1A-2

経時対応データに対するノンパラメトリック効果の評価

永久保太士† 後藤昌司‡  
†アスピオファーマ株式会社  
‡特定非営利活動法人 医学統計研究会

1 序に代えて

繰り返し測定値のなかで、その影響要因がランダムに割り付けられず、主として時間上で系統的に連結されるデータを経時対応データと呼ぶ。経時対応データの解析において、計量的に測定できる応答のとり扱いが主に検討されてきている。とくに、応答が連続で正規分布に従うと仮定できる場合、繰り返し測定分散分析が多くの実質科学分野で用いられる。経時対応データの解析において代表的な理論を模索している。この方法は、データの変動要因の主効果と交互作用について定量化に評価することができ、さらに、結果の解釈が容易であるという利点をもっている。しかし、応答が連続データではなくカテゴリカル・データで得られる場合があること、このとき、繰り返し測定分散分析は妥当性と最適性に疑問をもたれる。カテゴリカル・データ、とくに順序カテゴリカル・データの解析では順位に基づく接法がよく知られている。

なかでも、順位に基づく方法として順位経験分布法は、実験における効果を全観測値の経験分布の平均と相対的に評価される相対効果に焦点をあてている。相対効果は、一つの実験の中での相対的な効果であり、異なる実験においては比較ができない。異なる実験における包括的な解釈を実現するために、経験分布ではない、ある既知の分布を参照分布として用いるリジェット（Ridit）効果が提案されている（Munzel & Langer, 2004）。

本報では、経時に得られる順序カテゴリカル・データに対する相対効果に基づく方法および相対効果の解釈の助けとなるリジェット効果に基づいた方法を適用し、相対効果とリジェット効果の関係を評価する。

2 相対効果

2.1 相対効果の定義

経時対応データのなかでも複数の群をもつ場合を検討する。群の数がI、時点の数がTの場合、群i(i = 1, ..., I)における対象k(k = 1, ..., n_i)が時点t(t = 1, ..., T)において観測され、観測値X_{ikt}が得られる。これは、経時対応データで最も頻繁に用いられる形式である。これらの群iにおける対象kの観測値をまとめて

\[ X_{ik} = (X_{ikt_1}, ..., X_{ikt_T})^T \]

とベクトルで表す。すべての観測値数を \( N = \sum_{t=1}^{T} n_i T \) とする。各対象の観測値ベクトルX_{ik}はそれぞれ独立と仮定される。しかし、これらのベクトルのそれぞれの成分は、同一の対象から得られた観測値であり、任意に他の成分に従属している。周辺分布は群と時点によって定まり、観測値X_{ikt}は周辺分布F_iに従う。

パラメトリック・モデルもしくはセミパラメトリック・モデルにおいて、効果は平均もしくは周辺分布の変化によって定義される。ノンパラメトリック・モデルにおける効果は周辺分布の変化の代わりに周辺分布そのものが用いられる。2標本の比較に関してノンパラメトリック効果は、ある分布関数に従う確率変数が他方の分布関数に従う確率変数よりも大きくなる確率で表される（Brunner, 1991）。一組の順序カテゴリカル・データに関するノンパラメトリック効果がAgresti (1983)によって考察されている。このノンパラメトリック効果を二つ以上の分布をもつ場合に拡張すると、群i、時点tにおける効果が

\[ p_{it} = \int_{-\infty}^{\infty} H dF_i \]  

(2.1)
と定義される。ここに，$F_{it}$ は群$i$ 時点$t$における周辺分布であり、また

$$H = \frac{1}{N} \sum_{t=1}^{T} \sum_{i=1}^{I} n_i F_{it}$$

は周辺分布 $F_{it}$ の重みつき平均である。この効果は、すべての分布関数の平均の分布関数 $H$ と比較して、群$i$における時点$t$の効果を相対的に評価するので、相対効果と呼ばれる。相対効果 $p_{it}$ は周辺分布関数 $F_{it}$ に従う確率変数が、平均分布関数 $H$ に従う確率変数より大きくなる確率であり、平均分布関数 $H$ に関する周辺分布関数 $F_{it}$ の傾向を量で定める。$F_{it}$ に従う確率変数が $H$ に従う確率変数より小さくなる傾向にあれば，$p_{it} < \frac{1}{2}$ である。$F_{it}$ に従う確率変数が $H$ に従う確率変数より大きくなる傾向にあれば，$p_{it} > \frac{1}{2}$ である。さらに，$p_{it} < p_{iv}$ という関係は $F_{it}$ が $F_{iv}$ よりも小さい傾向にあることを意味している。これらの周辺効果 $p_{it}$ により主効果と交互作用が実験研究において定義できる。

2.2 相対効果の推定

経験分布関数と順位の関係を用いると，群$i$ 時点$t$における相対効果は

$$\hat{p}_{it} = \int H dF_{it} = \frac{1}{n_i} \sum_{k=1}^{n_i} \hat{H}(X_{ikt}) = \frac{1}{N} \left( R_{it} - \frac{1}{2} \right)$$

によって推定される。$R_{it} = n_i^{-1} \sum_{k=1}^{n_i} R_{ikt}$ は，群$i$ 時点$t$における順位の平均である。順位変換法では，観測値 $\{X_{ikt}\}$ とその順位 $\{R_{ikt}\}$ を単におきかえるだけであるが，この順位経験分布法では，相対効果を推定するために順位が用いられる。

推定量 $\hat{p}_{it}$ の漸近的な分布は Brunner et al. (2002) によって示されている。

3 リジット効果

3.1 リジット効果の定義

リジット効果は相対効果と関連しており，以下の定義からそれは明らかである。$H_{R}(x)$ は既知の参照分布関数を示す。

群$i$ 時点$t$のリジット効果は

$$\theta_{it} = \int H_{R} dF_{it}$$

と定義される。ここに，$Z$ は $H_{R}(x)$ に従う任意の確率変数である。$\theta_{it} > \theta_{jt}(\theta_{it} < \theta_{jt})$ ならば，群$i$の観測値は群$j$の観測値と比べて確率論的に大きい（小さい）傾向をもつといわれる。$H_{R}(x)$ は既知であるので，効果 $\theta_{it}$ は包括的な解釈をもち，複数の試験にわたってまとめることができる可能性がある。もちろん，効果と群間の差の解釈は参照分布の選択に依存する。

リジット効果の不偏推定量は

$$\hat{\theta}_{it} = \int H_{R} d\hat{F}_{it} = \frac{1}{n_i} \sum_{k=1}^{n_i} H_{R}(X_{ikt})$$

である。潜在的な尺度 $S_{t}$ が $m \geq 2$ 個の順序カテゴリ $c_{1}, c_{2}, \ldots, c_{m}$ で構成されるとする。$\pi_{l} = \Pr(Z = c_{l})$ は，確率変数 $Z \sim H_{R}(x)$ が $l$ 番目のカテゴリに観測される確率を示す。分布関数の平均値により以下の式を得る。

$$H_{R}(c_{l}) = \sum_{s=1}^{t} \pi_{s} + \frac{1}{2} \pi_{l}$$

これは $\hat{\theta}_{it}$ の計算に用いられる。$N_{it}$ は群$i$ 時点$t$においてカテゴリ$c_{l}$ に観測される観測値 $X_{ikt}$ の数を示す。式 (3.5) を用いると推定量 $\hat{\theta}_{it}$ は

$$\hat{\theta}_{it} = \frac{1}{n_i} \sum_{l=1}^{m} N_{it} H_{R}(c_{l}) = \frac{1}{n_i} \sum_{l=1}^{m} N_{it} \left( \sum_{s=1}^{t-1} \pi_{s} + \frac{1}{2} \pi_{l} \right)$$

で表される。
3.2 一様分布

参照分布 \( H_R(x) \) が一様分布である状況を調べる。潜在的な尺度 \( S_c \) は計量的であるとすると、尺度の点 \( c_1, c_2, \ldots, c_m \) は加算および減算が可能です。さらに、カテゴリを自然数と仮定する。つまり \( c_l = 1 \) であるとする。これは \( l = 1, \ldots, m \) について、\( \pi_l = \Pr(X_{lk} = l) = 1/m \) および \( H_R(l) = (l - 0,5)/m \) であることを意味する。したがって、処理群 \( i \) における原観測値 \( X_{ik} \) の算術平均は

\[
X_i = \frac{1}{n_i} \sum_{k=1}^{n_i} X_{ik} = \frac{1}{n_i} \sum_{l=1}^{m} N_{il} \cdot l
\]

で定められる。

対応するリッジ効果の推定量は

\[
\hat{\theta}_i = \frac{1}{n_i} \sum_{l=1}^{m} N_{il} \left( \frac{l - \frac{1}{2}}{m} \right) = \frac{1}{m} \left( \frac{1}{n_i} \sum_{l=1}^{m} N_{il} \cdot l - \frac{1}{2} \right) = \frac{1}{m} \left( X_i - \frac{1}{2} \right)
\]

である。リッジ効果は参照分布の選択によって検出力が異なり、対立仮説における分式に適合した分布を選択することで検出力が大きくなくなることが示されている (Munzel & Langer, 2004)。

4 事例検討

相対効果とリッジ効果を事例に適用する。リッジ効果の参照分布として一様分布を用いる。ここでは、肩の痛みに関する臨床試験の例を用いる (Lumley, 1996)。この臨床試験では、胸腔鏡手術に伴う肩の痛みが手術後 6 時間にわたり調査された。痛みのスコアは 1(低い)から 5(高い)の 5 カテゴリである。適格と判断された 41 例の患者が治療 \( Y \) に 22 例、治療 \( N \) に 19 例が割り付けられた。

肩の痛み試験における各群、各群での相対効果とリッジ効果の値を表 4.1 と表 4.2 に示し、経時的な推移を図 4.1 と図 4.2 に示す。治療 \( Y \) では、時点 1 からスコアが低く、そこからさらに低くなることがわかる。対して、治療 \( N \) では時点 1 から時間の経過によりスコアが上がって、時点 5 からスコアが低くなっていく。相対効果とリッジ効果ではスコアの推移について違いは見られないが、効果の大きさは相対効果の方が大きくなっている。

5 結びに代えて

本報告では、繰り返し順序カテゴリーク・データに対して相対効果とリッジ効果を評価した。相対効果とリッジ効果では効果の示す大きさが異なることが示唆された。リッジ効果の大さでは参照分布の選択によっても変化すると考えられ、参照分布の選択に関する検討が今後は必要であると考える。また、効果の大きさが検出力にどのように影響を与えるのかを検討したい。
表 4.1: 肩先の痛み試験における各群各時点の相対効果

<table>
<thead>
<tr>
<th>時点</th>
<th>相対効果</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>治療 Y</td>
<td>0.477</td>
<td>0.425</td>
<td>0.358</td>
<td>0.374</td>
<td>0.321</td>
<td>0.352</td>
<td></td>
</tr>
<tr>
<td>治療 N</td>
<td>0.578</td>
<td>0.697</td>
<td>0.682</td>
<td>0.722</td>
<td>0.600</td>
<td>0.522</td>
<td></td>
</tr>
</tbody>
</table>

表 4.2: 肩先の痛み試験における各群各時点のリジェット効果

<table>
<thead>
<tr>
<th>時点</th>
<th>リジェット効果</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>治療 Y</td>
<td>0.264</td>
<td>0.218</td>
<td>0.164</td>
<td>0.173</td>
<td>0.136</td>
<td>0.155</td>
<td></td>
</tr>
<tr>
<td>治療 N</td>
<td>0.384</td>
<td>0.489</td>
<td>0.479</td>
<td>0.521</td>
<td>0.395</td>
<td>0.289</td>
<td></td>
</tr>
</tbody>
</table>

表 4.3: 肩先の痛み試験における各群各時点の検定結果

<table>
<thead>
<tr>
<th>相対効果</th>
<th>リジェット効果</th>
<th>繰り返し測定分散分析</th>
</tr>
</thead>
<tbody>
<tr>
<td>要因</td>
<td>F 値</td>
<td>p 値</td>
</tr>
<tr>
<td>治療</td>
<td>17.86</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>時間</td>
<td>4.08</td>
<td>0.0074</td>
</tr>
<tr>
<td>治療 × 時間</td>
<td>3.77</td>
<td>0.0111</td>
</tr>
</tbody>
</table>

参考文献


