CSE 518, Human Computer Interaction Assignment 2

Mayukh Maitra
SBU ID: 112025294
NY, USA
mayukh.maitra@stonybrook.edu

ANOVA AND PAIRWISE T-TEST

The one-way analysis of variance (*ANOVA*), also known as one-factor *ANOVA*, is an extension of independent two-samples t-test for comparing means in a situation where there are more than two groups. In one-way *ANOVA*, the data is organized into several groups base on one single grouping variable (also called factor variable) [2].

ANOVA test hypotheses:

- *Null hypothesis*: the means of the different groups are the same
- *Alternative hypothesis*: At least one sample mean is not equal to the others.

How ANOVA test works

- Compute the common variance, which is called variance within samples S^2_{within}or residual variance.
- Compute the variance between sample means as follows:
  - Compute the mean of each group.
  - Compute the variance between sample means $S^2_{between}$.
  - Produce F-statistic as the ratio of $S^2_{between}/S^2_{within}$.

  From the results of *ANOVA* test, a lower ratio $(ratio < 1)$ indicates that there are no significant difference between the means of the samples being compared. However, a higher ratio implies that the variation among group means are significant.

Dataset

The dataset used for the experiment is a csv file consisting of 40 records. The dataset contains user id, type of menu and time. There are a total of 40 users, 10 each for a particular menu type. So, there are 4 groups of users and it's a between-group design.

From table 1 we can see how the four groups' data is distributed in the dataset.

The same can be seen from Fig 1 and Fig 2 below. It is observed that *toolpalette* has the highest mean and the least amount of standard deviation whereas *flowmenu* sees the maximum amount of standard deviation.

		High level descriptives	
Menu	*count*	*mean*	*standard deviation*
controlmenu	10	2.39	0.356
flowmenu	10	2.60	0.359
toolglass	10	2.45	0.245
toolpalette	10	2.78	0.178

Table 1. Data distribution by menu items

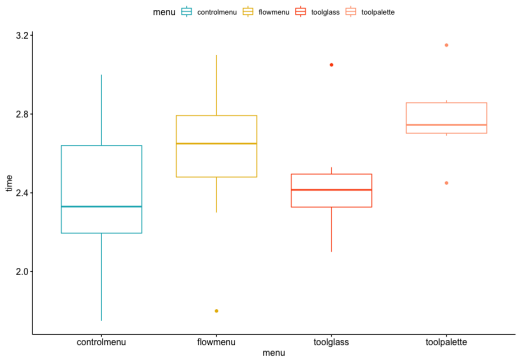

Figure 1. Distribution of data by groups

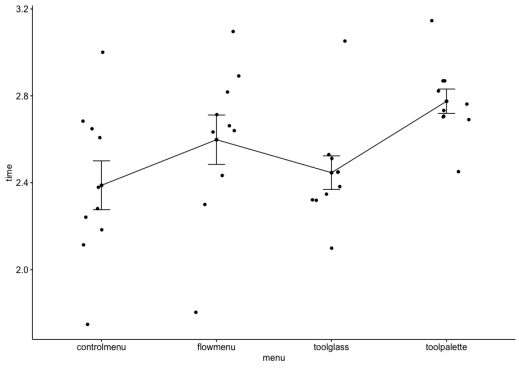

Figure 2. Distribution of data by mean of groups

ANOVA test experiment

The *R* inbuilt function *aov*() has been used to calculate the *ANOVA* between groups. The function *summary*() is used to summarize the analysis of variance model.

ACM ISBN 978-1-4503-2138-9.
DOI: 10.1145/1235

	ANOVA test results				
	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		

Table 2. Anova test results

From table 2 we can say that as the $p-value$ is less than the significance level 0.05, we can conclude that there are significant differences between the groups.

The residuals versus fits plot can be used to check the homogeneity of variances. In the plot below in Fig 3, there is no evident relationships between residuals and fitted values (the mean of each groups), which is good as we can assume the homogeneity of variances.

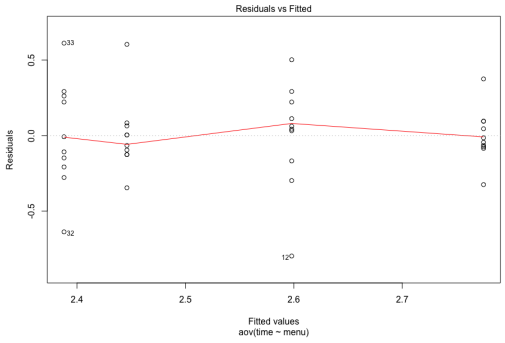

Figure 3. Residence vs Fitted observation

In the plot below in Fig 4, the quantiles of the residuals are plotted against the quantiles of the normal distribution. A 45-degree reference line is also plotted. The normal probability plot of residuals is used to check the assumption that the residuals are normally distributed. It should approximately follow a straight line. From Fig 1 we can see that most of the points fall approximately along this reference line, we can assume that there is some sort of normality.

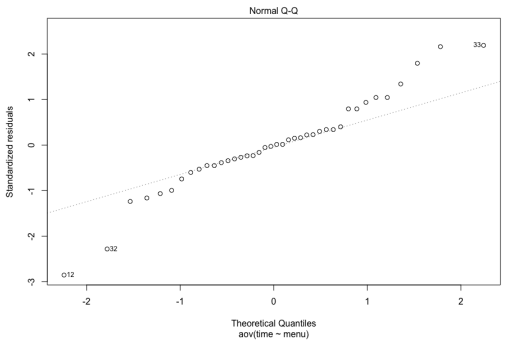

Figure 4. Normal QQ observation

Pairwise t-test

Since the p-value from anova test was significant, we were able to conduct the pairwise comparisons. To make pairwise comparisons between the menu groups, we have used the *pairwise.t.test*() function, which has the following major arguments.

1. *x*: the dependent variable
2. *g*: the independent variable
3. *p.adj*: the p-value adjustment method used to control for the family-wise Type I error rate across the comparisons; one of *"none", "bonferroni", "holm", "hochberg", "hommel", "BH",* or *"BY"*

For our experiment we have used the *bonferroni* adjustment method. The pairwise.t.test() function can be implemented as follows.

- use pairwise.t.test(x, g, p.adj) to test the pairwise comparisons between the treatment group means.

	Pairwise t-test results		
	controlmenu	*flowmenu*	*toolglass*
flowmenu	0.718	-	-
toolglass	1.000	1.000	-
toolpalette	0.034	-	0.103

Table 3. Pairwise t-test results

From table 3 above we can see that there is significant difference between *toolpalette* and *controlmenu* with $(p = 0.034)$. However, there is no difference between $controlmenu, toolglass$ and $flowmenu, toolglass$.

MANOVA AND PAIRWISE T-TEST

In the situation where there are multiple response variables, we can test them simultaneously using a multivariate analysis of variance (MANOVA)[1].

Assumptions of MANOVA

MANOVA can be used in certain conditions:

- The dependent variables should be normally distribute within groups. The *R* function *mshapiro.test*() [in the *mvnormtest* package] can be used to perform the $Shapiro-Wilk$ test for multivariate normality. This is useful in the case of MANOVA, which assumes multivariate normality.
- Homogeneity of variances across the range of predictors.
- Linearity between all pairs of dependent variables, all pairs of covariates, and all dependent variable-covariate pairs in each cell

Dataset

The dataset used for the experiment is a csv file consisting of 40 records. The dataset contains user id, type of menu, error and time. There are a total of 10 users, each user testing each menu type. It's a with-in subject design.

	High level descriptives		
Menu	*count*	*mean*	*standard deviation*
controlmenu	10	2.41	0.348
flowmenu	10	2.40	0.369
toolglass	10	2.50	0.248
toolpalette	10	2.73	0.182

Table 4. Data distribution by menu items with respect to time

From table 4 we can see how the four groups' data is distributed in the dataset with respect to time.

The same can be seen from Fig 5 and Fig 6 below. It is observed that *toolpalette* has the highest mean and the least amount of standard deviation whereas *flowmenu* sees the maximum amount of standard deviation.

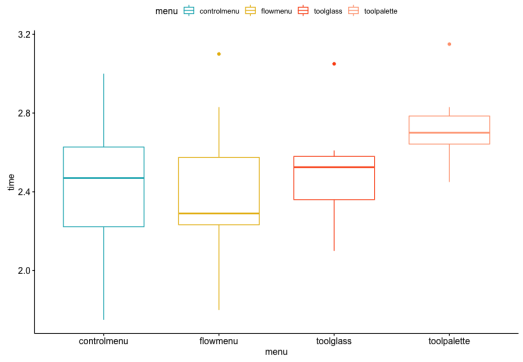

Figure 5. Distribution of data by groups with respect to time

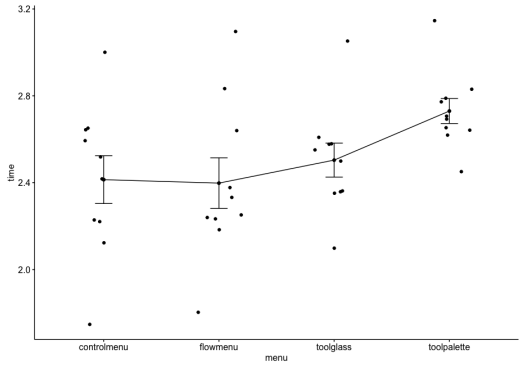

Figure 6. Distribution of data by mean of groups with respect to time

From table 5 we can see how the four groups' data is distributed in the dataset with respect to error.

The same can be seen from Fig 7 and Fig 8 below. It is observed that *toolglass* has the lowest mean error and the least amount of standard deviation whereas *flowmenu* and *controlmenu* has the maximum amount of mean error.

	High level descriptives		
Menu	*count*	*mean*	*standard deviation*
controlmenu	10	19.7	6.70
flowmenu	10	19.7	8.96
toolglass	10	4.1	2.02
toolpalette	10	12.8	7.94

Table 5. Data distribution by menu items with respect to time

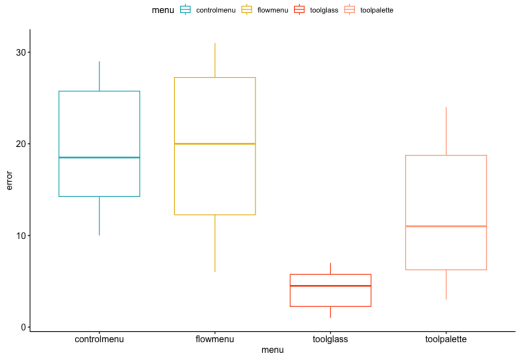

Figure 7. Distribution of data by groups with respect to error

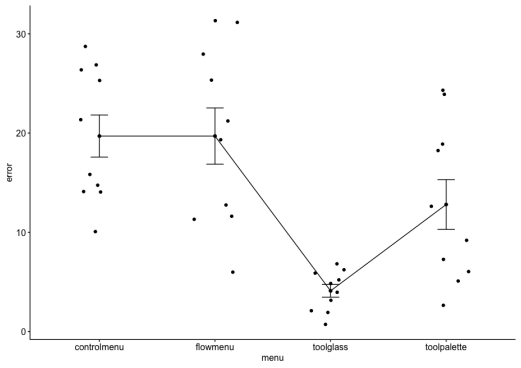

Figure 8. Distribution of data by mean of groups with respect to error

MANOVA test experiment

The *R* inbuilt function *manova*() has been used to calculate the *MANOVA* between groups. The function *summary*() is used to summarize the analysis of variance model.

	ANOVA test results					
	DF	*Pillai*	*approx F*	*num DF*	*den DF*	*Pr(>F)*
menu	3	0.6502	5.7804	6	72	5.918e-05
residuals	36					

Table 6. Manova test results

From table 6 we can say that the global multivariate test is significant as the $p-value$ is greater than the significance level 0 and we can conclude that the corresponding effect is significant.

Thus, we want to identify the specific dependent variables that contributed to the significant global effect *i.e* see how the time and error has contributed to the significance. To do that I did an *ANOVA* on top of the *MANOVA* result.

ANOVA test results for menu and time

	DF	*Sum Sq*	*Mean Sq*	*F value*	*Pr(>F)*
menu	3	0.7019	0.233957	2.6616	0.06267
residuals	36	3.1644	0.087901		

Table 7. Anova test results on manova

ANOVA test results for menu and error

	DF	*Sum Sq*	*Mean Sq*	*F value*	*Pr(>F)*
menu	3	1644.1	548.03	11.399	2.109e-05
residuals	36	1730.7	48.08		

Table 8. Anova test results on manova

From table 7 and 8 above it can be seen that the two variables *time* and *error* are highly significantly different among *menu* as the $p - value$ is greater than 0

Pairwise t-test

Since the p-value from *MANOVA* test was significant, we were able to conduct the pairwise comparisons. To make pairwise comparisons between the menu groups, we have used the $pairwise.t.test()$ function as we did for *ANOVA*. However, we did the pairwise t-test separately for $menu, time$ and $menu, error$.

For our experiment we have used the $bonferroni$ adjustment method.

Pairwise t-test results

	controlmenu	*flowmenu*	*toolglass*
flowmenu	1.000	-	-
toolglass	1.000	1.000	-
toolpalette	0.14	0.10	0.58

Table 9. Pairwise t-test results for menu and time

From table 9 above we can see that there is significant difference between *toolpalette* and *flowmenu* with $(p = 0.10)$. However, there is no difference between $controlmenu, toolglass, flowmenu$.

Pairwise t-test results

	controlmenu	*flowmenu*	*toolglass*
flowmenu	1.000	-	-
toolglass	8.2e-05	8.2e-05	-
toolpalette	0.195	0.195	0.048

Table 10. Pairwise t-test results for menu and error

From table 10 above we can see that there is significant difference between *toolglass*,*controlmenu* and *flowmenu* with $(p = 8.2e - 05)$. However, there is no difference between $controlmenu, flowmenu$.

REFERENCES

[1] http://www.sthda.com/english/wiki/manova-test-in-r-multivariate-analysis-of-variance.

[2] *http://www.sthda.com/english/wiki/one-way-anova-test-in-r*.