Assessing Univariate and Multivariate Homogeneity of Variance: A Guide For Practitioners

Kassu Mehari Beyene[1*] Sara Abera Bekele[2]
1. Hasselt University, Hasselt, Belgium.
2. Wollo University, Dessie, Ethiopia.

Abstract

Most statistical methods assume constant variance and the validity of result from some of the methods highly rely on constant variance. However, a very high number of practitioners and researchers publications do not check the constant variance assumptions and hence the results are very prone to error. With aim of reducing this, both graphical and formal methods of assessing constant variance assumption are presented and illustrated in this paper. In univariate data several methods have been proposed. The graphical methods of assessing constancy of variance include plot of residuals against the fitted values, residuals against the fitted value square, and residual versus predictor variable are widely used. In addition, formal tests of assessing this assumption are Bartlett's, Levene's, Breusch-Pagan, Brown and Forsythe, O'Brien's, White's and Fligner-Killeen are commonly used and also applicable in most of statistical software. For multivariate data, the two common tests in practice are Box's M teste and Bartlett's. Finally, when the constancy of variance assumption not satisfied, it is very important to find a variance-stabilizing transformation.

Keywords: Homogeneity of variance, Bartlett's test, Breusch-Pagan test, Brown and Forsythe test, Levene's test.

1. Introduction

In statistics, the homogeneity of variance is one of the main assumptions of liner models as it is presented in many literatures. The assumption of constancy of variance is required in many parametric univariate and multivariate statistical methods, such as two sample independent t-test, linear regression, Analysis of Variance (ANOVA), Multivariate Analysis of Variance (MANOVA), linear discriminant analysis, canonical correlation, etc.. (Neter *et al*., 2005; Johnson and Wichern, 2007). Therefore, it is important to assess multivariate normality in order to proceed with such statistical methods. Although, many researchers indicated the necessity of assessing the homogeneity of variance assumption to get efficient estimate of the parameters, many practitioners are reluctant to assess the homogeneity of variance assumption. The possible reasons for this could be: lack of awareness for the existence of tests, lack of software packages/procedures for calculating the test statistics and/or p value for a given test, and the practitioner may be not sure on the remedial measures, when there is departure from the homogeneity of variance assumption.

The non-homogeneity of error variance does not cause biased estimates, but it does cause problems for efficiency and the usual formulas for standard errors are inaccurate. The reason for loss of inefficiency of the ordinary least square estimates is equal weight is given to all observations regardless of the fact that those with large residuals contain less information about the regression. There are several statistical tools used for assessing the homogeneity of variance assumptions. The main purpose of this paper is to make available a concise overview of how to assess both univariate and multivariate homogeneity of variance assumption, using both graphical techniques and formal statistical test. To illustrate how to apply the methods for assessing univariate and multivariate homogeneity of variance, we used *rust inhibitor* and *technical dietary information* data, respectively.

2. Assessing Homogeneity of Variance

2.1. Assessing homogeneity of variance for univariate data

In the univariate data setting, several graphical and formal tests are available for checking the homogeneity of variance assumption. In case of assessing homogeneity of variance using graphical methods, direct examination of the data is usually not helpful in assessing non-constant error variance, especially if there are many predictors. We look to the distribution of the residuals to try and discover the distribution of the errors. Moreover, it is not helpful to plot the response against the residuals, because there is a built-in correlation between response variable and error term. The least squares fit ensures that the correlation between the estimated mean response and error term is zero, so a residual plot can help us uncover non-constant error variance. Therefore, plotting the residuals against the fitted values can prove to be very useful for assessing homogeneity of variance. In addition, one can use residual versus fitted value square and/or residual versus predictor variable. In this paper we use the most common graphical method, residual versus fitted value plot to illustrate the graphical method of assessing homogeneity of variance.

To have some impression about the graphical method of assessing homogeneity of variance let us consider the residual plots panel presented (Figure 1). From the Figure, the plot of residual versus fitted value presented in Figure 1a, one can clearly observe that residuals increase with fitted value in a pattern, this indicate that the error variances may not be constant. In contrast, Figure 1b, the vertical width of the scatter doesn't appear to increase or decrease across the fitted values, so we may assume that the variance in the error terms is constant. However, observing the scatter plots merely may not be enough to conclude that the error variances are constant and hence it is very crucial to use formal tests.

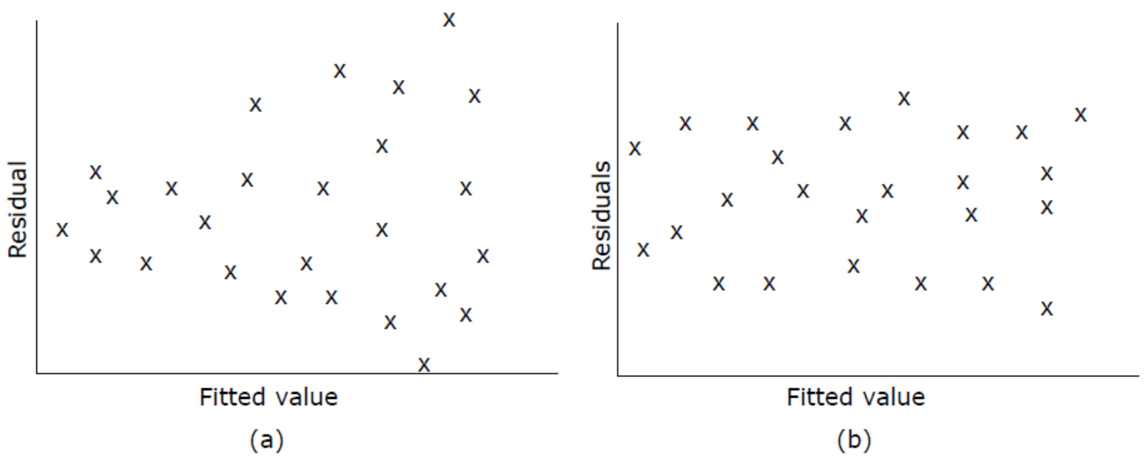

Figure 1. Sample residual plots for assessing constancy of variances.

2.1.1. Residual plot

In order to assess constancy of variance assumption, we plot residual versus fitted value. If the variances are constant, then we do not observe increasing or decreasing pattern in the vertical width across the fitted value. The data set we used for illustration has one continuous response variable and one categorical predictor variable with the total number of observations 40. Figure 2 below present the residual against fitted value plot. From the plot one can observe that the points are in vertical strips, this is due to the fact that the predictor variable is categorical with four categories and hence the are four vertical strips in the residual plot. From the figure, we can observe that the four vertical strips have almost equal length and hence this may suggest that the variances among the class are constant. However, graphical methods are very subjective, due to this; conclusions from the graphical methods should be made with cautious until the result supported by the formal tests.

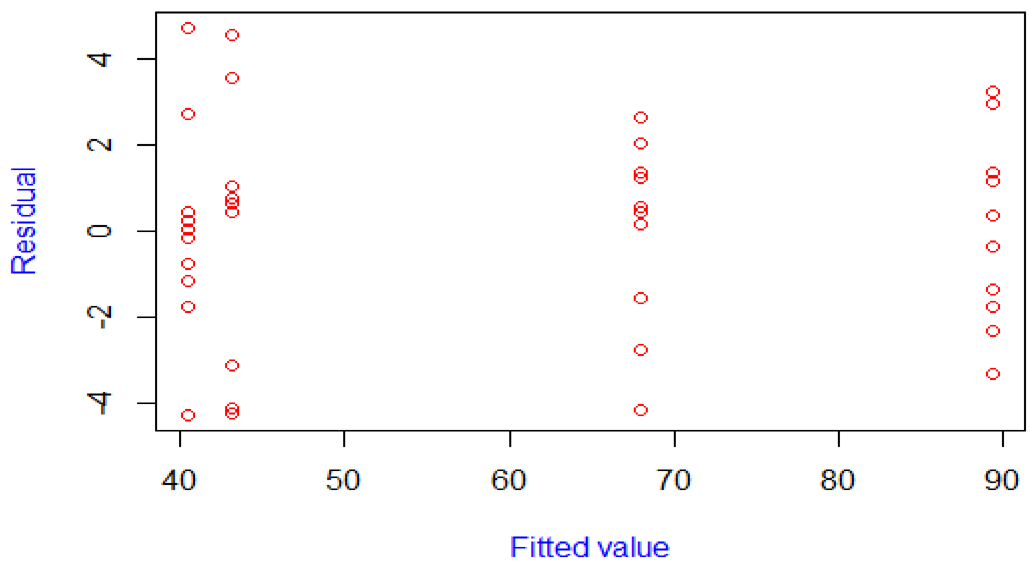

Figure 2. Residual against fitted value plot for assessing constancy of variances.

2.1.2. Formal tests for homogeneity of variances

To support the result of graphical methods one should apply formal tests. There are many existing tests for assessing homogeneity of variances, but the most common tests are Bartlett’s, Levene’s, Breusch-Pagan, Brown and Forsythe, O’Brien’s, White’s and Fligner-Killeen. The traditional test for homogeneity of variance is Bartlett’s (Barlett, 1937). While this is the optimal test under a normal model, it can be very inaccurate when the data deviate from normality (Box, 1953). Similarly, the Breusch-Pagan test assumes that the error terms are normally distributed (Neter et al., 2005). In contrast, Levene’s test, O’Brien’s test, and the Brown and Forsythe test are all based on an ANOVA for a dispersion variable derived from the dependent variable values. These tests are fairly robust to underlying distributions (Conover, Johnson, and Johnson, 1981). In this paper, the three main tests for homogeneity of variances: Bartlett's Test, Levene's Test, and the Brown-Forsythe Test were considered.

Bartlett's test is an older, alternative test. Bartlett's test is a chi-square statistic with (k-1) degrees of freedom, where k is the number of categories in the independent variable. The Bartlett's test is dependent on meeting the assumption of normality and therefore Levene's test has now largely replaced it. Levene's test of homogeneity of variance, which is the most common test, tests the assumption that each group (category) of one or more categorical independent variables has the same variance on an interval dependent. If the Levene statistic is significant at the .05 level or better, the researcher rejects the null hypothesis that the groups have equal variances. Levene's test is more robust in the face of non-normality than more traditional tests like Bartlett's test. Brown-Forsythe test is based on criticisms of the Levene test. It tests for equality of group means. The Brown-Forsythe test is more robust than the Levine test when groups are unequal in size and the absolute deviation scores (deviations from the group means) are highly skewed, causing a violation of the normality assumption and the assumption of equal variances (David, 2012).

The result for the three tests of homogeneity of variance is presented in Table 1. The null hypothesis for these tests is, the variance is constant (Homogeneous) against the alternative hypothesis the variance is not constant. At a 5% level of significance, we fail to reject the null hypothesis of homogeneity of variance since the p-value for all the three tests is much higher than the level of significance 5%. Therefore, the homogeneity of variance assumption is reasonably met.

Table 1. Tests for univariate constancy of variance.

Test	Test statistics	*P*-value
Levene's	0.42	0.7381
Bartlett's	1.199	0.7533
Brown and Forsythe's	0.23	0.8775

2.2. Assessing homogeneity of variance for multivariate data

Unlike homogeneity of variance tests in the univariate data case, the multivariate data has a very few tests available. In multivariate homogeneity of variance test, we test for the equality of variance-covariance matrix not a single numeric value of variance. The Box’s M and Bartlett's test are the most common tests available for testing homogeneity of variance-covariance matrix. Box's M tests the multivariate homogeneity of variances and covariance’s, as required by MANOVA and some other multivariate methods (IBM SPSS, 2012). When the result from this test is not significant, the researcher accepts the null hypothesis that groups do not differ. It has been shown to be a conservative test, failing to reject the null hypothesis too often. It also is highly sensitive to violations of multivariate normality (David, 2012). One can easily apply Box’s M test using available statistical software’s like SPSS and R. The alternative test for homogeneity of variance-covariance is Bartlett's test which can be in SAS using PROC DISCRIM procedure (SAS, 2010). Again if we were to reject the null hypothesis of homogeneity of variance-covariance matrices, then we would conclude that assumption is violated. MANOVA is not robust to violations of the assumption of homogeneous variance-covariance matrices and hence if the variance-covariance matrices are determined to be unequal then the solution is to find a variance-stabilizing transformation.

The result for both the Box’s M and Bartlett's tests of homogeneity of variance-covariance is presented in Table 2. The null hypothesis for these tests is, the variances and covariance’s are constant (Homogeneous) against the alternative hypothesis the variances and covariance’s are not constant. At a 5% level of significance, we fail to reject the null hypothesis of homogeneity of variance-covariance since the p-value for both tests is much higher than the level of significance 5%. Therefore, the homogeneity of variance and covariance assumption is reasonably met.

Table 2: Tests for multivariate constancy of variance.

Test	Test statistics	*P*-value
Box's M	0.8544	0.5938
Bartlett's	10.2858	0.5909

3. Conclusion

The homogeneity of variance assumption widely used practically for both univariate and multivariate data. The reliability of the results from both the univariate and multivariate statistical methods rely on this assumption. Several statistical tools and methods developed to help researchers in course of assessing this assumption. In univariate statistical methods, the residuals plot is very helpful graphical method for assessing homogeneity of variances. In addition, to supplement the graphical methods with formal test, statistical tests like Bartlett's, Levene's, Breusch-Pagan, Brown and Forsythe, O'Brien's, White's and Fligner-Killeen can be used. These tests are easily implemented in most statistical software including SAS and R. However, testes like Breusch-Pagan and Bartlett's are very sensitive to the violation of normality assumptions and hence care should be given when one select method out of all the alternatives.

For multivariate data the options are very limited as compared to univariate. Formal tests, Box's M and Bartlett's tests are widely available methods of testing the equality of variance and covariance. However, these tests are sensitive to the violation normality assumption and hence one should first check the normality assumptions.

Most of the univariate and multivariate methods not robust to violations of the assumption of constant variance and hence if the variances are determined to be unequal then the solution is to find a variance-stabilizing transformation (Neter et al., 2005; Johnson and Wichern, 2007).

References

Bartlett, M.S. (1937). Properties of Sufficiency and Statistical Tests, Proceedings of the Royal Society of London, Series A 160, 268–282.
Box, G.E.P. (1953), Non-normality and Tests on Variance, Biometrika, 40, 318–335.
Brown,M.B. and Forsythe, A.B. (1974). Robust Tests for Equality of Variances, Journal of the American Statistical Association, 69, 364–367.
Conover, W.J., Johnson, M.E., and Johnson, M.M. (1981). A Comparative Study of Tests for Homogeneity of Variances, with Applications to the Outer Continental Shelf Bidding Data, Technometrics, 23, 351–361.
David, G.G. (2012). Testing Statistical Assumptions. Statistical Associates Publishing.
IBM SPSS (2012). IBM SPSS Advanced Statistics 21. IBM Corporation.
Johnson, R. A., and Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis*. (6th ed.). Upper Saddle River, NJ: Prentice-Hall.
Neter, J., Kutner, M. H., Nachtsheim, C. J., and Wasserman, W. (2005). *Applied Linear Statistical Models*. (5th ed.). New York: McGraw-Hill Education.
SAS Institute. (2010). SAS 9.22 User's Guide. Cary, NC: SAS Institute Inc.