Štatistická kalibrácia a tolerančné oblasti

Martina Chvosteková

15. September (Sviatok Sedembolestnej Panny Márie - patrónky Slovenska), Jeseníky 2016

Ústav merania
Slovenská akadémia vied

Jednorozmerná absolútna kontrolovaná viacnásobná kalibrácia

Predpoklady

- dve jednorozmerné premenné: X, Y
- inferencia pre X, náročné získať jej hodnotu pre subjekt
- lineárna závislosť Y na X
- zistiť hodnotu Y pre subjekt je výrazne jednoduchšie ako získať hodnotu X pre subjekt

Jednorozmerná absolútna kontrolovaná viacnásobná kalibrácia

Úloha

Na základe *kalibračného experimentu* $\mathcal{E}_n$

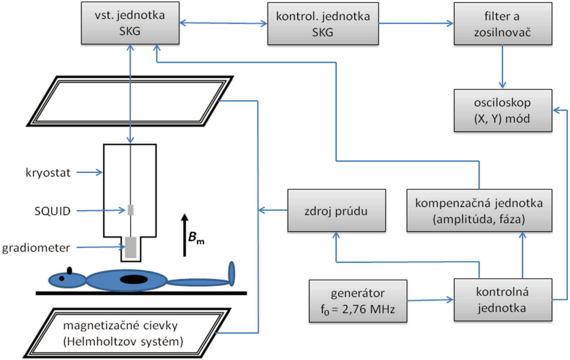
$$\mathcal{E}_n = \{(x_i, Y_i), i = 1, 2, \dots, n\}$$

- $x_i, i = 1, \dots, n$ - známe konštanty
- $Y_i = Y(x_i), i = 1, \dots, n$ - namerané hodnoty

pre každé budúce pozorovanie Y_{n+j} stanoviť simultánny kalibračný interval (viacnásobne použiteľný interval spoľahlivosti) pre prislúchajúcu hodnotu x_{n+j}, $j \to \infty$.

Jednorozmerná absolútna kontrolovaná viacnásobná kalibrácia

Ilustračný príklad



Jednorozmerná absolútna kontrolovaná viacnásobná kalibrácia

Ilustračný príklad - kalibračný experiment

Model trupu pozostával z dvoch polymetakrylátových nádob (25 cm $\times$ 30 cm $\times$ 25 cm) naplnených destilovanou vodou (diamagnetické prostredie), z dvoch identických vzduchom naplnených valcov (model pľúcnych lalokov) a modelu pečene, ktorý predstavoval polyetylénový valec s objemom 1000 cm^3 naplnený roztokom hexahydrát chloridu železitého ($FeCl_3 \cdot 6H_2O$).

Obr.: Model trupu a časť gradiometera umiestneného nad modelom pečene.

Jednorozmerná absolútna kontrolovaná viacnásobná kalibrácia

Kalibračný experiment

$$Y_i = Y(x_i) = \boldsymbol{f}^T(x_i)\boldsymbol{\beta} + \epsilon_i,\ \epsilon_i \sim N(0, \sigma^2)$$

- $x_1, x_2, \ldots, x_n$ - známe konštanty
- $\boldsymbol{f}^T(x)$ - známa funkcia, napr. $\boldsymbol{f}^T(x) = (1,\ x)$
- $\boldsymbol{\beta}^T = (\beta_0, \beta_1, \ldots, \beta_{q-1}), \sigma^2$ - neznáme parametre
- $\mathcal{X}$ - oblasť možných hodnôt x
- $\boldsymbol{f}^T(x)\boldsymbol{\beta}$ monotónna na $\mathcal{X}$
- $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)^T$
- $\boldsymbol{X} = (\boldsymbol{f}(x_1), \boldsymbol{f}(x_2), \ldots, \boldsymbol{f}(x_n))^T$

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y},\ \ S^2 = (\boldsymbol{Y} - \boldsymbol{f}^T(\boldsymbol{x})\hat{\boldsymbol{\beta}})^T(\boldsymbol{Y} - \boldsymbol{f}^T(\boldsymbol{x})\hat{\boldsymbol{\beta}})/(n-q)$$

Jednorozmerná absolútna kontrolovaná viacnásobná kalibrácia

Kalibračný interval pre hodnotu x

$$\mathcal{K} = \{x : \boldsymbol{f}^T(x)\hat{\boldsymbol{\beta}} - \lambda(x)S \leq Y \leq \boldsymbol{f}^T(x)\hat{\boldsymbol{\beta}} + \lambda(x)S\}$$

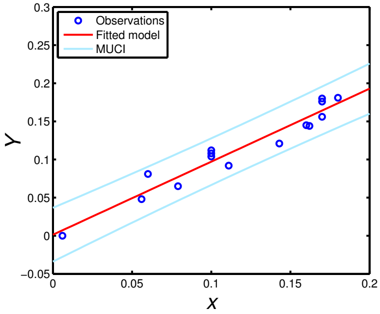
prislúchajúcu k budúcemu pozorovaniu $Y = \boldsymbol{f}^T(x)\boldsymbol{\beta} + \epsilon$, $\epsilon \sim N(0, \sigma^2)$.

Jednorozmerná absolútna kontrolovaná viacnásobná kalibrácia

Kalibračný interval pre hodnotu x

$$\mathcal{K} = \{x : \boldsymbol{f}^T(x)\hat{\boldsymbol{\beta}} - \lambda(x)S \leq Y \leq \boldsymbol{f}^T(x)\hat{\boldsymbol{\beta}} + \lambda(x)S\}$$

prislúchajúcu k budúcemu pozorovaniu $Y = \boldsymbol{f}^T(x)\boldsymbol{\beta} + \epsilon$, $\epsilon \sim N(0, \sigma^2)$.

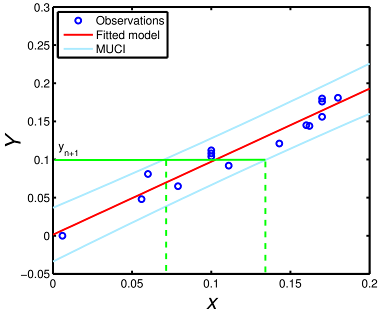


Jednorozmerná absolútna kontrolovaná viacnásobná kalibrácia

Kalibračný interval pre hodnotu x

$$\mathcal{K} = \{x : \boldsymbol{f}^T(x)\hat{\boldsymbol{\beta}} - \lambda(x)S \leq Y \leq \boldsymbol{f}^T(x)\hat{\boldsymbol{\beta}} + \lambda(x)S\}$$

prislúchajúcu k budúcemu pozorovaniu $Y = \boldsymbol{f}^T(x)\boldsymbol{\beta} + \epsilon$, $\epsilon \sim N(0, \sigma^2)$.

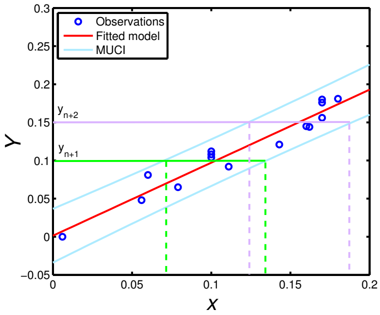


Jednorozmerná absolútna kontrolovaná viacnásobná kalibrácia

Kalibračný interval pre hodnotu x

$$\mathcal{K} = \{x : \boldsymbol{f}^T(x)\hat{\boldsymbol{\beta}} - \lambda(x)S \leq Y \leq \boldsymbol{f}^T(x)\hat{\boldsymbol{\beta}} + \lambda(x)S\}$$

prislúchajúcu k budúcemu pozorovaniu $Y = \boldsymbol{f}^T(x)\boldsymbol{\beta} + \epsilon$, $\epsilon \sim N(0, \sigma^2)$.

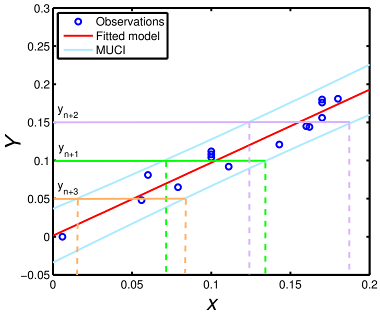


Jednorozmerná absolútna kontrolovaná viacnásobná kalibrácia

Kalibračný interval pre hodnotu x

$$\mathcal{K} = \{x : \boldsymbol{f}^T(x)\hat{\boldsymbol{\beta}} - \lambda(x)S \leq Y \leq \boldsymbol{f}^T(x)\hat{\boldsymbol{\beta}} + \lambda(x)S\}$$

prislúchajúcu k budúcemu pozorovaniu $Y = \boldsymbol{f}^T(x)\boldsymbol{\beta} + \epsilon$, $\epsilon \sim N(0, \sigma^2)$.



Jednorozmerná absolútna kontrolovaná viacnásobná kalibrácia

- konštrukcia intervalových odhadov pre $x_{n+1}, x_{n+2}, \ldots, x_{n+K}$ prislúchajúcich k K budúcim pozorovaniam $Y_{n+1}, Y_{n+2}, \ldots, Y_{n+K}$, $K \to \infty$
- intervalové odhady nazývame viacnásobne použiteľné intervaly spoľahlivosti (ang. multiple use confidence intervals - MUCI) alebo simultánne kalibračné intervaly
- MUCI sú skonštruované na základe jediného $\mathcal{E}_n$
- aspoň γ skonštruovaných MUCI pokrýva skutočnú prislúchajúcu hodnotu x so spoľahlivosťou $1 - \alpha$

Jednorozmerná absolútna kontrolovaná viacnásobná kalibrácia

Podmienka viacnásobne použiteľných intervalov spoľahlivosti (MUCI)

$$P_{\hat{\beta},S}\left[\frac{1}{K}\sum_{i=1}^{K} C(x_{n+i};\hat{\beta},S) \geq \gamma\right] = 1-\alpha$$

$$C(x;\hat{\beta},S) = P_{Y(x)}(\boldsymbol{f}^T(x)\hat{\beta} - \lambda(x)S \leq Y(x) \leq \boldsymbol{f}^T(x)\hat{\beta} + \lambda(x)S|\hat{\beta},S\}$$

Podmienka simultánnych tolerančných intervalov
(postačujúca podmienka pre MUCI)

$$P_{\hat{\beta},S}\left[\min_{x\in\mathcal{X}} C(x;\hat{\beta},S) \geq \gamma\right] = 1-\alpha$$

Jednorozmerná absolútna kontrolovaná viacnásobná kalibrácia

Podmienka simultánnych tolerančných intervalov
(postačujúca podmienka pre MUCI)

$$P_{\hat{\boldsymbol{\beta}},S}\left[\min_{x\in\mathcal{X}} C(x;\hat{\boldsymbol{\beta}},S)\geq\gamma\right]=1-\alpha$$

$$\begin{aligned}
C(x;\hat{\boldsymbol{\beta}},S) &= P_{Y(x)}(\boldsymbol{f}^T(x)\hat{\boldsymbol{\beta}}-\lambda(x)S\leq Y(x)\leq \boldsymbol{f}^T(x)\hat{\boldsymbol{\beta}}+\lambda(x)S|\hat{\boldsymbol{\beta}},S\} \\
&= P_{Y(\boldsymbol{x})}\left(\boldsymbol{f}^T(x)\frac{\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}}{\sigma}-\lambda(x)\frac{S}{\sigma}\leq\frac{Y(\boldsymbol{x})-\boldsymbol{f}^T(\boldsymbol{x})\hat{\boldsymbol{\beta}}}{\sigma}\leq\boldsymbol{f}^T(x)\frac{\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}}{\sigma}+\lambda(x)\frac{S}{\sigma}\right) \\
&= \Phi(\boldsymbol{f}^T(x)\boldsymbol{Z}_{\boldsymbol{X}}+\lambda(x)U)-\Phi(\boldsymbol{f}^T(x)\boldsymbol{Z}_{\boldsymbol{X}}-\lambda(x)U) \\
&= C(x;\boldsymbol{Z}_{\boldsymbol{X}},U)
\end{aligned}$$

$$\boldsymbol{Z}_{\boldsymbol{X}}=\frac{\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}}{\sigma}\sim N(\boldsymbol{0},(\boldsymbol{X}^T\boldsymbol{X})^{-1}),\quad U=\frac{S}{\sigma}\sim\sqrt{\frac{\chi^2_v}{v}},v=n-q$$

Jednorozmerná absolútna kontrolovaná viacnásobná kalibrácia

Podmienka simultánnych tolerančných intervalov
(postačujúca podmienka pre MUCI)

$$P_{\hat{\boldsymbol{\beta}},S}\left[\min_{x\in\mathcal{X}} C(x;\hat{\boldsymbol{\beta}},S)\geq\gamma\right]=1-\alpha$$

$$P_{\boldsymbol{Z_X},U}\left[\min_{x\in\mathcal{X}} C(x;\boldsymbol{Z_X},U)\geq\gamma\right]=1-\alpha$$

$$C(x;\boldsymbol{Z_X},U)=\Phi(\boldsymbol{f}^T(x)\boldsymbol{Z_X}+\lambda(x)U)-\Phi(\boldsymbol{f}^T(x)\boldsymbol{Z_X}-\lambda(x)U)$$

Pivotné premenné

$$\boldsymbol{Z_X}=\frac{\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}}{\sigma}\sim N(\mathbf{0},(\boldsymbol{X}^T\boldsymbol{X})^{-1}),\quad U=\frac{S}{\sigma}\sim\sqrt{\frac{\chi^2_v}{v}},v=n-q$$

Jednorozmerná absolútna kontrolovaná viacnásobná kalibrácia

Podmienka simultánnych tolerančných intervalov
(postačujúca podmienka pre MUCI)

$$P_{\boldsymbol{Z_X},U}\left[\min_{x\in\mathcal{X}} C(x;\boldsymbol{Z_X},U)\geq\gamma\right]=1-\alpha$$

$$C(x;\boldsymbol{Z_X},U)=\Phi(\boldsymbol{f}^T(x)\boldsymbol{Z_X}+\lambda(x)U)-\Phi(\boldsymbol{f}^T(x)\boldsymbol{Z_X}-\lambda(x)U)$$

Lieberman a Miller (1963), Lieberman et al. (1967), Wilson (1967), Scheffé (1973), Limam a Thomas (1988), Odeh a Mee (1990), Mee et al. (1991), Witkovský (2014), Han et al. (2015)

Jednorozmerná absolútna kontrolovaná viacnásobná kalibrácia

Podmienka simultánnych tolerančných intervalov
(postačujúca podmienka pre MUCI)

$$P_{\boldsymbol{Z_X},U}\left[\min_{x\in\mathcal{X}} C(x;\boldsymbol{Z_X},U)\geq\gamma\right]=1-\alpha$$

*Hahn, Y., et al., Statistical calibration and exact one-sided simultaneous tolerance intervals for polynomial regression. Journal of Statistical Planning and Inference, 2015.

Jednostranné simultánne tolerančné intervaly sú totožné s jednostranným pásom spoľahlivosti pre γ-kvantilovú krivku, t.j.

$\boldsymbol{f}^T(x)\boldsymbol{\beta}+z_\gamma\sigma$, z_γ je γ-kvantil $N(0,1)$

Jednorozmerná absolútna kontrolovaná viacnásobná kalibrácia

Podmienka simultánnych tolerančných intervalov
(postačujúca podmienka pre MUCI)

$$P_{\boldsymbol{Z_X},U}\left[\min_{x\in\mathcal{X}} C(x;\boldsymbol{Z_X},U)\geq\gamma\right]=1-\alpha$$

Modifikácia Hahn a kol.(2015) pre obojstranný prípad

$\lambda(x)=\lambda[z_{(1+\gamma)/2}+\sqrt{q+2}d(x)],\ \ d^2(x)=\boldsymbol{f}^T(x)(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{f}^T(x)$

1. vygenerovať 1000 000 realizácií $\boldsymbol{Z_X}, U$
2. voľba λ
3. pre každé i, $i=1,2,\ldots,1000\,000$ sa určí

$$W_i=\min_{x\in\mathcal{X}} C(x;\boldsymbol{Z}_{\boldsymbol{X},i},U_i)$$

4. určí sa pomer (R) $\boldsymbol{Z}_{\boldsymbol{X},i}, U_i$, $i=1,2,\ldots,1000\,000$, pre ktoré $W_i\geq\gamma$
5. body 2 - 4 sa opakujú, kým $R=1-\alpha$

Jednorozmerná absolútna kontrolovaná viacnásobná kalibrácia

Podmienka simultánnych tolerančných intervalov
(postačujúca podmienka pre MUCI)

$$P_{\boldsymbol{Z_x},U}\left[\min_{x\in\mathcal{X}} C(x;\boldsymbol{Z_X},U)\geq\gamma\right]=1-\alpha$$

$$P_{\boldsymbol{Z_x},U}\left[\frac{1}{K}\sum_{i=1}^{K} C(x_{n+i};\boldsymbol{Z_X},U)\geq\gamma\right]=1-\alpha$$

Jednorozmerná absolútna kontrolovaná viacnásobná kalibrácia

Podmienka simultánnych tolerančných intervalov
(postačujúca podmienka pre MUCI)

$$P_{\boldsymbol{Z_x},U}\left[\min_{x\in\mathcal{X}} C(x;\boldsymbol{Z_x},U)\geq\gamma\right]=1-\alpha$$

$$P_{\boldsymbol{Z_x},U}\left[\frac{1}{K}\sum_{i=1}^{K} C(x_{n+i};\boldsymbol{Z_x},U)\geq\gamma\right]=1-\alpha$$

$x\sim R(\mathcal{X})$

$$P_{\boldsymbol{Z_x},U}\left[\int_{\mathcal{X}} C(x;\boldsymbol{Z_x},U)f_{R(\mathcal{X})}dx\geq\gamma\right]=1-\alpha$$

Jednorozmerná absolútna kontrolovaná viacnásobná kalibrácia

$x \sim R(\mathcal{X})$

$$P_{\boldsymbol{Z_X},U}\left[\int_{\mathcal{X}} C(x;\boldsymbol{Z_X},U) f_{R(\mathcal{X})} dx \geq \gamma\right] = 1-\alpha$$

$$\begin{aligned} *C(x;\boldsymbol{Z_X},U) &= \Phi(\boldsymbol{f}^T(x)\boldsymbol{Z_X} + \lambda(\boldsymbol{x})U) - \Phi(\boldsymbol{f}^T(x)\boldsymbol{Z_X} - \lambda(\boldsymbol{x})U) \\ &\geq \Phi(d(x)W_q + \lambda(\boldsymbol{x})U) - \Phi(d(x)W_q - \lambda(\boldsymbol{x})U) \end{aligned}$$

$W_q^2 \sim \chi_q^2$, $\max_{\{x:\boldsymbol{f}^T(x)(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{f}(x)=d(x)\}}(\boldsymbol{f}^T(x)\boldsymbol{Z_X})^2 = d^2(x)\boldsymbol{Z_X}^T(\boldsymbol{X}^T\boldsymbol{X})\boldsymbol{Z_X}$

$$P_{W_q,U}\left[\int_{\mathcal{X}} C(x;W_q,U) f_{R(\mathcal{X})} dx \geq \gamma\right] = 1-\alpha$$

$*$ Mee et al. (1991) Calibration and simultaneous tolerance intervals for regression. Technometrics 38, 221-229.

Jednorozmerná absolútna kontrolovaná viacnásobná kalibrácia

Hondoty λ pre $\lambda(x) = \lambda(z_{(1+\gamma)/2} + \sqrt{q+2}d(x))$ odpovedajúce presným STI a navrhnutou metódou pre prípad $\boldsymbol{f}^T(x) = (1,\ x)$, $\alpha = 0.05$, $\gamma = 0.9$.

	STI*			nové		
	$\tau = 2$	$\tau = 3$	$\tau = 4$	$\tau = 2$	$\tau = 3$	$\tau = 4$
$n = 10$	1.3674	1.3787	1.3857	1.3382	1.3365	1.341
$n = 20$	1.1485	1.1575	1.1639	1.1192	1.1155	1.12
$n = 30$	1.0891	1.0964	1.1021	1.0593	1.052	1.054
$n = 40$	1.0619	1.0676	1.0725	1.0317	1.0217	1.0213
$n = 50$	1.0464	1.051	1.0553	1.0161	1.0041	1.0017

Za predpokladu $\bar{x} = 0$ platí $\boldsymbol{f}^T(x)(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{f}(x) = 1/n + x^2/S_{xx}$, kde $S_{xx} = \sum_{i=1}^{n} x_i^2$. Pre jednoduchosť sme uvažovali $\mathcal{X} = [-\tau, \tau]$ pričom $x_{max}^2/S_{xx} = x_{min}^2/S_{xx} = \tau^2/n$ ako v $*$Mee a kol. (1991).

Jednorozmerná absolútna kontrolovaná viacnásobná kalibrácia

Spoľahlivosť

$$P_{\boldsymbol{Z_X},U}\left[\frac{1}{K}\sum_{i=1}^{K} C(x_{n+i};\boldsymbol{Z_X},U)\geq\gamma\right]$$

navrhnutých MUCI sme numericky skúmali pre 3 rôzne postupnosti $\{x_{n+i}\}_{i=1}^{10000}$, $\mathcal{X}$.

$S1$ pozostávala x_{n+i}, $i=1,2,\ldots,10000$ generovaných z $R(\mathcal{X})$
$S2$ pozostávala x_{n+i}, $i=1,2,\ldots,10000$ generovaných z $Tr(\mathcal{X},0)$
$S3$ pozostávala x_{n+i}, $i=1,2,\ldots,10000$ generovaných z $Tr(\mathcal{X},\tau)$

- 1000 000 realizácií $\boldsymbol{Z_X}, U$

Jednorozmerná absolútna kontrolovaná viacnásobná kalibrácia

S_1: x_{n+i}, $i = 1, 2, \ldots, 10000$ generovaných z $R(\mathcal{X})$, $\mathcal{X} = [-\tau, \tau]$

$$1 - \alpha = 0.95$$

	$\tau = 2$	$\tau = 3$	$\tau = 4$
$n = 10$	0.9688	0.9725	0.9753
$n = 20$	0.9724	0.9776	0.9825
$n = 30$	0.9728	0.9797	0.9827
$n = 40$	0.9729	0.9824	0.9874
$n = 50$	0.9699	0.9780	0.9830

Jednorozmerná absolútna kontrolovaná viacnásobná kalibrácia

S_2: x_{n+i}, $i = 1, 2, \ldots, 10000$ generovaných z $Tr(\mathcal{X}, 0)$, $\mathcal{X} = [-\tau, \tau]$

$$1 - \alpha = 0.95$$

	$\tau = 2$	$\tau = 3$	$\tau = 4$
$n = 10$	0.9647	0.9684	0.9711
$n = 20$	0.9644	0.9711	0.9759
$n = 30$	0.9621	0.9700	0.9773
$n = 40$	0.9584	0.9678	0.9772
$n = 50$	0.9561	0.9643	0.9724

Jednorozmerná absolútna kontrolovaná viacnásobná kalibrácia

S_3: x_{n+i}, $i = 1, 2, \ldots, 10000$ generovaných z $Tr(\mathcal{X}, \tau)$, $\mathcal{X} = [-\tau, \tau]$

$$1 - \alpha = 0.95$$

	$\tau = 2$	$\tau = 3$	$\tau = 4$
$n = 10$	0.9678	0.9723	0.9755
$n = 20$	0.9709	0.9769	0.9810
$n = 30$	0.9717	0.9795	0.9828
$n = 40$	0.9708	0.9807	0.9852
$n = 50$	0.9692	0.9774	0.9825

Viacrozmerná absolútna kontrolovaná viacnásobná kalibrácia

Kalibračný experiment

$$\boldsymbol{Y}_i = \boldsymbol{B}\boldsymbol{x}_i + \epsilon_i, \ \epsilon_i \sim N_q(\mathbf{0}, \Sigma)$$

- $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ - známe $m \times 1$ vektory
- $\boldsymbol{B}$ $q \times m$ matica
- Σ $q \times q$ pozitívne definitná matica
- $\boldsymbol{B}$, Σ - neznáme parametre
- $\boldsymbol{Y} = (\boldsymbol{Y}_1, \boldsymbol{Y}_2, \ldots, \boldsymbol{Y}_n)$
- $\boldsymbol{X} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)$

$\hat{\boldsymbol{B}} = \boldsymbol{Y}\boldsymbol{X}^T(\boldsymbol{X}\boldsymbol{X}^T)^{-1}, \hat{\Sigma} = \boldsymbol{A}/(n-m), \ \boldsymbol{A} = \boldsymbol{Y}[\boldsymbol{I}_n - \boldsymbol{X}^T(\boldsymbol{X}\boldsymbol{X}^T)^{-1}\boldsymbol{X}]\boldsymbol{Y}^T \sim W_q(\Sigma, n-m)$

Viacrozmerná absolútna kontrolovaná viacnásobná kalibrácia

$$\boldsymbol{Y} \sim N(\boldsymbol{BX}, \boldsymbol{I}_n \otimes \Sigma), \;\; \boldsymbol{Y}(\boldsymbol{x}) \sim N(\boldsymbol{Bx}, \Sigma)$$

Ak by boli $\boldsymbol{B}, \Sigma$ známe $\rightarrow \widetilde{x} = (\boldsymbol{B}^T \Sigma^{-1} \boldsymbol{B})^{-1} \boldsymbol{B}^T \Sigma^{-1} \boldsymbol{Y}(\boldsymbol{x})$
s kovariančnou maticou $(\boldsymbol{B}^T \Sigma^{-1} \boldsymbol{B})^{-1}$ a vhodný pivot na konštrukciu oblasti pre $\boldsymbol{x}$ by bol $(\widetilde{\boldsymbol{x}} - \boldsymbol{x})^T \boldsymbol{B}^T \Sigma^{-1} \boldsymbol{B} (\widetilde{\boldsymbol{x}} - \boldsymbol{x})$.

Krishnamoorthy, K. and Mathew, T.(2008) Statistical Tolerance Regions: Theory, Applications, and Computation, Wiley

Mathew, T., and Zha, W. (1997) Multiple Use Confidence Regions in Multivariate Calibration, Journal of the American Statistical Association 92, 1141–1150.

Mathew, T., and Zha, W. (1996) Conservative Confidence Regions in Multivariate Calibration, The Annals of Statistics 24, 707–725.

Viacrozmerná absolútna kontrolovaná viacnásobná kalibrácia

$$\boldsymbol{Y} \sim N(\boldsymbol{BX}, \boldsymbol{I}_n \otimes \Sigma), \;\; \boldsymbol{Y}(\boldsymbol{x}) \sim N(\boldsymbol{BX}, \Sigma)$$

Ak by boli $\boldsymbol{B}, \boldsymbol{\Sigma}$ známe $\rightarrow \widetilde{x} = (\boldsymbol{B}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{B})^{-1} \boldsymbol{B}^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{Y}(\boldsymbol{x}) - \boldsymbol{B}_0)$
s kovariančnou maticou $(\boldsymbol{B}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{B})^{-1}$ a vhodný pivot na konštrukciu oblasti pre $\boldsymbol{x}$ by bol $(\widetilde{\boldsymbol{x}} - \boldsymbol{x})^T \boldsymbol{B}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{B} (\widetilde{\boldsymbol{x}} - \boldsymbol{x})$.

$\boldsymbol{B}, \boldsymbol{\Sigma}$ neznáme $\rightarrow \hat{x} = (\hat{\boldsymbol{B}}^T \boldsymbol{A}^{-1} \hat{\boldsymbol{B}})^{-1} \hat{\boldsymbol{B}}^T \boldsymbol{A}^{-1} \boldsymbol{Y}(\boldsymbol{x})$

$$\begin{aligned} T(\boldsymbol{x}) = & \frac{n-m-q+1}{m} (\hat{\boldsymbol{x}} - \boldsymbol{x})^T \hat{\boldsymbol{B}}^T \boldsymbol{A}^{-1} \hat{\boldsymbol{B}} (\hat{\boldsymbol{x}} - \boldsymbol{x}) \\ = & \frac{n-m-q+1}{m} [Y(\boldsymbol{x}) - \hat{\boldsymbol{B}}\boldsymbol{x}]^T \boldsymbol{A}^{-1} \hat{\boldsymbol{B}} (\hat{\boldsymbol{B}}^T \boldsymbol{A}^{-1} \hat{\boldsymbol{B}})^{-1} \hat{\boldsymbol{B}}^T \boldsymbol{A}^{-1} \times \\ & [Y(\boldsymbol{x}) - \hat{\boldsymbol{B}}\boldsymbol{x}] \end{aligned}$$

Mathew, T., and Zha, W. (1996) Conservative Confidence Regions in Multivariate Calibration, The Annals of Statistics 24, 707–725.

Viacrozmerná absolútna kontrolovaná viacnásobná kalibrácia

Pivotná premenná pre konštrukciu simultánnej kalibračnej oblasti

$$
\begin{aligned}
T(\boldsymbol{x}) =& \frac{n-m-q+1}{m}(\hat{\boldsymbol{x}}-\boldsymbol{x})^T\hat{\boldsymbol{B}}^T\boldsymbol{A}^{-1}\hat{\boldsymbol{B}}(\hat{\boldsymbol{x}}-\boldsymbol{x}) \\
=& \frac{n-m-q+1}{m}[Y(\boldsymbol{x})-\hat{\boldsymbol{B}}\boldsymbol{x}]^T\boldsymbol{A}^{-1}\hat{\boldsymbol{B}}(\hat{\boldsymbol{B}}^T\boldsymbol{A}^{-1}\hat{\boldsymbol{B}})^{-1}\hat{\boldsymbol{B}}^T\boldsymbol{A}^{-1}\times \\
& [Y(\boldsymbol{x})-\hat{\boldsymbol{B}}\boldsymbol{x}]
\end{aligned}
$$

Simultánna kalibračná oblasť

$$\{\boldsymbol{x}:\ T(\boldsymbol{x}) \leq \lambda(\boldsymbol{x})\}$$

Viacrozmerná absolútna kontrolovaná viacnásobná kalibrácia

Podmienka viacnásobne použiteľných konfidenčných oblastí

$$P_{\hat{\boldsymbol{B}},\boldsymbol{A}}\left[\frac{1}{K}\sum_{i=1}^{n} C(\boldsymbol{x}_{n+i};\hat{\boldsymbol{B}},\boldsymbol{A}) \geq \gamma\right] = 1-\alpha$$

$$C(\boldsymbol{x};\hat{\boldsymbol{B}},\boldsymbol{A}) = P_{Y(\boldsymbol{x})}\{T(\boldsymbol{x}) \leq \lambda(\boldsymbol{x})|\hat{\boldsymbol{B}},\boldsymbol{A}\}$$

Podmienka simultánnych tolerančných oblastí

$$P_{\hat{\boldsymbol{B}},\boldsymbol{A}}\left[\min_{\boldsymbol{x}} C(\boldsymbol{x};\hat{\boldsymbol{B}},\boldsymbol{A}) \geq \gamma\right] = 1-\alpha$$

Viacrozmerná absolútna kontrolovaná viacnásobná kalibrácia

Ak $q = m$,

$$
\begin{aligned}
*T(\boldsymbol{x}) &= \frac{n-2m+1}{m}[Y(\boldsymbol{x}) - \hat{\boldsymbol{B}}\boldsymbol{x}]^T \boldsymbol{A}^{-1}[Y(\boldsymbol{x}) - \hat{\boldsymbol{B}}\boldsymbol{x}] \\
&= \frac{n-2m+1}{m}(\boldsymbol{Z} - d\boldsymbol{H})^T \boldsymbol{V}^{-1}(\boldsymbol{Z} - d\boldsymbol{H})
\end{aligned}
$$

$\boldsymbol{Z} \sim N(\boldsymbol{0}, \boldsymbol{I}_q)$, $d^2 = \boldsymbol{x}^T(\boldsymbol{X}\boldsymbol{X}^T)^{-1}\boldsymbol{x}$, $\boldsymbol{H} \sim N(0, \boldsymbol{I}_q)$, $\boldsymbol{V} \sim W_q(\boldsymbol{I}_q, n-m)$

Krishnamoorthy, K. and Mathew, T.(2008) Statistical Tolerance Regions: Theory, Applications, and Computation, Wiley

Simultánne tolerančné oblasti

$$
P_{\boldsymbol{H},\boldsymbol{V}}\left[\min_d P_{\boldsymbol{Z}}\left\{\frac{n-2m+1}{m}(\boldsymbol{Z} - d\boldsymbol{H})^T \boldsymbol{V}^{-1}(\boldsymbol{Z} - d\boldsymbol{H}) \leq \lambda(d)|\boldsymbol{H}, \boldsymbol{V}\right\} \geq \gamma\right] = 1-\alpha
$$

Viacrozmerná absolútna kontrolovaná viacnásobná kalibrácia

$\lambda(d) = \lambda[z_{(1+\gamma)/2} + \sqrt{q+2}d]$, $d \in \mathcal{D}$

Algoritmus

1. $d_l \in \mathcal{D}$, $l = 1, 2, \ldots, m_1$ ($m_1 = 1000$)

2. vygenerujú sa $\boldsymbol{H}_i \sim N(0, \boldsymbol{I}_q)$, $\boldsymbol{V}_i \sim W_q(\boldsymbol{I}_q, n-m)$, $i = 1, 2, \ldots, m_2$; vygenerujú $\boldsymbol{Z}_{ij} \sim N_q(\boldsymbol{0}, \boldsymbol{I}_q)$, $j = 1, 2, \ldots, m_2$ ($m_2 = 100\,000$, $m_3 = 10\,000$)

3. zvolí sa λ

4. pre každé i sa pre každdú hodnotu d_l sa stanoví pomer ($R_{Zi}(d_l)$) Z_{ij}, $j = 1, 2, \ldots, m_3$ pre ktoré

$$\frac{n-2m+1}{m}(\boldsymbol{Z}_{ij} - d_l\boldsymbol{H}_i)^T \boldsymbol{V}_i^{-1}(\boldsymbol{Z}_{ij} - d_l\boldsymbol{H}_i) \leq \lambda(d_l)$$

5. stanoví sa pomer (R) $\boldsymbol{H}_i$, $\boldsymbol{V}_i$ pre ktoré $\min_{d_1,\ldots,d_l} R_{Zi}(d) \geq \gamma$ 6. zmeníme λ a zopakujem 3-5 kym $R = 1 - \alpha$

Viacrozmerná absolútna kontrolovaná viacnásobná kalibrácia

Ak $q \geq m$

$$
\begin{aligned}
U(\boldsymbol{x}) =& \frac{n-m-q+1}{m} \frac{[Y(\boldsymbol{x}) - \hat{\boldsymbol{B}}\boldsymbol{x}]^T \boldsymbol{A}^{-1} \hat{\boldsymbol{B}} (\hat{\boldsymbol{B}}^T \boldsymbol{A}^{-1} \hat{\boldsymbol{B}})^{-1} \hat{\boldsymbol{B}}^T \boldsymbol{A}^{-1} [Y(\boldsymbol{x}) - \hat{\boldsymbol{B}}\boldsymbol{x}]}{1 + Y^T(\boldsymbol{x}) [\boldsymbol{A}^{-1} - \boldsymbol{A}^{-1} \hat{\boldsymbol{B}} (\hat{\boldsymbol{B}}^T \boldsymbol{A}^{-1} \hat{\boldsymbol{B}})^{-1} \hat{\boldsymbol{B}}^T \boldsymbol{A}^{-1}] Y(\boldsymbol{x})} \\
\approx& F_{m,n-q-m+1} \left(\frac{d^2 G}{1 + (q-m)F/(n-q+1)} \right)
\end{aligned}
$$

$G \sim \chi^2_q, \quad F \sim F_{q-m,n-q+1}$

Mathew, T., and Zha, W. (1996) Conservative Confidence Regions in Multivariate Calibration, The Annals of Statistics 24, 707–725.

Viacrozmerná absolútna kontrolovaná viacnásobná kalibrácia

Ak $q \geq m$

$$P_{F,G}\left[\min_d P\left\{F_{m,n-q-m+1}\left(\frac{d^2 G}{1+(q-m)F/(n-q+1)}\right) \leq \lambda(d)|F,G\right\} \geq \gamma\right] = 1-\alpha$$

$$\lambda(d) = \lambda[z_{(1+\gamma)/2} + \sqrt{q+2}d],\, d \in \mathcal{D}$$

Ďakujem za pozoronosť!