# Key concepts for expectations and variances

⇨ *The <u>expected value</u> of X is defined to be the weighted average of all the possible values, where the weight assigned to each value is its probability:*

$$E(X) = 0P(X=0) + 1P(X=1) + 2P(X=2) + ... + nP(X=n)$$

$$= (0\ w_0) + (1\ w_1) + (2\ w_2) + ...+ (n\ w_n)$$

⇨ The expected value of *X*, is also called "the mean of *X* " or the "population mean."

⇨ When the expected value of an estimate equals the quantity that is being estimated, the estimate is said to be **<u>unbiased</u>**.

⇨ The <u>expected</u> value of *X* is not the same thing as the <u>most probable</u> value of *X*. That is, E(*X*) is not the value of *X* that has the greatest probability of occurrence.

The <u>variance</u> of a random variable, *X*, is defined as

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

That is, the variance is the expected value of the squared distance between *X* and its mean, E(*X*).

The <u>standard</u> <u>deviation</u> of *X* is the square root of its variance,

$$SD(X) = \sqrt{Var(X)}$$

⇨ Variances and standard deviations can <u>never</u> be negative.

⇨ **For any *constants* a and b, and any *random variable X***

$$\mathbf{E(\ a\mathit{X} + b\ ) = aE(\ \mathit{X}\ ) + b}$$

$$\mathbf{Var(\ a\mathit{X} + b\ ) = a^2\ Var(\ \mathit{X}\ )}$$

$$\mathbf{SD(\ a\mathit{X} + b\ ) = |a|\ SD(\ \mathit{X}\ )}$$

Key concepts for expectations and variances

⇨ If $X_1, X_2, ..., X_n$ are independent random variables with common mean μ and variance, $\sigma^2$, then for their average,

$$E(\overline{X}_n) = \mu$$

$$Var(\overline{X}_n) = \frac{\sigma^2}{n}$$

The standard deviation of the mean, called the standard error, is

$$SE = SD(\overline{X}_n) = \frac{\sigma}{\sqrt{n}}$$

AND the Key result:

$$\overline{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right) = N\left(E(X_i), \frac{Var(X_i)}{n}\right)$$

which tells us that the distribution of the sample mean is normal (with this mean and variance).

⇨ This result depends on the assumption that each individual observation is normal. Later we will learn that this result holds approximately in large samples regardless of the distribution generating the individual observations.