

TENTAMEN I 5B1504 MATEMATISK STATISTIK GRUNDKURS FÖR E3
LÖRDAGEN DEN 30 AUGUSTI 2003 KL 08.00–13.00.

Examinator: Gunnar Englund, tel. 790 7416.

Tillåtna hjälpmedel: Formel- och tabellsamling i Matematisk statistik. Beta Mathematics Handbook. Räknare.

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. Tentamen består av 6 uppgifter. Varje korrekt lösning ger 10 poäng. Gränsen för godkänt är preliminärt 24 poäng.

Resultatet anslås senast måndagen den 22 september 2003 på Matematisk statistiks anslags-tavla i entréplanet, Lindstedtsvägen 25, rakt fram innanför porten.

Tentamen kommer att finnas tillgänglig på elevexpeditionen sju veckor efter skrivnings-tillfället.

Uppgift 1

I faderskapsmål undersöks ofta blodgrupper hos moder, barn och den utpekade fadern. För enkelhets skull betraktar vi bara det s.k. $AB0$ -systemet (i verkligheten undersöks ett flertal sådana blodgruppssystem). För barn vars moder har blodgrupp 0 beror den av faderns blodgrupp betingade sannolikheten att barnet skall få blodgrupp $A0$:

Faderns blodgrupp	Sannolikheten att barnet skall få blodgrupp $A0$
$A0$	$1/2$
AA	1
AB	$1/2$
Övriga	0

Sannolikheten att en på måfå vald man skall ha blodgrupp $A0$, AA , AB respektive ”övriga” är 0.36, 0.08, 0.02 respektive 0.54. Beräkna sannolikheten att fadern har AA när barnet har $A0$ och modern 0. (10 p)

Uppgift 2

Fästingvaccin förpackas i lådor om tre provrör i varje. Varje provrör kontrolleras före nedpackning och sannolikheten att vaccinet i provröret uppfyller vissa krav är 0.90 oberoende av tidigare undersökta provrör. Låt X beskriva antalet provrör som måste undersökas innan en låda blir fylld. Bestäm $E[X]$ och $V(X)$ (5 p)

samt

sannolikheten att det krävs fler än 6 undersökta provrör innan lådan är full. (5 p)

Uppgift 3

I en veckotidning beskrivs en bantningskur som påstås ge en viktninskning av ungefär 7 kg på 14 dagar. En grupp på åtta personer bestämmer sig för att pröva kuren. Deras vikter (i kg) före och efter genomgången kur ges nedan:

Före	108	88	82	103	98	100	90	85
Efter	110	85	78	101	91	99	90	82

Vikterna kan anses vara observationer av oberoende normalfördelade stokastiska variabler. Det är dock inte rimligt att anta att samtliga personer har samma förväntad vikt före kuren, eller att de har samma förväntad vikt efter kuren.

- Beräkna ett konfidensintervall med konfidensgrad 95% för den förväntade viktändringen efter genomgången kur. (7 p)
- Testa på signifikansnivån 5% hypotesen att den förväntade viktninskningen är 7 kg. Slutsatsen av testet skall klart framgå. (3 p)

Uppgift 4

Den 14 september 2003 genomförs i Sverige en folkomröstning om EMU, där vi får ta ställning till frågan "anser du att Sverige ska införa euron som valuta". Följande är fritt tolkat från 1000 svar ur TEMO:s undersökning 16–18/6, 2003,

		Röstbenägenhet	
		Hög	Låg
Rösta ja	287	63	350
Vet ej/tveksam	93	48	141
Rösta nej	395	114	509
	775	225	1000

Undersök med ett lämpligt test ifall det på nivå 5% föreligger någon skillnad i röstbenägenheten baserat på hur personerna ställer sig till frågan om ett införande av euron som valuta. (10 p)

Uppgift 5

$\{N(t), t \in R\}$ är Gaussiskt vitt brus med väntevärdet 0 och med den konstanta spektraltätheten $1/2$. Sätt

$$X(t) = \int_{-0.5}^{0.5} N(t+u)du.$$

Bestäm för den stationära stokastiska processen $\{X(t), t \in R\}$

- spektraltäthet, (5 p)
- kovariansfunktion, (3 p)
- $P(X(0) > 1)$. (2 p)

Anm.: Man kan eventuellt ha nytta av sambandet

$$\max(0, 1 - |x|) = \int_{-\infty}^{\infty} e^{itx} \frac{1 - \cos t}{\pi t^2} dt, \quad x \in R$$

Uppgift 6

En stationär normalprocess $\{X(t), t \in R\}$ har väntevärdet 10 och kovariansfunktionen $e^{-|\tau|}$. Processen har observerats vid tiden 0 med $X(0) = 12$. Givet detta, beräkna sannolikheten att $X(1) > 11$. (10 p)

LÖSNING TENTAMEN I 5B1504 MATEMATISK STATISTIK GRUNDKURS FÖR E
03-08-30

Uppgift 1

Vi låter

Ω = modern har blodgrupp 0,
 H_1 = fadern har A0,
 H_2 = fadern har AA,
 H_3 = fadern har AB,
 H_4 = fadern har 'övriga' samt
 C = barnet har A0.

Vi söker $P(H_2 | C)$. Enligt texten är

$$P(H_1) = 0.36, \quad P(H_2) = 0.08, \quad P(H_3) = 0.02, \quad P(H_4) = 0.54,$$

och

$$P(C | H_1) = 1/2, \quad P(C | H_2) = 1, \quad P(C | H_3) = 1/2 \text{ och } P(C | H_4) = 0.$$

Bayes' sats ger då

$$\begin{aligned} P(H_2 | C) &= \frac{P(C | H_2)P(H_2)}{P(C | H_1)P(H_1) + P(C | H_2)P(H_2) + P(C | H_3)P(H_3) + P(C | H_4)P(H_4)} \\ &= \frac{1 \cdot 0.08}{0.5 \cdot 0.36 + 1 \cdot 0.08 + 0.5 \cdot 0.02 + 0 \cdot 0.54} = \frac{8}{27} \approx \underline{0.296}. \end{aligned}$$

Uppgift 2

Låt X_1 , X_2 och X_3 vara de stokastiska variabler som beskriver antalet provrör som måste undersökas tills ett provrör som uppfyller kraven hittas. Ur uppgiften fås att X_i , $i = 1, 2, 3$, är oberoende och ffg(p)-fördelade med $p = 0.90$, och

$$X = X_1 + X_2 + X_3.$$

Nu är

$$\mathbf{E}[X] = \mathbf{E}[X_1 + X_2 + X_3] = \mathbf{E}[X_1] + \mathbf{E}[X_2] + \mathbf{E}[X_3] = \frac{1}{p} + \frac{1}{p} + \frac{1}{p} = \frac{3}{p} = \frac{10}{3}$$

och, utnyttjandes oberoendet,

$$\mathbf{V}(X) = \mathbf{V}(X_1 + X_2 + X_3) = \mathbf{V}(X_1) + \mathbf{V}(X_2) + \mathbf{V}(X_3) = \frac{3(1-p)}{p^2} = \frac{10}{27}.$$

Med Y som antalet provrör som uppfyller kraven bland 6 undersökta, är $Y \sim \text{Bin}(n, p)$ med $n = 6$, $p = 0.90$. Händelsen $X > 6$ är samma som $Y \leq 2$ (tänk!) så

$$P(X > 6) = P(Y \leq 2) = \sum_{k=0}^2 \binom{n}{k} p^k (1-p)^{n-k} = 0.00127.$$

Uppgift 3

a) Beteckna person nr. i 's vikt före resp. efter genomgången kur med x_i resp. y_i , för $i = 1, \dots, 8$. Vikterna är observationer av stokastiska variabler $X_1, \dots, X_8$ resp. $Y_1, \dots, Y_8$, och lämpliga modellantaganden är att $X_i \sim N(m_i, \sigma_1)$ och $Y_i \sim N(m_i - \Delta, \sigma_2)$, för $i = 1, \dots, 8$, där Δ är den förväntade viktnedgången, och alla parametrar är okända (standardmodellen "stickprov i par"). Bilda de parvisa differenserna $Z_i = X_i - Y_i$, och motsvarande observerade värden $z_i = x_i - y_i$, för $i = 1, \dots, 8$. Då är $Z_i \sim N(\Delta, \sigma_z)$, där $\sigma_z = \sqrt{\sigma_1^2 + \sigma_2^2}$. De observerade värdena $z_1, \dots, z_8$ är

$$-2 \quad 3 \quad 4 \quad 2 \quad 7 \quad 1 \quad 0 \quad 3$$

Eftersom σ_z är okänt så är

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

ett konfidensintervall för Δ av grad $1 - \alpha$. I vårt fall gäller

$$\bar{z} = 2.25 \quad s_z = 2.7124 \quad n = 8 \quad t_{0.025}(7) = 2.36$$

vilket ger

$$I_\Delta = 2.25 \pm 2.36 \cdot \frac{2.7124}{\sqrt{8}} = \underline{2.25 \pm 2.26 = (-0.01, 4.51)}.$$

b) Eftersom $7 \notin I_\Delta$, så måste denna hypotes avvisas (på nivån 5%).

Uppgift 4

Bilda χ^2 -teststorheten för homogenitetstest

$$Q = \sum_{i,j} \frac{(X_{ij} - n_i p_j^*)^2}{n_i p_j^*},$$

där n_i är antalet observationer i serie i och X_{ij} beskriver antalet utfall i serie i som hamnar i kategori j . Under en hypotes H_0 om att de relativa storlekarna är samma i de tre mätserierna är Q approximativt $\chi^2(2)$ -fördelad. Hypotesen förkastas för stora värden på Q . Ur tabell erhåller vi att $\chi^2_{1-0.05}(2) = 5.99$. Vi observerar utfallet

$$q = \sum_{i,j} \frac{(x_{ij} - n_i p_j^*)^2}{n_i p_j^*} = 14.8 > \chi^2_{0.95}$$

och förkastar H_0 på signifikansnivån 5%. Det är säkerställt att röstbenägenheten skiljer sig åt mellan de olika mätserierna.

Uppgift 5

a) Frekvenssvarsfunktionen till filtret blir

$$H(\omega) = \int_{-0.5}^{0.5} e^{i\omega u} du = \frac{e^{i\omega 0.5} - e^{-i\omega 0.5}}{i\omega}$$

och filtersatsen ger spektraltätheten

$$S_X(\omega) = |(e^{i\omega 0.5} - e^{-i\omega 0.5})/i\omega|^2/2 = (1 - \cos \omega)/\omega^2.$$

b) Anmärkningen och sambandet mellan kovariansfunktion och spektraltäthet ger

$$r_X(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega\tau} S_X(\omega) d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega\tau} (1 - \cos \omega)/\omega^2 d\omega = \frac{1}{2} \max(0, 1 - |\tau|).$$

c) Eftersom inprocessen är Gaussisk och filtret är linjärt blir X -processen Gaussisk. Vi har $E(X(t)) = H(0)0 = 0$ och ur b) $V(X(t)) = 1/2$. Detta ger

$$P(X(0) > 1) = 1 - \Phi(\sqrt{2}) = 0.08.$$

Uppgift 6

Eftersom $\{X(t), t \in R\}$ är en normalprocess med väntevärdet 10 och kovariansfunktionen $e^{-|\tau|}$ så följer ur formlerna för betingad normalfördelning, att den betingade fördelningen för $X(1)$ givet $X(0) = 12$ är $N(10 + e^{-1}(12 - 10), \sqrt{1 - e^{-2}})$. Detta ger svaret $P(X(1) > 11 | X(0) = 12) = 1 - \Phi((11 - 10 - e^{-1}2)/\sqrt{1 - e^{-2}}) = 1 - \Phi(0.284) = 0.39$.