



Akademie věd České republiky
Ústav teorie informace a automatizace

Academy of Sciences of the Czech Republic
Institute of Information Theory and Automation

RESEARCH REPORT

JIŘÍ MICHÁLEK:

**Hustoty rozdělení pravděpodobnosti
pro odhady ukazatele C_{pk}**

číslo 2550

květen 2009

ÚTIA AV ČR, P. O. Box 18, 18208 Prague, Czech Republic
Telex: 122018 atom c, Fax: (+42) (2) 688 4903
E-mail: utia@utia.cas.cz

This report constitutes an unrefereed manuscript which is intended to be submitted for publication. Any opinions and conclusions expressed in this report are those of the author(s) and do not necessarily represent the views of the Institute.

Hustoty rozdělení pravděpodobnosti pro odhady ukazatele C_{pk}

Ukazatel C_{pk} , někdy též značený jako index C_{pk} , je spolu s ukazatelem C_p nejčastějším nástrojem pro hodnocení způsobilosti výrobního procesu vyjádřené zde chováním sledovaného znaku jakosti na výrobku, který je procesem produkován. Požadavek na ukazatele C_{pk} dává zákazník či konstruktér a tento požadavek se jednak vztahuje na parametr polohy μ a jednak na úroveň variability σ . V definici ukazatele C_{pk} je míněna tzv. inherentní variabilita, která představuje různorodost jednoho výrobku od druhého, které jsou odebírány ve formě tzv. logických podskupin. Tedy tato variabilita z pohledu času by se dala též charakterizovat jako okamžitá variabilita vyjadřující proměnlivost uvnitř podskupiny. Druhý parametr μ je parametr polohy a vyjadřuje umístění hodnot sledovaného znaku vůči specifikačnímu rozmezí. Parametr μ lze tedy chápat jako těžiště těchto hodnot, které jsou rozmístěny kolem něho právě s mírou variability či rozptýlenosti σ . Za předpokladu, že data lze považovat za normálně rozdělená, vzorec pro ukazatele C_{pk} je následující:

$$C_{pk} = \min(C_{pkU}, C_{pkL}), \quad \text{kde} \quad C_{pkU} = \frac{USL - \mu}{3\sigma}, \quad C_{pkL} = \frac{\mu - LSL}{3\sigma}.$$

USL a LSL jsou horní a dolní mezní hodnoty. Jestliže je dán požadavek na hodnotu ukazatele C_{pk} , např. $C_{pk} = 1,33$, je nutno si uvědomit, že tím nejsou jednoznačně určeny hodnoty parametrů μ a σ . To ale znamená, že zadané pouze hodnoty pro ukazatele C_{pk} nestačí, je nutno buď stanovit hodnotu parametru μ či naopak parametru σ . Hodnota parametru σ vyplývá např. z udání hodnoty ukazatele C_p , proto se v praxi nejčastěji udává dvojice C_p, C_{pk} pro úplnou charakterizaci požadavků na stav procesu. Dále je nutné si uvědomit, že pokud $C_p = C_{pk}$, je to požadavek na přesné centrování polohy znaku jakosti doprostřed specifikačního rozmezí. Dále při nerovnosti $C_{pk} < C_p$ to znamená, že parametr polohy μ nemusí ležet přesně uprostřed specifikačního rozmezí, ale z hodnoty ukazatele C_{pk} nevyplývá, zdali má proces sedět napravo či nalevo od středu specifikačního rozmezí. Pokud bychom požadovali, aby proces byl umístěn napravo od středu, pak jediné explicitně vyjádřením, že v tomto případě je $C_{pk} = C_{pkU}$. Pokud je tedy předepsána hodnota parametru σ , pak z hodnoty pro C_{pk} vyplývá existence dvou hodnot μ_- a μ_+ pro parametr polohy, totiž

$$C_{pkU} = \frac{USL - \mu_+}{3\sigma}, \quad C_{pkL} = \frac{\mu_- - LSL}{3\sigma};$$

když parametr polohy μ bude buď roven μ_- či μ_+ , pak C_{pk} bude roven předepsané hodnotě. Pokud parametr polohy μ bude mezi, tj.

$$\mu_- \leq \mu \leq \mu_+,$$

pak hodnota ukazatele C_{pk} nebude horší.

Z definice ukazatele C_{pk} vyplývá, že toto má smysl požadovat, když parametry μ a σ budou v čase neměnné, nebo v praxi téměř neměnné. Tomuto stavu odpovídá stabilita procesu v čase a je nutno toto zajistit, např. pomocí metod SPC, aby bylo možno odpovědně způsobilost procesu analyzovat. Pokud by tento stav tzv. statisticky zvládnutého procesu nebyl dosažen, hodnocení způsobilosti stojí na vodě a obvykle neodpovídá realitě.

Tedy na jednu stranu máme hodnotu pro C_{pk} , na druhé straně je reálný proces a je otázka, zdali tento proces je schopen požadavek splnit. Jak se to dozvíme? Jedině tak, že z procesu odebereme nějaké výrobky, na nich sledovaný znak přeměříme, ze získaných dat vypočteme vhodné odhady pro C_{pk} , a ty s danou požadovanou hodnotou zkonfrontujeme pomocí statistické analýzy. Abychom dovedli odhadovat úroveň inherentní variability, potřebujeme, aby data byla zorganizována v podskupinách. Pak přicházejí obvykle následující tři možnosti, jak parametr σ odhadovat:

1. $\hat{\sigma}_R = \frac{\bar{R}}{d_2(n)}$
2. $\hat{\sigma}_S = \frac{\bar{s}}{C_4(n)}$
3. $\hat{\sigma}_I = \left(\frac{1}{k(n-1)} \sum_{i=1}^k \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2 \right)^{1/2}$,

kde x_{ij} je j -té pozorování v i -té podskupině, $i = 1, 2, \dots, k$; $j = 1, 2, \dots, n$. \bar{R} je průměrné rozpětí, tedy

$$\bar{R} = \frac{1}{k} \sum_{i=1}^k R_i, \quad R_i = \max_j x_{ij} - \min_j x_{ij},$$

\bar{s} je průměrná směrodatná odchylka z podskupin, tedy

$$\begin{aligned} \bar{s} &= \frac{1}{k} \sum_{j=1}^k s_i, & s_i &= \left(\frac{1}{k(n-1)} \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2 \right)^{1/2}, \\ \bar{x}_i &= \frac{1}{n} \sum_{j=1}^n x_{ij}. \end{aligned}$$

Parametr μ se odhaduje nejčastěji pomocí tzv. celkového aritmetického průměru

$$\bar{\bar{x}} = \frac{1}{kn} \sum_{i=1}^k \sum_{j=1}^n x_{ij}.$$

Odhad ukazatele C_{pk} , označme ho \hat{C}_{pk} , pak má tvar:

$$\hat{C}_{pk} = \min \left(\frac{\text{USL} - \bar{\bar{x}}}{3\hat{\sigma}}, \frac{\bar{\bar{x}} - \text{LSL}}{3\hat{\sigma}} \right),$$

kde $\hat{\sigma}$ je jeden z výše uvažovaných odhadů směrodatné odchylky σ .

Jestliže data mají individuální charakter, tedy každá podskupina obsahuje pouze jedno pozorování, pak se problém podskupin obvykle řeší pomocí tzv. klouzavého rozpětí, což vlastně znamená vytváření umělých podskupin, např. ze dvou po sobě jdoucích pozorování. Pak lze taktéž uvažovat ony tři typy odhadů inherentní směrodatné odchylky, ale je nutno mít na paměti, že tentokrát se v součtu vyskytují stochasticky závislá data. Dalším problémem je, že tím do hry vstupuje výrazně čas ve formě časového odstupe mezi dvěma sousedními pozorováními. Budeme předpokládat, že data jsou organizována ve formě podskupin o počtu $n \geq 2$ kusů v podskupině, tento počet je neměnný a máme k dispozici k podskupin. Předpokládejme, že data jsou vzájemně stochasticky nezávislá a normálně rozdělená.

Abychom mohli studovat vlastnosti odhadů ukazatele C_{pk} , je nutno nejdříve odvodit distribuční funkce a hustoty rozdělení těchto odhadů. Je zřejmé, že tvar těchto funkcí bude záviset na použitém odhadu směrodatné odchylky, což v literatuře téměř není rozlišováno a v softwarech pro vyhodnocování způsobilosti rovněž.

Je několik přístupů, jak odvodit tvar rozdělení pro odhad \hat{C}_{pk} :

1. Postup založený na součinu dvou nezávislých náhodných veličin

Ukazatel C_{pk} se dá vyjádřit jako součin dvou veličin, z nichž jedna je funkcí pouze parametru polohy μ a druhá pouze parametru směrodatné odchylky σ . Platí totiž, že

$$C_{pk} = (1 - K) C_p,$$

kde $K = \frac{|\mu - T|}{\Delta}$, přičemž $T = \frac{USL + LSL}{2}$, $\Delta = \frac{USL - LSL}{2}$.

Odtud pro odhad \hat{C}_{pk} jasně platí, že

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

kde

$$\hat{K} = \frac{|\bar{x} - T|}{\Delta} \quad \text{a} \quad \hat{C}_p = \frac{USL - LSL}{6\hat{\sigma}} = \frac{\Delta}{3\hat{\sigma}}.$$

Protože vycházíme z předpokladu, že znak jakosti je normálně rozdělen, pak náhodné veličiny $1 - \hat{K}$ a \hat{C}_p musí být stochasticky nezávislé. Tento fakt je pak využit pro odvození hustoty rozdělení pravděpodobnosti pro odhad \hat{C}_{pk} . Hustota pro odhad \hat{C}_p závisí na volbě dohadu směrodatné odchylky a její odvození lze najít v [1].

Potřebujeme tedy odvodit hustotu pro veličinu $1 - \hat{K}$. Vyjděme z vyjádření pravděpodobnosti náhodného jevu $\{1 - \hat{K} < \lambda\}$, tedy pro λ reálné spočítejme

$$P\{1 - \hat{K} < \lambda\} = P\{1 - \lambda < \hat{K}\} = 1 - P\{\hat{K} \leq 1 - \lambda\} = 1$$

v případě, že $1 - \lambda < 0$, tedy pro $\lambda > 1$, neboť $\hat{K} \geq 0$. Pro $\lambda < 1$ máme pak

$$\begin{aligned} & P \{ \hat{K} \leq 1 - \lambda \} = P \{ |\bar{x} - T| \leq \Delta(1 - \lambda) \} = \\ & = P \{ T - \Delta(1 - \lambda) \leq \bar{x} \leq T + \Delta(1 - \lambda) \} = \\ & = P \left\{ (T - \mu - \Delta(1 - \lambda)) \frac{\sqrt{kn}}{\sigma} \leq \frac{\bar{x} - \mu}{\sigma} \sqrt{kn} \leq (T - \mu + \Delta(1 - \lambda)) \frac{\sqrt{kn}}{\sigma} \right\} = \\ & = \Phi \left(\frac{\sqrt{kn}}{\sigma} (T - \mu + \Delta(1 - \lambda)) \right) - \Phi \left(\frac{\sqrt{kn}}{\sigma} (T - \mu - \Delta(1 - \lambda)) \right), \end{aligned}$$

protože \bar{x} má rozdělení $N(\mu, \frac{\sigma^2}{kn})$. Po snadných úpravách máme, že

$$P\{1 - \hat{K} < \lambda\} = 1 - \Phi(3\sqrt{kn}(C_{pkU} - \lambda C_p)) + \Phi(3\sqrt{kn}(\lambda C_p - C_{pkL})),$$

kde $C_{pkL} = \frac{\mu - \text{LSL}}{3\sigma}$, $C_{pkU} = \frac{\text{USL} - \mu}{3\sigma}$.

Derivováním podle λ pak již snadno odvodíme hustotu pro veličinu $1 - \hat{K}$:

$$f_{1-\hat{K}}(\lambda) = 3C_p \sqrt{kn} \left(\varphi(3\sqrt{kn}(\lambda C_p - C_{pkL})) + \varphi(3\sqrt{kn}(C_{pkU} - \lambda C_p)) \right)$$

pro $\lambda \leq 1$, a pro $\lambda > 1$ je $f_{1-\hat{K}}(\lambda) = 0$. ($\Phi(\cdot)$ je distribuční funkce od $N(0, 1)$, $\varphi(\cdot)$ je její hustota.)

Nechť $f_{\hat{C}_p}(\cdot)$ je hustota rozdělení pravděpodobnosti odhadu \hat{C}_p , pak vzorec pro hustotu odhadu \hat{C}_{pk} založený na součinu dvou nezávislých náhodných veličin je:

$$\begin{aligned} f_{\hat{C}_{pk}}(x) &= \int_x^\infty \frac{1}{u} f_{\hat{C}_p}(u) f_{1-\hat{K}}\left(\frac{x}{u}\right) du \quad \text{pro } x > 0 \\ f_{\hat{C}_{pk}}(x) &= \int_0^\infty \frac{1}{u} f_{\hat{C}_p}(u) f_{1-\hat{K}}\left(\frac{x}{u}\right) du \quad \text{pro } x \leq 0. \end{aligned}$$

Volbou hustoty $f_{\hat{C}_p}(\cdot)$ podle různého odhadu parametru σ dostaneme tři případy pro hustotu $f_{\hat{C}_{pk}}(\cdot)$. Problém je ovšem v tom, že nelze z výše uvedeného vzorce hustotu $f_{\hat{C}_{pk}}(\cdot)$ explicitně vyjádřit, ale lze udělat numerickou studii a numericky vyjádřit odhady kvantilů, což se např. hodí pro konstrukci statisticky pokryvných intervalů pro odhad \hat{C}_{pk} .

2. Druhý přístup je založen na následující rovnosti:

$$\begin{aligned} & P \{ \hat{C}_{pk} > \lambda \} = P \{ \min(\hat{C}_{pkL}, \hat{C}_{pkU}) > \lambda \} = \\ & = P \left\{ \hat{C}_{pkL} > \lambda, \hat{C}_{pkU} > \lambda \right\} = P \left\{ \frac{\bar{x} - \text{LSL}}{3\hat{\sigma}} > \lambda, \frac{\text{USL} - \bar{x}}{3\hat{\sigma}} > \lambda \right\}. \end{aligned}$$

Hledáme tedy podmínky, za nichž jsou obě nerovnosti popisující dva náhodné jevy splněny. Podmínky splnění je nutno rozdělit na 4 možné případy:

a) $\lambda > 0, \hat{\sigma} > 0$

V tomto případě lze snadno ukázat, že podmínka splnění obou nerovností je

$$LSL + 3\lambda\sigma^2 < \bar{x} < USL - 3\lambda\hat{\sigma},$$

pokud bude $USL - LSL > 6\lambda\hat{\sigma}$, tedy $\hat{\sigma} < \frac{USL-LSL}{6\lambda}$.

b) $\lambda > 0, \hat{\sigma} < 0$

Tato situace vede k nerovnosti $USL - LSL < 6\lambda\hat{\sigma}$, což nemůže být splněno, neboť $USL - LSL > 0$, ale $6\lambda\hat{\sigma} < 0$.

c) $\lambda < 0, \hat{\sigma} > 0$

Tato situace vede k nerovnosti

$$LSL + 3\lambda\hat{\sigma} < \bar{x} < USL - 3\lambda\hat{\sigma},$$

kteřá platí pro každé $\hat{\sigma} > 0$, protože $\lambda < 0$.

d) $\lambda < 0, \hat{\sigma} < 0$

Zde dospějeme k nerovnosti

$$USL - 3\lambda\hat{\sigma} < \bar{x} < LSL + 3\lambda\hat{\sigma},$$

kteřá může být splněna jedině pro $\frac{USL-LSL}{6\lambda} > \hat{\sigma}$.

Pro $\lambda > 0$ shrnutím obou případů a), b) máme, že

$$\begin{aligned} P\{\hat{C}_{pk} \leq \lambda\} &= 1 - P\{\hat{C}_{pk} > \lambda\} = \\ &= 1 - P\{LSL + 3\lambda\hat{\sigma} < \bar{x} < USL - 3\lambda\hat{\sigma}\} = \\ &= 1 - P\left\{\sqrt{kn}\left(\frac{LSL + 3\lambda\hat{\sigma} - \mu}{\sigma}\right) < \frac{\bar{x} - \mu}{\sigma}\sqrt{kn} < \sqrt{kn}\left(\frac{USL - 3\lambda\hat{\sigma} - \mu}{\sigma}\right)\right\} = \\ &= 1 + \Phi\left(\sqrt{kn}\left(-3C_{pkL} + 3\lambda\frac{\hat{\sigma}}{\sigma}\right)\right) - \Phi\left(\left(-3C_{pkU} - 3\lambda\frac{\hat{\sigma}}{\sigma}\right)\sqrt{kn}\right). \end{aligned}$$

Nyní je nutné si uvědomit, že $\frac{\hat{\sigma}}{\sigma}$ je náhodná veličina, která má rozdělení pravděpodobnosti podle volby odhadu pro $\hat{\sigma}$. Lze snadno ukázat, že při volbě $\hat{\sigma} = \frac{\bar{R}}{d_2(n)}$ lze hustotu podílu $\frac{\hat{\sigma}}{\sigma}$ aproximovat velice dobře, viz [1], hustotou $N\left(1, \frac{\beta_n^2}{\alpha_n^2 k}\right)$, kde α_n, β_n jsou parametry charakterizující asymptotické chování výběrového rozpětí R z normálního rozdělení.

Označme $g(\cdot)$ hustotu rozdělení pravděpodobnosti veličiny $\frac{\hat{\sigma}}{\sigma}$, pak distribuční funkce odhadu \hat{C}_{pk} pro $\lambda > 0$ má tvar

$$\begin{aligned} F_{\hat{C}_{pk}}(\lambda) &= P\{\hat{C}_{pk} \leq \lambda\} = \\ &= 1 - \int_0^{C_p/\lambda} \left(-\Phi\left(3\sqrt{kn}(C_{pkU} - \lambda u)\right) + \Phi\left(3\sqrt{kn}(\lambda u - C_{pkL})\right)\right) g(u) du. \end{aligned}$$

Protože $2C_p = C_{pkL} + C_{pkU}$, pak lze distribuční funkci $F_{\hat{C}_{pk}}(\cdot)$ vyjádřit následovně

$$F_{\hat{C}_{pk}}(\lambda) = 1 - \int_0^{C_p/\lambda} \left(\Phi\left(3\sqrt{kn}(C_{pk} - \lambda u)\right) - \Phi\left(3\sqrt{kn}(\lambda u - C_p + C_{pk})\right) \right) g(u) du.$$

Vzorec pro odpovídající hustotu $f_{\hat{C}_{pk}}(\cdot)$ pro $\lambda > 0$ získáme derivováním integrálu podle parametru. Vyjděme z obecného vzorce

$$F(\lambda) = \int_0^{B(\lambda)} h(u, \lambda) du,$$

pak za předpokladů, které jsou splněny (viz [2]), derivace $F'(\lambda)$ podle parametru λ je

$$F'(\lambda) = \int_0^{B(\lambda)} \frac{dh(u, \lambda)}{d\lambda} du + h(B(\lambda), \lambda) B'(\lambda).$$

V našem případě $B(\lambda) = \frac{C_p}{\lambda}$, tedy $B'(\lambda) = -\frac{C_p}{\lambda^2}$,

$$h(u, \lambda) = \left\{ -\Phi\left(3\sqrt{kn}(C_{pk} - \lambda u)\right) + \Phi\left(3\sqrt{kn}(\lambda u - 2C_p + C_{pk})\right) \right\} g(u).$$

Odtud je snadno vidět, že pro $\lambda > 0$ je hustota pro \hat{C}_{pk} rovna

$$f_{\hat{C}_{pk}}(\lambda) = \int_0^{C_p/\lambda} 3\sqrt{kn} \cdot u \left\{ \varphi\left((\lambda u - 2C_p + C_{pk}) 3\sqrt{kn}\right) + \varphi\left(3\sqrt{kn}(C_{pk} - \lambda u)\right) \right\} g(u) du.$$

Pro $\lambda \leq 0$ vychází hustota rozdělení pravděpodobnosti pro \hat{C}_{pk} jako součet dvou složek, protože musíme oddělit $\hat{\sigma} > 0$ a $\hat{\sigma} < 0$. Pak

$$\begin{aligned} f_{\hat{C}_{pk}}(\lambda) &= \int_0^{\infty} 3\sqrt{kn} u \left\{ \varphi\left(3\sqrt{kn}(\lambda u - 3C_p + C_{pk})\right) + \varphi\left(3\sqrt{kn}(C_{pk} - \lambda u)\right) \right\} g(u) du \\ &+ \int_{-\infty}^{C_p/\lambda} 3\sqrt{kn} u \left\{ \varphi\left(3\sqrt{kn}(\lambda u - C_{pk})\right) + \varphi\left(3\sqrt{kn}(C_{pk} - 2C_p - \lambda u)\right) \right\} g(u) du. \end{aligned}$$

Pro praktické účely je možno na základě numerických studií položit $f_{\hat{C}_{pk}}(\lambda) = 0$ pro $\lambda \leq 0$.

3. Postup založený na podmíněné pravděpodobnosti

Vyjděme opět ze součinu $C_{pk} = (1-K)C_p$, což v řeči odhadů těchto ukazatelů znamená $\hat{C}_{pk} = (1 - \hat{K})\hat{C}_p$. Tedy

$$P\left\{\hat{C}_{pk} < \lambda\right\} = P\left\{(1 - \hat{K})\hat{C}_p < \lambda\right\} = \int_{-\infty}^{+\infty} P\left\{(1 - \hat{K})u < \lambda \mid \hat{C}_p = u\right\} g(u) du,$$

kde $g(\cdot)$ je hustota rozdělení pravděpodobnosti pro veličinu \hat{C}_p . Již víme, že lze uvažovat tři možnosti tvaru této hustoty podle volby odhadu parametru σ . Je nutno ve výše uvedeném vzorci rozlišit dva případy, a to $u > 0$, $u < 0$. Pro $u > 0$ se pak jedná o výraz

$$\int_0^{\infty} P\left\{(1 - \hat{K}) < \frac{\lambda}{u} \mid \hat{C}_p = u\right\} g(u) du,$$

a pro $u < 0$ se jedná o výraz

$$\int_{-\infty}^0 P \left\{ (1 - \hat{K}) > \frac{\lambda}{u} \mid \hat{C}_p = u \right\} g(u) du.$$

Celkem tedy pro $\lambda > 0$

$$\begin{aligned} F_{\hat{C}_{pk}}(\lambda) &= P \{ \hat{C}_{pk} < \lambda \} = \int_{\lambda}^{\infty} P \left\{ (1 - \hat{K}) < \frac{\lambda}{u} \right\} g(u) du + \\ &+ \int_{-\infty}^0 \left(1 - P \left\{ (1 - \hat{K}) < \frac{\lambda}{u} \right\} \right) g(u) du + \int_0^{\lambda} g(u) du, \end{aligned}$$

protože $P\{(1 - \hat{K}) < x\} = 1$ pro $x > 1$.

Potřebujeme tedy znát distribuční funkci veličiny $(1 - \hat{K})$, což již bylo odvozeno dříve:

$$P \{ (1 - \hat{K}) < x \} = 1 - \Phi \left(3\sqrt{kn}(C_{pkU} - xC_p) \right) + \Phi \left(3\sqrt{kn}(\lambda C_p - C_{pkL}) \right),$$

zde $x = \frac{\lambda}{u}$.

Derivováním podle proměnné λ dosáhneme implicitního vyjádření pro hustotu odhadu \hat{C}_{pk} , totiž pro $\lambda > 0$ (pro $\lambda < 0$ lze odvodit hustotu $f_{\hat{C}_{pk}}(\cdot)$ obdobně, pro praktické účely lze ji položit jako 0)

$$\begin{aligned} F_{\hat{C}_{pk}}(\lambda) &= \int_{\lambda}^{\infty} \frac{3C_p\sqrt{kn}}{u} \left(\left(\varphi \left(3\sqrt{kn}(C_{pkU} - \frac{\lambda}{u} C_p) \right) \right) + \right. \\ &\quad \left. + \varphi \left(3\sqrt{kn} \left(\frac{\lambda}{u} C_p - C_{pkL} \right) \right) \right) g(u) du + \\ &+ \int_{-\infty}^0 \frac{3C_p\sqrt{kn}}{u} \left(\left(\varphi \left(\frac{\lambda}{u} C_p - C_{pkL} \right) \right) - \varphi \left(2\sqrt{kn} \left(C_{pkU} - \frac{\lambda}{u} C_p \right) \right) \right) g(u) du. \end{aligned}$$

V případě odhadu \hat{C}_p založeném na “pooled” směrodatné odchylce část s integrálem pro $u < 0$ odpadá, protože hustota $g(\cdot)$ je odvozena od χ^2 -rozdělení.

Pro $\lambda > 1$ je hustota veličiny $(1 - \hat{K})$ rovna 0, tím $g(\cdot)$ ze vzorce vypadá. Samozřejmě by bylo možné roli veličin $(1 - \hat{K})$ a \hat{C}_p v podmínce otočit a dojít tak k vyjádření

$$\begin{aligned} P \{ \hat{C}_{pk} < \lambda \} &= \int_{-\infty}^{+\infty} P \left\{ \hat{C}_p < \frac{\lambda}{z} \mid (1 - \hat{K} = z) \right\} h(z) dz = \\ &= \int_{-\infty}^1 P \left\{ \hat{C}_p < \frac{\lambda}{z} \right\} h(z) dz + \int_1^{\infty} P \left\{ \hat{C}_p < \frac{\lambda}{z} \right\} h(z) dz, \end{aligned}$$

opět z toho důvodu, že $h(z) = 0$ pro $z > 1$, když $h(\cdot)$ bude hustota veličiny $1 - \hat{K}$.

V dalším bude provedena srovnávací numerická studie všech tří přístupů k vyjádření hustoty odhadu \hat{C}_{pk} a porovnání numerických odhadů příslušných kvantilů pro $\alpha = 0,01, 0,02, 0,05, 0,95, 0,98$ a $0,99$. Studie bude provedena pomocí softwaru MathCad.

Numerická studie v Dodatku se věnuje srovnání všech tří přístupů k aproximaci hustoty rozdělení pravděpodobnosti odhadu \hat{C}_{pk} .

Pro určení hranic konfidenčních intervalů a statistiky pokryvných intervalů je nutno odvodit rozdělení veličiny \hat{C}_{pk}/C_{pk} , tedy

$$\frac{\hat{C}_{pk}}{C_{pk}} = \frac{(1 - \hat{K})}{1 - K} \cdot \frac{\hat{C}_p}{C_p}.$$

Hodnoty ukazatelů C_p a C_{pk} musí být předem stanoveny. Tedy distribuční funkce pro podíl \hat{C}_{pk}/C_{pk} má tvar:

$$F(\lambda) = P \left\{ \hat{C}_{pk}/C_{pk} < \lambda \right\} = \int_{-\infty}^{+\infty} P \left\{ \frac{\hat{C}_p}{C_p} \cdot u < \lambda \mid \frac{1 - \hat{K}}{1 - K} = u \right\} q(u) du,$$

kde $q(\cdot)$ je hustota veličiny $(1 - \hat{K})/(1 - K)$. Tuto hustotu potřebujeme odvodit, proto

1. $\frac{1 - \hat{K}}{1 - K} > 0 \equiv 1 - \hat{K} > 0$, protože lze uvažovat, že $|\mu - T| < \Delta$,
tedy $\hat{K} < 1 \equiv \{\text{LSL} < \bar{x} < \text{USL}\}$.
2. $\frac{1 - \hat{K}}{1 - K} < 0 \equiv \{\bar{x} > \text{USL} \text{ či } \bar{x} < \text{LSL}\}$.

Z toho plyne, že

$$P \left\{ \frac{\hat{C}_{pk}}{C_{pk}} < \lambda \right\} = P \left\{ \frac{\hat{C}_p}{C_p} < \lambda \cdot \frac{1 - K}{1 - \hat{K}} \mid \bar{x} \in (\text{LSL}, \text{USL}) \right\} + \\ + P \left\{ \frac{\hat{C}_p}{C_p} > \lambda \cdot \frac{1 - K}{1 - \hat{K}} \mid \bar{x} > \text{USL} \text{ či } \bar{x} < \text{LSL} \right\}.$$

Je-li splněna podmínka $\bar{x} \in (\text{LSL}, \text{USL})$, pak to znamená, že

$$\frac{1 - \hat{K}}{1 - K} < 1,$$

pro podmínku, že $\bar{x} > \text{USL}$ či $\bar{x} < \text{LSL}$ to pak značí, že

$$\frac{1 - \hat{K}}{1 - K} > 1.$$

Tím se distribuční funkce pro podíl \hat{C}_{pk}/C_{pk} rozpadá na dvě části.

Nejdříve ale musíme odvodit hustotu rozdělení pravděpodobnosti pro veličinu $(1 - \hat{K})/(1 - K)$. Máme

$$Q(u) = P \left\{ (1 - \hat{K})/(1 - K) < u \right\} = P \left\{ 1 - \hat{K} < \frac{C_{pk}}{C_p} u \right\},$$

protože $1 - K = C_{pk}/C_p$. Odtud již snadno

$$Q(u) = 1 - P \left\{ |\bar{x} - T| < \Delta - u\Delta \frac{C_{pk}}{C_p} \right\},$$

protože $\hat{K} = \frac{|\bar{x} - T|}{\Delta}$. Jelikož aritmetický průměr \bar{x} má rozdělení $N(\mu, \frac{\sigma^2}{kn})$, pak

$$\begin{aligned} Q(u) &= 1 - P \left\{ \left(\frac{LSL - \mu}{\sigma} + \frac{u\Delta}{\sigma} \cdot \frac{C_{pk}}{C_p} \right) \sqrt{kn} < \frac{\bar{x} - \mu}{\sigma} \sqrt{kn} < \right. \\ &< \left. \left(\frac{USL - \mu}{\sigma} - \frac{u\Delta}{\sigma} \cdot \frac{C_{pk}}{C_p} \right) \sqrt{kn} \right\} = \\ &= 1 - \Phi \left(3\sqrt{kn}(C_{pkU} - C_{pk}u) \right) + \Phi \left(3\sqrt{kn}(C_{pk}u - C_{pkL}) \right), \end{aligned}$$

kde $\Phi(\cdot)$ je distribuce $N(0, 1)$.

Protože při odvozování je ve výrazu absolutní hodnota $|\bar{x} - T|$, pak musí být $\Delta - u\Delta \frac{C_{pk}}{C_p} > 0$, což vede k požadavku, aby

$$u < \frac{C_p}{C_{pk}}.$$

Pro $u \geq \frac{C_p}{C_{pk}}$ je pak automaticky $Q(u) = 1$.

Souhrnně tedy, bude-li $C_{pk} = C_{pkU}$, pak distribuce $Q(u)$ pro $u < \frac{C_p}{C_{pk}}$ má tvar

$$Q(u) = 1 - \Phi \left(3\sqrt{kn}(C_{pk} - u C_{pk}) \right) + \Phi \left(3\sqrt{kn}(u C_{pk} - 2C_p + C_{pk}) \right).$$

V případě, že $C_{pk} = C_{pkL}$, pak distribuce $Q(u)$ pro $u < \frac{C_p}{C_{pk}}$ má tvar

$$Q(u) = 1 - \Phi \left(3\sqrt{kn}(-C_{pk} + 2C_p - C_{pk}u) \right) + \Phi \left(3\sqrt{kn}(C_{pk}u - C_{pk}) \right).$$

Díky vlastnosti distribuce $\Phi(\cdot)$, totiž, že $\Phi(-z) = 1 - \Phi(z)$, je snadno vidět, že obě vyjádření pro hodnotu distribuce $Q(\cdot)$ jsou totožné, čili nezávisí na tom, zdali $C_{pk} = C_{pkU}$ či $C_{pk} = C_{pkL}$.

Odtud již snadno odvodíme výraz pro hustotu $q(\cdot)$:

$$\begin{aligned} q(u) &= 0 \quad \text{pro} \quad u \geq \frac{C_p}{C_{pk}}, \\ q(u) &= 3C_{pk}\sqrt{kn} \left(\phi \left(3\sqrt{kn}(C_{pk} - u C_{pk}) \right) + \phi \left(3\sqrt{kn}(C_{pk} + u C_{pk} - 2C_p) \right) \right) \quad \text{jinak.} \end{aligned}$$

Tím se dostáváme ke konečnému vyjádření distribuce $F(\cdot)$ pro podíl $\frac{C_{pk}}{C_p}$.

$\lambda > 0$:

$$\begin{aligned} F(\lambda) &= P \left\{ \frac{\hat{C}_{pk}}{C_{pk}} < \lambda \right\} = P \left\{ \frac{\hat{C}_p}{C_p} < \lambda \cdot \frac{1-K}{1-\hat{K}} \mid \hat{K} < 1 \right\} = \\ &= \int_0^\infty P \left\{ \frac{\hat{C}_p}{C_p} < \frac{\lambda}{u} \right\} q(u) du = \\ &= \int_0^{C_p/C_{pk}} P \left\{ \frac{\hat{C}_p}{C_p} < \frac{\lambda}{u} \right\} q(u) du. \end{aligned}$$

Pak tedy pro $\lambda > 0$

$$F(\lambda) = \int_0^{C_p/C_{pk}} P \left\{ \frac{\hat{C}_p}{C_p} < \frac{\lambda}{u} \right\} q(u) du,$$

protože $\frac{\hat{c}_p}{C_p} > 0$ s pravděpodobností 1.

Pro ilustraci uvažujme případ s odhadem parametru σ založeném na \bar{R} , tedy

$$\hat{\sigma} = \frac{\bar{R}}{d_2(u)}.$$

Pak distribuční funkce pro podíl \hat{C}_p/C_p může být velmi dobře aproximována funkcí

$$P \left\{ \frac{\hat{C}_p}{C_0} < \mu \right\} \doteq 1 - \Phi \left(\frac{\alpha_n \sqrt{k}}{\beta_n} \left(\frac{\hat{1}}{\mu} - 1 \right) \right), \quad \text{viz [1],}$$

pak tedy

$$P \left\{ \frac{\hat{C}_p}{C_p} < \frac{\lambda}{u} \right\} \doteq 1 - \Phi \left(\frac{\alpha_n \sqrt{k}}{\beta_n} \left(\frac{u}{\lambda} - 1 \right) \right)$$

a

$$P \left\{ \frac{\hat{C}_p}{C_p} \geq \frac{\lambda}{u} \right\} \doteq \Phi \left(\frac{\alpha_n \sqrt{k}}{\beta_n} \left(\frac{u}{\lambda} - 1 \right) \right).$$

Souhrnně pak tedy máme vyjádření pro distribuci $F(\cdot)$:

$$F(\lambda) = \int_0^{C_p/C_{pk}} \left(1 - \Phi \left(\frac{\alpha_n \sqrt{k}}{\beta_n} \left(\frac{u}{\lambda} - 1 \right) \right) \right) q(u) du \quad \text{pro } \lambda > 0$$

a

$$F(\lambda) = \int_{-\infty}^0 \Phi \left(\frac{\alpha_n \sqrt{k}}{\beta_n} \left(\frac{\hat{u}}{\lambda} - 1 \right) \right) q(u) du \quad \text{pro } \lambda < 0.$$

Odtud již snadno derivováním získáme hustotu rozdělení pravděpodobnosti pro veličinu \hat{C}_{pk}/C_{pk} :

pro $\lambda > 0$:

$$f(\lambda) = \int_0^{C_p/C_{pk}} \frac{\alpha_n \sqrt{k}}{\beta_n} \varphi \left(\frac{\alpha_n \sqrt{k}}{\beta_n} \left(\frac{u}{\lambda} - 1 \right) \right) \cdot \frac{u}{\lambda^2} q(u) du$$

a pro $\lambda < 0$

$$f(\lambda) = \int_{-\infty}^0 \frac{\alpha_n \sqrt{k}}{\beta_n} \varphi \left(\frac{\alpha_n \sqrt{k}}{\beta_n} \left(\frac{u}{\lambda} - 1 \right) \right) \cdot \frac{|u|}{\lambda^2} q(u) du$$

(pro praktické účely lze položit definitornicky $f(\lambda) = 0$ pro $\lambda < 0$, protože $f(\lambda)$ pro $\lambda < 0$ je zanedbatelná).

V dodatku jsou uvedeny tabulky kvantilů pro rozdělení pravděpodobnosti podílu \hat{C}_{pk}/C_{pk} pro velikost podskupiny $n = 2, 5$ a 10 a počet podskupin $k = 10, 20, 25$ a 50 . Kvantily lze využít následovně: vybereme konfidenční úroveň $1 - 2\alpha$, kde např. $\alpha = 0, 01$ či $0, 025$. Najdeme odpovídající kvantily q_α a $q_{1-\alpha}$, pak

$$P \left\{ q_\alpha \leq \frac{\hat{C}_{pk}}{C_{pk}} \leq q_{1-\alpha} \right\} = 1 - 2\alpha.$$

Odtud snadno odvodíme statistický pokryvný interval pro \hat{C}_{pk} , totiž

$$C_{pk} q_\alpha \leq \hat{C}_{pk} \leq C_{pk} q_{1-\alpha},$$

či konfidenční interval pro C_{pk} :

$$\frac{\hat{C}_{pk}}{q_{1-\alpha}} \leq C_{pk} \leq \frac{\hat{C}_{pk}}{q_\alpha}.$$

Hustoty rozdělení pravděpodobnosti pro ukazatele C_{pkL} a C_{pkU} .

Tyto ukazatelé představují základ pro definování ukazatele C_{pk} a používají se především při zadání pouze jediné specifikační meze. Je-li to USL, pak se použije C_{pkU} , při zadání LSL se pracuje s C_{pkL} .

$$C_{pkL} = \frac{\mu - \text{LSL}}{3\sigma}, \quad C_{pkU} = \frac{\text{USL} - \mu}{3\sigma},$$

kde μ je parametr polohy a σ je směrodatná odchylka inherentní variability sledovaného znaku jakosti. Definice obou ukazatelů v tomto pojetí předpokládá normální rozdělení $N(\mu, \sigma^2)$ znaku jakosti.

Reálný stav procesu pak vyjadřují jejich odhady \hat{C}_{pkL} a \hat{C}_{pkU} , kde parametr μ je odhadován aritmetickým průměrem \bar{x} ze všech dat a parametr σ je odhadován pomocí některého odhadu směrodatné odchylky σ , tedy

$$\hat{\sigma}_R = \frac{\bar{R}}{d_2(u)}, \quad \hat{\sigma}_S = \frac{\bar{s}}{C_4(u)}, \quad \hat{\sigma}_I = \left(\frac{1}{k(n-1)} \sum_{i=1}^k \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2 \right)^{1/2}.$$

Jde tedy o to odvodit tvar rozdělení těchto odhadů ukazatelů C_{pkL} a C_{pkU} podle použitého odhadu úrovně inherentní variability. Je předpokládáno, že data jsou organizována do

podskupin o rozsahu $n \geq 2$ a jsou mezi sebou navzájem nezávislá. Z předpokladu o normalitě dat vyplývá, že veličiny \bar{x} a $\hat{\sigma}$ jsou vzájemně nezávislé. Hustoty odhadů \hat{C}_{pkL} a \hat{C}_{pkU} odvodíme na základě vzorce pro hustotu podílu dvou nezávislých náhodných veličin

$$f(v) = \int_{-\infty}^{+\infty} |z| f_X(vz) f_Y(z) dz,$$

kde $f_X(\cdot)$ je hustota čitatele, $f_Y(\cdot)$ je hustota jmenovatele v podílu X/Y . Zde máme $X = \frac{\bar{x} - \text{LSL}}{3}$ a $Y = \hat{\sigma}_R$ či $\hat{\sigma}_S$ či $\hat{\sigma}_I$. Nejdříve probereme případ s $Y = \hat{\sigma}_R$. Je zřejmé, že na základě rozdělení pro \bar{x} :

$$\frac{\bar{x} - \text{LSL}}{3} \sim N\left(\frac{\mu - \text{LSL}}{3}, \frac{\sigma^2}{9kn}\right),$$

odhad $\hat{\sigma}_R$ má přibližné rozdělení $N\left(\sigma, \frac{\beta_n^2 \sigma^2}{k\alpha_n^2}\right)$.

Pak dosazením do obecného vzorce pro hustotu podílu dojdeme k integrálu

$$f(v) = \int_{-\infty}^{+\infty} |z| \frac{3k\alpha_n\sqrt{n}}{2\pi\beta_n\sigma^2} \exp\left\{-\frac{9kn}{2}\left(\frac{vz - \mu + \text{LSL}}{\sigma}\right)^2 - \frac{\alpha_n^2 k}{2\beta_n^2}\left(\frac{z - \sigma}{\sigma}\right)^2\right\} dz.$$

Integrál rozdělíme na dvě části $(-\infty, 0) \cup (0, +\infty)$ a po substituci $t = \frac{z}{\sigma}$ získáváme

$$f(v) = \int_{-\infty}^{+\infty} |t| \frac{3k\alpha_n\sqrt{n}}{2\pi\beta_n} \exp\left\{-\frac{9kn}{2}(vt - 3C)^2 - \frac{\alpha_n^2 k}{2\beta_n^2}(t - 1)^2\right\} dt,$$

kde hodnota $C_{pkL} = C$. Pro další výpočet integrálu rozdělíme opět na dvě části:

$$\begin{aligned} f_+(v) &= \frac{3k\alpha_n\sqrt{n}}{2\pi\beta_n} \int_0^{\infty} t \exp\left\{-\frac{3kn(vt - 3C)^2}{2} - \frac{\alpha_n^2 k}{2\beta_n^2}(t - 1)^2\right\} dt \\ f_-(v) &= \frac{3k\alpha_n\sqrt{n}}{2\pi\beta_n} \int_{-\infty}^0 t \exp\left\{-\frac{3kn(vt - 3C)^2}{2} - \frac{\alpha_n^2 k}{2\beta_n^2}(t - 1)^2\right\} dt. \end{aligned}$$

Pak $f(v) = f_+(v) - f_-(v)$.

V dalším použijeme následující vzorec:

$$\int_0^{\infty} t e^{-at^2 - 2bt} dt = \frac{1}{2a} - \frac{b}{2a} \sqrt{\frac{\pi}{a}} \exp\left\{\frac{b^2}{a}\right\} \left[1 - \text{erf}\left(\frac{b}{\sqrt{a}}\right)\right],$$

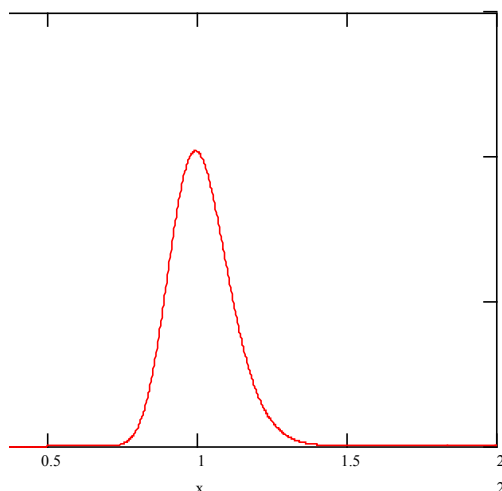
kde $a > 0$. V našem případě máme

$$a = \frac{k}{2} \left(2nv^2 + \frac{\alpha_n^2}{\beta_n^2}\right), \quad b = k \left(2Cnv^2 + \frac{\alpha_n^2}{\beta_n^2}\right),$$

označme ještě $d = \frac{k}{2} \left(9C^2n + \frac{\alpha_n^2}{\beta_n^2}\right)$. Pak hustota $f(\cdot)$ má explicitní vyjádření:

$$f(v) = \frac{3\alpha_n k \sqrt{n}}{2\pi\beta_n} e^{-d} \left[\frac{1}{a} + \frac{b\sqrt{\pi}}{2n^{3/2}} e^{b^2/4a} \left(2\Phi\left(\frac{b}{2\sqrt{a}}\right) - 1\right) \right].$$

Pro ilustraci uveďme grafické vyjádření hustoty pro C_{pkL} při $k = 30$, $n = 3$ a $C_{pkL} = 1$, $\alpha_n = 1,693$, $\beta_n = 0,888$.



Ihned je vidět, že rozdělení odhadu ukazatele C_{pkU} , tedy veličina \hat{C}_{pkU} , má tentýž tvar hustoty s tím, že $C = C_{pkU}$, protože rozdělení čitatele $\frac{USL - \bar{x}}{3}$ je v tom případě

$$N\left(\frac{USL - \mu}{3}, \frac{\sigma^2}{9kn}\right).$$

Pokud by nás zajímala hustota podílu \hat{C}_{pkL}/C_{pkL} , pak použijeme obecný vzorec pro hustotu podílu náhodné veličiny a nenulového čísla, a získáme ($g(\cdot)$ je hustota pro \hat{C}_{pkL}/C_{pkL}) $g(v) = C_{pkL} f(C_{pkL}v)$. Tento tvar hustoty se hodí pro numerické řešení nalezení konfidenčních mezí pro ukazatele C_{pkL} či odvození statistického pokrývného intervalu pro hodnoty odhadu \hat{C}_{pkL} při zadané hodnotě ukazatele C_{pkL} .

Jestliže bychom použili odhad $\hat{\sigma}_S$, pak se tvar výsledné hustoty pro odhad \hat{C}_{pkL} nezmění, jenom místo α_n, β_n budou konstanty a_{n-1} a b_{n-1} spojené s rozdělením výběrové směrodatné odchylky.

Jiná situace bude v případě použití tzv. pooled standard deviation $\hat{\sigma}_I$. Její rozdělení je odvoditelné od rozdělení χ^2 o $k(n-1)$ stupních volnosti a má tvar

$$r(x) = 2k(n-1) \frac{x}{\sigma^2} f_{\chi^2}\left(\frac{k(n-1)x^2}{\sigma^2}, k(n-1)\right),$$

kde $f_{\chi^2}(\cdot, k(n-1))$ je hustota χ^2 -rozdělení o $k(n-1)$ stupních volnosti. Podrobněji

$$r(x) = \frac{2k(n-1) \left(\frac{k(n-1)x^2}{\sigma^2}\right)^{\frac{k(n-1)}{2}-1}}{2^{\frac{k(n-1)}{2}} \Gamma\left(\frac{k(n-1)}{2}\right)} \exp\left\{-\frac{k(n-1)}{2\sigma^2} x^2\right\} \frac{x}{\sigma^2} \text{ pro } x \geq 0.$$

Na základě obecného vzorce pro hustotu podílu dvou nezávislých náhodných veličin, zde $\frac{\bar{x}-LSL}{3}$ a $\hat{\sigma}_I$, získáme vzorec pro hustotu odhadu \hat{C}_{pkL} při použití odhadu směrodatné odchylky založeném na “pooled deviation”.

$$f(v) = \frac{3\sqrt{kn}}{\sqrt{2\pi}} \frac{[k(n-1)]^{\frac{k(n-1)}{2}}}{2^{\frac{k(n-1)}{2}-1} \Gamma\left(\frac{k(n-1)}{2}\right)} \int_0^\infty t^{k(n-1)} \times \\ \times \exp\left\{-\frac{9kn}{2}(vt - C_{pkL})^2\right\} \exp\left\{-\frac{k(n-1)}{2}t^2\right\} dt.$$

Pro zjednodušení zavedme

$$\beta(v) = \frac{1}{2} (9knv^2 + k(n-1)), \quad \gamma(v) = -9kn C_{pkL}v,$$

pak lze hustotu $f(\cdot)$ přepsat do tvaru

$$f(v) = \frac{3\sqrt{kn}}{\sqrt{2\pi}} \cdot \frac{[k(n-1)]^{\frac{k(n-1)}{2}}}{2^{\frac{k(n-1)}{2}-1} \Gamma\left(\frac{k(n-1)}{2}\right)} \exp\left\{-\frac{9kn C_{pkL}^2}{2}\right\} \times \\ \times \int_0^\infty t^{k(n-1)} \exp\{-\beta(v)t^2 - \gamma(v)t\} dt = \\ = \frac{3\sqrt{kn}}{\sqrt{2\pi}} \frac{[k(n-1)]^{\frac{k(n-1)}{2}}}{2^{\frac{k(n-1)}{2}-1} \Gamma\left(\frac{k(n-1)}{2}\right)} \exp\left\{-\frac{9kn C_{pkL}^2}{2}\right\} \times \\ \times [2\beta(v)]^{-\frac{k(n-1)+1}{2}} \Gamma(k(n-1)+1) \exp\left\{\frac{\gamma^2(v)}{8\beta(v)}\right\} D_{-k(n-1)-1}\left(\frac{\gamma(v)}{\sqrt{2\beta(v)}}\right),$$

kde funkce $D_{-r}(\cdot)$ patří do rodiny tzv. funkcí parabolického válce a pro $r > 0$ je tato funkce definována jako:

$$D_{-r}(z) = \frac{e^{-\frac{z^2}{4}}}{\Gamma(r)} \int_0^\infty e^{-zx - \frac{x^2}{2}} x^{r-1} dx.$$

Hustotu $f(\cdot)$ lze napsat taky v poněkud jednodušším tvaru, a to

$$f(v) = \frac{3\sqrt{kn}[k(n-1)]^{\frac{k(n-1)}{2}} 2^{-\frac{k(n-1)+1}{2}} \beta(v)^{-\frac{k(n-1)+1}{2}}}{\sqrt{2\pi} 2^{\frac{k(n-1)}{2}-1} \Gamma\left(\frac{k(n-1)}{2}\right)} \times \\ \times \exp\left\{-\frac{9kn C_{pkL}^2}{2}\right\} \int_0^\infty e^{-\frac{x^2}{2} - \frac{\gamma(v)x}{2\beta(v)}} dx.$$

Aproximace hustoty rozdělení pravděpodobnosti odhadu \hat{C}_{pk} .

Z předchozího odvozování tvaru hustoty pro odhad \hat{C}_{pk} je vidět, že se velice těžko pracujeme explicitního vyjádření tvaru hustoty. Kdybychom takové explicitní vyjádření

ve tvaru vhodného vzorce měli, tak získání požadovaných kvantilů pro konstrukci statistického pokryvného intervalu by bylo možno snadno získat. V literatuře (např. [2]) lze najít odhady kvantilů či odkazy na další literaturu, kde jsou vhodné odhady kvantilů navrženy a studovány, ale aproximovat celou hustotu jinou vhodnou hustotou zatím studováno nebylo.

Na základě numerické studie lze předpokládat, že relativně dobrou aproximací hustoty odhadu \hat{C}_{pk} se jeví hustota odhadu \hat{C}_{pkL} , resp. \hat{C}_{pkU} , která byla odvozena v přechozím. Jestliže se bude lišit hodnota C_p od C_{pk} , tedy $C_p > C_{pk}$ více, tím spíše se chování odhadu \hat{C}_{pk} bude blížit chování odhadu \hat{C}_{pk} či \hat{C}_{pkU} podle toho, na které straně od centra tolerančního rozmezí bude umístěn parametr polohy. Horší aproximaci lze tedy očekávat v těch případech, kdy $C_p = C_{pk}$ (požadavek na přesné centrování) či C_p bude málo odlišné od C_{pk} . Necht' jsou dány požadované hodnoty ukazatelů C_p a C_{pk} . Dále máme k dispozici data organizovaná do podskupin rozsahu $n \geq 2$ a počet podskupin je k . Odhad inherentní variability (parametr σ) bude odhadován pomocí \bar{R} , resp. pomocí \bar{s} .

Výše byla odvozena hustota pro odhad \hat{C}_{pkL} , resp. \hat{C}_{pkU} na základě vzorce pro podíl dvou nezávislých náhodných veličin. Pro odhady směrodatné odchylky σ , a to $\bar{R}/d_2(u)$ a $\bar{s}/C_4(u)$, lze odvodit explicitní tvar hustoty pro odhady \hat{C}_{pkL} , resp. \hat{C}_{pkU} , a to ve formě

$$f_{\hat{C}_{pkL}}(x) = \int_{-\infty}^{+\infty} |y| f_Y(xy) f_Y(y) dy,$$

kde $f_X(\cdot)$ je hustota rozdělení $N\left(C_{pkL}, \frac{1}{9kn}\right)$, $f_Y(\cdot)$ je hustota rozdělení $N\left(1, \frac{\beta_n^2}{\alpha_n^2 k}\right)$; po dosazení těchto hustot do vzorce a potřebných výpočtech dojdeme k vyjádření ve tvaru:

$$f_{\hat{C}_{pkL}}(x) = \frac{3k\alpha_n\sqrt{n}}{2\pi\beta_n} \exp\{-d\} \left[\frac{1}{A(x)} + \frac{B(x)}{A(x)} \sqrt{\frac{\pi}{A(x)}} \exp\left\{\frac{B^2(x)}{A(x)}\right\} \right] \times \\ \times \left[\Phi\left(\frac{B(x)\sqrt{2}}{\sqrt{A(x)}}\right) - \Phi\left(-\frac{B(x)\sqrt{2}}{\sqrt{A(x)}}\right) \right]$$

kde

$$d = \frac{1}{2} \left(9kn C_{pkL}^2 + \frac{\alpha_n^2 k}{\beta_n^2} \right) \\ A(x) = \frac{1}{2} \left(9kn x^2 + \frac{\alpha_n^2 k}{\beta_n^2} \right) \\ B(x) = \frac{1}{2} \left(9kn C_{pkL} x + \frac{\alpha_n^2 k}{\beta_n^2} \right)$$

a $\Phi(\cdot)$ je distribuční funkce pro $N(0, 1)$.

Toto je aproximace pro případ odhadu parametru σ pomocí $\bar{R}/d_2(u)$. V případě odhadu σ pomocí $\bar{s}/C_4(n)$ se ve vzorci místo α_n , β_n použijí parametry a_{n-1} , b_{n-1} již dříve zmíněné.

Vhodnost aproximace ověříme na numerické studii provedené pomocí software Math-Cad. Studie je uvedena v Dodatku.

Literatura

- [1] J. Michálek: Odhady koeficientů způsobilosti a jejich vlastnosti. Výzkumná zpráva 2016, červen 2001, ÚTIA AV ČR.
- [2] S. Kotz, C. R. Lovelace: Process Capability Indices in Theory and Practice. Arnold, London 1998.

DODATEK

obsahující numerické studie:

- 1. porovnání tří přístupů k výpočtu hustoty odhadu indexu C_{pk}**
- 2. vliv volby odhadu pro směrodatnou odchylku na tvar hustoty odhadu indexu C_{pk}**
- 3. porovnání hodnot kvantilů rozdělení odhadu indexu C_{pk} s kvantily rozdělení odhadu indexu C_{pkL}**
- 4. tabulka hodnot důležitých kvantilů rozdělení podílu odhadu indexu C_{pk} a hodnoty indexu C_{pk}**

Numerická studie porovnávající 3 uvažované přístupy k výpočtu hustoty rozdělení pravděpodobnosti odhadu ukazatele Cpk

Studie provádí srovnání tří přístupů k vyjádření hustoty odhadu ukazatele Cpk, které jsou rozebírány v práci. Srovnání je provedeno na základě výpočtu vybraných kvantilů pro rozsah podskupiny $n=3,5$ a 10 a pro počet podskupin $k=10,15,20,25,50$ a 100. Jednotlivé přístupy jsou označeny Součin, Min a Podm.pst. podle základního principu odvození tvaru hustoty. Srovnání bylo provedeno pro hustotu založenou na odhadu inherentní směrodatné odchylky pomocí výběrového rozpětí R.

Součin $C_p=1,67$ $C_{pk}=1,33$

n=3	k=10	15	20	25	50	100
0,01	0,940	0,996	1,031	1,057	1,126	1,179
0,02	0,975	1,027	1,060	1,084	1,147	1,195
0,05	1,031	1,077	1,105	1,126	1,179	1,220
0,50	1,330	1,330	1,330	1,330	1,330	1,330
0,95	1,843	1,722	1,658	1,616	1,521	1,460
0,98	2,035	1,856	1,764	1,706	1,577	1,496
0,99	2,185	1,958	1,843	1,772	1,616	1,521

Min $C_p=1,67$ $C_{pk}=1,33$

n=3	k=10	15	20	25	50	100
0,01	0,940	0,996	1,031	1,057	1,126	1,179
0,02	0,975	1,027	1,060	1,084	1,147	1,195
0,05	1,081	1,077	1,105	1,126	1,179	1,220
0,50	1,330	1,330	1,330	1,330	1,330	1,330
0,95	1,842	1,722	1,657	1,616	1,521	1,460
0,98	2,034	1,856	1,764	1,706	1,577	1,496
0,99	2,185	1,958	1,843	1,771	1,616	1,521

Podm. Pst. $C_p=1,67$ $C_{pk}=1,33$

n=3	k=10	15	20	25	50	100
0,01	0,940	0,996	1,031	1,057	1,126	1,179
0,02	0,975	1,027	1,060	1,084	1,147	1,195
0,05	1,081	1,077	1,105	1,126	1,179	1,220
0,50	1,330	1,330	1,330	1,330	1,330	1,330
0,95	1,842	1,722	1,657	1,616	1,521	1,460
0,98	2,034	1,856	1,764	1,706	1,577	1,496
0,99	2,185	1,957	1,843	1,771	1,616	1,521

Součin $C_p=1,67$ $C_{pk}=1,33$

n=5	k=10	15	20	25	50	100
0,01	1,028	1,074	1,103	1,124	1,178	1,219
0,02	1,057	1,099	1,126	1,145	1,194	1,231
0,05	1,103	1,139	1,162	1,178	1,219	1,250
0,50	1,330	1,330	1,330	1,330	1,330	1,330
0,95	1,660	1,588	1,549	1,522	1,461	1,420
0,98	1,767	1,668	1,614	1,579	1,497	1,444
0,99	1,846	1,725	1,660	1,618	1,522	1,461

Min $C_p=1,67$ $C_{pk}=1,33$

n=5	k=10	15	20	25	50	100
0,01	1,028	1,074	1,103	1,124	1,178	1,219
0,02	1,057	1,099	1,126	1,145	1,194	1,231
0,05	1,103	1,139	1,162	1,178	1,219	1,250
0,50	1,330	1,330	1,330	1,330	1,330	1,330
0,95	1,660	1,588	1,549	1,522	1,461	1,420
0,98	1,767	1,668	1,614	1,579	1,497	1,444
0,99	1,846	1,725	1,660	1,618	1,522	1,461

Podm. Pst. $C_p=1,67$ $C_{pk}=1,33$

n=5	k=10	15	20	25	50	100
0,01	1,028	1,074	1,103	1,124	1,178	1,219
0,02	1,057	1,099	1,126	1,145	1,194	1,231
0,05	1,103	1,139	1,162	1,178	1,219	1,250
0,50	1,330	1,330	1,330	1,330	1,330	1,330
0,95	1,660	1,588	1,549	1,522	1,461	1,420
0,98	1,767	1,668	1,614	1,579	1,497	1,444
0,99	1,846	1,725	1,660	1,618	1,522	1,461

Součin $C_p=1,67$ $C_{pk}=1,33$

n=10	k=10	15	20	25	50	100
0,01	1,106	1,142	1,164	1,180	1,220	1,251
0,02	1,128	1,161	1,181	1,196	1,232	1,260
0,05	1,164	1,192	1,209	1,220	1,251	1,273
0,50	1,330	1,330	1,330	1,330	1,330	1,330
0,95	1,545	1,501	1,176	1,459	1,419	1,392
0,98	1,609	1,550	1,517	1,495	1,443	1,408
0,99	1,655	1,584	1,545	1,520	1,459	1,419

Min Cp=1,67 Cpk=1,33

n=10	k=10	15	20	25	50	100
0,01	1,106	1,142	1,164	1,180	1,220	1,251
0,02	1,128	1,161	1,181	1,196	1,232	1,260
0,05	1,164	1,192	1,209	1,220	1,251	1,273
0,50	1,330	1,330	1,330	1,330	1,330	1,330
0,95	1,545	1,501	1,476	1,459	1,419	1,392
0,98	1,609	1,550	1,517	1,495	1,443	1,408
0,99	1,655	1,584	1,545	1,520	1,459	1,419

Podm. Pst. Cp=1,67 Cpk=1,33

n=10	k=10	15	20	25	50	100
0,01	1,106	1,142	1,164	1,180	1,220	1,251
0,02	1,128	1,161	1,181	1,196	1,232	1,260
0,05	1,164	1,192	1,209	1,220	1,251	1,273
0,50	1,330	1,330	1,330	1,330	1,330	1,330
0,95	1,545	1,501	1,476	1,459	1,419	1,392
0,98	1,609	1,550	1,517	1,495	1,443	1,408
0,99	1,655	1,584	1,545	1,520	1,459	1,419

Na základě tohoto srovnání lze tvrdit, že pro numerické vyjádření průběhu hustoty rozdělení pravděpodobnosti pro odhad ukazatele Cpk jsou všechny tři přístupy uvedené v práci naprosto ekvivalentní a rozdíly v hodnotách příslušných kvantilů, pokud nějaké jsou, pak jsou v tisícinách.

Numerická studie o vlivu volby odhadu směrodatné odchylky

Srovnání kvantilů rozdělení pravděpodobnosti odhadu ukazatele C_{pk} vůči použitému odhadu inherentní směrodatné odchylky $\sigma_R, \sigma_s, \sigma_I$:

n =velikost podskupiny, k =počet podskupin, α = pravděpodobnost odpovídající příslušnému kvantilu

Výpočet kvantilů je založen na tvaru příslušné hustoty odhadu ukazatele C_{pk} odvozeném od součinu veličin $1-K$ a C_p

$C_p=1,33$ $C_{pk}=1$ $n=3$

k	10			15			20			25			50			100		
α	R	s	I	R	s	I	R	s	I	R	s	I	R	s	I	R	s	I
0,01	0,696	0,697	0,705	0,740	0,740	0,748	0,768	0,768	0,775	0,788	0,788	0,794	0,842	0,842	0,847	0,883	0,884	0,887
0,02	0,724	0,724	0,734	0,763	0,764	0,774	0,790	0,791	0,798	0,809	0,809	0,816	0,858	0,858	0,863	0,896	0,896	0,900
0,05	0,768	0,768	0,780	0,803	0,804	0,813	0,826	0,826	0,834	0,842	0,842	0,849	0,883	0,884	0,889	0,915	0,915	0,919
0,50	1	1	1,016	1	1	1,011	1	1	1,008	1	1	1,006	1	1	1,003	1	1	1,002
0,95	1,393	1,391	1,377	1,301	1,300	1,290	1,252	1,251	1,242	1,220	1,219	1,211	1,147	1,146	1,141	1,100	1,100	1,096
0,98	1,539	1,536	1,497	1,403	1,402	1,377	1,333	1,331	1,313	1,289	1,288	1,272	1,190	1,189	1,180	1,128	1,127	1,122
0,99	1,654	1,650	1,586	1,481	1,478	1,440	1,393	1,391	1,363	1,338	1,337	1,315	1,220	1,219	1,207	1,147	1,146	1,139

$C_p=1,33$ $C_{pk}=1$ $n=5$

k	10			15			20			25			50			100		
α	R	s	I	R	s	I	R	s	I	R	s	I	R	s	I	R	s	I
0,01	0,764	0,768	0,771	0,800	0,803	0,807	0,823	0,826	0,829	0,839	0,842	0,845	0,882	0,844	0,886	0,914	0,915	0,917
0,02	0,787	0,790	0,795	0,820	0,823	0,827	0,841	0,844	0,847	0,856	0,858	0,861	0,894	0,896	0,896	0,923	0,925	0,926
0,05	0,823	0,826	0,832	0,851	0,854	0,859	0,869	0,871	0,875	0,882	0,884	0,887	0,914	0,915	0,918	0,938	0,939	0,940
0,50	1	1	1,008	1	1	1,005	1	1	1,004	1	1	1,003	1	1	1,002	1	1	1,001
0,95	1,254	1,248	1,245	1,200	1,195	1,192	1,169	1,165	1,162	1,149	1,145	1,143	1,101	1,099	1,097	1,070	1,068	1,067
0,98	1,337	1,328	1,316	1,261	1,254	1,246	1,219	1,214	1,207	1,192	1,187	1,182	1,129	1,126	1,123	1,089	1,087	1,085
0,99	1,397	1,386	1,367	1,304	1,297	1,284	1,254	1,248	1,239	1,222	1,217	1,209	1,149	1,145	1,141	1,101	1,099	1,096

$C_p=1,33$ $C_{pk}=1$ $n=10$

k	10			15			20			25			50			100		
α	R	s	I	R	s	I	R	s	I	R	s	I	R	s	I	R	s	I
0,01	0,825	0,834	0,836	0,853	0,861	0,862	0,871	0,878	0,879	0,883	0,890	0,891	0,915	0,920	0,921	0,938	0,942	0,943
0,02	0,843	0,851	0,853	0,868	0,876	0,877	0,884	0,891	0,892	0,895	0,902	0,903	0,924	0,929	0,929	0,945	0,949	0,949
0,05	0,871	0,878	0,880	0,892	0,898	0,900	0,905	0,911	0,913	0,915	0,920	0,921	0,938	0,942	0,943	0,956	0,959	0,959
0,50	1	1	1,004	1	1	1,002	1	1	1,002	1	1	1,001	1	1	1,001	1	1	1
0,95	1,166	1,153	1,153	1,132	1,122	1,122	1,113	1,104	1,104	1,100	1,093	1,092	1,069	1,064	1,064	1,048	1,045	1,044
0,98	1,216	1,198	1,195	1,170	1,157	1,155	1,145	1,133	1,132	1,127	1,118	1,116	1,087	1,081	1,080	1,060	1,056	1,056
0,99	1,250	1,229	1,225	1,197	1,182	1,178	1,166	1,153	1,151	1,147	1,135	1,133	1,100	1,093	1,091	1,069	1,064	1,063

Cp=1,67 Cpk=1,33 n=3

k	10			15			20			25			50			100		
α	R	s	I	R	s	I	R	s	I	R	s	I	R	s	I	R	s	I
0,01	0,940	0,941	0,952	0,996	0,996	1,006	1,031	1,032	1,041	1,057	1,058	1,066	1,126	1,126	1,133	1,179	1,180	1,185
0,02	0,975	0,976	0,988	1,027	1,027	1,038	1,060	1,061	1,070	1,084	1,084	1,093	1,147	1,147	1,154	1,195	1,196	1,201
0,05	1,031	1,032	1,048	1,077	1,077	1,090	1,105	1,106	1,117	1,126	1,126	1,136	1,179	1,180	1,187	1,220	1,220	1,225
0,50	1,33	1,33	1,352	1,33	1,33	1,345	1,33	1,33	1,341	1,33	1,33	1,339	1,33	1,33	1,334	1,33	1,33	1,332
0,95	1,843	1,840	1,821	1,722	1,721	1,706	1,658	1,656	1,644	1,616	1,615	1,604	1,521	1,520	1,513	1,460	1,459	1,455
0,98	2,035	2,031	1,977	1,856	1,854	1,820	1,764	1,762	1,736	1,706	1,704	1,683	1,577	1,576	1,563	1,496	1,495	1,488
0,99	2,185	2,180	2,092	1,958	1,955	1,902	1,843	1,840	1,802	1,772	1,770	1,739	1,616	1,615	1,598	1,521	1,520	1,510

Cp=1,67 Cpk=1,33 n=5

k	10			15			20			25			50			100		
α	R	s	I	R	s	I	R	s	I	R	s	I	R	s	I	R	s	I
0,01	1,028	1,033	1,038	1,074	1,078	1,083	1,103	1,107	1,111	1,124	1,127	1,131	1,178	1,181	1,184	1,219	1,221	1,223
0,02	1,057	1,062	1,068	1,099	1,103	1,108	1,126	1,130	1,134	1,145	1,148	1,152	1,194	1,197	1,200	1,231	1,233	1,285
0,05	1,103	1,107	1,115	1,139	1,143	1,149	1,162	1,165	1,170	1,178	1,181	1,185	1,219	1,221	1,224	1,250	1,251	1,253
0,50	1,33	1,33	1,341	1,33	1,33	1,337	1,33	1,33	1,335	1,33	1,33	1,334	1,33	1,33	1,332	1,33	1,33	1,331
0,95	1,660	1,651	1,646	1,588	1,582	1,578	1,549	1,543	1,539	1,522	1,518	1,514	1,461	1,458	1,455	1,420	1,418	1,416
0,98	1,767	1,755	1,739	1,668	1,659	1,648	1,614	1,606	1,598	1,579	1,572	1,565	1,497	1,493	1,489	1,444	1,442	1,439
0,99	1,846	1,831	1,805	1,725	1,714	1,697	1,660	1,651	1,638	1,618	1,611	1,600	1,522	1,518	1,512	1,161	1,458	1,454

Cp=1,67 Cpk=1,33 n=10

k	10			15			20			25			50			100		
α	R	s	I	R	s	I	R	s	I	R	s	I	R	s	I	R	s	I
0,01	1,106	1,119	1,120	1,142	1,153	1,154	1,164	1,174	1,175	1,180	1,189	1,190	1,220	1,228	1,228	1,251	1,256	1,257
0,02	1,128	1,140	1,143	1,161	1,171	1,173	1,181	1,191	1,192	1,196	1,204	1,206	1,232	1,239	1,240	1,260	1,264	1,265
0,05	1,164	1,174	1,177	1,192	1,200	1,203	1,209	1,216	1,218	1,220	1,228	1,229	1,251	1,256	1,257	1,273	1,277	1,278
0,50	1,33	1,33	1,335	1,33	1,33	1,333	1,33	1,33	1,332	1,33	1,33	1,332	1,33	1,33	1,331	1,33	1,33	1,330
0,95	1,545	1,527	1,527	1,501	1,487	1,487	1,476	1,464	1,464	1,459	1,449	1,448	1,419	1,412	1,412	1,392	1,387	1,387
0,98	1,609	1,585	1,581	1,550	1,532	1,529	1,517	1,502	1,499	1,495	1,482	1,480	1,443	1,434	1,433	1,408	1,402	1,401
0,99	1,655	1,626	1,619	1,584	1,563	1,558	1,545	1,527	1,524	1,520	1,504	1,501	1,459	1,449	1,447	1,419	1,412	1,411

Shrnutí numerické studie: Na základě velikosti kvantilového rozpětí lze tvrdit, že nejkratší rozpětí vykazuje použití „pooled standard deviation“, zatím co obě dvě zbývající možnosti jsou prakticky rovnocenné pro malý rozsah podskupiny, ale začínají se lišit při větším rozsahu podskupiny ve prospěch klasické směrodatné odchylky. Zde ještě hraje roli počet uvažovaných podskupin a lze říci, že s rostoucím počtem podskupin se rozdíly mezi uvažovanými odhady inherentní směrodatné odchylky stírají.

**Numerická studie porovnání kvantilů rozdělení pravděpodobnosti
odhadu ukazatele Cpk a odhadu ukazatele CpkL
(aproximace hustoty odhadu Cpk pomocí hustoty odhadu CpkL)**

k=počet podskupin n=velikost podskupiny

α		Cpk	CpkL		Cpk	CpkL		Cpk	CpkL		Cpk	CpkL
0,01	Cp=1	0,746	0,765	Cp=1,33	0,764	0,765	Cp=1	0,809	0,824	Cp=1,33	0,823	0,824
0,02	Cpk=1	0,767	0,787	Cpk=1	0,787	0,787	Cpk=1	0,826	0,841	Cpk=1	0,841	0,841
0,05	CpkL=1	0,800	0,823	CpkL=1	0,823	0,823	CpkL=1	0,852	0,869	CpkL=1	0,869	0,869
0,50	k=10	0,963	1	k=10	1	1	k=20	0,974	1	k=20	1	1
0,95	n=5	1,198	1,255	n=5	1,254	1,255	n=5	1,131	1,169	n=5	1,169	1,169
0,98		1,275	1,337		1,337	1,337		1,178	1,220		1,219	1,220
0,99		1,332	1,398		1,397	1,389		1,212	1,256		1,254	1,256

α		Cpk	CpkL		Cpk	CpkL		Cpk	CpkL		Cpk	CpkL
0,01	Cp=1	0,872	0,881	Cp=1,33	0,882	0,881	Cp=1	0,907	0,920	Cp=1,33	0,914	0,920
0,02	Cpk=1	0,884	0,894	Cpk=1	0,894	0,894	Cpk=1	0,916	0,927	Cpk=1	0,923	0,927
0,05	CpkL=1	0,902	0,814	CpkL=1	0,914	0,914	CpkL=1	0,929	0,940	CpkL=1	0,938	0,940
0,50	k=50	0,983	1	k=50	1	1	k=100	0,988	1,001	k=100	1	1,001
0,95	n=5	1,079	1,101	n=5	1,101	1,101	n=5	1,054	1,073	n=5	1,070	1,073
0,98		1,105	1,129		1,129	1,129		1,072	1,099		1,089	1,099
0,99		1,123	1,148		1,149	1,148		1,084	1,101		1,101	1,101

α		Cpk	CpkL		Cpk	CpkL		Cpk	CpkL		Cpk	CpkL
0,01	Cp=1,33	0,764	0,765	Cp=1,67	0,764	0,765	Cp=1,33	0,823	0,824	Cp=1,67	0,823	0,824
0,02	Cpk=1,33	0,787	0,787	Cpk=1,33	0,787	0,787	Cpk=1,33	0,841	0,841	Cpk=1,33	0,841	0,841
0,05	CpkL=1,33	0,823	0,823	CpkL=1,33	0,823	0,823	CpkL=1,33	0,869	0,869	CpkL=1,33	0,869	0,869
0,50	k=10	1	1	k=10	1	1	k=20	1	1	k=20	1	1
0,95	n=5	1,254	1,255	n=5	1,254	1,255	n=5	1,169	1,169	n=5	1,169	1,169
0,98		1,337	1,337		1,337	1,337		1,219	1,220		1,219	1,220
0,99		1,397	1,398		1,397	1,398		1,254	1,256		1,254	1,256

α		Cpk	CpkL		Cpk	CpkL		Cpk	CpkL		Cpk	CpkL
0,01	Cp=1,33	0,882	0,881	Cp=1,67	0,882	0,881	Cp=1,33	0,914	0,920	Cp=1,67	0,914	0,920
0,02	Cpk=1,33	0,894	0,894	Cpk=1,33	0,894	0,894	Cpk=1,33	0,923	0,927	Cpk=1,33	0,923	0,927
0,05	CpkL=1,33	0,914	0,914	CpkL=1,33	0,914	0,914	CpkL=1,33	0,938	0,940	CpkL=1,33	0,938	0,940
0,50	k=50	1	1	k=50	1	1	k=100	1	1,001	k=100	1	1,001
0,95	n=5	1,101	1,101	n=5	1,101	1,101	n=5	1,070	1,073	n=5	1,070	1,073
0,98		1,129	1,129		1,129	1,129		1,089	1,099		1,089	1,099
0,99		1,149	1,148		1,149	1,148		1,101	*		1,101	*

Shrnutí numerické studie: Studie odhalila zajímavé zjištění, že studovaná aproximace hustoty odhadu ukazatele Cpk je lepší, čím je větší odstup mezi ukazateli Cp a Cpk. Když Cpk=Cp, pak jsou rozdíly řádově v 5-6ti setinách. Rovněž rozdíly se zmenšují, pokud rostou hodnoty ukazatelů a počet podskupin. Největší rozdíly se vyskytují, když Cp=Cpk=1 a při malém počtu podskupin. Lze ale říci, že pro praktické účely je navrhovaná aproximace uspokojivá.

Numerická studie kvantilů rozdělení pravděpodobnosti pro podíl odhad C_{pk}/C_p

k =počet podskupin n =velikost podskupiny

$C_p=1$ $n=3$ $C_{pk} = \rho C_p$

	k=10			k=20			k=25			k=50		
α	$\rho=1$	0,75	0,67	$\rho=1$	0,75	0,67	$\rho=1$	0,75	0,67	$\rho=1$	0,75	0,67
0,01	0,683	0,687	0,676	0,757	0,762	0,754	0,778	0,783	0,776	0,834	0,838	0,834
0,025	0,719	0,728	0,720	0,786	0,795	0,789	0,805	0,813	0,808	0,855	0,862	0,858
0,05	0,752	0,767	0,759	0,813	0,825	0,819	0,829	0,841	0,836	0,874	0,883	0,879
0,5	0,968	1,016	1,016	0,974	1,008	1,008	0,976	1,006	1,006	0,982	1,003	1,003
0,95	1,299	1,392	1,400	1,191	1,252	1,258	1,167	1,220	1,225	1,111	1,148	1,151
0,975	1,383	1,486	1,496	1,242	1,308	1,315	1,210	1,269	1,275	1,139	1,178	1,183
0,99	1,491	1,608	1,620	1,304	1,378	1,386	1,263	1,328	1,335	1,173	1,216	1,221

$C_p=1,33$ $n=3$ $C_{pk} = \rho C_p$

	k=10			k=20			k=25			k=50		
α	$\rho=1$	0,75	0,67	$\rho=1$	0,75	0,67	$\rho=1$	0,75	0,67	$\rho=1$	0,75	0,67
0,01	0,697	0,705	0,699	0,768	0,775	0,770	0,788	0,794	0,790	0,841	0,847	0,844
0,025	0,732	0,744	0,739	0,796	0,806	0,802	0,814	0,823	0,820	0,862	0,869	0,866
0,05	0,765	0,780	0,775	0,882	0,834	0,831	0,838	0,849	0,846	0,880	0,889	0,887
0,5	0,980	1,016	1,016	0,982	1,008	1,008	0,984	1,006	1,006	0,987	1,003	1,003
0,95	1,312	1,377	1,382	1,119	1,242	1,245	1,174	1,212	1,214	1,116	1,141	1,143
0,975	1,396	1,468	1,474	1,250	1,296	1,300	1,217	1,258	1,262	1,144	1,171	1,174
0,99	1,506	1,586	1,593	1,312	1,363	1,368	1,271	1,315	1,320	1,178	1,207	1,210

$C_p=1,67$ $n=3$ $C_{pk} = \rho C_p$

	k=10			k=20			k=25			k=50		
α	$\rho=1$	0,75	0,67	$\rho=1$	0,75	0,67	$\rho=1$	0,75	0,67	$\rho=1$	0,75	0,67
0,01	0,705	0,714	0,710	0,774	0,781	0,778	0,793	0,800	0,798	0,845	0,851	0,849
0,025	0,740	0,752	0,748	0,802	0,811	0,809	0,819	0,828	0,826	0,866	0,872	0,871
0,05	0,772	0,786	0,784	0,828	0,839	0,837	0,843	0,853	0,851	0,884	0,892	0,890
0,5	0,987	1,017	1,017	0,988	1,008	1,008	0,988	1,007	1,007	0,990	1,003	1,003
0,95	1,321	1,370	1,374	1,205	1,237	1,239	1,179	1,207	1,209	1,119	1,138	1,140
0,975	1,405	1,460	1,464	1,256	1,290	1,293	1,222	1,253	1,255	1,147	1,167	1,169
0,99	1,515	1,576	1,580	1,318	1,356	1,359	1,276	1,309	1,312	1,181	1,203	1,205

Cp=1 n=5 Cpk = ρ Cp

	k=10			k=20			k=25			k=50		
α	ρ=1	0,75	0,67	ρ=1	0,75	0,67	ρ=1	0,75	0,67	ρ=1	0,75	0,67
0,01	0,753	0,756	0,747	0,815	0,818	0,812	0,832	0,835	0,829	0,876	0,879	0,875
0,025	0,782	0,790	0,782	0,838	0,844	0,839	0,858	0,859	0,854	0,892	0,897	0,893
0,05	0,809	0,821	0,814	0,858	0,871	0,863	0,871	0,880	0,876	0,906	0,913	0,910
0,5	0,971	1,008	1,005	0,978	1,004	1,004	0,980	1,003	1,003	0,985	1,002	1,002
0,95	1,188	1,257	1,263	1,124	1,171	1,175	1,109	1,150	1,154	1,074	1,102	1,105
0,975	1,138	1,314	1,322	1,156	1,207	1,212	1,137	1,182	1,187	1,093	1,123	1,127
0,99	1,301	1,385	1,394	1,195	1,251	1,257	1,171	1,220	1,226	1,115	1,148	1,152

Cp=1,33 n=5 Cpk = ρ Cp

	k=10			k=20			k=25			k=50		
α	ρ=1	0,75	0,67	ρ=1	0,75	0,67	ρ=1	0,75	0,67	ρ=1	0,75	0,67
0,01	0,765	0,771	0,766	0,823	0,829	0,825	0,839	0,844	0,841	0,885	0,889	0,888
0,025	0,793	0,803	0,799	0,846	0,853	0,850	0,860	0,867	0,864	0,900	0,906	0,904
0,05	0,819	0,832	0,828	0,866	0,875	0,873	0,879	0,887	0,885	0,914	0,920	0,919
0,5	0,980	1,008	1,008	0,984	1,004	1,004	0,985	1,003	1,003	0,992	1,002	1,002
0,95	1,197	1,245	1,249	1,130	1,162	1,165	1,115	1,143	1,145	1,080	1,095	1,096
0,975	1,248	1,299	1,304	1,162	1,197	1,200	1,142	1,173	1,176	1,009	1,114	1,116
0,99	1,310	1,367	1,373	1,201	1,239	1,243	1,176	1,209	1,213	1,121	1,137	1,139

Cp=1,67 n=5 Cpk = ρ Cp

	k=10			k=20			k=25			k=50		
α	ρ=1	0,75	0,67	ρ=1	0,75	0,67	ρ=1	0,75	0,67	ρ=1	0,75	0,67
0,01	0,771	0,779	0,776	0,828	0,834	0,832	0,844	0,849	0,847	0,885	0,889	0,888
0,025	0,800	0,809	0,807	0,850	0,858	0,856	0,864	0,871	0,869	0,900	0,906	0,904
0,05	0,826	0,837	0,835	0,870	0,879	0,877	0,883	0,891	0,889	0,914	0,920	0,919
0,5	0,986	1,008	1,008	0,988	1,004	1,004	0,989	1,003	1,003	0,992	1,002	1,002
0,95	1,203	1,239	1,242	1,134	1,158	1,160	1,118	1,139	1,141	1,080	1,095	1,096
0,975	1,253	1,292	1,295	1,166	1,192	1,194	1,146	1,169	1,171	1,099	1,114	1,116
0,99	1,316	1,359	1,363	1,205	1,233	1,236	1,180	1,204	1,206	1,121	1,137	1,139

Cp=1 n=10 Cpk = ρ Cp

	k=10			k=20			k=25			k=50		
α	ρ=1	0,75	0,67	ρ=1	0,75	0,67	ρ=1	0,75	0,67	ρ=1	0,75	0,67
0,01	0,822	0,824	0,817	0,868	0,871	0,866	0,881	0,883	0,879	0,914	0,915	0,912
0,025	0,844	0,849	0,844	0,885	0,890	0,886	0,896	0,901	0,897	0,925	0,928	0,926
0,05	0,863	0,872	0,867	0,900	0,907	0,903	0,910	0,916	0,913	0,935	0,939	0,937
0,5	0,977	1,003	1,003	0,983	1,002	1,002	0,985	1,001	1,001	0,989	1,001	1,001
0,95	1,115	1,162	1,166	1,078	1,110	1,113	1,069	1,097	1,100	1,048	1,067	1,069
0,975	1,145	1,196	1,201	1,098	1,133	1,137	1,086	1,117	1,121	1,059	1,081	1,083
0,99	1,181	1,237	1,244	1,121	1,159	1,164	1,107	1,141	1,145	1,073	1,097	1,100

Cp=1,33 n=10 Cpk = ρ Cp

	k=10			k=20			k=25			k=50		
α	ρ=1	0,75	0,67	ρ=1	0,75	0,67	ρ=1	0,75	0,67	ρ=1	0,75	0,67
0,01	0,831	0,835	0,832	0,875	0,879	0,876	0,887	0,891	0,888	0,918	0,921	0,919
0,025	0,852	0,859	0,853	0,891	0,897	0,894	0,902	0,907	0,905	0,929	0,933	0,931
0,05	0,871	0,880	0,878	0,906	0,913	0,911	0,915	0,921	0,919	0,939	0,943	0,942
0,5	0,984	1,004	1,004	0,989	1,002	1,002	0,989	1,001	1,001	0,992	1,001	1,001
0,95	1,121	1,153	1,156	1,082	1,104	1,106	1,072	1,092	1,094	1,050	1,064	1,065
0,975	1,150	1,185	1,189	1,101	1,125	1,128	1,090	1,111	1,113	1,062	1,076	1,078
0,99	1,186	1,225	1,229	1,125	1,151	1,154	1,110	1,133	1,136	1,075	1,091	1,093

Cp=1,67 n=10 Cpk = ρ Cp

	k=10			k=20			k=25			k=50		
α	ρ=1	0,75	0,67	ρ=1	0,75	0,67	ρ=1	0,75	0,67	ρ=1	0,75	0,67
0,01	0,836	0,841	0,839	0,879	0,883	0,881	0,890	0,894	0,893	0,920	0,923	0,922
0,025	0,857	0,864	0,862	0,895	0,900	0,899	0,905	0,910	0,909	0,931	0,935	0,934
0,05	0,876	0,884	0,883	0,909	0,916	0,914	0,918	0,924	0,923	0,941	0,945	0,944
0,5	0,988	1,004	1,004	0,991	1,002	1,002	0,991	1,001	1,001	0,994	1,001	1,001
0,95	1,124	1,149	1,151	1,084	1,101	1,102	1,075	1,089	1,091	1,059	1,062	1,063
0,975	1,154	1,180	1,183	1,104	1,122	1,123	1,092	1,108	1,109	1,063	1,074	1,075
0,99	1,190	1,218	1,221	1,127	1,147	1,148	1,112	1,129	1,131	1,077	1,089	1,090