Národní informační středisko pro podporu kvality

Problémy s ukazateli způsobilosti a výkonnosti v praxi

Dr.Jiří Michálek, CSc.

Ústav teorie informace a automatizace AVČR

Ukazatel způsobilosti C_p

Předpoklady: • normální rozdělení $N(\mu , \sigma^2)$ znaku jakosti;
k podskupin o rozsahu n kusů
k*n = N celkový počet kontrolovaných kusů

.

$$C_p = \frac{USL - LSL}{6\sigma}$$

Ukazatel způsobilosti C_{pk}

Předpoklady: normální rozdělení $N(\mu , \sigma^2)$ znaku jakosti;
k podskupin o rozsahu n kůusů
k*n = N celkový počet kontrolovaných kusů

$$C_{pk} = \min\left\{\frac{USL - \mu}{3\sigma}, \frac{\mu - LSL}{3\sigma}\right\}$$

Ukazatel výkonnosti P_p

Předpoklady:
- normální rozdělení N(μ , σ^2) pro znak jakosti
- jeden náhodný výběr o N kusech

  sběr dat po podskupinách je ignorován

$$P_p = \frac{USL - LSL}{6\,\sigma_{tot}}$$

Ukazatel výkonnosti P_{pk}

Předpoklady: • normální rozdělení $N(\mu , \sigma^2)$ pro znak jakosti;
• jeden náhodný výběr o N kusech
sběr dat po podskupinách je ignorován

$$P_{pk} = \min\left\{\frac{USL-\mu}{3\sigma_{tot}}, \frac{\mu-LSL}{3\sigma_{tot}}\right\}$$

Odhad ukazatele C_p

$$\hat{\sigma} = \bar{R}/d_2; \quad \bar{s}/c_4; \quad \sqrt{\frac{1}{k}\sum_{j=1}^{k} s_j^2}$$

Je pouze využita informace o úrovni variability uvnitř podskupin- nejlepší odhad inherentní variability

$$s_i^2 = \frac{1}{n-1}\sum_{j=1}^{n}(x_{ij} - \bar{x}_i)^2$$

pro j = 1, 2, ..., k z j-té podskupiny

$$\bar{s} = \frac{1}{k}\sum_{j=1}^{k} s_j$$

.

Proces musí být statisticky zvládnut

Pozorování jsou organizována v podskupinách o rozsahu n,

data jsou x_{ij}, $i = 1, 2, ..., k$; $j = 1, 2, ..., n$.

Směrodatná odchylka σ je odhadována pomocí

$$s_R = \frac{\bar{R}}{d_2(n)} \;,$$

kde v i-té podskupině $R_i = \max x_{ij} - \min x_{ij}$.

Funkce hustoty pravděpodobnosti lze pak vyjádřit pomocí :

$$f_{\hat{C}_p}(x) \cong \frac{1}{\sqrt{2\pi}} e^{-\frac{k}{2}\frac{\alpha_n^2}{\beta_n^2}\left(\frac{C_p}{x}-1\right)^2} \frac{\alpha_n}{\beta_n}\frac{\sqrt{k}}{x^2} C_p \qquad \text{pro } x > 0 \;,$$

$$f_{\hat{C}_p}(x) = 0 \qquad \text{pro } x \le 0 \;.$$

α_n a β_n jsou tabelovány a závisí na rozsahu podskupiny n.

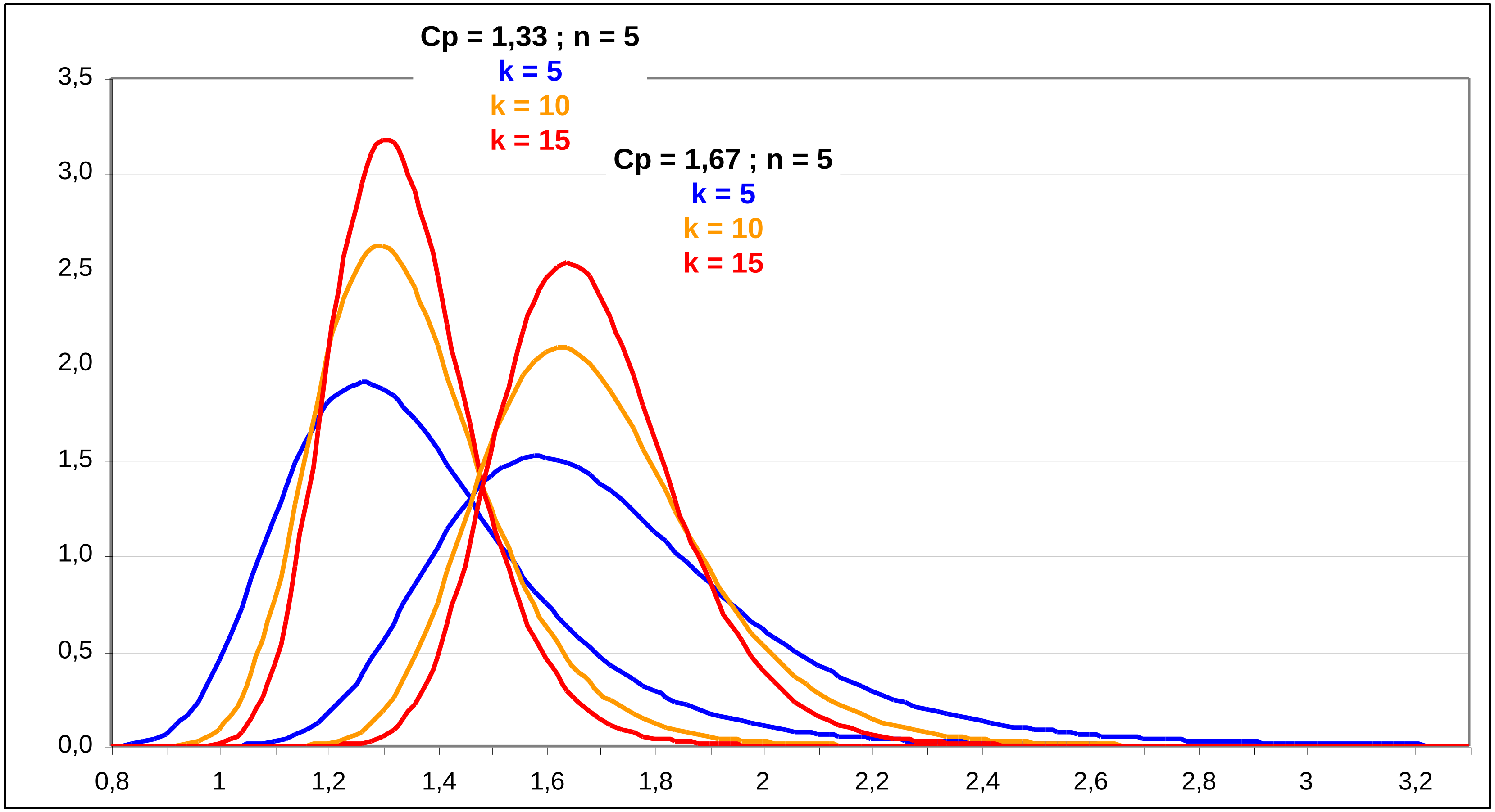
[illegible] .

Příklad: pro k = 10, n = 5, C_p = 1,33

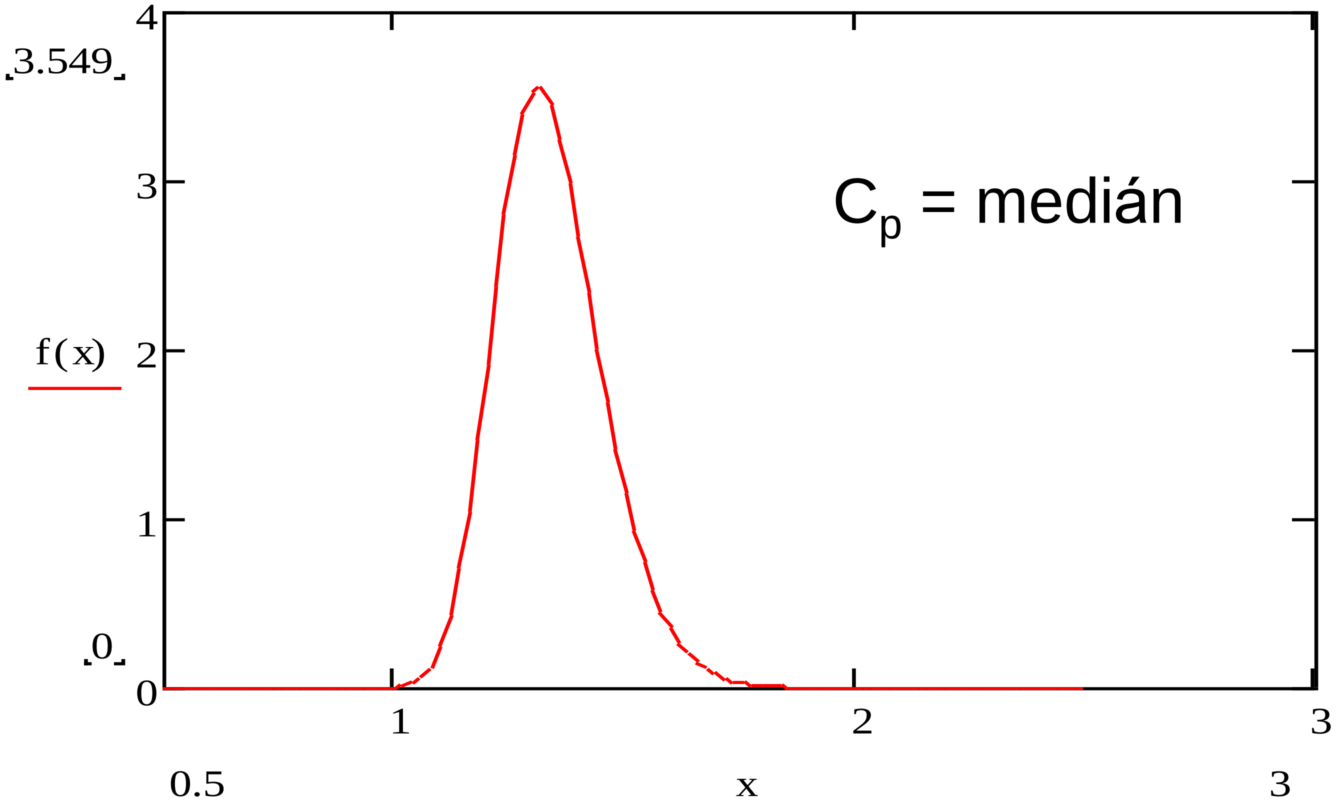
Odpovídající kvantily jsou:

α	$C_p(\alpha)$
0,01	1,045
0,025	1,081
0,05	1,115
0,5	1,330
0,95	1,649
0,975	1,728
0,99	1,830

Funkce hustoty pro odhad $\hat{C}_p$ založený na $\dfrac{\overline{R}}{d_2(n)}$



Hustota pro odhad ukazatele C_p (25 podskupin, 4 kusy, C_p=1,33)



Odhady pro C_{pk}

$$\sigma \approx \bar{R}/d_2;\quad \bar{s}/c_4;\quad \sqrt{\frac{1}{k}\sum_{j=1}^{k} s_j^2}$$

$$\mu \approx \bar{\bar{x}} = \frac{1}{k}\sum_{j=1}^{k} \bar{x}_j$$

Opět tyto odhady jsou založeny na informaci pouze v podskupinách. Proces musí být statisticky zvládnut jak pro parametr polohy tak pro úroveň inherentní variability

Jeden z přístupů k odvození tvaru hustoty
pro odhad ukazatele C_{pk},
data v podskupinách normální distribuce $N(\mu, \sigma^2)$.

Platí: $C_{pk} = (1-K)\, C_p$,

kde

$$\Delta = 0.5\,(USL - LSL)\ ,\quad K = \frac{\left|\mu - \frac{1}{2}(USL + LSL)\right|}{\Delta}$$

tudíž

$$\hat{C}_{pk} = (1 - \hat{K})\,\hat{C}_p$$

$$\hat{K} = \frac{\left|\bar{x} - \frac{1}{2}(USL + LSL)\right|}{\Delta} \ .$$

Důležité:

Proměnné $\hat{C}_p$ a $\hat{K}$ jsou nezávislé za předpokladu $N(\mu, \sigma^2)$.

Obecná formule pro odhad $\hat{C}_{pk}$

$$f_{\hat{C}_{pk}}(x) = \int_x^{\infty} \frac{1}{u} f_{\hat{C}p}(u)\, f_{1-\hat{K}}\left(\frac{x}{u}\right) du \qquad \text{pro } x > 0,$$

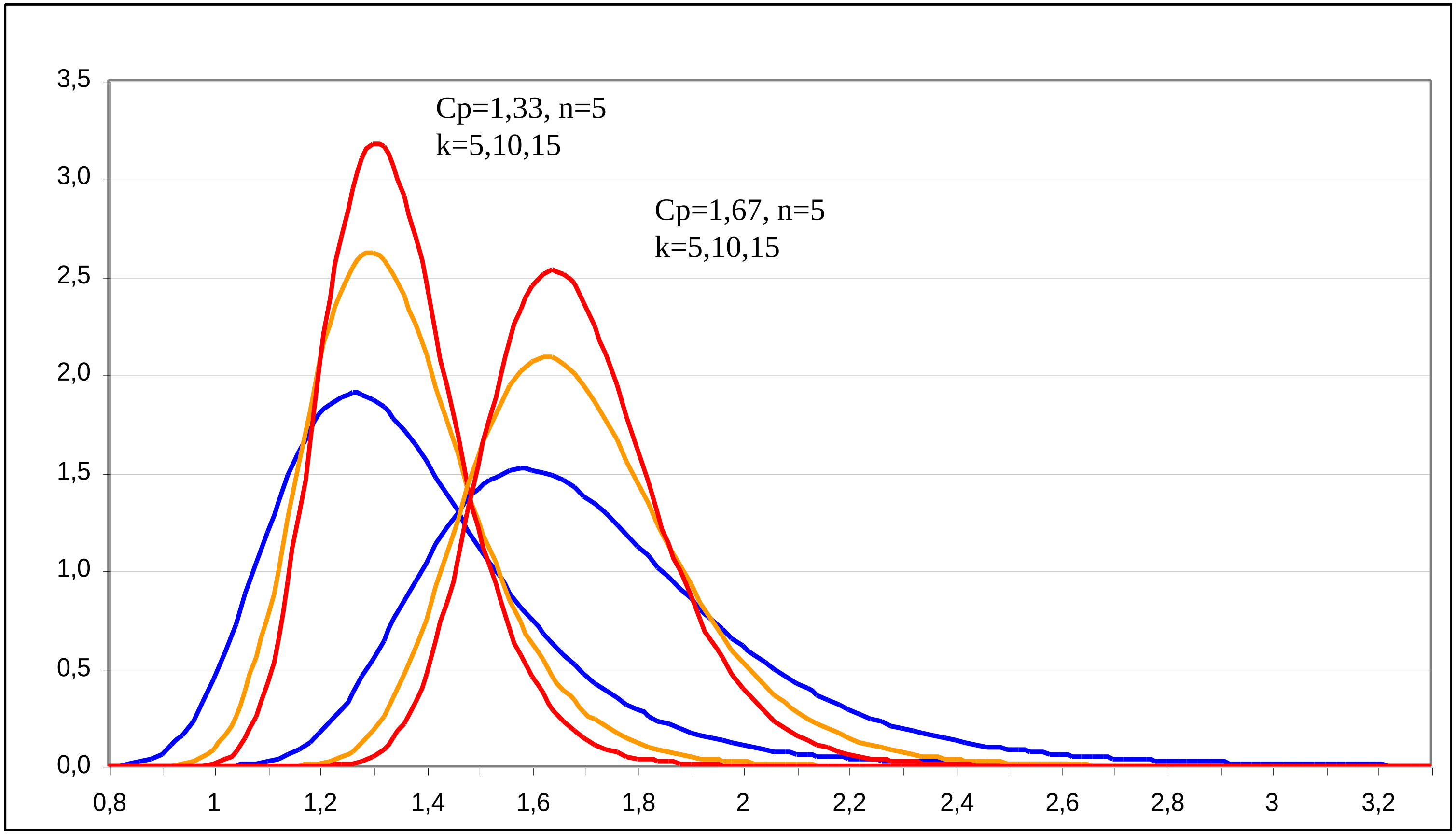
$$f_{\hat{C}_{pk}}(x) = \int_0^{\infty} \frac{1}{u} f_{\hat{C}_p}(u)\, f_{1-\hat{K}}\left(\frac{x}{u}\right) du \qquad \text{pro } x \le 0.$$

Bohužel tato formule nevede k explicitnímu vyjádření tvaru hustoty, a tudíž např. kvantily je nutno hledat numericky

Příklad: k = 10, n = 5, C_p = 1,33 odpovídající kvantily jsou:

α	$C_{pk}(\alpha)$
0,01	1,028
0,025	1,064
0,05	1,098
0,5	1,311
0,95	1,626
0,975	1,705
0,99	1,806

Funkce hustoty pro odhad $\hat{C}_{pk}$, proces je přesně centrován



Co obvykle zákazník vyžaduje?

Zákazník např. vyžaduje, aby $C_p = 1{,}33$, a tím automaticky předpokládá splnění nerovnosti

$$C_p < \hat{C}_p$$

s tím, že tato nerovnost mu již zaručuje očekávaných 64 ppm mimo specifikační meze.
Co ale ve skutečnosti tato nerovnost garantuje?

Co říká splnění předchozí nerovnosti ve skutečnosti?

Teprve konstrukce konfidenčního intervalu pro ukazatele C_p dává správnou odpověď:

$$\hat{C}_p \left(1 + \frac{\beta_n u_{\alpha/2}}{\alpha_n \sqrt{k}} \right) < C_p < \hat{C}_p \left(1 + \frac{\beta_n u_{1-\alpha/2}}{\alpha_n \sqrt{k}} \right)$$

Je-li např. $\hat{C}_p$ = 1,40 pro n=5, k=25 a riziko α=0,05 znamená to, že skutečná hodnota ukazatele C_p je v intervalu

(1,1566; 1,6834) s pravděpodobností 0,95.

Vlastnosti odhadů $\hat{C}_p$

Co lze říci o chování odhadů ukazatele C_p, je-li např.

$C_p = 1{,}33$?

Odpověď dává statistický pokryvný interval:

$$C_p \left(\frac{1}{1 + \frac{\beta_n u_{1-\alpha/2}}{\alpha_n \sqrt{k}}} \right) < \hat{C}_p < C_p \left(\frac{1}{1 + \frac{\beta_n u_{\alpha/2}}{\alpha_n \sqrt{k}}} \right)$$

Např. pro n=5, k=25 a α=0,05 to znamená, že odhady se budou s pravděpodobností 0,95 vyskytovat v intervalu (1,1219; 1,6329).

Co nerovnost požadovaná zákazníkem znamená pro vlastníka procesu?

Jak zajistit platnost nerovnosti v každém případě?

Odpověď: Je možno pouze zajistit její platnost s malým nenulovým rizikem, např. α=0,05.

Aby bylo např.1,33 < $\hat{C}_p$ splněno s pravděpodobností 1-α způsobilost procesu nesmí mít horší C_p nežli

$$1{,}33 < C_p \left(\frac{1}{1 + \dfrac{\beta_n u_{1-\alpha}}{\alpha_n \sqrt{k}}} \right)$$

Pak např. pro n=5, k=25 a α=0,05 proces musí vykazovat způsobilost minimálně na úrovni C_p=1,5944. To ale znamená daleko přísnější požadavek na přesnost procesu než požaduje ukazatel C_p = 1,33.

Požadavek na ukazatel C_p je požadavkem na úroveň inherentní variability,tj. na parametr σ

V předchozím případě s C_p=1,5944 to znamená, že parametr σ musí být maximálně

σ = (USL - LSL)/ 6*1,5944

Požadavek ne ukazatel C_p neříká nic o parametru polohy μ

Teprve μ je dáno ukazatelem C_{pk},

ale ne jednoznačně

Co znamená nerovnost $C_{pk} < \hat{C}_{pk}$?

Splnit tuto nerovnost s pravděpodobností 1 – 2α odhad parametru polohy μ

musí s touto pravděpodobností ležet uvnitř intervalu

$$(LSL + 3*C_{pk}*\sigma_0,\ USL - 3*C_{pk}*\sigma_0)$$

kde σ_0 maximal směrodatná odchylka daná požadavkem na ukazatele C_p

Toto bude zajištěno, když parametr polohy μ bude ležet uvnitř intervalu (μ_-, μ_+), kde

$$\mu_- = LSL + \sigma_0(3*Cpk – u_\alpha/(k*n)^{½}$$

$$\mu_+ = USL - \sigma_0(3*Cpk + u_{1-\alpha}/(k*n)^{½}$$

$u_{\alpha,}\ u_{1-\alpha}$ jsou kvantily rozdělení N(0,1)

Jak regulovat proces, aby plnil požadavky kladené na C_p a C_{pk}?

Pro ilustraci, regulační diagram (xbar,R) je použit

Regulační meze pro výběrové rozpětí R jsou klasické

Shewhartovy meze:

$$UCL(R) = D_2(n)*\sigma_o$$

$$LCL(R) = D_1(n)*\sigma_o$$

Ale regulační meze pro průměry musí být rozšířeny, protože

lze připustit jakýkoliv pohyb průměrů uvnitř intervalu (μ_-,μ_+):

$$LCL(xbar) = T-A(n)*\sigma_0 – (\mu_-+ \mu_+)/2$$

$$UCL(xbar) = T+A(n)*\sigma_0 + (\mu_-+ \mu_+)/2$$

kde T = (LSL+USL)/2

Poznámka: Klasické regulační meze pro průměry by mohly vyvolat příliš mnoho falešných poplachů

Odhady pro P_p a P_{pk}

$$\sigma_{tot} \approx s_{tot} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(x_i - \bar{x}_{tot}\right)^2}$$

$$\mu \approx \bar{x}_{tot} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

Totální výběrová směrodatná odchylka s_{tot} charakterizuje celkovou variabilitu v náhodném výběru z N kusů, tj. jak variabilitu v podskupinách, tak i mezi podskupinami, N = k*n. Data musí být popsatelná normálním rozdělením, jinak ukazatele výkonnosti ve výše uvedeném tvaru ztrácejí smysl.

Poznámka: *Zavedení ukazatelů P_p a P_{pk} je silně kritizováno ze strany matematických statistiků . Především z toho důvodu, že nic neříkají o chování znaku jakosti do budoucna*

Problémy s ukazateli P_p a P_{pk}

Původně tyto ukazatele byly zavedeny, aniž by se cokoliv předpokládalo o stabilitě a zvládnutelnosti procesu. Z jejich definice ale ihned vyplývá, že implicitně se data musejí dát popsat normálním rozdělení, jinak vlastně nic neříkají o stavu procesu a jejich aplikace je silně zavádějící. Navíc, pokud proces je pod kontrolou, pak odhady ukazatelů C_p, C_{pk} a P_p,P_{pk}.se příliš neliší.

Aplikace ukazatelů výkonnosti má své místo tam, kde na proces působí nějaká systematická příčina, o které víme, ale nelze ji z procesu odstranit. Proces ale musí vykazovat zvládnutelnost, např. v rámci rozšířených regulačních mezí, a data bez ohledu na sběr dat v podskupinách musí být popsatelná jako celek normálním rozdělením.

Reference:

S.Kotz, C.R.Lovelace: Process Capability Indices in Theory nad Practice, Arnold, London (1998)

J.Michálek: Capability indices estimates and their properties. Research report No.2016, UTIA, June 2001 (in Czech)

J.Michálek. Statistical tests of capability indices. Research report No.2154, UTIA, December 2005 (in Czech)