Lecture

INSTITUTE OF ACTUARIAL & QUANTITATIVE STUDIES

Class: FY BSc

Subject : Probability and Statistics 1

Chapter: Unit 4 Chapter 1

Chapter Name: Joint Distributions

Today's Agenda

1. Joint Distributions
    1. Joint Probability Density Functions
    2. Marginal Probability Functions
    3. Conditional probability Functions
    4. Independence of Random Variables

2. Expectation of function of two variables

3. Covariance and Correlation coefficient

4. Convolutions

5. Linear Combinations of Random Variables

1.1 Joint Probability (density) Functions

Defining several random variables simultaneously on a sample space gives rise to a multivariate distribution. In the case of just two variables, it is a bivariate distribution.

Discrete Case

The function ***f* (*x*,*y*) = *P*(*X* = *x*,*Y* = *y*)** for all values of (x,y) is the (joint/bivariate) probability function of (X,Y) – it specifies how the total probability of 1 is divided up amongst the possible values of (x,y) and so gives the (joint/bivariate) probability distribution of (X,Y) .

The requirements for a function to qualify as the probability function of a pair of discrete random variables are:

- $f(x,y) \geq 0$ for all values of x and y in the domain
- $\sum_x \sum_y f(x,y) = 1$

1.1 Joint Probability Functions

Continuous case

In the case of a pair of continuous variables, the distribution of probability over a specified area in the (*x*,*y*) plane is given by the (joint) probability density function *f* (*x*,*y*) . The probability that the pair (*X*,*Y*) takes values in some specified region *A* is obtained by integrating *f* (*x*,*y*) over *A* – this integral is a "double" integral.

Thus: $\boldsymbol{P(x1 < X < x2, y1 < Y < y2) = \int_{y1}^{y2}\int_{x1}^{x2} f(x,y)\,dx\,dy}$

The joint distribution function *F*(*x*,*y*) is defined by: $\boldsymbol{F(x,y) = P(X \leq x, Y \leq y)}$

and it is related to the joint density function by:

$$\mathbf{f(x,y)} = \frac{d^2}{dx\,dy}\,F(x,y)$$

1.1 Joint Probability Functions

Continuous case

The conditions for a function to qualify as a joint probability density function of a pair of continuous random variables are:

$f(x,y) \geq 0$ for all values of x and y in the domain

$$\int_x \int_y f(x,y)\,dx\,dy = 1$$

1.2 Marginal Probability Functions

Discrete case
The marginal distribution of a discrete random variable *X* is defined to be:

$$\boldsymbol{f_X(x) = \sum_y f(x, y)}$$

This is the distribution of *X* alone without considering the values that *Y* can take.

Continuous case
In the case of continuous variables the marginal density function of *X*, f_X(x) is obtained by "integrating over *y*" (for the given value of *x*) the joint PDF *f* (*x*,*y*) .

$$\boldsymbol{f_X(x) = \int_y f(x, y) dy}$$

The resulting $f_X(x)$ is a proper PDF – it integrates to 1. Similarly for $f_Y(y)$, by "integrating over *x*" (for the given value of *y*).

1.3 Conditional Probability Functions

The distribution of X for a particular value of Y is called the conditional distribution of X given y .

Discrete case

The probability function PX|Y =y (x | y) for the conditional distribution of X given Y = y for discrete random variables X and Y is:

$$\mathbf{P}_{X|Y=y}\,(x, y) = \mathbf{P(X=x \mid Y=y)} = \frac{\mathbf{P_{X,Y}\,(x,y)}}{\mathbf{P_Y\,(y)}}$$

for all values x in the range of X

This conditional distribution is only defined for those values of y for which $P_Y(y) > 0$

1.3 Conditional Probability Functions

Continuous case

The probability density function $f_{X|Y=y}\ (x, y)$ for the conditional distribution of *X* given *Y* = *y* for the continuous variables *X* and *Y* is a function such that:

$$\int_{x=x1}^{x2} f_{X|Y=y}\ (x, y)dx = \mathbf{P(x1 < X < x2 \mid Y = y)}$$

for all values x in the range of X.

This conditional distribution is only defined for those values of y for which $f_Y(y) > 0$.

1.4 Independence of Random Variables

Consider a pair of variables (X,Y) , and suppose that the conditional distribution of Y given $X = x$ does not actually depend on x at all. It follows that the probability function/PDF $f\,(y|\,x)$ must be simply that of the marginal distribution of Y, $fY\,(y)$.

So, if "conditional is equivalent to marginal", then:

$$\mathrm{f_Y(y)} = \mathrm{f}_{\mathrm{Y}|X=x}\,\mathrm{(y, x)} = \frac{\mathrm{f_{X,Y}(x, y)}}{f_X\ (x)}$$

i.e. $\mathbf{f_{X,Y}(x, y) = f_X\ (x).\ f_Y(y)}$

so "joint PF/PDF is the product of the marginals".

1.4 Independence of Random Variables

The random variables X and Y are independent if, and only if, the joint probability function/PDF is the product of the two marginal probability functions/PDFs for all (x,y) in the range of the variables, *ie*:

$\mathbf{f_{X,Y}(x, y) = f_X(x). f_Y(y)}$ for all (x,y) in the range.

Discrete case
It follows that probability statements about values assumed by (X,Y) can be broken down into statements about X and Y separately. So if X and Y are independent discrete variables then:
$P(X = x, Y = y) = P(X = x).\ P(Y = y)$

Continuous case
If X and Y are continuous, the double integral required to evaluate a joint probability splits into the product of two separate integrals, one for X and one for
Y, and we have:
$P(x1 < X < x2, y1 < Y < y2) = P(x1 < X < x2).\ P(y1 < Y < y2)$

Question

CT3 September 2016 Q6

Let X and Y be random variables with joint probability distribution:

$$f_{XY}(x,y) = \begin{cases} kx^2y^2, & 0 < x < y < 1 \\ 0, & \text{otherwise} \end{cases}$$

where k is a constant.

(i) Show that $k = 18$.

(ii) Determine $\mathbf{f_Y}(y)$, the marginal density function of Y.

(iii) Determine $P(X > 0.5 \mid Y = 0.75)$.

Solution

(i) $$\iint_{x\,y} f_{XY}(x,y)\,dydx = \int_0^1\int_x^1 kx^2y^2\,dydx = \int_0^1 \left[\frac{k}{3}x^2y^3\right]_{y=x}^{y=1} dx$$

$$= \frac{k}{3}\int_0^1 x^2 - x^5\,dx = \frac{k}{3}\left[\frac{x^3}{3} - \frac{x^6}{6}\right]_0^1 = \frac{k}{3}\left(\frac{1}{3} - \frac{1}{6}\right) = \frac{k}{18}$$

Want integral equal to $1 \Rightarrow k = 18$

Solution

(ii) $$f_Y(y) = \int_x f_{XY}(x,y)\,dx = \int_0^y 18x^2y^2\,dx = \left[6x^3y^2\right]_{x=0}^{x=y} = 6y^5$$

(iii) $$P(X > 0.5|Y = 0.75) = \int_{0.5}^{0.75} f_{(x|Y=0.75)}(x)\,dx = \int_{0.5}^{0.75} f_{XY}(x,0.75)/f_Y(0.75)\,dx$$

$$= \int_{0.5}^{0.75} 18x^2 0.75^2/(6\times 0.75^5)\,dx = 3\times\left(\frac{4}{3}\right)^3\left[\frac{x^3}{3}\right]_{0.5}^{0.75} = 0.7037$$

2 Expectations of functions of two Variables

Expectations

The expression for the expected value of a function $g(X,Y)$ of the random variables (X,Y) is found by summing (discrete case) or integrating (continuous case) the product:

Value * probability of assuming that value

over all values (or combinations of) (x,y) . The summation is a double summation, the integral a double integral.

Discrete case

$$E[g(X,Y)] = \sum_x \sum_y g(x,y)\, p_{X,Y}(x,y) = \sum_x \sum_y g(x,y)\, P(X = x, Y = y)$$

where the summation is over all possible values of x and y.

Continuous case

$$E[g(X,Y)] = \int_x^{\cdot} \int_y^{\cdot} g(x,y)\, f_{X,Y}(x,y)\,dx\,dy$$

where the integration is over all possible values of x and y .

2 Expectations of functions of two Variables

Expectation of a Sum

It follows that:

$$E\,[ag(X) + bh(Y)] = aE\,[g(X)] + bE\,[h(Y)]$$

where a and b are constants.

Expectation of a product

For independent random variables X and Y:

$$E\,[g(X).\,h(Y)] = E\,[g(X)].\,E[h(Y)]$$

since the joint density function factorises into the two marginal density functions.

Question

CT3 April 2015 Q8

The random variables *X* and *Y* have a joint probability distribution with density function:

$$f_{xy}(x,y)=\begin{cases}3x, & 0<y<x<1\\ 0, & \text{otherwise}\end{cases}$$

(i) Determine the marginal densities of *X* and *Y*.
(ii) State, with reasons, whether *X* and *Y* are independent.
(iii) Determine $E[X]$ and $E[Y]$.

Solution

(i) $$f_X(x) = \int_0^x 3x\,dy = [3xy]_{y=0}^{y=x} = 3x^2 \text{ for } 0 < x < 1$$

$$f_Y(y) = \int_y^1 3x\,dx = \left[\frac{3}{2}x^2\right]_{x=y}^{x=1} = \frac{3}{2}\left(1 - y^2\right) \text{ for } 0 < y < 1$$

(ii) Not independent because $f_X(x)\,f_Y(y) \neq f_{XY}(x, y)$

Solution

(iii) $$E[X] = \int_0^1 x f_X(x)\,dx = \left[\frac{3}{4}x^4\right]_0^1 = 0.75$$

$$E[Y] = \int_0^1 y f_Y(y)\,dy = \left[\frac{3}{2}\left(\frac{y^2}{2} - \frac{y^4}{4}\right)\right]_0^1 = \frac{3}{8}$$

3 Covariance and Correlation Coefficient

The covariance cov[X,Y] of two random variables X and Y is defined by:

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

which simplifies to:

Cov [X, Y] = E[X. Y] - E[X]. E[Y]

Useful results on handling covariance

(a) cov [$a.X + b$, $c.Y + d$]= $a.c.$ cov[X,Y]

(b) cov [X,$Y + Z$]= cov[X,Y] + cov[X,Z]

These two results hold for any random variables X, Y and Z (whenever the covariance exist).

The next result concerns random variables that are independent.

(c) If X and Y are independent, cov[X,Y] = 0 .

The correlation coefficient (X ,Y) of two random variables X and Y is defined by

$$\textbf{corr (X, Y)} = \frac{cov\,(X,Y)}{\sqrt{var\,(X).var(Y)}}$$

3 Variance of functions of two Variables

Variance of a Sum

For any random variables X and Y:

$\mathrm{var}[X + Y] = \mathrm{var}[X] + \mathrm{var}[Y] + 2\mathrm{cov}[X,Y]$

For independent random variables, this can be simplified:

$\mathrm{var}[X + Y] = \mathrm{var}[X] + \mathrm{var}[Y]$

since $\mathrm{cov}[X,Y] = 0$.

Question

CT3 April 2016 Q2

Consider two random variables X and Y.

(i) Write down the precise mathematical definition for the correlation coefficient $\rho(X, Y)$ between X and Y.

Assume now that $Y = aX + b$ where $a < 0$ and $-\infty < b < \infty$.

(ii) Determine the value of the correlation coefficient $\rho(X, Y)$.

Solution

(i) $$\rho(X,Y) = \frac{\text{Cov}(X,Y)}{\sqrt{V(X)V(Y)}}$$

(ii) $$\text{Cov}(X, aX + b) = aV(X)$$

$$V(Y) = V(aX + b) = a^2V(X)$$

For $a < 0$ we obtain $\rho(X,Y) = \dfrac{\text{Cov}(X,Y)}{\sqrt{V(X)V(Y)}} = \dfrac{aV(X)}{\sqrt{V(X)a^2V(X)}} = -1$

4 Convolutions

Much of statistical theory involves the distributions of sums of random variables. In particular the sum of a number of independent variables is especially important.

Definition

When a function P_Z can be expressed as a sum of this form, then P_Z is called the convolution of the functions P_X and P_Y . This is written symbolically as $P_Z = P_X * P_Y$. So here, the probability function of $Z = X + Y$ is the convolution of the (marginal) probability functions of X and Y.

4 Convolutions

Discrete Case

Consider the sum of two discrete random variables, so let $Z = X + Y$, where (X,Y) has joint probability function $P(x,y)$.

Then $P(Z = z)$ is found by summing $P(x,y)$ over all values of (x,y) such that

$\boldsymbol{x + y = z}$ ***i.e.*** $\boldsymbol{P_Z(z) = \sum_x P(x, z - x)}$

Now suppose that X and Y are independent variables, then $P(x,y)$ is the product of the two marginal probability functions, so

$\boldsymbol{P_Z(z) = \sum_x P_X(x)\, P_Y(z - x)}$

Continuous case

In the case where X and Y are independent continuous variables with joint probability density function $f(x,y)$, the corresponding expression is:

$\boldsymbol{f_Z(z) = \int_x f_X(x)\, f_Y(z - x)\, dx}$

5 Linear Combinations of Random Variables

Mean

If X_1, X_2,…., X_n are any random variables (not necessarily independent), then:

$$\mathbf{E[c_1X_1 + c_2X_2 + \dots + c_nX_n] = c_1E[X_1] + c_2E[X_2] + \dots. + c_nE[X_n]}$$

Variance

If X_1, X_2,…., X_n are pairwise uncorrelated (and hence certainly if they are independent) random variables, then:

$$\mathbf{Var[c_1X_1 + c_2X_2 + \dots + c_nX_n] = {c_1}^2var(X_1) + {c_2}^2var(X_2) + \dots + {c_n}^2var(X_n)}$$

5 Linear Combinations of Random Variables

In many cases generating functions may make it possible to specify the actual distribution of Y, where $Y = c_1X_1 + c_2X_2 + \dots + c_nX_n$

PGF

Let $Y = X_1 + X_2 + \dots + X_n$ where the Xi are independent and Xi has PGF $G_i(t)$, then:

$$\boldsymbol{G_Y(t) = G_1(t).\, G_2(t) \dots G_n(t)}$$

and if X_i in the sum is replaced by cX_i then $G_i(t)$ in the product is replaced by $G_i(t^c)$.

MGF

Let $Y = X_1 + X_2 + \dots + X_n$ where the Xi are independent and Xi has MGF $M_i(t)$, then:

$$\boldsymbol{M_Y(t) = M_1(t).\, M_2(t) \dots M_n(t)}$$

and if X_i in the sum is replaced by cX_i then $M_i(t)$ in the product is replaced by $M_i(\mathrm{ct})$.