


Qvintensen

NR 2 2024



Kommer maskiner ersätta statistiker?

SIDORNA 6-7

FRAMTIDENS ARBETSMARKNAD FÖR STATISTIKER

SIDORNA 4-5

HUR KAN EN STATISTIKER HJÄLPA EN ICKE-STATISTIKER ATT FÖRSTÅ?

SIDORNA 12-15

IQ- OCH EQ-TESTER - SÅ FUNGERAR DE VID REKRYTERING

SIDORNA 16-19

LIVET SOM STATISTIKER PÅ EUROSTAT - NÅGOT FÖR DIG?

SIDORNA 22-24

Innehåll

3	Redaktörens ruta Charlotte Ovesson	16	Tester i rekryteringar – så fungerar de! Mattias Strandberg
4	Framtidens arbetsmarknad för statistiker Karl Sigfrid och Jessica Franzén	20	Tjäna vetenskapen väl – bli statistisk sakkunniggranskare Erik Bülow
6	Stora språkmodeller, maskin- inlärning och statistik: Kommer statistiker att ersättas av maskiner? Måns Magnusson	22	Livet som statistiker på Eurostat – något för dig? Martin Karlberg
8	Demografiska förändringar, representativ data och AI: Ett föränderligt Sverige kräver nya angreppssätt för data- insamling Suzan Hourieh Lindberg	26	Givande kurs i läkemedels- epidemiologi Ida Hed Myrberg
10	Överförbara datadrivna lösningar inom svensk sjukvård: Möjligheter med att nyttja syntetiska data Patrik Rydén, Lars Mattsson, Sara D. Lundsten och Lenita Lindgren	27	Reflektioner kring fakta och feeling: Rapport från Statistikdagen med SCB Charlotte Ovesson
12	Statistisk litteracitet: Hur kan en statistiker hjälpa en icke-statistiker att förstå? Karin Landtblom	22	Två nya utgåvor av FENStatS nyhetsbrev Hans Alberg
		34	Minnesord över Lennart Bondesson Allan Gut
		35	Minnesord över Staffan Sollander Uno Davidsson

Föreningarna



28	Statistikfrämjandet Ordföranden har ordet	31	Cramér'sällskapet Ordföranden har ordet
29	Industriell statistik Ordföranden har ordet	32	Surveyföreningen Ordföranden har ordet
30	Föreningen för medicinsk statistik Ordföranden har ordet		



SVENSKA STATISTIKFRÄMJANDETS STYRELSE

Ordförande

Nancy Steinbach,
ordforande@statistikframjandet.se

Vice ordförande John Öhrvik

Kassör Annika Tillander
kassor@statistikframjandet.se

Sekreterare Mattias Strandberg
sekreterare@statistikframjandet.se

Ledamöter Magnus Pettersson, Hans Alberg,
Tea Unneback

Representant Surveyföreningen

Sara Scholtens

Representant FMS

Aaron Levine

Representant Industriell statistik

Hans Alberg

Representant Cramér'sällskapet

Anders Lundquist

E-post sekreterare@statistikframjandet.se

Webbplats www.statistikframjandet.se



Ansvarig utgivare

Nancy Steinbach

Redaktör

Charlotte Ovesson

Redaktion

Anders Sköllermo

Hans Alberg

Ingeborg Waernbaum

Marie Linder

E-post qvintensen@statistikframjandet.se

Produktion

Form och redigering: Mezzo Media AB

Tryckeri: Trydells Tryckeri AB

Annonser

Annonser i Qvintensen bokas med redaktören:
qvintensen@statistikframjandet.se.

Annonsskick på hemsida eller per e-post bokas med

Statistikfrämjandets sekreterare:

sekreterare@statistikframjandet.se



Charlotte Oveesson ny redaktör

Under hösten tillträdde jag som ny redaktör för Qvintensen. Som medlem i Svenska statistikfrämjandet har jag länge uppskattat tidningen och dess bredd. Vi statistiker arbetar inom olika fält och med olika metoder. Vår kompetens behövs verkligen inom många områden.

Jag är beteendevetaren som också råkade bli statistiker. Under min beteendevetenskapliga utbildning var jag inte nöjd med den kvantitativa metod vi fick lära oss. Jag ville veta mer, förstå mekanismerna bakom de siffror jag lärde mig att plocka fram i SPSS. Därför bestämde jag mig för att läsa en kurs i statistik vid Lunds universitet. Det blev inte en statistikkurs – det blev flera. Senare har jag förstått att jag inte är ensam om att ha upptäckt statistikämnet på det här sättet, att jag bara skulle läsa en liten kurs och att den öppnade mina ögon för hur otroligt spännande och användbart statistik är.

I mitt yrkesliv bygger jag ihop beteendevetenskapen med statistiken. Jag driver konsultföretaget Inclusive Insights som arbetar med jämlikhet och statistik. Främst arbetar jag med ideella organisationer som vill genomföra kartläggningar eller utvärderingar. Det blir en hel del enkäter, men även andra typer av data. Etik är ständigt närvarande och jag arbetar mycket med delaktighet, att de grupper som enkäten riktar sig till är närvarande i hela processen. Det bygger förtroende och leder till högre kvalitet på den data som samlas in.

Eftersom jag har en masterexamen i sociologi är jag självklart intresserad av statistikens roll i samhället. Vad vi väljer att mäta och vilken data vi väljer att samla in är ett intressant ämne. Varför mäter vi just detta? Vilka grupper osynliggörs i statistiken? Den statistik vi producerar som samhälle får stora konsekvenser för hur vi förstår världen och vilka politiska beslut som fattas. Statistik fyller en viktig funktion i ett demokratiskt samhälle, och här har vi som arbetar med statistik ett stort ansvar för att statistiken håller hög kvalitet.

Något annat som jag vurmar för är föreningsliv och folkbildning. För att dra mitt strå till stacken driver jag kontot statistikmedvetenhet på Instagram och Facebook. Där delar jag framför allt exempel på när statistik används på felaktiga eller vilseledande sätt. Eftersom statistik är viktigt för att vi ska kunna förstå världen och ta ställning i olika frågor behöver människor ha en god statistisk litteracitet. Det är viktigt att människor har både självförtroendet och verktygen för att ställa kritiska frågor till statistik de träffar på i sin vardag.

»Statistik fyller en viktig funktion i ett demokratiskt samhälle, och här har vi som arbetar med statistik ett stort ansvar för att statistiken håller hög kvalitet.»

Det här numret av Qvintensen är alltså mitt första nummer som redaktör. Temat är framtiden för statistiker. AI intar en central plats i detta nummer, av förklarliga skäl. Kommer statistiker att ersättas av AI? Jag ska redan nu avslöja svaret: nej. Det kommer fortfarande att behövas statistiker. När du läser artiklarna kommer det att bli tydligt varför.

Ett mål jag har som redaktör är att behålla bredden i Qvintensen. Att vi statistiker arbetar på olika sätt behöver speglas i vår medlemstidning. Alla medlemmar ska kunna hitta artiklar i tidningen som känns intressanta för dem. Jag vill också att tidningen ska inbjuda till nyfikenhet, att du ska vilja läsa artiklar om områden inom statistiken som du kanske inte är bekant med sedan tidigare. Det är alltid kul att lära sig något nytt.

Sist men inte minst vill jag konstatera att Qvintensen är en medlemstidning. Därför är din åsikt som medlem viktig. Jag vill att det ska vara låga trösklar för dig att höra av dig till Qvintensens redaktion med förslag och önskemål. Vad skulle du vilja läsa en artikel om? Har du tips på författare? Skulle du kanske vilja skriva något själv? Min inkorg står öppen.

Som Qvintensens nya redaktör tackar jag för förtroendet. Jag ska göra mitt bästa för att ge er en intressant medlemstidning som visar på bredden i statistikeryrket.

CHARLOTTE OVESSON
Redaktör

Framtidens arbetsmarknad

Utbildning i statistik har traditionellt varit en väg för den som vill bli statistiker. Att bli statistiker är fortfarande ett alternativ, men de flesta som har tagit en examen i statistik rekryteras numera till jobb med andra yrkestitlar. Det är tydligt att nya möjligheter har öppnats, och för att få en bättre bild av dagens arbetsmarknad har vi dragit nytta av öppen data från Arbetsförmedlingen. Med hjälp av AI-verktyg har vi kunnat sammanställa information ur jobbannonser, och därefter utforska arbetsmarknaden för statistiker i ett egenutvecklat webbaserat verktyg.

Sammanlagt publicerades drygt en miljon jobbannonser under 2023 i Arbetsförmedlingens platsbank. Den mest intressanta informationen finns i de löpande annonstexterna, som ostrukturerad data. För att sammanställa den sortens information hade vi för några år sedan varit hänvisade till tidskrävande manuellt arbete. I dag kan vi istället använda oss av stora språkmodeller som ChatGPT, som plockar ut specifika uppgifter ur en text och returnerar dem i ett strukturerat format.

En arbetsmarknad med nya yrkesroller

I webbverktyget som bearbetar och visualiserar vår data har vi kunnat skapa en överblick av arbetsmarknaden för studenter som läser statistik. I datamaterialet finns information till nytta för studenter som vill rusta sig för att möta den efterfrågan som finns på arbetsmarknaden. Vi har intresserat oss för de kompetenser som arbetsgivare vill att anställda ska ha i form av

kunskaper i programmering och andra digitala verktyg. Vi har också undersökt vilka yrkesroller som står öppna för den som har en statistikutbildning, och hur vanliga olika yrken är i olika delar av landet. Vi ville också veta vilka andra akademiska ämnen som är mest värdefulla att kombinera med statistik.

Delar av det vi ser stämmer väl överens med vad vi kunde förvänta oss, exempelvis i fråga om vilka yrkesroller som har växt fram. Ett par av de yrken som återkommer ofta i annonserna är dataanalytiker och data scientist, ofta med arbetsuppgifter som kräver både en utbildning i statistik och andra kunskaper, som programmering och databashantering.

Vid sidan av de nyare yrkesrollerna är det slående hur stor andel av jobbannonserna som kommer från den akademiska sektorn. Forskare med en ordentlig utbildning i statistik är eftertraktade som kollegor inom många olika forskningsfält. I Stockholm är Karolinska institutet den arbetsgivare som rekryterar flest statistikutbildade.

Arbetsmarknaden skiljer sig från region till region. I Stockholm står jobben som dataanalytiker för en betydligt större andel än i andra län, medan arbetsmarknaden i Örebro präglas av Statistiska Centralbyråns rekrytering av statistiker. I de flesta regioner fångar universiteten

och högskolorna upp en betydande andel av de statistikutbildade, vilket understryker ämnets betydelse för all typ av forskning.

Nya yrken ställer nya krav

Nya datayrken i näringslivet ger förstås spännande möjligheter för den som tar en examen inom statistik, men att förbereda studenterna ställer också krav på att utbildningen är moderniserad. I vårt webbverktyg ser vi vilka kompetenser arbetsgivare helst vill att studenterna ska ha med sig, och framför allt är det viktigt att kunna programmera. Python och R är de språk som oftast nämns i annonserna. En annan värdefull kompetens enligt många arbetsgivare är att kunna hantera databaser med hjälp av SQL.

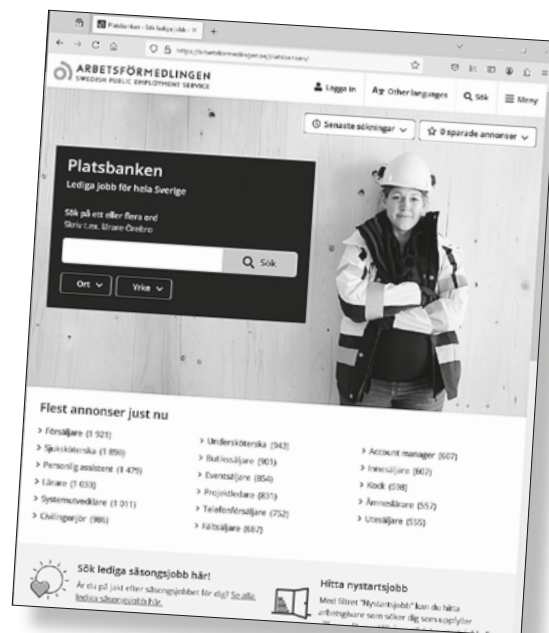
Det är inte alltid självklart vilka digitala verktyg som är viktigast för en student att lära sig. För den som vill ha ett mer traditionellt jobb som statistiker är SAS-kunskaper fortfarande en av de mest efterfrågade kompetenserna, tillsammans med

R. I andra änden av spektrumet hittar vi annonser för data scientists, och då dominerar Python helt. I många av annonserna för data scientists nämns också datavetenskap och ingenjörsutbildningar som lämpliga bakgrunder, vilket understryker fokuset på programmering.

Programmeringsspråket R dominerar framför allt inom forskar- och högskolesektorn, men är ett högt rankat språk oavsett vilken sektor

»Programmeringsspråket R dominerar framför allt inom forskar- och högskolesektorn, men är ett högt rankat språk oavsett vilken sektor vi talar om.»

Med hjälp av AI-verktyg har artikelförfattarna sammanställt information ur Platsbankens jobbannonser, och därefter utforskat arbetsmarknaden för statistiker i ett egenutvecklat webbaserat verktyg.



för statistiker

vi talar om. Att lära sig R är ett alternativ som håller dörren öppen till flera olika typer av jobb. Med en syntax som liknar den för många andra programmeringsspråk underlättar det dessutom för den som även vill lära sig ytterligare språk som Python eller Javascript.

Jobbannonserna som vi analyserar visar att statistik ofta nämns tillsammans med andra utbildningsbakgrunder. Vanligast är att annonserna nämner matematik, företagsekonomi, nationalekonomi eller datavetenskap. För en data scientist är matematik och datavetenskap särskilt viktigt. För dataanalytiker, som är en av de största yrkesgrupperna, ser arbetsgivarna gärna kunskaper i företagsekonomi eller nationalekonomi. Ekonomistudier vid sidan om statistiken är också värdefulla för den som vill jobba som statistiker eller utredare.

AI och framtiden för statistik

Vi kan utgå från att AI-utvecklingen på ett eller annat sätt kommer att förändra organisationers arbete med dataanalys. Hur det påverkar arbetsmarknaden och efterfrågan på statistisk kompetens är svårt att förutsäga, men enklare arbetsuppgifter som att visualisera deskriptiv statistik kan säkert komma att automatiseras oftare.

Yrkesverksamma med kunskaper i statistik kommer även fortsättningsvis att behövas, både för metodutveckling och för tillämpning. Metodutveckling förutsätter kunskaper om hur statistiska modeller fungerar under motorhuven, och att utveckla metoder för nya tillämpningar är en uppgift som kräver ett mått av kreativitet.

De som tillämpar statistisk måste fortfarande

förstå problemet som ska lösas, och därefter formulera det i statistiska termer. Sedan ska resultatet tolkas i ljuset av organisationens eller forskningsprojektets syften. Även om AI-system kan vara hjälpsamma för vart och ett av de här stegen behövs människor som kan föra projektet från start till mål.

AI-system är inte bara ett verktyg som kan användas för att analysera data, utan också ett verktyg som i stor skala kommer att ge tillgång till ny data att analysera. Det verktyg som vi använder för att kartlägga jobben för statistikstudenter är ett exempel på det. Den här utvecklingen ökar behovet av statistisk kompetens.

Kunskap om nya statistikyrken måste nå studenterna

Att arbetsmarknaden har förändrats har blivit uppenbart för de flesta som jobbar med statistik. En utmaning är att göra studenter medvetna om förändringarna, och om de dörrar som har öppnats. Allt för få känner till att statistik är kärnan i data science och dataanalys, och det är kanske inte så konstigt. Nya studenter från gymnasiet har sällan bred kunskap om vilka yrken som finns och vilka utbildningar som krävs. De studenter som ändå fördjupar sig i statistik har ofta halkat in på ämnet i samband med studier i andra ämnen. När de väl tar en kurs i statistik upptäcker många att det är ett kul och verklighetsnära ämne, och att de blir utrustade med handfasta färdigheter.

Samtidigt som nya arbetsmarknader öppnar sig för statistikutbildade är det relativt få som examineras inom ämnet. Siffror från Universitetskanslerämbetet visar att antalet studen-

ter som tagit kandidatexamen i statistik per år har legat ungefär i spannet mellan 100 och 150 studenter per år, medan 50 till 100 har tagit en masterexamen. Det kan jämföras med att 700 till 900 nationalekonomer och 3 000 till 5 000 företagsekonomer examineras ett vanligt år.

Vi ser framför oss att de som fördjupar sig i statistik kommer att få allt större valmöjligheter. Det är ett resultat av att organisationer inte längre frågar sig om de ska ägna sig åt dataanalys, utan snarare hur de ska utvinna insikter ur sin data.



KARL SIGFRID
JESSICA FRANZÉN

Referenser

- Universitetskanslerämbetet. (2024). Studenter och examinerade i högskoleutbildning på grundnivå och avancerad nivå. www.uka.se/vara-resultat/statistik

STORA SPRÅKMODELLER, MASKININLÄRNING OCH STATISTIK:

Kommer statistiker att ersä

Maskininläring och artificiell intelligens har under de senaste åren gått från teoretiska koncept till praktiska verktyg som kan analysera, förutse och till och med formulera svar på mänskliga frågor.

När generativa modeller som ChatGPT och andra stora språkmodeller introducerades för världen blev det snabbt klart att dessa skulle kunna påverka hur vi ser på dataanalys, statistik och, kanske mest spännande, hur vi använder dessa verktyg för att förstå allt från proteiner till sociala interaktioner.

Men innebär dessa framsteg att traditionella statistiska metoder och statistiker är på väg att bli överflödiga? Och hur mycket kan egentligen en språkmodell lära sig av data?

Tidigt i december 2022 stod jag framför mina studenter på Ekonomikum i Uppsala, redo att hålla en föreläsning om transformerarkitekturer för neurala nätverk. Denna gång var intresset i salen mer påtagligt än tidigare år. Några dagar innan min föreläsning hade ChatGPT lanserats, och plötsligt ville alla veta mer om hur denna teknik egentligen fungerade – särskilt transformerarkitekturen som låg bakom den nya modellen som precis tagit världen med storm.

Redan året därpå, hösten 2023, var kursens innehåll i behov av en uppdatering. För att möta studenternas ökade intresse och det nya modellernas genomslag lade jag till en föreläsning om stora språkmodeller. ChatGPT hade gjort att intresset för språkmodeller sträckte sig långt utanför forskarvärlden och började påverka allt från samhällsdebatten, med flera framstående forskare som uppmanade till ett stopp för forskning och utveckling av nya stora språkmodeller, till hur vi tänkte kring statistik och dataanalys. Och även om ChatGPT var en av de mer kända modellerna så var den teknologiska revolution som låg bakom den pågått flera år i forskningsmiljön. Sedan släppet i november 2022 hade det plötsligt blivit självklart för många att vi stod inför en ny era där dessa stora språkmodeller kunde komma att spela en avgörande roll.

Vad är AI, stora språkmodeller och maskininläring?

Artificiell intelligens, AI, är ett brett begrepp som innefattar allt från enkla sökalgoritmer till djupa neurala nätverk. Maskininläring, en gren inom AI, handlar om att dra slutsatser från data med hjälp av matematiska modeller – inte helt olik statistiska metoder som linjär och logistisk regression. Maskininläring skiljer sig dock från statistiken genom sitt starka fokus på prediktion: målet är inte nödvändigtvis att förstå vad som gör ett e-postmeddelande till ett spam-mejl, utan det viktiga är att kunna förut säga vilka mejl som är spam.

En annan distinktion ligger i den typ av data som brukar användas. Min vän och kollega Leif Jonsson, nu maskininläringsexpert på Ericsson, påpekade att statistiker ofta "tror att världen kommer i form av Excel-filer". Det ligger en hel del i den observationen. Inom maskininläring är data, som till exempel bilder, ljud och text, ofta högdimensionell och svåranalyserad med klassiska statistiska metoder. Djupa neurala nätverk har istället dominerat analysen inom dessa områden. Under 2010-talet revolutionerades datorseende, medan transformerarkitekturer, som utvecklades i slutet av decenniet, öppnade för stora och djupa neurala nätverk för textuell data.

Transformerbaserade språkmodeller utvecklades ursprungligen för att förut säga nästa ord i en text baserat på tidigare ord. Genom att fokusera på detta till synes enkla problem kunde modellerna läras upp att generera längre och mer sammanhängande sekvenser av ord. År 2018 släpptes den första GPT-modellen, "Generative Pretraining Transformer" med 117 miljoner parametrar, och utvecklingen tog fart. GPT-2 kom året därpå med 1,5 miljarder parametrar, och GPT-3, som lanserades 2020, innehöll 175 miljarder parametrar – tillräckligt för att skapa exempelvis artiklar som var svåra att skilja från artiklar skrivna av människor. År 2022 lanserades InstructGPT, en version av GPT-3 som tränats på mänskliga instruktioner för att kunna svara bättre på frågor och instruktioner. Med instruktionstränade modeller hade det sista problemet lösts, hur människor enkelt kan interagera med dessa stora modeller. Ett par månader senare släpptes ChatGPT.

Under de senaste åren har dessa modeller visat sig vara förvånansvärt bra på att sammanfatta texter och ge svar på även till synes komplexa frågor. Modellernas effektivitet tros bero på målet med träningen, att predicera enskilda ord i ett sammanhang. I en mening som "Joe Biden träffade Xi Jinping för att tala om ..." skulle ett ord som "Qvintensen" inte vara särskilt troligt. När jag ber ChatGPT att komplettera meningen, svarar modellen "relationerna mellan USA och Kina" (det sanna svaret i mitt exempel är "AI"). Så för att kunna predicera nästa ord, behöver modellen tränas för att skapa representationer av Joe Biden och Xi Jinping för effektiva prediktioner. Genom stora datamaterial utvecklar dessa modeller en representation av världen, vilket gör dem duktiga på att skapa text som känns naturlig och kontextuellt korrekt. Men det gör också att den text som genereras kan vara direkt felaktig.

Maskininläring och statistikämnet

Ur ett statistiskt perspektiv är de moderna maskininlärningsmodellerna inget annat än

Skattas av maskiner?

stora ickeinjära modeller som likt klassiska statistiska modeller skattas från data. Skillnaden ligger i målet: maskininlärningsmodeller fokuserar på prediktion, oavsett om det gäller nästa ord i en text, bostadspriser eller om det finns en katt i en bild. Detta prediktionsfokus har paralleller inom statistiken. Akaikes informationskriterium (AIC), en uppskattning av en modells prediktionsfel, är ett exempel på prediktionens betydelse för modelljämförelser inom klassisk statistik. Många maskininlärningsalgoritmer har därför funnit sin väg in i statistiken och används idag inom ekonomisk modeller, väderprognoser och datorseende i medicinska tillämpningar.

Djupa neurala nätverk innebär nya möjligheter för statistikämnet genom att de kan skatta komplexa modeller på stora och varierade datamängder, inklusive text och bilddata. Men de innebär också begränsningar: många av modellerna är så komplexa att det är svårt att förstå hur de kommer fram till sina prediktioner. Detta gäller särskilt stora språkmodeller. De kräver också stora mängder data för att tränas korrekt.

Ur ett strikt statistiskt perspektiv är det inte trivialt att använda stora språkmodeller. Ett exempel på detta en ESO-rapport om det svenska remissväsendet, som beräknas publiceras i slutet av 2024 eller början av 2025. Rapporten undersöker hur tydliga svenska myndigheter är i sina remissvar till regeringen, och språkmodeller användes för att klassificera remissvarens tydlighet. Denna språkmodellbaserade klassifikation möjliggör statistiska slutsatser om olika myndigheters remissförfarande, genom att dra slutsatser betingat på klassifikationerna. Problemet är bara att det många gånger kan bli fel i dessa klassifikationer, något också framgår av rapporten. Denna kombination av maskininläring och statistisk inferens har börjat kallas "prediction-powered inference" eftersom det krävs att den statistiska analysen även tar hänsyn till osäkerheten i maskininlärningsmodellen, särskilt vid användning av komplexa, djupa neurala nätverk. Att bortse från denna osäkerhet kan leda till skevheter i resultaten, något som

är extra tydligt när modellen tränats på textuell data som ofta innehåller sociala skevheter och "fördomar".

Att använda maskininlärningsmodeller för statistiska ändamål innebär ett utökat ansvar för statistiker, som måste säkerställa att modellerna används korrekt och transparent. Eftersom dessa modeller kräver stora mängder data för träning, är de ofta förtränade på omfattande datamaterial som därefter anpassas till specifika analysbehov. Exempelvis har Kungliga bibliotekets transformer-modell tränats på nyhetstext, svenska offentliga utredningar och Wikipedia, medan försök att träna en svensk GPT-3 även inkluderade material från Familjeliv och Flashback – vilket kan vara mer tveksamt. Därför är det avgörande att tydligt redovisa vilka modeller som används, vilken data de tränats på och hur de appliceras i statistiska sammanhang, och hur eventuella skevheter hanteras.

I ett projekt vid Brottsförebyggande rådet (Brå), finansierat av Länsförsäkringars forskningsfond, undersöker jag och min doktorand, Hannes Waldetoft, just hur djupa neuronnät kan användas i statistiska tillämpningar. Vi undersöker om vi kan använda transformermodeller för att identifiera hatbrott utifrån texten i polisanmälningar, något som idag är så pass svårt att det måste göras manuellt av experter på Brå. Genom att kombinera transformerarkitekturer med urvalsteori arbetar vi för att korrigera för de skevheter i skattningarna som uppstår när neurala nätverksmodeller används för klassificera hatbrott. En form av prediktionsbaserad inferens.

Hur ser framtiden ut för statistiker?

I år var två av Nobelprisen tydligt kopplade till maskininläring och neurala nätverk: fysikpriset belönade arbetet med neurala nätverk,

och kemipriset belönade AlphaFold, en djup neural nätverksmodell för prediktion av tredimensionella proteinstrukturer. Detta visar tydligt att maskininläring är här för att stanna. Vad innebär då dessa framsteg för statistiker?

Idag kan ChatGPT ge rimliga svar på statistiska frågor på avancerad nivå, och utvecklingen går mot en ökad automation av delar av de statistiska arbetsuppgifterna. Men statistisk analys sker sällan i ett vakuum; det krävs förståelse för kontexten och syftet med analyserna, och språkmodeller begår också ofta fel. Det betyder att exempelvis språkmodeller idag inte kan ersätta statistiker, men att statistiker kan använda dessa modeller som en partner

»... språkmodeller idag kan inte ersätta statistiker, men statistiker kan använda dessa modeller som en partner och inspiration i sitt arbete.»

och inspiration i sitt arbete. Språkmodeller kan vara till hjälp för att utveckla hypoteser, ge nya idéer och effektivisera analysarbetet.

Att själv börja använda AI och språkmodeller som statistiker kräver kunskaper i programmering, exempelvis i Python, R eller

Julia. Det kräver också att bekanta sig med verktyg för djupa neurala nätverk som TensorFlow och PyTorch. För den som är intresserad finns ett stort utbud av onlinekurser och universitetsutbildningar, de flesta masterutbildningar i statistik innehåller idag en eller flera kurser i maskininläring och djupa neurala nätverk.

Det sägs ibland att maskininlärningsmodeller inte kan ersätta radiologer, men att en radiolog som använder maskininläring kan ersätta flera kollegor. Detsamma kan nog sägas om statistiker. Med stöd av dessa typer av modeller kan vi snabbare hitta lösningar och utveckla nya metoder för statistisk analys. Dessa nya modeller blir ett kraftfullt verktyg för prediktioner, analys av komplex data, och inte minst, som stöd, i analysen och i kommunikation av resultaten.

Ett föränderligt Sverige kräver nya angreppssätt för datainsamling

Sverige har genomgått och går fortfarande igenom en av de mest betydande demografiska förändringarna i modern tid, vilket påverkar samhällsstrukturen på djupet. Det innebär enormt många möjligheter som vi behöver bli bättre på att ta tillvara. Idag är 20 procent av befolkningen utrikes födda, 26 procent har utländsk bakgrund (är utrikes födda eller födda i Sverige med två utrikes födda föräldrar) och 33 procent har minst en utrikes född förälder. Bland barnen är siffran ännu högre; hela 39 procent har minst en utrikes född förälder. Befolkningen blir äldre och födelsetalen minskar. Den snabbast växande demografiska

gruppen är svensken med en utrikes född förälder. Vi ser dessutom geopolitiska spänningar och klimatförändringar som leder till massflytt och omfattande demografiska rörelser.

Det Sverige vi lever i idag är inte detsamma som för tjugo år sedan, och det kommer att utvecklas och förändras ytterligare under de kommande decennierna. För att förstå och utvecklas med dessa förändringar behöver vi anpassa våra metoder för datainsamling och analys. Traditionella metoder som ofta innebär telefonintervjuer, pappersenkäter med i bästa fall QR-kod räcker inte längre till när befolkningen blir allt mer diversifierad och perspektivrik för att inte tala

om digital. Vi måste utveckla robusta och flexibla system som kan fånga upp den mångfald av röster och perspektiv som finns och som kommer att finnas i samhället. För guldets är just de olikheter som finns bland svenskarna – det är nämligen där insikterna hittas.

Underrepresentation – en utmaning för tillförlitlig data

Trots de demografiska förändringarna förblir flera grupper systematiskt underrepresenterade i undersökningar. Personer under 25 år och de över 60 år, som utgör 29 respektive 21 procent av befolkningen, nås sällan effektivt genom traditionella metoder. Likaså individer med läs- och skrivsvårigheter, som omfattar 20 till 25 procent av befolkningen. Svenskar med utländsk bakgrund är också en grupp som sällan inkluderas, där en vanlig missuppfattning är att anledningen är språkliga barriärer. Det handlar snarare om en brist på förtroende för undersökningsbolag och beställande institutioner. Detsamma gäller personer som bor i utsatta områden.

Sammantaget innebär detta att cirka 35 till 40 procent av befolkningen ofta inte inkluderas i statistiska undersökningar. Detta väcker allvarliga frågor om tillförlitligheten i data och statistik. Om en så stor del av befolkningen inte representeras, kan vi då verkligen lita på resultaten?

»Svenskar med utländsk bakgrund är också en grupp som sällan inkluderas, där en vanlig missuppfattning är att anledningen är språkliga barriärer. Det handlar snarare om en brist på förtroende för undersökningsbolag och beställande institutioner.»

Konsekvenserna av data-marginalisering

Vår organisation har genomfört skuggundersökningar som är representativa utifrån kön, ålder, geografisk spridning, härkomst, födelseland med mera. Resultaten visar betydande avvikelser jämfört med traditionella studier, särskilt när det gäller åsikter och behov hos utrikes födda och personer med utländsk bakgrund. Exempel är studier kopplat till mediekonsumtion samt trygghet där vi har sett helt felaktiga resultat. Dessa avvikelser pekar på allvarliga konsekvenser. Beslut fattas på underlag bestående av ofullständig information, vilket leder till felprioriteringar i resursfördelning och samhällsplanering. Företag och organisationer misslyckas med att tillgodose behoven hos en stor del av befolkningen, vilket resulterar i minskad effektivitet och förlorade möjligheter. När människor inte känner sig representerade minskar förtroendet för institutioner, vilket kan förstärka samhällsklyftor och leda till lägre medborgarligt engagemang såsom lägre valdeltagande, någonting man ser bland marginaliserade och underrepresenterade demografiska grupper.

Den mänskliga rösten: en ovärderlig resurs

Att inte inkluderas eller ha sitt perspektiv representerat i undersökningar lämnar djupa spår på individnivå. Känslor av utanförskap och osynlighet kan påverka social sammanhållning och öka klyftorna mellan olika grupper. Människans röst är därför ovärderlig i processen att förstå samhället. Den ger oss insikter som går bortom siffror och statistik, hjälper oss att förstå de bakomliggande orsakerna till beteenden och attityder och är avgörande för att fatta välgrundade beslut.



FOTO: SHUTTERSTOCK

AI kan fungera som ett kraftfullt verktyg för att förbättra användarvänlighet och tillgänglighet, anpassa undersökningar för personer med funktionsnedsättningar och underlätta snabbare analys. Men det är viktigt att betona att AI bör användas som ett stöd för människan, inte som en ersättning.

Samtidigt som vi står inför dessa demografiska utmaningar utvecklas teknologin i en rasande takt. Artificiell intelligens (AI) har potentialen att revolutionera sättet vi samlar in och analyserar data på. Genom avancerade algoritmer kan AI hantera stora datamängder, identifiera mönster och ge insikter som tidigare var svåra att upptäcka. Men här uppstår en kritisk fråga: kan AI ersätta människan i processen att förstå en så dynamisk och komplex verklighet?

Vissa startups har börjat ersätta mänskliga panelister med AI för att förenkla arbetet med enkätundersökningar. Även om detta kan verka som en effektiv lösning finns det flera problem med detta tillvägagångssätt. AI är endast så bra som den data den tränas på. Om träningsdatan är skev eller underrepresenterar vissa grupper kommer AI-modellerna att reproducera och förstärka dessa biaser. I en tid av snabba demografiska förändringar blir detta särskilt problematiskt. Framtiden kommer inte alltid att spegla historien, och AI som är baserad på historisk data kan ha svårt att förutse och anpassa sig till nya trender och mönster. Prediktivitet i maskininlärning är avancerad men långt från effektiv i denna form, inte minst i en tid av oförutsägbara förändringar i demografi, attityder och delvis även värderingar.

Människan och AI: en symbios snarare än en ersättning

Jag är personligen mycket positiv till AI och dess potential att förbättra många aspekter av vårt arbete. AI kan fungera som ett kraftfullt verktyg för att förbättra användarvänlighet och tillgänglighet, anpassa undersökningar för personer med funktionsnedsättningar och underlätta snabbare analys. Men det är viktigt att betona att AI bör användas som ett stöd för

människan, inte som en ersättning. Människans förmåga att förstå kontext, emotionella nyanser och kulturella subtiliteter är något som AI ännu inte kan matcha. Att ersätta människan med AI riskerar att avhumanisera processen och missa viktiga insikter som endast kan fås genom direkt mänsklig interaktion. Jag är ofta inne på Github och läser inte sällan skepticismen bland utvecklare kring just AI-paneler.

För att AI ska vara effektivt och etiskt måste det vara människocentrerad, det vill säga utformat med människan i centrum. Detta innebär att teknologin ska anpassas efter mänskliga behov och värderingar, inte tvärtom. Genom att sätta människan i fokus kan vi säkerställa att AI används på ett sätt som gynnar samhället som helhet.

En integrerad strategi för framtidens utmaningar

För att möta de komplexa utmaningar som följer med demografiska förändringar och teknologisk utveckling behöver vi en integrerad strategi som kombinerar mänsklig erfarenhet med teknologisk innovation. Vi bör utveckla inkluderande datainsamlingsmetoder genom att använda flerspråkiga undersökningar, alternativa kommunikationskanaler och anpassade metoder för att nå underrepresenterade grupper. Flerspråkiga undersökningar behövs för att skapa förtroende, inte för att man inte talar språket.

Vidare bör vi implementera etiska riktlinjer för AI för att säkerställa att AI-system är transparenta, ansvarstagande och fria från diskriminerande bias. Slutligen måste vi fortsätta att

prioritera mänsklig insikt och använda AI som ett verktyg för att stödja, inte ersätta, mänsklig analys och beslutsfattande.

Summa summarum: navigera framtiden med människan i centrum

För att möta de utmaningar som följer med en snabbt föränderlig demografi och teknologisk utveckling krävs en nyanterad och inkluderande ansats. Genom att säkerställa att alla röster hörs och att data verkligen speglar samhällets mångfald, kan vi skapa beslutsunderlag som är både tillförlitliga och rättvisa. Tekniken, inklusive AI, erbjuder kraftfulla verktyg för detta ändamål, men endast om den används på ett sätt som sätter människan och etiken i första rummet.

Att investera i representativ datainsamling och att värdesätta den mänskliga erfarenheten är inte bara nödvändigt för att fatta välgrundade beslut, det är också grunden för ett samhälle där varje individ känner sig sedd och inkluderad. Genom att förena mänsklig insikt med teknologisk innovation kan vi navigera mot en framtid där data inte bara är siffror, utan en reflektion av alla människors verklighet.

SUZAN HOURIEH LINDBERG
Medgrundare och vd, Perspektivo

Överförbara datadrivna lösningar inom svensk sjukvård

MÖJLIGHETER MED ATT NYTTJA SYNTETISKA DATA

Sverige har en vårdgaranti som innebär att patienter inte ska vänta längre än 90 dagar på operation. Trots detta tvingas många patienter vänta betydligt längre. Enligt statistik från SKR hade cirka 85 000 patienter i september 2024 väntat mer än 90 dagar (Skr.se, 2023). Operationsköerna har vuxit sedan pandemin, samtidigt som svensk sjukvård brottas med svårigheter att rekrytera och behålla personal. Situationen belyser ett akut behov av nya arbetssätt för att öka effektiviteten och optimera resursanvändning inom vården.

Ett potentiellt förbättringsområde är att använda de stora mängder data som varje patient genererar före, under och efter en operation. Genom att bygga maskininlärningsmodeller kan patientflöden och personalresurser optimeras. Problemet är dock att denna data inte alltid är lättillgänglig hos Sveriges regioner. Vissa regioner saknar helt lagring av relevant data, medan andra har komplicerade regelverk som begränsar åtkomst.

Ett utvecklingsarbete pågår inom Region Västerbotten för att förbättra patientflöden inom postoperativ vård. LAIs är ett AI-baserat system för att optimera patientflödet genom att använda maskininläring för att förutse vårdtider baserat på patientdata, vilket möjliggör effektivare planering och resursanvändning.

Ett delproblem är att prediktera patienternas behandlingstider på de postoperativa avdelningarna, d.v.s. hur länge patienterna behöver vårdas på avdelningarna. Maskininläring och historiska data från sjukhusen i Lycksele, Skellefteå och Umeå används för att bygga en modell för att prediktera patienternas behandlingstider. Problemet innehåller exempel på många av de statistiska utmaningar som finns när vi tar fram prediktiva kliniska modeller inom sjukvården.

Statistiska utmaningar med prediktiva modeller inom hälso- och sjukvården

Vi är övertygade om att utmaningarna som diskuteras här förekommer i många kliniska

tillämpningar, även om de inte alltid beaktas. Nedan listar vi ett flertal utmaningar, varav några kan lösas med hjälp av syntetiska data.

- Maskininlärningsmodeller kräver ett stort antal observationer vid träning, speciellt när antalet förklarande variabler är stort. Patientdata innehåller i regel känsliga personuppgifter (Digg.se, 2023), vilket gör det svårt att dela data mellan sjukhus. Anonymisering av data eller att arbeta med syntetiska data är två sätt att möjliggöra denna delning av data.
- Vilka variabler som observeras och samlas in varierar mellan sjukhusen, vilket förklaras av att ett sjukhus inte kan, eller har behovet av, att observera alla variabler. En analysstrategi är att endast inkludera variabler som observeras på samtliga sjukhus, men detta riskerar att leda till sämre prediktiva modeller. En alternativ lösning är att ersätta saknade data med syntetiska data.
- En relaterad utmaning är att olika insamlingsmetoder används på sjukhusen. Till exempel kan variabel observeras med metod 1 på några sjukhus och metod 2 på andra sjukhus, vilket kan introducera en systematisk skillnad. Problem kan lösas genom normalisering, till exempel kvantilnormalisering (Bolstad et al., 2003) baserat på samtliga observationer eller på en delmängd av homogena observationer.
- Behandlingstiden på de postoperativa avdelningarna är normalt skillnaden mellan inskrivningstiden och utskrivningstiden, men nattetid (20:00-08:00), när bemanningen är begränsad, skrivs färdigbehandlade patienter först ut på morgonen, vilket ger upphov till intervall-censurerade observationer. Intervall-censurerade data för maskininläring har fått viss uppmärksamhet (Wang, Li

Sjukhus	Antal	x1	x2	x3	x4	x5	x6	y
Umeå	50 000	data	data	data	NA	data-dc	data-m1	data-dc
Skellefteå	10 000	data	data	NA	data	data-dc	data-m2	data-dc
Lycksele	3 000	data	data	data	data	data-dc	data-m1	data-dc

Tabell 1. Tabellen beskriver historiska hälsosjukvårdsdata för sjukhusen i Umeå, Skellefteå och Lycksele. Antalet historiska observationer varierar mellan sjukhusen. Data består av de förklarade variablerna x och responsvariabeln y , som är behandlingstiden vid de postoperativa avdelningarna. De förklarade variablerna kan delas upp i sex principiella grupper $\{x_1, \dots, x_6\}$ med $\{p_1, \dots, p_6\}$ förklarande variabler. Gruppen x_1 består av känsliga data som inte kan flyttas mellan sjukhusen. Gruppen x_2 består av data som observeras på ett likvärdigt sätt på sjukhusen. För grupperna x_3 och x_4 observeras inte variablerna vid alla sjukhus (saknade variabler). För gruppen x_5 är en del av observationerna censurerade (censurerade data). För variablerna i gruppen x_6 finns det systematiska skillnader i hur observationerna har genererats (data med systematiska avvikelser). Responsvariabeln y innehåller också en blandning av ocensurerade och censurerade data. För respektive variabel förekommer det även att enskilda observationer saknas (saknade data) och att enskilda observationer är felaktiga (felaktiga data).



Artikelförfattarna, från vänster Patrik Rydén, Sara Lundsten, Lars Mattsson och Lenita Lindgren.

and Reddy, 2019), men är ett område där det behövs mer forskning.

- Slutligen kan enskilda variabler innehålla både saknade värden och felaktiga värden. Eftersom maskininlärning i regel kräver kompletta data, behöver saknade värden imputeras. Även här erbjuder syntetiska data en lösning.

Att använda maskininlärning på hälsodata för att ta fram prediktiva modeller har således flera utmaningar inklusive: inte tillräckligt med data, känsliga data, saknade variabler, data med systematiska avvikelser, censurerade data, saknade observationer och felaktiga observationer, se Tabell 1. Syntetiska data erbjuder en lösning på flera av utmaningarna, men konkurrerar samtidigt med andra alternativa lösningar.

Nedan beskriver vi kort idén bakom en copula-baserad metod för att generera syntetiska data. Vi har valt denna metod för att den är relativt enkel att beskriva men det finns fler alternativa metoder för att simulera syntetiska data, exempelvis metoder baserade på neurala nätverk.

Simulering av syntetiska data

Vår tanke är att ge en intuitiv beskrivning av en metod som används för att generera syntetiska data. Vi börjar med att införa några begrepp.

- Låt observationerna x vara observationer av de stokastiska variablerna X som har en multivariat fördelning F_X . Notera att X innefattar p antal variabler bestående av både de förklarade variablerna och responsvariabeln Y .
- Om fördelningen F_X är känd vet vi allt om X , inklusive alla tänkbara marginalfördelningar, vilket betyder att vi kan simulera saknade data, både observationer och variabler, och att vi kan generera data som är helt syntetiska. I praktiken känner vi inte till fördelningen F_X

men om vi kan göra bra skattningar av fördelningen så kan vi simulera syntetiska data med liknande egenskaper som originaldata. Här finns det emellertid flera utmaningar. En är att fördelningen och dess marginalfördelningar i regel varierar mellan olika sjukhus eftersom patientunderlaget och operationerna varierar. Detta medför att fördelningen måste skattas separat för respektive sjukhus, vilket är svårt att göra för sjukhus med relativt få observationer. En lösning på detta är att omformulera skattningsproblemet via en copula (Rüschendorf, 2009).

Tricket med metoden är att det existerar en multivariat fördelning C (en copula) som beskriver beroendestrukturen mellan X -variablernas rangordningsvärden. Fördelningen C kan användas för att simulera en slumpvektor $\{U_1, \dots, U_p\}$ i en enhetskub, där $[0 \leq U_i \leq 1], i = 1, \dots, p$. Av Sklars berömda sats från 1959 följer att

$$\{X_1, \dots, X_p\} = \{G_1(U_1), \dots, G_p(U_p)\}$$

där G_i betecknar inversen av F_i , $i = 1, \dots, p$. Det finns i regel tillräckligt många observationer för att skatta marginalfördelningarna och deras inverser individuellt för sjukhusen.

Problemet som kvarstår är att skatta fördelningen C . Även om marginalfördelningarna varierar mellan sjukhusen så kan ett rimligt antagande vara att beroendestrukturen är likartad, vilket gör att skattningen av C kan baseras på data från samtliga sjukhus. Skattningen av C kan göras parametrisk under olika antaganden eller icke-parametrisk. En utmaning här är att vi i regel har en blandning av kontinuerliga, diskreta och kategoriska variabler.

Avslutande kommentarer

Att det finns mängder av hälsodata i svensk sjukvård råder det ingen tvekan om, men att använda den till att bygga prediktiva modeller som kan ge verklig nytta ute i vården, är lättare sagt än gjort.

Ett stort hinder för att få spridning och stor effekt av sådana tekniska lösningar är just den stora datavariationen mellan olika regioner, sjukhus eller till och med kliniker inom samma sjukhus. Möjligheterna med syntetiska data som kan överbygga denna datavariationsskillnad är enorma, men det kräver både etiska förhållningssätt och klinisk kunskap lika mycket som tekniskt kunnande för att säkerställa kvaliteten.

PATRIK RYDÉN
LARS MATTSSON
SARA D. LUNDSTEN
LENITA LINDGREN

Referenser

- Bolstad, B.M., Irizarry, R.A., Astrand, M. and Speed, T.P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*, 19(2), pp.185–193. doi:<https://doi.org/10.1093/bioinformatics/19.2.185>.
- Digg.se. (2023). *Känsliga och extra skyddsvärda personuppgifter* | Digg. [online] Available at: <https://www.digg.se/kunskap-och-stod/metodstod-for-dataskydd-vid-innovation/rattslig-bedomning-av-personuppgiftsbehandlingen/kansliga-och-extra-skyddsvardva-personuppgifter> [Accessed 30 Oct. 2024].
- Rüschendorf, L. (2009). On the distributional transform, Sklar's theorem, and the empirical copula process. *Journal of Statistical Planning and Inference*, 139(11), pp.3921–3927. doi:<https://doi.org/10.1016/j.jspi.2009.05.030>.
- Skr.se. (2023). *Väntetider i vården* | SKR. [online] Available at: <https://skr.se/vantetiderivarden/> [Accessed 30 Oct. 2024].
- Wang, P., Li, Y. and Reddy, C.K. (2019). Machine Learning for Survival Analysis. *ACM Computing Surveys*, 51(6), pp.1–36. doi:<https://doi.org/10.1145/3214306>.

STATISTISK LITTERACITET:

Hur kan en statistiker hjälpa en icke-statistiker att förstå?

Att vara statistiskt litterat innebär att kunna tolka och kritiskt utvärdera statistisk information samt att kunna resonera om och reagera på den givna informationen. Det kan handla om att förmedla sina reaktioner på ett statistiskt innehåll, men också att ställa kritiska frågor huruvida given information är rimlig. Detta är något av stor vikt med tanke på det stora informationsflöde vi lever i. Den här artikeln börjar med en presentation av hur begreppet statistisk litteracitet ofta beskrivs inom forskning om undervisning och lärande i statistik. Därefter följer några beskrivningar av vad en elev i årskurs 4–6 kan möta för statistik i uppgifter om lägesmått samt hur en del elever, och även lärarstudenter och lärare, resonerar om

dessa. Avsikten är att ge en inblick i hur en icke-statistiker kan tänkas resonera. I slutet av denna text kommer jag att ge några konkreta exempel på hur statistiker kan hjälpa icke-statistiker att förstå statistik.

Statistisk litteracitet

Att vara statistiskt litterat omfattar att ha kunskaper inom matematik, statistik, kontext, litteracitet och att ställa kritiska frågor. Men det innebär också att våga resonera och ifrågasätta. De olika kunskapsområdena behöver därför kombineras med olika affektiva aspekter, till exempel självförtroende. För att markera vikten av att våga ställa kritiska frågor, resonera och ifrågasätta talar en del forskare idag även om kritisk statistisk litteracitet. Inom detta fält lyfts kritiska

aspekter såsom att reflektera över statistiska argument med koppling till olika samhällsfrågor. Utvecklingen av begreppet statistisk litteracitet tar en utgångspunkt i att det inte räcker med en intern eller objektiv syn på data som går ut på att förstå till exempel diagram, lägesmått och spridningsmått. Man förespråkar även en extern syn på data. Detta innebär att träna sin reflektionsförmåga över hur statistisk information om olika innehåll som politik eller ekonomi rapporteras via media i TV och tidningar. Att utveckla kunskaper kopplade till externt synsätt bidrar till insikt i både hur argument skapas men också förmåga att ifrågasätta argument. Detta innebär att ständigt ha en eftertanke om vem som ligger bakom presenterade data, hur den har presenterats, men också om det finns några

Figur 1. Några svar från en enkätundersökning som besvarats av elever i årskurs 6.

Enkätfråga	Exempel på svar
Beskriv begreppet typvärde?	Elev A: Det är typ hur mycket något är, ungefär alltså.
Vilket lägesmått är mest användbart?	Elev B (som valt typvärde): Typvärde, man kan ta reda på hur många som tycker bäst om chokladglass till exempel. Elev C (som valt typvärde): Typvärde, när man röstar ser man vad de flesta har valt.
Beskriv begreppet median?	Elev D: Ett tal i mitten.
Vilket lägesmått är mest användbart?	Elev E (som valt median): Om vi har ett extremvärde använder vi median istället för medelvärde.
Vilket lägesmått är minst användbart?	Elev F (som valt median): Median används bara i skolan på mattelektionerna, men är väl användbart för en mattelärare.

specifika motiv bakom detta. Att resonera om intern respektive extern kunskap är ett sätt att närma sig statistisk litteracitet och ämnet statistik som väl överensstämmer med hur ämnet statistik definieras. Statistik är ett ämne där en verklighet beskrivs i siffror.

Något om interna kunskaper från skolans värld

Jag kommer att presentera några insikter om vanliga kunskaper och missuppfattningar av intern karaktär. Denna del bygger på resultat från två olika studier. Den ena studien är en läromedelsanalys som omfattar sju vanligt förekommande läromedel för årskurs 4–6. I denna har 1392 uppgifter om medelvärde, median och typvärde analyserats utifrån datanivå (nominal-, ordinal-, intervall- och kvotnivå) och vilka kunskaper om matematiska egenskaper de erbjuder. Den andra studien är en enkätundersökning som omfattar 130 elever i årskurs 6 från fem olika skolor. Skolorna är både kommunala skolor och friskolor från olika sociala kontexter. I denna undersökning har eleverna beskrivit de tre lägesmåten, motiverat vilket lägesmått som är lättast respektive svårast att förklara, samt vilket som är mest respektive minst användbart. Elevsvaren är analyserade utifrån visade kunskaper i matematik (centralt innehåll i matematik inkluderar statistik som en del av ämnet), hur de kopplar lägesmåten till verkligheten (kontext), ordkunskap (littera-

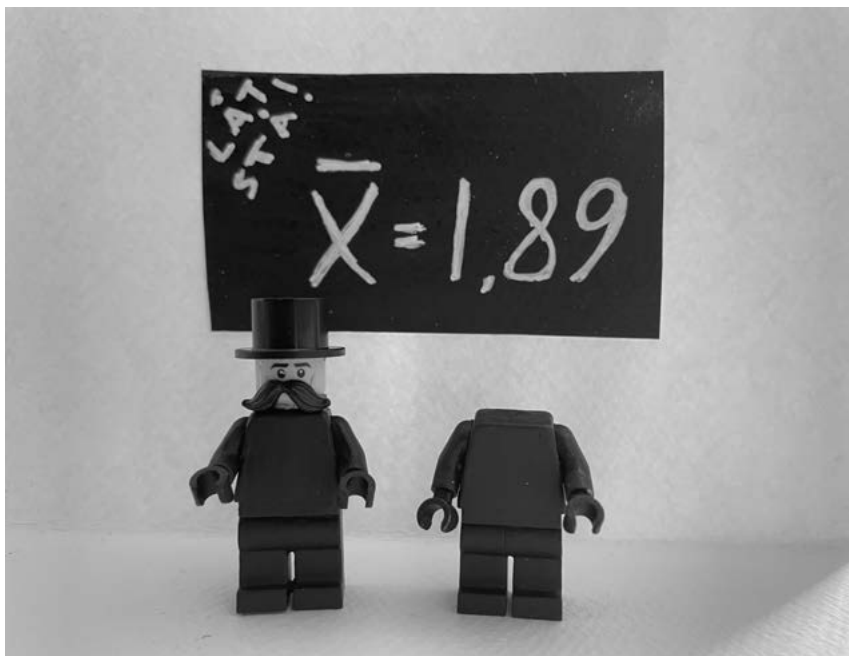
citet, här med fokus på orden/begreppen medelvärde, median och typvärde) samt affektiva beskrivningar (lätt/svårt, fattar inget, förstår inte osv.). Enkäten har också genomförts med lärarstudenter och lärare med inriktning mot årskurs 4–6. Här presenteras endast en liten del av resultaten från elevundersökningen.

Till att börja med kommer ett urval av frågor och svar från elevenkäten. Dessa har jag valt ut för att sedan koppla dessa till uppgifter i läromedel. Att man kan göra sådana kopplingar bygger på att läromedel i matematik har en så framträdande roll i undervisningen. Forskning visar dels att lärare förlitar sig i mycket hög utsträckning på läroboken. De använder exempel från boken som underlag för sin undervisning och elever arbetar större delen av lektionstiden med att lösa uppgifter i läroboken. Läroboken blir därmed källan till vilken kunskap eleverna möter och därmed får möjlighet att lära sig. Frågan är nu hur elever i årskurs 6 uppfattar lägesmåten. I Figur 1 presenteras några elevsvar från eleverna A–F gällande median och typvärde.

Resultatet visar att många elever, och även en del lärarstudenter, inte känner till begreppet typvärde. Det finns olika skäl till detta. En orsak är att uppgifter i läromedel i årskurs 4–6 frågar efter vad som är vanligast, vad flest väljer, vad som är mest populärt och så vidare utan att ordet typvärde nämns. En annan orsak är att många inte kan ange någon vardaglig koppling

till när typvärde är användbart. Ordet typvärde har ingen plats i flertalet av elevernas vokabulär. Att ca 20% av eleverna svarar att typvärdet är ungefär något (se elev A), kan tolkas som att de gör sitt yttersta att besvara en fråga i en enkät. De väljer den enda kopplingen de har till ordet ”typ” för att beskriva begreppet typvärde som de inte känner till. Trots att de har arbetat med många uppgifter och svarat på vad det finns flest av har inte frågan i uppgiften använt ordet typvärde och därmed har de inte kopplat uppgiftens innehåll till begreppet. Detta är ett exempel på varför litteracitet är en viktig kunskapsaspekt för att vara statistiskt litterat. Ett fåtal elever ger exempel på typvärde med verklighetsanknytning (kontext). Alla dessa kopplar till att rösta om något eller göra val av olika slag.

Median är ett ord med andra svårigheter. Om man inte känner igen begreppet så går det inte ens att gissa, som någon elev säger: meningen finns inte i ordet. Men trots detta visar många elever matematisk kunskap om att värdena ska rangordnas och svarar att medianen är talet i mitten. Flertalet ger exempel på ett udda antal värden, en del missar att rangordna. Alla tre grupperna, elever, lärarstudenter och lärare, ansåg dock att median var minst användbart. Det var endast några få av de vuxna som gav konkreta exempel, och då handlade det om medianlön. Det fanns också några i varje respondentgrupp, t.ex. Elev E, som menade att median var användbart utifrån den matematiska



Dr Frankenstein var statistiker på Statistiska Centralbyrån och trollade utan större ansträngning fram små svarta gubbar med mystisk innebörd.

egenskapen att medianen är ett mer representativt värde då vi har extremvärden. Denna visade kunskap återfinns i ett fåtal av uppgifterna som analyserats.

Om vi jämför resultaten gällande median och typvärde med medelvärde visar fler av respondenterna matematisk kunskap gällande medelvärdet. Detta är något som rapporterats även i internationella studier, och ett skäl som anges är att eleverna i högre utsträckning tränar begreppet i procedurella uppgifter. Även elever i min studie påtalar detta: ”Medelvärde för jag har arbetat mest med detta”. Uppgifter där syftet är att genomföra en beräkning, en procedur, en algoritm. Ett exempel är följande: Beräkna medelvärdet för följande tal 3, 9, 17, 21 och 5. Läromedelsanalysen visar att även svenska läromedel övervägande erbjuder procedurella uppgifter om lägesmått, med en påtagligt stor andel uppgifter om medelvärde. Att elever blandar ihop medelvärde och median är sannolikt en följd av detta. Svaret i uppgiften ovan, att medelvärdet är lika med 11, är endast ett numeriskt värde utan koppling till en reell kontext. I någon annan uppgift får eleverna fram ett svar att en median är 11. Detta, att i stor utsträck-

ning hantera numeriska svar utan sammanhang och utan att beröra vad lägesmålet säger om specifika data i en kontext är en orsak till denna sammanblandning.

Det finns också en koppling till litteracitet som gör att medelvärde är lättare att hantera för flertalet av eleverna. Begreppet används i vardagligt tal, och det finns ett flertal vardagliga synonymer, t.ex. snitt, genomsnitt och medelålder. Begreppet är därmed en del av elevernas språk och många kan genom ord som medelålder motivera användbarheten i en specifik kontext. Även ord som genomsnitt rymmer en uppfattning som kan kopplas till användbarhet och algoritmen. En elev uttrycker att ”då får man fram den genomsnittliga siffran”.

Avslutningsvis vill jag beröra den matematiska egenskapen att ett medelvärde ofta inte ger ett värde hämtat från den fysiska verkligheten. Till exempel att säga att medelvärdet för antalet syskon i en grupp av personer är lika med 2,19. Kognitivt är det svårt att förstå hur diskret data genom en beräkning omvandlas till ett kontinuerligt värde. Forskning visar att både elever och lärarstudierande men också i viss mån lärare har svårt att förstå och förklara vad ett sådant

överkligt värde innebär, ingen syskonskara kan ju bestå av 2,19 personer. Detta är ett exempel på när självförtroende behövs för att våga börja reda ut sina tankar. Denna svårighet berörs även av Tage Danielsson i Berättelsen om Sven-Erik Medeltal, statistikerns idealmänniska.

Dr Frankenstein var en man med övernaturliga egenskaper. Han var statistiker på Statistiska Centralbyrån och trollade utan större ansträngning fram små svarta gubbar med mystisk innebörd: ibland innehöll en sån där liten gubbe alla nykterister i Västerbottens län, ibland kunde den lilla svarta gubben föreställa alla svenskar som röstade på högerpartiet för två år sen, och bredvid stod då en annan liten gubbe och föreställde alla svenskar som röstade på högerpartiet nu senast, fast den sista lilla gubben hade inget huvu. Så ni förstår att Dr Frankenstein han var allt en riktig trollkarl! (Danielsson, 1964).

Nu är det sannolikt så att flertalet av er som läser denna artikel är statistiker, eller trollkarlar som Tage Danielsson skriver. Frågan är hur ni kan hjälpa gemene man att förstå den statistik ni sammanställer i syfte att informera vanliga människor om verkliga händelser. Hur kan ni skapa tillfälle och möjlighet för läsaren att tolka era resultat korrekt? Forskning visar att det är en utmaning att tolka data, och många tillfällen krävs för träna på att kommunicera sina resonemang för andra. Denna process involverar både matematisk kunskap, logiska resonemang och en koppling till den aktuella kontexten. Situationen kräver också kunskap att använda lämpliga ord för de begrepp som är involverade. Då finns det möjlighet att ett resultat förmedlas korrekt och kan mottas med begriplighet.

Några råd att tänka på

Avsikten med texten ovan är att beskriva vad vi kan förvänta oss att många elever får med sig från skolan. Det framgår att det är mer eller mindre interna kunskaper som snarare omfattar procedurell matematisk kunskap än konceptuell

Figur 2. Vad kan ni som statistiker bidra med?

Domän	Exempel
Kontexten Hjälpa till med tolkningar, vad är man ute efter.	Vad behöver beskrivas om kontexten? Notera om information saknas. Beröra eventuella alternativa tolkningar.
Data Beskriva den data som studien bygger på.	Beskriva urval. Jämförelser med andra uppsättningar av data.
Hur och vad har valts att beskrivas med utgångspunkt från data?	Vad den valda representationen kan förmedla. Vad den inte förmedlar. Saknas information? Vilka slutsatser kan dras av vald representation?
Hur är data vald att visualiseras?	Finns det andra sätt att redovisa data? I skolans värld brukar ett avsnitt inom statistiken handla om att luras med diagram.

Antal förfrågningar om förgiftning av lustgas



*Statistik t o m 25 september 2024

Fotnot: Statistiken avser antalet frågor till Giftinformationscentralens telefonjour om inträffade förgiftningstillbud av olika allvarlighetsgrad och inte antalet faktiska fall. Giftinformationscentralen kan få flera förfrågningar om samma tillbud/förgiftning. Källa: Giftinformationscentralen

Figur 3. Diagram från DN 2024-09-27 i en artikel om missbruk av lustgas.

statistisk kunskap om begreppen. Den procedurrella matematiska kunskapen bidrar inte till att stärka elevens kontextuella kunskaper. Likaså framgår att brister vad det gäller litteracitet inverkar på möjligheten att praktiskt kunna använda begreppen. I klartext, flertalet behöver få hjälp att fortsätta utveckla sin statistiska litteracitet ytterligare.

Vad kan ni som statistiker då bidra med, med målet att utveckla en kritisk handlingsförmåga? Målet att bli en kritisk konsument av statistiska budskap i den verkliga världen. I statistiska budskap ges en argumentation som utgår från ett visst sammanhang, vilken data man har valt att presentera, vad man har valt att tolka utifrån denna data samt hur man valt att presentera data. Era insatser kan då vara att hjälpa till med kommentarer inom fyra olika domäner vilka beskrivs ytterligare i Figur 2.

Ett mål i en värld av fake news är att med tydlighet träna allmänheten i att reflektera över om den som levererar ett budskap kan ha en dold agenda. Detta kan beröra själva innehållet. Vilka fakta har man valt att redogöra för? Vad händer om urvalet vore ett annat? Skulle man då få ett annat resultat? Skulle ett resultat kunna beskrivas med andra lägesmått? Är diagrammet ritat på ett sådant sätt att val av gradering inverkar på vad det visar? Som vi ser rör vi oss både i den interna såväl som den externa sektorn.

Ett exempel från verkligheten

Avslutningsvis ett exempel från DN 2024-09-27 (Figur 3). Artikeln har rubriken: *Larmet: Missbruket av lustgas har fördubblats*. Ett aktuellt ämne, och man blir av rubriken nyfiken på argumentet om en fördubbling. Ett viktigt innehåll och beskrivningen av kontexten har till att börja med fokus på skadeverkningar. Texten kompletteras med ett diagram.

Fördubblingen beskrivs ytterligare i texten som att den skett under ett par år, och tittar man i diagrammet kan man anta att det är mellan åren 2022 och 2024 som fördubblingen inträffat.

Samtidigt får vi en viktig inblick genom fotnoten i diagrammet. Det här urvalet stämmer nog inte med verkligheten. Påståendet fördubbling kopplas till hur många samtal som inkommit till Giftinformationscentralens jourtelefon under 2022 och 2024. Som framgår kan flera samtal ha koppling till samma person. Dessutom finns det ett mörkertal, eftersom alla tillbud inte har ringts in. Vi förstår att ordet fördubbling används felaktigt i rubriken. Det finns dock en återkoppling längre ner i texten som berör detta: ”Det går inte enbart utifrån statistiken att fastslå en ökad användning av lustgas. Men utifrån samlade rapporter från vården går det ändå att konstatera ett tydligt trendbrott, enligt Jenny Bång Arhammar.

– Det här är en form av missbruk som vi inte såg alls i princip före 2019, säger hon”.

Denna artikel följer vissa av råden i Figur 2. Framförallt ett hjälpsamt resonemang om urvalet som det valda diagrammet bygger på. Det är ett viktigt ämne och vi kan förstå att journalisten vill sprida kunskap om ett farligt ämne som missbrukas. Rubriken kanske är vald för att vi ska läsa artikeln. Man får en slags dementi av rubriken när man läser hela artikeln. Men hur många läser med dessa ögon. Kanske rubriken: ”En stor ökning av missbruket med lustgas” hade varit bättre. Och vad vet man, med annan data kanske vi kan konstatera att missbruket till och med har mer än fördubblats.

Utifrån mina korta beskrivningar från skolans värld framgår att skolan inte alltid lyckas med att ge möjlighet att utveckla statistisk litteracitet. Här ser jag en möjlighet att ni som statistiker, med stöd av lämpliga kommentarer, kan bidra till att fortsätta den utvecklingen hos era läsare.

KARIN LANDTBLOM



Referenser

- Büscher, C. (2024). Design principles for developing statistical literacy by integrating data, models, and context in a digital learning environment. I S. Podworny, D. Frischemeier, M. Dvir och D. Ben-Zvi (red), Reasoning with data models and modeling in the big data era (s. 49–60).
- Danielsson, T. (1964). Sagor för barn över 18 år. Wahlström & Widstrand.
- Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, 70(1), 1–25.
- Gal, I. (2024). What do citizens need to know about real-world statistical models and the teaching of data modeling. I S. Podworny, D. Frischemeier, M. Dvir och D. Ben-Zvi (red), Reasoning with data models and modeling in the big data era (s. 91–100).
- Landtblom, K. (2023). Mean, median, and mode in school years 4–6: A study about aspects of statistical literacy (Doctoral dissertation, Department of Teaching and Learning, Stockholm University).
- Landtblom, K. (2023). Opportunities to learn mean, median, and mode afforded by textbook tasks. *Statistics Education Research Journal*, 22(3), 6–6.
- Landtblom, K. & Sumpter, L. (accepted). Which measure of central tendency is most useful? Grade 6 students' expressed statistical literacy. *Statistics Education Research Journal*.
- Moore, D. S. (1990). Uncertainty. In L. Arthur (Red.), *On the shoulders of giants: New approaches to numeracy* (s.95–137). Mathematical Science Education Board.

Tester i rekryteringar – så fungerar de!

Tester har blivit ett allt vanligare inslag i rekryteringar till olika tjänster. Vanligast är att sökande till en tjänst får göra ett eller ett par olika tester, som sedan utgör grund för beslut om personen får komma på intervju eller inte. Olika tester ger olika typer av information och är dessutom olika bra för olika syften. I den här artikeln kommer jag därför gå igenom de olika typer av tester som finns, hur man kan avgöra om ett test är bra eller inte, samt vilken roll som testerna spelar i rekryteringssammanhang.

I grova drag finns det fem olika sorters tester. Dessa är IQ-tester, Personlighetstester, EQ-tester, Projektiva tester samt Specifika tester och arbetsprov.

IQ-tester

IQ-tester kallas också problemlösningstester och mäter en persons kognitiva intelligens. Det första IQ-testet togs fram i början av 1900-talet för att avgöra vilka skolelever som var i behov av särskilt stöd, och snart spreds testet för att bedöma vilka arbetare som var bäst att anställa. Tanken var att ju smartare man var desto bättre, en tes som dessvärre inte visat sig hålla.

Testerna mäter personers förmåga att använda logiskt tänkande. Syftet är att undersöka

g-faktorn eller ”generell intelligens”. Dock kan inte g-faktorn mätas isolerat så i stället mäts den genom andra typer av problem. Vanligast är att man undersöker personers förmåga att lösa verbala, matematiska och spatiala problem.

Verbala problem handlar ofta om att sätta ihop ordspråk eller bygga meningar av lösryckta ord. Det ställer vissa krav på befintlig språkförståelse och grammatik, vilket gör det beroende av att testpersonen kan språket.

Matematiska problem kan handla om rena matteuppgifter men även om sifferföljder och att uppskatta mängder och tider. Just förmågan att lösa matematiska problem kan tränas upp, och bra tester måste därför ta hänsyn till testpersonens akademiska bakgrund.

Spatiala problem mäter rumsuppfattning och förmåga att hantera mönster och former. Detta är den typ av problem som påverkas mest

»Ska man använda IQ-tester i rekrytering gäller det därför att inte bara leta efter personer med höga resultat, utan att se till vilka arbetsuppgifter som personen förväntas jobba med.»

av g-faktorn, och därför finns det tester som mäter IQ enbart genom spatiala uppgifter. Ett exempel är MENSAs test som enbart innehåller spatiala problem.

IQ-tester i rekrytering

Studier har visat att personer med hög IQ generellt sett tycker om att lösa problem och har ett ganska logiskt förhållningssätt till världen. Att lösa tankenötter eller andra utmaningar stimulerar personen. Däremot blir repetitiva eller monotona arbetsuppgifter snabbt tråkiga och personen riskerar då att göra uppgiften mer komplicerad eller hittar andra saker att sysselsätta sig med.

Ska man använda IQ-tester i rekrytering gäller det därför att inte bara leta efter personer med höga resultat, utan att se till vilka arbetsuppgifter som personen förväntas jobba med. Att ha ett



FOTO: SHUTTERSTOCK

IQ eller EQ? Mattias Strandberg reder ut hur IQ- och EQ-tester fungerar i rekryteringssammanhang.

område som personens resultat bör ligga inom (inte för högt och inte för lågt) är det som oftast är bäst i just rekryteringssammanhang.

En annan aspekt är hur detaljerad information man behöver ha om personens IQ. Finns det behov av att kunna avgöra personens förmåga inom de olika delområdena (verbal, matematisk och spatial), eller räcker det med att få en indikation på personens förmåga att lösa olika problem under tidspress? Ett avancerat IQ-test är därför inte alltid det bästa att använda.

Personlighetstester

Personlighetstester är den vanligaste typen av tester som används i företagssammanhang. De första personlighetstesterna kom redan för 2 500 år sedan, då de gamla grekerna började fundera på varför personer var olika. Deras observationer har levt kvar genom årtusendena och idag används personlighetstester i såväl rekrytering och ledarskapsutveckling, som grupparbeten och successionsplanering. Även mer "skämtsamma" test som "Vem är du i Star Trek?" är en typ av personlighetstest. En persons personlighet har i så gott som all forskning visat sig vara stabil över tid, och är enkelt förklarad den samlade beteenden som personen är mest bekväm med och föredrar att använda sig av.

Givetvis kan personen anpassa sig efter olika situationer, men gör då avkall på sin personlighet och betar sig på ett sätt som kan upplevas som mindre bekvämt. Oavsett vilka beteenden en person uppvisar så har människan alltid tyckt att det är intressant att jämföra olika personer, och för det behövs personlighetstester.

Den vanligaste typen av personlighetstest delar in människor i fyra olika grupper, några exempel på dessa är DiSC, IPU färganalys och Insight. Ett annat välkänt test som utgår från samma grundprincip är MBTI, som ofta används i Nordamerika. Den här typen av tester ger en grov uppfattning om hur olika individer vanligtvis betar sig. Därför lämpar de sig synnerligen dåligt i såväl rekrytering som individuell utveckling, och deras tillförlitlighet och användbarhet i psykologisk verksamhet har avfärdats av seriösa psykologer över hela världen.

I slutet av 1800-talet började forskare studera personlighetsdrag i stället för personlighet och på 40-talet kom det första fem-faktortestet. Det grundade sig i forskning som visade att människor generellt sett har fem olika personlighetsdrag som skiljer sig åt. Dessa fem är Öppenhet, Målmedvetenhet, Utåtriktning, Vänlighet och Känslomässig Instabilitet. Var och en av dessa fem kan hos en individ vara mer eller mindre

framträdande och det ger därför en mer nyanserad bild av personligheten än de tidigare nämnda fyrfälts-modellerna.

Flera tester har sedan brutit ner dessa fem faktorer i mindre beståndsdelar i syfte att få ett mer exakt redskap. Ett exempel är det mycket väl beforskade testet NEO-PI-R som bryter ner var och en av dessa fem i sex underkategorier och ger därför inte mindre än 35 olika personlighetsdrag att studera.

Personlighetstester i rekrytering

Tester som utgår från fyrfältsmodellen är alltför trubbiga för att med någon tillförlitlighet användas i rekrytering eller individuella utvecklingsprogram. De kan fylla ett syfte i vissa gruppssammanhang när man vill ha en enkel modell för att kunna prata om att människor är olika. Det finns dock en stor risk att den som gör ett sådant test identifierar sig allt för mycket med resultatet, och då blir personligheten i stället en mask att gömma sig bakom. Personen kan då ursäktas ett dåligt beteende med att "de är så", eller undvika arbetsuppgifter med ursäkten att "de inte kan det eftersom de är den här personligheten". Då har testet helt tappat sin funktion, och blir snarare en bromskloss än en möjliggörare för bättre gruppdynamik.

I rekryteringssammanhang är det därför mer relevant att använda sig av ett fem-faktortest. Ett sådant test som bryter ner de olika personlighetsdragen ytterligare ger dessutom en mycket detaljerad bild och därmed bättre möjlighet att förutse olika beteenden på arbetsplatsen.

EQ-tester

EQ-tester mäter Emotionell Intelligens, vilket enkelt kan förklaras som den buffert som finns mellan individens personlighet och det beteende som personen väljer för olika situationer. Med andra ord hur personen förhåller sig till olika situationer beroende på vad personen är mest bekväm med.

Studiet kring EQ startade på 1990-talet och tog sitt avstamp i en fundering kring vad som var avgörande för framgång. Att ha hög IQ hade inte visat sig vara avgörande och inte heller var någon särskild personlighet bättre än någon annan för att föreslå framgång i arbetslivet. Studierna visade att det som verkligen var avgörande var förmågan att kunna Förstå, Hantera och Fokusera emotionell information för att få bästa resultat i olika situationer.

Det visade sig även att EQ är något som utvecklas och förändras under en persons liv. Det innebär därför att individers EQ kan tränas upp för att personen ska bli mer bekväm eller effektiv i olika typer av sociala sammanhang. Ta exempelvis en person som tycker det är obekvämt att delta i olika sociala aktiviteter. Genom att delta och använda sig av några enkla knep och metoder kan det snart komma att kännas bekvämt, och personer kan till och med börja se fram emot att delta i olika sociala evenemang.

EQ-tester är därför utvecklade för att mäta personers förmåga att:

- **Förstå** emotionell information både från sig själv och sin omgivning.

- **Hantera** denna information på ett adekvat sätt.

- **Fokusera** sin tid och energi på det beteende som ger bäst effekt i situationen.

Det finns tre stora teorier om hur EQ faktiskt fungerar, vilket även påverkar hur olika tester är utformade:

- **Ability-modellen** ser EQ som individens förmåga att processa emotionell information för att navigera i sin sociala miljö. Här ligger fokus på att tolka signaler från andra.

- **Trait-modellen** menar att EQ handlar om vilka beteenden en individ föredrar samt vilka förmågor individen själv uppfattar att den har i sociala sammanhang. Här ligger fokus på individen själv.

- **Mix-modellen** är en kombination av de två, och ser EQ som en uppsättning förmågor och kunskaper som påverkar individens sociala och emotionella fungerande.

Vad som är genomgående för testerna är att de fokuserar på individens förmåga att prestera väl i olika sociala sammanhang. Förmågan att knyta an till andra, att hantera tankar och känslor som kan påverka relationen till andra samt

förmågan att föra fram sina åsikter på ett tydligt men respektfullt sätt är några av de områden som undersöks.

EQ i rekryterings-sammanhang

Eftersom EQ kan utvecklas och förändras över tid används ofta EQ-tester i utvecklingssammanhang. Många företag och organisationer använder även EQ i rekryteringssammanhang, och får då dels en bild av den sökandes potential för att bli framgångsrik i jobbet, dels en utgångsplan för ut-

veckling in i organisationen. EQ-tester används därför ofta i kombination med personlighets-tester, eftersom EQ och personlighet hör ihop. Det är ju individens personlighet som anpassas med hjälp av EQ, och då ges en fingervisning om hur mycket personen kan tänkas anpassa sig i olika situationer.

Projektiva tester

Projektiva tester (kallas även performancetester) mäter de saker som individen försöker dölja för andra och ibland även för sig själva. De är mycket kontroversiella, och därför mycket ovanliga i företagssammanhang. Många har hört talas om Rorschach-test där personen får se olika bläckplumpar och ska beskriva vad de ser. Teorin bakom detta är att när personer utsätts för en ny och oklar situation, utan tydliga riktlinjer på vad som är rätt eller vad som ska presteras, så tvingas de träda fram med sina verkliga jag.

Skämten om "bläckplumparnas betydelse" är många och det är lätt att avfärda dessa typer av test som bluff eller påhitt. Men faktum är att teorin bakom dessa tester går tillbaka till 1500-talet. Rorschach var däremot den första person som standardiserade bilderna och skapade ett system för hur tolkningen skulle ske.

Det finns idag mycket forskning som stöder användandet av den här typen av tester i vissa sammanhang. Tolkningsmetoderna har dessutom blivit alltmer digitala för att minska risken för att testledaren påverkar resultatet och tolkningen. Men trots nya, digitala stöd tar det ofta många år av utbildning och övning för en testledare eller psykolog att lära sig att administrera och tolka dessa tester på ett bra sätt.

Projektiva tester i rekryterings-sammanhang

Projektiva tester används främst i psykiatriska eller rättspsykiatriska sammanhang då testpersonen inte kan eller vill beskriva sig själv. De projektiva testerna blir då ett sätt att hantera denna svårighet. Genom att testpersonerna inte vet vad som förväntas så måste de helt enkelt utgå från sig själva, och det är resultatet av dessa prestationer som sedan tolkas.

Mycket av den information som de projektiva testerna ger är dock inte relevanta i de flesta affärssammanhang. I de fall där de används i rekryteringar handlar det huvudsakligen om att undersöka om en person kraftigt avviker från majoriteten av "normal fungerande" personer. Syftet är helt enkelt att hitta mytomaner och andra typer av psyko- och sociopater som annars är mycket skickliga på att manipulera självskattningstester och intervjuer.

Även när de används finns det dock en

överhängande risk för att testpersonen ogiltigt förklarar testresultatet och ifrågasätter hela rekryteringen. Det är trots allt väldigt kontroversiellt att hävda att en person inte skulle klara av sitt jobb för att den inte tyckte att en bläckplump såg ut som en fjäril.

Specifika tester och arbetsprov

Specifika tester kan vara exempelvis språktester eller kunskapstester, men även arbetsprov för exempelvis kreativa yrken eller för att testa om en person kan hantera uppgifter eller situationer som är typiska för arbetsplatsen. Dessa syftar till att mäta något specifikt och är därför ofta enkla att både administrera och rätta. De behöver dock oftast tas fram unikt för varje arbetsplats och det behöver finnas en tydlig koppling mellan testet och den specifika rollen.

Är testet bra eller dåligt?

För att avgöra om själva testet är bra eller dåligt är det två saker du behöver titta på: validitet och reliabilitet. Validitet undersöker om testet mäter vad det avser mäta. Reliabilitet undersöker om testet är pålitligt oberoende av kringfaktorer.

Fyra typer av validitet

Det finns fyra olika typer av validitet beroende på vad testet avser mäta.

- **Innehållsvaliditet** handlar om hur väl frågorna täcker av ett visst område. Ett exempel är kunskapstester som syftar till att undersöka om personen har de kunskaper som de säger. Täcker testet då hela kunskapsområdet eller bara delar?
- **Empirisk validitet** handlar om hur väl testresultatet möter ett uppsatt kriterium. Ett exempel på detta är högskoleprovet, där ett högt resultat förväntas kunna förutsäga framgång på universitet och högskola. Tyvärr har just högskoleprovet visat sig ha tämligen låg validitet, och en person med högt resultat kommer inte nödvändigtvis klara universitetsstudierna bättre än andra.
- **Begreppsvaliditet** handlar om hur väl ett test mäter ett abstrakt begrepp. Ett exempel på detta är personlighetstester som säger sig mäta drivkraft. Då måste begreppet "drivkraft" vara tydligt definierat, för annars är

»Vilket eller vilka tester som är bra att använda i en rekrytering beror i första hand på vilken information som behövs för att kunna fatta ett bra beslut om vem som ska anställas.»

det svårt att veta om testet mäter drivkraft eller inte.

- **Face validity** är den enklaste formen av validitet och handlar om hur väl testet ser ut att mäta vad det säger sig mäta. Här är det lätt att bli lurad av att testet "ser snyggt ut" och därför antas kunna ge bra resultat. Tester av låg kvalitet brukar ofta se väldigt proffsiga ut.

Två typer av reliabilitet

Det finns två huvudsakliga typer av reliabilitet, som visar på hur stabilt testet är beroende på kringfaktorer:

- **Test-retest reliabilitet** undersöker om samma person får liknande resultat om testet görs om med några veckors mellanrum. I de allra flesta fall hinner inte tillräckligt mycket ändras i en persons liv på några veckor för att det ska ge signifikanta utslag i ett test. Om resultatet därför ändras är det inte personen, utan testet det är fel på.
- **Interbedömarreliabilitet** undersöker om resultatet blir detsamma om två olika personer administrerar testet. Det kan dels handla om hur väl testpersonen förstår vad som ska göras, dels om hur resultatet tolkas. Idag avhjälpas detta ofta genom digital administration och resultatuttolkning. Men i slutändan är det ändå en person som ska ge återkoppling på testet, och då gäller det att resultatet kan tolkas lika oavsett vem som ger återkopplingen.

Mycket av den här informationen ska finnas i testets tekniska manual. Om det inte finns en teknisk manual så är testet sannolikt inte särskilt pålitligt eller användbart. Men även med en teknisk manual kan det vara svårt att veta om testet är bra eller dåligt. Då kan man vända sig till Stiftelsen för tillämpad Psykologi (STP) eller Buros Center for Testing (BCT).

STP lyder under Psykologförbundet och arbetar med att granska tester och testanvändares

certifieringar. BCT är ett oberoende forskningsinstitut inom University of Nebraska och är världens främsta center för att undersöka tester. Om ett test inte är undersökt av någon av dessa, eller har fått dålig kritik i granskningen, kan det vara värt att leta efter ett annat test, oavsett hur billigt eller enkelt det är att använda. Poängen med att använda ett test är ju att få information som är användbar i rekryteringen, inte göra testleverantören glad.

Testers roll i rekryteringar

Vilket eller vilka tester som är bra att använda i en rekrytering beror i första hand på vilken information som behövs för att kunna fatta ett bra beslut om vem som ska anställas. Testresultaten måste kompletteras med annan typ av information, exempelvis om personen har den erfarenhet eller utbildning som krävs för tjänsten. I en intervju går det sedan att få information om personens målsättning, inställning och förväntningar på arbetsplatsen, saker som är minst lika relevanta som informationen från eventuella tester.

Det gäller därför att se testerna som en del i ett pussel, där fler pusselbitar ger en mer helhetsbild. Det kan därför vara frestande att använda sig av alla testtyper för att få så mycket information som möjligt. Men att utsätta en person för 5-6 olika tester är inte rimligt, varken för personen eller den rekryterande organisationen. Efter 2-3 tester ger dessutom ytterligare test enbart marginellt med extra information, och är mest slöseri med både tid och pengar. Eftersom det aldrig går att få en helt komplett bild är det därför viktigt att fokusera på de bitar som tillsammans kan ge tillräcklig information för att fatta ett välinformerat beslut. Vilka tester som passar är därför olika beroende på olika rekryteringar, och i många sammanhang kan tester helt anses överflödiga.

Inom den akademiska världen finns ett stort behov av sakkunniggranskare ("peer reviewers") med statistisk kompetens. Detta kanske främst låter som en inom-akademisk aktivitet men faktum är att många statistiker (även utanför den direkta forskarvärlden) kan ha mycket att bidra med här. Arbetet är visserligen obetalt men ofta lärorikt!

Tjäna vetenskapens väl – bli statistisk sakkunnig

Den som någon gång stått som medförfattare till en vetenskaplig artikel har med viss sannolikhet därefter översköljts av inbjudningar till sakkunniggranskningar av andra författares verk. Nu för tiden kan sådana inbjudningar ske mer eller mindre automatiserat. Förlagens publiceringsverktyg söker av de inskickade manusens referenslistor och mailar dessa artiklars författare. Ibland är det i stället författarna själva som föreslår granskare och ibland (det traditionella sättet) är det redaktörens uppgift att söka upp, och därefter tillfråga, författarnas "jämlika granskare" (peer reviewers).

Min erfarenhet av sakkunniggranskning härrör främst från klinisk/medicinsk forskning. Detta fält skiljer sig från andra genom att forskare och doktorander ofta är parallellt verksamma inom stressiga vårdyrken. Därmed har de sällan samma möjligheter att närstudera statistiska metoder på samma sätt som "heltidsforskare" (Stanley S. Schor, Schor and Karten, 1966). Dock används ofta kvantitativa data varmed statistiska metoder tillämpas (Ercan and Demirtas, 2015). Fältet är därmed en lågt hängande frukt för den som söker förbättringspotential i inskickade manuskript (Altman, 1994; Matthew S. Thiese et al., 2015)!

Ärligt talat har jag oftast endast en rudimentär förståelse av artiklarnas egentliga/kliniska relevans och innehåll. Men då ges också desto större möjligheter att lära sig något nytt! Som statistiker fokuserar jag främst på att säkerställa (föreslå) att korrekta (bättre) statistiska metoder används (Skaik, 2015). Ibland handlar det om att upptäcka fel där författarna missat att kontrollera statistiska testers antaganden, eller övertolkat resultaten av sina egna beräkningar (Alexander Strasak et al., 2007). Ibland används å andra sidan helt korrekta och väldigt avancerade metoder. Då får jag själv fördjupa mig

på området och förstå vad just denna metod tillför utöver en enklare, mer väletablerad och redan allmänt erkänd metod. Har jag förslag på ändringar som jag tror är rimliga och konstruktiva så framför jag dessa. Å andra sidan är det sällan granskarens uppgift att helt "omdesigna" en redan genomförd studie. Då får man forska själv i stället!

En granskares lojalitet ligger i första hand hos samhället i stort. Första frågan man bör ställa sig är om framtida patienter riskerar få en sämre/farligare/dyrare behandling om ett felaktigt forskningsresultat implementeras efter en tveksam publicering. Därefter om en framtida läsare riskerar att missledas eller missförstå viktiga resultat på grund av felaktiga (eller bristfälligt beskrivna) statistiska metoder. Överväger risken för fel- och övertolkningar så rekommenderar man tidskriftens redaktör att inte publicera. Den tredje lojaliteten ligger alltså hos den initierande redaktören/tidskriften/förlaget som bjudit in till granskning (Altman, 1998). Redaktören har själv oftast varken tid eller kompetens att granska alla manus i detalj, varmed hen i stället ber exempelvis en statistiker om hjälp.

Den fjärde lojaliteten ligger hos manusförfattarna. Som granskare är man inte bara grindvakt utan kan i bästa fall också föreslå förändringar som leder till substantiella förbättringar inför publicering. Ibland sker det under en upprepad granskningsprocess med återkommande kommunikation mellan författare, redaktör och granskare. Ibland sker det via ett förslag som författarna tar till sig först efter att den specifika tidskriften avböjt publicering.

Vad får man då själv ut av det hela? Det altruistiska och ädla svaret är förstås att arbetet sker för mänsklighetens och vetenskapens bästa genom att kvalitén på publicerade artiklar höjs (Cobo et al., 2007). Den krassa verkligheten är samtidigt att stora kommersiella förlag på detta

sätt tjänar storkovan på gratis arbetskraft (Aczel, Szasz and Holcombe, 2021). En del tidskrifter publicerar både granskningsrapporterna och korrespondensen därkring (Ross-Hellauer, 2017). Men i de flesta fall sker processen anonymt och därmed utan möjlighet till formellt erkännande eller kvantifierbara meriter för granskaren.

Dock ses granskningar ofta som en bra och viktig merit, vilket bör lyftas fram i CV:n och i samband med utvecklingssamtal/nyanställning/befordran. Ibland erbjuds man kanske en present från förlaget till jul (en bok). Ofta summeras och tackas granskarna av redaktionen i slutet av året. Och för den tävlingsinriktade finns rankingslistor såsom Web of Science (www.webofscience.com, tidigare "Publons").

Tydligen har jag själv gjort cirka 100 granskningar för totalt 22 tidskrifter. Under min mest aktiva månad gjorde jag 12 stycken (pappaledig och uttråkad), även om jag på senare tid varit betydligt mer modest. "Världens toppgranskare" heter för övrigt Jonas Ranstam och är medicinsk statistiker i Lund. Redan 2016 hade han granskat 5 000 manus (Palus, 2016)! Hans bidrag till att höja den statistiska kvalitén inom medicinsk forskning omfattar även bland annat ett biträdande redaktörskap (med statistisk inriktning) för en i övrigt medicinsk tidskrift (Block, 2021).

Tyvärr upplever allt färre någon större entusiasm kring detta obetalda arbete. Förlagen får allt svårare att engagera granskare (Zupanc, 2024), processen att rekrytera folk tar därmed längre tid (Alves et al., 2024), och en del så kallade rovtidskrifter ("predatory journals") efterfrågar först granskning, men publicerar sedan ändå bristfälliga artiklar så länge författarna betalar (Asplund, 2021).

För egen del är jag numera selektiv och tackar bara ja i de fall (1) jag har tid över, (2) förlaget/tidskriften har ett gott rykte, (3) abstractet

granskare

(vilket man får del av på förhand) tycks relevant och intressant, (4) jag tror att jag kan bidra utan att behöva läsa in mig på mycket nytt, samtidigt som (5) jag ändå helst vill lära mig något nytt (en svår balansgång).

Att engagera sig som ideell sakkunniggranskare kräver egentligen ingen särskild arbetsgivare (även om det främst sker inom akademien och hos forskande företag). För den som vill vidga sina vyer, bidra till bättre (tillämpad, kvantitativ) forskning och besitter statistisk kompetens är detta därför en utmärkt möjlighet till egen fortbildning! De flesta redaktörer uppskattar en granskning av artiklars statistiska aspekter även om man som granskare i övrigt är öppen med att man inte har möjlighet att validera de mer ämnesspecifika delarna av ett manuskript.

Hur gör man då? Någon formell väg in i systemet finns inte men till en del förlag kan man höra av sig. Har man en synlig publik närvaro på nätet ökar också chansen att bli ”upptäckt” (även om detta också riskerar dra till sig en hel del mindre seriösa aktörer). Den som erbjuds bli granskare men tackar nej ges också ofta möjlighet att föreslå en annan granskare. Så berätta för de du känner att du är intresserad, varpå de i så fall kan föreslå ditt namn nästa gång! En mindre del av de mest prestigefulla tidskrifterna har också separata rutiner just för den statistiska granskningen. Några erbjuder till och med betalning, men detta är ovanligt (Hardwicke and Goodman, 2020).

När det väl är dags för en granskning är det ofta upp till dig själv vad du vill kommentera eller ge feedback på (även

om en del förlag använder standardformulär) (Pessin, 2015). Du skriver typiskt sett dina kommentarer direkt till manusförfattarna men adresserar ändå granskningen till redaktören (vilket minskar risken för en alltför konfrontativ dialog). Ofta ombeds man att rekommendera att tidskriften publicerar en artikel i befintligt skick (det händer i princip aldrig), att de kan överväga publicering efter större metodologiska förändringar, efter mindre redaktionella justeringar, eller att manuset refuseras. Rekommendationen ligger sedan till grund för redaktörens beslut (ofta har hen minst två oberoende granskare att förlita sig på). Om redaktören väljer att gå vidare med manuset så ges författarna därefter chansen att besvara kritiken/ta till sig förslagen. Ibland går detta snabbt men ibland skickas nya versioner fram och tillbaka ett par gånger. Ibland sker processen öppet eller dubbelblindat men min erfarenhet är att granskaren ofta känner till författarnas identitet, men inte tvärtom.

Den största behållningen med denna process, tycker jag, är att man ofta även får se de andra granskarnas kommentarer och därmed kan jämföra sina egna upplevelser av manuset med deras. Det är ofta väldigt intressant att se till exempel skillnaden i hur en läkare och statistiker bedömer kvalitén av olika aspekter i ett och samma manus.

ERIK BÜLOW

Referenser

- Aczel, B., Szasz, B. and Holcombe, A.O. (2021). ‘A billion-dollar donation: estimating the cost of researchers’ time spent on peer review’, *Research Integrity and Peer Review*, 6(1), p. 14. Available at: <https://doi.org/10.1186/s41073-021-00118-2>.
- Alexander Straszak *et al.* (2007). ‘Statistical errors in medical research--a review of common pitfalls’, *Swiss Medical Weekly*, 137(304), p. 44. Available at: <https://doi.org/10.4414/smw.2007.11587>.
- Altman, D.G. (1994). ‘The scandal of poor medical research’, *BMJ*, 308(6924), pp. 283–284. Available at: <https://doi.org/10.1136/bmj.308.6924.283>.
- Altman, D.G. (1998). ‘Statistical reviewing for medical journals’, *Statistics in Medicine*, 17(23), pp. 2661–2674. Available at: [https://doi.org/10.1002/\(sici\)1097-0258\(19981215\)17:23<2661::aid-sim33>3.0.co;2-b](https://doi.org/10.1002/(sici)1097-0258(19981215)17:23<2661::aid-sim33>3.0.co;2-b).
- Alves, T. *et al.* (2024). ‘Peer Review Innovations: Insights and Ideas from the Researcher to Reader 2024 Workshop’, *Science Editor*, 47(2). Available at: <https://doi.org/10.36591/SE-4702-04>.
- Asplund, K. (2021). ‘Exploaterar forskares strävan efter snabb granskning och publicering’, *Läkartidningen*, 11 August. Available at: <https://lakartidningen.se/klinik-och-vetenskap-1/artiklar-1/utbildning-och-forskning/2021/08/exploaterar-forskares-stravan-efter-snabb-granskning-och-publicering/> (Accessed: 8 May 2024).
- Block, J.A. (2021). ‘The reproducibility crisis and statistical review of clinical and translational studies’, *Osteoarthritis and Cartilage*, 29(7), pp. 937–938. Available at: <https://doi.org/10.1016/j.joca.2021.04.008>.
- Cobo, E. *et al.* (2007). ‘Statistical Reviewers Improve Reporting in Biomedical Articles: A Randomized Trial’, *PLOS ONE*, 2(3), p. e332. Available at: <https://doi.org/10.1371/journal.pone.0000332>.
- Ercan, I. and Demirtas, H. (2015). ‘Statistical Errors in Medical Publication’, *Biometrics & Biostatistics International Journal*, 2(1). Available at: <https://doi.org/10.15406/bbij.2015.02.00021>.
- Hardwicke, T.E. and Goodman, S.N. (2020). ‘How often do leading biomedical journals use statistical experts to evaluate statistical methods? The results of a survey’, *PLOS ONE*, 15(10), p. e0239598. Available at: <https://doi.org/10.1371/journal.pone.0239598>.
- Matthew S. Thiese *et al.* (2015). ‘The misuse and abuse of statistics in biomedical research’, *Biochemia Medica*, 25(1), pp. 5–11. Available at: <https://doi.org/10.11613/bm.2015.001>.
- Palus, S. (2016). ‘Peer Review Is a Thankless Task’, *Slate*, 23 September. Available at: <https://slate.com/technology/2016/09/top-peer-reviewers-are-motivated-by-their-dedication-to-science.html> (Accessed: 8 May 2024).
- Pessin, R. (2015). *Statistical Review of Manuscripts: View from the Trenches*, *Science Editor*. Available at: <https://www.csescienceeditor.org/article/statistical-review-of-manuscripts-view-from-the-trenches/> (Accessed: 8 May 2024).
- Ross-Hellauer, T. (2017). ‘What is open peer review? A systematic review’. F1000Research. Available at: <https://doi.org/10.12688/f1000research.11369.2>.
- Skaik, Y. (2015). ‘The panacea statistical toolbox of a biomedical peer reviewer’, *Pakistan Journal of Medical Sciences*, 31(4), pp. 999–1001. Available at: <https://doi.org/10.12669/pjms.314.8611>.
- Stanley S. Schor, S.S. and Karten, I. (1966). ‘Statistical evaluation of medical journal manuscripts’, *JAMA*, 195(13), pp. 1123–1128. Available at: <https://doi.org/10.1001/jama.1966.03100130097026>.
- Zupanc, G.K.H. (2024). ‘“It is becoming increasingly difficult to find reviewers”—myths and facts about peer review’, *Journal of Comparative Physiology A*, 210(1), pp. 1–5. Available at: <https://doi.org/10.1007/s00359-023-01642-w>.

Synpunkterna som framförs i denna artikel är artikelförfattarens egna, och representerar inte nödvändigtvis EU-kommissionens (Eurostats) ståndpunkter. Eventuella faktafel i texten är artikelförfattarens – för tillförlitlig information hänvisas till Eurostats webbplats.

Livet som statistiker på Eurostat – något för dig?

Hur hamnade jag på Eurostat?

Många av mina kollegor på Eurostat har en spännande, multidisciplinär, och ofta internationell karriär bakom sig (studier i ett flertal länder, diplomatiska poster i oroshärdar i tredje världen, forskning inom kärnfusion...).

I jämförelse med dem har jag haft en ganska rätlinjig (om än inte enahanda) yrkesbana, med statistikstudier i min födelsestad Stockholm fram till min doktorsexamen [1], och sedan arbete som biostatistiker på AstraZeneca (AZ) några mil söderut – för att därpå återvända till Stockholm som metodstatistiker på SCB. Redan under min tid på AZ hade jag börjat intressera mig för möjligheten att arbeta som EU-tjänsteman, inte minst på grund av vad som verkade vara väldigt goda anställningsvillkor, men snöpligt nog missat tidsfristen för att ansöka till 1998 års allmänna uttagsprov för statistiker (vi får komma ihåg att detta var innan EU-kommissionen hade någon lättill-

gänglig närvaro på nätet). Vis av den bittra erfarenheten så höll jag bättre uppsikt framöver, och år 2001 – faktiskt redan innan jag ens börjat på SCB! – så anmälde jag mig till nästa allmänna uttagsprov. Ungefär två år senare lyckades jag (som den ene av två svenskar) komma upp på den så kallade "listan" för statistiker. Det gjorde

»Eurostat är en unik arbetsplats och långt mer internationell än svenska myndigheter; i min egen korridor är kollegorna från cirka 20 olika länder.»

att jag kunde anställas av EU-kommissionen, där jag efter några månader blev rekryterad av Eurostat (EU:s statistikmyndighet) för att arbeta med kvalitets- och utvärderingsfrågor.

För att göra en lång historia kort så hamnade jag lite snett på min första tjänst, så redan ett drygt år efter min ankomst gick jag över från Eurostat till EU-kommissionens generaldirektorat för översättning [2]. Efter en sju år lång parentes där så kändes det som dags att byta till en organisation med statistik som kärnverksamhet – så jag återvände till Eurostat för att leda metodhetens forsk-

ningsgrupp. Där blev jag kvar längre än någon annanstans i mitt yrkesverksamma liv, innan jag för tre år sedan blev vice enhetschef på enheten för data- och metadatatjänster samt standarder.

Hur är det att jobba på Eurostat?

Eurostat är en unik arbetsplats och långt mer internationell än svenska myndigheter; i min egen korridor är kollegorna från cirka 20 olika länder. Även utanför arbetstid är umgänget vanligen rätt internationellt – i Luxemburg är vi inte bara många EU-anställda, vi har som invandrare även sällskap av anställda inom diverse internationella företag (inte minst ett flertal banker). Många lokalt anställda arbetspendlar dessutom från grannländerna (Tyskland, Frankrike och Belgien). Denna stora "bubbla" av utsocknes kan göra att man kan behöva jobba en del på att komma det luxemburgiska lokalsamhället in på livet.

Professionellt så har vi, i och med det europeiska statistiksystemet (ESS), en nära yrkesgemenskap med våra kollegor på de nationella statistikmyndigheterna. Det är i samarbete med dem som vi gemensamt utvecklar europeiska statistikprodukter, genom att vi lär oss av och



FOTO: SHUTTERSTOCK

lyssnar på varandra. Många andra internationella organisationer samlar också in data, men kan ha svårt att få ländernas myndigheter att skicka in rätt data i rätt tid. Till skillnad från dem så kan vi inom EU stifta lagar som medlemsländerna måste följa. Detta ger oss inom ESS helt andra möjligheter att åstadkomma offentlig statistik som är jämförbar mellan olika länder!

Givetvis har man som EU-tjänsteman en hel del administrativa uppgifter av den typ man kan se inom vilken offentlig förvaltning som helst – men jag tror inte att jag någon annanstans hade kunnat få leda eller delta i en sådan mängd av spännande och meningsfulla projekt som här på Eurostat.

Nedan berättar jag lite mer om vad jag faktiskt har arbetat med här genom åren.

Arbetsgrupper. När vi utvecklar idéer tillsammans med medlemsländerna så är det i arbetsgrupper som vi samarbetar, vanligen med alla medlemsländer representerade. Som Eurostat-kollega kan man förväntas få i uppdrag att lägga fram förslag till (eller presentera resultat för) sådana grupper. För att kunna arbeta mer intensivt med vissa uppgifter bildas ibland en undergrupp med ett mindre antal länder, som

kan mötas oftare och nå snabbare resultat. Självt har jag arbetat i sådana grupper med att bland annat:

- Dra igång samarbetet kring DIGICOM-projektet omanvändarorientering, moderniserad spridning och kommunikation [3]
- Samordna hela ESS-portföljen av innovationsprojekt, ESS Vision 2020 [4] (där jag klev in under portföljens slutskede för att hjälpa till att knyta ihop säcken).
- Samordna arbetet med att ta fram en lista över värdefulla datamängder [5] inom statistikområdet (där en undergrupp bildades strax innan Covid-19-utbrottet, och sedan arbetade kanske till och med mer intensivt än vi hade kunnat göra under normala förhållanden, eftersom vi kunde mötas så ofta på distans).

Jag har även haft förmånen att få leda delar av mötena för arbetsgruppen för standarder för att se till att kompromisslösningar som fungerar för alla medlemsländer kan nås – som sedan resulterat i bindande EU-förordningar om klassificeringar för ekonomiska aktiviteter.

Vi samarbetar inte bara med medlemsländerna, utan även med våra kollegor på EU-kommissionens olika politiska generaldirektorat.

Denna viktiga användargrupp har så klart ett intresse av att få så detaljerade data som möjligt som underlag för politiska förslag och beslut. Samtidigt har våra partners inom ESS ett intresse av praktisk mätbarhet, låga röjanderisker och uppgiftslämnarbröda, vilket typiskt leder till önskemål om en lägre detaljeringsgrad. Detta kan stundtals göra det ganska utmanande att finna kompromisslösningar.

R. Medan SAS och andra kommersiella statistikpaket förut var allenarådande, så måste alla statistikmyndigheter förhålla sig till att mycket utveckling numera sker med öppen programvara. Många nyutexaminerade studenter har rik erfarenhet av R och liknande programmeringsspråk. Även inom Eurostat ökar R-kompetensen stadigt, och ett antal R-paket har nu publicerats av Eurostat. Häromåret fick jag möjligheten att leda utvecklingen av ett R-paket för klassificeringar och korrespondenstabeller som vi sedan lade upp på CRAN, och nu håller på att vidareutveckla [6].

Som Eurostat-kollega med färdighet för R-programmering kan man numera hitta många tjänster där det är en fördel, och där man få utlopp för sitt intresse.

Metodarbeta. När jag återvände till Eurostat så drogs jag genast in i ett flertal olika initiativ. Bland annat fick jag chansen att ha en framträdande roll i ett projekt där vi, i samarbete med den berörda statistikavdelningen, utvecklade koncept [7] och metoder [8] för integrerade urvalsundersökningar.

Jag fick även chansen att delta i mindre traditionellt metodarbete när jag drogs med i gruppen av entusiaster som startade upp ESS-arbetet med så kallade ”Big Data” (Karlberg & Skaliotis [9]). Här skulle jag vilja påstå att vi låg i framkant, medan denna exotiska typ av data ännu var ett relativt okänt och kontroversiellt fenomen för de flesta offentliga statistiker. Nu-förtiden råder det en samsyn om att denna typ av nya datakällor är något som vi inom ESS ska satsa på, vilket man kan se i den planerade uppdateringen [10] av förordningen om europeisk statistik. Vi undersöker nu därför möjligheterna att göra detta inom många statistikområden, så Eurostat erbjuder allt fler möjligheter att arbeta med att utveckla europeisk statistik baserad på nya typer av data.

Främja forskning och innovation. Eurostat arbetar med att försöka stimulera forskning och innovation inom den offentliga statistiken, och som ledare för forskningsgruppen hade jag ganska många uppgifter med det syftet på mitt bord. Till exempel fick jag ansvaret för att anordna konferenserna om ny teknik och teknologi i den offentliga statistiken (NTTS). Troligen är NTTS världens största serie av konferenser för forskning inom den offentliga statistiken [11], med

större och större uppslutning för varje år [12].

Jag har även varit med om en hel del relativt originella initiativ för att stimulera forskning, innovation och statistikintresse i form av pristävlingar för innovation av olika slag. Exempelvis har Eurostat anordnat prognosmarkertävlingar för nuvärdesprognoser (”nowcasting”) [13] och tävlingar för nyttjande nya datakällor – t.ex. data på nätet [14] och transaktionsdata [15]). Inom DIGICOM-projektet bidrog jag även på ett litet hörn till den årliga europeiska statistiktävlingen för högstadi- och gymnasieelever [16] som har blivit mer framgångsrik än vi någonsin kunnat ana. Ska läsåret 2025/26 månne bli året då SCB hoppar på taget och drar igång en svensk deltävling? Det sista som dör är hoppet!

Här måste jag årligen säga att det är långt ifrån alla kollegor på Eurostat som har haft turen att arbeta med uppgifter med ett så stort allmänt inslag av forskning och innovation. Det är nog främst på metodenheten som sådana chanser erbjuds. Inom andra delar av Eurostat kan man dock också få chansen att arbeta med innovation – och då typiskt mer koncentrerat till ett specifikt område.

Data- och metadatatjänster och standarder. Min artikel om Eurostat [3] beskriver mer i detalj vad min nuvarande enhet gör – och som

man kan se där så ägnar vi det mesta av vår energi åt att stödja den dagliga statistikproduktionen, närmare bestämt att få överföringen av data från medlemsländerna, och uppläggningsen av data i vår publiceringsdatabas, att löpa friktionsfritt. Men samtidigt kan vi tjäna som ett exempel på hur utveckling och innovation kan ske inom ett

specifikt ansvarsområde. Till exempel så har jag bidragit till att formulera Eurostats nya policy för revideringar [17]. Dessutom håller vi just nu på att titta på hur något så otippat som blockkedjeteknik kan användas för offentlig statistik [18]. Här är idén att garantera reproducerbarhet i samband med revideringar av

redan publicerad offentlig statistik [19].

»Fram tills alldeles nyligen så var uttagningsproven för att bli EU-tjänsteman en fruktansvärt långdragen historia.»

Hur gör man för att komma till Eurostat?

Fram tills alldeles nyligen så var uttagningsproven för att bli EU-tjänsteman en fruktansvärt långdragen historia (precis som när jag sökte in 2001).

Nu har dock EU:s byrå för uttagningsprov (EPSO) beslutat att snabba upp uttagningsprovsförfarandet betydligt [20] med målet att det ska ta ett halvår från sista ansökningsdatum till dess att man hamnar på den så kallade ”reservlistan” på personer som kan anställas av EU-kommissionen (inklusive Eurostat). Det finns



Medalj efter 20 tjänsteår på EU-kommissionen.

flera vägar för dig att börja jobba på Eurostat:

- För att bli fast anställd på Eurostat behöver man ta ett allmänt uttagsprov för EU-tjänstemän. Alla sådana uttagsprov utanonseras på EPSO:s webbplats [21], där de även lägger ut sin planering för kommande uttagsprov [22] – inklusive eventuella uttagsprov för specialister inom statistikområdet.
- Eurostat utlyser även tjänster som tillfälligt anställd ("Temporary agent"), och här är rekryteringsförfarandet betydligt snabbare – det är bara att kolla in de tjänster som ligger ute [23] och ta chansen ifall du har de rätta kvalifikationerna.
- Om du arbetar på en svensk myndighet kan du även söka tjänst som utsänd nationell expert. Som nationell expert är du fortfarande anställd av din svenska arbetsgivare – men får ett särskilt traktamente. Även om de öppna jobben som nationell expert ligger ute på Eurostats webbplats [23] måste ansökan organiseras via din "hemmamyndighet".

Som redan nämnt så är en större uppdatering [10] av förordningen om europeisk statistik just nu på gång – och om du håller dig framme när anställningsmöjligheterna dyker upp så har du chansen att göra skillnad genom att komma hit till händelsernas centrum och vara med och ta fram nästa generation av europeisk statistik.

MARTIN KARLBERG,
Eurostat

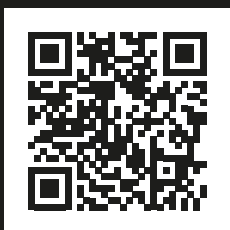
Referenser

- 1 Karlberg 1997. *Triad count estimation and transitivity testing in graphs and digraphs*. Edsbruk: Akademytryck <https://su.diva-portal.org/smash/record.jsf?pid=diva2%3A375405&dsid=3508>
- 2 M. Karlberg 2007. Livet som statistiker på EU-kommissionens generaldirektorat för översättning. *Qvartilen* 22(4) / *Qvintensen* 1(1). <https://statistikframjandet.se/wp-content/uploads/2021/02/Qvartilen-2007-4.pdf>
- 3 M. Karlberg. Europeisk statistik i samhället – med Eurostat. *Qvintensen* 2024(1). <https://statistikframjandet.se/wp-content/uploads/2024/06/Qvintensen-nr1-24-webb.pdf>
- 4 A. Ascheri, J-M Museux, A. Wirthmann, K. Giannakouris, M. Karlberg & E. Baldacci, 2022. "Innovation in the European Statistical System: Recent achievements and challenges ahead". *Statistical Journal of the IAOS* 38 (3), 805-813. <https://doi.org/10.3233/SJI-220053>
- 5 Kommissionens genomförandeförordning (EU) 2023/138 av den 21 december 2022 om fastställande av en förteckning över särskilda värdefulla dataset och arrangemangen för offentliggörande och vidareutnyttjande av dessa <https://eur-lex.europa.eu/legal-content/SV/TXT/?uri=CELEX:32023R0138>
- 6 M. Karlberg, V. Chasiotis, P. Stavropoulos, C. Laaboudi, M. Mészáros & D-A Nasiopoulou 2023. An R package for automatically generating candidate correspondence tables between classifications. *Statistical Journal of the IAOS* 39(4), 995-1009. doi: 10.3233/SJI-230039.
- 7 Karlberg M, F. Reis, C. Calizzani och F. Gras 2015. A toolbox for a modular design and pooled analysis of sample survey programmes, *Statistical Journal of the IAOS* 31(3), 447-462. <https://doi.org/10.3233/SJI-150913>
- 8 Ioannidis, E., T. Merkouris, L-C Zhang, M. Karlberg, M. Petrakos, F. Reis & P. Stavropoulos 2016. On a Modular Approach to the Design of Integrated Social Surveys. *Journal of Official Statistics*, 32(2), 259-286. <https://doi.org/10.1515/jos-2016-0013>
- 9 M. Karlberg & M. Skaliotis 2013. *Big Data for Official Statistics – Strategies and Some Initial European Applications*, WP30 in the proceedings of the UNECE Seminar on Statistical Data Collection. <https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.44/2013/mgt1/WP30.pdf>
- 10 *Förslag till Europaparlamentets och rådets förordning om ändring av förordning (EG) nr 223/2009 om europeisk statistik* <https://eur-lex.europa.eu/legal-content/SV/TXT/?uri=CELEX:52023PC0402>
- 11 M. Karlberg, S. Biffignandi, P.J.H. Daas, A. Holmberg, B. Hulliger, P. Jacques, R. Lehtonen, R.T. Münnich, N. Shlomo, R. Silberman & I. Stoop, 2015. "Preface." *Journal of Official Statistics* 31(2): 149-153. <https://doi.org/10.1515/jos-2015-0011>
- 12 M. Karlberg, S. Biffignandi, P.J.H. Daas, L. Di Consiglio, A. Holmberg, R. Lehtonen, R.T. Münnich, B. Nikic, M. Paasi, N. Shlomo, R. Silberman and I. Stoop, 2018. "Preface". *Journal of Official Statistics* 34(4): 797-809. <https://doi.org/10.2478/jos-2018-0040>
- 13 A. Zajac, M. Karlberg, & J.-M. Museux, 2024. The European Statistics Awards for Nowcasting: A New Approach to Engage with the Scientific Community and to Foster Improved Timeliness of Official Statistics. *Journal of Official Statistics*, 40(4). <https://doi.org/10.1177/0282423X241274614> <https://ec.europa.eu/eurostat/web/products-eurostat-news/-/edn-20220718-1>
- 14 Eurostat 2022. *Try your chances: Web Intelligence Competition!* (Nyhetsartikel 14 december 2022.) <https://ec.europa.eu/eurostat/web/products-eurostat-news/w/cn-20221213-1>
- 15 Eurostat: *European Big Data Hackathon 2023*. (webbsida) https://wayback.archive-it.org/12090/20231227181019/https://cross-legacy.ec.europa.eu/content/european-big-data-hackathon-2023_en.
- 16 A.F. Sanz, S. Luhmann, & A.Gálvez Moraleda 2019. "Official statistics through the eyes of students and teachers – the European Statistics Competition." *ASTA Wirtschafts- und Sozialstatistisches Archiv* 13: 245 255. <https://doi.org/10.1007/s11943-019-00249-5>
- 17 Eurostat: *Data revision policy* (webbsida) <https://ec.europa.eu/eurostat/en/web/main/data/data-revision-policy>
- 18 C. Tessitore. 2023 *The usage of blockchain technology for official statistics*. Book of abstracts of the 2023 conference on New Techniques and Technologies for Official Statistics (NTTS 2023); 449-451. https://wayback.archive-it.org/12090/20231230192209/https://cross-legacy.ec.europa.eu/sites/default/files/book_of_abstracts.pdf
- 19 Laaboudi, C., M. Karlberg and M. Islam 2024. Open data dissemination at Eurostat: State of the art. *Statistical Journal of the IAOS* 40 (1), 81-90.
- 20 EPSO: *En snabbare, smidigare och mer tillgänglig modell för uttagsproven!* (webbsida) <https://eu-careers.europa.eu/sv/introducing-faster-leaner-and-more-accessible-competition-model>
- 21 EPSO: *Dags att söka* (webbsida för pågående uttagsprov) <https://eu-careers.europa.eu/sv/job-opportunities/open-for-application>
- 22 EPSO *Upcoming selection procedures* (webbsida för kommande urvalsförfaranden) <https://eu-careers.europa.eu/en/job-opportunities/upcoming>
- 23 Eurostat: *Job opportunities* (webbsida som listar specifika utlysningar för tillfälligt anställda och nationella experter) <https://ec.europa.eu/eurostat/web/main/about-us/job-opportunities>



Har du flyttat?

Du kan själv ändra dina uppgifter genom att logga in på vår nya medlemsportal <https://stat.memlist.se/login/tb7LkmN>, användarnamn är din mailadress och lösenordet väljer du själv genom att klicka på "Glömt lösenord". Vid frågor kontakta Mattias Strandberg på sekreterare@statistikframjandet.se.



Givande kurs i läkemedels-epidemiologi

Varje år håller avdelningen för klinisk epidemiologi vid Århus universitet en sommarkurs i läkemedelsepidemiologi i det lilla danska kustsamhället Grenå på Nord-Jylland. Tack vare ett bidrag från FMS kunde jag delta i kursen tillsammans med omkring trettio andra kursdeltagare.

Läkemedelsepidemiologi, eller farmakoepidemiologi som det också heter, är ett sprudlande forskningsområde där man vill besvara frågeställningar som "hur ser användningen av ett visst läkemedel ut på befolkningsnivå?", och "medför läkemedel X en ökad risk för att utveckla sjukdom Y?". De nordiska läkemedelsregistren i kombination med andra nationella hälsoregister utgör en ovärderlig datakälla när det kommer till att besvara sådana frågeställningar. För att kunna dra korrekta slutsatser är det av yttersta vikt att data används på ett klokt sätt – annars finns risk att olika typer av bias byggs in i studiedesignen.

Under kursens gång presenterade erfarna och kunniga föreläsare ett flertal möjligheter och fallpropar inom läkemedelsepidemiologisk studiedesign och analys. Ett särskilt svårt problem när man vill studera sambandet mellan ett läkemedel och ett utfall är så kallad "confounding by indication", vilket innebär att förskrivningsindikationen för ett läkemedel har ett samband med utfallet – då blir det svårt att skilja på om det är läkemedlet eller enbart indikationen som har en effekt på utfallet. Ett sätt att komma åt detta problem är en studiedesign som kallas "active comparator, new user" där individer som startar behandling med läkemedel A jämförs med individer som startar läkemedel B som



Bildcollage från Grenå på Jylland.

har samma indikation som läkemedel A.

Ett annat ämne som togs upp i kursen var metaanalys av redan publicerade studier där varje inkluderad studie måste bedömas med avseende på risk för bias. Det är större sannolikhet att studier som visar statistiskt signifikanta resultat publiceras – denna publikationsbias riskerar att effektestimatet i en metaanalys blir starkt missvisande.

Sammantaget var kursen mycket lärorik och spännande, och inte minst var det en rolig upplevelse att få besöka Grenå och lära känna kursdeltagarna. Jag vill rikta ett stort tack till FMS som gjorde det möjligt för mig att delta i kursen.

IDA HED MYRBERG

Doktorand och biostatistiker vid Karolinska Institutet. Arbetar med registerforskning och kliniska studier inom tjock- och ändtarmscancer.

Reflektioner kring fakta och feeling

RAPPORT FRÅN STATISTIKDAGEN MED SCB

SCB arrangerade Statistikdagen på Kulturhuset stadsteatern i Stockholm 21 oktober och jag var där. Årets tema var fakta och feeling. När temat presenteras på det här sättet framträder den klassiska bilden av att fakta och känslor skulle vara två väsensskilda företeelser, kanske till och med extremer på en skala. Under dagen framkom att det är mer komplicerat än så. Jag tar fasta på Emma Stenströms, vid Handelshögskolan i Stockholm, presentation av en studie av Guilbeault m.fl. (2018).

Våra värderingar, åsikter och känslor påverkar hur vi tar till oss fakta och tolkar vår omvärld, inklusive hur vi förstår statistik. Guilbeault m.fl. visar att vi tolkar data om klimatet på olika sätt beroende på var vi står politiskt. I ett experiment fick respondenter med olika politiska inriktningar tolka ett diagram från NASA om tjockleken på den arktiska havsisen.

Hur vi tolkar data och statistik påverkas av vilka sociala sammanhang vi befinner oss i. Vad tycker de människor som finns runt oss? Det har också att göra med vår identitet. Om du identifierar dig som konservativ eller liberal (studien var amerikansk och det var dessa två politiska inriktningar som undersöktes) påverkar det hur du uppfattar världen. Du tar till dig information genom en viss lins, kan man säga, och det är fördelaktigt om den fakta du träffar på passar in i din identitet. Annars kan det bli lite obekvämt.

Vi tar alltså inte till oss fakta som om vi vore tomma kärl som fylls med kunskap. Vi är i allra högsta grad delaktiga i att tolka och förstå. Denna bild bekräftas av Ola Rosling från Gapminder. Inför Statistikdagen fick deltagarna svara på en undersökning där vi skulle gissa svaret på olika faktapåståenden. Ett av

dem var hur stor andel av all mat som produceras i världen som transporteras till ett annat land. Svartalternativen var 17, 37 eller 57 procent.

Det rätta svaret är 17 procent. Deltagarna på Statistikdagen gissade i hög utsträckning på 37 eller 57 procent, och enligt Rosling gör människor ofta det. För visst känns det som om massor av mat transporteras mellan länder? Det pratas en hel del om mattransporter och hur de påverkar klimatet. När du går in i matbutiken och ser tydliga skyltar om att vissa livsmedel är närproducerade är det lätt att få en bild av att resten av maten har transporterats väldigt långt.

Vi får en skev bild av världen, som inte beror på fake news utan på mer komplexa faktorer. Feeling spelar i allra högsta grad in när vi ska resonera kring fakta: det känns som om väldigt mycket mat transporteras mellan länder.

I ett allt mer polariserat politiskt klimat finns en stor risk att vi i allt högre utsträckning har olika världsbilder, vilket blir tydligt i det experiment som genomförts av Guilbeault m.fl. Att övertyga människor och få dem att ändra åsikt med hjälp av till exempel ett diagram kan bli svårt, för vår världsbild är mer svårubbad än så.

Finns det något vi kan göra åt problemet med bias? Något glädjande med studien av Guilbeault m.fl. är att deltagarnas tolkning av klimatdata blev mer korrekt när de fick diskutera med någon som har en annan politisk inriktning. Att samtala med någon som inte håller med tycks ha fått dem att ta in nya perspektiv och fundera ett varv till. Om du har en bild som är lite för influerad av feeling kan sådana här samtal bidra till att nyansera bilden.

Ett diagram om tjockleken på den ark-

tiska isen visar något faktamässigt, det är ingen åsikt. Trots det tolkar människor alltså diagrammet på olika sätt. Här blir det tydligt att det är förrädisk att presentera fakta utan förklaring och sammanhang och förutsätta att mottagaren kommer att ta till sig den på det sätt vi avser. Frågan för oss som statistiker är hur vi kan presentera och förklara statistik.

Men oavsett hur vi statistiker presenterar statistiken händer det mycket därefter som vi inte har kontroll över: var statistiken sprids, vem som sprider den, om någon med en viss politisk inriktning tolkar den på ett missvisande sätt och råkar ha ett stort nätverk, och hur vi pratar i samhället om den aktuella frågan. Det här är en större fråga än att vi som statistiker kan lösa den genom att kommunicera statistik tydligare.

Jag tror också att det är nyttigt att komma ihåg att du inte är immun mot att ha en bias. Vi behöver alla vara ödmjuka inför att vi tar till oss fakta inte som tomma kärl som fylls med kunskap, utan som tänkande och lärande varelser som aktivt konstruerar vår världsbild.

CHARLOTTE OVESSON

Referenser

- Guilbeault, D., Becker, J. & Centola, D. (2018). Social learning and partisan bias in the interpretation of climate trends. *PNAS* 115(39): s. 9714-9719. doi: 10.1073/pnas.1722664115



ORDFÖRANDE HAR ORDET

Fortsatt nyfikenhet är guld värd!

Oj, oj vilket år! Vi hittade vår nya redaktör, stort tack och välkommen önskar jag dig Charlotte Ovesson!

Och vilket höstmöte vi har haft! Årets tema handlade om statistikens framtid. Det var både inriktat på statistiska metoder som förändras med nya kraftiga verktyg men även vilka utmaningar och i vissa fall faror som kan vänta oss runt hörnet om vi inte är uppmärksamma.

Vi fick en grundlig genomgång av professor emeritus Yoav Benjamini från Tel Avivs universitet (och 2024 års vinnare av Rousseeuw-priset) om statistisk inferens i tillämpad statistik. Det framkom att även dr. Benjamini har haft sina utmaningar att förklara hur statistiska resultat kan och bör användas för beslutsfattare.

Genom dr. Olle Häggström vid Chalmers universitet, professor emeritus David Donoho vid Stanford universitet och Anders Bjurström vid företaget Tentify fick vi flera olika perspektiv kring våra möjligheter och ganska kraftiga utmaningar rörande AI.

Titta gärna på föreläsningarna som finns på vår hemsida att njuta av!

Under året har jag uppmärksamats på att statistik användares bruk (och missbruk) av resultat bör lyftas i diskussioner och att mer arbete i pedagogik och förklaringar av statistiken behövs. Den så kallade statistical literacy "statistisk läskunnighet" aktualiseras mer och mer. När jag började på SCB under tidigt 2000-tal fick jag veta att när statistiken publicerades hade vi möjlighet att skriva kort om resultaten, vilket var ett framsteg enligt alla jag jobbade med. Bara några år tidigare hade SCB haft en policy att

bara lägga ut tabeller med rubriker. SCB skulle inte förklara resultaten eftersom det kunde anses vara vinklat.

Men att bara lägga ut en tabell utan förklaringar leder till att bara de invigda kan använda resultaten korrekt. Och de som tror sig veta vad innehållet är missförstår ofta resultaten och kan ta helt fel beslut. Detta är något jag har fått erfara bara häromveckan. Jag själv beställde en färdig tabell från SCB och jag visste precis vad jag ville ha. Och fick det snabbt och snyggt. Men resultaten i tabellen förbryllade mig. Jag fick inte summor att stämma med andra resultat jag hade. Och tänka sig! Efter ett kort samtal med SCB hade jag svaret. Det gick ju inte att summera över grupperna eftersom jag då dubbelräknade (i mitt fall) företag! Så även om jag själv lade in specifikationen och visste vad jag ville ha missförstod jag tabellen i slutändan.

I fallet som dr. Benjamini tog upp handlade diskussionerna om hur man kan och bör använda konfidensintervall. Om användaren bara tittar på ena sidan intervallet för att motivera sina beslut kan stora medel användas i onödan på insatser som inte ger den effekt som förväntas. Det ger större konsekvenser på samhället än att jag med mitt enklare fall inte kunde se skogen för alla träd.

Vikten av en god kommunikation mellan beställare och leverantör är stor! Både vad gäller metadata över levererat material, där ingående parametrar, avgränsningar och eventuella antaganden beskrivs. Men även hur resultaten ska tolkas kan behöva diskuteras gemensamt.

Inom styrelsen har vi nu släppt vår programförklaring. Den ska hjälpa oss att hålla en ambitiös riktning för föreningens

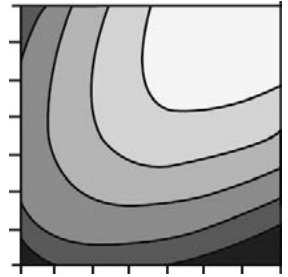
arbete. Denna förklaring har redan under förra året gett oss skjuts framåt där vi bland annat har påbörjat ett kunskaps- och erfarenhetsutbyte med andra länders statistikfrämjande och kommer att fortsätta träffa dem under detta år. Du hittar programförklaringen på vår hemsida!

Statistikfrämjandets ekonomi är tyvärr svag och vi går mot förlust i år igen. Vi kommer därför att söka nya vägar för att öka våra inkomster. Dessa inkomster gör att vi kan erbjuda ett fint och givande utbud av seminarier och aktiviteter. Bland annat tänkte vi erbjuda annonsplats i Qvintensen för att stärka inkomsterna. En annan mer repressiv åtgärd är att vi diskuterar att ta ut en avgift framöver för webinarier och konferenser. Det kan vara en anmälningsavgift eller en avgift för medlemmar som har anmält sig till våra event och inte meddelar i god tid innan om man inte kan delta. Under året har vi sett att drygt hälften av alla anmälda medlemmar till våra konferenser och webinarier inte dyker upp.

Någoting mycket roligare som vi gör tillsammans med de övriga statistikföreningarna är att diskutera om och hur vi kan modernisera oss. Denna diskussion syftar till att undersöka om vi kan få starkare synergier mellan våra föreningar och ge medlemmarna ett större utbud av våra tjänster. Vi har ett långsiktigt perspektiv i frågan så tålmod blir en dygd här!

Tack för ert engagemang i statistiken och vi hörs!

NANCY STEINBACH
Ordförande i Svenska
statistikfrämjandet



ORDFÖRANDE HAR ORDET

En bredare syn på industristatistik

När jag frågade mina kollegor i Indstats styrelse om tips för denna spalt fick jag följande återmatning från Ingemar Sjöström: ”Jag önskar ju alltid att vi har en bredare syn på industristatistik. Det finns flera viktiga näringar som vi inte kommer i kontakt med, typ skogen, lantbruket, fisket, transport, m.m. som säkert har intressanta frågor om mätningar och analys.” Kanske finns en del av svaret i ett äldre nummer av Qvintensen?

I Qvintensen 2019/1 publicerade Peter Guttorp och Georg Lindgren artikeln ”Är statistikforskningen industristyrd eller är det tvärtom?”. Författarna ”identifierade fyra industrigrenar: försäkring, skog och skogsindustri, teleindustri samt bil och annan transportindustri där specifika industriella problem har lett till ny och ibland oväntad statistisk forskning”.

Undertecknad publicerade i FENStatS nyhetsbrev 2024/2 artikeln ”The Importance of Statistical Literacy”. Artikeln hade en annan vinkling, nämligen att försöka påvisa att statistisk kompetens hos företrädare för olika discipliner kan leda till fruktbara resultat, men hade vissa be-

röringspunkter med Guttorp & Lindgrens artikel. Bland annat omnämns i bägge artiklarna följande personer:

- Walloddi Weibull (1887-1979), professor på KTH. Definierade i anslutning till forskning om metallers utmattningsfunktionerna.
- Agner Kraruo Erlang (1878-1929), dansk ingenjör och pionjär vad gällde mätningar och analys av belastningar på telenät.
- Conny Palm (1907-1951), svensk teleingenjör anställd på LM Ericsson. Hans doktorsavhandling räknas som ett pionjärbete inom teletrafikteori och köteori.

Några exempel förutom de som omnämns av Guttorp & Lindgren:

En intressant applikation inom gruvindustrin är kriging, en statistisk metod för att hitta mineralfyndigheter på ett optimalt sätt, uppkallad efter den sydafrikanske gruvingenjören Danie Gerhardus Krige (1919-2013).

En rätt stor statistisk tillämpning, i alla fall om man skall döma efter antal publikationer man hittar med Google Scholar, tycks vara studiet av partikel-

storleksfördelning. Ett intressant namn i detta sammanhang är Friedrich Kottler (1886-1965). Kottler var professor vid universitetet i Wien och bidrog till relativitetsteorin. På grund av sitt judiska påbrå tvingades han emigrera och arbetade på Kodak där han bland annat studerade fördelning av partikelstorlek.

Arvid Palmgren (1890-1971) studerade livslängden för kullager på SKF. Hans arbeten nyttjades av Stephen H. Crandall (1920-2013) som anses vara grundare till teorin för stokastiska vibrationer.

Guttorp & Lindgren utelämnar, fullt medvetet, ”den stora läkemedelsindustrin vilken förtjänar en egen behandling.” Detta område ligger också närmare en av våra systerföreningars, FMS, domäner.

Denna korta lista, liksom även Guttorps & Lindgrens artikel, indikerar att det borde finnas en uppsjö av ämnen inom industriell statistik som är värda att studera. Så svaret är ja, det borde vara av intresse att vidga synen på ”industristatistik”.

HANS ALBERG
Ordförande Industriell statistik



ORDFÖRANDEN HAR ORDET

Snabb utveckling av biostatistiska metoder

Medicinsk statistik omfattar många utmanande problem. Stora områden överlappar med eller liknar de frågeställningar som man har inom andra områden där statistik används, inom vetenskap, industri, förvaltning och andra samhällsområden. En klassisk skiljelinje inom medicinsk statistik och biostatistik går mellan randomiserade experiment och observationella studier. Jag tycker mig dock ana en trend att dessa områden mer och mer sammanstrålar och lånar idéer från varandra.

FMS uppsatspris går till Stina Zetterström för artikeln ”Bounds for selection bias using outcome probabilities”. Artikeln, som är fritt tillgänglig, är ett teoretiskt bidrag av primärt intresse inom observationella studier, t.ex. baserade på registerdata, men liknande problem dyker upp också i randomiserade studier, t.ex. för kausala estimander när vi har bortfall av data.

Vartannat år har FMS ett gemensamt möte med vår systerförening Dansk Selskab for Biofarmaceutisk Statistik (DSBS). Denna förening täcker bara läkemedelsstatistik och är därmed betydligt smalare än FMS. Å andra sidan blomstrar läkemedels-/biotech-industrin i Danmark. De har en handfull tämligen stora företag. Stjärnan är Novo Nordisk som är det högst värderade företaget i Europa tack vare deras läkemedel för obesitas och diabetes. I år hade vi ett utsålt möte i Köpenhamn som handlade dels om ny innovativ metodologi, dels om hur det är att arbeta som statistiker i läkemedelsindustrin.

Ett par av presentationerna på DSBS/FMS handlade om att använda historiska data i analysen av en randomiserad klinisk prövning (randomised clinical trial, RCT). Att jämföra läkemedelsbehandlade patienter med historiska kontroller kan ge bias, på samma sätt som comfounders kan störa registerstudier. För sällsynta sjukdomar kan mer information möjligen uppväga en bias-risk. Åtskilliga publikationer föreslår därför att observerade data för historiska kontroller och RCT-kontroller ska ”berätta” hur lika väntevärdena är, och därmed hur mycket vi kan lita på historiska kontroller. Men det ger ett mycket skakigt estimat som kan ge slumpmässiga beslut. Det är kanske bättre med en subjektiv Bayesiansk expertbedömning av skillnader mellan datakällor än att förlita sig på en förment objektiv Bayesiansk dynamisk procedur med låg transparens.

En rätt vanlig variant är att använda maskininlärning på historiska data för att hitta en god prediktor av effekt som funktion av alla tillgängliga baselinekovariater. Prediktorn blir typiskt icke-linjär och icke-transparent. Men vi kan plugga in denna pre-specifierade prediktor som en vanlig kovariat i en linjär ANCOVA och allt fungerar som vanligt. Om prediktorn kan förklara några procent av variationen, kan studien göras några procent mindre, och vi kan skära kostnader med miljoner. Detta är en superspännande idé, men det återstår att se om potentialen verkligen kan förverkligas i praktiska studier.

Ett annat tema var grupp-sekventiella design, en variant av sekventiell analys anpassad till RCT med väletablerade metoder för kontroll av Typ

1-felet för en hypotes. Men myndigheter kräver stark kontroll av family-wise error rate (FWER), risken att någon av alla nollhypoteser för olika primära och sekundära variabler (/doser/subpopulationer) felaktigt förkastas.

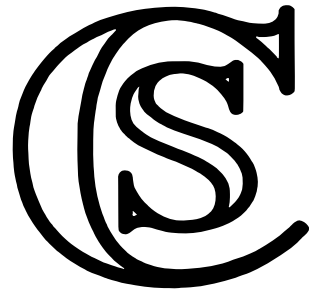
Sture Holm, i en av de mest citerade svenska statistikartiklarna genom tiderna (Scandinavian Journal of Statistics, 1975), banade väg för mycket forskning och vi har idag goda metoder för multipelinferens. Problemet är att kombinera multipelinferens med grupp-sekventiella design. En övning för den intresserade är att visa att man inte har FWER-kontroll om sekundär variabel testas på 5% närhelst den primära variabel är statistisk signifikant (vid interim eller final analys). Standardmetoder för FWER-kontroll är anti-konservativa. Kanske kan randomiseringstest hjälpa?

Det sker en snabb utveckling av biostatistiska metoder och jag tror att det finns stort utrymme för samarbete mellan universitetsforskare och tillämpade statistiker, liksom mellan olika områden inom och utom det medicinska tillämpningsområdet.

CARL-FREDRIK ”CAFFE” BURMAN
Ordförande FMS

Referenser

Zetterström, S. (2024). Bounds for selection bias using outcome probabilities. *Epidemiological Methods* 2024 13(1): 20230033. doi.org/10.1515/em-2023-0033



ORDFÖRANDEN HAR ORDET

Vad gör egentligen en modern statistiker?

Cramér'sällskapetets årliga höstmöte hölls under två halvdagar 23–24 oktober i ett höstsolsnigstrande Lund, närmare bestämt på Ekonomihögskolan vid Lunds universitet, med Statistiska institutionen som värd. Årets möte kretsade kring de senaste årens snabba utveckling inom AI och maskininläring och de utmaningar som denna innebär för våra utbildningsprogram i statistik och matematisk statistik. Temat för mötet, ”vad gör egentligen en modern statistiker?”, framstår som mer relevant än någonsin när våra studenters yrkesvägar blir allt mindre självklara och när nya roller, som dataanalytiker, växer fram.

För att fånga den moderna statistikers roll ur olika perspektiv bjöd programmet på en spännande mix av inbjudna talare från olika sektorer där statistiker är verksamma, däribland medicin- och läkemedelsindustrin, bank- och finanssektorn samt försäkringsbranschen. En representant från värdinstitutionen gav dessutom en inblick i hur dess utbildningsprogram anpassats för att möta de nya kraven. Mötet avslutades med ett uppskattat panelsamtal där några av talarna, tillsammans med representanter från akademien, diskuterade den moderna statistikers roll och dagens kompetenskrav.

Diskussionerna under mötet visade på en bred enighet kring att statistiker är fortsatt eftertraktade, och att denna efterfrågan förväntas bestå framöver. Det finns ett utbrett behov av personer med mer djupgående förståelse för statistisk inferens och modellering. Det konstaterades också att även om klassisk statistisk inferens förblir betydelsefull, särskilt inom läkemedelsindustrin, har fokus inom många andra sektorer förflyttats från validering till prediktion och klassificering, i och med det senaste decenniets stora framgångar inom djupinläring och maskininläring. Utgångspunkten för en statistiker förblir dock densamma: att utvärdera osäkerhet.

Samtidigt ställdes frågan om ”statistiker” är den mest relevanta benämningen på dagens yrkesroll. Kanske bör den ersättas med en som bättre speglar den teoretiska och metodologiska expertis som en modern statistiker förväntas ha?

Mötet gav även värdefulla insikter i vilka kompetenser som efterfrågas i dagens arbetsliv. I en tid då dataanalytiker är eftertraktade på arbetsmarknaden är det givetvis viktigt att en statistiker har en förståelse för de moderna metoder som efterfrågas. Förutom en gedigen grund inom både traditionella och moderna statistiska metoder framhölls

därför vikten av programmeringskunskaper och färdigheter inom optimering. Samtidigt betonades behovet av att inom kurser i statistisk maskininläring främja en teoretisk förståelse för dessa metoder. Bland de egenskaper som en framgångsrik modern statistiker bör ha framhölls återkommande kvalitativt tänkande och förmåga till riskavvägning, ett prediktivt fokus samt förmågan att kombinera teoretiska insikter med praktisk erfarenhet och implementeringsvana.

Sammanfattningsvis blev höstmötet en engagerande och värdefull plattform för diskussioner om statistikers roll idag och imorgon och om hur statistikutbildningen behöver utvecklas för att rusta våra studenter för ett arbetsliv där dataanalys och statistiska metoder blir allt mer centrala.

Då det finns en hel del kvar att avhandla får vi säkert anledning att ta upp detta tema i någon form på fler möten. Ett stort tack till alla medverkande för ert engagemang och för de insikter som ni delade med er av under dessa dagar!

JIMMY OLSSON
Ordförande Cramér'sällskapet



ORDFÖRANDEN HAR ORDET

Tillbakablickar och framtidspanande på föreningens jubileumskonferens

Surveyföreningen har den här hösten firat sitt 25-årsjubileum med en jubileumskonferens. När jag blickar tillbaka vill jag särskilt uppmärksamma tre händelser i föreningens historia.

Beslut om sektion och om förening. Vid årsmötet den 22 september 1999 i dåvarande Svenska statistikersamfundet fattades beslutet att bilda en sektion inom Samfundet med namnet Sektionen för Surveystatistik.

Hösten 2007 hade Surveysektionen ett årsmöte där namnet ändrades till det nuvarande Föreningen för surveystatistik (Surveyföreningen). Föreningen blev formellt självständig och kunde bli en egen juridisk person. När Statistiska föreningen och Svenska statistikersamfundet 2008 gick samman fortsatte Surveyföreningen som en självständig sektion inom Statistikfrämjandet.

Det heraldiska märket. Initiativ till Surveyföreningens heraldiska märke (logotyp) togs 2005. Det innehåller föreningens valspråk, Statistikersamfundets symbol och en bild för Surveyföreningen som ska associera till urval. Två gröna papegojor är sköldhållare. Papegojor brukar tjänstgöra som symboler för tålmod.

Utredande kommittéer. Survey-

föreningen har genom åren gjort flera stora insatser inom surveyområdet. Särskilt tre frågor har utretts, och de är fortfarande aktuella: bortfall, utbildning och webbpaneler.

Löpande aktiviteter. I styrelsens löpande arbete ingår ett årligt pris, en årlig utmärkelse, konferenser och seminarier samt en statistisk ordlista med inriktning på surveyområdet.

Det årliga priset delas sedan 2006 ut för bästa akademiska uppsats med surveyinriktning inom grundutbildningen.

Tore Dalenius (1917–2002) var en av portalfigurerna inom svensk surveyvetenskap. Kriterier för utmärkelsen Daleniustalare infördes formellt 2023. Det första talet hölls 2001.

Jubileumskonferensen. Onsdagen den 16 oktober var det jubileumskonferens i Stockholm med 50 deltagare. Föredragens titlar visas nedan för att illustrera surveyområdets bredd och antyda dess föränderlighet.

- Från sjunkande svarsfrekvenser till datadrivna insikter
- Is there a future for statistical surveys?
- Människorna bakom kryssen – då, nu och framåt
- Valforskningsprogrammet 70 år: framtidsutsikter för världens näst äldsta nationella väljarstudie

- Trender på surveyområdet
 - Framtidens statistikproduktion: utmaningar och möjligheter
- Föredragshållarna hade i uppdrag att peka på (i) det största/viktigaste under de senaste 25 åren (ii) det största/viktigaste under nästa 25 år samt (iii) faror, utmaningar och möjligheter för framtiden.

Det fanns likheter i bedömningarna, men spännvidden i infallsvinklar och exempel var stor. Den snabba teknikutvecklingen är ständigt betydelsefull, liksom omsorg om respondenter och etik. Sannolikhetsurval kvarstår men har blivit något annorlunda.

Framtiden – tre aspekter. Den första aspekten är den ämnesmässiga framtiden. Jubileumskonferensen var inspirerande.

Den andra aspekten är Surveyföreningens verksamhet i framtiden. Styrelsen räknar med att fortsätta med de löpande aktiviteterna. Dessutom vill vi synliggöra föreningen mer och få flera medlemmar.

Den tredje aspekten hänger ihop med Statistikfrämjandets initiativ att modernisera föreningen, tillsammans med de fyra sektionerna.

Framtiden är uppenbarligen spännande.

EVA ELVERS

Ordförande Surveyföreningen

Två nya utgåvor av FENStatS nyhetsbrev

Som nämndes i Qvintensen 2024/1 har FENStatS för avsikt att ge ut ett nyhetsbrev en gång per kvartal, med deadline för inläggen den 15/3, 15/6, 15/9 och 15/12 och distribution en vecka senare. Efter förra Qvintensens publicering har två nummer kommit ut planenligt.

Nyhetsbrevet har behållit sin ursprungliga struktur:

- Hälsning från ordförande Lola Ugarte
- FENStatS Announcements
- News from National Statistical Societies
- Conferences/Workshops
- Papers

I juni-numret kan man bland annat läsa att medlemmar i FENStatS (vilket torde innebära alla medlemmar i någon av FENStatS medlemsföreningar, exempelvis Svenska Statistikerfrämjandet) kan prenumerera på nätversionen av tidskriften *Significance* för £20 per år.

Under rubriken "News from National Statistical Societies" hittar man nyheter från finska, slovenska, brittiska, rumänska och spanska statistikersällskapen.

Undertecknad bidrog med ett "paper", "The importance of Statistical Literacy". Syftet med artikeln var att påvisa att forskare inom diverse olika discipliner har nytta av statistikkompeters. Detta inte bara för att tillämpa befintliga metoder utan även för att utveckla nya idéer. Man kan exempelvis betänka det enorma genomslag Hans Roslings diagram fått.

I **september-numret** kan man läsa att FENStatS General Assembly godkände FENStatS etiska kod.

De cypriotiska, slovenska, rumänska, spanska och belgiska statistikersällskapen bidrog med nyheter. Bland annat kan man notera att professor Peter Rousseeuw tilldelas ett hedersmedlemskap i det belgiska statistikersällskapet. Peter Rousseeuw är en belgisk statistiker, känd bland annat för att ha instiftat ett statistikpris liknande Nobelpriset, i alla fall med ungefär samma prissumma. Han har verkat vid flera framstående universitet, bland annat TU Delft och ETH, för att avsluta sin akademiska karriär som emeritus vid KU Leuven.

Slovenien stod 2024 som värd för European Statistics Competition där lag med gymnasieelever från 20 europeiska länder deltog. Sverige deltog inte, trots att vi har en del erfarenheter av vetenskapstävlingar för ungdomar sedan gammalt. Lag från bland annat Finland, Litauen, Polen, Tyskland och Island deltog, för att nämna några grannländer. Kanske borde Svenska Statistikerfrämjandet pusha för ett svenskt deltagande?

Undertecknad bidrog även till detta nummer av FENStatS nyhetsbrev med en artikel som triggades av upptäckten att de 25 högst rankade universiteterna enligt Shanghai ranking låg 19 i USA och bara ett i EU (Paris-Saclay University på plats 23). På andra plats inom EU kom ovan nämnda KU Leuven (plats 34). Man kan naturligtvis diskutera hur allvarligt man skall ta på

dylika rankingar, men det verkar som om många framstående universiteter tar rätt seriöst på det (se exempelvis på pressrelease från KI). För något år sedan gratulerade jag dåvarande rektorn för ett stort centraleuropeiskt universitet för att hans universitet klivit upp några pinnhål. Han tackade för detta men framhöll att han gärna sett att det hamnat bland de 100 främsta. Även om ranking inte är en exakt vetenskap bör man kanske ändå se det som en indikation på att EU ligger efter?

HANS ALBERG

Referenser

- Alberg, H. The Importance of Statistical Literacy. FENStatS Newsletter 2024/2 <https://drive.google.com/file/d/1cjLLkB6bzJF5BIcp7jdO80GoW2QyWzFV/view>
- Alberg, H. Can FENStatS do something about the status of Statistical Research in Europe? FENStatS Newsletter 2024/3 https://drive.google.com/file/d/1ku5iR_p1UOwFqyX-CK8Zm_MVCxILLcRI/view
- European Statistics Competition <https://www.esc2024.eu/>
- FENStatS Ethical Code Professional Ethics for Statistical Practice in Europe https://www.fenstats.eu/files/stranky/about_us/files/FENStatS_2024_Code_of_Ethics.pdf
- KI om ranking: <https://ki.se/om-ki/fakta-om-ki/ranking-och-karolinska-institutet>

Minnesord över Lennart Bondesson

Han skulle bara ta sig ett dopp och plötsligt så var han inte mer, nyss fyllda 80 år.

Lennart Bondesson tog studenten 1963, efter att bland annat ha placerat sig mycket högt i Svenska Dagbladets matematiktävling. Därefter studerade han framför allt matematiska ämnen i Lund och avlade licentiatexamen 1971. Läsåret 1972-73 var han stipendiat vid Institut Mittag-Leffler, där han delade rum med undertecknad.

I den magnifika villan i Djursholm finns statyer av Mittag-Leffler i alla rum och ett porträtt av modern utförd av Carl Larsson i nedre hallen. Där ägnade sig Lennart åt karaktäriseringsproblem i förlängningen av det klassiska att om \bar{x} och s^2 är oberoende så är fördelningen normal. Hans resultat var övervägande av typen ”om stickprovsstorleken är minst n och så och så och blablabla gäller så är fördelningen en viss given”. Han hamnade därvid i ett slags tävling med Abram Kagan i Leningrad (som det hette då) där antagandet om stickprovsstorleken blev allt mindre. ”Nu har jag kommit ner till minst 7” minns jag. Lennarts metod, som byggde på komplex analys, hade ett golv: den krävde minst $n=6$ för att fungera.

Från Mittag-Leffler-tiden minns jag speciellt två saker. Den ena var när Lennart, vid 17-tiden en dag, sa att han skulle gå ner till sin bostad som han hade på området och laga till en oxsvanssoppa. Dessvärre insåg han inte förrän han var igång att den tar sin tid, varför middagen avåts framåt midnatt. En annan gång gav han mig ett brev om några få rader i vilket han bad en Uppsalaprofessor om ett särtryck. Han undrade om ”man kan skriva så här” med tillägget ”du får bara säga att det är bra”.

Lennart disputerade i matematisk statistik i Lund 1974 på avhandlingen *Characterizations of Location and Scale Parameter Families of Distributions th-*

rough Optimality Properties of Certain Estimators. Efter en period på bland annat SCB i Stockholm tillträdde han 1983 professuren i skoglig matematisk statistik vid Skogshögskolan i Umeå. Våren 1994 ringde han mig – vi talades vid av och till genom åren. Han undrade plötsligt, kanske halvt på skämt, om jag hade något jobb åt honom. Just då var vi på väg att lysa ut ett lektorat, vilket jag uppmuntrade honom att söka. Det fick han förstås.

Från den tiden vill jag främst framhålla att han införde tentamina i de mer tillämpade delarna av statistiken vid datortorn. Studenterna fick datamaterial att analysera, som om det vore ”på riktigt”. Annorlunda var det på ”min tid” på 1960-talet då man kunde få ett datamaterial om 10 observationer tillsammans med $\sum x$, $\sum y$, $\sum x^2$, $\sum y^2$ och $\sum xy$ och ombads räkna ut korrelationskoefficienten.

Efter fem år i Uppsala återvände Lennart till Umeå som professor, då vid universitetet. Han hade även kunnat få en professur vid Lantbruksuniversitetet, SLU, men valde efter stora vändor Umeå. Där träffade han, den borne ungarlen (vad man vet, han var ju en smula hemlighetsfull), sedermera Britt-Marie som så småningom blev hans särbo.

Ett område Lennart ägnade sig åt under åtskilliga år var oändlig delbarhet, exempelvis av potenser av gammafördelade stokastiska variabler och av heltalsdelen av oändligt delbara stokastiska variabler, men även av simulering från oändligt delbara fördelningar. Han återvände även till karaktäriseringsproblemen där han bland annat studerade klasser av generaliserade gammafördelningar, klasser av blandade gammafördelningar och införde klassen hyperboliskt monotona tätheter (HCM). En höjdpunkt nåddes i en Springer Lecture Notes in Statistics 1992. Han publicerade även rejält inom

statistik, exempelvis sampling, samt lite allt möjligt annat.

Lennart hade en ovanligt stor beläsenhet inom sannolikhets teori och statistik med en stor bredd i sin publikationslista. Sålunda ägnade han sig även åt statistiska problem, speciellt sampling. Det var ett område han var väl bevandrad i, inte minst sedan sin tid vid SCB.

Lennart var, som många av oss minns, inte som alla andra, vilket för mig är en komplimang i den meningen att han hade styrkan att vara sig själv. Han var ingen konformist. Kepsen minns vi förstås. Han var en godmodig person, samtidigt som han kunde roas av andras motgångar. Från vårt gemensamma år minns jag när min första publicerade artikel hade blivit plagierad. Med ett leende på läpparna och ett kluckande skratt kunde han under många år fråga mig ”har du hört nåt från Basu än?” samtidigt som han höjde garden en smula och tog ett halvt steg bakåt.

Om Lennarts fritidssysselsättningar, om livet utanför den matematiska världen, må andra vittna. Jag minns dock att han var en ivrig och synnerligen kunnig svampplockare. Ett annat intresse var cykling, och då rejält. Ett speciellt minne är en konferens någonstans söderut då han, efter några dagars trampande, uppenbarar sig i en t-shirt och svarta tajta cykelshorts. Han var även en entusiastisk vinterbadare och inledde gärna sina dagar med att vandra ner till Umeälven och ta sig ett dopp.

Lennart var en alldeles speciell färgklick i vår krets som nu har fått en gråare nyans.

ALLAN GUT

Minnesord över Staffan Sollander

OCH HANS ARBETE MED SCB:S PARTISYMPATIUNDERSÖKNINGAR

Staffan Sollander, Bromma, har avlidit vid en ålder av 81 år. Hans närmaste är hustrun Christina och barnen Calle och Clara med familjer.

Vår vän och gamla arbetskamrat Staffan har lämnat oss efter en tids sjukdom. Vi var många 40-talister som fick jobb på Statistiska centralbyrån (SCB) i slutet av 60- och början på 70-talet och som kom att bli SCB trogna under hela vårt yrkesliv. Vi trivdes med arbetsplatsen och med varandra. Staffan var en av oss.

Staffan var i grunden metodstatistiker. Efter ett gediget utvecklingsarbete var han fram till sin pensionering ansvarig för SCB:s partisympatiundersökningar (PSU). Undersökningarna, som har genomförts sedan 1972 (med ett kort uppehåll i början på 80-talet) blev snabbt SCB:s populäraste. Stora urval och öppet redovisade metoder gjorde att PSU kom att inta en särställning bland politiska

opinionsundersökningar. Staffan var under alla år en stor auktoritet inom svensk opinionsundersökning.

Vi kände honom som en vänlig, omtänksam och humoristisk person. Med sin snabbtänkhets och associationsförmåga kunde han vara otroligt underhållande. Hans skarpa intellekt kom inte bara väl till pass i statistiska sammanhang, han var också en mycket praktiskt lagd person. Han mekade med bilar, diskmaskiner och alla upptänkliga tekniska prylar. Staffan drog in vattenburen värme i den omoderna Brommavillan, som givetvis behövde ett löpande underhåll som han själv fixade. När drivaxlarna på Calles folkabuss gick sönder ryckte Staffan ut.

Tiden räckte också till för att heja på Clara när hon framgångsrikt deltog i olika hästtävlingar. Staffan, Clara och hästvagnen tillbringade många timmar

på landets vägar. En lite mer udda aktivitet var Staffans munkförsäljning under många år vid jazz- och bluesfestivalen på Skeppsholmen.

Även om det är många år sedan vi alla gick i pension, har vi på olika sätt fortsatt att träffas och trivas med varandra, men nu är en lång gemensam tid av våra liv slut – Staffan fattas oss!

*För många gamla
arbetskamrater på SCB,
UNO DAVIDSSON*

Vill du annonsera i Qvintensen?

Nu öppnar vi upp för annonsering i Qvintensen! Både kommersiella och ideella aktörer kan annonsera. Det kan till exempel gälla ett evenemang eller en produkt som är av intresse för våra medlemmar.

Prislista för annons (exkl. moms)

	Ideella/ icke-vinstdrivande	Kommersiella/ vinstdrivande
1/2 sida	2 500 kr	12 500 kr
1/4 sida	1 500 kr	8 500 kr

Om du är intresserad av annonsering, kontakta Qvintensens redaktion på qvintensen@statistikframjandet.se.



Bli medlem i Svenska statistikfrämjandet

Svenska statistikfrämjandets syfte är bland annat att främja sund användning av statistik som beslutsunderlag och att väcka och sprida intresse för statistik i samhället.

För att bli medlem, gå till <http://www.statistikframjandet.se> och läs mer i högerspalten under "Bli medlem". Har du frågor kontakta Mattias Strandberg på sekreterare@statistikframjandet.se.

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.

