Avd. Matematisk statistik

KTH Matematik

TENTAMEN I SF2941 (f d 5B1541) SANNOLIKHETSTEORI OCH LINJÄRA MODELLER ONSDAGEN DEN 28 MAJ 2008 KL 08.00–13.00

Examinator: Timo Koski, tel. 790 71 34, e-post: timo@math.kth.se

Tillåtna hjälpmedel: *Appendix 2* ur A.Gut: An Intermediate Course in Probability (utdelas). *Formulas for probability theory and linear models* SF2941 (utdelas). Räknare.

Införda beteckningar skall förklaras och definieras. Resonemang och uträkningar skall vara så utförliga att de är lätta att följa. Numeriska svar skall anges med minst två siffrors noggrannhet. Resultat i deluppgift som inte lösts får användas i andra deluppgifter.

Tentamen består av 6 uppgifter. Varje korrekt lösning ger 10 poäng. Gränsen för godkänt är preliminärt 30 poäng. Möjlighet att komplettera ges för de tentander med 28–29 poäng. Tid och plats för komplettering kommer att anges på kursens hemsida. Det ankommer på dig själv att ta reda på om du har rätt att komplettera.

Lösningarna till tentamen kommer att införas på
`http://www.math.kth.se/matstat/gru/5b1541/`
onsdag den 28 maj 2008 tidigast kl. 13.05. Tentamen kommer att finnas tillgänglig på elevexpeditionen sju veckor efter skrivningstillfället.

Lycka till!

- -

Uppgift 1

$\mathbf{X} = (X_1, X_2)' \in N(\mu, C)$ där

$$\mu = \begin{pmatrix} 10 \\ 15 \end{pmatrix} \text{ och } C = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 4 \end{pmatrix}.$$

Bestäm den betingade fördelningen för X_1 givet $X_1 + X_2 = 30$
samt beräkna $P(X_1 \geq 10 | X_1 + X_2 = 30)$. (10 p)

Uppgift 2

Betrakta en Poissonprocess $\{N(t);\ t \geq 0\}$ med parameter λ.

Beräkna $P(N(1) < N(2) < N(3))$. (10 p)

Uppgift 3

Den stokastiska variabeln Θ är $\in U(0,1)$. För givet θ är den stokastiska variabeln X binomialfördelad, d.v.s. $X|\Theta = \theta \in \text{Bin}(n, \theta)$.

Beräkna den (obetingade) sannolikhetsgenererande funktionen till X och ange X:s sannolikhetsfunktion.

Uppgift 4

X är en diskret stokastisk variabel, vars värden är de icke-negativa heltalen $x = 0, 1, 2, \ldots,$. Vi vet att

$$E[X] = 1.$$

Låt B vara händelsen $B = \{X > 0\}$. Vi betraktar X trunkerad till de positiva hela talen eller $X|B$, dvs. X betingad på B. Vi vet också att

$$X|B \in \text{Fs}(p),$$

dvs. att sannolikhetsfunktionen för $X|B$ är

$$p_{X|B}(x) = (1-p)^{x-1} \cdot p, \quad x = 1, 2, 3, \ldots$$

Bestäm den momentgenererande funktionen (m.g.f.) för X. (10 p)

Uppgift 5

Vi antar att X är Ge(p) där $0 < p < 1$ dvs att $P(X = k) = p(1-p)^k, , k = 0, 1, 2, \cdots$. Visa att pX konvergerar i fördelning då $p \downarrow 0$ samt ange gränsfördelningen. (10 p)

Uppgift 6

$X \in N(0,1)$ och $Y \in N(0,1)$ och X och Y är oberoende. Vi bildar en stokastisk variabel Z enligt

$$Z = \begin{cases} Y, & \text{om } \lambda Y > X \\ -Y, & \text{om } \lambda Y \leq X. \end{cases}$$

Visa att sannolikhetstätheten $f_Z(z)$ för Z är

$$f_Z(z) = 2\phi(z)\Phi(\lambda z),$$

där $\phi(z)$ och $\Phi(z)$ är tätheten resp. fördelningsfunktionen för den standardiserade normalfördelningen $N(0,1)$. (10 p)

LÖSNING TENTAMEN I SF2941 SANNOLIKHETSTEORI OCH LINJÄRA MODELLER 08-05-28

Uppgift 1

Vi bestämmer först fördelningen för

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} X_1 \\ X_1 + X_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = B \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$$

som blir

$$N_2\left(B\mu, BCB^T\right) = N_2\left(\begin{pmatrix} 10 \\ 25 \end{pmatrix}, \begin{pmatrix} 1 & 1.5 \\ 1.5 & 6 \end{pmatrix}\right)$$

Fördelningen för Y_1 givet $Y_2 = 30$ ges enligt formelsamlingen av

$$N\left(\mu_1 + \frac{c_{12}}{c_{22}}(30 - \mu_2), c_{11} - \frac{c_{12}^2}{c_{22}}\right) = N(11.25, 0.625)$$

som ger att

$$P\left(X_1 \geq 10 | X_1 + X_2 = 30\right) = 1 - \Phi\left(\frac{10 - 11.25}{\sqrt{0.625}}\right) \approx \underline{0.9429}.$$

Uppgift 2

Vi har

$$P(N(1) < N(2) < N(3)) = (\text{processen är växande}) = P(N(2) - N(1) > 0,\ N(3) - N(2) > 0)$$

$$= (\text{oberoende tillskott}) = P(N(2) - N(1) > 0) P(N(3) - N(2) > 0) = \underline{\left(1 - e^{-\lambda}\right)^2}.$$

Uppgift 3

$$g_X(t) = E(t^X) = (\text{lagen om total förväntan}) = E(E(t^X | \Theta)) = E((1 - \Theta + \theta t)^n) =$$

$$= \int_0^1 (1 - \theta + \theta t)^n 1 \cdot d\theta = \left[\frac{(1 + \theta(t-1))^{n+1}}{(n+1)(t-1)}\right] =$$

$$= \frac{1 - t^{n+1}}{(n+1)(1-t)} = (\text{geometrisk serie}) = \frac{1}{n+1} \sum_0^n t^k$$

dvs $P(X = k) = 1/(n+1)$ för $k = 0, 1, 2, \cdots, n$ dvs att X är likformigt fördelad på $0, 1, \cdots, n$.

Uppgift 4

Definitionen på den momentgenererande funktionen innebär att

$$\psi_X(s) = E\left[e^{sX}\right] = \sum_{x=0}^{\infty} e^{sx} p_X(x) = p_X(0) + \sum_{x=1}^{\infty} e^{sx} p_X(x).$$

Vi har

$$p_{X|B}(x) = \begin{cases} \frac{p_X(x)}{P(B)} & x = 1, 2, \ldots \\ 0 & \text{för övrigt,} \end{cases}$$

dvs.

$$p_X(x) = P(B) \cdot p_{X|B}(x) = P(B) \cdot (1-p)^{x-1} \cdot p, \quad x = 1, 2, 3, \ldots$$

Detta ger

$$\psi_X(s) = p_X(0) + P(B) \sum_{x=1}^{\infty} e^{sx} \cdot (1-p)^{x-1} \cdot p.$$

Här känner vi igen summan $\sum_{x=1}^{\infty} e^{sx} p \cdot (1-p)^{x-1}$ som m.g.f. för $\mathrm{Fs}(p)$. Vi betecknar denna med

$$\psi_{\mathrm{Fs}}(s) = \sum_{x=1}^{\infty} e^{sx} \cdot (1-p)^{x-1} \cdot p,$$

så att

$$\psi_X(s) = p_X(0) + P(B)\psi_{\mathrm{Fs}}(s).$$

Derivering m.a.p. s ger

$$\frac{d}{ds}\psi_X(s) = P(B) \cdot \frac{d}{ds}\psi_{\mathrm{Fs}}(s) \Rightarrow \frac{d}{ds}\psi_X(0) = P(B) \cdot \frac{d}{ds}\phi_{\mathrm{Fs}}(0).$$

Men $\frac{d}{ds}\psi_{\mathrm{Fs}}(0) = \frac{1}{p}$, ty $\psi_{\mathrm{Fs}}(s)$ är m.g.f. för $\mathrm{Fs}(p)$ och väntevärdet av en $\mathrm{Fs}(p)$-fördelad stokastisk variabel är $\psi_{\mathrm{Fs}}(0) = \frac{1}{p}$ (v.s.v. och slå upp i kursens formelsamling). Vi har enligt uppgift

$$1 = E\left[X\right] = \frac{d}{ds}\psi_X(0) = P(B) \cdot \frac{d}{ds}\psi_{\mathrm{Ge}}(0) = P(B) \cdot \frac{1}{p},$$

dvs.

$$P(B) = p.$$

Dessutom fås

$$p_X(0) = 1 - P(B) = 1 - p.$$

Enligt det ovanstående har vi således

$$\psi_X(s) = (1-p) + p\psi_{\mathrm{Fs}}(s).$$

Enligt kursens formelsamling (eller med summering av en geometrisk serie) är m.g.f. för $\mathrm{Fs}(p)$

$$\phi_{\mathrm{Fs}}(s) = \frac{pe^s}{1-(1-p)e^s}.$$

SVAR: $\underline{\phi_X(s) = 1 - p + \frac{p^2 e^s}{1-(1-p)e^s}}.$

Uppgift 5

Den karaktäristiska funktionen för X är enligt formelsamlingen

$$\varphi_X(t) = p/(1-(1-p)e^{it}),$$

så vi får för pX karaktäristiska funktionen

$$\varphi_{pX}(t) = \varphi_X(pt) = p/(1-(1-p)e^{itp}) = (\text{Taylorutveckling}) =$$

$$p/(1-(1-p)(1+itp+o(p))) = 1/(1-it+o(p)/p) \to 1/(1-it)$$

som är karaktäristisk funktion för Exp(1)-fördelningen. Enligt kontinuitetssatsen gäller alltså att pX konvergerar i fördelning mot Exp(1).

Uppgift 6

Vi beräknar fördelningsfunktionen $F_Z(z)$ för ett godtyckligt z dvs.

$$F_Z(z) = P\left(Z \le z\right)$$

och enligt konstruktionen på Z fås händelsen $Z \le z$ som unionen av två disjunkta händelser och den sökta sannolikheten blir

$$P\left(Z \le z\right) = P\left(\{Y \le z\} \cap \{X < \lambda Y\}\right) + P\left(\{-Y \le z\} \cap \{X \ge \lambda Y\}\right). \tag{1}$$

Vi tar den forsta av dessa och får

$$P\left(\{Y \le z\} \cap \{X < \lambda Y\}\right) = \int_{-\infty}^{z} P\left(X < \lambda y \mid Y = y\right) f_Y(y)dy$$

och eftersom X och Y är oberoende erhåller vi

$$= \int_{-\infty}^{z} P\left(X < \lambda y\right) f_Y(y)dy =$$

$$= \int_{-\infty}^{z} \left(\int_{-\infty}^{\lambda y} \phi(u)du \right) f_Y(y)dy = \int_{-\infty}^{z} \Phi\left(\lambda y\right) f_Y(y)dy$$

där $\Phi(y)$ är fördelningsfunktionen för den standardiserade normalfördelningen, enär $X \in N(0,1)$. Detta kan vi ytterligare skriva som

$$= \int_{-\infty}^{z} \Phi\left(\lambda y\right) \phi(y)dy,$$

där $\phi(y)$ är täthetsfunktionen för den standardiserade normalfördelningen, eftersom $Y \in N(0,1)$.
För den andra termen på högra sidan i (1) gäller att

$$P\left(\{-Y \le z\} \cap \{X \ge \lambda Y\}\right) = P\left(\{Y \ge -z\} \cap \{X \ge \lambda Y\}\right)$$

På samma sätt som i det första fallet fås att

$$P\left(\{Y \ge -z\} \cap \{X \ge \lambda Y\}\right) = \int_{-z}^{\infty} P\left(X \ge \lambda y \mid Y = y\right) f_Y(y)dy$$

$$= \int_{-z}^{\infty} P\left(X \geq \lambda y\right) \phi(y) dy$$

$$= \int_{-z}^{\infty} \left(1 - P\left(X < \lambda y\right)\right) \phi(y) dy = \int_{-z}^{\infty} \left(1 - \Phi\left(\lambda y\right)\right) \phi(y) dy.$$

Med variabelbytet $y = -u$ fås

$$= -\int_{z}^{-\infty} \left(1 - \Phi\left(-\lambda u\right)\right) \phi(-u) du = \int_{-\infty}^{z} \Phi\left(\lambda u\right) \phi(u) du.$$

där vi använde oss av en känd grundregel för integraler, symmetrin hos ϕ, dvs. $\phi(-u) = \phi(u)$ och formeln $\Phi\left(-\lambda u\right) = 1 - \Phi\left(\lambda u\right)$. Med insättning i (1) fås således

$$P\left(Z \leq z\right) = \int_{-\infty}^{z} \Phi\left(\lambda y\right) \phi(y) dy + \int_{-\infty}^{z} \Phi\left(\lambda u\right) \phi(u) du =$$

dvs.

$$= 2 \int_{-\infty}^{z} \Phi\left(\lambda y\right) \phi(y) dy.$$

En derivering med avseende på z ger den sökta sannolikhetstätheten, dvs.

$$f_Z(z) = \frac{d}{dz} F_Z(z) = \frac{d}{dz} P\left(Z \leq z\right) = 2\phi(z)\Phi\left(\lambda z\right).$$