StaM-Bladet

Informationsblad för medlemmar i StaM (Statistisk Metodik), sektion inom SFK, Svenska Förbundet för Kvalitet

Juni 1997 **årgång 7 nummer 13**

Trettonde numret

I detta trettonde StaM-Bladet har vi återigen samlat några olika artiklar. En sådan diskuterar några egenskaper hos de populära måtten "min" och "max" och sammanhängande variationsvidd. Sture Rydin från Stora Cell svarar på frågan "Hur ofta skall man mäta?" och Mats Franzén på Ericsson Cables hjälper oss att hitta på internet.

Vi skriver också några rader om Tjebysjevs olikhet, en av många bra grejer att ha då man studerar statistisk metodik. Den nyfikne bör slå upp denna, och andra olikheter och tumregler, i någon bok om statistik.

Ordförandens ruta

Jag heter Susanna Weinberger och är ny ordförande i SFK-StaM och tillika bergsingenjör. Jag kom i kontakt med SFK-StaM i början på 90-talet när jag letade efter bättre sätt att genomföra försök i produktionen. Via Taguchis metoder kom jag att gå vidare till grunderna i statistisk försöksplanering och därmed Bo Bergmans grupp i Linköping och SFK-StaM:s seminariedagar. I mitt nuvarande arbete som driftsingenjör på Ovako Steel i Hofors har jag ännu inte hunnit ta initiativ till några statistiskt planerade försök, men det kommer säkert att bli tillfälle till det.

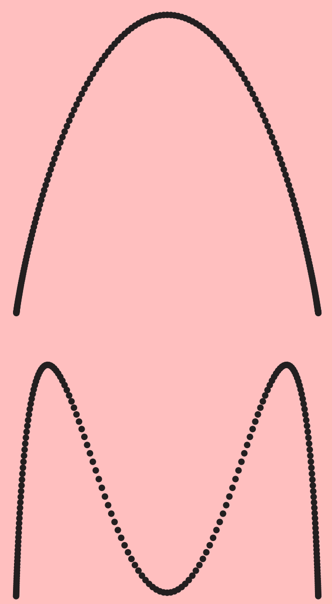
SFK-StaM är en liten idéell organisation men vi blev till vår förvåning ändå kontaktade sent i höstas av Royal Statistical Society i England. I brevet uttrycker de sitt intresse att organisera en seminarievecka tillsammans med oss under 1998 eller 1999 i Sverige! Vi hoppas att vi med andra organisationer, universitet och högskolor i Sverige skall kunna få till en kommitté som kan ta sig an detta.

Jag har tittat igenom alla StaM-Bladet som vi har givit ut. Det finns mycket matnyttigt i dem och många funderingar och tips som även personer utan statistisk examen kan ta till sig. Vi kommer även i fortsättning att balansera vidare på kanten mellan ”populärstatistik” och lite svårare inlägg. Vissa nummer väger över åt endera hållet men målet är som tidigare att intressera er medlemmar för olika typer av statistisk metodik. Tala gärna om för oss om ni har speciellt intresse inom ett område. Vi kan kopiera upp åt er om det finns inlägg i tidigare StaM-bladet som ni själva inte har.

Till sist: Låna ut era StaM-Bladet, låt dem gå på cirkulation till kollegor som kan ha nytta av dem!

Hälsningar

Susanna Weinberger, Ovako Steel AB



Översta figuren är inte en parabel dvs en andragradskurva, utan resultatet då man plottar $f(x)$ mot $F(x)$ för en normalfördelning N[0, 1].

Om man anpassar en andragradskurva med regressionsanalys till kurvan och sedan plottar residualerna, erhålles den undre kurvan.

Förteckning över styrelsen finns på sista sidan

Vaddå variation? Hos min och max?

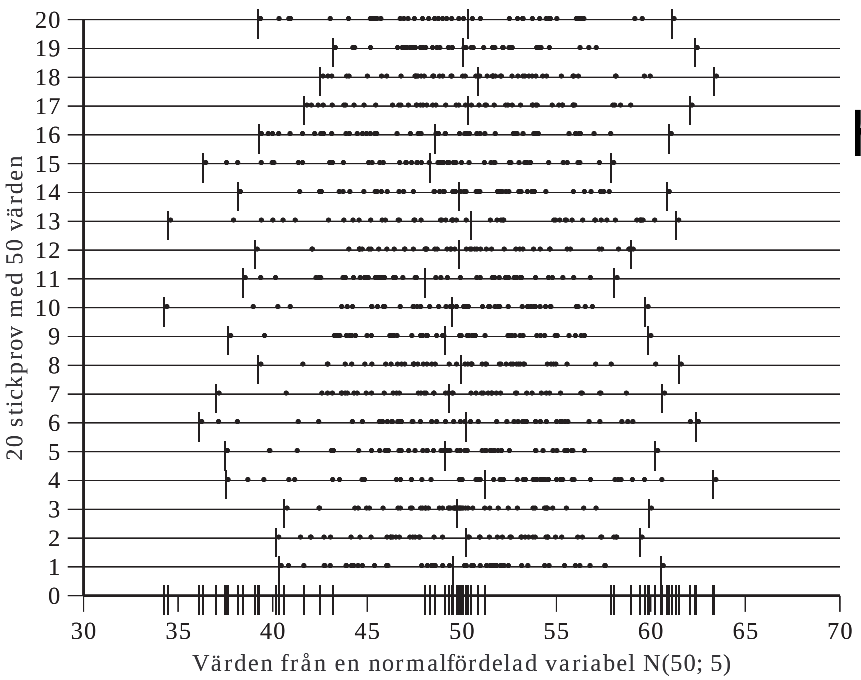
Den vanligaste uppgiften när man sammanfattar ett datamaterial är antagligen medelvärdet, och de flesta människor har inga svårigheter att vare sig beräkna eller förstå dess egenskaper. Dock har man ofta lättare att framhäva dess negativa sidor (jfr exemplet med frysboxen och kokplattan) än dess positiva sidor (väntevärdesriktighet m.m.).

Den som arbetar med numeriska data känner dock snart behovet av något slags spridningsmått, t.ex. variationsvidden som definieras som skillnaden mellan datamängdens max- och min-värden och betecknas ofta med *R* (Range).

Vi skall här inte diskutera variationsviddens olika egenskaper utan snarare titta på max- och min-värdena, deras spridning och deras förhållande till *standardavvikelsen* (*s*). Detta värde, som tidigare diskuterats ett antal gånger i StaM-Bladet, är ett vanligt mått på spridningen i ett datamaterial.

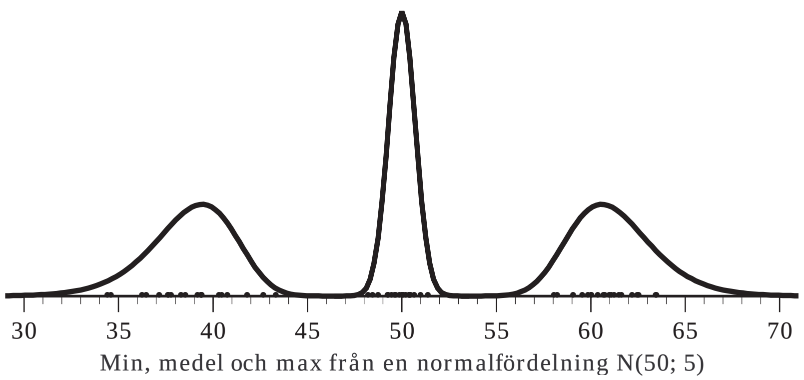
Nackdelen med *R* är att värdet ökar då antalet mätvärden (*n*) ökar. Förändras inte standard-avvikelsen med ökat *n*? Nej, bara skattningens noggrannhet förändras, dvs ju fler mätvärden desto bättre skattar *s* det sanna värdet σ (Se figur 3-4 eller StaM-Bladet nr 12).

Något om spridningen hos min och max



Figur 1. Diagrammet visar 20 simulerade stickprov från en normalfördelning. I varje stickprov har min-, medel- och maxvärdet markerats med ett vertikalt streck. Alla dessa har sedan sammanförts på diagrammets X-axel.

Det framgår tydligt att min- och maxvärdet har en mycket större spridning än medelvärdet. Eftersom min- och maxvärdena är ganska praktiska och lätta att förstå, används de ofta då man diskuterar eller analyserar spridningen hos en eller flera datamängder. Man bör då ha en viss förståelse för figur 1 och 2.



Figur 2. Diagrammet visar fördelningen för min-, medel- samt maxvärdet om man tar 50 värden ur en normalfördelning med väntevärde 50 och standardavvikelse 5.

Fördelningen för min- respektive maxvärdet är skeva (dvs ej symmetriska). Resultaten från figur 1 är också inritade i figur 2.

Antag att vår data kommer från en fördelning som vi betecknar med $f(x)$. Dess fördelningsfunktion kallar vi $F(x)$. Då kan vi beräkna fördelningen för minsta, största, näst minsta etc i stickprov med n värden. Detta brukar redovisas under *Order statistics* i läroböckerna i statistik.

Vi använder bokstaven k för att beteckna det värde vi är intresserade av:

$$f_k(x) = \frac{n!}{(k-1)! \cdot (n-k)!} \cdot f(x) \cdot [F(x)]^{k-1} \cdot [1-F(x)]^{n-k}$$

Om vi är intresserade av minsta värdet (min) dvs vi sätter $k = 1$, får vi följande uttryck:

$$f_1(x) = n \cdot f(x) \cdot (1-F(x))^{n-1}$$

Om vi är intresserade av största värdet (max) dvs vi sätter $k = n$, får vi i stället följande uttryck:

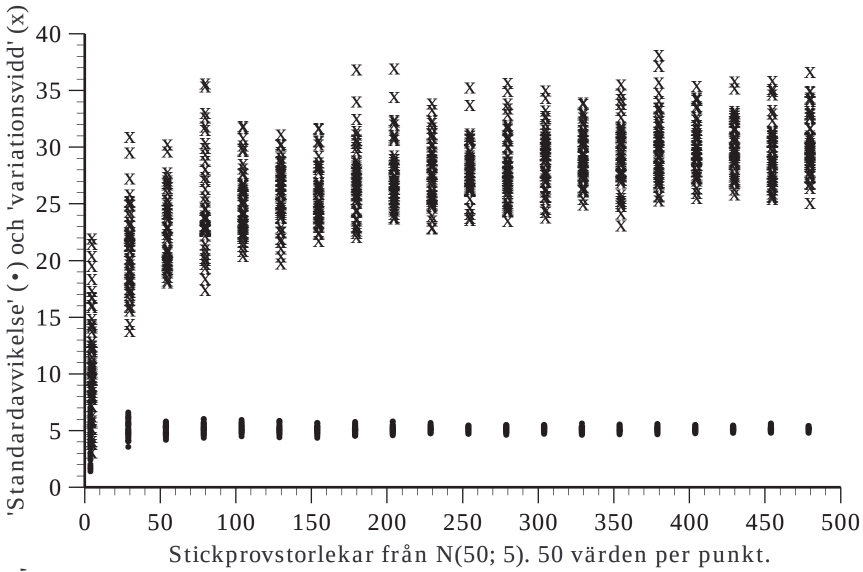
$$f_n(x) = n \cdot f(x) \cdot [F(x)]^{n-1}$$

Observera att det inte är lätt (eller ens alltid möjligt) att ge ett uttryck för $f_k(x)$. I figur 2 har vi använt en dator för att beräkna kurvorna för min- respektive maxvärdet men något enskilt uttryck för kurvorna har vi inte.

Något mer om max och min och standardavvikelsen

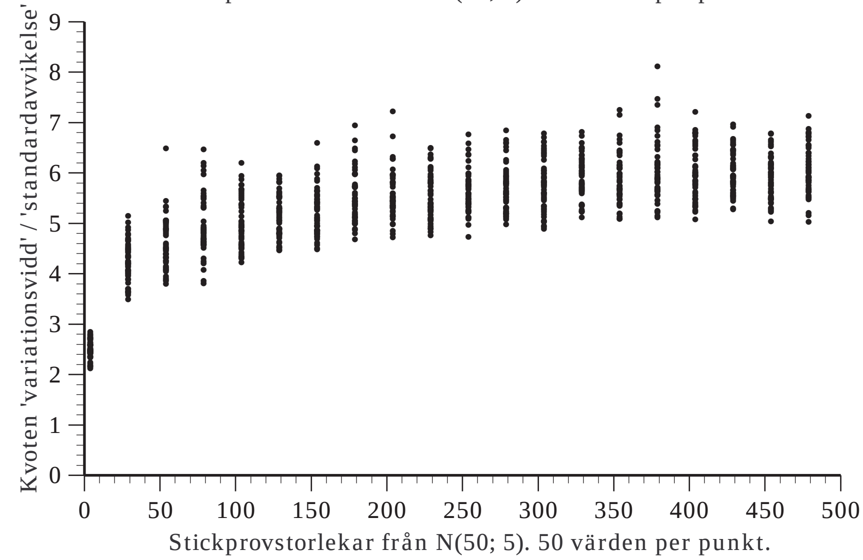
På de följande sidorna finns det ytterligare en betraktelse av förhållandet mellan min-, max och standardavvikelsen.

Något om variationsvidden



Figur 3. Diagrammet visar simulerade värden från en normalfördelning med stickprovstorlekar 5–480 värden. Varje stickprov storlek har simulerats 50 gånger.

Det framgår tydligt av figuren att variationsvidden R ökar då n ökar. Speciellt är detta tydligt då n ökar från 5 till t.ex. 80. Standardavvikelsen visar ingen sådan tendens. Dock ser man att variationen hos den beräknade standardavvikelsen minskar.



Figur 4. Ibland används tumregeln att förhållandet mellan variationsvidden och standardavvikelsen är ungefär 6. Figuren visar att denna tumregel kan användas då n är någorlunda stort t.ex. > 250 värden. Diagrammet visar simulerade värden från en normalfördelning med stickprovstorlekar 5 - 480 värden. Varje stickprov storlek har simulerats 50 gånger. Därefter har kvoten beräknats.

Se också den matematiska behanlingen av förhållandet mellan min, max och medelvärdet.

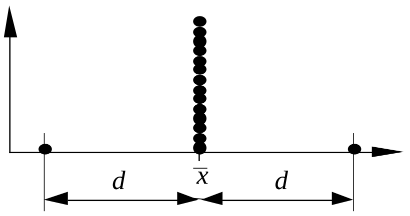
Antag $m = 5.3$ och $s = 0.34$ och $n = 11$. Vad blir då största möjliga variationsvidd? Vad blir minsta möjliga variationsvidd?

Antag att vi har beräknat medelvärdet och standardavikelsen i ett datamaterial. Vad kan vi säga om storleken på min- respektive maxvärdet?

Med endast de två datauppgifterna ovan, och utan de enskilda mätvärdena, kan vi naturligtvis inte säga exakt vad minimum eller maximum är. Vi kan dock ge ett intervall som stänger in dessa båda värden. Vi utgår ifrån uttrycket stickprovsvariansen och omvandlar det något (summan kallas ofta kvadratsumman eller kvadratavvikelsesumman):

$$s^2 = \frac{\Sigma(x_i - \overline{x})^2}{n-1} \quad \Rightarrow \quad \Sigma(x_i - \overline{x})^2 = (n-1)\cdot s^2$$

Största möjliga variationsvidd erhålles om datapunkterna fördelas enligt följande figur:



Formeln ovan kan då skrivas på följande sätt:

$$\underbrace{\left(x_{(1)} - \overline{x}\right)^2}_{d^2} + \underbrace{\left(x_{(2)} - \overline{x}\right)^2}_{0} + \ldots + \underbrace{\left(x_{(n)} - \overline{x}\right)^2}_{d^2} = (n-1)\cdot s^2$$

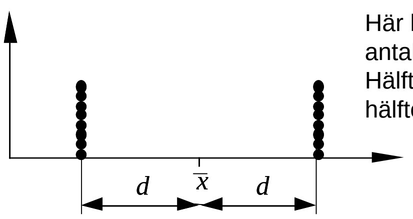
där $x_{(1)}$ och $x_{(n)}$ är minsta respektive största värdet. Vi kan nu beräkna värdet på avståndet d:

$$2d^2 = (n-1)\cdot s^2 \quad \text{vilket ger} \quad d = s\cdot\sqrt{\frac{n-1}{2}}$$

Största möjliga variationsvidd (R_{max}) blir nu:

$$\overline{x} \pm s\cdot\sqrt{\frac{n-1}{2}} \quad \Rightarrow \quad R_{max} = 2\cdot s\cdot\sqrt{\frac{n-1}{2}}$$

Minsta möjliga variationsvidd. Minsta möjliga variationsvidd erhålles då datapunkterna fördelas enligt följande figur:



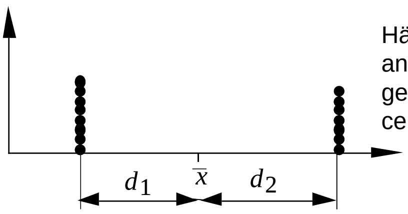
Här har vi ett jämnt antal datapunkter. Hälften vid min- och hälften vid maxvärdet.

Vi kan här skriva den totala kvadratsumman på följande sätt och få ett uttryck av d:

$$n\cdot d^2 = (n-1)\cdot s^2 \quad \Rightarrow \quad d = s\cdot\sqrt{\frac{n-1}{n}}$$

vilket ger minsta möjliga variationsvidd (R_{min}):

$$\overline{x} \pm s\cdot\sqrt{\frac{n-1}{n}} \quad \Rightarrow \quad R_{min} = 2\cdot s\cdot\sqrt{\frac{n-1}{n}}$$



Här har vi ett udda antal datapunkter. Det ger något mer komplicerade uttryck.

Med följande två formler, den första är uttrycket för kvadratsumman och den andra visar en "jämviktsformel" runt medelvärdet (medelvärdet är alltid datamängdens tyngdpunkt), får vi ett uttryck för d_1 och ett annat för d_2 (vi visar inte detaljerna):

$$\begin{cases} \left(\frac{n}{2}+0.5\right)\cdot d_1^2 + \left(\frac{n}{2}-0.5\right)\cdot d_2^2 = (n-1)\cdot s^2 \\ \left(\frac{n}{2}+0.5\right)\cdot d_1 = \left(\frac{n}{2}-0.5\right)\cdot d_2 \end{cases}$$

$$\begin{cases} d_1 = s\cdot\sqrt{\frac{(n-1)^2}{n\cdot(n+1)}} \\ d_2 = s\cdot\sqrt{\frac{n+1}{n}} \end{cases}$$

Dessa två uttryck gäller då vi har ett udda antal datapunkter.

Ett räkneexempel. Antag att $\overline{x} = 5.3$, $s = 0.34$ och $n = 11$ dvs udda. Största möjliga variationsvidd blir:

$$5.3 \pm 0.34\cdot\sqrt{\frac{11-1}{2}} \quad \Rightarrow \quad 6.0603 - 4.5397 = 1.5206$$

Minsta möjliga variationsvidd (R_{min}) då n är udda:

$$\begin{cases} \min = 5.3 - 0.34\cdot\sqrt{\frac{(11-1)^2}{11\cdot(11+1)}} = 5.0041 \\ \max = 5.3 + 0.34\cdot\sqrt{\frac{11+1}{11}} = 5.6551 \end{cases}$$

$$R_{min} = 5.6551 - 5.0041 = 0.6510$$

Kommentarer

Om man matar in 4.5397, 9 st 5.3-värden (dvs $\bar{x}$) samt 6.0603 erhålles $\bar{x} = 5.3$ och $s = 0.34$.

Hela variationen består alltså av de två extrempunkterna enligt figur och största möjliga variationsvidd blir 1.5206.

Om man matar in 6 stycken 5.0041-värden och 5 stycken 5.6551-värden, erhålles $\bar{x} = 5.3$ och $s = 0.34$ och minsta möjliga variationsvidd blir 0.6510.

Självklart kan de 11 datapunkterna fördelas på en oändligt många sätt och ändå ge $\bar{x} = 5.3$ och $s = 0.34$. Beräkningarna ovan ger bara gränsvärden. Observera att vi inte har blandat in någon diskussion av typ av fördelning.

Den som vill lära sig mer om ovanstående resonemang och intilliggande teorier bör studera någon bok i matematisk statistik och då bl.a. slå upp Tjebysjovs olikhet (se också sidan 9).

Ytterligare kommentarer

Om vi låter n bli stort är det ibland lättare att se vad formlerna ger. Vi ersätter $(n-1)$ och $(n+1)$ med bara n i formlerna (felet blir litet). Största möjliga variationsvidd, givet ett visst värde på standardavvikelsen, blir:

$$\bar{x} \pm s \cdot \sqrt{\frac{n}{2}}$$

Minsta variationsvidd blir $d = d_1 = d_2 = s$ och vi får:

$$\bar{x} \pm s$$

Exempel. Låt som tidigare $\bar{x} = 5.3$, $s = 0.34$, $n = 300$.

Vi matar in 150 värden 5.3 – 0.34 = 4.96 och 150 värden 5.3 + 0.34 = 5.64 i datorn.

En beräkning av medelvärde och standardavvikelse ger värdena 5.3 respektive 0.34.

Det är inte ofta som StaM-Bladet presenterar verk från en Nobelpristagare. Nedan finns dock ett

Bidrag till statistiken

Av hundra människor

är de som vet bäst
minst femtiotvå,

de som med tvekan tar ett steg
nästan hela resten,

de som är redo att hjälp till,
bara det inte tar för mycket tid,
så många som fyrtionio,

de som alltid är snälla,
för de kan inte annat,
fyra eller kanske fem,

de som kan beundra en utan avund
– arton stycken,

de som har förts vilse
av sin ungdom som förgått
– sextio, sextiofem,

de som inte är att leka med
– fyrtio plus fyra,

de som lever i ständig ångest
för någon eller någonting
– hela sjuttiosju,

de begåvade för lycka
– inte fler än tjugofem,

de harmlösa en och en,
som förvildas i hop,
säkert över hälften,

de grymma,
då läget tvingar dem därtill,
lika bra att inte veta
ens på ett ungefär,

de visa av skadan
– inte många fler än
de som var kloka förut,

de som inte får ut nånting
av livet utom tingen
– trettio stycken,
fast jag önskar att jag hade fel,

de hopkurade och pinade,
utan lyse ute i mörkret,
räknas till åttiotre
förr eller senare

de rättfärdiga
– ganska många, trettiofem,

men om det draget ska hänga ihop
med en strävan att förstå
– tre,

de som är värda att tyckas synd om
– nittionio av hundra,

de dödliga
–hundra av hundra
En siffra som hittills har stått sig.

Wislava Szymborska
Övers. Anders Bodegård

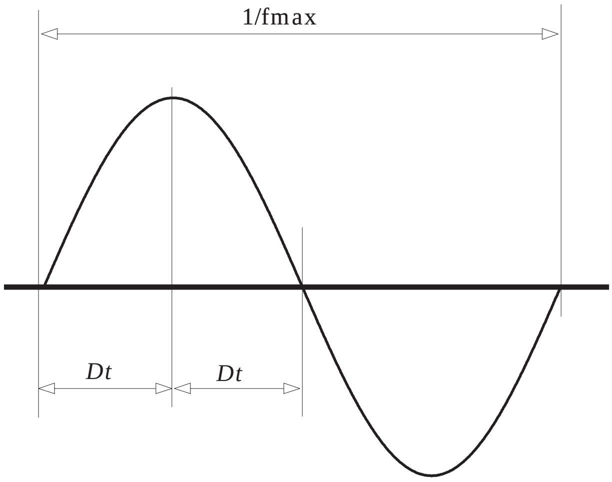
Hur ofta ska man mäta?

I alla tillverkningsprocesser förekommer intermittent provning och analys. Detta är kanske speciellt vanligt inom pappers och massaindustrin, där en mängd storheter mäts i processflödet för att kunna säkerställa processnivåer och i slutändan, kvaliten på den färdiga produkten. Helst skulle man vilja kunna mäta alla intressanta storheter online med hjälp av smarta instrument men så är ännu icke fallet, åtminstone inte inom skogsindustrin.

När man inom processindustrin gör upp scheman över provtagningsfrekvensen för olika storheter, så sker det i de flesta fall på basis av slentrian eller vad man tror sig hinna med att analysera och rapportera. Problemet togs upp inom STORA CORPORATE RESEARCH av två personer (Johan Wieslander och Mats Hiertner) som utvecklade en metod att bestämma det totala mätfelet vid intermittent provning och analys, som beskrivs nedan. Metoden har med framgång använts inom STORA-koncernens massafabriker och pappersbruk i syfte att få klarhet i om rätt provtagningsfrekvens använts. I samband med metodens utnyttjande är följande två frågor av central betydelse:

1. Kan man med det valda provtagningsintervallet följa processens variationer tillräckligt väl?
2. Hur stor är osäkerheten i mätningen i förhållande till processens variation till nästa mättidpunkt?

Eftersom mätningarna syftar till att ge en god beskrivning av en störning, så rekommenderas, att tiden mellan mätningarna väljs enligt det idealiserade sätt, som visas i figur 1.



Figur 1. Regel för val av mätfrekvens. *Dt* bör vara hälften av tiden mellan de snabbast kända förändringarna hos storheten (Nyquists samplingsteorem).

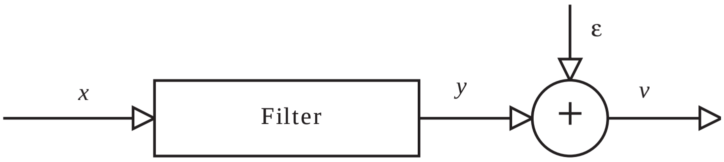
Metod

Metoden bygger på ett antagande om, att man kan dela upp de totala variationerna i mätdata i långsamma variationer, eller trender och brus. Med brus avses här:

1. de variationer, som är så snabba, att de med det valda provtagningsintervallet ej syns i två på varandra följande mätningar
2. de variationer, som beror på osäkerheten i själva mätningen.

I de långsamma variationerna finns i regel en korrelation mellan nuvärde och föregående mätning. Genom att beräkna mätseriens autokovariansfunktion, kan man också erhålla den önskade uppdelningen i processvarians och mätosäkerhet. En förutsättning är att data antas vara genererade av en autoregressiv process av första ordningen enligt följande statistiska modell som också visas schematiskt i figur 2:

$$y_t = kx_t + (1-k)y_{t-1}$$



Figur 2. Data antas genereras av en autoregressiv process av första ordningen, dvs AR(1).

x = vitt brus, som tillsammans med ett filter används för att beskriva processen
y = processens värde
k = filterkonstant
ε = brus
v = uppmätt värde

Under förutsättning, att x_t är en oberoende, stokastisk variabel med medelvärdet μ_x och standardavikelsen σ_x, så blir medelvärdet för y_t (om t är tillräckligt stort):

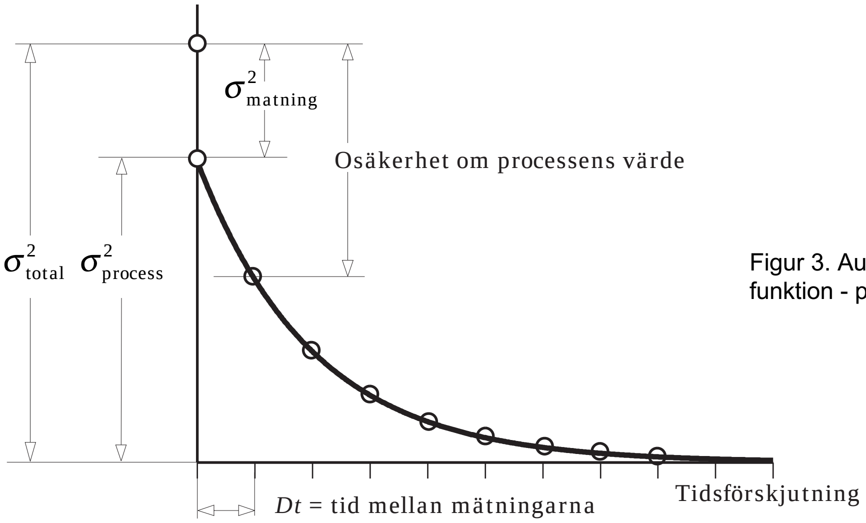
$\mu_y = \mu_x$ och variansen för y_t blir $\sigma_y^2 = \dfrac{k}{2-k}\sigma_x^2$

Autokovariansen, A, för y är kovariansen mellan de två tidserierna y_t och y_{t-1}, där T är tidsförskjutningen, dvs

$$A(T) = \frac{k(1-k)^T}{2-k}\sigma_x^2$$ Eftersom $(1 - k) < 1$ får vi då $T > 1$

$$A(T) = \frac{k}{2-k}\sigma_x^2 \cdot e^{-T/\tau}$$ τ = processens tidskonstant.

Processens tidskonstant är den tid, som processen använder för att återhämta sig från en störning. Det är alltså den snabbaste förändringen i processen, som är detekterbar med det mätintervall, som har valts. Autokovariansen för v är summan av autokovariansen för y och ε. Detta baseras på antagande att både x och ε är oberoende. Autokovariansfunktionens utseende framgår av figur 3.

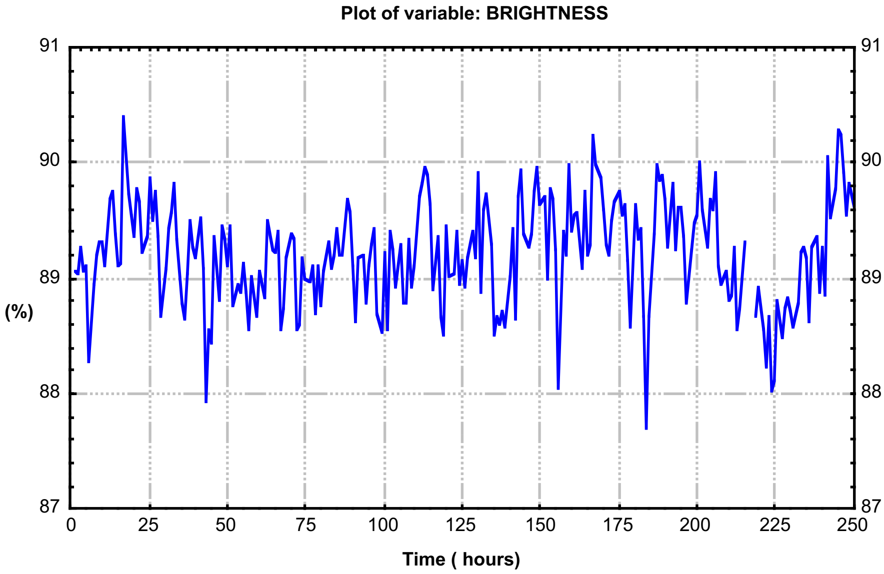


Figur 3. Autokovariansfunktion - principbild.

En tumregel för en bra mätning av en *ostyrd variabel* är att förhållandet mellan mätosäkerhet och totalvarians bör vara < 0.33. Om variabeln däremot är en *styrd* dito, ska mätosäkerheten, eller bruset, dominera.

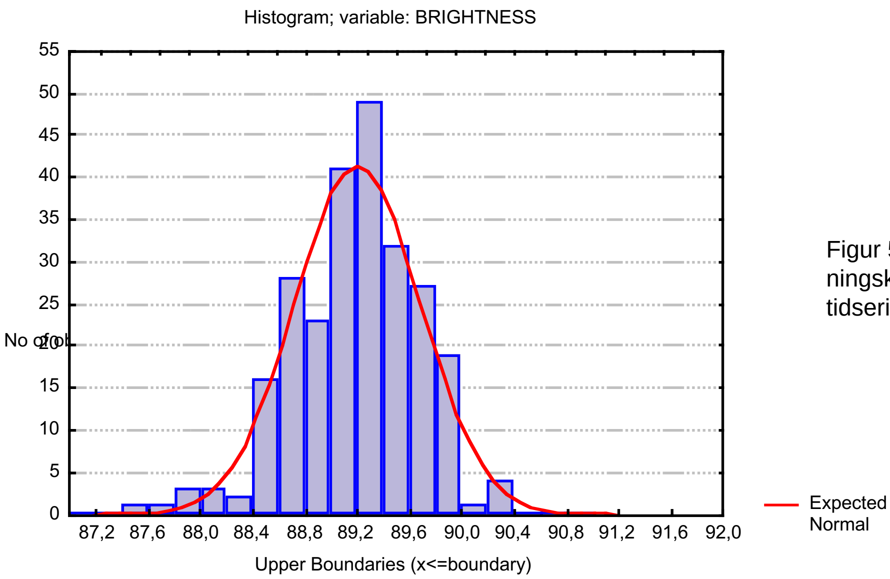
Några praktiska exempel

Nedanstående exempel är hämtade från en svensk massafabrik. Den variabel, som valts att studeras är den färdiga massans ljushet (Eng. ”Brightness”), som mäts genom intermittent provtagning och analys var 45:e minut. Figur 4 visar en tidserie på denna storhet.



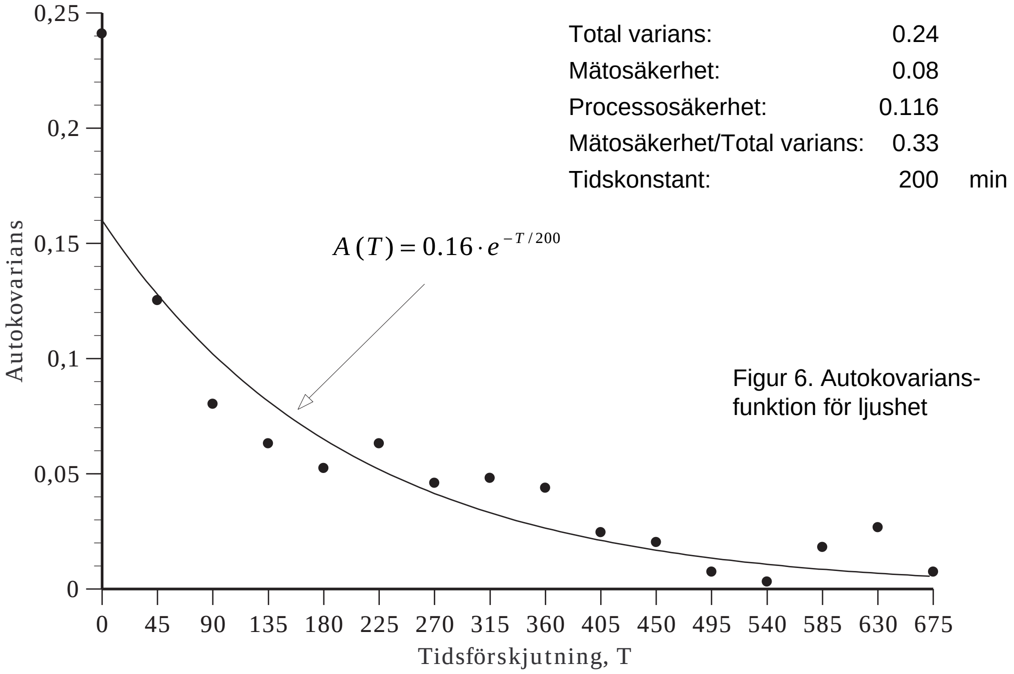
Figur 4. Tidskurva för ljushet på pappers-massa. En mätning var 45:e minut.

Motsvarande fördelningskurva för de ingående observationerna ges i figur 5.



Figur 5. Fördel-ningskurva för tidserien i figur 4.

Mätseriens autokovariansfunktion med de önskade, beräknade osäkerhetsmåtten samt med processens tidskonstant visas i figur 6.



Figur 6. Autokovarians-funktion för ljushet

De observationer på ljushet, som redovisas i det här exemplet, mäts i processens slutkedja och är i den delen av tillverkningsprocessen att anse som en ostyrd variabel. Kvoten mellan mätosäkerhet och totalvarians ligger på 0.33, vilket är på gränsen till att vara acceptabelt. Den snabbast detekterbara förändring, som kan mätas med det mätintervall som används, är 200 minuter. Också detta är på gränsen till vad som kan anses som acceptabelt.

Den angivna metoden utnyttjas inom STORA-koncernen för att då och då trimma in nya provtagnings och analysfrekvenser, vilket i regel blir fallet då nya analyser införs i processen eller då det har gjorts en större om- eller tillbyggnad i produktionsanläggningen.

Sture Rydin, Stora Cell AB

Tjebysjevs olikhet

På sidorna 4 och 5 visar vi att givet medelvärde, standardavvikelse och antal mätvärden, kan vi beräkna största möjliga variationsvidd. Resonemanget är allmängiltigt och genomfördes utan referens till någon speciell statistisk fördelning.

När man lär sig att hantera normalfördelningen lär man sig ofta att inom ± 1 standardavvikelse finns cirka 68 % av alla värden, inom ± 2 standardavvikelse finns cirka 95 % osv.

Med hjälp av Tjebysjevs olikhet (eng. Chebyshev's inequality) kan man göra liknande uttalanden men då gällande för alla fördelningar. Så här kan man skriva olikheten (obs att stavningen av namnet varierar mellan olika böcker):

Låt X vara en slumpvariabel med väntevärde μ och standardavvikelse σ. Då gäller följande:

$$P\left(|X-\mu| \geq t \cdot \sigma\right) \leq \frac{1}{t^2}$$

"Sannolikheten att vi får ett värde på slumpvariabeln som avviker mer än t standardavvikelser från medelvärdet, är mindre eller lika med $1/t^2$".

Sannolikheten att få ett värde som ligger mer än 2 standardavvikelser från medelvärdet är mindre än eller lika med 25 % (för t = 3 gäller 9 %).

Olikheten används ofta för att bevisa andra viktiga teoretiska aspekter av statistisk teori men är också en 'bra-att-ha'-grej i det praktiska arbetet.

Pröva teorin på någon ovanlig fördelning m.hj.a. simulering!

Statistik på Internet

Då det idag talas mycket om Internet, kan det vara på sin plats att visa exempel på de resurser som erbjuds inom området statistik. Alla som redan har erfarenhet av att surfa på Internet vet att det finns bra sökverktyg för att finna den information man söker. Det gäller ofta att finna en ingång till ett ämne och efter att den är funnen hittar man ofta en mängd information. Nedan visar jag några ingångar inom ämnet statistik. Adresserna är skrivna med kursiv stil.

Adresser till bibliotek

Det finns översiktliga samlingar med bl a adresser till andra resurser inom ämnet. Från nedanstående adresser kan man hitta mycket av intresse.
Juha Puranen, Dep of Statistics, University of Helsinki:
www.helsinki.fi/~jpuranen/links.html

WWW virtual library of statistics:
stat.ufl.edu/vlib/statistics.html

Statlib:
lib.stat.cmu.edu/

För den vetgirige finns en elektronisk lärobok i grundläggande statistik:

www.psychstat.smsu.edu/sbk00.htm

Adresser till programvara

De flesta vanliga statistikprogram har hemsidor på Internet t.ex. Minitab, Statistica, SPSS m.fl. Demonstrationsprogram brukar ibland finnas att kopiera för utvärdering. Det finns annan programvara att hämta hem via Internet både för Mac, DOS och Windows. Vissa är gratis, andra skall man betala för. En adress till samling av programvara för bl.a. statistik:

archives.math.utk.edu/software.html

Tre PC-program (DOS) värda att pröva

EPI-info (ver 6.04b) är ett gratis programpaket utvecklat av bl a WHO. Paketet är främst tänkt för insamling och analys vid medicinska undersökningar, men är flexibelt och ypperligt för bl.a. enkätundersökningar (inmatning och korstabuleringar och analyser) samt enklare statistisk analys. Data kan importeras och exporteras från/till olika dataformat och med bra flexibilitet.
EPI-info program:
www.cdc.gov/epo/epi/downepi6.htm.
"Online" manual:
www.cdc.gov/epo/epi/introepi.htm

SSS1 (ver 1) är ett gratis program för analys av tidsserier, t ex skattning av ARIMA-modeller. Det kan hantera datafiler från bl a EPI-info. SSS1 program:

www.cdc.gov/epo/epi/downsss1.htm

SPCEX är ett s.k. shareware program, utvecklat av Mark Shewhart. Efternamnet är bekant, eller hur? Programmet är mycket lätt att använda. Det presenterar och analyserar data med bl.a. styrdiagram, histogram, paretodiagram, duglighetsindex m m. Under samma adress som SPCEX finns mycket annat att pröva. SPCEX hittas under adress:

deming.eng.clemson.edu/pub/tqmbbs/software/

Lycka till och mycket nöje!!
Mats Franzén, Ericsson Cables

Multivariat dataanalys 28 oktober 1997 i Göteborg

Varför?

Jo, vi skall vid vår årliga konferens försöka förklara vad som menas med multivariat dataanalys och vad den kan användas till.

Under de senaste åren har intresset för att analysera sitt data med multivariata metoder ökat. Med multivariata metoder menar man då allt från multipel regression till diskriminantanalys, principalkomponentanalys, PLS eller kanonisk analys.

En forskargrupp inom kemometri vid Umeå universitet har under många år använt multivariata metoder för att analysera kemiska processer där man mätt många variabler men inte så många replikat. Inte minst de moderna analysinstrumenten har hjälpt till med den utvecklingen. Detta har visat sig vara en mycket framgångsrikt varför intresset för metoderna har vuxit inom t ex skogsindustrin och andra verksamheter med kemilaboratorier.

Som följd av ökat intresse har flera mycket bra programvaror för sådana analyser kommit ut på marknaden varför tillgängligheten för gemene man är stor. Därför har det uppstått ett stort behov av utbildning.

Denna konferens får ses som en början på den utbildningen där ni bl a får stifta kontakt med arbetsplatser som redan börjat använda dessa tekniker och ta lärdom av deras erfarenheter.

Planeringen för konferensen har just börjat varför ni har stora möjligheter att påverka innehållet. Om ni vill bidraga med ett inlägg eller har ett bra förslag på föredrag hör av er till Susanna Weinberger eller Lennart Nilsson, Matematisk statistik, Umeå universitet, 901 87 Umeå, tel 090-166077, e-post ln@matstat.umu.se

Styrelsen

Ordförande:
Susanna Weinberger
Ovako Steel AB
813 82 Hofors
0290 – 252 96

Sekreterare:
Clas Mellby
IVF
Argongatan 30
431 53 Mölndal
031 – 706 61 02

Kassör:
Anders Hynén
ABB Corporate Research
avd R
721 78 Västerås
021 – 32 32 24

Ledamöter:
Ann Brännström-Stenberg
Högskolan i Karlstad
Institutionen för teknik
651 88 Karlstad
054 – 83 82 36

Lennart Nilsson
Matematisk statistik
Universitetet
901 87 Umeå
090 – 16 60 77

Göran Lande
Ericsson Radio Systems AB

Sture Rydin
Stora Cell AB
814 81 Skutskär
026 – 851 07

Mats Franzén
Ericsson Cables AB
824 82 Hudiksvall
0650 – 361 61

Redaktionskommitté:
Lennart Nilsson
Ingemar Sjöström
Susanna Weinberger

Bidrag accepteras gärna via 3.5"-diskett med textmängden i format WordPerfect, Word e.d.

Man blir medlem i SFK–StaM genom att kontakta Svenska Förbundet för Kvalitet telefon 08 – 783 82 54 eller 783 01 71 (fax 661 1967). Kanslisekreterare är Berit Winnsjö.

SFKs internet-adress

- http://www.sfkvalitet.se
  är adressen till Svenska Förbundet för Kvalitet

Som vanligt välkomnar vi bidrag från läsarna!