Science World Journal Vol 11 (No 2) 2016
www.scienceworldjournal.org
ISSN 1597-6343

Full Length Research Article

ON APPLICATION OF MODIFIED F – STATISTIC: AN EXAMPLE OF SALES DISTRIBUTION OF PHARMACEUTICAL DRUGS.

Abidoye A. O*[1]., Jolayemi E. T[2]., Sanni O. O. M[3] and Oyejola B. A[4]

Department Of Statistics, University of Ilorin, Ilorin, Nigeria.

* [1] abidoye@unilorin.edu.ng, [2]tejujola@unilorin.edu.ng ,[3] omsanni@unilorin.edu.ng, [4] boyejola@unilorin.edu.ng

ABSTRACT

This research work often found in Medicine, Agriculture and Social sciences. Most of the earlier works that have addressed the problem of testing equality of means assume homogeneity of variances across the groups being compared but in this work we want to focus in the area of medicine and make use heterogeneity of variances across the group means. Comparison of mean measurements coming from various sources (groups) often is hampered or made difficult by the associated variances not being identical. This is particularly common, for example, in pharmaceutical products where various generic products exist and products have variances that are group dependent.

In this work, the use of harmonic mean of sample variances is demonstrated instead of Analysis of Variance (ANOVA) methodology, in order to remove the associated Behren's – Fisher's problem. This test statistic is more applicable in a situation when we have three or more samples to consider at a time, then F – test statistic is needed for testing the hypothesis that all the samples are drawn from the populations with the same means. The interest of this research work is to adopt a suitable test procedure to address heterogeneity of variance.

The result shows that the modified F – test statistic is found to be appropriate for the data set obtained from the specific example considered in demonstrating the use of the test procedure.

Keywords: Ordered alternative, Modified F – test statistic, Variances heterogeneity, Harmonic mean of variances

INTRODUCTION

Analysis of variance is one of the most popular models in statistics in which interest may be mainly in testing the homogeneity of the different group means using the classical Analysis of variance (ANOVA). In many practical situations the standard assumption of homogeneity of error variances are often violated. In general, ANOVA assumes that the sample data sets have been drawn from populations that have normal distributions.

One – way or more complex analysis of variance usually evaluates the effect of a factor (s) on a single response variable. For example, a pharmaceutical industry may be interested in comparing a newly introduced product into the market for treating a particular ailment e.g Drugs use for controlling Blood pressure in patients with high blood pressure.

To satisfy the assumptions, the sales of drugs are randomly monitored in some wholesale pharmaceutical shops, the value of the sales for the shops for each sampled within a given period. The observations can be assumed to have the normal distribution having variances that are group dependent. see Ott, 1984 and Abidoye (2016).

The analysis of variance supports unequal sample sizes. This is important because designers of experiments seldom have complete control over the ultimate sample sizes in their studies. Each of the factors (products) must have at least two or more readings or sample observations. However, it is recommended that the error degrees of freedom should not be below ten (10).

The conventional analysis of variance (ANOVA) is usually based on the assumption of normality, independence of errors and equality of the error variances. Past studies have shown that the F – test is not robust under the violation of equal error variances, especially if the sample sizes are not equal and some authors have developed an exact Analysis of variance for testing the means of independent normal populations by using one or two stage procedures. See Jonckhere (1954), Abidoye et. al (2016), (2015a) and (2015b), Barlow et al (1971), Dunnett (1980), Dunnett and Tamhane (1997), Gupta et al (2006). Behren – Fisher's problem is one of the notable complications that arise in this regard. The objective of this study is to demonstrate the use of harmonic mean instead of mean square error (MSE) in the Analysis of variance (ANOVA) in order to side track Behren – Fisher's problem resulting from non – homogeneity of the group variances .

The hypothesis of homogeneity of means is to be tested. In some medicine researches where the interest is to investigate the effectiveness of certain brands of a drug meant for a particular ailment , there might be a pre- conceived belief that certain drug (s) are more effective than others. Thus, if , where, i = 1, 2, …, h represents the mean measurement of ith brand of drug, the test hypothesis here would be of the ordered type given above. See Barlow et al (1971), Dunnett (1980), Dunnett and Tamhane (1997), Gupta et al (2006) .

In some instances, however, it may be of interest to investigate the claim that some drugs perform better than others. Hence, in an attempt to compare these differing performances, we resort to the use of hypothesis against non - directional alternative. See Abidoye 2016, Yahya and Jolayemi (2003), Ott (1984), Montgomery (1981), Jonckheere (1954) and Bartholomew (1959).

DEVELOPMENT OF THE TEST PROCEDURE

The One – way analysis of variance (ANOVA) model equation is :

ISSN 1597-6343

$$Y_{ij} = \alpha + \tau_i + e_{ij}$$

where

Y_{ij} is the j[th] observation for the i[th] drug

α is a constant value

τ_i is the effect of i[th] drug.

$e_{ij} \sim N(0, \sigma_i^2)$ j = 1, 2, …,n_i .

We are interested in adopting a suitable test procedure to test the hypothesis:

$H_0 : \tau_1 = \tau_2 = ... = \tau_h = \tau = 0$ against the alternative

$H_1 : \tau_i \neq \tau$ for at least one i, $i =$ 1, 2, 3, …, h …………………………………………………………………….(2.1)

The computational procedure are:

$$SST = \sum Y_{ij}^2 - \frac{(\sum Y_{ij})^2}{n} \quad ....(2.2)$$

$$SSt = \frac{\sum Y_{i.}^2}{n_i} - \frac{(\sum Y_{ij})^2}{n} \quad ....(2.3)$$

SSE = SST – SSt ……………………………………………...…(2.4)

Abidoye (2012) and Abidoye et. al (2016) proposed the use of S_H^2 to replace the usual MSE in the Analysis of Variance table (ANOVA):

$$MSE = S_H^2 \quad ....(2.5)$$

where

$$S_H^2 = \left[\frac{1}{h} \left(\frac{1}{S_1^2} + \frac{1}{S_2^2} + ... + \frac{1}{S_h^2} \right) \right]^{-1}$$

and S_i^2 is the sample variance of the i[th] group. S_H^2 has been shown to have the χ^2 distribution with r degrees of freedom (not necessary an integer) . Abidoye (2016) and Abidoye et. al 2015a demonstrated that S_H^2 is likened to the common MSE as defined in one- way Analysis of variance (ANOVA). S_H^2 is a stabilized "pooled" variance, unaffected by some outliers in group variances. The one – way ANOVA is presented in Table 2.1 below. In Table 2.1 the sum of squares between groups and adjusted sum of squares do not necessarily sum to total sum of squares.

Analysis of Variance Table

Table 2.1: One – way analysis of variance with unequal group variances

Source variation	d.f	SS	MS	F
Treatment(between groups)	t-1	$SSt = \frac{\sum Y_i^2}{n_i} - \frac{(\sum Y_{ij})^2}{n}$	SSt/(t-1)=MSSt	$MSt / MSE^*(S_H^2)$
Error	$r \doteq n - t$	$r \times S_H^2$	$MSE^*(S_H^2)$	
Total	n – 1	$SST = \sum Y_{ij}^2 - \frac{(\sum Y_{ij})^2}{n}$		

Simulation study show that without any loss of generality, nearest integer values can be used to approximate the r degrees of freedom.

Application of the Test Procedure

The data used in this study considered three drugs in use for treatment of Blood pressure (hypertension). The data is a secondary data, collected from pharmaceutical premises in Ilorin, Kwara State, covering the period of six months (April to September, 2014), see Table 3.1 below:

Table 3.1: One – way analysis of variance with unequal group variances: Wholesales of three drugs in five pharmaceutical companies in Ilorin, Kwara State, covering the same period.

Pharmaceutical House					
Drug Type	A	B	C	D	E
Lisinopil (old drug)	32	35	40	29	23
Losartan (old drug)	18	22	26	23	30
Cardesarten (new drug)	25	55	38	20	23

We need to verify the equality of the variances between these three drugs. This is shown in Table 3.2 using Levene's test. Thus the group variances cannot be assumed equal (P = 0.039). Thus the regular ANOVA procedure cannot be implemented in testing the hypothesis

Table 3.2: One – way analysis of variance with unequal group variances: Levene test for equality of variances

	Levene STATISTIC	df_1	df_2	P-value
Response	4.288	2	12	0.039

COMPUTATION ON SALES OF DRUGS' DATA ON BLOOD PRESSURE

From the data set on sales of drugs for treatment of hypertension, the following summary statistics were obtained:

Lisinopil: $\bar{Y}_{Lis}$ = 31.8, S^2_{Lis} = 40.7, n_{Lis} = 5

Losartan: $\bar{Y}_{Los}$ = 23.8, S^2_{Los} = 20.2, n_{Los} = 5

Cardesarten: $\bar{Y}_{Car}$ = 32.2, S^2_{Car} = 209.7, n_{Car} =5

In the above, data set, n_i = 5, h = 3, $n = \sum_{i=1}^{3} n_i = 15$;

$$S_H^2 = \left(\frac{1}{3}\sum_{i=1}^{3}\frac{1}{S_i^2}\right)^{-1}; \quad S_H^2 = 38.04,$$

$MSE = 90.20$ ....(see equation 2.5)

The main hypothesis is

$$H_0 : \mu_A = \mu_B = \mu_C = \mu \quad against$$
$$H_1 : \mu_i \neq \mu$$
$$for \;\; atleast \;\; one \;\; i, \, i.e \;\; i = A, B, C$$

....(see equation 2.1)

$$SST = \sum Y_{ij}^2 - \frac{(\sum_{i=1}^{n_i}\sum_{j=1}^{n_i} Y_{ij})^2}{n}$$

= $32^2 + 35^2 + 40^2 + 29^2 + 23^2 + \ldots + 20^2 + 23^2$ - 12,848.07

= 1,306.93

$$SSt = \frac{\sum Y_{i\cdot}^2}{n_i} - \frac{(\sum_{i=1}^{n_i}\sum_{j=1}^{n_i} Y_{ij})^2}{n}$$ ....(see equation 2.3)

= 13,072.6 – 12,848.07

= 224.53

SSE = SST – SSt.................................. (see equation 2.4)

= 1,306.93 - 224.53

= 1, 082.4

Table 3.3: Analysis of Variance Table (ANOVA)

Source of Variation	d.f	SS	MS	F	P – value
Treatment (between groups)	2	224.53	112.27	2.951	0.0907
Adjusted error	12	456.48	38.04		
Total	14	1,306.93			

where α = 0.05

Note that F- statistic using equal group variances produced the p – value of 0.323 which is higher than the one in the modified F – statistic obtained from harmonic mean of sample variances, showing that regular ANOVA test is conservative even though the same conclusion is obtained here.

Analysis of variance shows that the sales of the three drugs used for high blood pressure collected from some pharmaceutical premises in Ilorin, Kwara State are not significantly different.

Conclusion

In this work, we examined a test statistic for testing equality of means under unequal population variances in a comparative drug experiment. We adopted the proposed modified F – test statistic, where the harmonic mean of variances replaced the sample pooled variance in order to avoid the Beheren – Fisher's problem. Because the sample harmonic mean of variances has the chi – square distribution, the modified F –test is found to be appropriate for the data set used on the sale of hypertensive drugs. We can therefore conclude that the new drug (cardesartan) compete favourably well with the existing drugs for the treatment of Blood pressure (BP) as recommended by the caregivers.

REFERENCES

Abidoye, A. O. (2012). Development of Hypothesis Testing Technique for Ordered Alternatives under heterogeneous variances. Unpublished Ph.D Thesis submitted to Dept. of Statistics, University of Ilorin, Ilorin.

Abidoye, A.O, Jolayemi, E.T, Sanni, O.O.M and Oyejola, B. A. (2016). Development of Testing Ordered Mean Against a Control Under Heterogeneous variance. Journal of theNigerian Association of Mathematical Physics. Vol.33, pp 125 – 128.

Abidoye, A.O, Jolayemi, E.T, Sanni, O.O.M and Oyejola, B. A. (2015a) . Development of a Test Statistic for Testing Equality of Means Under Unequal Population Variances. Accepted for publication in Ilorin Journal of

Science. Published by Faculty Of Physical Sciences, University of Ilorin.

Abidoye, A.O, Jolayemi, E.T, Sanni, O.O.M and Oyejola, B. A. (2015b) . Development of Hypothesis Testing on Type one Error and Power Function. Accepted for publication in Ilorin Journal of Science. Published by Faculty Of Physical Sciences, University of Ilorin.

Barlow, R.E, Bartholomew, D.J, Bremmer,J.M and Brunk, H.D. (1971). "Statistical Inference under Order Restrictions: The Theory and Application of Isotonic Regression." Wiley, New York.

Bartholomew, D.J. (1959). " A test of homogeneity for ordered alternative." Biometrica, 46, 3648

Dunnett, C. W and Tamhane, A. C. (1997). Multiple testing to establish superiority / equivalence of a new treatment compare with k standard treatments. Statistics in Medicine 16, 2489 – 2506.

Dunnett, C.W. (1980). " Pairwise multiple comparison in the homogeneous variance, unequal sample size case" Journal of the American Statistical Association. 75, 789 – 795.

Gupta, A.K, Solomon, W.H and Yasunori, F. (2006). Asymptotics for testing hypothesis In some Multivariate variance components model Under non-normality. Journal of Multivariate Analysis archive. Vol. 97, pp 148-178.

Jonckheere, A.R. (1954). "A distribution-free k-sample test against ordered alternative." Biometrical, 41, 133-145.

Montgomery, D.C. (1981). "Design and Analysis of Experiment." Second Edition. John Wiley and sons Inc. New York.

Ott, L. (1984). "An Introduction To Statistical Methods and Data Analysis". Second edition. P.W.S Publisher. Boston.

Yahya, W.B and Jolayemi, E.T. (2003). Testing Ordered Means against a Control. JNSA,16, 40-51.