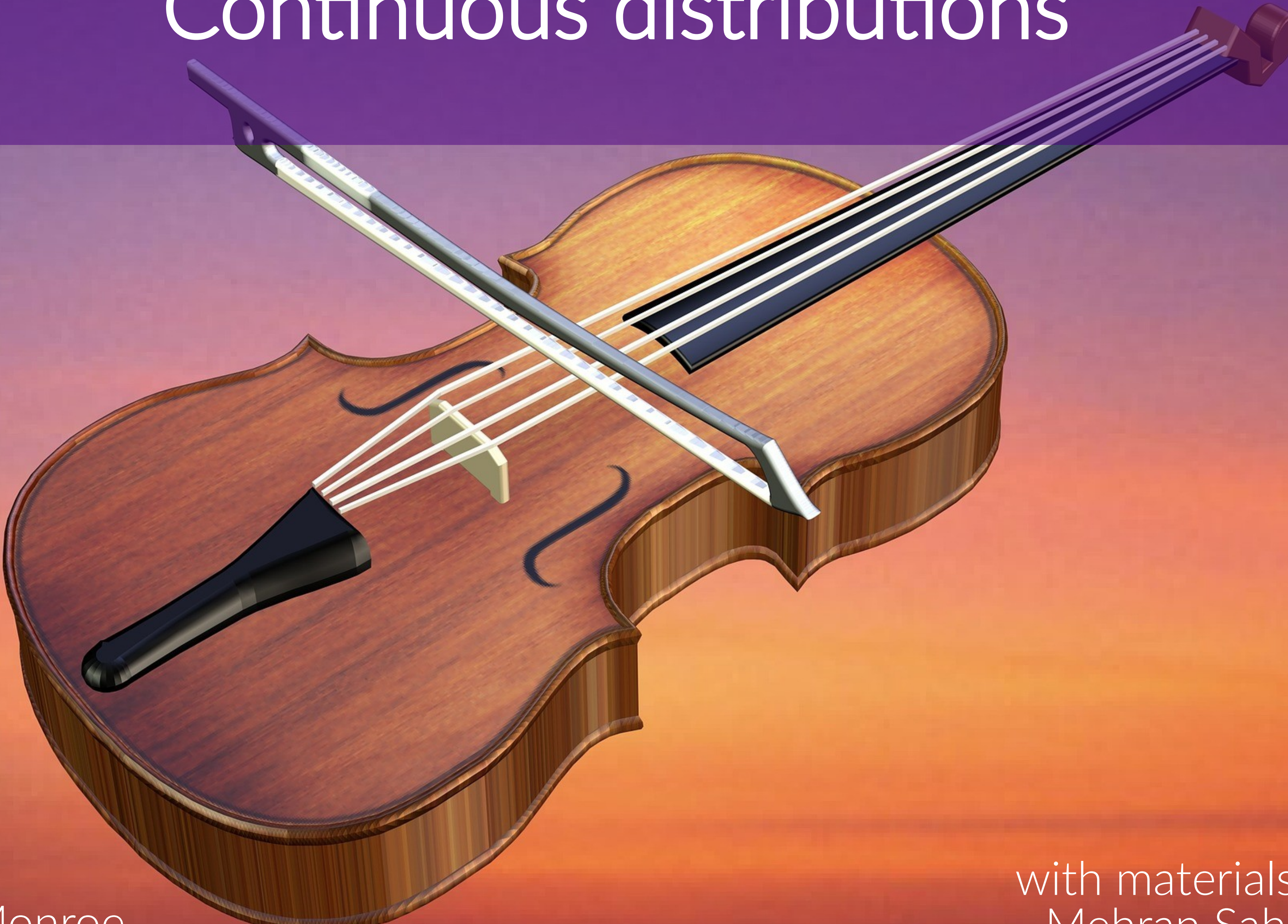
Continuous distributions

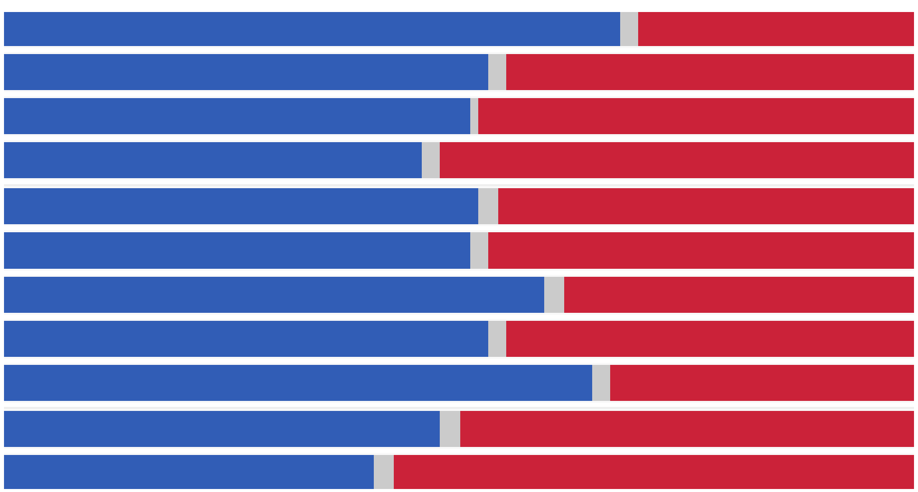


Will Monroe
July 17, 2017

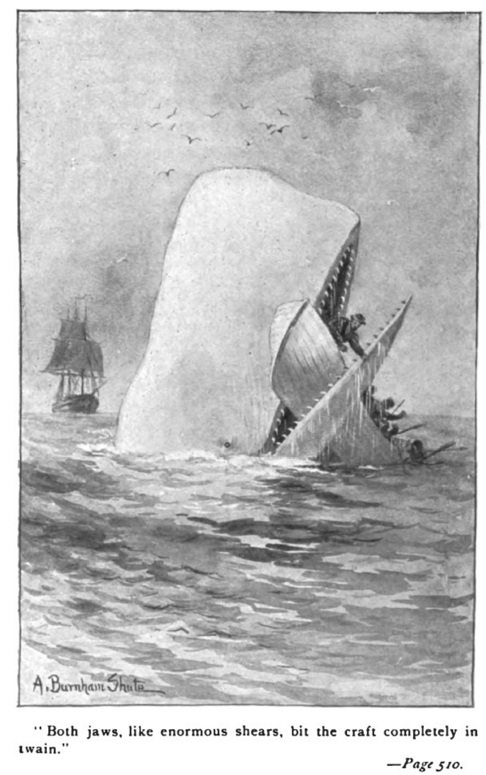
with materials by
Mehran Sahami
and Chris Piech

Announcements: Problem Set 3

Due **this Wednesday**, 7/19, at 12:30pm (before class).



(election prediction)



(Moby Dick)

Announcements: Midterm

A week from tomorrow:

Tuesday, July 25, 7:00-9:00pm

Review session:
This Thursday, July 20, 2:30-3:20pm

Geometric random variable

The **number of trials** it takes to get **one success**, if successes occur independently with probability *p*, obeys a **geometric distribution**.

$$X \sim \text{Geo}(p)$$

$$p_X(k) = \begin{cases} (1-p)^{k-1} \cdot p & \text{if } k \in \mathbb{Z}, k \geq 1 \\ 0 & \text{otherwise} \end{cases}$$

Geometric: Fact sheet

$$X \sim \text{Geo}(p)$$

probability of "success" (catch, heads, crash, ...)

PMF: $$p_X(k) = \begin{cases} (1-p)^{k-1} \cdot p & \text{if } k \in \mathbb{Z}, k \geq 1 \\ 0 & \text{otherwise} \end{cases}$$

CDF: $$F_X(k) = \begin{cases} 1-(1-p)^{k} & \text{if } k \in \mathbb{Z}, k \geq 1 \\ 0 & \text{otherwise} \end{cases}$$

expectation: $$E[X] = \frac{1}{p}$$

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

Negative binomial random variable

The **number of trials** it takes to get *r* successes, if successes occur independently with probability *p*, obeys a **negative binomial distribution**.

$$X \sim \text{NegBin}(r, p)$$

$$p_X(n) = \begin{cases} \binom{n-1}{r-1} p^r (1-p)^{n-r} & \text{if } n \in \mathbb{Z}, n \geq r \\ 0 & \text{otherwise} \end{cases}$$

Negative binomial: Fact sheet

number of **sucesses** (heads, crash, ...)

$$X \sim \mathrm{NegBin}(r, p)$$

number of **trials** (flips, program runs, ...)

probability of "success"

PMF: $$p_X(n) = \begin{cases} \binom{n-1}{r-1} p^r (1-p)^{n-r} & \text{if } n \in \mathbb{Z}, n \geq r \\ 0 & \text{otherwise} \end{cases}$$

expectation: $$E[X] = \frac{r}{p}$$

variance: $$\mathrm{Var}(X) = \frac{r(1-p)}{p^2}$$

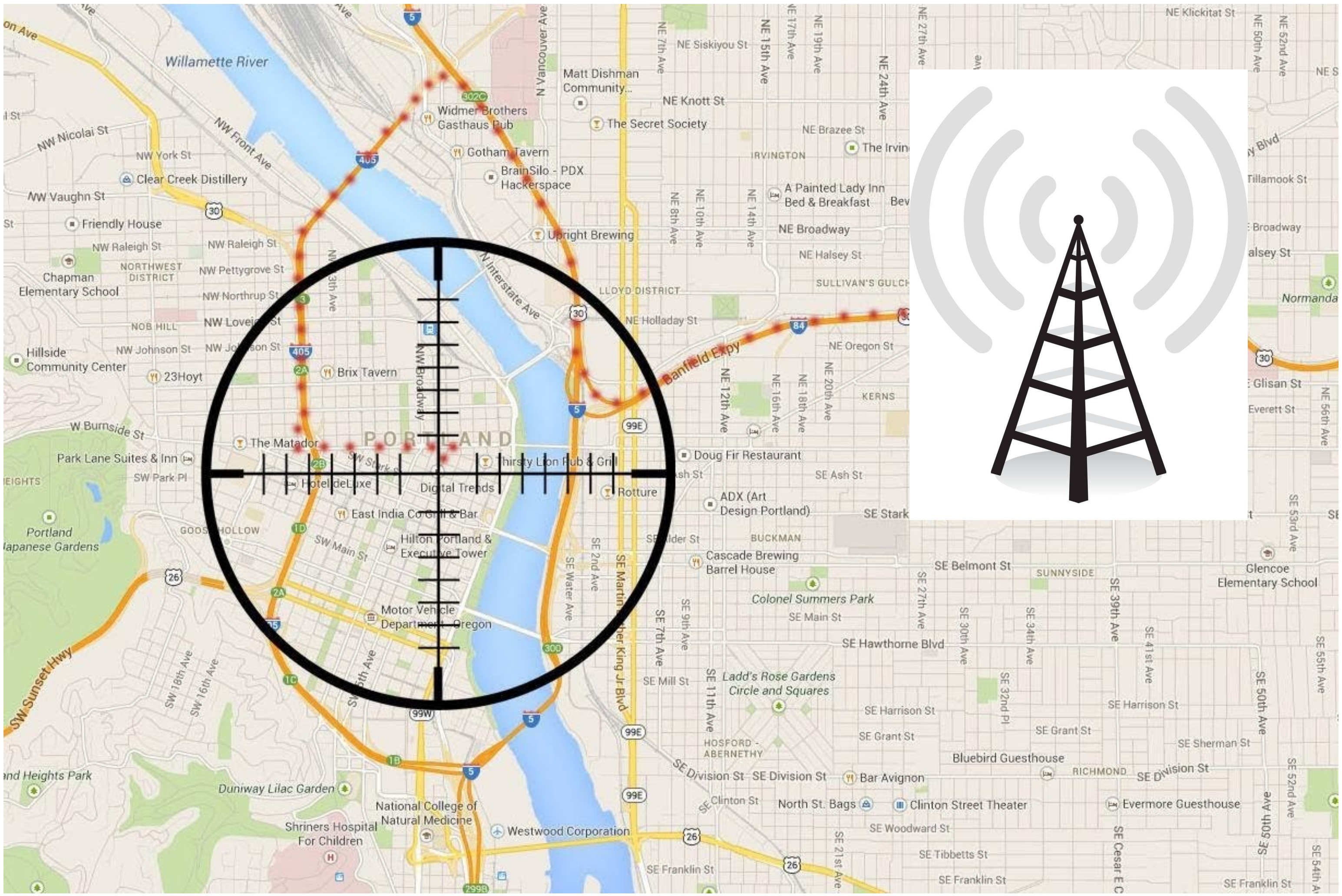
note:

$$\mathrm{Geo}(p) = \mathrm{NegBin}(1, p)$$

A grid of random variables

	number of **successes**	**time** to get successes	
One trial	$X \sim \mathrm{Ber}(p)$	$X \sim \mathrm{Geo}(p)$	One success
	↑ $n = 1$	↑ $r = 1$	
Several trials	$X \sim \mathrm{Bin}(n, p)$	$X \sim \mathrm{NegBin}(r, p)$	Several successes
Interval of time	$X \sim \mathrm{Poi}(\lambda)$	$X \sim \mathrm{Exp}(\lambda)$ (today!)	One success after interval of time

Not all values are discrete



Continuous random variables

A **continuous** random variable has a value that's a **real number** (not necessarily an integer).

Replace sums with **integrals**!

$$P(a < X \leq b) = F_X(b) - F_X(a)$$

$$F_X(a) = \int_{x=-\infty}^{a} dx \;\; f_X(x)$$

```
random.random()
```

Questions about `random`

```
X = random.random()
```

$$P(X = 0.6) = 0$$

$$P(X < 0.6) = 0.6$$

$$P(X \leq 0.6) = 0.6$$

$$P(X > 0.6) = 1 - 0.6 = 0.4$$

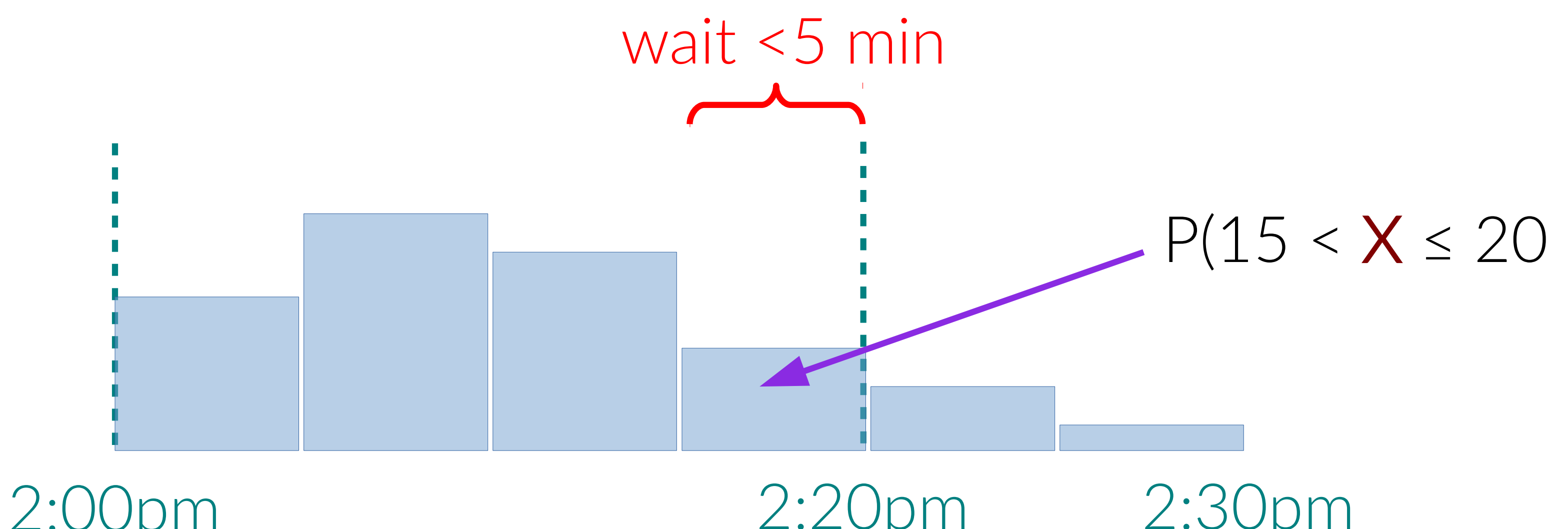
$$P(0.4 < X < 0.6) = 0.6 - 0.4 = 0.2$$

Riding the Marguerite

Marguerite stops at 20-minute intervals (2:00, 2:20, 2:40, ...).

You show up some time between 2:00 and 2:30.

What is P(wait < 5 minutes)?



X: number of minutes past 2pm

image: Haha169

Riding the Marguerite

Marguerite stops at 20-minute intervals (2:00, 2:20, 2:40, ...).

You show up some time between 2:00 and 2:30.

What is P(wait < 5 minutes)?

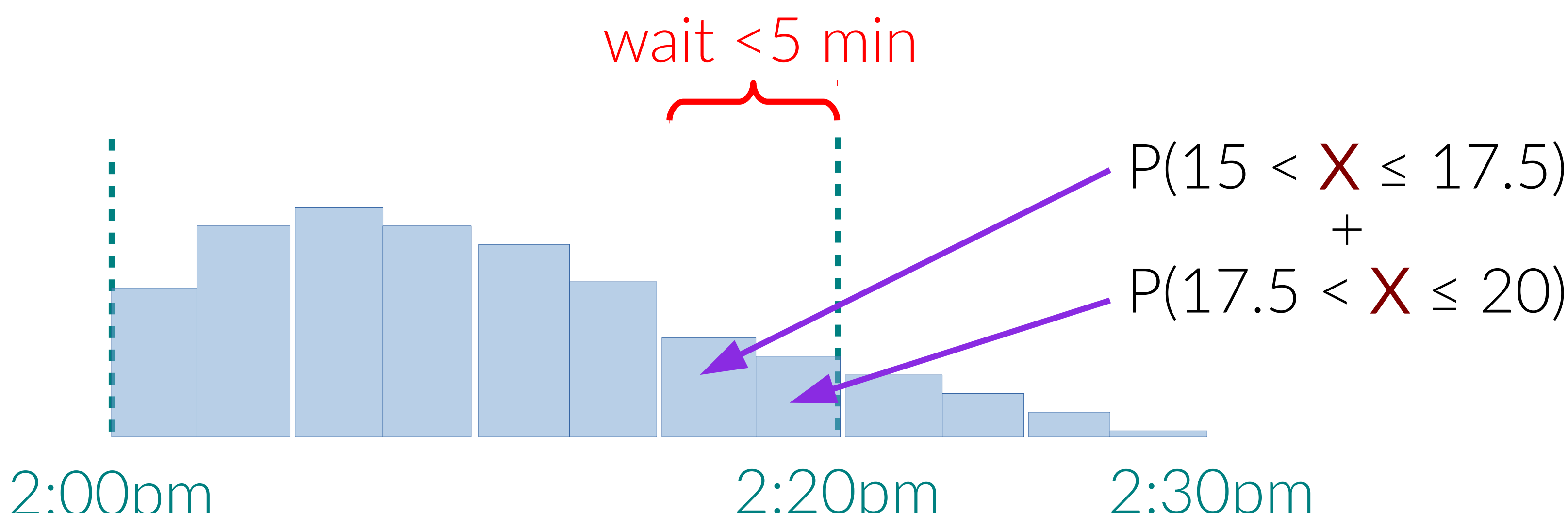


image: Haha169

Riding the Marguerite

Marguerite stops at 20-minute intervals (2:00, 2:20, 2:40, ...).

You show up some time between 2:00 and 2:30.

What is P(wait < 5 minutes)?

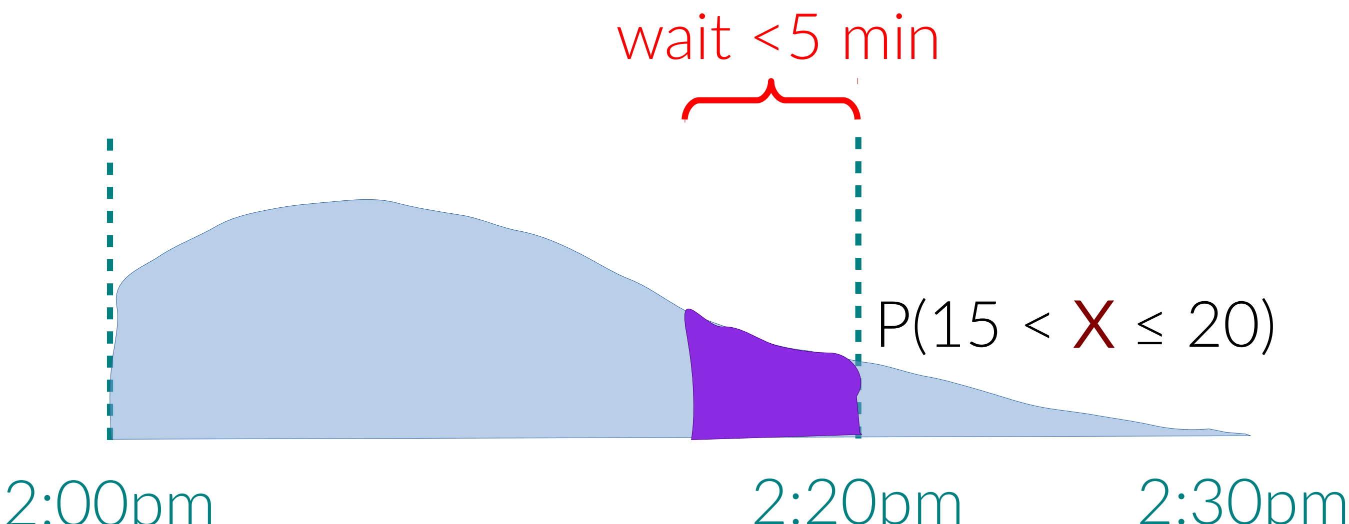


image: Haha169

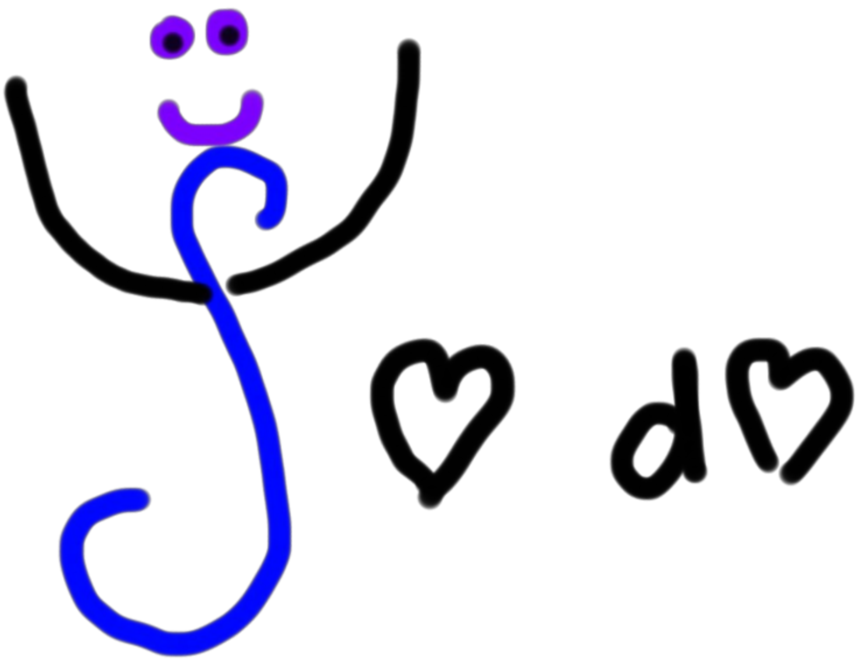
Probability density function

The **probability density function** (PDF) of a continuous random variable represents the relative likelihood of various values.

Units of probability *divided by units of X*.
Integrate it to get probabilities!

$$P(a < X \leq b) = \int_{x=a}^{b} dx \, \boxed{f_X(x)}$$

Integrals



(they just want to be your friend)

drawing by Chris Piech

Properties of PDFs

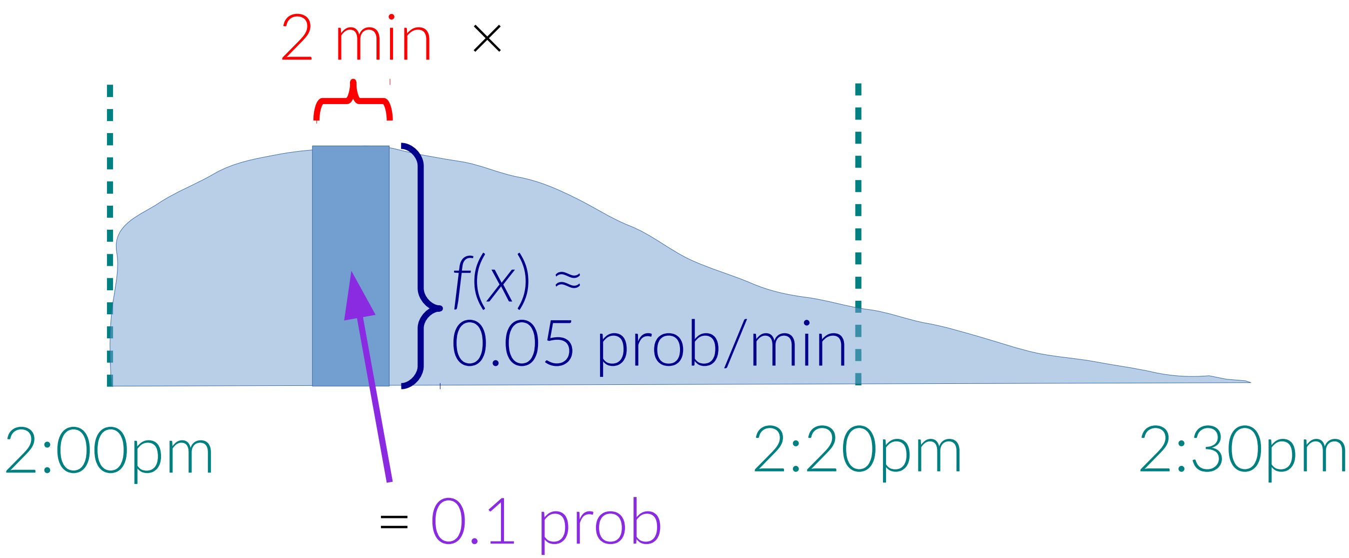
Integrals of PDFs must satisfy the axioms of probability:

$$0 \leq \int_{x=a}^{b} dx\, f_X(x) \leq 1$$

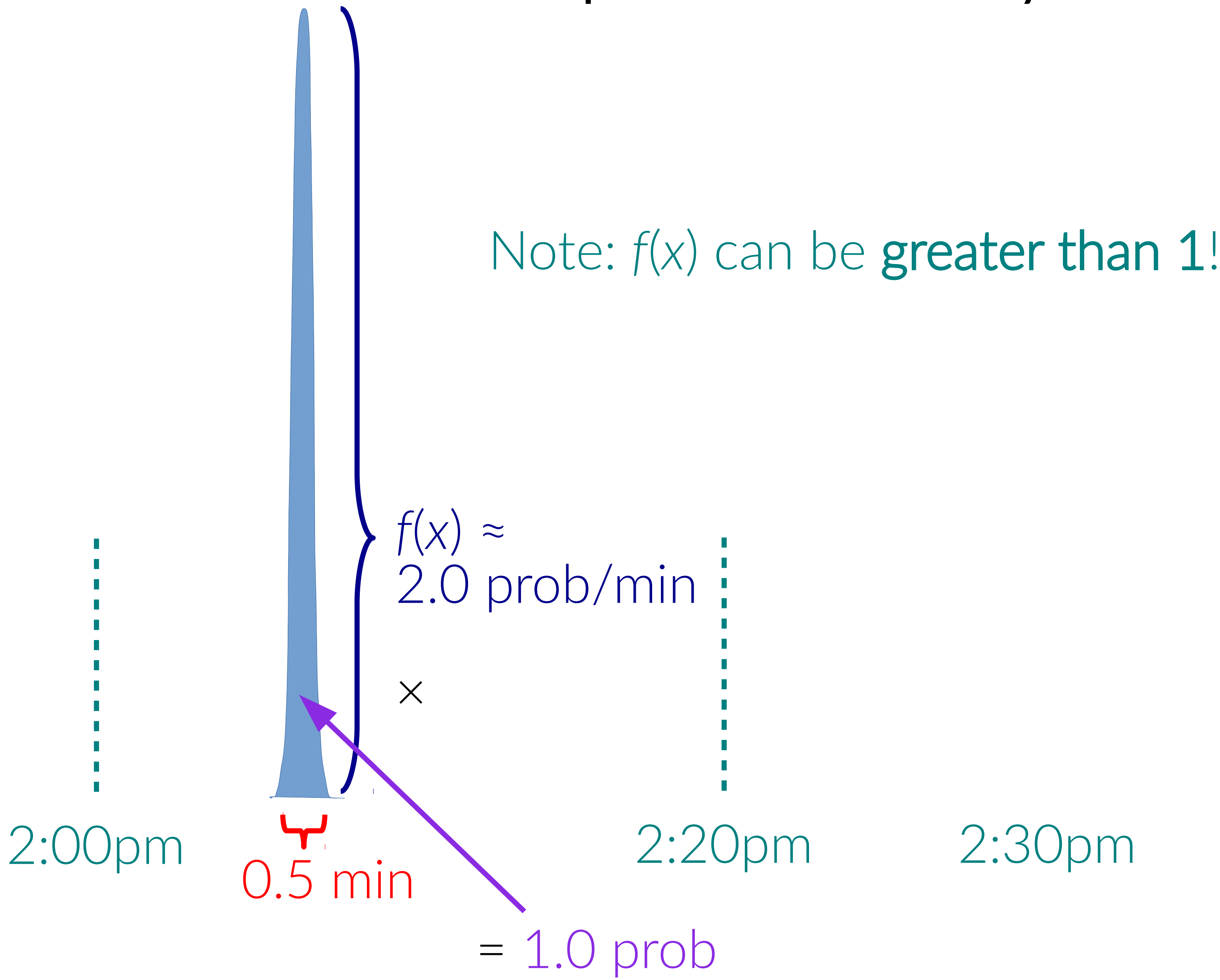
$$\int_{x=-\infty}^{\infty} dx\, f_X(x) = 1$$

$f(x)$ is not a probability

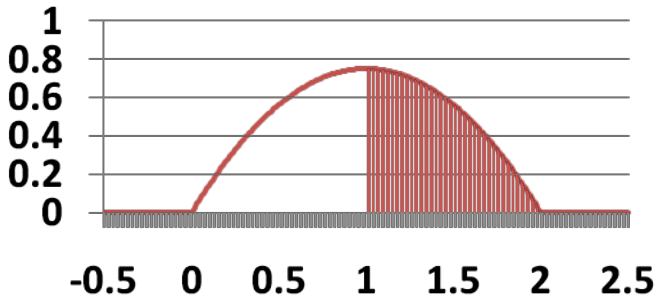
Rather, it has "units" of probability
divided by units of *X*.



$f(x)$ is not a probability



A simple example

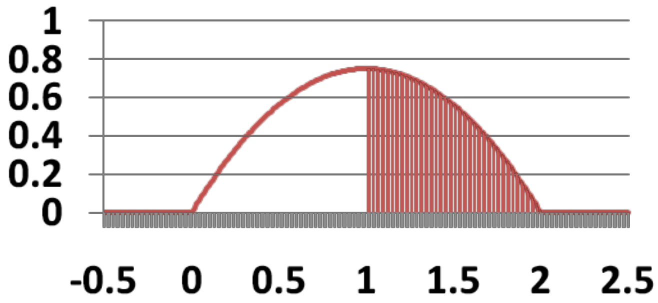


$$f_X(x) = \begin{cases} C(4x - 2x^2) & \text{if } 0 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

1) C = ?

2) P(X > 1) = ?

A simple example



$$f_X(x) = \begin{cases} C(4x - 2x^2) & \text{if } 0 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

1) C = ?

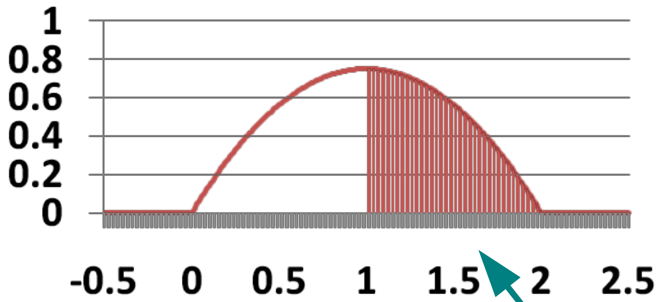
$$\int_{-\infty}^{\infty} dx\, f_X(x) = \int_0^2 dx\, f_X(x) = \int_0^2 dx\, C(4x - 2x^2)$$

$$= C\left[2x^2 - \frac{2}{3}x^3\right]_{x=0}^{2}$$

$$= C\left(2 \cdot 2^2 - \frac{2}{3}2^3\right) = \frac{8}{3}C = 1$$

$$C = \frac{3}{8}$$

A simple example



$$f_X(x) = \begin{cases} \frac{3}{8}(4x - 2x^2) & \text{if } 0 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

picture checks out!

2) P(X > 1) =

$$\int_1^{\infty} dx\, f_X(x) = \int_1^2 dx\, f_X(x) = \int_1^2 dx\, \frac{3}{8}(4x - 2x^2)$$

$$= \frac{3}{8}\left[2x^2 - \frac{2}{3}x^3\right]_{x=1}^{2}$$

$$= \frac{3}{8}\left[\left(2\cdot 2^2 - \frac{2}{3}2^3\right) - \left(2\cdot 1^2 - \frac{2}{3}1^3\right)\right]$$

$$= \frac{3}{8}\left[\frac{8}{3} - \frac{4}{3}\right] = \frac{1}{2}$$

Continuous expectation and variance

Remember: replace sums with **integrals**!

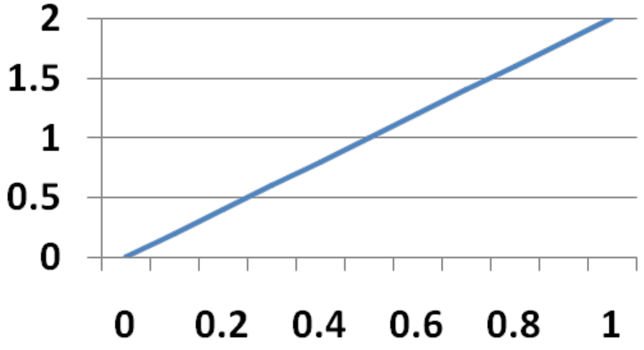
$$E[X] = \sum_{x=-\infty}^{\infty} x \cdot p_X(x) \longrightarrow E[X] = \int_{x=-\infty}^{\infty} dx\, x \cdot f_X(x)$$

$$E[X^2] = \sum_{x=-\infty}^{\infty} x^2 \cdot p_X(x) \longrightarrow E[X^2] = \int_{x=-\infty}^{\infty} dx\, x^2 \cdot f_X(x)$$

$$\text{Var}(X) = E[(X - E[X])^2] = E[X^2] - (E[X])^2$$

(still!)

Linearly increasing density

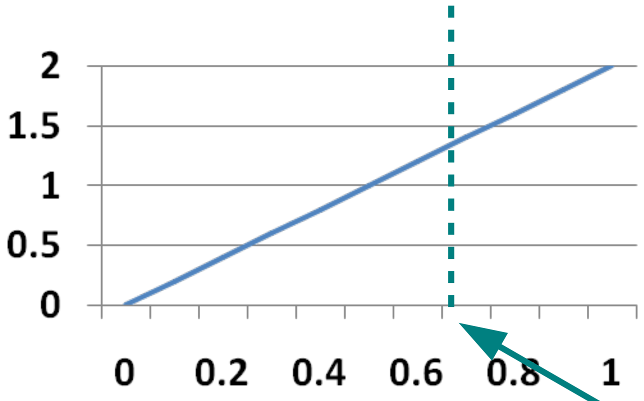


$$f_X(x) = \begin{cases} 2x & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

1) E[X] = ?

2) Var(X) = ?

Linearly increasing density



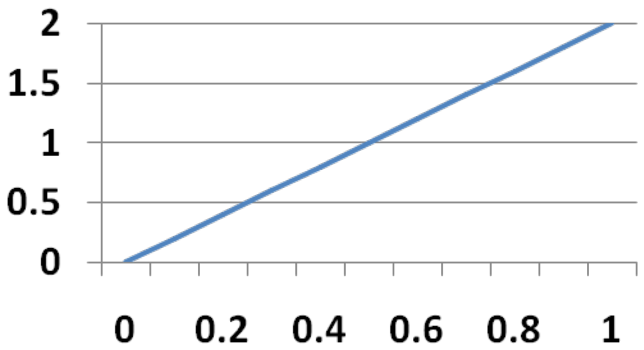
$$f_X(x) = \begin{cases} 2x & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

1) E[X] =

$$\int_{-\infty}^{\infty} dx\, x \cdot f_X(x) = \int_0^1 dx\, x \cdot f_X(x) = \int_0^1 dx\, x \cdot 2x$$

$$= \frac{2}{3} x^3 \Big|_{x=0}^{1} = \boxed{\frac{2}{3}}$$

Linearly increasing density



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

2) Var(X) = ?

$$E[X^2] = \int_{-\infty}^{\infty} dx\, x^2 \cdot f_X(x) = \int_0^1 dx\, x^2 \cdot f_X(x) = \int_0^1 dx\, x^2 \cdot 2x$$

$$= \frac{2}{4} x^4 \Big|_{x=0}^{1} = \frac{1}{2}$$

$$\mathrm{Var}(X) = E[X^2] - (E[X])^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \boxed{\frac{1}{18}}$$

Uniform random variable

A **uniform** random variable is **equally likely** to be any value in a single real number interval.

$$X \sim \text{Uni}(\alpha, \beta)$$

$$f_X(x) = \begin{cases} \dfrac{1}{\beta - \alpha} & \text{if } x \in [\alpha, \beta] \\ 0 & \text{otherwise} \end{cases}$$

Riding the Marguerite

Marguerite stops at 15-minute intervals (2:00, 2:15, 2:30, ...).

You show up some time between 2:00 and 2:30 **(all times equally likely)**.

What is P(wait < 5 minutes)?

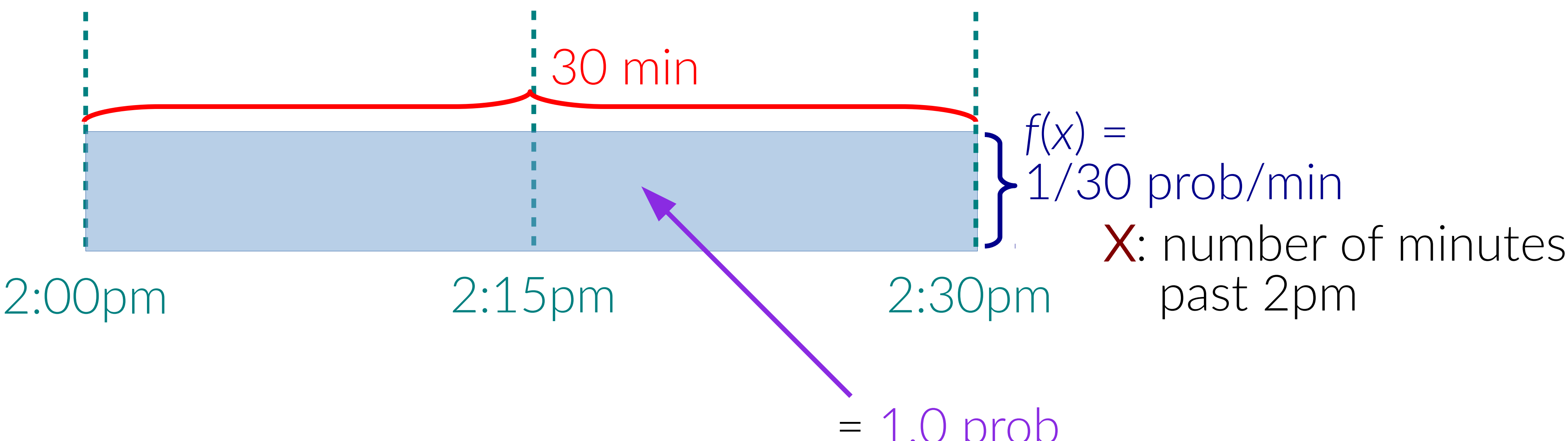


image: Haha169

Uniform: Fact sheet

minimum value

$$X \sim \text{Uni}(\alpha, \beta)$$

maximum value

PDF: $$f_X(x) = \begin{cases} \dfrac{1}{\beta - \alpha} & \text{if } x \in [\alpha, \beta] \\ 0 & \text{otherwise} \end{cases}$$

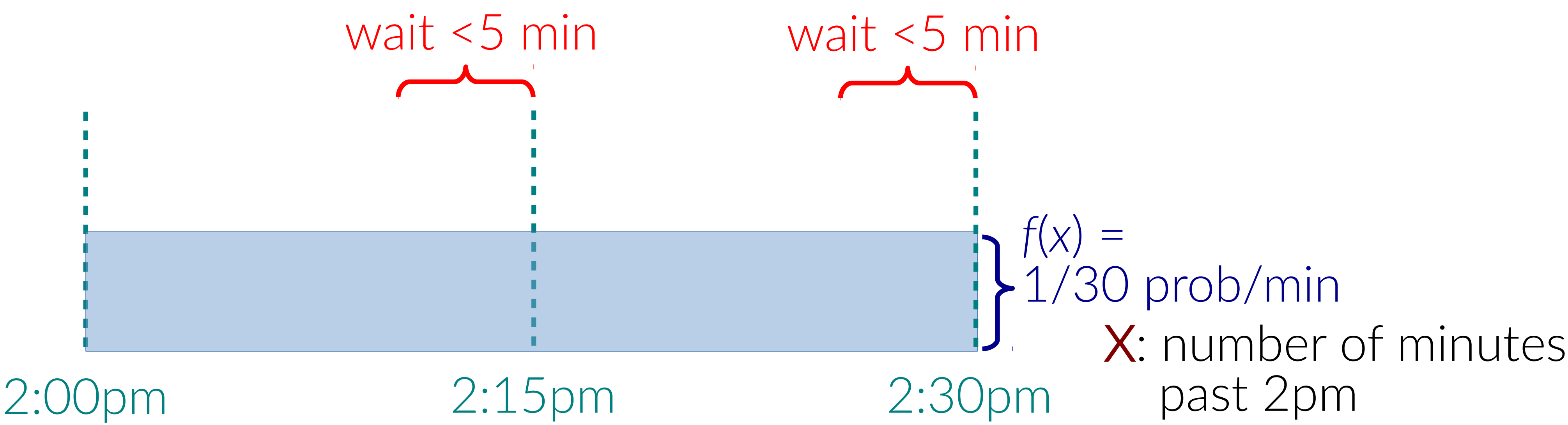
image: Haha169

Riding the Marguerite

Marguerite stops at 15-minute intervals (2:00, 2:15, 2:30, …).

You show up some time between 2:00 and 2:30 **(all times equally likely)**.

What is P(wait < 5 minutes)?



$$P(10<X<15)+P(25<X<30)=\int_{10}^{15} dx\,\frac{1}{30}+\int_{25}^{30} dx\,\frac{1}{30}=\frac{5}{30}+\frac{5}{30}=\boxed{\frac{1}{3}}$$

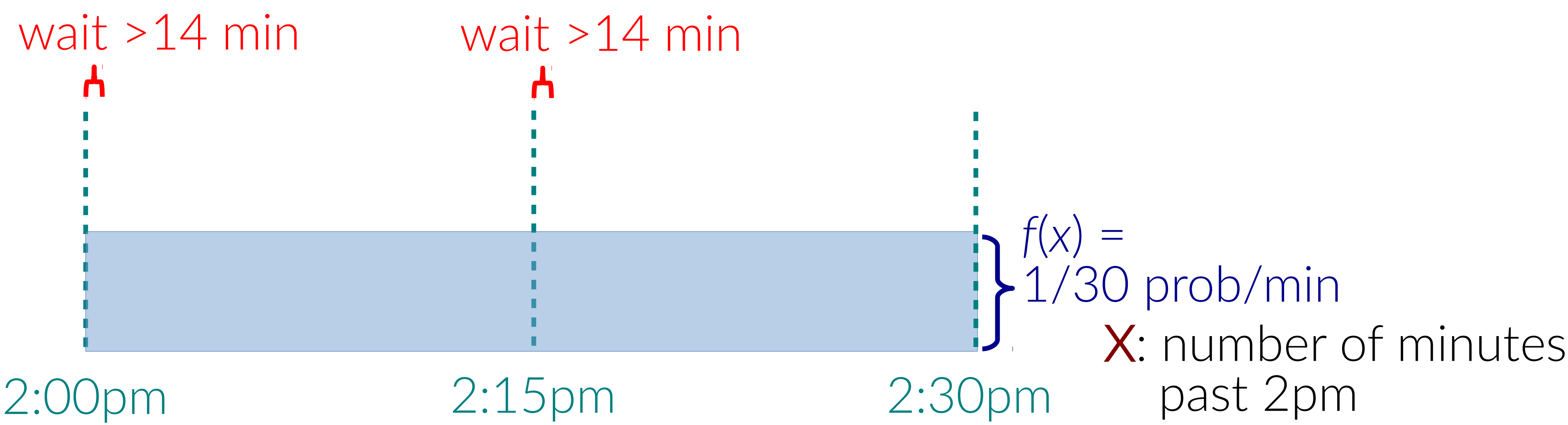
image: Haha169

Riding the Marguerite

Marguerite stops at 15-minute intervals (2:00, 2:15, 2:30, …).

You show up some time between 2:00 and 2:30 **(all times equally likely)**.

What is P(wait > 14 minutes)?



$$P(0<X<1)+P(15<X<16)=\int_0^1 dx\,\frac{1}{30}+\int_{15}^{16} dx\,\frac{1}{30}=\frac{1}{30}+\frac{1}{30}=\boxed{\frac{1}{15}}$$

image: Haha169

Uniform: Fact sheet

minimum value

$$X \sim \text{Uni}(\alpha, \beta)$$

maximum value

PDF: $$f_X(x) = \begin{cases} \dfrac{1}{\beta - \alpha} & \text{if } x \in [\alpha, \beta] \\ 0 & \text{otherwise} \end{cases}$$

CDF: $$F_X(x) = \begin{cases} \dfrac{x - \alpha}{\beta - \alpha} & \text{if } x \in [\alpha, \beta] \\ 1 & \text{if } x > \beta \\ 0 & \text{otherwise} \end{cases}$$

image: Haha169

Break time!

Expectation of uniform

$$X \sim \text{Uni}(\alpha, \beta)$$

$$E[X] = \int_{-\infty}^{\infty} dx\, x \cdot f_X(x) = \int_{\alpha}^{\beta} dx\, x \cdot \frac{1}{\beta - \alpha}$$

$$= \frac{1}{\beta - \alpha} \cdot \frac{1}{2} x^2 \Big|_{x=\alpha}^{\beta}$$

$$= \frac{1}{\beta - \alpha} \cdot \frac{1}{2} (\beta^2 - \alpha^2)$$

$$= \frac{1}{2} \cdot \frac{(\beta + \alpha)(\beta - \alpha)}{\beta - \alpha} = \boxed{\frac{1}{2}(\alpha + \beta)}$$

just average the start and end!

Uniform: Fact sheet

$$X \sim \text{Uni}(\alpha, \beta)$$

minimum value (α)

maximum value (β)

PDF: $$f_X(x) = \begin{cases} \frac{1}{\beta - \alpha} & \text{if } x \in [\alpha, \beta] \\ 0 & \text{otherwise} \end{cases}$$

CDF: $$F_X(x) = \begin{cases} \frac{x - \alpha}{\beta - \alpha} & \text{if } x \in [\alpha, \beta] \\ 1 & \text{if } x > \beta \\ 0 & \text{otherwise} \end{cases}$$

expectation: $$E[X] = \frac{\alpha + \beta}{2}$$

variance: $$\text{Var}(X) = \frac{(\beta - \alpha)^2}{12}$$

image: Haha169

Exponential random variable

An **exponential** random variable is the **amount of time until the first event** when events occur as in the Poisson distribution.

$$X \sim \mathrm{Exp}(\lambda)$$

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

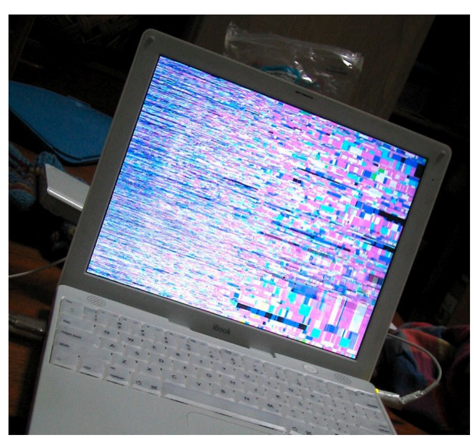


image: Adrian Sampson

Exponential: Fact sheet

rate of events per unit time

$$X \sim \text{Exp}(\lambda)$$

time until first event

PDF: $$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

image: Adrian Sampson

CDF of exponential

$$X \sim \mathrm{Exp}(\lambda)$$

$$F_X(a) = P(X \le a) = \int_0^a dx\, \lambda e^{-\lambda x}$$

$$= \lambda \frac{1}{-\lambda} e^{-\lambda x} \Big|_{x=0}^{a}$$

$$= -1\left(e^{-\lambda a} - e^{-\lambda(0)}\right)$$

$$= -1\left(e^{-\lambda a} - 1\right)$$

$$= 1 - e^{-\lambda a}$$

Exponential: Fact sheet

rate of events per unit time

$$X \sim \text{Exp}(\lambda)$$

time until first event

PDF: $$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

CDF: $$F_X(x) = \begin{cases} 1 - e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

image: Adrian Sampson

Laptop crashes

On average, a laptop dies after 5000 hours of use.

You use your laptop 5 hrs/day.

What is the probability the laptop will last 4 years?

X: lifetime of laptop (hours)

$$X \sim \text{Exp}\left(\frac{1}{5000}\right)$$

$$\begin{aligned} P(X > 4 \cdot 365 \cdot 5 = 7300) &= 1 - P(X \leq 7300) \\ &= 1 - F_X(7300) \\ &= 1 - \left(1 - e^{-7300/5000}\right) \\ &= e^{-1.46} \approx 0.2322 \end{aligned}$$

image: Adrian Sampson

Exponential: Fact sheet

rate of events per unit time

$$X \sim \mathrm{Exp}(\lambda)$$

time until first event

PDF: $$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

CDF: $$F_X(x) = \begin{cases} 1 - e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

expectation: $$E[X] = \frac{1}{\lambda}$$

variance: $$\mathrm{Var}(X) = \frac{1}{\lambda^2}$$

image: Adrian Sampson

Exponential and conditioning

$$X \sim \text{Exp}(\lambda)$$

$$P(X > t \mid X > s) = \frac{P(X > t, X > s)}{P(X > s)}$$

$$= \frac{P(X > t)}{P(X > s)}$$

$$= \frac{1 - \left(1 - e^{-\lambda \cdot t}\right)}{1 - \left(1 - e^{-\lambda \cdot s}\right)}$$

$$= \frac{e^{-t\lambda}}{e^{-s\lambda}} = e^{-(t-s)\lambda} = P(X > (t-s))$$

what happened in the past doesn't matter!
exponential is "memoryless"