Sample Proportion Sampling Dist: $\mu_{\hat{p}} = p$ and $\sigma_p = \sqrt{\frac{p(1-p)}{n}}$

Popular Z values:

Confidence	Error Probability	Z
.9	.1	1.65
.95	.05	1.96
.99	.01	2.58

Population Proportion Confidence Interval: $\hat{p} \pm z * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Test Statistic for Proportion Hypothesis Test: $z^* = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$

Hypothesis Test Decisions:

Alternative Hypothesis	Probability	P-Value
$H_a: \boldsymbol{p > p_0}$	Right Tail	P(Z>z*)
$H_a: \boldsymbol{p < p_0}$	Left Tail	P(Z<z*)
$H_a: \boldsymbol{p \neq p_0}$	Two Tail	2*P(Z<-|z*|)

1) John Lester, a Red Sox starting pitcher, throws two types of pitches – strikes and balls. In a random sample of 113 pitches, 73 were strikes.

$n = 113 \qquad \hat{p} = \frac{73}{113} = .6460 \qquad q = .3540$

a. Find a 99% confidence interval for the population proportion of strikes to pitches for John Lester and give a good interpretation of the interval.

• random sample • $n \cdot \hat{p} = 113(.6460) = 72.998 \geq 15$ • $n(1-\hat{p}) = 113(.3540) = 40.002 \geq 15$

$$\hat{p} \pm z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = .646 \pm 2.58\left(\sqrt{\frac{.646(1-.646)}{113}}\right) = .646 \pm .1161$$
$$= (.5299, .7621)$$

We are 99% confident that the true population proportion of strikes John Lester throws is between .5299 and .7621; he throws more strikes than balls.

b. Find a 95% confidence interval for the population proportion of strikes to pitches for John Lester and give a good interpretation of the interval.

The only thing that changes is $z = 1.96$

$$.646 \pm 1.96\left(\sqrt{\frac{.646(1-.646)}{113}}\right) = .646 \pm .0882 = (.5578, .7342)$$

We are 95% confident that the true population proportion of strikes John Lester throws is between .5578 & .7342; he throws more strikes than balls.

c. Name one of the two ways to make the confidence interval from part b narrower.

1) Increase n

2) decrease confidence $\rightarrow$ decrease z

d. Test, with 95% confidence that John Lester throws more strikes than balls. Give a good interpretation of your results. $\Rightarrow p_0 = .5$

i. State Hypothesis: $H_0: p \leq .5$

$H_a: p > .5$

ii. Check Assumptions:

- $np = 113 \cdot .5 = 56.5 \geq 15$
- $n(1-p) = 113 \cdot .5 = 56.5 \geq 15$
- random sample

iii. Calculate Test Statistic

$$z^* = \frac{.646 - .5}{\sqrt{\frac{.5(1-.5)}{113}}} = \frac{.146}{.0470} = 3.1064$$

iv. Find p-value $P(Z > z^*) \approx P(Z > 3.11)$

$$= 1 - P(Z < 3.11)$$
$$= 1 - .9991$$
$$= .0009$$

v. Interpret We reject the null hypothesis in favor of the alternative b/c $\underset{\text{p-value}}{.0009} < 1 - .95 = \underset{\alpha}{.05}$. There is sufficient evidence to suggest that the true population proportion is greater than .5.

e. Are the results the same at 99.99% confidence? Why or why not?

$\underset{\text{p-value}}{.0009} > 1 - .9999 = \underset{\alpha}{.0001}$ So we fail to reject here. The results are not the same

f. Test, with 95% confidence that the proportion of strikes John Lester throws differs from .75. Give a good interpretation of your results.

i. State Hypothesis: $H_0: p = .75$

$H_a: p \neq .75$

ii. Check Assumptions:

- $np = 113(.75) = 84.75 \geq 15$
- $n(1-p) = 113(.25) = 28.25 \geq 15$
- random sample

iii. Calculate Test Statistic

$$z^* = \frac{.646 - .75}{\sqrt{\frac{.75(1-.75)}{113}}} = \frac{-.104}{.040734} = -2.5531$$

iv. Find p-value

$$2P(z < -|z^*|) = 2P(z < -2.5531)$$

$$\approx 2P(z < -2.55)$$

$$= 2(.0054)$$

$$= .0108$$

v. Interpret

We reject the null hypothesis in favor of the alternative b/c $\underset{\text{pvalue}}{\underset{\uparrow}{.0108}} < 1 - .95 = \underset{\alpha}{\underset{\uparrow}{.05}}$. There is sufficent evidence to suggest that the true population proportion is not .75.

2) A random sample of 27 students shows that 18 rated themselves higher than they rated the class. $n = 27$ $\hat{p} = \frac{18}{27} = .\overline{66}$

a) Find a 99% confidence interval for the population proportion of students that rated themselves higher than the rest of the class.

- $n\hat{p} = 27(.\overline{66}) = 18 \geq 15$
- $n(1-\hat{p}) = 27(.\overline{33}) = 9$ ← this is not ≥ 15 so we proceed w/ caution
- random sample

$$\hat{p} \pm z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = .\overline{66} \pm 2.58\sqrt{\frac{.\overline{66}(1-.\overline{66})}{27}}$$
$$= .\overline{66} \pm .23406$$
$$= (.4326, .9007)$$

We are 99% confident that between .4326 and .9007 of the population of students rated themselves higher than the rest of the class.

b) Find a 95% confidence interval for the population proportion of students that rated themselves higher than the rest of the class.

Same as above except $z = 1.96$

$$.\overline{66} \pm 1.96\sqrt{\frac{.\overline{66}(1-.\overline{66})}{27}} = .\overline{66} \pm .17781$$
$$= (.48885, .84447)$$

c) **(7 points)** Name one of the two ways to make the confidence interval from part b narrower.

1) Increase n

2) Decrease confidense → decrease z

d) Test, with 95% confidence that more students rate themselves higher than the rest of the class. Note: You can assume normality in this case, despite n=27<30.

i. State Hypothesis: $H_0: p \le .5$

$H_a: p > .5$

ii. Check Assumptions:

- $np_0 = 27(.5) = 13.5$
- $n(1-p_0) = 27(.5) = 13.5$

We need to proceed w/ caution b/c 13.5 is not ≥ 15

- random sample

iii. Calculate Test Statistic

$$z^* = \frac{.\overline{66} - .5}{\sqrt{\frac{.5(1-.5)}{27}}} = \frac{.1\overline{66}}{.096225} = 1.73205$$

iv. Find p-value

$$P(z > z^*) \approx P(z > 1.73)$$
$$= 1 - P(z < 1.73)$$
$$= 1 - .9582$$
$$= .0418$$

v. Interpret We reject the null hypothesis in favor of the alternative because $\underset{\text{p-value}}{.0418} < 1 - .95 = \underset{\alpha}{.05}$.

There is sufficient evidence to suggest that more students rate themselves higher than the rest of the class.

e) **(6 points)** Are the results the same at 99.99% confidence? Why or why not?

$.0418 > 1 - .9999 = .0001$ So we would fail to reject. The results are different.