Hypothesis Testing

Four steps to test hypothesis

1. Formulate H_0 and H_1, and specify α
2. Using the sampling distribution of an appropriate test statistic, determine a critical region of size α
   a. Commonly used: normal $(z), t, \chi^2, F$
3. Determine the value of the test statistic from the sample data
4. Check whether the value of the test statistic falls into the critical region

Type I and Type II errors

1. Type I error (α): False positive (rejection) error

   Rejection of a null hypothesis when it is true
2. Type II error (β): False negative (acceptance) error

   Acceptance of the null hypothesis when it is false

Tests Concerning Mean

From one population

Assuming either the samples come from normal pop or the sample size is large enough to justify normal approximations (CLT).

Based on the LR (likelihood ratio test), use z-stat ($z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$) for test of $H_0: \mu = \mu_0$

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

If the number of observations less than 30, use t-stat ($t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$) with degree of freedom $df = n - 1$ for $H_0: \mu = \mu_0$

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

Concerning Difference between / within Means

With equal variance and independent assumption, For testing $H_0: \mu_1 - \mu_2 = \delta$, use z-stat

$$Z = \frac{\overline{x_1} - \overline{x_2} - \delta}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

If n_1 and n_2 are small (<30) , use t-stat with degree of freedom $df = n_1 + n_2 - 2$

$$t = \frac{\overline{x_1} - \overline{x_2} - \delta}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$S_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

If two samples are not independent, using paired t-test

Test concerning Variances

From one population

Use χ^2-stat $\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$ for test of $H_0: \sigma^2 = \sigma_0^2$, with critical regions as $\chi^2 \geq \chi^2_{\alpha/2,n-1}$ or $\chi^2 \leq \chi^2_{1-\alpha/2,n-1}$

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

Concerning the ratio of two variances

Use $\frac{s_1^2}{s_2^2} \sim f$ for $H_0: \sigma_1^2 = \sigma_2^2$ with critical regions

$$\frac{s_1^2}{s_2^2} \geq f_{a/2,n_1-1,n_2-1} \;\; if \; s_1^2 \geq \; s_2^2$$

$$\frac{s_2^2}{s_1^2} \geq f_{a/2,n_2-1,n_1-1} \;\; if \; s_1^2 \leq \; s_2^2$$

Test Concerning Proportions