UNIVERSITETET I OSLO
Matematisk Institutt

EKSAMEN I:	**ST 104 – Modellering ved stokastiske prosesser**
TID FOR EKSAMEN:	**Mandag 15. juni 1998 kl. 9:00–15:00**
HJELPEMIDLER:	**Kalkulator, abakus, regnestav**

Oppgavesettet inneholder fem oppgaver og er på fire sider

Oppgave 1

La $X_0, X_1, X_2, \ldots$ være en tidsstasjonær Markovkjede med utfallsrom $\{1, \ldots, K\}$, og la på vanlig måte $P_{i,j}$ betegne overgangssannsynlighetene.

(a) Anta at kjeden starter i $X_0 = i_0$. Vis at

$$\Pr\{X_1 = i_1, \ldots, X_n = i_n\} = P_{i_0,i_1} P_{i_1,i_2} \cdots P_{i_{n-1},i_n}.$$

(b) Forklar hvordan man kan beregne to-skritts-overgangssannsynligheter ut fra ett-skritts-overgangssannsynlighetene.

(c) Finn et uttrykk for

$$\Pr\{X_n = j \,|\, X_{n-1} = i, X_{n+1} = k\}.$$

Vis så at en Markovkjedes nå-tilstand, gitt hele fortiden og hele fremtiden, bare avhenger av tilstandene til de nærmeste naboene.
(Dette er et meget fruktbart utgangspunkt for generalisering av Markovmodeller til dimensjon 2 og 3, for eksempel for statistisk bildeanalyse; jfr. foredraget om forskning ved Norsk Regnesentral.)

Oppgave 2

Siméon Denis Poisson (1781–1840) skrev sitt viktigste arbeid om sannsynlighetsteori i 1837, *Recherches sur la Probabilité des Jugements*. Men den såkalte Poissonfordeling, som man kan hevde var kjent av de Moivre lenge før Poisson skrev om den, ble for alvor kjent for nøyaktig 100 år siden, da Ladislaus von Bortkiewics ga ut *Das Gesetz der Kleinen Zahlen* i Leipzig. Nærværende eksamenssett er derfor også en betimelig hyllest til Poisson og hans fordelings mangslungne generaliseringer.

von Bortkiewics illustrerte Poissonfordelingen med data over hvor mange soldater i den prøyssiske armé som var drept av hestespark. Sannsynligheten p for at en bestemt soldat skulle bli drept av hestespark i tjenesten i løpet av ett år var lykkeligvis meget liten, men når man betraktet et helt armékorps bestående av n mann, ville man likevel av og til oppleve slike ulykker. Den klassiske Poissontilnærmelse, som er og var utgangspunktet for betydelige generaliseringer, er den som tas opp i punkt (a).

(a) La X være binomisk (n, p), der n er stor og p er liten; mer formelt tenker vi oss at n vokser og p avtar på en slik måte at produktet $\lambda = np$ holdes fast. Vis at X i grensen blir Poissonfordelt med parameter λ.

Vi skal nå på tur i en Poissonskog. Hvis $N(A)$ betegner antall trær innenfor et område A, så er $N(A)$ Poissonfordelt med parameter $\lambda|A|$, hvor $|A|$ er arealet til A; dessuten er $N(A_1), \ldots, N(A_m)$ uavhengige hvis områdene $A_1, \ldots, A_m$ er disjunkte.

(b) Fra et gitt og tilfeldig valgt punkt i skogen måler jeg avstanden R til det nærmeste tre. Vis at R har sannsynlighetstetthet

$$g(r) = e^{-\lambda\pi r^2}\lambda 2\pi r \quad \text{for } r > 0.$$

(c) Vis at R har forventning $\frac{1}{2}\frac{1}{\sqrt{\lambda}}$.

Hvis vi får tak i et estimat $\widehat{\lambda}$ for λ får man automatisk et estimat for antall trær i hvert område; hvis totalområdet for skogen er $\mathcal{A}$ med areal $|\mathcal{A}|$ er for eksempel $\widehat{\lambda}|\mathcal{A}|$ et estimat for det totale antall trær i skogen.

(d) Jeg har målt distansene $R_1, \ldots, R_{100}$ til de nærmeste trær fra 100 forskjellige faste punkter i skogen. Foreslå en estimator for λ.

(e) Fra mitt gitte og tilfeldig valgte ståsted i skogen defineres, i tillegg til avstanden R til det nærmeste tre, også avstanden S til det nest-nærmeste tre. Finn et uttrykk for simultansannsynlighetstettheten $g(r,s)$ for (R,S).

(f) Bruk dette til å foreslå en mer presis estimator enn den som bare utnytter R_i-informasjon.

Oppgave 3

Vi skal se på en enkel modell for et fenomen av følgende type: et individ er en stor del av tiden i normaltilstand, kalt 0, men av og til rammes individet av et lite anfall, kjennetegnet ved en «opp»-periode, kalt tilstand 1, med påfølgende «ned»-periode, kalt tilstand 2, før individet igjen er tilbake i normaltilstand 0. Det antas at individets tilstandsprosess $X_0, X_1, X_2, \ldots$ over tidspunkter $0, 1, 2, \ldots$ følger en tidshomogen Markovkjedemodell med overgangsmatrise

$$\mathbf{P} = \begin{pmatrix} 1-p_0 & p_0 & 0 \\ 0 & 1-p_1 & p_1 \\ p_2 & 0 & 1-p_2 \end{pmatrix},$$

for visse parameterverdier p_0, p_1, p_2. (Man vil typisk ha p_0 betydelig mindre enn p_1 og p_2; jfr. hovedfagsstudentens foredrag om analyse av EEG-målinger for epilepsipasienter.)

I denne oppgaven kan du (om du trenger det) benytte at hvis M er geometrisk fordelt, med punktsannsynligheter $(1-p)^{m-1}p$ for $m = 1, 2, 3, \ldots$, så er forventning og varians lik henholdsvis $1/p$ og $(1-p)/p^2$.

(a) Anta $X_0 = 0$, og la T_0 være tiden til et anfall starter (altså det første tidspunkt der $X_n = 1$). Finn fordelingen til T_0 og sett opp dens forventningsverdi.

(b) La T være den tid det tar å tilbakelegge en «full syklus», fra start i tilstand 0 ved tid 0 til individet etter å ha gjennomgått anfallets del 1 og del 2 er tilbake i normaltilstanden. Finn forventning og varians for T.

(c) Finn stasjonærsannsynlighetene π_0, π_1, π_2 for kjeden (uttrykt ved p_0, p_1, p_2).

(d) La B_n være antall ganger individet har vært i opp-tilstanden 1 i løpet av de n første tidspunkter (det antas stadig at kjeden starter i $X_0 = 0$). Finn grensen for $\mathrm{E}(B_n/n)$ når n vokser. Finn også ut hva som skjer med $\mathrm{Var}\,(B_n/n)$, og kommenter resultatet.

Oppgave 4

En enkel modell for en innkommende pengestrøm til et forsikringsselskap er som følger. Innbetalinger $Y_1, Y_2, Y_3, \ldots$ skjer på tidspunktene $W_1 < W_2 < W_3 < \cdots$, der disse er modellert som hendelsestidspunkter i en Poissonprosess $\{N(t)\colon t \geq 0\}$ med fast rate λ. Innbetalingene er uavhengige av hverandre med samme fordeling, med $\mathrm{E}Y_i = \xi$ og $\mathrm{Var}\, Y_i = \tau^2$. Dessuten er innbetalingenes størrelser uavhengig av når de måtte ankomme, altså av Poissonprosessen. Den samlede innbetalte sum etter tid t blir altså

$$A(t) = \sum_{i=1}^{N(t)} Y_i.$$

Dette kalles en sammensatt Poissonprosess.

(a) Finn forventning og varians til $A(t)$ gitt at $N(t) = n$, og bruk dette til å finne ubetinget forventning og varians til $A(t)$.

(b) Forsikringsselskapet har muligheten til å investere disse innbetalingene etterhvert som de ankommer, med en avkastningsrate β. En sum av y kroner er altså verd $y\, e^{\beta s}$ kroner etter tid s. Spesielt er verdien av Y_i, som ankom etter tid W_i, blitt oppjustert til $Y_i\, e^{\beta(t-W_i)}$ ved tid t. Selskapets samlede inntjening av denne typen, etter tid t, blir altså

$$B(t) = \sum_{i=1}^{N(t)} Y_i\, e^{\beta(t-W_i)}.$$

Finn forventningen til $B(t)$.

(c) Finn også variansen til $B(t)$.

Oppgave 5

Studenter og kolleger oppsøker en viss professor Nilsen i henhold til en Poissonprosess med rate $\lambda = 3$ pr. døgn. Tiden det tar å få studentene eller kollegene ut av døren igjen varierer fra kasus til kasus, men samtalelengdene kan anses uavhengige og eksponensielt fordelte med parameter $\mu = 12$ pr. døgn. Lengden av den kø som befinner seg utenfor professorens kontor ved tidspunkt t kaller vi $X(t)$. Som en realistisk approksimasjon ser vi i denne oppgave bort fra pauser og arbeidstidsbeskrankninger (hr. Nilsen er dessuten skeptisk til streikevåpenet og dogmatisk uinteressert i lønnskamp).

(a) Forklar hvorfor $\{X(t)\colon t \geq 0\}$ danner en fødsels- og dødsprosess, og finn dens infinitesimale generator (matrisen A med de deriverte $P'_{i,j}(0)$).

(b) Vis at stasjonærfordelingen (eller likevektsfordelingen) for prosessen, altså fordelingen for $X(\infty)$, blir $\pi_j = (1-\rho)\rho^j$ for $j = 0, 1, 2, \ldots$, hvor $\rho = \lambda/\mu$. Hva er sannsynligheten for at en tilfeldig ankommende student eller kollega treffer Nilsen i et ledig øyeblikk?

(c) La W være den tid det tar for en ankommende student eller kollega å nå frem til og dernest samtale ferdig med professor Nilsen. Hvis det er j personer foran henne i samme ærend er altså $W = V_1 + \cdots + V_j + V_0$, der $V_1, \ldots, V_j$ er samtaletidene for de j personene og V_0 den nødvendige konferansetid for henne selv. Finn forventningen til W, ved først å betinge med hensyn på antallet personer foran henne ved ankomsttidspunktet.

(d) Den momentgenererende funksjon for en eksponensielt fordelt variabel med parameter μ er

$$M(t) = \mathrm{E}\, e^{tV} = \int_0^\infty \mu e^{-(\mu - t)v}\, \mathrm{d}v = \frac{\mu}{\mu - t} \quad \text{for } t < \mu.$$

Finn sannsynlighetsfordelingen for W fra forrige punkt, ved hjelp av momentgenererende funksjoner.

Den brøkdel av Nilsens tilmålte tid han kan anvende til egne forskningsprosjekter er altså $\pi_0 = 1 - \rho = 3/4$. Anta nå at det ledige professoratet ved avdelingen blir besatt av en person med analoge egenskaper, slik at Poissonstrømmen av studenter og kolleger nå får to like attraktive personer å henvende seg til. La $Y(t)$ være antall studenter eller kolleger som ved tidspunkt t venter på tur for samtale med enten Nilsen eller hans nye kollega.

(e) Skriv opp parametrene for denne nye fødsels- og dødsprosessen. Den får en ny likevektsfordeling med visse $\pi'_j = \Pr\{Y(\infty) = j\}$. Nilsen interesserer seg særlig for parameteren p_{alene}, den brøkdel av tiden han kan være alene på sitt kontor. Forklar hvorfor $p_{\text{alene}} = \pi'_0 + \frac{1}{2}\pi'_1$, og finn denne.

«Livet byr på kun to ekte gleder: oppdagelser innen matematikk og undervisning av matematikk.» — SIMÉON POISSON (1781–1840)

SLUTT