CS 237: Probability in Computing

Wayne Snyder
Computer Science Department
Boston University

Lecture 12:

- Continuous Distributions
- Exponential
- Uniform Distribution
- Normal Distribution (motivation)

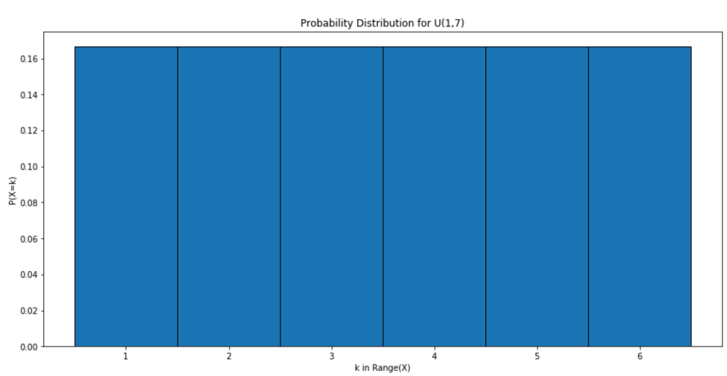
Review: Cumulative Distribution Functions

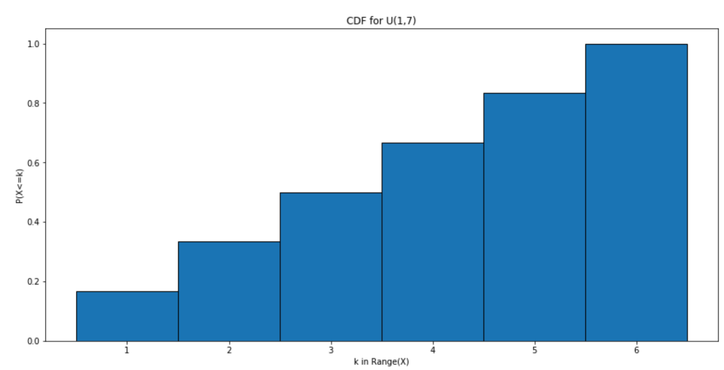
The Cumulative Distribution Function (CDF) for a random variable X shows what happens when we keep track of the sum of the probability distribution from left to right over its range:

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

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

Probability Distribution Function f_X Cumulative Distribution Function F_X





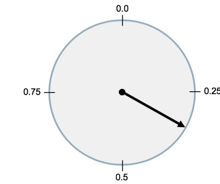
Continuous Distributions

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:

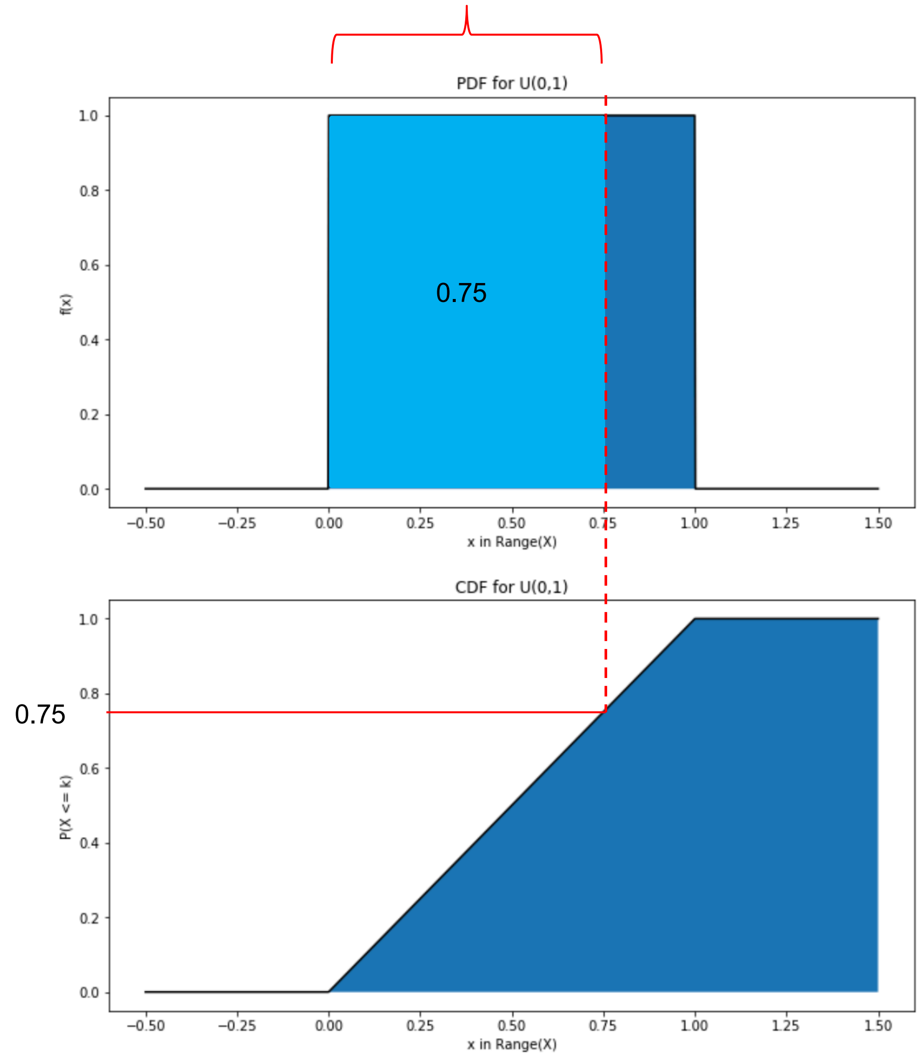
$$f(x) = \begin{cases} 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

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

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



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



Continuous Distributions

Discrete Random Variables

$$F_X(b) = P(X \leq b) =_{\text{def}} \sum_{x \leq b} f(x)$$

$$P(a \leq X \leq b) =_{\text{def}} \sum_{a \leq x \leq b} f(x)$$

$$E(X) = \sum_{x \in R_X} x \cdot f(x)$$

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

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

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

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

$$X^* =_{def} \frac{X - \mu_X}{\sigma_X}$$

Uniform Distribution

The simplest continuous distribution is similar to the spinner, but with arbitrary endpoints:

If X = “a random real number uniformly chosen from the interval [a..b]”

then X is a uniform random variable from a to b, denoted

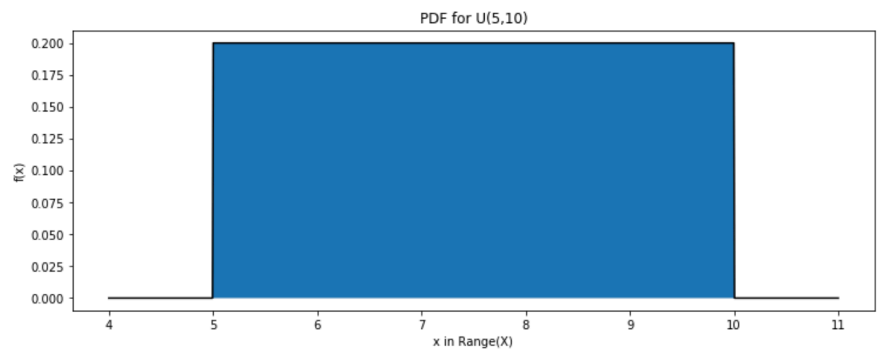
$$X \sim U(a, b)$$

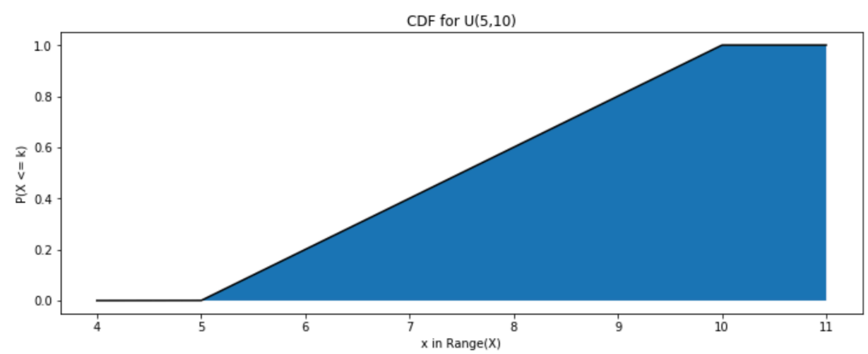
and where

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

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

$$X \sim U(5, 10)$$



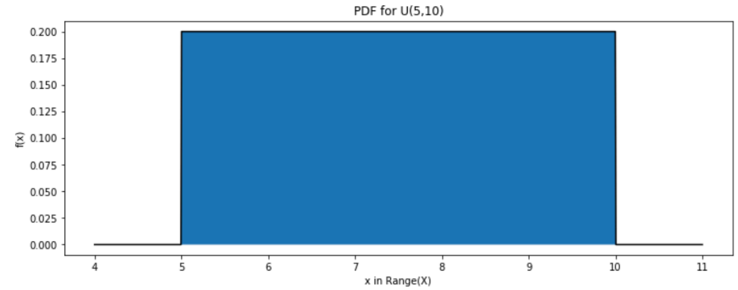


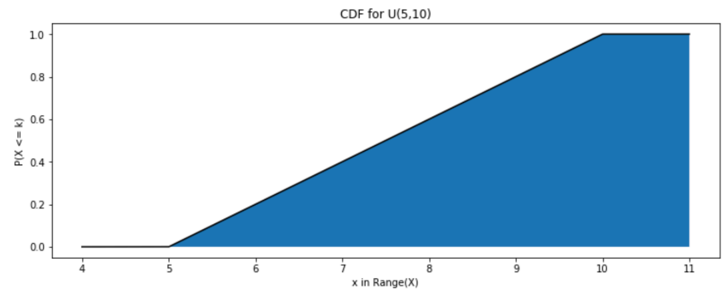
Uniform Distribution

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

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

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



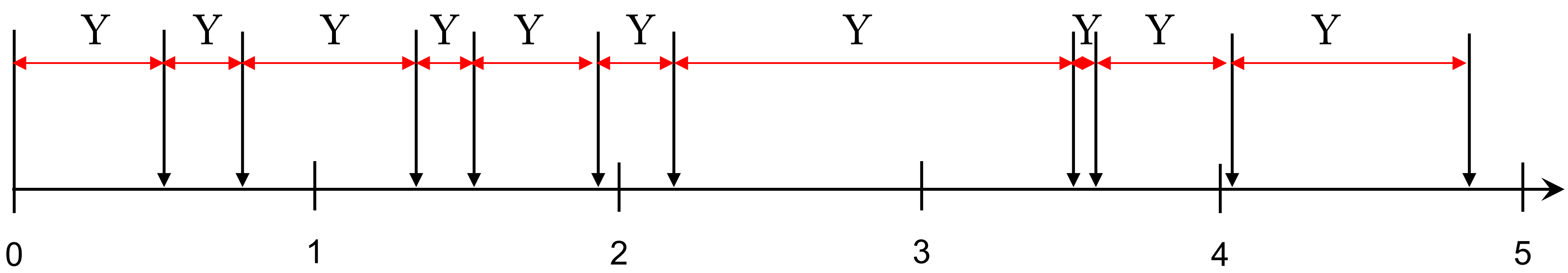


$$E(X) = \int_a^b x \cdot \frac{1}{b-a}\, dx = \frac{1}{b-a} \cdot \frac{x^2}{2}\Big|_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{b+a}{2}$$

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

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

Review: Inter-Arrival Times of a Poisson Process



What is the distribution of Y? Since

$$\lambda = E(N[0..1])$$

and the number of arrivals in an interval is proportional to its length, that is,
E(N[0..2]) = 2 * E(N[0..1]), etc., then $\lambda \cdot t = E(N[0..t])$

and so the probability that there are exactly n arrivals by time t is given by the Poisson:

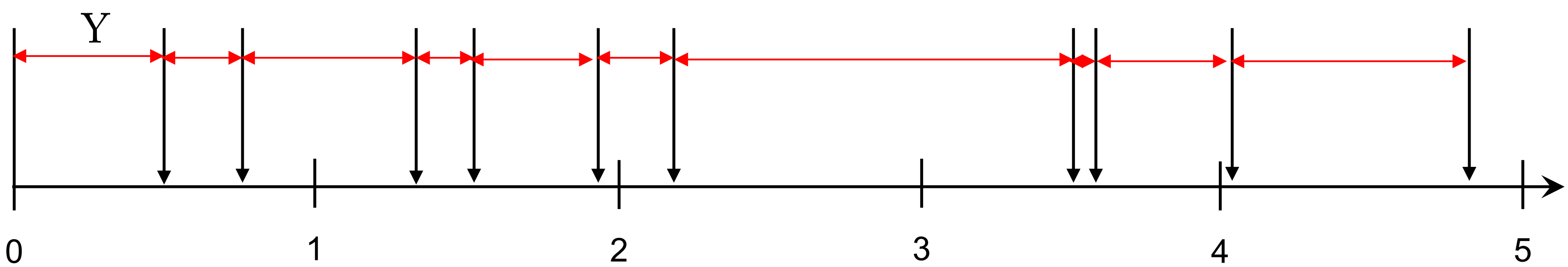
$$P(N[0..t] = n) = \frac{e^{-\lambda t}(\lambda t)^n}{n!}$$

and

$$P(Y > t) = P(N[0..t] = 0) = \frac{e^{-\lambda t}(\lambda t)^0}{0!} = e^{-\lambda t}$$

$$P(Y \leq t) = 1 - e^{-\lambda t}$$

Distribution of Interarrival Times of a Poisson Process



What is the distribution of Y?

$$P(Y \le t) = 1 - e^{-\lambda t}$$

Now, this is **the formula of a CDF**, that is,

$$F(t) = \begin{cases} 1 - e^{-\lambda t} & \text{if } t \ge 0 \\ 0 & \text{if } t < 0 \end{cases}$$

and so if we take the derivative $0 - (-\lambda)e^{-\lambda t} = \lambda e^{-\lambda t}$ we get the PDF:

$$f(t) = F'(t) = \begin{cases} \lambda e^{-\lambda t} & \text{if } t \ge 0 \\ 0 & \text{if } t < 0 \end{cases}$$

Recall the derivative of exponential:

$$\frac{d\,e^{cx}}{dx} = ce^{cx}$$

and the chain rule:

$$h(x) = f(g(x))$$

$$h'(x) = f'(g(x)) \cdot g'(x)$$

Exponential Distribution

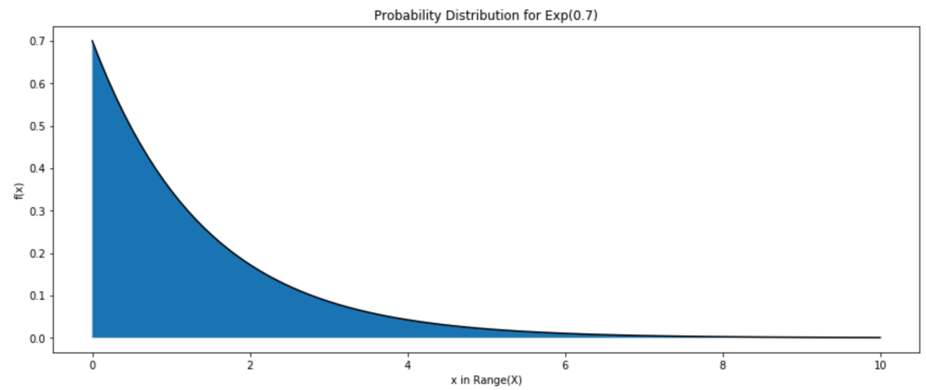
This is called the Exponential Distribution, and along with the Normal, is one of the most important continuous distributions in probability and statistics.

Formally, then, if the random variable

Y = "the interarrival time between events in a Poisson Process"

we say that Y is distributed according to the Exponential Distribution with rate parameter λ, denoted

$$Y \sim Exp(\lambda)$$



if

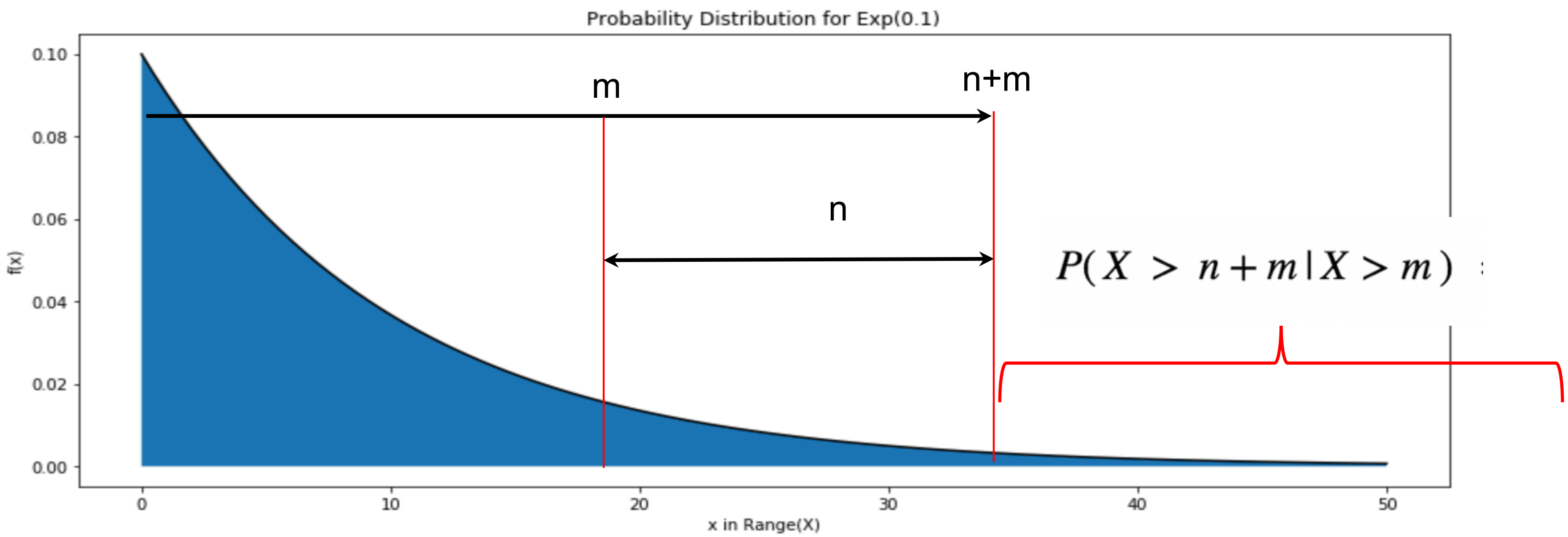
$$f(t) = \begin{cases} \lambda e^{-\lambda t} & \text{if } t \geq 0 \\ 0 & \text{if } t < 0 \end{cases}$$

$$F(t) = \begin{cases} 1 - e^{-\lambda t} & \text{if } t \geq 0 \\ 0 & \text{if } t < 0 \end{cases}$$

and where $E(X) = \frac{1}{\lambda}$ and $Var(X) = \frac{1}{\lambda^2}$

$$P(X > t) = e^{-\lambda t}$$
$$P(X \leq t) = 1.0 - e^{-\lambda t}$$

Exponential Distribution: The Memoryless Property

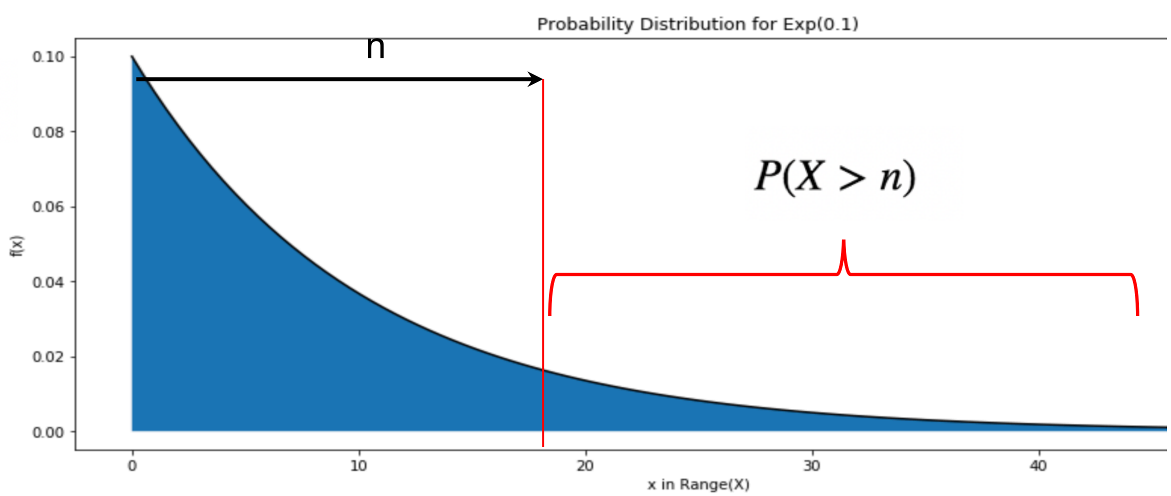


The exponential, like the geometric, has the memoryless property,

$$P(X > n+m \mid X > m) = P(X > n)$$

and the proof is the same!

$$\begin{aligned} P(X > n+m \mid X > m) &= \frac{P(X > n+m \text{ and } X > m)}{P(X > m)} \\ &= \frac{P(X > n+m)}{P(X > m)} \\ &= \frac{(1-p)^{(n+m)}}{(1-p)^m} \\ &= (1-p)^n \\ &= P(X > n) \end{aligned}$$



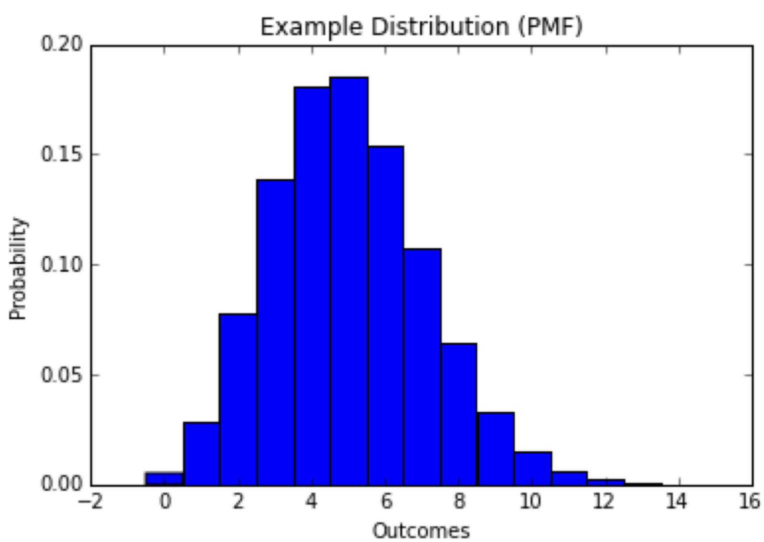
Optional: The Waiting Time Paradox:

- https://jakevdp.github.io/blog/2018/09/13/waiting-time-paradox/

Normal Distribution

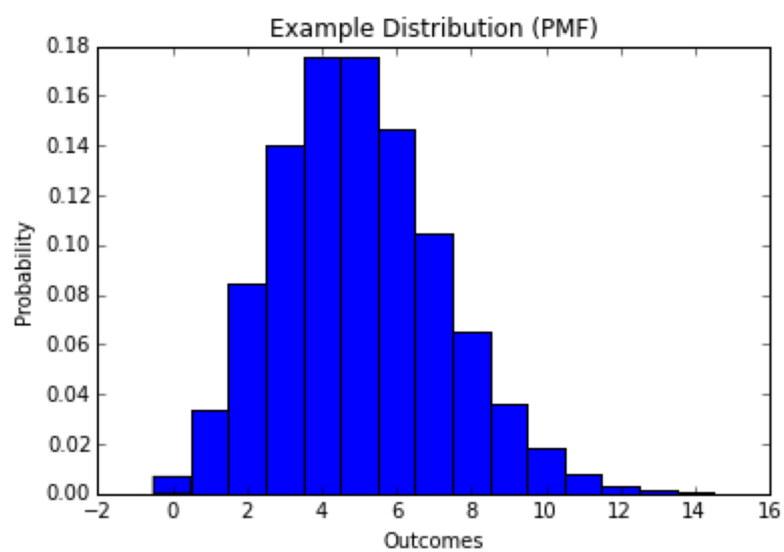
How to approximate the binomial? Under certain conditions (basically when p is small), the Poisson can be used to approximate the binomial (we will explore this in the next homework).

```
In [908]: drawDistribution(B[0][:15],B[1][:15])
```




```
E[X]     = 4.9989
Var(X)   = 4.4921
stdev(X) = 2.1195
Rng(X): [0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14]
P:      [0.0052, 0.0286, 0.0779, 0.1386, 0.1809, 0.1849, 0.1541, 0
0.0061, 0.0022, 0.0007, 0.0002]
```

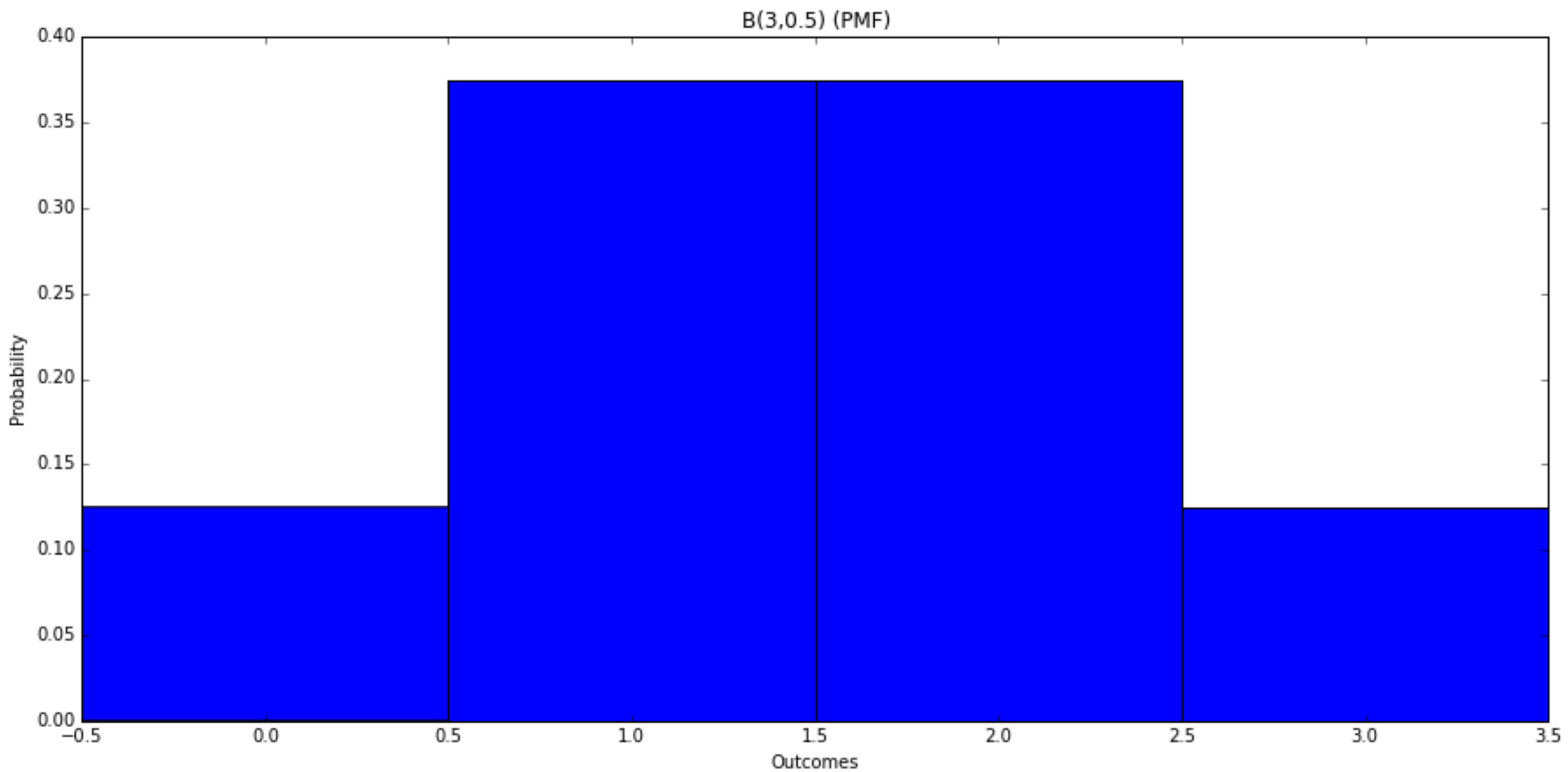
```
In [909]: drawDistribution(P[0][:15],P[1][:15])
```




```
E[X]     = 4.9965
Var(X)   = 4.9753
stdev(X) = 2.2305
Rng(X): [0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14]
P:      [0.0067, 0.0337, 0.0842, 0.1404, 0.1755, 0.1755, 0.1462, 0.1044,
0.0082, 0.0034, 0.0013, 0.0005]
```

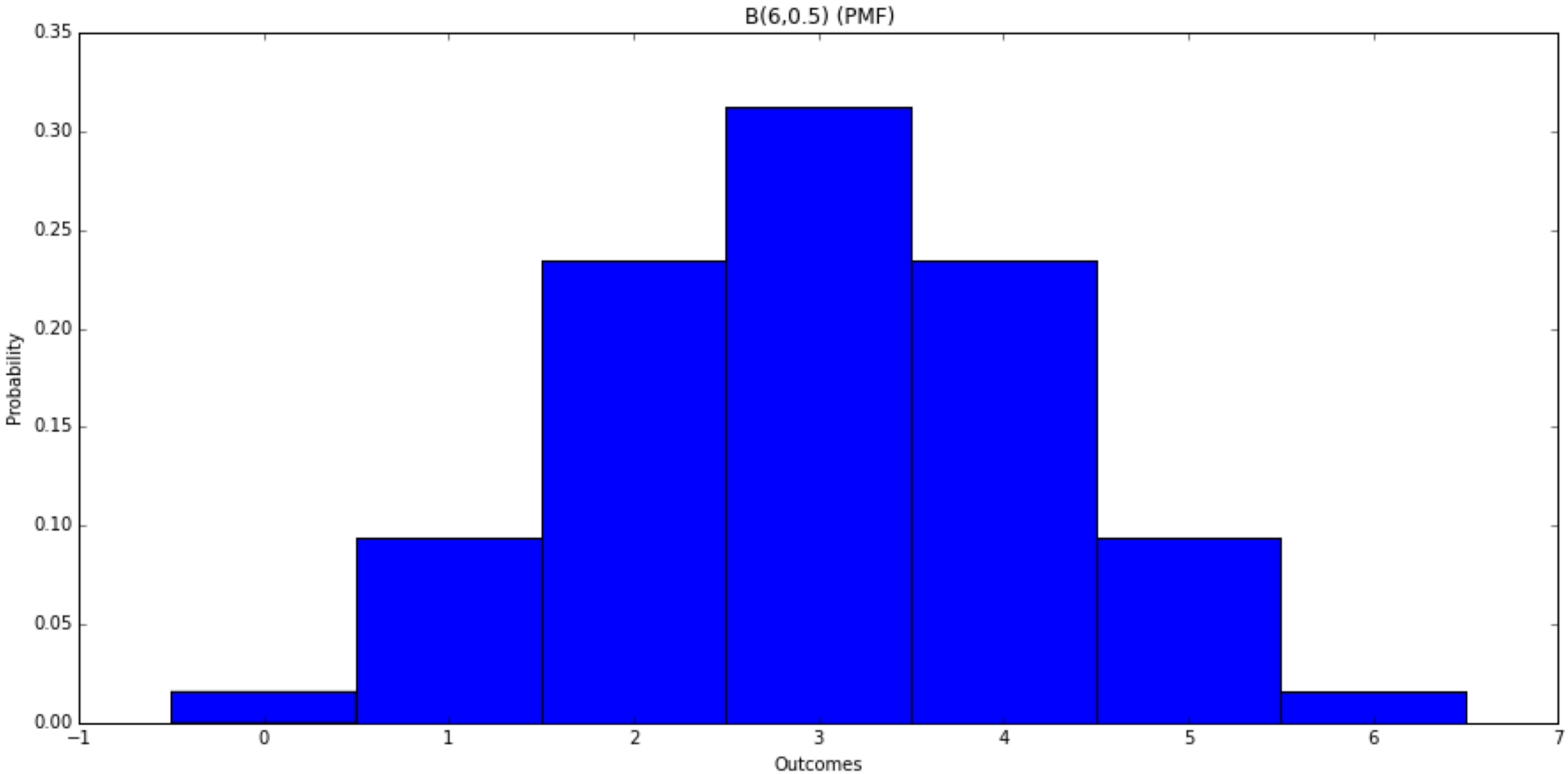
Normal Distribution

How to approximate the binomial when p is 1/2? When we observe the characteristic shape of the Binomial Distribution B(N,0.5) as N approaches Infinity, we see something interesting:



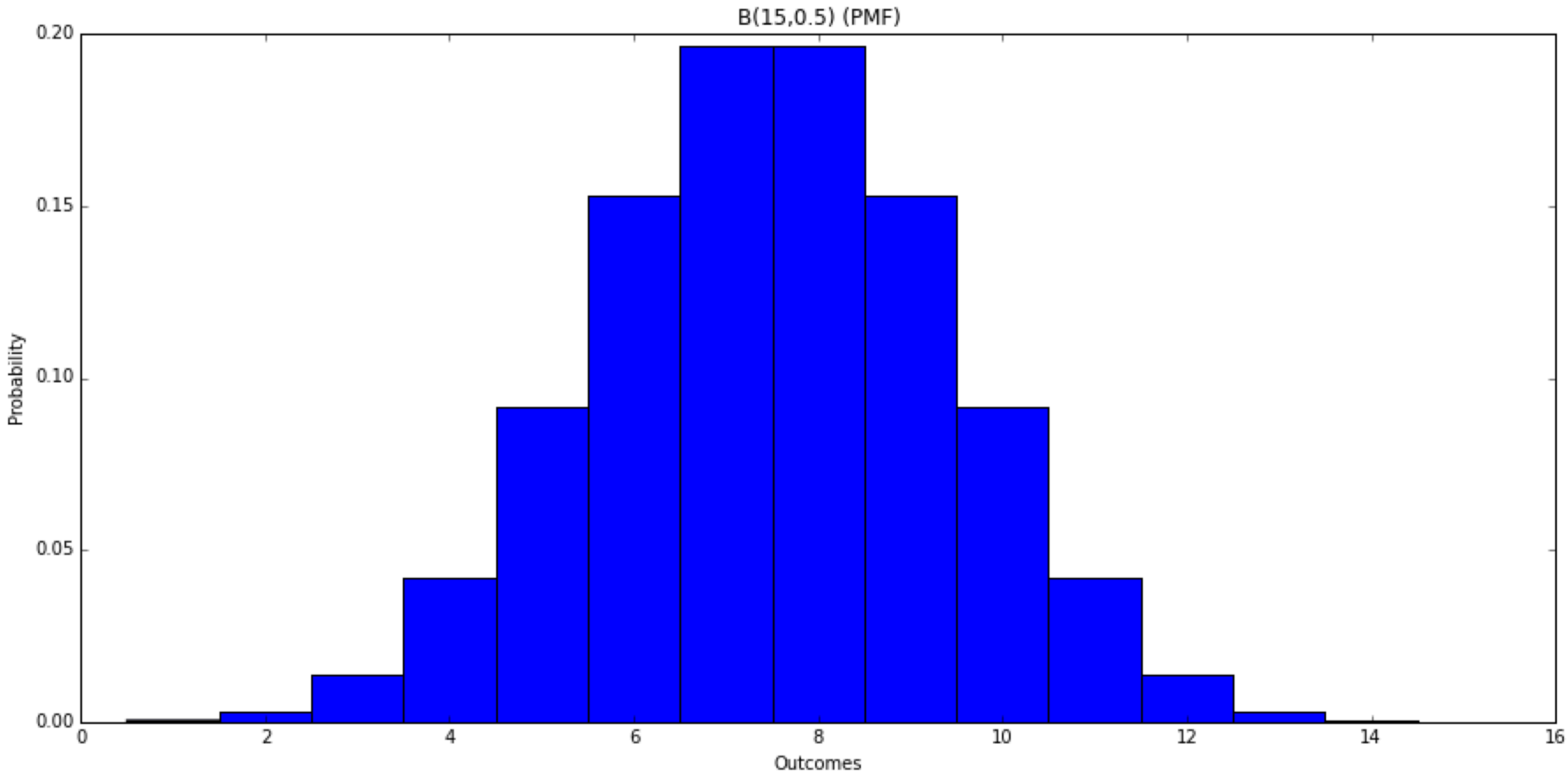
Normal Distribution

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:



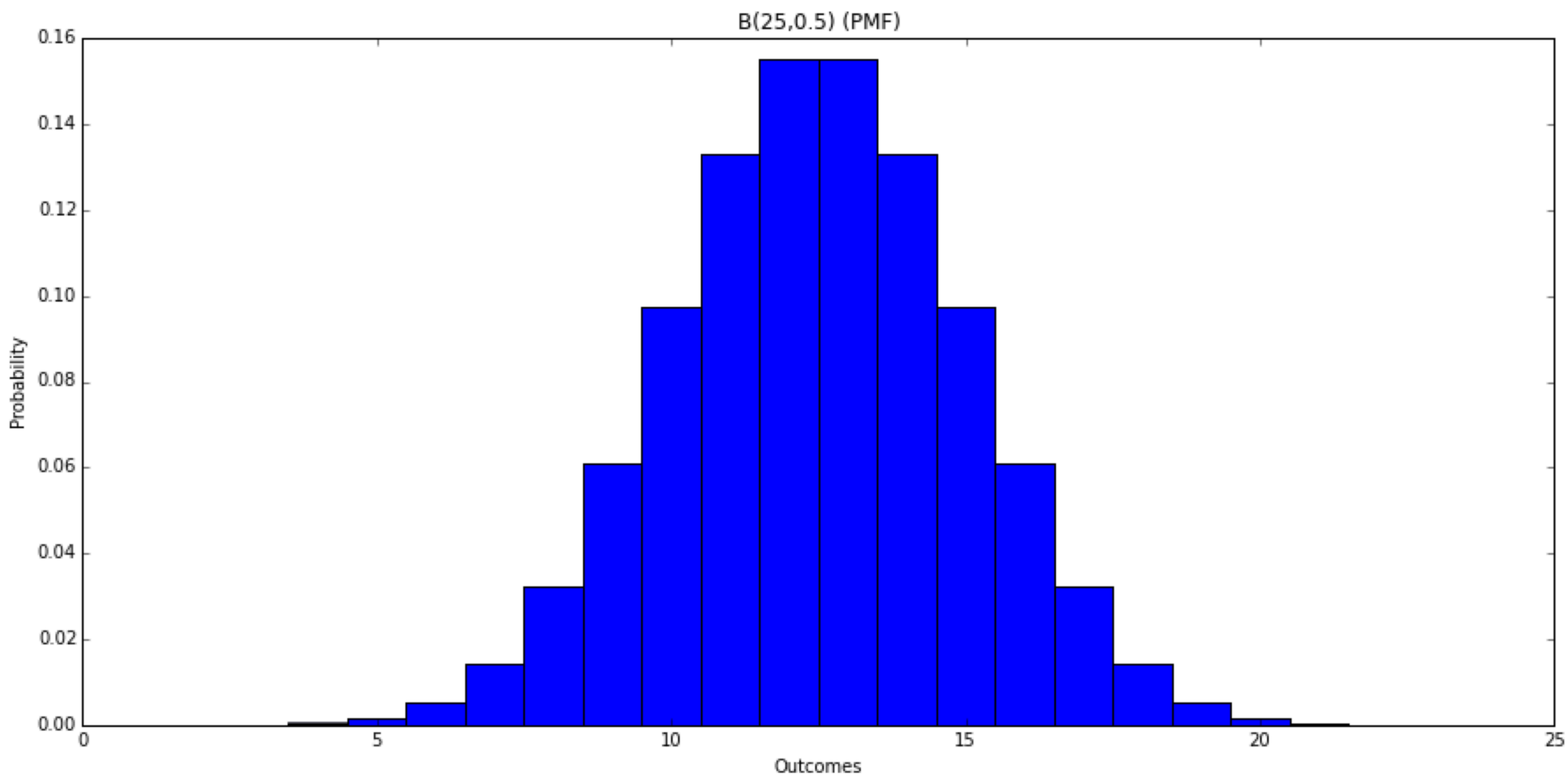
Normal Distribution

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:



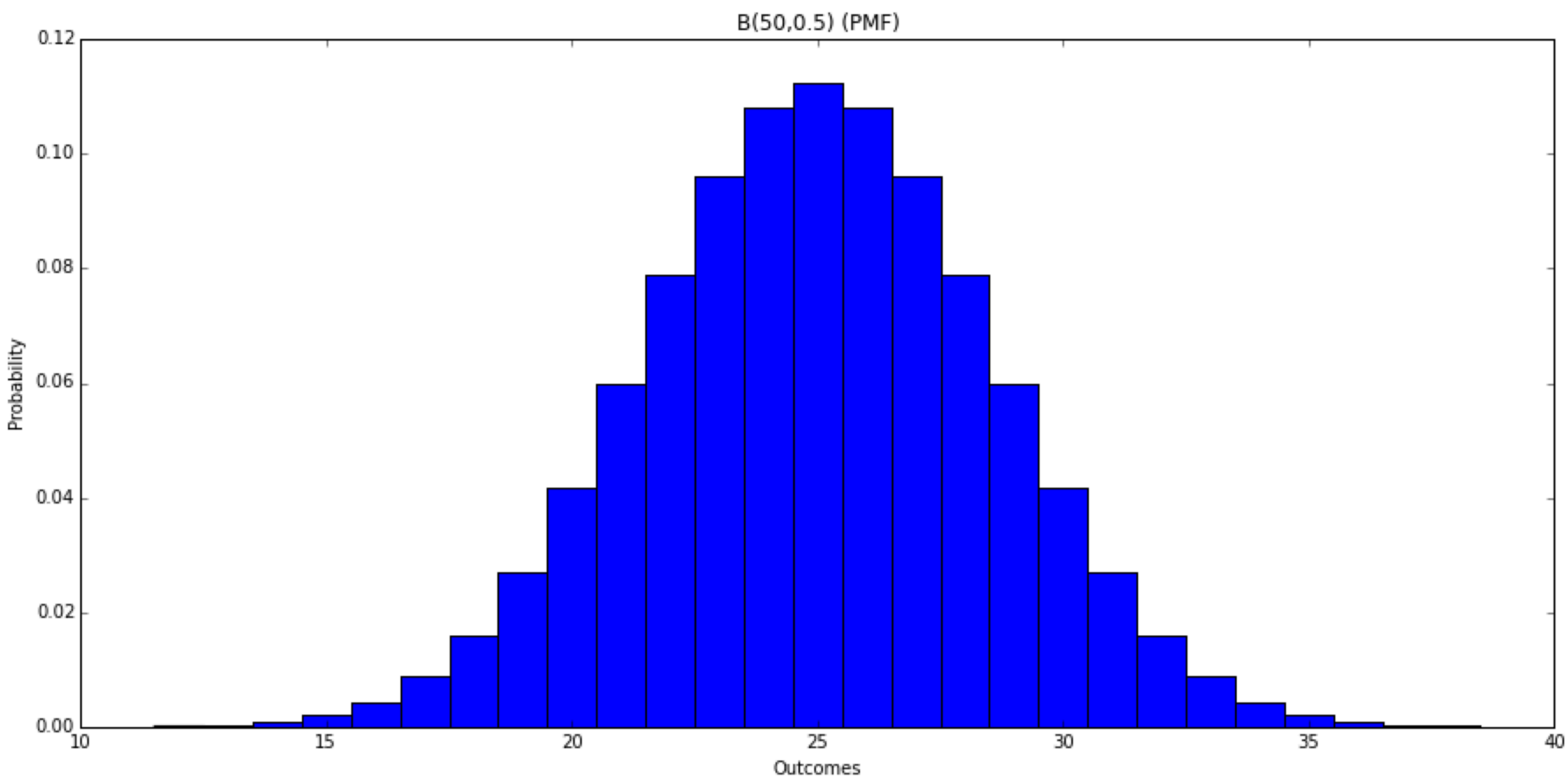
Normal Distribution

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:



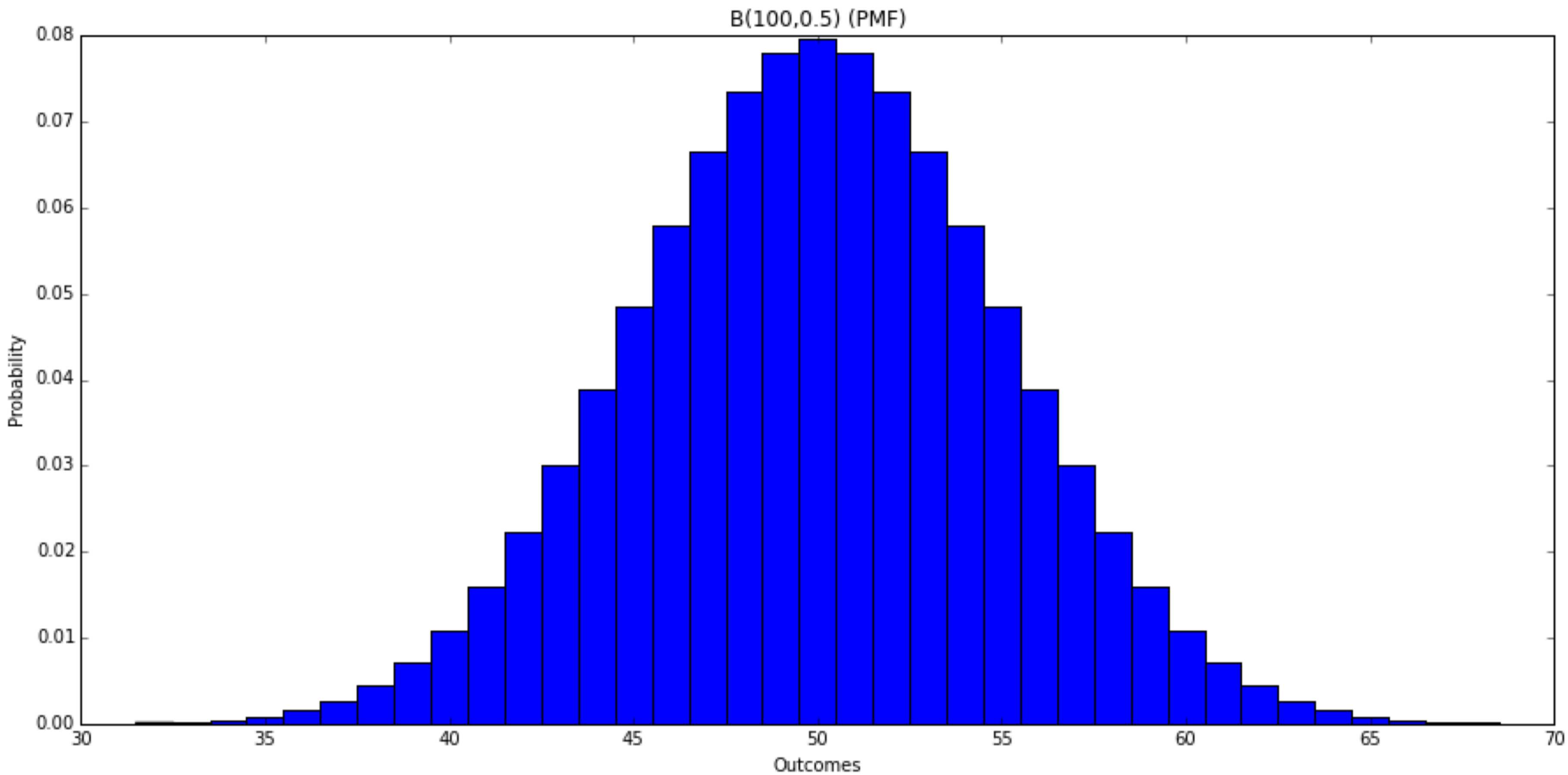
Normal Distribution

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:



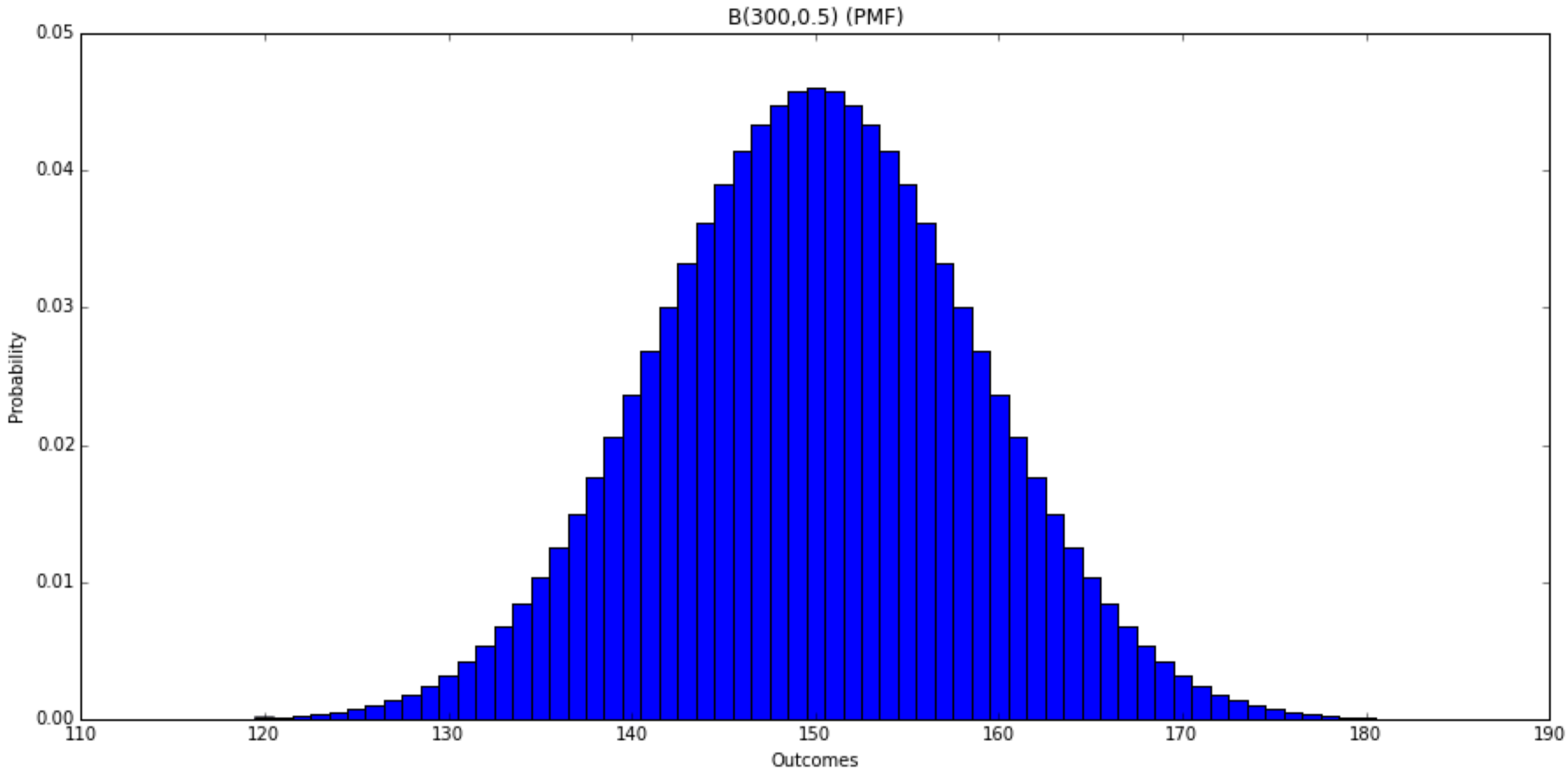
Normal Distribution

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:



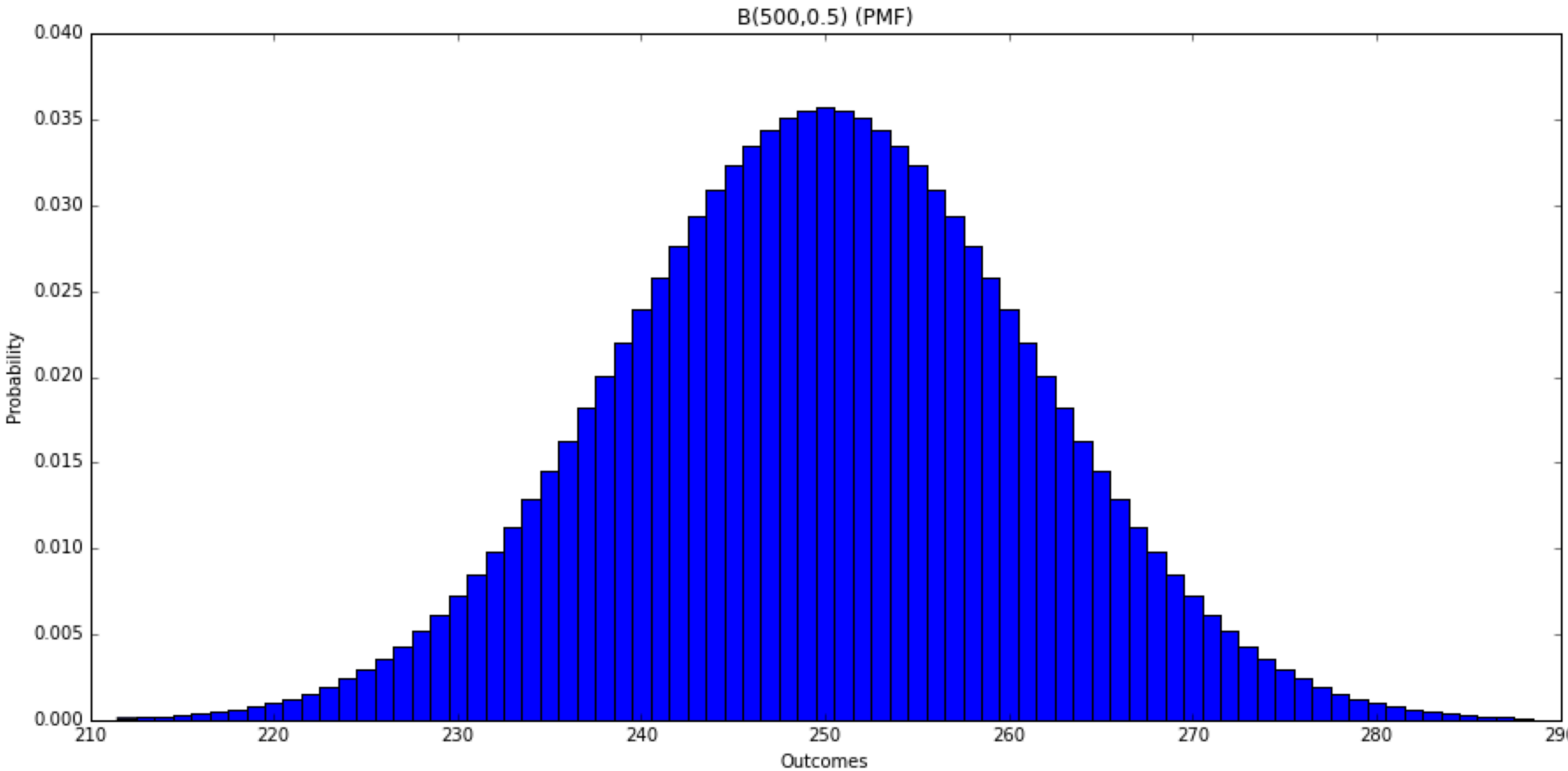
Normal Distribution

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:



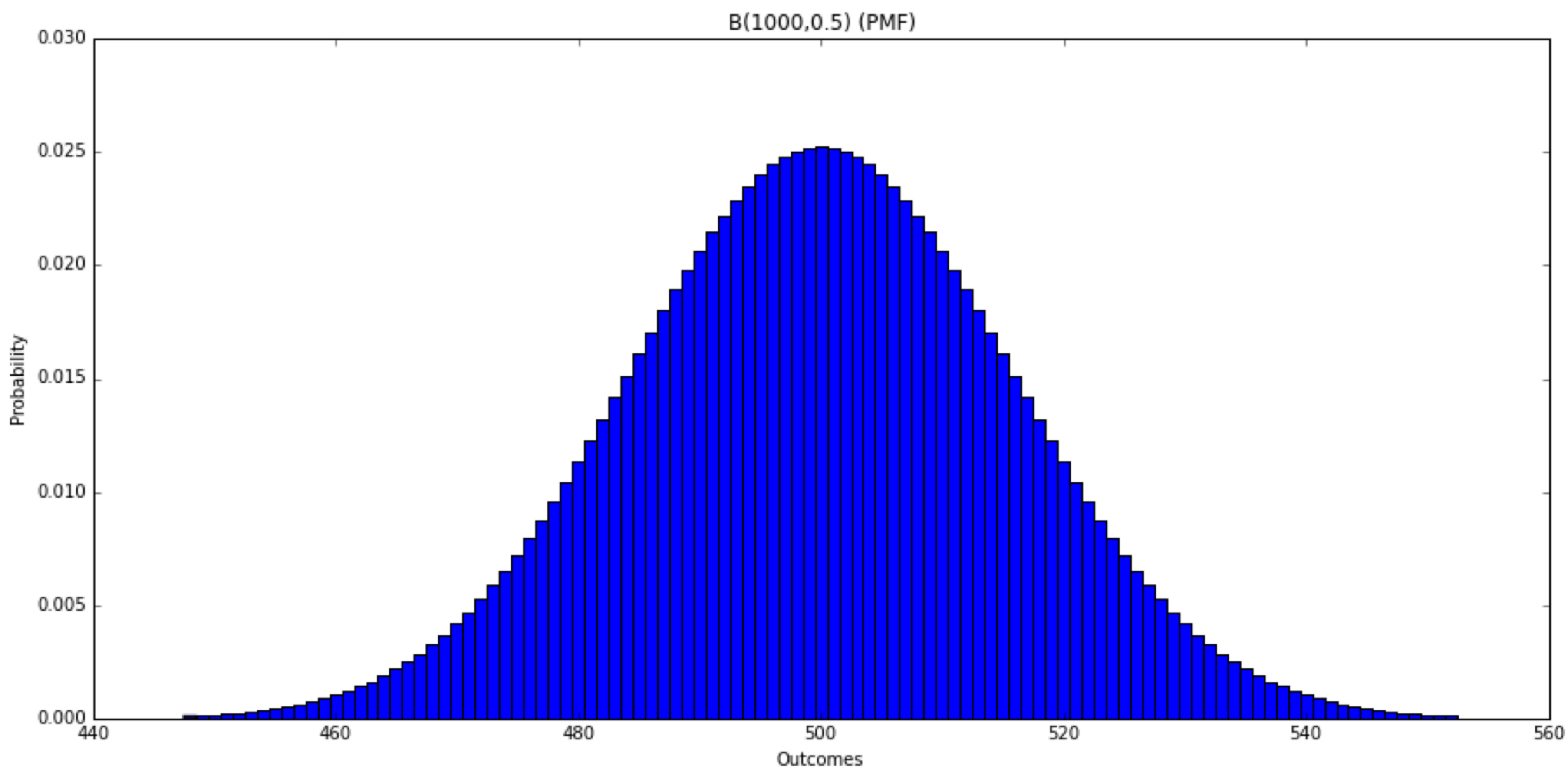
Normal Distribution

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:



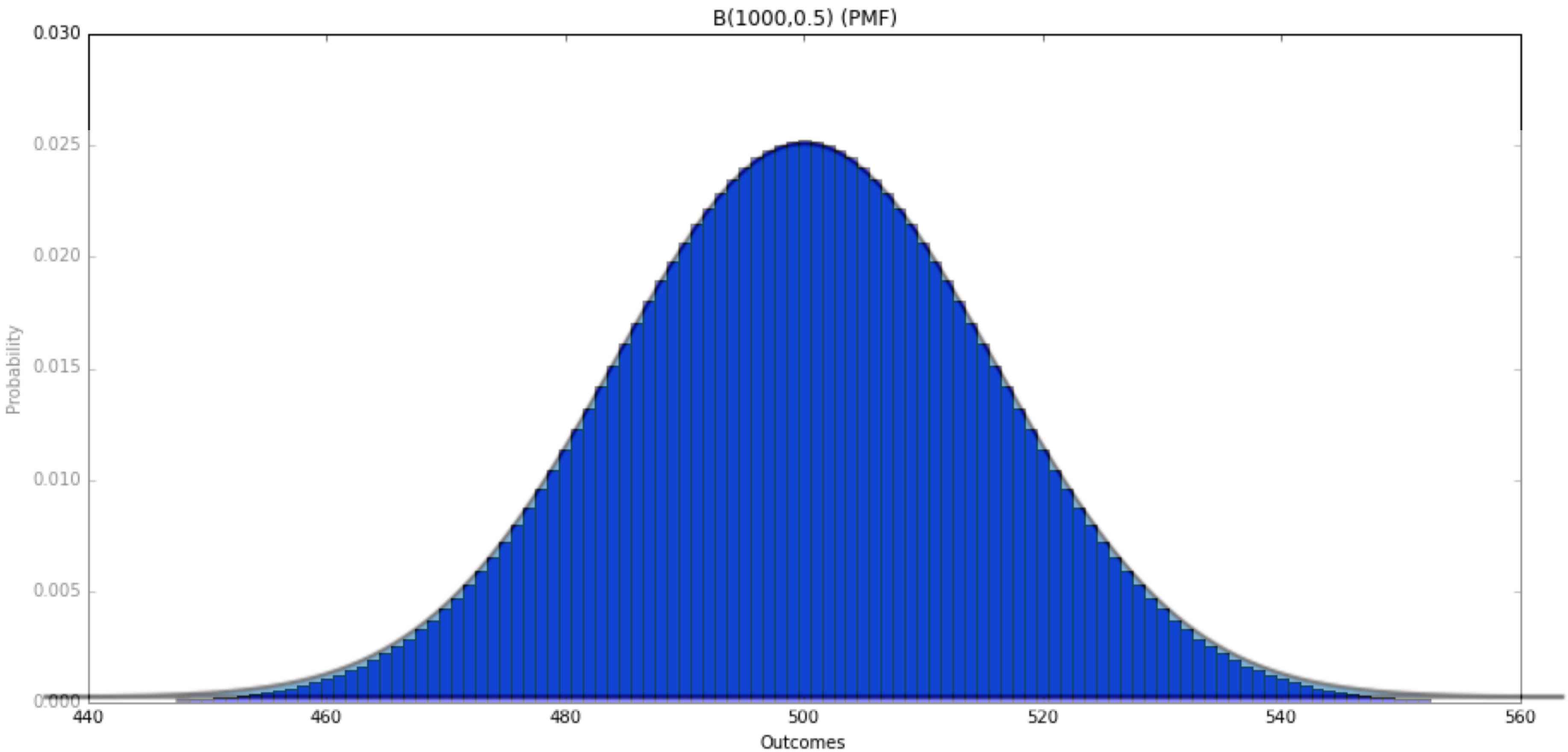
Normal Distribution

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

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

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: