

MATH 1342 – STATISTICS

SECTION 8.3: A POPULATION PROPORTION

of trials / Success rate

- Population proportion problem: no mention of a mean or average. Underlying distribution is binomial, $X \sim B(n, p)$.
- When n is large and p is not close to zero or one, can use the normal distribution to approximate the binomial: $X \sim N(np, \sqrt{npq})$. (RECALL: $X \sim N(\text{mean}, \text{std dev})$)
- New random variable called the estimated proportion: $P' = \hat{p} = \frac{X}{n}$ ← random variable for the number of successes / number of trials (or the sample size)
(read “P prime” or “P hat”)
- $\frac{X}{n} = P' \sim N\left(\frac{np}{n}, \frac{\sqrt{npq}}{n}\right)$. This simplifies to $\frac{X}{n} = P' \sim N\left(p, \sqrt{\frac{pq}{n}}\right)$

$$z = \frac{x - \mu}{\sigma}$$

$$z\text{-score formula: } z = \frac{p' - p}{\sqrt{\frac{pq}{n}}}$$

EBP EBP
← p' →

- Sample “mean” $p' = \frac{x}{n}$ is the estimated proportion of successes. (p' is a point estimate for the true proportion p .)

- The confidence interval: $(p' - EBP, p' + EBP)$
- EBP is error bound for the proportion (margin of error).
- $EBP = (z_{\alpha/2}) \left(\sqrt{\frac{p'q'}{n}} \right) = (z_{\alpha/2}) \left(\sqrt{\frac{p'(1-p')}{n}} \right)$

To Find Confidence Interval for a Population Proportion (Pg 460)

- Press [STAT] and arrow over to [TESTS].
- Arrow down to A:1-PropZint. Press [ENTER].
- Arrow down to x and enter the number of successes.
- Arrow down to n and enter the sample size.
- Arrow down to C-Level and enter a decimal.
- Arrow down to Calculate and press [ENTER].
- The confidence interval is displayed.

Ex. 1: Suppose 250 randomly selected people are surveyed to determine if they own a tablet. Of the 250 surveyed, 98 reported owning a tablet. Using a 95% confidence level, compute a confidence interval estimate for the true proportion of people who own tablets. “Success” is own tablet

(a) Using the t-table.

$$n = 250$$

$$x = 98$$

$$p' = \hat{p} = \frac{x}{n} = \frac{98}{250} = .392$$

$$q' = 1 - p' = 1 - .392 = .608$$

$$CL = .95$$

$$\alpha = .05$$

$$\frac{\alpha}{2} = \frac{.05}{2} = .025$$

$$EBP = ME = z_{\alpha/2} \left(\sqrt{\frac{p'q'}{n}} \right)$$

$$= 1.96 \sqrt{\frac{.392(.608)}{250}}$$

$$\approx .0605174876$$

$$z_{.025} = 1.96$$

$$\text{INVNorm}(.025) = +1.96$$

Conf. Int:

$$(.392 - .0605174876, .392 + .0605174876)$$

$$(.33148, .45252)$$

(b) Using the calculator.

A:1PropZInt:

$$x = 98$$

$$n = 250$$

$$CL = .95$$

$$(.33148, .45252)$$

$$p' = .68$$

Ex. 2: In a sample of 300 students, 68% said they own an iPod and a smart phone. Compute a 97% confidence interval for the true percent of students who own an iPod and a smartphone.

$$n = 300$$

$$x = 300(.68) = 204$$

$$p' = .68$$

$$CL = .97$$

$$CL (.62156, .73844)$$

We are 97% confident that
between 62.156% & 73.844% of
Students own an iPod & a
Smartphone.

- “Plus Four” Confidence Interval for p : There is a certain amount of error when calculating a confidence interval for a proportion. Studies have shown that the resulting estimation of the standard deviation can be flawed.
- Simple adjustment to produce more accurate confidence intervals: pretend we have 4 more observations (2 successes and 2 failures).
- $p' = \frac{x+2}{n+4}$ $\leftarrow$ new count of successes
 $\leftarrow$ new sample size
- Computer studies have demonstrated the effectiveness of this method.
- It should be used when the confidence level desired is at least 90% and the sample size is at least ten.
- When using the calculator, don't forget $x = \#$ of successes + 2 and $n = \text{sample size} + 4$

Ex. 3: Out of a random sample of 65 freshmen at State University, 31 students have declared a major. Use the “plus-four” method to find a 96% confidence interval for the true proportion of freshmen at State University who have declared a major.

$$x = 31 \Rightarrow 31 + 2 = 33$$

$$n = 65 \Rightarrow 65 + 4 = 69$$

$$CL = .96$$

1 PropZInt

$$x = 33$$

$$n = 69$$

$$CL = .96$$

Confidence Interval:

$$(.35476, .60177)$$

We are 96% confident that
between 35.476% & 60.177% of
freshmen stds have declared a
major

$$\text{w/ original #'s: } (.34969, .60416)$$

- If researchers desire a specific margin of error, they can use the error bound formula to calculate the required sample size.
- The error bound formula for a population proportion is $EBP = (z_{\alpha/2}) \left(\sqrt{\frac{p'q'}{n}} \right)$
- Solving for n gives you an equation for the sample size: $n = \frac{(z_{\alpha/2})^2(p'q')}{EBP^2}$
- We use invNorm to find the z -score for the appropriate $\alpha/2$

Ex. 4: Suppose an internet marketing company wants to determine the current percentage of customers who click on ads on their smartphones. How many customers should the company survey in order to be 90% confident that the estimated proportion is within five percentage points of the true population proportion of customers who click on ads on their smartphones?

$$n = ?$$

$$EBP = 0.05$$

$$CL = 0.9 \rightarrow \alpha = 1 - CL = 1 - 0.9 = 0.10 \rightarrow z_{\alpha/2} = z_{0.05} = 1.645 \quad p' = ?$$

$$q' = ?$$

NOTE: We want the largest n , so the product $p'q'$ must be as large as possible. Thus, let $p' = 0.5$ and $q' = 0.5$.

$$n = \frac{(z_{\alpha/2})^2(p'q')}{EBP^2}$$

$$n = \frac{(1.645)^2(0.5)(0.5)}{0.05^2} = 270.603$$

Round up: They need to survey 271 customers in order to be 90% confident that the estimated proportion is within five percentage points of the true proportion.

What if we want a smaller error: How many customers should the company survey in order to be 90% confident that the estimated proportion is within three percentage points of the true population proportion of customers who click on ads on their smartphones?

$$n = ?$$

$$EBP = 0.03$$

$$CL = 0.9 \rightarrow \alpha = 1 - CL = 1 - 0.9 = 0.10 \rightarrow z_{\alpha/2} = z_{0.05} = 1.645 \quad p' = ?$$

$$q' = ?$$

NOTE: We want the largest n , so the product $p'q'$ must be as large as possible. Thus, let $p' = 0.5$ and $q' = 0.5$.

$$n = \frac{(z_{\alpha/2})^2(p'q')}{EBP^2}$$

$$n = \frac{(1.645)^2(0.5)(0.5)}{0.03^2} = 751.7$$

Round up: They need to survey 752 customers in order to be 90% confident that the estimated proportion is within three percentage points of the true proportion.

(less error $\rightarrow$ bigger sample)