Joint Distributions II

Conditional Distributions

- The relationship between two random variables can often be clarified by consideration of the conditional distribution of one given the value of the other.

- Recall that for any two events E and F , the conditional probability of E given F is defined, provided that P(F) > 0, by

$$P(E/F) = P(EF)/P(F)$$

Conditional Distributions - Discrete Case

If X and Y are discrete random variables, it is natural to define the conditional probability mass function of X given that Y = y, by

$$p_{X|Y}(x \mid y) = P\{X = x \mid Y = y\}$$

$$= P\{X = x, Y = y\} / P\{Y = y\}$$

$$= p(x,y) / p_Y(y)$$

for all values of y such that $p_Y(y) > 0$

Conditional Distributions - Discrete Case

Suppose that $p(x, y)$, the joint probability mass function of X and Y , is given by

$$p(0, 0) = .4, \quad p(0, 1) = .2, \quad p(1, 0) = .1, \quad p(1, 1) = .3$$

Calculate the conditional probability mass function of X given that Y = 1.

Conditional Distributions - Discrete Case

Conditional Distributions - Continuous Case

If X and Y have a joint probability density function f (x, y), then the conditional probability density function of X , given that Y = y, is defined for all values of y such that $f_Y(y) > 0$, by

$$f_{X\,|\,Y}(x \mid y) = f(x,y)/ f_Y(y)$$

The use of conditional densities allows us to define conditional probabilities of events associated with one random variable when we are given the value of a second random variable. That is, if X and Y are jointly continuous, then, for any set A,

$$P\{X \in A|Y = y\} = \int_A f_{X\,|\,Y}(x \mid y)\, dx$$

Conditional Distributions - Continuous Case

The joint density function of X and Y is given by

$$f(x, y) = \begin{cases} 12/5\, x\,(2 - x - y) & 0 < x < 1,\ 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

Compute the conditional density of X , given that Y = y, where 0 < y < 1.

$6x(2 - x - y)/(4 - 3y)$

Conditional Expectation - Discrete Case

- More often than not, we are interested in questions such as determining a person's weight given that he is a particular height. Generally, we want to find the expected value of X given information about Y. If [X , Y] is a discrete random vector, the conditional expectations are

$$E(X \mid y) = \sum_{x} x \, p_{X \mid y}(x)$$

and

$$E(Y \mid x) = \sum_{y} y \, p_{Y \mid x}(y)$$

Conditional Expectation - Continuous Case

- If [X , Y] is a continuous random vector, the conditional expectations are

$$E(X \mid y) = \int_{-\infty}^{\infty} x \,.\, f_{X \mid y}(x)\, dx$$

and

$$E(Y \mid x) = \int_{-\infty}^{\infty} y \,.\, f_{Y \mid x}(y)\, dy$$

Conditional Expectation - Continuous Case

The joint density function of X and Y is given by

$$f(x, y) = \begin{cases} x^2 + xy/3 & 0 \leq x \leq 1,\ 0 \leq y \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

Find the conditional expectation E(Y | x)

Regression of the mean

You can observe that E (Y | *x*) is a value of the random variable E(Y | X) for a particular X = *x* and it is a function of *x*. The graph of this function is called the Regression of Y on X.

Alternatively, the function E(X | y) would be called the regression of X on Y.

Degree of Association

There are two important measures which describe the degree of association between two random variables X & Y

Covariance

Correlation

Covariance - Qualitative Definition

- Covariance signifies the direction of the linear relationship between the two variables. By direction we mean if the variables are directly proportional or inversely proportional to each other. (Increasing the value of one variable might have a positive or a negative impact on the value of the other variable).

- The values of covariance can be any number between the two opposite infinities. Also, it’s important to mention that covariance only measures how two variables change together, not the dependency of one variable on another one.

Covariance - Quantitative Definition

The covariance of two random variables X and Y , written Cov(X , Y) is defined by

$$Cov(X, Y) = E[(X - \mu_x)(Y - \mu_y)]$$

where

$$\mu_x = E[X]$$
$$\mu_y = E[Y]$$

It can further be written as

$$Cov(X, Y) = E[XY] - E[X]E[Y]$$

Covariance - Properties

- $Cov(X,Y) = Cov\,(Y,X)$
- $Cov(X,X) = Var(X)$
- $Cov(aX, Y) = a\,Cov(X,Y)$
- $Cov(X+Z, Y) = Cov(X,Y) + Cov(Z,Y)$

Correlation

In general, it can be shown that a positive value of Cov(X , Y) is an indication that Y tends to increase as X does, whereas a negative value indicates that Y tends to decrease as X increases. The strength of the relationship between X and Y is indicated by the correlation between X and Y , a dimensionless quantity obtained by dividing the covariance
by the product of the standard deviations of X and Y .

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}$$

Correlation - Exercise

The joint density function of X and Y is given by

$$f(x, y) = \begin{cases} 6x & 0 < x < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the correlation between X and Y