# Probability and Statistics (IT302)

# 26th August 2020 Wednesday 11:15 AM-11:45 AM

# Class 10

# Conditional Probability Distribution

It is extremely important that we make use of the special type of distribution of the form $f(x, y)/g(x)$ in order to be able to effectively compute conditional probabilities. This type of distribution is called a **Conditional Probability Distribution**; the formal definition follows.

Let $X$ and $Y$ be two random variables, discrete or continuous. The **conditional distribution** of the random variable $Y$ given that $X = x$ is

$$f(y|x) = \frac{f(x, y)}{g(x)}, \text{ provided } g(x) > 0.$$

Similarly, the conditional distribution of $X$ given that $Y = y$ is

$$f(x|y) = \frac{f(x, y)}{h(y)}, \text{ provided } h(y) > 0.$$

**Source :** Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.

# Conditional Probability Distribution Contd.

If we wish to find the probability that the discrete random variable $X$ falls between $a$ and $b$ when it is known that the discrete variable $Y = y$, we evaluate

$$P(a < X < b \mid Y = y) = \sum_{a<x<b} f(x|y),$$

where the summation extends over all values of $X$ between $a$ and $b$. When $X$ and $Y$ are continuous, we evaluate

$$P(a < X < b \mid Y = y) = \int_a^b f(x|y)\, dx.$$

**Source :** Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.

# Example 3.14

**Example 3.14:** Two ballpoint pens are selected at random from a box that contains 3 blue pens, 2 red pens, and 3 green pens. If $X$ is the number of blue pens selected and $Y$ is the number of red pens selected, find

(a) the joint probability function $f(x, y)$,

(b) $P[(X, Y) \in A]$, where $A$ is the region $\{(x, y) | x + y \leq 1\}$.

***Solution:*** The possible pairs of values $(x, y)$ are $(0, 0)$, $(0, 1)$, $(1, 0)$, $(1, 1)$, $(0, 2)$, and $(2, 0)$.

(a) Now, $f(0, 1)$, for example, represents the probability that a red and a green pens are selected. The total number of equally likely ways of selecting any 2 pens from the 8 is $\binom{8}{2} = 28$. The number of ways of selecting 1 red from 2 red pens and 1 green from 3 green pens is $\binom{2}{1}\binom{3}{1} = 6$. Hence, $f(0, 1) = 6/28$ $= 3/14$. Similar calculations yield the probabilities for the other cases, which

**Source :** Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.

# Example 3.14 Contd.

5, it will become clear that the joint probability distribution of Table 3.1 can be represented by the formula

$$f(x, y) = \frac{\binom{3}{x}\binom{2}{y}\binom{3}{2-x-y}}{\binom{8}{2}},$$

for $x = 0, 1, 2$; $y = 0, 1, 2$; and $0 \le x + y \le 2$.

(b) The probability that $(X, Y)$ fall in the region $A$ is

$$P[(X, Y) \in A] = P(X + Y \le 1) = f(0, 0) + f(0, 1) + f(1, 0)$$
$$= \frac{3}{28} + \frac{3}{14} + \frac{9}{28} = \frac{9}{14}.$$

Table 3.1: Joint Probability Distribution for Example 3.14

| | $f(x, y)$ | $x$ = 0 | 1 | 2 | Row Totals |
|---|---|---|---|---|---|
| $y$ | 0 | $\frac{3}{28}$ | $\frac{9}{28}$ | $\frac{3}{28}$ | $\frac{15}{28}$ |
| | 1 | $\frac{3}{14}$ | $\frac{3}{14}$ | 0 | $\frac{3}{7}$ |
| | 2 | $\frac{1}{28}$ | 0 | 0 | $\frac{1}{28}$ |
| Column Totals | | $\frac{5}{14}$ | $\frac{15}{28}$ | $\frac{3}{28}$ | 1 |

**Source :** Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.

# Conditional Probability Distribution Example 3.18

Referring to Example 3.14, find the Conditional Distribution of $X$, given that $Y = 1$, and use it to determine $P(X = 0 \mid Y = 1)$.

***Solution*:** We need to find $f(x|y)$, where $y = 1$. First, we find that

$$h(1) = \sum_{x=0}^{2} f(x,1) = \frac{3}{14} + \frac{3}{14} + 0 = \frac{3}{7}.$$

Now

$$f(x|1) = \frac{f(x,1)}{h(1)} = \left(\frac{7}{3}\right) f(x,1), \quad x = 0, 1, 2.$$

**Source :** Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.

# Conditional Probability Distribution Example 3.18 Contd.

Therefore,

$$f(0|1) = \left(\frac{7}{3}\right) f(0,1) = \left(\frac{7}{3}\right)\left(\frac{3}{14}\right) = \frac{1}{2}, \quad f(1|1) = \left(\frac{7}{3}\right) f(1,1) = \left(\frac{7}{3}\right)\left(\frac{3}{14}\right) = \frac{1}{2},$$

$$f(2|1) = \left(\frac{7}{3}\right) f(2,1) = \left(\frac{7}{3}\right)(0) = 0,$$

and the conditional distribution of $X$, given that $Y = 1$, is

| $x$ | 0 | 1 | 2 |
|---|---|---|---|
| $f(x\|1)$ | $\frac{1}{2}$ | $\frac{1}{2}$ | 0 |

Finally,

$$P(X = 0 \mid Y = 1) = f(0|1) = \frac{1}{2}.$$

Therefore, if it is known that 1 of the 2 pen refills selected is red, we have a probability equal to 1/2 that the other refill is not blue.

**Source :** Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.

# Conditional Probability Distribution Example 3.19

**Example 3.19:** The joint density for the random variables $(X, Y)$, where $X$ is the unit temperature change and $Y$ is the proportion of spectrum shift that a certain atomic particle produces, is

$$f(x, y) = \begin{cases} 10xy^2, & 0 < x < y < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Find the marginal densities $g(x)$, $h(y)$, and the conditional density $f(y|x)$.

(b) Find the probability that the spectrum shifts more than half of the total observations, given that the temperature is increased by 0.25 unit.

***Solution:*** (a) By definition,

$$g(x) = \int_{-\infty}^{\infty} f(x, y)\, dy = \int_{x}^{1} 10xy^2\, dy$$

$$= \frac{10}{3}xy^3\Big|_{y=x}^{y=1} = \frac{10}{3}x(1 - x^3),\ 0 < x < 1,$$

$$h(y) = \int_{-\infty}^{\infty} f(x, y)\, dx = \int_{0}^{y} 10xy^2\, dx = 5x^2y^2\Big|_{x=0}^{x=y} = 5y^4,\ 0 < y < 1.$$

**Source :** Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.

# Conditional Probability Distribution Example 3.19 Contd.

Now

$$f(y|x) = \frac{f(x,y)}{g(x)} = \frac{10xy^2}{\frac{10}{3}x(1-x^3)} = \frac{3y^2}{1-x^3}, \quad 0 < x < y < 1.$$

(b) Therefore,

$$P\left(Y > \frac{1}{2} \,\middle|\, X = 0.25\right) = \int_{1/2}^{1} f(y \mid x = 0.25)\, dy = \int_{1/2}^{1} \frac{3y^2}{1-0.25^3}\, dy = \frac{8}{9}.$$

**Source :** Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.

# Conditional Probability Distribution Example 3.20

**Example 3.20:** Given the joint density function

$$f(x,y)=\begin{cases}\dfrac{x(1+3y^2)}{4}, & 0<x<2,\ 0<y<1,\\ 0, & \text{elsewhere},\end{cases}$$

**Source :** Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.

# Conditional Probability Distribution Example 3.20 Contd.

find $g(x)$, $h(y)$, $f(x|y)$, and evaluate $P(\frac{1}{4} < X < \frac{1}{2} \mid Y = \frac{1}{3})$.

***Solution:*** By definition of the marginal density. for $0 < x < 2$,

$$g(x) = \int_{-\infty}^{\infty} f(x, y)\ dy = \int_0^1 \frac{x(1 + 3y^2)}{4} dy$$
$$= \left( \frac{xy}{4} + \frac{xy^3}{4} \right) \Big|_{y=0}^{y=1} = \frac{x}{2},$$

and for $0 < y < 1$,

$$h(y) = \int_{-\infty}^{\infty} f(x, y)\ dx = \int_0^2 \frac{x(1 + 3y^2)}{4} dx$$
$$= \left( \frac{x^2}{8} + \frac{3x^2y^2}{8} \right) \Big|_{x=0}^{x=2} = \frac{1 + 3y^2}{2}.$$

Therefore, using the conditional density definition, for $0 < x < 2$,

$$f(x|y) = \frac{f(x, y)}{h(y)} = \frac{x(1 + 3y^2)/4}{(1 + 3y^2)/2} = \frac{x}{2},$$

and

$$P\left( \frac{1}{4} < X < \frac{1}{2} \;\middle|\; Y = \frac{1}{3} \right) = \int_{1/4}^{1/2} \frac{x}{2} dx = \frac{3}{64}.$$

**Source :** Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.

# Statistical Independence Introduction

If $f(x|y)$ does not depend on $y$, as is the case for Example 3.20, then $f(x|y) = g(x)$ and $f(x, y) = g(x)h(y)$. The proof follows by substituting

$$f(x, y) = f(x|y)h(y)$$

into the marginal distribution of $X$. That is,

$$g(x) = \int_{-\infty}^{\infty} f(x, y)\ dy = \int_{-\infty}^{\infty} f(x|y)h(y)\ dy.$$

If $f(x|y)$ does not depend on $y$, we may write

$$g(x) = f(x|y) \int_{-\infty}^{\infty} h(y)\ dy.$$

Now

$$\int_{-\infty}^{\infty} h(y)\ dy = 1,$$

since $h(y)$ is the probability density function of $Y$. Therefore,

$$g(x) = f(x|y) \quad \text{and then} \quad f(x, y) = g(x)h(y).$$

**Source :** Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.

# Statistical Independence Definition

Let $X$ and $Y$ be two random variables, discrete or continuous, with joint probability distribution $f(x, y)$ and marginal distributions $g(x)$ and $h(y)$, respectively. The random variables $X$ and $Y$ are said to be **statistically independent** if and only if

$$f(x, y) = g(x)h(y)$$

for all $(x, y)$ within their range.

If you can find any point $(x, y)$ for which $f(x, y)$ is defined such that $f(x, y) \neq g(x)h(y)$, the discrete variables $X$ and $Y$ are not statistically independent.

**Source :** Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.

# Statistical Independence Example 3.21

Show that the random variables of Example 3.14 are not statistically independent.

***Proof*:** Let us consider the point $(0, 1)$. From Table 3.1 we find the three probabilities $f(0, 1)$, $g(0)$, and $h(1)$ to be

$$f(0,1) = \frac{3}{14},$$

$$g(0) = \sum_{y=0}^{2} f(0,y) = \frac{3}{28} + \frac{3}{14} + \frac{1}{28} = \frac{5}{14},$$

$$h(1) = \sum_{x=0}^{2} f(x,1) = \frac{3}{14} + \frac{3}{14} + 0 = \frac{3}{7}.$$

Clearly,

$$f(0,1) \neq g(0)h(1),$$

and therefore $X$ and $Y$ are not statistically independent. $\blacksquare$

**Source :** Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.

# Mutually Statistically Independent Random Variable Definition

Let $X_1, X_2, \ldots, X_n$ be $n$ random variables, discrete or continuous, with joint probability distribution $f(x_1, x_2, \ldots, x_n)$ and marginal distribution $f_1(x_1), f_2(x_2), \ldots, f_n(x_n)$, respectively. The random variables $X_1, X_2, \ldots, X_n$ are said to be mutually **statistically independent** if and only if

$$f(x_1, x_2, \ldots, x_n) = f_1(x_1) f_2(x_2) \cdots f_n(x_n)$$

for all $(x_1, x_2, \ldots, x_n)$ within their range.

**Source :** Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.

# Mutually Statistically Independent Random Variable Example 3.22

**Example 3.22:** Suppose that the shelf life, in years, of a certain perishable food product packaged in cardboard containers is a random variable whose probability density function is given by

$$f(x) = \begin{cases} e^{-x}, & x > 0, \\ 0, & \text{elsewhere.} \end{cases}$$

Let $X_1$, $X_2$, and $X_3$ represent the shelf lives for three of these containers selected independently and find $P(X_1 < 2, 1 < X_2 < 3, X_3 > 2)$.

***Solution:*** Since the containers were selected independently, we can assume that the random variables $X_1$, $X_2$, and $X_3$ are statistically independent, having the joint probability density

$$f(x_1, x_2, x_3) = f(x_1)f(x_2)f(x_3) = e^{-x_1}e^{-x_2}e^{-x_3} = e^{-x_1-x_2-x_3},$$

for $x_1 > 0$, $x_2 > 0$, $x_3 > 0$, and $f(x_1, x_2, x_3) = 0$ elsewhere. Hence

$$\begin{aligned} P(X_1 < 2, 1 < X_2 < 3, X_3 > 2) &= \int_2^\infty \int_1^3 \int_0^2 e^{-x_1-x_2-x_3}\, dx_1\, dx_2\, dx_3 \\ &= (1 - e^{-2})(e^{-1} - e^{-3})e^{-2} = 0.0372. \end{aligned}$$

**Source :** Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.