大規模統計処理に対する期待と現実、そしてSAS　SASユーザー総会 2015

経時眼圧データの統計解析法に関する研究

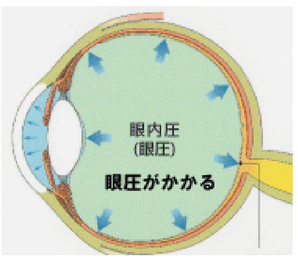
小川 岳人・佐野 雅隆・寒水 孝司・浜田 知久馬
東京理科大学大学院　工学研究科　経営工学専攻

背景・目的

➢ 眼圧と緑内障[4]

◆眼圧
- 眼内に生じている圧力
- 上昇すると，視神経が圧迫され，緑内障のリスクが上昇

◆緑内障
症状：視野障害，視力低下，失明等
一度失った視野や視力は改善されない
→ 眼圧を下げることによって，緑内障の悪化を防ぐ



(翠光会　高齢者の代表的な目の病気[4])

治療の評価 ＝ 眼圧降下作用の評価

- 緑内障の臨床試験では，各被験者の眼圧を反復測定
  →経時眼圧データ

➢ 経時眼圧データの特徴

◆各被験者に，左右両眼に関する眼圧
- 各被験者内での両眼の眼圧に相関
- 両眼の眼圧を独立とするとαエラー増大
- 片眼の眼圧のみ用いると検出力や解析精度低下

→両眼の眼圧を適切に用いたい

◆各被験者に反復測定 ＝ 経時データ
各眼内での時点ごとの眼圧に相関

これら2つの相関を考慮した解析が必要

➢ 想定される時点間の相関構造[3]

- AR(1) (Auto Regression)
  時点が離れると相関が小さくなる($0 < \rho < 1$)
  等間隔の経時データに対してよく用いられる
$$\begin{pmatrix} \sigma^2 & \rho\sigma^2 & \rho^2\sigma^2 \\ \rho\sigma^2 & \sigma^2 & \rho\sigma^2 \\ \rho^2\sigma^2 & \rho\sigma^2 & \sigma^2 \end{pmatrix}$$
- CS(Compound Symmetry)
  相関が全て等しい
$$\begin{pmatrix} \sigma_1^2+\sigma^2 & \sigma_1^2 & \sigma_1^2 \\ \sigma_1^2 & \sigma_1^2+\sigma^2 & \sigma_1^2 \\ \sigma_1^2 & \sigma_1^2 & \sigma_1^2+\sigma^2 \end{pmatrix}$$
- UN (Unstructured)
  特定の相関構造を仮定しない

緑内障の臨床試験を想定し，経時眼圧データの適切な統計解析法を明らかにする

方法

➢ 線形混合効果モデル（Linear Mixed Model;LMM）[3]

$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{\gamma}_i + \boldsymbol{\varepsilon}_i$$

- i：被験者
- $\boldsymbol{Y}_i$：反応変数ベクトル
- $\boldsymbol{X}_i$：固定効果に対するデザイン行列
- $\boldsymbol{Z}_i$：変量効果に対するデザイン行列
- $\boldsymbol{\beta}_i$：固定効果のパラメータベクトル
- $\boldsymbol{\gamma}_i \sim N(0, \boldsymbol{G})$：変量効果のパラメータベクトル
- $\boldsymbol{\varepsilon}_i \sim N(0, \boldsymbol{R})$：誤差ベクトル

➢ 分散の推定量[3]

◆ナイーブ分散
分散共分散行列の誤特定に関して頑健ではない
$$var(\hat{\boldsymbol{\beta}}) = \left(\sum_{i=1}^{N} \boldsymbol{X}_i^T \hat{\boldsymbol{V}}_i^{-1} \boldsymbol{X}_i\right)^{-1}$$

◆ロバスト(サンドウィッチ)分散
- 分散共分散行列の誤特定に関して頑健
- 保守的

$$var(\hat{\boldsymbol{\beta}}) = \left(\sum_{i=1}^{N} \boldsymbol{X}_i^T \hat{\boldsymbol{V}}_i^{-1} \boldsymbol{X}_i\right)^{-1} \left\{\sum_{i=1}^{N} \boldsymbol{X}_i^T \hat{\boldsymbol{V}}_i^{-1} var(\boldsymbol{Y}_i) \hat{\boldsymbol{V}}_i^{-1} \boldsymbol{X}_i\right\} \left(\sum_{i=1}^{N} \boldsymbol{X}_i^T \hat{\boldsymbol{V}}_i^{-1} \boldsymbol{X}_i\right)^{-1}$$

$$\boldsymbol{V}_i = \boldsymbol{V}_i(\alpha) = \boldsymbol{Z}_i \boldsymbol{G} \boldsymbol{Z}_i^T + \boldsymbol{R}_i$$
α：$\boldsymbol{V}_i$の分散成分

$$var(\boldsymbol{Y}_i) = (\boldsymbol{y}_i - \boldsymbol{X}_i\hat{\boldsymbol{\beta}})(\boldsymbol{y}_i - \boldsymbol{X}_i\hat{\boldsymbol{\beta}})^T$$

➢ 想定したLMM[3]

$$Y_{ijkt} = \beta_{1i} + \beta_{2t} + \beta_{12it} + \gamma_{jk} + \varepsilon_{ijkt}$$

◆固定効果
- β_{1i}：投与群
- β_{2t}：時点
- β_{12it}：交互作用

◆変量効果
- γ_{jk}：各個体の眼
- 両眼の間の相関を考慮
- 相関構造：CS

◆誤差
- ε_{ijkt}：各眼を単位とする
- 時点ごとの相関を考慮(経時誤差)

i：投与群
j：個体
k：右眼or左眼
t：時点

シミュレーション条件

◆2薬剤の眼圧降下作用を比較する臨床試験
- 総症例数：400眼(200人)
- 観測時点：0，3，6，9，12(ヶ月)
- 解析法：t検定，線形混合効果モデル(Linear Mixed Model;LMM)
- 最終観測時点での投与前との眼圧差を解析

◆シミュレーション回数：10000回

◆多変量正規乱数
- 時点0の母平均：19(mmHg)
- 標準偏差：2.5

	時点				
	0	3	6	9	12
群A	19	17	16	15.5	15
群B	19	16	15	14.5	14

(相関構造 AR(1) CS) × (各時点(0除く)の群間の差 [0, 0.2, …, 0.8, 1.0](mmHg))

- 紙面の都合上，相関構造AR(1)の結果のみ記載

◆反応変数：投与前(時点0)との眼圧差

◆データの欠測
- 欠測無し
- 観測値14mmHg以下→20%の確率で次の時点以降欠測

➢ 比較する解析法

◆ t検定
- 両眼の眼圧を独立とみなす場合(両眼t検定)
- 片眼の眼圧のみ用いる場合(片眼t検定)
- 両眼の平均値を用いる場合(両眼平均t検定)

◆ LMM(経時誤差項の相関構造，分散の推定量)
- CS，ナイーブ分散
- CS，ロバスト分散
- UN，ナイーブ分散
- UN，ロバスト分散

評価指標(有意水準両側5%)

αエラー，検出力，差の推定値の平均，標準偏差

結果

➢ 生データ相関構造:AR(1)，欠測無し

	αエラー	検出力 差0.4	検出力 差0.6	差の推定値(真値1.0)	標準偏差
両眼t検定	10.99	54.83	84.74	1.0006	0.2295
片眼t検定	5.10	32.32	61.13	1.0011	0.2681
両眼平均t検定	4.99	40.50	73.77	1.0006	0.2295
CS，ナイーブ分散	14.02	60.03	87.58	1.0006	0.2295
CS，ロバスト分散	5.17	41.21	74.31	1.0006	0.2295
UN，ナイーブ分散	8.91	51.21	82.40	1.0007	0.2291
UN，ロバスト分散	5.07	41.10	74.21	1.0007	0.2291

➢ 生データ相関構造:AR(1)，欠測有り

	αエラー	検出力 差0.4	検出力 差0.6	差の推定値(真値1.0)	標準偏差
両眼t検定	10.88	35.53	59.19	0.9977	0.3238
片眼t検定	5.25	18.52	35.56	0.9990	0.3751
両眼平均t検定	4.99	23.17	44.77	0.9977	0.3238
CS，ナイーブ分散	20.13	54.43	79.36	0.9977	0.2874
CS，ロバスト分散	5.28	29.62	55.29	0.9977	0.2874
UN，ナイーブ分散	10.03	41.61	68.93	0.9972	0.2780
UN，ロバスト分散	5.43	31.19	57.98	0.9972	0.2780

今後の課題

◆欠測値の補完法を使用した場合の性能評価
多重代入法など

◆結果の妥当性を実データ解析によって検討

まとめ

➢緑内障の臨床試験を想定し，t検定とLMMの性能を比較

- 片眼の眼圧のみ用いたt検定の検出力が低い
- 両眼を独立としたt検定のαエラー大
- 生データ相関構造がAR(1)の場合
  LMM(CS，ロバスト分散)の性能が良い
- 生データ相関構造がCSの場合
  - LMM(CS，ロバスト分散)のαエラーが多少大
  - データに欠測が無い場合は5%に保たれる
- 両眼平均t検定はαエラーが保たれるが，検出力と解析精度がLMM(CS，ロバスト分散)と比べて低い

LMM(CS，ロバスト分散)の性能が良い

参考文献

[1] R C Littell,G A Milliken,W W Stroup,and R D Wolfinger. SAS for Mixed Models,2nd edition.SAS Institute Inc,2006,pp.366-386.
[2] G Verbeke,G Molenberghs, Linear Mixed Models for Longitudinal Data.Springer Verlag New York,LLC,2009.
[3] SAS Institute Inc. 2013, SAS/STAT® 13.1 User's Guide – The MIXED Procedure. http://support.sas.com/documentation/onlinedoc/stat/131/mixed.pdf. (最終閲覧日：2015年1月17日)
[4]翠光会　上野毛クリニック　高齢者の代表的な目の病気 http://suikoukai-group.or.jp/html/page01.html. (最終閲覧日：2015年1月21日)
[5]日本眼科学会,目の病気　緑内障. http://www.nichigan.or.jp/public/disease/ryokunai_ryokunai.jsp. (最終閲覧日：2015年1月17日)
[6]日本緑内障学会,日本緑内障学会多治見緑内障疫学調査（通称：多治見スタディ）」報告. http://www.ryokunaisho.jp/general/ekigaku/tajimi.html. (最終閲覧日：2015年1月17日)