Central Limit Theorem

ECE 341: Random Processes

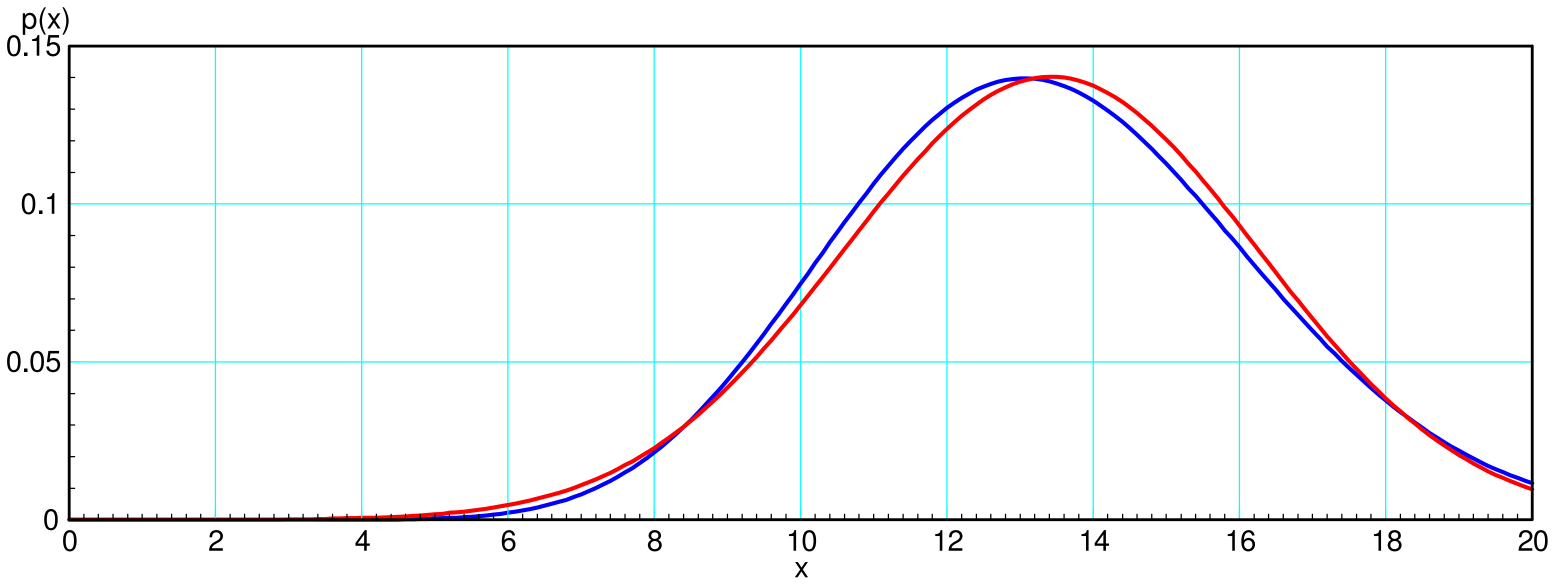
Lecture #19

note: All lecture notes, homework sets, and solutions are posted on www.BisonAcademy.com

Central Limit Theorem

- One of the most important theorems in statistics
- It basically says that all distributions coverage to a normal distribution.

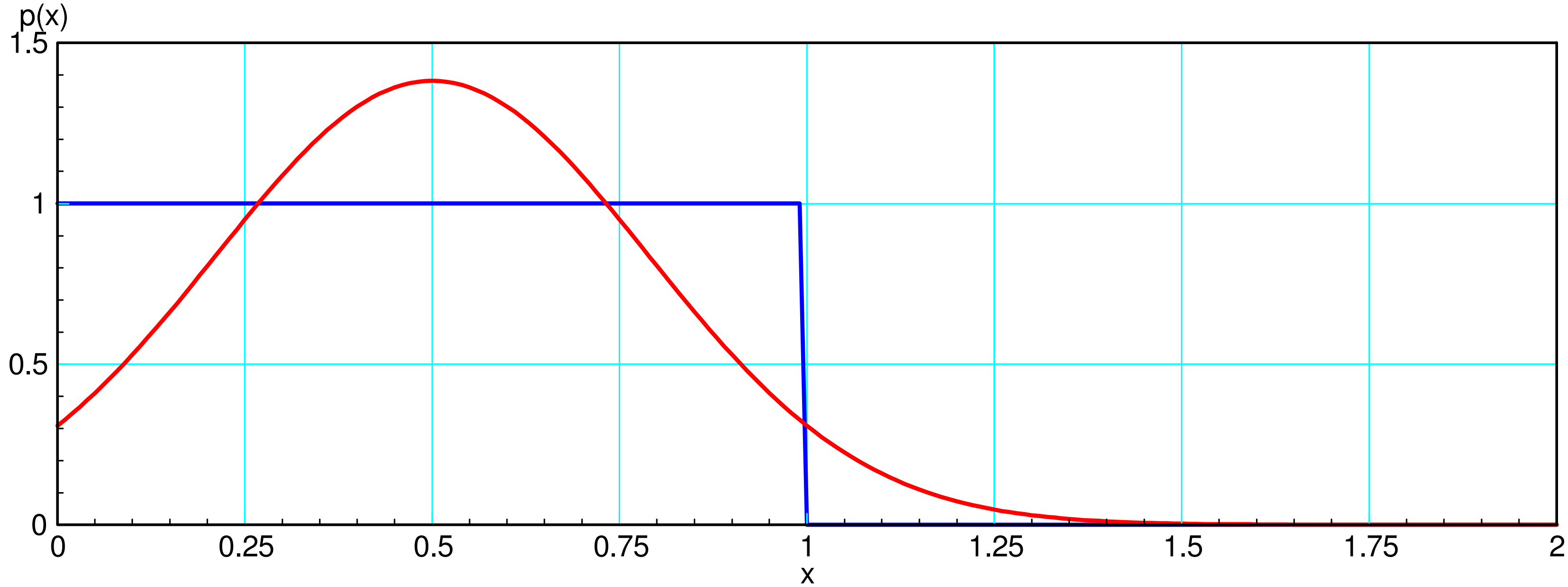
This is one of the reasons engineers tend to assume everything is described by a normal distribution (even when a Poisson distribution is more accurate). It also allows you to determine the probability for some fairly complex problems fairly accurately.



Example: Uniform Distribution.

Let A be a uniform distribution over the range of (0, 1).

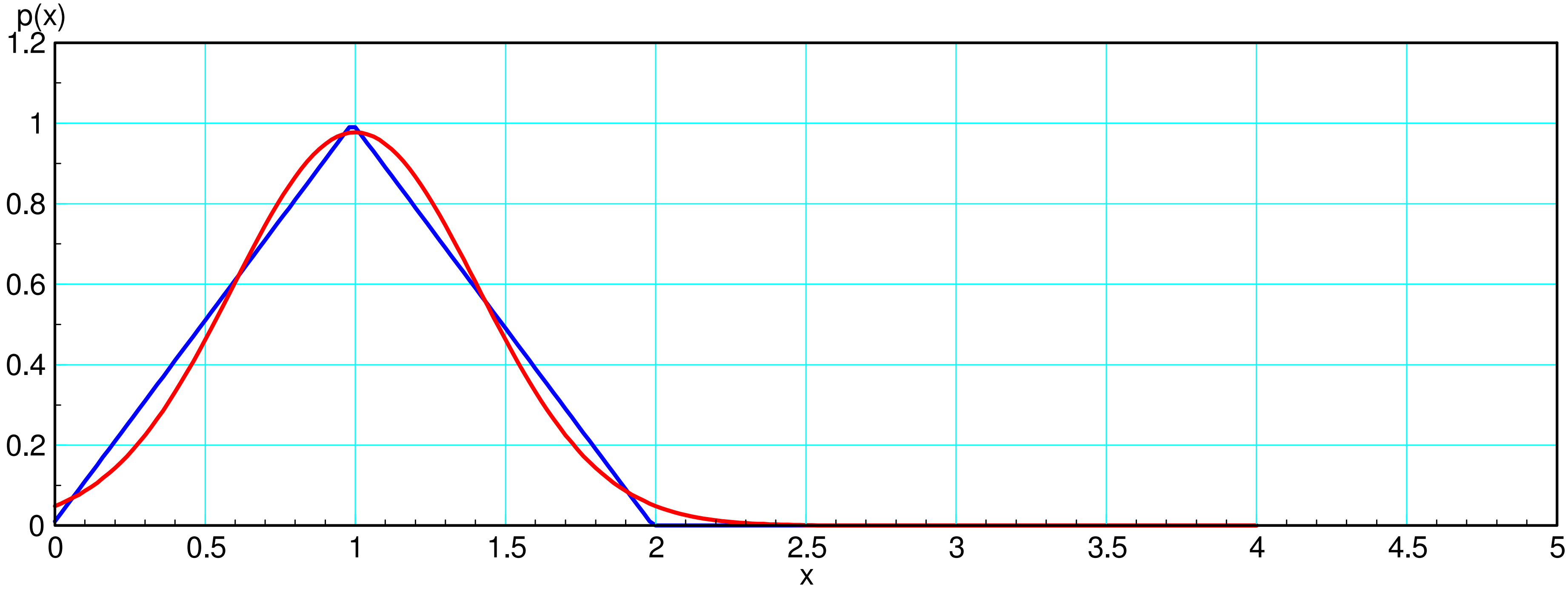
- Plot the pdf of A (blue)
- Plot the pdf of a Normal distribution with the same mean and variance (red)
- The two are different.



Uniform Distribution (blue) and it's normal approximation (red)

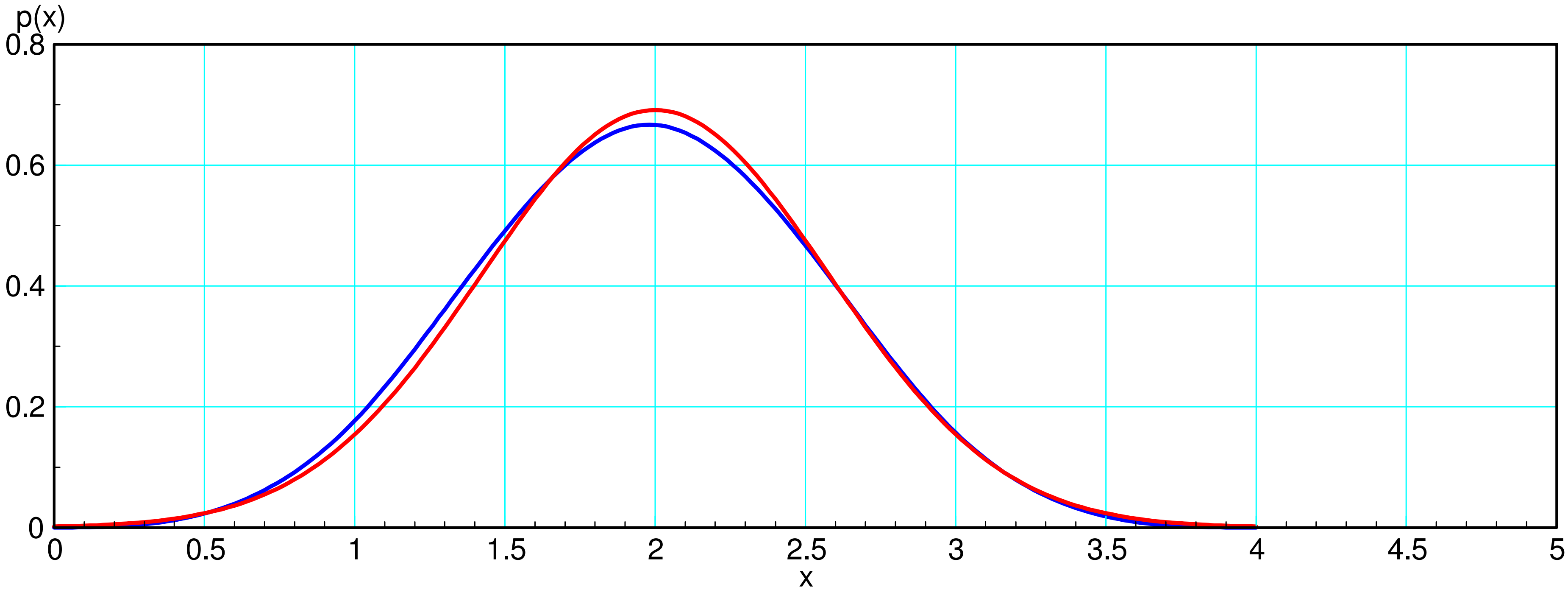
Summing two uniform distributions

- Blue = pdf of summing two uniform distributions
- Red = Normal distribution with the same mean and variance
- Getting closer after only two summations



Summing four uniform distributions

- Blue = pdf of summing two uniform distributions
- Red = Normal distribution with the same mean and variance
- Very close after only four summations



Example 2: Exponential pdf

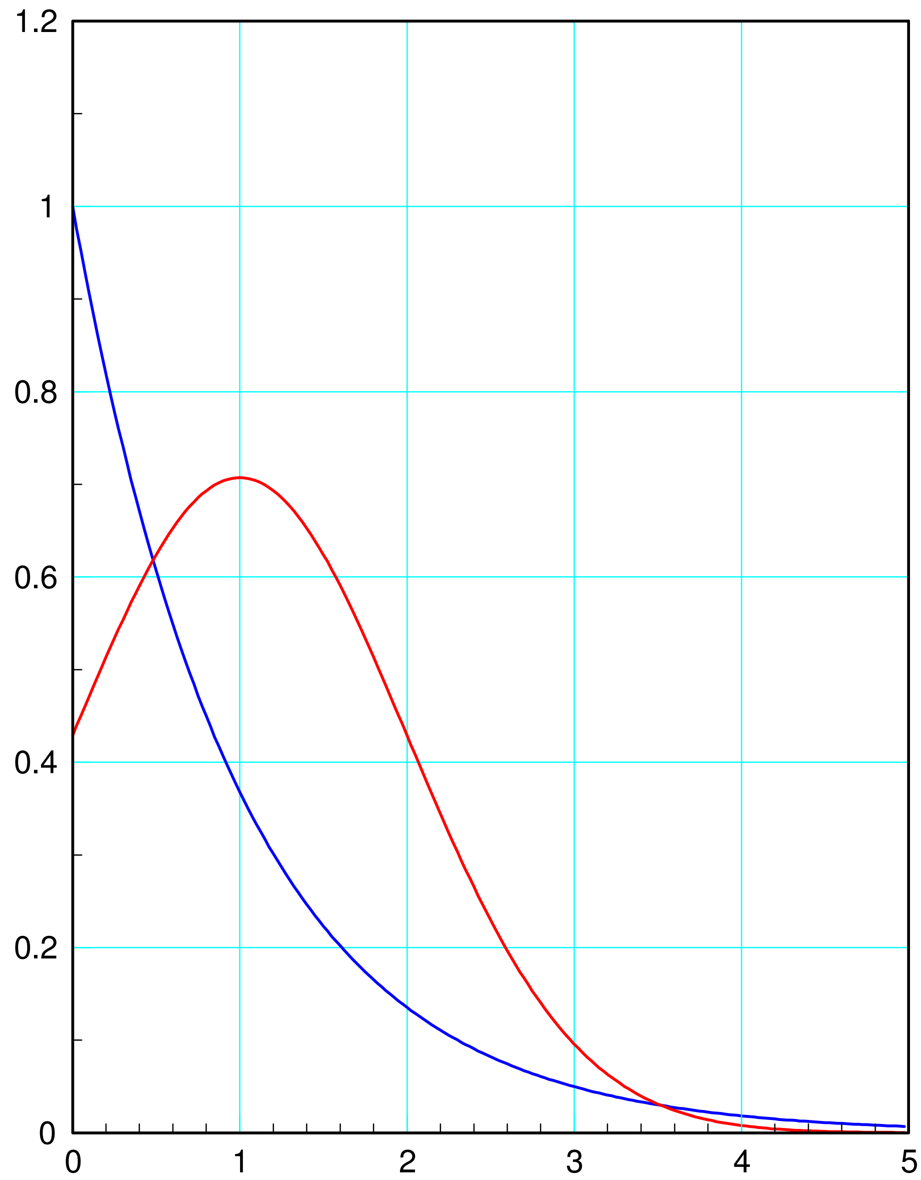
Exponential Distribution (blue)

- Mean = 1
- Variance = 1

Normal Distribution (red)

- Mean = 1
- Variance = 1

Not very close



Exponental pdf (cont'd)

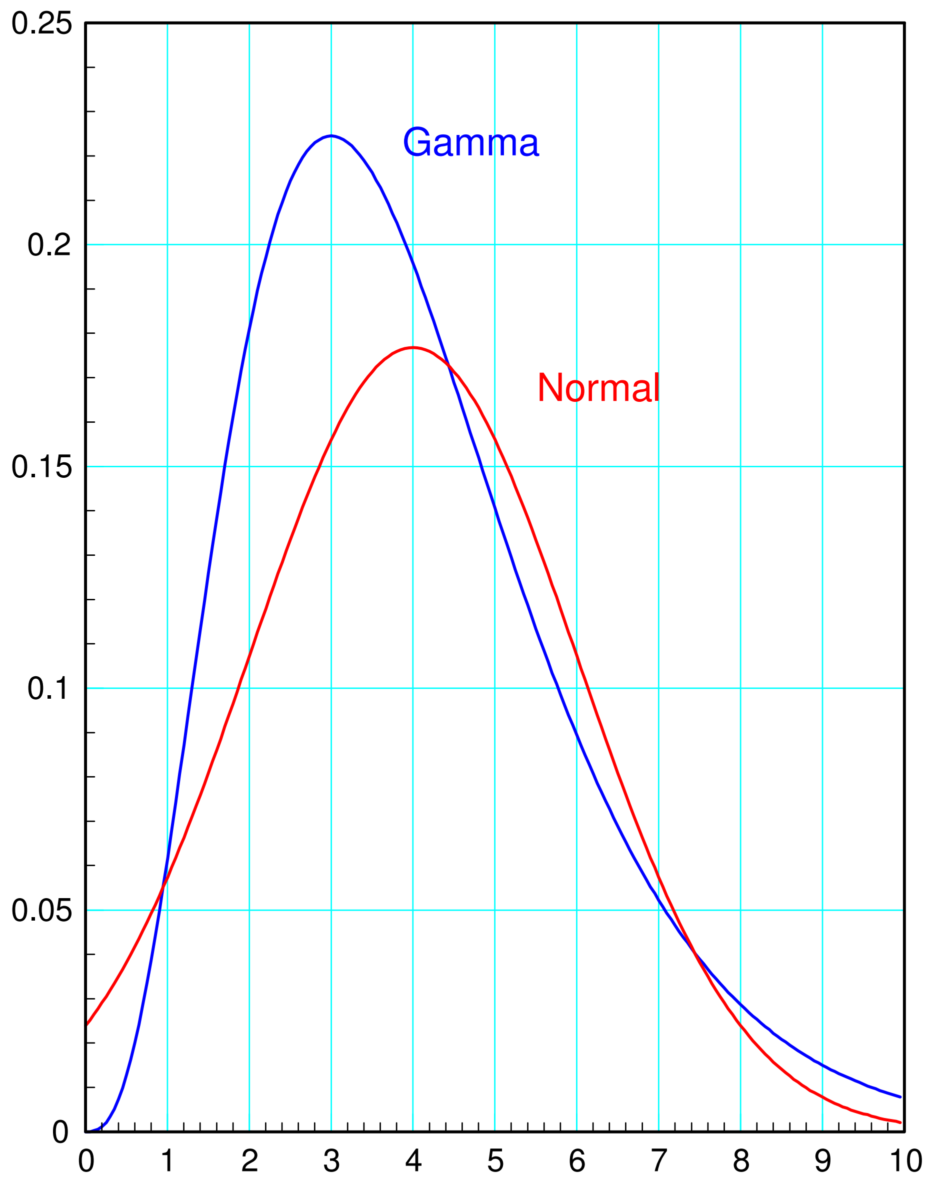
Gamma Distributions (blue)

- Sum of 4 exponential pdf's
- Mean = 4
- Variance = 4

Normal Distribution (red)

- Mean = 4
- Variance = 4

Getting closer, but still different



Exponental pdf (cont'd)

Gamma pdf (blue)

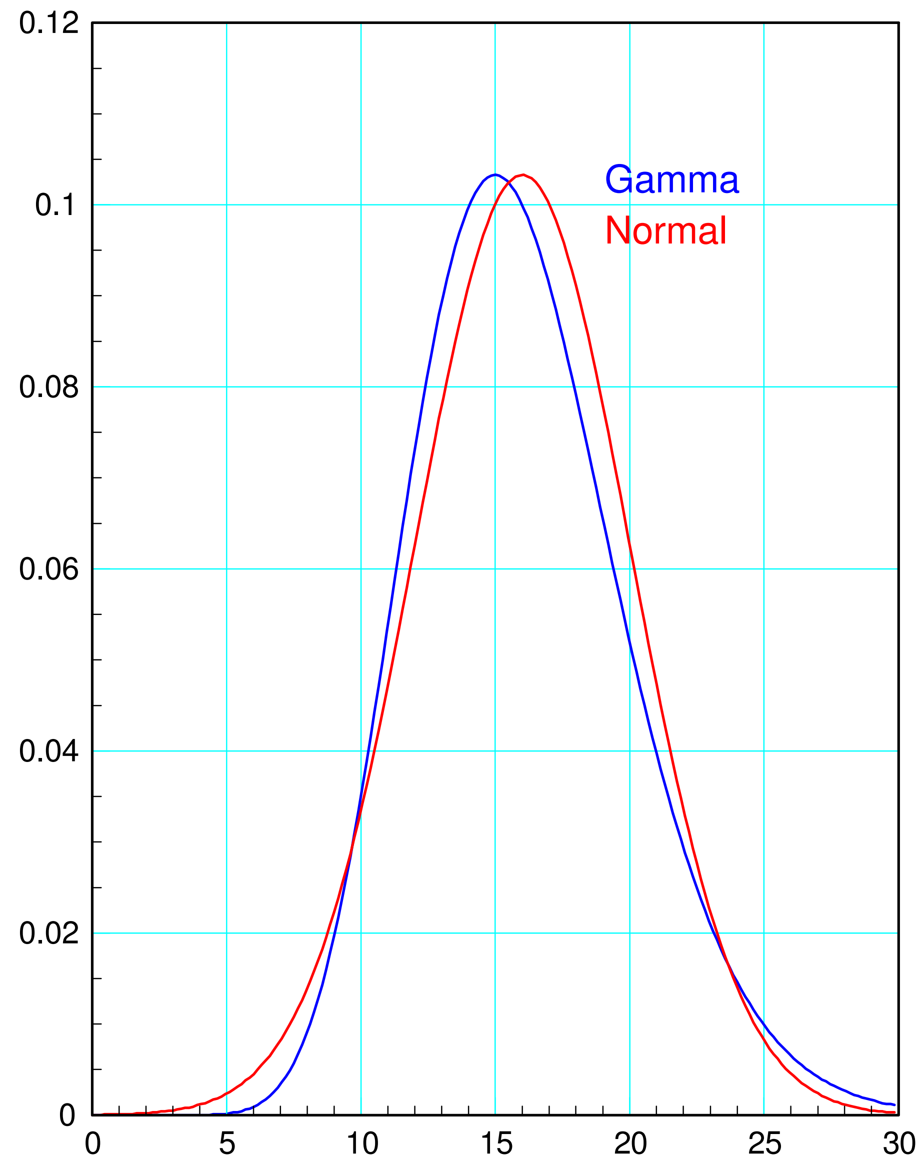
- Sum 16 exponential distributions
- Mean = 16
- Variance = 16

Normal Distribution (red)

- Mean = 16
- Variance = 16

Gamma is approaching a normal pdf

- Central Limit Theorem



Central Limit Theorem:

The sum of any distribution converges to a normal distribution.

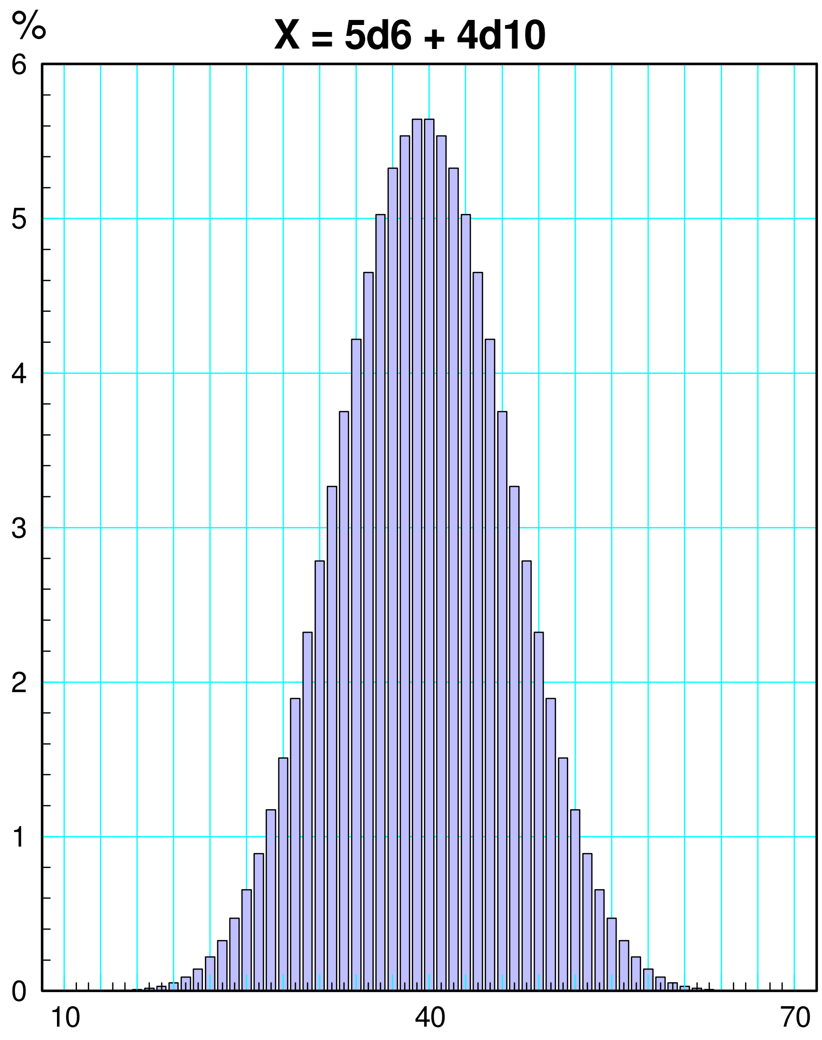
Example 1:   Central Limit Theorem with Dice

Let Y = 5d6 + 4d10

- Roll five 6-sided dice (5d6)
- Roll four 10-sided dice (4d10)
- Take the sum

What is the probability of

- rolling 45?
- rolling 45 or higher?
- rolling 50 or higher?



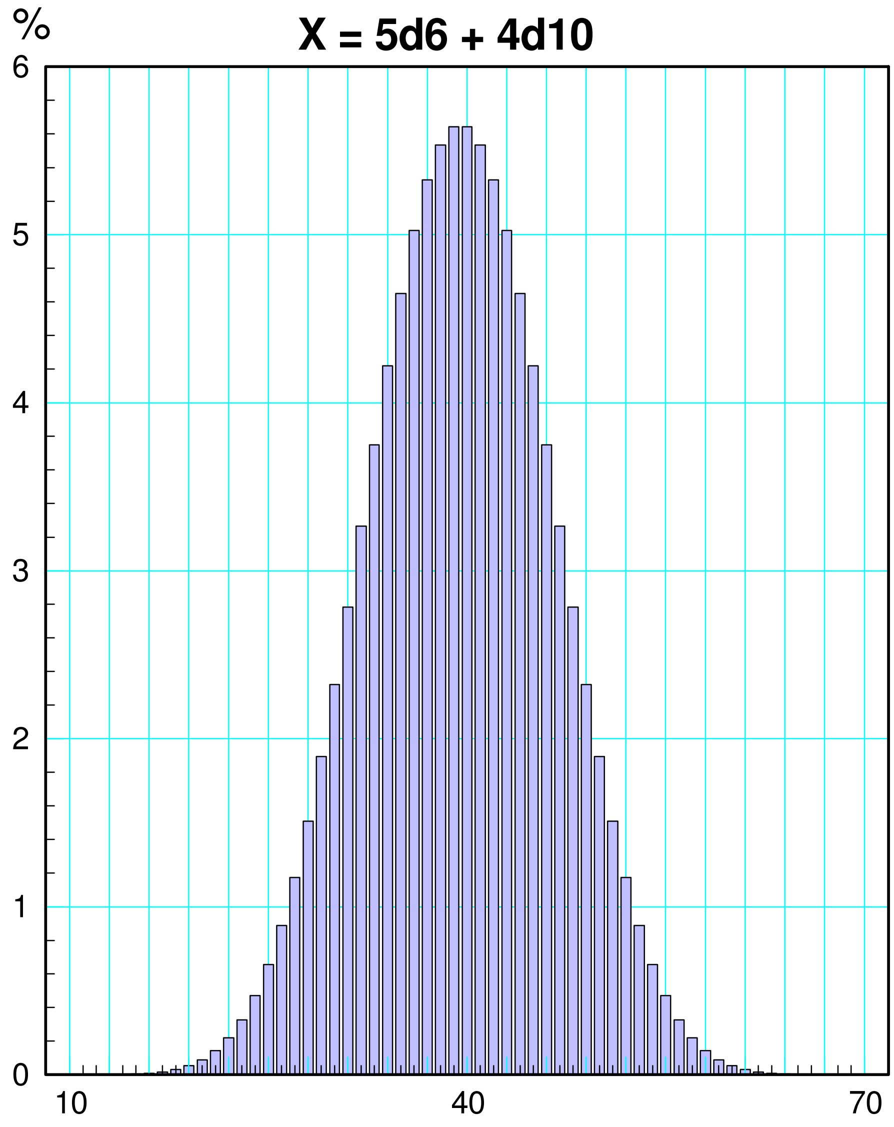
Exact Solution:

Use convolution

```
d6 = [0,ones(1,6)];
d10 = [0, ones(1,10)];
d6x2 = conv(d6, d6);
d6x4 = conv(d6x2,d6x2);
d6x5 = conv(d6x4, d6);
d10x2 = conv(d10, d10);
d10x4 = conv(d10x2, d10x2);
pdf = conv(d6x5, d10x4);
pdf = pdf / sum(pdf);
```

Exact Answers:

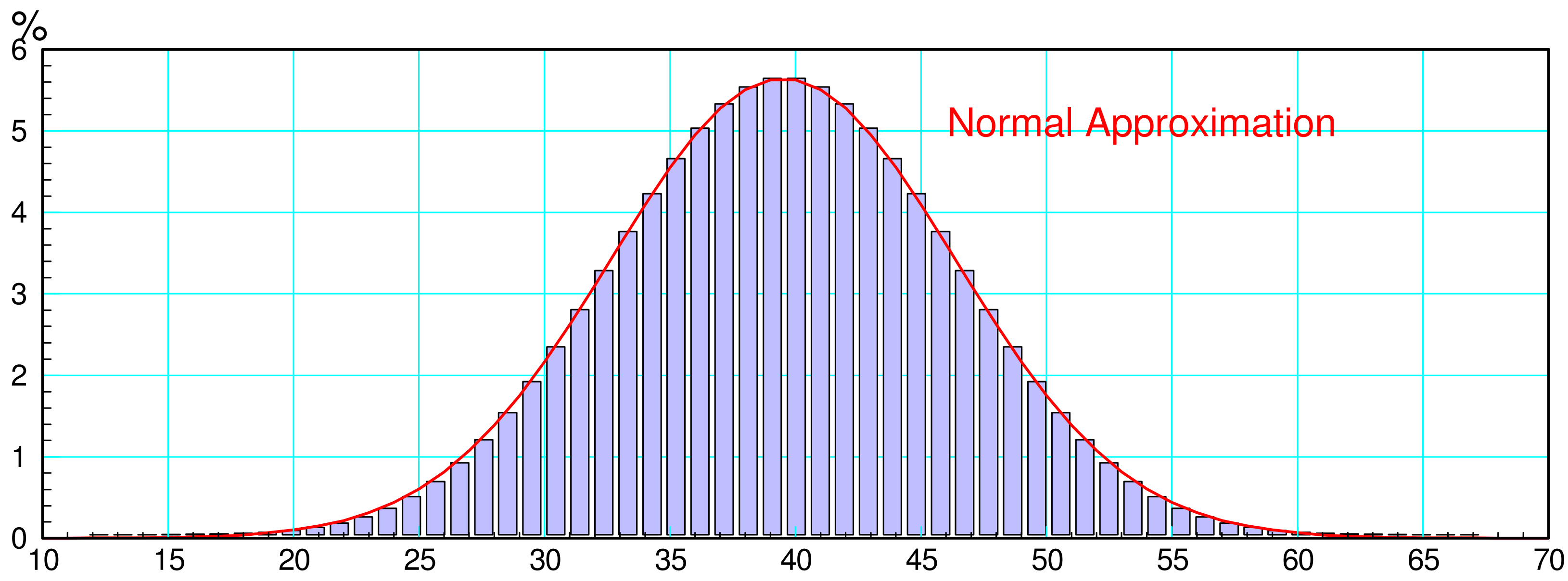
```
pdf(46)            ans =    0.0422

sum(pdf(46:71))    ans =    0.2382

sum(pdf(51:71))    ans =    0.0748
```



Solution using the Central Limit Theorem.

The mean and standard deviation of a 6 and 10 sided die are

	d6	5d6	d10	4d10	5d6+4d10
mean	3.50	17.50	5.50	22.00	39.50
variance	2.9167	14.08	8.250	33.00	47.08



What is the probability of rolling 45?

Normal Aproximation	Matlab Command Window

Interprit as: $p(44.5 < y < 45.5)$

Find the z-score for

- $y = 44.5$
- $y = 45.5$

Find the area of the tail in each case

- $p = 0.2331$
- $p = 0.1909$

Take the difference

- $p = 0.0422$

There is a 4.22% chance

- exact answer = 4.22%

```
>> x = 39.50;
>> s = sqrt(47.08);
>> z1 = (44.5 - x) / s
z1 =    0.7287

>> p1 = (1 - erf(z1/sqrt(2)))/2
p1 =    0.2331

>> z2 = (45.5 - x) / s
z2 =    0.8744

>> p2 = (1 - erf(z2/sqrt(2)))/2
p2 =    0.1909

>> p1 - p2
ans =    0.0422
```

What is the probability of rolling 45 or more?

Normal Aproximation	Matlab Command Window

Interprit as p(y > 44.5)
 	
Find the z-score for y = 44.5
• z = 0.7287
 	
Find the area of the tail
• p = 0.2331
 	
There is a 23.31% chance
• exact answer = 23.82%
	`>> x = 39.50;`
	`>> s = sqrt(47.08);`
	`>> z = (44.5 - x) / s`
	`z =    0.7287`
	
	`>> p = (1 - erf(z/sqrt(2)))/2`
	`p =    0.2331`

What is the probability of rolling 50 or more?

Normal Aproximation	Matlab Command Window

Interprit as p(y > 49.5)
 	
Find the z-score for y = 49.5
• z = 1.4574
 	
Find the area of the tail
• p = 0.0725
• StatTrek also works
 	
There is a 7.25% chance
• exact answer = 7.48%
	`>> x = 39.50;`
	`>> s = sqrt(47.08);`
	`>> z = (49.5 - x) / s`
	
	`z =    1.4574`
	
	`>> p = (1 - erf(z/sqrt(2)))/2`
	
	`p =    0.0725`

Example 2: Uniform Distribution.

- Let A1 .. A10 be uniform distributions over the interval (0, 1).
- Let X be the sum of A1 .. A10.

What is the probability that the sum is more than 6? More than 7?

Solution: Convolution with matlab.

```
dx = 0.01;
x = [0:dx:2]';
A = 1*(x < 1);
A2 = conv(A, A) * dx;
A4 = conv(A2, A2) * dx;
A8 = conv(A4, A4) * dx;
A10 = conv(A2, A8) * dx;

sum(A10(600:2000)) * dx
     ans =    0.1306

sum(A10(700:2000)) * dx
     ans =    0.0121
```

Solution: Monte-Carlo Simulation

```
N6 = 0;
N7 = 0;

for i=1:1e5

  X = sum(rand(1,10));
  if(X > 6)
      N6 = N6 + 1;
  end
  if(X > 7)
      N7 = N7 + 1;
  end

end

[N6, N7] / 1e5
```

p(y>6)	p(y>7)	
0.1388	**0.0137**	**monte-carlo**
0.1306	**0.0121**	**convolution**

Solution: Normal Approximation

	Uniform(0,1)	10 x Uniform
mean	1/2	10/2
variance	1/12	10/12

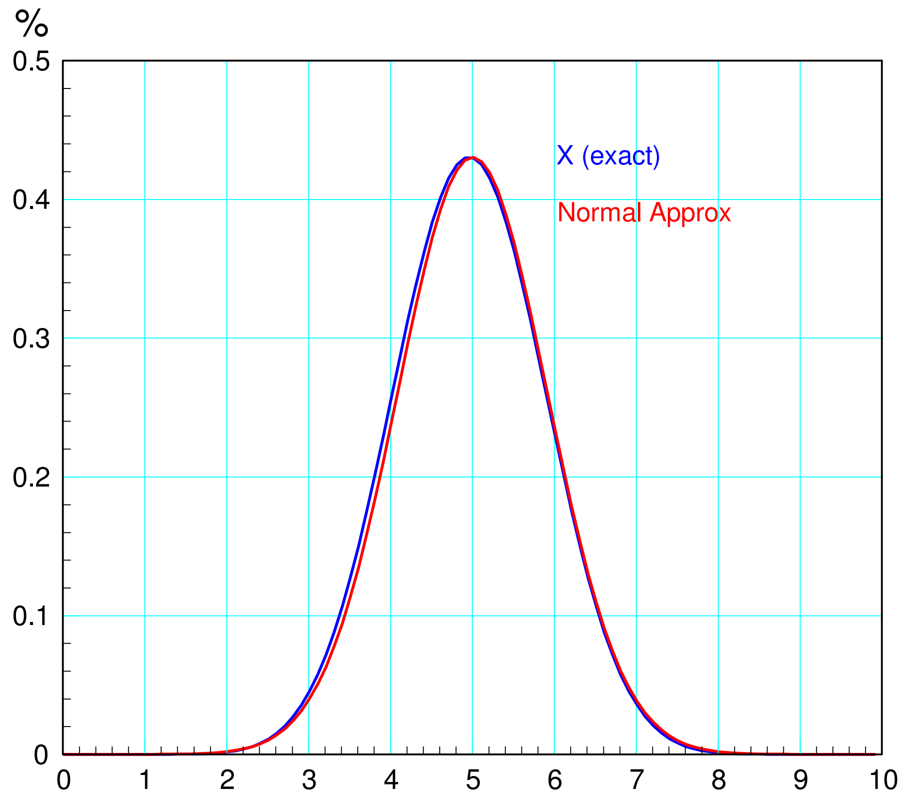
The z-score for 6.00 is

$$z = \left(\frac{6-5}{0.9129}\right) = 1.0954$$

```
>> z = (6-5)/sqrt(10/12);
>> p = (1 - erf(z/sqrt(2)))/2
p =    0.1367
```

This corresponds p = 0.1367

- Normal Approx: 0.1367
- Computed: 0.1306
- Monte Carlo: 0.1388

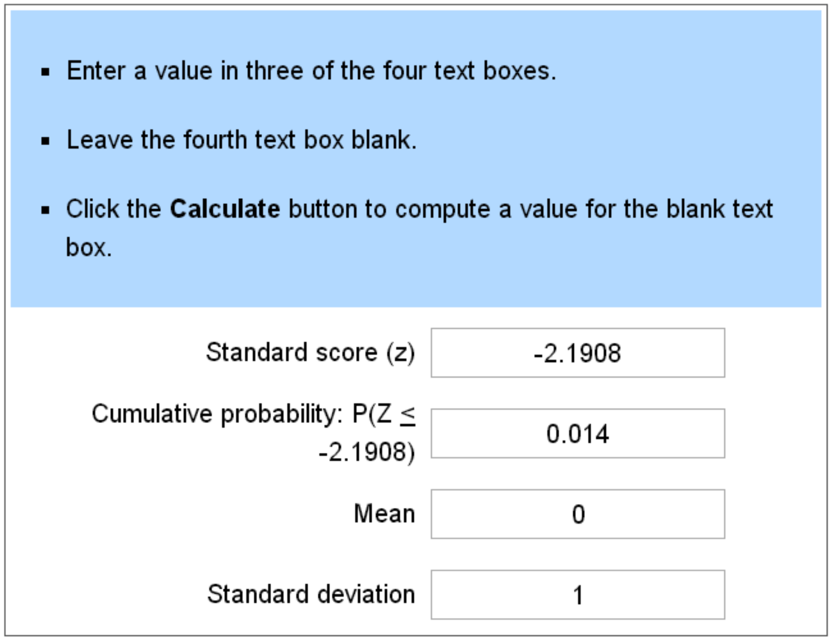


The z-score for rolling 7.00 or higher is

$$z = \left(\frac{7-5}{0.9129}\right) = 2.1908$$

This corresponds to a probability of 0.0142

- Normal Approx: 0.0142
- Computed: 0.0121
- Monte Carlo: 0.0137



In Matlab

```
>> z = (7-5)/sqrt(10/12);
>> p = (1 - erf(z/sqrt(2)))/2

p =    0.0142
```

Example 3: Uniform approximation for a Normal Distribution

- It is easy to compute random numbers over the range of (0,1)
- How do you generate a random number with a standard normal distribution?

Solution:

A uniform distribution has

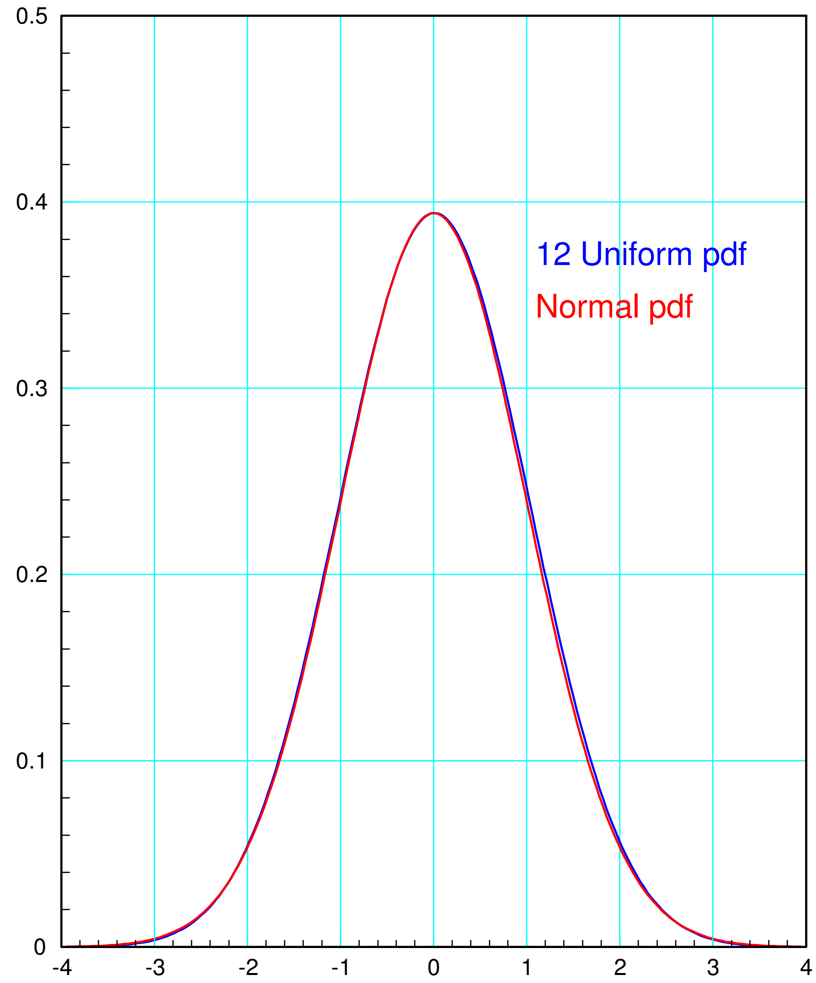
- mean = 1/2
- variance = 1/12

Sum twelve uniform distributions and subtract six

- mean = 0
- variance = 1

Its *very* difficult to tell the difference

- Central Limit Theorem in action



Summary

Summing pdf's converge to a normal distribution

- Central Limit Theorem

Likewise, you can approximate many pdf's with a normal distribution

This lets you determine probabilities fairly easily using a standard normal table

- or StatTrek
- or Matlab's erf() function

```
p = (1 - erf(z/sqrt(2)))/2
```