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)?

Solution to Question 1

Part (a) :

The null hypothesis is :

$$H_0 : \text{The lady has no ability to discriminate between the two processes for making tea.}$$

Of course different words can be used, e.g. 'discriminate' can be swapped with 'distinguish', etc. But the important point is that the null hypothesis is that the lady **no ability** to taste the difference between the cups of tea.

When the experiment is run, if the cups she identifies as being made with tea first actually have a high proportion of tea-first cups, this might provide evidence to reject the null hypothesis and therefore support her assertion that she is able to taste the difference between the two ways of making tea with milk.

Part (b) :

The number of ways of choosing 5 cups out of 10 is $\binom{10}{5}$.

The number of ways of choosing the 5 cups so that all 5 are tea-first, is the same as the number of ways of choosing 5 out of 5 tea-first cups, multiplied by the number of ways of choosing 0 out of 5 milk-first cups, i.e. $\binom{5}{5}\binom{5}{0}$.

Therefore, the probability of choosing 5 cups out of 10 so that all 5 cups are the 5 tea-first cups is

$$p = \frac{\binom{5}{5}\binom{5}{0}}{\binom{10}{5}} = \frac{1}{252}.$$

Part (c):

Since $\alpha = 0.01$ and

$$\frac{1}{252} = \frac{4}{1008} < \frac{4}{1000} = 0.004 < 0.01$$

then the null hypothesis would be rejected.

Part (d):

Using the explanation in Part (b),

$$p = \frac{\binom{5}{4}\binom{5}{1}}{\binom{10}{5}} = \frac{5 \cdot 5}{252} = \frac{25}{252}.$$

Part (e):

Adding the probabilities from Parts (d) and (e):

$$p = \frac{\binom{5}{4}\binom{5}{1}}{\binom{10}{5}} + \frac{\binom{5}{5}\binom{5}{0}}{\binom{10}{5}} = \frac{25+1}{252} = \frac{26}{252}.$$

Part (f):

Since,

$$p = \frac{26}{252} > \frac{26}{260} = 0.1.$$

or computing the p-value exactly,

$$p = \frac{26}{252} = 0.103,$$

there is only about a 10% chance that the lady would have chosen at least 4 out of the 5 tea-first cups by pure chance.

Since no significance threshold α was chosen at the start of the experiment, whether or not the result of the experiment is statistically significant enough for **YOU** is, literally, up to you.

For historical reasons, p-values above 0.05 are generally not regarded as statistically significant, so one could either say that the null hypothesis is not rejected at the $\alpha = 0.05$ (or even at the $\alpha = 0.1$) significance level, or one could reject the null hypothesis at the $\alpha = 0.104$ significance level, since $p < 0.104$.

Question 2

Suppose there is a class of 300 students with heights denoted by random variables $X_1, X_2, \ldots, X_n$, 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 $\theta > \theta_0 = 180$cm. A sample of 20 students volunteers to have their heights measured, and the sample mean is computed to be 182cm, 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?

Solution to Question 2

This question is very similar to Example 4.4.5 in the notes, except we use the t-distribution.

Part (a)

$$\begin{aligned} H_0 &: \theta \leq 180, \\ H_1 &: \theta > 180. \end{aligned}$$

Part (b)

Because the null hypothesis is one-sided, we need to take care when choosing our critical value. We first compute the critical value as $t_{19,0.01} = -t_{19,0.99} = -2.539$, using the symmetry of the t-distribution probability density function.

Part (c)

Suppose that we measured $\overline{x} = 181.5$. Then

$$t = \frac{\theta_0 - \overline{x}}{s/\sqrt{n}} = \frac{180 - 181.5}{3/\sqrt{20}} = -2.236.$$

Since $t \not< t_{19,0.01}$ (or $|t| \not> t_{19,0.01}$), the test statistic is not in the rejection region and we would fail to reject the null hypothesis, i.e. we would not declare the average heights of the students to be larger than 180cm.

Incidentally, we can compute the p-value exactly using R:

```
print(pt( (180-181.5)/(3/sqrt(20)), df=19))
#> [1] 0.01877
```

So the p-value is approximately $p = 0.019$. Not that while we would not reject the null hypothesis at the 0.01 level, we would have rejected the null hypothesis at the 0.05 level.

Part (d)

Since the class want to show that their average height is above 180cm, it is possible that mostly tall students volunteered to have their heights recorded.

Rather, the students to have their heights recorded should have been randomly selected from the class. Of course, each student would have to consent to having their height measured.

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 $x_1, x_2, \ldots, x_n$ with $n = 10$, sample mean $\overline{x} = 171.5$ and sample variance $s_x^2 = 2$. Group B consists of m people and their heights are recorded as $y_1, y_2, \ldots, y_m$, with $m = 12$, $\overline{y} = 170$ 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 $x_1, x_2, \ldots, x_n$ are observations of the independent random variables $X_1, X_2, \ldots, X_n$, respectively, which follow a normal distribution with unknown mean μ_1 and unknown variance σ_1^2. We also assume that the $y_1, y_2, \ldots, y_m$ are observations of the independent random variables $Y_1, Y_2, \ldots, Y_m$, 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.

Solution to Question 3

Part (a):

The null hypothesis is that the two means are equal, i.e.

$$H_0 : \mu_1 = \mu_2$$

Part (b):

Using the hint,

$$s_p^2 = \frac{1}{10 + 12 - 2} \left((9)2 + (11)3\right) = \frac{51}{20}$$

Furthermore,

$$\sqrt{\frac{1}{10} + \frac{1}{12}} = \sqrt{\frac{22}{120}} = \sqrt{\frac{11}{60}}$$

Under the null hypothesis $\mu_1 - \mu_2 = 0$. Then, the observed value of the statistic is

$$t = \frac{\overline{x} - \overline{y}}{s_p \sqrt{\frac{1}{n} + \frac{1}{m}}} = \frac{171.5 - 170}{\sqrt{\frac{51}{20}} \sqrt{\frac{11}{60}}}$$

$$= \frac{1.5}{\sqrt{\frac{17 \cdot 3}{20}} \sqrt{\frac{11}{3 \cdot 20}}} = \frac{1.5}{\frac{1}{20}\sqrt{17 \cdot 11}}$$

$$= \frac{30}{\sqrt{17 \cdot 11}} = 2.193817 \qquad \text{(using a calculator)}$$

Part (c):

In this case, we have not specified a significance threshold α, but want to use the table. If we look at the table for the cumulative distribution function of the t-distribution, in the row for $n+m-2=20$, we see that

$$\mathrm{P}\left(T < 2.086\right) = 0.975, \qquad \text{and} \qquad \mathrm{P}\left(T < 2.528\right) = 0.99.$$

Since $t = 2.193817$ falls between these two values, we can say that $0.02 < p < 0.05$.

(Remember, $1 - \alpha/2 = 0.975 \Rightarrow \alpha = 0.05$)

Therefore, we can reject the null hypothesis at significance level 0.05, but not at the level 0.02.

We can use R to get an exact p-value, after converting the two-sided 'p-value' to a one-sided p-value.

```r
ptilde <- pt(2.193817, df=20)
p <- 1 - 2 * abs(ptilde - 0.5)
cat("p-value is: ", p, "\n")
#> p-value is:  0.04023
```

We see the value of $0.04023 < 0.05$