Law of large numbers: The sample mean will tend to approach and stay close to the expected value.

The empirical rule: For approximately normal data sets:

- Approx 68% of the samples are in $\bar{X} \pm s$.
- Approx 95% of the samples are in $\bar{X} \pm 2s$.
- Approx 99.7% of the samples are in $\bar{X} \pm 3s$.

Central limit theorem: Assume i.i.d. samples $X_1, X_2, \ldots, X_n$ where X_i has mean μ (e.g. p) and variance σ^2 (e.g. $p(1-p)$).

Note: The aritmetic mean is:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

This aritmetic mean has variance:

$$\begin{aligned}
Var[\bar{X}] &= E\left[(\bar{X}-\mu)^2\right] \\
&= E\left[\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right)^2\right] \\
&= \frac{1}{n^2}E\left[\left(\sum_{i=1}^{n}(X_i - n\mu)\right)^2\right] \\
&= \frac{1}{n^2}E\left[\sum_{i=1}^{n}(X_i - n\mu)\sum_{j=1}^{n}(X_j - n\mu)\right] \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} E[(X_i - n\mu)(X_j - n\mu)] \\
&= \frac{1}{n^2}\sum_{i=1}^{n} E[(X_i - n\mu)(X_i - n\mu)] \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{\sigma^2}{n}
\end{aligned}$$

The theorem states:

$$X_1 + X_2 + \cdots + X_n$$

approximates a normal distribution with mean $n\mu$ and variance $n\sigma^2$.

Confidence intervals: A confidence interval is the interval in which we are 95% confident that the true mean value lies in this interval.