EXAMPLE OF WRITTEN EXAMINATION

1 The annual maximum sea-level, X (in meters), at Port Pirie, located in the southern part of Australia, has been recorded for a period of years. Gumbel distribution with the parameters a and b is used to model X. Based on data year 1923-1987, the ML-estimates are $a^* = 0.198$, $b^* = 3.87$.

(a) Estimate the hundred year sea-level, x_{100}, at Port Pirie. (5 p)

(b) Give an approximate 99% confidence interval for x_{100}. (5 p)

2 A new heating system has been installed. It is a modern system which can give a warning when reliability/efficiency is decreasing and then a major service is recommended. Suppose that service times form a stationary Poisson stream of events with an unknown intensity λ [year^{-1}]. Depending on quality of fuel and amount of needed energy, the intensity λ may vary. The dealer claims that on average, service is needed once in two years, i.e. $\lambda = 0.5$ year^{-1}.

(a) When ordering the system there is an option for a constant price of service, $c = 4000$ SEK. Since λ is intensity (frequency) of services needed, the average cost per year is $c\lambda$. What is the predicted yearly cost based on the information given by the dealer? (3 p)

(b) Suppose that you choose the option of constant service price and that service was needed once in the first 6 months of use. How does this information affect your predicted yearly service cost? (Hint: Use a suitable exponential prior, update and then compute $\mathsf{E}[c\Lambda]$.) (7 p).

3 In the field of life insurance standardized death-rates are often utilized. One example is the N-1963 standard according to which the death rate or "failure-intensity" for females is equal to

$$\lambda_{\mathrm{F}}(t) = \alpha + \beta \mathrm{e}^{(t-t_0)/c}, \qquad t > 0.$$

It was found the for a certain population the estimated parameters are $\alpha^* = 9 \cdot 10^{-4}$, $\beta^* = 4.4 \cdot 10^{-5}$, $c^* = 10.34$ and $t_0^* = 3$. Estimate the risk for death of 80 years old women in 10 years. (Hint: compute first probability that 80 years old women will live to be 90 years old.) (10 p)

4 Consider the data of failures of the air-condition system for a fleet of Boeing 720 plane. The time distance between failures (in hours) are

50	44	102	72	22	39	3	15	197	188	79	88
46	5	5	36	22	139	210	97	30	23	13	14

The hypothesis is that events of failures can be modelled as Poisson stream with constant intensity of failures denoted by λ. Test the hypothesis about the model (give the name of the test you are using for this purpose). If the test does not rejected the Poisson model estimate the intensity λ. (20 p)

5 In a 10-year period there were 57 events of large fires in industrial buildings of a certain type (similar sizes and utilization) in Scania region. In that period there were approximately 260 industial buildings of that type. Suppose that occurences of industry-fires follow Poisson process, with the same intensity for all buildings in that population, while losses in a fire X (SEK) are lognormal distributed, i.e. $\ln X \in \mathrm{N}(m_X, \sigma_X)$, with ML-estimates $m_X^* = 15$ and $\sigma_X^* = 1.32$.

(a) Estimate the intensity λ of the Poisson process. (5 p)

(b) What is the probability that the loss due to one fire in a building exceeds 10 million SEK, i.e. estimate $\mathsf{P}(X_i > 10^7)$? (5 p)

(c) Suppose that losses due to different fires are independent and identically distributed. What is the probability that a company that resides in a single industrial building experiences at least one fire with a loss exceeding 10 million SEK within next 10 years from now? (10 p)

6 Ekofisk is an oil field in the Norwegian sector of the North Sea. Discovered in 1969, it remains one of the most important oil fields in the North Sea. The present time horizon for the operation of Ekofisk is 2028. A challenging problem at Ekofisk has been the subsidence of the seabed. In 1986, deck structures of a number of platforms had to be elevated by 6 meters. When measured against the sea level at that time, it was 27 meters above the sea level. However, the subsidence today is 8.5 meters so the deck structure is now 18.5 meters above sea level. It is assumed that the annual maximum wave height, H (meters above sea level), is Gumbel distributed, with the estimated parameters $a^* = 1.95$ and $b^* = 5.2$.

(a) Estimate present yearly probability of wet deck, i.e. of waves reaching the platform deck structures in the next 12 months (assume that for this period the subsidence can be considered constant. (7 p)

(b) Estimate the Cornell safety index β_C for the risk of waves reaching the platform deck structures in the next 12 months. (3 p)

(c) Assume that the subsidence, S (in meters), of the seabed year 2028 can be modeled as the following function of a random variable X

$$S = 13.3 + 0.5\sqrt{X},$$

with $\mathsf{E}(X) = 1.5$ and $\mathsf{V}(X) = 0.5$. (The random variable X represents a certain parameter the nature of which is irrelevant for the problem.)
How much one need to increase the deck level to keep the same Cornell safety index for waves reaching the platform deck structures in year 2028?

(10 p)

Solutions:

1 Gumbel distribution is given by the following cumulative distribution function

$$F(x) = e^{-e^{-(x-b)/a}}, \ a > 0.$$

(a) By the definition of x_{100}, it is the solution in x of

$$F(x) = 99/100,$$

where $a = a^*$ and $b = b^*$. By a straightforward algebra

$$\begin{aligned} -e^{-(x-b^*)/a^*} &= \ln 99/100 \\ -(x-b^*)/a^* &\approx \ln 0.01 \\ (x-b^*)/a^* &\approx 4.61 \\ x &\approx 4.61 * a^* + b^* \\ x &\approx 4.78 \end{aligned}$$

Thus an estimated and approximated value the hundred year sea-level at Point Pirie is x_{100} is $x^*_{100} = 4.78$[m]. Alternatively, one can further approximate the ML estimator (see Section 10.3) by using

$$x^*_{100} = b^* + a^* \ln 100 = 3.87 + 0.198 * 4.6051 = 4.78181.$$

We observe that the both approximations yield the same value.

(b) For finding the confidence interval for x_{100}, we use the normal approximation to the estimation error distribution as given Section 10.3.1 of the Namely, the confidence interval is given by

$$\begin{aligned} x^*_{100} \pm \lambda_{0.005} \cdot a^* \sqrt{(1.11 + 0.61 * \ln^2 100 + 0.52 * \ln 100)/65} \\ \approx 4.78 \pm \lambda_{0.005} \cdot 0.1 \end{aligned}$$

or in other words $[4.5, 5.0]$[m].

2 Given λ, the number of failures X per year is following Poisson distribution with parameter λ, which has the mean $\mathsf{E}[X] = \lambda$.

(a) The average cost is $\mathsf{E}[c \cdot X] = c \cdot \mathsf{E}[X] = c \cdot \lambda = 2000$ SEK/year. We may also approach the problem through Bayesian methodology and assume the conjugate exponential prior for λ with the mean 0.5 which corresponds to Gamma(α, β) with $\alpha = 1$ and $\beta = 2$ so that $\mathsf{E}[\Lambda] = 1/\beta = 0.5$. This approach leads to the same answer $\mathsf{E}[c \cdot X] = \mathsf{E}[\mathsf{E}[c \cdot X|\Lambda]] = c \cdot \mathsf{E}[\Lambda] = 2000$ SEK/year.

(b) Our prior distribution for the problem as described above is Gamma(α, β) with $\alpha = 1$ and $\beta = 2$. because it is conjugate to the Poisson distribution, the posterior distribution of Λ is given by Gamma$(\alpha + x, \beta + t)$ =Gamma$(2, 2.5)$ hence using the mean value of Gamma distribution gives $\mathsf{E}[c\Lambda] = 4000\frac{2}{2.5} = 16000/5 = 3200$ SEK.

3 The standarized death-rates is another name for the failure intensity function. Let T denotes a life-time of a women. We are asked to compute the risk

$$\begin{aligned}\mathsf{P}(T < 90|T > 80) &= 1 - \mathsf{P}(T > 90)/\mathsf{P}(Y > 80) \\ &= 1 - R(90)/R(80) = 1 - \exp(-\int_{80}^{90} \lambda(t)\, dt).\end{aligned}$$

Given the estimated parameters, we compute the integral

$$\begin{aligned}\int_{80}^{90} \alpha^* + \beta^* e^{(t-3)/c^*}\, dt &= 10 * \alpha^* + \beta^* c^* e^{(87-77)/c^*} \\ &= 9 \cdot 10^{-3} + 4.4 \cdot 10.34 \cdot 10^{-5} \big(e^{87/10.34} - e^{77/10.34}\big) \\ &= 1.280636.\end{aligned}$$

Thus the estimated risk is approximately $1 - e^{-1.28} \approx 72\%$.

4 For the problem, we use Barlow-Proschan test that is specifically designed for testing Poisson model (using χ^2-test is also possible but not recommended for this problem). The test statistics (see Section 7.4 of the textbook) for the problem is

$$z = \sum_{k=1}^{n-1} \sum_{i=1}^{k} t_i / \sum_{k=1}^{n} t_k = \sum_{k=1}^{n-1} s_k/s_n,$$

where t_i are times between failures and s_k are failures time. Straightforward algebra gives s_k's as

50	94	196	268	290	329	332	347	544	732	811	899
945	950	955	991	1013	1152	1362	1459	1489	1512	1525	1539

Thus $z = 18245/1539 = 11.8551$. According to the test we would reject the model at the significance level $\alpha = 0.05$ if the value of z falls outside the interval

$$[23/2 - 1.96\sqrt{23/12}, 23/2 + 1.96\sqrt{23/12}] \approx [8.79, 14.21],$$

which is not the case, so we do not have evidence that using the Poisson model for these data is not appropriate.

The estimate of the intensity is then given as the reciprocal of the average of times between failures, i.e. $\lambda^* = 24/1539 \approx 0.0156$ per hour.

5 The problem combines two random factors: occurences of fires in the industrial buildings and losses due to these fires. In Part-a) we consider only the first factor, in Part b) we account only for the second factor, and finaly in Part c), we account for both.

(a) We assume that the we compute annual rate of fires and the obvious estimator (which is also the maximum likelihood estimator) of it is given as

$$\lambda^* = 57/(10 * 285) = 0.02.$$

(b) We use normal approximation to the distributon of the estimation error $\mathcal{E}$ that is discussed in Subsection 4.4.2 of the textbook. The

$$[5.7 - 1.96 * \sqrt{5.7/10}, 5.7 + 1.96 * \sqrt{5.7/10}] = [4.22, 7.17].$$

(c) Here we use the definition of lognormal random variable which states that $Y_i = \log X_i$ is normally distributed, so

$$\begin{aligned} \mathsf{P}(X_i > 10^7) &= \mathsf{P}(Y_i > 7 \log 10) \\ &\approx \mathsf{P}(Y_i > 16.12) = \mathsf{P}((Y_i - m_X)/\sigma_X > (16.12 - m_X)/\sigma_X) \\ &\approx \mathsf{P}(Z > 0.847) = 0.2, \end{aligned}$$

where in the last approximation we have used estimated values of m_X and σ_X, while Z is the standard normal variable.

(d) The risk is

$$1 - \exp(-t\lambda \cdot \mathsf{P}(X > 10^7)) \approx t\lambda \cdot \mathsf{P}(X > 10^7) = 0.02 * 0.2 * 10 = 0.04,$$

i.e. 4%.

6 The failure function is $h(H, S) = 27 - S - H$ where S is the subsidence for a given year.

(a) The failure function is $h(H, S) = 27 - S - H$, $S = 8.5$ hence $P(h(H, S) < 0) = P(H > 18.5) = 1 - e^{-e^{-(18.5-5.2)/1.95}} \approx 1/1000$.

(b) Cornells safety index $I = \mathsf{E}[h(H, S)]/\sqrt{\mathsf{V}(h(H, S))}$ and here $I = (18.5 - \mathsf{E}[H])/\sqrt{\mathsf{V}(h(H, S))} = 4.87$.

(c) Let x be the height one need to lift the deck. The failure function is now $h(H, S) = x + 27 - 13.3 - 0.5\sqrt{X} - H$. We use Gauss' approximation to compute the mean and variance of $h(H, S)$ and we obtain

$$\mathsf{E}[h(H, S)] \approx x + 13.7 - 5.2 - 1.95 * 0.5773 - 0.5\sqrt{1.5} = x + 6.76,$$

while the variance

$$\begin{aligned} \mathsf{V}(h(H, S) &= \mathsf{V}(H) + \mathsf{V}(S) \\ &\approx 1.95^2 * \pi^2/6 + 0.5^2 * (1/(4 * 1.5)) * 0.5 = 6.2757. \end{aligned}$$

The safety index $I \approx (x + 6.76)/\sqrt{6.276}$ should be equal to 4.87. Giving $x = -6.76 + 2.5 * 4.87 = 5.42$ meters.