

多変量解析 レポート課題

2016 年 5 月 18 日出題
高澤 兼二郎

注意. 以下を守ること. 守らないレポートは採点の対象としないことがある.

- A4 片面で作成すること.
- 表紙をつけること. 表紙には, 以下を明記せよ:
 - タイトル (例:「多変量解析レポート」),
 - 学生証番号と氏名,
 - 出題日と提出日.
- 右上をホチキスで止めること.
- 2016 年 5 月 31 日までに事務室前のレポート BOX へ提出すること.
- 手書き, L^AT_EX などの書式は問わないが, 読みやすい大きさ, 濃さの文字を用いること.
- 教科書, ウェブ上の資料などの各種文献を参考にしてもよい. その際は以下を守ること:
 - 参考にした文献, および, 解答中のどの部分の参考にしたかを明記すること.
 - 参考文献を理解した上で, 自分の言葉で解答を記述すること. 機械的に丸写しすることは厳禁である. もし文献が間違っていたとしても, レポートの責任は作成した自分にあることを理解せよ.
- 不正をしないこと.

課題. 以下にしたがって重回帰分析を行え.

(i) 自分の好きな実データを用意し, 表に記せ. ただし, データは説明変数 x_1, \dots, x_p と被説明変数 y からなる n サンプルのデータ $(y_i, x_{i1}, \dots, x_{ip})$ ($i = 1, \dots, n$) であり, 以下をみたすこと:

- 説明変数は, 個数 p が 2 以上で, すべて量的変数であること.
- サンプル数 n が 10 以上 20 以下であること.
- $n > p$ であること.

(ii) 説明変数 x を任意に一つ選び, 以下にしたがって単回帰分析を行え.

- (a) 最小二乗法により回帰直線 $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ を求め, xy 平面上に (x_i, y_i) ($i = 1, \dots, n$) および (\bar{x}, \bar{y}) とともに図示せよ. ただし, $\bar{x} = \sum_{i=1}^n x_i/n$, $\bar{y} = \sum_{i=1}^n y_i/n$ である.
- (b) $\hat{\beta}_0, \hat{\beta}_1$ の意味を説明せよ. また, 実際のデータと照らし合わせ, 妥当な値であるかどうか説明せよ.
- (c) 寄与率 R^2 および自由度調整済寄与率 R^{*2} を計算し, 分析結果の妥当性を論じよ.
- (d) 標準化残差 e'_i が大きいサンプル i がある場合, そのサンプルを除くべきかどうか検討せよ. 除くべきと判断した場合は, そのサンプルを除いた回帰直線を求め, (ii)(a) の図に追加せよ.

(iii) 以下にしたがって重回帰分析を行え.

- (a) 最小二乗法により, 重回帰式 $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \dots + \hat{\beta}_p x_p$ を求めよ.
- (b) $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p$ の意味を説明せよ. また, 実際のデータと照らし合わせ, 妥当な値であるかどうか説明せよ.
- (c) 寄与率 R^2 および自由度調整済寄与率 R^{*2} を計算し, 分析結果の妥当性を論じよ.
- (d) 標準化残差 e'_i が大きいサンプル i がある場合, そのサンプルを除くべきかどうか検討せよ. 除くべきと判断した場合は, そのサンプルを除いた重回帰式を求めよ.

(発展) 変数増加法により説明変数の選択を行え.

(以上)