Avd. Matematisk statistik

KTH Matematik

TENTAMEN I 5B1504 MATEMATISK STATISTIK GRUNDKURS FÖR E (gamlingar)
FREDAGEN DEN 16 APRIL 2004 KL 8.00–13.00

Examinator: Gunnar Englund, 790 7416

Tillåtna hjälpmedel: Formel- och tabellsamling i matematisk statistik för E. Beta Mathematics Handbook. Kalkylator.

Införda beteckningar skall förklaras och definieras. Resonemang och uträkningar skall vara så utförliga och väl motiverade att de är lätta att följa. Numeriska svar skall anges med minst två siffrors noggrannhet.

Varje korrekt lösning ger 10 poäng. Gränsen för godkänt är preliminärt 20 poäng.

Resultatet anslås senast fredagen den 7 maj 2004 på Matematisk statistiks anslagstavla i entréplanet, Lindstedtsvägen 25, rakt fram innanför porten och tentamen kommer att finnas tillgänglig på elevexpeditionen under sju veckor efter tentamensdagen.

Uppgift 1

a)
Låt den stokastiska variabeln X vara $R(-3, 4)$. Sätt $Y = 3X^2$. Beräkna $P(Y > 12)$. (5 p)

b) Antalet trafikolyckor per år på en viss väg kan anses vara oberoende och Poisson(4)-fördelade. Beräkna sannolikheten att det inträffar fler än 8 olyckor något (dvs minst ett) av de kommande 5 åren. (5 p)

Uppgift 2

Begreppet intelligens är operationellt definierat som resultatet av ett intelligenstest. För enkelhets skull antar vi att resultaten är oberoende och normalfördelade med samma standardavvikelse för alla individer i alla test. Man önskar undersöka om vana vid intelligenstest medför en förändring av testpoängen. För att pröva detta dras ett stickprov om 7 individer, vilka tidigare ej genomgått test. Dessa får genomgå prov på 6 test vardera, och nedan anges resultat på första och sista testet:

Individ nr	1	2	3	4	5	6	7
Första testet	108	142	121	138	112	102	91
Sista testet	106	146	122	140	117	104	94

Pröva hypotesen att ingen förändring i testpoäng sker mot hypotesen att en sådan förändring sker. 5% signifikansnivå. (10 p)

Uppgift 3

Erfarenhet med plaster visar att det existerar en relation mellan hårdhet (mätt i Brinell-enheter) och den tid (timmar) som förflutit sedan gjutningsprocessen. Resultat:

Mätning nr	1	2	3	4	5	6	7	8
x (Tid)	16	16	24	24	32	32	40	40
y (Hårdhet)	199	205	218	220	237	234	250	248

Hjälpsummor:
$\sum_i x_i = 224$, $\sum_i x_i^2 = 6912$, $\sum_i y_i = 1811$, $\sum_i y_i^2 = 412479$ och $\sum_i x_i y_i = 51968$.
Antag att en enkel regressionsmodell gäller.

a) Beräkna MK-skattningarna av α' och β i regressionslinjen $y = \alpha' + \beta x$. (4 p)

b) Bestäm ett 95 % konfidensintervall för förväntade hårdheten 40 timmar efter gjutning. (3 p)

c) Bestäm ett 95 % konfidensintervall för den förändring i förväntad hårdhet som sker under 1 timme. (3 p)

Uppgift 4

$(X(t);\, t \in R)$ är en stationär normalprocess med $m_X = 2$ och kovariansfunktion

$$r_X(\tau) = e^{-3|\tau|},\ \tau \in R.$$

Vi bildar processen $(Y(t);\, t \in R)$ där

$$Y(t) = \int_0^\infty e^{-3u} X(t-u)\,du.$$

a) Bestäm Y-processens kovariansfunktion och väntevärde. (5 p)

b) Beräkna $P(X(t) + Y(t) > 2)$. (5 p)

Uppgift 5

En Poissonprocess $(X(t);\, t \geq 0)$ har $E(X(t)) = 3t$. Låt T vara tiden då processen för första gången antar värdet 2. Bestäm $P\left(T \leq t | X(2) = 2\right)$ som funktion av t. (10 p)

KTH Matematik

Avd. Matematisk statistik

LÖSNINGSFÖRSLAG TENTAMEN I 5B1504 FREDAGEN DEN 16 APRIL 2004 KL 8.00–13.00

Uppgift 1

a) Eftersom $X \in R(-3, 4)$ så gäller (se formelsamlingen avsn 3)

$$f_X(x) = \begin{cases} \frac{1}{7} & -3 \leq x \leq 4, \\ 0 & \text{annars.} \end{cases}$$

$$\begin{aligned} P(Y > 12) &= P(3X^2 > 12) = P(X^2 > 4) = 1 - P(X^2 \leq 4) \\ &= 1 - P(-2 \leq X \leq 2) = 1 - \int_{-2}^{2} f_X(x)\,dx = 1 - \int_{-2}^{2} \frac{1}{7}\,dx = \frac{3}{7}. \end{aligned}$$

b) Låt $X_i =$ antal olyckor år i, $i = 1, 2, 3, 4, 5$.

$P(\text{fler än 8 olyckor minst ett år}) = 1 - P(\text{högst 8 olyckor varje år})$

$$= 1 - P(X_1 \leq 8 \cap X_2 \leq 8 \cap X_3 \leq 8 \cap X_4 \leq 8 \cap X_5 \leq 8) = \left\{ X_i\text{:na oberoende, likafördelade} \right\}$$

$$= 1 - P(X_1 \leq 8)^5 = \left\{ \text{formelsaml tab 7} \right\} = 1 - 0.97864^5 = 0.1023.$$

Uppgift 2

Parvisa jämförelser. Bilda differenser inom individer (sista–första).
Nya data. -2, 4, 1, 2, 5, 3
som är utfall av oberoende N(Δ,σ). $\widehat{\Delta} = \frac{1}{7}(-2 + 4 + 1 + 2 + 5 + 2 + 3) = 15/7 \approx 2.143$, $s^2 = \frac{1}{7-1}((-2)^2 + \cdots + 3^2 - 15^2/7) \approx 5.14$ som ger $s \approx 2.27$.

Hypotesen $\mathrm{H}_0 : \ \Delta = 0$ testas till exempel med konfidensmetoden. Ett 95 % konfidensintervall för Δ blir $\widehat{\Delta} \pm t_{0.025}(7-1) \cdot \frac{s}{\sqrt{7}} = 2.143 \pm 2.45 \cdot \frac{2.27}{\sqrt{7}} = 2.143 \pm 2.10$. Eftersom 0 ej tillhör intervallet. Alltså kan H_0 förkastas på nivån 5 %.

Uppgift 3

a) MK-skattningarna ges av $\beta^* = \sum_i (x_i - \bar{x})y_i / \sum_i (x_i - \bar{x})^2 = 1.96875$ och $\alpha^{'*} = \bar{y} - \beta^* \bar{x} = 171.25$. Vi har utnyttjat att $\sum_i (x_i - \bar{x})^2 = \sum_i x_i^2 - n\bar{x}^2$ och $\sum_i (x_i - \bar{x})y_i = \sum_i x_i y_i - n\bar{x}\bar{y}$.

b) Ett 95 % konfidensintervall ges av

$$I_{m_0} = \alpha^{'*} + \beta^* \cdot 40 \pm t_{0.025}(8-2)s\sqrt{\frac{1}{8} + \frac{(40 - \bar{x})^2}{\sum_i (x_i - \bar{x})^2}}$$

Vi har att $s^2 = \frac{1}{n-2}(\sum_i(y_i-\bar{y})^2 - \beta^{*2}\sum_i(x_i-\bar{x})^2 = 5.54$. Vi får ur tabell 3 att $t_{0.025}(6) = 2.45$ varför man erhåller intervallet $\underline{250 \pm 3.41}$

c) Konfidensintervall för β söks. Det ges av intervallet $I_\beta = \beta^* \pm t_{0.025}(8-2)s/\sqrt{\sum_i(x_i-\bar{x})^2}$ varför man erhåller intervallet $\underline{1.97 \pm 0.23}$

Uppgift 4

Vi ser att

$$E(Y(t)) = E\left(\int_0^\infty e^{-3u}X(t-u)du\right) = \int_0^\infty e^{-3u}E(X(t-u))\,du = \int_0^\infty e^{-3u}2du = \frac{2}{3}.$$

Eftersom $(X(t);\, t \in R)$ är en normalprocess är också $(Y(t);\, t \in R)$ en normalprocess.

Y-processens kovariansfunktion erhålls ut Y:s spektraltäthet. Integralen motsvarar en linjär filtrering med viktsfunktionen

$$h(u) = \begin{cases} e^{-3u} & \text{för } u \geq 0 \\ 0 & \text{för } u < 0 \end{cases}$$

som enligt Formelsamlingens transformpar för Laplacetransformer ger att $H(\omega) = 1/(3+i\omega)$ vilket ger

$$S_Y(\omega) = |H(\omega)|^2 S_X(\omega) = \frac{1}{3^2+\omega^2} \cdot \frac{2 \cdot 3}{(3^2+\omega^2)}.$$

Således har vi $S_Y(\omega) = 6/(9+\omega^2)^2$ vilket ger att

$$\underline{r_Y(\tau) = e^{-3|\tau|}\left(|\tau| + 1/3\right)/6}.$$

b) Eftersom Y-processen erhålls genom linjär filtrering av normalprocessen X är $X(t)+Y(t)$ normalfördelad. Dess väntevärde blir $E(X(t)+Y(t)) = E(X(t)) + E(Y(t)) = 2+2/3 = 8/3$ och variansen blir

$$V(X(t)+Y(t)) = V(X(t)) + V(Y(t)) + 2C(X(t),Y(t)) = r_X(0) + r_Y(0) + 2C(X(t),Y(t)).$$

Vi får

$$C(X(t),Y(t)) = C\left(X(t), \int_0^\infty e^{-3u}X(t-u)du\right) = \int_0^\infty e^{-3u}C(X(t),X(t-u))\,du =$$

$$= \int_0^\infty e^{-3u}e^{-3|u|}du = \int_0^\infty e^{-6u}du = \frac{1}{6}$$

Således har vi

$$V(X(t)+Y(t)) = r_X(0)+r_Y(0)+2C(X(t),Y(t)) = 1 + e^{-3\cdot 0}\left(|0|+1/3\right)/6 + 2\cdot 1/6 = 25/18.$$

Vi får då med $Z(t) = X(t)+Y(t)$ att $Z(t)$ är $N(8/3, \sqrt{25/18})$ som ger

$$P(Z(t) > 2) = P\left(\frac{Z(t)-8/3}{\sqrt{25/18}} > \frac{2-8/3}{\sqrt{25/18}}\right) = 1 - \Phi\left(\frac{2-8/3}{\sqrt{25/18}}\right) \approx 1-\Phi(-0.57) = \Phi(0.57) \approx \underline{0.72}.$$

Kommentar: Man kan också studera $Z(t) = X(t) + Y(t)$ betraktad som en linjär filtrering av X-processen där viktsfunktionen är

$$h(u) = \begin{cases} \delta(u) + e^{-3u} & \text{för } u \geq 0 \\ \delta(u) & \text{för } u < 0 \end{cases}$$

där $\delta(u)$ är Diracs delta-funktion. Ytterligare ett alternativ är att studera vad filtret gör med testfunktionen $e^{i\omega t}$.

Uppgift 5

Vi gör först observationen att $\{T \leq t\} = \{X(t) \geq 2\}$.

Med $t < 2$ får vi

$$P\left(T \leq t | X(2) = 2\right) = \frac{P\left(X(t) \geq 2 \cap X(2) = 2\right)}{P\left(X(2) = 2\right)} = \frac{P\left(X(t) = 2 \cap X(2) - X(t) = 0\right)}{P\left(X(2) = 2\right)} =$$

$$\frac{P\left(X(t) = 2\right) P\left(X(2) - X(t) = 0\right)}{P\left(X(2) = 2\right)} = \frac{\frac{(3t)^2}{2!} e^{-3t} \frac{(3(2-t))^0}{0!} e^{-3(2-t)}}{\frac{(3 \cdot 2)^2}{2!} e^{-3 \cdot 2}} = \frac{t^2}{4}$$

Om $t \geq 2$ är $P(T \leq t | X(2) = 2) = 1$ dvs

$$P(T \leq t | X(2) = 2) = \begin{cases} 0 & \text{för } t \leq 0, \\ t^2/4 & \text{för } 0 \leq t \leq 2, \\ 1 & \text{för } t \geq 2 \end{cases}$$