

Question 1

A lady asserts that by tasting a cup of tea made with milk she can determine whether the milk or the tea infusion was first added to the cup. Suppose that an experiment is designed to allow the lady to provide evidence in support her claim. 10 cups of tea are made and presented to the lady in a random order. 5 of the cups are made with tea first, and the other 5 are made with milk first. The lady is tasked with identifying which 5 of the 10 cups are made with tea first.

- (a) What is the null hypothesis in this experiment?
- (b) What is the probability of the lady correctly selecting all 5 tea first cups by pure chance?
- (c) If a significance threshold of $\alpha = 0.01$ had been chosen before conducting the experiment, would the null hypothesis be rejected?
- (d) What is the probability that, out the 5 cups the lady selects, exactly 4 are made with tea first, and 1 is made with milk first?
- (e) What is the probability that out of the 5 cups the lady select, at least 4 of the 5 cups are made with tea first?
- (f) Suppose the experiment takes place and, out of the 5 cups the lady selects, exactly 4 are made with tea first. If no significance threshold α had been chosen prior to conducting the experiment, would you accept the lady's claim that she can taste the difference between the two processes of making tea (i.e. would you reject the null hypothesis)?

Question 2

Suppose there is a class of 300 students with heights denoted by random variables X_1, X_2, \dots, X_n , which are assumed to follow a normal distribution with unknown mean θ and unknown variance σ^2 . This class of students seems particularly tall, and they wish to show that their average height $\theta > \theta_0 = 180\text{cm}$. A sample of 20 students volunteers to have their heights measured, and the sample mean is computed to be 181.5cm, and the sample variance is computed to be $s^2 = 9$.

- (a) Specify the null and alternative hypotheses for this experiment.
- (b) For a significance threshold of $\alpha = 0.01$, compute the relevant critical threshold value for this test.
- (c) Compute the appropriate test statistic and decide whether or not to reject the null hypothesis.
- (d) Notice that the students whose heights were measured were volunteers, i.e. they offered to have their heights recorded for the experiment. Are there any potential issues with this method of data collection? If so, how else could the data have been collected?

Question 3

Suppose the heights of two groups of people are recorded. Group A consists of n people and their heights are recorded (in cm) as x_1, x_2, \dots, x_n with $n = 10$, sample mean $\bar{x} = 171.5$ and sample variance $s_x^2 = 2$. Group B consists of m people and their heights are recorded as y_1, y_2, \dots, y_m , with $m = 12$, $\bar{y} = 170$ and sample variance $s_y^2 = 3$. We wish to test if the average heights of the two groups are significantly different or not. We start by assuming that the measurements x_1, x_2, \dots, x_n are observations of the independent random variables X_1, X_2, \dots, X_n , respectively, which follow a normal distribution with unknown mean μ_1 and unknown variance σ_1^2 . We also assume that the y_1, y_2, \dots, y_m are observations of the independent random variables Y_1, Y_2, \dots, Y_m , respectively, following a normal distribution with unknown mean μ_2 and unknown variance σ_2^2 . We also assume that although the variances are unknown, they are equal i.e. $\sigma_1^2 = \sigma_2^2 = \sigma^2$.

- (a) What is the null hypothesis for this test?
- (b) Assuming the null hypothesis is true, use Student's two-sample t -test to compute a test statistic.
- (c) Using Student's t table from your notes, obtain an upper bound for the p -value, and decide whether or not the average heights of the two groups are significantly different or not.