Formula Sheet

Standardized Normal Data

$$z = \frac{\text{statistic} - \text{parameter}}{\text{standard deviation of the statistic}}$$

Standard Deviations for Sample Means and Sample Proportions

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \qquad \sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

Addition and Multiplication Rules for Probability

$$\Pr(A \text{ or } B) = \Pr(A) + \Pr(B) - \Pr(A \text{ and } B) \qquad \Pr(B|A) = \frac{\Pr(A \text{ and } B)}{\Pr(A)}$$

Inference about Proportions

One Sample Hypothesis Test

$$H_0 : p = p_0 \qquad z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$
$$H_A : p \neq p_0$$

This is robust, as long as there are at least 10 expected successes and 10 expected failures in the sample.

One Sample Confidence Interval

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Confidence intervals for proportions are less robust than hypothesis tests (because you are using $\hat{p}$ instead of p_0 to estimate the standard error). On the other hand, plus-4 confidence intervals are actually more robust than hypothesis tests. To make a plus-4 confidence interval, add 2 fake successes and 2 fake failures to the sample.

Two Sample Hypothesis Test

$$H_0 : p_A = p_B \qquad z = \frac{\hat{p}_A - \hat{p}_B}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}}$$
$$H_A : p_A \neq p_B$$

where $\hat{p}$ is the pooled proportion. Robust, as long as their are at least 5 successes and 5 failures in each sample.

Two Sample Confidence Interval

$$\hat{p}_A - \hat{p}_B \pm z^* \sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{n_A} + \frac{\hat{p}_B(1-\hat{p}_B)}{n_B}}$$

To make a plus-4 confidence interval for two sample proportions, add 1 success and 1 failure to each sample.

Inference about Means

One Sample Hypothesis Test

$$H_0 : \mu = \mu_0 \qquad t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \qquad dF = n - 1$$
$$H_A : \mu \neq \mu_0$$

One Sample Confidence Interval

$$\bar{x} \pm t^* \frac{s}{\sqrt{n}}$$

These formulas are robust even if the population is not normal, as long as the sample is "nice" ($n \geq 40$ usually safe; $n \geq 15$ okay for samples w/ little skew & no outliers).