

無断使用をお断りします。日科技連出版社

Rで学ぶ マーケティングリサーチ

市場調査で収集したデータから消費者動向を知る

長畑秀和 著



日科技連

まえがき

本書は、マーケティングリサーチを学ぶ入門書です。多くの(手法を用いた)例題をRを使って解析・解釈しながらマーケティングリサーチを学ぶことを目的としています。本書ではまず、基本的なマーケティングリサーチの流れを説明します。調査を企画し、質問文・調査票を作成して、対象者・地域の選定後、調査を実施して、データを集計します。その後、分析において具体的に解析し、解釈します。本書は、コンピュータ上でフリーソフトであるRを利用して実際に計算し、解析手法を会得するための実習書にもなります。データ解析を理解するには具体例について計算し、実行してみることが必要です。これには複雑な計算を伴うので、コンピュータ利用がたいへん有効です。

本書を大学の講義で利用される場合、コンピュータが利用できる実習室で行われるときは例題を解説後、Rで実行されながら進めていただければと思います。主に講義で進められる場合は、解説後にコンピュータによる実行を提示しながら説明されたいと思います。受講された学生さんは自宅や情報実習室のコンピュータを動かして理解を深めていただければと思います。

独学でマーケティングリサーチを学ぼうとされている方は、まず第1章から第5章でどのようにデータを取り、集計していくかの流れを学んでいただき、実際にデータを集計した場合を第6章で学んでいただき、第7章でさまざまな具体例をRで読み込み解析しながら解釈・理解を進めていただければと思います。本書ではRの実行にRコマンドを利用して、実行するために入力するファイル(データ)を例題ごとに分けています。それらのファイルを読み込んで、実行したい部分をドラッグして実行し、結果を確認しながら読み進めてください。

本書の構成を以下に簡単に述べておきます。実際にマーケティングリサーチにより、市場を調査し、そのことから情報を得て、市場に生かす流れに沿って各章を書いています。第1章で、マーケティングリサーチの概要について書いています。次に、第2章では実際に必要とされる調査テーマの決定について述べています。そして調査をするにあたっての調査対象とサンプリングについて第3章で述べています。次に第4章では調査の方法について解説しています。そして調べる内容を具体化するため、第5章では調査票の作成について書いています。調査実施後の調査データの加工・分析に

については、第6章で、調査したデータを集計し、基本的な加工・分析する手法について述べています。具体的な解析・解釈について理解するため、第7章ではさまざまな例題について具体的に実行過程を表示しながら解説しています。さらに第8章では、発展的な例題について具体的に実行過程を表示しながら解説しています。そして解析した結果までをまとめて報告するため、第9章では、調査結果の報告についてまとめています。

R の version は、R-4.1.0 で、Rcmdr の version は、2.7-1 で実行しています。また、パッケージとして利用するのは以下です。

arules, FactoMineR, lavaan, MASS, pwr, ResourceSelection, semPlot, vcd

また、第6章～第8章までの例題・演習問題に使用したデータの一部は、日科技連出版社のウェブサイトからダウンロードできます。詳しくは目次の末尾を参照してください。

なお、思わぬ間違いがあるかもしれません。また解釈が不十分な箇所もありますが、ご意見をお寄せください。より改善していきたいと思っております。

本書の出版にあたって日科技連出版社の鈴木兄宏氏には細部にわたって校正いただき、たいへんお世話になりました。心より感謝いたします。最後に、日頃、いろいろと励ましてくれた家族に一言お礼を伝えたいと思います。

2022年10月

長 畑 秀 和

凡例(記号など)

以下に、本書で使用する文字、記号などについてまとめる。

① Σ (サメンション)記号

添え字とともに用いて、その添え字のある番地のものについて、 Σ 記号の下で指定された番地から Σ 記号の上で指定された番地まで足し合わせることを意味する。

[例] $\sum_{i=1}^n x_i = x_1 + x_2 + \cdots + x_n = x.$

② 順列と組合せ

異なる n 個のものから r 個をとって、1列に並べる並べ方は、

$$n(n-1)(n-2)\cdots(n-r+2)(n-r+1)$$

通りあり、これを ${}_n P_r$ と表す。これは階乗を使って、 ${}_n P_r = \frac{n!}{(n-r)!}$ とも表せる。なお、 $n! = n(n-1)\cdots 2 \cdot 1$ であり、 $0! = 1$ である (cf. Permutation)。異なる n 個のものから r 個とる組合せの数は(とったものの順番は区別しない)、順列の数をとってきた r 個の中での順列の数で割った

$$\frac{{}_n P_r}{r!} = \frac{n!}{(n-r)!r!}$$

通りである。これを、 ${}_n C_r$ または $\binom{n}{r}$ と表す (cf. Combination)。

[例] ${}_5 P_3 = 5 \times 4 \times 3 = 60, {}_5 C_3 = \frac{5 \times 4 \times 3}{3 \times 2 \times 1} = 10$

③ ギリシャ文字

表 ギリシャ文字の一覧表

大文字	小文字	読み	大文字	小文字	読み	大文字	小文字	読み
A	α	アルファ	I	ι	イオタ	P	ρ	ロー
B	β	ベータ	K	κ	カッパ	Σ	σ	シグマ
Γ	γ	ガンマ	Λ	λ	ラムダ	T	τ	タウ
Δ	δ	デルタ	M	μ	ミュー	Υ	υ	ユ(ウ)プシロン
E	ϵ	イプシロン	N	ν	ニュー	Φ	ϕ, φ	ファイ
Z	ζ	ゼータ(ツェータ)	Ξ	ξ	クサイ(クザイ)	X	χ	カイ
H	η	イータ	O	o	オミクロン	Ψ	ψ	サイ(プサイ)
Θ	θ, ϑ	テータ(シータ)	Π	π	パイ	Ω	ω	オメガ

なお、通常 μ を平均、 σ^2 を分散を表すために用いることが多い。

R のインストール方法

R のインストール方法については以下のウェブサイトが参考になります。なお、2022 年 10 月 20 日現在の情報です。

■Windows の場合

「R 言語インストール(2020 年 Windows)」

<https://qiita.com/FukuharaYohei/items/8e0ddd0af11132031355>



■Mac OS の場合

「臨床医のための R コマンダーによる医学統計解析マニュアル」

http://rcommanderdeigakutoukeikaiseki.com/support_install_mac.html



目 次

まえがき	iii
凡例(記号など)	v
Rのインストール方法	vi
第1章 マーケティングリサーチとは	1
1.1 マーケティング	1
1.2 調査で扱うデータ	5
1.3 調査の種類	5
1.3.1 調査の用途の面からの分類	6
1.3.2 調査方法の面からの分類	6
1.3.3 機能の面からの分類	6
1.3.4 調査対象の面からの分類	7
1.3.5 調査対象との接触面からの分類	7
1.4 マーケティングリサーチとは	7
1.4.1 市場調査	8
1.4.2 母集団	10
1.4.3 サンプリングとサンプル	11
1.4.4 測定とデータ	12
1.4.5 調査の手法(仕方)	12
1.4.6 集計・分析	13
1.4.7 調査結果の報告	13
1.4.8 基礎資料の収集	15
第2章 調査テーマの決定	17
2.1 テーマ決定にあたって	17
2.1.1 個人的関心	17
2.1.2 社会問題	17
2.1.3 学術的関心	18

2.1.4	必要性	18
2.1.5	外的刺激による選択	18
2.2	課題の明確化と概念規定	18
2.2.1	課題の明確化	18
2.2.2	概念と概念規定	18
2.2.3	概念の操作化	18
第3章 調査対象とサンプリング		23
3.1	調査対象	23
3.2	サンプリング(抽出)方法	23
3.2.1	確率抽出	24
3.2.2	有意抽出(非確率的サンプリング)	28
3.3	測定とデータ	29
3.3.1	サンプル数の決定	31
第4章 調査の方法		35
4.1	調査(実施)の手法(仕方)	35
4.1.1	量的調査法(定量調査)	35
4.1.2	質的調査法(定性調査)	37
4.2	調査の実施	39
第5章 調査票の作成		41
5.1	表紙の作成	41
5.2	フェイスシートの作成	42
5.3	質問項目の作成	43
5.4	調査票の設計	43
5.5	予備調査	46
5.6	調査票の例	47
5.6.1	ホテルにおけるアンケート調査	48
5.6.2	学生生活実態調査	51
第6章 調査データの加工・分析		59
6.1	調査データの分析	59

6.1.1	ローデータ (raw data : 生データ) の処理	59
6.2	データのまとめ方	61
6.2.1	グラフによるまとめ	61
6.2.2	単一回答の入力	62
6.2.3	複数回答の入力	63
6.2.4	順位回答の入力	64
6.3	データの集計	65
6.3.1	単純集計	66
6.3.2	クロス集計表 (クロス表)	75
第7章	さまざまな分析例	89
7.1	分散分析	89
7.2	回帰分析	109
7.3	数量化Ⅰ類	128
7.4	ロジスティック回帰分析	136
7.5	多項ロジット回帰分析	141
7.6	判別分析	145
7.7	数量化Ⅱ類	150
7.8	決定木分析	159
7.9	コンジョイント分析	165
7.10	主成分分析	173
7.11	因子分析	185
7.12	クラスター分析	196
7.12.1	階層的クラスター分析	196
7.12.2	非階層的クラスター分析	200
第8章	発展的な分析例	203
8.1	対応分析	203
8.2	数量化Ⅲ類	214
8.3	多次元尺度法	219
8.4	数量化Ⅳ類	228
8.5	ニューラルネットワーク分析	233
8.6	バスケット分析	237

8.7	正準相関分析	245
8.8	階層化意思決定法	251
8.9	共分散構造分析	254
第9章 調査結果の報告		277
9.1	調査報告書の構成	277
9.2	調査結果の報告	277
9.3	フィールドノートの作成	278
参考文献		279
索引		281

例題・演習問題のデータについて

例題・演習問題の一部は、日科技連出版社のウェブサイト (<https://www.juse-p.co.jp/>) からダウンロードできます。トップページ上部のタブ [ダウンロード] をクリックすると、検索画面が表示されますので、書名もしくは ISBN を入力してください。

注意事項

1. 例題・演習問題の著作権は、著者にあります。本データを無断で使用することを禁じます。
2. 著者および出版社のいずれも、本データをダウンロードしたことに伴い生じた損害について、責任を負うものではありません。

例題 7-8(数量化Ⅱ類)

ある製品の購買ブランドを数量化Ⅱ類により年代, 好み, 友人数で判別する方式を検討せよ。

《R(コマンダー)による解析》

(0) 予備解析

手順1 データの読み込み

操作手順 【データ】▶【データのインポート】▶【テキストファイルまたはクリップボード, URL から...】を選択し, ダイアログボックスで, フィールドの区切り記号としてカンマにチェックを入れて, をクリックする。フォルダからファイルを指定後, をクリックする。そしてをクリックすると, データが表示される。

【出力ウィンドウ】

```
> rei78 <- read.table("C:/RMRDAT/7syo/rei78.csv",
  header=TRUE, sep=",", na.strings="NA", dec=".",
  strip.white=TRUE)
> showData(rei78, placement='-20+200',
  font=getRcmdr('logFont'), maxwidth=80, maxheight=30)
```

手順2 基本統計量の計算

操作手順 【統計量】▶【要約】▶【アクティブデータセット】を選択し, ダイアログボックスで, フィールドの区切り記号としてカンマにチェックを入れて, をクリックする。そしてをクリックすると, データの要約が表示される。

【出力ウィンドウ】

```
> summary(rei78)
```

No	購買ブランド	年代	好み	友人数
Min.	:1.00 A:10	Min. :1.00	Min. :1.00	Min. :1.00
1st Qu.	:5.75 B:10	1st Qu. :1.75	1st Qu. :1.00	1st Qu. :1.75
Median	:10.50	Median :2.50	Median :2.00	Median :2.00
Mean	:10.50	Mean :2.50	Mean :1.95	Mean :2.10
3rd Qu.	:15.25	3rd Qu. :3.25	3rd Qu. :3.00	3rd Qu. :3.00
Max.	:20.00	Max. :4.00	Max. :3.00	Max. :3.00

```
> str(rei78)
'data.frame': 20 obs. of 5 variables:
 $ No      : int  1  2  3  4  5  6  7  8  9 10 ...
```

```

$ 購買ブランド : Factor w/ 2 levels "A","B": 1 1
$ 年代          : int  3  3  1  2  2  2  1  1  4  1 ...
$ 好み          : int  1  1  2  2  2  2  1  1  1  1 ...
$ 友人数        : int  3  3  2  2  2  2  1  3  1  1 ...
    
```

操作手順 図 7.99 のように、【データ】 ▶ 【アクティブデータセット内の変数の管理】 ▶ 【数値変数を因子に変換...】 を選択し、図 7.100 のように、ダイアログボックスで、変数を指定し、数値でにチェックを入れて、**OK** をクリックする。すると図 7.101 のように変数を上書きするか聞いてくるので、逐次**Yes** をクリックする。

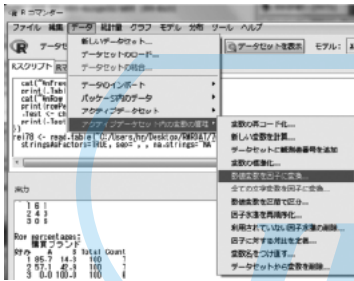


図 7.99 数値変数の因子化の指定



図 7.100 変数の指定

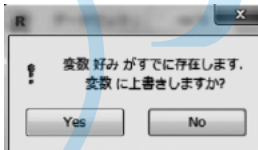


図 7.101 変数の上書き Yes

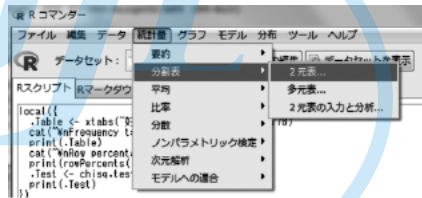


図 7.102 分割表の指定

【出力ウィンドウ】

```

# 数値の因子への変換
> rei78 <- within(rei78, {
+   好み <- as.factor(好み)
+   年代 <- as.factor(年代)
+   友人数 <- as.factor(友人数)
+ })
    
```

手順3 グラフの作成

操作手順 図 7.102 のように、【統計量】 ▶ 【分割表】 ▶ 【2元表...】 を選択し、図

7.103のように、行の変数で好み、列の変数で購買ブランドを選択する。さらに、【統計量】を選択し、図7.104のように行のパーセントと独立性のカイ2乗検定にチェックを入れて、**[OK]**をクリックすると、以下のように表示される。好みにより購買ブランドに違いがありそうである。同様に年代と購買ブランド、友人数と購買ブランドの分割表を作成する。

なお、購買ブランドと友人数の関係を見るモザイクプロットは図7.105であり、それほど関係があるとは見られない。

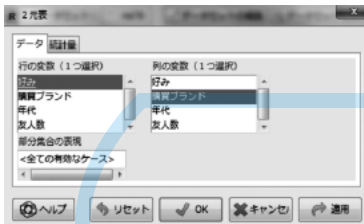


図 7.103 変数の指定

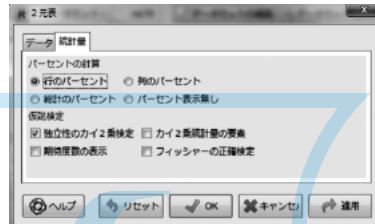


図 7.104 2変数での統計量の指定

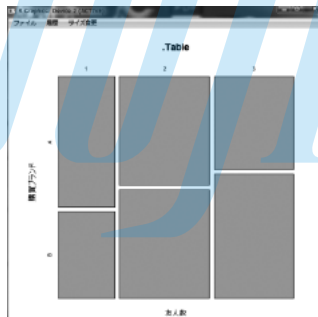


図 7.105 モザイクプロット

【出力ウィンドウ】

分割表の作成

```
> local{
```

```
+ .Table <- xtabs(~好み + 購買ブランド, data=rei78)
```

```
+ cat("\nFrequency table:\n")
```

```
+ print(.Table)
```

```
+ .Test <- chisq.test(.Table, correct=FALSE)
```

```
+ print(.Test)
```

```
+ }
```

Frequency table:

購買ブランド

```
好み A B
1 6 1
2 4 3
3 0 6
```

Pearson's Chi-squared test

data: .Table

X-squared = 9.7143, df = 2, p-value = 0.007773

分割表の作成

```
> local({
+   .Table <- xtabs(~年代+購買ブランド, data=rei78)
+   cat("\nFrequency table:\n")
+   print(.Table)
+   cat("\nRow percentages:\n")
+   print(rowPercents(.Table))
+   .Test <- chisq.test(.Table, correct=FALSE)
+   print(.Test)
+ })
```

Frequency table:

購買ブランド

```
年代 A B
1 4 1
2 3 2
3 2 3
4 1 4
```

Row percentages:

購買ブランド

年代	A	B	Total	Count
1	80	20	100	5
2	60	40	100	5
3	40	60	100	5
4	20	80	100	5

Pearson's Chi-squared test

data: .Table

X-squared = 4, df = 3, p-value = 0.2615

分割表の作成

```
> local({
+   .Table <- xtabs(~友人数+購買ブランド, data=rei78)
+   cat("\nFrequency table:\n")
```

```
+ print(.Table)
+ cat("\nRow percentages:\n")
+ print(rowPercents(.Table))
+ .Test <- chisq.test(.Table, correct=FALSE)
+ print(.Test)
+ })
```

Frequency table:

購買ブランド

友人数	A	B
1	3	2
2	4	4
3	3	4

Row percentages:

購買ブランド

友人数	A	B	Total	Count
1	60.0	40.0	100	5
2	50.0	50.0	100	8
3	42.9	57.1	100	7

Pearson's Chi-squared test

data: .Table

X-squared = 0.34286, df = 2, p-value = 0.8425

> mosaicplot(.Table)#モザイクプロットの表示 図7.105

(1) 本解析

手順1 判別関数の導出

【出力ウインドウ】

```
> library(MASS)#ライブラリMASSの利用
> summary(rei78)
      No      購買ブランド 年代 好み 友人数
Min.   : 1.00    A:10      1:5  1:7   1:5
1st Qu.: 5.75    B:10      2:5  2:7   2:8
Median :10.50                3:5  3:6   3:7
Mean   :10.50                4:5
3rd Qu.:15.25
Max.   :20.00

#rei78$購買ブランド<- as.factor(rei78$購買ブランド)
> rei78.lld<-lda(購買ブランド~年代+好み+友人数, data = rei78)
```

```

> rei78.ld
Call:
lda(購買ブランド ~ 年代 + 好み + 友人数, data = rei78)
Prior probabilities of groups:
  A    B
0.5  0.5
Group means:
  年代 [T.2]  年代 [T.3]  年代 [T.4]  好み [T.2]  好み [T.3]  友人数 [T.2]
A    0.3      0.2      0.1      0.4      0.0      0.4
B    0.2      0.3      0.4      0.3      0.6      0.4
  友人数 [T.3]
A    0.3
B    0.4
Coefficients of linear discriminants:
              LD1
年代 [T.2]   0.6449004
年代 [T.3]  -1.7907503
年代 [T.4]   2.6256661
好み [T.2]   4.4452066
好み [T.3]   7.1860335
友人数 [T.2] -3.2878406
友人数 [T.3]  3.7599999
># 定数項
> apply(rei78.ld$mean%*%rei78.ld$scaling, 2, mean)
      LD1
4.08245

```

そこで、判別式は以下ようになる。

$$\begin{aligned}
 y = -4.08245 + & \left\{ \begin{array}{l} 0(20代) \\ 0.6449004(30代) \\ -1.7907503(40代) \\ 2.6256661(50代) \end{array} \right\} + \left\{ \begin{array}{l} 0(嫌い) \\ 4.4452066(普通) \\ 7.1860335(好き) \end{array} \right\} \\
 & + \left\{ \begin{array}{l} 0(少ない) \\ -3.2878406(普通) \\ 3.7599999(多い) \end{array} \right\}
 \end{aligned}$$

手順2 予測

求めた判別関数により、予測を試みよう。判別関数にデータを代入した判別得点

著者紹介

長畑秀和 (ながはた ひでかず) 博士(理学)

1954年岡山県に生まれる。1979年九州大学大学院理学研究科数学専攻博士前期課程修了。1980年九州大学大学院理学研究科数学専攻博士後期課程中退。大阪大学、作陽短期大学、姫路短期大学、岡山大学教育学部を経て、岡山大学大学院社会文化科学研究科(経済学系)教授を歴任。現在、環太平洋大学(IPU)経済・経営学部教授。

[主な著書]

『Rで学ぶ多変量解析』、『Rで学ぶデータサイエンス』(以上、朝倉書店)、
『Rコマンドで学ぶ統計学』(共著、共立出版)

Rで学ぶマーケティングリサーチ

市場調査で収集したデータから消費者動向を知る

2022年11月30日 第1刷発行

著者 長畑秀和

発行人 戸羽節文

検印
省略

発行所 株式会社 日科技連出版社

〒151-0051 東京都渋谷区千駄ヶ谷5-15-5
DSビル

電話 出版 03-5379-1244

営業 03-5379-1238

Printed in Japan

印刷・製本 ㈱三秀舎

© Hidekazu Nagahata 2022

URL <https://www.juse-p.co.jp/>

ISBN 978-4-8171-9770-2

本書の全部または一部を無断でコピー、スキャン、デジタル化などの複製をすることは著作権法上の例外を除き禁じられています。本書を代行業者等の第三者に依頼してスキャンやデジタル化することは、たとえ個人や家庭内での利用でも著作権法違反です。