Sannolikhet, statistik och risk MVE300

Written examination - 28th May 2014, 08-12:30.

Allowed aids: Mathematical or statistical tables. The tables of formulæ of the course. Any calculator (not PC). Dictionaries for translation. Jour: Igor Rychlik 0707405575.

1 Let $X \in N(10, 2^2)$ and $Y \in N(3, 1^2)$ be independent random variables and define $Z = X - Y$.

(a) Give the distribution of Z. (5 p)

(b) Calculate $P(Z < 0)$. (5 p)

2 A Gumbel distribution is used to model the yearly maximum wind speed X in Lund,

$$F_X(x) = \mathrm{e}^{-\mathrm{e}^{-\frac{x-b}{a}}}.$$

Based on the available information the ML estimates of the model parameters are $a^* = 3$ and $b^* = 17$.

(a) Estimate the 50 year wind speed, i.e. x_{50}. (5 p)

(b) What is the probability that at least ones 50 years wind will hit Lund in the following five years? (5 p)

3 In a school a fire alarm is installed. Without service, the alarm system works for a random time, T [years], with $T \in \mathrm{Exp}(8)$.

(a) What is the probability that the fire alarm does not work after one year? (4 p)

(b) According to the local safety regulation the fire alarm should be functioning with a probability exceeding 0.95. Is this requirement fulfilled? (1 p)

(c) To increase the reliability of the alarm system, a second, identical, fire alarm is installed. The two fire alarms work independently. What is the probability that non of the fire alarms work one year? (5 p)

4 At a location on the coast one was measuring the frequencies of periods of high waves. The data are important for studies of risks for coastal erosion. During period 1941-1990 one have counted number of storms during a year having waves exceeding a critical threshold. The data are presented in the following table:

Number of storms with high waves during a year	0	1	2	3	4	5
Number of years	26	13	6	3	2	0

(a) Estimate intensity λ of storms having waves exceeding the critical threshold. (5 p)

(c) Test whether the number of storms during a year is Poisson distributed. (10 p)

(c) Give an asymptotic confidence interval for λ. (5 p)

5 Consider an oil pipeline. Suppose the number of imperfections $N(t)$ along a distance t [km] can be modeled by a Poisson process, that is $N(t) \in \mathrm{Po}(\lambda t)$ where λ is intensity (km^{-1}). The intensity is unknown. Use Bayesian approach to model the uncertainty pipelines is that in average the intensity of imperfections is 2 per km, i.e. $E(\Lambda) = 2$, km^{-1} while variation coefficient $R(\Lambda) = 1$.

a) Propose suitable prior density that expresses the common experience. (5 p)

b) One had studied a 500 meters of pipeline in details and found no imperfections. Include this information in the posteriori density of Λ. (5 p)

c) Compute the predicted probability of no imperfections in one km long part of a pipeline. (10 p)

6 The mean rate of oxygenation X from the atmospheric re-aeration process for a stream depends on the mean velocity V of stream flow and average depth of the stream bed H. (We do not specify units here.) Data from 12 experiments have been recorded. Suppose that X is given by the following relation

$$X = e^{\theta_0} V^{\theta_1} H^{\theta_2} = e^{\theta_0 + \theta_1 \ln(V) + \theta_2 \ln(H)}.$$

Using statistical methodology the parameters θ_0, θ_1, and θ_2 have been estimated giving the following values $\theta_0^* = 1.52$, $\theta_1^* = 1.08$, and $\theta_2^* = 1.5$. It can be assumed that estimation errors are independent and approximately normally distributed with means zero and variances 0.64, 0.25, 0.16, respectively.

Give an approximative 95% confidence interval for X if one knows that $V = 3.7$ and $H = 5.1$. (20 p)

Good luck!

Solutions: Written examination - 28 May 2013
Sannolikhet, statistik och risk MVE300.

1 **(a)** $Z = X - Y \in N(10-3, 2^2+1^2) = N(7,5)$

(b) $P(Z<0) = P(\frac{Z-7}{\sqrt{5}} < \frac{0-7}{\sqrt{5}}) = \Phi(-3.13) = 0.0008$.

2 a) Since $P(X > x_{50}) = 1/50$, $x^*_{50} = b^* - a^* \ln(-\ln(1-1/50)) = 28.7$ [m/s].

b) Y- maximum wind speed in 5 years. Y is Gumbel with $a^* = 3$, $b^* = 17 + \ln(5)\cdot 3$.

Answer $P(Y > x_{50}) = 1 - \exp(-\exp(-(28.7-21.83)/3)) = 0.096$

Alternative solution: $1-(1-1/50)^5 = 0.096$.

3 **(a)** $\mathsf{P}(T<1) = 1-\exp(-1/8) = 0.1175$

(b) No

(c) $\mathsf{P}(T_1 < 1 \cap T_2 < 1) = \mathsf{P}(T_1<1)\mathsf{P}(T_2<1) = (1-\exp(-1/8))^2 = 0.0138$

4 **(a)** $\lambda^* = (12 + 2\cdot 6 + 3\cdot 3 + 4\cdot 2)/50 = 0.82$.

(b) χ^2-test, 3-classes $N=0$, $N=1$, $N>1$. (We combine the last three cases in order to satisfy condition $np_i > 5$.) Since $\lambda^* = 0.82$, $p_i^* = (0.4404, 0.3612, 0.1984)$, $n_i = (26, 13, 11)$, $n = 50 \Rightarrow Q = 2.255 < \chi^2_{0.05}(3-1-1) = 3.841$, Do not reject.

(c) Approximative confidence interval is $[\lambda^* - 1.96\sqrt{\lambda^*/50}, \lambda^* + 1.96\sqrt{\lambda^*/50}] = [0.57, 1.07]$.

5 **(a)** The Gamma prior having mean 2 and coef. of variation 1 is Gamma$(1, 0.5)$. Hence $P(N(1)=0) = (0.5/1.5) = 1/3$.

(b) The information is included in posteriori density for Λ, which is Gamma$(1, 0.5+0.5)$.

(c) predictive probability

$$P(N(1)=0) = \int_0^\infty e^{-\lambda} f^{post}(\lambda)\, d\lambda = \int_0^\infty e^{-2\lambda}\, d\lambda = 1/2.$$

6 Since $\Theta_0^* + \Theta_1^* \ln(V) + \Theta_2^* \ln(H)$ is approximately normally distributed with mean $1.52 + 1.08\ln(3.7) + 1.5\ln(5.1) = 5.377$ and variance $0.64 + 0.25\ln(3.7)^2 + 0.16\ln(5.1)^2 = 1.4926$ consequently an approximative confidence interval for X is $[\exp(5.377 - 1.96\cdot\sqrt{1.493}), \exp(5.377 + 1.96\cdot\sqrt{1.493})] = [19.73, 2373]$. In practice only the lower bound is of interest.

One could also solve the problem by using Gauss formulas to compute $E[X^*]$, $V[X^*]$. Then 0.95 confidence interval would be $[E[X^*] - 1.96\sqrt{V[X^*]}, E[X^*] + 1.96\sqrt{V[X^*]}]$. For the computations of $E[X^*]$, $V[X^*]$ first define $h(x,y,z) = e^x\, 3.7^y\, 5.1^z$, then find partial derivatives $h_i(x,y,z)$. Then use formulas with x, y, z replaced by 1.52, 1.08, 1.5 .....