

What do I know already? (Do not look up the answers to these questions. The purpose is to assess your current level of knowledge on these topics.)

A. What does the standard error for the sample proportion tell us?

It tells us how the sample proportion varies from sample to sample.

B. A physician claims that average height for American males is now greater than 69 in. Express this claim symbolically.

$\mu > 69$ inches

C. In a hypothesis test, which hypothesis is initially assumed to be true?

The null hypothesis is initially assumed to be true.

D. Compare the hypothesis $\mu < \$25$ and the following 99% confidence interval which was constructed using a sample of data from the population in question:

$\$42.12 < \mu < \52.34 . Does the confidence interval provide support for the hypothesis?

No, the interval says the mean appears to be between \$42.12 and \$52.34. This range of values is higher than \$25, not less than \$25.

Learning Objectives: (Click the learning objectives below for a short clip on the topic.)

Calculate a Sample Proportion Using Summary Data (1)

Discuss the Standard Error of the Sample Proportion (2)

Review the Sample Size Guidelines for Estimating a Population Proportion (2)

Construct a Confidence Interval for the Population Proportion (3)

Interpret a Confidence Interval for the Population Proportion (3)

Know the Meaning of "Hypothesis" in the Context of Statistics (1)

List the Characteristics of the Null Hypothesis (1)

List the Characteristics of the Alternative Hypothesis (1)

Determine the Null and Alternative Hypothesis (2)

Explain the Logic of the Test Statistic (2)

Calculate the Test Statistics from Sample Data (3)

Exercises:

1. The survey of more than 4500 randomly selected adults was done by the Pew Research Center about attitudes toward political posts in social media. It found that 37% of social media users report feeling worn out by political posts. Use the sample data to construct a 90% confidence interval for the true proportion of people that feel worn out by political posts on social media.

$$\left(0.37 - 1.645 \sqrt{\frac{0.37 * 0.63}{4500}}, 0.37 + 1.645 \sqrt{\frac{0.37 * 0.63}{4500}} \right) = (0.358, 0.382)$$

We are 90% confident the true proportion is between 0.358 and 0.382.

2. Which of the following statements about the null hypothesis are valid? (select all that apply)

A. The null hypothesis is initially assumed to be true. ✓

B. The null hypothesis always contains the condition of equality (it uses one of these symbols: $\leq, \geq, =$). ✓

C. The null hypothesis is always the original claim in hypothesis testing. **Either hypothesis can be the original claim.**

D. In a hypothesis test, there has to be sufficient evidence to support the null hypothesis or it will be rejected. **The null is assumed to be true. It does not require support.**

3. A survey of a random sample of 955 adults over the age of 30 was conducted. Among those surveyed, 121 reported that they believed good luck had played a major role in their lives. Use the results of the survey to calculate the sample proportion for the number of adults over age 30 that believe good luck had

played a major role in their lives. $\hat{p} = \frac{x}{n} = \frac{121}{955} \approx 0.127$

4. Researchers want to determine the amount of time people in the US spend viewing media each day. A study of a random sample of 100 US residents found the average amount of time spent viewing media per day is 490 minutes with a standard deviation of 75 minutes. Use the sample of data to construct a 90% confidence interval to estimate the true mean amount of time people in the US spend viewing media each day.

$$\left(490 - 1.645 \frac{75}{\sqrt{100}}, 490 + 1.645 \frac{75}{\sqrt{100}} \right) = (477.7, 502.3)$$

We are 90% confident the true mean is between 477.7 and 502.3.

5. **True** or false: The test statistic for a hypothesis test about a population mean measures how many standard errors apart the sample mean is from the hypothesized mean. If the hypothesized mean is correct, the sample mean will typically be near the hypothesized mean.
6. Another finding from the survey of more than 4500 adults about attitudes toward political posts in social media was that 83% of Americans try to avoid posts from friends they disagree with politically. Use the sample data to construct a 95% confidence interval for the true proportion of Americans that try to avoid posts from friends they disagree with politically.

$$\left(0.83 - 1.960 \sqrt{\frac{0.83 * 0.17}{4500}}, 0.83 + 1.960 \sqrt{\frac{0.83 * 0.17}{4500}} \right) = (0.819, 0.841)$$

We are 95% confident the true proportion is between 0.819 and 0.841.

7. A survey of 500 male US soccer fans revealed that just 4% of them would watch the Women's World Cup matches if the US was not competing in the tournament. Is the sample size large enough to use the normal approximation? **Yes, both $n * p$ and $n * q$ are larger than 5. $500 * 0.04 = 20$.**
8. A 95% confidence interval for the true mean amount of money spent by Amazon Prime members per year was constructed. Interpret the interval: $\$2584 \pm \487 . Does the interval contradict the claim that the true mean amount of money spent by Amazon Prime members per year is greater than $\$2750$? **Since the interval ranges from $\$2097$ to $\$3071$, we believe the mean amount spent by Prime members the interval does not contradict the claim. The interval leaves open the possibility that the true mean is above $\$2750$.**
9. In an effort to determine the average amount of bulk candy purchased by customers at a local sweet shop, the owner records the weight in ounces for a random sample of 28 bulk candy purchases. These weights had an average of 10.2 ounces and a standard deviation of 3.5 ounces. Use the sample data to form a 98% confidence interval for the true mean weight of bulk candy purchased by customers visiting the sweet shop.

$$\left(10.2 - 2.473 \frac{3.5}{\sqrt{28}}, 10.2 + 2.473 \frac{3.5}{\sqrt{28}} \right) = (8.56, 11.84)$$

We are 98% confident the true mean is between 8.56 oz and 11.84 oz.

10. The head of a bank that provides auto loans claims that US households have an average amount of auto loan debt equal to \$29,000. A random sample of 100 US households had an average amount of auto loan debt of \$27,865 and a standard deviation of \$8,300. Calculate the test statistic for a test of this hypothesis.

$$z = \frac{27,865 - 29,000}{\frac{8,300}{\sqrt{100}}} \approx -1.37$$

11. A confidence interval for the true proportion of readers that buy more than two books per month was created from a large sample of readers. Interpret the interval: 0.17 ± 0.04 . If the CEO of Barnes & Noble claims that only 20% of readers buy more than two books per month, does the provided interval contradict his claim?

We believe the proportion of readers that buy more than two books per month is between 0.13 and 0.21. The interval leaves open the possibility that the true proportion is 0.20.

12. Which of the following statements about the alternative hypothesis are valid? (select all that apply)

A. The alternative hypothesis typically represents the status quo belief.

B. The alternative hypothesis always uses one of the following symbols: $<$, $>$, \neq .

C. If the null hypothesis is rejected, it is the alternative hypothesis that is supported.

D. In a hypothesis test, the null hypothesis is assumed true, so there needs to be sufficient evidence to support the alternative hypothesis. In other words, the burden of proof lies with the alternative hypothesis.

13. A pediatrician claims that the average age when children become potty trained is 30 months. State the null and alternative hypothesis that would be used to test the pediatrician's claim.

$$H_0 : \mu = 30$$

$$H_A : \mu \neq 30$$

14. Another pediatrician claims the average age girls become potty trained is less than 30 months. State the null and alternative hypothesis that would be used to test this pediatrician's claim.

$$H_0 : \mu \geq 30$$

$$H_A : \mu < 30$$

15. Market researchers want to know how much money households spend on cable TV in the US. From a previous study, the standard deviation for these spending amounts appears to be \$47. If the researchers want to create a 99% confidence interval to estimate the true mean to within \$5, on how many households must be randomly sampled?

$$n = \left[\frac{2.576 * 47}{5} \right]^2 = 586.3371... \approx 587 \quad \text{round up to 587}$$

16. An economist claims that the average US mortgage is less than \$180,000. A random sample of 40 US mortgages had an average value of \$172,341 and a standard deviation of \$106,312. Calculate the test statistic for a test of this hypothesis.

$$z = \frac{172,341 - 180,000}{\frac{106,312}{\sqrt{40}}} \approx -0.46$$

17. A survey of a random sample of 955 adults over the age of 30 was conducted. Among those surveyed, 121 reported that they believed good luck had played a major role in their lives. Use the results of the survey to **calculate the standard error** for the sample proportion for the number of adults over age 30 that believe good luck had played a major role in their lives.

$$\hat{p} = \frac{121}{955} \approx 0.1267015707, \quad \text{the standard error} = \sigma_{\hat{p}} = \sqrt{\frac{(0.1267)(0.8733)}{955}} \approx 0.0108$$

18. Most experts think that the ideal length of a pop song is 180 seconds (3 minutes). A 95% confidence interval for the true mean length of popular songs sold at iTunes was constructed. Interpret the interval: 225.1 seconds \pm 5.4 seconds. Based on this interval, does it appear that the true mean length of popular songs sold at iTunes is 180 seconds (3 minutes long)?

Since the interval ranges from 219.7 to 230.5, we believe the mean length of popular songs is between 219.7 seconds and 230.5 seconds. The interval contradicts the idea that the true mean song length for popular songs is 180 seconds.

19. **True** or false: A hypothesis in Statistics refers to a statement about a population parameter that can be verified or tested using sample evidence.