NTNU

Kunnskap for en bedre verden

Institutt for matematiske fag

Eksamensoppgave i **TMA4250 Romlig Statistikk**

Faglig kontakt under eksamen: Professor Henning Omre

Tlf: 90937848

Eksamensdato: 5. juni 2015

Eksamenstid (fra–til): 09:00-13:00

Hjelpemiddelkode/Tillatte hjelpemidler: C:

Tabeller og Formler i Statistikk, Tapir
NTNU godkjent kalkulator
Personlige, håndskrevne, gule titte-ark - A5-format

Målform/språk: bokmål

Antall sider: 3

Antall sider vedlegg: 0

Kontrollert av:

Dato Sign

Oppgave 1 KONTINUERLIGE STOKASTISKE FELT.

Betrakt et kontinuerlig stokastisk felt $\{R(x); x \in \mathbb{R}^1\}$, definert på et 1D referanserom $[-\infty, \infty]$. Modelparametrene er:

$$\mathrm{E}\{R(x)\} = \mu_R(x) = 10 - (x-a)^2 \; ; x \in \mathbb{R}^1$$
$$\mathrm{Cov}\{R(x'), R(x'')\} = c_R(\tau) = \exp\{-\frac{1}{10}\tau^2\} \; ; x', x'' \in \mathbb{R}^1$$

med $\tau = x'' - x'$.

Feltet har altså høyeste forventede verdi lik 10 i den ukjente lokasjonen $x = a$, og kovariansfunksjonen er gitt og translasjonsinvariant.

a) Spesifiser kravet som stilles til en gyldig romlig kovariansfunksjon.

Tegn en skisse av kovariansfunksjonen, $c_R(\tau)$ og forklar hvilke egenskaper ved feltet en kan tolke fra skissen.

La det stokastiske feltet være observert i lokasjonene $(x_1, \ldots, x_n)$, det vil si at observasjonene er $R_o = (R(x_1), \ldots, R(x_n))$.

b) Definer den lineære estimatoren for lokasjonen til høyeste forventet verdi a:

$$\hat{a} = \sum_{i=1}^{n} \beta_i R(x_i)$$

hvor $\beta = (\beta_1, \ldots, \beta_n)$ er ukjente vekter som skal bestemmes.

Utled et uttrykk for den beste lineære forventningsrette estimatoren under kvadratisk tap for a basert på observasjonssettet R_o. Bare minimeringsproblemet som må løses for å kunne bestemme vektene behøver å angis.

Definer det tilhørende deriverte feltet $\{R^*(x) = \frac{dR(x)}{dx}; x \in \mathbb{R}^1\}$.

c) Utled uttrykk for følgende modellparametre til det deriverte feltet $\{R^*(x); x \in \mathbb{R}^1\}$:

$$\mathrm{E}\{R^*(x)\} = \mu_{R^*}(x) \; ; x \in \mathbb{R}^1$$
$$\mathrm{Cov}\{R^*(x'), R^*(x'')\} = c_{R^*}(\tau) \; ; x', x'' \in \mathbb{R}^1$$

med $\tau = x'' - x'$.

Tegn en skisse av kovariansfunksjonen, $c_{R^*}(\tau)$ og forklar hvilke egenskaper ved det deriverte feltet en kan tolke fra skissen.

Anta at det stokastiske deriverte feltet også er observert i lokasjonene $(x_1, \ldots, x_n)$, det vil si at derivert-observasjonene er $R^*_o = (R^*(x_1), \ldots, R^*(x_n))$.

d) Definer den lineære estimatoren for lokasjonen til høyeste forventet verdi a:

$$\hat{a} = \sum_{i=1}^{n} \beta_i R(x_i) + \sum_{i=1}^{n} \gamma_i R^*(x_i)$$

med ukjente vekter $\beta = (\beta_1, \ldots, \beta_n)$ og $\gamma = (\gamma_1, \ldots, \gamma_n)$ som skal bestemmes.

Utled et uttrykk for den beste lineære forventningsrette estimatoren under kvadratisk tap for a basert på observasjonssettene R_o og R^*_o. Bare minimeringsproblemet som må løses for å kunne bestemme vektene behøver å angis.

Oppgave 2 HENDELSES STOKASTISKE FELT.

Betrakt et punkt stokastisk felt $\mathbb{F}_X : \{X_i; i = 1, \ldots, N; \mathcal{D} \subset \mathbb{R}^2\}$ definert over området $\mathcal{D} : [0, 10] \times [0, 10]$ $[m^2] \in \mathbb{R}^2$.

Anta først at feltet er et homogent Poisson stokastisk felt med konstant intensitetsparameter $\lambda \geq 0$ $[m^{-2}]$.

a) Spesifiser sannsynligheten for at det ikke finnes noen punkter i området $\mathcal{D}$. Begrunn svaret.

Spesifiser forventet antall punkter i området $\mathcal{D}$. Begrunn svaret.

Gitt at antallet punkter i området $\mathcal{D}$ er $N = 10$, spesifiser sannsynligheten for at det er akkurat 3 punkter i del-området $\mathcal{D}_\Delta : [0, 5] \times [0, 5]$ $[m^2]$. Begrunn svaret.

Anta heretter at feltet er et Cox stokastisk felt hvor, betinget på intensitetsparameteren λ, feltet er et homogent Poisson felt som definert over. Intensitetsparametren λ er nå stokastisk med sannsynlighetstetthet $f(\lambda) = \frac{1}{\mu} \exp\{-\frac{\lambda}{\mu}\}$; $\lambda \geq 0$.

b) Utled uttrykk for forventet antall punkter $\mathrm{E}\{N\}$ og variansen til antall punkter $\mathrm{Var}\{N\}$ for Cox feltet.

Definer S som korteste avstand mellom midtpunktet i området $\mathcal{D}$ og nærmeste punkt i punktfeltet $\mathbb{F}_X$. Anta at randeffekter forårsaket av at området er endelig kan ignoreres.

c) Utled et uttrykk for sannsynlighetstettheten til S for Cox feltet.

Oppgave 3 MOSAIKK STOKASTISKE FELT.

Betrakt et mosaikk stokastisk felt $L : \{L_x\, ; x \in \mathcal{L}_\mathcal{D}\}$ hvor $\mathcal{L}_\mathcal{D}$ er et regulært grid med n noder over $\mathcal{D} \subset \mathbb{R}^2$, og $L_x \in \Omega_l : \{W, B\}$. Det betyr at variablen L_x kan tilhøre en av klassene hvit (W) eller sort (B) for hver $x \in \mathcal{L}_\mathcal{D}$.

Definer følgende Gibbs modell for feltet:

$$\text{Prob}\{L = l; \beta\} = \text{const}_\beta \times \exp\{\beta \Sigma_{<u,v>} I(l_u = l_v)\}$$

hvor $u, v \in \mathcal{L}_\mathcal{D}$, $< u, v >$ representerer alle par av nærmeste naboer i griddet $\mathcal{L}_\mathcal{D}$ og $I(A)$ er en indikatorfunksjon som tar verdien 1 når A er sann og 0 ellers. Feltet er altså et Ising stokastisk felt. Modellparameteren β antas å være kjent.

Realisasjoner av feltet kan, teoretisk sett, genereres av en en-node ('single-site') Markov chain Monte Carlo (McMC) Metropolis-Hasting (M-H) simuleringsalgoritme. Denne algoritmen er iterativ og baserer seg på et forslags- og aksept/forkast-konsept i hver iterasjon. To mulige forslagsfordelinger er: trekk en node uniformt over $\mathcal{L}_\mathcal{D}$ og oppdater node-klassen ved enten å a) bytte klasse eller b) trekke klasse uniformt over Ω_l , uavhengig av klassene i naboskapet.

a) Bruk formell algoritmebeskrivelse og matematisk formalisme i besvarelsen, og vær særs nøye med notasjonen.

Spesifiser en-node McMC M-H simuleringsalgoritmen for feltet.

Spesifiser de to forslagsfordelingene skissert over og diskuter deres karakteristika.

Utled beregningseffektive uttrykk for akseptsannsynligheten for de to forslagsfordelingene.

Sammenlikn de to algoritmene definert av de to forslagsfordelingene ved å regne ut sannsynligheten for at realisasjonen av feltet endres i en iterasjon, og vurder hva dette betyr for konvergens- og miksings-egenskapene til algoritmene.