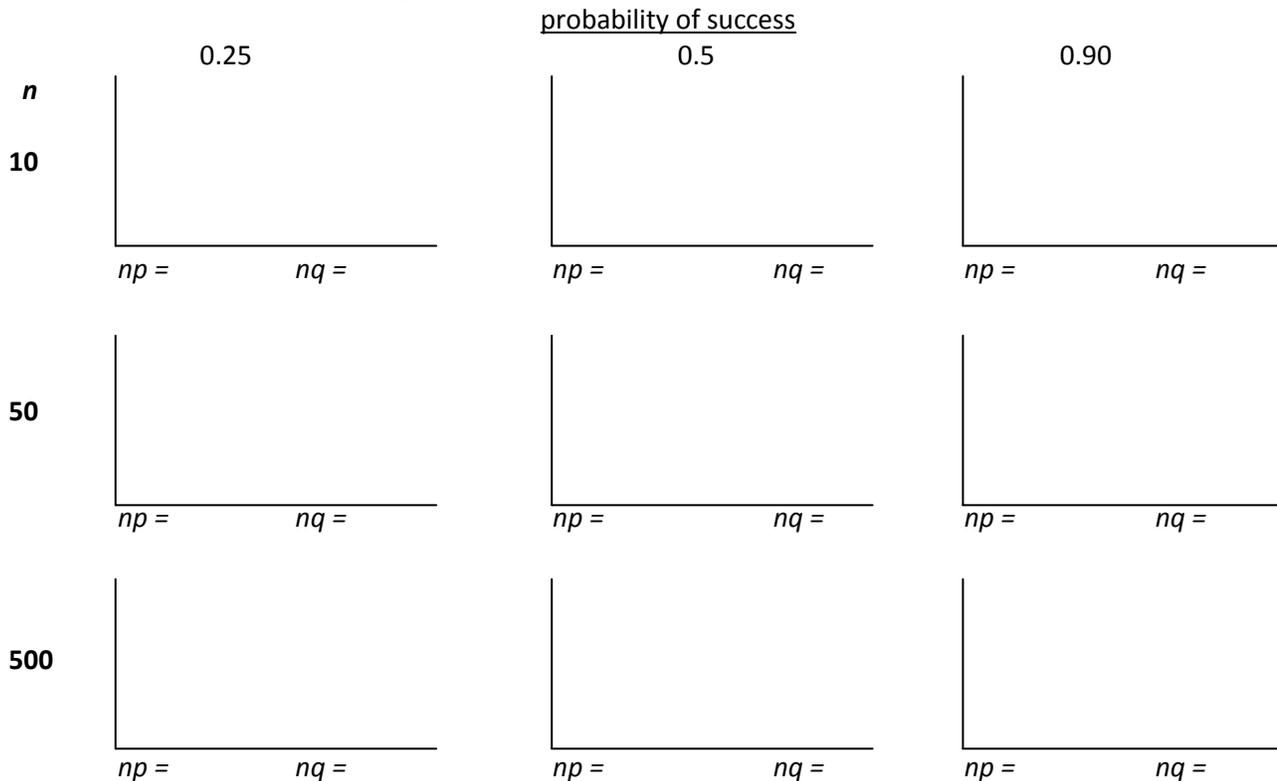


To show the number of successes in  $n$  trials we can create a **binomial** distribution. This **sampling** distribution model allows us to **predict** the natural sampling variability expected from samples of size  $n$ .

Use <http://www.stat.berkeley.edu/~stark/Java/Html/BinHist.htm> to model the following binomial distributions

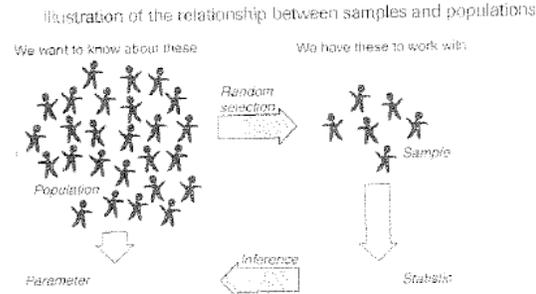
1) Quickly sketch each one and then describe the shape.

2) Then calculate the \_\_\_\_\_ condition to make a determination as to the appropriateness of approximating the distribution with a Normal Model.



We can easily transition from evaluating the number of successes in samples of size  $n$  to the proportion of successes by **dividing by  $n$** . So categorical data, **counts, percents, or proportions**, can be approximated with a normal model as long as certain assumptions and conditions are met:

<p>Assumptions / Conditions for using a Normal model as the SDM for a proportion:</p>	<p>Assumptions:</p> <ol style="list-style-type: none"> <li><b>Independent</b> - sampled values must be <b>independent</b> of each other. Conditions:             <ol style="list-style-type: none"> <li><b>Randomization</b> – <b>SRS</b> or at least representative and not <b>biased</b>.</li> <li><b>10% Condition</b> – If sampling w/o replacement Then <math>n \leq 10\%</math> of the population.</li> </ol> </li> <li>Sample Size - <math>n</math>, must be <b>large enough</b>. Conditions:             <ol style="list-style-type: none"> <li><b>Success/Failure</b> - <math>np \geq 10</math> and <math>nq \geq 10</math>.</li> </ol> </li> </ol>
---	--

Means summarize data	quantitative
We want to know the true population proportion (mean), __, but are often forced to work/estimate with a sample proportion (mean), __ .	<p><math>p</math> (<math>\mu</math>)</p> <p><math>\hat{p}</math> (<math>\bar{x}</math>)</p> 

We know that when we sample at random or randomize an experiment, the results we get will vary from sample-to-sample and from experiment-to-experiment. These results will not only be proportions (categorical variables) but means (quantitative variables) as well. We saw last class that the Normal model can often be utilized as the SDM for a proportion. But what kind of sampling distribution model is suitable for a mean? To investigate let's look at a simulation involving the mean of several dice.

<http://www.randomservices.org/random/apps/DiceExperiment.html>

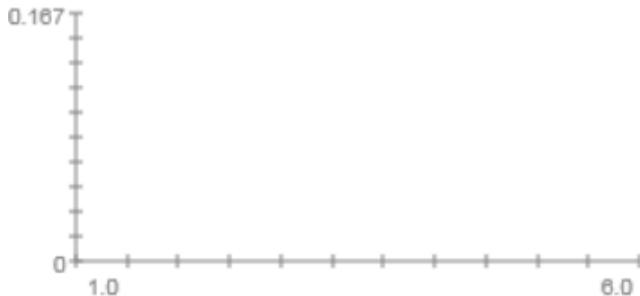
Goto: <http://www.math.uah.edu/stat/apps/DiceExperiment.html> 1) unclick "Show Distribution" 2) set slider so  $n$  = desired number 3) set dropdown boxes to Stop: 100, Fair, M: average score 4) click > a few times and note the results or your rolls 5) click >> to see the result of 100 rolls 6) perform several hundred rolls then click the "Show Distribution" box 7) Sketch the distribution and write down the mean and standard deviation.

$n = 1$  (1 die)

outcome	1	2				6
probability						

Using "1 variable stats" find the mean and standard deviation.

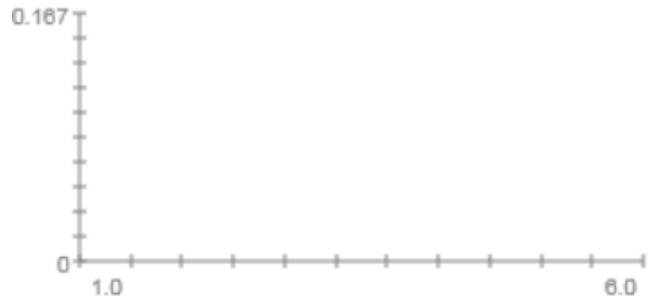
U =                      o =



$n = 2$  (2 die)

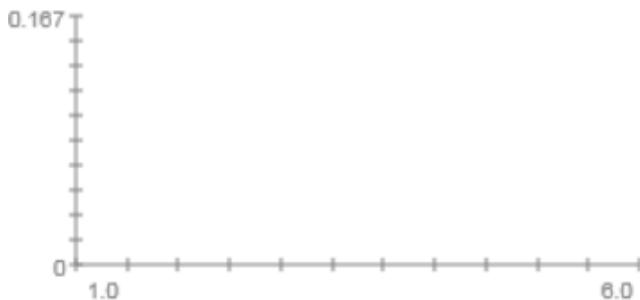
sum	2	3	4							12
Mean( $\div 2$ )										
Prob.										

U =                      o =



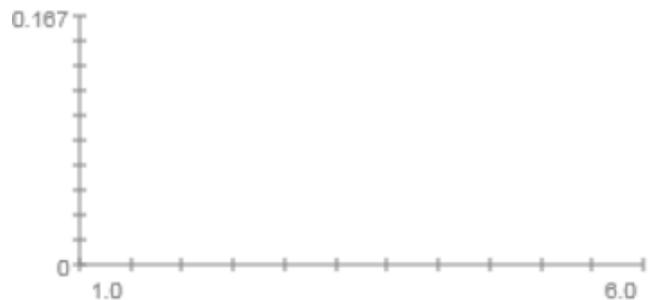
$n = 10$

U =                      o =



$n = 25$

U =                      o =



How does the distribution of the average change as the sample size increases?

Now let's investigate what the histogram of the sample means would look like for an initial population distribution that is not normal. Use the following applets:

Best, try this first [http://socr.ucla.edu/htmls/exp/Sampling\\_Distribution\\_CLT\\_Experiment.html](http://socr.ucla.edu/htmls/exp/Sampling_Distribution_CLT_Experiment.html)

<http://www.socr.ucla.edu/applets.dir/samplingdistributionapplet.html>

[http://onlinestatbook.com/2/sampling\\_distributions/clt\\_demo.html](http://onlinestatbook.com/2/sampling_distributions/clt_demo.html)



- 1) click and drag on the grey distribution to create the desired population distribution.
- 2) drag the bar next to Sample Size to the right to reveal then entire distribution.
- 3) set the drop down boxes to Mean N=2 and Mean N = 10
- 4) Click > and watch as the samples are selected and the mean calculated (blue)
- 5) Set Stop = 100 and click >> after taking a few hundred samples sketch the sampling distribution of means
- 6) Redo for N = 50

	Skewed Right	Bimodal	Your Creation
$n$			
2			
10			
50			
0			

What conclusions can you make? Explain below.

When you finish continue to page 4.