Avd. Matematisk statistik

KTH Matematik

TENTAMEN I 5B1504 MATEMATISK STATISTIK GRUNDKURS FÖR E (gamlingar)
LÖRDAGEN DEN 27 AUGUSTI 2005 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 måndagen den 19 september 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

Nollor och ettor sänds i en brusig miljö. Sannolikheten att en nolla resp. etta sänds är 0.4 resp. 0.6. Den mottagna signalen kan uppfattas som en stokastisk variabel X som är $N(0, 1)$ resp. $N(1, 1)$ om noll resp. ett sändes. Mottagaren använder regeln: om $X < 0.20$ anses en nolla ha sänts, annars har en etta sänts. Beräkna sannolikheten att en mottagen nolla verkligen har sänts som en nolla. (10 p)

Uppgift 2

För att undersöka effekten av ett rostskyddsmedel behandlade man 10 järnstavar med detta. På var och en av 10 olika platser grävdes därefter en av de behandlade stavarna jämte en obehandlad stav ner. Efter 3 månader togs alla stavar upp och rostgraden mättes. Resultat (i lämplig enhet):

Plats	1	2	3	4	5	6	7	8	9	10
Obehandlade	32.3	38.0	40.1	28.4	35.9	36.3	25.1	28.2	39.8	32.6
Behandlade	31.5	37.5	40.2	28.0	34.8	36.0	25.1	27.5	39.1	32.4

Resultatet från plats nr k för obehandlad resp. behandlad stav är observationer på $N(m_k, \sigma_{\text{Obeh.}})$- resp. $N(m_k + \Delta, \sigma_{\text{Beh.}})$-fördelade stokastiska variabler. Alla underliggande stokastiska variabler förutsätts oberoende. Δ (eller egentligen $-\Delta$) är ett mått på rostskyddsmedlets effekt.

a) Beräkna ett 95 % konfidensintervall för Δ. (7 p)

b) Anser Du att rostskyddsmedlet har effekt? Motivera utgående från Ditt resultat på a)-delen! (3 p)

Uppgift 3

Enligt en nyligen publicerad partisympatiundersökning angav 35.3 % av de tillfrågade att de sympatiserade med (s). Urvalet av tillfrågade kan ses som ett slumpmässigt stickprov av 1427 personer från populationen av alla röstberättigade.

a) Ge ett approximativt 95 % konfidensintervall av andelen (s)-sympatisörer av alla röstberättigade. (5 p)

b) Vid en tidigare undersökning hade 1532 personer valts ut och av dessa angav 37.5 % att de sympatiserade med (s). Kan man påstå med någon säkerhet att andelen (s)-sympatisörer i väljarkåren förändrats? Gör ett lämpligt approximativt 95 % konfidensintervall för förändringen och besvara frågan med hjälp av detta. (5 p)

Uppgift 4

$\{X(t); -\infty < t < \infty\}$ är en stationär normalprocess med väntevärdet 0 och kovariansfunktionen

$$r_X(\tau) = \begin{cases} 1 - |\tau|, & |\tau| < 1, \\ 0, & \text{annars.} \end{cases}$$

Processen filtreras genom ett filter givet av differentialekvationen

$$Y'(t) + Y(t) = X(t).$$

Beräkna

a) $P(X(t) > 1)$, (2 p)

b) $P(Y(t) > 1)$. (8 p)

Anm: Även delresultat kan ge poäng.

Uppgift 5

Punkter är slumpmässigt utspridda på tallinjen. Antalet punkter i disjunkta intervall är oberoende och antalet punkter i ett intervall är Poissonfördelat med väntevärdet givet av intervallängden. Låt $X(t)$ beteckna antalet punkter i $(t, t+1)$. Beräkna

a) $C(X(t), X(t+0.5))$, (5 p)

b) $C(X(t), X(t+s))$ för $-\infty < s < \infty$. (5 p)

Avd. Matematisk statistik

KTH Matematik

LÖSNINGSFÖRSLAG TENTAMEN I 5B1504 LÖRDAGEN DEN 27 AUGUSTI 2005 KL 8.00–13.00

Uppgift 1

Inför S_0 = "nolla sänds", S_1 = "etta sänds", M_0 = "nolla mottages". Med Bayes formel erhålls:

$$P(S_0 \mid M_0) = \frac{P(S_0)P(M_0 \mid S_0)}{P(S_0)P(M_0 \mid S_0) + P(S_1)P(M_0 \mid S_1)}$$

$$= \frac{0.4P(X < 0.2 \parallel X \in N(0,1))}{0.4P(X < 0.2 \parallel X \in N(0,1)) + 0.6P(X < 0.2 \parallel X \in N(1,1))}$$

$$= \frac{1}{1 + 1.5\Phi(-0.8)/\Phi(0.2)} = 0.65,$$

där $P(X < 0.2 \parallel X \in N(m,1))$ betyder sannolikheten för att $X < 0.2$ då X är $N(m,1)$-fördelad.

Uppgift 2

a) Låt resultatet från plats nr k för obehandlad resp. behandlad stav vara utfall av X_k resp. Y_k. Detta innebär att $Z_k = Y_k - X_k$ är en $N(\Delta, \sigma)$-fördelad stokastisk variabel. Vi har följande observationer:

Plats	1	2	3	4	5	6	7	8	9	10
y_k	31.5	37.5	40.2	28.0	34.8	36.0	25.1	27.5	39.1	32.4
x_k	32.3	38.0	40.1	28.4	35.9	36.3	25.1	28.2	39.8	32.6
$z_k = y_k - x_k$	− 0.8	− 0.5	0.1	− 0.4	− 1.1	− 0.3	0.0	− 0.7	− 0.7	− 0.2

Eftersom σ är okänt så är

$$I_\Delta = \bar{z} \pm t_{\alpha/2}(n-1)\frac{s_z}{\sqrt{n}}$$

ett konfidensintervall med konfidensgrad $1 - \alpha$. I vårt fall gäller

$$\bar{z} = -0.46 \qquad s_z = 0.375 \qquad n = 10 \qquad t_{0.025}(9) = 2.26$$

vilket ger $I_\Delta = -0.46 \pm 2.26 \cdot \dfrac{0.375}{\sqrt{10}} = \underline{-0.46 \pm 0.27 = (-0.73, -0.19)}$.

b) Eftersom hela intervallet ligger utanför 0 så har rostskyddsmedlet (statistiskt signifikant) effekt.

Uppgift 3

a) Antalet (s)-sympatisörer i urvalet kan ses som ett utfall av en $\mathrm{Bin}(n_1, p_1)$-fördelad stokastisk variabel, där $n_1 = 1427$ och $(p_1)^*_{\text{obs}} = 0.353$. Eftersom $n_1(p_1)^*_{\text{obs}}(1 - (p_1)^*_{\text{obs}}) = 325.9$, vilket är mycket större än 10 så kan vi normalapproximera. Det ger följande approximativa konfidensintervall:

$$I_{p_1} = (p_1)^*_{\text{obs}} \pm \lambda_{0.025}\sqrt{\frac{(p_1)^*_{\text{obs}}(1-(p_1)^*_{\text{obs}})}{n_1}} = 0.353 \pm 1.96\sqrt{\frac{0.353 \cdot 0.647}{1427}}$$

$$= 0.353 \pm 0.025 = \underline{(32.8\%, 37.8\%)}.$$

b) I den tidigare undersökningen hade vi $n_2 = 1532$ och $(p_2)^*_{\text{obs}} = 0.375$.
Eftersom $n_2(p_2)^*_{\text{obs}}(1-(p_2)^*_{\text{obs}}) = 359.1$ så kan vi även här normalapproximera. Vi är nu intresserade av följande approximativa konfidensintervall för $p_2 - p_1$:

$$I_{p_2-p_1} = (p_2)^*_{\text{obs}} - (p_1)^*_{\text{obs}} \pm \lambda_{0.025}\sqrt{\frac{(p_2)^*_{\text{obs}}(1-(p_2)^*_{\text{obs}})}{n_2} + \frac{(p_1)^*_{\text{obs}}(1-(p_1)^*_{\text{obs}})}{n_1}}$$

$$= 0.375 - 0.353 \pm 1.96\sqrt{\frac{0.375 \cdot 0.625}{1532} + \frac{0.353 \cdot 0.647}{1427}} = 0.022 \pm 0.035 = \underline{(-1.3\%, 5.7\%)}.$$

Eftersom $I_{p_2-p_1} \ni 0$ kan man inte påstå att andelen (s)-sympatisörer förändrats.

Uppgift 4

a) $X(t)$ är normalfördelad med väntevärde 0 och varians $r_X(0) = 1$. Alltså

$$P(X(t) > 1) = 1 - \Phi(1) = \underline{0.16}.$$

b) Insignalen $e^{i\omega t}$ till ett linjärt filter ger utsignalen $H(\omega)e^{i\omega t}$. För det givna filtret ger detta

$$H(\omega)i\omega e^{i\omega t} + H(\omega)e^{i\omega t} = e^{i\omega t},$$

och därmed $H(\omega) = 1/(1+i\omega)$. Spektraltätheten för insignalen är

$$S_X(\omega) = \int_{-1}^{1}(1-|\tau|)e^{-i\omega\tau}d\tau,$$

och spektraltätheten för utsignalen blir

$$S_Y(\omega) = |1+i\omega|^{-2}S_X(\omega).$$

Eftersom insignalen har väntevärdet 0 och är en normalprocess, så gäller detta också utsignalen. Variansen för utsignalen blir

$$r_Y(0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_Y(\omega)d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{1}{1+\omega^2}\int_{-1}^{1}(1-|\tau|)e^{-i\omega\tau}d\tau d\omega$$

$$= \int_{-1}^{1}(1-|\tau|)\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-i\omega\tau}}{1+\omega^2}d\omega d\tau = \int_{-1}^{1}(1-|\tau|)\frac{e^{-|\tau|}}{2}d\tau = \int_{0}^{1}(1-\tau)e^{-\tau}d\tau = e^{-1}.$$

Härav följer att

$$P(Y(t) > 1) = 1 - \Phi(\sqrt{e}) = \underline{0.05}.$$

Uppgift 5

a) Låt S, T, U vara antalet punkter i intervallen $(t, t+0.5)$, $(t+0.5, t+1)$, $(t+1, t+1.5)$. S, T, U är oberoende och $Po(0.5)$. Härav följer

$$C(X(t), X(t+0.5)) = C(S+T, T+U) = C(T,T) = V(T) = 0.5.$$

b) Om $s < -1$ eller $s > 1$ blir intervallen $(t, t+1), (t+s, t+s+1)$ disjunkta och därmed ger oberoende att $C(X(t), X(t+s)) = 0$. För $0 \leq s \leq 1$ fås som i a) med S, T, U $Po(s), Po(1-s), Po(s)$ resp. att $C(X(t), X(t+s)) = V(T) = 1-s$. Analogt fås för $-1 \leq s \leq 0$ att $C(X(t), X(t+s)) = 1+s$.