University of Tehran,
College of Engineering

Homework 4

Introduction to Statistical Interference

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

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

Problem 1. Hospital Recovery Status Analysis

The summary of the data is provided in Table 1.1:

Recovery Status	Private Hospital	Public Hospital	Total
Fully Recovered	30	25	**55**
Partially Recovered	20	15	**35**
Not Recovered	10	10	**20**
Total	**60**	**50**	**100**

Table 1.1. Observed Recovery Outcomes by Hospital Type

Hypothesis Formulation:

- **Null Hypothesis (H_0):** Recovery outcomes are independent of hospital type.
- **Alternative Hypothesis (H_A):** Recovery outcomes depend on hospital type.

To solve this problem, we need to perform a Chi-Square Test of Independence using the data provided in the contingency table.

Independence Test:

Expected frequency for a cell in a contingency table can be calculated using the formula:

$$E_{ij} = \frac{\text{Row total} \times \text{Column total}}{\text{Grand total}}$$

- For E_{11} (Fully Recovered, Private Hospital): $E_{11} = \frac{55\times60}{100} = 33$
- For E_{12} (Fully Recovered, Public Hospital): $E_{12} = \frac{55\times50}{100} = 22$
- For E_{13} (Partially Recovered, Private Hospital): $E_{21} = \frac{35\times60}{100} = 21$
- For E_{12} (Partially Recovered, Public Hospital): $E_{22} = \frac{35\times50}{100} = 17.5$
- For E_{12} (Not Recovered, Private Hospital): $E_{31} = \frac{60\times20}{100} = 12$
- For E_{12} (Not Recovered, Public Hospital): $E_{32} = \frac{20\times50}{100} = 10$

The calculated expected frequencies are shown in table 1.2.

Recovery Status	Private Hospital	Public Hospital
Fully Recovered	33	22
Partially Recovered	21	17.5
Not Recovered	12	10

Table 1.2. The Expected Frequency Table

Now I calculate the Chi-Square statistic:

$$\chi^2 = \sum \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}$$

Where O_{ij} is the observed frequency and E_{ij} is the expected frequency.

$$\chi^2 = \frac{(33-30)^2}{33} + \frac{(25-22)^2}{22} + \frac{(20-21)^2}{21} + \frac{(15-17.5)^2}{17.5} + \frac{(10-12)^2}{12} + \frac{(10-10)^2}{10}$$
$$= \frac{3}{11} + \frac{9}{22} + \frac{1}{21} + \frac{5}{14} + \frac{1}{3} \approx 1.42$$

$$\text{Degrees of Freedom} = (r-1)\times(c-1) = (3-1)\times(2-1) = 2$$

$$\boldsymbol{\chi^2 = 1.42}, \qquad \textbf{df} = \mathbf{2}$$

At a significance level $\alpha = 0.05$ and $\text{df} = 2$, the critical value from the Chi-Square distribution table is **5.991**.

Decision Rule: If $\chi^2 \leq 5.991$, fail to reject H_0; otherwise, reject H_0.

Result: $\chi^2 = 1.42$ is less than 5.991. Since $\chi^2 = 1.42 \leq \text{critical value } (5.99)$, we fail to reject the null hypothesis. Thus, the data does not provide sufficient evidence to conclude that recovery outcomes depend on hospital type.

Critical Evaluation:

The small expected frequency in the "Not Recovered" category for the private hospital may affect the reliability of the chi-square test. A larger sample size or alternative statistical methods (e.g., Fisher's exact test) could improve the analysis.

- The small expected frequencies for some cells (e.g., $E_{13}= 12$) do not violate assumptions of the test.
- Future studies could include a larger sample size or additional hospital types for more robust results.

Problem 2. Multiple Choice vs Written Response

The summary of the data is provided in Table 2.1:

Student	Multiple-Choice	Written-Response	Difference
1	25	20	5
2	30	27	3
3	15	12	3
4	32	29	3
5	28	25	3
6	24	22	2
7	29	30	-1
8	26	24	2
9	31	28	3
10	33	30	3
11	22	18	4
12	12	10	2

Table 2.1. Number of problems solved by students in two different formats.

Hypothesis Formulation:

- **Null Hypothesis (H_0):** There is no difference in performance between the two formats. ($\mu_{\text{diff}} = 0$)
- **Alternative Hypothesis (H_A):** There is a significant difference in performance between the two formats. ($\mu_{\text{diff}} \neq 0$)

To analyze this data, we perform a **paired t-test** because the data consists of paired observations (each student solves problems in both formats). This will test if there is a significant difference between the two formats (Multiple-Choice and Written-Response).

First, I compute the mean ($\bar{d}$) and standard deviation s_d of the difference:

$$\bar{d} = \frac{\sum d}{n} = \frac{5+3+3+3+3+2-1+2+3+3+4+2}{12} = \frac{32}{12} = 2.67$$

$$s_d = \sqrt{\frac{\sum(d_i - \bar{d})}{n-1}} = 1.42$$

Now, I calculate the t-statistic:

$$t = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}} = \frac{2.67}{\frac{1.42}{\sqrt{12}}} = 6.51$$

For a two-tailed test at a significance level $\alpha = 0.05$ with $n - 1 = 11$ degrees of freedom, the critical t-value is approximately ± 2.201.

Decision Rule: If $|t| > 2.201$, reject H_0; otherwise, fail to reject H_0.

Result: $t = 6.51$ is greater than 2.201. There is strong evidence to suggest that students perform differently depending on the format of the problems presented.

On average, students solve more problems in the multiple-choice format ($\bar{d} > 0$).

Problem 3. Teaching Methods and Mathematics Scores

Hypothesis Formulation:

- **Null Hypothesis (H_0):** There is no difference in effectiveness between the teaching methods.
- **Alternative Hypothesis (H_A):** The effectiveness of the teaching methods differs.

This problem involves using the **Mann-Whitney U test** (or a rank-sum test) to compare the two teaching methods (A and B). The steps are outlined below:

First, I combine the scores from both groups and assign ranks to the values in ascending order. Ties are assigned the average rank.

- **Group A:** $75, 78, 82, 88, 84, 90, 77, 85, 79, 83, 80, 81$
- **Group B:** $88, 86, 90, 92, 89, 85, 87, 91, 90, 88, 89, 86$

Combined Data (sorted and ranked):

$$75(1), 77(2), 78(3), 79(4), 80(5), 81(6), 82(7), 83(8), 84(9), 85(10.5), 85(10.5), 86(12.5), 86(12.5),$$

$$87(14), 88(16), 88(16), 88(16), 89(18.5), 89(18.5), 90(21), 90(21), 90(21), 91(23), 92(24)$$

Next, I sum the ranks for each group:

- **Ranks for Group A ($n_A = 12$):**

$$R_A = 1 + 2 + 3 + 7 + 9 + 21 + 4 + 10.5 + 5 + 8 + 6 + 16 = 91.5$$

- **Ranks for Group B ($n_B = 12$):**

$$R_B = 12.5 + 12.5 + 14 + 16 + 16 + 18.5 + 18.5 + 21 + 21 + 21 + 23 + 24 = 154.5$$

Now, I calculate the test statistic (W):

$$W = R_A = 91.5$$

1. Calculate the expected mean of W:

$$\mu_W = \frac{n_A(n_A + n_B + 1)}{2} = \frac{12(25)}{2} = 150$$

2. Calculate the standard deviation of W:

$$\sigma_W = \sqrt{\frac{n_A n_B (n_A + n_B + 1)}{12}} = \sqrt{\frac{12.12.25}{12}} = \sqrt{300} \approx 17.32$$

3. Calculate Z-Statistic:

$$Z = \frac{W - \mu_W}{\sigma_W} = \frac{91.5 - 150}{17.32} \approx -3.38$$

At a significance level of $\alpha = 0.05$, the critical Z-value for a two-tailed test is ± 1.96.

Decision Rule: If $|Z| > 1.96$, reject H_0; otherwise, fail to reject H_0.

Result: $Z = -3.38$ is less than -1.96. There is sufficient evidence to reject the null hypothesis. The effectiveness of the teaching methods differs significantly.

Group B (median rank sum: $\mathbf{154.5}$) appears to outperform Group A (median rank sum: $\mathbf{91.5}$).

Problem 4. Education vs. Income Relationship

First, I rank the values of both variables (education and income) separately. Assign average ranks in case of ties. The ranks, rank differences, and squared differences for education and income are displayed in the table 4.1. ($d_i = \text{Rank(Education)} - \text{Rank(Income)}$)

Individual	Education (Years)	Rank of Education	Monthly Income (x1000)	Rank of Income	d_i	d_i^2
1	10	9.5	2.5	9	0.5	0.25
2	12	7.5	3.0	8	-0.5	0.25
3	15	4	4.0	3.5	0.5	0.25
4	8	12	1.8	12	0	0
5	16	3	4.5	1	2	4
6	11	8	2.8	10	-2	4
7	14	5	4.2	2	3	9
8	9	11	2.1	11	0	0
9	13	6	3.5	6	0	0
10	12	7.5	3.0	8	-0.5	0.25
11	14	5	4.0	3.5	1.5	2.25
12	10	9.5	2.6	7	2.5	6.25

Table 4.1. Education and Income Data

This problem involves computing Spearman's rank correlation coefficient (r_s) to determine if a monotonic relationship exists between education and monthly income.

Hypothesis Formulation:

- **Null Hypothesis (H_0):** No monotonic relationship ($r_s = 0$).
- **Alternative Hypothesis (H_A):** A monotonic relationship exists ($r_s \neq 0$).

The formula for Spearman's rank correlation coefficient is:

$$r_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)} = 1 - 6 \times \frac{26.5}{12(144-1)} \approx 0.907$$

Use a significance level of $\alpha = 0.05$. For $n = 12$, the critical value of r_s for a two-tailed test is approximately ± 0.591 (from Spearman's rank correlation tables).

Decision Rule: If $|r_s| > 0.591$, reject H_0; otherwise, fail to reject H_0.

Result: Since $r_s = 0.907 > 0.591$, we reject H_0. There is sufficient evidence to suggest a strong positive monotonic relationship between education level and monthly income. This indicates that higher education levels are associated with higher monthly income.

Problems 5. Mean IQ of students

To solve this problem, we need to test whether the mean IQ of students in the population of interest is higher than 107, using the sample provided and a 0.05 level of significance.

The summary of the data is provided in Table 1.1:

Student	1	2	3	4	5	6	7	8
IQ	100	90	135	108	107	119	127	109

Student	9	10	11	12	13	14	15
IQ	105	112	95	123	98	110	120

Table 5.1. IQ scores of students from a rural area.

Hypothesis Formulation:

- **Null Hypothesis (H_0):** Recovery outcomes are independent of hospital type.
- **Alternative Hypothesis (H_A):** Recovery outcomes depend on hospital type.

To solve this problem, we need to perform a Chi-Square Test of Independence using the data provided in the contingency table.

First, we compute the sample mean ($\bar{x}$) and standard deviation (s).

$$\bar{x} = \frac{1}{n}\sum x_i = 110.53\,, \qquad s = \sqrt{\frac{\sum(x_i - \bar{x})}{n-1}} = 12.45$$

The t-statistic is calculated as:

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{110.53 - 107}{\frac{12.45}{\sqrt{15}}} = 1.1$$

Using a t-distribution table, the critical value for t at the 0.05 level of significance with $n - 1 = 14$ degrees of freedom is approximately 1.761.

Since the calculated t (1.1) is less than the critical value (1.761), **we fail to reject the null hypothesis**. There isn't enough evidence to conclude that the mean IQ of students aged 16 or older from the population of interest is higher than 107 at the 0.05 level of significance.

Problem 6. Difference in SmartWatches

To determine whether there is a statistically significant difference in step-counting accuracy between two smartwatch brands (Brand X and Brand Y). Data was collected from 10 participants, who used both brands, and their step counts (steps per minute) were recorded.

The step counts recorded for each participant using the two smartwatch brands are as table 6.1.

Participant	Brand X	Brand Y	Difference ($\Delta = X - Y$)	$\lvert\Delta\rvert$	Rank
Alex	120	118	2	2	2.5
Bailey	134	136	-2	2	2.5
Casey	128	130	-2	2	2.5
Drew	140	137	3	3	7
Ellis	145	143	2	2	2.5
Frankie	132	130	2	2	2.5
Gale	138	141	-3	3	7
Harper	125	122	3	3	7
Jordan	130	133	-3	3	7
Kelly	135	132	3	3	7

Table 6.1. Step-counting accuracy (steps per minute) for two smartwatch brands.

Since the step-counting accuracy data is not normally distributed, a non-parametric test was chosen. The **Wilcoxon Signed-Rank Test** is appropriate for comparing paired samples to evaluate whether their distributions differ.

Hypothesis Formulation:

- **Null Hypothesis (H_0):** There is no difference in step-counting accuracy between Brand X and Brand Y.
- **Alternative Hypothesis (H_A):** There is a difference in step-counting accuracy between Brand X and Brand Y.

$$\text{sum of positive ranks} = 2.5 + 7 + 2.5 + 2.5 + 7 + 7 = 28.5$$

$$\text{sum of negative ranks} = 2.5 + 2.5 + 7 + 7 = 19$$

The test statistic is the smaller of the two sums of ranks: $T = 19$

For 10 pairs, the critical value for the Wilcoxon Signed-Rank Test at the 5% significance level is approximately 8 (you can find critical values in statistical tables).

Since the test statistic (17) is greater than the critical value (8), we **fail to reject the null hypothesis**. This means there isn't enough evidence at the 5% significance level to conclude that there is a difference in accuracy between the two brands of smartwatches.

Problem 7. Coached and Independent group

To determine whether there is a statistically significant difference in the number of attempts required to achieve 10 successful free throws between the coached and independent practice groups.

The number of attempts required to achieve 10 successful free throws for each participant in both groups is as table 7.1.

Group	Number of Attempts
Coached Group	12, 15, 10, 18, 14
Independent Group	20, 22, 17, 19

Table 7.1. Number of attempts required to achieve 10 successful free throws.

Since the sample size is small and it is not stated whether the data follows a normal distribution, a non-parametric test is chosen. The **Mann-Whitney U Test** is appropriate for comparing two independent groups when the normality assumption may not hold.

Hypothesis Formulation:

- **Null Hypothesis (H_0):** There is no difference in the distribution of the number of attempts required between the coached and independent group.
- **Alternative Hypothesis (H_A):** There is a difference in the distribution of the number of attempts required between the coached and independent groups.

First, I combine the data from both groups and rank the values in ascending order. Ties are averaged. The results are shown in table 7.2.

Value	Group	Rank
10	Coached	1.0
12	Coached	2.0
14	Coached	3.0
15	Coached	4.0
17	Independent	5.0
18	Coached	6.0
19	Independent	7.0
20	Independent	8.0
22	Independent	9.0

Table 7.2. Number of attempts required to achieve 10 successful free throws.

$$\textbf{Coached Group}: \text{sum of ranks} = 1 + 4 + 2 + 6 + 3 = 16$$

$$\textbf{Independent Group}: \text{sum of ranks} = 8 + 9 + 5 + 7 = 29$$

Now, I compute the test statistic:

$$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1 = 19$$

$$U_2 = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2 = 1$$

The test statistic is the smaller of the two U values, so $U = 1$.

At a 5% significance level, the critical value for the Mann-Whitney U test for $n_1 = 5$ and $n_2 = 4$ is 2.

Since $U = 1$ is less than the critical value (2), we **reject the null hypothesis**. This means there is a significant difference in the number of attempts required between the coached and independent groups.