GLM
Repeated Measures

The GLM (general linear model) procedure provides analysis of variance when the same measurement or measurements are made several times on each subject or case (repeated measures). Algorithms for GLM not discussed in this chapter are in “GLM Univariate and Multivariate.” Distribution functions are discussed in Appendix 12.

Notation

The notation listed in the GLM chapter is used in this chapter. Additional notation conventions are defined below:

t	The number of within-subjects factors.
c	The number of measures.
r_k	The number of levels of the *k*th within-subjects factor. $r_k \geq 2, k = 1, \ldots, t$.
$\mathbf{M}_k$	The contrast matrix of the *k*th within-subjects factor, $k = 1, \ldots, t$. It is a square matrix with dimension r_k. Each element in the first column is usually equal to $1/r_k$. For a polynomial contrast each element is $1/\sqrt{r_k}$, or, for a user-specified contrast, a non-zero constant The other columns have zero column sums.

Number of Variables

It is required that $c \times \prod_{k=1}^{t} r_k = r$, the number of dependent variables in the model.

Covariance Structure

As usual in GLM, the data matrix is related to the parameter matrix $\mathbf{B}$ as $\mathbf{Y} = \mathbf{XB} + \mathbf{E}$. The rows of $\mathbf{E}$ are uncorrelated and the *i*th row has the distribution $N_r(0, w_i^{-1}\Sigma)$.

Repeated measures analysis has two additional assumptions:

- $\Sigma = \Sigma_C \otimes \Sigma_1 \otimes \cdots \otimes \Sigma_t$ where Σ_C is the covariance matrix of the measures and $\otimes$ is the Kronecker product operator.

- The Huynh and Feldt (1970) condition: Suppose $\sigma_{rs}^{(k)}$ is the (r,s)-th element of Σ_k $(k=1,\ldots,t)$; then $\sigma_{rr}^{(k)} + \sigma_{ss}^{(k)} - 2\sigma_{rs}^{(k)} =$ constant for $r \neq s$. Matrices satisfying this condition result in orthonormally transformed variables with spherical covariance matrices; for this reason, the assumption is sometimes referred to as the sphericity assumption. A matrix that has the property of compound symmetry (that is, identical diagonal elements and identical off-diagonal elements) automatically satisfies this assumption.

Tests on the Between-Subjects Effects

Procedure

The procedure for testing the hypothesis of no between-subjects effects uses the following steps:

1. Compute $\mathbf{M} = \mathbf{I}_c \otimes \mathbf{M}_{1;1} \otimes \cdots \otimes \mathbf{M}_{t;1}$ where $\mathbf{M}_{k;1}$ is the first column of the contrast matrix $\mathbf{M}_k$ of the *k*th within-subjects factors. Note that $\mathbf{M}$ is an $r \times c$ matrix.
2. For each of the between-subjects effects including the intercept, get the **L** matrix, according to the specified type of sum of squares.
3. Compute $\mathbf{S}_H = (\mathbf{L}\hat{\mathbf{B}}\mathbf{M})'(\mathbf{L}\mathbf{G}\mathbf{L}')(\mathbf{L}\hat{\mathbf{B}}\mathbf{M})$ and $\mathbf{S}_E = \mathbf{M}'\mathbf{S}\,\mathbf{M}$. Both are $c \times c$ matrices.
4. Compute the four multivariate test statistics: Wilks' lambda, Pillai's trace, Hotelling-Lawley trace, Roy's largest root, and the corresponding significance levels. Also compute the individual univariate *F* statistics.
5. Repeat steps 2 to 4 until all between-subjects effects have been tested.

Multivariate Tests on the Within-Subjects Effects

Procedure

The procedure for testing the hypothesis of no within-subjects effects uses the following steps:

1. For the *k*th within-subjects factor, compute $\mathbf{M} = \mathbf{I}_c \otimes \mathbf{A}_1 \otimes \cdots \otimes \mathbf{A}_t$ where $\mathbf{A}_k = \mathbf{M}_{k;2:r_k}$ which is the second-to-last column of $\mathbf{M}_k$ when the *k*th within-subjects factor is involved in the effect. Otherwise, $\mathbf{A}_k = \mathbf{M}_{k;1}$. Note that **M** is an $r \times cd$ matrix, where *d* is the number of columns in the Kronecker product $\mathbf{A_1} \otimes \ldots \otimes \mathbf{A_t}$. In general, $d > 1$.
2. For each of the between-subjects effects, get the **L** matrix, according to the specified type of sum of squares.
3. Compute $\mathbf{S}_H = (\mathbf{L\hat{B}M})'(\mathbf{LGL}')(\mathbf{L\hat{B}M})$ and $\mathbf{S}_E = \mathbf{M'SM}$. Both are $cd \times cd$ matrices.
4. Compute the four multivariate test statistics: Wilks' lambda, Pillai's trace, Hotelling-Lawley trace, Roy's largest root, and the corresponding significance levels. Also compute the individual univariate *F* statistics.
5. Repeat steps 2 to 4 for the next between-subjects effect. When all the between-subjects effects are used, go to step 6.
6. Repeat steps 1 to 5 until all within-subjects effects have been tested.

Averaged Tests on the Within-Subjects Effects

Procedure

The procedure for the averaged test of the hypothesis of no within-subjects effects uses the following steps:

1. Take $\mathbf{M}_k$ ($k = 1, \ldots, t$) the equally spaced polynomial contrast matrix.
2. Compute $\mathbf{M} = \mathbf{I}_c \otimes \mathbf{A}_1 \otimes \cdots \otimes \mathbf{A}_t$ where $\mathbf{A}_k = \mathbf{M}_{k;2:r_k}$ which is the 2nd to last column of $\mathbf{M}_k$ when the *k*th within-subjects factor is involved in the effect. Otherwise, $\mathbf{A}_k = \mathbf{1}_{r_k} \big/ \sqrt{r_k}$. Note that **M** is an $r \times cd$ matrix, where *d* is the number of columns in the Kronecker product $\mathbf{A_1} \otimes \ldots \otimes \mathbf{A_t}$. In general, $d > 1$.
3. For each of the between-subjects effects, get the **L** matrix, according to the specified type of sum of squares.
4. Compute $\mathbf{S}_H = (\mathbf{L\hat{B}M})'(\mathbf{LGL}')(\mathbf{L\hat{B}M})$ and $\mathbf{S}_E = \mathbf{M'SM}$. Both are $cd \times cd$ matrices.
5. Partition $\mathbf{S}_H$ into c^2 block matrices each of dimension $d \times d$. The (*k*,*l*)th block, denoted as $\mathbf{S}_{H;k,l}$, ($k = 1, \ldots, c$ and $l = 1, \ldots, c$), is a sub-matrix of $\mathbf{S}_H$ from row $(k-1)d+1$ to row *kd*, and from column $(l-1)d+1$ to column *ld*. Form the

$c \times c$ matrix, denoted by $\overline{\mathbf{S}}_H$, whose (k, l)th element is the trace of $\mathbf{S}_{H;k,l}$. The matrix $\overline{\mathbf{S}}_E$ is obtained similarly.

6. Use $\overline{\mathbf{S}}_H$ and $\overline{\mathbf{S}}_E$ for computing the four multivariate test statistics: Wilks' lambda, Pillai's trace, Hotelling-Lawley trace, Roy's largest root, and the corresponding significance levels. *Note:* Set the degrees of freedom for $\overline{\mathbf{S}}_H$ (same as the row dimension of $\mathbf{L}$ in the test procedure) equal to dr_{L} and that for $\overline{\mathbf{S}}_E$ equal to $d(n - r_X)$ in the computations. Also compute the individual univariate *F* statistics and their significance levels.
7. Repeat steps 3 to 6 for each between-subjects effect. When all the between-subjects effects are used, go to step 8.
8. Repeat steps 2 to 7 until all within-subjects effects have been tested.

Adjustments to Degrees of Freedom of the *F* Statistics

The adjustments to degrees of freedom of the univariate *F* test statistics are the Greenhouse-Geisser epsilon, the Huynh-Feldt epsilon, and the lower-bound epsilon.

Greenhouse-Geisser epsilon

$$\varepsilon_{GG} = \frac{\left(\mathrm{trace}(\mathbf{S}_E)\right)^2}{d \times \mathrm{trace}(\mathbf{S}_E^2)}$$

Huynh-Feldt epsilon

$$\varepsilon_{HF} = \min\left(\frac{nd\varepsilon_{GG} - 2}{d(n - r_{\mathbf{X}}) - d^2\varepsilon_{GG}}, 1\right)$$

Lower bound epsilon

$$\varepsilon_{LB} = 1/d$$

For any of the above three epsilons, the adjusted significance level is

$$1 - \mathtt{CDF.F}\left(F, \varepsilon dr_{\mathrm{L}}, \varepsilon d(n - r_{\mathbf{X}})\right)$$

where ε is one of the above three epsilons.

Mauchly's Test of Sphericity

Mauchly's test of sphericity is displayed for every repeated measures model.

Hypotheses

In Mauchly's test of sphericity the null hypothesis is $H_o\!:\mathbf{M}'\Sigma\mathbf{M} = \sigma^2\mathbf{I}_m$, versus the alternative hypothesis $H_1\!:\mathbf{M}'\Sigma\mathbf{M} \neq \sigma^2\mathbf{I}_m$, where $\sigma^2 > 0$ is unspecified, $\mathbf{I}$ is an $m \times m$ identity matrix, and $\mathbf{M}$ is the $r \times m$ orthonormal matrix associated with a within-subjects effect. $\mathbf{M}$ is generated using equally spaced polynomial contrasts applied to the within-subjects factors (see the descriptions in the section "Averaged Tests on the Within-Subjects Effects" on p. 3).

Mauchly's *W* Statistic

$$W = \begin{cases} \dfrac{|\Xi|}{\left(\mathrm{trace}(\Xi)/m\right)^m} & \text{if } \mathrm{trace}(\Xi) > 0 \\ \text{SYSMIS} & \text{if } \mathrm{trace}(\Xi) \le 0 \end{cases}$$

where $\Xi = \mathbf{M}'\mathbf{A}\mathbf{M}$ and $\mathbf{A} = \left(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}\right)'\mathbf{W}\left(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}\right)$ is the $r \times r$ matrix of residual sums of squares and cross products.

Chi-Square Approximation

When *n* is large and under the null hypothesis that for $n - r_X \ge 1$ and $m \ge 2$,

$$\Pr\left(-\rho\left(n - r_X\right)\log W \le c\right) = \Pr\left(\chi_f^2 \le c\right) + \omega_2\left(\Pr\left(\chi_{f+4}^2 \le c\right) - \Pr\left(\chi_f^2 \le c\right)\right) + O\left(n^{-3}\right)$$

where

$$f = m(m+1)/2 - 1$$

$$\rho = 1 - \left(2m^2 + m + 2\right) \Big/ \left(6m\left(n - r_X\right)\right)$$

$$\omega_2 = \frac{(m+2)(m-1)(m-2)\left(2m^3 + 6m^2 + 3m + 2\right)}{288m^2\left(n - r_X\right)^2 \rho^2}$$

Chi-Square Statistic

$$c = \begin{cases} -\rho\left(n - r_X\right)\log W & \text{if } W > 0 \\ \text{SYSMIS} & \text{otherwise} \end{cases}$$

Degrees of Freedom

$$f = m(m+1)/2 - 1$$

Significance

$$1 - \text{CDF.CHISQ}(c, f) - \omega_2\left(\text{CDF.CHISQ}(c, f+4) - \text{CDF.CHISQ}(c, f)\right)$$

where CDF.CHISQ is the SPSS function for cumulative chi-square distribution. The significance will be reset to zero in case the computed value is less than zero due to floating point imprecision.

References

Greenhouse, S. W., and Geisser, S. 1959. On methods in the analysis of profile data. *Psychometrika*, 24(2): 95–111.

Huynh, H., and Feldt, L. S. 1970. Conditions under which mean square ratios in repeated measurements designs have exact *F*-distributions. *Journal of the American Statistical Association*, 65: 1582–1589.

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

Mauchly, John W. 1940. Significance test for sphericity of a normal *n*-variate distribution. *Annals of Mathematical Statistics*, 11: 204–209.

Searle, S. R. 1982. *Matrix algebra useful for statistics*. New Work: John Wiley & Sons, Inc.