Journal of the Nigerian Association of Mathematical Physics
***Volume 33,* (*January, 2016*), pp 125 – 128**

Development of Testing Ordered Mean Against a Control Under Heterogeneous Variance

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

Department of Statistics, Faculty of Physical Science, University of Ilorin, Ilorin, Nigeria.

Abstract

Comparison of more than two population means based on independent random samples is common in research 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. However, when the variances are unequal, the pooled sample variance overestimates the appropriate variance and the test statistic become conservative. This is the well known Behrens – Fisher problem. There is therefore the need to seek for an alternative sample mean variance better than the pooled variance. The harmonic mean of variances was found to be better and therefore it being proposed as an alternative to the pooled sample variance when there is heterogeneity of variances. The interest of this research work is to develop a suitable test procedure based on the harmonic mean of variance to address heterogeneity of variance. We are proposing an hypothesis testing technique for testing ordered alternatives under heterogeneous variance for testing g ordered mean when one of this mean is being considered a control (standard) with the objective of a proposal of a test statistic for testing equality of means against directional alternative in the presence of heterogeneity of variance. In this study, the use of harmonic mean of sample variance is demonstrated with a practical application. The result shows that the proposed t – test statistic is found to be appropriate for the data set obtained from the specific example considered in demonstrating this test procedure.

Keywords: Ordered alternative, t – test statistic, variances heterogeneity, harmonic mean of variances, a control

1.0 Introduction

Application of directional alternative hypothesis testing may also be found in industry. Industry is one of the factors that contribute to the economy growth of the nation especially in Nigeria. One of the factors we want to look into in Nigeria is Coca – Cola bottling company. Coca – Cola drink is one of the oldest drink before other product is being produce by company, such product include Can Coca – Cola, Can - Fanta, Limca, Sprite, Schweppes , etc. In this work, we want to look at Can Coca – Cola products by comparing a set of mean sales, where they are expected to appear in a particular order, but specifically the situation when one of them (the mean sale) serves as a control (standard) for other. In particular, if μ_i represents the mean of population i,
i = 1,…, g, the problem is to test the hypothesis

$$H_0 : \mu_1 = \mu_2 = ... = \mu_{g-1} = \mu_g \text{ vs } H_1 : \mu_1 \le \mu_2 \le ... \le \mu_{g-1} < \mu_g$$

Where $\mu_g = \max_{1 \le i \le g} \mu_i$ is the control mean from the g^th^ population. By the term control, we mean that all other g-1 means are compared against μ_g In other words, μ_g serves as the basis of comparison against the other $\mu_i's, \quad 1 \le i \le g-1$.

Comparing more than two population means, based on independent random samples is a common occurrence in the fields of research. Several authors have provided adequate literature in this area via the use of Analysis of variance (ANOVA). [1 – 4].

However, in many experimental situations, one of the g treatments is a control and the analyst may be interested in comparing each of the other g – 1 treatment means with the control. Thus, there are only g – 1 comparisons to be made [4]. A procedure for making these comparisons was developed by Dunnett [5], which is a modification of the usual t – test.

Corresponding author: Abidoye A.O., E-mail: solatayo2003@yahoo.com, Tel.: +2348050663983

The test problem similar to the above could also be found in other disciplines such as Engineering and Technology, Economics, Politics and Humanities. A test procedure to investigate such cases is proposed in this paper. The test procedure to handle the control, which is the minimum, can be handled by the negation of the strict inequality. That is, μ_i would be the minimum in the hypothesis set above. Many authors have worked in this area[1, 6 – 14].

2.0 Methodology

In this section we are interested in developing a suitable test procedure to test the hypothesis testing g ordered mean when one this mean is being considered as a control (standard).

Let $X_{ij} \sim N(\mu_i, \sigma_i^2)$ where i = 1, 2, …, g and j =1, 2, …, n~i~ .... (2.1)

The hypothesis testing to be consider for this test procedure is

$$H_0 : \mu_1 = \mu_2 = ... = \mu_{g-1} = \mu_g \text{ vs } H_1 : \mu_1 \le \mu_2 \le ... \le \mu_{g-1} < \mu_g \qquad (2.2)$$

Where $\mu_g = \max_{1 \le i \le g} \mu_i$ being the control. Also consider a combination of any of the $\mu_i^{'} s$, i = 1, 2, …, g-1 and μ_g and define

$$\delta_i = \mu_i - \mu_g \qquad (2.3)$$

Then $\mu_i < \mu_g$ if and only if $\delta_i < 0 \ \ \forall_i, i = 1,2,..., g-1$

Therefore, the hypothesis of (2.2) now becomes

$$H_0 : \delta_i = 0 \qquad \forall_i, i = 1,2,..., g-1 \text{ against } H_1 : \max_{1 \le i \le g-1} \delta_i < 0 \quad ……………………………….$$

(2.4)

The unbiased estimate of δ_i is

$$\hat{\delta}_i = \overline{X}_i - \overline{X}_g$$

Therefore, $\hat{\delta}_i \sim N[\delta_i, V(\hat{\delta}_i)]$ ………………………………………………………………..

(2.5)

where, $V(\hat{\delta}_i) = V(\overline{X}_i - \overline{X}_g)$

$$= V(\overline{X}_i) + V(\overline{X}_g)$$

$$= \frac{\sigma_i^2}{n_i} + \frac{\sigma_g^2}{n_g}$$

$$Var(\hat{\delta}_i) = \left(\frac{n_i + n_g}{n_i n_g}\right)\sigma_H^2$$ ……………………………………………………………..

(2.6)

where $\sigma_i^2 \neq \sigma_g^2$ for at least one $i \neq g$

The test statistic for the hypotheses set in equation (2.2) is therefore

$$t^* = \frac{\lambda X}{Z}$$ ……………………………………………………………………………….....

(2.7)

Where,

$$X = \min(\bar{x}_i - \bar{x}_g)$$ ………………………………………………………………………….

(2.8)

and

$$Z = \sqrt{S_H^2\left(\frac{1}{n_i} + \frac{1}{n_g}\right)}$$ ………………………………………………………………….….

(2.9)

where λ is the appropriate normalization from order statistic . The null hypothesis is then rejected if

$$P\left(t_r^* = \lambda\frac{X}{Z} > t^*\right) = P\left(t_r^* = \lambda t_r > t^*\right)$$

$$= P\left(t_r > \frac{t^*}{\lambda}\right) < \alpha$$ …………………………………………………...…

(2.10)

where t_r is the regular t- distribution with r degrees of freedom is given by

$$t_r = \frac{\max(\overline{X}_i - \overline{X}_g)}{S_H\left(\frac{1}{n_i}+\frac{1}{n_g}\right)^{\frac{1}{2}}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.11)$$

because S_H^2 has approximately the χ^2 - distribution with the degree of freedom r to be determined from r = 22.096 + 0.266(n-g) - 0.000029 $(n-g)^2$ and explain it very well [7]. If the alternative hypothesis H_1 in equation (2.11) is reversed.

Then

$$t_r = \frac{\max(\overline{X}_g - \overline{X}_i)}{S_H\sqrt{\left(\frac{1}{n_i}+\frac{1}{n_g}\right)}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.. (2.12)$$

3.0 Data Analysis

Practical situations where this test procedure is applicable are presented in this section. The data of Coca – cola and other brands was found in Coca – cola bottling company Plc, Ilorin, Kwara State.

3.1 Data Presentation and Analysis

The data used in this work considered sales of five types of Can Coca- Cola. The data is a secondary data, collected from Coca – Cola bottling company Plc, in Kwara State, covering the period of four quarters of the year 2014 (January to December).

Table 3.1: Quarterly Sales of five different can Coca – Cola, in bottling company Plc in Kwara State, in the year 2014.

Type of can Coca-cola	First quarter(Q_1)	Second quarter(Q_2)	Third quarter(Q_3)	Four quarter(Q_4)
Coca- cola (control)(a)	7,814	10,904	11,616	12,032
Coca- cola zero(b)	01	293	1,441	236
Fanta orange (c)	4,155	4,846	4,202	5,719
Schweppes ok (d)	2,172	1,929	1,723	2,367
Sprite (e)	1,690	1,413	1,139	1,689

We first verified the equality of the variances between these quarterly sales. We adopted Leven's test. The result is shown in table 3.2. Our result indicated P – value of 0.034. Thus the Variances are not equal. Hence, the regular ANOVA procedure cannot be used for the analysis.

Table 3.2: Levene test for equality of variances

	Levene Statistic	df_1	df_2	P – value
Response	3.473	4	15	0.034

Computations

From the data set on sales of cans coca – cola for four quarters in one year (2014), the following summary statistics were obtained:

Coca cola (control): $\overline{Y}_{(a)} = 10{,}591.5, \quad S^2_{(a)} = 1{,}909.35, \quad n_{(a)} = 4$

Coca – cola zero: $\overline{Y}_{(b)} = 492.75, \quad S^2_{(b)} = 644.67, \quad n_{(b)} = 4$

Fanta Orange: $\overline{Y}_{(c)} = 4730.5, \quad S^2_{(c)} = 730.52, \quad n_{(c)} = 4$

Schweppes ok: $\overline{Y}_{(d)} = 2047.75, \quad S^2_{(d)} = 281.02, \quad n_{(d)} = 4$

Sprite: $\overline{Y}_{(e)} = 1482.75, \quad S^2_{(e)} = 263.64, \quad n_{(e)} = 4$

Then consider the following differences;

$$\overline{Y}_{(b)} - \overline{Y}_{(a)} = -10098.75$$
$$\overline{Y}_{(c)} - \overline{Y}_{(a)} = -5861$$
$$\overline{Y}_{(d)} - \overline{Y}_{(a)} = -8543.75$$
$$\overline{Y}_{(e)} - \overline{Y}_{(a)} = -9108.75$$

max $(\overline{X}_{(i)} - \overline{X}_{(g)}) = (\overline{X}_{(c)} - \overline{X}_{(a)}) = -5861$

with $n_i = 4$, $g = 5$, $n = \sum_{i=1}^{5} n_i = 20,$; $S_H^2 = \left(\frac{1}{5} \sum_{i=1}^{5} \frac{1}{S_i^2} \right)^{-1} = 463.16$; $S_H = 21.521$

$$t^* = \frac{-5861}{21.521\sqrt{\frac{1}{4} + \frac{1}{4}}} = 385.144$$

Now p- value = $P(t^* = t_r > t_{cal}) = P(t_r > t_{cal})$

$= P(t_r > 385.144) = 0.0001 < 0.05$

$r = 22.096 + 0.266(n - g) - 0.000029(n-g)^2$

$= 26.453$

This led to rejection of H_0 and concludes that the effect of the control on the sales of can coca- cola on the sales other products of coca – cola is higher as expected.

4.0 Conclusion

It has been established in this work that testing against ordered treatment means where one of the treatment is a control has practical application in many area of life. We can observed from the coca- cola bottle company that the convention product of can coca – cola still have highest sales compare with other product as we expected.

5.0 Reference

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

[2]. Ott, L (1984): " An Introduction To Statistical Methods and Data Analysis". 2nd Edition. P. W. S. Publishers. Boston.

[3]. Neter, J and Wassarman, W (1974): Applied Linear Statistical Model". Richard D. Irwin, Inc. IIIinois.

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

[5]. Dunnett , C. W (1964): New tables for multiple comparison with a control. Biometric, **20**, 482-491, (1964).

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

[7]. 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.
[8]. 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.
[9]. Dunnett, C.W (1980): " Pairwise multiple comparison in the homogeneous variance, unequal sample size case" Journal of the American Statistical Association. 75, 789 – 795.

[10]. 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.
[11]. 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**, 148-178.
[12]. Adegboye, O.S and Gupta, A.K "On Testing Against Restricted Alternative About The Means of Gaussian Models With common Unknown Variance." Sankhya, The Indian Journal of Statistics, series B, **48**, 331-341, (1986).
[13]. Bartholomew, D. J. A test of homogeneity for ordered alternatives". Biometrica, **46**, 36 – 48, (1959).
[14]. Chacko V. J. "Testing homogeneity against ordered alternatives". Annals of Mathematical Statistics, **34**, 945 – 956, (1963).