# Econ435 – Financial Markets and the Macroeconomy
# Math Review

Mircea Trandafir

Intuitively, a *random variable* is a measure that can take on different values depending on the *state of nature* or, in our case, on the *state of the economy*. Each state of the economy occurs with a certain *probability*. The sum of all probabilities has to be equal to one.

In the rest of this handout, uppercase letters (from the end of the alphabet) will denote random variables, the corresponding lowercase letters will denote their possible values, and lowercase letters from the beginning of the alphabet will denote constants. Probabilities will always be denoted by $p_i$, where $i$ is the state of the economy.

The first part of the handout includes an intuitive and very informal introduction to the concepts that will be used during the class. A numerical example is given in the second part.

## Mean

The *mean*, or *expected value*, of a random variable gives what you would expect to get, "on the average", if you were to observe the random variable many times. The mean of a random variable (like all the other measures described later) is a *constant* expressed in the same unit of measurement as the possible values of the random variable (for example, the expected rate of return is expressed in percentage points, just like the rate of return is).

**Formula and notation:** $E(X) = \sum_i p_i x_i$

**Properties:**

(i) $E(a) = a$,

(ii) $E(aX + bY) = aE(x) + bE(Y)$ (the property of *linearity*),

(iii) $E[X - E(X)] = 0$ (a direct application of the two properties above).

## Measure of uncertainty: Variance, Standard Deviation

The *variance* is a measure indicating how different from each other are the possible values that the random variable can take. It usually makes sense to interpret it only together with the mean because it depends on the units of measurement (for example, if we express the same variable in percentage points or in decimal values, we would get two different values for the variance).

**Formula and notation:** $\sigma_X^2 = E\Big([X - E(X)]^2\Big) = \sum_i p_i [x_i - E(X)]^2$

**Properties:**

(i) $\sigma_a^2 = 0$,

(ii) $\sigma_{aX+bY}^2 = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab \cdot \sigma_{XY}^2$, where $\sigma_{XY}$ is the covariance defined below.

The square root of the variance is called *standard deviation* (the notation for it is $\sigma_X$) and is also a widely used measure of uncertainty. Unlike the variance, it is expressed in *the same* units of measurement as the mean (e.g., percentage points in the case of rates of return).

## Measures of comovement: Covariance, Correlation Coefficient

The *covariance* between two random variables measures the extent to which the two variables move together (usually as compared to their means). As in the case of the other concepts above, it depends on the units of measurement of *both* variables, so it is extremely important that both variables are expressed in the same unit or in comparable units (for example, comparing a variable expressed in thousands of dollars to a variable expressed in dollars might not give the best result).

A positive covariance means that the two variables "usually" move in the same direction: when one of them takes on a high value (above its mean), so is the other. A negative covariance implies that they "usually" move in opposite directions: when one of them is large, the other is small. A covariance equal to zero means that there is no apparent relationship between the two variables, in which case they are called *uncorrelated*.

**Formula and notation:** $\sigma_{XY} = E\Big([X - E(X)][Y - E(Y)]\Big) = \sum_i p_i [x_i - E(X)][y_i - E(Y)]$

**Properties:**

(i) $\sigma_{XY} = \sigma_{YX}$,

(ii) $\sigma_{Xa} = 0$,

(iii) $\sigma_{XX} = \sigma_X^2$,

(iv) $\sigma_{aX+bY,Z} = a\sigma_{XZ} + b\sigma_{YZ}$.

The covariance is hard to interpret on its own, exactly because it depends on the units of measurement. An better alternative is the *correlation coefficient*, which always lies between $-1$ and 1. Intuitively, its absolute value gives the "strength" of the relationship between the two variables, with 0 being the weakest (uncorrelated) and 1 the strongest (perfectly correlated), while its sign gives the direction of the relationship.

**Formula and notation:** $\rho_{XY} = \frac{Cov(X,Y)}{\sqrt{Var(X) \cdot Var(Y)}} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$

**Properties:**

(i) $\rho_{XY} = \rho_{YX}$,

(ii) $\rho_{Xa} = 0$,

(iii) $\rho_{XX} = 1$.

## Independence

Two random variables are *independent* if there is no relationship between how they take on their possible values. This implies that they are uncorrelated, i.e. that their covariance and correlation coefficient are both equal to zero. However, the opposite is not true: just because two variables are uncorrelated it does not mean that they are independent (statistical literature abounds with examples).

## Examples

Suppose we have the following data on the random variable X, as shown in the first columns of table 1 (where $i$ denotes states of the economy). The calculation of the mean (see below the table) is needed for filling in columns 5–7. The first row in each column shows how the values in that column were obtained.

Table 1: Calculating the mean and variance of a random variable

| $i$ | $p_i$ | $x_i$ | $p_i x_i$ | $x_i - E(X)$ | $[x_i - E(X)]^2$ | $p_i[x_i - E(X)]^2$ |
|---|---|---|---|---|---|---|
| Bad | 0.3 | 15 | $0.3 \cdot 15 = 4.5$ | $15 - 20 = -5$ | $(-5)^2 = 25$ | $0.3 \cdot 25 = 7.5$ |
| Moderate | 0.4 | 20 | 8 | 0 | 0 | 0 |
| Good | 0.3 | 25 | 7.5 | 5 | 25 | 7.5 |

The expected value of $X$ is obtained by adding up the numbers in the fourth column:

$$E(X) = 4.5 + 8 + 7.5 = 20.$$

After filling in the last three columns of the table, the variance is calculated as the sum of the numbers in the last column:

$$\sigma_X^2 = 7.5 + 0 + 7.5 = 15.$$

The standard deviation is just the square root of the variance:

$$\sigma_X = \sqrt{15} = 3.87.$$

We can perform a similar exercise for a second random variable, $Y$. The calculations are shown in table 2.

Table 2: Calculating the mean and variance of another random variable

| $i$ | $p_i$ | $y_i$ | $p_i y_i$ | $y_i - E(Y)$ | $[y_i - E(Y)]^2$ | $p_i[y_i - E(Y)]^2$ |
|---|---|---|---|---|---|---|
| Bad | 0.3 | 18 | $0.3 \cdot 18 = 5.4$ | $18 - 15 = 3$ | $3^2 = 9$ | $0.3 \cdot 9 = 2.7$ |
| Moderate | 0.4 | 12 | 4.8 | $-3$ | 9 | 3.6 |
| Good | 0.3 | 16 | 4.8 | 1 | 1 | 0.3 |

Again, the expected value is given by the sum of the numbers in the fourth column:

$$E(Y) = 5.4 + 4.8 + 4.8 = 15$$

and the variance by the sum of the numbers in the last column:

$$\sigma_Y^2 = 2.7 + 3.6 + 0.3 = 6.6,$$
$$\sigma_Y = \sqrt{6.6} = 2.57.$$

We can also calculate the covariance between these two random variables, as in the table below (note that we use some of the columns in tables 1 and 2, and only the last column is actually calculated in this table):

Table 3: Calculating the covariance between two random variable

| $i$ | $p_i$ | $x_i$ | $y_i$ | $x_i - E(X)$ | $y_i - E(Y)$ | $p_i[x_i - E(X)][y_i - E(Y)]$ |
|---|---|---|---|---|---|---|
| Bad | 0.3 | 15 | 18 | $-5$ | 3 | $0.3 \cdot (-5) \cdot 3 = -4.5$ |
| Moderate | 0.4 | 20 | 12 | 0 | $-3$ | 0 |
| Good | 0.3 | 25 | 16 | 5 | 1 | 1.5 |

The covariance between $X$ and $Y$ is given by the sum of the numbers in the last column:

$$\sigma_{XY} = -4.5 + 0 + 1.5 = -3.$$

Now we can calculate the correlation coefficient:

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{-3}{3.87 \cdot 2.57} = -0.30.$$

So, the two variables tend to move in opposite directions (they are negatively correlated), but their relation is not that strong.