**Topic 13, 14, 15**

 ***Irrespective of the initial distribution*** from which the data is taken,

The **Distributions of Sample Statistics *approach* *Normal***  as n grows large.

* Mean is same
* Standard Deviation $^{SD}/\_{\sqrt{n}}$

Which mean and standard deviation to use depends on the initial variable type

|  |  |
| --- | --- |
|  | Sampling Distribution Variable Type |
|  | Categorical | Quantitative |
|  | Binary | Non-Normal | Normal |
| **Sample Statistic** | $$\hat{p}$$ | $$\overbar{x}$$ |
| Mean | π | µ |
| Standard Dev. | $$\sqrt{{π(1-π)}/{n}}$$ | $${σ}/{\sqrt{n}}$$ |
| n lower limit | $nπ\geq 10$ *and*$$n(π-1)\geq 10$$ | Depends on initial distribution. n~30 | No lower limit |
| Normality | Approximate | Approximate | Normal |

Note that the means do not change, but the standard deviations shrink as n grows. As n grows to infinity, then the *standard deviation of the sampling distribution* shrinks to zero.

 **Statistical Significance**

 How often a sampling result occurs by chance.

 **Statistically Significant**

If sampling result is unlikely due to random sampling variability, it is *statistically significant*

 Example 1: if you flip a coin twice, the statistical significance of two heads is .25

Example 2: If you flip a coin 30 times, and a head comes up each time with probability

P{H=30} = (.5)30=.0000000009, then you would say that event is statistically significant, and might suspect there is something funky about the coin (ie the coin is biased)

 **Statistical Confidence**

One can only calculate a statistic, but they are *Confident*the actual parameter is within a specific region near the statistic. For example, $\overbar{x}$ is calculated from the data. You are confident that the parameter µ is nearby. How confident you are depends on $σ$ and n.