Lecture 5: Conditional Distributions

20 March, 2018

Sunil Kumar Gauttam

Department of Mathematics, LNMIIT

Conditional probability $P(A|B)$, is probability of event A in the new universe (or sample space) B. Now we extend this idea to conditioning one random variable on another in order to give a quantification of dependence of one random variable over the other if the random variables are not independent. We first look at discrete random variables.

Conditional PMF

Definition 5.1 *Let X and Y be two discrete random variable associated with the same random experiment. Then the conditional pmf $f_{X|Y}$ of X given $Y = y$, is defined as*

$$f_{X|Y}(x|y) = \begin{cases} P\{X = x|Y = y\} & \text{if} \quad P\{Y = y\} > 0 \\ 0 & \text{if} \quad P\{Y = y\} = 0 \end{cases}$$

In the original sample space Ω, random variable X has some probability distribution. Now we are told that event $\{Y = y\}$ has occurred. Since X depend on Y, this new information provides partial knowledge about value of X. Hence the probability distribution of X in the new universe determined by the event $\{Y = y\}$ should change. This change is captured by conditional pmf.

A conditional pmf can be thought of as an ordinary pmf over a new universe determined by the conditioning event. For this, note that for fixed y, $f_{X|Y}(x|y) \geq 0$ for all $x \in R_X$. Also if $P(Y = y) > 0$ then

$$\sum_{x \in R_X} f_{X|Y}(x|y) = \sum_{x \in R_X} P(X = x|Y = y) = P\left(\bigcup_{x \in R_X} \{X = x\} \middle| Y = y\right) = P(\Omega|Y = y) = 1$$

If X, Y have joint pmf f, then using the definition of conditional probability we obtain

$$f_{X|Y}(x|y) = \begin{cases} \frac{f(x,y)}{f_Y(y)} & \text{if} \quad f_Y(y) > 0 \\ 0 & \text{if} \quad f_Y(y) = 0 \end{cases}$$

Conditional Distribution Function

Recall that the distribution function F_X of any random variable X (discrete, continuous or mixed) is defined as

$$F_X(x) = P\{X \leq x\}, \quad \forall x \in \mathbb{R}.$$

We define conditional distribution function of X given $Y = y$ as

$$F_{X|Y}(x|y) := P(X \leq x | Y = y).$$

So conditional distribution function is an ordinary (or unconditional) distribution function in new universe determined by the conditioning event.

Recall that if X is a discrete random variable with pmf f_X then

$$F_X(x) = \sum_{t \in R_X : t \leq x} f_X(t).$$

Similarly, if X is a discrete random variable with conditional pmf $f_{X|Y}$ then

$$F_{X|Y}(x|y) = \sum_{t \in R_X : t \leq x} f_{X|Y}(t|y).$$

Recall that if X is a discrete random variable with pmf f_X then and $A \subset \mathbb{R}$ then

$$P(X \in A) = \sum_{x \in A \cap R_X} f_X(x).$$

Similarly, if X is a discrete random variable with conditional pmf $f_{X|Y}$ and $A \subset \mathbb{R}$, then we have

$$P(X \in A | Y = y) = \sum_{x \in R_X \cap A} f_{X|Y}(x|y)$$

Example 5.2 *Let the joint pmf of X and Y is given as follows:*

X \ Y	*-1*	*0*	*1*
-1	0	$\frac{1}{4}$	0
0	$\frac{1}{4}$	0	$\frac{1}{4}$
1	0	$\frac{1}{4}$	0

Then compute the conditional pmf of X given $Y = 0$. Also compute the conditional distribution function of the same.

Solution: Note that $P(Y=0) = \frac{1}{2}$. Hence

$$f_{X|Y}(x|0) = \begin{cases} \frac{1}{2} & \text{if} \quad x = -1, 1 \\ 0 & \text{if} \quad x = 0 \end{cases}$$

Now the conditional distribution function

$$F_{X|Y}(x|0) = \begin{cases} 0 & \text{if} \quad x < -1 \\ \frac{1}{2} & \text{if} \quad -1 \le x < 1 \\ 1 & \text{if} \quad x \ge 1 \end{cases}$$

■

Remark 5.3 *We have said that conditional pmf is a pmf in the new universe determined by the conditioning event. In Example 5.2, the probability distribution of X is*

$$P(X=-1) = P(X=1) = \frac{1}{4}, P(X=0) = \frac{1}{2}.$$

Where as in new universe determined by the event $\{Y=0\}$, the probability distribution of X is revised as

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

Similarly, the distribution function of X is

$$F_X(x) = \begin{cases} 0 & \text{if} \quad x < -1 \\ \frac{1}{4} & \text{if} \quad -1 \le x < 0 \\ \frac{3}{4} & \text{if} \quad 0 \le x < 1 \\ 1 & \text{if} \quad x \ge 1 \end{cases}$$

F_X got revised as $F_{X|Y}(x|0)$ in the new universe determined by the event $\{Y=0\}$. Also note that $F_{X|Y}(x|0)$ satisfies all the properties of a distribution function:

1. $\lim\limits_{x\to-\infty} F_{X|Y}(x|0) = 0, \lim\limits_{x\to+\infty} F_{X|Y}(x|0) = 1.$
2. *$F_{X|Y}(\cdot|0)$ is non-decreasing on $\mathbb{R}$.*
3. *$F_{X|Y}(\cdot|0)$ is right-continuous on $\mathbb{R}$.*

The conditional PMF can also be used to calculate the marginal PMFs. In particular, we have by using the definitions,

$$f_X(x) = \sum_y f(x,y) = \sum_y f_{X|Y}(x|y) f_Y(y)$$

Example 5.4 *Suppose*

$$f_Y(y) = \begin{cases} \frac{5}{6} & \text{if} \quad y = 10^2 \\ \frac{1}{6} & \text{if} \quad y = 10^4 \end{cases}, \quad f_{X|Y}(x|10^2) = \begin{cases} \frac{1}{2} & \text{if} \quad x = 10^{-2} \\ \frac{1}{3} & \text{if} \quad x = 10^{-1} \\ \frac{1}{6} & \text{if} \quad x = 1 \end{cases}, \quad f_{X|Y}(x|10^4) = \begin{cases} \frac{1}{2} & \text{if} \quad x = 1 \\ \frac{1}{3} & \text{if} \quad x = 10 \\ \frac{1}{6} & \text{if} \quad x = 100 \end{cases},$$

Then find the pmf of X.

Solution: First of all by looking at conditional pmf $f_{X|Y}$ we see that X takes 5 values $10^{-2}, 10^{-1}, 1, 10, 100$. Now

$$\begin{aligned} f_X(10^{-2}) &= \frac{1}{2} \times \frac{5}{6} = \frac{5}{12} \\ f_X(10^{-1}) &= \frac{1}{3} \times \frac{5}{6} = \frac{5}{18} \\ f_X(1) &= \frac{1}{6} \times \frac{5}{6} + \frac{1}{2} \times \frac{1}{6} = \frac{8}{36} \\ f_X(10) &= \frac{1}{3} \times \frac{1}{6} = \frac{1}{18} \\ f_X(100) &= \frac{1}{6} \times \frac{1}{6} = \frac{1}{36} \end{aligned}$$

■

Definition 5.5 (Conditional Densities) *Let X and Y be two random variables with joint pdf f. The conditional density of X given $Y = y$ is defined as*

$$f_{X|Y}(x|y) = \begin{cases} \frac{f(x,y)}{f_Y(y)} & \text{if} \quad f_Y(y) > 0 \\ 0 & \text{if} \quad f_Y(y) = 0 \end{cases}$$

As the case of conditional pmf, conditional pdf can be thought of as an ordinary pdf over a new universe determined by the conditioning event. For this, note that for fixed y, $f_{X|Y}(x|y) \geq 0$ for all $x \in \mathbb{R}$. Also if $f_Y(y) > 0$ then

$$\int_{-\infty}^{\infty} f_{X|Y}(x|y)dx = \frac{1}{f_Y(y)} \int_{-\infty}^{\infty} f(x,y)dx = \frac{f_Y(y)}{f_Y(y)} = 1$$

Recall that if X is a continuous random variable with pdf f_X and B is any Borel subset of $\mathbb{R}$, then

$$P(X \in B) = \int_B f_X(x)dx.$$

The above motivated the following definition.

Definition 5.6 *Let X, Y be jointly continuous random variables and $f_{X|Y}(\cdot)$ denotes the conditional density of X given Y. Then for any Borel subset B of $\mathbb{R}$, we have*

$$P(X \in B|Y = y) = \int_B f_{X|Y}(x|y)dx$$

Remark 5.7 *Conditional probability $P(X \in B|Y = y)$ were left undefined if the $P\{Y = y\} = 0$. But the above formula provides a natural way of defining such conditional probabilities in the present context . In addition, it allows us to view the conditional PDF $f_{X|Y}$ (as a function of x) as a description of the probability law of X, given that the event $\{Y = y\}$ has occurred.*