Response Profile Analysis

- Are the mean response profiles similar in the groups, or in other words, are the mean response profiles parallel?
  This is a question that concerns the group × time **interaction effect**
- Assuming that the mean response profiles are parallel, are the means constant over time?
  This is a question that concerns the **time effect.**
- Assuming that the mean response profiles are parallel, are they at the same level?
  This is a question that concerns the **group effect.**

Generalized Least Squares (GLS)

GLS is an extension of ordinary least squares (OLS) since it seeks to minimize a weighted sum of squared residuals. Unlike OLS, GLS accommodates heterogeneity and correlation via the weights which correspond to the inverse of the variance-covariance matrix. The parameters are estimated by minimizing the GLS objective function:

$$Q_{GLS}(\boldsymbol{\beta}|\boldsymbol{\theta}) = \sum_{i=1}^{n} (\mathsf{Y}_i - \mathsf{X}_i\boldsymbol{\beta})' \, \boldsymbol{\Sigma}_i(\boldsymbol{\theta})^{-1} \, (\mathsf{Y}_i - \mathsf{X}_i\boldsymbol{\beta})$$

Theta is a vector of variance-covariance parameters. and $\boldsymbol{\Sigma}_i(\boldsymbol{\theta}) = \mathrm{Var}(\mathsf{Y}_i),$
We distinguish 2 cases, where theta is known and theta is not known.

Theta is Known

From calculus we know that in order to minimize the objective function we
1. Differentiate with respect to beta
2. Set the result to 0
3. Solve for beta

The solution to the minimization of the problem is:

$$\hat{\boldsymbol{\beta}}_{GLS} = \left(\sum_{i=1}^{n} \mathsf{X}_i' \boldsymbol{\Sigma}_i(\boldsymbol{\theta})^{-1} \mathsf{X}_i \right)^{-1} \sum_{i=1}^{n} \mathsf{X}_i' \boldsymbol{\Sigma}_i(\boldsymbol{\theta})^{-1} \mathsf{Y}_i$$

with

$$\mathrm{Var}(\hat{\boldsymbol{\beta}}_{GLS}) = \left(\sum_{i=1}^{n} \mathsf{X}_i' \boldsymbol{\Sigma}_i(\boldsymbol{\theta})^{-1} \mathsf{X}_i \right)^{-1}$$

Since assuming we know the variance-covariance matrix of Y is a rather strong assumption we should protect ourselves against model mis-specification. We can use for inference the empirical "sandwich" estimator which is robust:

$$\text{Var}^S(\hat{\boldsymbol{\beta}}_{GLS}) = \text{Var}(\hat{\boldsymbol{\beta}}_{GLS}) \left(\sum_{i=1}^{n} \mathsf{X}_i' \Sigma_i(\boldsymbol{\theta})^{-1} \hat{\mathsf{e}}_i \hat{\mathsf{e}}_i' \Sigma_i(\boldsymbol{\theta})^{-1} \mathsf{X}_i \right) \text{Var}(\hat{\boldsymbol{\beta}}_{GLS})$$

where $\hat{\mathsf{e}}_i = \mathsf{Y}_i - \mathsf{X}_i \hat{\boldsymbol{\beta}}_{GLS}$

Theta is Unknown

When theta is unknown we need to replace it in the GLS formulas with a consistent estimate theta_hat. Typically *theta_hat = theta_hat*beta_hat_0,* where *beta_hat_0* is an initial unbiased estimate of beta, such as the OLS estimate and *theta_hat*beta_hat_0* is a non-iterative method of moments (MM) type estimator that is consistent for theta. Then the estimated generalized least squares estimator (ELGS) is:

$$\hat{\boldsymbol{\beta}}_{EGLS} = \left(\sum_{i=1}^{n} \mathsf{X}_i' \Sigma_i(\hat{\boldsymbol{\theta}})^{-1} \mathsf{X}_i \right)^{-1} \sum_{i=1}^{n} \mathsf{X}_i' \Sigma_i(\hat{\boldsymbol{\theta}})^{-1} \mathsf{Y}_i$$

with

$$\text{Var}(\hat{\boldsymbol{\beta}}_{EGLS}) = \left(\sum_{i=1}^{n} \mathsf{X}_i' \Sigma_i(\hat{\boldsymbol{\theta}})^{-1} \mathsf{X}_i \right)^{-1}$$

Maximum Likelihood

The most common approach to estimation is the method of maximum likelihood. The idea is to use an estimate of beta with the values that are most likely. Given that Y is assumed to have a conditional distribution that is multivariate normal we must maximize the following log-likelihood function:

$$\begin{aligned} \ell(\boldsymbol{\beta}, \boldsymbol{\theta}; \mathsf{Y}) &= -\frac{np}{2} \log 2\pi - \frac{1}{2} \sum_{i=1}^{n} \log |\Sigma_i(\boldsymbol{\theta})| \\ &\quad -\frac{1}{2} \left\{ \sum_{i=1}^{n} (\mathsf{Y}_i - \mathsf{X}_i \boldsymbol{\beta})' \Sigma_i(\boldsymbol{\theta})^{-1} (\mathsf{Y}_i - \mathsf{X}_i \boldsymbol{\beta}) \right\} \\ &= -\frac{np}{2} \log 2\pi - \frac{1}{2} \sum_{i=1}^{n} \log |\Sigma_i(\boldsymbol{\theta})| - \frac{1}{2} Q_{GLS}(\boldsymbol{\beta} | \boldsymbol{\theta}) \end{aligned}$$

Maximum Likelihood Estimators (MLE)

When theta is known and fixed the ML and GLS estimators are equivalent. When theta is unknown ML and EGLS estimators are equivalent if theta_hat (in EGLS) is equal to the ML estimate of theta. The MLE is estimated iteratively (as opposed to non-iteratively by EGLS) by maximizing the profile log-likelihood:

$$\ell(\boldsymbol{\theta}; \mathsf{Y}) = constant - \frac{1}{2}\left\{\sum_{i=1}^{n}\left(\mathsf{r}'\Sigma_i(\boldsymbol{\theta})\mathsf{r} + log|\Sigma_i(\boldsymbol{\theta})|\right)\right\}$$

where

$$\mathsf{r}_i = \mathsf{Y}_i - \mathsf{X}_i\left(\sum_{j=1}^{n}\mathsf{X}_j'\Sigma_j(\boldsymbol{\theta})^{-1}\mathsf{X}_j\right)^{-1}\sum_{j=1}^{n}\mathsf{X}_j'\Sigma_j(\boldsymbol{\theta})^{-1}\mathsf{Y}_j$$

Variance of MLE

The variance of the MLE estimate beta_hat_MLE is:

$$\mathrm{Var}(\hat{\boldsymbol{\beta}}_{MLE}) = \left(\sum_{i=1}^{n}\mathsf{X}_i'\Sigma_i(\hat{\boldsymbol{\theta}}_{MLE})^{-1}\mathsf{X}_i\right)^{-1}$$

Similarly to GLS we can have a more robust "sandwich" estimator that can be used for inference:

$$\mathrm{Var}^S(\hat{\boldsymbol{\beta}}_{MLE}) = \mathrm{Var}(\hat{\boldsymbol{\beta}}_{MLE})\left(\sum_{i=1}^{n}\mathsf{X}_i'\Sigma_i(\hat{\boldsymbol{\theta}}_{MLE})^{-1}\hat{\mathsf{e}}_i\hat{\mathsf{e}}_i'\Sigma_i(\hat{\boldsymbol{\theta}}_{MLE})^{-1}\mathsf{X}_i\right)\mathrm{Var}(\hat{\boldsymbol{\beta}}_{MLE})$$

Restricted Maximum Likelihood (REML)

In small samples, ML estimation generally leads to small sample bias in the estimated variance components. To account for the fact that we have uncertainty from estimating beta, we can get unbiased estimates for the variance by maximizing the restricted profile log likelihood:

$$\begin{aligned}\ell_R(\boldsymbol{\theta}; \mathsf{Y}) &= constant - \frac{1}{2}\left\{\sum_{i=1}^{n}\left(\mathsf{r}'\Sigma_i(\boldsymbol{\theta})\mathsf{r} + log|\Sigma_i(\boldsymbol{\theta})|\right)\right. \\ &\left. + \log\left|\sum_{i=1}^{n}\mathsf{X}_i'\Sigma_i(\boldsymbol{\theta})^{-1}\mathsf{X}_i\right|\right\}\end{aligned}$$

Variance of REML

Similar to other estimation procedures the variance of the REML estimate of beta_hat_REML is:

$$\mathrm{Var}(\hat{\boldsymbol{\beta}}_{REML}) = \left(\sum_{i=1}^{n}\mathsf{X}_i'\Sigma_i(\hat{\boldsymbol{\theta}}_{REML})^{-1}\mathsf{X}_i\right)^{-1}$$

We also have a robust "sandwich" estimator that can be used for inference:

$$\mathrm{Var}^S(\hat{\boldsymbol{\beta}}_{REML}) = \mathrm{Var}(\hat{\boldsymbol{\beta}}_{REML})\left(\sum_{i=1}^{n}\mathsf{X}_i'\Sigma_i(\hat{\boldsymbol{\theta}}_{REML})^{-1}\hat{\mathsf{e}}_i\hat{\mathsf{e}}_i'\Sigma_i(\hat{\boldsymbol{\theta}}_{MLE})^{-1}\mathsf{X}_i\right)\mathrm{Var}(\hat{\boldsymbol{\beta}}_{REML})$$

Inference

Consider the following linear representation of a response profile model:

$$\begin{aligned} Y_{ij} &= \beta_0 + \beta_1 X_{1ij} + \beta_2 X_{2ij} + \beta_3 X_{3ij} + \beta_4 X_{4ij} + \\ &\quad \beta_5 X_{1ij} X_{4ij} + \beta_6 X_{2ij} X_{4ij} + \beta_7 X_{3ij} X_{4ij} + \epsilon_{ij} \end{aligned}$$

where

$$X_{1ij} = \begin{cases} 1 & \text{when measurement is observed in week 1} \\ 0 & \text{otherwise} \end{cases}$$

$$X_{2ij} = \begin{cases} 1 & \text{when measurement is observed in week 4} \\ 0 & \text{otherwise} \end{cases}$$

$$X_{3ij} = \begin{cases} 1 & \text{when measurement is observed in week 6} \\ 0 & \text{otherwise} \end{cases}$$

$$X_{4ij} = \begin{cases} 1 & \text{for measurement is in the treatment group} \\ 0 & \text{otherwise} \end{cases}$$

This gives us the following hypotheses:

1. Group x Time Interaction Effect
   Ex. H0: Beta_5, Beta_6, Beta 7 = 0 vs H1: At least one does not equal 0
2. Time Effect
   Ex. H0: Beta_1 = Beta_2 = Beta_3 vs H1: Beta_1, Beta_2, Beta_3 are not equal
3. Group Effect
   Ex. H0: Beta_4 = 0 vs H1: Beta_4 != 0

A large sample test for the general linear hypothesis:
H0: L*Beta = 0; H1: L*Beta != 0
can be carried out using the Wald chi-square test

$$W^2 = (\mathsf{L}\hat{\boldsymbol{\beta}})'(\mathsf{L}\widehat{\mathrm{Var}(\hat{\boldsymbol{\beta}})}\mathsf{L}')^{-1}(\mathsf{L}\hat{\boldsymbol{\beta}})$$

Where L is a r x p matrix of r contrasts or rank r <= p. If we want to test H0: Beta_5, Beta_6, Beta 7 = 0 for group*interaction effect:

$$\mathsf{L} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{pmatrix}$$

It can be shown that W^2 -> chi-squared(r) as n approaches infinity.

Another more robust approach is to replace var(beta_hat) with the robust sandwich estimator, var_r(beta_hat). However, there may be a loss in efficiency compared to the model-based estimator.

Contrasts from Baseline

In clinical trials, baseline response is independent from treatment, which means the groups have the same mean response at baseline. In this setting, we may be more interested in testing for equality of the difference between the average responses post-baseline and the baseline value in the two groups.

$$\left(\frac{\mu_{T1}+\mu_{T4}+\mu_{T1}}{3}-\mu_{T0}\right)-\left(\frac{\mu_{P1}+\mu_{P4}+\mu_{P6}}{3}-\mu_{P0}\right)$$

$$\begin{aligned}
&= \left(\frac{\beta_0+\beta_1+\beta_4+\beta_5+\beta_0+\beta_2+\beta_4+\beta_6+\beta_0+\beta_3+\beta_4+\beta_7}{3}-\beta_0-\beta_4\right)\\
&\quad -\left(\frac{\beta_0+\beta_1+\beta_0+\beta_2+\beta_0+\beta_3}{3}-\beta_0\right)\\
&= \beta_5+\beta_6+\beta_7
\end{aligned}$$

with

$$L = (\ 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1/3 \quad 1/3 \quad 1/3\ )$$

We can test it using the Wald test, which is W^2 -> $chi^2(1)$

Area Under the Curve

Instead of comparing the difference of the average post-baseline responses, we can compare the area under the curve minus baseline (AUC). This corresponds to a calculation of the area under the trapezoidal curve.

$$\begin{aligned}
&(w_1 Y_{\cdot 1} + w_2 Y_{\cdot 2} + w_3 Y_{\cdot 3} + \ldots + w_p Y_{\cdot})^{TRT}\\
-\ &(w_1 Y_{\cdot 1} + w_2 Y_{\cdot 2} + w_3 Y_{\cdot 3} + \ldots + w_p Y_{\cdot})^{PL}
\end{aligned}$$

$$w_j = \begin{cases} \frac{1}{2}(t_1 + t_2 - 2t_p) & \text{for } j = 1 \\ \frac{1}{2}(t_p - t_{p-1}) & \text{for } j = p \\ \frac{1}{2}(t_{j+1} - t_{j-1}) & \text{for } j \neq 1, p \end{cases}$$

Baseline as Covariate

An efficient way for testing group differences in a longitudinal clincal trial is to remove the baseline measure from the response and use it as a covariate in the post-baseline responses instead:

$$\begin{aligned}
Y_{ij} &= \beta_0 + \beta_1 Y_{1j} + \beta_2 X_{2ij} + \beta_3 X_{3ij} + \beta_4 X_{4ij} + \\
&\quad + \beta_6 X_{2ij} X_{4ij} + \beta_7 X_{3ij} X_{4ij}
\end{aligned}$$

where $i = 2, 3, 4$ and

$$X_{2ij} = \begin{cases} 1 & \text{when measurement is observed in week 4} \\ 0 & \text{otherwise} \end{cases}$$

$$X_{3ij} = \begin{cases} 1 & \text{when measurement is observed in week 6} \\ 0 & \text{otherwise} \end{cases}$$

$$X_{4ij} = \begin{cases} 1 & \text{for measurement is in the treatment group} \\ 0 & \text{otherwise} \end{cases}$$

Strengths and Weaknesses

+ Is conceptually straightforward
+ It allows arbitrary patterns in the mean response over time
+ It allows for missing values in the response
+ There are various ways to adjust for the baseline

- Repeated measures must be obtained at the same sequence of time for all participants
- It ignores the time ordering of the repeated measures in a longitudinal study
- It may have low power to detect group differences in specific patterns of the mean response over time
- The number of estimated parameters grows rapidly with the number of measurement occasions

---

Revision #9
Created 26 January 2023 19:03:45 by Elkip
Updated 29 January 2023 14:28:08 by Elkip