Norges teknisk–
naturvitenskapelige universitet
Institutt for matematiske fag

Bokmål

Faglig kontakt under eksamen:
Bo Lindqvist 73 59 35 32

EKSAMEN I EMNE SIF5072 STOKASTISKE PROSESSER

Onsdag 31. juli 2002
Tid: 09:00–14:00

Hjelpemidler: Kalkulator HP30S
Statistiske tabeller og formler, Tapir forlag.
K. Rottman: Matematisk formelsamling.
Ett gult ark (A5 med stempel) med egne formler og notater.

Sensur: 21. august 2002.

Oppgave 1

I brønner som bores i petroleumsreservoarer i Nordsjøen finnes ulike typer bergarter. Vi skal i denne oppgaven betrakte et reservoar hvor det finnes tre bergartstyper; 0: Skifer, 1: Skifrig sandstein og 2: Ren sandstein. La $X(t)$ betegne bergartstypen ved posisjon t (målt i meter), der $t = 0$ øverst i reservoaret øker nedover. Vi skal her anta at $\{X(t); t \geq 0\}$ kan modelleres som en kontinuerlig-tid markovprosess slik at

$$\mathrm{P}\{X(t+h) = j | X(t) = i\} = q_{ij}h + o(h) \quad \text{for} \quad i \neq j.$$

Vi skal videre anta følgende verdier for de instantane overgangsintensiter, q_{ij},

$$q_{01} = 4.0 \ \ , \ \ q_{02} = 1.0 \ \ , \ \ q_{10} = 0.5 \ \ , \ \ q_{12} = 1.5 \ \ , \ \ q_{20} = 0 \ \ \text{og} \ \ q_{21} = 0.5.$$

a) Tegn opp et skjema over prosessen.

Hva er gjennomsnittlig tykkelse av et lag med ren sandstein?

Hva er sannsynligheten for at et lag med skifer er minst 1 meter tykt?

Vi skal i første omgang fokusere på sekvensen av bergartstyper nedover langs en brønn og ignorerer dermed foreløpig tykkelsen på lagene. La Z_0 betegne bertartstypen ved $t = 0$, la Z_1 betegne den første bergartstypen som besøkes etter bergartstype Z_0, la Z_2 betegne bertartstypen som besøkes deretter, osv. Kjeden $\{Z_n\}_{n=0}^{\infty}$ blir da en diskret-tid markovkjede (som i læreboka kalles for "the embedded chain").

b) Regn ut overgangsmatrisen for Z_n og vis dermed at den er

$$\mathbf{P} = \begin{bmatrix} 0.00 & 0.80 & 0.20 \\ 0.25 & 0.00 & 0.75 \\ 0.00 & 1.00 & 0.00 \end{bmatrix}.$$

Bestem $\mathrm{P}\{Z_3 = 2 | Z_0 = 2\}$.

Bestem $\mathrm{P}\{Z_3 = 2 | Z_0 = 2 \cap Z_5 = 2\}$.

c) Bestem ekvivalensklassene for markovkjeden Z_n, klassifiser hver tilstand som transient eller rekurrent, samt finn perioden til hver tilstand (Begrunn svarene!).

Begrunn at markovkjeden har en grensefordeling og bestem denne.

d) Forklar med ord hva det vil si at en (diskret-tid) markovkjede er tidsreversibel.

Begrunn med ord eller vis matematisk at Z_n <u>ikke</u> er tidsreversibel.

I resten av oppgaven skal vi igjen betrakte den opprinnelige prosessen $X(t)$ definert i starten av oppgaven.

e) Bestem grensefordelingen for $X(t)$.

Hvor stor andel av reservoaret er skifer?

I gjennomsnitt, hvor mange ganger pr. meter vil det være en overgang fra ren sandstein til skifrig sandstein.

Skifer er tett og fungerer dermed som en barriere for flyt av olje og gass. Ren sandstein er porøs og har gode flytegenskaper. I skifrig sandstein er også flyt av olje og gass mulig, selv om flytegenskapene her er noe dårligere enn i ren sandstein. En viktig størrelse i vurderingen av et reservoar er derfor (vertikal) avstand mellom etterfølgende skiferbarrierer.

f) Bestem gjennomsnittlig tykkelse av en sone med ren sandstein og skifrig sandstein mellom to skiferbarrierer.

I tillegg til gjennomsnittlig vertikal avstand mellom to skiferbarrierer er det av interesse også å bestemme sannsynligheten for at avstanden er større enn eller lik d (der d er et gitt tall). Denne sannsynligheten er det ikke lett å finne ved analytisk regning, men den kan estimeres ved hjelp av simulering.

g) Ta utgangspunkt i at du har tilgjengelig en funksjon som kan generere (vilkårlig mange) uavhengige uniformfordelte tall på intervallet $[0, 1]$ og skriv ned en algoritme som benytter dette til å simulere markovprosessen $X(t)$.

Forklar hvordan den simulerte markovprosessen kan benyttes til å estimere den søkte sannsynlighet.

[*Vink*: Du kan i dette punktet uten bevis benytte at dersom U er uniformfordelt på intervallet $[0, 1]$, så blir $-\ln(U)/\lambda$ eksponensialfordelt med parameter λ.]

Oppgave 2

Anta at kunder ankommer en behandlingsstasjon ifølge en poissonprosess med intensitet λ. Anta videre at antall behandlingsenheter er ubegrenset, slik at kunder som ankommer umiddelbart påbegynner sin behandling. Vi antar at behandlingstidene for ulike kunder er uavhengige av hverandre og kommer fra en felles sannsynlighetsfordeling. La $F(t)$ være kumulativ fordelingsfunksjon for behandlingstidene. Kunder som er ferdigbehandlet forlater umiddelbart behandlingsstasjonen.

La $X(t)$ betegne antall kunder som ankommer i tidsintervallet fra 0 til t og la $Y(t)$ være antall kunder under behandling ved tid t. Anta at ingen kunder er under behandling til tid $t = 0$ slik at $X(0) = Y(0) = 0$.

a) Forklar betegnelsen $M/G/\infty$ for køsystemet beskrevet over.

For $\lambda = 0.1$ (pr. minutt), bestem sannsynligheten for at (nøyaktig) to kunder ankommer behandlingsstasjonen i løpet av en 15 minutters periode.

Hva er forventet tid til kunde nummer k ankommer behandlingsstasjonen? (Gi svaret som funksjon av λ.)

I neste punkt (og kun der) skal vi anta at sannsynlighetsfordelingen til behandlingstidene er en eksponensialfordeling med parameter μ, dvs. $F(t) = 1 - \exp\{-\mu t\}$. Dermed blir $Y(t)$ en fødsels- og dødsprosess.

b) Skriv ned fødsels- og dødsintensitetene for $Y(t)$.

Finn grensefordelingen for $Y(t)$.

I det lange løp, hvor stor andel av tiden vil det være tomt for kunder i behandlingsstasjonen?

For en $M/G/\infty$-kø kan man regne ut fordelingen til $Y(t)$ også for en vilkårlig sannsynlighetsfordeling $F(t)$, noe vi skal gjøre i det siste punktet i oppgavesettet.

c) Vis at $Y(t)$ er poissonfordelt med forventning lik $\lambda \int_0^t (1 - F(u))\mathrm{d}u$. (*Vink*: Benytt setningen om total sannsynlighet og beting på verdien til $X(t)$.)

Finn spesielt grensefordelingen for $Y(t)$ og vis at dette resultatet er konsistent med resultatet du fant i punkt **b**).

Formelsamling:

Setningene om total sannsynlighet og dobbelforventning

La $B_1, B_2, \ldots$ være parvis disjunkte hendelser med $P(\cup_{i=1}^{\infty} B_i) = 1$. Da gjelder

$$P(A|C) = \sum_{i=1}^{\infty} P(A|B_i \cap C)P(B_i|C),$$
$$E[X|C] = \sum_{i=1}^{\infty} E[X|B_i \cap C]P(B_i|C).$$

Markovkjeder i diskret tid

Chapman-Kolmogorov ligningene

$$P_{ij}^{(m+n)} = \sum_{k=0}^{\infty} P_{ik}^{(m)} P_{kj}^{(n)}.$$

For en irredusibel og ergodisk markovkjede eksisterer $\pi_j = \lim_{n\to\infty} P_{ij}^n$ og er gitt av ligningene

$$\pi_j = \sum_i \pi_i P_{ij} \quad \text{og} \quad \sum_i \pi_i = 1.$$

For transiente tilstander i, j og k er forventet tid i tilstand j gitt start i tilstand i, s_{ij},

$$s_{ij} = \delta_{ij} + \sum_k P_{ik} s_{kj}.$$

For transiente tilstander i og j er sannsynligheten for en eller annen gang å returnere til tilstand j gitt start i tilstand i, f_{ij},

$$f_{ij} = (s_{ij} - \delta_{ij})/s_{jj}.$$

Poissonprosess

Ventetid til n-te hendelse (n-te arrival time), S_n, har sannsynlighetstetthet

$$f_{S_n}(t) = \frac{\lambda^n t^{n-1}}{(n-1)!} e^{-\lambda t} \quad \text{for} \quad t \geq 0.$$

Gitt at antall hendelser $N(t) = n$, så har $S_1, S_2, \ldots, S_n$ simultantetthet

$$f_{S_1,S_2,\ldots,S_n|N(t)}(s_1, s_2, \ldots, s_n|n) = \frac{n!}{t^n} \quad \text{for} \quad 0 < s_1 < s_2 < \ldots < s_n \leq t.$$

Markovprosesser i kontinuerlig tid

Chapman-Kolmogorov ligningene

$$P_{ij}(t+s) = \sum_{k=0}^{\infty} P_{ik}(t)P_{kj}(s).$$

Kolmogorovs forover-ligninger

$$P'_{ij}(t) = \sum_{k\neq j} q_{kj}P_{ik}(t) - v_j P_{ij}(t).$$

Kolmogorovs bakover-ligninger

$$P'_{ij}(t) = \sum_{k\neq i} q_{ik}P_{kj}(t) - v_i P_{ij}(t).$$

Hvis $P_j = \lim_{t\to\infty} P_{ij}(t)$ eksisterer, er P_j gitt av

$$v_j P_j = \sum_{k\neq j} q_{kj} P_k \quad \text{og} \quad \sum_j P_j = 1.$$

Spesielt for fødsels- og dødsprosesser

$$P_0 = \frac{1}{\sum_{k=0}^{\infty} \theta_k} \quad \text{og} \quad P_k = \theta_k P_0 \quad \text{for} \quad k = 1, 2, \ldots$$

der

$$\theta_0 = 1 \quad \text{og} \quad \theta_k = \frac{\lambda_0 \lambda_1 \cdot \ldots \cdot \lambda_{k-1}}{\mu_1 \mu_2 \cdot \ldots \cdot \mu_k} \quad \text{for} \quad k = 1, 2, \ldots$$

Kø-teori

For gjennomsnittlig antall kunder i systemet, L, gjennomsnittlig tid pr. kunde i systemet, W, betjeningstid, S, og gjennomsnittlig gjenværende arbeid i systemet, V, gjelder

$$L = \lambda_a W.$$

$$L_q = \lambda_a W_q.$$

$$V = \lambda_a E[SW_q^*] + \lambda_a E[S^2]/2.$$