# Practice Exam #1 (Sta-209, F26)

Ryan Miller

### Directions

- Answer each question using *no more than specified number of sentences* and not attempt to avoid these guidelines by using run-on sentences. Answers that are unnecessarily verbose may result in point loss.
- Do not include superfluous information in your answers, you may be penalized if you make an inaccurate statement even if you go on to provide a correct answer. Your answers should be clear, concise, and include only what is needed to answer the question that was asked.

### Formulas and Definitions

Below are standard error formulas and large sample conditions used in hypothesis testing:

| Population Parameter | Sample Statistic | Standard Error | Conditions |
|---|---|---|---|
| $p$ | $\hat{p}$ | $\sqrt{\frac{p(1-p)}{n}}$ | $np \geq 10$ and $n(1-p) \geq 10$ |
| $\mu$ | $\bar{x}$ | $\frac{\sigma}{\sqrt{n}}$ | normal population or $n \geq 30$ |
| $p_1 - p_2$ | $\hat{p}_1 - \hat{p}_2$ | $\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$ | $n_i p_i \geq 10$ and $n_i(1-p_i) \geq 10$ for $i \in \{1,2\}$ |
| $\mu_1 - \mu_2$ | $\bar{x}_1 - \bar{x}_2$ | $\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ | normal populations or $n_1 \geq 30$ and $n_2 \geq 30$ |

Below are a few definitions you don't need to memorize:

- Type 1 error - rejecting $H_0$ when it is true
- Type 2 error - failing to reject $H_0$ when it is false
- Exact binomial test - a hypothesis test for a single proportion used for small samples
- Fisher's exact test - a hypothesis test for differences in proportions used for small samples
- Wilcoxon signed-rank test - a non-parametric alternative to the one-sample T-test
- Wilcoxon rank-sum test - a non-parametric alternative to the two-sample T-test

**Section 1 - True/False**

*Directions*: Clearly indicate whether each of the following statements is true or false. You do not need to explain your reasoning and you will not receive a better score for doing so.

1. A *large p-value*, such as $p = 0.90$, should be interpreted as strong evidence that the null hypothesis is *most likely correct.*

2. A *small p-value*, such as $p = 0.01$, should be interpreted as strong evidence that the null hypothesis is *most likely incorrect.*

3. In hypothesis testing, the null hypothesis is set up to reflect what the researchers suspect is true about the population they are studying.

4. The $T$-test is only appropriate when the sample size is relatively large.

5. Suppose we suspect a two-sided coin is biased to land on heads more often than it should. If we flip the coin 15 times and observe 12 heads, the one-sided $p$-value corresponding to this scenario is defined as $Pr(\hat{p} \geq 12/15|p = 0.5)$.

6. The $Z$-test is only appropriate for categorical data when we've observed a relatively large number of cases in each category involved in the test.

7. The $Z$-test compares a standardized test statistic to the $N(0, 1)$ distribution to compute the $p$-value.

8. A *very small p-value*, such as $p = 0.0001$ is suggestive of a very important scientific discovery.

9. In hypothesis testing, the null hypothesis is a claim we make about the sample data.

10. Consider $H_0 : p = 0.5$ and $H_a : p \neq 0.5$ and suppose we collect a sample of size $n = 50$, but we *fail to reject the null hypothesis* after finding a sample proportion of $\hat{p} = 0.6$. If we collect a larger sample, such as $n = 100$, and observe the same sample proportion, $\hat{p} = 0.6$, it is possible that we have found enough evidence in the new sample to reject the null hypothesis.

11. Consider $H_0 : p = 0.5$ and $H_a : p \neq 0.5$ and suppose we collect a sample of size $n = 50$ and *fail to reject the null hypothesis* $H_0 : p = 0.5$ after finding a sample proportion of $\hat{p} = 0.6$. If we collect another sample of size $n = 50$ and observe $\hat{p} = 0.58$ it is possible that we have found enough evidence in the new sample to reject the null hypothesis.

**Section 2 - Conceptual Questions**

*Directions*: Answer each question using no more than 3-sentences. Do not include unnecessary details, as you will be penalized for any inaccurate statements, regardless of whether they are relevant or not. Aim to clearly answer the question that was posed, not to demonstrate your knowledge of related topics.

1. Your friend collected data for their environmental studies course by taking $n = 10$ water samples from different locations in Arbor Lake and $n = 15$ samples from different locations in Lake Nyanza. Their research aim was to compare the nitrate levels (in mg/L) of each lake. Briefly describe the hypothesis test(s) you would suggest this friend use, including any additional information you'd request regarding the data to help you determine the most appropriate test.

2. In your own words, briefly explain *why* seeing a $p$-value meeting an established threshold for statistical significance (such as $\alpha = 0.05$) is a compelling reason to reject the null hypothesis.

3. In your own words, explain the concept of a *null distribution* in hypothesis testing. That is, what does a null distribution display? And how does a null distribution relate to the $p$-value?

4. For each of the following scenarios provide the name of an appropriate data visualization (graph). Be specific.

- A: One-sample categorical data
- B: Two-sample categorical data
- C: One-sample quantitative data
- D: Two-sample quantitative data

**Section 3 - Application**

In a 1997 research paper, psychologists Ruback and Juieng performed a series of studies investigating how quickly a person leaves their parking spot.

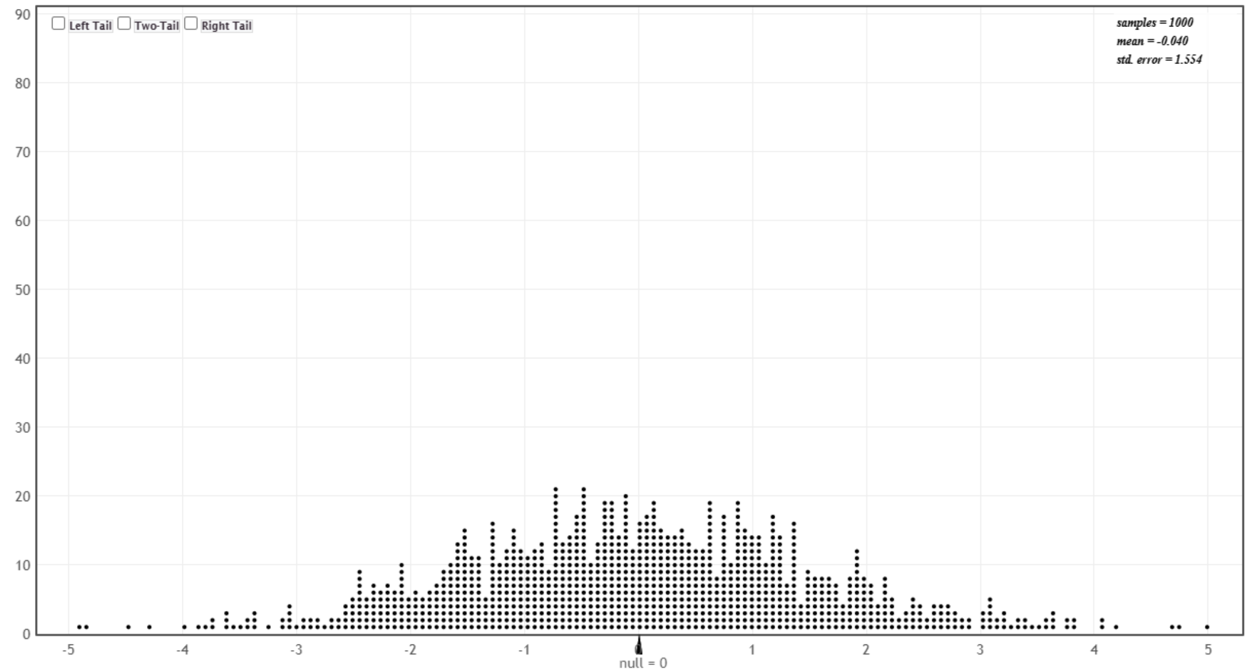
In one of these studies, Ruback and Juieng observed 200 drivers departing from a public parking lot. For each departing driver they recorded the time (in seconds) from when the driver first entered their vehicle to when they had fully exited their parking space. The main explanatory variable in the study was whether or not another car was waiting for the driver's parking space while they were exiting.

A summary of their data is given below:

| Waiting | Mean Time | Median Time | Std Dev | IQR | N |
|---|---|---|---|---|---|
| no | 45.55 | 46.13 | 10.53 | 13.93 | 132 |
| yes | 49.42 | 48.11 | 9.42 | 9.57 | 68 |
| total | 46.87 | 46.68 | 10.31 | 12.30 | 200 |

1. Suppose the researchers would like to evaluate whether the presence of another driver waiting influences the speed at which a person exits a public parking lot. State the null and alternative hypotheses these researchers should consider using the proper statistical notation.

2. Shown below are 1000 outcomes that were simulated using the null hypothesis you proposed in Part 1. Use this distribution to estimate the two-sided $p$-value. You may choose to annotate the figure to help me understand how you are estimating the $p$-value.



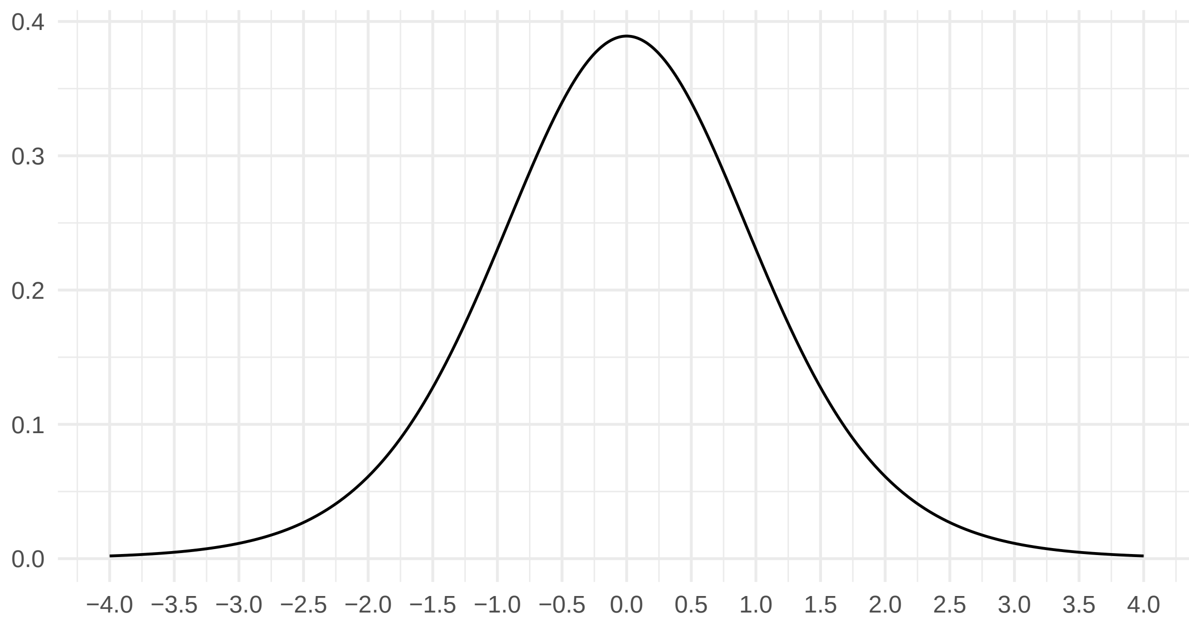

3. Express the $p$-value in this application as a formal probability statement, such as $Pr(\ldots)$.

4. Provide a one-sentence conclusion based upon the $p$-value and hypotheses you've provided in earlier parts of this question. Your conclusion should include all of the components of a proper conclusion discussed in our labs and examples.

5. Suppose you'd like to use a statistical test based upon a probability model (rather than a simulation) to evaluate the hypotheses you stated earlier. What is the name of the statistical test you should use? Provide the name of the test as well as a brief explanation of why it is an appropriate choice.

6. The proper test statistic in this scenario uses the standard error formula $SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$. Calculate this test statistic. Show your work.

7. Shown below is the probability model for the test statistic you calculated in the previous part. Shade the area of this distribution that represents the *two-sided p-value* in this application.




8. The researchers in this study balanced their sample by following an equal number of male and female drivers into the parking lot, but they could not control whether or not a vehicle was waiting for each of these drivers. For the male drivers, they observed 38 instances of a vehicle waiting, and for the female drivers they observed 30 instances of a vehicle waiting. Consider the null hypothesis that male and female drivers were equally likely to have a vehicle waiting. Fill in the missing components (question marks) in the `R` code given below to test this hypothesis:

```
prop.test(x = ?, n= ?, alternative = "two.sided")
```

9. The $p$-value produced by the `prop.test()` function from the previous part is $p = 0.296$. Provide an appropriate one-sentence conclusion based upon this $p$-value.

10. Is it possible that the conclusion you made in Part 9. was a decision error? If so, what type of error would it be? Briefly explain your rationale.

11. Suppose all 27 students in our class are motivated by this example and each go out and conduct their own replication in different public parking lots across Iowa. If each student conducted a hypothesis test using the $\alpha = 0.1$ significance level, and the null hypothesis were true for every test, how many students are expected to have a statistically significant result by chance alone?

12. Suppose you receive the $p$-values from each of the 27 tests described in Part 11. Briefly explain how would analyze these results in a way that limited the chances of making one or more Type 1 errors to 10%.