



Avd. Matematisk statistik

KTH Matematik

TENTAMEN I SF1903 SANNOLIKHETSLÄRA OCH STATISTIK FÖR 3-ÅRIG Media TIMEH
ONSDAGEN DEN 12 JANUARI 2011 KL 14.00–19.00.

Examinator: Gunnar Englund , tel. 790 74 16.

Tillåtna hjälpmedel: Läroboken. Formel- och tabellsamling i Matematisk statistik. Mathematics Handbook (Beta). 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 20 poäng. Möjlighet att komplettera ges för tentander med 18-19 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 vara rättad inom tre veckor från skrivningstillfället och kommer att finnas tillgänglig på elevexpeditionen minst sju veckor efter skrivningstillfället.

Uppgift 1

För händelserna A och B gäller att $P(A) = 0.5$ och $P(B) = 0.3$ samt $P(A \cup B) = 0.65$.

a) Beräkna $P(A \cap B)$ samt avgör om A och B är oberoende händelser. (6 p).

b) Beräkna $P(A|A \cup B)$. (4 p).

Uppgift 2

Vid en undersökning av livslängder av glödlampor testades 20 lampor och man erhöll följande observationer (timmar)

1045	963	1066	1111	995	1073	1093	950	958	1059
1010	1064	1071	1066	1095	1063	1090	970	1029	1022

Man antar att livslängden av en glödlampa är $N(\mu, \sigma)$ och att lampornas livslängder är oberoende av varandra. Parametrarna μ och σ är okända.

Man har ut datamaterialet ovan räknat ut $\bar{x} = 1039.65$ samt $s = 49.9$.

a) Ge en lämplig uppskattning av sannolikheten att en lampa fungerar mer än 1000 timmar. (6 p)

b) Beräkna ett 95 % konfidensintervall för den förväntade livslängden μ . (4 p)

Uppgift 3

I en produktionsprocess blir enheterna felaktiga med sannolikhet p oberoende av varandra.

a) Vad är sannolikheten att av 10 tillverkade enheter minst 3 blir felaktiga om $p = 0.1$? (6 p)

b) Om två tillverkade enheter i rad blir felaktiga, justeras tillverkningsprocessen. Vad är sannolikheten att tillverkningsprocessen justeras innan 5 enheter (dvs högst 4) har tillverkats om $p = 0.1$? (4 p)

Uppgift 4

Vid en tidpunkt har 1% av en population en sjukdom S. Låt B vara händelsen att en på måfå vald person har sjukdomen, dvs. $P(B) = 0.01$.

En läkare som undersöker personen ställer rätt diagnos med sannolikhet 80% om personen har sjukdomen S, och 95% om personen inte har sjukdomen S.

a) Bestäm sannolikheten för att läkaren ställer rätt diagnos. (5 p)

b) Bestäm sannolikheten för att en person med diagnosen "ej sjukdom S" ändå har sjukdomen, dvs. med A som händelsen "Diagnos: sjukdom S", bestäm $P(B|A^*)$. (5 p)

Uppgift 5

Ett bussbolag som har 81 bussar önskar dimensionera en reparationsverkstad. Man önskar därför få en uppfattning om det totala antalet större reparationer som behöver göras på en månad. Av erfarenhet vet man att antalet gånger som en buss behöver genomgå en större reparation under en månad kan beskrivas av en stokastisk variabel X med sannolikhetsfunktionen

$$p_X(k) = \begin{cases} 0.2 & \text{om } k = 0 \\ 0.5 & \text{om } k = 1 \\ 0.2 & \text{om } k = 2 \\ 0.1 & \text{om } k = 3 \\ 0 & \text{annars.} \end{cases}$$

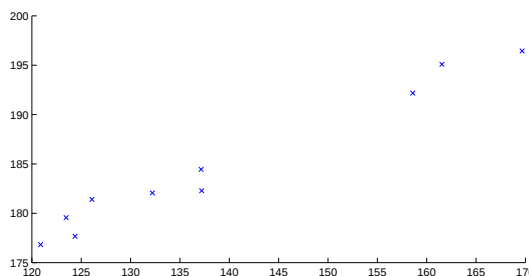
a) Beräkna $E(X)$ och $D(X)$. (5 p)

b) Beräkna med lämplig och välmotiverad approximation sannolikheten att minst 90 och högst 110 större reparationer totalt behöver göras under en månad. (5 p)

Uppgift 6

Tio mätningar av mängden spjälkat protein, $y_1, \dots, y_{10}$, är gjorda vid tio olika koncentrationer, $x_1, \dots, x_{10}$, av det aktiva enzymet. Ansätt regressionsmodellen att vid enzymkoncentration x_i beskrivs mängden protein av Y_i där Y_i är $N(\alpha + \beta x_i, \sigma)$.

De $n = 10$ mätningarna



sammanfattas av följande storheter

$$\bar{x} = 139.12 \quad \bar{y} = 184.79$$

$$\sum_{i=1}^n (x_i - \bar{x})^2 = 2834.9 \quad \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = 1131.4 \quad \sum_{i=1}^n (y_i - \bar{y})^2 = 463.36$$

a) Skatta regressionslinjens parametrar α och β . (4 p)

b) Skatta σ . (2 p)

c) Testa hypotesen $H_0 : \beta = 0.30$ mot $H_1 : \beta \neq 0.30$ på signifikansnivå 0.05. (4 p)



KTH Matematik

Avd. Matematisk statistik

LÖSNINGAR TILL

TENTAMEN I SF1903 (f d 5B2501) SANNOLIKHETSLÄRA OCH STATISTIK ONSDAGEN DEN 12 JANUARI 2011 KL 14.00–19.00.

Uppgift 1

a) Eftersom $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ erhålls $P(A \cap B) = P(A) + P(B) - P(A \cup B) = 0.5 + 0.3 - 0.65 = 0.15$. Vi får $P(A \cap B) = 0.15 = 0.5 \cdot 0.3 = P(A)P(B)$ inses att A och B är oberoende händelser.

b) Vi får

$$P(A|A \cup B) = \frac{P(A \cap (A \cup B))}{P(A \cup B)} = \frac{P(A)}{P(A \cup B)} = \frac{0.5}{0.65} = 10/13 \approx 0.77.$$

Uppgift 2

a) Om X är livslängden av en glödlampa så är

$$P(X > 1000) = 1 - P(X \leq 1000) = 1 - \Phi\left(\frac{1000 - \mu}{\sigma}\right) = \Phi\left(\frac{\mu - 1000}{\sigma}\right).$$

Detta värde uppskattar vi med

$$\Phi\left(\frac{\mu_{\text{obs}}^* - 1000}{s}\right) = \Phi\left(\frac{1039.65 - 1000}{49.9}\right) = \Phi(0.7946) = 0.7873.$$

Vi uppskattar alltså den sökta sannolikheten med 0.7873.

b) Ett 95 % konfidensintervall ges av $\bar{x} \pm t_{0.025}(n-1)s/\sqrt{n} = 1039.65 \pm 2.09 \cdot 49.9/\sqrt{20} = 1039.7 \pm 23.3$.

Uppgift 3

a) Låt X vara antalet felaktiga enheter av 10 tillverkade. Då är $X \in \text{Bin}(10, p)$ där p är sannolikheten att en enhet blir felaktig. Vi får från tabell 8, med $p = 0.1$, att $P(X \geq 3) = 1 - P(X \leq 2) = 1 - 0.9298 = \underline{0.0702}$.

b) Av fyra tillverkade enheter, kan två på varandra följande felaktiga enheter kan uppkomma på fyra sätt:

A : De två första enheterna är felaktiga.

B : Den första korrekt men den andra och den tredje enheten är felaktig.

C : De två första korrekta, men de två sista felaktiga.

D : Den första fel, den andra korrekt och de två sista felaktiga.

A , B , C och D är disjunkta. $P(A) = 0.1^2 = 0.01$ och $P(B) = 0.9 \cdot 0.1^2 = 0.009$, $P(C) = 0.9^2 \cdot 0.1^2 = 0.0081$ och $P(D) = 0.1 \cdot 0.9 \cdot 0.1^2 = 0.0009$ varför den sökta sannolikheten blir $P(A \cup B \cup C \cup D) = P(A) + P(B) + P(C) + P(D) = \underline{0.028}$.

Uppgift 4

Med A som händelsen "Diagnos sjukdom S " formuleras de givna sannolikheterna som

$$P(B) = 0.01 \quad P(A|B) = 0.80 \quad P(A^*|B^*) = 0.95$$

Alltså är

$$\begin{aligned} P(\text{rätt diagnos}) &= P(A \cap B) + P(A^* \cap B^*) \\ &= P(A|B)P(B) + P(A^*|B^*)P(B^*) \\ &= 0.80 \cdot 0.01 + 0.95 \cdot (1 - 0.01) = 0.9485. \end{aligned}$$

$$P(A) = P(A \cap B) + P(A \cap B^*) = P(A|B)P(B) + P(A|B^*)P(B^*) = 0.0575,$$

så

$$P(B|A^*) = \frac{P(A^* \cap B)}{P(A^*)} = \frac{P(A^*|B)P(B)}{1 - P(A)} = \frac{0.20 \cdot 0.01}{1 - 0.0575} = 0.02122.$$

Uppgift 5

a) Vi har att

$$\begin{aligned} E(X) &= \sum_{k=0}^3 k p_X(k) = 0 \cdot 0.2 + 1 \cdot 0.5 + 2 \cdot 0.2 + 3 \cdot 0.1 = \underline{1.2}. \\ E(X^2) &= \sum_{k=0}^3 k^2 p_X(k) = 0^2 \cdot 0.2 + 1^2 \cdot 0.5 + 2^2 \cdot 0.2 + 3^2 \cdot 0.1 = 2.2, \end{aligned}$$

vilket ger

$$V(X) = E(X^2) - (E(X))^2 = 2.2 - 1.2^2 = 2.2 - 1.44 = 0.76$$

och

$$D(X) = \sqrt{0.76} = \underline{0.872}.$$

b) Låt X_j vara antalet reparationer som buss nr j behöver genomgå under en månad, $j = 1, \dots, 81$. Modell: $X_1, \dots, X_{81}$ är oberoende och lika fördelade stokastiska variabler med samma fördelning som X . Sätt $Y = \sum_{j=1}^{81} X_j$. Vi har

$$E(Y) = \sum_{j=1}^{81} E(X_j) = 81 \cdot 1.2 = 97.2 = m$$

och

$$V(Y) = (\text{pga oberoendet}) = \sum_{j=1}^{81} V(X_j) = 81 \cdot 0.76 = 61.56 = \sigma^2.$$

Centrala gränsvärdessatsen ger att Y är approximativt normalfördelad $N(m, \sigma)$. Således fås

$$\begin{aligned} P(90 \leq Y \leq 110) &= P\left(\frac{90 - m}{\sigma} \leq \frac{Y - m}{\sigma} \leq \frac{110 - m}{\sigma}\right) \approx \Phi\left(\frac{110 - m}{\sigma}\right) - \Phi\left(\frac{90 - m}{\sigma}\right) \\ &= \Phi(1.63) - \Phi(-0.92) = \Phi(1.63) - (1 - \Phi(0.92)) \\ &= 0.9484 - (1 - 0.8212) = \underline{0.77}. \end{aligned}$$

Uppgift 6

a) Med beteckningar som i uppgiftslydelsen skattas α och β med

$$\beta_{\text{obs}}^* = \frac{S_{xy}}{S_{xx}} = \frac{1131.4}{2834.9} = 0.39909, \quad \alpha_{\text{obs}}^* = \bar{y} - \beta_{\text{obs}}^* \bar{x} = 184.79 - 0.39909 \cdot 139.12 = 129.27$$

b) Vi skattar σ^2 med

$$s^2 = \frac{1}{n-2} \left(\sum_{i=1}^n (y_i - \bar{y})^2 - (\beta_{\text{obs}}^*)^2 \sum_{i=1}^n (x_i - \bar{x})^2 \right) = 1.4777$$

dvs, σ skattas med $s = \underline{1.2156}$.

c) Vi förkastar hypotesen $H_0 : \beta = 0.30$ till förmån för $H_1 : \beta \neq 0.30$ på nivån 5% om 0.30 ej ingår i det 95%-iga konfidensintervallet för β . Detta intervall blir

$$\beta_{\text{obs}}^* \pm t_{0.025}(10-2) \frac{s}{\sqrt{S_{xx}}} = 0.39909 \pm 2.31 \frac{1.2156}{\sqrt{2834.9}} = 0.39909 \pm 0.0527.$$

Undre gränsen blir 0.3464 så $H_0 : \beta = 0.30$ förkastas.