Confidence Intervals

Confidence interval, single mean, variance known:

$$P\left(\bar{X}-\frac{Z_{\alpha/2}\sigma}{\sqrt{n}}<\mu<\bar{X}+\frac{Z_{\alpha/2}\sigma}{\sqrt{n}}\right)=1-\alpha \qquad n\geq\left(\frac{Z_{\alpha/2}\sigma}{e}\right)^2 \qquad e=\left(\frac{Z_{\alpha/2}\sigma}{n^{0.5}}\right)$$

Confidence interval, single mean, variance NOT known:

$$P\left(\bar{x}-\frac{t_{\alpha/2}S}{\sqrt{n}}<\mu<\bar{x}+\frac{t_{\alpha/2}S}{\sqrt{n}}\right)=1-\alpha \qquad \upsilon = n_1 - 1$$

Tolerance Limits: $(1-\gamma)$% confidence that the given limits contain at least the proportion $(1-\alpha)$ of the measurements: $\mu \pm k\,\sigma$ $\quad \underline{x} \pm k\,s$

Difference between 2 means, both variances known:

$$P\left[(\bar{x}_1-\bar{x}_2)-z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}<\mu_1-\mu_2<(\bar{x}_1-\bar{x}_2)+z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right]=1-\alpha$$

$$\sigma_{\bar{x}_1-\bar{x}_2}=\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$

Difference between 2 means, variances NOT known, but assumed EQUAL:

$$P\left((\bar{x}_1-\bar{x}_2)-t_{\alpha/2}\,s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}<(\mu_1-\mu_2)<(\bar{x}_1-\bar{x}_2)+t_{\alpha/2}\,s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)=1-\alpha$$

$$S_p^2=\frac{(n_1-1)\,S_1^2+(n_2-1)\,S_2^2}{n_1+n_2-2} \qquad \upsilon = n_1 + n_2 - 2$$

Difference between 2 means, variances NOT known, but assumed **UN**EQUAL:

$$P\left((\bar{x}_1-\bar{x}_2)-t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}<(\mu_1-\mu_2)<(\bar{x}_1-\bar{x}_2)+t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}\right)=1-\alpha$$

$$\upsilon=\frac{\left(s_1^2/n_1+s_2^2/n_2\right)^2}{\left(s_1^2/n_1\right)^2/(n_1-1)+\left(s_2^2/n_2\right)^2/(n_2-1)}$$

Confidence Interval for Mean of Differences of Paired

Data: $P\left\{\bar{d} - \frac{t_{\alpha/2} s_d}{\sqrt{n}} < \mu_d < \bar{d} + \frac{t_{\alpha/2} s_d}{\sqrt{n}}\right\} = 1 - \alpha$ $\mu_d = \underline{\mu}_1 - \underline{\mu}_2$ $\underline{d} = \underline{x}_1 - \underline{x}_2$

$$s_d = \sum_{i=1}^{n} \frac{\left(x_1 - x_2 - \mu_d\right)^2}{n-1} \qquad \upsilon = df = n_{pairs} - 1$$

Prediction Interval of a Single Value: $P\left\{ \bar{x} - Z_{\alpha/2}\sigma\sqrt{1+\frac{1}{n}} < x_o < \bar{x} + Z_{\alpha/2}\sigma\sqrt{1+\frac{1}{n}} \right\} = 1 - \alpha$

Or, $P\left\{ \bar{x} - t_{\alpha/2} s\sqrt{1+\frac{1}{n}} < x_o < \bar{x} + t_{\alpha/2} s\sqrt{1+\frac{1}{n}} \right\} = 1 - \alpha$ $\mu_d = \underline{x}_1 - \underline{x}_2$

Confidence interval for mean of a single proportion:

$$\mu_{\hat{p}} = p \qquad \sigma_{\hat{p}}^2 = \frac{pq}{n} \qquad Z = \frac{\hat{P} - p}{\sqrt{\frac{pq}{n}}} \qquad e = Z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$P\left(\hat{p} - Z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} < p < \hat{p} + Z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} \right) = 1 - \alpha$$

Confidence interval for mean of Two proportions:

$$\mu_{\hat{p}_1 - \hat{p}_2} = p_1 - p_2 \qquad \sigma_{\hat{p}_1 - \hat{p}_2}^2 = \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}$$

$$P\left((\hat{p}_1 - \hat{p}_2) - Z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} < (p_1 - p_2) < (\hat{p}_1 - \hat{p}_2) + Z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} \right) = 1 - \alpha$$

Confidence Interval of Variance: $P\left(\frac{(n-1)S^2}{\chi_{\alpha/2}^2} < \sigma^2 < \frac{(n-1)S^2}{\chi_{1-\alpha/2}^2} \right) = 1 - \alpha$

$$\chi^2 = \frac{(n-1)\,S^2}{\sigma^2} \qquad P\left(\sqrt{\frac{(n-1)S^2}{\chi_{\alpha/2}^2}} < \sigma < \sqrt{\frac{(n-1)S^2}{\chi_{1-\alpha/2}^2}} \right) = 1 - \alpha \qquad \upsilon = n - 1$$

Confidence Interval of Ratio of 2 Variances, assumed equal:

$$F = \frac{S_{num}^2 / \sigma_{num}^2}{S_{denom}^2 / \sigma_{denom}^2}, \qquad F = \frac{S_{num}^2}{S_{denom}^2} \quad if\ Ho: \sigma_{num} = \sigma_{denom}$$

$$P\left\{ \frac{s_{num}^2}{s_{denom}^2} \left(\frac{1}{f_{\alpha/2, \upsilon_{num}, \upsilon_{denom}}} \right) < \frac{\sigma_{num}^2}{\sigma_{denom}^2} < \frac{s_{num}^2}{s_{denom}^2} \left(\frac{1}{f_{1-\alpha/2, \upsilon_{num}, \upsilon_{denom}}} \right) \right\} = 1 - \alpha$$

$\upsilon_{num} = n_{num} - 1$ and $\upsilon_{denom} = n_{denom} - 1$

Or

$$P\left\{ \frac{s_{num}^2}{s_{denom}^2} \left(\frac{1}{f_{\alpha/2, \upsilon_{num}, \upsilon_{denom}}} \right) < \frac{\sigma_{num}^2}{\sigma_{denom}^2} < \frac{s_{num}^2}{s_{denom}^2} \left(f_{\alpha/2, \upsilon_{denom}, \upsilon_{num}} \right) \right\} = 1 - \alpha$$

HYPOTHESIS TESTING

Computing Type I and Type II Errors:

$\alpha = P(Z<Z_1) + P(Z>Z_2)$ $\beta = P(Z<Z_2) - P(Z<Z_1)$ when $\mu = \mu_{alt}$

For example, B could be c omputed using

$$Z_1 = \frac{x_1 - \mu_{alt}}{\sigma/\sqrt{n}} \quad and \quad Z_2 = \frac{x_2 - \mu_{alt}}{\sigma/\sqrt{n}}$$

Find α using x_1 and x_2.

Find B using same x_1 and x_2. Note that the Z_1 and Z_2 values for α and β are different.

Power = $1 - \beta$

Single mean, variance known: $Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$

Two means: variance known and equal: $Z = \frac{\left(\bar{X}_1 - \bar{X}_2\right) - (\mu_1 - \mu_2)}{\sigma\sqrt{1/n_1 + 1/n_2}}$

Two means: variance known and **un**equal: $Z = \frac{\left(\bar{X}_1 - \bar{X}_2\right) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$

Variance **un**known, single mean: $T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$

Variance unknown but assumed equal: $T = \frac{\left(\bar{X}_1 - \bar{X}_2\right) - (\mu_1 - \mu_2)}{s_p\sqrt{1/n_1 + 1/n_2}}$ df = n_1+n_2-2

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

Variances unknown and unequal:

$$T = \frac{\left(\bar{X}_1 - \bar{X}_2\right) - d_0}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} \qquad \upsilon = \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\left(s_1^2/n_1\right)^2/(n_1 - 1) + \left(s_2^2/n_2\right)^2/(n_2 - 1)}$$

Sample size for testing difference in means using both Type I and Type II errors,

one sided: $n \geq \frac{\left(z_\alpha + z_\beta\right)^2\left(\sigma_1^2 + \sigma_2^2\right)}{\delta^2}$ two-sided $n \geq \frac{\left(z_{\alpha/2} + z_\beta\right)^2\left(\sigma_1^2 + \sigma_2^2\right)}{\delta^2}$

Paired: $T = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}}$ $\upsilon = n_{pairs} - 1$ Remember WHEN you can pair

$\mu_d = \mu_1 - \mu_2$ $\bar{d} = \bar{x}_1 - \bar{x}_2$ $s_d = \sum_{i=1}^{n} \frac{(x_1 - x_2 - \mu_d)^2}{n-1}$

Normal approx for single proportion: $Z = \frac{x - n\,p_0}{\sqrt{n_0 p_0 q_0}}$, remember offset x

Two proportions: if assume $p_1 = p_2$, then: $\hat{p} = \frac{(x_1 + x_2)}{n_1 + n_2}$ $Z = \frac{(\hat{P}_1 - \hat{P}_2) - 0}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

Two proportions: if **cannot** assume $p_1 = p_2$, then: $Z = \frac{(\hat{P}_1 - \hat{P}_2) - (p_1 - p_2)}{\sqrt{\left(\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}\right)}}$

Single variance: $\chi^2 = \frac{(n-1)s^2}{\sigma_o^2}$

Critical region: Two-tailed: H_1: $\chi^2 < \chi_{1-\alpha/2}^2$ $\chi^2 > \chi_{\alpha/2}^2$

One-sided: H_1: $\sigma^2 < \sigma_o^2$ $\chi^2 < \chi_{1-\alpha}^2$

H_1: $\sigma^2 > \sigma_o^2$ $\chi^2 > \chi_{\alpha}^2$

Comparing two sample variances where assumed variances equal: $f = \frac{s_1^2}{s_2^2}$

Critical region:

- One-tailed: Critical region: $f < f_{1-\alpha(\upsilon 1,\upsilon 2)}$ or $f > f_{\alpha(\upsilon 1,\upsilon 2)}$
- Two-tailed: Critical region: $f < f_{1-\alpha/2(\upsilon 1,\upsilon 2)}$ & $f > f_{\alpha/2(\upsilon 1,\upsilon 2)}$

Goodness of fit test: $\chi^2 = \sum_{i=1}^{k} \frac{(Obs_i - Exp_i)^2}{Exp_i}$

