

無断使用をお断りします。日科技連出版社

Excelによる 多変量解析

豊富な例題ですぐに実務に活用できる

松本哲夫 [編著]

今野勤 [著]



日科技連

まえがき

生産，研究開発，マーケティング，経営企画などの部門においては，業務の革新，効率化，スピードアップのために多変量解析の活用が大切ということが叫ばれて久しい。しかし，多変量解析は数理統計学にその基礎を置くため難解であるという側面がある。統計解析のなかでも，多変量解析と実験計画法がその難しさにおいて双壁と感じている人が多いようである。

多変量解析とは，複数の変数に関するデータをもとに，これらの変数間の相互関連を分析する統計的手法の総称であり，特定の分析方法を指すものではない。

50年以上前になるが，多変量解析のセミナーが開催されるようになった当初は，電卓を使った手計算が基本となっていた時代であったから，行列の計算は簡単なものに限られ，肝心の多変量解析を理解する前に，行列の知識や基礎的な統計解析を理解する必要がある，高いハードルとなっていた。

奥野忠一，久米均，芳賀敏郎，吉澤正共著の『多変量解析法』（日科技連出版社，1971）をはじめとして，当時の多変量解析に関する書籍は極めて難解であり，数学の素養がないと読破は困難であった。近時，多変量解析を平易に解説しようとした書籍も多数見受けられるようになってきたが，理論面の取り扱い方は多種多様である。

本書では，実務への多変量解析の適用を念頭に置いている。よって，計算は手持ちのパソコンで簡単に実施できるようにし，そのための解析ソフトウェアの使用法と，計算結果(出力)の解釈方法について解説する。一方で，理論には深入りすることなく，数式部分を読み飛ばしても直観的に理屈を習得できるよう配慮した。

上記のようなねらいから，第1章では多変量解析の概要と必要最小限の関連知識を紹介し，幅広い場面で使用できる実務にすぐに役立つ解析ソフト，すなわち，Microsoft Excel(以下，単にExcelと呼ぶ)の機能を活用して，汎用的な

手順で多変量解析を定型的に行う方法を示す。そして、Excelの使い方と活用方法については、各章の例題への適用を通して習得できるようにした。

そのうえで、第2章では最も簡単な手法である単回帰分析、第3章では多変量解析の中心的な手法の1つである重回帰分析を取り上げ、第4章から第7章で展開する多変量解析の基礎となる考え方と解析方法を与える。

重回帰分析の発展系として、第4章で数量化理論I類、第5章でロジスティック回帰分析、第6章で曲線回帰分析を取り上げる。そして、第7章では、重回帰分析の理論・手順を用いて判別分析する方法を取り上げた。

第8章では、重回帰分析と並んで、多変量解析の中心的な手法の1つである主成分分析を取り上げ、第9章ではクラスター分析を解説し、第10章では、第9章までで取り上げなかった手法や多変量解析における注意点を示した。

本書の各章には以上のような関連性をもたせて構成してある。

手法の使い方だけでなく、ある程度は理論(数式)も理解したいと思っている読者のために、本文とは切り離し、巻末の付録として理論面について若干の補遺を示しているが、本文の内容はここを読まなくても理解できるので、安心していただきたい。さらなる理論面の習得を目指している読者は、専門書などで学んでいただきたい。

最後になるが、終始筆者らの活動をご支援くださった(一財)日本科学技術連盟大阪事務所の岡田拓治氏と山田ひとみ氏、および出版に当たって常に筆者らを励ましてくださった(株)日科技連出版社の鈴木兄宏氏と田中延志氏に深く感謝する。

2021年7月

編著者 松本哲夫

目 次

まえがき *iii*

第1章 多変量解析	1
1.1 多変量解析に当たっての前処理	2
1.2 データの尺度	2
1.3 多目的最適化	3
1.4 データマイニングとビッグデータ	4
1.5 多変量解析の目的	6
1.6 多変量解析の種類と理論的な基礎	6
1.7 Excel による計算方法	9
1.8 例題	12
1.9 Excel の分析ツールを使用するときの注意事項	18
第2章 単回帰分析	21
2.1 単回帰分析とは	21
2.2 単回帰分析の概要	21
2.3 解析の方法	22
2.4 繰返しのある場合の単回帰分析	33
2.5 おわりに	40
第3章 重回帰分析	41
3.1 重回帰分析とは	41
3.2 重回帰分析の概要	41

3.3	解析の方法	42
3.4	おわりに	64
第4章	数量化理論 I 類	65
4.1	適用場面と活用の仕方	65
4.2	数量化理論 I 類とは	65
4.3	適用例	66
4.4	解析の方法	67
4.5	おわりに	81
第5章	ロジスティック回帰分析	83
5.1	適用場面と活用の仕方	83
5.2	ロジスティック回帰分析とは	83
5.3	適用例	84
5.4	解析の方法	87
5.5	おわりに	91
第6章	曲線回帰分析	93
6.1	適用場面と活用の仕方	93
6.2	曲線回帰分析とは	93
6.3	解析の方法	93
6.4	区間予測について	100
6.5	おわりに	103
第7章	判別分析	105
7.1	適用場面と活用の仕方	105
7.2	判別分析とは	106
7.3	判別分析の考え方	107

7.4 適用例	110
7.5 解析の方法	110
7.6 おわりに	118
第8章 主成分分析	119
8.1 主成分分析とは	119
8.2 主成分分析の概要	119
8.3 解析の方法	122
8.4 おわりに	148
第9章 クラスタ分析	151
9.1 クラスタ分析とは	151
9.2 階層的クラスタ分析の概要	151
9.3 解析の方法	152
9.4 おわりに	160
第10章 その他の多変量解析, および留意事項	161
10.1 その他の多変量解析	161
10.2 正準相関分析	161
10.3 MT システム	163
10.4 多重共線性	164
付録	165
[付録 A] 直交多項式	165
[付録 B] マハラノビス(汎)距離	166
[付録 C] 最小2乗法とデータの直交性	168
[付録 D] 予測値の分散	172
[付録 E] 主成分分析の数理	173

参考文献	176
索引	179

例題・自由演習の数値データと標準解答のダウンロード方法

本書で使用する例題・自由演習の数値データと標準解答は、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。サンプルファイルと本書を併用すればより理解が深まるとともに、実務への適用が容易になります。

パソコンの環境

本書では、Windows 版 Excel がインストールされているパソコンを対象としています。

Excel のバージョンについては、現時点で、2007、2010、2013、2016、2019 での動作確認をしていますが、任意の環境で動作することを保証しているわけではありません。

免責事項

著者、および出版社のいずれも、Excel の解析ソフトウェアを利用した際に生じた損害についての責任、ならびにサポート義務を負うものではありません。

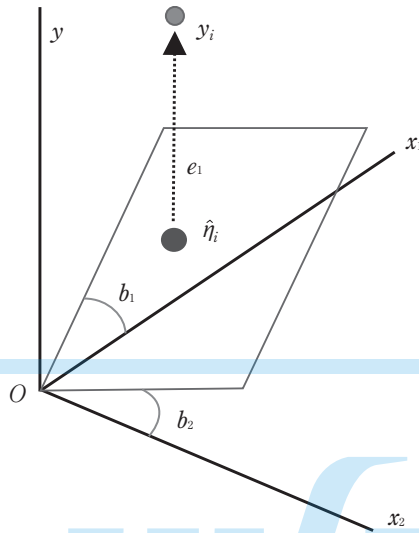


図 3.1 重回帰分析 ($p = 2$) の概念図

は回帰平面を表し、 $\hat{\eta}_i$ は予測値である。式(3.3)で、 $b_0, b_1, b_2, \dots, b_p$ は、それぞれ、 $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ の推定値を示す。また、 e_i は実測値と予測値の残差であり、式(3.4)で表され、 e_i の平方和は、残差の平方和となる。

3.3 解析の方法

以下では、Excelの分析ツールを使った重回帰分析について、2つの例題を解説する。[例題3.1]は、一般的な重回帰分析の例で、解析手順を中心に説明する。[例題3.2]では、変数選択、および多重共線性への対応を含む実務的な例であり、用語の意味などは[例題3.2]で説明する。

[例題 3.1]

表 3.1 は、あるプラスチック部品を製造している工場において、新しいプラスチック部品のライントライをしたときのデータである。プラスチック部品は製造条件の影響を受けやすく、外観や部品の反りに問題が発生す

表 3.1 プラスチック部品ライントライ時のデータ表

No.	x_1 射出速度(g/s)	x_2 冷却時間(min)	x_3 金型温度(°C)	y 反り(mm)
1	31	14	51	-13.1
2	33	16	49	-11.2
3	34	16	46	-18.7
4	39	16	42	-14.3
5	35	19	60	-1.3
6	38	19	58	7.4
7	37	17	52	-1.0
8	33	19	52	-9.4
9	39	15	47	-0.4
10	38	18	50	-1.0
11	34	15	55	-7.7
12	39	16	44	-8.7
13	34	19	57	-3.3
14	30	19	60	2.9
15	33	17	46	-13.5
16	36	17	42	-12.1
17	31	16	50	-9.5
18	36	15	54	-5.5
19	35	17	53	-4.8
20	37	19	48	-3.9
21	33	15	54	-1.0
22	36	18	44	-16.6
23	36	15	59	-1.9
24	35	15	59	-1.2
25	31	18	56	-6.8
26	33	18	48	-20.2
27	30	18	56	-9.7
28	33	18	49	-14.3
29	39	16	57	4.0
30	30	18	42	-26.2

ることがある。今回は部品の反りを低減するために、ライントライのデータを重回帰分析で解析する。

工程では、大型射出成形機に原料ポリマーを投入し、金型へ注入・冷却して成形した後、次の加工工程に搬送している。部品の反りに影響を与える要因として、 x_1 ：原料ポリマーの射出速度(g/s)、 x_2 ：金型における冷却時間(min)、 x_3 ：金型温度(°C)を取り上げた。

表 3.3 表 3.1 を重回帰分析した結果

回帰統計	
重相関 R	0.8906
重決定 R^2	0.7932
補正 R^2	0.7693
標準誤差	3.6979
観測数	30

$F(3, 26; 0.05) = 2.975$
$t(26, 0.05) = 2.056$

分散分析表

	自由度	変動	分散	観測された分散比	有意 F
回帰	3	1363.62	454.54	33.24	0.0
残差	26	355.54	13.67		
合計	29	1719.17			

	係数	標準誤差	t	P -値	下限 95%	上限 95%
切片	-116.019	14.055	-8.255	0.000	-144.909	-87.128
x_1 : 射出速度(g/s)	1.427	0.245	5.836	0.000	0.924	1.930
x_2 : 冷却時間(min)	0.201	0.448	0.447	0.658	-0.721	1.122
x_3 : 金型温度(°C)	1.087	0.123	8.856	0.000	0.835	1.340

ら、得られた回帰式は統計的に意味のある式である。しかし、 x_2 : 冷却時間についての偏回帰係数は、 t_0 値(t 値)が0.447で t 分布の5%点2.056より小さいので、意味があるとはいえない。

観測値のグラフと残差グラフについて考察する。 y と最も相関の高い x_3 と、最も相関の低い x_2 の観測値のグラフを、それぞれ図3.4と図3.5に示す(x_1 については省略する)。なお、分析ツールで出力される図は見づらいので、横軸の目盛りやマーカーなどを見やすいものに変更してある。

図3.4では、 x_3 と y との関係は予測値と観測値ともにプロットは右肩上がりに分布している。一方、図3.5では、 x_2 と y との関係には特段の目立った

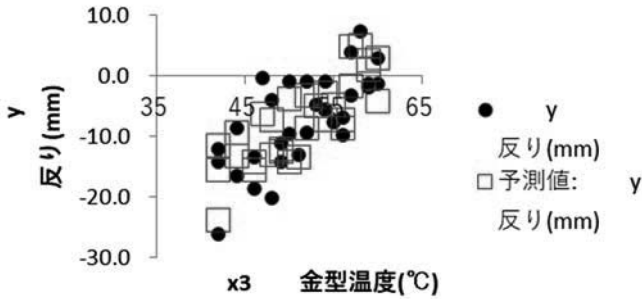


図 3.4 金型温度に関する観測値グラフ

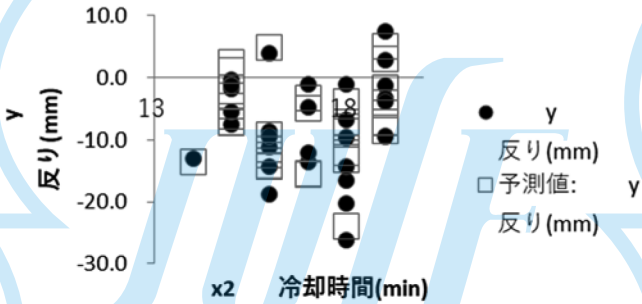


図 3.5 冷却時間に関する観測値グラフ

傾向は見えない。このことは、表 3.2 の相関係数における考察と矛盾しない。

次に残差(予測値と観測値との差)について考察する。ここでも、 y と最も相関の高い x_3 と、最も相関の低い x_2 の残差グラフを図 3.6 と図 3.7 に示す。

いずれの残差グラフも、打点に特別な傾向が見当たらないので問題はない。

(手順 4) (手順 3) で、 x_2 の冷却時間は有意と判定されなかったので、この変数を外し、説明変数を 2 つにして重回帰分析する。データは表 3.4 となる。

(手順 5) 表 3.4 のデータを重回帰分析すると、表 3.5 が得られる。 F_0 値(観測された分散比)は 51.28 で、 F 分布の 5% 点 3.354 より大きいから、得られた回帰式は統計的に意味のある式である。 F_0 値は、表 3.3 の 33.24 よりさらに大きくなり、かつ、残差も若干小さくなっている。 x_2 を説明変数から

●編著者紹介

松本哲夫(まつもと てつお)

1973年 大阪大学基礎工学部卒業

1975年 大阪大学大学院基礎工学研究科化学系修士課程修了

1975年 ユニチカ(株)入社。その後執行役員中央研究所長・技術開発本部長など歴任

1986年 技術士(経営工学部門)

2013年以降、同社顧問(現職)

[受賞歴] 科学技術賞開発部門 文部科学大臣表彰(2011)／(公社)高分子学会
フェロー表彰(2012)／(一社)日本品質管理学会 品質管理推進功労賞(2015)／(一社)日本品質管理学会 品質技術賞(2020)ほか

[著作] 『応用実験計画法』(日科技連出版社、共著、1995), 『実用実験計画法』(共立出版、共著、2005), 『実務に使える実験計画法』(日科技連出版社、共著、2012), 『実験計画法100問100答』(日科技連出版社、共著、2013)ほか多数

●著者紹介

今野勤(こんの つとむ)

1976年 早稲田大学理工学部卒業

1978年 早稲田大学院理工学研究科修士課程修了。(株)前川製作所入社

その後ヤマハ発動機(株)などを経て、

2000年 大阪大学大学院工学研究科博士後期課程修了。工学博士

龍谷大学経営学部特任教授を経て、

2008年 神戸学院大学経営学部教授(現職)

[その他] 日本科学技術連盟デミング賞実施賞委員会委員、日本経営システム学会理事、日本科学技術連盟クオリティフォーラム企画委員会委員長(2014/1～現在)など

[受賞歴] 日本経済新聞社 全国優秀先端事業所賞(1987)ーヤマハ発動機(株)エンジン組立工場建設プロジェクト FA コンピューターシステム担当(リーダー)／日経品質管理文献賞受賞(『商品企画七つ道具』全3巻(共著)、2001)／日本経営システム学会学会賞(「企画・開発力の同時実現による競争優位の実現」論文、2015)

[著作] 『商品企画七つ道具』(日科技連出版社、共著、1995), 『ヒット商品を生む商品企画七つ道具』(全3巻)(日科技連出版社、共著、2000), 『成功事例に学ぶCRMの実践手法』(日科技連出版社、共著、2003), 『実務に直結 Excelによる即効問題解決』(日科技連出版社、共著、2004年), 『QFD・TRIZ・タグチメソッドによる開発・設計の効率化』(日科技連出版社、共著、2005), 『経営系学生のための基礎統計学』(共立出版、共著、2011), 『ものづくりに役立つ経営工学の事典』(朝倉書店、共著、2014), 『文科系のための情報科学』(共立出版、共著、2017), 『データ解析による実践マーケティング』(日科技連出版社、単著、2019)ほか多数

無断使用をお断りします。日科技連出版社

Excel による多変量解析

豊富な例題ですぐに実務に活用できる

2021年8月30日 第1刷発行

編著者 松本哲夫
著者 今野勤
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 東港出版印刷

© Tetsuo Matsumoto, Tsutomu Konno 2021

ISBN 978-4-8171-9738-2

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

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で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。