Andrew Gard - equitable.equations@gmail.com

Conditional distributions

Suppose we have two random variables X and Y with some joint probability distribution. The **conditional distribution** of X when $Y = y$ (where y is a specific value of Y) is just the corresponding single-variable probability distribution of X when Y is fixed at the specified value.

Suppose we have two random variables X and Y with some joint probability distribution. The **conditional distribution** of X when $Y = y$ (where y is a specific value of Y) is just the corresponding single-variable probability distribution of X when Y is fixed at the specified value.

In this vid, we'll focus on the discrete case. The math is nearly identical for continuous random variables, with sums replaced by definite integrals.

If X and Y are discrete random variables with joint probability mass function $f(x, y)$, then the **conditional probability mass function** of X given $Y = y$ is

$$g(x \mid y) = \frac{f(x, y)}{f_Y(y)}$$

where $f_Y(y)$ is the marginal pmf of Y evaluated at $Y = y$. This is just $P(Y = y)$.

If X and Y are discrete random variables with joint probability mass function $f(x, y)$, then the **conditional probability mass function** of X given $Y = y$ is

$$g(x \mid y) = \frac{f(x, y)}{f_Y(y)}$$

where $f_Y(y)$ is the marginal pmf of Y evaluated at $Y = y$. This is just $P(Y = y)$.

Similarly, $h(y \mid x) = \dfrac{f(x, y)}{f_X(x)}$ is the conditional pmf of Y given $X = x$.

Example. Suppose the random variables X and Y have the following joint probability distribution. Compute $g(x \mid 3)$ and $h(y \mid 3)$.

$f(x,y)$	y			
x	1	2	3	4
1	-	0.12	0.10	0.09
2	-	0.18	0.07	-
3	0.02	0.14	0.11	-
4	0.05	0.12	-	-

$f(x,y)$	y			
x	1	2	3	4
1	-	0.12	0.10	0.09
2	-	0.18	0.07	-
3	0.02	0.14	0.11	-
4	0.05	0.12	-	-

First let's compute $g(x \mid 3)$. This means computing all $P(X = x \mid Y = 3)$ for all possible values of x.

$f(x,y)$	y			
x	1	2	3	4
1	-	0.12	0.10	0.09
2	-	0.18	0.07	-
3	0.02	0.14	0.11	-
4	0.05	0.12	-	-

First let's compute $g(x \mid 3)$. This means computing all $P(X = x \mid Y = 3)$ for all possible values of x.

$$g(1 \mid 3) = \frac{f(1,3)}{f_Y(3)} = \frac{P(X = 1 \text{ and } Y = 3)}{P(Y = 3)}$$

$f(x, y)$	y			
x	1	2	3	4
1	-	0.12	0.10	0.09
2	-	0.18	0.07	-
3	0.02	0.14	0.11	-
4	0.05	0.12	-	-

First let's compute $g(x \mid 3)$. This means computing all $P(X = x \mid Y = 3)$ for all possible values of x.

$$\begin{aligned} g(1 \mid 3) &= \frac{f(1,3)}{f_Y(3)} = \frac{P(X = 1 \text{ and } Y = 3)}{P(Y = 3)} \\ &= \frac{0.10}{(0.10 + 0.07 + 0.11)} \end{aligned}$$

$f(x,y)$	y			
x	1	2	3	4
1	-	0.12	0.10	0.09
2	-	0.18	0.07	-
3	0.02	0.14	0.11	-
4	0.05	0.12	-	-

First let's compute $g(x \mid 3)$. This means computing all $P(X = x \mid Y = 3)$ for all possible values of x.

$$
\begin{aligned}
g(1 \mid 3) &= \frac{f(1,3)}{f_Y(3)} = \frac{P(X = 1 \text{ and } Y = 3)}{P(Y = 3)} \\
&= \frac{0.10}{(0.10 + 0.07 + 0.11)} = .357
\end{aligned}
$$

$f(x,y)$	y			
x	1	2	3	4
1	-	0.12	0.10	0.09
2	-	0.18	0.07	-
3	0.02	0.14	0.11	-
4	0.05	0.12	-	-

Similarly,

$$P(X = 2 \mid Y = 3) \quad = \quad \frac{P(X = 2 \text{ and } Y = 3)}{P(Y = 3)}$$

$f(x, y)$	y			
x	1	2	3	4
1	-	0.12	0.10	0.09
2	-	0.18	0.07	-
3	0.02	0.14	0.11	-
4	0.05	0.12	-	-

Similarly,

$$P(X = 2 \,|\, Y = 3) \;=\; \frac{P(X = 2 \text{ and } Y = 3)}{P(Y = 3)} \;=\; \frac{0.07}{(0.10 + 0.07 + 0.11)} = .250$$

$f(x, y)$	y			
x	1	2	3	4
1	-	0.12	0.10	0.09
2	-	0.18	0.07	-
3	0.02	0.14	0.11	-
4	0.05	0.12	-	-

Similarly,

$$P(X = 2 \mid Y = 3) = \frac{P(X = 2 \text{ and } Y = 3)}{P(Y = 3)} = \frac{0.07}{(0.10 + 0.07 + 0.11)} = .250$$

$$P(X = 3 \mid Y = 3) = \frac{P(X = 2 \text{ and } Y = 3)}{P(Y = 3)}$$

$f(x,y)$	y			
x	1	2	3	4
1	-	0.12	0.10	0.09
2	-	0.18	0.07	-
3	0.02	0.14	0.11	-
4	0.05	0.12	-	-

Similarly,

$$
\begin{aligned}
P(X = 2 \,|\, Y = 3) &= \frac{P(X = 2 \text{ and } Y = 3)}{P(Y = 3)} = \frac{0.07}{(0.10 + 0.07 + 0.11)} = .250 \\
P(X = 3 \,|\, Y = 3) &= \frac{P(X = 2 \text{ and } Y = 3)}{P(Y = 3)} \\
&= \frac{0.07}{(0.10 + 0.07 + 0.11)} = .393
\end{aligned}
$$

$f(x, y)$	y			
x	1	2	3	4
1	-	0.12	0.10	0.09
2	-	0.18	0.07	-
3	0.02	0.14	0.11	-
4	0.05	0.12	-	-

Overall, the conditional probability mass function for X when $Y = 3$ is given by

x	1	2	3
$g(x \mid 3)$	35.7%	25.0%	39.3%

$f(x, y)$	y			
x	1	2	3	4
1	-	0.12	0.10	0.09
2	-	0.18	0.07	-
3	0.02	0.14	0.11	-
4	0.05	0.12	-	-

Overall, the conditional probability mass function for X when $Y = 3$ is given by

x	1	2	3
$g(x \mid 3)$	35.7%	25.0%	39.3%

These probabilities sum to 100%. A conditional pmf is a pmf in its own right.

$f(x, y)$	y			
x	1	2	3	4
1	-	0.12	0.10	0.09
2	-	0.18	0.07	-
3	0.02	0.14	0.11	-
4	0.05	0.12	-	-

Let's do $h(y \mid 3)$ a bit faster.

$f(x, y)$	y			
x	1	2	3	4
1	-	0.12	0.10	0.09
2	-	0.18	0.07	-
3	0.02	0.14	0.11	-
4	0.05	0.12	-	-

Let's do $h(y \mid 3)$ a bit faster. We need to compute $f(3, y)/f_X(3)$ for $y = 1, 2,$ and 3, which means looking at the $X = 3$ row of the table.

$f(x,y)$	y			
x	1	2	3	4
1	-	0.12	0.10	0.09
2	-	0.18	0.07	-
3	0.02	0.14	0.11	-
4	0.05	0.12	-	-

Let's do $h(y \mid 3)$ a bit faster. We need to compute $f(3, y)/f_X(3)$ for $y = 1, 2$, and 3, which means looking at the $X = 3$ row of the table. This row has total probability $f_X(3) = 0.02 + 0.14 + 0.11 = 0.27$.

$f(x,y)$	y			
x	1	2	3	4
1	-	0.12	0.10	0.09
2	-	0.18	0.07	-
3	0.02	0.14	0.11	-
4	0.05	0.12	-	-

Let's do $h(y \mid 3)$ a bit faster. We need to compute $f(3,y)/f_X(3)$ for $y = 1, 2$, and 3, which means looking at the $X = 3$ row of the table. This row has total probability $f_X(3) = 0.02 + 0.14 + 0.11 = 0.27$.

y	1	2	3
$h(y \mid 3)$	$0.02/0.27$	$0.14/0.27$	$0.11/0.27$

$f(x,y)$	y			
x	1	2	3	4
1	-	0.12	0.10	0.09
2	-	0.18	0.07	-
3	0.02	0.14	0.11	-
4	0.05	0.12	-	-

Let's do $h(y \mid 3)$ a bit faster. We need to compute $f(3,y)/f_X(3)$ for $y = 1, 2$, and 3, which means looking at the $X = 3$ row of the table. This row has total probability $f_X(3) = 0.02 + 0.14 + 0.11 = 0.27$.

y	1	2	3
$h(y \mid 3)$	$0.02/0.27$	$0.14/0.27$	$0.11/0.27$
	7.4%	51.9%	40.7%

Conditional means, variances, and standard deviations are computed in the natural way.

Conditional means, variances, and standard deviations are computed in the natural way. Here are the formulas when $Y = y$ is fixed and $g(X|y)$ is the corresponding conditional probability mass function for X.

$$E(X|y) \;=\; \mu_{X|y} \;=\; \sum_x x\, g(x|y)$$

Conditional means, variances, and standard deviations are computed in the natural way. Here are the formulas when $Y = y$ is fixed and $g(X|y)$ is the corresponding conditional probability mass function for X.

$$
\begin{aligned}
E(X|y) &= \mu_{X|y} &= \sum_{x} x\, g(x|y) \\
Var(X|y) &= \sigma^2_{X|y} &= \sum_{x} \left[x - \mu_{X|y}\right]^2 g(x|y)
\end{aligned}
$$

Conditional means, variances, and standard deviations are computed in the natural way. Here are the formulas when $Y = y$ is fixed and $g(X|y)$ is the corresponding conditional probability mass function for X.

$$\begin{aligned}
E(X|y) &= \mu_{X|y} &&= \sum_x x\, g(x|y) \\
Var(X|y) &= \sigma^2_{X|y} &&= \sum_x \left[x - \mu_{X|y}\right]^2 g(x|y) \\
& &&= E(X^2|y) - [E(X|y)]^2
\end{aligned}$$

Conditional means, variances, and standard deviations are computed in the natural way. Here are the formulas when $Y = y$ is fixed and $g(X|y)$ is the corresponding conditional probability mass function for X.

$$\begin{aligned} E(X|y) &= \mu_{X|y} &= \sum_x x\, g(x|y) \\ Var(X|y) &= \sigma^2_{X|y} &= \sum_x \left[x - \mu_{X|y}\right]^2 g(x|y) \\ & &= E(X^2|y) - [E(X|y)]^2 \end{aligned}$$

These are just the regular single-variable formulas applied to the variable X once the value of Y has been fixed.

$f(x, y)$		y		
x	1	2	3	4
1	-	0.12	0.10	0.09
2	-	0.18	0.07	-
3	0.02	0.14	0.11	-
4	0.05	0.12	-	-

Let's compute the conditional mean and variance of X when $Y = 3$.

$f(x, y)$	y			
x	1	2	3	4
1	-	0.12	0.10	0.09
2	-	0.18	0.07	-
3	0.02	0.14	0.11	-
4	0.05	0.12	-	-

Let's compute the conditional mean and variance of X when $Y = 3$.

x	1	2	3
$g(x \mid 3)$	.35.7%	25.0%	39.3%

$f(x,y)$	y			
x	1	2	3	4
1	-	0.12	0.10	0.09
2	-	0.18	0.07	-
3	0.02	0.14	0.11	-
4	0.05	0.12	-	-

Let's compute the conditional mean and variance of X when $Y = 3$.

x	1	2	3
$g(x \mid 3)$	.35.7%	25.0%	39.3%

$$\mu_{X|3} \quad = \quad (1)(0.357) + (2)(0.250) + (3)(0.393)$$

$f(x,y)$	y			
x	1	2	3	4
1	-	0.12	0.10	0.09
2	-	0.18	0.07	-
3	0.02	0.14	0.11	-
4	0.05	0.12	-	-

Let's compute the conditional mean and variance of X when $Y = 3$.

x	1	2	3
$g(x \mid 3)$	.35.7%	25.0%	39.3%

$$\mu_{X|3} \;=\; (1)(0.357) + (2)(0.250) + (3)(0.393) \;=\; 2.04$$

$f(x, y)$	y			
x	1	2	3	4
1	-	0.12	0.10	0.09
2	-	0.18	0.07	-
3	0.02	0.14	0.11	-
4	0.05	0.12	-	-

Let's compute the conditional mean and variance of X when $Y = 3$.

x	1	2	3
$g(x \mid 3)$	.35.7%	25.0%	39.3%

$$
\begin{aligned}
\mu_{X|3} &= (1)(0.357) + (2)(0.250) + (3)(0.393) = 2.04 \\
\sigma^2_{X|3} &= \sum_x (x - 2.04)^2 g(x|3)
\end{aligned}
$$

$f(x,y)$	y			
x	1	2	3	4
1	-	0.12	0.10	0.09
2	-	0.18	0.07	-
3	0.02	0.14	0.11	-
4	0.05	0.12	-	-

Let's compute the conditional mean and variance of X when $Y = 3$.

x	1	2	3
$g(x \mid 3)$	.35.7%	25.0%	39.3%

$$
\begin{aligned}
\mu_{X|3} &= (1)(0.357) + (2)(0.250) + (3)(0.393) = 2.04 \\
\sigma^2_{X|3} &= \sum_x (x - 2.04)^2 g(x|3) = 0.75
\end{aligned}
$$