ISSN 0803-8953

TILFELDIG GANG
Nr. 4, årgang 17
Desember 2000

Utgitt av Norsk Statistisk Forening

INNHOLD

	Side
Redaksjon:	1
Fra redaktøren. *Odd Aalen*	2
Fra lederen i NSF. *Geir Egil Eide*	4
Statistisk arbeidsliv	
Stokastisk modellering og statistikk i populasjonsbiologi. *Steinar Engen*	5
En forskningssjefs hverdag ved Norsk Regnesentral. *André Teigland*	8
Puslerispalten	
Pusleri nr. 5. Løsning av Pusleri nr. 4. *Jostein Lillestøl*	11
Helsefakta? *Redaktøren*	14
Milepæl	
Gaute Vik 50 år 26. juni 2000.	15
Undervisningsseminar	
Sammenkomst om statistikk-undervisning i Kristiansand. *Henrik Dahl*	16
Videreutdanning i statistikk for lærere. *Elinor Ytterstad*	17
Statistikk på økonomisk-administrative studier ved statlige høgskoler. *Tore Wentzel-Larsen*	18
Matematikk, statistikk og metode ved Siviløkonomutdanningen i Bodø. *Rolf Volden*	20
Bruk av statistisk programvare i elementærundervisningen: Må vi revidere kursoppleggene? *Ivar Heuch*	22
Nye læreplaner i sannsynlighetsregning og statistikk i videregående skole. *Ørnulf Borgan*	27
Eksamensoppgaver fra videregående skole, våren 2000.	30
Gjennomslag for statistikkurs på nett. *Klipp fra Uniforum*	33
Fra fagets historie	
Hva kan vi slutte av statistikk? Tredje foredrag. *Erling Sverdrup*	34
Faglig stoff	
Malignant melanoma, lung cancer, and prostate cancer in relation to diet. *Marit Bragelien Veierød*	45
Meldinger	
Her er sportsnytt! *Nils Lid hjort*	46
Nytt fra Norges Landbrukshøgskole.	46

Redaksjon:

Odd Aalen (redaktør)
Grete Usterud Fenstad (redaksjonssekretær)
Ingrid Glad
Nils Lid Hjort
Trond Anders Moger

Redaksjonens adresse:
Odd Aalen
Seksjon for medisinsk statistikk
Universitetet i Oslo
P.b. 1122 Blindern, 0317 Oslo

Tlf: 22 85 12 36
Fax: 22 85 13 13
e-mail: odd.aalen@basalmed.uio.no

Trykking og utsending av TG utføres av
Statistisk sentralbyrå

Bidrag til Tilfeldig Gang:

Alle slags bidrag til bladet er velkomne, enten det gjelder nye temaer av interesse for norske statistikere, eller man har kommentarer til innlegg som allerede har stått i bladet. Stoff til neste nummer må være redaksjonen i hende innen 1.2.2001.

Det greieste er at stoffet sendes inn som Word-fil (Word 97 eller lavere versjon) til redaktørens e-mail-adresse, se ovenfor. Alternativt kan vi ta en ren tekst-fil, men vi vil helst ikke ha TEX-format med mindre dette er nødvendig på grunn av matematikken.

Vi har etablert en nettversjon av TG, se følgende internettadresse:
www.math.uio.no/nsf/web/index.htm . De som leverer bidrag til TG og som ønsker å reservere seg mot at bidraget deres legges ut på nettet, må gi beskjed om dette ved innsendelse av stoff.

Fra redaktøren:

Så er det atter jul. Det virker som det var i går at jeg satt og skrev lederen i fjorårets siste TG-nummer. Det er kanskje en klisjé at tiden går raskere og raskere, men det er også en utvilsom sannhet.

Flere artikler i dette nummeret er viet undervisning. Bakgrunnen er et undervisningsseminar som ble arrangert i Kristiansand i sommer av Henrik Dahl, som er tilknyttet Høgskolen i Agder. Henrik har alltid hatt et sterkt engasjement i undervisning, og som det fremgår av artiklene er det mange som deler dette engasjementet.

En grundig gjennomgåelse av læreplanene i statistikk i videregående skole presenteres av Ørnulf Borgan. Ørnulf har forøvrig blitt lærebokforfatter, i det han inngår i forfatterteamet Erstad, Heir, Bjørnsgård, Borgan og Pålsgård. Dette læreverket i matematikk for videregående skole tok i år rundt 50% av markedet i 1. klasse, en meget betydelig oppgang fra året før. Uttalelser til denne redaktør fra lektorer i videregående skole indikerer at en årsak til dette var at en profesjonell statistiker var med som forfatter. Endelig hadde man tillit til at statistikkstoffet holdt faglig kvalitet etter alt rotet som det tidligere har vært rundt dette i enkelte lærebøker. Utvilsomt en seier for vårt fag.

Et hovedproblem for undervisningen i videregående skole er å utdanne lærerne (lektorene) i statistikk slik at de kan gjennomføre undervisningen med et minimum av innsikt. Jeg vil her fremheve det spennende initiativet fra Tromsø, som er beskrevet av Elinor Ytterstad i dette nummeret. Her satses det på fjernundervisning, men enkelte samlinger av deltagerne. Utgangspunktet er kurs som allerede gis i den ordinære undervisningen ved universitetet. Dette inititativet er utvilsomt et eksempel til etterfølgelse.

I det hele tatt er fjernundervisning i ferd med å bli noe helt nytt med dagens internett-teknologi. Det kanskje mest avanserte fjernundervisnings-tilbudet innen statististikk her i landet er det som gis av Petter Laake og Harald Fekjær ved Seksjon for medisinsk statistikk, Universitetet i Oslo. Dette har et meget profesjonelt internett-opplegg. Kurset er rettet mot medisinere, men tilsvarende opplegg skulle være like godt egnet til andre grupper, f.eks. lærere. Lærerne og legene har det til felles at de sitter spredt over hele landet.

Et annet tema på seminaret var økonomisk-administrative studier ved høgskolene. Det er her folkets nye helter får sin utdannelse, og det har blitt et noe omstridt spørsmål hvor mye statistikk man egentlig trenger for å følge aksjekursene, kjøpe (eller motta gratis!) opsjoner, lede bedrifter etc. Tore Wentzel-Larsen og Rolf Volden har fyldige innlegg om saken i dette nummeret.

Nok et høyst aktuelt tema tas opp av Ivar Heuch i en grundig artikkel, nemlig relasjonen mellom vår undervisning og statistisk programvare. Hvordan integrerer vi

programvaren i vår undervisning? Ivar peker på de store forandringer som skjer på dette feltet. Hvordan skal vi møte dem?

Ellers fortsetter vi vår serie om statistiske arbeidsplasser. André Teigland forteller, på en meget engasjerende måte, om sin virksomhet som forskningssjef ved Norsk Regnesentral. André har virkelig skjønt poenget med denne serien, han nytter sjansen til å drive entusiastisk PR for Norsk regnesentral som arbeidsplass. Vi mottar gjerne innlegg av denne typen fra andre hold også, litt spirituelle innlegg om arbeid og fag, med vekt på hvorfor det man driver med er *morsomt*!

Steinar Engen forteller om sin spennende virksomhet som populasjonsbiolog. Steinar nevner blant annet det veivalg han hadde, skulle han nærme seg fagstatistikerne eller biologene. Han valgte det siste, og jeg er personlig ikke i tvil om at han valgte rett. Er det noe vårt fag trenger, så er det friskt blod fra andre fag. Det bør i økende grad bli normen at statistikere har en tilleggs- (eller hoved-) utdannelse i et substansfag, eller at man senere søker omfattende samarbeid utenfor statistikkfaget. Er det noe populasjonsbiologer vet, så er det at innavl er usunt.

Steinar nevner R. A. Fisher. Med sine fundamentale idéer (likelihood, randomisering, variansanalyse etc etc) er det ikke tvil om at Fisher må oppfattes som grunnleggeren av statistikk-faget. Det mange statistikere ikke vet, er imidlertid at Fisher, som Steinar sier, også oppfattes som biolog, og faktisk som en meget betydelig person på dette fagfeltet. Fishers bok fra 1930: The Genetical Theory Of Natural Selection, regnes for et av de helt grunnleggende verker innen moderne evolusjonsbiologi, og bidro avgjørende til den moderne syntesen mellom Darwinismen og Mendelsk genetikk. Fishers navn ble sterkt fremhevet da den berømte evolusjonsbiolog Bill Hamilton, som regnet seg som en student av Fisher, døde tidligere i år.

Fisher utledet blant annet en diffusjonslikning for hvordan genfrekvensen endrer seg i en populasjon. Dette var før Kolmogorovs fundamentale verk om Markov- og diffusjonsprosesser, og det har vært hevdet at Fisher også kan ansees som en grunnlegger av denne teorien. Man skal forøvrig merke seg at den berømte astronom Sir Fred Hoyle i år utga en bok der han viderefører statistiske beregninger rundt evolusjon, basert nettopp på Fishers diffusjonsteori. Hoyles klare og kontroversielle konklusjon er at nydarwinismen ikke henger matematisk på greip. Naturlig seleksjon klarer ikke jobben med å skape den utviklingen som har skjedd. Så statistikk og sannsynlighetsteori er stadig i sentrum av fundamentale og uavklarte spørsmål.

Beste hilsen Odd Aalen

Fra lederen i NSF:

Kjære medlemmer.

På foreningsfronten har det vært relativt stille denne høsten. Styret har ikke hatt noen spesielle grunner til å møtes, men tar sikte på et møte på nyåret. Noen medlemmer savnet kanskje revidert regnskap for 1999 som var lovet til forrige TG. Dessverre kan vi ikke presentere dette denne gangen heller. Vi venter fortsatt på den endelige revisjonsrapporten, men kan forsikre at økonomien er god. Grimstad-møtet gikk med overraskende overskudd og jeg forventer at vi kan opprettholde ordningen med reisestipend for hovedfagstudenter også til neste års norske statistikermøte i Bergen. Annonsering av stipendier vil isåfall finnes i neste TG. Det er fint om aktuelle veiledere allerede nå kan informere sine studenter om mulighetene et studentmedlemskap i NSF kan gi.

For øvrig blir dette TG årets julenummer så jeg vil avslutte med å ønske dere alle en riktig god julehøytid og et faglig fruktbart nytt år!

Vennlig hilsen

Geir Egil Eide



TG's redaksjon ønsker alle lesere en

God Jul
og et
Godt Nytt År!

Stokastisk modellering og statistikk i populasjonsbiologi

Steinar Engen
Institutt for matematiske fag, NTNU

Personlig bakgrunn

I 1970 var jeg den første studenten ved NTH som tok diplomoppgave i statistikk. Min bakgrunn kunne ikke ha vært mer enn mellom 5 og 10 vekttall i statistikk, så Arnljot Høyland virket ganske usikker på om dette lot seg gjennomføre. Mitt første møte med biologifaget skjedde da jeg i 1971 ble sivilarbeider ved Vitenskapsmuseets zoologiske avdeling, hvor jeg etter hvert fikk en slags heltestatus fordi jeg kunne utføre en T-test og skrive ALGOL-programmer. Det hele endte opp med at senere miljøværnminister Olav Gjærevoll fikk dunket gjennom i forskningsrådet at jeg skulle ha et stipend. Det var behov for statistisk kompetanse i norsk biologisk forskning. Med god hjelp fra Høyland havnet jeg i Oxford ved Department of Biomathematics, der jeg jobbet under veiledning av M.S.Bartlett og M.G.Bulmer, og tok doktorgraden i 1975.

Da jeg fikk fast jobb som statistiker ved Universitetet i Trondheim, ved det som dengang het Norges Lærerhøgskole, sto jeg overfor et veivalg. Jeg kunne enten forsøke rent faglig å nærme meg Blindern-statistikerne, eller styrke min biologiske kompetanse via samarbeid med biologiske miljøer. Jeg valgte det siste, og er idag kanskje mer biolog enn statistiker og har publisert artikler med bortimot 50 forskjellige biologer. I en periode hadde jeg tett samarbeid med Nils Chr. Stenseth ved UiO, som lærte meg mye om populasjonsdynamikk, økologi og evolusjon. Idag arbeider jeg hovedsaklig med Bernt-Erik Sæther ved NTNU og våre felles internasjonale samarbeidsparnere med generelle problemstillinger innen populasjonsdynamikk, med spesiell vekt på bevaringsbiologi. Jeg har også bistilling som rådgiver ved Norsk Institutt for Naturforskning.

Modellering i populasjonsdynamikken

Økologi og populasjonsdynamikk er lite eksakte vitenskaper. Enhver naturlig populasjon er påvirket av en rekke fysiske og biologiske faktorer via kompliserte vekselvirkninger. Da de store datamaskinene gjorde sitt inntog tidlig på 70-tallet, var det mange som trodde at vi kunne oppnå den fullstendige forståelse for sammenhengene i naturen bare vi hadde nok data, store nok maskiner og programmer, og modeller med mange nok parametre. Men naturen oppfører seg stokastisk, uansett hvor detaljert man forsøker å beskrive den, så denne angrepsvinkelen, som delvis var basert på uforstand m.h.t. håndtering av stokastiske fenomener, var dømt til å mislykkes. Resultatene ligger begravet blant andre ikke-reproduserbare vitenskapelige oppdagelser.

Den forståelsen man idag har opparbeidet om økologiske sammenhenger og dynamikken i naturen kan sies å være basert på den motsatte tilnærmingsmåten. Via et nettverk av forenklede modeller, der man teoretisk studerer noen få sammenhenger, og samler alle andre faktorer i stokastiske ledd, bygger man opp en generell forståelse for sammenhengene i naturen. Stokastiske simuleringer brukes ofte for å studere modellenes egenskaper, selv om man tilstreber a finne generelle analytiske resultater med mer tyngde og anvendbarhet.

Individet er som regel den grunnleggende byggeklossen i modellene, og modeller kan grovt klassifiseres etter hvilke individer, og hvilke egenskaper ved individene man velger å beskrive. En stor del av teorien behandler enkle modeller, der alle individer er like slik at populasjonsstørrelsen er den eneste variable av interesse. Dette kan likevel gi interessante stokastiske modeller for studier av populasjonssvingninger, tetthetsregulerende effekter, og muligheten for at populasjonen dør ut.

Genetiske modeller er en annen stor og viktig klasse av modeller. Genetisk egenskaper er da individ-merkelapper som brukes i modelleringen. I motsetning til de enkle modellene nevnt ovenfor der endringene i populasjonsstørrelse står sentralt, så er de fleste populasjonsgenetiske modeller basert på den grove forenklingen at populasjonsstørrelsene faktisk er konstante. Det er genfrekvensene og endringene i disse over tid som settes under lupen.

For mange arter vil det være urealistisk å basere modeller på at individene er like. Spesielt vil individenes egenskaper, som evnen til å produsere avkom og til å overleve, ofte være sterkt avhengig av alder. Dette gir opphavet til såkalte aldersstrukturerte modeller, der individenes alder brukes i modellen, vanligvis ved å samle individer født i samme sesong i diskrete aldersklasser. Stokastiske aldersstrukturerte modeller blir fort kompliserte, og det gjenstår mye forskning før man kan si at betydningen av aldersstruktur er grundig forstått.

Et forskningsfelt der det idag skjer stor utvikling, er romlige populasjonsdynamiske studier. Individene gis da også en merkelapp som sier hvor de befinner seg. For å gjøre slike modeller realistiske må støyleddene ved de forskjellige lokalitetene være avhengige, migrasjon mellom lokaliteter må modelleres, og det må tas hensyn til forskjellige former for romlig heterogenitet mellom leveområdene. Ofte kan man konstruere realistiske romlige modeller der individenes posisjoner er en endelig diskret mengde. Dette er ikke bare en teoretisering man gjør for å slippe å arbeide med mer kompliserte kontinuerlige modeller. En av de viktigste praktiske problemstillingene er å studere effekter av menneskelige aktiviteter, og da kan spesielt fragmentering av dyrenes leveområder i diskrete delområder ved bygging av veier og bygningsmasse være en vesentlig trussel mot en bestands levedyktighet. Slike modeller kalles ofte metapopulasjonsmodeller, eller øymodeller. Dynamiske modeller i kontinuerlig rom er også et felt der der skjer nyere forskning.

Samspillet mellom arter studeres oftest ved å modellere to arter samtidig. De to hovedtypene av modeller er konkurranse-modeller, der artene konkurrerer om næring eller territorier, eller predator-byttedyr-modeller, hvor den ene arten er en del av næringsgrunnlaget til den andre. Vanligvis er slike modeller enkle i den forstand at det ikke tas hensyn til aldersstruktur, genetikk eller romlige koordinater. Men også her skjer det en del ny forskning på mer sammensatte modell-typer.

I studier av dyresamfunn bestående av et stort antall arter forsøker man ofte å finne modeller som gir realistisk beskrivelse av hvordan individene er fordelt mellom de forskjellige artene. Et nærliggende statistisk problem som mange har vist stor interesse for opp gjennom årene er å estimere antall arter i et samfunn. Det mest ekstreme eksempelet er å estimere antall arter på jorda. Estimatet er pr. idag temmelig usikkert. Antagelig er tallet av størrelsesorden 5-15 millioner. Et stort antall av disse forsvinner fra jordens overflate hvert år på grunn av forskjellige menneskelige aktiviteter.

I tillegg til fragmentering av leveområder er forurensning men medfølgende klimaendringer og høsting av populasjoner (f. eks. fiske) menneskelige inngrep man er interessert i å studere. Søking etter langsiktige optimale høstingsstrategier gir opphav til mange interessante matematiske, sannsynlighetsteoretiske og statistiske problemstillinger.

Statistikk og populasjonsdynamikk

Historisk sett er statistikkfaget nær knyttet til biologi. Blant biologer regnes for eksempel R.A.Fisher for å være biolog, og det er velkjent at begreper som korrelasjon, regresjons- og variansanalyse og forsøksplanlegging har sine utspring i biologiske problemstillinger. Fisher skrev også en rekke artikler innen genetikk.

Idag tror jeg det er riktig å si at det ikke finnes noen spesiell statistikk for biologer. Når modeller skal tilpasses til observasjoner er alle vanlige statistiske problemstillinger som estimering, usikkerhetsberegninger, testing og prediksjon aktuelle. Det er modellene man tilpasser som er spesielle, ikke statistikken. En nyutdannet statistiker som skal arbeide med problemer innen populasjonsbiologi vil være godt skolert når det gjelder statistikk, men må regne med å måtte sette seg inn i mye nytt stoff omkring biologisk tenkemåte, biologiske teorier og biologisk modellering.

Blant biologene har det skjedd en rivende utvikling i bruken av statistikk, noe som er en kombinasjon av datateknologien og tilgjengeligheten av lett anvendbar statistisk programvare, forbedring av utdanningsnivået i statistikk blant biologer, og behovet for analyse av stadig mer innsamlet datamatertiale. Bruken av statistikk som refereres i biologiske tidsskrifter blir stadig mer avansert. For eksempel er Bayesianske metoder og spesielt MCMC på full fart inn i biologisk forskning, og man ser stadig oftere at artikler i internasjonale tidsskrifter er samarbeidsprosjekter der personer med forskjellig bakgrunn deltar. Men, det forekommer også overtramp som skyldes at tilgjengeligheten av programvare kan bli for stor , både i forhold til statistisk kunnskap og kunnskapsnivået innen biologisk modellering.

Forskningsinstitusjoner og finansieringskilder

I tillegg til at universitetene, inkludert virksomheten på Svalbard, alle har relativt store biologiske miljøer som dekker mange forskjellige biologiske disipliner langt utenfor det som er beskrevet ovenfor, skjer det mye biologisk forskning ved forskningsinstitutter. De viktigste institusjonene er Havforskningsinstituttet, Norsk Institutt for Naturforskning, Norsk Polarforskningsinstituttet, Norsk Institutt for Skogforskning, og Norsk Institutt for Vannforskning. Mye av arbeidet som drives her er kontinuerlig datainnsamling som brukes til overvåkning av vårt dyre- og planteliv. Men institusjonene bidrar også med forskning og publisering på høyt internasjonalt nivå.

Norges Forskningsråd bevilger midler til biologisk forskning gjennom en rekke forskjellige forskningsprogrammer under områdene Naturvitenskap og Teknologi, Bioproduksjon og Foredling, og Miljø og Utvikling.

En forskningssjefs hverdag ved Norsk Regnesentral...

André Teigland

Kjære TG-leser!

"Kan du ikke skrive noen ord om ditt arbeid og din hverdag, gjerne med en personlig vri", lød mailen fra redaktør Odd. Jo, tenkte jeg, det kan jeg jo gjerne. Og så sitter jeg her og finner ut at da må jeg visst først finne ut hva jeg egentlig gjør...

Først noen ord om min bakgrunn: Jeg tok hovedfag i statistikk ved UiO i 1992, og ynder fortsatt å tro at jeg er temmelig ung. Med Erik Bølviken som veileder var kalmanfiltrering, dynamisk lineære modeller, metropolis, annealing, fortran, MCMC og Splus riktig fornøyelig! Etter endt hovedfag begynte jeg ved Norsk Regnesentral (NR). Jeg jobbet to år som forsker i statistikk før jeg fikk prøve meg som forskningssjefsvikar noen måneder. Å lede og styre ga mersmak, og i 1995 ble jeg forskningssjef for NRs bildeanalysemiljø. I 1998 slo vi sammen statistikk- og bildeanalysegruppene til en stor avdeling, som jeg har ledet siden. Vi teller i dag 26 forskere, og lyder det velklingende navnet SAMBA! (Statistisk Analyse, Mønstergjenkjenning og BildeAnalyse)

NR er et oppdragsbasert forskningsinstitutt. Vi får ca 10% av inntektene våre i mer eller mindre frie, grunnleggende midler fra Norges forskningsråd. Resten skaffes til veie gjennom konkrete oppdrag for kunder som skal ha løst et problem. Total omsetning er rundt 70 millioner. Uten inntekter, ingen jobb, så det blir mye økonomiprat ved siden av fag på et sted som NR. Instituttet står på to ben; statistikk og IT. Vi besitter det største oppdragsforskningsmiljø innen anvendt statistikk i Norge. Sammen med SAND (avdelingen som jobber med geostatistikk for petroleumsbransjen) er vi 40 forskere i og rundt statistikkfaget. I tillegg har NR rundt 40 rene IT-forskere.

For meg er det to virkelig store gleder ved å jobbe ved NR; 1) det sitter utrolig dyktige fagfolk rundt omkring på hvert eneste kontor! (usedvanlig hyggelige er de også!) og 2) jeg kommer hele tiden bort i ulike fagfelt, stadig nye oppdragsgivere, og spennende, flinke mennesker "der ute i den virkelige verden". Det er ikke rent få av TGs lesere og tidligere studievenner jeg har truffet igjen som oppdragsgivere eller samarbeidspartnere. Som anvendte statistikere, på oppdrag for kunder som er villige til å betale for oss, må vi hele tiden lære oss noe nytt. Vi må forstå (ihvertfall en del av) oppdragsgivers fagfelt, vi må skjønne hvor dataene stammer fra, hvordan de er samlet inn, vi må oversette kundens språk til matematikk og statistikk og vi må kunne implementere løsninger for dem.

Jeg har overhørt og deltatt i en del diskusjoner angående statistikerens rolle de senere årene. Er vi "bare" underdanige tjenere når vi arbeider anvendt? Ærlig talt forstår jeg ikke problemstillingen. JA, det er kunden/brukeren/oppdragsgiveren som besitter problemet, og ja, jeg opplever oss på NR som oppdragsgivers "tjener", men, nei, vi er aldeles ikke underdanige! Kunden er avhengig av vår kompetanse, og i de aller fleste prosjekter har vi et veldig godt samarbeid med våre kunder med gjensidig respekt for hverandres kompetanse og roller. For meg er en statistiker nyttig først når hun er til nytte...

Hva jobber vi så med? Svaret må bli "omtrent alt" der statistikkompetanse kreves. I SAMBA har vi anvendelser innen bransjer som industri, bank, forsikring, transport, logistikk, handel, offshore, internett og medisin. Vi lager prisprognoser, predikerer trafikkvolum, styrer smelteovner, beregner forsikringspremier, estimerer risikoavsetninger, deltar på utvalgsundersøkelser, hjelper leger osv. Kunsten (ja - det ER en kunst) er å beherske, eller vite hvem på nabokontorene som behersker, den metodikk som vil være mest mulig velegnet for å løse det konkrete problemet vi står overfor. I tillegg har vi egne spesielle satsingsområder i SAMBA: Satellittfjernmåling, bildeanalyse, mønstergjenkjenning og data mining, miljøstatistikk og marin ressursestimering. Rundt 100 store og små bedrifter og institusjoner står på vår kundeliste for de siste 3 årene.

Hverdagen på NR er preget av stor frihet under stort ansvar. Ingen kontrollerer når den enkelte forsker kommer og går. Ingen sjekker at artikkelen du ønsker å lese er relevant. Men, med arbeid for kunder, så er det hele tiden et eksternt press på at ting skal være ferdig til avtalt tid og til avtalt pris. Avtalt pris betyr for oss innen gitte timerammer. Prosjektene varierer veldig i størrelse og utstrekning. Alt fra noen ukesverk opp til flere års arbeid med mange personer involvert. De fleste arbeider med 1-2 prosjekter av gangen, og vi har som fast regel at ingen prosjekter skal utføres av bare én person. Rollen for en forsker ved NR utvikler seg over tid, fra at man i starten er mere en "ren utfører" til at man etterhvert også blir idégenerator, prosjektleder, og har mere utadrettet kontakt med oppdragsgivere. Enkelte har også et utvidet markeds-, fag- og strategisk ansvar for sine felt, i samarbeid med meg som forskningssjef. Heldigvis er NRs økonomi meget solid, noe som gjør at vi også har anledning til å drive mere egeninitiert "akademisk forskning". Vi publiserer en god del, vi deltar på konferanser både som foredragsholdere og for å lære mer, og vi er etterspurt i uformelle sammenhenger i bedrifter og bransjefora.

Hva gjør så en forskningssjef opp i alt dette? Vel, tidligere var en forskningssjef ved NR nettopp det tittelen skulle tilsi, en sjef (med stor S) for forskningen (med stor F). I dag har vi organisert oss i større enheter, og det er komplett umulig for én person å være en overordnet guru for så mye forskjellig som SAMBA bedriver. En forskningssjef ved NR er avhengig av å ha en solid fagkompetanse, men må sørge for å ansette folk som er dyktigere enn seg selv innen hvert eneste felt vi skal jobbe med! Konkret har jeg ansvar for mange biter som: strategi, planlegging, økonomi, markedsføring og profilering, forhandlinger med kundene, prosjektbemanning, ansettelser, samt personalansvar og overordnet prosjektansvar. Ikke minst har jeg det avgjørende ordet når deg gjelder hvilke felt vi skal bygge opp og hvilke vi skal bygge ned. Mye å holde styr på, men i praksis går det bra takket være dyktige medarbeidere det går an å delegere myndighet og ansvar til ☺. Som forskningssjef er jeg også en av 7 personer i NRs ledergruppe. Arbeidet der kan være både utfordrende og spennende. Man må kunne samarbeide (og samtidig sloss litt for sine folk..), man må takle fagforeninger, takle innspill fra et styre (uten statistikkunnskaper), og ikke minst kommer man en gang i året i den merkverdige situasjon at man skal forhandle lønn med sine gode kolleger mens man sitter på "den andre siden av bordet".

Min typiske hverdag? Hektisk. Først lese nattens mail - noe viktig skjedd?, en rask kopp kaffe, hei - god morgen, stikke hodet innom et par kontorer for å sjekke status på noen av våre 54 løpende prosjekter, forberede møte med mulig ny kunde klokken 10,

avbrutt av adm. dir. som må ha innspill til neste styremøte NÅ, svar på en haste-mail, en forsker kommer innom - ja, ok, han kan ta fri på mandag, telefonen ringer - ja, visst, det kan vi sikker bidra med - møte avtalt på tirsdag, skrive ut to foiler, kundemøte - lag prosjektforslag - max 1 million, lunch, ny telefon, maildiskusjon i ledelsen om neste års budsjett - nei, så mye vil jeg ikke påta meg at avdelingen skal tjene inn, diskusjon med en prosjektleder om mulig budsjettsprekk, telefonforhandling med kunde som synes vi er for dyre, avbrutt av adm.dir igjen, kaste meg over et av mine egne prosjekter (ja - jeg får brukt 10-20% av tiden min til vanlig prosjektarbeid), ny forsker innom - hvordan blir det med det store prosjektet? Har jeg glemt at hun skal ha en ukes ferie i desember?, gjennomgang av markedsoversikt for neste år, skrive et tilbud, attestere to regninger, sjekke innkommende mailer - ikke noe viktig denne gangen, jobbe litt på prosjekt igjen, planlegge morgendagens avdelingsmøte. Vips, så er enda en arbeidsdag over...
Alle muligheter for å bli stresset, men fryktelig gøy. Kjeder meg gjør jeg aldri! Og et par ting har jeg lært; 1) Går det dårlig med noe, så går det sikkert bra med noe annet - verden består, og 2) Rydd sakene unna der og da - ta heller en gal avgjørelse fra tid til annen enn å tenke så lenge at avgjørelsene bare hoper seg opp!

Konklusjonen etter å ha vært på NR noen år:
Er du en dyktig statistiker som vil ha en spennende hverdag med flinke mennesker rundt deg - søk jobb ved NR! (se annonse annet sted i bladet... ☺).
Vil du forbli fingerspissfagstatistikker, så påta deg aldri en forskningssjefsjobb...
På den annen side er det få forunt å få en såpass helhetlig oversikt over anvendelsesmuligheter for faget vårt, som jeg har fått anledning til å opparbeide meg. Og er du opptatt av strategi og ledelse, ønsker du å bidra til at dyktige mennesker jobber sammen og trekker i samme retning, ja da tror jeg ikke du kan få en mere utfordrende, spennende og lærerik jobb enn min!

Beste hilsner fra
André

andre.teigland@nr.no
www.nr.no

Puslerispalten

Pusleri nr. 5

Her er tre myntkastproblemer:

1. Hvordan kan du generere 50:50 sjanse med myntkast med en skjev mynt?

2. Hva er sannsynligheten for at KK forekommer i n rettferdige myntkast?

3. Hva er sannsynligheten for at KMK forekommer i n rettferdige myntkast?

Problemstillingen kan sikkert belyses på ulike måter, og *også denne gangen ønsker vi gode løsninger fra lesere for presentasjon i neste TG*, fra den enkle og intuitive til den matematisk sofistikerte under ulike antakelser. Send løsninger innen 4 uker fra mottatt TG til jostein.lillestol@nhh.no.

Jostein Lillestøl

Løsning Pusleri nr. 4

Vi minner om oppgaven:
Auksjoner kan foregå på flere måter, bl.a.

(1) Bud høynes inntil ingen flere bud kommer. Siste budgiver betaler budt beløp ("First prize - open cry", kalt engelsk auksjon).
(2) Alle budgivere byr i lukket konvolutt. Høyeste budgiver vinner og betaler budt beløp ("Sealed bid - first prize").
(3) Auksjonarius setter pris over høyest tenkelige og senker priser inntil første budgiver slår til (kalt hollandsk auksjon).
(4) Som (2), men høyeste budgiver betaler nest høyeste bud. ("Sealed bid - second prize").

Det er nokså lett å innse at (2) og (3) er ekvivalente. Vis (under rimelige forutsetninger) at (1) og (4) er ekvivalente.

Aha! I valget mellom de to vanlige auksjonstypene (1) og (2), må derfor (2) være bedre for selger enn (1), fordi (2) gir det høyeste lukkede bud, mens (1) gir beløp ekvivalent med nest høyeste lukkede bud. Sant eller galt?

Hint: Hvis du ikke greier å avgjøre generelt det siste spørsmålet, som ga William Vickrey nobelprisen i økonomi, kan du som statistiker få lov til å tenke deg n personer som vurderer auksjonsobjektet iht. uavhengige trekninger fra den rektangulære fordeling [0,1].

Løsning: La oss tenke oss bud på et hus, der hver budgiver har en bestemt oppfatning av hva huset er verdt. Ved hollandsk auksjon (3) lærer budgiverne ingen ting i budprosessen og kunne like gjerne ha skrevet ned budet på forhånd. Dette gir samme resultat som en lukket budrunde der høyeste bud vinner (2).

De to auksjonsformene (1) og (4) adskiller seg ved at engelsk auksjon (1) gir hver budgiver kunnskap om opponentenes bud og mulighet for å reagere på disse, noe som (2) ikke gir anledning til. Imidlertid vil det ved (1) for hver budgiver være optimalt å høyne budet (med så lite som mulig) så lenge den verdi vedkommende har satt på huset ikke er overskredet. Følgelig blir ved engelsk auksjon huset solgt til budgiveren med det høyeste bud marginalt over den med det nest høyeste bud.

Ved auksjonsform (4) vil optimal strategi for hver budgiver være å by den verdi en selv setter på huset i forvissning om at en kun trenger å betale det nest høyeste bud. Ingen budgiver har fordel av å by mindre, siden dette vil redusere sjansen for å få huset uten at prisen en må betale reduseres. Ei heller har en fordel av å by mer, siden dette bare øker sjansen for å vinne dersom også det nest høyeste bud er over ens egen verdiansettelse, og i så fall ender en opp med å betale mer enn hva huset er verdt (for en selv). Dersom alle byr sin private verdi vil altså huset bli solgt til høyeste byder til en pris for nest høyeste bud, som bare vil være marginalt forskjellig fra hva som oppnås ved (1).

For å belyse auksjonsform (2) tenker vi at det er n budgivere og at budgiver nr. i byr b_i på auksjonsobjektet som av budgiveren blir vurdert til verdi v_i. La p(b) være

sannsynligheten for at et bud lik b vinner budrunden. Resonnementene nedenfor innebærer en del betinging uten at dette blir synliggjort i notasjonen.

Forventet gevinst ved auksjonen for budgiver nr. i er da (alle budgivere er risikonøytrale!)

$$H_i = (v_i - b_i)\ p(b_i)$$

Siden alle budgivere står ovenfor det samme maksimeringsproblem, er det rimelig å anta at $b_i = g_i(v_i)$, der g har positiv derivert med invers h. Følgelig kan vi skrive

$$p(b_i) = P(\max_{j \neq i} g(v_j) < b_i) = P(\max_{j \neq i} v_j < h(b_i))$$

Antar vi at v_j'ene er stokastisk uavhengige trekninger fra en kumulativ fordeling F blir

$$p(b_i) = (F(h(b_i)))^{n-1} = (F(v_i))^{n-1}$$

Dersom budgiver nr. i vil maksimere forventet gevinst må vi ha

$$p'(v_i)\ (v_i - b_i) - p(b_i) = 0$$

Sammenholdes de to siste ligningene får vi

$$(n-1)\ f(h(b_i))\ h'(b_i)\ (v_i - b_i) - F(h(b_i)) = 0$$

For tilfellet at F er rektangulært fordelt [0, 1] blir dette

$$(n-1)\ h'(b_i)\ (h(b_i) - b_i) - h(b_i) = 0$$

Denne ligningen har løsning

$$v_i = h(b_i) = \frac{n}{n-1}\ b_i$$

Forventet pris blir da

$$\frac{n}{n-1}\ E \max v_i = \frac{n-1}{n} \cdot \frac{n}{n+1} = \frac{n-1}{n+1}$$

Hvis vi så sammenligner dette med (1) som var ekvivalent med (4), der det var optimalt å by egen verdsatt verdi, blir forventet pris i tilfellet med uavhengige R [0, 1] trivielt lik $(n-1)/(n+1)$, dvs. samme som ved (2). Altså er alle de fire auksjonsformer ekvivalente under de gitte forutsetninger.

Jostein Lillestøl

helsefakta

Gjennomsnittsalder ved død etter spesialitet, alle godkjente spesialister som døde 1989–99, antall døde og antall nålevende

	Alder ved død	Antall døde	Antall nålevende
Lungesykdommer	83,3	61	188
Kjevekirurgi og munnhulesykd.	80,5	4	34
Geriatri	80,1	10	103
Hud- og veneriske sykdommer	79,7	23	184
Urologi	79,6	9	146
Medisinsk mikrobiologi	79,1	11	113
Radiologi	78,5	63	523
Øyesykdommer	78,3	32	371
Øre-nese-halssykdommer	77,2	34	337
Generell kirurgi	77,1	109	1 227
Klinisk nevrofysiologi	76,3	3	28
Nevrologi	76,1	28	263
Nevrokirurgi	76,0	6	54
Fødselshjelp og kvinnesykdommer	75,8	42	646
Barnesykdommer	75,6	38	531
Indremedisin	75,3	102	1 461
Klinisk kjemi	75,2	6	115
Psykiatri	75,0	123	1 054
Revmatologi	75,0	24	150
Fysikalsk medisin og rehabilitering	73,3	9	122
Barne- og ungdomspsykiatri	72,6	9	146
Samfunnsmedisin	71,0	23	572
Ortopedisk kirurgi	70,4	16	461
Endokrinologi	70,3	3	62
Thoraxkirurgi	67,6	5	94
Gastroenterologisk kirurgi	67,0	5	148
Plastikkirurgi	67,0	8	91
Allmennmedisin	66,7	50	2 696
Fordøyelsessykdommer	66,6	5	173
Hjertesykdommer	65,8	9	279
Patologi	62,6	13	190
Anestesiologi	61,7	14	792

Kilde: Den norske lægeforenings legeregister per 1.8. 2000.

Helsefakta eller helsetøv?

Fra Tidsskrift for den norske legeforening.

Er dette verdifull informasjon?

Red.

Gjemt, men ikke glemt 50-åring: Gaute Vik.

I sommer stod denne jubileums-omtalen i "Bergens Tidende". Den viser en karriere som er noe uvanlig for en statistiker, og vi gjengir den derfor i TG. Gaute Vik tok sin candidatus realium-eksamen med hovedfag i statistikk ved Matematisk institutt, Universite-tet i Bergen høsten 1978. Hans veileder var Trygve Nilsen og tittelen på oppgaven er: "En tidskontinuerlig forgreiningsprosess i stokastisk miljø". Det var den tiden en var interessert i finurlige egenskaper for forgreiningsprosesser. De har en viss likhet med tidsrekker av typen AR(1)-modell der korrelasjons-koeffisienten er stokastisk. –Etter fullført eksamen ble Gaute ansatt på FFI på Kjeller og tok senere Ph.D. ved Oklahoma State University, Stillwater. Vedlagte utklipp fra Bergens Tidende skulle være bevis nok for at han også siden har gjort suksess i livet. Vi gratulerer!

Gamle fagfeller i Bergen.



jubilant BT 900623

Administrerende direktør GAUTE VIK i Norconsult Telematics, Kringv. 14 Oslo fyller 50 år 26. juni.

Gaute Vik er en ekte Vålandsgutt fra Stavanger og har vokst opp i Furrasgaten. Han er gift med Salhus-jenten Lisbeth og familien består i tillegg av barna Torstein og Gunn-vor.

Etter å ha tatt hovedfag i matematikk ved Universitetet i Bergen i 1978 fikk Gaute jobb som forsker ved forsvarets forskningsinstitutt på Kjeller. I perioden 1982 til 1984 hadde han permisjon fra sin forsker-stilling og studerte da ved Oklahoma State University hvor han tok doktorgrads-eksamen i statistikk.

Gaute Vik ble så tilsatt i bedriften Norconsult Tele-matic AS i Saudi Arabia i 1986. Dette verdensomfat-tende firmaet er han i dag administrerende direktør i. Lisbeth og barna som har fulgt med Gaute rundt i den store verden har vært hans beste støttespillere både under hans studieopphold i USA og hans krevende jobb i Midtøsten, hvor han nå har arbeidet i 14 år.

Gaute må kunne gis kar-akteristikken å være et enestående medmenneske. Han er kvikt med et utrolig godt humør og pågangsmot. Han er glad i mennesker og derfor betyr familien, ven-ner og medarbeidere svært mye for ham.

Han viser omsorg og set-ter krav til sine medmen-nesker, men de største kra-vene setter han til seg selv. Han må derfor være en enestående leder som får det beste ut av sine medar-beidere.

[illegible] setter svært stor pris på vår gode [illegible] nådde å «diveta» i middags-høyden. Vårt vennskap med Gaute betyr mye og har gitt oss mange uforglemmelige gode timer.

I dag på Sankthansaften skal Gaute feires på Lerby Gods i Lærenskog hvor fam-ilien og vennene fra nord til sør i landet skal hylle ham og gjøre stas på ham.

Vi ønsker Gaute lykke til [illegible] store dagen.

SAMMENKOMST OM STATISTIKK-UNDERVISNING KRISTIANSAND 16.8.2000

Henrik Dahl, Høgskolen i Agder

Norsk Statistisk Forening har i det siste engasjert seg stadig sterkere i undervisningsspørsmål. På Norsk Statistisk Forenings hovedmøte på Milde i 1987 ble det drøftet om foreningen skulle opprette undervisningsutvalg, men det ble da vedtatt å utsette saken. Det ble opprettet undervisningsutvalg i Norsk Statistisk Forening på hovedmøtet på Inderøy i 1989. I de første årene arbeidet medlemmene av undervisningsutvalget hovedsakelig individuelt, særlig med å forsyne TG med stoff om statistikk-undervisning. På 14. Nordiske konferanse i Matematisk Statistikk, Røros 1992 var det en egen sesjon for statistikkundervisning.

Det store engasjementet kom med Reform 94, der undervisningsutvalget med Jostein Lillestøl, Ivar Heuch og Henrik Dahl på oppfordring fra formannen i Norsk Statistisk Forening hadde en gjennomgang av de nye lærebøkene i VGS. Dette arbeidet er oppsummert i en artikkel i TG nr.1 1998. Etter dette har mange statistikere engasjert seg i undervisningsspørsmål. Det er gitt etterutdanningskurs og man har engasjert seg i statistikkundervisning i andre skoleslag. Tiden var derfor moden til å arrangere en sammenkomst om statistikkundervisning.

Takket være økonomisk støtte fra Norsk Statistisk Forening kunne vi få samlet folk fra Tromsø i nord til Kristiansand i sør. Det ble konsumert store mengder kaffe, spist kaker og sammenkomsten ble avrundet med reker. Deltakere var:

Elinor Ytterstad (Universitetet i Tromsø)
Tore Wentzel-Larsen og Rolf Volden (Høgskolen i Bodø)
Ivar Heuch (Universitetet i Bergen)
Ørnulf Borgan (Universitetet i Oslo)
Torfinn Torp (Norges Landbrukshøgskole, Ås)
Henrik Dahl (Høgskolen i Agder)
Hege Therese Syvertsen (Kristiansand)

Innlegget til Elinor dreide seg om etterutdanning. Slik situasjonen er i Norge i dag er det mange som underviser i statistikk med beskjeden bakgrunn i faget. Etterutdanning vil kunne bedre situasjonen for disse.

Tore og Rolf tok for seg angrep statistikk-faget utsettes fra og hvordan slike angrep kan imøtegås.

Ivar tok for seg hvordan vi skal trekke inn statistisk programvare i statistikkundervisningen. Å få slik undervisning til å virke godt i praksis er ikke noen enkel oppgave.

Ørnulf viste hvordan situasjonen er for statistikk-faget i VGS og hvordan man har konsolidert stillingen etter Reform 94 og turbulensen som fulgte. Statistikerne bør vise en fleksibel, men aktiv rolle.

Det var givende å få samlet statistikere som interesserer seg for undervisningsspørsmål i Norge. Jeg vet det er mange flere som deler våre interesser. Kanskje vi blir flere på neste møte.

Videreutdanning i statistikk for lærere
Erfaringer fra Tromsø

Elinor Ytterstad
Universitetet i Tromsø

Aktivitetene i Tromsø innen etter- og videreutdanning startet i etterkant av reform '94. Vi gjennomførte to femdagerskurs (todelt) i sannsynlighetsregning og statistikk med dekning av 2MX og 3MX pensum. Totalt deltok 51 matematikklærere fra Troms og Finnmark. Nordland arrangerte egne kurs i Bodø. Statens Lærerkurs fullfinansierte kursene.

Tre år senere arrangerte fakultetet et etterutdanningskurs i matematikk og statistikk. Også denne gang finansiert av Statens lærerkurs. 17 matematikklærere fra landsdelen deltok. Kurset gikk over en dag med forelesninger på formiddagen og diskusjoner etter lunsj.

Nå i høst startet S-101 (5 vt) som fjernundervisning. Ved hjelp av finansiering fra Statens Lærerkurs er begynnerkurset i statistikk ved Universitetet i Tromsø gjort mere tilgjengelig for lærere som ønsker ett fjernstudium. Bevilgningen er på 124 000 og skal hovedsakelig dekke reise og opphold i forbindelse med en samling i Tromsø samt eksamensreise. Kurset ble annonsert i Statens Lærerkurs sin kurskatalog for snart ett år siden. I tillegg har vi lagt ut informasjon på Internet lenket til fakultetets hjemmeside. 13 lærere fra videregående skoler er påmeldt, dvs fire fra Finnmark, tre fra Troms, fem fra Nordland og en fra Sogn og Fjordane.

Fjernstudiet startet med en samling over tre dager i oktober i Tromsø. Ni av de påmeldte møtte. De fleste hadde noe kunnskap innen sannsynlighetsregning og kombinatorikk fra tidligere, og vi kunne derfor gå forholdsvis raskt og overfladisk gjennom en tredel av pensum. Vi stoppet ofte opp for å regne enkle oppgaver. Mye tid ble brukt til å snakke om fagets egenart, det stokastiske og inferens.

I tiden fram til eksamen i mai/juni arbeider lærerne selvstendig etter en framdriftsplan med utvalgte oppgaver (med løsningsforslag) fra lærebok. De kan til enhver tid ta kontakt med kursansvarlig elektronisk eller per telefon, og informasjon fra oss går via e-post.

Til våren foreleses samme kurs på tradisjonell måte, med eksamen i mai/juni. Det er den samme eksamen fjernstudentene melder seg opp til og får reisestøtte til.

I etterkant av første samling ser vi at det burde ha vært søkt midler til to faglige sammenkomster i Tromsø. En samling f.eks. i mars måned. En annen modell er å reise ut til enkelte utvalgte videregående skoler noen ganger i løpet av perioden. Da vil flere lærere ved samme skole og eventuelt fra naboskoler, få anledning til å følge kurset. Dette krever flere undervisningstimer, men ikke nødvendigvis større busjetter.

Statistikk på økonomisk-administrative studier ved statlige høgskoler

Tore Wentzel-Larsen, Siviløkonomutdanningen i Bodø
Oppdatert og noe omarbeidet versjon av innlegg på møte i Kristiansand i august.

Det felles innslaget av statistikk på de økonomisk-administrative studiene har vært et tovekttallskurs i 2. semester, basert på 2 vekttall matematikk fra 1. semester. I tillegg har noen institusjoner generelle metodekurs og valgfag med varierende innslag av statistikk, og det er en god del bruk av statistikk på studentoppgaver. Gjeldende rammeplan (1995) inneholder følgende formuleringer om statistikkens plass i studiet:

"For å kunne gjennomføre grunnleggende analyser innenfor bedriftsøkonomi og samfunnsøkonomi er det behov for kunnskaper i metodefagene matematikk, statistikk og informasjonsbehandling. I tillegg har disse fagene egenverdi i kraft av å gi studentene trening i abstraksjon og stringens" (side 3)
"Metodefagene skal hjelpe studenten til å treffe hensiktsmessige valg av modeller og datasett. De skal gi oversikt over alternative framgangsmåter og hvilke konsekvenser disse får. Metodefagene skal være redskap for studenten til å øve opp evnen til kritisk vurdering av resultatet ved bruk av modeller." (side 8)
"Det er vesentlig å vurdere hvor godt en modell beskriver virkeligheten. For å kunne gjøre dette, er det nødvendig å beherske forskjellige utvalgsteknikker, eller forenklinger i behandling av data. På denne bakgrunn, skal studenten få grunnleggende innføring i statistiske metoder. En skal også kunne vurdere kvaliteten til innsamlede data og måten data presenteres på. Det er viktig at studenten kan vurdere validiteten i statistiske framstillinger både i faglig sammenheng og i dagliglivet. Studenten bør få innsikt i hvordan en nytter statistiske metoder i økonomi og samfunnsvitenskap." (side 9)
"Kursene i matematikk og statistikk bygger imidlertid på kunnskaper tilsvarende 2MY i studieretning for allmenne, økonomiske og administrative fag i videregående skole. For studenter som ikke har tilstrekkelige forkunnskaper kan høgskolene tilby forkurs i matematikk."

Nasjonalt råd for økonomisk-administrativ utdanning ved statlige høgskoler (NRØA, underlagt Universitets- og høgskolerådet) nedsatte i fjor et utvalg til å utrede tre-årig struktur for de økonomisk-administrative studiene. Utvalgets forslag, med endringer vedtatt etter NRØA-møte i mai, er oversendt Departementet med anbefaling. Utvalgets sammenfatting av behovet for ny struktur omfatter (side 4-7, mi sammenfatting av en del punkter)

- Toårig økadm. har utspilt sin rolle som selvstendig utdanning
  (endret arbeidsmarked, flere tar påbygningsstudier)
- Lav studentgjennomstrømning
- Konkurranse fra private høgskoler
- Internasjonale utviklingstrekk

Om målgruppe (side 7): Både heltid og deltid, også supplering av tidligere utdanning

Egenskaper (side 8-9, utvalgets overskrifter):

- Hovedmodellen skal fremstå som et treårig selvstendig bachelorprogram på heltid
- Bachelorprogrammet skal være direkte påbyggbart til mastergrad og doktorgrad
- Hele utdanningen, eller deler av den, skal lett kunne tilrettelegges for deltidsstudenter

- Heltidsmodellen skal kunne modulariseres
- Det skal være lett å ta moduler med seg fra en institusjon til en annen
- Strukturen i utdanningen må være robust og stabil over tid
- Modellen må være enkel å administrere
- Modellen må være lett å synliggjøre

Utvalget går inn for en 2+1-modell, med et forholdsvis fast opplegg for de to første studieårene, slik at det (vedlegg side 4) "skal være enkelt å bytte høgskole etter hvert fullført studieår". På tredjeåret har høgskolene (forslag til rammeplan side 4) "stor frihet til å framheve sin egenart og profil gjennom sine profileringsretninger". Med en slik modell blir nivået på de to første årene avgjørende for nivået på studiet, og også for nivået på påbygningsstudier.

For matematikk/statistikk var opprinnelig forslag ett felles kurs på 2 vekttall 1. studieår, basert på generell studiekompetanse. Det skulle inngå mer matematikk i en del økonomikurs, og noe mer statistikk i et kurs i samfunnsvitenskapelig metode 2. studieår. Eksamen var for de fleste 2-vekttallskurs 1 time nasjonal flervalgseksamen og 1 time utarbeidet lokalt. En del av disse forslagene møtte motstand innen flere fagmiljøer, også blant statistikere. Et resultat var Norsk statistisk forenings uttalelse, der samfunnsvitenskapelig metode i 2. studieår ble foreslått erstattet av et kurs "Statistisk analyse i bedriften".

På et NRØA-møte i mai gikk en, innenfor en 2+1-modell, inn for at "evalueringsformer ikke skal fastsettes i rammeplanen", at "det organiseres egne metodekurs på totalt 6 vt., og at disse utformes i samarbeid mellom fagmiljøene i metode- og økonomifagene" og det ble vedtatt at "NRØA gir arbeidsutvalget fullmakt til å nedsette fire faggrupper med representanter fra utvalgte høgskoler innen a) bedriftsøkonomisk analyse, b) administrasjonsfag og juridiske fag, c) samfunnsøkonomi og d) metodefag. Faggruppene gis i oppdrag å utarbeide mål og innhold i beskrivelsene av de ulike obligatoriske emnene i samsvar med den foreslåtte nye rammeplanen. Faggruppene bes også anbefale mulige undervisnings- og evalueringsformer." Arbeidsutvalget i NRØA sluttet seg i et møte seinere i mai til dette og vedtok at faggruppa i metodefag, som omfatter matematikk og statistikk, skulle oppnevnes av høgskolene i Buskerud, Sør-Trøndelag og Østfold. I arbeidsutvalgets reviderte forslag til rammeplan er følgende verd å merke seg: Innføringskursene i elementær matematikk og statistikk skal begge bygge på generell studiekompetanse alene. Til sammenlikning er innføringskurset i statistikk i nåværende rammeplan fra 1995 basert på grunnkurset i matematikk. Flervalgseksamen er nå tenkt som supplement til hjemmeeksamen. Videre kan evaluering i flere emner slås sammen, og det sies at enkelteksamener "bør normalt være på minst 4 vekttall." Dette kan innebære at det ikke anbefales en egen eksamen i statistikk på studiet.

Departementet har i høst blant annet (jfr arbeidsutvalgsmøte NRØA 12/9) meddelt NRØA at bruk av rammeplaner i høyere utdanning skal behandles politisk, så den videre framdrifta er usikker. Uansett er det viktig at Norsk statistisk forening fortsatt følger opp denne saken. Faggruppa i metodefag kan være en mulig kanal.

Referanser:
Gjeldende rammeplan fra 1995 kan lastes ned fra http://www.nnr.no/ En må leite en del, klikk først på Publikasjoner, så på Rammeplaner, Vedtatte og Grunnutdanning, og bla nedover.
Dokumentene fra NRØA finnes under www.uhr.no/nasjonale-rad/nrokad
(møtereferater, og forslag til ny struktur med link til foreslått rammeplan)

Matematikk, statistikk og metode ved Siviløkonomutdanningen i Bodø

Momenter til diskusjon i Kristiansand den 16. august ved Rolf Volden.
Innlegget er noe modifisert etter at forslaget til ny struktur i øk/adm-studiet ble revidert.

Nedenfor er listet opp ulike kurser og aktiviteter innen matematikk, statistikk og metode i ulike studier ved Siviløkonomutdanningen i Bodø (SiB). Vi har valgt å ta inn metodekursene i listen da de ofte inneholder en del statistikkrelatert stoff innen kvantitative metoder i tillegg til kvalitative metoder og vitenskapsteori. I tillegg inneholder listen en del aktiviteter i form av kandidatoppgave, siviløkonomoppgave, ledet selvstudium og hovedfagsoppgaver som kan inneholde nokså mye anvendt statistikk, alt avhengig av studentenes valg og interesser. Beskrivelser av studiene og kursene er delvis lagt ut under Høgskolen i Bodøs hjemmesider.

1. avdeling: Øk/adm-studiet (2 år):

11327 Matematikk for økonomer	2 vt. i 1. semester (obligatorisk)
11303 Grunnkurs i statistikk	2 vt. i 2. semester (obligatorisk)
10133 Metode	2 vt. i 3. semester (obligatorisk for adm-linje)
19903 Kandidatoppgave	2 vt. i 4. semester (obligatorisk)
22231 Grunnkurs i matematisk analyse	4 vt. i 3. og 4. semester (valgfag)

2. avdeling: Siviløkonomstudiet (2 år):

50002 Metodelære	2 vt. i 1. semester	(obligatorisk)
50027 Matematiske metoder	2 vt. i 1. semester	(valgfag)
50028 Statistiske metoder	2 vt. i 2. semester	(valgfag)
22231 Grunnkurs i matematisk analyse	4 vt. i 1. og 2. semester	(valgfag)
5060x Siviløkonomoppgaven	8 vt. i 4. semester	(obligatorisk)

Hovedfag i bedriftsøkonomi med vekt på innovasjon og iverksetting:

55002 Anvendt metode	2 vt. i 3. semester	(obligatorisk)
55040 Ledet selvstudium	4 vt. i 3. semester	(obligatorisk)
55100 Hovedfagsoppgave	10 vt. i 4. semester	(obligatorisk)
550xx Kurs i kvantitativ metode	(under planlegging)	

I matematikk er det bare kurset 11327 på 2 vekttall som er obligatorisk, som basis for de tre studiene. Resten av matematikk-kursene er valgfag. Enkelte studenter kan ha mer matematikk i opptaksgrunnlaget til disse studiene. I statistikk er 2-vekttallskurset 11303 obligatorisk og basis for de tre studiene, men i tillegg inneholder de ulike obligatoriske metodekursene på de tre studiene en del statistikkrelatert stoff innen kvantitative metoder. Det er få studenter på øk/adm-studiet som forsøker å ta valgfaget i matematikk, kurs 22231 (MA100-ekvivalent med vekttallsreduksjon) da dette kurset betraktes for å være arbeidskrevende og vanskelig.

Hvordan kunne så utvalget for vurdering av ny struktur for øk/adm-studiet forut for NRØA-møtet i mai foreslå å redusere den allerede begrensede obligatoriske basisdelen i matematikk og statistikk fra 4 til bare 2 vekttall i et felles metodekurs? Riktignok foreslo man å overføre deler av de innsparte vekttallene i matematikk og statistikk, hovedsaklig til kursene i finansiering og investering og til samfunnsvitenskaplig metode på 2. studieår. Det er et nokså illevarslende signal at administratorer og fagpersoner ved høgskolene virkelig tør foreslå et

slikt tiltak for å redusere frafallet i øk/adm-studiet. Forslaget ser nå ut til å være stoppet , trolig på grunn av de sterke innvendingene på mai-møtet fra flere høgskoler som var sterkt negative til de foreslåtte endringer på dette punktet. Kanskje har også innvendingene fra NSF spilt en viss rolle i denne sammenhengen. Det er i dag ikke klart hvilke endringer i rammeplanen for matematikk og statistikk-fagene som den nyutnevnte faggruppa i metodefag vil foreslå.

Matematikk og delvis statistikk har lenge hatt ord på seg for å være vanskelige fag med relativt høye strykprosenter og ikke-møtt prosenter. Dette er kurs som kommer på første året og som krever en god del systematisk jobbing og delvis endrede studieteknikker fra de benyttet i videregående skole. Spesielt matematikk som kommer i det første semestret, er et problem for mange studenter som ofte har glemt det lille de engang kunne fra videregående skole eller som der har valgt bort matematikk i den grad det var mulig Samtidig må kandidatene endre sine studievaner fra å være i en nokså kontrollert elevrolle til å være i en mer fri studentrolle uten nevneverdig kontroll av fremmøte, tilstedeværelse og innsats. I tillegg møter studentene et virvar av nye fritidstilbud og gode tilbud om deltidsjobber i høgskolenes nærmiljø, noe som kan være med på å dra interessen bort fra selve studiet. At studentene kommer dårlig ut i det første semestret på øk/adm-studiet kan derfor ha helt andre årsaker enn at "matematikk er vanskelig".

I de senere åra har strykprosentene i samfunnsøkonomi vanligvis vært noe lavere enn i matematikk, mens strykprosenten i bedriftsøkonomi og regnskap har vært tilnærmet lik den i matematikk. Forskjellene i strykprosenter mellom økonomifagene og matematikk er ikke større enn at de kan forklares ved at økonomifagene har eksamener noe lengre ut i studiene enn matematikk, og de mest umotiverte studentene har kanskje sluttet allerede da. I tillegg er det også betydelig høyere strykprosenter i fag som markedsføring og organisasjon enn det har vært tidligere. Dette siste kan tyde på at den lave gjennomføringen på øk/adm-studiet har flere og mer kompliserte årsaker enn et "vanskelig" matematikk-kurs. Den foreslåtte nedgraderingen av matematikk og statistikkfagene på øk/adm-studiet, som ble foreslått av NRØA-utvalget og som nå forhåpentligvis er stoppet, synes derfor å mangle en skikkelig begrunnelse.

Siviløkonomstudiet:
Som det fremgår av listen ovenfor, finnes det ingen obligatoriske kurs i matematikk, mens der er noe statistikk i det obligatoriske metodekurset 50002. De øvrige tilbud som finnes er i form av valgfag og en obligatorisk siviløkonomoppgave. Noe under 20% av studentene tar valgfag i matematikk eller statistikk. Enkelte studenter kan ha mer matematikk og statistikk i opptaksgrunnlaget til studiet. Det er uklart hva som vil skje ved SiBs kursporteføIge på siviløkonomstudiet om det skulle bli foreslått store endring i de obligatoriske kursene i matematikk og statistikk i den nye rammeplanen for øk/adm-studiet.

Hovedfaget i bedriftsøkonomi med vekt på innovasjon og iverksetting:
KursporteføIgen på hovedfaget er under stadig vurdering på grunn av endringer i hovedfaget ved NHH og konkurrerende mastergradsstudier i økonomi. I dag finnes det ingen obligatoriske kurser hverken i matematikk eller i statistikk.

Bruk av statistisk programvare i elementærundervisningen: Må vi revidere kursoppleggene?

Ivar Heuch
Matematisk institutt, Universitetet i Bergen

Bruk av programvare har inngått i undervisningen i statistikk på universitets- og høyskolenivå i mange år. Det er åpenbart stor enighet om at øvelser på datamaskin utgjør en naturlig del av undervisningen i faget, selv om omfang og organisering kan diskuteres. Når vi nå står foran et nytt trinn i utviklingen på dette feltet, er det naturlig å peke på at det ikke alltid har hersket slik samstemmighet. Da jeg i 1981 holdt et innlegg om saken på Norsk statistikermøte på Bø [1], møtte selve tanken om å legge anvendelser av programvare inn i brukerundervisningen til dels sterk motbør.

I dagens undervisningssystem inngår det ofte forskjellige typer statistisk programvare på ulike nivåer, avhengig av målgruppe og forutsetninger. Dette innlegget tar opp spørsmål i forbindelse med grunnleggende statistikkundervisning på høyskolenivå, i alt vesentlig beregnet på studenter uten særlig matematisk bakgrunn. Datakurs av denne typen forekommer i mange varianter ved lærestedene i Norge. Hensikten er i noen grad å gi studentene et ferdig redskap for analyse av datasett. Dessuten har undervisningen vanligvis et videre perspektiv som ikke er knyttet til noen bestemt programvare. Det gjelder for det første de generelle prinsippene for praktisk dataanalyse. I mange tilfeller utnyttes øvelsene på datamaskin også til å skape større forståelse for underliggende ideer i sannsynlighetsregning og statistikk. Kurs for brukergrupper på høyere nivå krever derimot ofte andre dataopplegg.

En klassifisering av dagens kursopplegg

Mange av kursene på dette nivået har fulgt samme mønster, men det har også forekommet vesentlige avvik. En fruktbar diskusjon av mulige fremtidige opplegg bør bygge på en klassifisering av kursoppleggene som allerede har vært i bruk. I det følgende skisserer jeg min egen grove inndeling:

1. 'Standardopplegget'. I mange kurs blir det gitt et par litt større oppgaver som forutsetter bruk av statistikkprogram på datamaskin. Oppgavene medfører vanligvis analyse av bestemte datasett, og det skal leveres besvarelser til godkjennelse innen oppgitte frister. I den grad det gis systematisk opplæring i bruk av det aktuelle programsystemet, skjer det gjerne uformelt ved siden av den vanlige undervisningen. Studentene benytter datamaskiner som er tilgjengelige på bestemte datarom, ofte i konkurranse med studenter i andre fag. Øvelsene faller omkring midten av undervisningssemesteret. Til eksamen blir det ikke gitt oppgaver som dreier seg spesifikt om programsystemet, men det forekommer ofte spørsmål om fortolkning av typiske utskrifter. Det finnes mange varianter av dette opplegget. For ti år siden hadde TG en serie som gikk over flere nummer, med beskrivelse av dataoppleggene som ble fulgt ved mange norske læresteder (innledet av et innlegg om forholdene ved det daværende NTH [2]). Artiklene byr stadig på interessant lesning, og mange av beskrivelsene gjenspeiler forskjellige utformninger av 'standardopplegget'. Selvsagt er det foretatt vesentlige forandringer i løpet av de siste ti årene, men det er grunn til å tro at det fortsatt eksisterer versjoner av dette opplegget mange steder.

2. 'Laboratorieopplegget'. Med studenter som ikke har trening i databehandling, kan 'standardopplegget' ha en tendens til å flyte ut, slik at tiden blir lite effektivt utnyttet. For å bøte på dette problemet innførte vi 1989 en mye strammere ordning i elementærkurs i statistikk for realfagsstudenter ved Universitetet i Bergen. Hver student ble plassert i en gruppe på maksimalt ti personer, tilsvarende antallet PC-er i et lokale brukt utelukkende for dette formålet. Hver gruppe fikk tilordnet en bestemt periode på 2 timer annenhver uke som 'laboratorietid', mer i tråd med øvelsespraksis i laboratoriefag som fysikk og kjemi. I løpet av denne tiden skulle hver student følge et forholdsvis stramt opplegg, med mer detaljert teknisk beskrivelse av hva som skulle utføres [3] enn det som har

vært vanlig i 'standardopplegget'. Under øvelsen er det hele tiden en gruppeleder til stede som gir hjelp når det dukker opp problemer. Gruppelederen bør også gripe inn på eget initiativ dersom en student tydeligvis faller utenfor øvelsesprogrammet. Oppgavene er formulert som spørsmål knyttet til en praktisk problemstilling, men det konsentrerte opplegget er valgt slik at det både skal illustrere statistiske metoder og noe mer tekniske sider ved programsystemet. Utskriften som fås gir grunnlag for å svare på spørsmålene i oppgaveteksten, selv om det kan bli påkrevd med mindre tilleggsutregninger. I alt har det inngått seks øvelser i løpet av semesteret i det aktuelle trevektallsemnet, mens det bare har vært krevd at fire av de seks besvarelsene skulle være godkjent for at en student kunne gå opp til eksamen. De statistiske problemene som tas opp, reflekterer stort sett progresjonen i behandlingen av pensum ellers. Det har også vært tilbudt ekstra tid etterpå for studenter som ville arbeide mer med de aktuelle oppgavene, men i de siste semestrene har dette ikke latt seg gjøre p.g.a. problemer med sikkerheten.

3. 'Læreboksopplegget'. Dette opplegget er på enkelte måter helt annerledes enn det som hittil har vært praktisert i Norge. Derimot ser det ut til å være ganske utbredt i USA, tatt i betraktning lærebøkene av denne typen som stadig kommer tilsendt fra forlagene. Disse bøkene gir forholdsvis enkle eksempler på hvordan bestemt programvare kan brukes til å illustrere begrep som nettopp er innført i teksten. Forfatterne tenker seg tydeligvis at leseren skal sitte med datamaskin ved siden av boken, slik at det straks er mulig å teste ut ideene ved hjelp av statistikkprogrammet. Et typisk eksempel er læreboken av N.E. Weiss [4], som til stadighet har avsnitt med referanser til Minitab innimellom den ordinære fremstillingen. Avsnittene er markert med et Minitab-symbol i margen, mens selve teksttypen er lik hele veien, slik at bruken av programmet blir integrert i den generelle forklaringen. Riktignok har boken også tilsvarende symboler for TI-83 og Excel, men disse forklaringene spiller en underordnet rolle. Fordelen med akkurat denne boken er at forklaringene til Minitab passer ganske godt til de nyeste programversjonene (men ikke helt, siden boken refererer til versjon 12, mens versjon 13 kom tidligere i år). Selvsagt finnes det også egne avsnitt om programvare f.eks. i norske elementærbøker i statistikk, men vanligvis er det mer snakk om forklaringer på siden av teksten enn illustrasjoner integrert i metodebeskrivelsene.

4. 'Opplegg basert på statistikkprogram som læremester'. Opplegget i punkt 3 er på mange måter ganske tradisjonelt i forklaringen av ideer og metoder, selv om bruk av datamaskin inngår i eksemplifiseringen. En annen og mer utfordrende ide er det å la selve utnyttelsen av programvaren spille rollen som læremester, i den forstand at alle grunnleggende begrep innføres ved illustrasjoner med statistikkprogrammet. Måten studentene nærmer seg stoffet på, blir annerledes, og opplegget kan tenkes å appellere til studenter som ellers neppe ville funnet seg til rette med en tradisjonell lærebokspresentasjon. Det mest utpregede eksempelet av denne typen i Norge er antagelig opplegget laget av Lillestøl og Ytrehus [5] for bruk ved kursvirksomheten ved Norges Handelshøyskole. Lillestøl har beskrevet positive og negative erfaringer fra dette opplegget [6].

5. 'Dataøvelser integrert i opplegg for datastøttet læring'. Det neste naturlige trinn i integrasjonen av datakraft med undervisningsopplegg representerer læreprogrammer utviklet nettopp med det formål å innøve begrep i statistikk. Det har lenge eksistert spesielle programmer som skulle gjøre det mulig for studenter å oppnå en bedre forståelse ved hjelp av skreddersydde simuleringsopplegg. Et system av denne typen, utviklet som en del av en hovedfagsoppgave, ble beskrevet i TG i 1990 [7]. På dette området har det skjedd en forholdsvis sterk utvikling utenfor vårt fag i de senere år. Det mest omfattende systemet for statistikkundervisning som jeg har sett av denne typen, er ActivStats [8], som jeg også har fått tilsendt fra forlaget. Straks man setter den vedlagte CD-en inn i en PC, får en tilgjengelig et omfattende læresystem med forklarende tekst for mange grunnleggende begrep. Systemet utnytter alle slags finesser på datamaskinen, med muligheter for små tester med svaralternativer, avspilling av videoer, oppkobling via web, og bruk av forfatterens egen pakke Data Desk for statistiske analyser. Det tekniske virker imponerende ved første gangs bruk, men hvor effektivt systemet er som læremiddel, er det verre å uttale seg om. Så vidt jeg vet, er ikke noe slikt system tatt i bruk i Norge, men forholdene kan raskt endres.

6. 'Programmeringsopplegget'. Før PC-enes og arbeidsstasjonenes tid var bruk av generelle programmeringsspråk mer utbredt innenfor statistiske miljøer. For enklere statistiske metoder kunne disse språkene bli brukt som pedagogisk hjelpemiddel, ved at studentene fikk i oppgave å

programmere passende algoritmer. I dag finner en noe av den samme tankegangen blant statistikere som anbefaler bruk av regneark i brukerkurs. Studentene får iallfall bedre innsikt i hva metoden går ut på når de må stille opp aktuelle regneuttrykk selv. Går en over til mer avanserte metoder, blir det derimot mindre rasjonelt å utføre slik programmering på egenhånd. Likevel skal en ikke undervurdere betydningen av populære system som regneark blant PC-brukerne i situasjoner der det trengs tallbehandling av moderat omfang.

De seks oppleggene skiller seg mye fra hverandre i motiveringen studentene får for å arbeide med programsystemene. I praktisering av 'standardopplegget' og 'laboratorieopplegget' henger motiveringen sammen med arbeidet frem mot godkjennelse av obligatoriske oppgaver. I opplegg 4 oppfordres studentene derimot til å arbeide med datamaskinen for i det hele tatt å forstå hva lærestoffet går ut på. 'Læreboksopplegget' virker delvis på samme måte, men her kan det være enklere for en student bare å hoppe over referansene til statistikkprogrammet. I opplegg 5 er det viktig at dataanalyseverktøyet er godt integrert i lærepakken, selv om dette opplegget også kan kombineres med oppgaveinnlevering. I 'programmeringsopplegget' kan noe av motiveringen ligge i tidligere positive erfaringer med dataverktøyet. Det er vanskelig å si hva som er mest effektivt!

Hvorfor er situasjonen blitt annerledes?

Selv om praktiske opplegg for dataøvelser i statistikkurs ikke har forandret seg fundamentalt i løpet av de siste ti årene, er det tydelig at studentene etter hvert har fått langt bedre databakgrunn. Tidligere var det ganske vanlig å treffe på studenter som var helt hjelpeløse i forhold til datamaskiner. De hadde store problemer med å komme noen vei med et ferdig programmert opplegg i løpet av et avmålt tidsrom. Holdningen er nå stort sett annerledes, også blant studenter som slett ikke orienterer seg i retning av datafag. I stor grad reflekterer dette forandringer som har foregått overalt i samfunnet. I kurssammenheng er det likevel viktig å ta hensyn til at studentmassen fortsatt er ganske heterogen, og ferdighetene til ferske studenter må derfor ikke overvurderes.

Samtidig er datamaskinene er blitt mye lettere tilgjengelige. Det er adskillig enklere å basere undervisning på bruk av dataprogrammer nå uten å måtte gå til store investeringer over trange universitetsbudsjetter. Mange studenter har også egen PC. I praksis er utbredelsen likevel enda ikke så stor at man kan unnvære løsninger basert på felles datarom. Styrende organer ved universitetene har til dels overdrevne forestillinger om hvor langt denne utviklingen er kommet. Ved Universitetet i Bergen ble f.eks. undervisningskatalogen fra høsten 2000 avskaffet som papirutgave, noe som ledet til stor forvirring blant studentene. Med det eksisterende utstyrstilbudet er det ganske enkelt ikke realistisk å tro at det store flertallet av nye studenter har noen mulighet til å slå opp og finne timeplaner på Internett. Om ganske få år kan situasjonen være en helt annen.

Generelt kan en merke store endringer i holdningen hos forfatterne av lærebøker til supplerende programvare. Som for ti år siden finnes det nærmest alle mulige avskygninger i lærebokstilbudet, men likevel må en kunne si at den typiske elementærboken i dag ser helt annerledes ut. Om ikke annet forekommer det iallfall hyppige henvisninger til bruk av programmer. Dessuten har det dukket opp nyere data-intensive statistiske metoder, men riktignok ikke så ofte på elementærnivået.

Samtidig har det foregått en utvikling av programvaren selv, så det er blitt lettere å gjennomføre kompliserte statistiske analyser med programsystem som tidligere ble ansett som relativt enkle. En student som blir vant til å bruke en bestemt statistikkpakke på et tidlig stadium, vil derfor også ha gode muligheter for å kunne gjennomføre f.eks. mer komplekse multivariable analyser senere med samme system. Det er ikke lenger påkrevd å sette seg inn i andre, mer omfattende programpakker. På denne måten blir det enda viktigere å gi en balansert innføring i anvendelser av programvare på grunnivå i undervisningen.

Det er tydelig at de forbedrede praktiske mulighetene for brukerne av statistikkfaget også har medført endringer i holdningen til faget generelt. På mer avansert nivå innenfor helt andre disipliner hører en langt oftere uttalelser om at fagfolk som har god bakgrunn i statistikk, stiller sterkt. I mange tilfeller *forventes* det praktiske ferdigheter med aktuell statistisk programvare for å bli en kvalifisert fagmann. Statistikerne bør møte denne utfordringen i sin undervisning.

Aktuelle diskusjonstemaer

Dagens situasjon med tanke på datatekniske hjelpemidler i statistikkundervisningen er langt fra stasjonær. De tilgjengelige hjelpemidlene blir stadig forbedret, samtidig med at det kommer nye ytre krav til undervisningen. Det virker ikke usannsynlig at utviklingen vil gå raskere i neste tiår enn det vi så i nitti-årene. Hvordan skal vi tilpasse oss? Det er naturlig å formulere en del åpne spørsmål som kan fremme diskusjonen. Forsøk på fullstendige svar er det ikke aktuelt å gi her.

Er 'standardopplegget' fortsatt et rimelig kompromiss i enkelte kurs? Dette opplegget forutsetter at det aktuelle kurset primært skal gi en innføring i statistisk tankegang langs mer tradisjonelle retningslinjer. Det kan dreie seg om enkel matematisk fremstilling eller uformell beskrivelse av begrep og metoder med andre konvensjonelle læringsteknikker. I hvilken grad dette fortsatt vil være et akseptabelt alternativ, avhenger av målene en setter for undervisningen. Tar en f.eks. sikte på å utdanne lærere som kan påta seg undervisning i matematikk i videregående skole, bør pensum ha et ikke uvesentlig teoretisk innslag. I slike kurs kan 'standardopplegget' fortsatt ha sin berettigelse. Kurs basert på mer problemorientert undervisning kan bli bygget opp rundt enda større og mer konsentrerte dataoppgaver. Men hvordan skal studentene i så fall trenes opp i bruk av datahjelpemidlene? Det kan bli nødvendig å ty til andre opplegg for å bringe studentene opp på riktig nivå. I slike tilfeller kan det også være nyttig med bøker som gir en selvstendig, uformell innføring i programvaren, som f.eks. [9].

Bør 'laboratorieopplegget' beholdes for svake studenter? I enkelte kurs er kravene til studentene blitt adskillig større etter at datamaskinene kom. Ideelt sett forventes det nemlig fortsatt de samme teoretiske kunnskapene, parallelt med nye tekniske ferdigheter. Trening i bruk av de tekniske hjelpemidlene behøves dersom studentene senere skal være i stand til å behandle praktiske problemer innenfor andre fagfelt. Bør en i denne situasjonen beholde en mulighet til et mer intensivt opplegg av laboratorietypen? Faller en student utenfor det praktiske dataopplegget, kan konsekvensene bli alvorligere enn hvis man misforstår perifer teori i et tradisjonelt opplegg. Det kan derfor fortsatt være viktig med individuell oppfølgning. Uansett er det ganske arbeidskrevende å utvikle nye 'laboratorieopplegg' og holde dem a jour.

Er det mulig i dag å tilpasse seg 'læreboksopplegget'? Slik forholdene er i øyeblikket, er svaret klart nei, hvis en f.eks. tar for seg realfagsstudenter. I stor grad vil stadige henvisninger til mulighetene i statistikkprogrammer virke som en forstyrrende faktor når en student uten datamaskin går gjennom en lærebok. Situasjonen kan likevel forandre seg raskt, p.g.a. faktorer som ligger helt utenfor statistikernes ansvarsområde. Lignende problemstillinger dukker også opp i andre fag, men nettopp fordi de grunnleggende spørsmålene i statistikkfaget handler om analyse og fortolkning av data, står vårt fag i en særstilling.

Er det berettiget å gi egne kurs etter 'opplegg basert på statistikkprogram som læremester'? Slike opplegg kan ha et potensial overfor nye grupper av studenter som ellers ikke ville lært statistikk. De kan være spesielt aktuelle for personer som er kommet forbi stadiet der man godtar undervisning på tradisjonell skolemaner. Med dagens brukervennlige programvare kan det være mer realistisk å få utarbeidet nyttige opplegg av denne typen, selv om dette også må bli svært arbeidskrevende. En kombinasjon med mer omfattende datastøttet læring kan være aktuell.

Når kan vi dra virkelig nytte av 'datastøttet læringsopplegg'? Dette er stadig et vanskelig generelt spørsmål! Svaret vil avhenge mye av utviklingen utenfor statistikkfaget. Man kan tenke seg at dataoppleggene blir et supplement til tradisjonell undervisning, slik at studentene vil veksle mellom forskjellige læremetoder. Øvelser med bruk av statistikkprogrammer på datasett kan styres av læreprogrammene, eller det kan være mer rasjonelt å holde de to anvendelsestypene av datamaskinene adskilt. Kommersielle forhold kan lett komme til å spille en stor rolle ved utviklingen av slike læretilbud, slik vi har sett det for lærebøker, der det amerikanske markedet ofte dominerer. På et senere stadium vil det trolig komme system som på en enkel måte gjør det mulig for lærerne å legge inn personlige preferanser om aktuelt stoff og oppgaver for kursene som det undervises i.

Bør studenter som f.eks. kjenner regneark, få egne tilbud i 'programmeringsopplegget'? Her spørs det om statistikerne ser seg tjent med å gi kurs innrettet mot grupper som insisterer på bruk av sitt eget favorittverktøy. Spørsmålet avhenger også av hvordan forholdet blir i fremtiden mellom

statistisk og annen programvare. Blir skillet mellom kategoriene utvisket, kan 'programmeringsopplegget' igjen bli mer aktuelt.

Skjer det reelle forbedringer?

Oppleggene for dataøvelser har hittil forandret seg gradvis i takt med endringene i statistisk praksis utenfor undervisningsinstitusjonene. Nye hjelpemidler vil sikkert tas i bruk i årene fremover. Internett gir nye muligheter også for statistikkundervisningen [10], spesielt i interaktive demonstrasjoner som utnytter potensialet i programmeringsspråket Java. Men hvor viktig dette blir for den daglige undervisningen i overskuelig fremtid, er det ikke lett å uttale seg om. At et bestemt læreopplegg f.eks. er basert på web-baserte hjelpemidler, er ingen garanti for kvalitet. For den saks skyld kan opplegget prinsipielt være riktig bra, men likevel uegnet for norske studenter.

Likevel ser en ofte satsing på svært generelle planer for bruk av nye tekniske hjelpemidler i undervisningen. Ved Universitetet i Bergen er det nylig lagt frem en uttalelse fra 'Havik-utvalget' om 'IKT-basert læring' [11]. Ideene som blir uttrykt her, viser stor underliggende entusiasme og legger opp til omfattende ressursbruk, men gir få anvisninger på konkrete tiltak som skal gjøre undervisningen bedre. Det er til og med vanskelig å oppfatte hvordan temaet som omtales skal avgrenses. På side 2 i innstillingen finner en denne presiseringen:

> 'Hva er IKT-basert læring (IKTL)? IKT-baserte læringsformer kjennetegnes ved fagdidaktisk bruk av informasjons- og kommunikasjonsteknologi til å understøtte individuelle og kollektive læringsprosesser, samt å bidra til å gjøre undervisning mer tilgjengelig, brukerorientert og fleksibel, dvs. mer uavhengig av tid og sted, og bedre tilpasset den enkeltes læringssituasjon.'

Kan statistikerne melde seg med noen beskjedne dataøvelser i en slik sammenheng? Jeg har ingen formening om hva de ansvarlige for innstillingen ville si til det spørsmålet. På mange måter fremstår likevel den gradvise utviklingen med dataøvelser som vi har sett over en lang tid i statistikkfaget, som et motstykke til de sentraldirigerte prosessene som det nå ivres for med IKTL. Den stadig mer utstrakte bruken av statistisk programvare har gjenspeilet den generelle bevegelsen i faget selv, og den har kommet brukernes ønsker i møte. Videre har undervisningen tilpasset seg det aktuelle tekniske nivået på forskjellige tidspunkter, og den har hatt direkte sammenheng med lærestoffet ellers i faget. Bevilgninger til datautstyr i statistikkundervisningen har ikke akkurat toppet prioriteringslistene i styrende organer, noe mange studenter også har merket, men iallfall i visse perioder har det vært vist en rimelig interesse lenger oppe i systemet. Hvis statistikerne fører videre den kontinuerlige reformprosessen som startet omkring 1980 med dataøvelser, vil faget fortsatt kunne bidra effektivt til universitets- og høyskoleutdannelser med mange slags siktemål.

Referanser

[1] Heuch I. *Ødelegger programpakkene statistikkundervisningen? Erfaringer fra emnet 'Anvendt statistikk' ved Universitetet i Bergen.* Sammendrag, Norsk statistikermøte 1981.

[2] Tyssedal JS. Litt om dataøvingar i statistikkundervisinga ved NTH, Trondheim. *Tilfeldig Gang* 7 (2), 30-31 (1990).

[3] Heuch I. *Øvelser i elementær statistikk på PC, med bruk av Minitab under Windows*. Matematisk institutt, Universitetet i Bergen, Bergen 1998.

[4] Weiss NA. Introductory statistics. 5. utg. Addison-Wesley, Reading, Mass. 1999.

[5] Lillestøl J, Ytrehus, LJ. *Statistisk dataanalyse med PC. Allmenn innføring med Minitab*. Rev. utg. Fagbokforlaget, Bergen 1993.

[6] Lillestøl J. Innføring i statistisk metode: Erfaringer med datastøttet læring ved Norges Handelshøyskole. *Tilfeldig Gang* 7 (3), 27-30 (1990).

[7] Lilleeng S. Et simuleringsprogram for illustrasjon av begreper i statistisk metodelære. *Tilfeldig Gang* 7 (3), 18-19 (1990).

[8] Velleman P. *ActivStats user's guide*. Addison Wesley Longman, Glenview 2000.

[9] Carver R. *Doing data analysis with Minitab 12*. Duxbury Press, Pacific Grove 1999.

[10] West RW, Ogden RT. Interactive demonstrations for statistics education on the world wide web. *Journal of Statistics Education* [online] 6 (3) (1998) (http://www.amstat.org/publications/jse/v6n3/ west.html).

[11] *Fleksibel læring på den Virtuelle Campus*. Innstilling fra utvalg oppnevnt av Universitetsdirektøren. Universitetet i Bergen 2000 .

Nye læreplaner i sannsynlighetsregning og statistikk i videregående skole

Ørnulf Borgan
Matematisk institutt, Universitetet i Oslo

16. august ble det avholdt en sammenkomst om statistikkundervisning i Kristiansand. Der snakket undertegnede om de nye læreplanene i statistikk og sannsynlighetsregning i videregående skole og om bakgrunnen for disse. Dette innlegget er en lett bearbeidet versjon av det jeg la fram i Kristiansand i august. Leserne får unnskylde at det er en del overlapp mellom dette innlegget og det jeg skrev i TG nr. 1 i år.

Med Reform-94 ble sannsynlighetsregning og statistisk en del av læreplanene i matematikk i videregående skole. På grunnkurset (dvs. første klasse) kom det med noe beskrivende statistikk for alle studieretninger, mens det i videregående kurs I og II (dvs. andre og tredje klasse) på studieretning for allmenne, økonomiske og administrative fag kom med en del sannsynlighetsregning og litt statistisk metodelære.

Noen år etter reformene i den videregående skolen, ble det foretatt en omfattende reform av hele grunnskolen fra barnetrinnet til ungdomstrinnet (Reform-97). Også her ble det endringer i læreplanene i matematikk. Blant annet kom det inn et mål om "behandling av data" som omfatter enkel beskrivende statistikk og litt sannsynlighetsregning.

Høsten 1998 kom matematikkfaget i videregående skole i fokus for den offentlige debatten. En rekke avisinnlegg tok opp faget og dets problemer. Årsaken til debatten var blant annet at de første kullene med Reform-94 elever oppnådde svake resultater i matematikk ved universiteter og høgskoler. Selv om det nok var mange grunner til dette, la flere av debattantene skylden på sannsynlighetsregningen og statistikken. Ved at denne var kommet med, ble det mindre tid til "ordentlig" matematikk, mente de.

I januar 1999 satte KUF i gang et arbeid med å revidere læreplanene i matematikk for videregående skole. Hovedgrunnen til at planene måtte revideres så kort etter Reform-94, var endringene i grunnskolen med Reform-97. Men i tillegg ønsket KUF å endre læreplanene i lys av erfaringene med Reform-94 og debatten om matematikkfaget. Til å lede arbeidet med læreplanene nedsatte KUF en gruppe på fem personer. For å få en bredere basis for diskusjonen, ble det i tillegg til læreplangruppa også oppnevnt en referansegruppe med seks medlemmer. Undertegnede var ett av disse.

Ved revisjonen av læreplanene arbeidet undertegnede - i samråd med NSFs leder og undervisningsutvalg - etter følgende tanker/momenter:

- Den beskrivende statistikken som kom med ved Reform-94 har ikke fungert godt. Det meste av dette stoffet dekkes videre av de nye planene for grunnskolen. Endringer er derfor nødvendig.
- Hovedvekten bør legges på sannsynlighetsregning. Denne må være spennende og inspirerende og ikke ha karakter av "julenøtter". Det er viktig å vise mange andre anvendelser enn spill.
- Litt statistisk metodelære bør med for å gi elevene en forståelse av forholdet mellom populasjon og utvalg og av at en fra et utvalg kan få informasjon om populasjonen. Det bør imidlertid ikke være med mye statistisk metodelære, delvis på grunn av den kompetansen matematikklærerne har.

- Mange lærere gjorde en stor innsats i forbindelse med Reform-94 ved å sette seg inn i sannsynlighetsregning og statistiske metodelære. Det er viktig å respektere denne innsatsen og unngå omfattende endringer i læreplanen der det er mulig.

Den nye læreplanen for grunnkurset ble vedtatt i juli 1999 og trådde i kraft i inneværende skoleår. I mitt innlegg i TG nr. 1 i år gjenga jeg hovedmålet om sannsynlighetsregning. For fullstendighetens skyld gjentar jeg dette her:

Mål 5: Elevene skal kjenne grunnbegrepene i sannsynlighetsregningen og kunne bruke dem til å utføre enkle beregninger.
Elevene skal
5a kjenne til begrepet sannsynlighetsmodell og kunne formulere og eksperimentere med enkle uniforme og ikke-uniforme sannsynlighetsmodeller
5b kunne regne ut sannsynligheter ved å telle opp alle gunstige og alle mulige utfall i enkle eksempler
5c kunne regne ut sannsynligheter ved hjelp av valgtrær, Venn-diagrammer og andre systematiske oppstillinger
5d* ha en intuitiv forståelse av uavhengighet og betinget sannsynlighet
5e* kunne bruke addisjonssetningen og produktsetningen
(De to stjernemerkete punktene er ikke obligatoriske for yrkesfagene.)

Da jeg skrev mitt innlegg i TG nr. 1 i år, hadde arbeidet med de nye læreplanene for videregående kurs I og II for allmenne, økonomiske og administrative fag (dvs. 2MX, 2MY, 3MX og 3MY) bare så vidt startet opp. I arbeidet med disse planene møtte jeg stor forståelse for synspunktene nevnt over. Spørsmålet om sannsynlighetsregningens og statistikkens plass i læreplanene ble derfor ikke et stridstema slik jeg hadde fryktet. De nye læreplanene i 2MX, 2MY, 3MX og 3MY ble vedtatt i juli i år etter en høringsrunde hvor også NSF kom med kommentarer. Planene trer i kraft høsten 2001 (for 2MX og 2MY) og høsten 2002 (for 3MX og 3MY). De har alle ett hovedmål i sannsynlighetsregning og statistikk. Målene for X- og Y-versjonene er forholdsvis like, men målene for X-versjonen er litt mer ambisiøse:

2MX
Kombinatorikk og sannsynlighetsregning
Elevene skal kjenne grunnbegrepene i kombinatorikk og sannsynlighetsteori, kunne vurdere utsagn om sjanser og sannsynligheter, og kunne løse praktiske problemer ved hjelp av sannsynlighetsregning
Elevene skal
6a kjenne begrepene uavhengighet og betinget sannsynlighet, og kunne bruke Bayes' setning på to hendelser
6b kunne behandle ordnede utvalg med og uten tilbakelegning, uordnede utvalg uten tilbakelegning, og kunne bruke dette til å beregne sannsynligheter
6c kunne regne med binomiske og hypergeometriske sannsynligheter

2MY
Mål 5: Sannsynlighetsregning
Elevene skal kjenne grunnbegrepene i sannsynlighetsteori, kunne vurdere utsagn om sjanser og sannsynligheter, og kunne løse praktiske problemer ved hjelp av sannsynlighetsregning
Elevene skal
5a kjenne begrepene uavhengighet og betinget sannsynlighet og kunne bruke Bayes' setning på to hendelser
5b være kjent med ordnede utvalg med og uten tilbakelegning, uordnede utvalg uten tilbakelegning, og kunne gjøre enkle sannsynlighetsberegninger knyttet til slike utvalg
5c kunne regne med binomiske og hypergeometriske sannsynligheter

3MX

Mål 7: Sannsynlighetsregning og statistikk

Elevene skal kunne regne med sannsynlighetsfordelinger, forventning og varians og kjenne til sammenhengen mellom modellparametere og parameterestimat.

Elevene skal

7a kjenne begrepene fordeling og stokastisk variabel for endelige utfallsrom og kunne finne forventning, varians og standardavvik

7b kjenne regnereglene for forventningen til en lineær kombinasjon av to stokastiske variable og variansen til en lineær kombinasjon av to uavhengige stokastiske variable

7c kjenne forventning og varians til en binomisk fordeling

7d kjenne normalfordelingene og kunne regne ut sannsynligheter knyttet til normalfordelinger

7e ha kjennskap til den sentrale grensesetningen og kunne beregne sannsynligheter ved å tilnærme binomiske fordelinger med normalfordelinger

7f forstå at utvalg fra en stor populasjon kan betraktes som uavhengige gjentak og kunne beregne punktestimat med standardfeil og tilnærmet konfidensintervall for en populasjonsandel og et populasjonsgjennomsnitt.

3MY

Mål 6: Sannsynlighetsregning og statistikk

Elevene skal kjenne begrepene fordeling og stokastisk variabel, kunne tilnærme binomiske fordelinger med normalfordelinger og vite litt om sammenhengen mellom modellparametere og parameterestimat.

Elevene skal

6a kjenne begrepene fordeling og stokastisk variabel for endelige utfallsrom og kunne finne forventning, varians og standardavvik

6b kjenne forventning og varians til en binomisk fordeling

6c kjenne normalfordelingene og vite litt om deres betydning

6d kunne tilnærme binomiske fordelinger med normalfordelinger

6e forstå at utvalg fra en stor populasjon kan betraktes som uavhengige gjentak og kunne beregne punktestimat med standardfeil og tilnærmet konfidensintervall for en populasjonsandel

Personlig er jeg godt fornøyd med de nye læreplanene både for grunnkurset og de videregående kursene. Ved at vi nå har fått planer med en bedre indre sammenheng og et klarere presisjonsnivå, håper jeg også at lærebokforfatterene får en lettere jobb enn de hadde da Reform-94 planene skulle settes ut i livet.

De nye og bedre planene må ikke fungere som en "sovepute" for oss statistikere. For læreplaner er en ting, hvordan de fungerer i praksis er ofte en annen. Det er derfor viktig at vi fortsatt engasjerer oss i det som skjer i skolen. Dette betyr at etterutdanning av lærere fortsatt bør være en prioritert oppgave. Etter de nye planene skal elvene hvert år gjennomføre et lite prosjekt i matematikk. Det hadde vært fint om NSF kunne få fram gode ideer til slike prosjekt og formidle disse til lærere og elever via web!

PS:

Hvis du er interessert i å se læreplanene i matematikk i sin helhet, kan du finne en "portal" til læreplaner i videregående skole på

`http://skolenettet.nls.no/prgdb/owa/lp.lp_forside`

Fra denne finner du grunnkurset i matematikk under "Felles: Felles allmenne fag" og de videregående kursene i under "Studieretninger: Allmenne, økonomiske og administrative fag".

Matematikk 2MX

30. mai 2000

OPPGAVE 2

a) Regn ut på hvor mange måter vi kan velge ut 5 barn av en gruppe på 12.

Vi antar at gruppen på 12 inneholder like mange gutter som jenter.

b) Hvor mange utvalg kan vi gjøre hvis utvalget skal inneholde 2 gutter og 3 jenter?

c) Hvor mange utvalg på 5 kan vi gjøre hvis utvalget skal inneholde minst 3 gutter?

Ved en skole spaserer $\frac{1}{5}$ av elevene til skolen hver dag. Vi velger ut 3 elever tilfeldig.

d) Regn ut sannsynligheten for at minst 2 av disse spaserer til skolen hver dag.

Matematikk 2MY

30. mai 2000

OPPGÅVE 1

d) Ved ein skule spaserer $\frac{1}{5}$ av elevane til skulen kvar dag. Vi vel ut 3 elevar tilfeldig. Rekn ut sannsynet for at minst 2 av desse spaserer til skulen kvar dag.

OPPGÅVE 4

I ein moropark får alle barn lov til å trekkje fem lodd ved inngangen. Ein dag blei parken besøkt av 4 518 barn. Tabellen nedanfor viser kor mange barn som ikkje vann eller som vann på eitt av lodda, to av lodda osb.

tal på gevinstar	0	1	2	3	4	5
tal på barn	2613	1556	298	50	0	1

a) Kor mange gevinstar fekk kvart barn i gjennomsnitt?

Fabrikken som lagar lodda gir opp at 10 % av lodda gir gevinst.

b) Rekn ut sannsynet for at eit barn som får fem tilfeldige lodd, vinn på 0, 1, 2, 3, 4 eller 5 av dei.

c) Samanlikn resultatet i b) med tabellen ovanfor. Synest du det er godt samsvar mellom dei?

Matematikk 3MX

Matematikk 3MY

30. mai 2000

OPPGAVE 1

c) En stokastisk variabel har sannsynlighetsfordeling som vist i tabellen nedenfor:

X	0	1	2	3	4
$P(X = x)$	$\frac{1}{6}$	$2k$	$\frac{1}{4}k$	$\frac{1}{3}$	$\frac{1}{4}k$

Regn ut k og finn $P(0 \leq X < 3)$.

OPPGAVE 4

I følge produsenten vil en viss type fyrverkeriraketter virke i 98 % av tilfellene. Ved et jubileum skal det sendes opp 250 slike raketter.

a) La X være antall raketter som virker. Finn forventningsverdien og standardavviket til X.

En annen type fyrverkeri blir brukt ved samme anledning. Av denne typen virket 189 av 200.

b) Finn et 95 % - konfidensintervall for det gjennomsnittlige antall raketter som virker i en stikkprøve på 200.

Fyrverkeriprodusenten overveier å endre litt på konstruksjonen av de rakettene som hittil har virket i 98 % av tilfellene. Han planlegger å lage en prøveproduksjon på 3000 av den nye rakettypen, og gir deg ansvaret for å utføre en hypotesetest for å undersøke om den nye typen virker i mer enn 98 % av tilfellene.

c) Forklar hvordan du vil gå fram. Bruk 5 % - signifikansnivå.

d) Det viser seg at 52 av de 3000 nye rakettene ikke virket. Hvilket resultat gir testen din?

Gjennomslag for statistikkurs på nett

– Mest ingen melde seg på den tradisjonelle klasseundervisninga i kurset i medisinsk statistikk for ingeniørar. Men då me tilbaud det same kurset som nettbasert kurs, melde over 30 seg på.

MARTIN TOFT
(TEKST OG FOTO)



LETT PÅ NETT: – Det er enklare å passa inn kurset i medisinsk statistikk i arbeidsdagen når det er nettbasert, er både kursleiarar og kursdeltakarar einige om. Frå venstre Petter Laake, Kristin Sand, Maje Siebke, Harald Fekjær og Aasa Schølling.

Det fortel dei kursansvarlege, forskar Harald Fekjær og professor Petter Laake, Seksjon for medisinsk statistikk og representant for Opplæringskomiteen ved UiO, avdelingsingeniør Maje Siebke.

Det internettbaserte kurset i medisinsk statistikk starta opp i haust med ein introduksjonsdag der kursleiarane gjekk gjennom opplegget og programpakka. På førehand hadde deltakarane fått tilsendt ein CD som sette dei i stand til å koma i gang med sjølve kurset, som elles er sett saman av ei blanding av skriftleg materiell, øvingar og oppgåver. Kursdeltakarane går gjennom åtte tema med ein arbeidsperiode på to veker for kvart tema.

Ingen ser kvarandre

– Alt det pedagogiske materiell ligg ute på Internett, og alle oppgåvene kursdeltakarane skal løysa, blir lagde ut ein gong i veka, fortel forskar Harald Fekjær, som er ansvarleg for å leggja oppgåvene ut på nettet. Han syter også for at alle kursdeltakarane får eit passord slik at dei kan koma inn på kurssidene på Internett. Laake synest det er heilt greitt å undervisa på denne måten, sjølv om verken han eller kursdeltakarane ser kvarandre.

– Alle studentane sender svara på oppgåvene til meg på e-post, og så sender eg dei tilbakemelding på det dei har gjort. Ved hjelp av mobiltelefonen min kan eg også sitja på eit hotellrom i Roma og halda kontakt med kursdeltakarane via e-post, seier ein oppglødd Laake, som saman med Fekjær har rekna ut at deltakarane brukar 5-7 timar i veka på kurset.

Er i full jobb

– Sidan det er eit nettbasert kurs, kan dei bruka fritida si på kurset og vera i full jobb samtidig. Dei gongene me hadde klasseundervisning, måtte deltakarane ta fri heile dagar for å kunna fylgja kurset. Det er nok også grunnen til at så mange melde seg på dette kurset etter at det blei nettbasert, seier Maje Siebke. Likevel var det naudsynt å driva ein del korridorpolitikk for å gjera folk merksame på dette kurstilbodet. Også kursdeltakarane Kristin Sand og Aasa Schølling er godt nøgde med at bioingeniørar får tilbod om eit slikt kurs.

– Det er jo oftast me ingeniørar som står i laboratoriet og utgreiar alle dei medisinske data til medisinprofessorane sine vitskaplege artiklar. Likevel er mange av dei imot at me skal leggja dei data me finn fram til inn i statistikken. Med eit slikt kurs i medisinsk statistikk vil me vera enda meir kvalifiserte til å gjera dette arbeidet, er både Sand og Schølling einige om. Ingen av dei saknar den tradisjonelle klasseromsundervisninga. Laake trekkjer også fram behovet for undervisning og vidareutdanning av ikkje-vitskapleg tilsette.

Høgt fagleg nivå

– Undervisning i statistikk for ingeniørane er viktig fordi dei arbeider i grupper med vitskapleg tilsette. Alle ingeniørane som deltar på dette kurset viser stor arbeidsinnsats, og dei har eit høgt fagleg nivå, slår Laake fast.

Det første nettbaserte kurset i medisinsk statistikk blei utvikla saman med Pedagogisk avdeling hos Lægeforeningen, opplyser Fekjær til Uniforum. Også neste år kjem Opplæringskomiteen ved UiO til å skipa til eit nytt kurs.

– Kurset gir ei kompetanseheving av deltakarane som kjem fagmiljøa ved UiO til gode. Det er viktig å hugsa på at kurset er gratis for alle som er tilsette ved UiO. Det er berre dei eksterne deltakarane som treng betala kursavgift, fortel Siebke.

Fra Uniforum 16. november 2000.

ERLING SVERDRUPS RADIOFOREDRAG FOR 45 ÅR SIDEN

I serien "Universitetenes radioforedrag" bidro Erling Sverdrup med tre foredrag under tittelen "Hva kan vi slutte av statistikk?". Manuskriptene til disse foredragene ble funnet da Knut Håkon Brox ryddet i Sverdrups etterlatte papirer for godt og vel ett år siden. De to første foredragene ble trykket i henholdsvis mai- og septembernummeret av TG. På de neste sidene i dette nummeret gjengis det siste av foredragene i sin originalform.

Hva kan vi slutte av statistikk ?

Erling Sverdrup.

Tredje foredrag.

Vi omtalte i siste foredrag hvilken rolle forutsetningene spiller ved statistiske slutninger. Og med forutsetninger tenkte vi på denslags opplysninger som er kjent på forhånd før vi begynner å innsamle eller studere det statistiske materiale. Ofte er denslags forhåndsviden almengyldige resultater som man har kommet frem til på grunnlag av tidligere erfaringer, erfaringer som ikke alltid er systematisk nedtegnet, men som kan være solide nok for det. De representerer gjerne den faglige innsikt på det området som statistikken er hentet fra. Det er en viktig oppgave for statistikeren i samarbeide med de som har en slik faglig innsikt, å vurdere hva denne forhåndsviden bør omfatte, d.v.s. hvilke forutsetninger som bør gjøres. Statistikeren må så oversette disse forutsetninger til det sproget som brukes i sannsynlighetsregningen. Først etter at dette er gjort kan man utnytte det statistiske materiale til å vinne ny innsikt. Vi skal i dette fordrag se litt på denne del av statistikerens arbeide. Vi skal m.a.o. se på fremgangsmåten ved tilveiebringelsen og tolkningen av statistikk.

Vi skal først si noen ord om den historiske utvikling som har ført frem til de metoder vi har i dag. Vi nevnte sist teoriene til Laplace og Gauss fra begynnelsen av det 19. århundre om den statistiske behandling av fysikalske måleresultater. Mot slutten av det 19. århundre skjedde det noe viktig når det gjaldt sannsynlighetsregningens anvendelse i fysikken. Sannsynlighetsregningen ble av fysikerne Boltzmann, Maxwell og Gibbs tatt i bruk til å forklare fenomener omkring gasser. I en gass er som bekjent molekylene i stadig bevegelse og kolliderer mot hverandre - samt mot gassens

begrensning hvis gassen er innelukket i en beholder. På grunn av det store antall molekyler i en gass ville det være uoverkommelig ~~oppgave~~ å behandle hvert enkelt molekyls bevegelser og kollisjoner. I stedet søker man å beregne den gjennomsnittlige veilengde molekylene tilbakelegger mellom kollisjonene, det gjennomsnittlige antall kollisjoner o.s.v. Ved disse beregninger kommer ganske naturlig sannsynlighetsregningen til anvendelse og det på en ganske anderledes ~~betydningsfull~~ grundig måte enn ved behandling av observasjonsfeil ved fysikalske målinger.

Det ligger utenfor rammen av disse foredrag å komme inn på dette. Det som er viktig å peke på i denne forbindelse er at inntil slutten av det 19. århundre ble sannsynlighetsregningen i det vesentlige anvendt på tre områder: For det første befolkningsstatistikk og livsforsikring, for det annet fysikalske måleresultater og for det tredje de fysikalske teorier for gasser.

På alle disse områder opererte man med store tall. Store befolkningsgrupper og store forsikringsbestander, et stort antall fysikalske målinger, mange molekyler i en gass. Det var derfor ikke så rart at den overbevisning festnet seg at forat det skulle ha noen hensikt eller mening å bruke sannsynlighetsregning måtte man nødvendigvis ha store tall.

Endringen av denne innstilling kam etterhvert gjennom en utvikling som begynte i England omkring år 1900. Det var en utvikling som i sannhet kan betraktes som oppsiktsvekkende, ikke minst fordi sannsynlighetsregning ble tatt i bruk ved behandling av observasjonsmaterialer fra andre områder enn de jeg nettopp nevnte. I første omgang gjaldt dette observasjonsmaterialer fra landbruk og biologi. Der oppsto en såkalt biometrisk retning hvis ledende personlighet var professor Karl Pearson. Innen denne skole vokste det etterhvert fram en rekke betydelige navn. En av disse var Gosset, kanskje mest kjent under pseudonymet Student. Han var ølbrygger ved de store Guinnes bryggerier i Storbrittania. Hans interesser og anlegg for matematikk gjorde at han ble interessert i bearbeidelse av de statistiske materialer fra bryggeriets laboratorier samt fra forsøksdriften ved bryggeriets landeiendommer. Hans statistiske undersøkelser har tjent som mønster på hvorledes slike undersøkelser skal legges an og de har betydd meget for ettertiden. Med sitt kjennskap til de underliggende realiteter, d.v.s. til produksjonsprossessen var det mulig for ham etter grundige overveielser og vurderinger å stille opp virkelighetstro matematiske modeller. Når han hadde latt sin praktiske erfaring fagkunnskap komme til uttrykk i en matematisk modell basert på sannsynlighetsregning, kunne han avlede statistiske metoder til å

vinne ny innsikt. Gosset brøt radikalt med den tradisjon at sannsynnlighetsregningen bare kunne anvendes når det statistiske materiale er stort. Som jeg ~~fremhevet~~ nevnte i første foredrag førte denne tradisjon til en unødig innskrenkning i mulighetene for bruk av sannsynlighetsregning. Det er nemlig klart at et lite statistisk materiale innsamlet ved omhyggelig planlegging ofte kan si adskillig mer enn et stort materiale mer vilkårlig innsamlet. Gosset grunnla teorien for den sannsynlighetsteoretiske behandling av små statistiske materialer og for moderne planlegging av statistiske forsøk. Det er særlig en avhandling fra 1908 som er viktig.

~~Gosset var forøvrig i sitt firmas tjeneste hele sitt liv bare avbrutt av tilfeldige gjesteopptredener ved universitetene. Fra 1935 ble han sjef for Guinness London-fabrikker og fikk da litbn tid til å arbeide med matematisk statistikk.~~

Gossets arbeide (i 1920 og 1930 årene) ble videreført av professor R.A. Fisher hvis omfattende og epokegjørende arbeider spenner over vidt forskjellige områder av statistikken. Fisher var opprinnelig sjefstatistiker ved den kjente Rothamsted landbruksstasjon og han er derfor inspirert av de statistiske problemer som oppstår innen landbruksforskning. Men for Fisher var det klart at de problemer som han behandlet ikke var enestående for landbruk. De oppstår i medisin, biologi, pedagogikk, psykologi, meteorologi, teknisk og industriell forskning og på mange andre felter. Etterhvert er derfor hans teorier blitt anvendt på alle disse områder. Det viser seg nemlig at det er visse typiske statistiske problemer som går igjen i mange forskjellige vitenskaper.

Man kunne nesten være fristet til å si at Laplace's over 100 år gamle drøm om sannsynlighetsregningen som et fundamentalredskap ved forskning i natur og samfunnsliv var blitt virkeliggjort og det av en mann som sto midt opp i det praktiske arbeide. Nå vel, det ville være å overdrive litt, men at Fisher sprengte alle tidligere rammer for sannsynlighetsregningens anvendelse er klart.

Vi skal ved noen eksempler antyde statistikerens arbeidsmåte når han utfra gitte forutsetninger søker å finne frem til gode metoder til å analysere statistiske materialer. Vi henter det første eksempel fra kontroll av industrielle produksjonsprosesser.

Vi vil tenke oss en maskin som masseproduserer en artikkel. Den produserer, la oss si, mange tusen artikler om dagen. Det kan f.eks. være skruer, nagler eller lignende. Det viser seg uoverkommelig eller ulønnsomt til stadighet å foreta omhyggelig kontroll av råstoffene eller maskin. Og det lønner seg heller ikke å kontrollere hvert enkelt produkt. I stedet vil fabrikken ved slutten av hver arbeidsdag velge tilfeldig noen av de produserte artikler og kontrollere dem. Formålet er å bringe på det rene om produksjonsprosessen bør avbrytes og maskiner og råstoffer kontrolleres. Nå viser det seg at når produksjonen løper helt mormalt vil likevel en viss prosent av artiklene være ubrukbare. Det kan f.eks. tenkes at 1 o/o vil være ubrukbare. Dette er altså hva produsenten må finne seg i og han ønsker i så fall at maskinen skal løpe uforstyrret. Statistikerens problem er å anbefale en slik fremgangsmåte at fabrikken i minst mulig utstrekning griper inn når proessen forløper normalt og i størst mulig utstrekning griper inn når noe er galt med prosessen. Statistikeren må altså for de fremgangsmåter som kan komme på tale beregne sjangsen for å gripe inn i normal produksjon og sjangsen for ikke å gripe inn når noe er galt. Han ønsker å holde begge disse sjangser små. Samtidig vil han naturligvis innskrenke testingen til det minst mulige, d.v.s. han ønsker å greie seg med så få prøveartikler som mulig. Det blir en vurderings sak å veie de forskjellige hensyn mot hverandre.

Slike overveielser er ikke enestående ved industriell prosesskontroll. Vi må resonnere på samme måte ved andre slags statistiske undersøkelser.

Ta f.eks. et medisinsk materiale til å belyse virkningen av en

bestemt slags vaksine ~~mto~~ mot tyfus. Det viser f.eks. at av de ikke-vaksinerte ble 2.3 o/o angrepet, mens av de vaksinerte ble 0.8 o/o angrepet. Spørsmålet er om forskjellen er så stor at den beviser at vaksinen mot tyfus er virkningsfull eller ikke. Anta nå at vi har fastlagt en bestemt metode til å avgjøre dette spørsmål. Hvorledes skal vi vurdere om metoden er fornuftig ? All den stund vi ikke på forhånd vet om vaksinen er virkningsfull eller ikke, det er jo det som skal avgjøres, er det naturlig å stille følgende spørsmål: Hva er sjangsen forat metoden skal føre til at vi påstår at vaksinen virker når den faktisk er virkningsløs ? Hva er på den annen side sjangsen for at vi skal si at vaksinen ikke har virket, når den faktisk er effektiv ? Begge disse sjangser for feiltolkning av materialet bør ~~naturligvis~~ helst være små. Nå har vi naturligvis mange metoder å velge mellom når vi skal analysere materialet. Vi må velge ut den som er gunstigst bedømt utfra sjangsene for feil-tolkning.

Som et tredje eksempel la oss se på et statistisk problem som har opptatt meteorologer meget. Det gjelder bestemmelsen av mulige perioder i nedbørmengde eller temperatur. La oss som eksempel tenke på variasjonene i gjennomsnittlig sommertemperatur fra år til år. Er det slik at vi har årrekker med kalde somre etterfulgt av årrekker med varme somre i regelmessig veksling ? Det kan f.eks. påståes at det er reglemessig veksling på en slik måte at varme somre kommer igjen med 30 års mellomrom. Da sier vi at det er en 30-års periode i sommertemperaturens variasjon. Men i tillegg til en slik periode kan vi ha andre kortere eller lengre perioder og vi kan ha tilfeldige variasjoner som forstyrrer det enkle billede. Spørsmålet er nå om statistikken over sommertemperaturens veksling tyder på at det er slike perioder. Det har vært brukt metoder til å avgjøre dette spørsmål som er slik at de høyst sannsynlig vil avsløre en periode hvis en slik eksisterer og det er jo vel og bra.

Men en nærmere undersøkelse av metoden har vist at hvis det ikke er noen periode og variasjonene bare er tilfeldige, vil metoden også anvise en periode, altså en falsk periode, og det er naturligvis ikke så bra. Når metoden anviser en periode vet vi altså ikke om den er falsk eller ekte og vi er kommet like langt. Straks man ble klar over dette ble naturligvis metoden forkastet. I stedet blir nå oppgaven stillet slik:

Finn en metode som er slik at den gjør både sjangsen for å påstå falske perioder og sjangsen for ikke å avsløre reelle perioder liten, når man har temperaturrekker over et tidsrum av rimelig lengde. Dette spørsmål er fremdeles uløst, men under visse mer spesielle forutsetninger har man kunnet løse slike spørsmål, men generelt må man si at spørsmålet ennå er uløst.

Vel, dette var tre eksempler til belysning av problemstillingen i statistikken. Statistikerne opererer altså med sjangser for feiltolkning. De innrømmer altså mulighetene av feiltolkning og noen vil kanskje påstå at det viser at de har rett de som sier at å trekke noen slutninger av statistikk det er en ytterst tvilsom affære. Til dette kan det svares at ethvert forsøk på å si noe almengyldig på grunnlag av et begrenset erfaringsmateriale er forbunnet med risiko. Det er alltid muligheter for feil, enten vi forbigår dette faktum i taushet eller ikke. Det faktum at vi så å si griper tyren ved hornene og uttrykkelig undersøker sjangsene for feiltolkninger, gjør naturligvis metodene påliteligere.

Vi har hittil omtalt statistikerens arbeidsmåte når det gjelder å tolke et gitt statistisk materiale. Men statistiske overveielser kommer naturligvis også inn når det gjelder planleggingen av hvorledes det målte eller talte skal bringes tilveie. Vi skal se på et eksempel for å illustrere hva slags betraktninger en statistiker må foreta under planleggingen.

Anta at man ønsker å vite hvilket av to gjødningsstoffer er det

beste. Vi vil utføre et dyrkingseksperiment og vi disponerer over en stripe land som er, la oss si, 10 meter bred og 100 meter lang. Denne striben inndeles i 10 like store ruter ved tvergående skillelinjer. Hver rute blir altså på 10 ganger 10 meter. Nå skal vi spre gjødningene i disse ruter, i 5 av rutene skal vi bruke den ene gjødning og i de 5 andre, den andre gjødning. Deretter skal vi så havre og høstutbytte skal så måles for hver enkel rute. Spørsmålet er nå i hvilke 5 ruter skal vi spre den ene gjødning og i hvilke 5 ruter den annen.

Den mest nærliggende tanke er vel å spre den ene sort gjødning i de 5 første ruter og den andre sort gjødning i de 5 etterfølgende ruter. Dette betyr i realiteten at vi ~~her~~ inndeler stripen på 100 meter i to striper, hver på 50 meter og lar de to striper få hver sin gjødning. Men nå er det klart at hvis høstutbytte i den første stripe var større enn i den andre, kunne vi ikke derav slutte at det skyldes gjødningen. Det kunne like gjerne skyldes egenskaper ved jorden som ikke har noe med gjødningen å gjøre. Nesten like store vanskeligheter ville vi være i hvis vi i annenhver rute i stripen brukte den ene gjødning og i annenhver den andre gjødning. Så lenge vi ikke har noe målt uttrykk for jordens kvalitet må vekslingen i jordens kvalitet betraktes som ukjent. Det kunne jo tenkes at jordens kvalitet vekslet i perioder i takt med rutene.

Men anta nå at vi trakk lodd om hvilke 5 ruter skulle ha den ene gjødning og hvilke 5 ruter den anden gjødning. Kanksje man vil påstå at man ikke har oppnådd noe mer ved det. Det kunne jo tenkes at loddtrekning falt så uheldig at vi kom til å dyrke den ene gjødning i det 5 ruter hvor jordens kvalitet var høyest og den annen gjødning i de 5 ruter hvor jordens kvalitet var lavest. Da ville man tilsynelatende være i samme vanskelighet som ved de andre fremgangsmåter som jeg nevnte. Likevel er det en viktig forskjell. Hvis man trekker lodd om hvor man skal bruke den ene og hvor man

skal bruke den andre gjødning vil det være tilfeldig om den ene gjødning ble spredt på gode eller dårlige akerlapper. Da ville altså innflytelsen av jordens kvalitet på resultatet være endel av de tilfeldige variasjoner. Men de tilfeldige variasjoner er noe vi behersker med det verktøy som sannsynlighetsregningen gir oss. For å si det på en annen måte. Vi behøver ikke lenger regne med noen systematisk virkning av jordens kvalitet ved de tallene vi kommer frem til.

Nå kan altså sannsynlighetsregningen brukes til å bearbeide tallene på samme måte som f.eks. ved kontroll av produksjons-prosesser.

Denne idé med loddtrekning synes enkel, men den bryter i virkeligheten ganske radikalt med de forestillinger som vært mange ~~mennesker~~ har om eksperimentell forskning. Hvor ofte hører man ikke følgende uttalelser: I den og den vitenskap er det umulig å slutte noe pålitelig om virkningen av de forskjellige faktorer. Vi kan nemlig ikke gjøre som i fysikerens laboratorium. Holde alle faktorer unntagen én fast. Til dette er å svare at også fysikerens målinger i laboratoriet er beheftet med såkalte tilfeldige feil. Ved loddtrekningen er det i virkeligheten lykkes å komme på linje med fysikeren. Høstutbytte i eksperimentet med jorddyrkningen er bare påvirket av en faktor, nemlig gjødningen. I tillegg hertil kommer bare de tilfeldige variasjoner.

Uten å komme nærmere inn på det kan jeg nevnte at også ved planleggingen må man se på sjangsene for feiltolkninger, slik som vi gjorde det ved analysen av et foreliggende statistisk materiale.

Som en avslutning på disse foredrag skal jeg si noe mer om hva det egentlig ligger i dette at vi ønsker sjangsene for feil-tolkning små.

At sjangsene for feiltolkning er små betyr at hvis vi rent

hypotetisk tenkte oss undersøkelsen gjentatt et meget stort antall ganger under samme omstendigher, vil vi som regel komme frem til en konklusjon som er i samsvar med realitetene. ~~Det er dette som vanlig~~ Vi kan kalle denne fremgangsmåte ~~vis kalles~~ "sampling-synspunktet". Det kommer av det engelske ord "sample" som betyr prøve eller utvalg. Det statistiske materiale er en prøve blandt de tenkte materialer ved de tenkte gjentagelser.

Det er klart at ~~for samfunnet er det~~ det er verdifullt at statistikerne ~~har denne innstilling~~ bruker samplingssynspunktet. Det ofres i dag mange arbeidstimer på praktisk-tekniske undersøkelser hvor statistiske metoder anvendes. Usikkerhetsmomenter ved slike undersøkelser kan aldri helt elimineres. Oppgaven må derfor være å sikre at man som regel får et fornuftig resultat uten for store omkostninger. Nå vil det naturligvis som regel ikke være slik som vi tenkte oss at en statistisk undersøkelse blir gjentatt under samme omstendigheter. Men ved å oppfatte hver enkelt undersøkelse som en av en rekke hypotetiske gjentakelser, sikrer vi oss i virkeligheten at vi blandt den brogede mangfoldighet av forskjellige statistiske undersøkelser som regel vil trekke fornuftige slutninger uten for store utgifter for samfunnet.

Det er analogt med et livsforsikringsselskap. Når aktuaren skal regne ut premien for en forsikringskontrakt, tenker han seg hypotetisk et selskap som tegner et stort antall identiske kontrakter og han ansetter premien slik at dette tenkte selskap er noenlunde sikret mot underskudd. I virkeligheten er det i et selskap ~~enn~~ en broget mangfoldighet av kontrakter og meget få som er identiske. Men ved for hver enkel kontrakt å resonere slik som jeg nevnte, har ~~han~~ aktuaren i virkeligheten sikret det konkrete selskap som han arbeider for mot underskudd.

Statistikeren må altså ved hver enkelt statistisk undersøkelse finne den metode som med små omkostninger gjør slutningene så sikre som mulig. Små omkostninger betyr lite statistikk. Han må også i hvert enkelt tilfelle vurdere hvorvidt statistikk kan si ham noe.

At det er en statistiker i farvannet betyr derfor ikke nødvendigvis flere tall, skjønt som oftest vil det vel det. Det er nemlig utrolig ofte slik at det man ønsker å vite ligger like rundt hjørnet, hvis man bare vil bruke litt tall og statistikk.

Malignant melanoma, lung cancer, and prostate cancer in relation to diet. A Norwegian prospective study emphasizing fat intake and methodological issues of measurement errors

Marit Bragelien Veierød
Sted: Seksjon for medisinsk statistikk, Det medisinske fakultet, UiO. *Disputas:* 01.09.2000
Grad: Dr. philos *Finansiering:* Den Norske Kreftforening, Den norske lægeforening
Veiledere: Petter Laake, Dag S. Thelle, Christian A. Drevon

Avhandlingen faller innenfor fagområdene kreft- og ernæringsepidemiologi og medisinsk statistikk. Sammenhengen mellom kosthold og malignt melanom (ondartet føflekkreft) og mellom fettinntak og kreft i lunger og prostata er studert i en norsk kohort bestående av 50 000 personer rekruttert i 16-56 års alder. Kostholdsdata vil alltid være beheftet med målefeil/feilklassifisering. Konsekvensen av dette for estimert sammenheng med kreftsykdom er studert ved hjelp av både stokastisk simulering og matematisk statistisk tilnærming til problemet.

Dataene er hentet fra Statens Helseundersøkelser sin hjertekar-undersøkelse i Finnmark, Sogn og Fjordane og Oppland i 1977-83. Alle deltagerne fikk utdelt et spørre-skjema om kosthold utviklet ved Institutt for ernæringsforskning, UiO. Skjemaet inne-holder spørsmål om konsum av vanlige mat- og drikkevarer med særlig vekt på de viktigste fettkildene. Kobling mot Kreftregisteret ga forekomst av kreft, mens opplysninger om død og emigrasjon ble hentet fra Statistisk Sentralbyrå. Kohorten er fulgt frem til 31/12-1992. Det ble registrert h.h.v. 108, 156 og 72 nye tilfeller av malignt melanom, lunge- og prostatakreft. Poisson regresjonsanalyse ble anvendt for å estimere sammen-hengen mellom kostholdsfaktorer og risikoen for kreft.

Noen få kostholdsfaktorer var assosiert med malignt melanom, men bare for kvinner. Vi tror dette kan forklares via sammenheng med andre ikke-registrerte faktorer som f.eks. solvaner. Vi fant ingen sammenheng mellom lungekreft og inntak av kolesterol og mettet fett, men redusert risiko ved bruk av skummet melk sammenlignet med helmelk og ved bruk tran. En beskyttende effekt av tran kan skyldes innholdet av omega-3 fettsyrer, men bør etterprøves i senere studier. For prostatakreft støtter denne studien ikke enkelte tidligere studier som har funnet positiv sammenheng med fettinntak, spesielt animalsk fett. Kohorten er fortsatt ung og andre studier har foreslått at kosthold i eldre år kan være av større betydning for risikoen for prostatakreft.

Ved analyse av sammenhengen mellom fettinntak og kreftsykdom deles vanligvis observert fettinntak inn i kvartiler. Ved hjelp av stokastisk simulering illustreres det hvordan de tre kvartil-koeffisientene estimert ved Poisson regresjon kan bli betydelig svekket på grunn av målefeil i fettinntaket. Svekning av sammenhengene er vanligst, men det illustreres også hvordan observerte sammenhenger kan bli styrket p.g.a. målefeil.

Det utledes en metode for å vurdere skjevhet i Poisson regresjonskoeffisienter for sammenhengen mellom sykdom og eksponering når sistnevnte er kategorisk og utsatt for feilklassifisering. Hver person i studien har en sann eksponeringskategori (typisk ukjent) og en observert eksponeringskategori. Kjennskap til førstnevnte ville gitt sanne regresjonskoeffisienter, mens analysen av observert eksponering gir såkalte naive regresjonskoeffisienter. Det utledes eksplisitte uttrykk for de naive koeffisientene som funksjon av de sanne koeffisientene og feilklassifiseringssannsynlighetene. Dersom vi antar uavhengighet mellom feilklassifiserings-sannsynlighetene og eventuelle justeringsvariable, er uttrykkene for de ujusterte og de justerte naive koeffisientene identiske. Tilsvarende utledes også for lineær regresjon. Utvalgte eksempler med både symmetrisk og asymmetrisk feilklassifissering samt forskjellige does-responsmønstre, illustrerer hvor vanskelig det er å forutsi konsekvensen av feilklassifisering. Både størrelsen og retningen på skjevheten i estimatene kan variere mellom eksponeringskategoriene.

Her er sportsnytt!

Hva er forskjellen på Edel Therese Høiseth og Jørund Gåsemyr? Vel: I tillegg til at frk. Høiseth aldri har fått en artikkel publisert i prestisjetidsskriftet *Annals of Statistics*, har Jørund deltatt i *flere* olympiske leker enn Edel Therese. Høiseth er med rette kjent for sine multiple olympiadiske sprintløp, med deltagelse i de siste fem vinter-OL (som kjent med plasseringene 24, 14, 20, 8, 20 på 500 m og 20, 15, 13, 26, 14 på 1000 m).

I dette lys er Jørunds prestasjoner enda mer multiple og enda mer imponerende – og hans plasseringer enda mer lysende. Han er nettopp tilbake fra Sydney Paralympics, der det ble en 4. plass i favorittøvelsen maraton med tiden 2:50:31. Gåsemyr er formodentlig (det internasjonale) statistikkfagets eneste olympiske mester. Han vant maraton i Barcelona 1988 med tiden 2:45:48. Ellers har Jørund som Grete Waitz en bakgrunn også i kortere løp, med sølv- og bronsemedaljer på 1500 og 5000 fra 1980 og 1984.

Jostein Lillestøls «puslerispalte» kan for denne anledning suppleres med følgende. Det norske statistikkmiljøet har såvidt redaksjonen vet fostret én idrettsverdensmester: hvem er dette? Og tilleggsspørsmålet, i tilknytning til Edel Therese Høiseth, må bli: Hvem er den ene skøyteløper i verden som har flere olympiske deltagelser enn hennes fem? Vedkommende har like mange OL bak seg som Jørund – men Jørund har fortsatt ikke lagt opp!

Nils Lid Hjort

Nytt fra Norges Landbrukshøgskole:

Nye Cand.scient. -er:

Guri Feten: '' Prediksjon i situasjoner med få observasjoner i forhold til tal på forklaringsvariable.''
Veiledere: Trygve Almøy og Are Aastveit.

Monica Kristiansen: ''Blanda modeller (mixed models) med anvendelse på repeterte målinger.''
Veiledere: Are Aastveit og Trygve Almøy

FORELØPIG ANNONSERING!

Norsk Statistisk Forening

KONFERANSE OG årsmøte

I Bergen

20.-22. juni, 2001

Det er planlagt et prekonferansekurs 18.-19. juni (eksakte metoder), med forelesning og PC-øvelser.

Hold av datoene -
og følg med i TG for videre informasjon og påmelding

Vel møtt i Bergen neste år!

Hilsen
Bergen Statistisk Forening

Norsk Regnesentral (NR) er et ledende forskningsinstitutt innen statistisk-matematisk modellering og analyse, mønstergjenkjenning, bildeanalyse og sentrale IKT-områder. Instituttets 100 ansatte arbeider med anvendt forskning på oppdrag for næringsliv og offentlig sektor i Norge og internasjonalt. Vår oppgave er å skape forskningsresultater som brukes og synes.



Landets ledende anvendte statistikkmiljø trenger flere

Forskere

NR utfører anvendt oppdragsforskning innen statistikk og matematisk modellering. Vi er over 40 forskere innen dette feltet. Med en stadig økende oppdragsmengde trenger vi å styrke kapasiteten ytterligere. NR jobber med bruk av avanserte modeller bl.a innen petroleum, finans, industri, medisin, miljø og ressursforvaltning for krevende oppdragsgivere!

Vi søker personer med universitets- eller høgskoleutdanning og som har:

- solide teoretiske og praktiske kunnskaper innen matematisk statistikk, matematisk modellering og numerisk analyse, gjerne med doktorgrad
- interesse for oppdragsforskning, markedskontakt og kundeoppfølging
- evne og lyst til å bli en ressursperson innen et av våre satsingsområder
- erfaring fra softwareutvikling (ingen betingelse)

Vi tilbyr spennende og varierte prosjekter og hyggelige kolleger i et unikt fagmiljø i sterk vekst. NR kan videre tilby kollektiv forsikrings- og pensjonsordning, egen bedriftshelsetjeneste og mulighet for barnehageplass.

For nærmere informasjon kontakt forskningssjef André Teigland, e-post Andre.Teigland@nr.no eller forskningssjef Lars Holden, e-post Lars.Holden@nr.no.

Kvinner og nyutdannede med stå-på-vilje oppfordres spesielt til å søke.

Skriftlig søknad med vitnemål og attester sendes personalkontoret.

Norsk Regnesentral, P.b. 114 Blindern, 0314 Oslo. Tlf: 22 85 25 00. Fax: 22 69 76 60.
For mer informasjon om NR, se vår URL: http://www.nr.no

Medlemsskap i Norsk statistisk Forening

Kontingent:
200 kr/år
50 kr/år for studenter (til og med hovedfag).
Kontonummer: 0530 2773299

Nummeret som er angitt på adresselappen på TG angir siste år kontingenten ble betalt.
Ved innbetaling må man opplyse om endret adresse (f.eks endret postnummer).

Korrespondanse om medlemsskap og innbetaling av kontingent skal skje til foreningens kasserer:

Norsk Statistisk forening v/Ågot Irgens
Kalfarlien 6A
5018 Bergen
Emailadresse: agot.irgens@haukeland.no

Nye medlemmer melder seg til kasserer og får så automatisk tilsendt Tilfeldig Gang fra første påfølgende nummer av.

Mailadresser:
Vi ønsker mailadresser til alle medlemmer i NSF:
Oppgi e-mailadressen din på merknadsfeltet på postgiroen, eller send en e-mail til agot.irgens@haukeland.no om mailadresse og evt adresseendring.

Annen korrespondanse med foreningen skal foregå med formannen:

Geir Egil Eide
Kompetansesenter for klinisk forskning
Haukeland sykehus
Armauer Hansens hus
5021 Bergen
Emailadresse: geir.egil.eide@haukeland.no

B-post

Retur:
Norsk Statistisk Forening
v/Ågot Irgens
Kalfarlien 6A
5018 BERGEN
Email: agot.irgens@haukeland.no