We start with a **null hypothesis** H_0, which we'll assume to be true until we have evidence to the contrary. Exactly what constitutes contrary evidence is determined by our choice of **alternative hypothesis** H_1. We **reject** H_0 in favor of H_1 if an appropriate test statistic falls in a **critical region**. If the test statistic does not fall in the critical region, then we fail to reject H_0.

1. Let $x_1, x_2, \ldots, x_n$ be a random sample from a population with mean μ. We will use the sample mean $\overline{x}$ to test the null hypothesis $H_0 : \mu = \mu_0$ against different alternative hypotheses. Describe the critical region for each alternative hypothesis (i.e. which values of $\overline{x}$ would lead us to reject H_0 in favor of H_1).

a) $H_1 : \mu \neq \mu_0$.

$\bar{X}$ much smaller than μ_0 or $\bar{x}$ much bigger than μ_0.

b) $H_1 : \mu > \mu_0$.

$\bar{X}$ much bigger than μ_0.

c) $H_1 : \mu < \mu_0$.

$\bar{X}$ much smaller than μ_0.

When testing a hypothesis, there are two kinds of mistake:

- **Type I error** is rejecting H_0 when it is true;
- **Type II error** is failing to reject H_0 when it is false.

The probability of a type I error is α. The probability of a type II error is β. Usually you should choose the largest acceptable value for α since this will minimize β. The critical region is chosen so that the test statistic lands in the critical region with probability α (when H_0 is true). It may also be useful to find the P-value (or **observed significance level**) of your data: this is the smallest value for α that leads you to reject H_0 with your data Different disciplines have different standards for what P-value is evidence against H_0, but generally a P-value of 0.05 (or less) is considered significant while a P-value of 0.01 (or less) is highly significant.

Example. We can use your height data to test the hypothesis that the mean height of a male Gonzaga undergraduate is 70 inches (because a web site told me that the mean height of a 20 year old American male is about 70 inches). Formally, we'll test $H_0 : \mu = 70$ against $H_1 : \mu \neq 70$. We assume that the population is normally distributed and that the class constitutes a random sample. The sample data give $\overline{x} = 72.08$ and $s = 2.19$ with a sample size of $n = 28$. Our test statistic is

$$t = \frac{72.08 - 70}{2.19/\sqrt{28}} \approx 5.02$$

The P-value of the test is $2P(T \geq 5.02) \approx 0.000029$ (where T has a t-distribution with 27 df; we multiply by 2 because this is a two-tailed test and $\overline{x}$ being either large or small constitutes evidence against H_0).

2. What is the conclusion of the test in the example (i.e. should you reject H_0)?

We should reject H_0. Such a small observed significance level means that the result is highly significant.

For **tests about the mean** ($H_0 : \mu = \mu_0$) test statistics are:

- $z = \dfrac{\overline{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$ for samples from a population with a known variance σ^2 (all sample sizes if the population is normal, otherwise just for large samples);
- $z = \dfrac{\overline{x} - \mu_0}{\frac{s}{\sqrt{n}}}$ for large samples;
- $t = \dfrac{\overline{x} - \mu_0}{\frac{s}{\sqrt{n}}}$ for small samples ($n \leq 30$ or as big as your t-table goes) from a normally distributed population. The test statistic has a t distribution with $n - 1$ degrees of freedom.

For **tests about a population proportion** ($H_0 : \theta = \theta_0$) use the sample proportion $\hat{\Theta}$ or the sample total $X = n\hat{\Theta}$. Test statistics:

- x (X is binomial with parameters n and θ_0, works best for small samples);
- $z = \dfrac{\hat{\theta} - \theta_0}{\sqrt{\frac{\theta_0(1-\theta_0)}{n}}} = \dfrac{x - n\theta_0}{\sqrt{n\theta_0(1 - \theta_0)}}$ for large samples (both $n\theta_0 \geq 10$ and $n(1 - \theta_0) \geq 10$).

For **tests about the variance** ($H_0 : \sigma^2 = \sigma_0^2$) the test statistic is

- $\chi^2 = \dfrac{(n-1)s^2}{\sigma_0^2}$ for samples from a normally distributed population. The test statistic has a chi-square distribution with $n - 1$ degrees of freedom.

For **tests about the standard deviation**, square everything and do a test for the variance.

3. Newly purchased tires are supposed to be filled to a pressure of 30 lb/in^2. Let μ denote the true average pressure. We wish to test the hypothesis $H_0 : \mu = 30$ against the alternative hypothesis $H_1 : \mu \neq 30$. Use the given sample mean and sample standard deviation of a random sample of size $n = 100$ to test H_0 against H_1 at significance level $\alpha = 0.05$.

a) $\overline{x} = 28.2,\ s = 8$

Test stat: $t = \dfrac{28.2 - 30}{8/\sqrt{100}} = -2.25$. coming from t dist with 99 df. (In all parts of this problem).

P-value: .0267. Reject Ho

b) $\overline{x} = 28.2,\ s = 4$

Test stat: $t = \dfrac{28.2 - 30}{4/\sqrt{100}} = -4.5$.

P-value: .000018549 Reject Ho.

c) $\overline{x} = 30.6,\ s = 4$

Test stat: $t = \dfrac{30.6 - 30}{4/\sqrt{100}} = 1.5$

P-value: .1368 Fail to reject Ho.

4. A random sample of 10 Black Angus steers is weighed, giving a sample standard deviation of 238 pounds. Test $H_0 : \sigma = 250$ against $H_1 : \sigma \neq 250$.

$H_0: \sigma^2 = 62500 \quad H_1: \sigma^2 \neq 62500$. Test stat $\chi^2 = \dfrac{9(238)^2}{(250)^2} = 8.156736$

The test stat is between $\chi^2_{.975,9}$ and $\chi^2_{.025,9}$ so we fail to reject at $\alpha = .05$.

5. It is known that roughly $\frac{2}{3}$ of all people have a dominant right foot and $\frac{2}{3}$ have a dominant right eye. Do people also kiss to the right? The article "Human Behavior: Adult Persistence of Head-Turning Asymmetry" reported that in a random sample of 124 kissing couples, 80 of the couples tended to lean more to the right than left. Does this result suggest that more than half of all couples lean right when kissing? Does this result provide evidence against the hypothesis that $\frac{2}{3}$ of all kissing couples lean right?

$H_0: \theta = \frac{1}{2}$ against $H_1: \theta > \frac{1}{2}$: $Z = \dfrac{80 - 62}{\sqrt{31}} \approx 3.23$ P-value: .0006 Reject Ho at $\alpha = .05$

$H_0: \theta = \frac{2}{3}$ against $H_1: \theta < \frac{2}{3}$: $Z = \dfrac{80 - 124(\frac{2}{3})}{\sqrt{\frac{2}{3}(\frac{1}{3})124}} \approx -.508$ P-value: .3057 Fail to reject at $\alpha = .05$