|  |  |  |  |
| --- | --- | --- | --- |
| **Data Analytics for Managers (MBA 135)** CIA 3, Data collection and application of suitable statistical techniques | | | |
| Hemanath Kumar J **(21121015)** | Prajwal Divate **(21121035)** | Submission Date: 16/10/2021 |
| **Hypothesis Testing of Academic Factors influencing MBA scores** | | | |

**Problems:**

* To find out which factors influenced a candidates MBA score.
* To find out whether previous scores matter for one to score higher in MBA degree.

**Source:**

<https://www.kaggle.com/benroshan/factors-affecting-campus-placement>

It includes secondary and higher secondary school percentage and specialization. It also includes degree specialization, type and Work experience and salary offers to the placed students.

**Hypothesis Testing**

Hypothesis testing in statistics is a way for you to test the results of a survey or experiment to see if you have meaningful results. You’re basically testing whether your results are valid by figuring out the odds that your results have happened by chance. If your results may have happened by chance, the experiment won’t be repeatable and thus won’t have any value.

Hypothesis testing is used to assess the plausibility of a hypothesis by using sample data. Such data may come from a larger population, or from a data-generating process. The word "population" will be used for both of these cases in the following descriptions.

# **Key Terms and Concepts**

* **Null hypothesis:** Null hypothesis is a statistical hypothesis that assumes that the observation is due to a chance factor.  The null hypothesis is denoted by; H0: μ1 = μ2, which shows that there is no difference between the two population means.
* **Alternative hypothesis:** Contrary to the null hypothesis, the alternative hypothesis shows that observations are the result of a real effect.
* **Level of significance:** Refers to the degree of significance in which we accept or reject the null-hypothesis.  100% accuracy is not possible for accepting or rejecting a hypothesis, so we, therefore, select a level of significance that is usually 5%.

**Analysis of variance** (**ANOVA)**

Analysis of variance is a collection of statistical models and their associated estimation procedures (such as the "variation" among and between groups) used to analyse the differences among group means in a sample. ANOVA was developed by the statistician Ronald Fisher. The ANOVA is based on the law of total variance, where the observed variance in a particular variable is partitioned into components attributable to different sources of variation. In its simplest form, ANOVA provides a statistical test of whether two or more population means are equal, and therefore generalizes the *t*-test beyond two means.

**ANOVA testing 1**

This ANOVA test is performed on the UG degree type and is being compared to MBA score.



From the above data,

**Null hypothesis:** There is no relationship between UG degree type and score in MBA.

**Alternative hypothesis:** There is a relationship between UG degree type and score in MBA.

**Inference:**

Since the P-value in the above data is greater than 0.05 and F value is lesser than F crit,

**Null hypothesis is true.**

**The alternate hypothesis is rejected.**

**Regression 1**



**Inference:**

**R² = 0.1619,** from R² value, we can say that change in degree type does some change in MBA score. But the rate of change is not very high. (16% of Y is dependent upon X)

**ANOVA testing 2:**

This ANOVA test is performed on the HSC group type and is being compared to MBA score.



From the above data,

**Null hypothesis:** There is no relationship between HSC group and score in MBA.

**Alternative hypothesis:** There is a relationship between HSC group and score in MBA.

**Inference:**

Since the P-value in the above data is greater than 0.05 and F value is lesser than F crit,

**Null hypothesis is true.**

**The alternate hypothesis is rejected.**

**Regression 2:**



**Inference:**

**R² = 0.1259,** from R² value, we can say that change in HSC group does some change in MBA score. But the rate of change is not very high.(12% of Y is dependent upon X)

**Overall summary:**

**Problem 1:** To find out which factors influenced a candidates MBA score.

From the hypothesis testing done, we can conclude that the external factors influenced MBA score are Degree type, Degree score, HSC group and HSC score. These factors slightly influenced the MBA score.

**Problem 2:** To find out whether previous scores matter for one to score higher in MBA degree.

From the hypothesis testing done, we can conclude that degree score and HSC score slightly contributed to the percentage scored in MBA.