**Notes on Confidence Intervals**

A common form for confidence intervals of a parameter:

point estimate of parameter $\pm$ margin of error

Typically, the point estimate is an unbiased estimate of the parameter. For example, a point estimate of population mean $\mu$ is the sample mean $\overline{X}$.

The margin of error is often a product of two factors:
(i) a critical value, that is, a percentile from the appropriate distribution with the percentile depending on the level of confidence
(ii) the standard deviation (or an estimate of the standard deviation) of the point estimate

**Example.** A 95% confidence interval for $\mu$ based on a random sample of size $n$ when sampling from a normal($\mu$, $\sigma$) population is given by

$$\overline{x} \pm t_{(n-1,.975)} \frac{s}{\sqrt{n}},$$

where the sample standard deviation $s$ is an estimate of $\sigma$ (and hence $s/\sqrt{n}$ is an estimate of the standard deviation of $\overline{X}$ ) and $t_{(n-1,.975)}$ is the 97.5th percentile from the $t(n-1)$ distribution since $\frac{\overline{X}-\mu}{s/\sqrt{n}}$ has a t distribution with $n-1$ degrees of freedom.

**Example.** An approximate 90% confidence interval for a population proportion $p$ based on a large random sample of size $n$ is given by

$$\widehat{p} \pm z_{.95}\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}$$

where the sample proportion $\widehat{p} = \frac{x}{n}$ with $x =$ number of population items with characteristic under study and $z_{.95}$ is the 95th percentile of the standard normal distribution. The factor $\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}$ is the estimated standard deviation of $\widehat{p}$. The standard normal distribution is used here since $X/n$ has an approximate normal distribution based on the central limit theorem.

Lagniappe ( a little extra).

We note that an exact confidence interval for $p$ can be derived using binomial probabilities.

**Exact Confidence Interval for $p$**

Let $X \sim$ binomial$(n,\ p)$ where $p$ is unknown. Observe $X = x$.
Then a $(1 - \alpha)100\%$ confidence interval for $p$ is given by $(p_{\mathrm{L}},\ p_{\mathrm{U}})$ where $p_{\mathrm{L}}$ is a solution to

$$\sum_{k=0}^{x-1} \binom{n}{k} p^k \, (1-p)^{n-k} = 1 - \alpha/2$$

and $p_{\mathrm{U}}$ is a solution to

$$\sum_{k=0}^{x} \binom{n}{k} \, p^k \, (1-p)^{n-k} = \alpha/2.$$

**Example.** Suppose $n = 20$ and $x = 3$, and we desire a 95% confidence interval for $p$.

To get $p_{\mathrm{L}}$ , we solve (for $p$):

$$\sum_{k=0}^{2} \binom{20}{k} \, p^k \, (1-p)^{20-k} = .975$$

**TI-83** code: solve (binomcdf(20, p, 2) − .975, p, .01) where .01 is a guess

solution for $p_{\mathrm{L}}$: $\boxed{.0320709372}$

To get $p_{\mathrm{U}}$, we solve (for $p$):

$$\sum_{k=0}^{3} \binom{20}{k} \, p^k \, (1-p)^{20-k} = .025$$

**TI-83** code: solve(binomcdf(20, p, 3) − .025, p, .5) where .5 is a guess

solution for $p_{\mathrm{U}}$: $\boxed{.3789268265}$