

Tilfeldig gang nr. 2/2018

Karl Ove Hufthammer · Anne Marie Fenstad

Innhold

Redaksjonelt

Frå leiaren

Frå redaksjonen

Medlemskontingent for 2018

Folk

Knut Kristian Aase 70 år og pensjonist!

Ti pionerer i skandinavisk statistikk – en digital utstilling

Statistikeromtale i media

På skrivebordet: Geir Aamodt

Bokmeldingar

Bokmelding: *Innføring i statistikk og dataanalyse for studenter i idretts- og helsefag*

Lesarbidrag

WISEflow – aldri meir!

Møte og konferansar

Invitasjon det 20. norske statistikermøtet

ISBA 2018

Slektstreff i Tartu

Meldingar

Nytt frå Noregs teknisk-naturvitskaplege universitet

Nytt frå Universitetet i Oslo

Nytt frå Nytt fra Oslo senter for biostatistikk og epidemiologi

Pusleri

Løsning på pusleri nummer 49

Pusleri nummer 50

Redaksjonelt

Frå leiaren

Marie Lilleborge

Velkommen tilbake fra høstferie! Jeg håper sommerhalvåret har vært inspirerende både på det faglige planet og i livet ellers. Norge er et godt land å være statistiker i, med tydelige årstider som hjelper til å holde rytme og variasjon i arbeidshverdagen. Kanskje har du fått reise på en workshop eller konferanse i vår eller i sommer, eller kanskje hadde du anledning til å trekke «kontoret» ut i solen en deilig sommerdag. Den varmende og sprudlende årstiden inviterer mange til å trekke ut og skaffe nye impulser og dele egen innsikt. Så kommer høsten, og det er godt å trekke seg tilbake med (kaffe og) ny inspirasjon til å sette seg dypere inn i noe nytt, flette sammen idéer, følge tankerekker helt ut, anvende og/eller konstruere. Og når dine nye ideer selvfølgelig tar deg lenger (tid å fullføre) enn du hadde tenkt, begynner du kanskje å glede deg til det 20. norske statistikermøtet på Sola Strand Hotel rett utenfor Stavanger, 17.–20. juni! Nettsiden til det kommende møtet har fått sine første byggesteiner på plass, og du kan lagre adressen <https://sites.google.com/site/statistiskforening/moeter/stavanger-2019> i nettleseren din allerede i dag. Jeg ser frem til vi møtes!



Marie Lilleborge.

Statistikermøtet er også arena for utdeling av NSF sine forskningspriser, nemlig Sverdrup-prisen «for unge forskere» og Sverdrup-prisen «for fremragende representant for statistikkfaget». Dersom du ikke er kjent med disse prisene, håper jeg du vil lese mer om dem på [våre nettsider](#). Uansett vil jeg anmode deg om å nominere aktuelle kandidater per e-post til NSF-styret. Ved årsskiftet vil styret

sende lister med kandidater til Sverdrup-komiteen, hvorpå komiteen går nøye gjennom disse og gir sin anbefaling tilbake til NSF-styret. Vi er altså avhengig av at du som medlem ser de gode statistikerne rundt deg, og gjerne nominerer en forsker eller to du er glad for at har funnet sin plass i statistikkverdenen.

Foruten å nyte sannsynlighet og statistikk, overraskelser og læring – og [Statistrikk](#)! – vil jeg anmode deg om å også bruke noen av høstens timer til å ta godt vare på deg selv. Nyt en varm kopp te, husk lue ute, planlegg for at renta går opp og vurder influensavaksine! Selv har jeg nylig gitt meg selv vannkoppvaksine i trettiårspresang. Og dersom det nærmer seg 10 år siden du sist tok DPT-vaksinen, [må du fylle på](#). Sjekk deg! («Livet er det fineste».) Og selvfølgelig, ta vare på dem rundt deg. Fortsatt god høst!

Smil fra Marie

Frå redaksjonen

Karl Ove Hufthammer

Velkommen til dette haustnummeret av TG!

Denne gongen har me ein noko amputert redaksjon, men heldigvis har me fine lesarar, som bidrar til å fylla TG med både tekst og foto – til glede for oss alle!



TG-redaksjonen. Foto: Roy Miodini Nilsen.

Og me tar sjølvsagt svært gjerne imot meir stoff, til framtidige nummer. Her er eksempel på *noko* av det me er interesserte i få meir av i bladet:

- «Jobben min som statistikar»/«Studiet mitt som statistikkstudent».
- Reisebrev (frå statistikarar på utanlandsopphald).
- Nytt frå det statistiske fagmiljøet ditt.
- Omtale av ei statistikkbok – eller ei anna bok som bør vera av særskild interesse for statistikarar – som du har lese.
- «Lesarinnlegg» om noko (statistisk) som du er opptatt av.
- Rapportar frå møte, konferansar, kurs eller liknande statistikkarrangement.
- Gratulasjonshelsingar for norske statistikarar som fyller runde dagar eller har vunne (statistikkrelevante) prisar.
- Minneord om norske statistikarar som har gått bort.
- Statistikkvitsar (eller -gåter/-oppgåver).
- Kommentatarar til tidlegare TG-stoff.

Send lesarbidrag til redaktøren på adressa <karl@huftis.org>.

Elles er det kjekt å sjå at statistikk og statistikarar faktisk viser seg å vera emne ikkje berre TG men òg *riksmedia* synest det er verdt å skriva om! Trygve Nilsen fortel for eksempel i dette nummeret om to tilfelle der statistikarar har fått [omtale i media](#) – både sterk ros og harselas.

Og nettavisens redaktør Gunnar Stavrum skreiv [i ein kommentar-artikkel](#) tidlegare i oktober om tolking av meningsmålingar: «Svært ofte slår man opp små bevegelser som er innenfor undersøkelsens usikkerhet [...] Men dagens «meningsmåling» i VG tar kaka.»

Det viste seg nemleg at VGs artikkel – om at eit fleirtal av KrF-veljarane syntest KrF burde gå inn ei *Støre-leia* regjering – var basert på berre 48 respondentar. Så konklusjonen var etter forfattarens syn «tatt ut av luften» og «fake news».

[Faktisk.no-redaksjonen](#) tok ein faktasjekk ([gjengitt i VG](#)), og kom til stikk motsett konklusjon. Dei ba blant anna to statistikarar ved Norsk

Regnesentral om hjelp til å rekna, og kom fram (basert på utvalet av dei 44 respondentane som faktisk valde side, så vidt eg kan sjå) til at «forskjellen mellom de to gruppene er klart signifikant».

Den grundige faktasjekken er verd å lesa. Ta òg ein tur innom tilhøyrande VG-kommentarfelt, om du tør. Der finn me fleire som er *overbeviste* om at 48 personar er *altfor lite* til å trekka konklusjonar av. Skal tru om dei same personane, viss dei for eksempel kasta ein mynt 48 (ev. 44) gongar og berre fekk kron på 8 av kasta, ville vore like avvisande til at det *var* ein god del som tydde på at det var mindre enn 50 % sjanse for kron med denne mynten. (Det er sjølvsagt òg *andre* feilkjelder ved meiningsmålingar enn dei reint sannsynsbaserte, men det var ikkje tema i artiklane.)

Elles finn me meir stoff om statistikarar i *vår eigen* [«folk»-spalte](#) – der den eldste statistikaren rett nok er over 300 år gammal! For yngre statistikarspirer har me bokmelding av ei ny [innføringsbok i statistikk](#). Og når det gjeld digital eksamen i statistikk for *lesarar* av slike bøker, har me eit innlegg om kva [teknologi som iallfall ikkje bør brukast](#).

Redaksjonen set alltid pris rapportar frå konferansar og møte for statistikarar, og i dette nummeret har me fått tilsendt reisebrev frå [to slike](#). Og til slutt har me [nytt frå dei som utdannar nye statistikarar](#), samt eit nytt [pusleri](#), som er vel verdt å prøva seg på! Med litt tenking bør det la seg løysa, òg av ferske statistikarar. (Men den mest openberre løysinga er kanskje ikkje den heilt riktige ...)

God lesnad!

Karl Ove Hufthammer (redaktør)
Anne Marie Fenstad

Medlemskontingent for 2018

Turid Follestad

Har du husket å betale kontingent i Norsk statistisk forening for 2018?

Medlemskontingenten i NSF for 2018 er kr 250,- per år for ordinære medlemmer. Kontingenten er fra og med i år hevet med 50 kroner, etter vedtak på hovedmøtet i juni 2017. Mastergradsstudenter får gratis medlemskap i to år fra innmelding.

Betalingsfrist: **1. juni 2018**

Hvordan betale kontingenten?

Kontingenten kan betales på en av følgende to måter:

- Ved overføring til kontonummer 0530.27.73299
- Ved bruk av Vipps, til Vipps-nummer 97112

Vennligst oppgi hvem og hvilket år kontingenten gjelder for ved innbetaling. Det første er spesielt viktig dersom arbeidsgiveren din betaler kontingenten.



Hva gjør vi med pengene?

De største utgiftspostene for NSF er overføringer til [lokallagene i Oslo, Bergen og Trondheim](#), utdeling av [Sverdrup-prisen](#), NSF's medlemskap i the [European Mathematical Society](#) (EMS) og [The International Statistical Institute](#) (ISI), og utgifter i forbindelse med diverse arrangementer (blant annet støtte til studenter ved deltakelse på [det norske statistikermøtet](#)).

Hvordan sjekke om kontingenten er betalt?

Alle medlemmene skal ha fått en individuell e-post rundt 10. mai i år, med beskjed om når kontingenten sist ble betalt. Dersom du ikke fikk en slik epost, eller du har spørsmål om kontingentbetalingen, kan du

[kontakte kassereren per e-post.](#)

Medlemsregister

For at medlemsregisteret skal være best mulig oppdatert er det viktig at du oppgir eventuell ny post-/e-postadresse ved å sende en e-post til kassereren på <turid.follestad@ntnu.no>. Det er opprettet en egen e-postliste for NSFs medlemmer, nsf-medlemmer@math.uio.no. Dersom du ikke mottar e-post fra denne listen, send den oppdaterte e-postadressen din til kassereren.

På forhånd takk!

*Turid Follestad,
kasserer i NSF*

Folk

Knut Kristian Aase 70 år og pensjonist!

Jostein Kåre Lillestøl

Professor Knut Aase ved Norges Handelshøyskole fylte 70 år 20. august, og gikk dermed formelt over i pensjonistenes rekke. Jubilanten har sin bakgrunn fra matematisk statistikk, men beveget seg tidlig over til forsikringsøkonomi, der han har blitt en internasjonal kapasitet.

Knut Aase er født i Meland i Nordhordland, og tok *examen artium* ved Bergen katedralskole i 1968. Han studerte realfag ved Universitetet i Bergen, der tilbudet og veiledning i matematisk-statistiske emner var i



Knut Kristian Aase.

sin spede begynnelse. Han tok hovedfag i anvendt matematikk i 1975 med Ivar Heuch som veileder. Hovedoppgaven var en anvendelse av sannsynlighetsteori på problemer innen populasjonsgenetikk. Deretter gikk turen med stipend til University of California, Berkeley, der han i 1979 tok sin PhD ved Department of Statistics med David Brillinger som veileder.

Tilbake i Norge arbeidet han noen år ved distriktshøyskolene i Agder (Kristiansand) og i Telemark (Bø), en pionértid for det nye skoleslaget. I 1982 ble Knut ansatt som førsteamanuensis i matematikk og statistikk ved Norges Handelshøyskole. Han orienterte seg da i retning av risikoteori og forsikring, og opparbeidet seg raskt kompetanse på dette feltet.

Ved Karl Borchs bortgang i 1986 var Knut godt kvalifisert for å ta over etter han. I 1989 ble han utnevnt til professor i bedriftsøkonomi, med ansvar for fagområdet forsikringsøkonomi og risikostyring. Da den nye finansmatematikken skyllet over oss på 1990-tallet, viste Knut at han var rett mann på rett plass, og et fruktbart samarbeid med professor Bernt Øksendal ved Matematisk institutt ved Universitetet i Oslo kom i stand, til beste for begge institusjonene. I denne sammenheng var Knut professor II ved Universitet i Oslo i årene 1993 til 1998, og tilknyttet dets *Centre of Mathematics for Applications* i årene 2005–2015, et av sentrene for fremragende forskning finansiert av Forskningsrådet.

Sammen med Bernt har Knut også engasjert seg i oppbygging av studieprogrammer på master- og doktorgradsnivå i det sørlige Afrika. Han har også hatt flere engasjementer som gjesteprofessor ved ledende amerikanske universiteter. Knut Aase har hatt flere redaktørverv i vitenskapelige tidsskrift og har blant annet vært norsk redaktør i *Scandinavian Actuarial Journal* i ti år.

Knut Aase kan vise til en omfattende vitenskapelig produksjon, med



I arbeid på kontoret.

nær 70 internasjonalt publiserte artikler. Han har hovedsakelig arbeidet med grunnforskning: Først innenfor statistikk, sannsynlighetsteori og økonometri, senere innenfor forsikringsøkonomi og matematisk finans. Han har vært en vital drivkraft for å utvikle forsikringsfaget ved å relatere det til moderne finansteori, og dermed bidratt til å forvalte NHHs internasjonale renommé som «The Norwegian School of Uncertainty».

I mange av sine senere arbeider kombinerer han avansert teori for stokastiske prosesser (typisk: hopp-diffusjoner) med økonomisk teori (typisk: likevektsantakelser). Han har fortsatt full fart i forskningen, og så sent som i 2016 presenterte han etter flere års arbeid en løsning på det såkalte «equity premium puzzle» som ble fremsatt i 1985 av Mehra og Prescott (nobelprisvinner). Arbeidet er publisert av *The Econometric Society* i tidsskriftet *Quantitative Economics*. Knut erstatter forventet nytte med såkalt rekursiv nytte, og viser at hans modell forklarer internasjonale konsum- og markedsdata for de siste 150 årene langt bedre enn standardmodellen. Artikkelen gir innsikt som har relevans

for viktige langsiktige beslutninger som livsforsikring, pensjonssparing og klimainvesteringer.

Det er en usedvanlig ungdommelig 70-åring som nå trer over i emeritusenes rekker. Gjennom hele sitt voksne liv har han dyrket sine fritidsinteresser løping, havkajakpadling og dristige fjellturer. Han har også tatt seg av gården på Litlebergen, der han vokste opp. Det statistiske fagmiljøet ønsker han alt godt som emeritus – og håper at hans brede engasjement vil fortsette.

Ti pionerer i skandinavisk statistikk – en digital utstilling

Ørnulf Borgan



ANDERS NICOLAI KIÆR
ADMINISTRATIVE STATISTICS
1838 – 1919

HEAD OF THE CENTRAL BUREAU OF STATISTICS
KIÆR DEVOTED HIS LONG LIFE TO THE DEVELOPMENT OF OFFICIAL STATISTICS IN NORWAY. AS A YOUNG LAWYER HE CAME TO THE DEPARTMENTAL STATISTICAL OFFICE. ON HIS INITIATIVE THIS OFFICE WAS IN 1878 REORGANIZED AS THE CENTRAL BUREAU OF STATISTICS AND KIÆR WAS ITS DIRECTOR. THIS POSITION HE HELD FOR 46 YEARS AND UNDER HIS LEADERSHIP THE BUREAU GREW INTO AN IMPORTANT INSTITUTION OF THE NATION.

OFFICIAL STATISTICS
MANY OF THE INVESTIGATIONS IN THE AREAS OF ECONOMICS, DEMOGRAPHY AND SOCIOLOGY WAS PERSONALLY LEAD BY KIÆR AND HE NORMALLY WROTE THE INTRODUCTIONS TO THE VARIOUS STATISTICAL PUBLICATIONS.

THE REPRESENTATIVE METHOD OF SAMPLE SURVEYS
IT WAS CLEAR THAT MANY IMPORTANT QUESTIONS WOULD REMAIN IN THE DARK IF ONLY THE METHOD OF TOTAL ENUMERATION SHOULD BE USED IN STATISTICAL INVESTIGATIONS. IN 1891 KIÆR DESIGNED AN INVESTIGATION OF PERSONAL INCOME AND CAPITAL BASED ON A QUASI RANDOM SAMPLE, AND IN 1895 HE USED A METHOD OF PROPORTIONAL STRATIFIED RANDOM SAMPLING. THE MAP PREPARED BY KIÆR AND DISPLAYED ABOVE INDICATES HIS SAMPLING PLAN.
KIÆR WAS NO MATHEMATICIAN AND DID NOT FORMALLY DEVELOP HIS IDEAS OF THE "REPRESENTATIVE METHOD OF SAMPLE SURVEYS". HE DESCRIBED HOWEVER IN DETAIL HOW THE SAMPLE WAS DRAWN AND NOTED IN ONE CASE THAT "THE SAMPLING MECHANISM IS SO HAPHAZARD AND RANDOM RELATIVE TO THE POPULATION THAT ONE MAY ASSUME THAT IT DOES FUNCTION IN THE SAME WAY AS A LOTTERY."

RANDOM SAMPLING VERSUS "LA STATISTIQUE SÉRIEUSE"
KIÆR FOUND IT HARD TO CONVINCE THE 1895 ISI-MEETING THAT THE REPRESENTATIVE METHOD WAS USEFUL. IT WAS ARGUED THAT ALTHOUGH THE METHOD NOW AND THEN HAD GIVEN INTERESTING INFORMATION, ITS PRINCIPLE WAS SO MUCH IN OPPOSITION TO "LA STATISTIQUE SÉRIEUSE" THAT THE IMPERFECT METHOD SHOULD NOT BE GRANTED EQUAL STATUS WITH THE STATISTICAL IDEAL.

INTERNATIONAL COOPERATION
KIÆR WAS AN ACTIVE MEMBER OF THE ISI, AND TOOK PART IN ITS FOUNDATION IN 1885. IN 1899 HE WAS THE HOST FOR THE FIFTH ISI MEETING IN KRISTIANIA (OSLO) - THE ONLY TRULY INTERNATIONAL STATISTICAL MEETING IN NORWAY BEFORE THE PRESENT ONE. A PROBLEM WHICH INTERESTED KIÆR WAS THAT OF ESTIMATING THE SIZE OF THE POPULATION IN COUNTRIES LIKE CHINA WHICH HAD NO REGULAR CENSUS.

Det 11. europeiske statistikermøtet fant sted i Oslo 14–18. august 1978. Til møtet laget Tore Schweder en utstilling av om ti av pionerene i skandinavisk statistikk:

- Pehr Wilhelm Wargentin (1717–1783)
- Eilert Sundt (1817–1875)
- Ludvig Henrik Ferdinand Oppermann (1817–1883)
- Thorvald Nicolai Thiele (1838–1910)
- Anders Nicolai Kiær (1838–1919)
- Jørgen Pedersen Gram (1850–1916)
- Carl Ludwig Charlier (1861–1934)
- Jarl Waldemar Lindeberg (1876–1932)
- Filip Lundberg (1876–1965)
- Agner Krarup Erlang (1878–1929)

De ti arbeidet innen ulike deler av faget vårt, fra beskrivende populasjons- og sosialstatistikk til matematisk statistikk og sannsynlighetsregning. Målet med utstillingen var å belyse arbeidet til de ti pionerene og sette det inn i en historisk sammenheng.

I forbindelse med 40-årsjubileet for møtet i Oslo, har vi fotografert utstillingsplakatene og lagd en [digital versjon av utstillingen](#). Teksten på plakatene er forfattet av Tore Schweder, mens faren hans, Carl Schweder, utformet plakatene og tegnet portrettene av de ti pionerene.

Statistikeromtale i media

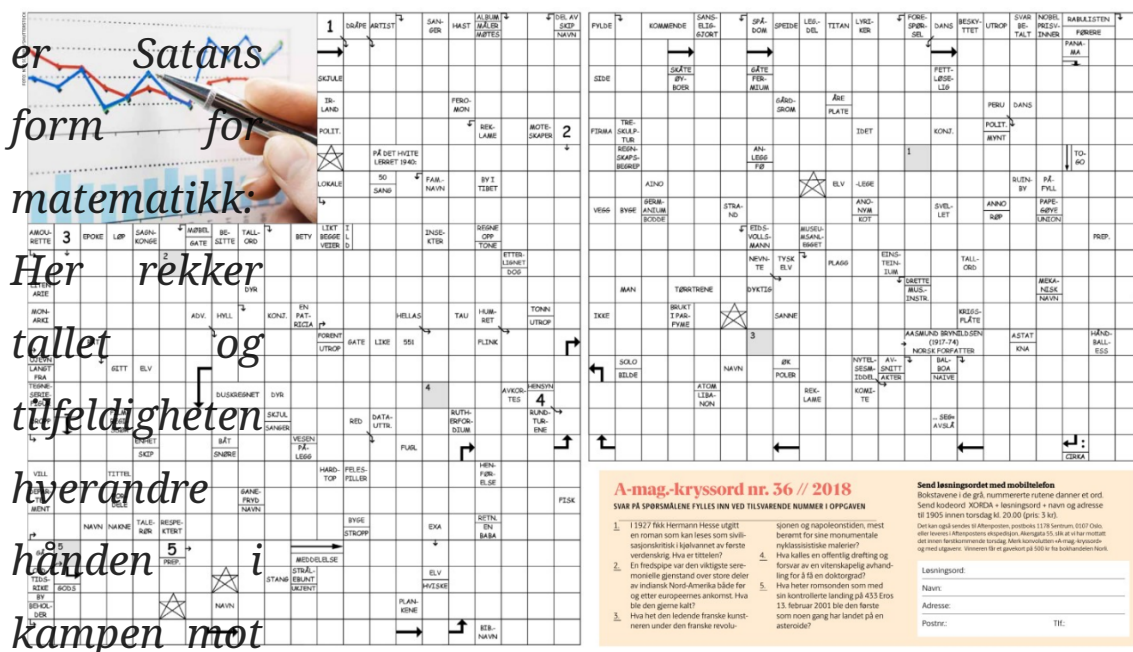
Tryve Nilsen

I høst er statistikeren både blitt harselert med (7. september) og takket for godt arbeid (26. september) i Bergens Tidende.

Statistikerharselas

Den lange setningen som det letes etter i kryssordet er:

Kryssord Statistikken



det
 Faksimile fra A-magasinet (vedlegg til blant annet Bergens Tidende)
 individuelle. 7. september 2018.

Statistikertakk

Bergens Tidende, «[Riddere av de kunstige ledd](#)» (krever abonnement):

Onsdag blir [Lasse] Engesæter og mangeårig kollega Leif Ivar Havelin begge utnevnt til Ridder av 1. klasse av Den Kongelige Norske St. Olavs Orden for sitt arbeid innen ortopedi. Særlig arbeidet med å bygge opp Nasjonalt register for leddproteser.

“Engescæter: Vi har førsteklasses statistikere og kontorpersonale i staben. Det er godt å se at flinke unge folk overtar etter oss.

På skrivebordet: Geir Aamodt

Anne Marie Fenstad

TGs faste spalte «På skrivebordet» – fritt bygget på Dagens Næringsliv
«På nattbordet» – har denne gangen snakket med Geir Aamodt fra



Institutt for folkehelsevitenskap, Norges miljø- og biovitenskapelige universitet (NMBU).

For de av leserne som er kommet opp i en moden alder, er NMBU kanskje bedre kjent som Universitetet for miljø- og biovitenskap og Norges landbrukshøgskole, og da kanskje aller helst Landbrukshøgskolen. NMBU er passende lokalisert i landlige omgivelser på Ås utenfor Oslo. Som den oppvakte leser muligens erindrer fra et [tidligere nummer av TG](#), ble Geir utfordret til å delta i «På skrivebordet» av Stein Atle Lie, professor ved institutt for klinisk odontologi, Universitetet i Bergen.

Geir har vært ansatt ved NMBU i 100 %-stilling siden 2014 og trives som «plommen i egget». Hans hovedinteresse innen forskning er å benytte GIS (geografiske informasjonssystemer) og romlig statistikk for å forstå hvordan miljø påvirker helse. Spesielt barns helse er viktig.

I NSF Bergen var vi så heldige å ha [Geir på besøk i 2013](#). Da jobbet han ved Folkehelseinstituttet og holdt et spennende foredrag om «Miljøepidemiologi og GIS». Miljøepidemiologi har fokus på det ytre miljøet og hvordan det påvirker helsa vår. Typiske eksponeringer er

luft og luftkvalitet, drikkevann, støy, stråling og tilgang til grønne arealer, og GIS blir brukt for å beregne disse eksponeringene for personer som er med i helseundersøkelser eller helseregistre.

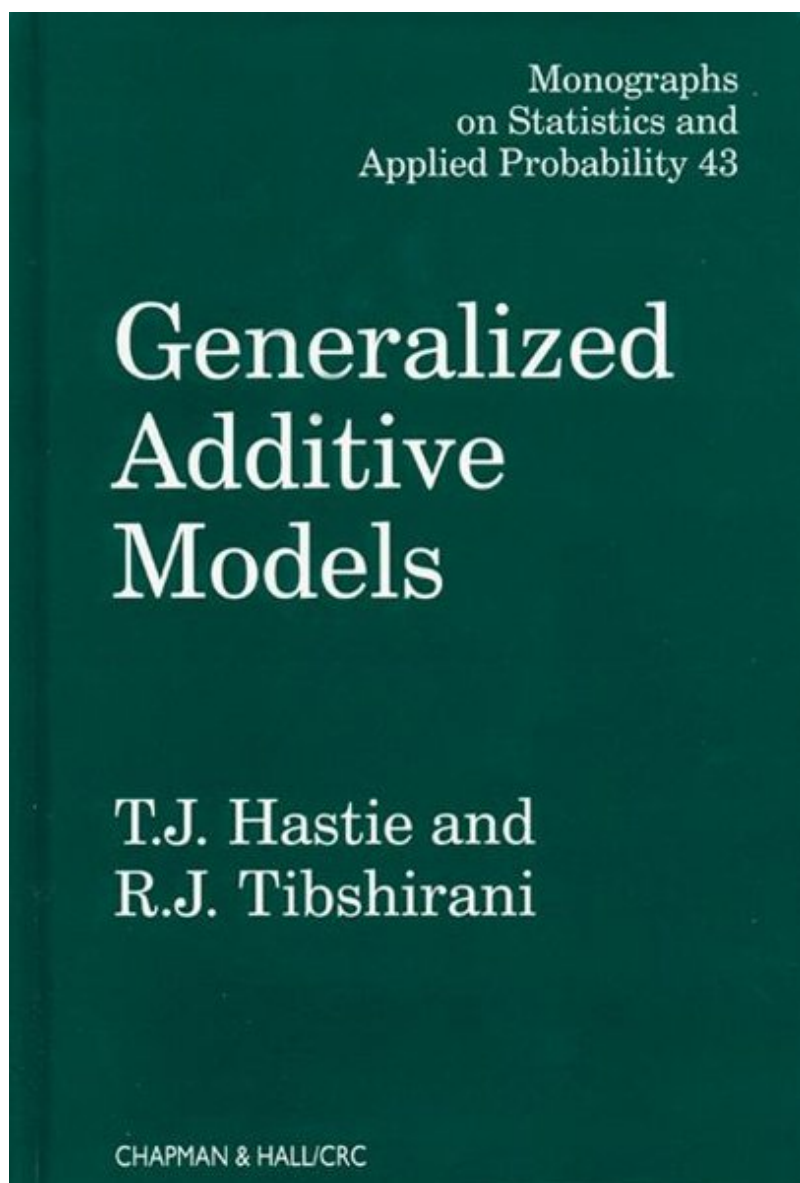
Når TGs reporter spør etter Geirs favoritt-fagbok, er han ganske tydelig på at det må være [*Generalized Additive Models*](#) av Hastie og Tibshirani, utgitt på Chapman & Hall i 1990. «Etter å ha lest denne boka, fikk jeg et helt nytt syn på modellering», sier Geir.

Geir har geologi som bakgrunn før han ble statistiker. Han studerte først geofag på NTNU: «Grunnkurset i geologi og den boka vi hadde i det faget var veldig spennende.

Friluftsliv ble noe

annet etter å ha lest den boka. Ja, geologifaget fikk en motsatt effekt på meg enn det jeg tenkte da jeg begynte å studere faget. Jeg forsto – litt for seint – at vi må ta vare på jorda og ikke ødelegge den.» Det med å forstå litt sent at man må ta vare på jorda kan sikkert mange relatere seg til. Husk, det er ikke for sent å begynne nå ...

Liv Høyland og Emil Spjøtvoll i Trondheim og seinere Tore Schweder



åpnet øynene til Geir for verdener som var spennende og vakre. «Kurset i multivariate metoder, som Spjøtvoll underviste, syntes jeg var magisk, og da jeg seinere ble kjent med Tore og fulgte hans og Erik Bølvikens kurs i redusert-rang-regresjon, ble jeg statistiker», forteller Geir.

Arnljot Høylands bok i elementær statistikk er den første fagboken Geir husker. Denne ble brukt i introduksjonskurset på NTNU (den gang NTH). Det at Geir gikk fra geofag til statistikk kom nok ikke helt ut av det blå; favorittfaget på skolen var alltid matte. Fritidsinteresser utover faglige diskusjoner er friluftsliv og musikk, opera og ballett. Han liker også å lese utenom faglitteratur, og favorittforfatteren er Sigrid Undset. Geir har flyttet til Vestby i Follo og bor nå midt i Sigrid Undsets rike. «Her har hun lagt handlingen til både Olav Audunssøn fra Hestviken, og her bodde Kristin Lavrandsdatter før Undset flyttet henne til Gudbrandsdalen. Undset bodde selv noen måneder både i Hvitsten og i Hølen, som er to vakre idyller i Vestby. Deler av handlingen i *Olav* foregår i Garder og Såner, som er andre vakre steder i mine nærområder», sier Geir begeistret.

Den morsomste boka Geir har lest er «Hemmelig dagbok for Adrian Mole, 13 $\frac{3}{4}$ ». Det er en britisk ungdomsbok som kom ut i 1982 og er skrevet av Sue Townsend. Boka handler om snart 14 år gamle Adrian Mole og hans hverdagsliv i Leicester tidlig på 1980-tallet. «Jeg lo fra første til siste side!»

Til neste nummer av «På skrivebordet» vil Geir Aamodt gjerne utfordre Solve Sæbø, som nå er ny prorektor på NMBU.

Bokmeldingar

Bokmelding: *Innføring i statistikk og dataanalyse for*

studenter i idretts- og helsefag

Eivind Hoffmann

Jon Ingulf Medbø:

[Innføring i statistikk og dataanalyse for](#)

[studenter i idretts- og helsefag](#). Cappelen

Damm Akademisk AS.

Oslo 2018. 413 sider.

Anmelderen av denne læreboka har lite kjennskap til målene for den undervisningen som gis til de

studentene som er målgruppen, men ser av forordet at den er ment å være «ei brukerretta lærebok for studenter som jobber med praktisk statistikk».

Det blir også

vist til at «i dag har vi dataprogram som utfører regnearbeidet for oss» og at «da er det viktigere at brukeren vet hvilket verktøy som skal brukes, enn å kjenne til de regnetekniske detaljene».

Gitt dette utgangspunktet er det både overraskende og skuffende at det flere steder understrekes at «statistikk er ei grein av matematikken» (blant annet på side 20), og at sitatet «Faget tar for seg innsamling, klassifisering, presentasjon og tolkning av data.» (R.R. Johnson: *Elementary Statistics*, Duxbury Press – 6. utgave 1992) sies å være «ei



kort og grei sammenfatning av hva statistisk bearbeiding av tallmateriale dreier seg om for brukere uten godt grunnlag i matematikk» uten at det reflekteres i teksten. Mange andre disipliner enn matematikk er relevante for en slik, korrekt, beskrivelse av statistikkfaget, men i realiteten er det bare det siste trinnet i denne listen som presenteres i lærebokas tekst, selv om kapittel 2, «Innføring i kvantitative metoder» (side 32–93, med vedlegg), og eksemplene på konkrete datasett ellers nok kan sies å inneholde stoff som er relevant for de tre andre trinnene.

Antagelig er det eksemplene som gjør at forfatter og forlag mener at læreboka vil være særlig egnet for «studenter i idretts- og helsefag», for ellers er det vanskelig å se hvordan teksten i kapitlene 3–11 (side 94–374) skiller seg fra andre elementære lærebøker i statistisk analyse for studenter som ikke forventes å ha tilegnet seg det matematiske grunnlaget for de metodene som presenteres.

Gitt det innledende sitatet om «dataprogram som utfører regnearbeidet» er det skuffende at læreboka i sin presentasjon av de ulike analysemetodene ikke sier noe tydelig om vilkårene for å velge én analysemetode framfor en annen. Å gi slik forståelse bør vel være en sentral oppgave for en lærebok for den oppgitte målgruppen. Også i presentasjonen av ulike spredningsmål savnet jeg tydelige beskrivelser av kriterier for valg.

Med min bakgrunn i arbeidet med innsamling av data for statistikk og presentasjon av statistikken, var det mye å irritere seg over i presentasjonene av konkrete eksempler. Det synes mye mot en regel om at for både figurer og tabeller bør teksten gi nok informasjon til at leseren fullt ut forstår det som beskrives. Figurene 3.5 og 3.6 (side 100–101) kan tjene som eksempel: Det som beskrives skal være resultatene for målt høyde for «tilfeldig utvalgte norske menn», uten at det er sagt noe om tidspunkt for målingene, selv om det i teksten

opplyses at slike målinger har vært foretatt av det norske forsvaret i om lag 100 år. Det er videre opplyst at «målingene er gjort som simulering i Excel med funksjonen `=NORM.INV(TILFELDIG();180;7)` ». Det er ikke angitt en referanse der til hva denne funksjonen er og gjør, men leseren finner en relativt god forklaring på side 186–187, som nok ikke er helt enkel å forstå. Det er heller ikke der angitt hva som ligger i ordet «simulering».

De aller fleste eksemplene med tilknytning til idrett gjelder prestasjoner (resultater) i idrett eller forhold med antatt direkte relevans for slike. For en amatør er det vanskelig å se for seg andre arbeidssituasjoner der denne typen eksempler vil være relevante enn ved opplegg av treningsprogram for en større gruppe idrettsutøvere. De som primært skal trene enkelte utøvere kan kanskje ha nytte av de eksemplene som beskriver endring over tid som følge av ulike treningsregimer, men det er vanskelig å tenke seg at for eksempel Gjert Ingebrigtsen har hatt (eller savnet) slik kunnskap i sitt arbeid med treningsopplegg for sønnene. Mangel på eksempler som er knyttet til administrasjon av idrettslag og -forbund, eller helseaktiviteter, er derfor påfallende for en lærebok for studenter som nok i hovedsak vil få slike arbeidsoppgaver – om de i det hele tatt får arbeid som studiene skal forberede dem til innenfor de mest relevante sektorene.

Det skal ikke underslås at boka har en god del fornuftige formuleringer og observasjoner, for eksempel om «Desimalsjuken» (side 66–68). Der savnet jeg imidlertid en indikasjon på at det har liten mening i å angi en prosentfordeling for undergrupper når antallet i populasjonen er meget mindre enn 100, og å angi gruppeprosenter med desimal når populasjonen er mindre enn 1 000.

Konklusjon: Etter min mening har forlaget ikke dekning for påstanden i baksideteksten om at dette er «en brukerrettet lærebok for studenter som jobber med praktisk statistikk», eller som skal jobbe med å hente

inn, forstå eller analysere slik statistikk etter studiet. Til det er for mange viktige emner utelatt eller bare summarisk behandlet. Det er heller ikke dekning for påstanden om at «boken også er aktuell i masterarbeidet for mange av [studentene som følger den nye femårige mastergraden ved lærerutdanningen]».

*Eivind Hoffmann, cand.oecon. med økonometri som spesialfag.
Arbeidserfaring med statistikk fra Statistisk sentralbyrå, Norsk
regnesentral, International Labour Office og Utlendingsdirektoratet.*

Lesarbidrag

WISEflow – aldri meir!

Lars Arne Jordanger

Høgskulen på Vestlandet (HVL) har ein avtale med [WISEflow](#) som gjer at vi kan bruke dette som verktøy for digitale eksamenar. WISEflow er utvilsamt eit verktøy som det kan vere fornuftig å bruke når det i den tradisjonelle skriftlege eksamensforma er slik at det som studentane skal levere inn primært er eit tekstsvar (dvs. minimalt behov for formalar/figurar), men det betyr slett ikkje at WISEflow vil fungere bra for fagfelt der svara til studentane bør innehalde element som likningar, utrekningar og figurar.

Eg har våren 2018 undervist statistikkemnet [ØMO003](#) ved HVL, der ein digital eksamen via WISEflow vart gjennomført. Eg har sett på tidlegare ØMO003-eksamenar gjennomført i WISEflow, og eg har i tillegg hatt eit sensoroppdrag for eit metodekurs ved NHH der WISEflow vart brukt. Med utgangspunkt i dei erfaringane dette har gitt meg, kan eg for min del oppsummere mitt syn som følgjer: **Aldri meir WISEflow!** I den forma WISEflow har per dags dato, er det fullstendig ueigna som eit verktøy for å teste studentane sine kunnskapar i eit statistikkemne.

La meg med ein gong understreke at hovudårsaka til mitt negative syn på WISEflow ikkje heng saman med at grensesnittet for produksjon av oppgåver er elendig (det er nesten imponerende kor fantastisk ineffektivt det er), og det har

heller ikkje samanheng med at det har vore nødvendig å omstrukturere måten vi må stille spørsmåla på (sjølv om det også har vore eit irritasjonsmoment). Det som ligg til grunn for min sterke motvilje mot WISEflow er at det ikkje gir oss ein god måte å testa kunnskapen til studentane på, i tillegg til at det på grunn av den underliggande strukturelle oppbygginga faktisk ikkje gjer det enklare og meir effektivt å sensurere studentane sine svar.

Det sentrale problemet er at vi ikkje kan forlange at studentane skal meistre korleis dei i WISEflow sine tekstfelt skal kunne skrive inn likningar og utrekningar – og om vi hadde lagt inn det som eit krav, så ville studentane truleg brukt vesentleg meir tid på skrive svara sine enn det ville tatt dei ved bruk av penn og papir. Kva skal vi i så fall gjere når det er studentane si evne til å gjere utrekningar som vi er interessert i å teste? Det er, slik eg kjenner det, primært to løysingar som er aktuelle i WISEflow:

- Enten kan vi bruke ei løysing med val av eitt alternativ ut av mange.
- Eller så kan vi gå for ei løysing der studentane skal skrive inn det talsvaret dei har rekna seg fram til.

○ Henter seksjoner

Sensur via WISEflow er basert på kontakt mot ein database på deira server, og det er ikkje mogeleg å få systemet til å laste ned ein lokal kopi av svara til studentane. Det vert dermed mykje venting når sensuren skal gjennomførast, og teksten «Henter seksjoner» må ein då sjå på medan ein ventar og ventar på at ein vonleg snart skal få tilgang til ein ny student sitt svar.

Ingen av desse løysingane gir etter mi vurdering eit godt nok vurderingsgrunnlag.

Lat oss først sjå på det alternativet der studentane skal få velje eitt svar frå fleire mogelege. Ei utfordring med denne oppgåvetypen er å inkludere «fornuftige» alternativ for dei feilaktige svara, og for å få til det må vi tenkje gjennom alle dei «standardfeila» som vi veit at fagleg svake studentar kan kome til å gjere. Eg har sett oppgåver der slike «fornuftige feilalternativ» ikkje har vore med, med den konsekvens at dei studentane som gjer ein feil dermed får ein ny sjanse til å prøve seg på oppgåva. Ei dårleg utforma oppgåve av denne typen vil gje ei vridning mot at ein del studentar får ei betre utteljing enn dei eigentleg skulle hatt. Eit anna potensielt problem er at denne typen spørsmål opnar for at studentane kan gjette seg fram til svaret. Dette burde normalt ikkje vere eit stort problem dersom det er mange alternativ å velje mellom, men eg har sett oppgåver der det er nokre svaralternativ som «opplagt» ikkje er realistiske, og det gir ein mykje større sjanse for at studenten kan gjette rett på dei resterande alternativa.

På grunnlag av desse argumenta mot fleirvalsoppgåver, så brukte eg for min eksamen ei løysing der studentane skulle skrive svara sine inn i eigne felt. WISEflow har ei løysing der vi kan legge inn automatisk testing av talsvar, der vi i tillegg kan spesifisere at vi vil akseptere at svaret kan ligge i eit gitt intervall rundt den korrekte verdien. Denne løysinga burde sikre at ei avrunding i ei mellomrekning ikkje medfører at studenten får null poeng for eit svar som heilt klart må vere resultatet av ei fornuftig utrekning. Det er også mogeleg å legge inn ei føring om at tekst ikkje er lov å skrive inn i dei felta der vi vil ha eit tal som svar, men det viste seg dessverre at det WISEflow aksepterer som «ikkje tekst» var ganske annleis enn det eg i utgangspunktet trudde. Spesielt: WISEflow nyttar standarden at eit punktum vert nytta i staden for eit komma som desimalskiljeteikn, og eg trudde diverre at «ikkje tekst» ville hindre bruk av komma i dei felta der tal skulle skrivast

Oppgave 1 / oppgave 1

I ei stor industribedrift arbeider 75 % av dei tilsette med produksjon og 25 % med administrasjon. Gjennomsnittleg sjukefråvær er 8 % for dei tilsette i produksjonen, medan det er 4 % for dei som er i administrasjonen.

Vi tenkjer oss at vi trekk ut ein tilfeldig tilsett i bedrifta, og lar S vere utfallet at den tilsette er fráværande på grunn av sjukdom, medan vi lar A vere at vedkommande arbeider i administrasjonen.

Rekn ut følgjande sannsyn, og forklar med ord kva desse sannsyna uttrykkjer. **NB: Skriv svara som desimaltal (ikkje prosentar) og ta med minst 3 siffer etter komma.**

I en stor industribedrift arbeider 75 % av de tilsette med produksjon og 25 % med administrasjon. Gjennomsnittlig sykefravær er 8 % for de tilsatte i produksjonen, mens det er 4 % for de som er i administrasjonen.

Vi tenker oss at vi trekker ut en tilfeldig tilsatt i bedriften, og lar S være utfallet at den tilsatte er fraværande på grunn av sykdom, mens vi lar A være at vedkommende arbeider i administrasjonen.

Regn ut følgende sannsynligheter, og forklar med ord kva disse sannsynlighetene uttrykker. **NB: Skriv svarene som desimaltall (ikke prosenter) og ta med minst 3 siffer etter komma.**

×

$P(S) =$

1

0,070

×

Correct answers:

1

0.07

Sannsynet for at vi trekk ut ein som er fráværande på grunn av sjukdom.

Forklaring av $P(S)$:

Dette er sannsynligheten for at en tilfeldig trukket ansatt i bedriften er fraværande på grunn av sykdom.

17 / 30 Word Limit

Sidan «,» her er brukt i staden for «.» som desimalskiljeteikn, var WISEflow si vurdering at alle svara var feil og programmet gav derfor null poeng på alt. Det som står i det blåfarga tekstfeltet er ein merknad som viser sensor ein variant av det som studenten burde ha skrive i tekstfeltet under.

inn – men det stemde ikkje. Dermed vart det til at det for mange studentar var slik at dei hadde skrive inn svar som var korrekte, men som WISEflow gav null poeng (dersom fasiten i WISEflow til dømes var «3.14», så ville «3,14» bli vurdert som null poeng).

Dersom studenten skriv sitt svar slik at WISEflow oppfattar det som eit tal, og dersom svaret er i samsvar med den fasiten som er lagt inn – så er det naturlegvis greitt at systemet gir poeng for dette. Men eg ser det som problematisk at vi for svar som ikkje er korrekte ikkje har eit grunnlag for å kunne sjå om det er på grunn av at studenten har gjort ein liten slurvefeil eller om det er snakk om ein alvorleg forståingsfeil. For ein skriftleg eksamen med penn og papir er dette vanlegvis enkelt å få avklart, og slik eg ser det er det eit sterkt argument mot å erstatte vanleg skriftleg eksamen med digital eksamen i eit reknefag.

I tillegg til dei problema som er nemnde ovanfor, så var det eit anna og til dels mykje meir alvorleg problem som dukka opp når eg skulle gje poeng til studentane sine svar – og det var problematikken rundt poenggjeving ved følgjefeil. Dersom studenten gjer ein reknefeil i delspørsmål a), og svaret frå dette delspørsmålet inngår i utrekninga av delspørsmål b), så er det ved ein vanleg skriftleg eksamen enkelt å sjå om studenten har brukt rett framgangsmåte og derfor bør ha poeng for det som er gjort i utrekninga av delspørsmål b). Når denne typen feil dukkar opp i WISEflow, så gir derimot systemet automatisk null poeng til alt, sidan ingen av delspørsmåla då har svar som er i samsvar med det som er lagt inn som fasit. Dersom den som skal sensurere desse oppgåvene ønskjer å undersøkje om det er følgjefeil som ligg til grunn i enkelte av oppgåvene, så må det gjennomførast utrekningar basert på dei verdiane som studenten har hatt som utgangspunkt (altså dei feilaktige svara frå tidlegare delspørsmål). Dette er ikkje akkurat noko som gjer sensurarbeidet meir effektivt ...

For min del vart det i vår nødvendig å få tak i ei [CSV-fil](#) som hadde med rådata for alle dei svara som studentane hadde lagt inn i WISEflow, og deretter måtte eg sjølv skrive eit program som sikra at alt det som skulle vere tal vart konvertert til tal. Eg måtte også få programmet til å sjekke om det var grunnlag for å gje studentane poeng for følgjefeil. Dette meirarbeidet medførte at den sensuren som normalt burde tatt

fire arbeidsdagar tok meir enn to fulle arbeidsveker (om ikkje vesentleg meir enn det også).

Det bør nemnast at vi ikkje slepp unna problematikken med følgjefeil om vi går for den løysinga der studenten skal velje eitt av fleire svaralternativ. Eit eksempel på dette er hypotesetesting, der studenten skal samanlikne ein testobservator (rekna ut på grunnlag av ei samling av observasjonar) med ein kvantil (som er avhengig av signifikansnivået til testen). Det sentrale poenget er at studenten skal sjå om testobservatoren ligg innanfor (eller utanfor) ein kritisk region, for det avgjer om vi skal forkaste (eller ikkje forkaste) nullhypotesen for det spesifiserte signifikansnivået. For ei slik oppgåve er det altså essensielt at studenten sin konklusjon (forkast/ikkje forkast) er i samsvar med den testobservatoren og den kvantilen som han/ho har funne fram til. Det vil seie at det kan skje at det for denne typen oppgåve (på grunn av følgjefeil) skal gjevast poeng for den omvendte konklusjonen av det som står i fasit. Slike nyansar er WISEflow ikkje i stand til å fange opp.

WISEflow opnar naturlegvis opp for at vi kan inkludere felt der studentane skal skrive tekst også, slik at vi kan be dei skissere korleis dei har gått fram for å løyse ei oppgåve – eller sjekke om dei forstår kva det dei har rekna ut betyr. I samband med dette har eg sett at studentar har valt ut rett svar i ei fleirvalsoppgåve, for deretter å kome med ei forklaring som viser at dei har brukt ein metode som er fullstendig feil (og som normalt slett ikkje skulle gje noko i nærleiken av det korrekte svaret).

Ein detalj som eg ikkje har gått inn på i diskusjonen ovanfor er at det også er utfordringar dersom vi ønskjer å inkludere spørsmål der studentane skal lage enkle figurar, noko som ofte kan vere ein grei måte å sjå om dei har forstått det sentrale i den delen av pensum som vi spør om. Eg har ikkje god nok kjennskap til kva løysingar (om nokon)

som finnest i WISEflow med omsyn til det å få inkludert slike spørsmål, men med utgangspunkt i det eg til no har sett av WISEflow, ville eg bli ekstremt overraska om det faktisk finnest ei løysing som eg ville vurdert som god nok.

Som nemnd innleiingsvis, så vurderer eg WISEflow (i si noverande form) til å vere noko som eg ikkje ønskjer å ha noko meir med å gjere. WISEflow er rett og slett ikkje eit verktøy som hjelper meg med å få eit godt bilete av kunnskapen til studentane. Eg for min del vil heretter insistere på at det skal vere ordinær skriftleg eksamen i dei emna som eg underviser. **WISEflow? ALDRI MEIR!**

*Lars Arne Jordanger, førsteamanuensis ved Institutt for data- og realfag,
Fakultet for ingeniør- og naturvitskap, Høgskulen på Vestlandet.*

Møte og konferansar

Invitasjon det 20. norske statistikermøtet

Ingvild Dalen

Velkommen til Stavanger og det 20. norske statistikermøtet 17.–20. juni 2019.

Vi møtes på vakre Sola strandhotell til
faglig påfyll og hyggelig samvær.

Hold av datoene, dette blir knallkjekt!



Mer informasjon kommer på
<http://norskstatistiskforening.no/>.



ISBA 2018

Ingeborg Gullikstad Hem

I slutten av juni dro vi – to professorer (Andrea Riebler og Ingelin

Steinsland) og fire stipendiater (Torstein Fjeldstad, Thea Roksvåg, Maria Selle og meg selv) fra statistikkgruppa på NTNU – til Edinburgh for å delta på [ISBA World Meeting 2018](#) (ISBA = International Society of Bayesian Analysis).

ISBA avholdes annethvert år, og neste gang blir i Kina i 2020. Det var mange hundre bayesianere som deltok, og enten holdt presentasjon eller stilte med poster. Med så mange deltakere var det hele fem parallelle sesjoner med foredrag, og de inviterte hovedforedragene ble streamet til flere rom så alle skulle få dem med seg.



Utsikt fra Arthur's Seat.



Edinburgh Castle, som ligger på Castle Rock.

Det var et stappfullt program fra morgen til kveld! Det var utrolig mange spennende foredrag, og jeg synes det var ekstra gøy at alle foredragene (naturlig nok) handlet om bayesiansk statistikk. I tillegg til

alle foredragene og postersesjonene rakk vi også å få med oss litt av den skotske kulturen, og fikk både sett Arthur's Seat med utsikt over hele byen, skotter iført kilt og Edinburgh Castle.

Alt i alt var det en veldig innholdsrik, interessant og lærerik uke!

Slektstreff i Tartu

Denne reiseskildringen fra nordisk statistikermøte i juni er skrevet for sampublisering i TG (Tilfeldig gang, statistikernes interne og uregelmessige tidsskrift) og på Kathrines populærvitenskapelige strikkeblogg, [Statistrikk](#).

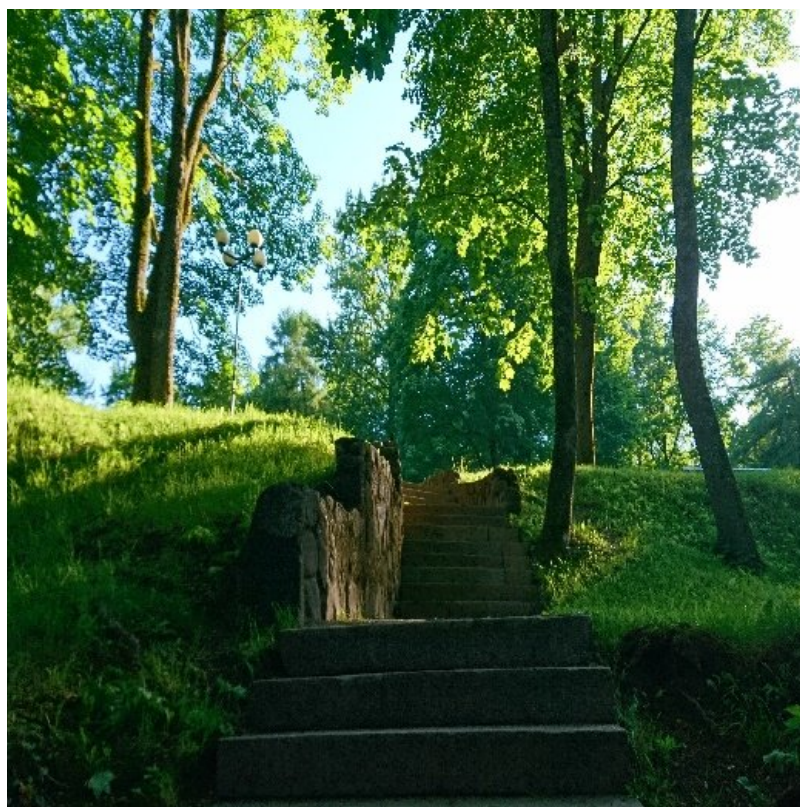
Med særdeles ikke-tilfeldige mellomrom møtes statistikere for å oppdatere hverandre om siste nytt på forskningsfronten:

- Har noen funnet på nye og lure metoder siden sist?
- Har noen utforsket nye sider ved gamle og lure metoder?
- Har noen brukt en velkjent, lur metode på en ny, lur måte?
- Eller har noen kommet på en lur måte å forklare lure metoder til dem som ikke er statistikere?

Og for å ha det hyggelig.

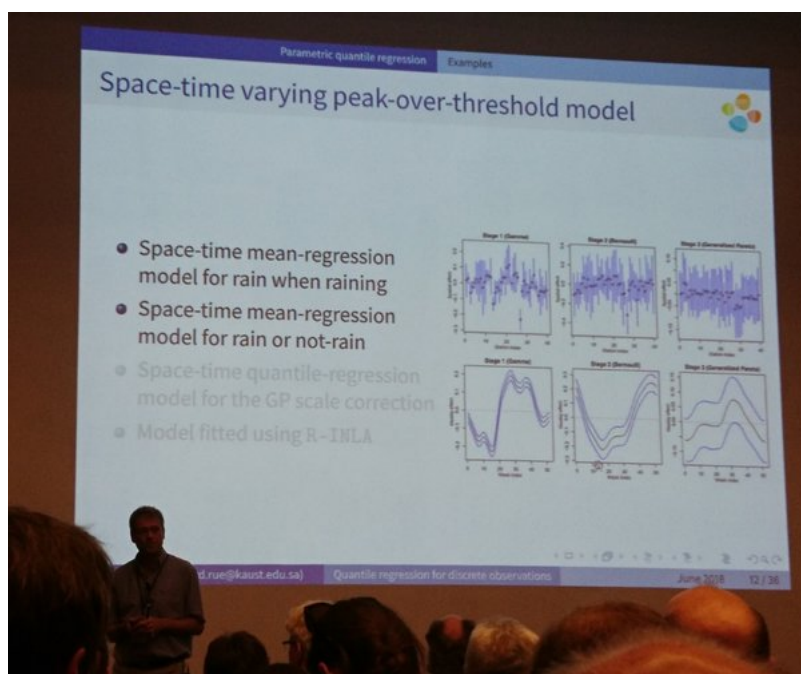
I fjor møttes den norske statistiske storfamilien i Fredrikstad. I år var det [nordisk slektstreff i den estiske byen Tartu](#).

I løpet av det første døgnet myldret statistikerne klyngevis til Tartu og Dorpat hotell (Dorpat er det gamle svenske og russiske navnet på Tartu). De smarteste



Bypark i Tartu.

kjørte buss. Tartu ligger godt inne i Estland, er Estlands nest største by med drøyt hundre tusen innbyggere, og har et universitet som ble opprettet allerede i 1632.



Håvard Rues foredrag.

Den norske familiegrenen var godt representert og utgjorde en signifikant del av deltakerlista, inkludert hovedtaler Håvard Rue (som vi fortsatt betrakter som vår, selv om han har valgt finansiering og midlertidig hjemmebane i Saudi-Arabia).

Håvard tok turen for å fortelle om bayesiansk kvantil-regresjon for diskrete observasjoner. Jeg drister meg til å prøve å formulere temaet med litt andre ord: Håwards data er telledata – antall barn, for eksempel. Eller antall koner. Eller antall elskere. For statistikere betyr *diskret* imidlertid ikke «i smug» (*diskré*), slik elskere ofte er, men derimot atskilte tall: én, to eller tre barn, koner eller elskere, og han ønsker å finne ut hvilke variabler som henger sammen med tallene han har observert.

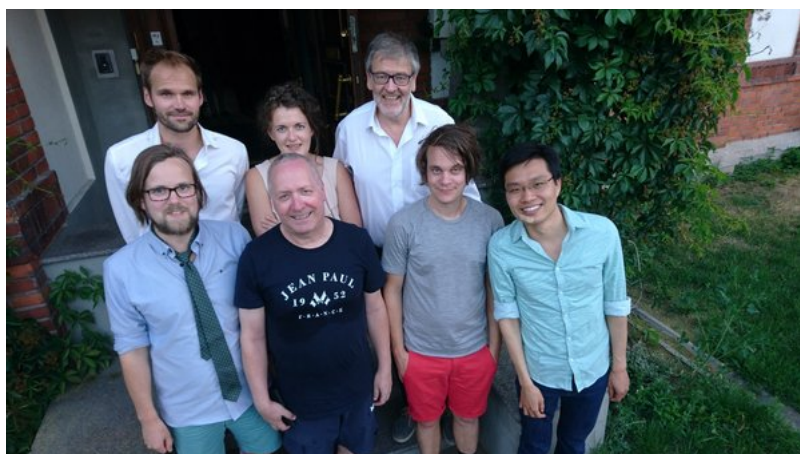
I motsetning til vanlige regresjonsanalyser som fokuserer på hvilke faktorer som kan forutsi gjennomsnittet for observasjonene (antall barn, koner, elskere), vil Håvard prøve å forutsi hva som gjør at de observerte tallene er spesielt høye eller spesielt lave.

Denne måten å beskrive kvantil-regresjon (med andre eksempler enn

dem Håvard brukte) er i utgangspunktet skrevet for lesere av Statistikk, men jeg velger å vise den til TG-lesere også, fordi:

1. Det er nyttig å prøve å forklare noe vanskelig med enkle ord, og
2. Jeg er slett ikke sikker på om dette er den beste, ei heller en tilstrekkelig god måte å oversette kvantil-regresjon til enkle ord.

Jeg oppfordrer derfor alle statistikkvenner til å søke opp kvantil-regresjon på egen hånd og finne en enda bedre forenklet beskrivelse, og deretter å sende den til [meg på mail](#) (du også, Håvard). Får jeg gode innspill, skriver jeg et innlegg om det i neste TG!



FocusStat-kjernefamilien.

Men NORDSTAT2018 hadde mer enn kvantil-regresjon å by på. Den sammensveisede kjernefamilie-forskningsgruppa FocuStat med hanefar Nils Lid Hjort i spissen var som vanlig sterkt til stede og presenterte en

rekke fokuserte nyvinninger.

Selv ble jeg ironisk nok så nervøs før mitt eget foredrag «How to build confidence» at jeg gikk glipp av Kjetil Røyslands inviterte kausalitetsforedrag, men ryktene sier at det gikk bra allikevel.

Jeg hadde stort utbytte av foredragene om statistikkundervisning, blant annet Inger Perssons Skype-overførte foredrag om [flippet klasserom](#). Dette er tema dere også gjerne kan søke opp og ta en titt på.

For min egen del fikk NORDSTAT2018 en ekstra dimensjon da jeg på

konferansens siste dag tok taxi til en lokal garngrossist, og presterte å møte en estisk, strikkende statistikerkollega der. Det var et møte av typen jeg vil karakterisere som et statistisk mirakel, og det kan dere lese om på bloggen.

Alt i alt var det en en svært hyggelig og vellykket konferanse. Om to år er [NORDSTAT i Tromsø](#). Skriv det opp i kalenderen!



«How to build confidence».

Meldingar

Nytt frå Noregs teknisk-naturvitskaplege universitet

Nye mastergradar

Kristian Aga

Modelling Neuronal Activity with Jittered Generalised Linear Models

Rettleiarar: Mette Langaas og Benjamin Dunn (Kavlisenteret)

Kristin B. Bakka

Changepoint model selection in Gaussian data by maximization

approximate Bayes Factors with the Pruned Exact Linear Time algorithm

Rettleiar: Mette Langaas

Anne-Line Evenstad Dahlen

Energy Efficient Monitoring using Kalman Filter – A case study on environmental variables collected from solar panels

Rettleiar: Jo Eidsvik

Ingunn Elisabeth Holen Dising

Pseudo-R² for Generalized Linear Models. A Study of Various Pseudo-R² for the Logit Model and the Log-Linear Poisson Model

Rettleiarar: Øyvind Bakke og Turid Follestad

Lene Finsveen

Time-series predictions with Recurrent Neural Networks

Rettleiar: Thiago G. Martins

Yngvild Hamre

Preserving projection properties when regular two-level designs are blocked

Rettleiar: John Tyssedal

Anine Harto

Predicting Recovery Rates of Defaulted Credit Card Accounts

Rettleiar: John Tyssedal

Pål Vegard Johnsen

Stochastic modelling and analysis of neural tumour evolution

Rettleiarar: Mette Langaas og Thea Bjørnland

Scott Macody Lund

Modelling and inference of cognitive parameters in jackdaws based on data from an investment problem

Rettleiar: Håkon Tjelmeland

Thibaut Malafosse

The use of extreme value statistics for calculating risk measures in finance

Rettleiar: Arvid Næss

Markus Mortensen

Tolerance Limit Communication Schemes and Nowcasting for Resource Constrained Sensor Networks (Utilizing NWP, Temporal and Spatial Dependencies for a Temperature Sensor Network at Spitsbergen)

Rettleiar: Ingelin Steinsland

Hans Olav Vogt Myklebust

Estimating the Value of Information Using Bayesian Optimization with Gaussian Process Surrogate Models – An Application to Failure Rates at Offshore Wind Farms

Rettleiarar: Jo Eidsvik og Iver Sperstad (Sintef Energy)

Kristian Amundsen Ruud

LIDAR Extended Object Tracking of a Maritime Vessel Using an Ellipsoidal Contour Model

Rettleiarar: Jo Eidsvik og Edmund Brekke (Institutt for kybernetikk, NTNU)

Runar Skagestad

Electricity Demand Forecasting with Gaussian Process Regression

Rettleiar: Bob O'Hara

Anders Sætherø

Backtesting counterparty credit exposure based on the Heath, Jarrow and Morton framework for simulation of interest rates

Rettleiar: Jacob Laading

Torfinn Tyvold

Short-Term Electric Load Forecasting Using Artificial Neural Networks

Rettleiar: Bob O'Hara

Markus Brabrand Urfjell

Pricing Asian Options by Numerical Path Integration with Advanced Lévy Dynamics

Rettleiar: Arvid Næss

Nytilsette ph.d.-kandidatar

Susan Anyosa

statistisk læring

Rettleiar: Jo Eidsvik

Michail Spitieris

i prosjektet «My Medical Digital Twin»

Rettleiar: Ingelin Steinsland

Jorge Armando Sicachá Parada

i prosjektet «Transforming Citizen Science for Biodiversity»

Rettleiar: Ingelin Steinsland

Pål Vegard Johnsen

statistisk læring med helsedata, samarbeid med SINTEF Digital

Rettleiar: Mette Langaas

Nytt frå Universitetet i Oslo

Nye mastergradar

Simon Boge Brant

Dynamic survival prediction for high-dimensional data

Rettleiar: Ørnulf Borgan

Jonas Christensen

Modelling of Total Claim Amounts with Regime Switching in the Framework of Generalized Cox Processes

Rettleiar: Frank Norbert Proske

Tristan Curteis

Focussed Model Selection for Longitudinal Data

Rettleiar: Ørnulf Borgan

John Dang

Internal models for reinsurance in Solvency II

Rettleiar: Erik Bølviken

Arman Færø Johnsen

Dødelighets- og Langtlivrisiko i Solvency II

Rettleiar: Erik Bølviken

Maria Kalitina

Reserve modelling in the presence of extreme insurance losses

Rettleiar: Ingrid Hobæk Haff

Håkon Nordtorp

Hvor gode er matching og volatility korreksjonene i Solvency II?

Rettleiar: Erik Bølviken

Wanjuan Ren

A Comparison of the Shrinkage Effects Between Boosting and Post-Selection Shrinkage Techniques

Rettleiar: Riccardo De Bin

Ingrid Emilie Sandmo

Optimal reinsurance. An examination of optimal stop-loss contracts for different lines of business

Rettleiar: Ingrid Hobæk Haff

Anniken Mjelde Sjø

Katastrofeforsikring i Solvency II: Evaluert mot en stokastisk modell

Rettleiar: Erik Bølviken

Roxicca Thirumeny

Simulation of Greeks of financial claims in Markets with Memory

Rettleiar: Frank N. Proske

Nytilsette

Li-Chun Zhang er ansatt som professor II ved Seksjon 2 (statistikk og biostatistikk).

Simon Boge Brandt og Riccardo Parviero er ansatt som stipendiater ved Seksjon 2. Michele Giordano og Alexander Lobbe ansatt er ansatt som stipendiater ved Seksjon 3 (stokastisk analyse, finans, forsikring og risiko).

Nytt frå Nytt fra Oslo senter for biostatistikk og epidemiologi

Nye ph.d.-grader

Rune Hoff

Causal inference in event history analysis with complex longitudinal data (prøveforelesning: Multi-state models vs. joint models of longitudinal and time-to-event: Pros and cons)

Hovudrettleiar: Jon Michael Gran, Avdeling for biostatistikk, UiO
Komité:

Research Statistician Brian Tom, Institute of Public Health, University of Cambridge

Dr. Jutta Gampe, Laboratory of Statistical Demography, Max Planck Institute for Demographic Research

Professor Tron Moger, Institutt for helse og samfunn, Det medisinske fakultet, Universitetet i Oslo

Disputas: 27. september 2018

Nytilsette

Valeria Vitelli og Jon Michael Gran er ansatt som førsteamenuensis.

Jesse Hemerik, Mats Julius Stensrud og Rune Hoff er ansatt som postdok-stipendiater.

Juri Kuronen, Leiv Rønneberg, Kusti Skytén og Niklas Maltzahn er ansatt som ph.d.-stipendiater.

Pusleri

Løsning på pusleri nummer 49

Jostein Kåre Lillestøl

Oppgaven var

A og B skal hver for seg gjennomføre rettferdige myntkast. B får n kast, mens A får lov til å ta et kast mer, dvs. $n + 1$ kast. Hva er sannsynligheten for at A får flere kron enn B etter de $2n + 1$ kastene?

Hint: Det fins en svært enkel løsning.

Er sannsynligheten for at B får flere kron enn A like lett å finne?

Tilleggsspørsmål: Generaliser til at A får ta m kast mer enn B, der $m > 1$.

Løsning

En mulighet er å skrive ned den simultane sannsynlighet for antallene kron ($X_A = i, X_B = j$), og så summere en dobbeltsum over alle $i > j$. Et alternativ er å tenke seg at A først gjør n kast, og så et ekstra kast, og så regne ut den søkte sannsynlighet ved å betinge med hensyn på

resultatet av de n første kastene.

Utrekningene gir begge etter litt kombinatorikk at den søkte sannsynlighet er $\frac{1}{2}$, og derfor noe overaskende ikke avhenger av n . Det bør derfor finnes en enklere løsning.

Her er den enkleste: Begivenhetene «A flere kron enn B» og «A flere mynt enn B» må, på grunn av rettfærdige kast og symmetri, være like sannsynlige. I tilfellet der den ene tar ett kast mer enn den andre, og bare da, er de to begivenhetene disjunkte og uttømmende, slik at hver må ha sannsynlighet $P_A = \frac{1}{2}$.

Sannsynligheten for at B får flere kron enn A blir imidlertid noe mer komplisert:

$$P_B = \frac{1}{2} - \binom{2n+1}{n} \left(\frac{1}{2}\right)^{2n+1}.$$

Dette kan finnes av $P_B = 1 - (P_A + P_C)$, der P_C er sannsynligheten for at A og B har like mange kron. Denne kan beregnes ved betinging mhp. antallet kron hver:

$$P_C = \sum_{i=0}^n \binom{n+1}{i} \left(\frac{1}{2}\right)^{n+1} \binom{n}{i} \left(\frac{1}{2}\right)^{2n} = \binom{2n+1}{n} \left(\frac{1}{2}\right)^{2n+1}.$$

Generaliseringen der A tar $n + m$ kast, løses trolig enklest ved betinging og rekursjon:

La $P_A(m)$ være sannsynligheten for «A får flere kron enn B» i situasjoner der A har m flere kast enn B. Anta at B har gjort sine n kast. For at A skal ha flere kron enn B etter sine $n + m$ kast, må dette enten være oppnådd etter A sine $n + m - 1$ første kast eller så må det være likhet etter $n + m - 1$ kast og en ny kron i A sitt neste kast. Dette gir for $m \geq 1$ rekursjonen

$$P_A(m) = P_A(m-1) + \binom{2n+m-1}{n} \left(\frac{1}{2}\right)^{2n+m}, \text{ der}$$

$$P_A(0) = \frac{1}{2} - \binom{2n}{n} \left(\frac{1}{2}\right)^{2n+1},$$

der vi også har benyttet sannsynligheten for likhet (og symmetri) i beregningen av $P_A(0)$. Herav følger for $m \geq 1$ at

$$P_A(m) = \frac{1}{2} + \left(\frac{1}{2}\right)^{2n+1} \sum_{i=1}^{m-1} \binom{2n+i}{n} \left(\frac{1}{2}\right)^i.$$

Pusleri nummer 50

Jostein Kåre Lillestøl

Per og Pål har kranget, og som straff har de fått husarrest for resten av kvelden. Begge ønsker imidlertid å gå på Brann-kamp, og går inn i forhandlinger med moren. Hun foreslår følgende:

Her har dere hver deres mynt. Gå til hvert deres rom, og bli der uten å ha kontakt med hverandre. Om 10 minutter kommer dere ned i stuen, enten med mynten eller lar den være igjen på rommet. Eventuelle mynter dere har med, blir knipset, og ved alt annet enn (bare) kron må dere forbli hjemme, som altså også vil skje dersom begge lar sin mynt bli igjen på rommet.

Hva er sannsynligheten for at Per og Pål kommer seg på kamp, gitt at begge gjør det optimale valg?

Send forslag til løsning til <Jostein.Lillestol@nhh.no>.

