# Sample Size Calculation *for the Population Proportion*

**How big a Sample to take?**

# Sample Size Calculation *for the Population Proportion*

- A pollster wanting to make a prediction about a particular candidate’s vote share in the US presidential election.

# Sample Size Calculation *for the Population Proportion*

- A pollster wanting to make a prediction about a particular candidate's vote share in the US presidential election.

**How many voters should the pollster survey?**

# Sample Size Calculation *for the Population Proportion*

- A pollster wanting to make a prediction about a particular candidate's vote share in the US presidential election.

**How many voters should the pollster survey?**

**The pollster in the prediction wants to have a *margin of error* +/- 3% with a *confidence level* of 95%**

RICE
JONES GRADUATE SCHOOL OF BUSINESS

# Sample Size Calculation *for the Population Proportion*

$$\hat{p} - |z_{\alpha/2}|\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \;<\; p \;<\; \hat{p} + |z_{\alpha/2}|\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

# Sample Size Calculation *for the Population Proportion*

$$\hat{p} - \left|z_{\alpha/2}\right|\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < \mathbf{p} < \hat{p} + \left|z_{\alpha/2}\right|\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

**Margin of error = 0.03**

# Sample Size Calculation *for the Population Proportion*

$$\hat{p} - \left|z_{\alpha/2}\right|\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p < \hat{p} + \left|z_{\alpha/2}\right|\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

**Margin of error = 0.03 =** $\left|z_{\alpha/2}\right|\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

# Sample Size Calculation *for the Population Proportion*

$$\hat{p} - |z_{\alpha/2}|\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p < \hat{p} + |z_{\alpha/2}|\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

**Margin of error = 0.03 =** $|z_{\alpha/2}|\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

**0.03 = |NORM.INV(0.05/2, 0, 1)| x** $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

**0.03 = 1.96 x** $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

# Sample Size Calculation *for the Population Proportion*

$$\hat{p} - |z_{\alpha/2}|\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p < \hat{p} + |z_{\alpha/2}|\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

**Margin of error = 0.03 =** $|z_{\alpha/2}|\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

**0.03 = |NORM.INV(0.05/2, 0, 1)| x** $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

**0.03 = 1.96 x** $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

**n =** $\frac{1.96^2}{0.03^2}$ **x** $\hat{p}(1-\hat{p})$

## Sample Size Calculation *for the Population Proportion*

$$\hat{p} - |z_{\alpha/2}|\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p < \hat{p} + |z_{\alpha/2}|\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

**Margin of error = 0.03 =** $|z_{\alpha/2}|\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

**0.03 = |NORM.INV(0.05/2, 0, 1)| x** $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

**0.03 = 1.96 x** $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

**n =** $\frac{1.96^2}{0.03^2}$ **x** $\hat{p}(1-\hat{p})$

What value do we use for $\hat{p}$ ?

## Sample Size Calculation *for the Population Proportion*

**Use a conservative estimate of $\hat{p}$**

# Sample Size Calculation *for the Population Proportion*

**Use a conservative estimate of $\hat{p}$**

$$n = \frac{1.96^2}{0.03^2} \times \hat{p}(1-\hat{p})$$

# Sample Size Calculation *for the Population Proportion*

**Use a conservative estimate of $\hat{p}$**

$$n = \frac{1.96^2}{0.03^2} \times \hat{p}(1-\hat{p})$$

For this to be maximum, $\hat{p}$ has to be 0.5

# Sample Size Calculation *for the Population Proportion*

**Use a conservative estimate of $\hat{p}$**

$$n = \frac{1.96^2}{0.03^2} \times \hat{p}(1-\hat{p})$$

For this to be maximum, $\hat{p}$ has to be 0.5

$$n = \frac{1.96^2}{0.03^2} \times 0.5(1-0.5)$$

$$n = 1067.11 \approx 1068$$

# Sample Size Calculation

- Many industries use rule-of-thumb strategies/heuristics

- We provided here some basis for choosing an appropriate sample size