

BREV
Ukonvoluteret



MEDDELELSER

Dansk Selskab for Teoretisk Statistik

HUSK: Julemøde i DSTS 8. december

Klimaændringer: Kan de overhovedet forudsiges



Returneres ved varig adresseændring

Næste nummer af "MEDDELELSER" udkommer 1. februar 2010.
Bidrag skal være redaktøren i hænde senest den 22. januar kl. 12.00.

Seminarer

Seminar in applied mathematics and statistics, H.C. Ørsted institute, KU side 3

Debatoplæg af Birger Stjernholm Madsen

Om at kommunikere statistik til ikke-statistikere, side 4

Stillinger

Associate Professor in Biostatistics, University of Copenhagen, side 8
Nybeskikkelse af censorkorps for matematik ved universiteterne, side 11

Dansk Selskab for Teoretisk Statistik Bestyrelse 2008	
Formand	frnd@dsts.dk
Niels Richard Hansen Afd. for Anvendt Matematik og Statistik Københavns Universitet Tlf: 35 32 07 83	
Næstformand	nfrnd@dsts.dk
Esben Agerbo National Centre for Register-based Research Faculty of Social Sciences University of Aarhus Tlf: 89 42 68 15	
Kasserer	knss@dsts.dk
Malene Højbjerg Novo Nordisk A/S Tlf: 30 79 62 09	
Sekretær	sekr@dsts.dk
Marc Andersen StatGroup ApS Tlf: 61 77 72 48	
Webmaster	web@dsts.dk
Klaus Kaae Andersen Informatics and Mathematical Modelling, Section for Statistics, DTU Tlf: 45 25 34 19	
Redaktør	red@dsts.dk
Charlotte Hindsberger Novo Nordisk Tlf: 30 79 65 92	

Meddelelser er medlemsblad for Dansk Selskab for Teoretisk Statistik (DSTS), se <http://www.dsts.dk>.

Selskabets formål er at fremme den statistiske videnskab og dens anvendelser.

Indmeldelse og adressændring i DSTS gøres via <http://www.dsts.dk/da/index.html>.

Selskabet har en elektronisk nyhedsliste E-Meddelelser, se <http://www.dsts.dk/da/index.htm>.

Bidrag og stillingsopslag til Meddelelser sendes til redaktøren - red@dsts.dk. Bidrag i elektronisk form modtages i PDF format med indlejrede fonte i sidestørrelse A4, egnet til sort/hvid tryk i A5 format. Alternativt modtages Word.

Annoncering af stillingsopslag i Meddelelser koster kr. 1.000,- pr. side. Opgiv venligst faktureringsoplysninger. Indstik til udsendelse i konvolut sammen med Meddelelser koster kr. 3.000,- pr. standard A4 side for første side og 500,- kr per efterfølgende side.

Meddelelser udkommer 9 gange om året, den første mandag eller tirsdag i måneden undtagen januar, juli og august måned.

Udgivelsesplan for Meddelelser 2010

Nr.	Bidrag senest	Udkommer
1	22. januar kl. 12	1. februar
2	19. februar kl. 12	1. marts
3	26. marts kl. 12	5. april
4	23. april kl. 12	3. maj
5	28. maj kl. 12	7. juni
6	27. august kl. 12	6. september
7	24. september kl. 12	4. oktober
8	22. oktober kl. 12	1. november
9	26. november kl. 12	6. december



SEMINAR IN APPLIED MATHEMATICS AND STATISTICS

Wednesday, December 9, 2009, 15:15, aud. 10, H.C. Ørsted Institute
Speaker: Enrico Bibbona, University of Copenhagen
Title: Estimation of the input parameters for the Feller (CIR) neuronal model

Abstract:

Neuronal data from intracellular recordings can be modeled as trajectories of a square root process (also known as CIR or Feller process) that stops the first time a threshold is reached. We propose a new estimation method for this process and we compare it with the known ones both in the presence and in the absence of the threshold in a parameter range compatible with biological observed data. The presence of the threshold brings a bias into the estimates of the drift while it does not affect the diffusion coefficient.

NOVEMBER 20, 2009

UNIVERSITETSPARKEN 5
DK-2100 COPENHAGEN
DENMARK

<http://www.math.ku.dk/>

Et debatoplæg: Om at kommunikere statistik til ikke-statistikere

Dette debatoplæg er et sammendrag af et foredrag, jeg holdt ved ENBIS konferencen i Gøteborg i september. Det er tanker, der blev formet, da jeg skrev bogen "Statistik for ikke-statistikere" (Samfundslitteratur 2008) og efterfølgende oversatte den til engelsk (ikke udkommet endnu).

Det er min opfattelse, at formidling (her tænker jeg primært skriftlig formidling) af statistik til ikke-statistikere (både i Danmark og i andre lande) dumper både på fremstillingen og på emnevalget.

Fremstillingen

Der bør efter min mening lægges vægt på en formulering, der er letforståelig frem for 100 % matematisk stringent. Nogle få eksempler:

1) Om terminologi.

Et af problemerne ved elementære publikationer om statistik er, at læseren præsenteres for *adskillige ord for samme fænomen*, ofte endda i samme publikation! Det gælder også helt elementære begreber.

F.eks. bruges følgende betegnelser synonymt: (Measures of) Dispersion, Spread, Scatter, Variability, Variation, etc. På samme måde finder vi: (Measures of) Location, Centre, Central Tendency, Position etc.

Hvorfor har så grundlæggende begreber så mange betegnelser?

Der er gjort forskellige forsøg på at standardisere statistiske betegnelser. Det mest vellykkede er nok ISO ISO 3534 "Statistics - Vocabulary and symbols" (en standard i 3 dele). På dansk er der dog i dette tilfælde til en vis grad enighed om at bruge betegnelserne (mål for) niveau og (mål for) spredning.

Omvendt er det vel et rimeligt forlangende, at *væsentlige begreber rent faktisk også har et navn!* Lad os betragte formen for et 95 % konfidensinterval for en middelværdi i normalfordelingen:

$$\bar{x} \pm t \cdot \frac{s}{\sqrt{n}}$$

Her er t som bekendt 0,975 fraktilen i en t -fordeling med $n-1$ frihedsgrader.

Størrelsen $t \cdot s / \sqrt{n}$ er af fundamental betydning for praktikerne! Den intuitive fortolkning er, at det er den "statistiske usikkerhed" hørende til estimatet af middelværdien.

En så fundamental størrelse har naturligvis et navn? Ikke i ISO eller andre internationale standarder... Heller ikke i de fleste lærebøger, hvor man kan finde betegnelser som "den halve længde af et 95 % konfidens interval" eller simpelthen "tallet efter \pm "...

Inden for måleusikkerhed **har** man faktisk (i internationale standarder) defineret et begreb svarende til denne størrelse, det kaldes "den ekspanderede usikkerhed". Denne betegnelse kan dog næppe overføres til statistikens verden uden at møde modstand både blandt statistikere såvel som ikke-statistikere.

Indenfor repræsentative undersøgelser kaldes denne størrelse faktisk mange steder simpelthen "den statistiske usikkerhed" (hvis stikprøven udgør en væsentlig del af populationen, skal der dog et korrektionsled på). At "ophøje" denne betegnelse til standard terminologi vil næppe møde den store modstand blandt ikke-statistikere. Det bliver måske værre med statistikerne!

2) Om græske bogstaver, symboler m.v.

Her finder vi en stor del af forklaringen på, at statistik ikke ligefrem er en populær disciplin blandt menigmand...

Åbner man den legendariske bog af Box, Hunter & Hunter (2. udgave, 2005), finder vi på p. viii (lige inden indholdsfortegnelsen): Greek Alphabet! Læseren efterlades med det indtryk, at dette virkelig må være væsentligt...

Men: Hvor mange græske bogstaver **har** vi egentlig brug for? I elementære publikationer efter min mening **højest** 4: Σ , μ , σ , og (evt.) χ .

Σ betyder som bekendt sum, og det har de fleste heldigvis styr på takket være regnearkene!

μ og σ er jo henh. middelværdi og spredning, de er svære at komme uden om. Takket være Six Sigma kender mange heldigvis i σ forvejen!

Hvad bogstavet χ angår, så er det strengt taget slet ikke nødvendigt. På den anden side vil mange læsere alligevel støde på betegnelsen χ^2 , hvilket tæller for at bruge dette græske bogstav også.

Hvad resten af de græske bogstaver angår: Vi har **ikke** brug for dem!

Også forskellige **symboler** gør læsningen af statistiske publikationer til en svær opgave for ikke-statistikere. F.eks. vil vi (professionelt) ofte bruge en "hat" som symbol for estimat og skrive

$$\hat{p} = \frac{x}{n}$$

Måske burde vi i stedet blot skrive

$$p = \frac{x}{n}$$

Jeg ved godt, det ikke er korrekt! Men: gør det noget? Hvad med at opgive den matematiske præcision for at opnå større læselighed?

3) Om matematiske formler

Hvad skal vi med **formler** som f.eks.

$$b = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

Denne formel findes i adskillige "elementære" bøger om statistik! **Hvorfor?** Hvad skal ikke-statistikere med den? Hvilken informationsværdi indeholder den?

Der er vel ingen, der udfører disse beregninger i hånden nu til dags? Og hvis man gør, så er ovenstående formel for resten ikke særlig velegnet...

I øvrigt er det også bemærkelsesværdigt, så få anstrengelser der gøres for at lette tilværelsen for læseren af de statistiske formler. Et eksempel:

I forbindelse med estimation af parameteren p i en binomialfordeling kan vi under forudsætning af simpel tilfældig udvælgelse opstille følgende formel for den statistiske usikkerhed på estimatet $p = x/n$:

$$u = 1.96 \cdot \sqrt{\frac{p \cdot (1-p)}{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

Ikke just noget kønt syn, hvis man ikke er statistiker...

Under forudsætning af, at populationen er meget større end stikprøven reduceres dette en hel del:

$$u = 1.96 \cdot \sqrt{p \cdot (1-p)} \cdot \frac{1}{\sqrt{n}}$$

Hvis vi nu yderligere antager, at p ikke er alt for langt fra 0,5 (f.eks. $0,3 < p < 0,7$) og vi ignorerer den ubetydelige forskel mellem 1,96 og 2, så reduceres dette til formlen:

$$u = \frac{1}{\sqrt{n}}$$

Denne formel må siges at være **bemærkelsesværdig simpel!** Og den er samtidig ekstremt nyttig! Alligevel er det ikke lykkedes mig at finde den i én eneste bog om statistik... **Hvorfor?**

Emnevalget

Dette bør styres af, at hvilke emner, der er nyttige i det praktiske arbejde med statistik. Dels er der efter min mening et par spørgsmål, der bør afvises, dels er der emner, der traditionelt lægges for lidt (eller ingen) vægt på i elementære publikationer om statistik. Nogle eksempler:

1) Sandsynlighedsregning

Selv i "elementære" bøger om statistik finder man ofte en introduktion til sandsynlighedsregning. Det hører naturligvis med til statistikerens tankegods, men hvad skal ikke-statistikerne med formler som

$$P(X = x) = \binom{n}{x} \cdot p^x (1-p)^{n-x}$$

Er der nogen praktikere, som har nytte af denne information? Hvad skal de bruge det til? Bidrager formelen til at forstå binomialfordelingen? Tja, måske hvis man er matematiker...

2) Styrkeberegninger

I forbindelse med planlægning af forsøg stilles statistikerne ofte spørgsmålet: Hvor stort skal mit forsøg være?

Nu begynder der så en længere dialog, hvor statistikerne sammen med praktikerne skal tage stilling til bl.a. valg af signifikansniveau og styrke. Dette er vanskelige begreber, og derfor ender det ofte med, at statistikerne blot beslutter dette!

Dernæst skal statistikerne bruge et bud på standardafvigelsen, det kan praktikerne som regel levere. Endelig stiller statistikerne så et spørgsmål i retning af "Hvad er den mindste forskel, du gerne vil være i stand til at opdage?". Dette er jo i bund og grund noget, der ikke er statistikerens domæne. Alligevel begynder praktikerne at få flakkende øjne...

3) Konfidensintervaller.

Da jeg for mange år siden læste faget Statistik 1, blev konfidensintervaller stort set ikke omtalt. Det kan jeg godt undre mig over i dag. Mange praktikere har svært ved at forholde sig til test af

hypoteser, det er svært stof for ikke-statistikere. Helt anderledes forholder det sig med konfidensintervaller.

Ikke-statistikerne, der skal sammenligne middelværdien i to grupper, kan vælge at udføre den klassiske t-test. Men han kan også beregne et konfidensinterval for forskellen mellem middelværdierne, det vil mange ikke-statistikere have langt lettere ved at forholde sig til.

Dette kan naturligvis også udnyttes i forbindelse med planlægning af forsøg. Ikke-statistikerne vil som regel godt kunne sige, hvor stor en statistisk usikkerhed, han kan leve med (f.eks. for forskellen mellem to middelværdier).

Den (maksimale) "statistiske usikkerhed" er for de fleste et langt mere forståeligt begreb end den før omtalte "mindste forskel". Så slipper man også for at tænke på styrken. Der er altså én ting mindre at forholde sig til! Det er bekvemt, da der er ret stor forskel på, om man vælger en styrke på 80 % eller 95 %...

4) Avancerede og utraditionelle emner

Man skal efter min mening ikke holde sig tilbage for utraditionelle emner. Et eksempel er en omtale af "skævhed" og "kurtosis". Dette er traditionelt henvist til avancerede lærebøger, på trods af, at de er nyttige redskaber for praktikerne. De er ét af flere nyttige værktøjer til at afgøre, om data er normalfordelt, og hvad der evt. kan gøres for at opnå dette (f.eks. logaritmisk transformation eller fjerne outliers).

Her bliver man naturligvis nødt til at have nogle vejledende grænser for, hvor store afvigelser fra 0, som maksimalt kan accepteres. Det er i bund og grund et meget **simpelt spørgsmål!** Men man skal lede meget længe for at finde et svar...

Hvis man leder længe i "Kendall's Advanced theory of statistics" (!!), volume 1, finder man (en del af) svaret gemt væk i et par øvelser i form af asymptotiske grænser. F.eks. er den maksimalt acceptable afvigelse fra 0 for skævheden givet ved udtrykket

For kurtosis er grænsen det dobbelte. Men man finder ingen vejledning om, hvor stor stikprøvestørrelsen n skal være, før disse grænser kan bruges i praksis. Og så er man jo som praktiker lige vidt...

Så må man ty til et af de værktøjer, som også kan formidles til ikke-statistikere. **Simulering!** Simuleringsstudier viser, at ovennævnte grænser gælder fint for små stikprøver (f.eks. $n=25$) for skævheden. Derimod skal vi væsentligt højere op, før de fundne grænser gælder for kurtosis, helt op i nabolaget af $n=1000$.

For mindre værdier af n kan man ved hjælp af simulering beregne vejledende grænser for kurtosis. Det viser sig i øvrig, at disse grænser ikke bliver symmetriske!

Dette er et godt eksempel på, at resultater af simuleringsstudier godt kan præsenteres for ikke-statistikere og **give nyttig information!**

Kommentarer til ovenstående betragtninger er velkomne!

Mit foredrag kan findes på www.enbis.org eller fås ved henvendelse til undertegnede.

Med venlig hilsen
Birger Stjernholm Madsen,
bsm@nqvozymes.com

The Faculty of Health Sciences
Associate Professor in Biostatistics

Section of Biostatistics, Department of Public Health at the Faculty of Health Sciences, University of Copenhagen offers a position as Associate Professor. The position is to be filled for the period 1 May 2010 to 31 January 2012.

Section of Biostatistics conducts general biostatistical research and methodological development in biostatistics, undertakes teaching of students at various educations, undergraduate as well as postgraduate and offers biostatistical advice to staff and students in all disciplines of the Faculty of Health Sciences. This position is partly sponsored by the Faculty's Core Facility in Biostatistics.

Qualifications and Job description

The position requires research in biostatistics, including publication and scientific communication as well as research based teaching (with associated conduct of exams). Part of the teaching obligation is replaced by biostatistical service consulting. The successful applicant must take part in teaching on one or more of the educations at the Faculty, including pregraduate teaching of students of Medicine, Dentistry, Public Health, Cand.Scient.San., Human Biology, Molecular Biomedicine, the Civil Engineer Education in Medicine and Technology, IT and health, the programmes of Master of Public Health and Master of International Health, as well as participate in the teaching and development of post graduate courses.

The position involves professional development, participation in the scientific development of the professional area, education of researchers, pedagogical guidance and supervision of Assistant Professors as well as professional review- and assessment tasks. Moreover, the candidate will handle duties within the purpose of the institution, including administrative duties. The University decides on the allocation between tasks, and the weighting of different duties may vary over time.

Appointment as Associate Professor requires research- and teaching qualifications corresponding to the level that can be achieved by a successfully accomplished appointment as Assistant Professor, although it may be obtained otherwise. Appointees as Associate Professor are presumed to have a documented scientific production on an international level as well as documented teaching qualifications.

The focus will be on scientific production at an international level in statistical methods that combine methodological depth with convincing applications as well as teaching experience (also of non-statisticians) and consulting and collaboration with empirical researchers. In the total assessment of each candidate the research potential and the interest in development and teaching will also be considered important.

For further information, reference is made to 'Cirkulære om stillingsstruktur for videnskabeligt personale ved universiteter' available in Danish at www.ku.dk/regel/2/2050.html/ or in English: 'Job Structure for Academic Staff at Universities 2007' at www.ku.dk/regel/. Click on "Personaleforhold" and scroll down to item 20.1.

As regards the scientific research requirement, reference is made to the applicable general rules, currently "*Københavns Universitets fælles regler for det videnskabelige personales forskningsindsats (forskningssikring)*" (only available in Danish) and to "KU-Academic Freedom and Responsibility at the Faculty of Health Sciences (51.04) (E)" both available at www.ku.dk/regel/; click "Forsikring" and scroll down to items 58.10 and 51.04.

Salary and employment terms

Salary and employment terms are subject to the collective agreement made between the Danish Ministry of Finance and the Danish Confederation of Professional Associations on the employment of academics in state service.

In addition to the seniority-based salary and an annual pensionable supplement of currently DKK 86.467 (1 April 2009 level), bonuses may be negotiated on an individual basis.

Further information

For further information, please contact Professor Niels Keiding (tel.: +45 35 32 79 03, email N.Keiding@biostat.ku.dk).

Procedure

The application must contain:

- a specification of the basis on which the applicant wishes his/her academic qualifications assessed
- a curriculum vitae
- a complete, consecutively numbered list of scientific publications, with an indication of the (max. 10) publications that the applicant wishes to submit for the assessment of scientific qualifications
- a thorough account of experience with research supervision
- a thorough account of teaching experience or material that may form a basis for an assessment of teaching qualifications as well as pedagogical level
- evaluation of the teaching qualifications obtained from the period as an Assistant Professor
- a thorough account of other qualifications of relevance for the position
- a description of the applicant's plans with respect to development of teaching in the professional area
- a research strategy
- copies of each of the above-mentioned (max. 10) publications

If an applicant for the Associate Professorship has not obtained a sufficient amount of teaching experience through an appointment as Assistant Professor, completed the pedagogical course for Assistant Professors or obtained analogous teaching experience elsewhere, the appointment as Associate Professor will be conditional upon a successful initial trial period of up to 1½ years.

The application, the material mentioned above and other documentation that the applicant may wish to refer to for the assessment of qualifications, has to be delivered in 4 copies collated in complete sets.

While general reference is made to the Danish Ministerial Order no. 284 of 25 April 2008 on the

employment of scientific staff at Danish universities –available in Danish and English at www.ku.dk/regel; click on “Personaleforhold” and scroll down to section 20.1.– please note that

- each single applicant will only be informed about the total number of applicants, not the names of these
- the applications will be assessed by an assessment committee appointed by the Faculty
- the applicants will be informed about the members of the assessment committee and be asked to consider possible ineligibility problems
- the applicants will be contacted in case the assessment committee should want any additional documentation and/or wish to include additional publications in the assessment
- each applicant will only be informed about his/her own assessment and asked to take a position on this assessment
- one or more applicants may be called upon to an interview and be invited to give trial lectures
- because of the undergraduate teaching obligations, it is a prerequisite for appointment that Non-Scandinavians become competent in Danish within two years

All interested parties are encouraged to apply for the post, regardless of age, sex, race, religion or ethnicity.

Applications should be marked “211-0660/09-3012”, addressed to the Rector and sent to the Faculty of Health Sciences, The Panum Institute, University of Copenhagen, Blegdamsvej 3 B, 2200 Copenhagen N, Denmark, and must be received by the Faculty by **10 December 2009, noon at the latest**. Material received after this time will not be considered.

If the material is delivered by hand, it must be delivered to the Faculty’s Service Center, which is situated on the first floor of the above address.

The Faculty of Health Sciences comprises 12 institutes and 9 study committees. The Faculty employs c. 1,870 people, including 920 teachers and scientists and 950 technical/administrative employees. The Faculty has c. 4,370 students, including c. 900 Ph.D. students.

Kalender 2009 - 2010

Dato	No.	Aktivitet
8 Dec. 09	8	Julemøde i DSTS <i>Klimaændringer: Kan de overhovedet forudsiges</i>
9 Dec. 09	9	SEMINAR IN APPLIED MATHEMATICS AND STATISTICS, H.C. Ørsted Institute <i>Estimation of the input parameters for the Feller (CTR) neuronal model</i>
24-28 Jan. 10	6	Workshop, Thiele Centre, University of Aarhus. <i>Workshop on Ambit Processes, Non-Semimartingales and Applications</i>

Nybeskikkelse af censorkorps for matematik ved universiteterne.

Beskikkelsesperioden for censorkorpset i matematik (inklusive forsikringsmatematik, matematik-økonomi og matematisk statistik) ved universiteterne i Aalborg, København, Roskilde, Odense og Århus udløber d. 31. marts 2010. Den relevante bekendtgørelse foreskriver, at mindst 25% af det eksisterende korps skal udskiftes, og at mindst 1/3 af al censur bør foretages af personer med hovedstilling uden for universiteterne (såkaldte aflagereensorer). Nærmere oplysninger om censorkorpset og muligheden for at søge om optagelse kan findes på internetadressen

www.sdu.dk/imada/censorjob

senest fra den 1. december 2009. Ansøgningsfristen bliver medio februar 2010.

Hans J. Munkholm,
formand for korpset.