

## The Expected Value and Variance of an Average of IID Random Variables

This is an outline of how to get the formulas for the expected value and variance of an average. Since most of the statistical quantities we are studying will be averages it is very important you know where these formulas come from. Below I will carefully walk you through each step, but you should absolutely work through this on your own!

Suppose we are looking at  $n$  independent and identically distributed random variables,  $X_1, X_2, \dots, X_n$ . Since they are iid, each random variable  $X_i$  has to have the same mean, which we will call  $\mu$ , and variance, which we'll call  $\sigma^2$ . Another way to say this is:

$$\mathbb{E}(X_i) = \mu \qquad \text{Var}(X_i) = \sigma^2 \qquad \text{for all } i = 1, 2, \dots, n$$

Let's suppose we want to look at the average value of our  $n$  random variables:

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} = \left(\frac{1}{n}\right)(X_1 + X_2 + \dots + X_n)$$

The last equality on the right is there to emphasize that the average is equal to the sum of all the  $X$ 's multiplied by  $\frac{1}{n}$ . We want to find the expected value and variance of the average,  $\mathbb{E}(\bar{X})$  and  $\text{Var}(\bar{X})$ . The easiest way to do this is in three steps.

Step 1. First recognize that the average equals  $\frac{1}{n}$  times the sum. Now look at slide 12 of notes 3: According to our linear formulas, when we multiply a random variable by a constant, the mean gets multiplied by the same constant and the variance gets multiplied by that constant *squared*. So we have

$$\begin{aligned} \mathbb{E}(\bar{X}) &= \mathbb{E}\left[\left(\frac{1}{n}\right)(X_1 + X_2 + \dots + X_n)\right] = \frac{1}{n}\mathbb{E}(X_1 + X_2 + \dots + X_n) \\ \text{Var}(\bar{X}) &= \text{Var}\left[\left(\frac{1}{n}\right)(X_1 + X_2 + \dots + X_n)\right] = \left(\frac{1}{n}\right)^2 \text{Var}(X_1 + X_2 + \dots + X_n) \end{aligned}$$

Step 2. Now we know how to deal with the expected value and variance of a sum. Look at notes 3, slide 30. The expected value of a sum is *always* the sum of the expected values:

$$\mathbb{E}(X_1 + X_2 + \dots + X_n) = \mathbb{E}(X_1) + \mathbb{E}(X_2) + \dots + \mathbb{E}(X_n) = \mu + \mu + \dots + \mu = n\mu$$

And since the  $X$ 's are independent, we the variance of the sum is the sum of the variances:

$$\text{Var}(X_1 + X_2 + \dots + X_n) = \text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_n) = \sigma^2 + \sigma^2 + \dots + \sigma^2 = n\sigma^2$$

Notice that because the variables are identically distributed all the means (and variances) have to be the same, so we are just adding  $\mu$  together  $n$  times (and similarly for  $\sigma^2$ ).

Now just put the two steps together:

$$\begin{aligned} \mathbb{E}(\bar{X}) &= \frac{1}{n}\mathbb{E}(X_1 + X_2 + \dots + X_n) = \left(\frac{1}{n}\right)(n\mu) = \mu \\ \text{Var}(\bar{X}) &= \left(\frac{1}{n}\right)^2 \text{Var}(X_1 + X_2 + \dots + X_n) = \left(\frac{1}{n}\right)^2 (n\sigma^2) = \frac{\sigma^2}{n} \end{aligned}$$