Conditional probability when conditioning on a RV

(Chapter 5.7)

Recall from Topic 2, Chapters 3.4 and 4.2.2, about conditioning on an event:

$$p_X(x|C) = \frac{P(X=x \cap C)}{P(C)} \quad \text{when } P(C)>0$$

$$F_X(x|C) = \frac{P(X\le x \cap C)}{P(C)} \quad \text{when } P(C)>0$$

→ These are general for any event C.

when C is an event that depends on X,

$$f_X(x|C) = \begin{cases} \frac{f_X(x)}{P(C)} & \text{when } x\in C \text{ and } P(C)>0 \\ 0 & \text{else} \end{cases}$$

This can be directly applied to joint RVs

ex: $$F_{XY}(x,y|C) = \frac{P(X\le x, Y\le y \text{ and } C)}{P(C)} \quad \text{when } P(C)>0$$

<u>But</u> what if the event $C=\{X=x\}$, and X is a continuous RV?

The basic concept is:
Given a sample (i.e., a specific outcome), and you know $X=x$, what's $f_Y(y|X=x)$?

Let's start with a discrete example,
to find a conditional pmf $p_X(x|y)$

This is $p_{XY}(x,y)$

$x \backslash y$	1	2	3
1	1/24	4/24	1/24
2	2/24	3/24	1/24
3	3/24	2/24	1/24
4	4/24	1/24	1/24

what's
$p_Y(y)$?
$p_X(x|Y=1)$?
$p_Y(y|X=4)$?
$p_X(x|y)$?

To find $p_Y(y)$:

this is the same $p_{XY}(x,y)$

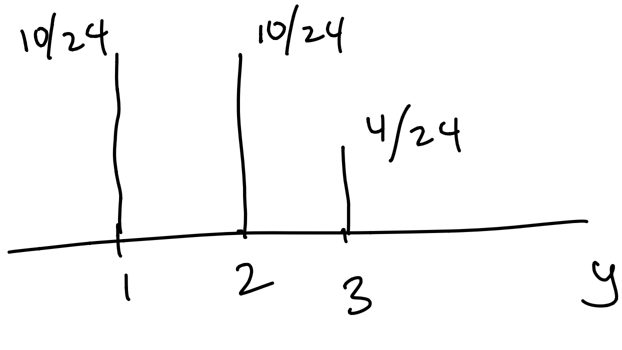
$x \backslash y$	1	2	3
1	1/24	4/24	1/24
2	2/24	3/24	1/24
3	3/24	2/24	1/24
4	4/24	1/24	1/24
	10/24	10/24	4/24

marginal pmf

$$p_Y(y) = \sum_x p_{XY}(x,y)$$

sum down each column

$p_Y(y)$



To find $p_X(x \mid y=1)$ conditional PMF

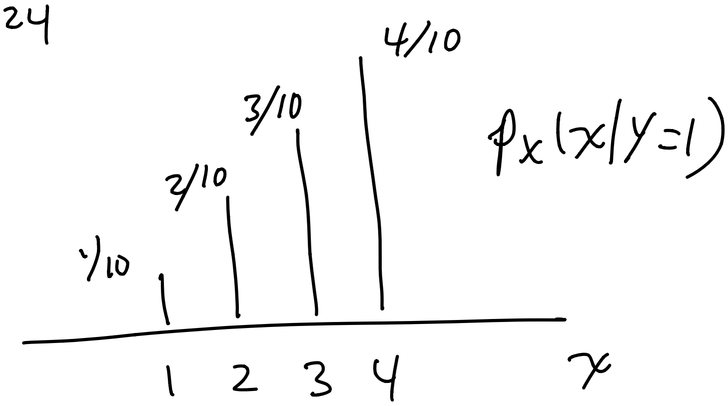
zoom in and scale by $p_Y(1)$

$X \backslash Y$	1	2	3
1	1/24	4/24	1/24
2	2/24	3/24	1/24
3	3/24	2/24	1/24
4	4/24	1/24	1/24

same $p_{XY}(x,y)$

$p_X(x \mid Y=1)$ is a PMF.

$$= \frac{p_{XY}(x, Y=1)}{p_Y(Y=1)}$$

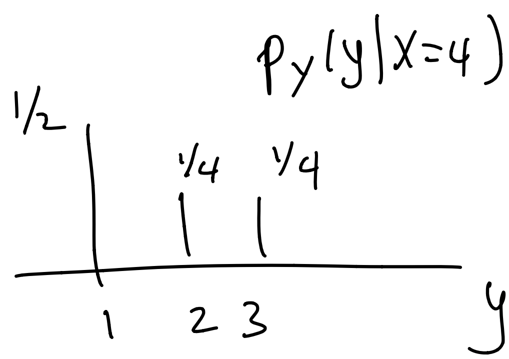


To find $p_Y(y \mid X=4)$ conditional pmf given $X=4$

$X \backslash Y$	1	2	3
1	1/24	4/24	1/24
2	2/24	3/24	1/24
3	3/24	2/24	1/24
4	4/24	1/24	1/24

same $p_{XY}(x,y)$

zoom in and scale to be sure it's a PMF



To find $p_X(x|y)$: $\boxed{p_X(x|y) = \dfrac{p_{XY}(x,y)}{p_Y(y)}}$

This is a function of both x <u>and</u> y.

general definition of conditional PMF

This is still $p_{XY}(x,y)$

$x \backslash y$	1	2	3
1	$1/24$	$4/24$	$1/24$
2	$2/24$	$3/24$	$1/24$
3	$3/24$	$2/24$	$1/24$
4	$4/24$	$1/24$	$1/24$

For a specific value of y, $p_X(x|y)$ is a PMF.

So $$\sum_x p_X(x|y) = 1$$

but $\sum_y p_X(x|y)$ could be anything

$p_X(x|y)$ (not $p_{XY}(x,y)$)

conditional pmf $p_X(x|y)$

$x \backslash y$	1	2	3
1	$1/10$	$4/10$	$1/4$
2	$2/10$	$3/10$	$1/4$
3	$3/10$	$2/10$	$1/4$
4	$4/10$	$1/10$	$1/4$

each is a pmf

Conditioning on $\{X = x\}$ (or $\{Y = y\}$) when X (or Y) is a continuous RV.

$P(X = x) = 0$ for a continuous RV.

punch line

$$\boxed{f_X(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} \quad \text{if } f_Y(y) > 0}$$

why?
use a small-interval approximation.

Recall when understanding the pdf "density",

we said $\Delta \cdot f_X(x) \approx P(x < X \leq x + \Delta)$
if Δ is a small increment

we can say similar things about

$$\Delta_y f_Y(y) \approx P(y < Y \leq y + \Delta_y)$$

$$\Delta_x \Delta_y f_{XY}(x, y) \approx P(x < X \leq x + \Delta_x, y < Y \leq y + \Delta_y)$$

for small Δ_y and Δ_x

And what we want to find is an expression for

$$\Delta_x f_X(x|y) \approx P(x < X \leq x + \Delta_x \mid y < Y \leq y + \Delta_y)$$

for small Δ_x and Δ_y

If we define A to be a slice of X,
and B to be a slice of Y

$$A = \{ x < X \le x + \Delta_x \}$$

$$B = \{ y < Y \le y + \Delta_y \}$$

then $\Delta_x f_X(x|y) \approx \dfrac{P(A \cap B)}{P(B)} = P(A|B)$

and $P(A \cap B) \approx f_{XY}(x, y)\, \Delta_x \Delta_y$

$$P(A) \approx f_X(x)\, \Delta_x$$

$$P(B) \approx f_Y(y)\, \Delta_y$$

Combining:

$$\Delta_x f_X(x|y) \approx \frac{f_{XY}(x, y)\, \Delta_x \Delta_y}{f_Y(y)\, \Delta_y}$$

(the probability X lies in a small slice near x, given that Y lies in a small slice near y)

(the probability X and Y lie in a small box near (x, y), divided by the probability Y lies in a small slice near y)

so
$$\boxed{f_X(x|y) = \frac{f_{XY}(x, y)}{f_Y(y)}}$$

Note: this has not been an exact proof by any stretch of the imagination, but it is meant to be conceptually useful

Conditional pdf definitions.

$$f_X(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} \qquad \text{if } f_Y(y) > 0$$

$$f_Y(y|x) = \frac{f_{XY}(x,y)}{f_X(x)} \qquad \text{if } f_X(x) > 0$$

Note: $\displaystyle\int_{-\infty}^{\infty} f_X(x/y)\,dx = 1$

but $\displaystyle\int_{-\infty}^{\infty} f_X(x/y)\,dy$ could be anything

For a specific value of y, $f_X(x/y)$ is a pdf.

But $f_X(x|y)$ <u>does</u> depend on both x and y.

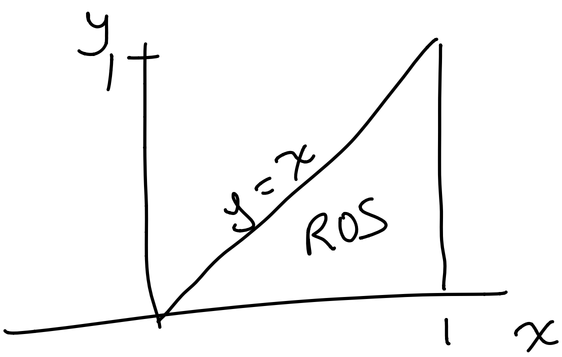
The ROS for $f_X(x|y)$ is the intersection of the ROS of $f_{XY}(x,y)$ and $f_Y(y)$

$\boxed{\text{Always indicate your ROS!}}$

Example (same as before)

Two random variables are uniformly distributed over the triangle formed by $(0,0)$, $(1,1)$, $(1,0)$. Find the conditional pdfs $f_X(x|y)$ and $f_Y(y|x)$.

$$f_{XY}(x,y) = \begin{cases} 2 & 0 \le y \le x \le 1 \\ 0 & \text{else} \end{cases}$$



$$f_X(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} \quad \text{if } f_Y(y) > 0$$

therefore we need to find $f_Y(y)$ by integrating over all x

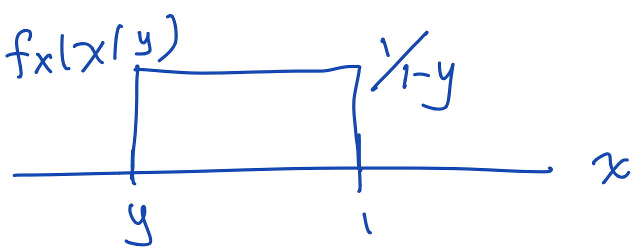
From before,

$$f_Y(y) = \begin{cases} 2(1-y) & 0 \le y \le 1 \\ 0 & \text{else} \end{cases}$$

strict inequality

And then $$f_X(x|y) = \frac{2}{2(1-y)} \quad 0 \le y \le x < 1$$

For a given value of y, x is a uniform RV:



Similarly, $f_Y(y|x) = \dfrac{f_{XY}(x,y)}{f_X(x)}$ if $f_X(x) > 0$

1st compute marginal. From before,

strict inequality

$$f_X(x) = \begin{cases} 2x & 0 \le x \le 1 \\ 0 & \text{else} \end{cases}$$

So $f_Y(y|x) = \dfrac{2}{2x} = \dfrac{1}{x}$ for $0 \le y < x \le 1$

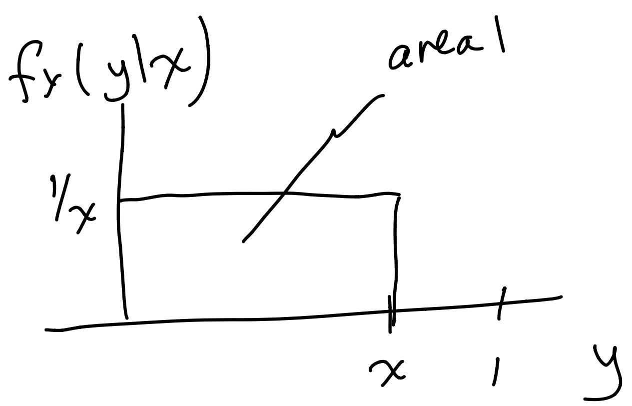
Again, for a given value of x,

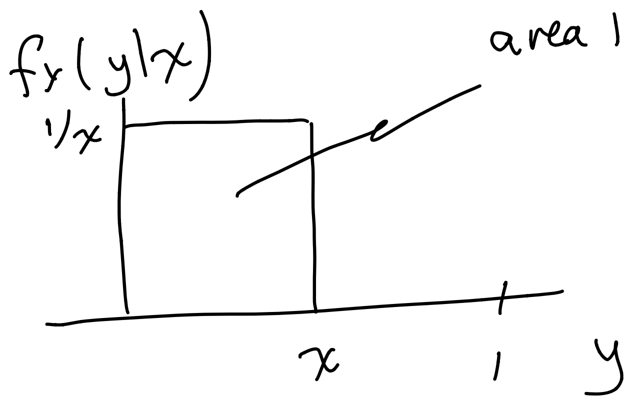
$f_Y(y|x)$ is a uniform RV

with width x and height $1/x$

As x gets smaller, $f_Y(y|x)$ gets

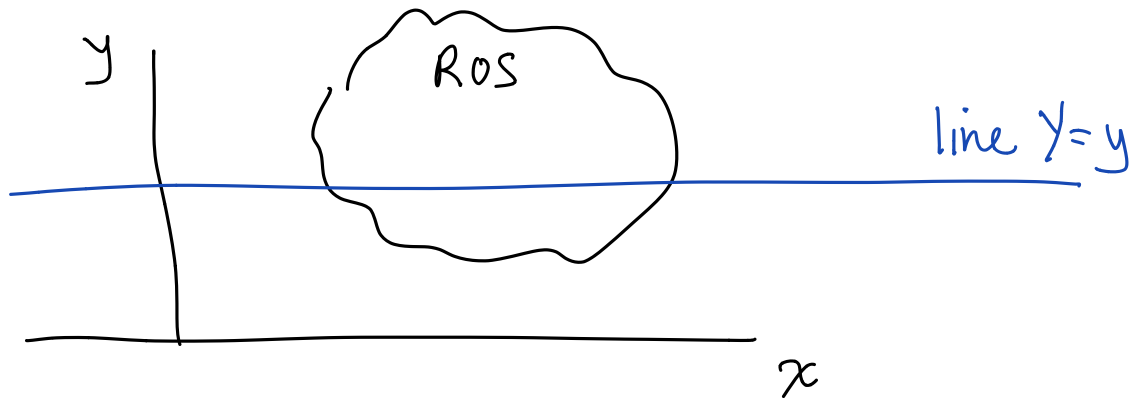
narrower and taller.





A "chop and scale" interpretation of

$$f_X(x \mid y) = f_X(x \mid Y=y)$$



Joint pdf $f_{XY}(x,y)$ is a 2D function that "sits on top of" the plane of the page.

$f_X(x \mid Y=y)$ is a 1D pdf that corresponds to the "slice" of this 2D function sliced at the value $Y=y$

But since $f_X(x \mid Y=y)$ is a PDF we need to scale it by something to ensure it integrates to one.

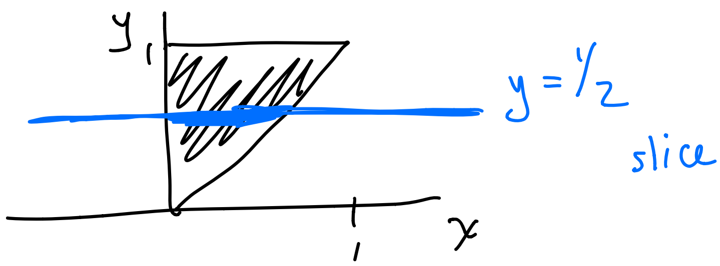
That scaling factor depends on the value of y, and just so happens to be $f_Y(y)$!

(To get marginal pdf, integrate over all y.
To get conditional pdf, choose one specific y.)

Another example of computing conditional from joint.

$$f_{XY}(x,y) = \begin{cases} 2 & 0 \le x < y \le 1 \\ 0 & \text{else} \end{cases}$$

Region of support
(flat height above plane)

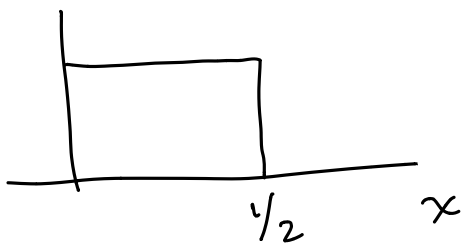


what is $f_X(x \mid Y = 1/2)$?

Intuitively, we take a slice at the line where $y = 1/2$

$f_X(x \mid Y = 1/2)$ takes the same shape as the joint pdf, and gets scaled to become a pdf across x.

$f_X(x \mid Y = 1/2)$

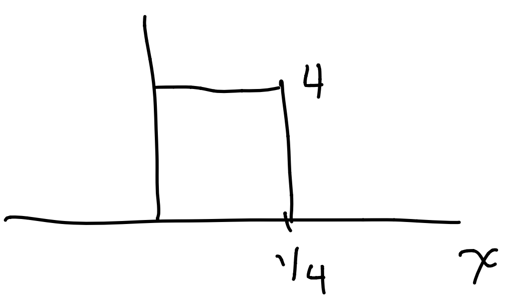


$\Rightarrow$ flat when $0 \le x \le 1/2$ and zero elsewhere

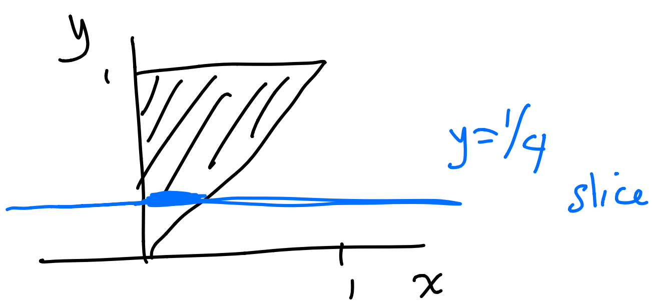
By inspection, height is 2 for $0 \le x \le 1/2$

$$f_X(x \mid Y = 1/2) = \begin{cases} 2 & 0 \le x \le 1/2 \\ 0 & \text{else} \end{cases}$$

Similarly, $f_X(x \mid Y = 1/4)$



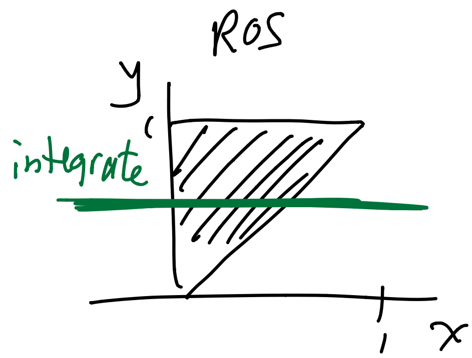
ROS



And in general, $f_X(x|y) = \begin{cases} \frac{1}{y} & 0 \le x \le y \\ 0 & \text{else} \end{cases}$

Note: Can also compute mathematically

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx$$



$$= \int_0^y 2\,dx = 2y$$

x varies from 0 to y.

Then $f_X(x|y) = \dfrac{f_{XY}(x,y)}{f_Y(y)} = \dfrac{2}{2y} = \dfrac{1}{y}$

But only when $0 \le x \le y$
and $0 < y \le 1$

note strict inequality

Interrelationships:

$$\begin{array}{lcc}
 & \text{Joint} & \text{marginal} \\
\text{CDF} & F_{XY}(x,y) \xrightarrow{\lim_{y\to\infty}} & F_X(x) \\
 & \downarrow \frac{\partial^2}{\partial x \partial y} & \downarrow \frac{d}{dx} \\
\text{PDF} & f_{XY}(x,y) \xrightarrow[\text{over } y]{\text{integrate}} & f_X(x)
\end{array}$$

$$\begin{tikzcd}
\text{Joint} & & \text{marginal} \\
f_{XY}(x,y) \arrow[rr, "\text{integrate}"] \arrow[drr] & & f_X(x) \text{ or } f_Y(y) \arrow[d] \\
 & & f_Y(y|x) \text{ or } f_X(x|y) \\
 & & \text{conditional}
\end{tikzcd}$$

need both
and divide

In general, given the two marginals, we cannot recover the joint. (need independence!)

However, given both a marginal and the corresponding complete conditional, (ie, not $f_X(x|Y=1)$ only, but for all y)

we can recover the joint:

$$\boxed{f_{XY}(x,y) = f_X(x|y)\, f_Y(y) = f_Y(y|x)\, f_X(x)}$$

Conditional probability and independence

From the definition of conditional pdf,

$$f_X(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} \quad \text{if } f_Y(y) > 0$$

So

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

what is $f_X(x|y)$ if X and Y are independent?

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

so $f_X(x|y) = f_X(x)$ provided $f_Y(y) > 0$.

As before, independence simplifies computation

Similarly $f_Y(y|x) = f_Y(y)$ if $f_X(x) > 0$ and X and Y are independent