

10

KONTROLA A ŘÍZENÍ JAKOSTI

10.1 Podstata úloh řízení jakosti

Pojem jakost (kvalita) se používá jak v celé hierarchii řízení podniků, tak v obchodní sféře a sféře spotřebitelů. Přitom se ukazuje, že již vlastní definice tohoto pojmu je často zmatená a neumožňuje jakost stanovit. Dobře je tento fakt demonstrován na výsledcích ankety mezi špičkovými britskými manažery. Na základě vyhodnocení této ankety bylo zjištěno, že:

- (a) Většina řídicích pracovníků cítí potřebu řízení jakosti;
- (b) Nikdo z řídicích pracovníků nemá pocit, že by měl za jakost osobně zodpovídat;
- (c) Většina řídicích pracovníků nedovede přesně vyjádřit, co to jakost vlastně je.

Závěrem uvedené ankety bylo konstatováno, že manažery zajímají tři problémy: *peníze, jak je získat a jak je neztratit.*

V této kapitole je učiněn pokus o naznačení základních přístupů k hodnocení jakosti s aplikací na průmysl¹⁻⁶. Je uveden základní matematicko-statistický aparát pro konstrukci regulačních diagramů, indexů způsobilosti procesu a souvisejících úloh. Konstrukce a použití regulačních diagramů je nejlépe známou statistickou technikou řízení jakosti. Přes velké rozšíření regulačních diagramů se často setkáváme s jejich nesprávným použitím, způsobeným nerespektováním některých zásad, anebo zanedbáním základních předpokladů o datech, jako je normalita, nezávislost, stabilita. Důsledkem je nesprávná interpretace diagramů, nedůvěra ke statistickým metodám, upouštění od statistické regulace, někdy i zkreslování výsledků a nedovolená manipulace s daty.

Definice jakosti: V literatuře lze nalézt celou řadu více či méně obecných definic jakosti. Populární Juranova učebnice jakosti⁷ uvádí definici: *"Jakost je vyjádřením vhodnosti k užívání"*. Podrobnější definici lze nalézt v normě ANSI/ASCQ z r. 1978: *"Jakost je souhrn rysů a charakteristik produktu nebo služby, který zajišťuje jeho schopnost vyhovět daným požadavkům"*. Velmi výstižná a obecná je definice z našich ČSN: *"Jakost výrobku je souhrnem vlastností podmiňujících způsobilost uspokojit potřeby odpovídající jeho účelu použití"*. Tato definice obsahuje v praxi často opomíjený fakt, že jakost je vždy spjatá s účelem použití. Nelze tedy říci, že se vyrábí jakostní výrobky, aniž je známo k čemu budou použity. To je např. v případě posuzování jakosti polotovarů často omezující.

Užitná hodnota: Existují jisté znaky jakosti, vyjádřené tzv. *užitnými vlastnostmi*, které mohou být buď jednoduše měřitelné (délkové rozměry, pevnost, tažnost, vlhkost, koncentrace), nebo přímo neměřitelné, většinou subjektivní (vůně, chuť, tvar, barevnost,

komfort při použití, vzhled), které jsou zejména pro výrobky spotřebního charakteru rozhodující. Předpokládejme, že lze obecně specifikovat K -tici užitečných vlastností R_1, \dots, R_K . Na základě přímých a nepřímých měření lze pak stanovit *ukazatele jakosti* (průměr, rozptyl, kvantily, podíl prvků mimo meze atd.) x_1, \dots, x_K . Tyto ukazatele charakterizují užitečné vlastnosti. Z hodnot x_i pro i -tou užitečnou vlastnost lze pomocí vhodné funkční transformace definovat dílčí úroveň jakosti

$$u_i = f(x_i, K_D, K_H),$$

kde K_D je předepsaná hodnota užitečné vlastnosti pro právě nevyhovující ($u_i = 0$) a K_H pro právě vyhovující ($u_i = 1$) výrobek. Celková úroveň jakosti, označovaná jako *užitná hodnota výrobku*, je pak vhodný vážený obecný průměr dílčích úrovní

$$u = \text{ave}(u_i, w_i),$$

kde w_i jsou váhy definující význam dané užitečné vlastnosti a související s účelem použití výrobku⁶. S ohledem na své vlastnosti (pro nulové u_i vychází také $u = 0$) se obvykle používá *vážený geometrický průměr*. Popsaný postup vycházející z obecné definice jakosti vyžaduje:

- (a) nalezení pokud možno úplné množiny významných užitečných vlastností,
- (b) stanovení jejich velikosti (měření),
- (c) nalezení vhodných vah.

Při konstrukci užitné hodnoty se projeví *hledisko hodnocení*.

(1) *Výrobce* bude zřejmě preferovat dodržení technologických parametrů výroby a snažit se omezovat variabilitu produktů.

(2) *Zpracovatel* bude hodnotit zpracovatelské vlastnosti vstupujícího "meziprojektu" a jejich vliv na jakost vyráběného produktu.

(3) *Spotřebitel* budou zřejmě zajímat užitečné vlastnosti, které nemusí přímo souviset s jakostí, vyjádřenou z hlediska výrobce a zpracovatele (organoleptické vlastnosti, vzhled, životnost atd.).

Protože je užitná hodnota u stanovena na základě experimentálních údajů, jde o náhodnou veličinu, pro kterou lze určit střední hodnotu $E(u)$, rozptyl $D(u)$ a interval spolehlivosti střední hodnoty. Na základě těchto údajů lze pak porovnávat rozdíly mezi užitečnými hodnotami výrobků s ohledem na přesnosti měření jednotlivých charakteristik. Komplexní charakteristika jakosti, *užitná hodnota*, se při řízení jakosti přímo ve výrobě uplatňuje velmi obtížně. Hodí se spíše pro komparaci finálních výrobků.

Ztrátová funkce: Moderní postupy "inženýrství kvality" využívají Taguchiho definice jakosti⁵: "*Jakost produktu je úměrná ztrátě způsobené společností odchylkou od předepsaných (cílových) hodnot*". Přitom ztráta, způsobená společností, zahrnuje různé opravy, čištění, přerušení výroby, likvidace, odpady, nespokojenost zákazníka, ztráty trhu, náklady na reklamaci, arbitráž atd. a vyjadřuje se v peněžních jednotkách. Pro jeden parametr jakosti x lze uvedenou ztrátu definovat jednoduše pomocí *kvadratické ztrátové funkce*

$$L(x) = K(x - T)^2,$$

kde T je předepsaná cílová hodnota a K je parametr určený z odchylky x od T právě o definovanou, předepsanou toleranci, stanovenou výrobcem nebo akceptovanou spotřebitelem. Ztrátovou funkci $L(x)$ je třeba pro některé parametry jakosti poněkud modifikovat:

a) pro případy kdy "nižší hodnota je lepší" (nestejnoměrnost, vady, obsah škodlivin) se používá jednoduchá *ztrátová funkce konvexně rostoucí*

$$L(x) = Kx^2,$$

b) pro případy, kdy "vyšší hodnota je lepší" (pevnost, stálost, zralost) se používá *konvexně klesající ztrátová funkce*, která má v nejjednodušším případě tvar

$$L(x) = \frac{K}{x^2},$$

c) pro případy, kdy "nominální hodnota je nejlepší" (rozměr, plošná hmotnost, tažnost atd.), se často místo funkce $K(x - T)^2$ používá asymetrická varianta ztrátové funkce

$$\begin{aligned} L(x) &= K_1(x - T)^2 && \text{pro } x \leq T, \\ L(x) &= K_2(x - T)^2 && \text{pro } x > T. \end{aligned}$$

Tyto tvary ztrátových funkcí odrážejí obecně to, jakým způsobem ovlivňují jednotlivé parametry jakosti ztrátu způsobenou společností. Uvedené definice ztrátové funkce vycházejí z předpokladu, že je parametr jakosti x stanoven přesně, čili jde o deterministickou proměnnou. Zejména ve zpracovatelském průmyslu, kde je variabilita složení a struktury vždy značná, je třeba uvažovat, že na základě výběru (z různých míst, v různých časech) lze získat pouze odhad střední hodnoty x_s (aritmetický průměr) a odpovídající rozptyl s^2 (výběrový rozptyl). Ztrátová funkce má pak tvar

$$L(x) = K[s^2 + (x_s - T)^2].$$

Při vlastním řízení jakosti pomocí ztrátové funkce lze na základě této rovnice určit, zda je výhodnější snižovat variabilitu s^2 nebo se blížit předepsané hodnotě T . Je také zřejmé, že i při dodržení předepsané hodnoty parametru jakosti může být ztráta způsobená variabilitou výroby nebo výrobku poměrně značná.

Techniky řízení jakosti: S ohledem na historii řízení jakosti lze specifikovat tři základní koncepce:

(a) *Přejímací plány*, které určují pravidla, podle nichž se na základě analýzy části výrobků usuzuje, zda je celá dodávka přijatelné kvality či nikoliv,

(b) *Statistické řízení procesů*, kdy se monitorují znaky jakosti s využitím regulačních diagramů (on-line regulace),

(c) *Inženýrství jakosti*, kdy se provádí off-line projektování hodnot procesních proměnných s využitím koncepce ztrátové funkce.

V současné době se využívá kombinace všech tří koncepcí řízení jakosti přesto, že zde existuje logický nesoulad. Přejímací plány a regulační diagramy vycházejí z předpokladu *skokové funkce jakosti*, tj. pokud jsou parametry jakosti v zadaných mezích, je *jakost přijatelná*, a pokud jsou mimo tyto meze, je *jakost nepřijatelná*. Inženýrství jakosti využívá *spojité funkce jakosti* (např. ztrátové funkce výše uvedené rovnice), kdy se každá odchylka od ideálního stavu ($x_s = T$, $s^2 = 0$) projeví ztrátou vyjádřenou finančně.

Teorie *přejímacích plánů* je v současné době zpracována jak pro přejímky srovnáváním (diskrétní znaky jakosti), tak i měřením (spojité znaky jakosti). Obecně je statistická přejímka soubor postupů výběrové kontroly jakosti, prováděné na dodávkách surovin, polotovarů a výrobků. Postup statistické přejímky je vlastně testem statistické hypotézy o

parametru jakosti. Pravděpodobnost zamítnutí vyhovující dodávky α se označuje jako *riziko dodavatele* (chyba prvního druhu), pravděpodobnost přijetí nevyhovující dodávky β se označuje jako *riziko odběratele* (chyba druhého druhu). Při statistické přejímce srovnáváním i měřením je třeba určit tzv. *rozhodné číslo* (odpovídá kritické hodnotě u testů) c a rozsah výběru N . Dvojice (c, N) se označuje jako *prejímací plán*. Demonstrujme si určení plánu např. u přejímky měřením. Jakostní znak x nechť je spojitá náhodná veličina s normálním rozdělením. Výrobek se považuje za vadný, pokud překročí veličina x horní toleranční mez USL . Označme P_1 jako přípustný podíl výrobků, pro které bude $x > USL$ a P_2 jako nepřipustný podíl takto definovaných vadných výrobků (obvykle se volí $P_1 = 0.001$ a $P_2 = 0.01$).

$$\text{Riziko dodavatele pak je } P(x_s \geq c + s \cdot USL \cdot P_1)$$

$$\text{a riziko odběratele je } P(x_s \geq c + s \cdot USL \cdot P_2),$$

kde x_s je výběrová střední hodnota a s^2 je výběrový rozptyl. Za předpokladu, že veličina $x_s + c + s$ má přibližně normální rozdělení, lze nalézt parametry přejímacího plánu

$$c = \frac{u_{1-P_1} + u_{1-P_2}}{u_{1-P_1} - u_{1-P_2}},$$

$$N = \frac{u_{1-P_1}^2 (2 + c^2)}{2 (u_{1-P_1} - c)^2},$$

kde u_r je 100 r %ní kvantil normovaného normálního rozdělení. Při volbě $\alpha = \beta = 0.05$ a $P_1 = 0.001$, $P_2 = 0.01$ vyjde $c = 2.71$ a $N = 174$. Při provedení přejímacího plánu se tedy vybere 174 výrobků, stanoví se hodnoty jakostních znaků x_i , $i = 1, \dots, 174$, a určí se odhady x_s , s^2 . Pokud je splněna nerovnost $x_s \geq 2.71 s + USL$, zásilka se přijímá. V opačném případě se zamítá. Kritickým je zde především předpoklad normality a nezávislosti prvků výběru. Obdobným způsobem se konstruují také přejímací plány pro přejímku srovnáváním.

Pro přímé řízení jakosti výroby se používají *toleranční meze*, LSL (dolní) a USL (horní), definující interval, ve kterém leží se zvolenou pravděpodobností předepsané procento výsledků, hodnot parametru jakosti. Na základě tolerančních mezí se definuje tzv. *parametr způsobilosti procesu* (process capability index)

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

kde σ je směrodatná odchylka parametru jakosti. Hodnota $C_p > 1$ svědčí o tom, že jakost je přijatelná čili v podstatě celá výroba je v tolerančních mezích. Je zajímavé, že i ve vyspělých státech, jako jsou USA, vychází často C_p menší než 1. Na druhé straně v Japonsku již v r. 1980 docílili průměrného $C_p = 1.33$, a u "high tech" produktů dokonce $C_p = 2$. Je zřejmé, že pro případ normálního rozdělení leží v mezích $\pm 3\sigma$ přibližně 99.73 % hodnot parametru jakosti. Hodnota $C_p = 1$ pak ukazuje že 99.73 % výrobků je v tolerančních mezích. Hodnota 6 ve jmenovateli rovnice pro C_p je obecně závislá na velikosti výběru, z něhož se odhaduje σ , a na rozdělení znaku jakosti.

Omezení problému s nenormalitou rozdělení znaku jakosti lze docílit vhodnou volbou procenta výrobků ležících v tolerančních mezích. Pokud zvolíme tento parametr 99 %, můžeme použít ve jmenovateli hodnotu 5.15, platnou pro řadu rozdělení, a to se šikmostí od 0 do 3.111 a špičatostí od 1 do 5.997.

Pro případ, že lze předpokládat normální rozdělení parametru jakosti, je také možno vypočítat LSL a USL (pokud nejsou zadány) relativně jednoduše (popsáno v ČSN 01 0230) a C_p lze pak snadno určit. V obecném případě je třeba použít komplikovanější postupy⁰. Je patrné, že přejímací plány a parametry způsobilosti jsou off-line charakteristiky umožňující sledovat jakost mimo vlastní proces výroby. Pro přímé ovlivňování výroby, monitorování kvality, se používají regulační diagramy. Jejich základní myšlenka vychází z intervalů spolehlivosti.

Účelem je ovlivňovat proces výroby tak, aby hodnoty parametrů jakosti ležely uvnitř těchto intervalů: *regulační diagram* je graf, v němž je obvykle vyznačena jistá předepsaná (průměrná) hodnota a regulační meze (příp. jiné kontrolní meze). Do tohoto grafu se postupně vyznačují hodnoty parametru jakosti určované přímo v procesu výroby. V případě, kdy se projeví nenáhodné trendy, následuje seřízení procesu nebo zásah do výroby.

Pro případ spojitého znaku jakosti se obvykle používají *regulační diagramy Shewhartovy*. Při jejich konstrukci se vychází z měřených dat (výběrů) a počítá se vhodná statistika S_i (což může být např. průměr \bar{x}_{si} , směrodatná odchylka s_i , rozpětí R_i atd.). Statistiky S_i se vynášejí do grafu, kde jsou znázorněny regulační meze, odpovídající obvykle intervalům spolehlivosti $E(S_i) \pm 3\sigma D(S_i)$. Zde $E(x)$ značí střední hodnotu a $D(x)$ rozptyl parametru x . Na základě znázorněných bodů se pak usuzuje, zda je proces *statisticky stabilní*. Pokud mají statistiky S_i přibližně normální rozdělení, je pravděpodobnost překročení těchto 3σ -mezí rovna pouze 0.0027. V ostatních případech je třeba stanovit regulační meze ze znalosti rozdělení statistiky S_i . Pokud je do regulačních diagramů vynášeno více hodnot současně (neprovádí se přímo on-line kontrola), je pravděpodobnost P toho, že jedna nebo více hodnot padne mimo regulační meze, závislá na počtu vynášených bodů N . Tato pravděpodobnost se dá modelovat pomocí binomického rozdělení. Platí, že

$$P = 1 - (1 - p)^N,$$

kde p je pravděpodobnost, že hodnota jednoho znaku padne mimo toleranční meze, tj. $p = 0.0027$. Pro malá N pak přibližně platí, že

$$P \approx 1 - (1 - p)^N \approx pN.$$

Je patrné, že při vynášení více bodů současně se odpovídající pravděpodobnost P zvyšuje. Rovnice pro P se dá použít pro určení p , pro které bude dodržena pravděpodobnost $P = 0.0027$ při zadaném N , např. pro $N = 20$ je $p = 0.00014$ a odpovídající $100(1-p/2)$ %ní kvantil normálního rozdělení je 3.81. Pro tuto situaci je tedy vhodné konstruovat regulační meze $LCL = \bar{x}_s - 3.81s$ a $UCL = \bar{x}_s + 3.81s$.

Regulační diagramy jsou velmi populární pro svoji jednoduchost a lze je konstruovat pro téměř všechny typy parametrů jakosti (spojité, diskrétní, ordinální, nominální). Při praktickém sestrojování činí potíže zejména nerobustnost odhadů \bar{x}_s a s^2 , nenormalita rozdělení znaku jakosti a závislost mezi jednotlivými měřeními.

Kromě diagramů Shewhartova typu jsou používány také diagramy kumulativních součtů (CUSUM), pohyblivých charakteristik polohy a různé kombinace¹. Regulační diagramy se dají použít také pro případ simultánního sledování více znaků jakosti, Hotellingovy karty¹. Přejímací plány a regulační diagramy vycházejí z představy, že proces, který sledujeme, je

v ustáleném stavu, tj. jeho říditelné parametry (teplota, koncentrace, rychlost atd.) jsou na optimální úrovni.

Inženýrství jakosti umožňuje off-line nastavení podmínek výroby tak, aby bylo dosaženo jakostní výroby. Využívá se principů plánování experimentů s několika modifikacemi:

- (a) Parametry výroby se dělí na *ovladatelné* (ty se optimalizují) a *šumy* (omezuje se jejich variabilita).
- (b) Měřítkem kvality experimentu je *poměr* signálu a šumu S/N , který se maximalizuje.
- (c) Používá se plánů ve tvaru ortogonálních polí.

Poměr signálu a šumu S/N souvisí s definicí ztrátové funkce $L(x)$. Pro případ, kdy "nominální je nejlepší", lze použít vztahu

$$S/N \sim 10 \log(x_s^2/s^2) .$$

Inženýrství jakosti využívá obecně technik plánovaných experimentů pro modelování chování procesů výroby s ohledem na nalezení optima. Při klasickém plánování experimentů se vychází z náhodné veličiny y (vysvětlující proměnná), která je výstupem procesu. Ta je funkcí vektoru vstupních parametrů x . Účelem je stanovit vhodnou regresní funkci $f(x, \beta)$ a odhadnout její parametry β tak, aby bylo dosaženo jistých kritérií optimality. Tato kritéria úzce souvisí s přesností parametrů a rozptylem σ_y^2 veličiny y . Plánované experimenty se s výhodou používají také pro hledání optimálních podmínek minimalizujících tzv. ztrátovou funkci.

Taguchi²⁷ používá jako vstupní tzv. *parametry plánu* a hledá optimální podmínky, aby byly splněny tyto požadavky:

a) Minimální rozptyly některých znaků jakosti, jejichž střední hodnota je rovna cílové (požadované) hodnotě.

b) Minimalizace citlivosti výstupu na externí a interní fluktuace neřízených faktorů.

Pokud má náhodná veličina y rozdělení se střední hodnotou μ_y a rozptylem σ_y^2 , lze poměr signálu a šumu S/N určit ze vztahu

$$S/N \sim 10 \log(\mu_y^2/F_y^2) \sim 10 \log(v^2) ,$$

kde $v = \sigma_y / \mu_y$ je variační koeficient. Podle vlivu na poměr S/N lze jednotlivé vstupní parametry rozdělit do tří kategorií:

1) *Řídící faktory*, ovlivňující variabilitu procesu vyjádřenou S/N .

2) *Nastavované parametry*, které mají zanedbatelný vliv na S/N , ale významně ovlivňují střední hodnotu procesu.

3) *Šumové faktory*, které neovlivňují ani S/N , ani střední hodnotu procesu.

Faktory prvních dvou skupin patří do *parametrů plánu*.

Klasický postup návrhu experimentů vychází z minimalizace očekávané ztrátové funkce

$$M(x) \sim E[(y(x) - T)^2] \sim F_y^2(x) / (\mu_y(x) - T)^2 ,$$

kde jak střední hodnota $\mu_y(x)$, tak i rozptyl $F_y^2(x)$ jsou funkcí parametrů plánů a T je požadovaná cílová hodnota. Využitím technik plánovaných experimentů lze nalézt podmínky x_0 , pro které je proces nejbližší požadovanému stavu. Tedy $M(x_0)$ je minimem. Speciálně *Taguchiho přístup* předpokládá, že systém má jisté specifické vlastnosti vedoucí ke zjednodušení problému. Tyto vlastnosti plynou ze vztahu mezi σ_y a μ_y .

Předpokládá se, že závislost mezi σ_y a μ_y je takového druhu, že lze nalézt funkci $f[\mu_y(x)]$, pro kterou je $\sigma_y^2(x) / \{f[\mu_y(x)]\}^2$ měřítkem rozptylu $P(x)$. Funkce $P(x)$ závisí pouze na

podmnožině x_1 všech parametrů plánu $x = (x_1, x_2)$. Pak je $P(x_1)$ nezávislá na μ , protože pro daná x_1 je μ závislé pouze na parametrech x_2 . Funkce $P(x_1)$ nezávisí také na x_2 . Pak x_1 jsou řídicí parametry ovlivňující rozptyl a x_2 jsou parametry nastavované, jejichž změny neovlivní rozptyl. Na základě těchto předpokladů lze vyjádřit ztrátovou funkci ve tvaru

$$M(x) = (f[\mu(x)])^2 P(x_1) + (\mu(x) - T)^2$$

Pro zvolené μ je pak $M(x)$ minimální, pokud se nalezne x_{10} , pro které je $P(x_{10})$ minimální. Po nalezení x_{10} lze určit absolutní minimum $M(x)$ změnami parametrů x_2 . Pak je možno vyjádřit $\mu_y(x_0) = \mu_0$, pro které nabývá $M(x)$ minima ve tvaru

$$\mu_0 = T + f^2(\mu_0) P(x_1)$$

Speciální případ funkce $f(\mu)$ je obecná mocnina μ^α . Pak platí, že

$$P(x_1) = F_y^2 / \mu_y^2 \quad \text{a} \quad \mu_0 = T / (1 + v_0^2)$$

kde v_0 je variační koeficient v minimu. Uvažujme dva speciální případy:

a) $\alpha = 0$, pak σ_y není funkcí μ_y , $P(x_1) = \sigma^2(x_1)$ a $\mu_0 = T$. V tomto případě se $M(x)$ minimalizuje nejdříve nalezením x_{10} minimalizujícího $\sigma(x_1)$ a pak se nalezne x_{20} tak, aby $\mu_0 = T$.

b) $\alpha = 1$, pak $P(x_1) = v^2$ a minimalizace $P(x_1)$ je shodná s maximalizací poměru signálu a šumu S/N . Při hledání x_{20} se hledá μ_0 , pro které je $\mu_0 = T / (1 + v_0^2)$. Je známo, že pokud je směrodatná odchylka lineární funkcí střední hodnoty, vede logaritmická transformace y ke zkonstantnění rozptylu. Minimalizace směrodatné odchylky v *logaritmické transformaci* je tedy ekvivalentní maximalizaci S/N v původních proměnných. Je patrné, že Taguchiho postup je možné převést na standardní úlohu plánovaných experimentů ve vhodné (logaritmické) transformaci. Podrobnosti lze nalézt např. v cit.²⁷.

Inženýrství jakosti je logicky prvním stadiem celého procesu řízení jakosti, protože umožňuje stanovení podmínek, které je třeba při výrobě dodržovat. Následně se používají regulační diagramy pro monitorování jakosti. Tento přístup k řízení jakosti úzce souvisí s restrukturalizací informačních a řídicích informací (total quality management) a začíná se uplatňovat při zabezpečení jakosti ve vyspělých státech⁹.

10.2 Regulační diagramy

Regulační diagramy patří k základním nástrojům pro regulaci jakosti při výrobních procesech. Dají se však použít zcela obecně všude tam, kde jsou postupně v čase získávány informace o jakosti. Umožňují pro procesy, které jsou *statisticky regulovatelné* a měřený znak jakosti má stejné v čase neměnné rozdělení, modifikovat výrobní procesy tak, aby procento zmetků, kdy znak jakosti leží mimo předepsané meze, bylo dostatečně malé. Hodí se velmi dobře zejména pro monitorování procesů pomocí počítače. V této kapitole jsou uvedeny základní typy regulačních diagramů včetně způsobů jejich konstrukce a statistických vlastností.

V květnu 1924 navrhl W. A. Shewhart z Bell Telephone Laboratory první regulační diagram pro posouzení, zda variabilita sledovaného procesního parametru je způsobena náhodným kolísáním, nebo speciálními příčinami (seřízení strojů, změna surovin atd.). V

současné době představují v praxi nejrozšířenější typ, i když nejsou zdaleka univerzální. Typický regulační diagram Shewhartova typu je diagram, kde je znázorněna centrální linie, značící standardní, očekávanou cílovou hodnotu charakteristiky znaku jakosti a obě regulační meze, tj. dolní *LCL* a horní *UCL*. Tyto meze určují interval, ve kterém se s velkou pravděpodobností pohybují charakteristiky znaku jakosti, pokud je proces v *požadovaném stavu*. Pokud padnou hodnoty charakteristiky znaku jakosti mimo regulační meze, je proces "*mimo*" *požadovaný stav* a je třeba provést korekce (seřízení strojů, atd.).

Jako charakteristiky znaků jakosti se používají zejména *průměr* \bar{x}_s , *směrodatná odchylka* s , *variační rozpětí* R , *podíl nestandardních výrobků* P , *počet defektních výrobků* C a *počet defektů na výrobek* u . Speciální typy regulačních diagramů jako jsou *kumulativní součty* CUSUM, *pohyblivé průměry* MA atd. vycházejí z *metod analýzy časových řad* a konstrukčně se poněkud liší od regulačních diagramů Shewhartova typu. Konečně je možné sestrojovat také vícerozměrné Hotellingovy regulační diagramy pro více znaků jakosti současně¹.

10.2.1 Regulační diagramy pro dílčí výběry

V řadě případů je měření znaku jakosti jednoduché a rychlé, takže lze provést opakovaná měření, téměř ve stejném čase, a získat v různých časech několik výběrů V_1, \dots, V_M . Uvažujme pro jednoduchost, že pro každý výběr V_j stejné velikosti N jsou určeny výběrové průměry \bar{x}_{sj} a rozptyly s_j^2 . Předpokládejme, že rozdělení výběrových průměrů \bar{x}_s je normální $N(d, \sigma^2/N)$ a výběrové průměry jsou vzájemně nezávislé. Odhadem střední hodnoty d je pak generální průměr \bar{d}^* , definovaný vztahem

$$\bar{d}^* = \frac{1}{M} \sum_{j=1}^M \bar{x}_{sj},$$

a odhadem směrodatné odchylky σ je např. průměrná směrodatná odchylka

$$F = \frac{1}{M C_4} \sum_{j=1}^M s_j,$$

kde C_4 je konstanta zajišťující nevychýlenost. S využitím známého faktu, že

$$\frac{(N-1)s^2}{\sigma^2} \sim P^2(N-1),$$

kde symbol $P^2(\cdot)$ označuje rozdělení χ^2 -kvadrát, lze určit, že střední hodnota výběrové směrodatné odchylky je

$$E(s) = C_4 \sigma,$$

kde $C_4 = \sqrt{\frac{2}{N-1} \frac{\Gamma(N/2)}{\Gamma(N/2-1/2)}} = 1 + \frac{1}{4N-4}$,

kde $\Gamma(\cdot)$ je gamma funkce. Odpovídající rozptyl je pak roven $D(s) = F^2(1 + C_4^2)$. Pro

odhad směrodatné odchylky σ se používá vztah

$$F\% = \sqrt{\frac{\sum_{j=1}^M s_j^2}{M}}$$

Odhad σ^* však již není nevychýlený.

10.2.2 Regulační diagramy typu "x s pruhem"

Tyto regulační diagramy využívají k posuzování stavu sledovaného procesu aritmetické průměry \bar{x}_{Si} (resp. obecněji *parametr polohy*). Vyžadují ke své konstrukci buď znalost parametrů d , σ^2 normálního rozdělení, ze kterého data pocházejí, resp. pouze znalost vhodných odhadů d^* a σ^{*2} . Regulační meze se konstruuji tak, aby bylo zajištěno, že pravděpodobnost jejich překročení pro proces v požadovaném stavu je dostatečně malá. Z vlastností normálního rozdělení plyne, s jakou pravděpodobností se vyskytuje veličina x_i v mezích

$$d \pm \frac{KF}{\sqrt{N}}$$

Pro $K = 1.96$ to je pravděpodobnost 0.95 a pro $K = 3.09$ je to 0.999. V praxi se pro konstrukci regulačních mezí volí běžně hodnota $K = 3$, které odpovídá pravděpodobnost 0.9973. Obyčejně však nejsou parametry d a σ známy a nahrazují se svými odhady d^* a σ^* .

Regulační diagram "x s pruhem" má centrální linii d^* a regulační meze

$$L\hat{C}L = d^* - \frac{3F}{\sqrt{N}},$$

$$U\hat{C}L = d^* + \frac{3F}{\sqrt{N}}.$$

Takto definované regulační meze jsou odhady regulačních mezí LCL a UCL . Je zřejmé, že veličina d^* má normální rozdělení a veličina σ^2 (jako průměr M nezávislých proměnných) má také přibližně normální rozdělení. Pak $L\hat{C}L$ a $U\hat{C}L$ jako lineární kombinace d^* a σ^* mají přibližně normální rozdělení. Lze ukázat, že pro $U\hat{C}L$ platí

$$E(U\hat{C}L) = d^* + \frac{3F}{\sqrt{N}} = UCL$$

$$a \quad D(U\hat{C}L) = \frac{F^2}{MN} \left[1 + \frac{9(1 + C_4^2)}{C_4^2} \right] = \frac{K_1 F^2}{MN}$$

Obdobně lze snadno určit střední hodnotu a rozptyl pro $L\hat{C}L$. Je zřejmé, že pravděpodobnost p , s jakou x_s překročí regulační meze $L\hat{C}L$ a $U\hat{C}L$, je závislá na M a N . Podle uvedeného výkladu totiž platí, že

$$p = 2 \left[1 - F_N \left(\frac{3}{\sqrt{1 + \frac{1}{M} K_1}} \right) \right]$$

Pomocí tohoto vztahu lze pak snadno určit zkrácení, ke kterému dochází vlivem použití odhadů $\hat{\sigma}$ a $\hat{\sigma}^*$. Tak např. pro případ, že $M = 30$ a $N = 5$, vyjde $C_4 = 0.94$ a $p = 0.00378$. Je tedy zřejmé, že pro tento případ je pravděpodobnost P , s jakou se x_s vyskytuje v regulačních mezích, rovna $P = 1 - p = 0.9962$. Pro menší M a N může být pokles výraznější, což negativně ovlivní použití regulačních diagramů, když totiž vzroste četnost nesprávného rozhodnutí o zásahu do výroby.

Pravděpodobnost p je jednou z charakteristik kvality regulačních diagramů. Souvisí přímo s rizikem I. druhu, tj. pravděpodobností, že extrém vyvolaný náhodným kolísáním bude interpretován jako důsledek speciálních příčin. Je zřejmé, že čím je p vyšší, tím je riziko I. druhu menší.

Riziko II. druhu je pravděpodobnost jevu, že x_s bude v regulačních mezích přesto, že došlo ke změně úrovně procesu, a to v důsledku speciálních příčin. Toto riziko je přímo úměrné velikosti p a nepřímo úměrné velikosti změny úrovně procesu.

Pro účely návrhu a posouzení regulačních diagramů je vhodné sledovat počet hodnot v regulačním diagramu L , který je třeba k tomu, aby bylo indikováno překročení regulačních mezí. Pokud jsou výběrové průměry nezávislé, má veličina L geometrické rozdělení s pravděpodobnostní funkcí

$$P(L = i) = p(1 - p)^{i-1}, \quad i = 1, 2, 3, \dots$$

kde p je pravděpodobnost toho, že jeden výběrový průměr x_s překročí regulační meze.

Střední hodnota $E(L)$ se označuje jako ARL a je rovna

$$E(L) = ARL = 1/p$$

a pro rozptyl pak platí $D(L) = D(ARL) = \frac{1 - p}{p^2}$. Pro regulační diagramy "x s

pruhem" je v případě exaktních regulačních mezí LCL a UCL veličina $ARL = 370.37$ a $D(L) = 136803.84$. Pro případ přibližných regulačních mezí, počítaných z odhadů střední hodnoty a rozptylu, jsou však

$$e_j = x_{sj} + U\hat{C}L$$

$$a \quad e_k = x_{sk} + U\hat{C}L$$

vždy pozitivně korelované a pro korelační koeficient platí

$$D(e_j, e_k) = [1 - M(1 - K_1)^{&1}]^{&1}.$$

Nejjednodušším typem změny stavu procesu vlivem speciálních příčin je posun střední hodnoty d na velikost d_p . Pro tento případ je možno snadno určit, že pravděpodobnost, s jakou bude ležet x_s v mezích LCL a UCL , je rovna

$$p = F_N(A + 3) - F_N(A - 3),$$

kde $A = ((d - d_p) / \sigma) / \sigma$ je standardizovaný posun střední hodnoty. S využitím p lze snadno určit ARL . Tak např. pro $A = 1$ vyjde $ARL = 43.89$ a pro $A = 2$ je $ARL = 6.3$. Také pro složitější změny stavu procesu lze buď analyticky, nebo na základě simulací určit ARL . Při konstrukci regulačních diagramů "x s pruhem" se vychází z předpokladů:

- Rozdělení dat je alespoň přibližně normální.
- Velikosti výběrů jsou stejné.
- Měření jsou nezávislá.
- V datech nejsou vybočující měření (hrubé chyby).

Obecně je třeba, jak ve fázi konstrukce, tak i ve fázi použití regulačních diagramů, testovat předpoklady o datech, podobně jako při statistické analýze jednorozměrných výběrů¹⁶.

Nonnormalita. V řadě případů má sledovaný parametr sešikmené rozdělení (např. pro cirkularitu, pevnost, koncentrace ve stopové analýze atd.). Pro menší a střední šikmosti se nonnormalita výrazně neprojeví, pokud je počet prvků v jednotlivých výběrech $N \geq 5$. Pro větší šikmosti je možno použít následujících technik:

1. Nalézt vhodnou normalizační transformaci (např. ve třídě Boxových-Coxových mocninných transformací¹⁶) a realizovat regulační diagramy v transformovaných proměnných. Jednoduché empirické pravidlo doporučuje použití logaritmické transformace dat, pokud jsou prvky ve výběrech řádově rozdílné.

2. Nalézt vhodnou teoretickou hustotu pravděpodobnosti, resp. distribuční funkci F_T a určit meze LCL a UCL tak, aby platilo

$$F_T(LCL) = 1 - p/2 \text{ a } F_T(UCL) = p/2,$$

kde standardně $p = 0.0027$. Pro hledání F_T je možné použít jak teoretických úvah, tak i celé řady exploratorních metod (např. $Q-Q$ grafy, atd.)¹⁶.

3. Použít heuristické techniky, vycházející z pravděpodobnosti P_x , že výběrové průměry x_{sj} leží pod generálním průměrem d^* . Tato pravděpodobnost se dá odhadnout z počtu výběrových průměrů x_s , ležících pod d^*

$$\hat{P}_x = \frac{1}{M} \sum_j I_{[x_{sj}]}$$

$$\text{kde indikátorová funkce } \begin{cases} I_{[x_{sj}]} = 1 & \text{pro } x_{sj} < d^* \\ I_{[x_{sj}]} = 0 & \text{pro } x_{sj} \geq d^* \end{cases}.$$

Pro rozdělení sešikmené k vyšším hodnotám je pak

$$UCL = \hat{d} + 3F \sqrt{\frac{2P_x}{N}}$$

a

$$LCL = \hat{d} - 3F \sqrt{\frac{2(1 - P_x)}{N}}$$

Meze UCL a LCL byly sestaveny na základě náhrady původní hustoty pravděpodobnosti dvěma segmenty v místě d . Každý segment je použit pro sestavení symetrické hustoty pravděpodobnosti¹⁸.

Nestejně velikosti výběrů. Pokud se počet prvků výběru N mění, je třeba pouze upravovat meze, definované rovnicí pro LCL a UCL . Regulační meze jsou pak tvořeny po částech konstantními úseky.

Autokorelace. Vlivem autokorelace dochází ke zkreslení regulačních diagramů "x s pruhem". Např. v případě pozitivní autokorelace roste počet případů, překračujících regulační meze, i když je proces v požadovaném stavu (falešný poplach). Pro případ autokorelace prvního řádu s autokorelačním koeficientem ρ platí pro střední hodnotu výběrového rozptylu vztah

$$E(s^2) = \frac{F^2}{(1 - 2\rho/N)}$$

Při použití s^2 tedy v případě pozitivní autokorelace vyjde rozptyl střední hodnoty podhodnocený, totiž nesprávně menší. Někteří autoři doporučují pro omezení vlivu autokorelace zvětšit regulační meze o faktor $1/\sqrt{1 - \rho}$. Obecně však nejsou vhodné, pokud jsou měření silně korelována. V některých případech lze zdroje autokorelace indikovat a odstranit.

Vybočující měření. Přítomnost vybočujících měření obecně zkresluje odhady \hat{d} a $\hat{\sigma}^*$ a vede k rozšiřování regulačních mezí. K velmi dobrým robustním výsledkům vede náhrada výběrových rozptylů interkvartilovým rozmezím

$$IQR = x_{(b)} - x_{(a)},$$

kde $a = \text{int}[N/4] + 1$ a $b = N - a + 1$. Zde symbol $\text{int}[x]$ označuje celočíselnou část čísla x a $x_{(i)}$ je i -tá pořádková statistika. Obecně se může zrobustňovat celá řada parametrů. Místo aritmetických průměrů \bar{x}_y lze použít robustní odhady polohy G_j , místo výběrových rozptylů lze použít robustní odhady rozptýlení s_{Rj} a místo aritmetických průměrů \hat{d} a $\hat{\sigma}^*$ lze použít robustní charakteristiky polohy $T(G)$, $T(s_R)$. Pro regulační meze lze pak použít vztahy

$$UCL = T(G) + 3 T(s_R) D(G) / E(T(s_R)),$$

$$LCL = T(G) - 3 T(s_R) D(G) / E(T(s_R)),$$

kde $E(.)$ a $D(.)$ označují střední hodnotu a rozptyl. Na základě detailnějšího testování bylo zjištěno, že dobré vlastnosti mají *mediánové regulační diagramy*, pro které jsou G_j mediány, s_{Rj} jsou interkvartilová rozmezí a $T(.)$ je aritmetický průměr. Místo mediánu lze také použít uřezaný průměr a stupeň uřezání je 0.25, cit.¹⁶. V případě, že data pocházejí z normálního rozdělení, jsou robustní regulační diagramy méně efektivní. To znamená, že regulační meze jsou široké. Tak např. pro mediánové diagramy a $N = 5$ je směrodatná odchylka o faktor 1.2 větší, tj. meze jsou o 20 % širší.

Klasické regulační diagramy "x s pruhem" jsou také málo citlivé na malé systematické změny střední hodnoty d (trend). Pro tyto účely se konstruuji ještě výstražné meze ve vzdálenostech $\pm 2\sigma^*/\%N$ a $\pm \sigma^*/\%N$ od generálního průměru d^* . Pro indikaci trendu se pak používá celá řada heuristických pravidel, např.:

1. Jedna hodnota x_s leží mimo regulační meze (3σ).
2. Dva ze tří po sobě následujících bodů leží mimo výstražné meze (2σ) na stejné straně od d^* .
3. Čtyři z pěti po sobě následujících hodnot x_{s_j} leží mimo výstražné meze (1σ) na stejné straně od d^* .
4. Osm po sobě následujících bodů leží na stejné straně od průměru.

Existuje ještě celá řada dalších heuristických pravidel, která mohou být pro speciální případy užitečná. Obvykle se do regulačních diagramů Shewhartova typu vynášejí x_{s_i} po konstantních časových intervalech. S ohledem na rychlost odhalení systematických změn je často výhodné použít také nekonstantní časové intervaly. Při konstrukci regulačních diagramů je třeba důkladně ověřit, zda použité výběry V_1, \dots, V_M netvoří nenáhodná seskupení či trendy. Podrobnosti o konstrukci a ověřování regulačních "diagramů x s pruhem" lze nalézt v ISO 8258.

10.2.3 Regulační diagramy pro posouzení variability

Tyto regulační diagramy umožňují posouzení úrovně variability procesu. Vycházejí opět z předpokladu normality nezávislosti výběru a konstantnosti rozptylů.

Regulační diagram "s" má centrální linii $s_p = \sigma^* C_4$ a regulační meze

$$LCL = s_p \sqrt{\frac{P_{0.001}^2(N-1)}{(N-1)}},$$

$$UCL = s_p \sqrt{\frac{P_{0.999}^2(N-1)}{(N-1)}},$$

kde $\chi_v^2(N-1)$ je 100 v %ní kvantil χ^2 -rozdělení s $N-1$ stupni volnosti. Volba $v = 0.001$ a $v = 0.999$ zajišťuje, že v regulačních mezích bude ležet (pokud je proces v požadovaném stavu) 99.8 % všech výběrových směrodatných odchylek.

V předpočítačové éře se často pro vyjádření variability používalo místo směrodatné odchylky s variační rozpětí R dat

$$R = x_{(N)} - x_{(1)},$$

kde $x_{(n)}$ je nejvyšší a $x_{(1)}$ nejmenší prvek výběru. Při konstrukci *regulačního diagramu "R"* lze pak použít průměrné rozpětí R_p z výběrů V_1, \dots, V_M . Pro regulační meze platí

$$LCL = D_3 R_p,$$

$$UCL = D_4 R_p.$$

Hodnoty D_3, D_4 souvisejí pouze s rozsahem výběru a jsou tabelovány např. v cit.¹.

Pro $N = 3$ je $D_3 = 0, D_4 = 2.575$, pro $N = 5$ je $D_3 = 0, D_4 = 2.115$ a pro $N = 10$ je $D_3 = 0.233, D_4 = 1.773$.

Je zřejmé, že snahou výrobců je snižování variability výroby. Z tohoto pohledu ztrácejí dolní meze pro regulační diagramy "s" a regulační diagramy "R" smysl a často se definuje pouze horní regulační mez. Při korektní konstrukci meze *UCL* pro regulační diagram "s" je nutné použít takový 100v% kvantil χ^2 -rozdělení, který odpovídá požadované pravděpodobnosti s jakou nemají výběrové směrodatné odchylky překročit regulační mez. Tedy pro případ, že minimálně 99.8 % výběrových směrodatných odchylek má ležet pod *UCL* se volí $v = 0.998$. Je zřejmé, že takto určené *UCL* bude nižší než *UCL*.

10.2.4 Regulační diagramy kumulativních součtů, CUSUM

Je zřejmé, že Shewhartovy diagramy jsou vlastně ekvivalentní opakovaným testům významnosti pro konstantní velikosti výběru. Jejich základní výhodou je jednoduchost a rychlá indikace velkých změn stavu procesu. Na střední a malé změny (resp. trendy) reagují pomalu. Základním omezením je předpoklad, že jednotlivé dílčí výběry jsou nezávislé. Pokud je účelem z co nejmenšího počtu výběrů zachytit nenáhodný trend (indikovaný např. posunem střední hodnoty), je vhodné použít regulačních diagramů typu CUSUM.

Page v roce 1954 navrhl konstrukci regulačních diagramů, založenou na součtu všech hodnot, takže *i*-tá hodnota v grafu obsahuje všechny "historické" informace, obsažené v hodnotě (*i* - 1) a změnu způsobenou přechodem ze stavu (*i* - 1) do stavu (*i*). Tyto regulační diagramy slouží tedy obvykle k rychlé detekci malých a středních změn průměrné úrovně znaku jakosti *x*. Ten má střední hodnotu $E(x) = d$ a rozptyl $D(x) = \sigma^2$. Nechť d_A označuje úroveň *přijatelné jakosti* (acceptable quality level, *AQL*) a d_R úroveň *nepřijatelné jakosti* (rejectable quality level, *RQL*). Veličina *AQL* se také často označuje jako *požadovaná, cílová hodnota*. Veličina $D = d_A - d_R$ pak určuje velikost změny průměrné úrovně znaku jakosti, kterou je nutno detekovat. Při klasickém postupu navrženém Pageem se v konstantních časových intervalech $i = 1, 2, \dots$ získávají výběry V_i a počítají se vhodné charakteristiky T_i (obvykle aritmetické průměry x_{Si}) a kumulativní součty těchto statistik se vynášejí proti indexu *i*.

Výpočet kumulativních součtů závisí na tom, zda je účelem detekovat systematickou změnu v jednom směru (pozitivní nebo negativní odchylky) nebo v obou směrech. Běžně se při výpočtu kumulativních součtů odečítá jistá *referenční hodnota K* a kumulativní součet po *j*-tém časovém intervalu (z výběru V_1, \dots, V_j) je roven

$$S_j = \sum_{i=1}^j (x_{S,i} - K).$$

Referenční hodnota K se volí s ohledem na to, jaké systematické odchylky se mají detekovat. Obyčejně se volí $K = (d_A + d_R)/2$. Regulační diagram CUSUM je pak graf S_j v závislosti na j . Pro detekci kladné odchylky ($d_{R1} > d_A$) se volí referenční hodnota $K_1 = (d_A + d_{R1})/2$ a podobně pro detekci záporné odchylky ($d_{R2} < d_A$) se volí $K_2 = (d_A + d_{R2})/2$. Zde d_{R1} a d_{R2} jsou úrovně nepřijatelné jakosti ve směru kladných a záporných odchylek. Kladná odchylka od průměrné úrovně je pak detekována, pokud

$$S_r \geq \min(S_j) \geq h^+ ,$$

a záporná odchylka je detekována, pokud

$$\max(S_j) \leq S_r \leq h^- .$$

Veličiny h^+ a h^- jsou tzv. *rozhodné meze*. Při výpočtu S_j a S_r se používá referenční hodnota K_1 a při výpočtu charakteristik se používá referenční hodnota K_2 . Tato verze diagramu CUSUM je vhodná zejména tam, kde je účelem pouze indikovat změnu úrovně. Vlastní postup je velmi blízký Waldově sekvenčnímu testu podílu věrohodností. Barnardova metoda konstrukce regulačních diagramů CUSUM spočívá v náhradě referenční hodnoty přímo veličinou d_A , takže se vynesou kumulativní součty

$$S_j = \sum_{i=1}^j (x_{Si} - d_A)$$

proti indexu j . Pokud zůstává proces na úrovni d_A , kolísají hodnoty S_j kolem nuly. Posun střední hodnoty se projeví trendem v diagramu. Pro jeho indikaci se často konstruuje tzv. *V-masky*, tj. výseč ve tvaru ">" umístěná u posledního vneseného bodu (S_r, r). Pokud nějaká S_j ($j < r$) leží mimo tuto výseč, došlo k posunu střední hodnoty u bodu (S_r, r). Pro použití V-masky je nezbytné zvolit vhodné měřítko na ose y kumulativních součtů. Obyčejně se předpokládá, že jednotková vzdálenost na této ose je rovna 2σ . V případě neznalosti směrodatné odchylky se využívá jejího odhadu F_e z rozptylu výběrových průměrů

$$F_e = \sqrt{\frac{1}{M+1} \sum_{i=1}^M (x_{Si} - d_A)^2} .$$

Pokud nejsou výrazné odchylky mezi výběrovými průměry a v datech je pouze náhodné kolísání σ , volí se $F_e = F / \sqrt{N}$. Konstrukce *V-masky* vyžaduje znalost vzdálenosti jejího vrcholu od posledního vneseného bodu w a úhlu θ , který svírají ramena této masky s její horizontální osou. Parametry V-masky úzce souvisí s posuzováním odchylek od standardního stavu s využitím referenčních hodnot K_1, K_2 a rozhodných mezí h^- a h^+ . Pro případ symetrických úrovní nepřijatelné jakosti lze psát

$$d_{R1} = d_A + F ,$$

$$d_{R2} = d_A - F ,$$

kde δ je standardizovaný posun střední hodnoty, který má být detekován. Je zřejmé, že velikost celkové změny procesu, která se má detekovat při simultánním sledování je

$$D = K_1 \& K_2 = (d_{R1} \& d_{R2}) / 2 = F^*$$

V případě neznalosti σ se používá jeho odhad, definovaný rovnicí pro F_e . Obyčejně se ještě zvolí $h^* = h$, tj. jedna rozhodná mez. Pak pro vzdálenost vrcholu V-masky od posledního bodu platí

$$w = \frac{h}{2 F \operatorname{tg} 1}$$

a pro úhel θ je možno psát

$$\operatorname{tg} 2 = \frac{D}{2 F} = \frac{*}{2}$$

Z těchto rovnic je zřejmé, že existuje úzká souvislost mezi referenčními hodnotami K_1 a K_2 , rozhodnou mezí h a parametry (w, θ) , charakterizujícími V-masku. Snadno lze určit, že rozhodná mez

$$h = w 2 F \operatorname{tg} 2$$

je vlastně průsečíkem a referenční hodnota $K = \delta \sigma$ je směrnice přímky v souřadném systému, jehož počátek je posunut do bodu (S, r) . Druhá přímka má úsek $-h$ a směrnici $-K$ a indikuje negativní odchylky. Existuje ještě celá řada různých variant V-masky (uříznutá, semiparabolická atp.), které vedou ke snížení průměrného počtu dat do indikace odchylky od referenčního stavu, tj. *ARL*.

Standardně začínají regulační diagramy CUSUM od nuly, tj. $S_0 = 0$. Modifikace FIR (fast initial response), která je citlivá na počáteční změny stavu procesu ve zvoleném směru, začíná s nenulovou hodnotou $S_0 = h / 2$. Zrobustnění lze docílit např. tím, že se hodnoty x_{sj} , převyšující specifikované meze, do výpočtu S_j nezahrnují. Obyčejně se x_{sj} považuje za silně vybočující, pokud

$$\frac{*x_{sj} \& d^{(*)}}{F} > 4$$

Velmi dobré se jeví pravidlo *dvou extrémů*, kdy se x_{sj} do výpočtu nezahrnuje, ale pokud vyjde totéž i pro x_{sj+1} , je indikována výrazná změna stavu procesu²⁰.

Odpovídající výpočet průměrného počtu hodnot $E(L)$ do indikace změny stavu procesu, tj. *ARL*, je dosti komplikovaný. Pokud je proces ve stavu se střední hodnotou $d_i(AQL)$ a měření mají normální rozdělení, může být rozdělení L považováno za geometrické. To však již neplatí pro změnu stavu, kdy má střední hodnota úroveň $d_r(RQL)$. Obecně je pro případ, kdy $S_0 = 0$,

$$ARL = \frac{N(0)}{1 \& P(0)}$$

kde $N(0)$ a $P(0)$ jsou Fredholmovy integrální rovnice druhého druhu, které je nutné pro zadané rozdělení veličiny x řešit numericky¹⁹.

Moderní regulační diagramy typu CUSUM využívají kumulativních součtů standardizovaných odchylek od generálního průměru d^* . Předpokládejme, že známe odhady d^* a σ^* . Pak můžeme pro výběrový průměr x_{sj} sestavit normalizovanou náhodnou veličinu

$$z_j = \frac{(x_{sj} - d^*) \sqrt{N}}{\sigma^*}$$

a určit dvojici kumulativních součtů

$$S_{H,j} = \max[0; (Z_j + K) \% S_{H,j-1}] ,$$

$$S_{L,j} = \max[0; (-Z_j + K) \% S_{L,j-1}] .$$

Hodnota K určuje polovinu průměrného posunu v Z -transformaci, která se má detekovat. Obvykle se volí $K = 0.5$, což odpovídá posunu o jedno σ . Začíná se od $S_{H,0} = S_{L,0} = 0$. Suma S_H slouží k detekci pozitivního posunu a suma S_L k detekci negativního posunu střední hodnoty. Obě sumy jsou kladné. Pokud $S_{H,j}$, resp. $S_{L,j}$ překročí rozhodnou mez h_j , došlo k indikaci nenáhodného posunu střední hodnoty. Běžně se volí $h_j = 4$ nebo 5 , i když se dá dokázat, že pro malé j by mělo být h_j menší. Přesnější funkce $h_j = f(j)$ lze nalézt v literatuře².

Zajímavou možností je kombinace CUSUM a regulačního diagramu "x s pruhem". Zde se počítají hodnoty $Z_j, S_{H,j}, S_{L,j}$. Indikací posunu střední hodnoty je:

- a) $Z_j > 3$, resp. $Z_j < -3$,
- b) $S_{H,j}$, resp. $S_{L,j} > h_j$.

Obecně má použití regulačních diagramů typu CUSUM následující přednosti:

1. Vyšší efektivnost oproti Shewhartovým diagramům pro případ standardizovaného posunu střední hodnoty $d_A - d_R = \delta \sigma$, $\delta \geq 0$ ($0.5 \div 2$).

2. Snadnou vizuální detekci posunu velikosti střední hodnoty z průměrné směrnice vynášených bodů.

3. Snadné určení místa, kde došlo k posunu střední hodnoty.

4. Vhodné pro případy, kdy náklady na získání experimentálních údajů jsou vysoké.

Na druhé straně lze nalézt také některé nevýhody. Mezi základní patří, že je třeba pro každý případ konstruovat *speciální V-masku*. Také regulační diagramy CUSUM vycházejí z celé řady předpokladů, jejichž nesplnění vede ke ztrátě efektivnosti. Rozhodující je především dobrý odhad σ . Pokud je tento odhad podhodnocený, snižuje se ARL , a pokud je nadhodnocený, ARL roste. Výrazný vliv má také nenormalita rozdělení veličiny x . Na RQL se efekt nenormality výrazně neprojeví. Kladná šikmost však výrazně sníží ARL na AQL . Negativní šikmost zvyšuje ARL . Výrazně zkreslující vliv má také autokorelace mezi původními daty. V případě kladné autokorelace se snižuje ARL a v případě záporné autokorelace se zvyšuje. Tedy ani diagramy CUSUM nejsou robustní vůči narušení předpokladů normality a nezávislosti. Podobně jako u Shewhartových diagramů lze i zde využít různých technik pro zlepšení, resp. eliminaci, těchto narušení, příp. modifikovat vlastní proces sestavení regulačních diagramů. Regulační diagramy CUSUM se dají použít také pro posouzení velikosti rozptýlení znaku jakosti x . Problémy činí zejména volba vhodné standardizace (modifikace rovnice pro z_j), která by zajistila alespoň přibližnou normalitu. Pro případ konstrukce diagramu CUSUM pro směrodatnou odchylku s se volí

$$Z_i = \left(s / F \right)^{0.625} , \text{ což zajišťuje přibližnou normalitu } Z \text{ pro velikosti výběrů } N = 3 \div 20.$$

10.2.5 Regulační diagramy na bázi lokálního vyhlazení

Existuje celá řada regulačních diagramů, využívajících různých typů vyhlazování. Mezi jednoduché patří *exponenciálně vážený pohyblivý průměr* (EWMA), označovaný také jako pohyblivý geometrický průměr

$$W_j = r x_{sj} + (1-r) W_{j-1},$$

kde pro $j = 0$ je $W_0 = d^*$ a r ($0 < r < 1$) je parametr definující váhu, obvykle se doporučuje $r = 0.25$. Rozptyl tohoto parametru je roven⁴

$$S_{W_j}^2 = \frac{F^2}{N} \frac{r}{2+r} (1 + (1-r)^{2j}).$$

Pro $r > 0.2$ a $j \geq 5$ lze použít aproximaci

$$S_{W_j}^2 \approx \frac{F^2}{N} \frac{r}{2+r}.$$

Regulační meze pro pohyblivý geometrický průměr pak jsou

$$LCL = d^* - 3 \sqrt{\frac{r F^2}{N(2+r)}},$$

$$UCL = d^* + 3 \sqrt{\frac{r F^2}{N(2+r)}}.$$

Regulační diagram pohyblivého geometrického průměru obsahuje pouze regulační meze LCL a UCL . Vynášejí se do ní hodnoty W_j . Tento regulační diagram je výhodný pro případy, kdy je třeba detekovat s vysokou přesností malé změny stavu procesu. Je opět nerobustní vůči odchylkám od normality a nezávislosti. Podobně jako u diagramů CUSUM, lze provést snadno kombinaci se Shewhartovými regulačními mezemi a provést zrobustnění s využitím identifikace odlehlých hodnot. Za jistých podmínek lze použít EWMA pro případy, kdy jsou data autokorelována. Tak např. pro proces popsany schématem

$$x_t = x_{t-1} + g + 2g_{t-1}$$

je predikce stavu $\hat{x}_{t+1}(t)$ v čase t (tzv. jednokroková predikce) rovná vztahu

$$W_t = \hat{x}_{t+1}(t).$$

Odpovídající chyby predikce $e_t = x_t - W_{t-1}$ jsou pak nezávislé náhodné veličiny s

$E(e_t) = 0$ a $D(e_t) = \sigma_p^2$. Regulační diagram pro tento případ má regulační meze

$$UCL = 3 \sigma_p,$$

$$\hat{LCL} = \bar{x} - 3F_p$$

Pro odhad směrodatné odchylky chyby jedнокrokové predikce σ_p lze použít průměrné absolutní odchylky $\Delta(t)$, počítané ze vztahu

$$\Delta(t) = \frac{1}{n} \sum_{i=1}^n |e_i^*| \quad (t \geq 1),$$

kde α je váhový parametr. Protože pro normální rozdělení je směrodatná odchylka o 25 % vyšší než $\Delta(t)$, odhaduje se F_p ze vztahu

$$F_p(t) = 1.25 \Delta(t)$$

Další možností je přímý výpočet vyhlazeného odhadu rozptylu chyby jedнокrokové predikce

$$F_p^2(t) = \frac{1}{n} \sum_{i=1}^n e_i^{*2} \quad (t \geq 1)$$

Doporučuje se použít rozmezí $0.03 \leq \alpha \leq 0.1$ s tím, že nižší hodnoty jsou vhodnější. Standardně se doporučuje $\alpha = 0.05$. Další možností je použití přímo hodnot W_t s tím, že se postupně upravují regulační meze s ohledem na to, že jde o jedнокrokovou predikci

$$U\hat{C}L_{t\%} = \bar{x}_t + 3F_p$$

$$L\hat{C}L_{t\%} = \bar{x}_t - 3F_p$$

Centrální linie tohoto diagramu pro časovou periodu $(t+1)$ je W_t a vynáší se do něho přímo x_{t+1} (příp. x_{St+1}). Pokud lze očekávat, že data tvoří spíše časové řady, je použití těchto modifikací regulačních diagramů EWMA výhodné pro posouzení dynamiky dějů²¹. Uvedené vztahy lze použít buď pro průměrné hodnoty x_{St} v čase t , resp. přímo pro jednotlivá pozorování x_t v čase t . Regulační diagramy EWMA se dají také snadno modifikovat pro neparametrické charakteristiky polohy, resp. variability, kdy je zajištěna necitlivost na typ rozdělení původních dat. Jednoduché je použití standardizovaného pořadí, kdy se jednotlivé hodnoty x_t nahrazují odpovídajícími pořadími R_t . Lze použít také Wilcoxonova pořadí se znaménky. Pro hodnoty $x_{t-g+1}, \dots, x_{t-1}, x_t$ je pořadí definováno vztahem

$$R_t^{(i)} = 1 + j_{(i)} I_{[x_t > x_j]}$$

kde indikátorová funkce $I_{[x_t > x_j]} = 1$ pro $x_t > x_j$ a v ostatních případech je nulová. Standardizované pořadí je pak

$$R_t = \frac{2}{g} \left(R_t^{(i)} - \frac{g+1}{2} \right)$$

Lze ukázat, že $E(R_t) = 0$ a $D(R_t) = \frac{g^2 + 1}{3g^2}$ pro všechna t .

Využitím rovnice pro W_t je pak možno definovat exponenciálně vážený průměr pořadí

$$T_t = (1 - r) T_{t-1} + r R_t,$$

kde pro $t = 0$ je $T_0 = 0$ a $0 < r < 1$ je parametr vyhlazení. Doporučená hodnota je $r = 0.3$. Pro případ větších g (obvykle $g \geq 30$) a r nepříliš blízkého k jedné je rozdělení T_t přibližně normální se střední hodnotou $E(T_t) = 0$ a rozptylem

$$D(T_t) = \frac{r(g^2 + 1)}{3g^2(2 + r)}.$$

Regulační meze pro tento typ regulačních diagramů jsou

$$LCL = -3 \sqrt{\frac{r(g^2 + 1)}{3g^2(2 + r)}},$$

$$UCL = 3 \sqrt{\frac{r(g^2 + 1)}{3g^2(2 + r)}}$$

a vynášejí se hodnoty T_t .

10.2.6 Regulační diagramy pro jednotlivé hodnoty

V některých případech není možné provést ve stejném čase N -tici měření pro sestavení výběru V . V řadě případů je také kolísání charakteristik sledovaného procesu příliš rychlé ve srovnání s měřením, takže "průměrování" postrádá smysl. Pak se konstruují regulační diagramy pro jednotlivá měření. Ty mají celou řadu nevýhod:

- Jsou citlivé na nenormalitu rozdělení znaku x (rozdělení průměru \bar{x}_s se více blíží normálnímu než rozdělení původních hodnot).
- Jsou málo citlivé na posun střední hodnoty.
- Jsou negativně ovlivněny trendy v datech.
- Jsou citlivé na velikosti výběru ze kterých se odhadují parametry rozdělení.

Vyděme z předpokladu, že znak jakosti x má normální rozdělení $N(d, \sigma^2)$. Parametry d a σ^2 se odhadují z výběru velikosti N . Vzhledem k tomu, že jde pouze o jeden výběr, je třeba aby $N \geq 50$ a před vlastní analýzou bylo provedeno ověření normality, resp. identifikace vybočujících měření.

Výhodné je použití $Q-Q$ grafů, kdy lze simultánně ověřit jak normalitu, tak i přítomnost vybočujících měření. Principem je hodnocení linearitu v grafu $x_{(i)}$ vs. $F_N^{-1}(P_i)$. Symbolem $x_{(i)}$ jsou označeny pořádkové statistiky, čili seřazené prvky výběru, a $P_i = i/(N + 1)$ je pořadová pravděpodobnost. Parametr d se odhaduje jako aritmetický průměr \bar{x}_s a odhadem parametru σ^2 je výběrový rozptyl s^2 . Vzhledem k tomu, že odpovídající směrodatná odchylka s je vychýleným odhadem, používá se místo ní nevychýlený odhad

$$s^* = s / C_4.$$

Regulační diagram "x" má centrální linii \bar{x}_s a regulační meze

$$LCL = \bar{x}_s - 3s$$

$$UCL = \bar{x}_s + 3s$$

Do tohoto grafu se vynášejí přímo naměřené hodnoty znaku x . Místo výběrové směrodatné odchylky se v praxi s oblibou používá *průměrného pohyblivého rozpětí MR*, definovaného vztahem

$$MR = \frac{1}{N+1} \sum_{i=2}^N (x_i - x_{i-1})$$

Pomocí MR lze definovat odhad směrodatné odchylky pro případ normálního rozdělení

$$F = \frac{\sqrt{B} MR}{2} = 0.8865 MR$$

Regulační meze jsou pak definovány vztahy

$$U\hat{C}L = \bar{x}_s + 3F = \bar{x}_s + 2.6595 MR$$

$$L\hat{C}L = \bar{x}_s - 3F = \bar{x}_s - 2.6595 MR$$

Doporučená velikost výběru je $N = 300$. Podobně jako u regulačních diagramů "x s pruhem", lze i zde v případě nenormality použít normalizační transformace a v případě vybočujících měření robustních odhadů. Pro jednotlivé hodnoty lze také snadno aplikovat diagramy CUSUM. Stačí pouze počítat veličinu Z_j ze vztahu

$$Z_j = \frac{x_j - \bar{x}_s}{s}$$

kde x_j je j -tá naměřená hodnota. Hawkins³ navrhuje pro regulaci variability na základě jednotlivých pozorování CUSUM regulační diagram. Pro tento diagram se volí Z_j ve tvaru

$$Z_j = \frac{\sqrt{(x_j - x_{j-1})^2 + F^2} + 0.82218}{0.34914}$$

Tato rovnice je motivována faktem, že $E \left(\frac{\sqrt{(x_j - x_{j-1})^2 + F^2}}{F} \right) = 0.82218$ a odpovídající

rozptyl je roven $(0.34914)^2$, pokud má znak jakosti x rozdělení $N(0, \sigma^2)$. Velmi snadné je také použití různých typů EWMA regulačních karet pro jednotlivé hodnoty.

10.2.7 Regulační diagramy pro distrétní znaky

V řadě případů lze sledovaný znak jakosti rozdělit pouze do dvou kategorií: *vyhovující* a *nevyhovující* (zmetek). V různých časech lze získat z výběru velikosti N celkový počet x (nevyhovujících). Podíl nevyhovujících výrobků je pak zřejmě

$$P_N^{(c)} = \frac{x}{N}.$$

Tento podíl je odhadem pravděpodobnosti výskytu nevyhovujících výrobků P_N . Je však třeba použít dostatečně vysoké N (obvykle $N > 500$). Lze ukázat, že pro P_N nepříliš vzdálené od 0.5 a střední N mají veličiny

$$Z_1 = \frac{x - NP_N}{\sqrt{NP_N(1 - P_N)}}$$

a

$$Z_2 = \frac{P_N^{(c)} - P_N}{\sqrt{P_N(1 - P_N)/N}}$$

přibližně normované normální rozdělení $N(0, 1)$.

Regulační diagramy "np" pro počet nevyhovujících jednotek mají centrální linii NP_N^* a regulační meze

$$LCL = NP_N^{(c)} - 3\sqrt{NP_N^{(c)}(1 - P_N^{(c)})},$$

$$UCL = NP_N^{(c)} + 3\sqrt{NP_N^{(c)}(1 - P_N^{(c)})}.$$

Vynáší se do nich počet nevyhovujících výrobků určený z výběrů velikosti N .

Regulační diagramy "p" pro podíl nevyhovujících výrobků mají centrální linii P_N^* a regulační meze

$$LCL = P_N^{(c)} - 3\sqrt{P_N^{(c)}(1 - P_N^{(c)})/N},$$

$$UCL = P_N^{(c)} + 3\sqrt{P_N^{(c)}(1 - P_N^{(c)})/N}.$$

Problém při použití těchto diagramů spočívá v tom, že rovnice pro Z_1 a Z_2 platí velmi přibližně, zejména pro malá N a P_N vzdálená od 0.5. Pro malé výběry je možné použít korekční faktor $(2N)^{-1}$ a pak

$$LCL = P_N^{(c)} - 3\sqrt{P_N^{(c)}(1 - P_N^{(c)})/N + 1/(2N)},$$

$$UCL = P_N^{(c)} + 3\sqrt{P_N^{(c)}(1 - P_N^{(c)})/N + 1/(2N)}.$$

Přiblížení k normalitě lze docílit například použitím arkussinové transformace

$$A(P_N) = \arcsin \sqrt{\frac{P_N \pm 3/8}{N \pm 3/4}},$$

pro kterou platí, že střední hodnota

$$E(A(P_N)) = \arcsin \sqrt{P_N}$$

a rozptyl

$$D(A(P_N)) = 1/4N.$$

Lze tedy sestavit regulační diagram p , do kterého se vynášejí $A(P_N)$ se střední linií arcsin $[\sqrt{P_N}]$ a regulačními mezemi

$$LCL = \arcsin(\sqrt{P_N}) - 1.5\sqrt{N},$$

$$UCL = \arcsin(\sqrt{P_N}) + 1.5\sqrt{N}.$$

Také pro počet, resp. podíl, nevyhovujících jednotek lze použít diagramy kumulativních součtů CUSUM. V některých případech je regulovanou veličinou počet vad výrobku C . Při konstrukci regulačního diagramu se vychází z předpokladu, že parametr C má Poissonovo rozdělení. Velikost C se v podstatě odhaduje jako průměrný počet vad C_p určený z N výrobků.

Regulační diagramy "c" mají centrální linii C_p a regulační meze jsou

$$LCL = C_p - 3\sqrt{C_p},$$

$$UCL = C_p + 3\sqrt{C_p}.$$

Tyto limity vycházejí opět z aproximace Poissonova rozdělení normálním. Zlepšení aproximace lze docílit vhodnou transformací. Jednoduchá je transformace

$$B(C) = \sqrt{C} + \sqrt{C+1},$$

kdy toleranční meze jsou $B(C_p) \pm 3$. Také pro počet vad lze využít diagramy CUSUM, kdy se dosazuje za Z_j veličina

$$Z = \frac{C - C_p}{\sqrt{C_p}}.$$

Další typy regulačních diagramů pro diskrétní znaky lze nalézt např. v cit.¹. Obecně je důležité ověřit, zda diskrétní data pocházejí z binomického nebo Poissonova rozdělení a podle toho volit další zpracování:

A. Pro Poissonovo rozdělení platí, že $E(x) = \lambda = D(x)$, tj. střední hodnota je totožná s rozptylem. Parametr λ se odhaduje jako aritmetický průměr $\hat{\theta} = \bar{x}_s$. Při konstrukci regulačních diagramů Shewhartova typu se pak používá buď normalizační aproximace, nebo vztahu mezi Poissonovým a χ^2 -rozdělením.

B. Pro binomické rozdělení platí, že

$$E(x) = np \quad \text{a} \quad D(x) = np(1-p).$$

To znamená, že rozptyl se nerovná střední hodnotě. Parametr p se odhaduje s pomocí výběrového aritmetického průměru $\hat{p} = x_s/n$. Při konstrukci regulačních diagramů se používá aproximace Poissonovým nebo normálním rozdělením. Jednoduše lze ověřit shodu rozdělení diskretních dat s Poissonovým a binomickým rozdělením využitím grafu poměru frekvencí¹⁶. Orientačně lze použít poměru rozptylů $V = s^2/D(x)$, kde $D(x)$ se dosazuje podle toho, které rozdělení se ověřuje. Pokud leží V mimo interval $0.8 \leq V \leq 1.25$, znamená to, že dané rozdělení neaproximuje dobře experimentální data.

10.2.8 Regulační diagramy pro více proměnných

V řadě reálných situací se kvalita procesu vyjadřuje vektorem znaků jakosti měřených simultánně (ve stejném čase), nebo se pro vyjádření (nepřímé) jednoho znaku používá více různých metod. Pro případ, že jsou jednotlivé znaky jakosti vzájemně nezávislé, je pro ně možné sestavit regulační diagramy individuálně, aniž dojde ke zkreslení. Čím více jsou znaky jakosti vzájemně korelovány, tím více jsou jednoduché regulační diagramy "zkresleny". Předpokládejme, že výsledkem sledování procesu jsou náhodné vektory $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$. Každý vektor obsahuje q složek, reprezentujících q -tici znaků jakosti. Složky těchto vektorů mohou být buď jednotlivé hodnoty $x_{ij}, j = 1, \dots, q$, nebo výběrové průměry \bar{x}_{sij} . Při konstrukci vícerozměrných regulačních diagramů se vychází z předpokladu, že vektory \mathbf{x}_j jsou nezávislé a mají vícerozměrné normální rozdělení $N(\boldsymbol{\mu}, \mathbf{C})$ s vektorem středních hodnot $\boldsymbol{\mu}$ a kovarianční maticí \mathbf{C} . Pokud je proces ve standardním stavu, je $\boldsymbol{\mu} = \boldsymbol{\mu}_0$ a $\mathbf{C} = \mathbf{C}_0$. Parametry $\boldsymbol{\mu}_0$ a \mathbf{C}_0 jsou buď zadány, nebo se odhadují z dat, kdy je proces ve standardním stavu. Přirozeným zobecněním Shewhartových regulačních diagramů typu "x s pruhem" jsou *Hotellingovy regulační diagramy*, které využívají T^2 -statistiky, která je ekvivalentní Mahalanobisově vzdálenosti

$$T_i^2 = (\mathbf{x}_i - \boldsymbol{\mu}_0)^T \mathbf{C}_0^{-1} (\mathbf{x}_i - \boldsymbol{\mu}_0).$$

Rozdělení náhodné veličiny T_i^2 souvisí s náhodnou veličinou mající F -rozdělení známým vztahem

$$T_i^2(N, q) = \frac{Nq}{Nq+1} F(q, Nq+1),$$

kde N je počet vektorů, použitý k určení odhadu kovarianční matice \mathbf{C}_0 , resp. vektoru středních hodnot $\boldsymbol{\mu}_0$. Jednodušší je aproximace rozdělení T_i^2 pomocí χ^2 -rozdělení s q stupni volnosti. Výsledný regulační diagram je nesymetrický. Začíná na úrovni $T = 0$ a má pouze *UCL* definovanou vztahem

$$UCL = P_{0.9975}^2(q).$$

Do tohoto grafu se vynášejí hodnoty T_i^2 . Pokud se pracuje s výběrovými vektory $x_{s_{ij}}$ lze definovat zobecněnou Hotellingovu statistiku²². Je možné také použít vícerozměrné analogie regulačních diagramů CUSUM. Jedna konstrukce diagramů CUSUM využívá veličiny

$$L_i = \sqrt{(s_{i&1} \% x_i)^T C_0^{-1} (s_{i&1} \% x_i)} .$$

Vektory s_i se generují podle schématu

$$s_i = \mathbf{0} , \text{ pokud } C_i \neq k_1 ,$$

nebo

$$s_i = (s_{i&1} \% x_i) (1 \& k_1 / L_i) .$$

Zde $\mathbf{0}$ je nulový vektor a $k_1 > 0$. Do této varianty diagramu CUSUM se vynášejí hodnoty

$$Y_i = \sqrt{s_i^T C_0^{-1} s_i} .$$

Podrobnosti o volbě k_1 a indikaci změny stavu procesu jsou popsány v práci²⁴. Další možností je použití vektorů *kumulativních součtů*

$$d_i = \sum_{j=1}^i x_j$$

a statistiky $MC_i = \max [0, \sqrt{d_i^T C_0^{-1} d_i} \& k_2 l_i]$. Zde $k_2 > 0$ a dále

$$l_i = l_{i&1} \% 1 , \text{ pokud } MC_{i&1} > 0 , \text{ resp. } l_i = 1 , \text{ pokud } MC_{i&1} \neq 0 , \text{ pro } i = 1,$$

2, 3, ... Do diagramu se vynášejí hodnoty MC_i . Podrobnosti o interpretaci tohoto diagramu lze nalézt v cit.²⁵. Mezi účinné vícerozměrné regulační diagramy patří *vícerozměrná varianta exponenciálně váženého pohyblivého průměru (MEWMA)*

$$z_i = \mathbf{R} x_i \% (\mathbf{E} \& \mathbf{R}) z_{i&1} ,$$

kde $z_0 = 0$, \mathbf{R} je diagonální matice s prvky r_i na hlavní diagonále a \mathbf{E} je jednotková matice. Pokud není důvod pro různé "vážení" rozličných složek sledovaného vektoru znaků jakosti, volí se prvky $r_i = r$, $i = 1, \dots, q$. Tato rovnice pak přechází na tvar

$$z_i = r x_i \% (1 \& r) z_{i&1} .$$

Odpovídající kovarianční matice má tvar

$$C_{Zi} = \frac{r [1 \& (1 \& r)^{2i}]}{2 \& r} C_0 .$$

Z těchto veličin se sestavuje statistika Hotellingova typu $T_i^2 = z_i^T C_{Zi}^{-1} z_i$, která se

vynáší do příslušného diagramu. Pokud vyjde $T_i^2 > h_4$, je proces mimo standardní stav. Vhodné volby parametrů, určené z analýzy *ARL*, jsou $r = 0.1$ a $h_4 = 8.79$.

Je také možné sestavovat vícerozměrné analogie regulačních diagramů "s" nebo "s²". Vychází se z předpokladu, že je k dispozici požadovaná kovarianční matice C_0 . Účelem je,

aby se kovarianční matice určená z dat C od C_0 výrazně nelišila. Je známo, že skalární míra vícerozměrného rozptylu je zobecněný rozptyl $\det(C)$, kde $\det(\cdot)$ označuje determinant.

Vícerozměrná analogie "s" diagramů pak využívá zobecněné směrodatné odchylky $s_g = \sqrt{\det(C)}$. Je možné odvodit, že $E(s_g) = s_{g0} b_1$ a $D(s_g) = s_{g0}^2 (b_1^2 + b_3^2)$, kde $s_g = \sqrt{\det(C_0)}$. Parametry b_1 a b_3 jsou dány vztahy

$$b_1 = \frac{\sum_{i=1}^g (N+1)}{(N+1)^g} \quad \text{a} \quad b_3 = \left(\frac{2}{N+1} \right)^{g/2} \cdot \frac{(N/2)}{(N+1)/2}.$$

Za předpokladu přibližné normality zobecněné směrodatné odchylky s_g lze konstruovat regulační diagramy se střední linií $CL = s_{g0} b_3$ a regulačními mezemi

$$UCL = s_{g0} (b_3 + 3\sqrt{b_1^2 + b_3^2}),$$

$$LCL = s_{g0} (b_3 - 3\sqrt{b_1^2 + b_3^2}).$$

Do tohoto diagramu se vynášejí přímo hodnoty s_g . Vícerozměrná analogie regulačních diagramů s^2 vychází ze statistiky

$$W = \frac{1}{q(N+1)} \ln s_g + \frac{1}{(N+1)} \ln s_{g0} - \frac{1}{q(N+1)} \text{tr}(C_0^{-1}C),$$

kde $\text{tr}(\cdot)$ značí stopu matice. Veličina W odpovídá testovací statistice testu shody kovariančních matic věrohodnostním poměrem. Pro konstrukci regulačních diagramů se využívá toho, že W má asymptoticky χ^2 -rozdělení s $q(q+1)/2$ stupni volnosti. Speciálním případem jsou regresní regulační diagramy, resp. regulační diagramy založené na sledování vstupní a výstupní jakosti. Uvažujme pro jednoduchost, že v *první fázi* sledovaného procesu je jakostním znakem veličina x (může to být jakostní znak suroviny) a ve *druhé fázi* sledovaného procesu je jakostním znakem veličina y (může to být jakostní znak produktu). Jednotlivá měření (x_i, y_i) jsou realizována na stejné poloze produkce (např. při kusové výrobě), takže je zřejmé, že $y = f(x, \mathbf{b})$, kde $f(x, \mathbf{b})$ je obecně neznámá modelová funkce.

Při sestavování regulačního diagramu se pak postupuje v těchto krocích²⁶:

a) Metodami regresní analýzy se z počáteční N -tice bodů (y_i, x_i) , $i = 1, \dots, N$ (kdy je proces ve standardním stavu) určí regresní model a jeho parametry $f(x, \mathbf{b})$,

b) stanoví se rezidua $e_i = y_i - f(x_i, \mathbf{b})$ a použije se analogie Shewhartova diagramu pro jednotlivá pozorování. Střední úroveň v tomto diagramu je

$$CL = \frac{1}{N} \sum_{i=1}^N e_i.$$

Při použití lineární regrese a metody nejmenších čtverců je vždy $CL = 0$. Pro regulační meze pak platí

$$UCL = CL + 2.66 MR,$$

$$LCL = CL - 2.66 MR,$$

kde MR je odhad směrodatné odchylky reziduí počítaný z pohyblivého rozpětí

$$MR = \frac{1}{N+1} \sum_{i=1}^{N+1} e_{i^*} \quad \& \quad e_i^*.$$

Do takto konstruovaného grafu se vynášejí rezidua $e_j = y_j - f(x_j, \mathbf{b})$, určená na základě dalších dvojic (y_j, x_j) získávaných postupně v průběhu sledování daného procesu. Tuto

variantu regresních regulačních diagramů lze bez potíží zobecnit na případ vektoru vstupních znaků jakosti x . Funkce $f(x, \mathbf{b})$ může být lineární i nelineární a při její konstrukci se využívá všech technik budování regresních modelů¹⁶.

V některých případech je výhodnější použít přímo regresní model $f(x, \mathbf{b})$ a jako regulační meze konfidenční pásy. Jejich konstrukce je popsána např. v cit.¹⁶ (intervaly spolehlivosti predikce). Při praktickém použití je vhodné kombinovat regresní regulační diagramy s regulačními diagramy typu "x s pruhem" pro obě proměnné. Také u vícerozměrných regulačních diagramů jsou kritickými předpoklady normalita a nezávislost vektorů x_i . Kromě uvedených regulačních diagramů lze využít hlavních komponent, resp. Andrewsovy techniky znázornění vícerozměrných dat²³.

10.2.9 Používání regulačních diagramů

Regulační diagram má obecně sloužit jako diagnostický nástroj k posouzení, zda se sledovaný proces (představovaný nějakou měřenou veličinou nebo veličinami, které jej charakterizují) chová tak, jak očekáváme, zvláště pak, nedošlo-li k nečekané změně procesu. Došlo-li k takové změně, je třeba ji interpretovat - vysvětlit a případně přistoupit k nějakému zásahu. Příkladem měřených veličin jsou spojité veličiny jako pevnost, koncentrace, rozměr, elektrický odpor nebo diskrétní veličiny jako podíl zmetků na 1000 výrobků, počet povrchových vad na laku nebo počet uzlíků na 1 m² tkaniny. Kromě samotné hodnoty je nutno v případě spojité veličiny sledovat také její variabilitu (míru kolísání či rozptylu), která je pro posouzení procesu stejně důležitá. Proto Shewhartův regulační diagram musí vždy obsahovat informace jak o sledované hodnotě samotné, tak o její variabilitě.

10.2.10 Konstrukce regulačních diagramů

Postup konstrukce regulačního diagramu

1. Zvolíme takovou část procesu, která odpovídá naší představě, předpisu nebo zkušenosti a připravíme příslušná procesní data.
2. Na základě těchto dat stanovíme jejich statistický model. Obyčejně však máme pouze střední hodnotu (aritmetický průměr) a směrodatnou odchylku a ověříme platnost statistických předpokladů Shewhartova diagramu.
3. Z těchto dvou parametrů se zkonstruuje vlastní regulační diagram, který má podobu základní linie *ZL* (angl. central line, *CL*) a horní a spodní regulační meze *LCL* a *UCL* (angl. lower control level a upper control level).
4. Do tohoto regulačního diagramu se pak vynášejí další data z procesu a sleduje se výskyt 'zvláštních případů', signalizujících nečekanou změnu chování procesu, z nichž základní je překročení regulační meze.
5. Výskyt zvláštních případů se eviduje a hledá se tzv. přiřaditelná příčina (pokud se ji podaří identifikovat) a opatření, které bylo přijato.

Základními předpoklady pro Shewhartův regulační diagramu měřením jsou:

- a) Normalita rozdělení dat, symetrie.

- b) Konstantní střední hodnota procesu.
- c) Konstantní rozptyl (směrodatná odchylka) dat.
- d) Nezávislost, nekorelovanost dat.
- e) Nepřítomnost vybočujících hodnot.
- f) Vhodně zvolené podskupiny.

Tyto předpoklady je nutno testovat před konstrukcí regulačního diagramu postupy pro analýzu jednorozměrného výběru. Pokud se nepodaří ověřit předpoklady pro použití diagramu v bodě b), je nutno zdroje porušení předpokladů ověřit. V případě, že je zdroj náhodný a není předpoklad, že by se měl opakovat, je možné ‘problematická data’ (např. vybočující hodnoty) ze souboru vyloučit a diagram konstruovat bez nich. Pokud je ale porušení předpokladů systematické, je inherentní vlastností procesu, nebo se jej nepodaří uspokojivě vysvětlit, není možné příslušná data vylučovat. Pak je třeba uvažovat o jiném typu regulačních diagramů. Zvláštní případy jsou takové situace, které jsou při optimálním průběhu procesu velmi nepravděpodobné. Historicky prvním případem je překročení regulačních mezí. Jsou-li regulační meze pro normálně rozdělená data stanoveny jako $\pm 3s$, je pravděpodobnost jejich překročení 0.27 %. To znamená, že k překročení dojde v průměru jednou z $1/0.0027 = 370$ případů, tedy přibližně jednou za rok, máme-li jedno měření denně.

10.2.11 Pravidla pro určování zvláštních případů

Zvláštní případy rozšiřují diagnostické možnosti Shewhartova diagramu a umožňují detekci poruch a změn, které se neprojeví překročením regulačních mezí, nebo by se projeví se zpožděním. Jedná se o osm nejpoužívanějších a ustálených situací. Pravděpodobnost jejich výskytu pro ideální data je srovnatelná s pravděpodobností překročení mezí. Pravděpodobnost výskytu následujících případů v normálně rozdělených nezávislých datech byla zjištěna výpočtem nebo simulacemi a je přibližně 0.25 %. V těchto osmi případech je nutno uvažovat o hledání přiřaditelné příčiny, případně regulačním zásahu. V případě jiného než normálního rozdělení může být tato pravděpodobnost o něco větší. Každý z těchto případů ukazuje na poruchu určitého druhu a lze jej použít jako užitečné vodítko při hledání přiřaditelné příčiny. Povahu možné poruchy uvádíme u jednotlivých případů. Grafické znázornění je na obr. 10.6.

Pravidla k odhalení zvláštních případů v diagramu “x s pruhem” a “x-individual” (podle ISO 8258) s komentářem

Pravidlo 1. *Jedna hodnota je mimo regulační meze.*

Lokální porucha procesu, chybné měření, výpadek. Chybně stanovené regulační meze, malá variabilita uvnitř podskupiny při konstrukci diagramu. Opakuje-li se na téže straně, může jít o posunutí střední hodnoty nebo o asymetrické rozdělení dat. Opakuje-li se na obou stranách, může jít o zvýšení nestability nebo rozptylu dat.

Pravidlo 2. *9 hodnot je na téže straně od centrální linie.*

Pravděpodobné posunutí střední hodnoty, snížení variability mezi podskupinami, asymetrie dat, příliš široké nebo neodpovídající regulační meze.

Pravidlo 3. *6 hodnot monotónně roste či klesá.*

Autokorelovaný proces, závislá měření. Lineární trend, způsobený opotřebením nebo výpadekem. Příliš široké regulační meze. Odstraněním přiřaditelné příčiny lze někdy zvýšit C_p .

Pravidlo 4. *14 alternujících hodnot.*

Přeregulovaný nebo nestabilní proces. Autokorelovaná měření se záporným r . Odstraněním přiřaditelné příčiny lze někdy zvýšit C_p . Podvádění operátorem, vymyšlená čísla.

Pravidlo 5. *2 ze 3 hodnot je mimo interval $\pm 2s$.*

Varování před možným překročením regulačních mezí.

Pravidlo 6. *4 z 5 hodnot mimo interval $\pm s$ na téže straně centrální linie.*

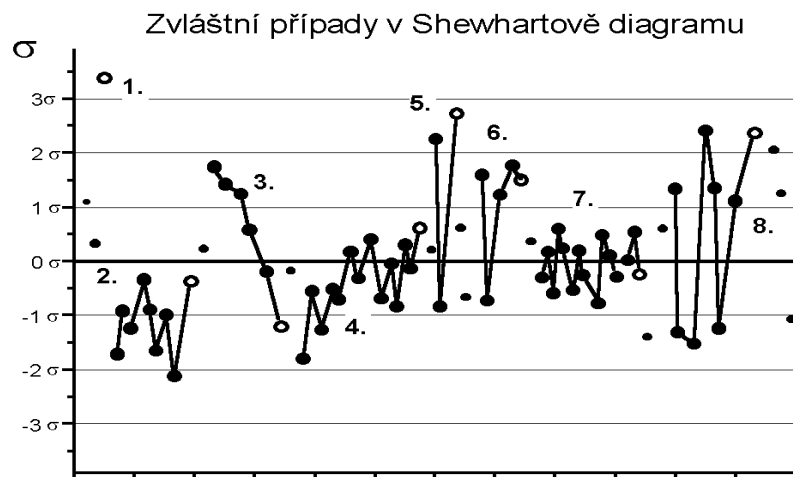
Pravděpodobné posunutí střední hodnoty. Varování před možným překročením regulačních mezí.

Pravidlo 7. *15 hodnot je uvnitř intervalu $\pm s$.*

Snížení variability mezi podskupinami. Při opakování uvažovat o nových regulačních mezích. Nesprávná volba regulačních mezí. Podvádění operátorem, vymyšlená čísla.

Pravidlo 8. *8 hodnot je mimo interval $\pm s$ na obou stranách centrální linie.*

Zvýšení variability mezi podskupinami. Varování před překročením regulačních mezí. Porucha procesu.



Obr. 10.6 Znárodnění pravidel pro určování zvláštních případů.

10.2.12 Porušení předpokladů o datech

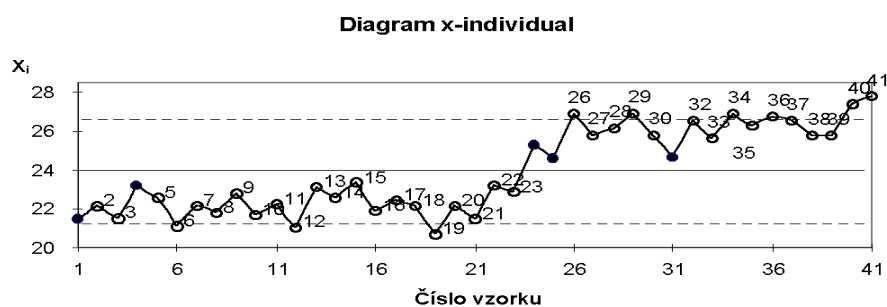
Regulační diagram pro individuální hodnoty je na různá porušení předpokladů o datech obvykle citlivější než ostatní Shewhartovy diagramy. Především to platí o normalitě dat. V praxi se často setkáváme s procesy a procesními veličinami, které nejsou nezávislé, nemají konstantní střední hodnotu nebo rozptyl, nejsou normálně rozdělené atd. Ze zkušeností víme, že statisticky “dobré” chování většinou vykazují strojírenské procesy, kde jsou měřeny veličinami rozměry nebo hmotnosti. V případě měření dalších fyzikálních veličin, jako pevnosti, viskozity, se setkáváme s asymetrickým rozdělením. Při sledování spojitých procesů v chemii, farmacii, potravinářství, metalurgii vykazují data často silnou závislost. Těžko ovlivnitelná jakost suroviny (např. horniny) může mít za následek kolísání nebo nekonstantnost střední hodnoty. Při sledování emisí a stopových koncentrací nečistot se setkáváme s logaritmicko-normálním, asymetrickým rozdělením atd. Ve všech těchto případech se jedná o vlastnost procesu, která se buď nedá ovlivnit, nebo se s ní v technologii počítá. Konstrukce Shewhartových diagramů může však v těchto případech selhat. Velmi

hrubě jsme se pokusili shrnout zkušenosti s několika stovkami typických reálných datových souborů z různých odvětví a technologií. Uvádí je následující tabulka, ve které je kroužkem označeno převažující splnění předpokladu a křížkem pak jeho porušení.

Typická porušení předpokladů v různých technologiích

Odvětví / technologie / veličina	Normalita	Nezávislost	Konstantnost střední hodnoty	Homogenita, vybočující body
Mechanické strojírenství, automobilový průmysl (rozměry)	o	o	o	o
Mechanické zkušebny (pevnost, pružnost, ...)	x	o	o	x
Chemie, metalurgie, hutnictví (koncentrace, obsahy)	o	x	x	x
Chemie, metalurgie, hutnictví (ostatní fyzikální parametry)	o	x	x	x
Životní prostředí, hygiena (nízké koncentrace)	x	x	x	x
Elektrické veličiny, součástky	o	o	o	x
Energetika	x	x	x	x
Plasty, polymery, textil, fyzikálně-mechanické veličiny	x	o	x	o
Biochemie, farmacie, potravinářství	x	x	o	o
Vnitropodnikové ekonomické a finanční ukazatele	x	x	x	o
Sociologie, lidské zdroje	x	x	x	x

Tato tabulka naznačuje, že ve většině případů je na místě určitý pesimismus vzhledem k jednoduchému mechanickému použití klasických regulačních diagramů. Je nutné mít k dispozici a používat takové typy a techniky konstrukce regulačního diagramu, které odpovídají reálným datům a povaze procesu či technologie. V opačném případě může vést použití regulačních diagramů ke chybné interpretaci a nedůvěře k technikám statistického řízení procesu.

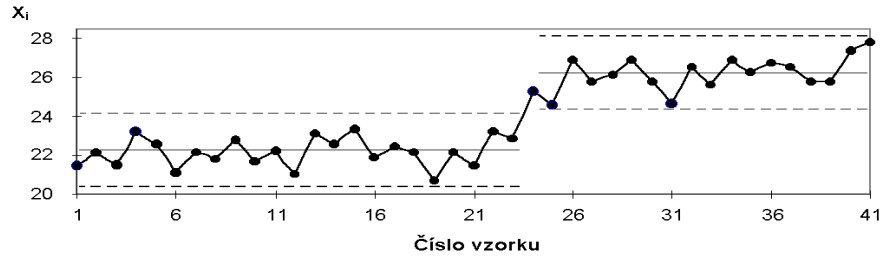


Obr. 10.8 Změna střední hodnoty ve Shewhartově regulačním diagramu typu x-individual.

Obr. 10.8, *nekonstantní střední hodnota*, ukazuje případ, kdy se změnila střední hodnota. Ke skokové změně dochází např. při přechodu na jinou surovinu nebo technologii. K posouzení významnosti rozdílu středních hodnot lze použít *t*-test.

Obr. 10.9, *trvalá změna střední hodnoty*, ukazuje, kdy je tato změna trvalá a je třeba konstruovat regulační diagram s novou základní linií.

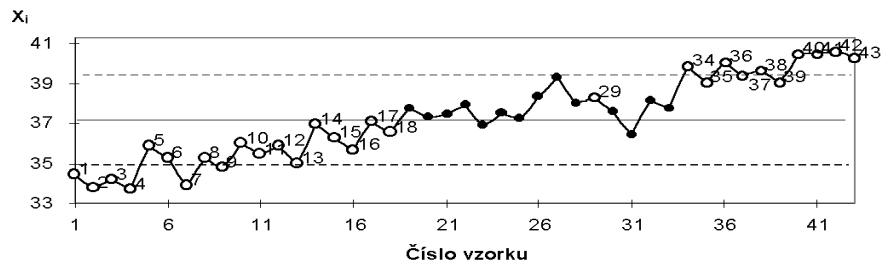
Diagram x-individual



Obr. 10.9 Data z obr. 10.8 o dvou úrovních základní linie.

Obr. 10.10, *data vykazují stálý statisticky významný trend*. Trend je možné prokázat testem významnosti směrnice.

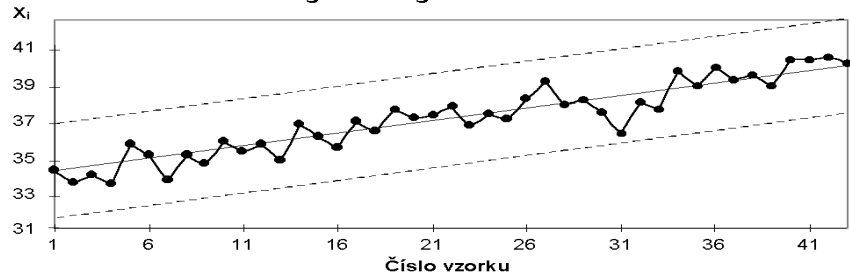
Diagram x-individual



Obr. 10.10 Lineární trend v datech.

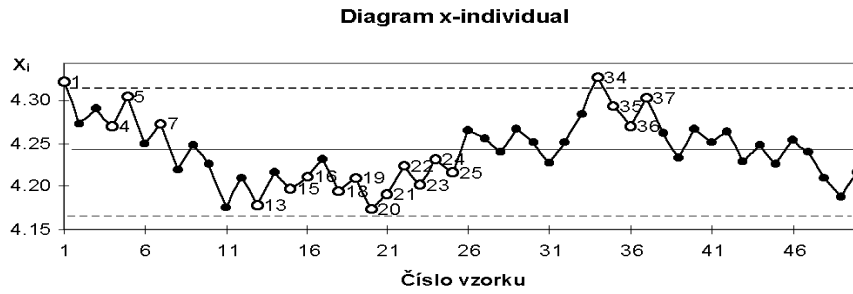
Obr. 10.11, *základní linie představuje regresní přímku* v případě lineárního trendu v datech. Alternativou je konstrukce regulačního diagramu pro rezidua (odchyly od regresní přímky).

Regresní diagram x-individual



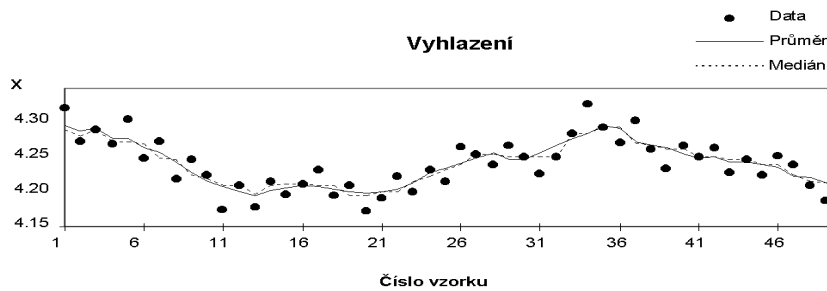
Obr. 10.11 Regrešní kontrolní diagram pro data z obr. 10.10 .

Obr. 10.12, *kolísání dat a autokorelace v datech* se dá prokázat např. znaménkovým testem nebo testem významnosti autokorelačního koeficientu.

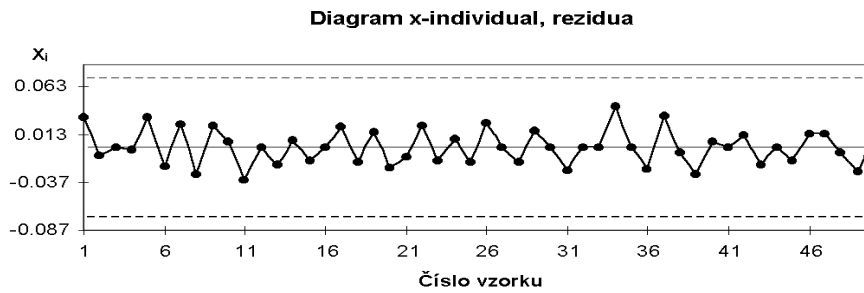


Obr. 10.12 Kolísání v datech, pokud o diagram “x-individual”.

Obrázky 10.13 a 10.14, vyhlazená data, lze zpracovat dynamickým regulačním diagramem EWMA, případně vyhladit pro konstrukci diagramu použít rezidua.

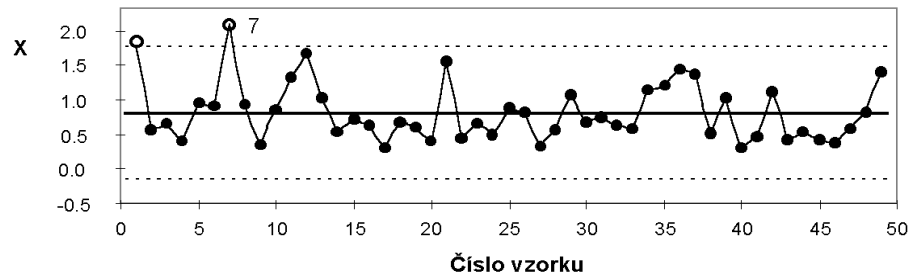


Obr. 10.13 Vyhlazená data z obr. 10.12.



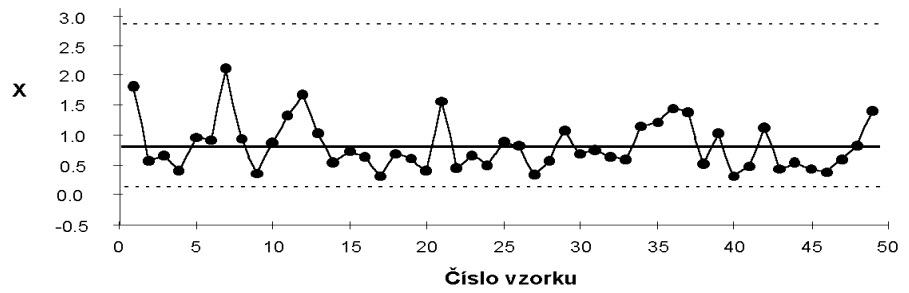
Obr. 10.14 Rezidua z vyhlazených dat na obr. 10.13 použitá pro diagram “x-individual”.

Obr. 10.15, *asymetrické rozdělení dat* (koncentrací). Jako kritérium lze použít test normality nebo test významnosti mocninné transformace. Spodní regulační mez vyšla záporná, což je v daném případě nesmyslné.



Obr. 10.15 Klařícký Shewhartův regulační diagram pro alymetrická data.

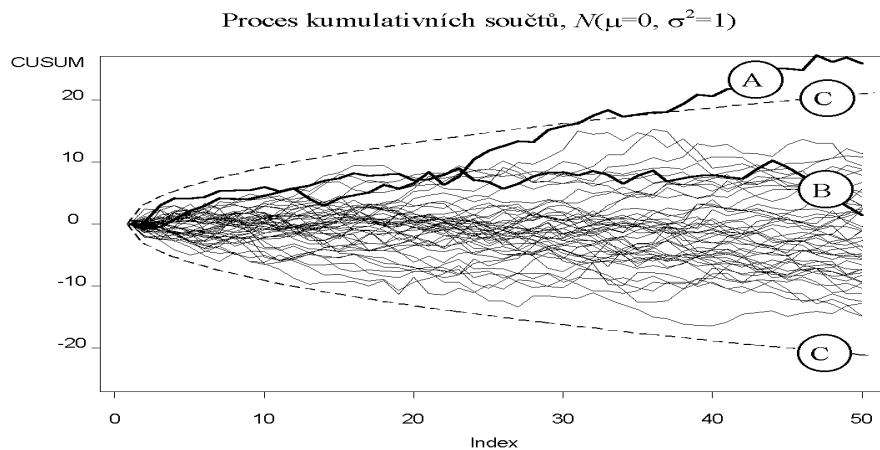
Obr. 10.16, *asymetrické regulační meze*: Regulační meze jsou konstruované pomocí retransformovaných kvantilů po nelineární Boxově-Coxově transformaci dat. Asymetrické regulační meze zde odpovídají povaze dat, spodní mez je kladná.



Obr. 10.16 Shewhartův regulační diagram s retransformovanými mezemi.

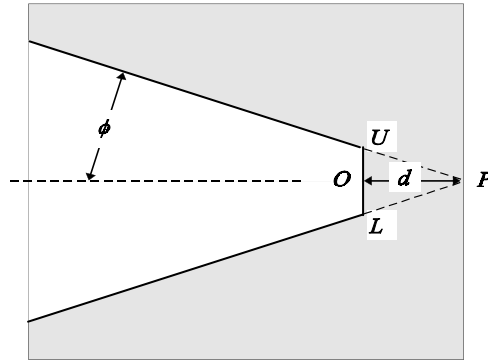
10.2.13 Pomůcky diagramů kumulativních součtů CUSUM

Diagramy, založené na kumulativních součtech (anglicky CUMulative SUMs) jsou rychlou detekcí relativně malého posunutí střední hodnoty procesu. Ve srovnání s Shewhartovými diagramy je tato detekce až o řád rychlejší. Ke konstrukci těchto diagramů se používá postupných součtů odchylek měřené veličiny od předepsané nebo očekávané konstantní cílové hodnoty K . Takový proces se ve statistice nazývá náhodné kráčení. Na obr. 10.17a je ukázka 50 diagramů, tvořených vygenerovanými posloupnostmi S_i . Má-li X_i konstantní střední hodnotu K a směrodatnou odchylku s , pak výsledná křivka na obr. 10.17a bude s velkou pravděpodobností ležet uvnitř parabolické oblasti C. Použijeme-li pravidla $3s$, bude totiž oblast, v níž leží 99.73 % bodů s_i , ležet uvnitř paraboly $\pm 3s_x \%V$. Dojde-li k odchylce střední hodnoty x od hodnoty K , (křivka A na obr. 10.17a), proces tyto meze velmi rychle překročí. Reakce směrodatné odchylky s_i na odchylku od hodnoty K je mnohem citlivější než u Shewhartových diagramů, proto jsou diagramy CUSUM velmi oblíbené a rozšířené.



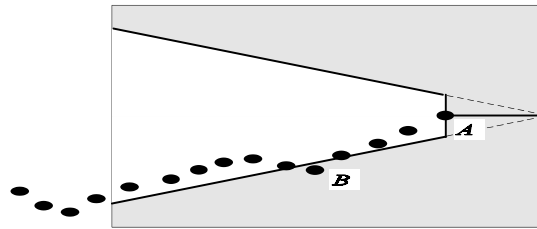
Obr. 10.17a Povaha procesu kumulativních součtů, $n=50$, A-vychýlená data, $m=+0.6$; B-nevychýlená data, $m=0$; C-parabolická mez.

Při praktickém použití se využívalo, a někde ještě využívá, speciální masky, která je zjednodušením paraboly, takzvané *V-masky*. V-masky, která je znázorněna na 10.17b, se vystříhla z papíru, a přikládala na graf kumulativních součtů.



Obr. 10.17b V-mařka.

Geometrické parametry V-masky d a f lze vypočítat podle vztahů upravených do tvaru $d = (2/\delta^2) \ln(1-\beta)/\alpha \varphi \operatorname{tg}^{-1}(\Delta x/2k)$, kde α , β jsou rizika I. resp. II. druhu, ΔX je odchylka od hodnoty K , kterou je třeba detekovat a d je její standardizovaná hodnota, $\delta = \Delta X/\sigma_X$, k je faktor měřítka v grafu, poměr velikosti jednotky na svislé a vodorovné ose. Hodnota δ (tedy citlivost na odchylku v jednotkách σ_X) se volí obvykle mezi 0.5 a 1.5. Pro menší hodnotu δ je diagram příliš citlivý a je nebezpečí, že bude hlásit falešné poplachy, pro větší hodnotu δ je diagram málo citlivý a ztrácí se jeho přednost před Shewhartovým diagramem. Hodnota ARL , tedy průměrný počet dat mezi dvěma následnými překročeními regulačních mezí je pro ideální data u diagramu CUSUM 500 oproti 370 u srovnatelného Shewhartova diagramu.



Obr. 10.18 Použití V-mařky, A-pořádní vynesení bod; B-bod mimo výseč.

Při použití pomůcky V-masky se postupuje podle obr. 10.18. Masku se přiloží vodorovně bodem na poslední vynesení bodu diagramu (A). Ocitne-li se některý z předchozích vynesení bodů mimo výseč (B), je bod A (nikoli B!) označen a proces se považuje za odchýlený od K . Výhodou tohoto postupu je jednoduchost výpočtu vynášených hodnot, což mělo význam na pracovištích bez výpočetní techniky. Nevýhodou je to, že vidíme vždy jen jediný bod (v našem případě bod A). Když chceme testovat jiný bod, musíme posunout masku. Nemáme tak přehled o delším úseku procesu, jak jsme na to byli zvyklí u Shewhartových diagramů. Tuto nevýhodu nemá např. rozšířený, modernější Lucasův postup konstrukce diagramu CUSUM.

10.3 Indexy způsobilosti procesu

Je přirozenou snahou řady výrobců charakterizovat nějakým způsobem kvalitu výrobního procesu s ohledem na rezervy při zvyšování jakosti výrobků. Obecně zahrnuje analýza způsobilosti procesů celou řadu fází:

1. Definici výrobního procesu jako systému (stanovení všech vazeb vývojového diagramu).
2. Specifikaci vstupních proměnných, procesních proměnných a charakteristik produktu a způsobu jejich měření.
3. Výběr znaků jakosti, které se budou sledovat, kde se budou sledovat a kdy.
4. Posouzení stavu výrobního procesu na základě indexů způsobilosti pro vybrané znaky jakosti.

Pro tento účel byla navržena v posledních 20 letech celá řada různých (snadno určitelných, ale hůře interpretovatelných) indexů. Nejznámější z nich je tzv. *index způsobilosti procesu* (proces capability index, *PCI*) označovaný symbolem C_p . V této kapitole je ukázán statistický charakter tohoto a souvisejících indexů způsobilosti.

Index C_p : Předpokládejme, že známe měřitelný spojitý znak x , který charakterizuje kvalitu výrobního procesu (např. nestejnomyšlnost příze, rozměry strojní součásti, elektrický odpor elektronické součástky atp.). Na základě požadované funkčnosti výrobku nebo polotovaru můžeme obvykle určit dolní *LSL* a horní *USL* meze specifikace jakosti. Pokud padne hodnota x do těchto mezí, je výrobek považován za jakostní, a pokud padne mimo tyto meze, jde o zmetek (nevyhovující výrobek, *NC*). Mírou toho, nakolik je výrobní proces schopen zajistit splnění požadavků $LSL < x < USL$, je index způsobilosti C_p . Je zřejmé, že čím bude C_p vyšší, tím bude proces výroby kvalitnější. Pro dobré výrobní procesy by mělo platit, že $C_p > 1$, tj. statistická variabilita procesu výroby je dostatečně malá ve srovnání s tolerančním intervalem $[USL, LSL]$. Pro případ, že znak x má normální rozdělení $N(d, \sigma^2)$, kde

$$d = (LSL + USL)/2,$$

je zřejmě očekávaný podíl zmetků, tj. hodnot mimo tolerance, roven $P_C = 2F_N(-m/\sigma)$,

kde $F_N(\cdot)$ označuje distribuční funkci normovaného normálního rozdělení a $m = (USL - LSL)/2$. Po dosazení vyjde

$$P_C = 2F_N(-3C_p).$$

Pro případ, že index způsobilosti $C_p = 1$, je z tabulek možno nalézt $F_N(-3) = 0.00135$. Pak $P_C = 0.0027$ a tedy maximálně 0.27 % výrobků budou tvořit zmetky. Při znalosti P_C lze za předpokladu normality definovat C_p z rovnice ve tvaru¹²

$$C_p = 0.33 F_N^{-1}((1 + P_C)/2),$$

kde $F_N^{-1}(\cdot)$ je kvantilová funkce normovaného normálního rozdělení. Na základě výběrového průměru x_S a rozptylu s^2 lze určit přirozený odhad P_C ve tvaru

$$P_C \approx 1 - F_N\left(\frac{USL - x_S}{s}\right) + F_N\left(\frac{LSL - x_S}{s}\right).$$

Tento odhad je však pro reálné výběry vychýlený. Poměrně komplikovaný nevychýlený odhad je publikován v práci¹¹. Pokud se směrodatná odchylka σ nahrazuje klasickou výběrovou směrodatnou odchylkou s , je odhad C_p^* náhodnou veličinou. Rozptyl tohoto odhadu je přibližně roven

$$\text{var}(C_p^*) = C_p^2 / (2(N - 3)),$$

kde N je velikost výběru, ze kterého byla počítána směrodatná odchylka s . Odpovídající 100(1- ν)%ní interval spolehlivosti pro populační hodnotu C_p má tvar

$$\frac{P_{</2}^2 (N \& 1) C_p^{\langle}}{\sqrt{N \& 1}} \# C_p \# \frac{P_{1\&</2}^2 (N \& 1) C_p^{\langle}}{\sqrt{N \& 1}},$$

kde $\chi_{\nu}^2(N - 1)$ je 100 ν %ní kvantil χ^2 -kvadrát rozdělení s $N - 1$ stupni volnosti. Pro malé výběry je C_p^* vychýlený odhad parametru C_p . Nevychýlený odhad C_p^N má pro $N > 15$ tvar

$$C_p^N = C_p^{\langle} (1 \& 0.75/(N \& 1))$$

a jeho rozptyl je roven $\text{var}(C_p^N) = \frac{8(N \& 1) \% 9}{16(N \& 1)(N \& 3)} C_p^2$.

Další indexy způsobilosti procesů: Dosavadní úvahy byly založeny na platnosti vztahu $d = (LSL + USL)/2$. Obecně však střední hodnota d_T znaku jakosti x není rovna d určenému z tohoto vztahu. Modifikovaný index C_{PK} má pak tvar¹³

$$C_{PK} = \min[(USL - d_T)/3\sigma, (d_T - LSL)/3\sigma],$$

resp. po úpravě $C_{PK} = C_p (1 - |d_T - d|/m)$. Je patrné, že C_{PK} je vždy menší nebo rovno C_p . Očekávaný podíl zmetků P_C je pak roven

$$P_C = F_N \left(\frac{LSL \& d_T}{F} \right) \% F_N \left[1 \& \left(\frac{USL \& d_T}{F} \right) \right].$$

Odhad C_{PK}^* lze určit snadno při znalosti výběrového průměru \bar{x}_S a rozptylu s^2

$$C_{PK}^* = (m - |\bar{x}_S - d|) / 3s.$$

Tento odhad je opět pro malé rozsahy výběru N vychýlený. Přibližný 100(1- ν)%ní interval spolehlivosti parametru C_{PK}^* je roven

$$\frac{(1 \& u_{1\&</2}) C_{PK}^{\langle}}{\sqrt{2N \& 2}} \# C_{PK} \# \frac{(1 \% u_{1\&</2}) C_{PK}^{\langle}}{\sqrt{2N \& 2}}.$$

Z uvedeného výkladu je patrné, že pokud lze přijmout předpoklad normality rozdělení znaku x , lze stanovit intervaly spolehlivosti pro C_p , resp. C_{PK} , podle uvedených vztahů. Komplikace činí především nutnost odhadu směrodatné odchylky a průměru z výběrových hodnot. V souvislosti se ztrátovou funkcí používanou při Taguchiho přístupu k hodnocení jakosti byl navržen index¹⁴

$$C_{Pm} = \frac{USL \& LSL}{6 \sqrt{F^2 \% (d_T \& T)^2}},$$

kde T je předepsaná (cílová hodnota) parametru x . Obyčejně platí, že $T = d$, tj. jde o střed tolerančního intervalu.

Ztrátová funkce $L(x)$ je standardně uvažována ve tvaru

$$L(x) = K (x - T)^2.$$

Výraz $\sigma^2 + (d_T - T)^2$ je pak *mírou průměrné ztráty*, způsobené nedodržením podmínek kvalitní výroby ($\sigma^2 = 0$, $T = d_T$). Je zřejmé, že $C_{Pm} \# C_p$. Lze odvodit, že

$$C_{Pm} = C_p \sqrt{1 + Z^2}, \text{ kde } Z = (d_T - T)/\sigma.$$

Při použití C_{Pm} je třeba brát v úvahu to, že je vhodný pro vyjádření způsobilosti procesů jen, pokud je cílová hodnota T rovna středu tolerančního intervalu d . Nevhodnost použití C_{Pm} pro obecné situace je patrná z tohoto jednoduchého příkladu. Mějme tři procesy s následujícími charakteristikami

- A $d_T = 50.00$ $\sigma = 5$,
 B $d_T = 57.50$ $\sigma = 2.5$,
 C $d_T = 61.25$ $\sigma = 1.25$.

Nechť $T = 50$, $USL = 65$ a $LSL = 35$. Pro tyto hodnoty jsou C_p , C_{PK} a C_{Pm} rovny

Proces	C_p	C_{PK}	C_{Pm}
A	1	1	1
B	2	1	0.64
C	4	1	0.44

Je patrné, že každý z uvedených indexů ukazuje zcela něco jiného. Základní rozdíl mezi C_{Pm} a C_{PK} spočívá v relativní významnosti mezi specifikace (LSL , USL) a cílové hodnoty T .

Funkcí indexu C_{PK} je indikace, do jaké míry je proces uvnitř mezi specifikace. Pokud $LSL < d_T < USL$, je pro $\sigma > 0$ hodnota $C_{PK} < 1$. Vysoké hodnoty C_{PK} však neukazují na rozdíly mezi T a d_T .

Index C_{Pm} ukazuje na míru, nakolik je proces blízký cílové hodnotě T . Pro případ, že $\sigma > 0$ a $d_T = T$ je pak $C_{Pm} = 1$. Zde meze specifikace slouží jen jako škála pro ztrátovou funkci. Lze ukázat, že pokud T není rovno d , vede posun průměrné hodnoty znaku x směrem k cílové hodnotě T (zvyšující C_{PK} a C_{Pm}) ke zvýšení podílu zmetků, počtu výrobků mimo meze specifikace.

Je patrné, že C_{PK} vzniklo modifikací čitatele a C_{Pm} modifikací jmenovatele vztahu pro C_p . Kombinací těchto dvou modifikací lze dospět k *indexu C_{PmK}* , definovanému výrazem⁵

$$C_{PmK} = \frac{m \& *d_T \& d^*}{3 \sqrt{F^2 \& (d_T \& T)^3}}.$$

Je zřejmé, že $C_{PmK} = C_{PK} C_{Pm} / C_p$ a pro případ, kdy $d_T = T = d$, jsou tyto indexy stejné. *Přirozený odhad indexu C_{PmK}* má tvar

$$C_{PmK}^{(c)} = \frac{m \& *x_s \& d^*}{3 \sqrt{\frac{\sum_{i=1}^N (x_i \& T)^2}{N}}}$$

Všechny dosavadní úvahy jsou založeny na předpokladu normality. Pokud je rozdělení znaku x silně nenormální, jsou odhady indexů způsobilosti vychýlené a mají nesprávný rozptyl. Pokud lze pro rozdělení znaku x specifikovat velikost šikmosti g_1 a špičatosti g_2 , lze použít tabulek (viz cit.¹¹) pro nahrazení čísla 6 ve jmenovateli vztahu pro C_p takovým číslem, pro které bude $P_c = 0.0027$. Pokud se místo $P_c = 0.0027$ použije vyšší hodnota $P_c = 0.01$ (tj. připouští se max. 1 % zmetků), je ve jmenovateli vztahu pro C_p nutno použít číslo 5.15. Zajímavé je, že tato hodnota vyhovuje také pro celou řadu dalších rozdělení ($0 < g_1 < 9$ a $1 < g_2 < 6$). Z uvedeného textu vidíme, že při použití indexů způsobilosti procesů je nezbytné

uvažovat také jejich meze a jejich statistickou povahu. Jen tak lze omezit jejich nevhodné aplikace. Pro případ více znaků, charakterizujících jakost procesu lze použít zobecněných parametrů, založených např. na Hotellingově T^2 statistice¹¹.

10.4 Software pro řízení jakosti

Jak je patrné, není vlastní použití regulačních diagramů a indexů způsobilosti procesu nikterak náročné. Velká většina metod byla konstruována tak, že nevyžaduje žádné složitější výpočty a tedy ani počítač. S tím, jak se i do provozů dostává výpočetní technika, sběr dat se provádí automaticky a řada strojů se jak monitoruje, tak i řídí počítačem, rostou i požadavky na vhodný software. Bohužel velká většina programů nevyužívá nikterak speciálně možností výpočetní techniky, takže pouze nahrazuje rutinní práci praktiků. Tento stav je často považován za vyhovující. Potíže však nastávají, jakmile nejsou splněny předpoklady, za kterých byly tyto metody odvozeny (normalita, nepřítomnost hrubých chyb, dostatečná velikost výběru, statistická stabilita procesu atd.). Lze říci, že řada speciálních programů pro řízení jakosti je orientována pouze na konstrukci regulačních diagramů a různých indexů, aniž ověřuje kvalitu dat. Z tohoto hlediska jsou výhodnější některé statistické programy nebo jejich kombinace, které umožňují komplexní statistickou analýzu dat a navíc konstrukci diagramů resp. dalších pomůcek pro řízení jakosti. Přehled současného softwarového vybavení pro oblast jakosti je uveden např. v cit.¹⁷ Pro komplexní řešení úloh souvisejících s jakostí a jejím řízením je nutná celá řada statistických metod. Mezi základní patří:

- (1) Analýza jednorozměrných výběrů.
- (2) Porovnání shody výběrového rozdělení se zvoleným teoretickým.
- (3) Analýza rozptylu.
- (4) Regresní analýza.
- (5) Konstrukce pravděpodobnostních modelů.

Ze všech statistických programových systémů, které umožňují řešení uvedených úloh, jsou dále diskutovány pouze ty, které se jeví jako pro praxi vhodné a jsou na našem trhu běžně k dispozici.

Jak bylo ukázáno, souvisí problém řízení jakosti velmi úzce s tvorbou intervalů spolehlivosti vybraných charakteristik jakostních znaků. Tyto intervaly vycházejí ze znalosti dobrého odhadu směrodatné odchylky a vyžadují splnění základních předpokladů o datech, jako jsou nezávislost, nepřítomnost vybočujících měření, normalita. Vhodný statistický software by měl tedy umožnit ověření těchto předpokladů nesložitými "robustními" technikami, které jsou pokud možno distribučně nezávislé.

Je zřejmé, že vlivem kolísání vlastností surovin, podmínek zpracování, stavu okolí atd., mají průmyslové procesy náhodný charakter. Variabilita procesů je jedním ze základních ukazatelů jeho stavu a její odhad se používá jak při konstrukci regulačních diagramů, tak i při výpočtech indexů způsobilosti. Základní příčiny variability průmyslových procesů lze rozdělit do těchto skupin:

- (a) *Náhodné šumy*, které se vyskytují i za podmínek, že je proces v optimálním (standardním) stavu. Jsou způsobeny nekontrolovanými příčinami a lze je obecně snížit pouze změnou procesu (strojního zařízení).
- (b) *Externí zdroje variability*, způsobené změnami podmínek okolí.
- (c) *Procesní zdroje variability*, způsobené nekonstantností procesních parametrů.

(d) *Přiraditelné zdroje variability*, způsobené např. změnou jakosti suroviny, špatným seřazením strojů, opotřebením pracovních orgánů, stárnutím materiálů a různými skokovými změnami externích, resp. procesních proměnných ovlivňujících stav procesu.

Techniky řízení jakosti se snaží postihnout právě *přiraditelné zdroje variability*, které lze provedením "regulačního" zásahu eliminovat a převést zpět do stavu, kdy proces ovlivňuje jen náhodné šумы. Standardně se při posuzování variability procesů vychází z (nezávislých) výběrů $V_1, \dots, V_p, \dots, V_M$, z nichž má každý velikost N a je charakterizován výběrovým průměrem \bar{x}_{Si} a rozptylem s_j^2 . Symbol x_{ij} označuje j -tý prvek v i -tém výběru V_i . Je účelné sestavit typické modely ukazující zdroje variability:

(a) *Nejjednodušší model* je model náhodného šumu, pro který platí

$$x_{ij} = d + \sigma_c G_{jp}, \quad i = 1, \dots, M, j = 1, \dots, N,$$

kde d je celkový průměr, σ_c^2 je rozptyl (uvnitř výběrů) a G_{ij} jsou nezávislé náhodné veličiny s normovaným normálním rozdělením $N(0, 1)$. Odhadem d je \bar{d} a odhadem σ_c je σ_c^* . Tento model je tedy standardně používán při konstrukci regulačních diagramů.

(b) Poněkud *komplexnější model* uvažuje, že kromě šumů se projeví i další zdroje variability, ovlivňující každý výběr jako celek (procesní, resp. externí zdroje variability). Pro tento model platí

$$x_{ij} = d + \sigma_B \omega_i + \sigma_c G_{jp}, \quad i = 1, \dots, M, j = 1, \dots, N,$$

kde σ_B^2 je rozptyl mezi výběry a ω_i jsou nezávislé náhodné veličiny s normovaným normálním rozdělením $N(0, 1)$. Pro tento model platí, že rozptyl měření je roven

$$D(x_{ij}) = \sigma_B^2 + \sigma_c^2$$

a rozptyl výběrových průměrů je roven

$$D(\bar{x}_{Si}) = \sigma_B^2 + \sigma_c^2/N.$$

Je patrné, že zvětšováním počtu prvků ve výběru N se sníží pouze část variability (uvnitř výběrů), ale neovlivní rozptyl mezi výběry. Odhad s_B^2 rozptylu mezi skupinami σ_B^2 se provádí obvykle podle vztahu pro klasický výběrový rozptyl pro jednotlivé průměry

$$s_B^2 = \frac{1}{M-1} \sum_{i=1}^M (\bar{x}_{Si} - \bar{d})^2.$$

Tento odhad je však vychýlený, protože obsahuje příspěvek z variability uvnitř výběrů. Platí, že

$$E(s_B) = \sqrt{F_B^2 + F_C^2/N}.$$

Často je výhodnější použít odhad F_B^2 , kde se odečte příspěvek rozptylu uvnitř skupin, tedy

$$F_B^2 = s_B^2 + F_C^2/N.$$

Pokud vyjde rozdíl na pravé straně této rovnice roven nule, dosadí se definitoricky $F_B^2 = 0$. Při konstrukci klasických Shewhartových regulačních diagramů se obvykle předpokládá, že $F_B^2 = 0$. To lze snadno ověřit na základě podílu

$$V = s_B^2 N / \sigma_c^2,$$

který má v případě platnosti nulové hypotézy $H_0: \sigma_B^2 = 0$ F -rozdělení s $(M-1)$ a $M(N-1)$ stupni volnosti. Pomocí F -testu lze tedy ověřit platnost hypotézy H_0 . Pro případ konstrukce diagramů CUSUM je třeba odhadnout směrodatnou odchylku skupinových průměrů σ_e . Pokud lze použít $\sigma_B = 0$, je $F_e = F_C/\sqrt{N}$, ale pokud je $\sigma_B > 0$, je nutné uvažovat, že $F_e = s_B$. Posouzení velikosti σ_B^2 je tedy pro řadu situací zcela *nezbytné*. Dalším zdrojem

variability může být chyba měření, vyjádřená rozptylem σ_M^2 , která se obvykle slučuje na rozptyl σ_C^2 . Celkový rozptyl je pak jejich součtem a odpovídající směrodatná odchylka je zvětšena o faktor $f = \sqrt{1 + \sigma_M^2 / \sigma_C^2}$. Pokud bude platit, že $\sigma_M \neq \sigma_C/3$, neprojeví se chyby

měření výrazně na úrovni rozptylu, vyjde maximální hodnota $f = \sqrt{1 + 1/9} = 1.054$.

(c) Posledním *frekventovaným modelem* je případ autokorelace prvního řádu. To vlastně znamená porušení předpokladu nezávislosti jednotlivých výběrů. To je typický model situace, kdy data tvoří časovou řadu, nebo když se některé procesní parametry resp. externí parametry mění zvolna. Model autokorelace prvního řádu má tvar

$$x_{ij} = d + W_i + \epsilon_{ij}, \text{ kde } W_i = D + W_{i-1} + F_B T_i.$$

Platí, že $W_0 = 0$. Autokorelačním koeficientem je korelační koeficient mezi dvojicemi x_{Si} a $x_{S(i+1)}$, $i = 1, \dots, M-1$. Z této rovnice lze snadno určit, že

$$D(W) = F_B^2 / (1 + D^2).$$

Využitím výsledků pro předchozí modely lze pak vyjádřit rozptyl měření

$$D(x_{ij}) = \frac{F_B^2}{1 + D^2} + F_C^2$$

a rozptyl výběrových průměrů $D(x_{Si}) = \frac{F_B^2}{1 + D^2} + F_C^2 / N$. Je patrné, že v případě

výrazné autokorelace dojde ke zvýšení rozptylu výběrových průměrů, což by mělo za následek zvětšení regulačních mezí a ztrátu citlivosti regulačních diagramů vůči změnám stavu procesu. Autokorelační koeficient ρ se obvykle odhaduje pomocí vztahu

$$D = \frac{\sum_{i=1}^{M-1} (x_{Si} - \bar{x}) (x_{S(i+1)} - \bar{x})}{\sum_{i=1}^{M-1} (x_{Si} - \bar{x})^2 + \sum_{i=1}^{M-1} (x_{S(i+1)} - \bar{x})^2}.$$

Orientačně platí, že pokud je $2/\sqrt{M} \neq D < 2/\sqrt{M}$, lze považovat ρ za nevýznamné. Přesnější testy významnosti jsou uvedeny v cit.¹⁶. Pro posouzení autokorelace lze použít také modifikovaný von Neumanův poměr

$$V^2 = \frac{\sum_{i=1}^{M-1} (x_{Si} - \bar{x})(x_{S(i+1)} - \bar{x})}{2 \sum_{i=1}^{M-1} (x_{Si} - \bar{x})^2}.$$

Lze ukázat, že pro velká M má veličina V^2 přibližně normální rozdělení s parametry

$$E(V^2) = 1, \quad D(V^2) = \frac{M + 2}{M^2 + 1} \approx \frac{1}{M}.$$

Orientačně platí, že pokud leží V^2 mimo interval $1 \pm 2/\sqrt{M+2}$, nejsou výběry nezávislé. Pokud vyjde V^2 menší než spodní mez tohoto intervalu, projevuje se v datech trend nebo pomalé cyklické změny. Pokud vyjde V^2 větší než horní mez tohoto intervalu, projevují se v datech rychlé cyklické změny.

Z uvedeného výkladu je patrné, že při určování směrodatné odchylky výběrových průměrů je třeba postupovat opatrně a nejdříve testovat autokorelaci, resp. významnost rozptylu σ_B^2 .