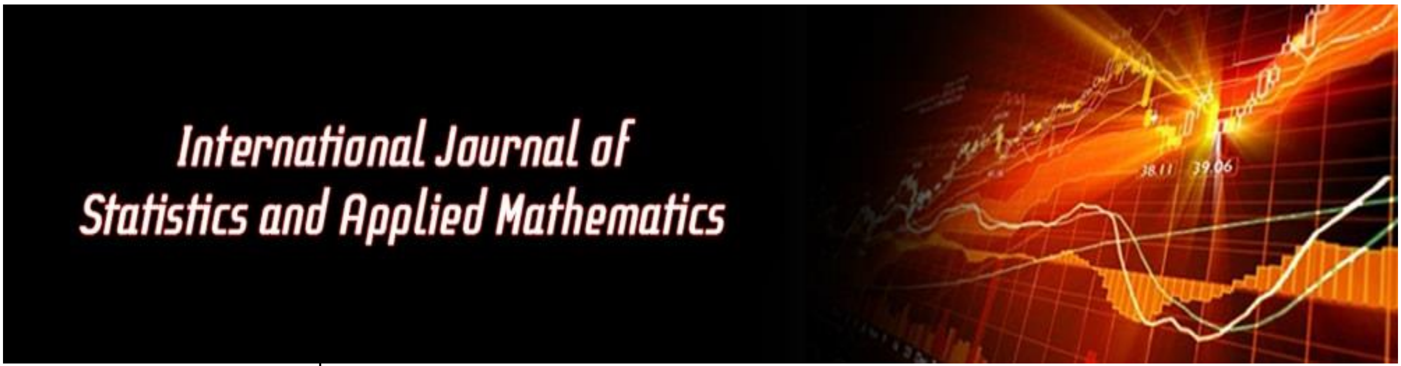
International Journal of Statistics and Applied Mathematics 2019; 4(1): 62-65



ISSN: 2456-1452
Maths 2019; 4(1): 62-65

www.mathsjournal.com
Received: 13-11-2018
Accepted: 21-12-2018

Samhita Pal
Department of Statistics,
Ballygunge Science College,
University of Calcutta
Kolkata, West Bengal, India

Correspondence
Samhita Pal
Department of Statistics,
Ballygunge Science College,
University of Calcutta
Kolkata, West Bengal, India

ANOVA in GLS transform vs. Welch's ANOVA under heteroscedasticity

Samhita Pal

Abstract
Heteroscedastic data arise from a population whose sub-populations have different variabilities from one another. In presence of heteroscedasticity, the LS estimators no longer remain the BLUEs, giving rise to wider acceptance regions and confidence intervals for the model parameters. The most conventional way of removing heteroscedasticity is the generalized least square technique. One can perform one-way ANOVA on this transformed data or use Welch's ANOVA. This paper compares the efficiencies, i.e., the power of these two methods.

Keywords: Heteroscedasticity, generalized least square, Welch's ANOVA, power analysis, effect size

Introduction
Let us consider the dataset named "Engel data" from R, consisting of 235 observations on income and expenditure on food for Belgian working class households (Belgian Francs). The scatterplot, obtained from R, is given below.

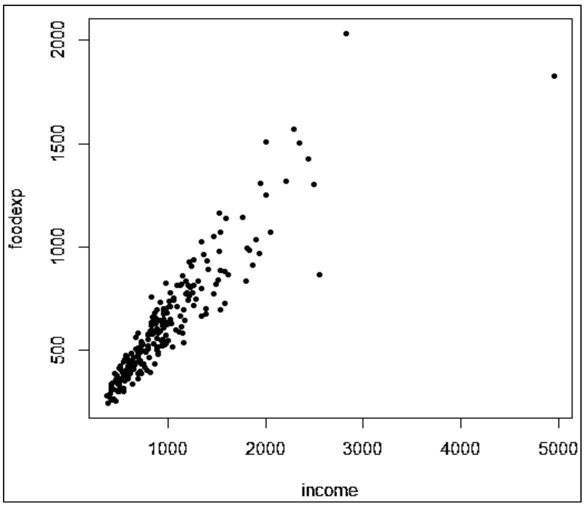


Fig 1: Scatterplot of incomes of Belgian Workers vs their expenditures.

The scatterplot in Figure 1 shows that as income increases, the expenditures of the Belgian workers become more dispersed, suggesting the presence of heteroscedasticity.
We now create a new column in "Engel" that converts the variable "income" into a categorical variable called "inc_gr" with three levels called "1", "2", and "3", denoting the low income group, the middle income group and the high income group. The lowest third of the data belongs to the first level, the middle third belongs to the second level, and the highest third of the data to the third level.
This is done by dividing the variable "foodexp" into three parts as per their quantiles corresponding to 0.333 and 0.666.

[R] Code

```
Quantile (engel $income, seq (0, 1, 1/3))
inc_gr = cut (engel $income, c (377, 718.3701, 1014.1540, 4958), labels = c (1, 2, 3))
Engel = cbind (engel, inc_gr)]
```

The assumptions of the one-way analysis of variance are

1. The data are continuous (not discrete).
2. The data follow the normal probability distribution. Each group is normally distributed about the group mean.
3. The variances of the populations are equal.
4. The groups are independent. There is no relationship among the individuals in one group as compared to another.
5. Each group is a simple random sample from its population. Each individual in the population has an equal probability of being selected in the sample.

On checking the variance or standard deviation of food expenditure corresponding to the three income groups, we fing them to differ widely; SD of food expenditure for lower income group is 75.83375, SD of food expenditure of middle income group is 89.53907 and that for the high income group is 282.972 Hence, the assumption of ANOVA is violated. The two remedies in such situations are stated next.

[R] Code

```
Plot (engel $inc_gr, engel $ foodexp, type = "p")]
```

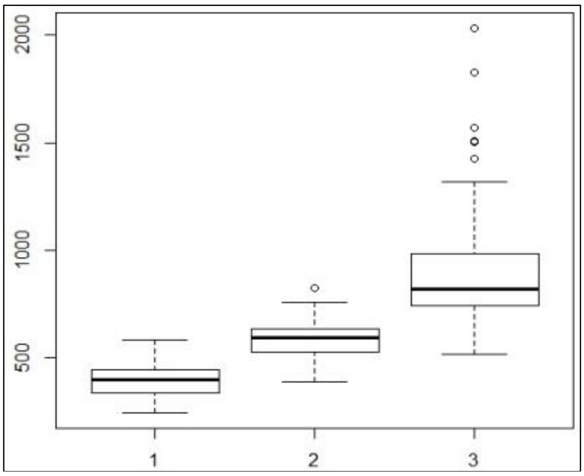


Fig 2: Boxplot of the standard deviations of the three income groups

2. Materials and Methods

2.1 Generalised Least Square Method

If the true nature of σ_i^2 is known, then, the GLS technique [1, 2] converts the heteroscedastic error to homoscedastic error, i.e., $\sigma_i^2 = \sigma^2 \ \forall$ i, by dividing the model by the known standard deviation σ_i.

$$\Rightarrow \frac{y_i}{\sigma_i} = \beta_1 \frac{1}{\sigma_i} + \beta_2 \frac{x_i}{\sigma_i} + \frac{u_i}{\sigma_i} \Rightarrow y_i* = \beta_1 x_{oi} * + \beta_2 x_i * + u_i *,$$

Where

$V(u_i*) = \sigma_i^2 / \sigma_i^2 = 1 \ \forall$ i.

The estimates of β_1 and β_2 obtained by the method of least squares with the transformed data on this new model are the *generalized least square (GLS) estimates*.

In our case, and in most practical cases, the true nature of σ_i^2 remains unknown. We seek help of the Glejser's test [3, 4] to identify the nature of σ_i^2 and then GLS method is applied as before.

Here, we consider the model $\sigma_i = \delta_o + \delta_1 x_i^{\delta_2} + v_i$, where v_i is the error term of this new model. σ_i being unknown, it is replaced by its estimate $|e_i|$, where e_i is the residual obtained by OLS. The transformed model is $|e_i| = \delta_o * + \delta_1 * x_i^{\delta_2 *} + v_i *$. The null hypothesis for Glejser's test is H_o: $\delta_1* = 0$. T-tests are carried out for some pre-fixed values of δ_2* (-1, -½, ½, 1). Rejection of H_o for any of these given values of δ_2*, suggests that $\sigma_i \propto x_i^{\delta_2 *}$.

For the engel data, we get

Table 1

δ_2*	-1	-½	½	1
|t-value|	6.624	7.95	11.704	13.981
p-value	0.00	0.00	0.00	0.00

As we can see from Table 1, the modulus of t-value for $\delta_2* = 1$ is maximum and therefore is considered to be the most significant. We thus conclude that $\sigma_i \propto x_i$.

The transformed data is obtained as $y_n = y/x$ and $x_n = 1/x$.

The scatterplot of the transformed data is-

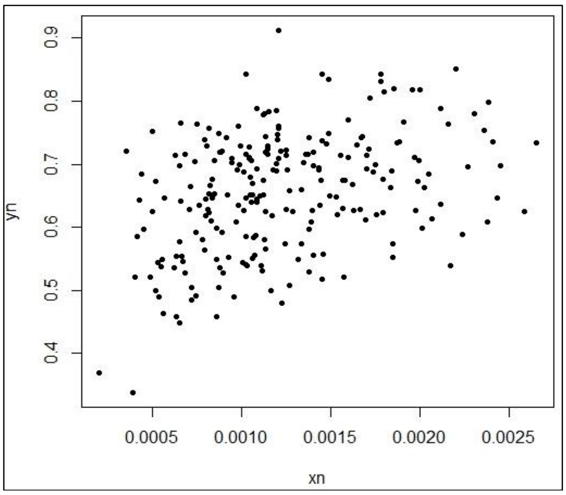


Fig 3: Scatterplot of the transformed data using GLS.

Clearly, the scatterplot shows no more sign of any heteroscedasticity. The transformed data thus satisfies all conditions required for conducting a one way ANOVA.

The transformed response variable y_n now has almost equal variances (close to 0.085) in the three groups. This is further represented by a boxplot below.

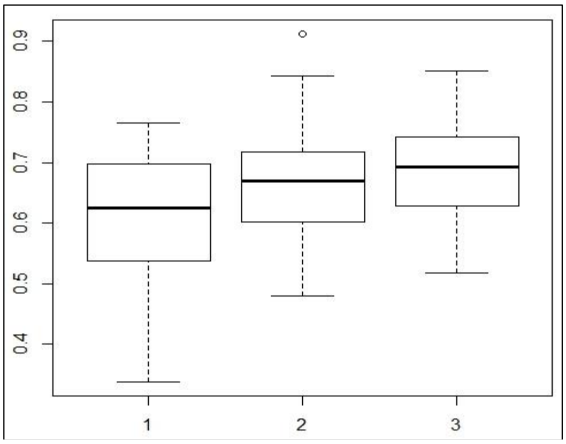


Fig 4: Boxplot of the standard deviations of the three income groups based on the transformed data.

The transformed data is now suitable for the one way ANOVA test. We obtain the ANOVA table through the lm (.) function in R.

Table 2

Source	df	SS	MS	F	p-value
Income group	2	0.26646	0.133232	17.301	9.916e-08
Error	232	1.78656	0.007701		
Total	234	2.05302	-		

We reject the null hypothesis of equal means (average food expenditure for all the three income groups are same) with 95% confidence.

Now we check the power of this ANOVA test [5, 6].

The power of any test of statistical significance is defined as the probability that it will reject a false null hypothesis. Statistical power is inversely related to beta or the probability of making a Type II error. In short, power = 1-β.

In other words, statistical power is the likelihood that a study will detect an effect when there is an effect there to be detected. If statistical power is high, the probability of making a Type II error, or concluding there is no effect when, in fact, there is one, goes down.

Statistical power is affected chiefly by the size of the effect and the size of the sample used to detect it. Bigger effects are easier to detect than smaller effects, while large samples offer greater test sensitivity than small samples.

The ANOVA model is $(y_n)_{it} = \mu_i + \varepsilon_{it}$ *;* i = 1,..,p *;* t = 1(1)n [Here p=3, n=78] where ε_{it} are i.i.d $N(0,\sigma^2)$. To test the hypothesis that the treatment means are all equal, we use F = MST/MSE (the ratio of the mean square between groups to the mean square within groups) and reject the null hypothesis if $F_{obs} > F_{2,\,232,\,\alpha}$ for a given probability of type I error α. When the null hypothesis of mean equality is rejected, the above ratio has a non-central *F* distribution which adds a non-centrality parameter [7], λ. This parameter is calculated as

$$\lambda = \frac{\sigma_m^2}{\sigma^2}\sum_{i=1}^{p} n_i$$

Where $\sigma_m^2 = \frac{\sum_1^p n_i(\mu_i-\bar{\mu}_w)^2}{\sum_1^p n_i}$ and $\bar{\mu}_w = \frac{\sum_1^p n_i\mu_i}{\sum n_i} = \frac{1}{p}\sum_1^p \mu_i$ (weighted mean of all groups).

Cohen (1988) [8] proposed an effect size parameter f which he defined as $f^2 = \frac{\sigma_m^2}{\sigma^2}$. Using this, the non-centrality parameter reduces to $\lambda = f^2\sum_{i=1}^{p} n_i$. Now, another effect size is given by $\eta^2 = \frac{f^2}{1+f^2}$, which is the proportion of the total variation in the dependent variable that is attributable to the groups. The power is computed as the probability of being greater than $F_{p-1,\,p(n-1)}$, α on a noncentral-*F* distribution with non-centrality parameter λ, where λ is calculated from η^2 using the formula given above.

In our example, we get the effect size as $\eta^2 = \frac{SS(treatment)}{SS(total)} = \frac{0.26646}{2.05302}$ = 0.129837994 => f = 0.386279.

Using R software, we compute power with the command pwr. ANOVA. Test (k=3, n=78, f=0.386279, sig. level=0.05) and obtain the power as 0.998.

2.2 Welch's ANOVA

For testing the equality of k sample means, when the variances are not equal, Welch (1951) [9-11] provided the following modification of the usual ANOVA F-test:

$$F_w = \frac{\sum_i w_i(\bar{y}_{io}-\hat{y})^2/(p-1)}{1+\frac{2(p-2)}{p^2-1}\sum_i \frac{1}{n_i-1}(1-\frac{w_i}{\sum_j w_j})^2}$$

Where $w_i = n_i / s_i^2$ and $\hat{y} = \frac{\sum_i w_i\bar{y}_{io}}{\sum_i w_i}$, s_i^2 being the variation in response values corresponding to the i^th^ group.

This Welch modification rejects the null hypothesis of equal means if the F_w statistic is larger than the critical value determined from an F distribution with degrees of freedom f_1^* and f_2^*,

Where-

$f_1^* = p-1$

$f_2^* = 1/[\frac{3}{p^2-1}\sum_i \frac{1}{n_i-1}(1-\frac{w_i}{\sum_j w_j})^2]$.

From R, we get $F_{w(obs)}$ = 194.41, numerator df = 2, denominator df = 138.91, p-value < 2.2e-16 using the oneway. Test function with var. equal=FALSE.

Here, we have $\frac{SSA}{SSE}$ = 194.41. Therefore, SSA = 194.41*SSE ⇒ TSS = 194.41*SSE + SSE = 195.41*SSE

Then, $\eta^2 = \frac{SS(treatment)}{SS(total)} = \frac{194.41*SSE}{195.41*SSE} = 0.9948826$ ⇒ f = 13.94316

Using R software, we compute power with the command pwr. ANOVA. Test (k=3, n=78, f= 13.94316, sig. level=0.05) and obtain the power as 1.

3. Results and Discussions

We observe that both one way ANOVA on transformed data and Welch's ANOVA test for the engel dataset gives the same decision, i.e., rejecting the null hypothesis of equal means of expenditure for food. The p-value for Welch's ANOVA test is much less than that of the classical one way ANOVA test, however power remains more or less same. As such, no discrepancy is noticed between the two results, which establishes the immense utility of both the methods.

One drawback of the classical one way ANOVA on the GLS transformed data is that the original data is lost and all results are based on a completely new data set obtained by dividing the original data by the explanatory variable. This creates difficulty in proper interpretations. The Welch's ANOVA surpasses the one way ANOVA on this ground.

4. References

1. Damodar NG. Heteroscedasticity. In: Basic Econometrics. Fourth edition, The McGraw-Hill Companies, US, 2004, 387-426.
2. Qingfeng L, Ryo O, Arihiro Y. Generalized Least Squares Model Averaging. Econometrics Reviews. 2016; 35(8-10):1692-1752.
3. Edgar PH, Clayton SL. A note on glejser's test for heteroscedasticity. Communications in Statistics. 1980; 9(12):1209-1220.
4. Jose AFM, Santos Silva JMC. Glejser's test revisited. Journal of Econometrics. 2000; 97(1):189-202.
5. Joel RL. Overcoming Feelings of Powerlessness in "Aging" Researchers: A Primer on Statistical Power in Analysis of Variance Designs, Psychology and Aging. 1997; 12(1):84-106.
6. Wilcox RR. ANOVA: A paradigm for low power and misleading measures of effect size, Review of Educational Research. 1995; 65(1):51-77.

7. Xiaofeng L, Stephen R. A Note on the Noncentrality Parameter and Effect Size Estimates for the F Test in ANOVA, Journal of Educational and Behavourial Statistics. 2004; 29(2):251-255.
8. Joseph LG, Yulia RG, Weiwen M. The Impact of Levene's Test of Equality of Variances on Statistical Theory and Practice, Statistical Science. 2009; 24(3):343-360.
9. Mehmet M, Soner Y. Comparison of ANOVA-*F* and ANOM tests with regard to type I error rate and test power, Journal of Statistical Computation and Simulation. 2013; 83(11):2093-2104.
10. Kenneth JL. An empirical comparison of the ANOVA *F*-test with alternatives which are more robust against heterogeneity of variance, Journal of Statistical Computation and Simulation. 1978; 8(1):49-57.
11. Mehmet M, Erkut A. Comparison of ANOVA F and Welch Tests with Their Respective Permutation Versions in Terms of Type I Error Rates and Test Power, Kafkas Universitesi Veteriner Fakultesi Dergisi. 2010; 16(5):711-716.