*International Journal of Latest Engineering and Management Research (IJLEMR)*
*ISSN: 2455-4847*
*www.ijlemr.com || Volume 05 - Issue 06 || June 2020 || PP. 30-35*

# Repeated Measurements ANOVA Method: An analysis on grades of students

Şenol Çelik*
**Department of Animal Science, Biometry Genetics Unit, Agricultural Faculty, Bingol University, Bingol, Turkey*

**Abstract:** The aim of the study was to evaluate the differences between the grades of the students who have taken the statistics course and mid-term, final, make-up examination during 2018-2019 academic year in Faculties of Agriculture, Health Sciences and Veterinary Science of Bingol University in Turkey using repeated measurements variance analysis.

According to the results of the study, Faculty of Veterinary Science students were found to be more successful in mid-term examination and Faculty of Health Sciences students were found to be more successful on make-up examination. Linear and quadratic trends were observed in grades. As a result of the parallelism test, the profiles presented according to Pillai's Trace, Wilks' Lambda, Hotelling's Trace, and Roy's Largest Root were not found to be parallel.

**Keywords:** Repeated measures, achievement, Parallelism Test

## Introduction

Some studies conducted in the field of education are planned to research whether there are any significant differences in terms of any variables between the measurements of individuals in various time and circumstances. For example, the effects of pilot tests applied to students at different times on success could be examined. Moreover whether there is any difference between the grades of the first, second, third, and fourth pilots tests could be examined. In other words, more than one measurement can be performed on the same subject (individual) in this type of studies.

In studies where many measurements are performed on the same subject, every individual becomes the control for himself/herself. In terms of experiment design, the logic of such experiments is to minimize the errors to be experienced due to individual differences (Winer, 1971).

In repeated measurements, variance analysis is used to examine whether there are any differences between measurement groups where more than two measurements (prior to implementation, after 1 month, 2 months... Etc.) are performed on individuals of a group.

As in other fields, there are also studies carried out using repeated measurements ANOVA method in the education field. The effects of Educational Software (DENIS) and game applications on success test (Kocaman and Cumaoglu, 2014), the effects of grades obtained from submissive behaviors scale and brief symptom inventory on student success (Anli and Sar, 2017) are some examples to these studies. Other examples are the effects of the blended learning environment on the academic success of a student (Ozerbas and Benli, 2015), student remarks on the measurement of Science of Technology course achievements with different approaches (Cetin and Cakan, 2010).

The aim of the study was to examine the change of the statistics grades of students studying in Faculties of Agriculture, Health Sciences and Veterinary Science according to examination times and the departments using repeated measurement variance analysis.

## Material and Method

The material of the study includes the grades of the students who have taken the statistics course and mid-term, final, make-up examination during 2018-2019 academic year in Faculties of Agriculture, Health Sciences and Veterinary Science of Bingol University in Turkey. Mid-term, final, make-up grades of the students in these faculties are recorded on the student information system by the academicians providing courses. As 3 different grades of every student who has taken the mid-term, final, make-up examination in 3 different times are examined, repeated measurements of dependent groups are considered. In other words, data on 3 measurements are obtained from every student taking the exams. The number of students who have taken mid-term, final, make-up examinations in Faculties of Agriculture, Health Sciences and Veterinary Science are respectively 19, 16, and 33 and 68 in total. As the same individual has taken 3 different examinations, multivariate variance analysis is used for repeated measurements. For this reason, Repeated ANOVA method is implemented.

Repeated-measures is a term used when the same participants participate in all conditions of an experiment.Sphericity refers to the equality of variances of the differences between treatment levels. SPSS produces a test known as Mauchly's test, which tests the hypothesis that the variances of the differences between conditions are equal. If, however, Mauchly's test statistic is non-significant (p >0.05) then it is reasonable to conclude that the variances of differences are not significantly different(Field, 2009). Multivariate tests are composed of Pillai's Trace, Wilks' Lambda, Hotelling's Trace and Roy's Largest Root tests.

Error terms ( $\varepsilon_{ijk}$ ) have a normal distribution that are independent of each other, with zero means and $\sum$ covariance matrix (Jeremy, 1974). $\bar{x}_i$ : is the mean vector of i.group, $\bar{x}$ : is the general mean vector, $n_i$: is the number of observations in i. group, $S_i$: is the variance-covariance vector in i. group

$$B = \sum_{i=1}^{k} n_i (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})'$$

$$W = \sum_{i=1}^{k} (n_i - 1) S_i$$

Eigenvalues of the $BW^{-1}$ matrix are $\lambda_i$ .

The largest root test statistics of Roy is the highest $\lambda_i$ value.

The modification of the Roy's test statistic is as given below. The Roy's greatest characteristic roots (Manly, 2005; Roy, 1957) test statistic is given by

$$\theta_{max} = \frac{\lambda_{max}}{1 + \lambda_{max}}$$

Pillai's Trace statistics(Manly, 2001) is given by

$$T = \sum_{i=1}^{n} \frac{\lambda_i}{(1 + \lambda_i)}$$

The Wilks statistic is a measure of the differences between groups of the centroid (vector) of means on the independent variables (Stevens, 1980). The test statistic in below can be expressed as

$$\Lambda = \prod_{i=1}^{n} (1 + \lambda_i)^{-1}$$

and using binomial theorem and Taylor series expansion (Dani, 2012) on below, it becomes;

$$\Lambda = \prod_{i=1}^{n} \left(1 - \lambda_i + \lambda_i^2 - \lambda_i^3 + \lambda_i^4 - \lambda_i^5 + \cdots\right)$$

Lawley-Hotelling trace test

$$T_0^2 = \sum_{i=1}^{s} \lambda_i$$

## Results and Discussion

Grade averages and standard deviation details of the students in Faculties of Agriculture, Health Sciences and Veterinary Science are given in Table 1.

Table 1. Descriptive statistics

| Exam | Department | $\bar{X}$ | s | N |
|---|---|---|---|---|
| Midterm | Agriculture | 48.26 | 20.99 | 19 |
| | Health Sciences | 72.50 | 14.56 | 16 |

| | | | | |
|---|---|---|---|---|
| | Veterinary | 75.36 | 7.54 | 33 |
| | General | 67.12 | 18.28 | 68 |
| Final | Agriculture | 28.63 | 20.20 | 19 |
| | Health Sciences | 23.75 | 15.36 | 16 |
| | Veterinary | 6.76 | 10.68 | 33 |
| | General | 16.87 | 17.83 | 68 |
| Make-up exam | Agriculture | 40.00 | 23.05 | 19 |
| | Health Sciences | 55.44 | 22.91 | 16 |
| | Veterinary | 23.88 | 11.33 | 33 |
| | General | 35.81 | 22.11 | 68 |

s: Standard Deviation, $\bar{X}$: Mean

Results of significance control of differences between the examination levels in terms of grades are presented in Table 2. As a result of Pillai's Trace, Wilks' Lambda, Hotelling's Trace, and Roy's Largest Root tests, the difference between grade averages of examination levels are observed to be significant (P<0.01).

$$\eta^2 = 1 - \Lambda = 1 - 0.167 = 0.833$$

This result is explained as the profiles over the groups of $1 - \Lambda$ of the total variance in combined total range will not be flat (Tabachnick and Fidell, 2013). 83.3% of the total variance in combined total range in this study is explained as the profiles on the groups will not be flat.

Table 2. Results of flatness test (between examination levels)

| Effect | | Value | F | Hypothesis df | Error df | p | Partial Eta Squared |
|---|---|---|---|---|---|---|---|
| Exam | Pillai's Trace | 0.833 | 159.617 | 2 | 64 | 0.001 | 0.833 |
| | Wilks' Lambda | 0.167 | 159.617 | 2 | 64 | 0.001 | 0.833 |
| | Hotelling's Trace | 4.988 | 159.617 | 2 | 64 | 0.001 | 0.833 |
| | Roy's Largest Root | 4.988 | 159.617 | 2 | 64 | 0.001 | 0.833 |

df: Degrees freedom

Whether different groups have different profiles is tested with the parallelism test. In repeated ANOVA method, the parallelism of the groups is examined primarily. In variance analysis with repeated measurements, exam x department interaction is the parallelism test in Repeated ANOVA. Parallelism test tests whether the difference between the exams in terms of grades changes according to the department.

Analysis results for parallelism of the profiles in terms of grades is given in Table 3. When the probability values of statistics used to test parallelism of profiles hypothesis of the grades are examined, it is observed that the profiles presented in terms of grades are not parallel according to Pillai's Trace, Wilks' Lambda, Hotelling's Trace and Roy's Largest Root. In other words, it is seen that the differences between the repeated measurements differ significantly according to the departments (p<0.001). Briefly, grades from the results of the same tests were found to be significantly different in terms of Exam x Department interaction (p<0.001). By calculating the effect portion of interaction between the grades and the department using partial eta squared,

$$partial\ \eta^2 = 1 - \Lambda^{1/2} = 1 - 0.377^{1/2} = 0.386$$

38.6% of the variance of the combined ranges is explained with the difference between the shapes of 3 different faculties. Profiles have different trends as they are not parallel to one another.

Table 3. Parallelism test

| Effect | | Value | F | Hypothesis df | Error df | p | Partial Eta Squared |
|---|---|---|---|---|---|---|---|
| Exam * Department | Pillai's Trace | 0.727 | 18.558 | 4 | 130 | 0.001 | 0.363 |
| | Wilks' Lambda | 0.377 | 20.105 | 4 | 128 | 0.001 | 0.386 |
| | Hotelling's Trace | 1.375 | 21.662 | 4 | 126 | 0.001 | 0.407 |
| | Roy's Largest Root | 1.131 | 36.772 | 2 | 65 | 0.001 | 0.531 |

Mauchly's Test of Sphericity is applied for the test of sphericity, one of the assumptions required for repeated ANOVA method (Table 4).

Tablo4. Mauchly's Test of Sphericity

| Within Subjects Effect | Mauchly's W | Approx. Chi-Square | df | p | Epsilon | | |
|---|---|---|---|---|---|---|---|
| | | | | | Greenhouse-Geisser | Huynh-Feldt | Lower-bound |
| Exam | 0.876 | 8.469 | 2 | 0.014 | 0.890 | 0.941 | 0.500 |

According to test of sphericity, sphericity condition was not metas per$Approx.\chi^2 = 8.469$and p<0.05. Therefore, Greenhouse-Geisser and Huynh-Feldt statistics are used (Table 5). As p obtained from F value is less than 0.001 in these statistics, there are significant differences between the success of students in 3 exams. Among the multiple comparison tests, the Tukey test was used to examine whether there are any differences between the grades of students in different faculties (Table 6).

Table 5. Tests of Within-Subjects Effects

| Source | | Sum of Squares | df | Mean Square | F | p | Partial Eta Squared |
|---|---|---|---|---|---|---|---|
| Exam | Sphericity Assumed | 64832.699 | 2 | 32416.349 | 178.52 | 0.001 | 0.733 |
| | Greenhouse-Geisser | 64832.699 | 1.779 | 36434.042 | 178.52 | 0.001 | 0.733 |
| | Huynh-Feldt | 64832.699 | 1.882 | 34442.52 | 178.52 | 0.001 | 0.733 |
| | Lower-bound | 64832.699 | 1 | 64832.699 | 178.52 | 0.001 | 0.733 |
| Exam * department | Sphericity Assumed | 19845.831 | 4 | 4961.458 | 27.323 | 0.001 | 0.457 |
| | Greenhouse-Geisser | 19845.831 | 3.559 | 5576.382 | 27.323 | 0.001 | 0.457 |
| | Huynh-Feldt | 19845.831 | 3.765 | 5271.572 | 27.323 | 0.001 | 0.457 |
| | Lower-bound | 19845.831 | 2 | 9922.915 | 27.323 | 0.001 | 0.457 |
| Error(exam) | Sphericity Assumed | 23605.846 | 130 | 181.583 | | | |
| | Greenhouse-Geisser | 23605.846 | 115.665 | 204.089 | | | |
| | Huynh-Feldt | 23605.846 | 122.352 | 192.933 | | | |
| | Lower-bound | 23605.846 | 65 | 363.167 | | | |

Table 6. Tukey test

| (I) department | (J) department | Mean Difference (I-J) | Std. Error | p |
|---|---|---|---|---|
| Agriculture | Health | -11.598* | 3.873 | 0.011 |
| | Veterinary | 3.632 | 3.287 | 0.515 |
| Health | Veterinary | 15.229* | 3.477 | 0.001 |

As seen in Table 6, there is a significant difference between the success of students in Faculties of Agriculture, Health Sciences and Veterinary Science ($p<0.05$ and $p<0.01$). Trend analysis summary of the exam and department is given in Table 7. There is a linear and quadratic trend. Linear and quadratic effects in the exam and exam-department interaction are considered important ($p<0.001$).

Table 7. Trend analysis test of the exam and exam-department interaction

| Source | Trend | Sum of Squares | df | Mean Square | F | p | Partial Eta Squared |
|---|---|---|---|---|---|---|---|
| Exam | Linear | 20283.528 | 1 | 20283.528 | 93.929 | 0.001 | 0.591 |
| | Quadratic | 44549.171 | 1 | 44549.171 | 302.599 | 0.001 | 0.823 |
| Exam * Department | Linear | 13385.825 | 2 | 6692.913 | 30.994 | 0.001 | 0.488 |
| | Quadratic | 6460.005 | 2 | 3230.003 | 21.940 | 0.001 | 0.403 |
| Error(exam) | Linear | 14036.432 | 65 | 215.945 | | | |
| | Quadratic | 9569.414 | 65 | 147.222 | | | |

In the study of Kocaman and Cumaoglu (2014), partial $\eta^2$ values were 0.717 and 0.724 and were very close to the result of this study. Ciftci et al. (2006), analyzed bean plant in terms of repeated randomized blocks experimental design by determining root lengths, shoot lengths, number of leaves, root and above the soil oven dry weights and under/over the soil dry weight rates after 15, 22, and 29 days following the sprout. In terms of characters examined by time, significant differences are observed among different types. Since the sphericity assumption was not met in the study of Anli and Sar (2017), variance analysis results were obtained using Greenhouse-Geisser and Huynh-Feldtcorrections. From this point of view, the study is similar to the present study in terms of method. Oner and Alpar (1990) applied three different pharmaceutical dosage forms consisting of the same active ingredient to 8 volunteers and performed repeated measurements variance analysis. As in this study, authors determined the difference in terms of three different dosage forms using the Tukey test.

## Conclusion

In this study, the difference between the grades averages related to examination levels was found to be significant according to the flatness test in variance analysis of repeated observations of the success of students in the exams. According to the parallelism test, the profiles presented in terms of grades as a result of Pillai's Trace, Wilks' Lambda, Hotelling's Trace, and Roy's Largest Root were not found parallel. In other words, it has been understood that the difference between the repeated measurements differs significantly with respect to the faculties. It is believed that using this method would be beneficial for further studies to be performed on dependent observations in many fields as well as education field.

## References

[1]. Anlı, G., Şar, A. H. 2017. Boyun Eğici Davranışları Azaltmaya Yönelik Bilişsel Davranışçı Psiko-Eğitim Programının Kişilerarası Duyarlılık, Öfke ve Düşmanlık (Hostilite) Üzerine Etkisi. Eğitim ve Bilim, 42(192): 383-405.

[2]. Çetin, M. O., Çakan, M. 2010. Fen ve Teknoloji Dersi Başarılarının Farklı Yaklaşımlarla Ölçülmesi ve Bu Yaklaşımlara İlişkin Öğrenci Görüşleri. Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi, 1(2): 93-99.

[3]. Çiftçi, V., Fırtına, D., Bildirici, N. 2006. The status of root and shoots of the dry bean (*Phaseolus vulgaris* L.) in the first development stage. KSU. Journal of Science and Engineering 9(1): 82-87.

[4]. Dani, S.G. 2012. Ancient Indian Mathematics-A conspectus. Resonance; 17(3): 236 - 46.

[5]. Field, A. 2009. Discovering statistics using SPSS. British Library Cataloguing in Publication data. Third edition, ISBN 978-1-84787-907-3

[6]. Jeremy, D. F. A. 1974. General Model for Multivariate Statistical Analysis. Prentice-Hall International, Inc, USA.
[7]. Kocaman, O.,Cumaoğlu, G.K. 2014. Eğitsel Yazılım (DENIS) ve oyunlarlaİngilizce kelime öğretiminin kelime öğrenme stratejilerine ve başarıya etkisi. Eğitim ve Bilim,39(176): 305-316.
[8]. Manly, B. F. J. 2001. Statistics for Environmental Science and Management: Chapman and Hall/CRC, USA.
[9]. Manly, B. F. J. 2005. Multivariate Statistical Methods: A primer. 3rd ed., Chapman and Hall/CRC, USA.
[10]. Öner, L., Alpar, R. 1990. Analysis of variance with repeated measures. FABAD J. Pharm. Sci., 15: 1113-1119.
[11]. Özerbaş, M. A., Benli, N. 2015. The Effect of Blended Learning on Students' Academic Achievement and Their Attitudes. GEFAD / GUJGEF 35(1): 87-108
[12]. Roy, S. N. 1957. Some aspect of multivariate analysis. John Wiley and Sons, New York.
[13]. Stevens, J. P. 1980. Power of the multivariate analysis of variance tests, Psychological Bulletin, 88:728-737
[14]. Tabachnick, B. G.,Fidell, L. S. 2013. Using Multivariate Statistics. Pearson Education, 6th Edition, ISBN-13: 978-0205849574.
[15]. Winer, B.J. 1971. Statistical Principles in Experimental Design. Second Edition, McGraw-Hill Book Company.