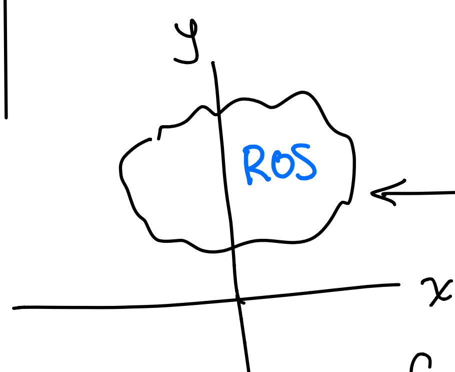
Overview of "Two Random Variables"

	one RV	two RVs	
experiment	one exper. one outcome one RV $X(\omega)$	one experiment one outcome ω two RVs $X(\omega), Y(\omega)$	ch 5.1
pmf (discrete RV)	$p_X(x) = P(X=x)$	$p_{XY}(x,y) = P(\{X=x\} \cap \{Y=y\})$ a probability	ch 5.2
cdf (any type)	$F_X(x) = P(X \le x)$	$F_{XY}(x,y) = P(\{X \le x\} \cap \{Y \le y\})$ a probability	ch 5.3
pdf	$f_X(x) = \frac{d}{dx} F_X(x)$ 	$f_{XY}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{XY}(x,y)$	ch 5.4



region of support where $f_{XY}(x,y) > 0$.

Two RVs: **Joint probability mass function**

$$p_{XY}(x,y) = P(X=x, Y=y)$$

Sample space

$$S_{XY} = \{(x,y) : p_{XY}(x,y) > 0\}$$

(the set of pairs (x,y) which have nonzero probability.)

Three ways to represent a pmf: list, table, graph

Example of a discrete pair of random variables

Test 2 ICs in sequence.
Each can be acceptable (a) or rejected (r).
Assume $P(a) = 0.9$, and
test outcomes are independent

Let X = # acceptable circuits
Y = # successful tests before 1^{st} rejection

$$S = \{rr, ra, ar, aa\}$$

$$S_{XY} = \{(x,y) : \underset{rr}{(0,0)}, \underset{ra}{(1,0)}, \underset{ar}{(1,1)}, \underset{aa}{(2,2)}$$

$S_X = \{0,1,2\}$ $S_Y = \{0,1,2\}$

mapping the outcome ω:

ω	X	Y	$P(\omega)$
rr	0	0	$(0.1)^2$
ra	1	0	$(0.1)(0.9)$
ar	1	1	$(0.9)(0.1)$
aa	2	2	$(0.9)^2$

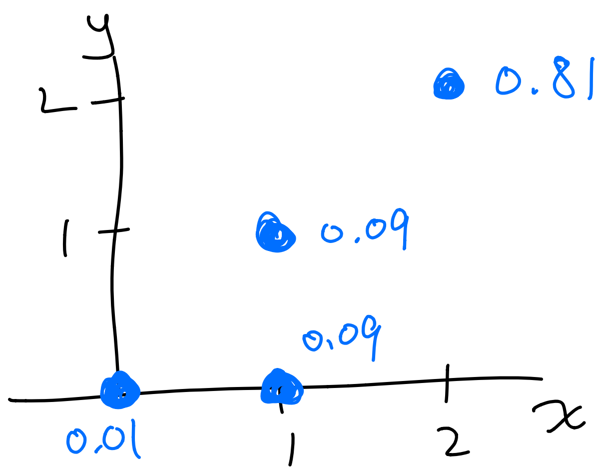
Joint PMF as a list / formula

$$p_{xy}(x,y) = \begin{cases} 0.01 & x=0,\ y=0 \\ 0.09 & x=1,\ y=0 \\ 0.09 & x=1,\ y=1 \\ 0.81 & x=2,\ y=2 \\ 0 & \text{else} \end{cases}$$

Joint PMF as a table

$x \backslash y$	0	1	2
0	0.01	—	—
1	0.09	0.09	—
2	—	—	0.81

Joint PMF as fully labeled graph



Computing probabilities from Joint PMF

All elements of S_{XY} are disjoint, so

$$P(B) = \sum_{(x,y)\in B} p_{XY}(x,y)$$

Marginal pmf

$$p_X(x) = P(X = x)$$

says nothing about value of Y

$$= P(X = x \text{ and } Y = \text{anything})$$

$$= \sum_{\substack{\text{all} \\ \text{possible} \\ y}} p_{XY}(x,y)$$

$$p_Y(y) = \sum_{\substack{\text{all} \\ \text{possible} \\ x}} p_{XY}(x,y)$$

Example of marginal pmfs

$x \backslash y$	0	1	2	$p_X(x)$
0	0.01	—	—	0.01
1	0.09	0.09	—	0.18
2	—	—	0.81	0.81
$p_Y(y)$	0.1	0.09	0.81	

summations "in the margins"

Joint CDF and Joint PDF (ch 5.3, 5.4)

$$\boxed{\begin{aligned} F_{XY}(x,y) &= P(X \le x, Y \le y) \\ f_{XY}(x,y) &= \frac{\partial^2}{\partial x \partial y} F_{XY}(x,y) \end{aligned}}$$

so $F_{XY}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{XY}(x',y')\,dy'\,dx'$

Marginal pdf

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

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

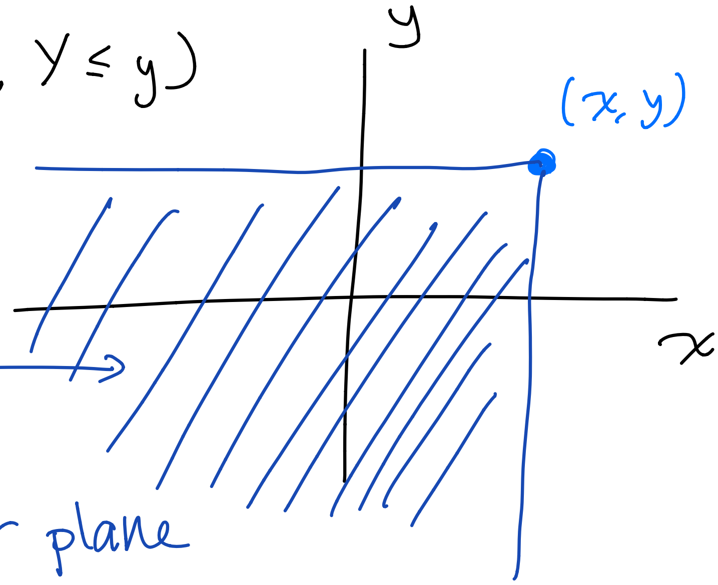
Marginal CDF

$$\begin{aligned} F_X(x) &= P(X \le x) = P(X \le x, Y \text{ anything}) \\ &= P(X \le x, Y \le \infty) \\ &= \lim_{y \to \infty} F_{XY}(x,y) \end{aligned}$$

$$F_Y(y) = \lim_{x \to \infty} F_{XY}(x,y)$$

Joint CDF and its properties

$$F_{XY}(x,y) = P(X \le x,\ Y \le y)$$



region of $\{X \le x,\ Y \le y\}$

– a quarter plane

Properties

① $\max\limits_{x,y} F_{XY}(x,y) = 1$ (when $\lim x \to \infty$ and $\lim y \to \infty$)

(the region becomes the entire sample space)

② $\min\limits_{x,y} F_{XY}(x,y) = 0$ (either x <u>or</u> y approaches $-\infty$)

(the region becomes the null set)

③ $\lim\limits_{y \to \infty} F_{XY}(x,y) = F_X(x)$

$\lim\limits_{x \to \infty} F_{XY}(x,y) = F_Y(y)$

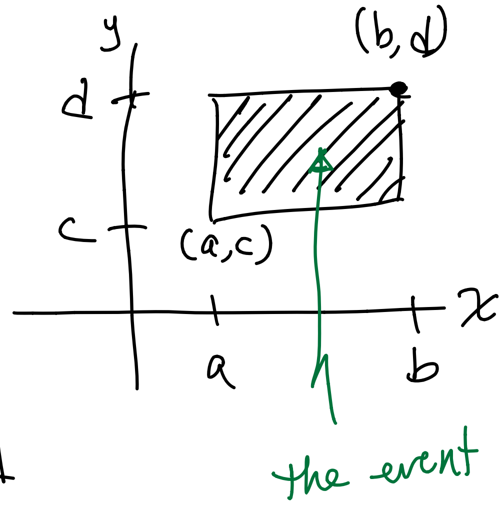
(marginal CDFs)

④ $P(a < X \le b,\ c < Y \le d)$

$$= F_{XY}(b,d) - F_{XY}(b,c) - F_{XY}(a,d) + F_{XY}(a,c)$$

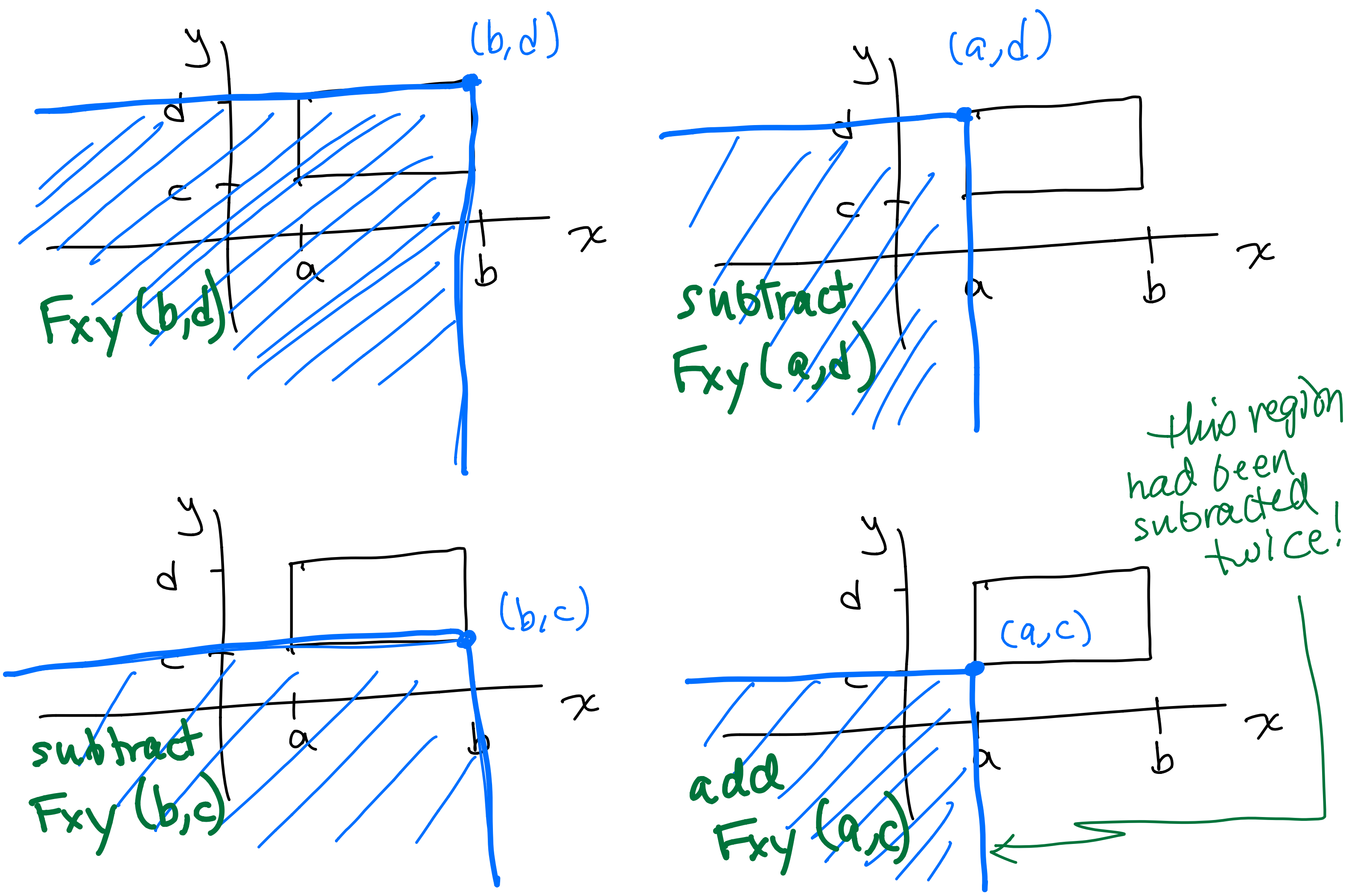
Compute probabilities of rectangular regions **directly** from the joint CDF

$$\boxed{\begin{aligned} P(a < X \le b,\ c < Y \le d) &= F_{XY}(b,d) - F_{XY}(b,c) \\ &\quad - F_{XY}(a,d) + F_{XY}(a,c) \end{aligned}}$$



(memorization tip:
plus F_{XY} evaluated at both big / both small,
minus F_{XY} evaluated at one big one small)

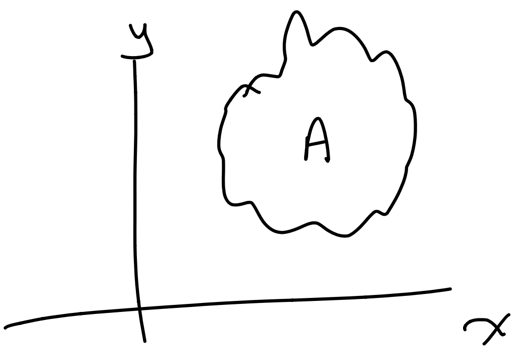
Pictorial explanation



Compute probabilities of any region by integrating the joint PDF

In general, $P((x,y) \in A) = \iint_{(x,y) \in A} f_{XY}(x,y)\,dx\,dy$

for an arbitrary event/region A



Note: the joint PDF is a "hill" rising above its Region of Support

If A is not completely inside the ROS of the PDF then you want to integrate only over $A \cap R$

For a rectangular region: $A = \{a \le X \le b,\ c \le Y \le d\}$

$$P(a \le X \le b,\ c \le Y \le d) = \int_{y=c}^{d} \int_{x=a}^{b} f_{XY}(x,y)\,dx\,dy$$

Note pay attention to setting up the integration correctly!

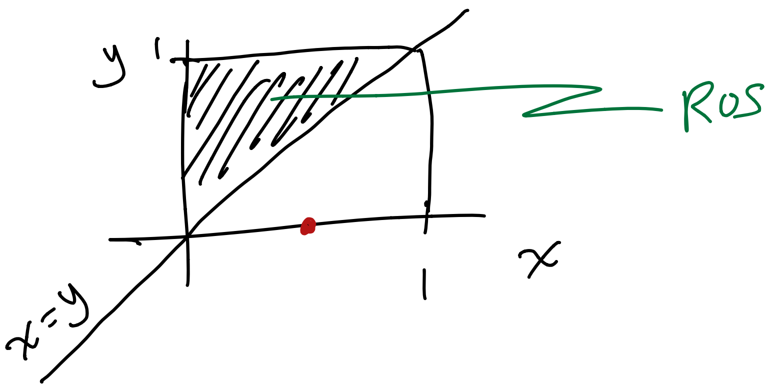
Example Computing a probability from a joint PDF

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

a) what is $P(X \le 1/2)$?

b) what is $P(X \le 1/4,\ Y \le 1/2)$?

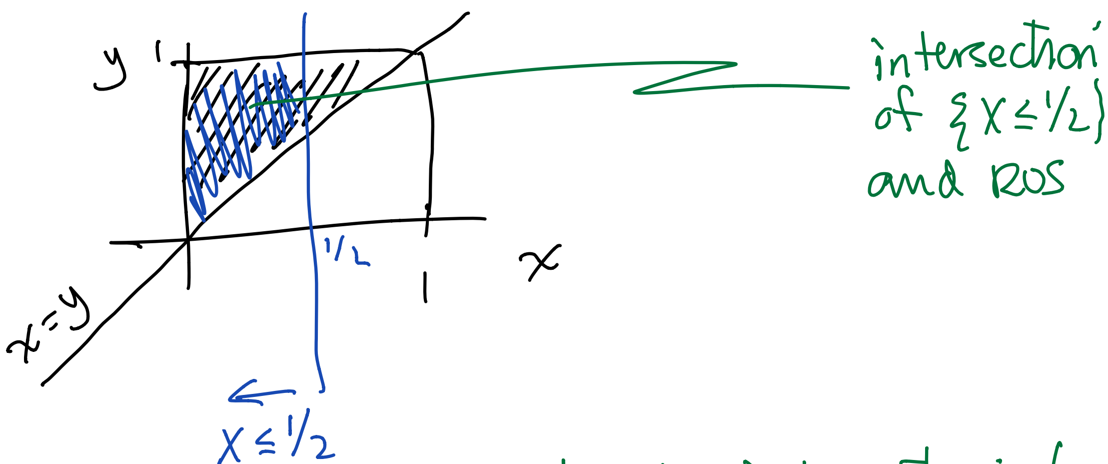
Draw the ROS! and each event



• which triangle is in ROS?
Pick a point and test.

$x = 1/2,\ y = 0$ is NOT in ROS

Now add the event of interest $\{X \le 1/2\}$

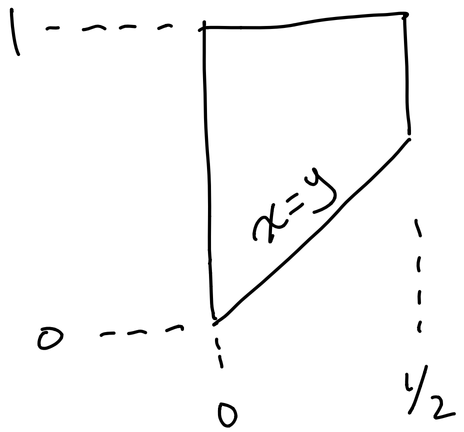


Set up and solve the double integration (see next page for reminder on how)

$$P(X \le 1/2) = \int_0^{1/2}\int_x^1 f_{XY}(x,y)\,dy\,dx = \int_0^{1/2}\int_x^1 2\,dy\,dx = \int_0^{1/2} 2(1-x)\,dx$$

$$= 3/4$$

- To set up the limits of integration, always always draw a picture and reason it out.



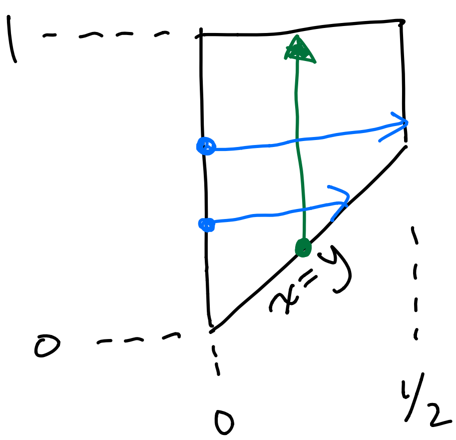
region for this integration

- The limits on the outer integral is *always* the full range for that variable.

Ex: if x is on outside, then $x \in [0, 1/2]$

Ex: if y is on outside, then $y \in [0, 1]$

- The limits on the inner integral must define the region. Draw arrows to see range.



- x on outside, for a given value of x

  $y \in [x, 1]$

- y on outside, there are 2 cases.

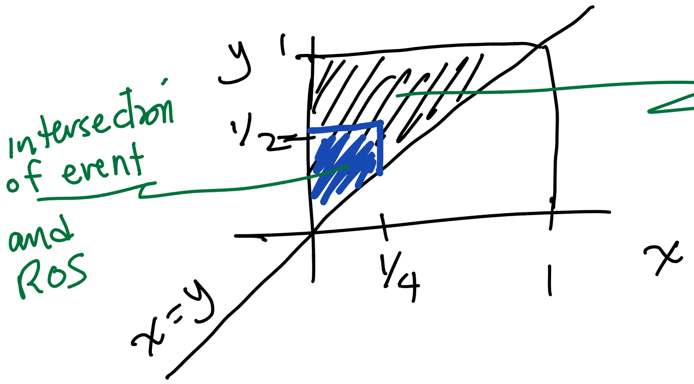
  if $0 \le y \le 1/2$ $\quad x \in [0, y]$

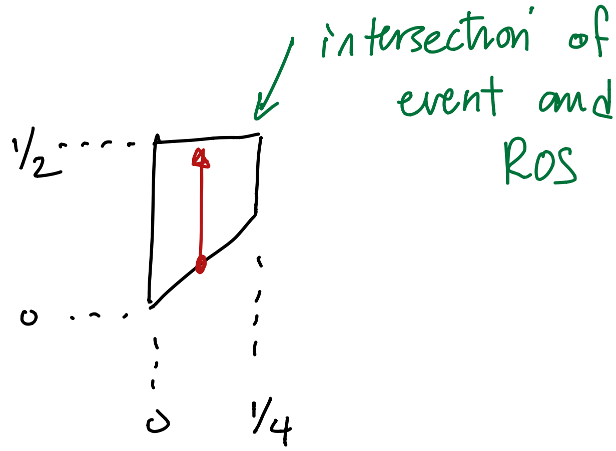
  if $1/2 \le y \le 1$ $\quad x \in [0, 1/2]$

For this example, x on outside is easier

y on outside would require 2 parts (but give the same answer)

Example part 2. What is $P(X \le 1/4 \text{ and } Y \le 1/2)$





if y inside and x outside,
draw arrow for a given x along y direction

$$P(X \le 1/4, Y \le 1/2) = \int_0^{1/4} \int_x^{1/2} 2 \, dy \, dx$$

$$= \int_0^{1/4} 2 \left(y \Big|_x^{1/2}\right) dy = \int_0^{1/4} 2\left(\tfrac{1}{2} - x\right) dx$$

$$= \left(x - \frac{2x^2}{2}\right)\Bigg|_0^{1/4} = \frac{1}{4} - \frac{1}{16} = \frac{3}{16}$$

Possible shortcuts

If the joint PDF is flat in the region of interest

$$P(A) = \frac{\text{area}(A \cap ROS)}{\text{area}(ROS)}$$

(don't do this if joint pdf is not flat over A or ROS)

Compute probabilities of rectangular regions directly from the CDF

Compute probabilities of any region by integrating the PDF

A joint CDF is useful because it expresses a probability.

However, it is most effective for computing probabilities of rectangular events only.

It's possible (but tedious) to compute a joint CDF from a joint PDF.

It's straightforward but requires a lot of attention to detail.

How many ways does the RUS intersect the event $\{X \leq x, Y \leq y\}$?

To compute the CDF you must consider them all.

Computing a joint CDF from a joint PDF

Example Joint uniform RVs.

(X, Y) is a randomly selected point in the unit square.

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

Find joint CDF.

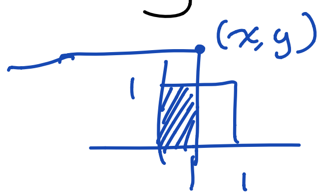
Method: divide 2-D plane into separate regions based on the intersection of the ROS with the region $\{X \le x, Y \le y\}$

Region 1: $x < 0$ <u>or</u> $y < 0$

No overlap $\Rightarrow$ $F_{XY}(x, y) = 0$

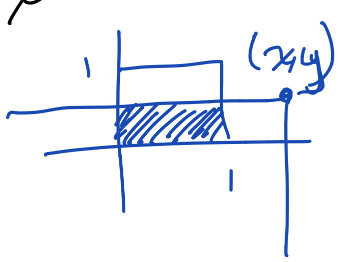
Region 2: (x, y) inside unit square

$$F_{XY}(x, y) = \int_0^x \int_0^y 1 \, dy' \, dx' = xy$$



Region 3: $0 \le x \le 1$ and $y > 1$

$$F_{XY}(x, y) = \int_0^1 \int_0^x 1 \, dx' \, dy' = x$$



Region 4: $x > 1$ and $0 \le y \le 1$

$$F_{XY}(x, y) = \int_0^1 \int_0^y 1 \, dy' \, dx' = y$$

Region 5: $x > 1$ and $y > 1$ $F_{XY}(x, y) = 1$

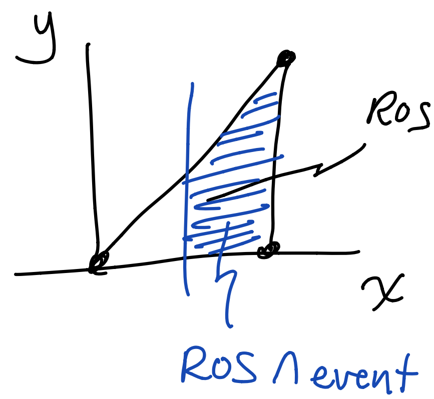
A few additional examples

Two random variables are uniformly distributed over the triangle formed by $(0,0)$, $(1,1)$, $(1,0)$.

- Find the constant c (i.e., height of PDF)
- Find $P(X > 1/2)$
- Compute the marginal PDFs

Height of PDF must be such that the joint pdf, integrated over its ROS, is 1.
The area of the triangle is $1/2$, so height is 2.

$$P(X > 1/2) = \int_{1/2}^{1} \int_{0}^{x} 2 \, dy \, dx$$



$$= 2 \int_{1/2}^{1} x \, dx = \frac{2x^2}{2} \Big|_{1/2}^{1} = \left(1 - \tfrac{1}{4}\right)$$
$$= 3/4$$

---

marginals:

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy = \int_{0}^{x} 2 \, dy = \boxed{2x \quad \text{when } 0 \le x \le 1 \quad (\text{and } 0 \text{ else})}$$

because for a given value of x, $f_{XY}(x,y)$ is non zero only for $0 \le y \le x$

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

$$= \int_{y}^{1} 2 \, dx = \boxed{2(1-y) \quad \text{when } 0 \le y \le 1 \quad \text{and } 0 \text{ else}}$$

Marginal PDF, PMF, CDF

- When you care only about the random variable and not the other.
- These are completely identical to what we studied in Topic 2.

Some properties (all are pmf /pdf)

joint pmf

$$\sum_{\text{all } x} \sum_{\text{all } y} P_{XY}(x,y) = 1$$

marginal pmfs

$$\sum_{\text{all } x} P_X(x) = 1$$

$$\sum_{\text{all } y} P_Y(y) = 1$$

joint pdf

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x,y)\, dx\, dy = 1$$

marginal pdfs

$$\int_{-\infty}^{\infty} f_X(x)\, dx = 1$$

$$\int_{-\infty}^{\infty} f_Y(y)\, dy = 1$$