

A Note on Cochran Test for Homogeneity in Two Ways ANOVA and Meta-Analysis

Pamphile Mezui-Mbeng

CIREGED, Department of Economics, Omar Bongo University, Libreville, Gabon
Email: pmezuibeng@gmail.com

Received 24 August 2015; accepted 27 December 2015; published 30 December 2015

Copyright © 2015 by author and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY).
<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In this paper, we generalize the proof of the Cochran statistic in the case of an ANOVA two ways structure that asymptotically follows a Chi-2. While construction of homogeneity statistics test usually resorts to the determination of the covariance matrix and its inverse, the Moore-Penrose matrix, our approach, avoids this step. We also show that the Cochran statistic in ANOVA two ways is equivalent to conventional homogeneity statistics test. In particular, we show that it satisfies the invariance property. Finally, we conduct empirical verification from a meta-analysis that confirms our theoretical results.

Keywords

Cochran Homogeneity Test, Chi-Square Distribution, Desimonian-Laird Test, Invariant, Two Ways ANOVA

1. Introduction

In ANOVA methodology, it is generally accepted that the error variance is unknown and is the subject of an estimate. However, in practice, these fundamental assumptions are rarely checked, forcing the use of Fisher statistic in the homogeneity test on the mean of the different groups ([1]-[5]).

According to the work of [6], the statistic test of homogeneity of two ways is of the form:

$$C = \sum_{i=1}^K \sum_{j=1}^L \omega_{ij} \left(\bar{y}_{ij} - \sum_{n=1}^K \sum_{m=1}^L h_{nm} \bar{y}_{nm} \right)^2 \quad (1)$$

where $\omega_{ij} = \frac{n_{ij}}{S_{ij}^2}$, \bar{y}_{ij} and S_{ij}^2 are respectively the mean and variance sample of the group (i, j) consisting of

n_{ij} observations; $h_{ij} = \frac{\omega_{ij}}{\omega}$, with $\omega = \sum_{i=1}^K \sum_{j=1}^L \omega_{ij}$. In the foregoing expression, the number of groups is equal to

KL, and Equation (1) is valid if $S_{ij} > 0, \forall (i, j) \in \{1, \dots, K\} \times \{1, \dots, L\}$.

[7] tests the homogeneity of medical treatment between both groups of patients, using the [6] statistic in a meta-analysis (e.g. see [11]). The above studies suggest that under the null hypothesis H_0 of equality of the means of different groups (i, j) , the Cochran statistic asymptotically follows a χ^2_{KL-1} . However, neither the work of [6], nor those of [7] offer a formal proof of this result.

Despite the existence of some attempts proposed by [8]-[10] and [11], in the literature, the construction of homogeneity statistical test on mean (or medians and percentiles) of various groups is generally based on a three-step methodology.

Step 1: The global average is estimated by a linear combination of the individual averages.

$$\bar{y} = \sum_{i=1}^K \sum_{j=1}^L p_{ij} \bar{y}_{ij}, \text{ where } \bar{y}_{ij} \text{ represents respectively the mean and the non-negative weight of group } (i, j),$$

and $\sum_{i=1}^K \sum_{j=1}^L p_{ij} = 1$.

Step 2: One assumes that the population variances of each group are unknown and estimated by the variances of the corresponding samples.

Let the vector q be given by: $q = (q_1, q_2, \dots, q_K)'$ where, $q_i = (\bar{y}_{i1} - \bar{y}, \bar{y}_{i2} - \bar{y}, \dots, \bar{y}_{iL} - \bar{y})'$. We have:

$$q = (\bar{y}_{11} - \bar{y}, \dots, \bar{y}_{1L} - \bar{y}, \bar{y}_{21} - \bar{y}, \dots, \bar{y}_{2L} - \bar{y}, \dots, \bar{y}_{K1} - \bar{y}, \dots, \bar{y}_{KL} - \bar{y})'$$

$$q = Q(\bar{y}_{11}, \dots, \bar{y}_{1L}, \bar{y}_{21}, \dots, \bar{y}_{2L}, \dots, \bar{y}_{K1}, \dots, \bar{y}_{KL})', \text{ with } Q = I_{KL} - P, \text{ } P = \begin{pmatrix} p_{11} & \dots & p_{1L} \\ \vdots & \ddots & \vdots \\ p_{K1} & \dots & p_{KL} \end{pmatrix}$$

The covariance matrix is then estimated by $\text{Var}(q) = Q\text{Var}(y)Q' = Q\Sigma Q'$, with

$$\Sigma = \begin{pmatrix} 1/\omega_{11} & 0 & \dots & 0 \\ 0 & 1/\omega_{12} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1/\omega_{KL} \end{pmatrix}$$

Step 3: One constructs the statistics test

$$G = q'\Psi^+ q$$

In explicit form,

$$G = q'\Psi^+ q = (\bar{y}_{11}, \dots, \bar{y}_{1L}, \bar{y}_{21}, \dots, \bar{y}_{2L}, \dots, \bar{y}_{K1}, \dots, \bar{y}_{KL})' [Q'(Q\Sigma Q')^+ Q] (\bar{y}_{11}, \dots, \bar{y}_{1L}, \bar{y}_{21}, \dots, \bar{y}_{2L}, \dots, \bar{y}_{K1}, \dots, \bar{y}_{KL})'$$

where Ψ^+ is the Moore-Penrose inverse matrix $(M - P)$ of Ψ .

In the case of one way ANOVA, [12] provides a faster method of building statistics homogeneity test, showing that this statistic is equivalent to Cochran. However, the authors offer no generalization of their result to the case of the two-factor ANOVA.

Following [12], this paper proposes to generalize the construction of the statistical homogeneity test in ANOVA two ways settings. To our knowledge, this issue has not been discussed in the literature. Beyond the theoretical importance, in practice it induces many applications, particularly in medicine, to compare the effectiveness of two methods of administration of a molecule to two different populations.

The remainder of the paper is organized as follows: Section 2 presents the main results. Section 3 provides an empirical evaluation of the proposed test; and Section 4 concludes the paper.

2. Main Results

In this section, we first show that statistics $T = d'\Gamma^{-1}d$ (see Equation (3) below) is asymptotically distributed according to a χ^2_{KL-1} . Then, we prove that the Cochran statistics in a two ways ANOVA is equivalent to T . Finally, we conclude that the C statistics also follows a χ^2_{KL-1} distribution. We thus have the following important results.

Proposition 1.

$$T_1 = d'\Gamma_1^{-1}d \approx \chi^2_{KL-1} \tag{2}$$

Proof.

We suppose that $\bar{y}_{ij} \approx N\left(\mu, \frac{\sigma_{ij}^2}{n_{ij}}\right)$ for the groups $(i, j) \in \{1, \dots, K\} \times \{1, \dots, L\}$ where the variance of the population σ_{ij}^2 is unknown.

Now let us consider $d_{ij} = \bar{y}_{ij} - \bar{y}_{11}$, for $(i, j) \in \{1, \dots, K\} \times \{1, \dots, L\} - \{1, 1\}$, then $E(d_{ij}) = 0$, $\text{var}(d_{ij}) = \text{var}(\bar{y}_{ij}) + \text{var}(\bar{y}_{11}) = v_{ij} + v_{11}$ and $\text{cov}(d_{ij}, d_{lm}) = v_{ij}\delta_{il}\delta_{jm} + (1 - \delta_{il}\delta_{1j} - \delta_{1l}\delta_{im})v_{11}$; with $v_{ij} = \sigma_{ij}^2/n_{ij}$.

Let us consider $d = (d_{12} - d_{11}, \dots, d_{1L} - d_{11}, d_{21} - d_{11}, \dots, d_{2L} - d_{11}, \dots, d_{K1} - d_{11}, \dots, d_{KL} - d_{11})'$, the variance-covariance matrix of d is written as follows:

$$\Gamma_1 = \begin{pmatrix} v_{12} + v_{11} & v_{11} & \dots & v_{11} \\ v_{11} & v_{13} + v_{11} & \dots & v_{11} \\ \vdots & \vdots & \ddots & \vdots \\ v_{11} & \dots & v_{11} & v_{KL} + v_{11} \end{pmatrix} = \begin{pmatrix} v_{12} & 0 & \dots & 0 \\ 0 & v_{13} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & v_{KL} \end{pmatrix} + v_{11} \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \dots & 1 & 1 \end{pmatrix}$$

According to [13], p.9, Theorem 1.7, the inverse of Γ_1 exists and it is given by:

$$\Gamma_1^{-1} = \begin{pmatrix} 1/v_{12} & 0 & \dots & 0 \\ 0 & 1/v_{13} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 1/v_{KL} \end{pmatrix} - \frac{1}{v} \begin{pmatrix} 1/v_{12} \\ 1/v_{13} \\ \vdots \\ 1/v_{KL} \end{pmatrix} (1/v_{12} \quad 1/v_{13} \quad \dots \quad 1/v_{KL}),$$

where $v = \sum_{i=1}^K \sum_{j=1}^L \frac{1}{v_{ij}}$. Therefore, one obtains the result. ■

Proposition 2.

$$T_1 = d'\Gamma_1^{-1}d \approx \chi^2_{KL-1} \tag{3}$$

Proof.

In practice, the variance σ_{ij}^2 is unknown and estimated from the variance of the sample s_{ij}^2 . Replacing in Equation (2) $v_{ij} = \sigma_{ij}^2/n_{ij}$ by $s_{ij}^2/n_{ij} = 1/\omega_{ij}$, and since,

$$\Gamma = \begin{pmatrix} 1/\omega_{12} + 1/\omega_{11} & 1/\omega_{11} & \dots & 1/\omega_{11} \\ 1/\omega_{11} & 1/\omega_{13} + 1/\omega_{11} & \dots & 1/\omega_{11} \\ \vdots & \vdots & \ddots & \vdots \\ 1/\omega_{11} & \dots & 1/\omega_{11} & 1/\omega_{KL} + 1/\omega_{11} \end{pmatrix} = \begin{pmatrix} 1/\omega_{12} & 0 & \dots & 0 \\ 0 & 1/\omega_{13} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 1/\omega_{KL} \end{pmatrix} + 1/\omega_{11} \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \dots & 1 & 1 \end{pmatrix}$$

Based on Slutsky Theorem, the statistic T is asymptotically distributed as a χ^2_{KL-1} distribution. ■

Proposition 3.

T and C are equivalent.

Proof.

Since,

$$\begin{aligned} C &= \sum_{i=1}^K \sum_{j=1}^L \omega_{ij} \left(\bar{y}_{ij} - \sum_{i=1}^K \sum_{j=1}^L h_{nm} \bar{y}_{nm} \right)^2 \\ C &= \sum_{i=1}^K \sum_{j=1}^L \omega_{ij} \left(\bar{y}_{ij} - \bar{y}_{11} + \bar{y}_{11} - \sum_{i=1}^K \sum_{j=1}^L h_{nm} \bar{y}_{nm} \right)^2 \\ C &= \sum_{i=1}^K \sum_{j=1}^L \omega_{ij} \left[(\bar{y}_{ij} - \bar{y}_{11})^2 + \left(\bar{y}_{11} - \sum_{n=1}^K \sum_{m=1}^L h_{nm} \bar{y}_{nm} \right)^2 + 2(\bar{y}_{ij} - \bar{y}_{11}) \left(\bar{y}_{11} - \sum_{n=1}^K \sum_{m=1}^L h_{nm} \bar{y}_{nm} \right) \right] \\ &= \sum_{(i,j) \neq (1,1)} \omega_{ij} (\bar{y}_{ij} - \bar{y}_{11})^2 + \left(\sum_{i=1}^K \sum_{j=1}^L \omega_{ij} \right) \left(\bar{y}_{11} - \sum_{n=1}^K \sum_{m=1}^L h_{nm} \bar{y}_{nm} \right)^2 + 2 \left(\bar{y}_{11} - \sum_{n=1}^K \sum_{m=1}^L h_{nm} \bar{y}_{nm} \right) \sum_{i=1}^K \sum_{j=1}^L \omega_{ij} (\bar{y}_{ij} - \bar{y}_{11}) \end{aligned}$$

and since,

$$\sum_{i=1}^K \sum_{j=1}^L \omega_{ij} (\bar{y}_{ij} - \bar{y}_{11}) = \sum_{i=1}^K \sum_{j=1}^L \omega h_{ij} (\bar{y}_{ij} - \bar{y}_{11}) = \omega \left(\sum_{i=1}^K \sum_{j=1}^L h_{ij} \bar{y}_{ij} - \bar{y}_{11} \right)$$

We obtain:

$$C = \sum_{(i,j) \neq (1,1)} \omega_{ij} (\bar{y}_{ij} - \bar{y}_{11})^2 + \omega \left(\bar{y}_{11} - \sum_{n=1}^K \sum_{m=1}^L h_{nm} \bar{y}_{nm} \right)^2 - 2\omega \left(\bar{y}_{11} - \sum_{n=1}^K \sum_{m=1}^L h_{nm} \bar{y}_{nm} \right)$$

Therefore, we get:

$$C = \sum_{(i,j) \neq (1,1)} \omega_{ij} (\bar{y}_{ij} - \bar{y}_{11})^2 - \omega \left(\bar{y}_{11} - \sum_{n=1}^K \sum_{m=1}^L h_{nm} \bar{y}_{nm} \right)^2$$

Also since, $T_1 = d' \Gamma_1^{-1} d$, i.e.

So that,

$$T = \sum_{(i,j) \neq (1,1)} \omega_{ij} (\bar{y}_{ij} - \bar{y}_{11})^2 - \omega \left[\bar{y}_{11} - \sum_{i=1}^K \sum_{j=1}^L h_{ij} \bar{y}_{ij} \right]^2$$

Therefore we obtain the equivalence between T and C . And as it was demonstrated that T is asymptotically distributed as a χ_{KL-1}^2 , then C also follows the same law. ■

Defining G by

$$\begin{aligned} T &= d' \begin{bmatrix} \omega_{12} & 0 & \dots & 0 \\ 0 & \omega_{13} & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \dots & 0 & \omega_{KL} \end{bmatrix} d - \frac{1}{\omega} d' \begin{bmatrix} \omega_{12} \\ \omega_{13} \\ \vdots \\ \omega_{KL} \end{bmatrix} \left[\omega_{12} \quad \omega_{13} \quad \dots \quad \omega_{KL} \right] d \\ &= \sum_{(i,j) \neq (1,1)} \omega_{ij} d_{ij}^2 - \frac{1}{\omega} \left(\sum_{(i,j) \neq (1,1)} \omega_{ij} d_{ij} \right)^2 = \sum_{(i,j) \neq (1,1)} \omega_{ij} (\bar{y}_{ij} - \bar{y}_{11})^2 - \frac{1}{\omega} \left[\sum_{(i,j) \neq (1,1)} \omega_{ij} (\bar{y}_{ij} - \bar{y}_{11}) \right]^2 \\ &= \sum_{(i,j) \neq (1,1)} \omega_{ij} (\bar{y}_{ij} - \bar{y}_{11})^2 - \omega \left[\sum_{(i,j) \neq (1,1)} h_{ij} (\bar{y}_{ij} - \bar{y}_{11}) \right]^2 \\ &= \sum_{(i,j) \neq (1,1)} \omega_{ij} (\bar{y}_{ij} - \bar{y}_{11})^2 - \omega \left[(1-h_{11}) \bar{y}_{11} - \sum_{(i,j) \neq (1,1)} h_{ij} \bar{y}_{ij} \right]^2 \end{aligned}$$

$$G = (\bar{y}_{11}, \dots, \bar{y}_{1L}, \bar{y}_{21}, \dots, \bar{y}_{2L}, \dots, \bar{y}_{K1}, \dots, \bar{y}_{KL}) \left[Q' (\mathcal{Q} \Sigma \mathcal{Q}')^+ \mathcal{Q} \right] (\bar{y}_{11}, \dots, \bar{y}_{1L}, \bar{y}_{21}, \dots, \bar{y}_{2L}, \dots, \bar{y}_{K1}, \dots, \bar{y}_{KL})' \quad (4)$$

G verifies the following invariance property.

Theorem 1.

The G statistics is invariant by the choice of the weights and $G = C$.

Proof of Theorem 1.

To prove this theorem, we need the following lemmas.

Lemma 1.

According to [14], p. 130, 7.11 (d) (ii), $X(X'X)^-X' = XX^+$ is invariant, where $(X'X)^-$ is the generalized inverse matrix of $X'X$ and X^+ is the inverse M-P matrix of X . Therefore, $X(X'X)^+X' = XX^+$.

Proof.

Straightforward. ■

Lemma 2.

According to [14] p. 144, 7.73, if A and B are compatible matrices, then $(AB)^+ = (A^+AB)^+(ABB^+)^+$.

Proof.

Straightforward. ■

Lemma 3.

For Q in (4), its singular value decomposition is, $Q = UDV'$, where $D = \text{diag} \left\{ \sqrt{(KL) \sum_{i=1}^K \sum_{j=1}^L p_{ij}^2}, 1, 1, \dots, 1, 0 \right\}$

and the k -th column of V is $v_k = \frac{1}{\sqrt{KL}}(1, 1, \dots, 1)'$

Proof.

It is easy to show that

$$Q'Q = I_{KL} - \begin{bmatrix} p_{11} & 1 \\ \vdots & \vdots \\ p_{1L} & 1 \\ \vdots & \vdots \\ p_{K1} & 1 \\ \vdots & \vdots \\ p_{KL} & 1 \end{bmatrix} \begin{bmatrix} 1 & \dots & 1 & \dots & 1 & \dots & 1 \\ p_{11} & \dots & p_{1L} & \dots & p_{K1} & \dots & p_{KL} \end{bmatrix} + \left(\sum_{i=1}^K \sum_{j=1}^L p_{ij}^2 \right) J_{KL}$$

where I_{KL} is the identity matrix of dimension $K \times L$ and J_{KL} is the squared matrix of dimension $K \times L$ whose elements are 1. The eigenvalues of $Q'Q$ are

$$(KL) \sum_{i=1}^K \sum_{j=1}^L p_{ij}^2; 1, 1, \dots, 1, 0.$$

Therefore, $D = \text{diag} \left\{ \sqrt{(KL) \sum_{i=1}^K \sum_{j=1}^L p_{ij}^2}, 1, 1, \dots, 1, 0 \right\}$ and the k -th column of V is $v_k = \frac{1}{\sqrt{KL}}(1, 1, \dots, 1)'$. ■

Lemma 4.

We have $Q^+Q = I_{KL} - \frac{1}{KL}J_{KL}$

Proof.

According to Lemma 3, $Q^+ = VD^+U'$, where $D^+ = \text{diag} \left\{ 1 / \sqrt{(KL) \sum_{i=1}^K \sum_{j=1}^L p_{ij}^2}, 1, 1, \dots, 1, 0 \right\}$. Therefore,

$$\begin{aligned} Q^+Q &= VD^+U'UDV' = VD^+DV' = V(I_{KL} - \text{diag}\{0, 0, \dots, 0, 1\})V' \\ &= VI_{KL}V' - V\text{diag}\{0, 0, \dots, 0, 1\}V' = I_{KL} - \begin{bmatrix} v_{1,KL}^2 & v_{1,KL}v_{2,KL} & \dots & v_{1,KL}v_{KL,KL} \\ v_{2,KL}v_{1,KL} & v_{2,KL}^2 & \dots & v_{2,KL}v_{KL,KL} \\ \vdots & \vdots & \ddots & \vdots \\ v_{KL,KL}v_{1,KL} & v_{KL,KL}v_{2,KL} & \dots & v_{KL,KL}^2 \end{bmatrix} = I_{KL} - \frac{1}{KL}J_{KL} \end{aligned}$$

■

From the above lemmas, we then can provide the proof of **Theorem 1**.

Proof of Theorem 1.

$$\begin{aligned}
 Q'(Q\Sigma Q')^+ Q &= \Sigma^{-1/2} \Sigma^{1/2} Q'(Q\Sigma Q')^+ Q \Sigma^{1/2} \Sigma^{-1/2} = \Sigma^{-1/2} \left(\Sigma^{1/2} Q'(Q\Sigma Q')^+ Q \Sigma^{1/2} \right) \Sigma^{-1/2} \\
 &= \Sigma^{-1/2} \left(\Sigma^{1/2} Q' (Q \Sigma^{1/2} \Sigma^{1/2} Q')^+ Q \Sigma^{1/2} \right) \Sigma^{-1/2} \\
 &= \Sigma^{-1/2} \left(\Sigma^{1/2} Q' (\Sigma^{1/2} Q')^+ \right) \Sigma^{-1/2} \quad (\text{Based on Lemma 1}) \\
 &= Q' \left((\Sigma^{1/2} Q')^+ \right) \Sigma^{-1/2} = Q' \left[\left((\Sigma^{1/2})^+ \Sigma^{1/2} Q' \right)^+ \left(\Sigma^{1/2} Q' (Q')^+ \right)^+ \right] \Sigma^{-1/2} \quad (\text{Based on Lemma 2}) \\
 &= Q' \left[(Q')^+ \left(\Sigma^{1/2} Q' (Q')^+ \right)^+ \right] \Sigma^{-1/2} = (Q^+ Q')' \left(\Sigma^{1/2} (Q^+ Q')' \right)^+ \Sigma^{-1/2} \\
 &= \left(I_{KL} - \frac{1}{KL} J_{KL} \right) \left[\Sigma^{1/2} \left(I_{KL} - \frac{1}{KL} J_{KL} \right) \right]^+ \Sigma^{-1/2} \quad (\text{Based on Lemma 4})
 \end{aligned}$$

Therefore, $Q'(Q\Sigma Q')^+ Q$ is invariant for Q and $G = T$. As $T = C$, one obtains $G = C$. In other words, the G statistics is determined by the variance and means of the sample in C . Finally, G is independent of the weights p_{ij} and Q . ■

3. Application Meta-Analysis

In this Section, we empirically verify equality between both statistics G and C from a meta-analysis. The data come from the Stael program base. Specifically, we want to compare the effectiveness of three different molecules and, at the same time, we want to appreciate the impact of administration mode of different molecules (orally or intravenously). However, we don't want to multiply experiments and number of subjects. In total, there are six possible combinations that means 6 series of measures (of different or identical subjects) on which is then measured a relevant quantitative parameter, sensible capture the influence of the decision of the molecules tested). The various combinations of two factors (molecules 3 and 2 modes of treatment) are the factorial design. Here the factor 1 has 3 modes: molecule A, B and C, while the factor 2 admits 2 modalities: Oral and injection.

Table 1 summarizes the distribution of the data used.

Table 2 and **Table 3** report the main statistical characteristics of the both factors.

Table 4 gives the estimation of different parameters and that of the Cochran statistic.

Thus, from the definition of Cochran statistics C :

$$C = \sum_{i=1}^K \sum_{j=1}^L w_{ij} \left(\bar{y}_{ij} - \sum_{n=1}^K \sum_{m=1}^L h_{nm} \bar{y}_{nm} \right)^2 \quad \text{with } K = 3 \text{ and } L = 2. \text{ After calculation, one obtains: } C = 44.5$$

Table 1. Data.

Mol. A	Mol. A	Mo. B	Mol. B	Mol. C	Mol. C
Oral	Injection	Oral	Injection	Oral	Injection
10	11	7	8	12	7
12	18	14	9	9	6
8	12	10	10	11	10
10	15	11	9	10	7
6	13	9	11	7	7
13	8	10	13	8	5
9	15	9	7	13	6
10	16	11	14	14	7
9	9	7	15	10	9
8	13	9	12	11	6

Table 2. Statistics of factor 1.

Factor 1	Mol. A	Mol. B	Mol. C
Nber subjects	20	20	20
Mean	11.25	10.25	8.75
Std deviat.	3.127	2.381	2.552
Median	10.5	10	8.5

Table 3. Statistics of factor 2.

Factor 2	Oral	Injection
Nber subjects	30	30
Mean	9.9	10.27
Std dvt.	2.057	3.503
Median	10	9.5

Table 4. Estimation of main parameters.

	Mol. A	Mol. A	Mol. B	Mol. B	Mol. C	Mol. C
	Oral	Injection	Oral	Injection	Oral	Injection
	10	11	7	8	12	7
	12	18	14	9	9	6
	8	12	10	10	11	10
	10	15	11	9	10	7
	6	13	9	11	7	7
	13	8	10	13	8	5
	9	15	9	7	13	6
	10	16	11	14	14	7
	9	9	7	15	10	9
	8	13	9	12	11	6
Size (n_{ij})	10	10	10	10	10	10
Mean (y_{ij})	9.5	13	9.7	10.8	10.5	7
Std deviation (s_{ij})	2.01	3.13	2.06	2.66	2.17	1.49
Variance (s_{ij}^2)	4.06	9.78	4.23	7.07	4.72	2.22
W_{ij}	2.47	1.02	2.36	1.42	2.12	4.50
h_{ij}(weights)	0.18	0.07	0.17	0.10	0.15	0.32
$y_{ij} * h_{ij}$	1.69	0.96	1.65	1.10	1.60	2.27
Q_{ij}	0.23	3.73	0.43	1.53	1.23	-2.27

Then we determine the G statistics as:

$$Q = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} -0.18 & -0.07 & -0.17 & -0.10 & -0.15 & -0.32 \\ -0.18 & -0.07 & -0.17 & -0.10 & -0.15 & -0.32 \\ -0.18 & -0.07 & -0.17 & -0.10 & -0.15 & -0.32 \\ -0.18 & -0.07 & -0.17 & -0.10 & -0.15 & -0.32 \\ -0.18 & -0.07 & -0.17 & -0.10 & -0.15 & -0.32 \\ -0.18 & -0.07 & -0.17 & -0.10 & -0.15 & -0.32 \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} 0.40 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.98 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.42 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.70 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.47 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.22 \end{pmatrix}$$

$$\Psi = Q'\Sigma Q = \begin{pmatrix} 0.33 & -0.071 & -0.100 & -0.07 & -0.073 & -0.07 \\ -0.07 & 0.913 & -0.100 & -0.07 & -0.069 & -0.07 \\ -0.07 & -0.07 & 0.350 & -0.07 & -0.072 & -0.07 \\ -0.07 & -0.068 & -0.100 & 0.634 & -0.071 & -0.07 \\ -0.07 & -0.069 & -0.100 & -0.07 & 0.400 & -0.07 \\ -0.07 & -0.069 & -0.100 & -0.07 & -0.071 & 0.151 \end{pmatrix}$$

The Moore-Penrose decomposition Ψ^+ of the matrix Ψ in pseudo-inverse is obtained by using Matlab program. Thus we get the singular decomposition matrix that provides a diagonal matrix (with positive values), and matrices I, S such that $\Psi = I * S * V'$. We obtains

$$I = \begin{pmatrix} -0.07 & -0.14 & -0.27 & 0.71 & 0.49 & 0.40 \\ 0.98 & 0.11 & 0.02 & -0.01 & 0.07 & 0.16 \\ -0.07 & -0.15 & -0.42 & -0.70 & 0.41 & 0.38 \\ -0.15 & 0.95 & 0.05 & -0.02 & 0.13 & 0.22 \\ -0.08 & -0.18 & 0.87 & -0.11 & 0.30 & 0.33 \\ -0.05 & -0.08 & -0.06 & 0.03 & -0.69 & 0.71 \end{pmatrix}$$

$$V = \begin{pmatrix} -0.07 & -0.14 & -0.27 & 0.71 & 0.49 & 0.40 \\ 0.98 & 0.11 & 0.02 & -0.01 & 0.07 & 0.16 \\ -0.07 & -0.15 & -0.42 & -0.70 & 0.41 & 0.38 \\ -0.15 & 0.95 & 0.05 & -0.02 & 0.13 & 0.22 \\ -0.08 & -0.18 & 0.87 & -0.11 & 0.30 & 0.33 \\ -0.05 & -0.08 & -0.06 & 0.03 & -0.69 & 0.71 \end{pmatrix};$$

$$S = \begin{pmatrix} 0.94 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.67 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.46 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.41 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.30 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Finally, we have $\Psi^+ = IS^+V'$. To obtain S^+ , we simply reverse the elements on the diagonal excepted those equal to zero. Thus,

$$S^+ = \begin{pmatrix} 1.06 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.50 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2.19 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2.41 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3.36 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The Moore-Penrose pseudo-inverse matrix of Ψ is then given by,

$$\Psi^+ = \begin{pmatrix} 2.22 & -0.01 & -0.22 & -0.05 & -0.16 & -1.03 \\ -0.01 & 1.05 & 0.00 & 0.03 & 0.01 & -0.24 \\ -0.22 & 0.00 & 2.16 & -0.04 & -0.14 & -0.95 \\ -0.05 & 0.03 & -0.04 & 1.45 & -0.01 & -0.41 \\ -0.16 & 0.01 & -0.14 & -0.01 & 2.02 & -0.78 \\ -1.03 & -0.24 & -0.95 & -0.41 & -0.78 & 1.63 \end{pmatrix}$$

Therefore, the G statistics is calculated according to the formula, $G = q\Psi^+q'$ where $q' = (0.23 \ 3.73 \ 0.43 \ 1.53 \ 1.23 \ -2.27)$; we get $G \cong 44.5$.

Finally, we can verify the invariance property of G statistics, compared to h_{ij} weights. It is assumed in this case that the weights are identical in all groups (i, j) , that means that

$$\forall (i, j) \in \{1, 2, 3\} \times \{1, 2\}, h_{ij} = 1/6.$$

We then obtain

$$Q = \begin{pmatrix} 0.83 & -0.17 & -0.17 & -0.17 & -0.17 & -0.17 \\ -0.17 & 0.83 & -0.17 & -0.17 & -0.17 & -0.17 \\ -0.17 & -0.17 & 0.83 & -0.17 & -0.17 & -0.17 \\ -0.17 & -0.17 & -0.17 & 0.83 & -0.17 & -0.17 \\ -0.17 & -0.17 & -0.17 & -0.17 & 0.83 & -0.17 \\ -0.17 & -0.17 & -0.17 & -0.17 & -0.17 & 0.83 \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} 0.40 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.98 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.42 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.70 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.47 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.22 \end{pmatrix}$$

Returning to the procedure described in the previous Section, the following results were obtained, $q = (-0.580 \ 2.92 \ -0.38 \ -0.72 \ 0.42 \ -3.08)$ and

$$\Psi = Q'\Sigma Q = \begin{pmatrix} 0.36 & -0.14 & -0.05 & -0.10 & -0.06 & -0.01 \\ -0.14 & 0.74 & -0.15 & -0.19 & -0.15 & -0.11 \\ -0.05 & -0.15 & 0.37 & -0.10 & -0.06 & -0.02 \\ -0.10 & -0.19 & -0.10 & 0.56 & -0.11 & -0.06 \\ -0.06 & -0.15 & -0.06 & -0.11 & 0.40 & -0.03 \\ -0.01 & -0.11 & -0.02 & -0.06 & -0.03 & 0.24 \end{pmatrix}$$

And the corresponding Moore-Penrose matrix is

$$\Psi^+ = \begin{pmatrix} 2.03 & -0.19 & -0.42 & -0.26 & -0.38 & -0.79 \\ -0.19 & 0.94 & -0.18 & -0.11 & -0.16 & -0.32 \\ -0.42 & -0.18 & 1.96 & -0.25 & -0.36 & -0.76 \\ -0.26 & -0.11 & -0.25 & 1.27 & -0.22 & -0.45 \\ -0.38 & -0.16 & -0.36 & -0.22 & 1.79 & -0.68 \\ -0.79 & -0.32 & -0.76 & -0.45 & -0.68 & 3.06 \end{pmatrix}$$

Once again, we can observe that

$$G = q\Psi^+q' = 44.47 \cong 44.5$$

Interpretation

According to the above results, we observe that $C = G$, and G is invariant whatever the choice of weights is. Finally, the null hypothesis H_0 that assume that all groups (i, j) have the same mean, can be tested based on the fact that $C \sim \chi(5)$. The tabulated statistics at the 5% level is 11.070. As a $C > C_{\text{Tabulated}}$, the null hypothesis H_0 of homogeneity between groups is rejected.

4. Final Remarks

The literature generally uses a multi-step method for determining homogeneity statistics test. It is based on a linear combination of individual mean of the sample to estimate the overall mean. Like the G statistic in (6), this approach involves determining a covariance matrix and its Moore-Penrose inverse. However, we show that **Theorem 1** generalizes the result of [12] in a two ways ANOVA and simplifies this process. We build a G statistic that is equivalent to C . In other words, the expression of C provides a simple formula for determining the statistic in the homogeneity test. Moreover, **Theorem 1** shows that G is asymptotically distributed according to a χ^2 distribution, and it checks certain properties of Cochran statistic. Finally, we also prove that the general form of the G statistic is invariant regardless of the choice of weights.

References

- [1] Hartung, J., Makambi, K. and Argaç, D. (2001) An Extended ANOVA F Test with Applications to the Heterogeneity Problem in Meta-Analysis. *Biometrical Journal*, **43**, 135-146. [http://dx.doi.org/10.1002/1521-4036\(200105\)43:2<135::AID-BIMJ135>3.0.CO;2-H](http://dx.doi.org/10.1002/1521-4036(200105)43:2<135::AID-BIMJ135>3.0.CO;2-H)
- [2] Hartung, J., Argaç, D. and Makambi, K. (2002) Small Sample Properties of Tests on Homogeneity in One-Way ANOVA and Meta-Analysis. *Stat Pap*, **43**, 197-235. <http://dx.doi.org/10.1007/s00362-002-0097-8>
- [3] Hartung, J., Knapp, G. and Sinha, B. (2008) Statistical Meta-Analysis with Applications. Vol. 738, Wiley, New York. <http://dx.doi.org/10.1002/9780470386347>
- [4] Asiribo, O. and Gurland, J. (1990) Coping with Variance Heterogeneity. *Communications in Statistics—Theory and Methods*, **19**, 4029-4048. <http://dx.doi.org/10.1080/03610929008830427>
- [5] De Beuckelaer, A. (1996) A Closer Examination on Some Parametric Alternatives to the ANOVA F Test. *Stat Pap*, **37**, 291-305. <http://dx.doi.org/10.1007/BF02926110>
- [6] Cochran, W. (1937) Problems Arising in the Analysis of a Series of Similar Experiments. *Journal of the Royal Statistical Society*, **4**, 102-118. <http://dx.doi.org/10.2307/2984123>
- [7] Der Simonian, R. and Laird, N. (1986) Meta-Analysis in Clinical Trials. *Controlled Clinical Trials*, **7**, 177-188. [http://dx.doi.org/10.1016/0197-2456\(86\)90046-2](http://dx.doi.org/10.1016/0197-2456(86)90046-2)
- [8] Biggerstaff, B. and Jackson, D. (2008) The Exact Distribution of Cochran's Heterogeneity Statistic in One-Way Random Effects Meta-Analysis. *Statistics in Medicine*, **27**, 6093-6110. <http://dx.doi.org/10.1002/sim.3428>
- [9] James, G. (1951) The Comparison of Several Groups of Observations When the Ratios of the Population Variances Are Unknown. *Biometrika*, **38**, 324-329. <http://dx.doi.org/10.1093/biomet/38.3-4.324>
- [10] Kulinskaya, E., Morgenthaler, S. and Staudte, R. (2008) Meta-Analysis: A Guide to Calibrating and Combining Statistical Evidence. Vol. 757, Wiley, New York.
- [11] Welch, B. (1951) On the Comparison of Several Mean Values: An Alternative Approach. *Biometrika*, **38**, 330-336. <http://dx.doi.org/10.1093/biomet/38.3-4.330>
- [12] Chen, Z., Ng, H.T. and Nadarajah, S. (2014) A Note on Cochran Test Homogeneity in One-Way ANOVA and Meta-Analysis. *Stat Pap*, **55**, 301-310. <http://dx.doi.org/10.1007/s00362-012-0475-9>
- [13] Schott, J. (1997) Matrix Analysis for Statistics. Wiley, New York.
- [14] Seber, G. (2008) A Matrix Handbook for Statisticians. Vol. 746, Wiley, New York.