

# 個人内共変と個人間相関の乖離を規定する要因 — 回帰係数に焦点をあてて —

北原瑞穂 (東京大学大学院)

## 問題と目的

2 変数の関連を考えるとき、「ある個人について、 $x$ の値が大きいほど $y$ の値も大きい」という個人内の共変関係(個人内共変)と、「集団において、 $x$ の値が大きい個人ほど $y$ の値も大きい」という個人間差に基づく相関関係(個人間相関)とを区別できる(南風原・小松, 1999)。しかし、実際の研究場面では、個人間相関によって個人内共変を模擬することが行われている(狩野, 2002)。このような模擬の当否を問う前提として、2 変数 $x, y$ の関連を回帰係数を用いて分析する場合を例に、個人内共変と個人間相関の乖離の大きさに影響する要因を検討する。

## 統計モデル

### 個人内共変

個人 $i$ の個人内での $x$ の平均(個人内平均)を $\mu_{xi}$ とし、分散(個人内分散)は全個人に共通で $\sigma_{xw}^2$ であるとする。また、個人間での個人内平均の平均を $\mu_x$ 、分散を $\sigma_{xb}^2$ とする。

個人 $i$ の時点 $t$ における $x, y$ の値の間には、

$$y_{it} = \beta_{0i} + \beta_{1i}x_{it} + e_{it}$$

という回帰モデルで表される共変関係があるとする。個人ごとに異なる回帰係数 $\beta_{0i}$ と切片 $\beta_{1i}$ (個人内回帰係数, 個人内切片とする)の平均をそれぞれ、 $\beta_{00}, \beta_{10}$ とし、これらの分散を $\tau_{00}, \tau_{11}$ , 共分散を $\tau_{10}$ とする。また、 $e_{it}$ は誤差で互いに独立であり、共通の分散をもつと仮定する。

### 個人間差データ

ある 1 時点 $t^*$ で横断データを収集することを考えると、得られる $x_{it^*}$ の個人間での分散について、 $\text{Var}(x_{it^*}) = \sigma_{xw}^2 + \sigma_{xb}^2$ が成り立つ。このうち個人間の分散が占める割合を $\eta_x^2$ とする。 $\eta_x^2 = \sigma_{xb}^2 / (\sigma_{xw}^2 + \sigma_{xb}^2)$ である。また、得られた $(x_{it^*}, y_{it^*})$ を用いて回帰分析を行うときに得られる回帰係数を $\beta_1^*$ とする。

## 個人内共変と個人間相関の乖離

個人内平均 $\mu_{xi}$ と個人内回帰係数 $\beta_{1i}$ が 2 変量正規分布に従うという仮定のもとでは、個人内回帰係数の平均である $\beta_{10}$ と個人間差に基づく回帰係数である $\beta_1^*$ との乖離 $\beta_1^* - \beta_{10}$ について、

$$\begin{aligned} \beta_1^* - \beta_{10} = & \text{Corr}(\beta_{0i}, \mu_{xi}) \sqrt{\tau_{00}} \frac{\eta_x^2}{\sigma_{xb}} \\ & + \mu_x \text{Corr}(\beta_{1i}, \mu_{xi}) \sqrt{\tau_{11}} \frac{\eta_x^2}{\sigma_{xb}} \end{aligned}$$

が成り立つ。したがって、乖離の大きさ $\beta_1^* - \beta_{10}$ は、 $\tau_{00}, \tau_{11}$ および $x$ の個人間分散 $\sigma_{xb}^2$ を一定とすれば、

(a)  $x$ の全変動のうち個人間分散の占める割合(分散比)  $\eta_x^2$ , (b)  $x$ の個人内平均 $\mu_{xi}$ と個人内切片 $\beta_{0i}$ の相関係数 $\text{Corr}(\beta_{0i}, \mu_{xi})$ , (c)  $\mu_{xi}$ と個人内回帰係数 $\beta_{1i}$ の相関係数 $\text{Corr}(\beta_{1i}, \mu_{xi})$ によって決まる。

(a)の分散比に注目した説明は生態学的誤謬の指摘でもなされてきたものである。また、(b)の個人内切片と $x$ の個人内平均の相関が第 3 の変数によって説明されるときは、狩野(2002)が指摘するような、乖離を変数の交絡の問題と解釈できる場合に相当する。(c)の個人内回帰係数と個人内平均の間に相関があることによる乖離は、この相関が他の変数によって説明される場合には、調整変数の問題として解釈できる。

## 課題と今後の展望

ここで示した乖離の大きさに影響する要因を踏まえて、実際のデータにおいてどの程度の乖離の大きさが見込まれるのか確認することが求められる。

個人内共変のモデルにおいては誤差間の相関を考慮したモデル等を考えることもでき、それらの要因が乖離に与える影響も検討する必要がある。また、2 変数の回帰分析だけでなく、より多様な分析手法についても同様の問題を考えることができ、検討の必要がある。