MAP531: Statistics
Homework assignment to submit on Friday 13/10/2017

EXERCISE 1 -Poisson models for reliability

We are interested in the lifespan X of some electronic equipment. It is reasonable to consider that the lifespan is random and distributed from an Exponential distribution. That is to say its distribution has the following p.d.f. :

$$f_\lambda(x) = \lambda e^{-\lambda x} \mathbb{1}_{\mathbb{R}^+}(x) \,,$$

with $\mathbb{E}[X] = 1/\lambda$.

Yet, we ignore the value of the parameter λ of this distribution.

1. Write the statistical model.
2. Give the maximum likelihood estimator associated to $(X_1, ..., X_n)$ a $n-$sample of lifespans of those equipments.
3. Propose a maximum likelihood estimator of the parameter $\alpha = \mathbb{P}(X > t_0)$, where t_0 is a fixed time.
4. Which estimators of λ and α can you propose using the method of moments ?

We now suppose that the observations of the lifespans are obtained thanks to the following experiment. At time $t = 0$, there is an equipment on a test bed. When this one breaks down, it is immediately replaced (then the replacement duration is zero) by an identical equipment but new. And so on until time t_0. Let K be the number of breakdowns during the time interval $[0, t_0]$.

5. Compute the probability that K is equal to zero.
6. Let T_k be the time that elapsed until the k^{th} observed breakdown. Then $T_k = X_1 + ... + X_k$. It can be shown that the random variable T_k is Gamma $\Gamma(k, \lambda)$ distributed, with associated p.d.f. :

$$f(x; k, \lambda) = \frac{\lambda^k}{\Gamma(k)} x^{k-1} \exp(-\lambda x) \,, \tag{1}$$

   with expectation $k\lambda$ and variance $k\lambda^2$.
   Write the event $\{K = k\}$ thanks to the random variables T_k and X_{k+1}. Deduce that the distribution of K is a Poisson distribution, and give its parameter.
   Reminder : the Poisson distribution with parameter θ is defined by : $\forall\ k \in\ \mathbb{N}$,

$$\mathbb{P}(K = k) = \frac{\theta^k}{k!} \exp(-\theta) \,. \tag{2}$$

We assume that we realize n times this experience. Let $K_1, ..., K_n$ be the numbers of breakdowns observed in each time interval $[0, t_0]$.

7. Give the statistical model associated with these observations $K_1, ..., K_n$.
8. Give another maximum likelihood estimator of λ, this time based on the observations $K_1, ..., K_n$.
9. Which estimator of λ do we obtain if we use the method of moments in this model ?

EXERCISE 2 -Statistics helping farmers

A farmer owns a square field and wants to estimate its area. A statistician friend told him that when he measures a side of the field, the experimental error of the measurement is a random variable distributed from a centred normal distribution with variance σ^2. His first measurement of the side is $x_1 = 510$ metres. From this first measurement, he deduces an area of $s_1 = 26.01$ hectares. He measures the side a second time and finds $x2 = 490$ meters, so a value of the area $s2 = 24.01$ hectares. He gives up its measurements and thinks. He wonders :

Which estimator of the area should he choose : s_1, s_2, or another estimator combining the two measures, such that :

$$s_3 = x_1 x_2 = 24.99, \tag{3}$$

$$s_4 = \frac{s_1 + s_2}{2} = 25.01, \tag{4}$$

$$s_5 = \left(\frac{x_1 + X_2}{2}\right)^2 = 25\,. \tag{5}$$

Should he go on its measurements until he finds two identical results, or cleverly combine the n measurements to build an estimator alike s_4 or s_5 (generalised to the n measurements) ?

We offer to help the farmer to solve his problem.

1. Precise the considered model along with the function $\theta \mapsto g(\theta)$ we want to estimate.
2. Study the first five proposed estimators. In particular, compute their bias, variance and mean square error. (Note : if $X \sim \mathcal{N}(m, \sigma^2)$ then $Var(X^2) = 2(\sigma^4 + 2m^2\sigma^2)$).
3. From these computations, help the farmer to choose an estimator which seems to be better than the others.
4. Give estimators that generalise s_4 and s_5 in the case where the farmer did n measurements of a side of his field. Do the same study as in Question 2. for these two estimators.
5. Give the maximum likelihood estimator. Study it if it is different from the ones previously studied.

EXERCISE 3 -

We assume that the weight of a newborn is a random variable with standard deviation 0.5 kg. The average weight of neonates, who were born during one month in some hospital is reported. The report says that 49 children were born with an average weight of 3.6 kg.

1. Give a 95%-confidence interval for the average weight of a newborn in this hospital.
2. Compute the confidence level of an interval centred at 3.6 kg, whose length is 0.1 kg.