STA286 Lecture 09

Neil Montgomery

Last edited: 2017-01-30 12:35

conditional distributions

Consider again the gas pipe joint (and marginal) distributions:

	X			
Y	1	1.5	1.75	Marginal
0.5	0.075	0.100	0.150	0.325
1	0.110	0.080	0.090	0.280
2	0.160	0.140	0.095	0.395
Marginal	0.345	0.320	0.335	1.000

conditional distributions

Consider again the gas pipe joint (and marginal) distributions:

	X			
Y	1	1.5	1.75	Marginal
0.5	0.075	0.100	0.150	0.325
1	0.110	0.080	0.090	0.280
2	0.160	0.140	0.095	0.395
Marginal	0.345	0.320	0.335	1.000

What are the probabilities of X taking on any of its three possible values *given* $Y = 2$, *say.*

conditional distributions

Consider again the gas pipe joint (and marginal) distributions:

	X			
Y	1	1.5	1.75	Marginal
0.5	0.075	0.100	0.150	0.325
1	0.110	0.080	0.090	0.280
2	0.160	0.140	0.095	0.395
Marginal	0.345	0.320	0.335	1.000

What are the probabilities of X taking on any of its three possible values *given* $Y = 2$, *say.*

$$P(X = 1|Y = 2) = fracP(X = 1, Y = 2)P(Y = 2) = \frac{0.16}{0.395}$$

conditional distributions

Consider again the gas pipe joint (and marginal) distributions:

	X			
Y	1	1.5	1.75	Marginal
0.5	0.075	0.100	0.150	0.325
1	0.110	0.080	0.090	0.280
2	0.160	0.140	0.095	0.395
Marginal	0.345	0.320	0.335	1.000

What are the probabilities of X taking on any of its three possible values *given* $Y = 2$, *say.*

$$P(X = 1|Y = 2) = fracP(X = 1, Y = 2)P(Y = 2) = \frac{0.16}{0.395}$$

$$P(X = 1.5|Y = 2) = \frac{0.14}{0.395} \qquad P(X = 1.75|Y = 2) = \frac{0.095}{0.395}$$

conditional pmf and conditional density

Given the joint pmf $p(x, y)$ for X and Y, the conditional pmf for X given $Y = y$ is:

$$p(x|y) = \frac{p(x, y)}{p_Y(y)},$$

provided $P(Y = y) > 0$.

conditional pmf and conditional density

Given the joint pmf $p(x, y)$ for X and Y, the conditional pmf for X given $Y = y$ is:

$$p(x|y) = \frac{p(x, y)}{p_Y(y)},$$

provided $P(Y = y) > 0$.

Given the joint density $f(x, y)$ for X and Y, the conditional density for X given $Y = y$ is:

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

provided $f_Y(y) > 0$.

conditional density example

Reconsider the two electronic components example:

$$f(x,y) \begin{cases} 2e^{-x-2y} & : x > 0, y > 0 \\ 0 & : \text{otherwise} \end{cases}$$

The marginal density for Y is $2e^{-2y}$ on $y > 0$.

conditional density example

Reconsider the two electronic components example:

$$f(x,y)\begin{cases}2e^{-x-2y} & : x>0, y>0\\ 0 & : \text{otherwise}\end{cases}$$

The marginal density for Y is $2e^{-2y}$ on $y>0$.

The conditional density for X given $Y=2$ is:

conditional density example

Reconsider the two electronic components example:

$$f(x,y) \begin{cases} 2e^{-x-2y} & : x > 0, y > 0 \\ 0 & : \text{otherwise} \end{cases}$$

The marginal density for Y is $2e^{-2y}$ on $y > 0$.

The conditional density for X given $Y = 2$ is:

$$e^{-x} \text{ on } x > 0$$

conditional density example

Reconsider the two electronic components example:

$$f(x,y) \begin{cases} 2e^{-x-2y} & : x > 0, y > 0 \\ 0 & : \text{otherwise} \end{cases}$$

The marginal density for Y is $2e^{-2y}$ on $y > 0$.

The conditional density for X given $Y = 2$ is:

$$e^{-x} \text{ on } x > 0$$

Heck, the conditional density for X given Y equals anything greater than 0 is still always e^{-x} on $x > 0$.

conditional density example

Reconsider the two electronic components example:

$$f(x,y)\begin{cases}2e^{-x-2y} & : x>0, y>0\\ 0 & : \text{otherwise}\end{cases}$$

The marginal density for Y is $2e^{-2y}$ on $y>0$.

The conditional density for X given $Y=2$ is:

$$e^{-x} \text{ on } x>0$$

Heck, the conditional density for X given Y equals anything greater than 0 is still always e^{-x} on $x>0$.

That's because X and Y have a special relationship...to be revisited.

conditional density example “uniform on a triangle”

Consider again X and Y with the following density.

$$f(x,y) = \begin{cases} 2 & : 0 < x < 1,\ 0 < y < 1,\ x + y < 1 \\ 0 & : \text{otherwise} \end{cases}$$

The marginal density for X was found to be $f_X(x) = 2x$ for $0 < x < 1$.

conditional density example "uniform on a triangle"

Consider again X and Y with the following density.

$$f(x,y) = \begin{cases} 2 & : 0 < x < 1,\ 0 < y < 1,\ x + y < 1 \\ 0 & : \text{otherwise} \end{cases}$$

The marginal density for X was found to be $f_X(x) = 2x$ for $0 < x < 1$.

The conditional density of Y given $X = 0.5$ will be 0 when y is outside $(0, 0.5)$. Otherwise it will be:

$$f(y|x = 0.5) = \frac{2}{2 \cdot 0.5} = 2$$

conditional density example "uniform on a triangle"

Consider again X and Y with the following density.

$$f(x,y) = \begin{cases} 2 & : 0 < x < 1,\ 0 < y < 1,\ x + y < 1 \\ 0 & : \text{otherwise} \end{cases}$$

The marginal density for X was found to be $f_X(x) = 2x$ for $0 < x < 1$.

The conditional density of Y given $X = 0.5$ will be 0 when y is outside $(0, 0.5)$. Otherwise it will be:

$$f(y|x = 0.5) = \frac{2}{2 \cdot 0.5} = 2$$

Unlike marginal densities, conditional densities can sometimes be visualized. They are just "slices" of the joint density.

independence

In the "two electronic component" example, the conditional density for X given $Y = y$ (no matter what $y > 0$) never changes:

$$f(x|y) = f(x)$$

Knowledge of the outcome of Y doesn't tell you anything about the distribution of X.

independence

In the “two electronic component” example, the conditional density for X given $Y = y$ (no matter what $y > 0$) never changes:

$$f(x|y) = f(x)$$

Knowledge of the outcome of Y doesn’t tell you anything about the distribution of X.

In the “uniform on a triangle” example, the conditional density for Y given $X = x$ is always going to be (for $0 < x < 1$):

$$f(y|x) = \frac{2}{2x} = \frac{1}{x}$$

for $0 < y < x$ and 0 otherwise. So knowledge of the outcome of X tells something about the distribution of Y.

independence

In the gas pipes pmf:

	X			
Y	1	1.5	1.75	Marginal
0.5	0.075	0.100	0.150	0.325
1	0.110	0.080	0.090	0.280
2	0.160	0.140	0.095	0.395
Marginal	0.345	0.320	0.335	1.000

the conditional distributions for X given $Y = y$ are all different (as well as those for Y given $X = x$.)

independence

In the gas pipes pmf:

	X			
Y	1	1.5	1.75	Marginal
0.5	0.075	0.100	0.150	0.325
1	0.110	0.080	0.090	0.280
2	0.160	0.140	0.095	0.395
Marginal	0.345	0.320	0.335	1.000

the conditional distributions for X given $Y = y$ are all different (as well as those for Y given $X = x$.)

Definition: X and Y are *independent* (or $X \perp Y$) if:

$$p(x, y) = p_X(x)p_Y(y) \qquad \text{(discrete)} \qquad\qquad f(x, y) = f_X(x)f_Y(y) \qquad \text{(continuous)}$$

gas pipe diameters and pressures “made” independent

Suppose the marginal distributions are the same. What would the joint pmf have to be?

	X			
Y	1	1.5	1.75	Marginal
0.5				0.325
1				0.280
2				0.395
Marginal	0.345	0.320	0.335	1.000

gas pipe diameters and pressures “made” independent

Suppose the marginal distributions are the same. What would the joint pmf have to be?

	X			
Y	1	1.5	1.75	Marginal
0.5	$0.345 \cdot 0.325$			0.325
1				0.280
2				0.395
Marginal	0.345	0.320	0.335	1.000

gas pipe diameters and pressures “made” independent

Suppose the marginal distributions are the same. What would the joint pmf have to be?

Y	X: 1	1.5	1.75	Marginal
0.5	$0.345 \cdot 0.325$			0.325
1				0.280
2		$0.320 \cdot 0.395$		0.395
Marginal	0.345	0.320	0.335	1.000

gas pipe diameters and pressures “made” independent

Suppose the marginal distributions are the same. What would the joint pmf have to be?

	X			
Y	1	1.5	1.75	Marginal
0.5	$0.345 \cdot 0.325$			0.325
1			$0.335 \cdot 0.280$	0.280
2		$0.320 \cdot 0.395$		0.395
Marginal	0.345	0.320	0.335	1.000

relationship with independence of events

For an event A from a sample space S, define the *indicator random variable*:

$$I_A(\omega) = \begin{cases} 1 & : \omega \in S \\ 0 & : \omega \notin S \end{cases}$$

This is the general "Bernoulli trial" and one of the most important random variables there is.

relationship with independence of events

relationship with independence of events

Suppose A and B are events in S. Consider the joint pmf for A and B:

	I_A		
I_B	0	1	Marginal
0			
1			
Marginal			

relationship with independence of events

Suppose A and B are events in S. Consider the joint pmf for A and B:

	I_A		
I_B	0	1	Marginal
0			
1			$P(B)$
Marginal		$P(A)$	

relationship with independence of events

Suppose A and B are events in S. Consider the joint pmf for A and B:

	I_A		
I_B	0	1	Marginal
0			$P(B')$
1			$P(B)$
Marginal	$P(A')$	$P(A)$	

relationship with independence of events

Suppose A and B are events in S. Consider the joint pmf for A and B:

	I_A		
I_B	0	1	Marginal
0			$P(B')$
1		$P(A \cap B)$	$P(B)$
Marginal	$P(A')$	$P(A)$	

relationship with independence of events

Suppose A and B are events in S. Consider the joint pmf for A and B:

	I_A		
I_B	0	1	Marginal
0		$P(A \cap B')$	$P(B')$
1		$P(A \cap B)$	$P(B)$
Marginal	$P(A')$	$P(A)$	

relationship with independence of events

Suppose A and B are events in S. Consider the joint pmf for A and B:

	I_A		
I_B	0	1	Marginal
0	$P(A' \cap B')$	$P(A \cap B')$	$P(B')$
1		$P(A \cap B)$	$P(B)$
Marginal	$P(A')$	$P(A)$	

relationship with independence of events

Suppose A and B are events in S. Consider the joint pmf for A and B:

	I_A		
I_B	0	1	Marginal
0	$P(A' \cap B')$	$P(A \cap B')$	$P(B')$
1	$P(A' \cap B)$	$P(A \cap B)$	$P(B)$
Marginal	$P(A')$	$P(A)$	

relationship with independence of events

Suppose A and B are events in S. Consider the joint pmf for A and B:

	I_A		
I_B	0	1	Marginal
0	$P(A' \cap B')$	$P(A \cap B')$	$P(B')$
1	$P(A' \cap B)$	$P(A \cap B)$	$P(B)$
Marginal	$P(A')$	$P(A)$	

Now suppose $A \perp B$. Recall that this is actually a "strong" statement equivalent to $A \perp B'$ and $A' \perp B'$ and $A' \perp B$.

relationship with independence of events

Suppose A and B are events in S. Consider the joint pmf for A and B:

	I_A		
I_B	0	1	Marginal
0	$P(A' \cap B')$	$P(A \cap B')$	$P(B')$
1	$P(A' \cap B)$	$P(A \cap B)$	$P(B)$
Marginal	$P(A')$	$P(A)$	

Now suppose $A \perp B$. Recall that this is actually a "strong" statement equivalent to $A \perp B'$ and $A' \perp B'$ and $A' \perp B$.

So it is also equivalent to the independence of the random variables I_A and I_B.

relationship with independence of events

Suppose A and B are events in S. Consider the joint pmf for A and B:

	I_A		
I_B	0	1	Marginal
0	$P(A' \cap B')$	$P(A \cap B')$	$P(B')$
1	$P(A' \cap B)$	$P(A \cap B)$	$P(B)$
Marginal	$P(A')$	$P(A)$	

Now suppose $A \perp B$. Recall that this is actually a "strong" statement equivalent to $A \perp B'$ and $A' \perp B'$ and $A' \perp B$.

So it is also equivalent to the independence of the random variables I_A and I_B.

The formal definition of independence is:

$$P(X \in C, Y \in D) = P(X \in C)P(Y \in D)$$

for any $C, D \subset \mathbb{R}$.

notes on verifying independence

First of all, in practice independence continues to be someone one generally *assumes*.

notes on verifying independence

First of all, in practice independence continues to be someone one generally *assumes*.

But one still needs to know how to verify independence, which is easy for continuous X and Y:

- the joint density factors:

$$f(x, y) = g(x)h(y)$$

notes on verifying independence

First of all, in practice independence continues to be someone one generally *assumes*.

But one still needs to know how to verify independence, which is easy for continuous X and Y:

- the joint density factors:
$$f(x, y) = g(x)h(y)$$
- and the non-zero region of f is a "rectangle"

notes on verifying independence

First of all, in practice independence continues to be someone one generally *assumes*.

But one still needs to know how to verify independence, which is easy for continuous X and Y:

- the joint density factors:
$$f(x,y) = g(x)h(y)$$
- and the non-zero region of f is a "rectangle"

notes on verifying independence

First of all, in practice independence continues to be someone one generally *assumes*.

But one still needs to know how to verify independence, which is easy for continuous X and Y:

- the joint density factors:
$$f(x, y) = g(x)h(y)$$
- and the non-zero region of f is a "rectangle"

Examples:

- $f(x, y) = x + y$ on $0 < x, y < 1$ and 0 otherwise.

notes on verifying independence

First of all, in practice independence continues to be someone one generally *assumes*.

But one still needs to know how to verify independence, which is easy for continuous X and Y:

- the joint density factors:
$$f(x, y) = g(x)h(y)$$
- and the non-zero region of f is a "rectangle"

Examples:

- $f(x, y) = x + y$ on $0 < x, y < 1$ and 0 otherwise.
- $f(x, y) = 24xy$ on $x > 0, y > 0, 0 < x + y < 1$ and 0 otherwise.

the spanish inquisition

expected value

Recall the sample average:

$$\overline{x} = \sum_{i=1}^{n} x_i \cdot \frac{1}{n}$$

which can be considered as a weighted sum with weights $w_i = 1/n$.

expected value

Recall the sample average:

$$\overline{x} = \sum_{i=1}^{n} x_i \cdot \frac{1}{n}$$

which can be considered as a weighted sum with weights $w_i = 1/n$.

A (discrete) random variable can have an “average”, which is a weighted sum of the outcomes with their probabilities as weights.

expected value

Recall the sample average:

$$\overline{x} = \sum_{i=1}^{n} x_i \cdot \frac{1}{n}$$

which can be considered as a weighted sum with weights $w_i = 1/n$.

A (discrete) random variable can have an "average", which is a weighted sum of the outcomes with their probabilities as weights.

BIG MONEY. We play a gambling game called BIG MONEY. Roll a die. This is your outcome after each play of BIG MONEY:

Roll	1, 2	3	4, 5	6
Outcome \$	-2	0	1	2

expected value

Recall the sample average:

$$\overline{x} = \sum_{i=1}^{n} x_i \cdot \frac{1}{n}$$

which can be considered as a weighted sum with weights $w_i = 1/n$.

A (discrete) random variable can have an "average", which is a weighted sum of the outcomes with their probabilities as weights.

BIG MONEY. We play a gambling game called BIG MONEY. Roll a die. This is your outcome after each play of BIG MONEY:

Roll	1, 2	3	4, 5	6
Outcome \$	-2	0	1	2
Probability	2/6	1/6	2/6	1/6

BIG MONEY is a "fair game"

Your *expected* winnings are (theoretically):

$$-2\frac{2}{6} + 0\frac{1}{6} + 1\frac{2}{6} + 2\frac{1}{6} = 0$$

BIG MONEY is a "fair game"

Your *expected* winnings are (theoretically):

$$-2\frac{2}{6} + 0\frac{1}{6} + 1\frac{2}{6} + 2\frac{1}{6} = 0$$

The *expected value* of a discrete random variable X is:

$$E(X) = \sum_x xp(x)$$

if the sum exists (actually it has to converge absolutely.)

expected number of coin tosses to first H

Denote by X the number of coin tosses until the first H. The pmf is:

$$p(x) = \left(1 - \frac{1}{2}\right)^{x-1} \frac{1}{2}$$

for $x \in \{1, 2, 3, \ldots\}$. For a moment replace $\frac{1}{2}$ with p.

expected number of coin tosses to first H

Denote by X the number of coin tosses until the first H. The pmf is:

$$p(x) = \left(1 - \frac{1}{2}\right)^{x-1} \frac{1}{2}$$

for $x \in \{1, 2, 3, \ldots\}$. For a moment replace $\frac{1}{2}$ with p.

$$E(X) = \sum_{x=1}^{\infty} x(1-p)^{x-1} p$$

expected number of coin tosses to first H

Denote by X the number of coin tosses until the first H. The pmf is:

$$p(x) = \left(1 - \frac{1}{2}\right)^{x-1} \frac{1}{2}$$

for $x \in \{1, 2, 3, \ldots\}$. For a moment replace $\frac{1}{2}$ with p.

$$\begin{aligned} E(X) &= \sum_{x=1}^{\infty} x(1-p)^{x-1} p \\ &= p \sum_{x=0}^{\infty} x(1-p)^{x-1} \end{aligned}$$

expected number of coin tosses to first H

Denote by X the number of coin tosses until the first H. The pmf is:

$$p(x) = \left(1 - \frac{1}{2}\right)^{x-1} \frac{1}{2}$$

for $x \in \{1, 2, 3, \ldots\}$. For a moment replace $\frac{1}{2}$ with p.

$$\begin{aligned} E(X) &= \sum_{x=1}^{\infty} x(1-p)^{x-1} p \\ &= p \sum_{x=0}^{\infty} x(1-p)^{x-1} \\ &= p \sum_{x=0}^{\infty} \frac{d}{dp}(1-p)^{x} \end{aligned}$$

expected number of coin tosses to first H

Denote by X the number of coin tosses until the first H. The pmf is:

$$p(x) = \left(1 - \frac{1}{2}\right)^{x-1} \frac{1}{2}$$

for $x \in \{1, 2, 3, \ldots\}$. For a moment replace $\frac{1}{2}$ with p.

$$\begin{aligned}
E(X) &= \sum_{x=1}^{\infty} x(1-p)^{x-1} p \\
&= p \sum_{x=0}^{\infty} x(1-p)^{x-1} \\
&= p \sum_{x=0}^{\infty} \frac{d}{dp}(1-p)^{x} \\
&= p \frac{d}{dp} \left(\sum_{x=0}^{\infty} (1-p)^{x} \right) = p \frac{d}{dp} \left(\frac{1}{1-(1-p)} \right) = \frac{1}{p}
\end{aligned}$$

expected value non-example

Let X have the following pmf (!):

$$p(x) = \frac{6}{\pi^2 x^2}, x \in \{1, 2, 3, \ldots\}$$

X does not have an expected value, because $\sum_x xp(x)$ does not converge.

expected value - continuous version

If X is continuous with density f, its expected value is:

$$E(X) = \int_{-\infty}^{\infty} xf(x)\,dx$$

provided the integral converges (absolutely).

expected value - continuous version

If X is continuous with density f, its expected value is:

$$E(X) = \int_{-\infty}^{\infty} x f(x) \, dx$$

provided the integral converges (absolutely).

Bus stop example. What is the expected waiting time?