Esssentials of Applied Data Analysis IPSA-USP Summer School 2017

Expected Value

Leonardo Sangali Barone
leonardo.barone@usp.br

jan/17

Expectation and mean value of a discrete random variable

Expectation and a simple game

Imagine a game where you toss a coin and get $1 if the result is head and 0 if the result is tail. How much should you expect to earn?

The expected value of a discrete random variable can be easily obtained by summing each result multiplied by probability of that result ocurring.

$$E[X] = 0.5 * 0 + 0.5 * 1 = 0.5$$

Now, imagine a game where you roll a dice and you can get $1 times the number you get on the dice. How much should you expect to earn?

$$E[X] = \frac{1}{6} * 1 + \frac{1}{6} * 2 + \frac{1}{6} * 3 + \frac{1}{6} * 4 + \frac{1}{6} * 5 + \frac{1}{6} * 6 + \frac{1}{6} * 1 = 3.666$$

Expectation and mean value of a discrete random variable

In more general term, the expectation or mean of a discrete random variable is:

$$E[X] = \sum_{i=1}^{n} x_i * P(X = x_i) = \sum_{i=1}^{n} x_i * f(x_i)$$

where x_i is an occurence of the variable X and $f(x_i)$ is the probability mass function (the probability that x_i will occur).

Note that, since the set of all x_i is the set of all possible values for X, then

$$E[X] = \sum_{i=1}^{n} P(X = x_i) = \sum_{i=1}^{n} f(x_i) = 1$$

Expectation and mean value of a discrete random variable

Example: (Made-up) survey with 2000 respondents in 2014 Brazilian presidential elections.

Candidate (x)	# of respondents	$P(X = x)$
Dilma	800	0.40
Aécio	500	0.25
Marina	400	0.20
Other/Null/DK	500	0.25
Total	2000	1.00

Can we calculate $E[X]$?

Expectation and mean value of a discrete random variable

When all $P(X = x_i)$ is the same for every x_i, we can simplify the expression of $E[X]$ to:

$$E[X] = \frac{1}{n} \sum_{i=1}^{n} x_i = \frac{x_1 + x_2 + ... + x_n}{n}$$

which is what we normally do to calculate avareges in daily life.

Expectation and mean value of a discrete random variable

Some properties of the mean:

$$E[a * X] = a * E[X]$$

$$E[X + b] = E[X] + b$$

So what? Well, if you multiply a variable by a number (a) to generate a new variable, the mean of the new variable is the mean of the old variable times a.

Also, if you sum a quantity b to a random variable, the mean of the result variable will be the mean of the original variable plus b.

Let's try it later on an statistical software!

Expectation and variance of a discrete random variable

Another important quantity of a random variable is the variance. The name is self-explanatory: the variance measures how spread-out a variable is.

Expectation and variance of a discrete random variable

The variance of a random variable is also an expectation:

$$Var[X] = \sum_{i=1}^{n} [x_i - E[X]]^2 * P(X = x_i) = \sum_{i=1}^{n} [x_i - E[X]]^2 * f(x_i)$$

where x_i is an occurence of the variable X, E[X] is the expected value of X and $f(x_i)$ is the probability mass function (the probability that x_i will occur).

Expectation and variance of a discrete random variable

Coin game (where heads pays off \$1 and tails pays off \$0):

$$Var[X] = \sum_{i=1}^{n} [x_i - E[X]]^2 * f(x_i) = [0.5 - 0.5]^2 * 0 + 0.5 * 1 = 0.5$$

Expectation and variance of a discrete random variable

Some properties of the variance:

$$Var[a * X] = a^2 * E[X]$$

$$Var[X + b] = Var[X]$$

So what? Well, if you multiply a variable by a number (a) to generate a new variable, the variance of the new variable is the variance of the old variable times a^2.

Also, if you sum a quantity b to a random variable, the variance of the result variable will equal to the variance of the original variable.

Let's try it later on an statistical software!

Expectation and variance of a discrete random variable

Some properties of the variance:

$$Var[a * X] = a^2 * E[X]$$

$$Var[X + b] = Var[X]$$

So what? Well, if you multiply a variable by a number (a) to generate a new variable, the variance of the new variable is the variance of the old variable times a^2.

Also, if you sum a quantity b to a random variable, the variance of the result variable will equal to the variance of the original variable.

Let's try it later on an statistical software!

Expectation, mean, variance and standard deviation

Notation:

$$E[X] = \mu[X] = \mu$$

$$Var[X] = \sigma^2[X] = \sigma^2$$

The standard deviation (σ) of a variable is

$$\sigma = \sqrt{Var[X]} = \sqrt{\sigma^2}$$

Another way to calculate the variance is simply doing:

$$Var[X] = E[X^2] - (E[X])^2$$