Inference for Proportions
Inference for a Single Proportion

Objectives (IPS Chapter 8.1)

Inference for a single proportion

- Large-sample confidence interval for $p$
- “Plus four” confidence interval for $p$
- Significance test for a single proportion
- Choosing a sample size
Example

How common is behavior that puts people at risk of AIDS? The National AIDS Behavioral Surveys interviewed a random sample of 2,673 adult heterosexuals. Of these, 170 admitted to having more than one sexual partner in the past year.

(a) What is a 95% confidence interval for \( p \), the true population proportion of all heterosexuals who admit to having multiple partners?

(b) Based on this sample, can we conclude that the true population proportion of all adult heterosexuals who admit to having multiple partners is different than 7.00%?

Sampling distribution of sample proportion

The sampling distribution of a sample proportion \( \hat{p} \) is approximately normal (normal approximation of a binomial distribution) when the sample size is large enough.
Conditions for inference on \( p \)

Assumptions:

1. The data used for the estimate are an SRS from the population studied.
2. The population is at least 10 times as large as the sample used for inference. This ensures that the standard deviation of \( \hat{p} \) is close to \( \sqrt{p(1-p)/n} \).
3. The sample size \( n \) is large enough that the sampling distribution can be approximated with a normal distribution. How large a sample size is required depends in part on the value of \( p \) and the test conducted. Otherwise, rely on the binomial distribution.

Large-sample confidence interval for \( p \)

Confidence intervals contain the population proportion \( p \) in \( C\% \) of samples. For an SRS of size \( n \) drawn from a large population, and with sample proportion \( \hat{p} \) calculated from the data, an approximate level \( C \) confidence interval for \( p \) is:

\[
\hat{p} \pm m, \ m \text{ is the margin of error }
\]

\[
m = z \times SE = z \times \sqrt{\hat{p}(1-\hat{p})/n}
\]

Use this method when the number of successes and the number of failures are both at least 15.

\( C \) is the area under the standard normal curve between \(-z^*\) and \( z^*\).
**Example**

How common is behavior that puts people at risk of AIDS? The National AIDS Behavioral Surveys interviewed a random sample of 2,673 adult heterosexuals. Of these, 170 admitted to having more than one sexual partner in the past year. What is a 90% confidence interval for \( p \), the true population proportion of all heterosexuals who admit to having multiple partners?

\[
\hat{p} = \frac{X}{n} = \frac{170}{2673} = 0.0636 \text{ or } 6.36\
\]

\[
m = z \times SE_{\hat{p}} = z \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}
\]

\[
= 1.645 \times \sqrt{\frac{0.0636(1-0.0636)}{2673}}
\]

\[
= 0.007765
\]

\[
\hat{p} \pm m = 0.0636 \pm 0.007765
\]

\[
= (0.0636 - 0.007765, 0.0636 + 0.007765)
\]

\[
= (0.055835, 0.071365) \approx (5.56\%, 7.13\%)
\]
Because we have to use an estimate of $p$ to compute the margin of error, confidence intervals for a population proportion are not very accurate.

$$m = z^* \sqrt{\frac{p(1 - p)}{n}}$$

Specifically, we tend to be incorrect more often than the confidence level would indicate. But there is no systematic amount (because it depends on $p$).

**Use with caution!**

### “Plus four” confidence interval for $p$

A simple adjustment produces more accurate confidence intervals. We act as if we had four additional observations, two being successes and two being failures. Thus, the new sample size is $n + 4$, and the count of successes is $X + 2$.

The “plus four” estimate of $p$ is: $\tilde{p} = \frac{\text{counts of successes} + 2}{\text{count of all observations} + 4}$

And an approximate level $C$ confidence interval is:

$$CI : \quad \tilde{p} \pm m, \text{ with }$$

$$m = z^* SE = z^* \sqrt{\frac{p(1 - p)}{(n + 4)}}$$

Use this method when $C$ is at least 90% and sample size is at least 10.
Example

**Instant versus fresh-brewed coffee.** A matched pairs experiment compares the taste of instant versus fresh-brewed coffee. Each subject tastes two unmarked cups of coffee, one of each type, in random order and states which he or she prefers. Of the 40 subjects who participate in the study, 12 prefer the instant coffee. Let \( p \) be the probability that a randomly chosen subject prefers fresh-brewed coffee to instant coffee. (In practical terms, \( p \) is the proportion of the population who prefer fresh-brewed coffee.)

What is a 95% confidence interval for the population proportion \( p \) who prefer **instant** coffee?

\[
\hat{p} = \frac{X + 2}{n + 4} = \frac{12 + 2}{40 + 4} = 0.3182 \quad \text{or} \quad 31.82\%
\]

\[
m = z\ast SE_{\hat{p}} = z\ast \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = 1.96 \times \sqrt{\frac{0.3182(1 - 0.3182)}{44}} = 0.13763
\]

\[
\hat{p} \pm m = 0.3182 \pm 0.13763 = (0.18057, 0.4558) \approx (18.1\%, 45.6\%)
\]
Sample size for a desired margin of error

You may need to choose a sample size large enough to achieve a specified margin of error. However, because the sampling distribution of \( \hat{p} \) is a function of the population proportion \( p \), this process requires that you guess a likely value for \( p: p^* \).

\[
p \sim N\left(p, \sqrt{p(1-p)/n}\right) \Rightarrow n = \left( \frac{z^*}{m} \right)^2 p^*(1-p^*)
\]

The margin of error will be less than or equal to \( m \) if \( p^* \) is chosen to be 0.5.

Remember, though, that sample size is not always stretchable at will. There are typically costs and constraints associated with large samples.

---

**SAMPLE SIZE FOR DESIRED MARGIN OF ERROR**

The level \( C \) confidence interval for a proportion \( p \) will have a margin of error approximately equal to a specified value \( m \) when the sample size satisfies

\[
n = \left( \frac{z^*}{m} \right)^2 p^*(1-p^*)
\]

Here \( z^* \) is the critical value for confidence \( C \), and \( p^* \) is a guessed value for the proportion of successes in the future sample.

The margin of error will be less than or equal to \( m \) if \( p^* \) is chosen to be 0.5. The sample size required when \( p^* = 0.5 \) is

\[
n = \frac{1}{4} \left( \frac{z^*}{m} \right)^2
\]
Example

Are the customers dissatisfied?
An automobile manufacturer would like to know what proportion of its customers are dissatisfied with the service received from their local dealer. The customer relations department will survey a random sample of customers and compute a 95% confidence interval for the proportion that are dissatisfied.

Find the sample size needed if the margin of error of the confidence interval is to be no more than 0.02.

\[
n = \frac{1}{4} \left( \frac{z^*}{m} \right)^2
\]

\[
= 0.25 \left( \frac{1.96}{0.02} \right)^2
\]

\[
= 0.25(98)^2
\]

\[
= 0.25(9604)
\]

\[
= 2401
\]

Example Cont.

From past studies, the dealer believes that the proportion of dissatisfied customers is about 0.15. Find the sample size needed if the margin of error of the confidence interval is to be no more than 0.02.

\[
n = \left( \frac{z^*}{m} \right)^2 \cdot p^* (1 - p^*)
\]

\[
= \left( \frac{1.96}{0.02} \right)^2 \times 0.15(1 - 0.15)
\]

\[
= (98)^2 \times 0.1275
\]

\[
= (9604) \times 0.1275
\]

\[
= 1224.51 \text{ or } 1225
\]
Significance test for $p$

The sampling distribution for $\hat{p}$ is approximately normal for large sample sizes and its shape depends solely on $p$ and $n$.

Thus, we can easily test the null hypothesis:

$H_0: p = p_0$ (a given value we are testing).

If $H_0$ is true, the sampling distribution is known →

The likelihood of our sample proportion given the null hypothesis depends on how far from $p_0$ our $\hat{p}$ is in units of standard deviation.

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

This is valid when both expected counts—expected successes $np_0$ and expected failures $n(1-p_0)$—are each 10 or larger.
Example

How common is behavior that puts people at risk of AIDS? The National AIDS Behavioral Surveys interviewed a random sample of 2,673 adult heterosexuals. Of these, 170 admitted to having more than one sexual partner in the past year. That's 6.36% of the sample. Based on these data, what can we say about the percent of all adult heterosexuals who admit to having multiple partners?

(a) Based on this sample, can we conclude that the true population proportion of all adult heterosexuals who admit to having multiple partners is different than 7.00%.

Example – Cont.

1. Set up the test
   - $H_0: p = 7\%$
   - $H_a: p \neq 7\%$

2. Compute the value of the Test Statistic

   \[
   z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \\
   = \frac{0.0636 - 0.07}{\sqrt{\frac{0.07(1-0.07)}{2673}}} \\
   = \frac{-0.0064}{0.00494} = -1.297 \text{ or } -1.30
   \]

3. Compute the P-Value

   \[
   P-Value = 0.1936
   \]

   \[
   P-Value = 2*P[Z > 1.30] = 2*\text{normalcdf}(1.30, 100, 0, 1) = 2*(0.0968) = 0.1936
   \]

Some large number far away from the mean of the curve 0.
Example – Cont.

(b) At the 5% level, can we reject the null hypothesis?

4. **Answer the question:**

   "Based on this sample, can we conclude that the true population proportion of all adult heterosexuals who admit to having multiple partners is different than 7.00%.

Since the P-value (19.36%) is not less than \( \alpha \) (5%), we cannot reject the null hypothesis.

At the 5% level, based on this sample, there is not enough evidence to conclude that the true population proportion of all adult heterosexuals who admit to having multiple partners is different than 7.00%.

A difference as large (or larger) between the hypothesized population proportion value (i.e., 7.00%) and the one we observed in the sample (i.e., 6.36%) is expected approximately in 1 out of every 5 samples of size 2673.

---

**P-values and one or two sided hypotheses—reminder**

\[ H_0: p > p_0 \text{ is } P(Z \geq z) \]

\[ H_0: p < p_0 \text{ is } P(Z \leq z) \]

\[ H_0: p \neq p_0 \text{ is } 2P(Z \geq |z|) \]

And as always, if the p-value is as small or smaller than the significance level \( \alpha \), then the difference is statistically significant and we reject \( H_0 \).
Example

A national survey by the National Institute for Occupational Safety and Health on restaurant employees found that 75% said that work stress had a negative impact on their personal lives.

You investigate a restaurant chain to see if the proportion of all their employees negatively affected by work stress differs from the national proportion \( p_0 = 0.75 \).

\[ H_0: p = p_0 = 0.75 \text{ vs. } H_a: p \neq 0.75 \] (2 sided alternative)

In your SRS of 100 employees, you find that 68 answered “Yes” when asked, “Does work stress have a negative impact on your personal life?”

The expected counts are \( 100 \times 0.75 = 75 \) and 25. Both are greater than 10, so we can use the z-test.

\[ z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}} \]

\( \hat{p} = \frac{68}{100} = 0.68 \) and \( n = 100 \)

\[ z = \frac{0.68 - 0.75}{\sqrt{\frac{0.75(0.25)}{100}}} = 1.62 \]

From Table A we find the area to the left of \( z = 1.62 \) is 0.9474.

Thus \( P(Z \geq 1.62) = 1 - 0.9474 \), or 0.0526. Since the alternative hypothesis is two-sided, the \( P \)-value is the area in both tails, and \( P = 2 \times 0.0526 = 0.1052 \).

\( \hat{p} \) is not significantly different from the national survey results \( (\hat{p} = 0.68, z = 1.62, P = 0.11) \).
Computing a CI for \( p \) Using the TI-83

8.1 Inference for a Single Proportion

Both the large-sample and plus-four level \( C \) confidence intervals can be calculated using the \texttt{1-PropZInt} feature from the \texttt{STAT TESTS} menu. Significance tests can be worked using the \texttt{1-PropZTest}.

A Large-Sample Confidence Interval

Example 8.1 Stress in Restaurant Workers. In a restaurant worker survey, 68 of a sample of 100 employees agreed that work stress had a negative impact on their personal lives. Find a 95\% confidence interval for the true proportion of restaurant employees who agree.

Computing a CI for \( p \) Using the TI-83

Solution. Bring up the \texttt{1-PropZInt} screen, enter 68 for \( x \), enter 100 for \( n \), and enter \( .95 \) for \texttt{C-Level}. Then press \texttt{ENTER} on \texttt{Calculate} to obtain a 95\% confidence interval of \((0.58857, 0.77143)\). Based on this sample, with 95\% confidence, between 58.9\% and 77.1\% of restaurant workers will agree that work stress has a negative impact on their personal lives.
Computing a CI for p Using the TI-83

A Plus-Four Confidence Interval

Example 8.2 Blind Medical Trials. Many medical trials randomly assign patients to either an active treatment or a placebo. The trials are supposed to be double-blind, but sometimes patients can tell whether or not they are getting the active treatment. Reports of medical research usually ignore the problem. Investigators looked at a random sample of 97 articles and found that only 7 discussed their success in blinding the trial. What proportion of all such studies discuss the success of blinding in their trials? Give a 95% confidence interval estimate.

Because we have a small number of studies that discussed the success of blinding in the trial (less than 10), we use the “plus-four” method: we simply add 4 to the number of studies (it becomes 101) and 2 to the number of “successes” (we’ll consider that 9 discussed the success in blinding the trial). We use 1-PropZInt as before, with the adjusted counts.

Computing a CI for p Using the TI-83

Solution. In the 1-PropZInt screen, enter 9 for x, which is 2 more than the actual number and enter 101 for n, which is 4 more than the actual sample size. Enter the desired C-level (95%) and press ENTER on Calculate to obtain the plus-four estimate \( \hat{p} = 0.089 \) and the confidence interval.

Based on this sample, we are 95% confident that between 3.4% and 14.5% of published medical studies will discuss the success of blinding in their trials.
Testing Hypotheses for $p$ with the TI-83

Example 8.4  More Stress in Restaurant Workers. In the restaurant worker survey, 68 of a sample of 100 employees agreed that work stress had a negative impact on their personal lives. Let $p$ be the true proportion of restaurant employees who agree. Test the hypothesis $H_0: p = 0.75$ versus $H_A: p \neq 0.75$.

Solution. Enter the data and alternative into the 1-PropZTest screen, then press ENTER on Calculate or Draw. We obtain a (two-sided) $p$-value of 0.106 from a $z$-statistic of $-1.61658$. If $p$ were equal to 0.75, then there would be a 10.6% chance of obtaining $\hat{p}$ as far away as 0.68 with a sample of size 100. We therefore believe the proportion of restaurant workers who would agree that work stress has a negative impact on their personal lives is not significantly different from 75%.

Testing Hypotheses for $p$ with the TI-83

Example 8.5 Who Likes Instant? In a taste test of instant versus fresh-brewed coffee, only 12 out of 40 subjects preferred the instant coffee. Let $p$ be the true probability that a random person prefers the instant coffee. Test the claim $H_0: p = 0.50$ versus $H_A: p < 0.50$ at the 5% level of significance.

Solution. Enter the data and alternative into the 1-PropZTest screen and calculate. We obtain a test statistic of $-2.53$ and a $p$-value of 0.0057. If $p$ were 0.50, then there would be only a 0.0057 probability of $\hat{p}$ being as low as 0.3 with 40 subjects. There is strong evidence to reject $H_0$, and conclude that those who prefer instant coffee are a minority of the population.
Problem to work in class

**Instant versus fresh-brewed coffee.** A matched pairs experiment compares the taste of instant versus fresh-brewed coffee. Each subject tastes two unmarked cups of coffee, one of each type, in random order and states which he or she prefers. Of the 40 subjects who participate in the study, 12 prefer the instant coffee. Let \( p \) be the probability that a randomly chosen subject prefers fresh-brewed coffee to instant coffee. (In practical terms, \( p \) is the proportion of the population who prefer fresh-brewed coffee.)

(a) Test the claim that a majority of people prefer the taste of fresh-brewed coffee. Report the large-sample \( z \) statistic and its \( P \)-value.

(b) Draw a sketch of a standard Normal curve and mark the location of your \( z \) statistic. Shade the appropriate area that corresponds to the \( P \)-value.

(c) Is your result significant at the 5% level?