TMA 4250 ROMLIG STATISTIKK
VÅREN 2008
Eksamen: 23. mai 2008
08.05.2008 H. OMRE

LØSNINGSFORSLAG

Oppgave 1 Kontinuerlig stokastisk felt

$$\{R(x);\ x \in \mathcal{D} \subset \mathbb{R}^2\}$$

$$E\{R(x)\} = \mu_R \quad \text{- ukjent}$$

$$\mathrm{Cov}\{R(x'), R(x'')\} = \sigma_R^2 \rho_R(x'-x'') \quad \text{- kjent}$$

a) $\rho_R(x'-x'')$ - pos. def korrelasjonsfunksjon

Krav:

$$\rho(0) = 1$$

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j \rho_R(x_i - x_j) > 0$$

$$\forall (\alpha_1, \dots, \alpha_n)$$

$$\forall n$$

b) Observasjoner: $R(x_1) = r_1$ og $R(x_2) = r_2$

Estimator: $\hat{\mu}_R = \alpha_0 + \alpha_1 R(x_1) + \alpha_2 R(x_2)$

Forv. rettet

$$E\{\hat{\mu}_R - \mu_R\} = 0$$

$$\alpha_0 + \alpha_1 E\{R(x_1)\} + \alpha_2 E\{R(x_2)\} = \mu_R$$

$$\alpha_0 + (\alpha_1 + \alpha_2)\mu_R = \mu_R$$

$$\alpha_0 = 0$$

$$\alpha_1 + \alpha_2 = 1$$

Estimeringsvarians:

$$\mathrm{Var}\{\hat{\mu}_R - \mu_R\} = \mathrm{Var}\{\alpha_1 R(x_1) + \alpha_2 R(x_2) - \mu_R\}$$

$$= \alpha_1^2 \sigma_R^2 + \alpha_2^2 \sigma_R^2 + 2\alpha_1\alpha_2 \sigma_R^2 \rho(x_1 - x_2)$$

Kriterium:

$$\hat{\alpha} = \underset{\substack{\alpha \\ \alpha_1+\alpha_2=1}}{\mathrm{argmin}} \mathrm{Var}\{\hat{\mu}_R - \mu_R\}$$

herav

$$\hat{\alpha} = \underset{\alpha_1}{\mathrm{argmin}}\{\alpha_1^2 \sigma_R^2 + (1-\alpha_1)^2\sigma_R^2 + 2\alpha_1(1-\alpha_1)\sigma_R^2 \rho(x_1-x_2)\}$$

$$\frac{d\cdot}{d\alpha_1} = 0 \Rightarrow 2\alpha_1\sigma_R^2 + 2(1-\alpha_1)(-1)\sigma_R^2 + 2\sigma_R^2 \rho(x_1-x_2)[(1-\alpha_1) + \alpha_1(-1)] = 0$$

$$[1-2\alpha_1][\rho(x_1-x_2) - 1] = 0 \Rightarrow \underline{\underline{\alpha_1 = 1/2}}$$

$$\underline{\underline{\alpha_2 = 1 - \alpha_1 = 1/2}}$$

Kan også sees ved symmetribetraktninger!!

$$\hat{\mu}_R = \frac{1}{2}[R(x_1) + R(x_2)]$$

$$\mathrm{Var}\{\hat{\mu}_R - \mu_R\} = \frac{1}{2}\sigma_R^2[1 + \rho(x_1 - x_2)]$$

c) Observasjoner: $R(x_1) = r_1$ og $\Delta_{23} = [R(x_2) - R(x_3)] = \delta_{23}$

Prediktor:

$$\hat{R}(x_0) = \alpha_0 + \alpha_1 R(x_1) + \alpha_2 \Delta_{23}$$

Forv. retthet

$$E\{\hat{R}(x_0) - R(x_0)\} = 0$$

$$\Rightarrow \alpha_0 + \alpha_1 \mu_R + \alpha_2 \cdot 0 = \mu_R$$

$\alpha_0 = 0$

$\alpha_1 = 1$

α_2 – vilkårlig

Pred. varians

$$\mathrm{Var}\{\hat{R}(x_0) - R(x_0)\} = \mathrm{Var}\{ R(x_1) + \alpha_2 [R(x_2) - R(x_3)] - R(x_0) \}$$

$$= \sigma_R^2 + \alpha_2^2 [\sigma_R^2 - 2\sigma_R^2 \rho_R(x_2 - x_3) + \sigma_R^2] + \sigma_R^2$$
$$+ 2\alpha_2 \sigma_R^2 [\rho_R(x_1 - x_2) - \rho_R(x_1 - x_3)] - 2\sigma_R^2 \rho_R(x_1 - x_0) - 2\alpha_2 \sigma_R^2 [\rho_R(x_0 - x_2) - \rho_R(x_0 - x_3)]$$

Kriterium:

$$\hat{\alpha}_2 = \operatorname{argmin} \mathrm{Var}\{\hat{R}(x_0) - R(x_0)\}$$

$$\frac{d\cdot}{d\alpha_2} = 0 \Rightarrow 2\alpha_2 [2\sigma_R^2 (1 - \rho_R(x_2 - x_3))]$$
$$+ 2\sigma_R^2 [\rho_R(x_1 - x_2) - \rho_R(x_1 - x_3)]$$
$$- 2\sigma_R^2 [\rho_R(x_0 - x_2) - \rho_R(x_0 - x_3)] = 0$$

$$\Rightarrow \hat{\alpha}_2 = \frac{[\rho_R(x_0 - x_2) - \rho_R(x_0 - x_3)] - [\rho_R(x_1 - x_2) - \rho_R(x_1 - x_3)]}{2[1 - \rho_R(x_2 - x_3)]}$$

$$\hat{R}(x_0) = R(x_1) + \hat{\alpha}_2 \Delta_{23}$$

$$\mathrm{Var}\{\hat{R}(x_0) - R(x_0)\} = \cdot \cdot \cdot \cdot$$ sett inn $\hat{\alpha}_2$!

Oppgave 2 Hendelses stokastiske felt

$$\{\underline{X}_i ; i=1,\dots,N ; \mathcal{D} \subset \mathbb{R}^2\}$$

$$\mathcal{A}, \mathcal{B} \subset \mathcal{D}$$

$N(\mathcal{A}) = \#$ punkter i $\mathcal{A}$

a) Homogent Poisson felt med int λ - kjent

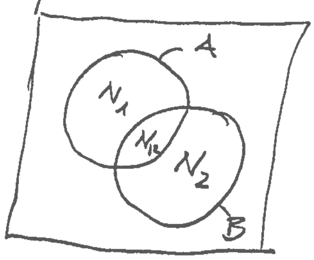
$$N(\mathcal{A}) \sim Poi(n; \lambda|\mathcal{A}|) = \frac{[\lambda|\mathcal{A}|]^n}{n!} e^{-\lambda|\mathcal{A}|}$$

herav

$$E\{N(\mathcal{A})\} = \lambda|\mathcal{A}|$$

$$Var\{N(\mathcal{A})\} = \lambda|\mathcal{A}|$$

b) Homogent Poisson felt med int λ - kjent



$$Cov\{N(\mathcal{A}), N(\mathcal{B})\}$$

$$= Cov\{(N_1 + N_{12}), (N_2 + N_{12})\}$$

$$= \underbrace{Cov\{N_1, N_2\}}_{=0} + \underbrace{Cov\{N_1, N_{12}\}}_{=0} + \underbrace{Cov\{N_{12}, N_2\}}_{=0} + \underbrace{Cov\{N_{12}, N_{12}\}}_{=Var\{N_{12}\}}$$

$$= Var\{N(A \cap B)\}$$

$$= \lambda|A \cap B|$$

Kommentarer:

- hvis $\mathcal{A} = \mathcal{B}$ $\quad Cov\{N(\mathcal{A}), N(\mathcal{A})\} = Var\{N(\mathcal{A})\} = \lambda|\mathcal{A}|$ ok!
- hvis $\mathcal{A} \cap \mathcal{B} = \emptyset$ $\quad Cov\{N(\mathcal{A}), N(\mathcal{B})\} = \lambda|\emptyset| = 0$ ok!
  fordi $[N(\mathcal{A}), N(\mathcal{B})]$ uavh.

c) Cox felt med stok. int $\lambda \sim f(\lambda) \begin{cases} E\{\lambda\} = \mu_\lambda \\ Var\{\lambda\} = \sigma_\lambda^2 \end{cases}$

HUSK $\quad E_X\{X\} = E_Y\{E_X\{X|Y\}\}$

$$Cov\{N(A), N(B)\} = Cov\{N_A, N_B\}$$

$$= E\{N_A \cdot N_B\} - E\{N_A\}E\{N_B\}$$

videre:

$$E_N\{N_A\} = E_\lambda\{E_N\{N_A|\lambda\}\} = E_\lambda\{\lambda|A|\} = \mu_\lambda|A|$$

$$E_N\{N_B\} = \qquad = \mu_\lambda|B|$$

$$E\{N_A \cdot N_B\} = E_\lambda\{E_N\{N_A N_B|\lambda\}\}$$

$$= E_\lambda\{Cov\{N_A, N_B|\lambda\} + E\{N_A|\lambda\}E\{N_B|\lambda\}\}$$

$$= E_\lambda\{\lambda|A \cap B| + \lambda^2|A||B|\}$$

$$= \mu_\lambda|A \cap B| + |A||B|E_\lambda\{\lambda^2\}$$

$$= \mu_\lambda|A \cap B| + |A||B|[\sigma_\lambda^2 + \mu_\lambda^2]$$

setter inn:

$$Cov\{N_A, N_B\} =$$

$$= \mu_\lambda|A \cap B| + |A||B|[\sigma_\lambda^2 + \mu_\lambda^2] - \mu_\lambda|A| \cdot \mu_\lambda|B|$$

$$= \underline{\underline{\mu_\lambda|A \cap B| + \sigma_\lambda^2|A||B|}}$$

Kommentar:

- $\sigma_\lambda^2 = 0$ ie $\lambda = \mu_\lambda$ $\quad Cov\{N_A, N_B\} = \lambda|A \cap B|$ ok
- $\sigma_\lambda^2 \neq 0$ aldri uavh mellom $[N_A, N_B]$ - koplet via int λ!

Oppgave 3 Mosaikk stokastisk felt

$L : \{L_s ; s \in \mathcal{L}_D\}$ $L_s \in \{-1, 1\}$

Naboskapssystem $\delta : \{\delta(s) ; s \in \mathcal{L}_D\}$

'Clique'-system $C : \{c_1, \ldots, c_n\}$

a) Gibbs formulering

$$\text{Prob}\{L = l\} = \text{const} \times \exp\{-\sum_{c \in C} V_c(l)\}$$

med

$V_c(l)$ - funksjon av $l_s \in c$ bare

Markovformulering

$$\text{Prob}\{L_s = l_s \mid L_r = l_r ; r \neq s, r \in \mathcal{L}_D\}$$

$$= \text{Prob}\{L_s = l_s \mid L_r = l_r ; r \in \delta(s)\} ; \forall s \in \mathcal{L}_D$$

MERK!

Gibbsformuleringen spesifiserer simultanfordelingen for hele L.

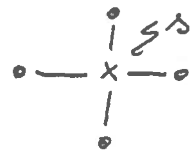
Markovformuleringen består av univariate betingete fordelinger for hver node $s \in \mathcal{L}_D$

Hammersley-Clifford Theorem impliserer at formuleringene er equivalente for tilhørende naboskaps/clique-systemer.

b) 'Clique' systemet består av alle:

C: 

tilhørende naboskap:

$\partial(s)$: 

Gitt observasjonene $L = l^o : \{l_s^o ; s \in \mathcal{L}_D\}$

<u>Full-likelihood:</u>

$$FL(\beta) = Prob\{L = l^o ; \beta\} =$$
$$= const(\beta) \times \exp\Big\{\beta \sum_{\substack{u,v \in \mathcal{L}_D \\ <u,v>}} l_u^o \cdot l_v^o\Big\}$$

Maksimum full-likelihood estimator:

$$\hat{\beta} = \underset{\beta}{\arg\max}\{\log FL(\beta)\} \quad \underbrace{\log const(\beta) + \beta \sum l_u^o l_v^o}$$

- det krever et uttrykk for const(β) som er svært krevende å bestemme.

<u>Pseudo-likelihood:</u>

$$PL(\beta) = \prod_{s \in \mathcal{L}_D} Prob\{L_s = l_s^o \mid L_r = l_r^o ; r \in \partial(s) ; \beta\}$$

Maksimum pseudo-likelihood estimator:

$$\tilde{\beta} = \underset{\beta}{\arg\max}\{\log PL(\beta)\} \quad \underbrace{\sum \log Prob\{\cdot\}}$$

- konstanten i den betingete fordelingen er avh av β, men kan beregnes lett som sum av to ledd. Selve optimeringen må gjøres numerisk