1. Identify the point estimate in a statement.
   1. Beets are rich in nitrates, which our bodies convert to nitric oxide. Circulating nitric oxide can promote blood flow, improve oxygen exchange, and strengthen muscle contractions. Your research question is whether drinking beet juice before exercise affects cardiorespiratory performance as compared to drinking water.   
        
      You randomized 100 patients each to a control arm, which drank 12 ounces of water 30 minutes before exercise, and to a treatment arm, which drank 12 ounces of beet juice 30 minutes before exercise. The outcome you measured was their maximum oxygen intake (VO2max) during a 5km run.  
        
      The average VO2max in the control arm was 33.0 mL/kg/min with a standard deviation of 7.0, and the VO2max was 4.0 mL/kg/min higher after drinking beet juice, with an average of 37.0 mL/kg/min and standard deviation of 10.0 mL/kg/min. You estimate that the standard error of the measurement is 1.2 mL/kg/min.   
        
      You’re using a Student’s t distribution with 99 degrees of freedom to infer the sampling distribution, and the critical value for a type I error rate of 0.05 is 1.98. With 95% confidence, the average VO2max during exercise after drinking beet juice is 1.6 to 6.4 mL/kg/min greater than it is after drinking water.  
        
      Your null hypothesis is that there is no difference in the average VO2max during exercise after drinking beet juice and after drinking water. Given the null hypothesis, the observed difference was 3.3 standard errors above the null value. Assuming that the null hypothesis is true, a 2-sample t-test indicated that the probability of observing a difference as extreme or more extreme than the one seen in your sample was 0.0014.  
        
      Which number below represents the point estimate for this hypothesis test?
   2. Options
      1. **4.0 mL/kg/min**
      2. 3.3 standard errors
      3. 1.2 mL/kg/min
      4. 0.0014
      5. 1.98
2. CLT assumptions – success/failure condition
   1. Your boss would like you to find out whether turnover rates are different between your company’s remote workers and in-person workforce. You took a random sample of 155 employees who worked at the company within the past 12 months from the records in Human Resources. 110 of the randomly sampled employees worked remotely, and 9.1% of those employees left the company during those 12 months. Of the 45 employees who worked in-person, 20% left the company during those 12 months.  
        
      Does your data meet the conditions for using the Central Limit Theorem to estimate the difference between these two proportions?
   2. Options
      1. **No**
      2. Yes
3. CLT assumptions – sample size / normality
   1. You are studying behavioral problems in juveniles, and you want to determine if the average blood lead level in your study population is above the 3.5 μg/dL threshold for lead poisoning risk. Unfortunately, you have a very limited budget. You take a random sample of 30 children and calculate the following summary statistics for your sample.   
        
      Does your data meet the conditions for using the Central Limit Theorem for hypothesis testing of a single mean?
   2. Options
      1. **No**
      2. Yes
4. Reliability check (observed = expected)
   1. You want to estimate the proportions of men and women who use recreational marijuana in your study population. In 1-2 sentences, state whether responses to a survey or the results of a urine test would provide more reliable sample statistics and why.
   2. Free response
5. Validity check (sample = population)
   1. You take a random sample of 100 days from 2023 and calculate a mean high temperature of 23 degrees Celsius. You also take a convenience sample of the last 100 days of 2023 and calculate a mean high temperature of 20 degrees Celsius. In 1-2 sentences, describe the validity of the sample statistics from the two sampling approaches.
   2. Free response.
6. T- vs Z- distribution
   1. Identify from picture which is T, explain why  
      A diagram of a normal distribution

      Description automatically generated
   2. Options
      1. Free response
7. Infer the sampling distribution for single p
   1. You run a small retail shop, and you are trying to write your budget for 2025. You’d like to infer the percentage of items that are returned from your 2024 data so that you can generalize this information to your 2025 projected revenue estimates.  
        
      In a random sample of 1000 items sold in 2024, 350 were on sale and 650 were full price. 119 of the items sold that were on sale were eventually returned, while 234 of the full price items sold were eventually returned.  
        
      Which of the following is the correct sampling distribution for the average proportion of items sold that are returned at your store?
   2. Options
      1. **N(0.353, 0.015)**
      2. N(0.340, 0.025)
      3. N(0.360, 0.019)
      4. N(0.353, 0.032)
8. T-distribution critical value
   1. You’ve been given a budget for targeted advertising, and you want to know if the average male customer or the average female customer spent more money at your business in 2024. You take a random sample of 500 customers, 300 of which are men and 200 of which are women.   
        
      Men spent an average of $247 dollars with a standard deviation of $37. Women spent an average of $266 dollars with a standard deviation of $64. With 95% confidence, the average woman spent between $9.13 and $28.87 more than the average man.  
        
      Which of the following is the correct critical value from the Student’s t distribution for a 95% confidence interval for the average spending by a female customer?
   2. Options
      1. **1.972**
      2. 1.968
      3. -1.960
      4. 1.960
9. Identify a type I error
   1. You want to determine the effectiveness of the abstinence-only sex education program at the high school you teach at. You conduct interviews with 75 randomly selected students about their experiences with sex. 49 out of the 75 students (65.3%) report that they have never had sex. You test the hypothesis that the majority of students (>50%) have not had sex yet and get a p-value of 0.003. Based on a significance threshold of alpha=0.05, you reject the null hypothesis that p=0.5.  
        
      However, you later learn that at least 10 of the students were not honest and lied about having had sex. In truth, only 40% of students had not yet had sex. What type of decision error did you make?
   2. Options
      1. **Type I Error**
      2. Type II Error
10. Identify a type II error
    1. You were feeling under the weather, so you took a COVID rapid test and made sure it was negative before going out to see friends. You felt worse the next day, and you had a COVID PCR test done at an urgent care facility. It came back positive. What is this an example of?
    2. Options
       1. **Type II Error**
       2. Type I Error
11. Identify correct code for a 1-sided Z-test p-value
    1. You are doing a 1-sample proportion test using a sample size of 100 and the null hypothesis H\_0: p=0.5. Your alternate hypothesis is that the average proportion is greater than the null value (H\_A: p > 0.5). Your observed proportion was 0.618 and your test statistic was 2.36. Identify any code which will generate the correct p-value for this hypothesis test.
    2. Options
       1. **pnorm(2.36, lower.tail = F)**
       2. **pnorm(-2.36)**
       3. **1-pnorm(2.36)**
       4. pnorm(2.36)
       5. pt(2.36, df = 99, lower.tail = F)
       6. pt(-2.36, df = 99)
       7. pt(2.36, df=99)
12. Proportion difference confidence interval
    1. Download the file bike\_lanes.xlsx from Canvas under “Take-Home Quiz” on the Modules page. Load it into R.   
         
       Calculate a 95% confidence interval for the difference in mean proportion of people who come to a complete stop before turning right when there is and is not a protected bike lane.   
         
       Interpret your confidence interval using a complete sentence.
    2. Free response
13. 2-sample proportion test
    1. Download the file bike\_lanes.xlsx from Canvas under “Take-Home Quiz” on the Modules page. Load it into R.  
         
       Conduct a 2-sided hypothesis test for a difference in the proportion of people who come to a complete stop before turning right when there is and is not a protected bike lane.   
         
       Using a significance threshold of alpha=0.05, report your test statistic, p-value, and conclusion of the hypothesis test.
    2. Free response
14. 2-sample t-test
    1. Download the file fast\_food.csv from Canvas under “Take-Home Quiz” on the Modules page. Load it into R.  
         
       Conduct a 2-sided hypothesis test for a difference in the mean levels of sodium in menu items from McDonald’s and Burger King.  
         
       Using a significance threshold of alpha=0.05, report your test statistic, p-value, and conclusion of the hypothesis test.
    2. Free response