

インサイト・ファクトリー 社内セミナー
2020年4月

マーケティング・ミックス・モデリング

IV. 静学的市場反応モデル

小野 滋

統計学・データ解析	マーケティング・ミックス・モデリングへの適用
I. イントロダクション	
	II. 市場反応モデルとは
III. 回帰分析の基礎	
	IV. 静学的市場反応モデル
V. 時系列分析の基礎	
	VI. 動学的市場反応モデル(1)
VII. 状態空間モデルの基礎	
	VIII. 動学的市場反応モデル(2)

目次

この章の内容

サンプルデータ

1. はじめに：市場反応モデリングにおける6個の挑戦
2. 変数の操作的定義
3. 関数形 I. 売上モデル
4. 関数形 II. シェアモデル
5. 説明変数の欠落
6. 多重共線性への対処
7. 交互作用
8. 分散の不均一性・非正規性
9. 異質性
10. 推定結果のプレゼンテーション

この章の引用文献

この章に登場したRの関数

静学的な市場反応モデルについて議論します

- 静学的 (static) とは
 - 目的変数と説明変数が、同一の期間でのみ関係を持つと考える
 - いいかえれば、時間という概念を無視する
- クロスセクション・データを想定します
 - 複数の主体の一時点(期間)の観察
 - 例) 地域ごとの観察, 店舗ごとの観察

本資料作成に用いたすべてのRコードを、以下で公開しています:

https://rpubs.com/shig_ono/MMM_4

ある消費財カテゴリ X に注目する。

このカテゴリの小売チャネルは、当該カテゴリの専門店 (全国200店) のみである。

市場は3つのブランドに寡占されている。

いずれのブランドでも、マーケティング活動は次の3種類である：

- 小売店への販売補助支出 (什器などに用いられる)
- 小売店への宣材送付
- 小売店への販売員訪問

なお、

- ある店舗に対するマーケティング活動は、その店舗の売上とは無関係に計画される。
- ある期の売上数量は、当期のマーケティング活動の影響を受けるが、前期までの活動の影響を受けない。
- ある小売店に対する活動は、他の小売店には影響しない。
- 小売店のブランド売上は、マーケティング活動の下で独立であり、空間相関はない(「近くにある店舗の売上は似ている」といった現象はない)。
- マーケティング活動は売上に影響するが、来店客数に影響しない。

(ずいぶん不自然な例ですが、何卒ご容赦ください)

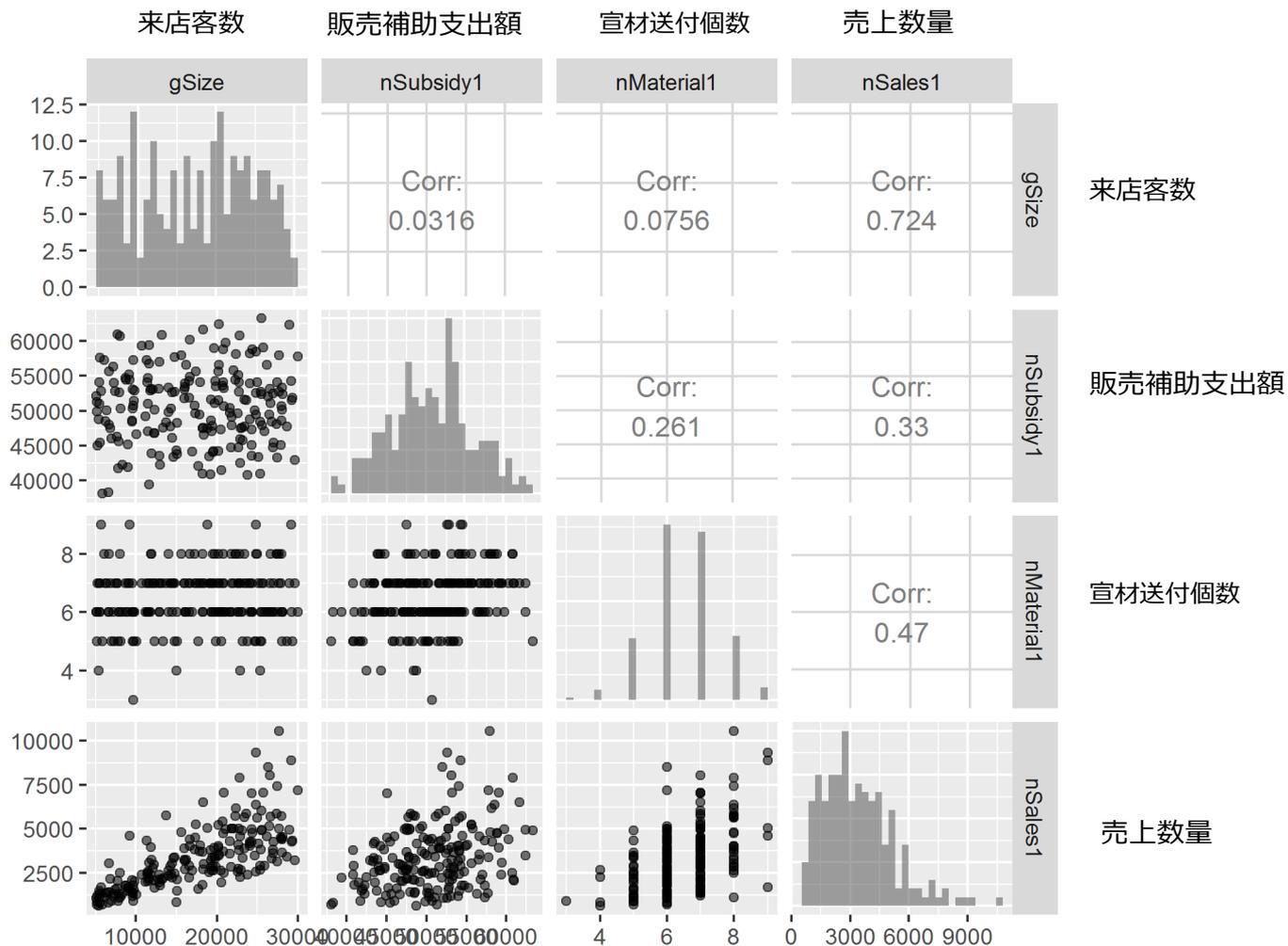
データに含まれている変数

Code 1

変数名	内容	有効な値の数	最小値	平均	最大値
nShopID	店舗番号	200	1.0	100.5	200.0
nAreaID	チェーン番号	200	1.0	7.0	12.0
gSize	来店客数	200	5130.3	17476.3	29980.0
nSales1	ブランド1の売上数量	200	624.0	3376.9	10548.0
nSales2	ブランド2の売上数量	200	195.0	6785.8	10000.0
nSales3	ブランド3の売上数量	200	175.0	620.4	1055.0
gShare1	ブランド1の売上数量シェア	200	0.1	0.3	0.8
gShare2	ブランド2の売上数量シェア	200	0.1	0.6	0.9
gShare3	ブランド3の売上数量シェア	200	0.0	0.1	0.1
nSubsidy1	ブランド1の販売補助支出額	200	38098.0	50725.6	63278.0
nSubsidy2	ブランド2の販売補助支出額	200	24531.0	39485.8	53891.0
nSubsidy3	ブランド3の販売補助支出額	200	60171.0	70169.5	79610.0
nSubsidy1A	ブランド1の販売補助支出額(タイプA)	200	23056.0	25246.4	27442.0
nSubsidy1B	ブランド1の販売補助支出額(タイプB)	200	13176.0	25479.2	38706.0
nMaterial1	ブランド1の宣材送付個数	200	3.0	6.5	9.0
nMaterial2	ブランド2の宣材送付個数	200	2.0	4.6	7.0
nMaterial3	ブランド3の宣材送付個数	200	3.0	8.8	15.0
nVisit1	(謎)	200	0.0	5.5	12.0

ブランド1の変数の散布図行列

Code 2



1. はじめに：売上反応モデリングにおける6個の挑戦

1. はじめに：売上反応モデリングにおける6個の挑戦

売上反応モデルの基礎となるのは回帰モデルである。

いっぽう、売上反応の特性は、回帰モデルの前提と多くの点で対立する。

本章では、そのうち6つの特性に注目し、対立を乗り越える方法について議論する。

回帰モデル

$$Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_K X_{Ki} + U_i$$

[1] X_i は確率変数でないか
 U_i と統計的に独立

[2] $E(U_i) = 0$

[3] $\text{Var}(U_i) = \sigma^2$

[4] $\text{Cov}(U_i, U_j) = 0, i \neq j$

[5] U_i は正規分布に従う

[6] X の列ベクトルは一次独立。
列数は行数より小さい

市場反応の特性

A) マーケティング活動と市場反応の関係は、**直線的ではない**かもしれない

B) 市場反応変数として、売上を用いる場合と**シェア**を用いる場合がある

C) 市場反応に影響する変数は多様であり、**データを入手できない変数**も多い

D) 広告・プロモーションの効果は、その**内容や媒体**によっても異なるかもしれない

E) 異なるマーケティング活動を同時に行うことにより、**シナジー**が生まれるかもしれない

F) マーケティング活動と市場反応の関係には、**異質性**があるかもしれない

本章との対応

3節

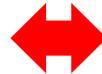
4節

5節

6節

7節

9節



2. 変数の操作的定義

2. 変数の操作的定義 (HPS pp.49-51)

本題に入る前に、モデリングに用いる変数の操作的定義について考える

- 操作的定義 ... 概念上の定義ではなく、実際の分析手続き上の定義

以下について述べる：

1. ダミー変数
2. 相対的変数

2-1. ダミー変数

カテゴリカルな説明変数を、複数の二値変数で表現すること

- Kカテゴリの質的変数は、K-1個の二値変数で表現できる

例) 店舗タイプの表現

店舗タイプ	大型店 ダミー 変数	専門店 ダミー 変数
大型	1	0
小型	0	0
専門店	0	1
小型	0	0
専門店	0	1
大型	1	0

2-2. 相対的変数

目的変数・説明変数を、なんらかの値に対する比で表現すること

相対的変数を使う理由：

- 市場反応・マーケティング活動を表す変数として、比が適切だから
 - 例) 地域の売上数量 → 地域の人口10万人当たり売上数量
 - 例) 月次の売上数量 → 月次の1日当たり売上数量
- 説明変数の数を減らすため
 - 例) 売上シェア

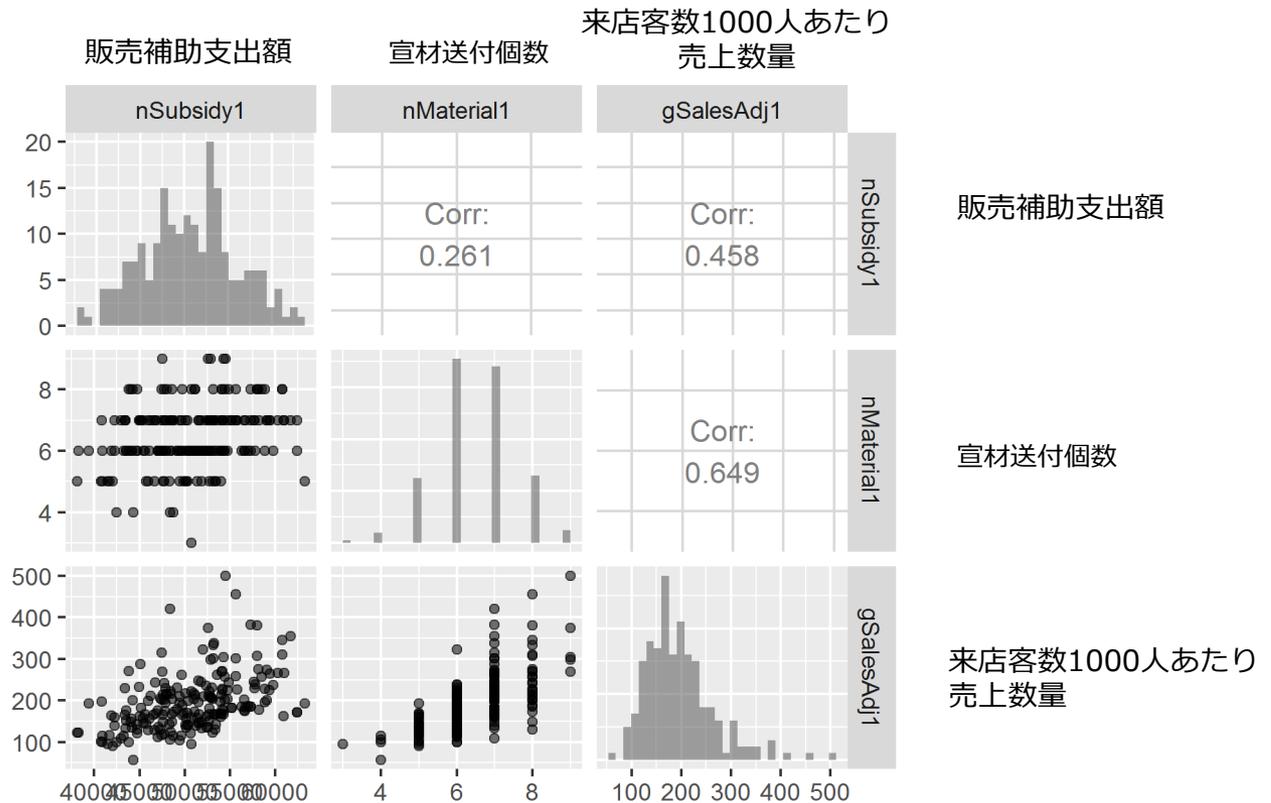
相対的変数のタイプ：

- カテゴリ合計に対する比
- 競合の値に対する比
- なんらかの「ベース」に対する比
 - 例) 通常価格に対する相対的価格
- その他、別の変数に対する比
 - 例) 地域の人口; 月の日数

店舗販売データ

Code 3

- マーケティング活動は「来店客数あたり売上数量」に影響するが、来店客数には影響しないと考えて...
- 売上数量を「来店客数1000人あたり売上数量」に変換しよう



3. 関数形 I. 売上モデル

回帰モデル

$$Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_K X_{Ki} + U_i$$



市場反応の特性

A) マーケティング活動と市場反応との関係は、**直線的ではない**かもしれない

この対立を乗り越える方法として、**関数形**の設定がある。

本節では、目的変数が売上である場合の関数形について検討する。

以下では、店舗 i ($= 1, \dots, N$) における自社ブランドの来店客1000人あたり売上数量を Y_i とする。

関数形

- 線形モデル
- 片対数モデル
- 両対数モデル (= 累乗モデル)
- 指数モデル
- 逆関数モデル
- ロジスティックモデル
- ...



市場反応の特性

- 規模に関して収穫一定
- 規模に関して収穫逓減
- 規模に関して収穫逓増
- S字型
- 飽和状態がある
- ...

3-1. 規模に関する収穫一定を表現する関数形

マーケティング変数が1単位変わると、売上が一定単位変わる

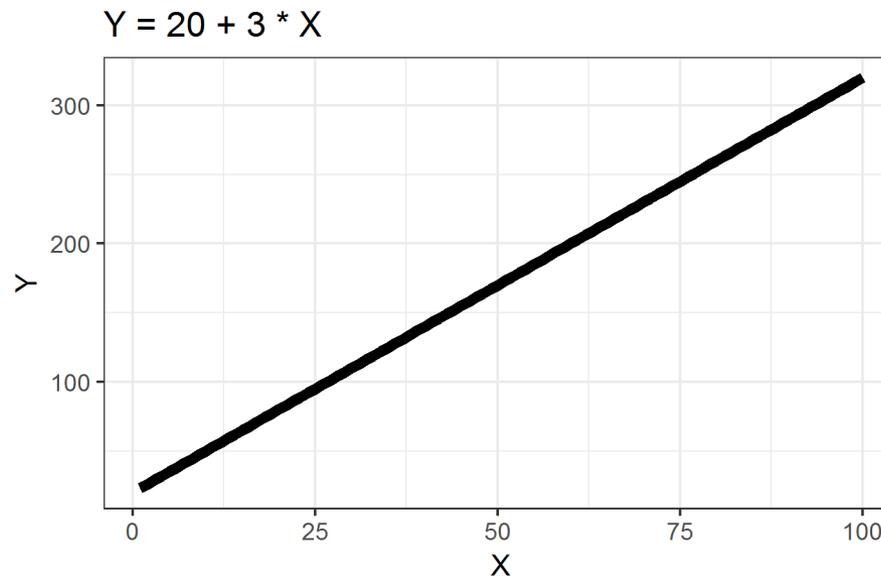
Code 4

- 現実的でない
- しかし、限られた範囲においては十分な近似であることが多い



線形モデル

$$Y_i = \beta_1 + \beta_2 X_{2i} + \dots$$



3-2. 規模に関する収穫逓減を表現する関数形

マーケティング変数を1単位ずつ増やすと、1単位増加に伴う売上増加は、その前の1単位増加に伴う売上増加よりも小さい

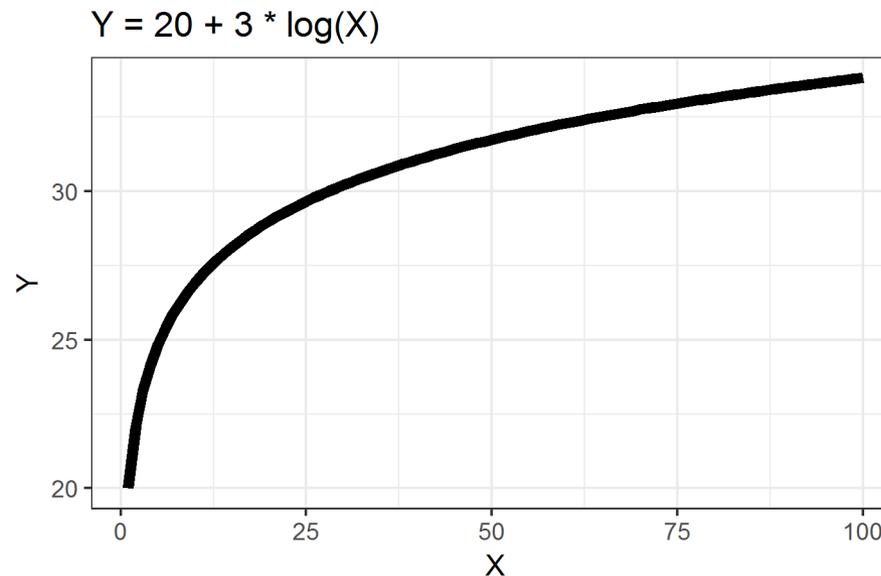
Code 4

- 広告において一般的にみられる特徴



- **片側対数モデル** (lin-logモデル)

$$Y_i = \beta_1 + \beta_2 \log X_{2i} + \dots$$



- 両側対数モデル (log-logモデル)

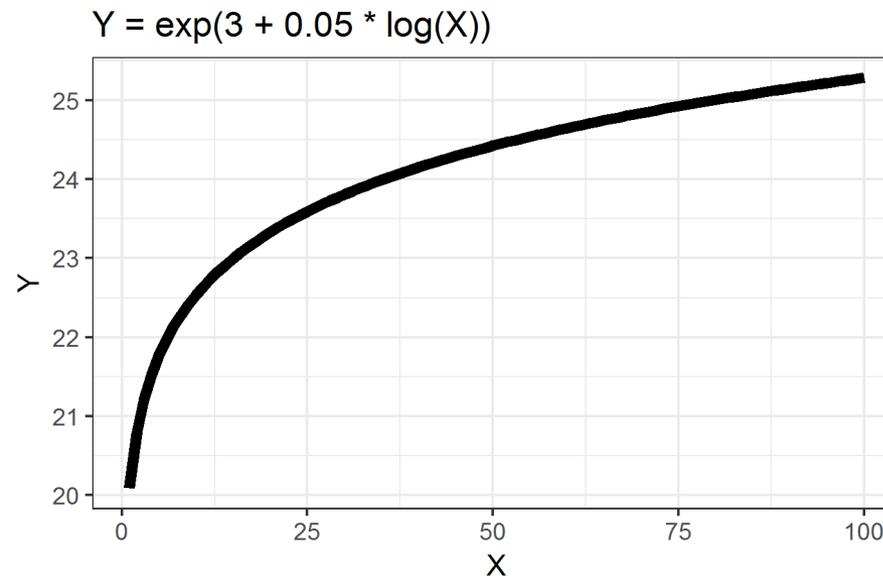
Code 4

$$\log Y_i = \beta_1 + \beta_2 \log X_{2i} + \dots$$

- すべての説明変数を対数変換するならば、このモデルは**乗法モデル**と同じ

$$Y_i = \exp(\beta_1) X_{2i}^{\beta_2} X_{3i}^{\beta_3} \dots$$

- β_2 は「 X_{2i} が1%増大すると Y は何%増大するか」すなわち**弾力性**を表す



3-3. 規模に関する収穫逓増を表現する関数形

マーケティング変数を1単位ずつ増やすと、1単位増加に伴う売上増加は、その前の1単位増加に伴う売上増加よりも大きい

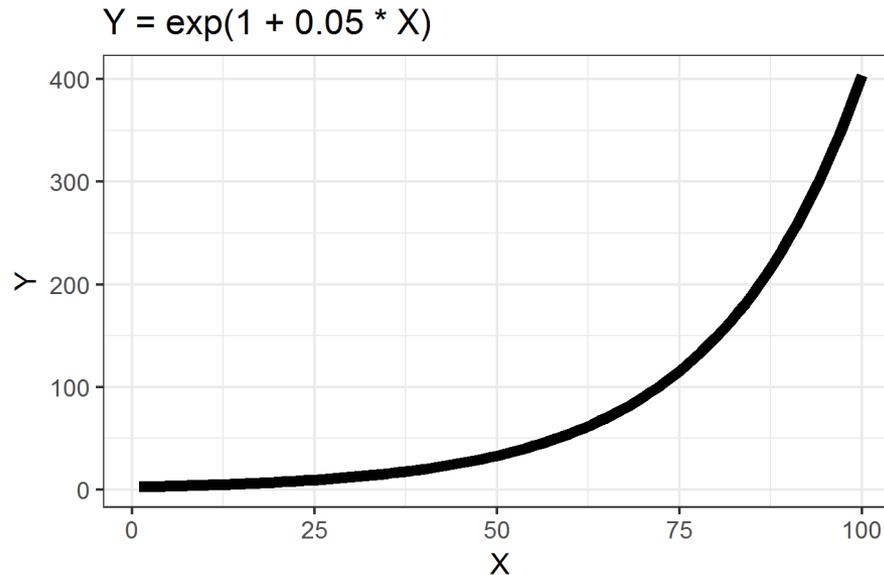
Code 4

- 価格プロモーション(値引き幅)について観察されることがある



指数モデル (log-linモデル)

$$\log(Y_i) = \beta_1 + \beta_2 X_{2i} + \dots$$



- 逆関数モデル

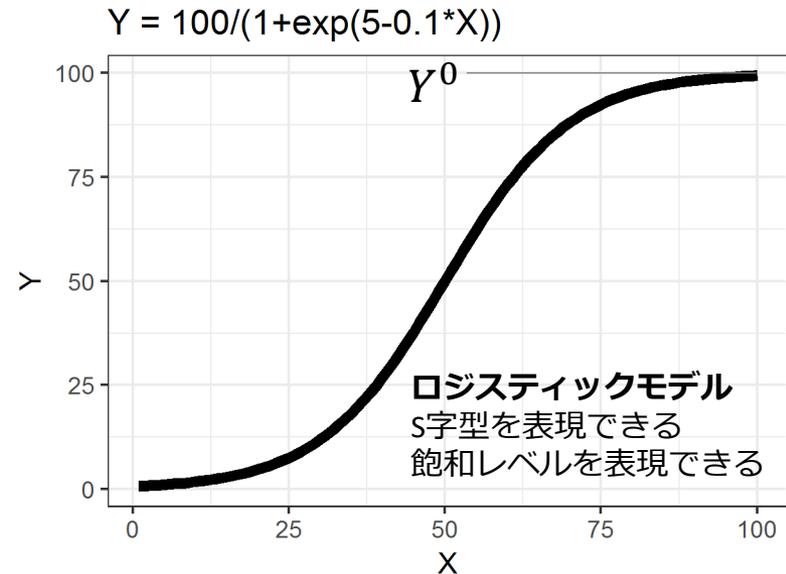
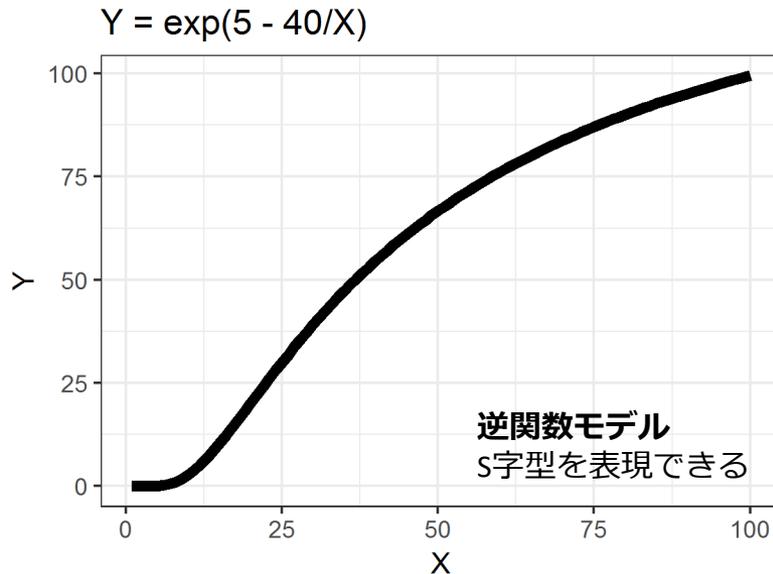
Code 4

$$\log(Y_i) = \beta_1 + \beta_2 \frac{1}{X_{2i}} + \dots$$

- ロジスティックモデル

$$\log\left(\frac{Y_i}{Y^0 - Y_i}\right) = \beta_1 + \beta_2 X_{2i} + \dots$$

その他、さまざまな関数形が提案されている



伝統的な回帰分析の枠組みから離れ、より柔軟な関数形を用いる

- ノンパラメトリック・モデル
- セミパラメトリック・モデル
- ニューラル・ネットワーク
- ...

実用性は高くない

- 大きなデータサイズを必要とする
- 売上とマーケティング・ミックス変数との関係が明示的でなくなる

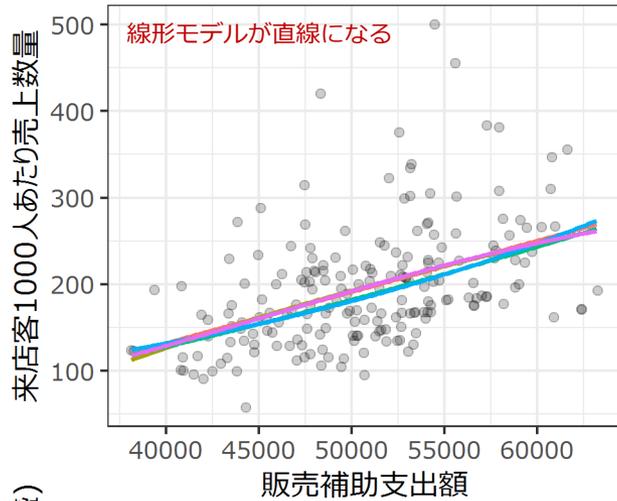
3-6. 推定方法

- 変数変換によって線形関数に変換し、最小二乗推定
 - 前項で紹介したモデルは、ロジスティックモデルを除き、変数変換によって線形関数に変換できる
 - 注意：変数変換によってモデルの意味が変わる (後述)
- 非線形関数のまま推定
 - 非線形最小二乗法 (NLS) が用いられることが多い
 - 一般に、NLS推定量は不偏性を持たない
 - 推定にはそれなりの修行と試行錯誤が必要

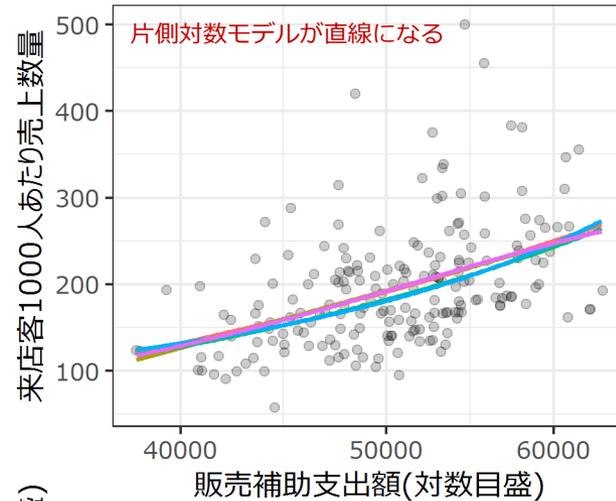
店舗販売データ

ブランド1について、販売補助支出額を説明変数とすると...

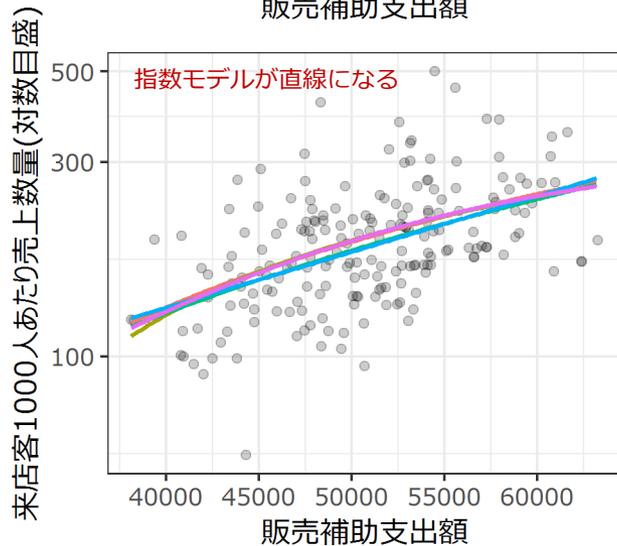
Code 5



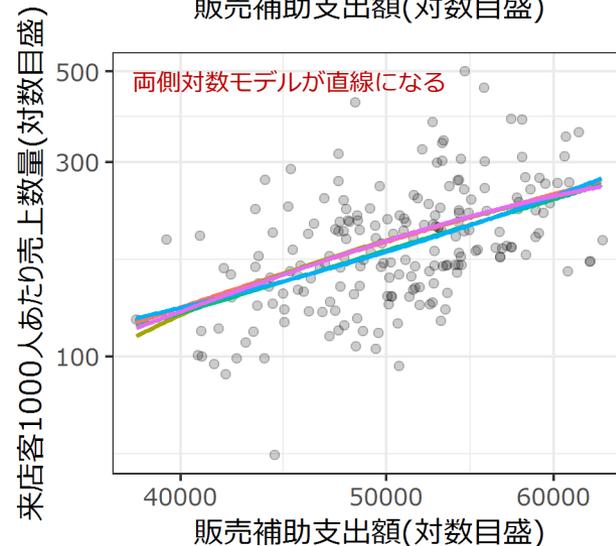
- 線形
- 片側対数
- 両側対数
- 指数
- ロジスティック



- 線形
- 片側対数
- 両側対数
- 指数
- ロジスティック



- 線形
- 片側対数
- 両側対数
- 指数
- ロジスティック



- 線形
- 片側対数
- 両側対数
- 指数
- ロジスティック

3-7. 変数変換に関する注意点

変数変換によって、モデルの意味が変わる！

例) 乗法モデル

$$Y_i = \beta_1 \times X_{2i}^{\beta_2} \times U_i$$

を、両側対数モデル

$$\log Y_i = \beta'_1 + \beta_2 \log X_{2i} + U'_i$$

に変換して最小二乗推定したとしよう。

このとき、

- 回帰分析の古典的仮定である [3]分散均一性, [5]正規性が、**変換後の攪乱項** $U'_i = \log(U_i)$ に対して求められる
 - もとの攪乱項 U_i についての仮定ではない
- モデルから得られる予測 $\widehat{\log(Y_i)} = \hat{\beta}'_1 + \hat{\beta}_2 \log(X_{2i})$ は、「 X_{2i} のもとでの $\log Y_i$ の条件付き期待値」 $E(\log Y_i | X_{2i})$ の推測値である
 - それを指数変換しても、それは「 X_{2i} のもとでの Y_i の条件付き期待値」 $E(Y_i | X_{2i})$ の推測値ではない
 - つまり、このモデルは**もはや Y_i を“あてる”ことを目指していない**

3-8. 関数形をどのように選ぶか？

- 一般的知識
 - 例) 広告は規模に対して収穫逓減となることが多い
- 当該カテゴリについての知識・経験
- データの観察
- モデルの適合度比較 (HPS pp.217-228)
- 交互作用についての検討 (後述)
- 残差分布の検討 (後述)
- 推定しやすさ
- 解釈しやすさ

店舗販売データ

Code 6

ブランド1の来店客数1000人あたり販売数量の対数 $\log Y_i$ を目的変数に設定。

以下の4つのモデルを推定した。 X_{2i} は販売補助支出額, X_{3i} は宣材送付個数を表す。

モデル1: $\log Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + U_i$ (指数モデル)

モデル2: $\log Y_i = \beta_1 + \beta_2 \log X_{2i} + \beta_3 X_{3i} + U_i$

モデル3: $\log Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 \log X_{3i} + U_i$

モデル4: $\log Y_i = \beta_1 + \beta_2 \log X_{2i} + \beta_3 \log X_{3i} + U_i$ (両側対数モデル)

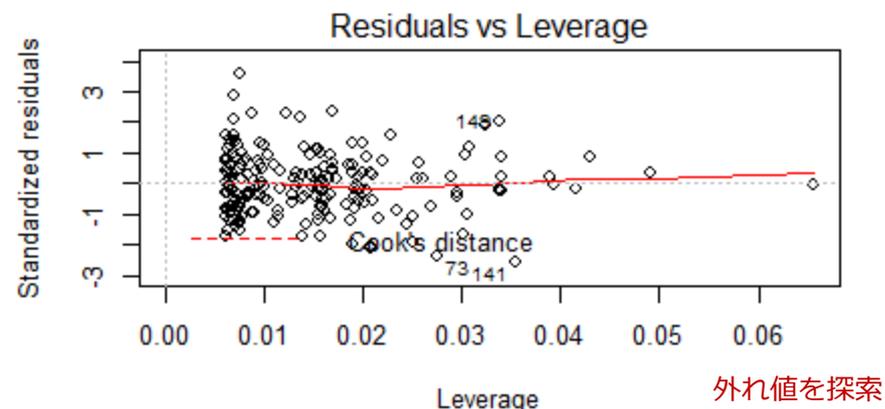
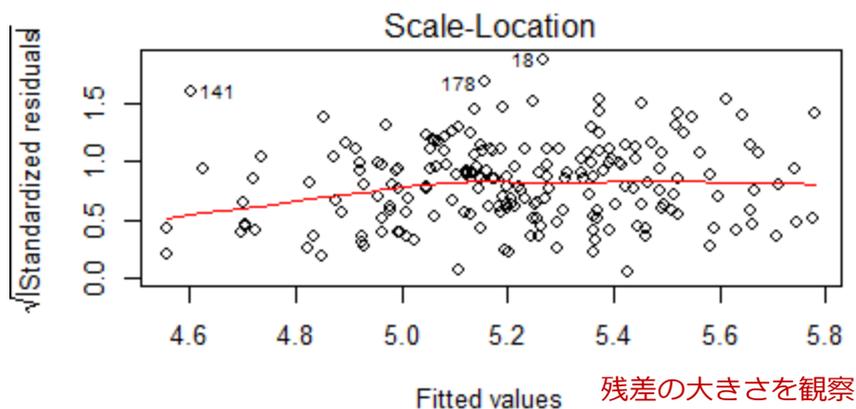
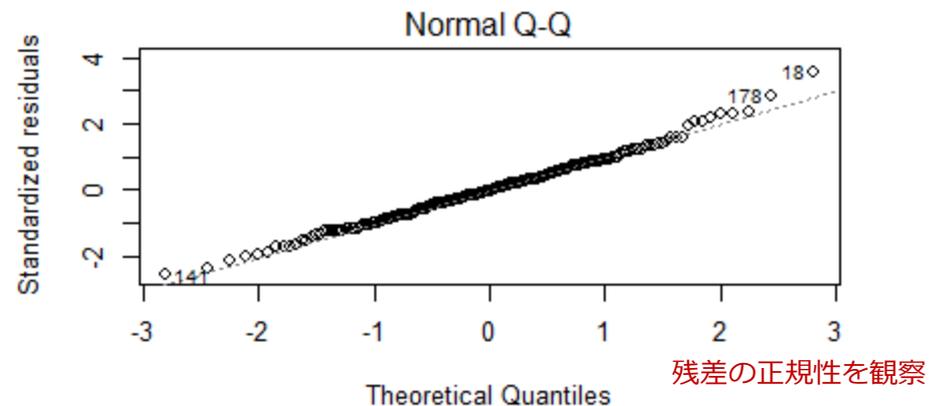
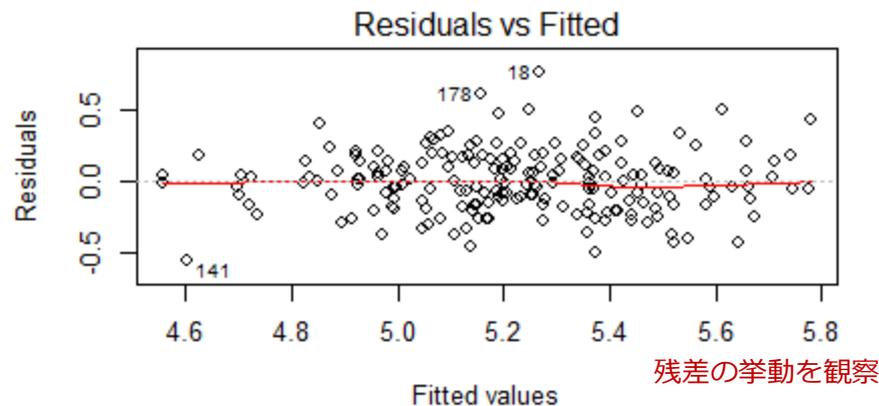
推定結果を示す。

	モデル1		モデル2		モデル3		モデル4	
	推定値	標準誤差	推定値	標準誤差	推定値	標準誤差	推定値	標準誤差
切片	2.88	0.16	-7.87	1.60	1.93	0.20	-8.85	1.59
販売補助支出額(10万円)	2.17	0.30			2.17	0.30		
販売補助支出額(対数)			1.09	0.15			1.10	0.15
宣材送付個数	0.19	0.02	0.19	0.02				
宣材送付個数(対数)					1.18	0.09	1.18	0.09
AIC	-37.23		-37.93		-36.66		-37.41	
RSQ	0.582		0.583		0.581		0.582	

データへのあてはまり
という観点からは、
モデル2がよい

モデル2の回帰診断図 (cf. 金, 2017 pp.139-141)

Code 6



4. 関数形 II. シェアモデル

回帰モデル

$$Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_K X_{Ki} + U_i$$



市場反応の特性

B) 市場反応変数として、売上を用いる場合と**シェア**を用いる場合がある

本節では、目的変数がシェアである場合の関数形について検討する。

以下では、店舗 i ($= 1, \dots, N$) におけるブランド b ($= 1, \dots, B$) の売上数量シェアを S_{bi} とする。

店舗販売データ

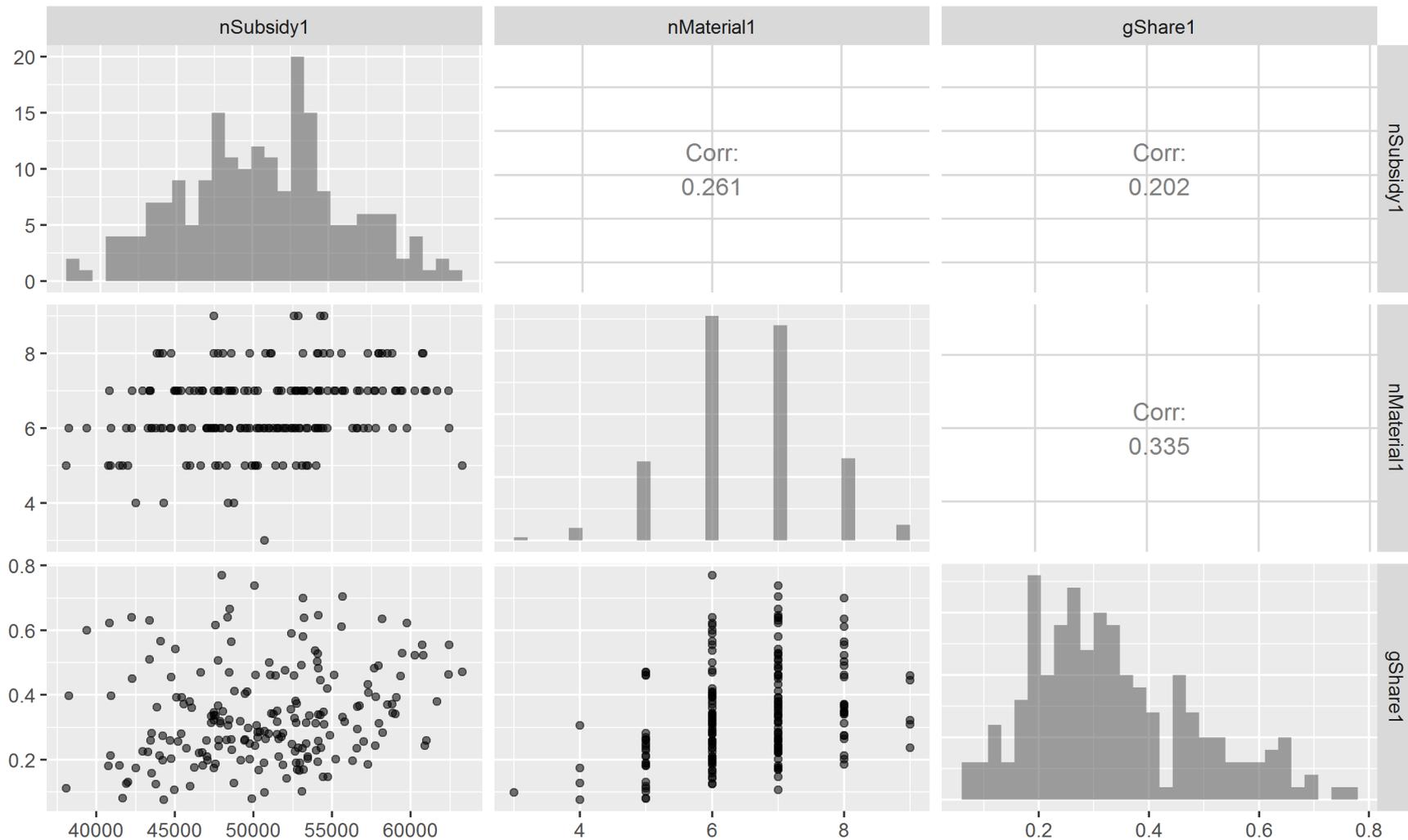
ブランド1

Code 7

販売補助支出額

宣材送付個数

売上数量シェア



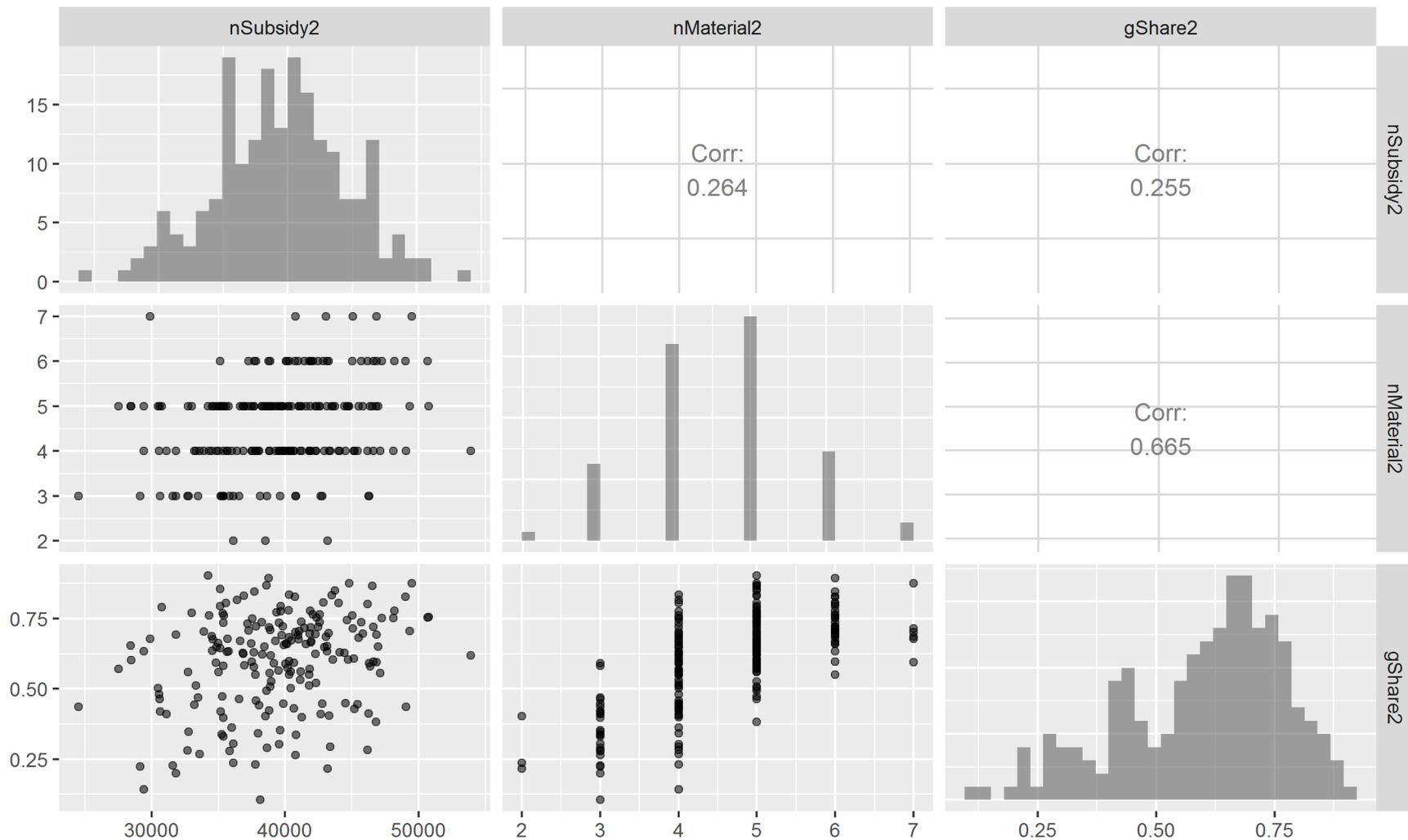
ブランド2

Code 7

販売補助支出額

宣材送付回数

売上数量シェア



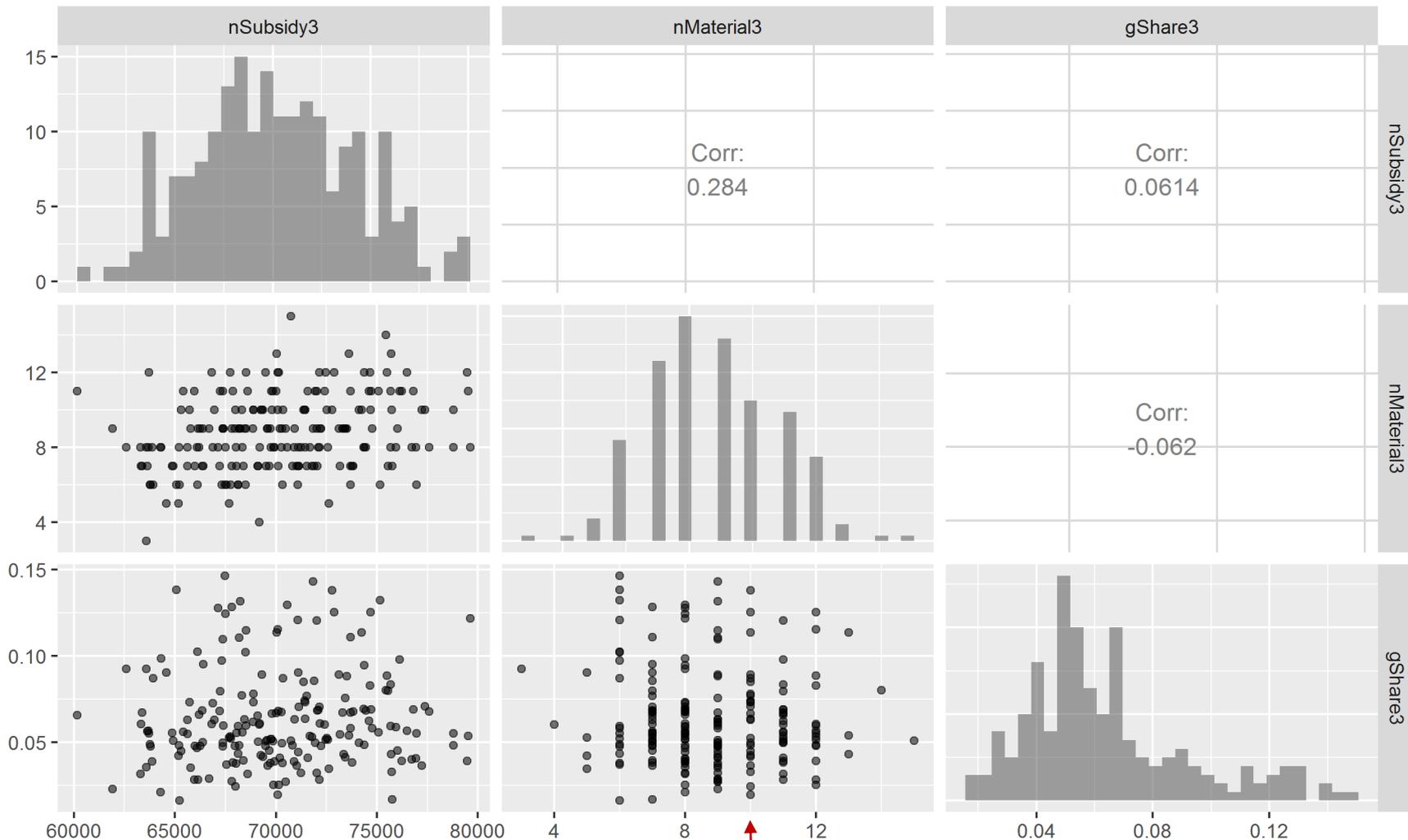
ブランド3

Code 7

販売補助支出額

宣材送付個数

売上数量シェア



4-1. シェアへの2つのアプローチ

シェアは次の特徴を持つ：

- 限界制約
 - シェアは0から1の間
- 合計制約
 - シェアのブランド間合計は1

いずれも、回帰モデルの想定に反している



アプローチ：

- 合計制約を無視した関数形
- 合計制約を取り入れた関数形

4-2. 合計制約を無視した関数形

Code 3

- シェア S_{bi} についての線形関数

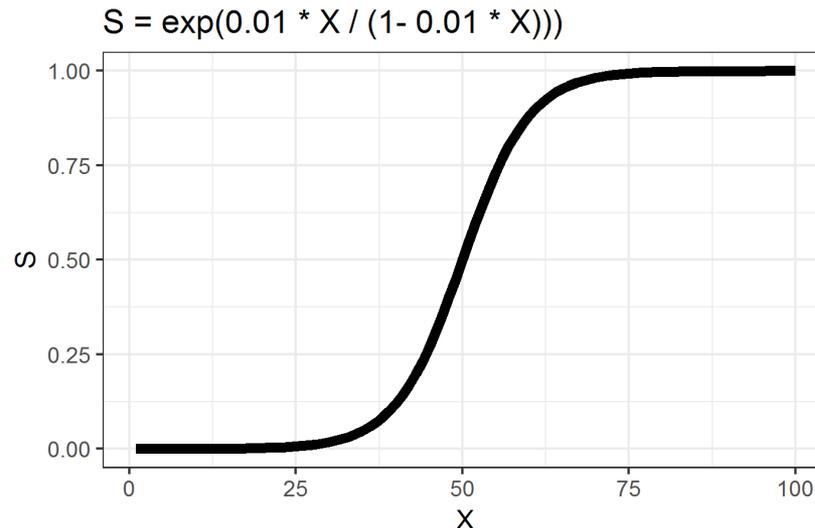
$$S_{bi} = \beta_{b1} + \beta_{2b}X_{2bi} + \dots$$

- シェアは本来 0以上1以下なのに、 S_{bi} は 0 を下回ったり 1を上回ったりする (限界制約の無視)

- シェアの**対数オッズ (ロジット)** についての線形関数

$$\log\left(\frac{S_{bi}}{1 - S_{bi}}\right) = \beta_{b1} + \beta_{2b}X_{2bi} + \dots$$

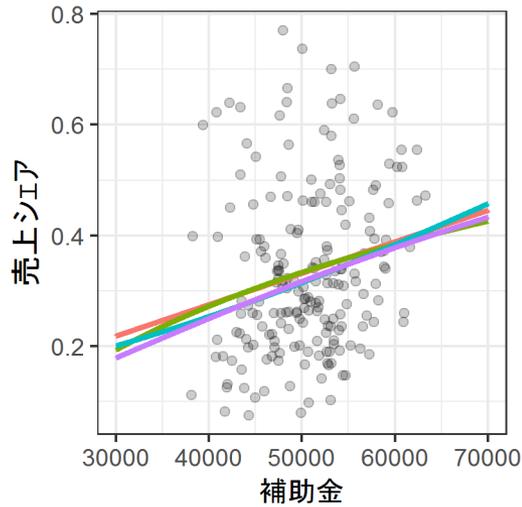
- S_{bi} は 0以上1以下となる
- 前項の「ロジスティック・モデル」の飽和レベルを1にした関数形



店舗販売データ

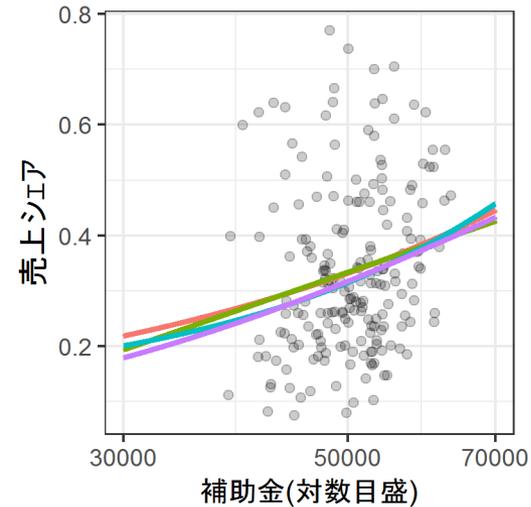
ブランド1について、販売補助支出額を説明変数とすると...

Code 8



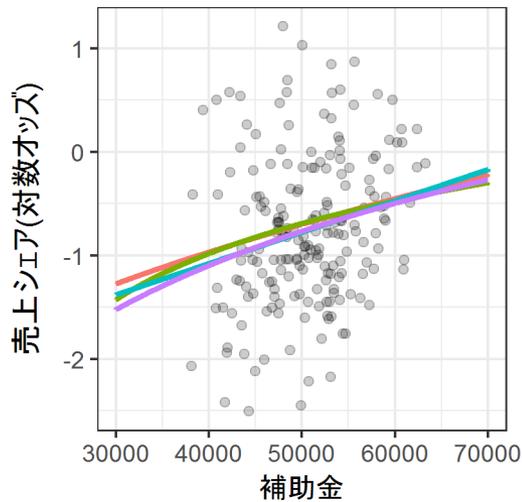
fModel

- 線形
- 線形, Xを対数変換
- 対数オッズ
- 対数オッズ, Xを対数変換



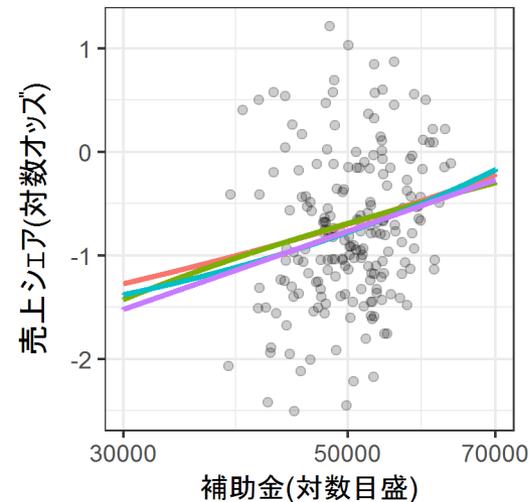
fModel

- 線形
- 線形, Xを対数変換
- 対数オッズ
- 対数オッズ, Xを対数変換



fModel

- 線形
- 線形, Xを対数変換
- 対数オッズ
- 対数オッズ, Xを対数変換



fModel

- 線形
- 線形, Xを対数変換
- 対数オッズ
- 対数オッズ, Xを対数変換

吸引力モデル (attraction model)

ブランド b の店舗 i における吸引力 (魅力) を $A_{bi} (> 0)$ とし、

$$S_{bi} = \frac{A_{bi}}{\sum_{j=1}^B A_{ji}}$$

とする。

さらに、 A_{bi} についてなんらかのモデルを想定する。

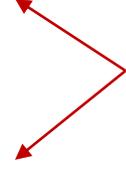
- MCIモデル (乗法競合交相互作用モデル)

$$\log(A_{bi}) = \beta_{1b} + \beta_2 \log(X_{2bi}) + \dots$$

- MNLモデル (多項ロジットモデル)

$$\log(A_{bi}) = \beta_{1b} + \beta_2 X_{2bi} + \dots$$

- ...



説明を簡単にするため、ここでは、偏回帰係数 β_2 がブランド間で同じだと仮定する(単一効果モデル)

- 吸引力モデルを線形モデルに変換する (cf. 里村, 2010 pp.58-63) :

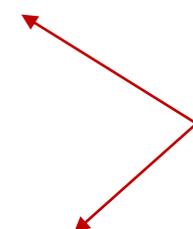
吸引力モデルの両側を対数にして

$$\log(S_{bi}) = \log(A_{bi}) - \log\left(\sum_{j=1}^B A_{ji}\right)$$

全ブランドについて合計し、ブランド数で割って

$$\frac{1}{B} \sum_{j=1}^B \log(S_{ji}) = \frac{1}{B} \sum_{j=1}^B \log(A_{ji}) - \log\left(\sum_{j=1}^B A_{ji}\right)$$

この項が消えるのがミソ



もとの式から引くと

$$\log(S_{bi}) - \frac{1}{B} \sum_{j=1}^B \log(S_{ji}) = \log(A_{bi}) - \frac{1}{B} \sum_{j=1}^B \log(A_{ji})$$

シェアの対数を求め、その平均を引いている。
これを**対数中心化**という

対数中心化したシェアを $S_{bi}^* = \log(S_{bi}) - \frac{1}{B} \sum_{j=1}^B \log(S_{ji})$ と略記する。

たとえば、 $\log(A_{bi}) = \beta_{1b} + \beta_2 \log(X_{2bi})$ とモデル化している場合(MCIモデル)、

$$\begin{aligned}
 S_{bi}^* &= \beta_{1b} + \beta_2 \log(X_{2bi}) - \frac{1}{B} \sum_j^B (\beta_{1j} + \beta_2 \log(X_{2ji})) \\
 &= \underbrace{\left(\beta_{1b} - \frac{1}{B} \sum_j^B \beta_{1j} \right)}_{\text{ひとつのパラメータと捉える}} + \beta_2 \underbrace{\left(\log(X_{2bi}) - \frac{1}{B} \sum_j^B \log(X_{2ji}) \right)}_{X_{2bi} \text{を対数中心化している}}
 \end{aligned}$$

従って、対数中心化後の説明変数を X_{2bi}^* とすれば

$$S_{bi}^* = \beta_{1b}^* + \beta_2 X_{2bi}^*$$

という線形関数となる。

MNLモデルも同様に線形関数に書き換えられる。

■ 吸引力モデルを推定する：

線形モデルに書き換え、攪乱項を追加して、

- 最小二乗推定法 (OLS)
- 一般化最小二乗推定法 (GLS)
 - ある店舗において、異なるブランドの攪乱項の間に相関があるはず
 - 片方が大きいと片方が小さくなるから
 - これは回帰モデルの古典的仮定のうち [4]独立性に反する
 - このとき、GLSはOLSより優れている (cf. 蓑谷 pp.252-263)

店舗販売データ

Code 9

次の吸引力モデルを想定した。店舗 i におけるブランド b のシェアを S_{bi} , 販売補助支出額を X_{2bi} , 宣材送付個数を X_{3bi} として、

$$S_{bi} = \frac{A_{bi}}{\sum_{j=1}^B A_{ji}}, \quad \log(A_{bi}) = \beta_{1b} + \beta_{2b} \log(X_{2bi}) + \beta_{3b} X_{3bi} + U_{bi}$$

$\sum \beta_{1b} = 0$ として、パラメータを最小二乗推定した。推定値を示す。

	推定値	推定量の 標準誤差
β_{11} : ブランド1の切片	-4.15	2.42
β_{12} : ブランド2の切片	3.43	2.08
β_{13} : ブランド3の切片	0.72	3.33
β_{21} : ブランド1の販売補助支出額の係数	1.01	0.23
β_{22} : ブランド2の販売補助支出額の係数	0.30	0.18
β_{23} : ブランド3の販売補助支出額の係数	0.53	0.43
β_{31} : ブランド1の宣材送付個数の係数	0.21	0.02
β_{32} : ブランド2の宣材送付個数の係数	0.49	0.02
β_{33} : ブランド3の宣材送付個数の係数	-0.01	0.01

ブランド3では宣材は効かない

吸引力モデルの長所:

- シェアが持つ論理的特性(限界制約, 合計制約)を正しく捉えている
- マーケティング活動の効果をブランド間で比較する際に便利
- 予測の観点からみて優れている？ ... 諸説ある

吸引力モデルの短所:

- ブランド数が多いと推定が大変
- 全ブランドで、マーケティング活動についての情報が同程度に揃っていないと、意義が乏しい
- ややこしい

5. 説明変数の欠落

回帰モデル

- [1] X_i は確率変数でないか
 U_i と統計的に独立
- [2] $E(U_i) = 0$



市場反応の特性

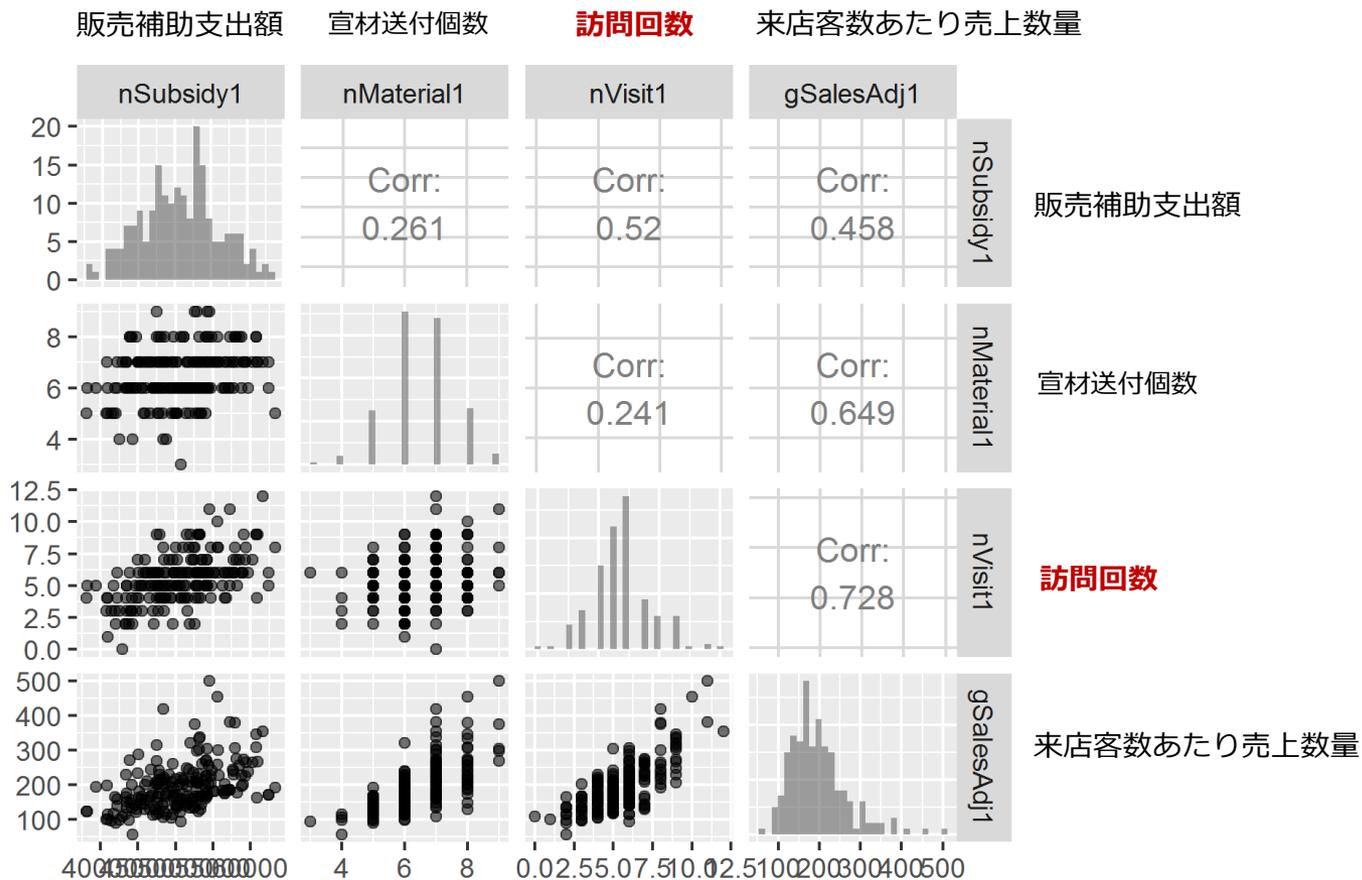
- D) 市場反応に影響する変数は多様であり、
データを手に入れない変数も多い

本節では、必要な説明変数がモデルから欠落していることがモデルにあたる影響について検討する。

店舗販売データ

Code 10

ブランド1について、新たに「小売店への販売員の訪問回数」を入手することができた。



3節で構築したモデルに訪問回数 X_{4i} を追加してみよう。

Code 11

追加前モデル: $\log Y_i = \beta_1 + \beta_2 \log X_{2i} + \beta_3 X_{3i} + U_i$

追加後モデル: $\log Y_i = \beta_1 + \beta_2 \log X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + U_i$

推定結果を示す。追加後モデルのほうが、データへのあてはまりがはるかに良い。追加後モデルを採用すべきだろう。

これによって、マーケティング活動の効果に対する私たちの理解は大きく修正されることになる。

	追加前モデル		追加後モデル	
	推定値	標準誤差	推定値	標準誤差
切片	-7.87	1.60	0.98	1.27
販売補助金支出(対数)	1.09	0.15	0.24	0.12
宣材送付数	0.19	0.02	0.17	0.01
訪問回数			0.10	0.01
決定係数	0.58		0.80	
AIC	-37.93		-183.31	

... 実際には、このような素晴らしい出来事は滅多に起こらない。私たちは通常、上で言う「追加前モデル」の段階で満足することしかできない。

5-1. 説明変数の欠落がもたらす影響

- 必要な説明変数が欠落していることにより、注目している説明変数の効果の推定量はどのような影響を受けるのか？
 1. 標準誤差が大きくなる
 2. 偏りが生じる

5-2. 標準誤差の増大

■ おさらい(III章8節) : 回帰モデル

$$Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_K X_{Ki} + U_i, \quad i = 1, \dots, n$$

のパラメータ β_j ($j = 2, \dots, K$) について、推定量の分散は

$$\text{Var}(\hat{\beta}_j) = \frac{\sigma^2}{\sum (X_{ji} - \bar{X}_j)^2} \times \frac{1}{1 - R_j^2}$$

攪乱項 U_i の分散。必要な説明変数が欠けていれば、当然大きくなる

■ おさらい：重回帰モデル

$$Y = X\beta + U$$

のパラメータ推定量は

$$\hat{\beta} = (X'X)^{-1}X'Y$$

■ 問題：真のモデルが

$$Y = X_1\beta_1 + X_2\beta_2 + U$$

であるときに、誤って

$$Y = X_1\beta_1 + U$$

と想定したとき、推定量 $b_1 = (X_1'X_1)^{-1}X_1'Y$ と、真のパラメータ β_1 のあいだにはどのような関係があるか？



$A_1 = (X_1'X_1)^{-1}X_1'$ とする。

$$b_1 = A_1Y$$

$Y = X_1\beta_1 + X_2\beta_2 + U$ を代入して

$$\begin{aligned} b_1 &= A_1(X_1\beta_1 + X_2\beta_2 + U) \\ &= A_1X_1\beta_1 + A_1X_2\beta_2 + A_1U \end{aligned}$$

$A_1X_1 = (X_1'X_1)^{-1}X_1'X_1 = I$ だから

$$b_1 = \beta_1 + A_1X_2\beta_2 + A_1U$$

$E(U) = 0$ より

$$E(b_1) = \beta_1 + (X_1'X_1)^{-1}X_1'X_2\beta_2$$

推定したいと思っている
真の偏回帰係数

回帰モデル
 $X_2 = X_1\gamma + V$
の γ の推定量！

欠落している説明変数の
真の偏回帰係数

図で表現すると ... (cf. 浅野・中村, 2009, pp.107-110)

想定しているモデル



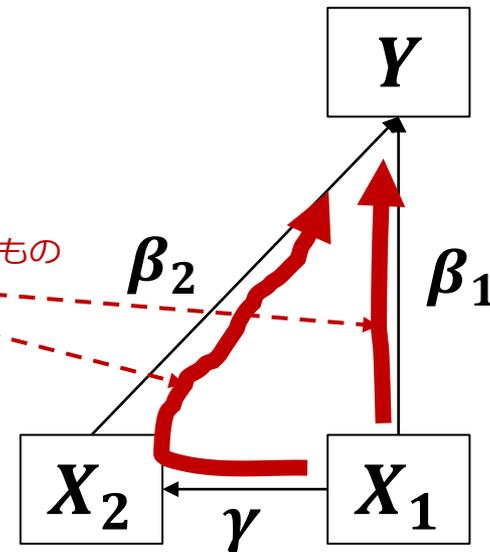
推定している
と思っているもの

b_1

本当に
推定しているもの

β_2

β_1



実線矢印は回帰モデルを表す。因果関係ではない

必要な説明変数が欠落していることにより、注目している説明変数の効果の推定量はどのような影響を受けるのか？

- 標準誤差が大きくなる (推定の信頼性が下がる)
 - その程度は、欠けている説明変数が目的変数の分散を説明する程度によって決まる
- **偏りを持った推定量になる**
 - その偏りの方向と大きさは、次の2つによって決まる：
 - **欠けている説明変数と、目的変数との関係**
 - **欠けている説明変数と、注目している説明変数の関係**
 - 逆にいうと... 欠けている変数が、注目している説明変数と相関を持っていない場合は、偏りは生じない

店舗販売データ

	追加前モデル	
	推定値	標準誤差
切片	-7.87	1.60
販売補助金支出(対数)	1.09	0.15
宣材送付数	0.19	0.02
決定係数	0.58	
AIC	-37.93	

上の推定結果をみると、私たちは次のことを考えなければならない。

- 売上数量に影響する変数として、ほかにどんなものがあるか？
- その変数は、販売補助金支出や宣材送付数と相関しているか？
 - その変数が、販売補助金と正の相関を持ち、売上とも正の相関を持つならば、上の推定値 (+1.09) は割り増しされていると推測できる

6. 多重共線性

回帰モデル

[6] X の列ベクトルは一次独立。
列数は行数より小さい



市場反応の特性

E) 広告・プロモーションの効果は、その
内容や媒体によっても異なるかもしれ
ない

通常、マーケティング・ミックス変数はなんらかの集計変数である(例, 広告支出)。

市場反応モデリングにおいては、マーケティング活動の効果を、できるだけ細分化したマーケティング・ミックス変数について調べたいという強いニーズがある(例, 媒体別広告支出)。

いっぽう回帰モデルにおいて、説明変数の細分化は深刻な問題 (多重共線性) を引き起こすことが多い。

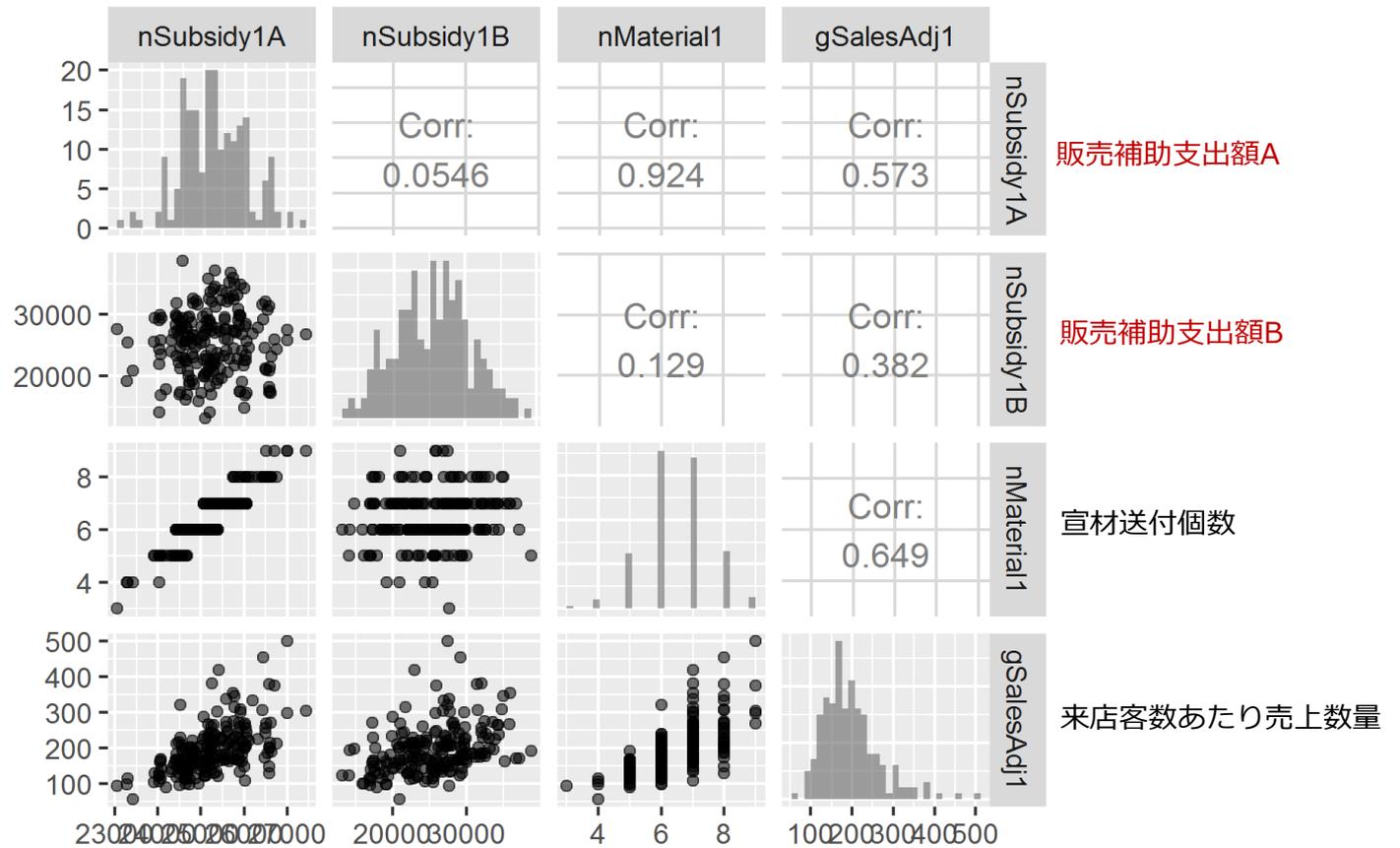
本節では、多重共線性の発見と対処について述べる。

店舗販売データ

Code 12

実は、ブランド1の販売補助支出額には二つの細目(A, B)がある。

販売補助支出額A 販売補助支出額B 宣材送付個数 来店客数あたり売上数量



そこで、説明変数としてタイプ別の販売補助支出額を投入し、モデルを推定してみると...

推定量の標準誤差が非常に大きくなってしまった。

	投入前			投入後		
	推定値	標準誤差	VIF(後述)	推定値	標準誤差	VIF(後述)
切片	-7.87	1.60		0.90	13.64	
販売補助支出額(対数)	1.09	0.15	1.07			
販売補助支出額A(対数)				-0.24	1.35	7.17
販売補助支出額B(対数)				0.53	0.07	1.05
宣材送付回数	0.19	0.02	1.07	0.21	0.04	7.27
決定係数	0.583			0.584		
AIC	-37.93			-36.50		

「販売支出補助額Aが大きいと売上が低い」という
推定値が得られているが、標準誤差が大きいので信頼できない

6-1. 多重共線性 (multicollinearity; マルチコ) とは

説明変数の間に強い線形的な関係があること

- X の列ベクトルの間の一次独立性が失われかけていること

多重共線性があるとなにが起きるか

- **推定量の信頼性が下がる**
- 逆にいうと... 推定量の性質(不偏性, 一致性, 有効性)は損なわれていない

おさらい (III章8節)

「 X_{ji} を目的変数、他のすべての説明変数を説明変数とした回帰モデル」の決定係数を R_j^2 とすると、

$$\text{Var}(\hat{\beta}_j) = \frac{\sigma^2}{\sum (X_{ji} - \bar{X}_j)^2} \times \frac{1}{1 - R_j^2}$$

VIF (variance inflation factor)
多重共線性によって生じた分散増大を表す

■ なぜ推定量の信頼性が下がるのか？

- モデル推定に必要なだけの情報が、データに含まれていないから

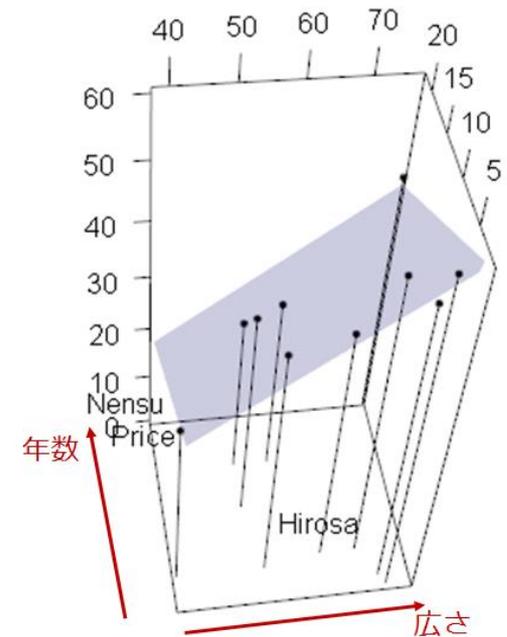
直観的にいうと...

- III章の中古マンションの事例を思い出そう
- 重回帰式

$$(\text{価格}) = (\text{切片}) + \beta_1 (\text{広さ}) + \beta_2 (\text{年数}) + (\text{攪乱項})$$

は、広さ x 年数 x 価格の3次元空間における平面となる

- 最小二乗法によるパラメータ推定とは、平面と点との鉛直方向の距離の二乗の合計が最小になるように、うまく平面を置くこと
- もし広さと築年数のあいだに強い相関があったら、この平面は不安定になるだろう



■ 多重共線性についてのよくある誤解

「多重共線性が生じるのは、説明変数間の単相関が高いときだ」 → ×

- 説明変数間の単相関では判断できない
- 例) X_2 と X_3 の相関は低く、 X_4 が $2X_2 + 3X_3$ に近いとしよう。 X_4 と X_2 の相関も、 X_4 と X_3 との相関もそれほど高くないかもしれないが、多重共線性は生じる

「多重共線性が生じるのは、説明変数間に因果関係があるときだ」 → ×

- 説明変数の間に強い線形的な関係があるかどうかは問題。関係がある理由は問題でない
- 説明変数間に因果関係があるとき、偏回帰係数の解釈はとても難しくなるが、それは多重共線性とは別の問題

「多重共線性とは、偏回帰係数の符号が相関係数の符号の逆になる現象だ」 → ×

- 推定の信頼性が下がった結果、偏回帰係数の推定値がどうなるかはわからない。符号は逆になるかもしれないし、ならないかもしれない
- 相関係数の符号と、真の偏回帰係数の符号が逆であることは全く珍しくない

「多重共線性は標本サイズとは無関係」 → ×

- データが十分に大きければ、説明変数間に線形的な関係があっても推定量は不安定になりにくい ($Var(\hat{\beta}_j)$ の式を参照)

VIF (variance inflation factor)

- 多重共線性によって生じた、パラメータ推定量の分散の増大

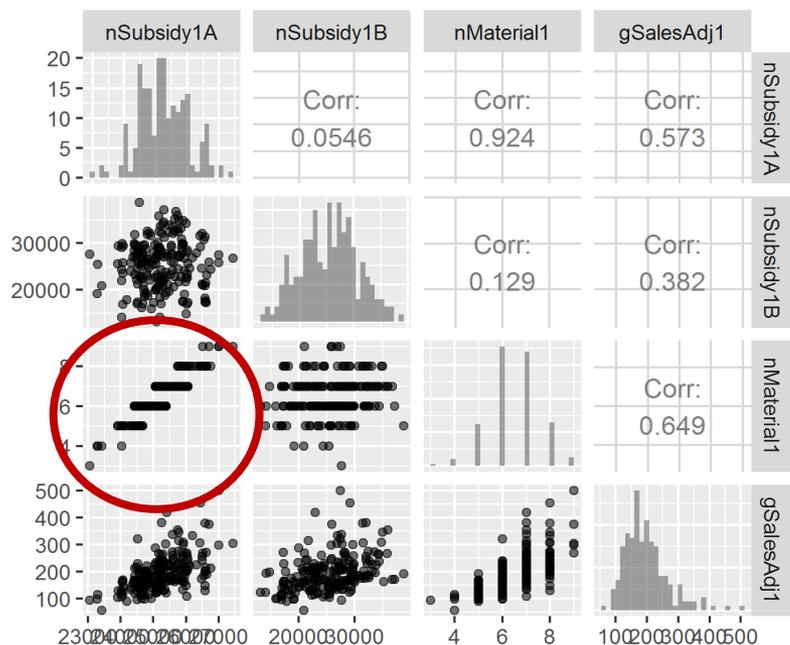
$$\text{Var}(\hat{\beta}_j) = \frac{\sigma^2}{\sum(X_{ji} - \bar{X}_j)^2} \times \frac{1}{1 - R_j^2} \leftarrow \text{VIF}$$

VIFについてのよくある誤解

- 「VIFが5以上(or 10以上)」のとき、多重共線性が起きている」→×
 - 5 or 10という値には根拠がない
- 「VIFが大きい時、推定量の信頼性は低い」→×
 - 推定量の信頼性を表現しているのは $\text{Var}(\hat{\beta}_j)$
 - 標本サイズが大きければ、VIFが高くてても $\text{Var}(\hat{\beta}_j)$ は小さいかもしれない
- 「VIFが高い変数のあいだで一次独立性が損なわれている」→×
 - VIFはその推定量の信頼性が多重共線性のせいで失われている程度を表す
 - 多重共線性がどの変数の間で生じているのか、VIFだけではわからないこともある

店舗販売データ

データをよく観察すると、販売補助支出額Aと宣材送付個数の間には高い相関がある。それにより、推定量の信頼性が損なわれているらしい



	投入後		
	推定値	標準誤差	VIF(後述)
切片	0.90	13.64	
販売補助支出額(対数)			
販売補助支出額A(対数)	-0.24	1.35	7.17
販売補助支出額B(対数)	0.53	0.07	1.05
宣材送付個数	0.21	0.04	7.27
決定係数	0.584		
AIC	-36.50		

6-3. 多重共線性への対処

多重共線性は“病気”ではなく“症状”

- 真の“病気”とは ... 必要な情報がデータに含まれていないこと

アプローチ1. 病気を治す

- データを追加する → 多くの場合、非現実的
- 変数を削除する → × (5節を参照)
- **変数を併合する** (i.e. 説明変数の細分化は断念する)

アプローチ2. 症状を抑える (cf. Hastie et al., 2014 3章4-5節)

- 変数縮約 ... 主成分回帰, PLS回帰など
- 正則化 ... リッジ回帰など
- 問題点
 - パラメータの解釈がさらに難しくなる
 - 推定量の分散は小さくなるが、偏りが生じる

7. 交互作用

回帰モデル

$$Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_K X_{Ki} + U_i$$



市場反応の特性

c) 異なるマーケティング活動を同時に行うことにより、**シナジー**が生まれるかもしれない

本章では、異なるマーケティング活動のシナジーを捉える方法について検討する。

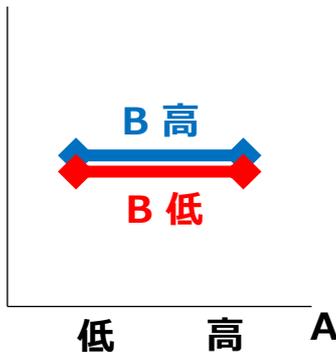
シナジーの例：

- 価格広告の増大によって、価格弾力性が高くなる (II章6節)
- TV CMと新聞広告を同時に出稿することで、それぞれの広告の効果を上回る高い効果が得られる (架空例)

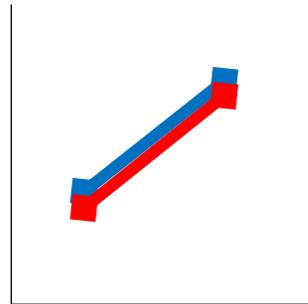
交互作用 (interaction)

- 統計学では、「目的変数に対するある説明変数の効果が、他の説明変数の値によって異なる」ことを指す用語

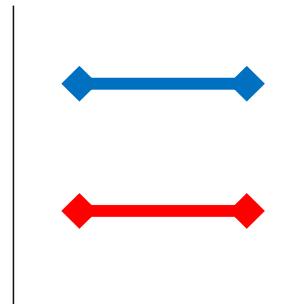
目的変数



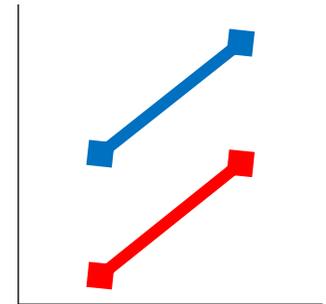
Aの主効果 なし
Bの主効果 なし
交互作用効果 なし



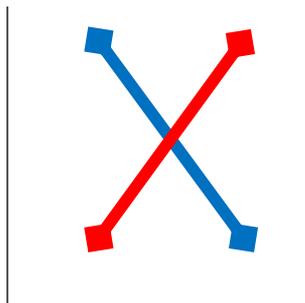
Aの主効果 **あり**
Bの主効果 なし
交互作用効果 なし



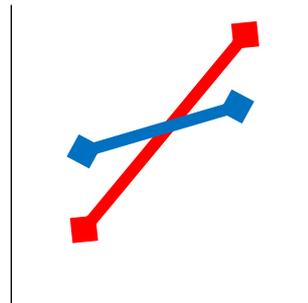
Aの主効果 なし
Bの主効果 **あり**
交互作用効果 なし



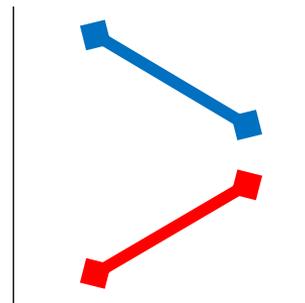
Aの主効果 **あり**
Bの主効果 **あり**
交互作用効果 なし



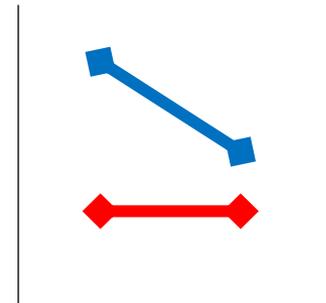
Aの主効果 なし
Bの主効果 なし
交互作用効果 **あり**



Aの主効果 **あり**
Bの主効果 なし
交互作用効果 **あり**



Aの主効果 なし
Bの主効果 **あり**
交互作用効果 **あり**

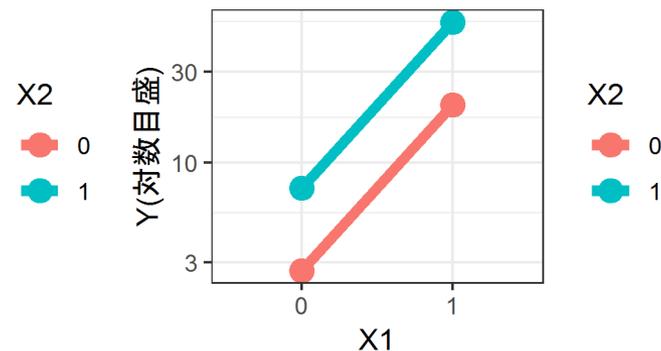
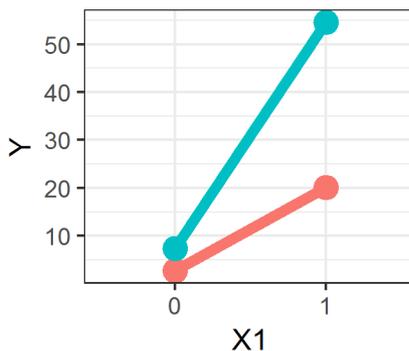


Aの主効果 **あり**
Bの主効果 **あり**
交互作用効果 **あり**

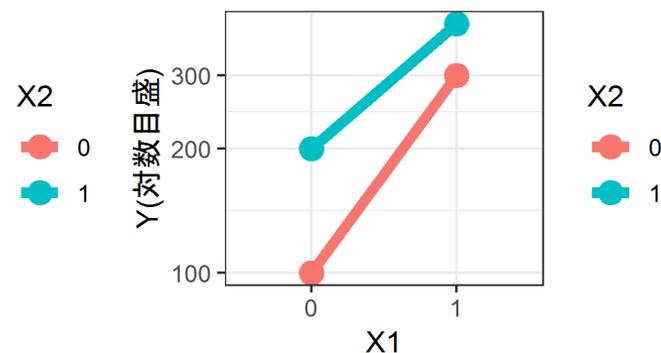
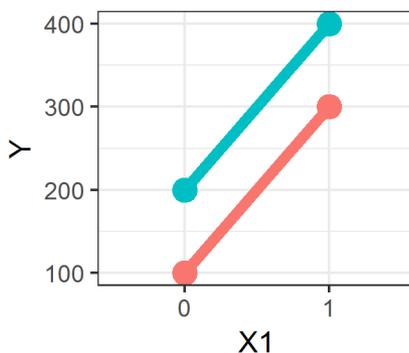
■ 交互作用の有無は関数形に依存する

- 関数形を変えると、交互作用が現れたり消えたりする
- つまり... マーケティング・ミックス変数間に“シナジー”が観察されているとしても、それは関数形の産物に過ぎず、モデルにおいて交互作用を表現する必要はないかもしれない

例1) 目的変数を対数変換することで、交互作用が消えるケース

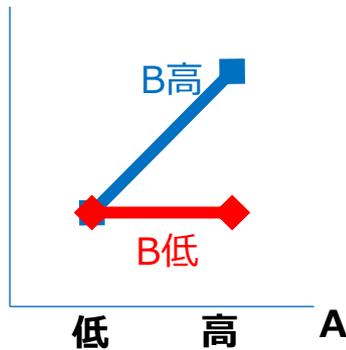


例2) 目的変数を対数変換することで、交互作用が出現するケース



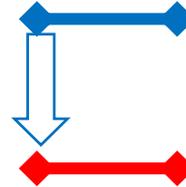
- AとBの交互作用が出現したとき、A, Bの効果の大きさをどのように捉えるか？
- いろいろな捉え方がある

現象

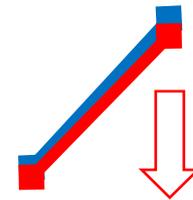


Aの主効果 **あり**
 Bの主効果 **あり**
 交互作用効果 **あり**

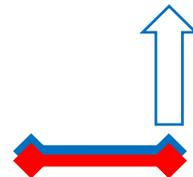
解釈



「Aに効果はない。ただし、A低&B高の場合に値が低くなる」



「Bに効果はない。ただし、A高&B低の場合に値が低くなる」

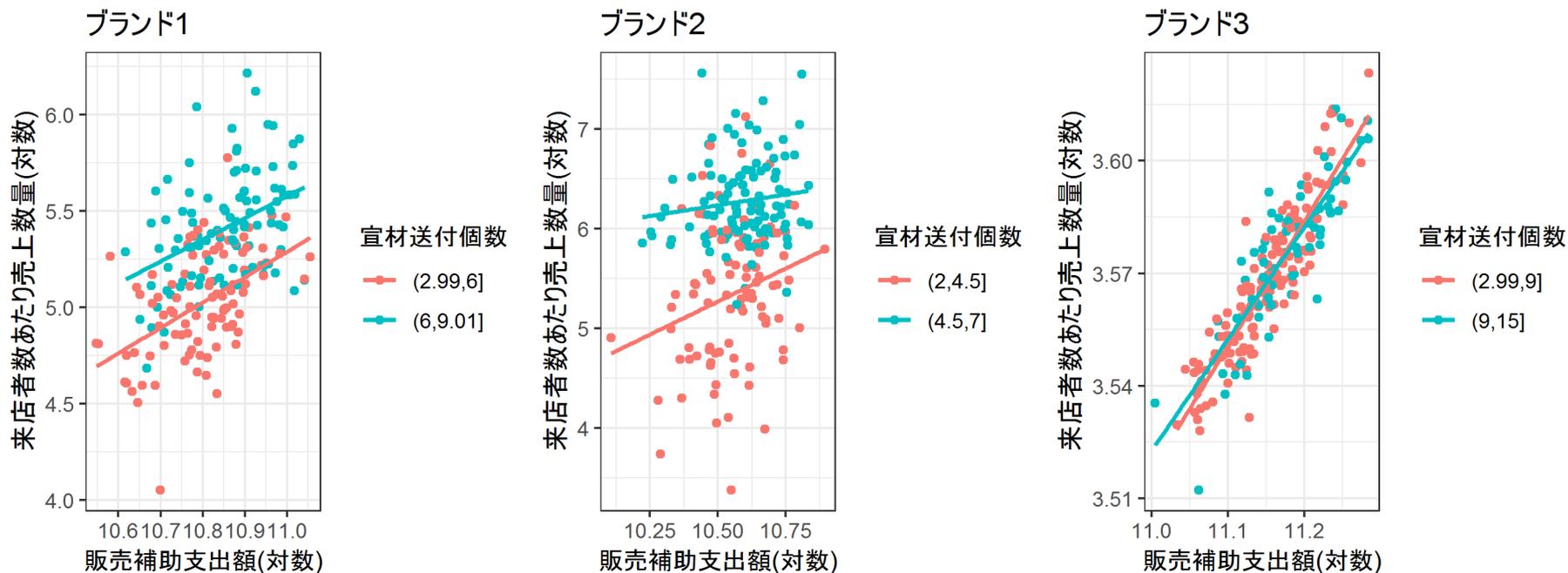


「AにもBも効果はない。ただし、A高&B高の場合に値が高くなる」

- 説明変数間の交互作用をどのように発見するか？
- データを探索し、交互作用がある組み合わせを発見する
 - 非現実的。説明変数の数が多い時、組み合わせの数は膨大になる
- 既存知識に基づき**仮説**を立て、データに基づいて検証する
 - 仮説：「どの変数間に交互作用があるかもしれないか」
 - 既存知識：
 - 市場反応についての一般的知識
 - 当該カテゴリについての知識・経験
 - 注意点：
 - マーケティング計画の視点ではなく、現象理解の視点に立つこと
 - 「交互作用が存在するかどうか」が問題
 - 「仮に存在したとして、それを活用したマーケティング計画が可能かどうか」はここでの問題ではない
 - 豊かな知識・経験を持つ人との密接な協働が必要だが...
 - たいていのは、「シナジー」「交互作用」という概念を正しく理解できていない
 - 深い聞き取りと、分析者による仮説構築が必要

ブランド2では、販売補助支出額と宣材送付個数の間に交互作用がある

- 宣材送付個数が多い場合、販売補助支出額の弾力性は低い



■ 説明変数間の交互作用をどのようにモデル化するか？

- 関数形を見直せるのなら、見直す
- モデルに交互作用を表現する項を投入する (cf. HPS p.107-109)
 - 説明変数(変換後) の積を用いることが多い
 - 線形モデル $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots$ に交互作用項 $X_{2i} X_{3i}$ を投入する場合は、あらかじめ説明変数 X_{2i}, X_{3i}, \dots をそれぞれ平均0に中心化しておくとい
 - 弊害：多重共線性
 - $X_2 X_3$ は、 X_2, X_3 のいずれとも高い相関を持つだろう
 - 交互作用という現象を捉える代償として、推定量の信頼性が損なわれる

店舗販売データ

Code 16

ブランド2について、来店客1000人あたり販売数量の対数 $\log Y_i$ を目的変数に設定。

以下の2つのモデルを推定した。 X_{2i} は販売補助支出額, X_{3i} は宣材送付個数を表す。

$$\text{モデル1: } \log Y_i = \beta_1 + \beta_2 \log X_{2i} + \beta_3 X_{3i} + U_i$$

$$\text{モデル2: } \log Y_i = \beta_1 + \beta_2 \log X_{2i} + \beta_3 X_{3i} + \beta_{23} \log(X_{2i})X_{3i} + U_i \quad (\text{交互作用項を追加})$$

推定結果を示す。

	交互作用項なし			交互作用項あり		
	推定値	標準誤差	VIF	推定値	標準誤差	VIF
切片	0.24	3.06		-26.83	12.89	
販売補助支出額(対数)	0.32	0.29	1.1	2.88	1.22	18.9
宣材送付個数	0.49	0.04	1.1	6.46	2.76	5836.3
交互作用				-0.56	0.26	6021.6
RSQ	0.487			0.499		
AIC	318.39			315.68		

交互作用項を追加することで、データに対するモデルのあてはまりは良くなった。

その反面、VIFは極端に増大し、推定量の標準誤差はきわめて大きくなっている。

8. 分散の不均一性・非正規性

回帰モデル

[3] $\text{Var}(U_i) = \sigma^2$

[5] U_i は正規分布に従う



データ

関数形を決めたうえで改めてデータを観察すると、説明変数で条件づけた目的変数の分散(残差の分散)が、不均一だったり、正規分布に従っていないだったりすることが少なくない。

これは回帰モデルの古典的仮定に反している。

本節では、残差の分散が不均一・非正規であるときの考慮点について、ごく簡単に紹介する。

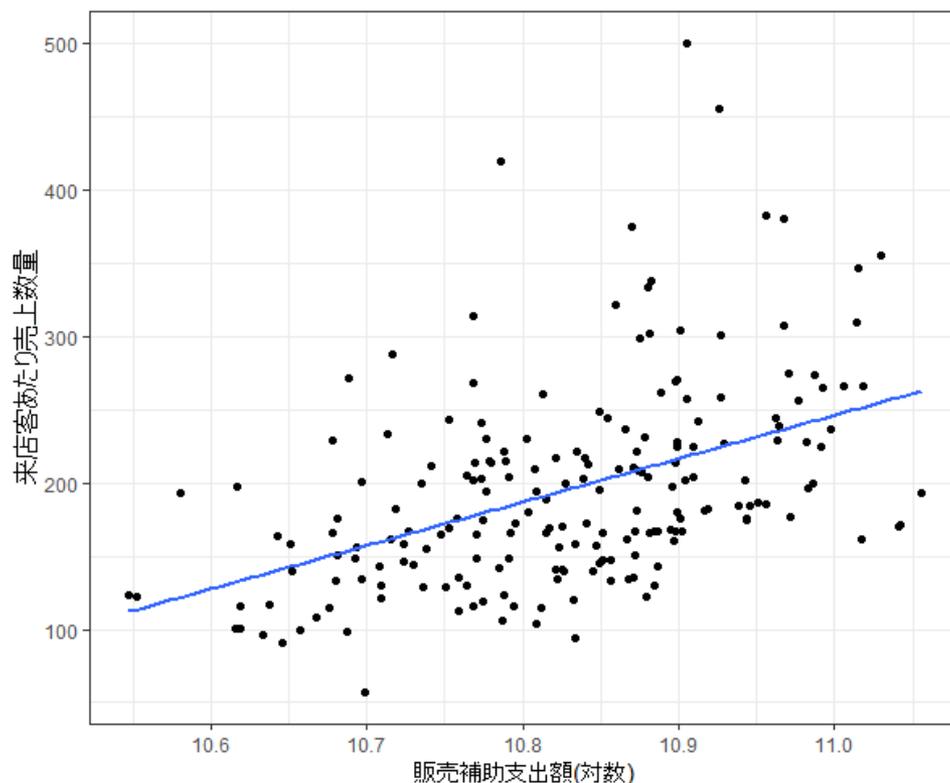
店舗販売データ

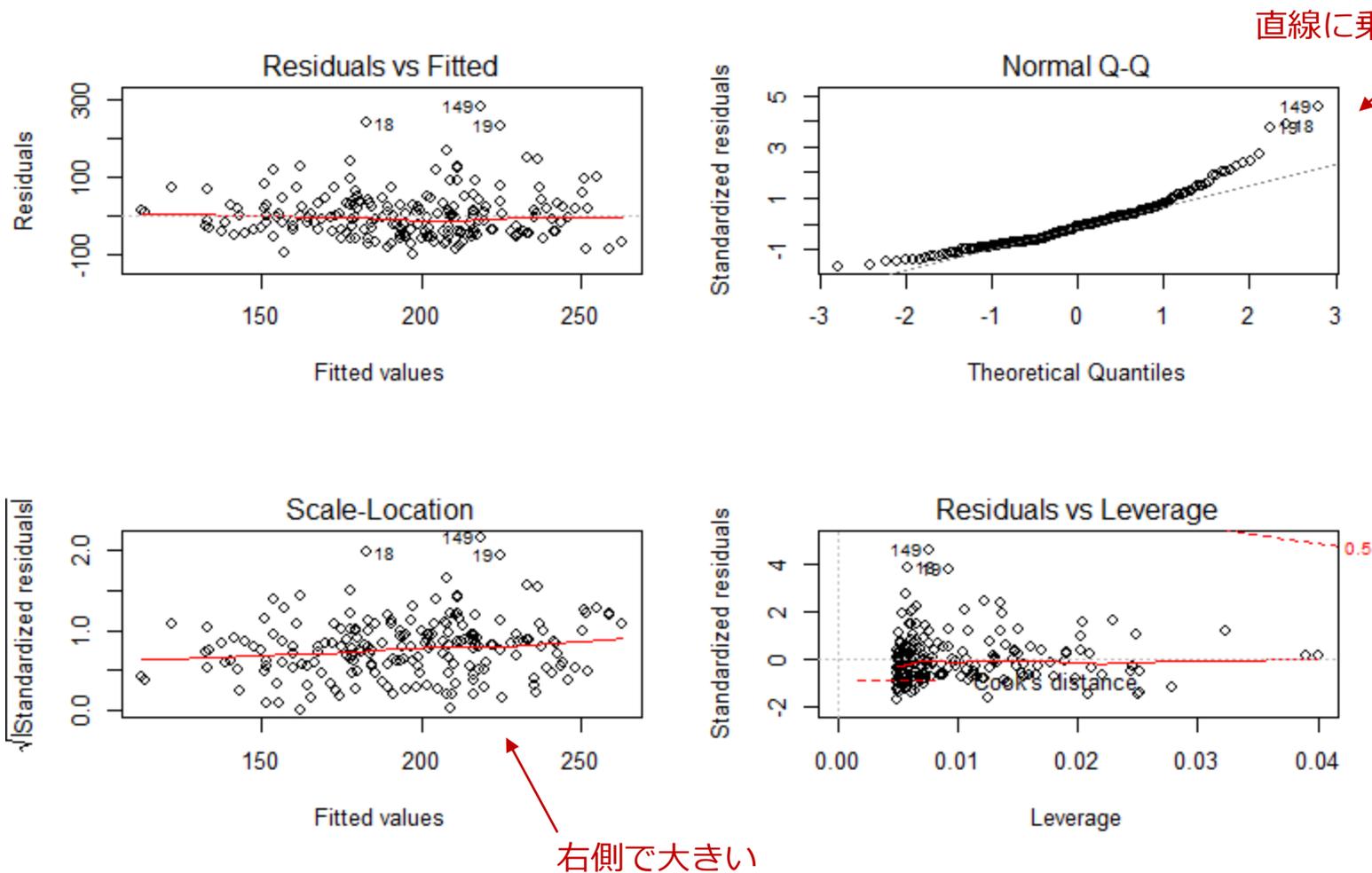
Code 17

ブランド1について、目的変数を来店客数あたり販売数量、説明変数を販売補助支出額(対数) としたとしよう (指数モデル)。

散布図を観察すると...

- 残差は右側のほうが大きいようにみえる
→ 均一に分布していない
- 残差の分布は、上側の裾が長いようにみえる
→ 正規分布に従っていない





攪乱項の均一性・正規性が疑われるから

- ポジティブにいうと：攪乱項が不均一・非正規であるならば、最小二乗法よりも良い推定方法があるかもしれないから

重回帰モデルの古典的な仮定

[1] X_i は確率変数でないか U_i と統計的に独立

[2] $E(U_i) = 0$

[3] $\text{Var}(U_i) = \sigma^2$ (均一性)

[4] $\text{Cov}(U_i, U_j) = 0, i \neq j$

[5] U_i は正規分布に従う (正規性)

[6] X の列ベクトルは一次独立で、列数は行数より小さい

推定量 $\hat{\beta}$ の性質：

- 仮定[1][2][6]のもとで、 β の**不偏推定量**
- 仮定[1][2][6] (と X に関するある仮定)の下で、 β の**一致推定量**
- 仮定[1][2][3][4][6]のもとで、 β の**最良線形不偏推定量 (BLUE)**
- 仮定[1][2][3][4][5][6]のもとで、 β の**有効推定量**
- 仮定[1][2][3][4][5][6]のもとで、 β の**最尤推定量**

- 原因を探る
 - マーケティング活動の効果や残差の分布が異なる、複数の群のデータが混ざっているのでは？
- 関数形を変える
 - 対数変換, 平方根変換, ボックス・コックス変換... (蓑谷, pp.228-252)
 - モデルの意味が変わってしまう
- “もっと良い推定量”を使う
 - 一般化最小二乗法 (GLS) (蓑谷, pp.252-256)
- 開き直す
 - 仮に攪乱項が不均一・非正規であっても、最小二乗推定量は依然として不偏推定量・一致推定量ではある

9. 異質性

回帰モデル

$$Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_K X_{Ki} + U_i$$



市場反応の特性

F) マーケティング活動と市場反応の関係には、**異質性**があるかもしれない

市場反応モデリングのためのデータは集計データである。その**集計主体**は、マーケティング意思決定の対象となる主体と異なることが多い。

意思決定の対象となる主体は、マーケティング活動への接触においても、また反応関数においても、**個体差**を持っている。これを**異質性** heterogeneityと呼ぶ。

本節では、異質性によって生じるバイアスについて検討する。

■ 主体集計バイアス

データの集計主体と、“本来の”主体が異なることによって生じるバイアス

“本来の”主体とは:

反応関数を持つと思われる主体

突き詰めていえば、マーケティング意思決定がなされるレベルの主体

- DM送付は顧客ごとに決定可能 → 顧客
- 店舗補助支出は店舗ごとに決定可能 → 店舗

例1) 顧客に対するDMの効果を知るために、月次売上とDM送付量との関係を調べる

例2) 店舗に対する活動の効果を知るために、チェーン別売上とチェーン別活動量との関係を調べる

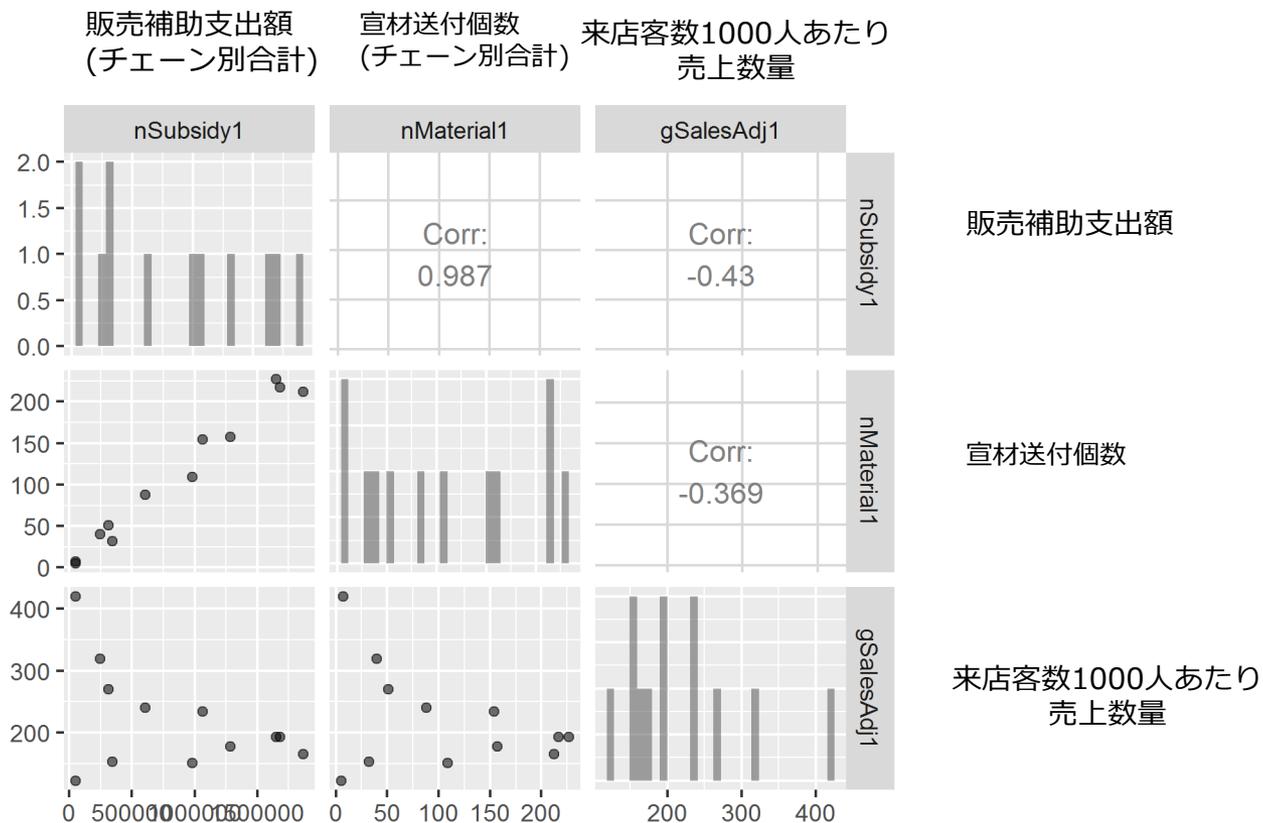


店舗販売データ

Code 18

200店舗は12個のチェーンに分かれている。

店舗レベルのデータではなく、チェーン・レベルのデータを分析してみよう。



店舗レベルのデータを用いて推定したモデル

$$\log Y_i = \beta_1 + \beta_2 \log X_{2i} + \beta_3 X_{3i} + U_i$$

を、チェーン・レベルの集計データにあてはめ、再度推定してみると、全く異なるパラメータ推定値が得られた。なぜだろう？

	店舗レベル		チェーン・レベル	
	推定値	標準誤差	推定値	標準誤差
切片	-7.87	1.60	6.66	2.40
販売補助支出額(対数)	1.09	0.15	-0.10	0.21
宣材送付個数	0.19	0.02	0.00	0.00
RSQ	0.583		0.094	

■ 主体集計バイアスはなぜ起きるか

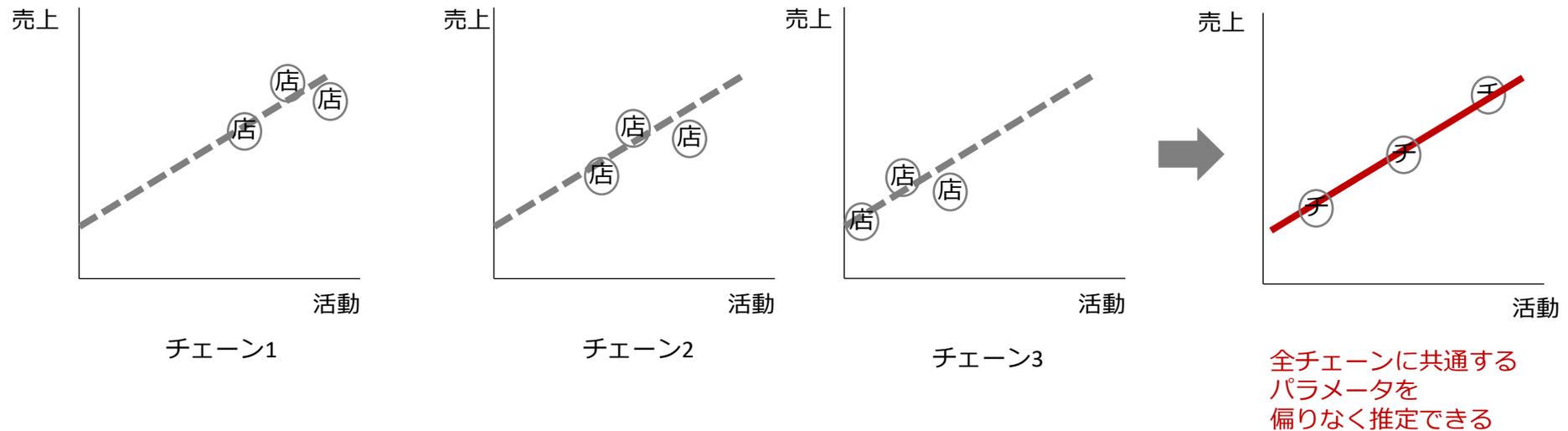
次の2つの要因によって生じる：

- マーケティング活動への接触レベルの異質性
例) 販売補助支出額が高いチェーンと低いチェーンがある
- パラメータの異質性
例) 販売補助支出の弾力性が高いチェーンと低いチェーンがある
切片項 (売上のベースライン)が高いチェーンと低いチェーンがある

売上反応関数が線形である場合について、いくつかの単純な例を考えてみよう。

■ シナリオ1.

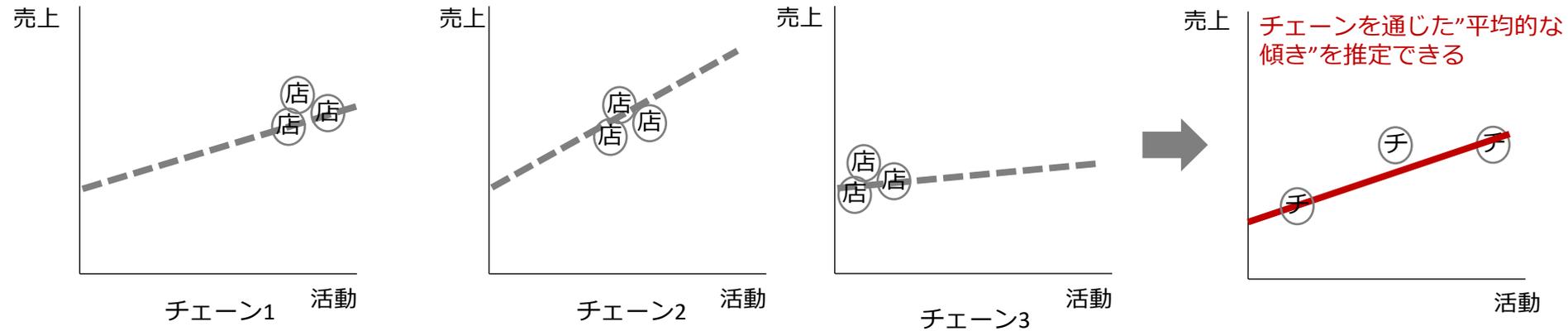
マーケティング活動への接触レベルはチェーン間で異なる
パラメータはチェーン間で等しい



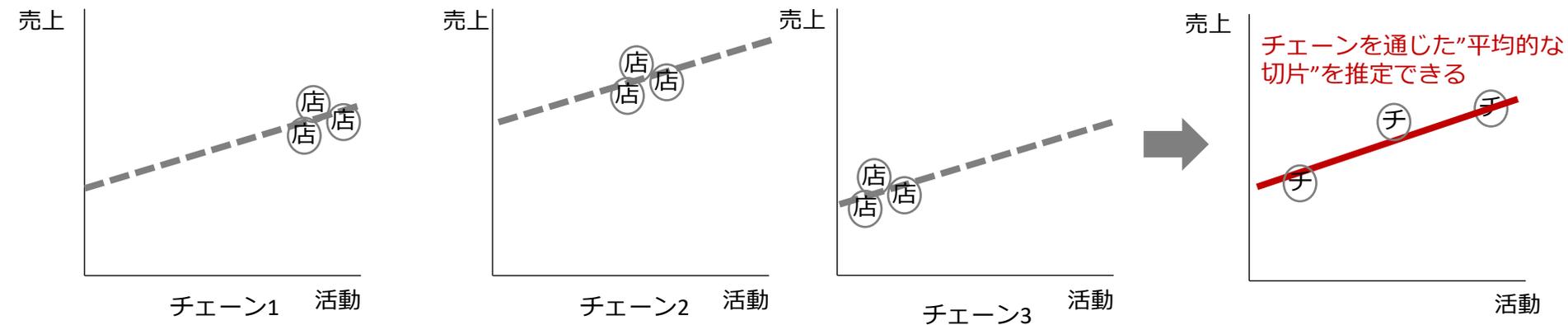
■ シナリオ2.

接触レベルもパラメータもチェーンによって異なる
接触レベルとパラメータに相関はない

チェーンによって傾きが異なる：



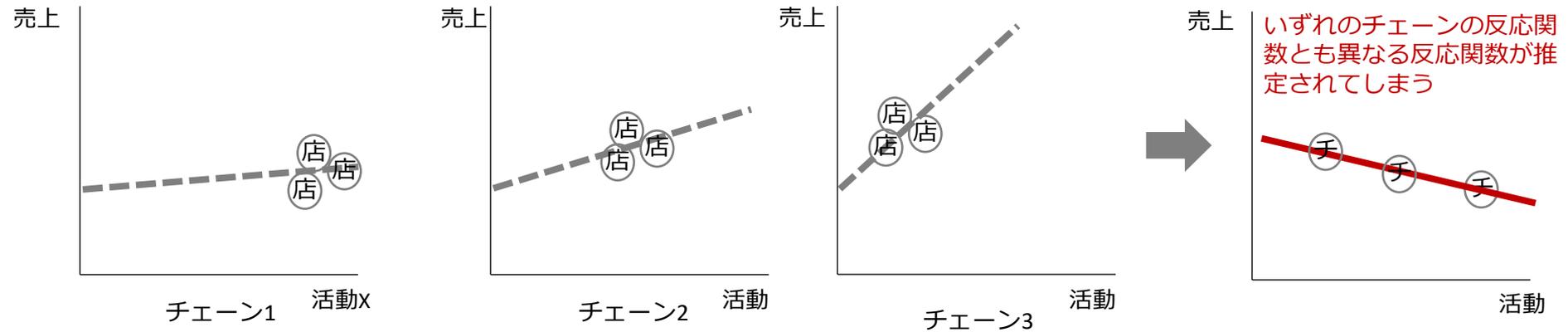
チェーンによって切片が異なる：



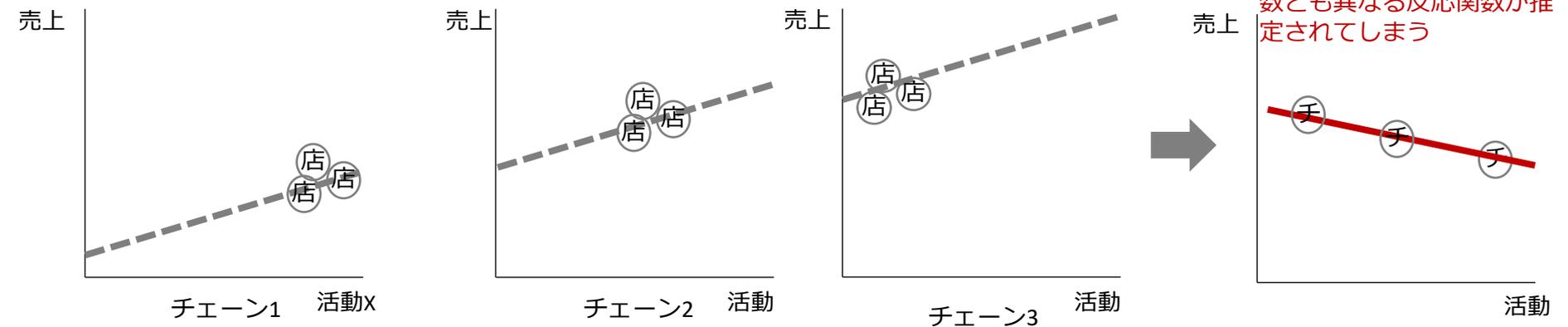
■ シナリオ3.

接触レベルもパラメータもチェーンによって異なる
接触レベルとパラメータに相関がある

各チェーンの傾きが、そのチェーンのマーケティング活動の水準と相関している場合：



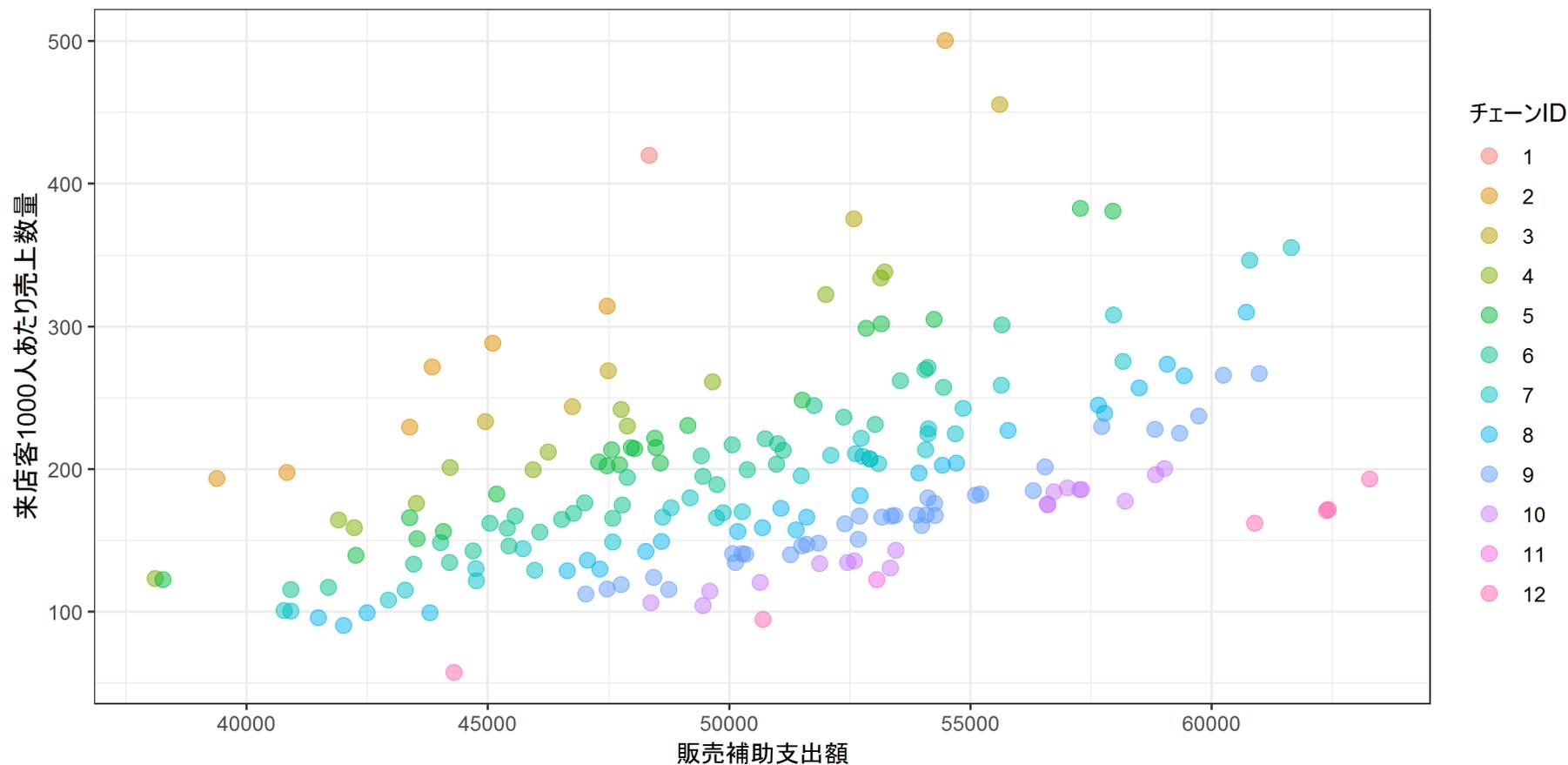
各チェーンの切片が、そのチェーンのマーケティング活動の水準と相関している場合：



店舗販売データ

Code 19

実は、売上が低いチェーンに対して、より多くの販売補助支出を行っていた
(シナリオ3に相当)



■ まとめ：主体集計バイアスはいつ起きるか

- 売上反応関数が線形である場合
接触レベルとパラメータの両方に異質性があり、さらに相関があるとき、バイアスが生じる
- 売上反応関数が非線形である場合
接触レベルとパラメータの両方に異質性があるとき、常にバイアスが生じる (cf. Christen, et al. 1997)

■ 異質性への対処

- “本来の”主体はなにかを考える
 - 理想的なデータとは？
- “本来の”主体における異質性について考える
 - マーケティング活動への接触レベルには、どんな個体差があるだろうか？
 - パラメータには、どんな個体差があるだろうか？
 - マーケティング活動への接触レベルは、パラメータと相関していないか？
- できるかぎり、“本来の”主体を単位とするデータを手にする
 - DM送付の効果 → 顧客レベルデータはないか？
 - 完全なデータが手に入ったら
 - 異なるレベルを同時に分析する... マルチレベルモデル
 - 限られたデータでも有益
 - 例) 一部の顧客について、顧客レベルのデータが入手できた
→ 集計レベルデータで得た知見を顧客レベルデータで確認できる

10. 推定結果のプレゼンテーション

この節では、売上反応モデルの推定結果をわかりやすく表現する方法について、次の2つの観点から議論する。

1. パラメータ推定値の表現
2. シミュレーションによる表現

次の2点に留意する：

- わかりやすく表現する

例: Y_1 は売上数量(1000個), X_2 は広告支出(百万円)

関数形	表現の例
$Y_i = \beta_1 + \beta_2 X_2 + \dots$	他の変数が変わらないとして、広告支出が百万円増えると、売上数量は 1000$\hat{\beta}_2$個 増える
$\log(Y_i) = \beta_1 + \beta_2 X_2 + \dots$	他の変数が変わらないとして、広告支出が百万円増えると、売上数量は exp($\hat{\beta}_2$) 倍に増える (注1)
$Y_i = \beta_1 + \beta_2 \log(X_2) + \dots$	他の変数が変わらないとして、広告支出が1パーセントポイント増えると、売上数量は 9.95 $\hat{\beta}_2$個 増える (注2)
$\log(Y_i) = \beta_1 + \beta_2 \log(X_2) + \dots$	他の変数が変わらないとして、広告支出が1パーセントポイント増えると、売上数量は $\hat{\beta}_2$ パーセントポイント増える

注1) $\log(Y_2) = \log(Y_1) + \beta$ であるとき $\exp(\beta) = Y_2/Y_1$

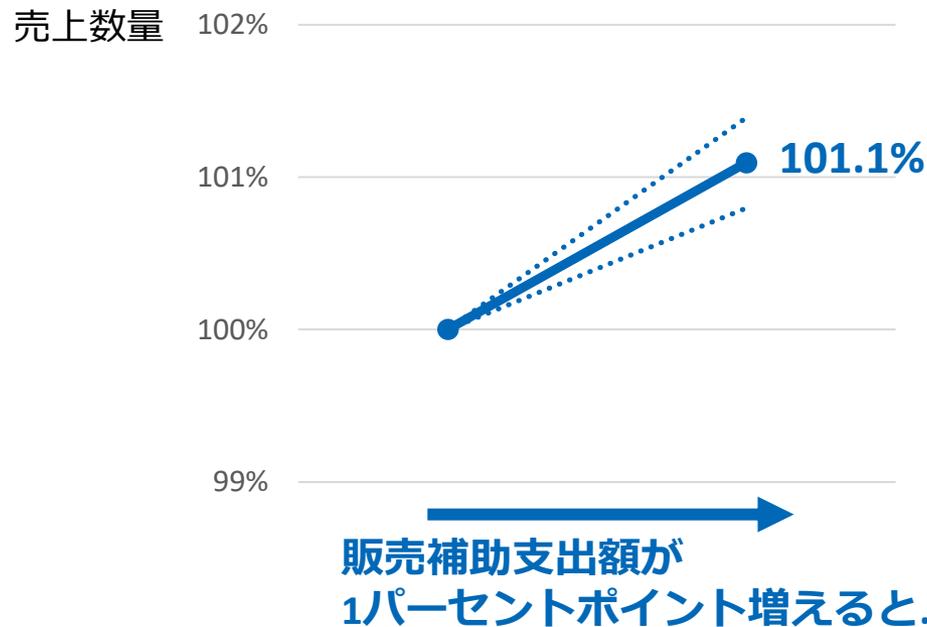
注2) $\log(1.01X) - \log(X) = \log(1.01) = 0.00995$

- 推定の不確実性を伝える

Code 20

例) 3節モデル2によって推定した販売補助支出額の効果

宣材送付個数が変わらないとして、
販売補助支出額が1パーセントポイント増えると、
売上数量は**1.1**パーセントポイント増える



■ シミュレーションによる表現の意義

- 目的変数の分布の任意の統計量に対して、説明変数の効果を示すことができる
 - 例) 「すべての店舗の販売補助支出金額を10000円増やしたら、売上数量10000個以上の店舗数はどうなる？」
- パラメータ推定の不確実性だけでなく、次の3つのばらつきを表現できる
 - パラメータ推定の不確実性
 - 説明変数の分散によって生じるばらつき
 - 攪乱項によって生じるばらつき

■ 方法 (King, Tomz, Wittenberg, 2000)

例) 3節モデル2による

- モデルは

$$\log(Y_i) = \beta_1 + \beta_2 \log(X_{2i}) + \beta_3 X_{3i} + U_i$$
$$U_i \sim N(0, \sigma^2)$$

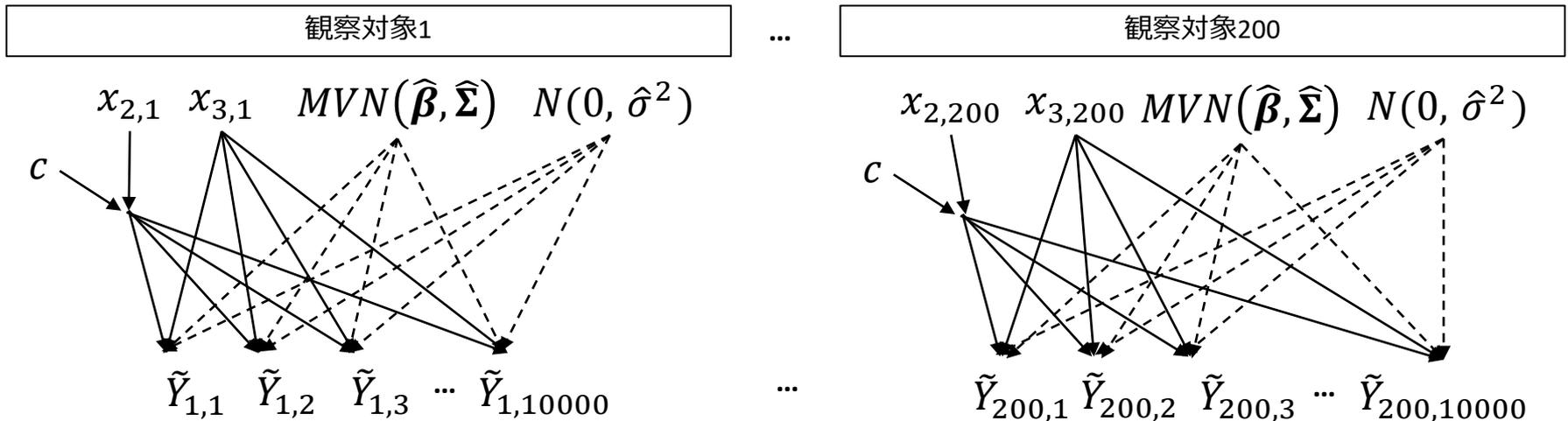
- すべての店舗において X_{2i} (販売補助支出金額)が c 増大したときの売上数量についてシミュレーションしたい

<記号>

- y_i, x_{1i}, x_{2i} : データ
- $\hat{\beta} = (\hat{\beta}_1 \quad \hat{\beta}_2 \quad \hat{\beta}_3)'$
- $\hat{\Sigma} = \begin{pmatrix} \widehat{Var}(\hat{\beta}_1) & \widehat{Cov}(\hat{\beta}_1, \hat{\beta}_2) & \widehat{Cov}(\hat{\beta}_1, \hat{\beta}_3) \\ \widehat{Cov}(\hat{\beta}_2, \hat{\beta}_1) & \widehat{Var}(\hat{\beta}_2) & \widehat{Cov}(\hat{\beta}_2, \hat{\beta}_3) \\ \widehat{Cov}(\hat{\beta}_3, \hat{\beta}_1) & \widehat{Cov}(\hat{\beta}_3, \hat{\beta}_2) & \widehat{Var}(\hat{\beta}_3) \end{pmatrix}$

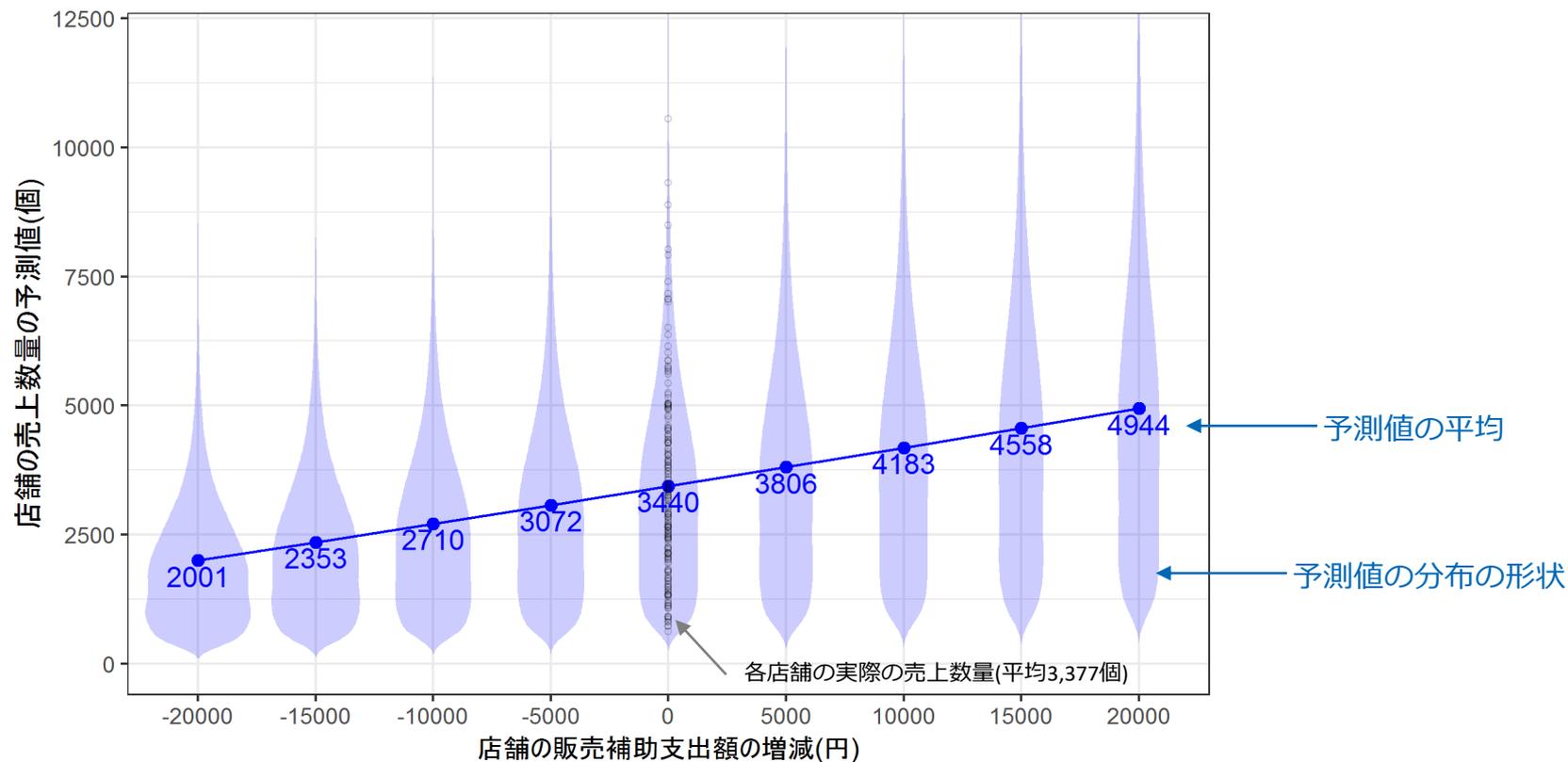
• 予測値のシミュレーション

- 架空のシナリオについての予測に適する
- 個々の*i*について、 $\tilde{Y}_{ij} = \exp(\tilde{b}_{1j} + \tilde{b}_{2j} \log(x_{2i} + c) + \tilde{b}_{3j} x_{3i} + \tilde{u}_j)$ を多数生成する
 - ただし、 $(\tilde{b}_{1j} \ \tilde{b}_{2j} \ \tilde{b}_{3j})' \sim MVN(\hat{\beta}, \hat{\Sigma})$, $\tilde{u}_j \sim N(0, \hat{\sigma}^2)$
- \tilde{Y}_{ij} に来店客数を掛けた値の分布を示す



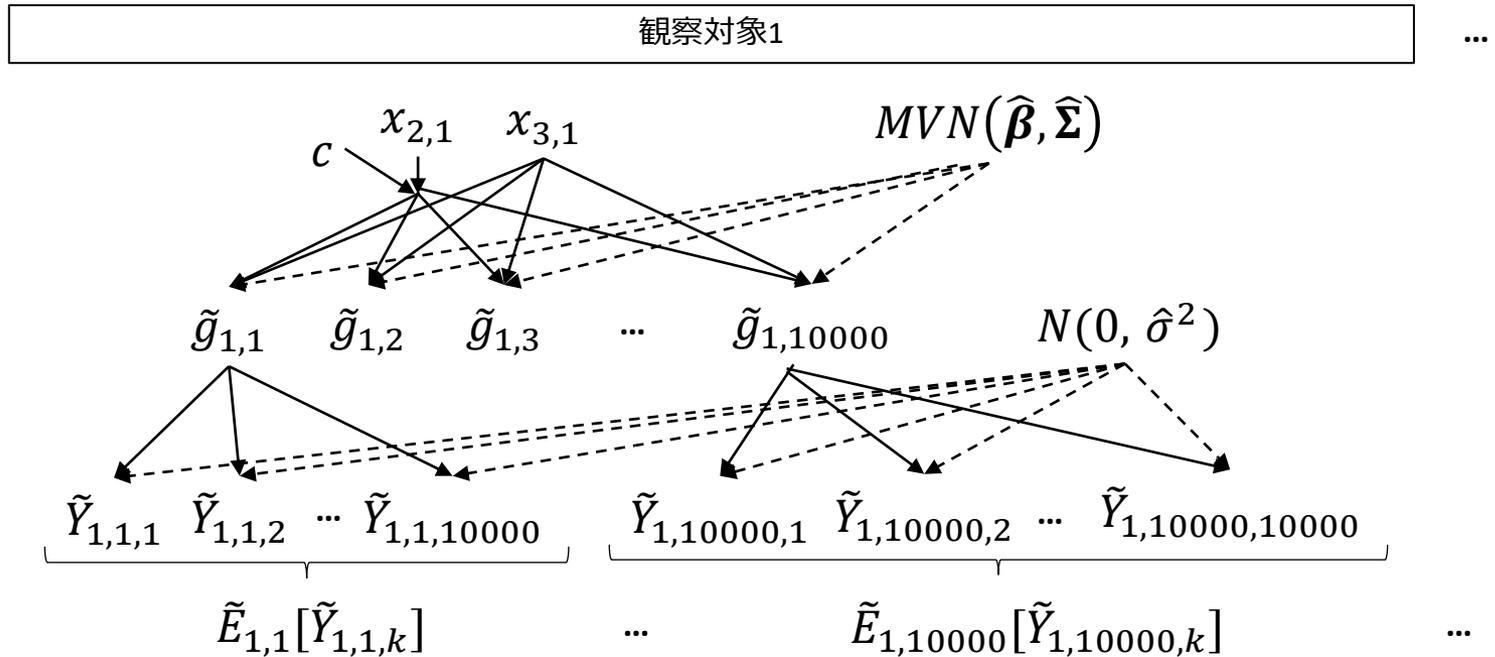
確率分布からのドローを点線で示す

各店舗の販売補助支出額を一律に x 円増減させたときの、売上数量の予測値の分布

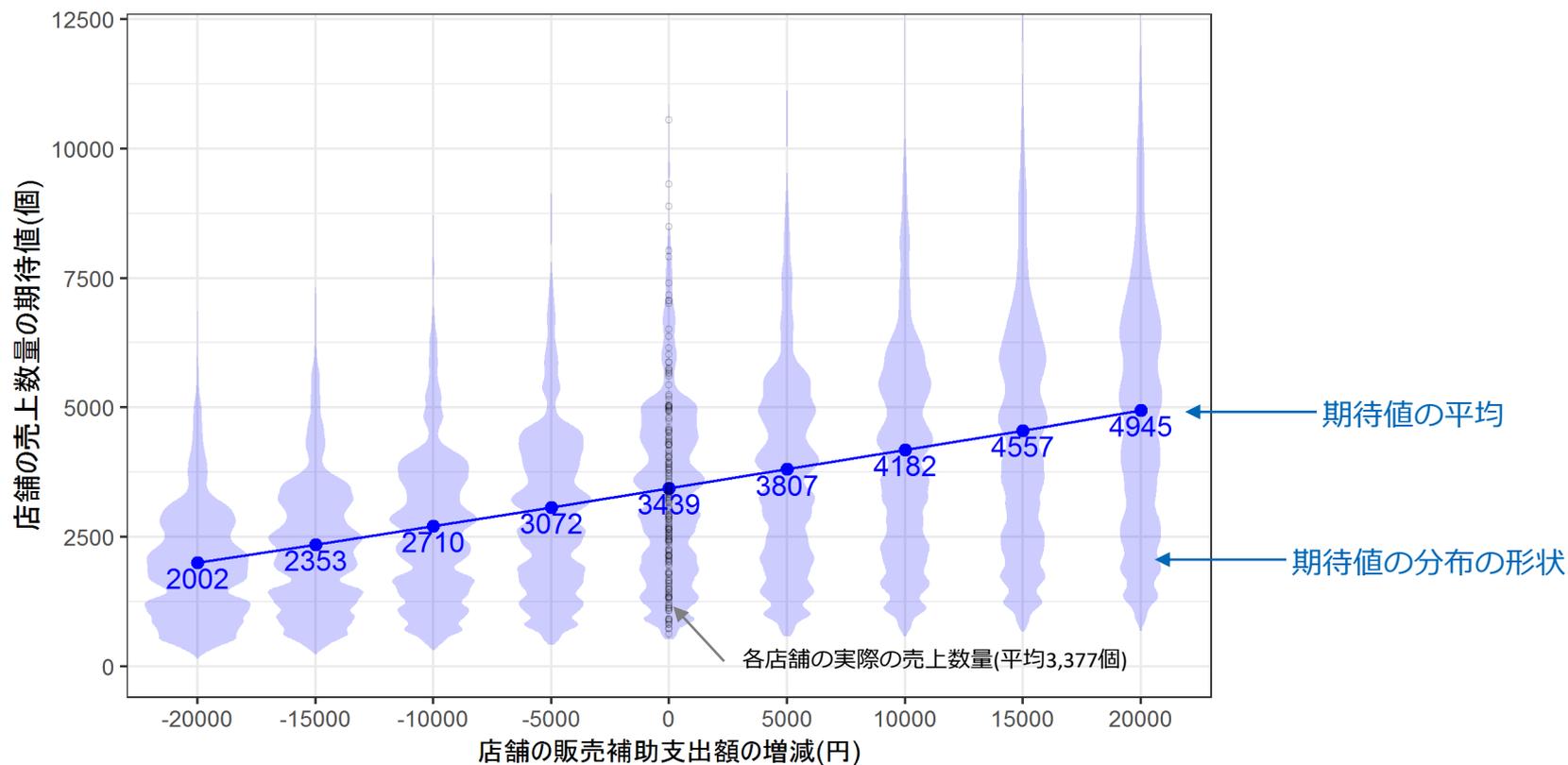


• 期待値のシミュレーション

- 説明変数の効果の表現に適する
- 個々の i について、 $\tilde{g}_{ij} = \tilde{b}_{1j} + \tilde{b}_{2j} \log(x_{2i} + c) + \tilde{b}_{3j} x_{3i}$ を多数生成する
- 個々の \tilde{g}_{ij} について、 $\tilde{Y}_{ijk} = \exp(\tilde{g}_{ij} + \tilde{u}_k)$ を多数生成し、その平均 $\tilde{E}_{ij}[\tilde{Y}_{ijk}]$ を求める
- $\tilde{E}_{ij}[\tilde{Y}_{ijk}]$ に来店客数を掛けた値の分布を示す



各店舗の販売補助支出額を一律に x 円増減させたときの、売上数量の期待値の分布



片平秀貴 (1987) 「マーケティング・サイエンス」, 東京大学出版会.

里村卓也 (2010) 「マーケティング・モデル」, 共立出版.

金明哲 (2017) 「Rによるデータサイエンス 第2版」, 森北出版.

浅野哲・中村二郎 (2009) 「計量経済学 第2版」, 有斐閣.

Hestie, T., Tibshirani, R., Friedman, J. (2014) 「統計的学習の基礎」, 杉山将ほか(監訳・訳), 共立出版. (原著 2011)

Christen, M., Gupta, S., Porter, J.C., Staelin, R., Wittink, D.R. (1997) Using Market-Level Data to Understand Promotion Effects in a Nonlinear Model. *Journal of Marketing Research*, 34(3), 322-334.

King, G., Tomz, M., Wittenberg, J. (2000) Making the most of statistical analysis: Improving interpretation and presentation. *American Journal of Political Science*, 44(2), 341-355.

- `nls(フォーミュラ, data=データフレーム, ...)`
 - 回帰モデルを非線形最小二乗法で推定する
 - さまざまなオプションをうまく指定する必要がある
 - いくつかの関数形については、簡単な指定方法が用意されている
 - 例) ロジスティック関数の場合
`nls(目的変数 ~ SSLogis(説明変数, ...), data=データフレーム)`