E600 Mathematics

Fall Semester 2023

---

# Self-study Exercises 5

to be solved by September 1

*Topics: Probability Theory And Statistics*

---

## Exercise 1: Bayes' Rule and Probabilities

### a.) Conditional Probability

Given two events A and B. Verbally state what the following objects mean?
$P(A)$, $P(A \cup B)$, $P(A \cap B)$, $P(A|B)$. What is the definition of $P(A|B)$ in terms of the other objects?

### b.) Bayes' Rule

Apply Bayes' rule to solve the following problems:
*(i)* In an economy, $\frac{1}{3}$ of all firms shut down, after a new regulation is introduced that affects $\frac{1}{3}$ of all firms. Among the firms shutting down, $\frac{2}{3}$ were regulated. What is the probability that a regulated firm had to shut down? Is this probability higher than the probability that an unregulated firm has to shut down?

*(ii)* Suppose 10 % of the male labor force and 5 % of the female labor force are unemployed. Suppose that the labor force consists to 30 % of males and the remaining 70 % are female. What is the probability that a randomly drawn member of the labor force will be unemployed?

*(iii)* Balls are drawn sequentially from two urns. The first ball is drawn from urn 1, which contains 50 % white and 50 % black balls. Ball 2 is drawn from urn 2, which contains 30 % white and 70 % black balls. What is the probability that ball 2 is white, given that ball 1 was white? What is the probability of drawing at least one white ball?

## Exercise 2: Expected Value, Variance and CDF

Take a random variable $X$ with the following density as given:

$$f_X(x) = \frac{1}{4} I_{[0,4]}(x)$$

$I_{[0,4]}(x)$ is an indicator function that returns the value 1 if $x \in [0,4]$ and zero otherwise.

### (i) Cumulative Distribution Function

Derive the cumulative distribution function $F_X(t)$ for any real number t.
*Hint: Look up the definition and calculate the integral.*

### (ii) Expected Value

Calculate the expected value of X. Use the definition of the expected value from chapter 5.

### (iii) Variance

What is the variance of X?

## Exercise 3: Properties of Variance and Covariance

Take random variables $X, Y, Z$ as given. Take $a, b \in \mathbb{R}$. Take also the following properties of the expected value as given:

- $\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$ (Linearity)
- For two independent random variables $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

### (i) Properties of the Variance

Using the above properties of expected values, show the following: $Var(aX+b) = a^2 Var(X)$, for independent $X, Y$: $Var(X+Y) = Var(X) + Var(Y)$.

### (i) Properties of the Covariance

Using the above properties of expected values, show the following: $Cov(aX+b, Y) = aCov(X,Y)$, $Cov(X+Y, Z) = Cov(X,Z) + Cov(Y,Z)$, $Cov(X,Y) = Cov(Y,X)$, for independent $X, Y$: $Cov(X,Y) = 0$.
*Note that conversely, $Cov(X,Y) = 0$ does not imply that $X$ and $Y$ are independent!*

## Exercise 4: Joint, Marginal and Conditional Density

Take two jointly continuous random variables $X, Y$ with the following joint density as given:

$$\psi_{X,Y}(x,y) = \frac{1}{2\pi\sqrt{(1-\rho)^2}} exp\left(-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)\right)$$

($X$ and $Y$ are jointly normal with mean $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and variance matrix $\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$)

### (i) Marginal Density

Derive the marginal density $\psi_Y(y)$.
*Hint: Look up the definition and note that* $\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi(1-\rho^2)}} exp\left(-\frac{1}{2(1-\rho^2)}(x-\rho y)^2\right) = 1$. *Once you have transformed the definition in a way that the previous term drops out (multiplication with 1), look at the density of a normal random variable.*
*Note: the result we have derived here holds more generally: any two jointly normal random variables $X, Y$ will be normal themselves. The reverse, that any two normal random variables will be jointly normal, is not true.*

### (ii) Conditional Density

Derive the conditional density $\psi_{X|Y=y}(x|y)$ for any $y \in \mathbb{R}$. *Hint: Look up the definition for the conditional and the marginal density. Plug in and solve until you have the density of a univariate normal random variable again.*

### (iii) Conditional Expectation

Argue that the conditional expectation given by

$$\int_{\mathbb{R}} x \, \psi_{X|Y}(x|y) \, dx$$

is in fact a function of y.

## Exercise 5: Working with Conditional Expectations

To practice working with conditional expectations, we will now prove that the conditional expectation is the best function to "predict" a random variable, when the measure for the goodness of fit is the "Least Squared Error". To set up the problem, we are looking for a function $h$, such that $h(y)$ minimizes $\mathbb{E}((h(Y) - X)^2)$ for two random variables $X, Y$.

### (i) Step 1

Use the rules for conditional expectations from chapter 5 to show the following holds:

$$\mathbb{E}((X - \mathbb{E}(X|Y))(\mathbb{E}(X|Y) - h(y))|Y) = 0$$

### (ii) Step 2

Given step 1, show that

$$\mathbb{E}((X-h(Y))^2|Y) = \mathbb{E}((X-\mathbb{E}(X|Y))^2|Y) + \mathbb{E}((\mathbb{E}(X|Y)-h(y))^2|Y)$$

*Hint: Add a "smart zero": $+\mathbb{E}(X|Y)^2 - \mathbb{E}(X|Y)^2$ on the left side of the above equation.*

### (ii) Step 3

Given step 2, show that

$$\mathbb{E}((X-h(Y))^2) = \mathbb{E}((X-\mathbb{E}(X|Y))^2) + \mathbb{E}((\mathbb{E}(X|Y)-h(y))^2)$$

### (iii) Step 4

Argue that this implies that for any function h, this implies that

$$\mathbb{E}((X-h(Y))^2) \geq \mathbb{E}((X-\mathbb{E}(X|Y))^2)$$

*Note: This shows the result stated above. As a technicality, we have to restrict ourselves to measurable functions h. (If you don't know what this means, don't worry!)*

## Exercise 6: Convergence of Random Variables

A good place to start working with convergence theorems for random variables is applying them to show that an estimator is consistent, i.e., that an estimator converges in probability to the object we intend to estimate.

### (i) Consistency of OLS

Use the convergence theorems from chapter 5 to show that the OLS estimator $\hat{\beta}$ is consistent, i.e. that:

$$\hat{\beta} = (X'X)^{-1}X'y \xrightarrow{p} \beta = \mathbb{E}(X_iX_i')^{-1}\mathbb{E}(X_iy_i)$$

where X is the regressor matrix introduced in chapter 5 and $X_i$ is a column vector containing the regressors for one individual i.

*Hint: First consider $X'X$ and $X'y$ separately, write them in sum notation, insert a "smart one": $\frac{1}{n} * n$. Show the convergence of the two averages arising. Slutsky's theorem implies that for two series of random variables $X_n, Y_n$ such that $X_n \xrightarrow{p} C$ and $Y_n \xrightarrow{p} T$, $X_nY_n \xrightarrow{p} CT$ if $|C|, |T| < \infty$.*