This is the ninth homework assignment. Students should tick in TUWEL problems they have solved and upload their detailed solutions by **20:00** on Wednesday **December 8, 2021**.

In the two sample situation throughout consider Welchâs statistics.

(1) **Two-sample t-test**

Create two independent samples from the normal distribution. The first sample of size 10 shall be taken from the $N(0, 1)$-distribution. The second sample of size 20 shall be taken from the $N(1, 1)$ distribution. Test the null hypothesis that the populations means are equal with a (two-sided) two-sample t-test at the 5%-significance level:

(a) Calculate the t-statistic (without t.test())

(b) Compare it to the output of t.test()

(c) Interpret the result of the test

(2) **Two-sample t-test using normal approximation**
Messages are frequently sent from a sender to either receiver 1 or receiver 2. For both receivers, several times for the transfer were measured (in seconds) and stored in the file waitingtimes2.Rdata.

(a) Visualize both data sets. Are the distributions approximately bell-shaped?

(b) Test the null-hypothesis of equal mean transfer times for both receivers on the 1%-level with a two sample t-test (using the normal approximation).

(c) Compare your result to the output of t.test()

(3) **Duality Theorem: equivalence of hypothesis testing and confidence intervals**
For a two-sided, two-sample test, show that we reject the null hypotheses $H_0 : d = d_0$ if and only if the confidence interval does not overlap d_0.

(4) **Paired Testing [harder]**
We have assumed until now that our X_i and Y_i samples are independent. This is often not the case. An extreme case of this is when the samples are *paired*. For example, a tire company is interested in comparing the longevity of their tires to a competitor. Therefore, they get 20 volunteers (drivers) to put the company's tires on the left side of their cars and a lead competitor's tires on the right side. The volunteers then drive for 6 months, after which the company measures the amount of wear on all of the tires. For simplicity, we will only be comparing the wear of front tires. Let X_i, $i = 1, 2, \ldots, n$ be the amount of wear of the left tire, and Y_i, $i = 1, 2, \ldots, n$ be the wear of the right tire. Data for this setting is available in the file tireData.Rdata, where column 1 contains the X_i and column 2 contains the Y_i. Wear is measured in terms of percent reduction in tread height.

(a) Visualize the distributions of wear for the left and right tires using a histogram.

(b) Conduct a two-sample t-test for the tires having equal longevity (ie, same wear for their tires and those of the competitor). Specify the null and alternative hypotheses. Perform the test at level .05.

(c) We know, however, that the wear of the tires is connected...they were on the same car! A better measure is actually to look at the difference, for each i, between the wear of the tires. Construct a new vector as $d_i = X_i - Y_i$ and conduct the one-sample hypothesis test that the mean difference is zero at level .05.

(d) Compute the standard deviation used in the test statistics of the previous two parts. Using these values, discuss the difference in power between the two tests and why one is preferred to the other.

(5) **Which statement is correct?**
For a two-sided, two-sample t-test at the 5%-level, suppose that we rejected the null hypothesis $H_0 : d = d_0$. Comment on the following statements.

(a) The null hypothesis would also be rejected at the 7%-level

(b) The equivalent 99% confidence interval does not contain d_0.

(c) If both sample sizes are increased by a factor 4, then the value of the t-statistic is halved (if all other estimates remain the same)

(d) If one of the sample sizes is increased, then the width of the 95%-confidence interval is increased (if all other estimates remain the same)

(e) There is a 5% chance that the null hypothesis is true

(f) If the null hypothesis is true, there is a 5% chance that we came to the wrong conclusion.

(6) **Simulation of test-power**

Simulate the test-power in the two-sample t-test: Let $X_1, \ldots, X_n$, $Y_1, \ldots, Y_n$ be independent random variables with $X_i \sim N(0, \sigma^2)$ and $Y_i \sim N(d, \sigma^2)$ for all $i = 1, 2, \ldots, n$. Let the null hypothesis be $H_0 : d = 0$ and the significance level be $\alpha = .05$. Simulate the test-power (relative frequency of rejections) for $d \in \{5, 4.5, 4, \ldots, 5\}$ in 1000 simulations each. Use the parameters

(a) $n = 10$ and $\sigma = 3$

(b) $n = 20$ and $\sigma = 3$

(c) $n = 20$ and $\sigma = 1$