Central Limit Theorem

ECE 341: Random Processes

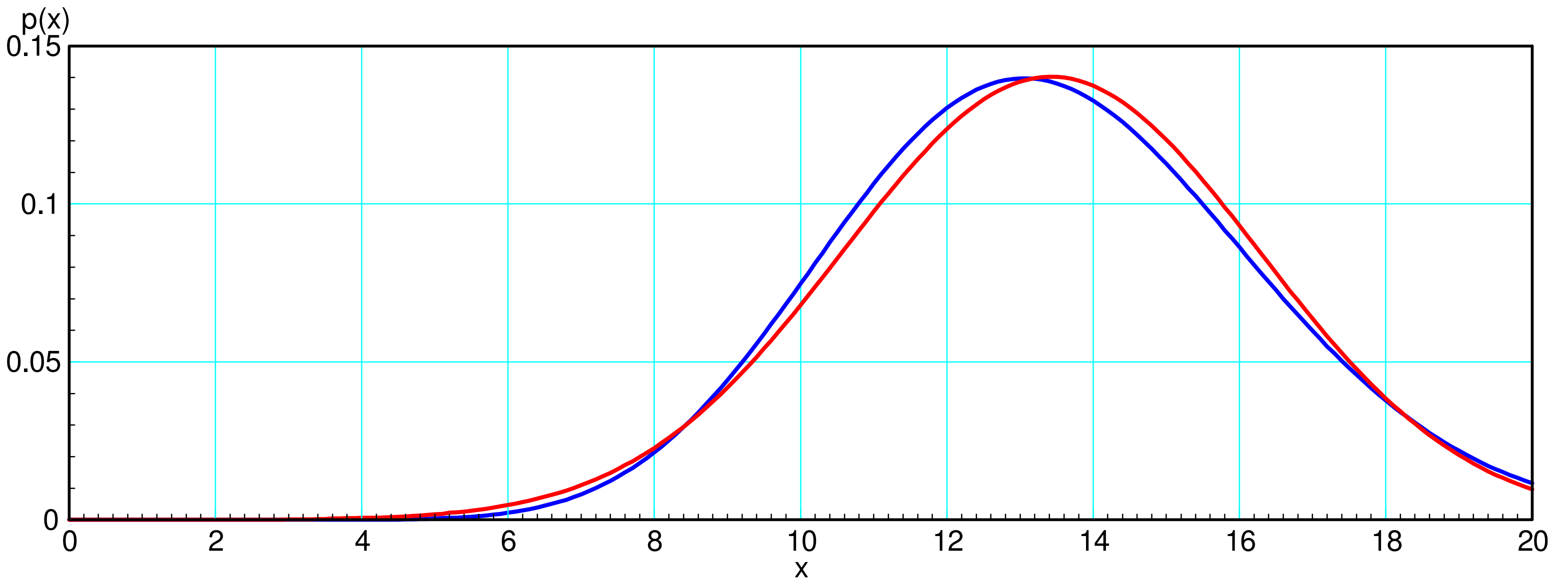
Lecture #17

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.

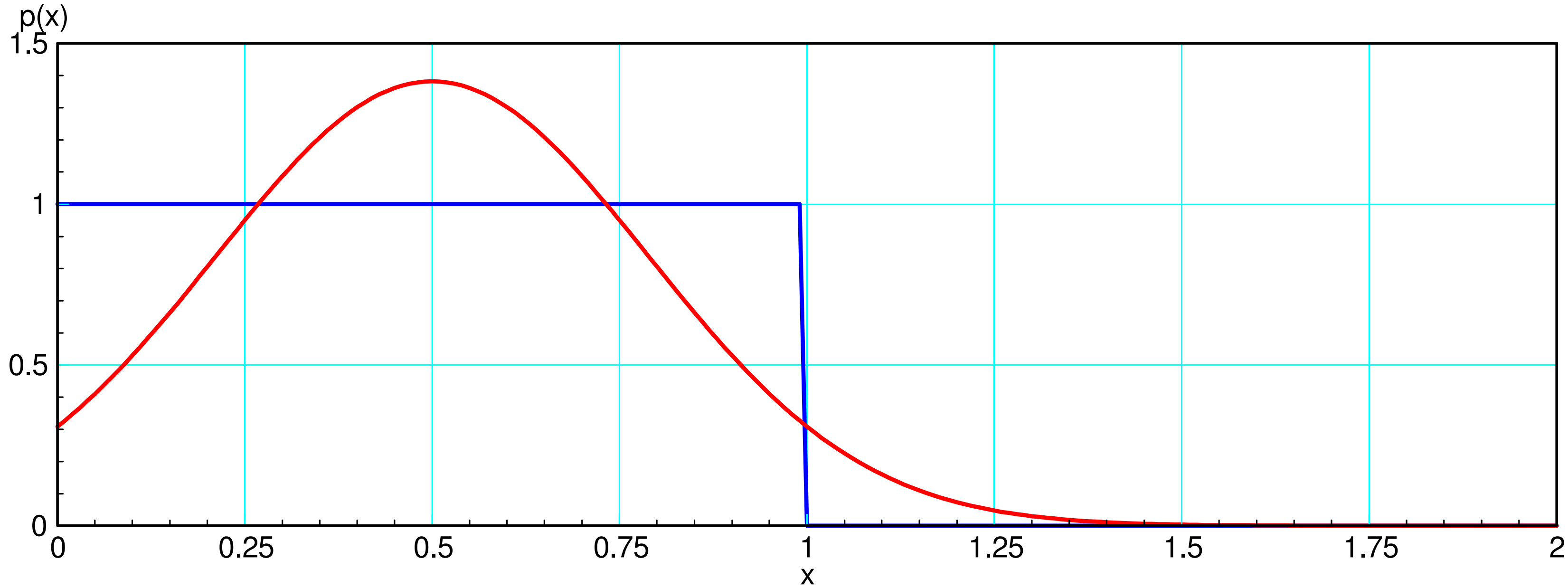
This is one of the reasons engineers tend to assume everything is described by a normal distribution (even when a Poisson distribution is more accurate). It also allows you to determine the probability for some fairly complex problems fairly accurately.



Example: Uniform Distribution.

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

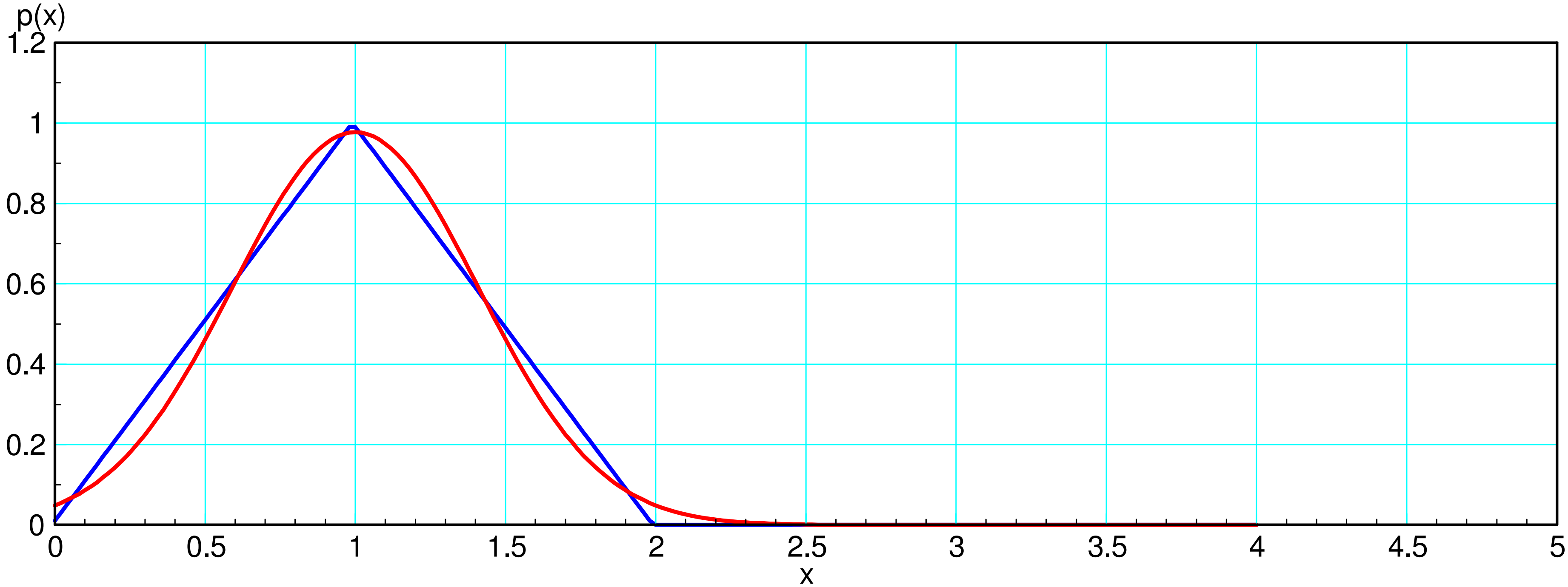
- Plot the pdf of A (blue)
- Plot the pdf of a Normal distribution with the same mean and variance (red)
- The two are different.



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

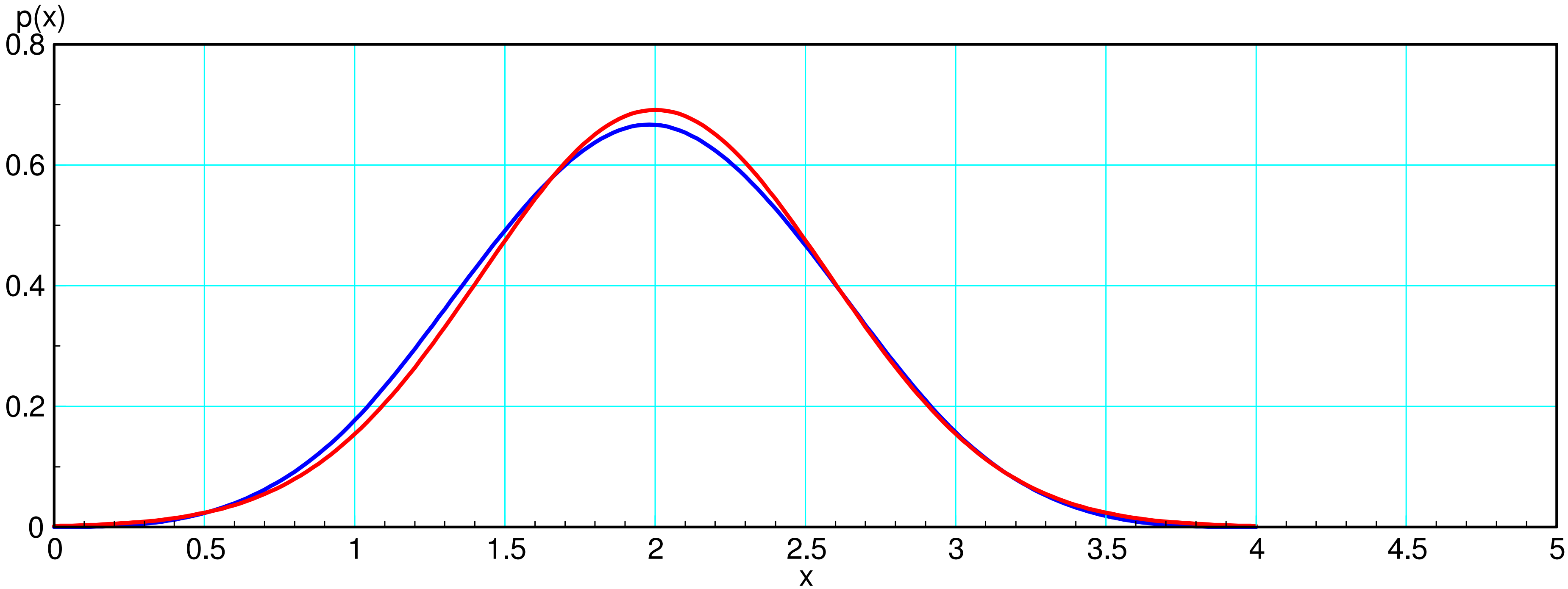
Summing two uniform distributions

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



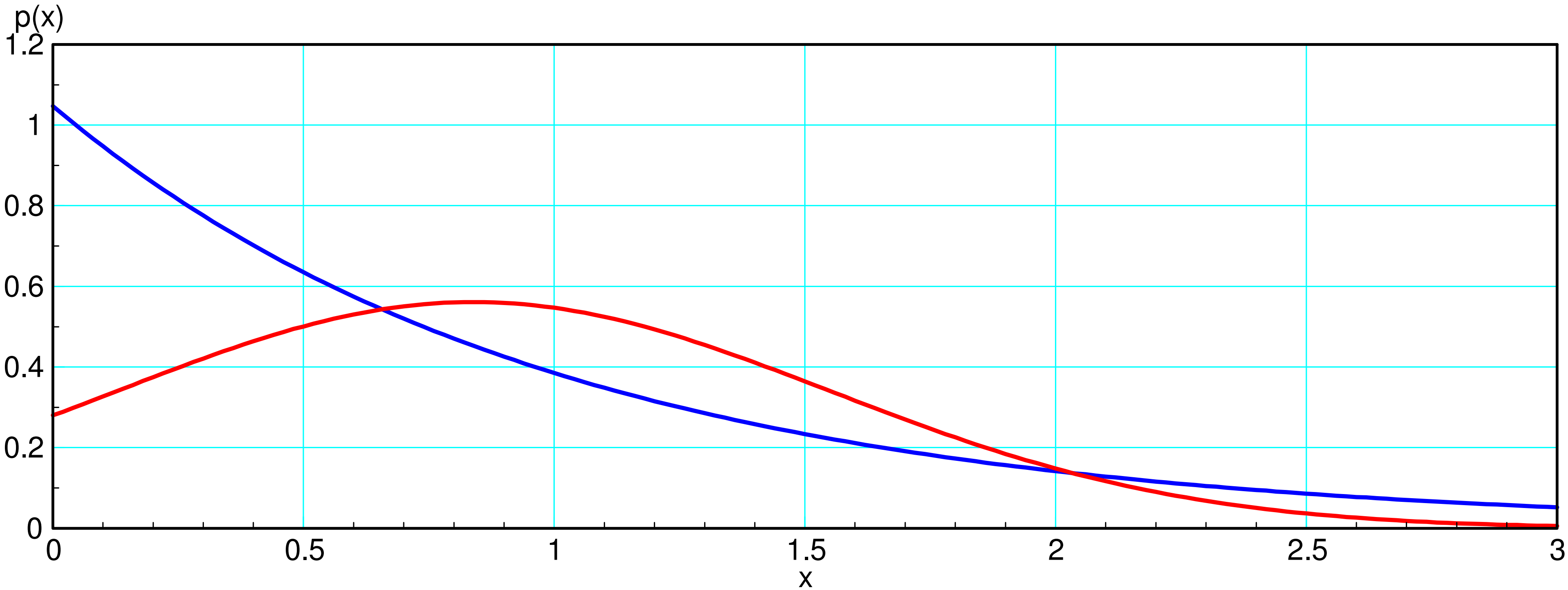
Summing four uniform distributions

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



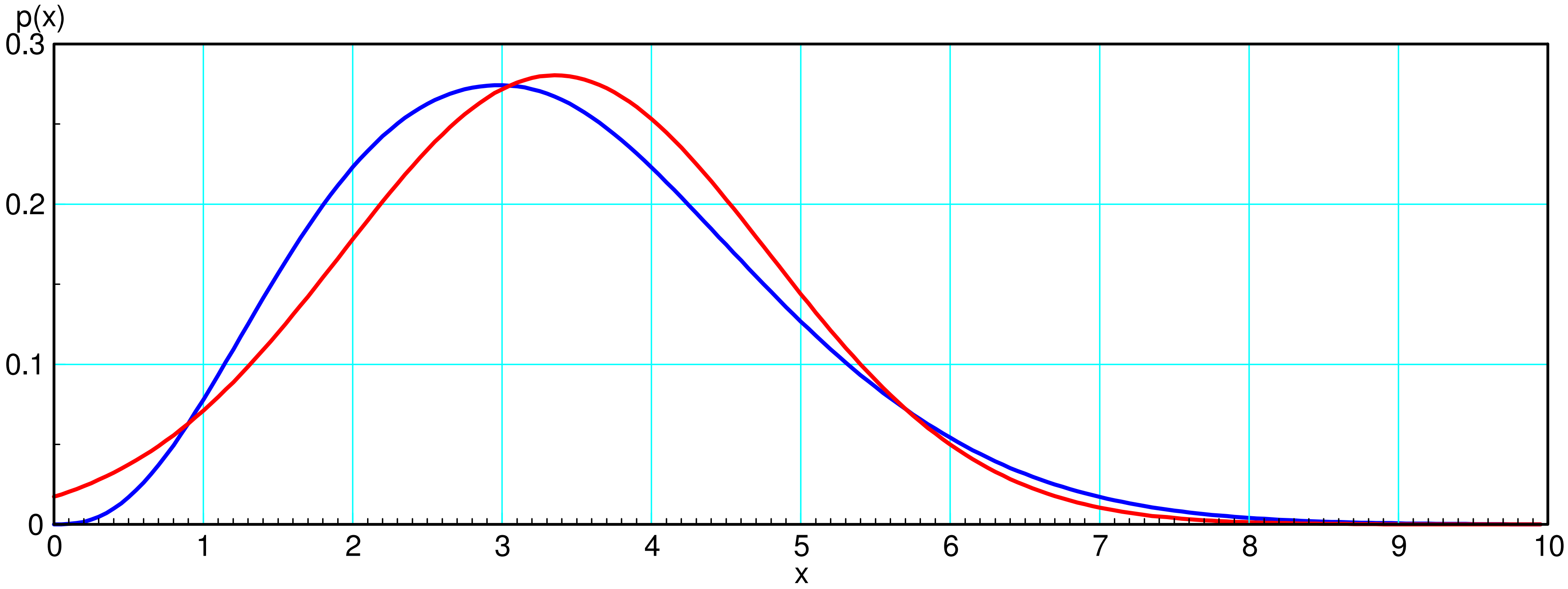
Example 2: Geometric Distribution

- Blue = geometric distribution with a mean of 1
- Red = Normal distribution with the same mean and variance
- They are not very close



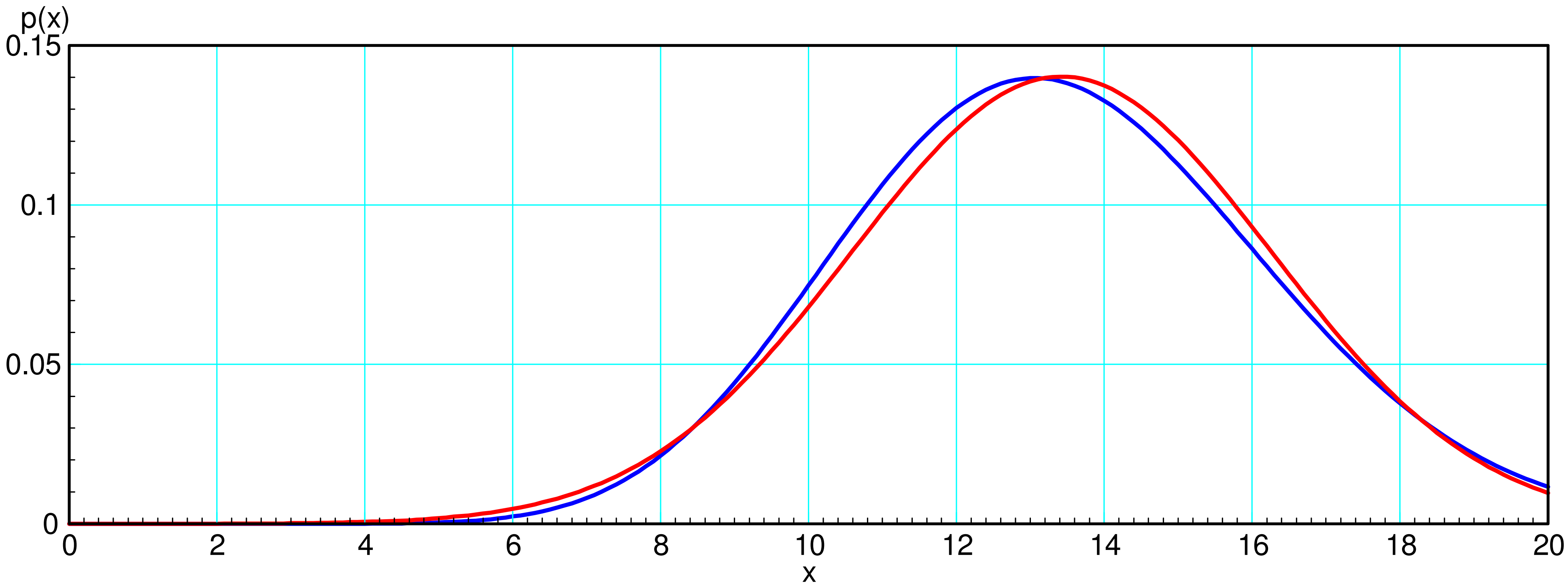
Sum of four geometric distributions (Poisson distribution)

- Blue = sum of four geometric distributions with a mean of 1
- Red = Normal distribution with the same mean and variance
- Closer...



Sum of 16 geometric distributions (Poisson distribution)

- Blue = sum of four geometric distributions with a mean of 1
- Red = Normal distribution with the same mean and variance
- Almost the same



Central Limit Theorem:

The sum of any distribution converges to a normal distribution.

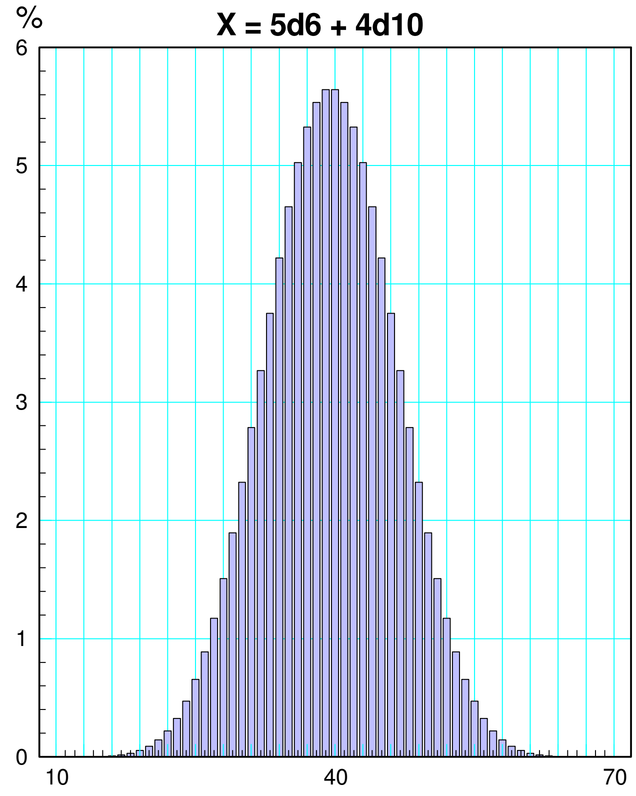
Example 1: Central Limit Theorem with Dice

Let X = 5d6 + 4d10

- The sum of rolling 5 six-sided dice (5d6) and four 10-sided dice (4d10).
- What is the probability of rolling 45 or higher? 55 or higher?

Exact Solution: Convolve the pdf for 5d6 and 4d10

```
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);

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

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



Monte-Carlo Simulation

```
N = 0;
for i=1:1e6
     d10 = sum( ceil(10*rand(1,4)) );
     d6 = sum( ceil(6*rand(1,5)) );
     X = d6 + d10;
     if(X >= 45)
          N = N + 1;
          end
     end
N/1e6
```

45 or higher: ans = 0.2384

- 0.2382 calculated

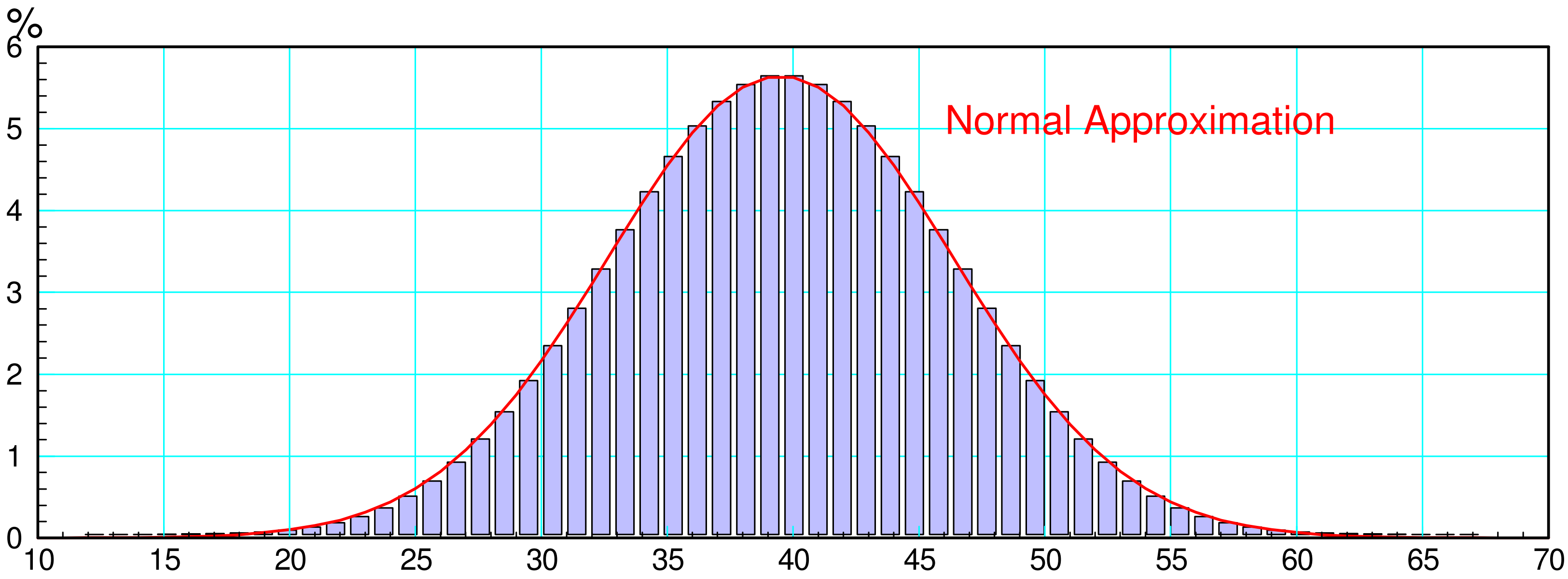
50 or higher ans = 0.07472

- 0.0748 calculated

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



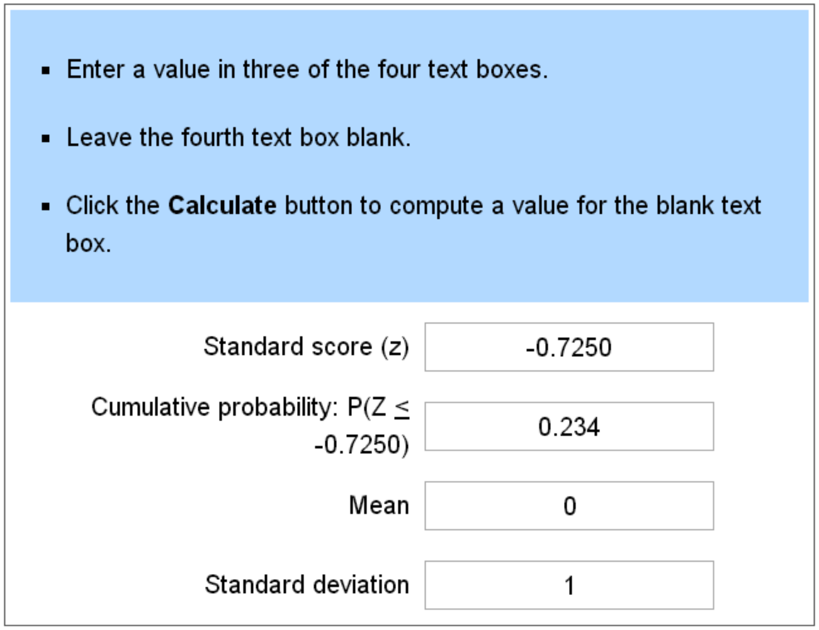
Converting z-score to a probability (StatTrek)

The z-score is the distance to the mean in terms of standard deviations

$$z = \left(\frac{44.5 - 39.5}{6.8981}\right) = 0.7250$$

A normal distribution converts this z-score to a probability

- p = 0.234 (vs. 0.2384 and 0.2382)



The z-score for 49.5 is (roll 50 or higher) is 1.450

$$z = \left(\frac{49.5-39.5}{6.8981}\right) = 1.450$$

The corresponds to a probability of 7.4% (vs. 7.472% and 7.48%)

- 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)	-1.450
Cumulative probability: P(Z ≤ -1.450)	0.074
Mean	0
Standard deviation	1

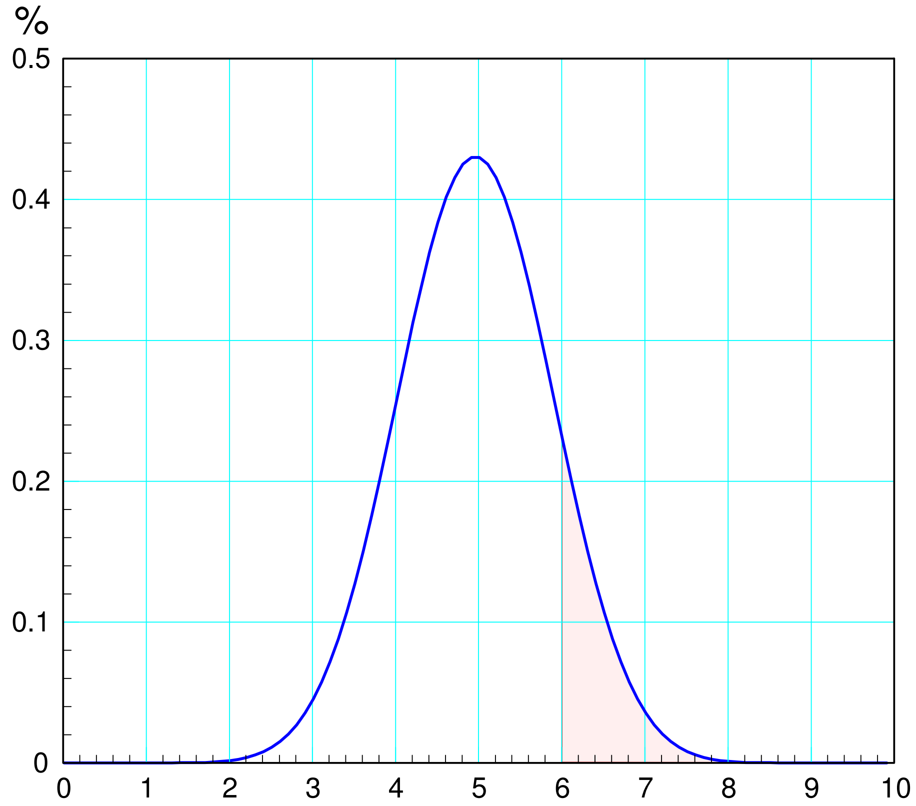
Example 2: Uniform Distribution.

- 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?

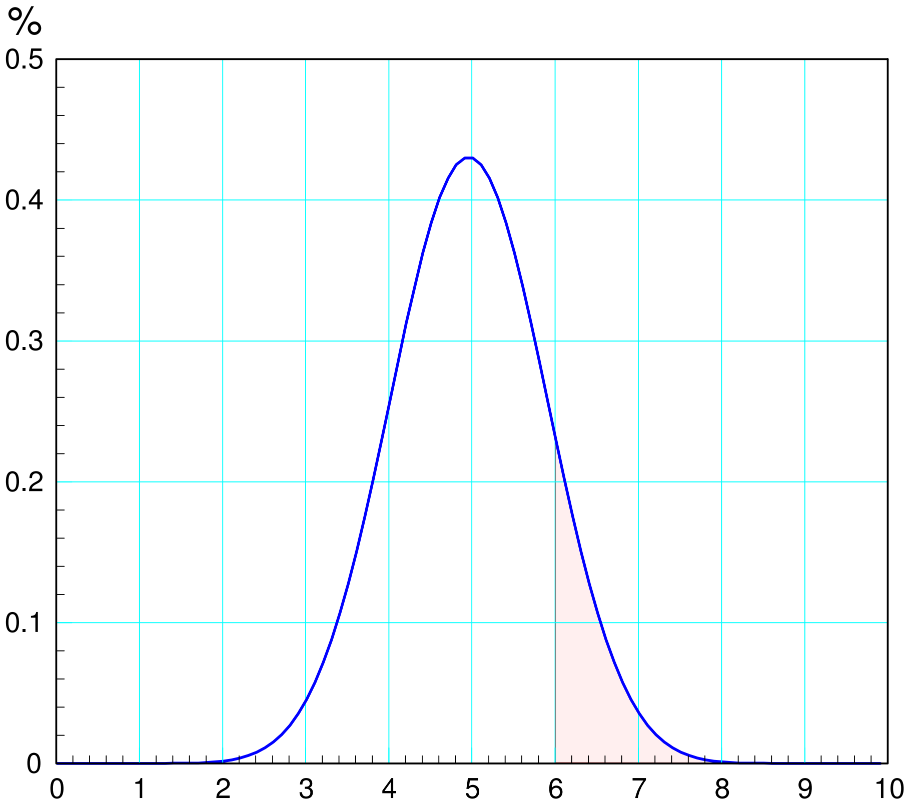
Solution: Convolution with matlab.

```
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


    0.1388     0.0137   monte-carlo
    0.1306     0.0121   convolution
```



Solution: Normal Approximation

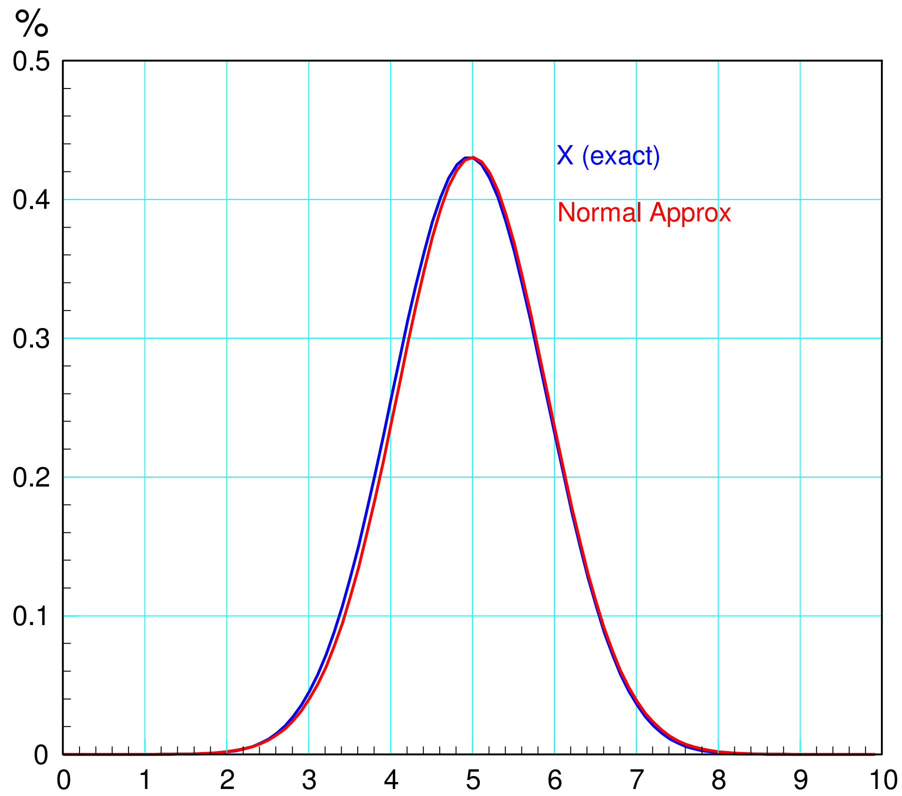
	Uniform(0,1)	10 x Uniform
mean	1/2	10/2
variance	1/12	10/12

The z-score for 6.00 is

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

This corresponds p = 0.137

- Normal Approx: 0.1370
- Computed: 0.1306
- Monte Carlo: 0.1388

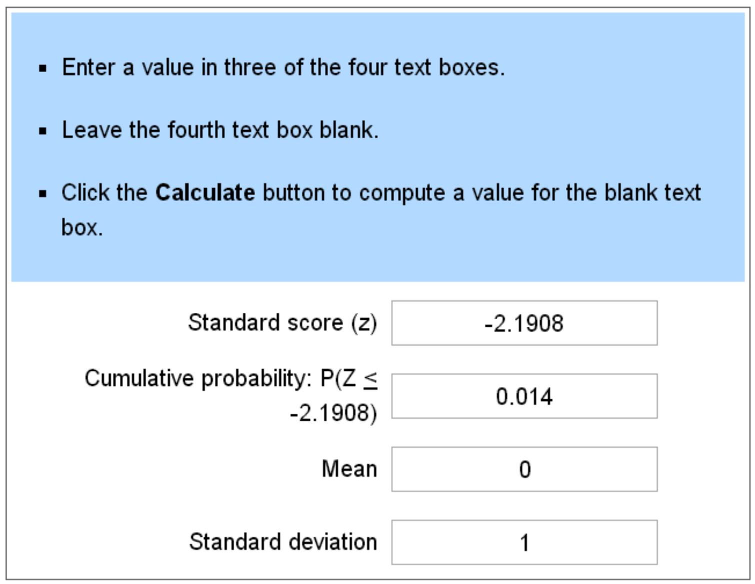


The z-score for rolling 7.00 or higher is

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

This corresponds to a probability of 0.014

- Normal Approx: 0.014
- Computed: 0.0121
- Monte Carlo: 0.0137



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?

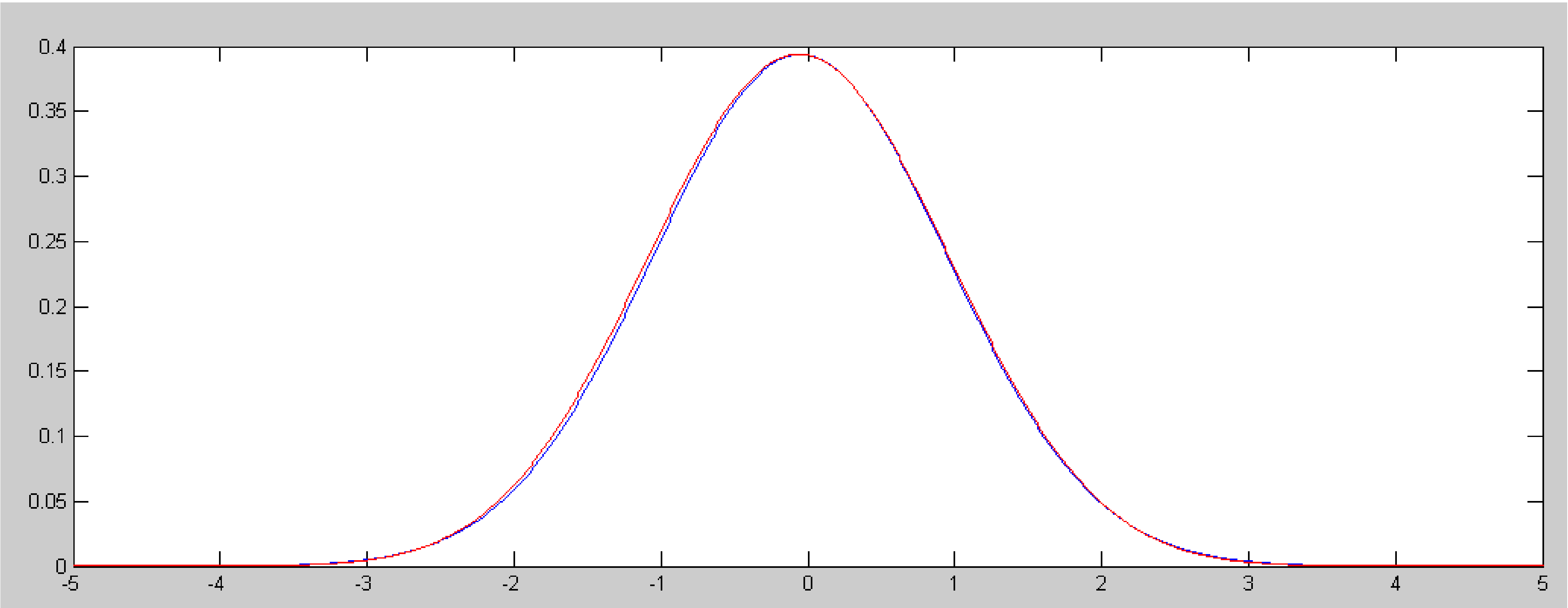
A uniform distribution has

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

Sum twelve uniform distributions and subtract six

- mean = 0
- variance = 1

Result is very close to a standard normal distribution



pdf for a Standard Normal Curve (red) and summing 12 uniform distributions and subtracting six

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