Section 5.5 Conditional distributions

Thursday, November 10, 2016 1:02 PM

Section 5.5
Condition...

5.5 Conditional distributions

RECALL: For events A and B in a non-empty sample space S, we defined

$$P(A|B) = \frac{P(A \cap B)}{P(B)},$$

for $P(B) > 0$. Now, suppose that (Y_1, Y_2) is a discrete random vector. If we let $B = \{Y_2 = y_2\}$ and $A = \{Y_1 = y_1\}$, we obtain

$$P(A|B) = \frac{P(Y_1 = y_1, Y_2 = y_2)}{P(Y_2 = y_2)} = \frac{p_{Y_1,Y_2}(y_1, y_2)}{p_{Y_2}(y_2)}.$$

This leads to the definition of a discrete conditional distribution.

TERMINOLOGY: Suppose that (Y_1, Y_2) is a discrete random vector with joint pmf $p_{Y_1,Y_2}(y_1, y_2)$. We define the **conditional probability mass function (pmf)** of Y_1, given $Y_2 = y_2$, as

$$p_{Y_1|Y_2}(y_1|y_2) = \frac{p_{Y_1,Y_2}(y_1, y_2)}{p_{Y_2}(y_2)},$$

whenever $p_{Y_2}(y_2) > 0$. Similarly, the conditional probability mass function of Y_2, given $Y_1 = y_1$, is

$$p_{Y_2|Y_1}(y_2|y_1) = \frac{p_{Y_1,Y_2}(y_1, y_2)}{p_{Y_1}(y_1)},$$

whenever $p_{Y_1}(y_1) > 0$.

Example 5.4. The joint pmf of (Y_1, Y_2) in Example 5.1 is depicted below (in the inner rectangular part of the table). The marginal distributions of Y_1 and Y_2 are catalogued in the margins of the table.

$p_{Y_1,Y_2}(y_1, y_2)$	$y_2 = 0$	$y_2 = 1$	$y_2 = 2$	$y_2 = 3$	$p_{Y_1}(y_1)$
$y_1 = 0$	0.12	0.06	0.05	0.02	0.25
$y_1 = 1$	0.13	0.15	0.12	0.03	0.43
$y_1 = 2$	0.05	0.15	0.10	0.02	0.32
$p_{Y_2}(y_2)$	0.30	0.36	0.27	0.07	1

$P(Y_1 = 1, Y_2 = 1) = 0.15$

$P(Y_2 = 1) = 0.36$

$P(Y_1 = 1 | Y_2 = 1)$

$= \frac{P(Y_1 = 1 \cap Y_2 = 1)}{P(Y_2 = 1)}$

$= \frac{0.15}{0.36}$

QUESTION: What is the conditional pmf of Y_1, given $Y_2 = 1$?

$P(Y_1 = 1 | Y_2 = 1) = \frac{15}{36}$

$P(Y_1 = 0 | Y_2 = 1) = \frac{6}{36}$

$P(Y_1 = 2 | Y_2 = 1) = \frac{15}{36}$

SOLUTION. Straightforward calculations show that

$$\begin{aligned}
p_{Y_1|Y_2}(y_1 = 0|y_2 = 1) &= \frac{p_{Y_1,Y_2}(y_1 = 0, y_2 = 1)}{p_{Y_2}(y_2 = 1)} = \frac{0.06}{0.36} = 2/12 \\
p_{Y_1|Y_2}(y_1 = 1|y_2 = 1) &= \frac{p_{Y_1,Y_2}(y_1 = 1, y_2 = 1)}{p_{Y_2}(y_2 = 1)} = \frac{0.15}{0.36} = 5/12 \\
p_{Y_1|Y_2}(y_1 = 2|y_2 = 1) &= \frac{p_{Y_1,Y_2}(y_1 = 2, y_2 = 1)}{p_{Y_2}(y_2 = 1)} = \frac{0.15}{0.36} = 5/12.
\end{aligned}$$

Thus, the conditional pmf of Y_1, given $Y_2 = 1$, is given by

y_1	0	1	2
$p_{Y_1\|Y_2}(y_1\|y_2 = 1)$	2/12	5/12	5/12

This conditional pmf tells us how Y_1 is distributed if we are given that $Y_2 = 1$.

EXERCISE. Find the conditional pmf of Y_2, given $Y_1 = 0$. □

TERMINOLOGY: Suppose that (Y_1, Y_2) is a continuous random vector with joint pdf $f_{Y_1,Y_2}(y_1, y_2)$. We define the **conditional probability density function (pdf)** of Y_1, given $Y_2 = y_2$, as

$$f_{Y_1|Y_2}(y_1|y_2) = \frac{f_{Y_1,Y_2}(y_1, y_2)}{f_{Y_2}(y_2)}.$$

$$P_{Y_1|Y_2}(y_1|y_2) = \frac{P_{Y_1,Y_2}(y_1,y_2)}{P_{Y_2}(y_2)}$$

Similarly, the conditional probability density function of Y_2, given $Y_1 = y_1$, is

$$f_{Y_2|Y_1}(y_2|y_1) = \frac{f_{Y_1,Y_2}(y_1, y_2)}{f_{Y_1}(y_1)}.$$

Example 5.5. Consider the bivariate pdf in Example 5.3,

$$f_{Y_1,Y_2}(y_1, y_2) = \begin{cases} 6y_1, & 0 < y_1 < y_2 < 1 \\ 0, & \text{otherwise.} \end{cases}$$

This model describes the distribution of the random vector (Y_1, Y_2), where Y_1, the proportion of a population with trait 1, is always less than Y_2, the proportion of a population with trait 2. Derive the conditional distributions $f_{Y_1|Y_2}(y_1|y_2)$ and $f_{Y_2|Y_1}(y_2|y_1)$.

SOLUTION. In Example 5.3, we derived the marginal pdfs to be

$$f_{Y_1}(y_1) = \begin{cases} 6y_1(1 - y_1), & 0 < y_1 < 1 \\ 0, & \text{otherwise} \end{cases}$$

$$\begin{aligned}
f_{Y_1}(y_1) &= \int_{y_1}^{1} f_{Y_1,Y_2}(y_1,y_2)\,dy_2 \\
&= \int_{y_1}^{1} 6y_1\,dy_2 \\
&= 6y_1 \times \left(y_2 \Big|_{y_1}^{1}\right) \\
&= 6y_1 \times (1-y_1)
\end{aligned}$$

and

$$f_{Y_2}(y_2) = \begin{cases} 3y_2^2, & 0 < y_2 < 1 \\ 0, & \text{otherwise.} \end{cases}$$

$\int_0^{y_2} f(y_1, y_2)\, dy_1$
$= \int_0^{y_2} 6y_1\, dy_1 = 3y_1^2 \Big|_0^{y_2}$
$= 3y_2^2$

First, we derive $f_{Y_1|Y_2}(y_1|y_2)$, so fix $Y_2 = y_2$. Remember, once we condition on $Y_2 = y_2$ (i.e., once we fix $Y_2 = y_2$), we then regard y_2 as simply a constant. **This is an important point to understand!** For values of $0 < y_1 < y_2$, it follows that

$$f_{Y_1|Y_2}(y_1|y_2) = \frac{f_{Y_1,Y_2}(y_1, y_2)}{f_{Y_2}(y_2)} = \frac{6y_1}{3y_2^2} = \frac{2y_1}{y_2^2},$$

$y_2 = \frac{1}{2}$. $Y_1 | Y_2 = \frac{1}{2}$

$f_{Y_1|Y_2}(y_1|\frac{1}{2}) = \frac{2y_1}{(\frac{1}{2})^2}$

$= 8y_1$

$0 < y_1 < \frac{1}{2}$

and, thus, this is the value of $f_{Y_1|Y_2}(y_1|y_2)$ when $0 < y_1 < y_2$. For values of $y_1 \notin (0, y_2)$, the conditional density $f_{Y_1|Y_2}(y_1|y_2) = 0$. Summarizing,

$$f_{Y_1|Y_2}(y_1|y_2) = \begin{cases} 2y_1/y_2^2, & 0 < y_1 < y_2 \\ 0, & \text{otherwise.} \end{cases}$$

To reiterate, in this (conditional) pdf, the value of y_2 is fixed and known. It is Y_1 that is varying. This function describes how Y_1 is distributed for y_2 fixed.

Now, to derive the conditional pdf of Y_2 given Y_1, we fix $Y_1 = y_1$; then, for all values of $y_1 < y_2 < 1$, we have

$$f_{Y_2|Y_1}(y_2|y_1) = \frac{f_{Y_1,Y_2}(y_1, y_2)}{f_{Y_1}(y_1)} = \frac{6y_1}{6y_1(1 - y_1)} = \frac{1}{1 - y_1}.$$

This is the value of $f_{Y_2|Y_1}(y_2|y_1)$ when $y_1 < y_2 < 1$. When $y_2 \notin (y_1, 1)$, the conditional pdf is $f_{Y_2|Y_1}(y_2|y_1) = 0$. Remember, once we condition on $Y_1 = y_1$, then we regard y_1 simply as a constant. Summarizing,

$$f_{Y_2|Y_1}(y_2|y_1) = \begin{cases} \frac{1}{1-y_1}, & y_1 < y_2 < 1 \\ 0, & \text{otherwise.} \end{cases}$$

That is, conditional on $Y_1 = y_1$, $Y_2 \sim \mathcal{U}(y_1, 1)$. **Again, in this (conditional) pdf, the value of y_1 is fixed and known. It is Y_2 that is varying. This function describes how Y_2 is distributed for y_1 fixed.** $\square$

RESULT: The use of conditional distributions allows us to define conditional probabilities of events associated with one random variable when we know the value of another random variable. If Y_1 and Y_2 are jointly **discrete**, then for any set $B \subset \mathcal{R}$,

$$\begin{aligned} P(Y_1 \in B | Y_2 = y_2) &= \sum_B p_{Y_1|Y_2}(y_1|y_2) \\ P(Y_2 \in B | Y_1 = y_1) &= \sum_B p_{Y_2|Y_1}(y_2|y_1). \end{aligned}$$

If Y_1 and Y_2 are jointly **continuous**, then for any set $B \subset \mathcal{R}$,

$$\begin{aligned} P(Y_1 \in B | Y_2 = y_2) &= \int_B f_{Y_1|Y_2}(y_1|y_2) dy_1 \\ P(Y_2 \in B | Y_1 = y_1) &= \int_B f_{Y_2|Y_1}(y_2|y_1) dy_2. \end{aligned}$$

Example 5.6. A health-food store stocks two different brands of grain. Let Y_1 denote the amount of brand 1 in stock and let Y_2 denote the amount of brand 2 in stock (both Y_1 and Y_2 are measured in 100s of lbs). The joint distribution of Y_1 and Y_2 is given by

$$f_{Y_1,Y_2}(y_1, y_2) = \begin{cases} 24y_1y_2, & y_1 > 0,\ y_2 > 0,\ 0 < y_1 + y_2 < 1 \\ 0, & \text{otherwise.} \end{cases}$$

(a) Find the conditional pdf $f_{Y_1|Y_2}(y_1|y_2)$.

(b) Compute $P(Y_1 > 0.5 | Y_2 = 0.3)$.

(c) Find $P(Y_1 > 0.5)$.

SOLUTIONS. (a) To find the conditional pdf $f_{Y_1|Y_2}(y_1|y_2)$, we first need to find the marginal pdf of Y_2. The marginal pdf of Y_2, for $0 < y_2 < 1$, is

$$f_{Y_2}(y_2) = \int_{y_1=0}^{1-y_2} 24y_1y_2\, dy_1 = 24y_2 \left(\left. \frac{y_1^2}{2} \right|_0^{1-y_2} \right) = 12y_2(1-y_2)^2,$$

and 0, otherwise. We recognize this as a beta(2, 3) pdf; i.e., $Y_2 \sim$ beta(2, 3). The conditional pdf of Y_1, given $Y_2 = y_2$, is

$$\begin{aligned} f_{Y_1|Y_2}(y_1|y_2) = \frac{f_{Y_1,Y_2}(y_1, y_2)}{f_{Y_2}(y_2)} &= \frac{24y_1y_2}{12y_2(1-y_2)^2} \\ &= \frac{2y_1}{(1-y_2)^2}, \end{aligned}$$

for $0 < y_1 < 1 - y_2$, and 0, otherwise. Summarizing,

$$f_{Y_1|Y_2}(y_1|y_2) = \begin{cases} \frac{2y_1}{(1-y_2)^2}, & 0 < y_1 < 1 - y_2 \\ 0, & \text{otherwise.} \end{cases}$$

(b) To compute $P(Y_1 > 0.5|Y_2 = 0.3)$, we work with the conditional pdf $f_{Y_1|Y_2}(y_1|y_2)$, which for $y_2 = 0.3$, is given by

$$f_{Y_1|Y_2}(y_1|y_2) = \begin{cases} \left(\frac{200}{49}\right) y_1, & \boxed{0 < y_1 < 0.7} \\ 0, & \text{otherwise.} \end{cases}$$

$f_{Y_1|Y_2}(y_1|0.3)$.

$0 < y_1 < 1 - y_2$

Thus,

$$P(Y_1 > 0.5|Y_2 = 0.3) = \int_{0.5}^{0.7} \left(\frac{200}{49}\right) y_1 dy_1 \approx 0.489.$$

(c) To compute $P(Y_1 > 0.5)$, we can either use the marginal pdf $f_{Y_1}(y_1)$ or the joint pdf $f_{Y_1,Y_2}(y_1, y_2)$. Marginally, it turns out that $Y_1 \sim \text{beta}(2, 3)$ as well (verify!). Thus,

$$P(Y_1 > 0.5) = \int_{0.5}^{1} 12y_1(1 - y_1)^2 dy_1 \approx 0.313.$$

REMARK: Notice how $P(Y_1 > 0.5|Y_2 = 0.3) \neq P(Y_1 > 0.5)$; that is, knowledge of the value of Y_2 has affected the way that we assign probability to events involving Y_1. Of course, one might expect this because of the support in the joint pdf $f_{Y_1,Y_2}(y_1, y_2)$. □

5.6 Independent random variables

TERMINOLOGY: Suppose (Y_1, Y_2) is a random vector (discrete or continuous) with joint cdf $F_{Y_1,Y_2}(y_1, y_2)$, and denote the marginal cdfs of Y_1 and Y_2 by $F_{Y_1}(y_1)$ and $F_{Y_2}(y_2)$, respectively. We say the random variables Y_1 and Y_2 are **independent** if and only if

$$F_{Y_1,Y_2}(y_1, y_2) = F_{Y_1}(y_1)F_{Y_2}(y_2)$$

for all values of y_1 and y_2. Otherwise, we say that Y_1 and Y_2 are **dependent**.

RESULT: Suppose that (Y_1, Y_2) is a random vector (discrete or continuous) with joint pdf (pmf) $f_{Y_1,Y_2}(y_1, y_2)$, and denote the marginal pdfs (pmfs) of Y_1 and Y_2 by $f_{Y_1}(y_1)$ and $f_{Y_2}(y_2)$, respectively. Then, Y_1 and Y_2 are independent if and only if

$$f_{Y_1,Y_2}(y_1, y_2) = f_{Y_1}(y_1)f_{Y_2}(y_2)$$

for all values of y_1 and y_2. Otherwise, Y_1 and Y_2 are dependent.

Proof. Exercise. □