Exercise Sheet 6

1. Let X, Y be two random variables of a random sample Ω. The expectation $E(X)$, the variance $\operatorname{Var}(X)$ of X and the covariance $\operatorname{Cov}(X,Y)$ can be expressed by

$$E(X) = \sum_x xP(X = x) := \mu_X,$$
$$\operatorname{Var}(X) = E\left((X - \mu_X)^2\right),$$
$$\operatorname{Cov}(X,Y) = E\left((X - \mu_X)(Y - \mu_Y)\right).$$

   Given that $E(X + Y) = E(X) + E(Y)$ show that:

   (a) $\operatorname{Var}(X) = E(X^2) - \mu_X^2$.

   (b) $\operatorname{Var}(aX + b) = a^2 \operatorname{Var}(X), \quad a, b$ constants.

   (c) $\operatorname{Cov}(X,Y) = E(XY) - \mu_X\mu_Y$.

2. Estimate the parameter θ using the method of moments and the method of maximum likelihood for a density of the form

$$f(x) = (\theta + 1)x^\theta, \quad x \in [0, 1], \quad \theta > 0$$

   and random sample with variables:

$$\frac{1}{5}, \frac{2}{5}, \frac{1}{2}, \frac{7}{10}, \frac{4}{5}, \frac{4}{5}, \frac{9}{10}, \frac{9}{10}.$$

3. The density of the Gamma distribution $G(a, \rho)$ is given by

$$f(x) = \frac{a^\rho}{\Gamma(\rho)} x^{\rho-1} e^{-ax}, \quad \text{for} \quad x > 0 \quad \text{and} \quad a, \rho > 0.$$

   Find the maximum likelihood estimator of a for $G(a, 2)$.

4. Let X, Y be two random variables with $\operatorname{Var}(X) > 0$, $\operatorname{Var}(Y) > 0$. Consider a linear function

$$g(X) = aX + b, \quad a, b \quad \text{parameters}.$$

   The function g is called the best linear predictor if g estimates Y by minimizing the expectation

$$E\left((Y - g(X))^2\right).$$

   Compute the parameters a, b such that the mean square error is minimum.

5. Consider X, Y two 1×200 vectors with random values from the standard uniform distribution on the interval $(0,1)$. Implement in MATLAB an algorithm with inputs X, Y and output the graph of X, Y and the best linear predictor $g(X)$, as found in exercise 4.

6. Consider a finite population of people with size N, containing exactly K musicians. We choose randomly n people without replacement. For a random variable X, the probability mass function for the hypergeometric distribution is given by

$$P(X = x) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$

and for the binomial distribution,

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x},$$

where p is the probability of success. For $N \to \infty$, the hypergeometric distribution approaches the binomial distribution. Let $N = 75.000$ and $K = 500$. We choose 25 people randomly. Implement in MATLAB an algorithm to compute the probabilities, using both distributions, that

(a) at most one musician is selected.

(b) two or three musicians are selected.