Tentamensskrivning i Matematisk statistik F1
Torsdagen den 11/1, 2001.

Jour: Mikael Knutsson, tel 7725380, examinator Urban Hjorth.

Hjälpmedel: Tabeller och fördelningstablå som utdelas vid skrivningen. Kalkylator med tomt textminne.

Varje väl löst uppgift ger 3 poäng och betygskraven blir 11, 15, 19 för betygen 3, 4, 5. Dessutom krävs minst 3 poäng på vardera delen.

Del 1:

1. Till en brandstation inkommer i genomsnitt 2 larm i veckan. Vi antar att denna siffra gäller alla årets veckor (dvs vi bortser från den extra stora risken vid storhelger som jul och nyår).

a) Vilken sannolikhetsfördelning bör man ansätta på antalet larm under en vecka?

b) Vad är sannolikheten för 5 eller fler larm under en vecka?

c) Vad är sannolikheten att någon av årets 52 veckor ger 5 eller fler larm?

2. Bestäm väntevärde och standardavvikelse för summan av 10 oberoende variabler som alla kan anta värdena 1, 2 eller 3 med sannolikheterna 0.6, 0.3 och 0.1 respektive.

3. Livslängden X för en fordonskomponent (del av bromssystemet) har modellerats som en stokastisk variabel med frekvensfunktionen

$$f(x) = abx^{b-1}e^{-ax^b},\ x > 0.$$

(Enhet 10000 km körsträcka.) Parametrarna har skattats till $a = 0.0024$, $b = 2.2$. För att bedöma sannolikheten att komponenter som körts felfritt sträckan t håller i ytterligare ett serviceintervall (1.5 enheter) så beräknas den betingade sannolikheten

$$P(X > t + 1.5 | X > t).$$

Finn denna sannolikhet för $t = 3$ och $t = 9$.

4. a) Låt X och Y vara oberoende med momentgenererande funktioner $m_X(t)$ och $m_Y(t)$. Härled momentgenererande funktionen för $aX + bY + c$ uttryckt i m_X och m_Y.

b) Låt nu X_i, $i = 1, \ldots, n$ vara oberoende $N(\mu, \sigma)$. Tillämpa a-delen (generaliserad till flera variabler) och härled momentgenererande funktionen för $(\bar{X} - \mu)/(\sigma/\sqrt{n})$ där $\bar{X} = \sum_1^n X_i/n$. Obs. det är härledningen och inte slutresultatet som efterfrågas.

Vänd!!

Del 2:

5. Andelen θ av alla tvillingfödslar är enäggstvillingar ocg dessa får säkert samma kön. Resten är tvåäggstvillingar och får manligt eller kvinnligt kön oberoende av varandra. (Sannolikheten för pojke är ca 0.514 men vi räknar här med lika sannolikheter 0.5 för flicka och pojk.) Av 1074 tvillingfödslar under ett år var 667 likkönade medan resten, dvs 407 stycken, gav olika kön. Skatta parametern θ.

6. För att jämföra fiskars tillväxt i olika miljöer så väger och märker man ett antal ungefär lika stora fiskar som sedan fördelas slumpvis och får leva i nätburar i respektive vatten. Efter en tid mäter man fiskarnas viktökning. Data (gram viktökning)

Miljö 1: 72, 35, 54, 68, 61, 49, 66, 70
Miljö 2: 39, 54, 50, 51, 46, 43, 33

Antag att alla fiskar tillväxer oberoende av varandra. Pröva med ett lämpligt test på signifikansnivån $\alpha = 0.10$ om väntevärdet av tillväxten kan vara lika i de båda miljöerna. (Lika varians får förutsättas).

7. Genom en ytbehandling ändras sannolikheten för defekter från p till θp där både p och θ är okända konstanter. Ställ upp likelihooden och finn maximum likelihoodskattningen av θ och p med hjälp av följande data från oberoende försök

utan ytbehandling: 1076 enheter gav 287 defekter;
med ytbehandling: 812 enheter gav 113 defekter.

8. Vi har gjort två regressionsanalyser med samma typ av modell $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ men på helt oberoende datamaterial. Standardavvikelsen σ för ε_i är samma i båda datamaterialen trots en rätt stor skillnad på de skattade standardavvikelserna. Gör ett konfidensintervall för att studera om lutningen β_1 är olika i de båda datamaterialen. Lämplig statistisk felrisk 0.01. Från en datorkörning får vi följande resultat:

Datamaterial 1: $n_1 = 20$ data. $\hat{\beta}_0 = 4.6761$, $\hat{\beta}_1 = 0.6249$, $s = 3.3926$, $X'X = \begin{pmatrix} 20 & 540 \\ 540 & 17538 \end{pmatrix}$, $(X'X)^{-1} = \begin{pmatrix} 0.2965 & -0.0091 \\ -0.0091 & 0.00034 \end{pmatrix}$.

Datamaterial 2: $n_2 = 30$ data. $\hat{\beta}_0 = 2.4258$, $\hat{\beta}_1 = 0.8819$, $s = 5.0514$, $X'X = \begin{pmatrix} 30 & 764 \\ 764 & 27672 \end{pmatrix}$, $(X'X)^{-1} = \begin{pmatrix} 0.1123 & -0.0031 \\ -0.0031 & 0.00012 \end{pmatrix}$.

Vi erinrar om att vid normalfördelning har vektorn $\hat{\beta}$ en $N(\beta, \sigma^2(X'X)^{-1})$-fördelning.

Tenta i Matematisk statistik för F1, 2001-01-11

UPPGIFT 1: (a) Poisson(2)

(b) $P(X \geq 5) = 1 - P(X \leq 4) = 1 - \sum_{x=0}^{4} P(X = x) =$

$= 1 - \sum_{x=0}^{4} \frac{e^{-2} 2^x}{x!} \approx 0{,}0527$

(c) $1 - (1 - 0.0527)^{52} \approx 0.94$

UPPGIFT 2: $Y = \sum_{i=1}^{10} X_i$, där $X_1, X_2, \ldots, X_{10}$ ober.

$E[X_i] = 1 \cdot 0.6 + 2 \cdot 0.3 + 3 \cdot 0.1 = 1.5$

$E[X_i^2] = 1^2 \cdot 0.6 + 2^2 \cdot 0.3 + 3^2 \cdot 0.1 = 2.7$

$\Rightarrow Var(X_i) = 2.7 - 1.5^2 = 0.45$

$\Rightarrow E[Y] = \sum_{i=1}^{10} E[X_i] = 10 \cdot 1.5 = 15$

$Var(Y) = \sum Var(Y_i) = 10 \cdot 0.45 = 4.5$,

dvs standardavvikelse $= \sqrt{4.5} \approx 2.12$

UPPGIFT 3:

$P(X > t) = \int_t^\infty abx^{b-1} e^{-ax^b} dx = \left[-e^{-ax^b}\right]_t^\infty =$

$= e^{-at^b}$

$\Rightarrow P(X > t+1.5 \mid X > t) = \frac{P(X > t+1.5)}{P(X > t)} = \frac{e^{-a(t+1.5)^b}}{e^{-at^b}} =$

$= \begin{cases} 0.9620, & \text{då } t = 3 \\ 0.8853, & \text{då } t = 9 \end{cases}$

UPPGIFT 4: (a) $m(t) = E[e^{t(aX + bY + c)}] =$

$= E[e^{atX}] E[e^{btY}] e^{ct} = e^{ct} m_X(at) m_Y(bt)$

(b) $E[e^{t[(\bar{X} - \mu)/(\sigma/\sqrt{n})]}] = e^{-\mu t/(\sigma/\sqrt{n})} \prod_{i=1}^{n} E[e^{t \cdot \frac{1}{n} \cdot \frac{X_i}{\sigma/\sqrt{n}}}]$

$= e^{-\mu t/(\sigma/\sqrt{n})} \prod_{i=1}^{n} E[e^{tX_i/(\sigma\sqrt{n})}] = e^{-\mu t/(\sigma/\sqrt{n})} \prod_{i=1}^{n} m_{X_i}\left(\frac{t}{\sigma\sqrt{n}}\right) =$

$= e^{-\frac{\mu t}{\sigma/\sqrt{n}}} \prod_{i=1}^{n} e^{\frac{\mu t}{\sigma\sqrt{n}} + t^2/2n} = e^{t^2/2}$

UPPGIFT 5: $P(\text{samma kön}) = \theta + (1-\theta)\cdot 0.5 = 0.5(1+\theta)$

$P(\text{olika kön}) = (1-\theta)\cdot 0.5$

Alternativ 1: Skatta θ med det θ som maximerar

$L(\theta) = [0.5(1+\theta)]^{667}[0.5(1-\theta)]^{407}$

$\Rightarrow \hat{\theta} = 260/1074 \approx 0.2421$

Alternativ 2: Skatta $0.5(1+\theta)$ med $667/1074$

$\Rightarrow$ skatta θ med $\hat{\theta} = 260/1074 \approx 0.2421$

UPPGIFT 6: Miljö 1: $n_1 = 8$, $\bar{x}_1 = 59.375$, $s_1^2 \approx 160.554$

Miljö 2: $n_2 = 7$, $\bar{x}_2 = 45.143$, $s_2^2 = 54.476$

$$S_p^2 = \frac{7s_1^2 + 6s_2^2}{13} \approx 111.595$$

Testa $H_0: \mu_1 = \mu_2$ mot $H_1: \mu_1 \neq \mu_2$

$$T = \frac{\bar{x}_1 - \bar{x}_2}{S_p\sqrt{1/n_1 + 1/n_2}} \approx 2.6 > 1.771 = t_{13, 0.05}$$

$\Rightarrow$ Förkasta H_0 på $\alpha = 0.10$-nivå.

UPPGIFT 7: Skatta p med det p som maximerar

$L_1(p) = p^{287}(1-p)^{789} \Rightarrow \hat{p} = \frac{287}{1076} \approx 0.2667$

Skatta θp med det värde som maximerar

$L_2(\theta p) = (\theta p)^{113}(1-\theta p)^{699} \Rightarrow (\widehat{\theta p}) = \frac{113}{812}$

$\Rightarrow$ Skatta θ med $\hat{\theta} = (113/812)/(287/1076)$
≈ 0.5217

UPPGIFT 8:

99% k.i. för $\beta_{11} - \beta_{12}$:

$$\hat{\beta}_{11} - \hat{\beta}_{12} \pm t^{(0.005)}_{(20-2)+(30-2)} \cdot S_p \sqrt{0.00034 + 0.00012} =$$

$$= 0.6249 - 0.8819 \pm 2.687 \cdot \sqrt{\frac{18 \cdot 3.3926^2 + 28 \cdot 5.0514^2}{46}} \cdot \sqrt{0.00046} =$$

$$= -0.257 \pm 0.258$$