# Handout

Dr. Gia Barboza-Salerno
Statistics I

September 15, 2023

**Problem 1.1.** The State University of Ohio (TSUO) is an elite university of 50,000 students. Nadia, a student studying at the university, wishes to determine the average IQ of students. She has decided that her sample size will be 50, and she is considering several different sampling methods. For each method, state the sampling frame and discuss whether the sampling method is random and whether it is independent.

1. Nadia chooses 50 names at random from the list of students at the university

   *Answer.*

   □

2. Nadia asks 50 of her acquaintances at the university if they would mind taking an IQ test

   *Answer.*

   □

3. Nadia chooses the first two students from the alphabetical list of each of the 25 university departments

   *Answer.*

   □

4. Nadia takes all 50,000 names and puts them into a hat. She draws out a name, writes it down, and then puts it back in the hat and draws again. This procedure is repeated 50 times

   *Answer.*

   □

**Problem 1.2.** Hale Prison is renowned for its poor internal discipline. The new prison governor wants to tackle this problem and decides to investigate whether removing prisoners' visiting privileges will act as a deterrent against future misbehaving. From 100 prisoners who recently took part in a violent prison riot, he selects the 25 inmates with the worst disciplinary records, removes their visiting privileges, and begins to monitor their progress relative to the others.

1. Does this method meet the criteria of independent random sampling? Why?

   *Answer.*

   □

2. Is independent sampling possible in this case? Why?

   *Answer.*

3. Describe a more appropriate sampling method

   *Answer.*

□

**Problem 1.3.** For each of the following hypotheses, state whether a one-tailed or a two-tailed test of statistical significance would be appropriate. In each case, explain your choice.

1. $H_1$: Citizens over the age of 50 are more likely to be the victims of assault than citizens under the age of 50.

   *Answer.*

□

2. $H_1$: Children raised by adopted parents have rates of delinquency different from those of children raised by their biological parents.

   *Answer.*

□

3. $H_1$: The experience of imprisonment has an impact on the chances of an ex-convict reoffending.

   *Answer.*

□

4. $H_1$: Male defendants are more likely to be held on bail than female defendants.

   *Answer.*

□

5. $H_1$: Teenagers have rates of crime that are different from adult rates of crime.

   *Answer.*

□

**Problem 1.4.** According to one estimate I read, 13% of adolescents will engage in delinquent behavior during adolescence. If we select five adolescents at random, find the probability that:

1. The first adolescent to engage in delinquent behavior is the fifth person chosen

   *Answer.*

□

2. Exactly 3 adolescents will engage in delinquent behavior?

   *Answer.*

□

3. There are some delinquent adolescents among the five chosen.
   *Answer.*

4. There are no more than 3 adolescents who engage in delinquent behavior.

   *Answer.*

□

**Problem 1.5.** A social work undergraduate student is interested in whether "successful" people are more likely to be born under some astrological signs than others? The data below shows the birthday signs of 256 executives of Fortune 400 companies. Notice there is some variation in the number of births per sign, and there are more people born under Pisces and Aquarius :). The research question we might want to address is whether successful people are more likely to be born under some signs than others? This requires that we test certain hypotheses. **Note%**: we will learn how to test this hypothesis later.

□

□

□

| Sign | Number of Births |
|---|---|
| Aries | 23 |
| Taurus | 20 |
| Gemini | 18 |
| Cancer | 23 |
| Leo | 20 |
| Virgo | 19 |
| Libra | 18 |
| Scorpio | 21 |
| Sagittarius | 19 |
| Capricorn | 22 |
| Aquarius | 24 |
| Pisces | 29 |

1. Define $H_0$ and $H_1$

   *Answer.*

   □

2. What is the sampling distribution for the statistical test you will conduct?

   *Answer.*

   □

3. Compute *df*

   *Answer.*

   □

4. Assume that you are willing to reject the $H_0$ even if the $H_0$ is true about 1 in 20 times. What is your $\alpha$?

   *Answer.*

   □

5. In words, state the Type I error in this problem.

   *Answer.*

   □

6. What is the number of births under each sign that you would expect if the $H_0$ is true?

□

**Problem 1.6.** Use the following binomial distribution showing the chances of success and failure for 12 trials.

| **Number of Successes** | **Probability** |
|---|---|
| 0 successes | 0.00118 |
| 1 successes | 0.01065 |
| 2 successes | 0.04418 |
| 3 successes | 0.11110 |
| 4 successes | 0.18857 |
| 5 successes | 0.22761 |
| 6 successes | 0.20032 |
| 7 successes | 0.12953 |
| 8 successes | 0.06107 |
| 9 successes | 0.02048 |
| 10 successes | 0.00463 |
| 11 successes | 0.00064 |
| 12 successes | 0.00004 |

□

Using $\alpha = .05$, what outcomes would lead you to reject the null hypothesis for each of the following pairs of hypotheses?

1. $H_0 : P = .05$
   $H_1 : P \neq .05$

   *Answer.*

2. $H_0 : P = .05$
   $H_1 : P < .05$

   *Answer.*

□

*Answer.*

□

□

**Problem 1.7.** In statistics, we usually refer to $x_1$ as the first observation, $x_2$ as the second observation, etc., and $x_n$ as the final observation when we write down our observations in the order they are obtained, where $n$ represents the total number of observations.

Often we prefer to work with ordered observations that are sorted from smallest to largest, such as when calculating the median. We can just sort them, but I wanted you to understand how to use notation, and how notation changes, depending on the problem at hand. So, when we describe ordered data we change the notation slightly as follows: Let $x_{(1)}$ refer to the smallest observation in a set of numbers, $x_{(2)}$ the second smallest, and $x_{(n)}$ the largest.

Suppose we have the following distribution:

$$3, 5, 8, 11, 0$$

Then

$$x_1 = 3, x_2 = 5, ...x_5 = 0.$$

To understand how we determine $x_{(1)}$, $x_{(2)}$ ... $x_{(n)}$ simply reorder the numbers... as follows

$$x_{(1)} = 0, x_{(2)} = 3, ...x_{(5)} = 11$$

Also, when we have a group of numbers recall that $\Sigma$ tells us to sum each number. For example, recall that the following notation means sum all observations starting from the first and ending with the last (i.e., from 1 to n). We have used this as our numerator when we calculate the mean.

$$\sum_{i=1}^{n} x_i$$

We do not have to start from the first observation. This formula says something slightly different. It says start with the second observation and sum the numbers from 2 to n.

$$\sum_{i=2}^{n} x_i$$

This formula says something different too. It says start with the second observation and sum the numbers from 2 to the second to the last $(n-1)$.

$$\sum_{i=2}^{n-1} x_i$$

Finally, this formula says to sum the **ordered** observations from the second to the second to last observation.

$$\sum_{i=2}^{n-1} x_{(i)}$$

1. Calculate the mean for the numbers above

   *Answer.*

   □

2. Calculate the following summation

$$\sum_{i=2}^{n} x_i$$

   *Answer.*

   □

3. Calculate the following ordered summation

$$\sum_{i=2}^{n-1} x_{(i)}$$