Confidence intervals and hypothesis testing

Confidence intervals

General form is

$$[(point\,estimator) - m, (point\,estimator) + m]$$

where the margin of error is

$$m = (t^* \, or \, z^*) SE$$

The formulae for the point estimator and for the standard error depend on the particular type of problem. Whether you use a critical t or z depends on the type of problem. The value of a critical z depends only on the confidence level. The value of a critical t depends only on the confidence level and the number of degrees of freedom.

Hypothesis testing

For a one population problem the null hypothesis is a statement that some parameter equals a specific value. The alternative hypothesis is one of three possibilities: that parameter is not equal to the specific value, greater than the specific value or less than the specific value.

For a two population problem we only consider problems where the null hypothesis is that two population parameters are equal. The alternative is either that the two parameters are not equal or that one is greater than the other. the difference of the two parameters To compute the p-value for your test statistic you assume the null hypothesis is true. Pay attention to whether it is a one or two tailed test.

If the p-value is less than the significance level α we reject the null hypothesis and accept the alternative hypothesis. If the p-value is greater than or equal to the significance level α, we do not reject the null hypothesis.

One population, mean of a quantitative RV, σ known

$$
\begin{aligned}
&point\,estimator = \overline{x},\\
&SE = \frac{\sigma}{\sqrt{n}},\\
&test\,statistic = z = \frac{\overline{x} - \mu_0}{SE}
\end{aligned}
$$

Two populations, means of quantitative RV's, $\sigma's$ known

$$
\begin{aligned}
&point\,estimator = \overline{x_1} - \overline{x_2},\\
&SE = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}},\\
&test\,statistic = z = \frac{\overline{x_1} - \overline{x_2}}{SE}
\end{aligned}
$$

We are assuming null hyp is that $\mu_1 = \mu_2$.

One population, mean of a quantitative RV, σ unknown

$$
\begin{aligned}
&point\,estimator = \overline{x},\\
&SE = \frac{s}{\sqrt{n}},\\
&test\,statistic = t = \frac{\overline{x} - \mu_0}{SE}, \quad df = n - 1
\end{aligned}
$$

Two populations, means of quantitative RV's, $\sigma's$ unknown

$$
\begin{aligned}
&point\,estimator = \overline{x_1} - \overline{x_2},\\
&SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}},\\
&test\,statistic = t = \frac{\overline{x_1} - \overline{x_2}}{SE}
\end{aligned}
$$

df is smaller of $n_1 - 1$ and $n_2 - 1$. We are assuming null hyp is that $\mu_1 = \mu_2$.

One population, proportion for a binary categorical RV

$$
\begin{aligned}
&point\,estimator = \hat{p}, \\
&SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \\
&test\,statistic = z = \frac{\hat{p} - p_0}{SE}
\end{aligned}
$$

Two populations, proportions for binary categorical RV's

$$
point\,estimator = \hat{p_1} - \hat{p_2},
$$

For confidence intervals the standard is taken to be

$$
SE = \sqrt{\frac{\hat{p_1}(1-\hat{p_1})}{n_1} + \frac{\hat{p_2}(1-\hat{p_2})}{n_2}}
$$

For hypothesis testing we use the pooled estimator

$$
\begin{aligned}
&\hat{p} = \frac{X_1 + X_2}{n_1 + n_2}, \\
&SE = \sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_1} + \frac{1}{n_2})}, \\
&test\,statistic = z = \frac{\hat{p_1} - \hat{p_2}}{SE}
\end{aligned}
$$