MAT1372, Classwork18, Fall2025

5.3 Hypothesis Testing of a Proportion (Conti.)

6. ***Method1***: Testing Hypotheses Using Confidence Intervals.

If, from a survey of 50 adults, 24% of respondents got the question correct about 1 year-old vaccinated status, then (a) does this data provide strong evidence that the proportion of all the adults who would answer this question correctly is different than 33.3%?

Not sure. Since we don't know that if this deviation of 24% from 33.3% is simply due to chance, or the data provide strong evidence that the population proportion is different from 33%

(b) Check whether it is reasonable to construct a confidence interval for p using the sample data, and if so, construct a 95% confidence interval.

sample proportion $\hat{p} = 24\%$, $n = 50$, Normal?

Central Limit Thm: ① the data come from a simple random sample (satisfies independence) and ② check success-failure condition: $n\hat{p} = 50 \cdot 24\% = 12 \geq 10$, $n \cdot (1-\hat{p}) = 50 \cdot 76\% = 38 \geq 10$,

Based on ①②, the confidence interval for p can be constructed

$SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.06$ $\qquad \hat{p} \pm z^{\star} \times SE =$ $0.24 \pm 1.96 \times 0.06 \Rightarrow (0.122, 0.358)$

(c) Is this 95% confidence interval of p able to reject null hypothesis?

NO, Because null value $p_0 = \frac{1}{3}$ from the hypothesis test falls within the range of plausible value from the confidence interval.

(d) Explain why we cannot conclude that the adults simply guessed on the infant vaccination question.

While we failed to reject H_0, that does not necessarily mean the null hypothesis is true.

7. Double Negative Can Sometimes Be Used in Statistics.

In many statistical explanations, we use double negative. For example, we might say H_0 is not implausible or we failed to reject H_0.

8. Decision Error.

In a hypothesis test, there are two competing hypotheses: the null (H_0) and alternative (H_A), and we make a statement about which one migh be true, but we might choose incorrectly.

There are four possible scenarios, which are summarized in Figure

		Test conclusion	
		do not reject H_0	reject H_0 in favor of H_A
Truth	H_0 true	okay	**Type 1 Error**
	H_A true	**Type 2 Error**	okay

A Type 1 Error is rejecting H_0 when H_0 is actually true

A Type 2 Error is failing to reject H_0 when H_0 is false (or H_A is true)

11. ***Method2***: Testing Hypotheses Using Significance Level.

Researcher asked a random sample of $\boxed{1000}$ American adults whether they supported the increased usage of coal to produce energy.

n

(a) Set up hypotheses to evaluate whether a majority of American support or oppose this idea.

H_0: p = 0.5 (no majority either way: half support and the other half oppose it)

H_A: p ≠ 0.5 (there is a majority support or oppose (though we don't know which one)

In this case, the null value $\boxed{p_0 = 0.5}$.

(b) What would the sampling distribution of $\hat{p}$ look like if the null hypothesis were true? $\hat{p} = p_0 = \frac{1}{2}$; $n = 1000$

If the null hypothesis were true, the population proportion would be the null value, 0.5. We previously learned that the sampling distribution of $\hat{p}$ will be normal when two conditions are met:

Independence. The poll was based on simple random sample, so independence is satisfied

Success-failure. Based on the poll sample size n = 1000, this condition is met, since

$np =$ $np_0 = 1000 \cdot \frac{1}{2} = 500 \geq 10$; $n(1-p) =$ $n(1-p_0) = 1000 \cdot \frac{1}{2} = 500 \geq 10$

The first procedural difference from confidence intervals: the condition checking by using $p_0 = \frac{1}{2}$

Since the condition is satisfied, this sample proportion would be normally distributed.

Next, we can compute the standard error by using $p_0 = \frac{1}{2}$ in the calculation:

$$SE_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} \overset{H_0}{=} \sqrt{\frac{p_0(1-p_0)}{1000}} = 0.016$$

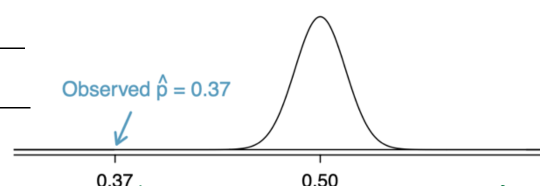
The other procedural difference from confidence intervals: the calculation using p_0, not $\hat{p}$.

When we identify the sampling distribution under H_0, it has a special name: null distribution.

(c) If 37% of respondents American adults support increased usage of coal, does 37% represent a real difference from the null hypothesis of 50%? $\hat{p} = 0.37$

If the null hypothesis were true, we can determine the chance of finding $\hat{p}$ at least as far into the tails as 0.37 under the null distribution, which is a normal distribution with mean μ = 0.5 and SE = 0.016.

First, we draw a null distribution to represent the situation



Second, to find the area below 0.37, we compute the Z-score using $\mu = \frac{1}{2}$, $\sigma = 0.016$

$$Z = \frac{x - \mu}{\sigma} = \frac{0.37 - 0.5}{0.016} = -8.125$$

and the area below 0.37 is $P(Z < -8.125) = 0.0002$ (in fact, it's way smaller than 0.0002).

(d) How do we reject or support $H_0: p = 0.5$ based this information (i.e. area below 0.37)?

By Significance Level in Hypothesis Testing and p-value.