Lecture 6.2

Multiple Random Variables:

Conditional distributions

Conditional Probability

Recall: IF $\mathbb{P}(B) > 0$, the conditional probability of A given B is

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Suppose now we have two discrete random variables X and Y. We can use the above concept to obtain conditional probability $\mathbb{P}(X = x \mid Y = y)$. By defining the events

$$A = \{X = x\} \quad \text{and} \quad B = \{Y = y\},$$

we use the definition of the conditional probability of A given B to get

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

Conditional Probability

Consider the joint probability mass function of (X, Y) below. Both X and Y only take the values $\{0, 1, 2\}$.

		y		
		0	1	2
	0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{16}$
x	1	0	$\frac{1}{4}$	$\frac{1}{8}$
	2	0	0	$\frac{1}{16}$

What is $\mathbb{P}(X = 0 \,|\, Y = 2)$?

$$\mathbb{P}(X = 0 \,|\, Y = 2) = \frac{\mathbb{P}(X = 0,\ Y = 2)}{\mathbb{P}(Y = 2)} = \frac{\frac{1}{16}}{\frac{1}{16} + \frac{1}{8} + \frac{1}{16}} = \frac{1}{4}.$$

Conditional Probability Mass Function

Similarly, we can show

$$\mathbb{P}(X = 1 \,|\, Y = 2) = \frac{\frac{1}{8}}{\frac{1}{16} + \frac{1}{8} + \frac{1}{16}} = \frac{1}{2},$$

and

$$\mathbb{P}(X = 2 \,|\, Y = 2) = \frac{\frac{1}{16}}{\frac{1}{16} + \frac{1}{8} + \frac{1}{16}} = \frac{1}{4}.$$

These three conditional probabilities form the **conditional pmf** of X given $Y = 2$:

x	0	1	2
$\mathbb{P}(X = x \,\|\, Y = 2)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Conditional Probability Mass Function

As we saw, for jointly discrete (X, Y), the concept of a conditional pmf is an extension of the concept of conditional probability of an event.

Suppose X and Y are both discrete with joint pmf $f_{X,Y}$, and suppose $f_Y(y) > 0$. The **conditional pmf** of X given $Y = y$ is defined as

$$f_{X\,|\,Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}, \quad \text{for all } x.$$

Conditional Probability Mass Function

For each y, the conditional pmf $f_{X\,|\,Y}(x|y)$ is a genuine pmf, i.e., it is non-negative and its sum over all x is one.

Rewriting this, we find the "product rule" for pmfs:

$$f_{X,Y}(x,y) = f_Y(y) f_{X\,|\,Y}(x|y).$$

Summing over y we obtain the marginal pmf of X:

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

which is the "the law of total probability" for pmfs.

Example

Suppose $Y \sim \text{Bin}(2, \frac{1}{2})$, and conditional on $Y = y$, X has a $\text{Bin}(y, \frac{1}{2})$ distribution. What is the joint pmf of (X, Y)?

For $x, y \in \{0, 1, 2\}, x \leq y$, the joint pmf of (X, Y) is

$$\begin{aligned} f_{X,Y}(x,y) &= f_{X\,|\,Y}(x|y) f_Y(y) \\ &= \binom{y}{x}(0.5)^x(1-0.5)^{y-x} \cdot \binom{2}{y}(0.5)^y(1-0.5)^{2-y} \end{aligned}$$

		y		
		0	1	2
	0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{16}$
x	1	0	$\frac{1}{4}$	$\frac{1}{8}$
	2	0	0	$\frac{1}{16}$

Conditional Probability Density Functions

For continuous random variable (X, Y), we can no longer use the concept of a conditional probability of an event since $\mathbb{P}(Y = y) = 0$ for any y.

So, instead we directly extend the concept of conditional pmf by considering conditional pdf.

Suppose X and Y are both *continuous* with joint *pdf* $f_{X,Y}$, and suppose $f_Y(y) > 0$. The **conditional pdf** of X given $Y = y$ is defined as

$$f_{X\,|\,Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}, \quad \text{for all } x.$$

Conditional Probability Density Functions

For each y, the conditional pdf $f_{X\,|\,Y}(x|y)$ is a genuine pdf.

For values of y such that $f(y) = 0$, we are free to define $f_{X\,|\,Y}(x|y)$ however we wish, so long as $f_{X\,|\,Y}(x|y)$ is a pdf as a function of x.

Rewriting conditional pdf, we find the "product rule" for pdfs:

$$f_{X,Y}(x, y) = f_Y(y) f_{X\,|\,Y}(x|y).$$

Integrating over y we obtain the marginal *pdf* of X:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dy = \int_{-\infty}^{\infty} f_Y(y) f_{X\,|\,Y}(x|y)\, dy,$$

which is again a reminiscent of the "the law of total probability".

Example

A pair of random variables (X, Y) has a joint distribution in which Y has marginal pdf given by

$$f_Y(y) = \begin{cases} 3y^2, & y \in [0,1] \\ 0, & \text{else,} \end{cases}$$

and the conditional pdf of X given $\{Y = y\}$ is uniform on the interval $[0, y]$.

Q1: Write down the joint probability density function of (X, Y), clearly specifying where the joint pdf is non-zero.

Q2: Determine the marginal pdf of X.

Example (continued)

Q1: Write down the joint probability density function of (X, Y), clearly specifying where the joint pdf is non-zero.

The conditional pdf of X given $\{Y = y\}$ is

$$f_{X|Y}(x \mid y) = \begin{cases} \dfrac{1}{y}, & 0 \leq x \leq y \\ 0, & \text{else.} \end{cases}$$

So, the joint probability density function of (X, Y) is given by

$$f_{X,Y}(x, y) = f_{X|Y}(x \mid y) f_Y(y) = \begin{cases} 3y = 3y^2 \times \dfrac{1}{y}, & 0 \leq x \leq y \leq 1 \\ 0, & \text{else.} \end{cases}$$

Example (continued)

Q2: Determine the marginal pdf of X.

The marginal pdf of X is

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}y.$$

So, for $x \in [0,1]$,

$$f_X(x) = \int_{-\infty}^{x} 0 \, \mathrm{d}y + \int_{x}^{1} 3y \, \mathrm{d}y + \int_{1}^{\infty} 0 \, \mathrm{d}y = \frac{3}{2}(1 - x^2).$$

Also, for $x \notin [0,1]$, $f_X(x) = 0$.

Conditional Expectation

As conditional distributions are genuine distributions, we can also consider expectations with respect to these conditional distributions.

Conditional expectation of X given $Y = y$:

- If X and Y are both discrete, then

$$\mathbb{E}(X \,|\, Y = y) = \sum_{x} x \, f_{X\,|\,Y}(x|y).$$

- If X and Y are both continuous and have a joint pdf, then

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

Example

A pair of random variables (X, Y) has a joint distribution in which Y has marginal pdf given by

$$f_Y(y) = \begin{cases} 3y^2, & y \in [0, 1] \\ 0, & \text{else}, \end{cases}$$

and the conditional pdf of X given $\{Y = y\}$ is uniform on the interval $[0, y]$.

Q: What is the conditional expectation of X given $Y = y$?

Example (continued)

For any $y \in [0, 1]$,

$$
\begin{aligned}
\mathbb{E}(X \mid Y = y) &= \int_{-\infty}^{\infty} x \, f_{X|Y}(x \mid y) \, \mathrm{d}x \\
&= \int_{-\infty}^{0} x \cdot 0 \, \mathrm{d}x + \int_{0}^{y} x \cdot \frac{1}{y} \, \mathrm{d}x + \int_{y}^{\infty} x \cdot 0 \, \mathrm{d}x = \frac{y}{2}.
\end{aligned}
$$

Tower Property

The conditional expectation $\mathbb{E}(X \,|\, Y = y)$ is a function of y. Call this function $h(y)$. So, $h(Y)$ is a random variable which takes on values $h(y)$. This random variable is denoted as $\mathbb{E}(X \mid Y)$ and is called the **conditional expectation of X given Y**.

What is $\mathbb{E}(\mathbb{E}(X \mid Y))$? Suppose X and Y are discrete. Then

$$h(y) = \mathbb{E}(X \,|\, Y = y) = \sum_x x \, f_{X \,|\, Y}(x, y).$$

So

$$\begin{aligned}
\mathbb{E}(\mathbb{E}(X \mid Y)) = \mathbb{E}h(Y) = \sum_y h(y) \, f_Y(y) &= \sum_y \left[\sum_x x \, f_{X \,|\, Y}(x, y) \right] f_Y(y) \\
&= \sum_y \sum_x x \, f_Y(y) \, f_{X \,|\, Y}(x|y) = \sum_y \sum_x x \, f_{X,Y}(x, y) \\
&= \sum_x x \sum_y f_{X,Y}(x, y) = \sum_x x \, f_X(x) = \mathbb{E}X.
\end{aligned}$$

Tower Property

More generally, for <u>any</u> two random variable X and Y such that X has finite expectation, we have

$$\mathbb{E}(\mathbb{E}(X \mid Y)) = \mathbb{E}(X).$$

This is sometimes also referred to as *Law of Total Probability for Expectations.*

Example

A pair of random variables (X, Y) has a joint distribution in which Y has marginal pdf given by

$$f_Y(y) = \begin{cases} 3y^2, & y \in [0, 1] \\ 0, & \text{else}, \end{cases}$$

and the conditional pdf of X given $\{Y = y\}$ is uniform on the interval $[0, y]$.

Q: What is the expectation of X?

Example (continued)

$$
\begin{aligned}
\mathbb{E}(X) &= \mathbb{E}(\mathbb{E}(X \mid Y)) \\
&= \int_{-\infty}^{\infty} \mathbb{E}(X \mid Y = y) f_Y(y) \, \mathrm{d}y \\
&= \int_0^1 \frac{y}{2} \cdot 3y^2 \, \mathrm{d}y = \dots
\end{aligned}
$$