Lesson X Practical Exercise

STUDENT NAME

School of Public Affairs: Pennsylvania State University

PADM 504, Section 001: Data Analysis for Policy and Administration

Professor Lauren Azevedo

DATE

**Purpose**

Orient the reader to the purpose of your analysis. Provide the research question and how you are answering it.

**Method**

**Equality of two variance (F-test)**

The F-test (Snedecor and Cochran, 1983) is used to evaluate if the variances of two populations are equal. This test can be either two-tailed or one-tailed. The two-tailed version compares the variances to the alternative of not being equal. The F hypothesis test is defined as follows:

: vs :

Where, are population variance of two groups.

The test statistics for this test procedure is,

F = ~ if is true.

Where, are sample variance of two groups.

If the p-value of the corresponding F statistic is less than 0.05, the F test shows that there is sufficient evidence to reject the null hypothesis that the two group variances are equal at the 5% level of significance.

**Independent sample t-test**

The independent sample t-test is a statistical technique used to compare the means of two independent groups. We may compare the means of two samples from the same population by using the independent samples t-test. When samples are drawn from two independent populations, the sample mean may differ. It is used in this case to draw conclusions about the means of two populations and assess whether they are similar.

**Assumptions in independent samples t-test:**

1. Assumes that the dependent variable is normally distributed.
2. Observations are independent of each other.

The t test hypothesis is,

(The means of the two groups are not significantly different)

(The means of the two groups are significantly different)

Where, are two populations mean.

If the variance of the two groups is equal, then the test statistic for the above hypothesis is defined as,

t = ~ if the is true.

Where, = is the pooled variance of two sample and are sample means of two groups.

If the variance of the two groups is not equal, then the test statistic for the above hypothesis is defined as,

t = ~ if the is true.

where, v = .

The t test indicates that there is enough evidence to reject the null hypothesis that the two group means are equal at the 0.05 significance level if the p-value of the corresponding t statistic is less than 0.05 (significance level).

**Missing value**

In the presence of missing value, the t-test can produce misleading results. To achieve a correct result, the missing values are removed using the R programming language (version 4.2.2). “IAP”, “DK” & “NA” values are excluded from the HRS1 variable as they are missing value.

**Results**

To compare the average working hours between male workers & female workers, firstly, it should be determined that the variance of the two groups are equal or not. Based on the result of variance test, independent sample t test should be performed.

Table-01: F test to compare the variances of male & female group

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variables | Sex | Variances | df | F statistic | p-value |
| Working hours | Male | 224.058 | 669 | 1.2713 | 0.0016\*\* |
| Female | 176.2493 | 710 |  |  |

Note: \*\*\*p-value < 0.0001. \*\*p-value < 0.01. \*p-value < 0.05.

It is observed from the above table that the variance of the working hours in male & female groups is significantly different at 5% level of significance.

Since it is obtained that the variance of the two groups is significantly different at 5% level of significance, independent sample t-test with unequal variance should be used to compare the average working hours between male & female groups.

Table-02: Independent sample t-test for equality of male & female group.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variables | Sex | Means | t statistic | df | p-value |
| Working hours | Male | 44.60 | 8.4463 | 1336 | 0.0000\*\*\* |
| Female | 38.15 |  |  |  |

Note: \*\*\*p-value < 0.0001. \*\*p-value < 0.01. \*p-value < 0.05.

It is observed from the above table that mean of the working hours in male & female group is significantly different at 5% level of significance. That means, the male work, on average, more hours than female in 2018.

**Discussion and Policy Implications/Recommendations**

Most of the practical exercises ask you to offer policy implications or recommendations based on your findings. First discuss the overall meaning of your analysis and then present your policy implications and recommendations to the reader.

**References** (if used)