混合効果分散分析モデルを適用した
自動車販売台数に関する統計的分析

2017SS024 各務歌倫

指導教員：白石高章

1　はじめに

私はレンタカーのアルバイトを通し、自動車の貸し出しやカーリースの予約受付を行っている。車種の決定からお客様のご要望をヒアリングしそれに適した自動車を提供する必要がある。また様々な自動車を洗車や点検，運転できる機会もあり自動車は普段から私の身近な存在である。そのような経験から自動車に興味を持つようになった。そして自動車についてより深く知りたいという気持ちが強くなり最も登録台数が多いトヨタ自動車はじめホンダ，日産自動車の販売台数を統計的分析していくことにする．

2　データについて

本研究では、一般社団法人自動車販売協会連合会　新車販売台数『統計データ』の 2016 年以降フルモデルチェンジが行われたトヨタ，ホンダ，日産の新車乗用車販売台数の 9 車種 (プリウス，パッソ，カローラ，クラウン，RAV4，フィット，フリード，セレナ，リーフ)(文献 [1]　参照　) を使用する．フルモデルチェンジした月から 1 ヵ月後，4 か月後，8 ヵ月後のデータを収集し，以下の表 1 に示す．

表 1　フルモデルチェンジにおける販売台数の経時変化

群	車種	1 ヵ月後	4 か月後	8 ヶ月後
群 1：トヨタ	プリウス	22447	17946	10014
	パッソ	6784	6476	5269
	カローラ	11190	8480	3802
	クラウン	7225	6715	6540
	RAV4	6817	6277	5759
群 2：ホンダ	フィット	14845	9016	9001
	フリード	9153	9020	8626
群 3：日産	セレナ	6488	6068	4354
	リーフ	3629	3678	1765

3　多群の混合効果分散分析モデル

混合効果分散分析モデルは，時点効果と個体効果を共に含む統計的モデルである．データが 1 群の場合は目的変数となる測定値を時点，個体の因子で説明しこれで説明できない部分を誤差と考えるモデルである．このとき，時点要因は変量因子であると考える．データがいくつかの群に分かれている (多群) の場合には，群効果を加味し，さらに時点と群との交互作用を考慮したモデルである．

4　モデルの設定

藤越ら [2] より多群の場合の混交効果分散分析モデルは次のように表すことができる．

$$y_{ijk} = \mu + \alpha_j + \delta_{ij} + \beta_k + \gamma_{jk} + \epsilon_{ijk} \qquad (1)$$

$$i = 1, \cdots, n_j \ ; \ j = 1, \cdots, q \ ; \ k = 1, \cdots, p$$

μ は一般効果，α_j は第 j 群の効果，β_k は k 時点での効果，γ_{jk} は第 j 群と第 k 時点との交互作用，δ_{ij} は第 j 群の第 i 番目の個体変動を表す変量 (確率変数) である．誤差項 ϵ_{ijk} は互いに独立でそれぞれ $N(0, \sigma^2)$ に従い，個体変動 δ_{ij} は互いに独立でそれぞれ $N(0, \sigma_\delta^2)$ に従い，個体変動と誤差項は互いに独立とする．また未知母数に対して，$\sigma_\delta^2 \geq 0,\ \sigma^2 > 0$ を仮定し，ウエイト付き制約条件

$$C_1 : \sum_{j=1}^{q} n_j \alpha_j = 0,\ \sum_{\mathrm{k}=1}^{\mathrm{p}} \beta_k = 0$$

$$\sum_{j=1}^{q} n_j \gamma_{jk} = 0,\ (k = 1, \cdots, p),\ \sum_{j=1}^{p} \gamma_{jk} = 0,\ (j = 1, \cdots, q)$$

をみたすものとする．

5　効果の推定

群効果 α_j の最尤推定量あるいは最小 2 乗推定量は交互作用の有無に関係なく

$$\hat{\alpha}_j = \bar{y}_{.j.} - \bar{y}_{...},\ j = 1, ..., q$$

時点間効果 β_k の最尤推定量あるいは最小 2 乗推定量も交互作用の有無に関係なく

$$\hat{\beta}_k = \bar{y}_{..k} - \bar{y}_{...}, k = 1, ..., p$$

5.1　分析方法

フルモデルチェンジ経過月数の変化に伴い販売台数に変化があるか，車種を 3 つのブランド (群) に分け，群間の有意性を明らかにする．時点，個体，群，交互作用，誤差の変動要素についての分散分析表から，各効果の有意性を検証する. 時点要因と群要因が有意であれば，どの水準が有意であるかを多重比較法を用いて検証する．このとき，群要因と時点要因との交互作用も考えられるので，以下では交互作用がある場合とない場合に分けて分散分析を考えることにする．

時点要因による平方和を s_β^2，個体要因による平方和を s_δ^2，群要因による平方和を s_α^2，交互作用による平方和を s_γ^2，誤差 (残差) による平方和を s_e^2 とし，全体平方和を s_t^2 と表す．このときこれらは次のように表せる．

$$s_\beta^2 = n \sum_{k=1}^{p} (\bar{y}_{..k} - \bar{y}_{...})^2,\ s_\delta^2 = p \sum_{j=1}^{q} \sum_{i=1}^{n_j} (\bar{y}_{ij.} - \bar{y}_{.j.})^2,$$

$$s_\alpha^2 = p \sum_{j=1}^{q} n_j (\bar{y}_{.j.} - \bar{y}_{...})^2,\ s_\gamma^2 = \sum_{j=1}^{q} \sum_{k=1}^{p} n_j (\bar{y}_{.jk} - \bar{y}_{.j.} - \bar{y}_{..k} + \bar{y}_{...})^2,$$

$$s_e^2 = \sum_{j=1}^{q}\sum_{i=1}^{n_j}\sum_{k=1}^{p}(\bar{y}_{ijk} - \bar{y}_{.jk} - \bar{y}_{ij.} + \bar{y}_{.j.})^2.$$

このとき全平方和は

$$s_t^2 = \sum_{j=1}^{q}\sum_{i=1}^{n_j}\sum_{k=1}^{p}(\bar{y}_{ijk} - \bar{y}_{...})^2 = s_\alpha^2 + s_\beta^2 + s_\delta^2 + s_\beta^2 + s_\gamma^2 + s_e^2$$

となる．

以下 (1) ～ (4) の効果の F 検定は，藤越ら [2] の p.58 の方法を使う．

(1) 帰無仮説： $H_\alpha; \alpha_1 = \cdots = \alpha_q = 0$ に対する群間の検定

(2) 帰無仮説： H_γ ; すべての $\gamma_{ik} = 0$ に対する群と時点との交互作用の検定

(3) 帰無仮説： $H_\beta; \beta_1 = \cdots = \beta_p = 0$ に対する時点間の検定

(4) 帰無仮説： $H_\delta; \sigma_\delta^2 = 0$ に対する個体差の検定

これらの結果から表 1 のデータに混合分散分析モデルを適用すると，次の分散分析表が導かれる．

表 2 フルモデルチェンジにおける分散分析表

変動	偏差平方和	自由度	分散比	p 値
全体変動	512327140.5	26	-	-
時点間変動	62402154.74	2	5.75	0.0176
個体間変動	262630836.1	6	8.073	0.0011
群間変動	112615831.7	2	1.28	0.3117
交互作用	9618276.72	4	0.44	0.7751
誤差変動	65059328.42	12	-	-

分散分析表の時点間変動に着目すると，p 値 $= 0.0176$ より棄却された．

6 Tukey 型の多重比較検定

表 2 より，時点間の一様性が有意水準 2 %で棄却されたので，時点間 { 帰無仮説 $H_{(k,k')} : \beta_k = \beta_{k'}$ vs. 対立仮説 $H^A_{(k,k')} : \beta_k \neq \beta_{k'} \mid 1 \leq k < k' \leq p\}$ に対する多重比較検定を考える．検定統計量は，

$$T_{kk'} = \frac{\hat{\beta}_k - \hat{\beta}_{k'}}{\sqrt{(\frac{1}{n} + \frac{1}{n})\hat{\sigma}^2}} \qquad (2)$$

表 2 の場合，交互作用はないと考え，$\hat{\sigma}^2 = (s_e^2 + s_\gamma^2)/\{(p-1)(n-1)\}$ とした．

6.1 検定方式

$m = (p-1)(n-1),\ n = n_1 + \cdots + n_q$ とし，

$$TA(t) \equiv p\int_0^\infty \left[\int_{-\infty}^{\infty}\{\Phi(x) - \Phi(x - \sqrt{2}\cdot ts)\}^{p-1} d\Phi(x)\right] g(s|m)ds,$$

ただし、

$$g(s|m) \equiv \frac{m^{m/2}}{\Gamma(m/2)2^{(m/2-1)}} s^{m-1}\exp(-ms^2/2)$$

とする．

方程式 $TA(t) = 1 - \alpha$ を満たす t の解を $ta(p, m; \alpha)$ とおく．

$$|T_{kk'}| \geq ta(p, m; \alpha) \Rightarrow H_{(k,k')} \text{ を棄却する}$$

と判定する．

6.2 時点間変動の解析

表 1 から $n = 9$，$p = 3$ とし，$\alpha = 0.05$ とする．$ta(p, m; \alpha) = ta(3; 16; 0.05) = 2.580$ (白石・杉浦 [4] 参照)

(2) 式よりそれぞれ検定統計量を求めると

$|T_{1,2}| = 1.63 < 2.580, \quad |T_{1,3}| = 3.667 \geq 2.580$

$|T_{2,3}| = 2.034 < 2.580$

となり帰無仮説 $H_{(1,3)}$ が棄却された．

7 トヨタ自動車に限定した解析

表 1 の群 1 だけを使って 1 群の混合効果分散分析モデルでの多重比較解析を行う．この方法は Shiraishi and Matsuda [3] で論じた乱塊法モデルにおける多重比較法と同じである．y がすべての個体に対して同じ p 個の時点 (あるいは条件) t_1，$\cdots$，t_p で測定されているとき，個体 i の t_j 時点での y の測定値を y_{ij} とし，y_{ij} は，一般効果 μ，個体 i のランダム効果 δ_i，時点 t_j の母数効果 β_j，および誤差項 ε_{ij} の和として 1 群の

$$y_{ij} = \mu + \delta_i + \beta_j + \varepsilon_{ij},\ i = 1, \cdots, n\ ;\ j = 1, \cdots, p$$

と表せる．ただし $\sum_{j=1}^{n} \beta_j = 0$ とする．時点間に対する多重比較検定を行ったところ，棄却されなかった．

8 おわりに

本研究を通してフルモデルチェンジによって販売台数に影響を及ぼしていることがわかった. またデータを 1 群と多群に分け実際に解析したところ，1 群では棄却されなかったが多群では棄却された．このように結果の違いを可視することができた．

参考文献

[1] 一般社団法人日本自動車販売協会連合会 :『統計データ』．http://www.jada.or.jp/data/month/m-brand-ranking/ (2020 年 9 月閲覧)

[2] 藤越康祝, 管民郎, 土方裕子『経時データ分析』オーム社, 東京, 2008

[3] Shiraishi, T. and Matsuda，S. Closed testing procedures for all pairwise comparisons in a randomized block design．Commun. Statist., SerA. 47, p3571-3587.(2018)．

[4] 白石高章, 杉浦洋『多重比較法の理論と数値計算』共立出版, 東京, 2018