Sannolikhet, statistik och risk MVE300
Written examination - 21th May 2012, 14 - 18.

Allowed aids: The course book ”Probability and Risk Analysis - An Introduction for Engineers”, copies of the e-book, earlier versions of the book (compendium). Mathematical or statistical tables. The tables of formulæ of the course. Any calculator (not PC). Dictionaries for translation. Jour: Roza Maghsood 0737747320.

1 The water supply to a small town comes from two sources: from a reservoir and from pumping underground water. Suppose during a summer month, the amount of water available from each source is normally distributed $N(30, 9)$, $N(15, 16)$, million liters, respectively. Suppose the demand during the month is also variable and can be modeled as a normally distributed variable with mean 35 millions of liters and coefficient of variation 10%.

(a): Determine the probability Pf that there will be insufficient supply of water during the summer month. Assume that demand and supply vary independently. (7p)

(b): Determine the probability Pf that there will be insufficient supply of water during the summer months, under the assumption that demand and the total supply of water have correlation coefficient $\rho = -0.8$. (3p)

2 At a determination of the yield Z (unit: %) when oxidating ammonia in a converter, the following quantities are measured: X = the amount NH_3 in the reactant gas; Y = the amount NO in the product gas. The yield Z was calculated as

$$Z = Y(100/X - 1.25).$$

The following standard deviations have been found:

$$\sigma_X = 7,7 \cdot 10^{-2}, \quad \sigma_Y = 9,6 \cdot 10^{-2},$$

while expected values are $E[X] = 12.0$, $E[Y] = 13.5$. Suppose that X and Y are independent then calculate approximately the standard deviation of the yield. (10 p)

3 The yearly maximal water level in a harbor has been observed for 100 years. A preliminary analysis indicates that a Gumbel distribution is reasonable and that the observations can be considered independent. Computer software gives the ML estimates $a^* = 192$ and $b^* = 806$ of the scale and location parameter, respectively (unit: cm).

(a) Estimate the 50-year return level, h_{50}. (7p)

(b) Find the probability that the yearly maximum at least once during the following five years will exceed h_{50}. (3p)

4 In order to plan the emergency level in a regional hospital during weekends one was investigated the time distribution of heart attacks. For the last 202 registered heart attacks 36 occurred on Monday, 27 on Tuesday, 26 on Wednesday, 32 on Thursday, 26 on Friday, 30 on Saturday and 25 on Sunday. Test the hypothesis that time (day of the week) for heart attack is uniformly distributed. (20p)

5 An offshore windmill farm contains 60 wind turbines. The turbines are of the same model and are exposed to a similar wind climate. Suppose that the time of gearbox's failures in the farm form Poisson point process with unknown intensity λ. Manufacturers claim that expected lifetime of a gearbox is 15 years.

(a) One has estimated that a failure of a wind turbine gearbox costs, in average, 2 millions SEK (reparation costs, productions losses). Using information provided by the manufacturer estimate the average yearly cost for gearbox failures. (3p)

(b) Uncertainty in the value of the intensity λ can be described by means of a prior density $f^{prior}(\lambda)$. Choose an suitable prior density using the information provided by the manufacturer. (5p)

(c) Suppose that the first year two gearboxes have failed. Use this information to reevaluate the possible values of the intensity (find $f^{post}(\lambda)$). (5p)

(d) Use $f^{post}(\lambda)$ to compute an predictive yearly cost for gearbox reparations. (7p)

6 In a 12-year period a Swedish insurance company experienced 93 events with wind storm losses. Suppose that storm events form a stationary Poisson stream of events with intensity λ.

(a) Estimate the intensity λ of the Poisson stream of events. (3 p)

(b) Give a 90% confidence interval for λ. (7 p)

The insurance company is especially interested in the intensity (frequency) of storms with damages exceeding 100 million SEK, however not all storms are causing such large damages. The historical data indicates that the damages X, say, are independent and lognormally distributed. The mean and standard deviation of the logarithms of the damages are $m_X^* = 2$ and $\sigma_X^* = 1.32$ (milion SEK), i.e. $\ln X$ is estimated to be $\mathrm{N}(2, 1.32^2)$.

(c) What is the intensity of storms with damages exceeding 100 million SEK? (5 p)

(d) What is the probability that the company experience at least one storm event with a loss exceeding 100 million SEK in a time period of 5 years? (5 p)

Good luck!

Solutions: Written examination - 21 May 2012
Sannolikhet, statistik och risk MVE300.

1 Supply of water R is normally distributed with mean 45 million liter and variance $9+16=25$. The demand S is also normal with mean 35 million liters and variance 3.5^2.

(a) Pf$= P(R-S<0) = \Phi(-10/\sqrt{25+3.5^2}) = 0.0507$.

(b) $V(R-S) = 25+3.5^2+2\cdot(-1)\cdot 5\cdot 3.5\cdot(-0.8) = 65.25$ hence Pf$= \Phi(-10/\sqrt{65.25}) = \Phi(-1.24) \approx 0.11$

2 We will use Gauss' approximation. Introduce $g(x,y) = y(100/x - 1.25)$. Then

$$\frac{\partial g}{\partial x} = -\frac{100y}{x^2}, \quad \frac{\partial g}{\partial y} = \frac{100}{x} - 1.25$$

and Gauss' approximation gives

$$\mathsf{V}(Z) = \sigma_X^2\left(-\frac{100\mathsf{E}(Y)}{(\mathsf{E}(X))^2}\right)^2 + \sigma_Y^2\left(\frac{100}{\mathsf{E}(X)} - 1.25\right)^2 = 0.98.$$

It follows that $\mathsf{D}(Z) = 0.99$.

3 (a) Gumbel distribution:

$$\mathsf{P}(X \le x) = \exp(\mathrm{e}^{-(x-b)/a}), \quad x \in \mathbb{R}.$$

The 50-year return value h_{50} is obtained by solving for h_{50} in $\mathsf{P}(X > h_{50}) = 1/50$, which results in

$$h_{50} = b - a\ln(-\ln(1-1/50)).$$

An estimate h_{50}^* is then given by

$$h_{50}^* = b^* - a^*\ln(-\ln(1-1/50)) \doteq 1555 \text{ (cm)}$$

(b) Introduce N = "The mumber of years with yearly maximum exceeding h_{50}". From (a), we know that $\mathsf{P}(X \ge h_{50}) = 1/50$. Hence $N \in \mathrm{Bin}(5, 1/50)$ and we can compute

$$\mathsf{P}(N \ge 1) = 1 - \mathsf{P}(N=0) = 1 - (1-1/50)^5 \doteq 0.096 \approx 5/50.$$

4 Use χ^2-test. For all i, $np_i = 202/7$ and

$$\begin{aligned} Q &= \sum_{i=1}^{7} \frac{(n_i - np_i)^2}{np_i} = 1.7857 + 0.1207 + 0.2857 + 0.3457 \\ &\quad + 0.2857 + 0.0457 + 0.5207 \doteq 3.39. \end{aligned}$$

Since $\chi^2_{0.05}(7-1) \doteq 12.59$ then the hypothesis can not be rejected on significance level 0.05.

5 a) Average yearly cost for gearboxes reparations is $60 * 2/15 = 8$ millions SEK is predicted by manufacturer.

b) Basically the manufacture claims that $E[\Lambda] = 4$ years^{-4}. As in Example 6.2 one could choose exponential priors with expected value 4, viz. gamma pdf with parameter $a = 1$ and $b = 1/4$.

c) The posteriori density is gamma with parameters $a = 1 + 2$ and $b = 1/4 + 1$.

d) The predictive yearly costs is $E[\Lambda] \cdot 2 = 2\frac{3}{1.25} = 4.8$ millions SEK.

6 a) $\lambda^* = \frac{x}{n} = \frac{93}{12} = 7.75$

b) Two possible solutions

approx.: $\sigma_\varepsilon^* = \sqrt{\lambda^*/n} = 0.8$

$\lambda \in [\lambda^* - \lambda_{\alpha/2}\sigma_\varepsilon^*, \lambda^* + \lambda_{\alpha/2}\sigma_\varepsilon^*] = [7.75 - 1.64 \cdot 0.8, 7.75 + 1.64 \cdot 0.8] = [6.44, 9, 06]$

exact: $\lambda \in [\frac{\chi^2_{1-\alpha/2}(2n\lambda^*)}{2n}, \frac{\chi^2_{\alpha/2}(2n\lambda^*+2)}{2n}] = [6.7, 9.48]$

c) $\hat{\lambda}^* = \lambda^* \cdot \mathsf{P}(X > 100) = 7.75 \cdot (1 - \Phi(\frac{\ln(100)-2}{1.32})) = 0.19$ year^{-1}

d) $N \in \mathrm{Po}(\hat{\lambda} \cdot 5)$, $\mathsf{P}(N \geq 1) = 1 - \mathsf{P}(N = 0) = 1 - \exp(-\hat{\lambda}^* \cdot 5) = 0.61$