4.4. Conditional Distributions

1. *Definition.* Suppose that X and Y are discrete random variables, and they have a joint pmf $f(x, y)$. Let S_1 be the range for X, and let S_2 be the range for Y.

   Let $f_1(x)$ be the marginal pmf of X, and let $f_2(y)$ be the marginal pmf of Y.

   For each fixed $b \in S_2$, the conditional pmf of X given that $Y = b$, is

$$g_1(x|Y = b) = g_1(x \mid b) = \frac{f(x, b)}{f_2(b)}, \quad x \in S_1.$$

Since

$$f(x,b) = P(X = x, Y = b), \quad f_2(b) = P(Y = b),$$

we have

$$g_1(x \mid Y = b) = \frac{P(X = x, Y = b)}{P(Y = b)} = P(X = x \mid Y = b).$$

Note

$$f_2(b) = \sum_{x \in S_1} f(x,b).$$

Dividing by $f_2(b)$, we get

$$1 = \sum_{x \in S_1} \frac{f(x,b)}{f_2(b)} = \sum_{x \in S_1} g_1(x \mid b).$$

So $g_1(x \mid b)$, when considered as a function of x, qualifies as a pmf.

So, $g_1(x \mid b)$ is the conditional pmf of X under the condition that $Y = b$. The mean of this conditional distribution is called the conditional mean of X given that $Y = b$, and the variance of this conditional distribution is called the conditional variance of X given that $Y = b$. The formal definitions are:

$$E(X|Y = b) = \sum_{x \in S_1} x g_1(x|b),$$

$$Var(X \mid b) = E(X^2|b) - (E(X|b))^2,$$

or, equivalently,

$$Var(X \mid b) = E[(X - E(X|b))^2|Y = b].$$

Similarly, for each fixed $a \in S_1$, the conditional pmf of Y given that $X = a$, is

$$g_2(y|X = a) = g_2(y|a) = \frac{f(a, y)}{f_1(a)}, \quad y \in S_2.$$

The conditional mean of Y, given that $X = a$, is

$$E(Y|X = a) = \sum_{y \in S_2} y g_2(y|a).$$

The conditional variance of Y, given that $X = a$, is

$$Var(Y|a) = E(Y^2|a) - (E(Y|a))^2$$

2. *Example.* Suppose that X and Y are discrete random variables, and suppose that their joint pmf is given by the table below.

$X\backslash f\backslash Y$	1	2	3
1	0.2	0.1	0.1
2	0.1	0.2	0.3

(a) Find the conditional pmf of Y given that $X = 2$.

(b) Find $E(Y|X = 2)$.

(c) Find $Var(Y|X = 2)$.

— *Solution.* First, we read the following information from the table:

$$f(2,1) = 0.1, \quad f(2,2) = 0.2, \quad f(2,3) = 0.3. \tag{1}$$

It follows that

$$f_1(2) = 0.1 + 0.2 + 0.3 = 0.6. \tag{2}$$

Divide the three numbers in (1) by $f_1(2) = 0.6$, we get

$$g_2(1 \mid 2) = \frac{1}{6}, \quad g_2(2 \mid 2) = \frac{1}{3}, \quad g_2(3 \mid 2) = \frac{1}{2}.$$

We summarize this information in a table

y	1	2	3
$g_2(y\|2)$	$1/6$	$1/3$	$1/2.$

This is $g_2(y|X=2)$, the conditional pmf of Y given that $X=2$.

The rest is routine:

$$E(Y|2) = 1 \times \frac{1}{6} + 2 \times \frac{1}{3} + 3 \times \frac{1}{2} = \frac{7}{3},$$

$$E(Y^2|2) = 1 \times \frac{1}{6} + 4 \times \frac{1}{3} + 9 \times \frac{1}{2} = 6,$$

$$Var(Y|2) = 6 - (7/3)^2 = \frac{5}{9}.$$

3. *Definition.* Suppose that X and Y are continuous random variables with a joint pdf $f(x, y)$.

   Let $f_1(x)$ be the marginal pdf of X, and let $f_2(y)$ be the marginal pdf of Y.

   If $b \in \mathbb{R}$ is such that $f_2(b) > 0$, then the conditional pdf of X, given that $Y = b$, is

$$g_1(x|b) = \frac{f(x, b)}{f_2(b)}.$$

   If $a \in \mathbb{R}$ is such that $f_1(a) > 0$, then the conditional pdf of Y, given that $X = a$, is

$$g_2(y|a) = \frac{f(a, y)}{f_1(a)}.$$

4. *Definition.* (continued) The conditional mean of Y, given that $X = a$, is

$$E(Y|a) = \int_{-\infty}^{\infty} y g_2(y|a) dy.$$

The conditional variance of Y, given that $X = a$, is

$$Var(Y|a) = E(Y^2|a) - [E(Y|a)]^2.$$

Similarly, we can define the following: The conditional mean of X, given that $Y = b$, is

$$E(X|b) = \int_{-\infty}^{\infty} x g_1(x|b) dx.$$

The conditional variance of X, given that $Y = b$, is

$$Var(X|b) = E(X^2|b) - [E(X|b)]^2.$$

5. *Example.* Suppose that X and Y are continuous random variables, and suppose that their joint pdf is

$$f(x, y) = x + y, \quad 0 < x < 1, 0 < y < 1.$$

(a) Find the conditional pdf of Y given that $X = 1/3$.

(b) Find $E(Y|X = 1/3)$.

(c) Find $Var(Y|X = 1/3)$.

— *Solution.* We have

$$f_1(x) = x + \frac{1}{2}, \quad 0 < x < 1.$$

Hence,

$$f_1(1/3) = \frac{5}{6}.$$

If follows that, for $0 < y < 1$,

$$g_2(y|\tfrac{1}{3}) = \frac{f(\frac{1}{3}, y)}{f_1(\frac{1}{3})} = \frac{\frac{1}{3} + y}{5/6} = \frac{2 + 6y}{5}.$$

This is the conditional pdf of Y given that $X = \frac{1}{3}$.

The rest is routine:

$$E(Y|\tfrac{1}{3}) = \int_0^1 y \cdot \frac{2+6y}{5} dy = \tfrac{3}{5},$$

$$E(Y^2|\tfrac{1}{3}) = \int_0^1 y^2 \cdot \frac{2+6y}{5} dy = \tfrac{13}{30},$$

$$Var(Y|\tfrac{1}{3}) = \tfrac{13}{30} - \left(\tfrac{3}{5}\right)^2 = \frac{11}{150}.$$