STA286 Lecture 08

Neil Montgomery

Last edited: 2017-01-26 16:27

a note on "identically distributed"

The distribution is all we care about.

So if X_1 and X_2 have the same distributions, they are effectively the same (even if they are not the same functions.)

a note on "identically distributed"

The distribution is all we care about.

So if X_1 and X_2 have the same distributions, they are effectively the same (even if they are not the same functions.)

For example, roll a fair die so that $S = \{1, 2, 3, 4, 5, 6\}$.

a note on "identically distributed"

The distribution is all we care about.

So if X_1 and X_2 have the same distributions, they are effectively the same (even if they are not the same functions.)

For example, roll a fair die so that $S = \{1, 2, 3, 4, 5, 6\}$.

Define:

$$X_1 = \begin{cases} 1 & : \text{3 or 4 appears} \\ 0 & : \text{otherwise} \end{cases} \qquad \text{and} \qquad X_2 = \begin{cases} 1 & : \text{5 or 6 appears} \\ 0 & : \text{otherwise} \end{cases}$$

a note on "identically distributed"

The distribution is all we care about.

So if X_1 and X_2 have the same distributions, they are effectively the same (even if they are not the same functions.)

For example, roll a fair die so that $S = \{1, 2, 3, 4, 5, 6\}$.

Define:

$$X_1 = \begin{cases} 1 & : \text{3 or 4 appears} \\ 0 & : \text{otherwise} \end{cases} \qquad \text{and} \qquad X_2 = \begin{cases} 1 & : \text{5 or 6 appears} \\ 0 & : \text{otherwise} \end{cases}$$

X_1 and X_2 are not the same functions. But the have the same p.m.f.:

$$p(x) = \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{1-x}, \; x \in \{0, 1\}.$$

We say X_1 and X_2 are *identically distributed*.

joint distributions

more than one random variable at a time

A dataset will usually have more than one variable. They might be modeled at the same time.

more than one random variable at a time

A dataset will usually have more than one variable. They might be modeled at the same time.

Certainly a dataset will have more than one observation, so considering multiple random variables is essential.

more than one random variable at a time

A dataset will usually have more than one variable. They might be modeled at the same time.

Certainly a dataset will have more than one observation, so considering multiple random variables is essential.

The approach is to consider a random vector such as (X, Y), (X_1, X_2), $(X_1, X_2, \ldots X_n)$ and so on. At first we'll consider two at a time: (X, Y).

more than one random variable at a time

A dataset will usually have more than one variable. They might be modeled at the same time.

Certainly a dataset will have more than one observation, so considering multiple random variables is essential.

The approach is to consider a random vector such as (X, Y), (X_1, X_2), $(X_1, X_2, \ldots X_n)$ and so on. At first we'll consider two at a time: (X, Y).

As usual we are interested in the *distribution* of, say, (X, Y), which now means the collection of things like $P((X, Y) \in A)$ for $A \in \mathbb{R}^2$.

more than one random variable at a time

A dataset will usually have more than one variable. They might be modeled at the same time.

Certainly a dataset will have more than one observation, so considering multiple random variables is essential.

The approach is to consider a random vector such as (X, Y), (X_1, X_2), $(X_1, X_2, \ldots X_n)$ and so on. At first we'll consider two at a time: (X, Y).

As usual we are interested in the *distribution* of, say, (X, Y), which now means the collection of things like $P((X, Y) \in A)$ for $A \in \mathbb{R}^2$.

Main interest is in statements like:

$$P(X = x, Y = y) \qquad \text{or} \qquad P(a < X < b, c < Y < d)$$

where the "comma" notation is a compact way of writing, say:

$$P(\{X = x\} \cap \{Y = y\})$$

cdf, pmf, pdf

cdf, pmf, pdf

The natural ways to characterize joint distributions are still probability mass functions and probability density functions.

cdf, pmf, pdf

The natural ways to characterize joint distributions are still probability mass functions and probability density functions.

They also have cdfs:

$$F(x, y) = P(X \leqslant x, Y \leqslant y)$$

but we won't focus much on these.

a joint pmf

A gas distribution company has pipes with diameters 1, 1.5, and 1.75 inches. The pipes are used at pressures of 2, 1, and 0.5 pounds per square inch.

Pick a pipe at random and denote its diameter by X and its pressure by Y.

a joint pmf

A gas distribution company has pipes with diameters 1, 1.5, and 1.75 inches. The pipes are used at pressures of 2, 1, and 0.5 pounds per square inch.

Pick a pipe at random and denote its diameter by X and its pressure by Y.

X and Y might have the following *joint probability mass function*:

	X		
Y	1	1.5	1.75
0.5	0.075	0.100	0.150
1	0.110	0.080	0.090
2	0.160	0.140	0.095

a joint pmf

A gas distribution company has pipes with diameters 1, 1.5, and 1.75 inches. The pipes are used at pressures of 2, 1, and 0.5 pounds per square inch.

Pick a pipe at random and denote its diameter by X and its pressure by Y.

X and Y might have the following *joint probability mass function*:

	X		
Y	1	1.5	1.75
0.5	0.075	0.100	0.150
1	0.110	0.080	0.090
2	0.160	0.140	0.095

e.g. the probability that a randomly selected pipe has diameter $X = 1.5$ and pressure $Y = 0.5$ is:

$$P(X = 1.5, Y = 0.5) = 0.1$$

joint pmf properties; joint density

A joint pmf is still just a pmf. It must be non-negative, and its positive values must add up to 1.

joint pmf properties; joint density

A joint pmf is still just a pmf. It must be non-negative, and its positive values must add up to 1.

(X, Y) are (jointly) continuous if they have a joint density f such that:

$$P(a < X < b, c < Y < d) = \int_c^d \int_a^b f(x, y)\, dxdy$$

joint pmf properties; joint density

A joint pmf is still just a pmf. It must be non-negative, and its positive values must add up to 1.

(X, Y) are (jointly) continuous if they have a joint density f such that:

$$P(a < X < b, c < Y < d) = \int_c^d \int_a^b f(x, y)\, dxdy$$

A joint density also must be non-negative and integrate to 1 over all $\mathbb{R}^2$.

joint density example - I

Two electronic components fail at times X and Y according to the following joint density (measured in years):

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

joint density example - I

Two electronic components fail at times X and Y according to the following joint density (measured in years):

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

Is this actually a valid joint density? It is non-negative, and:

$$\iint\limits_{\mathbb{R}^2} f(x,y)\,dxdy = \int\limits_0^\infty \int\limits_0^\infty 2e^{-x-2y}\,dxdy = \int\limits_0^\infty e^{-x}\,dx \int\limits_0^\infty 2e^{-2y}\,dy = 1$$

joint density example - II

What is the probability that the first component fails before the second? In other words, what is $P(X < Y)$?

joint density example - II

What is the probability that the first component fails before the second? In other words, what is $P(X < Y)$?

Answer: integrate the joint density over the region where $x < y$.

$$\int_0^\infty \int_0^y 2e^{-x-2y}\,dx\,dy = \int_0^\infty \left[-2e^{-x-2y}\right]_{x=0}^{y} dy$$

joint density example - II

What is the probability that the first component fails before the second? In other words, what is $P(X < Y)$?

Answer: integrate the joint density over the region where $x < y$.

$$\int_0^\infty \int_0^y 2e^{-x-2y}\,dx\,dy = \int_0^\infty \left[-2e^{-x-2y}\right]_{x=0}^{y} dy$$

$$= \int_0^\infty 2e^{-2y} - 2e^{-3y}\,dy$$

joint density example - II

What is the probability that the first component fails before the second? In other words, what is $P(X < Y)$?

Answer: integrate the joint density over the region where $x < y$.

$$
\begin{aligned}
\int_0^\infty \int_0^y 2e^{-x-2y}\,dx\,dy &= \int_0^\infty \left[-2e^{-x-2y}\right]_{x=0}^{y} dy \\
&= \int_0^\infty 2e^{-2y} - 2e^{-3y}\,dy \\
&= \left[-e^{-2y} + \frac{2}{3}e^{-3y}\right]_{y=0}^{\infty}
\end{aligned}
$$

joint density example - II

What is the probability that the first component fails before the second? In other words, what is $P(X < Y)$?

Answer: integrate the joint density over the region where $x < y$.

$$\int_0^\infty \int_0^y 2e^{-x-2y}\,dx\,dy = \int_0^\infty \left[-2e^{-x-2y}\right]_{x=0}^{y} dy$$
$$= \int_0^\infty 2e^{-2y} - 2e^{-3y}\,dy$$
$$= \left[-e^{-2y} + \frac{2}{3}e^{-3y}\right]_{y=0}^{\infty}$$
$$= \frac{1}{3}$$

marginal distributions

A joint distribution contains *all* information about both X and Y together.

"*All*" includes information about X alone, and about Y alone. And it's easy to get this information from the joint distribution.

marginal distributions

A joint distribution contains *all* information about both X and Y together.

"*All*" includes information about X alone, and about Y alone. And it's easy to get this information from the joint distribution.

For example, look again at the joint distribution for diameter X and pressure Y of the randomly selected pipe:

	X		
Y	1	1.5	1.75
0.5	0.075	0.100	0.150
1	0.110	0.080	0.090
2	0.160	0.140	0.095

marginal distributions

A joint distribution contains *all* information about both X and Y together.

"*All*" includes information about X alone, and about Y alone. And it's easy to get this information from the joint distribution.

For example, look again at the joint distribution for diameter X and pressure Y of the randomly selected pipe:

	X		
Y	1	1.5	1.75
0.5	0.075	0.100	0.150
1	0.110	0.080	0.090
2	0.160	0.140	0.095

A statement about X alone could be something like $P(X = 1)$, which is just $0.075 + 0.11 + 0.16 = 0.345$

marginal pmf

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

$$p_X(x) = \sum_y p(x, y)$$

marginal pmf

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

$$p_X(x) = \sum_y p(x, y)$$

Here are the "marginals" for both X and Y in the gas example:

	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

marginal density

Reconsider the two electronic components example:

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

A statement about, say, Y alone might be $P(Y>1)$, which would be the integral on the entire half plane $y>1$:

$$\iint\limits_{y>1} f(x,y)\,dxdy$$

marginal density

Reconsider the two electronic components example:

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

A statement about, say, Y alone might be $P(Y>1)$, which would be the integral on the entire half plane $y>1$:

$$\iint\limits_{y>1} f(x,y)\,dxdy = \int\limits_1^\infty\int\limits_0^\infty 2e^{-x-2y}\,dxdy$$

But in the end, the calculation only ever involved $2e^{-2y}$ once x was “integrated out.”

marginal density

Reconsider the two electronic components example:

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

A statement about, say, Y alone might be $P(Y > 1)$, which would be the integral on the entire half plane $y > 1$:

$$\iint\limits_{y>1} f(x,y)\,dxdy = \int\limits_{1}^{\infty}\int\limits_{0}^{\infty} 2e^{-x-2y}\,dxdy = \int\limits_{1}^{\infty} 2e^{-2y}\,dy$$

But in the end, the calculation only ever involved $2e^{-2y}$ once x was "integrated out."

marginal density

Reconsider the two electronic components example:

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

A statement about, say, Y alone might be $P(Y > 1)$, which would be the integral on the entire half plane $y > 1$:

$$\iint\limits_{y>1} f(x,y)\, dxdy = \int_1^\infty \int_0^\infty 2e^{-x-2y}\, dxdy = \int_1^\infty 2e^{-2y}\, dy = e^{-2}$$

But in the end, the calculation only ever involved $2e^{-2y}$ once x was "integrated out."

marginal density

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

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$$

marginal density

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

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$$

Marginal densities can't really be visualized. At best they are a "projection" onto one or the other axis.

marginal density

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

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$$

Marginal densities can't really be visualized. At best they are a "projection" onto one or the other axis.

e.g. Consider the following function:

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

marginal density

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

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$$

Marginal densities can't really be visualized. At best they are a "projection" onto one or the other axis.

e.g. Consider the following function:

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

This is a valid density. Let X and Y have this joint density.

marginal density

The marginal density for X will be 0 outside $x \in (0, 1)$. Otherwise:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$$

marginal density

The marginal density for X will be 0 outside $x \in (0, 1)$. Otherwise:

$$
\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f(x, y)\, dy \\
&= \int_{0}^{1-x} 2\, dy
\end{aligned}
$$

marginal density

The marginal density for X will be 0 outside $x \in (0, 1)$. Otherwise:

$$
\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f(x, y)\, dy \\
&= \int_{0}^{1-x} 2\, dy \\
&= 2(1 - x)
\end{aligned}
$$

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

Consider again the following function:

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

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

conditional density example

Consider again the following function:

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

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:

$$f_Y(y) = \frac{2}{2 \cdot 0.5} = 2$$

conditional density example

Consider again the following function:

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

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:

$$f_Y(y) = \frac{2}{2 \cdot 0.5} = 2$$

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