Lecturer: Prof. Dr. Mete SONER
Coordinator: Yilin WANG

ETH Zürich

Exercise Series 7

Q1. Suppose two random variables X_1 and X_2 have a continuous joint distribution for which the joint p.d.f. is as follows:

$$f(x_1, x_2) = \begin{cases} 4x_1x_2 & \text{for } 0 < x_1, x_2 < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Calculate $Cov(X_1, X_2)$. Determine the joint p.d.f. of two new random variables Y_1 and Y_2, which are defined by the relations

$$Y_1 = \frac{X_1}{X_2} \text{ and } Y_2 = X_1X_2.$$

Let α and β be positive numbers. A random variable X has the *gamma distribution with parameters α and β* if X has a continuous distribution for which the p.d.f. is

$$f(x|\alpha, \beta) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} & \text{for } x > 0, \\ 0 & \text{for } x \leq 0. \end{cases}$$

Where Γ is the function defined as: for $\alpha > 0$,

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx.$$

Q2. Let X have the gamma distribution with parameters α and β.

(a) For $k = 1, 2, \ldots$, show that the k-th moment of X is

$$\mathbb{E}(X^k) = \frac{\Gamma(\alpha + k)}{\beta^k \Gamma(\alpha)} = \frac{\alpha(\alpha + 1) \cdots (\alpha + k - 1)}{\beta^k}.$$

What are $\mathbb{E}(X)$ and $Var(X)$?

(b) What is the moment generating function of X?

(c) If the random variables $X_1, \cdots, X_k$ are independent, and if X_i (for $i = 1, \ldots, k$) has the gamma distribution with parameters α_i and β, show that the sum $X_1 + \cdots + X_k$ has the gamma distribution with parameters $\alpha_1 + \cdots + \alpha_k$ and β.

Q3. Service Times in a Queue. For $i = 1, \cdots, n$, suppose that customer i in a queue must wait time X_i for service once reaching the head of the queue. Let Z be the rate at which the average customer is served. A typical probability model for this situation is to say that, conditional on $Z = z$, $X_1, \ldots, X_n$ are i.i.d. with a distribution having the conditional p.d.f. $g_1(x_i|z) = z \exp(-zx_i)$ for $x_i > 0$. Suppose that Z is also unknown and has the p.d.f. $f_2(z) = 2 \exp(-2z)$ for $z > 0$.

Table 1: Data for Q4.(b)

i	x_i	y_i
1	0.5	40
2	1.0	41
3	1.5	43
4	2.0	42
5	2.5	44
6	3.0	42
7	3.5	43
8	4.0	42

(a) What is the joint p.d.f. of $X_1, \ldots, X_n, Z$.

(b) What is the marginal joint distribution of $X_1, \ldots, X_n$.

(c) What is the conditional p.d.f. $g_2(z|x_1, \ldots, x_n)$ of Z given $X_1 = x_i, \ldots, X_n = x_n$? For this we can set $y = 2 + \sum_{i=1}^n x_i$.

(d) What is the expected average service rate given the observations $X_1 = x_1, \cdots, X_n = x_n$?

Q4. LEAST-SQUARES LINE.

(a) Let $(x_1, y_1), \ldots, (x_n, y_n)$ be a set of n points of $\mathbb{R}^2$ and x_is are not all the same. Show that the straight line defined by the equation $y(x) = \hat{\beta}_0 + \hat{\beta}_1 x$ that minimizes the sum of the squares of the vertical deviations of all the points from the line has the following slope and intercept, i.e. $(\hat{\beta}_0, \hat{\beta}_1)$ minimizes

$$I(\beta_0, \beta_1) := \sum_{i=1}^{n} (\beta_0 + \beta_1 x_i - y_i)^2$$

over all choices of $(\beta_0, \beta_1) \in \mathbb{R}^2$:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2},$$
$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x},$$

where $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ and $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$.
The minimizing line is called the *least-squares line*. Remark that the least-squares line passes through the point $(\bar{x}, \bar{y})$.

(b) Fit a straight line of the form $y = \beta_0 + \beta_1 x$ to these values by the method of least squares (with your calculator or Excel).