Joint, Marginal, and Conditional Densities

Joint densities tell us information about the distribution of two variables in conjunction:

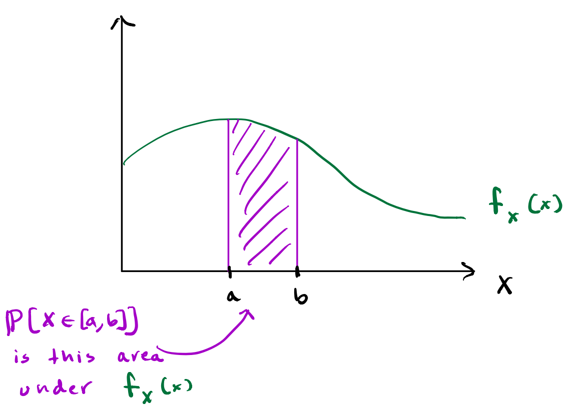
- If X, Y are continuous RVs, their **joint density** is a function $f_{X,Y} : \mathbb{R}^2 \to \mathbb{R}$ satisfying

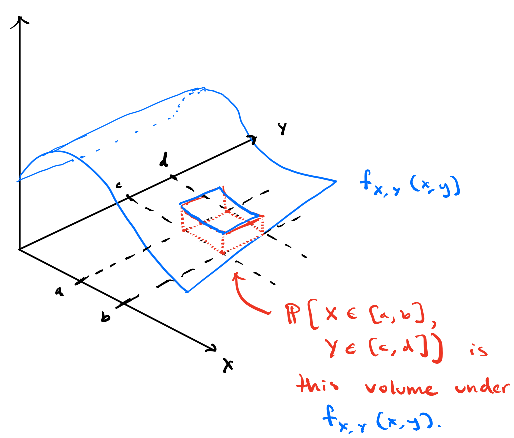
$$\mathbb{P}[X \in [a,b], Y \in [c,d]] = \int_a^b \int_c^d f_{X,Y}(x,y)\, dy\, dx$$

- Note the similarity w/ the single variable case:

$$\mathbb{P}[X \in [a,b]] = \int_a^b f_X(x)\, dx$$

- Geometric Intuition:



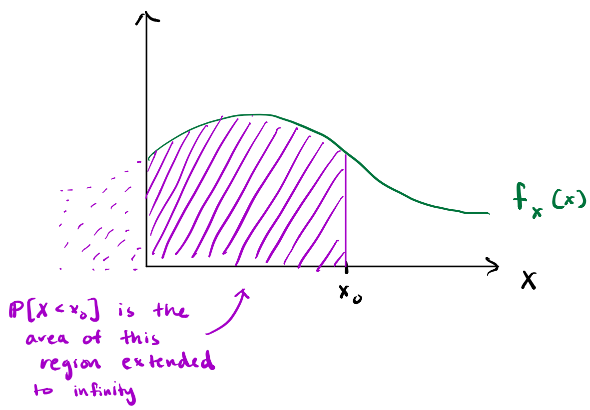


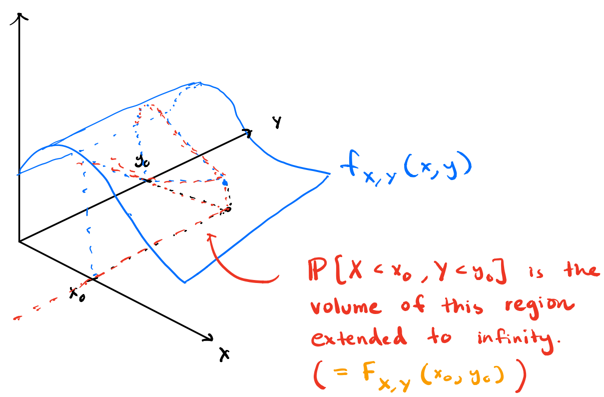
- Joint densities must also satisfy
  - $f_{X,Y}(x,y) \geq 0 \quad \forall x, y$
  - $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx\, dy = 1$

We can also consider a joint CDF, which is a function $F_{X,Y} : \mathbb{R}^2 \longrightarrow \mathbb{R}$ that satisfies

$$\mathbb{P}[X < x_0, Y < y_0] = F_{X,Y}(x_0, y_0)$$

- Again, note the similarity to the single variable case.
- Geometric intuition:





- We can say that

$$F_{X,Y}(x_0, y_0) = \int_{-\infty}^{x_0} \int_{-\infty}^{y_0} f_{X,Y}(x,y)\, dx\, dy$$

Equivalently:

$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x\, \partial y} F_{X,Y}(x,y)$$

- We have

$$\lim_{x \to \infty} \lim_{y \to \infty} F_{X,Y}(x,y) = 1, \quad \lim_{x \to -\infty} \lim_{y \to -\infty} F_{X,Y}(x,y) = 0$$

Given a joint density, we can recover a marginal density via integration

- If X, Y have joint density $f_{X,Y}(x,y)$, then we can isolate the marginal $f_X(x)$ via

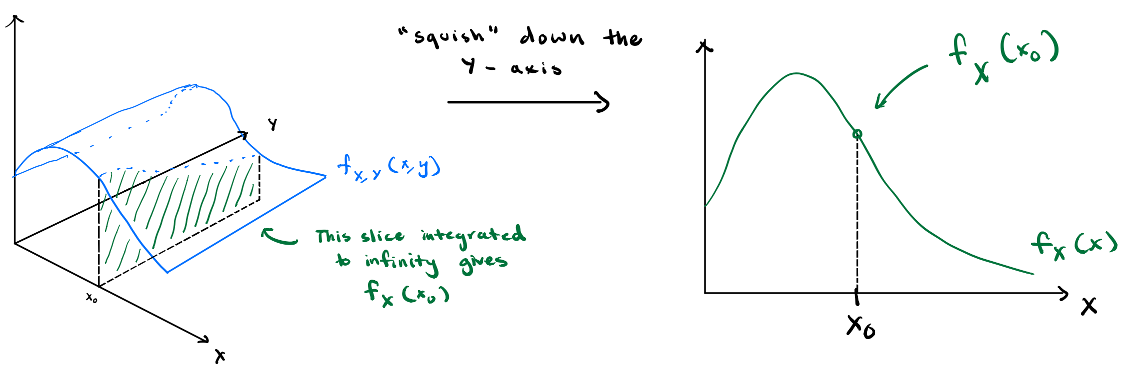
$$f(x_0) = \int_{-\infty}^{\infty} f(x_0, y)\, dy$$

same constant value x_0

- Note the similarity to the discrete case:

$$p_X(x_0) = \sum_y p_{X,Y}(x_0, y)$$

- Geometric intuition:



- In general for two variables,

$$\int \text{joint} \longrightarrow \text{marginal}$$

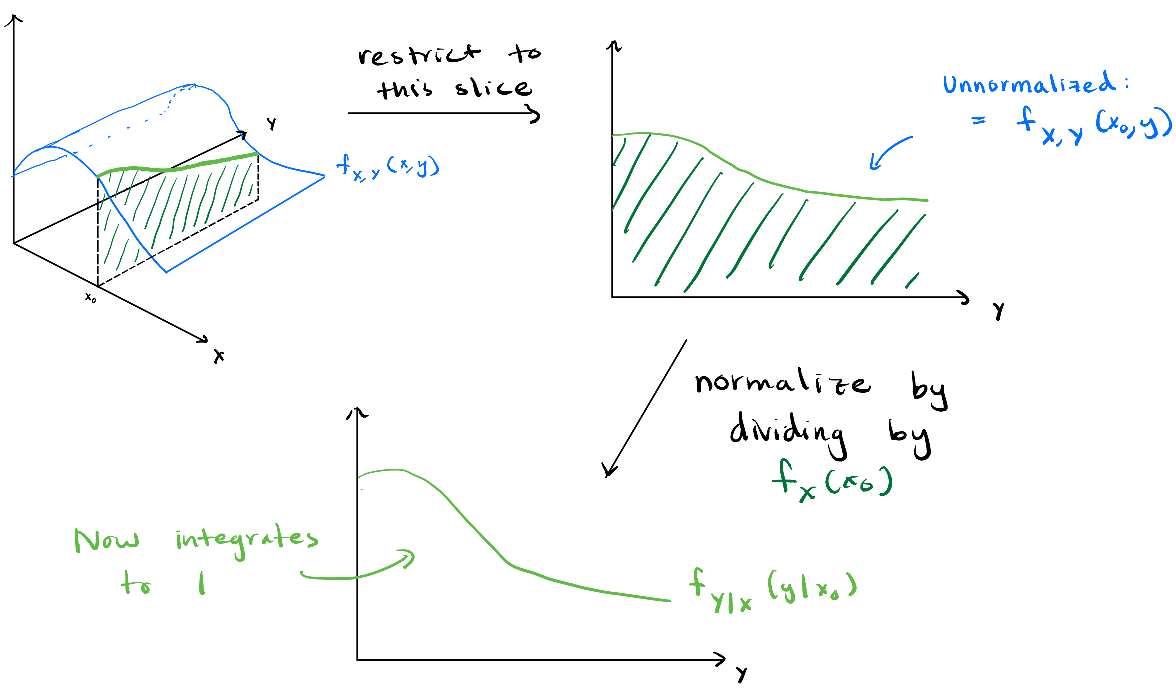
$$\iint \text{joint} \longrightarrow \text{probability}$$

We can also condition continuous RVs on each other:

- Let X, Y have joint density $f_{X,Y}(x,y)$. Then the conditional density $f_{Y|X}(y|x_0)$ of Y given $X = x_0$ is

$$f_{Y|X}(y|x_0) = \frac{f_{X,Y}(x_0, y)}{f_X(x_0)}$$

- Geometric Intuition:



- This is a special case where conditioning reduces dim.
  ↳ In general, (1) restrict domain, (2) normalize.

- The following are equivalent:
  (1) X, Y are independent
  (2) $f_{Y|X}(y|x_0) = f_Y(y) \quad \forall y, x_0$
  (3) $f_{X,Y}(x,y) = f_X(x)\, f_Y(y) \quad \forall x, y$

Reference:

Total Probability:

Discrete

D: $\mathbb{P}[A] = \sum_i \mathbb{P}[A \mid B_i]\, \mathbb{P}[B_i]$

C: $f_X(x) = \sum_i f_{X \mid B_i}(x)\, \mathbb{P}[B_i]$

Continuous

D: $\mathbb{P}[A] = \int_{-\infty}^{\infty} \mathbb{P}[A \mid X = x]\, f_X(x)\, dx$

C: $f_X(x) = \int_{-\infty}^{\infty} f_Y(y)\, f_{X \mid Y}(x \mid y)\, dy$

Bayes' Theorem

Discrete

D: $$\mathbb{P}[A_i \mid B] = \frac{\mathbb{P}[A_i]\, \mathbb{P}[B \mid A_i]}{\sum_j \mathbb{P}[A_j]\, \mathbb{P}[B \mid A_j]}$$

C: $$f_{X \mid A}(x) = \frac{f_X(x)\, \mathbb{P}[A \mid X = x]}{\int_{-\infty}^{\infty} f_X(t)\, \mathbb{P}[A \mid X = t]\, dt}$$

Continuous

D: $$\mathbb{P}[A_i \mid X = x] = \frac{\mathbb{P}[A_i]\, f_{X \mid A_i}(x)}{\sum_j \mathbb{P}[A_j]\, f_{X \mid A_j}(x)}$$

C: $$f_{X \mid Y}(x \mid y) = \frac{f_X(x)\, f_{Y \mid X}(y \mid x)}{\int_{-\infty}^{\infty} f_X(t)\, f_{Y \mid X}(y \mid t)\, dt}$$

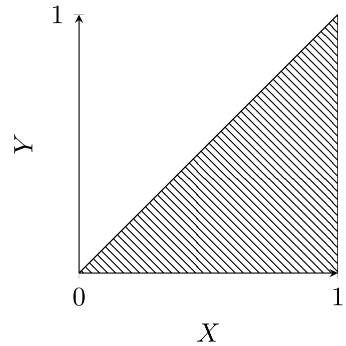
Problem 1 (Spring 2019 Final 8.2)**:** Consider continuous random variables X, Y with joint density function $f_{X,Y}(x, y) = cxy$ for all $x, y \in [0, 1]$ and 0 everywhere else.

(a) What is the value of c?

(b) What is $\mathbb{P}\left[|X - Y| \leq \frac{1}{2}\right]$?

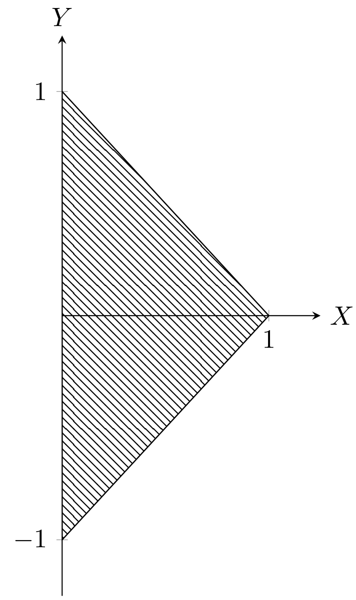
Problem 3 (Summer 2019 Final 2w)**:** Let X, Y be independent uniform random variables over the interval $[0, 1]$. What is the CDF of $|X - Y|$?

Problem 2 (Spring 2018 Final 5.14)**:** Consider continuous random variables X, Y with uniform joint density over the region $\{(x, y) \mid 0 \leq y < x \leq 1\}$ (shown below).



Suppose someone takes a sample of either X or Y with equal probability, then announces the value is $\frac{2}{3}$. What is the probability that the sample is from X?

Problem 4 (Summer 2019 Final 4 (Edited))**:** Suppose we have a triangle ABC in the plane with coordinates $A = (0.1), B = (1, 0)$, and $C = (0, -1)$, (see the figure below) and we choose a point uniformly at random from this triangle.



Let X be the random variable corresponding to the x coordinate of the point chosen, and let Y be the random variable corresponding to its y coordinate.

(a) Find the joint density $f_{X,Y}(x, y)$.

(b) Find $\mathbb{E}[X]$ and $\mathbb{E}[Y]$.

(c) Let $Z = |X| + |Y|$. Find the PDF of Z.

Problem 5 (Spring 2017 Final 7.2)**:** You pick a real number from the range $[0, 1]$ using the uniform distribution. Then Alvin independently picks a real number uniformly at random from the range $[0, 2]$.

(a) What is the probability that your numbers differ by no more than 1?

(b) Suppose you pick your number from the same range except now with PDF $f_X(x) = 2x$, with Alvin still picking uniformly from $[0, 2]$. What is the probability that your numbers differ by no more than 1?

Problem 6 (Spring 2016 Final 3.5)**:** You play a game of darts with a friend. You are better than they are and the distances of your shots from the center of the dartboard are i.i.d. $U[0,1]$ while theirs are $U[0,2]$. To make the game fair, you agree to throw one dart, while your friend throws two. The dart closest to the center wins the game. What is the probability that you win the game?

Problem 7 (Fall 2018 Final 8)**:** For $n \geq 2$, let $X_1, \ldots, X_n$ be independent $U[0,1]$ random variables and for $i \in \{1, \ldots, n\}$, let Y_i denote the ith smallest value of $\{X_1, \ldots, X_n\}$. For example, $Y_1 = \min\{X_1, \ldots, X_n\}$ and $Y_n = \max\{X_1, \ldots, X_n\}$.

(a) Find the PDF of Y_2.

(b) If $n = 2$, find the joint PDF of Y_1 and Y_2.

(c) Assume again that $n = 2$ and let $G = Y_2 - Y_1$ be the gap size between Y_1 and Y_2. Find the PDF of G.

(d) What is $\mathbb{P}\left[G > \frac{1}{2}\right]$?