# PITMANOVO POROVNÁNÍ ODHADŮ

Michal Friesl
Jiří Reif

*Katedra matematiky*
*Fakulta aplikovaných věd*
*Západočeská univerzita v Plzni*

•Definice •Příklady •Bayesovsky •Regrese

# Pitman (1937)

- Pozorování $X$, jehož rozdělení závisí na parametru $\theta$
- Pitman (1937) navrhuje: $\delta_1 = \delta_1(X)$ je lepší/bližší (closer) odhad $\theta$ než $\delta_2 = \delta_2(X)$, když

$$\mathrm{P}_\theta[|\delta_1 - \theta| < |\delta_2 - \theta|] > 1/2 \qquad \text{pro všechna } \theta$$

  Nejlepší odhad je takový, který je lepší než kterýkoli jiný (pokud existuje).
- Neměří se kvalita jednotlivých odhadů, ale porovnávají se dvojice (nemusí se podařit). Zjišťuje se pouze, kdy $\delta_1$ je lepší (horší) než $\delta_2$, ne už ale jak moc je lepší (horší).
- Není třeba předpokládat existenci momentů.

# Definice

- Řekneme, že $\delta_1$ je lepší (bližší $\theta$) než $\delta_2$, když

$$\mathrm{P}_\theta[|\delta_1-\theta|<|\delta_2-\theta|] \geqq \mathrm{P}_\theta[|\delta_1-\theta|>|\delta_2-\theta|] \quad \text{pro vš. } \theta$$

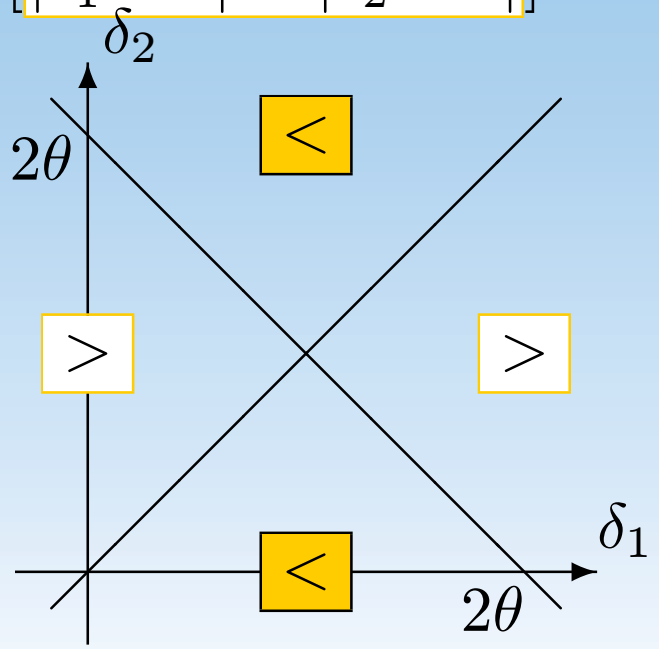
s ostrou nerovností alespoň pro jedno $\theta$.

- Používá se (Pitman's closeness/nearness measure)

$$\mathrm{PC}_\theta(\delta_1,\delta_2) = \mathrm{P}_\theta[\boxed{|\delta_1-\theta|<|\delta_2-\theta|}] - \mathrm{P}_\theta[\boxed{|\delta_1-\theta|>|\delta_2-\theta|}]$$

a porovnává s $0$.

- Místo $L(\delta,\theta)=|\delta-\theta|$ lze použít i jinou ztrátovou funkci, kritérium vyjde stejně pro všechny ztrátové funkce, které jsou rostoucí funkcí vzdálenosti mezi $\delta$ a $\theta$.



# Medián a postačující statistika

- Uvažujme odhady, které jsou funkcí postačující statistiky $T$. Je-li $\psi(\theta) = \operatorname{med}_\theta T$ monotónní, pak

$$\operatorname{med}_\theta \psi^{-1}(T) = \theta$$

- Pitman (1937): $T_1, T_2$ spojitě rozdělené veličiny, $T_1$ s mediánem $\theta$. Existuje-li veličina $Z$ (buď kladná, nebo záporná), že $T_1$ a $Z(T_2 - T_1)$ jsou nezávislé, pak $T_1$ je lepším odhadem $\theta$ než $T_2$.
- Pro problémy měřítka a polohy se hodí $Z = 1/T_1$ (předpoklad znamená $T_1$ a $T_2/T_1$ nezávislé) a $Z = 1$ ($T_1$ a $T_2 - T_1$ nezávislé).
- Např. odhad $\sigma^2$ z $X_1, \ldots, X_n \sim \mathrm{N}(\mu, \sigma^2)$. Postačující statistika $T_1 = (n-1)s^2 / \operatorname{med} \chi^2_{n-1}$ má medián $\sigma^2$. Každý jiný odhad $T_2$ invariantní vzhledem k měřítku a poloze (tedy $T_2/T_1$ nezávislé s $T_1$) je horší. Při známém např. $\mu = 0$ je takovým $\sum X_i^2 / \operatorname{med} \chi^2_n$.

# Další příklady

- Odhad střední hodnoty z $X_1, \ldots, X_n \sim \mathrm{N}(\mu, 1)$. Postačující statistika $T_1 = \overline{X}$ má medián $\mu$ a každý jiný odhad $T_2$ invariantní vůči posunutí (tudíž $T_2 - T_1$ nezávislé s $T_1$) je horší. Podobně při neznámém rozptylu.
- Mezi invariantními (vzhledem k měřítku) odhady $\theta$ v $\mathrm{G}(1/\theta, m)$ se známým $m$ je $2\sum X_i / \operatorname{med} \chi^2_{2nm}$ nejlepším.
- V $\mathrm{Exp}(a, 1/\theta)$ vychází $\widehat{\theta} = 2\sum(X_i - X_{(1)}) / \operatorname{med} \chi^2_{2n-2}$ a

$$\widehat{a} = X_{(1)} - (2^{1/(n-1)} - 1) \sum (X_i - X_{(1)})/n$$

- V $\mathrm{Ro}(a - \theta/2, a + \theta/2)$

$$\widehat{a} = (X_{(1)} + X_{(n)})/2, \qquad \widehat{\theta} = (X_{(n)} - X_{(1)})/m,$$

kde $nm^{n-1} - (n-1)m^n = 1/2$.

# Nejlepší odhady

- $T$ je mediánově nestranný odhad $\theta$, když $\mathrm{P}_\theta[T \leqq \theta] = \mathrm{P}_\theta[T \geqq \theta]$.
- Ghosh, Sen (1989): Mediánově nestranný $T_1$ je lepší odhad $\theta$ než $T_2 = T_1 + Z$, kde $T_1$ a $Z$ jsou nezávislé.
- Při hustotě $f(\frac{x-\theta}{\sigma})/\sigma$ je $T_1 > 0$, mediánově nestranný odhad $\sigma$, lepší než $T_2 = T_1(1+Z)$, kdykoli $T_1$ a $Z$ jsou nezávislé.
- Nayak (1990, 1998): Pro $\theta$ parametr polohy — $(X_1, \ldots, X_n)$ s hustotou $f(x_1 - \theta, \ldots, x_n - \theta)$, $\theta \in \mathbf{R}$, (a ztrátovou funkci s $\delta - \theta$ nejprve klesající k 0 pro $\delta = \theta$ a pak rostoucí) — je nejlepší ekvivariantním odhadem $\theta$

$$X_n - \mathrm{med}_{\theta=0}(X_n \mid X_1 - X_n, \ldots, X_{n-1} - X_n)$$

- Pro elipticky symetrické hustoty $f(x-\theta) = h((x-\theta)'B(x-\theta))$ je nejlepším ekvivariantním odhadem $\sum_i (x_i \sum_j b_{ij}) / \sum_{ij} b_{ij}$.

# Kritika

- Netranzitivita. Může být $\delta_1$ lepší než $\delta_2$, $\delta_2$ lepší než $\delta_3$, a přitom $\delta_3$ lepší než $\delta_1$. Např. při odhadu $\theta$ z $X \sim \mathrm{Ro}(-\theta, \theta)$ je $X$ stejně dobrý jako $0,9|X|$, což je lepší než $3,2|X|$, a ten je lepší než $X$.
- Pokud nám jde o nejlepší odhad a nacházíme ho, je nám jedno, jak se mezi sebou řadí poražení.
- Kritérium závisí na sdruženém rozdělení dvojice odhadů.
- Dokonce dva odhady mohou mít stejné (marginální) rozdělení a přesto jeden být lepší než druhý, např. pro $X_1, X_2$ nezávislé se stejným Cauchyovým rozdělením, je $(X_1 + X_2)/2$ lepším odhadem parametru polohy, než $X_1$.
- $\mathrm{PC}$ nezapadá do teorie rozhodování.

# Aposteriorní PC

- Ghosh, Sen (1991), Bose (1991, 1998): Předpokládáme apriorní rozdělení $\pi(\theta)$ a místo $\mathrm{PC}_\theta(\delta_1, \delta_2)$ použijeme rozdíl aposteriorních pravděpodobností

$$\mathrm{PPC}(\delta_1, \delta_2) = \mathrm{P}[|\delta_1 - \theta| < |\delta_2 - \theta| \mid X] - \mathrm{P}[|\delta_1 - \theta| > |\delta_2 - \theta| \mid X]$$

- Aposteriorní medián $\mathrm{med}(\theta \mid X)$ je nejlepší odhad, stejně rozdělené odhady jsou stejně preferovány, tranzitivita $\iff$ $\pi(\theta|x)$ má jediný medián $\iff$ nejlepší odhad je jediný.
- Pro $\theta$ vektor je vícerozměrný medián $m$ ($a'm = \mathrm{med}\, a'\theta$ pro vš. $a$) nejlepší (pokud existuje) a je jediný (nebo je rozdělení koncentrováno na přímce).
- Podmínky pro tranzitivitu ve vícerozměrném případě.

# Teorie rozhodování

- Rukhin (1996). Mnoho v teorii rozhodování závisí jen na rozdílech rizikových funkcí $R_\theta(\delta_0) - R_\theta(\delta_1)$, kde $R_\theta(\delta) = \mathrm{E}_\theta\, L(\delta, \theta)$.
- Srovnávací riziko ($C$omparative risk)

$$R_\theta(\delta_0, \delta_1) = \mathrm{E}_\theta\, L(\theta, \delta_0, \delta_1)$$

  kde $L(\theta, \delta_0, \delta_1)$ je srovnávací ztrátová funkce — čím větší, tím hůře si $\delta_0$ vede ve srovnání s $\delta_1$.
- Předpokládá se symetrie $R_\theta(\delta_0, \delta_1) = -R_\theta(\delta_1, \delta_0)$ (a konečnost), např. když $L(\theta, \delta_0, \delta_1) = -L(\theta, \delta_0, \delta_1)$ (a tedy $L(\theta, \delta, \delta) = 0$).
- Odhad $\delta_0$ je $C$-přípustný (comparative admissible), když neexistuje lepší, tj. $R_\theta(\delta_0, \delta) \geqq 0$ pro vš. $\theta$ implikuje rovnost.
- Při apriorním rozdělení $\pi(\theta)$ nazveme $\delta_0$ $C$-bayesovský, když

$$\sup_\delta \int R_\theta(\delta_0, \delta)\, \mathrm{d}\pi(\theta) \leqq 0 \qquad \text{(tedy rovno 0).}$$

# Pitmanovo kritérium

- Volíme-li $L(\theta, \delta_0, \delta_1) = \operatorname{sign}(L(\delta_0, \theta) - L(\delta_1, \theta))$, dostáváme
$$R_\theta(\delta_0, \delta_1) = -\operatorname{PC}_\theta(\delta_0, \delta_1)$$
- Srovnávací ztrátová funkce sice tranzitivní, tranzitivita kritéria ale není. Nepřípustný odhad nemusí mít přípustné zlepšení.
- Pro $\theta$ reálné a $L(\delta, \theta)$ pro každé $\delta$ jakožto funkci $\theta$ unimodální, spojitou, s minimem v $\delta$, je $C$-bayesovský odhad $\delta_0$ dán
$$\int_{-\infty}^{\delta_0} f(x \mid \theta)\, \mathrm{d}\pi(\theta) = \int_{\delta_0}^{\infty} f(x \mid \theta)\, \mathrm{d}\pi(\theta)$$
tedy $\delta_0$ je aposteriorní medián (a nezávisí na $L(\delta, \theta)$).

# Regrese

- Standardní lineární model $y = X\beta + e$, $X$ s plnou hodností $k$, $e \sim \mathrm{N}_n(0, \sigma^2 I)$
- Maximálně věrohodný odhad $b = (X'X)^{-1}X'y$
- Lepší odhady — Srivastava, Srivastava (1993): Steinovy odhady

$$b^S = \Big(1 - c\frac{(y - X\beta)'(y - X\beta)}{(n - k + 2)b'X'Xb}\Big)b$$

při $0 < c < 2(k-1)$, $k > 1$ a $L(b, \beta) = (b - \beta)'X'X(b - \beta)$ (asymptoticky, za předpokladu konvergence $X'X/n$ ke konečné nesingulární matici). Lze dále zlepšit — Hoa, Chaturvedi (1999) při $\sigma \to 0$.

# Menší model

- Oproti modelu s $X = (X_1, x_2)$ a $\beta = (\beta_1, \beta_2)$, kde $x_2$ je poslední sloupec $X$ a $\beta_2$ poslední složka $\beta$, uvažme

$$y = X_1\beta_1 + e^*, \qquad e^* = x_2\beta_2 + e$$

- Maximálně věrohodný odhad v modelu s omezením $\beta_2 = 0$

$$b^O = ((b_1^O)', 0)', \qquad \text{kde } b_1^O = (X_1'X_1)^{-1}X_1'y.$$

- Rozhodujeme se mezi $b^O$ a $b$ na základě testu o $\beta_2 = 0$, pre-test odhad

$$b_a^P = I_{[|T|<a]}b^O + I_{[|T|\geqq a]}b,$$

kde $T$ je testovací statistika a $a$ její kritická hodnota.

# Označení

- Ztrátová funkce $L(b,\beta) = (b-\beta)'Q(b-\beta)$, kde $Q = A'A$ a $A = (A_1, a_2)$
- Rozlišíme případy podle hodnot konstant

$$c = (a_2 - z)'(a_2 - z), \qquad \text{kde } z = A_1(X_1'X_1)^{-1}X_1'x_2$$

$$d = \sigma\sqrt{(a_2 - z)'A_1(X_1'X_1)^{-1}A_1'(a_2 - z)}$$

nezávisejících na volbě rozkladu $Q = A'A$.

- Např. pro model $y_i = \beta_1 + \beta_2 x_i + e_i$, $i = 1, \ldots, n$, se ztrátovou funkcí $\sum_i (b_1 + b_2 x_i - (\beta_1 + \beta_2 x_i))^2$ je $c = \sum_i (x_i - \bar{x})^2$ a $d = 0$, při $(b_1 + b_2 x_0 - (\beta_1 + \beta_2 x_0))^2$ je $c = (x_0 - \bar{x})^2$, $d = |x_0 - \bar{x}|\sigma/\sqrt{n}$.
- Obecně pro nesingulární $Q$ je $c \neq 0$, zatímco $d = 0$ např. pro $Q = X'X$ (také pro $Q = I$, když $x_2'X_1 = 0$).

# Odhad z chudšího modelu

- Pro $c = 0$ je $\mathrm{PC}_{\beta,\sigma^2}(b^O, b) = 0$, pro $c \neq 0$ a $d = 0$ je
$$\mathrm{PC}_{\beta,\sigma^2}(b^O, b) = 3 - 4\Phi(|\theta|),$$
kde $\theta = \beta_2/\sqrt{\operatorname{var} b_2} = \beta_2/\sqrt{\sigma^2[(X'X)^{-1}]_{22}}$. Při zúžení parametrického prostoru na $|\theta| < u_{0,75} \doteq 0,773$ je $b^O$ lepší než $b$, při $|\theta| > u_{0,75}$ naopak. Není tedy žádný z odhadů lepší.
- Při $c \neq 0$, $d \neq 0$ je $b^O$ lepší pro $|\theta| < \theta_1 = u_{0,75}$, zatímco $b$ pro $|\theta| > \theta_2 \doteq 0,877$, kde $\theta_2$ je jediné kladné řešení $\theta + 2\varphi(\theta) - 2\theta\Phi(\theta) = 0$. Pro $\theta \in (\theta_1, \theta_2)$ porovnání závisí i na $Q$.
- MSE preferuje $b^O$ až do $|\theta| \leqq 1$.

# Pre-test odhad

- Známe-li $\sigma^2$, použijeme $T = b_2/\sqrt{\sigma^2[(X'X)^{-1}]_{22}}$. Pro $d = 0$ je
$$\mathrm{PC}_\beta(b_a^P, b) = g_a(|\theta|),$$
kde $g_a(x) = \Phi(x+a) - \Phi(x-a) - 2\Phi(x) + 2I_{[x<a/2]}(\Phi(a-x) - \Phi(x))$.
- $b_a^P$ je preferován před $b$ při $|\theta| < \theta^*$, kde $\theta^*$ závisí jen na $a$ a je menší než $a/2$ i $u_{0,75}$.
- Neznáme-li $\sigma^2$, použijeme $T = b_2/\sqrt{s^2[(X'X)^{-1}]_{22}}$. Pro $d = 0$ je
$$\mathrm{PC}_{\beta,\sigma^2}(b_a^P, b) = \int_0^\infty g_{ax}(|\theta|) f(x)\,\mathrm{d}x$$
kde $f(x)$ je hustota $\sqrt{\chi^2(n-k)/(n-k)}$.
- Při $\theta = 0$ je preferováno $b_a^P$ před $b$, při $|\theta| > u_{0,75}$ naopak.

# PITMANOVO POROVNÁNÍ ODHADŮ

Michal Friesl
Jiří Reif

*Katedra matematiky*
*Fakulta aplikovaných věd*
*Západočeská univerzita v Plzni*

•http://home.zcu.cz/~friesl/Archiv/FriRob02.html