Svenska statistikfrämjandet
SSFr

SEX ÅR MED QVINTENSEN
SIDORNA 11-13

DISKUSSION/ WEBBPANELER
SIDORNA 20-24

Qvintensen

SVENSKA STATISTIKFRÄMJANDET
NR 2-3 2015

Valla skidor – konst eller vetenskap?
SIDORNA 34-35

Skrivråd för textbyggare
SIDORNA 14-17

Genetik och statistik

– ETT STRÄVSAMT PAR SOM LYSSNAR PÅ VARANDRA
SIDORNA 5-10

»SCB ska respekteras som den instans i samhället som vet bäst»
Carl-Erik Särndal, sidorna 24-26

STATISTICAL SCIENCE AIDING SPORTS

STATISTICAL SIGNIFICANCE

Sports statistics such as a batting average don't involve the scientific discipline of statistics at all, but are merely numbers determined by simple arithmetic. Statistics in the sense of the scientific discipline of collecting, analyzing, and understanding data can yield powerful insights and advantages for those who employ it for sports of any kind. The use of statistical science in sports is still in its early stages, but showing its power and utility, especially with ever-increasing amounts of data.

Statistical Science Applied to Sports

DISCOVERING "HIDDEN" TALENT

Baseball teams are increasingly using a more statistical approach in their decision-making and talent evaluations. Thanks to the vast amount of data collected during games, statisticians can apply sophisticated analyses to assess a player and identify factors that contribute to his success. Teams use this "Moneyball" method to uncover overlooked young talent, as well as players in the major leagues undervalued by their teams. This approach of identifying "hidden gems" is expanding to other sports and influencing non-sports businesses.

RANKING TOP TEAMS IN COLLEGE SPORTS

For NCAA championship tournaments in baseball, basketball, and hockey, only a select number of teams are chosen. Because of the large number of eligible teams and their varying opponents, it is a challenge to select and rank the teams for the playoff tournaments. Statistical models play a key role in this process through the analysis of the outcomes of all games while accounting for such differences as the quality of the opponents faced.

ANTICIPATING OPPONENT BEHAVIOR

Coaches and managers constantly are seeking an advantage over their team's opponent. For example, using the frequency of specific plays a team runs inside the 20-yard line, NFL teams can use statistical analysis to predict the likelihood of specific plays being run for specific situations. Teams also are developing more complicated statistical models to predict the frequency of the type of plays an opponent calls at any place on the field, regardless of score and time remaining. Statistics is used widely in other sports, too. Volleyball coaches track serving tendencies or determine the best way to block a specific attack.

JUDGING PLAYER PERFORMANCE

Decisions in sports—Who to put in the starting lineup? Who to take the last shot?—are increasingly based on performance data, but are not without their challenges. For example, a baseball manager choosing to rest a player with limited success against that day's opponent may understand the judgment often is based on small amounts of data and compounded by the constantly changing nature of human performance. In hockey, judging performance is complicated by the infrequency of goals scored and the practice of players subbing as "lines." The statistical perspective is crucial for addressing these challenges, framing them through the statistical lens of dealing with uncertainty and variability.

ADDRESSING SPORTS CLICHÉS

When a basketball player is called "hot," it implies she likely will score the next basket because she is "in the zone." Yet, studies show it is common for skilled college or professional players to make several consecutive shots. Similarly, in baseball, they say, "Never make the first or third out at third base." Further, a player is called "clutch" when he hits better with runners in scoring position in late-game situations. With scant data for these situations, the application of rigorous statistical techniques is essential to better understand what wisdom, if any, these sayings hold.

MAINTAINING INTEGRITY IN SPORTS

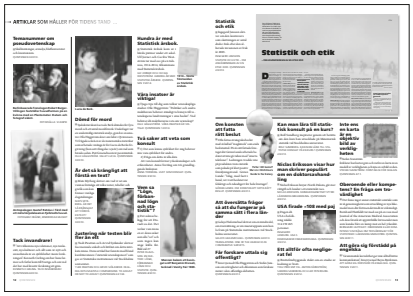
Statistics also plays a key role in ensuring the integrity of sports ranging from baseball and basketball to cycling and sumo wrestling. Sporting events can be tainted by many factors, such as tanking, discrimination, and doping. Demonstrating such breaches can be extremely challenging, however. Statistical analysis can help demonstrate the integrity—or lack thereof—of a sport through the sophisticated examinations of the data, disentangling the many intricate factors.

"Statistical Science Aiding Sports" is part of Statistical Significance, a series from the American Statistical Association documenting the contributions of statistics to our country and society. For more in this series, visit *www.amstat.org/policy/statsig.cfm*. The American Statistical Association is the foremost professional society of statisticians, representing 19,000 scientists in industry, government, and academia: *www.amstat.org*. This Statistical Significance was produced under the supervision of the ASA Section on Statistics in Sports.

Qvintensen
SVENSKA STATISTIKFRÄMJANDET
NR 2-3 2015

Konst eller vetenskap?

Edgar Bueno.



Ett axplock – sex år med Qvintensen.

Genetik och statistik.

Innehåll

Varken myndigheten eller rättsinstanserna har haft sakkunskap att bedöma uttalanden av detta slag.»

Jan Wretman, sidorna 18-20

Fasta spalter

Våra föreningar

I HUVUDET PÅ EN REDAKTÖR

Spelar det någon roll om det man säger är sant?

Frågan ”vad är sanning?” kan användas som en pinne att smaska till den som hävdar att sanning spelar roll. ”Sanning är överensstämmelse med verkligheten”, som filosofiprofessorn Ragnar Ohlsson uttryckte det.

Sammanhanget var sanning i påståenden om verkligheten, sagt i en offentlig debatt med öppen inbjudan. Det var som om publiken blev paff av att han fick en stor fråga att bli enkel.

Publiken blev paff av att han fick en stor fråga att bli enkel.

Vi alla har fel ibland. Men om man som många klimatförnekare inte bryr sig om sanningen, hur tänker man då?

Om en lögndetektor visar i ett experiment p-värdet 0,04, är det då visat sant att apparaten fungerar? Naturligtvis inte. (Se t.ex. min diskussion med Lars-Olof Bygren i Qvintensen nr 3/14).

Om apparaten kan avslöja en lögnare med sannolikheten 90%, är den användbar för Försäkringskassan? En person som får sjukpenning visar sig berätta på Facebook om sitt filmprojekt. Hen får komma in till lokalkontoret och sätta på sig lögndetektorns manschetter. Till att börja med är det bara halva sidan av saken; det är också viktigt att veta hur ofta apparaten har fel om en person som talar sanning. Den kan också vara 90%. En ofta förbisedd sannolikhet.

Efter att ha frågat många användare av statistik om hur de gör när de använder statistik har jag funnit att man ofta formar en helhetsbild, en gestalt, av de siffror och intryck man har. Det kallar jag gestaltinferens. Många beslutsfattare kallar det för ”en samlad bedömning”. De flesta bryr sig åtminstone om sanningen i sin gestaltinferens. Men det kan bli hur fel som helst. Göran Lande skrev (sidan 13 i det här numret) om det svenska företag som la ner projektet med smarta mobiltelefoner eftersom ledningen inte trodde på produkten.

Hur gör proffsen en ”samlad bedömning”? Anders Norgaard har skrivit om det (sidan 13).

Det här är mitt sista nummer som redaktör för Qvintensen. Dags att lämna över till nya krafter.

Jag vill verkligen tacka alla skribenter som ställt upp med sin tid, sitt engagemang och sin sanningslidelse. Tack också till Erika och Sven Åke på MezzoMedia som layoutat alla fina nummer (ett till exempel på skillnaden mellan proffs och amatörer), tack så väldigt mycket till redaktionskommittén, och inte minst de biträdande redaktörerna under de här åren: först Helen Lindkvist och sen Ingegerd Jansson.

FOTO: LIESELOTTE VAN DER MEIJS

Qvintensen

Ansvarig utgivare
John Öhrvik
Redaktör
Dan Hedlin, 070 297 10 27
Ingegerd Jansson (biträdande)
Redaktion
Marie Linder, FMS
Jonas Ortman, Surveyföreningen
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 statistikfrämjandet

SVENSKA STATISTIKFRÄMJANDETS STYRELSE
Ordförande
John Öhrvik, 070-666 08 08,
John.Ohrvik@ki.se
Vice ordförande
Inger Eklund, 08-506 947 13
Sekreterare
Carl Åkerlindh
Klubbmästare Karin Lindgren
Skattmästare Can Tongur
Surveyföreningen
Åke Wissing
FMS
Frank Miller, 08-16 29 76
Industriell statistik
Göran Lande
Cramérsällskapet
Maria Karlsson, 090-786 55 96
Övriga ledamöter
Marie Wiberg, Ylva Tingström
Adress
Svenska statistikfrämjandet
c/o: Carl Åkerlindh
Matematisk statistik, Lunds universitet
Box 118, 221 00 Lund
E-post sekrfram@gmail.com
Webbplats www.statistikframjandet.se

ISSN 2000-1819

GENETIK OCH STATISTIK

– ett strävsamt par som lyssnar på varandra

Som bekant mår förhållanden bra av ömsesidig kommunikation, där båda sidor lyssnar in och tar lärdom av varandra. Det gäller inte bara mänskliga relationer, utan även vetenskapliga discipliner. Genetik och statistik har under lång tid inspirerat varandra. Utan att göra anspråk på att täcka alla delar av denna interaktion, vill jag med några nedslag i historien och belysa händelser som påverkat ämnenas utveckling, och även peka på några aktuella forskningsområden.

Man kan säga att den moderna genetiken föddes då Georg Mendel på 1860-talet upptäckte att anlag ärvs ned från föräldrar till barn i kvantiteter som i matematisk mening är diskreta. Mendel undersökte olika egenskaper hos ärtväxter. Deras fröer var till exempel antingen släta eller veckade; fröerna uppvisade inte en kontinuerlig skala från att vara släta till att vara veckade. Idag kallar vi de diskreta kvantiteter som ger upphov till sådana egenskaper för gener. Mendels experiment med korsning av ärtväxter ledde bland annat fram till två insikter, den att varje individ bär på en dubbel uppsättning anlag eller genvarianter som erhålls från respektive förälder, och den att endast en av individens två anlag, med samma sannolikhet 0,5, ärvs vidare till nästa generation. Mendel menade att dessa principer borde gälla för alla arter. Eftersom varje gen har ett ändligt antal olika varianter innebar hans teorier att den ärftliga variationen är diskret i matematisk mening.

Två synsätt

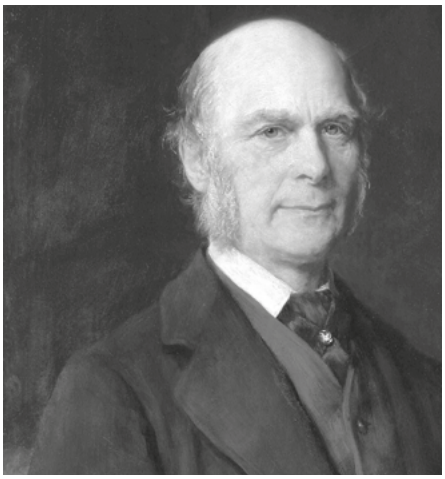
Trots dessa upptäckter dominerade det biometriska synsättet i slutet av 1800-talet, där ärftlig variation antogs vara kontinuerlig, genom att barnet tar emot en blandning av ”anlagsdoser” från respektive förälder, och att dessa doser blandas ungefär som man blandar vätskor. Det synsättet företräddes av Francis Galton. För att kvantifiera i vilken utsträckning kroppslängd var ärftlig införde han korrelationskoefficienten, genom regressionsanpassning av barnens kroppslängd mot föräldrarnas. Han inspirerade i sin tur Karl Pearson, Udny Yule och många andra att anamma kvantitativa metoder inom genetiken, vilket påskyndade statistikämnets utveckling.

Ola Hössjer.

Jag ska använda begreppet fenotyp i fortsättningen. Med fenotyp menas en organisms observerbara egenskaper, som till exempel kroppslängd.

Mendels bortglömda teorier återupptäcktes efter sekelskiftet. Snart uppstod en kontrovers mellan de mendelska och biometriska synsätten, eftersom genbegreppet verkade oförenligt med ärftlig variation av kontinuerliga fenotyper, medan blandningshypotesen skulle medföra att ärftlig variation försvann efter några generationer eftersom blandningen då blivit praktiskt taget homogen. Mendel själv och senare Yule hade visserligen föreslagit att kontinuerliga fenotyper kan ses som polygena, dvs att de orsakas av flera gener som ärvs ned oberoende av varandra, var och en med liten effekt.

Det blev dock en ung brittisk forskare vid namn Ronald Fisher som i ett berömt arbete 1918 på ett mer generellt sätt »»»



Den stenrike Francis Galton verkade först ha fel om generna, men i modern tid har man förstått att han delvis hade rätt ändå.

Georg Mendel, den moderna genetikens fader.

visade hur kontinuerlig anlagsvariation gick att kombinera med det mendelska synsättet. Han delade upp fenotypvariansen i en miljökomponent och ett antal additiva och dominanta genetiska komponenter, och grundlade därmed den kvantitativa genetiken, som fick stor betydelse för bland annat husdjursavel. Lars Rönnegård skrev i Qvintensen nr 1/2013 hur statistiska metoder används idag inom husdjursavel. Men dessa idéer låg även till grund för den variansanalys han senare på 1920-talet utvecklade för försöksplanering inom jordbruket.

Fisher räknas kanske som 1900-talets främste statistiker. Han verkade under en tid när statistikämnet växte fram som en självständig disciplin, både internationellt och i Sverige. Själv var han stora delar av sitt liv professor i genetik, och många av hans statistiska upptäckter, såsom ML-skattningen och det vi numera kallar Fisherinformationen, inspirerades av eller tillämpades på genetiska frågeställningar.

Överkorsningar

På 1910-talet hade man förstått att gener ligger utplacerade längs kromosomer som för människor utgör 23 par, av vilka ett består av två könskromosomer. Avkomman får ena kromosomen i ett par från sin mamma och den andra kromosomen i paret från pappa. För en icke könskromosom har föräldern i sin tur fått de anlag han eller hon ärver vidare till barnet från sin mamma eller pappa. För gener på olika kromosomer stämmer Mendels hypotes att olika anlag ärvs ned oberoende av varandra. För två gener på samma kromosom är däremot sannolikheten för att barnet får anlag från olika mor- eller farföräldrar (en så kallad rekombination) en funktion av generas avstånd från varandra. Att barnet inte säkert får alla anlag på en kromosom från mamma från samma morförälder beror på att delar av den kromosom som härrör från morfar kan byta plats med motsvarande del från mormors kromosom när en äggcell bildas. På samma sätt kommer kromosomstycken från farmor och farfar byta plats när en spermiecell bildas hos pappan. Effekten blir att kromosomer i praktiken ärvs ned styckvis från respektive mor- eller farförälder. Om två gener ligger nära varandra är sannolikheten stor att avkomman får båda av dem från samma mor- eller farförälder. Brytpunkterna mellan kromosomstyckena kallas överkorsningar, och J.B.S. Haldane föreslog 1919 att de lämpligen beskrivs av en Poissonprocess.

Haldanes och snarlika modeller för överkorsningar låg till grund för flera av de genletningsmetoder som utvecklades. Gener lokaliserades successivt längs olika arters kromosomer, genom att Fisher och andra utvecklade metoder för att testa om en nyfunnen gen ligger på en viss kromosom, och givet det, bestämma läget för genen genom att ML-skatta sannolikheten att den rekombinerar, dvs har ett udda antal överkorsningar mellan sig och ett kromosomavsnitt som redan positionsbestämts, en så kallad markör. Detta förfarande kallas kopplingsanalys, eftersom markören och genen sägs vara kopplade till varandra om de inte ärvs ned oberoende. För att hitta anlagen för ärftliga mänskliga sjukdomar samlade man in fenotyper och genvarianter för allt större släktträd. Mendels lagar medför att generas nedärvning i ett släktträd på varje kromosomposition kan beskrivas av en markovsk graf, eftersom vilka gener avkomman får bara beror på de gener som finns i föräldrargenerationen. På 1970-talet utvecklade Robert Elston, Kenneth Lange, Chris Cannings, Elisabeth Thompson och andra sannolikhetsberäkningar på sådana grafer (en föregångare till dagens grafiska modeller och bayesianska nätverk) för att skatta genpositioner.

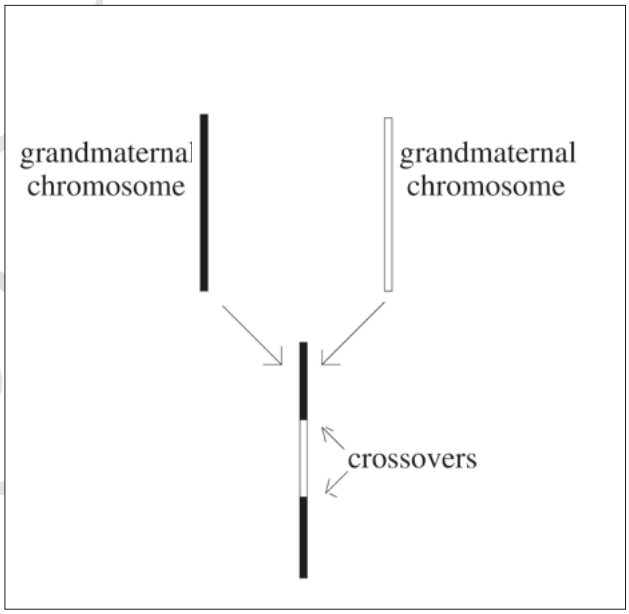
»Den kanske viktigaste användningen av denna kunskap är att skräddarsy medicin, genterapi och andra behandlingsmetoder för olika människor.»

Upptäckten av DNA-molekylen

Med molekylärbiologins landvinningar ökade samtidigt förståelsen för cellens kemiska uppbyggnad. På 1950-talet hade man kommit till insikt om att kromosomerna mellan celldelningar var belägna i cellkärnan, som en dubbelsträngad stege av kvävebaser (DNA-molekylen), där båda strängarna innehåller samma information. I cellerna byggs proteiner upp från beståndsdelar som är aminosyror. Proteiner i kosten spjälkas upp i aminisyror för att senare sättas samman till nya proteiner som behövs i kroppen för en mängd olika uppgifter. Det måste alltså finnas något recept eller instruktion som styr vilka proteiner som kommer till. Några år efter upptäckten av kvävebaserna i DNA hade man förstått att generna utgör kromosomavsnitt som via RNA ger instruktioner för (”kodar för”) cellernas proteinsyntes. Varje instruktion kan beskrivas med trebokstaviga ord som svarar mot olika aminosyror. Dessa insikter möjliggjorde i sin tur en snabb utveckling av gentekniken. Man upptäckte att restriktionsenzymer kan klippa ut DNA-avsnitt, som sedan kunde särskiljas från varandra med elektroforesmetoder. Detta gjorde att data från en snabbt växande karta av genetiska markörer kunde samlas in.

2 gånger Markov

Medan Mendels ärftlighetslagar ger en Markovstruktur över hur gener förs ner genom generationerna, ger Haldanes modell för överkorsningar en annan Markovegenskap för nedärvning, längs kromosomerna. Eftersom



Schematisk bild av hur en kromosom i förälderns könscell (äggcell eller spermie) bildas från ett par av kromosomer från de två far- eller morföräldrarna, i detta fall genom två överkorsningar (crossovers).

de observerade genetiska markörerna ger en bild av denna nedärvning som innehåller stora mått av statistisk osäkerhet, visade bland annat Eric Lander på 1980-talet hur gömda (ej direkt observerbara) Markovkedjor kunde tillämpas för att på ett bättre och mer fullständigt sätt utnyttja informationen från alla markörer. Under 1980-talet och en bit in på 90-talet hittade man de genvarianter som ökar risken för cystisk fibros, Huntingtons sjukdom och ett stort antal andra monogena sjukdomar.

I början av 90-talet sjöd optimismen, snart skulle man finna de genetiska komponenterna hos många betydligt vanligare folksjukdomar, exempelvis åldersdiabetes, Alzheimers sjukdom, multipel skleros, vissa psykiska sjukdomar och olika typer av cancer. För att åstadkomma detta beslutades att hela det mänskliga genomet skulle kartläggas, det så kallade HUGO-projektet. Det tog ett decennium att sekvensera ett prototypgenom med cirka 3 miljarder baspar genom att först sönderdela DNA-strängarna i delvis överlappande små fragment. Med Poissonapproximation uppskattades hur många fragment som krävs för att med stor sannolikhet täcka hela genomet, och sedan användes olika datalogiska metoder för att pussla ihop dem i rätt ordning. Efter HUGO-projektets slutförande i början av 2000-talet har man under mer än ett decennium i de så kallade HapMap och 1000 Genomesprojekten kartlagt de miljontals baspar som är polymorfa, dvs varierar mellan individer och gör oss genetiskt olika som människor.

Enorma mängder av data medför nya problem

Med dessa nya data uppstod redan på 1990-talet behov av ny teoriutveckling för genletning, inte minst olika simuleringstekniker som MCMC och vägd simulering, för att hantera stora släktträd, många markörer och ofullständiga data. Trots denna avancerade statistiska teori visade det sig vara mycket svårare att genbestämma folksjukdomarna, eftersom de svarar mot det gamla biometriska synsättet, där ingen enskild gen har en stor effekt. Detsamma gäller många andra komplexa fenotyper, såsom kroppslängd, BMI och olika typer av kognitiv förmåga.

När man genomför många hypotestest i samma studie kommer med stor sannolikhet någon andel av testen visa falska signifikanser. I takt med att markörkartorna för hela genomet blev tätare, började tyvärr alltfler falska positiva fynd med kandidatregioner för sjukdomsgener publiceras. Det blev därför nödvändigt att i kopplingsanalysen korrigera för multipel testning på ett strikt sätt, och eftersom nedärvningen av närliggande markörer är korrelerad, är Bonferronimetoden alldeles för konservativ. Istället använde David Siegmund, Eric Lander och Leonid Kruglyak extremvärdesteori för gaussiska processer för att ange kriterier för när ett kromosomavsnitt kunde rapporteras som ett äkta fynd. Dessa formler för ”familywise error rate” (FWER) är dock approximativa, speciellt när informationen från markörer är ofullständig, och idag används ofta simuleringsmetoder.

»Under 1980-talet och en bit in på 90-talet hittade man de genvarianter som ökar risken för cystisk fibros, Huntingtons sjukdom och ett stort antal andra monogena sjukdomar.»

Analys i annan ledd

I stället för att som i kopplingsanalysen använda data från familjer, kan data bestå av fall och kontroller utan nära släktskap. Man låter markörerna vara kovariater i en logistisk regressionsmodell och testar om de själva (inte deras nedärvning) är associerade med den oberoende variabeln, som anger om individen är sjuk eller ej. Det kallas associationsanalys. Förutom att data för associationsanalysen är lättare att samla in än de data som behövs i kopplingsanalysen, upptäckte Neil Risch och hans medarbetare mot slutet av 90-talet att associationsanalysen ofta har högre styrka än kopplingsanalysen för polygena sjukdomar, där varje riskgen tagen för sig har liten effekt. Detta gäller trots att problemet att välja rätt modell i associationsanalysen är besvärligt, eftersom fler markörer måste undersökas.

Trots associationsstudiernas högre styrka blev det ändå nödvändigt att forskargrupper från många länder bildade konsortier och slog ihop sina dataset, och att markörer som inte var med i en viss studie imputerades från referensgenomen i HapMap eller 1000 Genomes. Detta har många gånger varit framgångsrikt, och man har identifierat tusentals genvarianter som har en statistiskt signifikant (men oftast marginell) association med totalt flera hundra komplexa sjukdomar eller andra fenotyper. Statistisk association är dock inte samma sak som kausalitet, och för de flesta av dessa genvarianter återstår att förklara varför de höjer risken för sjukdomen. Dessutom försvåras sådana internationella metastudier av att gener kan i olika grad vara associerade med sjukdomen i olika populationer (heterogenitet), och att associationsanalysen har svårigheter att upptäcka ovanliga genvarianter. För att även utnyttja de insamlade familjedata som faktiskt finns, används därför kopplingsanalysen ibland som ett komplement till eller tillsammans med associationsanalysen.

Men data från HUGO-projektet användes också tillsammans med dess efterföljare för andra arter för att gruppera DNA-sekvenser »»»

(eller aminosyresekvenser) med liknande utseende hos olika arter, för att bland annat förstå funktionen hos de proteiner de kodar för. Detta kan åstadkommas med gömda Markovmodeller, medan programpaketet BLAST använder andra stokastiska metoder (utvecklade av bland annat Samuel Karlin och Stephen Altschul) för att avgöra om likhet mellan olika sekvenser är signifikant. Detta innefattar extremvärdesteori för exkursioner av slumpvandringar med negativ drift, förnyelseteori och sekventiella tester. Jag skrev om detta i Qvartilen år 2004.

Det hänger inte bara på vad genen gör utan även på hur aktivt den jobbar

En annan viktig utveckling var när man mot slutet av 1990-talet hittade storskaliga metoder för så kallad mikroarrayanalys, med syftet att bestämma vilka gener som är aktiva i olika celler med att uttrycka sin proteinkod. För att hitta riskgener för sjukdomar är det framför allt skillnaden i genaktivitet mellan sjuka och friska individer i den för sjukdomen aktuella cellvävnaden som är intressant. Dessa nya data ledde till vidareutveckling av Robert Tibshiranis lasso och andra metoder för att skatta effekter i regressionsmodeller där antal prediktorer (t ex antalet undersökta gener) är större än antal observationer (exempelvis antal individer). Även teorin för multipla tester fördes framåt. Eftersom FWER är ett alldeles för konservativt kriterium för mikroarrayanalys, används istället ofta bayesianska metoder eller False Disovery Rates (FDR), och genom arbeten av bland annat Sandrine Dudoit, Terry Speed och John Storey har förståelsen för multipel testning i allmänhet och FDR i synnerhet ökat. En modern efterföljare till mikroarrayanalys, så kallad RNA-sekvensering, har ibland visat sig vara ett mer flexibelt sätt att upptäcka fler typer av genaktivitet, och även deras dynamik. För denna typ av data består genuttrycken av en multivariat tidsserie, och att analysera dem innebär nya statistiska utmaningar.

Genetikern RA Fishers originella idéer är idag standargods i utbildningen i statistik

Skräpgenerna har visat sig vara viktiga

Bara de senaste 15 åren har den moderna biologin utvecklats mycket snabbt. ENCODE-projektet i början på 2000-talet visade att icke-kodande DNA, som finns mellan generna och som upptar drygt 98 % av genomet, verkar ha en mängd viktiga funktioner, bland annat för genreglering med hjälp av så kallade transkriptionsfaktorer eller mikro-RNA. Det är möjligt att flera av de genvarianter som hittats med associationsstudier snarare ligger nära men utanför en gen eller i en icke-kodande del av genen. I båda fallen är det genens grad av aktivitet som påverkas, snarare än vilket protein den kodar för. Här kan även stokastisk automatteori och andra matematiska verktyg få betydelse för att i mer detalj beskriva informationen hos den icke-kodande DNA-strängen. Den transkription av DNA-strängar som ligger till grund för proteinsyntesen är mycket mer raffinerad och komplicerad än man tidigare trott, eftersom en gen kan koda för flera proteiner, beroende på varifrån och i vilken riktning avläsningen sker. Dessutom har epigenetiken visat att förmågor som förvärvats under livet i vissa fall kan ärvas vidare till nästa generation, bland annat genom förändringar i de proteinkomplex (histoner) som DNA-molekylerna är virade kring.

»Varje cell kan liknas vid en hel stad av aktivitet»

Den nya kunskapen om DNA-strängar integreras alltmer med ökad förståelse av hur en cell fungerar. Det eller de proteiner en gen kodar för ingår i ett antal kemiska reaktioner. Varje cell kan liknas vid en hel stad av aktivitet, där systembiologin ger utmärkta verktyg för att beskriva de processer för ämnesomsättning, kraftproduktion, lagring av restprodukter och transport som pågår. Även här har statistiken en viktig roll att spela, t ex associationsstudier mellan genetiska markörer och olika multivariata fenotyper på cellnivå, såsom genuttryck, proteinsekvenser och metabolism. När sådana ”big data” analyseras får man ännu större problem med multipla tester, inte minst när olika typer av data kombineras. Men även stokastisk reglerteori, dynamiska bayesianska nätverk och olika hierarkiska modeller är väl lämpade för att beskriva sådana komplexa system. För vissa bayesianska modeller är likelihoodfunktionen så svår att beräkna att den måste approximeras med simuleringar (så kallade ABC-metoder). Men ibland är approximativa och enklare modeller mer praktiskt användbara, dels för att de är mindre beräkningsintensiva, dels för att de är robustare mot felaktiga modellantaganden.

Det jag hittills beskrivit är hur den genetiska kunskapen ökat under 150 år, både vad gäller kartläggning av gener, vilka varianter av dem som finns och vilken funktion de har. Den kanske viktigaste användningen av denna kunskap är att skräddarsy medicin, genterapi och andra behandlingsmetoder för olika människor. Vi är fortfarande bara i början på denna utveckling, där man snabbt kommer in på etiska frågeställningar när den personliga integriteten ska avvägas mot användande av mer eller mindre kodad genetisk information. Rent statistiskt kan DNA, genaktivitet, metabolism och andra biomarkörer i olika celler ses som kovariater eller prediktorer i en regressionsmodell, där effekten av en viss medicin blir responsvariabel. För att öka förståelsen av sambandet mellan riskfaktorerna kan kausal inferens vara användbart, ett område där bland annat Donald Rubin gjort viktiga bidrag.

Förändringar av den genetiska variationen i en hel population

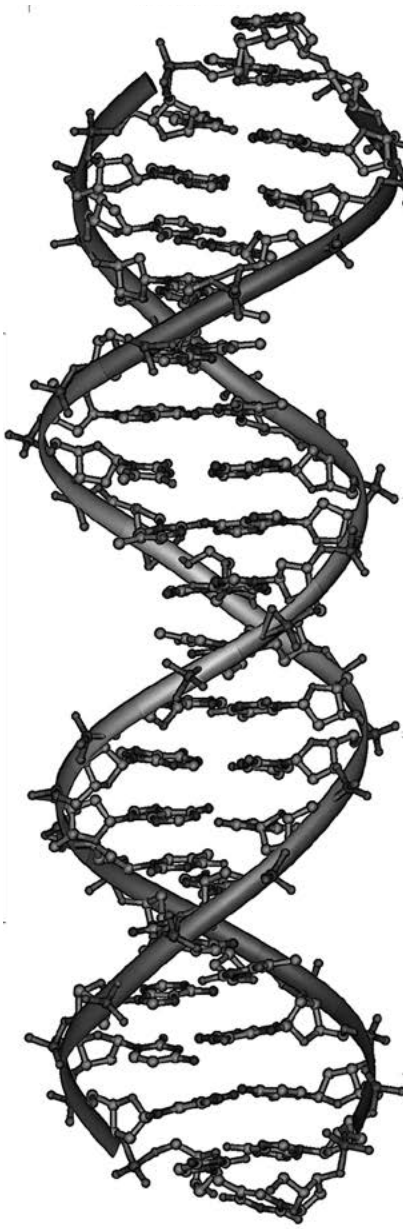
En annan viktig frågeställning är hur den genetiska variationen i en hel population eller art *förändras* över tid. Populationsgenetiken började utvecklades redan på 1920- och 1930-talen för att besvara just denna fråga. Avgörande insatser gjordes av Fisher, Haldane och Sewall Wright för att kartlägga de byggstenar som driver fram förändringar i form av 1) genetisk drift, dvs slumpvariation med avseende på vilken av de två mor- eller farföräldrarna som ärver ned ett anlag till ett barnbarn, 2) mutationer, dvs förändringar

i DNA då könsceller bildas, 3) naturligt urval, dvs en selektiv reproduktiv fördel för individer med vissa genvarianter, 4) rekombination av DNA genom överkorsningar, och 5) migration, dvs genutbyte mellan individer från olika delpopulationer. Av dessa mekanismer är det bara mutationerna, och i viss mån rekombinationerna, som kan medföra att nya genvarianter bildas. Eftersom de allra flesta mutationerna dessutom är neutrala eller skadliga, föreslog Motoo Kimura på 1970-talet att genetisk drift är en viktigare mekanism för att förklara genetiska förändringar än naturligt urval. Observera att genetiker och sannolikhetsteoretiker lägger in helt olika betydelser i ordet drift, något som till en början är förvillande för den som är driftig nog att läsa arbeten från båda disciplinerna.

Populationsgenetiken använder flitigt redskap från sannolikhetsteori och stokastiska processer. Jag nämnde ovan att nedärvning i ett släktträd kan beskrivas av en markovsk graf över generationer. Men då släktförhållandena sällan är kända i en hel population, brukar man istället använda förenklade Markovkedjor för att beskriva hur den genetiska sammansättningen ändras framåt i tiden t. I den enklaste enkönade Wright-Fisher modellen (WF) är genetisk drift den enda förändringsmekanismen. Den utgår från en selektivt neutral gen eller genetisk markör med två varianter (I och II), utan nya mutationer. Populationsstorleken N antas vara konstant och tiden räknas i generationer, där antalet kopior

$$X_{t+1} \sim \text{Bin}\,(N, \tfrac{X_t}{N}) \qquad (1)$$

av I ändras från en generation till nästa genom att barnen slumpmässigt med återläggning drar sina föräldrar, så att X_t bildar en Markovkedja. Samuel Karlin och Chris Cannings utvecklade sedan mer allmänna modeller, där antalet ”barn” till föräldrarna i generation t också är utbytbara stokastiska variabler, men med större variation än den som följer av (1). Man kan generalisera modellen och låta populationsstorleken variera över tid. Selektion införs i WF-modellen genom att låta genvariant I ha en reproduktiv fördel $s > 0$ och låta sannolikheten i (1) bli proportionell mot $(1 + s)\, X_t/N + (1 - X_t/N)$. För att bygga in migration måste man införa geografisk struktur, exempelvis i form av delvis isolerade öar med ett visst genutbyte emellan. Men man kan även införa andra typer av struktur, såsom



Schematisk bild av DNA, en dubbelspiral uppbyggd av de fyra kvävebaserna A (adenin, grön), T (tymin, lila), C (cytosin, rosa) och G (guanin, blå). De två strängarna innehåller samma information eftersom A alltid binds till T och C alltid till G. Se bild i färg på framsidan.

ålder (genom att definiera livslängdstabeller, på liknande sätt som i livförsäkring), eller kön (genom de mendelska nedärvningslagarna och olika parningsmönster).

Att få in mutationer i modellerna

För modeller med mutationer blir tillståndsrummet mer komplicerat, eftersom nya genvarianter tillkommer. Men delar in de N individerna i ekvivalensklasser beroende på deras genvariant, och låter X_t beskriva tidsdynamiken hos klassstorlekarnas fördelning. Warren Ewens angav i början av 1970-talet en samplingformel, som under vissa antaganden ger jämviktsfördelningen hos selektivt neutrala genvarianter i närvaro av mutationer. Den uppstår som en balans mellan de två mekanismer som strävar efter att minska (genetisk drift) respektive öka (mutationer) den genetiska variationen. Rekombinationer, slutligen, införs när den genetiska variationens tidsutveckling studeras för flera markörer samtidigt, så att X_t blir en vektor där varje element svarar mot en markör.

Hittills har ”tiden” t bara varit en räknare för generation. För stora populationer är det ofta en fördel att införa kontinuerlig tid och approximera Markovkedjan med en stokastisk differentialekvation. Om vi undersöker flera genpositioner samtidigt blir denna SDE vektorvärd, där A) den genetiska driften och rekombinationerna ger elementen i diffusionstermens kovariansmatris, och B) det naturliga urvalet och migrationen orsakar den systematiska driften. Om vi dessutom utökar modellen till att innehålla mutationer blir tillståndsrummet mer komplicerat, som i det tidsdiskreta fallet. Medan Fisher lade grunden för SDE-approximationerna, utvecklade Wright denna teknik vidare, och sedan tog Kimura vid och löste på 1950-talet Kolmogorovs framåt- och bakåtekvationer explicit i flera viktiga fall.

Att vända på tiden

I slutet av 1970-talet skrev Ewens en bok som sammanfattade mycket av den dittills utvecklade matematiska teorin för populationsgenetik. Kort därefter introducerade John Kingman i början av 1980-talet en idé som visade sig vara mycket fruktbar för populationsgenetiken. Hans koalescensteori kan sägas vara en vidareutveckling av att vända på tiden för en Markovkedja. Vi får då en ny Markovkedja. Visserligen hade Charles Cotterman och Gustave Malécot mer än 30 år tidigare infört begreppet identisk härkomst för »»»

gener, men Kingman visade hur ett helt genetiskt släktträd är fördelat bakåt i tiden. Detta släktträd är inte en vanlig individbaserad stamtavla, utan ett Markovträd som visar hur olika individers kopior av en gen har ärvts ned längs olika släktlinjer som så småningom sammanstrålar hos en gemensam anfader. Dess utseende beror inte bara på den kromosom vid vilken genen är belägen, utan även (på grund av överkorsningar) på positionen längs denna kromosom. En av de största fördelarna med koalescensteorin är att selektivt neutrala modeller blir mycket enklare att analysera, eftersom endast de mutationer som överlevt till dagens generation tas med, och de kan slumpas ut som en Poissonprocess på släktträdet.

Koalescensteorin har idag utvecklats till en egen disciplin inom sannolikhetsteorin, där Jean Bertoin, Peter Donnelly, Richard Durrett, Simon Tavaré och andra gjort viktiga bidrag, men även flera svenskar, såsom Magnus Nordborg, Peter Jagers, Ingemar Kaj och Serik Sagitov. Den har många tillämpningsområden, och inom populationsgenetiken är den ett fruktbart verktyg för att studera mänsklighetens historia och andra processer bakåt i tiden, där moderna dataset med hundratusentals markörer per individ ger nya möjligheter att jämföra olika modeller. Något mer överraskande är att koalescensteorin även är mycket användbar för att studera förändringar framåt i tiden. Den kan exempelvis ge framåtrekursioner för den predikterade inavelskoefficienten

$f_t = P$(samma genvariant för två slumpmässigt valda individer, tid t)

och andra av Wright införda storheter. Liksom i filmen Tillbaka till framtiden kan vi med tidsmaskinens hjälp bygga upp (sannolikhetsfördelningen för) ett släktträd bakåt, med start i framtiden.

Risken för inavel

Inom bevarandebiologin studerar man hur den biologiska mångfalden i olika ekosystem kan bevaras. Bara de senaste decennierna har många djurarter antingen dött ut eller blivit starkt hotade, bland annat på grund av ökad inavel som ökar förekomsten av recessiva sjukdomar. För att förhindra detta vill man kunna förutsäga risken för att olika anlagsvarianter går förlorade och hur olika strategier påverkar inavelsförändringen (som djurreservat, fiske, fiskodling, anläggning av dammar, avskjutning av älgar, migration av vargar från Finland och Ryssland). Under det korta tidsperspektiv det här är fråga om är det framför allt genetisk drift, migration och rekombinationer som påverkar. Dessutom verkar de kontinuerliga SDE-approximationerna ofta över för lång tid för att vara användbara. Eftersom risken för ökad inavel är större för små populationer än för stora, används den effektiva populationsstorleken N_e som ett mått på inversen av den takt med vilken inaveln ökar. En vanlig tumregel är $N_e \geq 50$ för kortsiktigt bevarande av en population 5-10 generationer framåt i tiden.

»Man vill kunna förutsäga risken för att olika anlagsvarianter går förlorade och hur olika strategier påverkar inavelsförändringen»

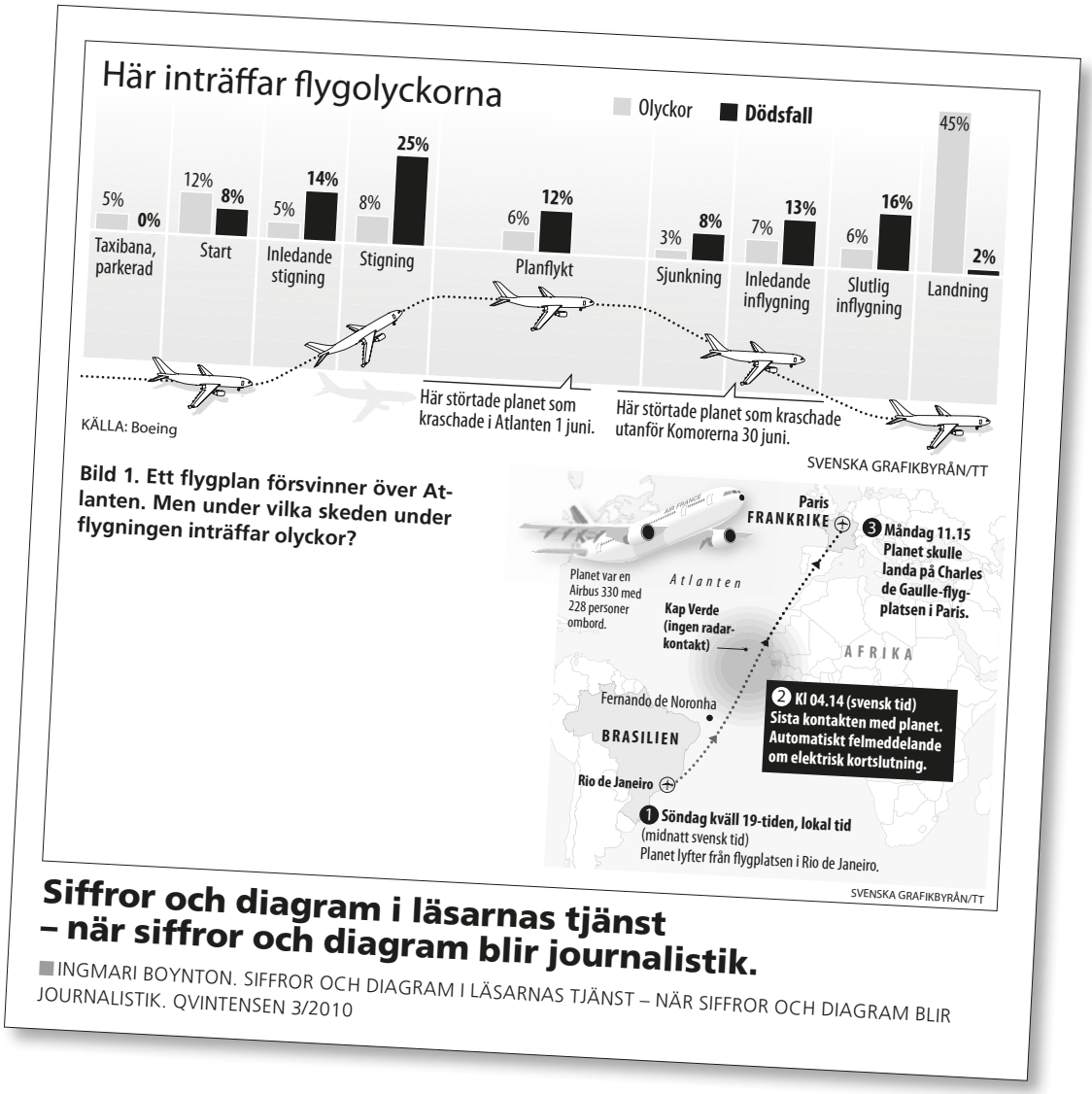
Man kan ange N_e på flera olika sätt. En metod är att utgå från hur snabbt sannolikheten f_t ökar mot 1. Alternativt kvantifierar man hur snabbt genetisk drift sker. För en genetisk markör med två varianter, blir N_e då inversen av en normaliserad varians

$$\sigma^2 = V \alpha r \left(\frac{x_{t+1}}{x_t(N-x_t)} | X_t \right) = \frac{1}{N_e},$$

som kan liknas vid volatiliteten för de processer som studeras i finansiell matematik. En tredje metod utgår från största egenvärdet mindre än 1 hos övergångsmatrisen för Markovkedjan X_t. Det anger hur snabbt ett absorberande tillstånd nås: I ett absorberande tillstånd stannar Markovkedjan där den är och därmed har anlag förlorats. Ett fjärde sätt är att med hjälp av koalescensteorin ange hur snabbt grenarna i trädet går ihop bakåt i tiden. Det visar sig att alla dessa definitioner av N_e överensstämmer med den verkliga populationsstorleken N för WF-modellen. Men oftast är N_e betydligt mindre, speciellt om reproduktiviteten mellan föräldrar varierar mer än vad (1) anger.

Under de senaste åren har jag i min forskning arbetat med att få fram en enhetlig matematisk teori för populationer som av någon orsak är strukturerade. Såväl Markovkedjan X_t som den deterministiska inavelsprocessen f_t blir då vektorvärda, och definitionerna av effektiv storlek ger olika värden på N_e. Till exempel visar det sig att alla komponenter hos f_t -processen går mot 1 med samma hastighet, och den ges av största egenvärdet hos den matris som ligger till grund för tidsrekursionen från f_t till f_{t+1}. Men för kortare tidsperspektiv är det mer intressant att ange hela tidsprofilen för inavelsökningen, och då tar man med matrisens hela spektrum av egenvärden. En del av detta, och annat, beskrivs i den avhandling som Fredrik Olsson försvarade i maj 2015, ett arbete som utförts i samarbete med två populationsgenetiker från Stockholms universitet, Linda Laikre och Nils Ryman.

OLA HÖSSJER
STOCKHOLMS UNIVERSITET



Siffror och diagram i läsarnas tjänst – när siffror och diagram blir journalistik.

INGMARI BOYNTON. SIFFROR OCH DIAGRAM I LÄSARNAS TJÄNST – NÄR SIFFROR OCH DIAGRAM BLIR JOURNALISTIK. QVINTENSEN 3/2010

SEX ÅR MED QVINTENSEN

Artiklar som håller för tidens tand…

Efter sex år lämnar jag posten som redaktör för Qvintensen. Jag ville göra ett collage med artiklar från de sex åren, för att tacka alla skribenter som ställt upp med sin fritid och sitt engagemang. Att samla ihop spännande artiklar, som håller för tidens tand, är en enkel uppgift, för spännande, intressanta och roliga artiklar finns det ett överflöd av i Qvintensen.

Att välja var svårare, eller helt omöjligt.

Alla goda artiklar har inte kommit med, förstås, för det skulle kräva enormt många sidor. Men alla nummer av Qvintensen kan hämtas från Svenska statistikfrämjandets webbplats, och de flesta av de artiklar som nämns här finns dessutom som separata pdf-filer på http://statistikframjandet.se/qvintensen.

Sprid dem på sociala medier!

Tack för den här tiden!

DAN HEDLIN

»»»

Temanummer om pseudovetenskap

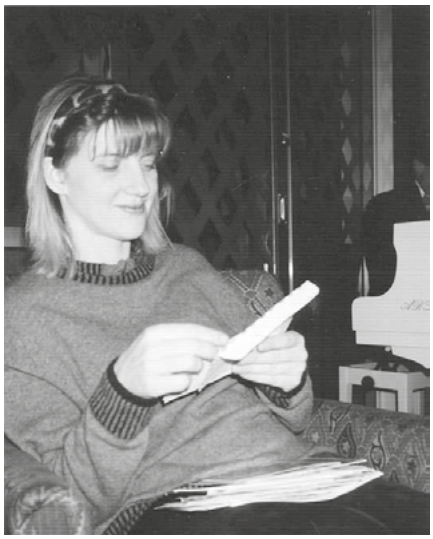
■ Skallmätningar, ormolja, bluffuniversitet och kreationism.
QVINTENSEN 3/2010.

Berlinbaserade frenologen Robert Burger-Villingen fastställer huvudformen på en kvinna med en Plastometer. Datum och fotograf okänt.
FOTO/KÄLLA: SCANPIX

Antropologen Gustaf Retzius i färd med att mäta härjedalssamen Fjellstedts huvud.
FOTOGRAF OKÄND, ©NORDISKA MUSEET

Tack invandrare!

■ ”Att välkomna nya relationer, nya tankesätt, nya kulturer och allt som är nytt och annorlunda är en självklarhet inom forskningen”. Kenneth Carling om hur hans farmor och farfar kom till Sverige och om vad det har med kreativ forskning att göra.
KENNETH CARLING. TACK INVANDRARE! QVINTENSEN 4/2014.



Lucia de Berk.

Dömd för mord

■ Sjuksköterskan Lucia de Berk dömdes för fyra mord och ett antal mordförsök. Underlaget var en undermålig statistisk analys, gjord av en amatör. Olle Häggström skrev om fallet i Qvintensen. Vi frågade också en av de matematiska statistiker som arbetade i många år för Lucia de Berks frigivning (hon satt i fängelse i sju år) om vad som hände sedan. Peit Groeneboom är pessimistisk.
OLLE HÄGGSTRÖM. FALLET LUCIA. QVINTENSEN 3/2010.

Är det så krångligt att förstå en text?

■ Mats Myrberg skriver om vad vi vet om vuxnas förmåga att tolka texter, tabeller och grafik som kan finnas i vanliga dagstidningar.
MATS MYRBERG. ÄR DET SÅ KRÅNGLIGT ATT FÖRSTÅ EN TEXT? QVINTENSEN 3/2010.

Justering när testen blir fler än ett

■ Yudi Pawitan och Arvid Sjölander skriver fascinerande enkelt och lättläst om detta intrikata ämne. Deras artikel har funnits med bland kurslitteraturen i ”statistisk vetenskapsteori” som ges av Statistiska institutionen vid Stockholms universitet.
YUDI PAWITAN OCH ARVID SJÖLANDER. DEALING WITH MULTIPLE COMPARISONS: TO ADJUST OR NOT TO ADJUST. QVINTENSEN 2/2014.

Hundra år med Statistisk årsbok.

■ Statistisk årsbok kom ut i hårda pärmar under ett sekel. Ulf Jorner och Cecilia Weststrom tar med oss på en tidsresa, 1914-2014, tillsammans med Statistisk årsbok.
ULF JORNER OCH CECILIA WESTSTRÖM. HUNDRA ÅR MED STATISTISK ÅRSBOK. QVINTENSEN 1/2014.

1914 – första förstasidan av Statistisk årsbok.

Våra insatser är viktiga!

■ Tjugo tips till dig som tolkar vetenskapliga studier. Olle Häggström: ”Politiker och andra makthavare behöver ständigt ta hänsyn till vetenskapens landvinningar i sina beslut”. Vad behöver då makthavarna veta om vetenskap?
OLLE HÄGGSTRÖM. VÅRA INSATSER ÄR VIKTIGA! QVINTENSEN 1/2014.

Två saker att veta som konsult

■ 1) Det som känns självklart för mig behöver inte vara självklart för andra.

2) Fråga om detta är alla data

Att vara konsult kräver yrkeskunskaper och erfarenheter. Anna Torrång om två grundläggande lärdomar.
ANNA TORRÅNG. LIVET SOM KONSULT. QVINTENSEN 3/2014.

Vem sa ”Lögn, förbannad lögn och statistik”?

■ Det saknas belägg för att Disraeli sa det. Det verkar vara ännu en av dessa saker som alla ”vet” och som ingen kan ange källa för.
PETER OLAUSSON. EN FAKTOID: LÖGN, FÖRBANNAD LÖGN OCH STATISTIK. QVINTENSEN 4/2013.

Mannen bakom ett bevingat ord? Benjamin Disraeli, tecknad i Vanity Fair 1869.

Statistik och etik

■ Ingegerd Jansson skriver om den kontrovers som orsakades av skattningen av antal döda i Irak efter den allierade invasionen av Irak år 2003.
INGEGERD JANSSON. STATISTIK OCH ETIK – OM ANVÄNDNINGEN AV EN ETISK KOD. QVINTENSEN 2/2010.

ETIK

Den 4 februari 2009 gick organisationen American Association for Public Opinion Research (AAPOR) ut med ett pressmeddelande där man offentligt kritiserade Gilbert Burnham, John Hopkins Bloomberg School of Public Health, för att ha brutit mot AAPOR:s etiska kod.

Statistik och etik

– OM ANVÄNDNINGEN AV EN ETISK KOD

Det är ingenting som AAPOR gör slentrianmässigt, senast det hände var 12 år tidigare, och beslutet att kritisera Burnham föregicks av åtta månaders utredning. Vad var det då som låg bakom beslutet?

Redan 2004 publicerade Burnham och en grupp kollegor en studie i the Lancet. Syftet med studien var bland annat att skatta antalet civila döda i Irak ett och ett halvt år efter den allierade invasionen av landet i mars 2003. Enligt studien hade dödsfrekvensen bland civila ökat efter invasionen, och framför allt risken att dö en våldsam död. Det var speciellt skattningen av antalet döda som väckte uppmärksamhet. 98 000 dödsfall skulle enligt studien vara orsakade av invasionen. Skattningen var behäftad med stor osäkerhet, det rapporterade konfidensintervallet var (8 000;194 000). Uppmärksamheten i media var intensiv och studien fick en hel del kritik som bland annat rörde mätmetoden. Det hävdades också att man inte räknat korrekt och att konfidensintervallet i själva verket täcker noll. Men man fick även erkännande för att man lyckats genomföra undersökningen i ett krigshärjat land och för att man bidragit till att belysa krigets tragiska konsekvenser. USA tar inte fram officiella siffror över antalet civila döda i Irak, men det finns andra rapporterade siffror som dokumenterats på till exempel sjukhus, bårhus och genom projektet Iraq Body Count. Dessa siffror var betydligt lägre än skattningen i the Lancet och insamlade med helt andra metoder.

I oktober 2006 publicerade Burnham och hans kollegor en ny artikel i the Lancet, Ytterligare en studie hade då genomförts, på ungefär samma sätt som tidigare och i lite större omfattning, 50 kluster fördelades proportionellt mot befolkningen över 18 strata motsvarande de 18 provinserna i Irak. Inom varje provins gjordes ett urval av administrativa områden proportionellt mot befolkningsstorleken. En startpunkt valdes slumpmässigt i varje område genom att först välja en huvudgata och därefter en gata med bostadshus som korsade huvudgatan. Därefter valdes ett hushåll som tillsammans med de 39 närmaste hushållen utgjorde ett kluster. Data om födslar och dödsfall (inklusive dödsorsak) sedan januari 2002 samlades in genom besksintervjuer.

I den tidigare studien från 2004 valdes startpunkterna ut genom att slumpa fram koordinater och GPS används för att lokalisera det första hushållet i ett kluster. Det fanns ett viktigt skäl till att metoden ändrades; att vina sig med en GPS i Irak var förenat med livsfara eftersom den kunde misstas för en fjärrutlösare eller tolkas som att intervjuaren bedrev spaningsarbete inför ett luftanfall.

Det var speciellt skattningen av antalet döda som väckte uppmärksamhet. 98 000 dödsfall skulle enligt studien vara orsakade av invasionen.

I Burnhams studie från 2006 blev skattningen av antalet dödsfall orsakade av invasionen cirka 655 000 och konfidensintervallet ungefär (393 000; 794 000). Precis som i den tidigare studien var skattningen väldigt mycket högre än de dödsfallssiffror som rapporterades av andra källor, även med hänsyn tagen till den stora osäkerheten. Uppmärksamheten i media och mer eller mindre politiskt färgad kritik från olika håll lät inte vänta på sig, till exempel dömde föga förvånande den dåvarande presidenten George Bush ut studien vid en presskonferens.

Det har även gjorts vetenskapliga granskningar av artiklarna. David Marker gjorde en grundlig genomgång av studien från 2006. Den publicerades i Public Opinion Quarterly 2008, vol. 72, nr 2, pp. 345–363. Artikeln finns fritt tillgänglig på AAPOR:s webbplats. Marker diskuterar metodologiska problem med studien som kan leda till både över- och underskattningar av antalet döda och som indikerar att osäkerheten i skattningarna egentligen är ännu större. Utbildning och kvalitetskontroll av intervjuare, hur man borde ha tagit hänsyn till migration inom och ut ur landet (centralt i ett krigshärjat land), täckningsproblem, hur man hade kunnat minska slumpfelet genom valet av design och att man inte tagit hänsyn till att urvalsannolikheterna varierar är några exempel på vad Marker tar upp.

AAPOR valde att granska studien på begäran av en medlem, enligt en fastställd procedur för hur föregivna brott mot den etiska koden ska behandlas. Man tillsatte en grupp som skulle granska studien och som bland annat begärde information från Burnham, till exempel en kopia av frågeformuläret och en detaljerad beskrivning av hur urvalet gått till. Det gruppen begärde att få veta var endast sådan information som en forskare enligt AAPOR:s etiska regler ska dela med sig av. Burnham svarade att man använt standardmetoder och att all nödvändig information fanns i artikeln. Någon annan information fick inte gruppen, och som det verkar inte heller någon närmare förklaring av varför. AAPOR valde då att kritisera Burnham för att ha brutit mot den etiska koden genom att vägra lämna ifrån sig den begärda informationen. AAPOR:s kritik fick stor uppmärksamhet i amerikansk media, trots att kritiken inte rörde resultaten i studien utan endast Burnhams agerande.

Efter AAPOR:s kritiska uttalande förekom en hel del diskussioner bland medlemmarna kring principiella frågor. Varför la man så mycket tid på just det här fallet när det finns åtskilliga dåliga undersökningar att fördöma där ute? En förklaring till det är förstås att frågan är politiskt känslig och att studien därför blev intensivt debatterad. Och sist men inte minst: Burnham är inte medlem i AAPOR och har aldrig förbundit sig att följa AAPOR:s etiska kod. Är det då ett korrekt förfarande att offentligt kritisera honom för brott mot densamma?

INGEGERD JANSSON

American Association for Public Opinion Research

■ AAPOR är en organisation för alla som har intresse av samhällsundersökningar. Den grundades 1947.

Medlemmarna kommer bland annat från universitet och forskningsinstitut, myndigheter, privata företag och ideella organisationer. Läs mer om AAPOR på www.aapor.org

QVINTENSEN 17

Om konsten att fatta rätt beslut

Fattar rätt beslut? Marvin i Hitchhiker´s Guide to the Galaxy.

■ Ofta fattas strategiska beslut med så kallad ”magkänsla” som enda beslutsstöd. På så sätt fattade företaget där Göran Lande arbetade beslutet att inte gå vidare med ”smarta telefoner”. Ledningen trodde inte på produkten trots statistik som pekade på klart positiv försäljningstrend. Göran Lande: ”Idag, med facit i hand, vet vi att beslutet var felaktigt och ödesdigert för hela företaget”.
GÖRAN LANDE. OM KONSTEN ATT FATTA RÄTT BESLUT. QVINTENSEN 4/2013.

Att översätta frågor så att de fungerar på samma sätt i flera länder

■ Gelaye Hailemichael skriver om en mindre del, om översättning, av sin masteruppsats som hon la fram på Statistiska institutionen vid Stockholms universitet.
GELAYE HAILEMICHAEL. QVINTENSEN 4/2013. TRANSLATIONS: ONE OF THE HEADACHES IN COMPARATIVE SURVEYS.

Får forskare uttala sig offentligt?

■ Intervju med Olle Häggström och Stefan Jansson om svårigheter och dilemman som forskare möter i den offentliga debatten.
QVINTENSEN 4/2013.

Vad är teknisk bevisning i en rättegång?

■ Anders Norgaard förklarar hur analyser görs som leder fram till ”teknisk bevisning”.
QVINTENSEN 1/2013.

Niclas Eriksson visar hur man skriver populärt om en doktorsavhandling

■ Niclas Eriksson börjar i North Dakota, går över till råttgift och landar i ett statistiskt race.
NICLAS ERIKSSON. FARMAKOGENETISKT STYRD LÄKEMEDELSBEHANDLING. QVINTENSEN 4/2012.

USA firade $\pi 10^8$ med paj

■ Dagen då USA:s befolkning nådde 314 159 265 personer.
INGEGERD JANSSON. USA:S DEMOGRAFER FIRAR NERDVANA. QVINTENSEN 3/2012.

Ett alltför ofta negligerat fel

■ Brottsförebyggande rådet om en studie av kodning av brott.
TOLV PROCENT FEL. ANTON FÄRNSTRÖM, SANDRA SANDBERG OCH LEIF PETERSSON. QVINTENSEN 3/2012.

Inte ens en karta är en objektiv bild av verkligheten

■ Stefan Svanström förklarar hur kartor görs och varför en karta är en delvis subjektiv modell av verkligheten och inte en avbild av den.
STEFAN SVANSTRÖM. FÖRSTÅELSE FÖR SKALAN. QVINTENSEN 4/2011.

Oberoende eller kompetens? En fråga om trovärdighet

■ Det finns inget annat statistiskt område som är så genomreglerat som utveckling av nya läkemedel men där förtroendet ändå är så känsligt. Bernhard Huitfeldt tar med oss på en resa med Journal of the American Medical Association och dess försök att upprätthålla förtroendet men som kanske blev en kamp mot en hel yrkeskår.
BERNHARD HUITFELDT. OBEROENDE ELLER KOMPETENS? EN FRÅGA OM TROVÄRDIGHET FÖR STATISTIKER I LÄKEMEDELSINDUSTRIN. QVINTENSEN 4/2011.

Att göra sig förstådd på engelska

■ ”Grammatisk korrekthet ger inte alltid bättre kommunikation”. Intervju med Beyza Björkman.
ATT GÖRA SIG FÖRSTÅDD PÅ ENGELSKA. QVINTENSEN 1/2011.

Texter, texter, texter – vi översköljs av dem, såväl på jobbet som privat. Vissa läser vi med intresse, andra suckar vi över: ”Vad vill de säga egentligen?” ”Varför ska jag läsa det här?”. Men hur undviker man att själv producera sådana texter? Språkvårdaren Nathalie Parès ger oss här några råd om vad man bör tänka på när man skriver.

Nathalie Parès.

Skrivråd för text

”Jag kan inte skriva” bekände en kollega för mig, redan första gången vi träffades. Jag hade precis fått tjänsten som språkvårdare på Brottsförebyggande rådet (Brå) och blev extra nyfiken.

Vad var det min kollega menade? Ja, det var naturligtvis inte att det är så ruskigt svårt att stava rätt, eller att det fanns några praktiska, tekniska orsaker till den där känslan. (Kollegan är en medelålders akademiker med svenska som modersmål, utan tendenser till dyslexi eller andra relevanta handikapp.) Nej, det handlade snarare om dåligt självförtroende, men också om svårigheten att skriva så att läsaren förstår, att skriva så att texten blir läst, att skriva så att själva syftet med texten blir uppfyllt.

Hur gör man då, för att skriva begripligt, intresseväckande och effektivt? På ett sätt är utmaningen densamma som när man ska föreläsa – att fånga gruppens intresse och förmedla det man vill på ett tydligt sätt utan att bli för långrandig – men med den avgörande skillnaden att man i skrift inte löpande kan anpassa kommunikationen till gruppens reaktioner och frågor. Dem måste man gissa sig till och bemöta i förväg. Någon patentlösning på hur man lyckas gör jag inte anspråk på att presentera här, men några goda råd vill jag gärna bjuda på!

Fyra grundläggande frågor

Det man bör ta tag i allra först, men många glömmer, är att tänka till ordentligt på vilka man vänder sig till, vad de kan ha för intresse av att läsa texten, i vilket sammanhang den uppträder och vad man själv vill uppnå. Det utgör sedan utgångspunkten för själva skrivandet, som jag återkommer till lite längre fram…

Vi börjar från början, när man sitter där med ett tomt papper eller en ny, tom fil i datorn. För vissa känns det då omöjligt att ens få ur sig en första mening, innan den måste omformuleras och slipas till, gång på gång. Tvivlet på att det någonsin ska bli en text infinner sig. För andra flödar orden ur dem, utan egentlig svårighet, men liksom planlöst, utan riktning och struktur. Även då kan resultatet bli mindre lyckat.

Innan man ens börjar skriva finns det därför några frågor man bör ställa sig:

1) Vilka vänder jag mig till?
2) Vad motiverar dem att läsa min text?
3) När, var och hur kommer texten att läsas?
4) Vad är mitt syfte med texten?

Om man inte är klar över svaren, är risken större att den text man skriver mest bara mal på, och att den som läser den ledsnar, någonstans halvvägs igenom.

Vilka vänder du dig till?

Fråga dig vilka som troligen kommer att läsa texten, och vad du vet om dem. Vad är de intresserade av och vad har de för bakgrundskunskaper? Ska du betrakta textens samtliga potentiella läsare som din målgrupp, eller ska du välja ut en delmängd att fokusera på? Om läsarna är många, och det dessutom inte rör sig om en särskilt enhetlig församling, kan den här typen av frågor vara extra svåra att besvara. Men detta förberedande tankearbete blir därmed inte desto mindre

byggare

TECKNING: JAN STJERNKVIST

viktigt att göra – liksom att påminna sig om svaren under skrivandets gång.

Vad motiverar dem att läsa din text?

Som läsare vill man naturligtvis gärna uppleva texten som matnyttig, att man får ut någonting av den – och vad det är läsaren vill få ut, det växlar. Ibland vill man bara ha en stunds tidsfördriv, ibland vill man lära sig någonting nytt, ibland är man ute efter svaret på en bestämd fråga. Det gäller att du som skribent är klar över vad som motiverar dina tänkta läsare att ta sig an din text, vad som är deras syfte med läsningen. Och sedan att underlätta läsningen för dem.

När, var och hur kommer texten att läsas?

I vilken form kommer din text att möta läsarens blick? Rör det sig om en rapport, en vetenskaplig artikel, en kvartalsskrift, en webbtext eller kanske ett brev? Är det en artikel i Metro eller en ”understreckare” i Svenska Dagbladet? Och var befinner sig läsaren? Hemma i soffan, på jobbet eller på bussen? Under stress eller i lugn och ro? Lässituationen, texttypen och sammanhanget i stort har betydelse för läsarens tålamod och förväntningar på texten. Vissa texter bör vara rappa och kortfattade, andra inte.

Vad är ditt syfte med texten?

Vad vill du uppnå med din text?
Är det kanske bara att förmedla en viss information, att läsaren ska få veta någonting? Eller vill du påverka din läsare till att känna någonting specifikt inför det du berättar? Vad ska läsaren tycka om informationen? Vill du få läsaren att dela din åsikt om det du beskriver? Vill du kanske rent av påverka din läsare till att göra någonting, som att delta i en enkätundersökning eller ändra ett invant arbetssätt? Om du inte är klar över det själv, är risken större att du inte uppnår ditt mål.

Bearbeta texten, på olika nivåer

När du väl ringat in svaren på dessa grundläggande frågor, kan själva skrivandet och textbearbetningen ta vid. Och nu gäller det att låta svaren påverka textens utformning, på olika nivåer i texten – såväl vad gäller helheten som på detaljnivå!

Omfång och urval

Hur omfattande bör texten vara? Ofta ger sammanhanget avgränsningen. En generell rekommendation, när det gäller facktext, är dock att hålla sig kort. Att inte berätta allt man kan och vet. Bara välja ut det som är relevant, med tanke på de grundläggande frågorna jag nämnde. Vad behöver alltså tas med för att texten ska motsvara den aktuella målgruppens behov och förväntningar, i det aktuella sammanhanget, och för att du ska uppnå ditt syfte med texten? Skala bort allt annat – stryk det helt eller lägg det i en bilaga!

Disposition och struktur

I vilken ordning ska du då ta upp det du vill berätta? Fråga dig allra först vad som intresserar målgruppen. Vilken information söker de? Se till att de får svar på sina frågor tidigt – facktexter är inte deckare! Bakgrundsbeskrivningar kan

visserligen vara viktiga för att läsaren ska kunna sätta in en information i sitt sammanhang och värdera den, men i regel kan man ändå ge en första blänkare om nyheten – det för läsaren mest intressanta – redan tidigare.

Ge också din läsare möjlighet att överblicka innehållet. Var generös med mellanrubriker, så att man kan skumma igenom din text på ytan, greppa hur den är uppbyggd och välja de bitar man eventuellt vill fördjupa sig i. Komplettera gärna med en innehållsförteckning i början (om det passar texttypen), så att man dessutom kan granska strukturen lättare, utan att ens bläddra igenom texten.

I många sammanhang finns en mer eller mindre given disposition, till exempel när man skriver en vetenskaplig artikel. Men även då finns det mycket att tänka på. Vad ska du ta upp i det introducerande avsnittet? Vad kan läsaren till just den här tidskriften tänkas känna till om ditt ämne?

När man skriver om statistik finns det dessutom ytterligare faktorer som påverkar hur din text ska se ut, som regeln att individbaserad officiell statistik ska vara könsuppdelad, enligt förordningen om officiell statistik.

Styckeindelning och meningslängd

Innan jag ens nämnde mellanrubriker borde jag kanske ha påtalat vikten av att stycka upp texten också (om inte annat blir det svårt att peta in mellanrubrikerna annars ...). En kompakt textmassa utan luft kan ge ett så ogenomträngligt intryck att man som läsare tappar lusten att ge sig i kast med den. Eller att man i vilket fall inte kommer igenom den. Läsaren behöver andrum!

Hur långa styckena bör vara beror helt på innehållet. Bryt för nytt stycke när du kommer in på ett nytt ämne eller en ny tankegång. "En tanke, ett stycke" brukar vara ett bra ledord. Inled stycket med det som är kärnmeningen, det viktigaste, i det stycket. Därefter utvecklar eller fördjupar du resonemanget. Kärnmeningen bjuder in läsaren till det som avhandlas i stycket.

Bryt för nytt stycke när du kommer in på ett nytt ämne eller en ny tankegång. "En tanke, ett stycke" brukar vara ett bra ledord.

Generellt kan sägas att texten blir tyngre att läsa ju längre både styckena och meningarna är. Långa meningar, som sträcker sig över flera rader, är naturligtvis tyngre att läsa än korta. Än tyngre blir det om läsarens tvingas hålla alltför mycket information i minnet innan predikatet kommer! Det är det som talar om vad som händer i meningen, och det är först när man fått veta det som resten av informationen har någonting att landa på.

Ordval

Vissa ord, som kan försvåra läsningen, bör man vara sparsam med. Men vilka är de? Många tänker kanske i första hand på facktermer. De kan visserligen vara det absolut mest effektiva sättet att kommunicera, experter emellan, men utestänger ju dem som inte hör till skrået. Undvik dem gärna om du har en blandad målgrupp, där bara vissa behärskar fackterminologin, eller se till att förklara termerna första gången de förekommer i texten.

Detta gäller även eventuella viktiga förkortningar som du vill kunna använda i texten, såsom ANOVA eller πps. Behöver de finnas med i texten, gäller det att förklara dem, i själva texten eller i en separat termlista. En del förkortningar kan dessutom stå för olika saker, såsom KI, vilket kan leda till missförstånd. Överlag bör du alltså undvika förkortningar i möjligaste mån, även om det går snabbare för dig att skriva dem. För läsaren går det ju till exempel knappast snabbare att läsa ut förkortningen m.a.o. än frasen med andra ord, och för den som inte ens förstår förkortningen tar det definitivt längre tid.

Ett tredje exempel på sådant som kan göra din text onödigt svår att ta till sig är ord som påverkar tonen i texten. Ålderdomliga eller byråkratspråkliga ord (som ehuru, avseende, bifalles eller därvid) ger texten en prägel som håller läsaren på armslängds avstånd.

Vissa ord kan i gengäld underlätta läsningen av en text, genom att bidra till att göra sambanden tydliga. Så kallade sambandsmarkörer kan exempelvis vara eftersom, däremot, trots att, det tyder på eller samtidigt. Tack vare dem visar du läsaren hur du resonerar.

Underlätta skrivandet

Hur väl man än tycker sig ha planerat sin text är det lätt hänt att köra fast under skrivandets

Exempel 1: Månadsrapport »Anmälda brott»

Du ska skriva en ny månadsrapport om ”Anmälda brott”. Vilka vänder sig rapporten till? Är det läsare som vill ha en mycket detaljerad redogörelse över anmälda brott, uppdelade i olika kategorier? Och varför vill de ha det? Eller är det läsare som bara vill veta något om de övergripande trenderna och de stora dragen?

Behöver den tänkta läsaren motiveras? I så fall kanske rapporten ska lyfta fram något överraskande.

Är läsaren en person som läser ”Anmälda brott” varje månad och vill att den ska ha samma disposition varje gång? En läsare som bara vill kolla ”sina” siffror, rutinmässigt, varje månad? Eller finns det flera olika grupper av läsare, som ställer olika krav på rapporten?

En kanske inte ovanlig tanke är att en månadsrapport om anmälda brott inte ska ha något uttalat syfte, utan bara redovisa brotten som de är. Men vad betyder egentligen det? Låt oss säga att fylleri till sjöss förs till Blekinge län om domen fälldes i Blekinge, även om inte brottet utfördes i Blekinge. Då kanske Blekinge får fler fall än andra län. Vad är syftet med statistiken i rapporten? Att redovisa antalet fall som administrativt hänför sig till olika län, eller antalet fall som utfördes i länet? Ett annat exempel är den statistik som tycks visa att fler människor dör på nyårsafton än på andra dagar under året. Det beror på alla de fall då ett specifikt dödsdatum faktiskt saknas, eftersom den 31 december då skrivs in i rutan för dödsdatum. Vad är syftet med att redovisa den statistiken? Kanske är det då mer intressant att redovisa statistiska skattningar av verkliga dödsdatum?

gång. Därför vill jag gärna avsluta med några råd som kan underlätta textbearbetningen.

Se styckena som byggklossar

Ett knep om du vill ändra textens disposition, är att först sätta beskrivande mellanrubriker över varje stycke, oavsett om du sedan vill behålla rubrikerna i den färdiga texten. Det blir lättare att stuva om bland ”byggklossarna” (styckena) när du snabbt kan överblicka vad de handlar om.

Om det inte funkar att beskriva ett stycke med en kortfattad rubrik, är problemet förmodligen att stycket är för spretigt innehållsligt och behöver bearbetas. Denna lilla övning kan alltså också hjälpa dig att slipa på styckeindelningen.

Om det inte funkar att beskriva ett stycke med en kortfattad rubrik, är problemet förmodligen att stycket är för spretigt...

Överblicka helheten

Att ta ett helhetsgrepp och bedöma om texten håller är inte alltid det lättaste. Ett effektivt verktyg för att bedöma om du skapat en logisk struktur i texten är att skapa en innehållsförteckning, där du enkelt kan överblicka rubrikerna. Om texten som slutprodukt inte ska ha en innehållsförteckning är det bara att plocka bort den när du inte behöver den längre. Så gjorde jag själv när jag skrev den här artikeln, eftersom jag ville överblicka helheten under skrivandets gång.

Glöm inte frågorna

Återvänd emellanåt till de grundläggande frågorna under skrivandets gång. Använd dem som måttstock: ”Funkar detta till den målgrupp jag vill nå? I det sammanhanget? Uppnår jag mitt syfte?”

Låt skrivandet ta tid, men sätt igång genast

Inspiration kommer sällan på beställning. Vänta inte på den! När du väl börjat skriva gäller det dock att låta skrivandet få ta sin tid, så att du kan lägga ifrån dig texten och sova på saken för att lättare se din text med färska ögon. Låt också någon annan läsa din text – innan du betraktar den som färdig! Be inte bara om korrekturläsning på slutet, utan om tidiga synpunkter på upplägget som helhet. Det är idealiskt om du hittar någon som också känner till den målgrupp som texten vänder sig till, men även en annan "hjälpläsare" kan komma med värdefulla synpunkter och peka ut textens svaga punkter.

Varsågod, nu är det din tur!

Kanske har jag med denna text påverkat dig att börja se skrivandet som ett bygge, som delvis går att underlätta genom god planering. Den lilla missionären i mig vill gärna belysa att ”välskrivna texter” är någonting mer än en hop rättstavade ord i grammatiskt korrekta meningar. De växer fram i händerna på medvetna skribenter, som strävar efter att anpassa sina texter till målgruppen, sammanhanget och syftet.

NATHALIE PARÈS,
BROTTSFÖREBYGGANDE RÅDET

Exempel 2: Vem skriver jag för?

Ett vanligt nybörjarfel bland nya forskare är att inte tänka på vad som motiverar läsaren. Svaret på frågorna ”Varför är det här vetenskapliga området intressant?” och ”Vad i just den här artikeln är intressant?” göms undan långt in i artikeln. Att stimulera läsarens motivation är faktiskt nästan lika viktigt som det vetenskapliga innehållet i artikeln.

Frågan ”Vilka jag vänder mig till?” kan tyckas enkel när man skriver för en vetenskaplig tidskrift. Men "forskare" [eller något annat vagt] duger inte som svar. Ringa in mer exakt, på en mer detaljeratd nivå, vilka det är som läser till exempel Journal of the Royal Statistical Society, Series C.

Många som skriver vetenskapliga artiklar skriver först artikeln och funderar sedan på vilken tidskrift den ska skickas till. Ibland kan man tänka tvärtom: "Vilken typ av tidskrift ska vi vända oss till? Vill vi skriva för en tidskrift som är inriktad på kvantitativ kriminologi eller en som är inriktad på statistiska metoder i specifika tillämpningar, eller kanske en som vill lyfta fram de generella dragen i statistiska metoder?"

Vad är mitt syfte med den här artikeln som jag håller på att skriva? Vill jag bara redovisa några vetenskapliga resultat? Eller vill jag verka för en förändring? Svaret på de frågorna påverkar vilka resultat jag väljer att föra fram i artikeln.

Har du flyttat?

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

Svenska statistikfrämjandet

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 Carl Åkerlindh 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.

Det här handlar om ett verkligt fall. Jag är inte själv inblandad, men jag har av en slump råkat få lite inblick i fallet, och jag vill här återge vissa resonemang som jag lagt märke till.

För att inte chikanera de inblandade har jag i denna framställning avidentifierat dem. Detta gör att jag inte kan ange källor till de citat jag ger.

JAN WRETMAN:

Ett upphandlings

Bakgrunden är följande: En statlig myndighet, *M*, behöver för sin verksamhet viss månatlig statistik som ska belysa vissa saker i samhället. Man kan inte samla in statistiken själv utan går ut med en så kallad upphandling. Flera anbud från olika undersökningsföretag kommer in. Efter granskning kvarstår anbuden från de båda undersökningsföretagen *A* och *B* som intressanta för myndigheten.

Myndigheten *M* bestämmer sig slutligen för att anta anbudet från företaget *A*, vilket framkallar invändningar från företaget *B*, som överklagar i diverse rättsliga instanser. Myndigheten *M*:s beslut kvarstår dock. Det är skriftväxlingen i det här ärendet, som jag fått viss inblick i.

Anbuden från *A* och *B* ligger ekonomiskt nära varandra. Det är alltså inte av ekonomiska skäl man fastnat för *A*, utan valet sker på andra grunder. Det kan vara intressant (ibland häpnadsväckande) för en statistiker att se hur resonemanget har gått.

Metodskillnader mellan de två anbuden

Företaget *A* erbjuder sig att leverera statistik genom en så kallad *webbpanelundersökning*. Generellt sker urvalet vid en sådan undersökning i två faser: (1) Först rekryteras en panel bestående av personer, som förklarar sig villiga att medverka som respondenter i flera framtida undersökningar via webben, ifall de blir utvalda. (2) För varje specifik undersökning görs sedan ett urval av personer från den givna panelen. I *A*:s fall köper man en färdig panel från företaget CINT, som är en välkänd leverantör av sådana paneler. Från denna köpta panel tänker man sedan varje månad göra ett urval av personer.

Hur den köpta panelen har rekryterats har man ingen detaljerad kunskap om, men av de papper jag sett har jag förstått att panelen har bildats genom hopslagning av omkring hundra olika paneler. Sådana här paneler brukar vanligtvis erhållas genom självrekrytering eller annat icke-sannolikhetsbaserat urvalsförfarande, och det finns anledning att tro att så är fallet även här. Själva datainsamlingen sker genom webbenkät till de personer som väljs ut från panelen.

Företaget *B* erbjuder sig att leverera statistik genom en undersökning av mer traditionellt slag. De personer som ska tillfrågas varje månad är här ett *personurval från SPAR* (Statens personadressregister, vilket omfattar alla som är folkbokförda i Sverige). Man skaffar sedan telefonnummer för de utvalda personerna; i första hand fasta abonnemang, i andra hand mobila. Datainsamlingen tänks sedan ske genom telefonintervju.

Ur statistisk metodsynpunkt skiljer sig de båda anbudsgivarnas förslag från varandra i två viktiga avseenden, nämligen i fråga om *urvalsmetod* och i fråga om *datainsamlingsmetod*. Vad gäller urval är det webbpanelundersökning gentemot personurval från ett personregister. Vad gäller datainsamling är det webbenkät gentemot telefonintervju.

Ett viktigt begrepp när man ska diskutera urvalsförfarande är begreppet *sannolikhetsurval*. Med detta menas (1) att det finns en entydigt definierad population från vilken urval ska göras med användande av en viss urvalsram, och (2) att urvalet görs med användande av en av oss själva kontrollerad slumpmekanism, så att varje individ i urvalsramen får en känd sannolikhet, större än noll, att bli utvald. Urvalsförfaranden som inte uppfyller villkoren för sannolikhetsurval faller under rubriken *icke-sannolikhetsurval*. Huruvida ett urvalsförfarande kan betraktas som sannolikhetsurval eller ej är en viktig sak, eftersom det bara är i fallet med sannolikhetsurval som det går att beräkna konfidensintervall, såvida man inte har en välgrundad statistisk modell för undersökningsvariablerna.

Hur är det nu i de två aktuella fallen, kan något av de två urvalsförfarandena ses som exempel på sannolikhetsurval? I fallet *A* kan man direkt säga att det *inte* är fråga om sannolikhetsurval. Visserligen säger *A* i en skrivelse till rättsliga instanser att urvalet av personer från

ärende

panelen är ”slumpmässigt”, men eftersom själva panelrekryteringen antagligen inte skett genom sannolikhetsurval, så kan urvalsförfarandet som helhet inte betraktas som sannolikhetsurval från en klart definierad population. Vidare brukar vanligen bortfallet vid panelrekrytering och urval från panelen vara så stort att urvalet även av denna anledning inte kan betraktas som sannolikhetsurval. Det går då inte att beräkna konfidensintervall, såvida man inte har en välgrundad statistisk modell för hur bortfallet uppkommer. Det verkar som om man i detta fall inte har någon sådan modell.

Även i fallet *B* är det osäkert om man kan tala om sannolikhetsurval. Visserligen är urvalet från början ett sannolikhetsurval, men om man därefter har ett stort bortfall, så kan den kvarstående mängden av respondenter inte utan vidare betraktas som erhållen genom sannolikhetsurval. Jag har tolkat de uppgifter jag sett som att man räknar med en svarsfrekvens på cirka 45 procent, alltså ett bortfall på cirka 55 procent. Det är alltså även här mycket tveksamt ifall det är meningsfullt att försöka beräkna konfidensintervall.

Uttalanden från myndigheten M

De två citat som här följer är hämtade från papper där myndigheten *M* kommenterar sitt val av *A* som statistikleverantör. Sin allmänna syn på statistisk undersökningsmetodik ger myndigheten i följande uttalande:

> Alla urvalsundersökningar är alltid behäftade med kvalitetsbrister. Beroende på val av undersökningsmetod uppstår kvalitetsproblem på lite olika sätt, men de finns alltid där.
>
> Webbpanelundersökningar är behäftade med vissa problem, telefonintervjuundersökningar med andra. Den tekniskt mest riktiga metoden är troligen postala undersökningen mot totalbefolkningen, men där uppstår andra komplikationer, bl a kostnaden.

Vi kan konstatera att de för en statistiker viktigaste frågorna, de om urvalsförfarande och möjligheter att dra slutsatser från stickprovet till populationen, inte nämns vare sig i citatet ovan eller någon annanstans i det citerade dokumentet. Varför postenkäter skulle vara den ”tekniskt mest riktiga metoden” förstår jag inte.

I information inför sin kommande upphandling skrev myndigheten att ”varje leverantör får ta ansvar för att levererade data håller önskad kvalitet, vilket i anbud specificeras med avseende på konfidensintervall”. Vid jämförelse mellan de anbud, som sedan kom in, lade myndigheten stor vikt vid anbudsgivarnas beskrivning av konfidensintervall. Där gav myndigheten högst betyg åt *A*, högre betyg än vad man gav åt *B*. Myndigheten sade:

> Vår bedömning är att *A* gjort en omfattande, noggrann och pedagogisk sammanställning av beräkningar av konfidensintervall […]. Vår bedömning är att *B* gjort en mer kortfattad sammanställning med färre exempel.

Den viktigaste frågan, den om det överhuvudtaget är *möjligt* att beräkna tillförlitliga konfidensintervall i de undersökningar som det här är frågan om, var man från myndighetens sida uppenbart omedveten om.

De båda uttalanden som här citerats tyder enligt min mening på brist på statistisk medvetenhet hos myndigheten *M*. Detta ska inte uppfattas som kritik. Vad jag menar är att när en myndighet saknar egen kompetens i en bedömningsfråga, så kan man söka hjälp av utomstående experter, i detta fall till exempel Statistiska centralbyrån.

Uttalanden från undersökningsföretaget A

Som redan nämnts kom *B* med invändningar mot att myndigheten valt *A* som statistikleverantör. Den ur statistisk metodsynpunkt mest intressanta invändningen har att göra med om det är möjligt att beräkna tillförlitliga konfidensintervall vid den typ av undersökningar som *A* föreslår. Som också sagts ovan hänger denna fråga ihop med frågan om data från en webbpanelundersökning kan betraktas som erhållna genom sannolikhetsurval eller ej.

I skrivelse till Förvaltningsrätten hävdar *A* att urvalet till deras webbpanelundersökning är ett sannolikhetsurval:

> Respondenterna är INTE självrekryterade […] utan enkäten skickas slumpmässigt ut till ett urval från de 450 000 individer som är panelmedlemmar i vår samarbetspartners CINTs paneldatabas.

A glömmer här att urvalet vid en webbpanelundersökning görs i två faser. I första fasen »»»

»»» ETT UPPHANDLINGSÄRENDE

rekryteras en panel från populationen, i andra fasen görs ett urval från panelen till den specifika undersökningen. Om panelen erhållits genom självrekrytering, så blir urvalsförfarandet i sin helhet ett icke-sannolikhetsurval. Då hjälper det inte att urvalet från panelen till den specifika undersökningen gjorts med sannolikhetsurval.

»Jag finner detta fullständigt obegripligt. Och uppenbarligen har varken myndigheten eller rättsinstanserna haft sakkunskap att bedöma uttalanden av detta slag.»

Det faktum att panelen som *A* erhållit från CINT utgör en hopslagning av cirka hundra paneler, som tillkommit på olika sätt kommenteras av *A* på följande sätt:

> Detta ger att respondenterna i en undersökning kommer från olika paneler som ägs av olika företag med olika rekryteringssätt och olika ersättningslösningar. Detta medför att det naturliga bias som finns i alla paneler reduceras.

Att eventuell skevhet i en panel skulle bli mindre, ifall panelen är hopsatt av ett stort antal mindre paneler, är ett påstående som, såvitt jag förstår, saknar empirisk grund.

Beträffande sannolikhetsurval och konfidensintervall säger *A* slutligen:

> Eftersom hela 36 000 respondenter på årsbasis väljs ut slumpmässigt från hela 450 000 svenska medlemmar kan urvalet sägs vara nära sannolikhetsurval. Med ett så stort urval som 36 000 st så är de variabler som mäts normalfördelade och därför möjligt att beräkna konfidensintervall på. Urvalet blir därför inte skevt i den utsträckning som den kritik säger som riktas mot självrekryterade paneler (se Stora talens lag).

I det citerade uttalandet sägs, såvitt jag förstår, följande:

- Om ett urval är stort, så blir det ”nära” ett sannolikhetsurval (oavsett hur urvalet gjorts).
- Om ett urval är stort, så blir undersökningsvariablerna normalfördelade.
- Då blir det möjligt att beräkna konfidensintervall.
- Då kommer självrekrytering inte att orsaka skevhet i större utsträckning.
- Allt detta följer av Stora talens lag.

Jag finner detta fullständigt obegripligt. Och uppenbarligen har varken myndigheten eller rättsinstanserna haft sakkunskap att bedöma uttalanden av detta slag.

En avvikelse från ämnet

På några ställen i de dokument jag studerat har jag sett hänvisning till en skrivelse från Eurostat, Riktlinjer för europeisk statistik, där det i den svenska översättningen bland annat sägs att

> […] insamlingen av statistiska uppgifter ska uppfylla höga krav på opartiskhet, insyn, tillförlitlighet, objektivitet, vetenskapligt oberoende, kostnadseffektivitet och insynsskydd för statistiska uppgifter.

Här finns en för mig obegriplig term, nämligen ”vetenskapligt oberoende”. Vad är det för någonting? Det är tydligen inte detsamma som opartiskhet och objektivitet, för dessa saker har ju redan räknats upp. Jag gissar att man kanske menar att insamlingen av uppgifter ska vila på någon sorts vetenskaplig grund. Men då borde väl datainsamlingen vara i någon mening *beroende* av vetenskapen snarare än oberoende?

JAN WRETMAN

Inte helt på pricken

Det citerade stycket, som börjar "insamlingen av statistiska uppgifter ska uppfylla höga krav på opartiskhet” kommer i stort sett från Princip 6 i Riktlinjer för europeisk statistik. Det verkar vara en översättning som inte är helt på pricken. I den version som fastställdes 28 september 2011 står det ”De statistikansvariga myndigheterna utvecklar, framställer och sprider europeisk statistik på ett sätt som tar hänsyn till vetenskapligt oberoende, objektivitet, yrkesmässighet, öppenhet och insyn samt jämlik behandling av alla användare”. Jämför med den engelska versionen: ”Statistical authorities develop, produce and disseminate European Statistics respecting scientific independence and in an objective, professional and transparent manner in which all users are treated equitably”.

DAN HEDLIN

Om web

Webbpanelundersökningar har på senare år fått stort genomslag bland marknads- och opinionsundersökare. Med en webbpanel avses ett register eller en databas över personer som förklarat sig villiga att delta i webbundersökningar. En webbpanelundersökning är en undersökning med urval från en webbpanel. Sådana undersökningar kan synas ha praktiska fördelar framför traditionella undersökningar baserade på sannolikhetsurval. Men en fråga är om deras kvalitet räcker.

Webbpaneler ger ett relativt enkelt, snabbt och billigt sätt att samla in stora mängder data. Inga intervjuare är inblandade, och det blir inga kostnader för porto eller hantering av postförsändelser med pappersblanketter. Uppgiftslämnarbördan kan i en mening sägas bli lindrig, genom att respondenterna på eget bevåg väljer att gå med i webbpanelerna och får rutin på att medverka i undersökningar. En stor webbpanel ger därtill en möjlighet att få tag i en liten grupp, t.ex. båtägare, utan att behöva filtrera fram denna i den specifika undersökningen. Det förutsätter dock att den uppgiften samlats in för alla i webbpanelen.

Är det då dags att över lag ta till webbpanelundersökningar för officiell statistik? Nej, knappast i närtid, eftersom metodproblemen med webbpanelundersökningar inte kan anses ha klarats ut väl nog för det. Huvudproblemet är hur slutsatser ska kunna dras från dem som deltar i undersökningen upp till målpopulationen, dvs. möjligheten att göra inferens från svarsmängd till population. Webbpaneler kan inte företräda människor som inte är ute på webben. Denna

bpanelundersökningar

undertäckning kan snedvrida resultaten av undersökningar om företeelser som är särskilt frekventa i grupper av äldre, lågutbildade eller utrikes födda. I det följande fokuseras dock på selektionsproblematiken.

Två typer av webbpanelundersökningar

Det förekommer att webbpaneler rekryteras utifrån sannolikhetsurval. Webbpanelundersökningen har då två slumpmekanismer: dels en välkontrollerad där urvalet görs, dels en bortom effektiv kontroll där bortfallet uppstår. Den s.k. kumulativa deltagarandelen, som grovt kan sägas motsvara svarsandelen i en traditionell undersökning, ligger vanligen under 10 procent, dvs. väsentligt lägre än svarsandelarna i traditionella undersökningar för samhällsstatistik. Det mycket stora bortfallet riskerar leda till betydande systematiska fel i statistikresultaten. Denna typ av webbpanelundersökning kan ändå i princip uppfattas som en traditionell undersökning och är då hanterbar med statistisk metodik för urval och bortfallskompensation. För att detta ska kunna hålla behöver dock det vanligen mycket stora bortfallet hållas under någorlunda kontroll. Den nederländske professorn Jelke Bethlehem [1] menar att denna undersökningstyp kan vara tänkbar för officiell statistik under en del ambitiösa förutsättningar, såsom att bortfall förebyggs i möjlig mån både vid panelrekrytering och i den specifika undersökningen, att hjälpinformation används för att justera för bortfallsfel, och att panelen förnyas regelbundet. Han betonar att det är en komplex uppgift att uppfylla förutsättningarna och att det kan vara resurskrävande.

»Samhällsstatistik ligger till grund för viktiga beslut av myndigheter och andra aktörer, och då behöver den vara trovärdig genom att hålla vetenskapligt. »

Webbpanelundersökningar baseras dock ofta på icke-sannolikhetsurval, t.ex. genom rekrytering via reklam på internet. Panelen som erhålls är då självselekterad: respondenterna har själva valt att delta i panelen efter en bred inbjudan, och urvalsförfarandet är okontrollerat. Information saknas både om vilka personer som inte varit medvetna om att valet fanns att delta i panelen och om vilka personer som valt att inte vara med. Bortfallet är alltså okänt, eller egentligen odefinierat. Risken är att panelen kommer att bestå av ”proffstyckare” eller en högljudd minoritet. Denna typ av webbpanelundersökning med självrekrytering är alltså principiellt annorlunda och kan knappast ses som ens i princip jämförlig med en traditionell undersökning. Det hjälper inte att ett sannolikhetsurval dras från panelen för den enskilda undersökningen. Undersökningar utifrån självrekryterade paneler saknar i nuläget en sannolikhetsteoretisk grund, och tilltron till dem bygger snarare på empiri från enskilda undersökningar.

Kvalitetsmått

För icke-sannolikhetsurval saknar det vanliga osäkerhetsmåttet (skattat urvalsfel) tolkning, varför det inte heller kan användas för att konstruera konfidensintervall. Följden blir att signifikanta skillnader mellan grupper eller över tid inte kan urskiljas trovärdigt. Detta gör att det för statistik baserad på icke-sannolikhetsurval normalt inte är möjligt att säkra en tillräcklig kvalitet för att statistiken ska hålla som trovärdigt beslutsunderlag.

Ett alternativ till konfidensintervall är att ta fram ett slags bedömningsintervall. Detta måste dock baseras på en modell, ofta utifrån en bayesiansk ansats, och blir starkt beroende av vilka antaganden som görs om urvalet och populationen.

Jag var med och tillsatte Surveyföreningens webbpanelkommitté 2009. Efter ett gediget arbete avlämnade kommittén i fjol en mycket läsvärd rapport [2]. Den belyser kvalitetsfrågor och presenterar mått som på olika sätt kan beskriva kvalitet i webbpanelundersökningar. Några numeriska beskrivningsmått föreslogs för kvalitetsbedömning, och därutöver rekommenderades verbala beskrivningar av rekrytering, urval och skattning m.m. Det finns dock inte underlag ännu för att säga hur väl de olika måtten predicerar kvaliteten i en webbpanelundersökning.

Empiriska resultat

Den teoretiska grunden är alltså svag för den typ av icke-sannolikhetsurval som ofta görs för webbpanelundersökningar. Hur ser då empirin ut? Ett väsentligt empiriskt resultat om webbpaneler finns i den s.k. Stanford-studien [3]. Där jämfördes resultat från några amerikanska webbpanelundersökningar och från sannolikhetsbaserade undersökningar med riktmärken »»»

»»» OM WEBBPANEL-UNDERSÖKNINGAR

från tillförlitliga källor. Det visade sig att webbpanelundersökningarna konsekvent hade sämre tillförlitlighet, även efter bortfallsjusterande vägning, och att deras tillförlitlighet varierade mer. Undersökningarna som byggde på sannolikhetsurval fungerade däremot väl, även vid relativt stora bortfall. Det verkar alltså även vid stort bortfall ändå vara en avgörande fördel att utgå från sannolikhetsurval. I [4] jämfördes resultat från webbpanelundersökningar med kända riktmärken. Jämförelserna avsåg dels en sannolikhetsbaserad webbpanelundersökning, dels en specifik undersökning genomförd utifrån nitton olika självrekryterade webbpaneler. Det systematiska felet var för fem av sex riktmärken mindre för den sannolikhetsbaserade undersökningen än för genomsnittet av de nitton andra undersökningarna. I [5] presenterades en metaanalys av åtta självrekryterade webbpaneler avseende effekten av vägning för att reducera skevhet på grund av undertäckning och självselektion. Justeringarna eliminerade som högst ca 60 procent av det systematiska felet, och för vissa variabler ökade felet.

Webbpanelbaserade väljarbarometrar har dock många gånger gett relativt träffsäkra resultat, både i Sverige och utomlands [6]. Men väljarbarometrar är en speciell typ av undersökningar. De har en enda huvudfråga, vilken kan betraktas som enkel, väsentligen vilket parti respondenten skulle rösta på om det vore val i dag. Justerande vägning kan göras med hjälp av de vanliga bakgrundsvariablerna, men det finns även möjlighet att fråga hur respondenten röstade i förra riksdagsvalet och använda detta för s.k. partivägning. I Sverige är dessutom valdeltagandet högt, vilket gör det lättare att jämföra och kalibrera väljarbarometrar mot valresultatet. Därtill finns det många andra konkurrerande barometrar att jämföra med. Att lyckas med en väljarbarometer, där det finns en tydlig huvudvariabel och gott om hjälpinformation, ger därför inte någon garanti för framgång med helt andra typer av undersökningar. Många av de undersökningar som genomförs för att ta fram samhällsstatistik avser dessutom kvantitativa variabler, som kan vara snedfördelade, vilket ökar risken för stora systematiska fel.

Metodutveckling

En intressant frågeställning är om det går att dra nytta av fördelarna med en webbpanelundersökning utan att behöva ge upp en kontrollerad inferenssituation, se [7]. Kombinationer av webbpanelundersökningar och traditionella sannolikhetsurval kan då möjligen vara en lösning. Tyngdpunkten flyttas då från en enskild undersökning som en fristående enhet till integration av data från flera olika källor. En sådan utveckling finns när det gäller kombinationer av urvalsundersökningar och registerbaserade undersökningar, och kan skönjas även för kombinationer av nya datakällor (under benämningen Big Data) och urvalsundersökningar eller registerbaserade undersökningar. Den vetenskapliga grunden är dock relativt svag för hur dessa kombinationer ska göras på bästa sätt. Olika kombinationsansatser kan vara tänkbara att pröva. I sammanhang där webbpanelundersökningar kan tänkas komma ifråga, kan två parallella undersökningar genomföras: en traditionell, sannolikhetsbaserad undersökning och en webbpanelundersökning. Resultaten kan användas för att utröna graden av skillnad i skattningarna mellan de två ansatserna. Dessutom skulle den väletablerade, sannolikhetsbaserade undersökningen kunna utnyttjas för att ”överföra information” till webbpanelundersökningen eller kommande sådana. Notera dock att ansatserna kan komma att reducera fördelarna med webbpanelundersökningar, att de är enkla, snabba och billiga.

Samhällsstatistik ligger till grund för viktiga beslut av myndigheter och andra aktörer, och då behöver den vara trovärdig genom att hålla vetenskapligt. För att detta ska gå med webbpanelundersökningar förutsätts att ett passande vetenskapligt ramverk tas fram. Relativt mycket forskning har bedrivits på senare år inom webbpanelmetodik; för en sammanställning av forskningsläget, se [8]. Mest lovande är troligen metoder som beaktar inferensproblematiken både i urvals- och skattningsfaserna. Bland annat prövas olika matchningsmetoder och val av hjälpinformation för att effektivt minska selektionseffekter.

Det saknas dock fortfarande en solid vetenskaplig grund för webbpanelundersökningar, i synnerhet för dem som grundas helt på icke-sannolikhetsurval. Även om webbpanelundersökningar kan variera mycket i kvalitet, blir min slutsats i nuläget ändå att de normalt ska undvikas när målet är att skatta populationsstorheter tillförlitligt, i synnerhet för officiell statistik. Att traditionella, sannolikhetsbaserade individundersökningar numera drabbas av höga bortfall (ofta mellan 20 och 60 procent) räcker inte som argument för att överge dem, eftersom bortfallet i webbpanelundersökningar är ännu mycket högre (vanligen över 90 procent för sannolikhetsbaserade webbpaneler). För självrekryterade webbpaneler kan bortfallet inte ens definieras. Ett alternativ till webbpanelundersökningar är ökad användning av webbenkäter inom traditionella, sannolikhetsbaserade undersökningar. Förhoppningsvis kan också nya framsteg göras det närmaste decenniet inom forskningen på webbpanelundersökningarnas metodik.

JÖRGEN BREWITZ (F.D. SVENSSON),
STATISTISKA CENTRALBYRÅN

Referenser

[1] Bethlehem, J.G., och Cobben, F. (2013). *Web Panels for Official Statistics?* Invited paper presented at the 59th ISI World Statistical Congress, 2013, Hong Kong, China.

[2] Surveyföreningen (2014). *Kvalitet i webbpanelundersökningar.* ISBN 978-91-637-3193-8.

[3] Yeager, D.S., Krosnick, J.A., Chang, L., Javitz, H.S., Levendusky, M.S., Simpser, A., och Wang, R. (2011). Comparing the Accuracy of RDD Telephone Surveys and Internet Surveys Conducted with Probability and Non-Probability Samples. *Public Opinion Quarterly,* Vol. 75, pp. 709–747.

[4] Scherpenzeel, A., och Bethlehem, J.G. (2011). How representative are online panels? Problems of coverage and selection and possible solutions. In M. Das, P. Ester, and L. Kaczmirek (Eds), *Social and behavioral research and the internet: Advances in applied methods and research strategies* (pp. 105–132). New York: Routledge.

[5] Tourangeau, R., Conrad, F.C., och Couper, M.P. (2013). *The science of web surveys.* Oxford: Oxford University Press.

[6] Baker, R., Blumberg, S.J. , Brick, J.M., Couper, M.P., Courtright, M., Dennis, J.M., Dillman, D., Frankel, M.R., Garland, P., Groves, R.M., Kennedy, C., Krosnick, J., Lavrakas, P.J., Lee, S., Link, M., Piekarski, L., Rao, K., Thomas, R.K., och Zahs, D. (2010). Research Synthesis: AAPOR Report on Online Panels. *Public Opinion Quarterly,* Vol. 74, pp. 711–781.

[7] Svensson, J. (2014). *Web panel surveys – a challenge for official statistics.* Proceedings of Statistics Canada Symposium 2014, Gatineau, Canada.

[8] Callegaro, M., Baker, R., Bethlehem, J.G., Goritz, A., Krosnick, J.A., och Lavrakas, P.J. (editors) (2014). *Online panel research: a data quality perspective.* John Wiley & Sons, Ltd.

Webbpaneler, kvalitet och sannolikhetsurval

Allt fler undersökningar genomförs idag med webbpaneler. Som svar på detta har Surveyföreningen kommit med en rapport om kvalitet i webbpaneler. Dock lägger denna rapport ett allt för stort fokus på huruvida rekryteringen har skett med självrekryterade urval eller med sannolikhetsurval. Med så pass låga (kumulativa) deltagarandelar som dagens webbpaneler har, blir frågan om sannolikhetsurval mindre intressant. Det som istället spelar roll är vilka kalibreringsmodeller som används.

I februari presenterade Surveyföreningen sitt arbete avseende kvalitet i webbpaneler. Ett av de huvudsakliga problemen som lyftes i samband med seminariet är att många webbpaneler använder självselekterade urval, istället för de i undersökningssammanhang klassiska sannolikhetsurvalen. Problemet med självselekterade urval i webbpaneler lyfts i Surveyföreningens rapport liksom i Torbjörn Sjöströms artikel i senaste Qvintensen. Det konstateras ofta att det inte finns någon statistisk teori då självselekterade urval används och att dessa därför ”saknar vetenskaplig grund”. Även i Surveyföreningens arbete ”Kvalitet i webbpanelundersökningar: metoder och mått” lyfts bristen på teori för självselekterade urval fram gång på gång. För självselekterade urval menar författarna att ”någon teoretisk grund för att beräkna osäkerhetstal finns inte” (sidan 19). Osäkerhetsintervall bör därför inte produceras eller publiceras.

Jag menar att dessa slutsatser är fel väg att gå för att bedöma kvaliteten i webbpanelundersökningar, dels för att det finns statistisk teori – även om den inte bygger på teorin för sannolikhetsurval - och dels för att osäkerhetsintervall är det självklara sättet vi alltid bör presentera statistisk osäkerhet på.

Idag är det stora (och ökande) bortfallet ett stort problem för de flesta moderna undersökningar. Vid seminariet konstaterades att det inte är ovanligt med bortfall på över 60 % för vanliga undersökningar och (kumulativa) bortfall på mer än 95 % i webbpanelundersökningar. Bortfallet påverkar våra skattningar och introducerar alltid bortfallsfel – en felkälla som inte heller minskar med antalet observationer. Med hjälp av kalibrering kan vi även vid relativt stora bortfall minska bortfallsfelet – givet att vi har en god bortfallskompensatorisk modell med bra hjälpvariabler. Ju större vårt bortfall är, desto tyngre vilar våra resultat på denna kompensatoriska modell, denna ”magiska viktning” som Torbjörn Sjöström kallar det. Idag betyder det att webbpaneler, sannolikhetsurval eller ej, i praktiken är beroende av den kompensatoriska modellen.

Att i fallet med webbpaneler då göra en stor skillnad på sannolikhetsurval (med ex. 96% kumulativt bortfall) och självselekterade urval (som vi kan se som ett urval med 99,99…% bortfall) blir missvisande. Det centrala, i båda fallen, är vilken bortfallskompensatorisk modell som används – inte huruvida den lilla grupp som rekryteras har kända inklusionssannolikheter eller inte.

»...osäkerhetsintervall är det självklara sättet vi alltid bör presentera statistisk osäkerhet på.»

Det är också här jag inte delar synsättet i Surveyföreningens rapport, eller i de resonemang som fördes fram vid seminariet och i Torbjörn Sjöströms artikel i förra Qvintensen. Att det inte finns någon teori eller vetenskaplig grund för självselekterade urval är ett felaktigt påstående. Det finns en lång tradition av modellbaserad stickprovsteori för ändliga populationer, både frekventistisk och bayesiansk, »»»

»»» WEBBPANELER, KVALITET OCH ...

som kan användas. Ett allt vanligare exempel där denna metodik används är skattningar för små undersökningsgrupper (small area estimation). Med dessa metoder kan vi beräkna konfidensintervall, medelkvadratfel eller tillhandahålla beslutsunderlag i form av posteriorifördelningar med många goda teoretiska egenskaper. Vi blir dock modellberoende – liksom i fallet då vi stöder oss på vår bortfallskompensatoriska modell. Men genom att bygga vår teori för webbpaneler på modellbaserad skattningsmetodik får vi ett naturligt sätt att presentera vår osäkerhet med osäkerhetsintervall, samtidigt som vi kan använda oss av kunskaperna för bortfallskompensation som kalibrering från stickprovsteorin. Särskilt den bayesianska metodiken har här fördelar som kan utnyttjas då vi kan inkludera metodologisk kunskap och resultat som stöd i våra skattningar.

I Surveyföreningens text avråds nu från osäkerhetsintervall vid självselekterade urval. Det vi då ger intryck av att säga är att det inte finns någon osäkerhet i våra estimat. Det är ju helt fel – med så starka modellberoenden är vi **mer** osäkra på våra skattningar än om vi hade ett sannolikhetsurval med 100 % svarsfrekvens. Hur ska personer som inte är statistiskt kunniga kunna omvandla kvalitativa mått på en webbpanel till kunskap om enskilda skattningar? Här ligger ett ansvar hos oss statistiker att presentera och utveckla modeller för att fånga in och presentera denna osäkerhet – inte att låta bli att försöka eller hävda att det inte finns teori för osäkerhetsmått.

Vad finns det då för ytterligare kvalitetsindikatorer för webbpaneler som skulle behövas? Jag skulle hävda att den mest centrala kvalitetsdimensionen är öppenhet och reproducerbarhet. Vilka modellantaganden görs och hur görs beräkningarna, det är inte ovanligt att det bara konstateras att omviktning görs efter ett antal variabler. Inte hur denna viktning görs.

Istället för att lägga vår kraft på om de få procent som deltar i webbpaneler är insamlade med sannolikhetsurval eller inte bör vi istället fokusera på de bortfallskompensatoriska modellerna och den teori som redan finns och skapa metoder för att presentera osäkerhet från självselekterade urval. Vi bör fokusera på öppenhet som en kvalitetsindikator i en värld där det inte finns någon statistisk inferens från webbpaneler som inte är modellberoende.

MÅNS MAGNUSSON

DEBATTEN OM BORTFALL; ETT SVAR TILL TORBJÖRN SJÖSTRÖM

»SCB ska respek i samhället som

Mitt syfte i artikeln i Qvintensen nr 4/2014 var att kontrastera tänkesätten, SCB:s gentemot de privata statistikproducenternas, med utgångspunkt i trippeln bortfall – viktning – representativitet som hade präglat två artiklar, en från SCB (Källa SCB nr 2/2014), en där de privata instituten porträtterades som grupp (Qvintensen nr 2/2014). En kärnverksamhet i vår profession, att göra en statistisk undersökning, som det heter i titeln på en populär lärobok av Karin Dahmström – är polariserad i två läger som iakttar varandra, liksom genom en delvis transparent tjock glasvägg. Aspekter av det vill jag belysa för Qvintensens läsare.

Mitt val av Torbjörn Sjöströms företag Novus för att exemplifiera skaran av privata aktörer var inte slumpmässigt. Utan närmre kännedom om marknaden kollade jag hemsidorna hos några av företagen. Novus föreföll ”en markant profil”, en bjärt utmanare i en bransch vars svaghet i mångas ögon är bristen på deklarerad, övertygande linje i metodfrågor. Det var ett bra val, inte minst genom att Torbjörn Sjöström, Qvintensen nr 1/2015, sätter fingret på saker som jag själv vill tillägga något om, i raderna som följer. Jag tackar honom för hans bidrag till debatten.

Qvintensen var lämplig plats för mitt bidrag, för Svenska statistikfrämjandet omfattar olika inriktningar inom fältet statistik, från starkt teoretiska till klart praktiska. Många teoretiker – jag använder inte gärna ordet ”akademin” men menar något ditåt – undrar över turerna kring bortfall i undersökningar. En del frågar sig: Vad finns den teoretiska strukturen, vi har ju i alla fall en statistikvetenskap med solida matematiska grunder?

En debatt blossade upp i Dagens Nyheter 2015-01-18 kring bortfall, speciellt i SCB:s undersökningar. Det hände så vitt jag vet oberoende av min publicerade artikel, genom att media noterat ökat bortfall i den viktiga arbetskraftsundersökningen (AKU). Diverse inlägg följde, bland andra ett från Torbjörn Sjöström. Intervjuinslag kom i Sveriges Radio. Blandade synpunkter, merkantila, vetenskapliga, eller rent känslomässiga, delgavs allmänheten. Med den bakgrunden svarar jag i raderna som följer, med mitt sätt att se.

Utgångspunkterna är olika. Torbjörn Sjöström presenterar sig och sitt företag, talar med övertygelse utifrån sin horisont som företagschef. Hans uppfattning kontrasterar inte så lite mot konkurrenternas i den privata undersökningsbranschen. För min del är jag teoretiker, i matematisk/statistisk tradition, inte beteendevetenskaplig. Min referensvärld är den nationella statistikbyrån när den är som bäst och tar ledarskap i användning och vidareutveckling av metoderna, hos eliten av byråerna även av teorin. Jag är inte praktiker. Att jag under åren har haft att ta ställning ”på papperet” till metodik i hundratals undersökningar, i olika länders statistikbyråer, uttrycker att surveyteorin som jag bidragit till är viktig. En erfarenhet jag saknar men har beundrat är den som finns hos centralbyråstatistiker som ”levt med sin undersökning” kanske i decennier, som de kan allt om; det är imponerande att prata med dem. Bland undersökningar intresserar jag mig inte ”handgripligen” för någon. Jag intresserar mig för alla ur ett perspektiv där matematisk/statistisk teori ska ge lösningar och struktur. En sådan syn från ovan är relevant: I mer än hundra år har den teorin konsoliderat och ”stadfäst” i stort sett alla de viktiga metoderna i urvalsundersökningar. Mellan Torbjörn Sjöström och mig finns

teras som den instans vet bäst»

skillnader såväl som likheter i statistiskt synsätt.

Domedagsprofetior. Torbjörn Sjöström ser tendens till att bortfallsproblemet överförstoras; han säger att SCB har ”vid upprepande tillfällen utmålat det som det största hotet för officiell statistik”, i linje med en hotbild uttryckt i DN-rubriker 2015-01-18: ”SCB larmar om statistikproblem” och ”Sveriges officiella statistik hotar att bli missvisande.” Torbjörn Sjöström säger: ”Domedagsprofetior om bortfallet är … kraftigt överdrivna. Det skall inte ignoreras, men man skall i inte heller utropa undersökningarnas nära förestående död.” SCB fruktar inte en ”nära förestående död” för officiell statistik. Samhällsviktiga undersökningars existens är inte ifrågasatt. Att sättet att göra dem kan komma att ändras är inte uteslutet, inte ens ”grundvalarna”.

Kan man ”lita” på siffrorna? Frågan tränger sig på, tacksam för journalister, ställd i DN 2015-01-18: ”går det att lita på den officiella arbetslöshetsstatistiken när bortfallet är så högt?” Frågan är försåtlig. Bortfallet som akilleshäl för statistikens tillförlitlighet är sedan länge uppmärksammat inom SCB. Att det nu framställs i media som ett sensationellt ”avslöjande” är olyckligt. Det blir ett slag mot det som nationella statistikbyråer runt om i världen – de bättre av dem i alla fall – vet nödvändigheten av: att åtnjuta tillit (”trust”) hos allmänheten. I landets intresse bör media vara försiktiga med att dela ut onödiga slag. SCB bör vinnlägga sig om att förtjäna ”trust”.

Tilliten i siffrorna är långt högre än om man inte känt till metoderna (de som SCB använder, se nedan) utan agerat som att ”bortfall gör inget”, att det skulle vara slumpmässigt. Frågan ”kan vi lita på statistiken” var teoretiskt sett lika befogad för 30 år sedan som idag fast bortfallet, i procent, var klart mindre. De insatta visste det då. Att den blossar upp igen 2015 är en ny episod i en gammal följetång. Nyhetsvärdet ska falna denna gång också.

Om handfallenhet: Pratet om ett hot mot offentlig statistik ger ”en känsla av handfallenhet”, menar Torbjörn Sjöström, där man ”fokuserar på ett problem utan att varken sätta det i en relevant kontext eller diskutera de lösningar som finns”. Hans ord ”handfallenhet” innehåller en reprimand till SCB:s ledning. Det förmildras bara delvis av att SCB är en komplex organisation med expertfunktioner av olika slag på olika avdelningar. Kompetenta managers behöver inte ha vetenskapligt grundad insikt. Pinsamt var att se SCB, ledarskapet i svensk officiell statistik, bli utsatt i DN-debatten för mer eller mindre amatörmässiga råd om hur datainsamling och bortfall kanske borde skötas i stället. Allmänheten fick veta att SCB krisreagerat med ett bortfallsprojekt i syfte att ”vända utvecklingen”. Intervjuad i Sveriges Radio beskrev projektledaren målet som att komma tillbaks till 70- och 80-talens bortfallsnivåer. Även om det svävande målet skulle uppnås är det otillfredsställande. SCB borde klart låta förstå att oönskad bortfallsökning har vi levt med länge, vi kan en hel del om det där, och så här går det till, med våra metoder, internationellt erkända som bästa möjliga i dagsläget och vetenskapligt grundade. Var fanns ledarna som skulle auktoritativt ha klargjort detta omedelbart? SCB kom i kläm i debatten, i medborgarnas ögon. Jag är inte ensam om att beklaga det.

Nyckelfrågan ”lita eller inte lita på” är upp till varje användare, men SCB ska respekteras som den instans i samhället som vet bäst hur man tar fram de officiella siffrorna. Det finns inget bra alternativ i ett demokratiskt samhälle

Befintliga lösningar. Att ”diskutera de lösningar som finns” efterlyser Torbjörn Sjöström. Det gjorde SCB, fast ”på efterkälken” (Svenska Dagbladet 2015-03-03) när fyra av verkets metodexperter tycks ha fått i uppdrag gå ut och redovisa att bortfallsmetodiken i t.ex. AKU ”vilar på en vetenskaplig grund om stickprovsundersökningar” och att ”under de senaste decenniernas oönskade … utveckling med ett allt större bortfall … har ny forskning kompletterat den ursprungliga teorin med metoder för att motverka bortfallets negativa effekter”. Att ”motverka negativa effekter” genom viktning skedde för 50 år sedan eller mer med enkla medel som ”post-stratifiering” och ”raking ratio”. Från 80-talet och framåt etablerades viktning som generellt flexibelt verktyg på vetenskaplig grund. Till den nya forskningen har SCB bidragit bland annat genom att flera anställda har skrivit licentiat- eller doktorsavhandlingar kring bortfall. SCB:s beredskap är långt bättre än om dessa metoder inte funnes.

Generaliteten medför risker som att varianter kallas för ”knådning” (Inizio, i DN 2015-01-27). Det är sant att man inte kräver av en bagare att han avslöjar hemligheten, knådningen och annat, bakom sina goda bakverk. Inizios anspråk på yrkeskompetens ligger till en del i att anse sig själv nog, att tillåta sig vikta lite hur som helst, en ”knådning”, naivt beskriven av dem som ”en liten hemlighet från undersökningsvärldens insida”, men ”poängen är att vi talar tyst om den”. »»»

»»» SCB SKA RESPEKTERAS SOM DEN...

Viktning i surveyer är verkligen ingen hemlighet; har varit vetenskapligt erkänd i minst 50 år; kompetenta användare talar inte tyst om den tekniken. Dess detaljer ska givetvis redovisas. Torbjörn Sjöström (DN 2015-01-28) ges tillfälle att beskylla Inizio för ”magi och svarta lådor”.

Nyckelfrågan ”lita eller inte lita på” är upp till varje användare, men SCB ska respekteras som den instans i samhället som vet bäst hur man tar fram de officiella siffrorna. Det finns inget bra alternativ i ett demokratiskt samhälle: Nationella statistikbyrån ska ha (och ges) den kompetens, vetenskaplig och annan, som vet att ”göra det bästa möjliga”. Att ”rikspolitiker … är bekymrade” (DN 2015-01-18) om bortfallet låter gråtmilt. Samhället Sverige, som rikspolitikerna har ansvar för, kännetecknas alltmer av att folk inte kan nås eller fås att lämna information. De som är bekymrade nu borde varit det redan för länge sedan. Bortfall i undersökningar utpekades som ett hot på 1950-talet. DN med flera inser nu, 60 år senare, att det är ett hot.

Vad skulle SCB gjort, nu och tidigare, i ett proaktivt tänkande? I deras roll borde ligga att frimodigt och transparent deklarera ”på synliga ställen” att så här gör vi i en oönskad situation. Det finns ingen anledning att dölja hur det går till. Viktning ”i efterskott” är inget fult. ”Beskrivning av statistiken”, tillgänglig från hemsidan för varje SCB-produkt, läses inte av allmänheten och är ibland varken uppdaterad eller klargörande om bortfallet och om metoderna att behandla det. Försummas den ”populariserande” aspekten uppstår kompromitterande lägen, som i detta fall.

Slumpmässigheten i bortfallet kan jag se som ett önsketänkande från Torbjörn Sjöström, en hägring.

Bortfall egentligen inget problem? Bortfall behöver inte vara ett problem är Torbjörn Sjöströms tes. Principiellt, eller snarare i utopin, håller jag med. I praktiken är problemet oundvikligt, utom kanske för Torbjörn Sjöström, om han övertygande (matematiskt/statistiskt) visar hur slumpmässigt bortfall garanteras (så att viktning onödiggörs) i hans undersökningar. Som utmanande exempel citerar Torbjörn Sjöström, i DN 2015-01-28, en studie vid Stanford University där tillförlitliga resultat uppnåtts med 95 % bortfall (5% svar) förutsatt att först urvalet, sedan bortfallet från urvalet, är slumpmässigt. Med 5% garanterat slumpmässiga eller representativa svar utifrån ett 20-falt större-än-vanligt urval (så att i alla fall tillräckligt många observationer ger låg varians i skattningarna), så kan det så klart bli tillförlitligare statistik (tack vare mycket mindre skevhet, eller bias) än med 70% svarsandel i en dåligt balanserad, traditionellt insamlad datamängd. I detta banala konstaterande är ”5% svarsandel” kärnpunkten. Att minimal svarsandel kan ge bra statistik är givetvis utmanande för ”the establishment”. Man kan sitta ner och konstruera frapperande exempel. Men vad Torbjörn Sjöström kritiserar är den fixering som många har på bortfallsprocenten. Den är inte centralpunkten.

När jag håller med Torbjörn Sjöström i det är min syn så här: Att bortfallet i en undersökning (i den klassiska meningen ”andel icke-svar i ett sannolikhetsurval”) är 49% snarare än 42% är likgiltigt, både i ett forskningsperspektiv och i praktiken. SCB:s nuvarande metoder justerar automatiskt, ”rättar till” siffrorna med kanske stora belopp jämfört med om man inte hade brytt sig om det, och detta ungefär lika bra i båda fallen, fast troligen otillräckligt bra i båda. För små delgrupper av populationen, till exempel på låg regional nivå, tillkommer att antalet svar, oavsett andelen svar, är avgörande för estimatens kvalitet.

Bortfall som vägledande begrepp har sprängt sig själv inifrån, imploderat. Det har tjänat ut, för mig i alla fall, som ledstjärna för forskning och metodutveckling. Framtida undersökningar (med eller utan sannolikhetsurval) måste rikta in sig på fullödigare kriterier på kvalitet i insamlad datamängd än den torftiga och banala bortfallsprocenten. Som jag sa i min artikel, diskussionen (kring officiell statistikframställning) kommer i framtiden att röra sig kring andra nyckelfrågor, när konservatismen övervunnits.

Det är sant att begreppet bortfall (som andel av sannolikhetsurval) vägledde oss för 30 år sedan när vi sa: ”Då får vi utveckla metoderna där bakgrundsfakta (även kallat hjälpinformation, eller kompletterande information) från register och andra källor sätts in för att bortfallsjustera genom viktning”. Då hoppades vi ännu på återgång till nära hundra procent svar. Även om *metodiken* fortfarande är nyttig (som när SCB nu säger att den fortfarande fungerar hyfsat bra) så är, med blick på framtiden, *tänkesättet* ”old hat”.

Torbjörn Sjöström betonar ”slumpmässighet”, i strävan att profilera sig mot de övriga instituten. I sitt svar till Inizio, DN 2015-01-28 säger han: ”I en korrekt genomförd undersökning mot allmänheten är även bortfallet slumpmässigt. Med andra ord blir resultatet tillförlitligt.” Det är ett kollektivt underkännande av alla undersökningar, inklusive SCB:s, där bortfallet inte blir slumpmässigt. Jag har liktänkande kollegor som skulle ha en hel del att säga om det underkännandet. Slumpmässigheten i bortfallet kan jag se som ett önsketänkande från Torbjörn Sjöström, en hägring. Att han jobbar mot det idealet är föredömligt. Man vill veta hur han uppnår det i sina undersökningar, med webbpanel eller på annat sätt, och den matematiska innebörd han lägger in i ”slumpmässigt bortfall”.

En roll för akademin. Om de privata undersökningsföretagen som ”prispressad förtroendebransch” säger Torbjörn Sjöström: ”Tyvärr finns det också flera aktörer som tär på detta förtroende”, att metodutveckling efterfrågas, att ”Novus har välkomnat alla samarbeten med akademin vi har blivit inbjudna till” och ”även själva aktivt sökt samarbete, men tyvärr ser vi att den metodutveckling branschen så väl behöver snarare har fått ske helt i vår egen regi”. Det är synd att sådant intresse är svalt. Mycket metodutveckling – dessförinnan mer grundläggande forskning - finns att göra kring panelundersökningar, inklusive att formulera och bygga in den ”slumpmässighet” som Torbjörn Sjöström värnar om i bedömningen av dessa allt viktigare undersökningar. Statistiker med matematisk inriktning borde tillmötesgå Torbjörn Sjöström och tackla frågorna som han ställer.

CARL-ERIK SÄRNDAL

Debatten om bortfall i DN och SvD

Debatt i Dagens Nyheter och Svenska Dagbladet om bortfall i statistiska undersökningar:

”Sveriges officiella statistik hotar att bli missvisande”, Kristoffer Örstadius, Dagens Nyheter, www.dn.se, 2015-01-18.

”SCB larmar om statistikproblem”, Sveriges radio, www.sr.se, 2015-01-18.

”Det är i justeringen av siffrorna som opinionen kommer fram”, Karin Nelsson och Anders Lithner, Inizio. Dagens Nyheter, www.dn.se, 2015-01-27.

”Bortfallet är inte så problematiskt som Inizio påstår”, Torbjörn Sjöström, Novus. Dagens Nyheter, www.dn.se 2015-01-28.

”Bortfallet är inte slumpmässigt”, Karin Nelsson och Anders Lithner, Inizio. Dagens Nyheter, www.dn.se 2015-02-04.

”Färre svar problem för undersökningar”, Martin Axelson, Joakim Malmdin, Gustaf Strandell och Marianne Ängsved, SCB, www.svd.se, 2015-02-09.

DAN HEDLIN

ORDFÖRANDEN HAR ORDET

Kvalitet i webbpanel-undersökningar

Diskussionen om bortfall i urvalsundersökningar går vidare. Dagens Nyheter och Svenska Dagbladet (se referenser på sidan 26) har under det här året ägnat betydande utrymme åt artiklar och debattinlägg i frågan och i föreliggande nummer av Qvintensen får vi fler synpunkter. På årsmötet i mars framkom önskemål om att vi även från Surveyföreningens sida i ett seminarium eller i en workshop borde ta upp problematiken kring bortfall och kanske kan vi redan till hösten tillgodose detta.

Årsmötet bjöd på intressanta föredrag med tillhörande diskussioner av en kunnig och engagerad publik. Anders Eklund presenterade justerade jämförelsemått för möjlig tillämpning i organisationers verksamhetsutveckling. Dan Hedlin följde upp med ”irrläriga” synpunkter kring sannolikhetsbaserade urval och därefter presenterade Edgar Bueno sitt vinnande bidrag i den årliga pristävlingen för bästa uppsats inom surveyområdet. Edgar har föreslagit en ny urvalsmetod med tillhörande estimator (se sidan 28).

Åke Wissing: årets Tore Dalenius-föreläsare.

Åke Wissing avslutade sedan med årets Tore Dalenius-föredrag. Åke spekulerade bland annat kring hur Tore skulle ha resonerat angående webbpanelundersökningars användbarhet och kvalitet.

Åke Wissing var också primus motor för Surveyföreningens första seminarium för i år: ”Kvalitet i webbpanelundersökningar”. Evenemanget lockade ett 80-tal deltagare och inleddes av Gösta Forsman som kort presenterade bakgrunden till rapporten betitlad som seminariet. Åke presenterade sedan de föreslagna beskrivningsmåtten med referensvärden från några undersökningar som utförts av olika undersökningsföretag. Därefter vidtog föredrag av representanter både från beställare (köpare) av undersökningar och undersökare (säljare). Konferencieren Arne Modig fick avslutningsvis agera moderator i en panel-(vad annars?!) diskussion. Det framkom av presentationerna och avslutningsdebatten att det finns ett behov av att träffas för diskussion av policy- och kvalitetsfrågor.

»...det finns ett behov av att träffas för diskussion av policy- och kvalitetsfrågor»

Seminariet har följts upp av en enkätundersökning bland deltagarna (med glädjande litet bortfall!) och denna ska nu analyseras för att vi på bästa sätt ska kunna följa upp det lyckade seminariet med nya evenemang.

Föreningens medlemmar uppmuntras förstås också att komma med förslag till Surveyföreningens utåtriktade verksamhet.

PER GÖSTA ANDERSSON

Surveyföreningens styrelse 2014-2015

Per Gösta Andersson (ordförande), Jonas Ortman (sekreterare), Maria Varedian (kassör), Marcus Berg (webbansvarig).
Övriga ledamöter: Fredrik Jonsson (nyval), Karin Schneller och Åke Wissing

FREDRIK JONSSON, NY LEDAMOT

Jag jobbar på Statistiska centralbyrån, med huvudsaklig inriktning mot ekonomisk statistik. Min grundutbildning fick jag vid matematiska institutionen, Uppsala universitet. Det hela mynnade ut i doktorsavhandlingen *Self-Normalized Sums and Directional Conclusions*, Allan Gut som huvudhandledare, för vilken jag erhöll Statistikfrämjandets Cramérpris. Våren 2012 lämnade jag studentlivet i Uppsala till förmån för huvudstaden. Jobbmässigt först en kortare anställning som post doc på Karolinska Institutet, MEB närmare bestämt, innan jag hamnade där jag befinner mig nu. Sedan ett par månader tillbaka är jag nybliven far. Mellan jobb och familj snörar jag gärna på mig löpardojorna. I skrivande stund börjar det bli dags att trappa ned på träningen och toppa formen inför Stockholm Marathon!

Fredrik Jonsson, ny ledamot.

Edgar Bueno vann Surveyföreningens pris för bästa uppsats på surveyområdet. Här skriver han om sin uppsats.

One of the most important stepping stones in survey statistics was when researchers realized that a lot is gained by taking a random sample, as opposed to taking a non-random sample or making a census.

q-sampling shows good

In the 1950s more elaborated sampling designs became common. One set of sampling designs that use auxiliary information to make the sample more efficient and save resources is probability proportional to size designs, or simply PPS designs.

PPS designs are generally more efficient than designs that do not use auxiliary information, but this efficiency is affected by different factors such as the variable under study, the other properties of the design and the estimator that is used. For example, PPS with replacement is often used together with the so-called Hansen-Hurwitz estimator and a small variance is achieved when the study variable is well correlated with the selection probabilities, but its with replacement nature makes it less efficient than desired; PPS designs without replacement are often coupled with the Horvitz-Thompson estimator and a small variance is again achieved when the study variable is well correlated with the inclusion probabilities.

However, it has not been possible to obtain a PPS design that simultaneously satisfies a set of desired properties, e.g. having a simple selection method that works for any sample size, having second order inclusion probabilities that are positive and easy to calculate, being of fixed size and without replacement.

To try to find a PPS design that satisfies all the desired properties at the same time, I wanted to explore an idea of mine that I call q-sampling.

Defining q-sampling and the q-estimator

The q-sampling is a design defined as

$$p(s) = \frac{1}{\binom{N-1}{n-1}} \sum\nolimits_s q_k$$

where p(s) is the probability that sample s is drawn. The sum is taken over a support defined as all the without-replacement combinations of size n out of the N elements of the finite population, where the q values are such that

$$t_q = \sum_U q_k = 1 \text{ and } \sum_{i=1}^{n} q_{(i)} > 0,$$

$q_{(i)}$ are the order statistics of an auxiliary variable q.

The q-estimator of the total is defined as

$$\hat{t}_q = \frac{\sum_s y_k}{\sum_s q_k}$$

It is unbiased under q-sampling.

A sampling strategy is the combination of a sampling design and an estimator. The sampling design may be interpreted as how to select the sample, the estimator as how to use the selected sample in order to estimate a given parameter. The strategy (q-sampling – q-estimator) can be considered as a member of the family of PPS sampling, where the proportionality is with respect to the q-values.

Assessing q-sampling

The ideal situation for using PPS sampling in practice is when there is a linear association without intercept between the auxiliary variable and the study variable as well as a high correlation between them. A population following these properties was simulated as follows: two thousand observations from a chi-square distribution with one degree of freedom were simulated and considered as the auxiliary variable, denoted by x; using the simulated x-values, observations for the study variable, denoted by y, were obtained using the following equation:

$$y_k = 2.39x_k + \varepsilon_k \text{ with } \varepsilon_k \sim N(0, x_k)$$

Using the simulated population, four aspects of the strategy (q-sampling – q-estimator) were assessed.

An expression for the variance is not available, so an approximate expression was obtained by the Taylor linearization method. In order to assess how close the approximation is to the

coverage

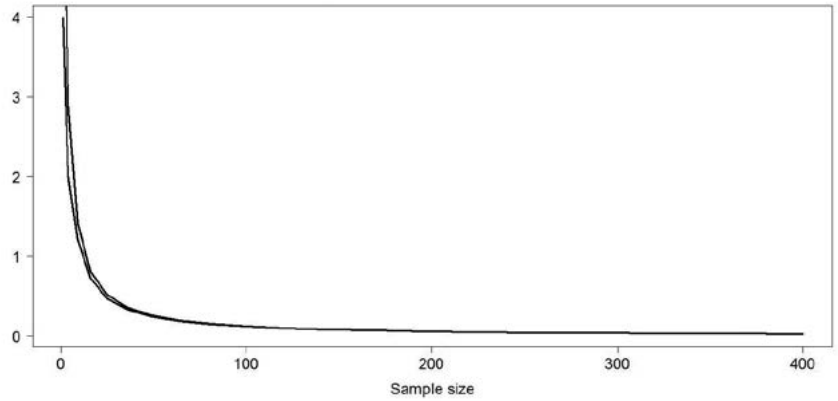


Figure 1. Comparison of simulated variance with the Taylor approximated variance for different sample sizes. They are very close.

(unknown) variance, the latter was obtained by simulation. A comparison of the approximated variance and the simulated variance is shown in Figure 1. It can be observed that, as the sample size increases, the approximated variance gets closer to the simulated variance. A sample size of n = 100 is enough for both lines to be visually indistinguishable from each other.

In practice it is not possible to calculate the variance or its approximation, it is only possible to calculate an estimate of it. A biased estimator of the variance was proposed for the strategy (q-sampling – q-estimator). However, the magnitude of the bias is unknown. In order to assess this magnitude, the expectation of the variance estimator was calculated by simulation and then compared with the approximation to the variance (Figure 2, on the last page of this issue of Qvintensen). It is observed that, as the sample size increases, the expected value of the variance estimator gets closer to the approximate variance. However, medium to large samples are needed for the bias to be negligible.

One criterion to make comparisons between sampling strategies is Mean Square Error, which is reduced to the variance when the strategies are unbiased. An unbiased strategy is said to be more efficient than another when its variance is smaller. In order to assess the efficiency of the strategy (q-sampling – q-estimator), its variance was compared with those of some known strategies: (Simple Random Sampling, SRS – HT-estimator), (Pareto – HT-estimator) and (systematic PPS – HT-estimator). All the strategies are unbiased or nearly unbiased. Figure 3 shows the simulated variance of the strategies under comparison for different sample sizes. It can be seen that q-sampling is more efficient than SRS, but systematic PPS and Pareto sampling are even more efficient than q-sampling.

Usually, using the central limit theorem, the distribution of the estimates is approximated by a normal distribution and confidence intervals are estimated using this assumption. It is expected that confidence intervals built in this way will cover the true parameter with a probability of approximately 100(1-α)% for a given confidence level, α. This coverage was also assessed for q-sampling and compared with the same strategies mentioned above (SRS, Pareto sampling and systematic PPS). Figure 4 shows that, for all the strategies, as the sample size increases, the coverage tends to the desired 95% level. Nevertheless, SRS shows coverage below the desired level, q-sampling performs slightly better but also below the desired level. On the other hand, Pareto sampling shows coverage very close to the desired 95%, whereas systematic sampling lies above the desired level.

As mentioned above, the assessment presented so far is based on a simulated variable that satisfies the ideal situation in practice. The same analysis was carried out for different simulated variables that are different, to some extent, from an ideal situation. In general the conclusions are:

- The variance of q-sampling tends to be in between SRS and Pareto sampling, i.e. when q-sampling is more (less) efficient than SRS, then Pareto sampling is, in turn, more (less) efficient than q-sampling.
- When the study variable departs from the ideal situation, coverage for Pareto sampling falls below the desired level, whereas q-sampling always performs well in regard to this criterion.

Conclusion

The sampling strategy (q-sampling – q-estimator) has some interesting properties, among them it is unbiased, a simple expression for the inclusion probabilities of any order is available, also a selection scheme that works for any sample size is at hand, and its coverage seems to be very close to the desired one. On the other hand, the main drawback of the strategy is that, if powerful auxiliary information is available, it is inefficient compared to other strategies like Pareto sampling and systematic πPS. It would be interesting to develop further research in order to find an optimal way to define the q-values.

EDGAR BUENO

Nya ledamöter i styrelsen för FMS

På årsmötet valdes till FMS styrelse för 2015: Marcus Thuresson som ordförande (omval), Annica Dominicus som kassör (nyval), Caroline Weibull som sekreterare (omval), Aldana Rosso som arkiv- och hemsideansvarig (nyval), Anna Lindgren som FMS representant i Statistikfrämjandets styrelse (nyval) samt Nasser Nuru Mahmud och Anna Berglind som ledamöter (omval).

Som representanter till EFSPI, European Federation of Statisticians in the Pharmaceutical Industry, valdes Magnus Kjaer och Mattis Gottlow.

Till styrelsen finns sedan tidigare även Marie Linder knuten som kontakt med Qvintensen.

På sidan 33 berättar Annica Dominicus och Aldana Rosso lite mer om sig själva,

ORDFÖRANDEN HAR ORDET

En "nära statistikhändelse"

Vårmötet med efterföljande årsmöte hölls i år på Karolinska Institutet och mina tankar om vårmötet finns härintill.

Förra året lades en del krut på medlemsfrågan, och med tanke på att vi tyvärr ser att medlemsantalet i FMS minskar år för år så finns det definitivt skäl att fortsätta jobba med den frågan. En första uppmaning till alla som läser detta är att komma ihåg att betala medlemsavgiften (50 kronor är en struntsumma i sammanhanget) om du redan är medlem, och om du inte är medlem att kontakta **fmsstyrelse@gmail.com** och genast ansöka om medlemskap.

Jag tänkte jag skulle dela med mig av en ”nära statistikhändelse” från verkliga livet. En av mina söner var i början av mars med i Schackfyran, en tävling mellan fjärdeklassare från hela Sverige. Att schack legat i framkant när det gäller att rangordna spelare hade jag sedan tidigare koll på men att de skulle kunna göra det i detta format var för mig väldigt imponerande. I Uppsala där denna deltävling ägde rum var det ca 700 deltagare. Varje match spelades i maximalt 20 minuter. Efter 20 minuter samlades alla resultat in och baserat på hur det gick i matchen (och tidigare matcher) matchades man mot en jämbördig motståndare. Det matchningssystem som används heter för övrigt Monrad om någon är intresserad av detaljer kring detta förfarande. Efter avslutad omgång lyckades man samla in, läsa in resultaten, göra ny matchning samt skriva ut och lägga ut de nya matcherna på ca 20 minuter. Jag tror det finns många tävlingar som har mycket att lära av detta.

Den stora snackisen inom biostatistikområdet under våren var definitivt beslutet av *Basic and Applied Social Psychology* att förbjuda såväl p-värden som konfidensintervall för artiklar som publicerades i tidskriften. Detta debatterades flitigt av stora institutioner som Royal Statistical Society och American Statistical Association och fick väldigt stor medial uppmärksamhet i den statistiska världen. Om vi bortser från det symboliska värdet kan man fundera på hur stor betydelse det kan komma att ha för oss statistiker. Enligt Wikipedia fanns det år 2013 totalt 11 427 tidskrifter inkluderade i *Journal Citation Reports*.

Sett i det perspektivet känns det som man kan fokusera på att det trots allt är 11 426 (99,99 %) tidskrifter som fortfarande accepterar p-värden och konfidensintervall och inte fokusera allt för mycket på en avvikande observation.

MARCUS THURESSON

Lästips: Dan Hedlin skriver om den ”snackis” som Marcus Thuresson nämner. Se ”Tidskrift ogillar statistisk inferens” på http://statistikframjandet.se/qvintensen /red

...”big data” är ett område där biostatistiker skulle kunna göra mycket nytta, men där vi nu hamnar i skuggan av datavetare och personer med andra kompetenser.

FRAMTIDENS BIOSTATISTIK

TEMA PÅ FMS VÅRMÖTE 2015

Interpunktionen efter titeln Framtidens biostatistik var utelämnad men givet intrycken från dagen kan man mycket väl tänka sig att avsluta temat med såväl ett utropstecken som ett frågetecken.

Till en början gavs ett flertal exempel som fokuserade på just framtidens biostatistik. Vi bjöds på exempel avseende biostatistikens historiska, nutida och naturligtvis framtida betydelse inom såväl traditionella som framtida områden såsom läkemedelsutveckling (Carl-Fredrik Burman från AstraZeneca/Chalmers och Anna Törner från Scandinavian Development Services), big data (Yudi Pawitan från Karolinska Institutet, KI), bioinformatik (Ingrid Lönnstedt från Statisticon) och spatial statistik (Aila Särkkä från Chalmers). Utan att gå in på detaljer så kan presentationerna ändå generellt sett sammanfattas som väldigt inspirerande och intressanta.

Sista punkten på agendan under vårmötet var en paneldiskussion där några av Sveriges mest meriterade biostatistiker deltog i panelen. Moderator för diskussionen var Paul Dickman från KI och deltagare i panelen var Tom Britton från Stockholms universitet, Johan Bring från Statisticon, Mattias Rantalainen från KI, Aila Särkkä samt Carl-Fredrik Burman. Om temat för den tidigare delen av dagen kunde avslutas med ett utropstecken så hade definitivt temat för denna diskussion ett stort frågetecken efter sig. I diskussionerna om dagens och framtidens biostatistik kom det fram ett antal farhågor inom flera olika områden.

Utbildning

Allt för få biostatistiker utbildas och de statistiker som kommer direkt från utbildningarna saknar vitala kompetenser såsom problemlösning och programmering. Det skulle behövas specificerade biostatistikavdelningar eller biostatistikcentrum där kompetensen kan samlas. Om man skall undervisa i ämnen relaterade till läkemedelsutveckling skulle det behövas att statistiker med gedigen erfarenhet inom området kommer in i utbildningarna och delar med sig av sin kunskap.

Läkemedelsindustrin

Läkemedelsindustrin har förändrats avsevärt de senaste decennierna och idag kan man se en tydlig trend att allt färre arbetstillfällen finns för biostatistiker inom detta område i Sverige. En tydlig trend är också att antalet patienter som ingår i läkemedelsprövningar har sjunkit drastiskt de senaste tio åren. Vissa farhågor framfördes att en allt större del av den statistiska programmeringen hos framför allt AstraZeneca numera sköts från Indien. Detta bemöttes dock med att det framför allt är mindre kvalificerade uppgifter som flyttats medan mer kvalificerade uppgifter ännu finns kvar. Problemet med var nyutexaminerade statistiker skall få erfarenhet inom området kvarstår dock.

Arbetsplatser

Om man strävar efter att utbilda fler biostatistiker måste man kunna garantera jobb. Idag är det mer lockande att söka sig till t.ex. försäkringsstatistiken, där det är lätt att få jobb för en nyutexaminerad statistiker. Biostatistiker är idag inte en naturlig del av en forskargrupp (till skillnad mot många andra länder) utan tas in vid akuta behov. Om detta synsätt förändrades skulle det finnas en stor marknad som behövde fyllas.

Big data

Det som ganska löst kallas ”big data” är ett område där biostatistiker skulle kunna göra mycket nytta, men där vi nu hamnar i skuggan av datavetare och personer med andra kompetenser.

Epidemiologi

I jämförelse med till exempel Storbritannien ligger Sverige långt efter med avseende på statistisk kunskap inom det epidemiologiska området.

Som slutsats kan man definitivt konkludera att biostatistik som ämne och yrke i Sverige står inför stora utmaningar, men att det också finns många ljuspunkter. Alltså endera ett utropstecken eller ett frågetecken, eller kanske båda.

MARCUS THURESSON

Fem DAGar i London:

»Det tog en stund innan jag kunde ta ordet kausalitet i min mun...»

Som nybliven epidemiolog, efter många år inom läkemedelsindustrin, åkte jag till London för att lära om kausal inferens. Som erfaren statistiker trodde jag mig veta att statistiken inte ägnar sig åt att klarlägga kausalitet, utan åt att

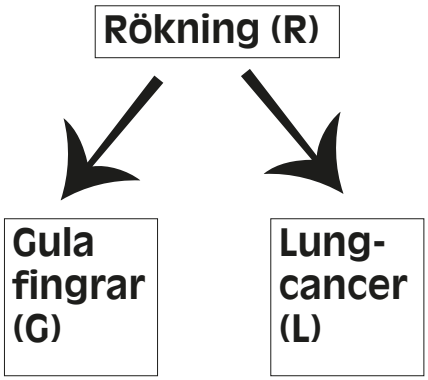
1. testa variablers inverkan på något intressant och mätbart tillstånd
2. prediktera något som vi vill veta om framtiden

I min statistiska utbildning har det aldrig hetat att vi har klarlagt orsakssamband, utan det har till exempel betonats att korrelation inte mäter kausalitet. Att klarlägga orsakssammanhang lät i mina öron som något deterministiskt. Så det tog en stund innan jag kunde ta ordet kausalitet i min mun.

Jag anlände till ett novembergrått London och tog in på ett enkelt hotell nära London School of Hygiene and Tropical Medicine, som gav kursen.

Det var en välrenommerad lärarkår som mötte upp: Bianca De Stavola, Simon Cousens, Richard Silverwood, Rhian Daniel, Karla Diaz-Ordaz och Stijn Vansteelandt. Deltagarna var i huvudsak från brittiska öarna, men där fanns också långväga deltagare från Japan och USA.

Något som omedelbart vädjade till min intuition var de Directed Acyclic Graphs (DAGs) som föreläsarna ritade från första stund. Jag har lagt mig till med vanan att rita upp orsakssamband mellan variabler som underlag för diskussioner med mina biovetenskapliga kollegor. Begreppet confounding, eller förväxlingseffekter, går det nu att samtala om på ett nytt intuitivt sätt.



Den som vill kan tänka på kopplade ekvationer (Structural Equation Models) för att förstå en DAG. Antag att egenskapen Gula fingrar är oberoende av Lungcancer hos en individ givet Rökning, se grafen. Antag för enkelhets skull att variablerna är kontinuerliga. Exempelvis kan grafen översättas till $G = \beta_1 R + \varepsilon_G$ och $L = \beta_2 R + \varepsilon_L$, där ε_G och ε_L representerar feltermer. I sådana modeller där de gemensamma orsakerna anses inkluderade antas feltermerna vara oberoende. Det visar sig att i så fall gäller att varje variabel är oberoende av alla andra variabler (dess effekter undantagna) givet dess direkta orsaker. I vårt exempel gäller alltså att G och L är oberoende givet R.

När det gäller det kausala har jag nu läst Dawid, som säger att kausala teorier inte är något annat än ambitiösa probabilistiska teorier, som postulerar vissa gemensamma beteendemönster tvärsöver en rad sammanhang. Jag är en praktiker med massor av mylla under naglarna, så den typen av filosofiska utgjutelser begriper jag inte. Men jag tänker som så: en okontroversiell teori för betingade fördelningar tolkas i termer av kausalitet. En modell för betingade oberoenden kan anses beskriva en orsakskedja. Men det krävs ett leap of faith för att känna visshet.

Ett annat begrepp som var relativt nytt för mig var propensity score, vilket är sannolikheten att få den behandling som studeras. Om analysen görs med denna som kovariat, eventuellt efter någon transformation, så kan man reducera bias. Den som tidigare stött på Little och Rubin och deras idé att modellera sannolikheten för saknade värden i samband med imputeringar kan känna igen resonemangen.

All dataanalys genomfördes i programpaketet STATA, som är nischat mot denna sorts epidemiologiska tillämpningar. Både övningsdata och kod gicks igenom på djupet.

För mig som inlett en ny karriär som epidemiolog var denna kurs gefundenes Fressen.

PER BROBERG

ANNICA DOMINICUS, NY KASSÖR

Annica Dominicus, ny kassör i FMS

Jag arbetar som biostatistiker och tycker att det är fantastiskt! Med en magister- och forskarutbildning i matematisk statistik från Stockholms universitet i ryggen tog jag klivet ut i arbetslivet 2006. Jag hade börjat så smått redan innan med deltidsarbete som statistiker på institutionen för medicinsk epidemiologi och biostatistik (MEB) på Karolinska Institutet dit även min egen forskning var knuten.

Sedan blev det arbete på AstraZeneca där jag var kvar fram till nedläggningen av forskningskontoret i Södertälje 2012. Sedan dess har jag arbetat som konsult – sedan förra sommaren på Scandinavian Development Services (SDS).

Som konsult får jag chans att jobba med många olika frågeställningar och typer av data (och därmed statistiska metoder) inom medicinsk forskning. Mestadels är det kliniska prövningar, observationsstudier och registerstudier. Det jag uppskattar mest med mitt jobb är det stora engagemanget och den enorma kunskapen hos de jag arbetar med – kunder och kollegor. Tack vare den stora bredden på uppdrag kommer det också alltid nya utmaningar och chanser till att lära sig något nytt.

På fritiden tillbringar jag mycket tid tillsammans med familjen, pysslar i trädgården och dansar flamenco.

ALDANA ROSSO, NY LEDAMOT

Aldana Rosso, ny ledamot i styrelsen för FMS

Jag arbetar som statistiker på Epidemiologi och Registercentrum Syd, som är en enhet inom Skånes universitetssjukhus (SUS). Det är mycket spännande att jobba på SUS eftersom jag har kontakt både med kliniska forskare och grundforskare, som arbetar med allt från väldigt jordnära problem till rent teoretiska studier. På min enhet tar vi hand om cirka en tredjedel av alla svenska kvalitetsregister.

Huvudsakligen arbetar jag med kvalitetsregister inom ortopedi med allt från databasutveckling till forskning. Jag hjälper också doktorander och forskare inom SUS och på Lunds universitet i olika forskningsprojekt: allt från t-test till stora kliniska studier med tiotusentals deltagare.

Tillsammans med en grupp mycket duktiga statistiker från SUS ger och utvecklar jag kurser inom statistik för doktorander vid Medicinska fakulteten i Lund. Jag ansvarar för en nivå-II kurs med inriktning mot biomedicin och labbmedicin.

Min väg till statistiken började egentligen som ett starkt intresse för experimentell fysik. Jag flyttade från Buenos Aires under 2004 för att doktorera vid Maxlab och Uppsala universitet. Jag ville doktorera inom synkrotronljusfysik och då fanns det endast ett trettiotal labb i hela världen som hade utrustningen jag ville jobba med, bland dessa Maxlab. En studiekamrat hittade en tjänst i Lund och tipsade mig om den. Jag sökte och fick den. Under doktorandutbildningen blev jag mer intresserad av dataanalys än av fysik och påbörjade en masterutbildning i statistik i Lund vid sidan om. Efter disputationen har jag jobbat som statistiker inom läkemedelsindustrin innan jag landade på SUS.

Jag trivs mycket bra i Sverige. Jag har upptäckt saker som längdskidor och skridskor, och all natur med skog, sjöar och fjäll. Jag bor i Södra Sandby (en liten by 10 km från Lund) tillsammans med min man Henrik. En härlig bonus med mitt jobb: Jag kan cykla eller åka rullskidor till jobbet året runt!

HÄNT OCH HÖRT

DISPUTATIONER 2015

- 3 mars 2015 **Johan Nykvist** *Topics in importance sampling and derivatives pricing*, tillämpad matematik och beräkningsmatematik, KTH. Opponent: Bert Zwart, Centrum Wiskunde & Informatica, Amsterdam.
- 14 april 2015 **Karin Stål** *Identifying Influential Observations in Nonlinear Regression*, statistik, Stockholms universitet. Opponent: Sune Karlsson, Örebro universitet.
- 8 maj 2015 **Shaobo Jin** *Essays on Estimation Methods for Factor Models and Structural Equation Models*, statistik, Uppsala universitet. Opponent: Li Cai, UCLA.
- 11 maj 2015 **Yuli Liang** *Contributions to Estimation and Testing Block Covariance Structures in Multivariate Normal Models,* statistik, Stockholms universitet. Opponent: Ivan Žežula, Pavol Jozef Šafárik University, Košice, Slovakien
- 22 maj 2015 **Fredrik Olsson** *Inbreeding, Effective Population Sizes and Genetic Differentiation : A Mathematical Analysis of Structured Populations,* matematisk statistik, Stockholms universitet. Opponent: Steinar Engen, Norges teknisk-naturvitenskapelige universitet.

LICENTIATSEMINARIER 2014

- 24 februari 2015 **Jonas Hallgren** *Continuous time Graphical Models and Decomposition Sampling*, tillämpad matematik och beräkningsmatematik, KTH. Diskutant: Błazej Miasojedow, University of Warsaw.
- 24 april 2015 **Henrike Häbel** *Characterization of Micro-Structures in Materials*, matematisk statistik, Chalmers. Diskutant: Mats Kvarnström, AstraZeneca.

CRAMÉRSÄLLSKAPET

ORDFÖRANDEN HAR ORDET

Att valla skidor: konst eller vetenskap?

I början av maj konstituerades Cramérsällskapets nya styrelse för innevarande verksamhetsår. Tre ledamöter är nya, nämligen Johan Lindström, Tatjana Pavlenko och Tatjana von Rosen där den sistnämnda efterträder Leif Ruckman som kassör. Samtidigt lämnar Leif Ruckman, Lars Rönnegård och Anna Lindgren styrelsen och till dem vill jag rikta ett stort tack för de insatser de bidragit med under det gångna året (och även dessförinnan).

Valet av ny styrelse ägde som sig bör rum på årsmötet i mars, ett möte där jag själv inte hade möjlighet att delta. Jag var tjänstledig under vintern och ägnade mig åt skidor på heltid i två avseenden. Jag var tidtagningschef på skid-VM i Falun under februari och det finns många statistiska infallsvinklar på tidtagning – mättekniker, datafångst, bearbetning, nyttjande av sekundärdata och distribution i realtid för att tillfredsställa tävlande, åskådare på plats, var och en med eller utan diverse mobilappar, tv-bolag och deras tittare samt forskare. Ja, det fanns en mindre armé av idrottsforskare på plats också, ständigt efterfrågande allsköns observationsdata och med önskemål om att få rigga egna experiment och egen datafångst under tävlingarna.

Men jag tänkte i stället skriva om skidvallning som jag i vinter ägnat hundratals timmar. Vetenskaplig metodik och skidvallning borde väl hänga lika intimt samman som vetenskaplig metodik och framtagning av läkemedel, pedagogiska undervisningsmetoder och växtförädling. Men tro mig, så är det inte. Funnes det en motsvarighet till den forna kvacksalverilagen för skidvallning så skulle antalet svenska fängelseplatser skyndsamt behöva mångfaldigas. Vallaindustrin, understödd av TV-experter och andra förståsigpåare, har främst ägnat sig åt att mystifiera skidvallningen och därigenom förringat värdet av ett vetenskapligt angreppssätt. Ett undantag i sammanhanget är den nyligen avgångna ordförande i Vetenskapsrådets beredningsgrupp för matematik, Lars-Erik Persson, som genom eget testande och argumenterande under ett decennium ger ett handfast råd till skidmotionärer: bästa, billigaste, miljömässigt sundaste och säkraste alternativet är att skippa glidvallan och i stället stålsickla, det vill säga hyvla av skidbelaget så att det är rent, jämt och plant.

Jag tror att ett viktigt skäl till att vetenskaplig metodik inte används i vallning (och inte heller inom en mängd andra, viktigare, områden i samhället) är att så få personer har grundläggande kunskaper i försöksplanering.

Vallningsproblemet innefattar, något förenklat och begränsat till glid, följande faktorer: val av skidor (S), permanent struktur (X) i skidbelaget som stenslipas ut, basparaffin (B) som värms in på skidbelaget, topping (H) ovanpå basparaffinet som består av fluorokarbon och slutligen den temporära struktur (R) som rullas in i belaget med en mönstrad stålrulle. Modellen är således S*X*B*H*R.

En första tanke för att ta reda på vilken kombination som ger bäst glid vore att testa alla kombinationer, och välja den bästa. Men det är inte praktiskt möjligt. En anledning är att ett skidpar bara kan ha en kombination av skidvalla i taget

GRATTIS MÅNS THULIN!

2015 års Cramérpris har tilldelats Måns Thulin som disputerade i matematisk statistik vid Uppsala universitet.

Måns Thulin har både en djup och bred kunskap inom ämnet och visar detta på flera sätt. I sin avhandling 'On Confidence Intervals and Two-Sided Hypothesis Testing' har Måns Thulin visat en mycket hög vetenskaplig kvalitet med fem publicerade artiklar under eget namn samt ett manuskript. Utöver den välarbetade avhandlingen har Måns Thulin även visat upp en bred kunskap inom ämnet genom ytterligare fyra publikationer varav två inom tillämpningsområden.

»Vallaindustrin, understödd av TV-experter och andra förståsigpåare, har främst ägnat sig åt att mystifiera skidvallningen.»

vilket omöjliggör en relativ samtida jämförelse. En andra anledning är mängden kombinationer. Det finns minst tjugo alternativ för var och en av faktorerna och därför överstiger antalet kombinationer $20^4 = 160\,000$ av skidvallning för ett skidpar. En vallning tar runt två timmar och det skulle ta minst 100 år att valla med alla kombinationer om man vallar alla dagar året om, men tillåter sig matrast och normal sömn. Och när man väl är klar med det är det dags att sätta igång med alla andra skidpar.

De som inte använder försöksplanering brukar tackla ett beslutsproblem av den här typen på ungefär det här sättet: man testar de kombinationer man tror är mest intressanta att testa, antingen baserat på egen erfarenhet eller vallatillverkarens rekommendationer. Om de testerna är väl valda hänger alltså helt på den subjektiva bedömningen. Dessutom är det vanligt att bortse från interaktionseffekter eftersom de är svåra att identifiera och ingen vallatillverkare lämnar rekommendationer om hur olika faktorer kan tänkas samverka. Men åkaren är för att nå en topplacering beroende av bästa tänkbara skidor under tävlingen och eventuella högre ordningens interaktioner kan vara avgörande. Det mest olyckliga med att inte använda försöksplanering är att vid nästa tävling är man tillbaka på ruta ett. Man måste igen förlita sig på den subjektiva bedömningen eftersom inget systematiskt lärande har ägt rum.

Hur angriper vårt vallateam beslutsproblemet? För det första bortser vi från X eftersom det tar lång tid att lägga ny struktur och normalt sker det bara inför en ny säsong. För det andra har vi testat för interaktioner i hopp om att drastiskt kunna reducera antalet kombinationer. Om v är fart och en liten bokstav med efterföljande siffra för en faktor benämner nivå på faktorn, så har vi först valt ut fem likvärdiga skidpar, s1, s2, …, s5, så att v(s1)=...=v(s5). Farttestet görs med två lika tunga och lika stora åkare i nedförsbackar med hastigheter runt 20, 35 och 50 km/h med många upprepningar. Därefter jämför vi val av basparaffin mot obehandlade skidpar, med andra ord testar vi exempelvis om v(s1,B)>v(s2) för skidpar 1 mot skidpar 2. Låt säga att det tredje basparaffinet visade sig ge bättre glid så att v(s1,b3)>v(s2). I så fall lägger vi på tredje basparaffinet även på andra skidparet och undersöker om par 1 och 2 glider lika bra det vill säga om v(s1,b3)=v(s2,b3) och fortsätter på samma sätt för samtliga skidpar. Vi har med denna metodik kommit fram till att det inte finns någon nämnvärd interaktion mellan S och B. Och vidare, med en generalisering av detta förfarande, har vi nu slutit oss till att högre ordningens interaktioner kan negligeras. I alla fall för de cirka tjugotalet väder- och snöförhållanden (nogsamt dokumenterade) som vi haft möjlighet att testa i.

Hur angriper vårt vallateam beslutsproblemet?

Nästa fråga är antalet nivåer på faktorerna. Våra åkare har tre par skidor för respektive tävlingsform och det är ganska lätt att undersöka v(S|F) där F är väder- och snöförhållanden. Utbudet på basparaffin är dock stort. Genom att testa om v(s1,b1)=v(s2,b2)=v(s3,b3) och så vidare har vi dock slutit oss till att de etablerade vallaproducenternas basparaffin är likvärdiga och vi använder nu alltid ett och samma basparaffin. När det gäller faktorerna topping, H, och temporär struktur, R, är det lite mer komplicerat. Vid varje tävling testar vi ett tiotal kombinationer på H och R för att få fram bästa glidvallning. Jag tror dock att det är en mindre uppsättning vallor än vad de flesta skidmotionärer har i sitt garage eller var de nu förvarar vallan.

Den som likafullt misstänker att jag inte begriper komplexiteten i glidvallning av skidor och som accepterar vallaindustrin och diverse experters påstående om vallaprodukternas unikhet och den exceptionella kompetensen hos professionella vallare, bör begrunda följande. I alpin skidåkning tävlar man på skidor av samma plastmaterial, vallar med samma produkter och har minst lika stort behov av bra glid som i längdåkning. Hur många sändningstimmar i tv har ägnats åt de alpina skidlagens vallainsatser?

KENNETH CARLING

Qvintensen

POSTTIDNING B

SVERIGE
PORTO BETALT
PORT PAYÉ

Avsändare/Sender: Svenska statistikfrämjandet
c/o Carl Åkerlindh, Matematisk statistik,
Lunds universitet, Box 118, 221 00 Lund, Sweden

q-sampling shows good coverage

**Figurerna hör till artikeln "q-sampling shows good coverage" av Edgar Bueno.
Edgar vann Surveyföreningens pris för bästa uppsats på surveyområdet. Hans artikel finns på sidorna 28 och 29.**

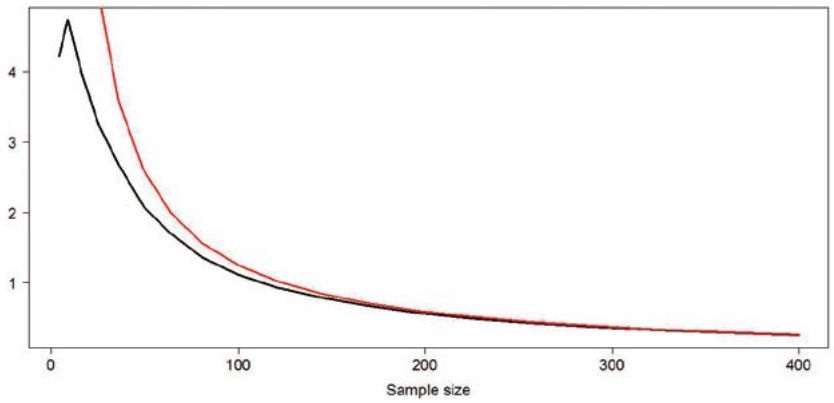


Figure 2. Comparison of simulated expectation of the variance estimator (black) with a variance approximation (red) for different sample sizes

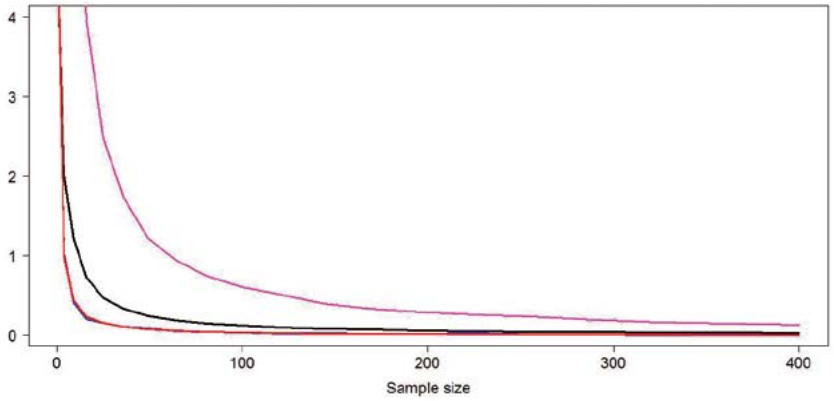


Figure 3. Comparison of the variance of four sampling strategies under different sample sizes: SRS (magenta), q-sampling (black), Pareto sampling (red) and systematic PPS (blue). The Pareto and systematic PPS designs are very close.

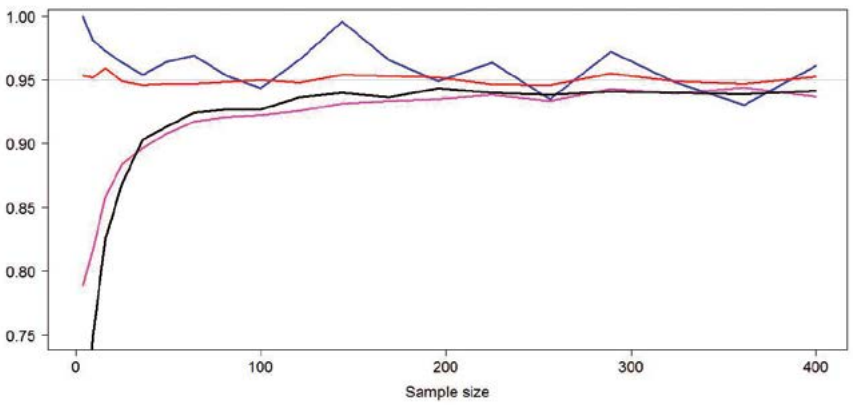


Figure 4. Comparison of the coverage of four sampling strategies under different sample sizes: SRS (magenta), q-sampling (black), Pareto sampling (red) and systematic PPS (blue).