STAT 515 fa 2023 Lec 06 slides

Continuous random variables

Karl B. Gregory

University of South Carolina

These slides are an instructional aid; their sole purpose is to display, during the lecture, definitions, plots, results, etc. which take too much time to write by hand on the blackboard. They are not intended to explain or expound on any material.

random variable

all the values it could take.

Recall: A random variable is a *continuous* rv if its support is an interval[1].

Examples:

1. Let X be the waiting time until a bus arrives.
2. Let Y be the time until you drop your new phone.
3. Let Z be the body temperature of a randomly selected student.

How do we assign probabilities to the values a continuous rv can take?

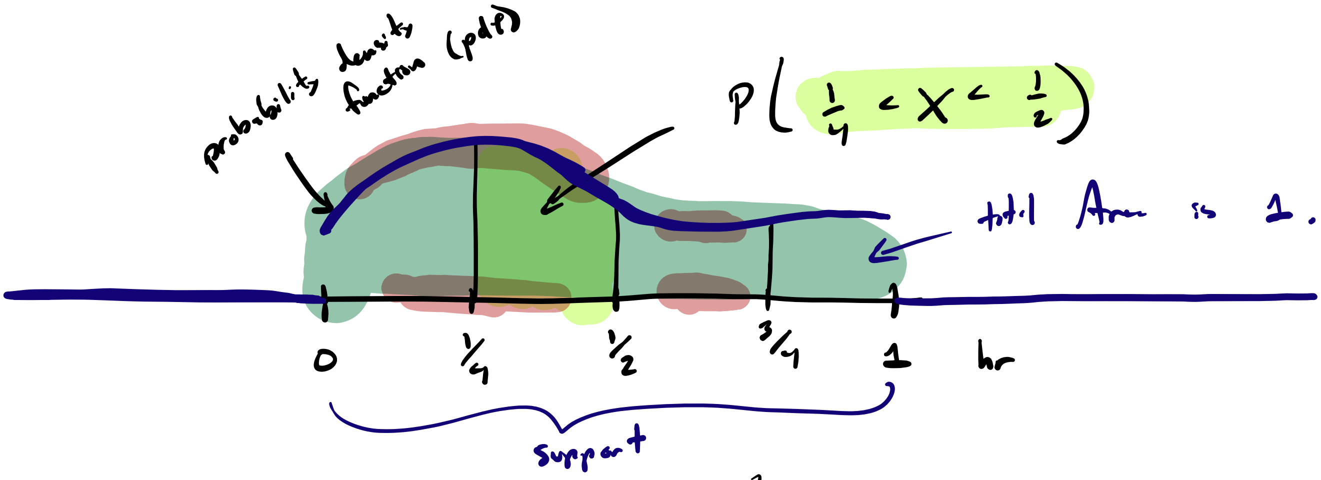
- Cannot list vals $x_1, x_2, x_3, \ldots$ and assign to them $p_1, p_2, p_3, \ldots$ as for discrete.
- Cannot represent the distribution with a table as we have done...

x	x_1	x_2	$\ldots$	x_k
$P(X = x)$	p_1	p_2		p_k

[1] or a union of intervals

X = Waiting time for a bus that comes every hr.

$P\left(\frac{1}{4} < X < \frac{1}{2}\right)$



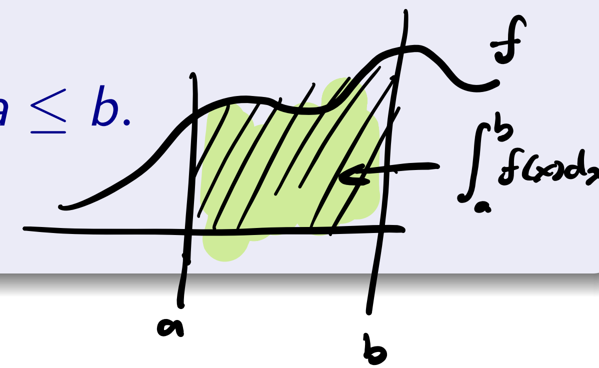
$P\left(X = \frac{3}{4}\right) = 0.$

$P\left(0 \le X \le 1\right) = 1$

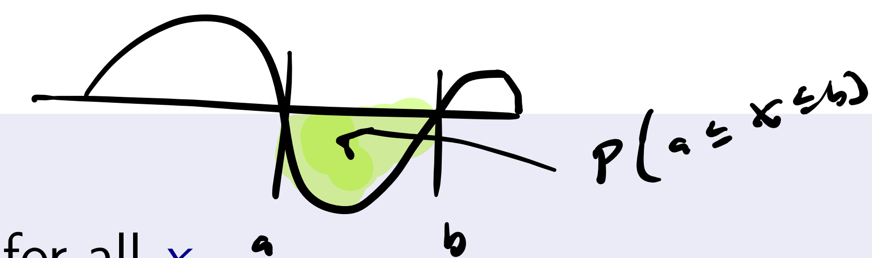
Probability density function of a continuous rv

The *probability density function (pdf)* of a cont. rv X is the function f satisfying

$$P(a \leq X \leq b) = \int_a^b f(x)dx \quad \text{for all } a \leq b.$$



Note: $\int_a^b f(x)dx =$ Area between f and the horizontal axis on the interval $[a, b]$.



Every pdf f has these two properties:

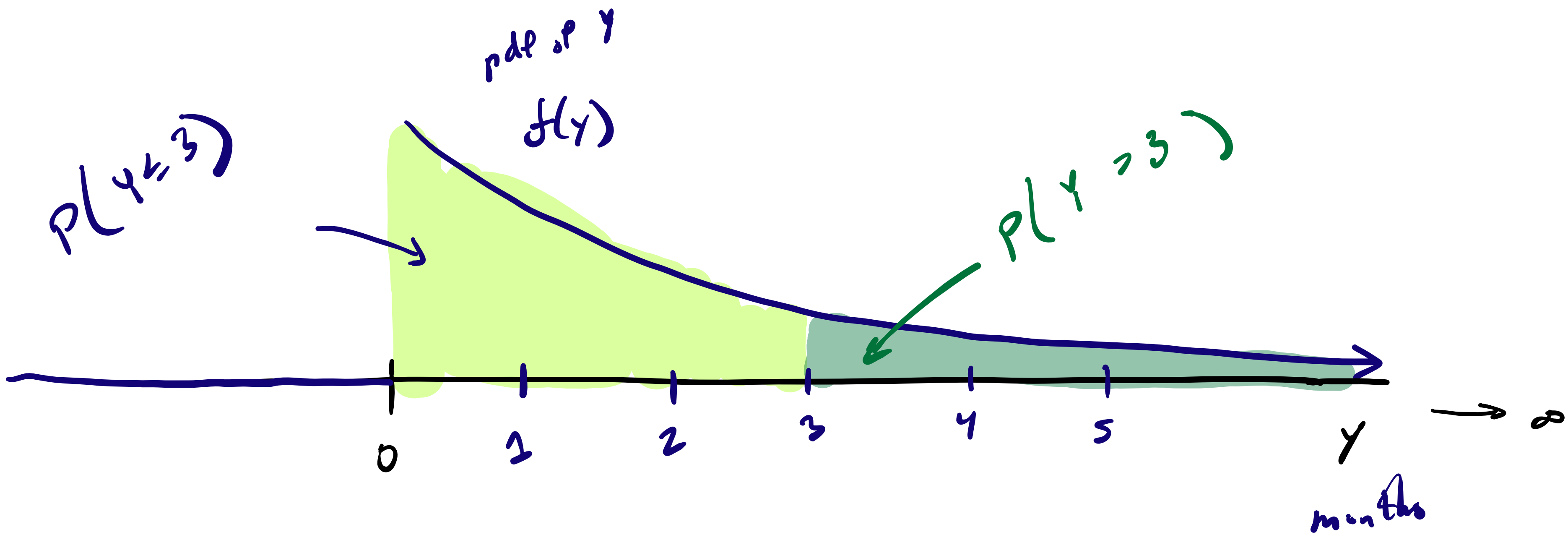
1. It is always non-negative, i.e. $f(x) \geq 0$ for all x.
2. The total area between f and the horizontal axis is 1, i.e. $\int_{-\infty}^{\infty} f(x)dx = 1$.

If X is an rv with pdf f, then we write $X \sim f$.

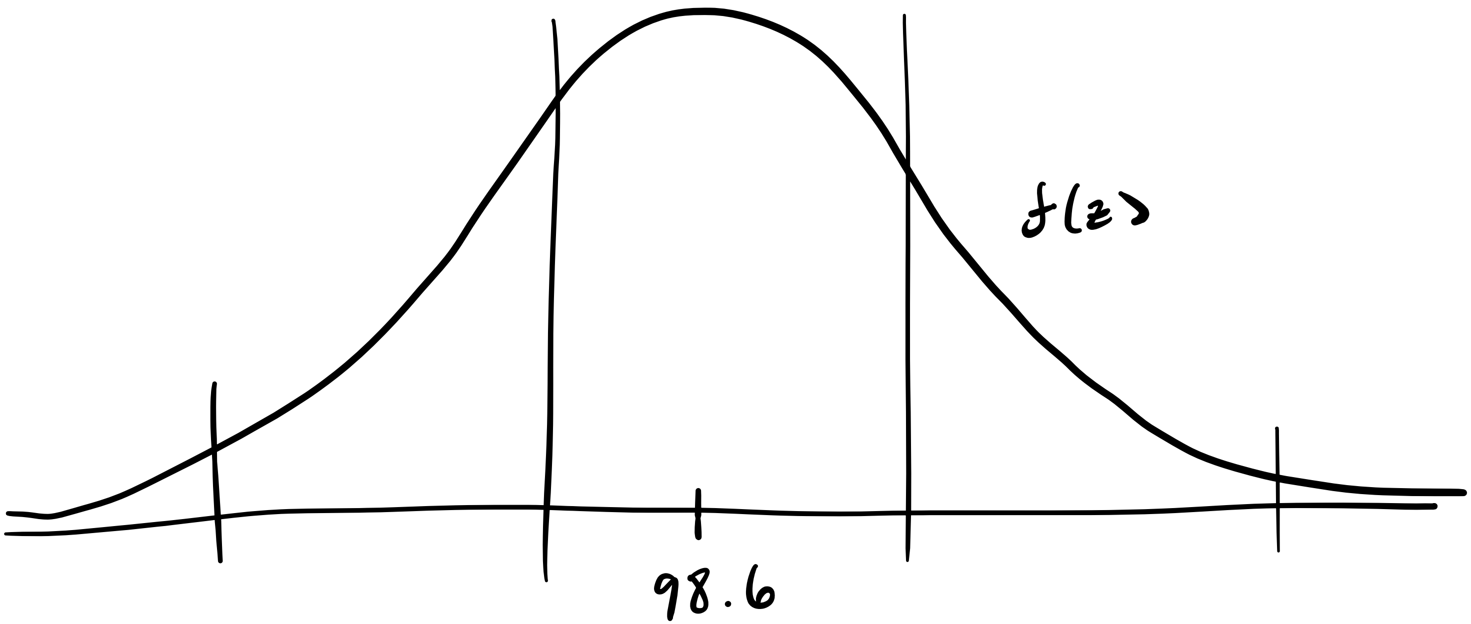
Discuss: Let X = wait time for hourly arriving bus.

$\mathcal{Y} = [0, \infty)$

Discuss: Let Y be the time until you drop your new phone.



Discuss: Let Z be the body temperature of a randomly selected student.



Point probabilities are equal to zero for continuous rvs

For a continuous rv X, for any value c we have

$$\boxed{P(X = c) = 0,}$$

and, as a consequence, for any values a and b such that $a < b$, we have

$$\boxed{P(a < X < b) = P(a \leq X \leq b)} = P(a < X \leq b) = P(a \leq X < b).$$

$P(X=a) + P(a<X<b) + P(X=b)$

Exercise: Let X = mpg on next tank of gas.

1. What is $P(X = 24)$?
2. What is $P(X = 24.0000000000)$?
3. What is $P(24 < X < 25)$ versus $P(24 \leq X < 25)$?
4. Say you got $X = 24.2349022301$ last time. Did you observe a 0-prob event??
5. But how often will that value recur if you fill your tank again and again?

Cumulative distribution function

The *cumulative distribution function (cdf)* F of a rv X is the function F given by

$$F(x) = P(X \leq x) \quad \text{for all } x.$$

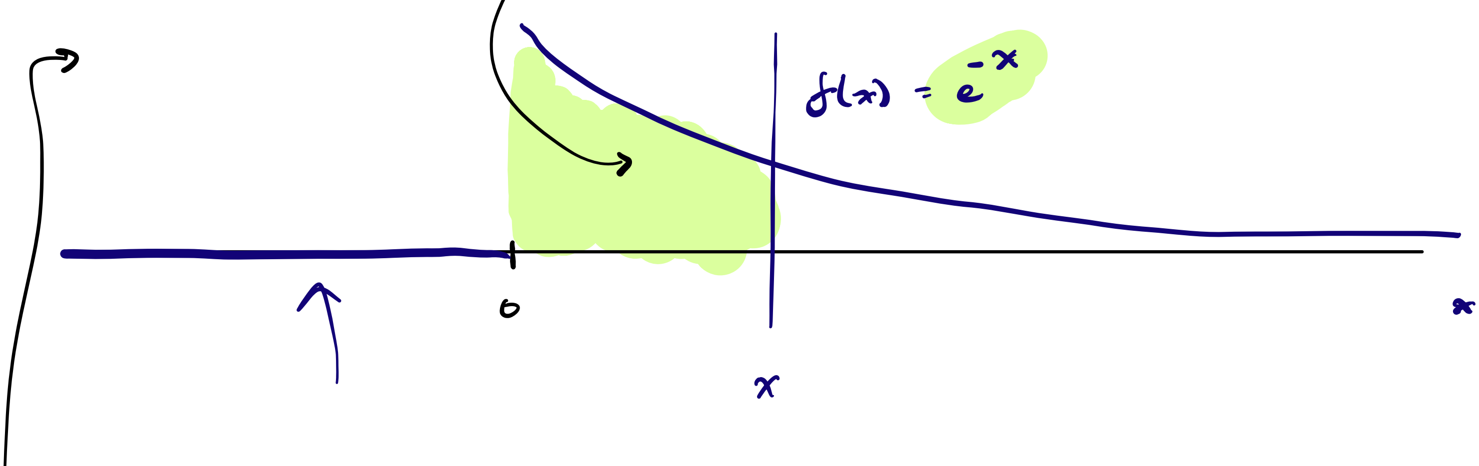
If X is an rv with cdf F, then we write $X \sim F$.

We have the following expressions:

$$F(x) = \begin{cases} \displaystyle\int_{-\infty}^{x} f(t)dt & \text{if } X \text{ is continuous with pdf } f \\ \displaystyle\sum_{t \in \mathcal{X} : t \leq x} p(t) & \text{if } X \text{ is discrete with support } \mathcal{X} \text{ and pmf } p. \end{cases}$$

Area under the pdf to the left of x.

$$F(x) = P(X \leq x) = \int_0^x e^{-t}\,dt = 1 - e^{-x}$$



Examples:

1. If X = time until you drop your phone, we might have

$$f(x) = \begin{cases} e^{-x}, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad \text{and} \quad F(x) = \begin{cases} 1 - e^{-x}, & x > 0 \\ 0, & x \leq 0. \end{cases}$$

Draw pictures.

2. If $X = \#$ jellyfish washed up on the beach, we might have

x	0	1	2	3	4	5	$\cdots$
$P(X = x)$	0.050	0.149	0.224	0.224	0.168	0.101	$\cdots$
$F(x) = P(X \leq x)$	0.050	0.199	0.423	0.647	0.815	0.916	$\cdots$

0.05 + 0.149

discrete

$\mathbb{E}X = p_1 x_1 + p_2 x_2 + \cdots$

$\mathrm{Var}\,X = p_1 (x_1 - \mu)^2 + p_2 (x_2 - \mu)^2 + \cdots$

Expected value and variance of a continuous rv

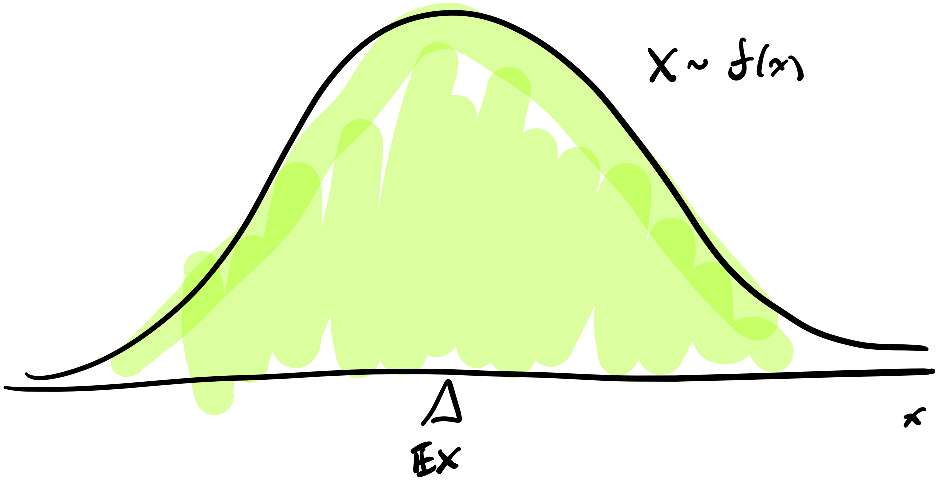
If $X \sim f(x)$ is a continuous rv, then

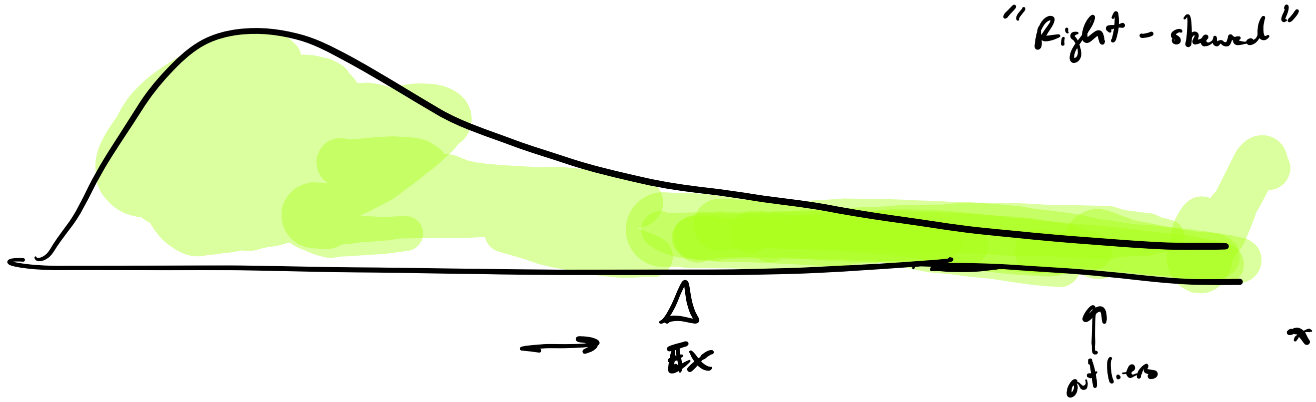
- $\mathbb{E}X = \int_{-\infty}^{\infty} x \cdot f(x) dx$
- $\mathrm{Var}\,X = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x) dx$, where $\mu = \mathbb{E}X$.

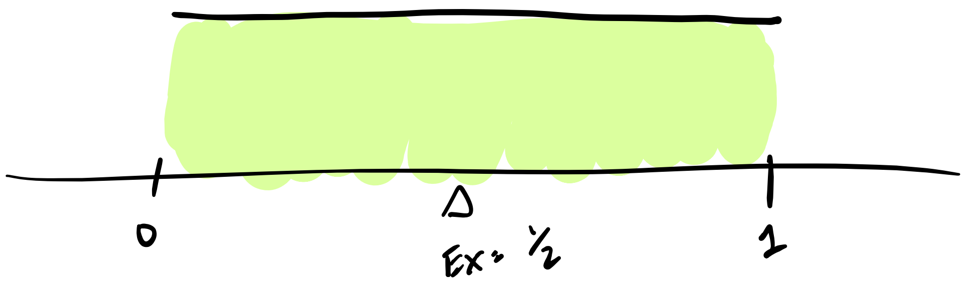
Need to use calculus to get these.

$\mathbb{E}X$ is the "balancing point" of the pdf (illustrate).

$\mathrm{Var}\,X$ describes spread.


X ~ f(x)
x
EX


"Right - skewed"
EX
outliers
x


0
EX = 1/2
1