



ISSN 0803-8953

Tilfeldig gang

Nr. 1, årgang 40

Juli 2023

Utgitt av Norsk statistisk forening

Redaksjonelt

Leder	1
Fra redaksjonen	2

Artikler

Nils Lid Hjort 70 år	3
Første doktorgrad i biostatistikk ved Norges idrettshøgskole taklet statistisk skandale i idrettsforskning	5

Tankenøtter

Pusleri nr. 58	9
Pusleri nr 57 (løsning)	9

Meldinger

Nytt fra Universitetet i Stavanger	14
Nytt fra Norges Idrettshøgskole	14
Nytt fra Oslo senter for biostatistikk og epidemiologi	14
Nytt fra Tromsø	14
Nytt fra NTNU	14
Nytt fra UiB	15
Nytt fra UiO	16

Leder

Jacob Skauvold | Norsk Regnesentral

NSF har nytt styre etter hovudmøtet tysdag 13. juni. Du finn møtereferat på [nettsida til NSF](#). Før eg takka ja til å stille til val som leiar vart eg lova kontinuitet i styret. Basert på ei rask samanlikning av gammalt og nytt styre (sjå tabell) ser dette ut til å vera oppfylt. Her er fleire som har vore med på å styre NSF før, og ein av dei sit til overmål på eit kontor eg kan gå tørrskodd til. I tillegg er revisorane og medlemmane av Tilfeldig Gang-redaksjonen dei same som før.

	Førre styre	Nytt styre
Leiar	Ingrid Hobæk Haff, UiO	Jacob Skauvold, NR
Kasserar	Sondre Hølleland, HI	Eirik Myrvoll-Nilsen, UiT
Styremedlemmar	Ragnar Bang Huseby, NR Yushu Li, UiB Alf Harbitz, HI	Ragnar Bang Huseby, NR Yushu Li, UiB Sara Martino, NTNU
Varamedlemmar	Tore Wentzel-Larsen, RBUP Clara-Cecilie Günther, Gjensidige	Tore Wentzel-Larsen, RBUP Clara-Cecilie Günther, Gjensidige

Nordstat 2023 gjekk føre seg førre veke på Chalmers i Gøteborg, og me kan no sjå fram til det 21. norske statistikarmøtet i Tønsberg 17. til 20. juni 2024. Eg håper alle som les dette har tenkt seg dit. Det vert på dagen fem år etter det 20. møtet, som var på Sola i 2019. Opphaldet kom først på grunn av pandemien, og vart ekstra langt på grunn av utakt med Nordstat. Frå neste år er me vonleg inne i ei ny takt.

Det er ei spesiell tid for oss talknusarar og kunnskapsarbeidarar. Mange av oss tek i bruk nye verktøy som ChatGPT og GitHub Copilot, og lurar på korleis me kjem til å arbeide i framtida. NSF vart stifta 28. april 1936, så hundreårsjubileet er om snaue 13 år. Skal me framleis skriva forskningsartiklar og fare på konferansar då? Skal me framleis halda forelesningar og laga eksamensoppgåver?

Når maskinlærarane springer føre og søker det nye får me statistikarar ei mogleg rolle som dei som held att og stiller naudsynte spørsmål: Korleis fungerer det egentleg? Kva for føresetnader kviler det på? Kan det forklarast? Kan det tolkast? Kort sagt, kva trur de at de veit, og kvifor trur de det? Me har kanskje ikkje bedt om denne rolla, men likevel bør me ta ho på alvor. Viss eg skal seie det med ei omskriving av ordspråket be the change you want to see in the world så må det bli ver det blyet du vil ha i kjølen.

Takk for tilliten, og stor takk for innsatsen til alle i det førre styret og spesielt til avtroppande styreleiar Ingrid Hobæk Haff.

Fra redaksjonen

Jarle Tufto, Gunnar Taraldsen | Trondheim

Redaksjonen ønsker å takke alle som har bidratt med stoff til denne utgaven. Vi tar sikte på ett nytt TG-nummer rundt juletider.

Nils Lid Hjort 70 år

Ingrid K. Glad & Geir Storvik | UiO

Den 12. januar fylte Nils Lid Hjort 70 år. For de fleste ville et slikt jubileum markere slutten på et arbeidsliv, men ikke for Nils. I inneværende akademiske år leder Nils, sammen med Håvard Hegre fra PRIO, prosjektet *Stability and Change* ved Senter for grunnforskning (CAS) ved Det Norske Videnskaps-Akademi. En prominent gruppe av nasjonale og internasjonale forskere sitter sammen i Drammensveien og kombinerer statistisk metodikk med fred-og-konflikt-forskning, et svært tidsaktuelt tema! Også på andre felt fortsetter Nils sin aktivitet. Han veileder, vil fortsette å undervise, er med i Integreat (nytt SFF), skriver bok og, ikke minst, er aktiv på sosiale medier med formidling og folkeopplysning.

Nils har hatt det meste av sin yrkeskarriere ved Matematisk institutt, Universitetet i Oslo. Han ble førsteamanuensis i 1988 og professor i 1991. Før det var han en periode på Norsk Regnesentral. Nils har hatt (og har fremdeles) en imponerende forskningskarriere innenfor mange fagområder og med mange medforfattere (over 70 er listet i Web-of-Science!). Mange av oss i generasjonen bak Nils ble kjent med ham gjennom hans grønne bok om sannsynlighetsregning, som han skrev allerede som 25-åring. Allerede da viste han sine fremragende pedagogiske og matematisk konsise evner. Boken *Model selection and model averaging* sammen med Gerda Claeskens og de relaterte artiklene *The focused information criterion* og *Frequentist model average estimators* er de mest siterte arbeidene til Nils, men de dekker bare en liten del av hans omfangsrike bidrag. Sammen med Tore Schweder utga han i 2016 boken *Confidence, likelihood, probability* basert på Fisher's fidusielle fordelinger. Andre områder innkluderer (Bayesiansk) ikkeparametrisk regresjon, tetthetsestimering, levetidsanalyse, asymptotikk og romlig statistikk. Nils mottok Sverdrup-prisen i 2013 for sin faglige innsats.

Nils har også hatt stor innflytelse utenfor statistikkverdenen. Vi siterer fra TG i 2013 (av Ørnulf Borgan og Tore Schweder):

Nils flytter fjell. Han maktet å få Den internasjonale olympiske komité (IOC) til å endre reglene for 500-meter skøyteøp i vinterolympiadene fra og med Nagano-ol i 1998. Alle må nå gå to løp, ett med første ytre og ett med første indre. De som har greie på skøyteøp visste jo at første indre var en fordel, men Nils påviste det «signifikant», og da er det så sant at det ikke er til å komme utenom. Fjellet IOC akter Nils å flytte nok en gang. Nå er det 1 000 meter det gjelder. På grunnlag av de 12 sprintverdensmesterskapene 2001–2012 og 3 World Cupkonkurranser, hvor hver løper må gå både første ytre og første indre, finner han en konfidensfordeling for fordelene ved å starte i



indre bane som svarer til om lag to meter, med all konfidens til høyre for null.

Sitt engasjement for omverdenen har han også vist i mange andre sammenhenger, ikke minst gjennom senere års samarbeid med PRIO. I tillegg bruker han sin tid og sitt gode hode på språkstudier, Donald Duck, musisering, korsynging og god litteratur. Blant annet!

Gjennom karrieren har Nils veiledet et stort antall hovedfags-/master- og PhD-studenter. Nils tar seg alltid god tid til studentene, både til faglig veiledning og diskusjon om alle mulige andre temaer. Som underviser har Nils alltid vært fremragende og engasjerende, og han har undervist en imponerende rekke av ulike kurs. Det går rykter om at Nils forbereder de beste forelesningene sine i heisen på vei ned til forelesningssalen. Nils kan (nesten) alt, og alle som har jobbet seg gjennom hans innholdsrike hjemmeoppgaver, har lært for livet.

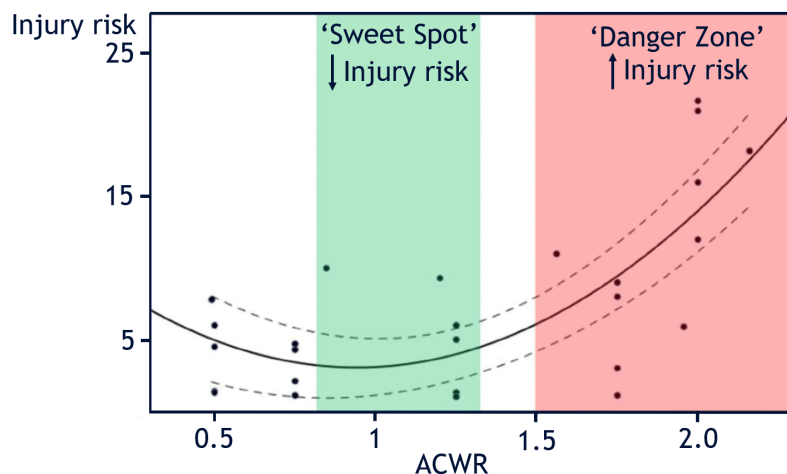
Nils har alltid vært en god kollega, engasjert i både de små og de store saker, alltid positiv og interessert, med tid til en prat og gode kommentarer.

Det vil bli arrangert et 2-dagers seminar 4.-5. desember på Blindern for å feire Nils sine utallige og viktige bidrag til statistikkfeltet. For mer detaljer og påmelding, se [egen web-side](#). Dog, med sitt fortsatt levende engasjement for faget, vil dette neppe markere de siste bidragene vi ser fra ham. Vi gleder oss til fortsettelsen!

Første doktorgrad i biostatistikk ved Norges idrettshøgskole taklet statistisk skandale i idrettsforskning

Lena Kristin Bache-Mathiesen | Norges Idrettshøgskole

I 2016 kom en ny metode for å beregne treningsbelastning. Den nye metoden, den såkalte «acute:chronic workload ratio» (ACWR), var ikke bare en måte å beskrive belastning, den kunne også brukes til å redusere idrettsskader (figur 1, [1]). Ved hjelp av en enkel figur kunne en idrettsutøver, mosjonist eller trener vurdere om man har trent for lite, for mye, eller akkurat passe. Idrettsforskning tok imot figuren med storm, og den ble publisert i konsensus-artikler og spredd via sosiale medier [2]. Den var nesten for god til å være sann.



Lena Kristin Bache-Mathiesen

Figur 1: Ved hjelp av en ny beregning til treningsbelastning, «Acute:Chronic Workload Ratio» (ACWR), kan man monitorere treningsbelastning og redusere skaderisiko, ifølge Gabbett.

Forskere fant grove statistiske feil i utarbeidelse av figuren [3], og store begrensninger med den nye metoden for å beregne treningsbelastning [4]. I tillegg hadde studier i treningsbelastning- og skadefeltet mange andre metodologiske problemer [5]. Utvalgsstørrelsene var små, statistiske antagelser ignorert, og manglende data var håndtert på ymse vis, om de var håndtert i det hele tatt. Antagelser i sentrale hypoteser indikerte at tradisjonelle metoder ikke var egnet for å studere treningsbelastning og skade. Ekspertene var flinke til å påpeke utfordringer og problemer med de tradisjonelle metodene [6], men det var ikke noe klart svar til gyldige alternativer. Ingen hadde utført simuleringer eller andre sammenligninger av statistiske metoder. Som konsekvens, har studiene i feltet brukte et drøss med metoder for å beregne treningsbelastning [7]. Resultatene mellom studiene har vært inkonsekvente, og systematiske oversikter har slitt med å sammenligne studier med så forskjellig metodikk. Hvilke statistiske metoder burde man bruke for å studere sammenhengen mellom treningsbelastning og risiko for skade? Hvis man ikke skal bruke ACWR, hva burde man bruke da? Ved Senter for idrettsskadeforskning satte man i gang et doktorgradsprosjekt for å finne svar på nettopp dette.

Kandidat Lena Kristin Bache-Mathiesen utførte prosjektet

og fullførte den første doktorgraden i biostatistikk ved Norges idrettshøgskole. Veiledere var Dr. Morten Wang Fagerland fra Oslo senter for biostatistikk og epidemiologi, og Prof. Thor Einar Andersen ved Senter for idrettsskadeforskning, Norges idrettshøgskole. Prosjektet omfattet fire sentrale utfordringer i feltet:

- 1 Statistisk styrke i treningsbelastning- og skadefeltet er ofte begrenset av antall skader som dukker opp i løpet av en longitudinell kohort studie. Manglende data i treningsbelastningsmålinger har ført til at forskere har slettet skader og dermed redusert styrken i studien. Ved hjelp av en simulering basert på data fra menn i Eliteserien, har prosjektet funnet at multipl imputering fungerer bra for å håndtere manglende data i treningsbelastnings-spesifikke målinger [8]. I tillegg viste prosjektet at å imputere variabler før man beregner sammensatte målinger var bedre enn andre tilnærminger.
- 2 Hypoteser foreslo at ikke bare for mye treningsbelastning, men også for lite treningsbelastning, kan øke risiko for en idrettsskade. Denne antagelsen tilsier en ikke-lineær, parabolisk «U»-sammenheng mellom treningsbelastning og sannsynlighet for idrettsskade. Forskere i feltet har tidligere brukt kategorisering for å håndtere denne antagelsen. Til tross for en simuleringsstudie som viser faren med denne tilnærmingen [9], hadde senere artikler anbefalt dette [10, 11], og det har gjort studier enda mindre sammenlignbare [7]. I doktorgradsprosjektet simulerte man ulike ikke-lineære og lineære sammenhenger mellom treningsbelastningsdata på U-19 fotballspillere, og sannsynligheten for en skade [12]. Kategorisering med kvartiler kunne ikke engang detektere den lineære sammenhengen. Fraktale polynomer og begrensede kubiske spliner kunne derimot detektere alle sammenhenger, og var de anbefalte metodene for å utforske ikke-lineæritet. I denne studien fant man også ut at ved skjeve fordelinger, kan det være lurt å subjektivt velge plasseringen av knutene i de kubiske splinene, i stedet for å forholde seg til data-drevne metoder.
- 3 En stor utfordring i feltet har vært antagelser om en kumulativ sammenheng mellom langtidstreningsbelastning og skaderisiko. Skaderisikoen skal teoretisk sett være en samlet sum basert på alle treningsaktiviteter utført i fortiden. Aktiviteter utført nylig påvirker sannsynligvis skaderisiko i større grad enn aktiviteter utført for flere uker siden [13]. I tillegg har teorier påstått at nylig utførte aktiviteter kan øke risiko for skade, mens aktiviteter utført i fortiden kan forebygge og redusere risikoen for skade [1]. Kort sagt, retningen til effekten av treningsbelastning kan være avhengig av hvor lenge siden aktiviteten var utført. Tredje studie simulerte ulike scenarier i sammenhengen mellom langtidstreningsbelastning og skaderisiko [14]. Resultatene viste at statistiske modeller brukt i miljøepidemiologi, til å studere langtidseffekten av forurensning og radioaktivitet

på ulike utfall, også kan brukes til å ta høyde for disse antagelsene i treningsbelastning [15]. Såkalte «distributed lag non-linear models» kan ta høyde for forsinkede effekter, men kompleksiteten til slike modeller vil kreve store mengder data for å ha tilstrekkelig med frihetsgrader.

- 4 Til slutt har prosjektet vurdert alternativer til den berømte «Acute:Chronic Workload Ratio» (ACWR). Brå endringer fra langtidstreningsbelastning («chronic load») til nåtidstreningsbelastning («acute load») kan betraktelig øke risikoen for en skade. ACWR delte ukentlig sum av «acute load» på 4-uker glidende gjennomsnitt av «chronic load». Dette hadde en rekke problemer. Definisjon av 1 vs. 4 uker var arbitrær og innebar alle ulemper som kategorisering innebar. En ratio gjorde det umulig å håndtere en treningsbelastning på 0. Siste del av prosjektet demonstrerte hvordan man kan inkludere «acute load» og «chronic load» som separate, uavhengige variabler i en modell, og slippe problemene med en ratio [16]. På denne måten kunne man også utforske interaksjonen mellom nåtids- og langtidstreningsbelastning. Studien argumenterte at dagens treningsbelastning har sannsynligvis såpass ulik effekt på skaderisiko i forhold til treningsbelastning fra alle ukene før, at det er trygt å kategorisere disse tidsdefinisjonene.

Avhandlingen er et dypdykk i tema og diskuterer de fire delprosjektene i lys av hverandre [17]. Forhåpentligvis vil tilgjengeligheten til egnede statistiske verktøy gjøre det lettere for forskere å studere effekten av treningsbelastning på idrettsskader, og unngå at mosjonister og idrettsutøvere blir pålagt intervensjoner basert på utilstrekkelig evidens.

Referanser

1. Gabbett, T.J., The training—injury prevention paradox: should athletes be training smarter and harder? *British Journal of Sports Medicine*, 2016. 50(5): p. 273-280.
2. Soligard, T., et al., How much is too much?(Part 1) International Olympic Committee consensus statement on load in sport and risk of injury. *British Journal of Sports Medicine*, 2016. 50(17): p. 1030-1041.
3. Impellizzeri, F., et al., The acute-chronic workload ratio-injury figure and its 'sweet spot' are flawed. *SportRxiv*, 2019.
4. Wang, C., et al., Analyzing Activity and Injury: Lessons Learned from the Acute: Chronic Workload Ratio. *Sports Medicine*, 2020: p. 1-12.
5. Windt, J., et al., Getting the most out of intensive longitudinal data: a methodological review of workload–injury studies. *BMJ open*, 2018. 8(10): p. e022626.
6. Menaspà, P., Are rolling averages a good way to assess training load for injury prevention? *British journal of sports medicine*, 2017. 51(7): p. 618-619.

7. Dalen-Loretsen, T., et al., A Cherry, Ripe for Picking: The Relationship Between the Acute-Chronic Workload Ratio and Health Problems. *Journal of Orthopaedic & Sports Physical Therapy*, 2021. 51(4): p. 162-173.
8. Bache-Mathiesen, L.K., et al., Handling and reporting missing data in training load and injury risk research. *Science and Medicine in Football*, 2022. 6(4): p. 452-464.
9. Carey, D.L., et al., Modeling Training Loads and Injuries: The Dangers of Discretization. *Medicine and science in sports and exercise*, 2018. 50(11): p. 2267-2276.
10. Ruddy, J.D., et al., Modeling the risk of team sport injuries: a narrative review of different statistical approaches. *Frontiers in physiology*, 2019. 10: p. 829.
11. Nielsen, R.O., et al., Time-to-event analysis for sports injury research part 1: time-varying exposures. *British Journal of Sports Medicine*, 2019. 53(1): p. 61-68.
12. Bache-Mathiesen, L.K., et al., Not straightforward: modelling non-linearity in training load and injury research. *BMJ Open Sport & Exercise Medicine*, 2021. 7(3): p. e001119.
13. Williams, S., et al., Better way to determine the acute: chronic workload ratio? *British Journal of Sports Medicine*, 2017. 51(3): p. 209-210.
14. Bache-Mathiesen, L.K., et al., Assessing the cumulative effect of long-term training load on the risk of injury in team sports. *BMJ Open Sport & Exercise Medicine*, 2022. 8(2): p. e001342.
15. Gasparrini, A., Modeling exposure–lag–response associations with distributed lag non-linear models. *Statistics in medicine*, 2014. 33(5): p. 881-899.
16. Bache-Mathiesen, L.K., et al., A new statistical approach to training load and injury risk: separating the acute from the chronic load. *Biology of Sport*, 2023, in press.
17. Bache-Mathiesen, L.K., Improving the Methodology of Training Load and Injury Risk Research: An Analysis of Analyses, in *Institute of Sports Medicine*. 2023, Norwegian School of Sport Sciences.

Pusleri nr. 58

Jostein Lillestøl | Norges Handelshøyskole

I et TV-program, som TV2's Kompani Lauritzen, konkurrerer 14 kjendiser som rekrutter, 7 menn og 7 kvinner. Etter noen uker er 5 rekrutter dimittert, og 9 gjenstår, 5 menn og 4 kvinner. Anta at disse blir kommandert til å stille opp på linje i tilfeldig rekkefølge med ansiktet vendt samme vei. Så kommer ordren: «Menn: Høyre om! Kvinner: Venstre om!». Resultatet blir at noen står ansikt til ansikt med en av motsatt kjønn, mens noen ser inn i en rygg til en av samme kjønn eller ser ingen. Hvert par som står ansikt til ansikt skal fortsette som oppdragsteam, mens de øvrige tas ut til en drittjobb (rengjøring av latrine).

Hva er sannsynlighetsfordelingen til antall team? Hva er forventet antall rekrutter tildelt drittjobben? Hva kan sies mer generelt?

Løsning til Jostein.lillestol@nhh.no

Pusleri nr. 57 (løsning)

Oppgaven var: I en sjakkklubb står kampen om klubbmesterskapet mellom A og B. For å redusere risikoen for mange kjedsomme remis-partier er det foreslått at mesterskapet avgjøres ved at mester blir den som først vinner to partier på rad. Totalt antall partier er derfor ikke kjent på forhånd, men det er enighet om at opp til ti partier kan aksepteres. Som medlem av klubben blir du bedt om å vurdere forslaget ut fra en sannsynlighetsteoretisk synsvinkel. Ta utgangspunkt i følgende antakelser:

For hvert parti er det 50% sjanse for remis, 30% sjanse for at A vinner og 20% sjanse for at B vinner (dvs. uavhengig av hvem som spiller svart og hvit).

Hva er sannsynligheten for at A blir mester?

Hva er forventet antall partier som må spilles for å kåre mesteren?

Vil det være et bedre forslag at mester blir den første som har vunnet minst en gang med hvit og minst en gang med svart, ikke nødvendigvis etter hverandre?

Løsning:

Det skal først innrømmes at antakelser om uavhengighet og konstante sannsynligheter er nokså urealistisk i denne konteksten, men beregninger under disse forutsetninger kan likevel belyse interessante problemstillinger.

La sannsynligheten for at A vinner/B vinner et parti være henholdsvis a/b , slik at sannsynligheten for remis er $r = 1 - (a + b)$. Betegn de mulige utfall av hvert parti med R, A og B.

La p_i være sannsynligheten for at A blir mester gitt uavklart situasjon der forrige parti hadde utfall i ($i=R,A,B$). Med et fiktivt 0'te remis-parti er den søkte sannsynlighet det samme som p_R . Betinging mhp. utfallet av neste parti gir

$$\begin{aligned} p_R &= r \cdot p_R + a \cdot p_A + b \cdot p_B \\ p_A &= r \cdot p_R + a \cdot 1 + b \cdot p_B \\ p_B &= r \cdot p_R + a \cdot p_A + b \cdot 0 \end{aligned}$$

Ved eliminasjon av p_A og p_B får vi den søkte sannsynlighet

$$p_R = \frac{a^2(1+b)}{a^2(1+b) + b^2(1+a)}$$

Innsetting av $a = 0.30$, $b = 0.20$ gir $p_R = 0.675$.

Som ventet er denne sannsynligheten noe høyere enn sannsynligheten 0.60 for at A er vinner av et enkelt ikke-remis parti.

Forventet antall partier som trengs til vinneren er kåret kan utledes på tilsvarende måte. La f_i være forventet antall partier gitt siste parti med utfall i ($i=R,A,B$). Med et fiktivt 0'te remis-parti har vi tre ligninger

$$\begin{aligned} f_R &= 1 + r \cdot f_R + a \cdot f_A + b \cdot f_B \\ f_A &= 1 + r \cdot f_R + a \cdot 1 + b \cdot f_B \\ f_B &= 1 + r \cdot f_R + a \cdot f_A + b \cdot 1 \end{aligned}$$

som gir løsningen

$$E(N) = f_R = 1 + \frac{1 + a + b + ab}{a^2(1 + b) + b^2(1 + a)}$$

Innsetting av $a = 0.30$, $b = 0.20$ gir $E(N) = 10.75$. Altså: Allerede med disse antakelser er vi i overkant av det omfang som kan aksepteres. Dersom remishyppigheten er høyere og/eller A og B er tilnærmet like gode vil det ta lengre tid. Eksempelvis vil $a = b = 0.20$ og dermed $r = 0.60$ gi $E(N) = 16.0$, mens $a = b = 0.10$ og dermed $r = 0.80$ vil gi $E(N) = 56.0$. I vurderingen må en selvfølgelig også ta med risikoen for at antall partier overgår forventningen. Dette kan studeres ved etablering av tilsvarende formler for varians eller ved simulering. I den foreliggende kontekst har man vel allerede nok skyts til å fraråde forslaget.

Alternativ løsning: Problemstillingen kan alternativt studeres som en Markovkjede med fem tilstander: R, A, B, A*, B*, der de tre første er utfallet av sist spilte parti, og de to siste er de absorberende tilstander når henholdsvis A og B har vunnet matchen. Overgangsmatrisen blir

$$\begin{array}{ccccc} & R & A & B & A^* & B^* \\ \begin{array}{c} R \\ A \\ B \\ A^* \\ B^* \end{array} & \begin{bmatrix} r & a & b & 0 & 0 \\ r & 0 & b & a & 0 \\ r & a & 0 & 0 & b \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

Matrisen har følgende struktur

$$\begin{bmatrix} Q & S \\ 0 & I \end{bmatrix}$$

der Q representerer overgangene mellom de tre transiente tilstandene. Våre resultater kan avledes av fundamentalmatrisen

$$F = (I - Q)^{-1} = \frac{1}{d} \begin{bmatrix} 1 - ab & a(1 + b) & b(1 + a) \\ r(1 + b) & 1 - r(1 + b) & b \\ a(1 + b) & a & 1 - r(1 + a) \end{bmatrix}$$

der nevneren er determinanten, som alternativt kan skrives

$$d = (1 - r)^2 - (1 + r)ab = a^2 + b^2 + (1 - r)ab = a^2(1 + b) + b^2(1 + a).$$

Første linje i matrisen F angår start i den fiktive starttilstand R . Sannsynligheten for absorpsjon i hver av de to tilstandene A^* og B^* finnes som elementene i første linje i matrisen FS , mens forventet antall trinn før absorpsjon i en av de to tilstandene finnes som toppelementet i vektoren Fe , der e er kolonnevektoren med 1-tall. Dette gir resultatene ovenfor, der en for forventningen må legge til 1 for selve absorpsjonen.

Alternativt opplegg: Å spille partier inntil en spiller har vunnet en gang med hvit og en gang med svart, ikke nødvendigvis i rekkefølge. Dette gir tilsynelatende en mer komplisert utregning enn ovenfor. Presisering: La den værende mester begynne som hvit, og la denne forbli mester dersom denne når målet med et hvit-parti, slik at utfordrerer ikke får spille et følgende hvit-parti for å oppnå uavgjort og dermed uavklart mesterskap. Med antakelsen om like gode spillere med sannsynlighetene r =remis, p =hvit vinner og q =svart vinner kan en som ovenfor etablere et omfattende ligningssystem. Etter mye strev framkom formelen:

$$E(N) = \frac{1}{1-r} \left\{ 1 + \frac{p}{1-r(p+r)} \left(2 + \frac{2qr}{1-r} + \frac{p}{q} \right) + \frac{q}{1-r(q+r)} \left(2 + \frac{2pr}{1-r} + \frac{q}{p} \right) \right\}$$

Med $r=0.50$, $p=0.30$ og $q=0.20$ blir forventet antall partier $E(N) = 7.91$, et resultat som ble bekreftet ved simulering (og dermed gir tro på at formelen er riktig). Også her kan situasjonene formuleres som en Markovkjede, nå med ni transiente og to absorberende tilstander. Dette opplegget gir altså et noe mindre forventet omfang enn forslaget i oppgaven, og virker muligens mer rimelig å praktisere.

Kommentarer:

Reglene for kåring av verdensmesteren i sjakk har variert mye gjennom tidene. I hovedsak har det vært best av et forhåndsbestemt antall partier eller spill inntil et visst antall seire, oftest uten øvre grense på antall partier. Eksempelvis ble det i matchen mellom Alekhin og Capablanca i 1927 spilt inntil 6 vunne partier, noe som medførte 34 spilte partier. I rekken av Botvinniks seire på 1950-tallet ble det spilt best av 24 partier. Ved den legendariske VM-matchen i 1972 mellom Bobby Fischer og Boris Spasskij ble det også spilt best av 24 partier. Fischer ønsket i forbindelse med en mulig returmatch i 1975 at det, for å unngå lange rekker med remis, skulle spilles inntil 10 vunne partier, med tilleggsregelen at ved 9-9 ble regjerende mester sittende. Med våre utfalls sannsynligheter (0.30, 0.20, 0.50) vil det (uten tilleggsregelen) gi en vinstsannsynlighet for A på 0.8139, og forventet antall partier for å kåre en vinner lik 32.4. For å redusere denne forventningen til under 24, er det med våre data nok å kreve 7 vunne partier, men forventningen blir selvsagt større med mer jevnbyrdige spillere. Ønsket ble avvist, og Fischer trakk seg kort etter fra turneringspill, og returmatch ble det aldri noe av. Ved VM-matchen i 1984 mellom den da sittende mester Anatolij Karpov og den unge utfordrerer Garri Kasparov satset man likevel på et slikt opplegg, da det skulle spilles inntil 6 vunne partier. Dette gir med utfalls sannsynligheter (0.30, 0.20, 0.50) som ovenfor et forventet antall partier lik 18.7, og med jevn gode spillere (0.25, 0.25, 0.50) lik 19.1. Likevel gikk det skeis, og matchen ble avbrutt etter 48 partier uten å ha kåret ny mester. Opplegget

reduserte slett ikke rekken av remis, og slikt opplegg er ikke forsøkt senere. Magnus Carlsen spilte best av 12 partier i 2013, 2014, 2016 og 2018, de to siste avgjort med «tie break», og deretter i 2021 best av 14 partier, som var avgjort etter 11 partier på stillingen $7\frac{1}{2} - 3\frac{1}{2}$.

jostein.lillestol@nhh.no

Nytt fra Universitetet i Stavanger

Nyansatte

- Jimmy Huy Tran er ansatt som stipendiat i statistikk.

Nye Mastergrader

- Arezo Shakeri: «Monitored Health Data Analysis Based on State Intensity Regression and Functional Data Analysis» (Veileder: Jörn Schulz)

Nytt fra Norges Idrettshøyskole

Nye PhDer

- Lena Kristin Bache-Mathiesen: “Improving the methodology of training load and injury risk research: an analysis of analyses“ (Veiledere: Morten Wang Fagerland, Thor Einar Andersen)

Nytt fra Oslo senter for biostatistikk og epidemiologi

Nyansettelser

- Permanente forskerstillinger: Henri Pesonen, William Denault, Alvaro Köhn-Luque
- Postdoktorer: Niklas Maltzahn, Maren-Helene Langeland Degnes
- Universitetslektor: Chi Zhang

Nytt fra Tromsø

Nye Mastergrader

- Adrian Daniel Renz: A Price Prediction for Guarantees of Origin. Veileder: Sigrunn H. Sørbye

Nytt fra NTNU

Nyansatte

- John Paige: Førsteamanuensis i statistikk ved Institutt for matematiske fag.
- Khanh Truong: PhD-student.

Nye PhD-grader

- Umat Altay: “Geostatistical Analysis of DHS Data: Accounting for Random Displacement of Survey Locations” Supervisor: Associate Professor Geir-Arne Fuglstad. Co-supervisor has been Professor Andrea Riebler. Committee: 1. opponent: Professor Christel Faes, Hasselt University, Belgium 2. opponent: Dr. Edson Utazi, University of Southampton, UK Internal member: Associate Professor Stefanie Muff, Department of Mathematical Sciences, NTNU Trial lecture: “Inference for model-based geostatistics & geostatistical design” Defence: June 14 2023
- Michail Spitieris: “Physics-Informed Statistical Machine Learning and Methods for Digital Twins” Assessment Committee: 1. opponent: Associate Professor Maurizio Filippone, EURECOM, France 2. opponent: Reader in Mathematics of Data Science Aretha Teckenrump, University of Edinburgh, UK Internal member: Professor Håkon Tjelmeland, Department of Mathematical Sciences, NTNU Trial lecture: “Approximate inference for Bayesian neural network Defence: May 30 2023
- Jorge Sicacha Parada: «Spatial statistical models using citizen science data and professional surveys for biodiversity insight» Supervisor: Professor Ingelin Steinsland, NTNU. Co-supervisors Professor Jan Ketil Rød and Professor Bob O’Hara. Committee: 1. opponent: Prof. Alan Gelfand, Duke University, USA 2. opponent: Prof. Roger Bivand, NHH, Bergen, Norway Member and Administrator: Professor Andrea Riebler, Department of Mathematical Sciences, NTNU Trial lecture: «The Role of Spatial Scale in Spatial Covariance Function Specification» Defense: 20 March 2023

Nytt fra UiB

Nyansatte

- Postdoktor Kazeem Adesina Dauda

Nye Mastergrader

- Mathias Einarsen Ostnes: Modern Variable Selection Methods with Empirical Analysis. Veiledere: Yushu Li og Ingvild M. Helgøy
- Hilde Haugland: Forsikringsprising: En sammenligning av Tweedie generalisert additiv modell og Boosted Tweedie sammensatt poissonmodell. Veileder: Bård Støve
- Thomas F. Mullally: Generative Adversarial Networks for Art Images. Veileder: Bård Støve

Nye PhD-grader

- Ingvild M. Helgøy: Sparse Bayesian learning methods and statistical survival models. Veiledere: Yushu Li og Hans Skaug

Nytt fra UiO

Nyansatte

- Thordis L. Thorarinsdottir. Førsteamanuensis ved Seksjon for statistikk og data science.

Nye Mastergrader

- Håkon Nortvedt: Unit-linked Insurance Policies in the Presence of Credit Risk. Veileder: Frank N. Proske.
- Torgeir Ladstein Waagbøl: Autoregressive Generative Models with Applications to Super Resolution. Veileder: Odd Kolbjørnsen.
- Amund Bjørnevik: Computation of Life Insurance Reserves under Fractional Hull-White Interest Rates. Veileder: Frank N. Proske.
- Jenny Kolstad: Å maksimere utbyttet av en ressurs. Veileder: Tom Lindstrøm.
- Juan Pablo Villa Jiménez: Portfolio optimisation in exponential Lévy markets. Veileder: Fred E. Benth.
- Trym Clementz: Brownske bevegelser på fraktale strukturer. Veileder: Tom Lindstrøm.
- Andreas Krakereid Slåttelid: Advancements in Risk-Free Reference Rates and ESG-linked interest rate swaps. Veileder: Fred E. Benth.
- Sander Finnset Ørnes: SDEs driven by VMLV processes and ambit fields A study of existence and uniqueness. Veileder: Fred Espen Benth.
- Léandre Jean Jules Clemen Hansen: Financial Markets with Delay. Veileder: Giulia Di Nunno.
- Frida Josefine Emberland Straum: Nonstandard Stochastic Analysis in Infinite. Veileder: Tom Lindstrøm.
- Haakon Mugerud: Ranked-based Bayesian clustering and variable selection for high-dimensional data. Veileder: Ida Scheel, Valeria Vitelli.
- Philip Sebastian Sommerfelt: Combining Variational Bayes and GMJMCMC for Scalable Inference on Bayesian Generalized Nonlinear Models. Veileder: Aliaksandr Hubin
- Rohullah Akbari: NLP-Based Automated Conspiracy Detection for Massive Twitter Datasets Large-scale inference with Graphcore Intelligence Processing Units. Veileder: Johannes Langguth, Konstantin Pogorelov, Daniel Schroeder, Andrei Kutuzov.

- Anders Kielland: Integrating Biological Domain Knowledge in Machine Learning Models for Cancer Precision Medicine. Veiledere: Alvaro Köhn-Luque, Ingrid Glad.
- Rayyan Ahmad Shah Syed: Causal discovery with Bayesian networks. Veileder: Johan Pensar.
- Leif-Martin Sæther Sunde: Spherical Priors for Bayesian Deep Learning. Veileder: Geir Storvik.
- Andreas Christensen: Optimal Reinsurance in a Multivariate Case. Veileder: Arne Bang Huseby.
- Sigurd Føreid Fladby: Optimising Multivariate Reinsurance Contracts Under Uncertainty. Veileder: Arne Bang Huseby.
- Alaa Ayoub: Discovering interactions. Veileder: Ingrid Hobæk Haff.

Nye PhD-grader

- Alise Danielle Midtfjord: Machine learning methods for safety-critical systems with time dependency. Prøveforelesning: Bayesian machine learning. Veiledere: Professor Arne Bang Huseby og førsteamanuensis Riccardo De Bin. Komité: Professor Pieter van Gelder, TU Delft, Professor Tore Selland Kleppe, Universitetet i Stavanger, Professor Tom Lindstrøm, Universitetet i Oslo.