Probability in Computing

Announcements

- HW4 is out
- Expect Nut Bars for the final exam
- Bring your bottle of water

- **LECTURE 13**
- Estimation by sampling II
- Markov inequality
- Chebyshev inequality
- Applications of Markov and Chebyshev's inequalities
- Continuous distributions:
  - Uniform, Normal
  - Exponential, Poisson

Prof. Tiago Januario

Slides by Prof. Sofya Raskhodnikova and Alina Ene

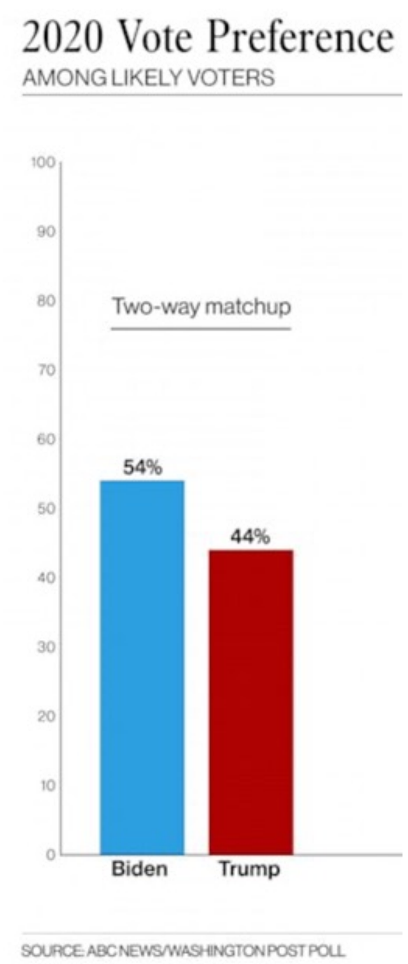
Simulations

- In lecture and in the homework, we have used simulations to empirically estimate probabilities, expectations, and distributions
- In each of those cases, we used many repeated trials (typically 10,000 trials)
- Our goal this week is to understand questions like:
  – How accurate were our estimates?
  – How do we decide how many trials/samples we need?
  – Were 10,000 trials too few, too many, or just right?

Estimation by Sampling

We will use polling as a running example

- Suppose we have an upcoming election with two candidates: candidate A and candidate B
- Let p be the fraction of the voters that support candidate A
- The job of a pollster is to estimate this unknown fraction p
- How can the pollster proceed?



Estimation by Sampling

- Approach 1: Pollster calls every voter and asks them which candidate they support
  - **Pro:** the pollster will get a perfect estimate
  - **Con:** the pollster will need to call hundreds of millions of people, this is nearly impossible and could take years
- Approach 2: Pollster calls a small sample of voters and asks them which candidate they support
  - **Pro:** the number of people is much smaller
  - **Con:** the estimate could be very inaccurate

Estimation by Sampling: Polling

The pollster uses the following polling algorithm:

1. Choose a sample size n
2. Sample n people independently and uniformly at random with replacement from the entire population
3. For each sampled person, ask them which candidate they support (we assume they answer truthfully)
4. Use the fraction of people in the sample that support candidate A as the estimate for the true fraction

Estimation by Sampling: RVs

- For each $i \in \{1,2,\ldots,n\}$, define:

$$X_i = \begin{cases} 1 \text{ if } i^{\text{th}} \text{ person in the sample supports A} \\ 0 \quad \text{otherwise} \end{cases}$$

- The pollster's estimate is:

$$P = \frac{1}{n}\sum_{i=1}^{n} X_i$$

- Our goal is to understand how ``close'' the estimate P is to the actual fraction of voters that support candidate A

Expectation of Estimate P

- Fraction of the population supporting candidate A is $\boldsymbol{p}$
- The pollster's estimate is: $\boldsymbol{P} = \frac{1}{n}\sum_{i=1}^{n} X_i$,

  where $X_i = \begin{cases} 1 \text{ if } i^{\text{th}} \text{ person in the sample supports A} \\ 0 \quad \text{otherwise} \end{cases}$
- What are the distribution, expectation and variance of X_i?

  $X_i \sim$ Bernoulli(p) $\quad \mathbb{E}(X_i) = p \quad \mathrm{Var}(X_i) = p(1-p)$
- By linearity of expectation,

$$\mathbb{E}(P) = \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(X_i) = \frac{1}{n}\sum_{i=1}^{n} p = \frac{p}{n}\sum_{i=1}^{n} 1 = \frac{np}{n} = p$$

$\mathbb{E}(P)$ is exactly the unknown fraction p we wanted to estimate

Variance of Estimate P

- Fraction of the population supporting candidate A is $\boldsymbol{p}$
- The pollster's estimate is: $\boldsymbol{P} = \frac{1}{n}\sum_{i=1}^{n} X_i$,

  where $X_i = \begin{cases} 1 \text{ if } i^{\text{th}} \text{ person in the sample supports A} \\ 0 \quad \text{otherwise} \end{cases}$

  $X_i \sim$ $\qquad\qquad$ $\mathbb{E}(X_i) =$ $\qquad\qquad$ $\text{Var}(X_i) =$

- Since X_i are independent,

$$\text{Var(P)} = \text{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) =$$

The larger n is, the smaller Var(P) is.

Estimation by Sampling

- The pollster's estimate is correct in expectation
- Next, we introduce the tools we will need to understand how much it can deviate from its expectation
- Developing the intuition: What fraction of students can get **at least twice the average** on an exam?

at least thrice the average?

Markov Inequality

Theorem (Markov Inequality)

Let X be a random variable taking only nonnegative values. Then, for all $a > 0$,

$$\Pr(X \geq a) \leq \frac{\mathbb{E}(X)}{a}.$$

Useful when $a > \mathbb{E}(X)$

Andrei Markov [1856-1922]

Proof: Let $a > 0$.

Want to prove: $\mathbb{E}(X) \geq a \cdot \Pr(X \geq a)$

Proof for continuous random variables

For discrete random variables, we can use a similar idea

just replace sums with integrals

image source https://en.wikipedia.org/wiki/Andrey_Markov

Markov Inequality: Proof

- Want to prove: $\mathbb{E}(X) \geq a \cdot \Pr(X \geq a)$
- $\mathbb{E}(X) =$

Markov Inequality: Corollary

Theorem (Markov Inequality)

Let X be a random variable taking only nonnegative values. Then, for all $a > 0$,

$$\Pr(X \geq a) \leq \frac{\mathbb{E}(X)}{a}.$$

Corollary (Variant of Markov Inequality)

Let X be a random variable taking only nonnegative values. Then, for all $b > 1$,

$$\Pr(X \geq b \cdot \mathbb{E}(X)) \leq \frac{1}{b}.$$

Andrei Markov
[1856-1922]

Markov Inequality: Polling

Corollary (Markov Inequality)

Let X be a RV taking only nonnegative values. Then, for all $b > 1$,

$$\Pr(X \geq b \cdot \mathbb{E}(X)) \leq \frac{1}{b}.$$

Andrei Markov
[1856-1922]

- By Markov inequality, applied to the polling estimate,

$$\Pr(P \geq 1.1p) \leq$$
$$\Pr(P \geq 2\,p) \leq$$
$$\Pr(P \geq b\,p) \leq$$

Markov Inequality: Assumption

Theorem (Markov Inequality)

Let X be a RV taking only nonnegative values.
Then, for all $a > 0$,

$$\Pr(X \geq a) \leq \frac{\mathbb{E}(X)}{a}.$$

Andrei Markov
[1856-1922]

- The nonnegativity assumption is necessary.
- Consider, for example,

Does not satisfy the nonnegativity assumption

$$X = \begin{cases} 1 & \text{with probability } 1/2 \\ -1 & \text{with probability } 1/2 \end{cases}$$

Then $\Pr(X \geq a) \nleq \frac{\mathbb{E}(X)}{a} \Rightarrow \Pr(X \geq 1) = \frac{1}{2} \nleq \mathbb{E}(X) = 0$, for $a = 1$.

Markov Inequality: Polling

- The guarantees we obtained using Markov inequality seem rather weak
- The guarantee is the same regardless of whether we poll 1 person or 225,000,000 people
- Can we do better?

Top Hat question (Join Code: 258641)

BU Terriers win each game independently with probability 2/3.

Let X be the number of their losses in n games.

- What bound on $\mathbf{Pr}\left(X \geq \frac{n}{2}\right)$ can you get using Markov inequality?

A. $\leq 1/3$

B. $\leq 1/2$

C. $\geq 1/2$

D. $\leq 2/3$

E. None of the above

Chebyshev's Inequality

$\mu_X = E(X)$

Theorem (Chebyshev's Inequality)

Let X be a random variable. For all $a > 0$,

$$\Pr(|X - \mathbb{E}(X)| \geq a) \leq \frac{\text{Var}(X)}{a^2}.$$

Pafnuty Chebyshev
[1821-1894]

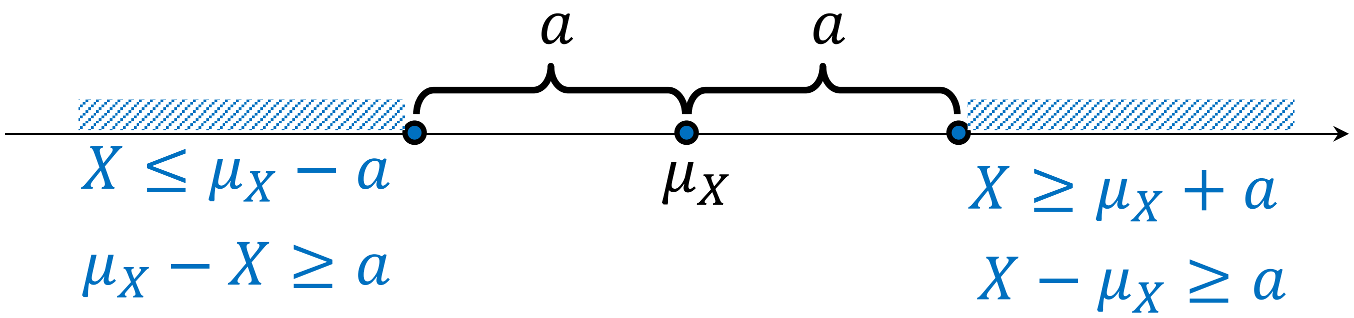


image source https://www.britannica.com/biography/Pafnuty-Lvovich-Chebyshev

Chebyshev's Inequality

Theorem (Chebyshev's Inequality)

Let X be a random variable. For all $a > 0$,

$$\Pr(|X - \mathbb{E}(X)| \geq a) \leq \frac{\mathrm{Var}(X)}{a^2}.$$

Proof: $\Pr(|X - \mathbb{E}(X)| \geq a) = \Pr\big(\underbrace{(X - \mathbb{E}(X))^2}_{Y \geq 0} \geq a^2\big)$

$$\leq \frac{\mathbb{E}(Y)}{a^2} \quad \text{(by Markov)}$$

$$= \frac{\mathbb{E}((X - \mathbb{E}(X))^2)}{a^2}$$

$$= \frac{\mathrm{Var}(X)}{a^2}$$

Top Hat question (Join Code: 258641)

BU Terriers win each game independently with probability 2/3.
Let X be the number of their losses in n games.
What bound on $\mathbf{Pr}\left(\boldsymbol{X} \geq \frac{\boldsymbol{n}}{\mathbf{2}}\right)$ can you get using Chebyshev's inequality?

A. $\leq 1/3$
B. $\leq const/n$
C. $\geq const/n^2$
D. $\leq const/n^3$
E. None of the above

For $n > 12$, Chebyshev is much better than Markov!

Estimation by Sampling: RVs

- For each $i \in \{1,2,\ldots,n\}$, define:

$$X_i = \begin{cases} 1 \text{ if } i^{\text{th}} \text{ person in the sample supports A} \\ 0 \quad \text{otherwise} \end{cases}$$

- The pollster's estimate is:

$$P = \frac{1}{n}\sum_{i=1}^{n} X_i$$

- Our goal is to understand how ``close'' the estimate P is to the actual fraction of voters that support candidate A

CS 237 Sampling Estimate from Polling

- Fraction of the population supporting candidate A is $\boldsymbol{p}$
- The pollster's estimate is: $\boldsymbol{P} = \frac{1}{n}\sum_{i=1}^{n} X_i$, where $X_i = \begin{cases} 1 \text{ if } i^{\text{th}} \text{ person in the sample supports A} \\ 0 \quad \text{otherwise} \end{cases}$

$$X_i \sim Bernoulli(p) \qquad \mathbb{E}(X_i) = p \qquad \mathrm{Var}(X_i) = p(1-p)$$

- Expectation and variance of our estimate:

$$\mathbb{E}(P) = \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(X_i) = \frac{1}{n} \cdot np = \boldsymbol{p}$$

$$\mathrm{Var}(P) = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}(X_i) = \frac{\boldsymbol{p(1-p)}}{\boldsymbol{n}}$$

Chebyshev's Inequality: Polling

Theorem (Chebyshev's Inequality)

Let X be a random variable. For all $a > 0$,

$$\Pr(|X - \mathbb{E}(X)| \geq a) \leq \frac{\mathrm{Var}(X)}{a^2}.$$

Pafnuty Chebyshev
[1821-1894]

- By applying Chebyshev's inequality to the polling estimate,

$\Pr(|P - p| \geq a) \leq$

Chebyshev's Inequality: Polling

$\Pr(|P-p| \geq a) \leq \frac{1}{4na^2}$

- The Chebyshev inequality bound allows us to determine how many people to poll
- Suppose we want the estimate to be within 0.04 of p with probability at least 0.95

Want:

Good

$$\Pr(|P-p| \leq 0.04) \geq 0.95$$

$$\Updownarrow$$

Bad

$$\Pr(|P-p| > 0.04) \leq 0.05$$

Chebyshev:

$\Pr(|P-p| > 0.04)$

$$\leq \Pr(|P-p| \geq 0.04) \leq \frac{1}{4 \cdot (0.04)^2 \cdot n} = 0.05$$

$$\frac{1}{4 \cdot (0.04)^2 \cdot n} = 0.05$$

Thus it suffices to poll $n = 3125$ voters

Estimation by Sampling

- The approach we studied in the context of polling can be used in a wide range of settings
- A common scenario in CS and beyond is that we want to estimate the expectation of a distribution using sampling
- Our earlier analysis works equally well for this setting: we take several samples from the distribution and we use their average as our estimate for the expectation
- Chebyshev's inequality then tells us how close our estimate is to the actual expectation, and it also tells us how many samples we need

Random Hats: Using Markov

Number of fixed points of a permutation. Let X be the number of students that get their hats back when n students randomly switch hats, so that every permutation of hats is equally likely.

Find a bound on $\Pr(X \geq x)$ using Markov inequality.

Solution: X_i= the indicator R.V. for person i getting their hat back.

$$X = X_1 + \cdots + X_n$$

Then
$$\begin{aligned}\mathbb{E}(X) &= \mathbb{E}(X_1 + \cdots + X_n) \\ &= \mathbb{E}(X_1) + \cdots + \mathbb{E}(X_n) \\ &= n \cdot \mathbb{E}(X_1) \\ &= n \cdot \Pr(X_1 = 1) \\ &= n \cdot \frac{1}{n} = 1\end{aligned}$$

by linearity of expectation

by symmetry

since X_1 is an indicator

By Markov inequality, $\Pr(X \geq x) \leq$

Random Hats: Variance

Find a bound on $\Pr(X \geq x)$ using Chebyshev's inequality.

Solution: We know: $X = X_1 + \cdots + X_n$ and $\mathbb{E}(X) = 1$

To compute $\mathrm{Var}(X)$:

$$\mathbb{E}(X^2) = \mathbb{E}[(X_1 + \cdots + X_n)(X_1 + \cdots + X_n)]$$

$$= \mathbb{E}(X_1^2) + \cdots \mathbb{E}(X_n^2) + \mathbb{E}(X_1 \cdot X_2) + \cdots + \mathbb{E}(X_{n-1} \cdot X_n)$$

by linearity of expectation

$$= n \cdot \mathbb{E}(X_1^2) + n(n-1) \cdot \mathbb{E}(X_1 \cdot X_2)$$

by symmetry

$$= n \cdot \Pr(X_1 = 1) + n(n-1) \cdot \Pr(X_1 \cdot X_2 = 1)$$

since X_1 and $X_1 \cdot X_2$ are indicators

$$= n \cdot \frac{1}{n} + n(n-1) \cdot \frac{1}{n(n-1)} = 2$$

Random Hats: Using Chebyshev

Find a bound on $\Pr(X \geq x)$ using Chebyshev's inequality.

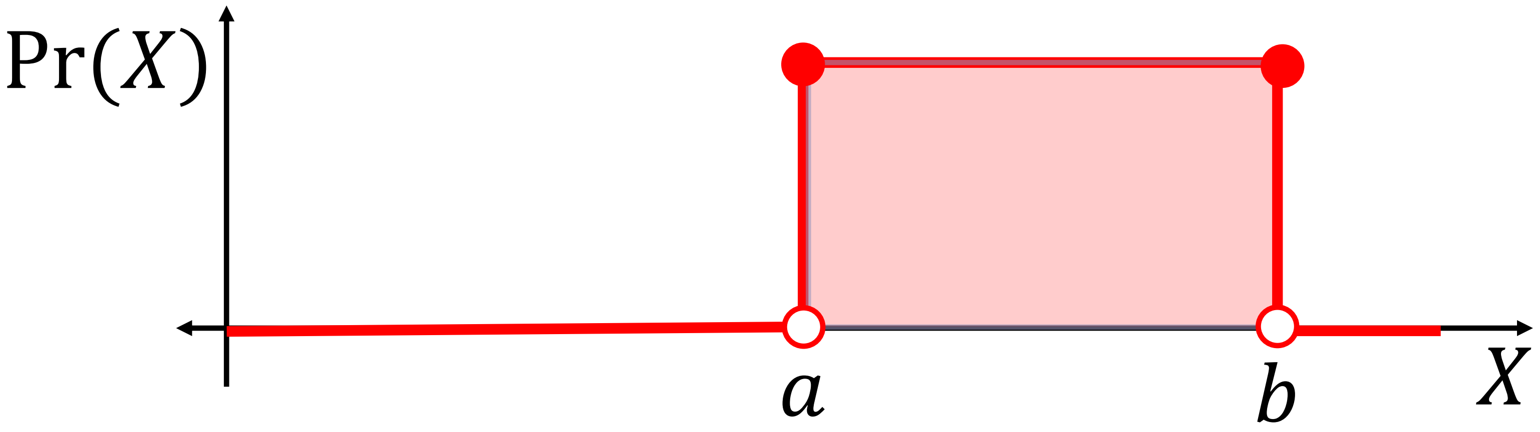
Solution: We know: $\mathbb{E}(X) = 1$ and $\mathbb{E}(X^2) = 2$

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = 2 - 1 = 1$$

$$\begin{aligned}\Pr(X \geq x) &= \Pr(X - 1 \geq x - 1) \\ &= \Pr(X - \mathbb{E}(X) \geq x - 1) \\ &\leq \Pr(|X - \mathbb{E}(X)| \geq x - 1)\end{aligned}$$

Continuous Uniform Distribution

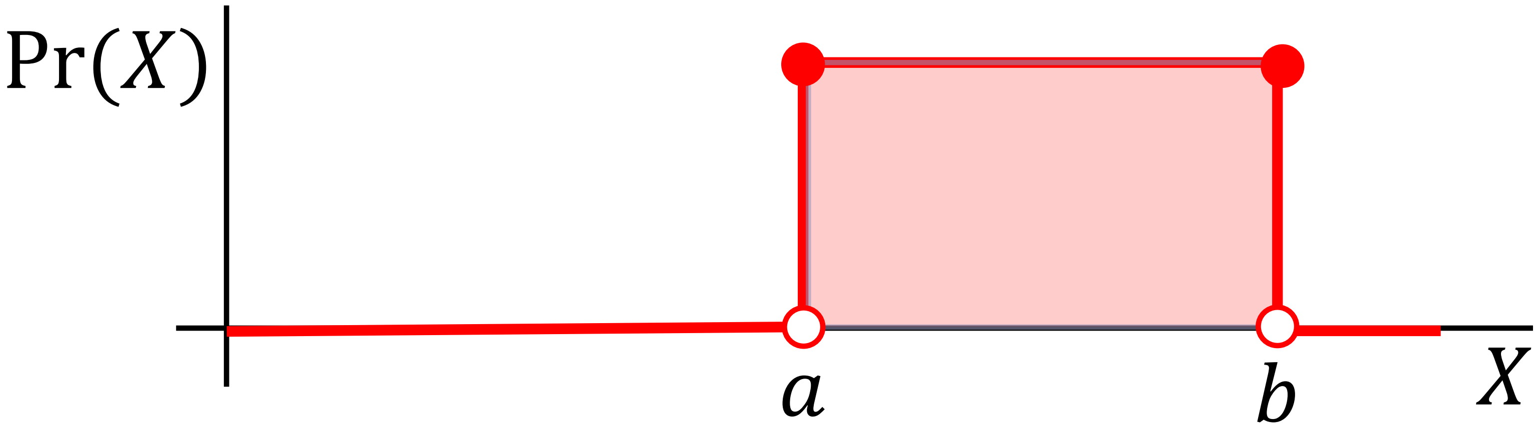
- Also known as **Rectangular Distribution**



- $\Pr(a \leq X \leq b) = 1 = height \cdot widht$

Continuous Uniform Distribution

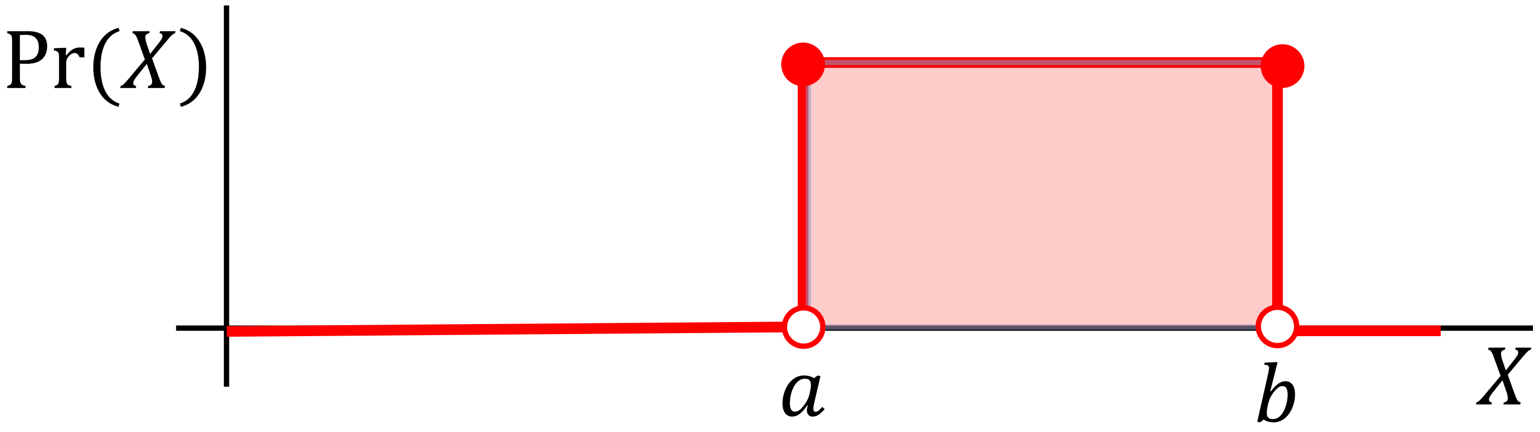
- Also known as **Rectangular Distribution**



- $\Pr(a \leq X \leq b) = 1 = height \cdot (b - a)$

Continuous Uniform Distribution

- Also known as **Rectangular Distribution**

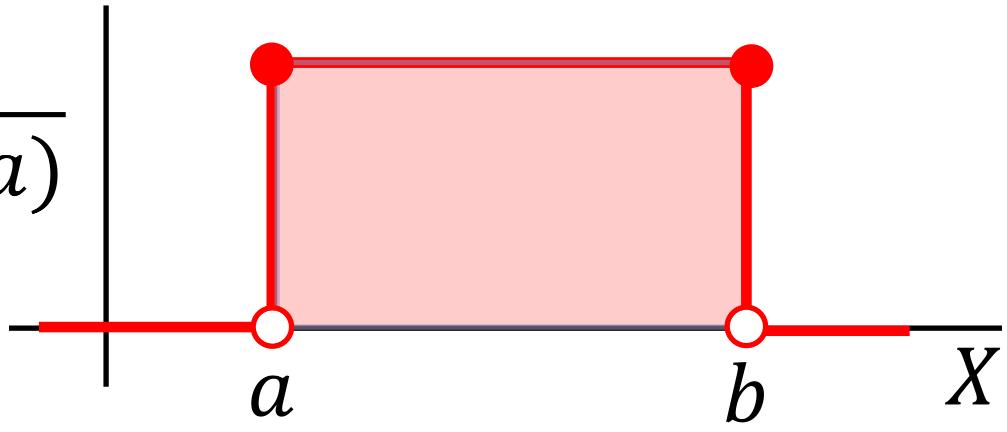


- $\Pr(a \leq X \leq b) = 1 = height \cdot (b - a)$
- $1 = \Pr(X) \cdot (b - a)$

Continuous Uniform Distribution

- Also known as **Rectangular Distribution**

$$\Pr(X) = \frac{1}{(b-a)}$$



- $\Pr(a \le X \le b) = 1 = height \cdot (b-a)$
- $1 = \Pr(X) \cdot (b-a)$
- $\Pr(X) = \frac{1}{(b-a)}$

PDF: Uniform Distribution

- A continuous random variable X is said to have a Uniform distribution over the interval $[a, b]$, shown as $X \sim Uniform(a, b)$, if its **PDF** is given by

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

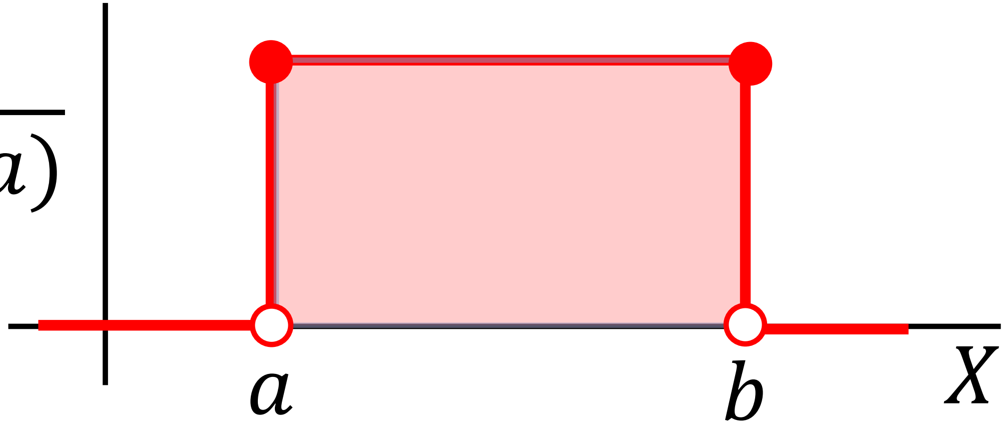
Top Hat question (Join Code: 033357)

- A student waits for a T between zero and 10 minutes, uniformly distributed. What is the probability that a student waits between 4 to 6 minutes?

Top Hat question (Join Code: 033357)

- A student waits for a T between zero and 10 minutes, uniformly distributed. What is the probability that a student waits between 4 to 6 minutes?

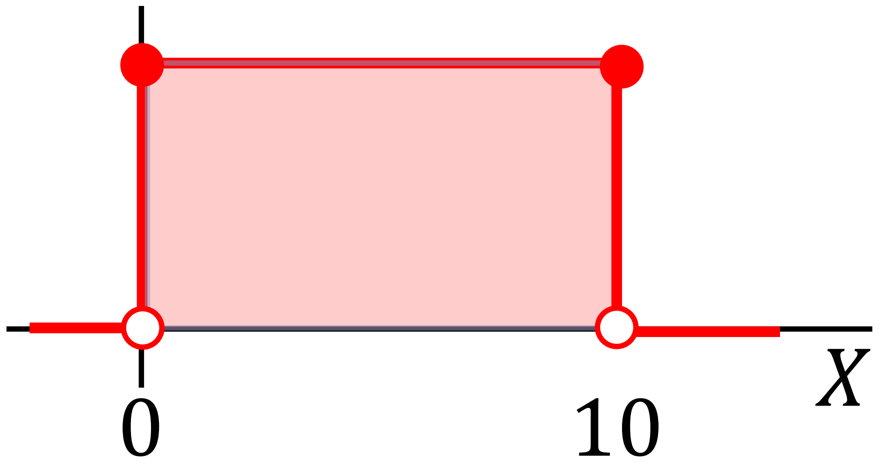
$$\Pr(X) = \frac{1}{(b-a)}$$



Top Hat question (Join Code: 033357)

- A student waits for a T between zero and 10 minutes, uniformly distributed. What is the probability that a student waits between 4 to 6 minutes?

$$\Pr(4 \leq X \leq 6) = ?$$

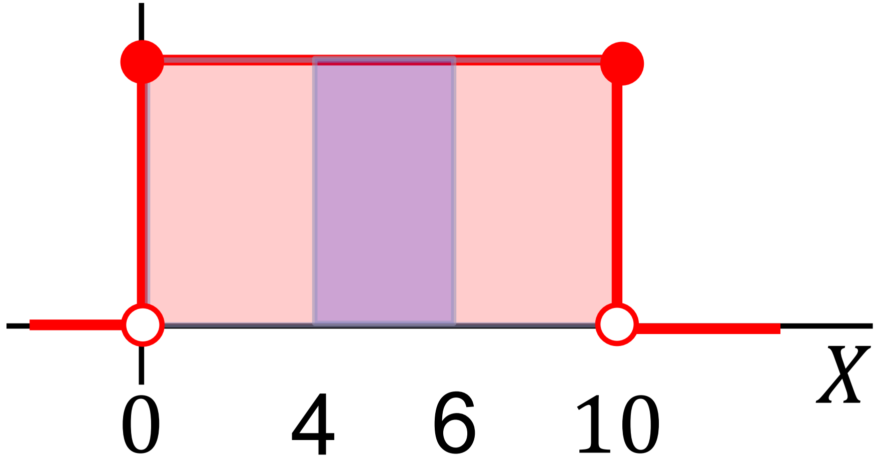


CS 237

Top Hat question (Join Code: 033357)

- A student waits for a T between zero and 10 minutes, uniformly distributed. What is the probability that a student waits between 4 to 6 minutes?

$$\Pr(4 \le X \le 6) = ?$$

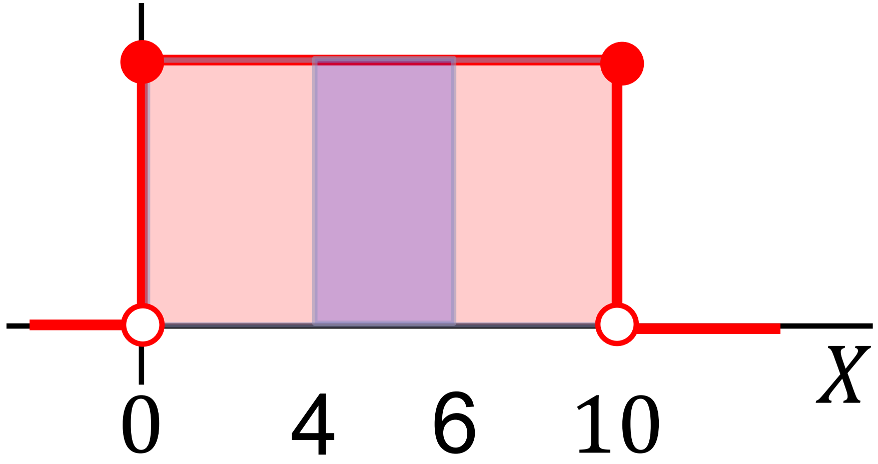


Top Hat question (Join Code: 033357)

- A student waits for a T between zero and 10 minutes, uniformly distributed. What is the probability that a student waits between 4 to 6 minutes?

$$\Pr(4 \leq X \leq 6) = ?$$

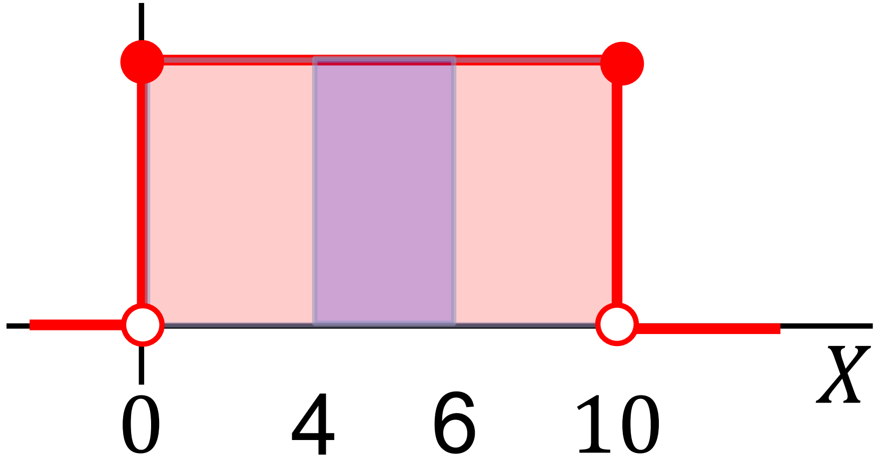
$$= \frac{6-4}{10-0}$$



Top Hat question (Join Code: 033357)

- A student waits for a T between zero and 10 minutes, uniformly distributed. What is the probability that a student waits between 4 to 6 minutes?

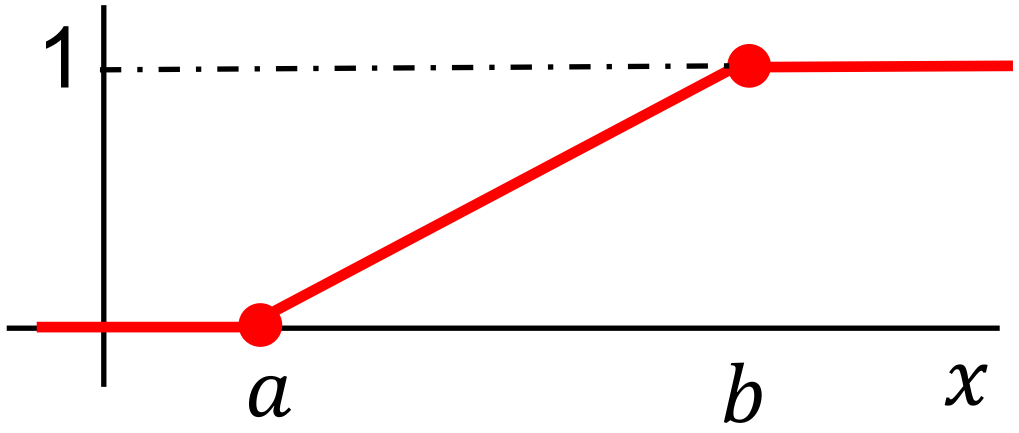
$$\Pr(4 \le X \le 6) = ?$$
$$= \frac{6-4}{10-0}$$
$$= 0.2$$



CDF: Uniform Distribution

- By definition $F_X(x) = \Pr(X \leq x)$. We have $F_X(x) = 0$, for $x < a$ and $F_X(x) = 1$, for $x > b$
- For $a \leq x \leq b$ we have

$$F_X(x) = \Pr(X \leq x)$$
$$= \Pr(X \in [a, x])$$
$$= \frac{x-a}{b-a}$$



CDF: Uniform Distribution

- By definition $F_X(x) = \Pr(X \leq x)$. We have $F_X(x) = 0$, for $x < a$ and $F_X(x) = 1$, for $x > b$
- For $a \leq x \leq b$ we have
- $F_X(x) = \frac{x-a}{b-a}$
- To summarize

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

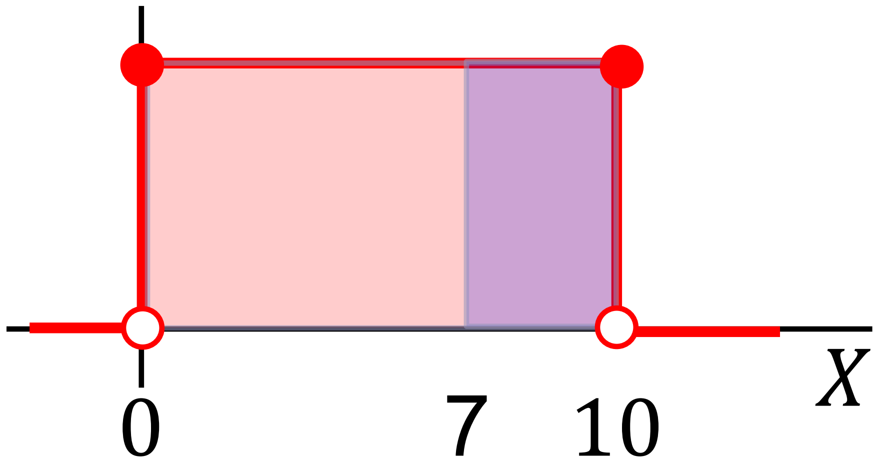
Top Hat question (Join Code: 033357)

A student waits for a T between zero and 10 minutes, uniformly distributed. What is the probability that a student waits at least 7 minutes?

Top Hat question (Join Code: 033357)

A student waits for a T between zero and 10 minutes, uniformly distributed. What is the probability that a student waits at least 7 minutes?

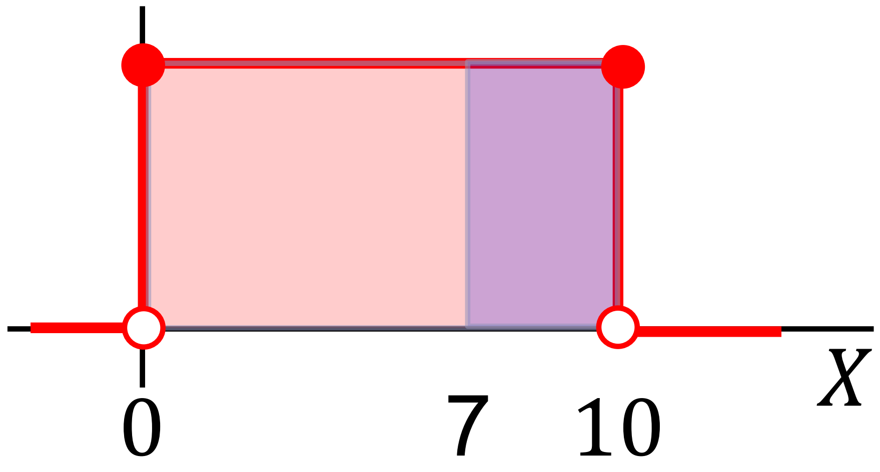
$$\Pr(X \geq 7) = ?$$



Top Hat question (Join Code: 033357)

A student waits for a T between zero and 10 minutes, uniformly distributed. What is the probability that a student waits at least 7 minutes?

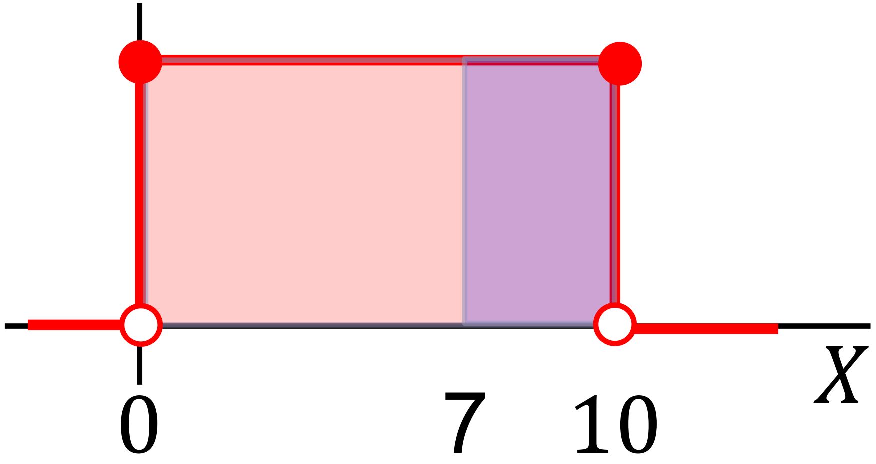
$$1 - \Pr(X \leq 7) = ?$$



Top Hat question (Join Code: 033357)

A student waits for a T between zero and 10 minutes, uniformly distributed. What is the probability that a student waits at least 7 minutes?

$$1 - \Pr(X \leq 7) =$$
$$= 1 - F_X(7)$$
$$= 1 - \frac{7}{10}$$
$$= 1 - 0.7$$
$$= 0.3$$



Top Hat question (Join Code: 033357)

A student waits for a T between zero and 10 minutes, uniformly distributed. What is the expected waiting time?

Top Hat question (Join Code: 033357)

A student waits for a T between zero and 10 minutes, uniformly distributed. What is the expected waiting time?

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x \cdot f(x)dx = \int_a^b \frac{x}{b-a}dx$$

$$= \frac{1}{b-a} \cdot \frac{x^2}{2}\bigg|_a^b = \frac{b^2 - a^2}{2 \cdot (b-a)}$$

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

Variance and Standard Deviation

A student waits for a T between zero and 10 minutes, uniformly distributed. What is the variance and standard deviation?

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x)dx = \frac{a^2 + a \cdot b + b^2}{3}$$

$$Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

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

$$\sigma = \sqrt{Var(X)} = \frac{b-a}{\sqrt{12}}$$

Conditional Uniform Probability

What is the probability that a student waits at most 8 minutes given that they waited at least 6 minutes?

$$\Pr(X \leq 8 | X \geq 6) = \frac{\Pr(X \leq 8 \cap X \geq 6)}{\Pr(X \geq 6)}$$

$$= \frac{\Pr(6 \leq X \leq 8)}{1 - \Pr(X \leq 6)} = \frac{F_X(8) - F_X(6)}{1 - F_X(6)}$$

$$= \frac{{}^{2}/_{10}}{4/10} = \frac{1}{2} = 0.5$$

Continuous Uniform Distribution

$$X \sim \text{Uniform}(a, b)$$

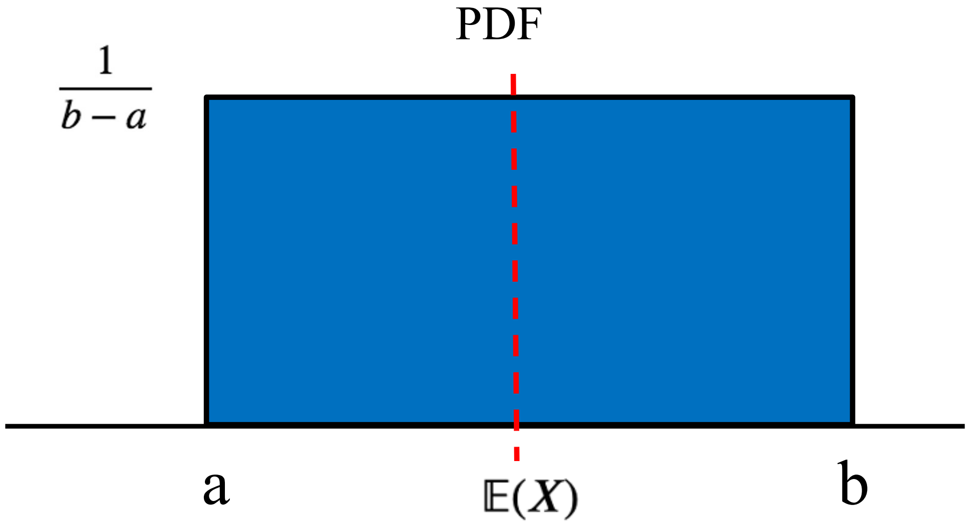
PDF: $$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

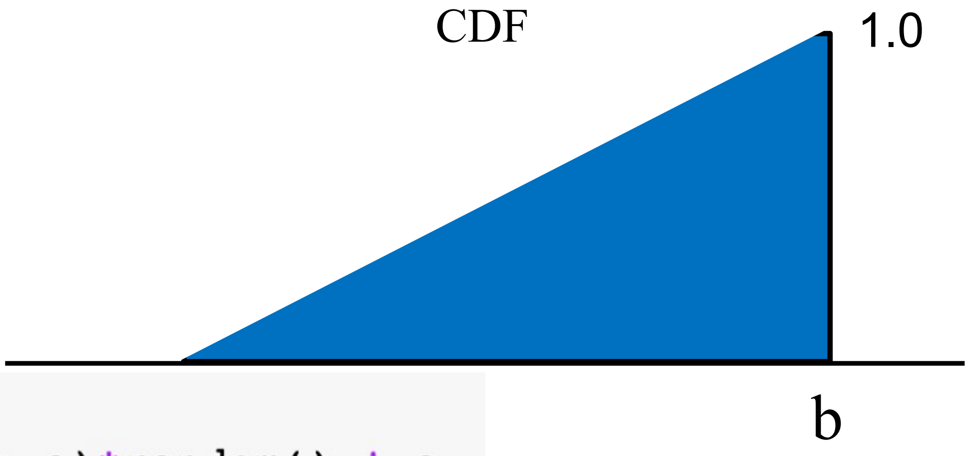
CDF: $$F_X(x) = \begin{cases} 0 & \text{if } x < a \\ \frac{x-a}{b-a} & \text{if } a \le x \le b \\ 1 & \text{if } x > b \end{cases}$$

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

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

$$\sigma_X = \frac{b-a}{\sqrt{12}}$$






```
def X(a,b):
    return (b-a)*random() + a
```

Continuous Uniform Distribution

Reminder: The ONLY way to calculate probabilities for continuous RVs is with intervals, usually using the CDF.

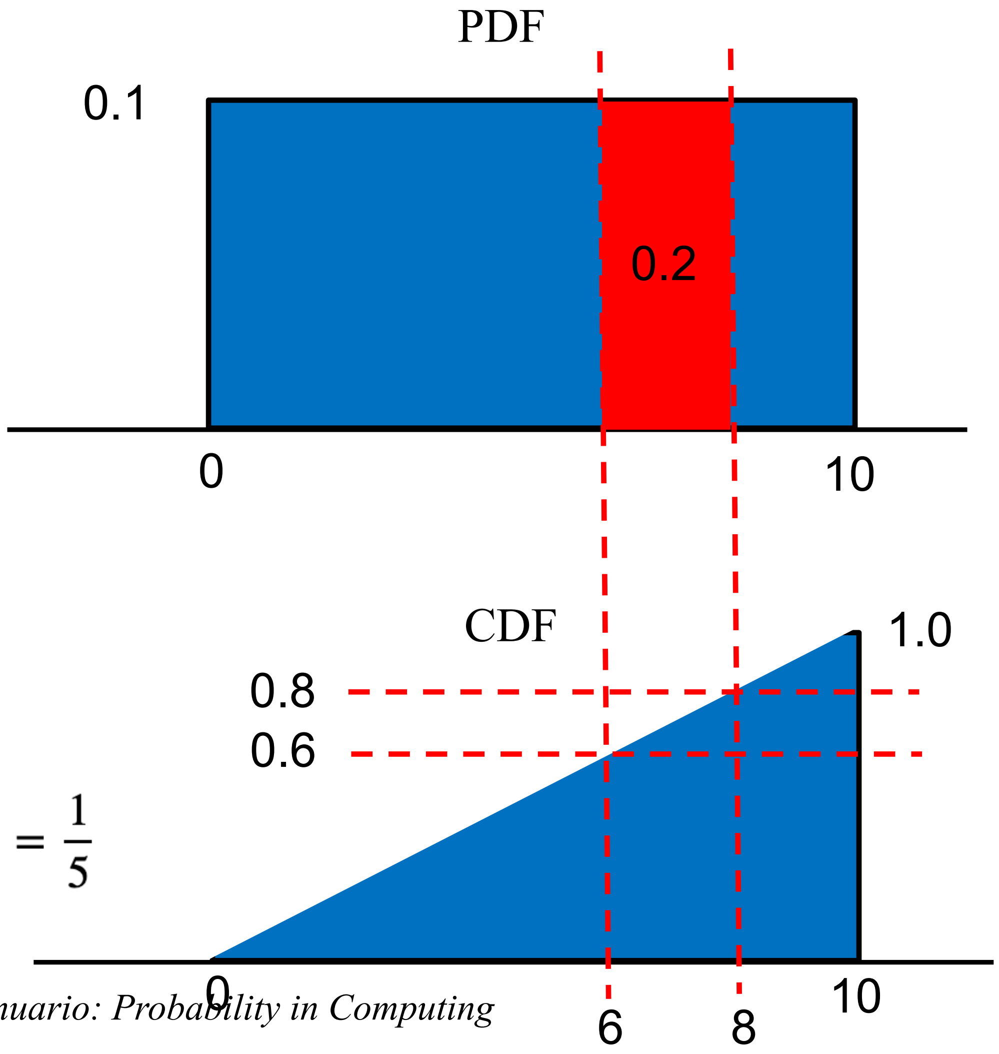
Example:

$$X \sim \text{Uniform}(0, 10)$$

$$f_X(x) = \begin{cases} \frac{1}{10} & 0 \le x \le 10 \\ 0 & \text{otherwise} \end{cases}$$

$$F_X(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{10} & 0 \le x \le 10 \\ 1 & x > 10 \end{cases}$$

$$\Pr(6 \le X \le 8) = F_X(8) - F_X(6) = \frac{8}{10} - \frac{6}{10} = \frac{1}{5}$$

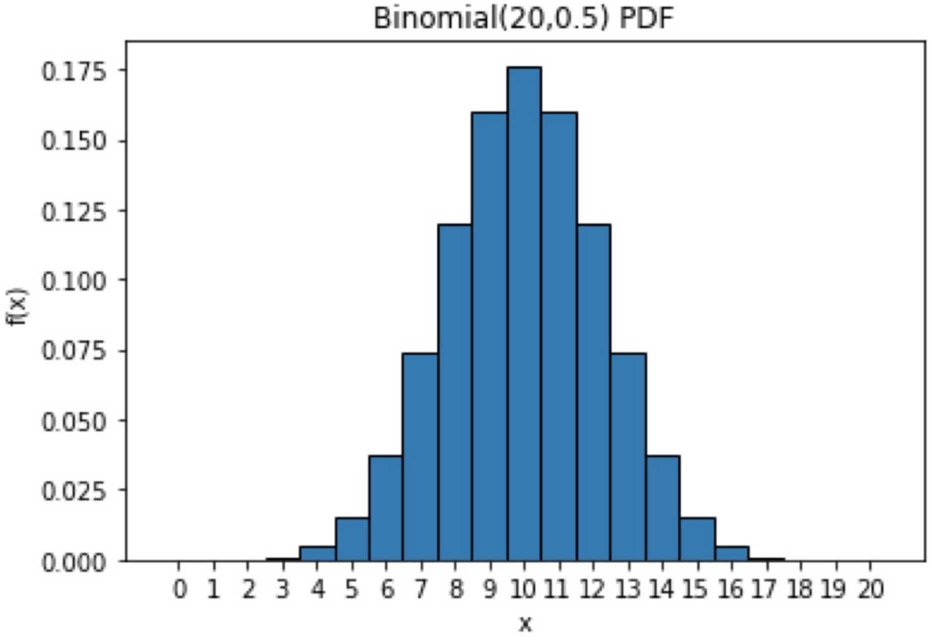


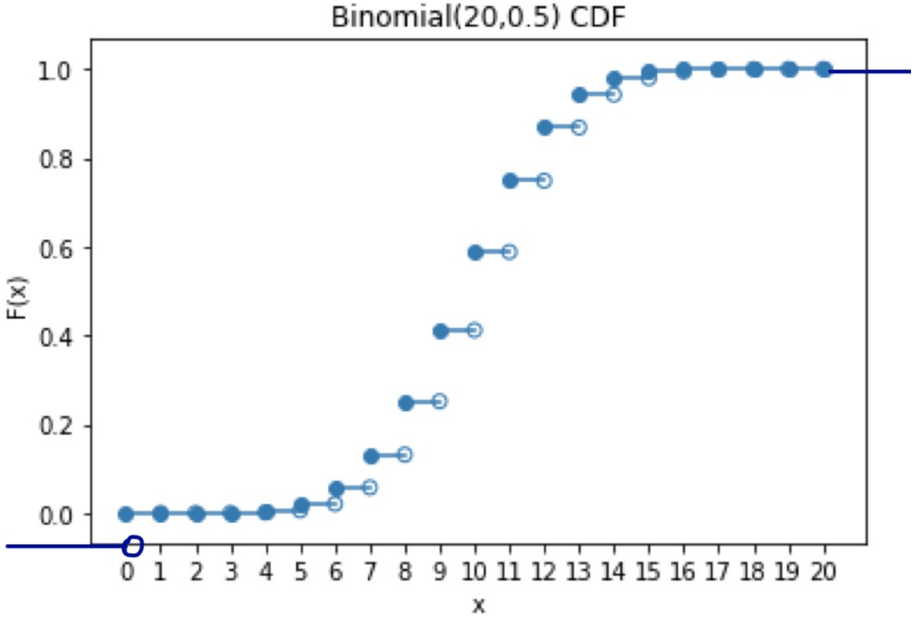
Continuous Uniform Distribution

https://www.causeweb.org/cause/resources/fun/cartoons/uniform-distribution

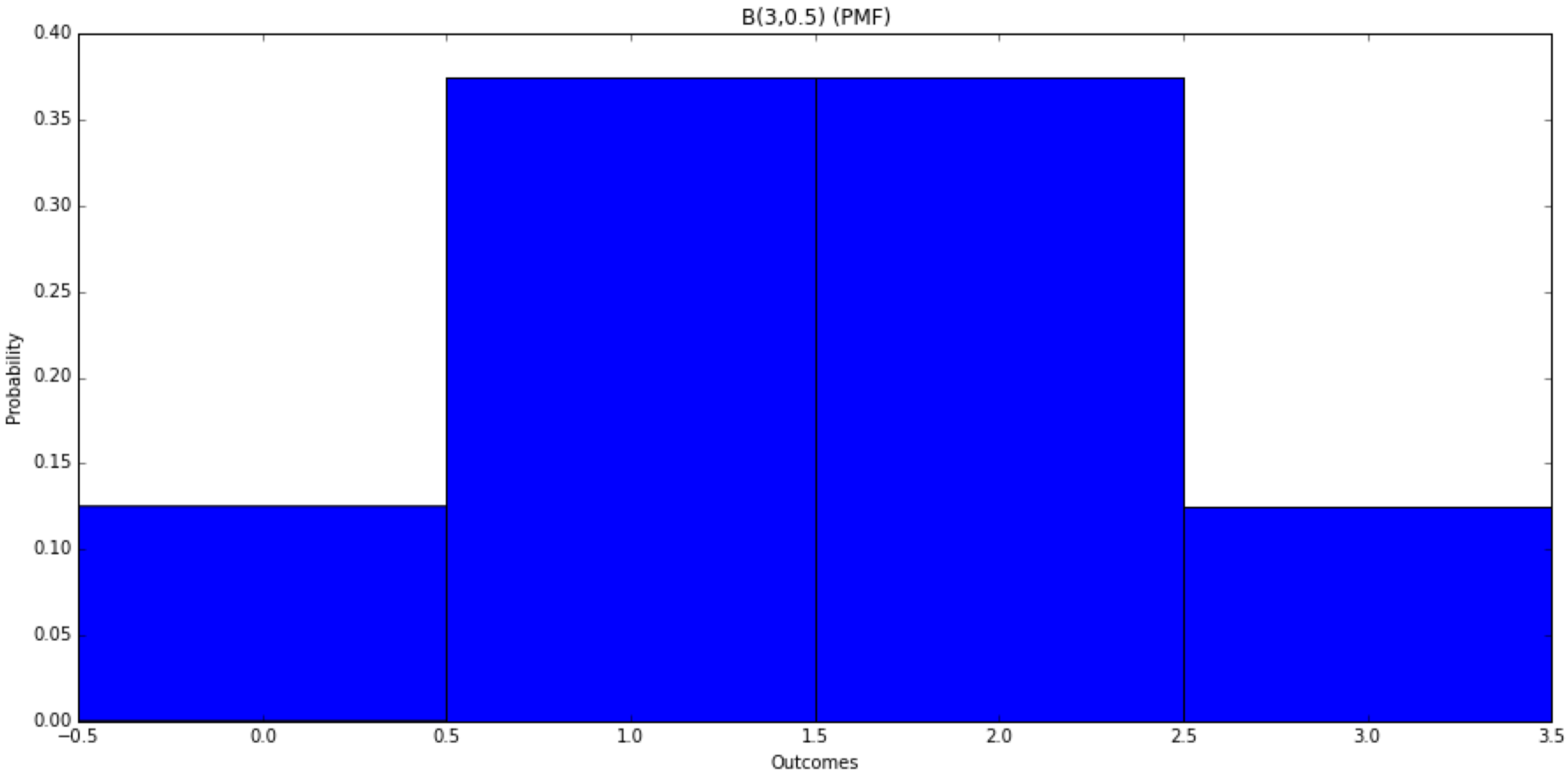
Recall: Binomial Distribution

- X ~ Binomial(n, p) example

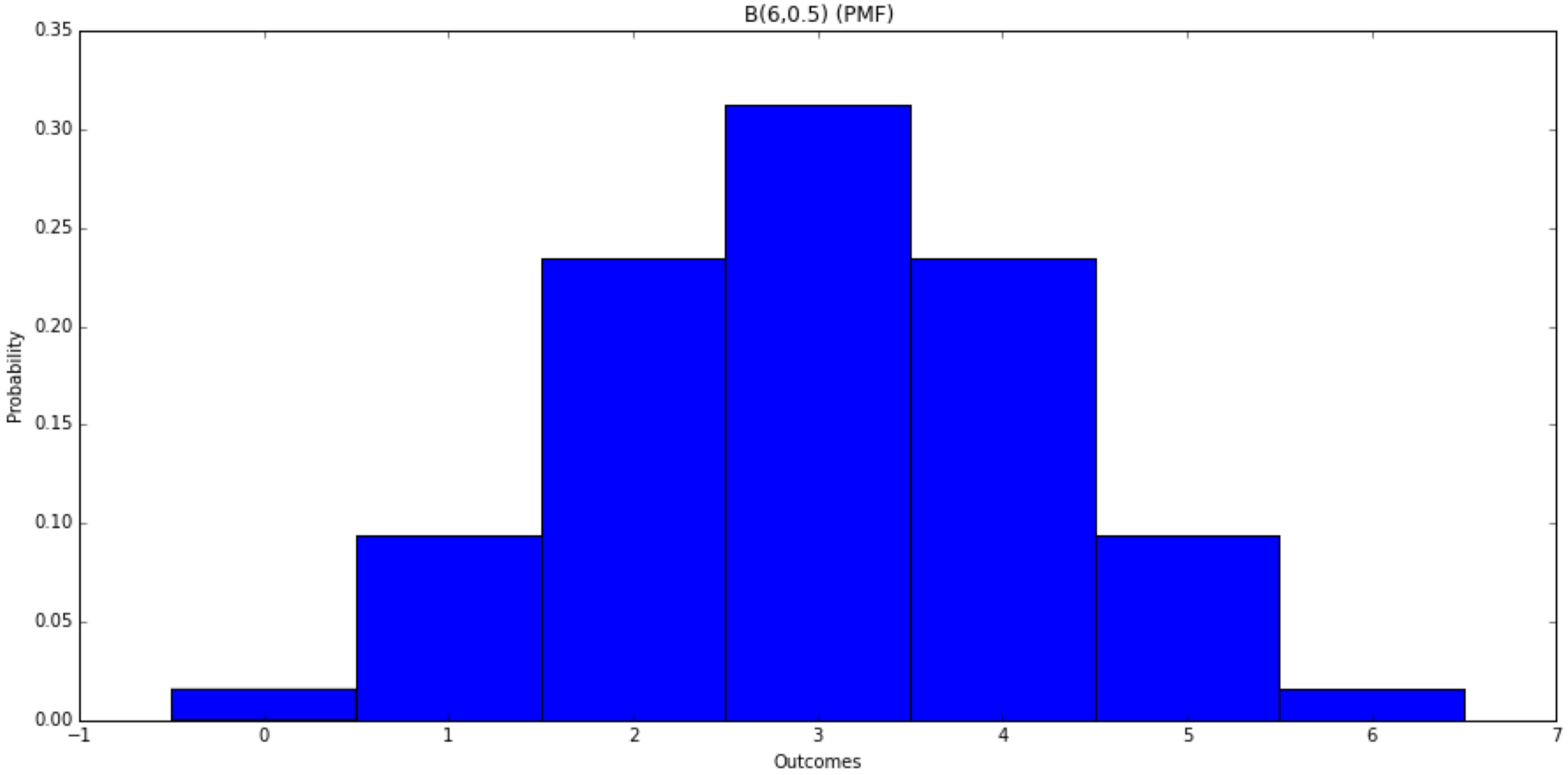




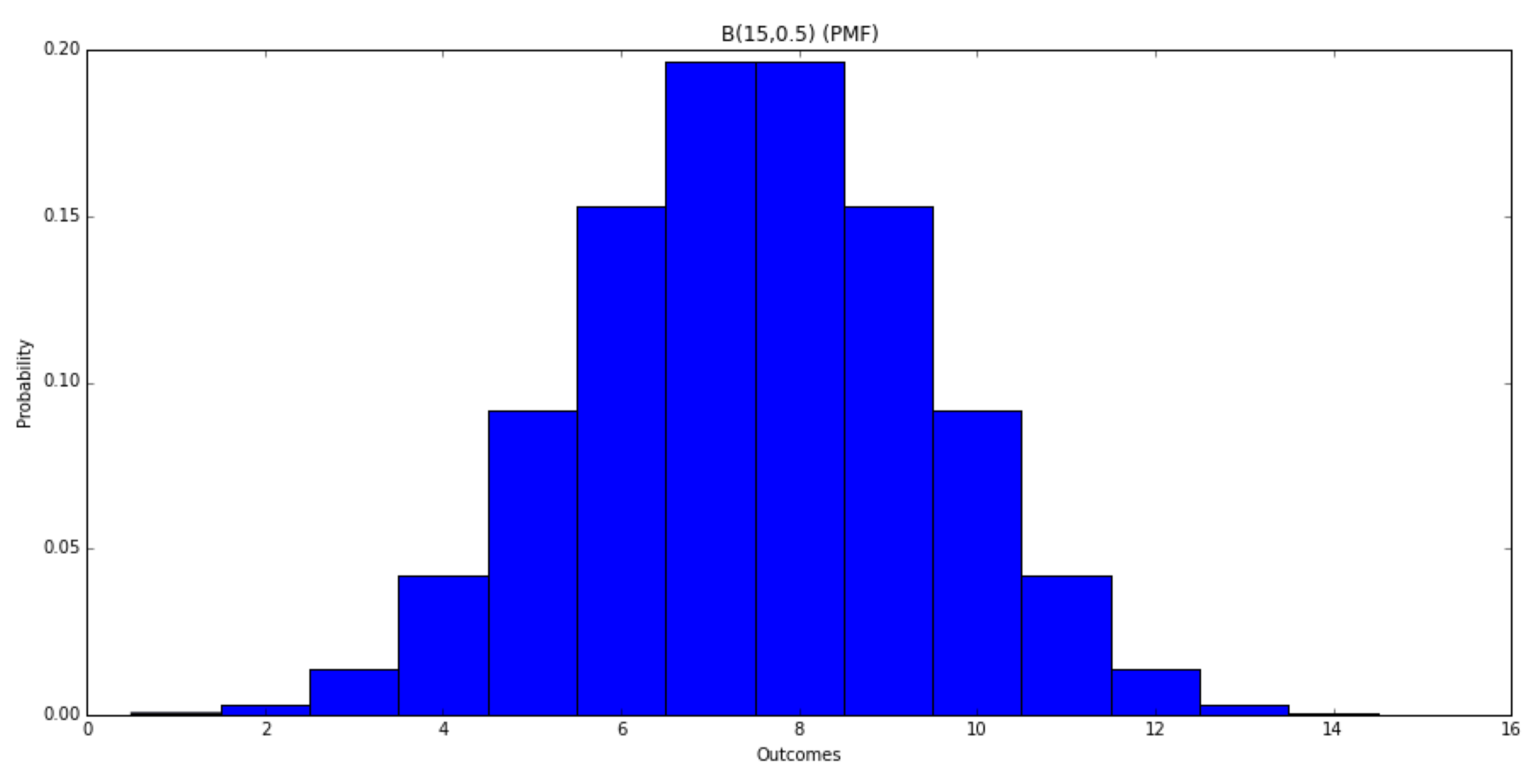
$X \sim Binomial(n, p)$ for $p = 0.5$ as $n \to \infty$



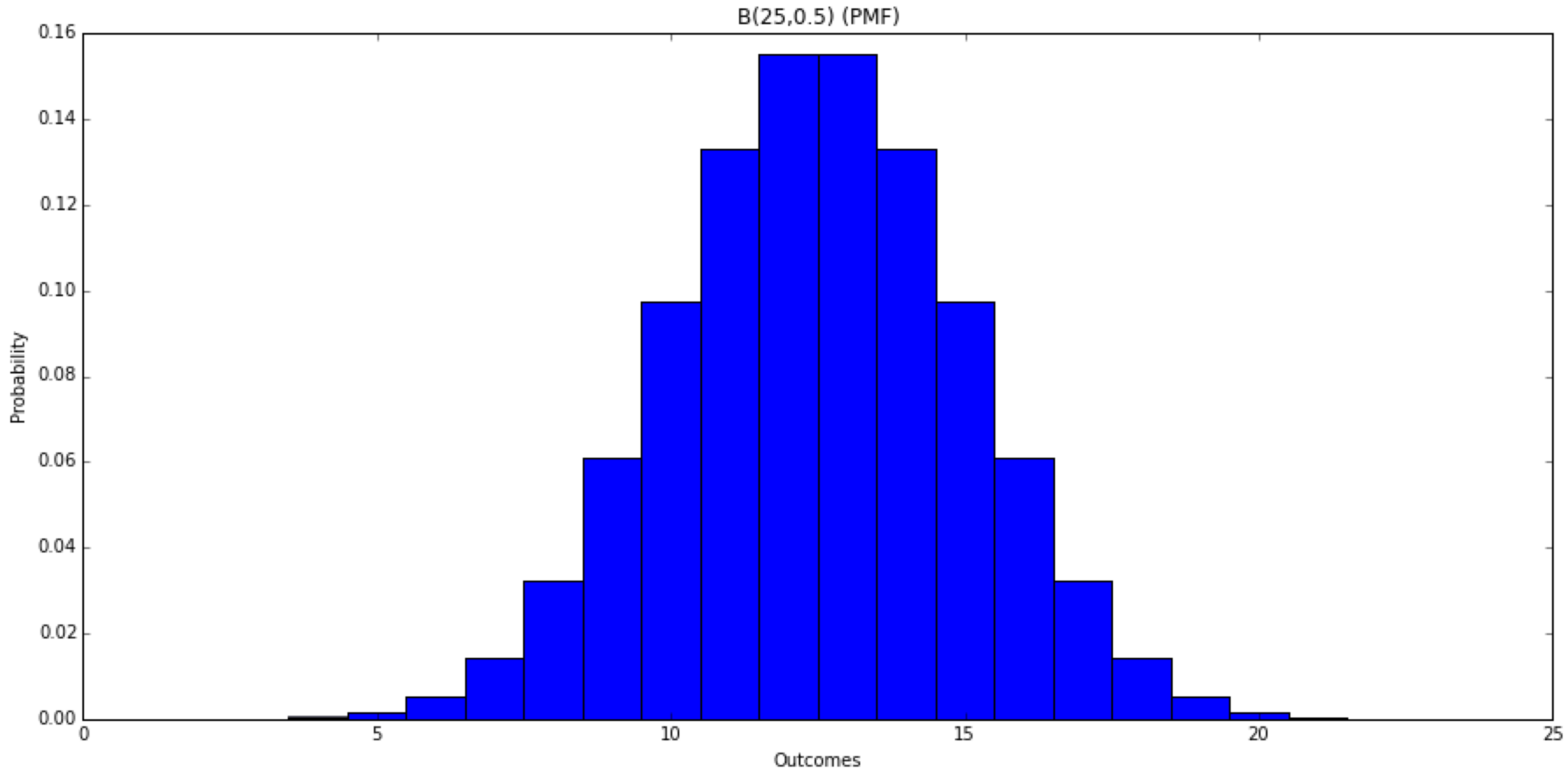
$X \sim Binomial(n, p)$ for $p = 0.5$ as $n \to \infty$



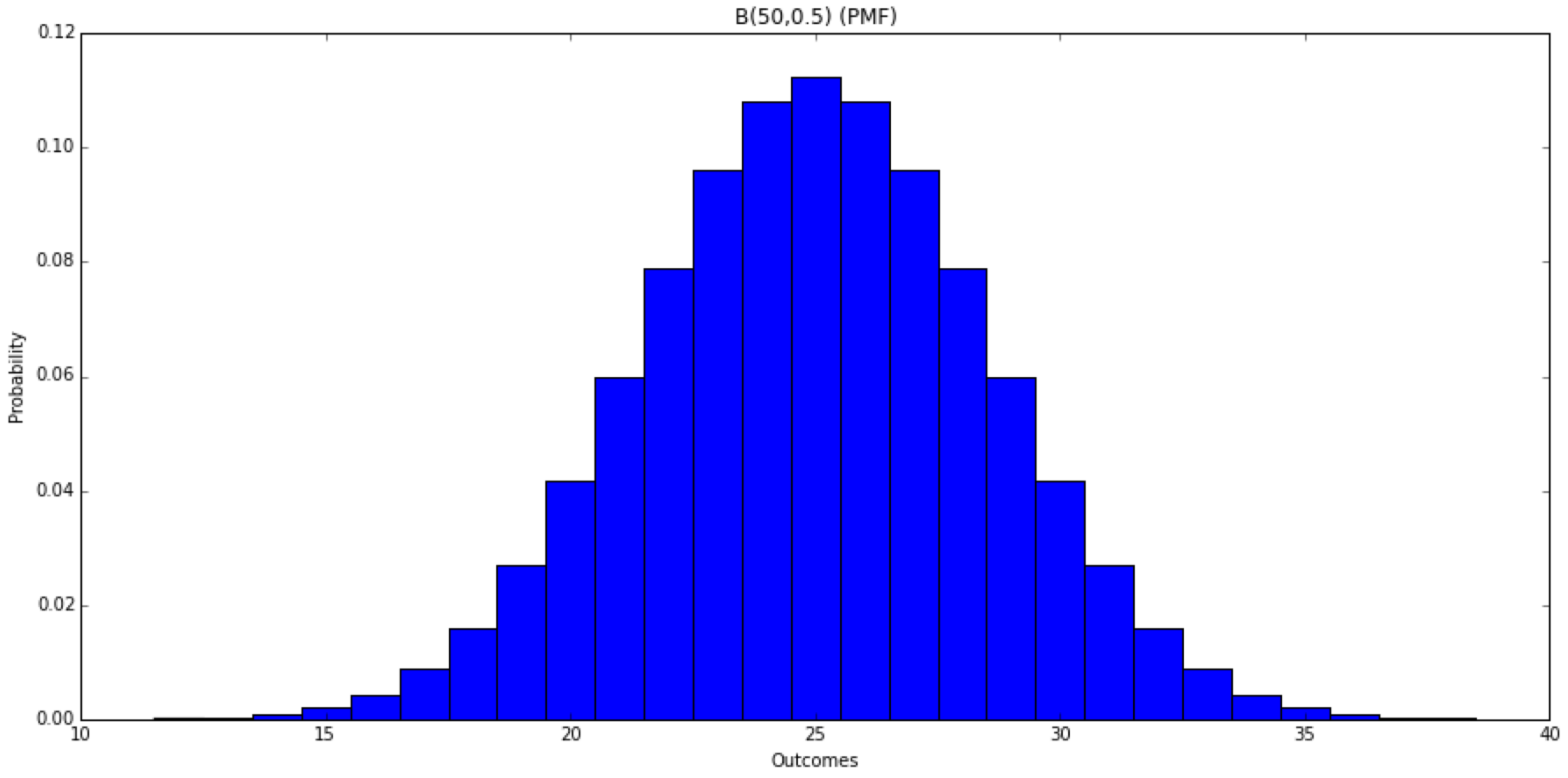
$X \sim Binomial(n, p)$ for $p = 0.5$ as $n \to \infty$



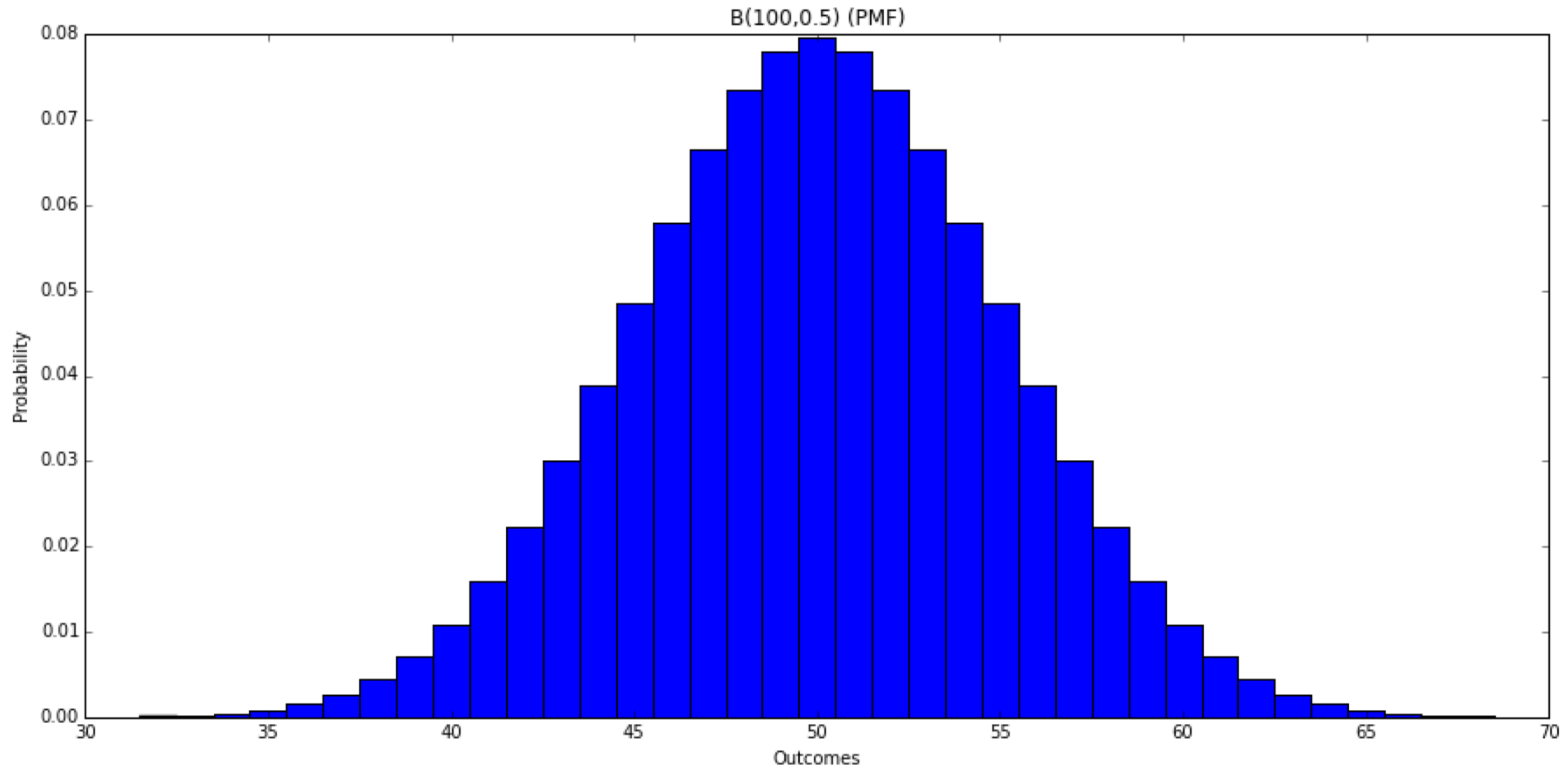
$X \sim Binomial(n, p)$ for $p = 0.5$ as $n \to \infty$



$X \sim Binomial(n, p)$ for $p = 0.5$ as $n \to \infty$

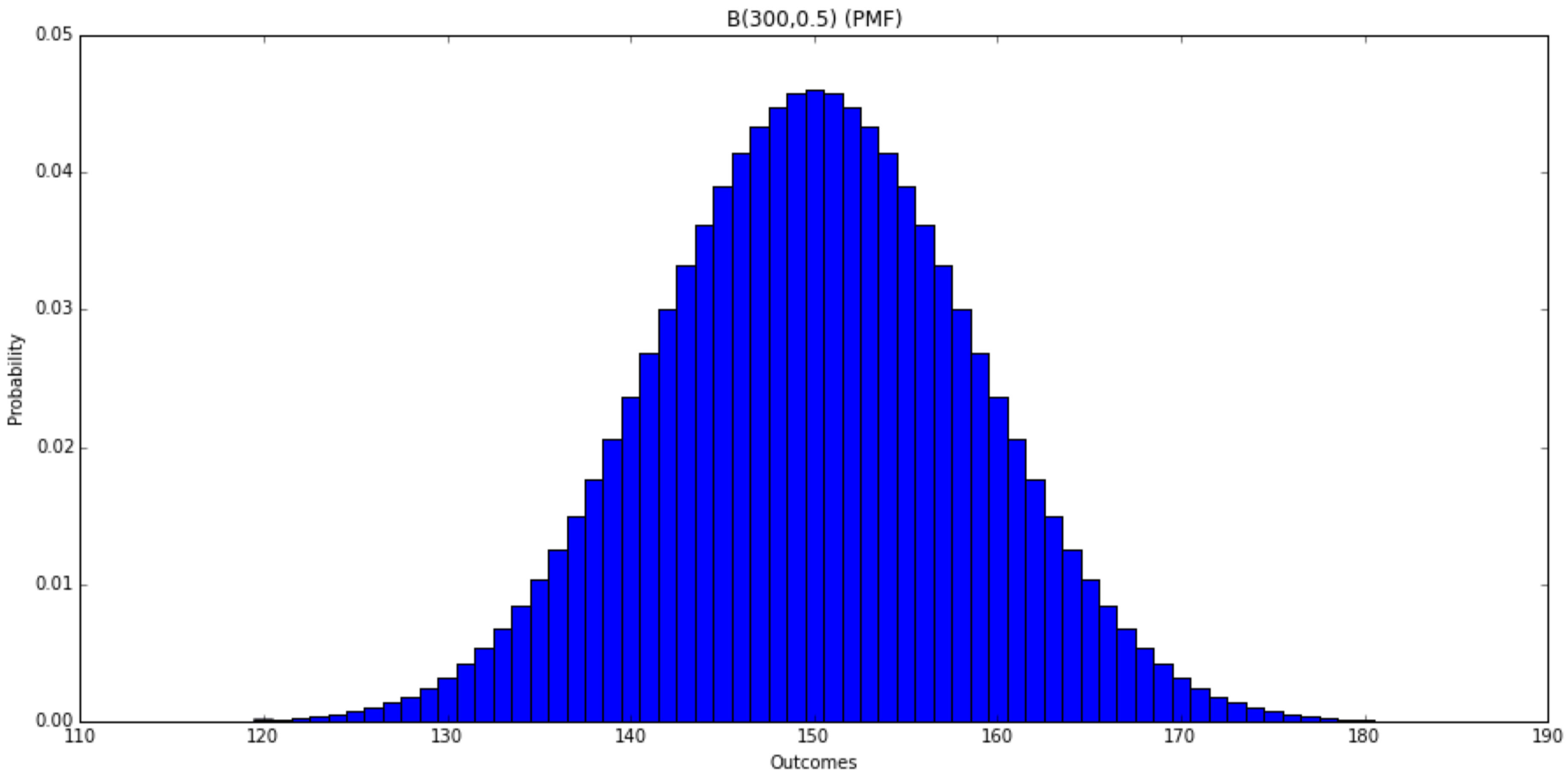


$X \sim Binomial(n, p)$ for $p = 0.5$ as $n \to \infty$

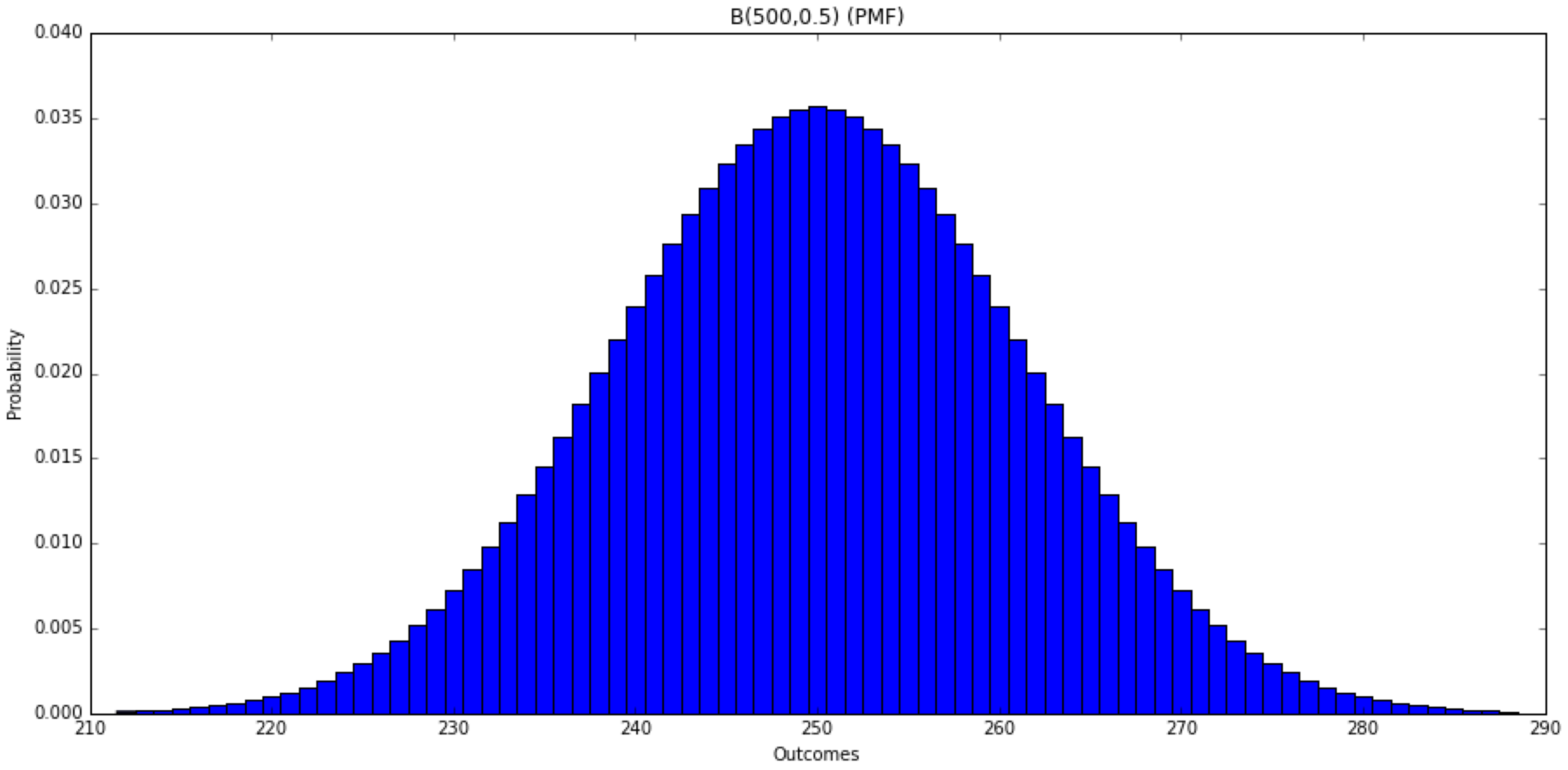


CS 237

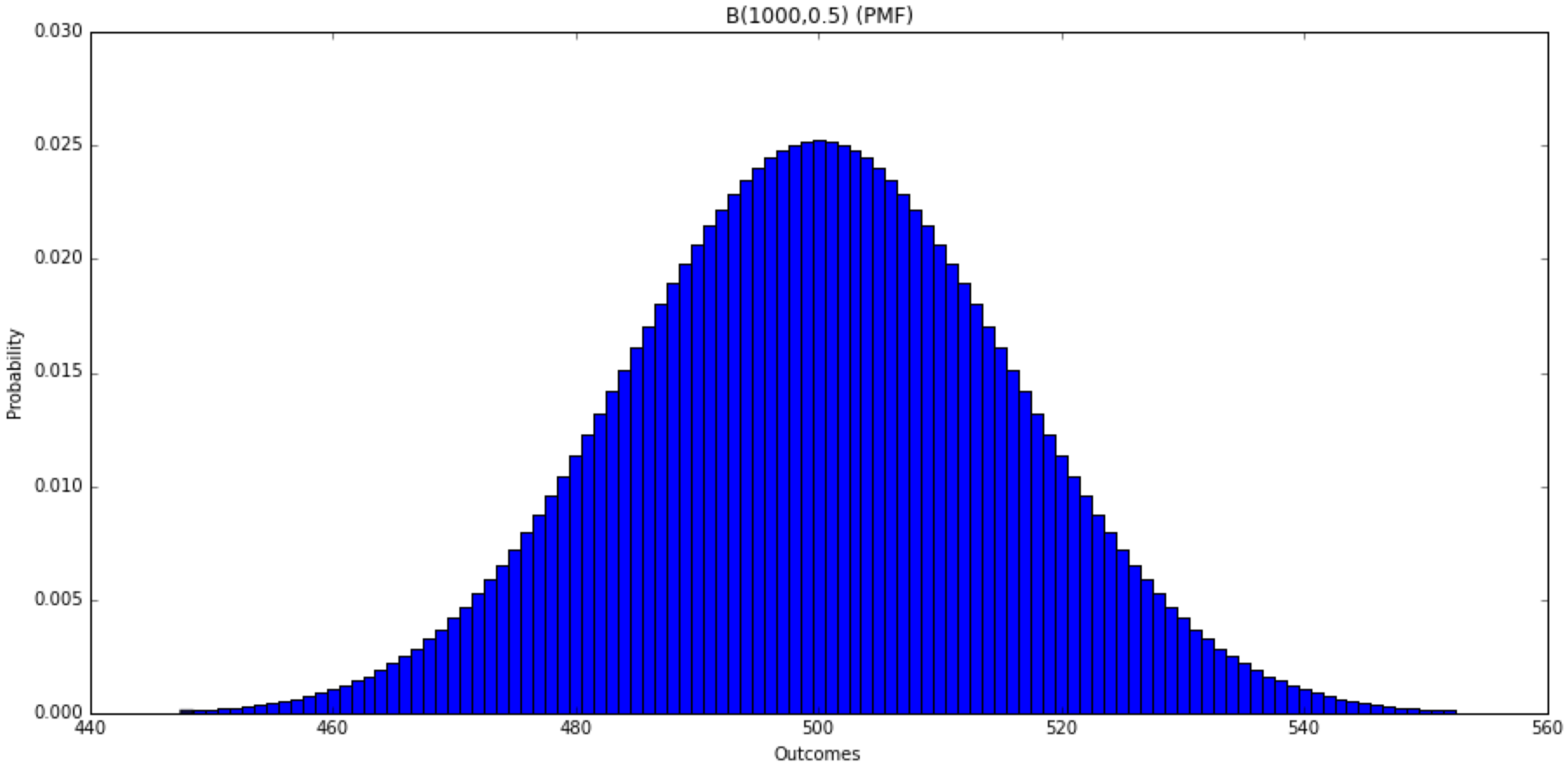
$X \sim Binomial(n, p)$ for $p = 0.5$ as $n \to \infty$



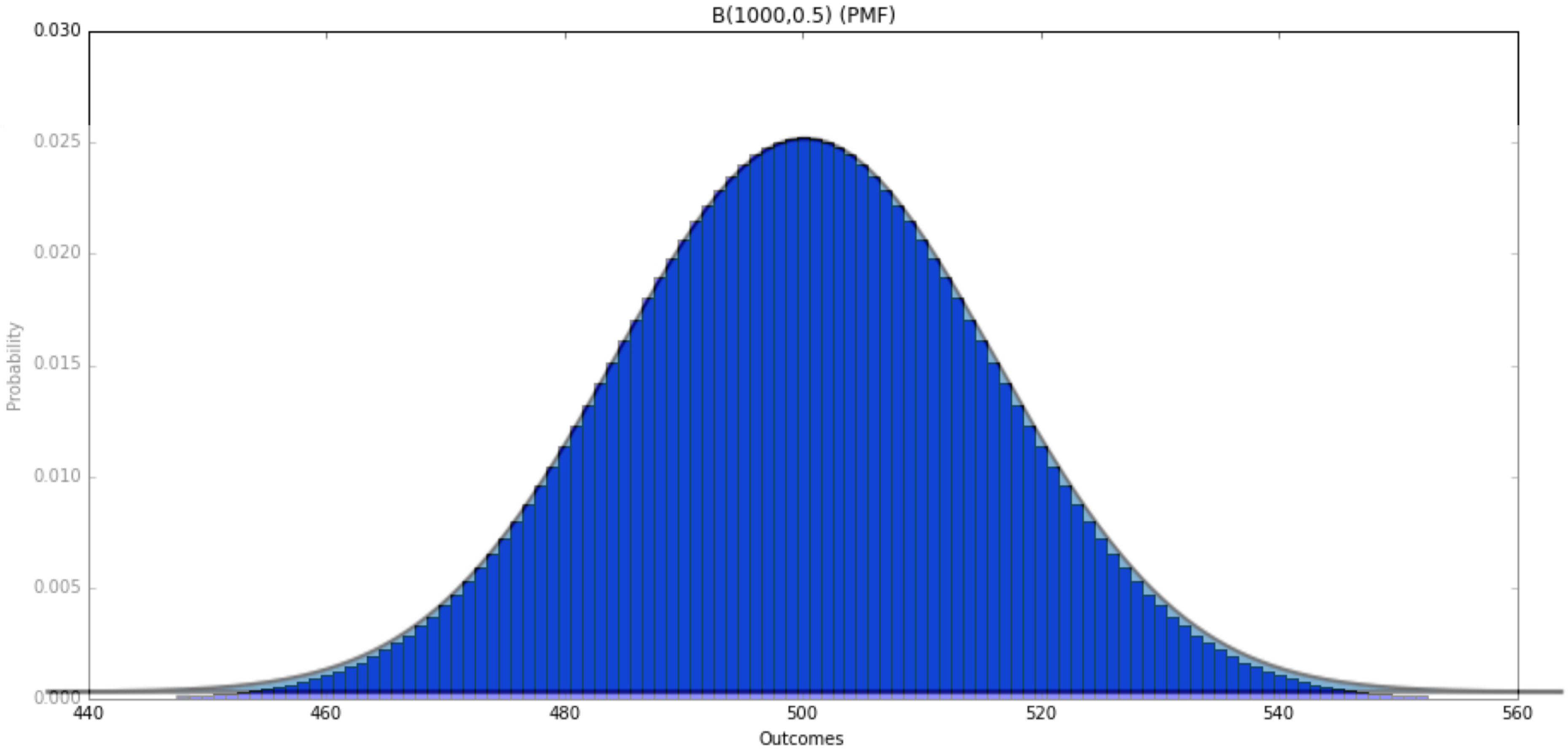
$X \sim Binomial(n, p)$ for $p = 0.5$ as $n \to \infty$



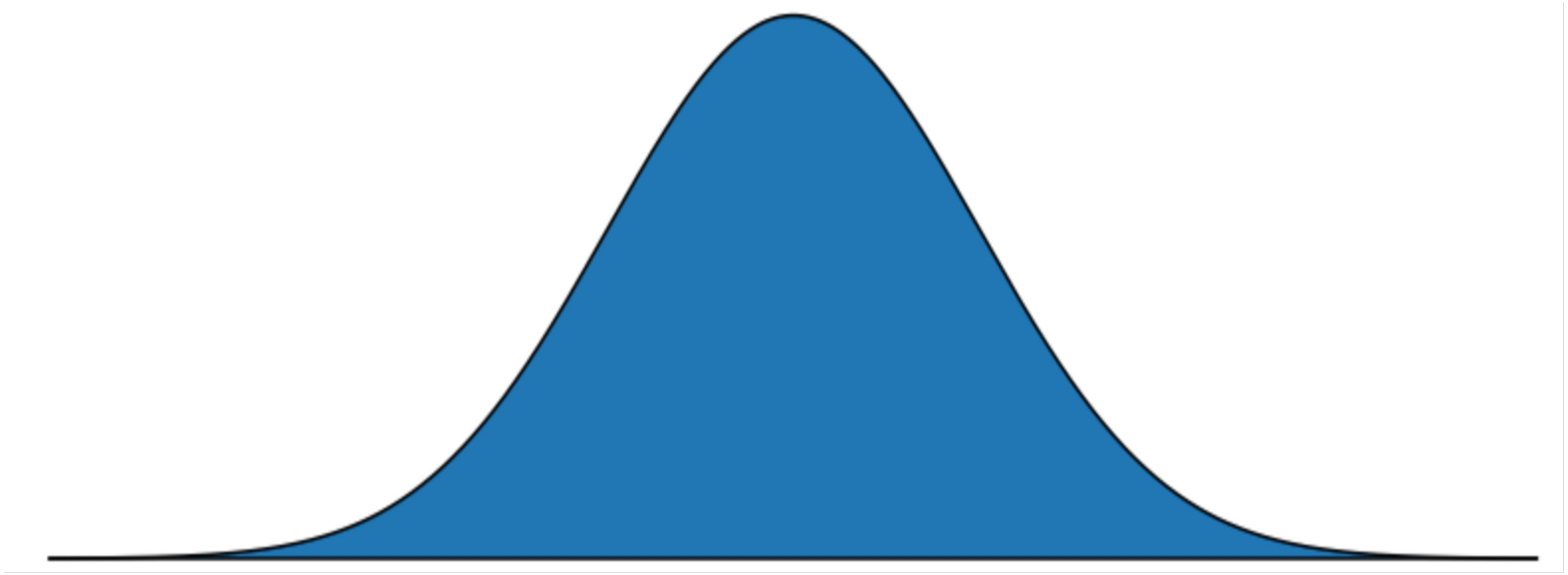
$X \sim Binomial(n, p)$ for $p = 0.5$ as $n \to \infty$



$X \sim Binomial(n, p)$ for $p = 0.5$ as $n \to \infty$

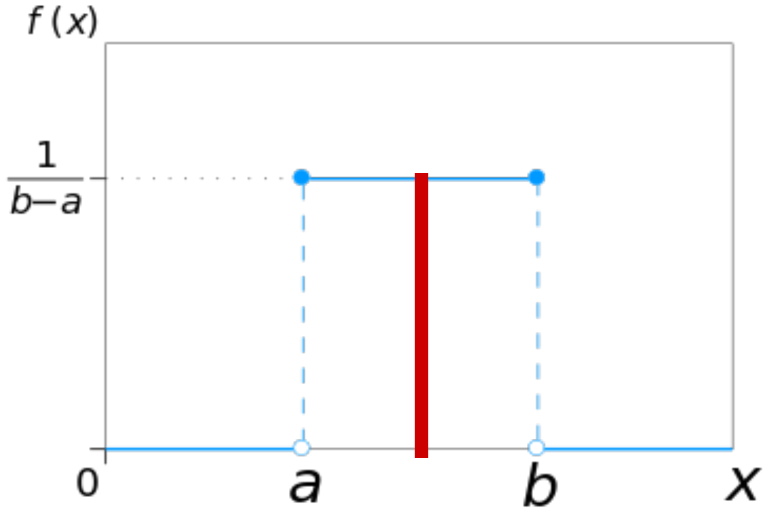


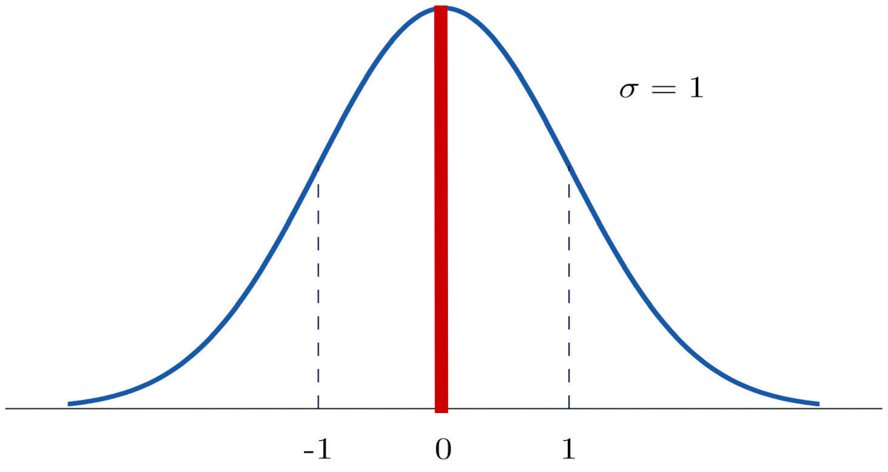
$X \sim Binomial(n, p)$ for $p = 0.5$ as $n \to \infty$



Symmetric probability distributions

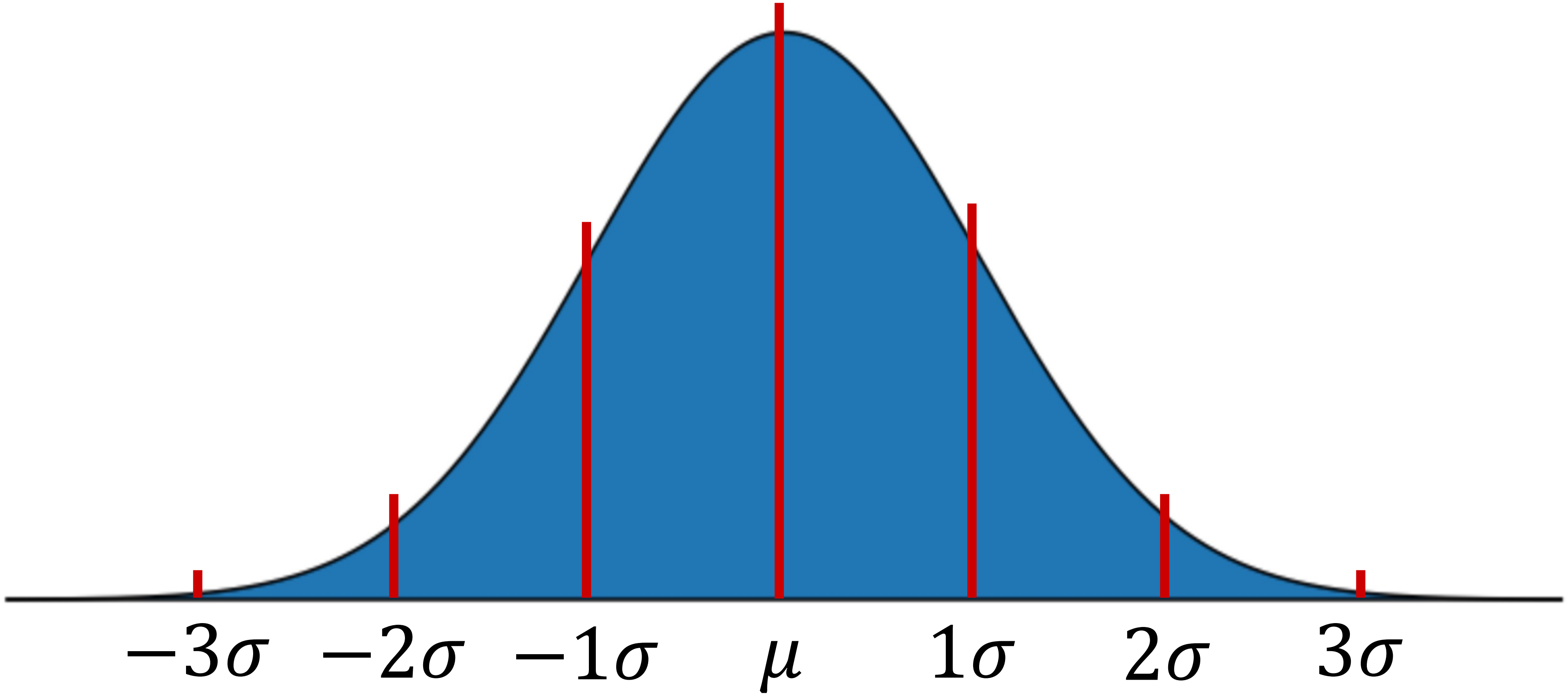
- A symmetric probability distribution is a probability distribution which is unchanged when its PDF is reflected around a vertical line at some value of the random variable represented by the distribution.



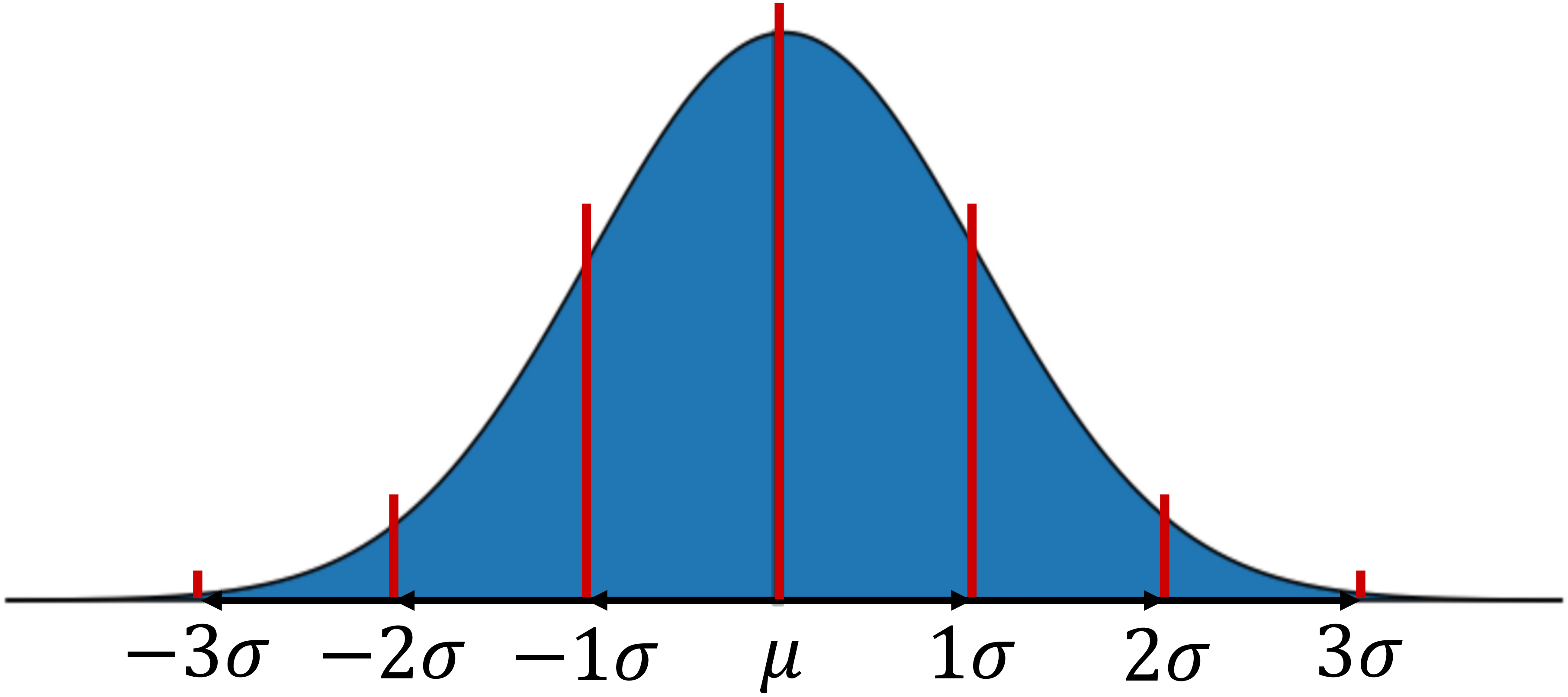


Normal Distribution

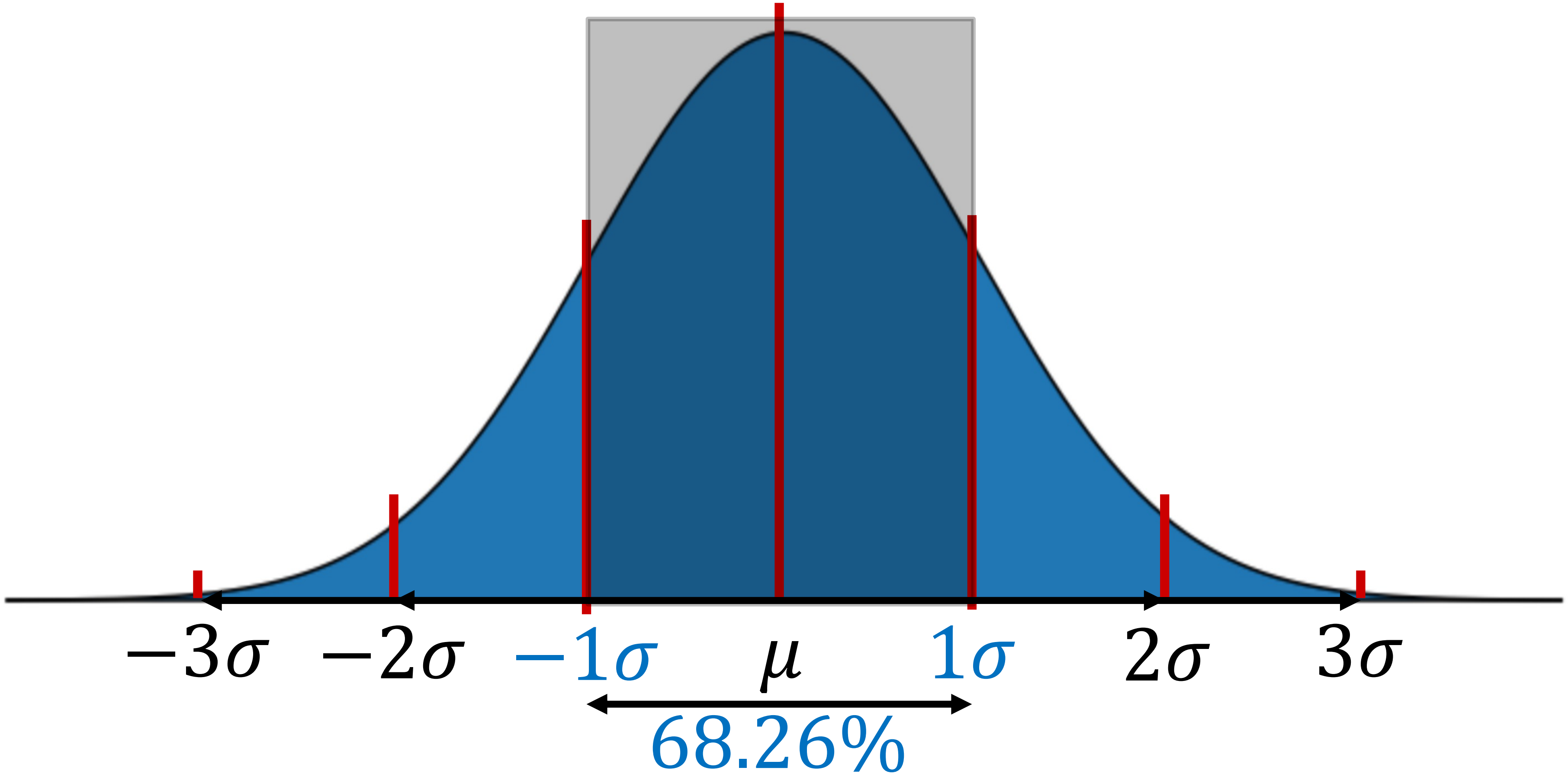
- The most important probability distribution
- Also known as Gaussian or Bell-shaped curve



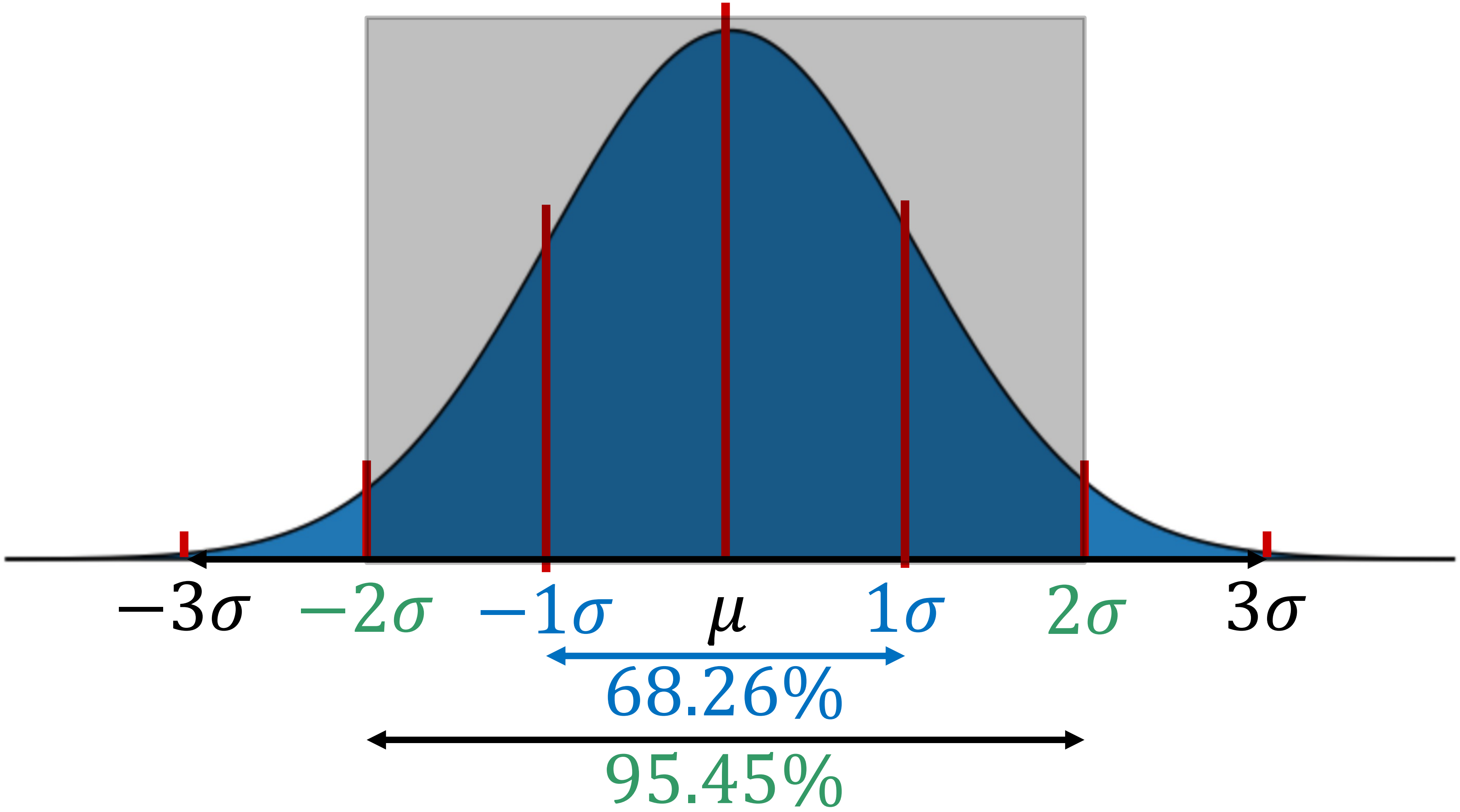
68-95-99.7% rule



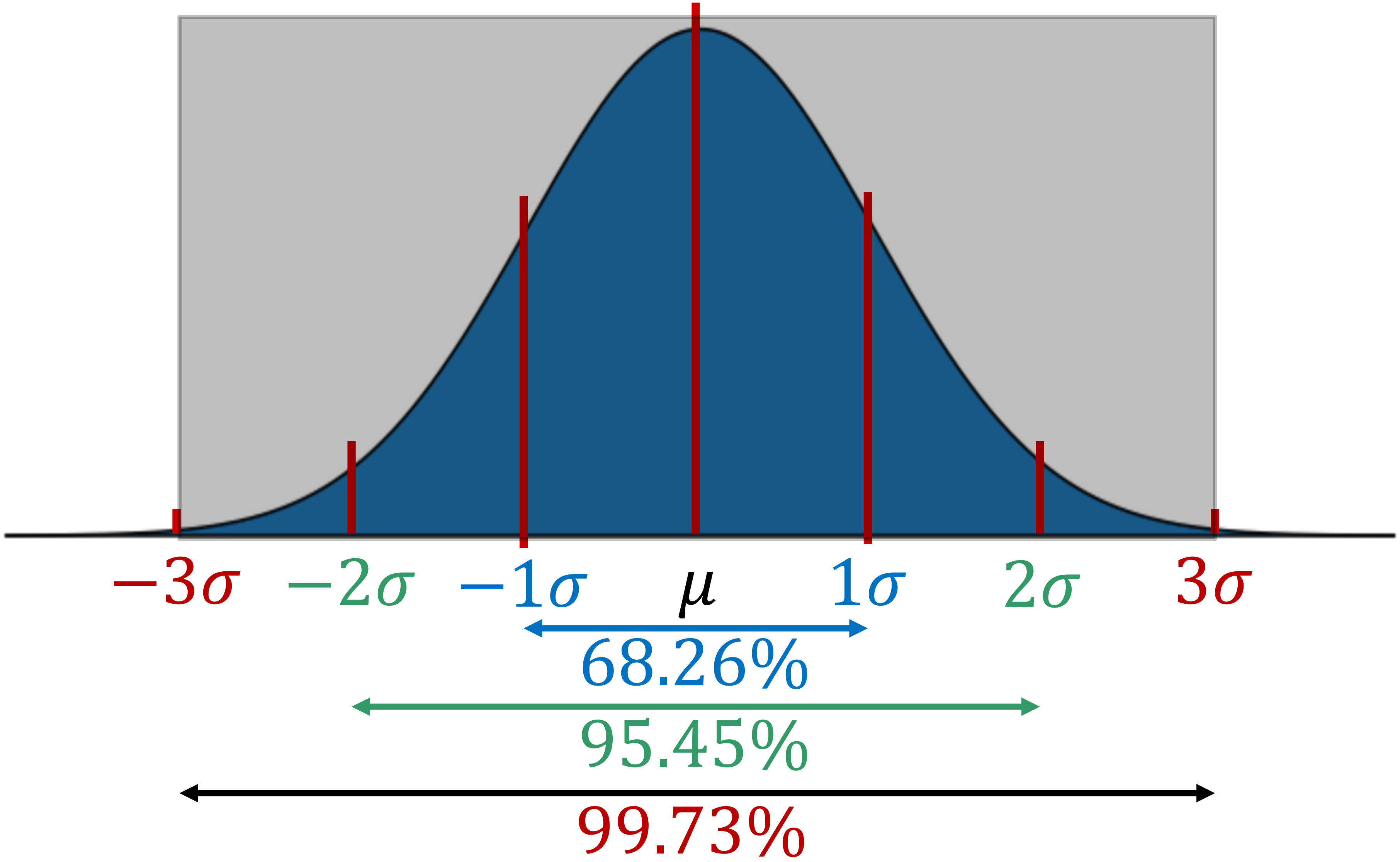
68-95-99.7% rule



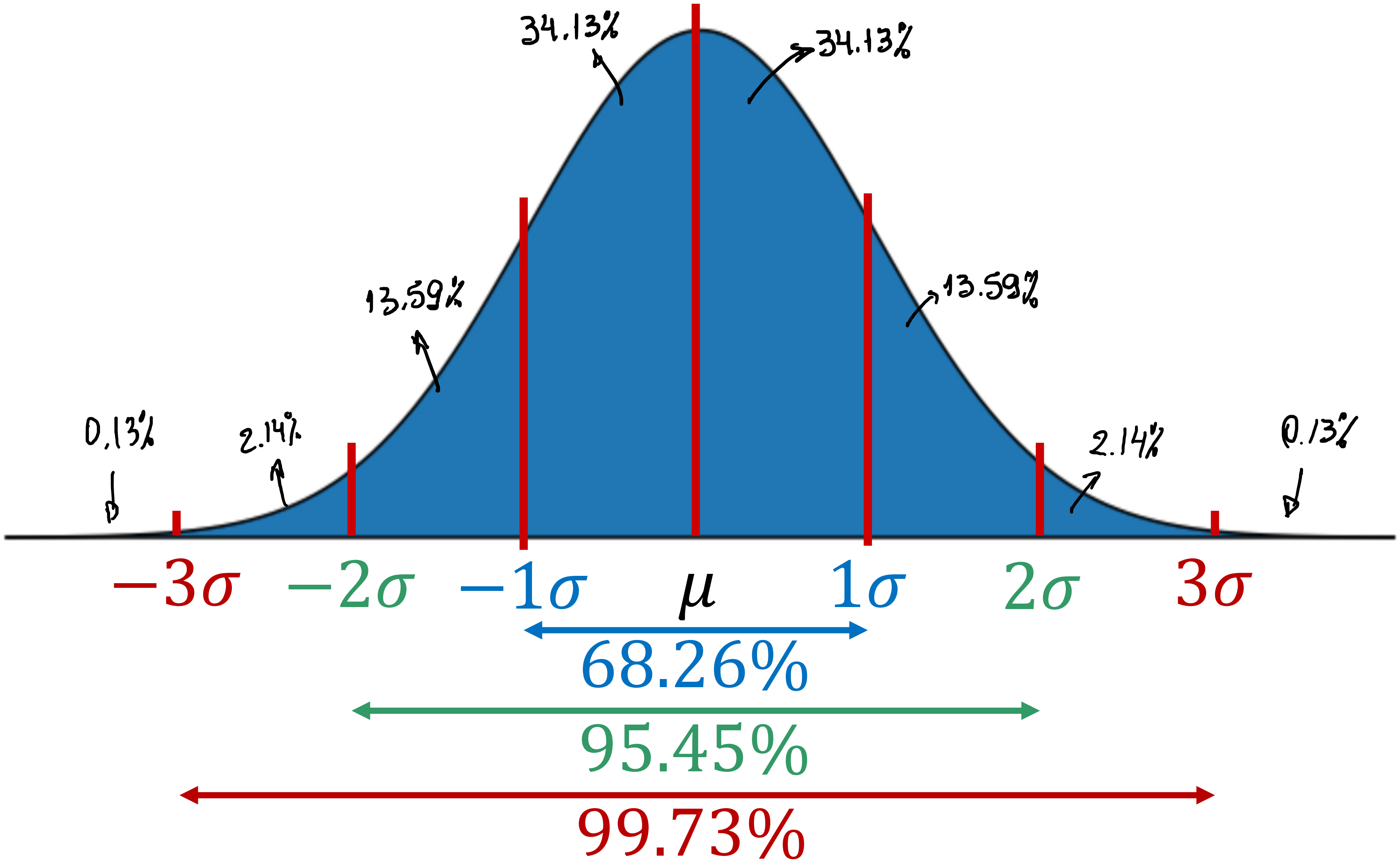
68-95-99.7% rule



68-95-99.7% rule



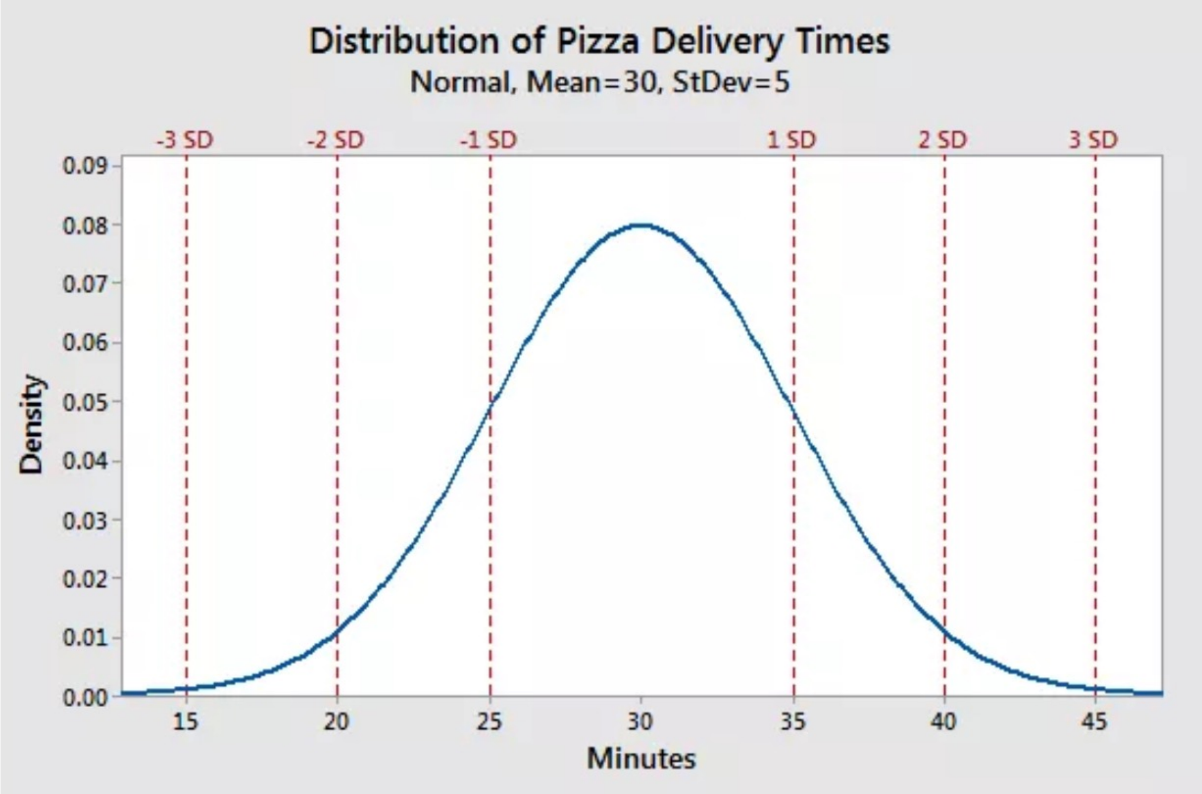
68-95-99.7% rule



Normal Distribution: Examples

The normal distribution occurs when a very large number of factors add together to create some random phenomenon.

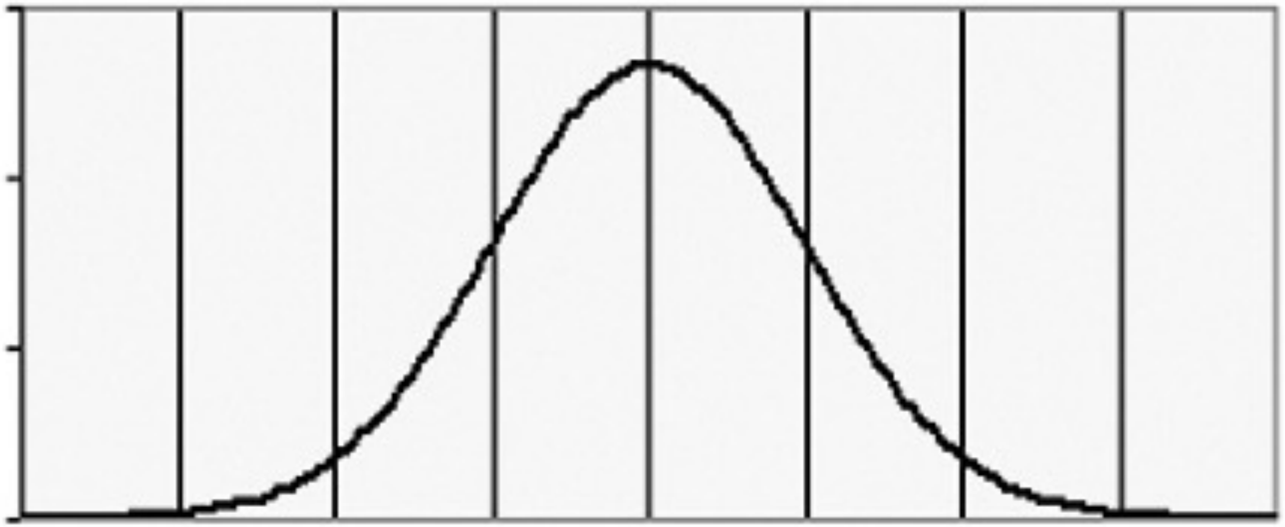
Example: How long until I get my Pizza?



Normal Distribution: Examples

The normal distribution occurs when a very large number of factors add together to create some random phenomenon.

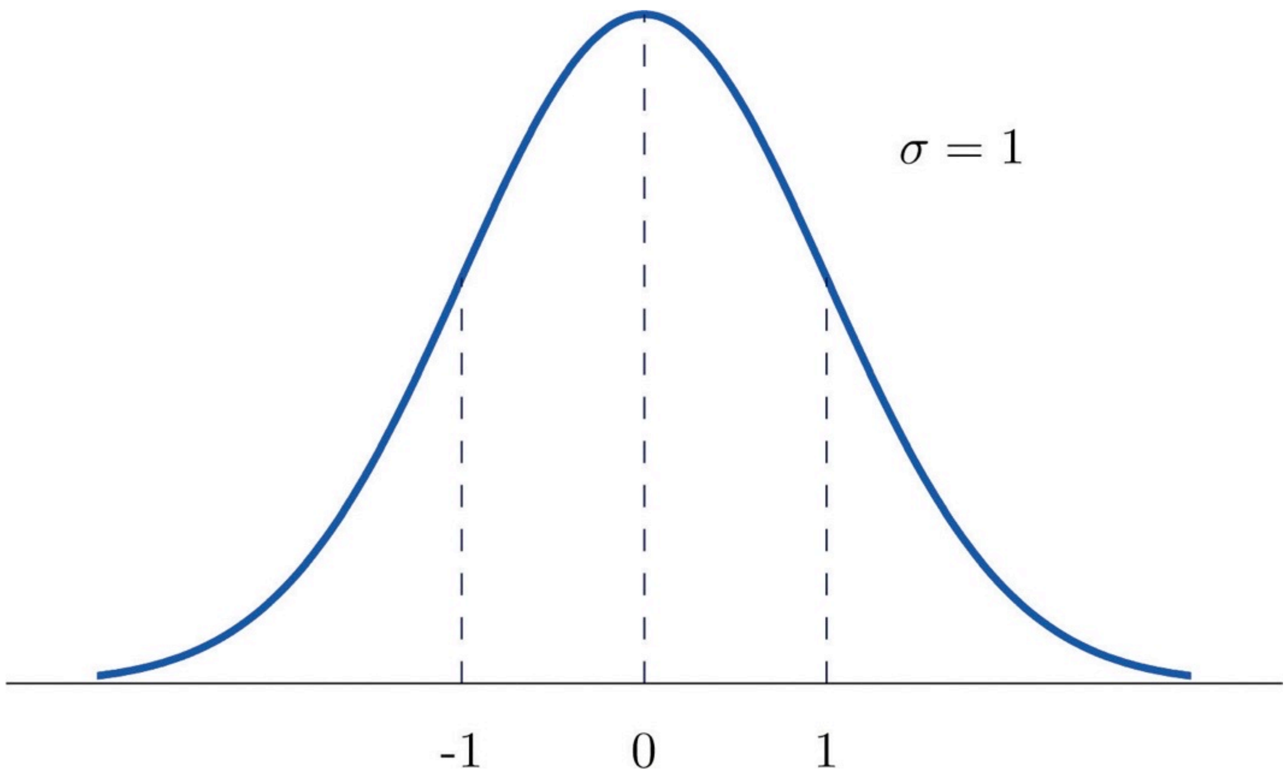
Example: What is the IQ of a human being?



Standard Deviations	-4	-3	-2	-1	0	1	2	3	4
Wechsler IQ	40	55	70	85	100	115	130	145	160
Stanford-Binet IQ	36	52	68	84	100	116	132	148	164
Cumulative %	0.003	0.135	2.275	15.866	50.000	84.134	97.725	99.865	99.997

Standard Normal Distribution

Since there are countless Normal Distributions, we focus on a normalized version, simply called the Standard Normal Distribution

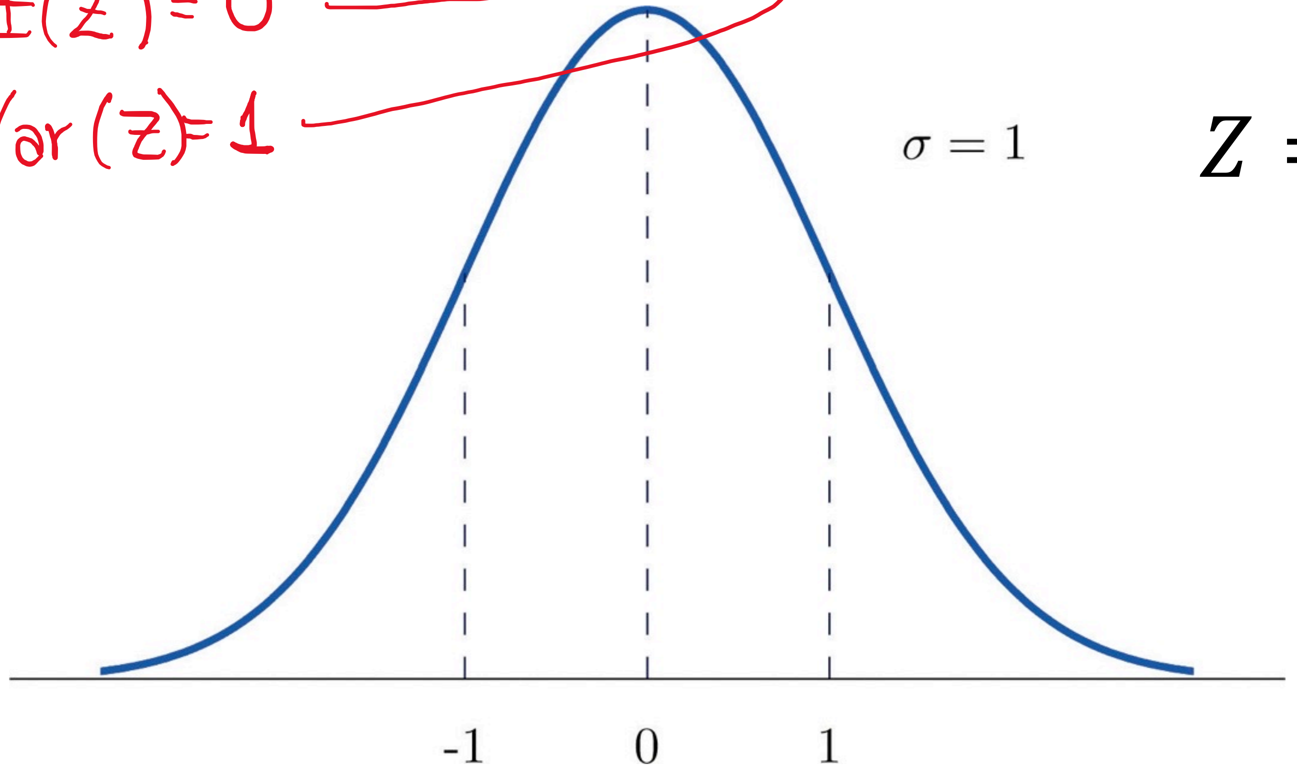


Standard Normal Distribution

We can convert any random variable which has a normal distribution $X \sim N(\mu, \sigma^2)$ into a standardized random variable $Z \sim N(0,1)$ by computing its z-score

$\mathbb{E}(Z) = 0$

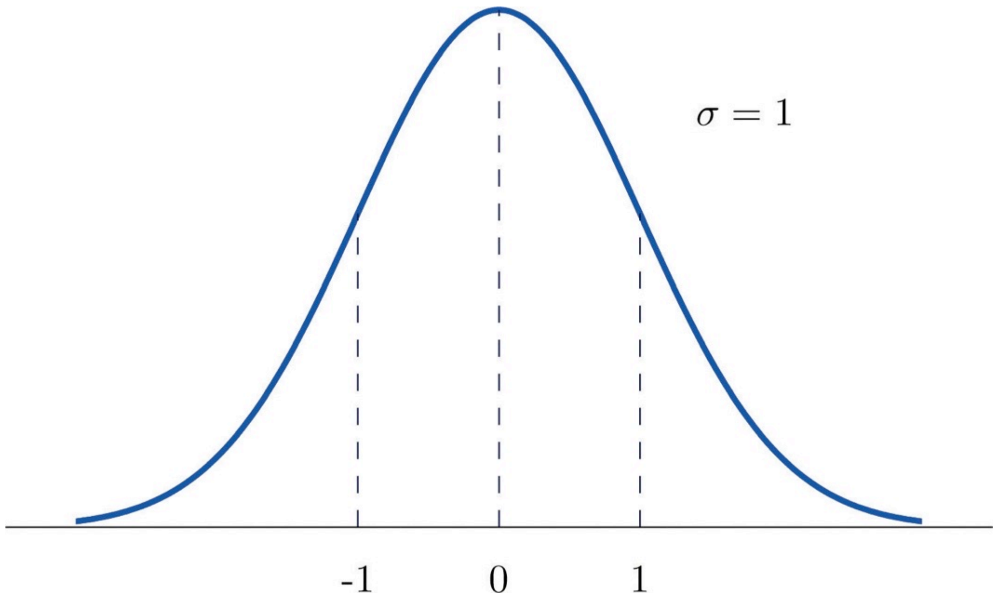
$Var(Z) = 1$



$$Z = \frac{X - \mu_x}{\sigma_X}$$

Standard Normal Distribution

We can convert any random variable which has a normal distribution $X \sim N(\mu, \sigma^2)$ into a standardized random variable $Z \sim N(0,1)$ by computing its z-score

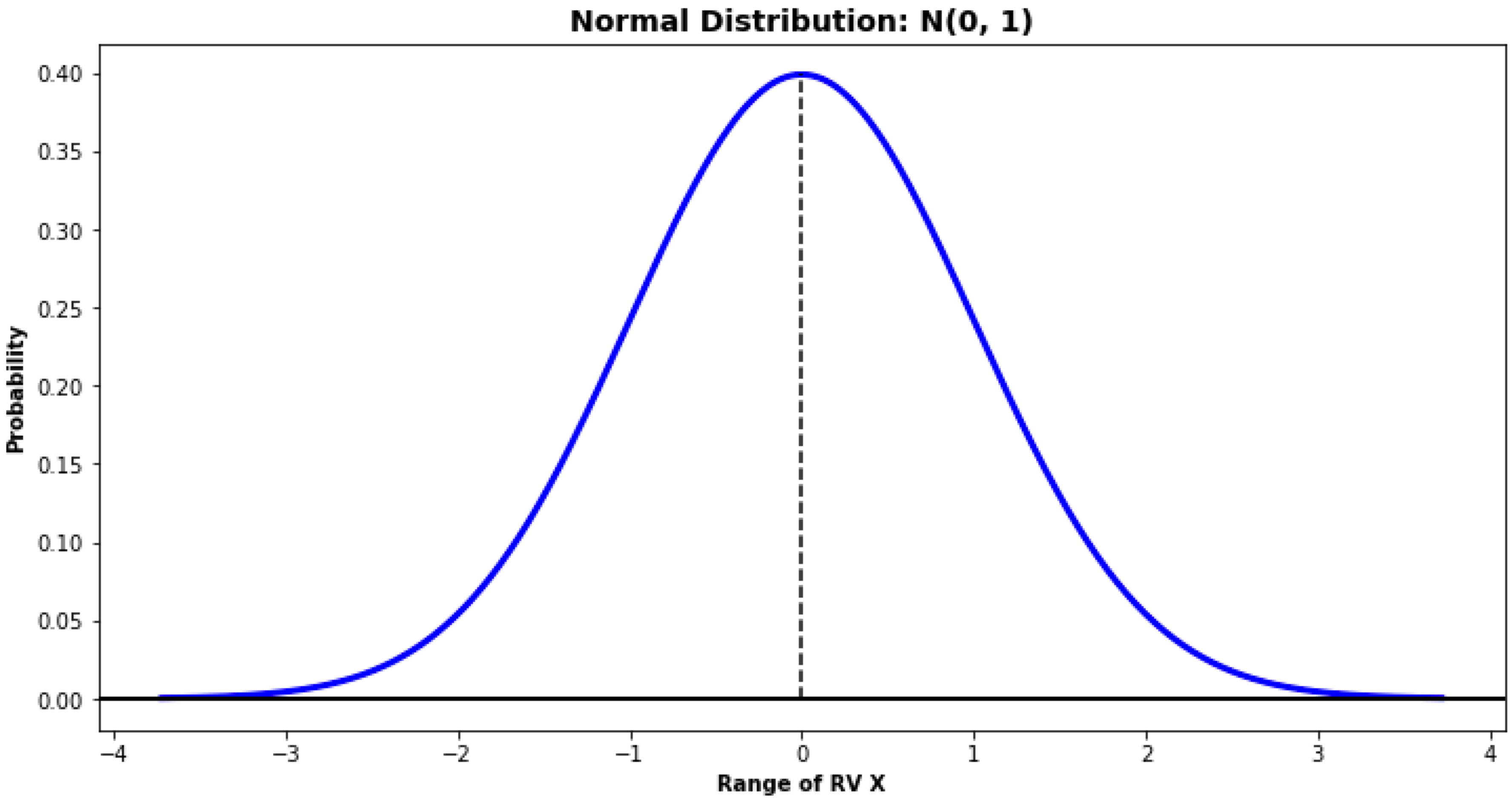


$$X = \sigma_X \cdot Z + \mu$$

Top Hat Activity (Join Code 033357)

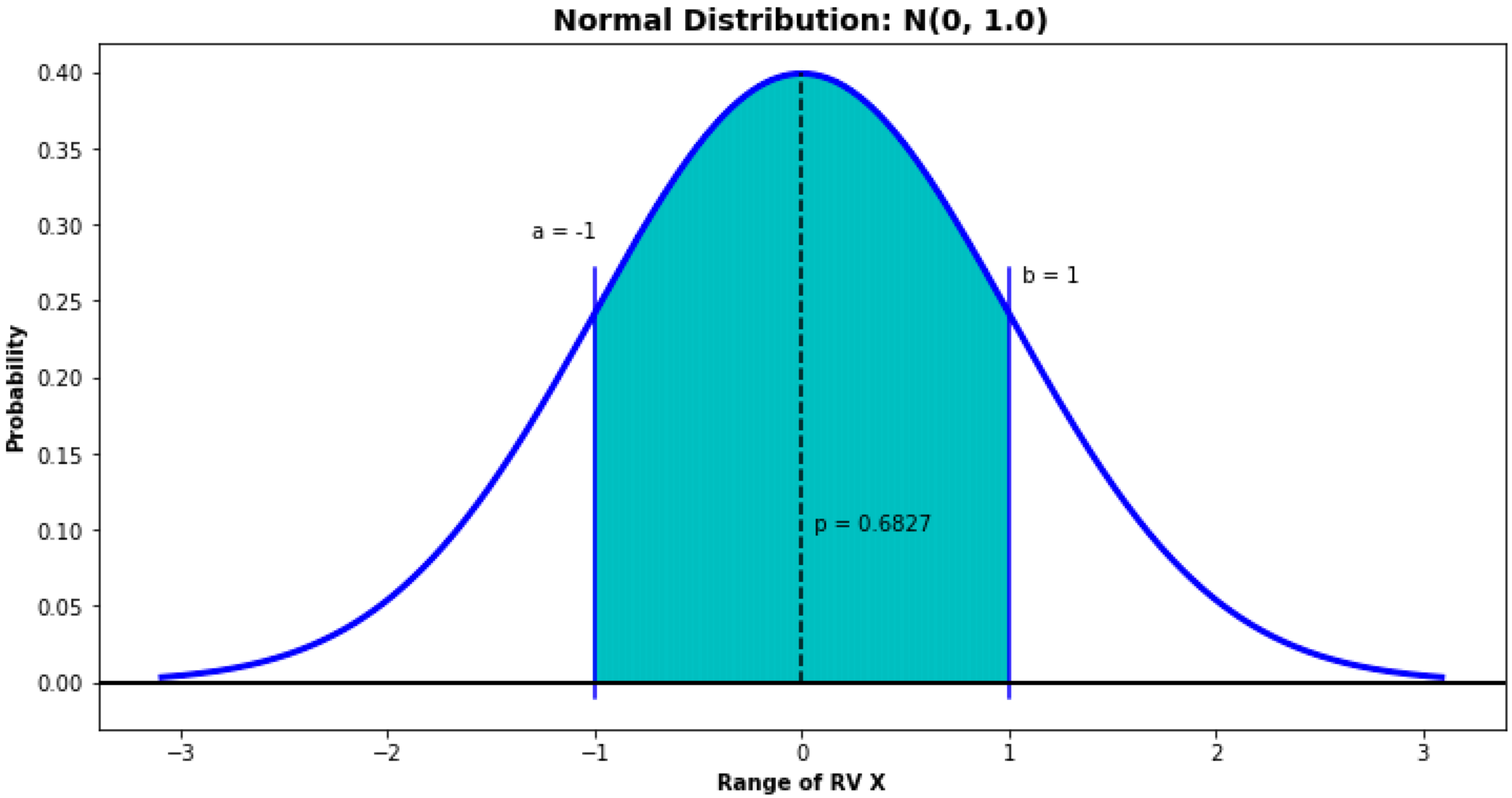
- The heights of adult men in the United States are approximately normally distributed with a mean of 70 inches and a standard deviation of 3 inches.
- Heights of adult women are approximately normally distributed with a mean of 64.5 inches and a standard deviation of 2.5 inches.
- Compute the z-score of your height?

Normal Distribution



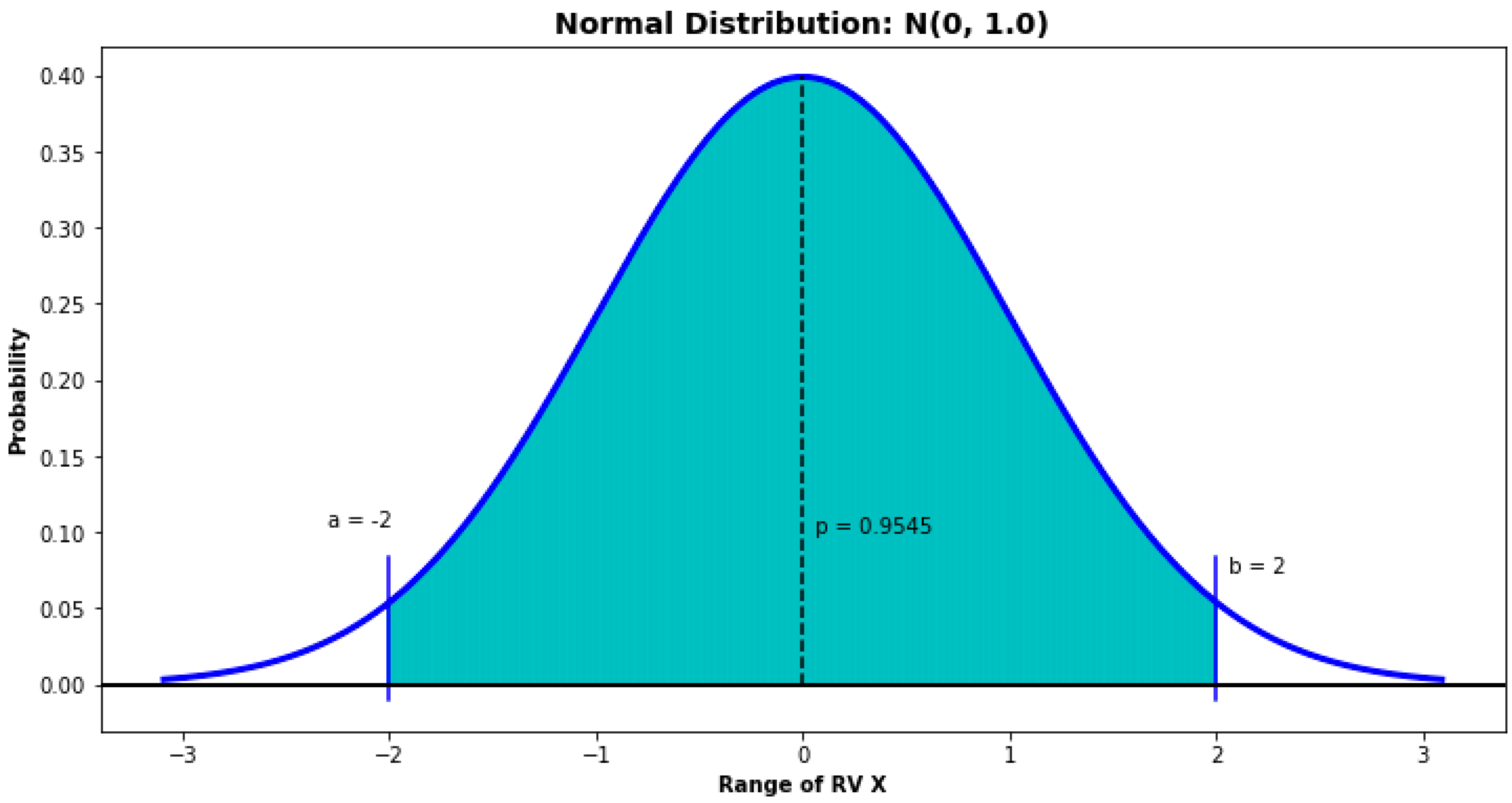

```
mean = 0
var = 1
stdev = 1
```

Normal Distribution



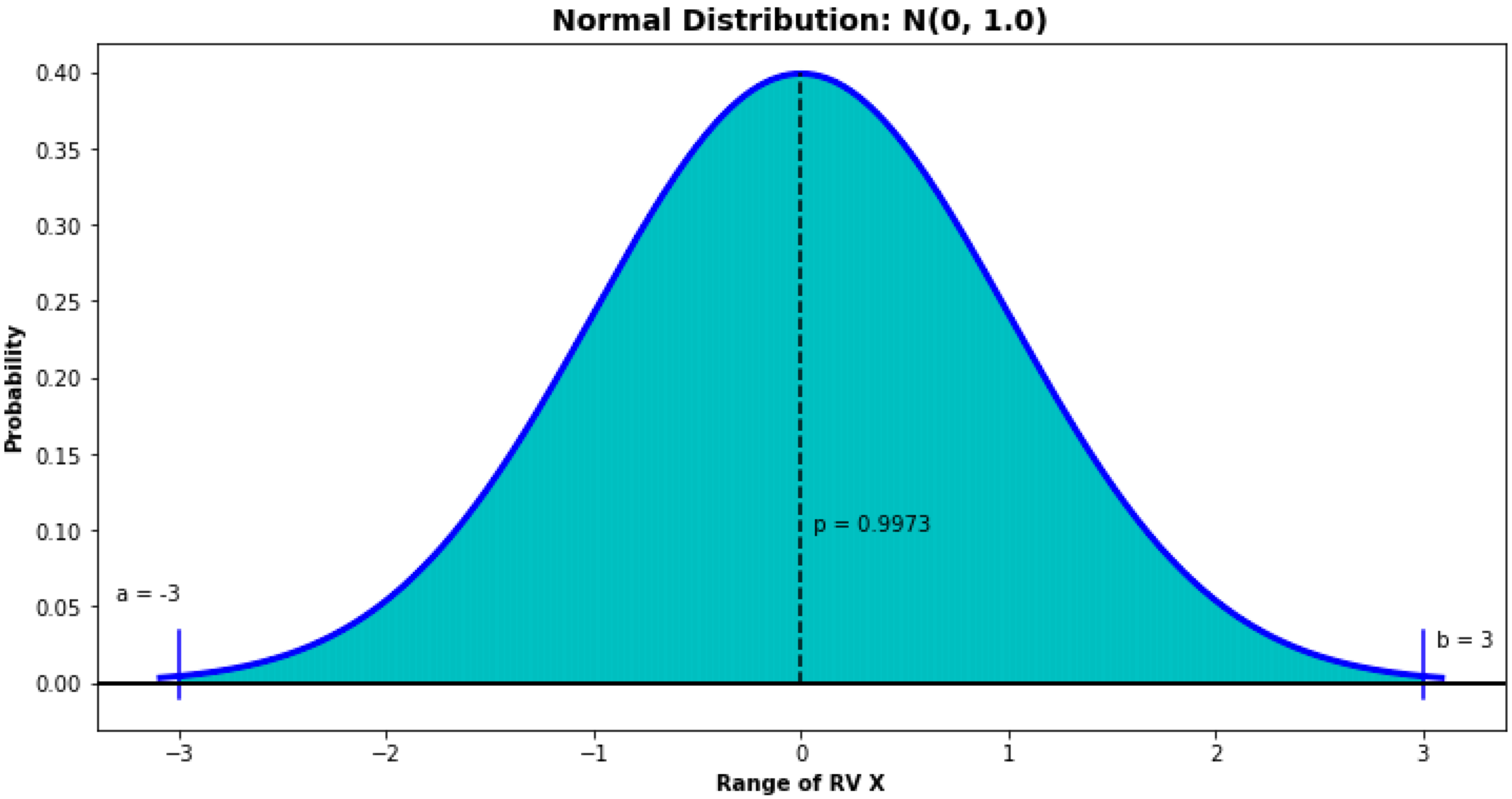

```
mean = 0
var = 1
stdev = 1.0
P(−1 < X < 1) = P(X < 1) − P(X < −1) = 0.8413 − 0.1587 = 0.6827
```

Normal Distribution




```
mean = 0
var = 1
stdev = 1.0
P(−2 < X < 2) = P(X < 2) − P(X < −2) = 0.9772 − 0.0228 = 0.9545
```

Normal Distribution




```
mean = 0
var = 1
stdev = 1.0
P(−3 < X < 3) = P(X < 3) − P(X < −3) = 0.9987 − 0.0013 = 0.9973
```

Example: Normal Distribution

Your local pizza shop claims that their large is at least 16in or it is free. Their pizza is normally distributed with $\mu = 16.3in$ and $\sigma = 0.2in$. What is the probability of getting free pizza?

CS 237 PDF: Normal Distribution

- The PDF of a *standard normal* random variable, $Z \sim N(0,1)$ is given by

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \text{for all } z \in \mathbb{R}.$$

- and its CDF is given by

$$F_Z(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{u^2}{2}} du$$

Unfortunately, this integral does not have a closed form solution ☹

PDF: Normal Distribution

- The PDF of a *standard normal* random variable, $Z \sim N(0,1)$ is given by

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \text{for all } z \in \mathbb{R}.$$

- and its CDF is given by

$$\Phi(x) = \Pr(Z \leq x) = F_Z(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du$$

Unfortunately, this integral does not have a closed form solution ☹

CS 237 Properties of $\Phi(x)$

Here are some cool properties of ϕ

1. $\lim_{x\to\infty} \Phi(x) = 1$ and $\lim_{x\to-\infty} \Phi(x) = 0$
2. $\Phi(0) = \frac{1}{2}$
3. $\Phi(-x) = 1 - \Phi(x)$, for all $x \in \mathbb{R}$

If Z is a standard normal variable and $X = \sigma Z + \mu$ is a normal random variable, then

$$X \sim N(\mu, \sigma^2)$$

Getting Free Pizza

Your local pizza shop claims that their large is at least 16in or it is free. Their pizza is normally distributed with $\mu = 16.3in$ and $\sigma = 0.2in$. What is the probability of getting free pizza?

Getting Huge Pizza

Your local pizza shop claims that their large is at least 16in or it is free. Their pizza is normally distributed with $\mu = 16.3in$ and $\sigma = 0.2in$. What is the probability of getting a pizza over $16.5in$?

Getting Huge Pizza

Your local pizza shop claims that their large is at least 16in or it is free. Their pizza is normally distributed with $\mu = 16.3in$ and $\sigma = 0.2in$. What is the probability of getting a pizza between $15.95in$ and $16.63in$?

Normal Distribution

Recall that the only way we can analyze probabilities in the continuous case is with the CDF:

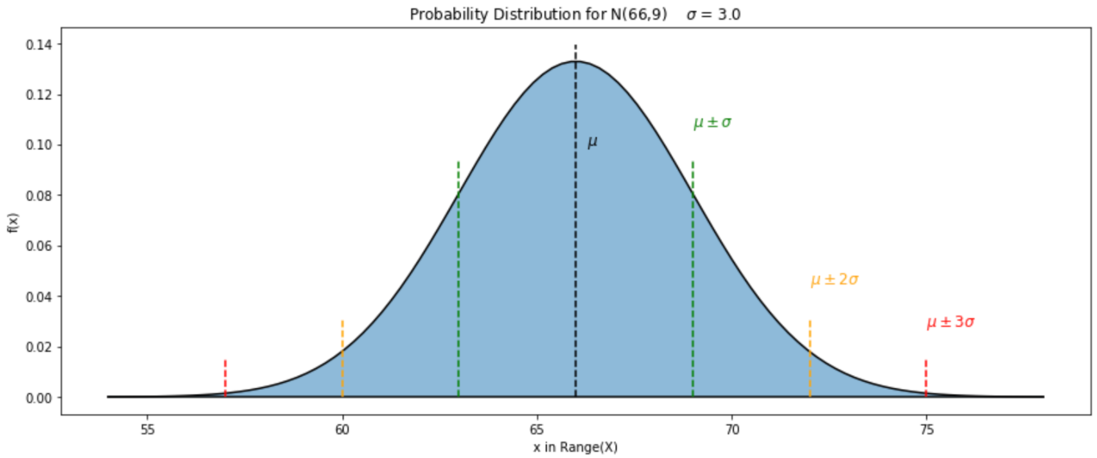
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

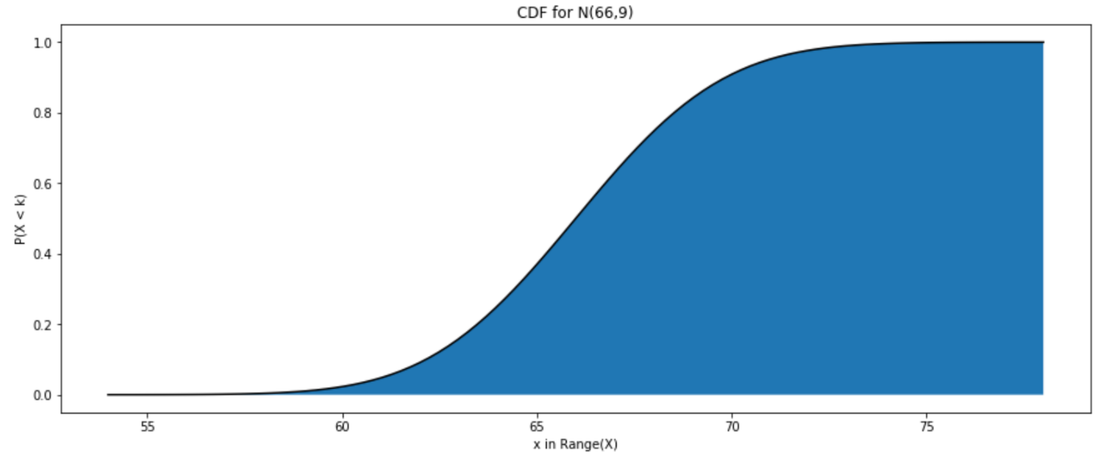
$$F(a) = \int_{-\infty}^{a} f(x)\, dx$$

$$\Pr(X < a) = F(a)$$

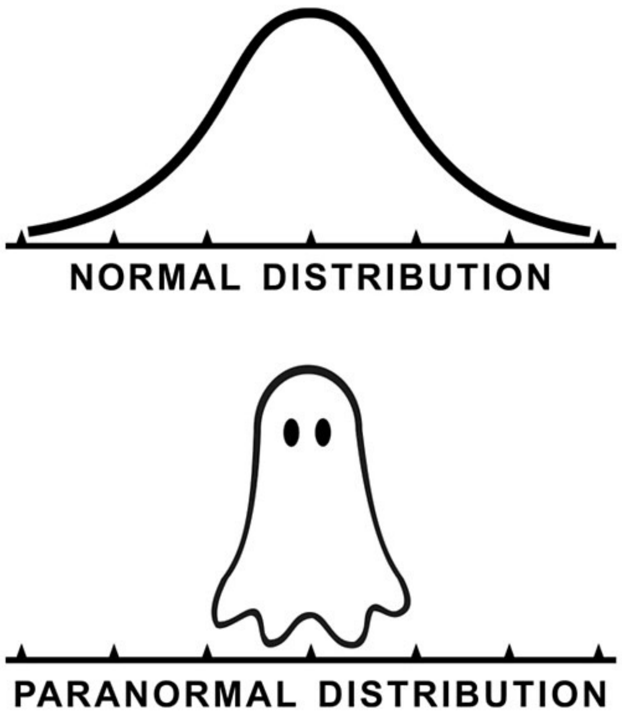
$$\Pr(X > a) = 1.0 - F(a)$$

$$\Pr(a < X < b) = F(b) - F(a)$$





Normal Distribution



https://www.redbubble.com/i/ipad-skin/Normal-Paranormal-Distribution-by-Dreambase/19202537.MHP6F