

## CS2065 Spring 2022

### Discussion: A/B Testing

One special kind of hypothesis test we do in this class is called an A/B test. The steps used to run an A/B test are the same as a general hypothesis test, but A/B tests have a specific null hypothesis (that two samples were drawn from the same distribution), which we test by performing a *permutation*.

1. Choose True/False for each of the statements below, and explain your answer.

- a) A/B testing is used to determine whether or not we believe two samples come from the same underlying distribution.
- b) To conduct a permutation test, you should sample your data with replacement with a sample size equal to the number of rows in the table.
- c) A/B testing is the same as using total variation distance as a test statistic for a hypothesis test.
- d) A/B testing is about comparing two samples, whereas regular hypothesis testing is about testing properties of one sample.

2. Natalia, a museum curator, has recently been brought specimens of caddisflies collected from various parts of Northern California. The scientists who collected the caddisflies think that caddisflies collected at higher altitudes tend to be bigger. They tell her that the average length of the 560 caddisflies collected at high elevation is 14mm, while the average length of the 450 caddisflies collected slightly lower down is 12mm. She's not sure that this difference really matters, and thinks that this could just be the result of chance in sampling.

- a) How could you test the null hypothesis in the A/B test from above? What assumption would you make to test the hypothesis, and how would you simulate under that assumption?
- b) What would be a useful test statistic for the A/B test?

c) Assume `original_table` refers to the following table :

A/B label	Specimen length (mm)
High elevation	12.3
Low elevation	13.1
High elevation	12.0

...  
(1007 rows omitted)

Fill in the blanks in this code to generate one value of the test statistic under the null hypothesis.

```
def permutation_test():
    shuffled_labels =
    original_table._____ (_____) .column('A/B label')

    original_with_shuffled_labels = original_table.drop('A/B
    label').with_columns(_____, _____)

    grouped =
    original_with_shuffled_labels._____ (_____, _____)

    means = grouped.column('Specimen length mean')
    statistic = _____
    return statistic
```

d) Fill in the code below to simulate 10000 trials of our permutation test

```
test_stats = _____
repetitions = _____

for i in np.arange(_____):

    one_stat = _____

    test_stats = _____

test_stats
```

e) Given that the p-value of the test was 0.1, draw a possible histogram of the test statistics