



# 文化学園リポジトリ

Academic Repository of BUNKA GAKUEN

服飾文化共同研究拠点／文化ファッション研究機構

Joint Research Center for Fashion and Clothing Culture / Bunka Fashion Research Institute

文化学園大学

*Bunka Gakuen University*

文化服装学院

*Bunka Fashion College*

文化ファッション大学院大学

*Bunka Fashion Graduate University*

文化外国語専門学校

*Bunka Institute of Language*

Title	模擬実験と回帰診断
Author(s)	高橋, 武則; 楊, 国林
Citation	文化女子大学紀要. 服装学・生活造形学研究 27 (1996-01) pp.47-63
Issue Date	1996-01-31
URL	<a href="http://hdl.handle.net/10457/2355">http://hdl.handle.net/10457/2355</a>
Rights	

# 模擬実験と回帰診断

高橋武則\* 楊 国林\*\*

## Simulative Experiment and Regression Diagnostics

Takenori Takahashi and Koulin Yoh

**要 旨** コンピュータの発展のめざましい現在、統計教育においてパーソナルコンピュータの利用は当たり前のこととなりつつある。そして、このことによりかつてはごく一部の人のものと言われていた高度な解析が、多くの人に可能になってきた。中でも重回帰分析の利用の機会は飛躍的に増えている。しかしながら、とりあえず手持ちのデータをコンピュータで処理して、そのまま結論とするケースが少なくない。このため、得られた結論に信頼のおけないものや妥当でないものがしばしば登場する。本研究では、模擬実験を用いて回帰診断を行い、重回帰分析のスムーズな導入と、回帰診断を体験学習することにより、重回帰分析の使いこなし方を確実に身につけるアプローチを提案し、実施した事例を紹介する。

### 1. はじめに

重回帰分析は高度な統計手法のうちで最もよく用いられる一つである。しかしながら、意味のある重回帰式を得るためには回帰診断は不可欠である。近年回帰診断の可能なソフトが多く提供されているが、回帰診断を実際に行っている人の数は決して多くはない。本研究は教室内で手軽に出来る模擬実験を用いて回帰診断を体験学習方式で学ぶアプローチを提案し、合わせて実施事例を紹介する。回帰診断は単なる分析に終わらせてはならない。診断の過程で明らかになった外れ値についてはその原因を解析し、再発防止のための対策が必要である。また、外れ値であることが分かった場合には、そのデータをキャンセルする。しかし、そのままではデータ数が減り、あるいはバランスのとれた実験では欠測値となる。したがって、データを取り直す必要がある。その際安易な取り直しは慎むべきで、ここでは実践的で合理的な診断とデー

タの取り直しのためのアプローチを提案し、具体例を紹介する。

模擬体験（実技演習）により統計のもつ意味とその有効性を実感した学生は、統計のさらに深い理解と高度な応用のために統計理論の学習へと進むのがよい。有効性を実感した経験を有する学生の勉学意欲はきわめて高い。

### 2. 体験学習方式の重回帰分析教育

#### 2.1 基本方針

多くの人にとって自然に学ぶことが出来かつ実践力が養われる教育方式として、体験学習方式がある。できれば本物を体験することが望ましいが、多くの場合それは費用と手間の点で困難である。仮に可能であったとしても、本物による体験を通しての教育は、受講生の側にとってはしばしば苦痛になる。時間が掛かるのと、固有技術を理解しなければならないからである。このとき、模擬実験を用いて教育を実施することは大いに有効な方法である。

このような教育方式を設計するにあたり、その基本方針は以下の通りである。

(1) GLM (Graduated Level Method) とす

\* 非常勤講師 数理統計学

\*\* 本学助教授 生体情報工学

る。すなわち、前の実験はそれ自体が目的であると同時に後の実験の準備になるように設計する。ただし、スキップが可能な場合にはスキップすることもある。

(2) 必ず回帰診断を行い、外れ値の原因を追求し、再発防止をはかる。

(3) micro-CDA を基本ソフトとする。

(4) 計算および作図は原則としてコンピュータで処理し人手では行わない。

本書では計算や作図はコンピュータで行うことを前提に構成している。このため東京理科大学芳賀敏郎教授の開発された「micro-CDA」というプログラムを使用する。このプログラムはたいへん優れたプログラムであり、かつソースプログラム（言語はBASIC）が公開されている。

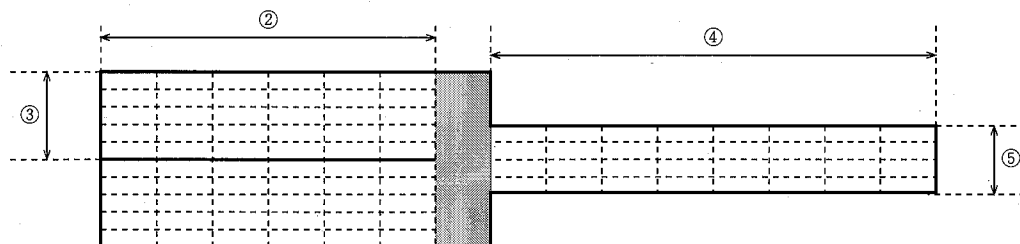
## 2.2 模擬実験

本研究では模擬実験モデルとして「紙ヘリコ

プター降下」(以後、降下のことを飛行と呼ぶ)を用いて、重回帰分析の実践的回帰診断のアプローチを展開する。このアプローチは他の模擬実験モデルでも行うことができる。著者の開発したオリジナルモデルには「コイン滑走」と「紙グライダー飛行」がある。しかし、今回の報告では重回帰分析の実践的回帰診断のアプローチそのものに焦点を合わせるので、初心者にとって実施の容易な紙ヘリコプターを用いる。なお、紙ヘリコプターそのものは著者のオリジナルではないが、これを用いてのカリキュラムの構成と演習の進め方については著者が独自に創案し設計している。

### 2.2.1 6つの説明変数

本研究では、質的変数も含む多数の説明変数がある場合の重回帰分析を扱うために、以下に示す6つの条件を取り上げる。各々の条件は図2.1に示している。



②: 翼の長さ ③: 翼の幅 ④: 足の長さ ⑤: 足の幅

: この部分は切ってはならない  
この部分に機体番号を記入する

図 2.1 機体の各部位の名称とサイズの定義

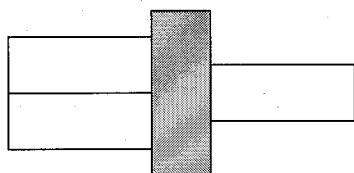


図 2.2 翼の幅が狭い場合

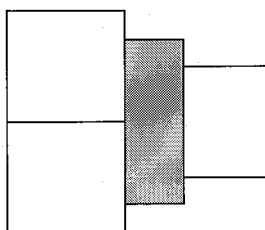


図 2.3 翼の幅が広い場合

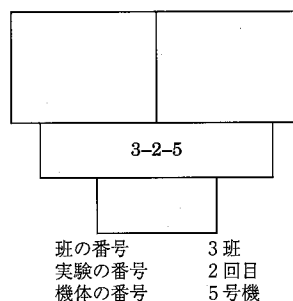


図 2.4 機体への番号の記入例

- ① 紙質
- ② 翼の長さ
- ③ 翼の幅
- ④ 足の長さ
- ⑤ 足の幅
- ⑥ 重さ

この実験では、機体の条件が複雑になるので、以下のようにして機体を準備する。

- (1) 機体番号を書く。
- (2) 用紙の上の条件欄にそれぞれの条件を記入する。
- (3) 条件に合わせてマーカー（ラインマーカー）で切るべき線を書く。
- (4) 条件とマーカーが合っているかどうか別の人がチェックする。
- (5) 1枚に3機あるので、各々の機体ごとに分離する。（多人数で作業ができる。）
- (6) マーカーの線に合わせて切る。

### 2.2.2 実験上の注意

- (1) 1回の飛行を3台のストップウォッチで計測し、メジアンを採用する。
- (2) 1機につき3回飛行させて、メジアン（メジアンのメジアン）を採用する。
- (3) 飛行の順番は必ずランダムにする。その方法として、紙を小さく切って番号を書いてくじを作り、よく混ぜてから抜いて順番を決める。
- (4) 重回帰の解析が終了するまでは、実験に用いたすべての機体をそのままの状態で保存しておく。このことにより、解析の途中で外れ値や、おかしいことが見いだされた場合に、その原因を現物に基づいて探ることができる。
- (5) 実験条件に不審な点があれば、条件が合っているかどうかを実物で確認する。
- (6) 飛行あるいは計測に不審な点があれば、再度飛行させて検討する。
- (7) 検証のための実験にあたっては、対象の実物機の複製機を作成し、これを利用することにより、実物機はできる限りその状態を保存する。

## 3. 教育カリキュラム

本章では教育全体の構造について述べる。この教育のゴールは説明変数が6個の重回帰分析である。そこへスムーズに到達するために段階的なカリキュラムを用意している。しかし、実施にあたっては大きく三つのタイプで実施することができる。

- (1) すべてのカリキュラムを実施する。
- (2) いくつかのカリキュラムをスキップする。
- (3) 必要最小限のもので実施する。

本章ではカリキュラム全体の概要を紹介し、4章と5章においては重要な実験について紹介する。なお、いくつかのカリキュラムの中では時間短縮法を紹介しているが、できればこれを用いずに正式にデータを取った方がよい。

### 3.1 実験誤差の管理：全ての基礎

実験誤差を管理することは最も重要な基礎である。この実験における誤差は大きく3種類の誤差（製作誤差、飛行誤差、計測誤差）に分類できる。これらのばらつきを分解し、中でも大きいものを低減することによって最終的に実験誤差を十分に小さなものにしてから実験を始めなければならない。誤差のばらつきが大きいと以後の実験においてしばしば次のような問題が発生する。

- (1) 異常な機体を検出することができない。
- (2) 本来有意差のあるものが有意差がないという結論になる。
- (3) 得られた回帰式が使いものにならない。（区間が広くて推定が役に立たない。）

なお、ばらつきの分解にあたっては、多くのデータをとってヒストグラムを作成するのがよいが、少ないデータで行わざるを得ない場合も考えられるので、正規確率紙を用いる場合について紹介する。

### 3.2 1つの分布：定数項のみの回帰

一つのタイプの紙ヘリコプターを多数機作成して飛行させたデータは、一つの正規分布に従

うことが期待される。これは、回帰分析としては定数項のみの回帰式と考えることができる。この段階から回帰診断を行う習慣を身につけることは重要である。また、この段階での回帰診断は直感的にも分かりやすいので学習上の効果も高い。この場合の回帰診断では以下の3点を吟味する。

- (1) 残差の吟味 (t 値でチェック)
- (2) 正規性の確認 (正規確率紙でチェック)
- (3) データの独立性 (ダービンワトソン比でチェック)

### 3.3 2つの分布：1変数の回帰（単回帰分析の基礎）

#### 3.3.1 質的変数が一つで水準が2の回帰

紙質を2種類用いて、同一のデザインの紙ヘリコプターをn機ずつ作成して飛行させたデータは、分散の等しい2つの正規分布に従うことが期待される。ただし、平均値に関しては同じ場合もあるし異なる場合もある。実験としては平均値が異なるような紙質を用いるとよい。このときの回帰診断では以下の点を吟味する。

[A] 各々の分布に関しての吟味

- (1) 残差の吟味 (t 値でチェック)
- (2) 正規性の確認 (正規確率紙でチェック)
- (3) データの独立性 (ダービンワトソン比でチェック)

[B] 2つの分布に関しての吟味

- (1) 等分散
- (2) 質的変数が一つ（2水準）の回帰のもとでの残差の吟味

【時間短縮法】3.2の実験の結果を利用して（これを第1水準のデータとして）、これとは別のもう一つの紙質のデータをとったら両者を合わせてこの実験とすることができる。

#### 3.3.2 量的変数の場合には直線回帰

重り（クリップ）を2水準（第1水準はクリップ1個、第2水準はクリップ3個）用いて、同一のデザインの紙ヘリコプターをn機ずつ作成して飛行させたデータは、分散の等しい2つの正規分布に従うことが期待される。この場合滞空時間の平均値は必ず異なり、重い場合は

短く軽い場合は長くなる。このときの回帰診断では以下の点を吟味する。

[A] 各々の分布に関しての吟味

- (1) 残差の吟味 (t 値でチェック)
- (2) 正規性の確認 (正規確率紙でチェック)
- (3) データの独立性 (ダービンワトソン比でチェック)

[B] 2つの分布に関しての吟味

- (1) 等分散
- (2) 量的変数が一つの回帰のもとでの残差の吟味

[C] 推定と確認実験

各水準の平均値を平均してクリップが2個の場合の平均値を推定する。改めてn機を作成し、クリップを2個付けて飛行させて平均値を求め、確認する。

【時間短縮法】重りのみが異なるので、n機の機体を1セットだけ作り、これをクリップの数だけ変更して3回利用する。

### 3.4 1変数実験：単回帰分析

#### 3.4.1 直線回帰（範囲が狭い場合）

##### 3.4.1.1 内挿による推定が成功する実験

翼長が6と10の場合に対して各n機ずつの実験を行い、この結果得られた回帰式で翼長8を推定し、その上で確認実験を行う。

[A] 量的変数が一つの回帰のもとでの残差の吟味

- (1) t 値による吟味
- (2) 正規確率紙による吟味
- (3) 時系列の観点の吟味 (ダービンワトソン比を含む)

[B] 推定と確認実験

得られた回帰式に翼長8を代入して推定する。推定の後で改めて翼長が8のn機を作成して飛行させ平均値を求め、確認する。

【時間短縮法】翼長10の機体がn機あるので、これをカットして翼長8の機体をn機作成する。

##### 3.4.1.2 外挿による推定が失敗する実験

すでに前の実験で翼長が6, 8, 10の結果が得られているので、このデータから求めた回帰

式を用いて翼長が15の場合の滞空時間を推定する。推定の後で改めて翼長が15の  $n$  機を作成して飛行させ平均値を求め、確認する。この実験によって外挿が危険であることを理解する。

【時間短縮法】この実験方法自体が時間短縮法になっている。正式には翼長が6, 8, 10の実験から始めなければならないのである。

### 3.4.2 曲線回帰 (範囲が広い場合)

すでに前の実験で翼長が6, 8, 10の結果が得られているので、翼長が16の実験のみを行う。翼長が6, 8, 10, 16のデータで2次の曲線回帰式を求める。

この段階でも前述の回帰診断は必ず行う。以後、回帰診断は必ず行うこととして、回帰診断に関する記述は省略する。

この式に翼長15を代入して推定する。すでに前の実験で得られている翼長15の結果と推定との比較を行う。

【時間短縮法】この実験方法自体が大幅な時間短縮法になっている。正式には翼長が6, 8, 10, 16の実験から始め、翼長15の確認実験も改めて機体を製作して行わなければならないのである。

## 3.5 2変数実験：基本的な重回帰分析/等高線図の作図が可能な場合

2変数の実験の解析は等高線図を用いて行うことができる。そして、実験の範囲が広くなければ、同じデータを用いて重回帰分析を行うことができる場合も少なくない。重回帰分析が適用できれば、分析が楽になるとともに、回帰診断が可能になる。したがって、等高線図を作成して概形を確認し単純な形状であった場合には重回帰分析を適用すべきである。しかし、範囲をかなり広くとるとははや低次の式で近似することは困難になり、その場合には等高線図を用いる方が望ましくなる。

### 3.5.1 等高線図による解析

翼長と翼幅を各々4水準ずつ取り上げて、16種類の紙ヘリコプターを製作し、飛行させて等高線図を作成する。これより、指定した条件(翼長, 翼幅)のもとでの滞空時間を予測し、

確認実験を行う。誤差が小さければ十分精度のよい推定が可能である。この場合、等高線の作図は誤差を考慮して描くことがポイントである。もし誤差が十分小さくない場合には、機体を複数作成したり、飛行を複数回試みたりして、そのもとでの平均値を用いるようにする。また、実験に不安(製作や飛行や計測にたまたま異常なことが起こる)を感じた場合にはメジアンを採用する。

### 3.5.2 重回帰分析

等高線図で用いた同じデータで重回帰分析(必ず2次の項まで用いる)を行い、両者の比較をするとよい。重回帰分析を適用すると、回帰診断が可能になり、有効な情報を手に入れることができる。

### 3.5.3 重回帰分析の適用が危険で等高線図の方が望ましい場合

応答局面の様子が複雑に変化している場合には重回帰分析の適用を避けなければならない。実験条件の範囲をかなり広くとるとこの状態になる。あえて重回帰分析を適用した場合と、等高線図で推測した場合を比較するとよい。重回帰分析の適用限界を知るためにも、これは一度トライすべきである。

## 3.6 3変数実験：重回帰分析

### 3.6.1 質的変数が入った3変数実験：2枚の等高線図の作図が可能な場合

前節で述べた2変数(翼長と翼幅)各4水準で16機の実験を2種類の紙のもとで実施する。この場合には、各々の紙ごとに等高線図で解析するとともに、各々に対して重回帰分析も実施する。2種類の紙のもとでの結果が似ている場合には、紙質を質的変数として3変数の重回帰分析を行う。この分析では32機のデータがすべて用いられる。

2つの等高線図がおおむね同じで、平均だけが異なるという場合には質的変数の意味が直感的に理解できる。

### 3.6.2 量的変数のみの3変数実験：全体を一望する等高線図が作図できない場合

3変数がすべて量的変数の（翼長，翼幅，足幅）の実験を行う。この場合3枚の等高線図で表現することも可能であるが，作図が大変であることと，3枚の等高線図から全体を読みとることも大変なので，等高線図の作図は行わない。したがって，最初から重回帰分析を適用する。ただし，範囲については広くふらないようにする。

### 3.7 6変数実験：最終目的の重回帰分析

6変数（紙質，翼長，翼幅，足長，足幅，重り）の実験を行う。各々の水準を2水準とし，L32の直交表を用いて割付て実施する。この実験では変数の数が多いので，機体製作や重りの数そして飛行の順番などを間違えないようにするために実験の管理を十分慎重に行う必要がある。

### 3.8 実施方法

以上のカリキュラムのうちで最も重要な「実験誤差の管理」と「6変数実験」の具体的な実施の方法について，前者を第4章で後者を第5章で紹介する。なお，他のカリキュラムについては紙数の都合により割愛する。

## 4. 実験誤差の管理：全ての基礎

以後の実験では各種の条件を変更して結果を観測するわけであるが，同一条件のもとでの繰り返し（製作+飛行+計測）を行った場合のばらつきは実験誤差であり，これが十分管理されないと実験で何を調べているのかが分からなくなる。紙ヘリコプターは製作を慎重に行わないとばらついてしまうし，とても軽くて脆弱なので飛行方法を注意しないと飛行がばらついてしまう。計測は滞空時間をストップウォッチで測るので一瞬のタイミングの狂いで大きなばらつきを生じてしまう。

実験を開始する前に，実験誤差を管理することはとても重要である。

## 4.1 ばらつきの分解

実験のばらつきは大きく3つの要素に分解することができる。

$$\sigma_T^2 = \sigma_P^2 + \sigma_F^2 + \sigma_M^2$$

(Total)(Production)(Flight)(Measurement)

$\sigma_T^2$ ：全体のばらつき

$\sigma_P^2$ ：製作のばらつき

$\sigma_F^2$ ：飛行のばらつき

$\sigma_M^2$ ：計測のばらつき

### 4.1.1 3種類の実験

実例で紹介する実験では，経験者が慎重に作成して飛行したものであるため，各々のばらつきの値は比較的小さい。

この実験では全部で10機作成する。

(1) 第1の実験  $\sigma_1^2$ ：n機の機体を1回飛行（1号機から9号機を使用）

$$\sigma_1^2 = \sigma_T^2 = \sigma_P^2 + \sigma_F^2 + \sigma_M^2$$

9機を各1回飛行した結果の正規確率紙を図4.1に示している。図より  $\sigma_1^2$  の推定値は  $5.7^2 = 32.49$  となっている。

(2) 第2の実験  $\sigma_2^2$ ：1機の機体をn回飛行（10号機を使用）

$$\sigma_2^2 = \sigma_F^2 + \sigma_M^2$$

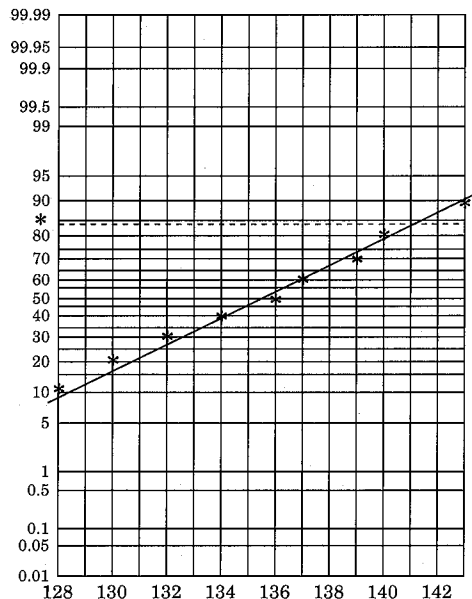


図4.1 第1の実験の正規確率紙

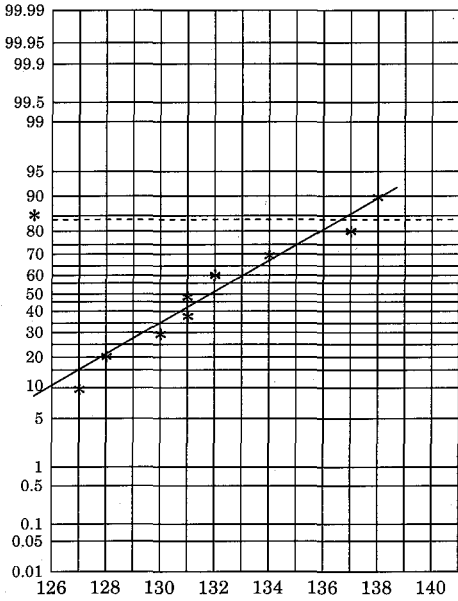


図 4.2 第 2 の実験の正規確率紙

1 機を 9 回飛行した結果の正規確率紙を図 4.2 に示している。図より  $\sigma_2^2$  の推定値は  $4.9^2 = 24.01$  となっている。

(3) 第 3 の実験  $\sigma_3^2$ : 振り子の振幅を  $n$  回計測

$$\sigma_3^2 = \sigma_M^2$$

時間間隔が一定でかつ真値が分かるものとして糸と重りで作った振り子を用いて計測誤差を評価する。振り子の一往復の時間を計測した 9 回の結果の正規確率紙を図 4.3 に示している。図より  $\sigma_3^2$  の推定値は  $4.6^2 = 21.16$  となっている。なお、振り子の周期は糸の長さを  $m$  (メートル) 単位で  $L$  とすると、次式で求めることができる。

$$T = 2\pi\sqrt{L/g} \quad g = 9.8 \text{ (m/s}^2\text{)}$$

この実験では  $L = 0.54$  (m) で、 $T = 1.48$  (s: 秒) である。図 4.3 より正規確率紙による平均は 1.49 (s) なので理論値に対してバイアスがないことも分かる。また、9 個のデータのメジアンも 1.49 (s) である。

#### 4.1.2 n 機の 2 回飛行 (1 号機-9 号機を再飛行)

飛行行為は飛行時と着床時に機体にダメージ

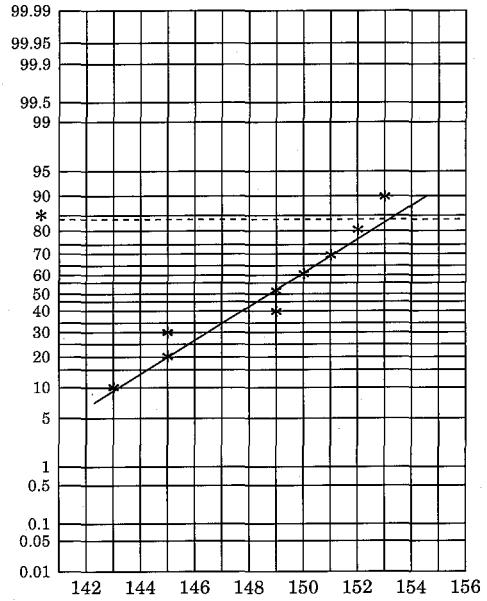


図 4.3 第 3 の実験の正規確率紙

を与える。また、飛行回数  $n$  が大きいと飛行のための取扱いの中で機体の状態を変更してしまう危険が高い。そこで、 $n$  機の各々を 2 回飛行させ各々の 2 回の結果の差の絶対値である範囲  $R$  (Range) から

$$\sigma_2^2 = \sigma_F^2 + \sigma_M^2$$

を求めることが考えられる。 $\sigma^2$  は

$$(\bar{R}/d_2)^2 = (\bar{R}/1.128)^2$$

で求められる。これで求めた結果と、1 機  $n$  回飛行で求めた結果を比較するとよい。

第 1 の実験で飛行させた 9 機を再度飛行させて求めた結果は表 4.1 である。

表 4.1 より、 $(\bar{R}/d_2)^2 = 4.7^2 = 22.28$  となっている。この結果は第 2 の実験より正規確率紙で求めた値と整合している。したがって、この実験では機体は慎重に扱われ、多数回の繰り返し

表 4.1 9 機各 2 回飛行の範囲 R

機体 No.	1	2	3	4	5	6	7	8	9	計
1 回目	131	128	138	127	131	137	130	134	132	
2 回目	138	134	135	132	122	131	127	138	137	
範囲 R	7	6	3	5	9	6	3	4	5	48



飛行にもかかわらず問題となるようなダメージを受けていないと判断できる。

#### 4.1.3 各ばらつきの把握

3種類のばらつきの値は以下のようにして求めることができる。なお、すべてを正規確率紙から求めた値で揃えて計算している。

$$\sigma_P^2 = \sigma_1^2 - \sigma_2^2 = 5.7^2 - 4.9^2 = 2.9^2$$

$$\sigma_F^2 = \sigma_2^2 - \sigma_3^2 = 4.9^2 - 4.6^2 = 1.7^2$$

$$\sigma_M^2 = \sigma_3^2 = 4.6^2$$

寄与率とは各ばらつきが全体の中で占める割合のことであり、以下のようになる。

$$\begin{aligned} \text{製作のばらつきの寄与率} &= 2.9^2 / 5.7^2 \\ &= 0.26 \text{ (26\%)} \end{aligned}$$

$$\begin{aligned} \text{飛行のばらつきの寄与率} &= 1.7^2 / 5.7^2 \\ &= 0.09 \text{ (9\%)} \end{aligned}$$

$$\begin{aligned} \text{計測のばらつきの寄与率} &= 4.6^2 / 5.7^2 \\ &= 0.65 \text{ (65\%)} \end{aligned}$$

これより、計測のばらつきが問題であることが分かる。

#### 4.1.4 ばらつきの低減

寄与率の大きなばらつきの低減をはかる。今回の結果では計測のばらつきが問題である。これに対して以下の対策を行った。

#### [A] 計測に関する対策

(1) 3人がストップウォッチで同時に計測して、メジアン(大きさの順に並べて2番目の値)を採用する。

この対策のもとで振り子の実験を行ったところ図4.4の正規確率紙の結果を得た。図より $\sigma_3^2$ の推定値は $2.4^2 = 5.76$ となっている。したがって、計測のばらつきが $4.6^2 = 21.16$ から $2.4^2 = 5.76$ になり、もとのばらつきの約2/3 ( $5.76 / 21.16 = 0.27$ )へと低減することができた。したがって、ばらつきは約1/4に低減されている。また、飛行においては以下の(2)と(3)を、製作に関しては(4)と(5)の対策を行った。

#### [B] 飛行に関する対策

(2) 飛行の際にカウントダウンする。(例えば「さんに いち はい!」)

(3) 床にお盆を置き、着床の瞬間クリップがお盆に当たり音がでるようにする。

#### [C] 製作に関する対策

(4) 紙をカットする際に鋏ではなくカッターと定規を用いて正確な直線に切る。

このことは製作時間を短縮できかつ紙に皺や歪みを与えない。

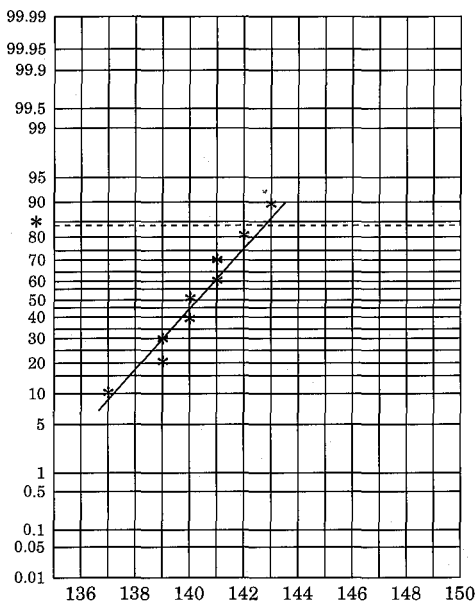


図4.4 改善された計測に関する正規確率紙

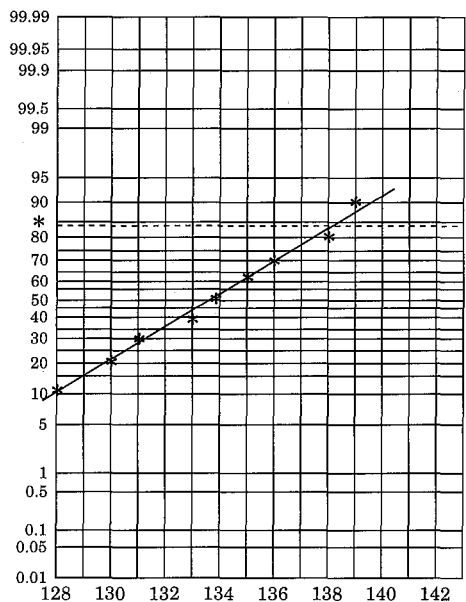


図4.5 改善された誤差に関する正規確率紙

(5) 翼を一端折りつぶしてから戻して90度にするのは紙にストレスを加えることになる。これを避けるため、直角になっている角(フロッピーの箱, 積み木など)を用いて最初から90度に折る。

以上の対策のもとで, 改めて9機作成し飛行・計測した結果の正規確率紙を図4.5に示している。全体のばらつきが $5.7^2=32.49$ から $4.7^2=22.09$ になり, もとのばらつきの約 $2/3$  ( $22.09/32.49=0.68$ ) へと低減することができた。

## 5. 6 変数の重回帰分析

重回帰分析のゴールとして, 6変数の重回帰分析を体験する必要がある。本章では, 実際に行われた重回帰分析を紹介する。なお, 分析のプロセスで重要な結果を示す。また, 「micro-CDA」の使用状況を明らかにするため, 各々の分析における対象のメニューとそのなかのどのメニューを選択したかについても合わせて示す。

### 5.1 実験の条件と実験結果

実験で取り上げる変数とその水準を表5.1に示す。この実験では質的変数を1, 量的変数を5の合計6変数を取り上げている。水準はいずれの変数も2であり, その幅は狭い範囲で取っている。したがって, 応答局面に関しては平面近似を考えている。また, 重回帰分析にあつ

ては交互作用項を取り上げずに分析している。

### 5.2 紙ヘリコプターの滞空時間に関する重回帰分析の解析例

本節では micro-CDA で解析した結果を要約して示す。ここではまず実験で得られたデータをそのまま用いて変数選択し, 推定と確認実験まで行う。その上でこの結果に基づいて次節で回帰診断を行う。

### 5.3 確認実験の実施

上記の解析の中で行った予測(データを取った範囲の内側, すなわち内挿での予測であることに注意)に対して, 実際に5機の紙ヘリコプターを作り飛行させて確認した。

紙質=コピー用紙, 翼長=7, 翼幅=6, 足長さ=9, 足幅=4, 重さ=1 → 予測値=1.67

1号機1.66, 2号機1.63, 3号機1.60, 4号機1.61, 5号機1.65 → 平均値=1.63

確認の結果は予測と比較的よく整合していることが確認できた。

### 5.4 残差の検討

全体としては解析はおおむね満足できるものであるが, 図5.8の残差で気になる機体が5機ある。また, 確認実験で飛行させた5機はすべて予測値を下まわっている点も気になる。問題となる機体のうち15号機と25号機はt値の絶対値が2を越えている。また, 4号機と8号機はt値の絶対値が極めて2に近い。したがって, この4機について検討してみた。

表 5.1 9機各2回飛行の範囲 R

No.	変数の記号	変数の内容	単位	コ ー ド
1	$X_1$	紙質	種類	①: コピー用紙, ②: レポート用紙
2	$X_2$	翼の長さ	目盛り	①: 6, ②: 9
3	$X_3$	翼の幅	目盛り	①: 5, ②: 7
4	$X_4$	足の長さ	目盛り	①: 8, ②: 10
5	$X_5$	足の幅	目盛り	①: 4, ②: 6
6	$X_6$	重さ	クリップ	①: クリップ1個, ②: クリップ2個
	Y	滞空時間	1秒	有効数字は0.01秒(計測最小単位)

表 5.2 データ表

No.	X <sub>1</sub> 紙質	X <sub>2</sub> 翼長	X <sub>3</sub> 翼幅	X <sub>4</sub> 足長	X <sub>5</sub> 足幅	X <sub>6</sub> 重量	Y 滞空時間	飛行 順番	備 考
1	①コ	①6	①5	① 8	①4	①1	1.50	28	
2	①コ	①6	①5	① 8	②6	②2	1.30	16	
3	①コ	①6	①5	②10	①4	①1	1.56	12	
4	①コ	①6	①5	②10	②6	②2	1.16	8	
5	①コ	①6	②7	① 8	①4	②2	1.47	30	
6	①コ	①6	②7	① 8	②6	①1	1.49	14	飛行ミスで再飛行
7	①コ	①6	②7	②10	①4	②2	1.33	7	
8	①コ	①6	②7	②10	②6	①1	1.58	18	タイミングが微妙
9	①コ	②9	①5	① 8	①4	①1	1.86	4	
10	①コ	②9	①5	① 8	②6	②2	1.65	15	
11	①コ	②9	①5	②10	①4	①1	1.85	23	
12	①コ	②9	①5	②10	②6	②2	1.59	25	
13	①コ	②9	②7	① 8	①4	②2	1.67	32	
14	①コ	②9	②7	① 8	②6	①1	1.78	21	
15	①コ	②9	②7	②10	①4	②2	1.84	27	
16	①コ	②9	②7	②10	①6	①1	1.67	17	
17	②レ	①6	①5	① 8	①4	①1	1.47	1	
18	②レ	①6	①5	① 8	②6	②2	1.26	24	
19	②レ	①6	①5	②10	①4	①1	1.41	10	計測ミスで再飛行
20	②レ	①6	①5	②10	②6	②2	1.29	22	
21	②レ	①6	②7	① 8	①4	②2	1.25	9	
22	②レ	①6	②7	① 8	②6	①1	1.36	3	
23	②レ	①6	②7	②10	①4	②2	1.37	11	
24	②レ	①6	②7	②10	②6	①1	1.33	20	
25	②レ	②9	①5	① 8	①4	①1	1.93	6	ドアの開閉あり
26	②レ	②9	①5	① 8	②6	②2	1.43	13	
27	②レ	②9	①5	②10	①4	①1	1.85	5	
28	②レ	②9	①5	②10	②6	②2	1.55	31	
29	②レ	②9	②7	① 8	①4	②2	1.53	29	
30	②レ	②9	②7	① 8	①6	①1	1.73	2	
31	②レ	②9	②7	②10	①4	②2	1.61	19	
32	②レ	②9	②7	②10	②6	①1	1.70	26	

Data File: KAMIHERI (n=32 p=7) 95/05/07 13:12:05

1 KAMI 16 16

) List, Fund. -stat, Graph, Var., Monit, Reg or End  
L

	紙質	翼の長さ	翼の幅	足の長さ	足の幅	重さ	滞空時間
	1	2	3	4	5	6	7
	KAMI	TNAG	THAB	ANAG	AHAB	OMOS	TIME
1	1	6.000	5.000	8.000	4.000	1.000	1.500
2	1	6.000	5.000	8.000	6.000	2.000	1.300
3	1	6.000	5.000	10.000	4.000	1.000	1.560
4	1	6.000	5.000	10.000	6.000	2.000	1.160
5	1	6.000	7.000	8.000	4.000	2.000	1.470
6	1	6.000	7.000	8.000	6.000	1.000	1.490
7	1	6.000	7.000	10.000	4.000	2.000	1.330
8	1	6.000	7.000	10.000	6.000	1.000	1.580
9	1	9.000	5.000	8.000	4.000	1.000	1.860
10	1	9.000	5.000	8.000	6.000	2.000	1.650
11	1	9.000	5.000	10.000	4.000	1.000	1.850
12	1	9.000	5.000	10.000	6.000	2.000	1.590
13	1	9.000	7.000	8.000	4.000	2.000	1.670
14	1	9.000	7.000	8.000	6.000	1.000	1.780
15	1	9.000	7.000	10.000	4.000	2.000	1.840
16	1	9.000	7.000	10.000	6.000	1.000	1.670
17	2	6.000	5.000	8.000	4.000	1.000	1.470
18	2	6.000	5.000	8.000	6.000	2.000	1.260
19	2	6.000	5.000	10.000	4.000	1.000	1.410
20	2	6.000	5.000	10.000	6.000	2.000	1.290
21	2	6.000	7.000	8.000	4.000	2.000	1.250
22	2	6.000	7.000	8.000	6.000	1.000	1.360
23	2	6.000	7.000	10.000	4.000	2.000	1.370
24	2	6.000	7.000	10.000	6.000	1.000	1.330
25	2	9.000	5.000	8.000	4.000	1.000	1.930
26	2	9.000	5.000	8.000	6.000	2.000	1.430
27	2	9.000	5.000	10.000	4.000	1.000	1.850
28	2	9.000	5.000	10.000	6.000	2.000	1.550
29	2	9.000	7.000	8.000	4.000	2.000	1.530
30	2	9.000	7.000	8.000	6.000	1.000	1.730
31	2	9.000	7.000	10.000	4.000	2.000	1.610
32	2	9.000	7.000	10.000	6.000	1.000	1.700

図 5.1 データのリスト

) List, Fund. -stat, Graph, Var., Monit, Reg or End  
F

variable	mean	minimum	maximum	std. dev.	C. V.	skewness	kurtosis
2 TNAG	7.500	6.000	9.000	1.524	0.203	0.000	-2.138*
3 THAB	6.000	5.000	7.000	1.016	0.169	0.000	-2.138*
4 ANAG	9.000	8.000	10.000	1.016	0.113	0.000	-2.138*
5 AHAB	5.000	4.000	6.000	1.016	0.203	0.000	-2.138*
6 OMOS	1.500	1.000	2.000	0.508	0.339	0.000	-2.138*
7 TIME	1.543	1.160	1.930	0.207	0.134	0.132	-0.905

Correlation Matrix

	2 TNAG	3 THAB	4 ANAG	5 AHAB	6 OMOS	7 TIME
2 TNAG	1.0000					
3 THAB	0.0000	1.0000				
4 ANAG	0.0000	0.0000	1.0000			
5 AHAB	0.0000	0.0000	0.0000	1.0000		
6 OMOS	0.0000	0.0000	0.0000	0.0000	1.0000	
7 TIME	0.7831	0.0077	0.0015	-0.2498	-0.4245	1.0000

図 5.2 基本統計量

) List, Fund. -stat, Graph, Var., Monit, Reg or End  
R

var.	St	D(S)	F	b	df(t)	sig(t)
	1.331				31	0.207
0 CON	77.499	76.169	1774.497	1.543		
1 KAMI	1.283	-0.047	1.105	(-)		
2 TNAG	0.515	-0.816	47.567	(+)		
3 THAB	1.331	-0.000	0.002	(+)		
4 ANAG	1.331	-0.000	0.000	(+)		
5 AHAB	1.248	-0.083	1.996	(-)		
6 OMOS	1.091	-0.240	6.594	(-)		

図 5.3 定数項のみの回帰

R) Sel., B-&-Cat. Sc., Cat. Pool, Residual, Pred., List, Matrix or End  
S 2

Enter	Se	R <sup>2</sup>	R* <sup>2</sup>	R** <sup>2</sup>	df(e)	sig(e)
2 TNAG	0.515	61.32%	60.03%	58.82%	30	0.131
var.	Se	D(S)	F	b		
0 CON	1.197	0.682	39.753	0.744		
1 KAMI	0.467	-0.047	2.934	(-)		
2 TNAG	1.331	0.816	47.567	0.106		
3 THAB	0.515	-0.000	0.004	(+)		
4 ANAG	0.515	-0.000	0.000	(+)		
5 AHAB	0.432	-0.083	5.579	(-)		
6 OMOS	0.275	-0.240	25.298	(-)		

図 5.4 第1の変数選択 (翼の長さ選択)

模擬実験と回帰診断

R) Sel., B-&-Cat. Sc., Cat. Pool, Residual, Pred., List, Matrix or End S 6						
Enter	Se	R <sup>2</sup>	R* <sup>2</sup>	R** <sup>2</sup>	df(e)	sig(e)
6 OMOS	0.275	79.34%	77.92%	76.58%	29	0.097
var.	Se	D(S)	F	b		
0 CON	1.197	0.922	97.248	1.004		
1 KAMI	0.228	-0.047	5.817	(-)		
2 TNAG	1.091	0.816	86.093	0.106		
3 THAB	0.275	-0.000	0.008	(+)		
4 ANAG	0.275	-0.000	0.000	(+)		
5 AHAB	0.192	-0.083	12.119	(-)		
6 OMOS	0.515	0.240	25.298	-0.173		

図 5.5 第 2 の変数選択 (重り選択)

R) Sel., B-&-Cat. Sc., Cat. Pool, Residual, Pred., List, Matrix or End S 5						
Enter	Se	R <sup>2</sup>	R* <sup>2</sup>	R** <sup>2</sup>	df(e)	sig(e)
5 AHAB	0.192	85.58%	84.04%	82.59%	28	0.083
var.	Se	D(S)	F	b		
0 CON	1.037	0.845	123.340	1.259		
1 KAMI	0.145	-0.047	8.830	(-)		
2 TNAG	1.008	0.816	119.101	0.106		
3 THAB	0.192	-0.000	0.011	(+)		
4 ANAG	0.192	-0.000	0.000	(+)		
5 AHAB	0.275	0.083	12.119	-0.051		
6 OMOS	0.432	0.240	34.997	-0.173		

図 5.6 第 3 の変数選択 (足の幅選択)

R) Sel., B-&-Cat. Sc., Cat. Pool, Residual, Pred., List, Matrix or End S 1						
Enter	Se	R <sup>2</sup>	R* <sup>2</sup>	R** <sup>2</sup>	df(e)	sig(e)
1 KAMI	0.145	89.14%	87.53%	86.01%	27	0.073
var.	Se	D(S)	F	b		
0 CON	1.027	0.883	164.871	1.297		
1 KAMI	0.192	0.047	8.830	-0.077		
2 TNAG	0.961	0.816	152.409	0.106		
3 THAB	0.144	-0.000	0.014	(+)		
4 ANAG	0.145	-0.000	0.001	(+)		
5 AHAB	0.228	0.083	15.508	-0.051		
6 OMOS	0.384	0.240	44.784	-0.173		

図 5.7 第 4 の選択 (紙質選択)

R) Sel., B-&Cat. Sc., Cat. Pool, Residual, Pred., List, Matrix or End  
R

i	Y(obs)	Y(est)	resid.	t-value	Leverage	Pred. err.
4	1.160	1.284	-0.124	-1.938	0.156	-0.147
8	1.580	1.457	0.123	1.915	0.156	0.146
15	1.840	1.705	0.135	2.131*	0.156	0.160
16	1.670	1.777	-0.107	-1.634	0.156	-0.126
25	1.930	1.802	0.128	2.017*	0.156	0.152

PRESS=0.203

図 5.8 残差の検討

R, R) Graph., Time-plot, Normal-plot, List or End  
T

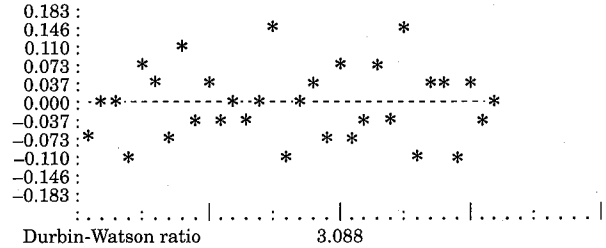


図 5.9 時系列の検討

R, R) Graph., Time-plot, Normal-plot, List or End  
N

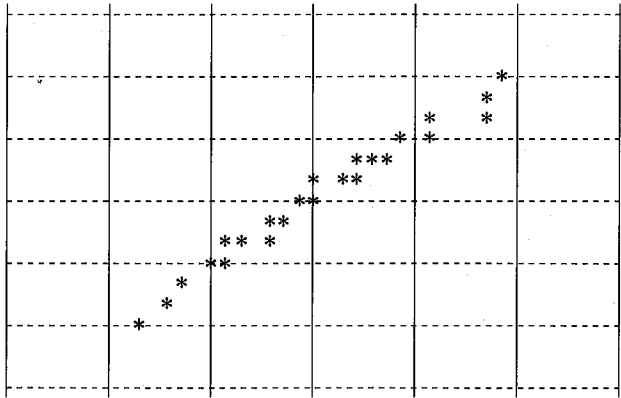


図 5.10 正規確率紙

R, R) Graph., Time-plot, Normal-plot, List or End E
R) Sel., B-&-Cat. Sc., Cat. Pool, Residual, Pred., List, Matrix or End P
R, P) セツメイ ヘンスウ ノ アタイ ヲ クウハク デ クギッテ ニュウリョク スル 1 KAMI 2 TNAG 5 AHAB 6 OMOS 1 7 4 1

Predic.	Coeff	SIGO	Y(L)	Y(U)	SIGI	Y(LL)	Y(UU)
1.666	0.128	0.026	1.612	1.719	0.078	1.506	1.825

図 5.11 予測

### 5.5 再飛行による検討

機体の問題を検討する前に、飛行・計測で問題があるかどうかの検討が必要である。最初にこの点を検討するために再飛行を試みた。飛行回数は3回で、3つの値のメジアンをとった。その結果、以下のことがわかった。

(1) 15号機と25号機の再飛行の結果は理論値(予測値)に近く、機体に問題はなく飛行・計測に問題があった。その原因は以下のように推論した。

- ① 15号機の飛行の際にカウントダウン(「3, 2, 1, はい」)の「はい」と実際の飛行開始のタイミングが若干ずれた(関係者のディスカッション)。
- ② 25号機の飛行の際に、ドアが開いて人が入室し室内の気流が乱れた。

#### (1) 変数選択の結果

R) Sel., B-&-Cat. Sc., Cat. Pool, Residual, Pred., List, Matrix or End S1
--

Enter	Se	R <sup>2</sup>	R* <sup>2</sup>	R** <sup>2</sup>	df(e)	sig(e)
1 KAMI	0.076	93.38%	92.40%	91.48%	27	0.053
var.	Se	D(S)	F	b		
0 CON	0.920	0.844	300.198	1.269		
1 KAMI	0.121	0.045	15.997	-0.075		
2 TNAG	0.832	0.756	268.916	0.103		
3 THAB	0.075	-0.001	0.277	(-)		
4 ANAG	0.076	-0.000	0.004	(-)		
5 AHAB	0.141	0.065	23.036	-0.045		
6 OMOS	0.281	0.205	72.806	-0.160		

図 5.12 2回目の重回帰分析の変数選択の結果

(2) 4号機と8号機の再飛行の結果は最初の実験結果とほとんど同様で理論値(予測値)から外れている。そこで、飛行・計測の問題ではなく機体の問題であると考えて、機体を吟味したところ以下のことがわかった。

- ③ 4号機の右翼の角度が90度よりもやや小さい(鋭角)状態であった。
- ④ 8号機は翼長が少し長めになっていた(切り残しがあった)。

したがって、15号機と25号機に関しては改めて再飛行させてデータを取り、4号機と8号機に関しては改めて新しく機体を製作した上で飛行させてデータをとった。4つのデータを直した上で、再度同様の回帰分析を行った。

### 5.6 再分析

再分析の結果の要点を以下に示す。



選択された変数や偏回帰係数の符号や値はほとんど同じであり、基本的には同様の結果であるといえよう。なお、標準偏差が小さくなり、寄与率が約6%向上した。しかし、もともと標準

偏差はたいぶ小さく、寄与率も十分高かったので大きな変化があったとはいえない。誤差が十分小さいのでちょっとした問題も見逃さないことがわかる。

(2) 残差

R) Sel., B-&-Cat. Sc., Cat. Pool, Residual, Pred., List, Matrix or End						
R						
i	Y(obs)	Y(est)	resid.	t-value	Leverage	Pred. err.
5	1.470	1.384	0.086	1.848	0.156	0.102
16	1.670	1.761	-0.091	-1.970	0.156	-0.108
26	1.430	1.526	-0.096	-2.096*	0.156	-0.114
27	1.850	1.776	0.074	1.553	0.156	0.087
29	1.530	1.616	-0.086	-1.848	0.156	-0.102

PRESS=0.107

図 5.13 2 回目の重回帰分析の残差

今度は26号機が問題となっている。しかし、誤差の標準偏差が十分小さいこと、寄与率が十分高いこと、そしてt値の絶対値がぎりぎり2を越えている程度であることから考えると、こ

こで止めてもよいと考える。

したがって、時系列プロットと正規確率紙については省略する。

(3) 予測

改めて予測を試みたら、以下の結果を得た。

R, P) セツメイ ヘンスウ ノ アタイ ヲ クウハク デ クギッテ ニュウリョク スル							
1 KAMI 2 TNAG 5 AHAB 6 OMOS							
1741							
Predic.	Coeff	SIGO	Y(L)	Y(U)	SIGI	Y(LL)	Y(UU)
1.646	0.128	0.019	1.607	1.685	0.056	1.531	1.762

図 5.14 2 回目の重回帰分析のもとでの予測

2 回目の重回帰分析の結果の方が、5.3で行った確認実験の結果により整合していることが分かった。

6. お わ り に

本研究は教室内で手軽に出来る模擬実験を用いて回帰診断を体験学習方式で学ぶアプローチを提案し、合わせて実施事例を紹介した。

【参 考 文 献】

- 1) 及川忠雄・斉藤 淳 (1992): “紙ヘリコプターによる課題解決手順の体験学習の開発—教育研修でPDCAを回す—”, 品質管理, Vol. 43, 5月臨時増刊号 pp. 410-414
- 2) Okuno, T. (1991): “Education and Training of Quality Control in the Japanese Industry”, Rep. Stat. Appl. Res., JUSE Vol. 37, No. 3-4, pp. 1-4
- 3) Kadoyama, M. (1991): “Academic QC Education”, Rep. Stat. Appl. Res., JUSE Vol. 37, No. 3-4, pp. 49-51

- 4) Kume, H., Takahashi, T., and et al (1985): Statistical Methods for Quality Improvement, The Association for Overseas Technical Scholarship
- 5) Søren Bisgaard (1989): "Teaching Statistics to Engineers", CQPI Report No. 38, The College of Engineering, University of Wisconsin-Madison
- 6) 高橋武則, 池澤辰夫 (1975): "Pole Design & Analysis (I) (競技スキー・トレーニングと品質管理)", JSQC 第9回研究発表会発表要旨集, pp. 12-15
- 7) 高橋武則, 池澤辰夫 (1976): "システムティック・スキー・トレーニング (I)", 品質管理第27巻5月臨時増刊号 pp. 167-173
- 8) 高橋武則 (1981): 「応用スキー解析」, WISE 出版局
- 9) 高橋武則 (1986): 「統計的推測の基礎」, 文化出版局
- 10) T. TAKAHASHI (1990): QC-way of Problem Solving No. 1~No. 12, Japan International Cooperation Agency
- 11) 高橋武則 (1991): "問題解決を構成する要素と構造", 品質, Vol. 21, No. 2, pp. 23-33
- 12) 高橋武則 (1991): "問題解決とPDCAサイクルの関係に関する考察", 品質, Vol. 21, No. 2, pp. 68-78
- 13) 高橋武則 (1992): 「統計的方法と管理・改善」, 品質月間委員会
- 14) 高橋武則 (1992): 「経営幹部としての統計的考え方・手法」, 日科技連QCセミナー経営幹部特別コーステキスト
- 15) Takahashi, T., and Fujino, K. (1993): "Analysis of Process Based on Fraction Defective", Proceedings of Asia Quality Control Symposium-1993 Tokyo, pp. 143-146
- 16) 高橋武則 (1993): 「統計モデルとQC的問題解決法」, 日本規格協会
- 17) Takahashi, T. (1994): "Production Game for Statistical Process Control", Proc. of International Conference on Statistics in Industry, Science and Technology, pp. 185-190
- 18) 高橋武則, 永田裕二, 佐藤浩司, 伊藤龍太郎 (1994): "統計モデルとQC的問題解決の教育に関する研究—製造工程のモデルとしてのコイン打撃ゲーム—", 日本品質管理学会第46回研究発表会発表要旨集, pp. 12-15
- 19) 高橋武則, 伊藤龍太郎 (1995): "紙グライダーの製作と飛行を用いたSQC教育の設計—統計モデルとQC的問題解決の教育に関する研究(第2報)—", 日本品質管理学会第49回研究発表会
- 20) Takahashi, T.: "Education for Quality Improvement and Statistical Process Control by Simulative Production", Proc. of Asia Quality Management Symposium-1995 Seoul, pp. 436-441
- 21) Takahashi, T.: "Statistical Games and Software Tools for Quality Assurance Based on Statistical Process Control", Proc. of 18th International Conference on Computers & Industrial Engineering (1995)
- 22) 芳賀敏郎・橋本茂司 (1990): 「実験データの解析(2)」, 日科技連出版社
- 23) 芳賀敏郎 (1991): "東京理科大学における統計教育の実態", 品質, Vol. 21, No. 1, pp. 48-53
- 24) 山田典夫 (1992): "富士ゼロックスにおける品質工学の教育・普及", 合成樹脂, Vol. 38, No. 12, pp. 39-45