University of Paris 1 Pantheon - Sorbonne — 2020 - 2021
QEM FINAL EXAM — December 10 2020

Probability and Statistics (10h00-12h00)

NO DOCUMENTS, NO PHONE, NO COMPUTER, NO CALCULATOR

• In this exam, $\mathcal{N}(m, \sigma^2)$ denotes a Gaussian distribution with mean $m \in \mathbb{R}$ and variance $\sigma^2 \in \mathbb{R}_+^*$.

•For $\lambda > 0$ and $\alpha > 0$ we say that $X \hookrightarrow \Gamma(\alpha, \lambda)$ if the density of X is given by

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} I_{]0,+\infty[}(x).$$

We remind that in this case

$$E[X] = \frac{\alpha}{\lambda} \text{ and } Var[X] = \frac{\alpha}{\lambda^2}.$$

EX1, EX2 and EX3 *are mandatory*. Students have to choose *ONE* exercise among EX4 and EX5.

The BONUS questions are more difficult and of scale. Please do them only at the end !

Exercise 1 : (20 mins) Let $X = (X_1, X_2, X_3)$ be a Gaussian vector with mean zero and variance-covariance matrix $\begin{pmatrix} 3 & -2 & 1 \\ -2 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$.

a) Prove that X_2 and X_3 are independent.

b) Find the distribution of $X_1 + X_2$.

c) Compute $E[X_2 + X_3 \mid X_3]$.

Exercise 2 : (30 mins) Let (X, Y) be a pair of random variables having the density

$$f(x, y) = \frac{1}{4\pi} e^{-\frac{1}{2}(\frac{x^2}{2} - xy + y^2)}.$$

a) Prove that $\frac{x^2}{2} - xy + y^2 = \frac{1}{2}(x-y)^2 + \frac{y^2}{2} = (y - \frac{x}{2})^2 + \frac{x^2}{4}$.

b) Find the conditional density of Y given $X = x$. Compute $E[Y|X]$ and $E[Y^2|X]$.

c) Prove that X and $Y - \frac{X}{2}$ are independent. Find the distribution of $Y - \frac{X}{2}$.

d) Compute $Var[E[Y \mid X]]$, $E[Var[Y \mid X]]$ and $Var[Y]$. Is there a general link between these quantities ?

Exercise 3 : (40 mins) We consider n independent observations $(X_1, ..., X_n)$ of X following a $\Gamma(\alpha, \lambda)$ and we suppose that $\alpha = \alpha_0$ **is known**.

a) Find the maximum likelihood estimator (it is denoted by $\hat{\theta_n}$) of $\theta = \frac{1}{\lambda}$.

b) Is it the best unbiased estimator of θ in the sense of the mean squared error ?

c) (**BONUS**) Prove that $\hat{\theta_n}$ follows a $\Gamma(n\alpha_0, n\alpha_0\lambda)$.

d) Find the maximum likelihood estimator (it is denoted by $\hat{\lambda_n}$) of λ.

e) (**BONUS**) Compute the bias of $\hat{\lambda_n}$ when $\alpha_0 > 1$.

CHOOSE ONE EXERCISE AMONG EX4 and EX5

Exercise 4 : (30 mins) Let $X_1, ..., X_{n_1}$ and $Y_1, ..., Y_{n_2}$ be two independent samples from distributions $\mathcal{N}(\mu, \sigma_X^2)$ and $\mathcal{N}(\mu, \sigma_Y^2)$ respectively. We denote by $\overline{X_{n_1}}$ and $\overline{Y_{n_2}}$ the corresponding sample means.

a) For an estimator T of μ, prove that the mean squared error (MSE) fulfills

$$MSE(T) = Var[T] + Bias^2.$$

b) For $c \in [0,1]$, we define $T_c = c\overline{X_{n_1}} + (1-c)\overline{Y_{n_2}}$ as an estimator of μ. Find $c^* \in [0,1]$ that minimizes $MSE(T_c)$.

c) Doing the computations, prove that

$$MSE(T_{c^*}) \leq Min(MSE(\overline{X_{n_1}}), MSE(\overline{Y_{n_2}})).$$

d) What happens when $n_1 = n_2$ and $\sigma_X^2 = \sigma_Y^2$?

Exercise 5 : (30 mins) Let $X_1, ...X_n$ be a random sample from a $\mathcal{N}(\mu, 1)$ and consider testing for $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$.

Indication : We remind that $P(\mathcal{N}(0,1) \leq 1.96) = 0.975$, $P(\mathcal{N}(0,1) \leq 0.96) = 0.83147$ and $P(\mathcal{N}(0,1) \leq 2.96) = 0.99856$.

1) Find the LRT (likelihood ratio test) statistics for this problem.
2) When $n = 100$, define the rejection area to obtain a test of level (or size) 0.05.
3) When $n = 100$ and $\mu_0 = 0$, find the power function for $\mu = 0.1$ when the size is 0.05 .