University of Pennsylvania
BIOL4536 Fall 2023

HW#1
(Hypothesis Testing)

Assigned August 30th
Due September 6th, 3:30pm

Problem 1. (3 points)

Suppose you have a coin and you want to test the following null hypothesis about it:

H_0: The probability of a head is 0.1

We'll flip the coin 11 times and reject if we observe more heads than we would expect if H_0 were true, where we're willing to be wrong with probability no more than 0.05. In other words we want to see if the probability of the observed number of heads *or more* is less than 0.05, assuming H_0 were true.

Suppose you flip the coin 11 times and record the following sequence of heads and tails. This is "the data".

$$T, T, H, T, T, T, T, H, T, H, T \tag{1}$$

The number of heads out of M flips is a binomial random variable with probability function:

$$\text{Prob(flipping } k \text{ heads in } M \text{ flips} \,|\, H_0) = \binom{M}{k} p^k (1-p)^{M-k} \tag{2}$$

where p is the probability of getting a head in one flip under H_0 and $\binom{M}{k}$ is the "binomial coefficient"

$$\binom{M}{k} = \frac{M!}{k!(M-k)!}$$

Use formula (2) to calculate the probability of flipping the number of heads observed in data set (1), assuming the null hypothesis is true ($p = 0.1$). You will need to use a calculator, or use an online binomial coefficient calculators, for example this one:
`https://www.gigacalculator.com/calculators/binomial-probability-calculator.php`

The p-value of the event of flipping k heads under H_0 is the probability of flipping k *or more* heads under H_0. Compute the p-value for the data (1) using formula (2).

Would you accept or reject the null hypothesis?

Problem 2. (3 points)

Suppose we collect data, six values from each of two experimental conditions:

Condition 1	23.17	19.23	24.4	21.22	14.17	28.2
Condition 2	13.3	18.88	19.21	24.31	15.3	13.8

Consider two null hyptheses:

H_0: The means in the two conditions are equal
H_0': The means in the two conditions are different

Use the following online calculator to perform a T-test:
`https://www.graphpad.com/quickcalcs/ttest1/?format=C`
You want to use an 'unpaired' test here.

What is the reported p-value?

Is this the p-value for H_0, H_0' or both?

Based on the p-value, would you reject H_0?

Based on the p-value, would you reject H_0'?

Problem 3. (4 points)

Suppose we want to test if a coin is fair. In other words we want to test the null hypothesis:

H_0: The probability of a head is 0.5

Suppose you flip the coin 12 times.
- Don't actually do it, just assume you did and that you got k heads out of the 12 flips.

We're going to reject H_0 if the number of heads is surprisingly small or surprisingly large. The goal is to determine these two cutoffs for rejection.

If we flip $k \geq 6$ heads. Then the p-value is

$$\text{Prob(flipped} \geq k \text{ heads or} \leq 12 - k \text{ heads).}$$

And if we flipped $k \leq 6$ heads. Then the p-value is

$$\text{Prob(flipped} \leq k \text{ heads or} \geq 12 - k \text{ heads).}$$

So for example if we flipped $k = 11$ heads then the p-value is Prob(flipped either $0, 1, 11$ or 12 heads).

Note, this is the exact same p-value as you'd get if you observed $k = 1$ heads.

This is called a two-tailed test which accounts for the fact that the observed number could be inordinately small or large if the null is true. Strictly speaking, we should have done a two-sided test also in exercise #1, but it gets a bit complicated when $p \neq 0.5$. Doing it one-sided is not wrong, but loses the ability to detect if things go the other way. For example had we flipped zero heads out of 11 that's also evidence that the probability of a head is not 0.1, but our one-sided test would fail to see that.

Suppose we decide reject H_0 if $p < 0.05$

What are the p-values for the values for k for which you would reject? Give them to at least three significant digits (that is, the first non-zero digit and the subsequent two, whatever they are).

Hint: Use formula (2) to calculate the p-value for various values of k (start with $k = 12$ and work your way down, and use symmetry to infer the p-values for the small values of k from the large, e.g. 0 from 12, 1 from 11, etc.).