Chapter 8
Continuous Random Variables:
Joint distributions

Joint Densities

<u>Defn</u>: The **joint probability density function** between continuous random variables X and Y is a non-negative function $f_{X,Y}(x,y)$, where

$$\int_c^d \int_a^b f_{X,Y}(x,y)dxdy = P\{a \leq X \leq b \ \& \ c \leq Y \leq d\}$$

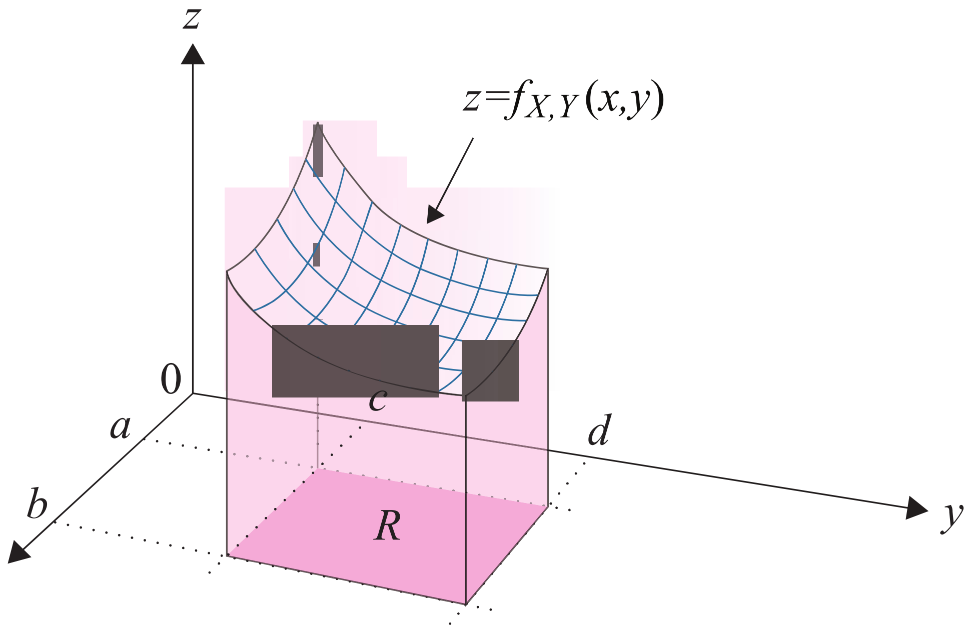
and where

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy = 1$$

Joint Densities

Volume under the curve equals:

$$\int_c^d \int_a^b f_{X,Y}(x,y)dxdy = P\{a \le X \le b \ \ \& \ \ c \le Y \le d\}$$



Example

Two-year-olds range in weight from 15 – 35 pounds.
They range in height from 25 – 40 inches.

$f_{W,H}(w,h)$ denotes the joint p.d.f. of weight and height.

Q: What is the fraction of two-year-olds with weight > 30 pounds but height < 30 inches?

A:

$$\int_{h=-\infty}^{h=30} \int_{w=30}^{w=\infty} f_{W,H}(w,h)dwdh = \int_{25}^{30} \int_{30}^{35} f_{W,H}(w,h)dwdh$$

Marginal densities

Defn: The **marginal densities** $f_X(x)$ and $f_Y(y)$ are defined as:

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

Note that $f_X(x)$ and $f_Y(y)$ are densities and not probabilities.

Q: If $f_{W,H}(w,h)$ is the joint p.d.f. of weight and height in two-year-olds, what is the fraction of two-year-olds whose height is exactly 30 inches?

A: $$\int_{w=-\infty}^{w=\infty} f_{W,H}(w,30)dw = f_H(30)$$

This is a zero-probability event!

Independence

<u>Defn</u>: Continuous random variables X and Y are **independent**, written $X \perp Y$, if:

$$f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y) \quad \forall x, y$$

Let's consider some joint p.d.f.s to determine whether X and Y are independent.

Example

$$f_{X,Y}(x,y) = \begin{cases} x+y & \text{if } 0 \le x,y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Q: (a) What is $\boldsymbol{E}[X]$? (b) Is $X \perp Y$?

A: part (a)

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = \int_0^1 (x+y)dy = x + \frac{1}{2}$$

$$\boldsymbol{E}[X] = \int_{-\infty}^{\infty} f_X(x) \cdot x dx = \int_0^1 \left(x + \frac{1}{2}\right) \cdot x\, dx = \frac{7}{12}$$

Example

$$f_{X,Y}(x,y) = \begin{cases} x + y & \text{if } 0 \leq x, y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Q: (a) What is $\boldsymbol{E}[X]$? (b) Is $X \perp Y$?

A: part (b)

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = x + \frac{1}{2}$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx = y + \frac{1}{2}$$

Clearly, $f_{X,Y}(x,y) \neq f_X(x) \cdot f_Y(y)$

Example

$$f_{X,Y}(x,y) = \begin{cases} 4xy & \text{if } 0 \le x, y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Q: Is $X \perp Y$?

A:

$$f_X(x) = \int_0^1 4xy\,dy = 2x$$

$$f_Y(y) = \int_0^1 4xy\,dx = 2y$$

Clearly, $f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$

Example: Which Exponential happens first?

- The time until server 1 crashes is $X \sim Exp(\lambda)$
- The time until server 2 crashes is $Y \sim Exp(\mu)$

$X \sim Exp(\lambda)$

$Y \sim Exp(\mu)$

Q: What is the probability that server 1 crashes before server 2? Assume $X \perp Y$.

A:

$$\boldsymbol{P}\{X < Y\} = \int_{x=0}^{\infty} \int_{y=x}^{\infty} f_{X,Y}(x,y)dydx$$

$$= \int_{x=0}^{\infty} \int_{y=x}^{\infty} f_X(x) \cdot f_Y(y)\, dydx$$

$$= \int_{x=0}^{\infty} \lambda e^{-\lambda x} \int_{y=x}^{\infty} \mu e^{-\mu y} dydx \quad = \frac{\lambda}{\lambda + \mu}$$

What happens when $\lambda = \mu$?

Conditional p.d.f. and Bayes' Law

<u>Defn</u>: Given two continuous random variables, X and Y, we define the **conditional p.d.f.** of r.v. X given event $Y = y$ as:

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{f_{Y|X=x}(y) \cdot f_X(x)}{f_Y(y)} = \frac{f_{Y|X=x}(y) \cdot f_X(x)}{\int_x f_{X,Y}(x,y)\, dx}$$

This is the definition of the conditional p.d.f., where we're conditioning on a zero-probability event $Y = y$

Here we've used the same definition but this time applied it to conditioning on $X = x$, resulting in a Bayes' Law for two continuous r.v.s

Here we've simply expanded out $f_Y(y)$

Observe that the conditional p.d.f. is still a proper p.d.f., i.e., $\int_{-\infty}^{\infty} f_{X|Y=y}(x)\, dx = 1$

Law of Total Probability Generalized

Recall the Law of Total Probability, repeated below:

Theorem: Let A be an event and Y be a continuous r.v.
Then we can compute $\boldsymbol{P}\{A\}$ by conditioning on the value of Y as follows:

$$\boldsymbol{P}\{A\} = \int_{-\infty}^{\infty} f_Y(y \cap A)dy = \int_{-\infty}^{\infty} \boldsymbol{P}\{A \mid Y = y\}\, f_Y(y)dy$$

Using the definition for the conditional p.d.f. from the prior slide, we can similarly express $f_X(x)$ by conditioning on the value of Y:

Theorem: Let X and Y be continuous random variables. Then, from the definition of the conditional p.d.f., we have:

$$f_X(x) = \int_y f_{X,Y}(x,y)dy = \int_y f_{X|Y=y}(x) f_Y(y)dy$$

Example: Which Exponential happens first?

- The time until server 1 crashes is $X \sim Exp(\lambda)$
- The time until server 2 crashes is $Y \sim Exp(\mu)$

$X \sim Exp(\lambda)$

$Y \sim Exp(\mu)$

Q: What is the probability that server 1 crashes before server 2? Assume $X \perp Y$.

A:

$$\boldsymbol{P}\{X < Y\} = \int_0^\infty \boldsymbol{P}\{X < Y \mid X = x\} \cdot f_X(x)dx$$

$$= \int_0^\infty \boldsymbol{P}\{Y > x \mid X = x\} \cdot \lambda e^{-\lambda x} dx$$

$$= \int_0^\infty \boldsymbol{P}\{Y > x\} \cdot \lambda e^{-\lambda x} dx$$

$$= \int_0^\infty e^{-\mu x} \cdot \lambda e^{-\lambda x} dx = \frac{\lambda}{\lambda + \mu}$$

Where did we use independence?

From Midterm 2020

Random variables X and Y are NOT independent. Their joint density is:

$$f_{X,Y}(x, y) \text{ where } 0 \leq x, y \leq \infty$$

All your answers should be in terms of $f_{X,Y}(x, y)$ or prior results.

a. Write an expression for $f_Y(y)$

b. Write an expression for $f_{X|Y=5}(x)$

c. Write an expression for $\boldsymbol{P}\{X + Y < 10 \mid Y = 5\}$

d. Write an expression for $f_{Y|X<6}(y)$

e. Write an expression for $f_{Y|Y<6}(y)$

From Midterm 2020

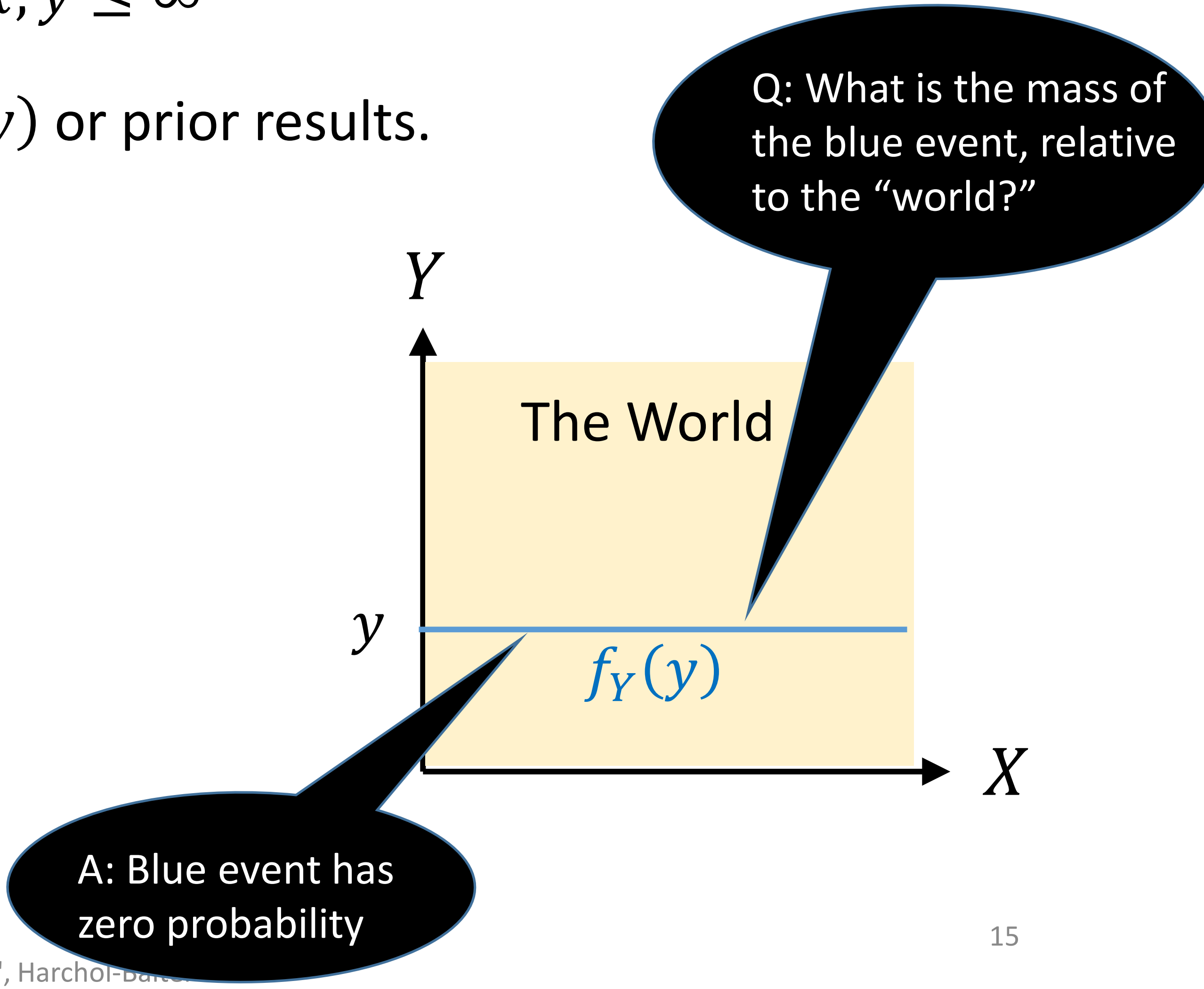
Random variables X and Y are NOT independent. Their joint density is:

$$f_{X,Y}(x,y) \text{ where } 0 \leq x, y \leq \infty$$

All your answers should be in terms of $f_{X,Y}(x,y)$ or prior results.

a. Write an expression for $f_Y(y)$

$$f_Y(y) = \int_0^\infty f_{X,Y}(x,y)dx$$



From Midterm 2020

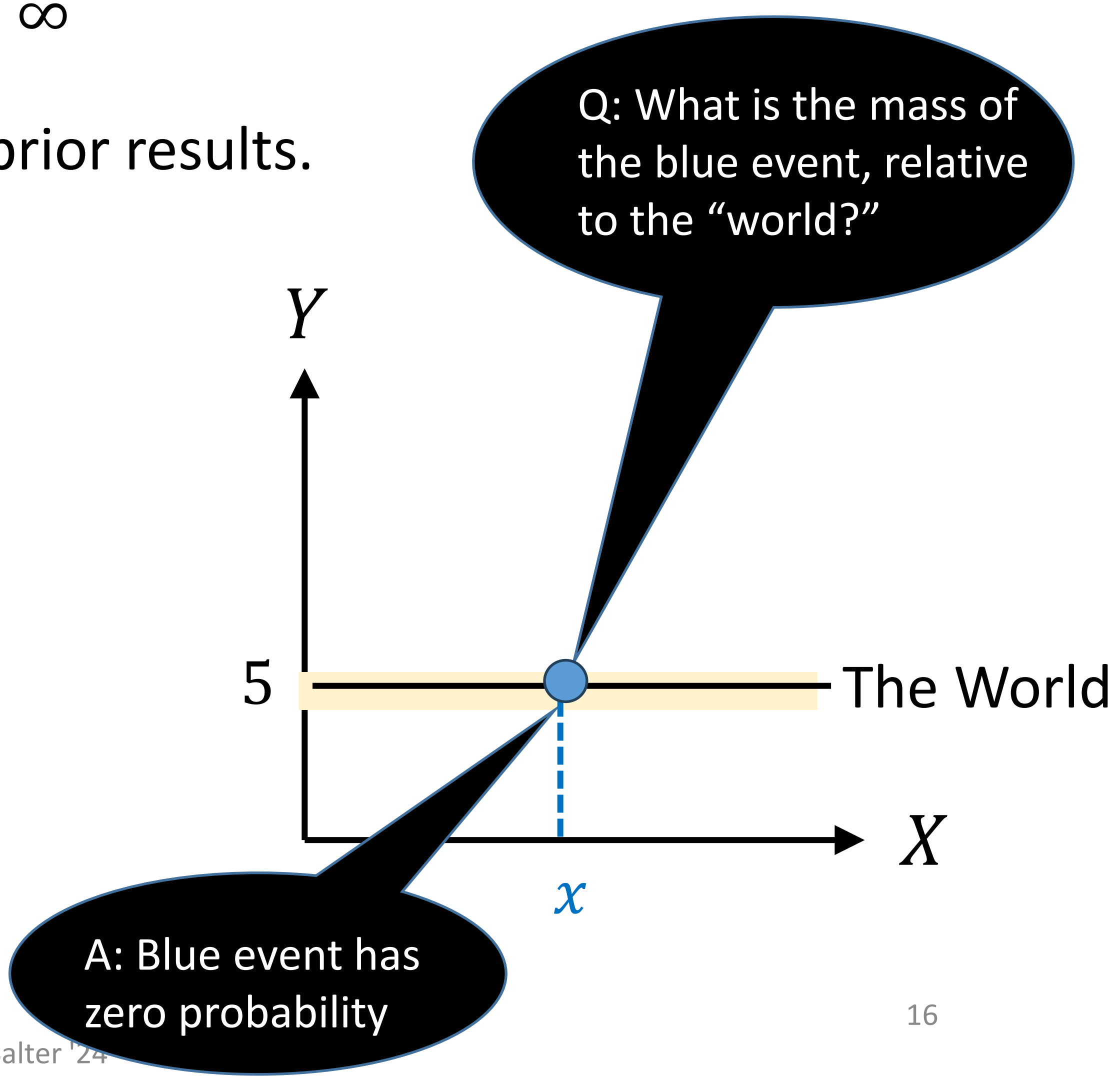
Random variables X and Y are NOT independent. Their joint density is:

$$f_{X,Y}(x,y) \text{ where } 0 \leq x, y \leq \infty$$

All your answers should be in terms of $f_{X,Y}(x,y)$ or prior results.

b. Write an expression for $f_{X|Y=5}(x)$

$$f_{X|Y=5}(x) = \frac{f_{X,Y}(x,5)}{f_Y(5)},$$



From Midterm 2020

Random variables X and Y are NOT independent. Their joint density is:

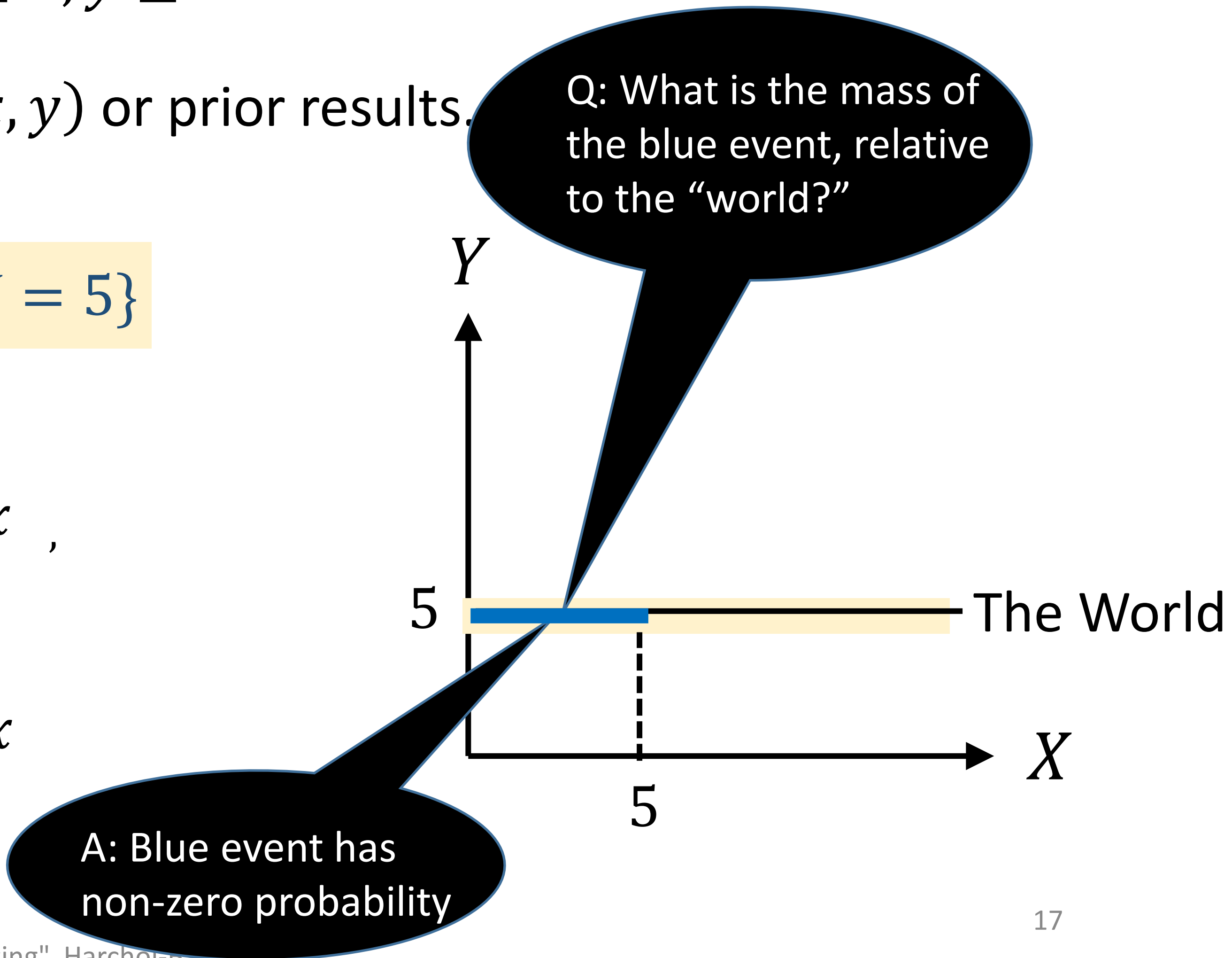
$$f_{X,Y}(x,y) \text{ where } 0 \leq x, y \leq \infty$$

All your answers should be in terms of $f_{X,Y}(x,y)$ or prior results.

c. Write an expression for $\boldsymbol{P}\{X+Y<10 \mid Y=5\}$

$$\boldsymbol{P}\{X+Y<10 \mid Y=5\} = \int_{x=0}^{5} f_{X|Y=5}(x)dx$$

$$= \int_{x=0}^{5} \frac{f_{X,Y}(x,5)}{f_Y(5)} dx$$



From Midterm 2020

Random variables X and Y are NOT independent. Their joint density is:

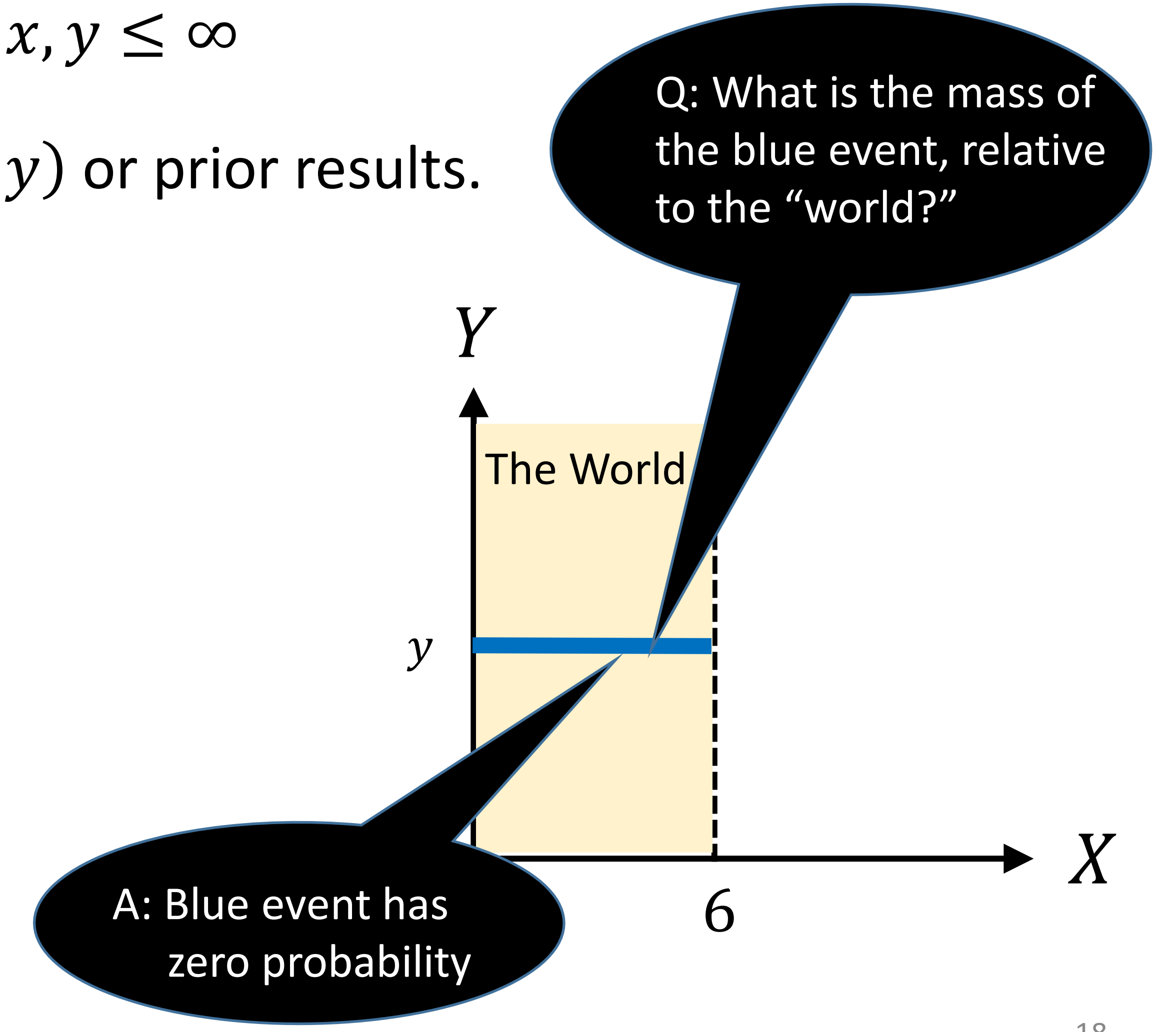
$$f_{X,Y}(x, y) \text{ where } 0 \leq x, y \leq \infty$$

All your answers should be in terms of $f_{X,Y}(x, y)$ or prior results.

d. Write an expression for $f_{Y|X<6}(y)$

$$f_{Y|X<6}(y) = \frac{f_Y(y \cap X < 6)}{\boldsymbol{P}\{X < 6\}},$$

$$= \frac{\int_{x=0}^{6} f_{X,Y}(x, y)dx}{\int_{x=0}^{6} f_X(x)dx}$$



From Midterm 2020

Random variables X and Y are NOT independent. Their joint density is:

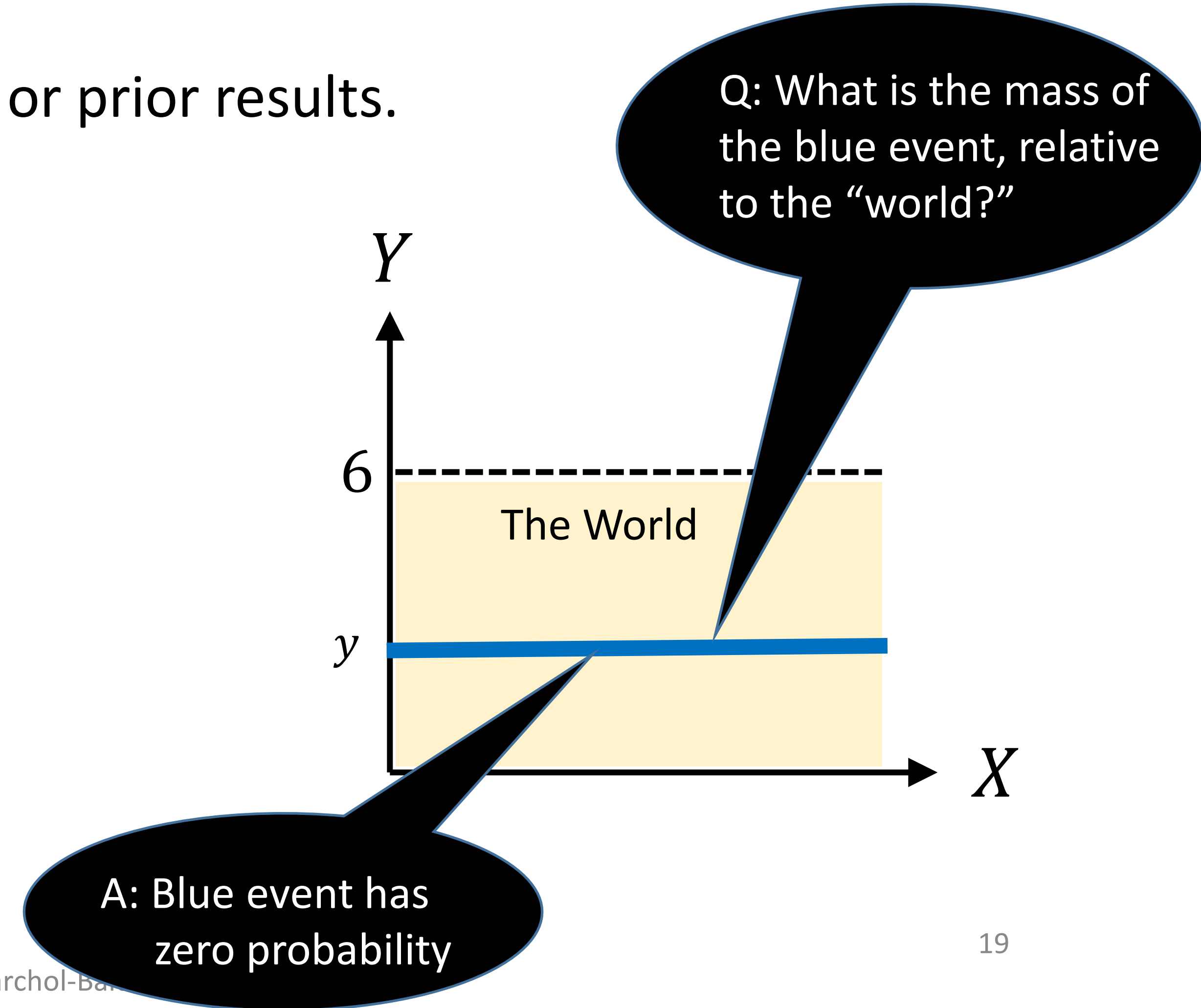
$$f_{X,Y}(x,y) \text{ where } 0 \le x, y \le \infty$$

All your answers should be in terms of $f_{X,Y}(x,y)$ or prior results.

e. Write an expression for $f_{Y|Y<6}(y)$

$$f_{Y|X<6}(y) = \frac{f_Y(y \cap Y < 6)}{\boldsymbol{P}\{Y < 6\}}$$

$$= \begin{cases} \dfrac{f_Y(y)}{\boldsymbol{P}\{Y < 6\}} & \text{if } y < 6 \\ 0 & \text{otherwise} \end{cases}$$



Expectation with multiple r.v.s

Defn: Let X and Y be continuous random variables with joint p.d.f. $f_{X,Y}(x, y)$.
Then, for any function $g(x, y)$, we have

$$\boldsymbol{E}[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) \cdot f_{X,Y}(x, y) dx dy$$

Conditional expectation with multiple RVs

Recall Defn: For a **continuous** r.v. X and an event A, where $\boldsymbol{P}\{A\} > 0$,the **conditional expectation of $\boldsymbol{X}$ given $\boldsymbol{A}$** is:

$$\boldsymbol{E}[X|A] = \int_x x \cdot f_{X|A}(x)dx$$

Defn: For **continuous** r.v.s X and Y

$$\boldsymbol{E}[X|Y = y] = \int_x x \cdot f_{X|Y=y}(x)dx = \int_x x \cdot \frac{f_{X,Y}(x,y)}{f_Y(y)}dx$$

Theorem: We can derive $\boldsymbol{E}[X]$ by conditioning on the value of continuous r.v. Y :

$$\boldsymbol{E}[X] = \int_y \boldsymbol{E}[X\ |Y = y] \cdot f_Y(y)dy$$

Example

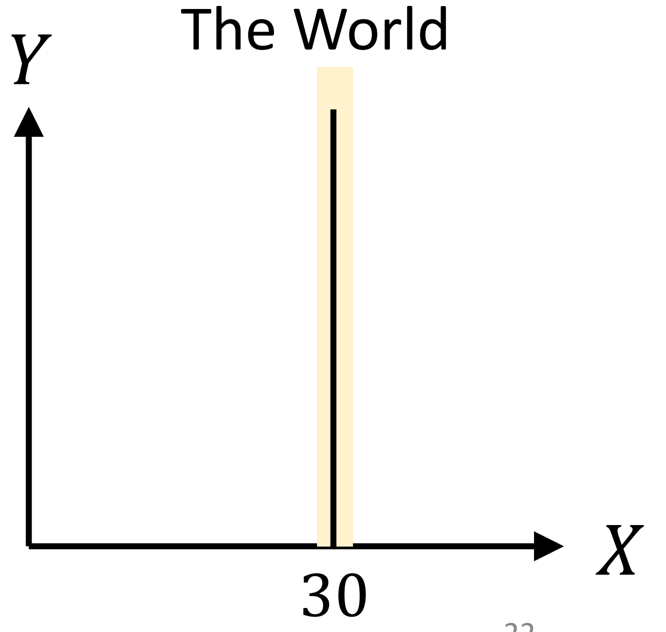
Two-year-olds range in weight from 15 – 35 pounds.
They range in height from 25 – 40 inches.

Q: My 2-year old is 30 inches tall.
What is their expected weight?

A:

$$\boldsymbol{E}[W \mid H = 30] = \int_{w=15}^{35} w \cdot f_{W|H=30}(w)dw$$

$$= \int_{w=15}^{35} w \cdot \frac{f_{W,H}(w,30)}{f_H(30)} dw$$



Example

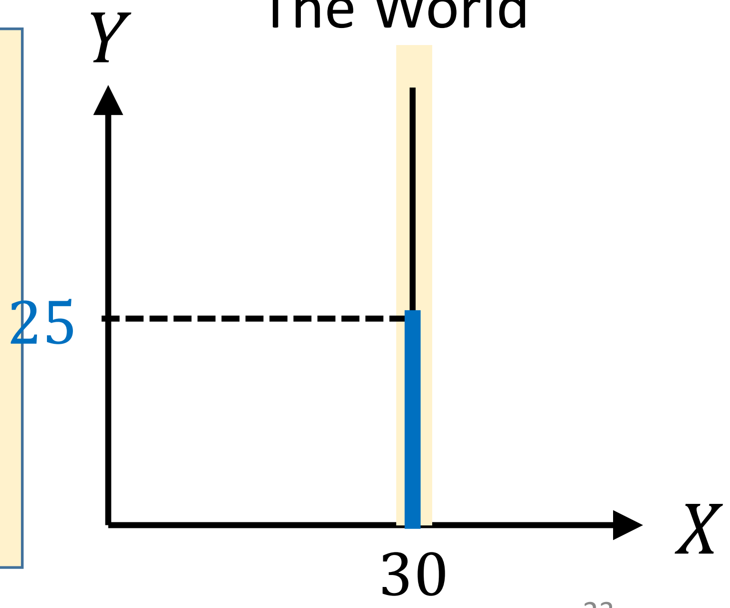
Two-year-olds range in weight from 15 – 35 pounds.
They range in height from 25 – 40 inches.

Q: What fraction of 2-year olds with height 30 inches have weight $<$ 25 pounds?

A:

$$\boldsymbol{P}\{W < 25 \,|H = 30\} = \int_{w=15}^{25} f_{W|H=30}(w)dw$$

$$= \int_{w=15}^{25} \frac{f_{W,H}(w,30)}{f_H(30)} dw$$



Example: Hand-in Time versus Grade

- T = number of days early that homework is submitted: $0 \leq T \leq 2$
- G = grade on homework (as a percentage): $0 \leq G \leq 1$
- Joint density function: $0 \leq t \leq 2,\ 0 \leq g \leq 1$:

$$f_{G,T}(g,t) = \frac{9}{10}\mathrm{tg}^2 + \frac{1}{5}$$

Q: What is the probability that a random student gets a grade above 50%?

A:

$$f_G(g) = \int_0^2 f_{G,T}(g,t)dt = \int_0^2 \left(\frac{9}{10}\mathrm{tg}^2 + \frac{1}{5}\right)dt = \frac{9}{5}g^2 + \frac{2}{5}$$

$$\boldsymbol{P}\left\{G > \frac{1}{2}\right\} = \int_{0.5}^1 f_G(g)dg = \int_{0.5}^1 \left(\frac{9}{5}\mathrm{g}^2 + \frac{2}{5}\right)dt = 0.725$$

Example: Hand-in Time versus Grade

- T = number of days early that homework is submitted: $0 \leq T \leq 2$
- G = grade on homework (as a percentage): $0 \leq G \leq 1$

Q: Given that a student submitted less than a day before the deadline, does the probability of getting a grade >50% go down?

A: $$\boldsymbol{P}\{G > 0.5 \mid T < 1\} = \frac{\boldsymbol{P}\{G > 0.5 \;\&\; T < 1\}}{\boldsymbol{P}\{T < 1\}} = \frac{\int_{g=0.5}^{g=1} \int_{t=0}^{t=1} f_{G,T}(g,t)dt\,dg}{\int_{t=0}^{t=1} f_T(t)dt}$$

$$= \frac{\int_{g=0.5}^{g=1} \int_{t=0}^{t=1} \left(\frac{9}{10}\text{tg}^2 + \frac{1}{5}\right) dt\,dg}{\int_{t=0}^{t=1} \left(\frac{3}{10}t + \frac{1}{5}\right) dt} = 0.66$$

$$f_{G,T}(g,t) = \frac{9}{10}\text{tg}^2 + \frac{1}{5}$$

$$f_T(t) = \int_0^1 f_{G,T}(g,t)dg = \int_0^1 \left(\frac{9}{10}\text{tg}^2 + \frac{1}{5}\right) dg = \frac{3}{10}t + \frac{1}{5}$$

Example: Hand-in Time versus Grade

- T = number of days early that homework is submitted: $0 \leq T \leq 2$
- G = grade on homework (as a percentage): $0 \leq G \leq 1$

Q: A student submits at $T = 0$, i.e., exactly when the homework is due. What is their expected grade?

A:

$$\boldsymbol{E}[G \mid T = 0] = \int_{g=0}^{1} g \cdot f_{G|T=0}(g)dg = \int_{g=0}^{1} g \cdot \frac{f_{G,T}(g,0)}{f_T(0)} dg$$

$$= \int_{g=0}^{1} g \cdot \frac{\frac{1}{5}}{\frac{1}{5}} dg = 0.5$$

$$f_{G,T}(g,t) = \frac{9}{10}\mathrm{tg}^2 + \frac{1}{5}$$

$$f_T(t) = \int_0^1 f_{G,T}(g,t)dg = \int_0^1 \left(\frac{9}{10}\mathrm{tg}^2 + \frac{1}{5}\right) dg = \frac{3}{10}t + \frac{1}{5}$$

Example: Hand-in Time versus Grade

- T = number of days early that homework is submitted: $0 \leq T \leq 2$
- G = grade on homework (as a percentage): $0 \leq G \leq 1$

Q: By contrast, what is the expected grade of a student who submits > 1 day early?

A: $$\boldsymbol{E}[G \mid 1 < T < 2] = \int_{g=0}^{1} g \cdot f_{G|1<T<2}(g) dg = \int_{g=0}^{1} g \cdot \frac{f_G(g \ \cap \ 1 < T < 2)}{\boldsymbol{P}\{1 < T < 2\}} dg$$

$$f_{G,T}(g,t) = \frac{9}{10} \text{tg}^2 + \frac{1}{5}$$

$$= \int_{g=0}^{1} g \cdot \frac{\int_1^2 f_{G,T}(g,t)dt}{\int_1^2 f_T(t)dt} dg$$

$$f_T(t) = \int_0^1 f_{G,T}(g,t)dg = \frac{3}{10} t + \frac{1}{5}$$

$$= \int_{g=0}^{1} g \cdot \frac{\int_1^2 \left(\frac{9}{10} \text{tg}^2 + \frac{1}{5}\right) dt}{\int_1^2 \left(\frac{3}{10} t + \frac{1}{5}\right) dt} dg = 0.673$$