University of Tehran,
College of Engineering

Homework 5 *(Mandatory Questions)*

Introduction to Statistical Interference

Instructor: **Dr. Mohammadreza A. Dehaqani**

Erfan Panahi (*Student Number:* **810103084**)

Problem 2. Comparison of Flashlight Battery Lifetimes

In this question we should compare the lifetimes of four brands of flashlight batteries (A, B, C, and D) using ANOVA to determine if there is a significant difference in the mean lifetime among the brands at the 5% significance level.

The lifetimes of the batteries, to the nearest hour, are shown in Table 2.1:

Brand	D	C	B	A
1	20	24	28	42
2	32	36	36	30
3	38	28	31	39
4	28	28	32	28
5	25	33	27	29

Table 2.1. Lifetimes of flashlight batteries (in hours)

1. **Basic conditions of using this test and checking them.**
   Independence: Each brand is tested in separate flashlights (no overlap).
   Normality: Stated in the problem (data from normal populations).
   Equal Variances: Stated in the problem (equal standard deviations).

2. **State the null and alternative hypotheses.**
   Null Hypothesis (H_0): $\mu_A = \mu_B = \mu_C = \mu_D$
   Alternative Hypothesis (H_1): At least one mean differs.

3. **Construct the ANOVA table.**
   First, we should calculate the sample mean and sample variance of each brand. The results are shown in table 2.2.

Brand	D	C	B	A
Sample mean	28.6	29.8	30.8	33.6
Sample variance	45.3	16.2	11.2	36.3

Table 2.2. the sample mean and sample variance of each brand

Then, we should calculate the degrees of freedom and mean squares between and within groups:

Between groups: $\mathrm{df} = 4 - 1 = 3$ $\quad$ $\mathrm{MSB} = 22.83$
Within groups: $\mathrm{df} = 20 - 4 = 16$ $\quad$ $\mathrm{MSW} = 27.25$

The final ANOVA table is shown in table 2.3.

Source	DF	Mean Squares	F-value	p-value
Sample mean	3	22.83	0.838	0.493
Sample variance	16	27.25		

Table 2.3. the ANOVA table

The critical value for F (from F-table) is $F_{0.05}(3,16) = 3.24$. Since $F = 0.838 < 3.24$, we **fail to reject H_0**. **No significant difference in mean battery lifetimes among the four brands** $(p > 0.05)$.

Problem 5. Controlling False Discovery Rate in Multiple Comparisons

In this question, given p-values from an experiment, we should use the Benjamini-Hochberg (BH) method to control the false discovery rate (FDR) at 5%.

Table 5.1 shows the p and critical value for each rank.

Rank	p-value	Critical value (BH threshold)= $\frac{Rank}{8} \times p$
1	0.005	0.00625
2	0.009	0.0125
3	0.019	0.01875
4	0.022	0.025
5	0.051	0.03125
6	0.101	0.0375
7	0.361	0.04375
8	0.387	0.05

Table 5.1. Compute Critical Values

Part 1. Identify Significant p-values

In this part we should find ranks (i) that $p_i < \text{Critical Value}_{\text{i}}$. As it is shown in the table 5.1, **The first four p-values ($\mathbf{0.005}$, $\mathbf{0.009}$, $\mathbf{0.019}$, and $\mathbf{0.022}$) are significant** as they fall below their respective critical values (BH thresholds).

Part 2. p-Value Chart

The p-value chart according to their rank and the cut-off line are shown in figure 5.1.

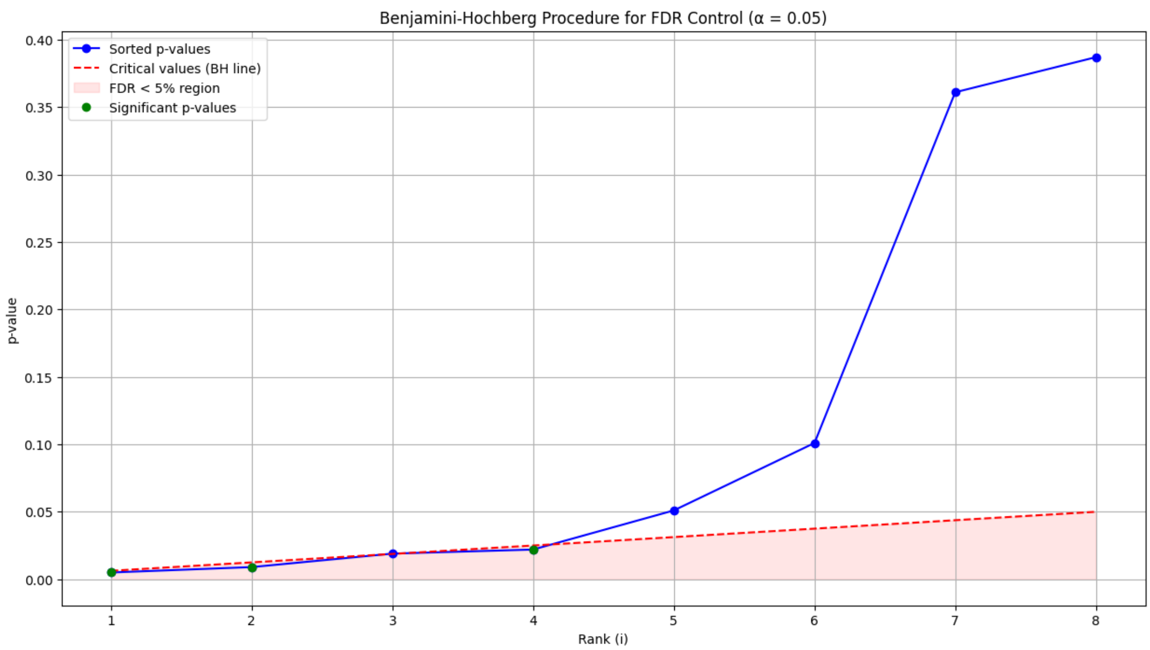

Figure 5.1. P-value Chart

Part 3. FDR vs. FWER

- **FDR (Benjamini-Hochberg):** Controls the proportion of false discoveries among the rejected hypotheses.
- **FWER (Bonferroni):** Controls the probability of making any false discovery. It is more conservative and reduces power.

Problem 6. Effect of Aromas on Anagram Solving Efficiency

In this question, we should analyze the effect of four different aroma treatments on anagram-solving efficiency using one-way ANOVA and Tukey's method.

Part 1. Hypotheses (One-Way ANOVA)

Null Hypothesis (H_0): All aroma treatments have the same mean efficiency.

$$\mu_{\text{Lemon}} = \mu_{\text{Floral}} = \mu_{\text{Fried}} = \mu_{\text{Control}}$$

Alternative Hypothesis (H_1): At least one aroma treatment has a different mean efficiency.

To solving this part, we use **Python** code. The result is:

$$F = 17.73684, \qquad p = 0.00068$$

Since the p-value is less than the significance level ($\alpha = 0.05$), **we reject the null hypothesis**. There is a statistically significant difference in the mean number of anagrams solved across the aroma treatments.

Part 2. Tukey's method

To determine which specific groups are different, we will use Tukey's Honestly Significant Difference (HSD) test. To solving this part, we use **Python** code. The result is shown in figure 6.1.

```
Multiple Comparison of Means - Tukey HSD, FWER=0.05
====================================================
 group1 group2 meandiff p-adj   lower   upper  reject
----------------------------------------------------
Control Floral      5.0 0.0054  1.7099  8.2901   True
Control  Fried  -1.3333  0.589 -4.6234  1.9568  False
Control  Lemon      4.0 0.0192  0.7099  7.2901   True
 Floral  Fried  -6.3333 0.0012 -9.6234 -3.0432   True
 Floral  Lemon     -1.0 0.7679 -4.2901  2.2901  False
  Fried  Lemon   5.3333 0.0037  2.0432  8.6234   True
----------------------------------------------------
```

Figure 6.1. Multiple Comparison of Means - Tukey HSD

As observed, by calculating the P-value, if this value is less than 0.05, the null hypothesis is rejected, and there is a significant difference in odor between the specified pair of groups. Otherwise, the null hypothesis cannot be rejected. According to the problem's data, only the groups **Floral** and **Lemon**, as well as **Lemon** and **Fried Food**, have a $p > 0.05$. The remaining pairs of groups have a significant difference from each other because they have $p < 0.05$.

Result: Significant differences were found between the following pairs: **<u>Fried Food vs Lemon</u>, <u>Fried Food vs Floral</u>.**

Part 3. Limitations of Tukey's Procedure

- **Assumption of Homogeneity of Variances:** Tukey's HSD assumes that the variances of the groups are equal. If this assumption is violated, the results may not be reliable.
- **Sensitivity to Outliers:** Tukey's HSD can be sensitive to outliers, which can affect the results.

Problem 7. Bootstrapping for a Problem of Ants' Population

In this question, through bootstrapping, we want to estimate the mean number of ants and construct a confidence interval to understand the variability and reliability of our estimate.

The number of ants counted in each experiment is as follows (table 7.1):

Experiment Number	1	2	3	4	5	6	7	8
Number of Ants	43	5	77	70	36	42	15	71

Table 7.1. The number of ants counted in each experiment

Part i. Mean and Standard Deviation of the Sample

Mean (μ): $\mu = \frac{53+5+77+70+36+42+15+71}{8} = \frac{359}{8} = \mathbf{44.875}$

Standard Deviation (σ): $\sigma = \sqrt{\frac{\sum_i (x_i - \mu)}{n}} = \sqrt{\frac{4918.625}{8}} = \sqrt{614.8281} \approx \mathbf{24.8}$

Part ii. Creating Bootstrap Statistics

To create a Bootstrap Statistics using 8 pieces of paper:

1. Write each of the 8 observed ant counts on separate pieces of paper.
2. Place all pieces in a container and randomly draw one, record the number, and replace it.
3. Repeat this process 8 times to form one bootstrap sample.
4. Calculate the mean of this bootstrap sample.
5. Repeat the entire process many times (e.g., 1000 times) to create a bootstrap distribution of means.

Part iii. Shape, Distribution, and Center of Bootstrap Distribution

- **Shape:** The bootstrap distribution is expected to be approximately normal due to the Central Limit Theorem, especially with a large number of bootstrap samples.
- **Distribution:** The distribution will center around the sample mean (44.875).
- **Center:** The center of the bootstrap distribution will be close to the original sample mean.

Part iv. Parameter and Best Estimate

- **Parameter:** The population mean (μ) is a good parameter to represent the population of interest.
- **Best Estimate:** The best estimate for the population mean is the sample mean, which is 44.875.

Part v. 95% Confidence Interval Using Standard Error

Given a standard error (SE) of 4.85: $(\alpha = 5\%)$

$$\text{Confidence Interval} = \left(\mu \pm z_{\frac{\alpha}{2}} \cdot \mathrm{SE}\right) = (44.875 \pm 1.96 \times 4.85) = (44.875 \pm 9.506) = (35.37, 54.38)$$

We are $\mathbf{95\%}$ **confident** that the true population mean number of ants that would gather on a piece of sandwich near an ant nest lies between approximately $\mathbf{35.37}$ and $\mathbf{54.38}$.

Problems 9. Simple Linear Regression

In this question, given data for alloy attributes vs temperature, perform a simple linear regression.

Considering the table 9.1, which pertains to a simple linear regression problem where the values of Y correspond to a specific attribute of several types of alloys at temperatures X,

i	1	2	3	4	5	6	7	8
x_i	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
y_i	40	41	43	42	44	42	43	42

Table 9.1. Alloy Attribute vs Temperature

Part i. Parameter Estimation using Maximum Likelihood Estimation

We will estimate the parameters β_0 (intercept), β_1 (slope), and σ^2 (variance) using maximum likelihood estimation (MLE).

- $\bar{x} = \frac{\sum x_i}{n} = 4 \times \frac{4+0.5}{8} = 2.25$
- $\bar{y} = \frac{\sum y_i}{n} = 40 + \frac{1+3+2+4+2+3+2}{8} = 42.125$

Calculating β_1 and β_0:

- $\beta_1 = \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum(x_i-\bar{x})^2} = \frac{5.75}{10.5} \approx \mathbf{0.5476}$
- $\beta_0 = \bar{y} - \beta_1\bar{x} = 412.125 - 0.5476 \times 2.25 \approx \mathbf{40.8929}$

For calculating σ^2 we use $\hat{y}_i = \beta_0 + \beta_1 x_i$. The table 9.2 shows the $\hat{y}_i$ and residuals $e_i = y_i - \hat{y}_i$.

i	1	2	3	4	5	6	7	8
x_i	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
y_i	40	41	43	42	44	42	43	42
$\hat{y}_i$	41.17	41.44	41.71	41.99	42.26	42.54	42.81	43.08
e_i^2	1.36	0.19	1.65	0.00	3.02	0.29	0.04	1.17

Table 9.2. $\hat{y}_i$ and residuals e_i^2

Finally, we can calculate the variance (σ^2):

$$\sigma^2 = \frac{1}{n}\sum(y_i - \hat{y}_i)^2 = \frac{1}{8} \times 7.73 \approx \mathbf{0.97}$$

Now, we want to calculate the variance related to β_0, β_1:

$$\sigma_{\beta_1}^2 = \frac{\sigma^2}{\sum(x_i-\bar{x})^2} \approx \mathbf{0.092}$$

$$\sigma_{\beta_0}^2 = \sigma^2\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum(x_i-\bar{x})^2}\right) \approx \mathbf{0.586}$$

Part ii. Correlation Between β_0 and β_1

$$\mathrm{corr}(\hat{\beta}_0, \hat{\beta}_1) = -\frac{\bar{x}}{\sqrt{\sum(x_i-\bar{x})^2}} = -\frac{2.25}{10.5} \approx \mathbf{-0.6942}$$

Part iii. Unbiased Estimator for θ ($\hat{\theta} = 5 + \hat{\beta}_0 - 4\hat{\beta}_1$)

Using the estimated values:

$$\hat{\theta} = 5 + \hat{\beta}_0 - 4\hat{\beta}_1 = 5 + 40.8929 - 4 \times 0.5476 = 43.7025$$

Mean Squared Error (MSE) of $\hat{\boldsymbol{\theta}}$:

$$\begin{aligned}\operatorname{var}(\hat{\theta}) = \operatorname{var}(\hat{\beta}_0) + 16\operatorname{var}(\hat{\beta}_1) - 8\operatorname{cov}(\hat{\beta}_0, \hat{\beta}_1) = \operatorname{var}(\hat{\beta}_0) + 16\operatorname{var}(\hat{\beta}_1) - 8\operatorname{corr}(\hat{\beta}_0, \hat{\beta}_1) \cdot \sigma_{\beta_1} \cdot \sigma_{\beta_2} \\ = 0.586 + 16 \times 0.092 - 8 \times \underbrace{(-0.6942) \times 0.586 \times 0.092}_{-0.161} = 3.346\end{aligned}$$

So, the **MSE of $\hat{\boldsymbol{\theta}}$ is approximately $\mathbf{3.346}$**.

Part iv. Optimal c for Minimum MSE ($\hat{\theta} = 5 + \hat{\beta}_0 - c\hat{\beta}_1$)

$$\begin{aligned}\text{MSE} = \operatorname{var}(\hat{\theta}) = \operatorname{var}(\hat{\beta}_0) + c^2 \cdot \operatorname{var}(\hat{\beta}_1) - 2c \cdot \operatorname{cov}(\hat{\beta}_0, \hat{\beta}_1) \rightarrow \frac{d}{dc}\operatorname{var}(\hat{\theta}) \\ = 2c \cdot \operatorname{var}(\hat{\beta}_1) - 2\operatorname{cov}(\hat{\beta}_0, \hat{\beta}_1) = 0\end{aligned}$$

$$c = \frac{\operatorname{cov}(\hat{\beta}_0, \hat{\beta}_1)}{\operatorname{var}(\hat{\beta}_1)} = \frac{-0.161}{0.092} \approx -1.75$$

So, the MSE reaches its minimum when $c \approx -1.75$.

Part v. Prediction for New Sample ($x = 3.25$)

Predicted value: $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x = 40.8929 + 0.5476 \times 3.25 \approx 42.6726$

MSE of Prediction: The MSE of the prediction is the variance of the prediction error.

$$\operatorname{var}(\hat{y}) = \sigma^2 \left(1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum (x_i - \bar{x})^2}\right) = 0.97 \left(1 + \frac{1}{8} + \frac{(3.25 - 2.25)^2}{10.5}\right) \approx 0.97 \times 1.2202 = 1.1836$$

Part vi. Optimal x for Minimum MSE

The MSE of the **prediction is minimized when $\boldsymbol{x = \bar{x}}$**, because the term $(x - \bar{x})^2$ is minimized (zero) at this point. Therefore, the predicted sample of alloy will reach the minimum MSE when $x = 2.25$.