Conditional distributions and conditional expectations

Discrete case
Suppose that $X$ and $Y$ are discrete random variables with joint mass function $p(x, y)$. If $y$ is a real number with $p_Y(y) > 0$, then for any real number $x$,

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

It is easy to see that, the function (of $x$)

$$
p_X|Y(x|y) = \frac{p(x, y)}{p_Y(y)}, \quad x \in \mathbb{R}
$$

is a probability mass function and it is called the conditional mass function of $X$ given $Y = y$. The conditional expectation of $X$ given $Y = y$ is defined to be

$$
E[X|Y = y] = \sum_x p_X|Y(x|y).
$$

Example 1 Suppose that $X$ and $Y$ are independent Poisson random variables with parameters $\lambda_1$ and $\lambda_2$ respectively. Suppose that $n$ is a positive integer. Find the conditional mass function of $X$ given $X + Y = n$ and the conditional expectation of $X$ given $X + Y = n$.

Example 2 Suppose that $X$ and $Y$ are independent geometric random variables with parameter $p$. Suppose that $n > 2$ is a positive integer. Find the conditional mass function of $X$ given $X + Y = n$ and the conditional expectation of $X$ given $X + Y = n$.

Solution In a previous example, we have already found that

$$
P(X + Y = n) = (n - 1)p^2(1 - p)^{n-2}.
$$

Thus for any positive integer $x < n$,

$$
p_{X|X+Y}(x|n) = \frac{P(X = x, X + Y = n)}{P(X + Y = n)} = \frac{P(X = x, Y = n - x)}{P(X + Y = n)} = \frac{1}{n - 1}.
$$

In other words, given $X + Y = n$, $X$ is uniformly distributed in $\{1, 2, \ldots, n - 1\}$. Thus

$$
E[X|X + Y = n] = \frac{n}{2}.
$$

Absolute continuous case
Suppose that $X$ and $Y$ are jointly absolutely continuous with joint mass function $p(x, y)$. If $y$ is a real number with $f_Y(y) > 0$, then the following function of $x$

$$
f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)}, \quad x \in \mathbb{R}
$$

is a probability density function and it is called the conditional density function of $X$ given $Y = y$. The conditional expectation of $X$ given $Y = y$ is defined to be

$$
E[X|Y = y] = \int_x f_X(x|y) dx.
$$
is a probability density function and it is called the conditional density function of $X$ given $Y = y$. Conditional densities can be used to define conditional probabilities. Thus we define

$$P(a \leq X \leq b | Y = y) = \int_a^b f_{X|Y}(x|y) dx.$$  (0.1)

Alternatively, we could attempt to define the conditional probability appearing in (0.1) by means of the following limit:

$$P(a \leq X \leq b | Y = y) = \lim_{h \to 0} P(a \leq X \leq b | y - h \leq Y \leq y + h).$$  (0.2)

The right hand side above can be written in terms of $f$ as

$$\lim_{h \to 0} \frac{\int_{y-h}^{y+h} f(a \leq x \leq b | y) \, dv}{\int_{y-h}^{y+h} \left( \int_{-\infty}^{\infty} f(x, v) \, dx \right) \, dv} = \lim_{h \to 0} \frac{(1/2h)f_{Y|X}(y) \int_{y-h}^{y+h} f(x, v) \, dv}{(1/2h)f_Y(y) \int_{y-h}^{y+h} f_Y(v) \, dv}.$$  

If $\int_a^b f(x, v) \, dv$ is continuous in $v$ at $v = y$, the numerator of the last limit converges to $\int_a^b f(x, y) \, dy$ as $h \to 0$. If $f_Y$ is continuous at $y$ the denominator converges to $f_Y(y)$ as $h \to 0$. We are thus lead from (0.2) to

$$P(a \leq X \leq b | Y = y) = \int_a^b f_{X|Y}(x|y) dx$$

which agrees with (0.1). In summary, we have defined conditional densities and conditional probabilities in the absolutely continuous case by analogy with the discrete case. We have also noted that, under some assumptions, a limiting process would yield the same definition of conditional probabilities. It turns out that such limiting procedures are difficult to work with and will not be used any further.

The conditional expectation of $X$ given $Y = y$ is defined to be

$$E[X|Y = y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y) \, dx.$$  

For any function $\phi$ on $\mathbb{R}$, we define

$$E[\phi(X)|Y = y] = \int_{-\infty}^{\infty} \phi(x) f_{X|Y}(x|y) \, dx.$$  

The conditional variance of $X$ given $Y = y$ is defined to be

$$\text{Var}(X|Y = y) = E[X^2|Y = y] - (E[X|Y = y])^2.$$  

It follows from the definition of conditional density that

$$f(x, y) = f_Y(y) f_{X|Y}(x, y), \quad x, y \in \mathbb{R}.$$  

If $X$ and $Y$ are independent and

$$f(x, y) = f_X(x) f_Y(y), \quad x, y \in \mathbb{R}$$  

then

$$f_{X|Y}(x|y) = f_X(x), \quad x \in \mathbb{R}, f_Y(y) > 0.$$  

Conversely, if the equation above holds, then $X$ and $Y$ are independent.
Example 3 Let $X$ and $Y$ be jointly absolutely continuous with joint density

$$f(x, y) = \frac{\sqrt{3}}{4\pi} e^{-\frac{(x^2 - xy + y^2)/2}{2}}, \quad x, y \in \mathbb{R}.$$ 

The marginal density of $Y$ is a normal density with mean 0 and variance $4/3$. Thus for all $x, y \in \mathbb{R},$

$$f_{X|Y}(x|y) = \frac{\sqrt{3}}{2\sqrt{2\pi}} e^{-\frac{(x - y/2)^2}{2}}.$$

In other words, conditional on $Y = y$, $X$ is a normal random variable with parameters $(y/2, 1)$. Thus $E[X|Y = y] = y/2$.

Example 4 Let $Y$ be uniformly distributed over $(0, 1)$ and let $X$ be uniformly distributed over $(0, Y)$. Find the joint density of $X$ and $Y$ and the marginal density of $X$.

Example 5 Let $X$ and $Y$ be jointly absolutely continuous with joint density

$$f(x, y) = \begin{cases} \frac{1}{y} e^{-\frac{x}{y}} e^{-y}, & x > 0, y > 0 \\ 0, & \text{otherwise}. \end{cases}$$

Find $E[X|Y = y]$ for $y > 0$.

**Computing Expectations by Conditioning**

Let us denote by $E[X|Y]$ the function of the random variable $Y$ whose value at $Y = y$ is $E[X|Y = y]$. Note that $E[X|Y]$ is itself a random variable. An extremely important property of condition expectation is the following

**Theorem 6**

$$E[E[X|Y]] = E[X].$$

**Proof.**

Example 7 A miner is trapped in a mine containing 3 doors. The first door leads to a tunnel that will take him to safety after 3 hours of travel, The second door leads to a tunnel that will return him to the mine after 5 hours of travel. The third door leads to a tunnel that will return him to the mine after 7 hours of travel. If we assume that the miner is at all times equally likely to choose any one of the doors, find the expected time for him to get to safety.

**Solution** Let $Y$ denote the door the minor initially chooses. Then we have

$$E[X|Y = 1] = 3, \quad E[X|Y = 2] = 5 + E[X], \quad E[X|Y = 2] = 7 + E[X].$$

Thus by the above theorem we have

$$E[X] = E[E[X|Y]] = E[X|Y = 1] \mathbb{P}(Y = 1) + E[X|Y = 2] \mathbb{P}(Y = 2) + E[X|Y = 3] \mathbb{P}(Y = 3)$$

$$= \frac{1}{3} (E[X|Y = 1] + E[X|Y = 2] + E[X|Y = 2])$$

$$= 5 + \frac{2}{3} E[X].$$

Therefore $E[X] = 15.$
Example 8 Suppose that the number of people entering a department store on a given day is a random variable $N$ with mean 50. Suppose further that the amount of money spent by these customers are independent random variables with a common mean $8$. Suppose also that the amount of money spent by a customer is independent of $N$. Find teh expected amount of money spent in the store on a given day.

Solution Let $X_i$ be the amount spent by the $i$-th customer. Then the total amount of money spent in the store on a given day is $\sum_{i=1}^{N} X_i$. Now

$$E[\sum_{i=1}^{N} X_i] = E \left[ E[\sum_{i=1}^{N} X_i | N] \right].$$

But

$$E[\sum_{i=1}^{N} X_i | N = n] = E[\sum_{i=1}^{n} X_i | N = n] = E[\sum_{i=1}^{n} X_i] = nE[X_1]$$

by the independence of the $X_i$ and $N$. Consequently

$$E[\sum_{i=1}^{N} X_i | N] = NE[X_1].$$

Therefore

$$E[\sum_{i=1}^{N} X_i] = E[NE[X_1]] = E[N]E[X_1] = 400.$$

Remark on Notations