REPORT ON STATISTICAL ANALYSIS

Nikhil Doifode
SB ID#: 112714121

INTRODUCTION

The purpose of this assignment is to understand the basics of statistical significance testing. We are analyzing given datasets and performing Analysis of Variance (ANOVA) and Multivariate analysis of variance (MANOVA) tests.

EXPERIMENT 1

Task - In the 1st test dataset we have 4 group of users each with 10 users who navigate in the given menu type.

Number of Conditions - The types of menus are Tool glass, Flow menu, Tool palette, and Control menu. We have given time each user takes to navigate the certain menu type.

Variables:

Independent Variables (IV) - Menu Types are independent variables.

Dependent Variables (DV) - Time taken by a user to navigate the menu is dependent variable.

Hypothesis Testing: We consider Null hypothesis to be true unless statistics allow us to reject it.

Alternate Hypothesis - Manipulation of IV (menu type) affects DV (time) in some way. Not all mean time are equal and they do not occur by chance.

Null hypothesis - Manipulation of IV (menu type) has no effect on DV (time). Mean time of all the menu types are equal

Experimental Design - We have to follow "Between-group design" for 1st dataset in which different group of user use different type of menu. Also we have to make sure that we avoid carry-over learning effect between users.

Statistical Analysis

Descriptive summary statistics:

We compute central tendencies each independent variable (Menu Type) by visualizing Mean and Standard Deviation against time taken for user to navigate them. We will use R base graphs and **ggpubr** R package for data visualization.

In the figure 1, each point shows the data point which is time taken by user to navigate on that menu type, **center point of the bar for each menu type shows mean time and the distance between top and bottom of the bar indicates the Standard Deviation (SD)**. Therefore, for Tool Glass the mean of time taken is 2.446 and SD is 0.24518474 whereas for Flow menu mean is 2.598 and SD is 0.35894289. For Tool Palette and Control Menu mean is 2.775 and 2.388 and SD is 0.17815412 and 0.35574023 respectively.
This also shows that the control menu took the least time to navigate the menu 1.75s, which took 1.4 seconds quicker than the Tool Palette which took greatest amount of time (3.15s) to navigate.

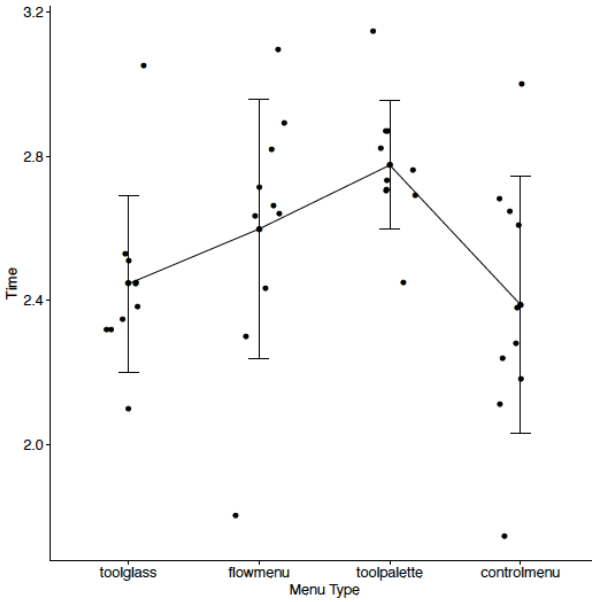

Figure 1: Data points, Mean and Standard Deviation for time taken

In the figure 2, the **line between the box is median** of the time taken for each menu type and the bottom line of box indicates the 1st Quartile and the top line of box indicates the 3rd Quartile of time.

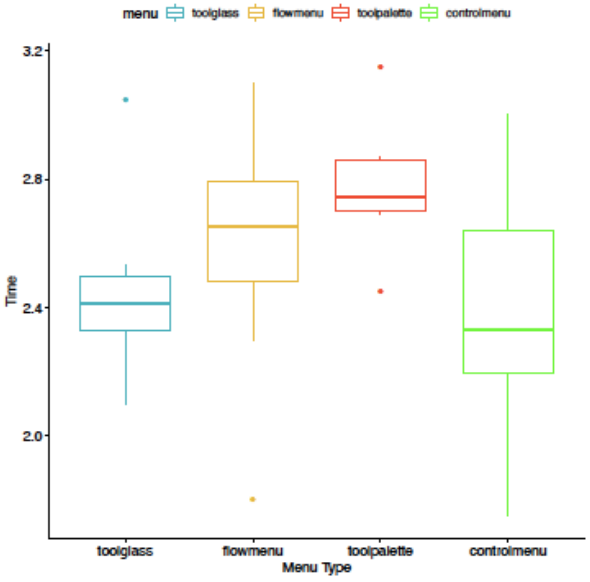

Figure 2: Median of time taken by users on each menu type

STATISTICAL TEST

ANOVA

We follow ANOVA for the analysis of variance, the data is organized into several groups based on menu type. Here we want to find the effect of independent variable on one continuous dependent variable. We use the inbuilt R function **aov**() to find results. We consider all the menu types once for testing. We use the function **summary.aov**() is used to summarize the analysis of variance model.

```
> summary(resultAnova)
            Df Sum Sq Mean Sq F value Pr(>F)
menu         3 0.8998 0.29992   3.455 0.0264 *
Residuals   36 3.1252 0.08681
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
```

Figure 3: Summary of results of ANOVA test

In the following figure we show the histogram of the residuals from the ANOVA tests.

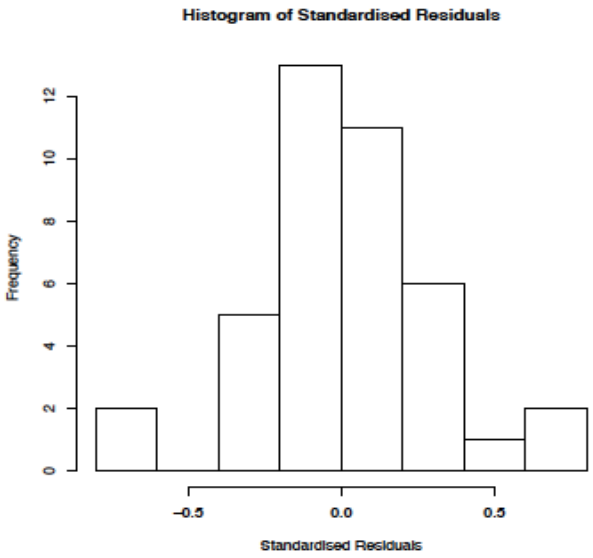

Figure 4: Histogram of Standardized Residuals

Interpreting the result of ANOVA test

The output includes the columns F value and Pr(>F) corresponding to the p-value of the test. P Value shows the likelihood that the results are due to chance variation. P Value < 0.05 usually considered insignificant. It means that there is < 5% chance that null hypothesis is true.

As the F(3,36 = 3.455, p = 0.0264) is less than the significance level 0.05, we can conclude that there are significant differences between the menus highlighted with “*" in the model summary hence NULL hypothesis is not true and changing of Independent Variable (Menu Type) has significant effect on Dependent Variable (Time taken) to navigate the menu type. Hence the time mean values for all the menu type are not equal and they do not occur by chance as there is a significant difference among them.

Pairwise T - Testing

ANOVA tests whether means differ between menu types, but does not tell us by which means differ, for this we must perform pairwise t-tests. We first make use of the pairwise.t.test() function from R, which has the following arguments. x: DV, g: IV and p.adjust.method: the p-value adjustment method used to control. We use Bonferroni Adjustment method.
The Bonferroni adjustment simply divides the Type I error rate (.05) by the number of tests. The Bonferroni adjustment can be made using p.adjust.method = “bonferroni” in the pairwise.t.test() function.

Null Hypothesis: The mean values of time are equal between two Menu items.

Alternative Hypothesis: The mean values of time are not equal between two Menu items.

```
> pairwise.t.test(dsSet1$time, dsSet1$menu, p.adjust.method = "bonferroni")

        Pairwise comparisons using t tests with pooled SD

data:  dsSet1$time and dsSet1$menu

            controlmenu flowmenu toolglass
flowmenu    0.718       -        -
toolglass   1.000       1.000    -
toolpalette 0.034       1.000    0.103

P value adjustment method: bonferroni
```

Figure 5: Pairwise T-Testing using Bonferroni

Using the Bonferroni adjustment, only the tool palette – control menu comparison is statistically significant (0.034 < 0.05) hence we can reject null hypothesis for them. Hence, there is a significant difference in the time mean values between control menu-tool palette where as we cannot reject the null hypothesis for the other pairs as their p-value is greater than 0.05 significance value.

EXPERIMENT 2

Task - In the 2nd dataset each user uses each menu type. We are given data of 10 users who are tested for different menu types. In total we have 40 results where each user’s time and error is recorded for each menu type.

Number of Conditions - The types of menus are Tool glass, Flow menu, Tool palette, and Control menu.

Variables:
Independent Variables (IV) - Menu Types.

Dependent Variables (DV) - Time taken by a user to navigate the menu and Error Rate are dependent variables.

Hypothesis Testing: We consider Null hypothesis to be true unless statistics allow us to reject it.

Alternate Hypothesis - Manipulation of IV (menu type) affects DV (time) in some way. The means of at least 2 groups are different and there is significant difference between the means between or within groups.

Null hypothesis - Manipulation of IV (menu type) has no effect on DV (time). The means within groups and between groups are same and there is no significant difference between the groups.

Experimental Design - We have to follow "With-in subject design" for 2nd dataset in which all users try all menu types. We compare one user across all menu types to isolate effects of individual diffs.

Statistical Analysis

Descriptive summary statistics:

We compute central tendencies each independent variable (Menu Type) by visualizing Mean and Standard Deviation (SD) against time taken for user to navigate them. We will use R base graphs and **ggpubr** R package for data visualization. Since we have two dependent variables hence we have to calculate mean, SD and median for Time and Error Separately.

In the figure 6, each point shows the data point, center point of the bar for each menu type shows mean time and the distance between top and bottom of the bar indicates the Standard Deviation (SD). Therefore, for Tool Glass the mean of time taken is 2.504 and SD is 0.24753 whereas for Flow menu mean is 2.398 and SD is 0.3686. For Tool Palette and Control Menu mean is 2.73 and 2.414 and SD is 0.182 and 0.3482 respectively. This also shows that the control menu took the least time to navigate the menu 1.75s, which took 1.4 seconds quicker than the Tool Palette which took greatest amount of time (3.15s) to navigate.

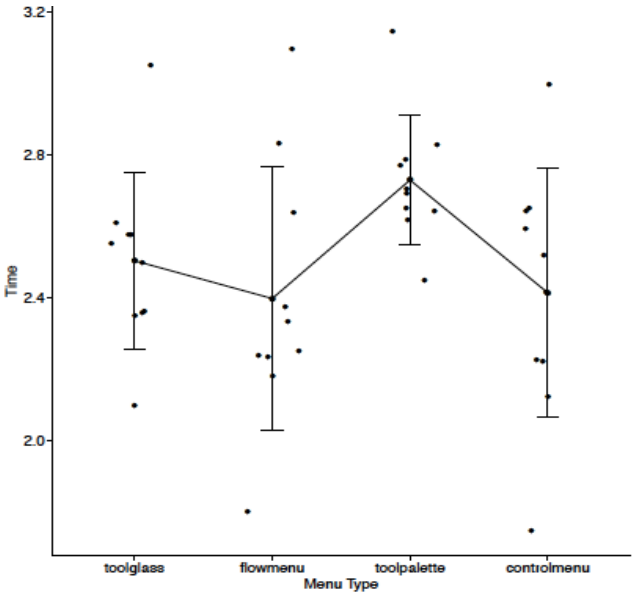

Figure 6: Mean and SD for Time and Menu

In the figure 7, the line between the box in median of the time taken for each menu type and the bottom line of box indicates the 1st Quartile and the top line of box indicates the 3rd Quartile of time.

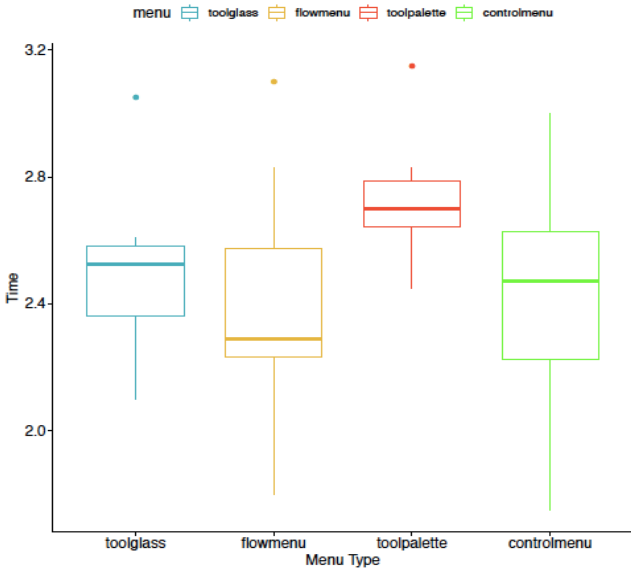

Figure 7: Median for Time and Menu

In the figure 8, each point shows the data point, center point of the bar for each menu type shows mean error and the distance between top and bottom of the bar indicates the Standard Deviation (SD). Therefore, for Tool Glass the mean of error is 4.1 and SD is 2.02 whereas for Flow menu mean is 19.7 and SD is 8.95. For Tool Palette and Control Menu mean is 12.8 and 19.7 and SD is 7.94 and 6.7 respectively. This also shows that the tool glass has least error of 1, which is 30 higher than the Flow menu which has highest (31).

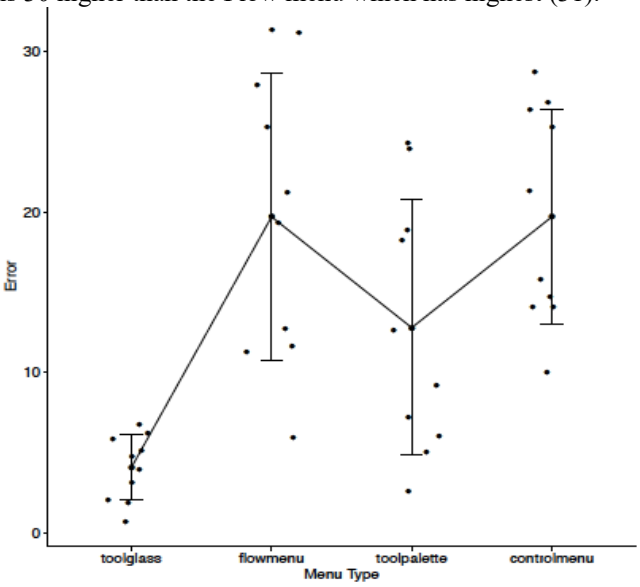

Figure 8: Mean and SD for Error and Menu

In the figure 9, the line between the box is median of the error for each menu type and the bottom line of box indicates the 1st Quartile and the top line of box indicates the 3rd Quartile of time.

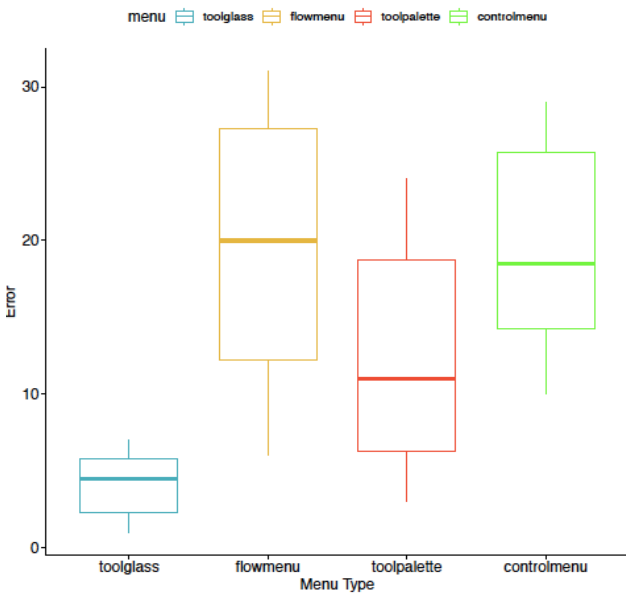

Figure 9: Median for Error and Menu

STATISTICAL TEST

MANOVA

We follow MANOVA for the multivariate analysis of variance. It is an extension of Analysis of Variance (ANOVA) to several dependent variables. The approach to MANOVA is similar to ANOVA in many regards and requires the same assumptions. Here we have multiple response variables and we can test them simultaneously using a multivariate analysis of variance (MANOVA). We use the inbuilt R function **manova**() to find results.

```
> dsSet2$user <- factor(dsSet2$user)
> resultManova <- manova(cbind(time, error) ~ menu + Error(user/menu), data = dsSet2)
> summary(resultManova)

Error: user
          Df Pillai approx F num Df den Df Pr(>F)
Residuals  9

Error: user:menu
          Df Pillai approx F num Df den Df    Pr(>F)
menu       3 1.0924   10.834      6     54 6.978e-08 ***
Residuals 27
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 10: Summary of results of MANOVA test

Interpreting the result of MANOVA test

The output includes the columns F value and Pr(>F) corresponding to the p-value of the test. P Value shows the likelihood that the results are due to chance variation. P Value < 0.05 usually considered insignificant. It means that there is < 5% chance that null hypothesis is true.

As the F(3,36 = 10.834, p = 0.00000006978) is way less than the significance level 0.05, we can conclude that there are significant differences between the menus highlighted with “*" in the model summary hence NULL hypothesis is not true and changing of Independent Variable (Menu Type) has significant effect on Dependent Variable (Time taken and Error) to navigate the menu type. Therefore, on the basis of this result we reject the Null Hypothesis and accept the Alternative Hypothesis.

Pairwise T - Testing

Since, the same users are used for different menu types we will perform Paired Pairwise t-test. The paired t-test reduces inter-subject variability (because it makes comparisons between the same subject), and thus is theoretically more powerful than the unpaired t-test. Here since we have two dependent variables hence we have to do pairwise t-testing 2 times. Once for Time-Menu and for Error-Menu.

Null Hypothesis: The mean values of time are equal between two Menu items.

Alternative Hypothesis (Ha) : The mean values of time are not equal between two Menu items.

```
> pairwise.t.test(dsSet2$time, dsSet2$menu, p.adjust.method = "bonferroni", paired = TRUE)

        Pairwise comparisons using paired t tests

data:  dsSet2$time and dsSet2$menu

            controlmenu flowmenu toolglass
flowmenu    1.0000      -        -
toolglass   1.0000      0.4600   -
toolpalette 0.0030      0.0099   0.0026

P value adjustment method: bonferroni
```

Figure 11: Pairwise T-Testing for Time-Menu using Bonferroni

Figure 11 shows the table of paired t-test result for the Time Variable and Menu Types. From the table we can say that tool palette-control menu, tool palette-tool glass and tool palette-flow menu have significant difference between them and reject the Null Hypothesis as the p-value = 0.003, 0.0026 and 0.0099 are < 0.05 respectively. For all the other groups of Menu types the p-values are greater than 0.05 and hence there is no significant difference between the mean time values for these menu types.

Now, we perform the similar test with the Error variable

Null Hypothesis: The mean values of error are equal between two Menu items.

Alternative Hypothesis: The mean values of error are not equal between two Menu items.

```
> pairwise.t.test(dsSet2$error, dsSet2$menu, p.adjust.method = "bonferroni", paired = TRUE)

        Pairwise comparisons using paired t tests

data:  dsSet2$error and dsSet2$menu

            controlmenu flowmenu toolglass
flowmenu    1.00000     -        -
toolglass   0.00018     0.00140  -
toolpalette 0.57649     0.54757  0.08495

P value adjustment method: bonferroni
```

Figure 12: Pairwise T-Testing for Error-Menu using Bonferroni

Figure 12 shows the table of paired t-test result for the Error Variable and Menu Types. From the table we can say that tool glass-control menu and flow menu-tool glass have significant differences between them and reject the Null Hypothesis as the p-value = 0.00018 and 0.0014 < 0.05 respectively. For all the other groups of Menu types the p-values are greater than 0.05 and hence there is no significant difference between the mean error values for these menu types.