STA-1380 – Elementary Statistics
Week #9

Hello! Welcome to the additional online **Weekly Resources** for the course of **STA-1380**. Following a traditional calendar semester, these will be some of the topics your professors will go over. If you do not see material your section is going over for the week, please look at the other resources listed for this course. In addition to these resources, there might be **Group Tutoring** for this course, please see our website for more details. These sessions will go over these materials in more detail as well as any questions about the material.

Any additional help or services can be found through the **Baylor Tutoring Website**. Visit to schedule a free 30-minute private tutoring session, drop-in times for your course, the Baylor Tutoring YouTube channel, or any additional tutoring resources.

Contacts: Sid Rich M-Th 9am-8pm (Fall and Spring class days)   Office Phone: 254-710-4135

Topic of the Week:
"Hypothesis Testing"

Key Points:

- Symbols and Definitions
- Type I and II Errors
- Structure and Notation

In the previous guides we put all of the knowledge about Sampling Distributions, sample statistics, and the new T-Distribution to create a Confidence Interval. Designed to show the range in which a true population parameter could lie in, the CI works wonders to show where the data could be but has little in terms of definitive results. This is where the process of Hypothesis Testing comes in.

Suppose that a scientist is working in the lab and is expecting the same yield results as a study done back in 1940. These were the last parameters known, so they must be treated as fact until proven false. But how does one prove that a preconceived notation is now wrong? With a test of course.

Hypothesis Testing is the process in which a statistician attempts to reject a preconceived notation in favor of an alternative result. In the scientist's example, the yield rate could be higher than before due to the advancements of modern medicine. In other cases, like in population proportions, the true population parameter could be much lower than initially believed.

We use Hypothesis Tests as a clearer and concise way of representing sample data results. Sometimes a small difference from the previous parameter can lead to two wildly different outcomes if the original parameter is not proven false. Hypothesis Testing allows us to show the significance of a result at any level and makes the results of studies easier for everyone outside of statistics to understand.

Highlight #1
"Symbols and Definitions"

Definition: The nomenclature associated with Hypothesis Testing necessary to understand the processes of computation and results of the data.

Symbols: Ho = Null Hypothesis, Ha= Alternative Hypothesis
p = Population Proportion, μ = Population Mean
$\hat{p}$ = Sample Proportion, $\overline{x}$ = Sample Mean
α = Alpha, β = Beta, (1- β) = Power

Many of the symbols for Hypothesis Testing are shared with the CI notations. For example, the parameters tested are the same as well as the sample data symbols. However, the new symbols revolve around the Hypothesis statements and the errors of a test.

The Null Hypothesis, Ho, (sometimes pronounced as H-Oh, or H-naught) is the predetermined conception of the population parameter. This can be either stated from previous censuses or known by manufacturing plants. The Ho will be the population parameters of p and μ and will be the value we wish to disprove or reject through our studies. Until proven otherwise, always assume Ho is correct during computations. In a properly done Hypothesis Test, Ho will either be rejected or, failed to reject. NEVER accept Ho as the reality.

The Alternative Hypothesis, Ha, is the expectation we plan to have. Ha will never be stated as an exact value, (Saying that Ha might be 55 if Ho is 50). The purpose of Ha is to state what the expected result will be, either higher, lower, or not equal to the assumed parameter. You do not accept nor reject Ha, only use it to reject or fail to reject Ho. You will use the population parameters for the Ha, not the sample statistics.

If the symbol of Alpha looks familiar, it should! This is the same a that we have used in the CI chapter. This will determine the Level of Significance. This is the portion of probability research will consider not likely to occur assuming Ho is true and thus, the probability associated with a is 'significant' if it appears. In the CI guides, (1- α) was the probability we wished to account for. This value will be set and known by the researchers.

The symbol of β, Beta, is the term denoted for the probability a certain error is to occur. Beta has a complicated formula, but it is important to know that Beta and Alpha have inverse relationships. The smaller Alpha will be, the larger Beta will be and vice versa.

Power is the probability a test is capable of rejecting Ho when it is false (a correct conclusion). This is one of the concepts of Hypothesis Testing that truly tests the understanding of how statistics and certain variables react with one another. Beta is the inverse probability of the power of a test. Power is manipulated by 3 values, β, α, and n. When α increases, so does the power of a test. When n increases, power increases. When β increases, power decreases. A .01 change in α results in a higher change in power than a 1 increase in sample size, meaning that power is more susceptive to changes in α than n.

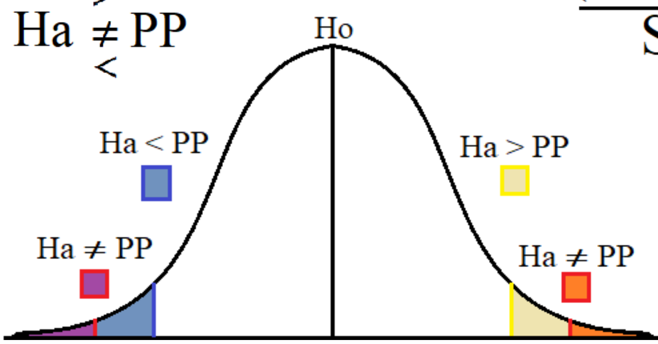
Highlight #2
"Structure and Notation"

Definition: The form in which a Hypothesis Test is computed and interpreted.

Formulas:

$$\text{Ho} = \text{PP}$$
$$\text{Ha} \; \begin{matrix} > \\ \neq \\ < \end{matrix} \; \text{PP}$$

$$\text{T.S.} = \frac{(\text{SS} - \text{PP})}{\text{SE}} \qquad \sqrt{\frac{p*q}{n}}$$



*A Normal Curve depicting 3 potential Rejection Regions

When formatting a Hypothesis Test, Ho will always be an exact value. Ha will always be represented in an inequality. Never use combination inequality signs that include the bar on the bottom. In order to properly be an Ha, the formula must always use an inequality sign. Both of the hypotheses must use the population parameter in question. Never use Sample statistic in your statements. In most cases, you will see the Hypothesis Statements as Ho: PP = # vs Ha: PP >≠< #. This means that the Ho states that the PP is equal to a number, and that the Ha is either greater than, less than, or not equal to the established number.

There are 3 potential Hypothesis tests that can be done with a single parameter. These are shown in the first formula as the potential places Ha can be, if Ho is rejected. Often called

'Tailed'-Tests, these determine what values are considered significant enough to reject Ho.

A 'One-tailed' test is the result of choosing either above or below Ho. These are shown above as Ha < PP (population parameter) and Ha > PP. With an α of 5%, this will mean that both the area to the left of the blue bar (including the purple) and the area above the yellow bar (including the orange) are both 5% exactly. Reminder that these are only 1-tailed, so all of the area to the right of the blue bar is the region in which Ho will fail to reject.

A 'Two-tailed' test is similar to a CI in which, when there is no guarantee which way the PP could truly be, the probability is halved and is placed on either tail. In the graph above, both the orange and purple areas are apart of the same test (Ha ≠ PP). If α = 5% for this example, then both sides would be 2.5% and together, they make up the 5% of probability that can reject Ho.

In addition to formulas, there are several terms unique to the 'Tests' of statistics that tie visual representations to their formulas and results. These are: Rejection Regions, Test Statistic, Observed Value, and P-value.

The Rejection Region (RR) is a predetermined area under a probability curve to which an Ho is rejected in favor of a Ha. In the graph above, the Blue bar represents the upper bound to Ha < PP, meaning that to reject Ho and accept Ha, the values determined must be lower than the blue bar. The same is true for above the yellow bar (Ha > PP), and the area towards the tails on either side of the red bars (Ha ≠ PP).

The Test Statistic (TS) is similar to Standardization in that it is a formula designed to compute a value used for statistical inference. It is found by subtracting the Sample Statistic (SS) from the Population Parameter (PP) and being divided by the Standard Error (SE). Instead of producing a Z-score, a Test Statistic produces an Observed Value. The Observed Value (OV) is compared against one of the rejection region's bounds to see if it is within or outside of said region. The Region's bounds will be given by a Z-score in which the Observed Value must be more extreme in order to reject Ho.

The P-value of a Hypothesis test is a probability given to describe the likelihood that Ha occurs under the assumption that Ho is true. In other words, the probability of getting the result assuming that the population parameter previously determined is indeed the middle of the distribution. Having a smaller p-value is better because it means that the Ha is less likely to occur given the assumptions.

The p-value of .08 means that there is an 8% chance that the resulting observed value can occur from random sampling derived from a population with a parameter of Ho. The p-value will have a cut off point between being statistically significant and failing to reject Ho. This value is the exact probability that α was set at. If α was .05, then 8% would fail to reject. If α was .1, then 8% would be statistically significant and we would reject Ho. There will never be a point in which the p-value is less than α, but the OV is outside of the RR. The p-value determines what Z_a the RR's bound will be and thus, both must agree.

Computing a Hypothesis Test for proportions uses a different Standard Error than the Confidence Intervals. Unlike in the Confidence Intervals, we used the known σ for the computations. In a Hypothesis Test, we use an altered version of the SE form that uses the predetermined PP instead of the SS. The Formula for a Hypothesis Test SE is shown in the 3rd Formula.

Confidence Intervals can also be used to describe the results of a Hypothesis Test if Ho is rejected. Since the previous parameter has been disproven, statisticians like to create a CI for a range based off of the sample data. This range often includes a different value than the OV calculated for the Hypothesis Test. The Formula for a Hypothesis CI is, **CI = (PE) ± (|RF|*SE)**. Where PE is the Point Estimate (Sample Statistic), the Reliability Factor (RF) is the RR bound associated with α, and SE the Standard Error (Std. Dev. for the Sampling Distribution). Depending on the type of test, the Confidence Interval can either be one region or two on the graph. If Ha > Ho CI = [x,∞), If Ha < Ho CI = (-∞, x], and If Ha ≠ Ho CI = [x, x]. For the Two-Tailed Test, the CI will be on one of the tails and uses the RF of that tail (α/2 rather than just α).

Highlight #3
"Type I and II Errors"

Definition: The errors associated with incorrect conclusions derived from the study and its results.

Example: Rejecting Ho when it's true or Failing to reject when Ho is false.

	Fail to Reject	Reject
Ho True	Correct Choice Probabiltiy: (1 - α)	TYPE I ERROR Probability (α)
Ho False	TYPE II ERROR Probability (β)	Correct Choice Probability (1- β)

When firmly declaring whether or not Ho can be rejected always leaves the chance the determined outcome could be incorrect. These chances are called the Type I and II Errors.

Besides the errors, there are the two chances that the choice to reject Ho or Fail to reject is the correct choice.

Type I Errors are when a False rejection occurs, claiming Ho to be false when it is true.

Type II Errors are when a False failure occurs, failing to reject Ho when it is false.

Let is go back to the researcher with a new yield ratio than previously recorded and that the truth is that Ho is false. In this case, the only error that can occur is a Type II, in which the researcher does not have enough statistically significant data to reject Ho. The research has committed a Type II error in that they have incorrectly continued to keep the old yield ration when a new one has been established. This means that the test of his power, or the probability of not committing this error, was low enough that this error occurred.

Another example would be if the Ho was true, but out of sheer luck, the random sample collected was skewed enough that the Ho is rejected in favor of the Ha. Say a dog breeder has been noticing that a dog breed has been having larger birth weights than expected. They run a test and find that there is significant results and reject the previous claims. If the original Ho is still true, then this is a Type 1 error. This means that the chosen alpha level is the reason that the error occurred, and it was not small enough to account for these small but possible chances.

Depending on the circumstances, one error might be more favorable than the other. If a Scientist fails to reject a wrong conclusion, serious consequences could occur from the unbalanced chemical compounds and result in catastrophic results. If the Dog breeder expects larger birth weights, she could accidentally over feed the puppies trying to reach that weight for all other dogs and could accidentally hurt them.

In the same way, some errors would be better than others. If a company that focuses on safety tests the previously determined weight limit of an elevator, it would be better to reject Ho and rework the elevator's mechanics than to fail to reject a faulty piece of equipment. In each instance, look to see which is the 'better' of the errors for the circumstances.

Whether it is a Type I, II, or correct conclusion, assess the conclusion from a Hypothesis Test and determine what it means in the context of the problem.

Check Your Learning

1. Little Timmy walks on a bridge every morning to throw out bird feed and enjoy the company of the neighborhood pigeons. He notices that after he throws the feed, there are less pigeons than he normally sees. Assuming that the birds that have flown in today are a random sample. Timmy is used to seeing roughly 45% of the pigeons in his neighborhood

every morning, but today he only sees 23%. He wonders if the true proportion of the 200 pigeons that live in his neighborhood that like to snack on his feed has changed. Compute a Hypothesis test at the 5% level.

- **a.** What is the Observed Value? Compare it to the Z_α of 5%. What does this tell you about Timmy's Pigeons?
- **b.** Describe a Type II Error in the context of this problem. Can he have committed it given your results?

2. Amber has been taught from a young age how to shoot arrows from her bow. Before getting her new bow, she was used to 89% of her arrow shots being on the target. However, she recently got a new bow for her birthday from her friend Collei and is curious if her accuracy has changed, but she's not sure if its for better or worse. She fires 32 shots at a target at random times during the day to keep them independent. She finds that 24 of them land on target.

- **a.** Compute a Hypothesis Test at the 1% Confidence.
- **b.** Compute the CI associated with this test, or, explain why this cannot be done.

Things Students Struggle With

1. $<$ vs $>$ vs $\neq$:

- **a.** Unless the question clearly states which to compute, it can seem difficult to determine what test to run. One way is to look for word queues. If the question gives words such as 'fewer, less, little,' or any other synonyms, then you will check to see if Ha is < Ho. In contrast, if words such as 'greater, more, many' etc. are said within a question, then test for Ha > Ho. If the question only states if they want to see if the true parameter is different than Ho, than test for ≠. Do not use the numbers alone as your reasoning for a certain test.

2. Z_α vs $Z_{\alpha/2}$:

- **a.** After using $Z_{\alpha/2}$ for so long during the CI chapter, it can be difficult to discern when to use $Z_{\alpha/2}$ and when to use Z_α. Since two of the Hypothesis Tests are 'One-Tailed' Tests, this means that all of the α probability will have to be on one tail. This is when you use Zα. When the Test is asking for either higher or lower, that is when you use $Z_{\alpha/2}$, when the test needs to account for both tails. You are still checking for α% of probability, only now it is between both tails.

Concluding Comments

That's it for this week! Please reach out if you have any questions and don't forget to visit the Tutoring Center website for further information at https://www.baylor.edu/tutoring

---

Answers to CYL

1. **a.** Ho: $p = .45$ vs Ha: $p < .45$ (Ha < Ho). OV = -6.25. (-6.25 < 1.64) Timmy should reject Ho and accept that fewer pigeons are coming to him than normal.
   b. A Type II Error is when you Fail to reject a False Ho. In this instance, Timmy's pigeon attendance would have truly been less, but the sample data could not defend it. Timmy cannot make a Type II Error because he chose to reject Ho. The only Error he could have made would be a Type I.

2. **a.** Ho: $p = .89$ vs Ha: $p \neq .89$. OV = -2.53, P-value = .0114. RR Bound: -2.58. We fail to reject Ho at the 1% level.
   b. Because Ho could not be refuted, it stands to reason that Amber's aim has not changed and thus, there is no need to create a CI when the population parameter of .89 has not been disproven.