Parameter: A number describing a population.

Statistic: A numer describing a sample.

Null/Alternative Hypotheses: The null hypothesis, H_0, is some statement about a parameter being tested. The alternative hypothesis, H_a, is some other statement against which the null hypothesis is being tested. They are compared using a significance level, α, and a p-value from a sample. If $p < \alpha$, the alternative hypothesis is more likely so we reject H_0 in favor of H_a. If $p > \alpha$, the sample data was consistent with H_0 and the null hypothesis is more likely. Note that rejecting the null hypothesis means H_a is more likely but *not* that the explanation for H_a is correct. Similarly, failing to reject H_0 does not mean that H_0 is true, just that the data is consistent with H_0.

Significance Level: The probability at which you call the likelihood of obtaining a particular sample (or one even more unusual) 'significant'.

Test Statistic: Measures the compatibility of the null hypothesis against the data. The test statistic can be used to find the probability of obtaining a sample as 'unusual' (or more 'unusual') as the one obtained. Examples include z, t, X^2, F, etc.

p-value: The probability associated to a particular test statistic. This is the probability of obtaining a sample as 'unusual' (or more 'unusual') as the one obtained.

Critical Value: The test statistic(s) at which one would begin to reject H_0, i.e. the 'smallest' possible test statistics that would correspond to a p-value which would cause one to reject H_0.

Type I Error: The event where you reject a null hypothesis that is true, i.e. a false positive.

Type II Error: The event where you fail to reject a false null hypothesis, i.e. a false negative.

Power: The power of a significance test is a measurement of the ability of the test to detect some alternative hypothesis. Specifically, the power is the probability that the test will reject H_0 when some other alternative is true.

Confidence Level: The probability that the confidence interval has of containing the estimated parameter.

Margin of Error: A numerical measure of spread of a sampling distribution. It is a number of standard deviations times a standard error. The margin of error only accounts for random sampling error.

Unbiased Estimator: A statistic used to estimate a parameter is unbiased if the mean of the sampling distribution is equal to the true value of the parameter being estimated. Otherwise, the statistic is called biased.

Robust/Resistant: A statistical test is robust if the probability calculations are insensitive to violations of assumptions for the test, i.e. the probabilities do not change 'much' if the assumptions are violated 'a bit'.

Sampling Distribution: For a distribution with mean μ and standard deviation σ, then the distribution of sample means of size n is given by $N(\mu, \sigma/\sqrt{n})$ if $n \geq 30$ or the original distribution sampled from is normal. The test statistic for a sample mean is then $z = \dfrac{\overline{x} - \mu}{\sigma/\sqrt{n}}$ with confidence interval $\overline{x} \pm z^* \dfrac{\sigma}{\sqrt{n}}$.

One-Sample t-test: For a sample of size n drawn from a population with mean μ and unknown standard deviations, but with sample standard deviation s, the test statistic is $t = \dfrac{\overline{x} - \mu}{s/\sqrt{n}}$ with confidence interval $\overline{x} \pm t^* \dfrac{s}{\sqrt{n}}$. If $n < 15$, the sample must be normally distributed, if $15 \leq n < 40$, the sample must have no skewness or outliers, and no assumption need be made if $n \geq 40$. One-sample t-procedures are robust. The degrees of freedom for this test are $\text{dof} = n - 1$.

Two-Sample t-test: For a samples of size n_1, n_2 drawn from a population with mean μ and unknown standard deviation, but with sample standard deviation s_1, s_2, the test statistic is

$$t = \frac{(\overline{x}_1 - \overline{x}_2) - D}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}},$$

where D is the hypothesized difference for $\mu_1 - \mu_2$, with confidence interval $(\overline{x}_1 - \overline{x}_2) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$. If $n_1 + n_2 < 15$, the samples must be normally distributed, if $15 \leq n_1 + n_2 < 40$, the samples must have no skewness or outliers, and no assumption need be made if $n_1 + n_2 \geq 40$. The degrees for freedom for this test are approximately $\text{dof} = \min\{n_1 - 1, n_2 - 1\}$. Two-sample t-procedures are robust. They are even more robust than one-sample t-procedures.

Pooled Two-Sample t-test: For a samples of size n_1, n_2 drawn from a population with mean μ and unknown but equal standard deviations and sample standard deviations s_1, s_2, the test statistic is

$$t = \frac{(\overline{x}_1 - \overline{x}_2) - D}{s_p \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}},$$

where D is the hypothesized difference for $\mu_1 - \mu_2$ and $s_p^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$, with confidence interval $(\overline{x}_1 - \overline{x}_2) \pm t^* \, s_p \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$. If $n_1 + n_2 < 15$, the samples must be normally distributed, if $15 \leq n_1 + n_2 < 40$, the samples must have no skewness or outliers, and no assumption need be made if $n_1 + n_2 \geq 40$. The degrees for freedom for this test are approximately $\text{dof} = n_1 + n_2 - 2$. Pooled two-sample t-procedures are even more robust than two-sample t-procedures. Note that a pooled two-sample t-procedure is inappropriate if s_1/s_2 is not between 0.5 and 2.

One-Sample Proportion: For a sample proportion $\hat{p} = X/n$ taken from a population with assumed proportion p_0, the test statistic is $z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)}} n$ with confidence interval $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$. One need assume that the sample contains at least 10 successes and 10 failures. One can also determine the sample sizes needed for a desired margin of error, m, using $n = (z^*/m) p^*(1-p^*)$ if p^* is some predicted/known/assumed value for p, or $n = \frac{1}{4}(z^*/m)^2$ if there is no prior estimate value for p.

Two-Sample Proportion: For sample proportions $\hat{p}_1 = X_1/n_1, \hat{p}_2 = X_2/n_2$, the test statistic is $z = \frac{(\hat{p}_1 - \hat{p}_2) - D}{\text{SE}_{D_p}}$, where D is the assumed difference,

$$\text{SE}_{D_p} = \sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

and $\hat{p} = \frac{X_1 + X_2}{n_1 + n_2}$, with confidence interval $(\hat{p}_1 - \hat{p}_2) \pm z^* \text{SE}_D$, where

$$\text{SE}_D = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

One need assume that each sample contains at least 5 successes and 5 failures.