

ESS011: Matematisk statistik och signalbehandling – Tid: 14.00-18.00, Datum: 2009-08-25

Examinator: Ottmar Cronie (Serik Sagitov)

Jour: Ottmar Cronie, tel. 031-772 35 44

Hjälpmedel: Formelblad, Formelsamling "Beta", Räknedosa.

Betygsgränser: För betyg 3 krävs 16 poäng, för betyg 4 krävs 24 poäng, för betyg 5 krävs 32 poäng, max. 40 poäng.

Fullständig och välmotiverad lösning på uppgift krävs för att full poäng skall erhållas.

-
1. Visa mha definitioner vilka av relationerna nedan som är sanna för händelserna A , B , och C , dels då de är disjunkta och dels då de är oberoende (svar som i varje deluppgift ej tydligt anger dina argument i de två separata fallen ger inga poäng).
 - a) $P(A \cap B) = P(A)P(B)$ **(1p)**
 - b) $\mathbb{P}(A|B) = P(A)$ **(1p)**
 - c) $\mathbb{P}(A \cap B|C) = P(A|C) + P(B|C)$ **(1p)**
 - d) $\frac{P(A|B)}{P(B)} = \frac{P(B|A)}{P(A)}$ **(1p)**
 2. Antag att ξ är normalfördelad och att $P(\xi \leq 0) = 0.1587$ samt $P(\xi \leq 3) = 0.6915$.
 - a) Bestäm väntevärdet och variansen för ξ . **(3p)**
 - b) Låt $\eta \sim N(-2, 6)$ och ξ vara oberoende. Vad för fördelning kommer $\eta + \xi$ att ha (ange både fördelningens parametrar och fördelningsfamiljen samt motivera varför så är fallet)? **(2p)**
 3. Ett lager består av 20 exemplar av en viss elektronisk komponent. Man vet att 3 av dessa har smärre defekter men vet inte vilka (defekten upptäcks genom tidskrävande mätningar). Till en kund vill man leverera 4 komponenter utan defekt och chauffören chansar på att bara ta med sig fem (slumpvis) valda komponenter utan att undersöka dem. Låt X beteckna antalet defekta komponenter som chauffören får med sig.
 - a) Vad för typ av fördelning har X ? **(1p)**
 - b) Ge den exakta täthetsfunktionen för X (och specificera därmed vad parametrarna i fördelningen tar för värden och vilka värden X kan antaga). **(1p)**
 - c) Vad är sannolikheten att kunden blir nöjd (dvs att hon kommer att plocka ut minst 4 defekt-fria komponenter)? **(1p)**
 - d) Hur många defekta komponenter förväntas levereras? **(1p)**
 - e) Vad har X för standardavvikelse? **(1p)**
 4. Små resultat för de stokastiska variablerna X och Y .
 - a) Visa med hjälp av definitionen av (ko)varians att $\text{Cov}(X, Y) = \text{Var}(X) = \text{Var}(Y)$ om $X = Y$. **(0.5p)**

- b) Visa med hjälp av definitionen av (ko)varians att $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$. **(1.5p)**
- c) Finn $\mathbb{E}[X]$, $\mathbb{E}[X^2]$ och $\text{Var}(X)$ då den momentgenererande funktionen för X ges av $m_X(t) = e^{\lambda(e^t - 1)}$. **(1.5p)**
- d) Låt de momentgenererande funktionerna ges av $m_X(t) = e^{\lambda(e^t - 1)}$ och $m_Y(t) = e^{\gamma(e^t - 1)}$. Härled $m_{X+Y}(t)$ då X och Y är oberoende och använd denna för att finna $\mathbb{E}[X + Y]$. **(1.5p)**
5. Möbelformgivaren AKIE har fått klagomål från en del kunder som hävdar att det är krångligare (och därmed mer tidsödande) att montera AKIEs bokhylla än motsvarande hylla från konkurrenten OMI. För att undersöka detta anlitar AKIE 10 försökspersoner som var och en får montera en hylla från vardera fabrikanter. Ordningen slumpas för varje person. Resultaten (tid i minuter):

Typ\Person	1	2	3	4	5	6	7	8	9	10
AKIE	66	63	78	48	41	79	53	55	81	69
OMI	55	48	78	43	39	73	59	49	86	55

Normalfördelningsantagandet kan antagas gälla i alla situationer i uppgiften.

- a) För att motivera huruvida det finns fog för kundernas klagomål, utför ett lämpligt test där du drar din slutsats utifrån testets p-värde (tänk på hur du formulerar din hypotes). **(2.5p)**
- b) Motivera utifrån ett dubbelsidigt 95%igt konfidensintervall (som skall tagas fram) huruvida det finns en skillnad i monterings tiden mellan de två fabrikerna. **(2p)**
- c) Motivera varför du har valt just de metoder som du har använt i a) och b). **(0.5p)**
6. Låt $X_1, \dots, X_k$ vara ett stickprov från $Poi(\lambda)$ -fördelningen.
- a) Härled $\hat{\lambda}_{ml}$; maximum-likelihood-skattaren för λ . **(2p)**
- b) Vad blir maximum-likelihood-skattningen då vi har följande 10 observationer? **(1p)**
- 2 5 1 3 2 3 0 8 2 5
- c) Är skattaren $\hat{\lambda}_{ml}$ en väntevärdesriktig (unbiased) skattare för λ ? Vad har skattaren $\hat{\lambda}_{ml}$ för standardfel? Visa. **(2p)**
7. Följande data representerar koldioxidutsläpp (CO_2) mellan åren 1965 och 1977 från koleldade kraftverk som har använts för vattenuppvärmning (enheten är 1000 ton). Den förklarande variabeln (år) har standardiserats för att ge följande tabell:

År (x)	0	5	8	9	10	11	12
CO_2 -utsläpp (y)	910	680	520	450	370	380	340

- a) Finn skattningen av den linjära regressions ekvationen $\mu_{Y|x}$ och visa dess slutgiltiga form. **(1p)**

- b) Tolka de skattade koefficienterna i ord. **(1p)**
- c) Finns det en signifikant linjär trend i CO_2 -utsläpp över den uppmätta tidsperioden (dvs utför lämpligt test på signifikansnivå $\alpha = 0.01$)? **(2.5p)**
- d) Ge de exakta antaganden som vi måste göra beträffande $Y_1, \dots, Y_7$, $Y_i = \beta_0 + \beta_1 x_i + E_i$, för att vi skall kunna utföra testet i uppgift b). **(1p)**
- e) Vad kan man ge för kommentar kring användandet av den skattade regressionslinjen för att skatta det förväntade CO_2 -utsläppet år 2000? **(0.5p)**

Följande beräkningar kan komma till användning:

$$\begin{array}{ll} \sum_i x_i = 55 & \sum_i y_i = 3650 \\ \sum_i x_i^2 = 535 & \sum_i y_i^2 = 2160300 \\ \sum_i x_i y_i = 23570 & \end{array}$$

8. Man använder sig av två olika typer av metalldetektorer på flygplatser jorden över. Den ena kallas för kontinuerlig vågdetektor och den andra kallas pulsfältsvågdetektor. Båda detektorerna är lika effektiva vad gäller att upptäcka stora metallobjekt såsom skjutvapen och knivar. Dock tror man att den kontinuerliga vågdetektorn tenderar att vara mindre effektiv i det att den lättare triggas av små objekt som t.ex. mynt, nycklar och andra små harmlösa metallföremål. För att besvara frågan huruvida den kontinuerliga vågdetektorn triggas av en högre andel passagerare observerar man, för vardera detektorn, 175 slumpvis utvalda passagerare som går igenom vardera detektortypen. I fallet av den kontinuerliga vågdetektorn triggades den 113 gånger emedan pulsfältsvågdetektorn endast triggades för 4 av de passerande.
 - a) Utför ett lämpligt ensidigt test som på signifikansnivån $\alpha = 0.001$ besvarar huruvida den kontinuerliga vågdetektorn triggas av en högre andel passagerare. **(2.5p)**
 - b) Vad skulle du säga att (det approximativa) p-värdet för testet som utfördes i a) blir? **(1p)**
 - c) Vad är det för teoretiskt resultat som rättfärdigar de metoder som du använder dig av i a) och b)? **(1p)**
 - d) Vad för statistiskt fel finns det risk att vi begår, baserat på vår utsaga i a)? **(1/2p)**