro všechna $t \in T$ jen při tzv. rozděleních spojitého typu.

Dále budeme užívat jen kritických oborů, jež mají ještě vlastnost, kterou nazveme důsledností, a kterou definujeme takto:

Jestliže systém kritických oborů $w(t)$ splňuje podmínku $t_1 > t_2 \Rightarrow w(t_1) \subset w(t_2)$ pro každou dvojici čísel t_1, t_2 z T, $t_1 > t_2$, nazveme jej systémem důsledným vůči pravostranným alternativám.

Jestliže systém kritických oborů $w(t)$ splňuje $t_1 < t_2 \Rightarrow w(t_1) \subset w(t_2)$, pro každou dvojici $t_1 < t_2$, nazveme jej důsledným vůči levostranným alternativám.

Jestliže systém kritických oborů $w(t)$ splňuje podmínku, že pro každou trojici $t_1 < t_0 < t_2$ jest

$$w(t_0) \subset [w(t_1) \cup w(t_2)]$$

při čemž oba průniky $w(t_1) \cap w(t_0)$ i $w(t_2) \cap w(t_0)$ jsou neprázdné, nazveme jej důsledným vůči oboustranným alternativám. Vlastnost důslednosti, jak byla právě definována, představuje přijatelný požadavek na kritické obory, určené k testování hypotézy $\Theta(\omega) = \Theta_0$ proti skupinám alternativ $\Theta(\omega) > t$, resp. $\Theta(\omega) < t$, resp. $\Theta(\omega) \neq t$, to znamená kritické obory určené k testování hypotézy $\Theta(\omega) = t$ v situaci, kdy je pokládáno za chybu nezamítnutí $\Theta(\omega) = t$ jen když ve skutečnosti $\Theta(\omega) > t$, ale nikoli když $\Theta(\omega) < t$, či naopak, resp. když je chybou nezamítnout $\Theta(\omega) = t$ ať je odchylka na kteroukoliv stranu.

Požadavek důslednosti vůči pravostranným alternativám totiž znamená: vede-li bod x ve výběrovém prostoru k zamítnutí hypotézy $\Theta(\omega) = t_1$, vede tím spíše k zamítnutí hypotézy $\Theta(\omega) = t_2$ pro jakékoliv $t_2 < t_1$, tj. jestliže při pozorování x zamítáme hypotetickou hodnotu t_1, pokládáme za nepřijatelné i všechny menší. Důslednost vůči oboustranným alternativám znamená: vede-li určitý bod x ve výběrovém prostoru k zamítnutí hypotetické hodnoty t_0 pak vede buď také k zamítnutí všech hodnot nižších, nebo také k zamítnutí všech hodnot vyšších.

Příklad: Budiž X vektor navzájem nezávislých náhodných veličin s rozdělením normálním se střední hodnotou μ a s rozptylem 1. Systém kritických oborů velikosti α, $w(t) = \left\{\mathrm{x} : \sum_{i=1}^{n} x_i - t > u_{1-\alpha}\sqrt{n}\right\}$ kde $u_{1-\alpha}$ je 100 $(1-\alpha)$ % kvantil normálního rozdělení s parametry 0 a 1, je důsledný pro test hypotézy $\mu = t$ vůči pravostranným alternativám; je-li totiž $t_1 > t_2$, je $\sum_{i=1}^{n} x_i - t_2 > \sum_{i=1}^{n} x_i - t_1$ a tudíž je-li $\mathrm{x} \in w(t_1)$, tj. $\sum_{i=1}^{n} x_i - t_1 > u_{1-\alpha}\sqrt{n}$, je tím spíše $\sum_{i=1}^{n} x_i - t_2 > u_{1-\alpha}\sqrt{n}$, tj. $\mathrm{x} \in w(t_2)$. Systém oborů $w(t) = \left\{\mathrm{x} : \left|\sum_{i=1}^{n} x_i - t\right| > u_{1-\frac{\alpha}{2}}\sqrt{n}\right\}$ je důsledný pro test $\mu = t$ vůči oboustranným alternativám; $\mathrm{x} \in w(t_0)$ znamená totiž, že buď $\sum_{i=1}^{n} x_i - t_0 > u_{1-\frac{\alpha}{2}}\sqrt{n}$ nebo $\sum_{i=1}^{n} x_i - t_0 < -u_{1-\frac{\alpha}{2}}\sqrt{n}$. Vezměme čísla t_1, t_2 s $t_1 < t_0 < t_2$. Je-li $\sum_{i=1}^{n} x_i - t_0 > u_{1-\frac{\alpha}{2}}\sqrt{n}$, je možná i $\sum_{i=1}^{n} x_i - t_2 > u_{1-\frac{\alpha}{2}}\sqrt{n}$, určitě však $\sum_{i=1}^{n} x_i - t_1 > u_{1-\frac{\alpha}{2}}\sqrt{n}$, a tedy $\mathrm{x} \in w(t_1) \subset w(t_1) \cup w(t_2)$; je-li $\sum_{i=1}^{n} x_i - t_0 < -u_{1-\frac{\alpha}{2}}\sqrt{n}$, je možná i $\sum_{i=1}^{n} x_i - t_1 > u_{1-\frac{\alpha}{2}}\sqrt{n}$, v kterémžto případě je $\mathrm{x} \in w(t_1)$, určitě však $\sum_{i=1}^{n} x_i - t_2 < -u_{1-\frac{\alpha}{2}}\sqrt{n}$, takže $\mathrm{x} \in w(t_2)$. Tedy každý bod $w(t_0)$ patří aspoň do jedné z množin $w(t_1)$ nebo $w(t_2)$.

Nyní definujeme na výběrovém prostoru W funkce $\underline{\Theta}(\mathrm{x})$ a $\overline{\Theta}(\mathrm{x})$ vztahy $\underline{\Theta}(\mathrm{x}) = \inf\{t : \mathrm{x} \in W - w(t)\}$ a $\overline{\Theta}(\mathrm{x}) = \sup\{t : \mathrm{x} \in W - w(t)\}$ (to znamená, že $\underline{\Theta}(\mathrm{x})$ je infimum hodnot $\Theta(\omega)$, přijatelných přihlédnutím k tomu, že pokus dal výsledek x, $\overline{\Theta}(\mathrm{x})$ jejich supremum).

Pro konstrukci intervalů spolehlivosti platí věta: Je-li $\{w(t) : t \in T\}$ systém kritických oborů velikosti α pro test hypotézy $\Theta(\omega) = t$, který má některou z uvedených vlastností důslednosti),* pak $\underline{\Theta}(\mathrm{X})$, $\overline{\Theta}(\mathrm{X})$ je interval spolehlivosti pro $\Theta(\omega)$ s koeficientem spolehlivosti $1 - \alpha$, tj.

$$\inf_{w \in \Theta^{-1}(\Theta_0)} \mathrm{P}\{\underline{\Theta}(\mathrm{X}) \leqq \Theta_0 \leqq \overline{\Theta}(\mathrm{X})\} = 1 - \alpha,$$

kde Θ_0 je skutečná hodnota $\Theta(\omega)$.

Je-li systém $\{w(t)\}$ důsledný vůči pravostranným alternativám, má interval tvar $(\underline{\Theta}(\mathrm{X}), \sup \mathrm{T})$, tj. $\underline{\Theta}(\mathrm{X})$ je tzv. dolní $100(1 - \alpha)\%$ hranicí spolehlivosti pro $\Theta(\omega)$, je-li systém $\{w(t)\}$ důsledný vůči levostranným alternativám, má interval spolehlivosti tvar $(\inf T, \overline{\Theta}(\mathrm{X}))$ tj. $\overline{\Theta}(\mathrm{X})$ je tzv. horní $100(1 - \alpha)$ % hranicí spolehlivosti pro $\Theta(\omega)$; je-li obor $\{w(t)\}$ důsledný vůči oboustranným alternativám, jde o dvojstranný interval spolehlivosti.

Horní hranice spolehlivosti se zřejmě užije tenkrát, když bude podle povahy úlohy záležet jen na odhadu hodnoty $\Theta(\omega)$ shora a za chybu bude považováno jen podcenění této hodnoty, proti němuž se budeme chtít zabezpečit tak, aby jeho pravděpodobnost nepřekročila α. Dolní hranice spolehlivosti se užije tenkrát, když bude záležet na odhadu hodnoty $\Theta(\omega)$ zdola, a chybou bude jen její přecenění. Konečně, obyčejného intervalu oboustranného se užije tenkrát, když jak podcenění tak přecenění skutečné hodnoty bude pokládáno za vadu.

Důkaz věty: Je-li $\mathrm{x} \in W - w(t)$, pak $\underline{\Theta}(\mathrm{x}) \leqq t$ a $\overline{\Theta}(\mathrm{x}) \geqq t$ podle definice suprema a infima, takže

$$W - w(t) \subset \{\mathrm{x} : \underline{\Theta}(\mathrm{x}) \leqq t \leqq \overline{\Theta}(\mathrm{x})\}. \tag{2}$$

Je-li systém $\{w(t)\}$ důsledný vůči pravostranným alternativám, je navíc

a) $\sup \{t : \mathrm{x} \in W - w(t)\} = \sup \mathrm{T}$ pro každé x,

b) $\inf \{t : \mathrm{x} \in W - w(t)\} < t' \Rightarrow \mathrm{x} \in W - w(t')$, neboť z $\inf \{t : \mathrm{x} \in W - w(t)\} < t'$ plyne, že existuje $t'' < t'$, pro které $\mathrm{x} \in W - w(t'')$ a z důslednosti plyne hned tvrzení b).

Protože množina $\{\mathrm{x} : \underline{\Theta}(\mathrm{x}) = \Theta_0\}$ má podle předpokladu pravděpodobnost 0 je tedy

$$P\{\underline{\Theta}(\mathrm{X}) \leqq \Theta_0 \leqq \overline{\Theta}(\mathrm{X})\} = P\{\underline{\Theta}(\mathrm{X}) \leqq \Theta_0\} = P\{\mathrm{X} \in W - w(\Theta_0)\} \geqq 1 - \alpha$$

neboť kritické obory systému $\{w(t), t \in T\}$ mají velikost α.* Je-li systém $\{w(t) : t \in T\}$ důsledný vůči levostranným alternativám, ukáže se podobně, že

$$P\{\overline{\Theta}(\mathrm{X}) \geqq \Theta_0 \geqq \underline{\Theta}(\mathrm{X})\} = P\{\overline{\Theta}(\mathrm{X}) \geqq \Theta_0\} = P\{\mathrm{X} \in W - w(\Theta_0)\} \geqq 1 - \alpha .$$

---

*) a množina bodů výběrového prostoru $\{\mathrm{X} : \underline{\Theta}(\mathrm{X}) = \Theta_0\} \cup \{\mathrm{X} : \overline{\Theta}(\mathrm{X}) = \Theta_0\}$ má pravděpodobnost 0 (což prakticky znamená, že $\underline{\Theta}(\mathrm{X})$ a $\overline{\Theta}(\mathrm{X})$ mají rozdělení spojitého typu).

Konečně budiž $\{w(t),\ t \in T\}$ důsledný vůči oboustranným alternativám; pak, je-li $\underline{\Theta}(\mathrm{x}) < t < \overline{\Theta}(\mathrm{x})$, existuje nějaké $t' < t$, pro které $\mathrm{x} \in W - w(t')$ a nějaké $t'' > t$, pro které $\mathrm{x} \in W - w(t'')$. Z důslednosti systému plyne, že $w(t) \subset w(t') \cup \cup\, w(t'')$, čili $[W - w(t')] \cap [W - w(t'')] \subset W - w(t)$. Platí tedy implikace $\underline{\Theta}(\mathrm{x}) < t < \overline{\Theta}(\mathrm{x}) \Rightarrow \mathrm{x} \in W - w(t)$.

Vzhledem k předpokladu, že pravděpodobnost množiny

$$\{\mathrm{x} : \underline{\Theta}(\mathrm{x}) = \Theta_0\} \cup \{\mathrm{x} : \overline{\Theta}(\mathrm{x}) = \Theta_0\}$$

je rovna 0, je tedy

$$P\{\underline{\Theta}(\mathrm{X}) \leqq \Theta_0 \leqq \overline{\Theta}(\mathrm{X})\} = P\{\mathrm{X} \in W - w(\Theta_0)\} \geqq 1 - \alpha\,.$$

Přesná hodnota pravděpodobnosti, že interval spolehlivosti pokryje skutečnou hodnotu $\Theta(\omega)$ závisí na tom, které ω z $\Theta^{-1}\,(\Theta_0)$ je skutečným parametrem rozdělení. V případě, že $\Theta(\omega)$ je jedno- jednoznačná funkce, nebo v případě, že pro test hypotézy $\Theta(\omega) = t$ je zvolen kritický obor podobný výběrovému prostoru, (existuje-li takový), je pravděpodobnost pokrytí skutečné hodnoty právě rovna $1 - \alpha$.

Uvedeme dva příklady konstrukce intervalu spolehlivosti podle uvedené věty:

Příklad 1: Budiž X náhodná veličina s normálním rozdělením se střední hodnotou μ_X a s rozptylem σ_X^2, Y náhodná veličina se střední hodnotou μ_Y a s rozptylem σ_Y^2. Střední hodnoty μ_X a μ_Y jsou neznámy, rozptyly jsou známá čísla. Je žádána dolní $100\,(1 - \alpha)$ % hranice spolehlivosti pro absolutní hodnotu rozdílu středních hodnot, $d = |\mu_X - \mu_Y|$. Za účelem odhadu se provádí n_X navzájem nezávislých pokusů s veličinou X a n_Y navzájem nezávislých pokusů s veličinou Y.

Výběrový prostor W tedy v tomto případě je množina všech skupin $n_X + n_Y$ reálných čísel $\mathrm{X} = (X_1, \ldots, x_{n_Y}, y_1, \ldots, y_{n_Y})$, parametrickým prostorem Ω množina všech dvojic reálných čísel, $\omega = (\mu_X, \mu_Y)$, funkcí $\Theta(\omega)$ je $d(\mu_X, \mu_Y) = |\mu_X - \mu_Y|$. Jako kritický obor pro test hypotézy d $(\mu_X, \mu_Y) = t$ se nabízí množina $w\,(t) = \left\{\mathrm{x} : \frac{|\bar{x} - \bar{y}|}{\sigma} > l\,(t, \alpha)\right\}$, kde $\bar{x}$ je aritmetický průměr čísel

$$x_1, \ldots, x_{n_X},\ \bar{x} = \frac{1}{n}\sum_{i=1}^{n_x} x_i\,,$$

$$\bar{y} = \frac{1}{n_Y}\sum_{j=1}^{n_y} y_j\,, \quad \sigma = \left[\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}\right]^{1/2},$$

a $l\,(t, \alpha)$ je číslo, závislé na t a na α, stanovené tak, aby

$$P\left\{\frac{|\bar{X} - \bar{Y}|}{\sigma} > l(t, \alpha)\ \middle|\ d(\mu_X, \mu_Y) = \mathrm{t}\right\} = \alpha\,.$$

Číslo $l(t, \alpha)$, splňující tuto podmínku, je kořenem rovnice

$$\Phi\left(\frac{t}{\sigma} + l(t, \alpha)\right) - \Phi\left(\frac{t}{\sigma} - l(t, \alpha)\right) = 1 - \alpha ,$$

kde Φ je distribuční funkce normálního rozdělení s parametry 0 a 1. K ověření důslednosti systému kritických oborů $w(t) = \left\{\mathbf{x} : \frac{|\overline{x} - \overline{y}|}{\sigma} > l(t, \alpha)\right\}$ vůči pravostranným alternativám stačí si všimnout, že $l(t, \alpha)$ je rostoucí funkcí t v intervalu $T = \langle 0, \infty)$, a jiné hodnoty t nemají smyslu. Dolní 100 $(1 - \alpha)$ % hranicí spolehlivosti pro $d(\mu_X, \mu_Y)$ tedy je $\underline{d}(\mathbf{x}) = \inf\left\{t : \frac{|\overline{X} - \overline{Y}|}{\sigma} > l(t, \alpha)\right\}$; protože $l(t, \alpha)$ je rostoucí funkce t, je $\underline{d}(\mathbf{X})$ rovno kořeni rovnice $l(\underline{d}(\mathbf{X}), \alpha) = \frac{|\overline{X} - \overline{Y}|}{\sigma}$, to znamená, že lze $\underline{d}(\mathbf{X})$ určit z rovnice

$$\Phi\left(\frac{\underline{d}(\mathbf{X})}{\sigma} + \frac{|\overline{X} - \overline{Y}|}{\sigma}\right) - \Phi\left(\frac{\underline{d}(\mathbf{X})}{\sigma} - \frac{|\overline{X} - \overline{Y}|}{\sigma}\right) = 1 - \alpha .$$

K tomu by ovšem byly účelné tabulky funkce $\Phi(x + y) - \Phi(x - y)$ ([4]), jinak by bylo nutno řešit rovnici vhodnou numerickou metodou.

Příklad 2: Budiž $\mathbf{X} = (X_1, \ldots, X_n)$ náhodný výběr ze souboru s normálním rozdělením s neznámým rozptylem σ^2 a s neznámou střední hodnotou μ. Sestrojíme 100 $(1 - \alpha)$ % horní hranici spolehlivosti pro funkci těchto parametrů $\Theta(\mu, \sigma^2) = \Phi\left(\frac{H - \mu}{\sigma}\right)$, kde H je dané číslo a Φ značí distribuční funkci normálního rozdělení s parametry (0, 1). Odhadujeme tedy horní hranici spolehlivosti pro podíl základního souboru s hodnotou znaku menší než H nebo nejvýše rovnou H. Vycházíme z testu hypotézy $\Theta(\mu, \sigma^2) = t$ proti alternativě $\Theta(\mu, \sigma^2) < t$. Testovaná hypotéza je rovnocenná s hypotézou $\frac{H - \mu}{\sigma} = u_t$, alternativní hypotéza odpovídá hypotéze $\frac{H - \mu}{\sigma} < u_t$, kde u_t jest $100t$ % kvantil normálního rozdělení s nulovou střední hodnotou a jednotkovým rozptylem. Kritický obor pro testování $\frac{H - \mu}{\sigma} = u_t$ proti $\frac{H - \mu}{\sigma} < u_t$ je

$$w(t) = \left\{\mathbf{x} : \frac{H - \overline{x}}{s} < k(t)\right\}$$

kde $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, $s^2 = \frac{1}{n - 1}\sum_{i=1}^{n}(x_i - \overline{x})^2$ a $k(t)$ je číslo, stanovené tak, aby

$$P\left\{\frac{H - \overline{X}}{s} < k(t) \,\middle|\, \frac{H - \mu}{\sigma} = u_t\right\} = \alpha .$$

Tato poslední rovnice vlastně je

$$P\{t(n-1, u_t\sqrt{n}) < \sqrt{n}k(t)\},$$

kde $t(n-1, u_t\sqrt{n})$ je náhodná veličina s necentrálním rozdělením t s $n-1$ stupni volnosti a s parametrem necentrality $u_t\sqrt{n}$ (viz [6], [7]). Tudíž

$$k(t) = \frac{1}{\sqrt{n}}\, t(n-1, u_t\sqrt{n}, 1-\alpha)$$

$$= -\frac{1}{\sqrt{n}}\, t(n-1, -u_t\sqrt{n}, \alpha),$$

kde $t(n-1, -u_t\sqrt{n}, \alpha)$ je 100 α % kritická hodnota necentrálního t — rozdělení s $n-1$ stupni volnosti a s parametrem necentrality $-u_t\sqrt{n}$. Lze ji stanovit z tabulek, jež sestavili N. L. Johnson a B. L. Welch ([5], [6]).

Horní hranice spolehlivosti pro $\Theta(\mu, \sigma^2) = \Phi\left(\frac{H-\mu}{\sigma}\right)$ pak jest

$$\bar{\Theta}(\mathbf{X}) = \sup\left\{t : \frac{H-\bar{X}}{s} \geqq k(t)\right\} =$$

$$= \sup\left\{t : \frac{H-\bar{X}}{s} \geqq -\frac{1}{\sqrt{n}}\, t(n-1, -u_t\sqrt{n}, \alpha)\right\} =$$

$$= \sup\left\{t : t(n-1, -u_t\sqrt{n}, \alpha) \geqq \frac{\bar{X}-H}{s}\sqrt{n}\right\}.$$

Kvantil u_t je rostoucí funkce t, tedy $-u_t\sqrt{n}$ je klesající funkcí t a $t(n-1, -u_t\sqrt{n}, \alpha)$ je také klesající funkcí t, takže hranice $\bar{\Theta}(\mathbf{X})$ je rovna hodnotě t, splňující rovnici

$$t(n-1, -u_t\sqrt{n}, \alpha) = \frac{\bar{X}-H}{s}\sqrt{n},$$

tj. rovnici

$$-u_t\sqrt{n} = \delta\left(n-1, \frac{\bar{X}-H}{s}\sqrt{n}, \alpha\right).$$

Zde $\delta\left(n-1, \frac{\bar{X}-H}{s}\sqrt{n}, \alpha\right)$ značí tu hodnotu parametru necentrality, při níž 100 α % kritická hodnota necentrálního t - rozdělení je právě rovna $\frac{\bar{X}-H}{s}\sqrt{n}$. Tyto hodnoty jsou tabelovány v [5], [6]. Horní hranice spolehlivosti pro $\Theta(\mu, \sigma^2)$ pak je

$$\bar{\Theta}(\mathbf{X}) = \Phi\left(-\frac{1}{\sqrt{n}}\,\delta\left(n-1, \frac{\bar{X}-H}{s}\sqrt{n}, \alpha\right)\right).$$

Pro hypotézu $\Theta(\omega) = t$ existuje ne jeden, nýbrž řada různých testů, a z nich se volí ten, který má nejvýhodnější silofunkci. Právě tak existuje řada možných konstrukcí intervalu spolehlivosti pro $\Theta(\omega)$ a je třeba mezi nimi zvolit tu, která dá interval s nejvýhodnějšími vlastnostmi. Je přirozené, zvolit ze všech intervalů spolehlivosti s týmž koeficientem spolehlivosti ten, jehož koncové body jsou nejblíže skutečné hodnotě parametru. Tedy mezi všemi dolními hranicemi spolehlivosti, odpovídajícími témuž koeficientu spolehlivosti $1 - \alpha$ budeme za nejlepší pokládat tu, která bude „stochasticky největší", to jest tu hranici $\underline{\Theta}^*(\mathbf{X})$ pro kterou bude platit

$$P\{\underline{\Theta}^*(\mathbf{X}) > t\} \geqq P\{\underline{\Theta}(\mathbf{X}) > t\} \tag{3}$$

ať je t jakékoliv číslo menší než skutečná hodnota parametru, $t < \Theta_0$, a $\underline{\Theta}(\mathbf{X})$ jakákoliv jiná dolní hranice spolehlivosti s týmž koeficientem spolehlivosti. Hranici $\underline{\Theta}^*(\mathbf{X})$, splňující (3) nazveme *nejtěsnější* 100 $(1 - \alpha)$ % *dolní hranicí spolehlivosti*. Podobně hranici $\overline{\Theta}^*(\mathbf{X})$, splňující podmínky

$$P\{\overline{\Theta}^*(\mathbf{X}) \geqq \Theta_0\} = 1 - \alpha$$

a

$$P\{\overline{\Theta}^*(\mathbf{X}) < t\} \geqq P\{\overline{\Theta}(\mathbf{X}) < t\}$$

pro všechna $t > \Theta_0$ a jakoukoliv jinou horní hranici spolehlivosti s týmž koeficientem spolehlivosti $1 - \alpha$, nazveme *nejtěsnější* 100 $(1 - \alpha)$ % *horní hranicí spolehlivosti*.

Je-li $\Theta(\omega)$ jedno- jednoznačná funkce, pak každá hypotéza $\Theta(\omega) = t$ je jednoduchá hypotéza. V takovém případě často existuje kritický obor $w(t)$ pro test hypotézy $\Theta(\omega) = t$, který je nejsilnější vůči všem alternativám $\Theta(\omega) = t'$, $t' > t$, tj. takový, že pro jakékoliv $t' > t$ a jakýkoliv jiný kritický obor $w_1(t)$, splňující

$$P\{\mathbf{X} \in w_1(t)/\Theta(\omega) = t\} = \alpha$$

platí

$$P\{\mathbf{X} \in w(t)/\Theta(\omega) = t'\} \geqq P\{\mathbf{X} \in w_1(t)/\Theta(\omega) = t'\}$$

čili

$$\int_{w(t)} \mathrm{d}F_{\Theta^{-1}(t')}(\mathbf{x}) \geqq \int_{w_1(t)} \mathrm{d}F_{\Theta^{-1}(t')}(\mathbf{x}).$$

Takový kritický obor nazveme oborem stejnoměrně nejsilnějším vůči pravostranným alternativám. Podobně je-li $\Theta(\omega)$ jedno- jednoznačná, často existuje pro jakékoliv t obor $w(t)$, stejnoměrně nejsilnější vůči levostranným alternativám, tj. takový, že pro jakýkoliv jiný obor $w_1(t)$, splňující

$$P\{\mathbf{X} \in w_1(t)/\Theta(\omega) = t\} = P\{\mathbf{X} \in w(t)/\Theta(\omega) = t\}$$

platí

$$P\{\mathbf{X} \in w(t)/\Theta(\omega) = t'\} \geqq P\{\mathbf{X} \in w_1(t)/\Theta(\omega) = t'\},$$

ať je t' jakékoliv číslo menší než t.

Mezi nejtěsnějšími hranicemi spolehlivosti a stejnoměrně nejsilnějšími kritickými obory platí vztah, vyjádřený následující větou:

Je-li $\{w^*(t)\}$ důsledný systém kritických oborů velikosti α, stejnoměrně nejsilnějších vůči pravostranným (resp. levostranným) alternativám, pak $\underline{\Theta}^*(\mathbf{x}) = \inf \{t : \mathbf{x} \in W - w^*(t)\}$ resp. $\overline{\Theta}^*(\mathbf{X}) = \sup \{t : \mathbf{x} \in W - w^*(t)\}$ je nejtěsnější dolní (resp. horní) 100 $(1 - \alpha)$ % hranice spolehlivosti pro $\Theta(\omega)$.

Správnost věty vyplyne z této úvahy: s každou konstrukcí dolní hranice spolehlivosti $\underline{\Theta}(\mathbf{X})$ je sdružen kritický obor $w_{\underline{\Theta}}(t)$ pro test $\Theta(\omega) = t$ proti pravostranným alternativám, totiž obor, definovaný vztahem

$$w_{\underline{\Theta}}(t) = \{\mathbf{x} : \underline{\Theta}(\mathbf{x}) > t\}.$$

Zvolme nějaké $t' < \Theta_0$. Protože $\{w^*(t)\}$ je důsledný systém, jev $\underline{\Theta}^*(\mathbf{X}) > t'$ nastane tehdy a jen tehdy, když $\mathbf{X} \in w^*(t')$. Protože pak každý obor ze systému $\{w^*(t)\}$ je stejnoměrně nejsilnější vůči pravostranným alternativám, je pravděpodobnost tohoto jevu při $\Theta = \Theta_0$ nejméně rovna pravděpodobnosti, že X padne do jakékoliv jiného kritického oboru, speciálně toho, který je sdružen s intervalem $\underline{\Theta}(\mathbf{X})$, tedy

$$P\{\mathbf{X} \in w^*(t')/\Theta = \Theta_0\} \geqq P\{\mathbf{X} \in w_{\underline{\Theta}}(t')/\Theta = \Theta_0\}.$$

(Systém oborů $\{w_{\underline{\Theta}}(t')\}$ je důsledný samozřejmě podle své konstrukce). Pro horní hranici spolehlivosti je postup úplně stejný. Z věty plyne už zcela jasně užití Neyman-Pearsonovy základní věty ke konstrukci hranic spolehlivosti a význam suficientních charakteristik pro konstrukci hranic spolehlivosti. Není-li $\Theta(\omega)$ jedno- jednoznačná funkce, takže $\Theta(\omega) = t$ je složená hypotéza, lze na konstrukci hranic spolehlivosti přenést princip podobnosti kritického oboru výběrovému prostoru (viz [3]), a hledat nejtěsnější hranici mezi těmi, u kterých riziko nepokrytí skutečné hodnoty $\Theta(\omega)$ je konstantní pro všechna $\omega \in \Theta^{-1}(\Theta_0)$.

Nejtěsnější hranice spolehlivosti, jak byly prve definovány, mají ještě následující vlastnost, jež dále ospravedlňuje jejich název: je-li $\underline{\Theta}^*(\mathbf{X})$ (resp. $\overline{\Theta}^*(\mathbf{X})$) nejtěsnější dolní (resp. horní) hranice spolehlivosti, pak pro jakoukoliv jinou dolní hranici spolehlivosti $\underline{\Theta}(\mathbf{X})$ (resp. jakoukoliv jinou horní hranici spolehlivosti $\overline{\Theta}(\mathbf{X})$) s týmž koeficientem spolehlivosti platí

$$\mathrm{E}\{\underline{\Theta}^*(\mathbf{X})/\underline{\Theta}^*(\mathbf{X}) \leqq \Theta_0\} \geqq E\{\underline{\Theta}(\mathbf{X})/\underline{\Theta}(\mathbf{X}) \leqq \Theta_0\}$$

resp.

$$\mathrm{E}\{\overline{\Theta}^*(\mathbf{X})/\overline{\Theta}^*(\mathbf{X}) \geqq \Theta_0\} \leqq E\{\overline{\Theta}(\mathbf{X})/\overline{\Theta}(\mathbf{X}) \geqq \Theta_0\}$$

tj. nejtěsnější dolní hranice má největší střední hodnotu, nejtěsnější horní hranice má nejnižší střední hodnotu ze všech hranic, odpovídajících témuž koeficientu spolehlivosti podmíněnou jevem, že hranice je „na správné straně skutečné hodnoty“.

Při oboustranném intervalu spolehlivosti nelze dosáhnout toho, aby dolní jeho hranice byla „stochasticky největší“ a zároveň horní hranice stochasticky nejmenší ze všech možných oboustranných intervalů spolehlivosti, tj. nelze dosáhnout toho, aby dolní i horní hranice byly nejtěsnější ve smyslu definice, zavedené při jednostranných intervalech spolehlivosti. Proto upravíme požadavek, který budeme klást na optimální interval spolehlivosti, a vyhledáme nikoliv interval nejlepší (ve smyslu tohoto požadavku) ze všech možných intervalů spolehlivosti (mezi nimiž vůbec takový nemusí existovat), nýbrž ze třídy užší, ze třídy oboustranných intervalů spolehlivosti, splňujících ještě nějaký další vhodný požadavek. (To je stejný postup, jakého se užívá při ověřování statistických hypotéz; neexistuje-li kritický obor, nejlepší ze všech kritických oborů dané velikosti, hledá se nejlepší mezi obory nějaké užší třídy.)

Takovým vhodným požadavkem je nestrannost. Interval spolehlivosti ($\underline{\Theta}(\mathbf{X})$, $\overline{\Theta}(\mathbf{X})$) se nazývá nestranným, jestliže

$$P\{\underline{\Theta}(\mathbf{X}) \leqq \Theta_0 \leqq \overline{\Theta}(\mathbf{X})\} \geqq P\{\underline{\Theta}(\mathbf{X}) \leqq t \leqq \overline{\Theta}(\mathbf{X})\}$$

pro jakékoliv $t \neq \Theta_0$ tj., jestliže skutečná hodnota parametru je pokryta častěji než kterákoliv jiná možná hodnota nebo alespoň stejně často.

Za nejlepší z nestranných intervalů spolehlivosti budeme pokládat ten interval ($\underline{\Theta}^*(\mathbf{X})$, $\overline{\Theta}^*(\mathbf{X})$), pro který

$$P\{\underline{\Theta}^*(\mathbf{X}) \leqq t \leqq \overline{\Theta}^*(\mathbf{X}\} \leqq P\{\underline{\Theta}\mathbf{X}) \leqq t \leqq \overline{\Theta}(\mathbf{X})\}$$

pro jakékoliv t a pro jakýkoliv jiný nestranný interval spolehlivosti ($\underline{\Theta}(\mathbf{X})$, $\overline{\Theta}(\mathbf{X})$) s týmž koeficientem spolehlivosti. Tedy za optimální nestranný interval spolehlivosti se pokládá ten, který pokryje každou z možných hodnot parametru s pravděpodobností menší než kterýkoliv jiný nestranný interval spolehlivosti, nebo v nejhorším případě stejně často. Takový interval nazveme „*nejužší*“ (protože slovo „nejkratší“ je vyhrazeno pro intervaly s minimální střední hodnotou rozdílu mezi horní a dolní hranicí). Také se jim říká „nejselektivnější“. Ukážeme, že nejužší intervaly spolehlivosti jsou odvozeny ze stejnoměrně nejsilnějších nestranných testů statistických hypotéz.

Předpokládejme opět pro zjednodušení, že funkce je jedno-jednoznačná, takže hypotéza $\Theta(\omega) = t$ je při každém $t \in T$ jednoduchá. Kritický obor $w(t)$ pro test hypotézy $\Theta(\omega) = t$ se nazývá nestranný, jestliže splňuje podmínku

$$P\{\mathbf{X} \in w(t)/\Theta(\omega) = t\} \leqq P\{\mathbf{X} \in w(t)/\Theta(\omega) = t'\}$$

pro všechna $t' \neq t$. ([7], [8]).
To znamená, že při užití nestranného kritického oboru má silofunkce testu minimum v testované hypotéze, tj. testovaná hypotéza je zamítána nejméně často, jestliže je správná, což je zajisté rozumný požadavek. Jestliže mezi nestrannými kritickými obory pro test $\Theta(\omega) = t$ existuje takový, který je nejsilnější vůči všem alternativám zároveň, nazýváme jej stejnoměrně nejsilnějším nestranným kritickým oborem. J. Neyman a E. S. Pearson uvedli podmínky, za kterých stejnoměrně nejsilnější nestranný kritický obor existuje a našli metodu jeho konstrukce [7]. Mezi stejnoměrně nejsilnějšími nestrannými kritickými obory a nejužšími nestrannými intervaly spolehlivosti platí vztah, vyjádřený následující větou:

Je-li $\{w^*(t) : t \in T\}$ systém stejnoměrně nejsilnějších nestranných kritických oborů pro test hypotézy $\Theta(\omega) = t$, důsledný vůči oboustranným alternativám, pak interval ($\underline{\Theta}^*(\mathbf{X})$, $\overline{\Theta}^*(\mathbf{X})$

$$\underline{\Theta}^*(\mathbf{X}) = \inf \{t : \mathbf{X} \in W - w^*(t)\}$$

$$\overline{\Theta}^*(\mathbf{X}) = \sup \{t : \mathbf{X} \in W - w^*(t)\}$$

je nejužším nestranným intervalem spolehlivosti pro $\Theta(\omega)$.

Označme Θ_0 skutečnou hodnotu $\Theta(\omega)$, vezměme nějakou hodnotu $t \neq \Theta_0$, a ukažme, že je

$$P\{\underline{\Theta}^*(\mathbf{X}) \leqq t \leqq \overline{\Theta}^*(\mathbf{X})\} \leqq P\{\underline{\Theta}^*(\mathbf{X}) \leqq \Theta_0 \leqq \overline{\Theta}^*(\mathbf{X})\}.$$

To platí, protože vzhledem k důslednosti systému $\{w^*(t) : t \in T\}$ je $\underline{\Theta}^*(\mathbf{X}) \leqq$ $\leqq t \leqq \overline{\Theta}^*(\mathbf{X})$ tehdy a jen tehdy, když $\mathbf{X} \in W - w^*(t)$, a

$$P\{\mathbf{X} \in W - w^*(t)/\Theta(\omega) = \Theta_0\} = 1 - P\{\mathbf{X} \in w^*(t)/\Theta(\omega) = \Theta_0\} \leqq$$
$$\leqq 1 - P\{\mathbf{X} \in w^*(\Theta_0)/\Theta(\omega) = \Theta_0\}$$

vzhledem k nestrannosti kritického oboru $w^*(t)$. Tím je dokázáno, že interval, odvozený z nestranného kritického oboru, je nestranný.

Že je nejužší, plyne z této úvahy: s každou metodou konstrukce intervalu spolehlivosti $(\underline{\Theta}(\mathbf{X}), \overline{\Theta}(\mathbf{X}))$ je sdružen kritický obor pro test hypotézy $\Theta(\omega) = t$, totiž obor

$$w(t) = \{\mathbf{x} : \underline{\Theta}(\mathbf{x}) > t\} \cup \{\mathbf{X}\overline{\Theta}(\mathbf{x}) < t\}.$$

Je-li interval $(\underline{\Theta}(\mathbf{X}), \overline{\Theta}(\mathbf{X}))$ nestranný, je s ním sdružený obor také nestranný, protože

$$P\{\mathbf{X} \in w(\Theta_0)/\Theta(\omega) = \Theta_0\} = 1 - P\{\mathbf{X} \in W - w(\Theta_0)/\Theta(\omega) = \Theta_0\} =$$
$$= 1 - P\{\underline{\Theta}(\mathbf{X}) \leqq \Theta_0 \leqq \overline{\Theta}(\mathbf{X})/\Theta(\omega) = \Theta_0\} \leqq$$
$$\leqq 1 - P\{\underline{\Theta}(\mathbf{X}) \leqq t \leqq \overline{\Theta}(\mathbf{X})/\Theta(\omega) = \Theta_0\}$$

pro jakékoliv $t \neq \Theta_0$ vzhledem k nestrannosti intervalu spolehlivosti. A nemůže existovat nestranný interval spolehlivosti $(\underline{\Theta}(\mathbf{X}), \overline{\Theta}(\mathbf{X}))$ „užší" než $\underline{\Theta}^*(\mathbf{X}), \overline{\Theta}^*(\mathbf{X}))$, protože s ním sdružený kritický obor by byl silnější než $w^*(t)$, což není možné, protože $w^*(t)$ je stejnoměrně nejsilnější nestranný kritický obor. Tvrzení, že $(\underline{\Theta}(\mathbf{X}), \overline{\Theta}(\mathbf{X}))$ je „užší" než $(\underline{\Theta}^*(\mathbf{X}), \overline{\Theta}^*(\mathbf{X}))$, znamená totiž, že existují hodnoty t a t′ pro které platí:

$$P\{\mathbf{X} \in w^*(t)/\Theta(\omega) = t'\} < P\{\underline{\Theta}(\mathbf{X}) > t/\Theta(\omega) = t'\} +$$
$$+ P\{\overline{\Theta}(\mathbf{X}) < t/\Theta(\omega) = t'\} = P\{\mathbf{X} \in w(t)/\Theta(\omega) = t'\};$$

pak by ovšem obor $w(t)$, sdružený s intervalem $(\underline{\Theta}(\mathbf{X}), \overline{\Theta}(\mathbf{X}))$ byl proti alternativě $\Theta(\omega) = t'$ silnější než $w^*(t)$.

Formule pro hranice nejužších nestranných intervalů spolehlivosti lze odvodit z Neyman Pearsonovy metody konstrukce stejnoměrně nejsilnějších nestranných testů.

Podobně je možno aplikovat na problém intervalového odhadu v případě, že $\Theta(\omega)$ není jedno-jednoznačné, tj. v přítomnosti „nežádoucích" parametrů metody konstrukce tzv. kritických oborů typu B [9] a Scheffého zobecnění této metody.

Poznamenejme ještě, že stejného přístupu, jakého jsme užili ke studiu jednostranných intervalů spolehlivosti, lze užít ke studiu tzv. mediánových odhadů. Mediánový odhad není totiž ničím jiným, než 50 % dolní hranicí spolehlivosti (resp. 50 % horní hranicí spolehlivosti).

Tím uzavíráme tuto poznámku, jejímž účelem bylo upozornit na možnost výkladu teorie intervalového odhadu s využitím výsledků teorie ověřování statistických hypotéz. Tento výklad se zdá výhodným. Zároveň je možno užít této souvislosti k diskusi Fisherovy metody fiduciálních pravděpodobností.

ЗАМЕТКА К ТЕОРИИ ДОВЕРИТЕЛЬНЫХ ИНТЕРВАЛОВ

Резюме

В настоящей статье рассмотрен метод получения доверительных интервалов на основе критериев значимости с определенными свойствами.

Пусть $\mathbf{F} = \{F_\omega(\mathbf{x}) : \omega \in \Omega\}$ есть система функций распределения n-мерной случайной величины $\mathbf{X} = (X_1, X_2, \ldots, X_n)$, так называемой выборки, принимающей значения в некотором подмножестве W n-мерного Евклидова пространства. Элементы $\mathbf{F}$ отличаются друг от друга значением параметра ω, принимающего значения в невырожденном интервале Ω (может быть, больше чем одномерном). На Ω задана вещественная функция $\Theta(\omega)$, отображающая Ω на (открытый) интервал T.

Систему критических областей $\{w(t) : t \in T\}$ для проверки гипотезы $\Theta(\omega) = T$, $t \in T$, назовем состоятельной против правых альтернативных гипотез, если для каждой пары $t_1\, t_2$ значений из T неравенство $t_1 > t_2$ влечет за собой отношение $w(t_1) \subset w(t_2)$, состоятельной против левых альтернативных гипотез, если $t_1 > t_2$ влечет за собой $w(t_2) \subset w(t_1)$, и состоятельной против двусторонних алтернативных гипотез, если для каждой тройки элементов t_0, t_1, t_2 интервала T выполняющих неравенство $t_1 > t_0 > t_2$ имеет место соотношение $w(t_0) \subset w(t_1) \cup \cup w(t_2)$, причем пересечения $w(t_1) \cap w(t_0)$ и $w(t_2) \cap w(t_0)$ не пусты. Определим на W функции $\underline{\Theta}(\mathbf{x})$ и $\overline{\Theta}(\mathbf{x})$ соотношениями

$$\underline{\Theta}(\mathbf{x}) = \inf \{t : \mathbf{x} \in W - w(t)\}, \quad \overline{\Theta}(\mathbf{x}) = \sup \{t : \mathbf{x} \in W - w(t)\}.$$

Справедлива теорема: Если система критических областей $\{w(t) : t \in T\}$ имеющих один и тот же уровень значимости α обладает некоторым из перечисленных више свойств состоятельности, и если множества $\{\mathbf{x} : \underline{\Theta}(\mathbf{x}) = \Theta_0\}$ и $\{\mathbf{x} : \overline{\Theta}(\mathbf{x}) = \Theta_0\}$ имеют вероятсность нуль (Θ_0 — истинное значение $\Theta(\omega)$), то пара случайных величин $(\underline{\Theta}(\mathbf{X}), \overline{\Theta}(\mathbf{X}))$ будет доверительным интервалом для $\Theta(\omega)$ с коеффицентом доверия равным $1 - \alpha$. Елси система критических областей, лежащая в основе конструкции доверительного интервала, состоятельна против правых альтернатив, то доверительный интервал будет иметь форму $\langle \underline{\Theta}(\mathbf{X}), \sup T)$, если система критических областей состоятельна против левых альтернатив, то получится интервал формы $(\inf T, \overline{\Theta}(\mathbf{X})\rangle$ и если $\{w(t) : t \in T\}$ состоятельна против двусторонних альтернатив, то из теоремы полулится обыкновенный двусторонний интервал.

В качестве примера теорема применена к решению задачи конструкции доверительных границ для простого значения разности средних двух нормальных совокупностей и доверительных границ для значения функции распределения нормальной совокупности с неизвестными параметрами в данной точке.

На заключение изучены свойства интервалов полученных из систем критических областей, алвяющихся оптимальными с точки зрения мощности.

A NOTE ON THE THEORY OF CONFIDENCE INTERVALS

Summary

This note is concerned with the approach to estimation by confidence intervals via the tests of statistical hypotheses with some specified properties. Conditions are studied for the resulting confidence set to be an interval.

Let $\mathbf{F} = \{F\omega(\mathbf{x}) : \omega \in \Omega\}$ be a system of distribution functions over a sample space W (assumed to be n-dimensional euclidean). The elements of $\mathbf{F}$ are labeled by an index ω, assuming values in a set Ω called parameter space. Let $\Theta(\omega)$ be a real function taking Ω onto an (open) interval T.

We shall call a system $\{w(t) : t \in T\}$ of critical regions for testing the hypothesis

$\Theta(\omega) = t$, $t \in T$, consistent against right-sided alternatives, if for any pair of values t_1, t_2 from T the inequality $t_1 > t_2$ implies $w(t_1) \subset w(t_2)$, consistent against left-sided alternatives, if $t_1 > t_2$ implies $w(t_2) \subset w(t_1)$ and consistent against two-sided alternatives if for any triple t_0, t_1, t_2 of elements of T the inequality $t_1 > t_0 > t_2$ implies $w(t_0) \subset w(t_1) \cup w(t_2)$ and the intersections $w(t_1) \cap w(t_0)$ and $w(t_2) \cap w(t_0)$ are both non-empty. (It must be stressed that this definition of consistency applies to systems of critical regions and should not be confused with the generally adopted concept of consistency of a test of statistical hypothesis.) Denote by $\underline{\Theta}(\mathbf{x})$ and $\overline{\Theta}(\mathbf{x})$ respectively the greatest lower bound and the least upper bound respectively of the set $\{t : \mathbf{x} \in W - w(t)\}$, i. e.

$$\underline{\Theta}(\mathbf{x}) = \inf \{t : \mathbf{x} \in W - w(t)\}, \quad \overline{\Theta}(\mathbf{x}) = \sup \{t : \mathbf{x} \in W - w(t)\}.$$

Then it holds: If the system of size-α critical regions $\{w(t) : t \in T\}$ possesses some of the properties of consistency listed above and if the sets $\{\mathbf{x} : \underline{\Theta}(\mathbf{x}) = \Theta_0\}$ and $\{\mathbf{x} : \overline{\Theta}(\mathbf{x}) = \Theta_0\}$ have probablity zero (where Θ_0 is the true value of $\Theta(\omega)$), then the pair of random variables $(\underline{\Theta}(\mathbf{X}), \overline{\Theta}(\mathbf{X}))$ constitutes a confidence interval for $\Theta(\omega)$ with the confidence coefficient $1 - \alpha$. If the system of critical regions employed in the construction is consistent against right-sided alternatives, the theorem yields a confidence interval of the form $<\underline{\Theta}(\mathbf{X}), \sup T)$, i. e. a lower confidence limit; if it is consistent against left-sided alternatives, we obtain an upper confidence limit, and if it is consistent against two-sided alternatives, a common two-sided confidence interval results.

The theorem is applied to the construction of confidence limits for the absolute value of the difference between means of two normal populations and confidence limits for the proportion of elements in a normal population with unknown mean and variance that have a value less than a specified constant.

The relation of various optimum properties of tests to optimum properties of confidence intervals are studied next.

REFERENCES

[1] S. S. Wilks: Mathematical Statistics, Princeton, New Jersey, 1950.
[2] L. Schmetterer: Einführung in die mathematische Statistik, Wien 1956.
[3] D. A. S. Fraser: Non-parametric methods of statistics, New York 1957.
[4] W. J. Dixon: Table of normal probabilities for intervals of various lengths and locations, Annals of Mathematical Statistics 19 (1948), str. 424—426.
[5] N. L. Johnson, B. L. Welch: Some applications of the noncentral t-distribution, Biometrika 31 (1939), str. 362—389.
[6] J. Janko: Statistické tabulky, Praha 1958.
[7] J. Neyman, E. S. Pearson: Contributions to the theory of testing statistical hypotheses, Statistical Research Memoirs, 1 (1936), str. 1—32.
[8] J. Janko: Teorie ověřování statistických hypotéz, Pokroky matematiky, fysiky a astronomie, 3 (1958) str. 9—21, str. 125—139.
[9] Jerzy Neyman: Bulletin de la Société Mathématique de France, 63 (1935), str. 246.