CHARACTERISTICS OF PROBABILITY DISTRIBUTIONS

(Appendix B, section B.1 and B.3)

Reminder

Discrete random variable: a random variable that takes on *countably many* values (they can be finite or infinite, but we can count them);

Continuous random variable: a random variable that takes on *uncountably many* values.

E.g. The computer time (in seconds) required to process a certain program

I. Measures of central tendency

1. The expected value (population mean)

Notation: E(X) or $\boldsymbol{\mu_X}$

a) Definition

The **expected value** of a random variable X is *a weighted average of all possible values* of X, weighted by the probability of each outcome occurring. In other words, each possible value that the random variable can assume is multiplied by its probability of occurring, and the resulting products are then added together to find the expected value.

Example: rolling a die

E(X)=1/6*1+1/6*2+1/6*3+1/6*4+1/6*5+1/6*6=(1+2+3+4+5+6) * 1/6 =3.5

Two things to note:

1) The *expected value need not be a possible outcome of the random variable*: it is evident that we cannot get 3.5 as an outcome of a die roll.

2) Intuitively, *the expected value is simply the average outcome*: if we perform a large number of die-rolling experiments (e.g. we roll a die 100,000 times), on average, we are going to obtain 3.5.

b) Properties

b1) if C=const $\Rightarrow$ E(C)=C

b2) for A, B=const $\Rightarrow$ E(AX+BY) =AE(X) + BE(Y)

In other words: *the expected value passes through linear functions*!

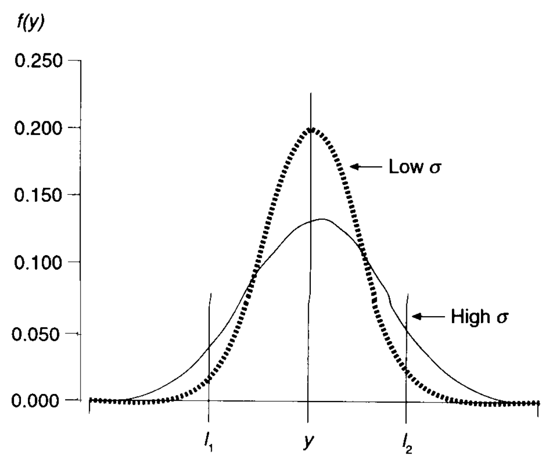
This is a property we are going to use a lot in this course.

3. Measures of Variability/Dispersion: The Variance and Standard deviation

Notation: var(X) or $\boldsymbol{\sigma}^2$

a) Definition

Aside: The probability density function (pdf) of a random variable, is a function that describes the relative likelihood for this random variable to take on a given value.



The **variance** of a variable X tells us the *how far is each value of X from the mean, on average*:

$$\mathbf{Var\ (X) = E(X-\boldsymbol{\mu}_x)^2 = E(X^2) - \boldsymbol{\mu}_x^2}$$

The standard deviation SD(X) is the positive square root of the variance.

b) Properties

b1) if C=const $\Leftrightarrow$ Var(C)=0

(on average, it is at distance 0 from its mean). Otherwise, Var(X)>0.

b2) for A, B = const $\Rightarrow$ Var (AX+B) = A^2var(X)

III. Measures of association

1. Covariance

Cov (X,Y) = E[(X - $\boldsymbol{\mu}_X$)(Y- $\boldsymbol{\mu}_Y$)] = E(XY) - $\boldsymbol{\mu}_X \boldsymbol{\mu}_Y$

The covariance between two variables X and Y is a measure of how the two variables X and Y are related – a positive covariance indicates they move in the same direction, a negative – that they move in the opposite direction. Put differently, a positive covariance means that when X is above its average, so is Y, and vice versa.

b) Properties

b1) If X=Y, the Cov(X,Y)=Var(X). This follows by the definition of variance and covariance.

b2) If X and Y are independent then Cov(X,Y)=0.

b3) Cov (AX + B,CY + D) =AC Cov(X,Y).

Note: The covariance *depends* on the units of measurement of variables X and Y, which is not very convenient. For this reason we often calculate the correlation, which is just the standardized covariance.

2. Correlation coefficient

Corr (X, Y) = Cov (X, Y) / SD(X)*SD(Y)

b) Properties

b1) -1<= Corr (X, Y)<=1

The magnitude of the correlation between two variables shows the strength of the linear relationship between X and Y – the closer to 1 in absolute value the correlation, the stronger the relationship between the variables.

b2) Corr (A_1X + B_1, A_2Y + B_2)=Corr(X,Y)

The covariance does *not* depend on the units of measurement.

3. Variance of a sum of random variables

$Var(AX+BY)=A^2Var(X) + B^2Var(Y) + 2ABCov(X, Y)$

If X and Y are independent then: $Var(AX+BY)=A^2Var(X) + B^2Var(Y)$

I. Measures of central tendency

2. The Median

Notation: med(X)

a) Definition

A general definition of the median is too complicated for our purposes. Just think of it the following way:

- if X is **continuous**, the med(X) is the value of X s.t. ½ of the area under the PDF is to the left and ½ is to the right of it
- if X is **discrete**, the med(X) is the value which is in the middle when ordering all possible values of X in increasing order.