Chapter 20/21 Notes
Testing Hypotheses about Proportions

Flipper bands have been used for decades to identify individual penguins so they can be tracked on land and sea. Studies have shown that only 20% ($p = 0.2$) of penguins tagged with these bands survive past 10 years. French researchers random selected 50 king penguins (pop. size is 100,000) and had tiny transponders implanted under their skin. After a decade of observation, 18 of the 50 (36%) penguins were alive.

We would like to determine whether the transponders actually increased the proportion of penguins surviving or if the observed proportion is due to sampling variability.

- If no change has occurred:

- If a change has occurred:

We need to be able to answer the question: *Are we likely to get a sample proportion of 36% or more if no change has occurred (we expect 20%)?*

If no change has occurred then:

- $p =$

- The sampling distribution for $\hat{p}$

But Wait! We need to check the conditions necessary to use the normal model before proceeding with analysis.

Now that we have verified the conditions, we can use the sampling distribution to determine *how likely it is to observe a sample proportion of 0.36 or more if there has been no change (i.e. $p = 0.2$)?*

Do you think the proportion of penguins surviving, p, has increased when using transponders?

Would it be unethical to still use Flipper tagging?

Parts of a Hypothesis Test: Null Hypothesis

The **null hypothesis**, denoted H_0, proposes a value for the population proportion (parameter).

Notation:

The null hypothesis is used to specify the sampling distribution for $\hat{p}$.

Parts of a Hypothesis Test: Alternative Hypothesis

The **alternative hypothesis**, denoted H_A, specifies parameter values that we consider plausible when we *reject* the null hypothesis.

Notation: There are three possible alternative hypotheses

1.

2.

3.

Parts of a Hypothesis Test: Assumptions

In order to use the normal model for the sampling distribution of $\hat{p}$, we need to check three conditions (these should look familiar).

1. **randomization condition**

2. **10% condition**

3. **success/failure condition**

Parts of a Hypothesis Test: Sampling Distribution

The sampling distribution of $\hat{p}$ *assuming the null hypothesis is true*

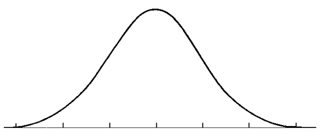
Parts of a Hypothesis Test: Test Statistic

We want to compare the observed data to what we would expect if the null hypothesis is true. We can do this by finding out how many standard deviations away from the proposed value the observed data falls.

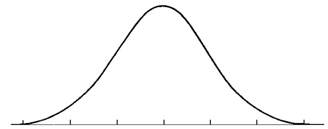
Parts of a Hypothesis Test: p-value

We now ask ourselves how likely it is to obtain the results that we have observed or more extreme results **if the null hypothesis is true.**
More extreme means

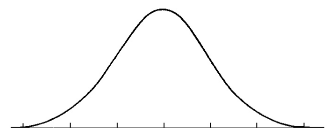
1. If $H_A : p < p_0$, p-value $= \Pr(Z < z)$



2. If $H_A : p > p_0$, p-value $= \Pr(Z > z)$



3. If $H_A : p \neq p_0$, p-value $= 2 \times \Pr(Z > |z|)$



Parts of a Hypothesis Test: Decision

Recall: the p-value is the probability of the observed test statistic *given* that the null hypothesis is **true**.

When the p-value is *small*

When the p-value is *large*

But what is the difference between small and large p-values?

The **α-level** (a.k.a. the significance level) of a hypothesis test defines a *rare event* by setting a threshold for the p-value. Thus, we have the rejection rule

- if $\mathrm{p-value} < \alpha$

- if $\mathrm{p-value} \geq \alpha$

Parts of a Hypothesis Test: Conclusion

- The **conclusion** in a hypothesis test is a statement about the null hypothesis stating either that we reject or fail to reject the null hypothesis.

- THE CONCLUSION SHOULD ALWAYS BE STATED IN CONTEXT

A Trial as a Hypothesis Test Consider the logic of a jury trial.

- To prove someone is guilty, we start by assuming they are innocent.
- Only when the evidence makes innocence unlikely beyond a reasonable doubt can we reject the hypothesis of innocence, and declare a person guilty.

What would a p-value be?

- If the defendant is innocent, how likely is it that we would see the evidence?

Example 1

The survival rate for King penguins tagged with flipper bands is 20% after 10 years. Tiny transponders implanted under the skin are designed to increase this survival rate. 50 randomly selected 50 king penguins (pop. size is 100,000) and had tiny transponders implanted under their skin. After a decade of observation, 18 of the 50 (36%) penguins were alive. Did the transponders *increase* the survival rate? Use $\alpha = 0.01$.

Step 1: State the null and alternative hypotheses.

Step 2: Check the assumptions necessary for the use of the normal model.

Step 3: Identify the sampling distribution for $\hat{p}$ if the null hypothesis is true.

Step 4: Calculate the test statistic.

Step 5: Obtain the p-value.

Step 6: Reach a decision.

Step 7: State your conclusion in the context of the problem.

Example 2

The article Statistical Evidence of Discrimination (JASA, 1982: 773-783) discusses the court case Swain vs. Alabama (1965), in which it was alleged that there was discrimination against blacks in grand jury selection. Census data suggested that 25% of those eligible for jury service were black, yet a random sample of 1050 called to appear for possible jury duty yielded only 177 blacks. Do these data argue strongly for a conclusion of discrimination? Use $\alpha = 0.05$.

Step 1: State the null and alternative hypotheses.

Step 2: Check the assumptions necessary for the use of the normal model.

Step 3: Identify the sampling distribution for $\hat{p}$ if the null hypothesis is true.

Step 4: Calculate the test statistic.

Step 5: Obtain the p-value.

Step 6: Reach a decision.

Step 7: State your conclusion in the context of the problem.

Example 3 An advertisement for a headache remedy claims that 90% or more of headache sufferers get relief if they use the remedy. A truth in advertising agency is considering a suit for false advertising and obtains a sample of 100 individuals, which shows that 88 indicate that the remedy gave them relief. Can the suit be justified using $\alpha = 0.07$.

Step 1: State the null and alternative hypotheses.

Step 2: Check the assumptions necessary for the use of the normal model.

Step 3: Identify the sampling distribution for $\hat{p}$ if the null hypothesis is true.

Step 4: Calculate the test statistic.

Step 5: Obtain the p-value.

Step 6: Reach a decision.

Step 7: State your conclusion in the context of the problem.

Statistical Significance vs. Practical Significance

A test is **statistically significant** if the p-value is lower than the α-level. This does not mean that statistical significance carries with it any sense of practical importance/impact.

Confidence Intervals and Hypothesis Tests

Confidence intervals and hypothesis tests are built from the same calculations and require the same assumptions, so we can approximate a hypothesis test by examining a confidence interval.

Making Errors

There are two ways to make errors in hypothesis testing