EEF 271E
Probability and Statistics

Week #4

CONTINUOUS RANDOM VARIABLES

We introduced the concept of a random variable as a real-valued function defined over the points of a sample space with a probability measure .

In the continuous case, where random variables can take on values on a continuous scale, the procedure is very much the same.

PROBABILITY DENSITY FUNCTIONS

The definition of probability in the continuous case presumes for each random variable the existence of a function, called a **probability density function**, such that areas under the curve give the probabilities associated with the corresponding intervals along the horizontal axis.

In other words, a probability density function, integrated from a to b (with $a \leqq b$), gives the probability that the corresponding random variable will take on a value on the interval from a to b.

DEFINITION A function with values $f(x)$, defined over the set of all real numbers, is called a **probability density function** of the continuous random variable X if and only if

$$P(a \leqq X \leqq b) = \int_a^b f(x)\,dx$$

for any real constants a and b with $a \leqq b$.

Probability density functions are also referred to, more briefly, as **probability densities, density functions, densities**, or **p.d.f.'s**.

Note that $f(c)$, the value of the probability density of X at c, does not give $P(X = c)$ as in the discrete case.

In connection with continuous random variables, probabilities are always associated with intervals and $P(X = c) = 0$ for any real constant c.

$f(x)$ is the value of *a* probability density, not *the* probability density of the random variable X at x.

THEOREM A function can serve as a probability density of a continuous random variable X if its values, $f(x)$, satisfy the conditions†

1. $f(x) \geqq 0 \quad \text{for } -\infty < x < \infty;$

2. $\displaystyle\int_{-\infty}^{\infty} f(x)\,dx = 1.$

EXAMPLE

If X has the probability density

$$f(x) = \begin{cases} k \cdot e^{-3x} & \text{for } x > 0 \\ 0 & \text{elsewhere} \end{cases}$$

find k and $P(0.5 \leqq X \leqq 1)$.

Solution To satisfy the second condition we must have

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{0}^{\infty} k \cdot e^{-3x}\,dx$$

$$= k \cdot \lim_{t \to \infty} \frac{e^{-3x}}{-3}\Big|_0^t = \frac{k}{3} = 1$$

and it follows that $k = 3$.

For the probability we get

$$P(0.5 \leqq X \leqq 1) = \int_{0.5}^{1} 3e^{-3x}\,dx = -e^{-3x}\Big|_{0.5}^{1} = -e^{-3} + e^{-1.5} = 0.173$$

As in the discrete case, there are many problems in which it is of interest to know the probability that the value of a continuous random variable X is less than or equal to some real number x. Thus, let us make the following definition

Definition If X is a continuous random variable and the value of its probability density at t is $f(t)$, then the function given by

$$F(x) = P(X \leqq x) = \int_{-\infty}^{x} f(t)\, dt \qquad \text{for } -\infty < x < \infty$$

is called the **distribution function**, or the **cumulative distribution**, of X.

The properties of distribution functions hold also for the continuous case; that is, $F(-\infty) = 0$, $F(\infty) = 1$, and $F(a) \leqq F(b)$ when $a < b$.

THEOREM If $f(x)$ and $F(x)$ are the values of the probability density and the distribution function of X at x, then

$$P(a \leqq X \leqq b) = F(b) - F(a)$$

for any real constants a and b with $a \leqq b$, and

$$f(x) = \frac{dF(x)}{dx}$$

where the derivative exists.

EXAMPLE

Find the distribution function of the random variable X of the previous Example and use it to reevaluate $P(0.5 \leqq X \leqq 1)$.

Solution For $x > 0$,

$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{0}^{x} 3e^{-3t}dt$$

$$= -e^{-3t}\Big|_{0}^{x} = 1 - e^{-3x}$$

and since $F(x) = 0$ for $x \leqq 0$, we can write

$$F(x) = \begin{cases} 0 & \text{for } x \leqq 0 \\ 1 - e^{-3x} & \text{for } x > 0 \end{cases}$$

To determine the probability $P(0.5 \leqq X \leqq 1)$, we use

$$P(0.5 \leqq X \leqq 1) = F(1) - F(0.5)$$

$$= (1 - e^{-3}) - (1 - e^{-1.5})$$
$$= 0.173$$

This agrees with the result obtained by using the probability density directly .

MULTIVARIATE DISTRIBUTIONS

In this section we shall be concerned first with the **bivariate case**, that is, with situations where we are interested at the same time in a pair of random variables defined over a joint sample space.

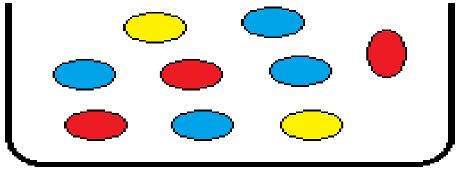
Later, we shall extend this discussion to the **nultivariate case**, covering any finite number of random variables.

If X and Y are discrete random variables, we write the probability that X will take on the value x and Y will take on the value y as $P(X = x, Y = y)$.

Thus, $P(X = x, Y = y)$ is the probability of the intersection of the events $X = x$ and $Y = y$.

As in the **univariate case**, where we dealt with one random variable we can now, in the bivariate case, display the probabilities associated with all pairs of values of X and Y by means of a table.

EXAMPLE



Two balls are selected at random from a box containing 3 Red 2 Yellow and 4 Blue balls. If X and Y are respectively, the numbers Red and Yellow balls, find the probabilities associated with all possible pairs of values of X and Y.

Solution

The possible pairs are

(0, 0), (0, 1), (1, 0), (1, 1), (0, 2), and (2, 0).

$$P = \frac{n}{N}$$

and the total number of ways in which two of the nine caplets can be selected is

$$N = \binom{9}{2} = 36.$$

To find the probability associated with (1, 0), The number of ways in which this can be done is

$$n = \binom{3}{1}\binom{2}{0}\binom{4}{1} = 12,$$

$$P = \frac{n}{N} = \frac{12}{36} = \frac{1}{3}$$

Similarly, the probability associated with (1, 1) is

$$\frac{\binom{3}{1}\binom{2}{1}\binom{4}{0}}{36} = \frac{6}{36} = \frac{1}{6}$$

and, continuing this way, we obtain the values shown in the following table:

		$x=0$	$x=1$	$x=2$
y	0	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{12}$
	1	$\frac{2}{9}$	$\frac{1}{6}$	
	2	$\frac{1}{36}$		

it is preferable to express the probabilities by means of a function with the values

$$f(x, y) = P(X = x, Y = y)$$

for any pair of values (x, y) within the range of the random variables X, Y

$$f(x, y) = \frac{\binom{3}{x}\binom{2}{y}\binom{4}{2-x-y}}{\binom{9}{2}}$$

for $x = 0, 1, 2;$
$0 \leqq x + y \leqq 2$

DEFINITION If X and Y are discrete random variables, the function given by $f(x, y) = P(X = x, Y = y)$ for each pair of values (x, y) within the range of X and Y is called the **joint probability distribution** of X and Y.

THEOREM A bivariate function can serve as the joint probability distribution of a pair of discrete random variables X and Y if and only if its values, $f(x, y)$, satisfy the conditions

1. $f(x, y) \geqq 0$ for each pair of values (x, y) within its domain;

2. $\sum_x \sum_y f(x, y) = 1$, where the double summation extends over all possible pairs (x, y) within its domain.