EXAMINATION: Tentamensskrivning i Matematisk Statistik (TMS061)

Time: Thursday 21 August 2008

Jour: Erik Jacobsson, mobile: 0737353153

Aid: You are allowed to use a scientific calculator and a one page (both sides) of hand written notes

Grade: You need 32 out of (30+10) points for 5, 24 points for 4 and 17 points for 3.

Motivate all your answers. Good Luck!

1) The random variable X has probability distribution function $f(x) = p(1-p)^{x-1}, \quad x = 1, 2, \ldots$

a) Compute $P(X \leq 2)$. (1p)

b) Describe a situation where this probability distribution could be used. (1p)

2) The time when a technical system fails follow a Poisson distribution with intensity $\lambda = 1$ per month. The system is operated for 2 months. What is the probability that

a) the system fails once during each of the two months? (2p)

b) the system fails twice in the first month and does not fail at all during the second month? (1p)

c) the system fails twice during one month and does not fail at all during the other month? (1p)

(3) The continuous random variable X has probability density function

$$f(x) = \theta x^{\theta - 1}, \quad 0 \leq x \leq 1,$$

where $\theta > 0$ is an unknown parameter. Find the maximum likelihood estimator of θ based on the sample $(0.3, 0.5, 0.8)$ of X. (4p)

VAR GOD VÄND!

4) The values 14.9, 15.0, 15.1 come from a normal distribution with unknown mean μ and unknown variance σ^2. Construct a 95% confidence interval for the mean and a 95% confidence interval for the variance σ^2. (4p)

5) Of 1000 numbers that are randomly generated in the interval $[0, 1]$, 226 of them fall in the interval $(0, 0.25)$, 270 in the interval $(0.25, 0.50)$, 278 in $(0.5, 0.75)$ and 226 in $(0.75, 1)$. Test at the 95% level, the hypothesis that the generator is an unbiased one, that is the numbers generated inside each interval are independent of the interval. (5p)

6) In order to study the relation between the temperature and the melting index of a material, we have measured the melting index at 5 different temperatures of the same material. The results are given below:

Temperature x_i:	30	40	50	60	70
Melting point y_i:	0.71	0.73	0.77	0.79	0.83

Assume that y_i are observations from the random variable Y_i that are normally distributed.

a) Calculate the least squares estimate of the slope β_1 and intercept β_0. (2.5p)

b) Test the hypothesis $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$ at the 99% confidence level. (2.5p)

($\sum x_i = 250, \sum x_i^2 = 13500, \sum y_i = 3.82, \sum y_i^2 = 2.9264, \sum x_i y_i = 193.8$)

7) Let X be the number of heads and Y the number of tails in 10 independent throws of a fair coin. What is the correlation coefficient between X and Y? Motivate your answer! (3p)

8) Let A and B be two events such that $P(A) = 0.4$, $P(B) = p$ and $P(A \cup B) = 0.8$. For what value of p are the events A and B
a) disjoint? (1.5p)

b) independent? (1.5p)

21 August 2008. Matematisk statistik TMS 061

#1 a) $f(x) = p(1-p)^{x-1}$, $x = 1, 2, 3 \ldots$

$$P(X \le 2) = P(X=1) + P(X=2) = p(1-p)^{1-1} + p(1-p)^{2-1} =$$
$$= p + p(1-p) = p\cdot(1+1-p) = \underline{\underline{p\cdot(2-p)}}$$

b) This distribution counts number of trials until the first success.

#2 a) $f(x) = \dfrac{e^{-\lambda}\lambda^x}{x!}$

Let X_1 = # of failures during 1^{st} month $\sim Po(1)$
X_2 = # of failures — " — 2^{nd} month $\sim Po(1)$

a) $P(X_1 = 1, X_2 = 1) \overset{\text{indep}}{=} P(X_1 = 1)\cdot P(X_2 = 1) =$

$$= \frac{e^{-1}\cdot 1^1}{1!}\cdot\frac{e^{-1}\cdot 1^1}{1!} = e^{-1}\cdot e^{-1} \approx 0.1353$$

b) $P(X_1 = 2, X_2 = 0) = P(X_1 = 2)\cdot P(X_2 = 0) = \dfrac{e^{-1}1^2}{2!}\cdot\dfrac{e^{-1}1^0}{0!} =$

$$= \frac{e^{-1}}{2}e^{-1} \approx \frac{0.1353}{2}$$

c) $P(X_1 = 2, X_2 = 0) + P(X_1 = 0, X_2 = 2) = 2\,P(X_1 = 0)\,P(X_2 = 2)$

$$= 2\cdot\frac{1}{2}e^{-2} = e^{-2} \approx 0.1353.$$

#3) $f(x) = \theta x^{\theta-1}$, $0 \le x \le 1$

$$L(\theta) = \prod_{i=1}^{n} f(x_i) = \prod_{i=1}^{n} \theta \cdot x_i^{\theta-1} = \theta^n \prod_{i=1}^{n} x_i^{\theta-1}$$

$$\log L(\theta) = \log \theta^n \prod_{i=1}^{n} x_i^{\theta-1} = n \log\theta + \sum_{i=1}^{n} \log x_i^{\theta-1} =$$

$$= n \cdot \log\theta + (\theta-1) \cdot \sum_{i=1}^{n} \log x_i$$

$$\frac{\partial \log L(\theta)}{\partial\theta} = \frac{n}{\theta} + \sum_{i=1}^{n} \log x_i = 0 \Rightarrow \sum_{i=1}^{n} \log x_i = -\frac{n}{\theta} \Rightarrow$$

$$\Rightarrow \frac{1}{\sum_{i=1}^{n} \log x_i} = -\frac{\theta}{n} \Rightarrow \boxed{\hat{\theta} = -\frac{n}{\sum_{i=1}^{n} \log x_i}}.$$

$$\therefore \hat{\theta} = -\frac{3}{\log 0.3 + \log 0.5 + \log 0.8} = \underline{\underline{1.41}}$$

#4). A 95% C.I for μ from $N(\mu, \sigma^2)$ with μ and σ^2 unknown is given by

$$\mu \in \left[\bar{x} \pm \frac{s \cdot t_{0.05, n-1}}{\sqrt{n}}\right] = \left[15 \pm 4.3027 \cdot \frac{0.1}{\sqrt{3}}\right] =$$

$$= 15 \pm 0.2484 = [14.7516, 15.2484]$$

#4 continue) A 95% CI for the unknown variance is given by

$$\sigma^2 \in \left[\frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}}\right] = \left[\frac{2\cdot 0.1^2}{\chi^2_{0.025,2}}, \frac{2\cdot 0.1^2}{\chi^2_{0.975,2}}\right]$$

$$= \left[\frac{2 * 0.1^2}{7.38}, \frac{2 * 0.1^2}{0.05}\right] = [0.0163, 2.4]$$

#5)

Int.	Observed freq.	Theor. freq
0 - 0.25	226	250
0.25 - 0.5	270	250
0.5 - 0.75	275	250
0.75 - 1	226	250

H_0: same prob. for each interval
H_1: NOT H_0.

$$\chi_0^2 = \frac{(226-250)^2}{250} + \frac{(270-250)^2}{250} + \frac{(275-250)^2}{250} + \frac{(226-250)^2}{250} =$$

$= 9.34$

#5.) $\chi^2_{0.05,3} = 7.815$. $\quad n-k+1 = 4-0+1 = 3$.

Since $\chi^2_{0.05,3} = 7.815 < 9.34 = \chi_0^2$, we reject the hypothesis of independence and conclude that the generator is not unbiased

#6) $S_{xx} = \sum x_i^2 - n\bar{x}^2 = 13500 - 5 \cdot (250/5)^2 = 1000$

$S_{yy} = \sum y_i^2 - n\bar{y}^2 = 2.9024 - 5\,(3.82/5)^2 = 0.04392$

$S_{xy} = \sum x_i y_i - n\bar{x}\bar{y} = 193.8 - 5 \cdot \left(\frac{250}{5}\right) \cdot \left(\frac{3.82}{5}\right) = 2.8$

$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{2.8}{1000} = 0.0028$, $\quad \hat{\beta}_0 = 0.6280$

$S^2 = \frac{1}{3}\left(0.04392 - \hat{\beta}_1 \cdot 2.8\right) = 0.0121 \Rightarrow$

$S = \sqrt{S^2} = 0.1097$

$\beta_1 \in \left[\hat{\beta}_1 \pm t^{(3)}_{0.005} \cdot S/\sqrt{S_{xx}}\right] = (0.0028 \pm 5.841 * 0.1097/\sqrt{1000}) = (0.0028 \pm 0.020) = (-0.0172, 0.0228)$

#7) $X+Y=10 \Rightarrow Y=-X+10 \Rightarrow \rho=-1$

#8) $P(A)=0.4$, $P(B)=p$, $P(A\cup B)=0.8$

b) A, B independent $\Rightarrow P(A\cap B)=P(A)\cdot P(B)$

$P(A\cup B)=P(A)+P(B)-P(A\cap B) \overset{\text{indep}}{\Rightarrow}$

$P(A\cup B)=P(A)+P(B)-P(A)\cdot P(B) \Rightarrow$

$0.8 = 0.4 + p - p * 0.4 \Rightarrow$

$0.4 = p - 0.4 * p = 0.6p \Rightarrow$

$p = \frac{0.4}{0.6} \Rightarrow \underline{\underline{p \simeq 0.667}}$

a) A, B disjoint $\Rightarrow P(A\cap B)=0$. Then

$P(A\cup B)=P(A)+P(B) \Rightarrow$

$0.8 = 0.4 + p \Rightarrow \underline{\underline{p=0.4}}$