CS 237: Probability in Computing

Wayne Snyder
Computer Science Department
Boston University

Lecture 15:

- Continuous Distributions
  - Basic Definitions
  - Importance of the CDF
  - Calculation of probabilities using integration
- Uniform Continuous Distribution
- Introduction to Normal Distribution

Discrete vs Continuous Distributions

Recall: A Random Variable X is a function from a sample space S into the reals:

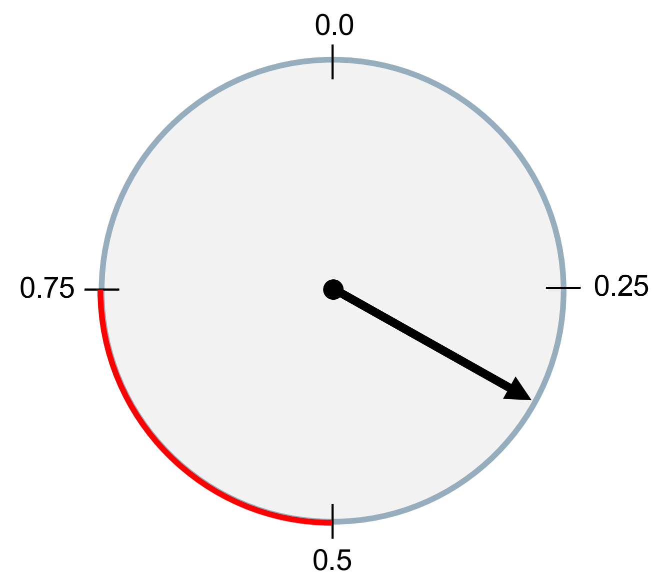
$$X : S \rightarrow \mathcal{R}$$

A random variable is called continuous if Rx is uncountable.

What needs to change when working with continuous as opposed to discrete distributions?

Recall: The probability of a random experiment such as a spinner outputting any particular, exact real number is 0:

$$f_X(a) = P(X = a) = 0$$



This result extends to any countable collection of real numbers!

So we can **only** think about (countable unions of) intervals:

$$P(0.5 < X < 0.75) = 0.25$$

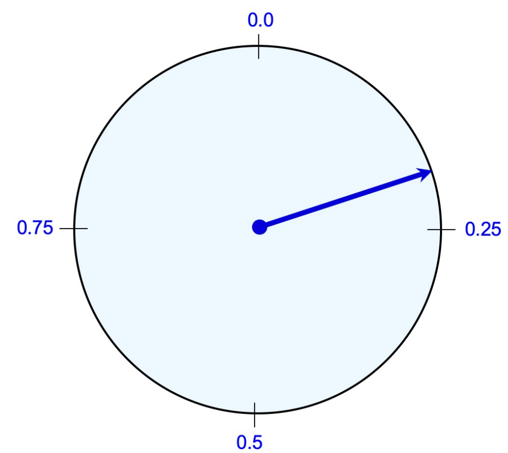
Probability Functions: Equiprobable vs Not Equiprobable

When the sample space is uncountable, say with the spinner, it is possible for the probability function to be equiprobable or non-equiprobable.

Uncountable and Equiprobable:

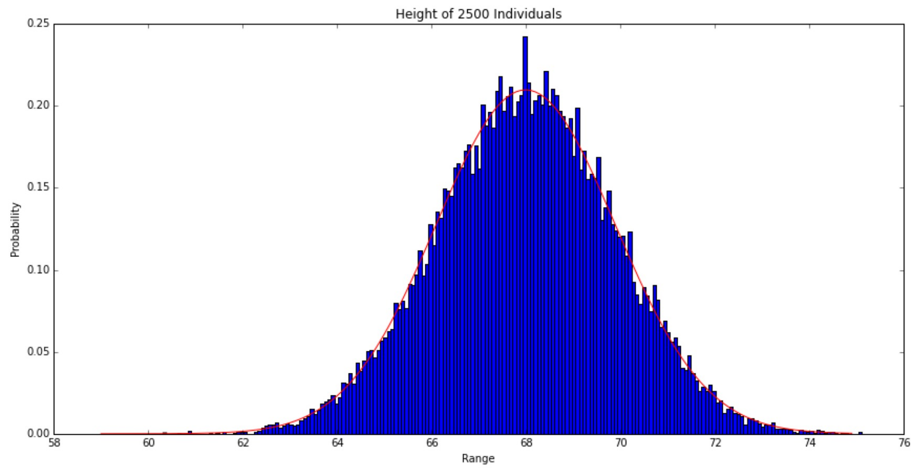
Example: Spin the spinner and report the real number showing.

S = [0..1) **Any point is equally likely**



Uncountable and NOT Equiprobable:

Example: Heights of Human Beings:



People are more likely to be close to the average height than at the extremes!

Review: Cumulative Distribution Functions

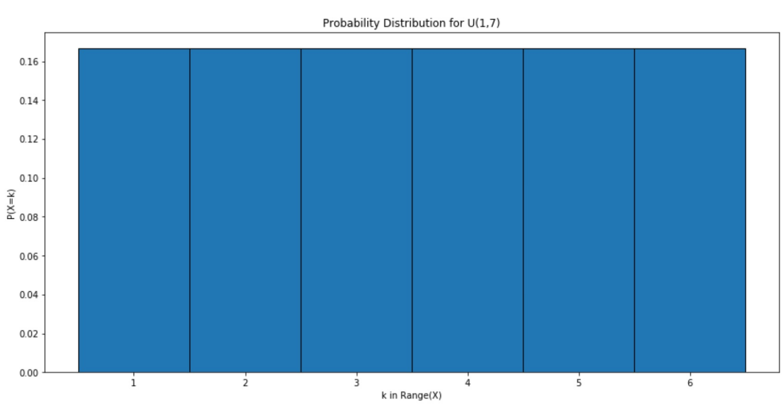
The Cumulative Distribution Function (CDF) for a random variable X shows what happens when we keep track of the sum of the probability distribution from left to right over its range:

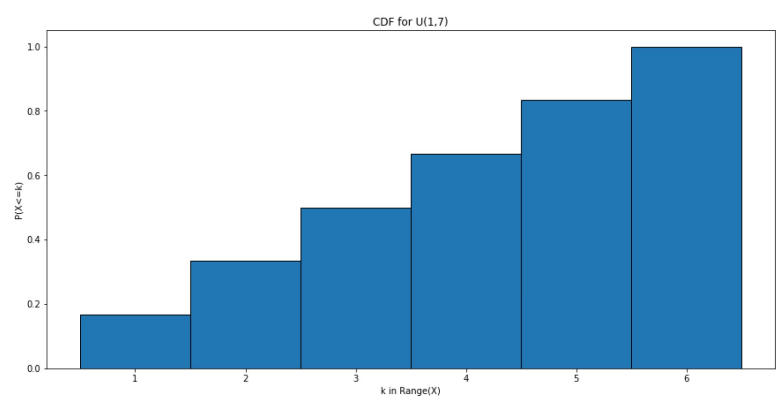
$$F_X(k) = P(X \leq k) = \sum_{a \leq k} \mathrm{P}_X(a)$$

Example: X = “The number of dots showing on a thrown die”

Probability Distribution Function P_X

Cumulative Distribution Function F_X





Discrete vs Continuous Distributions: PDF vs PMF

Because of the anomolies having to do with continuous probability, we need to keep the following important points in mind:

(A) We will no longer be able to use a discrete Probability Mass Function, but instead a Probability Density Function (PDF), **$f_X(a)$.**

(A) The probability function f_X does NOT represent the probability of a point in the domain, since as we know:

$$f_X(a) = P(X = a) = 0$$

therefore we can ONLY work with intervals $P(X \leq a)$, $P(X > a)$, $P(a \leq X \leq b)$, etc. and f_X is not as important as the CDF F_X .

(B) In calculating F_X and working with intervals, we can not use discrete sums $\sum_{x=a}^{b}$ as we did in the discrete case, but will have to use integrals: $\int_a^b$

(C) The range R_X will be all the reals $(-\infty..\infty)$ and so we don't specify it each time.

Discrete vs Continuous Distributions

Discrete Random Variables

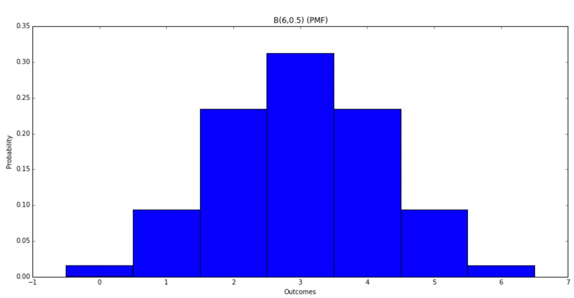
The Probability Mass Function (PMF) of a discrete random variable X is a function from the range of X into $\mathcal{R}$:

$$P_x : R_X \mapsto \mathcal{R}$$

such that

(i) $\forall y \in R_X \;\; P_X(y) \geq 0.0$

(ii) $\sum_{y \in R_X} P_X(y) = 1.0$



Continuous Random Variables

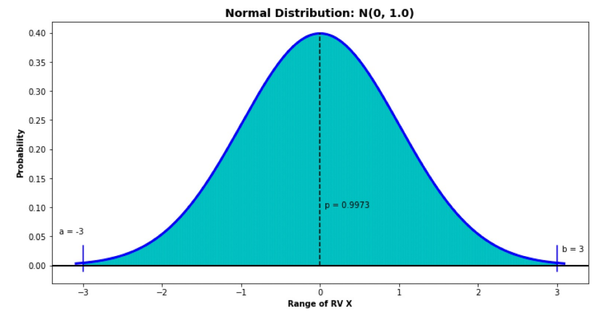
The Probability Density Function (PDF) of a continuous random variable X is a function from $\mathcal{R}$ to $\mathcal{R}$:

$$f_x : \mathcal{R} \mapsto \mathcal{R}$$

such that

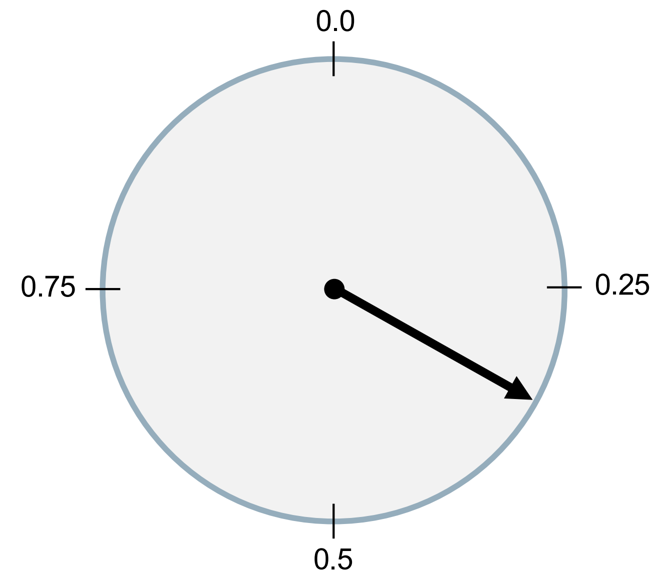
(i) $\forall y \;\; f_X(y) \geq 0.0$

(ii) $\int_{-\infty}^{\infty} f_X(y)\, dy = 1.0$



Continuous Distributions

Let's clarify these ideas with an example....



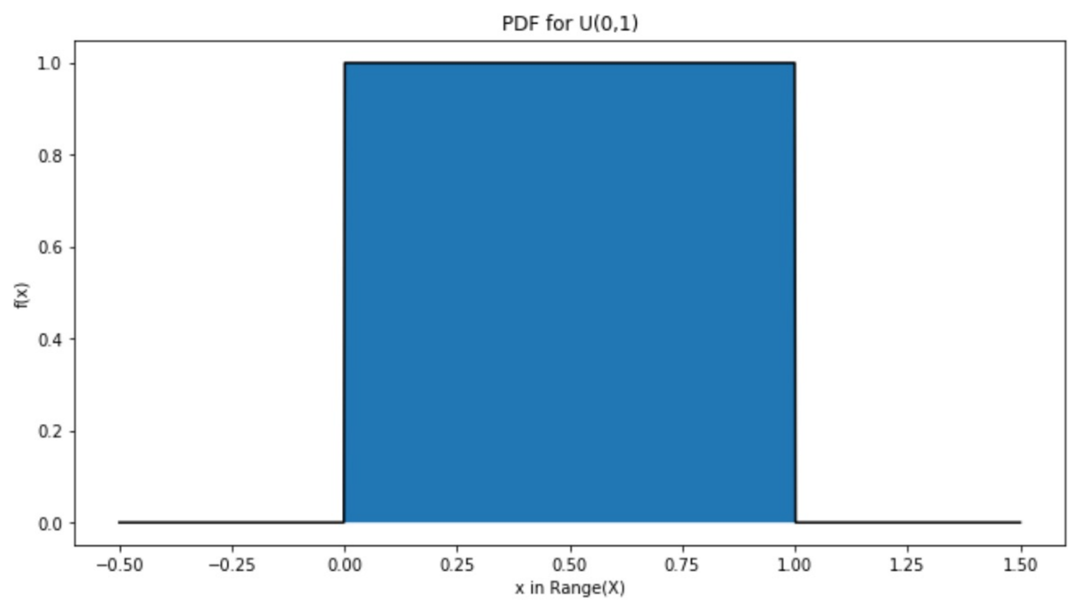
Consider the spinner example from way back when:

X = "the real number in [0..1) that the spinner lands on"

The probability density function is:

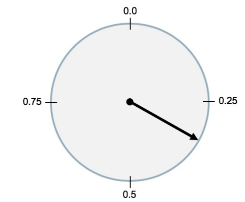
$$f(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Note that the area is 1.0 and for any $0 \leq a \leq 1$, we have $f(a) = 1.0$, so it is uniform across [0..1). But clearly $P(X = a) = 0.0$.



Continuous Distributions

Now recall that the ONLY way to deal with continuous probability is to use intervals and to use area (or extent) for the probability. Hence we will calculate probabilities of intervals using the CDF:

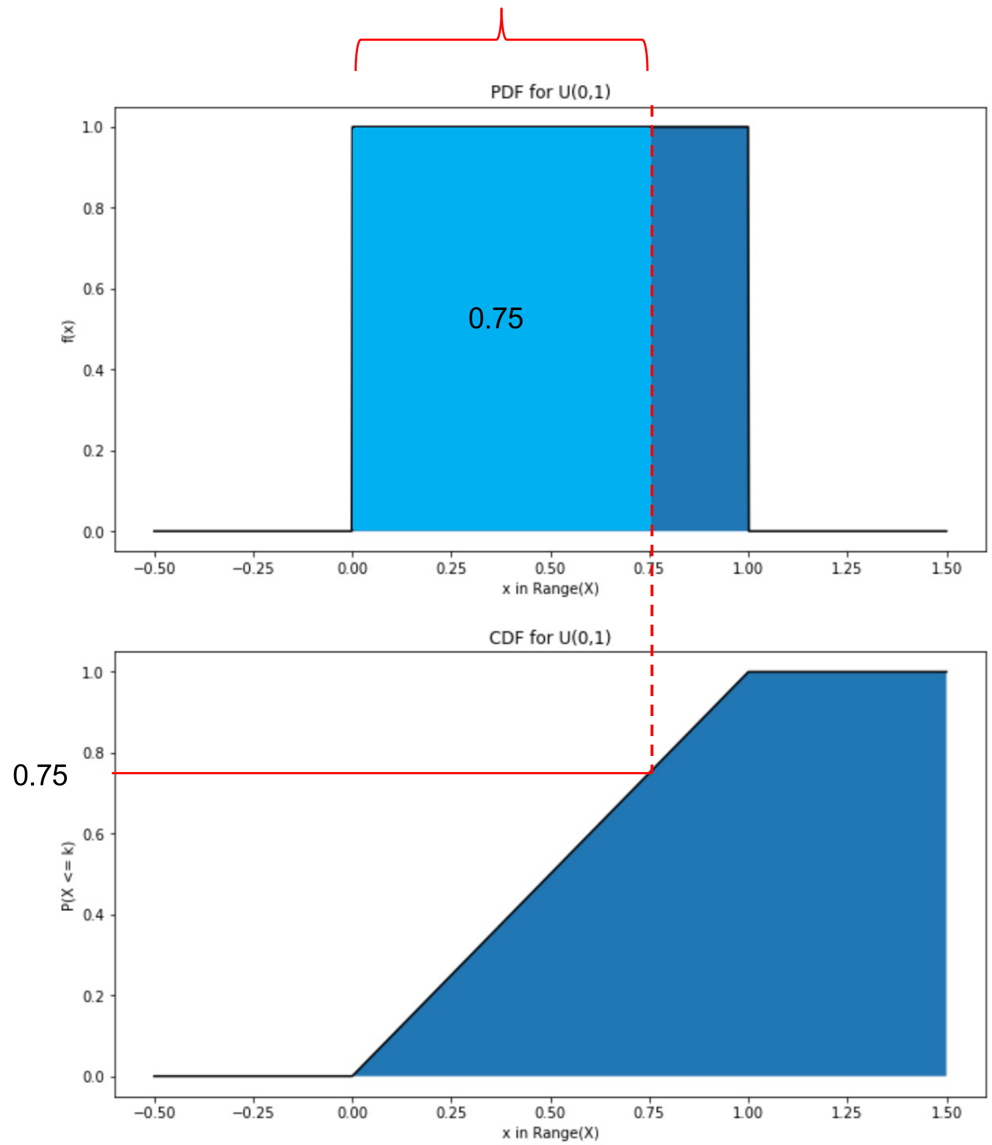


$$P(X < 0.75) = F(0.75) = 0.75$$

$$f(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

$$F(a) = \int_0^a 1 \; dx = x \Big|_0^a = a$$

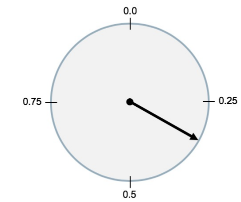
$$F(a) = \begin{cases} 0 & \text{if } a < 0 \\ a & \text{if } 0 \leq a \leq 1 \\ 1 & \text{if } a > 1 \end{cases}$$



$$F(a) = \int_{-\infty}^{a} 1 \, dx = a$$

A brief review of integration is on the YT channel!

Continuous Distributions

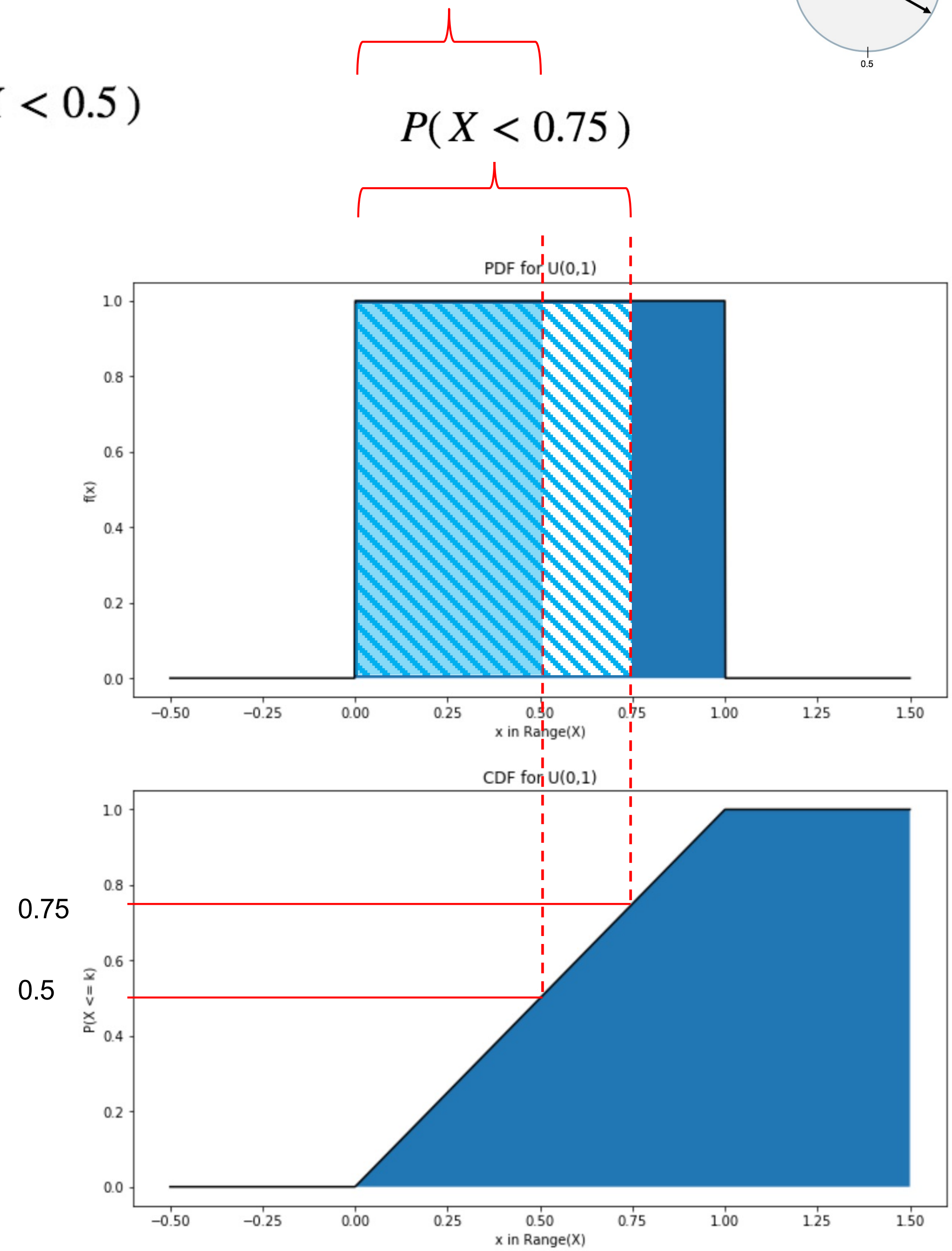


$$P(0.5 < X < 0.75) = P(X < 0.75) - P(X < 0.5)$$
$$= F(0.75) - F(0.5)$$
$$= 0.75 - 0.5$$
$$= 0.25$$

$$f(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

$$F(a) = \int_0^a 1\, dx = x \Big|_0^a = a$$

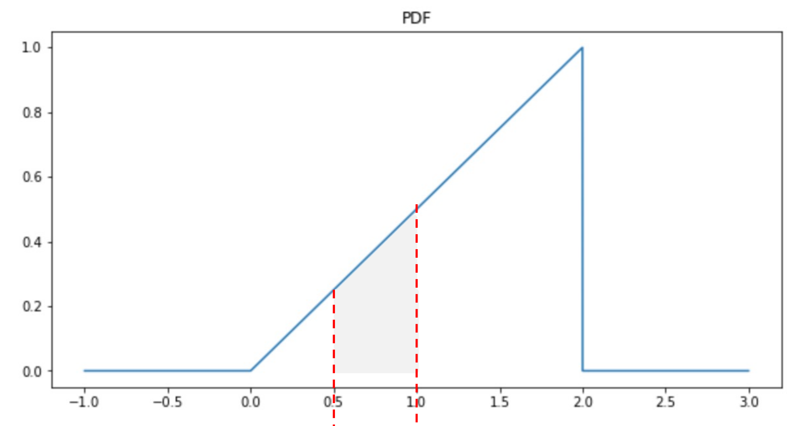
$$F(a) = \begin{cases} 0 & \text{if } a < 0 \\ a & \text{if } 0 \leq a \leq 1 \\ 1 & \text{if } a > 1 \end{cases}$$



Continuous Distributions

Bottom Line: In order to deal with continuous distributions, you have to either calculate areas using geometric techniques, or do integrals.

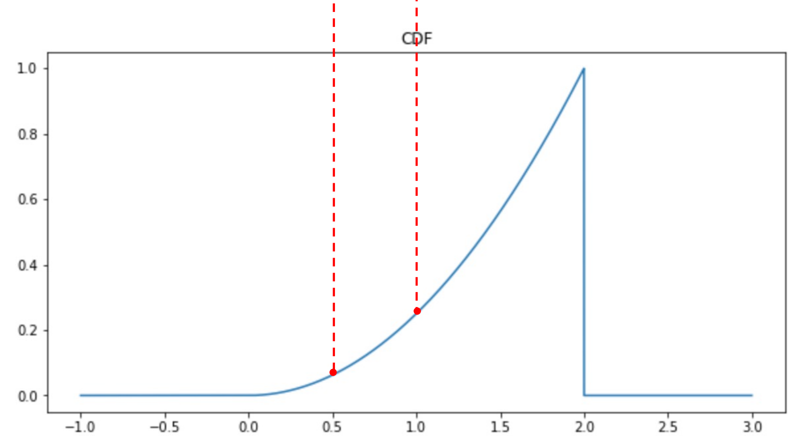
Example: Suppose our PDF looked like this:



$$f(x) = \begin{cases} \dfrac{x}{2} & \text{if } 0 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

To calculate the probability of intervals, we need to determine the CDF, which means doing the following integral:

$$F(a) = \int_{-\infty}^{a} \frac{x}{2}\, dx = \frac{a^2}{4}$$



So for example,

$$P(0.5 \leq X \leq 1) = F(1) - F(0.5) = \frac{1^2}{4} - \frac{0.5^2}{4} = \frac{1}{4} - \frac{1}{16} = \frac{3}{16} = 0.1875$$

$$F(a) = \int_{-\infty}^{a} \frac{x}{2}\, dx = \frac{a^2}{4}$$

Continuous Distributions

Discrete Random Variables

$$F_X(b) = P(X \leq b) =_{def} \sum_{y \leq b} P_X(y)$$

$$P(a \leq X \leq b) =_{def} \sum_{a \leq y \leq b} P_X(y)$$

$$E(X) =_{def} \sum_{y \in R_X} y \cdot P_X(y)$$

Continuous Random Variables

$$F_X(b) = P(X < b) =_{\text{def}} \int_{-\infty}^{b} f(x)\,dx$$

$$P(a < X < b) =_{\text{def}} \int_{a}^{b} f(x)\,dx$$

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

Same for both Discrete and Continuous Random Variables

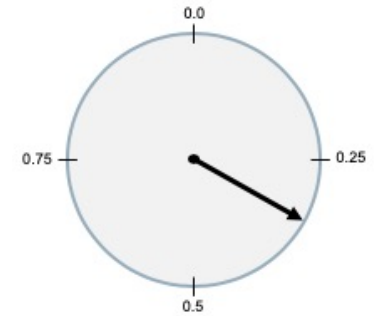
$$Var(X) =_{def} E[(X - \mu_X)^2]$$

$$\sigma_X =_{def} \sqrt{Var(X)}$$

$$Var(X) = E(X^2) - (\mu_X)^2$$

All previous theorems about E(X) and Var(X) still hold, it does not matter whether X is continuous or discrete!

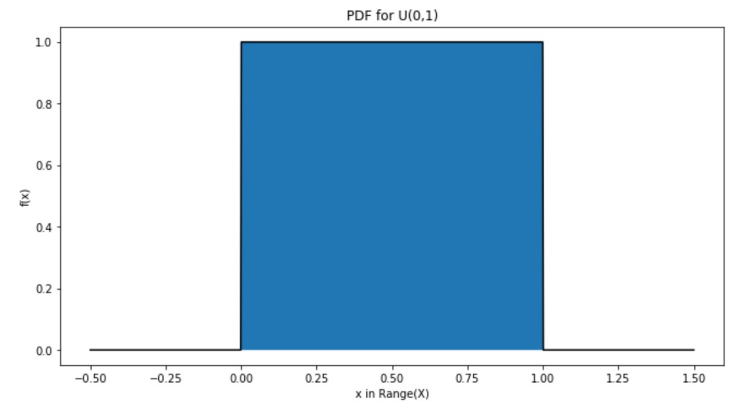
Example: Calculate the **expected value** of the uniform distribution over the interval [0..1):



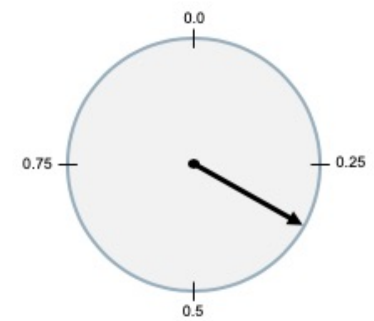
$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\, dx$$

$$R_X = [0..1)$$

$$f(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$



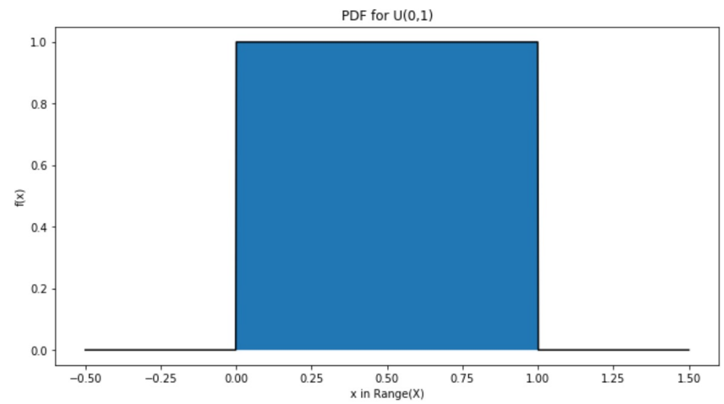
Example: Calculate the **variance** of the uniform distribution over the interval [0..1):



$$Var(X) = E(X^2) - (\mu_X)^2$$

$$R_X = [0..1)$$

$$f(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

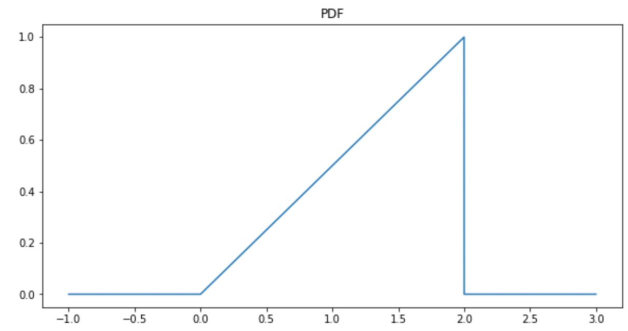


Example: Calculate the **expected value** of the following distribution over the interval [0..2):

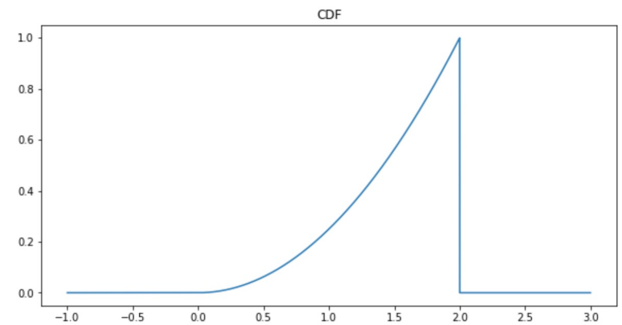
$$R_X = [0..2)$$

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

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



$$F(a) = \int_{-\infty}^{a} \frac{x}{2}\, dx = \frac{x^2}{4}$$

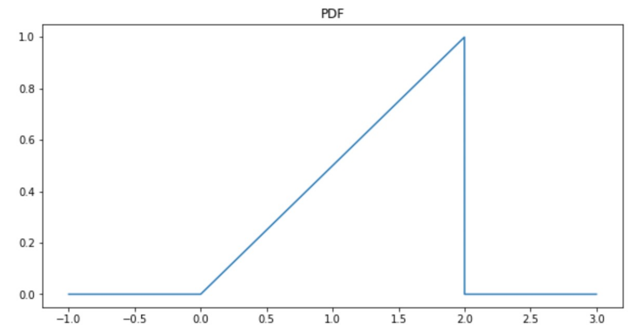


Example: Calculate the **variance** of the following distribution over the interval [0..2):

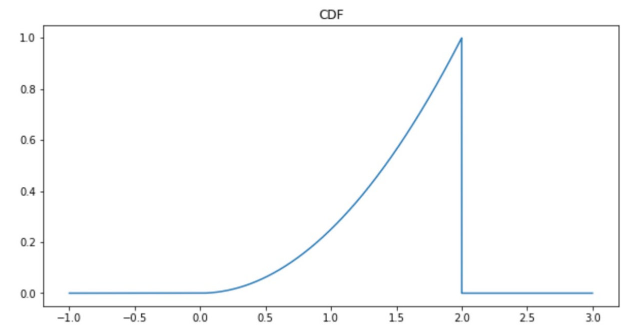
$$R_X = [0..2)$$

$$Var(X) = E(X^2) - (\mu_X)^2$$

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



$$F(a) = \int_{-\infty}^{a} \frac{x}{2}\, dx = \frac{x^2}{4}$$

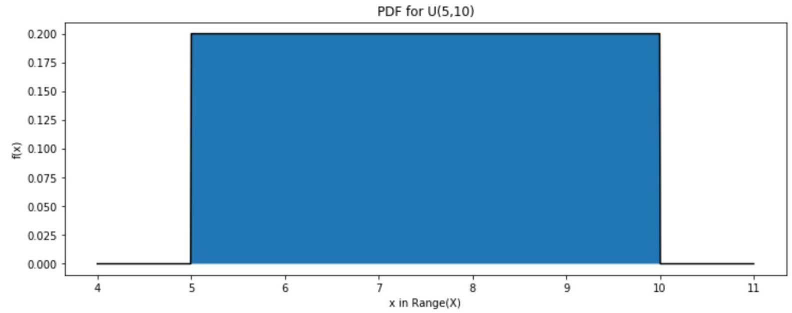


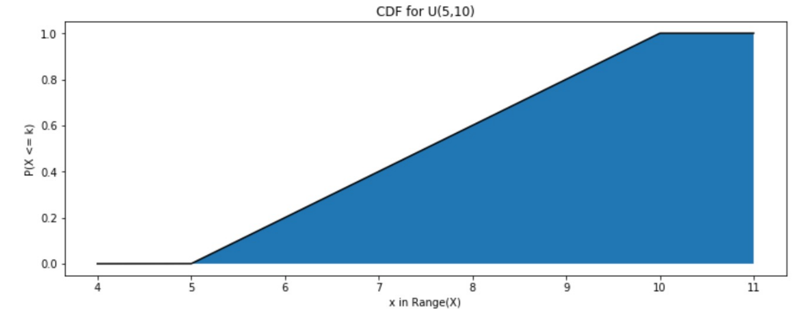
Uniform Distribution

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

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

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





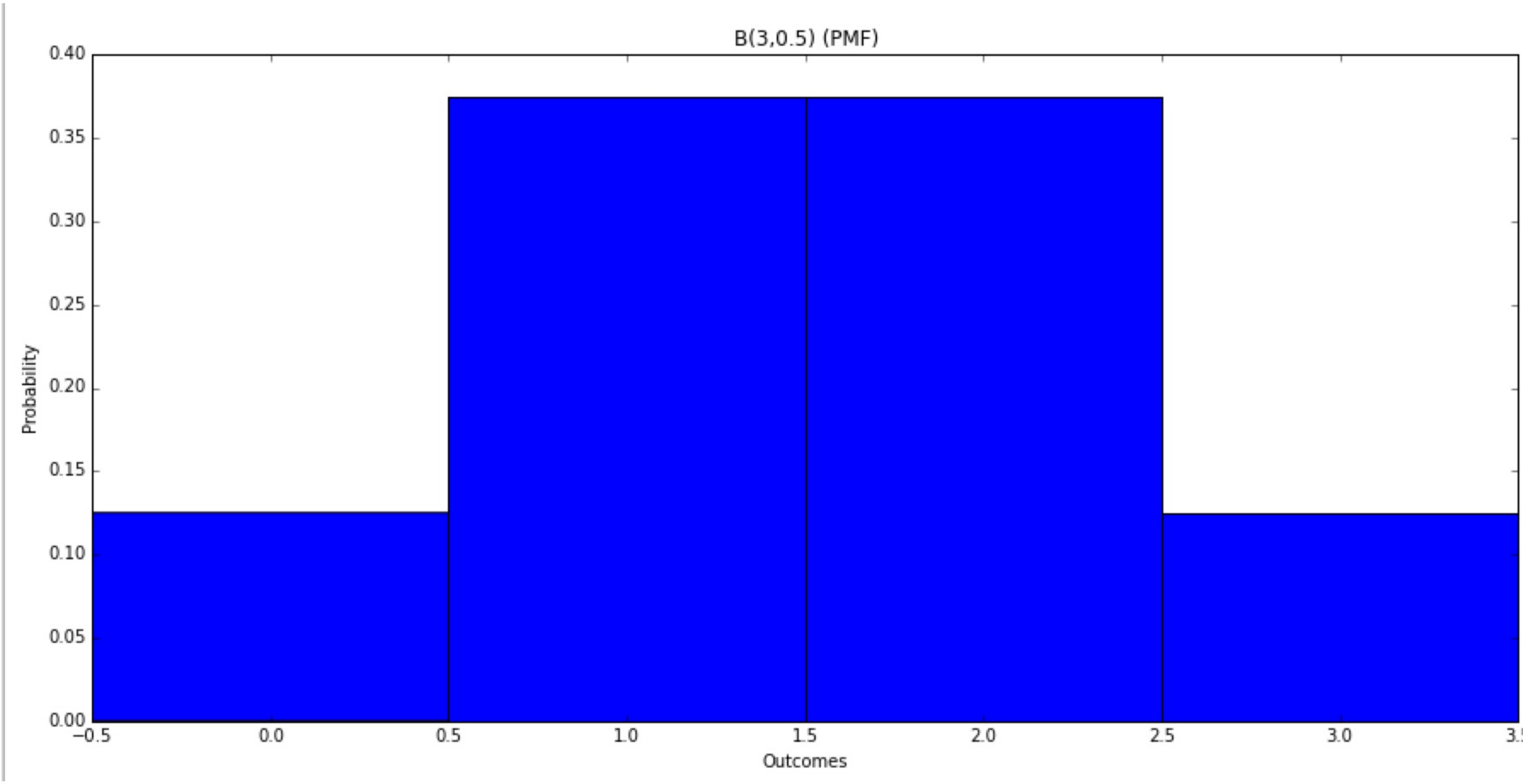
$$E(X) = \int_a^b x \cdot \frac{1}{b-a}\, dx = \frac{1}{b-a} \cdot \frac{x^2}{2}\Big|_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{b+a}{2}$$

$$E(X^2) = \int_a^b x^2 \cdot \frac{1}{b-a}\, dx = \frac{1}{b-a} \cdot \frac{x^3}{3}\Big|_a^b = \frac{b^3 - a^3}{3(b-a)} = \frac{a^2 + ab + b^2}{3}$$

$$Var(X) = E(X^2) - E(X)^2 = \frac{a^2 + ab + b^2}{3} - \frac{a^2 - 2ab + b^2}{4} = \frac{a^2 + 2ab + b^2}{12} = \frac{(b-a)^2}{12}$$

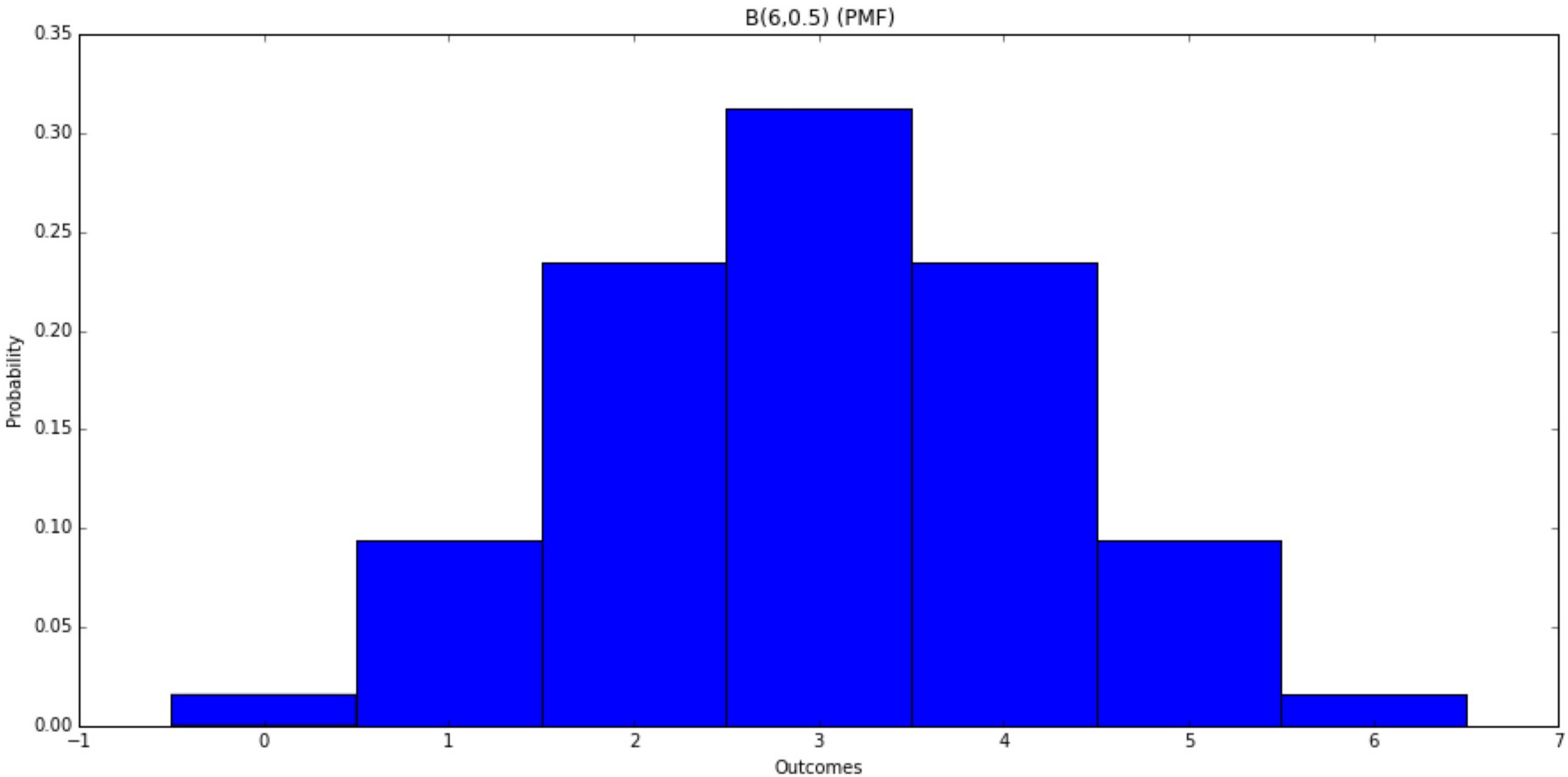
Normal Distribution as Limit of Binomial

When we observe the characteristic shape of the Binomial Distribution B(N,0.5) as N approaches Infinity, we see something interesting:



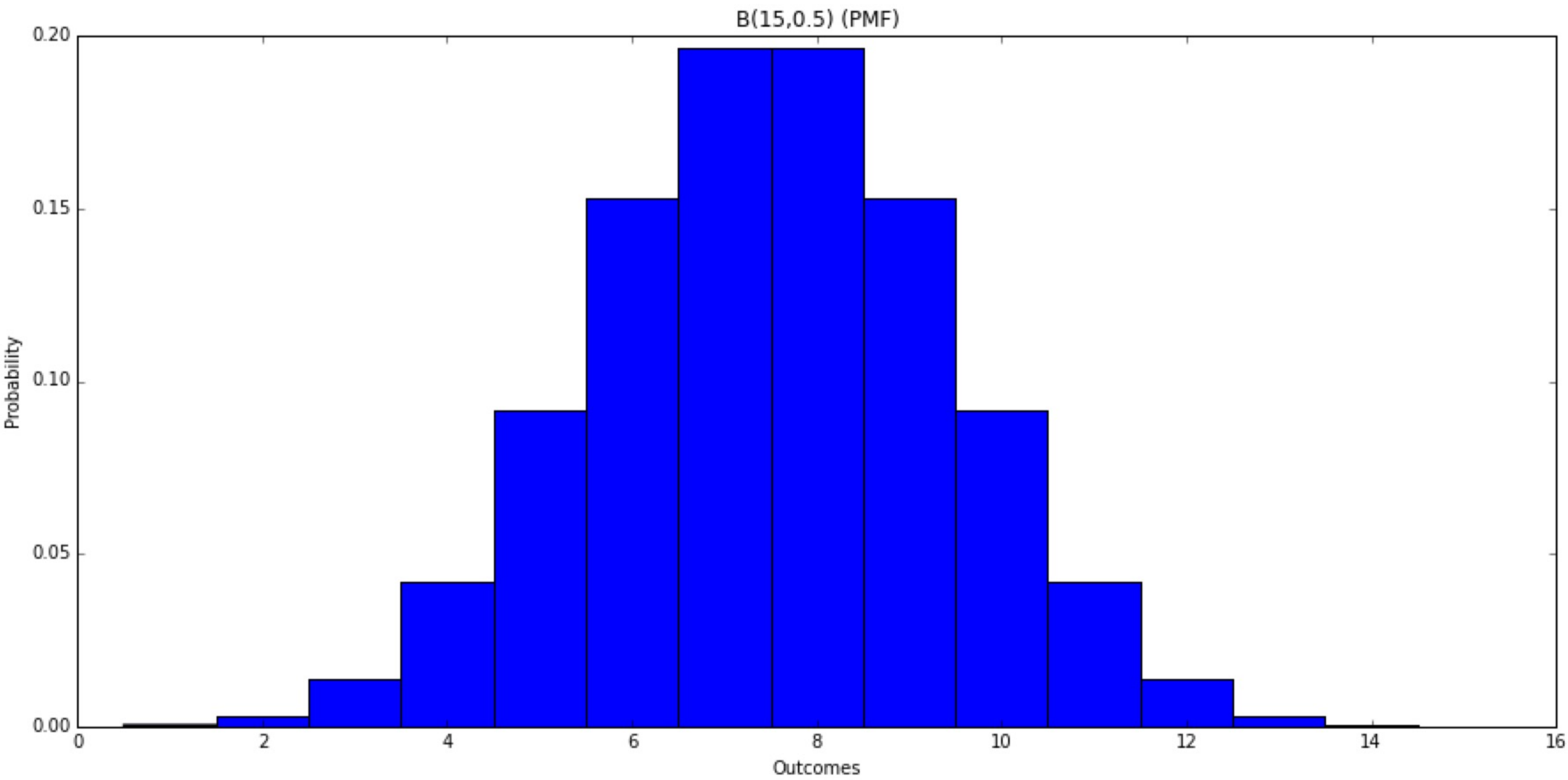
Normal Distribution as Limit of Binomial

How to approximate the binomial? When we observe the characteristic shape of the Binomial Distribution B(N,0.5) as N approaches Infinity, we see something interesting:



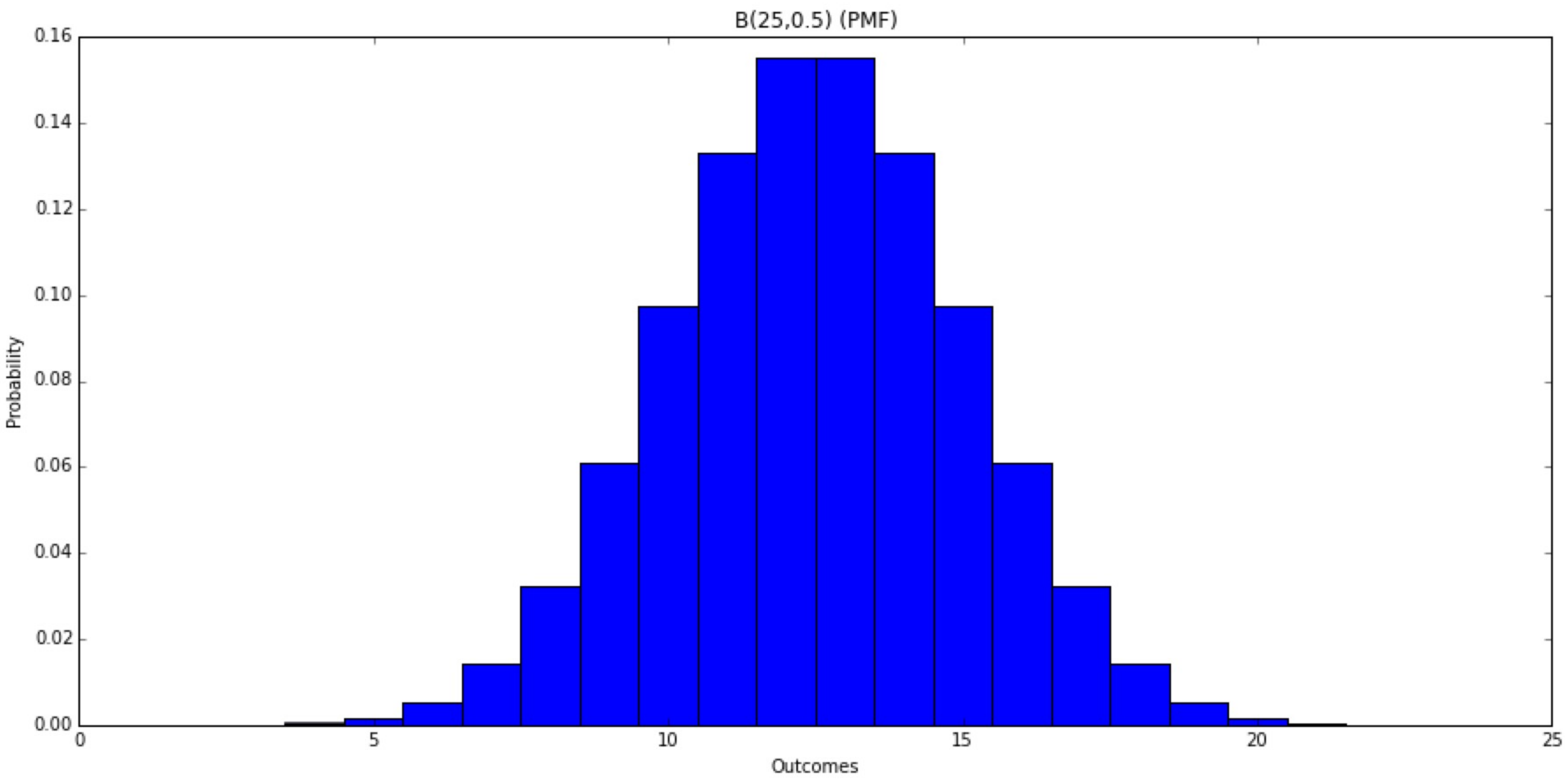
Normal Distribution as Limit of Binomial

How to approximate the binomial? When we observe the characteristic shape of the Binomial Distribution B(N,0.5) as N approaches Infinity, we see something interesting:



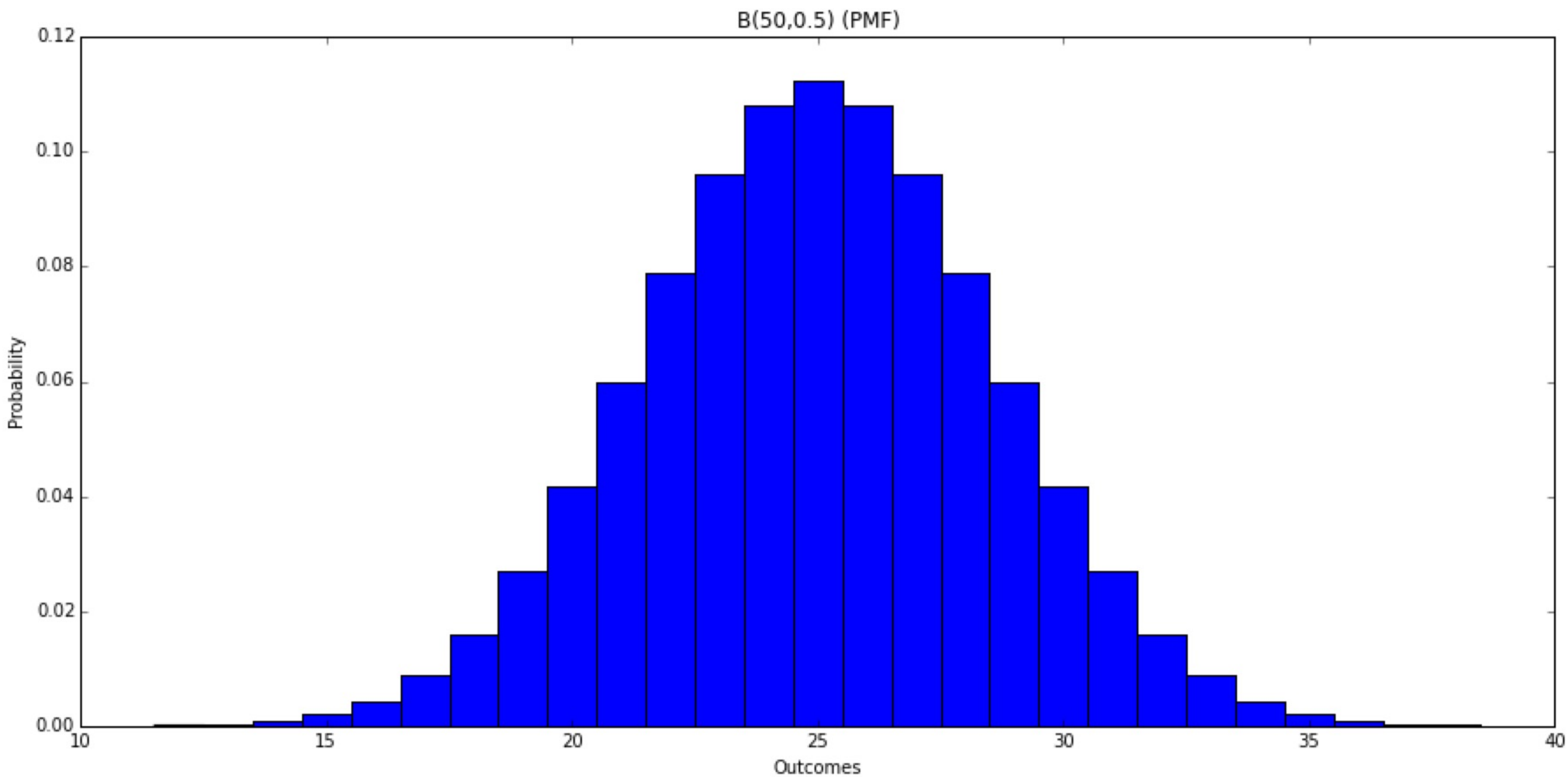
Normal Distribution as Limit of Binomial

How to approximate the binomial? When we observe the characteristic shape of the Binomial Distribution B(N,0.5) as N approaches Infinity, we see something interesting:



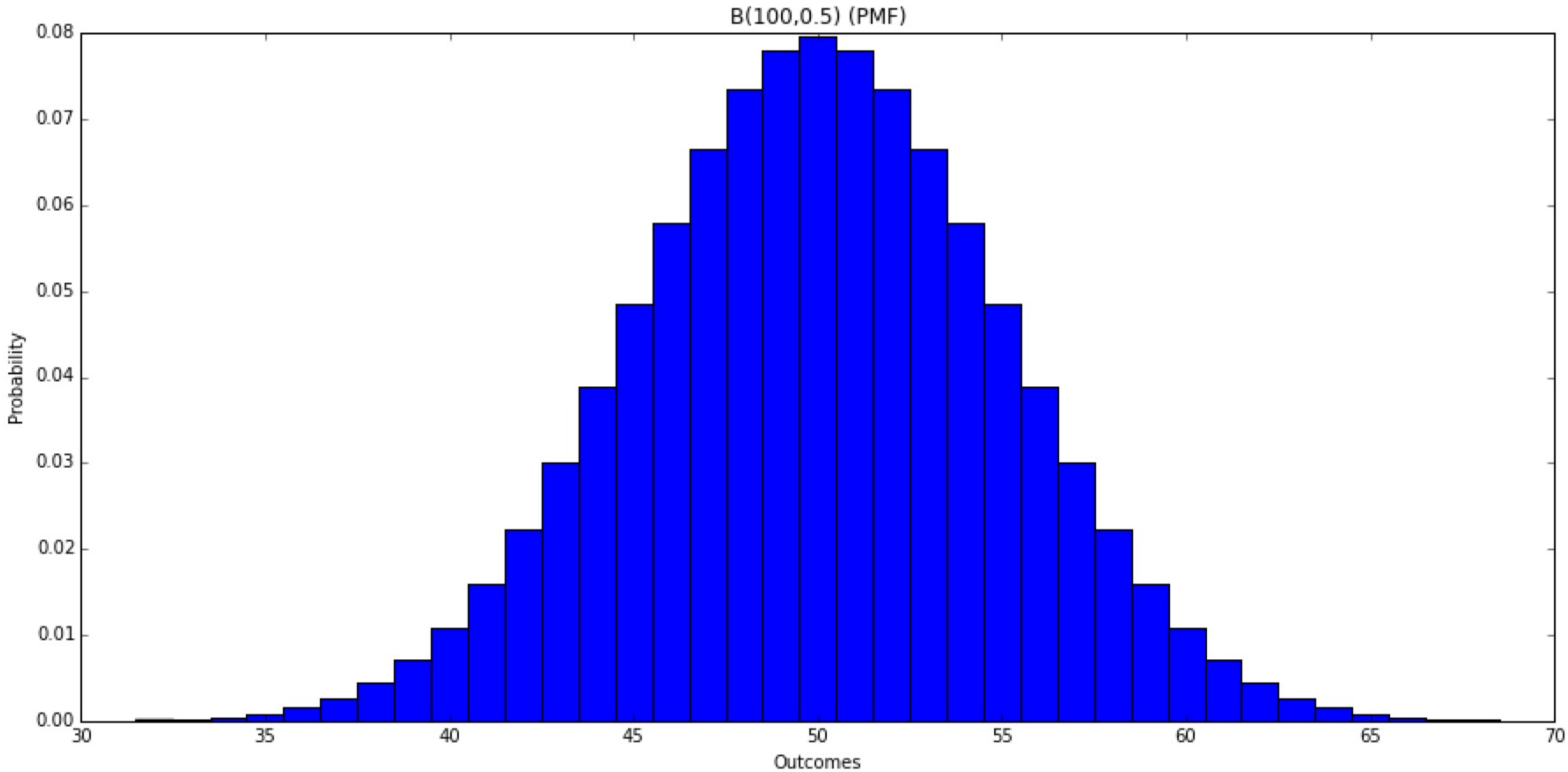
Normal Distribution as Limit of Binomial

How to approximate the binomial? When we observe the characteristic shape of the Binomial Distribution B(N,0.5) as N approaches Infinity, we see something interesting:



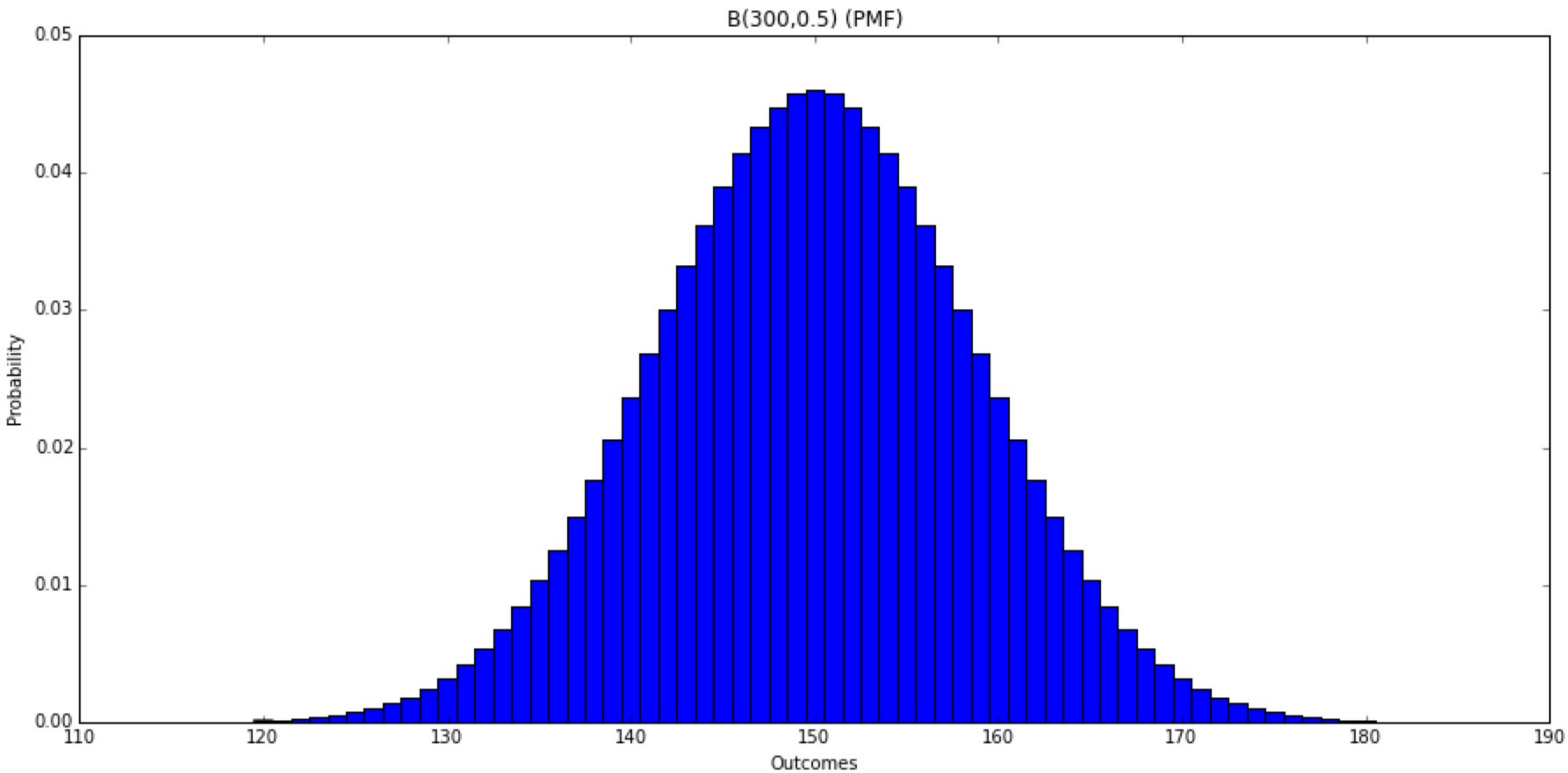
Normal Distribution as Limit of Binomial

How to approximate the binomial? When we observe the characteristic shape of the Binomial Distribution B(N,0.5) as N approaches Infinity, we see something interesting:



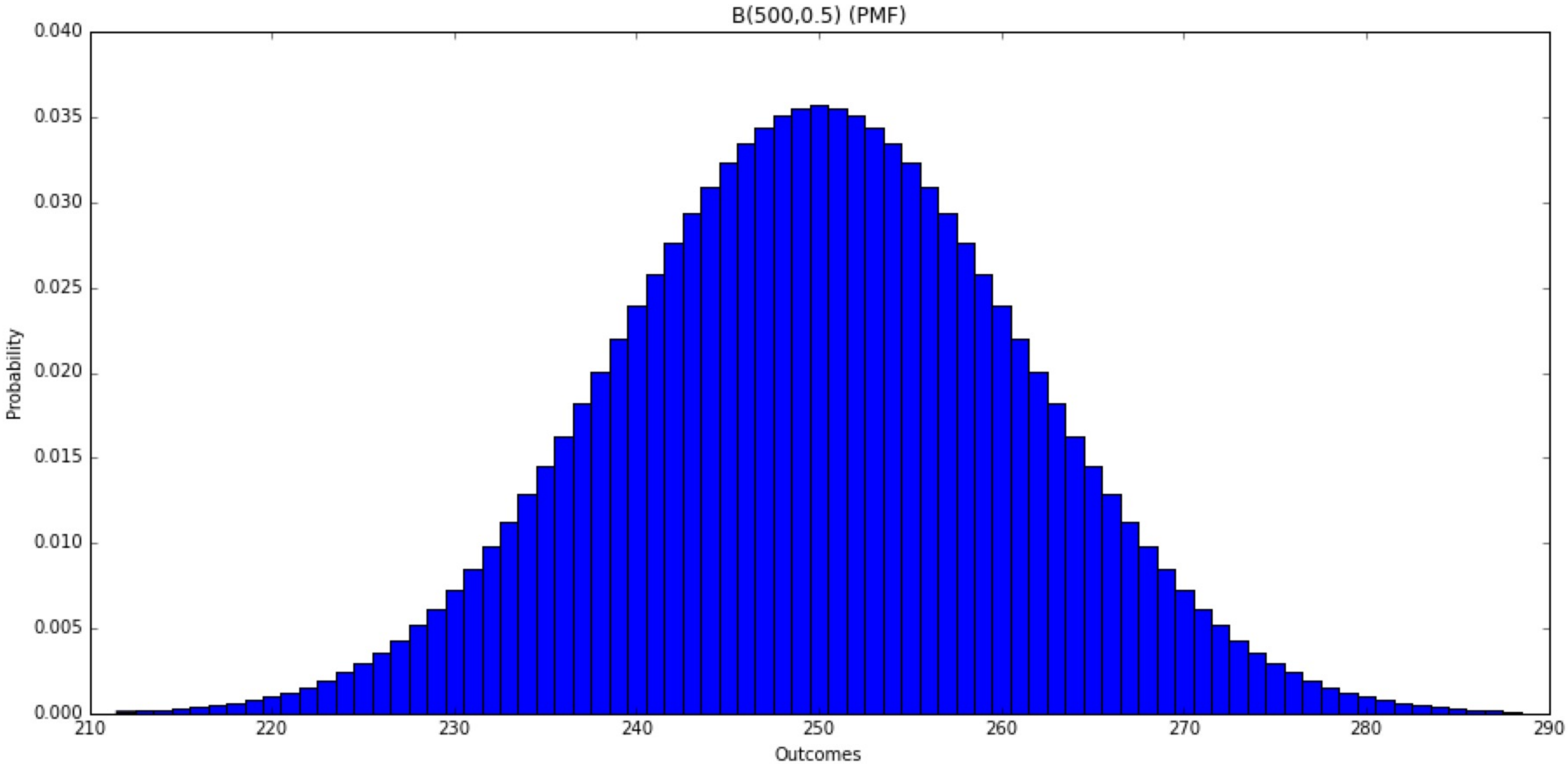
Normal Distribution as Limit of Binomial

How to approximate the binomial? When we observe the characteristic shape of the Binomial Distribution B(N,0.5) as N approaches Infinity, we see something interesting:



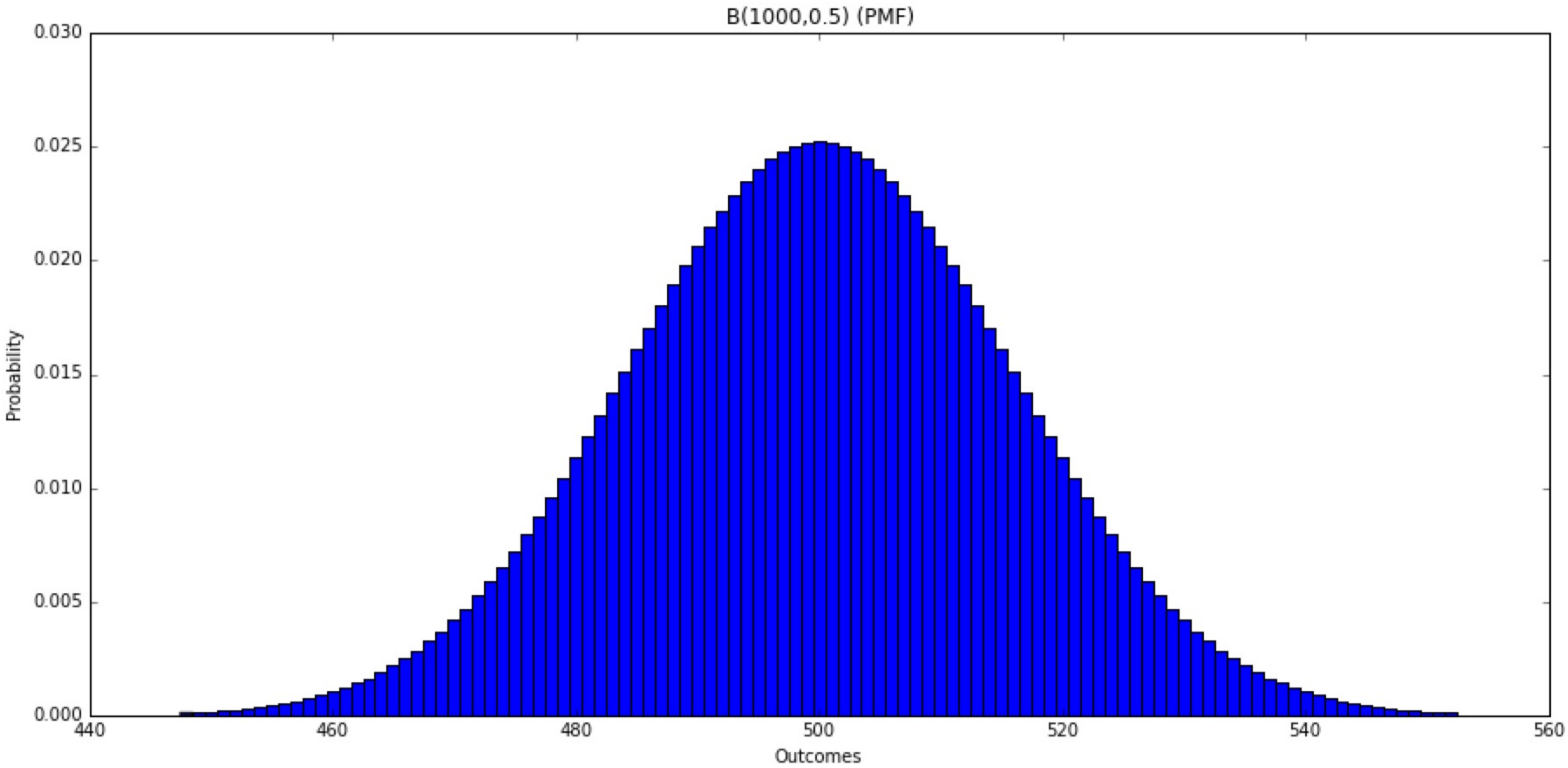
Normal Distribution as Limit of Binomial

How to approximate the binomial? When we observe the characteristic shape of the Binomial Distribution B(N,0.5) as N approaches Infinity, we see something interesting:



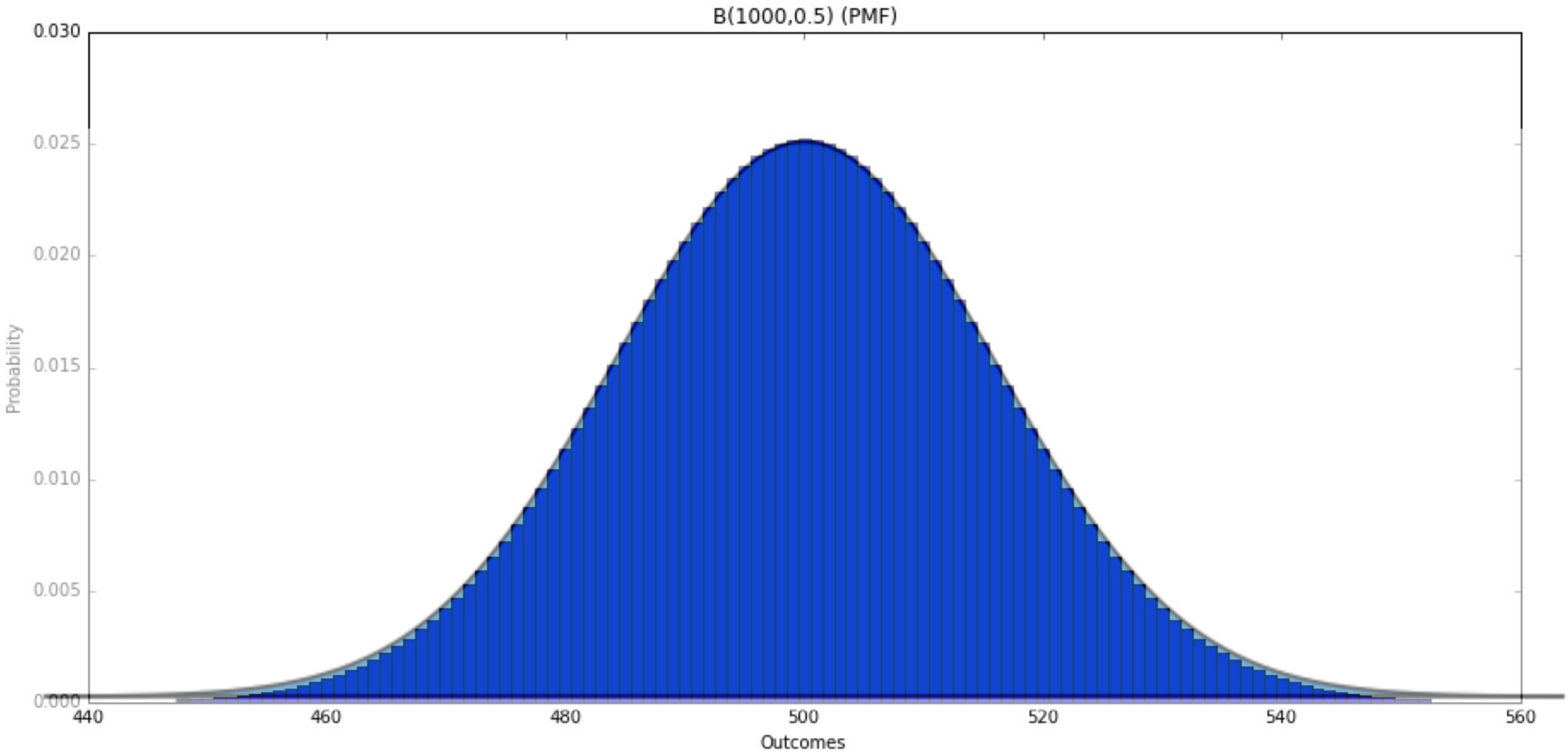
Normal Distribution as Limit of Binomial

How to approximate the binomial? When we observe the characteristic shape of the Binomial Distribution B(N,0.5) as N approaches Infinity, we see something interesting:



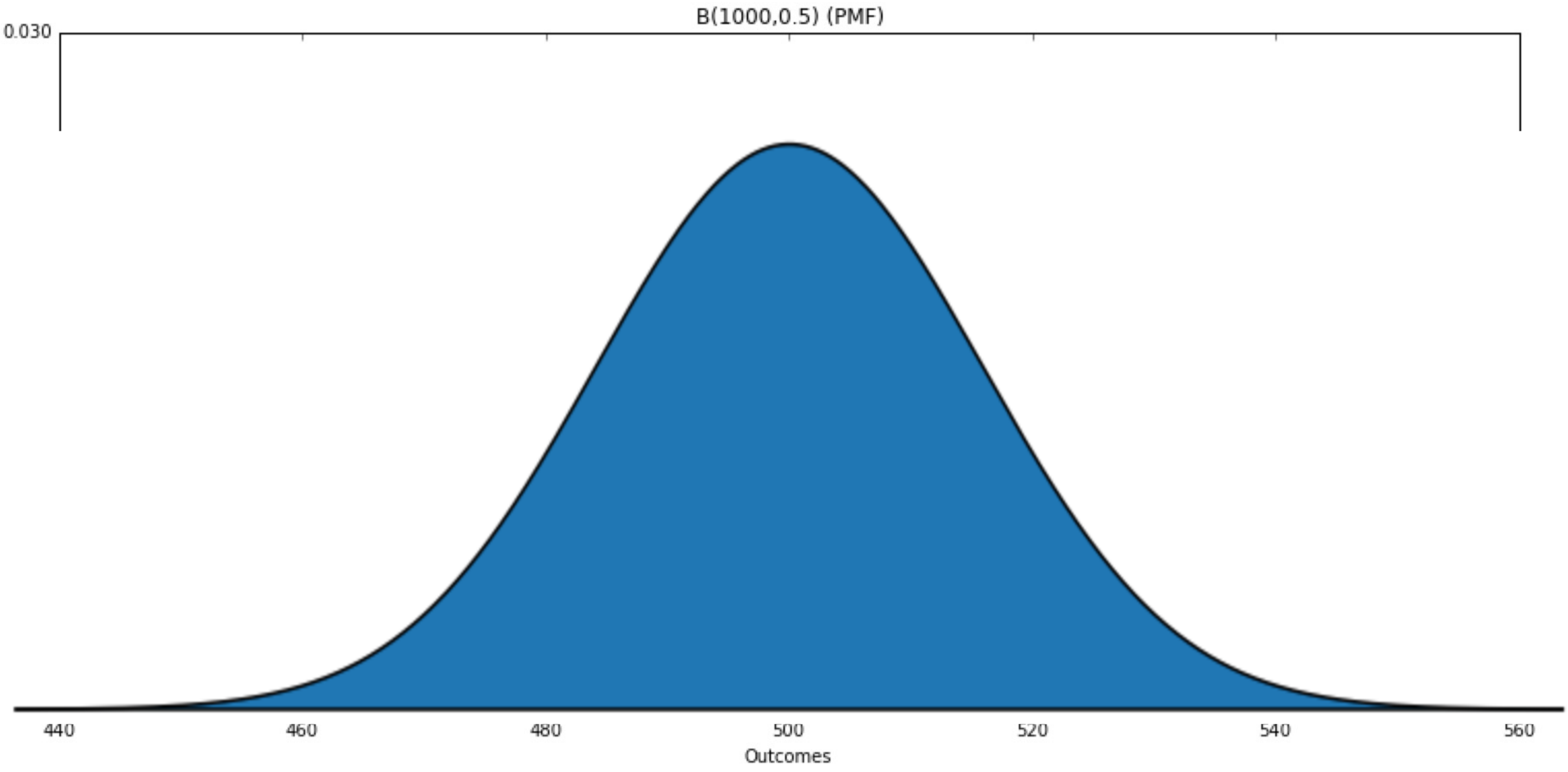
Normal Distribution as Limit of Binomial

How to approximate the binomial? When we observe the characteristic shape of the Binomial Distribution B(N,0.5) as N approaches Infinity, we see something interesting:



Normal Distribution as Limit of Binomial

How to approximate the binomial? When we observe the characteristic shape of the Binomial Distribution B(N,0.5) as N approaches Infinity, we see something interesting:



Normal Distribution

By using parameters to fit the requirements of probability theory (e.g., that the area under the curve is 1.0), we have the formula for the Normal Distribution, which can be used to approximate the Binomial Distribution and which models a wide variety of random phenomena:

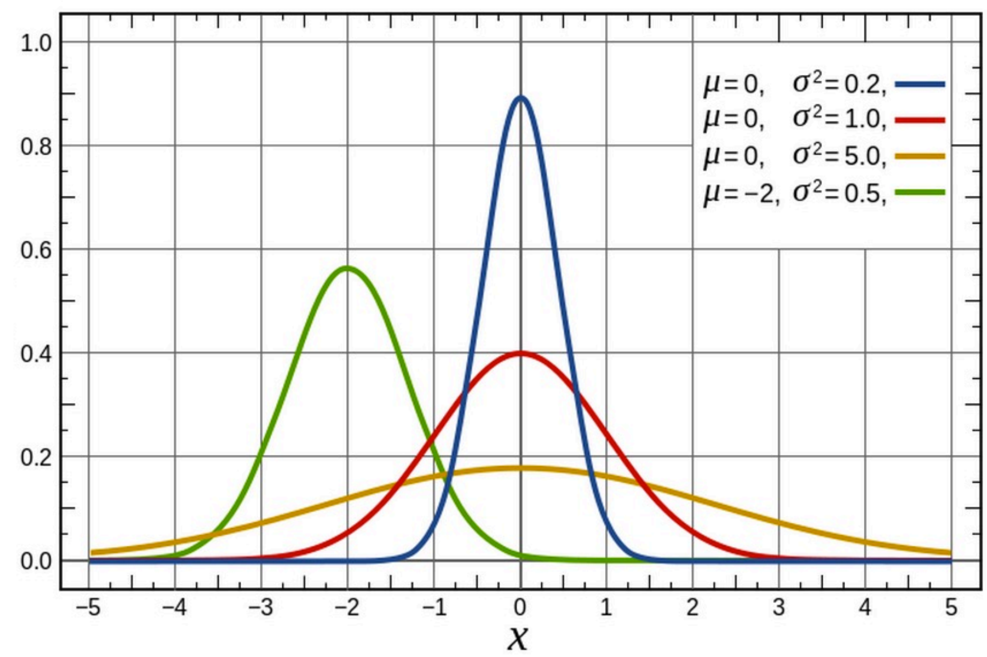
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

where

μ = mean/expected value

σ = standard deviation

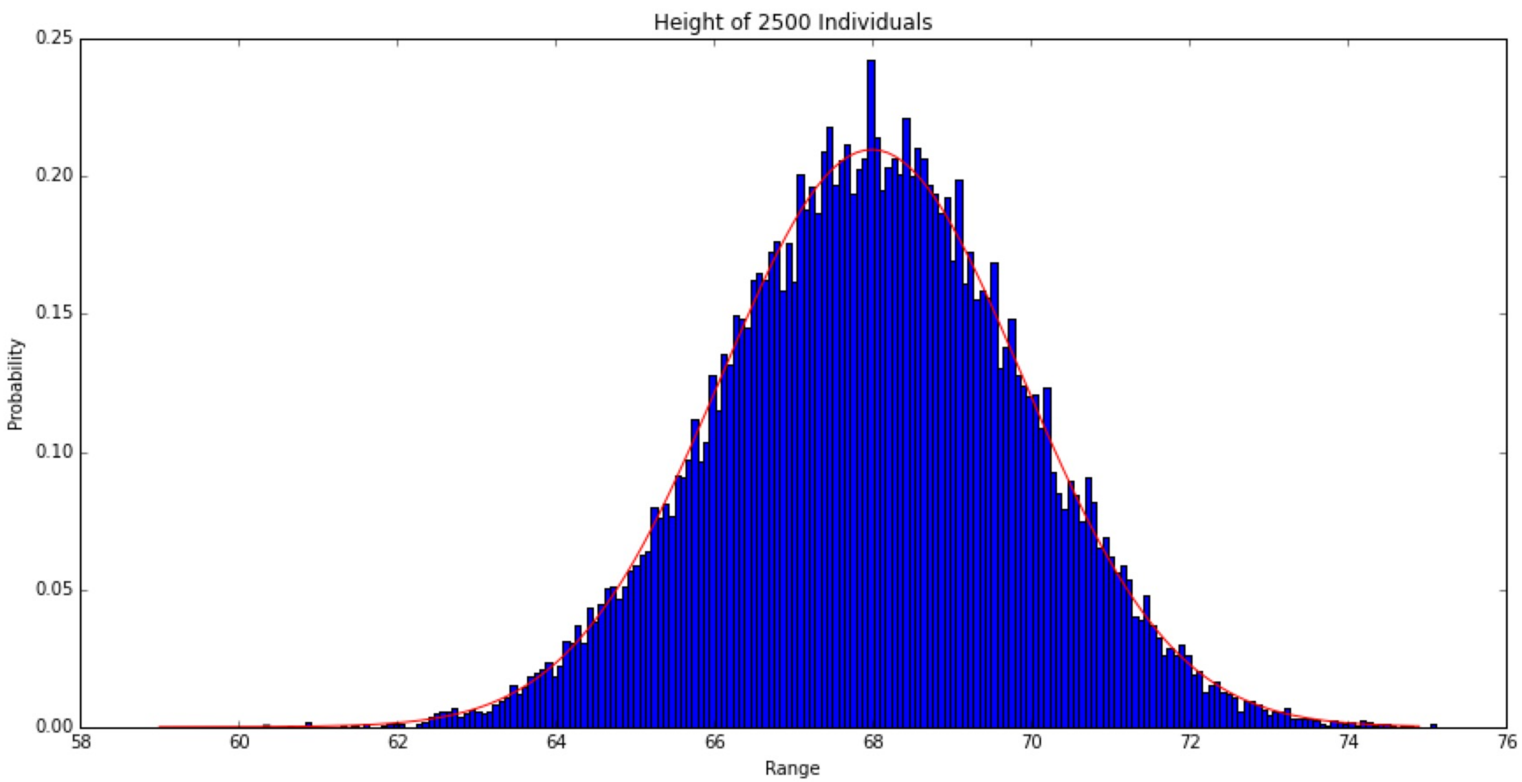
σ^2 = variance



Normal Distribution

The normal distribution, as the limit of B(N,0.5), occurs when a very large number of factors add together to create some random phenomenon.

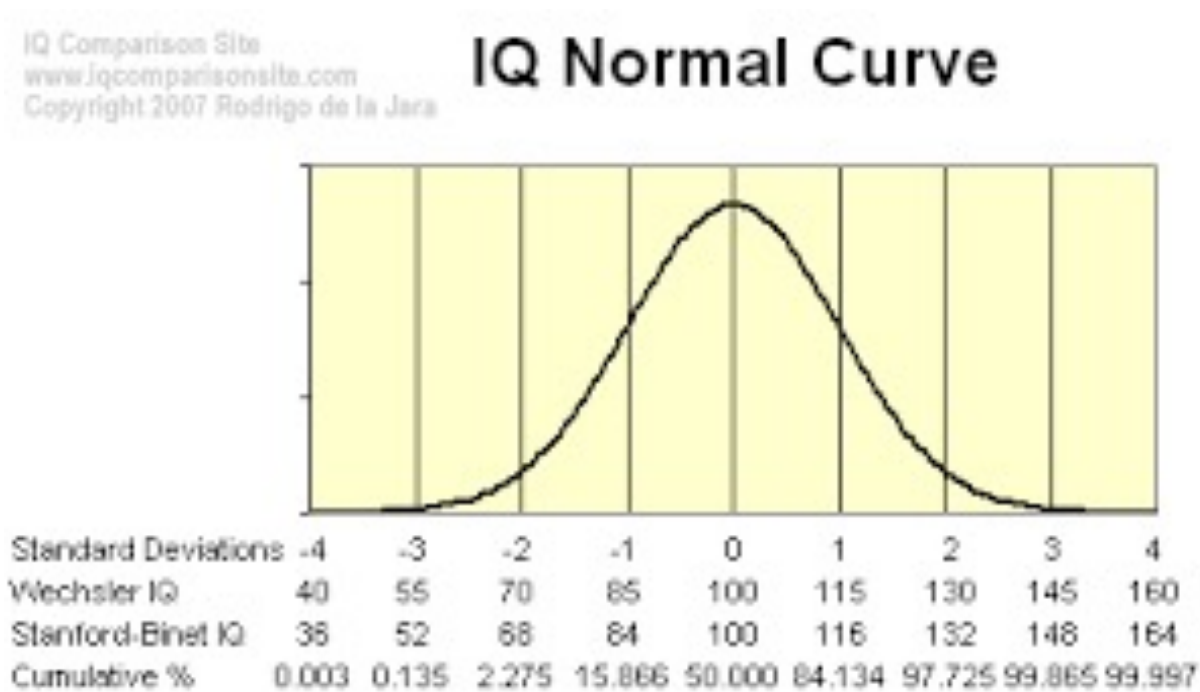
Example: What is the height of a human being?



Normal Distribution

The normal distribution, as the limit of B(N,0.5), occurs when a very large number of factors add together to create some random phenomenon.

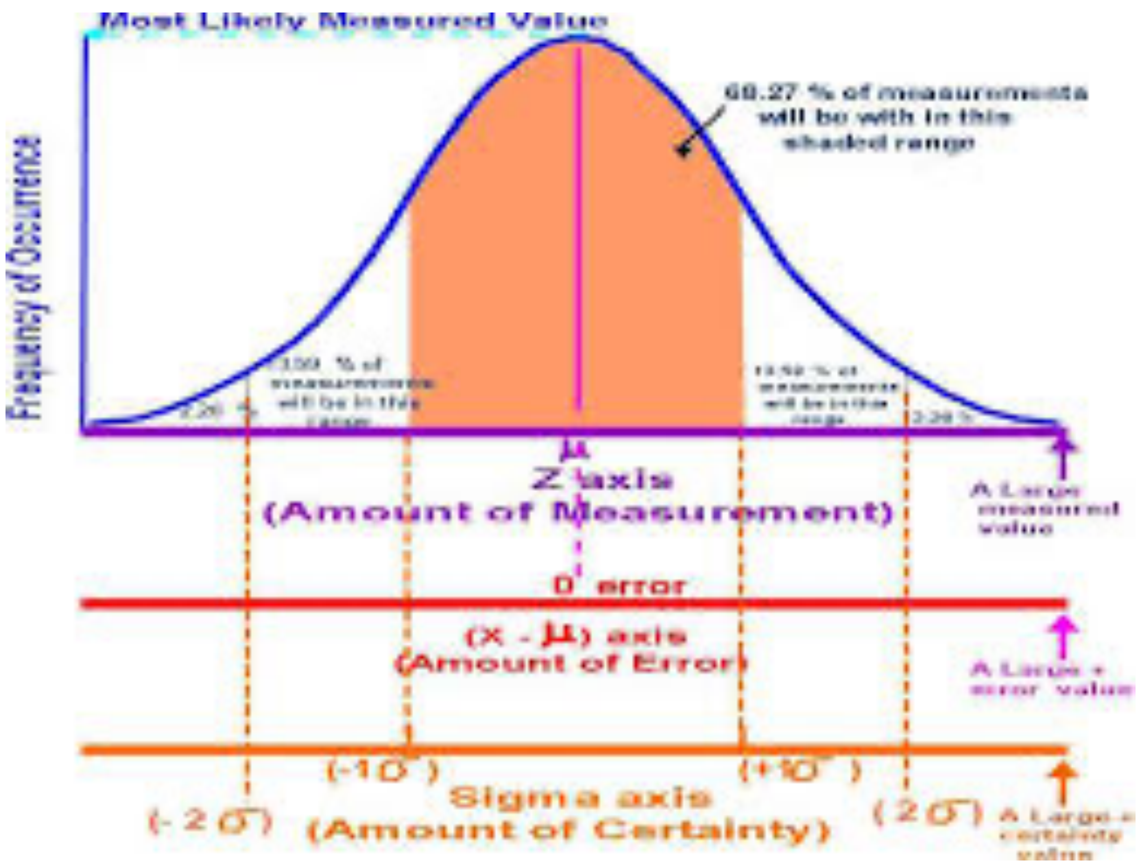
Example: What is the IQ of a human being?



Normal Distribution

The normal distribution, as the limit of B(N,0.5), occurs when a very large number of factors add together to create some random phenomenon.

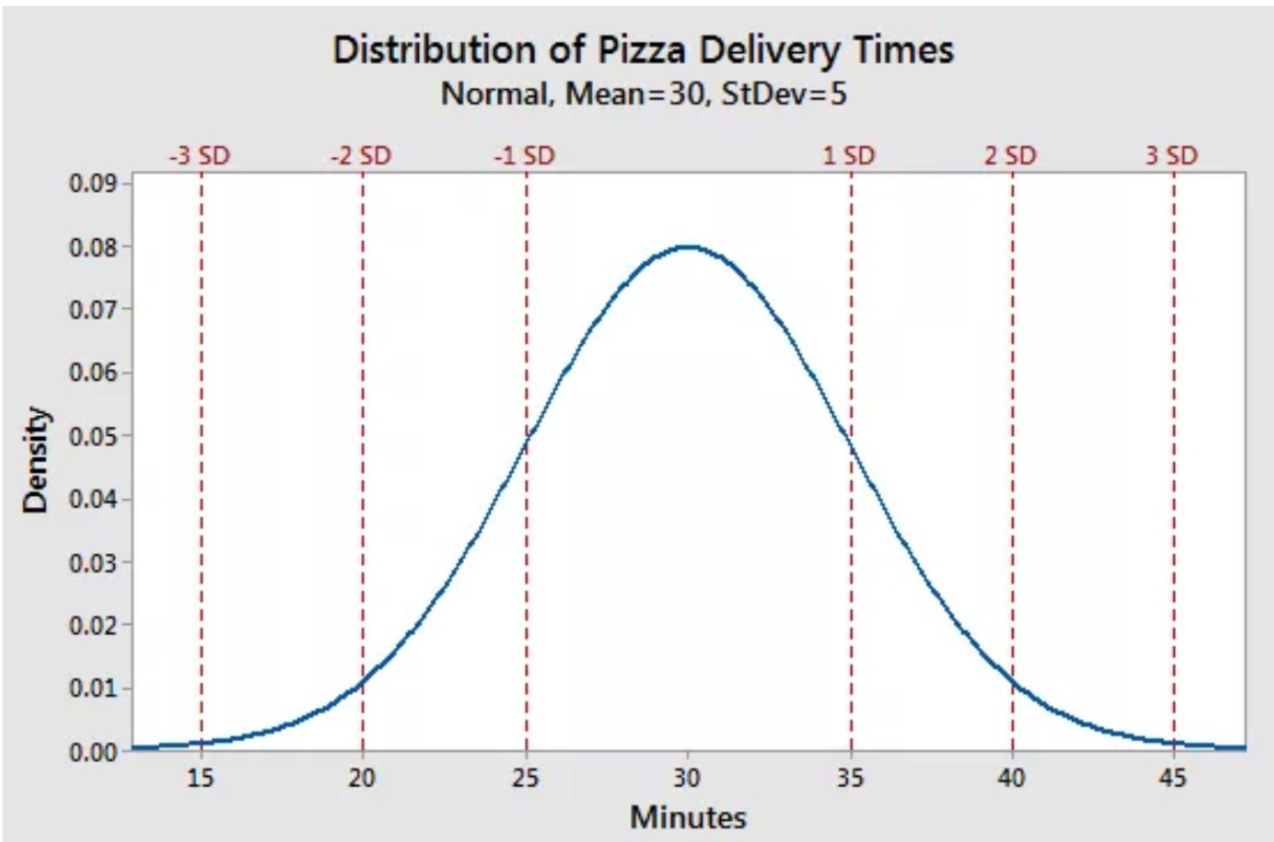
Example: What is the distribution of measurement errors?



Normal Distribution

The normal distribution, as the limit of B(N,0.5), occurs when a very large number of factors add together to create some random phenomenon.

Example: Even REALLY IMPORTANT things are normally distributed!



Normal Distribution

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

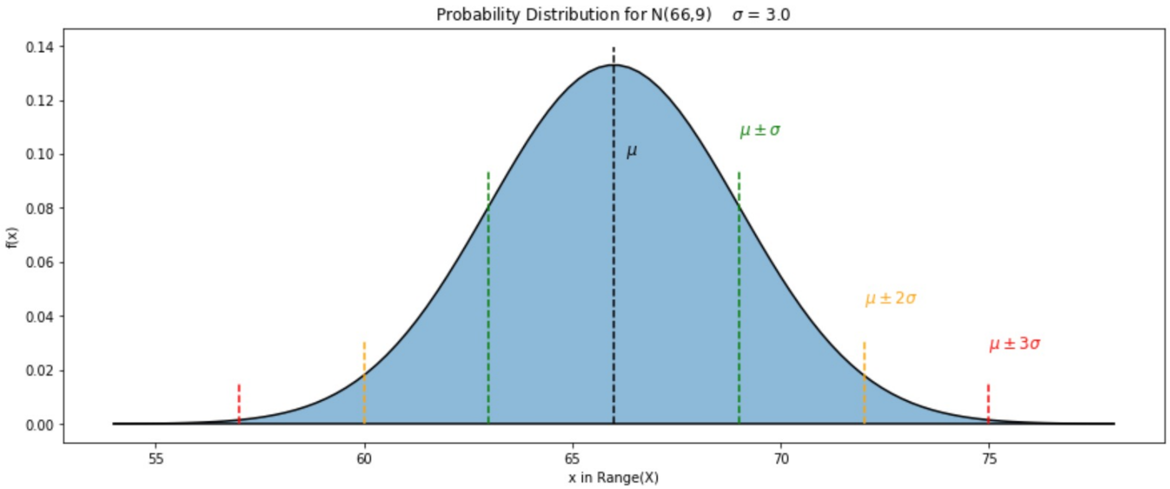
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

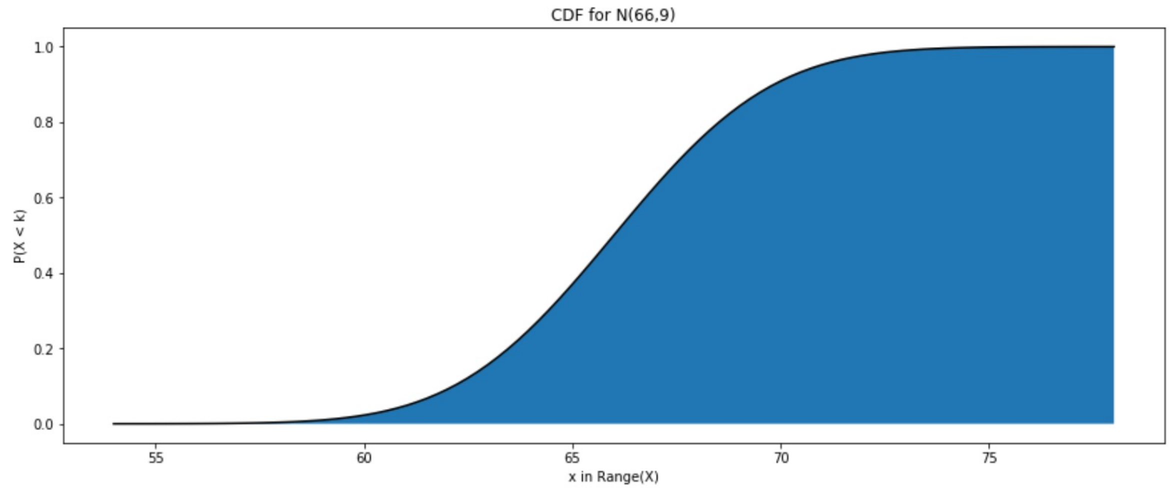
$$F(a) = \int_{-\infty}^{a} f(x)\, dx$$

P(X < a) = F(a)

P(X > a) = 1.0 – F(a)

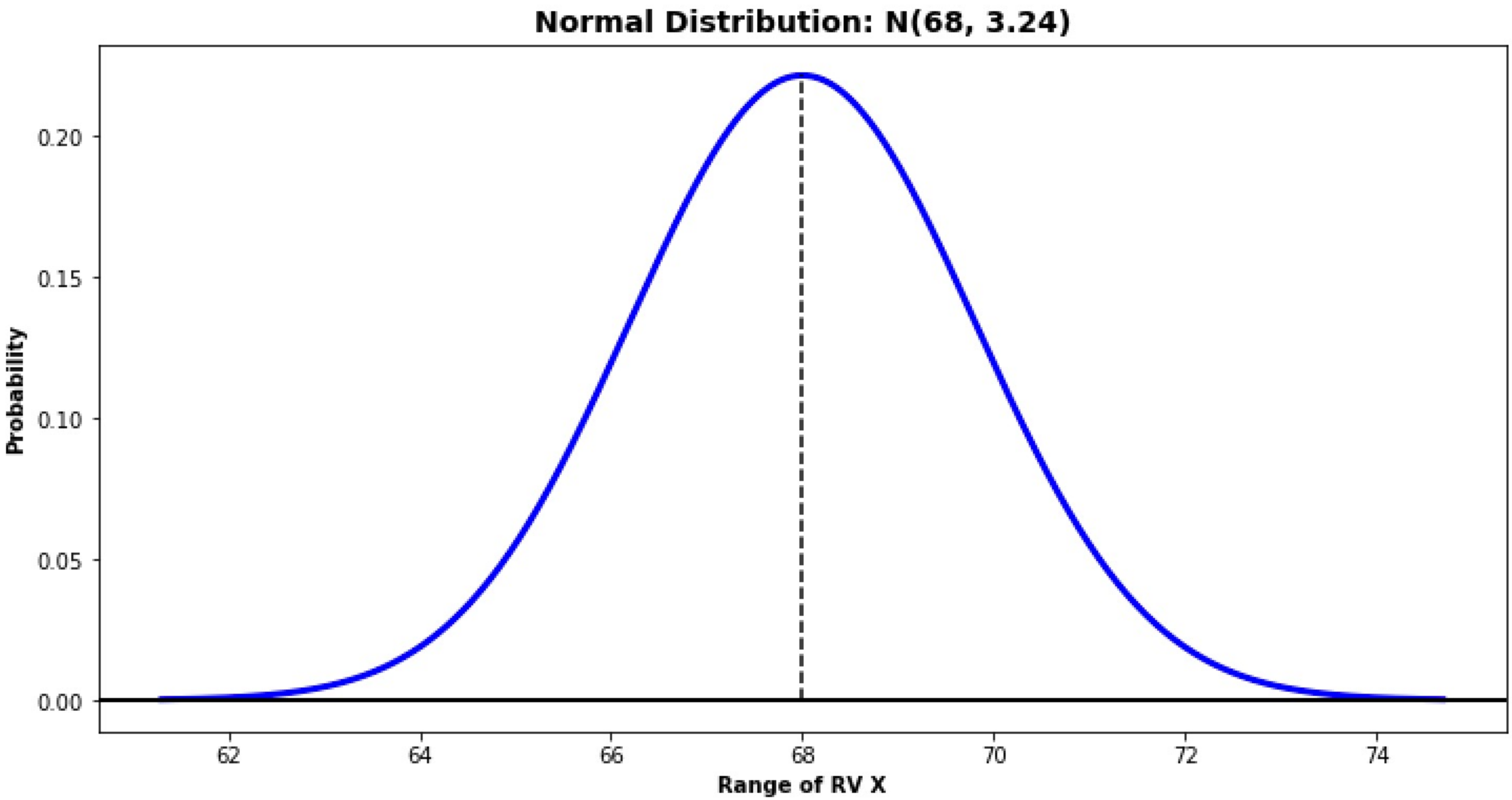
P(a < X < b) = F(b) – F(a)





Normal Distribution

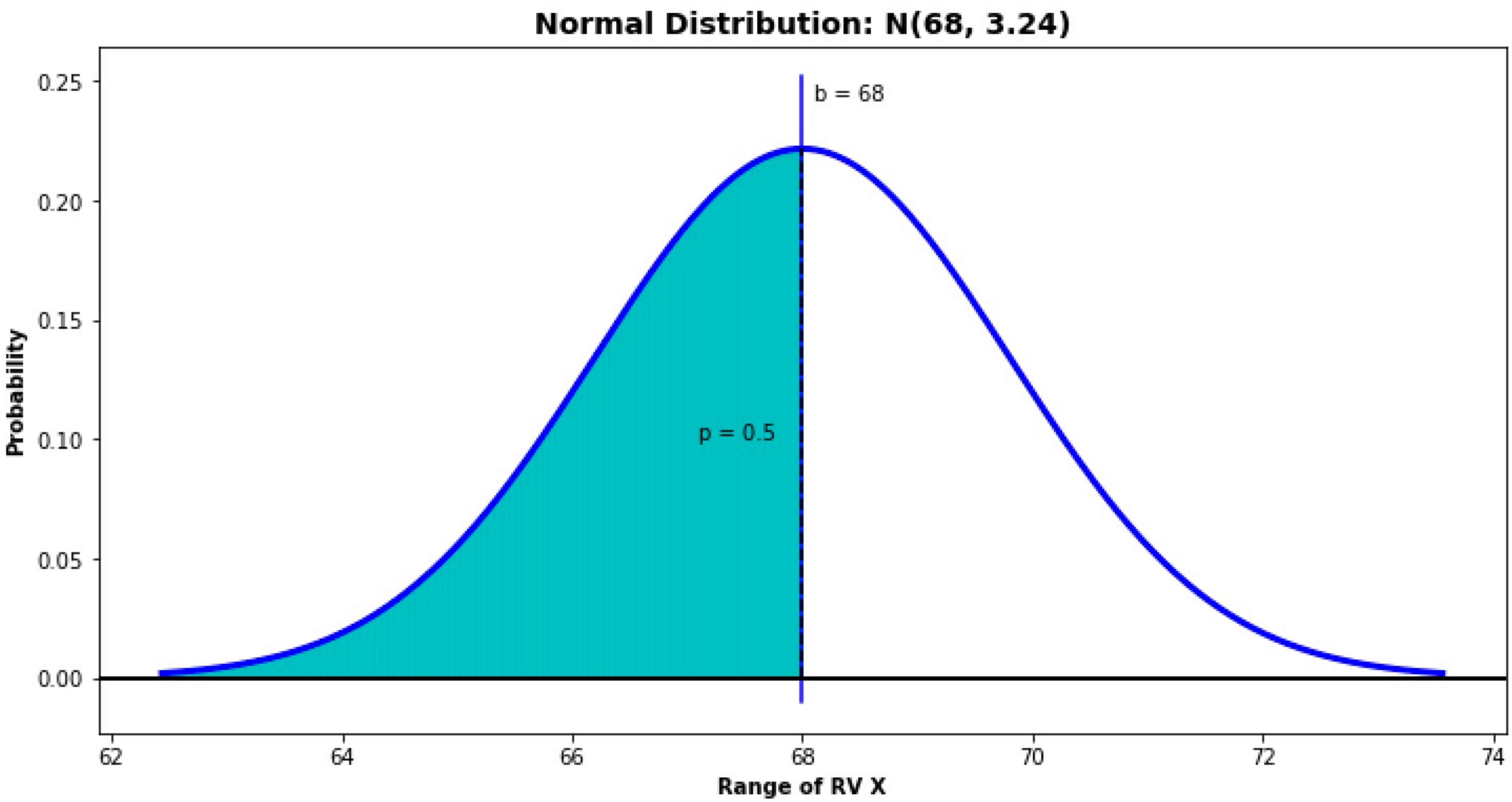
Suppose heights at BU are distributed normally with a mean of 68 inches and a standard deviation of 1.8 inches.




```
mean = 68
var = 3.24
stdev = 1.8
```

Normal Distribution

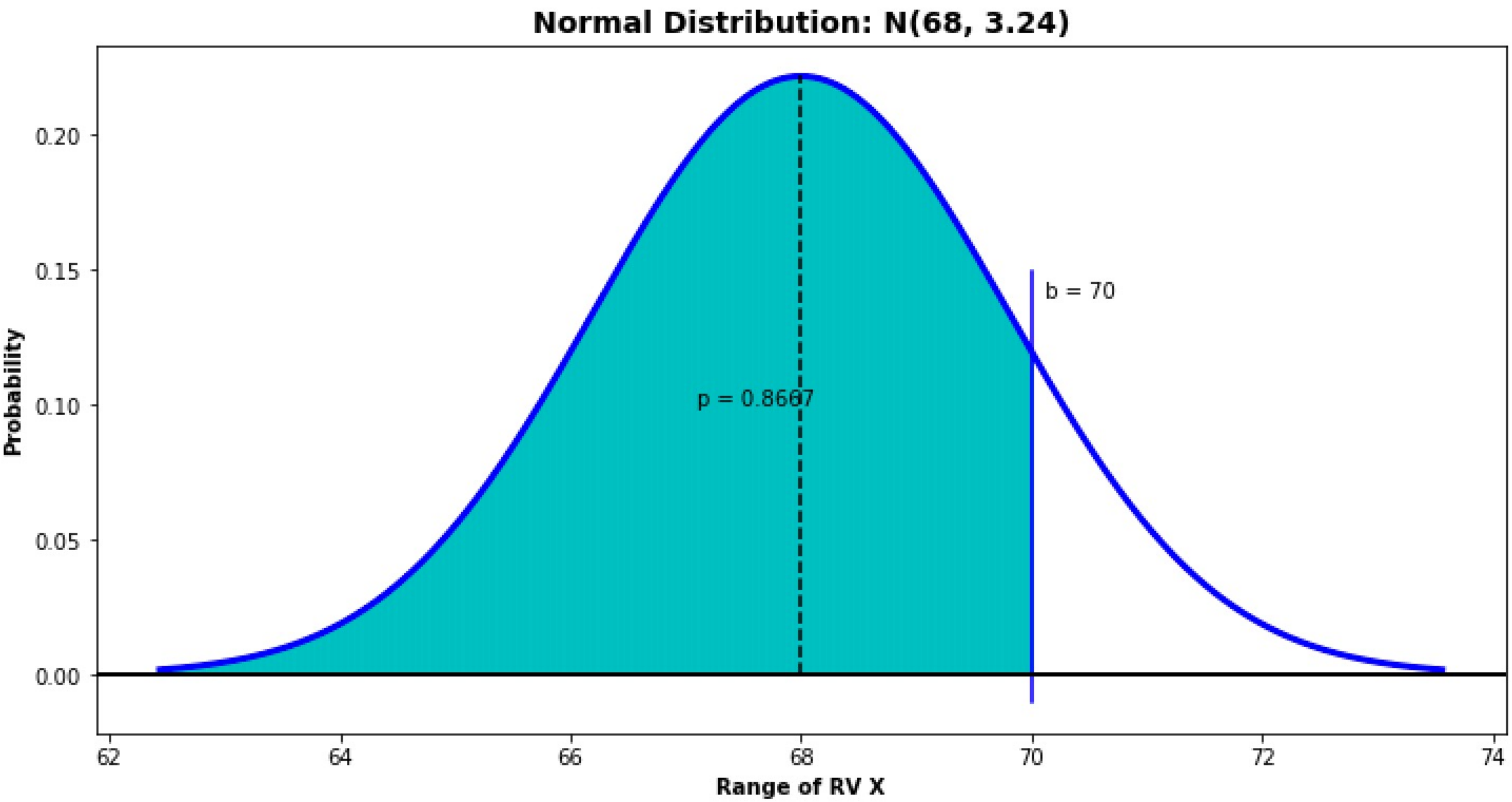
How many people are of less than average height?




```
mean = 68
var = 3.24
stdev = 1.8
P(X < 68) = 0.5
```

Normal Distribution

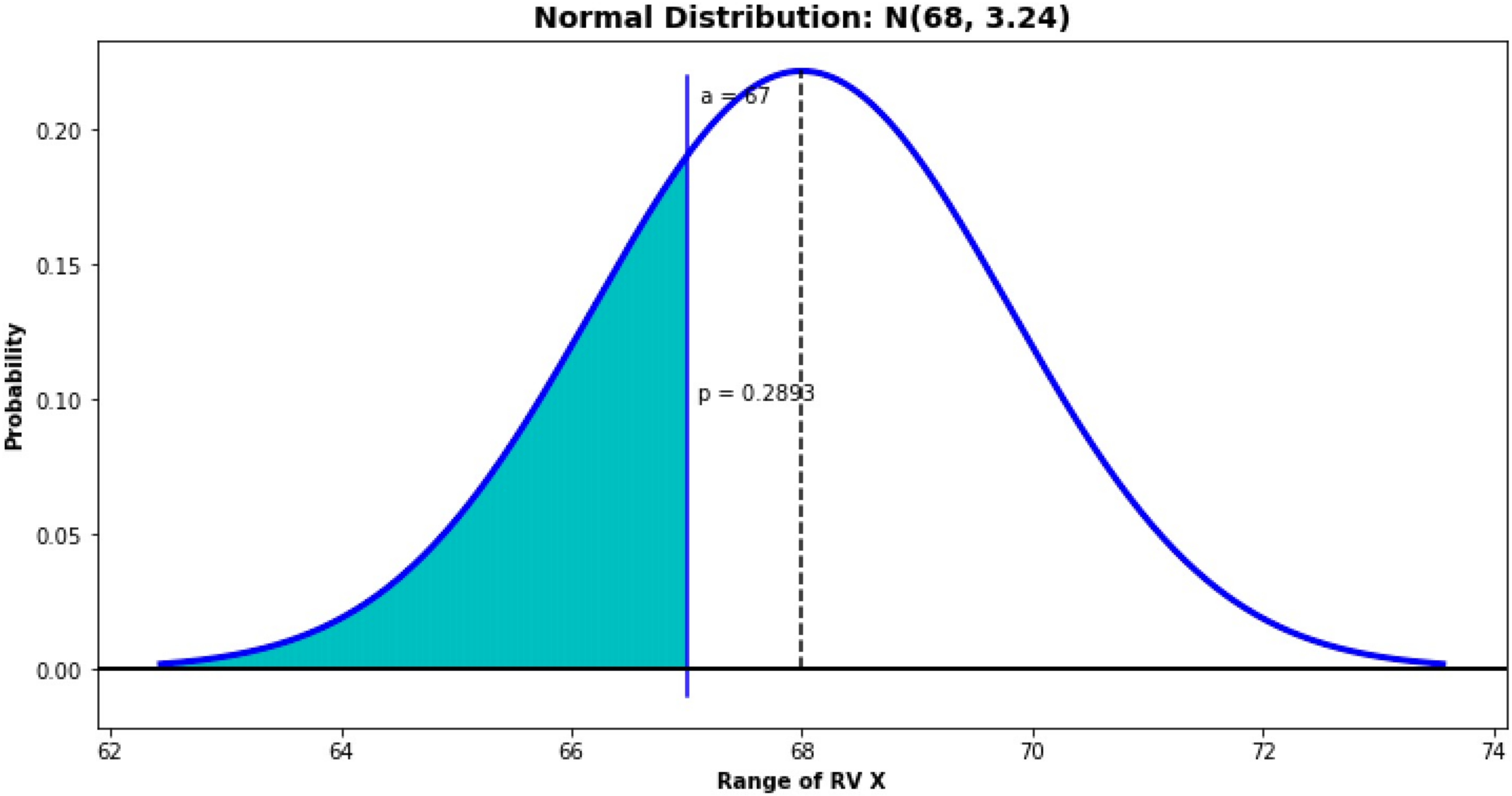
How many people are less than 70 inches?




```
mean = 68
var = 3.24
stdev = 1.8
P(X < 70) = 0.8667
```

Normal Distribution

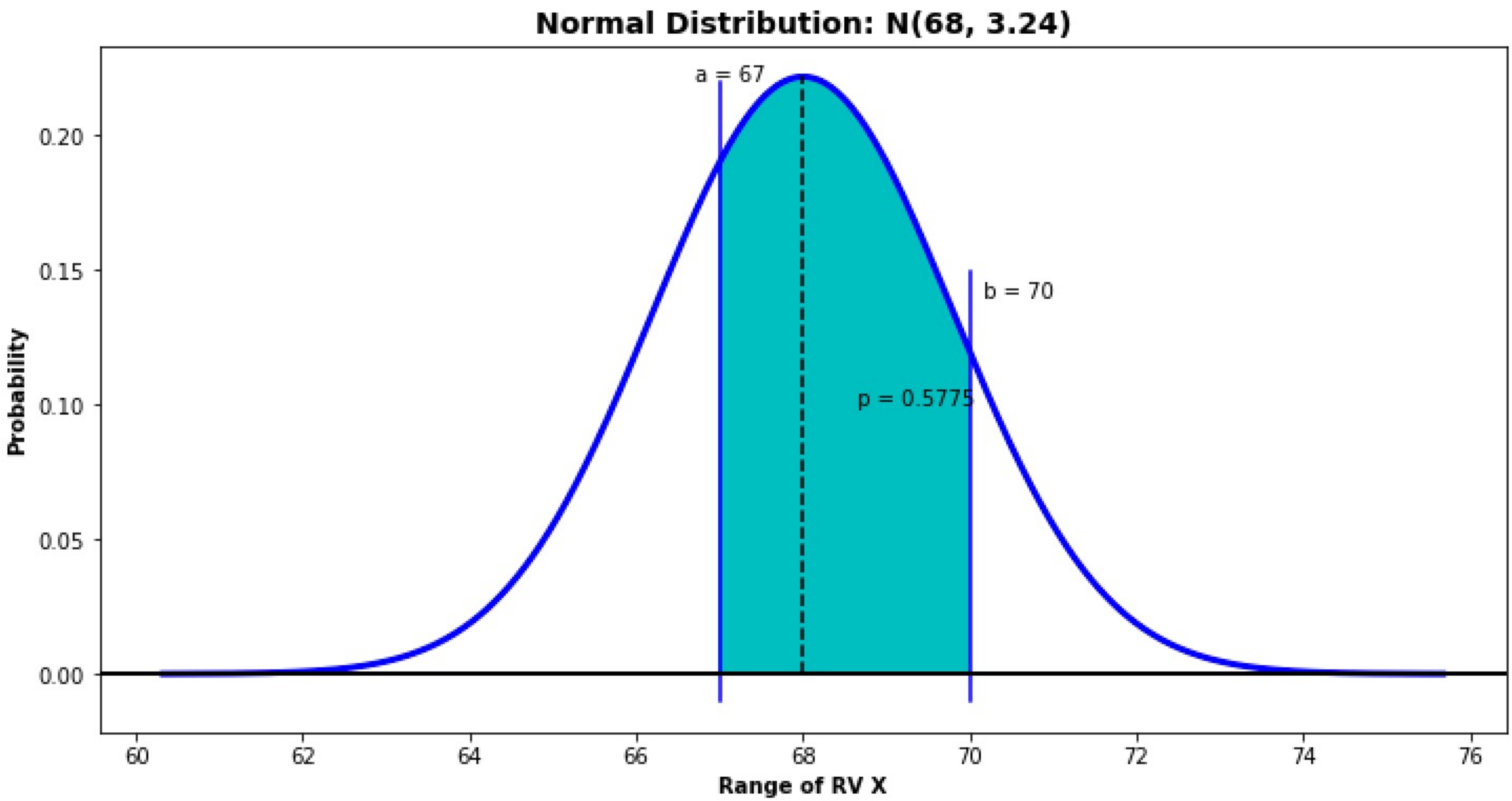
How many people are less than 67 inches?




```
mean = 68
var = 3.24
stdev = 1.8
P(X < 67) = 0.2893
```

Normal Distribution

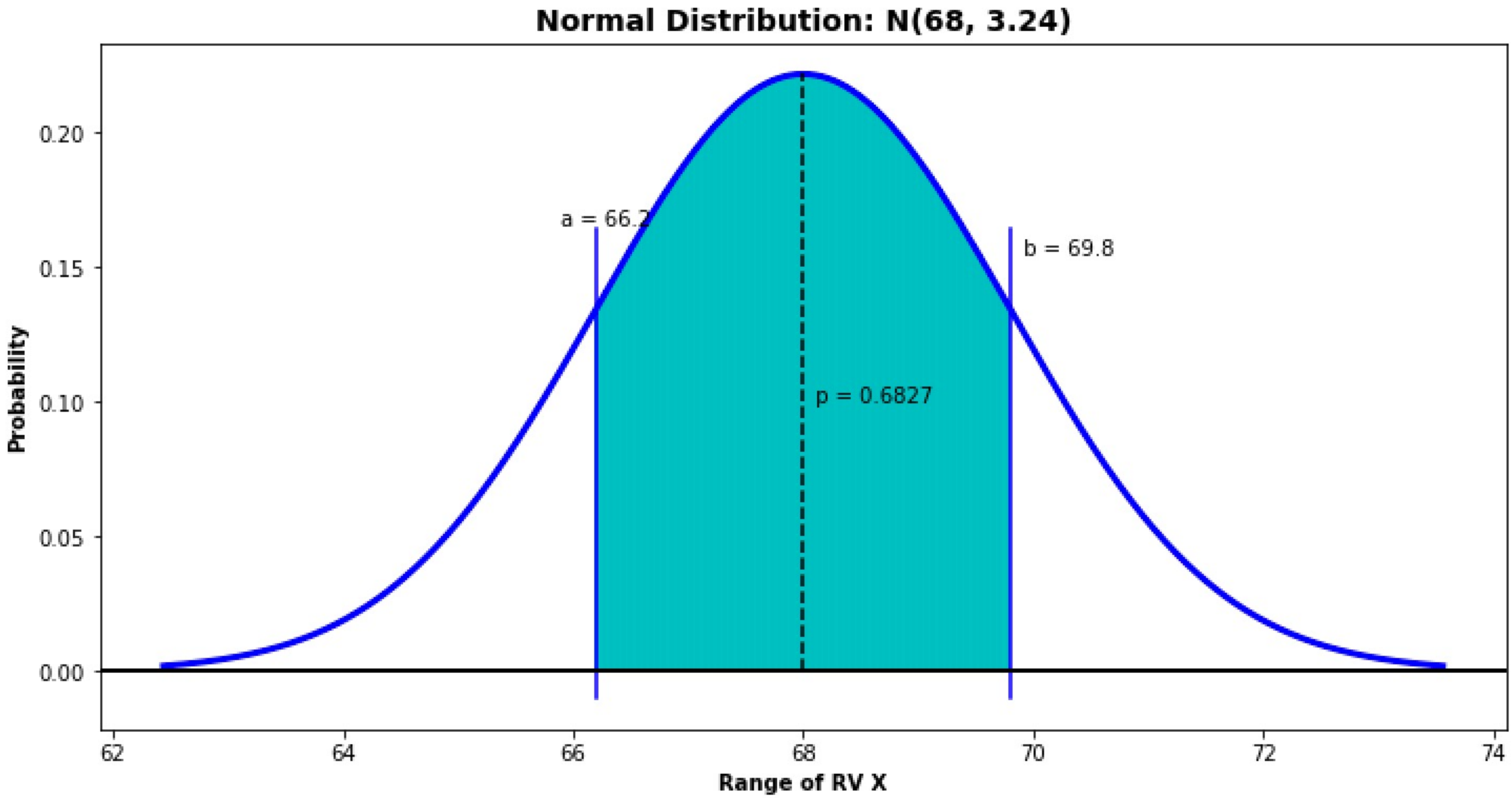
How many people are between 67 and 70 inches?




```
mean = 68
var = 3.24
stdev = 1.8
P(67 < X < 70) = P(X < 70) - P(X < 67) = 0.8667 - 0.2893 = 0.5775
```

Normal Distribution

How many people are within one standard deviation of the mean height?




```
mean = 68
var = 3.24
stdev = 1.8
P(66.2 < X < 69.8) = P(X < 69.8) - P(X < 66.2) = 0.8413 - 0.1587 = 0.6827
```

Normal Distribution

Modern people use the appropriate formulae:

```
def f_normal(mu,var,x):
    return (1/(math.sqrt(var*2*math.pi))) * math.exp(-(x-mu)*(x-mu)/(2*var))

def F_normal(mu,var,x):
    return (1 + math.erf((x-mu)/(var**0.5 * 2.0**0.5))) / 2

def normalRange(mu,var,low,high):
    return F_normal(mu,var,high) - F_normal(mu,var,low)


# OR use the scipy.stats.norm functions given at the top of the notebook:

# Loc = mean, scale = standard deviation

norm.pdf(x=50,loc=40,scale=5)

norm.cdf(x=50,loc=40,scale=5)

norm.rvs(loc=40,scale=5)      # random variates
```

Normal Distribution

Or a calculator or a web site:

- Enter a value in three of the four text boxes.
- Leave the fourth text box blank.
- Click the **Calculate** button to compute a value for the blank text box.

Standard score (z)	1.5
Cumulative probability: $P(Z \leq 1.5)$	0.933
Mean	0
Standard deviation	1

Calculate