

◎「R」を使って練習問題を解きながら
ビッグデータの活用に必須知識を学ぶ。



Rによる 統計的学習 入門

(原著) G. James, D. Witten, T. Hastie, R. Tibshirani

(訳) 落海浩 (南カリフォルニア大学ビジネススクール)

首藤信通 (神戸大学大学院海事科学研究科)

A5判 416頁 オールカラー

定価 (本体 6,800円+税) (12224-4) C3041

- ・南カリフォルニア大学ビジネススクールの人気講義のテキスト“An Introduction to Statistical Learning: with Applications in R”の全訳。
- ・道具として使えるようにRで実践的に統計解析・機械学習を学ぶ。
- ・統計・情報・計算機科学・データサイエンス・経済・経営・金融 (FinTech)などを専攻する学部学生, ビジネススクールの学生, 若手研究者, 実務者。ビッグデータの解析を期待される担当者。

〔目次〕

1. 導入
2. 統計的学習
3. 線形回帰
4. 分類
5. リサンプリング法
6. 線形モデル選択と正則化
7. 線形を超えて
8. 木に基づく方法
9. サポートベクターマシン
10. 教師なし学習

きりとり線

【お申し込み書】この申し込み書にご記入のうえ、最寄りの書店にご注文下さい。

Rによる統計的学習入門

A5判 416頁 定価 (本体 6,800円+税)

ISBN978-4-254-12224-4 C3041

冊

取扱書店

●お名前 公費 / 私費

●ご住所 (〒) TEL

 朝倉書店

〒162-8707 東京都新宿区新小川町 6-29 / 振替 00160-9-8673
電話 03-3260-7631 / FAX 03-3260-0180
<http://www.asakura.co.jp> / eigyo@asakura.co.jp

・ 数学的な厳密さより、わかりやすさ、使えることを重視！

education	-0.020852	0.019613	-1.06	0.28836
UrbanYes	0.140160	0.112402	1.25	0.21317
USYes	-0.157557	0.148923	-1.06	0.29073
Income:Advertising	0.000751	0.000278	2.70	0.00799
Price:Age				

Rによる 統計的学習入門

本文組見本

オールカラーで
見やすい本文

2.1 統計的学習とは

統計的学習について学ぶにあたり、まずは簡単な例から始める。例えば私たちは統計コンサルタントで、ある商品のセールス^{訳注1}を増やすための提案を行うためにクライアント企業に雇われたとする。Advertisingデータセットの中に、200の市場におけるセールスデータと、それぞれの市場でのTV, radio, そして newspaper につ

クライアント企業が直接
いくら使うかを決
めがあることを示せれ
ばより、間接的にセー

7.8 実習：非線形モデリング

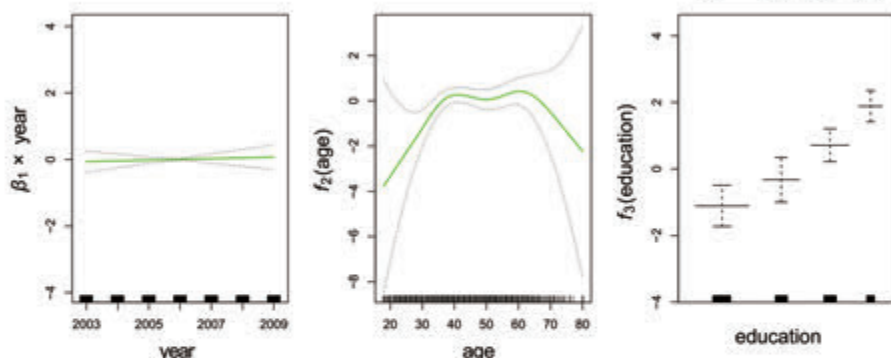
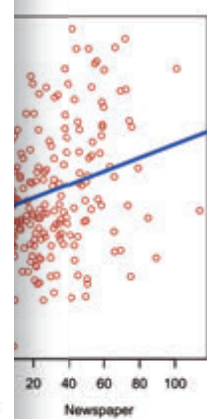


図 7.14 図 7.13 と同じモデルであるが、今回は education が<HS となる観測値を除いている。最終学歴のレベルが高いと、より高収入となる様子がわかる。

は図 7.13 の通りである。最後の図では、<HS のレベルの信頼区間幅が非常に広がっており、疑わしい結果となっている。実際、このカテゴリーには誰も属していない。つまり、最終学歴が高卒未満の人で年収\$250,000 以上となっている人はいない。したがって、高卒未満の人達を除いて、再度 GAM を当てはめる。その結果は、図 7.14 に示されている。図 7.11, 7.12 と同様、3つの図はいずれも縦軸のスケールが類似しており、age と education は year よりも大きい影響をもつことがわかる。



によって sales
それぞれのグラフ
して得られた直
paper を使って

7.8 実習：非線形モデリング

この実習では、多くの複雑な非線形モデルの当てはめが R で簡単に実行可能であることを例示するために、本章の例として扱った Wage データを再度解析する。はじめに、このデータをもつ ISLR ライブラリを読み込む。

```
> library(ISLR)
> attach(Wage)
```

7.8.1 多項式回帰と階段関数

図 7.1 がどのようにして得られるかを体験する。初めに、以下のコマンドによってモデルを当てはめる。

```
> fit=lm(wage~poly(age,4),data=Wage)
> coef(summary(fit))
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   111.704      0.729   153.28  <2e-16
poly(age, 4)1  447.068     39.915    11.20  <2e-16
poly(age, 4)2 -478.316     39.915   -11.98  <2e-16
```