Statistical Theory: Exercise Sheet 1

September 25, 2018; 8:15-10:00 in MA 12.

Tomáš Rubín, tomas.rubin@epfl.ch

Exercise 1. If X is exponentially distributed with intensity λ, what is the distribution of $Y = \lfloor X \rfloor$? ($\lfloor x \rfloor$ stands for the integer part of x.)

Exercise 2. Suppose that $S \sim \operatorname{Exp}(\lambda)$, $C \sim \operatorname{Exp}(\gamma)$ are independent. Define $T = \min(S, C)$ and $D = \mathbf{1}_{[T=S]}$.

(a) Find the joint ditribution of T and D and their marginal distributions.

(b) Are T and D independent?

Exercise 3. Consider a random vector $(X, Z)^\mathsf{T}$. Let the marginal distribution of Z be exponential with parameter γ, i.e., with density

$$f_Z(z) = \gamma e^{-\gamma z} 1_{(0,\infty)}(z).$$

Suppose that the conditional distribution of X given $Z = z$ is Poisson with parameter λz, that is,

$$f_{X|Z}(x|z) = P(X = x|Z = z) = \begin{cases} \frac{(\lambda z)^x}{x!} e^{-\lambda z}, & x = 0, 1, \ldots, \quad z > 0, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the joint density $f_{X,Z}(x, z)$ (for all (x, z)).

(b) Compute the marginal (unconditional) distribution of X. Which known distribution is it?

(c) Find $\mathsf{E}[X|Z]$.

(d) Find $\mathsf{E}\, X$ using the formula $\mathsf{E}\, X = \mathsf{E}\{\mathsf{E}[X|Z]\}$.

(e) Find $\operatorname{var}[X|Z]$.

(f) Compute $\operatorname{var} X$ using the formula

$$\operatorname{var} X = \mathsf{E}\{\operatorname{var}[X|Z]\} + \operatorname{var}\{\mathsf{E}[X|Z]\}.$$

(g) Compute the conditional density $f_{Z|X}(z|x)$. Which known distribution is it?

Remark: The variable X can be interpreted as the value of a Poisson process $N(t)$ with intensity λ at random $\mathrm{Exp}(\gamma)$-distributed time $t = Z$ (independent of $N(\cdot)$), that is, $X \sim N(Z)$; also, X follows a Poisson model with an exponential frailty (random effect).

Exercise 4. Suppose that the observation X follows the model of the previous exercise. Is this model for X (parametrised by (λ, γ)) identifiable?

Exercise 5. Let the observations $Y_1, \ldots, Y_n$ satisfy the regression model

$$Y_i = \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + e_i,$$

where x_{ij} are some constants and e_i are independent variables with distribution $\mathrm{N}(0, \sigma^2)$. Give a necessary and sufficient condition on $x_{ij}, i = 1, \ldots, n, j = 1, \ldots, p$ for identifiability of this model (parametrised by $(\beta_1, \ldots, \beta_p, \sigma^2)$).

Exercise 6. Consider a young couple sharing an apartment, and the random variables:

- T_b … the waiting time until the couple have a baby,
- T_m … the waiting time until the couple move,
- T_d … the waiting time until the couple break up,
- T … the waiting time until the first of the above-mentioned events takes place,
- N … the number of people living in the apartment at time T:

$$N = \begin{cases} 3 & \text{if } T = T_b, \\ 0 & \text{if } T = T_m, \\ 1 & \text{if } T = T_d. \end{cases}$$

Suppose that the variables T_b, T_m, T_d are independent, and $T_b \sim \mathrm{Exp}(\beta)$, $T_m \sim \mathrm{Exp}(\mu)$, $T_d \sim \mathrm{Exp}(\delta)$.

(a) Find the distribution of T.

(b) Find the joint distribution of T and N.

(c) Find the distribution of N.

(d) Is the model for T parametrized by (β, μ, δ) identifiable? If not, re-parametrize the model (i.e. parametrize it by a function of the three parameters) so that it becomes identifiable.

(e) Is the model for T and N parametrized by (β, μ, δ) identifiable? If not, re-parametrize the model so that it becomes identifiable.