Making an Inference

Single Population **Mean**

Population normal

σ known — n large or small
$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$$
$$\overline{X} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

σ unknown

- n small
$$t = \frac{\overline{X} - \mu_0}{S / \sqrt{n}}$$
$$\overline{X} \pm t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$$

- n large
$$t \to Z$$
$$\overline{X} \pm t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$$

Population non-normal or unknown

σ known
- n small — nonparametric test
- n large
$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$$
$$\overline{X} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

σ unknown
- n small — nonparametric test
- n large
$$t = \frac{\overline{X} - \mu_0}{S / \sqrt{n}} \to Z$$
$$\overline{X} \pm t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$$
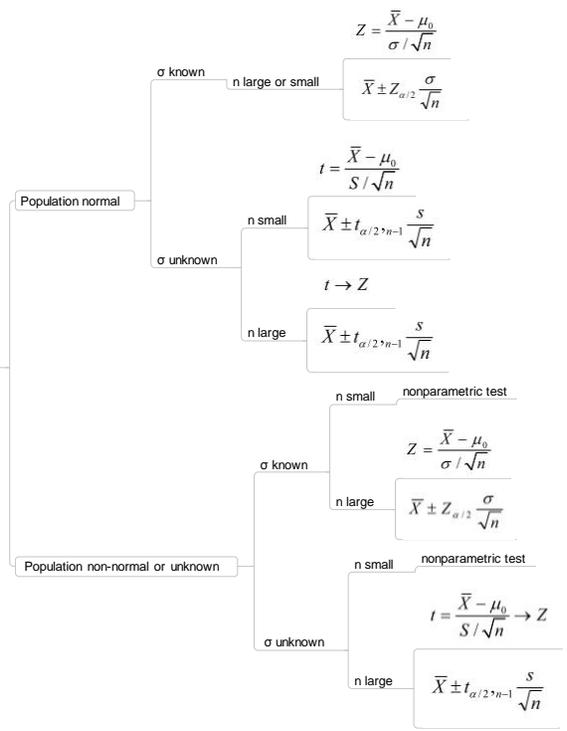
Single Population **Proportion**

np > 5, nq > 5

$$Z = \frac{\hat{p} - p_o}{\sqrt{\frac{p_o q_o}{n}}}$$

$$\hat{p} \pm Z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

Difference between Two Population **Proportions**

$n_1 p_1 > 5, \; n_2 p_2 > 5$

$$Z = \frac{\hat{p}_1 - \hat{p}_2 - 0}{\sqrt{\frac{\overline{p}\,\overline{q}}{n_1} + \frac{\overline{p}\,\overline{q}}{n_2}}}$$

$$\hat{p}_1 - \hat{p}_2 \pm Z_{\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

$$\overline{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

Difference between Two Population **Means**

σ1, σ2 known — n_1, n_2 large or Populations normal; n_1, n_2 small
$$Z = \frac{(\overline{X}_1 - \overline{X}_2) - \mu_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$
$$(\overline{X}_1 - \overline{X}_2) \pm z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

σ1, σ2 unknown

$\sigma_1^2 \neq \sigma_2^2$ — n_1, n_2 large or Populations normal; n_1, n_2 small
$$t = \frac{(\overline{X}_1 - \overline{X}_2) - \mu_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$
$$(\overline{X}_1 - \overline{X}_2) \pm t_{v, \alpha/2} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$
Where v = modified degrees of freedom.

When n_1, n_2 large — $t \to Z$

$\sigma_1^2 = \sigma_2^2 = \sigma^2$ — $n_1, n_2 < 50$
$$t = \frac{(\overline{X}_1 - \overline{X}_2) - \mu_0}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}}$$
$$(\overline{X}_1 - \overline{X}_2) \pm t_{n_1 + n_2 - 2, \alpha/2} \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}$$
Potential benefit of pooling

Decision Map - Making an Inference