Likelihood ratio test
2020-07-21

1 Model settings

The GEM model is specified as

$$\begin{aligned} \boldsymbol{y_{ki}} =& \boldsymbol{X_i}(\boldsymbol{\beta_k} + \boldsymbol{b_{ki}} + \boldsymbol{\Gamma_k}(\boldsymbol{\alpha' x_i})) + \boldsymbol{\epsilon_{ki}} \\ =& \boldsymbol{X_i}\boldsymbol{\beta_k} + (\boldsymbol{X}_i\boldsymbol{\Gamma}_k\boldsymbol{\alpha}')\boldsymbol{x}_i + \boldsymbol{X_i}\boldsymbol{b_{ki}} + \boldsymbol{\epsilon_{ki}} \end{aligned} \tag{1}$$

The multi-GEM model can be sepcified as

$$\begin{aligned} \boldsymbol{y_{ki}} =& \boldsymbol{X_i}(\boldsymbol{\beta_k} + \boldsymbol{b_{ki}} + \boldsymbol{\Gamma_k}(\boldsymbol{\alpha'_k x_i})) + \boldsymbol{\epsilon_{ki}} \\ =& \boldsymbol{X_i}\boldsymbol{\beta_k} + (\boldsymbol{X}_i\boldsymbol{\Gamma}_k\boldsymbol{\alpha}'_k)\boldsymbol{x}_i + \boldsymbol{X_i}\boldsymbol{b_{ki}} + \boldsymbol{\epsilon_{ki}} \end{aligned} \tag{2}$$

The unrestricted model is then

$$\begin{aligned} \boldsymbol{y_{ki}} =& \boldsymbol{X_i}(\boldsymbol{\beta_k} + \boldsymbol{b_{ki}} + \gamma_k \boldsymbol{x}_i) + \boldsymbol{\epsilon_{ki}} \\ =& \boldsymbol{X_i}\boldsymbol{\beta_k} + (\boldsymbol{X}_i\gamma_k)\boldsymbol{x}_i + \boldsymbol{X_i}\boldsymbol{b_{ki}} + \boldsymbol{\epsilon_{ki}} \end{aligned} \tag{3}$$

where

- $\boldsymbol{X}_i$ is the $n_t \times t$ dimension design matrix for subject i;
- $\boldsymbol{x}_i = \begin{pmatrix} x_{i1} & ... & x_{ip} \end{pmatrix}'$ is the $p \times 1$ vector of predictors for subject i; x_{ij} presents the jth predictor for the ith subject;
- $\boldsymbol{\beta}_k$ is the fixed effect coefficient with dimension $t \times 1$;
- $\boldsymbol{\Gamma}_k$ is the fixed effect coefficient associated with predictors.
- $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_p)'$ and $\boldsymbol{\alpha}_k = (\alpha_{k1}, ..., \alpha_{kp})'$ are $p \times 1$ vectors.
- $\boldsymbol{\gamma}_k$ is a $t \times p$ matrix
- $\boldsymbol{b}_{ki} \sim N(0, \boldsymbol{D}_k)$, $\boldsymbol{D}_k$ is a $t \times t$ matrix, with $\frac{t(t+1)}{2}$ parameters.
- $\boldsymbol{\epsilon}_{ki} \sim N(0, \sigma_k^2)$

Whether the it is a GEM model or an unrestricted model, it depends on the term $(\boldsymbol{X}_i\boldsymbol{\Gamma}_k\boldsymbol{\alpha}')\boldsymbol{x}_i$.

The GEM model means that, all subjects in each of the k treatment group have a shared $\boldsymbol{\alpha}$ vector.

If the subjects in different treatment group have different $\boldsymbol{\alpha}$, however, they have the same α within their treatment group and have the formula as $\boldsymbol{\Gamma}_k\boldsymbol{\alpha}'_k$, I call it as a multi-GEM model.

If there do not have any restrictions, the model has formula as equation (3), that is, the coefficients are different for different groups and they do not have the $\boldsymbol{\Gamma}_k\boldsymbol{\alpha}'_k$ form.
To make it more clear, let's assume

- $\boldsymbol{X}_i = (1, t, t^2)$; p = 2; k = 2

Then the term $(\boldsymbol{X}_i\boldsymbol{\Gamma}_k\boldsymbol{\alpha}')\boldsymbol{x}_i$ in equation (1) can be written as

- GEM model

$$\begin{aligned} (\boldsymbol{X}_i\boldsymbol{\Gamma}_k\boldsymbol{\alpha}')\boldsymbol{x}_i &= \begin{pmatrix} 1 & t & t^2 \end{pmatrix} \begin{pmatrix} \Gamma_{k1} \\ \Gamma_{k2} \\ \Gamma_{k1} \end{pmatrix} \begin{pmatrix} \alpha_1 & \alpha_2 \end{pmatrix} \begin{pmatrix} x_{i1} \\ x_{i2} \end{pmatrix} \\ &= \begin{pmatrix} 1 & t & t^2 \end{pmatrix} \begin{pmatrix} \Gamma_{k1}\alpha_1 & \Gamma_{k1}\alpha_2 \\ \Gamma_{k2}\alpha_1 & \Gamma_{k2}\alpha_2 \\ \Gamma_{k3}\alpha_1 & \Gamma_{k3}\alpha_2 \end{pmatrix} \begin{pmatrix} x_{i1} \\ x_{i2} \end{pmatrix} \\ &= \Gamma_{k1}\alpha_1 x_1 + \Gamma_{k2}\alpha_1 t x_1 + \Gamma_{k3}\alpha_1 t^2 x_1 + \Gamma_{k1}\alpha_2 x_2 + \Gamma_{k2}\alpha_2 t x_2 + \Gamma_{k3}\alpha_2 t^2 x_2 \end{aligned} \tag{4}$$

Therefore the degree of freedom for this term is

$$k \times 3 + (p-1) \quad (\alpha \text{ with df} = p-1) \tag{5}$$

- multi-GEM model

$$
\begin{aligned}
(\boldsymbol{X}_i\boldsymbol{\Gamma}_k\boldsymbol{\alpha}_k')\boldsymbol{x}_i &= \begin{pmatrix} 1 & t & t^2 \end{pmatrix} \begin{pmatrix} \Gamma_{k1} \\ \Gamma_{k2} \\ \Gamma_{k1} \end{pmatrix} \begin{pmatrix} \alpha_{k1} & \alpha_{k2} \end{pmatrix} \begin{pmatrix} x_{i1} \\ x_{i2} \end{pmatrix} \\
&= \begin{pmatrix} 1 & t & t^2 \end{pmatrix} \begin{pmatrix} \Gamma_{k1}\alpha_{k1} & \Gamma_{k1}\alpha_{k2} \\ \Gamma_{k2}\alpha_{k1} & \Gamma_{k2}\alpha_{k2} \\ \Gamma_{k3}\alpha_{k1} & \Gamma_{k3}\alpha_{k2} \end{pmatrix} \begin{pmatrix} x_{i1} \\ x_{i2} \end{pmatrix} \\
&= \Gamma_{k1}\alpha_{k1}x_1 + \Gamma_{k2}\alpha_{k1}tx_1 + \Gamma_{k3}\alpha_{k1}t^2x_1 + \Gamma_{k1}\alpha_{k2}x_2 + \Gamma_{k2}\alpha_{k2}tx_2 + \Gamma_{k3}\alpha_{k2}t^2x_2
\end{aligned} \tag{6}
$$

Therefore the degree of freedom for this term is

$$
k \times (3 + p - 1) \quad (\alpha_k \text{ with df} = p - 1) \tag{7}
$$

- unrestricted model

$$
\begin{aligned}
(\boldsymbol{X}_i\boldsymbol{\gamma}_k)\boldsymbol{x}_i &= \begin{pmatrix} 1 & t & t^2 \end{pmatrix} \begin{pmatrix} \gamma_{k11} & \gamma_{k12} \\ \gamma_{k21} & \gamma_{k22} \\ \gamma_{k31} & \gamma_{k32} \end{pmatrix} \begin{pmatrix} x_{i1} \\ x_{i2} \end{pmatrix} \\
&= \gamma_{k11}x_1 + \gamma_{k21}tx_1 + \gamma_{k31}t^2x_1 + \gamma_{k12}x_2 + \gamma_{k22}tx_2 + \gamma_{k32}t^2x_2
\end{aligned} \tag{8}
$$

Therefore the degree of freedom for this term is

$$
k \times 3 \times p \tag{9}
$$

The total df is $k \times (3 + 6 + 1 + 3p)$

Summary Table

DF	GEM model	multi-GEM model	unrestricted model
Predictor term	$3k + p - 1$	$k \times (p + 2)$	$3kp$
$k = 2, p = 3$	8	10	18
$k = 2, p = 10$	15	24	60
$k = 2, p = 20$	25	44	120

2 Likelihood ratio test

2.1 GEM model v.s. multi-GEM model

The hypothesis is

$$
\begin{aligned}
H_0 &: \ \boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_2 = \boldsymbol{\alpha} \\
H_1 &: \ \boldsymbol{\alpha}_1 \neq \boldsymbol{\alpha}_2
\end{aligned}
$$

The test statistics is

$$
T = -2(L_0 - L_1) \sim \mathcal{X}^2(df)
$$

The degree of freedom is $k(p + 2) - 3k - p + 1 = kp - p - k + 1$.

If the true data is generated from a GEM model, we should accept the H_0. And by repeating the simulation for 500 times, the p-values should have a (0,1) uniform distribution, the test statistics should follows the $\mathcal{X}^2(df)$ distribution.

If the true data is generated from a multi-GEM model, we should reject the H_0 hypothesis. We should get small p-values.

2.2 GEM model v.s. unrestricted model

The hypothesis is

$$
\begin{aligned}
H_0 &: \ \boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_2 = \boldsymbol{\alpha} \\
H_1 &: \ \text{the data does not have this structure}
\end{aligned}
$$

The test statistics is

$$
T = -2(L_0 - L_1) \sim \mathcal{X}^2(df)
$$

The degree of freedom is $3kp - 3k - p + 1$.

If the true data is generated from a GEM model, we should accept the H_0. And by repeating the simulation for 500 times, the p-values should have a (0,1) uniform distribution, the test statistics should follows the $\mathcal{X}^2(df)$ distribution.

If the true data is generated from a asa-GEM model, we should reject the H_0 hypothesis. We should get small p-values.

	Data generated from		
Comparison	GEM model	multi-GEM model	unrestricted model
GEM v.s. multi-GEM	H_0, large p-value	H_1, small p-value	H_1, small p-value
GEM v.s. unrestricted	H_1, small p-value	H_0, large p-value	
multi-GEM v.s. unrestricted		H_0, large p-value	H_1, small p-value

3 Likelihood calculation

To calculate the likelihood under GEM model,

- Use purity/likelihood method to find the optimal α
- Fit LME to calculate the loglikelihood

To calculate the likelihood under multi-GEM model, for each treatment group k

- Use likelihood method to find α that maximizes the likelihood of

$$\boldsymbol{y_{ki}} = \boldsymbol{X_i}\boldsymbol{\beta_k} + (\boldsymbol{X}_i\boldsymbol{\Gamma}_k\boldsymbol{\alpha}_k')\boldsymbol{x}_i + \boldsymbol{X_i}\boldsymbol{b_{ki}} + \boldsymbol{\epsilon_{ki}}$$

- Fit LME to calculate the loglikelihood

To calculate the likelihood under unrestricted model,

- Fit LME with interaction of $\boldsymbol{S}_i$ and predictors $\boldsymbol{x}_i$. Calculate the loglikelihood