| | |
|---|---|
| **Matematisk statistik** | **Tentamen: 2009–01–08 kl $08^{00}$–$13^{00}$** |
| **Matematikcentrum** | **FMS 012 — Matematisk statistik för I, CDE, FΠN 9 hp** |
| **Lunds tekniska högskola** | **MAS B03 — Matematisk statistik för fysiker, 9 hp** |
| **Lunds universitet** | **(Även FMS022, FMS121, FMS233 för CDE, I, resp. fysiker)** |

Korrekt, väl motiverad lösning på uppgift 1–3 ger 10 poäng vardera medan uppgift 4–6 ger 20 poäng vardera. Totalt kan man få 90 poäng. Gränsen för godkänd är 40 poäng.

Institutionens papper används både som kladdpapper och som inskrivningspapper. Varje lösning skall börja överst på nytt papper. Rödpenna får ej användas. Skriv fullständigt namn på alla papper.

Tillåtna hjälpmedel: Matematiska och statistiska tabeller som ej innehåller statistiska formler, Formelsamling i matematisk statistik AK 2001 eller senare, samt miniräknare.

**Resultatet anslås senast torsdag den 15 januari i matematikhusets entréhall.**

---

1. I en trädgård planteras tulpaner i olika färger. Färg-fördelningen på tulpanblommorna är $\mathsf{P}(\text{gul}) = 0.3$, $\mathsf{P}(\text{röd}) = 0.5$, och $\mathsf{P}(\text{svart}) = 0.2$. Ett rådjur som bor i området gillar tulpaner, och äter upp tulpaner det träffar på, med olika sannolikhet beroende på vilken färg tulpanen har:

$$\begin{aligned} \mathsf{P}(\text{rådjuret äter en tulpan} \mid \text{tulpanen är gul}) &= 1.0, \\ \mathsf{P}(\text{rådjuret äter en tulpan} \mid \text{tulpanen är röd}) &= 0.6, \\ \mathsf{P}(\text{rådjuret äter en tulpan} \mid \text{tulpanen är svart}) &= 0.1 \end{aligned}$$

I ett försök undvika att rådjuret äter upp alla tulpaner placeras en godtycklig av tulpanerna utanför trädgårdens staket, i hopp om att rådjuret ska nöja sig med denna tulpan.

   (a) Beräkna sannolikheten att rådjuret äter upp tulpanen om det träffar på den. (4p)

   (b) En viss kväll passerar rådjuret tulpanen med sannolikhet 0.8. Beräkna den betingade sannolikheten att rådjuret dök upp denna kväll, givet att tulpanen står kvar följande dag. (6p)

2. Den tid som behövs för att betjäna en kund som anländer till lager A kan betraktas som en summa av tre stokastiska variabler $X$, $Y$ och $Z$, som är oberoende och exponentialfördelade med väntevärdena $\mathsf{E}(X) = 2$, $\mathsf{E}(Y) = 3$ respektive $\mathsf{E}(Z) = 6$ minuter. Tiden för att betjäna en kund som kommer till lager B är däremot en enda stokastisk variabel $W$, som har en okänd fördelning men där vi känner väntevärde och standardavvikelse, $\mathsf{E}(W) = 10$ respektive $\mathsf{D}(W) = 6$ minuter.

   (a) Beräkna väntevärde och standardavvikelse för den sammanlagda tid det tar att betjäna en kund som kommer till lager A. (4p)

   (b) Beräkna approximativt sannolikheten att det går snabbare att betjäna 100 kunder vid lager A än det gör att betjäna 100 kunder vid lager B. (6p)

3. Låt $X$ och $Y$ vara oberoende exponentialfördelade stokastiska variabler, med väntevärden $\mathsf{E}(X) = a$ respektive $\mathsf{E}(Y) = 2a$. Beräkna sannolikheten att $2X \leq Y$. (10p)

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

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

Låt $\Delta$ beteckna den systematiska rostskyddseffekten.

---

Var god vänd!

Hjälpberäkningar:

$$\bar{x} = \frac{1}{10}\sum_{i=1}^{10} x_i = 33.67 \qquad \bar{y} = \frac{1}{10}\sum_{i=1}^{10} y_i = 33.21 \qquad \bar{z} = \frac{1}{10}\sum_{i=1}^{10} (y_i - x_i) = -0.4600$$

$$s_x^2 = \frac{1}{9}\sum_{i=1}^{10}(x_i - \bar{x})^2 = 27.08 \quad s_y^2 = \frac{1}{9}\sum_{i=1}^{10}(y_i - \bar{y})^2 = 26.82 \quad s_z^2 = \frac{1}{9}\sum_{i=1}^{10}(y_i - x_i - \bar{z})^2 = 0.1404$$

(a) Ange lämplig modell, och beräkna ett tvåsidigt 95 % konfidensintervall för $\Delta$. (10p)

(b) Antag nu att man blandat ihop de obehandlade stavarna, så att man inte vet med vilka behandlade stavar de grävts ner med. Ange lämplig modell, och beräkna ett tvåsidigt 95 % konfidensintervall för $\Delta$. Jämför med svaret i (a). (10p)

5. Läget som funktion av tiden för en kula som rullar utför ett svagt sluttande plan följer approximativt det kvadratiska uttrycket $\mu(t) = \beta_0 + \beta_1 \cdot t + \beta_2 \cdot t^2$, där $\beta_0$ är utgångsläget, $\beta_1$ är initialhastigheten och $2\beta_2$ är horisontalaccelerationen. I en försöksuppställning mättes kulans läge vid $n = 21$ tidpunkter, $x_i$. På grund av osäkerhet i lägesmätningarna kan vi anta en modell för mätningarna, $y_i = \mu(x_i) + \varepsilon_i$, där $\varepsilon_i \in N(0,\ \sigma)$ är oberoende mätfel.

Mätresultat och hjälpberäkningar:

$$\mathbf{X} = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0.1 & 0.01 \\ 1 & 0.2 & 0.04 \\ \vdots & \vdots & \vdots \\ 1 & 2.0 & 4.0 \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} -0.2154 \\ 0.0313 \\ 0.3663 \\ \vdots \\ 3.8431 \end{bmatrix}$$

$$(\mathbf{X}^T\mathbf{X})^{-1} = \begin{bmatrix} 0.3429 & -0.6374 & 0.2470 \\ -0.6374 & 1.6693 & -0.7411 \\ 0.2470 & -0.7411 & 0.3529 \end{bmatrix} \qquad \mathbf{X}^T\mathbf{y} = \begin{bmatrix} 45.44 \\ 65.29 \\ 104.84 \end{bmatrix}$$

$$Q_0 = 0.48391$$

(a) Skatta de fyra okända parametrarna i modellen. (4p)

(b) Testa på nivån 99% om planet verkligen lutar, dvs om accelerationen är lika med 0. (8p)

(c) Oavsett svaret i (b), använd den skattade modellen för att konstruera ett konfidensintervall för läget vid tiden $x_0 = 1.5$. (8p)

6. Livslängderna hos en viss typ av bultar kan anses vara exponentialfördelade med väntevärde $\mu$. Om man monterar två bultar fördelas belastningen så att livslängden fördubblas och blir exponentialfördelad med väntevärde $2\mu$. Man har en observation $x = 11$ från en enkelmonterad bult och en observation $y = 25$ från en dubbelmonterad bult, och vill skatta $\mu$.

(a) Härled Maximum Likelihood-skattningen $\mu^*_{ML}$ av $\mu$. (6p)

(b) Sätt $Z = \frac{1}{2}\left(X + \frac{Y}{2}\right)$ och beräkna väntevärde och varians för $Z$. (6p)

(c) Man har också noterat att den bult som satt monterad tillsammans med $y$ hade livslängden $y_2 = 20$. Man vill utnyttja denna mätning också, och föreslår skattningen $\mu^*_{ny} = \frac{1}{2}\left(X + \frac{Y+Y_2}{4}\right)$. Livslängderna för de två dubbelmonterade bultarna är inte oberoende. De är båda exponentialfördelade med väntevärde $2\mu$ men har kovariansen $\mathsf{C}(Y, Y_2) = \mu^2$. Beräkna variansen för $\mu^*_{ny}$ och avgör om skattningen blir bättre när vi tar med observationen $y_2$ på detta sättet. (8p)

---

Lycka till!