Summation, Expectation, Variance, Covariance, and Correlation

2018-08-23

Assume the X and Y are random variables and c is a constant, such that:

$$
\begin{aligned}
X &= \{x_1, x_2, x_3, \ldots, x_n\} \\
Y &= \{y_1, y_2, y_3, \ldots, y_n\} \\
c &= \{c_1, c_2, c_3, \ldots, c_n\} \qquad \text{where } c_1 = c_2 = c_3 = \ldots = c_n
\end{aligned}
$$

The mean of these random (and constant) variables is denoted as the expected value, namely, $\mathbb{E}(X)$, $\mathbb{E}(Y)$, and $\mathbb{E}(c)$.

Formula for Variance

One very useful measure that we will work with a lot in the course is the variance. Here are several formulas to compute the variance of a random variable, X. We denote the variance of X using σ^2_X or $\mathrm{Var}(X)$. The most common formula for variance is:

$$
\sigma^2_X = \mathrm{Var}(X) = \frac{\sum_{i=1}^n \bigg(X_i - \mathbb{E}(X)\bigg)^2}{n}
$$

We can also compute variance as an expected value of the squared mean deviations:

$$
\sigma^2_X = \mathrm{Var}(X) = \mathbb{E}\Bigg(\big[X_i - \mathbb{E}(X)\big]^2\Bigg)
$$

Lastly, it can sometimes be helpful to express the variance as the difference between the expected value of X^2 and the squared expected value of X:

$$
\sigma^2_X = \mathrm{Var}(X) = \mathbb{E}(X^2) - \big[\mathbb{E}(X)\big]^2
$$

Lastly, we note that the standard deviation is the square root of the variance:

$$
\sigma_X = \sqrt{\sigma^2_X} = \sqrt{\mathrm{Var}(X)} = \mathrm{SD}(X)
$$

Formula for Covariance

Another useful measure that we will be working with in the course is the covariance. We denote the covariance between X and Y using σ_{XY} or $\text{Cov}(X, Y)$. The most common formula for covariance is:

$$\sigma_{XY} = \text{Cov}(X, Y) = \frac{\sum_{i=1}^{n} \Big(X_i - \mathbb{E}(X)\Big)\Big(Y_i - \mathbb{E}(Y)\Big)}{n}$$

The covariance can also be expressed as an expectation:

$$\sigma_{XY} = \text{Cov}(X, Y) = \mathbb{E}\Bigg(\big[X - \mathbb{E}(X)\big]\big[Y - \mathbb{E}(Y)\big]\Bigg)$$

Lastly, we can also express the covariance as a difference of expectations.

$$\sigma_{XY} = \text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

Formula for Correlation Coefficient

The correlation coefficient is a standardized covariance value. We denote the correlation between X and Y using ρ_{XY} or $\text{Cor}(X, Y)$. The most common formula for correlation is:

$$\rho_{XY} = \text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}$$

Rules for Working with Sums

The sum of X is defined as,

$$\sum_{i=1}^{n} X_i = x_1 + x_2 + x_3 + \ldots + x_n$$

To keep the notation simpler, we will just denote this as $\sum X$.

Rule 1: When a summation is itself a sum or difference, the summation sign may be distributed among the separate terms of the sum. That is:

$$\sum(X + Y) = \sum X + \sum Y$$

Rule 2: The sum of a constant, c, is n times the value of the constant.

$$\sum(c) = nc$$

Rules for Working with Expectations (Means)

Rule 1: The expectation of a constant, c, is the constant.

$$\mathbb{E}(c) = c$$

Rule 2: Adding a constant value, c, to each term in a random variable, X, increases the expected value (or mean) of X by the constant.

$$\mathbb{E}(X + c) = \mathbb{E}(X) + c$$

Rule 3: Multiplying a random variable, X, by a constant value, c, multiplies the expected value (or mean) of X by that constant.

$$\mathbb{E}(cX) = c\bigg(\mathbb{E}(X)\bigg)$$

Rule 4: The expected value (or mean) of the sum of two random variables, X and Y is the sum of the expected values (or means). This is also known as the *additive law of expectation*.

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$$

Rules for Working with Variances

Rule 1: The variance of a constant, c, is zero.

$$\text{Var}(c) = 0$$

Rule 2: Adding a constant value, c, to a random variable, X does not change the variance of X.

$$\text{Var}(X + c) = \text{Var}(X)$$

Rule 3: Multiplying a random variable, X by a constant, c increases the variance of X by the square of the constant.

$$\text{Var}(cX) = c^2 \times \text{Var}(X)$$

Rule 4: The variance of the sum of two random variables, X and Y is equal to the sum of their variances and the covariance between them.

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + \mathrm{Cov}(X,Y)$$

Rules for Working with Covariances

Rule 1: The covariance of two constants, c and k, is zero.

$$\mathrm{Cov}(c,k) = 0$$

Rule 2: The covariance of two *independent* random variables is zero.

$$\mathrm{Cov}(X,Y) = 0$$

Rule 3: The covariance is a combinative.

$$\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$$

Rule 4: The covariance of a random variable, X, with a constant, c is zero.

$$\mathrm{Cov}(X, c) = 0$$

Rule 5: Adding a constant to either or both random variables does not change their covariances.

$$\mathrm{Cov}(X + c, Y + k) = \mathrm{Cov}(X, Y)$$

Rule 6: Multiplying a random variable by a constant multiplies the covariance by that constant.

$$\mathrm{Cov}(cX, kY) = c \times k \times \mathrm{Cov}(X, Y)$$

Rule 7: The additive law of covariance holds that the covariance of a random variable with a sum of random variables is just the sum of the covariances with each of the random variables.

$$\mathrm{Cov}(X + Y, Z) = \mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z)$$

Rule 8: The covariance of a variable with itself is the variance of the random variable.

$$\mathrm{Cov}(X, X) = \mathrm{Var}(X)$$

Rules for Working with Correlation Coefficients

Rule 1: Adding a constant to a random variable does not change their correlation coefficient.

$$\mathrm{Cor}(X + c, Y + k) = \mathrm{Cor}(X, Y)$$

Rule 2: Multiplying a random variable by a constant does not change their correlation coefficient.

$$\mathrm{Cor}(cX, dY) = \mathrm{Cor}(X, Y)$$

Rule 3: Because the square root of the variance is always positive, the correlation coefficient can be negative only when the covariance is negative. This implies that:

$$-1 \leq \mathrm{Cor}(X, Y) \leq 1$$