Activity #1: Introduction to Probability & Statistical Inference

Suppose your statistics professor on the first day of class proclaims:

> "Over the summer, I developed telekinesis. To demonstrate, I will stand up front while a student in the back of the room tosses 100 pennies. Through the sheer power of my mind, I will cause those pennies to come up heads. Now I admit that I have not yet perfected this skill – I cannot make **all** the coins come up heads. What I can do is make an **impressive** number come up heads, to the point where you will believe that I may be telling the truth."

1) How many of the 100 pennies would need to come up heads for you to be impressed enough to take my claim seriously? Why?

2) Would you be impressed if 50 pennies came up heads? Why or why not?

3) You probably would not be impressed with 51, 52, or 53 heads, but what if I got 65, 75, or 85 heads? At what point would you begin to think that the results were not likely to have come from 100 normal (fair) pennies? Mark this on the following number line:

0 25 50 75 100 (Number of Heads)

The logic we will use to determine if the professor has telekinesis is:

a) We will assume nothing special or unusual is happening – a *null hypothesis*. In this example, we will assume the professor does not have telekinesis.

b) We will perform the experiment. In this example, we toss 100 pennies and record the number of heads.

c) We compare the results we **observed** from our experiment with the results we would **expect** under our null hypothesis. If our null hypothesis were true, then we would *expect* to have 50 heads and 50 tails. If the results from our experiment differ *significantly* from this expectation, then we would conclude our null hypothesis must be incorrect. We would reject our null hypothesis and conclude another hypothesis (the professor has telekinesis) is correct.

4) This logic is called statistical inference, or *Statistics*. What do the terms statistics, Statistics, and probability represent?

statistics:

Statistics:

Probability:

5) Suppose we conduct our experiment and observe 52 heads. What do we conclude? How certain are we of this conclusion?

6) Assuming the coin is fair and that the professor does not have telekinesis, what is the probability that the student will toss 50, 52, 70, or 87 heads? We'll learn how to calculate these probabilities in the first few weeks of class – for now, see how (in)accurate your intuition is.

P(exactly 50 heads) = _________ P(50 or more heads) = _________ P(49 or fewer heads) = __________

P(52 or more heads) = _________ P(70 or more heads) = __________ P(87 or more heads) = __________

7) Let's say we conduct this experiment and the student does toss more than 70 heads. From this, we conclude our assumption (no telekinesis, fair coin) is incorrect. How certain are we in our conclusion? What is the probability that our conclusion is incorrect?

8) This time, suppose the student tosses only 10 pennies and gets 7 heads. Would you be as impressed with 7 heads out of 10 pennies than you would be with 70 heads out of 100 pennies?

P(exactly 5 heads from 10 pennies) = __________ P(6 or more heads from 10 pennies) = __________

P(7 or more heads from 10 pennies) = __________ P(9 or more heads from 10 pennies) = __________

9) Let's say the student tosses 10 pennies and gets 7 heads. What is our conclusion? Explain your reasoning.

Statistics is the analysis of data to make effective decisions. To do this, we will need to understand probability, descriptive statistics, and inferential statistics (MATH 300). Once we understand these concepts, we can use a variety of methods to analyze all sorts of data (MATH 301). From there, we can apply our skills to real data sets to make real decisions (MATH 305).

Scenario: Rosen, B. & Jerdee, T. (1974). Influence of sex role stereotypes on personnel decisions. *Journal of Applied Psychology,* 59: 9-14.

In 1972, 24 male bank supervisors were each given the same personnel file and asked to judge whether the person (applicant) should be promoted to a branch manager position. While they were given the same personnel file, supervisors were randomly assigned into one of two groups:

Group A: The 12 supervisors in this group were told that the personnel file was from a **male applicant**
Group B: The 12 supervisors in this group were told that the personnel file was from a **female applicant**

Of the 24 personnel files reviewed, 18 were selected for promotion.

10) We can classify studies into *experimental studies* (where we randomly assign subjects to groups) or *observational studies* (where we do not randomly assign subjects to groups). Why did the researchers **randomly** assign supervisors to the two groups? Why didn't they just put the oldest 12 supervisors in the first group? How could you **randomly** assign subjects to groups?

11) We know 18 of the 24 applicants were selected for promotion. Let's assume that gender had **no** impact on promotion decisions. We'll call this the *null hypothesis* (a statement that, in essence, says "nothing happens"). Assuming this null hypothesis is true, how many males and females should have been promoted? Complete the following table:

EXPECTED RESULTS	Promoted	Not Promoted	Total
Group A (Male)			12
Group B (Female)			12
Total	18	6	24

12) The actual results from this experiment are shown below. What proportion of male and female applicants were promoted? From this observed data, can we conclude that male bank supervisors are sexist?

OBSERVED RESULTS	Promoted	Not Promoted	Total
Group A (Male)	11	1	12
Group B (Female)	7	5	12
Total	18	6	24

Our **null hypothesis**: Gender has no impact on promotion decisions

Under this null hypothesis, we expected to find 9 males and 9 females promoted.

From our experiment, we actually observed 11 males and 7 females were promoted.

Key question: Assuming our null hypothesis is true, how likely were we to observe the results we actually got from our experiment? If gender had no impact on promotion decisions, how likely were we to obtain these results?

In MATH 300, we'll learn how to approximate answers to these questions using randomization methods. In a few weeks, we'll learn how to calculate an exact answer to these questions using the hypergeometric distribution. In MATH 301, we'll learn yet another method (chi-square test for independence) that we could use to address these questions.

For now, let's conduct a simulation to estimate the likelihood of observing 11 male and 7 female promotions (assuming gender has no impact on promotion decisions.)

SIMULATION METHODS

We will assume that supervisors make their promotion decisions randomly (gender had no impact on promotion decisions).

We'll also assume that if we could go back in time and conduct the experiment again (with supervisors likely being randomly assigned to different groups), we would still have had 18 promotion decisions and 6 non-promotion decisions.

We're interested in estimating the likelihood of observing 11 or more male promotions under these assumptions. This will tell us the likelihood of observing results as (or even more) extreme as what we actually observed.

a) Take 24 index cards and write **PROMOTED** on **18** of them. Write **NOT PROMOTED** on the other **6** cards.
b) Shuffle the cards and randomly select **12 cards** to represent the 12 "male" personnel files assigned to Group A.
c) Count the number of promotions you get in this sample of 12 cards. Write this number in the table below.
d) Repeat this simulation 5 times and complete the table

Simulation Number:	**1**	**2**	**3**	**4**	**5**
Number of promotions in Group A:	_______	_______	_______	_______	_______
Did you find 11 or more promotions?	_______	_______	_______	_______	_______

Combine the results from everyone in the class and complete the following table:

# of promoted males	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
# of repetitions												
Percentage												
% of repetitions with 11 or more male promotions												

Null hypothesis: Gender had no impact on promotion decisions

Assuming this hypothesis is true, we would expect to find 9 promoted males and 9 promoted females.

We actually observed 11 promoted males and 7 promoted females.

The likelihood of this happening, assuming the null hypothesis is true (from our simulation), is **_______________**

Therefore, we conclude that our null hypothesis is probably: TRUE or FALSE

In a few weeks, we'll be able to calculate the following likelihoods of obtaining a specific number of male promotions:

Number of promoted males	Likelihood	Cumulative Likelihood
1	0.0%	0.0%
2	0.0%	0.0%
3	0.0%	0.0%
4	0.0%	0.0%
5	0.0%	0.0%
6	0.6%	0.6%
7	7.1%	7.7%
8	24.3%	32.0%
9	36.0%	68.0%
10	24.3%	92.3%
11	7.1%	99.4%
12	0.6%	100.0%

QUESTIONS:

1. **Why do we assume the null hypothesis is true?**

2. **Suppose 10 males were promoted in this study. What's the likelihood of observing 10 or more promoted males if gender has no impact on promotion decisions?**