4.2 Conditional Distributions and Independence

Definition 4.2.1 Let (X, Y) be a discrete bivariate random vector with joint pmf $f(x, y)$ and marginal pmfs $f_X(x)$ and $f_Y(y)$. For any x such that $P(X = x) = f_X(x) > 0$, the conditional pmf of Y given that $X = x$ is the function of y denoted by $f(y|x)$ and defined by

$$f(y|x) = P(Y = y|X = x) = \frac{f(x, y)}{f_X(x)}.$$

For any y such that $P(Y = y) = f_Y(y) > 0$, the conditional pmf of X given that $Y = y$ is the function of x denoted by $f(x|y)$ and defined by

$$f(x|y) = P(X = x|Y = y) = \frac{f(x, y)}{f_Y(y)}.$$

It is easy to verify that $f(y|x)$ and $f(x|y)$ are indeed distributions. First, $f(y|x) \geq 0$ for every y since $f(x, y) \geq 0$ and $f_X(x) > 0$. Second,

$$\sum_y f(y|x) = \frac{\sum_y f(x, y)}{f_X(x)} = \frac{f_X(x)}{f_X(x)} = 1.$$

Example 4.2.2 (Calculating conditional probabilities)

Define the joint pmf of (X, Y) by

$$f(0, 10) = f(0, 20) = \frac{2}{18}, \quad f(1, 10) = f(1, 30) = \frac{3}{18}, \quad f(1, 20) = \frac{4}{18}, \quad f(2, 30) = \frac{4}{18}.$$

The conditional probability

$$f_{Y|X}(10|0) = \frac{f(0, 10)}{f_X(0)} = \frac{f(0, 10)}{f(0, 10) + f(0, 20)} = \frac{1}{2}.$$

Definition 4.2.3

Let (X, Y) be a continuous bivariate random vector with joint pdf $f(x, y)$ and marginal pdfs $f_X(x)$ and $f_Y(y)$. For any x such that $f_X(x) > 0$, the conditional pdf of Y given that $X = x$ is the function of y denoted by $f(y|x)$ and defined by

$$f(y|x) = \frac{f(x, y)}{f_X(x)}.$$

For any y such that $f_Y(y) > 0$, the conditional pdf of X given that $Y = y$ is the function of x denoted by $f(x|y)$ and defined by

$$f(x|y) = \frac{f(x,y)}{f_y(y)}.$$

If $g(Y)$ is a function of Y, then the conditional expected value of $g(Y)$ given that $X = x$ is denoted by $E(g(Y)|x)$ and is given by

$$E(g(Y)|x) = \sum_y g(y)f(y|x) \quad \text{and} \quad E(g(Y)|x) = \int_{-\infty}^{\infty} g(y)f(y|x)dy$$

in the discrete and continuous cases, respectively.

Example 4.2.4 (Calculating conditional pdfs)

Let the continuous random vector (X,Y) have joint pdf

$$f(x,y) = e^{-y}, \quad 0 < x < y < \infty.$$

The marginal of X is

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_x^{\infty} e^{-y}dy = e6{-}x.$$

Thus, marginally, X has an exponential distribution. The conditional distribution of Y is

$$f(y|x) = \frac{f(x,y)}{f_X(x)} = \begin{cases} \frac{e^{-y}}{e^{-x}} = e^{-(y-x)}, & \text{if } y > x, \\ \frac{0}{e^{-x}} = 0, & \text{if } y \leq x \end{cases}$$

The mean of the conditional distribution is

$$E(Y|X = x) = \int_x^{\infty} ye^{-(y-x)}dy = 1 + x.$$

The variance of the conditional distribution is

$$\begin{aligned} \text{Var } (Y|x) &= E(Y^2|x) - (E(Y|x))^2 \\ &= \int_x^{\infty} y^2 e^{-(y-x)}dy - (\int_x^{\infty} ye^{-(y-x)})^2 \\ &= 1 \end{aligned}$$

In all the previous examples, the conditional distribution of Y given $X = x$ was different for different values of x. In some situations, the knowledge that $X = x$ does not give us any more information about Y than we already had. This important relationship between X and Y is called independence.

Definition 4.2.5 Let (X, Y) be a bivariate random vector with joint pdf or pmf $f(x, y)$ and marginal pdfs or pmfs $f_X(x)$ and $f_Y(y)$. Then X and Y are called independent random variables if, for **EVERY** $x \in \mathbb{R}$ and $y \in mR$,

$$f(x, y) = f_X(x) f_Y(y).$$

If X and Y are independent, the conditional pdf of Y given $X = x$ is

$$f(y|x) = \frac{f(x, y)}{f_X(x)} = \frac{f_X(x) f_Y(y)}{f_X(x)} = f_Y(y)$$

regardless of the value of x.

Lemma 4.2.7 Let (X, Y) be a bivariate random vector with joint pdf or pmf $f(x, y)$. Then X and Y are independent random variables if and only if there exist functions $g(x)$ and $h(y)$ such that, for every $x \in \mathbb{R}$ and $y \in \mathbb{R}$,

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

PROOF: The "only if" part is proved by defining $g(x) = f_X(x)$ and $h(y) = f_Y(y)$. To proved the "if" part for continuous random variables, suppose that $f(x, y) = g(x)h(y)$. Define

$$\int_{-\infty}^{\infty} g(x)dx = c \quad \text{and} \quad \int_{-\infty}^{\infty} h(y)dy = d,$$

where the constants c and d satisfy

$$\begin{aligned} cd &= (\int_{-\infty}^{\infty} g(x)dx)(\int_{-\infty}^{\infty} h(y)dy) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)dxdy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)dxdy = 1 \end{aligned}$$

Furthermore, the marginal pdfs are given by

$$f_X(x) = \int_{-\infty}^{\infty} g(x)h(y)dy = g(x)d$$

and

$$f_Y(y) = \int_{-\infty}^{\infty} g(x)h(y)dx = h(y)c.$$

Thus, we have

$$f(x,y) = g(x)h(y) = g(x)h(y)cd = f_X(x)f_Y(y),$$

showing that X and Y are independent. Replacing integrals with sums proves the lemma for discrete random vectors. □

<u>Example 4.2.8 (Checking independence)</u>

Consider the joint pdf $f(x,y) = \frac{1}{384}x^2y^2e^{-y-(x/2)}$, $x > 0$ and $y > 0$. If we define

$$g(x) = \begin{cases} x^2e^{-x/2} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

and

$$h(y) = \begin{cases} y^4e^{-y}/384 & y > 0 \\ 0 & y \leq 0 \end{cases}$$

then $f(x,y) = g(x)h(y)$ for all $x \in \mathbb{R}$ and all $y \in \mathbb{R}$. By Lemma 4.2.7, we conclude that X and Y are independent random variables.

<u>Theorem 4.2.10</u> Let X and Y be independent random variables.

(a) For any $A \subset \mathbb{R}$ and $B \subset \mathbb{R}$, $P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$; that is, the events $\{X \in A\}$ and $\{Y \in B\}$ are independent events.

(b) Let $g(x)$ be a function only of x and $h(y)$ be a function only of y. Then

$$E(g(X)h(Y)) = (Eg(X))(Eh(Y)).$$

PROOF: For continuous random variables, part (b) is proved by noting that

$$\begin{aligned} E(g(X)h(Y)) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f(x,y)dxdy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f_X(x)f_Y(y)dxdy \\ &= (\int_{-\infty}^{\infty} g(x)f_X(x)dx)(\int_{-\infty}^{\infty} h(y)f_Y(y)dy) \\ &= (Eg(X))(Eh(Y)). \end{aligned}$$

The result for discrete random variables is proved bt replacing integrals by sums.

Part (a) can be proved similarly. Let $g(x)$ be the indicator function of the set A. let $h(y)$ be the indicator function of the set B. Note that $g(x)h(y)$ is the indicator function of the set $C \in \mathbb{R}^2$ defined by $C = \{(x,y) : x \in A, y \in B\}$. Also note that for an indicator function such as $g(x)$, $Eg(X) = P(X \in A)$. Thus,

$$P(X \in A, Y \in B) = P((X,Y) \in C) = E(g(X)h(Y))$$
$$= (Eg(X))(Eh(Y)) = P(X \in A)P(Y \in B).$$

□

<u>Example 4.2.11 (Expectations of independent variables)</u>
Let X and Y be independent exponential(1) random variables. So

$$P(X \geq 4, Y \leq 3) = P(X \geq 4)P(Y \leq 3) = e^{-4}(1 - e^{-3})/$$

Letting $g(x) = x^2$ and $h(y) = y$, we have

$$E(X^2Y) = E(X^2)E(Y) = (2)(1) = 2.$$

<u>Theorem 4.2.12</u> Let X and Y be independent random variables with moment generating functions $M_X(t)$ and $M_Y(t)$. Then the moment generating function of the random variable $Z = X + Y$ is given by

$$M_Z(t) = M_X(t)M_Y(t).$$

PROOF:

$$M_Z(t) = Ee^{t(X+Y)} = (Ee^{tX})(Ee^{tY}) = M_X(t)M_Y(t).$$

□

Theorem 4.2.14 Let $X \sim N(\mu, \sigma^2)$ and $Y \sim N(\gamma, \tau^2)$ be independent normal random variables. Then the random variable $Z = X + Y$ has a $N(\mu + \gamma, \sigma^2 + \tau^2)$ distribution.

PROOF: Using Theorem 4.2.12, we have

$$M_Z(t) = M_X(t)M_Y(t) = \exp\{(\mu + \gamma)t + (\sigma^2 + \tau^2)t^2/2\}.$$

Hence, $Z \sim N(\mu + \gamma, \sigma^2 + \tau^2)$. □