OFFICE HOURS

- Instructor: Dr. Letao Zhang
  Office Hour: Tuesday 1:30pm-2:30pm
  Location: Math Learning Center (Math Tower Room S-235)

- TA: Mu Zhao (mu.zhao - at - stonybrook.edu)
  Office hour: Tuesday 4-5pm
  Location: Math Tower 2-122
  Office hour: Tuesday 11:00am-11:30am
  Location: Math Learning Center (Math Tower Room S-235)

MATH 331

PROBABILITY REVIEW

RECAP

Definition. A random variable is a mapping

$$X : \Omega \to \mathbb{R}$$

that assigns a real number $X(\omega)$ to each outcome ω.
- Ω is the sample space. Points
- ω in Ω are called sample outcomes, realizations, or elements.

- Note X only assigns ONE and only ONE real number to each element in the sample space Ω.
- The set of all possible values of the random variable X is called the support, or space, of X.

PDF

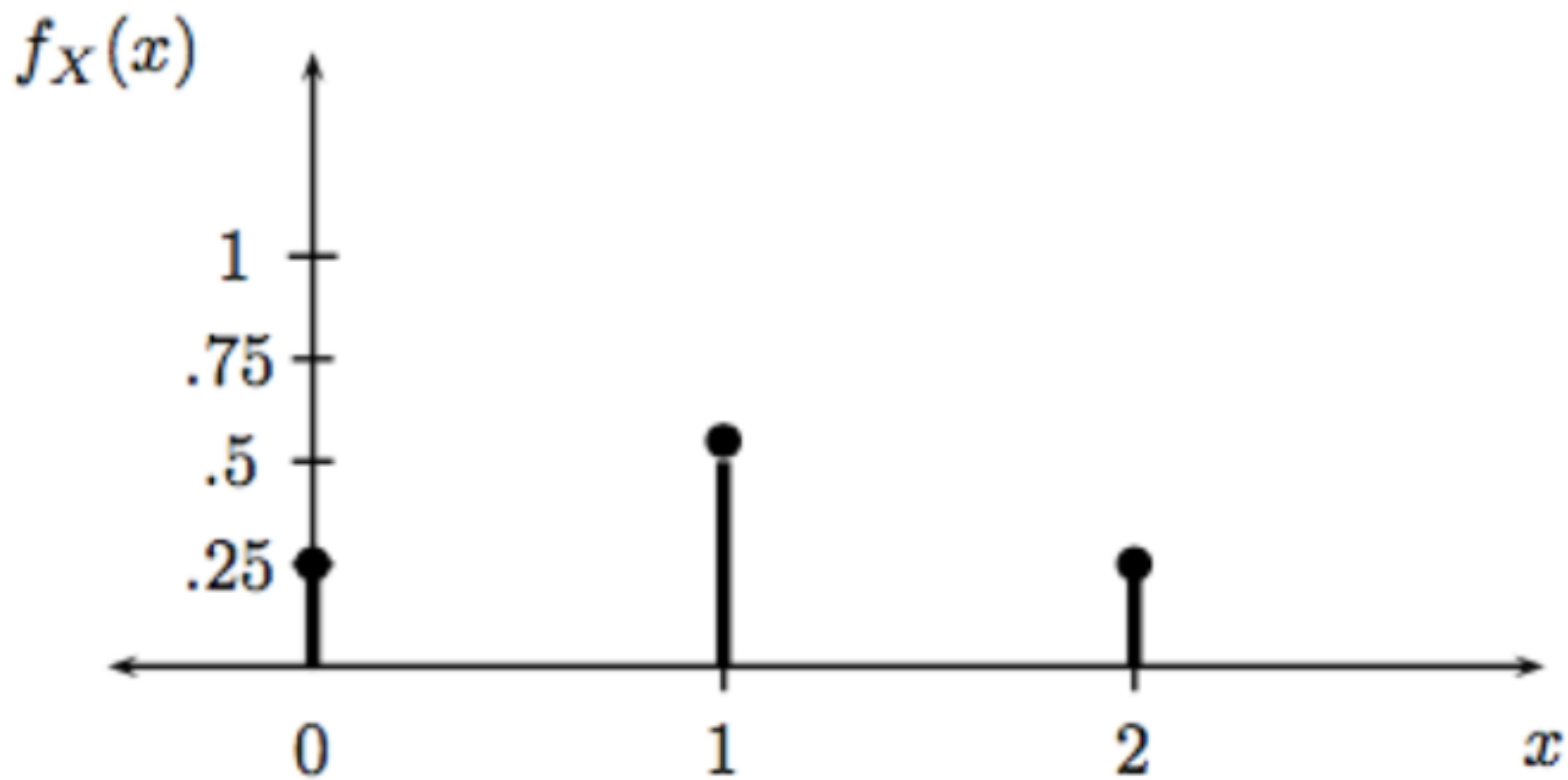
X is called a **discrete random variable** if it takes countably many values $\{x_1, x_2, \ldots\}$.

We define the **probability function** or the **probability mass function (pmf)** for X by:

$$f_X(x) = p(X = x)$$

PROBABILITY MASS AND DISTRIBUTION FUNCTION

The pmf for the number of heads in two coin tosses (taken from All of Stats) looks like this:



CUMULATIVE DISTRIBUTION FUNCTION

The **cumulative distribution function**, or the **CDF**, is a function

$F_X : \mathbb{R} \to [0, 1],$

defined by

$$F_X(x) = p(X \leq x).$$

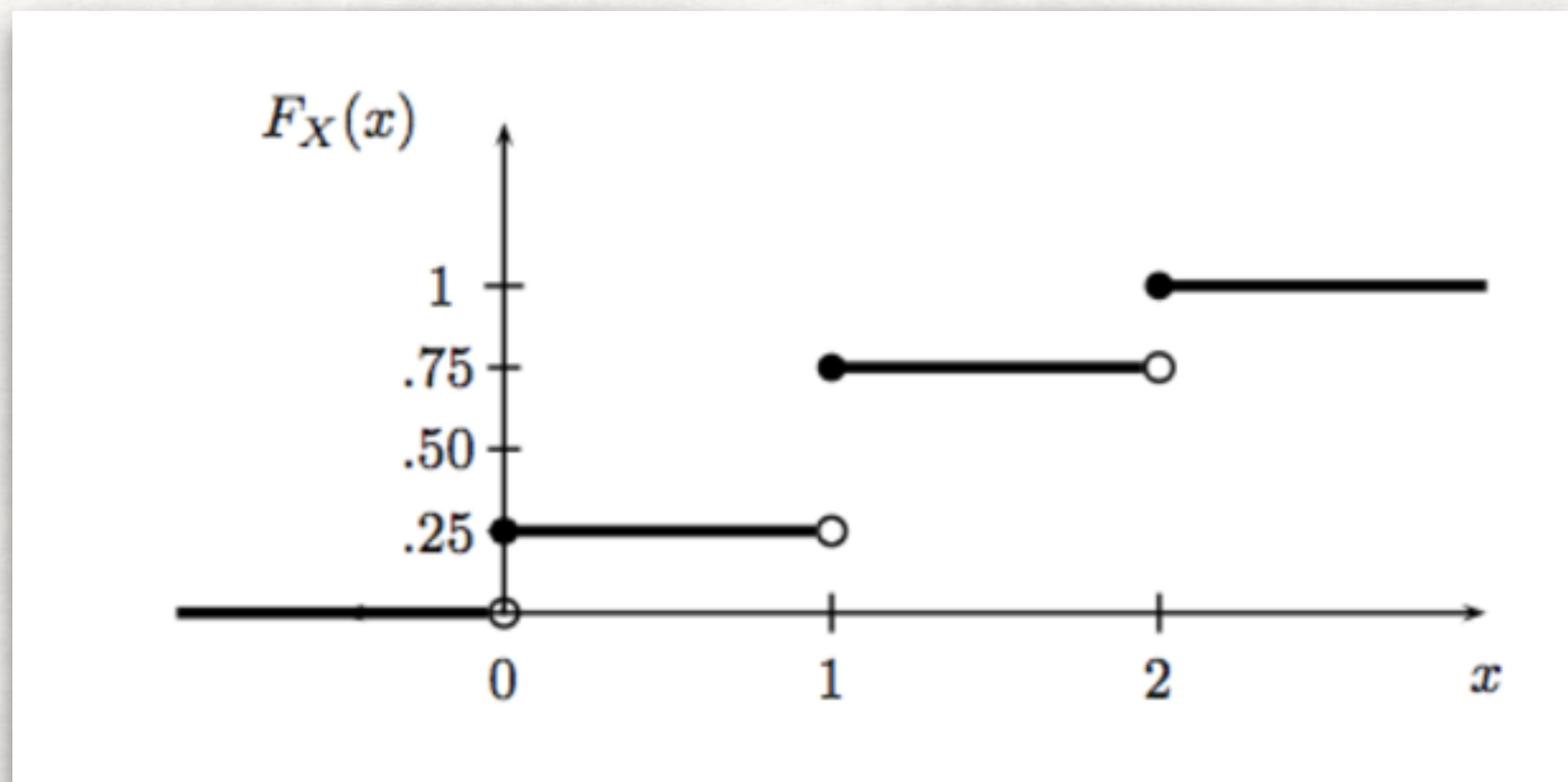
We also call this function: Distribution

CDF

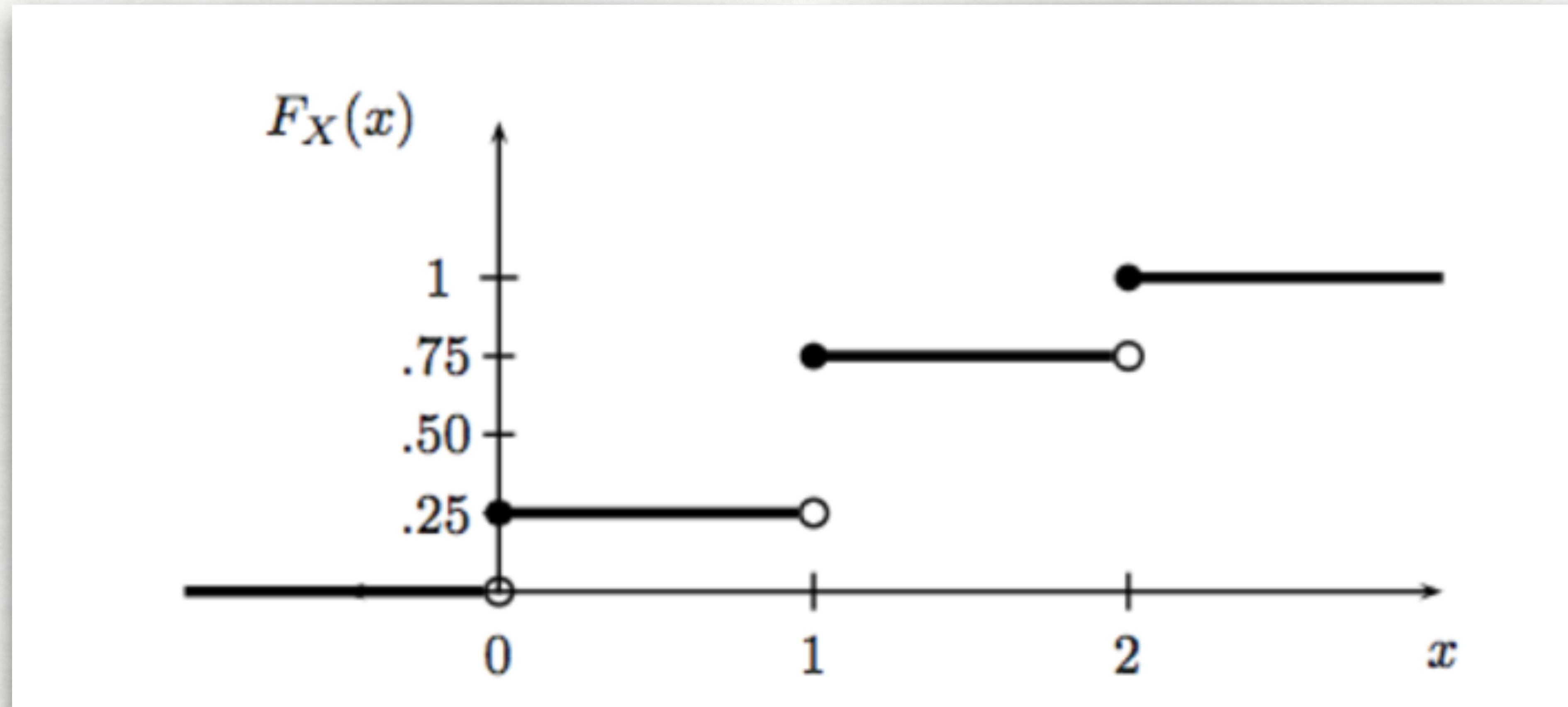
CUMULATIVE DISTRIBUTION FUNCTION

- Let X be the random variable representing the number of heads in two coin tosses.
  P(X = 2) = 1/4
  P(X<1.5) = P(X = 0) + P(X = 1) = 0.75

- CDF for this random variable:



CDF



- Note: This function is **right-continuous** and defined for all real numbers x.

RECAP

DISCRETE

The Rules of Probability

Sum rule

$$P(x) = \sum_y P(x, y)$$

Product rule

$$P(x, y) = P(y|x)P(x)$$

Bayes' theorem

$$P(y|x) = \frac{P(x|y)P(y)}{P(x)}$$

Denominator

$$P(x) = \sum_y P(x|y)P(y)$$

Bayes' theorem

$$P(x, y) = P(y|x)P(x)$$

likelihood

prior

$$P(y|x) = \frac{P(x|y)P(y)}{P(x)}$$

posterior

Prior – belief before making a particular obs.

Posterior – belief after making the obs.

Posterior is the prior for the next observation

– Intrinsically incremental

CONTINUOUS RANDOM VARIABLES

CONTINUOUS PROBABILITY DISTRIBUTION

On the other hand, a random variable is called a **continuous random variable** if there exists a function f_X such that $f_X(x) \geq 0$ for all x, $\int_{-\infty}^{\infty} f_X(x)dx = 1$ and for every a ≤ b,

$$p(a < X < b) = \int_a^b f_X(x)dx$$

The function f_X is called the probability density function (pdf). We have the CDF:

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt$$

CONTINUOUS PROBABILITY DISTRIBUTION

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt$$

$f_X(x) = \frac{dF_X(x)}{dx}$ at all points x at which F_X is differentiable.

- For continuous random variables, we use:
  Probability density function for f(x)
- For discrete random variables, we use:
  Probability MASS function for f(x)

DISCRETE VS CONTINUOUS DISTRIBUTIONS

- Discrete:
  1) The probability distribution is defined by a probability mass function (or simply probability function) denoted by f(x), which provides the probability for each value of the random variable
  2) Required conditions for discrete probability function are:

$$f(x) \geq 0$$

$$\Sigma f(x) = 1$$

- Continuous:
  1) It's not possible to talk about probability for each value of the random variable (or it is always zero). Instead we use intervals
  2) A probability distribution or probability density function (pdf) or (or simply probability function) of X satisfies for any a<b,

$$P(a \leq x \leq b) = \int_a^b f(x)dx \leq 1$$

A DISCRETE EXAMPLE: THE BERNOULLI DISTRIBUTION

- The Bernoulli Distribution represents the distribution a coin flip.
- Let the random variable X represent such a coin flip (like before), where X(H)=1, and X(T) = 0.
- Let us further say that the probability of heads is p (p=0.5 is a fair coin).
- In general, the number p is the success probability.

A DISCRETE EXAMPLE: THE BERNOULLI DISTRIBUTION

- The pmf or probability function associated with the Bernoulli distribution is:

$$f(x) = \begin{cases} 1-p & x = 0 \\ p & x = 1. \end{cases}$$

- p is between 0 and 1. This pmf, for x in the set {0,1}, is also written as:

$$f(x) = p^x(1-p)^{1-x}$$

- f(0) + f(1) = 1

- p is called a parameter of the Bernoulli distribution.

A DISCRETE EXAMPLE: THE BERNOULLI DISTRIBUTION

- We then write:

$$X \sim Bernoulli(p)$$

- which is to be read as "X has distribution Bernoulli(p)" or "X follows a Bernoulli distribution of probability p".

CONTINUOUS PROBABILITY DISTRIBUTION

UNIFORM DISTRIBUTION

- A continuous example: the Uniform Distribution
- Uniform Distribution is saying that:

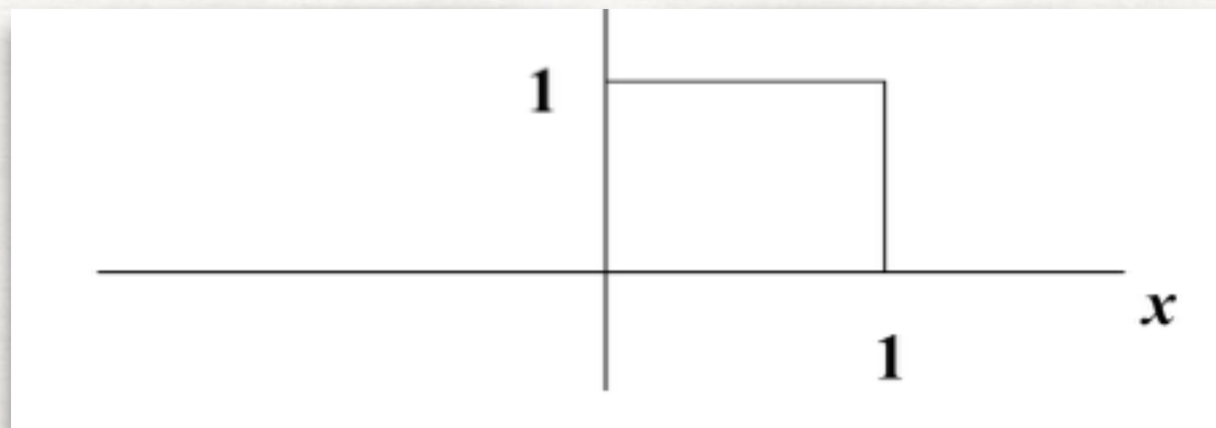
  Fix an interval [a,b] on the real line, any point show up inside the interval is equal likely.

CONTINUOUS PROBABILITY DISTRIBUTION

UNIFORM DISTRIBUTION

- The uniform distribution: all values are equally likely. The uniform distribution:

- $f(x) = 1$, for $1 \geq x \geq 0$
  $f(x) = 0$, elsewhere



- We can see it's a probability distribution because it integrates to 1 (the area under the curve is 1):

$$\int_0^1 1 = x \Big|_0^1 = 1 - 0 = 1$$

CONTINUOUS PROBABILITY DISTRIBUTION

UNIFORM DISTRIBUTION

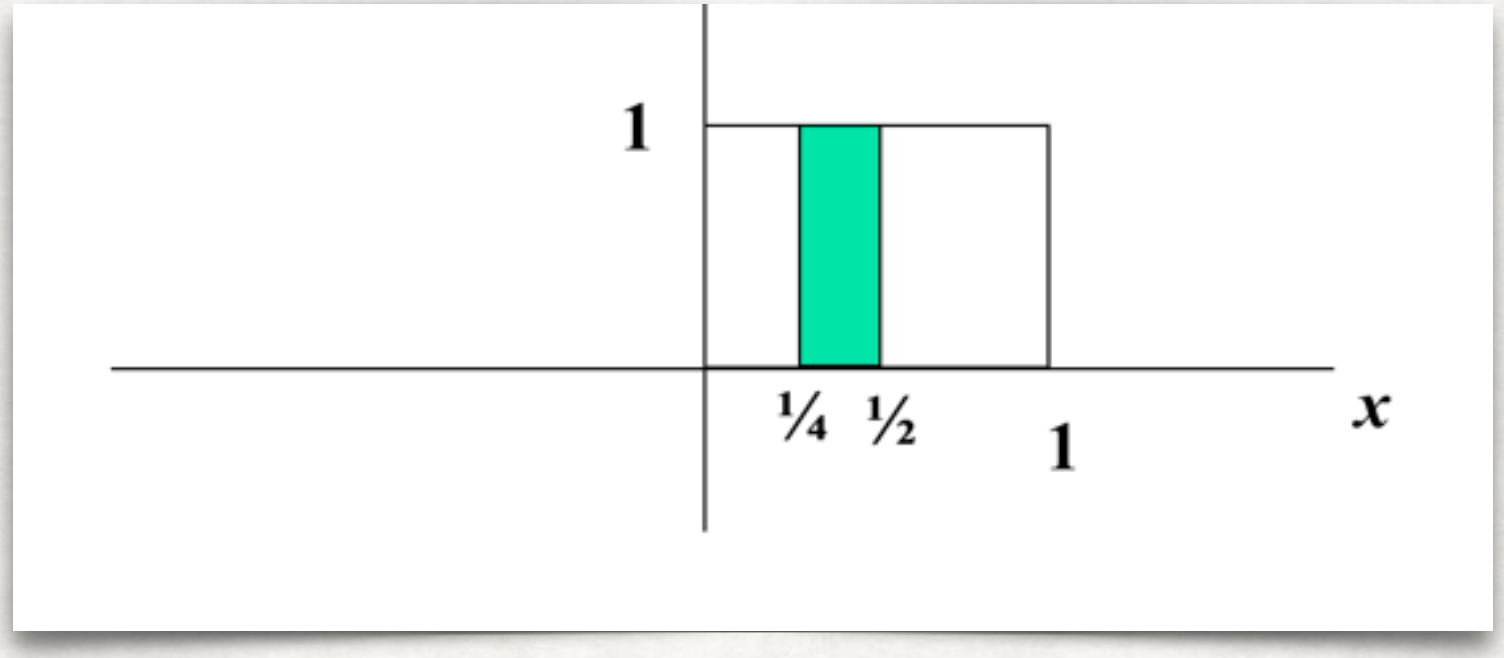
- What's the probability that x is between $\frac{1}{4}$ and $\frac{1}{2}$?

CONTINUOUS PROBABILITY DISTRIBUTION

UNIFORM DISTRIBUTION

- What's the probability that x is between ¼ and ½?

- P(½ ≥x≥ ¼)= ¼



CONTINUOUS PROBABILITY DISTRIBUTION

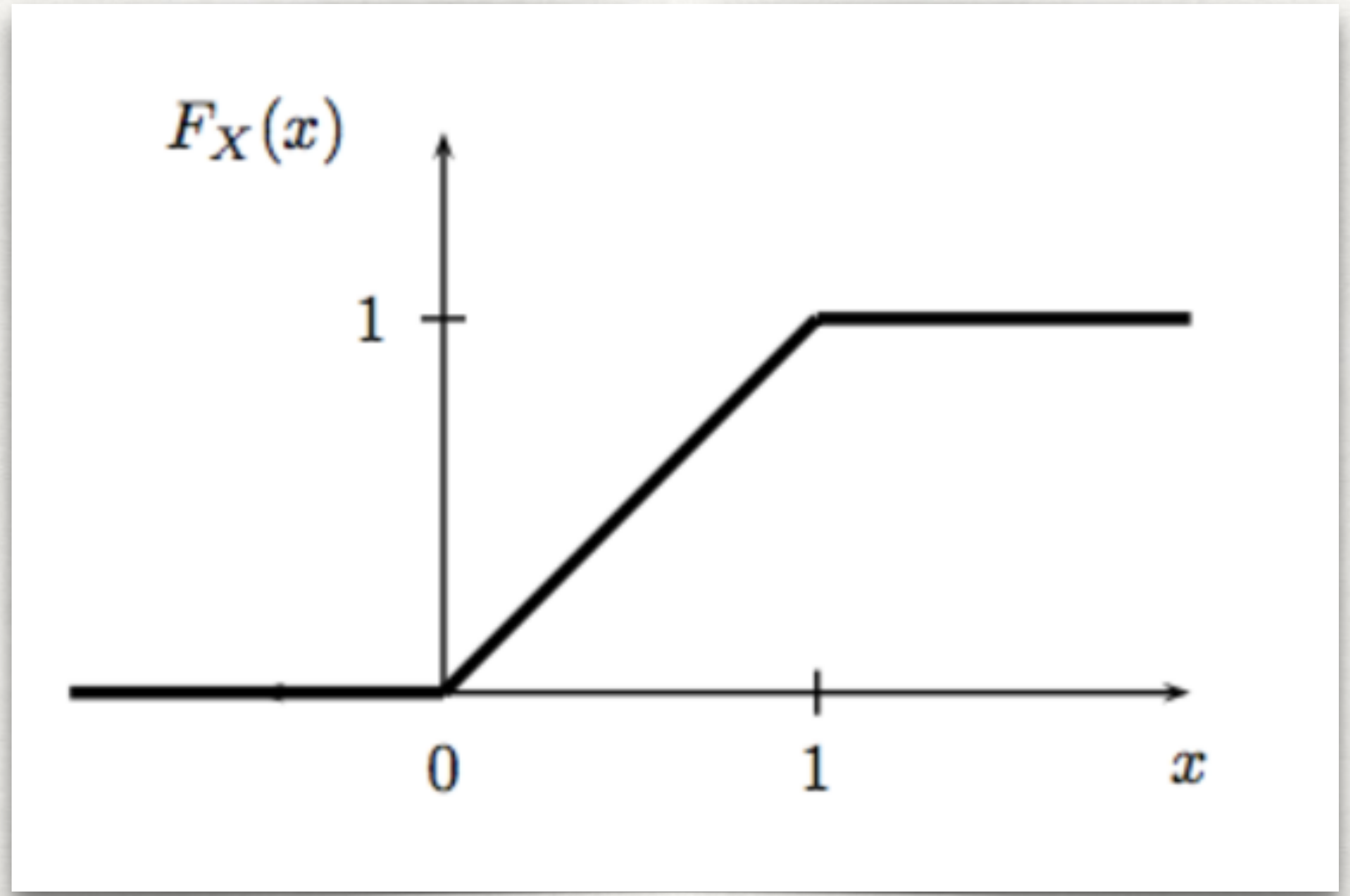
UNIFORM DISTRIBUTION

- What is the CDF?

CONTINUOUS PROBABILITY DISTRIBUTION

UNIFORM DISTRIBUTION

- The CDF is:



CONTINUOUS PROBABILITY DISTRIBUTION

EXPONENTIAL DISTRIBUTION

- For example, the negative exponential function with parameter 1 (in probability, this is called an "exponential distribution"):

- For x nonnegative: $f(x) = e^{-x}$
  For other x, f(x) = 0

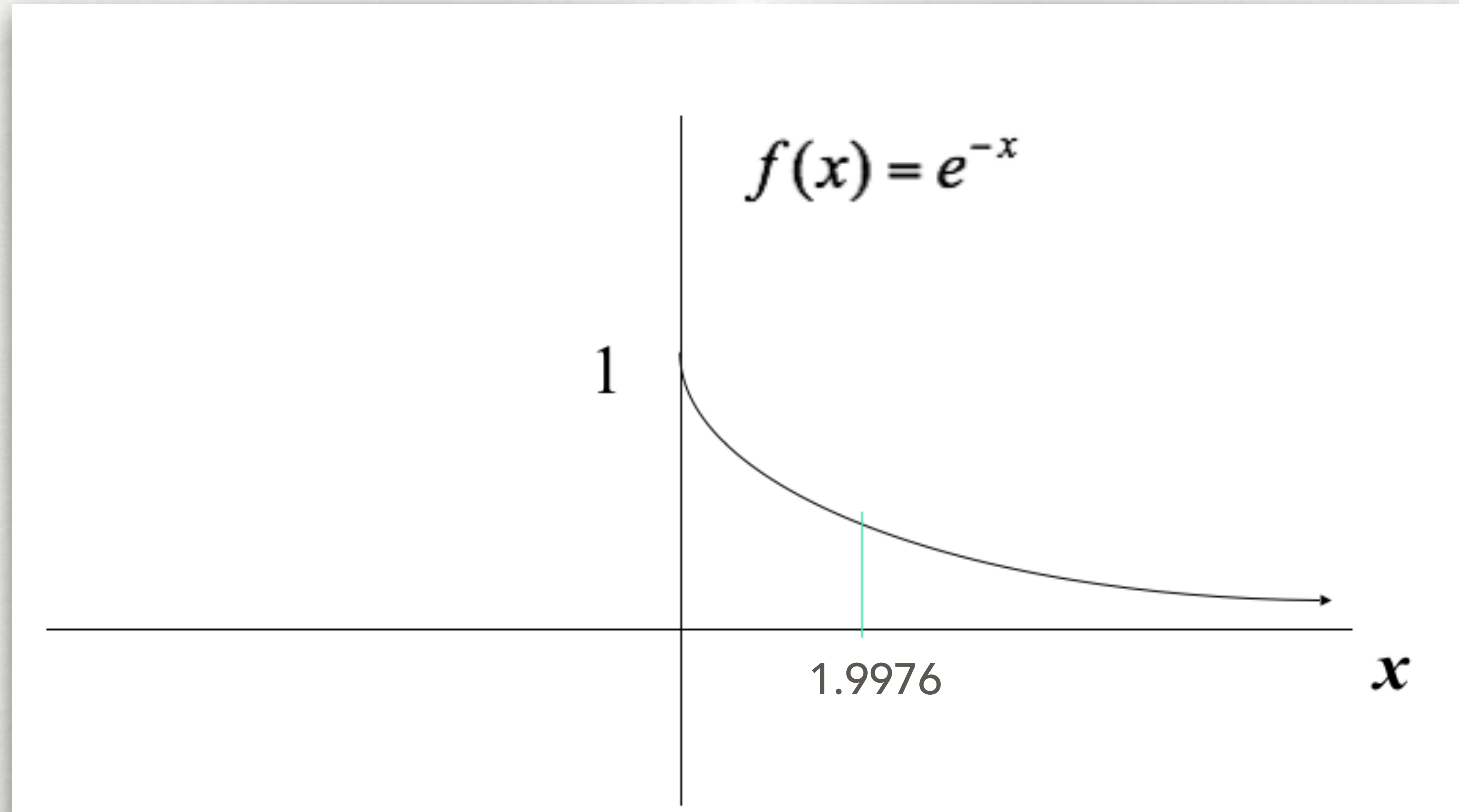
- This function integrates to 1:

$$\int_0^{+\infty} e^{-x} = -e^{-x} \Big|_0^{+\infty} = 0 + 1 = 1$$

CONTINUOUS PROBABILITY DISTRIBUTION

EXPONENTIAL DISTRIBUTION

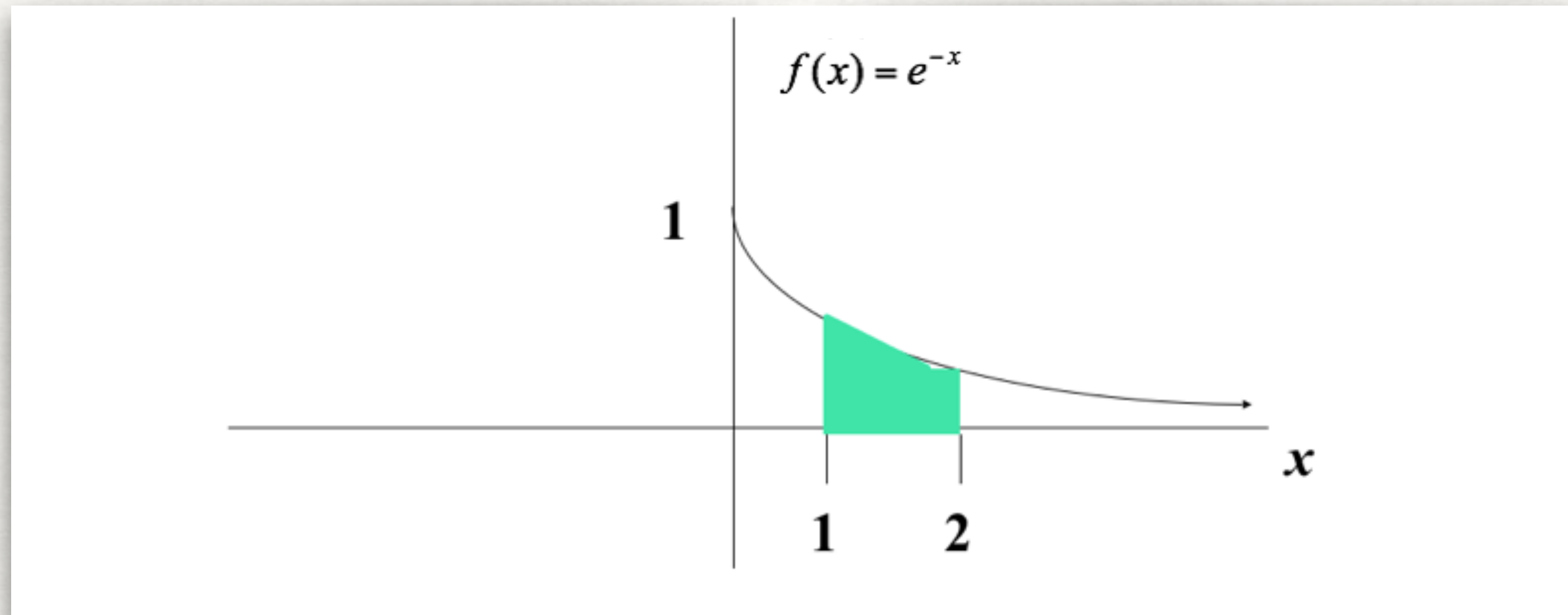
- The probability that x is any exact particular value (such as x = 1.9976) is 0;



CONTINUOUS PROBABILITY DISTRIBUTION

EXPONENTIAL DISTRIBUTION

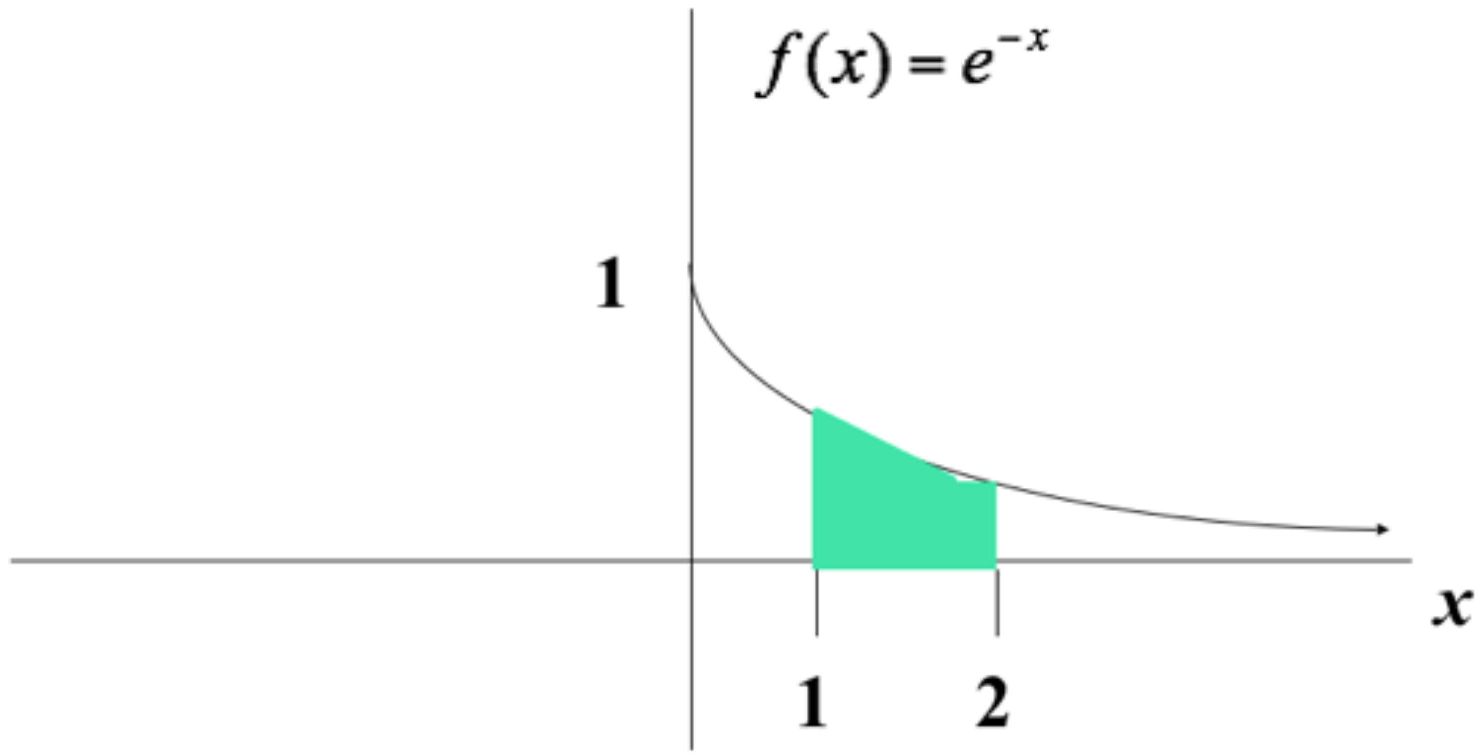
- we can only assign probabilities to possible ranges of x.



CONTINUOUS PROBABILITY DISTRIBUTION

EXPONENTIAL DISTRIBUTION

the probability of x falling within 1 to 2



$$P(1 \le x \le 2) = \int_1^2 e^{-x} = -e^{-x} \Big|_1^2 = -e^{-2} - -e^{-1} = -.135 + .368 = .23$$

CONTINUOUS PROBABILITY DISTRIBUTION

EXPONENTIAL DISTRIBUTION

Cumulative distribution function:

As in the discrete case, we can specify the “cumulative distribution function” (CDF):

The CDF here = P($x \leq$ A) if 0 ≤ A
= 0 if 0 > A

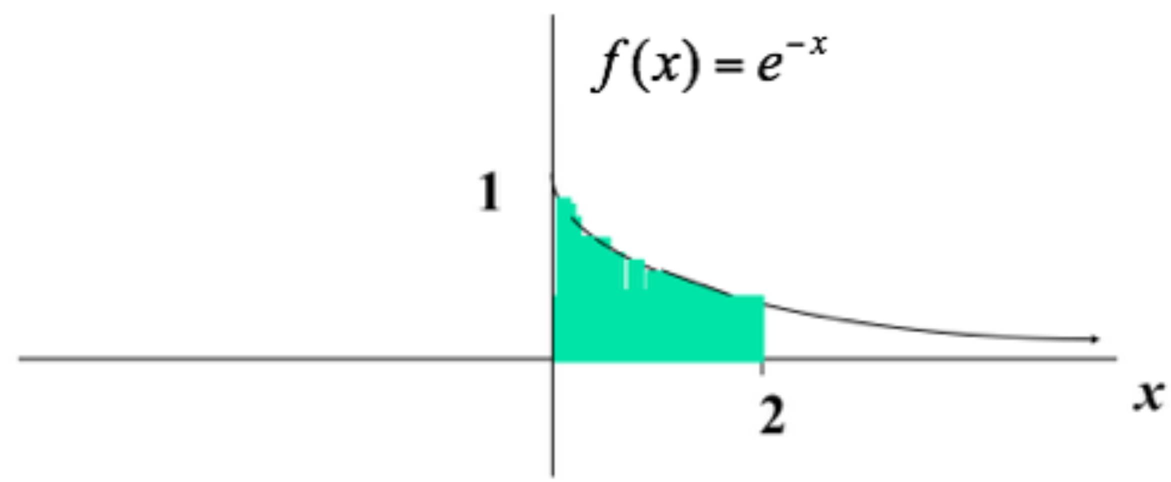
$$\int_0^A e^{-x} = -e^{-x} \Big|_0^A = -e^{-A} - -e^{0} = -e^{-A} + 1 = 1 - e^{-A}$$

CONTINUOUS PROBABILITY DISTRIBUTION

EXPONENTIAL DISTRIBUTION

By knowing CDF explicitly, we can compute:

P($x \le 2$) by plug in $x = 2$



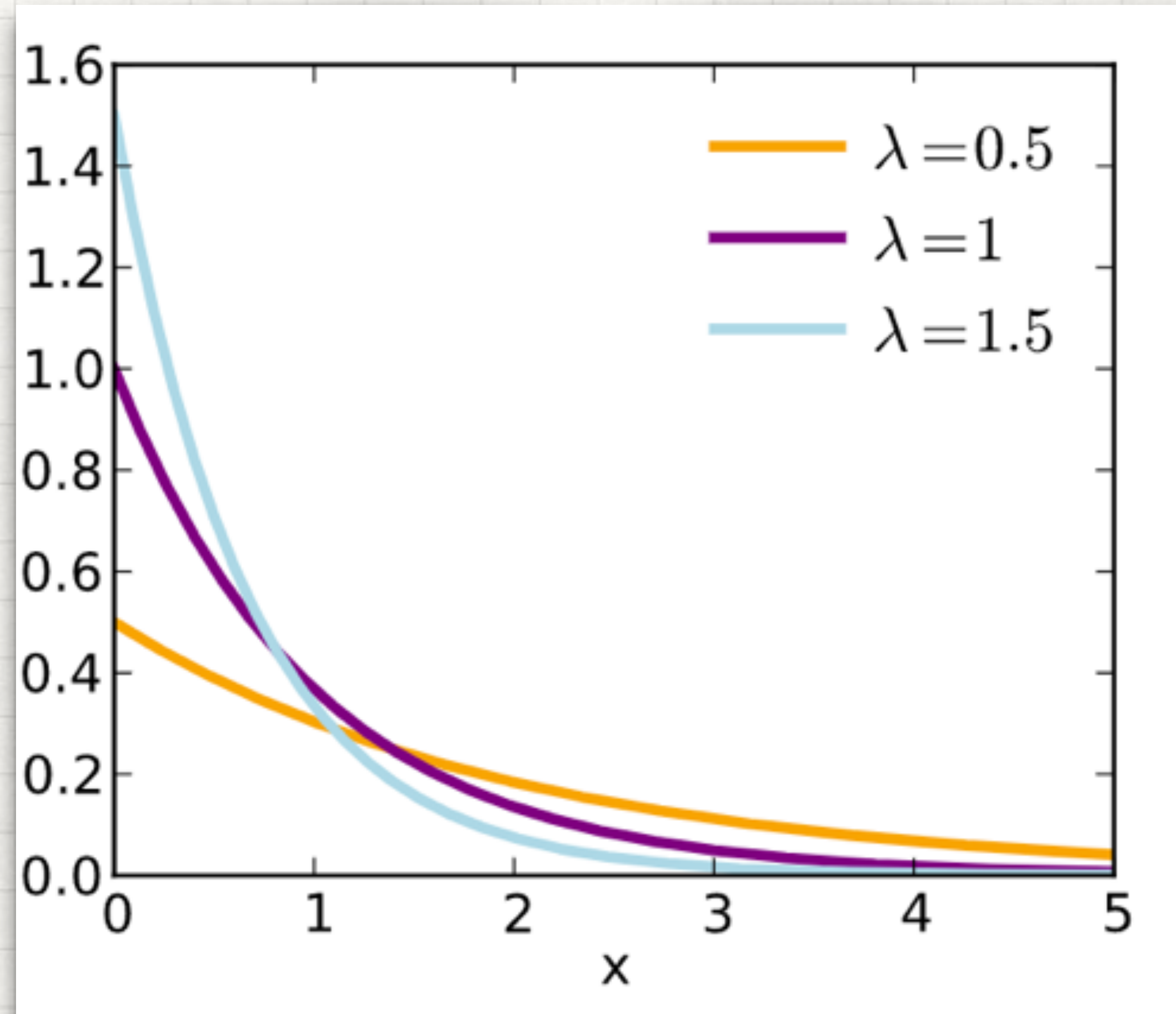
$$P(x \le 2) = 1 - e^{-2} = 1 - .135 = .865$$

CONTINUOUS PROBABILITY DISTRIBUTION

EXPONENTIAL DISTRIBUTION

- Probability density of exponential distribution with parameter **λ**

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

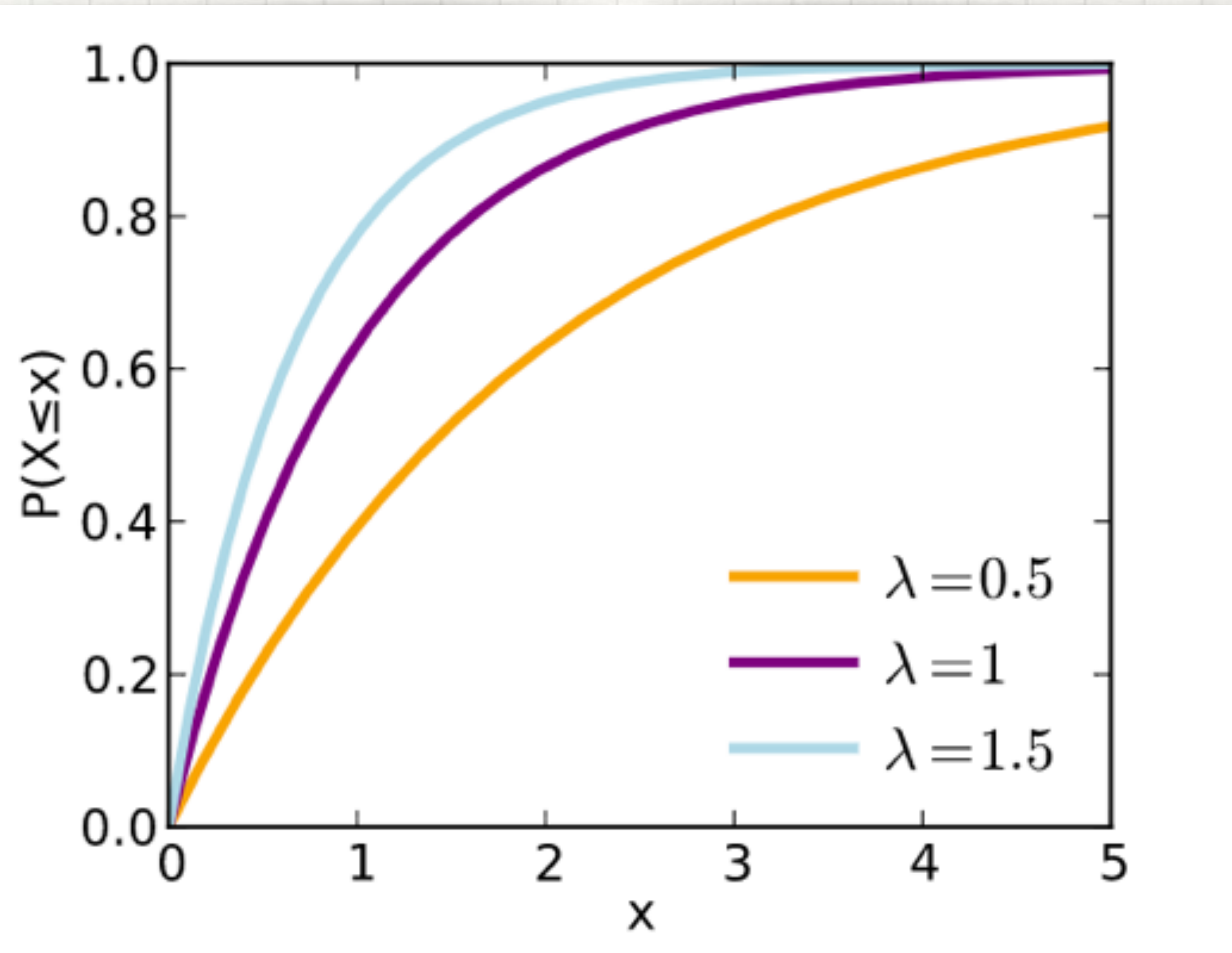


CONTINUOUS PROBABILITY DISTRIBUTION

EXPONENTIAL DISTRIBUTION

- Cumulative distribution function of exponential distribution with parameter λ

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



CONTINUOUS PROBABILITY DISTRIBUTION

EXPONENTIAL FUNCTION

Suppose that survival drops off rapidly in the year following diagnosis of a certain type of advanced cancer.
Suppose that the length of survival (or time-to-death) is a random variable that approximately follows an **exponential distribution with parameter 2** (makes it a steeper drop off):

pdf here is: $2e^{-2T}$

$$[note: \int_0^{+\infty} 2e^{-2x} = -e^{-2x} \Big|_0^{+\infty} = 0 + 1 = 1]$$

What's the probability that a person who is diagnosed with this illness survives a year?

CONTINUOUS PROBABILITY DISTRIBUTION

EXPONENTIAL FUNCTION

The probability of dying within 1 year can be calculated using the cumulative distribution function:
Cumulative distribution function is:

$$P(x \le T) = -e^{-2x} \Big|_0^T = 1 - e^{-2(T)}$$

The chance of surviving past 1 year is: $P(x \geq 1) = 1 - P(x \leq 1)$

$$1 - (1 - e^{-2(1)}) = .135$$

CONTINUOUS PROBABILITY DISTRIBUTION

OTHER IMPORTANT DISTRIBUTIONS

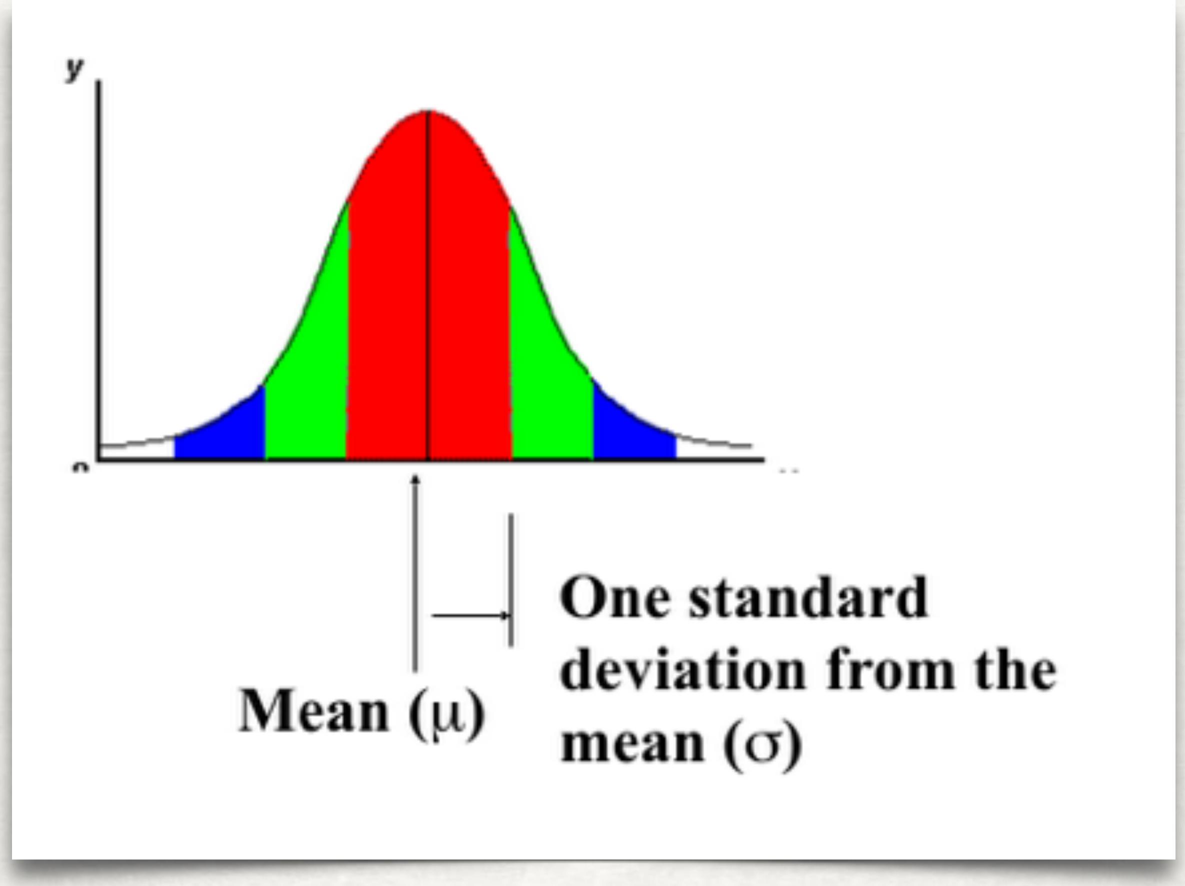
- Poisson
- Gaussian and multivariate Gaussian
- Gamma distribution
- Beta distribution
- Binomial (discrete)

EXPECTED VALUE AND VARIANCE

EXPECTED VALUES AND VARIANCE

ALL RANDOM VARIABLES

- All probability distributions are characterized by an expected value and a variance (standard deviation squared).
- For example, bell-curve (normal) distribution:



EXPECTED VALUES AND VARIANCE

EXPECTED VALUE, OR MEAN

- If we understand the underlying probability function of a certain phenomenon,
  then we can make informed decisions based on **how we expect x to behave on-average over the long-run**…(so called "frequentist" theory of probability).

- Expected value is just the **weighted average** or mean (μ) of random variable x.

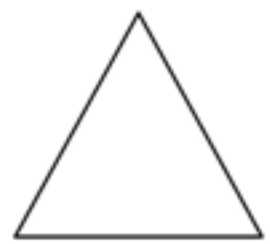
- Imagine placing the masses p(x) at the points X on a beam; the balance point of the beam is the expected value of x.

EXPECTED VALUES AND VARIANCE

EXPECTED VALUE, OR MEAN

- eg:
  Discrete random variable X with following probability distribution

x	10	11	12	13	14
P(x)	.4	.2	.2	.1	.1



$$\sum_{i=1}^{5} x_i p(x) = 10(.4) + 11(.2) + 12(.2) + 13(.1) + 14(.1) = 11.3$$

- Mean or expectation = probability based weighted average

EXPECTED VALUES AND VARIANCE

EXPECTED VALUE, OR MEAN

- Discrete case:

$$E[X] = \sum_i x_i f(x_i)$$

- Continuous case:

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx$$

WHAT IS AVERAGE IN REALITY?

- Empirical Mean is a special case of Expected Value…

Sample mean, for a sample of n subjects: =

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n} = \sum_{i=1}^{n} x_i \left(\frac{1}{n}\right)$$

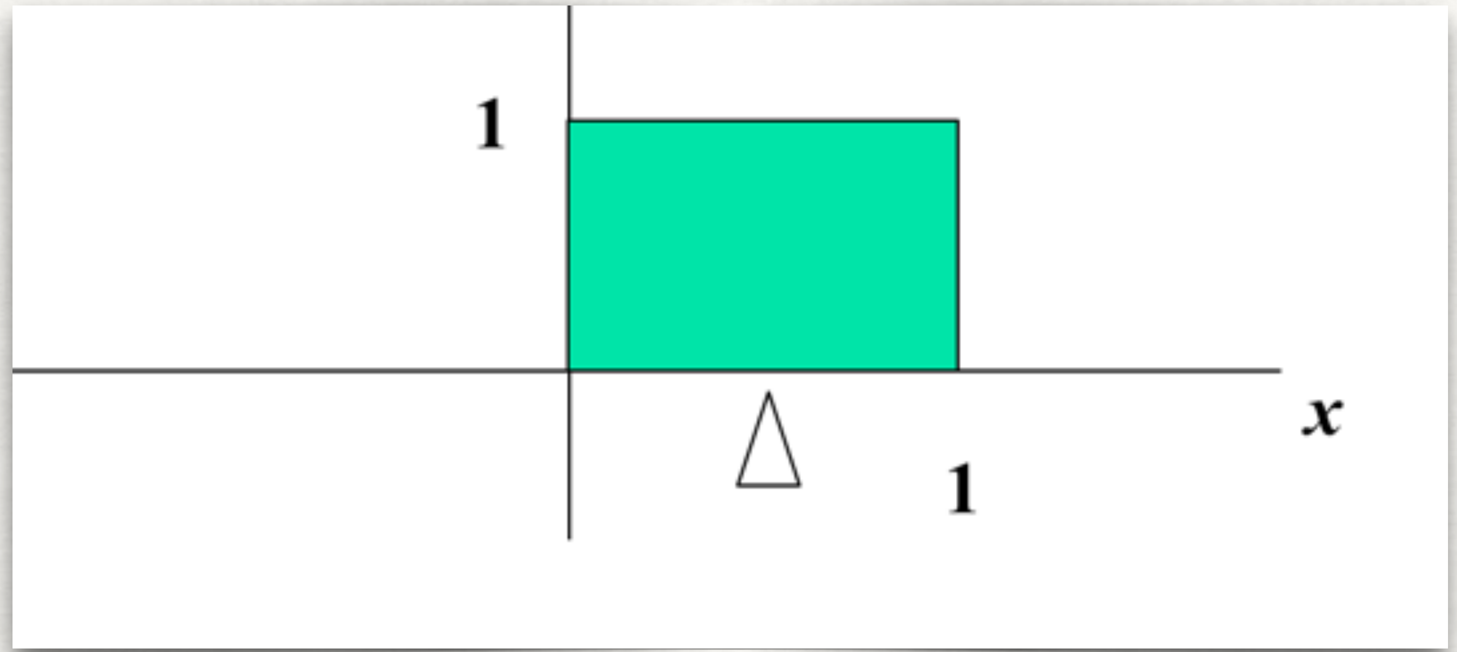
The probability (frequency) of each person in the sample is 1/n.

EXPECTED VALUES: CONTINUOUS CASE

UNIFORM DISTRIBUTION

- uniform distribution on [0,1]

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

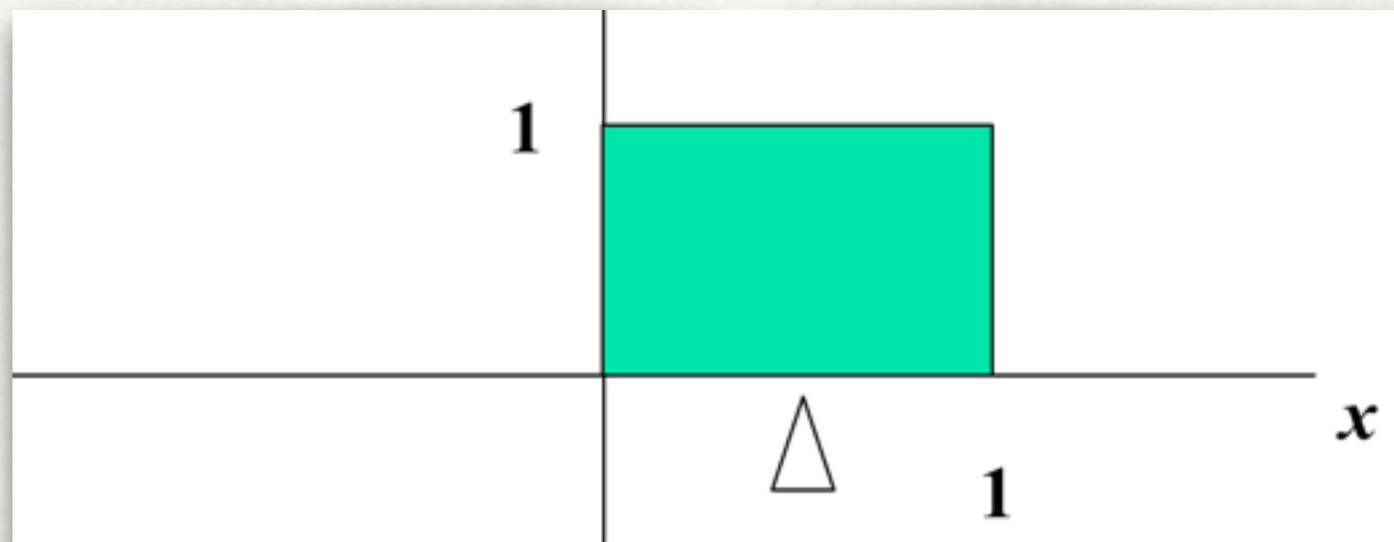


EXPECTED VALUES: CONTINUOUS CASE

UNIFORM DISTRIBUTION

- uniform distribution on [0,1]

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



$$E(X) = \int_0^1 x(1)dx = \frac{x^2}{2} \Big|_0^1 = \frac{1}{2} - 0 = \frac{1}{2}$$

EXPECTATION SYMBOL

EXPECTED VALUE, OR MEAN

- $E(X) = \mu$
  these symbols are used interchangeably

- Expected value is an extremely useful concept for good decision-making! Or the idea of machine learning to predict the future

- Predicting future
  = Expected Value +/- (uncertainty…)

EXAMPLE: THE LOTTERY

EXPECTED VALUE, OR MEAN

- example about The simple Lottery:
- A certain lottery works by the following
    - Picking 6 numbers from 1 to 49.
    - It costs $1.00 to play the lottery
    - if you win, you win $2 million after taxes.
- If you play the lottery once, what are your expected winnings or losses?

EXAMPLE: THE LOTTERY

EXPECTED VALUE, OR MEAN

- Calculate the probability of winning in 1 try:

EXAMPLE: THE LOTTERY

EXPECTED VALUE, OR MEAN

- Calculate the probability of winning in 1 try:

$$\frac{1}{\binom{49}{6}} = \frac{1}{\frac{49!}{43!6!}} = \frac{1}{13{,}983{,}816} = 7.2 \times 10^{-8}$$

"49 choose 6"

Out of 49 numbers, this is the number of distinct combinations of 6.

EXAMPLE: THE LOTTERY

EXPECTED VALUE, OR MEAN

- Calculate the probability of winning in 1 try:

$$\frac{1}{\binom{49}{6}} = \frac{1}{\frac{49!}{43!6!}} = \frac{1}{13,983,816} = 7.2 \times 10^{-8}$$

"49 choose 6"

Out of 49 numbers, this is the number of distinct combinations of 6.

The probability function (note, sums to 1.0):

x$	p(x)
-1	.999999928
+ 2 million	$7.2 \times 10^{--8}$

EXAMPLE: THE LOTTERY

EXPECTED VALUE, OR MEAN

- Rules:
  - Picking 6 numbers from 1 to 49.
  - It costs $1.00 to play the lottery
  - if you win, you win $2 million after taxes.
- The probability function

x$	p(x)
-1	.999999928
+ 2 million	$7.2 \times 10^{--8}$

EXAMPLE: THE LOTTERY

EXPECTED VALUE, OR MEAN

- Rules:
  - Picking 6 numbers from 1 to 49.
  - It costs $1.00 to play the lottery
  - if you win, you win $2 million after taxes.

- The probability function

x$	p(x)
-1	.999999928
+ 2 million	7.2×10^{-8}

- Expected Value:
  E(X) = P(win)*$2,000,000 + P(lose)*-$1.00
  = .144 - .999999928 = -$.86

EXAMPLE: THE LOTTERY

EXPECTED VALUE, OR MEAN

- The probability function

x$	p(x)
-1	.999999928
+ 2 million	$7.2 \times 10^{--8}$

- Expected Value:
  E(X) = P(win)*$2,000,000   + P(lose)*-$1.00
  = .144 - .999999928 = -$.86

- Negative expected value means?

EXAMPLE: THE LOTTERY

EXPECTED VALUE, OR MEAN

- The probability function

x$	p(x)
-1	.999999928
+ 2 million	7.2×10^{-8}

- Expected Value:
  E(X) = P(win)*$2,000,000 + P(lose)*-$1.00
  = .144 - .999999928 = -$.86

- Negative expected value means?

- You shouldn't play if you expect to lose money!

EXAMPLE: THE LOTTERY

EXPECTED VALUE, OR MEAN

- Lottery
- Expected Value:
  $E(X) = P(win)*\$2,000,000 \quad + \quad P(lose)*-\$1.00$
  $= .144 - .999999928 = -\$.86$
- Negative expected value is never good!
  You shouldn't play if you expect to lose money!
- But statistics/probability is a summary from a big population…
  So, there is always some lucky person…
- **Luck is always for someone else and expected value is always for a group of population.**

EXAMPLE: THE LOTTERY

EXPECTED VALUE, OR MEAN

- If you play the lottery every week for 10 years, what are your expected winnings or losses?

EXAMPLE: THE LOTTERY

EXPECTED VALUE, OR MEAN

- If you play the lottery every week for 10 years, what are your expected winnings or losses?

- 52 weeks per year

- 520 x (-.86) = -$447.20

EXPECTED VALUES

- A few notes about Expected Value as a mathematical operator:
- c is a constant, X and Y are two random variables
- $E(c) = c$
- $E(cX)=cE(X)$
- $E(c + X)=c + E(X)$
- $E(X+Y)= E(X) + E(Y)$

EXPECTED VALUES

- E(c) = c

- Example:
  If you cash in water bottles in NY, you always get 5 cents per bottle.

- Therefore, there's no randomness. You always expect to (and do) get 5 cents.

EXPECTED VALUES

- E(X+Y)= E(X) + E(Y)

- Example: If you play the lottery twice, you expect to lose: -\$.86 + -\$.86. (E(X+X))

-

- NOTE: This works even if X and Y are dependent!! Does not require independence!! Proof left for later…

EXPECTED VALUES

- Expected value isn't everything though…
- What does uncertainty mean?

EXPECTED VALUES

- Expected value isn't everything though…
- What does uncertainty mean?
- eg: Let's say there are two identical doors one with \$1 and the other with \$400,000.
  While, the banker offers you \$200,000.
  You can choose to open door (pay 0.5\$ to do so) or to accept bankers offer, but not both.

- So, Deal or No Deal?

EXPECTED VALUES

- We have two cases here

-0.5

x$	p(x)
+1	.50
+$400,000	.50

x$	p(x)
+$200,000	1.0

EXPECTED VALUES

x$	p(x)
+1	.50
+$400,000	.50

$$\mu = E(X) = \sum_{\text{all } x} x_i p(x_i) = +1(.50) + 400{,}000(.50) = 200{,}000$$

-0.5

x$	p(x)
+$200,000	1.0

$$\mu = E(X) = 200{,}000$$

EXPECTED VALUES

- Expected values are the same, how to decide?

EXPECTED VALUES

- Expected values are the same, how to decide?
- Well, I will just take the 200000, feel safe…
- What is the “safe feeling”?

EXPECTED VALUES

Variance!

- **If you take the deal, the variance/standard deviation is 0.**
- **If you don't take the deal, what is average deviation from the mean?**
- **What's your gut guess?**

VARIANCE/STANDARD DEVIATION

VARIANCE/STANDARD DEVIATION

MEASURE UNCERTAINTY

- "The average (expected) squared distance (or deviation) from the mean"
- Discrete version:

$$\sigma^2 = Var(x) = E[(x-\mu)^2] = \sum_{\text{all x}} (x_i - \mu)^2 p(x_i)$$

- Continuous version:

$$\mathrm{Var}(X) = \sigma^2 = \int (x-\mu)^2 f(x)\, dx$$

VARIANCE/STANDARD DEVIATION

MEASURE UNCERTAINTY

- "The average (expected) squared distance (or deviation) from the mean"

$$\sigma^2$$

- We square because squaring has better properties than absolute value. Take square root to get back linear average distance from the mean (="standard deviation").
- more like a square of distance
- standard deviation: **σ** — distance

VARIANCE/STANDARD DEVIATION

MEASURE UNCERTAINTY

- Similarity to empirical variance

The variance of a sample: $s^2 =$

$$\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{n-1} = \sum_{i=1}^{N}(x_i - \bar{x})^2\left(\frac{1}{n-1}\right)$$

Division by n-1 reflects the fact that we have lost a "degree of freedom" (piece of information) because we had to estimate the sample mean before we could estimate the sample variance.

VARIANCE/STANDARD DEVIATION

MEASURE UNCERTAINTY

- symbol
- $Var(X) = \sigma^2$
  - these symbols are used interchangeably
- **σ** — standard deviation

VARIANCE/STANDARD DEVIATION

MEASURE UNCERTAINTY

- Back to the question: accept banker's offer?
  Both cases has expected value 200,000

x$	p(x)
+1	.50
+$400,000	.50

x$	p(x)
+$200,000	1.0

Now you examine your personal risk tolerance…

VARIANCE/STANDARD DEVIATION

MEASURE UNCERTAINTY

- Back to the question: accept banker's offer?

$$\sigma^2 = \sum_{\text{all x}} (x_i - \mu)^2 \, p(x_i) =$$

$$= (1 - 200{,}000)^2 (.5) + (400{,}000 - 200{,}000)^2 (.5) = 200{,}000^2$$

$$\sigma = \sqrt{200{,}000^2} = 200{,}000$$

Now you examine your personal risk tolerance…

VARIANCE/STANDARD DEVIATION

MEASURE UNCERTAINTY

- Handy calculation formula — discrete random variable

$$Var(X) = \sum_{\text{all x}} (x_i - \mu)^2 p(x_i) = \sum_{\text{all x}} x_i^2 p(x_i) - (\mu)^2$$

Intervening algebra!

$$= E(x^2) - [E(x)]^2$$

VARIANCE/STANDARD DEVIATION

MEASURE UNCERTAINTY

- For continuous random variable we have the same formula (need proofs)

$$\mathbf{Var(x) = E(x-\mu)^2 = E(x^2) - [E(x)]^2}$$

VARIANCE/STANDARD DEVIATION

MEASURE UNCERTAINTY

- c= a constant number (i.e., not a variable)
- X and Y are random variables, then
  - Var(c) = 0
  - Var (c+X)= Var(X)
  - Var(cX)= c2Var(X)
  - Var(X+Y)= Var(X) + Var(Y)+2Cov(X,Y)

VARIANCE/STANDARD DEVIATION

MEASURE UNCERTAINTY

$\text{Var}(c) = 0$

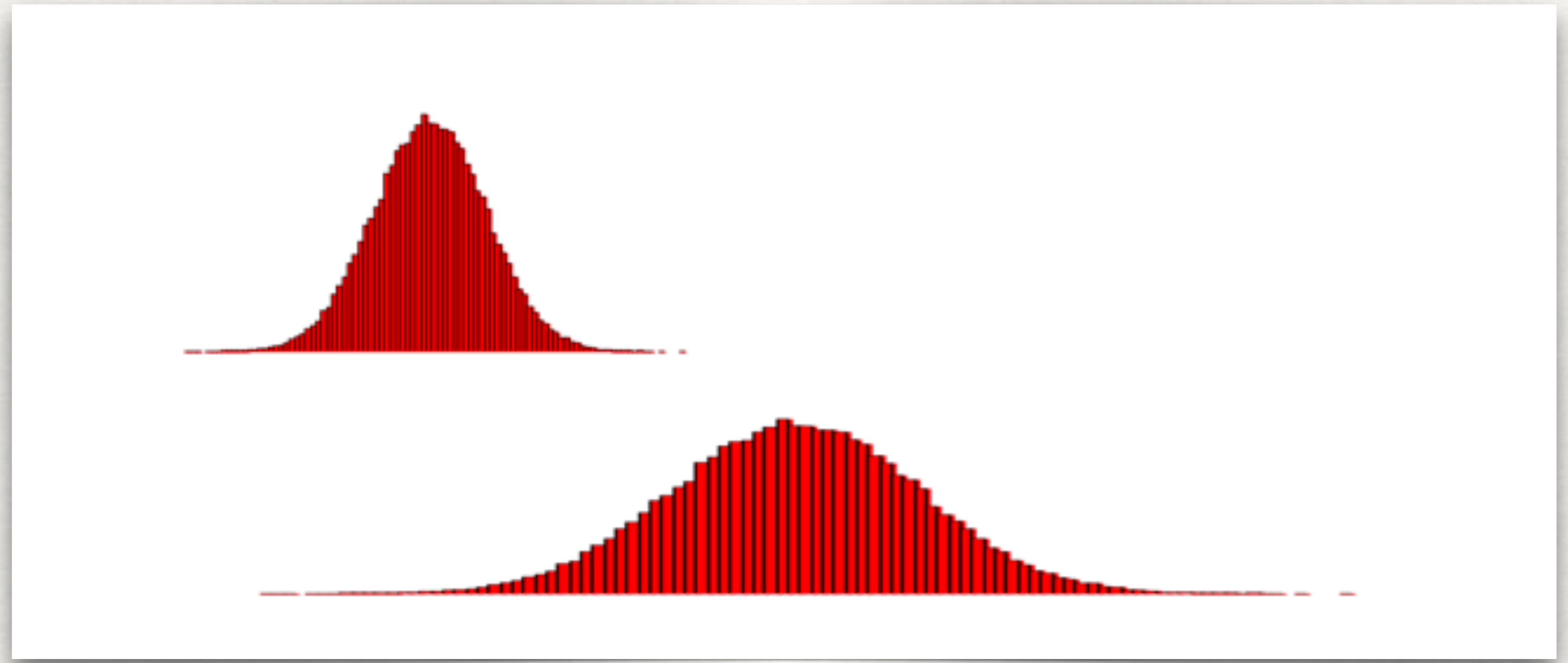
Constants, no vary!

VARIANCE/STANDARD DEVIATION

MEASURE UNCERTAINTY

$Var(cX) = c^2 Var(X)$

Multiplying each instance of the random variable by c makes it c-times as wide of a distribution, which corresponds to c^2 as much variance (deviation squared). For example, if everyone suddenly became twice as tall, there'd be twice the deviation and 4 times the variance in heights in the population.



Thus, we use standard deviation that make a factor impact the same

STANDARD DEVIATION

- square root of variance…
  universal symbol:
- σ

COVARIANCE: JOINT PROBABILITY

COVARIANCE

- The covariance measures the strength of the linear relationship between two variables

- The covariance:

$$E[(x - \mu_x)(y - \mu_y)]$$

$$\sigma_{xy} = \sum_{i=1}^{N} (x_i - \mu_x)(y_i - \mu_y)\, P(x_i, y_i)$$

COVARIANCE

- sample covariance

$$\operatorname{cov}(x,y) = \frac{\sum_{i=1}^{n}(x_i - \bar{X})(y_i - \bar{Y})}{n-1}$$

COVARIANCE

- Covariance between two random variables:

- cov(X,Y) > 0 X and Y are positively correlated
- cov(X,Y) < 0 X and Y are inversely correlated
- cov(X,Y) = 0 X and Y are independent

COVARIANCE

- sample covariance

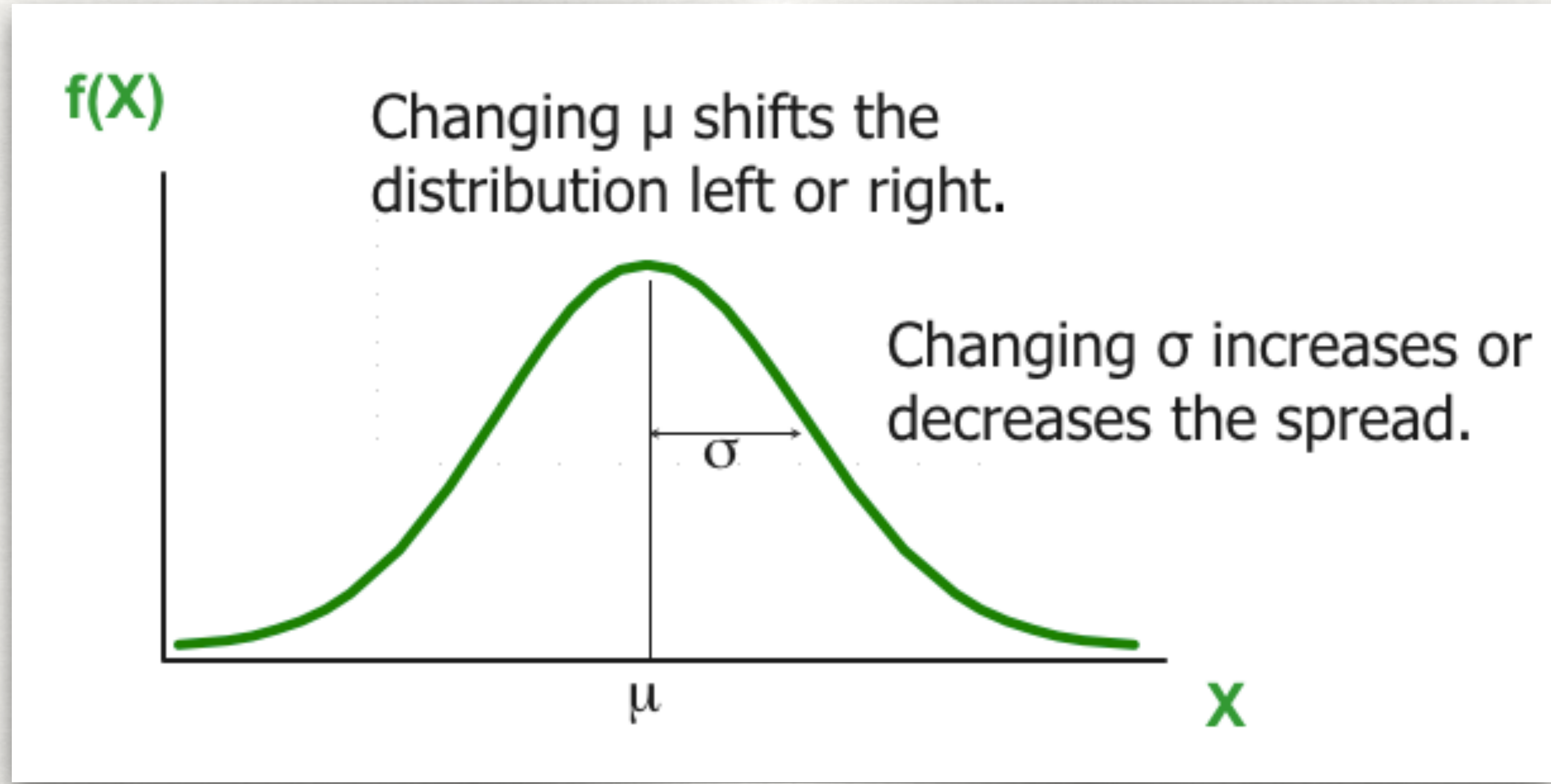
$$\operatorname{cov}(x, y) = \frac{\sum_{i=1}^{n} (x_i - \bar{X})(y_i - \bar{Y})}{n-1}$$

THE NORMAL AND STANDARD NORMAL

THE NORMAL DISTRIBUTION

GAUSSIAN

- Probability density function of a Gaussian distribution



THE NORMAL DISTRIBUTION

GAUSSIAN

- The Normal Distribution:
  as mathematical function (pdf)

$$f(x) = \frac{1}{\boxed{\sigma}\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-\boxed{\mu}}{\sigma}\right)^2}$$

<u>Note constants:</u>
π=3.14159
e=2.71828

This is a bell shaped curve with different centers and spreads depending on μ and σ

THE NORMAL DISTRIBUTION

GAUSSIAN

- It's a probability function, so no matter what the values of μ and **σ**, must integrate to 1!

$$E(X)=\mu = \boxed{\int_{-\infty}^{+\infty} x \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx}$$

$$Var(X)=\sigma^2 = \boxed{\int_{-\infty}^{+\infty} x^2 \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx) - \mu^2}$$

$$\text{Standard Deviation}(X)=\sigma$$

THE NORMAL DISTRIBUTION

GAUSSIAN

- Normal curve:

  No matter what μ and $\boldsymbol{\sigma}$ are,

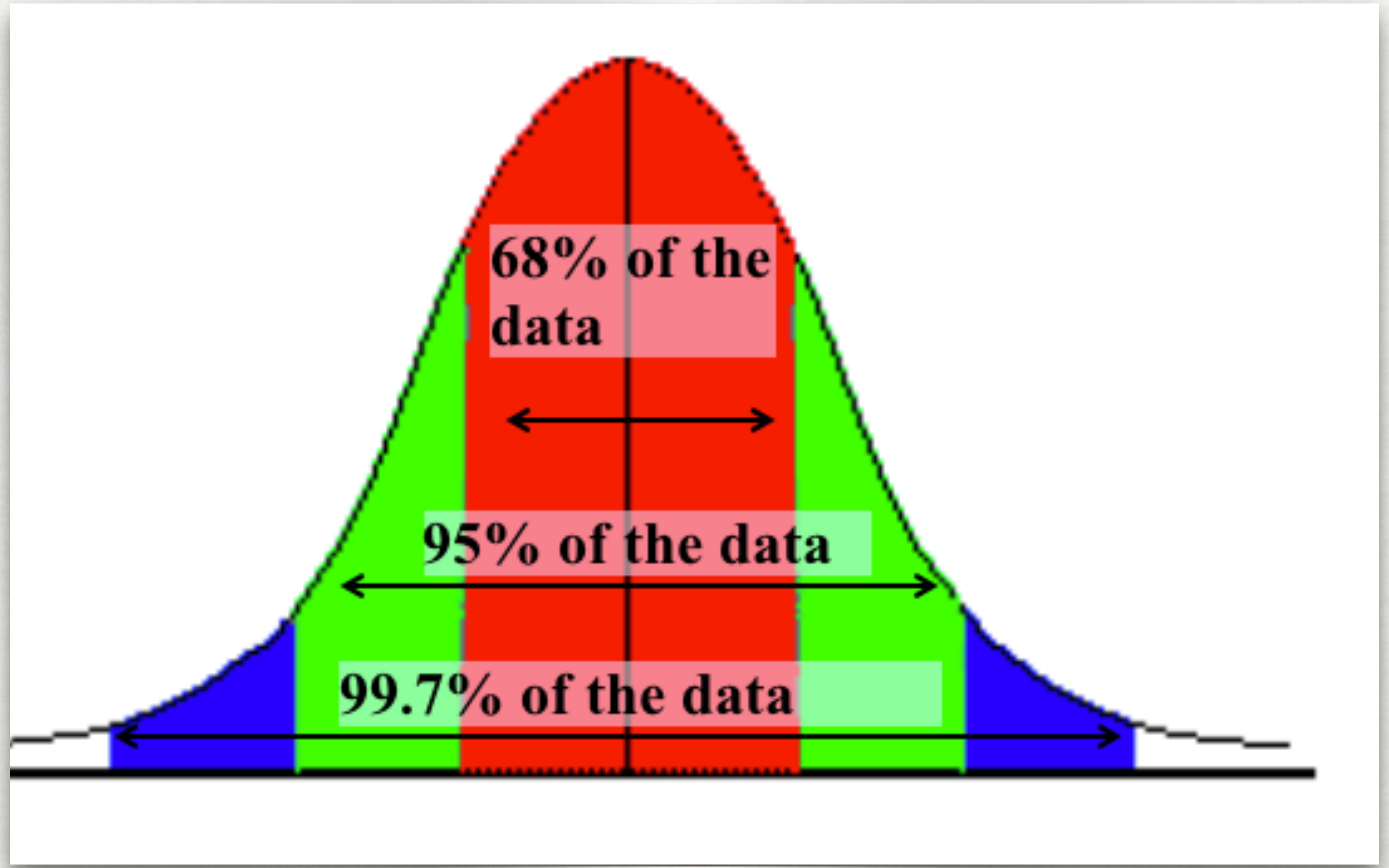
  - the area between μ-$\boldsymbol{\sigma}$ and μ+$\boldsymbol{\sigma}$ is about 68%;
  - the area between μ-2$\boldsymbol{\sigma}$ and μ+2$\boldsymbol{\sigma}$ is about 95%;
  - and the area between μ-3$\boldsymbol{\sigma}$ and μ+3$\boldsymbol{\sigma}$ is about 99.7%.

- Almost all values fall within 3 standard deviations.

- 68-95-99.7 Rule

THE NORMAL DISTRIBUTION

GAUSSIAN



THE NORMAL DISTRIBUTION

GAUSSIAN

- 68-95-99.7 Rule (more data later)
  in Math terms…

$$\int_{\mu-\sigma}^{\mu+\sigma} \frac{1}{\sigma\sqrt{2\pi}} \bullet e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = .68$$

$$\int_{\mu-2\sigma}^{\mu+2\sigma} \frac{1}{\sigma\sqrt{2\pi}} \bullet e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = .95$$

$$\int_{\mu-3\sigma}^{\mu+3\sigma} \frac{1}{\sigma\sqrt{2\pi}} \bullet e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = .997$$