

Qvintensen

SVENSKA STATISTIKFRÄMJANDET

NR 3 2014

Att lära ut statistik med vinprovning

SIDORNA 20-22

DISKUSSION/

Om farmödrar har genomgått en drastisk förändring i tillgången på mat kan det påverka risken för hjärt-kärlsjukdomar i *tredje* generationen. Hur övertygande är evidensen för det?

SIDORNA 15-18

Det finns ofta värde i det som inte hörs eller syns.»

Anna Torráng, sidan 19

Spelar det roll vilken opinionsundersökning man tar del av?

SIDORNA 8-9



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 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

Girig industri.



Yuli Liang på väg.



Intressant pippi?



Jakob Bergman.



Anna Torrång.

Qvintensen

SVENSKA STATISTIKFRÄMJANDET

NR 3 2014

Innehåll

PÅ VÄG ATT LÄRA SIG STATISTIK..... 6-7

OPINIONSUNDERSÖKNINGAR

– spelar det någon roll vilken man tar del av? 8-9

HÖGSKOLEUTFORSKAREN UNDER-

LÄTTAR FÖR LÄROSÄTENA..... 10-11

LIVET SOM KONSULT 19

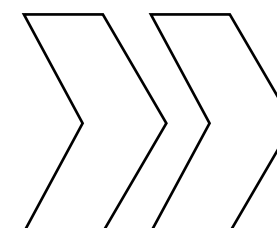
PRIS FÖR BÄSTA PRESENTATION 29

ORSAK TILL ETT VANLIGT FELTÄNK 30

DISKUSSION/

Kursen som jag nätt och jämnt fick godkänt på 12-14

En studie i multipla test med repliker..... 15-18



Uttalandet ger ändå en bild av hur statistik kan uppfattas ... fakta som man kan slå varandra i huvudet med ..., men som inte leder till någon konstruktiv utveckling.»

Hans Nyquist, sidorna 12-14



Våra föreningar

INDUSTRIELL STATISTIK

Statistisk vinprovning..... 20-22

CRAMÉRSÄLLSKAPET

Att lära sig visualisering och presentation.....22

Vad är vitsen med sannolikhetslära?23

SURVEYFÖRENINGEN

Ständigt dessa siffror!24

Att klassificera ord med siffror..... 25-27

FÖRENINGEN FÖR MEDICINSK STATISTIK

Läkemedelsindustrin och de sju dödssynderna 28-29

Fasta spalter

I HUVUDET PÅ EN REDAKTÖR4

ORDFÖRANDEN HAR ORDET5

HÄNT & HÖRT29



I HUVUDET PÅ EN REDAKTÖR

Qvintensen

Ansvarig utgivare
John Öhrvik
Redaktör
Dan Hedlin, 070 297 10 27
Ingegerd Jansson (biträdande)

Redaktion
Marie Linder, FMS
Jonas Ortman, Surveyföreningen
Maria Blixt, Industriell statistik
Lars Rönnegård, Cramérskallskapet
Hans Ekholm
Alf Fyhrlund
Anna Torrång

Adress
Qvintensen, c/o Dan Hedlin
Dammtorps allé 6, 170 62 Solna
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 6 000 kronor för institutionsmedlem eller motsvarande, och 8 500 kronor för övriga annonsörer. På alla priser tillkommer 25 % moms.



SVENSKA STATISTIKFRÄMJANDETS STYRELSE

Ordförande
John Öhrvik, 070-666 08 08,
John.Ohrvik@ki.se
Vice ordförande
Inger Eklund, 08-506 947 13
Sekreterare
Thorbjörn Gudmundsson
Klubbmästare Anders Hallberg
Skattmästare Jonas Selling
Surveyföreningen
Maria Öman
FMS
Frank Miller, 08-16 29 76
Industriell statistik
Jenny Jonasson
Cramérskallskapet
Maria Karlsson, 090-786 55 96
Övriga ledamöter
Marie Wiberg, Stig Johan Wiklund

Adress
Svenska statistikfrämjandet
c/o: T Gudmundsson
Institutionen för matematik
KTH, 100 44 Stockholm
E-post sekfram@gmail.com
Webbplats www.statistikframjandet.se



Möten med människor

Det bästa med statistisk är att man får leka på alla andras bakgårdar. Det sa den teoretiska matematikern John Tukey som skolade om sig till statistiker eftersom han tyckte att de statistiska problemen var intressantare. Jobbet som statistiker innebär möten med människor.

Qvintensen handlar den här gången om möten med människor. Det finns ofta värde i det som inte hörs eller sägs, skriver Anna Torrång. Man ska vara frågvis. Det gäller både i mötet med människor och i mötet med data. En del av Tukeys mångfacetterade insatser gäller konsten att ställa frågor till data och få dem att svara. I det här numret skriver Hans Nyquist, Göran Lande och Kenneth Carling om att levandegöra undervisning i statistik när man möter människor som i förstone inte är lika intresserade som man själv. Ett sätt att väcka intresset är att ta upp sådant som deltagarna kan relatera till.

Ett annat sätt är att ta upp tankegator. Ta fotbollstränaren som har för vana att skälla ut dem som inte lyckats bra med de moment på träningen som han tycker att de borde behärska. Han menar bestämt att beröm inte fungerar, däremot visar erfarenheten att utskällningar gör det. De som han berömmar brukar lyckas sämre nästa gång, men dem han kritiserar lyckas nästan alltid bättre.

Det här feltänket har fått ett namn: regression mot medelvärde – läs artikel på sidan 30.

För många är statistiska metoder en verk-

»Man ska vara frågvis. Det gäller både i mötet med människor och i mötet med data.»

tygslåda, inte ett sätt att tänka. Det är som att dra i en skruv. Om man använder rätt skruv blir resultatet gott, även om man inte är lika händig som en snickare. För att välja rätt skruv behöver man kunskaper. Resten gör man själv. Det är en mycket dålig analogi med statistiskt arbete.

Feltänket ovan visar varför analogin är så dålig. Erfarenheten visar ju att beröm inte fungerar, men att utskällningar gör det. Eller? Analogin med skruven blir bättre om det inte för ett otränat öga går att se om skruven kommer att lossna så fort man vänder ryggen till.

En läkares arbete ligger närmare: att lyssna, försöka avgöra vad det är som inte syns, och hitta någon behandling som kan förbättra livet för patienten.

DAN HEDLIN



FOTO: LIESELOTTE VAN DER MEIJ

ISSN 2000-1819

Författare och andra som utan ersättning bidrar med material för publicering behåller upphovsrätten till sina verk. Svenska statistikfrämjandet äger rätten att publicera dessa verk i valfri form och upplaga samt att behålla alla eventuella intäkter från sådan publicering, om inget annat avtalas. Författare och andra som utan ersättning bidrar med material för publicering garanteras rätten att själva få sprida sitt verk under förutsättning att Qvintensen anges som originalkälla med angivande av nummer och år. Svenska statistikfrämjandet äger upphovsrätten till publicerat material då det är framställt av redaktionsmedlem, styrelsemedlem, någon till Främjandet knuten förenings styrelse eller person som erhållit arvode för sitt bidrag. © Svenska statistikfrämjandet 2014.

ORDFÖRANDEN HAR ORDET

Den svåra tolkningen av statistik

Efter en sommar som startade kyligt, vi minns alla midsommar, kom äntligen juli med värme. Nu har en ny termins arbete startat efter en skön och vederkvickande semester. För Svenska statistikfrämjandets del inleddes hösten med ett medlemmöte i Stockholm. På mötet fick vi höra om de statistiska metoderna bakom de olika måtten som tas fram i SCB:s partisympatiundersökning, och hur förutsättningarna för dessa har förändrats över åren.

Jag skriver detta några dagar inför valet den 14 september. Vi översköljs av undersökningar av väljarsympatier från de stora opinionsinstitut. Inte sällan med disparata resultat – lördagen den 6 september presenterade Sifo en undersökning som gav de rödgröna 45,0 % och alliansen 40,5 % och Novus en som gav de rödgröna 48,0 % och alliansen 38,1 %, det vill säga en skillnad mellan blocken på 4,5 respektive 9,9 procentenheter. I sanning en remarkabel skillnad då undersökningarna innehöll cirka 1 250 respektive drygt 1 600 intervjuer; vad ska man tro? Detta reser krav på en tydligare redovisning av de metoder som opinionsinstitut använder och den kvalitet deras undersökningar håller, inte minst vad gäller bortfallet.

I förra numret av Qvintensen beskrevs en metod utvecklad av Jakob Bergman och Björn Holmquist för att väga ihop olika opinionsundersökningar till ett sammanfattande resultat. I det här numret utvecklar Jakob Bergman detta. Behovet av tillförlitliga metoder för sammanvägning av olika mätningar visades tydligt av väljarbarometrarna inför valet.

Ett citat som tankarna sökt söka sig till i dessa valdebattstider



»Detta reser krav på en tydligare redovisning av de metoder som opinionsinstitut använder.»

av effekten i de starkaste fynden som är känd under namnet "winner's curse" eller "regression towards the mean". Begreppet "winner's curse" har speciellt kommit till användning inom helgenom och epidemiologiska studier och syftar på att den förste som gör ett signifikant fynd oftast redovisar en betydligt starkare effekt än vad som senare kommer att upptäckas i liknande replikationsstudier.

Med förhoppning om en skön brittsommar.

John Öhrvik, ordförande



FRÅN TIANJIN TILL STOCKHOLM

På väg att lära sig statistik

Höstterminen 2014 är elfte året jag läser statistik. Det är ett ämne som jag har ett oerhört stort intresse för. Här ser jag tillbaka på, och sammanfattar, min läroprocess.

Höstterminen 2003 började jag studera på kandidatprogrammet i statistik i en av Kinas största städer, Tianjin, som ligger vid kusten i nordöstra Kina. På mitt universitet, Tianjin University of Finance and Economics, krävdes många matematiska övningar för att lägga en bra grund, särskilt i matematik. Jag minns att på kursen i matematisk analys arbetade vi med fyra övningsböcker skrivna av den välkände matematikern Demidovich [1], och vi löste minst 130 olika problem med integration. Jag har deltagit i många kurser i grundläggande statistik, linjär algebra, ekonometri, sannolikhetsteori o.s.v. De fyra årens utbildning var fulla av både svett och glädje. Alla dessa matematiska studier har jag haft stor nytta av senare, i master- och doktorandutbildningen.

2007 var vändpunkten i mitt liv. Jag flyttade till Sverige och fortsatte med masterprogrammet i statistik vid både Högskolan Dalarna och Uppsala universitet. Det var en helt ny miljö för mig, både med avseende på studierna och livsstilen. Kinesisk undervisning lägger mer vikt på kunskapsförmedling, vilket innebär att på föreläsningarna försöker läraren förmedla så mycket kunskap som möjligt till studenterna.

Men i många kurser i Sverige är föreläsningarna i form av seminarier och läraren uppmuntrar oftare studenterna att studera själva och kommunicera och diskutera med varandra. I Sverige har min förmåga till självständigt tänkande och presentation förbättrats. Samtidigt har min matlagning blivit mycket bättre eftersom på svenska universitet brukar studenterna laga mat själva istället för att lita mycket till universitetets matsal, som man gör på Kinas universitet.

Förutom många kurser på avancerad nivå, till exempel statistisk inferensteori, generaliserade linjära modeller, multivariat analys, programmering i R, analys av upprepade mätningar och strukturella ekvationsmodeller, har jag skrivit två uppsatser. Efter det började jag få lite känsla för statistik som ett språk för att beskriva osäkerhet.

»I Sverige har min förmåga till självständigt tänkande och presentation förbättrats.»

Den första uppsatsen var en statistisk modellering av unga skidåkares framsteg. Vi använde en typ av icke-linjära blandade modeller för att undersöka skillnader i pubertetens effekter mellan pojkar och flickor. Jag och mina tre medförfattare har fått artikeln publicerad i tidskriften Chance [2].

Den andra uppsatsen var en statistisk analys av neurokirurgi vid en ryggkirurgisk klinik i Sverige. Vi använde logistisk regression för att undersöka om patienterna hade hög blodning under specifik ryggkirurgi och hur det samvarierade med olika variabler. Jag besökte ett sjukhus och tittade på riktigt kirurgi för att förstå olika

FOTO TJEIWAŠAN: STEVEN HALE/WIKIMEDIA



faktorer bättre. Problem från en tillämpning kan vara intressant även för en statistiker.

Under de sex åren fick jag inte bara bred kunskap i statistik och utvecklade min analytiska förmåga, jag bestämde mig också för att fortsätta forska i statistik. 2009 blev jag doktorand vid Statistiska institutionen vid Stockholms universitet. Forskning är hjärtat i doktorandutbildningen och det viktigaste är att lära sig ett tänkesätt för att svara på nya forskningsfrågor, vilket är utmanande. Mitt avhandlingsämne är flernivåmodeller under antagande av blocksymmetriska kovariansstrukturer. Flernivåmodeller tillhandahåller möjligheter att analysera effekter av bakgrundsvariabler som ligger på olika nivåer såsom elev- respektive skolnivå eller patient- respektive sjukhusnivå [3]. När data är hierarkiska blir beroenden ofta mer komplicerade. Avhandlingen behandlar problemet med att skatta och testa en speciell blockkovariansmatris där blocken är utbytbara men där det inom block



Allt vad en statistiker kan behöva lära sig.

föreligger rotationsinvarians [4] när observationerna antas följa en normalfördelning.

Samtidigt är jag också intresserad av ett annat område inom statistik, överlevnadsanalys, vilket fokuserar på hur exponering för olika faktorer kan påverka t.ex. senare sjukdomsrisk. Jag har skrivit en uppsats som använder bayesianska beräkningar av överlevnadsfunktioner för att undersöka hur sociala faktorer kan påverka människors civilstånd i Sverige.

Livet som doktorand har gett mig en perfekt kombination av forskning och undervisning som är bästa sättet att lära sig ett ämne. Jag ska disputera i början av nästa år men det är absolut inte slutet på mitt statistiska lärande. Jag ser fram emot att arbeta med mer tillämpade områden och utveckla min statistiska förmåga i framtiden. Jag fortsätter alltid att lära mig mer om statistik.

YULI LIANG
STOCKHOLMS UNIVERSITET

Referenser

- [1] Boris Pavlovich Demidovich (Novogradok, 2 mars 1906 - Moskva, 23 april 1977) var en sovjetisk/vitrysk matematiker.
- [2] Alam, M., Carling, K., Chen, R., Liang, Y. (2008). *How to Determine the Progression of Young Skiers?* CHANCE 21, 4, 13-19.
- [3] Hox, J. J., Kreft, I. G. (1994). Multilevel analysis methods. *Sociological Methods and Research*, 22, 283-299.
- [4] Liang, Y., von Rosen, D., von Rosen, T. (2014), *On Estimation in Hierarchical Models with Block Circular Covariance Structures*. Annals of the Institute of Statistical Mathematics. DOI 10.1007/s10463-014-0475-8.

Yuli Liang är doktorand vid Stockholms universitet, här framför Teleborgs slott i Växjö.

Som utlovat i förra numret av Qvintensen skriver Jakob Bergman här om skillnader mellan de olika opinionsinstituten mätningar.

Spelar det någon roll vilken opinionsundersökning man tar del av?

Opinionsinstitut använder telefonintervjuer eller webbenkäter. De använder metoder som skiljer sig åt i insamlingsperiodens längd, kontaktstrategi, grad av bearbetning, t.ex. efterstratifiering. Deras urvalsstorlekar skiljer sig också åt. Men alla har som mål att spegla partisympatierna i väljarkåren vid en viss tidpunkt. Olika mediaföretag samarbetar med olika opinionsinstitut och därför presenteras vanligen olika mätningar i olika media. En fråga som man då ställer sig är: spelar det någon roll vilken mätning man tar del av?

En viktig sak att observera när det gäller undersökningar av partisympatier är att de anger andelen väljare (den relativa frekvensen väljare) och inte antalet väljare (den absoluta frekvensen). Andelen som sympatiserar med ett parti är beroende av de andelar som sympatiserar med andra partier. Om ett parti ökar måste ökningen tas från andra partier, som alltså minskar. Detta fenomen påpekades redan av Pearson (1897), som kallade det ”spurious correlation”, men det har dessvärre inte alltid observerats av sentida efterföljare.

Björn Holmquist och jag har utvecklat en metod som utgår från att andelarna är ”compositions”; en composition är en summa av slumpvariabler som alltid summerar till en och samma konstant (Bergman och Holmquist 2013). I vårt fall är den konstanten 1, eftersom andelarna alltid summerar till 1. ”Sammansättning” är kanske en lämplig svensk term för composition.

Vi har valt att väga samman mätningarna från olika opinionsinstitut med en variant av en lokalt vägd regressionsmodell, loess (Cleveland 1979). Skillnaden är att vi har anpassat loess till

sammansättningar. I ”vanlig” regression väljer man att beräkna predikterade värden $\hat{x}_i$ så att summan av kvadraterna mellan observerat och predikterat värde, $\sum_{i=1}^n (x_i - \hat{x}_i)^2$, minimeras (OLS, ordinary least squares). Ibland vill man ha med en vikt, till exempel för att justera för heteroskedasticitet: då minimeras $\sum_{i=1}^n w_i (x_i - \hat{x}_i)^2$ (WLS, weighted least squares). I loess minimerar man samma summa men med vikter som gör att ju närmare x_i en observation är, desto mer räknas den med i beräkningen av $\hat{x}_i$. Utanför ett fönster som omfattar de q närmaste grannarna är vikten noll (med något lämpligt valt q).

För att kunna anpassa loess till sammansättningar behövs två operationer, en transformation och ett avståndsmått. Låt y vara resultatet av ett opinionsinstituts mätning, $y = (y_1, y_2, \dots, y_9)'$ om vi fokuserar på de åtta partierna i Riksdagen samt y_9 för övriga. Summan av $y_1, y_2, \dots, y_9$ är 1.

Först definierar vi en operation som kombinerar två sammansättningar:

$x \otimes y = C(x_1, y_1, x_2, y_2, \dots, x_9, y_9)'$, där C en normering som säkerställer att $x \otimes y$ också är en sammansättning. Som symbolen antyder är det en motsvarighet till addition.

Sedan definierar vi en operation för att ”multiplicera” en sammansättning med en konstant: $x \otimes y = C(x_1^a, x_2^a, \dots, x_9^a)'$. Man kan visa att de här två operationerna bildar ett vektorrum på vektorer av typen $y = (y_1, y_2, \dots, y_9)'$ där elementens summa är 1. Den kurva som ska anpassas till ”punkterna” (t, y) har formen $\beta_0 \otimes (t_i \otimes \beta_1)$ där t_i är tiden (det går förstås att ha andra förklarande variabler) och parametrarna β_0 och β_1 är sammansättningar. Det här uttrycket motsvarar $\beta_0 + \beta_1 t_i$ i ”vanlig” regression. En transformation som är lämplig i det här sammanhanget är $clr(y) = (\log(\frac{y_1}{g(y)}), \dots, \log(\frac{y_9}{g(y)}))$ där $g(y)$ är det geometriska medelvärde av $y_1, y_2, \dots, y_9$.

Bokstäverna *clr* står för centered log ratio.

Som avstånd mellan två sammansättningar, x och y , har vi valt kvadraten av det euklidiska avståndet mellan $clr(x)$ och $clr(y)$. Detta avståndsmått kallas Aitchison-avståndet (Aitchison 1983). Vi betecknar avståndet mellan x och y med $d^2(x, y)$.

Loess anpassat för sammansättningar är den kurva som minimerar summan

$$\sum_{i=1}^n w_i d^2[y_i, \beta_0 \otimes (t_i \otimes \beta_1)].$$

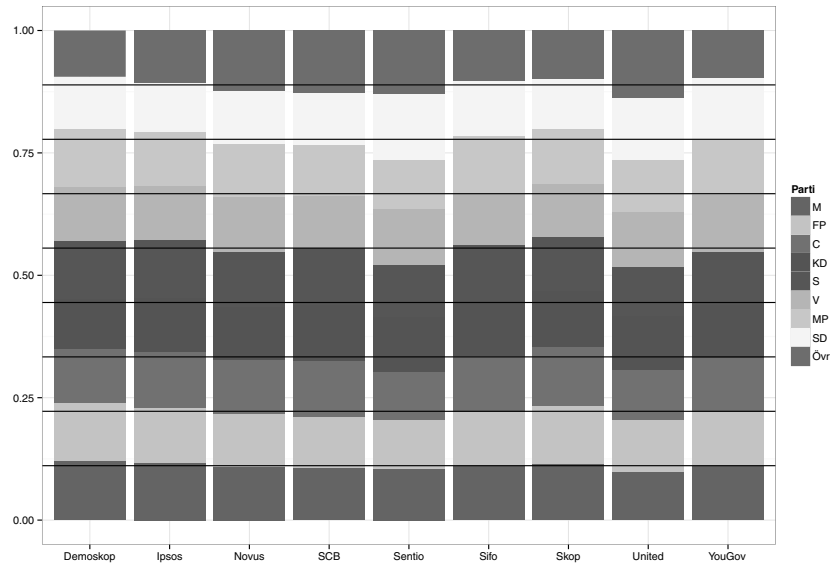
Att beteckna vikten med w_i är oegentligt eftersom den beror av hur långt det är från observation i till vardera av de närmaste grannarna. Summan ser komplicerad ut men den är en analogi till summan i WLS; det kvadratiske avståndet mellan det observerade y_i och det predikterade $\beta_0 \otimes (t_i \otimes \beta_1)$ med en vikt w_i .

Om syftet är att få en så bra ”poll of polls” som möjligt bör man ta hänsyn till att instituten drar olika stora urval. De med störst urval bör få störst vikt i den sammansatta mätningen. Att anpassa metoden till det är en fråga för fortsatt forskning. Men om syftet är att se om det verkar finnas systematiska skillnader mellan institutens opinionsmätningar, kan vår loessskattning användas som den är.

I juni 2013 hade vi resultatet av 509 partisympatiundersökningar, med start september 2006. De kom från nio olika opinionsinstitut. Data har vi fått från Novus (2014). Ett diagram över sammanvägda skattningar fanns på sidan 31 i nummer 2/2014 av Qvintensen. Nu ska vi se hur instituten förhåller sig till varandra.

För var och en av de 509 opinionsundersökningarna mäter vi hur långt ifrån den sammanvägda skattningen den ligger, med avståndet mätt som Aitchisonavstånd. Tabell 1 visar medelvärde för nio opinionsinstitut. Man kan notera

Figur 1. Genomsnittliga avvikelser för respektive opinionsinstitut i förhållande till den sammanvägda skattningen. De svarta linjerna anger 1/9, 2/9, etc. Se figuren i färg på sista sidan.



att Novus och Sifo i genomsnitt ligger närmast den sammanvägda skattningen och att United Minds och Sentio Research ligger längst ifrån. Skillnaderna i medelavstånd mellan de olika instituten är för stora för att rimligen ha kunnat uppstå om det bara var slump. Det betyder förstås inte att Novus är bäst. Tyvärr finns ju inte möjligheten att jämföra resultaten med de sanna partisympatierna annat än vid val och då är bortfallet ca 20 procent.

Olika partier gynnas av olika institut

Vi kan undersöka om olika institut systematiskt tenderar att över- respektive underskatta vissa partier i förhållande till vår sammanvägda skattning. Detta systematiska fel brukar på engelska kallas för house effects. För varje mätning tittar vi på avvikelserna u mellan det faktiska värdet x och den sammanvägda skattningen $\hat{x}$:

$$u = x - \hat{x} = C(x_1/\hat{x}_1, x_2/\hat{x}_2, \dots, x_9/\hat{x}_9)'$$

Därefter beräknar vi komponentvis det geo-

metriska medelvärdet (normerat så summan blir 1) av dessa avvikelser uppdelat på de olika instituten. På så sätt får vi en institutens medelavvikelse från den sammanvägda skattningen. I Figur 1 på tidningens baksida visas de genomsnittliga avvikelserna för de olika instituten. Man kan tänka på avvikelserna som ett slags vikter. Om ett institut har avvikelsen 1/9 för alla partier innebär det att inget parti över- eller underskattas i genomsnitt (dvs. alla partier får lika vikt). Om ett parti får mindre än 1/9 så innebär detta att institutet i genomsnitt underskattar partiet och om det får mer än 1/9 så överskattar institutet partiet i förhållande till den sammanvägda skattningen.

Vi kan se att Novus, SCB, Sifo och YouGov ligger mycket nära de svarta linjerna (1/9, 2/9, etc.) och således i genomsnitt nära den sammanvägda skattningen. Sentio Research, Skop och United Minds ligger längst ifrån. United Minds tenderar att underskatta M, FP, C och S samt överskatta SD och Övriga partier, medan Skop tenderar att överskatta de borgerliga och underskatta SD och Övriga partier i förhållande till den sammanvägda skattningen.

En invändning skulle kunna vara att institut med många mätningar har större inflytande på sammanvägningen och alltså skulle gynnas vid jäm-

förelse med övriga institut. Detta tycks dock inte vara fallet, då både institut med många mätningar (t.ex. Novus) och väldigt få mätningar (t.ex. SCB) klarar sig mycket bra.

Oavsett vilket institut som är närmast sanningen kan man dra slutsatsen att det finns systematiska skillnader mellan instituten och att det således kan ha betydelse vem man lyssnar till.

JAKOB BERGMAN
LUNDS UNIVERSITET



Referenser

- Aitchison, J. (1983). Principal component analysis of compositional data. *Biometrika*, 70, 57-65.
- Bergman, J. och Holmquist, B. (2013). Poll of polls: A compositional loess model. *Scandinavian Journal of Statistics*, 41, 301-310.
- Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74, 829-837.
- Novus (2014). Samtliga svenska väljarbarometrar. Webbadress: <http://www.novus.se/vaeljaropinionen/samtliga-svenska-vaeljarbarometrar.aspx> (Excellfil hämtad i sep 2013).
- Pearson, K. (1897). Mathematical contributions to the theory of evolution. On a form of spurious correlation which may arise when indices are used in the measurement of organs. *Proceedings of the Royal Society of London*, LX, 489-498.

Högskoleutforskaren är ett analysverktyg som alla kan använda. Det är även möjligt att ladda upp egna data. Användarna kan, förutom att titta på de färdiga historierna, själva välja vilka variabler de vill jämföra och vilken typ av diagram de vill se.

Artikelförfattarna, från vänster: Niklas Karlsson, Anders Viberg, Andrea Amft och Max Kesselberg



FOTO: KARIN WRETE

Högskoleutforskaren underlättar för lärosätena att själva analysera sin verksamhet

Hur använder egentligen universiteten och högskolorna sina resurser? Granskning av effektivitet är en ny uppgift som gavs till Universitetskanslersämbetet (UKÄ) vid starten av myndigheten 2013. Myndigheten ska granska hur effektivt verksamheten bedrivs vid universitet och högskolor.

– Hela högskolesektorn omsätter 74 miljarder kronor och är en av de största statliga sektorerna. Det är rimligt att se vad pengarna används till. Samtidigt är det här en uppgift som har inneburit ett stort utvecklingsarbete och som kommer att innebära ett fortsatt kontinuerligt utvecklingsarbete. Vad betyder "effektivitet" och hur ska vi mäta den? Det är frågor som vi fortfarande arbetar med och som vi har täta kontakter med högskolesektorn om säger Annika Pontén, chef för UKÄ:s analysavdelning.

Utmaning att hitta lämpligt mått

Anders Viberg, Andrea Amft, Max Kesselberg och Niklas Karlsson, utredare på analysavdelningen, leder arbetet med att utveckla presentationsformerna för ämbetets granskning av effektivitet och ta fram lämpliga underlag för Högskoleutforskaren som är namnet på ämbetets nya webbaserade interaktiva verktyg.

– "Effektivitet" är ett omfattande begrepp och det har varit en utmaning att hitta lämpliga mått. Vi lägger även ner mycket arbete på att studera hur andra länder mäter, men det är få som verkligen mäter effektiviteten i resursutnyttjandet. När det gäller effektiviteten inom högskolan tror vi dock att Sverige ligger bra till i ett internationellt perspektiv, säger Anders Viberg.

På UKÄ:s webbplats har vi valt att vid sidan av längre effektivitetsanalyser även

satsa på en mer kortfattad interaktiv presentation av uppgifter som tas fram av UKÄ.

– Vi kallar det för "effektivitetsfakta", berättar Andrea Amft. Vi presenterar ett faktaunderlag på ett kortfattat och lättillgängligt sätt tillsammans med rörliga och interaktiva diagram som gör det enklare att visa utvecklingen över tid. I vissa fall finns även längre och djupare effektivitetsanalyser.

– I det nya interaktiva verktyget som vi använder oss av har vi ringat in några olika områden som fokuserar på lärosätenas produktivitet och resursutnyttjande. Vissa områden är utpekade av regeringen medan annat är initierat av UKÄ, säger Max Kesselberg.

"Vi ser det som benchmarking mellan lärosätena"

UKÄ bedömer inte hur väl universiteten och högskolorna hushållar med sina pengar.

– Vi respekterar lärosätenas självständighet. Vi redovisar genomsnittliga kostnader och tillgängliggör olika data som lärosätena kan använda för att själva analysera hur de använder sina resurser. Det handlar mer om benchmarking. Kvalitet måste ju få kosta, men pengarna ska naturligtvis användas på ett befogat sätt. Här behöver lärosätena komma överens om en rimlig nivå. Det är även uttryckt i högskolelagen att resurserna ska utnyttjas effektivt – lärosätena ska alltså ta ansvar för detta, säger Andrea Amft.

En del av arbetet med effektivitet handlar om att tillhandahålla uppgifter så att lärosätena kan jämföra sig med varandra när det gäller resursutnyttjande och produktivitet och själva bedöma rimligheten i nivåerna.

– Vi tycker inte att det är vår roll att bedöma detta, utan vi vill snarare bistå med relevanta underlag för både lärosätenas och

regeringens bedömningar. Vi gör också analyser som är mer tematiskt djupgående medan annat publiceras som effektivitetsfakta – en kortare visuell presentation med rörliga diagram som enkelt visar utvecklingen över tid.

– Våra effektivitetsfakta har tagits emot väl. Det är uppskattat att få en snabb överblick på ett visst tema på ett visualiserat sätt, säger Andrea Amft.

Högskoleutforskaren underlättar jämförelser

För den som vill gå djupare i data än de effektivitetsfakta vi presenterar är Högskoleutforskaren ett analysverktyg som är fritt att använda. Statistik som har relevans för resursutnyttjande och produktivitet kommer att läggas in efter hand, men det är även möjligt att ladda upp egna data och spara lokalt. Användarna kan, förutom att titta på de färdiga historierna, själva välja vilka

»Vi ser gärna att lärosätena använder sig av Högskoleutforskaren för att analysera hur de själva använder sina resurser och jämföra med andra lärosäten.»

– Det ska poängteras att UKÄ inte kan se vilka filer eller information som laddas upp eller vilka analyser eller beräkningar som externa användare gör i Högskoleutforskaren, säger Niklas Karlsson.

– Vi ser gärna att lärosätena använder sig av Högskoleutforskaren för att analysera hur de själva använder sina resurser och jämföra med andra lärosäten. Förhoppningsvis blir verktyget ett hjälpmedel för lärosätena i deras egen analys och planering av resursutnyttjandet, säger Andrea Amft.

– Vi höll en kurs om Högskoleutforskaren i våras, som blev fulltecknad på två dagar. Det var framför allt personer som jobbar med ekonomi och planering på lärosätena som deltog. Vi anordnade två nya kurstillfällen i september, även dessa platser tog slut på en gång.

Här följer utdrag från tre effektivitetsfakta: Högskolans lokaler, Högskolans personalkostnader och Etablering på arbetsmarknaden.

Högskolans lokaler

Under perioden 1994–2013 har en del av lärosätenas ökade intäkter använts för finansiering av ökat lokalbehov, men majoriteten av lärosätena har ändå minskat sin andel av lokalkostnaderna. Detta avspeglas i Diagram 1, på sidan 31, som visar att den hyrda arean för de flesta lärosäten ökar i takt med hyreskostnaderna i löpande priser. Detta visas också av cirkelns färg. I huvudsak har dock skillnaderna i års

hyror mellan lärosätena sin grund i vilken typ av utbildning och forskning som utförs. Konstnärlig utbildning och idrott är mer än genomsnittligt lokalkrävande och det gäller även teknik och naturvetenskap, medan resurserna och lokalbehoven är mindre inom

de samhällsvetenskapliga och humanistiska områdena.

Högskolans personalkostnader

UKÄ har studerat löneutvecklingen för professorer, lektorer och adjunkter. Vi har räknat fram hur lärosätenas lönekostnader för professorer, lektorer respektive adjunkter har förändrats genom att multiplicera medianlönen med antalet helårspersoner i respektive grupp. Förändringen beror på att

1. antalet anställda har ökat.
2. lönerna har ökat.
3. andelen professorer, lektorer respektive adjunkter har förändrats.

Ökningen av lönekostnaderna som beror på förändrad personalsammansättning ligger strax under 5 procent, se Diagram 2 på sidan 31. Det motsvarar en kostnad på nära 29 miljoner kronor i 2013 års priser. Räknar vi med sociala avgifter som i allmänhet är cirka 50 procent av lönen blir lärosätenas merkostnad cirka 43 miljoner kronor.

Etablering på arbetsmarknaden

Högskoleutbildade kan räkna med jobb. Nästan 80 procent av de som tog examen från högskolan under läsåret 2009/10 hade etablerat sig på arbetsmarknaden under år 2011. De högsta andelarna etablerade hade bland annat läkare, psykoterapeuter, yrkeslärare samt högskole- och civilingenjörer med bygginriktning.

Det är stora skillnader i etableringen mellan olika examina. I Diagram 3 visas utvecklingen för några examina mellan åren 1996–2011. På x-axeln visas antal examinerade och på y-axeln andel etablerade.

Storbritannien visar intresse för UKÄ:s arbete

UKÄ studerar även hur andra länder jobbar med effektivitetsfrågan. Andrea Amft har tillsammans med två andra kolleger besökt Universities UK (Storbritanniens motsvarighet till Sveriges universitets- och högskoleförbund, SUHF) för att lära av dem och utbyta erfarenheter. I Storbritannien genomförs just nu stora besparingar och det finns därmed ett stort fokus på kostnadseffektivitet.

– De brittiska lärosätena ska göra mer för mindre pengar. Dessutom har de studieavgifter, vilka höjts kraftigt, vilket innebär att studenterna mer ses som "kunder" som kan kräva att deras

»»»

pengar används på ett bra sätt som gynnar deras utbildningar. Det är inte alls samma förutsättningar som i Sverige. De brittiska lärosätena är mer företagslika och äger bland annat sina lokaler, framhåller Andrea Amft.

– Genom att studera andra länders system mer djupgående blir det tydligare vilka förutsättningar och drivkrafter vi har i Sverige. Det är mycket värdefullt för oss i vårt fortsatta utvecklingsarbete.

Andrea Amft berättar att Universities UK ville höra mycket om hur UKÄ arbetar med effektivitetsuppgiften.

– De var väldigt engagerade och gillade särskilt våra visualiseringar på webbplatsen, det vill säga Effektivitetsfakta och Högskoleutforskaren. Vi tog det som en bekräftelse på att vi är på rätt väg. Universities UK driver webbplatsen Efficiency Exchange som är en mötesplats där brittiska lärosäten kan dela med sig av goda exempel för att främja effektiviteten.

– Vi blev ombedda att skriva en artikel om hur UKÄ arbetar med effektivitetsgranskningen och våra erfarenheter. Artikeln har precis blivit publicerad på Efficiency Exchange. Även tidningen the Guardian har publicerat artikeln, säger Andrea Amft.

– Vi kommer att hålla kontakten med Storbritannien. Efter sommaren ska vi också titta närmare på hur Finland tar sig an den här uppgiften. Kostnadseffektivitet är i fokus även där och Finland är ett land som liknar oss mer än Storbritannien.

ANDREA AMFT,
NIKLAS KARLSSON,
MAX KESSELBERG
OCH ANDERS VIBERG,
UNIVERSITETSKANSLERSÄMBETET

Se samtliga effektivitetsanalyser och effektivitetsfakta:

<http://www.uka.se/effektivitet>

Till Högskoleutforskaren – gör din egen analys:

<http://www.uka.se/hogskoleutforskaren>

Läs artikeln i the Guardian om UKÄ:s arbete:

<http://www.theguardian.com/higher-education-network/efficiency-exchange-partner-zone/monitoring-efficiency-swedish-higher-education>

Efficiency Exchange :

<http://www.efficiencyexchange.ac.uk/4644/swedish-higher-education-authority-visit-to-universities-uk-and-hesa/>

DISKUSSION

Hans Nyquist diskuterar undervisning i statistik. Artikeln är baserad på ett föredrag på Cramérsällskapets och Föreningen industriell statistiks gemensamma vårkonferens i mars i år.

Kursen som jag nätt och jämnt fick godkänt på

En journalist på tv refererade för en liten tid sedan en debatt mellan politiker om skolfrågan som debatterades intensivt före valet. Han kallade

den ena politikern för ”det levande Excel-dokumentet” och sa att ”det är inte statistik vi behöver i skolfrågan, utan visioner.” För mig finns det inget motsatsförhållande mellan statistik och visioner. Det kanske till och med tvärt om är så att statistik kan hjälpa till att skapa och realisera visioner. Uttalandet ger ändå en bild av hur statistik kan uppfattas av många ute i samhället; oantastliga fakta som man kan slå varandra i huvudet med i politiska debatter, men som inte leder till någon konstruktiv utveckling. Den som kan presentera sifferuppgifter, i stort sett oavsett vilka, har ett övertag i en argumentation.

Journalistens uttalande har fått mig att fundera på hur statistikämnet uppfattas ute i samhället och hur den uppfattningen kan ha skapats.

Förra året besökte jag en statistikinstitution i USA där man förde en diskussion om statistikämnet, hur det uppfattas och vad vi ska ha det till. Man hade gjort ett litet presentationsblad av ämnet där man skrev:

When many people hear the word “statistics” they think on either sports-related numbers or the college class they took and barely passed. While statistics can be thought of in these terms, there is much more to the relationship between you and statistics than you can imagine.

Sedan fortsätter man med informella definitioner av statistik som ”the science of learning from (or making sense out of) data” och ”the theory and methods of extracting information from observational data for solving real-world problems”.

Jag tror att alla av oss är övertygade om att statistik är en viktig och kraftfull verktygslåda för att belysa och hjälpa till att lösa problem i de flesta delar av vårt samhälle, men för den breda massan, inklusive journalisten som gjorde sig lustig över skolpolitikern, är inte den uppfattningen självklar. Många lever sina liv helt omedvetna om statistikens nytta och användning.

Jag inser att uppfattningen om statistik som ”the college class I barely passed” passar allt för väl in på många före detta studenter man träffar. Detta gäller vare sig det är studenter från vårdprogram, ingenjörsprogram eller samhällsvetenskapligt inriktade program. Många studenter har, trots genomgången statistikkurs, ingen aning om att statistik kan användas för att lösa problem i det verkliga livet, än mindre hur det faktiskt skulle gå till att göra det. För dessa studenter framstår statistikkursen som en samling mer eller mindre begripliga lösningar på konstruerade och världsfrånvända problem som saknar praktiskt intresse.

Ett viktigt mål för oss som undervisar i statistik är att medvetandegöra den kraftfulla och viktiga verktygslåda som statistik faktiskt är. Den kan ge oss insikter om olika förhållanden, bidra till att ge oss förklaringar till fenomen, bidra till att utveckla nya produkter och bidra till ett bättre samhälle. För många av våra studenter har detta mål reducerats till ”the college class I barely passed”.

Jag tror det kan finnas många orsaker till varför det har blivit så. En förklaring är att det lätt blir vattentäta skott mellan å ena sidan de begrepp som bär upp vår teori och de analystekniker vi vill lära ut och å andra sidan ”making sense out of data” och ”solving real-world problems”. Det är så lätt att stanna vid definitioner av histogram, sufficiens, oblique rotation och alla andra begrepp vi anser viktiga, utan att sätta in begreppen i en verklighetsnära och verklighetsrelaterad kontext.

Ett viktigt första steg för att visa hur vår teori fungerar i verkligheten sker genom de exempel och övningsuppgifter vi presenterar för våra studenter. Jag har därför bläddrat i några av de böcker som används idag i våra introduktionskurser och hittade då övningsuppgifter där man får veta att det ligger två tioöringar och en tjugofemöring i en börs, man drar på målfå utan återläggning två mynt och har i uppgift att beräkna kovariansen mellan de dragna mynten.

I en annan uppgift får man veta att Per och Pål har elva frukter varav tre är giftiga. Per äter fyra, Pål sex och hunden den kvarvarande frukten. Uppgiften är att beräkna sannolik-

heten att hunden klarar sig, den betingade sannolikheten att både Per och Pål blir förgiftade givet att hunden klarar sig, och så vidare.

Floran av övningar med tveksam verklighetsrelevans kan göras mycket lång. Här ska jag bara nämna ytterligare ett exempel som jag inte hittat i läroböcker, men som förekommer i undervisningen. Man får veta att det finns två pubar, A och B, och man får veta den betingade sannolikheten att man besöker A-puben nästa gång givet att man just besökt den och motsvarande betingade sannolikhet för B-puben. Uppgiften är att bestämma med vilka sannolikheter man kommer att besöka pubarna efter ett mycket stort antal pubkvällar. Att besöka pubar är förvisso viktigt för många, inte minst för många studenter, men är det verkligen frågan med vilken sannolikhet man kommer att besöka A-puben efter 1000 pubkvällar som är den man funderar på?

Till alla problemkonstruktörers försvar ska jag säga att det finns olika syften med övningar. En del övningar behöver bara vara färdighetsträning på så sätt att när man just har lärt sig marginalfördelning kan man behöva summera några rader och några kolumner i en tabell över en bivariat fördelning, utan några större kopplingar till näringsliv eller samhälle i övrigt.

För att färdighetsträningen skall vara meningsfull, måste man emellertid också få veta vad färdigheterna skall användas till.

På våra kurser har vi i allmänhet många saker som skall gås igenom på en mycket begränsad tid. Eftersom verklighetsnära problem ofta är komplicerade och kräver mycket tid att behandla, behöver övningarna ofta förenklas och läggas till rätta. Det finns helt enkelt ingen tid för mer sammansatta tillämpade övningar. Det blir då uppenbart att den traditionella undervisningen behöver kompletteras med moment där det finns möjlighet att ta ett bredare perspektiv på ämnet, moment där man inte fördjupar sig i hur man räknar ut minsta kvadratskattningar, utan om, och i så fall när, det är lämpligt att räkna ut minsta kvadratskattningar. Moment som innehåller mer av ax till kaka-diskussioner och mindre av mjölets beståndsdelar, alltså.

Vi har idag både kandidatuppsatser och masteruppsatser där det finns möjlighet att

arbeta med mer ostrukturerade problem och problem som kan vara av intresse för världen utanför statistikinstitutionerna. Tyvärr kommer uppsatserna relativt sent i utbildningen och det är en nästan försvinnande liten andel av studenterna som kommer dit. Det är bara en liten andel studenter som läser mer än 30 hp (en termin) och de flesta faktiskt bara upp till 7,5 hp (en fjärdedels termin).

Det är därför önskvärt, rentav nödvändigt, att arbeta med mindre projekt och med kursdelar som tar upp ett bredare perspektiv och placerar in ämnet i ett större sammanhang redan i introduktionskurser.

Här blir det också nödvändigt att ta in en del vetenskapsteoretiska aspekter på ämnet. Med det menar jag att vi behöver kunna föra en jordnära diskussion om begrepp som kausalitet och förklaringar och varför kausalitet och förklaringar inte kommer som ett resultat från en datorkörning, inte ens om man använder de senaste tipp topp programmen.

Jag ska ta ett exempel som visar behovet att ta ett bredare perspektiv på vårt ämne. För en tid sedan skrev Karin Bojs i Dagens Nyheter under rubriken ”Falsk matematik och farmors matvanor” om en tveksamt utförd undersökning om hur farmors diet påverkade barnbarnens hjärt- och kärlsjukdomar. Hade vi haft moment i vår utbildning som hade tagit upp problemet med att ”fiska

vilt bland alla upptänkliga samband”, som Karin Bojs skrev, och om forskarna som gjorde undersökningen hade genomgått och tagit till sig en statistikkurs med vad jag menar med ett bredare perspektiv, så hade problemet inte behövt uppstå. Notera här också att problemet inte berodde på att man hade valt minsta kvadratmetoden i stället för maximum likelihood eller på att man hade strukturerat en kovariansmatris på det ena eller det andra sättet. Problemet hade sin grund i bristen på helhetssyn om statistiska metoder eller bristen på ett vetenskapsteoretiskt förhållningssätt. [Redaktörens anmärkning: den diskussionen fortsätter på sidan 15]

Datorerna och datorernas utveckling samt utvecklingen av programvara har förändrat för-

utsättningarna för statistisk analys. Vi kan idag med lätthet hantera modeller och genomföra analyser som vi för 35 år sedan inte ens kunde tänka var möjliga. Till detta har kommit möjligheter att göra observationer på helt nya sätt som är både enklare och ger oss tillgång till ofantligt mycket större datamängder än vad som fanns tidigare. Jag vill inte påstå att vår utbildning håller jämna steg med den tekniska utvecklingen, men den tekniska utvecklingen har påverkat undervisningen, åtminstone vad gäller undervisningens innehåll på så sätt att vi idag kan undervisa om både fler och mer avancerade analyser. Den ökade tillgången på datorbaserade statistikpaket och omfattande mängder av data är emellertid inte enbart av godo. Det ökar också möjligheterna till tanklös och okritisk användning av statistiska metoder. Vi har, så att säga, idag möjlighet att göra fler fel på kortare tid än vad vi tidigare har kunnat.

Listan på undersökningar med tanklös och okritisk användning av statistiska metoder kan göras mycket lång. Självklart skulle floran av tveksamt utnyttjade data och vilseledande slutsatser kunna begränsas om vi kunde skapa en större medvetenhet om vetenskaplig metod och om undersökningsmetodik i vår utbildning. Vikten av insikt i vetenskapliga metoder poängteras också av Olle Häggström i hans artikel i Qvintensen nr 1/2014.

Vilka hinder finns det då för att öka medvetenheten om ett vetenskapligt förhållningssätt till statistisk analys? Ett hinder är att en del av produktionen av statistiska undersökningar utförs av personer med otillräcklig statistikutbildning, inte sällan av personer med en helt annan grundutbildning än i statistik.

Ett medel att råda bot på det är att arbeta för att kurser i statistik ingår i fler utbildningsprogram på universitet och högskolor.

Dessvärre är utrymmen på utbildningsprogram begränsade samtidigt som flera andra ämnen också arbetar för att expandera. Därför behövs goda argument för att kunna få gehör för ambitionen att komma in på fler program. Som jag ser det är det helt nödvändigt att vi får möjlighet att delta i fler program, samtidigt som

»»» STATISTICS – THE COLLEGE CLASS...

detta också utgör en av de största farorna. Det är nämligen av största vikt att det vi lär ut på kurserna är relevant för kursdeltagarna. Antag att vi ges tillfälle att ge en kurs för sociologer och där berättar om hur sannolikt det är att hunden dör om Per och Pål har elva äpplen. Den sortens kursinnehåll ger helt säkert en kurs som ”you barely passed” och genererar inga som helst positiva avtryck när kursdeltagarna ska använda statistiska metoder i sin framtida yrkesverksamhet.

De tekniska utbildningsprogrammen brukar vara duktiga på att ge statistikkurser, men hur används kunskaperna senare i utbildningen? Förmodligen vill inte civilingenjörer heller lära sig om Per och Påls äpplen. De vill veta hur sannolikt det är att en konstruktion kommer att fungera tillfredsställande och om det är möjligt att designa konstruktionen på något sätt så att sannolikheten för en tillfredsställande funktion ökar.

Precis som Ghazi Shukur påpekar i sin krönika i Qvintensen nr 2/2013 måste våra kurser fyllas med ett innehåll som är relevant för kursdeltagarna, annars blir vi snart, och med rätta, utkastade från utbildningsprogrammen.

Om vi blir utkastade, blir följden att vi får ännu fler aningslösa tillämpningar av statistiska metoder och ännu fler undersökningar med vilseledande resultat. Det är därför viktigt att vi arbetar för att statistik kunskaperna kommer till användning i olika delar av ett utbildningsprogram.

Ett sätt, kanske en förutsättning, att uppnå det är att samarbeta med andra ämnen. Andra ämnen kan komma in och berika våra kurser med relevanta problemställningar och vi kan komma in i andra ämnens kurser och berika dem med goda lösningar på empiriska problem.

Det finns ett annat stort hinder för en positiv utveckling av utbildningen i statistik, nämligen vi själva, eller vår oförmåga eller ovilja att arbeta tillsammans med andra och ta till oss av deras problem. Jag inser att många av oss vill fortsätta leva i en skyddad värld där de stora världsproblemen består av tärningskast och Per och Pål som råkar äta förgiftade frukter givet att hunden inte åt en sådan, eller hur det nu var. Att

inte vilja samarbeta med andra ämnen isolerar oss, fjärrar våra problem från verkligheten och bidrar till att reducera våra kurser till ”the college class I barely passed”.

Jag ska nämna två andra exempel på oförmåga som hinder för ämnets utveckling. Det första handlar om avsaknad av insikt om vad som är kärnan i vårt ämne. När man utformar ett utbildningsprogram föreställer jag mig att en utgångspunkt är att definiera vad som är viktigt för programmet, vad man vill att deltagarna i utbildningen ska kunna när de har gått igenom den. Nästa steg är att konstruera kurser så att kunskapsmålen uppfylls. Det kan nu inträffa att de kurser som leder fram till målen i sin tur kräver förkunskaper, vilket i så fall får lösas genom att konstruera kurser med målsättningen att ge de nödvändiga förkunskaperna. För att till exempel kunna tillgodogöra sig en kurs i sannolikhetsteori kan studenterna behöva ett visst mått av kunskaper i matematik. Vad som behövs är en progression av kunskap som leder till programmets mål. Dessvärre är detta inte helt uppenbart för alla. Härförleden stötte jag på ett utbildningsprogram i statistik som förvisso innehåller ett antal olika kurser om statistiska analystekniker, men, som det verkade utgjorde dessa kurser snarare förkunskapskrav till en senare kurs om numerisk lösning av ekvationer, numerisk optimering av funktioner, simuleringstekniker etc. Ett sådant upplägg av ett utbildningsprogram signalerar att målet för en god statistisk analys är att ha många korrekt uträknade decimaler i p-värdet då man bestämmer signifikansen av farmors diet som orsak till barnbarnens hjärt- och kärlsjukdomar. Om modeller som beskriver samband mellan variabler och statistisk metodik reduceras från huvudämne till att blott vara förkunskaper för något annat, och beräkningstekniker därvid hamnar i förarsätet, då är det lätt hänt att inferensens grundvalar, inklusive problemet med multipel signifikans, glöms bort till förmån för den falska matematik som Karin Bojs skriver om. Upplägget av den här utbildningen fjärrar sig således från de vetenskapliga problemställningarna.

Det andra exemplet på oförmåga handlar om ren och skär okunskap. För en tid sedan såg jag en students lösning på en tentamensuppgift.

Låt $Y = \min(X_1, X_2)$. Då är

$$F_Y(y) = P(X_1 \leq y \cup X_2 \leq y) = P(X_1 \leq y) + P(X_2 \leq y) - P(X_1 \leq y \cap X_2 \leq y) \\ = (1 - e^{-y/\lambda}) + (1 - e^{-y/\lambda}) - (1 - e^{-y/\lambda})(1 - e^{-y/\lambda}) = 1 - e^{-2y/\lambda}$$

Uppgiften bestod i att bestämma sannolikhetsfördelningen för den kortaste av två oberoende, exponentialfördelade livslängder, X_1 och X_2 , båda med samma parameter, λ . Ett enkelt standardproblem, kan man tycka. Studenten i fråga hade gjort en måhända ovanlig, men innovativ lösning genom att gå utanför textbokens standardlösning och använda additionssatsen (se formel i rutan här ovan).

De rättande lärarna kände tydligen inte till något annat än textbokens standardlösning och beslutade att den här lösningen är fel. Visserligen blev det rätt svar, men det måste, menade man, ha varit en tillfällighet. Den slutliga förklaringen till att lösningen bedömdes som felaktig var att symbolerna $\cup$ och $\cap$ fanns med i lösningen och ”eftersom $\cup$ och $\cap$ inte är symboler som finns i statistikteorin så måste lösningen per definition vara fel”!

Jag menar att det är väldigt sällan som okunskap och rena felaktigheter leder till en positiv utveckling. Den här typen av oförmåga ger ytterligare näring åt uppfattningen om statistik som the college class I barely passed. Däremot kan, menar jag, innovativa lösningar och nytänkande bidra till en positiv utveckling av ämnet, speciellt om innovationerna kommer från studenter, men vi måste ha mod och insikt och kunskap för att släppa fram dem.

Sammanfattningsvis ser jag det som mycket viktigt att statistikämnet får en större förankring i samhället och får inta den kunskapsgenererande funktion som det kan ha. En förutsättning för att uppnå detta är att undervisningen i statistikämnet utvecklas varvid det är speciellt viktigt att:

- Kurserna fylls med tillämpbart och verklighetsnära innehåll.
- Statistikämnet framställs som en vetenskaplig metod för att studera empiriska problem.
- Statistiska metoder lärs ut av kvalificerade statistiker i fler utbildningsprogram.
- Moment att resonera statistiskt förs in parallellt med moment med statistiska tekniker (att resonera fritt i stället för att blint och ångestfyllt följa kokboksrecept).
- Lärarkåren har hög statistisk kompetens.

HANS NYQUIST

Det är sällan som den statistiska frågan om hantering av multipla test når dagstidningarna. Men det hände faktiskt i mars i år.

EN STUDIE I MULTIPLA TEST

Om man testar en nollhypotes på nivån 5% är risken att man felaktigt förkastar nollhypotesen högst 5%. Nollhypotesen kan vara att ”far- och morföräldrars period av svält i barndomen inte påverkar risken för barnbarnen att drabbas av hjärt-kärlsjukdomar”. Om det är osannolikt att man skulle ha fått de data som man har samlat in, eller att man skulle ha fått ännu mer extrema data, om nollhypotesen verkligen vore sann, då förkastar man nollhypotesen. Om man testar två nollhypoteser vid samma tillfälle växer risken att man felaktigt förkastar någon av dem till ungefär 10%. Tre nollhypoteser och det blir 14%. Sexton nollhypoteser – då är risken 56%. Det går att justera för denna risktillväxt.

En artikel (Bygren et al. 2014) med intressanta resultat i området epigenetik (ärfylliga förändringar som inte är associerade med förändring i själva DNA:t) omtalades i flera dagstidningar och även i Forskning & Framsteg. Sveriges Radios vetenskapsredaktion bestämde sig för inte ta upp studien eftersom man inte litade på den. Studien sa att om farmödrar har genomgått en drastisk förändring i tillgången på mat, vare sig det var från bra till dålig tillgång eller vice versa, kan det påverka risken för hjärt-kärlsjukdomar i tredje generationen. Forskarna har studerat 317 slumpmässigt valda personer födda 1890, 1905 eller 1920 i Överkalix enligt folkbokföringen. De kallas nedan för barnbarn. Personer med icke fullständiga data uteslöts och kvar blev 277 personer. Vid två tillfällen under den studerade perioden, 1812–13 and 1821–22, följdes ett års missväxt av ett gott år. Vid tre tillfällen var det tvärtom; det hände 1799–1800, 1876–77 och 1880–81. Om far- eller morföräldrar levde och var högst 13 år vid någon av de fem perioderna (även om de var ännu ofödda foster), då räknade forskarna in dem som utsatta för drastisk förändring i tillgång på mat, annars ej.

Man kan förstås inte veta vad folk åt under de fem tillfällena men eftersom Överkalix var så isolerat kan man anta att dålig skörd betydde knapp tillgång på mat och eftersom det var svårt för bönderna att sälja

och transportera bort överskottet av en god skörd blev det antagligen gott om mat kvar i byn. Inte heller kan man exakt veta hur skördarna var i Överkalix. En av de källor som använts för att bedöma skörden är Statistisk Tidskrift från 1871 (Bygren et al. 2001).

Den modell som valdes var en proportional hazards model, stratifierad efter barnbarnets kön. En av de potentiella oberoende variablerna sällades bort eftersom ett lograng-test gav ett p-värde större än 0.25. Kvar blev kohort, moderns läskunnighet och om fadern till barnbarnet dog innan barnbarnet fyllt 13 år. (De av er som är särskilt intresserade av lograngtest kan titta tillbaka på Oskar Hagbergs artikel i Qvintensen nr 1/2014) . Utfallsvariabeln var om barnbarnet dött i hjärt-kärlsjukdom eller ej före uppföljningstidpunkten 2011-12-31. Tabell 2 på sidan 17 återger några av studiens resultat.

Karin Bojs skrev i en krönika i Dagens Nyheter om studien (Bojs, 2014). Men den fick rubriken ”Falsk matematik om farmors matvanor”. Varför det?

Efter att ha rådfrågat Olle Häggström, som är professor i matematisk statistik, hade Karin Bojs funnit att artikeln testar åtskilliga hypoteser och lyft fram den av dem med lägst p-värde och deklarerat denna som statistiskt signifikant. För en statistiker ligger en alternativ tolkning närmare till hands: fyndet är en statistisk chimär och skälet till det är just att åtskilliga hypoteser testas, alltså det statistiska spratt som multipla test kan utsätta forskaren eller utredaren för. Även Sveriges Radio hade kontaktat Olle Häggström och SR valde som sagt att ligga lågt.

Det var bara hos sondötrarna som studien visade signifikant effekt av drastisk förändring i tillgång på mat. Tabell 2 på sidan 17 visar visserligen bara signifikant effekt om förändringen gick från bra till dålig tillgång, men studien förtäljer att författarna även har provat att kontrollera för hur många de drastiska förändringarna var, och då blev även riktningen dåligt till bra signifikant.

Artikelförfattarna skriver att det är märkligt att effekten när den går över tre generationer bara syns när farmodern genomgått en drastisk förändring på

»»» EN STUDIE I ...

mat, inte till exempel mormodern, och bara när barnbarnet är en kvinna. De diskuterar tänkbara förklaringar till dessa ’rather peculiar phenomena’. För en statistiker är det som sagt inte alls märkvärdigt, utan tvärtom något man förväntar sig i en studie med många signifikanstest. Tabell 2 innehåller 16 test. Studien omtalar ytterligare test och det är svårt att räkna hur många de är.

Karin Bojs avslutar sin krönika med rådet att olika forskare bör samarbeta mer med professionella statistiker.

Här har jag skrivit om tolkningen av p-värden och att de lätt kan övertolkas. Man kan i empiriska vetenskaper se en sorts jakt på p-värden som är mindre än den ”magiska” gränsen 0,05. Personligen tycker jag att det är olyckligt att det har blivit så, och jag är säker på att många av forskarna som ”jagar” signifikans tycker det också. Den artikel jag har skrivit om här är verkligen inte ovanlig på det sättet. Det är egentligen väldigt orättvist att lyfta fram just den här artikeln som problematisk i sin statistiska (över-)tolkning: den befinner sig i ett stort sällskap av andra artiklar i många empiriska vetenskaper.

Den betydligt större frågan, vilka allmänna råd man ska ge till forskare och utredare som måste ta ställning till om och hur de ska justera sina p-värden i multipla test, är synnerligen intrikat. Varför det blir så besvärligt diskuterar Yudi Pawitan och Arvid Sjölander i sin spännande artikel i Qvintensen nr 2/2014.

DAN HEDLIN

Referenser

Bojs, K. (2014). Falsk matematik om farmors matvanor. Dagens Nyheter. Publicerad 2014-03-02 i DN:s webbutplaga.

Bygren, L-O., Kaati, G. och Edvinsson, S. (2001). Longevity determined by paternal ancestors’ nutrition during their slow growth period. Acta Biotheoretica, 49, 53-59.

Bygren, L-O., Tinghög, P., Carstensen, J., Edvinsson, S., Kaati, G., Pembrey, M.E. och Sjöström, M. (2014). Change in paternal grandmothers’ early food supply influenced cardiovascular mortality of the female grandchildren. BMC Genetics, 15:12. <http://www.biomedcentral.com/1471-2156/15/12>

»Läsarna är fullt kapabla att själva värdera fynden och kan förhoppningsvis stimuleras till reproduktion»

Angående transgenerationella svar på miljön och bedömningen av vår artikel om detta.

Jag tackar för att ha blivit inbjuden att kommentera vår artikel just i den här kretsen. Jag har haft mycket att göra med statistiker särskilt i mina funktioner i dödsorsaksstatistiken och tidigare i Vetenskapliga rådet vid Statistiska centralbyrån. Faktiskt var jag en gång medlem i en av era ursprungsföreningar genom förmedling av dess dåvarande ordförande, Tore Dalenius.

Bakgrund

Artikeln kommenterad av Dan Hedlin utgick från studier inom generation av nutritionsförändringar i tidig utveckling från embryo och framåt. David Barker har avhandlat långsiktiga följder av nutritionsbrister under tiden i moderlivet för vuxenlivets hjärt-kärlsjuklighet och hade tanken att moderkakans storlek kunde vara bakomliggande. Den styrs delvis av förändring av nutritionen under livet i uterus och också av prägling av genomet. Vi tog upp frågan i en studie av Demografiska databasens material i Skellefteå och kunde se att förändring av tillgången på mat i omgivningen från stor till liten och från liten till stor under tiden i moderlivet var följd av hög dödlighet i slaganfall, jämfört med när tillgången var kontinuerligt liten, medelstor eller stor. Under årens lopp har vi sedan arbetat med frågan om miljöpåverkan på spermier och ägg kan ha följder för efterkommande generationer. Den senaste är den nu kommenterade artikeln där vi går tillbaka till tanken att förändring av födans energiinnehåll kan vara främsta stimulus till de metabola förändringar som man antar ligger bakom också människans svar på förfäders tillgång på mat under tidig utveckling. Vi använde oss av de naturliga experiment som vädret orsakade under 1800-talet.

Om transgenerationella svar på miljön retiska mekanismer och forskningen på området som huvudsakligen utförs på djur och kommenterar resultaten därifrån utifrån våra studier på människor. Att skilja på exponeringens nivå och dess förändring kan nog visa sig fruktbar.

Kommentaren av Dan Hedlin

För en beskrivning av vår nu kommenterade artikel hänvisar jag till Dan Hedlins berättelse i Qvintensen i detta nummer och tar bara upp tre inslag.

Dan Hedlin har fastnat på vårt uttryck ”peculiar phenomenon”. Det var en inledning till en diskussion där traditionell genetik mötte socialvetenskap. Fenomenet kunde inte förklaras utifrån vanlig genetisk ärftlighet och inte heller med inlärning eller imitation mellan generationer utan möjligen med andra biologiska mekanismer som vi sedan resonerade om. Det var skälet till att vi lade alla kort på bordet. Vi redovisade alla punktskattningar och konfidensintervaller och den stora osäkerheten i skattningarna. Sett i relation till studierna på djur och växter var det viktigt att redovisa dessa möjliga effekter på människor. Det var ingen jakt på signifikanser. Våra reservationer genomsyrade artikeln. Det är nästan tröttsamt när jag nu läser artikeln. Kolleger i detta fält förstår det epigenetiska resonemanget och kan lätt värdera vårt möjliga belägg för icke-genetiskt transgenerationellt svar på mattillgång respektive möjligheten att det var en tillfällighet. Massignifikans nämns tyvärr inte uttryckligen men den finns att se för var och en som förstår statistik och kan göra en avvägd bedömning.

Dan Hedlins tolkning är att det var slumpmässigt genererade resultat. Även om det inte skulle ha varit det vill också vi se fler studier på människor av förändring av mattillgången eftersom djurexperimenten är övertygande. Dan Hedlins sannoliketskalkyl är ytlig och

tar upp en fördjupning, inte grundanalysen. Grundanalysen gällde könsbundet transgenerationellt svar på förändrad mattillgång. Fördjupningen var ett försök att besvara en förmodad fråga från läsaren. Hade riktningen på förändringen betydelse?

Den tabell som Dan Hedlin redovisar¹ är alltså inte rättvis i förhållande till frågan om massignifikans. Tanken var att förändring av mattillgång under tidig utveckling också kunde ha transgenerationella följder, inte bara inom generation. Det är därför vår Tabell 1² som inte tar upp riktningen på förändringen, bara förändring, bör bedömas utifrån frågan om massignifikans. Där var det naturligt att undersöka långtidseffekten av föräldrars, far- och morföräldrars förändrade mattillgång.

Varför då tabellen med uppdelning på riktning av förändringen? Också utan testning och utan signifikans har den läsvärde för initierade forskare.

Dan Hedlin tar upp en källa från 1800-talet bakom vår oberoende variabel, mattillgången. Det ska nog vara underförstått att den var osä-

ker. Men det var en mycket imponerande föregångare till er på 1800-talet, Hörstadius. Våra historiska källor var flera och en av mina medförfattare är demografisk historiker. Hörstadius visade sig ha robusta uppskattningar för stora delar av 1800-talet skördar när de ställdes mot matpriser och andra källor som Hushållningssällskapet, mötesreferat och lokalhistoriker som beskrivit skördarna.

Dan Hedlin tar också upp journalist-kommentarer men de är knappast intressanta. Karin Bojs krönika i Dagens Nyheter var ett angrepp på hela den epigenetiska forskningen på grund av Stalins förföljelser av övriga genetiker. Till omfånget ägnade hon halva krönikan åt detta. Det känns långt ifrån Överkalix.

Slutord

Artikeln är väl mottagen i fältet. Den är publicerad i en för ämnet passande tidskrift av hygglig ställning och med hygglig impact score (2.8). Läsarna är fullt kapabla att själva värdera fynden och kan förhoppningsvis stimuleras till reproduktion. Materialet är litet vilket återspeglas i

Tabell 1. Far- och morföräldrars utsatthet för drastisk förändring på tillgång till mat och dess eventuella betydelse för risken för deras barnbarn att dö i hjärt-kärlsjukdom. Tabellen visar hasardkvoter och deras konfidensintervall. (Fet stil från källan, Bygren et al. 2014)

	Män	Kvinnor
Farfar	0.87 (0.46-1.64)	0.91 (0.43-1.96)
Farmor	0.64 (0.32-1.29)	2.69 (1.05-6.92)
Morfar	1.26 (0.68-2.34)	1.32 (0.58-3.04)
Mormor	0.69 (0.35-1.36)	0.56 (0.22-1.49)

Tabell 2. Far- och morföräldrars utsatthet för drastisk förändring på tillgång till mat, uppdelad på förändringens riktning, och dess eventuella betydelse för risken för deras barnbarn att dö i hjärt-kärlsjukdom. Tabellen visar hasardkvoter och deras konfidensintervall. (Fet stil från källan, Bygren et al. 2014)

	Dålig tillgång till god		God tillgång till dålig	
	Män	Kvinnor	Män	Kvinnor
Farfar	0.90 (0.45-1.80)	1.42 (0.64-3.14)	1.08 (0.29-4.00)	0.15 (0.02-1.41)
Farmor	0.54 (0.24-1.20)	1.62 (0.58-4.52)	0.86 (0.22-3.3)	21.27 (2.24-202.47)
Morfar	1.17 (0.58-2.35)	1.22 (0.46-3.23)	1.28 (0.31-5.34)	3.37 (0.59-19.22)
Mormor	0.60 (0.24-1.49)	0.91 (0.30-2.75)	0.75 (0.22-2.58)	0.16 (0.02-1.32)

vida konfidensintervall. I experiment på växter och djur har man emellertid funnit transgenerationella svar på omgivningen motsvarande dem vi tycker oss se. Men antalet studier på människa av experimentell natur som tar upp transgenerationella miljösvår i form av sjukdomar är bara någon handfull. Poängen är att det var viktigt att redovisa analysen inte bara utifrån sannolikhetläran utan också utifrån relaterad forskning om grundläggande biologiska mekanismer. Det behövs fler sådana studier på människa.

LARS OLOV BYGREN,
KAROLINSKA INSTITUTET

Referens

Barker DJ och Clark PM. (1997). Fetal undernutrition and disease in later life. Reviews of Reproduction, 2, 105-112

Redaktörens fotnoter

1. Lars Olov Bygren syftar på Tabell 2, som återges här bredvid.
2. Tabell 1 återges här bredvid.

Replik från Karin Bojs

■ Extraordinära påståenden kräver extraordinära belägg. Några sådana belägg har inte Lars Olof Bygren levererat.

KARIN BOJS

Replik på ”Angående transgenerationella svar på miljön och bedömningen av vår artikel om detta”.

»Författarna bryter mot en viktig regel som gäller forskare: den att vara självkritisk»

■ **Lars Olov Bygren har rätt i** att jag bara tar upp artikelns Tabell 2. Han skriver att grundanalysen gällde könsbundet transgenerationellt svar på förändrad mattillgång. I Tabell 2 delas de skattade hasardkvoterna upp på förändringens riktning. I Tabell 1 spelar inte förändringens riktning någon roll, det räcker att far- och morföräldern har varit med om en drastisk förändring i tillgången på mat, vare sig det är från knapp tillgång till god eller tvärtom. Lars Olov Bygren fortsätter: ”Det är därför vår Tabell 1 som inte tar upp riktningen på förändringen, bara förändring, bör bedömas utifrån frågan om masssignifikans.”

Det är helt ok att dela upp tester i ”primära” och ”sekundära”, där frågan om masssignifikans bara avser de primära testerna. Det borde dock ha framgått tydligare i artikeln Bygren et al. (2014) att det är så testerna ska uppfattas, att det är så de tolkas i artikeln och att det var så de planerades från början.

Tabell 1 redovisar åtta hasardkvoter, varav den största är 2,69 med 95% konfidensintervall 1,05 – 6,92. Men även om frågan om masssignifikans begränsas till Tabell 1 försvinner den inte. Artikeln redovisar inte något exakt p-värde men jag skattar p-värdet för det test som motsvarar hasardkvotens konfidensintervall 1,05 – 6,92 till ungefär 0,04. Tar man hänsyn till att Tabell 1 innehåller åtta test, blir slutsatsen att artikeln inte innehåller något övertygande argument för att detta konfidensintervall utgör evidens på att hasardkvoten är signifikant skild från ett (och därmed skulle utgöra evidens för långtidseffekt av far- eller morföräldrars förändrade mattillgång).

Det jag vill kritisera i Bygren et al. (2014) är framför allt att författarna bryter mot en viktig regel som gäller forskare: den att vara självkritisk mot sina egna resultat, och då främst detta: Den uppenbara, alternativa tolkningen diskuteras inte i artikeln: att sannolikheten att få dessa testresultat om nollhypotesen vore sann (p.g.a. masssignifikans) inte alls är särskilt låg. Ändå framförs följande i abstraktet som artikelns slutsats ”Conclusion: The shock of change in food availability seems to give specific transgenerational responses.” Det står också i abstraktet: ”The confidence intervals were very wide but we found a striking TGR” (transgenerational response). Vi kan också läsa artikelns titel en gång till: Change in paternal grandmothers’ early food supply influenced cardiovascular mortality of the female grandchildren.

Jag upplever att två helt olika saker blandas ihop i frågan om man bör justera p-värden när man utför många test:

1. Bör man justera?
 2. Hur ska testvärdena tolkas?
- Naturligtvis kan en forskare välja att inte justera, men då ska hen inte heller tolka resultaten som vore de justerade.

Lars Olov Bygren skriver: ”Dan Hedlin tar också upp journalist-kommentarer men de är knappast intressanta. Karin Bojs krönika i Dagens Nyheter var ett angrepp på hela den epigenetiska forskningen på grund av Stalins förföljelser av övriga genetiker.” Jag har svårt att se hur Karin Bojs krönika kan tas som ett angrepp på hela den epigenetiska forskningen. Hon skriver t.ex. att ”epigenetik är ett mycket hett och dynamiskt fält som kan få stora konsekvenser för exempelvis cancerforskning och diabetesforskning.” Är det ett angrepp? Det hon kritiserar är övertolkning på grund av att testresultat tolkas utan hänsyn till masssignifikans.

DAN HEDLIN

Slutord i replikskiftet med Dan Hedlin om artikeln om transgenerationella svar på miljön

»Slutsatsen grundades på statistisk inferens, biologisk teori och djurförsök»

■ **Jag tackar för** Dan Hedlins inbjudan till slutord.

Han har hedervärt nyanserat sin kritik i någon mån men anser att vi har brutit mot en grundläggande regel. Vi har inte varit självkritiska mot våra egna resultat och inte diskuterat den alternativa tolkningen utifrån sannolikhetsläran.

Jag måste avvisa den kritiken. Vi visade svart på vitt svagheter t.ex. höga p-värden och vida konfidensintervall och vi skrev också explicit: ”The sample size was however small and the confidence intervals large”. Inom epidemiologien är masssignifikans-problemet välkänt men normalt redovisar man punktskattningar och konfidensintervall. Vi tror att man i debatten om artikeln möjligen underskattar det internationella vetenskapssamhällets förmåga att läsa och tolka de resultaten på ett rimligt sätt. Man kan fixera sig vid siffror och p-värden men forskare som kan något om studieobjektet väger in tidigare experimentell evidens

Bloggposten som genererade diskussionen

■ **Qvintensen har frågat** Olle Häggström, som uppmärksammade Dagens Nyheters första omnämnande av Bygren et al. (2014), om han vill skriva ett inlägg i diskussionen. Han hänvisar till de två bloggposter (Häggström 2014a och 2014b) där han diskuterar Bygren et al. (2014) samt till den diskussion som den första bloggposten genererade.

Referenser

Häggström, O. (2014a). Om statistisk signifikans, epigenetik och de norrbottniska farmödrarna. Ur bloggen Häggström hävdar, 140223. <http://haggstrom.blogspot.se/2014/02/om-statistisk-signifikans-epigenetik.html>. (Den här referensen är upprättad 140828).

Häggström, O. (2014b). Om några av reaktionerna på Bygrenincidenten. Ur bloggen Häggström hävdar, 140313. <http://haggstrom.blogspot.se/2014/03/om-nagra-av-reaktionerna-pa.html>. (Den här referensen är upprättad 140828).

och grundläggande biologiska mekanismer i tolkningen.

När man tolkar testresultat måste man förstås ta hänsyn till sannolikhetsläran och frågan om masssignifikans. Hedlin medger visserligen att det inte räcker men kritiserar att att vi skriver att det ser ut (”seems”) som att förändringen av mattillgång ger ett specifikt transgenerationellt svar. Jag avvisar också den kritiken på våra vägnar. Slutsatsen grundades på statistisk inferens, biologisk teori och djurförsök.

Till slut vill jag säga att vi hade en skyldighet mot den stora grupp forskare som arbetar med frågan om transgenerationella svar på växters och djurs miljö att publicera dessa sällsynta resultat från våra mänskliga kohorter. Att det kan komma att förklara paradoxala fynd i svältkatastrof-forskningen är ytterligare ett skäl.

LARS OLOV BYGREN

Livet som konsult

Hon borde spela basket! Förslaget fick mig att le. En förälder i min äldsta dotters klass tittade allvarligt och strängt på mig. Vad nu, skämtar hon inte? Jag hade precis försökt byta samtalsämne eftersom jag var olidligt trött på att höra att min dotter var så kortväxt och så liten för sin ålder. Nja, hon är inte särskilt intresserad av bollsport. Gymnastik och hästar är hennes grej försökte jag förklara. Basket borde hon absolut syssla med vidhöll den andra mamman. Se bara hur långa spelarna blir. Varpå jag ler ännu mer men jag möts fortfarande av denna allvarliga uppsyn. Jag avhöll mig från att kommentera att det kausala sambandet sannolikt är det omvända.

Lärdom 1: det som känns så självklart för mig behöver inte vara självklart för andra.

Sedan drygt ett år tillbaka jobbar jag som konsult. Det är mycket fascinerande och utvecklande. De flesta uppdragen är hyfsat specificerade och tydliga men det finns undantag. Vad skulle du göra om du fick ett datamaterial av en kund där specifikationen på uppdraget avslutas med ”Gör något statistiskt av det här”. Troligen skulle du börja undra över frågeställning, hypoteser, primära variabler. Varför, vad, hur....?

Ibland kan viktig information finnas i det som inte sägs och inte förmedlas till statistikern.

Precis som med pausen i musiken, luft-rummet mellan kroppens konturer i ett porträtt eller plötslig tystnad i talet. Det finns ofta värde i det som inte hörs eller syns. Som konsult gäller det att vara frågvis!

Nyligen presenterade jag resultat från en akademisk klinisk studie. Jag kände mig nöjd med mötet och kunden verkade vara helt med på noterna. Precis när vi skulle bryta upp visade jag datamaterialet forskaren skickat till mig. Forskaren sa: ”Åh, du vad jobbigt...nu kom jag på...du måste nog göra om analyserna... tyvärr. Vi brukar ju ta bort alla värden som är under detektionsgränsen för den variabeln och de har blivit kvar”. Han skrollade ner i datamängden.

”Oj, oj, det är ju minst tio observationer som alla är <0,2”. Varpå jag lugnade honom med att det var jät-tebra att han inte hade tagit bort dom. Helt rätt!

Lärdom 2: fråga om detta är alla data.

Häromdagen var det dags för ”slutprov”. En duktig pre-klinisk forskare som dessutom är mycket intresserad av statistik lämnade något motvilligt sina data till mig. Jag som utgett mig för att vara den professionella expertkonsulten kände mig nästan lite besvrad. Överlevnadsanalys var redan gjord av forskaren själv. Ma-



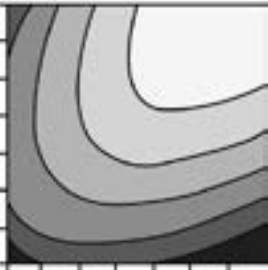
Anna Torrång.

terialiet var relativt litet som sig bör när det handlar om försöksdjur. Jag tänkte i mitt stilla sinne – vad mer kan jag göra? Lograng-testet ger nog samma resultat när jag kör det. Kaplan-Meierkurvorna kommer se likadana ut. Då ställde jag FRÅGAN. Är det här alla data? Forskaren bläddrade mellan flikarna i arket. Nja, vi har ju exkluderat de djur som dog tidigt av helt andra orsaker. Vi ska ju göra så enligt vår

försöksplan. Jag hade klarat testet. Jag hade ställt den rätta frågan och kunde därför göra nytta. Jag tillförde något när jag lät samtliga djur bidra med den tid de varit med i studien. Det som känns självklart för mig behöver inte vara självklart för andra!

Jag vill absolut inte göra narr av våra kunder. De är alla väldigt duktiga och engagerade i det de håller på med. Alla kan tänka fel ibland – även jag. Det är ofta i mötet mellan människor, i mötet mellan olika experter som det stora sker. De viktigaste lärdomarna får vi i mötet mellan tillämpning och metod. Det är just det som gör statistikeryrket så svårt men också så fantastiskt spännande och intressant!

ANNA TORRÅNG, SCANDINAVIAN DEVELOPMENT SERVICES



Det är viktigt för industrin att kunna erbjuda sina kunder kvalitetsprodukter till ett lågt pris. Därför är det angeläget att ha goda kunskaper om sin tillverkningsprocess och hålla nere kostnaderna för produktionen. Det kan uppnås med hjälp av statistisk försöksplanering, en kraftfull metodik, som vid flera tillfällen har beskrivits på Föreningen industriell statistik webbinarier samt i Qvintensen nummer 2/2009 av Magnus Arnér.

En bra lösning för processtyrning kräver kunskaper i industriell statistik och ett lyckat resultat ger processtabilitet och produktkvalitet och därmed ekonomiska fördelar. Försöksplanering används i viss omfattning inom industrin för att identifiera viktiga faktorer som påverkar processen, optimera systemet samt för att skapa robusta produkter och processer. Metodiken ger uppenbara konkurrensfördelar och därför är det angeläget att öka och popularisera kunskaper om användningen.

Erfarenheter visar att industrin tjänar på att tillämpa statistisk försöksplanering. Man får en bredare kunskap om processen som gör det lättare att förbättra och styra processen till det optimala. För vissa företag innebär detta att produktionspersonal är duktiga på statistik, för andra företag att man dessutom har välutbildade statistiker. Men det räcker inte. Man måste även ha insiktsfulla mottagare av informationen som är genererad av statistik.

I min verksamhet som rådgivare för ledningsgrupper ser jag ofta behovet av statistikutbildning. Kunskaperna är nödvändiga för att kunna fatta rätt beslut och bedöma om till exempel analys av marknadsdata är rätt utförd och därmed tillförlitlig. Att kunna sortera, filtrera, bearbeta och förstå sambanden mellan rapporterade uppgifter, som verksamhetssystemen levererar, upplevs av många som svårt. Statistisk analys har en strategisk betydelse eftersom det leder till förståelse av just detta, och därmed utgör resultaten av analyserna viktiga beslutsunderlag.

Ledningsgrupper för både små och stora företag samlas till möten, men då finns sällan tid för en intern utbildning. En framgångsrik lösning kan vara att erbjuda ledningsgrupper intressanta och givande kurser i statistik som skapar en berikande upplevelse utan formler

Statistisk vinprovning



eller Power Point-bilder. Populärast är kurser som lär deltagarna någonting nytt och socialt användbart. Här tänker jag på en introduktion till försöksplanering med hjälp av vinprovning, där deltagarna lär sig varför de tycker om vissa viner och hur de kan identifiera de faktorer som mest påverkar deras smakupplevelse.

Självklart görs alla vinförsök i samstämmighet med företagets gällande alkoholpolicy. Det positiva med provningen är att man deltar i ett socialt event samtidigt som man lär sig principer för försöksplanering, utan dator och bara med hjälp av papper och penna. Målet är inte att deltagarna ska lära sig de allra bästa och mest avancerade metoderna eller tillägna sig djupa insikter i statistisk teori. Snarare är

målet att de ska se kraften och nyttan i statistiska metoder. Och ha roligt under tiden.

Försöket organiseras med hjälp av viner från Systembolaget där varje flaska har en beskrivning av faktorerna fyllighet, strävhet och fruktsyra i form av tre klockor som i Figur 1.

1884 Reservado Pinot Noir (nr 6561) Rött vin, Argentina, Cuyo, Mendoza Årgång 2012



Figur 1. Klockan för fyllighet står på 5, strävhet på 4 och fruktsyra på 10.

Tabell 2. En tablå över responsen för deltagaren i Figur 3.

Testnivå								Testnivå				Samverkan			
Vin nr	Fyllighet	Strävhet	Fruktsyra	Re-spons	Fyllighet	Strävhet	Fruktsyra	F*S	F*Fr	S*Fr	F*S*	Fyll-Str	Fyll-Frukt	Str-Frukt	Fyll-Str-Frukt
1	1	1	1	10	10	10	10	1	1	1	1	10	10	10	10
2	-1	1	1	5	-5	5	5	-1	-1	1	-1	-5	-5	5	-5
3	1	-1	1	8	8	-8	8	-1	1	-1	-1	-8	8	-8	-8
4	-1	-1	1	5	-5	-5	5	1	-1	-1	1	5	-5	-5	5
5	1	1	-1	9	9	9	-9	1	-1	-1	-1	9	-9	-9	-9
6	-1	1	-1	2	-2	2	-2	-1	1	-1	1	-2	2	-2	2
7	1	-1	-1	3	3	-3	-3	-1	-1	1	1	-3	-3	3	3
8	-1	-1	-1	1	-1	-1	-1	1	1	1	-1	1	1	1	-1
Medelvärde				5,4	2,1	1,1	1,6					0,9	-0,1	-0,6	-0,4

Vin	Fyllighet - A	Strävhet - B	Fruktsyra - C
Vin nr 1	+1	+1	+1
Vin nr 2	-1	+1	+1
Vin nr 3	+1	-1	+1
Vin nr 4	-1	-1	+1
Vin nr 5	+1	+1	-1
Vin nr 6	-1	+1	-1
Vin nr 7	+1	-1	-1
Vin nr 8	-1	-1	-1

Tabell 1. Åtta viners stiliserade egenskaper.

Det finns tre faktorer som kan variera mellan 0 och 12 och som vi kan testa på två nivåer: "lite", motsvarande -1 och "mycket" motsvarande +1.

Följaktligen behöver vi välja och köpa 2³ = 8 stycken vinflaskor med egenskaper som i Tabell 1.

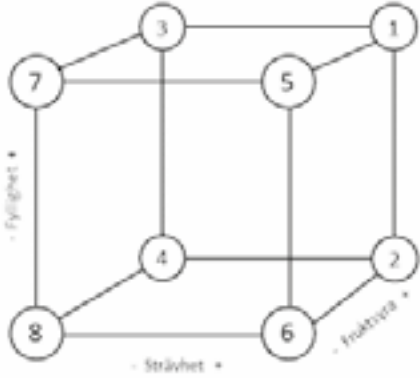
Beroende på egenskaper och testnivå kan vinerna placeras i en kub som i Figur 2, där siffrorna anger vinets nummer.

Nästa steg i försöket är att låta deltagarna smaka på samtliga viner och skatta hur nöjda de är med smaken. Till skattningen används en skala från 1 = "inte gott" till 10 = "mycket gott". Vinerna är slumpmässigt placerade på bordet (randomisering) för att förebygga risken för systematiska fel. Deltagarna vet bara vilket vinnummer de avsmakar men ser inte klockorna, det vill säga uppgifterna om egenskaperna

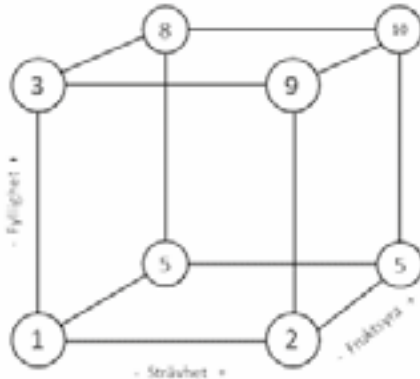
fyllighet - A, strävhet - B och fruktsyra - C.

Efter flera avsmakningar, gärna med olika tilltugg, visas resultatet för var och en av deltagarna i form av en kub som i Figur 3. Varje hörn visar medelvärdet av deltagarens skattning av nöjdhet med vinet. Deltagaren har smakat på varje vin flera gånger, och för till exempel vin nummer 3 har deltagaren i genomsnitt satt en åtta på en tiogradig skala.

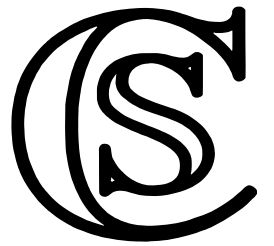
Nästa steg i försöket är att räkna fram effekten av till exempel fyllighet. Effekten blir i figuren skillnaden mellan värdena i kubens nedre och övre plan. Det räknas fram som ett medelvärde för "fyllighet på hög nivå", (3 + 9 + 8 + 10)/4 = 7,5 minus ett medelvärde för "fyllighet på låg nivå", (1 + 5 + 2 + 5)/4 = 3,25. Skillnaden mellan "hög" och "låg" nivå är effekten av vinets fyllighet, d.v.s. 7,5 - 3,25 = 4,25.



Figur 2. Vinernas egenskaper åskådliggjorda i ett diagram. Vin nummer 1 är fylligt och strävt med mycket syra.



Figur 3. En deltagares genomsnittliga upplevelse på en tiogradig skala (10 är bäst). Den här deltagaren gillade bäst det sträva och fylliga vinet med mycket syra.



»»» STATISTISK VINPROVNING

Motsvarande effekt av strävhet blir $(2+5+9+10)/4 - (1+5+8+3)/4 = 2,25$ och för fruktsyra $(5+5+8+10)/4 - (1+2+3+9)/4 = 3,25$.

Den här deltagarens "nöjdhet" verkar påverkas mest av vinets fyllighet och därefter av fruktsyra respektive strävhet. Intressanta är också slutsatser om faktornas eventuella samverkan, d.v.s. interaktion mellan (fyllighet * strävhet), (fyllighet * fruktsyra), (strävhet * fruktsyra) och (strävhet * fyllighet * fruktsyra). Tabell 2 redovisar resultaten.

Sammanfattningsvis kan vi konstatera att effekten av samverkan, d.v.s. interaktionen, för (fyllighet * strävhet) är $2*0,9=1,8$, (fyllighet * fruktsyra) $2*(-0,1)=-0,2$, (strävhet * fruktsyra) $2*(-0,6)=-1,2$ och (strävhet * fyllighet * fruktsyra) $2*(-0,4)=-0,8$.

Deltagarens "nöjdhet" med valda viner kan också beskrivas som en ekvation där vi använder koefficienterna från Tabell 2:

"Nöjdhet" = $5,4(\text{konstant}) + 2,1* \text{fyllighet} + 1,1* \text{strävhet} + 1,6* \text{fruktsyra} + 0,9*(\text{fyllighet-strävhet}) - 0,1*(\text{fyllighet-fruktsyra}) - 0,6*(\text{strävhet-fruktsyra}) - 0,4*(\text{strävhet-fyllighet-fruktsyra})$

Vi kan således rekommendera den här deltagaren att köpa huvudsakligen fylliga viner som har lägre värden för strävhet än för fruktsyra. Samverkan mellan fyllighet och strävhet kan ha betydelse för hur "extra" nöjd deltagaren blir med vinets smak. Samverkan mellan strävhet och fruktsyra kan försämra smakupplevelsen. Men det klart, deltagaren kan ju även välja att köpa ett vin som passar någon annans smak.

Exemplet med kuben som vi använder här är ett specifikt fall av försöksplanering. Syftet är att illustrera ett trevligt sätt att introducera

»En framgångsrik lösning kan vara att erbjuda ledningsgrupper intressanta och givande kurser i statistik som skapar en berikande upplevelse utan formler eller Power Point-bilder. Populärast är kurser som lär deltagarna någonting nytt och socialt användbart.»

fortsätta med diskussioner av validitet och slump. Om de skulle upprepa precis samma experiment med samma viner nästa vecka, får de knappast exakt samma resultat. Troligen avviker deras bedömning något från den bedömning de nyss gjort. Hur kan de statistiska metoderna inkorporera sådan "slumpmässig" variation?

Vad tror deltagarna att experimentet säger om ett antal andra viner med samma nivåer av fyllighet, strävhet och syra? De kommer inte att smaka exakt likadant som de viner de nyss provat.

Min erfarenhet är att deltagarna uppskattar statistisk vinprovning eftersom de får goda kunskaper om de egna preferenserna vad gäller valet av viner. De får också en bra förklaring till varför de tycker om specifika viner och insikt om nytta av statistisk metodik vid identifiering av påverkande, styrande faktorer.



GÖRAN LANDE

FOTO: LINA LANDE

■ Hadley Wickham visar ggvis i den här videon: <http://vimeo.com/94184685> (googla på hadley wickham ggvis så hittar du den). De som är intresserade av visualisering av data och explorativ dataanalys kan också ta del av Hadley Wickhams doktorsavhandling: <http://had.co.nz/thesis/>.

Att lära sig visualisering och presentation

Cramérsällskapets sommarskola med temat *Scientific Visualization and Presentation* lockade drygt femtio deltagare till Falun i mitten av juni.

Deltagarna gavs möjlighet att lära sig mycket om visualisering och vetenskaplig presentation. Ett annat viktigt syfte med sommarskolan var att ge ett tillfälle för doktorander från olika universitet att träffas och presentera sin forskning för varandra. Jag hoppas och tror att de sjutton doktorander som deltog i sommarskolan tyckte att upplägget med "presentationer av doktorander för doktorander" var lyckat. Det gör i alla fall jag och de andra i organisationskommittén, som imponerades både av presentationerna och av de efterföljande kommentarerna, frågorna och förslagen på förbättringar från åhörarna.

De fyra inbjudna talarna, Tom Britton, Allan Gut, Jo Røislien och Hadley Wickham, gav också högt uppskattade presentationer. Det jag personligen uppskattade allra mest var den tjuvtitt på R-paketet *ggvis* som Hadley Wickham bjöd oss på, Jo Røisliens noggrant utvalda och vackra bilder, Allan Guts olika exempel på hur fel det ibland kan bli när journalister sätter rubriker och, sist men inte minst, erfarenheterna av att vara redaktör för en vetenskaplig tidskrift som Tom Britton delade med oss. Hur ska man tänka och göra när man skriver en kollegial bedömning (peer review) av en artikel och hur ska man bemöta kritik på sina egna artiklar?

På Cramérsällskapets webbplats finns mer information om sommarskolan och ett fotogalleri.

MARIA KARLSSON



Sylvia borin – hur intressant är det?

ORDFÖRANDEN HAR ORDET

Vad är vitsen med sannolikhetslära?

skrivande stund inleder vi ett nytt akademiskt år. De flesta studenter kommer tillbaka efter sommaren för att fortsätta sina studier, men runt en fjärdedel av studenterna är nybörjare. De försöker nu greppa högskolestudiernas kodex och anpassa sitt beteende till den. Det är svårt och förmodligen något skrämmande för många av dem.

En liten, men inte försumbar, grupp av studenterna stöter nu på begreppet sannolikhetslära för första gången i livet. Det är inget de valt att komma i kontakt med, det råkade bara vara medtaget i studieplanen för något mera begripligt som ett ingenjörsprogram eller en ekonomutbildning.

Vilken erfarenhet från livet och vilka abstraktioner ska studenten associera till när hen försöker omfatta begrepp som sannolikhet, slumpvariabel och centrala gränsvärdessatsen? Associationerna är nog mycket få för lejonparten av studenterna och sannolikhetsläran ter sig svårbegriplig och onödig för dem. Svårbegriplig för att associationer saknas och onödig för att det till synes gått utmärkt att leva livets första fjärdedel utan kännedom om sannolikhetslära.

Så vad är vitsen med sannolikhetslära? Studenten har synnerligen liten chans att få svar på den frågan om hen läser inledningen i en

kursbok i sannolikhetslära. En högrävaning mening av slaget – *Sannolikhetsläran är fundamentalt integrerad i vår vardag* – följt av ytterligare ett par floskelspäckade meningar övergår snabbt till en redogörelse för utfallsrum följt av satser och definitioner i en strid ström. En ornitolog kan rätteligen säga till mig – *Fåglar är fundamentalt integrerade i vår vardag* – men jag förblir likafullt kallsinnig till fåglar. Och än mer iskallt blir mitt sinne om jag förevisas bild efter bild på olika fåglar med olika löjligen, svärmemorerade namn och därtill får veta att någon kuf hittat på att fågelskrällena ska ha ytterligare namn, på latin!

Jag ska snart träffa en ny grupp studenter och prata om sannolikhetslära. Då kommer jag att hålla i minnet att en del av dem har samma känslor inför detta ämne som jag har för fåglar.

Jag kommer att tillstå att sannolikhetsläran är svårbegriplig. Men jag kommer att insistera på att den är behövlig. Jag tänker påstå att sannolikhetsläran håller oss på vägen under livets krokiga resa, både privat och i arbetsrollen. Jag kommer att avsätta flera dagar åt att övertyga studenterna

»Den vetenskapliga evidensen för att vi människor är urusla på att intuitivt välja smart bland olika alternativ är överväldigande.»

om påståendets giltighet, och jag kommer att gödsla med sedelärande exempel hämtade från Daniel Kahnemans bästsäljare *Tänka, snabbt och långsamt* (Thinking, fast and slow i originalets titel). Den vetenskapliga evidensen för att vi människor är urusla på att intuitivt välja smart bland olika alternativ är överväldigande. Vitsen med sannolikhetslära är att dess formalitet ger oss människor ett redskap för att fatta smarta val. Ingen tackar nej till ett sådant kraftfullt redskap, och min, och kanske din, tacknämliga uppgift i höst är att få visa det redskapet till en ny grupp unga människor.

Avslutningsvis vill jag tacka Lars Rönnegård, Leif Ruckman och Maria Karlsson i Cramérsällskapets styrelse för en välorganiserad och lyckad sommarskola. Maria skriver om sommarskolan här bredvid.



KENNETH CARLING



ORDFÖRANDEN HAR ORDET

Ständigt dessa siffror!

Tidigare i den här spalten har problematiken kring redovisningen av felmarginaler för opinionsundersökningar tagits upp och nu är det dags för en modest skärskådning av hur massmedia presenterar resultat från sådana undersökningar.

Den 16 augusti publicerade Göteborgs-Posten och Svenska Dagbladet TNS Sifo:s väljarbarometer för mätperioden 11-14 augusti. I pappersversionen av Göteborgs-Posten presenterades resultaten för de olika partierna med de sedvanliga staplarna, tillsammans med angivelse av hur stora uppgångar/nedgångar de haft jämfört med tidigare mätning. För att få reda på felmarginaler hänvisades finstilt till Göteborg-Postens nätupplaga. Om procenttalen summeras fås ett väljarstöd som uppgår till 100,2% för de åtta riksdagspartierna plus kategorin som inkluderar bland andra Feministiskt initiativ. Detta är förstas en effekt av adderade avrundningsfel, men läsaren kan ju här också få uppfattningen att alla 1254 intervjuade har uppgett partisympati.

För att få mer fullständig information behöver man vända sig till TNS Sifo:s webbsida och där finner man att av de 1254 intervjuade har 13% uppgett att de ska rösta blankt eller ej uppgett parti. I det här fallet är det enbart Socialdemokraterna och Moderaterna som får ett högre väljarstöd än denna grupp potentiella väljare. Utan den här informationen kan man lätt få intrycket av ett renodlat nollsummespel, där partierna enbart måste erövrast röster från varandra. Någon skulle här kunna hävda att resultatet återspeglar valsituationen ganska väl: Cirka 87% av valmanskåren kommer att rösta

och resterande 13% i undersökningen motsvarar ungefär ”soffliggarna”. Mot detta kan man anföra att vi inte har någon information om bortfallet. Troligen är det betydande och om ”bortfallspersonernas” inställning till frågan “Vilket parti skulle du rösta på i riksdagsvalet om det var val i dag?” vet vi ingenting!

Förändringar i förhållande till föregående mätning är också av stort intresse och i den i Göteborgs-Posten till väljarbarometern vidhäftande artikeln får man leta länge tills man hittar uppgiften att det är juni-mätningen som man jämför med. I den mätningen var det för övrigt 1922 intervjuade och i maj-mätningen 2888.

Sammanfattningsvis måste konstateras att relevanta fakta skulle kunna presenteras på ett tydligare sätt än vad som är fallet i Göteborgs-Postens redovisning.

Opinionsundersökningarna kommer tätt. De görs av olika företag och redovisas i tidningar på olika sätt. Den 20 augusti publicerade Dagens Nyheter IPSO:s opinionsundersökning (mätperiod 7-18 augusti) och där redovisades att 1252 av 1512 respondenter angav partisympati, dvs 17,2% gjorde ej det. I undersökningen fick till exempel Moderaterna stöd av 20,8%, vilket omräknat betyder att av samtliga intervjuade sympatiserade 17,2% med detta parti. Andelen bland alla intervjuade utan angiven partisympati är således här precis lika hög som andelen moderatanhängare. Detta framgår dock tyvärr ej uttryckligen i artikeln, utan var något som jag själv måste räkna fram. Detta trots att det bärande temat i artikeln är den stora andelen osäkra väljare. I Göteborgs-Postens redovisning 16 augusti har Moderaterna stöd av 24,6% av de intervjuade, men då har som sagt de som inte

uppgett någon partisympati undantagits från procentberäkningen.

I **TNS Sifo:s väljarbarometer** intervjuades alltså personer från 11 till 14 augusti. I detta nummer av Qvintensen nämns tekniken med sammanvägda skattningar från olika undersökningar och då kommer oundvikligen också ett större mått av tidsdynamik in i bilden. Skolexemplet av en urvalssituation är att man drar ett urval från en fix population vid en viss given tidpunkt. Sammanvägningar av flera undersökningars resultat torde innebära att man behöver sammanföra resultat från mät- och referensperioder som inte helt sammanfaller och då behöver frågan ställas: Vad mäts? eller omformulerat: Vad är det man skattar? Vidare: Om det tydligt är angivet vad som skattas, finns det tillräckligt med information från de olika undersökningarna för att åstadkomma en relevant sammanvägd skattning?

För att inte den här spalten ska få rykte om sig att vara skriven av en riktig gnällspik, är det här på sin plats att konstatera att till exempel TNS Sifos väljarbarometer uppvisar väldigt stabila och trovärdiga skattningar av väljar-sympatier över tid och visar kanske därmed att tillämpad urvalsteori fungerar trots alla praktiska bekymmer!



PER GÖSTA
ANDERSSON

Bästa uppsats

■ Surveyföreningens pris för bästa uppsats på surveyområdet vanns i år av Andrea Gustafsen (till vänster) och Bergdís Sigurjónsdóttir som skrev kandidatuppsats på Statistiska institutionen vid Stockholms universitet. Uppsatsen har titeln Coding Process and Quality Control: Literature overview and the Case of Statistics Sweden. Här skriver de om sitt arbete.



Att klassificera ord med siffror

Uppsatsen “Coding Process and Quality Control. Literature Overview and the Case of Statistics Sweden” är en litteraturstudie över ämnet kodning och kodningskvalitet samt en analys av kodningsfelen i Statistiska centralbyråns Arbetskraftsundersökning 2012. Uppsatsen vann Surveyföreningens pris för bästa uppsats inom surveyområdet. Här skriver författarna Andrea Gustafsen och Bergdís Sigurjónsdóttir om sin uppsats som de lade fram 2013 på Statistiska institutionen vid Stockholms universitet.

En survey kan innehålla en fråga som ”Vad arbetar du med?”. Respondenten besvarar den med egna ord. Sedan klassificerar undersökaren svaret på något sätt och det kallas kodning.

Kodning innebär att svar på öppna svarsalternativ tilldelas en specifik kod enligt ett kodschema eller en klassifikationsstandard. Kodningsprocessen är kostsam, tidskrävande och felbenägen. Kodningsfel försämrar datakvaliteten och påverkar de statistiska resultaten. Litteraturen om kodning är mycket begränsad. De studier som har gjorts har dock påvisat höga felfrekvenser som resultat av kodningsprocessen.

Kodning kan utföras manuellt, helt automatiserat eller med datorstöd. Manuell kodning utförs utan hjälp av datorer eller mjukvara. Den kan genomföras inom en organisation av mer eller mindre kodningsutbildad personal eller utanför organisationen, till exempel av

personer som genomför intervjuer. Detta är en av de mest kostsamma delarna i en surveyundersökning. Automatisk kodning utförs av mjukvara som försöker matcha respondentens svar med ett databaserat lexikon. Vi kan få exakt matchning när svaret är identiskt med definitionen enligt lexikon eller inexakt matchning när svaret är tillfredsställande nära. Datorstödd kodning är en blandning av manuell och automatisk kodning där kodaren använder en kodningsmjukvara. Kodaren skriver in respondentens svar i mjukvaran och får då en lista med förslag över möjliga koder. Den slutgiltiga koden är ett resultat av samverkan mellan kodaren och mjukvaran.

Vid kvalitetskontroll av kodningsprocessen används olika mått på kodningsfel. Det vanligaste måttet är felfrekvensen som är kvoten mellan antalet felkodade objekt och totala antalet kodade objekt. Felfrekvensen kan presente-

»»»

Tabell 1. Kodning av uppgiftslämnarnas socioekonomiska klass.

SEI, 1-digit level	Original code								
Correct code	1	2	3	4	5	7	8	B	Total
1 Unskilled manual workers	110	7	4	1	1	1	0	1	125
2 Skilled manual workers	2	80	0	3	0	0	0	0	85
3 Lower managerial, professionals	5	5	144	12	1	0	0	1	168
4 Intermediate managerial, professionals	2	2	13	261	15	0	0	4	297
5 Higher managerial, professionals	0	0	2	17	198	0	0	3	220
7 Self-employed, excl. farmers	0	0	0	1	0	2	0	0	3
8 Farmers	0	0	0	0	0	0	3	0	3
b Missing values	2	1	5	1	3	0	0	36	48
Original codes	121	95	168	296	218	3	3	45	949
Correct codes	125	85	168	297	220	3	3	48	949
Originally correct	110	80	144	261	198	2	3	36	834
Proportion correct	0,91	0,84	0,86	0,88	0,91	0,67	1,00	0,80	0,88
Gross error	26	20	48	71	42	2	0	21	115
Gross error %	21,5	21,1	28,6	24,0	19,3	66,7	0,0	46,7	12,1
Net error	-4	10	0	-1	-2	0	0	-3	.
Net error %	-3,3	10,5	0,0	-0,3	-0,9	0,0	0,0	-6,7	.

»»» ATT
KLASSIFICERA
ORD MED
SIFFROR

ras brutto och netto. Bruttofelet för en specifik kategori eller kod är antalet objekt som felaktigt inkluderats i kategorin plus antalet objekt som felaktigt exkluderats. Nettofelet är antalet felaktigt inkluderade objekt minus antalet felaktigt exkluderade objekt. Nettofelet är oftast betydligt mindre än bruttofelet då felklassifikationerna tenderar att ta ut varandra. Felfrekvenser kan beräknas på olika sätt såsom separata frekvenser för varje kodare eller för varje specifik kod eller för olika nivåer i hierarkiskt ordnade kodsyst

Kodningsfelens konsekvenser skiljer sig åt beroende på om man använder aggregerade data eller ej. Aggregerade data kallas ofta makrodata medan mikrodata avser data på individnivå. Vid användning av makrodata är kodningsfel inte ett påtagligt problem eftersom nettofelet ofta är små. Kodningsfelen är mer allvarliga vid användning av mikrodata eftersom objekt som felaktigt inkluderats eller exkluderats i en kategori används för vidare analys. Kodningsfelen påverkar även precisionen av estimaten

då de ökar variansen för estimaten för specifika kodningskategorier.

Kodningsprocessen kan kontrolleras genom beroende eller oberoende verifikation. Vid beroende verifikation inspekterar en andrakodare den kod som en förstakodare har valt. Om andrakodaren bedömer att den ursprungliga koden är felaktig ändrar han eller hon koden. Beroende verifikation har visat sig vara mycket ineffektivt då endast en liten del av de felaktiga koderna ändras av andrakodaren. Istället förespråkas oberoende verifikation där samma objekt kodas av två eller fler kodare som inte har vetskap om hur de andra har kodat. Om koden skiljer sig åt mellan olika kodare lämnas objektet till reconciliering där ytterligare en kodare avgör vilken den korrekta koden är.

2008 beslutade generaldirektören för Statistiska centralbyrån, SCB, att myndigheten skulle arbeta mot en certifiering (uppnådd 2014) enligt den internationella standarden ISO 2052. SCB har därför infört kontroller och förbättrat

kvaliteten av kodningen för att uppfylla kraven. Standardrutinen som används är oberoende verifikation. SCB presenterar kodningsfelen dels genom andelen fel för enskilda kodare, dels genom beräkningar av brutto- och nettofel.

I vår uppsats presenterar vi beräkningar av kodningsfel för Arbetskraftsundersökningen 2012. De totala bruttofelen för socioekonomisk indelning SEI, standard för svensk yrkesklassificering SSYK och International Standard Classification of Occupations ISCO-08 var mellan 10 och 16%. Andelen fel ökade inte markant vid fler siffernivåer i koden, alltså när kodarna satte koder enligt en finare klassificering i ett hierarkiskt ordnat kodsyst

I Tabell 1 visas som ett exempel beräkningar för socioekonomisk indelning, SEI, på ensifernivå. I den första omgången kodades 296 individer som kod 4 fastän 297 tillhörde kod 4. Av dessa 296 individer blev 261 rätt kodade. 35 individer kodades felaktigt som kod 4 och 36 individer som tillhörde kod 4 gavs felaktigt



Figur 1. Andel rätt kodade individer för elva kodare (A, B, ... K), uppdelade på den grövsta indelningen (ensifernivå; den vänstra stapeln i varje par) och en finare indelning (tvåsifternivå; den högra stapeln i varje par).

andra koder. Andelen individer som kodats rätt är 88%. Bruttofelet är 24%, men nettofelet -0,3%. Detta på grund av att antalet kodade med kod 4 var nära det sanna antalet.

Andel rätt kodade poster för de individuella kodarna låg på mellan 80 och 95%. Felfrekvenserna var likartade för enskilda kodare för olika klassifikationer samt siffernivåer. En kodare kodade exempelvis rätt i 80% av posterna i samtliga klassifikationer. I Figur 1 visas andelen rätt kodade poster för individuella kodare för SEI, på en- och tvåsifternivå (med tvåsifternivå menas att t.ex. kod 1 är uppdelad i finare grupper, vars koder har två siffror).

Förslag till fortsatta förbättringar av kvaliteten i kodningen hos SCB är att ännu mera utnyttja framtagen information och analys om kodningsfelen, i syfte att eliminera felkällor i kodningsprocessen. Kodarna bör mera regelbundet få återkoppling samt involveras i processen att hitta och åtgärda felkällorna.

ANDREA GUSTAFSEN OCH
BERGDIS SIGURJÓNSDÓTTIR



– 2122.

Yrkesområde 2 – Yrken med krav på fördjupad högskolekompetens

Huvudgrupp 21 inom yrkesområde 2. Yrken med krav på fördjupad högskolekompetens inom naturvetenskap och teknik

- I arbetet ingår att:
- utveckla teorier, modeller och metoder samt utforma praktiska tillämpningar av dessa inom fysik, kemi, matematik, statistik, ingenjörsvetenskap, farmakologi, miljö samt inom jord- och skogsbruk
 - studera fenomen, analysera och värdera data
 - ge råd och utfärda anvisningar
 - rita byggnader m.m. och leda bygg- och anläggningsprojekt
 - planera och kontrollera utvinnings- och tillverkningsprocesser
 - designa och utveckla produkter
 - kontrollera att hälso- och miljöskyddsföreskrifter efterlevs.

Yrkesgrupp 212 Matematiker, aktuarier och statistiker

Utvecklar och tillämpar matematiska och statistiska teorier, modeller och metoder. Utvecklar beräkningsmodeller och löser beräkningsproblem. Utformar pensions- och andra försäkringssystem m.m. Planerar, utvecklar, bearbetar, framställer, analyserar och presenterar statistiska undersökningar.

Undergrupp 2121: matematiker och aktuarier

Utvecklar och tillämpar matematiska teorier, modeller och metoder. Utvecklar beräkningsmodeller och löser beräkningsproblem. Utformar och förvaltar premie-, pensions- och andra försäkringssystem m.m.

- Hit hör bland annat:
- Försäkringsaktuarie
 - Hit hör inte:
 - Underwriter (3321)

Undergrupp 2122: statistiker

Utvecklar och tillämpar statistiska teorier, modeller och metoder. Planerar, utvecklar, bearbetar, framställer, analyserar och presenterar statistiska undersökningar. Gör prognoser. Bearbetar administrativa uppgifter till statistik.

Källa: SCB, MIS 2012:1, SSYK 2012 Standard för svensk yrkesklassificering. Tillgänglig på http://www.scb.se/sv_/Hitta-statistik/Publiceringskalender/Visa-detaljrad-information/?Publobjid=18597 (googla på yrkesklassificering 2012).

ORDFÖRANDEN
HAR ORDETLäkemedels-
industrin och de sju
dödssynderna

När detta skrivs har sommaren så smått börjat avrundas efter en av de hetaste somrarna i mannaminne. Skolorna har precis satt igång och i skogarna i Västmanland har man äntligen fått kontroll på den stora skogsbranden, en brand som med sin rök som brett ut sig över landet fått oss att påminnas om vår sårbarhet. När detta kommer i tryck är hösten kommen och sommarens värme ett minne blott, höstlovet börjar närma sig, men i skogarna i Västmanland pyr sannolikt fortfarande elden. Det känns onekligen en aning egendomligt att skriva med ett sådant tidsperspektiv mellan datum för inlämning och färdig produkt.

Som en följd av detta är de två höstevenemangen Statistikerträffen i Göteborg och Höstmötet i Köpenhamn redan passerade då detta nummer av Qvintensen kommer ut, men när detta skrivs är det ännu mer än en månad kvar. Därför kan jag i skrivande stund varken göra reklam för dessa arrangemang, eller ge några omdömen om dem, men i nästa nummer kan vi kanske säga något om dem.

En kort rapport från styrelsen i FMS är att vi

så smått kommit igång. Personligen har jag alltid tyckt det är en utmaning att få till ett fungerande möte då man enbart hörs över telefon, framför allt om man aldrig tidigare träffats fysiskt. Förhoppningsvis får vi möjlighet att ses i samband med något av höstens evenemang. I förra numret beskrev jag en förhoppning om att kunna se över medlemsstatistiken, det är en av de punkter styrelsen för närvarande arbetar med. Bland annat avser vi att skicka ut en kortare medlemsenkät som vi hoppas alla FMS medlemmar svarar på. Om du inte fått någon e-post från FMS om detta kan det vara så att vi har fel e-postadress, och då föreslår jag att du kontaktar oss och meddelar den korrekta så att du i fortsättningen får information från oss. Syftet med enkäten är att få en bättre kännedom om var våra medlemmar finns geografiskt och inom vilka områden de arbetar. Detta för att bättre kunna anpassa vår verksamhet till medlemmarnas behov.

Då jag blev vald till ordförande för FMS visste jag, med de förväntningar som mina skrivskickliga företrädare skapat, att det kom på min lott att skriva denna spalt. Som en synnerligen ovan skribent tänkte jag att det kunde vara en bra idé att hänga upp varje spalt på ett särskilt tema. Då

jag räknade med att det skulle bli åtta spalter att skriva (förutsatt att jag blir omvald), tänkte jag mig sju sammanhängande ämnen och en summerande/förklarande spalt. Efter att ha funderat på ämnen som de sju underverken, dvärgarna, veckodagarna och haven, låg det närmast till att jag skulle välja att koppla varje spalt till någon av de sju dödssynderna (högmod, girighet, vållust, avund, frosseri, vrede samt lättja). Nu kom det aldrig längre än till en tanke då jag i slutändan inte kunde komma på hur jag skulle koppla dessa synder till medicinsk statistik.

När jag började jobba inom området (medicinsk forskning/läkemedelsutveckling) hade jag den idealistiska tanken att all utveckling sker med den enskilda patientens välgång som huvudmål. Med tiden har jag dock insett att så inte alltid är fallet. Visst är det så att läkemedel tas fram med syfte att gagna en enskild patient, men det läkemedel som inte kommer kunna betala tillbaka utvecklingskostnaderna kommer aldrig att komma så långt i utvecklingsprocessen att det till slut gynnar patienterna. Detta kan jag dock acceptera, för troligen är alternativet att läkemedel inte utvecklas alls.

Under våren har vi dock i media fått följa

Girighet – en av de sju dödssynderna.

hur ett stort läkemedelsföretag har försökt köpa upp ett annat. Då det huvudsakliga syftet varit att minimera den skatt man skall behöva betala, utan minsta tanke på att gynna patienterna, kan jag inte annat än konstatera att det trots allt går att hitta lämpliga kopplingar till åtminstone någon av dödssynderna.

För att avsluta med något lite mer avslappnat så läste jag på Twitter häromdagen: "Statistically, you are more likely to be killed by a statistician than by a shark". Nu visade det sig att det kom från en högst osäker källa, och syftet var att det skulle vara humoristiskt snarare än korrekt, men meningar som innehåller både statistiker och haj är alltid roliga att lyfta fram.



MARCUS THURESSON



Matias Quiroz.



Yuli Liang.

Två unga statistiker får pris för bästa presentation

Young Scientists Awards, for best oral presentation

■ Konferensen International Conference on Trends and Perspectives in Linear Statistical Inference (LinStat) 2014 i Linköping utsåg två pristagare för bästa presentation:

YULI LIANG

Hypothesis Testing in Multilevel Models with Block Circular Covariance Structures

MATIAS QUIROZ

Speeding up MCMC by Efficient Subsampling

Disputationer

2014

28 februari **JONAS WALLIN** *Stochastic Models Involving Second Order Lévy Motions*, matematisk statistik, Lunds universitet. Opponent: Håvard Rue, Norges teknisk-naturvitenskapelige universitet, Trondheim.

16 april **MEHRDAD JAFARI MAGHANI** *A Treatise on Measuring Wiener-Granger Causality*, matematisk statistik, Stockholms universitet. Opponent: Sebastiano Stramaglia, University of Bari.

14 maj **PIERRE NYQUIST** *On large deviations and design of efficient importance sampling algorithms*, tillämpad matematik och beräkningsmatematik, KTH. Opponent: Sandeep Juneja, School of Technology and Computer Science, Tata Institute of Fundamental Research, Mumbai.

26 maj **XIJIA LIU** *On Non Parametric Regression and Panel Unit Root Testing*, statistik, Uppsala universitet. Opponent: Niklas Ahlgren, Hanken Svenska handelshögskolan, Helsingfors.



HÄNT OCH HÖRT

30 maj **BERGRÚN TINNA MAGNÚSDÓTTIR** *Estimation and optimal designs for multi-response Emax models*, statistik, Stockholms universitet. Opponent: Nancy Flournoy, University of Missouri-Columbia.

2 juni **SHUTONG DING** *Model choice in Bayesian VAR models*, statistik, Örebro universitet. Opponent: Matias Villani, Linköpings universitet.

5 juni **ERLAND EKHEDEN** *Catastrophe, ruin and death – some perspectives on insurance mathematics*, matematisk statistik, Stockholms universitet. Opponent: Steven Haberman, Cass Business School, City University, London.

(från sidan 4) **ORSAK TILL ETT VANLIGT FELTÄNK:**

REGRESSION MOT MEDELVÄRDET

Tänk på det här experimentet: Hundra personer singlar slant tolv gånger var. Inför varje slantsingling förutsäger de vad det ska bli. De hundra personerna rangordnas efter hur många gånger de gissat rätt. De fem bästa får ett fint pris och de fem sämsta får en strut på huvudet.

Alla hundra gör om experimentet. Vem tror att de fem bästa i första omgången kommer att fortsätta att vara bäst och vice versa? Om det vore så att de var bäst i kraft av skicklighet i sina förutsägelser, och denna skicklighet består, då skulle man teoretiskt kunna tänka sig att de som var bäst fortsätter att vara bäst.

Daniel Kahneman berättar i sin bok [1] en historia från sitt möte med pilotlärare i det israeliska flygvapnet. Kahneman hävdade att beröm fungerar bättre än bestraffning. En av lärarna invänder. Hans erfarenhet är att en utskällning av dem som borde ha presterat bättre gör att de är bättre nästa gång, och de som fått beröm för att de lyckades alldeles utmärkt med övningen, de brukar inte lyckas lika bra nästa gång.

Fenomenet kallas regression mot medelvärdet. Om det finns ett inte obetydligt element av slump i prestationen, har de som hamnade högt i rangordningen från första omgången gjort det med en del tur. Nästa gång är det andra personers tur att ha tur och de som var bäst första gången har en tendens att halka ner i rangordningen. Om det också finns ett betydande element av skicklighet så är det sannolikt att flera av dem som var bäst i första omgången placerar sig ganska högt även i andra omgången, men kanske inte i toppen. Om de inte har tur igen, eller "medflyt" som det heter i sportsammanhang.

Pilotläraren som har för vana att skälla ut dem som presterat dåligt i en övning är väl inte en så oförstående lärare att han ger sig på dem som misslyckats därför att de inte är tillräckligt skickliga. Dem ger han väl möjlighet att tillägna sig den skickligheten. Men de som han bedömer borde kunna lyckas bra, men inte gör det i en viss övning, de får en rejäl utskällning. Men om det mest var slumpens spel, då skulle pilotläraren få lika god prestationshöjning om han klappar eleven på axeln och säger: "Bättre lycka nästa gång".

Fenomenet regression mot medelvärdet dyker upp i många kläde-dräkter. John Öhrvik nämner en av dem, "vinnarens förbannelse", på sidan 5. (Se till exempel [2] för en beskrivning av vinnarens förbannelse). Det finns mycket skrivet om regression mot medelvärdet. Den här artikeln är inspirerad av [4].

En annan variant är denna: I en skola ger man stöd till dem som lyckas sämst med något. Nästa gång lyckas de bättre. Slutsatsen att de lyckades bättre just därför att de fick stödet är inte välgrundad. Det kan naturligtvis vara sant ändå. Men för att säkert kunna dra den slutsatsen krävs det mer på fötterna.

Sådana slutsatser drar man tyvärr alltför lätt i vardagslivet, i samhället och faktiskt även i forskningen. Faigel [3] redovisar ett experiment som involverade 32 elever som hade så stark ångest eller nervositet under skrivningar i skolan att de presterade illa på skrivningar. De fick beta-blockerare inför en skrivning, och då presterade de bättre. Men experimentdesignen saknade kontrollgrupp, alltså en grupp elever med liknande bakgrund som inte fick beta-blockerare. Så det gick inte att skilja den eventuellt positiva effekten av beta-blockerare från regression mot medelvärdet. En bättre experimentdesign hade medgett det.

DAN HEDLIN

Referenser

- [1] Kahneman, D. (2013). Tänka, snabbt och långsamt. Månpocket.
- [2] Nakaoka, H. och Inoue, I. (2010). The Winner's Curse. In: Encyclopedia of Life Sciences (ELS). Chichester: Wiley.
- [3] Faigel, H.C. (1991). The effect of beta blockage on stress-induced cognitive dysfunction on adolescents. Clinical pediatrics, 30: 441-445.
- [4] Lane, D.M. Regression toward the mean. Online Statistics Education: An Interactive Multimedia Course of Study. Rice University. <http://onlinestatbook.com/>



Har du flyttat?

■ Meddela din nya adress genom att skicka ett e-brev till Thorbjörn Gudmundsson på sekrfram@gmail.com.

Så blir du 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 bli medlem i Svenska statistikfrämjandet skicka ett e-brev till

Svenska statistikfrämjandet

Thorbjörn Gudmundsson 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.

Illustrationer till artikeln:

Högskoleutforskaren underlättar för lärosätena att själva analysera sin verksamhet

Diagram 1: Sambandet mellan årshyran (lodrät axel) och hyrd area (vågrät axel) för några av de stora lärosätena för perioden 1994-2013.

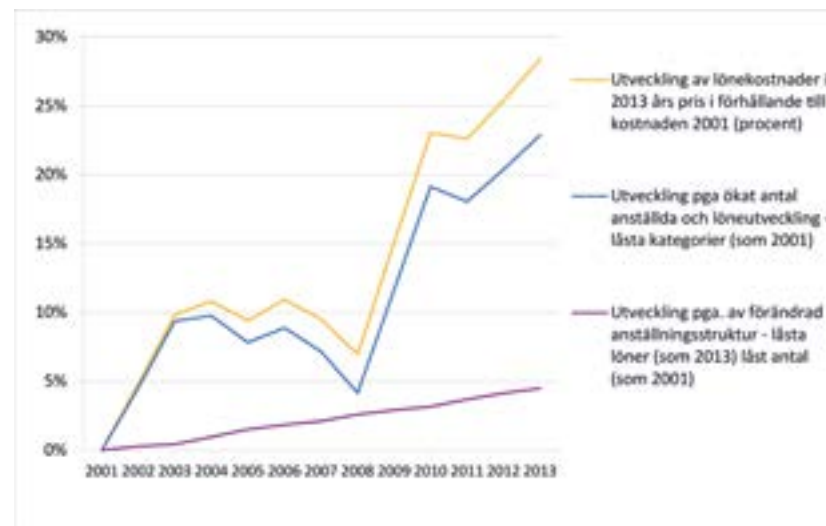
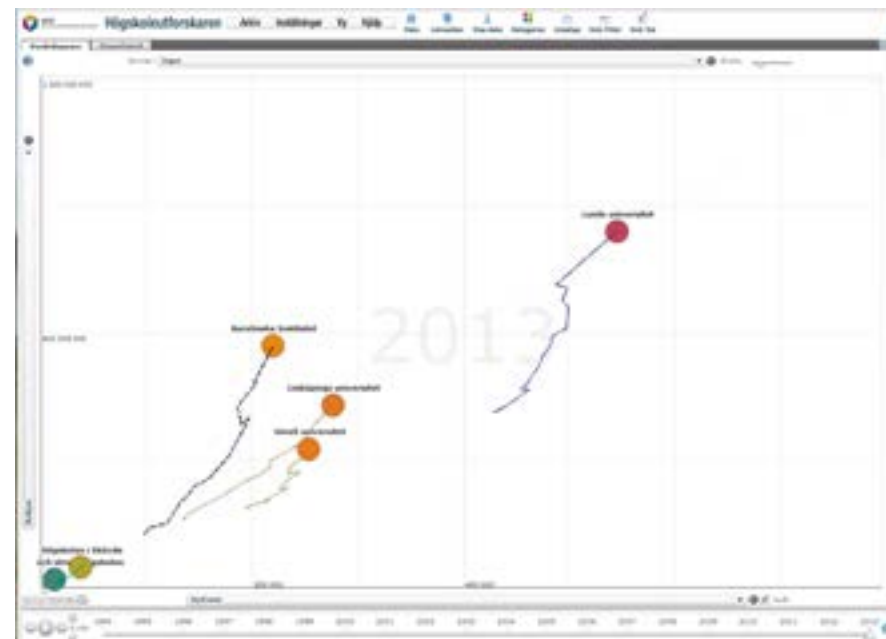


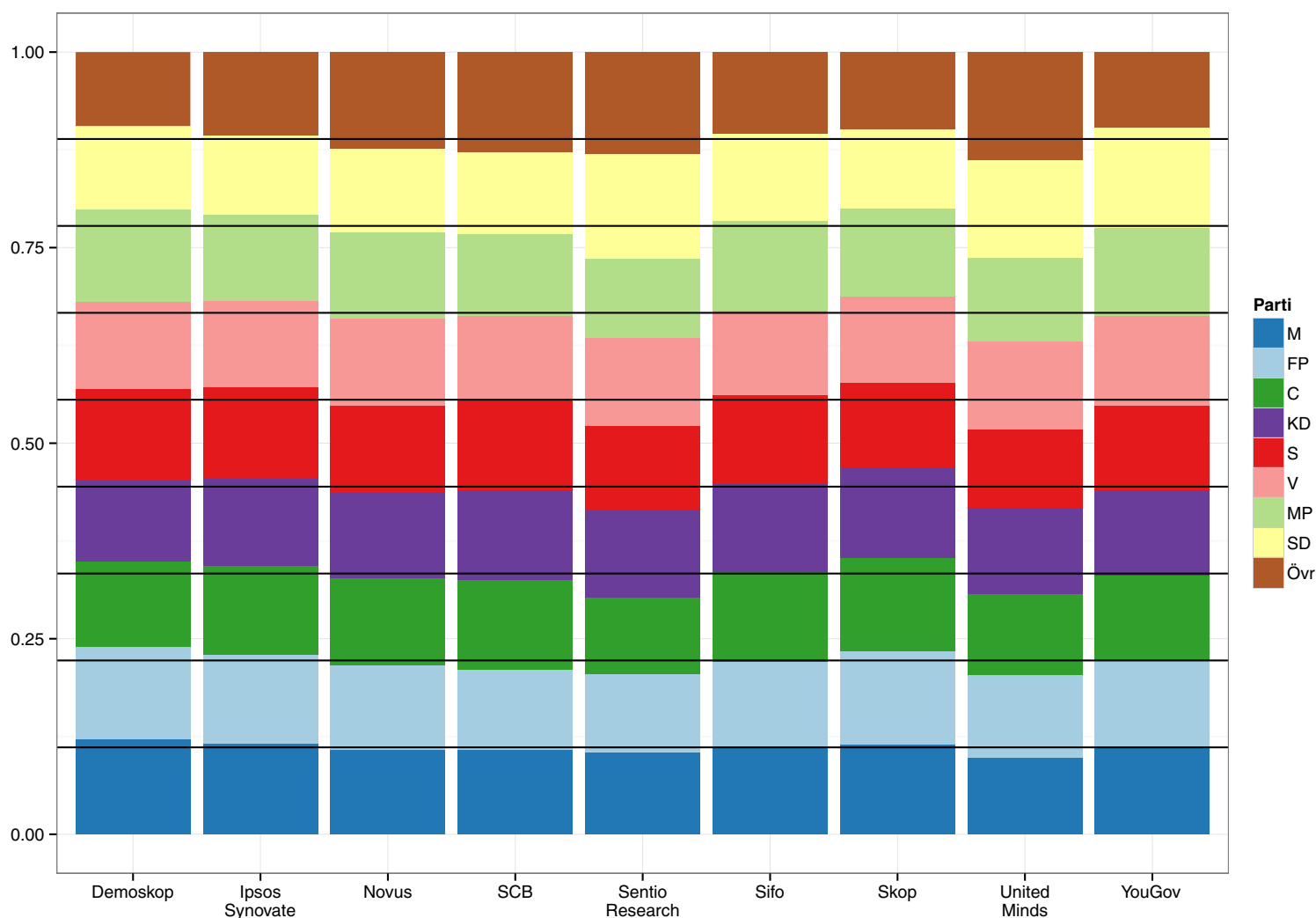
Diagram 2. Lönekostnader 2001–2013 för professorer, lektorer och adjunkter (gul linje), beroende på ökat antal anställda och löneutveckling - lästa kategorier (som 2001) (blå linje) och enbart beroende på den förändrade sammansättningen av professorer, lektorer och adjunkter (lila linje).



Diagram 3. Andelen etablerade (lodrät axel) och antalet examinerade (vågrät axel) för några yrkes- och generella examinationsprogram perioden 1996-2011.

Illustration till artikeln:

**Spelar det någon
roll vilken opinions-
undersökning man
tar del av?**



Figur 1. Genomsnittliga avvikelser för respektive opinionsinstitut i förhållande till den sammanvägda skattningen. De svarta linjerna anger 1/9, 2/9, etc. Novus, SCB, Sifo och YouGov ligger mycket nära de svarta linjerna (1/9, 2/9, etc.) och således i genomsnitt nära den sammanvägda skattningen. Sentio Research, Skop och United Minds ligger längst ifrån.