Properties of the Expected Value

This lecture discusses some fundamental properties of the expected value operator.

Some of these properties can be proved using the material presented in previous lectures. Others are gathered here for convenience but can be fully understood only after reading the material presented in subsequent lectures.

It may be a good idea to memorize these properties as they provide essential rules for performing computations that involve the expected value.

Table of contents

If X is a random variable and a is a constant, then

$$E[aX] = aE[X]$$

This property has been discussed in the lecture on the Expected value. It can be proved in several different ways, for example, by using the transformation theorem or the linearity of the Riemann-Stieltjes integral.

Example

Let X be a random variable with expectation $E[X]$ and define $Y = aX$. Then, $E[Y] = aE[X]$.

If X1, X2, ..., Xn are random variables, then

$$E[X1 + X2 + \ldots + Xn] = E[X1] + E[X2] + \ldots + E[Xn]$$

See the lecture on the Expected value. The same comments made for the previous property apply.

Example

Let X and Y be two random variables with expected values $E[X]$ and $E[Y]$, and define $Z = X + Y$. Then, $E[Z] = E[X] + E[Y]$.

If $X1, X2, \ldots, Xn$ are random variables and $a1$, $a2$, ..., an are constants, then

$$E[a1X1 + a2X2 + \ldots + anXn] = a1E[X1] + a2E[X2] + \ldots + anE[Xn]$$

This can be trivially obtained by combining the two properties above (scalar multiplication and sum).

Consider $a = (a1, a2, \ldots, an)$ as the entries of a vector and $X = (X1, X2, \ldots, Xn)$ as the entries of a random vector.

Then, we can also write

$$E[a'X] = a'E[X]$$

which is a multivariate generalization of the Scalar multiplication property above.

Example

Let X and Y be two random variables with expected values $E[X]$ and $E[Y]$, and define $Z = aX + bY$. Then, $E[Z] = aE[X] + bE[Y]$.

Expected value of a constant

A perhaps obvious property is that the expected value of a constant is equal to the constant itself:

$$E[a] = a$$

for any constant a. This rule is again a consequence of the fact that the expected value is a Riemann-Stieltjes integral and the latter is linear.

Expected value of a product of independent random variables

Let X and Y be two random variables. In general, there is no easy rule or formula for computing the expected value of their product.

However, if X and Y are statistically independent, then

$$E[XY] = E[X]E[Y]$$

See the lecture on statistical independence.

Expected value of a function of a random variable

Let g be a non-linear function. In general,

$$E[g(X)] \neq g(E[X])$$

However, Jensen's inequality tells us that

$$E[g(X)] \geq g(E[X])$$

if g is convex and

$$E[g(X)] \leq g(E[X])$$

if g is concave.

Example

Since $g(x) = x^2$ is a convex function, we have

$$E[X^2] \geq (E[X])^2$$

Expected value of a random matrix

Let X be a random matrix, that is, a matrix whose entries are random variables.

If A is a matrix of constants, then

$$E[A + X] = A + E[X]$$

This is easily proved by applying the linearity properties above to each entry of the random matrix X.

Note that a **random vector is just a particular instance of a random matrix**. So, if X is a **random vector** and a is a **vector of constants**, then

$$E[a + X] = a + E[X]$$

Example

Let X be a random vector such that its two entries $X1$ and $X2$ have expected values $E[X1]$ and $E[X2]$. Let a be the following constant vector:

$$a = [1, 2]$$

Define

$$Y = a + X = [1 + X1, 2 + X2]$$

Then,

$$E[Y] = [1 + E[X1], 2 + E[X2]]$$

Multiplication of a constant matrix and a random matrix

Let X be a random matrix.

If A is a matrix of constants, then

$$E[AX] = AE[X]$$

If B is a a matrix of constants, then

$$E[XB] = E[X]B$$

These are immediate consequences of the linearity properties above.

By iteratively applying these properties, if A is a matrix of constants and B is a a matrix of constants, we obtain

$$E[AXB] = AE[X]B$$

Example

Let X be a random vector such that

$$E[X] = [E[X1], E[X2]]$$

where $X1$ and $X2$ are the two components of X. Let A be the following matrix of constants:

$$A = [[1, 2], [3, 4]]$$

Define

$$Y = AX = [X1 + 2X2, 3X1 + 4X2]$$

Then,

$$E[Y] = AE[X] = [E[X1] + 2E[X2], 3E[X1] + 4E[X2]]$$

Expectation of a positive random variable

Let X be an integrable random variable defined on a sample space Ω.

Let $X(\omega) \geq 0$ for all $\omega \in \Omega$ (i.e., X is a positive random variable).

Then,

$$E[X] \geq 0$$

Intuitively, this is obvious. The expected value of X is a weighted average of the values that X can take on. But X can take on only positive values. Therefore, also its expectation must be positive. Formally, the expected value is the Lebesgue integral of X, and can be approximated to any degree of accuracy by positive simple random variables whose Lebesgue integral is positive. Therefore, also the Lebesgue integral of X must be positive.

Monotonicity

Let X and Y be two integrable random variables defined on a sample space Ω.

Let $X(\omega)$ and $Y(\omega)$ be such that

$$X(\omega) \leq Y(\omega)$$

almost surely. In other words, there exists a zero-probability event A such that

$$X(\omega) \leq Y(\omega) \forall \omega \in A^c$$

Then,

$$E[X] \leq E[Y]$$

Proof

Let A be a zero-probability event such that $X(\omega) \leq Y(\omega) \forall \omega \in A^c$. First, note that

$$E[X] = E[X * 1A] + E[X * 1Ac]$$

where $1A$ is the indicator of the event A and $1Ac$ is the indicator of the complement of A. As a consequence, we can write

$$E[Y] - E[X] = E[Y * 1A] + E[Y * 1Ac] - E[X * 1A] - E[X * 1Ac]$$
$$= E[(Y - X) * 1A] + E[(Y - X) * 1Ac]$$

By the properties of indicators of zero-probability events, we have

$$E[(Y - X) * 1A] = 0$$

Thus, we can write

$$E[Y] - E[X] = E[(Y - X) * 1Ac]$$

Now, when $\omega \in A$, then $1Ac(\omega) = 0$ and $(Y(\omega) - X(\omega)) * 1Ac(\omega) = 0$. On the contrary, when $\omega \in Ac$, then $1Ac(\omega) = 1$ and $Y(\omega) - X(\omega) \geq 0$. Therefore, $(Y(\omega) - X(\omega)) * 1Ac(\omega) \geq 0$ for all $\omega \in \Omega (i.e., (Y - X) * 1Ac$ is a positive random variable). Thus, by the previous property (expectation of a positive random variable), we have

$$E[(Y - X) * 1Ac] \geq 0$$

which implies

$$E[Y] - E[X] \geq 0$$

By the linearity of the expected value, we get

$$E[Y - X] = E[Y] - E[X] \geq 0$$

Therefore,

$$E[Y] \geq E[X]$$

Exercises

Below you can find some exercises with explained solutions.

Exercise 1

Let X and Y be two random variables, having expected values:

$$E[X] = 2$$

$$E[Y] = -4$$

Compute the expected value of the random variable defined as follows:

$$Z = 3X - Y$$

Solution

Using the linearity of the expected value operator, we obtain

$$E[Z] = 3E[X] - E[Y] = 3 * 2 - (-4) = 10$$

Exercise 2

Let X be a random vector such that its two entries $X1$ and $X2$ have expected values

$$E[X1] = 8$$

$$E[X2] = 2$$

Let A be the following matrix of constants:

$$A = [[1, 5], [3, 4]]$$

Compute the expected value of the random vector defined as follows:

$$Y = AX$$

Solution

The linearity property of the expected value applies to the multiplication of a constant matrix and a random vector:

$$E[Y] = AE[X] = [[1, 5], [3, 4]] * [8, 2] = [18, 32]$$

Exercise 3

Let X be a matrix with random entries, such that all its entries have expected value equal to 1.

Let a be the following constant vector:

$$a = [2, 4]$$

Compute the expected value of the random vector defined as follows:

$$Y = Xa$$

Solution

The linearity property of the expected value operator applies to the multiplication of a constant vector and a matrix with random entries:

$$E[Y] = E[X]a = [2, 4]$$