Prosessin suorituskykyindeksit

Jarkko Rosnell

Tilastotieteen pro gradu -tutkielma
25. huhtikuuta 1997

Jyväskylän yliopisto
tilastotieteen laitos

Tiivistelmä

Jarkko Rosnell: *Prosessin suorituskykyindeksit* Tilastotieteen pro gradu -tutkielma, Jyväskylän yliopisto, 25.huhtikuuta 1997. Sivuja 61, liitteitä -.

Tässä tutkielmassa käsitellään prosessin suorituskykyindeksejä, ne ovat oikein käytettynä erittäin käyttökelpoisia suhdelukuja kuvaamaan teollisten prosessien suorituskykyä. Indeksejä on käytetty, muunneltu ja jopa useasti kritisoitu viimeisen vuosikymmenen aikana. Yksimuuttujaisten ja monimuuttujaisten indeksien laskemisen perusedellytys ja luotettavuuden tae on, että prosessi on hallinnassa, ja havainnot noudattavat normaalijakaumaa. Epänormaaleihin tapauksiin on myös kehitetty monenlaisia ja -muotoisia indeksejä, myös Bootsrap-menetelmillä saadaan oivia sovelluksia indekseistä. Indekseihin, varsinkin yksinkertaisimpiin yksimuuttujaisiin, kannattaa suhtautua kuitenkin pienellä varauksella. Arvoihin ei sinänsä kannata tuijottaa orjallisesti, terve harkinta yleensä auttaa ongelmatilanteissa. Vielä kun pitää mielessä mitä indeksit merkitsevät ja mitä ne oikeastaan mittaavat. Kysehän on loppujen lopuksi siitä, että tuotteet pyritään tekemään toleranssien mukaisiksi, ja indekseillä mitataan kuinka hyvin tämä ominaisuus toteutuu.

Avainsanoja: suorituskyky, indeksi, toleranssialue, tavoitearvo, hajonta, luottamusväli, suhde

Sisällysluettelo

Luku 1

Johdanto

Ishikawan (1990) mukaan tuotantoprosesseissa syntyvien tuotteiden vaihtelu aiheutuu kahdesta erilaisesta syystä, ja siksi on myös kahdenlaista vaihtelun tyyppiä. Ensimmäinen vaihtelun tyyppi on sellaista, joka aiheutuu huolimatta siitä, että jokainen prosessin kanssa tekemisissä oleva henkilö toimii samalla tavoin kuten aina annettujen ohjeiden ja standardien mukaan. Nämä ovat syitä, jotka eivät ole teknisesti vielä hallinnassa, mutta on nykyään melkein pystytty määrittelemään teoreettisesti. Näitä syitä on kutsuttu myös väistämättömiksi (*unavoidable causes*) tai satunnaisiksi syiksi (*chance causes*), ja näistä johtuvaa vaihtelua on kutsuttu hallituksi vaihteluksi (*controlled variability*). Kun kaikki työ on tehty olemassa olevien normien mukaan, johtajat eivät voi syyttää työntekijöitään näistä syistä. Vain johdon puuttuminen asioihin ja niiden resursointi voi vähentää näistä yleisistä syistä johtuvaa vaihtelua.

Vastakohtana on toinen syiden tyyppi, erityiset syyt (*special causes*), jotka aiheuttavat jotain poikkeavuutta prosessissa. Tuloksena on erityisen suuri vaihtelu, kun jokin tapahtuma ei täytä työn standardeja, tai annettuja työn normeja ei ole noudatettu. Tällaiset syyt voidaan eliminoida käyttämällä teknologiaa, jos jokainen mukana olija tekee omalta osaltaan ponnisteluja sen suhteen. Niitä kutsutaan myös väistettävissä oleviksi (*avoidable*) tai määrätyiksi syiksi (*assignable causes*). Näistä syistä johtuvaa vaihtelua prosessissa on kutsuttu hallitsemattomaksi (*uncontrolled*) vaihteluksi. Jos työn standardit ja muut normit ovat virheettömiä, niin hallitsematon vaihtelu voisi merkitä sitä, että työntekijät eivät ole tehneet niin kuin heitä on käsketty. Tällainen vaihtelun tyyppi osoittaa, että esimerkiksi normin mukaisia raaka-aineita ei ole käytetty, laitteet ja työkalut ovat merkittävästi kuluneet, tai käytetyt mittalaitteet ovat menettäneet luotettavuutensa. Kun hallitsematonta vaihtelua esiintyy, on johdon ammattitaitoisuudesta riippuvaista tekevätkö he toimenpiteitä, jotka varmistaisivat, että työ tehdään normien mukaan. Kuinka ollakaan nämä normit eivät yleensä ole täydellisiä, ja näiden syiden vähentämiseen tarvitaan johtajien ja työntekijöiden yhteisiä ponnisteluja.

Kuten jo aikaisemmin on mainittu nämä kaksi syiden tyyppiä aiheuttavat kahdenlaista vaihtelua valmiissa tuotteissa. Kun työ tehdään normien mukaan, ja luotettavat mittaukset on tehty, tuotteiden laadun vaihtelu aiheutuu satunnaisista syistä. Näille syille voidaan löytää jokin kiinteä jakauma, kuten normaalijakauma. Tämä tilanne tunnetaan hallittuna vaihteluna. Prosessin tilaa, joka tuottaa tuloksia, joissa vain hajonta on hallittua vaihtelua on kutsuttu hallituksi tilaksi (*controlled state*). Prosessi on silloin hallinnassa (*in control*), kun prosessin taso ja hajonta säilyvät samoina eli prosessi on stationaarinen.

Toisaalta, kun määrätyt syyt kasvavat, ja poikkeavuutta ilmenee prosessissa, niin prosessin hajonta tulee epänormaalin suureksi. Tällaista epänormaalin suuruista hajontaa on kutsuttu hallitsemattomaksi vaihteluksi, ja prosessin tilaa joka tuottaa tuloksenaan sellaisen vaihtelun on kutsuttu hallitsemattomaksi tilaksi.

Prosessin suorituskyvyn tutkimukset ovat laadunohjauksen kulmakivi. Kun tutkimme suorituskykyä olemme kiinnostuneita tuotteiden tärkeiden ominaisuuksien vaihtelusta. Prosessin suorituskyky voidaan määritellä suorituskyvyksi tiettynä ajanjaksona samalla, kun prosessi on tilastollisesti hallitussa tilassa. Se on tavallisesti ilmaistu laadun jakaumana, tai virheiden lukumääränä. Suorituskyky ilmaisee ennemminkin prosessin tuloksia tai luonteenomaisten arvojen jakaumaa. Kun prosessi koostuu yksittäisestä koneesta, sen laadullisena suorituskykyä on kutsuttu koneen suorituskyvyksi tai koneen tarkkuudeksi. Hallinnassa olevan prosessin suorituskyky on muunmuassa silloin sitä parempi mitä suurempi on muuttujien vaihtelualueen ja toleranssialueen suhde.

Termiä tuotannon kapasiteetti (*production capacity*) on käytetty yleensä määrällisen suorituskyvyn vastineena ja tätä termiä täytyy varoa sekoittamasta prosessin suorituskykyyn, joka merkitsee laadullista suorituskykyä. Niin kauan kuin laadunohjausta toteutetaan, on tärkeää tutkia prosessin laadullista suorituskykyä. Suorituskyvyn suhteellisessa tutkimisessa prosessin suorituskykyindeksit toimivat oivina apuvälineinä.

Tässä tutkielmassa käsitellään siis prosessin suorituskykyindeksejä. Yksimuuttujaisista suorituskykyindekseistä esitellään aluksi tavallisimmin käytetyt indeksit ja niiden muunnelmat, mutta tuodaan esille myös harvemmin käytettyjä indeksejä. Sen jäkeen tuodaan esille niiden yhteys normaalijakauman todennäköisyysalueeseen. Suorituskykyindeksejä on hyvä käsitellä myös satunnaismuuttujina, tavallisimmille indekseille esitetään jakaumat ja luottamusvälit. Tutkielman ehkä haasteellisin alue on tilanne, jossa prosessin jakauma ei ole enää normaali. Mikä on prosessin suorituskyky silloin ? Entä indeksit ja mikä on niiden luotettavuus ? Hämärän verhoa pyritään raottamaan muunnelmilla ja jakaumien sovittamisella. Tutkielman havainto-osassa on käytetty Vammas Defencetec Oy:ltä saatua 40 millisen ilmatorjuntatykin hylsyjen ominaisuuksien aineistoa.

Luku 2

Historiallista taustaa laadunohjaukseen

Sanaa laatu käytetään nykyään turhankin usein eri tilanteissa, tuntuu siltä kuin laadulla olisi jokin maaginen olemus, mitä se sitten oikein todella on. Määritelmiä on monia riippuen siitä kenen näkökannasta on kyse. Lause tehdä hyvänlaatuisia tuotteita ymmärretään usein niin, että se on sama kuin tehdä tuotteelle paras mahdollinen laatu, näin ei kuitenkaan ole. Kun puhumme laadusta laadunvalvonnassa, hyvä laatu tarkoittaa parasta laatua, jota yritys voi tuottaa nykyisen tuotannon, teknologian ja prosessin suorituskyvyn puitteissa, tyydyttäen samalla kuluttajien tarpeet. Laatua voidaan ajatella myös tilastollisena jakaumana, kun on mahdotonta valmistaa tuotteita ilman vaihtelua. Laadun valvonnassa asetamme rajat tuotteiden ominaisuuksien vaihtelulle, ja valvomme että rajat eivät ylity.

Jo varhaisen antiikin aikana olivat ihmiset valvoneet laatua, ilman tällaista toimenpidettä olisi tuskin kyetty rakentamaan massiivisia temppeleitä ja palatseja vuosisatoja kestäviksi. Erilaisten rakennusmateriaalien koko ja laatu oli jo tuolloin tarkoin määrätty, esimerkiksi egyptiläisten käyttämien rakennustiilien muoto ja valmistusprosessi oli standardisoitu.

1200-luvulla eurooppalaiset saman ammattikunnan käsityöläiset järjestäytyivät killoiksi. Në pitivät kurissa jäsentensä kilpailun ja tekivät vaatimuksia laadun ja taidon suhteen. Kiltojen ideana oli, että nuori mies voisi oppia käsityöammatin salat mestarin oppipoikana. Mestarin arvon sai vain, jos täytti killan vaativat kriteerit. Varhaiset siirtolaiset veivät myös kiltasysteemin mukaanaan meren yli uudelle mantereelle nykyiseen Pohjois-Amerikkaan.

1700-luvun lopulla muutosten tuulet puhalsivat läpi vanhan mantereen. Liberalismin aatteen vaikutuksesta kiltojen merkitys väheni olennaisesti. Tuotteiden valmistus siirtyi pienistä pajoista suuriin tehtaisiin. Tätä historiallisesti merkittävää ajanjaksoa on kutsuttu teolliseksi vallankumoukseksi. Käsityöläisistä tuli tehtaiden työväkeä ja entisistä killan mestareista laadunvalvojia tai osuvammin sanottuna laaduntarkkailijoita. Tuotannon suunnittelu ja toteutus erkaantuivat toisistaan itsenäisiksi yksiköiksi.

Massatuotantoon siirtyminen toi esiin uusia laadullisia ongelmia. Kun yksittäiset käsityöläiset olivat valmistaneet tuotteensa alusta loppuun, niin tehtaiden työntekijät olivat vastuussa vain pienistä elementeistä. Työväeltä puuttui käsityöläisten ammattiylpeys. Oli vaikea tuntea mielihyvää tekemästään tuotteesta tai sen osasta varsinkaan, jos ei nähnyt työnsä tulosta valmiina. Kun työ oli vielä usein saman vaiheen toistamista, niin se aiheutti motivoitumattomissa tekijöissään usein kyllästymistä. Näiden asioiden seurauksena valmiiden tuotteiden laatu kärsi, mikä voidaankin nimetä massatuotannon yhdeksi suureksi haitaksi.

Alkusysäys tilastolliseen laadunohjaukseen (*Statistical Quality Control*) saatiin 1920-luvun puolivälissä. Aikaisemmin laatua oli valvottu pelkästään tarkastamalla näytteitä. Nyt havaittiin, että on tärkeämpää tutkia käynnissäolevaa prosessia ja sen muuttujia, kuin tarkastaa pelkästään valmiita tuotteita. Ottamalla huomioon tämän asian Walter A. Shewhart kehitti Bell Telephone Labs'issa 1924 ensimmäisen modernin luonnoksen valvontakortista (Wetherill & Brown 1991). Siitä tuli merkittävä työkalu teolliseen laadunvalvontaan. Vuonna 1931 Shewhart julkaisi kuuluisan klassikkonsa *Economic Control of Quality of Manufactured Product.* Teoksessaan Shewhart esitteli tilastollisen prosessin ohjauksen (*Statistical Process Control*) perusperiaatteet.

Laadunvalvonta teki suuren maihinnousun Japaniin 1950 (Devor, Chang & Sutherland 1992). Japanilaiset tiedemiehet ja teollisuusyritysten johtajat kutsuivat W. E. Demingin opettamaan kyseistä menetelmää heille. Demingin oppi oli, että laadun parannus vähentää kustannuksia ja kasvattaa samalla työn tuottavuutta. Japanilaiset ottivat tämän filosofian todella omakseen. Heidän taloudellisesti vahva asemansa ja huipputeknologinen osaamisensa voidaan ainakin osittain selittää sillä tosiasialla, että teollisuusyritysten johtajisto sisäisti laadun ohjauksen merkityksen. Kolme sanaa "*Made in Japan*" saattoi merkitä muutama vuosikymmen sitten heikkolaatuista tuottetta (Ryan 1989), mutta viimeisen vuosikymmenen aikana se on ollut merkkinä erittäin korkealaatuisille tuotteille.

Vasta 1980-luvun alussa länsimaat ymmärsivät kokonaisvaltaltaisen laadun hallinnan (*Total Quality of Managemet*) merkityksen. Siitä lähtien ja tähän päivään saakka laatu on ollut kuuma puheenaihe. Vuonna 1987 ISO (*International Organization for Standardization*) julkaisi laadun vakuuden normit sisällään pitävän standardin ISO 9000 (Wierda 1994). Kiinnostus laatuun, laadunohjaukseen ja laadun hallintaan (*quality management*) säilyy varmasti jatkossakin, ja sen mitä tulevaisuus tuo mukanaan, voi vain aavistaa.

Tilastollinen laadunohjaus jakaantuu kahteen osaan valvontakortteihin (*control charts*) ja suorituskyvyn tutkimiseen (*capability analyses*). Suorituskyvyn tutkimisessa prosessin suorituskykyindeksit ovat erittäin käyttökelpoisia välineitä. Ensimmäiset suorituskykyindeksit otettiin käyttöön 1980-luvun puolivälissä, ja ne ovat vakiintuneet prosessin suorituskyvyn mitaksi määrittely rajoihinsa nähden. Nykyään suorituskykyindeksit ovat itseään arvostavissa teollisuuslaitoksissa ja yrityksissä arkipäivää, ne toimivat tavallaan yhtiön näyteikkunana asiakkaille, jotka asettavat vaatimuksia tilaamiensa tuotteiden, sekä niiden valmistuksessa käytettyjen prosessien suhteen. Kritiikkiäkin on indeksien suhteen esitetty, eikä aina aivan syyttä.

Luku 3

Yksimuuttujaiset suorituskykyindeksit

Prosessin suorituskykyindeksit (*Process Capability Indices*) ovat eräänlaisia suhdelukuja, joilla hallinnassa olevan prosessin suhteellista suorituskykyä voidaan kuvata numeromuodossa. Suorituskykyindeksin yksinkertaisena määritelmänä voidaan pitää toleranssivälin suhdetta prosessin hajontaan. Pystyvässä eli suorituskykyisessä prosessissa, jonka hajonta on pientä, syntyy vain vähän viallisia ja käyttökelvottomia, toleranssivälin ulkopuolelle päätyviä tuotteitta. Milloin tuotannon virhe on taloudellisesti niin suuri, että huomattava osa tuotannosta asettuu määriteltyjen toleranssien ulkopuolelle? Sen ilmaisee suorituskykyindeksi. Mitä suurempi indeksin arvo on, sitä parempi on prosessin suorituskyky ,ja tuotannon hävikki on vastaavasti pienempi.

Banksin (1993) mukaan indeksit palvelevat kolmea erilaista päämäärää.

1. Indeksien avulla voidaan tutkia prosessiin vaikuttavien tekijöiden kuten paineen, lämpötilan tai erilaisten materiaalikoostumusten vaikutuksia prosessin suorituskykyyn.

2. Indeksien avulla voidaan verrata eri prosesseja keskenään.

3. Indekseillä voidaan karkeasti arvioida prosessissa syntyvien viallisten tuotteiden lukumäärää.

Tässä luvussa esitettyjen suorituskykyindeksien laskemisen perustana on seuraavat asiat:

- Vaatimuksena on, että prosessi pitää olla hallinnassa.
- Oletuksena on, että prosessin muuttujat noudattavat normaalijakaumaa.
- Prosessissa on ennalta määrätyt toleranssirajat (*specification limits*) L alempi ja U ylempi.
- Toleranssirajat muodostavat toleranssivälin (L,U), jolla mitatun muuttujan X arvojen sallitaan vaihdella.
- Tuotteita, joiden kiinnostuksen kohteena olevien ominaisuuksien havaitut arvot eivät asetu toleranssirajojen sisälle (*nonconformings products*) kutsutaan hävikiksi ja merkitään jatkossa lyhyesti NC tuotteiksi.

Prosessin hajontaa merkitään parametrilla σ ja prosessin odotusarvoa parametrilla ξ. Parametrit ξ ja σ eivät käytännössä ole aina tunnettuja, ja niin ne joudutaankin estimoimaan prosessista poimituista satunnaisotoksista. Parametrin μ estimaattorina käytetään otoksesta laskettua prosessin keskiarvoa $\bar{X}$ ja vastaavasti prosessin hajonnan σ estimaattina käytetään otoshajontaa S, jolle

$$S^2 = \frac{1}{n-1}\sum_{j=1}^{n}(X_j - \bar{X})^2. \tag{3.1}$$

3.1 Indeksi C_p

Yleisin prosessin suorituskyvyn mitta on indeksi

$$C_p = \frac{U-L}{6\sigma}. \tag{3.2}$$

Indeksi C_p riippuu siis toleranssivälin lisäksi prosessin hajonnasta. Suuri indeksin arvo merkitsee suorituskykyistä prosessia ja vastaavasti pieni indeksin arvo suorituskyvytöntä. Kun prosessin suorituskykyä pyritään parantamaan, niin käytännössä ainoa mahdollisuus siihen on prosessin hajonnan pienentäminen. Toleranssivälin leventäminen tietenkin kasvattaa indeksien arvoja, mutta tämä keinotekoinen arvojen nostaminen tuottaa prosessissa kuitenkin tuotteita, joissa on entistä suurempi vaihtelu. [1]

[1] Mainittakoon tämän tavan olevan kaukana tilastollisen laadunohjauksen filosofiasta. Varsin selkeä syy näiden keinottelujen käyttämiseen on sokea tulosjohtaminen, ja huonosti suunniteltu palkitsemisjärjestelmä.

Jos prosessin jakauma on normaali ja muuttujan X odotusarvo $\xi = \frac{1}{2}(L + U)$ eli toleranssivälin keskipiste, niin NC tuotteiden osuus on $2\Phi(-d/\sigma)$, missä $d = \frac{1}{2}(L + U)$ on puolet toleranssivälin pituudesta.

Indeksin jakaja 6σ eli väli $(-3\sigma, +3\sigma)$ peittää 99.73% normaalijakauman $X \sim N(\xi, \sigma)$ tiheysfunktiosta, sillä kun $X \sim N(0, 1)$

$$
\begin{aligned}
P(-3\sigma \leq X \leq 3\sigma) &= P(-3 \leq Z \leq 3) \\
&= \Phi(3) - \Phi(-3) \\
&= \Phi(3) - (1 - \Phi(3)) \\
&= 2\Phi(3) - 1 \\
&= 2 \cdot 0.9965 - 1 \\
&= 0.9973.
\end{aligned}
$$

Joten suorituskykyindeksin $C_p = 1$ arvo tarkoittaa, että valmiiden tuotteiden ominaisuuksista 99.73% mahtuu toleranssivälille (L, U). NC tuotteiden osuus on vastaavasti 0.27%. Kun prosessin suorituskykyindeksin C_p arvo saadaan nostettua 1.33, niin käytännössä prosessin hajontaa on saatu pienennettyä. Tuolloin toleranssiväli peittää 99.994% normaalijakauman tiheysfunktiosta, prosessin NC tuotteiden osuus on todennäköiseti enää vain 0.006%. Tämä tarkoittaa sitä, että miljoonan kappaleen tuotantoerässä olisi vain noin 60 tuottetta, jotka eivät täyttäisi niille asetettuja kriteerejä.

Ishikawa (1990) jakaa indeksin C_p arvot viiteen eri luokkaan

- **Erityisluokka** $C_p > 1.67$: erittäin korkea luotettavuuden taso, tällainen laatu saattaa olla liian korkea yleisiin tarkoituksiin.
- **Luokka A** $1.33 < C_p \leq 1.67$: Erittäin hyvä laatu. Tarkastamista voidaan vähentää.
- **Luokka B** $1.0 < C_p \leq 1.33$: Melko hyvä laatu. Näytteiden tarkastus on riittävä.
- **Luokka C** $0.67 < C_p \leq 1.00$: Joitakin viallisia tuotteita valmistuu. C_p tulisi nostaa 1.00 tai sen yläpuluolelle.
- **Luokka D** $C_p \leq 0.67$: Todella huono.

Montgomery(1985) suosittelee Kotz & Johnsonin (1993) mukaan seuraavia minimiarvoja indeksille C_p: Olemassa oleville prosesseille 1.33 ja uusille 1.50. Tuotteiden, joiden ominaisuuksiin liittyy olennaisena osana turvallisuus suosituksena on minimiarvo 1.50 olemassa oleville prosesseille, ja 1.67 uusille prosesseille. Kotz & Johnson (1993) painottavat, että on parempi käyttää indeksien arvojen suhdetta NC tuotteiden osuuteen, milloin vain tämä on mahdollista. Yleisesti ottaen indeksin C_p arvon pitää olla vähintään 1, jotta prosessia voidaan pitää jossain määrin suorituskykyisenä.

Prosessin keskiarvo pyritään aina saamaan mahdollisimman lähelle ennalta määrättyä tavoitearvoa T (*target value*), mutta indeksin C_p arvoon tavoitearvo ei kuitenkaan vaikuta. Kane (1986) määritteli parannuksen indeksille C_p huomioimalla tavoitearvon sijainnin toleranssivälillä

$$C_p^* = \min\left(\frac{T-L}{3\sigma}, \frac{U-T}{3\sigma}\right), \tag{3.3}$$

eli kun $U - T \neq T - L$. Jos $U - T = T - L$ niin silloin $C_p^* = C_p$.

3.2 Indeksi C_{pk}

Indeksi C_p riippuu toleranssivälin lisäksi ainoastaan prosessin varianssista, eikä sen laskemisessa tarvita tietoa prosessin odotusarvosta ξ. Tämän seurauksena on johdonmukainen virhe indeksin C_p arvoon, kun prosessi ei ole keskistetty. Jos prosessin odotusarvo ei ole toleranssivälin keskellä, indeksin C_p arvot eivät vastaa enää odotettujen NC tuotteiden osuutta tarkasti. Indeksi C_{pk} kehitettiin antamaan tietoa parametrin ξ arvojen vaikutuksesta suorituskykyindeksiin. Tämä indeksi määritellään

$$C_{pk} = \frac{\min(U-\xi, \xi-L)}{3\sigma}. \tag{3.4}$$

Koska

$$\min(a,b) = \frac{1}{2}(|a+b| - |a-b|) \tag{3.5}$$

niin saadaan

$$\begin{aligned} C_{pk} &= \frac{\min(U-\xi, \xi-L)}{3\sigma} \\ &= \frac{\frac{1}{2}(|(U-\xi)+(\xi-L)| - |(U-\xi)-(\xi-L)|)}{3\sigma} \end{aligned}$$

$$\begin{aligned}
&= \frac{\frac{1}{2}(|U-L| - |U-\xi-\xi+L|)}{3\sigma} = \frac{d - |\frac{1}{2}(L+U)-\xi|}{3\sigma} \\
&= C_p - \frac{|\frac{1}{2}(U+L)-\xi|}{3\sigma} = \left(1 - \frac{3\sigma|\frac{1}{2}(U+L)-\xi|}{3\sigma d}\right) C_p \\
&= \left(1 - \frac{|\frac{1}{2}(L+U)-\xi|}{d}\right) C_p. \qquad (3.6)
\end{aligned}$$

Lisäksi sovitaan, että indeksi $C_{pk} = 0$, jos $U - \xi \leq 0$ tai $\xi - L \leq 0$. Tämä tarkoittaa tilannetta, jolloin prosessin odotusarvo on toleranssivälin ulkopuolella, tällöinhän kyseisen prosessin suorituskyky on olematon. Kun $C_p = d/(3\sigma)$, niin $C_{pk} \leq C_p$. Indeksit ovat yhtäsuuria vain, kun $\xi = \frac{1}{2}(L+U)$.

Jos muuttujan X jakauma on normaali, niin NC tuotteiden osuus on

$$\Phi\left(\frac{L-\xi}{\sigma}\right) + \left\{1 - \Phi\left(\frac{U-\xi}{\sigma}\right)\right\}. \qquad (3.7)$$

Merkitsemällä

$$k = \frac{\frac{1}{2}(L+U)-\xi}{d} \qquad (3.8)$$

saadaan indeksi C_{pk} kaavan (3.6) avulla myös muotoon

$$C_{pk} = C_p(1-k). \qquad (3.9)$$

Teollisuudessa sattuu usein tilanteita, joissa tarvitaan vain yksipuoleinen toleranssiraja. Indeksit tällaisiin tapauksiin määritellään alemmalle toleranssirajalle

$$C_{pkl} = \frac{\xi - L}{3\sigma} \qquad (3.10)$$

ja ylemmälle toleranssirajalle

$$C_{pku} = \frac{U-\xi}{3\sigma}. \qquad (3.11)$$

Eli C_{pkl} ja C_{pku} ovat suorituskykyindeksin C_{pk} osatekijöitä.

Prosessin keskiarvo pyritään aina saamaan mahdollisimman lähelle ennalta määrättyä tavoitearvoa T (*target value*). Indeksit C_p ja C_{pk} eivät kuitenkaan huomioi prosessin keskiarvon läheisyyttä tavoitearvoon T. Kane (1986) muunsi myös indeksin C_{pk} tilanteeseen $U - T \neq T - L$

$$C^*_{pk} = \min(C^*_{pl}, C^*_{pu}), \qquad (3.12)$$

missä

$$C^*_{pl} = \frac{T-L}{3\sigma}\left(1 - \frac{|T-\xi|}{T-L}\right) \tag{3.13}$$

ja

$$C^*_{pu} = \frac{U-T}{3\sigma}\left(1 - \frac{|T-\xi|}{U-T}\right). \tag{3.14}$$

Indeksi C^*_{pk}=0, jos $|T-\mu| > T-L$ tai $|T-\mu| > U-T$. Se voidaan myös esittää muodossa

$$C^*_{pk} = C^*_p(1-k^*), \tag{3.15}$$

missä $k^* = |T-\xi|/\min(T-L, U-T)$. Myös $C^*_{pk} = 0$ jos $k^* \geq 1$.

3.3 Indeksi C_{pm}

Chan, Cheng & Spring (1988) esittivät prosessin suorituskykyindeksin C_{pm} käyttämällä varianssin $\sigma^2 = E(X-\mu)^2$ sijasta keskineliövirhettä $E(X-T)^2 = \sigma^2 + (\xi - T)^2$ eli **varianssi + harhan neliö**. Indeksin C_{pm} alkuperäinen määritelmä on

$$C_{pm} = \frac{U-L}{6(\sigma^2 + (\xi - T)^2)^{\frac{1}{2}}} = \frac{d}{3(\sigma^2 + (\xi - T)^2)^{\frac{1}{2}}}. \tag{3.16}$$

Usein tavoitearvo T on toleranssivälin keskipiste eli $m = \frac{1}{2}(U+L)$.

Kun merkitään $\tau = E[(X-T)^2] = \sigma^2 + (\xi - T)^2$ saadaan indeksistä C_{pm} sievempi muoto

$$C_{pm} = \frac{U-L}{6\tau} = \frac{d}{3\tau}. \tag{3.17}$$

Indeksi C_{pm} huomioi prosessin vaihtelun sekä odotusarvon sijainnin tavoitearvoon T nähden. Jos merkitään

$$\zeta = \frac{\xi - T}{\sigma}, \tag{3.18}$$

niin indeksi C_{pm} voidaan esittää myös muodossa

$$C_{pm} = \frac{C_p}{\sqrt{1+\zeta^2}}, \tag{3.19}$$

sillä

$$
\begin{aligned}
C_{pm} &= \frac{U-L}{6\tau} = \frac{U-L}{6\sqrt{\sigma^2+(\xi-T)^2}} \\
&= \frac{(U-L)/\sigma}{(6\sqrt{\sigma^2+(\xi-T)^2})/\sigma} = \frac{\frac{U-L}{6\sigma}}{\frac{\sqrt{\sigma^2+(\xi-T)^2}}{\sigma}} \\
&= \frac{C_p}{\sqrt{\frac{\sigma^2+(\xi-T)^2}{\sigma^2}}} = \frac{C_p}{\sqrt{\frac{\sigma^2}{\sigma^2}+\frac{(\xi-T)^2}{\sigma^2}}} \\
&= \frac{C_p}{\sqrt{1+\left(\frac{\xi-T}{\sigma}\right)^2}} \\
&= \frac{C_p}{\sqrt{1+\zeta^2}}.
\end{aligned}
$$

Kun tavoitearvo ei ole toleranssialueen keskipisteenä eli kun $U-T \neq L-T$, määrittelivät Chan ym. (1988) indeksin C_{pm} muunnelman C^*_{pm} käyttämällä apuna kaavaa (3.5)

$$
C^*_{pm} = \frac{\min(U-T, T-L)}{3(\sigma^2+(\xi-T)^2)^{\frac{1}{2}}} = \frac{d-|T-\frac{1}{2}(U+L)|}{3(\sigma^2+(\xi-T)^2)^{\frac{1}{2}}}. \tag{3.20}
$$

Jos $T = \frac{1}{2}(U+L)$ kuten usein on, niin $C^*_{pm} = C_{pm}$.

$$
C^*_{pm} = \min\left(\frac{T-L}{3\tau}, \frac{U-T}{3\tau}\right). \tag{3.21}
$$

Choin & Owenin (1990) mukaan Boyles (1990) teki Taguchin (1986) innoittamana seuraavan muunnelman indeksistä C_{pm}

$$
C^+_{pm} = \frac{U-L}{6\sqrt{E(L)}}, \tag{3.22}
$$

missä $E(L)$ merkitsee hävikkifunktion

$$
L(X,T) = \begin{cases} k_1(X-T)^2 & \text{,kun } X \le T, \\ k_2(X-T)^2 & \text{,kun } X \ge T \end{cases}
$$

odotusarvoa. Näin saadaan

$$
E(L) = \sigma^2((1+\zeta^2)(k_1(1-\Phi(\zeta)) + k_2\Phi(\zeta)) - (k_1-k_2)\zeta\phi(\zeta)), \tag{3.23}
$$

missä $k_1 = (r_2/r_1)k_0$, $k_2 = (r_1/r_2)k_0$, $r_1 = (T-L)/(U-L)$, $r_2 = (U-T)/(U-L)$, $k_0 = \max(r_1/r_2, r_2/r_1)/(2(r_1^2+r_2^2))$, Φ on standardoidun normaalijakauman kertymäfunktio ja ϕ on vastaavan jakauman tiheysfunktio. Symmetrisessä tapauksessa $r_1 = r_2 = 1/2$ ja $E(L) = \zeta$ joten $C_{pm}^+ = C_{pm}$. Kannattaa panna merkille, että k_0 on määritelty niin, että $C_p = C_{pm}^+$, kun $\xi = T$.

3.4 Indeksi C_{pn}

Indeksit C_{pm}, C_{pm}^* ja C_{pm}^+ ottavat huomioon prosessin vaihtelun lisäksi myös prosessin odotusarvon sijainnin tavoitearvoon T nähden. Ne eivät kuitenkan huomioi prosessin odotusarvon sijaintia välillä (L, U). Suure τ kuvaa yhtäaikaa prosessin odotusarvon vaihtelua ja prosessin keskiarvon läheisyyttä tavoitearvoon T. Pysyvän vaihtelun alaisena prosessin viallisten tuotteiden osuus on silloin pienin, kun odotusarvo on välin (U, L) keskellä. Silloin $\min(\xi - L, U - \xi)$ kuvastaa prosessin keskiarvon keskistyneisyyttä. Näin ollen sillä voidaan ilmaista viallisten NC tuotteiden lukumäärää eli niiden tuotteiden osuutta, joiden mitatut ominaisuudet eivät mahdu toleranssialueelle (U, L). Nämä seikat huomioimalla Choi & Owen (1990) määrittelivät prosessin suorituskykyindeksin

$$C_{pn} = \min(C_{pnl}, C_{pnu}), \tag{3.24}$$

missä

$$C_{pnl} = \frac{\xi - L}{3\tau}, \tag{3.25}$$

$$C_{pnu} = \frac{U - \xi}{3\tau}. \tag{3.26}$$

Lisäksi he määrittelivät myös, että $C_{pn} = 0$, jos $C_{pnl} \leq 0$ tai $C_{pnu} \leq 0$. Indeksille C_{pn} voidaan helposti löytää yhteys indeksiin C_{pm}

$$C_{pn} = C_{pm}(1-k), \tag{3.27}$$

missä k on määritelty aikaisemmin. Samoin kuin indeksit C_{pk} ja C_{pk}^*, niin indeksi $C_{pn} = 0$ kun $k > 1$. Voimme ilmaista indeksin C_{pn} myös hävikkifunktion avulla

$$C_{pn} = \frac{U - L}{6\sqrt{E(L)}}, \tag{3.28}$$

missä $E(L)$ merkitsee hävikkifunktion

$$L(X,T) = w(\xi)(X-T)^2 \tag{3.29}$$

painotettua odotusarvoa ja $w(\xi) = (1-k)^{-2}$, jos $k < 1$, ja $w(\xi) = \infty$, jos $k \geq 1$. Edellinen määritelmä on sama kuin C^+_{pm}, mutta eroaa hieman hävikkifunktion osalta. Indeksillä C^+_{pm} funktio $w(\xi)$ on pelkästään vakio k_1 tai k_2 kuten edellä on annettu. Voimme nähdä painofunktion $w(\xi)$ kuvaavan prosessin odotusarvon keskistyneisyyttä suureen $|M-\xi|$ avulla. Toisaalta funktion $w(\mu)$ avulla voidaan arvioida tuotteiden hävikkiä kokonaistuotannossa.

Luku 4

Suorituskykyindeksit satunnaismuuttujina

Tässä luvussa oletetaan, että luvussa 3 annetut perusteet suorituskykyindeksien laskemiselle, kuten normaalisuus, ovat edelleen voimassa. Suorituskykyindeksien laskemisessa tarvitaan tietoa prosessin odotusarvosta ja hajonnasta. Useimmiten näitä tunnuslukuja ei voida suoraan laskea, vaan ne joudutaan estimoimaan prosessista poimitusta otoksesta. On syytä muistaa, että indeksien luotettavuus riippuu prosessin hajonnan ja odotusarvon hyvistä estimaateista. Oletetaan, että mitattu satunnaismuuttuja X on normaalisti jakautunut ja sen odotusarvo on ξ ja keskihajonta on σ. Olkoon $(X_1, X_2, ..., X_n)$ n suuruinen satunnaisotos prosessista. Otoskeskiarvo $\bar{X}$ ja otosvarianssi S^2 ovat otoksesta laskettuja prosessin parametrien estimaattoreita.

4.1 Indeksin C_p estimointi

Ainoa parametri, joka indeksin C_p kaavassa (3.2) tarvitsee estimoida on σ, eli muuttujan X keskihajonta. Parametrin σ luonnollinen estimaattori on

$$\hat{\sigma} = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(X_j - \bar{X})^2}, \tag{4.1}$$

missä

$$\bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_j \tag{4.2}$$

eli muuttujan X keskiarvo. Jos muuttujan X jakauma on normaali, niin parametrin $\hat{\sigma}^2$ jakauma on

$$\frac{1}{n-1}\sigma^2\chi^2_{n-1}. \tag{4.3}$$

Tällöin indeksille

$$C_p^{-1} = \frac{6}{U-L}\sigma = \frac{3}{d}\sigma$$

saadaan suoraan

$$\hat{C}_p^{-1} = \frac{6}{U-L}\hat{\sigma} = \frac{3\hat{\sigma}}{d} \sim C_p^{-1}(\chi_{n-1}/\sqrt{(n-1)}).$$

Analyysi indeksin C_p estimaatorille

$$\hat{C}_p = \frac{U-L}{6\hat{\sigma}} = \frac{d}{3\hat{\sigma}} = \frac{\sigma}{\hat{\sigma}}C_p, \tag{4.4}$$

joka noudattaa jakaumaa $C_p\sqrt{n-1}/\chi_{n-1}$, on monimutkaisempi. Tätä on käsitelty useissa julkaisuissa muun muassa Kane (1986) ja Chou & Owen (1989). Seuraavana tutkimustulosten lyhyt esittely estimaattorin $\hat{C}_p$ jakaumasta olettaen, että muuttuja X on normaalisti jakautunut. Kaavoista (4.3) ja (4.4) saadaan

$$P\left(\frac{\hat{C}_p}{C_p} > c\right) = P(\chi^2_{n-1} < (n-1)c^{-2}). \tag{4.5}$$

Merkitsemällä $P(\chi^2_{n-1} \leq \chi^2_{n-1,\varepsilon}) = \varepsilon$, saadaan

$$P\left(\frac{\sigma^2}{n-1}\chi^2_{n-1,\alpha/2} \leq \hat{\sigma}^2 \leq \frac{\sigma^2}{n-1}\chi^2_{n-1,1-\alpha/2}\right) = 1-\alpha. \tag{4.6}$$

Väli

$$\left(\frac{(n-1)\hat{\sigma}^2}{\chi_{n-1,1-\alpha/2}}, \frac{(n-1)\hat{\sigma}^2}{\chi_{n-1,\alpha/2}}\right) \tag{4.7}$$

on $100(1-\alpha)\%$ luottamusväli parametrille σ^2, joten väli

$$\left(\frac{6}{U-L}\frac{(n-1)^{\frac{1}{2}}}{\chi_{n-1,1-\alpha/2}}\hat{\sigma}, \frac{6}{U-L}\frac{(n-1)^{\frac{1}{2}}}{\chi_{n-1,\alpha/2}}\hat{\sigma}\right) \tag{4.8}$$

on $100(1-\alpha)\%$ luottamusväli indeksille C_p^{-1}, ja

$$\left(\frac{U-L}{6\hat{\sigma}}\frac{\chi_{n-1,\alpha/2}}{(n-1)^{\frac{1}{2}}}, \frac{U-L}{6\hat{\sigma}}\frac{\chi_{n-1,1-\alpha/2}}{(n-1)^{\frac{1}{2}}}\right) \tag{4.9}$$

$$\equiv \left(\frac{\chi_{n-1,\alpha/2}}{(n-1)^{\frac{1}{2}}}\hat{C}_p, \frac{\chi_{n-1,1-\alpha/2}}{(n-1)^{\frac{1}{2}}}\hat{C}_p\right) \tag{4.10}$$

on $100(1-\alpha)\%$ luottamusväli indeksille C_p.

Tarvittaville prosenttipisteille kaksi yleisemmin käytettyä approksimaatiota on Fisherin

$$\chi_{\nu,\alpha} \cong (\nu - \frac{1}{2})^{\frac{1}{2}} + \frac{z_\alpha}{\sqrt{2}} \tag{4.11}$$

ja Wilson-Hilfertyn

$$\chi_{\nu,\alpha} \cong \nu^{\frac{1}{2}}\left(1 - \frac{2}{9\nu} + z_\alpha\left(\frac{2}{9\nu}\right)^{\frac{1}{2}}\right)^{\frac{1}{2}} \tag{4.12}$$

approksimaatiot missä $\Phi(z_\alpha) = \alpha$.

Näillä approksimaatioilla voidaan muodostaa likimääräisen $100(1-\alpha)\%$ luottamusväli indeksille C_p. Käyttämällä kaavoja (4.7),(4.8) ja (4.10) saadaan Fisherin approksimaation avulla

$$\frac{1}{(n-1)^{\frac{1}{2}}}\hat{C}_p\left(\left(n-\frac{3}{2}\right)^{\frac{1}{2}} - \frac{z_{1-\alpha/2}}{\sqrt{2}}, \left(n-\frac{3}{2}\right)^{\frac{1}{2}} + \frac{z_{1-\alpha/2}}{\sqrt{2}}\right) \tag{4.13}$$

ja Wilson-Hilfertyn approksimaation avulla väli

$$\hat{C}_p\left(1 - \frac{2}{9(n-1)} - z_{1-\alpha/2}\left(\frac{2}{9(n-1)}\right)^{\frac{1}{2}},\right.$$
$$\left. 1 - \frac{2}{9(n-1)} + z_{1-\alpha/2}\left(\frac{2}{9(n-1)}\right)^{\frac{1}{2}}\right)^{\frac{3}{2}}. \tag{4.14}$$

Heavlin (1988) julkaisemattomassa teknisessä raportissaan kävi Kotzin & Johnsonin (1993) mukaan käsiksi muuttujan S^{-1} jakaumaan. Hän kehitti approksimaationsa perustuen normaalijakautuneisuusoletukseen.

$$E[S^{-1}] \cong \left(1 + \frac{3}{4(n-1)}\right)\sigma^{-1} \tag{4.15}$$

$$var(S^{-1}) \cong \frac{1}{2(n-3)\sigma^2}. \tag{4.16}$$

Kaavasta (4.16) saamme approksimaation

$$var(\hat{C}_p) \cong \frac{1}{2}\frac{C_p^2}{(n-3)} \tag{4.17}$$

ja ottamalla mukaan korkeamman asteen termi

$$var(\hat{C}_p) \cong \frac{C_p^2}{2(n-3)}\left(1+\frac{6}{n-1}\right). \tag{4.18}$$

Niinpä väliä

$$\hat{C}_p\left(1-\left(\frac{1}{2(n-3)}\left(1+\frac{6}{n-1}\right)\right)^{\frac{1}{2}} z_{1-\alpha/2}\,,\right.$$
$$\left. 1+\left(\frac{1}{2(n-3)}\left(1+\frac{6}{n-1}\right)\right)^{\frac{1}{2}} z_{1-\alpha/2}\right) \tag{4.19}$$

voidaan käyttää $100(1-\alpha)\%$ luottamusvälinä indeksille C_p.

Vaihtoehtoisia varianssiestimaattoreita saadaan seuraavasti Uusi yleistys saattaa olla hyödyllinen jos oletetaan, että $\hat{\sigma}$ on tilastollisesti muuttujasta $\bar{X}$ riippumaton, ja jakautunut mahdollisesti likimääräisesti kuin $(\chi_f/\sqrt{f})\sigma$. (Yllä $f = n-1$.) Tämä sisältää tilanteen, missä σ on estimoitu usean otoksen vaihteluväleistä (tai otoksen keskiarvon poikkeamana).

Kaavan (4.4) mukaan $\hat{C}_p$ on jakautunut kuten

$$\frac{U-L}{6\sigma}\frac{\sqrt{f}}{\chi_f} = \frac{\sqrt{f}}{\chi_f}C_p, \tag{4.20}$$

joten $G = (\hat{C}_p/C_p)^2$ on jakautunut kuten $f|\chi_f^2|$, ja muuttujan G tiheysfunktio on

$$f_{(G)}(g) = \frac{f^{\frac{1}{2}f}}{2^{\frac{1}{2}f}\Gamma\left(\frac{1}{2}f\right)} g^{-\frac{1}{2}f-1}\exp\left(-\frac{1}{2}fg^{-1}\right), \tag{4.21}$$

missä $g > 0$. Estimaattorin $\hat{C}_p$ r:s momentti on

$$\begin{aligned} E[\hat{C}_p^r] &= f^{\frac{1}{2}r}C_p^r E[\chi_r^{-r}] = f^{\frac{1}{2}r}C_p^r E[(\chi_f^2)^{-\frac{1}{2}r}] \\ &= \left(\frac{f}{2}\right)^{\frac{1}{2}r}\frac{\Gamma(\frac{1}{2}(f-r))}{\Gamma(\frac{1}{2}f)}C_p^r. \end{aligned} \tag{4.22}$$

Erityisesti

$$E[\hat{C}_p] = \left(\frac{f}{2}\right)^{\frac{1}{2}} \frac{\Gamma\left(\frac{f-1}{2}\right)}{\Gamma\left(\frac{f}{2}\right)} C_p = \frac{1}{b_f} C_p \tag{4.23}$$

$$E[\hat{C}_p^2] = \frac{f}{f-2} C_p^2, \tag{4.24}$$

ja

$$var(\hat{C}_p) = \left(\frac{f}{f-2} - b_f^{-2}\right) C_p^2, \tag{4.25}$$

missä

$$b_f = \left(\frac{2}{f}\right)^{\frac{1}{2}} \frac{\Gamma(\frac{1}{2}f)}{\Gamma(\frac{1}{2}(f-1))}. \tag{4.26}$$

Estimaattorin

$$\hat{C}_p' = b_f \hat{C}_p \tag{4.27}$$

odotusarvo on C_p ja niin se on indeksin C_p harhaton estimaattori. Sen varianssi on

$$var(\hat{C}_p') = \left(\frac{f b_f^2}{f-2} - 1\right) C_p^2. \tag{4.28}$$

Kun parametri f on suurempi kuin 14 (Kotz & Johnson 1993), niin tarkka arvio kertoimelle b_f on

$$b_f \cong 1 - \frac{3}{4} f^{-1}. \tag{4.29}$$

Tällä approksimaatiolla

$$var(\hat{C}_p) \cong \frac{f(8f+9)}{(f-2)(4f-3)^2} C_p^2. \tag{4.30}$$

4.2 Indeksin C_{pk} estimointi

Indeksin C_{pk} luonnollinen estimaattori on

$$\hat{C}_{pk} = \frac{d - |\bar{X} - \frac{1}{2}(L+U)|}{3\hat{\sigma}}, \tag{4.31}$$

missä $\hat{\sigma}$ parametrin σ estimaattori.

Jos muuttujan X jakauma on normaali niin estimaattori $\hat{\sigma}$ on jakautunut kuten $\chi_f \sigma / \sqrt{f}$ sekä:

- a) Muuttuja X on normaalijakautunut odotusarvolla ξ ja keskihajonnalla $\sigma/\sqrt{n}$.

- b) $\bar{X}$ ja $\hat{\sigma}$ ovat toisistaan riippumattomia.

Kaavasta (4.31) sekä käyttämällä ylläolevia kohtia (a) ja (b) apuna saadaan

$$
\begin{aligned}
E[\hat{C}^r_{pk}] &= \frac{1}{3^r} E[\hat{\sigma}^{-r}] \sum_{j=0}^{r} (-1)^j \binom{r}{j} d^{r-j} E[|\bar{X} - \frac{1}{2}(L+U)|^j] \\
&= \left(\frac{d\sqrt{f}}{3\sigma}\right)^r E[\chi_f^{-r}] \sum_{j=0}^{r} (-1)^j \binom{r}{j} \\
&\quad \cdot \left(\frac{\sigma}{d\sqrt{n}}\right)^j E\left[\left|\frac{\sqrt{n}(\bar{X} - \frac{1}{2}(L+U))}{\sigma}\right|\right] \qquad (4.32)
\end{aligned}
$$

Termillä $\sqrt{n}|\bar{X} - \frac{1}{2}(L+U)|/\sigma$ on taitettu normaalijakauma (*folded normal distribution*), jonka ovat määritelleet Leone, Nelson & Nottingham (1961). Se on jakautunut kuten

$$|U + \delta|, \qquad (4.33)$$

missä U on standardoitua normaalijakaumaa noudattava muuttuja ja

$$\delta = \frac{\sqrt{n}}{\sigma}(\xi - \frac{1}{2}(L+U)). \qquad (4.34)$$

Kaavasta (4.32) saadaan

$$
\begin{aligned}
E[\hat{C}_{pk}] &= \frac{1}{3}\left(\frac{f}{2}\right)^{\frac{1}{2}} \frac{\Gamma\left(\frac{f-1}{2}\right)}{\Gamma\left(\frac{f}{2}\right)} \Bigg[\frac{d}{\sigma} - \left(\frac{2}{\pi n}\right)^{\frac{1}{2}} \\
&\quad \exp\left\{-\frac{n(\xi - \frac{1}{2}(L+U))^2}{2\sigma^2}\right\} - \frac{|\xi - \frac{1}{2}(L+U)|}{\sigma} \\
&\quad \cdot \left\{1 - 2\Phi\left(\frac{-\sqrt{n}|\xi - \frac{1}{2}(L+U)|}{\sigma}\right)\right\}\Bigg] \qquad (4.35)
\end{aligned}
$$

ja

$$
var(\hat{C}_{pk}) = \frac{f}{9(f-2)}\left[\left(\frac{d}{\sigma}\right)^2 - 2\left(\frac{d}{\sigma}\right)\left[\left(\frac{2}{\pi n}\right)^{\frac{1}{2}}\right.\right.
$$

$$\exp\left\{-\frac{n[\xi-\frac{1}{2}(L+U)]}{2\sigma^2}\right\}+\frac{|\xi-\frac{1}{2}(L+U)|}{\sigma}$$
$$\left\{1-2\Phi\left(\frac{-\sqrt{n}|\xi-\frac{1}{2}(L+U)|}{\sigma}\right)\right\}\Bigg]$$
$$+\frac{(\xi-\frac{1}{2}(L+U))^2}{\sigma^2}+\frac{1}{n}\Bigg]-(E[\hat{C}_{pk}])^2 \tag{4.36}$$

Zhang, Stenback & Wardrop (1990) Kotz & Johnsonin (1993) mukaan. Tässä $f = n - 1$ käytettäessä hajonnan tavanomaista estimaattoria

$$\hat{\sigma} = \left[\frac{1}{n-1}\sum_{j=1}^{n}(X_i - \bar{X})^2\right]^{\frac{1}{2}}.$$

Estimaattori $\hat{C}_{pk}$ on indeksin C_{pk} harhainen estimaattori. Harha kasvaa kahdesta lähteestä (Kotz & Johnson 1993):

1. $$E[\hat{\sigma}^{-1}] = b_f^{-1} \neq \sigma^{-1}.$$

   Tämä harha on positiivinen, koska $b_f < 1$.

2. $$E\left[\frac{\sqrt{n}|\bar{X}-\frac{1}{2}(L+U)|}{\sigma}\right] - \frac{\sqrt{n}|\xi-\frac{1}{2}(L+U)|}{\sigma} \geq 0.$$

   Tämä johtaa negatiiviseen harhaan, koska osoittajalla $\sqrt{n}|\bar{X}-\frac{1}{2}(L+U)|$ on negatiivinen etumerkki.

Chou & Owen (1989) käyttivät kaavaa (3.4) estimaattorin $\hat{C}_{pk}$ johtamiseen. Indeksin C_{pk} estimaattori voidaan kirjoittaa muodossa

$$\hat{C}_{pk} = \min(\hat{C}_{pkl}, \hat{C}_{pku}),$$

missä

$$\hat{C}_{pkl} = \frac{\bar{X}-L}{3\hat{\sigma}} \text{ ja } \hat{C}_{pku} = \frac{U-\bar{X}}{3\hat{\sigma}}$$

ovat indeksien C_{pkl} ja C_{pku} luonnollisia estimaattoreita. Muuttujien $3\sqrt{n}\hat{C}_{pkl}$ ja $3\sqrt{n}\hat{C}_{pku}$ jakaumat ovat epäkeskeisiä t-jakaumia vapausastein f ja epäkeskeisyysparametrein $\sqrt{n}(\xi - L)/\sigma$ ja $\sqrt{n}(U - \xi)/\sigma$.

Kotz & Johnson (1993) ovat tarkastelleet estimaattorin $\hat{C}_{pk}$ jakaumaa. Kun $\hat{C}_{pk} < c$, niin $d - |\bar{X} - m| < 3c\hat{\sigma}$, joten $|\bar{X} - m| + 3c\hat{\sigma} > d$. Tällöin normaalisuusoletuksin

$$\sqrt{n}|\bar{X} - m| \sim \chi_1 \sigma \tag{4.37}$$

ja

$$\hat{\sigma}\sqrt{n-1} \sim \chi_{n-1}\sigma. \tag{4.38}$$

Yhtälöt (4.37) ja (4.38) ovat riippumattomia toisistaan, joten

$$P(\hat{C}_{pk} < c) = P\left(\frac{1}{\sqrt{n}}\chi_1 + \frac{3c}{(n-1)^{\frac{1}{2}}}\chi_{n-1} > \frac{d}{\sigma}\right), \tag{4.39}$$

missä χ_1 ja χ_{n-1} ovat toisistaan riippumattomia.

Bissell (1990) käytti Kotzin & Johnsonin (1993) mukaan muunneltuja estimaattoreita

$$\hat{C}_{pk'} = \begin{cases} \frac{U-\bar{X}}{3\hat{\sigma}}, & \text{kun } \xi \geq \frac{L+U}{2} \\ \frac{\bar{X}-L}{3\hat{\sigma}}, & \text{kun } \xi \leq \frac{L+U}{2} \end{cases} \tag{4.40}$$

Erona indeksiin C_{pk} on se, että prosessin keskiarvon ξ paikalla käytetään otoskeskiarvoa $\bar{X}$. On huomattava, että estimaattoria (4.40) ei voida laskea, jos prosessin keskiarvon sijaintia ei tiedetä. Jos prosessin keskiarvo tunnetaan niin sitä ei tietenkään tarvitse estimoida, ja ~~joilloin~~ voidaan käyttää estimaattoreita

$$\hat{C}^*_{pk} = \min(\hat{C}^*_{pkl}, \hat{C}^*_{pku}), \tag{4.41}$$

$$\hat{C}^*_{pku} = \frac{U-\xi}{3\sigma} \text{ ja } \hat{C}^*_{pkl} = \frac{\xi - L}{3\sigma}. \tag{4.42}$$

Indeksin C_{pk} luottamusvälien rakentaminen on monimutkaisempaa kuin indeksille C_p, koska yhtälössä esiintyy kaksi tuntematonta parametria (ξ, σ). Tässä tapauksessa pyritään ensin luomaan eri luottamusvälit kahdelle parametrille ja sen jälkeen muodostetaan yhdistetty luottamusväli indeksille C_{pk}. Vaikka tämä menetelmä antaisikin tarkkoja indeksin C_{pk} estimaatteja, niin on tärkeä muistaa kaksi asiaa Kotz & Johnson (1993)

1. Kummankin parametrin luottamusvälin luottamuskerroin on $100(1-\alpha)\%$, joten indeksin C_{pk} johdetun luottamusvälin luottamuskerroin on $100(1-\alpha)^2$. Tämä merkitsee sitä, että arvolla $\alpha = 0.05$ indeksin C_{pk} luottamuskerroin on 90.25%. Joten käytännössä kahden luottamusvälin yhdistämisessä saadaan yksi laajempi ja vähän epäluotettavampi väli.

2. Toisaalta vain osa tämän suorakulmaisen alueen pisteistä vaikuttaa indeksin C_{pk} luottamusväliin, ja loput saattavat kompensoida kohdan 1. vaikutusta.

Heavlin (1988) Kotz & Johnsonin (1993) mukaan esitti indeksille C_{pk} $100(1-\alpha)\%$ luottamusvälin

$$\left(\hat{C}_{pk} - z_{1-\alpha/2}\sqrt{\frac{n-1}{9n(n-3)} + \hat{C}_{pk}^2\frac{1}{2(n-3)}\left(1+\frac{6}{n-1}\right)}\right.$$
$$\left.\hat{C}_{pk} + z_{1-\alpha/2}\sqrt{\frac{n-1}{9n(n-3)} + \hat{C}_{pk}^2\frac{1}{2(n-3)}\left(1+\frac{6}{n-1}\right)}\right). \tag{4.43}$$

Kotz & Johnsonin (1993) mukaan Franklin & Wasserman (1992) osoittivat, että

$$\hat{C}_{pk} - z_{1-\alpha}\left(\frac{1}{9n} + \frac{C_{pk}^2}{2(n-1)}\right) \tag{4.44}$$

tuottaa tarpeeksi suurella otoskoolla ($n \geq 30$) huomattavan tarkkoja arvoja $100(1-\alpha)\%$ luottamusvälin alemmalle rajalle. Lisää aproksimaatioita löytyy kirjallisuudesta.

4.3 Indeksin C_{pm} ja muunnelmien estimointi

Indeksin C_{pm} luonnollinen estimaattori saadaan, kun korvataan lauseke $E[(X-T)^2]$ harhattomalla estimaattorillaan

$$\tilde{\sigma}'^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - T)^2. \tag{4.45}$$

Parametri $\tilde{\sigma}'$ ei ole tietenkään muuttujan σ' harhaton estimaattori, vaan sillä on negatiivinen harha, koska $E[\tilde{\sigma}'^2] > \{E[\tilde{\sigma}']\}^2$. Myöskään $\frac{1}{3}d/\tilde{\sigma}'$ ei ole indeksin C_{pm} harhaton estimaattori.

Määritellään

$$\tilde{C}_{pm} = \frac{d}{3\tilde{\sigma}'} = \frac{d}{3}\frac{1}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - T)^2}}. \tag{4.46}$$

Vastaava estimaattori indeksille C_{pm}^* on

$$\tilde{C}_{pm}^* = \frac{(d-|T-m|)}{3}\frac{1}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - T)^2}}. \tag{4.47}$$

Tämä on myös harhainen estimaattori. Molempien estimaattoreiden $\tilde{C}_{pm}$ ja $\tilde{C}^*_{pm}$ harhat tulevat yleensä positiivisiksi, koska estimaattorin $\tilde{\sigma}'$ harha on negatiivinen. Niille tapauksille, kun kullakin $X_1, ..., X_n$ on normaalijakauma odotusarvolla ξ ja keskihajonnalla σ, voidaan johtaa estimaattoreiden $\tilde{C}_{pm}$ ja $\tilde{C}^*_{pm}$ tarkat jakaumat. Kun

$$n\tilde{\sigma}'^2 = \sum_{i=1}^{n}(X_i - T)^2,$$

niin sen jakauma on

$$\sigma^2 \chi_n'^2(\lambda), \tag{4.48}$$

missä epäkeskeisyysparametri $\lambda = n(\xi - T)^2\sigma^{-2}$. Tämä on keskeisen χ^2-jakauman sekoitus vapausastein $(n + 2j)$ ja vastaavin Poisson-painoin

$$P_j(\frac{1}{2}\lambda) = \exp(-\frac{1}{2}\lambda)\frac{(\frac{1}{2}\lambda)^j}{j!}, \tag{4.49}$$

missä $j = 0, 1, ..., n$.

Satunnaismuuttujan $\tilde{C}^r_{pm}$ odotusarvo on

$$\begin{aligned} E[\tilde{C}^r_{pm}] &= \left(\frac{d\sqrt{n}}{3\sigma}\right)^r \exp(-\frac{1}{2}\lambda)\sum_{j=0}^{\infty}\frac{(\frac{1}{2}\lambda)^j}{j!}E[(\chi^2_{n+2j})^{-\frac{1}{2}}] \\ &= \left(\frac{d\sqrt{n}}{3\sigma}\right)^r \exp(-\frac{1}{2}\lambda)\sum_{j=0}^{\infty}\frac{(\frac{1}{2}\lambda)^j}{j!}\frac{\Gamma(\frac{1}{2}(n-r)+j)}{2^{\frac{1}{2}r}\Gamma(\frac{1}{2}n+j)}, \end{aligned} \tag{4.50}$$

kun $r \geq n$, niin $E[\tilde{C}_{pm}]$ on ääretön.

Estimaattorin momentit saadaan korvaamalla kaavan (4.50) muuttuja d termillä $d - |T - m|$. Erityisesti saadaan

$$E[\tilde{C}_{pm}] = \frac{d\sqrt{n}}{3\sigma}\exp(-\frac{1}{2}\lambda)\sum_{j=0}^{\infty}\frac{(\frac{1}{2}\lambda)^j}{j!}\frac{\Gamma(\frac{1}{2}(n-1)+j)}{2^{\frac{1}{2}}\Gamma(\frac{1}{2}n+j)}, \tag{4.51}$$

ja

$$var(\tilde{C}_{pm}) = \left(\frac{d\sqrt{n}}{3\sigma}\right)^2 \exp(-\frac{1}{2}\lambda)\sum_{j=0}^{\infty}\frac{(\frac{1}{2}\lambda)^j}{j!}\frac{1}{n-2+2j} - \{E[\tilde{C}_{pm}]\}^2. \tag{4.52}$$

Kun $\xi = T$, niin $\lambda = 0$ ja

$$C_{pm} = \frac{d}{3\sigma}(= C_p) \tag{4.53}$$

$$E[\tilde{C}_{pm}] = \frac{d\sqrt{n}}{3\sigma}\frac{\Gamma(\frac{1}{2}(n-1))}{2^{\frac{1}{2}}\Gamma(\frac{1}{2}n)} \tag{4.54}$$

$$var(\tilde{C}_pm) = \left(\frac{d\sqrt{n}}{3\sigma}\right)^2 \left[\frac{1}{n-2} - \frac{1}{2}\left\{\frac{\Gamma(\frac{1}{2}(n-1))}{\Gamma(\frac{1}{2}n)}\right\}^2\right] \tag{4.55}$$

Koska

$$C_{pm} = \frac{1}{3}d\{\sigma^2 + (\xi - T)^2\}^{-\frac{1}{2}}$$

niin luottamusväli indeksille C_{pm} saadaan rakentamalla luottamusväli lausekkeelle $\sigma^2 + (\xi - T)^2$.

Jos käytämme normaaliapproksimaatiota jakaumalle

$$\sum_{i=1}^{n}(X_i - T)^2,$$

niin suurille otoksille saamme Kotzin & Johnsonin (1993) mukaan $100(1-\alpha)\%$ luottamusvälin

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - T)^2 \pm z_{1-\alpha/2}\sigma\frac{1}{2n^{-1}(\sigma^2 + 2(\xi - T)^2)},$$

missä $\Phi(z_{1-\alpha}) = 1 - \alpha/2$

4.4 Indeksin C_{pn} estimointi

Choin & Owenin (1990) mukaan indeksin luonnollinen estimaattori on

$$\hat{C}_{pn} = \min\left(\frac{\bar{X} - L}{3\hat{\tau}}, \frac{U - \bar{X}}{3\hat{\tau}}\right). \tag{4.56}$$

Edelleen Choin & Owenin (1990) mukaan positiivisella vakiolla w

$$P(\hat{C}_{pn} \geq w) = P\left(\frac{Z + \sqrt{n}d_1}{3\sqrt{\frac{n}{n-1}Y + (Z + \sqrt{n}\zeta)^2}} \geq w, \frac{-Z + \sqrt{n}d_2}{3\sqrt{\frac{n}{n-1}Y + (Z + \sqrt{n}\zeta)^2}} \geq w\right), \tag{4.57}$$

missä

$$d_1 = \frac{\xi - L}{\sigma},\ d_2 = \frac{U - \xi}{\sigma}. \tag{4.58}$$

ζ on määritelty aikaisemmin (3.18), tässä Z ja Y ovat toisistaan riippumattomia satunnaismuuttujia, jotka liittyvät indeksiin C_{pn}, Z noudattaa standardoitua normaalijakaumaa ja Y noudattaa χ^2-jakaumaa vapausastein $n-1$.

Jos $Z \leq -\sqrt{n}d_1$ tai $Z \geq \sqrt{n}d_1$ niin kaavan (4.57) todennäköisyys on 0, koska $C_{pn} \geq 0$; muulloin se on

$$\begin{aligned} P(\hat{C}_{pn} \geq w) &= P\left(\frac{(Z+\sqrt{n}d_1)^2}{\frac{n}{n-1}Y + (Z+\sqrt{n}\zeta)^2} \geq 9w^2, \right. \\ &\qquad \left. \frac{(Z-\sqrt{n}d_2)^2}{\frac{n}{n-1}Y + (Z+\sqrt{n}\zeta)^2} \geq 9w^2\right). \end{aligned}$$

Tämä todennäköisyys voidaan esittää myös muodossa

$$\begin{aligned} P(\hat{C}_{pn} \geq w) &= P\left(\frac{9nw^2}{n-1}Y \leq (1-9w^2)Z^2 + 2\sqrt{n}(d_1 - 9w^2\zeta)Z + nd_1^2 - 9nw^2\zeta^2, \right. \\ &\qquad \left. \frac{9nw^2}{n-1}Y \leq (1-9w^2)Z^2 - 2\sqrt{n}(d_2 - 9w^2\zeta)Z + nd_2^2 - 9nw^2\zeta^2\right). \end{aligned}$$

Luku 5

Suorituskykyindeksit, kun prosessin jakauma ei ole normaali

Luvussa 3 esitetyt suorituskykyindeksit perustuvat oletukseen muuttujan X normaalijakautuneisuudesta. Entä sitten, kun tämä ehto ei enää toteudu? Käytännössä perin harvoin tutkittava muuttuja on jakautunut puhtaasti teoreettisen normaalijakauman mukaan: jakaumassa saattaa esiintyä eriasteista vinoutta ja huipukkuutta. Mitä voimakkaampina nämä ominaisuudet esiintyvät prosessin jakaumassa, sitä epäluotettavampia ovat normaalijakaumaoletuksen avulla lasketut suorituskykyindeksien arvot. Niinpä epänormaaleihin tapauksiin tarvitaan uusia luotettavampia indeksejä. Epänormaalisesta tilanteesta voi päästä eroon myös käyttämällä jakauman muunnoksia.

Tarkasteltaessa indeksejä ja epänormaaliutta käsittely jakaantuu kahteen osaan (Kotz & Johnson (1993):

1. Suorituskykyindeksien ja niiden estimaattoreiden ominaisuuksien tutkimiseen, kun muuttujan X jakauma on epänormaali.

2. Epänormaaliuden huomioivien menetelmien kehittäminen ja uusien luotettavampien indeksien muodostaminen.

5.1 Luotettavien suorituskykyindeksien muodostaminen

Tässä kappaleessa esitettävät suorituskykyindeksit eivät ole täydellisesti jakaumasta riippumattomia indeksejä, mutta niillä on tarkoitus vähentää epänormaalisuuden vaikutuksia.

Clementsin menetelmä: Clements (1989) esitti Kotzin & Johnsonin (1993) mukaan menetelmän perustuen prosessin jakauman esittämiseen sopivan Pearsonin jakauman

$$\frac{d(\log(f(x))}{dx} = \frac{-(a+x)}{c_0 + c_1 x + c_2 x^2} \tag{5.1}$$

avulla.

Menetelmässä korvataan alkuperäisen indeksin C_p nimittäjän kertoja '6' parametrilla θ, joka antaa todennäköisyyden

$$P(\xi - \frac{1}{2}\theta\sigma \leq X \leq \xi + \frac{1}{2}\theta\sigma) = 0.0027.$$

Annetuilla vinouden ja huipukkuuden kertoimilla $\sqrt{\beta_1}$ ja β_2 saadaan Pearsonin jakaumaperheestä (5.1) määrättyä jakauman f estimaatti $\tilde{f}$. Tästä määrätään parametrien θ_l ja θ_u arvot siten, että

$$P(X \leq \xi - \theta_l\sigma) = 0.135\% = P(X \geq \xi + \theta_u\sigma),$$

josta saadaan $\theta = \theta_u - \theta_l$. Normaalijakauman tilanteessa parametri $\theta = 6$, joka johtaa tuttuun tilanteeseen $C_p = \frac{1}{3}d/\sigma$.

Clements (1989) käytti Kotzin & Johnsonin (1993) mukaan samaa parametrin arvoa θ indeksin C_{pk} laskemiseksi epänormaalissa tilanteessa, mutta hän korvasi keskiarvon $\bar{X}$ mediaanilla M.

Johnson-Kotz-Pearn menetelmä: Clementsin menetelmässä tarvitaan tietoa vinouden ja huipukkuuden kertoimista $\sqrt{\beta_1}$ ja β_2, mutta näiden ominaisuuksien tarkka estimoiminen vaatii melko suuria otoskokoja. Johnson, Kotz & Pearn (1992) ehdottivat käytettäväksi Kotzin & Johnsonin (1993) mukaan indeksiä

$$C_{p(\theta)} = \frac{U-L}{\theta\sigma} = \frac{d}{\frac{1}{2}\theta\sigma}, \tag{5.2}$$

missä θ on nyt valittu niin kutsutuksi kyvykkyydeksi, johon prosessin jakauman muodolla ei ole kovin suurta vaikutusta.

Kotz & Johnson (1993) suosittelevat käytettäväksi indeksiä

$$C_{p(5.15)} = \frac{2d}{5.15\sigma} = \frac{d}{2.575\sigma}. \tag{5.3}$$

Indeksin luotettavuuden paranemisesta tosin seuraa, että siinä menetään 0.27% taso viallisten NC tuotteiden osuudesta.

Kotz & Johnson (1993) mainitsevat erikseen myös indeksin C_{pk} muunnoksen

$$C_{pk(5.15)} = \frac{d - |\xi - \frac{1}{2}(U + L)|}{2.575\sigma}. \tag{5.4}$$

Momenttiestimaattorit indekseille saadaan (5.3) ja (5.4) korvaamalla σ sopivalla estimaattorilla. Menetelmässä oletetaan, että prosessin jakauma on lähellä gammajakaumaa, jolloin ei ole enää takuita jakauman häntätodennäköisyyksien yhtäsuuruuksista.

Jakaumasta riippumattomat suorituskykyindeksit: (Chan, Chen & Spring 1988) esittivät menetelmän jakaumasta riippumattomien suorituskykyindeksien laskemista varten. Estimaattoria $\hat{\sigma}$ on käytetty indeksin $\hat{C}_p$ nimittäjässä arvioimaan välin 6σ pituutta. Tämä välihän peittää normaalijakautuneen ja täydellisesti keskistyneen ($\xi = \frac{1}{2}(U+L)$) prosessin muuttujan X arvoista 99.73%. He ehdottivat käyttettäväksi jakaumasta riippumattomia toleranssivälejä, jotka on muotoiltu pitämään sisällään vähintään 100β% jakaumasta ennalta annetulla todennäköisyydellä $100(1-\alpha)$% parametrien α ja β arvoilla, joista β on lähellä arvoa yksi ja α lähellä nollaa.

Perustana suorituskykyindekseille luonnolliset parametrien arvot olisivat $\beta = 0.9973$ ja ehkä $\alpha = 0.05$. Valitettavasti tällaisten toleranssivälien muodostaminen vaatii suuria otoskokoja: suosituksena on, että otoskoko olisi yli 1000.

Chan ym. (1988) esittivät tämän ongelman voittamiseksi toleranssivälien rakentamista pienemmällä parametrin β arvoilla, ($\alpha = 0.05$ edelleen), joka vaati pienemmän otoskoon (vähemmän kuin 300). He suosittelevat seuraavaa käytäntöä

- $\beta = 0.9546$ ja käytetään termin $6\hat{\sigma}$ sijasta termiä $\frac{3}{2}\cdot$ (toleranssivälin pituus), tai
- $\beta = 0.6826$ ja käytetään termin $6\hat{\sigma}$ sijasta termiä $3\cdot$ (toleranssivälin pituus).

Bootstrap menetelmät: Bootstrap on parametriton otoksen uudelleenkäyttömenetelmä, jonka avulla pyritään saamaan virheetön otosestimaatti hyödyntäen empiiristä otosjakaumaa. Tehdään otos, jonka koko on n ja arvot ovat $x_1, x_2, \ldots, x_n$, tämän jälkeen valitaan kokoa n oleva satunnaisotos [1] $X_{[1]1}, X_{[1]2}, \ldots, X_{[1]n}$ ja lasketaan $\hat{C}_{[1]pk}$ tästä uudesta otoksesta. Tämä toistetaan g kertaa, jolloin saadaan indeksien arvojen sarja $\hat{C}_{[1]pk}, \hat{C}_{[2]pk}, \ldots, \hat{C}_{[g]pk}$, jota pidetään indeksin $C_{pk(kok)}$ jakauman approksimaationa otoskoolla n, tämä estimaatti on bootstrap-jakauma. Käytäntö on osoittanut, luotettavan bootstap-luottamusvälilaskelmaan tarvitaan ainakin 1000 bootstrap-otosta.

Franklin & Wasserman (1991) jakavat bootstrap-luottamusvälit indeksille C_{pk} kolmeen luokkaan:

1. *Standardiluottamusväli*, jossa 1000 bootstrap-otosta on tehty takaisin sijottamalla ja havaitut arvot ovat $X_1, X_2, \ldots, X_n$. Ensin lasketaan bootstrap-estimaatin keskiarvo

$$\hat{C}^*_{pk}(\cdot) = \sum_{i=1}^{1000} \frac{C_{pk}(i)^*}{1000}, \tag{5.5}$$

ja hajonta

$$S_{C_{pk}} = \sqrt{\sum_{i=1}^{1000} \frac{(C^*_{pk}(i) - C^*_{pk}(\cdot))^2}{1000 - 1}}. \tag{5.6}$$

Bootstrap-estimaatille $100(1-\alpha)\%$ luottamusväli on

$$(\hat{C}^*_{pk} - z_{1-\alpha/2} S^*(C_{pk}), \hat{C}^*_{pk} + z_{1-\alpha/2} S^*(C_{pk})), \tag{5.7}$$

missä $\Phi(z_{1-\alpha/2}) = 1 - \alpha/2$.

2. *Prosenttiluottamusväli*, jossa 1000 $\hat{C}_{pk}$ indeksiä on järjestetty pienimmästä suurimpaan $C_{pk}(1) \leq C_{pk}(2) \leq \ldots \leq C_{pk}(1000)$. Luottamusväli on

$$\left(\hat{C}^*_{pk}(\alpha 1000), \hat{C}^*_{pk}((1-\alpha)1000)\right), \tag{5.8}$$

3. *Harhakorjattu prosenttiluottamusväli* on kehitetty korjaamaan mahdollista bootstrap-jakauman harhaa. Menetelmässä ensimmäinen askel on löytää havaitun indeksin $\hat{C}_{pk(kok)}$ arvon paikka järjestetystä bootstrapjonosta $C_{pk}(1) \leq C_{pk}(2) \leq \ldots \leq C_{pk}(1000)$, esimerkiksi $\hat{C}_{pk(kok)} = 1.63$ ja järjestetyt indeksien $C^*_{pk}(i)$ arvot $\hat{C}^*_{pk}(412) = 1.61$ ja $\hat{C}^*_{pk}(413) = 1.66$,

niin $p_0 = P(\hat{C}^*_{pk} \leq 1.63) = 412/1000 = 0.412$. Kun Φ on standardoidun normaalimuuttujan kertymäfunktio, niin lasketaan

$$Z_0 = \Phi^{-1}(p_0).$$

Esimerkissä $Z_0 = \Phi^{-1}(0.412) = -0.222$, koska $P(Z \leq -0.222) = 0.412$. Tämän jälkeen lasketaan

$$\Phi(2Z_0 - Z_\alpha) \text{ ja } \Phi(2Z_0 + Z_\alpha),$$

joten 90% luottamusrajat ovat

$$p_l = \Phi(2(-0.222) - 1.645) = \Phi(-2.089) = 0.0183$$

$$p_u = \Phi(2(-0.222) + 1.645) = \Phi(1.201) = 0.8849.$$

Joten luottamusvälin muoto on

$$\left(\hat{C}_{pk}(1000p_l), \hat{C}_{pk}(1000p_u)\right).$$

Tätä menetelmää voidaan soveltaa myös muihin suorituskykyindekseihin.

5.2 Joustavat suorituskykyindeksit

Johnson ym.(1992) määrittelivät Kotzin & Johnsonin (1993) mukaan joustavan suorituskykyindeksin ottamalla huomioon mahdolliset vaihteluerot kohdearvon T ylä- ja alapuolella. He määrittelivät yksipuoleisen suorituskykyindeksin seuraavasti

$$CU_{jkp} = \frac{1}{3\sqrt{2}} \frac{U - T}{\sqrt{E_{X>T}[(X-T)^2]}} \tag{5.9}$$

ja

$$CL_{jkp} = \frac{1}{3\sqrt{2}} \frac{T - L}{\sqrt{E_{X<T}[(X-T)^2]}}, \tag{5.10}$$

missä

$$E_{X>T}[(X-T)^2] = E[(X-T)^2 | X > T] P(X > T) \tag{5.11}$$

ja

$$E_{X<T}[(X-T)^2] = E[(X-T)^2 | X < T] P(X < T). \tag{5.12}$$

Kertojan $1/(3\sqrt{2})$ käyttö aikaisemmin käytetyn suhteen $\frac{1}{3}$ sijasta, johtuu siitä, että symmetrinen jakauma varianssilla σ^2 ja odotusarvolla T olisi

$$E_{X>T}[(X-T)^2] = E_{X<T}[(X-T)^2] = \frac{1}{2}\sigma^2. \tag{5.13}$$

Lopuksi Johnson ym.(1992) määrittelivät Kotz & Johnsonin (1993) mukaan

$$\begin{aligned} C_{jkp} &= \min(CU_{jkp}, CL_{jkp}). \\ &= \frac{1}{3\sqrt{2}} \min\left(\frac{U-T}{\sqrt{E_{X>T}[(X-T)^2]}}, \right. \\ &\qquad \left. \frac{T-L}{\sqrt{E_{X<T}[(X-T)^2]}} \right) \end{aligned} \tag{5.14}$$

On huomattava, jos $T = \frac{1}{2}(U+L) = m$, eli $U - T = T - L = d$ niin

$$C_{jkp} = \frac{d}{3\sqrt{2}} \max\left(\frac{1}{\sqrt{E_{X>T}[(X-T)^2]}}, \frac{1}{\sqrt{E_{X<T}[(X-T)^2]}} \right). \tag{5.15}$$

Indeksin C_{jkp} luonnollinen estimaattori on

$$\hat{C}_{jkp} = \frac{1}{3\sqrt{2}} \min\left(\frac{U-T}{\sqrt{S_+/n}}, \frac{T-L}{\sqrt{S_-/n}} \right), \tag{5.16}$$

missä

$$S_+ = \sum_{X_i>T} (X_i - T)^2 \text{ ja } S_- = \sum_{X_i<T} (X_i - T)^2.$$

Huomattava on, että

$$E[S_+] = nE[(X-T)^2 | X > T]P(X > T),$$

joten S_+/n on yhtälön (5.11) ja S_-/n yhtälön (5.12) harhaton estimaattori.

Tilanteessa $L < T < U$, indeksin $\hat{C}_{jkp}$ estimaattori voidaan esittää muodossa

$$\hat{C}_{jkp} = \frac{d\sqrt{n}}{(3\sqrt{2})\sigma} \min\left(\frac{U-T}{d\sqrt{\frac{S_+}{\sigma^2}}}, \frac{T-L}{d\sqrt{\frac{S_-}{\sigma^2}}} \right). \tag{5.17}$$

Sen jakauman tutkimiseksi merkitään

$$\hat{D} = \frac{n}{18}\left(\frac{d}{\sigma}\right)^2 \frac{1}{\hat{C}^2_{jkp}} = \max\left(\frac{a_1 S_+}{\sigma^2}, \frac{a_2 S_-}{\sigma^2}\right), \tag{5.18}$$

missä

$$a_1 = \left(\frac{d}{U-T}\right)^2 \text{ ja } a_2 = \left(\frac{d}{T-L}\right)^2.$$

Merkittävää on, että $a_1^{-\frac{1}{2}} + a_2^{-\frac{1}{2}} = 2$, ja

$$E[\hat{C}^r_{jkp}] = \left\{\frac{n}{18}\left(\frac{d}{\sigma}\right)^2\right\}^{\frac{1}{2}r} E[\hat{D}^{-\frac{1}{2}r}].$$

Yleensä estimaattori $\hat{C}_{jkp}$ on siis harhainen. Harha on negatiivinen kun $T = \frac{1}{2}(U+L)$, mutta kasvaa kun $(U-T)/d$ pienenee. Se on aika huomattava, kun $(U-T)/d = 0.4$. Indeksin $\hat{C}_{jkp}$ varianssi suurenee, kun $(U-T)/d$ pienenee, eli kun tavoitearvo T lähenee toleranssialueen ylärajaa. Harha vähenee myös, kun n suurenee. Tämän vaikutus on erityisen huomattava pienillä $(U-T)/d$. Kun $n \geq 25$ ja $(U-T)/d \leq 0.60$, niin odotusarvo ja keskihajonta ovat huomattavan stabiileja.

5.3 Asymptoottiset ominaisuudet

Tässä esitellään joitain asymptoottisia, suuriin otoksiin liittyviä tuloksia, joita voidaan käyttää tutkittaessa suorituskykyindeksien ominaisuuksia epänormaaleissa tapauksissa.

Chan, Xiong & Zhang (1990) ovat tutkineet approksimaatioita indeksien $\hat{C}_p$, $\hat{C}_{pk}$ ja $\hat{C}_{pn}$ jakaumille. Kun

$$\hat{C}_p^{-1} = \left(\frac{3}{d}\right)\hat{\sigma}$$

estimaattorin $\hat{C}_p$ asymptoottista jakaumaa voidaan arvioida tarkastelemalla parametrin $\hat{\sigma}$ asymptoottista jakaumaa.

Chan ym. (1990) ovat todistaneet, että $\sqrt{n}(\hat{C}_p - C_p)$ on asymptoottisesti normaalijakautunut odotusarvolla 0 ja varianssilla

$$\sigma_p^2 = \frac{1}{4}(\beta_2 - 1)C_p^2. \tag{5.19}$$

Jos $\beta_2 = 3$ (eli kyseessä on normaalijakautunut prosessi), niin $\sigma_p^2 = \frac{1}{2}C_p^2$.

Estimaattorille $\hat{C}_{pk}$ Chan ym. (1990) havaitsivat, että jos $\xi \neq m$, niin muuttujan $\sqrt{n}(\hat{C}_{pk} - C_{pk})$ asymptoottinen jakauma on normaali odotusarvolla 0 ja varianssilla

$$\sigma_{pk}^2 = \frac{1}{9} + \frac{\varphi}{2}\sqrt{\beta_1}C_{pk} + \frac{1}{4}(\beta_2 - 1)C_{pk}^2, \tag{5.20}$$

missä $\varphi = 1$ jos $\xi > m$ ja $\varphi = -1$, jos $\xi < m$.

Jos prosessin jakauma on normaali niin $\sqrt{\beta_1} = 0$ ja $\beta_2 = 3$, joten

$$\sigma_{pk}^2 = \frac{1}{9} + \frac{1}{2}C_{pk}^2. \tag{5.21}$$

Erittäin harvinaisessa tapauksessa, kun $\xi = m$, niin muuttujan $\sqrt{n}(\tilde{C}_{pk} - C_{pk})$ rajajakauma on muotoa

$$-\frac{1}{3}\{|U_1| + \frac{1}{2}C_{pk}(\beta_2 - 1)^{\frac{1}{2}}U_2\}, \tag{5.22}$$

missä U_1 ja U_2 ovat standardoituja kaksiulotteisia normaalimuuttujia korrelaatiokertoimella $\sqrt{\beta_1}/(\beta_2 - 1)^{\frac{1}{2}}$. Jos prosessin jakauma on normaali, niin jakauma on silloin muotoa

$$-\frac{1}{3}\left\{|U_1| + \frac{1}{\sqrt{2}}C_{pk}U_2\right\}, \tag{5.23}$$

missä U_1 ja U_2 ovat riippumattomia standardoituja normaalimuuttujia.

Estimaattorille $\hat{C}_{pm}$ on $\sqrt{n}(\hat{C}_{pm} - C_{pm})$ asymptoottisesti normaalijakautunut odotusarvolla 0 ja varianssilla

$$\sigma_{pm}^2 = \frac{\left(\frac{\xi-m}{\sigma}\right)^2 + \sqrt{\beta_1}\left(\frac{\xi-m}{\sigma}\right) + \frac{1}{4}(\beta_2 - 1)}{\left\{1 + \left(\frac{\xi-m}{\sigma}\right)^2\right\}^2}C_{pm}^2. \tag{5.24}$$

Jos prosessi on normaali, niin $\sqrt{\beta_1} = 0$ ja $\beta_2 = 3$, ja

$$\sigma_{pm}^2 = \frac{\left(\frac{\xi-m}{\sigma}\right)^2 + \frac{1}{2}}{\left\{1 + \left(\frac{\xi-m}{\sigma}\right)^2\right\}^2}C_{pm}^2, \tag{5.25}$$

ja jos $\xi = m$, niin

$$\sigma_{pm}^2 = \frac{1}{2}C_{pm}^2. \tag{5.26}$$

Luku 6

Monimuuttujaiset suorituskykyindeksit

Yksimuuttujaisessa tilanteessa prosessin suorituskyky on siis sitä parempi, mitä vähemmän viallisia tuotteita syntyy. Toisin sanoen prosessin suhteellinen suorituskyky on sitä parempi, mitä suurempia indeksien arvot ovat. Useimmiten tilastollisen laadunohjauksen soveltamisessa teollisiin prosesseihin on kyse monimuuttujaisesta laadunohjauksesta. Tuotteella on useita mitattuja ominaisuuksia, joiden jokaisen pitää olla määritelmien mukainen. Monimuuttujaisessa tilanteessa tuotteiden ennaltamäärätty laatu on paljon enemmän riippuvaisempi näiden ominaisuuksien yhteisistä vaikutuksista kuin pelkästään yksittäisistä arvoista. Nykyaikaiset tietokoneet ovat tehneet useiden tunnuslukujen samanaikaisen tallentamisen ja käsittelemisen mahdolliseksi, ja samalla nopeuttaneet analyysien tekoja. Näin monimuuttujaisista valvontamenetelmistä, jotka saivat alkunsa Hotellingin (1947) töistä, on vasta viime aikoina tullut aktiivinen ja tuottoisa tutkimuksen kohde. Kotz & Johnson (1993) ovat sitä mieltä, että suorituskykyindeksien soveltamisessa monimuuttujaisiin mittoihin saa olla huomattavasti varovaisempi kuin yksimuuttujaisessa tapauksessa. Ehkä juuri siksi teollisuus ei ole juurikaan hyödyntänyt monimuuttujaisten suorituskykyindeksien suomaa mahdollisuutta prosessin suorituskyvyn tutkimisessa.

6.1 Monimuuttujaisten suorituskykyindeksin muodostaminen

Kotzin & Johnsonin (1993) mukaan yksimuuttujaisessa tapauksessa tarvitsee huomioida vain mitattujen muuttujien vaihtelu, mutta monimuuttujaisessa tilanteessa täytyy varianssin lisäksi huomioida myös korrelaatiot, sekä tavoitevektorin T hajonta.

Vaihtelun rakenne on riipuvainen toleranssialueen R yhteydestä mitattuihin muuttujiin $X_1, X_2, ..., X_\nu$. Toleranssialue R voi olla periaatteessa minkä muotoinen tahansa, mutta käytännössä sen muoto on suorakulmainen suuntaissärmiö (ν-dimensioinen laatikko), joka määritellään

$$L_i \leq X_i \leq U_i, i = 1, ..., \nu. \tag{6.1}$$

Jos toleranssialue on tätä muotoa, Chan, Xiong & Zhang (1990) käyttivät monimuuttujaisena indeksinä yksittäisten indeksien C_{pm} arvojen tuloa. Tällä sovelluksella on omat puutteensa, vaikka indeksien arvot olisivat riippumattomia toisistaan, voidaan joutuaa ristiriitaisiin tilanteihin: Huomattavan korkea indeksin C_{pm} arvo voi kompensoida erittäin pienen arvon, mikä taas sattaa johtaa kohtuuttoman hyvään arvoon monimuuttujaiselle indeksille. Esimerksi, jos yksi komponentti on vain harvoin toleranssirajojen sisäpuolella, mutta muut komponetit ovat aina toleranssialueella, niin tuloksena on mitä suurimmassa määrin viallinen tuote - tähän ei tietenkään pyritä.

Yksimuuttujaisilla suorituskykyindekseillä on kaksi mahdollista lähestymistapaa: ensimmäinen perustuu virheellisten tuotteiden odotettuun määrään, toinen taas perustuu pääasiallisesti hävikkifunktion käyttöön. Ensimmäisen lähestymistavan perustana on jakauman muoto, kun toisen on vastaavasti hävikkifunktion muoto. Molemmat taas perustuvat normaalijakautuneisuus oletukseen. Näitä enemmän tai vähemmän mielivaltaisia oletuksia on monimuuttujaisessa tilanteessa vielä enemmän, koska muuttujiakin on käytetty enemmän.

Kun yksimuuttujaisessa tilanteessa oletuksena on normaalisuus, niin monimuuttujaisessa on vastaavasti oletuksena multinormaalisuus, missä

$$E[X_i] = \xi_i$$

$$var(X_i) = \sigma_i^2$$

$$corr(X_i, X_{i'}) = \rho'_{ii} \text{ ja } i, i' = 1, ..., \nu$$

missä odotusarvovektori $\xi = (\xi_1, ..., \xi_\nu)$ ja kovarianssimatriisi $\mathbf{V}_0 = (\rho_{ii'}, \sigma_i, \sigma_i')$ ja $\rho_{ii} = 1 \ \forall \ i$. Joten yhdistetyn satunnaisvektori $\mathbf{X} = (X_1, ..., X_\nu)$ tiheysfunktio on

$$f_{\mathbf{X}}(\mathrm{x}) = (2\pi)^{-\nu/2}|\mathbf{V}_0|^{-\frac{1}{2}} \exp\{-\frac{1}{2}(\mathrm{x}-\xi)'\mathbf{V}_0^{-1}(\mathrm{x}-\xi)\}. \tag{6.2}$$

Jos on käytetty hävikkifunktiolähestymistapaa, niin yksimuuttujaisen hävikkifunktion $k(X-T)^2$ yleistys on kvadrattinen muoto

$$L(\mathbf{X}) = (\mathbf{X}-\mathbf{T})'\mathbf{A}^{-1}(\mathbf{X}-\mathbf{T}), \tag{6.3}$$

jossa toleranssialue $L(\mathbf{X}) \leq c^2$.

Kotz & Johnson (1993) painottavat, että prosessin suorituskyvyn mittarina käytettyjen monimuuttujaisten indeksien yksittäiset arvot ovat vaikeampia ymmärtää ja helpompi käyttää väärin kuin yksimuuttujaisten indeksien tapauksessa. Vanhat kunnia-arvoiset mitat ja klassiset tekniikat tilastollisissa monimuuttujaanalyyseissa kuten Hotellingin T^2, Mahalanobisin etäisyys ja pääkomponenttianalyysi tarjoavat mahdollisuuden prosessin vaihtelun yksityiskohtaisempaan tarkasteluun. Jos monimuutujaista ilmiötä pilkotaan yksittäisiin komponentteihin, niin siitä saattaa aiheutua merkittävää informaatiokatoa. Jos prosessi on stabiili, niin monimuuttujaisella indeksillä on paljon helpompi valvoa suorityskykyä kuin käyttämällä monia $\bar{X}/R$-keskiarvoja vaihteluvälikortteja (Kotz & Johnson 1993).

6.2 Indeksin C_p monimuuttujainen vastine

Jos multinormaalisuusoletus (6.2) hyväksytään, Kotzin & Johnsonin (1993) mukaan on luonnollista käyttää suorituskykyindeksin perustana kvadraattista muotoa

$$W = (\mathbf{X}-\xi)\mathbf{V}_0^{-1}(\mathbf{X}-\xi). \tag{6.4}$$

Muuttujan W pitäisi olla χ^2-jakautunut vapausastein ν. Jos toleranssialue R on, kuten aiemmin (6.1) on määritelty, niin

$$W \leq \chi^2_{\nu,0.9973} = c_\nu^2,$$

ja viallisten NC tuotteiden osuus pitäisi noin 0.27% kokonaistuotannosta.

Viallisten NC tuotteiden odotusarvon estimaattori $\hat{p}$ perustuu havaittuun aineistoon. Estimaattoria $\hat{p}$ on käytetty myös sinällään suorituskykyindeksinä, kun prosessin jakaumalla on oletettu, yleensä multinormaalinen ominaisuus (Kotz & Johnson 1993). On erittäin tärkeää ymmärtää, että viallisten tuotteiden riippuvuus prosessin jakauman oikeasta muodosta on merkityksellisempää monimuuttujaisessa kuin yksimuuttujaisessa tapauksessa. Littig, Lam & Pollock (1992) ovat Kotzin & Johnsonin (1993) mukaan esittäneet käytettäväksi lauseketta $\Phi^{-1}(\frac{1}{2}(\hat{p}+1))$ yksimuuttujaisena suorituskykyindeksinä.

Palataan jälleen yksimuuttujaisen indeksin C_p määritelmään

$$C_p = \frac{\text{toleranssialueen pituus}}{6 \cdot (\text{ muuttujan } X \text{ keskihajonta})}. \tag{6.5}$$

Taam, Subbaiah & Liddy (1991) esittivät Kotzin & Johnsonin (1993) mukaan monimuuttujaisen indeksin C_p luonnollisen yleistyksen perustuen hyväksyttyjen tuotteiden klassiseen 99.73% osuuteen

$$\mathbf{C}_p = \frac{\text{toleranssialueen tilavuus}}{\text{alueen tilavuus, jolla on 99.73\% muuttujan } X \text{ arvoista}}. \tag{6.6}$$

Indeksin $\mathbf{C}_p$ jakaja on sen ellipsoidin tilavuus, jolle

$$(\mathbf{x} - \xi)'\mathbf{V}_0^{-1}(\mathbf{x} - \xi) \leq \chi^2_{\nu,0.9973}. \tag{6.7}$$

Tämä on

$$\frac{(\pi\chi^2_{\nu,0.9973})^{\frac{1}{2}\nu}}{\Gamma(\frac{1}{2}\nu + 1)}|\mathbf{V}_0|^{\frac{1}{2}}. \tag{6.8}$$

Saadaan siis

$$\mathbf{C}_p = \frac{\text{toleranssialueen } R \text{ tilavuus}}{a_\nu|\mathbf{V}_0|^{\frac{1}{2}}}, \tag{6.9}$$

missä $a_\nu = (\pi\chi^2_{\nu,0.9973})^{\frac{1}{2}\nu}/\Gamma(\frac{1}{2}\nu + 1)$.

Jos $\xi = \mathbf{T}$ ja $\mathbf{T}$ on toleranssialueen keskellä, niin indeksin $\mathbf{C}_p = 1$ arvo varmistaa, että 99.73% muuttujan $\mathbf{X}$ arvoista asettuu toleranssialueen R sisäpuolelle. Kotzin & Johnsonin (1993) mukaan jos käytetään aluetta $L_i \leq X_i \leq U_i$ $(i = 1, ..., \nu)$, niin suuntaissärmiön tilavuus on

$$\prod_{i=1}^{\nu}(U_i - L_i), \tag{6.10}$$

mutta todennäköisyys $P(U_i \leq X_i \leq L_i)$ kaikilla $i = 1, ..., \nu$ ei ole välttämättä sama kuin 99.73%, vaikkakin

$$\prod_{i=1}^{\nu}(U_i - L_i) = \frac{(\pi\chi^2_{0.9973})^{\frac{1}{2}\nu}|\mathbf{V}_0|}{\Gamma(\frac{1}{2}\nu + 1)}. \tag{6.11}$$

Jos $\mathbf{V}_0 = \theta^2\mathbf{A}$, niin

$$\begin{aligned}
\mathbf{C}_p &= \frac{\text{tilavuus } \mathbf{x}\prime\mathbf{A}^{-1}\mathbf{x} \leq K^2}{\text{tilavuus } \mathbf{x}\prime\mathbf{V}_0^{-1}\mathbf{x} \leq \chi^2_{\nu,0.9973}} \\
&= \frac{\text{tilavuus } \mathbf{x}\prime\mathbf{V}_0^{-1}\mathbf{x} \leq (K/\theta)^2}{\text{tilavuus } \mathbf{x}\prime\mathbf{V}_0^{-1}\mathbf{x} \leq \chi^2_{\nu,0.9973}} \\
&= \left(\frac{K^2}{\theta^2\chi^2_{\nu,0.9973}}\right)^{\nu/2}.
\end{aligned} \tag{6.12}$$

Pearn, Kotz & Johnson (1992) muotoilivat samantapaisen indeksin ottamalla huomioon $\chi^2_{\nu,0.9973}$ yleistetyn pituuden (*generalized length*) vastaavuus viallisten NC tuotteiden todennäköiseen osuuteen 0.27%, ja termiä K^2/h^2 he käyttivät yleistettynä sallittuna pituutena (*generalized allowable length*). He määrittelivät myös indeksin

$$_{\nu}\mathbf{C}_p = \frac{K}{\theta\chi_{\nu,0.9973}} = \mathbf{C}_p^{1/\nu}. \tag{6.13}$$

Indeksien $\mathbf{C}_p$ tai $_{\nu}\mathbf{C}_p$ estimaatit vastaavat parametrin θ estimaattia. Jos $\mathbf{V}_0 = \theta^2\mathbf{A}$, ja arvot $\mathbf{X}_j = (x_{1j}, ...x_{\nu j})$ ovat käytettävissä muuttujan n yksittäisillä $(j = 1, ..., n)$ arvoilla, niin

$$S = \sum_{j=1}^{n}(\mathbf{X}_j - \bar{\mathbf{X}})\prime\mathbf{A}^{-1}(\mathbf{X}_j - \bar{\mathbf{X}}), \tag{6.14}$$

on $\theta^2\chi^2_{(n-1)\nu}$ jakautunut, ja $S/\{(n-1)\nu\}$ on parametrin θ^2 harhaton estimaattori.

Indeksin $_{\nu}\mathbf{C}_p$ luonnollinen estimaattori on

$$_{\nu}\hat{\mathbf{C}}_p = \frac{K}{\chi_{\nu,0.9973}}\left(\frac{(n-1)\nu}{S}\right)^{\frac{1}{2}}, \tag{6.15}$$

joka on jakautunut kuten

$$\frac{K}{\chi_{\nu,0.9973}}\frac{((n-1)\nu)^{\frac{1}{2}}}{\theta\chi_{(n-1)\nu}} =_\nu \mathbf{C}_p\frac{((n-1)\nu)^{\frac{1}{2}}}{\chi_{(n-1)\nu}}. \tag{6.16}$$

Indeksille ${}_\nu\mathbf{C}_p$ $100(1-\alpha)\%$ luottamusväli on siis muotoa

$$\left(\frac{\chi_{(n-1)\nu,\alpha/2}}{((n-1)\nu)^{\frac{1}{2}\nu}}\hat{\mathbf{C}}_p, \frac{\chi_{(n-1)\nu,1-\alpha/2}}{((n-1)\nu)^{\frac{1}{2}\nu}}\hat{\mathbf{C}}_p\right). \tag{6.17}$$

6.3 Indeksin C_{pm} monimuuttujainen vastine

Taam ym. (1991) määrittelivät suorituskykyindeksin C_{pm} monimuuttujaisen vastineen

$$\mathbf{C}_{pm} = \frac{\text{tilavuus } R^*}{\text{ellipsoidin } \mathbf{x}'\mathbf{V}_\mathbf{T}^{-1}\mathbf{x} < c^2 \text{ tilavuus}}, \tag{6.18}$$

missä matriisi

$$\mathbf{V_T} = \mathbf{V}_0 + (\xi - \mathbf{T})(\xi - \mathbf{T})'.$$

$$\begin{aligned} |\mathbf{V_T}| &= |\mathbf{V}_0|\frac{1}{1+(\xi-\mathbf{T})'\mathbf{V}_0^{-1}(\xi-\mathbf{T})} \\ \mathbf{C}^*_{pm} &= \mathbf{C}^*_p\frac{1}{1+(\xi-\mathbf{T})'\mathbf{V}_0^{-1}(\xi-\mathbf{T})}. \end{aligned} \tag{6.19}$$

Kun tällä indeksillä on arvo 1, ja prosessin keskiarvovektori ξ on tavoitevektorin (*target vector*) **T** kanssa yhtäläinen, niin prosessin muuttujien arvoista 99.73% on sallitulla toleranssialueella R. Tämä on yhdenmukainen ominaisuus yksimuuttujaisen indeksin C_{pm} kanssa.

Matriisin $\mathbf{V_T}$ harhaton estimaattori on

$$\hat{\mathbf{V}}_\mathbf{T} = \frac{1}{n}\sum_{j=1}^{n}(\mathbf{X}_j - \mathbf{T})(\mathbf{X}_j - \mathbf{T})' \tag{6.20}$$

ja matriisin $\mathbf{V}_0$

$$\hat{\mathbf{V}}_0 = \frac{1}{n-1}\sum_{j=1}^{n}(\mathbf{X}_j - \bar{\mathbf{X}})(\mathbf{X}_j - \bar{\mathbf{X}})', \tag{6.21}$$

missä

$$\bar{\mathbf{X}} = \frac{1}{n}\sum \mathbf{X}_j.$$

Merkille pantavaa on, että

$$\hat{\mathbf{V}}_{\mathbf{T}} = \frac{n-1}{n}\hat{\mathbf{V}}_0 + (\bar{\mathbf{X}} - \mathbf{T})(\mathbf{X} - \mathbf{T})'. \tag{6.22}$$

Chan, Cheng & Spring (1990) olettivat, että toleranssialue R on muotoa

$$(\mathbf{X} - \mathbf{T})'\mathbf{V}_0^{-1}(\mathbf{X} - \mathbf{T}) \leq K,$$

tällöin he määrittelivät monimuuttujaisen indeksin

$$\begin{aligned} \mathbf{C}_{pm} &= \left(\frac{\nu}{E[(\mathbf{X} - \mathbf{T})'\mathbf{V}_0^{-1}(\mathbf{X} - \mathbf{T})]}\right)^{\frac{1}{2}} \\ &= (1 + \nu^{-1}(\xi - \mathbf{T})\mathbf{V}_0^{-1}(\xi - \mathbf{T}))^{-\frac{1}{2}}. \end{aligned} \tag{6.23}$$

Jos matriisi $\mathbf{V}_0$ on tunnettu, niin indeksin $\mathbf{C}_{pm}^2$ jakajan harhaton estimaattori on

$$\frac{1}{n}\sum_{j=1}^{n}(\mathbf{X}_j - \mathbf{T})'\mathbf{V}_0^{-1}(\mathbf{X} - \mathbf{T}), \tag{6.24}$$

ja indeksin $\mathbf{C}_{pm}$ luonnollinen estimaattori on muotoa

$$\hat{\mathbf{C}}_{pm} = \left(\frac{n\nu}{\sum_{j=1}^{n}(\mathbf{X}_j - \mathbf{T})'\mathbf{V}_0^{-1}(\mathbf{X}_j - \mathbf{T})}\right)^{\frac{1}{2}}. \tag{6.25}$$

On syytä kuitenkin huomata, että indeksillä $\mathbf{C}_{pm}$ on samat huonot puolet kuin vastaavalla yksimuuttujaisella indeksillä C_{pm}. Indeksien arvot voivat vastata dramaattisen erilaisia viallisten NC tuotteiden määriä, jos tavoitearvovektori T ei ole toleranssialueen keskellä. Estimaattori $\hat{\mathbf{C}}_{pm}$ vaatii myös tietoa matriisin $\mathbf{V}_0$ oikeasta muodosta (Chen (1992) Kotzin & Johnsonin (1993) mukaan).

Jos $\xi = \mathbf{T}$, niin $\hat{\mathbf{C}}_{pm}$ on $(n\nu)^{-1}\chi^2_{n\nu}$ jakautunut. Tällöin

$$E[\hat{\mathbf{C}}_{pm}] = \left(\frac{n\nu}{2}\right)^{\frac{1}{2}} \frac{\Gamma(\frac{1}{2}(n\nu - 1))}{\Gamma(\frac{1}{2}n\nu)}, \tag{6.26}$$

ja

$$var(\hat{\mathbf{C}}_{pm}) = \frac{1}{2}n\nu\left(\frac{\Gamma(\frac{1}{2}\Gamma(\frac{1}{2}n\nu - 1) - (\Gamma(\frac{1}{2}(n\nu - 1)))^2}{(\Gamma(\frac{1}{2}n\nu))^2}\right) \tag{6.27}$$

(Chan, Xiong & Zhang 1990).

Prosenttipisteet estimaattorille $\hat{\mathrm{C}}_{pm}$ oletuksella $\xi = \mathbf{T}$ saadaan yhtälöstä

$$
\begin{aligned}
P(\hat{\mathrm{C}}_{pm} \leq G_\alpha) &= P(\hat{\mathrm{C}}_{pm}^{-2} \geq G_\alpha^2) \\
&= P((n\nu)^{-1}\chi_{n\nu}^2 \geq G_{\alpha^{-2}}) \\
&= P(\chi_{n\nu}^2 \geq n\nu G_\alpha^{-2}),
\end{aligned}
$$

josta

$$
G_\alpha = \frac{(n\nu)^{\frac{1}{2}}}{\chi_{n\nu,1-\alpha}}. \tag{6.28}
$$

Kun $n\nu$ on pieni, niin voidaan käyttää approksimaatiota

$$
G_\alpha \cong \frac{(2n\nu)^{\frac{1}{2}}}{z_{1-\alpha} + (2n\nu - 1)^{\frac{1}{2}}}, \tag{6.29}
$$

missä $\Phi(z_{1-\alpha} = 1 - \alpha)$, tai jopa

$$
G_\alpha \cong \left(\frac{z_{1-\alpha}}{(2n\nu)^{\frac{1}{2}}} + 1 - \frac{1}{4n\nu} \right)^{-1}. \tag{6.30}
$$

antaa arvoja, jotka ovat riittävän tarkkoja käytnnöllisiin tarkoituksiin.

6.4 Wierdan monimuuttujainen indeksi

Wierda (1994) että lauseketta $\frac{1}{3}\Phi^{-1}$ käytettäisiin suorituskykyindeksinä, jossa viallisten NC tuotteiden odotettu osuus on θ, se on käyttökelpoinen niin monimuuttujaisissa kuin yksimuuttujaisissa tilanteissa. Sitä voidaan pitää monimuuttujaisen indeksin C_{pk} yleistyksenä. Muuttujan θ estimointi perustuu prosessin jakauman oletettuun muotoon, eli p-dimensioinen satunnaisvektori $X \sim N_p(\mu, \Sigma)$. satunnaismuuttuja. Vektorin X tarkkuus riippuu luonnollisesti tämän oletuksen oikeellisuudesta. Wierda (1994) määritteli suorakulmaisen toleranssialueen

$$
[L, U] = \{x \in R^p | L \leq x \leq U\}, \tag{6.31}
$$

näillä oletuksilla NC tuotteiden osuus on

$$
\theta = \int_L^U n_p(x|\mu, \Sigma)dx. \tag{6.32}
$$

Tästä saadaan monimuuttujainen indeksi

$$
MC_{pk} = \frac{1}{3}\Phi^{-1}(\theta). \tag{6.33}
$$

Luku 7

Suorituskykyindeksit ja Vammaksen aineisto

Tässä tutkielman havainto-osassa on käytetty Vammas Defencetec Oy:n aineistoa, joka sisältää 1772 mittauskertaa 40 millisen ilmatorjuntatykin ammusten hylsyjen pituuksia ja läpimittoja. Tarkalleen aineistossa on mitattu hylsyn 11 eri ominaisuutta; esimerkkiesityksien mittayksikkö on $10^{-3}mm$, jos toisin ei ole mainittu. Indeksien laskut on suoritettu käyttämällä otoskokoa $n = 250$, eli on käytetty vain 250 viimeistä mittauskertaa. Koska tuolloin vasta hylsyjen valmistusprosessi näyttää savuttaneen eri kehitysvaiheiden ansiosta stationaarisen tilan, joka on perusedellytys luotettaville indeksien arvoille. Tämän havainnollistavat kuvat 7.1, 7.2, 7.3 ja 7.4.

Kuvissa 7.1 ja 7.2 on kuvattu muuttujan $X4$ vaihtelua ensimmäisestä mittauskerrasta viimeiseen. Kuvassa 7.1 on kaksi poikkeavaa havaintoa 227 ja 241, jotka aiheuttavat selvät piikit muuttujan $x4$ arvoissa. Kuvassa 7.2 poikkeavat havainnot on poistettu ja prosessin tilaa tarkastellaan hieman tarkemmin. Kuvissa 7.3 ja 7.4 on havainnollistettu muuttujan vaihtelua. Kuten kuvista käy ilmi, muuttujan $X6$ tason vakiintuminen ja vaihtelun pienentyminen on selkeämpää kuin muuttujan $x4$. Noin 500 mittauskerran jälkeen muuttujan $x6$ arvot ovat radikaalisti pienentyneet kohti tavoitearvoa 13800, vielä tämän jälkeen arvoissa on muutamia suuria hyppäyksiä, kunnes viimeissä 250 havainnoissa vaihtelu näyttää pienentyneen merkittävästi.

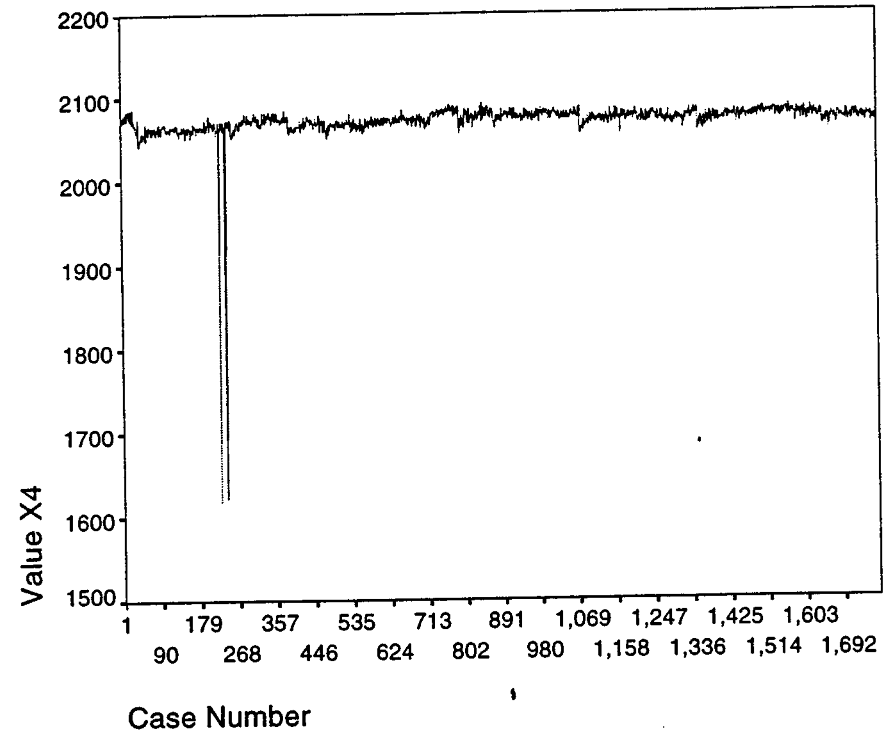
Muuttujien tavoitearvot ovat lähes kaikki toleranssialueen päätepisteitä, poikkeuksena on muuttuja $x4$, jonka tavoitearvo on toleranssialueen keskellä; ks. taulukko 7.1.

Taulukko 7.1: Muuttujien toleranssirajat ja tavoitearvot

	U	L	T
$x1$	57940	57750	57940
$x2$	65000	64700	65000
$x3$	8100	7880	8100
$x4$	2200	1800	2000
$x5$	5120	5000	5000
$x6$	14230	13800	13800
$x7$	25822	25492	25492
$x8$	33890	33500	33500
$x9$	55190	55000	55000
$x10$	365000	364430	365000
$x11$	980	830	980

Taulukko 7.2: Havaintojen tunnuslukuja, kun $n = 250$

Muuttuja	$\bar{X}$	S	Vinous	Huipukkuus
$x1$	57925.00	11.92	-2.88	20.05
$x2$	64923.70	47.02	-13.26	238.11
$x3$	7971.62	10.11	7.12	89.37
$x4$	2075.20	5.19	-0.23	-0.33
$x5$	5062.13	12.05	-2.26	13.79
$x6$	13983.70	3.59	0.54	0.99
$x7$	2724.6	25.14	1.32	1.37
$x8$	33750.4	29.08	1.50	1.06
$x9$	55156.1	9.39	-2.35	9.94
$x10$	364815	18.07	0.70	6.45
$x11$	855.5920	13.94	0.39	-0.24

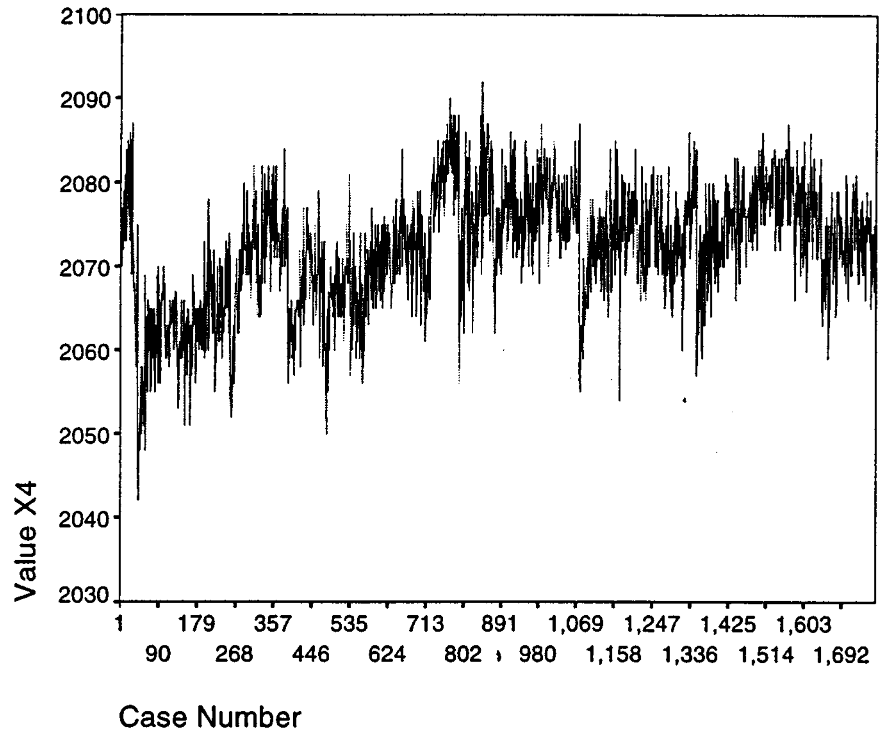


Kuvio 7.1: Muuttujan $x4$ havainnot

7.1 Yksimuuttujaisten suorituskykyindeksien vertailua

Taulukossa 7.3 on laskettu muuttujille $x4$, $x5$, $x6$, $x8$ ja $x11$ eri indeksien arvoja. Indeksin C_p arvojen mukaan kaikkien laskettujen muuttujien prosessien luotettavuuden taso on vähintään erittäin korkea. Korkeimpana indeksin arvona voidaan mainita muuttujan $x6$ huikea arvo $C_p = 19.97$. Indeksi C_p mittaa siis toleranssivälin suhdetta vaihteluun. Indeksi C_{pk} ilmoittaa myös prosessin odotusarvon ξ sijainnin toleranssivälillä, mitä pienempi indeksin C_{pk} arvo on indeksiin C_p nähden, sitä kauempana on prosessin odotusarvo toleranssivälin keskipisteestä.

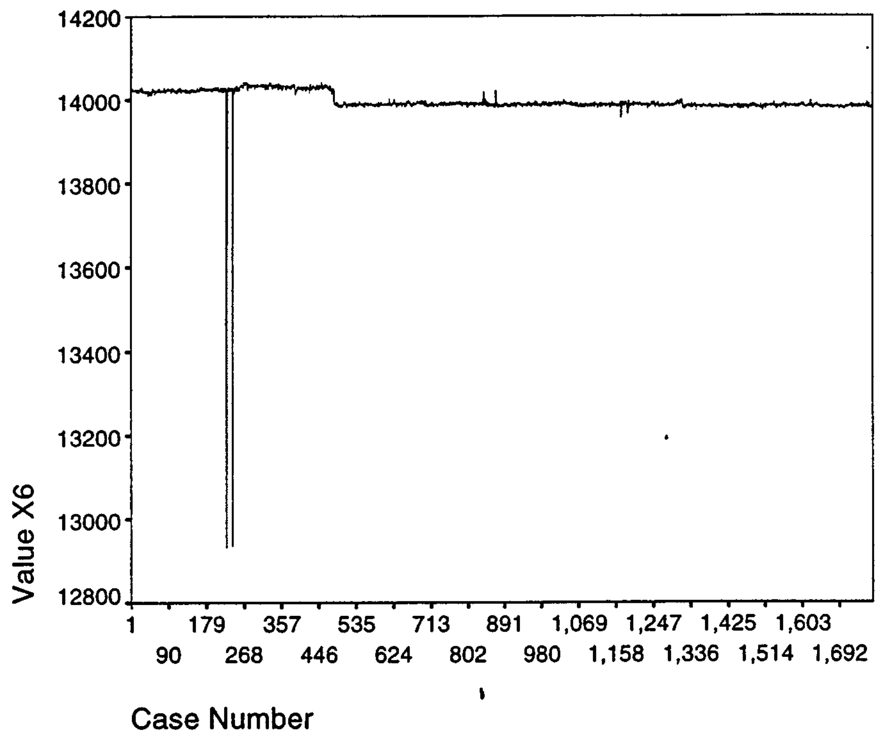
Indeksi C_{pm} mittaa luonnollisen vaihtelun σ lisäksi myös prosessin odotusarvon ξ läheisyyttä tavoitearvoon T, mitä kauempana odotusarvo on tavoitearvosta, sen pienempi on indeksin C_{pm} arvo. Jos mietitään miksi muuttujien $x4$ ja $x6$ indeksin C_{pm} arvot ovat pienempiä kuin vastaavat indeksit



Kuvio 7.2: Muuttujan $x4$ havainnot, kun poikkeavat havainnot 227 ja 241 on poistettu

C_p ja C_{pk} (taulukko 7.3), niin syykin on päivän selvä, prosessien odotusarvot ovat suhteellisesti kaukana tavoitearvoista. Esimerkiksi muuttujan $x6$ indeksi $C_{pm} = 0.39$, keskiarvo $\xi = 13983.70$ ja tavoitearvo $T = 13800$. Tässä tapauksessa prosessin suorituskykyä voisi parantaa pienentämällä prosessin odotusarvoa kohti tavoitearvoa $T = 13800$. Samalla on syytä varoa radikaaleja muutoksia, koska T=L ja muuttujien arvot saattavat joutua toleranssivälin ulkopuolelle radikaalin muutoksen vaikutuksesta, tämähän laskee prosessin suorituskykyä.

Prosessissa pyritään saamaan viallisten tuotteiden osuus ja vaihtelu mahdollisimman pieneksi sekä keskiarvo mahdollisimman lähelle tavoitearvoa. Määritelmien mukaan prosessissa syntyy mahdollisimman vähän viallisia tuotteita, kun prosessin keskiarvo on välin (L, U) keskellä. Indeksi C_{pn} mittaa muiden ominaisuuksien lisäksi tätä suuretta. Indeksi C_{pn} sopii erinomaisesti tilanteeseen, jossa tavoitearvo on toleranssivälin keskellä. Esimerkiksi



Kuvio 7.3: Muuttujan $x6$ havainnot

muuttujan $x4$ indeksin arvo $C_{pn} = 0.55$. Kun tavoitearvo on toleranssivälin päätepisteessä, kuten muuttujan $x6$ tilanteessa, ei indeksin C_{pn} arvoa kannata tuijottaa vakavasti. Jos näin kuitenkin tehdään, syntyy ristiriita, jossa prosessin odotusarvo pyritään saamaan kahtaalle, eli toleranssivälin keskelle sekä sen päätepisteeseen, tällöin indeksin C_{pn} arvo ei voi nousta korkealle. Tämä pätee myös muuttujien $(x5, x8, x11)$ indeksin C_{pn} arvoihin (0.32, 0.36 ja 0.20), ne ovat järjestään pienemmät kuin muuttujan $x4$.

Yksimuuttujaisten suorituskykyindeksien vertaileminen tai laittaminen edes jonkinlaiseen paremmuusjärjestykseen ei ole suorituskykyindeksien relevantein asia. Tärkeintä on ymmärtää, että jokainen indeksi mittaa eri suuretta, ja jokainen indeksi tuo omalta osaltaan tietoa prosessista ja sen suhteellisesta suorituskyvystä. Oli kyse sitten indekseistä C_p, C_p^*, C_{pk}, C_{pk}^*, C_{pm}, C_{pm}^*, C_{pm}^+ tai C_{pn}, ne kannattaa aina laskea.

Taulukko 7.3: Muuttujien $x4, x5, x6, x8$ ja $x11$ indeksit

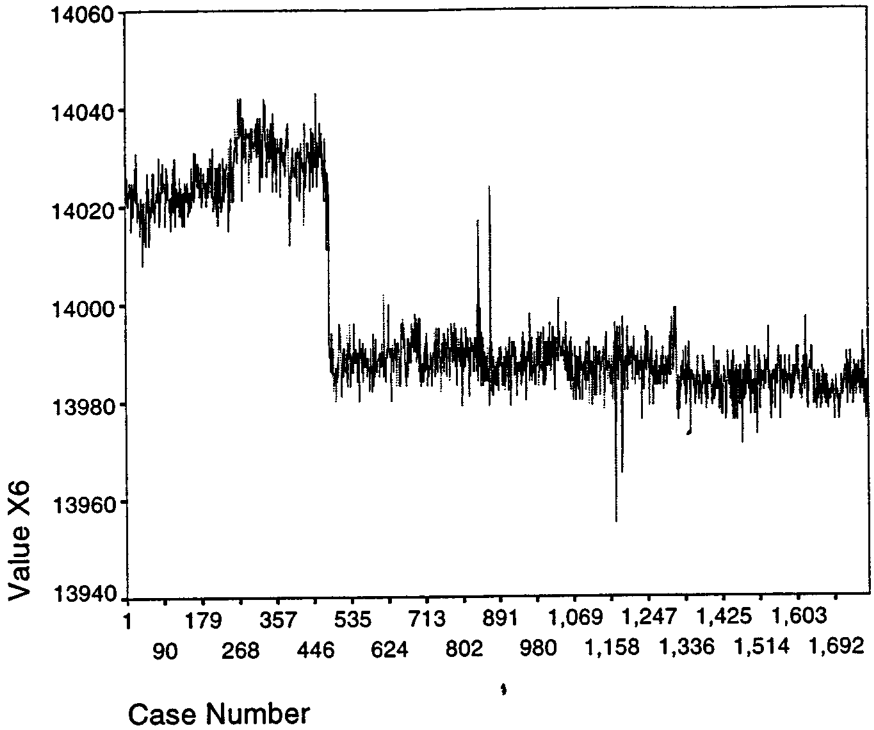
muuttujat	C_p	C_p^*	C_{pk}	C_{pk}^*	C_{pm}	C_{pm}^*	C_{pn}
x4	12.86	12.86	8.02	8.02	0.88	0.88	0.55
x5	1.66	-	1.60	-	0.32	-	0.30
x6	19.97	-	17.06	-	0.39	-	0.33
x8	2.24	-	1.60	-	0.26	-	0.18
x11	1.79	-	0.61	-	0.20	-	0.07

Taulukko 7.4: Suorituskykyindeksien 95% luottamusvälit

luottamusväli	x4	x5	x6	x8	x11
C_p					
Fisherin	11.73 , 13.98	1.51 , 1.81	18.22 , 21.73	2.03 , 2.43	1.64 , 1.95
Wilson-Hilfertyn	12.09 , 13.60	1.56 , 1.76	18.79 , 21.12	2.10 , 2.36	1.69 , 1.90
Heavlinin	11.71 , 14.00	1.51 , 1.81	18.19 , 21.75	2.04 , 2.43	1.63 , 1.95
C_{pk}					
Heavlinin	7.30 , 8.74	1.45 , 1.75	15.54 , 18.59	1.45 , 1.75	0.54 , 0.68
Franklinin alempi	7.32	1.45	15.57	1.45	0.54

Taulukko 7.5: Suorituskykyindeksien 99% luottamusvälit

luottamusväli	x4	x5	x6	x8	x11
C_p					
Fisherin	11.37 , 14.34	1.47 , 1.85	17.67 , 22.28	1.98 , 2.49	1.59 , 2.00
Wilson-Hilfertyn	11.85 , 13.83	1.53 , 1.79	18.42 , 21.49	2.06 , 2.41	1.65 , 1.93
Heavlinin	11.35 , 14.36	1.64 , 1.85	17.63 , 22.31	1.97 , 2.50	1.58 , 2.00
C_{pk}					
Heavlinin	7.08 , 8.96	1.41 , 1.80	15.06 , 19.07	1.40 , 1.80	0.52 , 0.70
Franklinin alempi	7.09	1.41	15.09	1.41	0.52

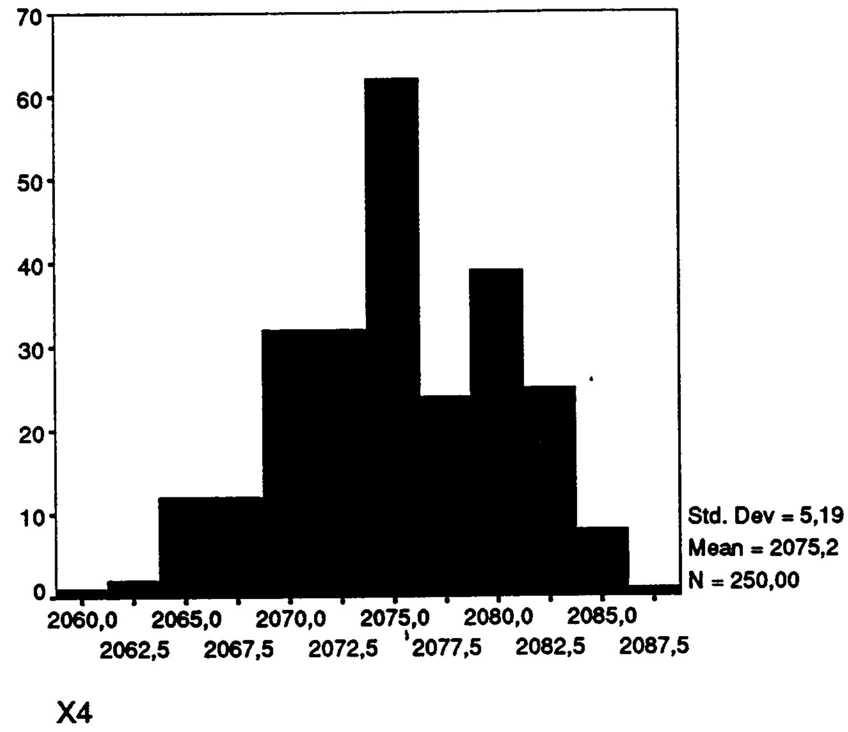


Kuvio 7.4: Muuttujan $x6$ havainnot, kun poikkeavat havainnot 227 ja 241 on poistettu

7.2 Monimuuttujainen hylsy

Ilmatorjuntatykin ammuksen hylsyllä on 11 eri muuttujaa, joiden käyttäytymistä prosessissa pitää valvoa. Hylsyn monimuuttujaisen luonteen johdosta on syytä käyttää myös monimuuttujaista laadunohjausta. Seuraavassa on laskettu kaksi monimuuttujaisen indeksin arvoa muuttujille $x4$, $x8$ ja $x11$. Monimuuttujaisessa tilanteessa hajonnan lisäksi pitää tietää myös muuttujien väliset korrelaatiot, tarvittavat arvot ovat taulukossa 7.7. Monimuuttujaisen indeksin $\mathbf{C}_p$ määritelmä on

$$\mathbf{C}_p = \frac{\prod_{i=1}^{\nu}(U_i - L_i)}{a_\nu|\mathbf{V}_0|^{\frac{1}{2}}},$$



Kuvio 7.5: Muuttujan $x4$ histogrammi

missä $a_\nu = (\pi\chi_{\nu,0.9973})^{\frac{1}{2}\nu}/\Gamma(\frac{1}{2}+1)$. Kun esimerkkitilanteessa

$$\prod_{i=1}^{\nu}(U_3 - L_3) = 400 \cdot 390 \cdot 150$$

ja

$$|\mathbf{V}_0^{\frac{1}{2}}| = 2076.25$$

sekä

$$a_\nu = \frac{(\pi \cdot \chi_{3,0.9973})^{\frac{3}{2}}}{\Gamma(\frac{1}{2}\cdot 3+1)} = \frac{(\pi \cdot 14.16)^{\frac{3}{2}}}{1.33}.$$

Joten

$$\mathbf{C}_p = 50.52$$

ja

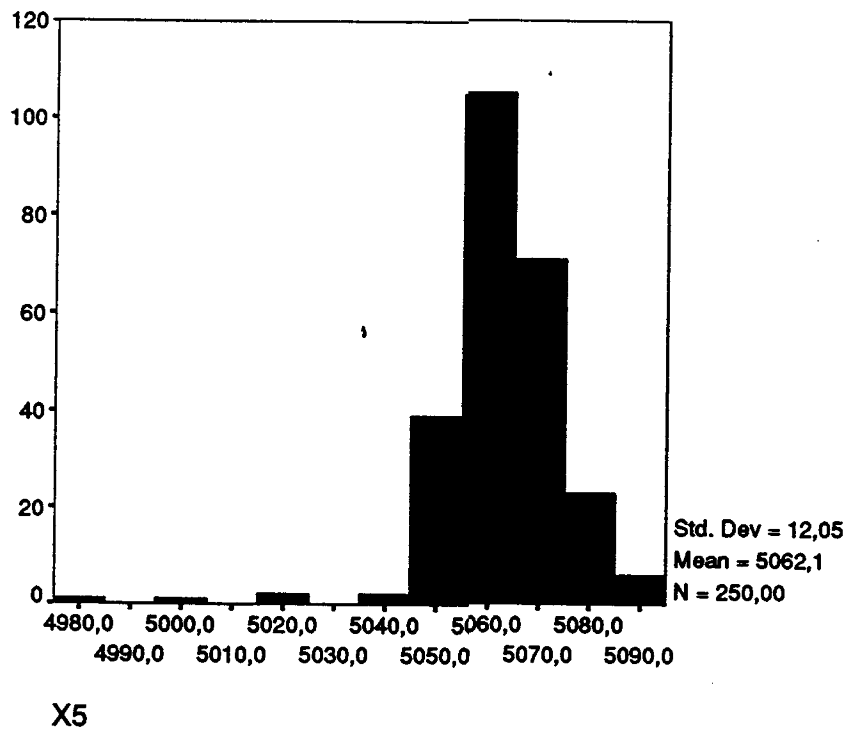
$$_\nu\mathbf{C}_p = \mathbf{C}_p^{\frac{1}{3}} = 3.70.$$

Taulukko 7.6: Suorituskykyindeksien 99.9% luottamusvälit

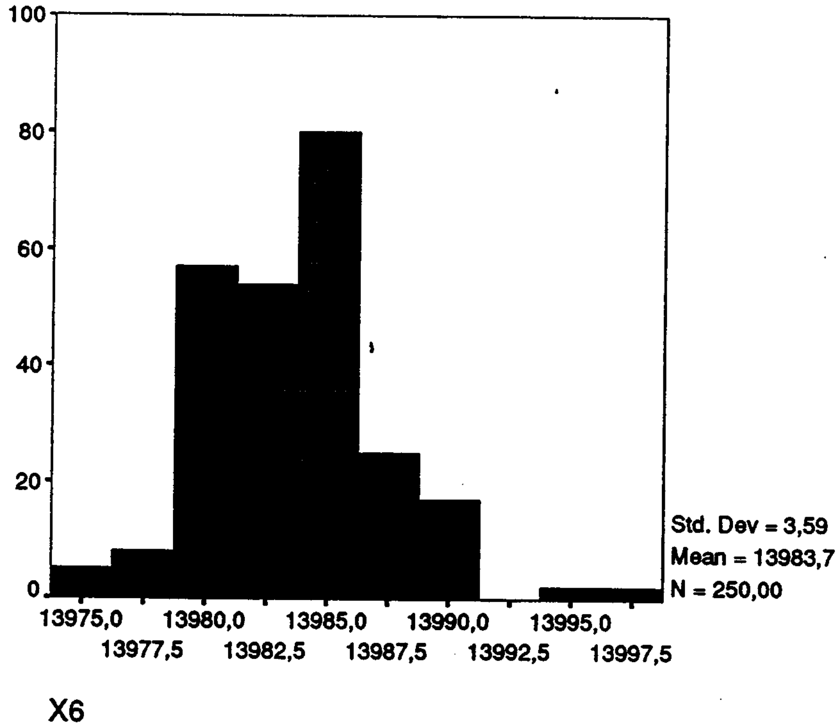
luottamusväli	x4	x5	x6	x8	x11
C_p					
Fisherin	10.96 , 14.75	1.41 , 1.90	17.03 , 21.92	1.91 , 2.56	1.53 , 2.06
Wilson-Hilfertyn	11.58 , 14.11	1.49 , 1.82	17.99 , 21.92	2.01 , 2.45	1.63 , 1.97
Heavlinin	10.93 , 14.78	1.41 , 1.91	16.98 , 22.96	1.90 , 2.57	1.53 , 2.06
C_{pk}					
Heavlinin	6.82 , 9.23	1.35 , 1.85	14.51 , 19.62	1.35 , 1.85	0.50 , 0.73
Franklinin alempi	6.64	1.35	14.55	1.35	0.50

Taulukko 7.7: Muuttujien väliset korrelaatiot ja kovarianssit

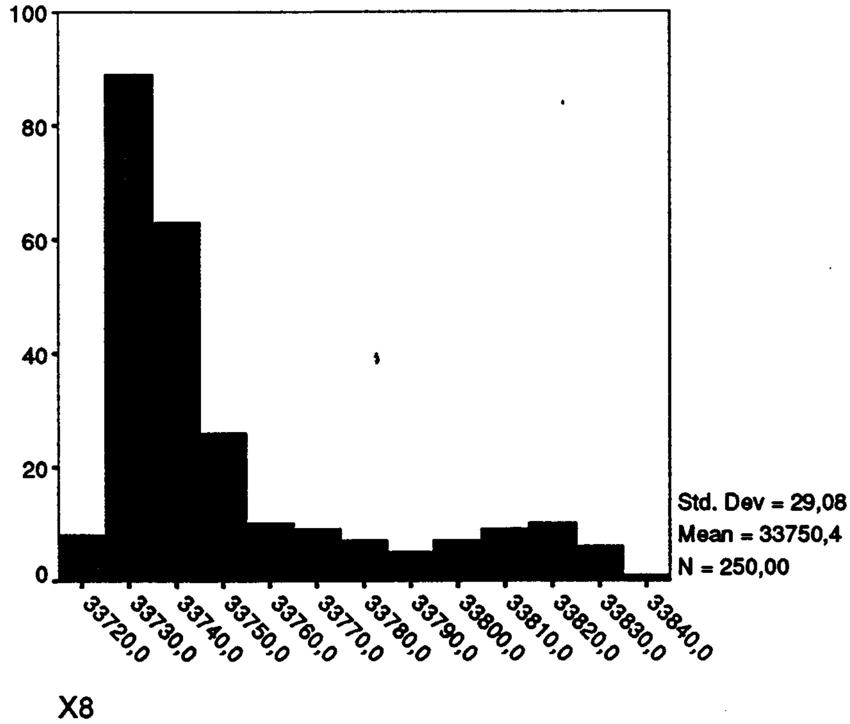
	$x4$	$x8$	$x11$
Pearsonin korrelaatiot			
$x4$	1.00	-0.13	0.02
$x8$	-0.13	1.00	0.08
$x11$	0.02	0.08	1.00
kovarianssit			
$x4$	26.89	-19.79	1.71
$x8$	-19.79	845.61	30.84
$x11$	1.71	30.84	194.22



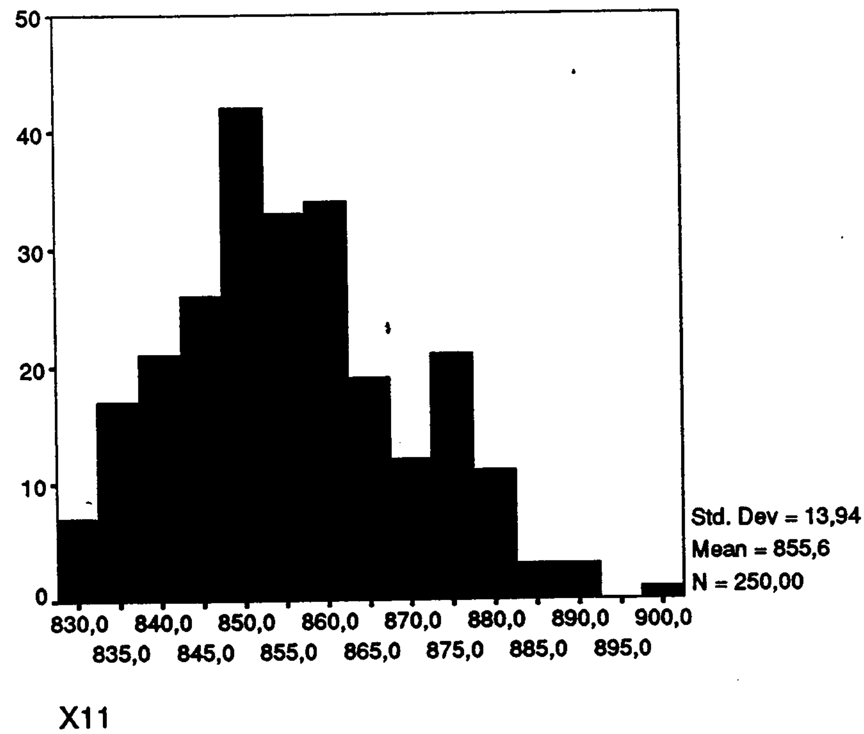
Kuvio 7.6: Muuttujan $x5$ histogrammi



Kuvio 7.7: Muuttujan $x6$ histogrammi



Kuvio 7.8: Muuttujan $x8$ histogrammi



Kuvio 7.9: Muuttujan $x11$ histogrammi

Bibliography

Banks, D. (1993): Is the industrial statistics out of control? *Statistical Science* **8**, 356–377

Bissell, A. F. (1990): How reliable is your capability index? *Appl. Statist.* **39**, 331–340

Boyles, R. A.(1991): The Taguchi capability index. *Journal of Quality Technology* **23**, 17–26

Chan, L. K, Xiong, Z. & Zhang, D. (1990): On the asymptotic distributions of some process capability indices. *Communications in Statistics - Theory and Methods* **19**, 11–18

Chan, L. K., Cheng, S. W. & Spring, F. A. (1988): A new measure of process capability:C_{pm}. *Journal of Quality Technology* **20**, 162–175

Chan, L. K. (1992), Cheng, S. W. & Spring, F. A. (1990): A multivariate measure of process capability. *J. Modeling Simul.* **11**, 1–6. [Abridged version..]

Chen, H. F. (1992): A multivariate process index over a rectancular paralellopiped. *Tecn. Report* Math. Statist. Dept, Bowling Green University, Bowling Green, Ohio

Choi, B-C. & Owen, D. P. (1990): A study of a new process capability index. *Communications in Statistics - Theory and Methods* **19**, 1231–1245

Chou, Y-M. & Owen, D. P. (1989): On the distributions of the estimated process capability indices. *Communications in Statistics - Theory and Methods* **18**, 4549–4560

Clements, J. A (1989): Process capability calculations for nonnormal calculations. *Quality Progress* **22**(2), 49–55

Devor, R. E., Chang, T. H., & Sutherland, J. W. (1992): *Statitical quality design and control: contemporary concepts and methods.* Singapore: Maxwell Macmillan International

Franklin, L. A. & Wasserman, G. S. (1991): Bootstarp confidence interval estimates of C_{pk}: An Introduction *Communication in Statistics - Simulation and Computation* **20**, 231–242

Franklin, L. A. & Wasserman, G. S. (1992): A note on the conservative nature of the tables of lower confidence limits for C_{pk} with a suggested correction. *Communication in Statistics - Simulation and Computation* **21**, 1165-1169

Heavlin, W. D. (1988): Statistical properties of capability indices. *Technical Report No. 32*

Hotelling, H. (1947): In Techniques of Statistical Analysis. McGraw-Hill: New York, 111–184

Ishikawa K. (1990): *Introduction to quality control.* Tokio: 3A Corporation

Johnson, N. L., Kotz, S. & Pern, W. L.(1992): Flexible process capability indices. (Submitted for publication)

Kane, V. E. (1986): Process capability indices. *Journal of Quality Technology* **18**, 41–52

Kotz, S. & Johnson, N. L. (1993): *Process Capability Indices.* Lontoo: Chapman & Hall

Leone, F. C., Nelson, L. S. & Nottingham, R. R. (1961): The folded normal distribution. *Technometrics* **3**, 543–550

Littig, J. L., Lam, C. T. & Pollock, S. M. (1992): Process capability measurements for a bivariate characteristic over an elliptical tolerance zone. *Tech. Rep.* 92–42 Dept. Industrial and Operations Engineering, University of Michigan, Ann Arbor.

Montgomery, D. C. (1985): *Introduction to Statistical Quality Control.* New York: Wiley

Pearn, W. L., Kotz, S. & Johnson, N. L. (1992): Distributional and inferential properties of process capability. *Journal of Quality Technology* **24**, 215–231

Ryan, T. P. (1989): *Statistical methods for quality improvement.* New York: John Wiley & Sons

Taam, W., Subbaiah, P. & , Liddy J. W. (1991): A note on multivariate capability indices. *Working paper* , Dept. Mathematical Sciences, Oakland University, Rochester MI.

Taguchi, G. (1986): *Introduction to quality engineering:Designing quality into products and processes.* New York: UnipubKraus

Wetherill, G. B., & Brown, D. W. (1991): *Statistical process control:theory and practice.* London: Chapman and Hall.

Wierda, S. J. (1994): *Multivariate statistical process control.* Groningen: Wolters-Noordhoff

Zhang, N. F., Stenback, G. A. & Wardrop, D. M. (1990): Interval estimation on process capability index C_{pk}. *Communications in Statistics - Theory and Methods* **19**, 4455–4470