Journal of Statistical Modelling: Theory and Applications
Vol. 3, No. 2, 2022, pp. 31-38
Yazd University Press 2022

JSMTA

Research Paper

Inference on linear models with unequal variances

Ali Akbar Jafari*
Department of Statistics, Yazd University, Yazd, Iran

Received: March 31, 2023/ Revised: June 15, 2023/ Accepted: June 16, 2023

Abstract: In this paper, we consider a diagonal form for the variances of errors in linear models. This form contains the homogeneous and heterogeneous for the errors. First, an estimation for the variances is given, and then a method is introduced for the hypothesis test of parameters in linear models. Some applications of this method are presented.

Keywords: Behrens-Fisher; Generalized p-value; Heterogeneous; Linear model; One-way ANOVA.
Mathematics Subject Classification (2010): 62J05, 62J10.

1 Introduction

In the linear model, one of the problems is finding an exact test when the variances of errors are unequal. It is known as the Behrens-Fisher problem. Tsui and Weerahandi (1989) introduced the concept of generalized p-value (GPV) and proposed a test for the Behrens-Fisher problem to compare the means of two normal populations. Weerahandi (1987) and Koschat and Weerahandi (1992) gave tests for comparing several multiple linear regression models. Under unequal variances, an exact test for one-way ANOVA was proposed by Weerahandi (1995a) and for two-way ANOVA by Ananda and Weerahandi (1994) and Bao and Ananda (2001). A method to test the common mean of several normal populations has been introduced by Lin and Lee (2005).

In this article, we consider a diagonal matrix form for variances of errors in linear models. This form contains many forms for the variance of errors. First, we will obtain an estimation of unknown parameters of the models and nuisance parameters with some considerations. Then, we will introduce a GPV to test the linear models with unequal variances.

*Corresponding author: aajafari@yazd.ac.ir

In Section 2, the theory of GPV will be introduced. Section 3 is devoted to introducing the model. Then, an estimation of the unknown parameters and a method to test the hypothesis are given. The applications of our method are presented in Section 4.

2 Generalized p-value

The concept of GPV was first introduced by Tsui and Weerahandi (1989) to deal with the statistical testing problem in which nuisance parameters are present and it is difficult or impossible to obtain a nontrivial test with a fixed level of significance. The setup is as follows. Let X be a random variable having a probability density function $f(x|\boldsymbol{\zeta})$, where $\boldsymbol{\zeta} = (\theta, \boldsymbol{\eta})$ is a vector of unknown parameters, θ is the parameter of interest, and $\boldsymbol{\eta}$ is a vector of nuisance parameters. Suppose that we are interested in testing

$$H_\circ : \theta \geq \theta_\circ \;\; vs \;\; H_1 : \theta > \theta_\circ, \tag{1}$$

where $\theta_\circ$ is a specified value.

Let x denote the observed value of X. The random variable $T(X; x, \boldsymbol{\zeta})$ is called a generalized test variable if it has the following requirements:
(i) For fixed x and $\boldsymbol{\zeta} = (\theta_\circ, \boldsymbol{\eta})$, the distribution of $T(X; x, \boldsymbol{\zeta})$ is free of the nuisance parameters $\boldsymbol{\eta}$.
(ii) $t_{obs} = T(x; x, \boldsymbol{\zeta})$ does not depend on unknown parameters.
(iii) For fixed x and $\boldsymbol{\eta}$, the random variable $T(X; x, \boldsymbol{\zeta})$ is either stochastically increasing or decreasing in θ for any given t.

Under the above conditions, if $T(X; x, \boldsymbol{\zeta})$ is stochastically increasing in θ, then the GPV to test the hypothesis in (1) can be defined as

$$p = \sup_{\theta \geq \theta_\circ} P(T(X; x, \theta, \boldsymbol{\eta}) \geq t^*) = P(T(X; x, \theta_\circ, \boldsymbol{\eta}) \geq t^*),$$

where $t^* = T(x; x, \theta_\circ, \boldsymbol{\eta})$.

For further details and several applications based on the GPV, we refer to the book by Weerahandi (1995a). The GPV has been used for inference on ANOVA, regression and quantiles for example see Weerahandi (1995b), Sadooghi-Alvandi et al. (2012, 2015) and Malekzadeh and Jafari (2018).

3 Model and method

Consider the linear model

$$\boldsymbol{Y} = A\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon} \sim MVN(\boldsymbol{0}, D),$$

where $\boldsymbol{Y}$ is an $n \times 1$ vector of observations such that $\boldsymbol{Y} = (\boldsymbol{Y}_1', \ldots, \boldsymbol{Y}_t')'$ and $\boldsymbol{Y}_i$ is an $n_i \times 1$ vector ($i = 1, \ldots, t$, $n_i \geq 2$, $\sum_{i=1}^t n_i = n$) , A is an $n \times k$ matrix of rank $r < k$, $\boldsymbol{\beta}$ is an $k \times 1$ vector of parameters, $\boldsymbol{\varepsilon}$ is an $n \times 1$ vector of errors and D is an $n \times n$ unknown matrix

$$D = \mathrm{diag}(D_1, D_2, \ldots, D_t),$$

where $D_i = \sigma_i^2 I_i$ $(i = 1, \dots, t)$ and I_i is an $n_i \times n_i$ identity matrix. Therefore

$$Cov(\boldsymbol{Y}_i, \boldsymbol{Y}_j) = \boldsymbol{0}, \quad Var(\boldsymbol{Y}_i) = \sigma_i^2 I_i = D_i, \quad i \neq j, i = 1, 2, .., t.$$

This model contains linear regression, one-way and two-way ANOVA, ANCOVA, common mean problem, Chow test, and other linear models with equal and unequal variances.

Consider the hypothesis test

$$H_\circ : H\boldsymbol{\beta} = \boldsymbol{C} \quad vs \quad H_1 : H\boldsymbol{\beta} \neq \boldsymbol{C}, \tag{2}$$

where H is an $m \times k$ matrix $(m < k)$, $\boldsymbol{C}$ is an $m \times 1$ vector, $rank(H) = m < k$, and $H_\circ : H\boldsymbol{\beta} = \boldsymbol{C}$ is testable.

The general least square estimation for $\boldsymbol{\beta}$ is given as

$$\hat{\boldsymbol{\beta}} = (A'D^{-1}A)^{-}A'D^{-1}\boldsymbol{Y} + \left[I - (A'D^{-1}A)^{-}(A'D^{-1}A)\right]\gamma, \;\; \gamma \in R,$$

where I is an $k \times k$ identity matrix and B^{-} is generalized inverse of B. Since $H\boldsymbol{\beta}$ is estimable then the least square estimation of $H\boldsymbol{\beta}$ is unique for any choice of a least square estimation for $\boldsymbol{\beta}$ (see Kshirsagar, 1983), that is

$$\widehat{H\boldsymbol{\beta}} = H(A'D^{-1}A)^{-}A'D^{-1}\boldsymbol{Y}.$$

Let the matrix A be partitioned as

$$A = [a_{ij}]_{i,j=1}^{n} = (A_1^{*'}, A_2^{*'}, \dots, A_t^{*'})',$$

where $A_i^* = (A'_{i1}, A'_{i2}, \dots, A'_{in_i})'$, $A_{ij} = (a_{v1}, a_{v2}, \dots, a_{vk})$ and $v = j + \sum_{l=1}^{i-1} n_l$, $i = 1, 2, \dots, t$. It is clear that the MLE for σ_i^2, $i = 1, 2, \dots, t$ is

$$S_i^2 = \hat{\sigma}_i^2 = \frac{1}{n_i} Y_i^{'}(I_i - A_i^{*'}(A_i^{*'}A_i^*)^{-}A_i^*)\boldsymbol{Y}_i,$$

and $U_i = \frac{n_i \hat{\sigma}_i^2}{\sigma_i^2} \sim \chi^2_{(n_i - p_i)}$, with $p_i = rank(A_i^*)$. We remark that $\hat{\sigma}_i^2$ is not necessarily independent from the estimation of $\hat{\boldsymbol{\beta}}$.

Suppose that

$$Q = (\widehat{H\boldsymbol{\beta}} - H\boldsymbol{\beta})'V^{-1}(\widehat{H\boldsymbol{\beta}} - H\boldsymbol{\beta}),$$

where $V = H(A'D^{-1}A)^{-}H'$. Threfore, Q has a chi-square distribution with m degrees of freedom, since $\widehat{H\boldsymbol{\beta}} \sim MVN_m(H\boldsymbol{\beta}, V)$.

Theorem 3.1. *Let $\tilde{V} = H(A'\tilde{D}^{-1}A)^{-}H'$, where $\tilde{D} = \mathrm{diag}(\tilde{D}_1, \tilde{D}_2, \dots, \tilde{D}_t)$, $\tilde{D}_i = \sigma_i^2 \frac{s_i^2}{S_i^2} I_i = \frac{n_i s_i^2}{U_i} I_i$, and s_i^2 is the observed value of S_i^2, $i = 1, 2, \dots, t$. Let*

$$T = \frac{(\widehat{H\boldsymbol{\beta}} - H\boldsymbol{\beta})'V^{-1}(\widehat{H\boldsymbol{\beta}} - H\boldsymbol{\beta})}{(\widehat{H\boldsymbol{\beta}}_{obs} - H\boldsymbol{\beta})'\hat{V}^{-1}(\widehat{H\boldsymbol{\beta}}_{obs} - H\boldsymbol{\beta})}, \tag{3}$$

where $\widehat{H\boldsymbol{\beta}}_{obs} = H(A'\tilde{D}^{-1}A)^{-}A'\tilde{D}^{-1}\boldsymbol{y}$ and $\boldsymbol{y}$ is the observed value of $\boldsymbol{Y}$. Then T is a generalized test variable for $H\boldsymbol{\beta}$.

Proof. It is simple to show that the observed value of T, is equal to 1 and the distribution of T is free from nuisance parameters, σ_i^2, $i = 1, .., t$. Therefore T is a generalized test variable for $H\boldsymbol{\beta}$. □

The GPV for (2) using the generalized test variable in (3) is

$$p = P_{H_\circ}(T \geq t_{obs}) = 1 - E\left[H_m\left((\widehat{H\boldsymbol{\beta}}_{obs} - C)'\left[H(A'\tilde{D}^{-1}A)^{-}H'\right]^{-1}(\widehat{H\boldsymbol{\beta}}_{obs} - C)\right)\right], \tag{4}$$

where $H_m(\cdot)$ is the cumulative distribution function (cdf) of a chi-square distribution with m degrees of freedom, and expectation is taken with respect to independent chi-square random variables with $n_i - p_i$, $i = 1, \ldots, t$ degrees of freedom. The test rejects the null hypothesis if the GPV is smaller than the significance level.

This GPV can be computed by numerical integration but can also be well approximated by a Monte Carlo simulation method. In this method, a large and equal number of a random number from each chi-square random variable with $n_i - p_i$, $i = 1, \ldots, t$ degrees of freedom is generated. The cdf H_m is evaluated at each set, and the expectation is estimated by their sample mean (see Weerahandi, 1995b).

4 Applications

In this section, we show some applications of the given GPV in Section 3.

4.1 One Way ANOVA

Consider the one-way ANOVA model

$$Y_{ij} = \mu_i + \varepsilon_{ij}, \;\; i = 1, \ldots, k, \;\; j = 1, \ldots, n_i.$$

Testing the equality of means

$$H_\circ : \mu_1 = \cdots = \mu_k. \tag{5}$$

is equivalent to $H_\circ : H\boldsymbol{\mu} = 0$, where $H = [\mathbf{1} : H^*]$, $\mathbf{1} = (1, \ldots, 1)'$, $H^* = \text{diag}(-1, \ldots, -1)$ and $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_k)'$. We can show that rank($H$)$= k - 1$.

If $\sigma_1^2 = \cdots = \sigma_k^2 = \sigma^2$, then by notations given in Section 3, we have $t = 1$ and $A = A_1^*$. We obtain $\hat{\mu}_i = \hat{\beta}_i = \bar{Y}_{i.} = \frac{1}{n_i}\sum_{j=1}^{n_i} Y_{ij}$ and $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij} - \bar{Y}_{i.})^2 = \frac{1}{n}SSE$. Therefore, the GPV in (4) is the regular p-value for one-way ANOVA based on F test.

If σ_i^2's are not equal then $t = k$ and the generalized test variable in (3) for the hypothesis test (5) is

$$T = \frac{(H\bar{\boldsymbol{Y}} - H\boldsymbol{\mu})'\left[HVH'\right]^{-1}(H\bar{\boldsymbol{Y}} - H\boldsymbol{\mu})}{(H\bar{\boldsymbol{y}} - H\boldsymbol{\mu})'\left[H\hat{V}H'\right]^{-1}(H\bar{\boldsymbol{y}} - H\boldsymbol{\mu})}, \tag{6}$$

where $\bar{Y} = (\bar{Y}_1, \ldots, \bar{Y}_k)$, $V = \text{diag}(\frac{\sigma_1^2}{n_1}, \ldots, \frac{\sigma_k^2}{n_k})$, $\hat{V} = \text{diag}(\frac{\sigma_1^2 s_1^2}{n_1 S_1^2}, \ldots, \frac{\sigma_k^2 s_k^2}{n_k S_k^2})$, $S_i^2 = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij} - \bar{Y}_{i.})^2$ and $\bar{y}$ and s_i^2 are the observed values of $\bar{Y}$ and S_i^2, respectively.

Therefore, the GPV for (5) is

$$p = P(T \geq 1) = 1 - E\left[F_{\chi^2_{(k-1)}}\left((H\bar{\boldsymbol{y}})'\left[H\hat{V}H'\right]^{-1}(H\bar{\boldsymbol{y}})\right)\right],$$

where $F_{\chi^2_{(k-1)}}(.)$ is the cdf of chi-square distribution with $k-1$ degrees of freedom. This GPV is given by Witkovsky (2002). We can show that this GPV is equal to the GPV given by Weerahandi (1995a) using the following theorem.

Theorem 4.1. *For given T in* (6)*, we have*

$$(H\bar{\boldsymbol{Y}})'\left[HVH'\right]^{-1}(H\bar{\boldsymbol{Y}}) = \sum_{i=1}^{k}\frac{n_i\bar{Y}_i^2}{\sigma_i^2} - \frac{(\sum_{i=1}^{k}\frac{n_i l\bar{Y}_i}{\sigma_i^2})^2}{\sum_{i=1}^{k}\frac{n_i}{\sigma_i^2}}.$$

Proof. We can show that $HVH' = B + \boldsymbol{CC}'$, where $\boldsymbol{C} = \frac{\sigma_1}{\sqrt{n_1}}(1,1,\ldots,1)'$ and $B = \text{diag}(\frac{\sigma_2^2}{n_2},\ldots,\frac{\sigma_k^2}{n_k})$. Therefore, $(B+\boldsymbol{CC}')^{-1} = B^{-1} - \frac{B^{-1}\boldsymbol{CC}'B^{-1}}{1+\boldsymbol{C}'B^{-1}C}$, (see Rencher, 2000) and $(HVH')^{-1} = \text{diag}(\frac{n_2}{\sigma_2^2},\ldots,\frac{n_k}{\sigma_k^2}) - \frac{1}{\sum_{i=1}^{k}\frac{n_i}{\sigma_i^2}}(\frac{n_2}{\sigma_2^2},\ldots,\frac{n_k}{\sigma_k^2})'(\frac{n_2}{\sigma_2^2},\ldots,\frac{n_k}{\sigma_k^2})$. Then, we have

$$\begin{aligned}(H\bar{\boldsymbol{Y}})'[HVH']^{-1}(H\bar{\boldsymbol{Y}}) &= \sum_{i=1}^{k}(\bar{Y}_1-\bar{Y}_k)^2\frac{n_i}{\sigma_i^2} - \frac{\left[\sum_{i=1}^{k}(\bar{Y}_1-\bar{Y}_k)\frac{n_i}{\sigma_i^2}\right]^2}{\sum_{i=1}^{k}\frac{n_i}{\sigma_i^2}} \\ &= \sum_{i=1}^{k}\frac{n_i\bar{Y_k}^2}{\sigma_i^2} - \frac{\left[\sum_{i=1}^{k}\frac{n_i\bar{Y}_i}{\sigma_i^2}\right]^2}{\sum_{i=1}^{k}\frac{n_i}{\sigma_i^2}}\end{aligned}$$

□

Remark 4.2. *For two-way ANOVA, the GPV in* (4) *is equivalent the GPVs introduced by Bao and Ananda (2001).*

4.2 Common mean

Consider k different normal populations with a common mean μ but possibility different variances σ_i^2, $i = 1,..,k$, that is

$$Y_{ij} = \mu + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim N(0,\sigma_i^2), \quad i = 1,\ldots,k, \;\; j = 1,\ldots,n_i.$$

Suppose that we are interested in the hypothesis test

$$H_\circ : \mu = \mu_\circ \;\; vs \;\; H_1 : \mu \neq \mu_\circ.$$

Then by using (3), the generalized test variable for this hypothesis test is

$$T = \frac{Z^2}{\sum_{i=1}^{k}\frac{U_i}{s_i^2}\left(\frac{\sum_{i=1}^{k}\frac{U_i\bar{y}_i}{s_i^2}}{\sum_{i=1}^{k}\frac{U_i}{s_i^2}} - \mu\right)^2},$$

where $Z^2 = \sum_{i=1}^{k} \frac{n_i}{\sigma_i^2} \left(\frac{\sum_{i=1}^{k} \frac{n_i \bar{Y}_i}{\sigma_i^2}}{\sum_{i=1}^{k} \frac{n_i}{\sigma_i^2}} - \mu \right)$ and $U_i = \frac{n_i S_i^2}{\sigma_i^2}$, and $S_i^2 = \frac{1}{n_i} \sum_{i=1}^{k} (Y_{ij} - \bar{Y}_{i.})^2$. Therefore, the GPV is

$$p = 1 - E\left[F_{\chi^2_{(1)}} \left(\frac{\left(\sum_{i=1}^{k} \frac{U_i \bar{y}_i}{s_i^2} - \mu_\circ \sum_{i=1}^{k} \frac{U_i}{s_i^2} \right)^2}{\left(\sum_{i=1}^{k} \frac{U_i}{s_i^2} \right)} \right) \right].$$

This GPV is equivalent to the GPV introduced by Lin and Lee (2005).

4.3 Chow-Test

Consider the test of equality of sets of coefficient in t linear regressions, that is

$$Y_{ij} = \boldsymbol{x}'_{ij} \boldsymbol{\beta}_i + \varepsilon_{ij}, \;\; \varepsilon_{ij} \sim N(0, \sigma_i^2), \; i = 1, 2, \ldots, t, \; j = 1, 2, \ldots, n_i,$$

where Y_{ij} is jth observation in ith regression, $\boldsymbol{x}_{ij}$ is an $(k \times 1)$ vector, and $\boldsymbol{\beta}_i$ is an $(k \times 1)$ vector of coefficient. Our problem of interest is the hypothesis test

$$H_\circ : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \cdots = \boldsymbol{\beta}_t.$$

This hypothesis test can be written as

$$H_\circ : H\boldsymbol{\beta} = 0,$$

where $\boldsymbol{\beta} = (\boldsymbol{\beta}'_1, \boldsymbol{\beta}'_2, \ldots, \boldsymbol{\beta}'_t)'$ and H is an $k(t-1) \times kt$ matrix with rank $k(t-1)$,

$$H = \begin{bmatrix} I & -I & 0 & \cdots & 0 \\ I & 0 & -I & & 0 \\ \vdots & \vdots & & \ddots & \\ I & 0 & 0 & \cdots & -I \end{bmatrix},$$

where I is $k \times k$ identity matrix.

We can show that $\hat{\sigma}_i^2 = \frac{1}{n_i} \boldsymbol{Y}'_i (I_i - X_i (X'_i X_i)^{-1} X'_i) \boldsymbol{Y}_i = \frac{1}{n_i} SSE_i$, where $\boldsymbol{Y}_i = (Y_{i1}, \ldots, Y_{in_i})'$ and $X_i = (\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, \ldots, \boldsymbol{x}_{in_i})'$, $i = 1, 2, \ldots, t$. Therefore, the GPV for hypothesis test in (3) is

$$p = 1 - E\left[H_{k(t-1)} \left\{ (\widehat{H\boldsymbol{\beta}}_{obs})' \left[H (X' \hat{D}^{-1} X)^{-1} H' \right]^{-1} (\widehat{H\boldsymbol{\beta}}_{obs}) \right\} \right] \quad (7)$$

where $X = \text{diag}(X_1, X_2, \ldots, X_t)$.

If $t = 2$ then the GPV in (7) is equivalent to the GPV given by Koschat and Weerahandi (1992).

5 Discussion and conclusions

The linear model $\boldsymbol{Y} = A\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, contains linear regression, one-way and two-way ANOVA, ANCOVA, common mean, and Chow test. Here, we consider diagonal form

$D = \text{diag}(D_1, D_2, \ldots, D_t)$ for the variances of errors, where $D_i = \sigma_i^2 I_i$ $(i = 1, \ldots, t)$. When σ_i^2 are equal, the classical F test is an exact test for inference on the parameters. When σ_i^2 are unequal, inference on the linear model becomes to Behrens-Fisher problem which is one of the old problems in statistics. There is no exact method for this problem. Here, we proposed a GPV for inference on this model. In special case, for one-way ANOVA, this GPV becomes the GPV proposed by Weerahandi (1995a). For two-way ANOVA, common mean, and Chow-test this GPV becomes the GPVs proposed by Bao and Ananda (2001), Lin and Lee (2005), and Koschat and Weerahandi (1992). However, this GPV can be applied to other models such as ANCOVA and three-way ANOVA

Acknowledgement

The authors are grateful to anonymous referee for their helpful comments and suggestions to improve this manuscript.

References

Ananda, M.M.A. and Weerahandi, S. (1994). Two-way ANOVA with unequal cell frequencies and unequal variances. Technical Report, Department of Mathematics, University of Nevada, Las Vegas.

Bao, P. and Ananda M.M.A. (2001). Performance of two-way ANOVA procedures when cell frequencies and variances are unequal. *Communication in Statistics- Simulation and Computation*, **30**(4):805–829.

Koschat, M. and Weerahandi, S. (1992). Chow-Type test under heteroscedasticity. *Journal of Business and Economic Statistics*, **10**(2):221–228.

Kshirsagar, M.A. (1983). *A Course in Linear Model*. New York: Marcel Decker.

Lin, S.H. and Lee, J.C. (2005). Generalized inferences on the common mean of several normal populations. *Journal of Statistical Planning and Inference*, **134**:568–582.

Malekzadeh, A. and Jafari, A.A. (2018). Testing equality of quantiles of two-parameter exponential distributions under progressive Type II censoring. *Journal of Statistical Theory and Practice*, **12**:776–793.

Rencher, A.C. (2000). *Linear Models in Statistics*. New York: John Willy and Sons.

Sadooghi-Alvandi, S.M., Jafari, A.A. and Mardani-Fard, H.A. (2012). One-way ANOVA with unequal variances. *Communications in Statistics-theory and Methods*, **41**(22):4200–4221.

Sadooghi-Alvandi, S.M., Jafari, A.A. and Mardani-Fard, H.A. (2016). Comparing several regression models with unequal variances. *Communications in Statistics-Simulation and Computation*, **45**(9):3190–3216.

Tsui, K.W. and Weerahandi, S. (1989). Generalized p-value in significance testing of hypothesis in the presence of nuisance parameters. *Journal of American Statistical Association*, **84**:602–607.

Weerahandi, S. (1987). Testing regression with unequal variances. *Econometrics*, **55**:1211–1215.

Weerahandi, S. (1995a). ANOVA under unequal error variances. *Biometrics*, **51**:589–599.

Weerahandi, S. (1995b). *Exact Statistical Methods for Data Analysis*. New York: Springer-Verlag.

Witkovsky, V. (2002). On the Behrens-Fisher distribution and its generalization to pairwise comparisons. *Discussiones Mathematicae-Probability and Statistics*, **22**:73–104.