UNIVERSITY OF LUGANO
PROF. CLAUDIO ORTELLI
PROF. DAVIDE LA VECCHIA

FACULTY OF INFORMATICS
SS 2011

Statistics

ALaRI Exam — 15 September 2011

- Duration: 2 hours and 30 minutes
- Open book exam
- Solve all exercises

Exercise 1

Suppose that for a certain device the instantaneous failure rate has the form

$$h(t) = a + bt + c\, sin(2\pi t).$$

where t denotes the time in years and a, b and c are given parameters such that $h \geq 0$.

1. What is the interpretation of $h(t)$ in general? Discuss this particular failure rate and give examples of devices which potentially could possess this kind of failure rate.

2. Derive the corresponding lifetime distribution function $F_X(t)$. Is it an increasing or decreasing failure rate distribution? Motivate your answer.

Exercise 2

We assume that the lifetime X of a particular device has a Gamma distribution. Recall that the Gamma distribution has density function given by

$$f(t) = \frac{t^{k-1} e^{-t/\lambda}}{(k-1)!\lambda^k},\ t \geq 0.$$

1. Compute the expectations $E(X)$ and $E(X^2)$. Show the computations!

2. Derive the variance of X.

3. The results of 10 independent realizations of X are summarized in the following table:

i	1	2	3	4	5	6	7	8	9	10
x_i	17.6	13.8	8.7	11.5	9	8.7	8	4.3	8.7	8.2

Tabella 1: Independent realizations of X.

For $k = 4$ compute the method of moments estimate of λ using

(a) the first moment of X, denoted $\hat{\lambda}_1$;

(b) the second centered moment (the variance) of X, denoted $\hat{\lambda}_2$.

Exercise 3

For the same problem as in Exercise 2

1. derive the Maximum-Likelihood estimator of λ, denoted $\hat{\lambda}_{ml}$ when k is known.

2. derive the variance of the Maximum-Likelihood estimator and show that the variance of $\hat{\lambda}_{ml}$ depends on the parameter λ.

3. Compute the Maximum-Likelihood estimate.

4. How would you estimate the variance of $\hat{\lambda}_{ml}$ once you have the estimate of point 3?

5. An important theorem in Maximum-Likelihood estimation theory says that the quantity

$$\sqrt{n}(\hat{\lambda}_{ml} - \lambda)$$

for large n is distributed according to the normal distribution $N(0, 1/I(\lambda))$ where the quantity $I(\lambda)$ is called the Fisher information and is equal to

$$I(\lambda) = E\left[\left(\frac{\partial}{\partial\lambda} log(f(X;\lambda))\right)^2\right]. \quad (1)$$

Under regularity conditions on the underlying density function the Fisher information can be computed by applying the following alternative formula

$$I(\lambda) = -E\left[\frac{\partial^2}{\partial\lambda^2} log(f(X;\lambda))\right] \quad (2)$$

(a) Compute $\frac{\partial^2}{\partial\lambda^2} log(f(X;\lambda))$.

(b) Compute its expectation and verify that

   i. for the Gamma distribution the equality between (1) and (2) holds;

   ii. the variance of $\sqrt{n}(\hat{\lambda}_{ml} - \lambda)$ corresponds (after the appropriate transformations) to the variance calculated in points 2 and 4.

6. Suppose that the estimate of $I(\lambda)$ is equal to 0.64 and that we want to test the null hypothesis $H_0 : \lambda = 2.5$ versus the alternative $H_A : \lambda \neq 2.5$ at a 5% significance level.

   (a) Is H_0 a simple or composite hypothesis? And H_A?

   (b) Perform the test and explain the conclusions. Is the null hypothesis rejected?

Exercise 4

We define X as the number of Bernoulli trials needed by a client to access in write mode a particular database table (the table may be locked by another client). X is distributed according to the geometric distribution

$$P(X = k) = (1 - p)^{k-1}p,\ k = 1, 2, \ldots$$

where p denotes the probability of success in one trial.

1. Compute the $z-$transform $G_X(z)$ of the random variable X.

2. Similarly to the exponential distribution, the geometric distribution is *memoryless. Given that the first success has not yet occurred, the conditional probability distribution of the number h of additional trials required to obtain the first success does not depend on the number of observed failures. In other words, show that the conditional distribution*

$$P(X = n + h \mid X > n) = (1 - p)^{h-1}p.$$

3. Show that if $X_1, \ldots, X_n$ are independent, geometrically distributed random variables with parameter p, then

$$X_{min} = min(X_1, \ldots, X_n)$$

   is also geometrically distributed with parameter p_{min}. Derive explicitly the relation between p and p_{min}. In particular

(a) Explain the following inequality

$$P(X_{min} > k) = \prod_{i=1}^{n} P(X_i > k).$$

(b) Compute explicitly as a function of p the probabilities

$$P(X_i \leq k) \text{ and } P(X_i > k).$$

(c) Compute the probability

$$P(X_{min} > k)$$

as a function of p and show that X_{min} is geometrically distributed with probability of success p_{min}. Derive explicitly p_{min} as a function of p. Show the computations!