Avd. Matematisk statistik

KTH Matematik

TENTAMEN I SF2941 (f d 5B1541) SANNOLIKHETSTEORI OCH LINJÄRA MODELLER TISDAGEN DEN 19 AUGUSTI 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.

Tentamen kommer att finnas tillgänglig på elevexpeditionen sju veckor efter skrivningstillfället.

LYCKA TILL!

- - - - - - - - - - - - - - - - - - - -

Uppgift 1

En stokastisk variabel Z har sannolikhetstätheten $f_Z(z)$ given med

$$f_Z(z) = 2\phi(z)\Phi(\lambda z), \quad -\infty < z < \infty,$$

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

a) Visa att

$$P(| Z | \leq t) = \int_0^t 2\phi(z)dz.$$

(5 p)

b) Visa att $Z^2 \in \chi^2(1)$. *Ledning:* Bestäm tätheten för Z^2. (5 p)

Uppgift 2

Låt $N = \{N(t) \mid t \geq 0\}$ vara en poissonprocess med intensitet $\lambda = 2$. Vi bildar en ny heltalsvärd process genom

$$X(t) = \left\lfloor \frac{N(t)}{2} \right\rfloor, \quad t \geq 0,$$

där $\lfloor y \rfloor$ är heltalsdelen av det reella talet y, dvs. om k är ett heltal,

$$\lfloor y \rfloor = k, \quad \text{för } k \leq y < k+1.$$

t.e.x $\left\lfloor \frac{3}{2} \right\rfloor = 1$.
Bestäm sannolikheten

$$P\left(X(1) = 1, X(2) = 1\right).$$

(10 p)

Uppgift 3

$X \in \mathrm{Po}(\lambda)$ och den stokastiska variabeln Z är

$$Z = X | X > 0.$$

a) Beräkna den sannolikhetsgenererande funktionen $g_Z(t)$ till Z. (5 p)

b) Använd $g_Z(t)$ för att bestämma $E(Z)$. (5 p)

Uppgift 4

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

$$\mu = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ och } C = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix}.$$

Bestäm den betingade fördelningen för X_1 givet $X_1 - 4X_2 = 5$ samt beräkna

$$P(X_1 \leq 1 | X_1 - 4X_2 = 5).$$

(10 p)

Uppgift 5

$X_1, X_2, \ldots$ är oberoende och identiskt fördelade diskreta stokastiska variabler med sannolikhetsfunktionen $P(X_k = -1) = P(X_k = +1) = \frac{1}{2}$. Vi bildar

$$Y_n := \frac{1}{\sqrt{n \ln \ln n}} \sum_{k=1}^{n} X_k$$

Visa att

$$Y_n \xrightarrow{d} 0, \quad \text{då } n \to \infty.$$

(10 p)

Uppgift 6

Låt $X_1, X_2, \cdots, X_n$ vara oberoende likafördelade med $E(X_i) = 0$ och $V(X_i) = \sigma^2$ med karaktäristisk funktion $\varphi(t)$.
Låt vidare $\{N(t); t \geq 0\}$ vara en av X:n oberoende Poisson-process med intensitet λ med $\lambda > 0$. Definiera processen

$$S(t) = h \sum_{k=1}^{N(t)} X_k, \quad t \geq 0,$$

där $h \neq 0$.

a) Bestäm för varje $t > 0$ den karaktäristiska funktionen för $S(t)$ som funktion av λ, h och den karaktäristiska funktionen för X. (5 p)

b) Om h är litet men λ är stort så är $S(t)$="summan av många oberoende små hopp" och det verkar troligt att $S(t)$ är approximativt normalfördelad. Visa mer precist att om $\lambda \to \infty$ och $h \to 0$ så att $\lambda h^2 = 1$ så gäller för varje fixt t att $S(t)$ konvergerar i fördelning mot $N(0, \sigma^2 t)$. (5 p)

LÖSNING TENTAMEN I SF2941 SANNOLIKHETSTEORI OCH LINJÄRA MODELLER 08-18-19

Uppgift 1

a) Z är en kontinuerlig stokastisk variabel, vi har

$$P(|Z| \leq t) = P(-t \leq Z \leq t) = P(-t < Z \leq t) = \int_{-t}^{t} 2\phi(z)\Phi(\lambda z)\,dz$$

$$= \int_{-t}^{0} 2\phi(z)\Phi(\lambda z)\,dz + \int_{0}^{t} 2\phi(z)\Phi(\lambda z)\,dz.$$

Vi omskriver den första av integralerna i högra ledet med variabelbytet $z = -u$,

$$\int_{-t}^{0} 2\phi(z)\Phi(\lambda z)\,dz = -\int_{t}^{0} 2\phi(-u)\Phi(-\lambda u)\,du = \int_{0}^{t} 2\phi(-u)\Phi(-\lambda u)\,du = \int_{0}^{t} 2\phi(-z)\Phi(-\lambda z)\,dz.$$

Detta ger från ovan

$$P(|Z| \leq t) = 2\int_{0}^{t} \left(\phi(-z)\Phi(-\lambda z) + \phi(z)\Phi(\lambda z)\right)dz.$$

Men för den standardiserade normalfördelningen gäller $\phi(-z) = \phi(z)$ och $\Phi(-\lambda z) = 1 - \Phi(\lambda z)$ och detta ger

$$= 2\int_{0}^{t} \phi(z)\left(1 - \Phi(\lambda z) + \Phi(\lambda z)\right)dz = 2\int_{0}^{t} \phi(z)dz,$$

vilket är det påstådda resultatet.

b) Vi tar $x \geq 0$ och återkallar i minnet att $\sqrt{x^2} = |x|$.

$$P\left(Z^2 \leq x\right) = P\left(|Z| \leq \sqrt{x}\right) = P\left(-\sqrt{x} < Z \leq \sqrt{x}\right) = 2\int_{0}^{\sqrt{x}} \phi(z)dz,$$

där vi utnyttjade uppgiftens del a). Detta leverar tätheten $f_{Z^2}(x)$ för Z^2 som

$$f_{Z^2}(x) = \frac{d}{dx}P\left(Z^2 \leq x\right) = \frac{d}{dx}2\int_{0}^{\sqrt{x}} \phi(z)dz = 2 \cdot \phi(\sqrt{x})\frac{d}{dx}\sqrt{x}$$

$$= 2 \cdot \frac{e^{\frac{-x}{2}}}{\sqrt{2\pi}}\frac{1}{2x^{1/2}} = \frac{x^{-1/2}}{\sqrt{2\pi}}e^{\frac{-x}{2}}.$$

Detta uttryck återger ingenting annat än tätheten för $\chi^2(1)$. (För att jämföra med uttrycket för tätheten för $\chi^2(n)$ i Appendix 2 av A.Gut: An Intermediate Course in Probability, observera att $\Gamma(1/2) = \sqrt{\pi}$ enligt den utdelade formelsamligen.)

Uppgift 2

För att bestämma den sökta sannolikheten

$$P\left(X(1) = 1, X(2) = 1\right)$$

måste vi beskriva händelsen $\{X(1) = 1, X(2) = 1\}$ med hjälp av den bakomliggande poissonprocessen $N = \{N(t) \mid t \geq 0\}$.
Då inses p.g.a definitionen på $\lfloor \cdot \rfloor$ och att poissonprocessen har de icke-negativa heltalen som värdeförråd, att $X(1) = 1$, om $N(1) = 2$ eller $N(1) = 3$, och att $X(2) = 1$, om $N(2) = 2$ eller $N(2) = 3$. Eftersom poissonprocessen är icke-avtagande, innebär dessa samband att

$$\{X(1) = 1, X(2) = 1\} = A \cup B \cup C,$$

där

$$A = \{N(1) = 2, N(2) = 2\},$$
$$B = \{N(1) = 2, N(2) = 3\},$$
$$C = \{N(1) = 3, N(2) = 3\}.$$

Eftersom N är en poissonprocess med intensitet $\lambda = 2$ får vi att

$$P(A) = P\left(N(2) - N(1) = 0 \mid N(1) = 2\right) P\left(N(1) = 2\right) =$$
$$= e^{-2} \cdot \left(e^{-2}\frac{2^2}{2!}\right) = e^{-4}2,$$
$$P(B) = P\left(N(2) - N(1) = 1 \mid N(1) = 2\right) P\left(N(1) = 2\right)$$
$$= e^{-2} \cdot 2 \cdot \left(e^{-2}\frac{2^2}{2!}\right) = e^{-4}\frac{2^3}{2!}$$
$$P(C) = P\left(N(2) - N(1) = 0 \mid N(1) = 3\right) P\left(N(1) = 3\right)$$
$$= e^{-2} \cdot \left(e^{-2}\frac{2^3}{3!}\right) = e^{-4}\frac{2^3}{3!}$$

Eftersom händelserna A, B och C är disjunkta

$$P\left(X(1) = 1, X(2) = 1\right) = P(A) + P(B) + P(B)$$
$$= e^{-4}\left(2 + 2^3\left(\frac{1}{2!} + \frac{1}{3!}\right)\right) = e^{-4}\left(2 + 8\frac{4}{6}\right) =$$
$$= e^{-4} \cdot \frac{44}{6} = 0.1343.$$

SVAR: $\underline{P\left(X(1) = 1, X(2) = 1\right) \approx 0.13.}$

Uppgift 3

a) Om Z är X givet $X > 0$ ges dess sannolihetsfunktion av (see Formelsamlingen)

$$p_Z(Z = k) = \begin{cases} \frac{p_X(k)}{p_X(X>0)} & \text{if } k = 1, 2, \ldots \\ 0 & k = 0 \end{cases}$$

Om $X \in \text{Po}(\lambda)$, så är $p_X(X > 0) = 1 - e^{\lambda}$. Därmed gäller

$$p_Z(Z = k) = \begin{cases} \frac{e^{-\lambda}}{1-e^{\lambda}} \frac{\lambda^k}{k!} & \text{if } k = 1, 2, \ldots \\ 0 & k = 0. \end{cases}$$

Definitionen på sannolihetsgenererande funktion ger

$$g_Z(t) = E\left[t^Z\right] = \sum_{k=1}^{\infty} t^k p_Z(k) = \frac{e^{-\lambda}}{1-e^{-\lambda}} \sum_{k=1}^{\infty} t^k \frac{\lambda^k}{k!}$$

$$= \frac{e^{-\lambda}}{1-e^{-\lambda}} \left(\sum_{k=0}^{\infty} \frac{\left(t^k \lambda\right)^k}{k!} - 1 \right) = \frac{e^{-\lambda}}{1-e^{-\lambda}} \left(e^{t\lambda} - 1\right).$$

SVAR a): $\underline{g_Z(t) = \frac{e^{-\lambda}}{1-e^{-\lambda}} \left(e^{t\lambda} - 1\right).}$

b) Vi vet att sannolikhetsgenererande funktionens första derivata i punkten 1 är det sökta väntevärdet, dvs. $E(Z) = g_Z'(1)$.
Vi har

$$g_Z'(t) = \frac{d}{dt} \left(\frac{e^{-\lambda}}{1-e^{-\lambda}} \left(e^{t\lambda} - 1\right) \right) = \frac{e^{-\lambda}}{1-e^{-\lambda}} \lambda e^{t\lambda}.$$

Detta ger

$$g_Z'(1) = \frac{e^{-\lambda}}{1-e^{-\lambda}} \lambda e^{\lambda} = \frac{\lambda}{1-e^{-\lambda}}.$$

SVAR b): $\underline{E(Z) = \frac{\lambda}{1-e^{-\lambda}}.}$

Uppgift 4

Vi behöver fördelningen för

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} X_1 \\ X_1 - 4X_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & -4 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = B \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$$

som blir

$$N_2\left(B\mu, BCB^T\right) = N_2\left(\begin{pmatrix} 0 \\ -4 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ -1 & 13 \end{pmatrix} \right)$$

Fördelningen för Y_1 givet $Y_2 = 5$ ges enligt formelsamlingen av

$$N\left(\mu_1 + \frac{c_{12}}{c_{22}}(5 - \mu_2), c_{11} - \frac{c_{12}^2}{c_{22}} \right) = N\left(0 + \frac{-1}{13}(5 - (-4)), 1 - \frac{(-1)^2}{13} \right) =$$

$$= N\left(\frac{-9}{13}, \frac{12}{13} \right)$$

som ger att

$$P(X_1 \leq 1 | X_1 + X_2 = 5) = \Phi\left(\frac{1+\frac{9}{13}}{\sqrt{\frac{12}{13}}}\right) = \Phi\left(\frac{\frac{22}{13}}{\sqrt{\frac{12}{13}}}\right) \approx \Phi(1.76) \approx 0.96$$

SVAR b): $\underline{P(X_1 \leq 1 | X_1 + X_2 = 5) \approx 0.96.}$

Uppgift 5

En enkel strategi är att först visa konvergens i sannolikhet med tillhjälp av Tjebysjovs olikhet. För detta behöver vi väntevärde och varians för Y_n. Vi har att

$$E(X_k) = (-1)\frac{1}{2} + (+1)\frac{1}{2} = 0$$

och således är

$$\mathrm{Var}(X_k) = E\left(X_k^2\right) - (E(X_k))^2 = (-1)^2\frac{1}{2} + (+1)^2\frac{1}{2} = 1.$$

Detta ger

$$E(Y_n) = \frac{1}{\sqrt{n \ln\ln n}} \sum_{k=1}^{n} E(X_k) = 0$$

samt eftersom X_k:na är oberoende att

$$\mathrm{Var}(Y_n) = \frac{1}{n \ln\ln n} \sum_{k=1}^{n} \mathrm{Var}(X_k) = \frac{1}{n \ln\ln n} \sum_{k=1}^{n} 1 = \frac{1}{\ln\ln n}$$

Tjebysjovs olikhet säger att $P(|Y_n - E(Y_n)| > \epsilon) \leq \frac{1}{\epsilon^2}\mathrm{Var}(Y_n)$. Således har vi från ovan att

$$P(|Y_n| > \epsilon) \leq \frac{1}{\epsilon^2 \ln\ln n}$$

och detta visar att

$$P(|Y_n| > \epsilon) \to 0, \qquad \text{då } n \to \infty,$$

eller med andra ord

$$Y_n \xrightarrow{p} 0, \qquad \text{då } n \to \infty.$$

Eftersom konvergens i sannolikhet medför konvergens i fördelning, har vi därmed ådagalagt att

$$Y_n \xrightarrow{d} 0, \qquad \text{då } n \to \infty,$$

vilket ju skulle visas.

Uppgift 6

Vi får för $S(t)$ den karaktäristiska funktionen

$$\varphi_{S(t)}(u) = E(e^{iuS(t)}) = E(E(\exp(iuh \sum_{1}^{N(t)} X_k | N(t)))) =$$

$$= \sum_{n=1}^{\infty}(E(\exp(iuh\sum_{1}^{n} X_k)))e^{-\lambda t}\frac{(\lambda t)^n}{n!} = e^{-\lambda t}\sum_{n=0}^{\infty}(\varphi(uh))^n(\lambda t)^n/n! =$$

$$= \exp(\lambda t(\varphi(uh) - 1)).$$

b) Vi har

$$\varphi(s) = 1 - \frac{s^2\sigma^2}{2} + o(s^2)$$

då $s \to 0$. Med $h^2 = 1\lambda$ får vi

$$\varphi(uh) = \varphi(u/\sqrt{\lambda}) = 1 - \frac{u^2\sigma^2}{2\lambda} + o(1/\lambda).$$

då $\lambda \to \infty$. Detta ger

$$\varphi_{S(t)}(u) = \exp(\lambda t(1 - \frac{u^2\sigma^2}{2\lambda} + o(1/\lambda) - 1)) =$$

$$= \exp(-\frac{u^2\sigma^2 t}{2} + \lambda t o(1/\lambda)) \to \exp(-\frac{u^2\sigma^2 t}{2})$$

då $\lambda \to \infty$ och detta är ju karaktäristisk funktion för $N(0, \sigma^2 t)$ och kontinuitetssatsen ger det önskade resultatet.