14.31/14.310 Lecture 6

Probability---example

Two weeks ago, we ended with an example involving computing probabilities from a joint PDF. It generated a lot of questions. So let's do another example*.

Suppose we have $f_{XY}(x,y) = \begin{cases} cx^2y & \text{for } x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$

*Taken from DeGroot and Schervish.

Probability---example

Two weeks ago, we ended with an example involving computing probabilities from a joint PDF. It generated a lot of questions. So let's do another example*.

Suppose we have $f_{XY}(x,y) = \begin{cases} cx^2y & \text{for } x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$

First, let's draw a picture of the support of this distribution, or the region of the xy-plane over which there is positive probability.

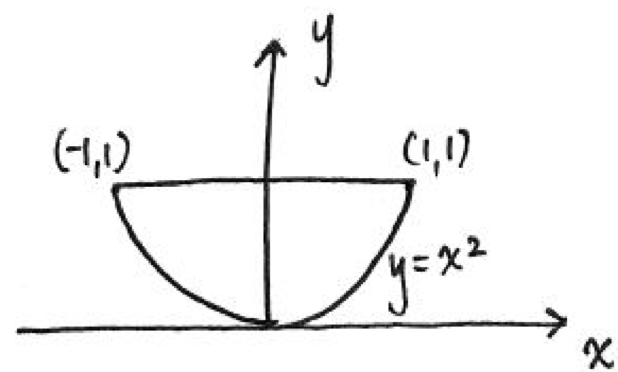
*Taken from DeGroot and Schervish.

Probability---example

Two weeks ago, we ended with an example involving computing probabilities from a joint PDF. It generated a lot of questions. So let's do another example*.

Suppose we have $f_{XY}(x,y) = \begin{cases} cx^2y & \text{for } x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$

First, let's draw a picture of the support of this distribution, or the region of the xy-plane over which there is positive probability.

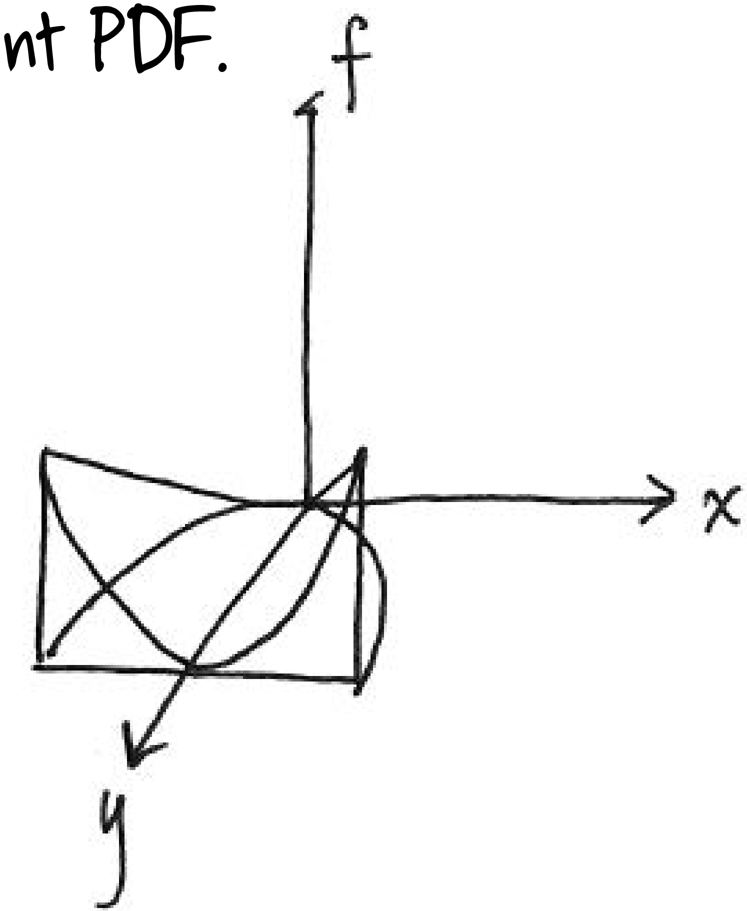


*Taken from DeGroot and Schervish.

Probability---example

Suppose we have $f_{XY}(x,y) = \begin{cases} cx^2y & \text{for } x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$

Now here's a 3D drawing of the joint PDF.



Probability---example

Suppose we have $f_{XY}(x,y) = \begin{cases} cx^2y & \text{for } x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$

Let's figure out what c is. How?

Probability---example

Suppose we have $f_{XY}(x,y) = \begin{cases} cx^2y & \text{for } x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$

Let's figure out what c is. How? We know this joint PDF has to integrate to 1. So let's integrate this thing and solve for what c must be.

Probability---example

$$f_{XY}(x,y) = \begin{cases} cx^2y & \text{for } x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$$

Note that we only need to integrate over the support of the distribution because the PDF is 0 elsewhere.

$$\int_{-1}^{1}\int_{x^2}^{1} cx^2y\,dy\,dx = \frac{4}{21}c$$

So $c = 21/4$.

Probability---example

$$f_{XY}(x,y) = \begin{cases} cx^2y & \text{for } x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$$

Note that we only need to integrate over the support of the distribution because the PDF is 0 elsewhere.

$$\int_{-1}^{1} \int_{x^2}^{1} cx^2y\,dy\,dx = \frac{4}{21}c$$

How did I get these limits of integration?

So c = 21/4.

Probability---example

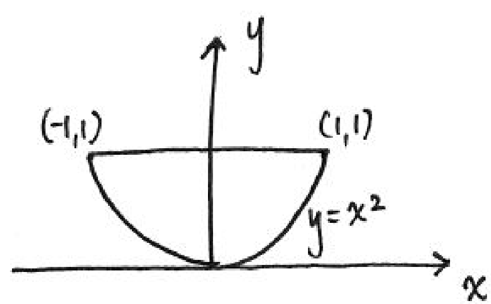
$$f_{XY}(x,y) = \begin{cases} cx^2y & \text{for } x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$$

Note that we only need to integrate over the support of the distribution because the PDF is 0 elsewhere.

$$\int_{-1}^{1} \int_{x^2}^{1} cx^2 y \, dy \, dx = \frac{4}{21} c$$

How did I get these limits of integration?

So c = 21/4.



Probability---example

So $f_{XY}(x,y) = \begin{cases} (21/4)x^2y & x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$

What is $P(X>Y)$?

Probability---example

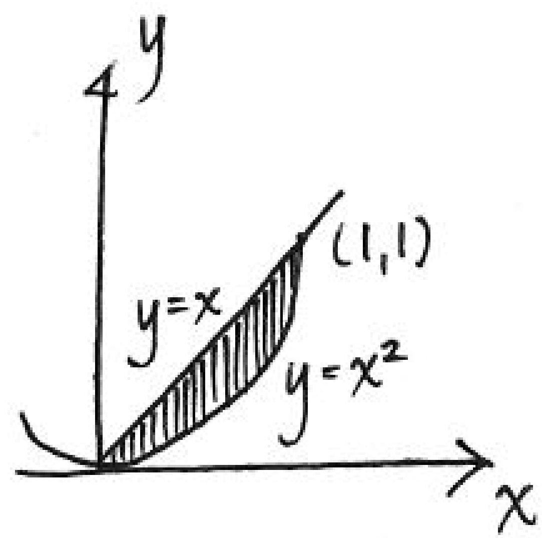
So $f_{XY}(x,y) = \begin{cases} (21/4)x^2y & x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$

What is $P(X>Y)$? Have to figure out the region of the xy-plane over which we integrate.

Probability---example

So $f_{XY}(x,y) = \begin{cases} (21/4)x^2y & x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$

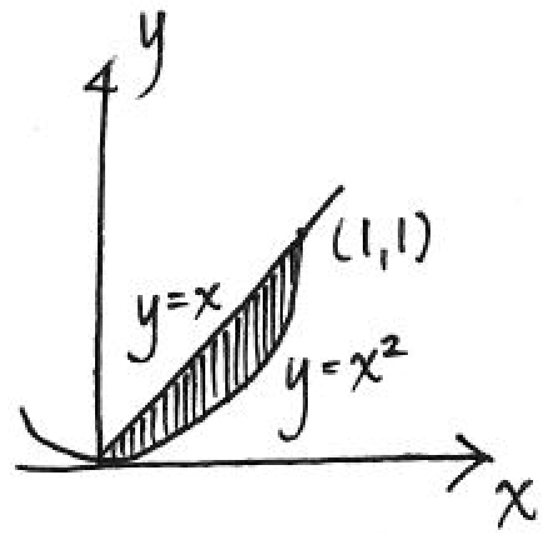
What is $P(X>Y)$?



Probability---example

So $f_{XY}(x,y) = \begin{cases} (21/4)x^2y & x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$

What is P(X>Y)?

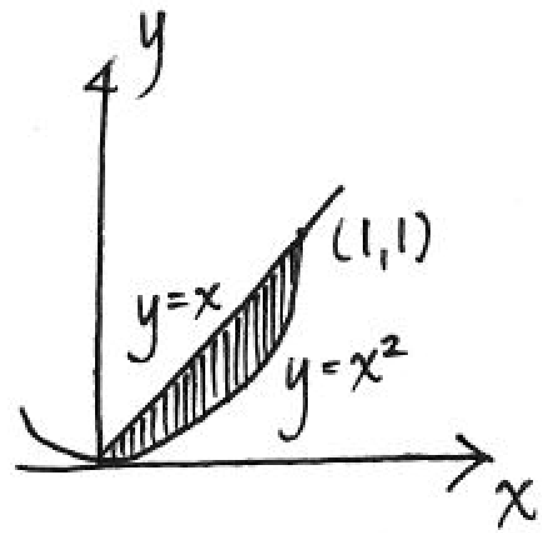


This is the intersection between the support of the joint PDF and the half-plane where x > y.

Probability---example

So $f_{XY}(x,y) = \begin{cases} (21/4)x^2y & x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$

What is P(X>Y)?



$$\int_0^1 \int_{x^2}^{x} \tfrac{21}{4} x^2 y \, dy \, dx = \tfrac{3}{20}$$

Probability---joint, marginal, conditional dst^{n}s

Ok, so now we're comfortable with the notion of a joint distribution being a surface (or set of point masses) over the xy-plane that describe the probability with which the random vector (X,Y) is in certain regions of the xy-plane. We saw examples of how to calculate probabilities by integrating the PDF f_{XY} over the relevant regions.

Probability---joint, marginal, conditional $dst^n s$

Ok, so now we're comfortable with the notion of a joint distribution being a surface (or set of point masses) over the xy-plane that describe the probability with which the random vector (X,Y) is in certain regions of the xy-plane. We saw examples of how to calculate probabilities by integrating the PDF f_{XY} over the relevant regions.

Now, we'll see some other things we can do with joint distributions. To start, we are going to see how to recover individual, or marginal, distributions from the joint.

Probability---joint, marginal, conditional dst^{n}s

For discrete:

$$f_X(x) = \sum_y f_{XY}(x,y)$$

For continuous:

$$f_X(x) = \int_y f_{XY}(x,y)dy$$

For discrete random variables, the intuition is clearer, perhaps. For a particular value of x, just sum up the joint distribution over all values of y to obtain the marginal distribution of X at that point.

For continuous random variables, just do the continuous analog.

Probability---discrete example

Esther and I play tennis. Like many sports, tennis players tend to rise (or fall) to the level their opponent is playing, so it would not be surprising to learn that we're more likely to both be playing well or both playing poorly.

What would the observation above suggest about the shape of a joint PF of our unforced errors by game? (By the way, a game is completed when a player wins four points by at least two points, but the score-keeping has this strange vestigial character: Love, 15, 30, 40, game.)

Probability---discrete example

This is the MIT Tennis Bubble

Esther and I play tennis. Like many sports, tennis players tend to rise (or fall) to the level their opponent is playing, so it would not be surprising to learn that we're more likely to both be playing well or both playing poorly.

What would the observation above suggest about the shape of a joint PF of our unforced errors by game? (By the way, a game is completed when a player wins four points by at least two points, but the score-keeping has this strange vestigial character: Love, 15, 30, 40, game.)

Probability---discrete example

Here's the joint PF of our unforced errors in a game:

f_{XY}	$Y=0$	1	2	3	4
$X=0$	1/4	1/8	0	0	0
1	1/8	1/8	1/16	0	0
2	0	1/16	1/16	1/32	0
3	0	0	1/32	1/16	1/64
4	0	0	0	1/64	1/32

Note the pattern---we either both have few unforced errors or both make a lot.

Probability---discrete example

Here's the joint PF of our unforced errors in a game:

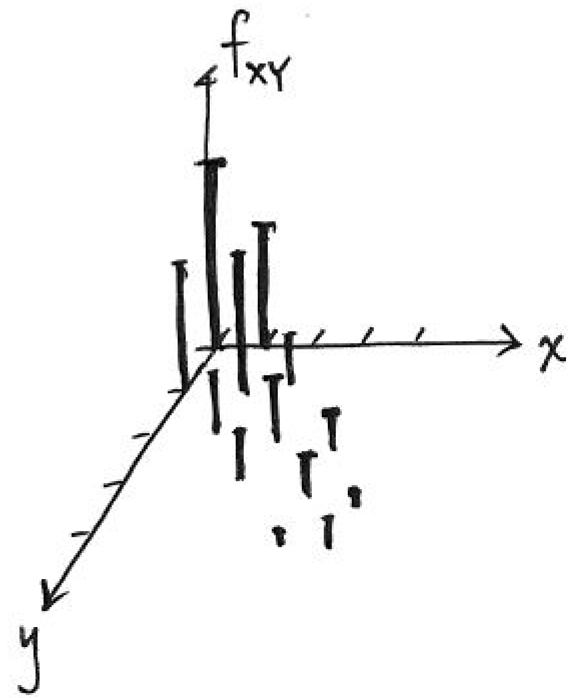
f_{XY}	Y = 0	1	2	3	4
X = 0	1/4	1/8	0	0	0
1	1/8	1/8	1/16	0	0
2	0	1/16	1/16	1/32	0
3	0	0	1/32	1/16	1/64
4	0	0	0	1/64	1/32

Apparently we're pretty good.

Note the pattern---we either both have few unforced errors or both make a lot.

Probability---discrete example

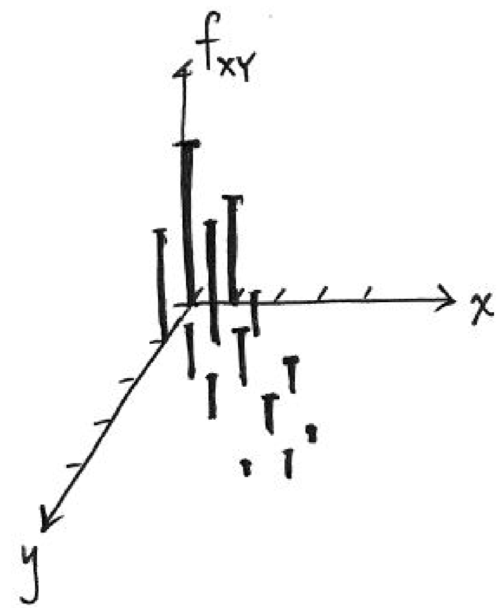
Here's the joint PF of our unforced errors in a game:



Note the pattern---we either both have few unforced errors or both make a lot.

Probability---discrete example

To calculate the marginal distributions, we just add up over values of the other random variable. Specifically, the probability that I make 2 unforced errors in a game is the probability that I make 2 and Esther makes 0 + the probability that I make 2 and Esther makes 1 + . . .



For a particular value of x, add up over all possible values of y.

Probability---discrete example

To calculate the marginal distributions, we just add up over values of the other random variable. Specifically, the probability that I make 2 unforced errors in a game is the probability that I make 2 and Esther makes 0 + the probability that I make 2 and Esther makes 1 + . . .

Since I have set up the joint PF to be symmetric,

$$f_X(x) = f_Y(y) = \begin{cases} 3/8 & x^* = 0 \\ 5/16 & x = 1 \\ 5/32 & x = 2 \\ 7/64 & x = 3 \\ 3/64 & x = 4 \end{cases}$$

* or y

Probability---discrete example

To calculate the marginal distributions, we just add up over values of the other random variable. Specifically, the probability that I make 2 unforced errors in a game is the probability that I make 2 and Esther makes 0 + the probability that I make 2 and Esther makes 1 + . . .

Since I have set up the joint PF to be symmetric,

You will never know the real truth.

$$f_X(x) = f_Y(y) = \begin{cases} 3/8 & x^* = 0 \\ 5/16 & x = 1 \\ 5/32 & x = 2 \\ 7/64 & x = 3 \\ 3/64 & x = 4 \end{cases}$$

* or y

Probability---discrete example

To calculate the marginal distributions, we just add up over values of the other random variable. Specifically, the probability that I make 2 unforced errors in a game is the probability that I make 2 and Esther makes 0 + the probability that I make 2 and Esther makes 1 + . . .

Since I have set up the joint PF to be symmetric,

$$f_X(x) = f_Y(y) = \begin{cases} 3/8 & x^* = 0 \\ 5/16 & x = 1 \\ 5/32 & x = 2 \\ 7/64 & x = 3 \\ 3/64 & x = 4 \end{cases}$$

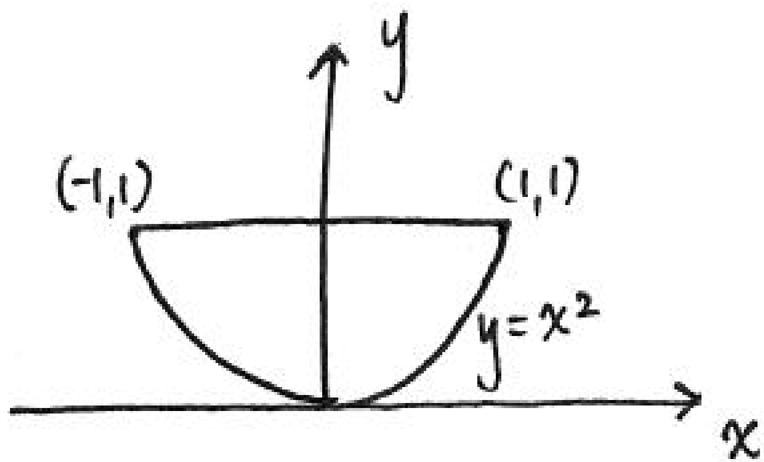
* or y

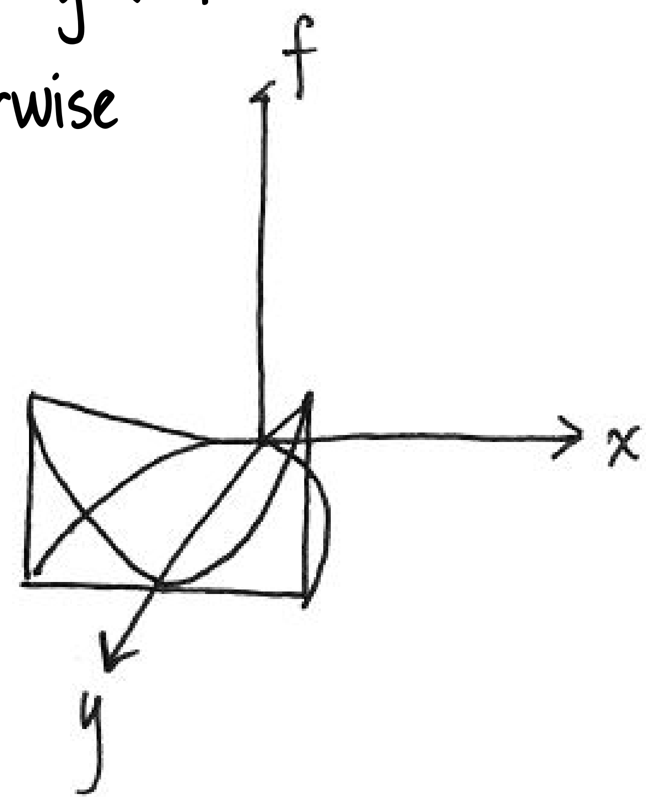
Probability---continuous example

We'll do something similar using a joint PDF. Let's return to our example from the beginning of lecture.

So $f_{XY}(x,y) = \begin{cases} (21/4)x^2y & x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$

Recall:

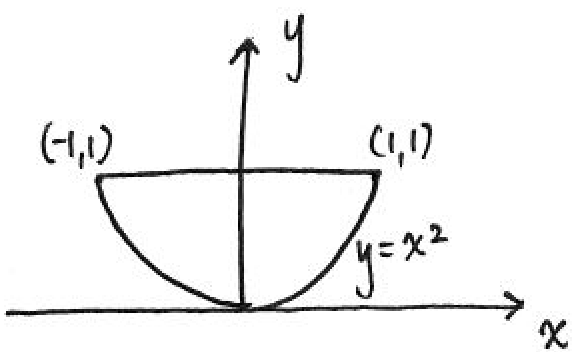




Probability---continuous example

$$f_{XY}(x,y) = \begin{cases} (21/4)x^2y & x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$$

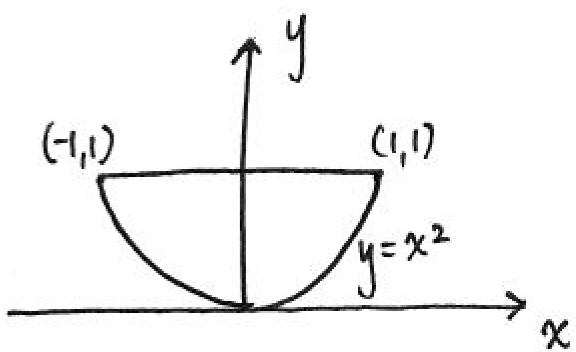
$$\text{So, } f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy = \int_{x^2}^{1} \tfrac{21}{4} x^2 y\,dy$$



Probability---continuous example

$$f_{XY}(x,y) = \begin{cases} (21/4)x^2y & x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$$

So, $f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy = \int_{x^2}^{1} \frac{21}{4} x^2 y\,dy \; I\{-1 \leq x \leq 1\}$



Probability---continuous example

$$f_{XY}(x,y) = \begin{cases} (21/4)x^2y & x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\text{So, } f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy = \int_{x^2}^{1} \tfrac{21}{4} x^2 y\,dy \; I\{-1 \leq x \leq 1\}$$

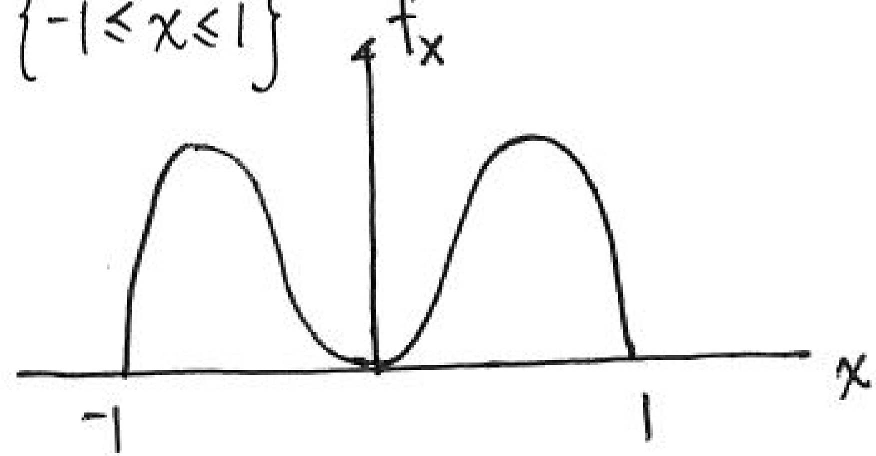
$$= \tfrac{21}{8} x^2 (1 - x^4) \; I\{-1 \leq x \leq 1\}$$

Probability---continuous example

$$f_{XY}(x,y) = \begin{cases} (21/4)x^2y & x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\text{So, } f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy = \int_{x^2}^{1} \tfrac{21}{4} x^2 y\,dy \; I\{-1 \leq x \leq 1\}$$

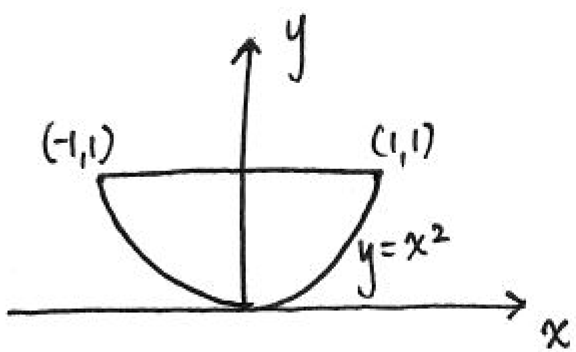
$$= \tfrac{21}{8} x^2 (1 - x^4) \; I\{-1 \leq x \leq 1\}$$



Probability---continuous example

$$f_{XY}(x,y) = \begin{cases} (21/4)x^2y & x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$$

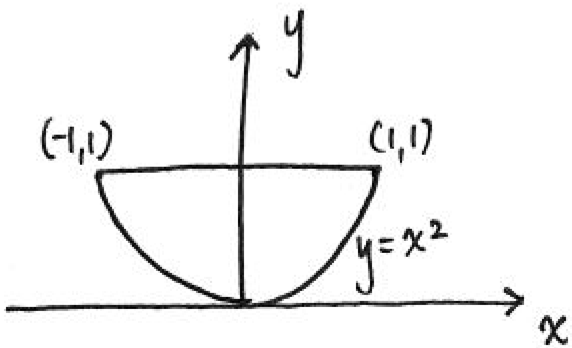
$$\text{So, } f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx = \int_{-ry}^{ry} \tfrac{21}{4} x^2 y\,dx$$



Probability---continuous example

$$f_{XY}(x,y) = \begin{cases} (21/4)x^2y & x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\text{So, } f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx = \int_{-\sqrt{y}}^{\sqrt{y}} \tfrac{21}{4}\, x^2 y \, dx \;\; I\{0 \le y \le 1\}$$



Probability---continuous example

$$f_{XY}(x,y) = \begin{cases} (21/4)x^2y & x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$$

So, $f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{21}{4} x^2 y \, dx \; I\{0 \leq y \leq 1\}$

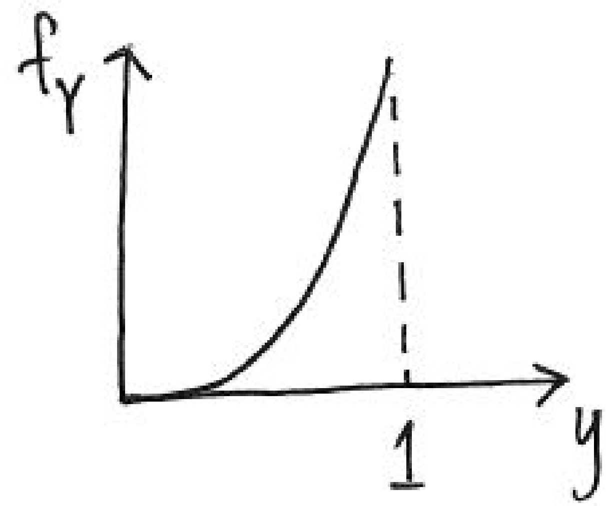
$$= \frac{7}{2} y^{5/2} \; I\{0 \leq y \leq 1\}$$

Probability---continuous example

$$f_{XY}(x,y) = \begin{cases} (21/4)x^2y & x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\text{So, } f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx = \int_{-\sqrt{y}}^{\sqrt{y}} \tfrac{21}{4} x^2 y \,dx \; I\{0 \le y \le 1\}$$

$$= \tfrac{7}{2} y^{5/2} \; I\{0 \le y \le 1\}$$



Probability---joint, marginal, conditional dst^{n}s

We have seen: if you know the joint distribution, you can recover the marginal distributions of the constituent random variables.

If you know the marginals, can you construct the joint?

Probability---joint, marginal, conditional dst^{n}s

We have seen: if you know the joint distribution, you can recover the marginal distributions of the constituent random variables.

If you know the marginals, can you construct the joint?

In general, no. We need another crucial piece of information: the relationship between the random variables.

Probability---independence of RVs

X&Y are independent if $P(X \subset A$ & $Y \subset B) = P(X \subset A)P(Y \subset B)$ for all regions A&B.

Well, that could certainly be hard to check.

However, that definition does imply that $F_{XY}(x,y) = F_X(x)F_Y(y)$, which could be useful. Furthermore, one can prove that X&Y are independent iff $f_{XY}(x,y,) = f_X(x)f_Y(y)$. This condition is easy to check and useful.

In fact, if X&Y are both continuous with joint PDF f_{XY}, X&Y are independent iff $f_{XY}(x,y,) = g(x)h(y)$ where g is a non-negative function of x alone and h the same with y.

Example from previous lecture

Probability---example

Suppose after hours of writing lecture notes, I develop a splitting headache. I rummage around in my drawer and find one tablet of naproxen and one of acetaminophen. I take both. Let X be the effective period of naproxen. Let Y be the effective period of acetaminophen. Suppose

$$f_{XY}(x,y) = \lambda^2 \exp\{-\lambda(x+y)\} \quad \text{for } x,y >= 0$$

What is the probability that my headache comes back within three hours?

Are X&Y independent here?

Example from previous lecture

Probability---example

Suppose after hours of writing lecture notes, I develop a splitting headache. I rummage around in my drawer and find one tablet of naproxen and one of acetaminophen. I take both. Let X be the effective period of naproxen. Let Y be the effective period of acetaminophen. Suppose

$$f_{XY}(x,y) = \lambda^2 \exp\{-\lambda(x+y)\} \quad \text{for } x,y >= 0$$

What is the probability that my headache comes back within three hours?

Are X&Y independent here? Yes! This PDF can be factored, and X&Y have same dist^n, in fact.

Example from earlier in this lecture

Probability---example

Last time we ended with an example that generated a lot of questions. We were computing probabilities from a joint PDF. Let's do another example.

Suppose we have $f_{XY}(x,y) = \begin{cases} cx^2y & \text{for } x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$

First, let's figure out what c is. How? We know this joint PDF has to integrate to 1.

Are X&Y independent here?

Example from earlier in this lecture

Probability---example

Last time we ended with an example that generated a lot of questions. We were computing probabilities from a joint PDF. Let's do another example.

Suppose we have $f_{XY}(x,y) = \begin{cases} cx^2y & \text{for } x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$

First, let's figure out what c is. How? We know this joint PDF has to integrate to 1.

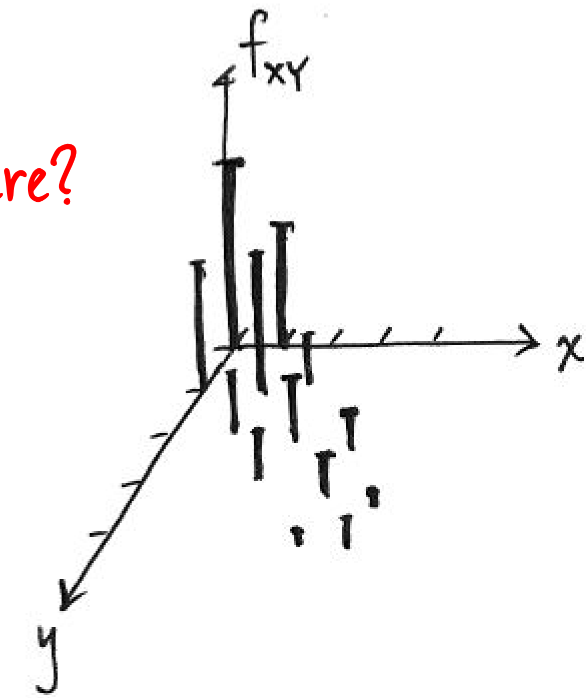
How about the support? The values that Y can take on depend on the value of x. So, no.

Example from earlier in this lecture

Probability---discrete example

Here's the joint PF of our unforced errors in a game:

Are X&Y independent here?

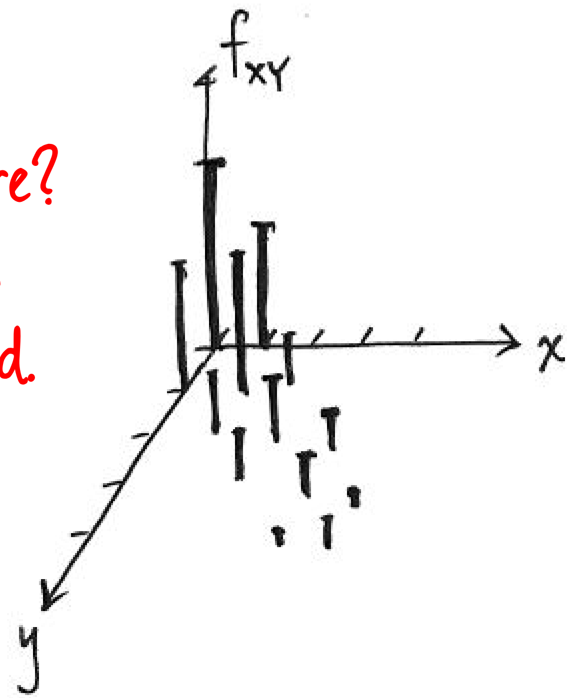


Note the pattern---we either both have few unforced errors or both make a lot.

Example from earlier in this lecture

Probability---discrete example

Here's the joint PF of our unforced errors in a game:



Are X&Y independent here?
Easy to think of a region where definition is violated.
So, no.

Note the pattern---we either both have few unforced errors or both make a lot.

Probability---independence of RVs

For discrete random variables, if you have a table representing their joint PF, the two variables are independent iff the rows of the table are proportional to one another (linearly dependent) iff the columns of the table are proportional to one another.

Why?

Probability---independence of RVs

For discrete random variables, if you have a table representing their joint PF, the two variables are independent iff the rows of the table are proportional to one another (linearly dependent) iff the columns of the table are proportional to one another.

Why?

Independence means that the product of the marginals is equal to the joint so each column of the table is just a multiple of every other column, the multiple being the ratio of marginal probabilities associated with the two columns.

Probability---independence of RVs

An example of independent discrete RVs

f_{XY}	Y: 0	1	2	3	4
X: 0	1/128	1/64	1/64	1/64	1/128
1	1/32	1/16	1/16	1/16	1/32
2	3/64	3/32	3/32	3/32	3/64
3	1/32	1/16	1/16	1/16	1/32
4	1/128	1/64	1/64	1/64	1/128

Probability---independence of RVs

An example of independent discrete RVs

f_{XY}	$Y=0$	$Y=1$	$Y=2$	$Y=3$	$Y=4$
$X=0$	$1/128$	$1/64$	$1/64$	$1/64$	$1/128$
$X=1$	$1/32$	$1/16$	$1/16$	$1/16$	$1/32$
$X=2$	$3/64$	$3/32$	$3/32$	$3/32$	$3/64$
$X=3$	$1/32$	$1/16$	$1/16$	$1/16$	$1/32$
$X=4$	$1/128$	$1/64$	$1/64$	$1/64$	$1/128$

We can think of this as a possible joint PF of unforced errors if my and Esther's unforced errors were independent, instead of having the character that we either both made a lot or both made few.

Probability---joint, marginal, conditional dst^{n}s

Similar to the idea of conditional probability, we want to introduce the conditional distribution, which allows one to "update" the distribution of a random variable, if necessary, given relevant information.

The conditional PDF of Y given X is

$$f_{Y|X}(y|x) = f_{XY}(x,y)/f_X(x)$$

$$(= P(Y=y|X=x) \text{ for } X,Y \text{ discrete })$$

Note the conditional PDFs are often written as a function of both x and y. For a particular value of the conditioning variable, though, they behave just like a marginal PDF.

Probability---joint, marginal, conditional dst^{n}s

Similar to the idea of conditional probability, we want to introduce the conditional distribution, which allows one to "update" the distribution of a random variable, if necessary, given relevant information.

The conditional PDF of Y given X is

$$f_{Y|X}(y|x) = f_{XY}(x,y)/f_X(x)$$

$$(= P(Y=y|X=x) \text{ for } X,Y \text{ discrete })$$

Take the relevant slice of the joint PDF and blow it up so that it is a PDF itself (i.e., integrates to 1).

Note the conditional PDFs are often written as a function of both x and y. For a particular value of the conditioning variable, though, they behave just like a marginal PDF.

Probability---example

Recall the original joint PF of unforced errors:

f_{XY}		Y				
		0	1	2	3	4
X	0	1/4	1/8	0	0	0
	1	1/8	1/8	1/16	0	0
	2	0	1/16	1/16	1/32	0
	3	0	0	1/32	1/16	1/64
	4	0	0	0	1/64	1/32

Probability---example

Recall the original joint PF of unforced errors:

Take a slice, say, Y=2.

f_{XY}	Y = 0	1	2	3	4
X = 0	1/4	1/8	0	0	0
1	1/8	1/8	1/16	0	0
2	0	1/16	1/16	1/32	0
3	0	0	1/32	1/16	1/64
4	0	0	0	1/64	1/32

Probability---example

Recall the original joint PF of unforced errors:

f_{XY}	Y = 0	1	2	3	4
X = 0	1/4	1/8	0	0	0
1	1/8	1/8	1/16	0	0
2	0	1/16	1/16	1/32	0
3	0	0	1/32	1/16	1/64
4	0	0	0	1/64	1/32

Take a slice, say, Y=2.

That slice is not a PDF on its own (doesn't integrate to 1), so we need to blow it up.

Probability---example

Recall the original joint PF of unforced errors:

f_{XY}	Y = 0	1	2	3	4
X = 0	1/4	1/8	0	0	0
1	1/8	1/8	1/16	0	0
2	0	1/16	1/16	1/32	0
3	0	0	1/32	1/16	1/64
4	0	0	0	1/64	1/32

Take a slice, say, Y=2.

That slice is not a PDF on its own (doesn't integrate to 1), so we need to blow it up.

The factor we blow it up by is P(Y=2), or 5/32.

Probability---example

Recall the original joint PF of unforced errors:

f_{XY}	Y=0	1	2	3	4
X=0	1/4	1/8	0	0	0
1	1/8	1/8	1/16	0	0
2	0	1/16	1/16	1/32	0
3	0	0	1/32	1/16	1/64
4	0	0	0	1/64	1/32

So,

$$f_{X|Y}(x|y=2) = \begin{cases} 2/5 & \text{for } x = 1,2 \\ 1/5 & \text{for } x = 3 \\ 0 & \text{otherwise} \end{cases}$$

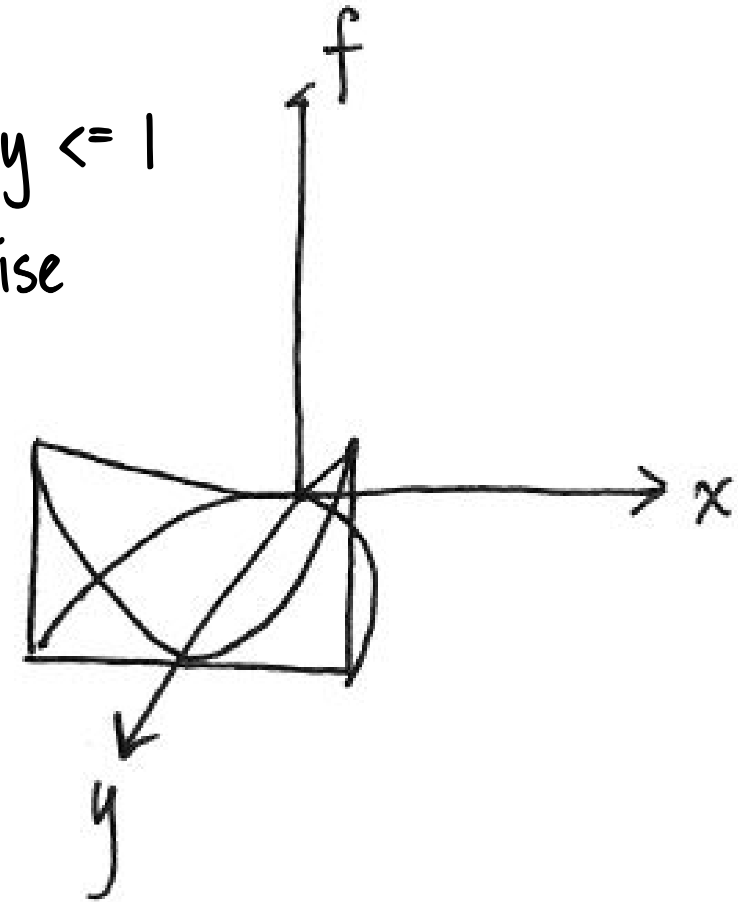
Probability---example

Let's also compute a conditional PDF, using the earlier example.

Probability---example

Let's also compute a conditional PDF, using the earlier example.

$$f_{XY}(x,y) = \begin{cases} (21/4)x^2y & x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$$

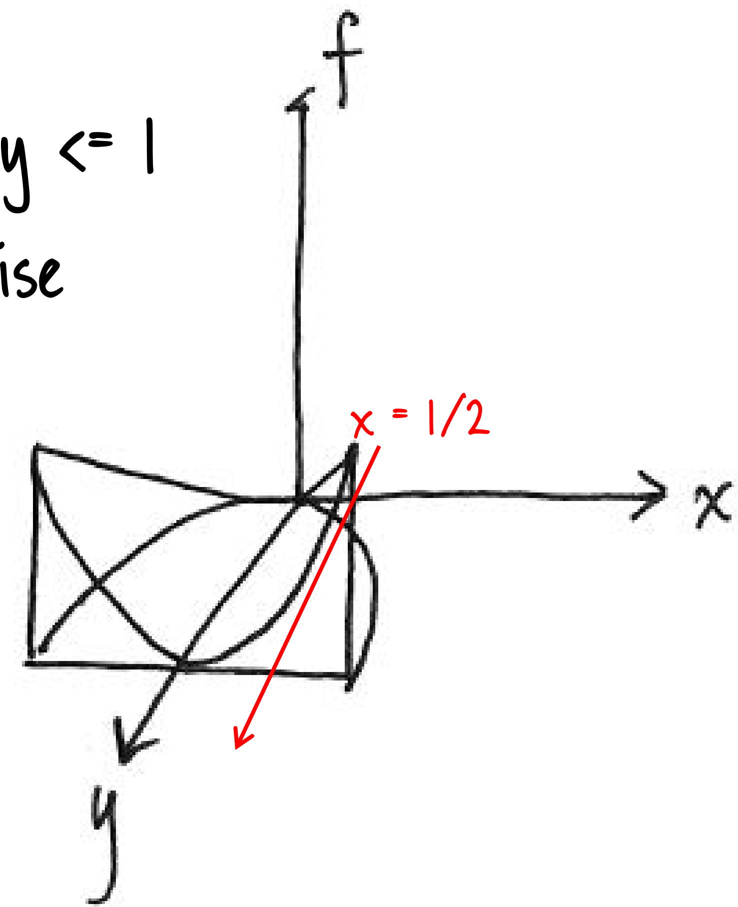


Probability---example

Let's also compute a conditional PDF, using the earlier example.

$$f_{XY}(x,y) = \begin{cases} (21/4)x^2y & x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$$

We will calculate the conditional PDF of Y as a function of x, but for a particular value of x, think of this function as taking a cross-sectional slice of the joint PDF and suitably normalizing it.



Probability---example

Let's also compute a conditional PDF, using the earlier example.

$$f_{XY}(x,y) = \begin{cases} (21/4)x^2y & x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$$

Recall $f_X(x) = \frac{21}{8} x^2(1-x^4)\ I\{-1 \leq x \leq 1\}$

(Note that this PDF is non-zero for all x in [-1,1] except 0. A PDF conditional on x will only be defined for non-zero values of x.)

Probability---example

Let's also compute a conditional PDF, using the earlier example.

$$f_{XY}(x,y) = \begin{cases} (21/4)x^2y & x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$$

Recall $f_X(x) = \frac{21}{8} x^2(1-x^4)\ I\{-1 \leq x \leq 1\}$

Now recall that $f_{Y|X}(y|x) = f_{XY}(x,y)/f_X(x)$.

Probability---example

Let's also compute a conditional PDF, using the earlier example.

$$f_{XY}(x,y) = \begin{cases} (21/4)x^2y & x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$$

Recall $f_X(x) = 21/8\, x^2(1-x^4)\, I\{-1 \le x \le 1\}$

Now recall that $f_{Y|X}(y|x) = f_{XY}(x,y)/f_X(x)$.

So $f_{Y|X}(y|x) = \begin{cases} 2y/(1-x^4) & x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$

Probability---example

Let's also compute a conditional PDF, using the earlier example.

$$f_{XY}(x,y) = \begin{cases} (21/4)x^2y & x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$$

Recall $f_X(x) = \frac{21}{8} x^2(1-x^4)\ I\{-1 \le x \le 1\}$

Now recall that $f_{Y|X}(y|x) = f_{XY}(x,y)/f_X(x)$.

So $f_{Y|X}(y|x) = \begin{cases} 2y/(1-x^4) & x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$

Only defined for non-zero values of x.

Probability---example

So $f_{Y|X}(y|x) = \begin{cases} 2y/(1-x^4) & x^2 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$

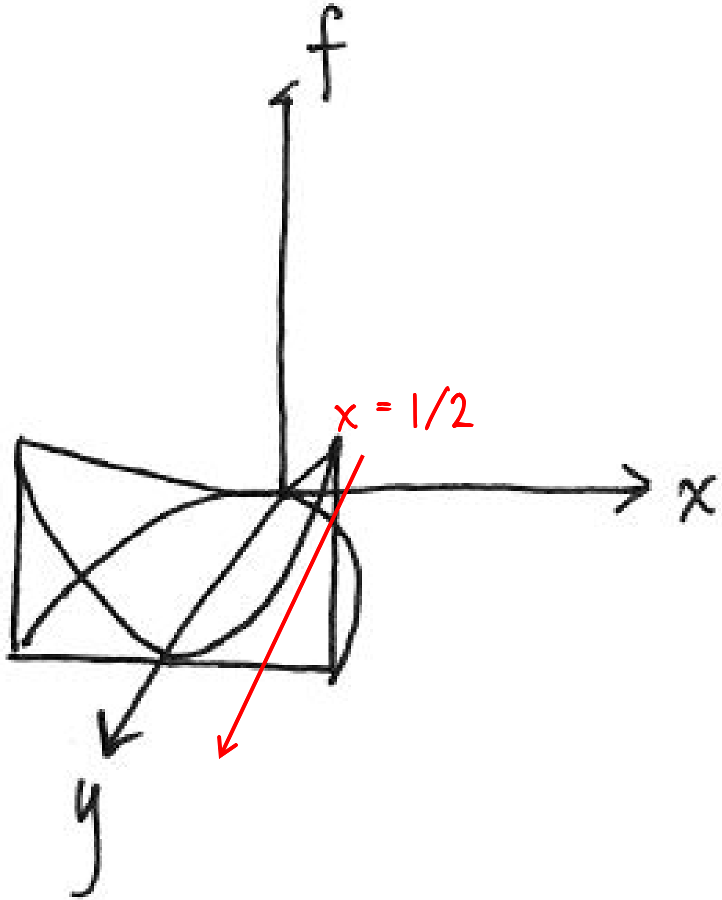
Only defined for non-zero values of x.

Plug in any legitimate value of x, say x = 1/2.

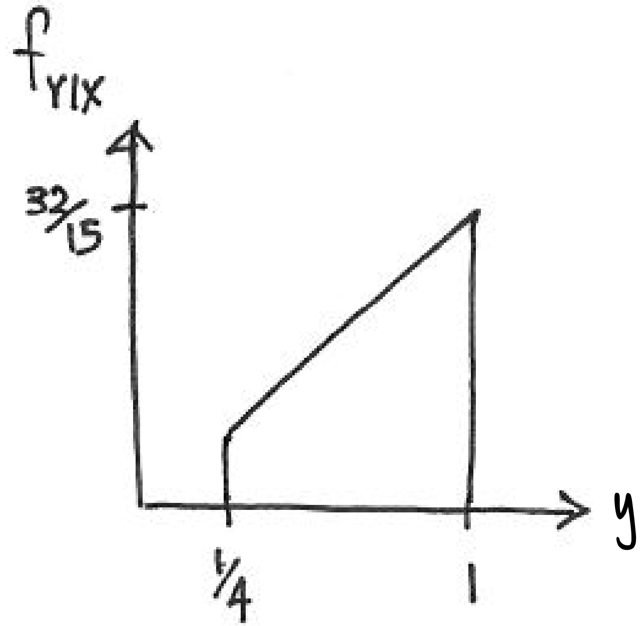
We get $f_{Y|X}(y|x) = \begin{cases} (32/15)y & 1/4 <= y <= 1 \\ 0 & \text{otherwise} \end{cases}$

Probability---example

Here's the picture from before:



Here is the conditional PDF at $x = 1/2$:



Probability---joint, marginal, conditional dst^{n}s

Not surprisingly, there is a relationship between conditional distributions and independence.

$$f_{Y|X}(y|x) = f_Y(y) \text{ iff } f_{XY}(x,y) = f_X(x)f_Y(y)$$
$$\text{iff } X \,\&\, Y \text{ independent}$$

If two random variables are independent, knowing something about the realizations of one doesn't tell you anything about the distribution of the other.

Probability---functions of RVs

As I've emphasized before, we need to start with a foundation in probability because we can't talk about how functions of random variables behave until we know about how random variables behave. And we can't talk about statistics, such as the sample mean, until we know about functions of random variables because that's precisely what a statistic is.

So now we start our discussion of functions of random variables.

Probability---functions of RVs

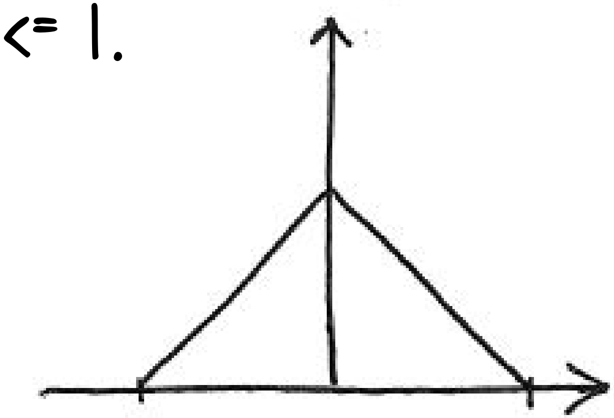
Basic idea: we have a random variable X and its PDF. We want to know how a new random variable $Y = h(X)$ is distributed. (More complicated: we have random variables $X_1, X_2, X_3, \ldots$, and we want to know how $h(X)$, a function of the entire random vector X, is distributed.)

Probability---functions of RVs

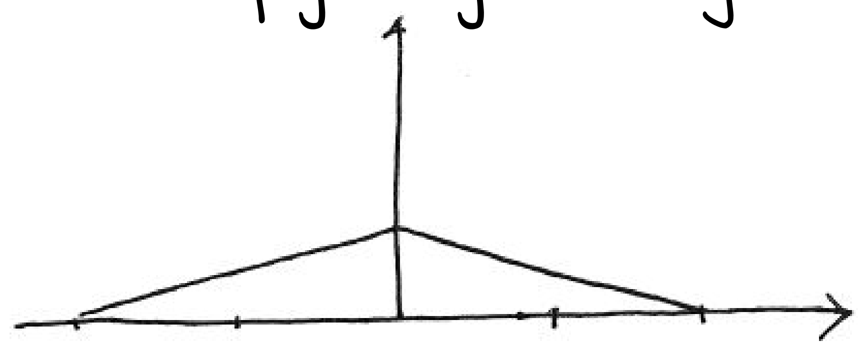
Let's start with a graphical example.

We want the distribution of $Y = |2X| + 3$, where X has PDF $f_X(x) = 1 - |x|$ for $-1 <= x <= 1$.

Here's the PDF of X:



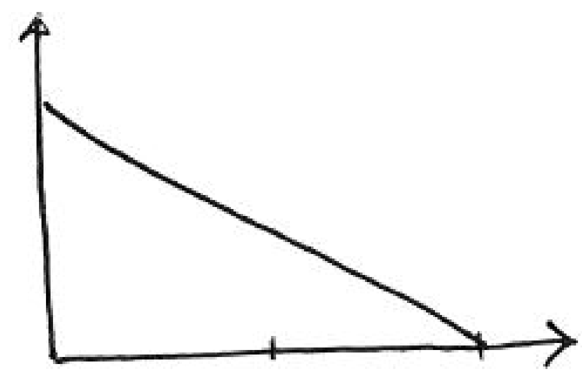
Here's what happens when we multiply it by 2---it gets stretched out:



Probability---functions of RVs

Let's start with a graphical example.

We want the distribution of $Y = |2X| + 3$, where X has PDF $f_X(x) = 1 - |x|$ for $-1 <= x <= 1$.



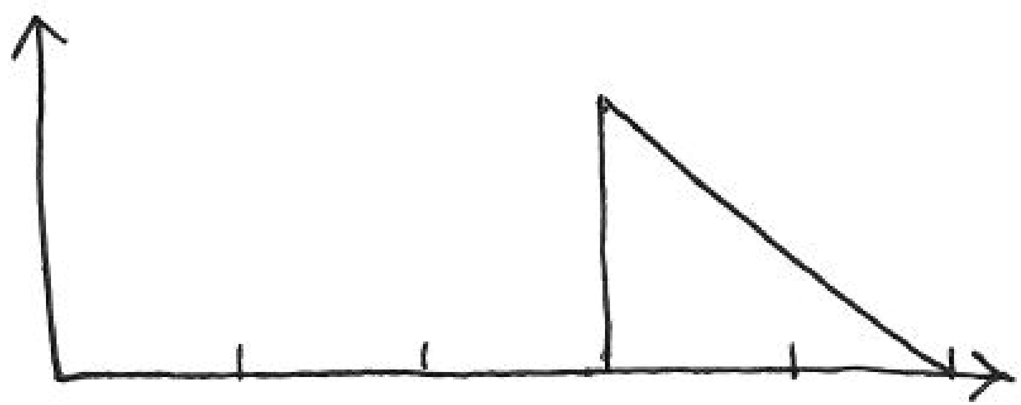
Now take the absolute value---all of the density over negative values gets folded over onto the positive values:

Probability---functions of RVs

Let's start with a graphical example.

We want the distribution of $Y = |2X| + 3$, where X has PDF $f_X(x) = 1 - |x|$ for $-1 <= x <= 1$.

Now finally let's add 3---shifts entire distribution over:

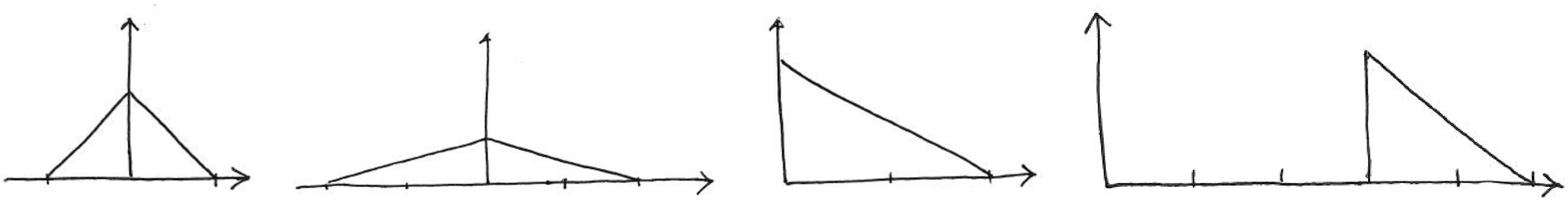


Probability---functions of RVs

Let's start with a graphical example.

We want the distribution of $Y = |2X| + 3$, where X has PDF $f_X(x) = 1 - |x|$ for $-1 <= x <= 1$.

Keep in mind that throughout this process, the distribution always retained the properties of a PDF, in particular, it integrated to 1.



Probability---functions of RVs

One more example that should help firm up your intuition of what a function of a RV does:

Suppose we have $X \sim U[0,1]$. What function g can transform X to a $B(2,.5)$?

Probability---functions of RVs

One more example that should help firm up your intuition of what a function of a RV does:

Suppose we have $X \sim U[0,1]$. What function g can transform X to a $B(2,.5)$?

$f_X(x) = 1 \quad 0 <= x <= 1$

$$f_Y(y) = \begin{cases} 1/4 & x = 0,\ 2 \\ 1/2 & x = 1 \end{cases}$$

Probability---functions of RVs

One more example that should help firm up your intuition of what a function of a RV does:

Suppose we have $X \sim U[0,1]$. What function g can transform X to a $B(2,.5)$?

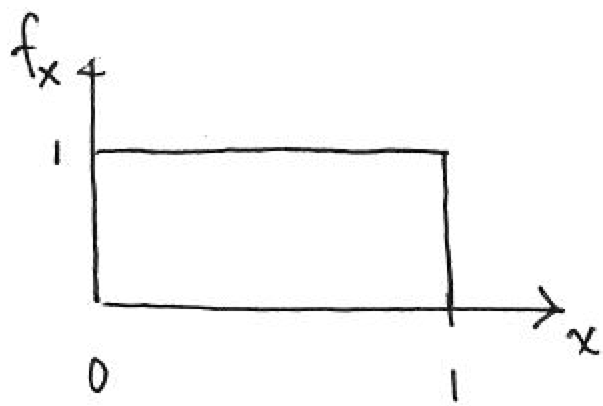
$f_X(x) = 1 \quad 0 <= x <= 1$

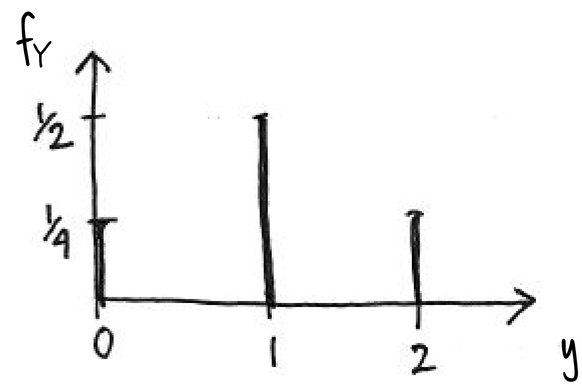
$$f_Y(y) = \begin{cases} 1/4 & x = 0, 2 \\ 1/2 & x = 1 \end{cases}$$

How about just chopping up the unit interval and mapping appropriate sized sub-intervals to each point mass?

Probability---functions of RVs

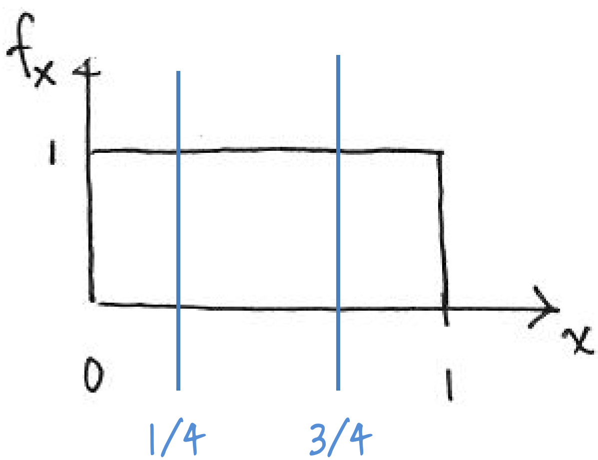
How about just chopping up the unit interval and mapping appropriate sized sub-intervals to each point mass?

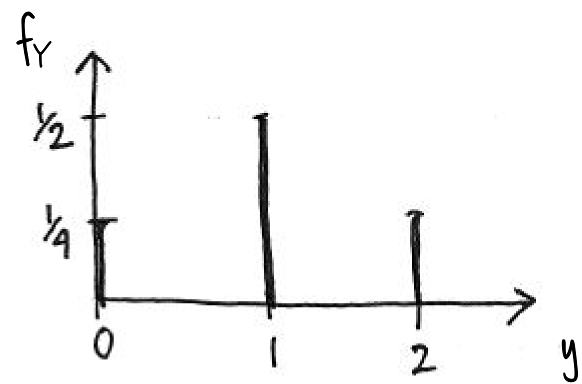




Probability---functions of RVs

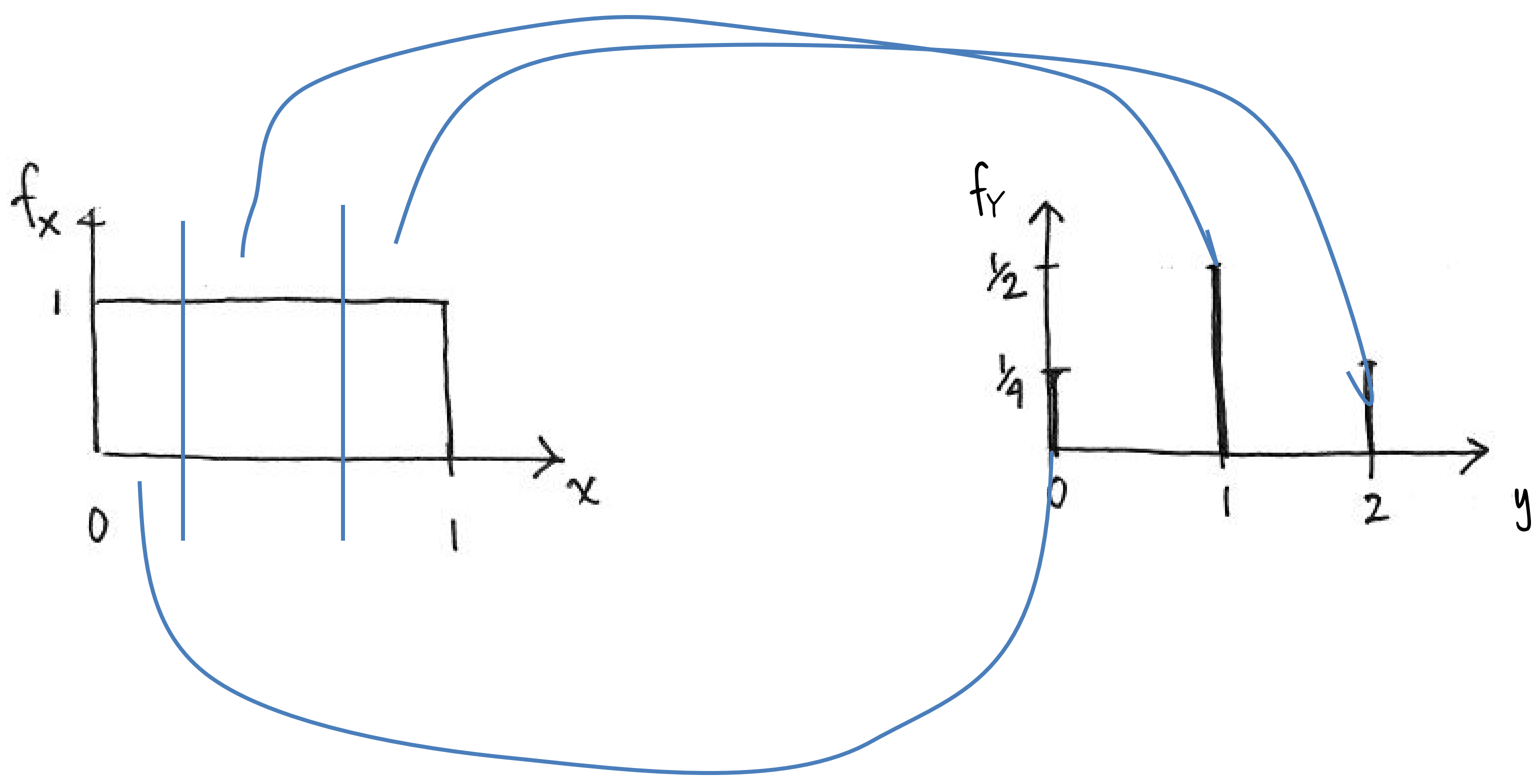
How about just chopping up the unit interval and mapping appropriate sized sub-intervals to each point mass?





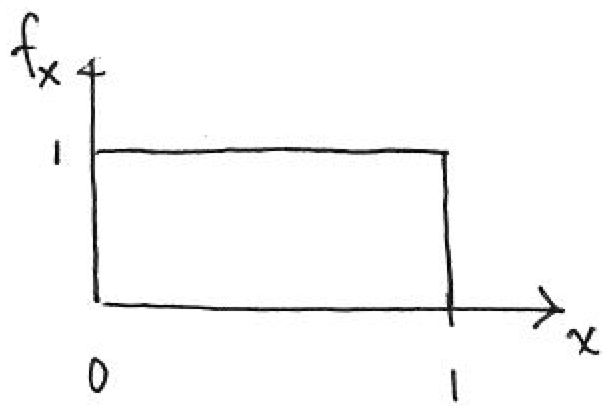
Probability---functions of RVs

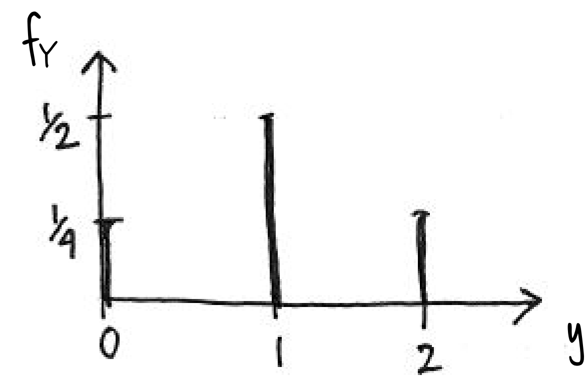
How about just chopping up the unit interval and mapping appropriate sized sub-intervals to each point mass?



Probability---functions of RVs

How about just chopping up the unit interval and mapping appropriate sized sub-intervals to each point mass?

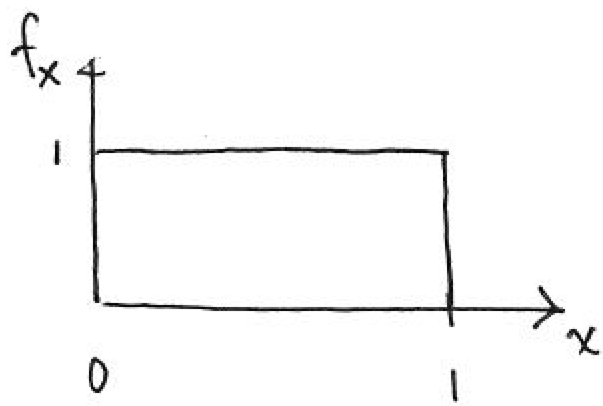


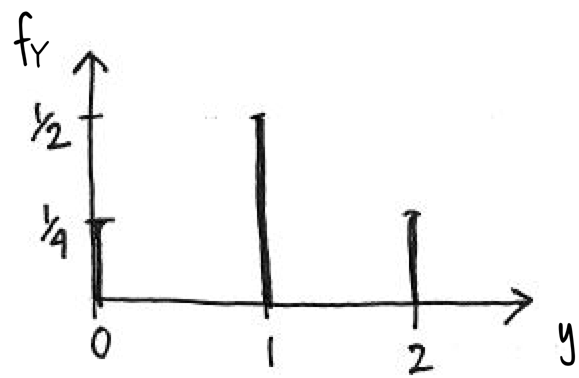


$$\text{So } Y = \begin{cases} 0 & x <= 1/4 \\ 1 & 1/4 <= x <= 3/4 \\ 2 & 3/4 <= x \end{cases}$$

Probability---functions of RVs

How about just chopping up the unit interval and mapping appropriate sized sub-intervals to each point mass?





$$
\text{So } Y = \begin{cases} 0 & x <= 1/4 \\ 1 & 1/4 <= x <= 3/4 \\ 2 & 3/4 <= x \end{cases}
$$

This is one possible function---it is certainly not unique.

Probability---functions of RVs

There are various methods one can use to figure out the distribution of a function of random variables. Which methods one can use on a particular problem depend on whether the original random variable is discrete or continuous, whether there is just one random variable or a random vector, and whether the function is invertible or not. We will not learn all of the methods here. Instead we'll learn one important method and also see a lot of examples that can be applied somewhat generally.