

## HW 06 Written Work

● Graded

Student

Sangwon Ji

Total Points

20 / 20 pts

Question 1

Q1.2

4 / 4 pts

✓ + 4 pts Valid Null (e.g. assignment of vaccines is like tossing a coin that lands heads with chance 3/5)

+ 0 pts Incorrect/blank

Question 2

Q1.3

4 / 4 pts

✓ + 4 pts Valid Alternative (e.g. assignment of vaccines is not like tossing a coin)

+ 0 pts Incorrect/blank

Question 3

Q1.7

4 / 4 pts

✓ + 0.5 pts Uses sample\_proportions in one\_simulated\_statistic

✓ + 0.5 pts Correct model\_proportions (3/5, 2/5)

✓ + 1.5 pts Returns test statistic (absolute difference between percent heads and 60)

✓ + 0.5 pts Iterates through for loop num\_simulations amount of times

✓ + 1 pt Appends correctly to simulated\_statistics array

+ 0 pts Incorrect/blank

#### Question 4

Q2.1

Resolved 4 / 4 pts

✓ + 1 pt Valid null (each factor contributes equally)

✓ + 1 pt Valid alternative (each factor does not contribute equally)

✓ + 2 pts Identifies TVD as test statistic

+ 0 pts Incorrect/blank

🔄 Regrade Request

Submitted on: Aug 02

It is marked as blank but I do have the answer written:

Null hypothesis: Each factor accounts for 1/6 of the total happiness score  
Alternative Hypothesis: Each factor does not account for 1/6 of the total happiness score  
Test Statistic: Using the TVD, total variation difference, the difference in distribution,  
 $\text{sum}(\text{abs}(\text{observed\_distribution} - \text{null\_distribution}))/2$

The reason I think it showed up like this is because question 2.1 didn't have the proper blank to answer this so I have answered it in a different blank, which is right below.

Gotcha, next time double check your formatting! I'll make an exception but generally you'd lose points for this formatting mistake.

Reviewed on: Aug 03

#### Question 5

Q2.5

4 / 4 pts

✓ + 1.5 pts Correctly identifies that the p-value is less than the cutoff value

✓ + 2.5 pts Correctly concludes that the data is more consistent with the alternative hypothesis

+ 0 pts Incorrect/blank

**Question 1.2.** State the null hypothesis. It should reflect the position of either Dr. DeNero or Dr. Sahai.  
(4 points)

*Note:* Check out [11.3](#) for a refresher on hypotheses.

The probability of getting vaccine 1 is  $3/5$ .  $P = 3/5$ . Probability of getting vaccine 2 is  $2/5$ .



**Question 1.3.** State the alternative hypothesis. It should reflect the position of the doctor you did not choose to represent in Question 1.2. (4 points)

*Note:* Check out [11.3](#) for a refresher on hypotheses.

The probability of getting vaccine 1 is not  $3/5$ .



**Question 1.7.** Simulate 20,000 values of the test statistic under the assumption you picked in Question 1.6. (4 points)

As usual, start by defining a function that simulates one value of the statistic. Your function should use `sample_proportions`. (You may find a variable defined in Question 1.1 useful here!) Then, write a `for` loop to simulate multiple values and collect them in the array `simulated_statistics`.

Use as many lines of code as you need. We have included the code that visualizes the distribution of the simulated values. The red dot represents the observed statistic you found in Question 1.5.

```
In [11]: def one_simulated_statistic():
          sample_distribution = abs(sample_proportions(sample_size, [3/5,2/5]).item(0))
          return 100 * abs(sample_distribution - 0.6)

          num_simulations = 20000
          simulated_statistics = make_array()
          for i in np.arange(num_simulations):
              simulated_statistics= np.append(simulated_statistics, one_simulated_statistic())

          # Run the this cell a few times to see how the simulated statistic changes
          one_simulated_statistic()
```

```
Out[11]: 3.836477987421383
```





**Question 2.1.** Suppose we want to test whether or not each factor contributes the same amount to the overall Happiness Score. Define the null hypothesis, alternative hypothesis, and test statistic in the cell below. Feel free to check your work with another student or course staff. **(4 points)**

*Note:* Please format your answer as follows: - Null Hypothesis: ...

- Alternative Hypothesis: ...

- Test Statistic: ...



Null hypothesis: Each factor accounts for  $1/6$  of the total happiness score  
Alternative Hypothesis: Each factor does not account for  $1/6$  of the total happiness score  
Test Statistic: Using the TVD, total variation difference, the difference in distribution,  $\text{sum}(\text{abs}(\text{observed\_distribution} - \text{null\_distribution}))/2$

**Question 2.5.** What can you conclude about how each factor contributes to the overall happiness score in the US? Explain your answer using the results of your hypothesis test. Assume a p-value cutoff of 5%. (4 points)

The result, test statistic is lower than the cutoff of the p-value which is 5%. So, we reject the null hypothesis. We can conclude and indicate that factors do not account for the overall happiness score at the same proportion, which is proportion of  $1/6$ .

