INTRODUCTION TO STATISTICS

LECTURE 4

LAST TIME

- Continuous random variables
  - $P(X \leq x)$, $P(X > x)$, $P(X = x)$
  - CDFs
- Uniform distribution

TODAY

- Probability density function (PDF)
- PDF and CDF
- Properties of continuous random variables
  - Mean
  - Variance
- + 1 more continuous distribution

SHORT QUIZ

CDFs

SHORT QUIZ

CDF $F(x) = P(X \leq x)$

Probability mass function

1. F(5)

SHORT QUIZ

CDF $F(x) = P(X \leq x)$

1. F(5)

Probability mass function

1. $P(X \leq 5)$

SHORT QUIZ

CDF $F(x) = P(X \leq x)$

1. F(5)

Probability mass function

1. $P(X \leq 5)$
2. $P(X > 3)$

SHORT QUIZ

CDF $F(x) = P(X \leq x)$

1. F(5)

2. 1 – F(3)

Probability mass function

1. $P(X \leq 5)$

2. $P(X > 3)$

SHORT QUIZ

CDF $F(x) = P(X \leq x)$

1. F(5)

2. 1 – F(3)

3. F(2) – F(-1)

Probability mass function

1. $P(X \leq 5)$

2. $P(X > 3)$

SHORT QUIZ

CDF $F(x) = P(X \leq x)$

1. F(5)
2. 1 – F(3)
3. F(2) – F(-1)

Probability mass function

1. $P(X \leq 5)$
2. $P(X > 3)$
3. $P(-1 < X \leq 2)$

SHORT QUIZ

CDF $F(x) = P(X \leq x)$

1. F(5)
2. 1 – F(3)
3. F(2) – F(-1)

Probability mass function

1. $P(X \leq 5)$
2. $P(X > 3)$
3. $P(-1 < X \leq 2)$
4. $P(-5 < X \leq -3)$

SHORT QUIZ

CDF $F(x) = P(X \leq x)$

1. $F(5)$
2. $1 – F(3)$
3. $F(2) – F(-1)$
4. $F(-3) – F(-5)$

Probability mass function

1. $P(X \leq 5)$
2. $P(X > 3)$
3. $P(-1 < X \leq 2)$
4. $P(-5 < X \leq -3)$

SHORT QUIZ

CDF $F(x) = P(X \leq x)$

1. F(5)
2. 1 – F(3)
3. F(2) – F(-1)
4. F(-3) – F(-5)
5. F(5) – F(3) + F(-1) – F(-5)

Probability mass function

1. $P(X \leq 5)$
2. $P(X > 3)$
3. $P(-1 < X \leq 2)$
4. $P(-5 < X \leq -3)$

SHORT QUIZ

CDF $F(x) = P(X \leq x)$

1. $F(5)$
2. $1 – F(3)$
3. $F(2) – F(-1)$
4. $F(-3) – F(-5)$
5. $F(5) – F(3) + F(-1) – F(-5)$

Probability mass function

1. $P(X \leq 5)$
2. $P(X > 3)$
3. $P(-1 < X \leq 2)$
4. $P(-5 < X \leq -3)$
5. $P(3 < X \leq 5 \text{ or } -5 < X \leq -1)$

PROBABILITY DENSITY FUNCTION

PDFs

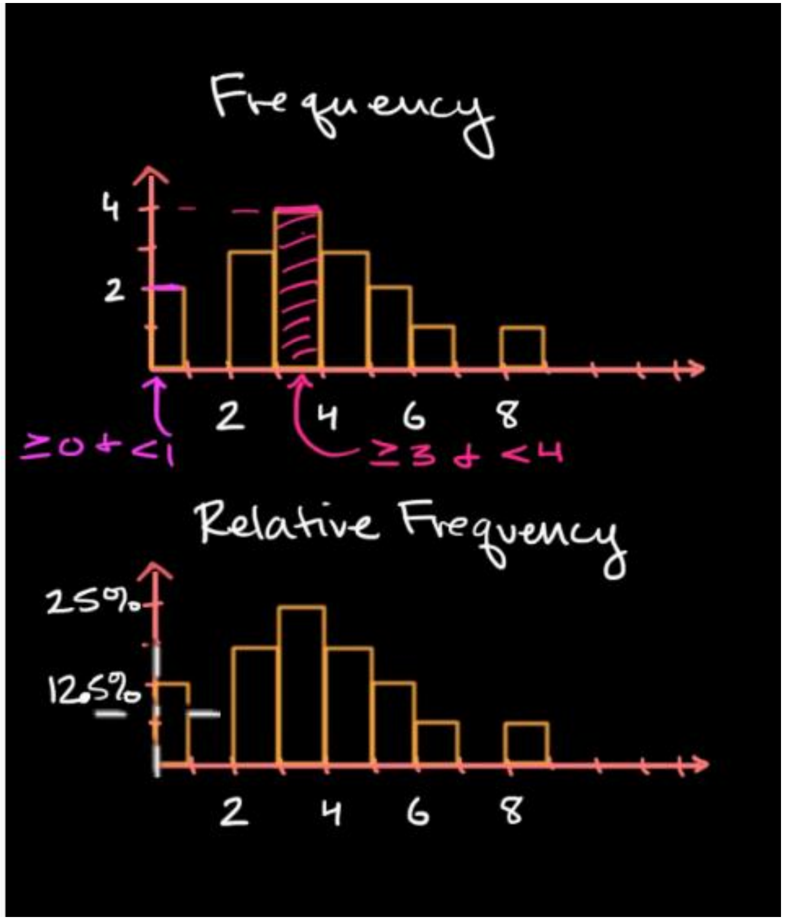
WATCH THE VIDEO

https://youtu.be/PUvUQMQ7xQk

NOW, LET'S DISCUSS IT STEP BY STEP

HISTOGRAMS

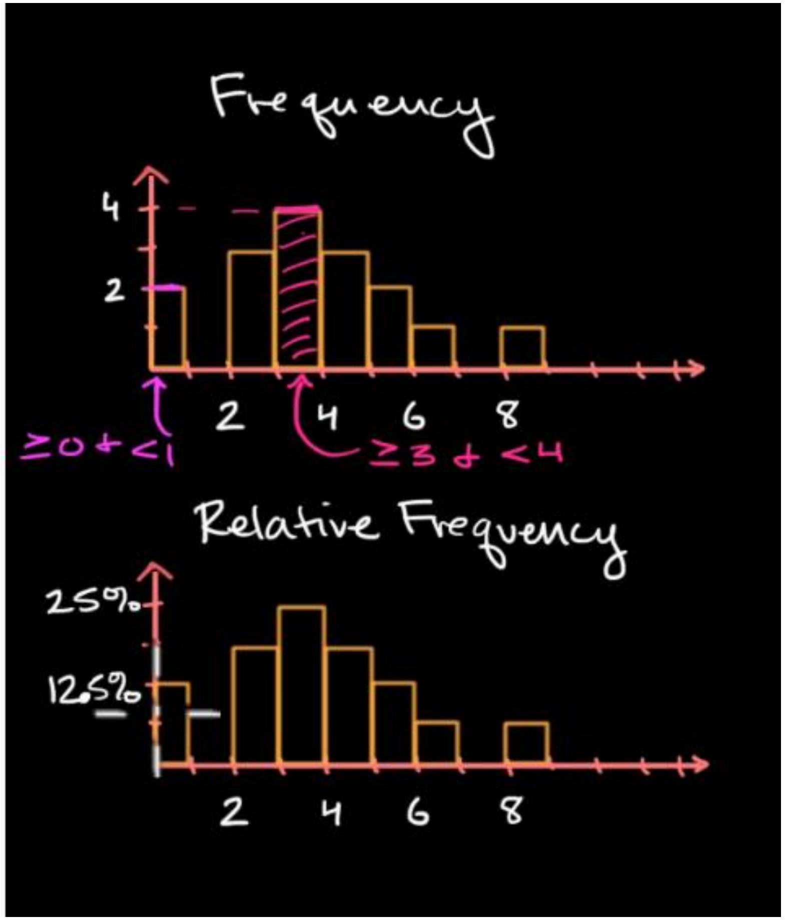
What's the difference between 'frequency' and 'relative frequency' histograms?



HISTOGRAMS

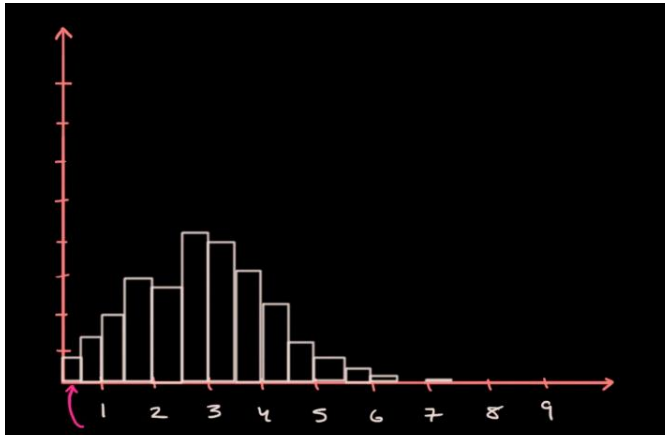
What's the difference between 'frequency' and 'relative frequency' histograms (as stated in the video)?

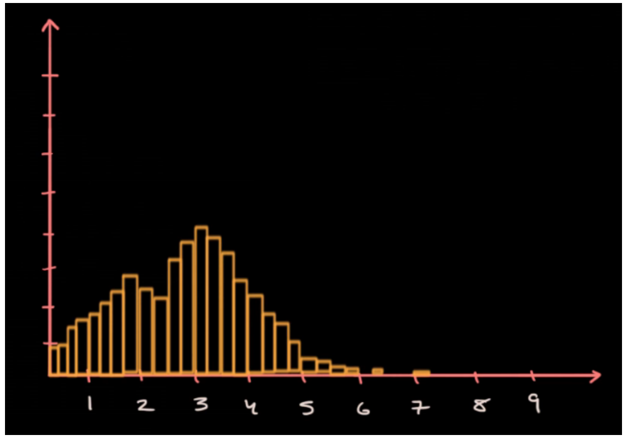
When one is better than another?

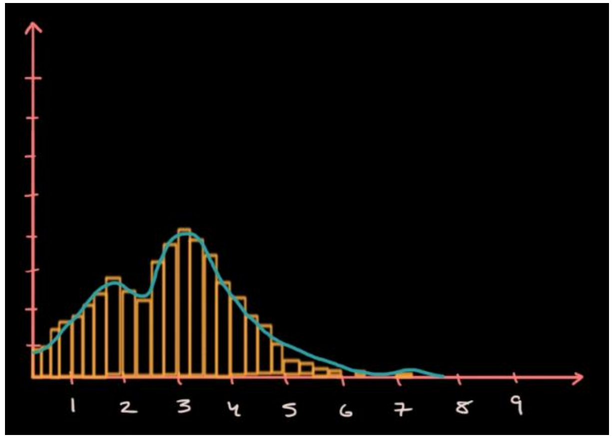


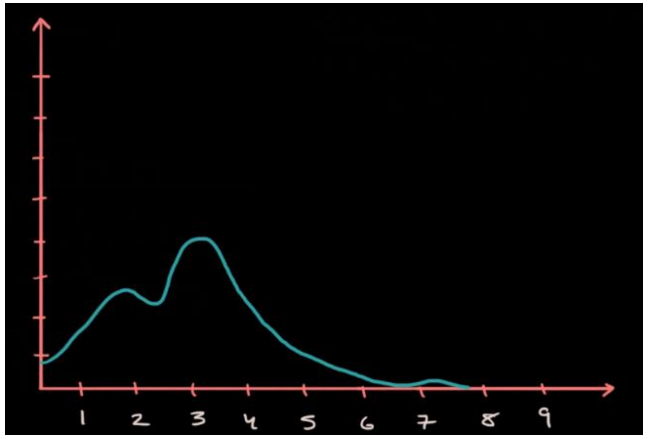
DENSITY: ESTIMATION

- **How did they obtain the density?**









DENSITY: BASIC PROPERTIES

- *Which important properties of a density function were mentioned in the video?*

DENSITY: BASIC PROPERTIES

- *Which important properties of a density function were mentioned in the video?*

1. area under the curve is 1;

DENSITY: BASIC PROPERTIES

- *Which important properties of a density function were mentioned in the video?*

1. area under the curve is 1;
2. $p(x) \geq 0$.

DENSITY: AREA UNDER THE CURVE

- *Which probability does the area of the region corresponds to?*

DENSITY: AREA UNDER THE CURVE

- *Which probability does the area of the region corresponds to?*

$$P(2 < X \leq 4)$$

DENSITY: AREA UNDER THE CURVE

- *Which probability does the area of the region corresponds to?*

$$P(2 < X \leq 4)$$

- *How to formulate it with CDF?*

DENSITY: AREA UNDER THE CURVE

- *Which probability does the area of the region corresponds to?*

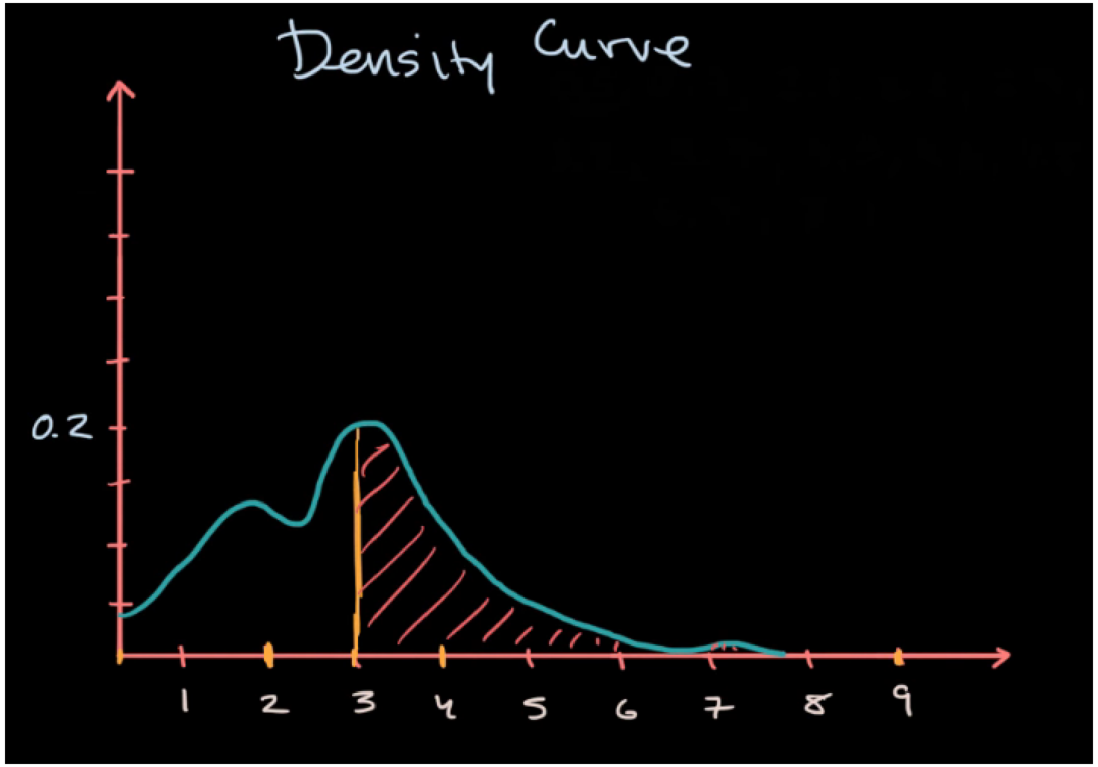
$$P(2 < X \leq 4)$$

- *How to formulate it with CDF?*

$$F(4) - F(2)$$

DENSITY: AREA UNDER THE CURVE

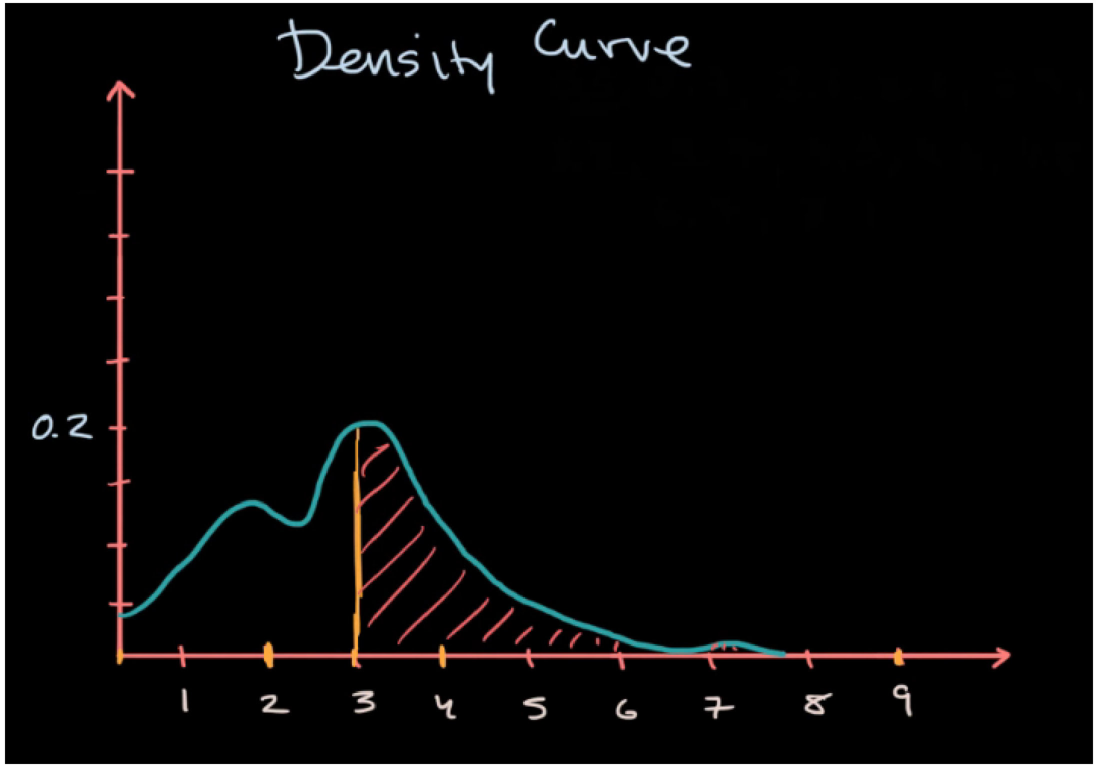
- *Which probability does the area of the region corresponds to?*



DENSITY: AREA UNDER THE CURVE

- *Which probability does the area of the region corresponds to?*

P(X > 3)

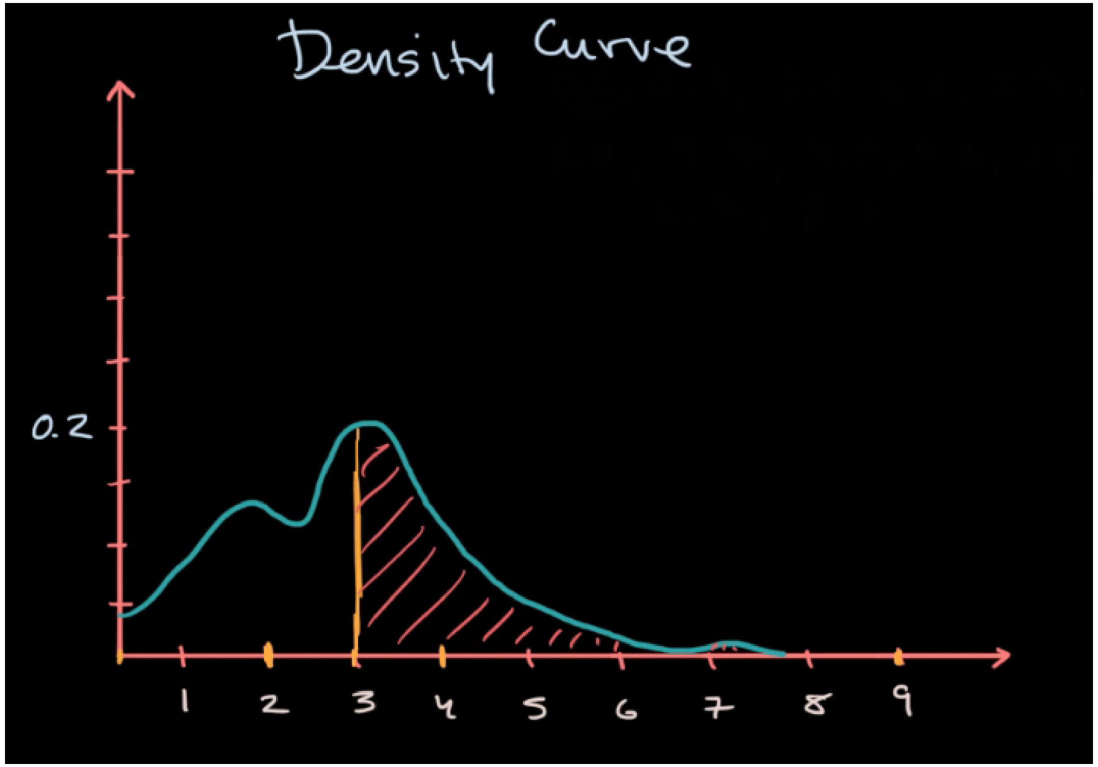


DENSITY: AREA UNDER THE CURVE

- *Which probability does the area of the region corresponds to?*

$$P(X > 3)$$

- *How to formulate it with CDF?*



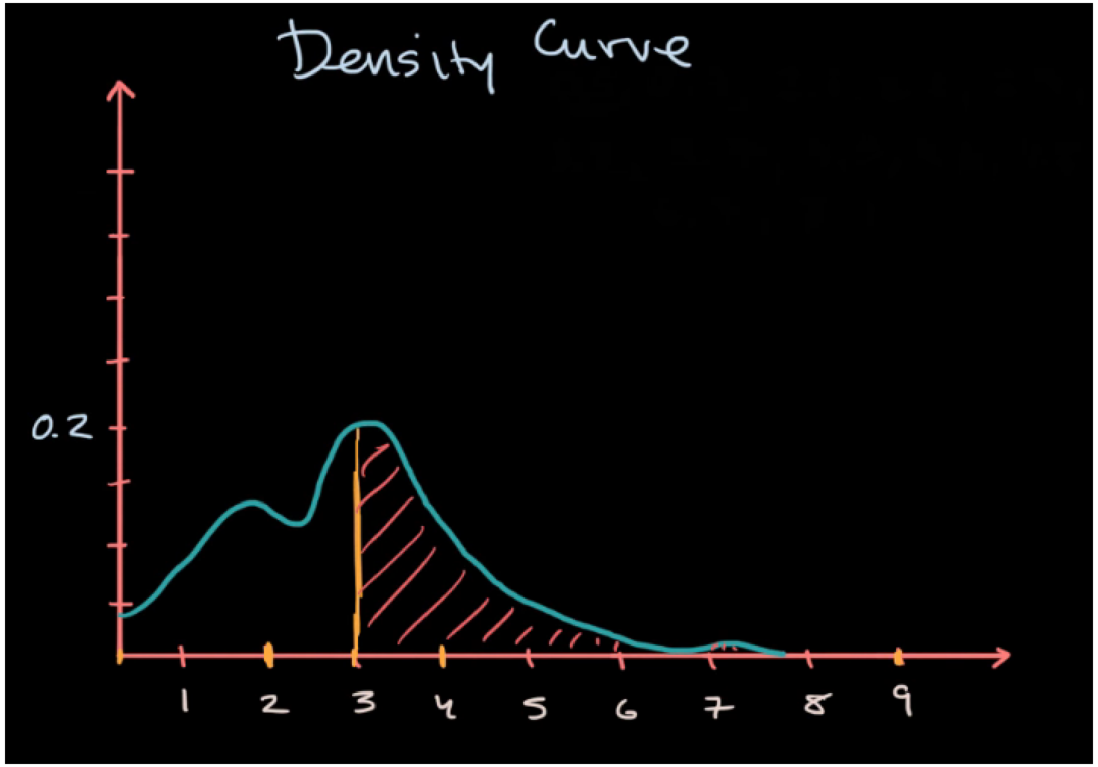
DENSITY: AREA UNDER THE CURVE

- *Which probability does the area of the region corresponds to?*

$P(X > 3)$

- *How to formulate it with CDF?*

$1 - F(3)$



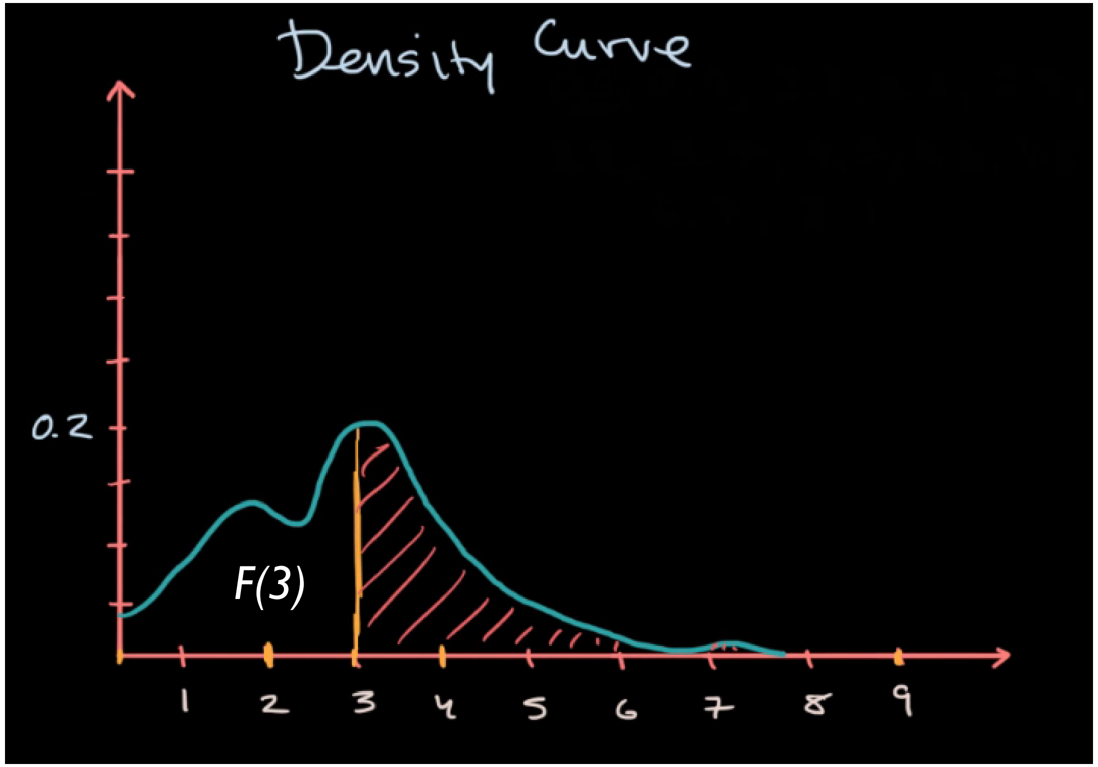
DENSITY: AREA UNDER THE CURVE

- *Which probability does the area of the region corresponds to?*

$P(X > 3)$

- *How to formulate it with CDF?*

$1 - F(3)$



THERE IS A MISTAKE IN A VIDEO THOUGH…

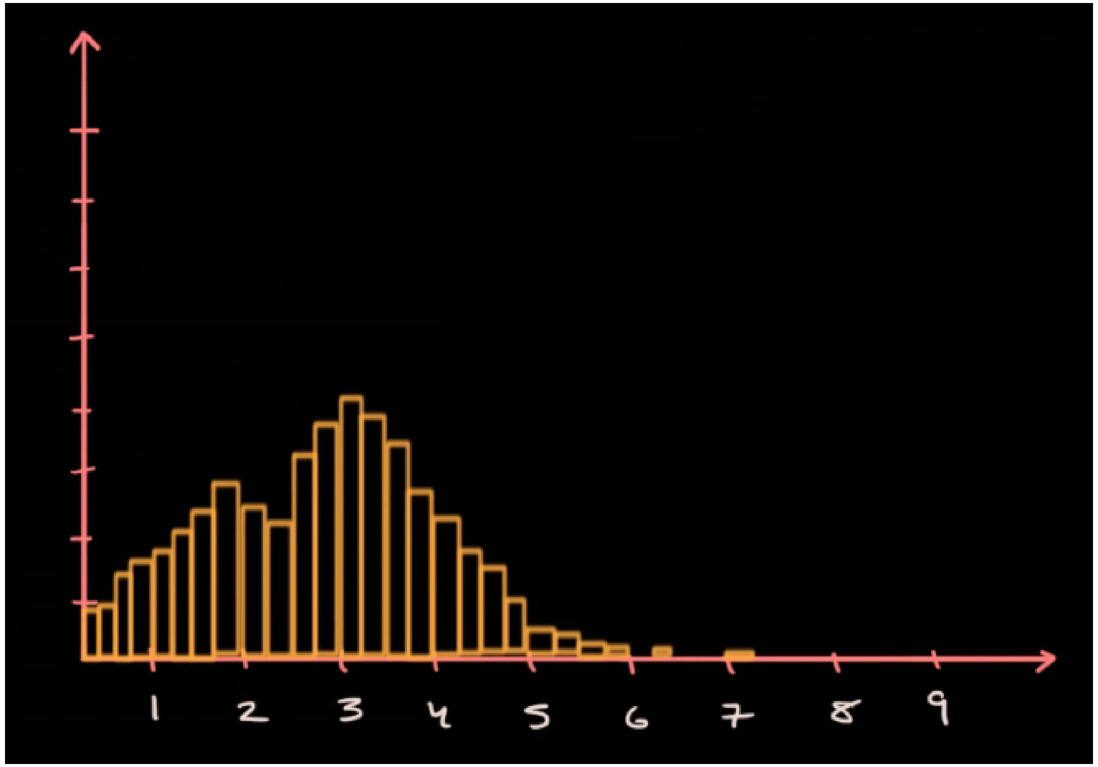
Google Classroom -> [Python] From histograms to density

NORMALIZING BIN HEIGHTS

- Total area must sum up to 1:

$$\sum_i area(bar_i) = 1$$

- What is the area of each bar for such a histogram?



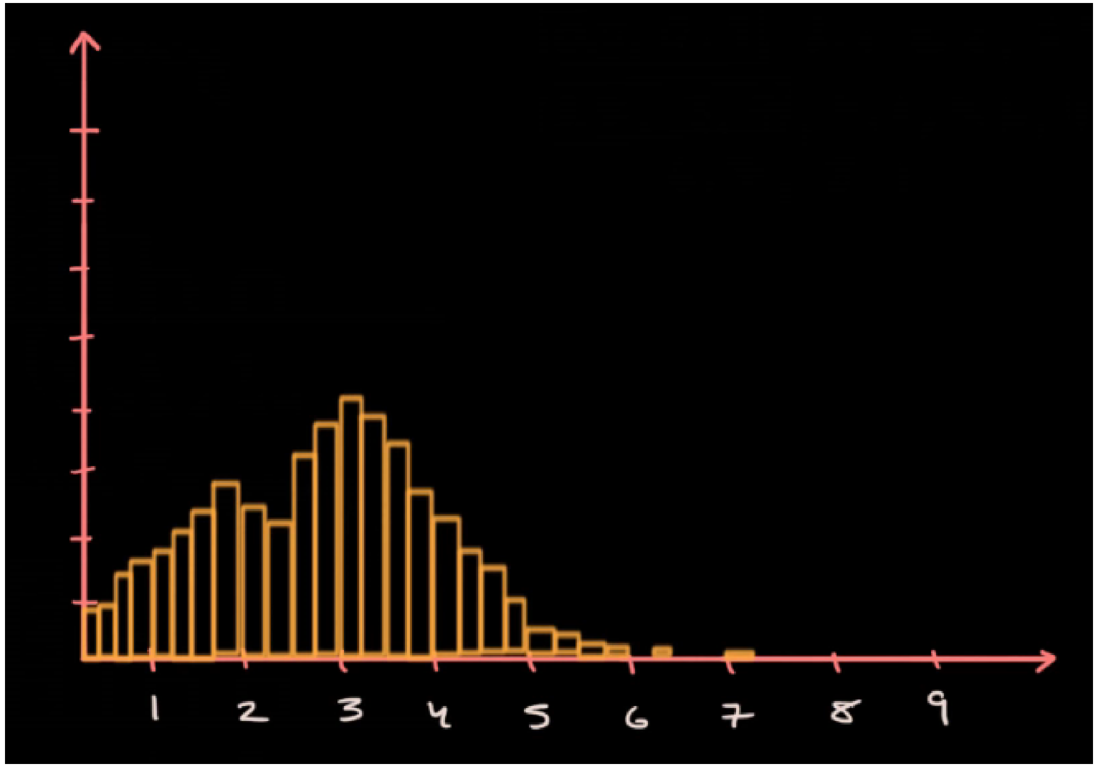
NORMALIZING BIN HEIGHTS

- Total area must sum up to 1:

$$\sum_i area(bar_i) = 1$$

- What is the area of each bar for such a histogram?

bin height x bin width



NORMALIZING BIN HEIGHTS

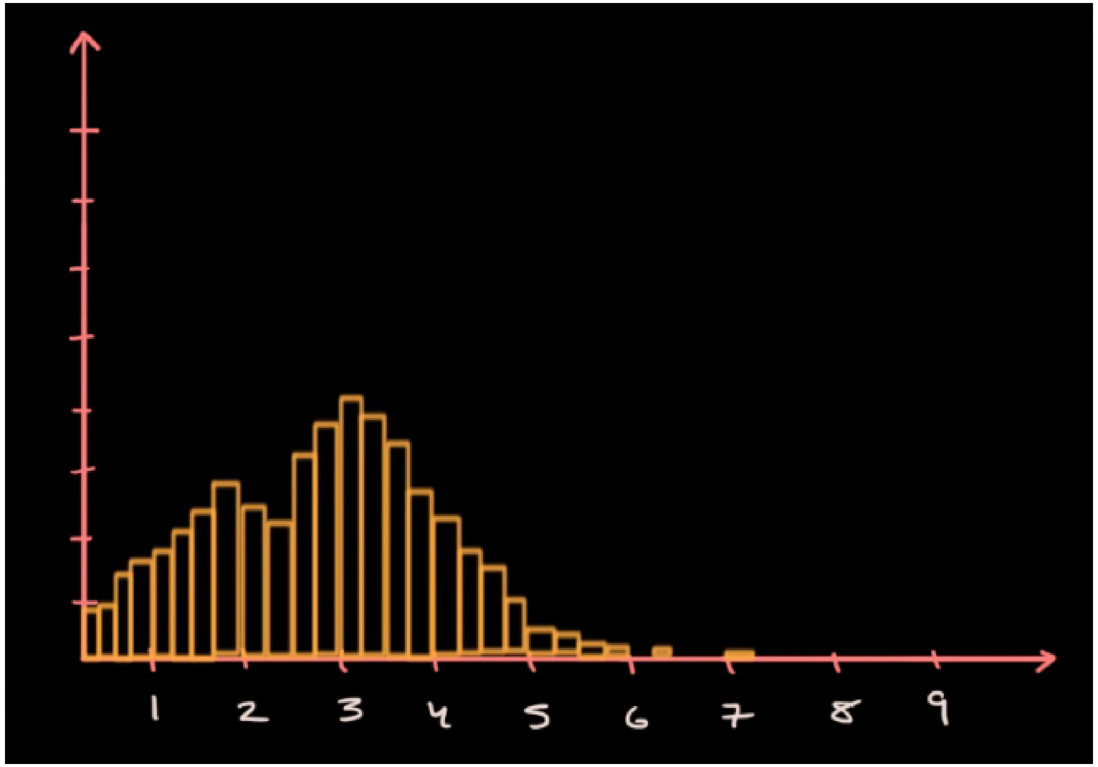
- Total area must sum up to 1:

$$\sum_i area(bar_i) = 1$$

- What is the area of each bar for such a histogram?

bin height x bin width

relative frequency x bin width



NORMALIZING BIN HEIGHTS

- Total area must sum up to 1:

$$\sum_i area(bar_i) = 1$$

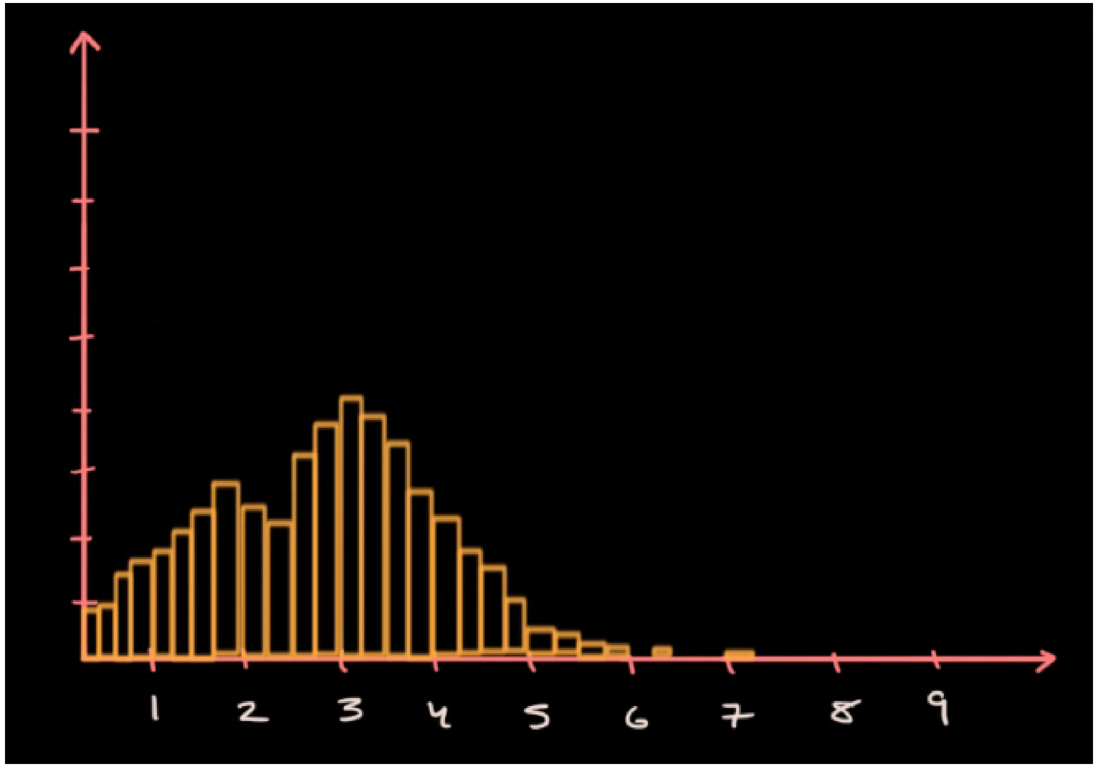
- What is the area of each bar for such a histogram?

bin height x bin width

relative frequency x bin width

$\rightarrow \mathbf{0}$

when bin width $\rightarrow \mathbf{0}$



NORMALIZING BIN HEIGHTS

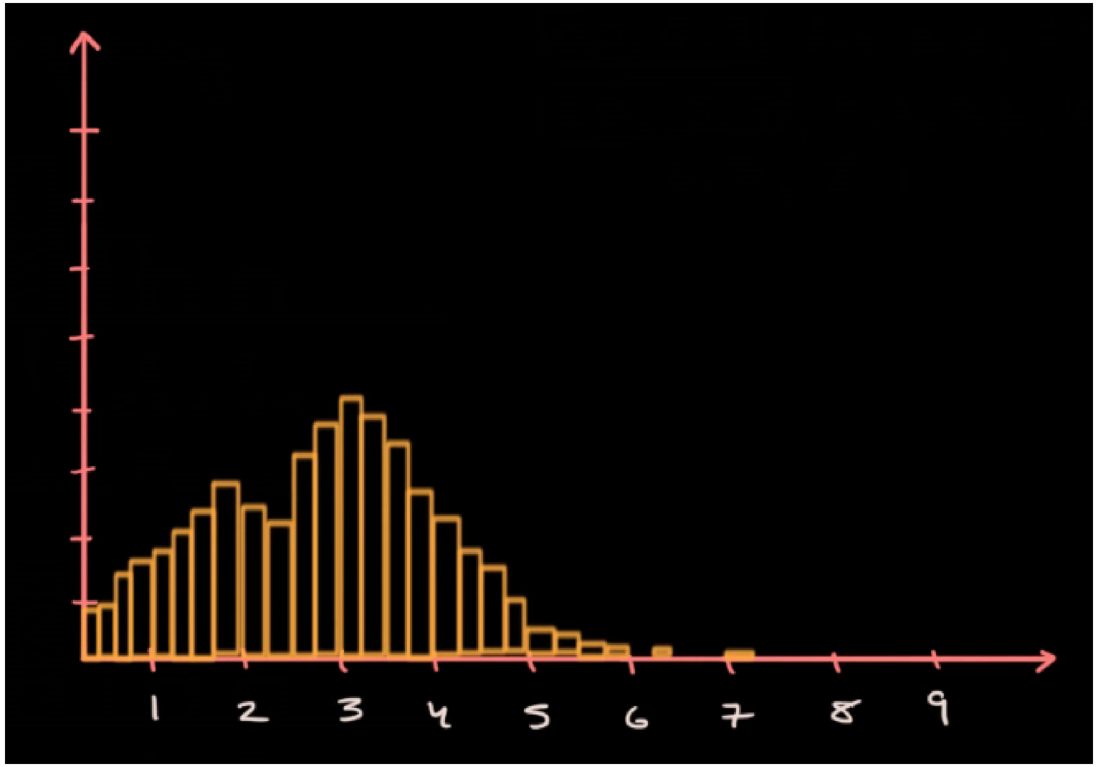
- Total area must sum up to 1:

$$\sum_i area(bar_i) = 1$$

- What is the area of each bar for such a histogram?

bin height x bin width

(relative frequency/**bin width**) x bin width



NORMALIZING BIN HEIGHTS

- Total area must sum up to 1:

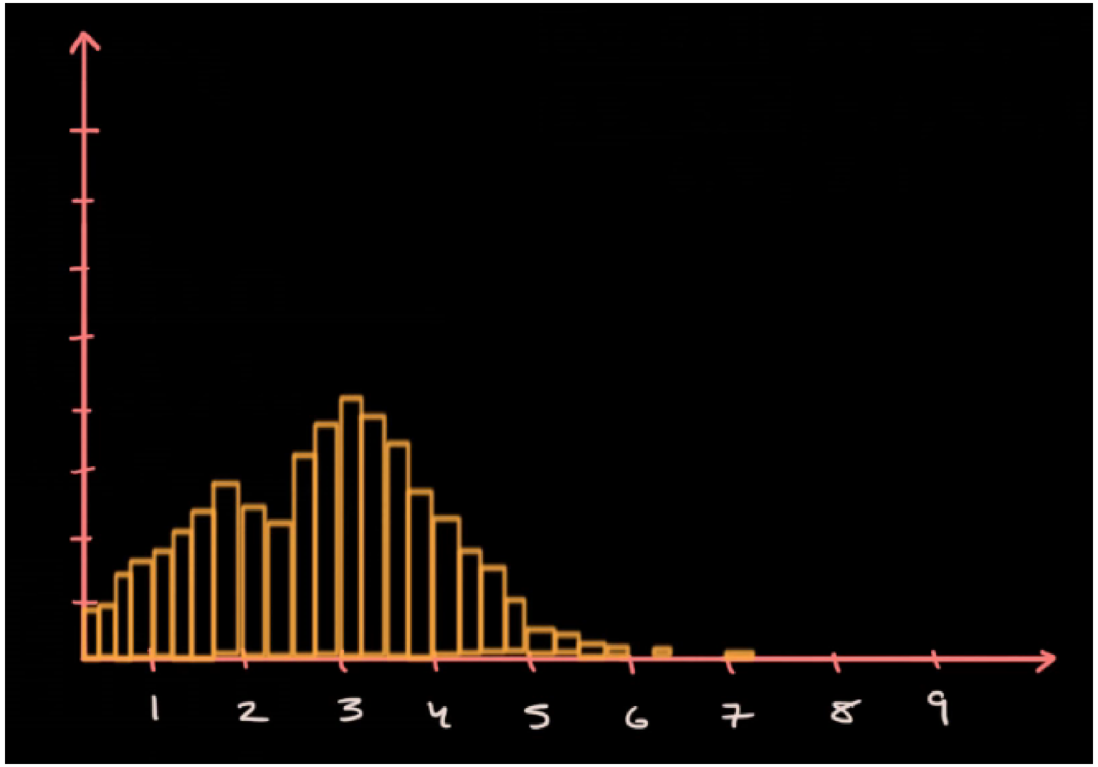
$$\sum_i area(bar_i) = 1$$

- What is the area of each bar for such a histogram?

bin height x bin width

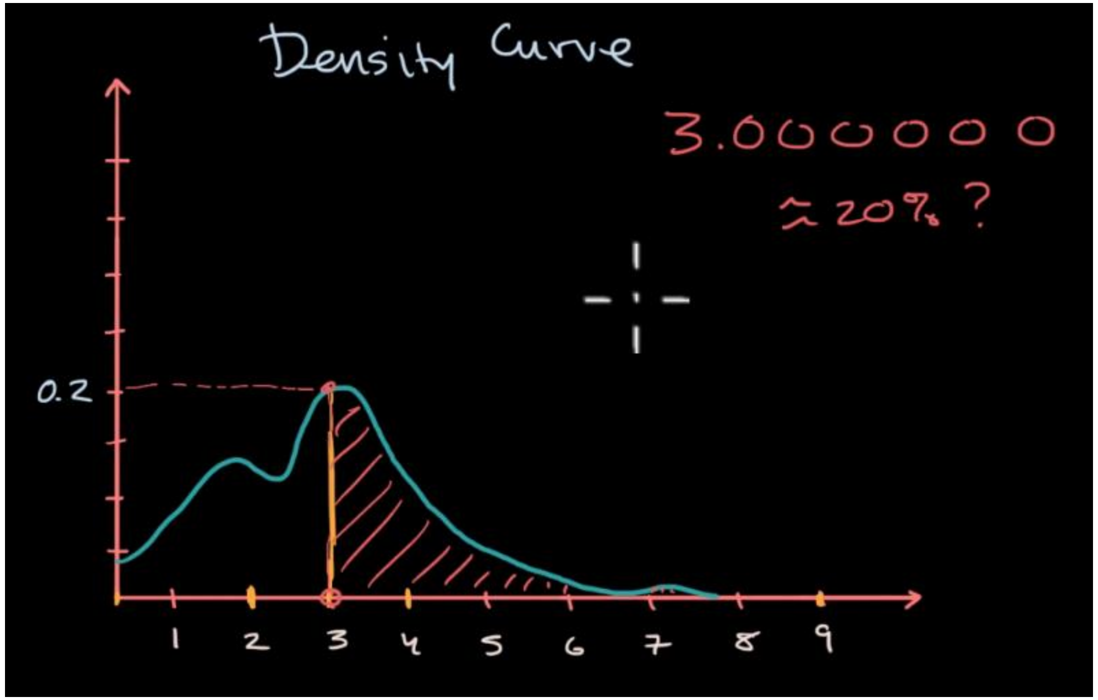
(relative frequency/**bin width**) x bin width

always sums up to 1



DENSITY: INTERPRETATION

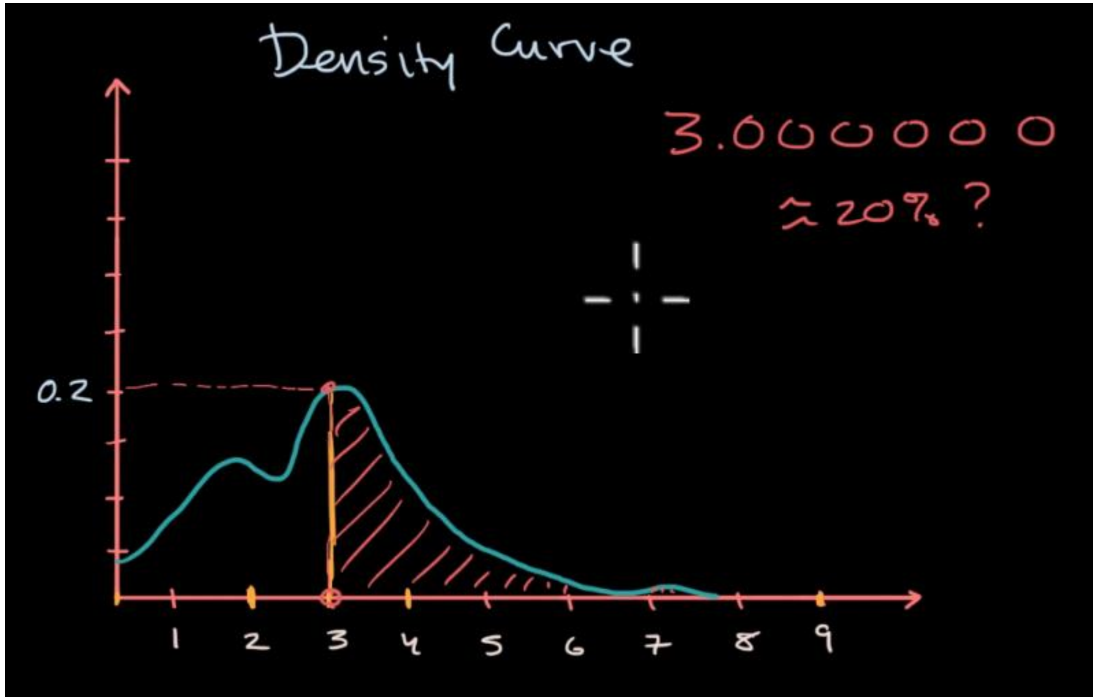
What is P(X = 3)?



DENSITY: INTERPRETATION

What is $P(X = 3)$?

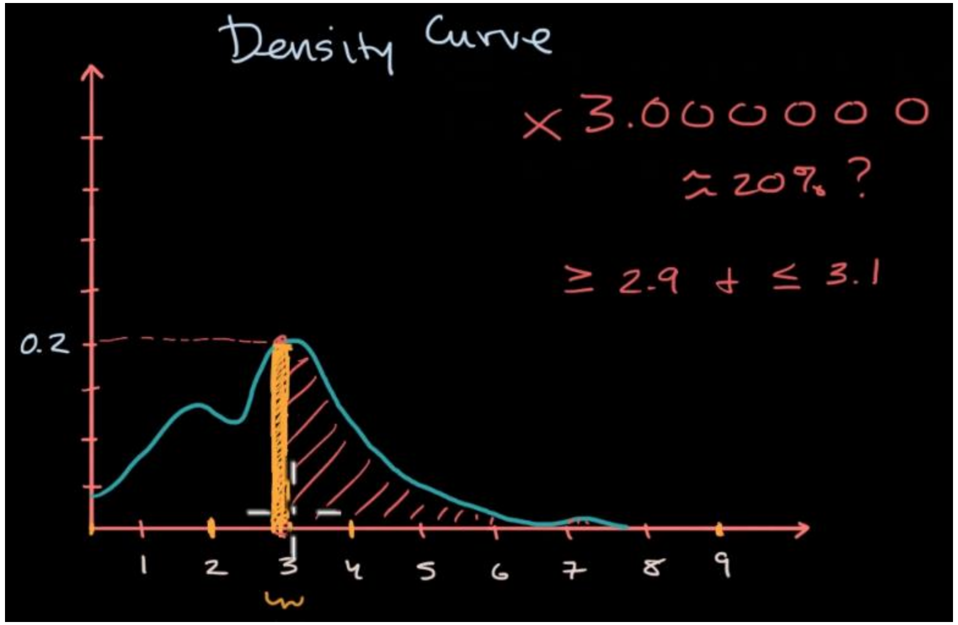
$$P(X = 3) = 0$$



RELATION BETWEEN CDF AND DENSITY

DENSITY: INTERPRETATION

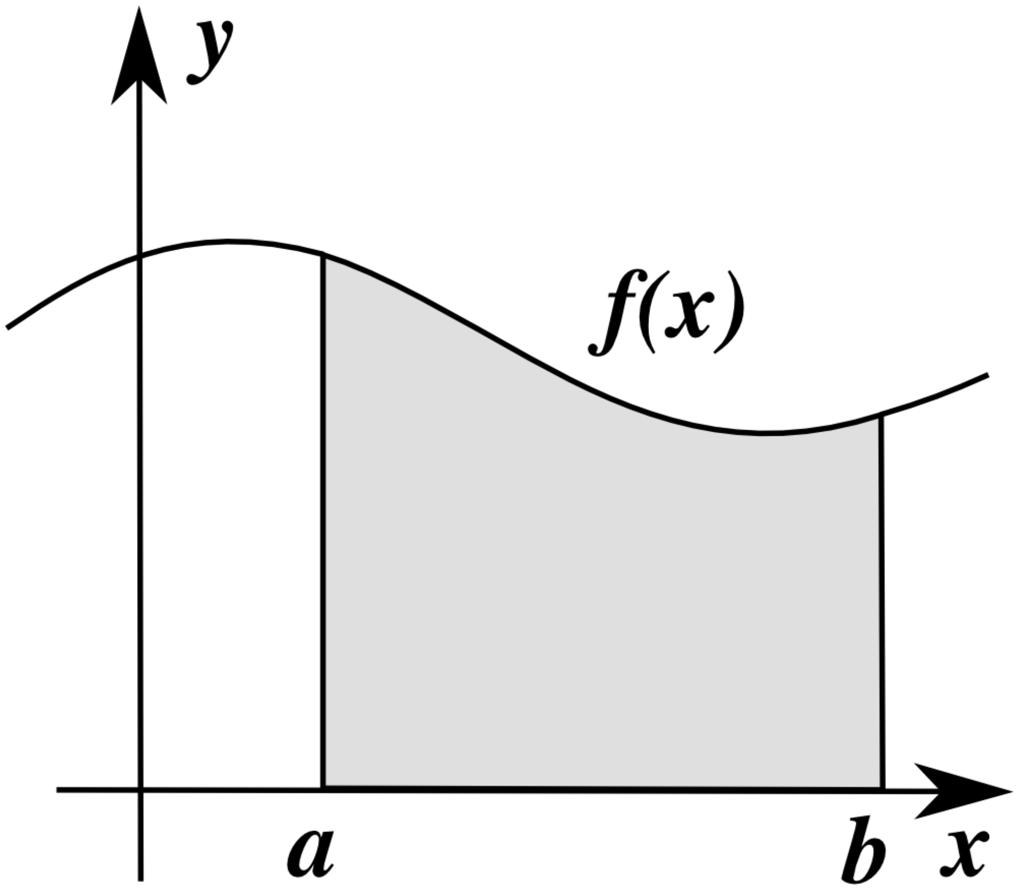
How did they estimate the area under the curve?



DENSITY: INTERPRETATION

How did they estimate the area under the curve?

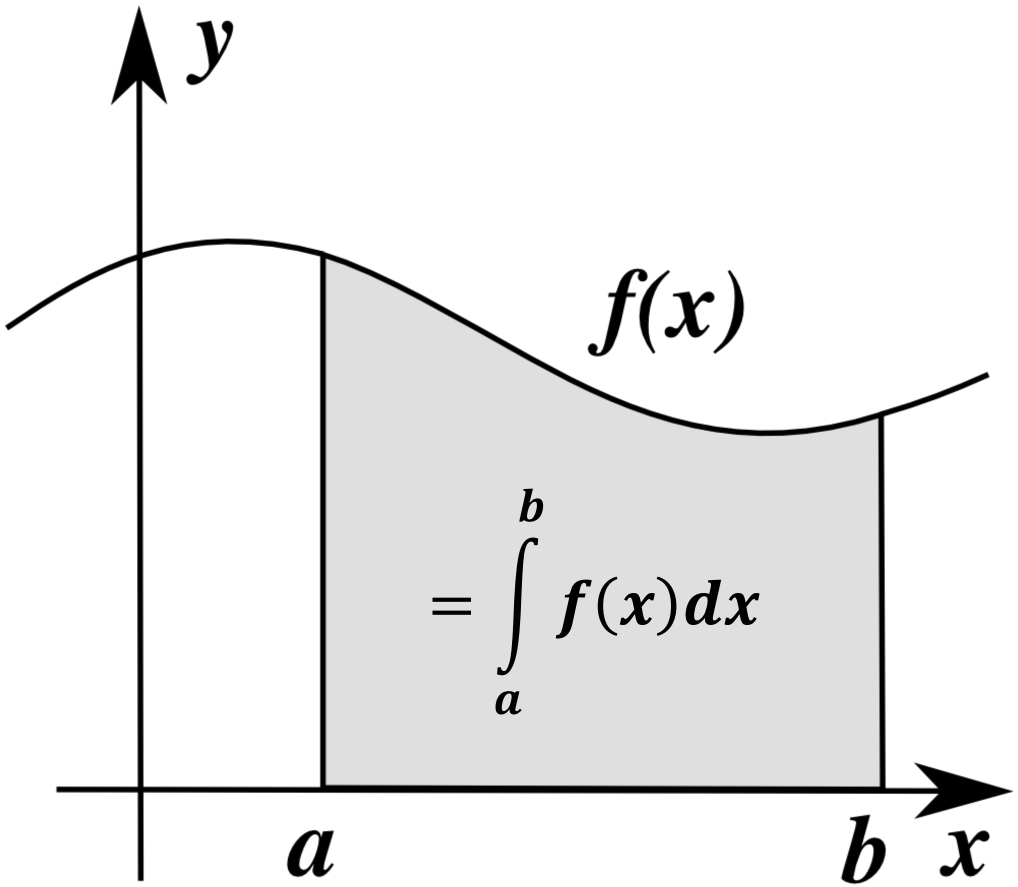
Do you remember how to compute it exactly?



DENSITY: INTERPRETATION

How did they estimate the area under the curve?

Do you remember how to compute it exactly?



CALCULUS 101

$$\int f(x)dx = F(x) + C \iff$$

$$\frac{d}{dx}F(x) = f(x)$$

CALCULUS 101

$$\int f(x)dx = F(x) + C \iff$$

$$\frac{d}{dx}F(x) = f(x)$$

$$\int x dx = \frac{x^2}{2} + C$$

CALCULUS 101

$$\int f(x)dx = F(x) + C \iff$$

$$\frac{d}{dx}F(x) = f(x)$$

$$\int x dx = \frac{x^2}{2} + C$$

$$\int \frac{1}{x} dx = \log x + C$$

CALCULUS 101

$$\int f(x)dx = F(x) + C \iff$$

$$\frac{d}{dx}F(x) = f(x)$$

$$\int x dx = \frac{x^2}{2} + C$$

$$\int \frac{1}{x} dx = \log x + C$$

$$\int_a^b f(x)dx = F(b) - F(a)$$

CALCULUS 101

$$\int f(x)dx = F(x) + C \;\; \Leftrightarrow$$

$$\frac{d}{dx}F(x) = f(x)$$

$$\int x dx = \frac{x^2}{2} + C$$

$$\int \frac{1}{x} dx = \log x + C$$

$$\int_a^b f(x)dx = F(b) - F(a)$$

$$\int_0^1 x dx = \frac{1^2}{2} - \frac{0^2}{2} = \frac{1}{2}$$

DENSITY: AREA UNDER THE CURVE

- *Which probability does the area of the region corresponds to?*

$$P(2 < X \leq 4)$$

- *How to formulate it with CDF?*

$$F(4) - F(2)$$

DENSITY: AREA UNDER THE CURVE

- *Which probability does the area of the region corresponds to?*

$$\int_2^4 \boldsymbol{p(x)dx} = \mathrm{P}(2 < X \leq 4)$$

- *How to formulate it with CDF?*

F(4) – F(2)

DENSITY: AREA UNDER THE CURVE

- *Which probability does the area of the region corresponds to?*

$$\int_2^4 p(x)dx = \mathrm{P}(2 < X \leq 4)$$

- *How to formulate it with CDF?*

$$\int_2^4 p(x)dx = \mathrm{F}(4) - \mathrm{F}(2)$$

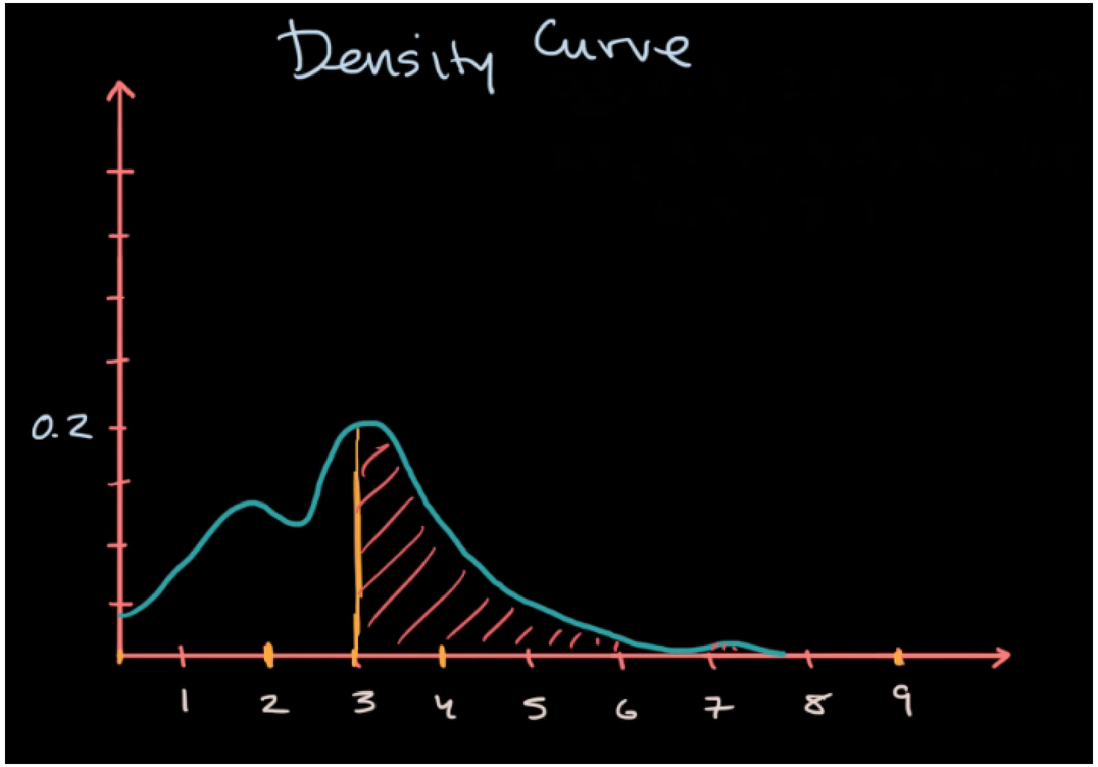
DENSITY: AREA UNDER THE CURVE

- *Which probability does the area of the region corresponds to?*

$$P(X > 3)$$

- *How to formulate it with CDF?*

$$1 - F(3)$$



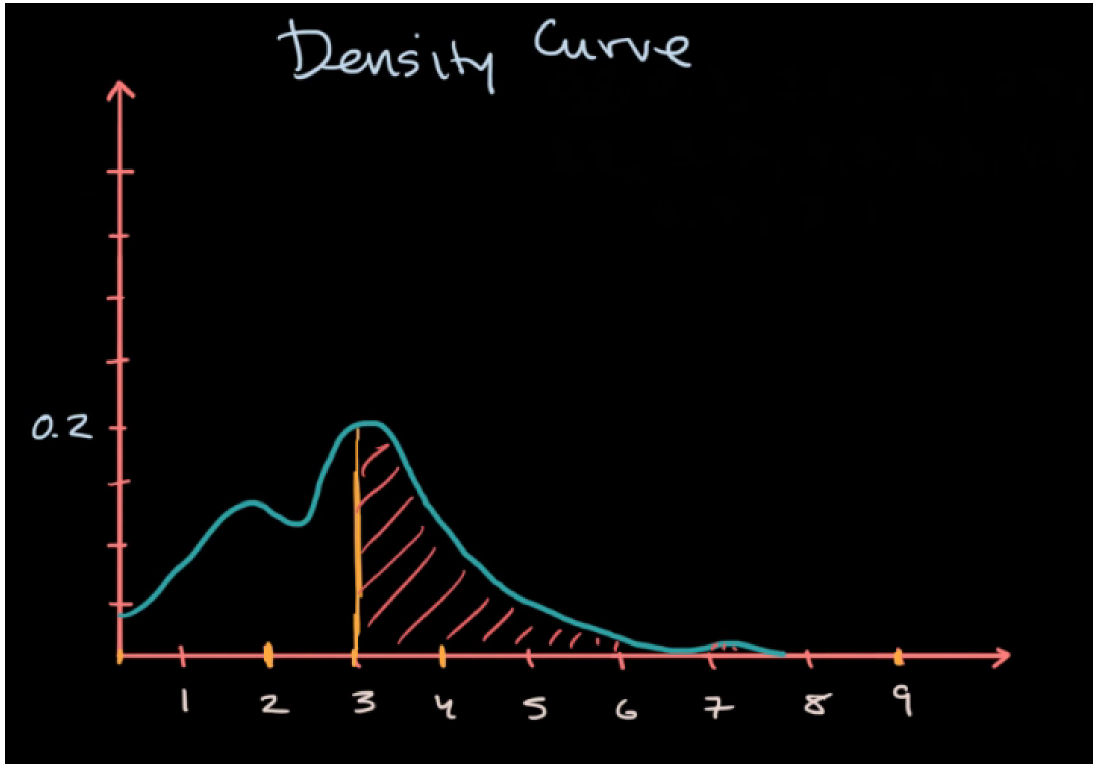
DENSITY: AREA UNDER THE CURVE

- *Which probability does the area of the region corresponds to?*

$$\int_3^{+\infty} p(x)dx = P(X > 3)$$

- *How to formulate it with CDF?*

$$1 - F(3)$$



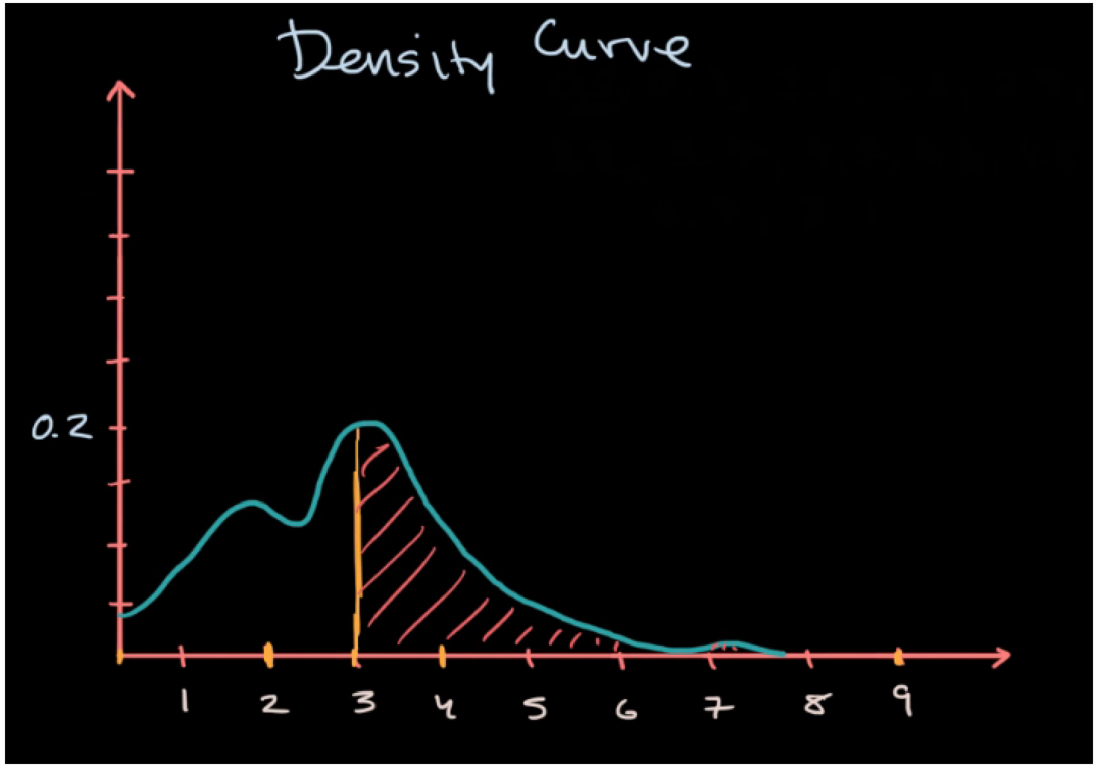
DENSITY: AREA UNDER THE CURVE

- *Which probability does the area of the region corresponds to?*

$$\int_{3}^{+\infty} p(x)dx = P(X > 3)$$

- *How to formulate it with CDF?*

$$\int_{3}^{+\infty} p(x)dx = 1 - F(3) =$$



DENSITY: AREA UNDER THE CURVE

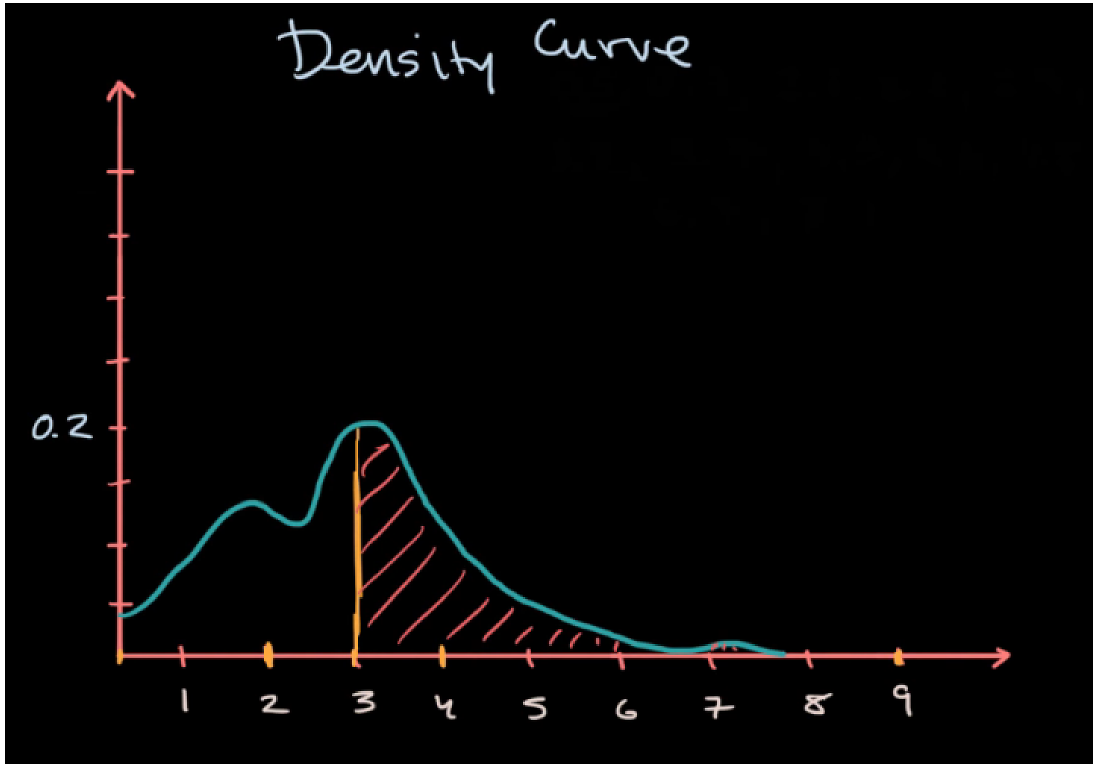
- *Which probability does the area of the region corresponds to?*

$$\int_{3}^{+\infty} p(x)dx = P(X > 3)$$

- *How to formulate it with CDF?*

$$\int_{3}^{+\infty} p(x)dx = 1 - F(3) =$$

$$= 1 - \int_{-\infty}^{3} p(x)dx$$



CDF AND PDF

- $\int_2^4 p(x)dx$ = F(4) – F(2)
- $\int_{-\infty}^3 p(x)dx = \mathrm{F}(3)$

CDF AND PDF

- $\int_2^4 p(x)dx$ = F(4) – F(2)
- $\int_{-\infty}^{3} p(x)dx = \mathrm{F}(3)$

$$F(x) = P(X \le x) = \int_{-\infty}^{x} p(t)dt$$

$$\frac{d}{dx}F(x) = \mathrm{p}(x)$$

PDF OF THE UNIFORM DISTRIBUTION

$X \sim U(a, b)$

CDF: *F(x)* = *P(*X $\leq$ *x)*

$$\mathrm{F(x)} = \begin{cases} 0, & x < a \\ \dfrac{\mathrm{x - a}}{\mathrm{b - a}}, & a \leq x \leq b \\ 1, & x > b \end{cases}$$

- **PDF:**

PDF OF THE UNIFORM DISTRIBUTION

$X \sim U(a, b)$

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} p(t)dt$$

$$\frac{d}{dx} F(x) = \mathrm{p}(x)$$

CDF: $F(x) = P(X \leq x)$

- **PDF:**

$$\mathrm{F(x)} = \begin{cases} 0, & x < a \\ \dfrac{\mathrm{x - a}}{\mathrm{b - a}}, & a \leq x \leq b \\ 1, & x > b \end{cases}$$

$$p(x) = \begin{cases} \\ \end{cases}$$

PDF OF THE UNIFORM DISTRIBUTION

$X \sim U(a, b)$

CDF: $F(x) = P(X \leq x)$

$$F(x) = \begin{cases} 0, & x < a \\ \dfrac{x - a}{b - a}, & a \leq x \leq b \\ 1, & x > b \end{cases}$$

- **PDF:**

$$\boxed{\begin{gathered} F(x) = P(X \leq x) = \int_{-\infty}^{x} p(t)dt \\ \tfrac{d}{dx} F(x) = \mathrm{p}(x) \end{gathered}}$$

$$p(x) = \begin{cases} & x \leq a \text{ or } x \geq b \\ & a < b < b \end{cases}$$

PDF OF THE UNIFORM DISTRIBUTION

$X \sim U(a, b)$

CDF: $F(x) = P(X \leq x)$

$$F(x) = \begin{cases} 0, & x < a \\ \dfrac{x-a}{b-a}, & a \leq x \leq b \\ 1, & x > b \end{cases}$$

- **PDF:**

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} p(t)dt$$

$$\frac{d}{dx}F(x) = p(x)$$

$$p(x) = \begin{cases} 0, & x \leq a \text{ or } x \geq b \\ & a < b < b \end{cases}$$

PDF OF THE UNIFORM DISTRIBUTION

$X \sim U(a, b)$

CDF: $F(x) = P(X \leq x)$

$$F(x) = \begin{cases} 0, & x < a \\ \dfrac{x-a}{b-a}, & a \leq x \leq b \\ 1, & x > b \end{cases}$$

- **PDF:**

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} p(t)dt$$

$$\frac{d}{dx}F(x) = p(x)$$

$$p(x) = \begin{cases} 0, & x \leq a \text{ or } x \geq b \\ \dfrac{1}{b-a}, & a < b < b \end{cases}$$

SHORT QUIZ

CDFs and PDFs

WHAT IS DIFFERENT FROM THE REST

1. $P(X \leq 3)$
2. $\int_{3}^{+\infty} p(x)dx$
3. $F(3)$
4. $\int_{-\infty}^{3} p(x)dx$

WHAT IS DIFFERENT FROM THE REST

1. $P(X \leq 3)$
2. $\boldsymbol{\int_{3}^{+\infty} p(x)dx}$
3. $F(3)$
4. $\int_{-\infty}^{3} p(x)dx$

WHAT IS DIFFERENT FROM THE REST

1. $P(X > 5)$
2. $\int_{5}^{+\infty} p(x)dx$
3. $\int_{-\infty}^{5} p(x)dx$
4. $1 - F(5)$

WHAT IS DIFFERENT FROM THE REST

1. $P(X > 5)$
2. $\int_{5}^{+\infty} p(x)dx$
3. $\boldsymbol{\int_{-\infty}^{5} p(x)dx}$
4. $1 - F(5)$

WHAT IS DIFFERENT FROM THE REST

1. $\int_3^5 p(x)dx$

2. $P(3 < X \leq 5)$

3. $F(3) - F(5)$

4. $\int_{-\infty}^5 p(x)dx - \int_{-\infty}^3 p(x)dx$

WHAT IS DIFFERENT FROM THE REST

1. $\int_3^5 p(x)dx$

2. $P(3 < X \leq 5)$

3. $\boldsymbol{F(3) - F(5)}$

4. $\int_{-\infty}^5 p(x)dx - \int_{-\infty}^3 p(x)dx$

PROPERTIES OF CONTINUOUS RANDOM VARIABLES

UNIFORM DISTRIBUTION: WHAT WE KNOW SO FAR

- $X \sim U(a, b)$

UNIFORM DISTRIBUTION: WHAT WE KNOW SO FAR

- $X \sim U(a, b)$
- **CDF:**

$$\mathrm{F}(\mathrm{x}) = \begin{cases} 0, & x < a \\ \dfrac{\mathrm{x} - \mathrm{a}}{\mathrm{b} - \mathrm{a}}, & a \le x \le b \\ 1, & x > b \end{cases}$$

UNIFORM DISTRIBUTION: WHAT WE KNOW SO FAR

- $X \sim U(a, b)$

- **CDF:**

$$\mathrm{F(x)} = \begin{cases} 0, & x < a \\ \dfrac{\mathrm{x-a}}{\mathrm{b-a}}, & a \le x \le b \\ 1, & x > b \end{cases}$$

- **PDF:**

$$p(x) = \begin{cases} 0, & x \le a \; or \; x \ge b \\ \dfrac{1}{b-a}, & a < b < b \end{cases}$$

UNIFORM DISTRIBUTION: WHAT WE KNOW SO FAR

- $X \sim U(a, b)$

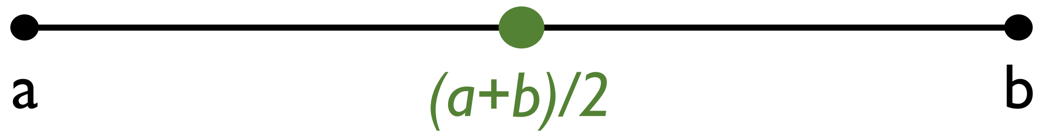
- **CDF:**

$$F(x) = \begin{cases} 0, & x < a \\ \dfrac{x - a}{b - a}, & a \leq x \leq b \\ 1, & x > b \end{cases}$$

- **PDF:**

$$p(x) = \begin{cases} 0, & x \leq a \; or \; x \geq b \\ \dfrac{1}{b - a}, & a < b < b \end{cases}$$

- $E(X) = \frac{(a+b)}{2}$



UNIFORM DISTRIBUTION: WHAT WE KNOW SO FAR

- $X \sim U(a, b)$

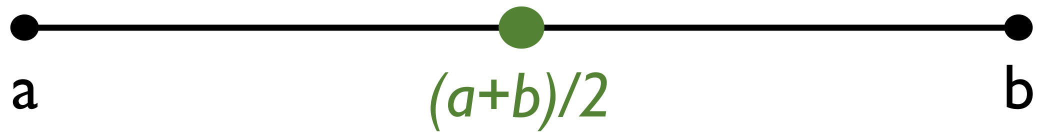
- **CDF:**

$$\mathrm{F(x)} = \begin{cases} 0, & x < a \\ \dfrac{\mathrm{x - a}}{\mathrm{b - a}}, & a \le x \le b \\ 1, & x > b \end{cases}$$

- **PDF:**

$$p(x) = \begin{cases} 0, & x \le a \text{ } or \text{ } x \ge b \\ \dfrac{1}{b - a}, & a < b < b \end{cases}$$

- $E(X) = \frac{(a+b)}{2}$
- $Var(X) = ?$



UNIFORM DISTRIBUTION: WHAT WE KNOW SO FAR

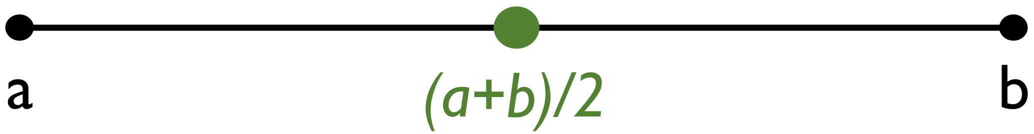
- $X \sim U(a, b)$
- **CDF:**

$$F(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \le x \le b \\ 1, & x > b \end{cases}$$

- **PDF:**

$$p(x) = \begin{cases} 0, & x \le a \text{ or } x \ge b \\ \frac{1}{b-a}, & a < b < b \end{cases}$$

- $E(X) = \frac{(a+b)}{2}$
- $Var(X) = ?$



But how to compute expected value and variance of a continuous variable?

EXPECTED VALUE

DISCRETE RANDOM VARIABLE

- Sum up all the values a random variable can take, multiplying them by their probabilities:

CONTINUOUS RANDOM VARIABLE

EXPECTED VALUE

DISCRETE RANDOM VARIABLE

- Sum up all the values a random variable can take, multiplying them by their probabilities:

$$E(X) = \sum_{X_i} X_i \cdot P(X = X_i)$$

CONTINUOUS RANDOM VARIABLE

EXPECTED VALUE

DISCRETE RANDOM VARIABLE

- Sum up all the values a random variable can take, multiplying them by their probabilities:

$$E(X) = \sum_{X_i} X_i \cdot P(X = X_i)$$

CONTINUOUS RANDOM VARIABLE

- Same principle:

EXPECTED VALUE

DISCRETE RANDOM VARIABLE

- Sum up all the values a random variable can take, multiplying them by their probabilities:

$$E(X) = \sum_{X_i} X_i \cdot P(X = X_i)$$

CONTINUOUS RANDOM VARIABLE

- Same principle:

$$\boxed{\boldsymbol{E(X) = \int_{-\infty}^{+\infty} x \cdot p(x) dx}}$$

MEAN OF UNIFORM DISTRIBUTION

$X \sim U(a, b)$

$$p(x) = \begin{cases} 0, & x \le a \; or \; x \ge b \\ \dfrac{1}{b-a}, & a < b < b \end{cases}$$

MEAN OF UNIFORM DISTRIBUTION

$X \sim U(a, b)$

$$E(X) = \int_{-\infty}^{+\infty} x \cdot p(x) dx =$$

$$p(x) = \begin{cases} 0, & x \leq a \; or \; x \geq b \\ \dfrac{1}{b-a}, & a < b < b \end{cases}$$

MEAN OF UNIFORM DISTRIBUTION

$X \sim U(a, b)$

$$p(x) = \begin{cases} 0, & x \leq a \; or \; x \geq b \\ \dfrac{1}{b-a}, & a < b < b \end{cases}$$

$$E(X) = \int_{-\infty}^{+\infty} x \cdot p(x) dx =$$

$$= \int_{a}^{b} x \cdot \frac{1}{b-a} dx =$$

MEAN OF UNIFORM DISTRIBUTION

$X \sim U(a, b)$

$$p(x) = \begin{cases} 0, & x \leq a \; or \; x \geq b \\ \dfrac{1}{b-a}, & a < b < b \end{cases}$$

$$E(X) = \int_{-\infty}^{+\infty} x \cdot p(x) dx =$$

$$= \int_{a}^{b} x \cdot \frac{1}{b-a} dx =$$

$$= \frac{1}{b-a} \cdot \frac{b^2 - a^2}{2} =$$

MEAN OF UNIFORM DISTRIBUTION

$X \sim U(a, b)$

$$p(x) = \begin{cases} 0, & x \leq a \; or \; x \geq b \\ \dfrac{1}{b-a}, & a < b < b \end{cases}$$

$$E(X) = \int_{-\infty}^{+\infty} x \cdot p(x) dx =$$

$$= \int_{a}^{b} x \cdot \frac{1}{b-a} dx =$$

$$= \frac{1}{b-a} \cdot \frac{b^2 - a^2}{2} = \frac{a+b}{2}$$

MEAN OF UNIFORM DISTRIBUTION

- Imagine that $X \sim U(0, 10)$

MEAN OF UNIFORM DISTRIBUTION

- Imagine that $X \sim U(0, 10)$
- $\mathrm{E}(\mathrm{X}) = \frac{10+0}{2} = 5$

MEAN OF UNIFORM DISTRIBUTION

- Imagine that $X \sim U(0, 10)$
- $\mathrm{E(X)} = \frac{10+0}{2} = 5$
- $\mathrm{E(X - 1)} =$

MEAN OF UNIFORM DISTRIBUTION

- Imagine that $X{\sim}U(0, 10)$
- $\mathrm{E(X)} = \frac{10+0}{2} = 5$
- $\mathrm{E(X - 1) = E(X) - 1 = 5}$

MEAN OF UNIFORM DISTRIBUTION

- Imagine that $X{\sim}U(0, 10)$
- $\mathrm{E(X)} = \frac{10+0}{2} = 5$
- $\mathrm{E(X-1)} = \mathrm{E(X)} - 1 = 5$
- $\mathrm{E(2 \cdot X)} =$

MEAN OF UNIFORM DISTRIBUTION

- Imagine that $X \sim U(0, 10)$
- $\mathrm{E}(\mathrm{X}) = \frac{10+0}{2} = 5$
- $\mathrm{E}(\mathrm{X} - 1) = \mathrm{E}(\mathrm{X}) - 1 = 4$
- $\mathrm{E}(2 \cdot \mathrm{X}) = 2 \cdot \mathrm{E}(\mathrm{X}) = 10$

MEAN OF UNIFORM DISTRIBUTION

- Imagine that $X \sim U(0, 10)$
- $\mathrm{E(X)} = \frac{10+0}{2} = 5$
- $\mathrm{E(X-1)} = \mathrm{E(X)} - 1 = 5$
- $\mathrm{E(2 \cdot X)} = 2 \cdot \mathrm{E(X)} = 10$
- $\mathrm{E(X^2)} =$

MEAN OF UNIFORM DISTRIBUTION

- Imagine that $X{\sim}U(0, 10)$
- $\mathrm{E}(\mathrm{X}) = \frac{10+0}{2} = 5$
- $\mathrm{E}(\mathrm{X} - 1) = \mathrm{E}(\mathrm{X}) - 1 = 5$
- $\mathrm{E}(2 \cdot \mathrm{X}) = 2 \cdot \mathrm{E}(\mathrm{X}) = 10$
- $\mathrm{E}(\mathrm{X}^2) =$

To compute mean of some f(X):

$$\boldsymbol{E}\big(\boldsymbol{f}(\boldsymbol{X})\big) = \int_{-\infty}^{+\infty} \boldsymbol{f}(\boldsymbol{x}) \cdot \boldsymbol{p}(\boldsymbol{x}) \boldsymbol{d}\boldsymbol{x}$$

MEAN OF UNIFORM DISTRIBUTION

- Imagine that $X \sim U(0, 10)$
- $\mathrm{E}(\mathrm{X}) = \frac{10+0}{2} = 5$
- $\mathrm{E}(\mathrm{X} - 1) = \mathrm{E}(\mathrm{X}) - 1 = 5$
- $\mathrm{E}(2 \cdot \mathrm{X}) = 2 \cdot \mathrm{E}(\mathrm{X}) = 10$
- $\mathrm{E}(\mathrm{X}^2) = \int_0^{10} x^2 \cdot \frac{1}{10} dx =$

To compute mean of some f(X):

$$\boldsymbol{E}\big(\boldsymbol{f}(\boldsymbol{X})\big) = \int_{-\infty}^{+\infty} \boldsymbol{f}(\boldsymbol{x}) \cdot \boldsymbol{p}(\boldsymbol{x}) \boldsymbol{dx}$$

MEAN OF UNIFORM DISTRIBUTION

- Imagine that $X \sim U(0, 10)$
- $\mathrm{E}(\mathrm{X}) = \frac{10+0}{2} = 5$
- $\mathrm{E}(\mathrm{X} - 1) = \mathrm{E}(\mathrm{X}) - 1 = 5$
- $\mathrm{E}(2 \cdot \mathrm{X}) = 2 \cdot \mathrm{E}(\mathrm{X}) = 10$
- $\mathrm{E}(\mathrm{X}^2) = \int_0^{10} x^2 \cdot \frac{1}{10} dx = \frac{1}{10} \cdot \frac{10^3}{3} =$

To compute mean of some f(X):

$$\boldsymbol{E}\big(\boldsymbol{f}(\boldsymbol{X})\big) = \int_{-\infty}^{+\infty} \boldsymbol{f}(\boldsymbol{x}) \cdot \boldsymbol{p}(\boldsymbol{x}) \boldsymbol{d}\boldsymbol{x}$$

MEAN OF UNIFORM DISTRIBUTION

- Imagine that $X \sim U(0, 10)$
- $\mathrm{E(X)} = \frac{10+0}{2} = 5$
- $\mathrm{E(X-1)} = \mathrm{E(X)} - 1 = 5$
- $\mathrm{E(2 \cdot X)} = 2 \cdot \mathrm{E(X)} = 10$
- $\mathrm{E(X^2)} = \int_0^{10} x^2 \cdot \frac{1}{10} dx = \frac{1}{10} \cdot \frac{10^3}{3} = \frac{100}{3}$

To compute mean of some f(X):

$$\boldsymbol{E\big(f(X)\big) = \int_{-\infty}^{+\infty} f(x) \cdot p(x) dx}$$

VARIANCE

DISCRETE RANDOM VARIABLE

CONTINUOUS RANDOM VARIABLE

- *Expected squared distance between a value and the mean:*

$$Var(X) = E\left(\left(X - E(X)\right)^2\right)$$

VARIANCE

DISCRETE RANDOM VARIABLE

- *Expected squared distance between a value and the mean:*

$$Var(X) = E\left(\left(X - E(X)\right)^2\right) =$$
$$= E(X^2) - \left(E(X)\right)^2$$

CONTINUOUS RANDOM VARIABLE

VARIANCE

DISCRETE RANDOM VARIABLE

CONTINUOUS RANDOM VARIABLE

- *Expected squared distance between a value and the mean:*

$$Var(X) = E\left(\left(X - E(X)\right)^2\right) =$$
$$= E(X^2) - \left(E(X)\right)^2$$

E.g., for Bernoulli

VARIANCE

DISCRETE RANDOM VARIABLE

CONTINUOUS RANDOM VARIABLE

- *Expected squared distance between a value and the mean:*

$$Var(X) = E\left(\left(X - E(X)\right)^2\right) =$$
$$= E(X^2) - \left(E(X)\right)^2$$

E.g., for Bernoulli

Var(X) = $p - p^2 = p(1\text{-}p)$

VARIANCE

DISCRETE RANDOM VARIABLE

- *Expected squared distance between a value and the mean:*

$$Var(X) = E\left(\left(X - E(X)\right)^2\right) =$$
$$= E(X^2) - \left(E(X)\right)^2$$

E.g., for Bernoulli

Var(X) = $p - p^2 = p(1\text{-}p)$

CONTINUOUS RANDOM VARIABLE

- Same principle.

VARIANCE

DISCRETE RANDOM VARIABLE

- *Expected squared distance between a value and the mean:*

$$Var(X) = E\left(\left(X - E(X)\right)^2\right) =$$
$$= E(X^2) - \left(E(X)\right)^2$$

E.g., for Bernoulli

Var(X) = $p - p^2 = p(1\text{-}p)$

CONTINUOUS RANDOM VARIABLE

- Same principle.
- X ~ U(a, b), Var(X) = ?

VARIANCE OF A UNIFORM DISTRIBUTION

- $X \sim U(a, b)$

- $E(X) = \frac{a+b}{2}$

VARIANCE OF A UNIFORM DISTRIBUTION

- $X \sim U(a, b)$
- $E(X) = \frac{a+b}{2}$
- $Var(X) = E\left(\left(X - E(X)\right)^2\right) = E(X^2) - \left(E(X)\right)^2$

VARIANCE OF A UNIFORM DISTRIBUTION

- $X \sim U(a, b)$

- $E(X) = \frac{a+b}{2}$

- $Var(X) = E\left(\left(X - E(X)\right)^2\right) = \boldsymbol{E\left(X^2\right) - \left(E(X)\right)^2}$

- $Var(X) = \boldsymbol{E\left(X^2\right)} - \left(\boldsymbol{\frac{a+b}{2}}\right)^2 =$

VARIANCE OF A UNIFORM DISTRIBUTION

- $X \sim U(a, b)$
- $E(X) = \frac{a+b}{2}$
- $Var(X) = E\left(\left(X - E(X)\right)^2\right) = E(X^2) - \left(E(X)\right)^2$
- $Var(X) = E(X^2) - \left(\frac{a+b}{2}\right)^2 = \frac{b^3 - a^3}{3} - \frac{(a+b)^2}{4} =$

VARIANCE OF A UNIFORM DISTRIBUTION

- $X \sim U(a,b)$
- $E(X) = \frac{a+b}{2}$
- $Var(X) = E\left(\left(X - E(X)\right)^2\right) = E(X^2) - \left(E(X)\right)^2$
- $Var(X) = E(X^2) - \left(\frac{a+b}{2}\right)^2 = \frac{b^3-a^3}{3} - \frac{(a+b)^2}{4} = \frac{(a-b)^2}{12}$

UNIFORM DISTRIBUTION

- $X \sim U(a, b)$

- **CDF:**

$$\mathrm{F(x)} = \begin{cases} 0, & x < a \\ \dfrac{\mathrm{x - a}}{\mathrm{b - a}}, & a \leq x \leq b \\ 1, & x > b \end{cases}$$

- **PDF:**

$$p(x) = \begin{cases} 0, & x \leq a \; or \; x \geq b \\ \dfrac{1}{b - a}, & a < b < b \end{cases}$$

- **MEAN:**

$$E(X) = \frac{(a + b)}{2}$$

- **VARIANCE:**

$$Var(X) = \frac{(a-b)^2}{12}$$

UNIFORM DISTRIBUTION

EXAMPLE: $X \sim U(0,1)$

$\mathrm{F}(\mathrm{x}) =$

$E(X) =$

$p(x) =$

$Var(X) =$

UNIFORM DISTRIBUTION

EXAMPLE: $X \sim U(0,1)$

$\mathrm{F}(\mathrm{x}) =$

$$E(X) = \frac{1-0}{2} = 0.5$$

$$Var(X) =$$

$p(x) =$

UNIFORM DISTRIBUTION

EXAMPLE: $X \sim U(0, 1)$

$\mathrm{F(x)} =$

$$E(X) = \frac{1-0}{2} = 0.5$$

$p(x) =$

$$Var(X) = \frac{(1-0)^2}{12} = \frac{1}{12}$$

UNIFORM DISTRIBUTION

EXAMPLE: $X \sim U(0, 1)$

$$\mathrm{F(x)} = \begin{cases} & x < 0 \\ & 0 \leq x \leq 1 \\ & x > 1 \end{cases}$$

$$E(X) = \frac{1 - 0}{2} = 0.5$$

$$p(x) =$$

$$Var(X) = \frac{(1 - 0)^2}{12} = \frac{1}{12}$$

UNIFORM DISTRIBUTION

EXAMPLE: $X \sim U(0,1)$

$$\mathrm{F(x)} = \begin{cases} 0, & x < 0 \\ x, & 0 \le x \le 1 \\ 1, & x > 1 \end{cases}$$

$$E(X) = \frac{1-0}{2} = 0.5$$

$$p(x) =$$

$$Var(X) = \frac{(1-0)^2}{12} = \frac{1}{12}$$

UNIFORM DISTRIBUTION

EXAMPLE: $X \sim U(0, 1)$

$$\mathrm{F(x)} = \begin{cases} 0, & x < 0 \\ x, & 0 \le x \le 1 \\ 1, & x > 1 \end{cases}$$

$$E(X) = \frac{1-0}{2} = 0.5$$

$$p(x) = \begin{cases} & x \le 0 \ or \ x \ge 1 \\ & 0 < x < 1 \end{cases}$$

$$Var(X) = \frac{(1-0)^2}{12} = \frac{1}{12}$$

UNIFORM DISTRIBUTION

EXAMPLE: $X \sim U(0,1)$

$$\mathrm{F(x)} = \begin{cases} 0, & x < 0 \\ x, & 0 \le x \le 1 \\ 1, & x > 1 \end{cases}$$

$$E(X) = \frac{1-0}{2} = 0.5$$

$$p(x) = \begin{cases} 0, & x \le 0 \ or \ x \ge 1 \\ 1, & 0 < x < 1 \end{cases}$$

$$Var(X) = \frac{(1-0)^2}{12} = \frac{1}{12}$$

EXPONENTIAL DISTRIBUTION

EXPONENTIAL DISTRIBUTION

- **Poisson distribution (discrete):**
  #of events that occur in a given period of time.

EXPONENTIAL DISTRIBUTION

- **Poisson distribution (discrete):**
  #of events that occur in a given period of time.

- **Exponential distribution (continuous)**:

EXPONENTIAL DISTRIBUTION

- **Poisson distribution (discrete):**
  #of events that occur in a given period of time.

- **Exponential distribution (continuous)**:
  time between two Poisson events.

EXPONENTIAL DISTRIBUTION

- **Poisson distribution (discrete):**
  #of events that occur in a given period of time.
  # of phone calls received every hour

- **Exponential distribution (continuous)**:
  time between two Poisson events.
  Time between the two phone calls

EXPONENTIAL DISTRIBUTION

- $X \sim Exp(\lambda), \qquad \lambda > 0$ – rate parameter.

EXPONENTIAL DISTRIBUTION

$$F(x) = P(X \le x) = \int_{-\infty}^{x} p(t)dt$$

$$\frac{d}{dx}F(x) = \mathrm{p}(x)$$

- $X \sim Exp(\lambda), \qquad \lambda > 0$ – rate parameter.

- **CDF:**

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \ge 0 \\ 0, & x < 0 \end{cases}$$

- **PDF:**

EXPONENTIAL DISTRIBUTION

$$F(x) = P(X \le x) = \int_{-\infty}^{x} p(t)dt$$

$$\frac{d}{dx}F(x) = \mathrm{p}(x)$$

- $X \sim Exp(\lambda)$, $\lambda > 0$ – rate parameter.

- **CDF:**

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \ge 0 \\ 0, & x < 0 \end{cases}$$

- **PDF:**

$$p(x) = \begin{cases} & x \ge 0 \\ & x < 0 \end{cases}$$

EXPONENTIAL DISTRIBUTION

$$F(x) = P(X \le x) = \int_{-\infty}^{x} p(t)dt$$

$$\frac{d}{dx}F(x) = \mathrm{p}(x)$$

- $X \sim Exp(\lambda)$, $\lambda > 0$ – rate parameter.

- **CDF:**

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \ge 0 \\ 0, & x < 0 \end{cases}$$

- **PDF:**

$$p(x) = \begin{cases} \lambda \cdot e^{-\lambda x}, & x \ge 0 \\ 0, & x < 0 \end{cases}$$

EXPONENTIAL DISTRIBUTION

- $X \sim Exp(\lambda), \qquad \lambda > 0$ – rate parameter.

EXPONENTIAL DISTRIBUTION

- $X \sim Exp(\lambda)$, $\lambda > 0$ – rate parameter.
- **MEAN:**

$$E(X) = \int_{-\infty}^{+\infty} x \cdot p(x)dx =$$

EXPONENTIAL DISTRIBUTION

- $X \sim Exp(\lambda)$, $\lambda > 0$ – rate parameter.
- **MEAN:**

$$E(X) = \int_{-\infty}^{+\infty} x \cdot p(x)dx = \int_{0}^{+\infty} x \cdot \lambda e^{-\lambda x} dx =$$

EXPONENTIAL DISTRIBUTION

- $X \sim Exp(\lambda)$, $\lambda > 0$ – rate parameter.
- **MEAN:**

$$E(X) = \int_{-\infty}^{+\infty} x \cdot p(x)dx = \int_{0}^{+\infty} x \cdot \lambda e^{-\lambda x} dx = \frac{1}{\lambda}$$

EXPONENTIAL DISTRIBUTION

- $X \sim Exp(\lambda), \qquad \lambda > 0$ – rate parameter.
- **MEAN:**

$$E(X) = \int_{-\infty}^{+\infty} x \cdot p(x) dx = \int_{0}^{+\infty} x \cdot \lambda e^{-\lambda x} dx = \frac{1}{\lambda}$$

- **VARIANCE:**

EXPONENTIAL DISTRIBUTION

- $X \sim Exp(\lambda), \qquad \lambda > 0$ – rate parameter.
- **MEAN:**

$$E(X) = \int_{-\infty}^{+\infty} x \cdot p(x)dx = \int_{0}^{+\infty} x \cdot \lambda e^{-\lambda x} dx = \frac{1}{\lambda}$$

- **VARIANCE:**

$$E(X^2) =$$

EXPONENTIAL DISTRIBUTION

- $X \sim Exp(\lambda)$, $\lambda > 0$ – rate parameter.

- **MEAN:**

$$E(X) = \int_{-\infty}^{+\infty} x \cdot p(x)dx = \int_{0}^{+\infty} x \cdot \lambda e^{-\lambda x} dx = \frac{1}{\lambda}$$

- **VARIANCE:**

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 \cdot p(x)dx =$$

EXPONENTIAL DISTRIBUTION

- $X \sim Exp(\lambda), \qquad \lambda > 0$ – rate parameter.
- **MEAN:**

$$E(X) = \int_{-\infty}^{+\infty} x \cdot p(x)dx = \int_{0}^{+\infty} x \cdot \lambda e^{-\lambda x}dx = \frac{1}{\lambda}$$

- **VARIANCE:**

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 \cdot p(x)dx = \int_{0}^{+\infty} x^2 \cdot \lambda e^{-\lambda x}dx =$$

EXPONENTIAL DISTRIBUTION

- $X \sim Exp(\lambda)$, $\qquad \lambda > 0$ – rate parameter.
- **MEAN:**

$$E(X) = \int_{-\infty}^{+\infty} x \cdot p(x)dx = \int_{0}^{+\infty} x \cdot \lambda e^{-\lambda x} dx = \frac{1}{\lambda}$$

- **VARIANCE:**

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 \cdot p(x)dx = \int_{0}^{+\infty} x^2 \cdot \lambda e^{-\lambda x} dx = \frac{2}{\lambda^2}$$

EXPONENTIAL DISTRIBUTION

- $X \sim Exp(\lambda), \qquad \lambda > 0$ – rate parameter.
- **MEAN:**

$$E(X) = \int_{-\infty}^{+\infty} x \cdot p(x)dx = \int_{0}^{+\infty} x \cdot \lambda e^{-\lambda x}dx = \frac{1}{\lambda}$$

- **VARIANCE:**

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 \cdot p(x)dx = \int_{0}^{+\infty} x^2 \cdot \lambda e^{-\lambda x}dx = \frac{2}{\lambda^2}$$

$$Var\,(X) =$$

EXPONENTIAL DISTRIBUTION

- $X \sim Exp(\lambda)$, $\lambda > 0$ – rate parameter.
- **MEAN:**

$$E(X) = \int_{-\infty}^{+\infty} x \cdot p(x)dx = \int_{0}^{+\infty} x \cdot \lambda e^{-\lambda x} dx = \frac{1}{\lambda}$$

- **VARIANCE:**

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 \cdot p(x)dx = \int_{0}^{+\infty} x^2 \cdot \lambda e^{-\lambda x} dx = \frac{2}{\lambda^2}$$

$$Var\,(X) = E(X^2) - \big(E(X)\big)^2 =$$

EXPONENTIAL DISTRIBUTION

- $X \sim Exp(\lambda)$, $\lambda > 0$ – rate parameter.
- **MEAN:**

$$E(X) = \int_{-\infty}^{+\infty} x \cdot p(x)dx = \int_{0}^{+\infty} x \cdot \lambda e^{-\lambda x} dx = \frac{1}{\lambda}$$

- **VARIANCE:**

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 \cdot p(x)dx = \int_{0}^{+\infty} x^2 \cdot \lambda e^{-\lambda x} dx = \frac{2}{\lambda^2}$$

$$Var\,(X) = E(X^2) - \big(E(X)\big)^2 = \frac{1}{\lambda^2}$$