MER ÄN BARA ETT VÄRDE

SIDAN 12

TID FÖR NYA METODER I TILLÄMPAD STATISTIK

SIDAN 14

SSFr
Svenska statistikfrämjandet

# Qvintensen

MEDLEMSTIDNING FÖR SVENSKA STATISTIKFRÄMJANDET NR 2 2011

## $\pi=4$ – ett lagförslag

SIDAN 30

**Man har blivit slav under sina egna krav som man själv satt men inte förstått syftet med.»**

Magnus Arnér och Kerstin Vännman

SIDAN 7

SURVEYFÖRENINGEN/

## Lastbilsmätningar – bästa uppsats på surveyområdet

SIDAN 24

## Webbpaneler – vad är det?

SIDAN 22

# PRICKSÄKRA KUNSKAPER OM IMORGON. BARA SÅ DU VET.

**DET FINNS TRE SLAGS LÖGNER SA MARK TWAIN EN GÅNG: LÖGN, FÖRBANNAD LÖGN OCH STATISTIK.**

Vi håller inte med. I vår värld finns det tre slags sanningar: fakta, avancerad analys och statistik.

Det är med de grunderna som SAS Institute skapar kunskap och beredskap inför framtidens scenarion. Just lösningar för beslutsstöd med avancerad analys är vi världsledande på.

På sas.com/sweden hittar du alltid aktuella seminarier, webb-seminarier och utbildningar, vare sig du är befintlig eller blivande kund.

**Välkommen in.**

sas | THE POWER TO KNOW.

**SAS Institute** är ledande på lösningar för beslutsstöd med avancerad analys. SAS Institute, även världens största privatägda mjukvaruföretag. omsatte 2009 2,31 miljarder dollar i 118 länder och har sedan 1976 erfarenhet av att utveckla verktyg och metoder som låter stora organisationer lära av sin historia, mäta och kommunicera pågående aktiviteter och inte minst att skapa insikt om framtiden. Världen runt har SAS Institute totalt gjort 45 000 kundinstallationer, bland annat i 92 procent av Fortune 500-företagen. I Sverige startade SAS Institute AB år 1986 och har idag drygt 100 anställda på kontoret i Stockholm. Bland de svenska kunderna finns landets mest betydande företag och organisationer. För mer information besök: *www.sas.com/sweden*

Qvintensen
MEDLEMSTIDNING FÖR
SVENSKA STATISTIKFRÄMJANDET
NR 2 2011

Linda Ederyd.

Enligt ritning? Första appen?

Sophia Olofsson vann Surveyföreningens uppsatstävling.

## Innehåll



»Vilken av Svenska statistikfrämjandets föreningar blir först ut med en app?»

Joakim Malmdin, sidan 21

## Fasta spalter



## Våra föreningar

# Enkla, stora hjälpmedel


Framgång i en verksamhet uttrycks ofta som att ”minst x%” av produkterna ska uppfylla något krav. Det är ett simpelt kvalitetsmått. Jag tror att många suckar när miljöminister Andreas Carlgren begär ”85%” av SMHI och generaldirektör Lena Häll Eriksson lovar att myndigheten ska skärpa sig.

**– Jag har gett** ett uppdrag till dem som står för prognoserna att de faktiskt måste förbättra kvaliteten. De jobbar med det nu, säger SMHI:s GD. (DN:s webbversion 2011-05-18)

**Som om de** inte alltid gjort det. Många av er läsare kommer att känna igen sig i Magnus Arnérs och Kerstin Vännmans artiklar om det som ofta kallas duglighet. Tänk, en siffra, ett duglighetsindex, som talar om hur bra en produkt är. Vår längtan efter ett trollspö. Tänk att leta efter nycklarna, vifta avmätt med spöt, uttala ”accio nycklar”, och de kommer vinande. Eller ännu bättre: ”accio det där som jag glömt om det var något jag skulle komma ihåg”. Samtidigt behöver vi ett duglighetsindex. Men vi kanske inte kan trolla fram det. Och det kommer inte att vara alldeles enkelt att tolka. Men det går att göra: det finns ju forskningsresultat som visar statistiska egenskaper hos olika duglighetsindex.

**»Samtidigt behöver vi ett duglighetsindex.»**

Ett annat enkelt mått är p-värdet som ibland avgör skillnaden mellan misslyckande och framgång. Linda Ederyd skriver om p-värdens slumpmässighet.

**Det är fascinerande** med tekniska hjälpmedel. En del försvinner, en del blir praktiska hjälpmedel – och en del förändrar världen. Jag minns de ”datorintensiva metoderna”. Idag används de metoderna dagligen. Webbpaneler är också ”datorintensiva”. Folk sitter i sina hem och kommunicerar med frågeställaren. Webbpaneler använder det som förändrat världens maktstrukturer, det globala nätet som vanligt folk kan bli en liten del av. Läs mer om webbpaneler i Surveyföreningens sektion.

**Det är de** successiva förändringarna som förändrar världen. Sophia Olofsson skriver om det dagliga arbetets utredningar. Den typ av metoder som Sophia Olofsson använder har funnits i hundra år, senare förädlade av bl.a. Herman Otto Hartley under 60-talet. Numer finns de i läroböckerna och på arbetsplatserna. Kanske visionären Hartley kunde ana att det skulle bli så.

DAN HEDLIN

FOTO: MONICA HOLMBERG

Qvintensen

**Ansvarig utgivare**
Hans Nyquist
**Redaktör**
Dan Hedlin, 070-578 4334
Ingegerd Jansson (biträdande)
**Redaktion**
Hans Ekholm, Främjandet
Alf Fyhrlund, Främjandet
Marie Linder, FMS
Johan Lyhagen, Främjandet
Lennart Nordberg, Surveyföreningen
Lars Söderström, Industriell statistik
Ann-Marie Flygare, Cramérsällskapet
**Adress**
Qvintensen, c/o Dan Hedlin
Runbyvägen 31, 194 44 Upplands Väsby
**E-post** qvintensen@gmail.com
**Produktion**
Form och redigering: Mezzo Media AB
Tryckeri: Trydells Tryckeri AB

**Annonser**
Annonser i Qvintensen bokas med redaktören.
Annonsutskick per e-post bokas med Svenska statistikfrämjandets sekreterare.
Priset är 5 000 kronor för institutionsmedlem eller motsvarande, och
7 500 kronor för övriga annonsörer. På alla priser tillkommer 25 % moms.

Svenska statistikfrämjandet

**SVENSKA STATISTIKFRÄMJANDETS STYRELSE**
**Ordförande**
Hans Nyquist, 08-16 11 32,
hans.nyquist at stat.su.se
**Vice ordförande**
Jun Yu, 090-78 68 468
**Sekreterare**
Chandra Adolfsson
Pernilla Andersson
**Skattmästare** Jonas Selling
**Klubbmästare** Therese Andersson
**Surveyföreningen**
Joakim Malmdin, 08-506 949 45
**FMS** Mats Rudholm, 031-703 73 72
**Industriell statistik**
Göran Lande, 070-526 26 18
**Cramérsällskapet**
Jan-Eric Englund 040-41 52 36
**Övriga ledamöter**
Henrik Hult, Mathias Lindholm
**Adress**
Svenska statistikfrämjandet
c/o Chandra Adolfsson
Statistiska centralbyrån
701 89 Örebro
**E-post** sekrfram@gmail.com
**Webbplats** www.statistikframjandet.se

ISSN 2000-1819

SSFr
Svenska statistikfrämjandet

ORDFÖRANDEN HAR ORDET

# ÅRETS STATISTIKFRÄMJARE ÄR GUNNAR KULLDORFF

**Ännu ett verksamhetsår** för Svenska statistikfrämjandet lades till handlingarna i och med att Främjandets tredje årsmöte genomfördes. Årsmötet, vårkonferensen och medlemsföreningarnas årsmöten hölls på Örebro universitet. Mandatperioden för Sune Karlsson (vice ordförande), Ann Lindqvist (representant för Industriell statistik) och Karin Dahmström (representant för Surveyföreningen) hade gått ut och de lämnade styrelsen. Tidigare under året har Andreas Nordvall Lagerås lämnat styrelsen på grund av stor belastning i hans ordinarie arbete. Under den senare delen av verksamhetsåret har Paulina von Rahmel varit adjungerad till styrelsen som sekreterare. I och med årsmötet lämnar även hon styrelsen. Jag vill tacka de avgående ledamöterna för deras arbete med att forma och leda vår förening. De nya ledamöter som valdes in i styrelsen är Jun Yu (vice ordförande), Pernilla Andersson (sekreterare), Mathias Lindholm (ledamot), Göran Lande (representant för Industriell statistik) och Joakim Malmdin (representant för Surveyföreningen).

Hans Nyquist, ordförande.

**»Jag vill tacka de avgående ledamöterna för deras arbete med att forma och leda vår förening.»**

**Vid det föregående** årsmötet antogs en etisk kod som skulle utvärderas under året. Efter förslag från den arbetsgrupp som har utvärderat koden under året fastställde årsmötet den etiska koden. Den finns att läsa på Svenska statistikfrämjandets webbplats.

**Enligt nuvarande stadgar** beslutar årsmötet om medlemsavgiftens storlek för det kommande året. Det innebär att utskick om de nya medlemsavgifterna inte kan ske förrän efter årsmötet och då kommer utskicket att vartannat år administreras av en nytillträdd sekreterare och skattmästare. Det kan innebära att utskicken av räkningar för medlemsavgifterna försenas högst avsevärt. Vid årsmötet diskuterades ett förslag om att årsmötet i stället skall besluta om nästföljande års medlemsavgift. Det skulle då vara möjligt för mer rutinerade sekreterare och skattmästare att kunna skicka ut räkningarna tidigt på året. Eventuellt kan beslut i denna fråga tas vid nästa årsmöte.

**Den stora arbetsuppgiften** under det gångna året har varit att skapa en fungerande webbplats. Denna till synes enkla uppgift visade sig vara långt mer komplicerad och mödosam än vad vi trodde. Vi har dock nu en webbplats och medlemsföreningarna arbetar för fullt med att arbeta upp sina respektive webbplatser. Den stora uppgift som nu ligger framför oss är att skapa ett enhetligt medlemsregister. Visionen för medlemsregistret är att det skall vara lätt att arbeta med samt att det skall vara gemensamt för de olika medlemsföreningarna och tillgängligt för föreningarnas styrelser.

**I samband med** årsmötet hölls också vårkonferensen, där temat berörde ämnets utveckling. Hans Stenlund berättade om diskussionerna som föregick inrättandet av de första professurerna i statistik och Torbjörn Thedéen berättade om några problemområden där statistisk teori har kunnat ge viktiga bidrag. Vidare berättade Tommy Johnsson om statistikämnets kopplingar till vetenskapsteori. I det sista föredraget presenterade Jef Teugels sina synpunkter på behovet av statistiska föreningar och deras roll.

**Årets statistikfrämjare är** en viktig utmärkelse som har inrättats av Främjandet och delas ut i samband med årsmöteskonferensen. Årets statistikfrämjare är Gunnar Kulldorff. Gunnar har under en lång tid främjat statistiken. Några poster i hans meritlista är att han är professor i båda våra ämnen, statistik och matematisk statistik, han var mycket aktiv i utvecklingen av Svenska statistikersamfundet och har vid två tillfällen varit samfundets ordförande. Han har också varit president i International Statistical Institute. För närvarande arbetar Gunnar med att utveckla ett samarbete inom statistikområdet mellan de nordiska länderna, Baltikum och Ukraina. Jag vill gratulera Gunnar till utmärkelsen och för hans ytterst värdefulla arbete med utvecklingen av vårt ämne.

## Uppsala först ut

De första professurerna i statistik (ej matematisk statistik):

| | | |
|---|---|---|
| Uppsala universitet | Gustav Sundbärg | 1910 |
| Lunds universitet | Sven Dag Wicksell | 1926 |
| Stockholms högskola | Sten Wahlund | 1938 |
| Göteborgs högskola | Hannes Hyrenius | 1951 |
| Umeå universitet | Gunnar Kulldorff | 1965 |

Gustav Sundbärg.

Sven Dag Wicksell.

# ÄMNETS UTVECKLING

**På temat Ämnets utveckling talade Hans Stenlund och Torbjörn Thedéen under Svenska statistikfrämjandes vårkonferens i Örebro. Hans Stenlund redogjorde för tillkomsten av de första professurerna i statistik i Sverige, varefter Torbjörn Thedéen speglade ämnets utveckling från 1950-talet till idag genom sin karriär.**

Hans Stenlund påminde om att Sverige kring förra sekelskiftet var ett land i stark industriell utveckling, där parlamentarismen men ännu inte den allmänna rösträtten gjort sitt intåg. SCB och dess föregångare hade samlat in data i omkring 150 år inom en stor mängd områden. Den första motionen om att inrätta professurer i statistik i Lund och Uppsala skrevs 1906 av högermannen Nils Trolle och liberalen Julius Centerwall. Vilket intresse dessa båda herrar personligen hade av statistik är något oklart, men deras argument för motionen var att det fanns brister i den svenska statistikproduktionen beroende på brister i utbildningen av tjänstemännen. Hans refererade den vältaliga debatten i kammaren 1906 där bland andra Branting yrkade bifall. Motionen avslogs och nya motioner lämnades in 1907 och 1908. Även de avslogs. 1909 kom dock en proposition i samma ämne som bifölls. Gustav Sundbärg kunde därför 1910 tillträda den första professuren i statistik i Uppsala. I Lund inrättades dock ingen professur och inte heller delade man på den professur som fanns i statskunskap och statistik trots förslag från universitetet. Däremot kom det förslag att omvandla professuren i Uppsala till en professur i nationalekonomi och statistik, häpnadsväckande nog från dåvarande professorn själv, Nils Wohlin. Detta avslogs dock och 1926 inrättades en professur i statistik i Lund som tillträddes av Sven Dag Wicksell.

Rubriken för Torbjörn Thedéens anförande var *Statistikämnet, modeller och tillämpningar*. Med många underhållande anekdotiska utvikningar fick åhörarna följa hans karriär från operationsanalys under militärtjänsten till matematisk statistik, vidare till trafiksäkerhetsforskning, som var stort på 1950- och 1960-talen, över till kriminalstatistik på BRÅ. Energikommissionens arbete under 1970-talets senare hälft ledde till arbete kring riskanalys (numera säkerhetsanalys) av kärnkraft, ett område med mycket modeller men lite data. När Erland von Hofsten behövde en assistent till valvakorna i tv påbörjades en mångårig sidokarriär som expert i tv. Torbjörns slutord om att ”det finns så många intressanta arbeten med så spännande arbetsuppgifter” är det nog många som kan instämma i.

JAKOB BERGMAN, LUNDS UNIVERSITET

# Att mäta duglighet hos en tillverknings-process

En av författarna minns ett e-brev som kom för några år sedan och som handlade om dugligheten hos en tillverkningsprocess. Det nådde författarens ögon efter att många läsare inom organisationen under några få dagar hade tillfogat sina synpunkter och kommentarer. E-brevet väckte förundran över hur enkla tumregler kan överskugga det ändamål de tillkommit för och hur lätt det ibland kan vara att koppla bort sitt sunda förnuft.

På en inköpsavdelning i bilindustrin, som det i detta fall handlade om, arbetar såväl ekonomer som ingenjörer. Ingenjörernas uppgifter är till stora delar inriktade på kvalitetsteknik. De ska avgöra om leverantörerna jobbar med statistisk processtyrning på det vis de ska, om de mätmetoder de använder håller tillräckligt hög kvalitet och om produktionsprocessen och den levererade produkten uppfyller de krav som ställs. Bland dessa krav finns kapabilitet (även kallad duglighet), d.v.s. produktionsprocessens förmåga att producera enheter som ligger inom satta toleransgränser. Den som skrivit det första meddelandet hade riktat sig till leverantören. Leverantören hade ett år tidigare sänt in en kapabilitetsrapport där man hade ett s.k. kapabilitetsindex på 7,53. Nu hade ännu en kapabilitetsrapport skickats in, men denna gång var indexet 6,14. Det hade tolkats som att leverantören hade blivit sämre och måste förbättra sig.

**När sedan e-brevet** cirkulerat inom företaget anmärkte den ena efter den andra att förändringen av ett kapabilitetsindex från 7,53 till 6,14 kanske inte är så alarmerande, trots allt tal om ständiga förbättringar. Men paragrafryttarna var inte av samma åsikt. Det var minst sagt intressant att följa diskussionen och ta del av argumenten. Men låt oss lämna denna brevväxling ett tag för att se varför kapabilitetsindex har fått en så betydande roll och vad det medfört. Trots att begreppet fått till följd att statistikens roll i tillverkningsindustrin stärkts och slumpmässig variation hamnat mer i fokus, så är vår erfarenhet av dess användning inom industrin inte odelat positiv. Vi kommer därför att ta upp ett antal erfarenheter vi har av rena missbruk av kapabilitetsindex.

För den som inte har så stor erfarenhet av kapabilitetsindex finns en beskrivning på sidan 10 om vad det är. Vi använder i fortsättningen begrepp som toleransgränser och beteckningar som $C_{pk}$; dessa förklaras i den beskrivningen.

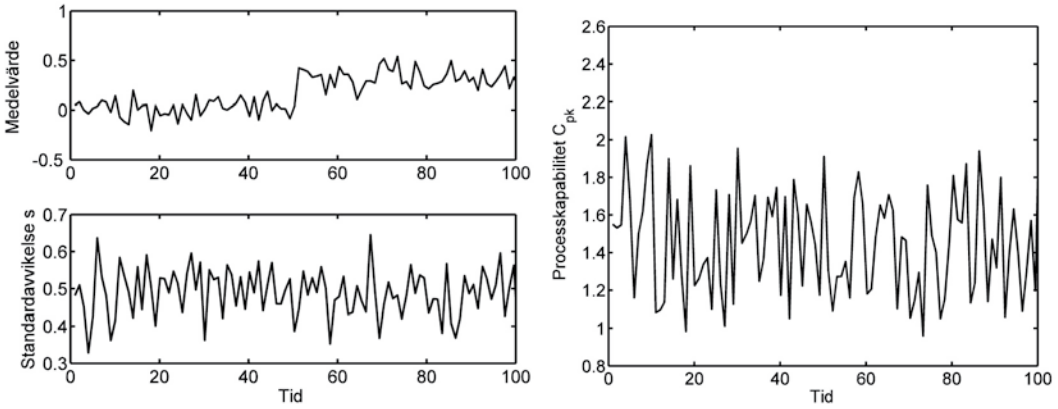
### *Ett skenbart enkelt kvalitetsmått*

Ett index av den typ som $C_{pk}$ och alla dess släktingar utgör är den typ av mätetal som mången inom industrin söker efter med ljus och lykta. Eftersom det är enhetslöst och normerat så tycker sig alla förstå vad det är. En standardavvikelse är svårare att tolka. Vad innebär det att $\sigma = 0{,}7$ mm? Är det stort eller litet? Har man däremot ett $C_{pk}$ på 0,87 vet alla att det är för litet. Det står nämligen i regelverket. Gränsen anges där kanske till 1,33. Understigs denna gräns så är det för dåligt. Denna enkelhet ger oanade möjligheter. Man kan t.ex. ställa krav på leverantörerna ("kan ni inte påvisa att ni har ett $C_{pk}$ på minst 1,33 så får ni inte leverera").

Det finns fler mätetal som är av denna typ, till synes enkla att förstå för var och en, men försåtligt svåra vid en närmare anblick. Just kapabilitetsindex utmärker sig även bland dessa. Den skenbara enkelheten har lett till missbruk, som gör att man många gånger blir förvånad och ibland också upprörd.

### *Slav under kravet Cpk ≥ 1,33*

Ett kapabilitetsindex förutsätter att det finns toleransgränser. Dessa ska sättas med utgångspunkt i produktens funktion. Men sådana gränser är ibland svåra att sätta, och de som finns kan dessutom vara satta utan eftertanke. Om man då har ett krav på $C_{pk} \geq 1{,}33$ så kan man »»»




**Figur 1. Medelvärde och standardavvikelse respektive processkapabilitet plottade mot tidpunkten då stickprovet tagits.**

få betala mycket stora summor till leverantören eller i investeringar i sin egen tillverkningsprocess utan att det kanske skulle behövas. För att undvika dessa kostnader händer det att man räknar ut vilken standardavvikelse man har och sedan sätter toleransgränserna på så sätt att $C_{pk} \geq 1{,}33$ uppnås. Det innebär att toleransgränserna sätts helt utan någon koppling till att produkten ska tillhandahålla avsedd funktion. Arbetet med kapabilitetsindex blir bara en byråkratisk verksamhet. Man har blivit slav under sina egna krav som man själv satt men inte förstått syftet med.

### *Kapabilitetskrav då toleransgränser saknas*

Att man skulle ha kapabilitetskrav då toleransgränser saknas eller ännu inte satts kan tyckas befängt och är det också. Men inom vissa företag används begreppet kapabilitet i tid och otid, så pass mycket att de som inte förstått vad det är faktiskt efterfrågar kapabilitetsindex även när toleransgränser saknas. Att efterfråga en standardavvikelse hade i detta fall varit det rimliga.

### *Kapabilitetsindex kan dölja viktig information*

Kapabilitetsindex som $C_{pk}$ och $C_{pm}$ (se sidan 10) väger ihop information om processens läge (μ) och processens spridning (σ) och relaterar den till toleransgränser och eventuellt satta målvärden. Om man då glömmer bort att följa upp hur processens läge och spridning varierar var för sig så kan man missa viktig information om sin tillverkningsprocess. Innan man ens börjar mäta kapabilitet bör man via styrdiagram över läge och spridning ha förvissat sig om att processen är stabil eller i s.k. statistisk jämvikt. Om så inte är fallet så blir slutsatserna från ett kapabilitetsindex inte speciellt användbara eftersom man då bara får en ögonblicksbild och inte ett mått på vad som gäller i långa loppet. Men om man nu anser sig ha en stabil process så bör man inte sluta med styrdiagrammen över läge och spridning och ersätta dem med enbart ett diagram som visar hur kapabilitetsindexet varierar. Det framgår av följande exempel.

Exempel: Antag att man för en viss egenskap $X$ hos en produkt har toleransgränserna $LSL = -0{,}7$ och $USL = 1{,}0$. Vid tidpunkterna 1, 2, … tar man ett stickprov av storlek $n = 30$ från tillverkningsprocessen. Fördelningen för $X$ är $N(0, 0{,}5)$ fram till tiden 50, därefter händer det något. Man får ett skifte i väntevärdet och $X$ blir $N(0{,}3, 0{,}5)$. Skiftet i väntevärdet syns väldigt tydligt om man studerar medelvärdesdiagrammet i Figur 1. Om man däremot studerar enbart $C_{pk}$ i den högra figuren så upptäcks inte förändringen.

Det finns alltså en risk att man missar processförändringar om man enbart studerar kapabiliteten utan att samtidigt studera styrdiagram för medelvärden och standardavvikelser. Tyvärr förekommer det i praktiken.

### *Kapabilitet och icke-stabila processer*

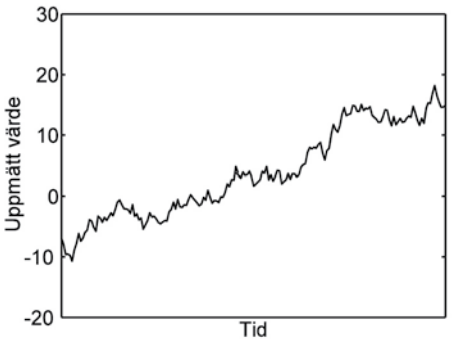
Kravet på att en process bör kunna antas vara stabil innan man drar slutsatser om processens kapabilitet nämndes ovan. Ett typiskt exempel på en icke-stabil process är när man har drift i tillverkningssystemet, som illustreras i Figur 2. Om man skattar $C_{pk}$ okritiskt genom att beräkna medelvärde och standardavvikelse för värdena över tidsperioden så får man ett skattat värde på indexet. Man vad säger egentligen det värdet? Trots det uppenbart ointressanta med ett sådant värde så ser man ibland kapabilitetsindex som räknats ut trots att processen inte är stabil.

### *Hur har man tagit sitt stickprov?*

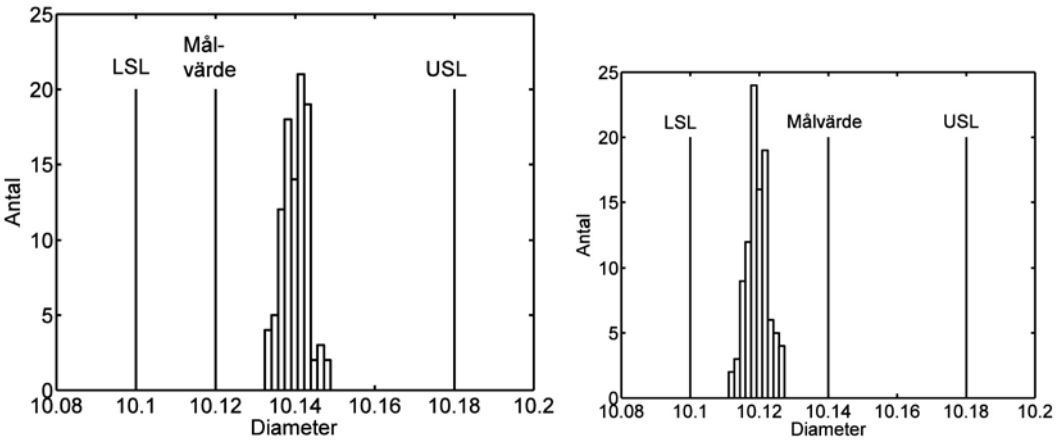
Mönstret man ser i Figur 2 skulle kunna inträffa även om tillverkningsprocessen inte uppvisar någon drift utan därför att man har en autokorrelerad process, så att två intilliggande observationer tenderar att vara ganska lika. Har man då tagit sitt stickprov under en alltför kort tidsperiod fångar man inte upp hela tillverkningsprocessens variation. Man kommer då att underskatta standardavvikelsen och typiskt få alltför optimistiska värden på sitt kapabilitetsindex. Svårigheten att få ett stickprov som är representativt för populationen underskattas ofta.

### *Skattade index är osäkra*

När man pratar om kapabilitetsindex skiljer man inte




**Figur 2. Tillverkningsprocess med drift.**




**Figur 3: Om man schablonmässigt ställer alla krav i form av $C_{pk}$ kommer man ibland att trigga ett felaktigt beteende. Att tänka efter vilket sorts kapabilitetsindex som är lämpligt är nödvändigt. I dessa bägge fall är $C_{pm}$ att föredra framför $C_{pk}$.**

alltid på de teoretiska indexen och deras skattningar. Man har sällan så stora stickprov att man kan bortse från att skattningarna är osäkra på grund av att de är beräknade från ett stickprov. Dessutom blir konfidensintervallen för indexen ganska breda om man har små stickprov. Tyvärr är det inte ovanligt att man har en stickprovsstorlek på 10 vid beräkningen av indexet. Blir det då skattade $C_{pk}$-värdet, säg, 1,4 så innebär det på intet sätt att man kan påstå med någon rimligt låg felrisk att det sanna $C_{pk}$-värdet överstiger 1,33. Man måste alltså alltid tolka det skattade indexet med hjälp av konfidensintervall eller test, vilket alltför sällan görs.

### *Diagram är viktiga*

När man rapporterar om kapabilitetsindex bör man också bifoga lämpliga diagram. Bland annat förutsätter de vanligaste kapabilitetsindexen att observationer är normalfördelade. Ett histogram eller ett normalfördelningsdiagram bör därför göras – och tolkas. Ser data inte normalfördelade ut så måste man tänka sig för. Man ska dessutom se upp och inte bara rapportera de diagram som den statistiska programvaran producerar utan också reflektera över det som kommer fram. Man kan ibland få se rapporterade histogram av den typ som visas i Figur 3, där man systematiskt missar målvärdet. Det första fallet ser man ofta då kontraktet mellan leverantör och kund bara innehåller ett krav på $C_{pk}$ men inget på $C_{pm}$. Det är faktiskt ganska vanligt och kan bero på en okunskap hos beställaren om att indexet $C_{pm}$ existerar. I det andra fallet har man kanske medvetet lagt sig nära den undre gränsen. Det kan bero på det ålderdomliga synsättet att allting innanför toleransgränserna är lika bra. Kan man då t ex minska på materialåtgången genom att göra godset tunnare så gör man det. Men så ska det naturligtvis inte vara utan man ska sikta på målvärdet. Dessutom ökar risken för att hamna utanför toleransgränsen vid en liten, plötslig ändring av processens läge. Då är det bättre att använda ett index som ställer krav på att tillverka mot målvärdet, exempelvis $C_{pm}$.

### *Åter till inledningens e-brev*

Hur var det nu med e-brevet som cirkulerade, det brev där leverantören fick krav på att förbättra sig? Vad var felet med inköparens reaktion? Man kan väl i stort dela upp det i tre delar. Den första är att sannolikheten att någonsin få en enhet som ligger utanför toleransgränserna är försvinnande liten, i vilket fall som helst med så höga värden på $C_{pk}$.

Den andra delen har att göra med att det värde som rapporteras från leverantören är en skattning baserad på ett stickprov och därigenom en observation av en stokastisk variabel. Man får då fråga sig om försämringen är statistiskt signifikant. Det är vanligt att man i industrin glömmer detta.

Den tredje frågan som man ställer sig är hur stickprovet tagits. Är det ett slumpmässigt stickprov från den avsedda fördelningen? Vår erfarenhet säger oss att anledningen till att man får kolossalt stora värden på $C_{pk}$ vanligen har sin förklaring i annat än att processen är extremt duglig. Ibland har man så vida toleransgränser att leverantören vet att de måste vara snävare för att avsedd funktion ska kunna tillhandahållas och bryr sig således inte om toleransintervallen vid tillverkningen. Men ännu vanligare är kanske att stickprovet bara fångar en liten del av variationen. Hela stickprovet kan komma från en och samma batch eller så kan det finnas tio gjutformar i fabriken och hela stickprovet är taget från en enda.

Dessa tre tillsammans gör att vi, om vi suttit som inköpstekniker, hade avstått från att kritisera leverantören. Däremot är det fullt möjligt att vi begärt in bättre uppgifter om hur stickprovet tagits.

### *Varför har kapabilitetsbegreppet blivit så populärt?*

Det finns företag, även ganska stora, där kapabilitetsbegreppet är fullständigt okänt. Trots det så vågar vi påstå att användningen av kapabilitetsindex vuxit sig ganska omfattande under de senaste åren. Man kan fråga sig varför, och vi ser tre besläktade orsaker.

**Sex sigma** (Six Sigma), som är en metodik och filosofi för kvalitetsarbete. Sedan sex sigma slog rot har det spritt sig som en löpeld och är troligen den främsta orsaken till den spridning som kapabilitetsbegreppet fått.

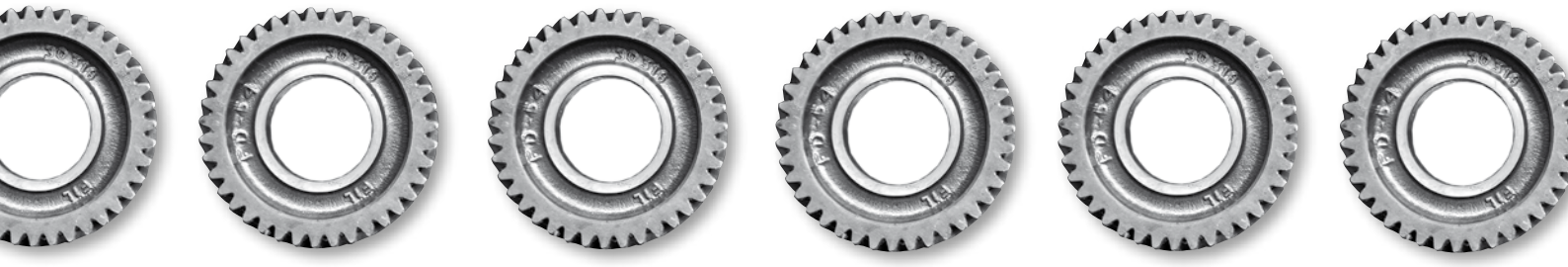
**QS 9000.** I den nordamerikanska bilindustrins gemensamma kvalitetskrav på dess underleverantörer, QS 9000, fanns en hel del om just kapabilitet. Eftersom man där generellt har kravet $C_{pk} \geq 1{,}33$ är det möjligt att detta, som blivit gängse i stora delar av industrin, kommer just därifrån. Numera har QS 9000 uppgått i ett internationellt regelsystem.

**Programvaror.** Numera finns moduler för kapabilitetsberäkning inbyggda i många programvaror. En särställning bland dessa intar Minitab. Kapabilitetsmodulen i Minitab tycks ha hämtat hela sin vokabulär och sitt beteckningssystem från QS 9000. Det är dock inte därför som Minitab intar en särställning; programvaran behandlas rätt noggrant i de flesta Six Sigma-utbildningar och har därmed fått en särställning i industrin.

Visserligen leder användningen av kapabilitetsindex ibland till konstigheter och premierar ibland brist på eftertanke. Men det innebär ändå större fokus på slumpmässig variation i industriell produktion, vilket vi tycker är bra.

Trots all den kritik vi framfört mot användningen av kapabilitetsindex är vi båda stora anhängare av begreppet. Men det måste användas med förstånd, och det är tyvärr inte alltid fallet. Vi kommer att följa upp den här artikeln i ett kommande nummer av Qvintensen och ge förslag på hur man kan komma till rätta med några av de brister i hanteringen av kapabilitetsindexen som vi har lyft fram.

MAGNUS ARNÉR, TETRA PAK
KERSTIN VÄNNMAN,
LULEÅ TEKNISKA UNIVERSITET OCH
UMEÅ UNIVERSITET



# VAD ÄR KAPABILITETSINDEX?

■ **Människor har i alla tider** förstått att man inte kan tillverka produkter exakt enligt de mått som finns på en ritning. Det kommer alltid att finnas en slumpmässig variation runt målvärdet. När den industriella revolutionen kom med sin massproduktion försvann möjligheten att fila på detaljer tills dess de passade ihop och gick att montera. Enheterna skulle kunna passa ihop utan justering. Man insåg också att en liten avvikelse kan tolereras och inte innebär att enheterna blir obrukbara. Toleransgränserna hade därmed sett dagens ljus. Toleransgränser finns för alla möjliga mått i tillverkningsindustrin. Det kan röra sig om allt från styvheten hos gummibussningar till diametern på stansade hål. Men hur väl lyckas vi producera enheter vars mått ligger inom toleransgränserna? För att kunna ange detta har begreppet kapabilitet uppkommit. På svenska används ibland duglighet som synonym till kapabilitet. Det engelska uttrycket är capability.

**Kapabilitet brukar mätas** med ett kapabilitetsindex. Det är ett mått, eller snarare en familj av mått, på hur väl en produktionsprocess kan producera enheter som ligger innanför de satta toleransgränserna. Det allra enklaste av dessa index betecknas $C_p$ och introducerades på 1970-talet. Det definieras som

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

där *USL* (Upper Specification Limit) och *LSL* (Lower Specification Limit) betecknar övre respektive nedre toleransgräns och σ står för processens standardavvikelse. Indexet $C_p$ jämför alltså toleransintervallets längd med 6σ, det som kommit att kallas för processens naturliga variation. Ju större värde på $C_p$, desto dugligare process. Detta kapabilitetsindex har troligtvis uppkommit med normalfördelningen i åtanke. Om normalfördelning gäller, processen är centrerad och indexet = 1 (d.v.s. $\pm 3\sigma$ ryms inom toleransintervallet) så förväntas det stora flertalet (99,7%) av de tillverkade enheterna finnas inom toleransgränserna. Ursprungligen var också en kapabilitet på minst 1,00 vanligen det man krävde, medan det numera är väl så vanligt att man vill uppnå 1,33 eller 1,5 eller t.o.m. 2,0.

**En klar nackdel** med kapabilitetsindexet $C_p$ framgår av Figur 1. Det tar inte hänsyn till processens läge i förhållande till toleransintervallets läge. En annan nackdel är att $C_p$ är odefinierad vid enkelsidiga toleranser.

Trots dessa nackdelar används $C_p$ rätt flitigt. Indexet brukar ofta beskrivas som processens potentiella kapabilitet. Orsaken är att tillverkningsprocessers läge ofta är rätt så lätta att justera medan deras spridning är besvärligare. Så potentialen hos en process ges av $C_p$ bara man lyckas skruva in den rätt.

**För att komma** bort från de brister som $C_p$ har introducerades indexet $C_{pk}$ på 1980-talet. Det definieras som

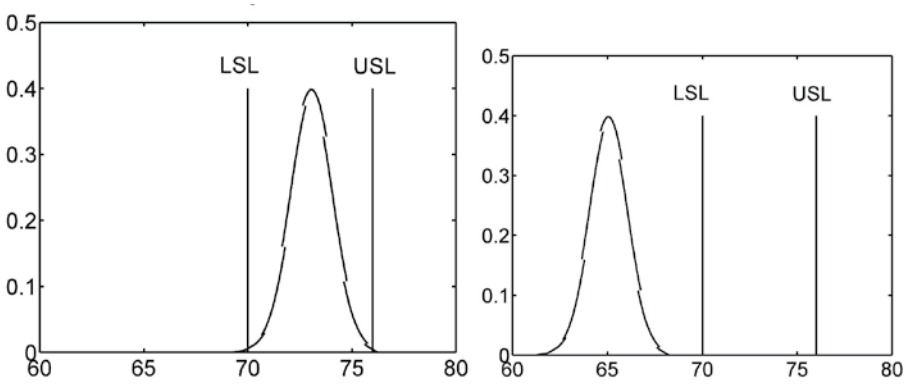
$$C_{pk} = \min\left\{\frac{USL - \mu}{3\sigma}, \frac{\mu - LSL}{3\sigma}\right\}.$$

Med detta index tar man, som Figur 2 visar, hänsyn till utfallets läge i förhållande till toleransintervallet.

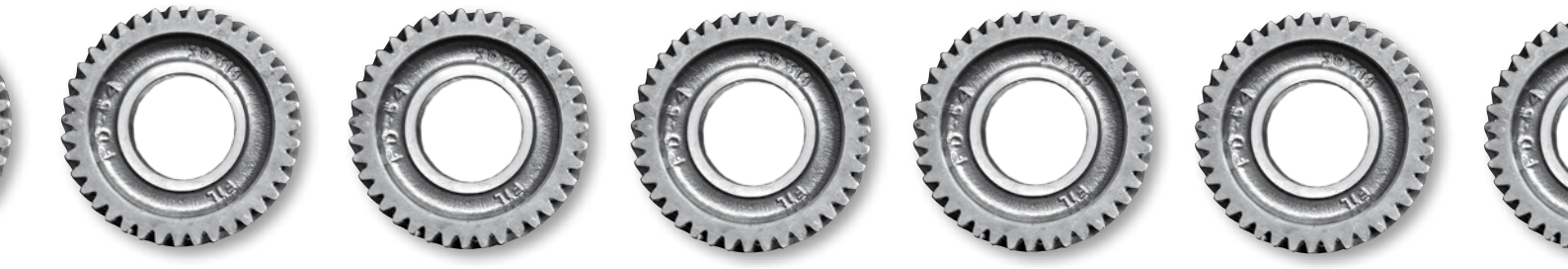
Användningen av indexen $C_p$ och $C_{pk}$ tog fart under 1980-talet då Ford Motor Company satsade på kvalitetsutveckling. De spreds sedan snabbt till övriga bilindustrin och vidare till många andra industriområden. Indexet $C_{pk}$ förbättrade $C_p$, men en stor nackdel återstår. Det tar inte hänsyn till målvärdet utan betraktar alla utfall inom toleransgränserna som lika bra. Så är det naturligtvis inte. Ju längre ifrån målvärdet man kommer desto sämre är det, även om man ligger kvar inom toleransgränserna. Om en bult får aningen för stor ytterdiameter så kanske den ändå passar till muttern, men då krävs istället ett något större moment för att skruva ihop dem, vilket är en försämring av processen. Nästa kapabilitetsindex som introducerades, också på 1980-talet, tar hänsyn till målvärdet *M* genom att väga in både processens spridning och avståndet till målvärdet *M*. Det betecknas $C_{pm}$ och definieras som

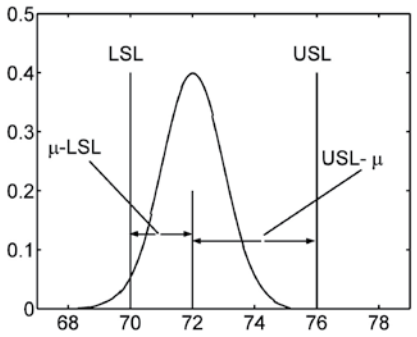
$$C_{pm} = \frac{USL - LSL}{6\sqrt{(\mu - M)^2 + \sigma^2}}.$$

Om man har en situation som i Figur 3 så kommer indexet $C_{pm}$ att vara lågt och signalera

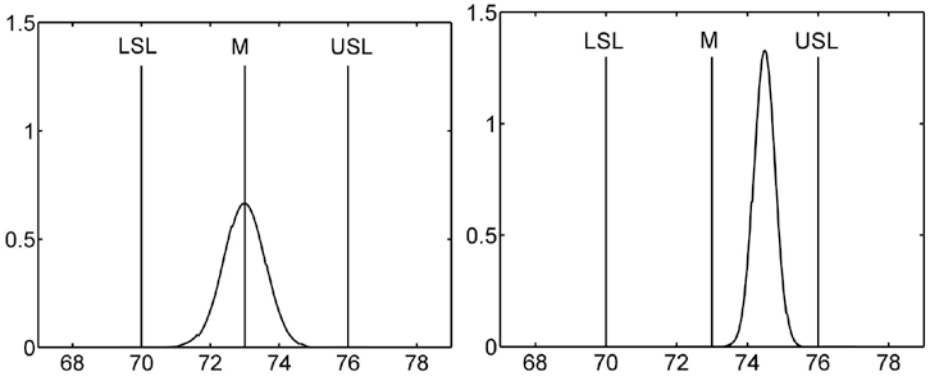



**Figur 1: Dessa bägge tillverkningsprocesser har samma värde på indexet $C_p$ fast det är helt tydligt att den vänstra är bra mycket bättre eftersom nästan alla produkter i den högra ligger utanför toleransgränserna. Längs x-axeln visas värdet på egenskapen hos den produkt som mäts.**






**Figur 2: Eftersom väntevärdet μ i detta fall ligger närmare nedre än övre toleransgräns så blir $C_{pk}$ = (μ-LSL)/(3σ).**
**Här är μ=72 och σ =1 vilket ger**
**$C_{pk}$ = (72-70)/3=0,67**




**Figur 3: Indexet $C_{pk}$=1,67 i bägge dessa fall. I det vänstra diagrammet gäller det även att $C_{pm}$= 1,67 medan $C_{pm}$=0,65 i fallet till höger.**

att processen inte är duglig eftersom den ligger långt från målvärdet medan indexet $C_{pk}$ indikerar att allt är OK. Men även om processens spridning är liten i detta fall så är det vanskligt att ligga så nära en toleransgräns. Det behövs ju bara en liten ändring av processens läge för att man ska börja producera en stor andel enheter utanför toleransgränserna. Det är inte heller bra att systematiskt hamna bredvid målvärdet. Målvärdet finns ju där för att det är just det bästa värdet som storheten kan anta. Varje avvikelse därifrån innebär en försämring. Ännu så länge har inte indexet $C_{pm}$ kommit att användas så mycket i industrin trots att det har tilltalande egenskaper. Bortsett från de här nämnda indexen så finns det ett stort antal andra som föreslagits i litteraturen.

**En annan aspekt** på kapabilitetsindex har att göra med vad man menar med processens standardavvikelse $\sigma$. Vilken population är det som har denna standardavvikelse? Man brukar då skilja mellan kort och lång tidshorisont. Den korta tidshorisonten kan vara inom en tillverkningsbatch eller inom ett produktionsskift. Man antar att alla batcher har samma standardavvikelse, $\sigma_{ST}$, där ST står för Short Term. Om man däremot betraktar dels variationen inom batcher, dels mellan batcher, får man ett större värde på standardavvikelsen, $\sigma_{LT}$ (Long Term). Man skiljer ibland på dessa till beteckningen (och även till namnet) och låter $C_p$ enbart beteckna den korta tidshorisonten. Man brukar då skriva

$$C_p = \frac{USL - LSL}{6\sigma_{ST}} \text{ och } P_p = \frac{USL - LSL}{6\sigma_{LT}}$$

och motsvarande för indexen $C_{pk}$ och $C_{pm}$. Bokstaven *P* uttalas då ibland *performance*. Oftast är $P_p$ intressantare än $C_p$ och $P_{pk}$ intressantare än $C_{pk}$. Man brukar därför lägga större vikt vid långtidskapabilitet. I denna artikel har vi valt att genomgående beteckna kapabiliteten med bokstaven *C* eftersom detta är gängse praxis.

**Det vi presenterat** hittills är teoretiska mått baserade på okända parametrar. Vår kunskap om processens kapabilitet kommer alltid att grunda sig på ett stickprov. Vi får därmed endast en skattning $\hat{C}_p$ av kapabiliteten. Under förutsättning att processen är normalfördelad så är det lätt att räkna ut sannolikhetsfördelningen för

$$\hat{C}_p = \frac{USL - LSL}{6s}.$$

Sannolikhetsfördelningen för

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

är däremot lite besvärlig eftersom det rör sig om ett minimum av två stokastiska variabler som var och en är erhållna som kvoter av två andra stokastiska variabler. När man drar slutsatser från skattade index ska man använda konfidensintervall eller hypotestest. Tyvärr bortser man ofta från detta i industriella sammanhang. Det kan leda till felaktiga beslut.

**Magnus Arnér.**

**Kerstin Vännman.**

**Även om ett** kapabilitetsindex rent formellt går att räkna ut med de uttryck som presenterats oavsett vilken sannolikhetsfördelning som tillverkningsprocessen än uppvisar så förutsätter tolkningen av indexen normalfördelning. Eftersom normalfördelningsantagandet inte alltid uppfylls så har det utvecklats en ganska omfattande teoribildning för kapabilitetsindex för andra fördelningar.

MAGNUS ARNÉR, TETRA PAK
KERSTIN VÄNNMAN,
LULEÅ TEKNISKA UNIVERSITET
OCH UMEÅ UNIVERSITET

**Linda Ederyd.**

**ARTIKELN BYGGER PÅ LINDA EDERYDS KANDIDATUPPSATS I MATEMATISK STATISTIK, "BESTÄMD SIGNIFIKANSNIVÅ ELLER P-VÄRDE, VILKET ÄR ATT FÖREDRA?"**

# Mer än bara ett

Ett p-värde är mer än bara ett värde, mer än bara en sannolikhet, det innehåller massvis av information. Ett p-värde är nämligen en slumpvariabel som kan följa en mängd olika fördelningar. Det är även viktigt att komma ihåg att en slumpvariabel inte bara är en variabel utan faktiskt en funktion.

Sådana här p-värden fås som resultat av tester av olika hypoteser, oftast en nollhypotes och en alternativ hypotes. Ett p-värde är enligt definitionen i dessa fall sannolikheten att få ett minst lika extremt värde som utfallet på den testvariabel som använts, givet att nollhypotesen är sann. Testvariabeln är en funktion av stickprovet i fråga. Detta medför att test av olika slag beror av slumpen i och med att de beror på stickprovets värden. Det man är intresserad av är att se om stickprovet verkar stämma överens med nollhypotesens antagande eller inte.

I de allra flesta test används p-värdet som ett mått på evidens för att en nollhypotes inte stämmer. Om p-värdet är litet så förkastas nollhypotesen. Vi vill veta med så stor säkerhet som möjligt att en alternativ hypotes verkar vara sann innan vi tror på den.

Ett p-värde kan ses som en observation av en slumpvariabel P, vilken då måste ha en fördelning. Denna beror dock på vilken hypotes som är den sanna, det vill säga den beror på stickprovets bakomliggande fördelning. Det enklaste fallet är då nollhypotesen är sann. Då kommer P alltid att vara likformigt fördelad mellan värdena 0 och 1. Detta går att bevisa (Donahue 1999), men följande är en simulering som stödjer resonemanget.

**Bild 1 visar** resultatet av 1000 stycken t-test på stickprov av storlek 100 där varje observation är ett slumpmässigt värde från en normalfördelning med väntevärde 0 och standardavvikelse 1. Nollhypotesen är att $\mu=0$.

Vi ser att i de fall som nollhypotesen är sann så är det lika sannolikt att få ett högt p-värde som att få ett lågt. Ett högre p-värde behöver alltså inte betyda att nollhypotesen verkar vara mer sann än vad som indikeras av ett lägre värde. (Här talar vi om signifikanstest.) Det är således viktigt att komma ihåg att p-värdet inte är sannolikheten att nollhypotesen är sann!

I de fall då den alternativa hypotesen är sann beror p-värdets fördelning på stickprovets bakomliggande fördelning. Vi tar här upp två exempel. Vi utför z-test och låter stickproven bestå av 100 stycken oberoende värden från en $N(\mu,1)$-fördelning. Vi ansätter hypoteserna $H_0$:$\mu\leq 0$ och

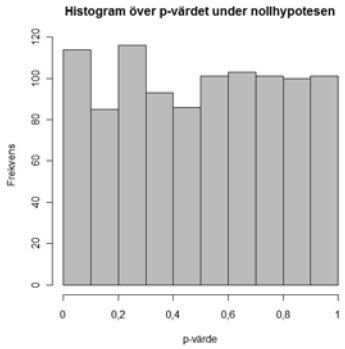



**Bild 1. Histogram över p-värdet då nollhypotesen är sann**

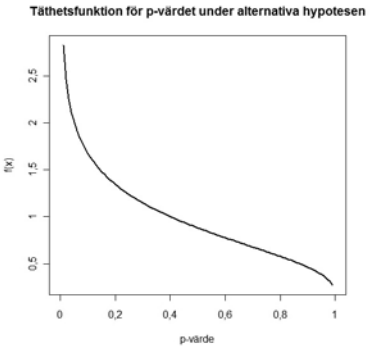



**Bild 2. N(0,05, 1), stickprovs-storlek 100**

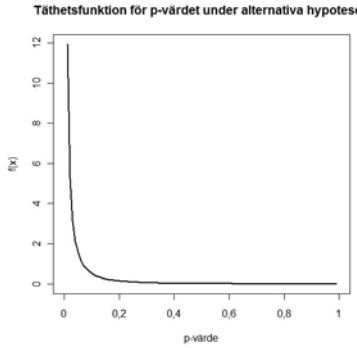



**Bild 3. N(0,3, 1), stickprovs-storlek 100**

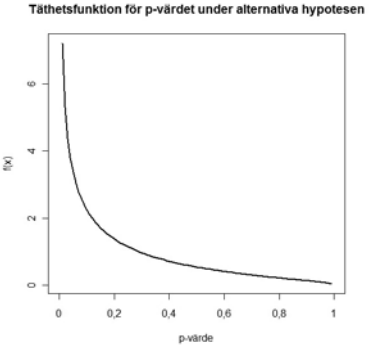



**Bild 4. N(0,05, 1), stickprovs-storlek 500**

# värde

$H_1$:$\mu>0$. I de fall då $\mu$ är 0,05 respektive 0,3 ser p-värdets fördelning ut som i Bild 2 och 3.

Vi ser att ju närmare nollhypotesens värde det sanna värdet ligger, desto mer lik blir fördelningen för p-värdet den likformiga fördelningen. Det är större sannolikhet att få ett signifikant p-värde då det sanna värdet ligger längre ifrån nollhypotesens värde.

I praktiken är det inte enkelt att finna p-värdets fördelning, men om man har många p-värden så kan man se om de verkar vara likformigt fördelade eller ej.

Nollhypotesen förkastas om p-värdet är lågt och då tror vi mer på den alternativa hypotesen än på nollhypotesen. Om p-värdet däremot är högt så kan vi inte anta att nollhypotesen är sann. Det kan nämligen finnas andra hypoteser, ofta liknande nollhypotesen, som kan vara den sanna. Vi kan dock få en aning om att nollhypotesen faktiskt ligger nära verkligheten.

**I och med** att p-värdet beror på stickprovet så är det bland annat viktigt att tänka på hur stort stickprov man behöver dra. Antag att den alternativa hypotesen är sann. Om vi har ett litet stickprov så är det svårare att förkasta nollhypotesen då slumpen spelar stor roll angående om de få värdena vi valde representerar en hel population på ett tillförlitligt vis. Om vi däremot tar ett större stickprov så förkastas nollhypotesen oftare då stickprovet mer kommer att likna populationen. Tar vi dock riktigt stora stickprov så kommer vi kunna förkasta nollhypotesen ofta, även i de fall då denna är mer eller mindre sann. Skillnader mellan parametrarna i nollhypotesens fördelning och parametrarna för den sanna fördelningen kan komma att detekteras hur små de än är. Om man därför bara är intresserad av om skillnaden är stor så räcker ett relativt litet stickprov för att finna denna skillnad.

För att påvisa att p-värdet beror på stickprovets storlek visas här åter en bild av p-värdets fördelning då den alternativa hypotesen är sann och z-test genomförs, se Bild 4. Stickprovet består nu av 500 stycken oberoende värden (istället för 100) från en $N(\mu,1)$-fördelning med $\mu = 0{,}05$. Hypoteserna är desamma: $H_0$:$\mu\leq 0$ och $H_1$:$\mu>0$.

Vi har nu med hjälp av simuleringar och exempel sett att p-värdet är en slumpvariabel. I och med detta kan det vara intressant att dela upp ett större slumpmässigt stickprov i flera små. På så vis kan flera test utföras och många p-värden fås fram. Då kan man försöka finna p-värdets fördelning för att se om denna verkar vara likformig eller inte. Denna metod kan då användas istället för att se om endast ett p-värde är signifikant.

Om du vill läsa mer om p-värden så kan artiklarna i referenslistan rekommenderas.

LINDA EDERYD, UPPSALA UNIVERSITET

**Referenser**

Cox, D. R. (1982), Statistical Significance Tests, British Journal of Clinical Pharmacology, Vol. 14, pp. 325-331.

Donahue, R. M. J. (1999), A Note on Information Seldom Reported Via the P Value, The American Statistician, Vol. 53, pp. 303-306.

Lambert, D., Hall, W. J. (1982), Asymptotic Lognormality of P-values, The Annals of Statistics, Vol. 10, pp. 44-64.

Statistiskt förberedelsearbete med variabelkonstruktion och editering av data kan ha nytta av en allmän metodik baserad på entropier som är förhållandevis okänd inom tillämpad statistik. Entropibegreppet kommer från statistisk mekanik och informations- och kommunikationsvetenskap och används också i teoretisk statistik. **Det är min uppfattning att begreppet entropi lämpar sig för tillämpad statistik och borde förklaras och användas i elementära läroböcker i statistik.**

Ove Frank.

# Tid för nya metoder i tillämpad statistik!

Jag vill i den här artikeln argumentera för ökad användning av begreppet entropi i tillämpad statistik genom att belysa hur entropimått kan användas för att bedöma variabilitet hos en- eller flerdimensionella variabler på alla datanivåer från nominalskala till kvotskala. Olika sammansatta entropimått kan också användas för att bedöma samband mellan variabler. Sambanden kan vara funktionella samband i matematisk mening men också associativa samband i statistisk mening. Entropimetodikens styrka är att den utan modellantaganden kan göra förutsättningslösa analyser av multivariata data. Den passar också utmärkt för explorativ analys för att gruppera utfall och variabler på ändamålsenligt sätt.

### *Statistikens metoder och data*

När det numera är möjligt att genomföra snabba och omfattande bearbetningar av stora datamaterial är det en risk att man underskattar behovet av bra data och tror att man kan pröva sig fram med olika variabler och metoder tills man hittar några lämpliga resultat. Det räcker dock inte med datorvana utan man måste också ha datavana för att lyckas, d.v.s. man måste ha känsla för valet av variabler och kunna bedöma kvalitet och eventuellt behov av kvantifiering av observationer och mätningar. Det statistiska förberedelsearbetet med variabelkonstruktion, selektion av variabler, val av datanivåer och variabelvärden, justeringar för bortfall och annan editering av data är lika viktigt som den bearbetning man gör för att pröva teorier eller för att skatta förekomsten av effekter och trender om sakförhållanden av skilda slag.

**»Det räcker dock inte med datorvana utan man måste också ha datavana»**

Variabelkonstruktion nämns inte ofta i statistiska läroböcker, utan det tas i regel för givet att man har tillgång till lämpliga diskreta eller kontinuerliga numeriska variabler. Ändå är alla medvetna om att det är långt från givet hur man på bra sätt mäter t.ex. arbetsmiljö, bostadsstandard eller hälsostatus för individer och befolkningsgrupper. Diskreta kvalitativa variabler behandlas som kategoriska variabler med etikettvärden som ibland väljs som heltal och till och med behandlas som sådana. I samhällsvetenskapliga tillämpningar brukar man skilja på datanivåer i termer av nominalskala, ordinalskala, intervallskala och kvotskala och anpassa metodiken därefter. Vid behandling av data från enkätundersökningar brukar det nämnas att frågornas konstruktion är väsentlig för vilka möjligheter man har att bearbeta svaren. Variabelkonstruktion kan också innefatta val av index och olika sätt att kombinera flera variabler till ett sammanfattningsmått.

Speciell statistisk metodik har utvecklats inom många områden och statistikern kan bidra med att sprida kunskap om metoder som visat sig framgångsrika på något område och visa hur de kan nyttiggöras också inom andra områden. Statistiska metoder för interaktiva partikelsystem i fysiken har också kunnat tillämpas inom psykometri och sociometri för att studera interaktion mellan individer och sociala grupper. Inom informations- och kommunikationsteori har entropibaserade metoder utvecklats som också har spritts till flera andra områden. De har använts inom teoretisk statistik och för kodning och datasäkerhet, men deras potential har inte uppmärksammats tillräckligt i allmän tillämpad statistik.

### *Kort om entropi*

Entropi är ett begrepp som används för att beskriva ordning och oordning (t.ex. bland gasmolekyler) eller koncentration och utspridning (t.ex. av värme). Oordning och utspridning motsvarar stor entropi. I statistiska sammanhang kan man säga att variabler med liten entropi är koncentrerade till få vanliga eller typiska värden som har stora sannolikheter, medan variabler med stor entropi är utspridda på många ovanliga

**Ökande entropi.**

eller perifera värden som har små sannolikheter. Det kan bara finnas ett begränsat antal värden med stora sannolikheter men obegränsat många med små sannolikheter. Svårigheten är att veta vilka sannolikheter som är så små att värdena kan anses perifera. Entropin mäter i viss mening detta, d.v.s. hur många värden som är centrala och som därför inte bör bortses ifrån. Entropin H är logaritmen för ett centralitetsmått c som på visst sätt kombinerar oddsen för olika möjliga värden. T.ex. kan c=7,75 tolkas som att det finns nästan 8 centrala värden och att entropin är nästan log(8)=3 om man mäter osäkerheten i bitar, d.v.s. med basen 2 för logaritmen. Se artikeln bredvid för en något fylligare beskrivning av entropibegreppet.

### *Centralitet och spridning*

För en kvalitativ variabel X med r olika utfall på nominalskala eller ordinalskala är ett naturligt centralvärde det s.k. typvärdet, d.v.s. det utfall som har störst sannolikhet eller de utfall som har den största sannolikheten om det finns flera med samma. Om det finns många utfall med små sannolikheter kan det vara lämpligt att bortse från de mest perifera. De utfall som är så centrala att man inte bör bortse ifrån dem bildar en s.k. centralregion. Dess storlek ges av det minsta heltal som är större än eller lika med centralitetsmåttet c. Typvärden ingår i centralregionen men den innehåller i allmänhet även andra utfall. »»»

# Entropi-begreppet

**Betrakta ett experiment** som har r olika möjliga utfall numrerade med heltal 1,…,r. Låt först utfallen ha samma sannolikhet 1/r för k=1,…,r. Vid n oberoende sådana experiment finns $r^n$ möjliga utfall som alla har samma sannolikhet $1/r^n$. Låt nu istället experimentets sannolikheter $p_k$ vara godtyckliga positiva tal med summan 1 för k=1,…,r. Då finns fortfarande $r^n$ möjliga utfall vid n experiment men deras sannolikheter varierar beroende på hur mycket experimentets utfallssannolikheter varierar. Man kan visa att när n växer så blir sannolikhetsfördelningen alltmera lik en likformig fördelning över ungefär $c^n$ centrala (typiska, vanliga) utfall, där c är ett tal mellan 1 och r som beror av sannolikheterna men inte av n. De $r^n$ utfallen vid n upprepningar av experimentet kan uppdelas i en grupp perifera utfall som tillsammans har liten sannolikhet och en grupp centrala utfall som tillsammans har en sannolikhet som konvergerar mot 1 när n växer.

För att få en uppfattning om hur utfallssekvenserna utvecklas när deras längd ökar kan man resonera på följande sätt. Sannolikheten för en sekvens av längd n där utfall k inträffar $n_k$ gånger är en produkt över alla k av $p_k$ upphöjt till $n_k$. Förväntade värdet för $n_k$ är $np_k$. Sekvenssannolikheten är därför ungefär lika med produkten över alla k av $p_k$ upphöjt till $np_k$, $[\prod_k p_k^{p_k}]^n$ .

Detta är ett genomsnittligt sannolikhetstal upphöjt till n. Genomsnittstalet ligger mellan 0 och 1 och kan därför betecknas som exp(-H) för något positivt tal H eller som 1/c för något tal c större än 1. Med en aning algebra kan man visa att dessa tal ges av formeln

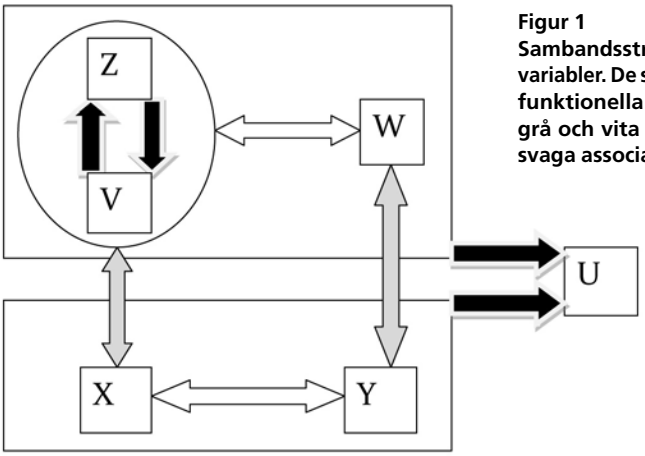
$H = \log(c) = \Sigma p_k \log(1/p_k)$ där summan tas över k=1, ..., r.

Talet H kallas entropin och c centraliteten för sannolikhetsfördelningen $(p_1, \ldots, p_r)$ eller för variabeln X. Talen H och c kan tolkas som mått som styr tillväxten av centrala utfallssekvenser vid upprepade experiment. En region som består av minst c utfall som väljs efter avtagande (icke-växande) utfallssannolikheter kan generera de centrala utfallssekvenserna av godtycklig längd. Ingen mindre region kan generera utfallssekvenser som tillsammans har en sannolikhet som konvergerar mot 1. Talen H och c kan också tolkas som genomsnittliga osäkerhetsmått. Man definierar osäkerheten i ett utfall som dess inverterade sannolikhet, vilket väsentligen är det oddsbegrepp som förekommer i vissa spelsammanhang. Entropin H är då den förväntade logaritmiska osäkerheten, och c är det med sannolikheterna vägda geometriska medelvärdet av osäkerheterna.

En annan tolkning av H, som gäller kodteori, innebär att om man vill representera vanliga utfall med korta kodord och ovanliga utfall med längre kodord, så utgör H en gräns för hur liten den förväntade längden kan bli med användning av trädgenererade s.k. prefixkoder. Tre utfall med sannolikheter 1/2, 1/4, 1/4 kan t.ex. binärkodas enligt 0, 10, 11, varvid förväntade längden blir minimal och lika med entropin H=1,5 uttryckt i bitar (bits, binary digits). Om sex utfall med sannolikheterna 1/2, 1/4, 1/16, 1/16, 1/16, 1/16 binärkodas enligt 00, 01, 100, 101, 110, 111 så blir förväntade längden 2,25 bitar, men entropin är H=2 och en bättre kodning är 0, 10, 1100, 1101, 1110, 1111 om man vill ha minimal förväntad längd.

Om en stokastisk variabel X har sannolikheten p(x) för utfall x i ett ändligt utfallsrum R så kan man bilda en ny positiv stokastisk variabel som mäter den logaritmiska osäkerheten $U = \log[1/p(X)]$. Entropin för X kan då uppfattas som väntevärdet för U vilket ibland är ett praktiskt sätt att enkelt få fram entropin. Entropin kan definieras på samma sätt för variabler X med uppräkneligt oändliga utfallsrum R om väntevärdet för U existerar. Också för kontinuerliga stokastiska variabler X med täthetsfunktion p(x) för reella tal x kan entropin definieras som väntevärdet för U när det existerar. Om X är en flerdimensionell variabel och p(x) den flerdimensionella sannolikheten eller sannolikhetstätheten kan också samma definition tillämpas. I samtliga fall kan c uppfattas som storleksmått för den centrala regionen i utfallsrummet R.

OVE FRANK




**Figur 1**
**Sambandsstruktur mellan sex variabler. De svarta pilarna visar funktionella samband och de grå och vita visar starka resp. svaga associationer.**

## »»» TID FÖR NYA METODER I TILLÄMPAD STATISTIK!

För godtyckliga diskreta variabler med r möjliga utfall kan man använda H och c som spridningsmått och H/log(r) och c/r som relativa spridningsmått. Entropin H kan variera mellan gränserna 0 och log(r). Undre gränsen antas enbart när det finns bara ett utfall och övre gränsen enbart när man har en likformig fördelning över r utfall. På motsvarande sätt varierar centraliteten c mellan gränserna 1 och r.

Om X är normalfördelad med väntevärde μ och standardavvikelse σ kan man visa att entropin kan approximeras av $H = \log\sqrt{(2\pi e\sigma^2)}$, och den centrala regionens storlek är ungefär $c = 4{,}13\sigma$. Eftersom centrala utfall har större sannolikhetstäthet än perifera, är den centrala regionen ett intervall placerat symmetriskt kring väntevärdet, dvs. $\mu \pm c/2 = \mu \pm 2{,}07\sigma$. Intervallsannolikheten är ungefär 96 %. Den centrala regionen ringar alltså in variationen i en normalfördelning med cirka två standardavvikelser kring väntevärdet på konventionellt sätt. Liknande resultat fås också för flerdimensionella normalfördelningar. Den centrala regionen blir ellipsformad på konventionellt sätt för den tvådimensionella normalfördelningen.

### *Samband*

De entropibaserade måtten som gör det möjligt att behandla centralitet och spridning på ett enhetligt sätt för alla datanivåer och variabeltyper är också synnerligen lämpliga för att studera samband av olika slag. Korrelationskoefficienten mäter som bekant enbart graden av linjära samband och förutsätter alltså numeriska variabler. Andra statistiska associationsmått finns av flera olika slag men inga erbjuder samma flexibilitet och enhetlighet som entropimåtten. Det som gör entropimåtten så flexibla är att de kan lokalisera och mäta graden av samband mellan variabler av godtyckliga dimensionstal genom att på ett systematiskt sätt jämföra olika variabelgruppers centralregioner. Det är möjligt att specificera entropimått som mäter exakta och approximativa funktionella samband såväl som betingade och obetingade statistiska oberoenden eller associativa samband.

När man vill jämföra två variabler X och Y beräknar man deras entropier H(X) och H(Y) men också entropin för simultanfördelningen H(X,Y). Beräkningen av multivariata entropier sker enligt den allmänna formeln som ett viktat medelvärde av logaritmerna för de inverterade sannolikheterna för alla möjliga kombinationer av värden på variablerna. Man kan visa att H(X) är högst lika med H(X,Y) som är högst lika med summan av H(X) och H(Y). Om H(X) och H(X,Y) är lika, d.v.s. om X har samma centralitet som (X,Y), så betyder det att centrala utfallssekvenser av (X,Y) också kan genereras enbart från X. Intuitivt betyder detta att Y är överflödig och kan härledas från X. Detta är också fallet. Man kan visa att likhet mellan H(X) och H(X,Y) bara gäller om Y är en matematiskt given funktion av X, d.v.s. varje utfall av X leder till exakt ett speciellt utfall av Y.

Man kan också visa att den övre gränsen för H(X,Y), som är H(X)+H(Y), antas endast när X och Y är stokastiskt oberoende. Om man uttrycker den simultana fördelningen för (X,Y) med hjälp av marginalfördelningen för X samt den av X betingade fördelningen för Y så kan man definiera en betingad entropi som H(Y|X)=H(X,Y)-H(X). Denna betingade entropi för Y antar sitt maximala värde H(Y) endast vid oberoende. Avvikelsen från det maximala värdet utgör ett mått på associationen mellan X och Y.

För datamaterial med många observationer på flera diskreta variabler, t.ex. variabler som kodar svaren på ett batteri enkätfrågor, kan det vara svårt att överblicka sambandsstrukturen. Ett praktiskt sätt är att systematiskt jämföra alla univariata, bivariata och trivariata entropier bland de olika variablerna. Liksom man med de bivariata entropierna kan undersöka om en variabel är funktionellt beroende av en annan, kan man med trivariata entropier H(X,Y,Z) konstatera när Z är en funktion av (X,Y). På liknande sätt kan man också fortsätta och kartlägga funktionella samband av högre grad. Trivariata entropier avslöjar också oberoenden och betingade oberoenden. T.ex. är X och Y oberoende betingat av Z när den trivariata entropin är lika med H(X,Z)+H(Y,Z)-H(Z).

Eventuellt kombinerar man sökandet efter funktionella samband och oberoenden med sökande efter redundanta variabler och slopande av dessa. Redundanta variabler kan vara sådana som är konstanta eller nästan konstanta och har entropi 0 eller nära 0. Man kan också betrakta variabler som redundanta om de är funktioner av andra variabler och ingen av dessa har slopats.

När samband och associationer är kartlagda kan man skapa överskådliga diagram för variabelstrukturen. Man plottar variablerna som noder, funktionella beroenden som kraftiga pilar från en inringad grupp av variabler till den av dem beroende variabeln, och parvisa associationer som pilar mellan variabler. Eventuellt skiljer man på starka och svaga associationer genom pilar av olika tjocklek. Detta kan ske med hjälp av de gränsvärden för associationerna som svarar mot olika signifikansnivåer gentemot oberoende. En illustration av hur ett sådant diagram kan se ut för en struktur bland sex variabler visas i Figur 1. En utförligare genomgång av den här metodiken har jag beskrivit i en översikt kallad ”Statistical information tools for multivariate discrete data” som ingår i en volym från Springer som snart publiceras i serien Understanding Complex Systems. Metodikens potential behöver nyttiggöras och bli lättillgänglig, och det kräver utveckling av ändamålsenlig programvara.

OVE FRANK,
STOCKHOLMS UNIVERSITET

Statistiska test väcker ständigt nya diskussioner. I förra numret frågade sig Ulf Jorner om vi behöver enkelsidiga test. Martin Ribe för här den frågan vidare, med svar av Ulf. Efter denna omgång drar vi streck i debatten – för den här gången.

# Enkelsidiga test för stabila hypoteser

Ulf Jorner ifrågasätter i sitt inlägg i Qvintensen nr 1/2011 (s. 21-22) om det ”någonsin i praktiken är motiverat att göra enkelsidiga test”. Något djärvt kan det tyckas, när läroböcker har med enkelsidiga test i verktygslådan. Ulf Jorners argumentering har dock onekligen en väsentlig poäng.

Jag undrar om inte en skiljelinje kan dras mellan flyktig och stabil verklighet, som jag här väljer att kalla det. Flyktiga hypoteser avser en viss tidsperiod och kan röra t.ex. opinioner, konjunkturer eller atmosfären. Efteråt är det flyktiga känt enbart genom en given mängd efterlämnade data.

Stabila hypoteser å andra sidan rör mera tidlösa drag. Det kan vara effekter under specificerade förutsättningar av t.ex. läkemedel, miljögifter, formulärdesign eller undervisningsformer. Det stabila går att undersöka igen och skaffa nya data om. (Distinktionen är en lite annan än Wolds mellan icke-experimentella och experimentella data [2] [4].)

För test av flyktiga hypoteser håller jag med om att dubbelsidigt test är rimligt att ha som regel. Med enkelsidigt test i planen så är det kört om effekten som ska testas visar sig gå åt andra hållet än det väntade. Strikt sett går det då inte att dra någon slutsats.

För test av stabila hypoteser däremot bör enkelsidigt test kunna vara adekvat om det finns starka skäl att vänta sig en effekt åt ett visst håll, t.ex. på naturvetenskaplig kunskapsgrund. En fördel för enkelsidigt är då att testets styrka (power) vanligen blir bättre i relevanta områden än med dubbelsidigt.

Har man nu lagt fast en plan med enkelsidigt test, så visar sig baksidan om utfallet ändå blir en effekt åt andra hållet än det väntade. Då blir det en förlust i tempo och resursinsats. Men knorren är att möjligheten till framtida kunskap står orubbad, om testet handlar om stabil verklighet.

Jag kan i det läget tänka mig följande recept. Rapportera och informera om utfallet som planerat. Analysera explorativt vad som kan ligga bakom utfallet, på både befintliga och nya data. När man på nytt känner sig mogen att försöka dra en kontrollerat säkerställd slutsats (oavsett vilken), kör då i skarpt läge, alltså konfirmativt gällande, ett väl valt test på helt nya data. Processen ska naturligtvis skötas med statistikerförstånd, så att man klart redovisar turerna och tar hänsyn till problematiken som följer av multipla test [1] [3].

Eller tänk på ett $\chi^2$-test. Där kan det hända att utfallet på teststatistikan hamnar osannolikt nära noll, troligen orsakat av förbisedda beroenden. Inte skulle väl någon för den skull tänka sig att göra ett vanligt $\chi^2$-test dubbelsidigt (vilket formellt skulle gå, genom att kritiska området skulle få innehålla även ett intervall invid noll med sannolikheten p/2 under nollhypotesen). Inträffar ett utfall av det nämnda slaget gör man rimligen rätt i att vara öppen om det, reda ut vad som förbisågs, och sedan börja om men inte glömma.

MARTIN RIBE, STATISTISKA CENTRALBYRÅN

**Referenser**

[1] Multiple comparisons, www.wikipedia.org , English.

[2] A. Taube, Mest medvind, Uppsala: Uppsala Publishing House, 2006 (se s. 160).

[3] A. Torrång, 3 tsk framtid och 1 msk dåtid, Qvintensen nr 1/2011 (se s. 27).

[4] H. Wold, Causality and econometrics, Econometrica, 22 (1954), 162-177 (se s. 168).

Slutreplik av Ulf Jorner →

# Statistiska förvillelser i Qvintensen

Vad accepteras för publicering och vilket ansvar tar redaktionen för innehållet i införda artiklar? Materialet i Qvintensen kan ju vara av olika typ, t ex information om aktiviteter, referat från konferenser, recension av böcker, presentation av egen forskning, debattartiklar etc. I dessa exempel är det troligen uppenbart att det är författaren som har ansvar för innehållet och redaktionens uppgift är att bedöma om innehållet är av intresse för läsarna eller ej. Mer problematiskt blir det med artiklar som presenterar statistiska metoder och synsätt. Här tror jag det behövs en kritisk granskning så att inte vilka stolligheter som helst publiceras, åtminstone inte utan kommentarer. Risken är annars att en publicering i Qvintensen kan uppfattas som om innehållet i artikeln sanktionerats av statistisk expertis. Ett exempel är en artikel av Sven Ove Hansson, *Signifikansmissbruket i pseudovetenskap*, som publicerades i Qvintensen nr 3/2010. Avsikten med artikeln är tydligen att diskutera vanliga missförstånd om statistisk signifikans, men artikeln presenterar huvudsakligen en synnerligen märklig syn på statistisk hypotesprövning. I stort sett samma artikel har sedan publicerats även i nr 4/2010 av Folkvett, organ för Vetenskap och Folkbildning, vars uppgift är att främja folkbildning om vetenskapens metoder och resultat. I artikeln i Folkvett hänvisas till att den redan är publicerad i Statistikfrämjandets tidskrift Qvintensen och jag har förstått att detta uppfattas som ett starkt stöd för att den är statistiskt korrekt. På detta sätt kommer alltså Statistikfrämjandet att bidra till spridningen av statistiska förvillelser och det var väl inte riktigt meningen.

**Vad kan man** då göra? Mitt förslag är att redaktionen utnyttjar granskare på liknande sätt som vetenskapliga tidskrifter. Jag menar inte att allt ska granskas utan tror att redaktörens fingertoppskänsla kan avgöra när det kan vara befogat. Efter en granskning avgör man sedan om en artikel ska refuseras eller publiceras med kommentarer. Vad som behövs är en lista på personer, som representerar olika kompetensområden och är villiga att ställa upp som granskare.

GÖRAN NILSSON

# SLUTREPLIK OM TEST

Den distinktion som Martin Ribe för in, mellan flyktiga och stabila situationer, är intressant. Vi är nog helt överens om att enkelsidiga test finns, t.ex. i läroböckerna. Vi är också överens om att i det som Martin kallar flyktiga situationer är det inte meningsfullt med enkelsidiga test, åtminstone inte av skälet att man anser sig ha förhandsinformation om att endast en ”svans” är intressant.

Jag kan dock tänka mig ett slags situation där enkelsidiga test ändå är motiverade i en flyktig situation. Det skulle vara situationer där man endast är intresserad av att gardera sig mot fel åt det ena hållet. Exempel kan vara inom rättsväsendet (har en person minst x ‰ i blodet) eller inom myndighetsutövning (är en förorening över y ppm). I dessa fall skulle jag t.o.m. kunna tänka mig enkelsidiga konfidensintervall, ett annars ganska ovanligt redskap i statistikerns verktygslåda.

Martins huvudpoäng när det gäller stabila situationer är att man då kan upprepa undersökningen och signifikanstestet. Och resonemanget om hur man hanterar en stor avvikelse åt ”fel” håll är tänkvärt. Jag ser dock vissa svårigheter att i praktiken skilja mellan de två situationerna. Dessutom befarar jag att det enkelsidiga testet i praktiken kan få ”för bra” power även i en stabil situation.

Till sist $\chi^2$-test. Är de inte i praktiken dubbelsidiga? Tag t.ex. ett test av oberoende. Här testar man oberoende mot alla slags beroenden. Den lägre svansen ger en helt annan typ av information, Martin är vänlig nog att peka på förbisedda beroenden. En mindre vänlig person kan börja misstänka datas kvalitet. Om jag inte minns fel finns det klassiska exempel, både ärtor och ihjälsparkade kavallerister, som uppvisar osannolikt bra överensstämmelse med teorin.

ULF JORNER

# SVAR: ROLIGT ATT QVINTENSEN VÄCKER DEBATT

**Vi vill gärna** ha mer diskussion i Qvintensen. Sven Ove Hanssons artikel Signifikansmissbruket i pseudovetenskap i Qvintensen nr 3/2010 väckte debattlust och det är roligt. Göran Nilsson har skrivit ett kritiskt inlägg; detta och svar från Hansson finns på Svenska statistikfrämjandets webbplats under fliken Qvintensen. Där finns också en artikel av Lars Ängquist som diskuterar tolkningar av Hanssons ”missförstånd 3”, med svar från Sven Ove Hansson.

**Göran Nilsson ställer** olika frågor och framför synpunkter i sin insändare. Jag ska försöka besvara dem i min roll som redaktör för Qvintensen.

Göran Nilsson har själv satt rubriken ”statistiska förvillelser i Qvintensen” och syftar på Sven Ove Hanssons artikel. Rubriken och insändaren kan dock bygga på ett missförstånd. Det är lätt att läsa ”nollhypotes” där det står ”hypotes”. Jag tycker att Hanssons artikel är utmärkt. Hanssons ”missförstånd 3” har föranlett en del diskussion, bland andra av som sagt Lars Ängquist. Göran Nilsson diskuterar även detta i sitt inlägg som finns på webbplatsen.

**Göran Nilsson frågar** vad som accepteras för publicering i Qvintensen. Vi ”accepterar” i stort sett allt ej tidigare publicerat material som faller inom Qvintensens område, omfång, nivå och stil. Till exempel åtta sidor långa, tekniska artiklar passar inte in i Qvintensen. Texter som har förbättringspotential försöker vi hjälpa till med. Erfarenheten visar att en del material i praktiken inte når publicering eftersom författaren drar tillbaka sin text eller redaktionen väntar på svar på frågor eller på förtydliganden från författaren.

Göran Nilsson frågar vilket ansvar redaktionen tar för innehållet för införda artiklar. Vi läser noggrant och det brukar gå många vändor mellan redaktion och författare innan en text är klar. Målet är alltid att artiklar ska vara av värde för en bred grupp läsare. Qvintensen ska vara vederhäftig. Skulle ”stolligheter” (Göran Nilssons ordval) ändå publiceras i Qvintensen bemöts de lämpligen antingen i kommande nummer eller på Svenska statistikfrämjandets webbplats.

**Göran Nilsson anför** att ”det som behövs är en lista på personer, som representerar olika kompetensområden och är villiga att ställa upp som granskare.” Jag tycker att vi har det. Den listan finns att läsa på sidan 4 i varje tidning under biträdande redaktör och redaktion. Det händer att vi frågar ytterligare personer om hjälp av olika slag.

Göran Nilsson anför även att Qvintensen ska utnyttja granskare på liknande sätt som vetenskapliga tidskrifter. Det har vi inga planer på. Qvintensen ska inte vara en vetenskaplig tidskrift. I en sådan görs bedömningar i olika steg om ett insänt förslag på artikel ska accepteras, gå vidare i granskningsprocessen eller refuseras. Framgångsrika vetenskapliga tidskrifter refuserar långt fler artiklar än vad de publicerar. I Qvintensen arbetar vi på det sätt som mycket kortfattat beskrivs ovan.

DAN HEDLIN

## Mer på Qvintensens webbplats

Sven Ove Hanssons artikel ”Signifikansmissbruket i pseudovetenskap” i Qvintensen nr 3/2010 väckte debatt. På Svenska statistikfrämjandets webbplats under fliken Qvintensen finns inlägg av Göran Nilsson och Lars Ängquist med svar från Sven Ove Hansson.

## Har du flyttat?

Meddela din nya adress genom att gå in på Svenska statistikfrämjandets webbplats www.statistikframjandet.se, välj Medlemsregister och följ instruktionerna. Ditt användar-id är tryckt ovanför din adress på tidningens baksida.

## Bli medlem

Svenska statistikfrämjandets syfte är bland annat att främja sund användning av statistik som beslutsunderlag och att väcka och sprida intresse för statistik i samhället. Om du önskar gå med som medlem i Svenska statistikfrämjandet skicka ett e-brev till Chandra Adolfsson och Pernilla Andersson på sekrfram@gmail.com. Du får Qvintensen i brevlådan och platsannonser via e-post. Det ställs inga krav för att bli medlem; alla som är intresserade av statistik och vill stödja statistikens roll i samhället är välkomna.

# NYA I SURVEYFÖRENINGENS STYRELSE

## Åsa Rosendahl, kassör

– Jag arbetar som statistiker på Försäkringskassans analys- och prognosavdelning. Innan dess har jag jobbat på SCB och Skogsstyrelsen med bland annat arbetskraftsundersökningen och undersökningar bland skogsägare och entreprenörer. Surveyundersökningar har alltid varit min hjärtefråga och det område som intresserat mig mest. Jag trivs bäst när jag får arbeta med helheten, allt från att dra urval och skapa enkäter till att sedan analysera resultatet. Jag har nyss gått på föräldraledighet och jag och familjen har passat på att flytta upp till Luleå där jag ser fram emot att fortsätta i samma bana.

## Lennart Nordberg, ledamot

– Som nybliven pensionär tillhör jag väl knappast Surveyföreningens stall av påläggskalvar trots att jag är nyinvald i styrelsen. Under perioden 1996-98 var jag med i styrelsen för Statistikersamfundet så jag har åtminstone viss erfarenhet av sådant arbete. I styrelsen för Surveyföreningen kommer jag bland annat att vara föreningens representant i redaktionskommittén för Qvintensen. Jag arbetade vid SCB i 28 år fram till pensioneringen, den mesta delen av tiden som metodstatistiker, bland annat med jordbruks-, energi- och ekonomisk statistik och även med mera generell metodutveckling. I sammanlagt sex år var jag chef för olika metodgrupper. Innan jag började vid SCB var jag ett antal år vid KTH där jag disputerade i matematisk statistik 1981.

## Silke Burestam, ledamot

– Jag jobbar på Stockholms stads utrednings- och statistikkontor som statistiker/ projektledare och IT-chef. USK AB är ett kommunalägt bolag som är specialiserat på kommunal och offentlig verksamhet och som framför allt tar fram underlag för verksamheten inom Stockholms stad. Jag arbetar huvudsakligen med varierande uppdrag inom utbildning på uppdrag av stadens utbildningsförvaltning men även med metodstöd och IT-frågor. Jag har en masterexamen i statistik från Stockholms universitet. Privat ägnar jag nästan all min tid åt friidrott. Jag sitter i styrelsen för DN Galan som bland annat kommer att arrangera Lag-EM på Stockholms stadion i juni.

## ORDFÖRANDEN HAR ORDET

# Vem blir först ut med en egen app?


Var tid har sina kvantitativa mått. Just nu mäts allt, från tillförlitlighet och sanningshalt till popularitet, med antal träffar på någon av de populära sökmotorerna, vanligtvis Google. Något nedslående är det då att "Surveyföreningen" föräras med endast 260 träffar med detta sökverktyg. Mer upplyftande är det att sådant som engagerar oss i föreningen resulterar i fler träffar, "bortfall" till exempel, får hela 331 000 träffar. Bortfallsproblemet är tillsammans med webbpanelundersökningar de två områden som tycks intressera oss surveystatistiker mest just nu. Men vare sig ett stort bortfall eller en självvald webbpanel var något som pionjärerna inom surveyteorin hade i åtanke när riktlinjerna för hur en statistisk undersökning ska genomföras ritades upp en bit in på 1900-talet. Problematiken med bortfall stöts och blöts från två håll – dels från det statistiskt teoretiska, dels från det beteendevetenskapliga. Inom statistikteorin görs stora ansträngningar för att vi ska kunna använda oss av information som nödvändigtvis inte direkt handlar om det vi frågar efter eller är intresserade av, så kallad hjälpinformation. Ett exempel ges av Sophia Olofsson i det här numret av Qvintensen. Från det mättekniska och beteendevetenskapliga hållet försöker man vrida och vända på det faktum att medborgarna inte längre är lika villiga att delta i undersökningar. Kan vi ställa frågorna på andra sätt och i andra sammanhang? Varför är det så att många i den yngre generationen gladeligen lämnar ut information om sig själva via sociala medier, men inte är lika benägna att fylla i enkäter?

**»Varför är det så att många i den yngre generationen gladeligen lämnar ut information om sig själva via sociala medier, men inte är lika benägna att fylla i enkäter?»**

**Vad gäller webbpanelundersökningar** är inställningen tudelad. Å ena sidan kan vi se stora möjligheter i och med att insamlingssättet är så kostnadseffektivt och, för var dag som går, allt mer lättillgängligt. Den yngre generationen är hemtam när det gäller datorer, smarta mobiltelefoner och allehanda e-tjänster. Här gäller det för oss att hänga med och dra nytta av de nya kommunikationsmöjligheterna. Vilken av Svenska statistikfrämjandets föreningar blir först ut med en egen app?

**Det problematiska är** å andra sidan förstås att personer i en webbpanel i hög grad själva väljer om de vill vara med i en undersökning eller inte. De är alltså inte slumpmässigt utvalda från en väl definierad population. Men vem vet vad som händer i framtiden. Ponera att varje lägenhet förses med en egen, och unik, elektronisk brevlåda som har samma funktion som dagens fysiska brevlåda. Kanske sker det i samband med att vår nyckel byts ut mot ett chip och våra lägenheter får en elektronisk identitet motsvarande de lägenhetsnummer vi nyligen fått oss tilldelade? Fram till dess arbetar undersökningsföretag och myndigheter vidare med frågor som har att göra med webbpanelundersökningarnas utmaningar. Surveyföreningens egen kommitté på området höll som bekant ett välbesökt seminarium i början av februari i år. Ingegerd Janssons referat från det finns att läsa på nästa sida. Om allt går vägen har vi att se fram emot en skrift i ämnet i början av nästa år.

**Det går fort** nu och vi får inte glömma bort att bara två generationer tillbaka i tiden såg det stora flertalet det som en plikt att svara på enkäter (från myndigheter). Även om det inte längre är på det viset och enkäterna har ändrat skepnad så har vi som arbetar med samhällsstatistik exakt samma uppgift som statistiker då hade – att bidra med vederhäftig information för analys och beslutsfattande.

Surveyföreningens nya styrelse har tagit form och vi ser framför oss ett år då vi får ordning på föreningens webbplats, håller ytterligare ett par seminarier med aktuella surveyteman och mot slutet av året utlyser tävlingen om bästa uppsats med inriktning mot surveystatistik.

Hör gärna av er med frågor och synpunkter!

JOAKIM MALMDIN
joakim.malmdin at scb.se

FOTO: MONICA HOLMBERG

# Webbpaneler

På senare år har det blivit allt vanligare bland svenska undersökningsföretag att ha egna webbpaneler. Ett mer generellt begrepp är accesspanel. En accesspanel används som en sorts ram ur vilken urval görs när någon beställer en undersökning, en webbpanel är som namnet antyder en accesspanel för undersökningar som ska göras via webben. Skillnaden mot det vi normalt kallar urvalsram är att en webbpanel inte är lika med undersökningspopulationen, i det här fallet Sveriges befolkning. Tanken är ändå att webbpanelen ska vara uppbyggd så att den i någon mening representerar hela landets befolkning. Sverige har bra förutsättningar: bland dem som är under 65 år är det mer vanligt att ha tillgång till Internet än att ha fast telefon.

**Webbpanelen kan beskrivas** som ett register över personer som förklarat sig villiga att delta i webbundersökningar. Registret måste minst innehålla en aktuell e-postadress för varje person, och dessutom gärna en uppsättning bakgrundsvariabler, t ex ålder. En viktig fråga är hur dessa personer rekryterats till panelen. Eftersom panelen endast är ett urval ur populationen vore idealet att personerna i panelen rekryteras genom ett sannolikhetsurval ur hela befolkningen. Ett krav på sannolikhetsurval är att alla personer i populationen ska ha en känd sannolikhet som är större än noll att komma med i panelen. Oftast sker dock rekryteringen med metoder som inte kan kallas för sannolikhetsurval, till exempel genom självrekrytering via annonser eller uppmaningar på Internet eller genom att man vänder sig till medlemmar i föreningar.

När någon beställer en undersökning görs i sin tur ett urval ur webbpanelen. Om urvalsramen i sig är ett urval ur populationen, skapad genom ett icke-sannolikhetsurval, så kan det slutliga urvalet för undersökningen inte kallas ett sannolikhetsurval. Då gäller inte de teoretiska grundvalar som en surveystatistiker är van att förlita sig på. Men vad gäller istället? Går det i praktiken att få resultat från en sådan webbpanelundersökning som kan generaliseras till hela populationen?

**Dessa frågor och** en hel del annat togs upp på ett seminarium som anordnades av Surveyföreningen i februari i år. För ett par år sedan tillsatte Surveyföreningen en kommitté med uppgift att närmare utreda frågan hur man ska se på resultat från undersökningar baserade på webbpaneler. När kommittén nu arbetat ett tag var det dags att stämma av vad man hittills kommit fram till och få synpunkter på hur arbetet borde fortsätta. Gösta Forsman, kommitténs ordförande, förklarade att kommitténs avsikt är att föreslå kvalitetsmått för webbpanelundersökningar och ta fram en skrift som kan användas av producenter, köpare, användare av resultaten och helst också i utbildningar av surveystatistiker. Åke Wissing från ISI Wissing

# – vad är det?

gav en översikt över diverse mått som föreslagits i olika sammanhang och som kommittén hittills har tittat på, men man har ännu inte tagit fram någon egen rekommendation. Befintliga kvalitetsmått beskriver främst tillvägagångssätt och dokumenterar t ex urvalsprocessen. Det kan tyckas vara svaga indikatorer på kvalitet, men även undersökningsföretagets förmåga att göra beskrivningarna kan ses som en indikator i sig.

**Webbpaneler i teori** och i praktik behandlades av Jan Wretman från Stockholms universitet respektive Henrik Kronberg från Norstat. Jan ville inte svara på frågan om icke-sannolikhetsurval aldrig kan vara representativa, utan diskuterade på ett generellare sätt vad representativitet och kvalitet kan vara. I surveysammanhang brukar man med representativitet mena att urvalet i någon mening liknar populationen och att resultat kan generaliseras från urvalet till populationen. Men det förekommer olika tolkningar av vad det egentligen betyder. Det kan till exempel vara att population och urval har samma fördelning m a p vissa bakgrundsvariabler, eller att alla intressanta grupper finns representerade i urvalet eller att varje enskild individ i urvalet är en typisk representant för populationen. En undersöknings kvalitet beror dessutom av många saker utöver urvalsmetoden; mättekniska aspekter, bearbetningar och skattningsförfarande bland annat. Det ger upphov till osäkerhet av olika slag. Vid ett sannolikhetsurval kan man skatta den del av osäkerheten som beror på urvalet. Vid ett icke-sannolikhetsurval finns inte alltid motsvarande möjlighet, då får man istället beskriva tillvägagångssättet. Jan berörde även problemet med att bortfallet i många individundersökningar baserade på sannolikhetsurval gör att även dessa i praktiken blir icke-sannolikhetsurval.

Henrik berättade hur arbetet med rekrytering till och skötsel av en webbpanel går till i praktiken. E-postadresser, den initiala och helt avgörande informationen i rekryteringen till en webbpanel, är färskvara. Därför skickar Norstat numera en så kallad profilundersökning till dem som anmält intresse att ingå i panelen redan inom en timme efter att de fått en e-postadress. I profilundersökningen frågar man till exempel om demografiska uppgifter, hushållets utseende, beteenden och innehav av produkter. Information som är viktig när man drar urval till undersökningarna. I likhet med andra webbpanelföretag har man belöningar (incentives), som man ger ”panelisterna” för att delta i enskilda undersökningar. Den allra mest populära belöningen är trisslotter. I samband med profilundersökningen delges paneldeltagaren regler för sitt deltagande i undersökningarna.

**Det är viktigt** för Norstat att behålla folk i panelen så länge det går. Norstat lägger därför mycket arbete på att vårda sin panel, och undersöka dess sammansättning och vad som motiverar folk att delta. Panelen jämförs ständigt med kontrolltotaler som fås från officiell statistik. Sammansättningen av panelen blir ofta skev, det är lättast att rekrytera medelålders kvinnor på landsbygden och svårast att nå unga killar i storstäder. För att kompensera för skevheten stratifieras urvalet. Individer från svagt representerade grupper kan bli utvalda oftare än panelister i grupper där man har stark representation. Henrik beskrev webbpanelen som en amöba: den ändrar sig hela tiden. Faktorer som gör att personer stannar kvar i panelen är till exempel bra belöningssystem, varierande och bra undersökningar, att undersökningarna inte är för långa och att man inte blir utvald till undersökningar alltför ofta. De enskilda paneldeltagarnas insatser följs upp. Man tittar till exempel på hur lång tid de tar på sig att besvara frågorna i en undersökning, konsistens i svaren mellan undersökningar och typ av svar på öppna frågor.

Det finns även fördelar med en webbundersökning via webbpanel jämfört med en traditionell undersökning via telefon- eller besöksintervju: det blir billigare, man har inga intervjuareffekter och folk kan svara när de själva vill.

**Kvalitet i webbpanelundersökningar** är förstås inte enbart en fråga i Sverige. Bengt Larsson, ordförande i branschorganisationen SMIF, presenterade arbete som gjorts internationellt. Det finns etiska koder, forskningsöversikter, förhållningsregler och liknande från stora internationella organisationer, bland andra ESOMAR (European Society for Opinion and Marketing Research) . På SMIF:s webbplats finns tips på frågor som en köpare av marknadsundersökningar ska ställa sig och ESOMAR:s råd till kunden. Det finns även ISO-standarder som ställer krav på surveyundersökningar i allmänhet och accesspaneler i synnerhet. De krav som formulerats internationellt handlar mycket om dokumentation av processen. I den standard som handlar om accesspaneler, ISO 26362, finns bland annat regler för hur man räknar antalet medlemmar i en panel (bara de aktiva) och man får inte använda termen svarsfrekvens utan endast tala om deltagandefrekvens (participation rate) eftersom det i strikt mening inte kan finnas bortfall när det rör sig om accesspaneler.

**Kommittén har valt** att inrikta sig på sannolikhetsurval, men denna begränsning var det flera i publiken som hade invändningar mot. Icke-sannolikhetsurval är så vanligt förekommande att det behövs riktlinjer även för det och det vore olyckligt att begränsa kommitténs arbete. Denna synpunkt lovade Gösta speciellt att kommittén ska överväga i det fortsatta arbetet. Nicklas Källebring från Synovate, som hade till uppgift att sammanfatta synpunkter i den avslutande diskussionen, skickade även med kommittén en del andra synpunkter. Icke-sannolikhetsurval borde dessutom beaktas i en finare uppdelning eftersom det finns så många olika varianter. Man får inte heller komma med för många mått, då blir det praktiska arbetet med att ta fram dessa ogörligt för undersökningsföretagen.

INGEGERD JANSSON

FOTO: JOHANNA NORIN

**Sophia Olofsson.**

Sophia Olofsson vann i hård konkurrens med andra förnämliga uppsatser Surveyföreningens årliga tävling om bästa uppsats på surveyområdet. Här skriver Sophia en sammanfattning av sin D-uppsats i statistik som hon la fram vid Uppsala universitet.

# Effektivare estimation med hjälpinformation?

Uppsatsen baseras på undersökningen ”Lastbilstrafik – Inrikes och utrikes trafik med svenska lastbilar” och undersöker om det är möjligt att öka precisionen i skattningarna genom att använda hjälpinformation. Undersökningen om lastbilstrafik ingår i Sveriges officiella statistik och publiceras av myndigheten Trafikanalys. Producent av statistiken är Statisticon AB. Undersökningen syftar till att ge en bild av godstransporter med tunga svenska lastbilar, som underlag till bland annat transport- och infrastrukturplanering.

Undersökningen om lastbilstrafik publiceras varje kvartal samt i en sammanfattande årsrapport. Rapporterna redovisar antal transporter, antal körda kilometer, lastad godsmängd och transportarbete (produkten av lastad godsmängd och sträcka, mäts i tonkilometer). Undersökningen är EU-reglerad och precisionskrav finns för årsskattningarna av totalt antal körda kilometer, total lastad godsmängd och totalt transportarbete. Det är dessa tre parametrar som ligger i fokus för uppsatsen.

Undersökningspopulationen består av svenskregistrerade lastbilar med en maxlastvikt på minst 3,5 ton och är stratifierad i 57 strata, bland annat efter region, typ av lastbil (karosstyp) och om lastbilen huvudsakligen kör inrikes eller utrikes trafik. Varje kvartal dras ett urval av lastbilsveckor (en lastbil under en vecka) och enkäter skickas ut till lastbilens ägare. Ägaren uppmanas rapportera samtliga körningar under mätveckan. Urvalet kan hanteras som ett stratifierat klusterurval där klustren utgörs av lastbilsveckor. Man kan se det som om ett obundet slumpmässigt urval av kluster dras inom strata, även om urvalsdesignen strikt sett inte uppfyller kraven för detta (se Rosén och Zamani 1993).

Idag används Horvitz-Thompson-estimatorn (HT-estimatorn) vid estimationen. HT- skattningen är en vägd summa av de observerade värdena i urvalet där varje värde ges en vikt som (i detta sammanhang) är inversen av sannolikheten att just den mätveckan valdes. I allmänhet kan man uppfatta vikten i en HT-skattning som att den säger hur många objekt det utvalda objektet representerar inklusive sig själv.

Med en urvalsstorlek om 3000 lastbilsveckor per kvartal uppnås EU:s precisionskrav, dock ligger parametern lastad godsmängd på gränsen av den tillåtna osäkerhetsmarginalen. En ökad precision i skattningarna är dels önskvärd i sig, dels skulle det kunna leda till att man kan uppnå precisionskraven med ett minskat urval – något som i sin tur innebär en minskad uppgiftslämnarbörda.

**Tanken med hjälpinformation** är att man kan förbättra precisionen i skattningen av en parameter genom att använda en annan variabel som är korrelerad med undersökningsvariabeln. Detta kan illustreras genom följande exempel. Vi är intresserade av att skatta total körsträcka. För att göra detta drar vi ett stickprov av lastbilar och mäter körsträckan. Vi har också tillgång till körsträckan för föregående år för alla lastbilar i urvalsramen genom ett register. För stickprovet kan vi både skatta körsträckan för innevarande år och körsträckan för föregående år. Om körsträckan detta år och körsträckan föregående år är korrelerade kan skattningen av total körsträcka föregående år ge en indikation på om det i stickprovet finns en överrepresentation av lastbilar som kör kort (eller långt) genom att den skattade totalen är lägre (eller högre) än den sanna totalen. Om vi använder denna information i skattningarna kan vi öka precisionen. Detta kan sägas vara principen för regressionsestimatorn. Denna kan skrivas som HT- estimatorn plus en justeringsterm där justeringstermen avspeglar dels hur bra HT- estimatorn skattar hjälpvariablernas kända populationstotaler, dels hur undersökningsvariabeln samvarierar med hjälpvariablerna i en regressionsmodell. Eller uttryckt i formler:

$$\hat{t}_{y,REG} = \hat{t}_{y,HT} + \sum_{j=1}^{J} \hat{B}_j (t_{x_j} - \hat{t}_{x_j,HT})$$

där $t_{x_j}$ är den kända populationstotalen för hjälpvariabel $x_j$ och $\hat{B}_j$ är skattade regressionskoefficienter i modellen

$$y_i = \sum_{j=1}^{J} B_j x_{j,i} + \varepsilon_i$$

Tillgänglig hjälpinformation är olika registervariabler, dels från fordonsregistret, dels från SCB:s körsträckedatabas. I Fordonsregistret finns exempelvis uppgifter om maxlastvikt, antal axlar och karosskod. SCB:s körsträckedatabas innehåller uppgifter om senast kända årliga körsträcka för fordon. Vissa av dessa variabler används i stratifieringen, men då i en grövre in-

## Tävlingskommitténs motivering

■ **”Arbetet är väl** genomfört med utförliga och pedagogiska förklaringar av tankegången och alla införda estimatorer. Språket, dispositionen och framställningen håller hög klass.

Studien bygger på en verklig undersökning där författaren utvecklar en skattningsmetod och studerar dess egenskaper. Metoden är en simuleringsstudie där datamaterialet får utgöra hela populationen, där tillgänglig hjälpinformation från populationen används. Genom att resultaten jämförs mot de sanna värdena finns det fog att tro att simuleringarna är rätt genomförda och att slutsatserna håller mer generellt. För att värdera olika metoder har ett nytt rankningssystem konstruerats. Uppsatsen innehåller också en genomtänkt bortfallsanalys med imputering.”

*Sophia Olofsson*, Lastbilstrafik – inrikes och utrikes trafik med svenska lastbilar. – Effektivare estimation med hjälpinformation? D-uppsats, 15 poäng. Uppsatsen i sin helhet finns på http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-152034

delning. Utifrån kända samband, till exempel att vissa fordonstyper ofta körs tungt och långt, och från korrelations- och regressionsanalyser valdes fyra olika kombinationer av hjälpvariabler för varje parameter. Regressionsestimatorerna undersöktes därefter i en simuleringsstudie.

**I simuleringarna skapades** en population av svarsdata från de fyra kvartalen år 2009. Simuleringspopulationen innehöll cirka 8500 lastbilsveckor. Från populationen drogs sedan upprepade stickprov och parametrarna skattades. Eftersom stickprovsstorleken 3000 lastbilsveckor utgjorde en betydande del av simuleringspopulationen användes även tre mindre stickprovsstorlekar (2000, 1500 respektive 1000 lastbilsveckor).

För varje skattad parameter testades fem estimatorer; HT-estimatorn och fyra regressionsestimatorer, detta för vardera av de fyra olika stickprovsstorlekarna. Skillnaden mellan regressionsestimatorerna ligger i deras olika krav på tillgänglighet av hjälpinformation. För att underlätta jämförelsen mellan de olika estimatorerna skapades ett rankningsystem där estimatorerna bedömdes för varje stickprovsstorlek utifrån bland annat förväntningsriktighet och standardavvikelse. HT-estimatorn presterade bra, men det fanns ändå en skillnad mot regressionsestimatorerna som i nästan alla fall presterade bättre. Skillnaden mellan HT-estimatorn och regressionsestimatorerna som helhet var större än skillnaderna mellan regressionsestimatorerna.

Slutligen upprepades skattningarna i kvartals- och årsrapporterna med riktiga data. I en jämförelse av de relativa osäkerhetsmarginalerna mellan HT-estimatorn och regressionsestimatorn syntes en förbättring i precision för alla tre parametrar, men för lastad godsmängd var förbättringen liten.

**Slutsatsen av uppsatsen** blev att man kan uppnå ökad precision i estimationen med hjälpinformation, men att det inte går att minska urvalet enbart baserat på resultaten i denna uppsats. I ett fortsatt arbete skulle man kunna undersöka om det finns andra (kombinationer av) hjälpvariabler som kan användas i skattningarna för lastad godsmängd.

SOPHIA OLOFSSON, STATISTICON AB

**Referens**

Rosén, B. och Zamani, M (1993). Översyn av Undersökningen av lastbilstransporter i Sverige (UVAV), SCB R&D Report 1993:2. Statistiska centralbyrån.

# Folk-räkningar

Den röda tråden för förmiddagspasset vid Surveyföreningens årskonferens var folkräkningar, som i de tre föredragen belystes ur olika perspektiv.

**Anders Ahlbom vid** Institutet för miljömedicin på KI använder folkräkningsdata i sin forskning. Han tog bland annat upp en studie av hjärtinfarkter i Stockholms län där han använt socioekonomiska data från en folk- och bostadsräkning (FoB). Vidare nämnde han en studie av tandvårdspersonal och förekomsten av en sällsynt cancerform där man använt uppgifter om yrke från FoB. Anders avslutade med att ta två exempel från sportens värld. I den första kunde man se att golfare i genomsnitt lever längre än sina icke-golfande medmänniskor. Detta var en effekt som visade sig gälla oavsett socioekonomisk tillhörighet. I det sista exemplet visades samma effekt gälla även vasaloppsåkare även om tiden närmast före, under och efter själva loppet möjligen utgör en viss riskfaktor.

FOTO: LINNEA HEDLIN

**Mugg med samlarvärde?**

**Hur genomförs då** folkräkningar? Claes Andersson från SCB redogjorde för arbetet med den kommande svenska undersökningen, den första sedan 1990, och som i egentlig mening inte är en folkräkning utan snarare en hushålls- och bostadsräkning (HoB). För första gången kommer undersökningen att helt baseras på registerdata, något som möjliggjorts genom att Riksdagen efter viss politisk vånda beslutat att ett lägenhetsregister ska upprättas. Genom att samtliga personer via Skatteverkets försorg folkbokförs på lägenhet kan man registervägen föra samman personer till hushåll, något som tidigare inte varit möjligt för boende i flerbostadshus. Claes avslutade sitt anförande med att lyfta fram styrkor respektive svagheter med en registerbaserad ansats jämfört med en traditionell folk- och bostadsräkning. Bland styrkorna lyfte han fram möjligheten att ta fram årlig detaljerad statistik medan det till svagheterna hör att man blir hänvisad till folkbokföringsadressen, vilken kanske inte alltid är densamma som den faktiska bostadsadressen.

**»Golfare lever i genomsnitt längre än sina icke-golfande medmänniskor»**

**Mats Berggren från** Riksarkivet avslutade förmiddagspasset med att redogöra för Riksarkivets arbete med att digitalisera historiska folkräkningar. Det långsiktiga målet är att digitalisera samtliga folkräkningar från 1860. De efter 1950 och vissa tidigare folkräkningar finns redan i digital form. De finns tillgängliga via Riksarkivets webbplats. Vidare berättade Mats om en databas, upprättad av North Atlantic Population Project, NAPP, under ledning av Minnesota Population Center vid University of Minnesota i Minneapolis, där historiska befolkningsdata från sex länder samlats. För att kunna göra jämförande analyser mellan länderna har NAPP-projektet harmoniserat datastruktur, kodning och dokumentation för de olika ländernas folkräkningar. Totalt innehåller databasen information om 93 miljoner personer och 19 miljoner hushåll!

Samtliga föredragshållare tackades av ordförande Daniel Thorburn och fick för sina insatser en exklusiv mugg försedd med Surveyföreningens emblem.

STEFAN BERG, STATISTISKA CENTRALBYRÅN

## NYA I CRAMÉR-SÄLLSKAPETS STYRELSE

### Martin Sköld, sekreterare

– Jag är sedan i höstas lektor i matematisk statistik vid Stockholms universitet. På vägen hit har jag bland annat passerat universiteten i Lund, Lancaster, Canberra och Örebro som doktorand, postdoc, och nu senast som lektor i statistik. I min doktorsavhandling studerade jag teoretiska egenskaper hos icke-parametriska kurvskattningar. På senare år har jag även forskat om Monte-Carlometoder, bayesiansk statistik och tillämpningar inom biovetenskaperna, speciellt ekologi. Med två barn under två går fritiden mest åt till familjen och till att drömma om att få lite mer tid i löparspåret.

### Jakob Bergman, ledamot

– Jag arbetar sedan våren 2010 vid Statistiska institutionen vid Lunds universitet. Utöver undervisning och forskning ägnar jag mycket tid åt den flod av utvärderingar och granskningar som sköljer över universitetsvärlden för närvarande. Jag disputerade våren 2010 på en avhandling som behandlar hur man kan beskriva korrelation mellan vektorer av andelar (s.k. sammansättningar eller kompositioner). Min forskning idag handlar till stor del om olika aspekter av tidsserier av vektorer av andelar.

Anton Grafström – årets Cramérpristagare.

## ORDFÖRANDEN HAR ORDET

# Grattis Anton Grafström!

Temat för Cramérsällskapets årsmöteskonferens i Örebro var forskarutbildningarna i statistik och matematisk statistik, och diskussionsinledare var undertecknad, Tom Britton och Anders Grimvall. Något som kom fram var att det finns en tendens, särskilt inom matematisk statistik, att avhandlingarna får en mer och mer tillämpad karaktär. Därför skiljer sig avhandlingarna i matematisk statistik och statistik ofta inte så mycket åt med avseende på teori längre, utan mer ifråga om inriktning på tillämpningar (natur- eller samhällsvetenskapliga). Men ett rimligt minimikrav kan vara att minst en av avhandlingens artiklar ska vara publicerbar i en statistisk/sannolikhetsteoretisk tidskrift.

Anders Grimvall underströk behovet av att modernisera forskarutbildningarna, till förmån för analys av stora datamaterial och eventuellt på bekostnad av mer ”traditionell” inferensteori. Här nådde vi nog ingen konsensus på mötet, men jag kan ändå konstatera att vi har tagit intryck av dessa strömningar när det gäller vår sommarskola i juni, se vidare nedan.

**Så avslöjades vem** som var årets Cramérpristagare. Det var Anton Grafström från matematisk statistik i Umeå, som har skrivit en avhandling med titeln ”Unequal Probability Sampling Designs”. Anton presenterade på ett mycket förtjänstfullt och pedagogiskt sätt sin avhandling, speciellt en artikel där han föreslår entropi som ett kvalitetsmått på en urvalsdesign.

**Sedan avslutades mötet** som sig bör med årsmötesförhandlingar. Som så ofta förr innehöll dessa en diskussion om ekonomi, och vi kom bland annat fram till att en framtida höjning av medlemsavgiften är nödvändig. Men vi ska naturligtvis också försöka att bli bättre på att rekrytera nya medlemmar. Medlemsavgifterna är för övrigt i nuläget 50 kr per år för enskilda medlemmar och 500 kr för institutionsmedlemmar. Vidare valdes en ny styrelse bestående av undertecknad som ordförande, Martin Sköld som sekreterare, Mikael Möller som kassör, Jan-Eric Englund som representant i Svenska statistikfrämjandets styrelse och Jakob Bergman, Ann-Marie Flygare och Aila Särkää som ledamöter.

**I dagarna arrangerar** vi vår första sommarskola, vilket är en direkt fortsättning på vad Svenska statistikersamfundet tidigare gjorde. Senast (2007) var ämnet data mining och platsen Linköping. Den här gången har vi ett snarlikt ämne, statistical learning, och sommarskolan äger just nu rum på Asa herrgård utanför Växjö. Inbjudna talare är John Shawe-Taylor, Mattias Villani och Jennifer Castle, och utöver deras föredrag kommer ett par stycken deltagarbidrag att presenteras. Till sist vill jag tacka de vid årsmötet avgående styrelsemedlemmarna Ulla Blomqvist, Jessica Franzén och Erik Lindström för sitt engagemang och förtjänstfulla arbete under deras tid i styrelsen.

ROLF LARSSON

## NY I CRAMÉRSÄLLSKAPETS STYRELSE

### Aila Särkkä, ledamot

– Jag är född i Pirkkala (nära Tammerfors) i Finland. Jag disputerade i statistik vid Jyväskyläs universitet 1994. Efter ett och ett halvt år i Avignon i Frankrike som postdoc flyttade jag till Göteborg. Först hade jag en forskarassistenttjänst på Statistiska institutionen på Handelshögskolan, och sedan 1998 har jag jobbat på Matematiska vetenskaper på Chalmers och Göteborgs universitet, först som lektor och numera som docent.

– Mitt forskningsområde är spatial och spatio-temporal modellering. Speciellt är jag intresserad av spatiala punktprocesser, där data är i form av positioner av punkter (t.ex. träd i skogen) och märken som tillhör till punkterna (t.ex. diametern av trädet). Mina viktigaste tillämpningsområden är skogsbruk, materialvetenskap och neurologi.

## ORDFÖRANDEN HAR ORDET

# Den som väntar på något gott ...

Nej, nej, nej. Vad har hänt? Vi tittar förskräckt på varandra. Det är lördagskväll när vi upptäcker att världens utan överdrift kraftigaste spik har borrat sig in i ett av däcken. Den ser ut att ha suttit där ett bra tag men luften har ännu inte pyst ut helt. Tävlingssäsongen är igång på allvar. Ibland är det lite tufft att vara ”ponny-mamma”. Reservdäckets skick verkar inte vara mycket bättre. Jag ringer transportägaren som tack och lov har ett annat däck som vi kan byta till. Sagt och gjort – min händige man får rycka in och strax innan mörkret börjar lägga sig lyckas han få loss det trasiga däcket och dra fast det andra. Några timmar senare är det dags att stiga upp (strax före kl 6), äta frukost, fylla en termos med starkt kaffe och hämta hästen i stallet. Nästa utmaning är att övertyga hästen om att det är en bra idé att gå in i transporten och när det väl lyckats är det dags för mig, tävlingsryttaren (yngsta dottern) och coachen (äldsta dottern) att köra iväg en fullastad bil med utrustning, hästtillbehör (checklistan är lång) och ett släp som väger över ett ton. Ungefär så här ser en helt vanlig helg ut för en hästägarfamilj med måttliga tävlingsambitioner i ponnyallsvenskans division III. Jag hittar numera rätt bra på de smala, slingriga vägarna till de flesta av ridanläggningarna på den vackra sörmländska landsbygden.

**Nu till vår** förening FMS, som inte råkat ut för några punkteringar vad jag vet. Tack för ett mycket trevligt vår- och årsmöte på Örebro universitet i mars! Föreläsarna gjorde ett kanonjobb, deltagarna var aktiva och allt det praktiska fungerade väl. På förmiddagen fick vi höra Sharon Kühlmann Berenzon från Smittskyddsinstitutet berätta om en spatial epidemiologisk studie av sporadiska fall av VTEC (Verocytotoxin-producing Escherichia coli). Målet med studien var att identifiera geografiska områden med hög incidens och testa om avstånd till djurfarm är associerat med risk för infektion. Caroline Dietrich från MEB på Karolinska Institutet föredrog om blod från blodgivare och flexibel parametrisk modellering av lagringstidens eventuella påverkan på överlevnaden hos patienter som genomgått en blodtransfusion. Min kollega Pär Karlsson på AstraZeneca, tillika ledamot i FMS:s styrelse, höll ett väldigt givande föredrag om en ”enkel” modellering av läkemedelsutvecklingsprogram. Pär ville inspirera oss statistiker till att medverka till en förbättrad utvecklingsprocess, vilket han lyckades väl med.

**»Jag tackar för det förnyade förtroendet.»**

**På eftermiddagen fick** vi besök från SCB i Örebro och Örebro universitet. Presentationerna om nationella hälsoräkenskaper och användning av SCB:s register i medicinsk forskning var väldigt allmänbildande och återanvändning av gamla personnummer var ett av de intressanta registerproblem som diskuterades. Kaisa Ahola, Marie Glanzelius, Christina Liwendahl och Peter Abrahamsson tog oss galant igenom dessa spännande områden. Lennart Bodin från Örebro universitet talade om tillväxtkurvor för barns längd och vikt, ett område som många av oss som är föräldrar har en särskild relation till. Maria Grünewald, Smittskyddsinstitutet, nyligen disputerad vid Stockholms universitet, diskuterade effektiva urval i teori och praktik. Dagen var som ni förstår mycket innehållsrik och med en unik ämnesbredd inom medicin och hälsa.

**Jag tackar också** för det förnyade förtroendet. Trevligt att få vara FMS ordförande ännu ett år. Marita Olsson och Peter Wessman har nu lämnat styrelsen och avtackats efter flera aktiva år. De ersätts av Linda Hartman och Emil Rhenberg. Åsa Vernby (sekreterare), Hanna Svensson (kassör), Pär Karlsson, Mats Rudholm samt EFSPI-representanterna Carl-Fredrik Burman och Marie Göthberg omvaldes.

**Hur länge tål den att lagras?**

**När ni läser** dessa ord har endagskonferensen ”Model-Based Drug Development” på KTH nyligen ägt rum där FMS var en av arrangörerna. Statistikerträffen i Göteborg är en återkommande aktivitet att se fram emot efter sommaren. Vi har även i år utlyst stipendium och stipendiaterna ska utses nu i juni. Har ni medlemmar något förslag på tema eller område till kommande möten eller kurser? I så fall skicka

FOTO: SIEMENS PRESS PICTURE

gärna idéer till fmsstyrelse@gmail.com. Ni som har glömt betala medlemsavgiften de senaste åren, önskar anmäla ändrad mejladress eller önskar bli ny medlem, kontakta oss. Vi arbetar på att föra över vår webbplats från Samfundets gamla till Främjandets nya. Den som väntar på något gott – ja ni vet.

Ha nu en riktigt skön sommar!

ANNA TORRÅNG

## NYA I FMS STYRELSE

### Emil Rehnberg, ledamot

– Jag arbetar som biostatistiker på Karolinska Institutet vid Institutionen för medicinsk epidemiologi och biostatistik (MEB). Under tiden som jag studerade till en magister i biostatistik påbörjade jag mitt examensarbete på MEB. Det handlade om imputering av genotyper, och efter det har jag fortsatt att arbeta här på MEB, nu i ett och ett halvt år. Jag är involverad i flera forskningsprojekt med doktorander och forskare. Mestadels är projekten inriktade på genetiska exponeringar men jag arbetar även med en del registerbaserade studier.

Efter arbetet pluggar jag japanska och ägnar mig åt klättring (inom grenen bouldering). Än så länge klättrar jag enbart inomhus men snart även utomhus i Stockholm med omnejd.

### Linda Hartman, ledamot

– Jag arbetar sedan i höstas som statistiker på Cancerregistret och Onkologens forskningsavdelning vid Lunds universitet. Tack vare en delad tjänst får jag arbeta med perspektiv från preklinik ända till epidemiologi.

– Jag läste teknisk fysik i Lund följt av doktorandstudier i matematisk statistik. Som doktorand tog jag steget från teknik mot biostatistik, och disputerade 2007 på en avhandling med främsta fokus på algoritmer för kartläggning av sjukdomsgener. Några av er kanske var med då jag på FMS höstmöte 2006 presenterade denna forskning. Utöver forskning lade jag under doktorandtiden stort fokus och engagemang på undervisning och kursutvecklingsprojekt.

– Efter disputationen arbetade jag med utveckling av läkemedel mot astma och KOL (”rökarsjuka”) på AstraZeneca R&D i Lund, främst med fas IV-studier och publikationsarbete.

– Jag bor i Lund, med man och tre döttrar. Jag är friluftsmänniska, som de senaste åren fått byta veckoturerna över vita fjällvidder mot mer barnanpassade turer, ofta som mullefröken. Annars ägnar jag gärna min lediga tid åt kolonilott, god mat och umgänge. Och en god bok på huvudkudden.

# Lagen om π

## OM EN AMERIKANSK DELSTAT SOM STIFTADE EN LAG SOM SADE ATT π = 4

**FAKTOID**

En uppgift som ser ut som ett faktum men som inte är det; en känd (hos allmänheten eller i någon specifik grupp) men sällan ifrågasatt osanning, halvsanning och/eller missuppfattning.

Det rör sig verkligen om ett historiskt faktum, eller åtminstone bra nära. Det lades fram som ett lagförslag (på fullt allvar?) som gick genom byråkratin utan att någon reagerade nämnvärt; det var inte förrän förslaget skulle krönas till stiftad lag som det fick det bemötande det förtjänade.

Delstaten var Indiana, året 1897. En viss Edwin J. Goodwin (Edward i en källa, Goodman i en) lade fram ett lagförslag där ett antal nya värden på π definierades, att fritt användas i delstaten Indiana för utbildningsändamål. (Det har påpekats att resten av världen skulle få betala royalty, vilket måhända var det verkliga målet med det hela.)

Texten är såpass tilltrasslad att det inte är självklart vilka värdena är, eller ens deras antal; en matematiker kom fram till sex olika värden, en annan källa anger fyra, och det är även oklart huruvida just värdet 4 finns med. (Man passar även på att definiera värdet på roten ur 2 till 10/7).

**Matematiken eller ens** logiken i det hela var tydligen inget som stack i ögonen på ledamöterna i någon av de olika kommittéer som förslaget, *House Bill No. 246*, passerade. Inte heller verkar tidningarna ha insett exakt hur absurt det hela var. Då kommer en viss C. A. Waldo, professor i matematik, in i berättelsen - hans insats beskrivs en smula olika i källorna, men det mest troliga verkar vara att han fick syn på förslaget i generalförsamlingen, och lyckades varna senaten (i Indiana, inte i Washington DC) om vad de hade att vänta sig.

**När förslaget så** kom till senaten så blev reaktionen den följande:

> *... The bill was brought up and made fun of. The Senators made bad puns about it, ridiculed it and laughed over it. The fun lasted half an hour. Senator Hubbell said that it was not meet for the Senate, which was costing the State $250 a day, to waste its time in such frivolity. He said that in reading the leading newspapers of Chicago and the East, he found that the Indiana State Legislature had laid itself open to ridicule by the action already taken on the bill. He thought consideration of such a proposition was not dignified or worthy of the Senate. He moved the indefinite postponement of the bill, and the motion carried.*
>
> Indianapolis News, 13 februari 1897

## Läs mer

■ Från Peter Olaussons Faktoider, *www.faktoider.nu*, där det även finns referenser, inklusive en länk till själva lagförslaget, återgivet in extenso. ”Lagen om π” finns under rubriken Sant! eftersom berättelsen om lagförslaget faktiskt stämmer. Denna berättelse finns också i Peter Olaussons bok ”Fler faktoider” (Forum 2009).

# Få full kontroll över din statistiska undersökning

*Från datainsamling till rapport – att göra en statistisk undersökning* beskriver och diskuterar planering och konstruktion av frågeformulär, inklusive hur svarsprocessen går till när enskilda personer respektive företag svarar. Boken tar även upp kodning, olika datainsamlingsmetoder, val av urvalsmetod, hur fel i undersökningar kan undvikas samt hur rapporten kan struktureras och utformas.

Denna femte upplaga är en kraftigt utökad och aktualiserad genomgång av de olika stegen i en statistisk undersökning. Begreppet statistisk undersökning är nu utvidgat till att även omfatta grunderna i experiment, kvasiexperiment och epidemiologi samt vissa kvalitativa undersökningar. Övningsuppgifterna är till största delen aktualiserade och i kapitlet om urvalsmetoder har ett stort antal nya uppgifter tillkommit.

**Du beställer enkelt *Från datainsamling till rapport – att göra en statistisk undersökning* via www.studentlitteratur.se/3423 och där finner du också mer information om boken.**


Studentlitteratur
VINNANDE VETANDE

Qvintensen

POSTTIDNING B

SVERIGE
PORTO BETALT
PORT PAYÉ

Avsändare/Sender: Svenska statistikfrämjandet,
c/o Chandra Adolfsson, Statistiska centralbyrån
701 89 Örebro, Sweden

# KALENDARIUM

## DETTA HÄNDER 2011

| | KONFERENS | PLATS | WEBBPLATS |
|---|---|---|---|
| 4-8 juli | PROBASTAT 2011 | Smolenice Castle, Slovakien | www.um.sav.sk/en/probastat2011.html |
| 6-8 juli | 16th INFORMS Applied Probability Society Conference | Stockholm | http://meetings.informs.org/APS2011 |
| 6-8 juli | The 2011 International Conference of Computational Statistics and Data Engineering | London, England | www.iaeng.org/WCE2011/ICCSDE2011.html |
| 30 juli – 4 aug | 2011 Joint Statistical Meetings | Miami Beach, USA | www.amstat.org/meetings/ |
| 8-10 aug | ECAS Small Area in Statistics: Theory and Practice | Trier, Tyskland | www.uni-trier.de/index.php?id=33500 |
| 11-13 aug | Small Area Estimation | Trier, Tyskland | www.uni-trier.de/index.php?id=30790 |
| 15-17 aug | SCB:s och Örebro universitets sommarskola, Small Area Estimation | Örebro, Sverige | www.oru.se/hh/summerschool2011 |
| 21-26 aug | International Statistical Institute, 58th ISI World Statistics Congress | Dublin, Irland | www.isi2011.ie/ |
| 26-29 aug | 65th European meeting of the Econometric Society | Oslo, Norge | eea-esem2011oslo.org/ |
| 4-8 sep | ENBIS 11 | Coimbra, Portugal | www.enbis.org |
| 5-9 sep | 17th European Young Statisticians Meeting | Lissabon, Portugal | www.fct.unl.pt/17eysm/ |
| 12-14 sep | 2nd European Establishment Statistics Workshop | Neuchâtel, Schweitz | www.enbes.org |
| 5-7 okt | 9th International Conference on Health Policy Statistics | Cleveland, USA | www.amstat.org/meetings/ichps/2011/index.cfm |
| 1-4 nov | Strategies for Standardization of Methods and Tools | Ottawa, Kanada | www.statcan.gc.ca/conferences/symposium2011/ |
| 28-31 dec | International Conference on Advances in Probability and Statistics | Hong Kong, Kina | http://faculty.smu.edu/ngh/icaps2011.html |

## ...OCH LITE LÄNGRE FRAM

| | | | |
|---|---|---|---|
| 10-14 juni 2012 | Nordic Conference on Mathematical Statistics | Umeå | www.nordstat2012.se |