In Neil Salkind (Ed.), *Encyclopedia of Research Design.*
Thousand Oaks, CA: Sage. 2010

The Greenhouse-Geisser Correction

Hervé Abdi

1 Overview and background

When performing an analysis of variance with a one factor repeated measurement design, the effect of the independent variable is tested by computing an F statistic which is computed as the ratio of the of mean square of effect by the mean square of the interaction between the subject factor and the independent variable. For a design with S subjects and A experimental treatments, when some assumptions are met, the sampling distribution of this F ratio is a Fisher distribution with $\nu_1 = A - 1$ and $\nu_2 = (A-1)(S-1)$ degrees of freedom.

In addition to the usual assumptions of normality of the error and homogeneity of variance, the F test for repeated measurement designs assumes a condition called *"sphericity."* (Huynh & Feldt, 1970; Rouanet & Lépine, 1970). Intuitively, this condition indicates that the ranking of the subjects does not change across experimental treatment. This is equivalent to stating that the population correlation (computed form the subjects' scores) between two treatments

Hervé Abdi
The University of Texas at Dallas
Address correspondence to:
Hervé Abdi
Program in Cognition and Neurosciences, MS: Gr.4.1,
The University of Texas at Dallas,
Richardson, TX 75083–0688, USA
E-mail: herve@utdallas.edu http://www.utd.edu/∼herve

Table 1: A data set for a repeated measurement design.

	a_1	a_2	a_3	a_4		$M_{.s}$
S_1	76	64	34	26		50
S_2	60	48	46	30		46
S_3	58	34	32	28		38
S_4	46	46	32	28		38
S_5	30	18	36	28		28
$M_{a.}$	54	42	36	28	$M_{..} =$	40

is the same for all pairs of treatments. This condition implies that there is no interaction between the subject factor and the treatment.

If the sphericity assumption is not valid, then the F test becomes too liberal (*i.e.,* the proportion of rejections of the null hypothesis is larger than the α level when the null hypothesis is true). In order to minimize this problem, Greenhouse and Geisser (1959) elaborating on early work by Box (1954) suggested to use an index of deviation to sphericity to correct the number of degrees of freedom of the F distribution. We first present this index of non sphericity (called the Box index, denoted ε), then we present its estimation and its application known as the Greenhouse-Geisser correction. We also present the Huyhn-Feldt correction which is a more efficient procedure. Finally, we explore tests for sphericity.

2 An index of sphericity

Box (1954a & b) has suggested a measure for sphericity denoted ε which varies between 0 and 1 and reaches the value of 1 when the data are perfectly spherical. We will illustrate the computation of this index with the fictitious example given in Table 1 where we collected the data from $S = 5$ subjects whose responses were measured for $A = 4$ different treatments. The standard analysis of variance of these data gives a value of $F_A = \frac{600}{112} = 5.36$, which, with $\nu_1 = 3$ and $\nu_2 = 12$, has a p value of .018.

Table 2: The covariance matrix for the data set of Table 1.

	a_1	a_2	a_3	a_4		
a_1	294	258	8	−8		
a_2	258	294	8	−8		
a_3	8	8	34	6		
a_4	−8	−8	6	2		
$\bar{t}_{a.}$	138	138	14	−2	$\bar{t}_{..} =$	72
$\bar{t}_{a.} - \bar{t}_{..}$	66	66	−58	−74		

In order to evaluate the degree of sphericity (or lack thereof), the first step is to create a table called a *covariance matrix*. This matrix comprises the variances of all treatments and all the covariances between treatments. As an illustration, the covariance matrix for our example is is given in Table 2.

Box (1954) defined an index of sphericity, denoted ε, which applies to a *population covariance* matrix. If we call $\zeta_{a,a'}$ the entries of this $A \times A$ table, the Box index of sphericity is obtained as

$$\varepsilon = \frac{\left(\sum_a \zeta_{a,a}\right)^2}{(A-1)\sum_{a,a'} \zeta^2_{a,a'}} . \tag{1}$$

Box also showed that when sphericity fails, the number of degrees of freedom of the F_A ratio depends directly upon the degree of sphericity (*i.e.,* ε) and are equal to $\nu_1 = \varepsilon(A-1)$ and $\nu_1 = \varepsilon(A-1)(S-1)$.

2.1 Greenhouse-Geisser correction

Box's approach works for the *population* covariance matrix, but, unfortunately, in general this matrix is not known. In order to estimate ε we need to transform the sample covariance matrix into an *estimate* of the population covariance matrix. In order to compute this estimate, we denote by $t_{a,a'}$ the sample estimate of the covariance between groups a and a' (these values are given in Table 2), by $\bar{t}_{a.}$ the mean of the covariances for group a and by $\bar{t}_{..}$ the grand mean of the covariance table. The estimation of the population covariance matrix will have for general term $s_{a,a'}$ which is computed as

$$s_{a,a'} = (t_{a,a'} - \bar{t}_{..}) - (\bar{t}_{a,.} - \bar{t}_{..}) - (\bar{t}_{a',.} - \bar{t}_{..}) = t_{a,a'} - \bar{t}_{a,.} - \bar{t}_{a',.} + \bar{t}_{..} \ . \quad (2)$$

(this procedure is called "double-centering").

Table 3 gives the double centered covariance matrix. From this matrix, we can compute the estimate of ε which is denoted $\widehat{\varepsilon}$ (compare with Equation 1):

$$\widehat{\varepsilon} = \frac{\left(\sum_{a} s_{a,a}\right)^2}{(A-1)\sum_{a,a'} s_{a,a'}^2} \ . \quad (3)$$

In our example, this formula gives:

$$\widehat{\varepsilon} = \frac{(90+90+78+78)^2}{(4-1)\left(90^2+54^2+\cdots+66^2+78^2\right)^2} = \frac{336^2}{3 \times 84,384} = \frac{112,896}{253,152}$$
$$= .4460 \ . \quad (4)$$

We use the value of $\widehat{\varepsilon} = .4460$ to correct the number of degrees of freedom of F_A as $\nu_1 = \widehat{\varepsilon}(A-1) = 1.34$ and $\nu_2 = \widehat{\varepsilon}(A-1)(S-1) = 5.35$. These corrected values of ν_1 and ν_2 give for $F_A = 5.36$ a probability of $p = .059$. If we want to use the critical value approach, we need to round the values of these corrected degree of freedom to the nearest integer (which will give here the values of $\nu_1 = 1$ and $\nu_2 = 5$).

2.2 Greenhouse-Geisser Correction and eigenvalues

The Box index of sphericity is best understood in relation to the eigenvalues (see, e.g., Abdi, 2007 for an introduction) of a covariance matrix. Recall that covariance matrices belong to the class of

Table 3: The double centered covariance matrix used to estimate the population covariance matrix.

	a_1	a_2	a_3	a_4
a_1	90	54	−72	−72
a_2	54	90	−72	−72
a_3	−72	−72	78	66
a_4	−72	−72	66	78

positive semi-definite matrices and therefore always has positive of null eigenvalues. Specifically, if we denote by $\mathbf{\Sigma}$ a population covariance, and by λ_ℓ the ℓ-th eigenvalue of $\mathbf{\Sigma}$, The sphericity condition is equivalent to having all eigenvalues equal to a constant. Formally, the shericity condition states that:

$$\lambda_\ell = \text{constant} \quad \forall \ell \ . \tag{5}$$

In addition, if we denote by V (also called β, see Abdi 2007; or ν, see Worsley & Friston, 1995) the following index:

$$V = \frac{\left(\sum \lambda_\ell\right)^2}{\sum \lambda_\ell^2} \ . \tag{6}$$

then the Box coefficient can be expressed as

$$\varepsilon = \frac{1}{A-1} V \ , \tag{7}$$

Under sphericity, all the eigenvalues are equal and V is equal to $(A-1)$. The estimate of ε is obtained by using the eigenvalues of the estimated covariance matrix. For example, the matrix from Table 3, has the following eigenvalues:

$$\lambda_1 = 288, \quad \lambda_2 = 36, \quad \lambda_3 = 12 \ . \tag{8}$$

This gives:

$$V = \frac{\left(\sum \lambda_\ell\right)^2}{\sum \lambda_\ell^2} = \frac{(288+36+12)^2}{288^2+36^2+12^2} \approx 1.3379 \ , \tag{9}$$

which, in turn, gives

$$\widehat{\varepsilon} = \frac{1}{A-1} V = \frac{1.3379}{3} \approx .4460 \tag{10}$$

(this matches the results of Equation 4).

2.3 Extreme Greenhouse-Geisser correction

A *conservative* (*i.e.,* increasing the risk of Type II error: the probability of not rejecting the null hypothesis when it is false) correction for sphericity has been suggested by Greenhouse and Geisser (1959). Their idea is to choose the largest possible value of $\widehat{\varepsilon}$, which is equal to $A-1$. This leads to consider that F_A follows a Fisher distribution with $\nu_1 = 1$ and $\nu_2 = S-1$ degrees of freedom. In this case, these corrected values of $\nu_1 = 1$ and $\nu_2 = 4$ give for $F_A = 5.36$ a probability of $p = .081$.

2.4 Huynh-Feldt correction

Huynh and Feldt (1976) suggested a better (more powerful) approximation for ε denoted $\widetilde{\varepsilon}$ and computed as

$$\widetilde{\varepsilon} = \frac{S(A-1)\widehat{\varepsilon} - 2}{(A-1)\left[S - 1 - (A-1)\widehat{\varepsilon}\right]} \ . \tag{11}$$

In our example, this formula gives:

$$\widetilde{\varepsilon} = \frac{5(4-1).4460 - 2}{(4-1)\left[5 - 1 - (4-1).4460\right]} = .5872 \ .$$

We use the value of $\widetilde{\varepsilon} = .5872$ to correct the number of degrees of freedom of F_A as $\nu_1 = \widetilde{\varepsilon}(A-1) = 1.76$ and $\nu_2 = \widetilde{\varepsilon}(A-1)(S-1) = 7.04$. These corrected values of give for $F_A = 5.36$ a probability of $p = .041$. If we want to use the critical value approach, we need to round these corrected values for the number of degree of freedom to the nearest integer (which will give here the values of $\nu_1 = 2$

and $\nu_2 = 7$). In general, the correction of Huynh and Feldt is to be preferred because it is more powerful (and Greenhouse-Geisser is too conservative).

2.5 Stepwise Strategy for sphericity

Greenhouse and Geisser (1959) suggest to use a stepwise strategy for the implementation of the correction for lack of sphericity. If F_A is not significant with the standard degrees of freedom. there is no need to implement a correction (because it will make it even less significant). If F_A is significant with the extreme correction [*i.e.*, with $\nu_1 = 1$ and $\nu_2 = (S-1)$], then there is no need to correct either (because the correction will make it more significant). If F_A is not significant with the extreme correction but is not significant with the standard number of degree of freedom, then use the ε correction (they recommend using $\widehat{\varepsilon}$, but the subsequent $\widetilde{\varepsilon}$ is a better estimate and should be preferred).

3 Testing for sphericity

One incidental question about using a correction for lack of sphericity is to decide when a sample covariance matrix is *not* spherical. There are several tests that can be used to answer this question. The most well known is Mauchly's test, the most powerful is John, Nagao and Suguiara's test.

3.1 Mauchly's test for sphericity

Mauchly (1940) constructed a test for sphericity based on the following statistics which uses the eigenvalues of the estimated covariance matrix

$$W = \frac{\prod \lambda_\ell}{\left[\frac{1}{A-1}\sum \lambda_\ell\right]^{A-1)}} \ . \tag{12}$$

This statistics varies between 0 and 1 and reaches 1 when the matrix is spherical. For our example, we find that

$$W = \frac{\prod \lambda_\ell}{\left[\frac{1}{A-1}\sum \lambda_\ell\right]^{A-1)}} = \frac{228 \times 36 \times 12}{\left[\frac{1}{3}\left(228+36+12\right)\right]^3} = \frac{124,416}{1,404,928} \approx .0886. . \tag{13}$$

Tables for the critical values of W are available in Nagarsenker and Pillai (1973), but a good approximation (Pillai & Nagarsenker, 1971) is obtained by transforming W into

$$X^2_W = -(1-f) \times (S-1) \times \ln\{W\} \tag{14}$$

where

$$f = \frac{2(A-1)^2 + A + 2}{6(A-1)(S-1)} . \tag{15}$$

Under the null hypothesis of sphericity, X^2_W is approximately distributed as a χ^2 with degrees of freedom equal to

$$\nu = \frac{1}{2}A(A-1) . \tag{16}$$

For our example, we find that

$$f = \frac{2(A-1)^2 + A + 2}{6(A-1)(S-1)} = 2 \times 3^2 + 4 + 26 \times 3 \times 4 = \frac{24}{72} = .33 , \tag{17}$$

and

$$X^2_W = -(1-f) \times (S-1) \times \ln\{W\} = -4(1-.33) \times \ln\{.0886\} \approx 6.46 . \tag{18}$$

with $\nu = \frac{1}{2}4 \times 3 = 6$, we find that $p = .38$ and we cannot reject the null hypothesis. Despite its relative popularity, the Mauchly test is not recommended because it lacks power (see Boik, 1981; Cornell, Young, Seaman, & Kirk, 1992; Keselman, Algina, & Kowalchuk, 2001). A more powerful alternatice is the John, Suguira & Nagao test for sphericity described below.

3.2 John, Nagao & Sugiura's test for sphericity

According to Cornell, Young, Seaman, and Kirk, (1992) the best test for sphericity uses V (John, 1972; Nagao, 1973; Suguira, 1972).

Tables for the critical values of W are available in Grieve (1984), but a good approximation (Suguira, 1972) is obtained by transforming V into

$$X_V^2 = \frac{1}{2}S(A-1)^2\left(V - \frac{1}{A-1}\right) . \tag{19}$$

Under the null hypothesis, X_V^2 is approximately distributed as a χ^2 distribution with $\nu = \frac{1}{2}A(A-1) - 1$. For our example, we find that

$$X_V^2 = \frac{1}{2}S(A-1)^2\left(V - \frac{1}{A-1}\right) = \frac{5 \times 3^2}{2}\left(1.3379 - \frac{1}{3}\right) = 22.60 . \tag{20}$$

with $\nu = \frac{1}{2}4 \times 3 - 1 = 5$, we find that $p = .004$ and we can reject the null hypothesis with the usual test. The discrepancy between the conclusions reached from the two tests for sphericity illustrates the lack of power of Mauchly's test.

References

1. Abdi, H. (2007). The R_v coefficient and the congruence coefficient. In N.J. Salkind (Ed.): *Encyclopedia of Measurement and Statistics*. Thousand Oaks (CA): Sage. pp. 849–853.
2. Boik, R.J. (1981). A priori tests in repeated measures designs: Effects of non-sphericity. *Psychometrika*, **46**, 241–255.
3. Box, G.E.P. (1954). Some theorems on quadrative forms applied to the study of analysis of variance problem; I: Effect of inequality of variance in the one way-classification. *Annals of Mathematical Statistics*, **25**, 290–302.
4. Box, G.E.P. (1954). Some theorems on quadrative forms applied to the study of analysis of variance problem; II: Effect of inequality of variance and of correlation in the a way-classification. *Annals of Mathematical Statistics*, **25**, 484–498.
5. Collier, R.O., Baker, F.B., Mandeville, & G.K., Hayes, H.T. (1967). Estimates of test size for several test procedures based on variance ratios in the repeated measures design. *Psychometrika*, **32**, 339–353.
6. Cornell, J.E., Young, D.M., Seaman, S.L., Kirk, R.E. (1992). Power comparisons of eight tests for sphericity in repeated measures designs *Journal of Educational Statistics*, **17**, 233–249.
7. Geisser, S., & Greenhouse, S.W. (1958). An extension of Box's result on the use of F distribution in multivariate analysis. *Annals of Mathematical Statistics*, **29**, 885–891.
8. Greenhouse, S.W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, **24**, 95–112.
9. Grieve, A.P (1984). Tests of sphericity of normal distribution and the analysis of repeated measure designs. *Psychometrika*, **49**, 257–267.
10. Huynh, H., & Feldt, L.S. (1970). Conditions under which mean square ratios in repeated measurement designs have exact F–distributions. *Journal of the American Statistical Association*, **65**, 1582–1589.

11. Huynh, H., & Feldt, L.S. (1976) Estimation of the Box correction for degrees of freedom from sample data in randomized block and split=plot designs. *Journal of Educational Statistics*, **1**, 69–82.

12. Huynh, H., & Mandeville, G. (1979). Validity conditions in repeated measures designs. *Psychological Bulletin*, **86**, 964–973.

13. John, S. (1972). The distribution of a statistic used for testing sphericity of normal distributions. *Biometrika*, **59**, 169-173.

14. Keselman, H.J., Algina, J., & Kowalchuk, R.K. (2001). The analysis of repeated measures designs: A review. *British Journal of Mathematical & Statistical Psychology*, **54**, 1–20.

15. Mauchly, J.W. (1940). Significance test for sphericity of n-variate normal population. *Annals of Mathematical Statistics*, **11**, 204-209.

16. Nagao, H. (1973). On some test criteria for covariance matrix. *The Annals of Statistics*, **1**, 700–709.

17. Rouanet, H., & Lépine, D. (1970). Comparison between treatments in a repeated–measures design: ANOVA and multivariate methods. *British Journal of Psychology*, **8**, 166–84.

18. Suguira, N. (1972). Locally best invariant test for sphericity anr the limiting distribution. *Annals of Mathematical Statistics*, **43**, 1312–1326.

19. Worsley K.J., & Friston, K.J (1995). Analysis of *f*MRI Time-series revisited—Again. *NeuroImage*, **2**, 173-181.