

College Analysis レファレンスマニュアル

－ O R －

## 目次

1. 線形計画法 .....	1
2. 多目的線形計画法 .....	8
3. D E A.....	13
4. 待ち行列シミュレータ .....	24
<b>5. Q C 7つ道具.....</b>	<b>31</b>
6. 在庫管理シミュレータ .....	44
7. P E R T .....	52
8. システムダイナミクス .....	56
9. 不良品診断 .....	65
10. パラメータ設計 .....	70
11. オンライン品質工学.....	79
12. 異常検知 .....	88

## 1. 線形計画法

線形計画法は、一般に (2), (3) 式の制約条件のもとで、(1) 式の目的関数を最大化（最小化）する問題を解決するための手法である。

$$\text{目的関数} \quad z = \mathbf{c}^o \mathbf{x}^o \quad \text{最大化（最小化）}, \quad (1)$$

$$\text{制約条件} \quad \mathbf{A}_1 \mathbf{x}^o \leq \mathbf{b}_1, \quad (2a)$$

$$\mathbf{A}_2 \mathbf{x}^o \geq \mathbf{b}_2, \quad (2b)$$

$$\mathbf{A}_3 \mathbf{x}^o = \mathbf{b}_3, \quad (2c)$$

$$\mathbf{x}^o \geq \mathbf{0}. \quad (3)$$

ここに、 $\mathbf{x}^o (n_0 \times 1)$  は (3) 式によって各要素が非負に制限された変数行列であり、 $\mathbf{c}^o (n_0 \times 1)$ ,  $\mathbf{b}_i (m_i \times 1) \geq \mathbf{0}$ ,  $\mathbf{A}_i (m_i \times n)$  は係数行列である。

制約条件が (2a) だけの場合は、スラック変数を加え、容易に初期実行可能基底解が与えられるので、すぐにピボット操作による計算が実行出来るが、(2b) または (2c) を含む場合は、一般に実行可能基底解を得るために、2 段階法を適用する。即ち、スラック変数  $\mathbf{x}_1^s (m_1 \times 1) \geq \mathbf{0}$ ,  $\mathbf{x}_2^s (m_2 \times 1) \geq \mathbf{0}$  と人為変数  $\mathbf{x}_2^a (m_2 \times 1) \geq \mathbf{0}$ ,  $\mathbf{x}_3^a (m_3 \times 1) \geq \mathbf{0}$  を加え、(2) 式を変形する。

$$\mathbf{A}_1 \mathbf{x}^o + \mathbf{x}_1^s = \mathbf{b}_1, \quad (4a)$$

$$\mathbf{A}_2 \mathbf{x}^o - \mathbf{x}_2^s + \mathbf{x}_2^a = \mathbf{b}_2, \quad (4b)$$

$$\mathbf{A}_3 \mathbf{x}^o + \mathbf{x}_3^a = \mathbf{b}_3. \quad (4c)$$

第 1 段階の目的関数は、

$$w = \mathbf{e}_2^t \mathbf{x}_2^a + \mathbf{e}_3^t \mathbf{x}_3^a \quad \text{最小化}, \quad (5)$$

のように書ける。ここに、 $\mathbf{e}_i (m_i \times 1)$  は、全ての成分が 1 の行列である。

(5) 式は基底形式でないので、(4b), (4c) 式を加え、基底形式に変形する。

$$\text{第 1 段階目的関数} \quad w + \mathbf{d} \mathbf{x} = \mathbf{e} \mathbf{b}, \quad w : \text{最小化} \quad (6)$$

$$\text{第 2 段階目的関数} \quad z - \mathbf{c} \mathbf{x} = 0, \quad z : \text{最大化（最小化）} \quad (7)$$

$$\text{制約条件} \quad \mathbf{A} \mathbf{x} = \mathbf{b}, \quad \mathbf{x} \geq \mathbf{0}. \quad (2.8)$$

ここに、記号については以下にまとめる。

$$\mathbf{x} (n \times 1) = \begin{pmatrix} \mathbf{x}^o \\ \mathbf{x}_1^s \\ \mathbf{x}_2^s \\ \mathbf{x}_2^a \\ \mathbf{x}_3^a \end{pmatrix}, \quad \mathbf{c} (n \times 1) = \begin{pmatrix} \mathbf{c}^o \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \quad \mathbf{d} (n \times 1) = \begin{pmatrix} \mathbf{e}_2^t \mathbf{A}_2 \mathbf{e}_2 + \mathbf{e}_3^t \mathbf{A}_3 \mathbf{e}_3 \\ \mathbf{0} \\ -\mathbf{e}_2 \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix},$$

$$\mathbf{A}(m \times n) = \begin{pmatrix} \mathbf{A}_1 & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{A}_2 & \mathbf{0} & -\mathbf{I} & \mathbf{I} & \mathbf{0} \\ \mathbf{A}_3 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} \end{pmatrix}, \quad \mathbf{b}(m \times 1) = \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \mathbf{b}_3 \end{pmatrix}, \quad \mathbf{e}(m \times 1) = \begin{pmatrix} \mathbf{0} \\ \mathbf{e}_2 \\ \mathbf{e}_3 \end{pmatrix},$$

$$n = n_0 + m_1 + 2m_2 + m_3, \quad m = m_1 + m_2 + m_3. \quad (9)$$

ピボット操作を繰り返し、第 1 段階で最適解が存在すれば、 $\mathbf{x}_2^a = \mathbf{0}$ ,  $\mathbf{x}_3^a = \mathbf{0}$ ,  $w = 0$  の解を得る。このとき、その最適解についての基底行列を  $\mathbf{B}_1$ 、それに対応する係数  $\mathbf{c}$  と  $\mathbf{d}$  の基底部分を  $\mathbf{c}_{B_1}$ ,  $\mathbf{d}_{B_1}$  とすると、(6)~(8)式は以下となる。

$$\text{第 1 段階目的関数} \quad w + (\mathbf{d} - {}^t\mathbf{d}_{B_1}\mathbf{B}_1^{-1}\mathbf{A})\mathbf{x} = {}^t\mathbf{e}\mathbf{b} - {}^t\mathbf{d}_{B_1}\mathbf{B}_1^{-1}\mathbf{b} (= 0), \quad (10)$$

$$\text{第 2 段階目的関数} \quad z - ({}^t\mathbf{c} - {}^t\mathbf{c}_{B_1}\mathbf{B}_1^{-1}\mathbf{A})\mathbf{x} = {}^t\mathbf{c}_{B_1}\mathbf{B}_1^{-1}\mathbf{b}, \quad (11)$$

$$\text{制約条件} \quad \mathbf{B}_1^{-1}\mathbf{A}\mathbf{x} = \mathbf{B}_1^{-1}\mathbf{b}, \quad \mathbf{x} \geq \mathbf{0}. \quad (12)$$

第 1 段階の最適解は第 2 段階の初期可能基底解になっている。そこで、人為変数を消去して次元を縮小し、新たに、以下の問題を考える。

$$\text{目的関数} \quad z - {}^t\mathbf{c}^{(1)}\mathbf{x} = {}^t\mathbf{c}_{B_1}\mathbf{B}_1^{-1}\mathbf{b}, \quad z : \text{最大化 (最小化)} \quad (13)$$

$$\text{制約条件} \quad \mathbf{A}^{(1)}\mathbf{x} = \mathbf{b}^{(1)}, \quad \mathbf{x} \geq \mathbf{0}. \quad (14)$$

記号については改めて以下にまとめる。

$$\mathbf{A}^{(1)} = \mathbf{B}_1^{-1}\mathbf{A}, \quad \mathbf{b}^{(1)} = \mathbf{B}_1^{-1}\mathbf{b}, \quad {}^t\mathbf{c}^{(1)} = {}^t\mathbf{c} - {}^t\mathbf{c}_{B_1}\mathbf{B}_1^{-1}\mathbf{A},$$

$$\mathbf{A}(m \times n) = \begin{pmatrix} \mathbf{A}_1 & \mathbf{I} & \mathbf{0} \\ \mathbf{A}_2 & \mathbf{0} & -\mathbf{I} \\ \mathbf{A}_3 & \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad \mathbf{x}(n \times 1) = \begin{pmatrix} \mathbf{x}^o \\ \mathbf{x}_1^s \\ \mathbf{x}_2^s \end{pmatrix}, \quad \mathbf{c}(n \times 1) = \begin{pmatrix} \mathbf{c}^o \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix},$$

$$n = n_0 + m_1 + m_2, \quad m = m_1 + m_2 + m_3. \quad (15)$$

(13), (14)の問題に対して、最適解が存在する場合には、ピボット操作を行った後、最終的に以下の式を得る。

$$\text{目的関数} \quad z - {}^t\mathbf{c}^{(2)}\mathbf{x} = {}^t\mathbf{c}_{B_1}\mathbf{b}^{(1)} + {}^t\mathbf{c}_{B_2}^{(1)}\mathbf{b}^{(2)}, \quad (16)$$

$$\text{制約条件} \quad \mathbf{A}^{(2)}\mathbf{x} = \mathbf{b}^{(2)}, \quad (17)$$

$$\mathbf{A}^{(2)} = \mathbf{B}_2^{-1}\mathbf{A}^{(1)}, \quad \mathbf{b}^{(2)} = \mathbf{B}_2^{-1}\mathbf{b}^{(1)}, \quad {}^t\mathbf{c}^{(2)} = {}^t\mathbf{c}^{(1)} - {}^t\mathbf{c}_{B_2}^{(1)}\mathbf{B}_2^{-1}\mathbf{A}^{(1)}, \quad (18)$$

ここに、 $\mathbf{B}_2$  は  $\mathbf{A}^{(1)}$  についての基底行列である。

これらの式は、最終的な基底変数に対応する基底行列  $\mathbf{B} (= \mathbf{B}_1\mathbf{B}_2)$  を用いると、以下のように書けることにも注意しておく必要がある。

$$\text{目的関数} \quad z - {}^t\hat{\mathbf{c}}\mathbf{x} = {}^t\mathbf{c}_B\mathbf{B}^{-1}\mathbf{b}, \quad (19)$$

$$\text{制約条件} \quad \hat{\mathbf{A}}\mathbf{x} = \mathbf{B}^{-1}\mathbf{b}, \quad (20)$$

$${}^t\hat{\mathbf{c}} \equiv {}^t\mathbf{c}^{(2)} = {}^t\mathbf{c} - {}^t\mathbf{c}_B\mathbf{B}^{-1}\mathbf{A}, \quad \hat{\mathbf{A}} \equiv \mathbf{A}^{(2)} = \mathbf{B}^{-1}\mathbf{A}. \quad (21)$$

これより基底変数  $\mathbf{x}_B$  と非基底変数  $\mathbf{x}_N$  に対して、最適解は、 $\hat{\mathbf{x}}_B = \mathbf{B}^{-1}\mathbf{b}$ 、 $\hat{\mathbf{x}}_N = \mathbf{0}$  となる。その際、目的関数は  $\hat{z} = {}^t\mathbf{c}_B\mathbf{B}^{-1}\mathbf{b}$  となる。

線形計画問題 (1)~(3) に対して、一般性を失わず、制約式右辺の非負性  $\mathbf{b} \geq \mathbf{0}$  を除き、また変数の非負性  $\mathbf{x} \geq \mathbf{0}$  も一部取り除いて、

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11}(m_1 \times n_1) & \mathbf{A}_{12}(m_1 \times n_2) \\ \mathbf{A}_{21}(m_2 \times n_1) & \mathbf{A}_{22}(m_2 \times n_2) \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} \mathbf{b}_1(m_1 \times 1) \\ \mathbf{b}_2(m_2 \times 1) \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} \mathbf{c}_1(n_1 \times 1) \\ \mathbf{c}_2(n_2 \times 1) \end{pmatrix},$$

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_1(n_1 \times 1) \\ \mathbf{x}_2(n_2 \times 1) \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} \mathbf{y}_1(m_1 \times 1) \\ \mathbf{y}_2(m_2 \times 1) \end{pmatrix}, \quad (22)$$

の記号を用いると、主問題と双対問題はお互いに以下のように表現することが出来る。

<p>主問題（双対問題）</p> <p><math>z = {}^t\mathbf{c}\mathbf{x}</math> 最大化,</p> <p><math>\mathbf{A}_{11}\mathbf{x}_1 + \mathbf{A}_{12}\mathbf{x}_2 \leq \mathbf{b}_1,</math></p> <p><math>\mathbf{A}_{21}\mathbf{x}_1 + \mathbf{A}_{22}\mathbf{x}_2 = \mathbf{b}_2,</math></p> <p><math>\mathbf{x}_1 \geq \mathbf{0}.</math></p>	$\Leftrightarrow$	<p>双対問題（主問題）</p> <p><math>z' = {}^t\mathbf{b}\mathbf{y}</math> 最小化,</p> <p><math>{}^t\mathbf{A}_{11}\mathbf{y}_1 + {}^t\mathbf{A}_{21}\mathbf{y}_2 \geq \mathbf{c}_1,</math></p> <p><math>{}^t\mathbf{A}_{12}\mathbf{y}_1 + {}^t\mathbf{A}_{22}\mathbf{y}_2 = \mathbf{c}_2,</math></p> <p><math>\mathbf{y}_1 \geq \mathbf{0}.</math></p>
---	-------------------	--

(23)

これらの問題の間には、最適解  $\Leftrightarrow$  最適解，無限解  $\Rightarrow$  不能，不能  $\Rightarrow$  無限解または不能，の関係が存在するが、最適解の場合、双方の目的関数の値は等しくなる。

$$\hat{z} = {}^t\mathbf{c}_B\hat{\mathbf{x}}_B = {}^t\mathbf{c}_B\mathbf{B}^{-1}\mathbf{b} = {}^t\hat{\mathbf{y}}\mathbf{b} = \hat{z}'. \quad (24)$$

ここに、変数  $\mathbf{y}$  の最適解  $\hat{\mathbf{y}} = {}^t\mathbf{B}^{-1}\mathbf{c}_B$  は双対価格と呼ばれる。

次に、パラメータの変化に対する最適解の安定性を吟味する感度分析について考える。目的関数の係数の微小変化によって、最適解の係数は  ${}^t\hat{\mathbf{c}} + {}^t\Delta\hat{\mathbf{c}} = {}^t\hat{\mathbf{c}} + {}^t\Delta\mathbf{c} - {}^t\Delta\mathbf{c}_B\mathbf{B}^{-1}\mathbf{A}$  のように変化する。最大化の場合これが非正（最小化では非負）のままだと基底変数の組が維持される。即ち、非基底変数の係数について  ${}^t\hat{\mathbf{c}}_R + {}^t\Delta\mathbf{c}_R - {}^t\Delta\mathbf{c}_B\mathbf{B}^{-1}\mathbf{R} \leq \mathbf{0}$ （最小化では非負）を満たしていなければならない。ここに、 $\mathbf{R}$  は行列  $\mathbf{A}$  の非基底部分とする。この関係より、係数変化の範囲として以下の条件を得る。但し、最小化の場合は不等号の向きが逆になることに注意する。

$$\Delta\mathbf{c}_R \leq -\hat{\mathbf{c}}_R,$$

$$(\Delta\mathbf{c}_B)_i \geq (\hat{\mathbf{c}}_R)_j / (\mathbf{B}^{-1}\mathbf{R})_{ij}, \text{ for } (\mathbf{B}^{-1}\mathbf{R})_{ij} > 0,$$

$$(\Delta\mathbf{c}_B)_i \leq (\hat{\mathbf{c}}_R)_j / (\mathbf{B}^{-1}\mathbf{R})_{ij}, \text{ for } (\mathbf{B}^{-1}\mathbf{R})_{ij} < 0,$$

$$\text{制限なし} \quad \text{for } (\mathbf{B}^{-1}\mathbf{R})_{ij} = 0. \quad (26)$$

次に、係数  $\mathbf{b}$  の変化に対して、基底変数の組が変化しない範囲は基底変数の最適解が非負を維持する範囲  $\hat{\mathbf{x}}_B + \Delta\hat{\mathbf{x}}_B = \hat{\mathbf{x}}_B + \mathbf{B}^{-1}\Delta\mathbf{b} \geq \mathbf{0}$  である。これより以下の条件を得る。

$$(\Delta\mathbf{b})_i \geq -(\hat{\mathbf{x}}_B)_j / (\mathbf{B}^{-1})_{ji}, \text{ for } (\mathbf{B}^{-1})_{ij} > 0,$$

$$\begin{aligned}
 (\Delta \mathbf{b})_i &\leq -(\hat{\mathbf{x}}_B)_j / (\mathbf{B}^{-1})_{ji}, \quad \text{for } (\mathbf{B}^{-1})_{ij} < 0, \\
 &\text{制限なし} \quad \quad \quad \text{for } (\mathbf{B}^{-1})_{ij} = 0.
 \end{aligned}
 \tag{27}$$

メニュー[分析－OR－線形計画法]を選択すると、図 1 のようなメニューが表示される。その際、変数と制約式の部分は「確認」ボタンによって記入される。これは、エディターの行数と列数から求められるので、データに空白行（列）があっても構わない。

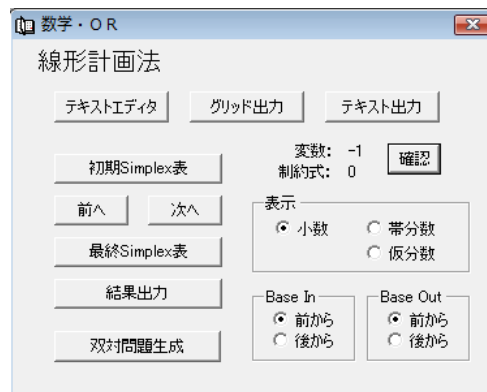


図 1 線形計画法のメニュー

データ入力は、直接グリッドエディタからでも、テキストエディタで一旦作成してグリッドエディタに移してもよい。グリッドエディタ内のデータ形式は図 2 に示す。

	x1	x2	x3	x4		
▶ 1	45	30	60	50		min
2	23	10	30	20	>=	25
3	245	150	350	380	>=	320
4	68	45	102	80	>=	75
5	1	1	1	1	=	1

1/1 (1,1)      分析: 線形計画法      備考: min,4,4

図 2 線形計画法のデータ形式

表の最上部（0 行目）は変数名、最左部（0 列目）は行番号であるが、これがなくても特に問題はない。1 行目は目的関数の係数である。最大化と最小化の区別は 1 行目最右部に、MAX または MIN の記号（大文字・小文字は区別しない）を入力して決定する。2 行目以降は制約式である。不等号・等号については BASIC で利用されるものを用いるが、不等号に等号が付いたものと付かないものは

区別しない。分析の欄には線形計画法、備考の欄には最小化を表す記号と変数の数及び制約式の数を表すものが入力されているが、これらは覚書であって何も書かなくても差し支えない。

同じものをテキストエディタで作成する場合は図 3 のようになる。

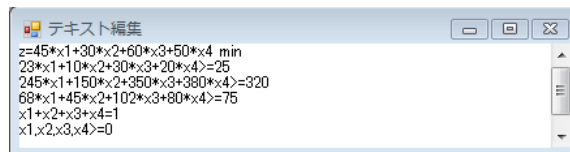


図 3 線形計画法テキスト入力データ形式

このデータは「グリッド出力」ボタンによって、図 1 のようなグリッドデータとなる。

メニューの「初期 simplex 表」は、図 2 の線形計画問題から、標準的なシンプレックス表を作成するためのもので、以下の規則に従う。

- 1) 最小化問題は目的関数に  $-1$  を掛けて最大化問題に直す。
- 2) 右辺が負の場合、両辺に  $-1$  を掛けて不等号の向きは反対にする。
- 3) 変数に非負の条件がない場合、変数を「変数名\_1」と「変数名\_2」の 2 つの非負変数に分け、元の変数をこれらの引き算で表す。
- 4) 制約式の不等号が「 $\leq$ 」の場合、スラック変数を加える。
- 5) 制約式の不等号が「 $\geq$ 」の場合、スラック変数を引き、人為変数を加える。
- 6) 制約式が等号「 $=$ 」の場合、人為変数を加える。
- 7) 人為変数を加えた場合は、2 段階法の目的関数を加える。

後に例が示されるが、変数が非負かどうかを表すために、非負条件が付かない場合、代表的な線形計画法パッケージソフトウェア LINDO<sup>®</sup> に習って変数名の最後に「!」記号を加えることにする。実行結果は、図 4 に示す。

第 1 段階シンプレックス表 STEP 0 (No. 0)														
	x1	x2	x3	x4	SL1	SL2	SL3	AR1	AR2	AR3	AR4			
▶ <-W>	-337	-206	-483	-481	1	1	1	0	0	0	0		=	-421
<-Z>	45	30	60	50	0	0	0	0	0	0	0		=	0
AR1	23	10	30	20	-1	0	0	1	0	0	0		=	25
AR2	245	150	350	380	0	-1	0	0	1	0	0		=	320
AR3	68	45	102	80	0	0	-1	0	0	1	0		=	75
AR4	1	1	1	1	0	0	0	0	0	0	1		=	1

図 4 初期シンプレックス表

ここにスラック変数には「SL 番号」、人為変数には「AR 番号」のように、自動的に変数名が付けられる。目的関数は Z で表し、2 段階法の場合は、第 1 段階の目的関数として W を導入する。ここ

では、第1段階も第2段階も最小化問題であるので、最大化問題にするために  $W$  と  $Z$  両方に  $-1$  が掛かっている。最小化問題はそのままでも計算上特に問題はないが、ここでは統一性を重視した。

図 2.2 のメニュー画面で「Pivot 操作の順次出力」を選択すると、シンプレックス法の計算過程を 1 ステップ毎表示する。ピボット操作のステップは、初期を 0 として、2 段階法の場合も継続して数えている。続けて選択することによって、利用者は基底変数の移り変わり等を簡単に見ることが出来る。最終的なシンプレックス表が表示された後は最初の段階に戻る。

計算の途中で、解の存在しない場合や、無限大の解が存在する場合はメッセージにより知らせる。その際、利用者はシンプレックス表により、その理由を見ることが出来る。

メニューで「結果出力」を選択すると、最終的なシンプレックス表と実行結果が図 5、図 6 のように表示される。

	x1	x2	x3	x4	SL1	SL2	SL3		
▶ <-Z>	0	6.507937	0	0	1.190476	0.063492	0	=	-52.142857
SL3	0	-25.380952	0	0	-2.642857	-0.147619	1	=	9.357143
x4	0	0.555556	0	1	0.083333	-0.005556	0	=	0.25
x3	0	-1.619048	1	0	-0.107143	-0.002381	0	=	0.392857
x1	1	2.063492	0	0	0.023810	0.007937	0	=	0.357143

図 5 最終結果のシンプレックス表

LP 最適解 TOTAL STEP 5 最小値 52.1429			
変数	値	被約費用	
x1	0.3571	0.0000	
x2	0.0000	6.5079	
x3	0.3929	0.0000	
x4	0.2500	0.0000	
行	スラック	双対価格	
1	0.0000	1.1905	
2	0.0000	0.0835	
3	0.3571	0.0000	
4	0.0000	2.0635	
係数範囲			
変数	現在値	増加上限	減少下限
x1	45.0000	3.1538	-infinity
x2	30.0000	infinity	-6.5079
x3	60.0000	infinity	-4.0196
x4	50.0000	11.7143	-11.4286
右辺範囲			
行	現在値	増加上限	減少下限
1	25.0000	3.0000	-3.5405
2	320.0000	45.0000	-45.0000
3	75.0000	3.3571	-infinity
4	1.0000	0.1289	-0.1023

図 6 最終結果の出力

最終結果の出力は、LINDO の出力に習った。以下にその定義を与える。

「最適解」は変数  $\mathbf{x}^0$  の最適値であり、「被約費用」は目的関数における  $\mathbf{x}^0$  の係数である。「余剰」は制約式の左辺と右辺との差を表している。これは、初期可能基底変数の値であるため、スラック変



数の最適値と等号制約の人為変数の値 ( $= 0$ ) である。「双対価格」は  ${}^t\pi = {}^t\mathbf{c}_B\mathbf{B}^{-1}$  で定義され、双対問題の最適解でもあるが、制約式の右辺値の微小変化における目的関数の変化率としても解釈される。

感度分析において、基底変数の組が維持される目的関数の係数の範囲は、(26)式を用いて図 1.5 の「係数範囲」の部分に、また右辺係数の範囲は、(27)式を用いて「右辺範囲」の部分に表示される。

主問題に対する双対問題は、図 1 のメニュー画面で「双対問題生成」を選択すると、エディター上に 2 ページ目として生成される。双対問題は同じメニュー上のオプションボタンを選択することにより、主問題と同じように解くことが出来る。これによって実際の数値で、主問題と双対問題の関係を確認することが出来る。

図 7 や図 8 のように変数名  $x2!$ ,  $y2!$  のように後に感嘆符が付いた変数があるが、これは以前に説明したように、非負条件の付かない変数である。等号条件に対応する双対問題の変数がこれに相当する。

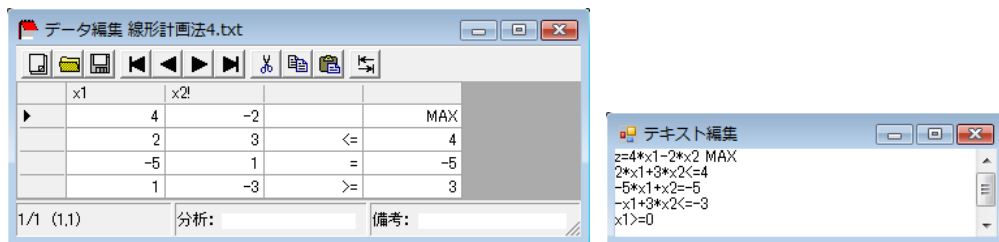


図 7 主問題

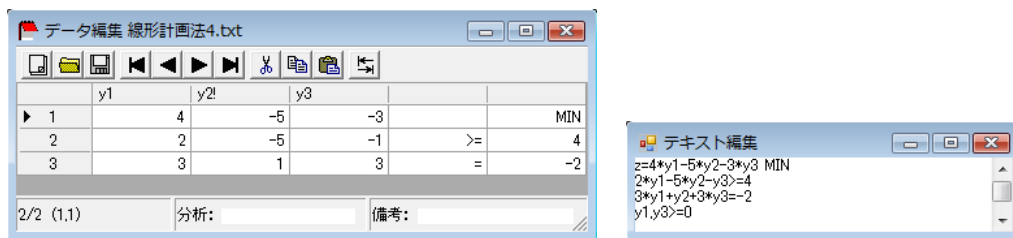


図 8 双対問題

## 2. 多目的線形計画法

線形計画法は与えられた線形の制約条件のもとで、線形の目的関数を最大化（最小化）する問題を解決するための手法であり、一般に以下の形で書くことができる<sup>5)</sup>。

目的関数  $z = \mathbf{c}^T \mathbf{x}$  最大化（最小化）

制約条件  $\mathbf{A}_1 \mathbf{x} \leq \mathbf{b}_1, \mathbf{A}_2 \mathbf{x} \geq \mathbf{b}_2, \mathbf{A}_3 \mathbf{x} = \mathbf{b}_3, \mathbf{x} \geq \mathbf{0}$

ここに、 $\mathbf{x}(n \times 1)$  は各要素が非負に制限された変数行列であり、 $\mathbf{c}(n \times 1)$ ,  $\mathbf{b}_i(m_i \times 1) \geq \mathbf{0}$ ,  $\mathbf{A}_i(m_i \times n)$  ( $i=1,2,3$ ) は係数行列である。

この問題は生産計画や輸送計画など様々な分野で利用されているが、現実には目的関数が唯一であるようなケースはむしろ珍しく、いくつかの目的関数ある程度最適化するような解を求めることが多い。このような問題に答えるのが多目的線形計画法である。多目的線形計画法では線形計画法の制約条件はそのまま、目的関数の数を2つ以上考える。ここではこの問題に対するグローバル評価法と呼ばれる解法を説明する<sup>4)</sup>。

今我々は以下の多目的線形計画問題を考える。

目的関数  $z^a = \mathbf{c}^a \mathbf{x}$  ( $a=1,2,\dots,p$ ) 最大化（最小化）

制約条件  $\mathbf{A}_1 \mathbf{x} \leq \mathbf{b}_1, \mathbf{A}_2 \mathbf{x} \geq \mathbf{b}_2, \mathbf{A}_3 \mathbf{x} = \mathbf{b}_3, \mathbf{x} \geq \mathbf{0}$

最初にまず目的関数  $a$  の最適解を考えて、その解を用いた  $z^a$  の最適値を  $z^{a*}$  とする。目的関数  $a$  について、ある解  $\mathbf{x}$  の充足率  $r^a$  を以下のように定義する。

$$r^a = 1 - \left| (z^{a*} - \mathbf{c}^a \mathbf{x}) / z^{a*} \right|$$

ここで最適解の目的関数の値が0に近い場合、充足率が負になる場合があるので気を付ける。目的関数の値は目的関数に定数項を加えれば自由に変更できるが、これによって充足率の値も変わってくるのでこの方法にも問題がある。今後充足率の定義についても検討しなければならない。

我々は各目的関数の  $1 - \text{充足率}$  の値の合計を新しい目的関数にして、上の制約条件で線形計画問題を考える。

目的関数  $z = \sum_{a=1}^p (1 - r^a)$  最小化

制約条件  $\mathbf{A}_1 \mathbf{x} \leq \mathbf{b}_1, \mathbf{A}_2 \mathbf{x} \geq \mathbf{b}_2, \mathbf{A}_3 \mathbf{x} = \mathbf{b}_3, \mathbf{x} \geq \mathbf{0}$

またこの目的関数にウェイト  $w^a$  ( $a=1,2,\dots,p$ ) を付けてもよい。

目的関数  $z = \sum_{a=1}^p (1 - w^a r^a)$  最小化

制約条件  $\mathbf{A}_1 \mathbf{x} \leq \mathbf{b}_1, \mathbf{A}_2 \mathbf{x} \geq \mathbf{b}_2, \mathbf{A}_3 \mathbf{x} = \mathbf{b}_3, \mathbf{x} \geq \mathbf{0}$

これらの線形計画問題を最良化問題と呼び、その最適解  $\mathbf{x}^*$  を最良化妥協解と呼ぶ。

次に具体的な実行例について説明する。図1に多目的線形計画法のメニュー画面を示す。

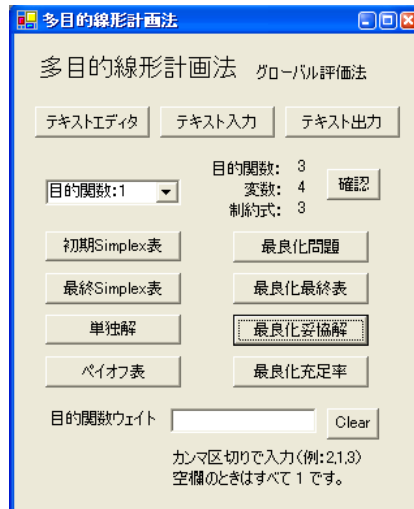


図 1 多目的線形計画法メニュー画面

線形計画問題の式を表すデータは基本的に表形式で入力するが、必要に応じてテキスト形式でも入力できる。メニューの「テキストエディタ」ボタンでテキスト入力用のエディターが表示され、式をそこに書き込むことができる。「テキスト入力」ボタンはテキスト形式を表形式に変換し、「テキスト出力」ボタンは表形式をテキスト形式に変換する役割を持つ。表形式のデータの例を図 2 に示す。

	x1	x2	x3	x4	
	4	5	3	2	
	2	2	2	1	max
	1	3	3	2	max
	1	5	3	4	<= 50
	3	3	2	5	= 30
	2	-4	6	1	>= 20

図 2 表形式データ画面

また、テキスト形式のデータの例を図 3a、図 3b に示すが、どちらの形式でも同じ表形式のデータを作り出す。テキスト出力のデフォルトは図 3a の形式である。

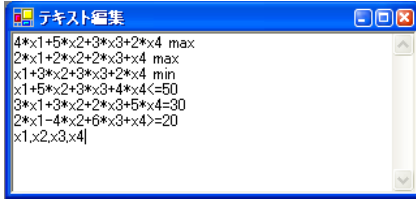


図 3a テキスト形式データ画面 1

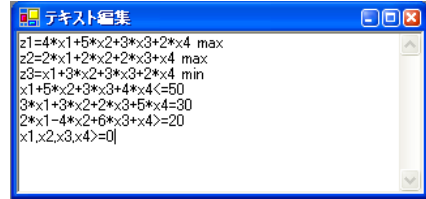


図 3b テキスト形式データ画面 2

上のデータから各目的関数についての単独解を求めるには、まず左上のコンボボックスで目的関数を選択し、「初期 Simplex 表」、「最終 Simplex 表」、「単独解」ボタンを利用する。これらの表示画面をそれぞれ図 4、図 5、図 6 に示す。これらは線形計画法における出力形式に等しい。

	x1	x2	x3	x4	SL1	SL2	AR1	AR2		
<W>	-5	1	-8	-6	0	1	0	0	=	-50
<Z>	-4	-5	-3	-2	0	0	0	0	=	0
SL1	1	5	3	4	1	0	0	0	=	50
AR1	3	3	2	5	0	0	1	0	=	30
AR2	2	-4	6	1	0	-1	0	1	=	20

図 4 目的関数 1 の初期 Simplex 表画面

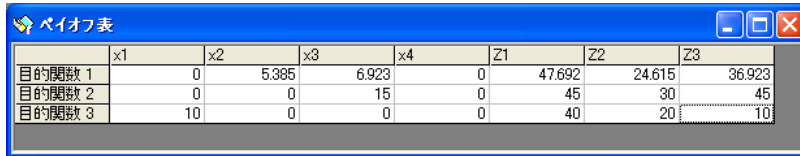
	x1	x2	x3	x4	SL1	SL2			
<Z>	0.769231	0	0	6.038462	0	0.038462	=		47.692308
SL1	-3.769231	0	0	-4.038462	1	-0.038462	=		2307.692
x2	0.538462	1	0	1.076923	0	0.076923	=		5.384615
x3	0.692308	0	1	0.884615	0	-0.115385	=		6.923077

図 5 目的関数 1 の最終 Simplex 表画面

LP 最適解 TOTAL STEP 4 最大値 47.6923			
変数	値	節約費用	
x1	0.0000	0.7692	
x2	5.3846	0.0000	
x3	6.9231	0.0000	
x4	0.0000	6.0385	
行	スラック	双対価格	
1	2.3077	0.0000	
2	0.0000	1.6154	
3	0.0000	0.0385	
係数範囲	現在値	増加上限	減少下限
x1	4.0000	0.7692	-infinity
x2	5.0000	infinity	-0.5000
x3	3.0000	0.3333	-1.1111
x4	2.0000	6.0385	-infinity

図 6 目的関数 1 の単独解画面

それぞれの単独解で他の目的関数がどのような値になるのかを表すものがペイオフ表である。メニューの「ペイオフ表」ボタンをクリックすると図 7 のようなペイオフ表画面が表示される。

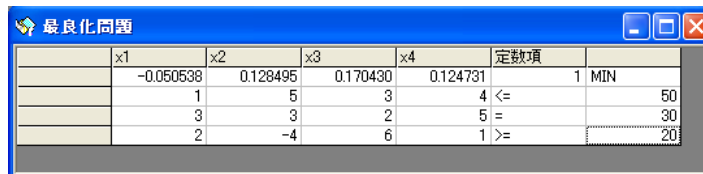


	x1	x2	x3	x4	Z1	Z2	Z3
目的関数 1	0	5.385	6.923	0	47.692	24.615	36.923
目的関数 2	0	0	15	0	45	30	45
目的関数 3	10	0	0	0	40	20	10

図 7 ペイオフ表画面

例えばここで行名が目的関数 1 の行は目的関数 1 の単独解を使った 3 つの目的関数の値を表しており、逆に列名が Z1 の列は 3 つの目的関数の最適解を使った目的関数 1 の値を表している。

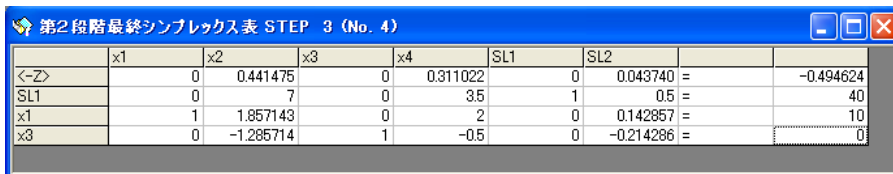
これらの単独解を使った最良化問題のデータは「最良化問題」ボタンをクリックして図 8 のように得られる。



	x1	x2	x3	x4	定数項	
	-0.050538	0.128495	0.170430	0.124731	1	MIN
	1	5	3	4	<=	50
	3	3	2	5	=	30
	2	-4	6	1	>=	20

図 8 最良化問題画面

この最良化問題の解は、「最良化最終表」ボタンと「最良化妥協解」ボタンによって得られる。それらを図 9 と図 10 に示す。この最良化妥協解による各目的関数の充足率はメニューの「最良化充足率」ボタンで求めることができる。その例を図 11 に示す。



	x1	x2	x3	x4	SL1	SL2		
<-Z>	0	0.441475	0	0.311022	0	0.043740	=	-0.494624
SL1	0	7	0	3.5	1	0.5	=	40
x1	1	1.857143	0	2	0	0.142857	=	10
x3	0	-1.285714	1	-0.5	0	-0.214286	=	0

図 9 最良化問題最終シンプレックス表画面



LP 最適解 TOTAL STEP 4 最小値 0.4946		
変数	値	被約費用
x1	10.0000	0.0000
x2	0.0000	0.4415
x3	0.0000	0.0000
x4	0.0000	0.3110
行	スラック	双対価格
1	40.0000	0.0000
2	0.0000	-0.0480
3	0.0000	0.0437

図 10 最良化妥協解画面



目的関数	最大/最小	最適化値	妥協値	充足率
目的関数 1	最大化	47.692	40	0.839
目的関数 2	最大化	30	20	0.667
目的関数 3	最小化	10	10	1

図 11 最良化妥協解と充足率画面

### 3. DEA

DEA (Data Envelopment Analysis) は事業体に関して、得意な分野を評価するという姿勢で、その効率性を求める手法である。ここでは効率性を検討する各事業体を DMU (Decision Making Unit) と呼び、効率性は  $r$  個の入力変数の線形結合と  $s$  個の出力変数の線形結合の比として表わされる。今、DMU の全数を  $n$  として、DMU <sub>$i$</sub>  ( $i$  番目の DMU) の入力と出力をそれぞれ、

$${}^t\mathbf{x}_i = (x_{i1} \quad x_{i2} \quad \cdots \quad x_{ir}), \quad {}^t\mathbf{y}_i = (y_{i1} \quad y_{i2} \quad \cdots \quad y_{is}),$$

入力と出力に掛かるパラメータをそれぞれ、

$${}^t\mathbf{v} = (v_1 \quad v_2 \quad \cdots \quad v_r), \quad {}^t\mathbf{u} = (u_1 \quad u_2 \quad \cdots \quad u_s),$$

として、その効率性を  $\theta_i = {}^t\mathbf{u}\mathbf{y}_i / {}^t\mathbf{v}\mathbf{x}_i$  で与える。但し、効率性を計算している DMU を  $o$  として、 $0 \leq \theta_i \leq 1$  の範囲で  $\theta_o$  を最大化するようにパラメータ  $\mathbf{v}$ ,  $\mathbf{u}$  を決定する。それ故、効率性を計算する DMU 毎にパラメータの値も変わってくる。このパラメータの決定方法が、最初に述べた得意な分野を評価する姿勢を表わしている。さて、ここまで述べたことを分数計画問題として以下のようにまとめておく。

分数計画問題

$$\text{目的関数} \quad z = {}^t\mathbf{u}\mathbf{y}_o / {}^t\mathbf{v}\mathbf{x}_o \quad \text{最大化}$$

$$\text{制約式} \quad {}^t\mathbf{u}\mathbf{y}_i / {}^t\mathbf{v}\mathbf{x}_i \leq 1 \quad (i = 1, \dots, n), \quad \mathbf{u} \geq \mathbf{0}, \quad \mathbf{v} \geq \mathbf{0}$$

この分数計画問題は、以下の線形計画問題として考えることができる。

線形計画問題 (主問題)

$$\text{目的関数} \quad z = {}^t\mathbf{u}\mathbf{y}_o \quad \text{最大化}$$

$$\text{制約式} \quad {}^t\mathbf{v}\mathbf{x}_o = 1, \quad -{}^t\mathbf{v}\mathbf{X} + {}^t\mathbf{u}\mathbf{Y} \leq \mathbf{0}, \quad \mathbf{u} \geq \mathbf{0}, \quad \mathbf{v} \geq \mathbf{0}$$

但し、 $\mathbf{X} = (\mathbf{x}_1 \quad \mathbf{x}_2 \quad \cdots \quad \mathbf{x}_n)$ ,  $\mathbf{Y} = (\mathbf{y}_1 \quad \mathbf{y}_2 \quad \cdots \quad \mathbf{y}_n)$  である。

この線形計画問題は通常以下の双対問題から解が求められる。

線形計画問題 (双対問題)

$$\text{目的関数} \quad z' = \theta \quad \text{最小化}$$

$$\text{制約式} \quad \theta \mathbf{x}_o - \mathbf{X}\boldsymbol{\lambda} \geq \mathbf{0}, \quad -\mathbf{y}_o + \mathbf{Y}\boldsymbol{\lambda} \geq \mathbf{0}, \quad \boldsymbol{\lambda} \geq \mathbf{0}$$

ここに、双対問題の変数を、 $\theta$  と  ${}^t\boldsymbol{\lambda} = (\lambda_1 \quad \lambda_2 \quad \cdots \quad \lambda_n)$  で与えた。ところでこの双対問題において、 $\theta = 1$ ,  $\lambda_o = 1$ ,  $\lambda_i = 0$  ( $i \neq o$ ) は制約式を満たすので解は必ず存在することが分り、このことから必ず  $\theta \leq 1$  となる。

さて、以下のような集合  $P$  を考える。

$$P = \{(\mathbf{x}, \mathbf{y}) \mid \mathbf{x} - \mathbf{X}\boldsymbol{\lambda} \geq \mathbf{0}, -\mathbf{y} + \mathbf{Y}\boldsymbol{\lambda} \geq \mathbf{0}, \boldsymbol{\lambda} \geq \mathbf{0}\}$$

今、効率を測定する DMU の入力と出力  $\mathbf{x}_o$ ,  $\mathbf{y}_o$  について、 $(\theta \mathbf{x}_o, \mathbf{y}_o) \in P$  であれば、双対問題の

制約式を満たすことが分かる。 $\mathbf{x}_o$ を $\theta \mathbf{x}_o$ として、集合 $P$ の境界まで縮めたときの倍率 $\theta^*$ が最小の目的関数値となっている。この集合を生産可能集合と呼ぶ。

$\theta$ を最小化する最適解でも余剰の自由度は残る。そこで余剰 $\mathbf{s}_x = \theta^* \mathbf{x}_o - \mathbf{X}\boldsymbol{\lambda}$ 及び、 $\mathbf{s}_y = -\mathbf{y}_o + \mathbf{Y}\boldsymbol{\lambda}$ の成分の合計が最大となるように再度線形計画問題を解く。

線形計画問題（余剰の最大化）

目的関数  $w = {}^t \mathbf{e}_x \mathbf{s}_x + {}^t \mathbf{e}_y \mathbf{s}_y$  最大化

制約式  $\mathbf{s}_x = \theta^* \mathbf{x}_o - \mathbf{X}\boldsymbol{\lambda}$ ,  $\mathbf{s}_y = -\mathbf{y}_o + \mathbf{Y}\boldsymbol{\lambda}$ ,  $\boldsymbol{\lambda} \geq \mathbf{0}$ ,  $\mathbf{s}_x \geq \mathbf{0}$ ,  $\mathbf{s}_y \geq \mathbf{0}$

ここに、 ${}^t \mathbf{e}_x = (1 \ 1 \ \cdots \ 1) [r \text{ 成分}]$ ,  ${}^t \mathbf{e}_y = (1 \ 1 \ \cdots \ 1) [s \text{ 成分}]$ である。この解 $\boldsymbol{\lambda}^*$ ,  $\mathbf{s}_x^*$ ,  $\mathbf{s}_y^*$ を最大スラック解と呼ぶ。 $\theta^* = 1$ ,  $\mathbf{s}_x^* = \mathbf{0}$ ,  $\mathbf{s}_y^* = \mathbf{0}$ のとき、観測しているDMU<sub>o</sub>を効率的であるといい、これ以外るとき非効率的であるという。

DMU<sub>o</sub>の改善点を求めるために、入力過剰量と出力不足量 $\Delta \mathbf{x}$ ,  $\Delta \mathbf{y}$ を求める。

$$\begin{aligned} \Delta \mathbf{x} &= \mathbf{x}_o - \mathbf{X}\boldsymbol{\lambda}^* = (1 - \theta^*)\mathbf{x}_o + \mathbf{s}_x^* \\ \Delta \mathbf{y} &= -\mathbf{y}_o + \mathbf{Y}\boldsymbol{\lambda}^* = \mathbf{s}_y^* \end{aligned}$$

これによって効率性改善の示唆を得ることができる。

対象となるDMUの特徴と改善点を考える際に、似たDMUで自分より優れたものを知ることは意味がある。DMU<sub>o</sub>が非効率的であるとき、以下の $E_o$ をDMU<sub>o</sub>に対する優位集合という。

$$E_o = \{j \mid \lambda_j^* > 0, j = 1, \dots, n\}$$

優位集合に属する活動は効率的であることが知られている。

生産可能集合を直感的に理解するために1入力、1出力の場合を図で表わしてみる。この場合、生産可能集合は以下となる。

$$P = \{(x, y) \mid x \geq x_1 \lambda_1 + \cdots + x_n \lambda_n, y \leq y_1 \lambda_1 + \cdots + y_n \lambda_n, \boldsymbol{\lambda} \geq \mathbf{0}\}$$

この範囲を図で表わすと、図1の網掛けの部分になる。また、DMU<sub>o</sub>の効率 $\theta^*$ は図1に示した $x$ 座標の値を用いて、 $\theta^* = x'/x$ で与えられる。

図 2.1 CCR モデルの生産可能集合



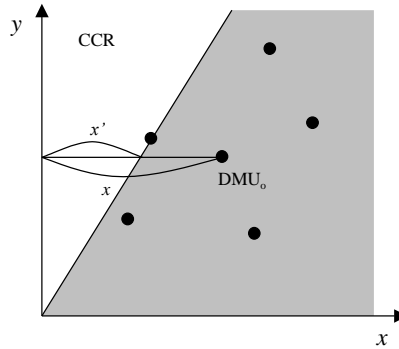


図 1 CCR モデルの生産可能集合

様々な基本的なモデルはこの生産可能集合に以下の条件を付けて得られる。

$$L \leq {}^t \mathbf{e} \boldsymbol{\lambda} \leq U \quad (0 \leq L \leq 1, U \geq 1)$$

#### CCR モデル ( $L=0, U=\infty$ )

元々の生産可能集合による効率決定モデルを CCR モデルと呼ぶ。これは生産規模によって効率に優劣が生じない、規模の収穫が一定のモデルである。

#### BCC モデル ( $L=1, U=1$ )

生産可能集合に  ${}^t \mathbf{e} \boldsymbol{\lambda} = 1$  の条件を付けたものが BCC モデルである。この生産可能集合は図 2 で表わされる。

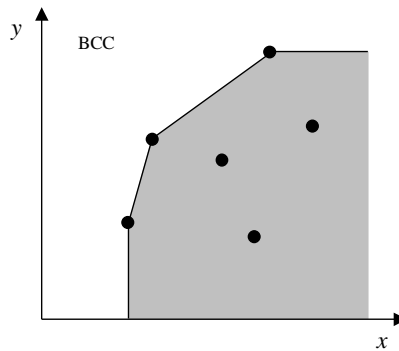


図 2 BCC モデルの生産可能集合

#### IRS モデル ( $L=1, U=\infty$ )

生産可能集合に  ${}^t \mathbf{e} \boldsymbol{\lambda} \geq 1$  の条件を付けたものが IRS (Increasing Returns to Scale) モデルで、規模の収穫が増加することを想定したモデルである。この生産可能集合は図 3 で表わされる。

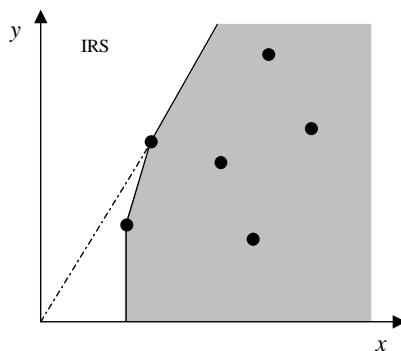


図 3 IRS モデルの生産可能集合

**DRS モデル ( $L=0, U=1$ )**

生産可能集合に  $\sum \lambda_j = 1$  の条件を付けたものが DRS (Decreasing Returns to Scale) モデルで、規模の収穫が減少することを想定したモデルである。この生産可能集合は図 4 で表わされる。

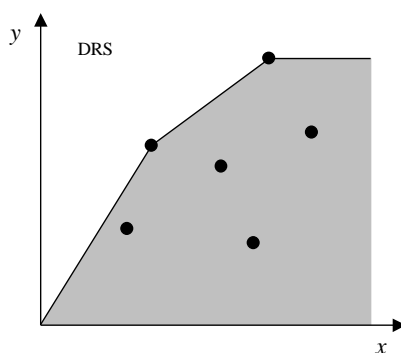


図 4 DRS モデルの生産可能集合

**GRS モデル ( $0 \leq L \leq 1, U \geq 1$ )**

下限と上限に上記の範囲で任意の値を取ったものを GRS (General Returns to Scale) モデルという。これは一般的なモデルで、利用者が  $L$  と  $U$  の値を与える。

モデル毎に、効率の測り方として、図 5 のように  $y$  軸を用いた方法も考えられる。これは出力型モデルと呼ばれ、それぞれのモデル名の後に O という文字を付け、例えば CCRO モデルのように表わす。

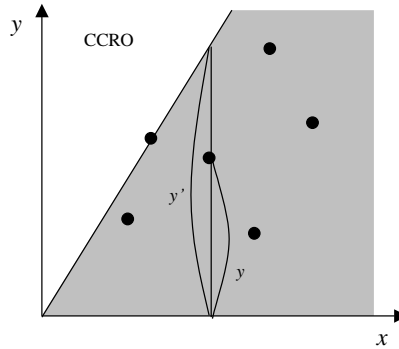


図 5 CCRO モデルの生産可能集合

CCRO モデルの場合、線形計画問題は以下のように与えられる。

線形計画問題（主問題）

目的関数  $z = \mathbf{v}^t \mathbf{x}_o$  最小化

制約式  $\mathbf{u}^t \mathbf{y}_o = 1, -\mathbf{v}^t \mathbf{X} + \mathbf{u}^t \mathbf{Y} \leq \mathbf{0}, \mathbf{u} \geq \mathbf{0}, \mathbf{v} \geq \mathbf{0}$

線形計画問題（双対問題）

目的関数  $z' = \eta$  最大化

制約式  $\mathbf{x}_o - \mathbf{X}\boldsymbol{\lambda} \geq \mathbf{0}, -\eta \mathbf{y}_o + \mathbf{Y}\boldsymbol{\lambda} \geq \mathbf{0}, \boldsymbol{\lambda} \geq \mathbf{0}$

出力型モデルの効率は  $1/\eta$  で与えられる。特に CCRO モデルの場合に限り、効率は CCR モデルと一致する。他のモデルでは双対問題に  $\boldsymbol{\lambda}$  についての制約が付く。

### DEA の解釈と制御不能変数

新しく制御不能変数を導入した場合の DEA についてプログラムを作成したので、理論的な背景を再度説明しておく。

#### 制御不能変数を含まないモデル

刀根先生の DEA の教科書ではなかなか分かりづらいところを 2 入力 1 出力の問題で考えてみる。

CCR モデルでは、問題は以下の通りである。

目的関数

$$z = \frac{u_1 y_{o1}}{v_1 x_{o1} + v_2 x_{o2}} \quad \text{最大化}$$

制約条件

$$\frac{u_1 y_{i1}}{v_1 x_{i1} + v_2 x_{i2}} \leq 1$$

$$u_1, v_1, v_2 \geq 0$$

これに  $v_1x_{o1} + v_2x_{o2} = 1$  の条件を付けて (入力モデル)、以下の線形計画問題に書き換える。

目的関数

$$z = u_1y_{o1} \quad \text{最大化}$$

制約条件

$$\begin{aligned} v_1x_{o1} + v_2x_{o2} &= 1 \\ -v_1x_{i1} - v_2x_{i2} + u_1y_{i1} &\leq 0 \quad (i=1, \dots, n) \\ u_1, v_1, v_2 &\geq 0 \end{aligned}$$

これを線形計画問題の主問題として、双対問題を作成する。

目的関数

$$z' = \theta \quad \text{最小化}$$

制約条件

$$\begin{aligned} \theta x_{oj} - \sum_{k=1}^n \lambda_k x_{kj} &\geq 0 \\ -y_{o1} + \sum_{k=1}^n \lambda_k y_{k1} &\geq 0 \\ \lambda_j &\geq 0 \end{aligned}$$

刀根先生の図に合わせるために、この問題の制約条件の上から2つを以下のように書き換える。

目的関数

$$z' = \theta \quad \text{最小化}$$

制約条件

$$\begin{aligned} \theta x_{oj} / y_{o1} - \sum_{k=1}^n \lambda_k (y_{k1} / y_{o1}) (x_{kj} / y_{k1}) &\geq 0 \\ -1 + \sum_{k=1}^n \lambda_k (y_{k1} / y_{o1}) &\geq 0 \\ \lambda_j &\geq 0 \end{aligned}$$

新しく  $\lambda'_k \equiv \lambda_k (y_{k1} / y_{o1})$  ,  $x'_{kj} \equiv x_{kj} / y_{k1}$  とおくと、上の式は以下ようになる。

目的関数

$$z' = \theta \quad \text{最小化}$$

制約条件

$$\theta x'_{oj} - \sum_{k=1}^n \lambda'_k x'_{kj} \geq 0$$

$$\sum_{k=1}^n \lambda'_k \geq 1$$

$$\lambda'_j \geq 0$$

この式から、CCR モデルでは図 6 のような効率的フロンティアが示される。

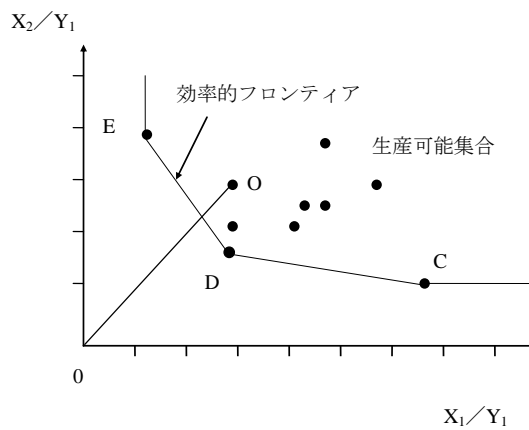


図 6 制御不能変数を含まない CCR モデル

では、BCC モデルではどうなのだろうか。BCC モデルの線形計画問題は以下のようにになる。

目的関数

$$z' = \theta \quad \text{最小化}$$

制約条件

$$\theta x_{oj} - \sum_{k=1}^n \lambda_k x_{kj} \geq 0$$

$$-y_{01} + \sum_{k=1}^n \lambda_k y_{k1} \geq 0$$

$$\sum_{k=1}^n \lambda_k = 1$$

$$\lambda_j \geq 0$$

これに先ほどの変換を行うと以下のようにになる。

目的関数

$$z' = \theta \quad \text{最小化}$$

制約条件

$$\theta x'_{oj} - \sum_{k=0}^n \lambda'_k x'_{kj} \geq 0$$

$$\sum_{k=1}^n \lambda'_k \geq 1$$

$$\sum_{k=1}^n \lambda'_k (y_{o1} / y_{k1}) = 1$$

$$\lambda'_j \geq 0$$

3 番目の条件  $\sum_{k=1}^n \lambda'_k (y_{o1} / y_{k1}) = 1$  から、効率的フロンティアは、単純に要素を結んだものではなく、要素ごとに異なったものになる。

**制御不能変数を含むモデル**

2 変数のうち 1 変数（変数 2）が制御不能変数であった場合、問題は以下ようになる。

目的関数

$$z' = \theta \quad \text{最小化}$$

制約条件

$$\theta x_{o1} - \sum_{k=1}^n \lambda_k x_{k1} \geq 0$$

$$x_{o2} - \sum_{k=1}^n \lambda_k x_{k2} = 0$$

$$-y_{o1} + \sum_{k=1}^n \lambda_k y_{k1} \geq 0$$

$$\lambda_j \geq 0$$

さらに、先ほどの変換を用いると、問題は以下のように変換される。

$$z' = \theta \quad \text{最小化}$$

制約条件

$$\theta x'_{o1} - \sum_{k=1}^n \lambda'_k x'_{k1} \geq 0$$

$$x'_{o2} - \sum_{k=1}^n \lambda'_k x'_{k2} = 0$$

$$\sum_{k=1}^n \lambda'_k \geq 1$$

$$\lambda'_j \geq 0$$

2 番目の制約条件から、この最小化問題は生産可能集合中、 $x'_2 = x'_{2o}$  の平面の中の問題になる。

以上より、効率の測定は図 7 のようになる。

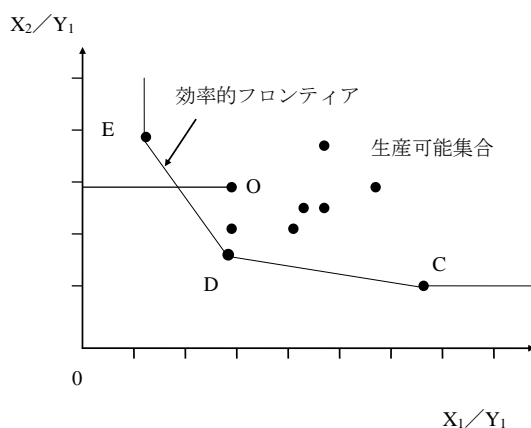


図 7 制御不能変数を含む CCR モデル

BCC モデルは、先ほどと同じく、効率的フロンティアが描けないため、この図で表現することは難しい。

### CCR モデルと BCC モデルの 3 次元表現

最後に CCR モデルと BCC モデルについて、制御不能変数を含まない場合と含む場合の 3 次元グラフを考えてみる。図 8 と図 9 は、それぞれ College Analysis の 3 D モデルビューアで作った CCR モデルと BCC モデルである。一番上の要素が効率かどうか大きな違いである。

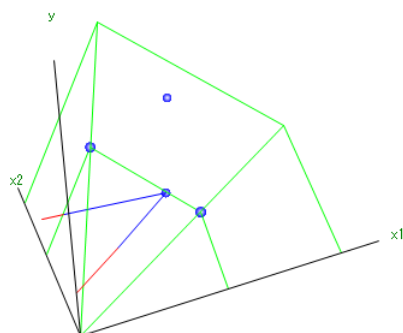


図 8 CCR モデル

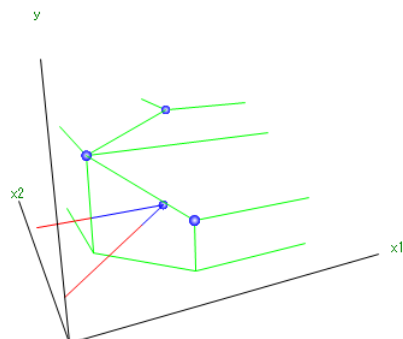


図 9 BCC モデル

測定している要素から伸びた線が効率を測る線で、全体の長さと赤い部分の比で測定する。

実際のプログラム実行画面は図 10 に示される。

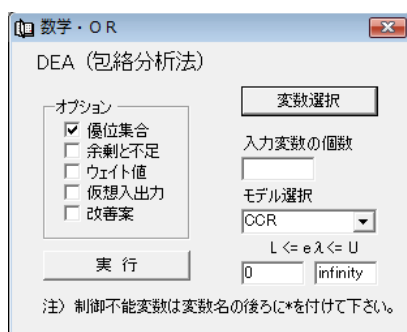


図 10 DEA 実行画面

利用されるデータは、通常の統計分析のデータと同じ、フィールドとレコードによって表わされる形式のものである。「変数選択」により、どの変数を使用するかを指定し、入力変数の個数を入力する。但し、変数選択の順番は、入力変数を先に、出力変数を後に選ばなければならない。出力変数の個数は、全部の変数の数から入力変数の数を引いたものとして認識される。

モデルとしては、CCR, BCC, IRS, DRS, GRS モデルとそれぞれの出力モデル CCRO, BCCO, IRSO, DRISO, GRSO モデルが用意されている。GRS モデルの場合、変数  $\lambda$  についての制約式の上限と下限を入力しなければならない。「実行」ボタンをクリックすると分析結果が表示されるが、表示のオプ



ションとして、優位集合の表示、余剰と不足（スラック変数の値）、ウェイト値（ $\mathbf{v}$ ,  $\mathbf{u}$  の値）、仮想入出力（どの変数を評価したかを表す指標）、改善案がある。図 11 にウェイト値を除く変数を指定した場合の出力結果の例を示す。改善案については別のテキスト画面に 1 つの提言が示される。図 12 にその例を示す。

	D効率値	優位集合く	Sx1×10 <sup>-3</sup>	Sx2	Sy1	Sy2×10 <sup>-3</sup>	v1×蔵書数	v2×職員数	u1×登録者	u2×貸出冊
▶ 千代田	0.226	世田谷(0.02...	0.003537	0.000000	0.000000	0.013840	0.000000	1.000000	0.226006	0.000000
中央	0.638	世田谷(0.09...	0.107170	0.000000	0.000000	0.073293	0.000000	1.000000	0.637738	0.000000
台東	0.540	世田谷(0.13...	0.000000	0.798052	8.812326	0.000000	1.000000	0.000000	0.000000	0.540055
荒川	0.593	世田谷(0.20...	0.000000	4.444691	8.758464	0.000000	1.000000	0.000000	0.000000	0.593021
港	0.911	板橋(0.256)...	0.000000	0.000000	0.000000	0.341054	0.481606	0.518394	0.911285	0.000000
文京	0.745	世田谷(0.35...	0.000000	13.977343	1.006353	0.000000	1.000000	0.000000	0.000000	0.744964
墨田	0.650	杉並(0.170)...	0.073602	0.000000	0.000000	0.000000	0.000000	1.000000	0.263896	0.385774
渋谷	0.539	板橋(0.322)...	0.000000	1.862262	0.000000	0.009582	1.000000	0.000000	0.539130	0.000000
目黒	0.897	杉並(0.166)...	0.000000	0.000000	3.741332	0.000000	0.339039	0.660961	0.000000	0.896643
豊島	0.705	杉並(0.027)...	0.000000	0.000000	3.827265	0.000000	0.327905	0.672095	0.000000	0.705144

図 11 結果表示

DMU: 千代田(0.2260)
入力を 0.2260 倍します。
蔵書数(3.537) を除去します。
貸出冊数(13.840) を追加します。
DMU: 中央(0.6377)
入力を 0.6377 倍します。
蔵書数(107.170) を除去します。
貸出冊数(73.293) を追加します。
DMU: 台東(0.5401)
入力を 0.5401 倍します。
職員数(0.798) を除去します。
登録者数(8.812) を追加します。
DMU: 荒川(0.5930)
入力を 0.5930 倍します。
職員数(4.446) を除去します。
登録者数(8.758) を追加します。

図 12 改善案の提示

## 4. 待ち行列シミュレータ

待ち行列理論は、サービスシステムにおいてサービスを行う人あるいは機械(サービス窓口)が有限個である場合に、サービス待ちの個体数(待ち行列長さ)の変動やシステムの運転効率などについて議論するための理論である(図1)。

サービスを受ける側にとっては待ち時間(システムに到着してからサービス窓口に入るまでの時間)が短いほうがよいのは当然であるし、供給側から見てもサービスした人数が利益に換算される場合には回転率を上げたいであろう。それには、サービス窓口の数を増やすのが手っ取り早い解決策であるが、人件費、設備費等を考えるとむやみにそれを増やすのは得策ではない。そこで、サービスを受ける側がストレスを感じず、しかも供給側には利益が上がるように“ほどほど”の人員、設備を配置することが問題となる。この“ほどほど”に対して指針を与えるのが待ち行列理論である。

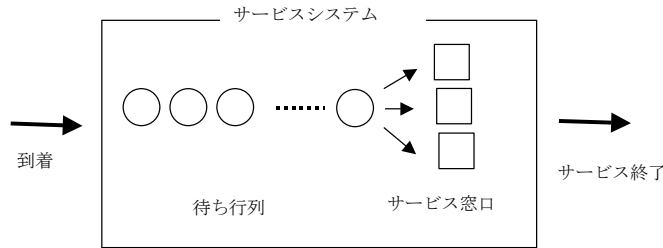


図1 サービスの流れ

待ち行列の長さは客のシステムへの到着時間間隔、サービス窓口の数それから各窓口のサービス時間で決定される。到着時間間隔やサービス時間はオートメーション作業のようにきっちりと決められている場合もあるが、一般には分布を持つ。もちろん、待ち行列の長さはこの分布にも依存するが、ここでは典型的な分布である指数分布とアーラン分布について考えることにする。前述の到着時間間隔やサービス時間が決まっている場合をレギュラー分布ということもある。ここでそれぞれの分布の分布関数を与えておく。 $F(x)$ を時刻  $x$  までに客の到着する確率とすると、

$$\text{レギュラー分布: } F(x) = \begin{cases} 1 & (x \geq 1/\lambda) \\ 0 & (x < 1/\lambda) \end{cases} \quad (1)$$

$$\text{指数分布: } F(x) = \begin{cases} 1 - e^{-\lambda x} & (x \geq 0) \\ 0 & (x < 0) \end{cases} \quad (2)$$

$$\text{k 次アーラン分布: } F(x) = \begin{cases} \int_0^x \frac{(\lambda k)^k t^{k-1}}{(k-1)!} e^{-\lambda k t} dt & (x \geq 0) \\ 0 & (x < 0) \end{cases} \quad (3)$$

ただし、 $\lambda$  は単位時間あたりの到着数（サービス数）である。

待ち行列理論を分類するためにケンドールの記号  $X/Y/s(N)$  を導入しておく。ここで、 $X$  は到着時間間隔の分布を、 $Y$  はサービス時間の分布をそれぞれあらわし、 $X$ （あるいは  $Y$ ）=  $D$  のときレギュラー分布、 $M$  のとき指数分布、 $E_k$  のとき  $k$  次のアーラン分布をそれぞれあらわす。また、 $s$  はサービス窓口の数で、 $N$  はシステムへの入場制限をあらわす。例えば、 $M/M/3(\infty)$  は到着時間間隔が指数分布でシステムへの入場数に制限はなく、窓口 3 つでサービス時間が指数分布であるような待ち行列理論をあらわす、という具合である。

（I） $M/M/c(N)$  到着時間間隔、サービス時間ともに指数分布でサービス窓口の数が  $c$ 、システムへの入場数の制限が  $N$  であるような場合を考える。これは、システムへの到着やサービス終了がまったくランダムに起こる場合に相当する。このことからランダム到着、ランダムサービスの問題と呼ばれる。またこの場合、単位時間あたりの到着数やサービス数はポアソン分布に従うのでポアソン到着と呼ばれることもある。

さて、このシステムの平衡状態を考えよう。単位時間あたりの平均到着数を  $\lambda$ 、平均サービス数を  $\mu$  とする。このとき、平均到着時間間隔は  $1/\lambda$ 、平均サービス時間は  $1/\mu$  となる。システム内に  $n$  人の客がいる確率を  $p_n$  とすると次のような漸化式が成立する。

$$\begin{aligned} -\lambda p_0 + \mu p_1 &= 0 \\ \lambda p_{n-1} - (\lambda + n\mu) p_n + (n+1)\mu p_{n+1} &= 0 \quad (0 < n < c) \\ \lambda p_{n-1} - (\lambda + c\mu) p_n + c\mu p_{n+1} &= 0 \quad (c \leq n < N) \\ \lambda p_{N-1} - c\mu p_N &= 0 \end{aligned} \quad (4)$$

これらを解いて、

$$\begin{aligned} p_n &= \frac{(c\rho)^n}{n!} p_0 \quad (0 \leq n \leq c) \\ p_n &= \frac{c^c}{c!} \rho^n p_0 \quad (c \leq n \leq N) \end{aligned} \quad (5)$$

を得る。ここで、 $\rho = \frac{\lambda}{c\mu}$  と置いた。平衡状態は  $\rho < 1$  の場合にのみ存在する。 $p_0$  は規格化条件

$\sum_{n=0}^N p_n = 1$  より次のように決定される。

$$p_0 = \frac{1 - \rho}{(1 - \rho) \sum_{n=0}^c \frac{(c\rho)^n}{n!} + \frac{(c\rho)^c}{c!} \rho - \frac{(c\rho)^c}{c!} \rho^{N-c+1}} \quad (6)$$

式 (5)、(6) よりシステムの諸量を計算することが出来る。例えば、待ち行列の平均長さ  $L_q$  とシステム平均滞在者数  $L$  はそれぞれ、

$$L_q = \sum_{n=c+1}^N (n-c) p_n = \frac{p_0 \rho (c\rho)^c}{c!(1-\rho)^2} \{ 1 - (N-c+1)\rho^{N-c} + (N-c)\rho^{N-c+1} \} \quad (7)$$

$$L = \sum_{n=1}^N n p_n = L_q + c\rho - \frac{p_0 c (c\rho)^c}{c!} \rho^{N-c+1} \quad (8)$$

となる。また、システムに入ってからサービスを受けるまでの待ち時間の平均値  $W_q$ 、システム内の平均滞在時間  $W$  は  $L_q$ 、 $L$  と次のような関係式が成り立つ。

$$W_q = \frac{L_q}{\lambda} \quad (9)$$

$$W = \frac{L}{\lambda} \quad (10)$$

(II) M/M/c( $\infty$ ) 待ち行列の長さに制限のない場合は、(I) の結果から  $N \rightarrow \infty$  の極限を取ることにより得られ、

$$L_q = \sum_{n=c+1}^{\infty} (n-c) p_n = \frac{p_0 \rho (c\rho)^c}{c!(1-\rho)^2} \quad (11)$$

$$L = \sum_{n=1}^{\infty} n p_n = L_q + c\rho \quad (12)$$

となる。ただし、

$$p_0 = \frac{1-\rho}{(1-\rho) \sum_{n=0}^c \frac{(c\rho)^n}{n!} + \frac{(c\rho)^c}{c!} \rho} \quad (3.13)$$

である。

(III) M/ $E_k$ /1( $\infty$ ) ポアソン到着で、 $k$  次のアーランサービス、窓口 1 つで待ち行列に制限のない場合を考える。この場合は、サービスが  $k$  段階に別れていて、各段階のサービスが平均サービス時間  $1/k\mu$  の指数分布であるようなものを考えれば良い。もちろん同時に一人しかサービスを受けることは出来ない。 $k=1$  の場合には指数サービスに一致する。結果のみ示すと、平均待ち行列長さ  $L_q$ 、平均滞在者数  $L$  はそれぞれ、

$$L_q = \frac{\rho^2}{2(1-\rho)} \left(1 + \frac{1}{k}\right) \quad (14)$$

$$L = L_q + \rho \quad (15)$$

となる。また、この場合も関係式 (9) (10) が成り立つ。

シミュレーションは、単位時間を適当な数に分割し、離散時間での待ち行列の時間発展を数値的に追う。これまでの待ち行列シミュレーションプログラムは様々な条件の下に定常解を求めるシミュレータであった。我々はこれに機能を追加し、単位時間当たりの平均到着数（到着数と略す）、単位時間・窓口数当たりの平均サービス数（サービス数と略す）、窓口数の時間的な変化に対応できるようにした。メニュー画面を図 2 に示す。

The screenshot shows a software window titled "待ち行列シミュレータ" (Waiting Line Simulator) with a menu bar containing "数学・OR". The interface includes the following elements:

- Input Fields:**
  - 平均到着数 (Average Arrival Rate): 3
  - 平均サービス数 (Average Service Rate): 4
  - 初期滞在数 (Initial Stay Count): 0
  - 窓口の数 (Number of Counters): 1
  - Seed: 1
  - 確率調整 到着 (Probability Adjustment Arrival): 1
  - サービス (Service): 0.98
- Checkboxes and Options:**
  - ☐ ファイルから (Load from File)
  - ☐ 複数列 (Multiple Lines)
  - ☐ 待ち行列長さ制限 (Waiting Line Length Limit)
  - ☐ 自動 (Automatic)
- Distribution Selection:**
  - 到着分布 (Arrival Distribution):** Radio buttons for D: 確定 (Fixed), M: ポアソン(指数) (Poisson/Exponential) [selected], and E: アーラン (Erlang).
  - サービス分布 (Service Distribution):** Radio buttons for D: 確定 (Fixed), M: 指数(ポアソン) (Exponential/Poisson) [selected], and E: アーラン (Erlang).
- Execution Parameters:**
  - 実行時間 (Execution Time): 100
  - 時間当たり分割数 (Divisions per Time Unit): 100
  - 実行回数 (Number of Runs): 100
- Graphing Options:**
  - グラフ描画 (Graph Drawing):** Radio buttons for 待ち数 (Waiting Count) [selected], 滞在数 (Stay Count), and サービス (Service).
  - 表示間隔 (Display Interval):** 分割数 (Divisions)
  - Buttons: 入力書式 (Input Format), 実行 (Execute), 再描画 (Redraw)
- Status Bar:** 進行状況 (Progress Status)

図 2 待ち行列分析メニュー画面

入力する条件は、単位時間あたりの客の到着数、サービス数、到着分布とサービス分布、初期行列の長さ、サービス窓口の数、行列長さの制限、実行時間、実行回数、単位時間の分割数である。到着分布は（レギュラー、ポアソン（指数））分布から、サービス分布は（レギュラー、ポアソン（指数）、アーラン）分布から選択できるようにした。レギュラー分布では客の到着時間間隔やサービス時間は一定であるのに対して、それ以外の分布では現象は確率に支配される。

複数窓口の場合について、これまでは全体の窓口に対し 1 列の待ち行列を考えており、サービスはそこから振り分けて行うものとしたが、今回窓口 1 つずつの待ち行列も考えることにした。この場合新しく来た人は最も短い列の後ろに並ぶものとしている。また、待ち人数の制限がある場合は窓口 1 つ当たりの制限人数で与える。

シミュレーションの再現性を重んじ、乱数のシードを自分で入力できるようにしている。また、到着分布にアーラン分布を加え、理論とシミュレーションの微妙な違いを埋めるため、到着とサービス

の確率を微調整する機能を付けている。これは、結果の分からないシミュレーションに対して、実際に近い値を出すための補正用である。ここでは計算結果を表示する画面を図3に示す。

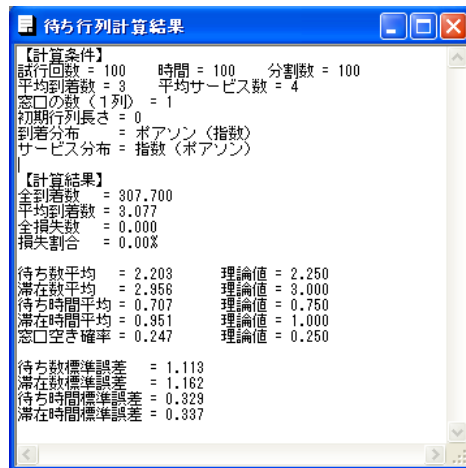


図3 待ち行列結果表示画面

次に待ち行列のグラフ表示については、これまで独自のものをを用いていたが、システム標準の折れ線グラフを利用することにした。そのため、ウィンドウの伸張に連動する機能やコピー、元データを表示する機能などが利用できる。また、グラフの種類も待ち数、滞在数、サービス数の3種類から選択でき、表示間隔も変更することができる。M/M/1( $\infty$ )で、平均到着数3、平均サービス数4、初期滞在数0、単位時間当たり分割数100の場合の待ち数の平均的推移についてのグラフ(表示時間間隔を1/10とした場合)を図4に示す。

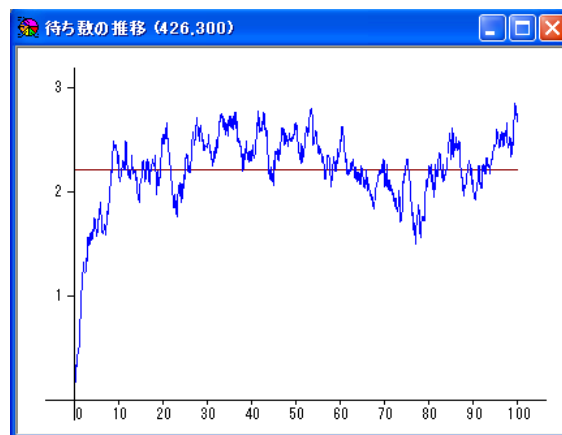


図4 待ち行列グラフ表示画面1

また、D/D/1( $\infty$ ) で、平均到着数 3、平均サービス数 4、初期滞在数 11、単位時間当たり分割数 12 の場合の待ち数の推移は図 5 のようになる。

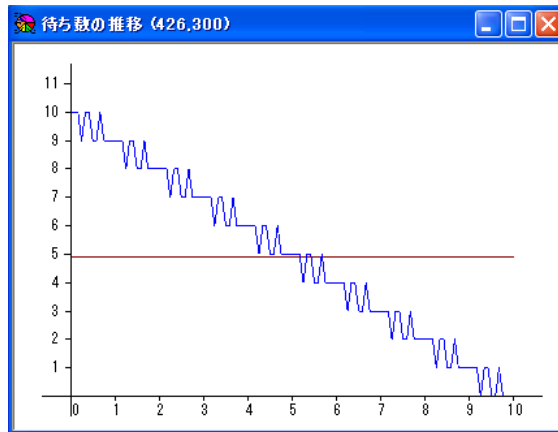


図 5 待ち行列グラフ表示画面 2

図 2 のメニュー画面は以前のものに「時間変化データから」チェックボックスと「入力書式」ボタンの機能を追加したものである。時間変化を与えるデータの例を図 6 に示す。

時刻	到着数	サービス数	窓口数
0	3	4	1
50	6	4	1
55	6	4	2
100			

図 6 時間変化データ

これは、シミュレーションの時間  $0 \leq t < 50$  (単位時間) の間、到着数 3、サービス数 4、窓口数 1、 $50 \leq t < 55$  の間、到着数 6、サービス数 4、窓口数 1、 $55 \leq t < 100$  の間、到着数 6、サービス数 4、窓口数 2 とするものである。さらに細かな設定は時刻の分割を多くするだけである。この状況は最初定常状態だった待ち行列が、到着数が増えたことで行列が伸び、すぐに窓口数を増やして対応したということになる。

メニューの「時間変化データから」チェックボックスをチェックして、「実行」ボタンをクリックすると図 7a と図 7b の実行結果が得られる。

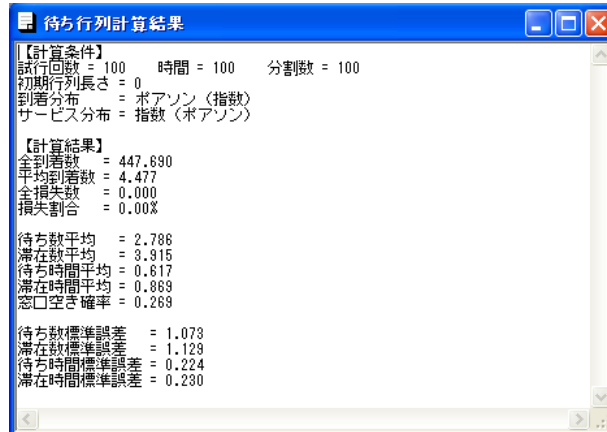


図 7a 待ち行列結果表示画面

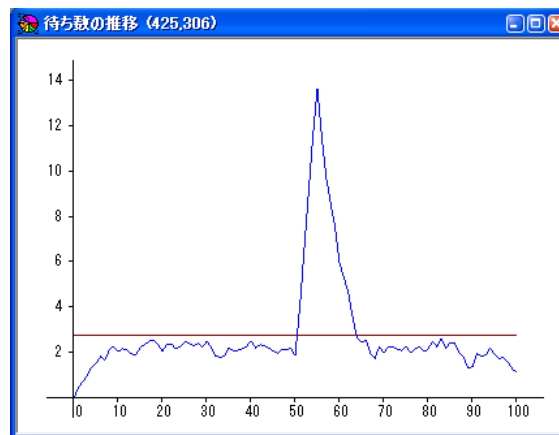


図 7b 待ち行列待ち数グラフ画面

このシミュレーションに関して、到着数とサービス数については条件の変化するところですぐさま変化させるが、窓口数については以下のように取り扱っている。窓口を開ける場合、条件の変化するところでサービスの空いた窓口を開ける。客は待ち数の少ない窓口に行く設定になっているので、その窓口の客の数は自然に増えてゆく。窓口を閉める場合、条件の変化するところで窓口への客の到着を止め、サービスを待っている客はそのままサービスを待つ。この場合、窓口が完全に終了するのは、指定時間より遅れることになる。



## 5. QC 7つ道具

品質管理では QC 7つ道具及び新 QC 7つ道具と呼ばれる代表的な手法があるが、特にグラフ・管理図、ヒストグラム、チェックシート、パレート図、特性要因図、散布図、層別からなる QC 七つ道具は、古くから多くの企業で利用されており、特に生産部門においてこの手法の知識は必要不可欠なものである。今回我々はこれらのうちで、チェックシートを除く 6 つの手法について、コンピュータでサポートするプログラムを作成した。グラフ、ヒストグラム、散布図、層別はこれまでのプログラム<sup>2)</sup>を品質管理の目的に合うように作り変え、管理図、パレート図、特性要因図は新しく作成した。特に特性要因図はグラフィックエディタの特殊な形として、参考文献 3) でも取り上げている。

ここでは品質管理の QC 七つ道具に従って新しいプログラムの作成や、既存のプログラムの活用を行っている。メニュー「分析－数学・OR－品質管理」を選択すると図 1 のようなメインメニューが表示される。

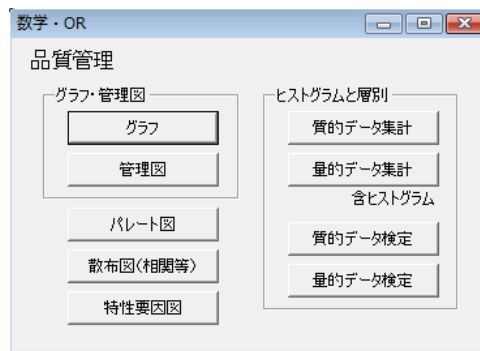


図 1 メインメニュー

メインメニューでは、QC 七つ道具に従い、「グラフ・管理図」、「パレート図」、「散布図」、「特性要因図」、「ヒストグラムと層別」のメニューが並んでいるが、ヒストグラムはこれまでの統計分析の分類から、「層別」と同じグループボックス内の「量的データの集計」の中にある。「層別」は個別な分析手法というよりも、全体を通した分析の考え方と手順であるので、「パレート図」や「散布図」の中にも存在する。七つ道具の 1 つ、「チェックシート」はプログラムで対応するものではないので、ここには含まれていない。以後順を追って個別の機能について紹介する。

### 5.1 グラフ・管理図

図 1 の「グラフ」ボタンをクリックすると、図 1.1 のようなグラフメニューが表示される。

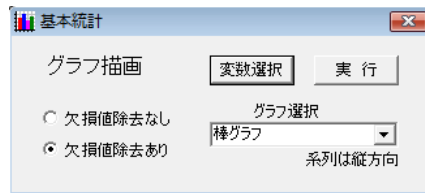


図 1.1 グラフ描画メニュー

これはメニュー「分析－基本統計－グラフ描画」で表示されるものと同じメニューである。変数を選択して、グラフを選び、「実行」ボタンをクリックすると指定されたグラフが表示される。グラフの種類には「棒グラフ」、「積み重ね棒グラフ」、「横棒グラフ」、「積み重ね横棒グラフ」、「帯グラフ」、「3D棒グラフ」、「折れ線グラフ」、「横折れ線グラフ」、「円グラフ」、「散布図」、「レーダーチャート」、「比較レーダーチャート」があるが、今回新しく最後の2つを作成した。ここでは2つのレーダーチャートについて説明する。

図 1.2 にレーダーチャート用のデータを示す。

	目標値	評価1
▶ 価格	10	9
性能	10	7
デザイン	10	8
燃費	10	8
安全性	20	9

図 1.2 レーダーチャート用データ

レーダーチャートは、各目標値にデータがどれだけ近付いたのかを表すグラフであるので、データには目標値が必要である。また、殆どの場合目標値の多角形と1つの評価値の多角形が表示されるが、評価値は複数ある場合もある。変数選択で目標値を最初を選び、「レーダーチャート」または「比較レーダーチャート」を選択して、「実行」ボタンをクリックすると図 1.3 または図 1.4 のようなグラフが表示される。

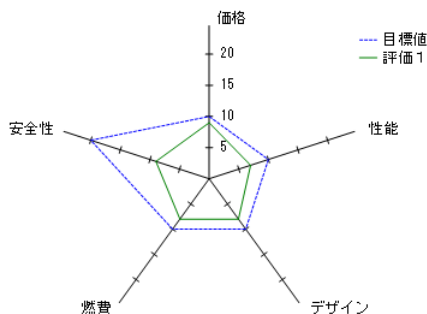


図 1.3 レーダーチャート

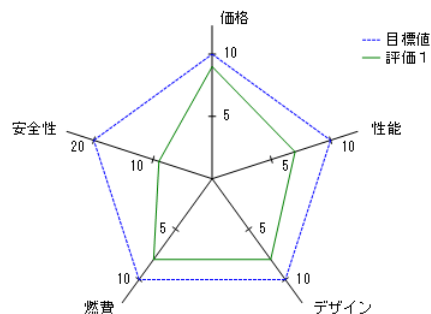


図 1.4 比較レーダーチャート

レーダーチャートは目盛の間隔が同じで、比較レーダーチャートは目標値を同じ半径に設定している。その他のグラフについては、他で説明済みのものが多いので、ここでは省略する。

図1のメインメニューで「管理図」ボタンをクリックすると図1.5のような管理図メニューが表示される。

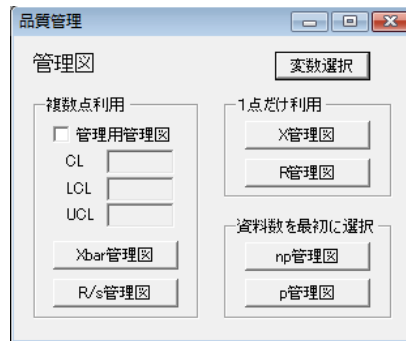


図 1.5 管理図メニュー

管理図には様々な形式があるが、よく使われる形式は $\bar{x}-R$ 管理図である。これには、解析用管理図と管理用管理図の2種類あり、解析用管理図は中心線（CL）と管理限界線（LCL, UCL）をすべてデータから計算するものであり、管理用管理図はこれらを安全設計計画に基づいて利用者が与えるものである。

$\bar{x}-R$ 解析用管理図は、一般に $n$ 個のデータ $x_{it}$ （ $i=1,\dots,n \mid t=1,\dots,T$ ）の時系列的な集合によって求められる。ここでは $\bar{x}-R$ について、 $\bar{x}$ 管理図と $R$ 管理図に分けて見て行く。 $\bar{x}$ 管理図について、CL, LCL, UCLの値は以下で与えられる。

$$CL = \tilde{x}, \quad LCL = CL - A_2 \bar{R}, \quad UCL = CL + A_2 \bar{R}$$

ここに、 $\tilde{x} \equiv \frac{1}{T} \sum_{t=1}^T \bar{x}_t$ ,  $\bar{x}_t \equiv \frac{1}{n} \sum_{i=1}^n x_{it}$ ,  $\bar{R} \equiv \frac{1}{T} \sum_{t=1}^T R_t$  であり、 $R_t$ は $t$ 時点のデータの最大値から最小値を引いた値（レンジ）である。また $A_2$ はJISで定められた $n$ の値による表1.1の数値である。

表 1.1  $\bar{x}$  管理図用係数

群の大きさ	2	3	4	5	6	7	8	9	10
$A_2$	1.880	1.023	0.729	0.577	0.483	0.419	0.373	0.337	0.308

$n > 10$  の場合、CL, LCL, UCLの値は以下で与えられる。

$$CL = \tilde{x}, \quad LCL = CL - 3s, \quad UCL = CL + 3s$$

ここに、 $s = \sqrt{\frac{1}{T} \sum_{t=1}^T (\bar{x}_t - \tilde{x})^2}$  である。

次に  $\bar{x}-R$  管理図における  $R$  管理図について、 $CL, LCL, UCL$  の値は以下になる。

$$CL = \bar{R}, LCL = CL - D_3 \bar{R}, UCL = CL + D_4 \bar{R}$$

$D_3, D_4$  の値については JIS で定められた  $n$  の値による表 1.2 の数値である。

表 1.2  $R$  管理図用係数

群の大きさ	2	3	4	5	6	7	8	9	10
$D_3$	0	0	0	0	0	0.076	0.136	0.184	0.223
$D_4$	3.267	2.574	2.282	2.114	2.004	1.924	1.864	1.816	1.777

具体的には、 $\bar{x}-R$  管理図は図 1.6 のようなデータから描く。この場合  $n=5$  である。

	data1	data2	data3	data4	data5
▶ 1	48.9	53.0	50.0	57.0	53.0
2	47.8	47.1	54.2	51.4	54.9
3	44.2	50.0	55.4	45.7	45.9
4	47.8	46.3	53.1	49.8	50.7
5	49.6	57.1	58.3	44.6	50.2
6	51.5	49.2	43.1	46.9	41.4
7	51.0	50.2	54.8	59.4	46.7
8	41.1	63.0	51.5	55.0	49.2
9	59.6	50.5	56.0	56.5	41.7

図 1.6  $\bar{x}-R$  管理図用データ

図 1.5 のメニューで、「xbar 管理図」ボタンと「R/s 管理図」ボタンをクリックすると、図 1.7 と図 1.8 のような  $\bar{x}-R$  管理図が描かれる。

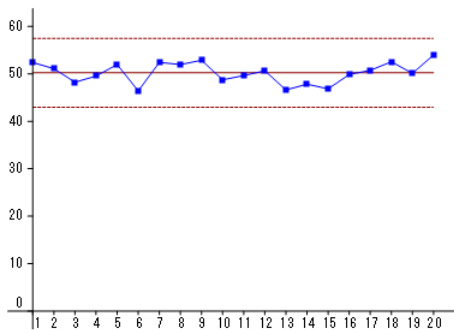


図 1.7 解析用  $\bar{x}$  管理図

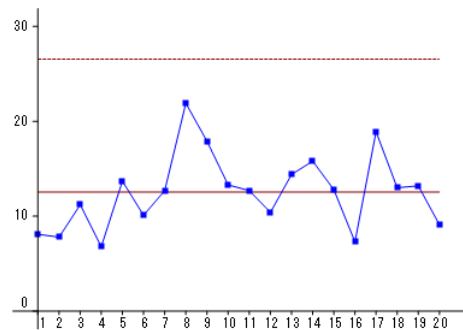
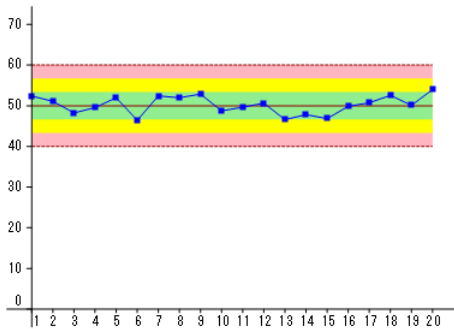
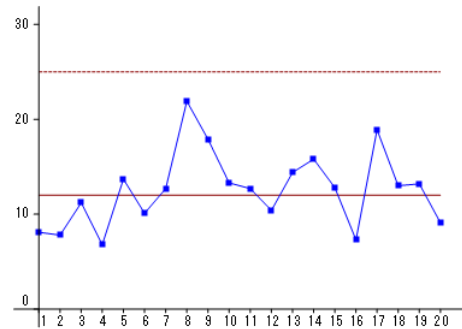


図 1.8 解析用  $R$  管理図

管理用  $\bar{x}-R$  管理図の管理限界線は、解析用管理図から得られたデータを元に利用者が指定する。指定データをそれぞれ  $CL, LCL, UCL$  テキストボックスに記入し、「xbar 管理図」、「R 管理図」ボタンをクリックすると図 1.9、図 1.10 のように管理用管理図が描かれる。

図 1.9 管理用  $\bar{x}$  管理図

(CL = 50, LCL = 40, UCL = 60)

図 1.10 管理用  $R$  管理図

(CL = 12, LCL = 0, UCL = 25)

ここで管理用  $\bar{x}$  管理図には、CL から LCL、CL から UCL を 3 等分して CL に近い方から、緑色、黄色、赤色と色分けしてある。これは、JIS の不安定状態の判定基準に用いられるものである。

次に np 管理図は、図 1.11 のような資料数と不良品数のデータから描かれる。その際、資料数は常に一定でなければならない。これを用いて図 1.5 のメニューで「np 管理図」ボタンをクリックすると、図 1.12 のような np 管理図が描かれる。

	試料数	データ1
▶ 1	100	8
2	100	6
3	100	8
4	100	9
5	100	10
6	100	10
7	100	4
8	100	15
9	100	12
10	100	5
11	100	9
12	100	9
13	100	14
14	100	12
15	100	17

図 1.11 np 管理図用データ

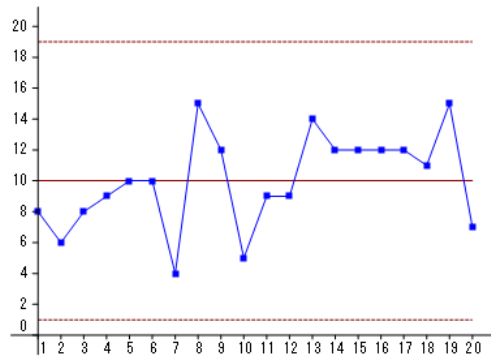


図 1.12 np 管理図

np 管理図の CL, LCL, UCL の値は資料数を  $d$  として以下ようになる。

$$CL = \bar{x}, LCL = CL - \sqrt{dp(1-p)}, UCL = CL + \sqrt{dp(1-p)}$$

ここに、 $\bar{x} = \sum_{i=1}^n x_i / d$ 、 $p = \bar{x} / n$  である。

図 2.1.13 のように資料数に変化がある場合は、p 管理を利用する。p 管理の CL, LCL, UCL の値は  $\lambda$  番目のデータの資料数を  $d_\lambda$ 、データ数を  $x_\lambda$  として以下ようになる。

$$CL = p = \sum_{\lambda=1}^n x_{\lambda} / \sum_{\lambda=1}^n d_{\lambda}, LCL = CL - \sqrt{p(1-p)/d_{\lambda}}, UCL = CL + \sqrt{p(1-p)/d_{\lambda}}$$

グラフは  $p_{\lambda} = x_{\lambda} / d_{\lambda}$  の値で描画する。

具体的な p 管理図は図 1.13 のようなデータを利用して、図 1.14 のようになる。その際限界線は一般に資料数の変わる点で連続でなくなる。

	試料数	データ1
▶ 1	100	8
2	100	6
3	100	8
4	100	9
5	100	10
6	100	10
7	100	4
8	100	15
9	100	12
10	100	5
11	200	23
12	200	17
13	200	20
14	200	16
15	200	27

図 1.13 p 管理図用データ

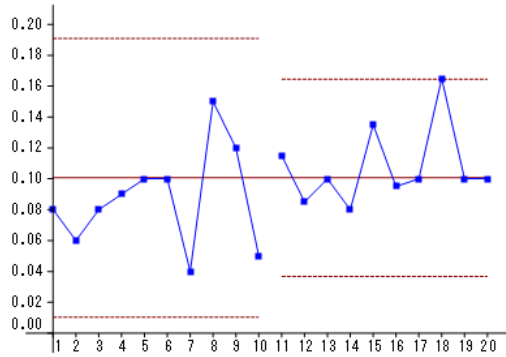


図 1.14 p 管理図

1 点だけ利用する場合の X 管理図と R 管理図について説明する。この場合、X 管理図には 1 つずつのデータの値、R 管理図には前のデータとの差の絶対値を用いる。X 管理図の限界線は、データの平均を  $\bar{x}$ 、データの差の絶対値の平均を  $\bar{R}$  とすると、以下のように与えられる。

$$LCL = \bar{x} - 2.659 \times \bar{R}, UCL = \bar{x} + 2.659 \times \bar{R}$$

R 管理図の限界線は LCL は利用せず、UCL は以下で与えられる。(JIS による)

$$UCL = 3.267 \times \bar{R}$$

データとしては図 1.6 のデータ 1 のみを利用する。データ 1 を変数選択し、「X 管理図」ボタンをクリックすると図 1.15 の実行結果が得られる。

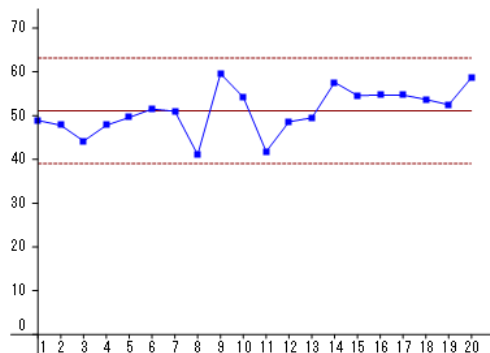


図 1.15 X 管理図

また、「R 管理図」ボタンをクリックすると図 1.16 の実行結果が得られる。

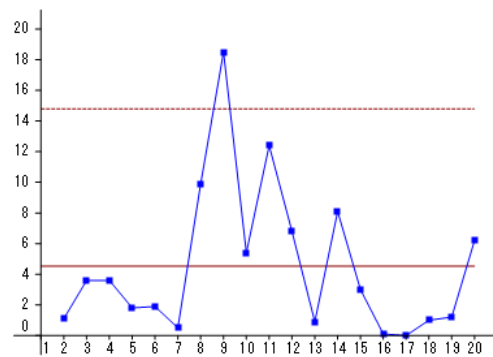


図 1.16 R 管理図

## 5.2 パレート図

図 1 のメインメニューで「パレート図」ボタンをクリックすると、図 2.1 のパレート図メニューが表示される。

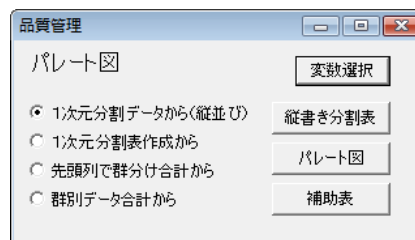


図 2.1 パレート図メニュー

パレート図はいくつかの形式のデータから表示することができる。標準的には図 2.2 のように与えられた縦書きの分割表から、図 2.3 のように表示される。

	不良数
外径不良	73
外觀不良	3
全長不良	2
穴位置不良	56
その他	10
▶ 真円度不良	34
打痕キズ	20

図 2.2 パレート図データ

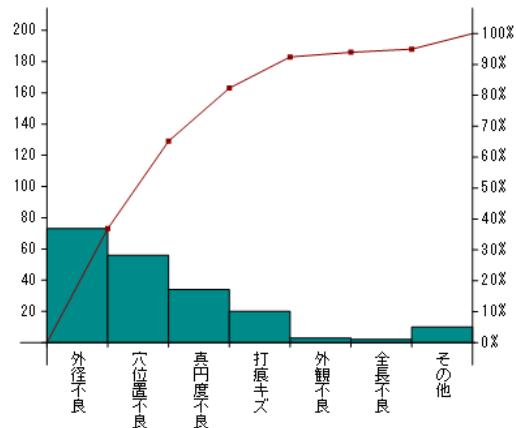


図 2.3 パレート図

ここで図 2.3 の x 軸の名称について、標準では横書きになっているが、図のメニューで縦書きにも書き換えられる。また、名称で「その他」、「他」、「else」とすると、大きさに関わらず最後に表示されるようになっている。また、図 2.1 のメニューでパレート図と同様にして、「補助表」のボタンをクリックすると、補助表として図 2.4 のようなパレート図を少し詳しくしたデータが表示される。

パレート図補助表				
	値	比率(%)	累積値	累積比率(%)
▶ 外径不良	73	36.87	73	36.87
穴位置不良	56	28.28	129	65.15
真円度不良	34	17.17	163	82.32
打痕キズ	20	10.1	183	92.42
外觀不良	3	1.52	186	93.94
全長不良	2	1.01	188	94.95
その他	10	5.05	198	100

図 2.4 パレート図補助表

その他に「1次元分割表作成から」ラジオボタンを選ぶと、元データから1次元分割表を作成して、それを元にパレート図を描く。また、日々の売上高のデータのような場合には、「先頭列で群分け合計から」ラジオボックスを選択する。最初に選択した列でまとめた合計を用いてパレート図を作成する。最後に、売上高を列ごとに合計してデータとしたい場合は、「群別データ合計から」を選択すればよい。

### 5.3 ヒストグラムと箱ひげ図

データの分布を調べる場合はヒストグラムを利用するが、層別に分布を比べるときは、箱ひげ図が用いられることがある。ヒストグラムはすでに作成しているので、今回は新しく箱ひげ図を作成した。箱ひげ図はヒストグラムと同じように利用されるので、図 3.1 のように量的データの集計メニューの中にボタンを加えた。



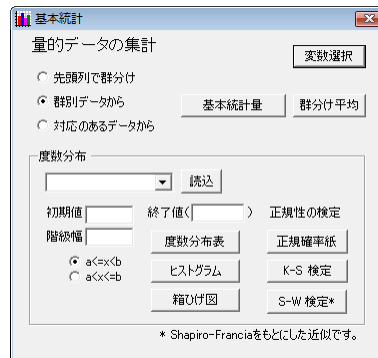


図 3.1 量的データの集計メニュー

箱ひげ図の表示法は、箱の中央を平均値、下端を 25%分位点、上端を 75%分位点、ひげの下端はデータの最小値または  $-3\sigma$  点の大きい方、ひげの上端はデータの最大値または  $+3\sigma$  点の小さい方とする。ひげの範囲を超えるデータについては○印で、具体的にプロットする。図 3.1 の量的データの集計メニューで変数を 1 つずつ選んで、「箱ひげ図」ボタンをクリックすると図 3.2 のような結果が表示される。ここでは、外れ値の多い例と少ない例を示している。

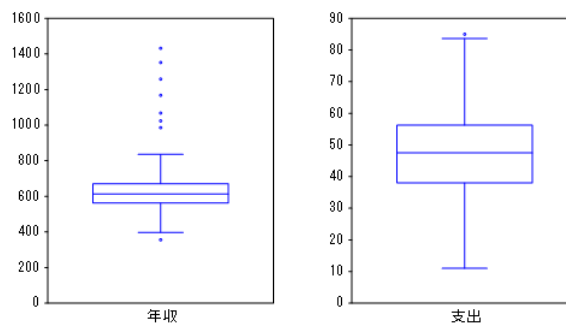


図 3.2 単独の箱ひげ図

これに対し、「先頭列で群分け」ラジオボタンをチェックして、2 つの群で比較するように表示すると図 3.3 のような結果となる。

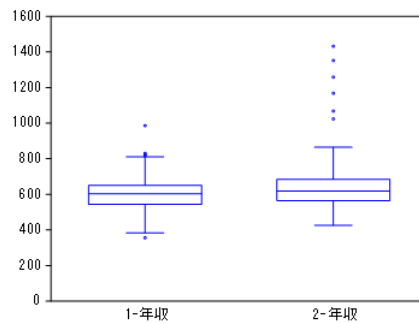


図 3.3 比較のための箱ひげ図

## 5.4 散布図

散布図は以前から作られていたが、品質管理用に層別の考え方を強化し、「相関係数」、「順位相関係数」、「散布図」、「回帰分析」、「予測値と残差」、「予測実測散布図」で、層別（先頭列で群分け）の機能が利用できるようにした。これまでは2つの変数を選んで、分析を実行していたが、層別では、「先頭列で群分け」ラジオボタンを選択し、最初に群分けする変数を選び、その後2つの変数を選択して利用する。メニューは図 4.1 のようにこれまでのものと殆ど変わらない。

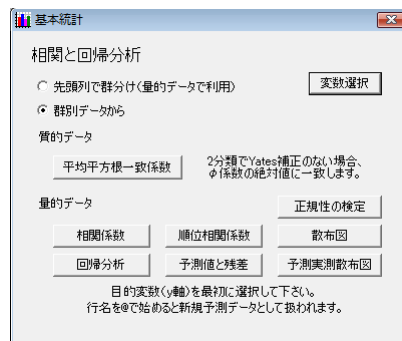


図 4.1 相関と回帰分析メニュー

先頭列で群分けで、「相関係数」ボタンをクリックすると、図 4.2 のように群分けされたデータから相関係数が計算される。順位相関係数も同様であるので、ここでは省略する。

同様に散布図では、図 4.3 のように群分けされたドットで散布図が表示され、それぞれの回帰直線が描かれる。回帰直線は、図の上下端以内なら、x 座標のデータの最小値から最大値まで引くようになっている。グラフのメニューによって回帰直線を非表示にすることもできる。

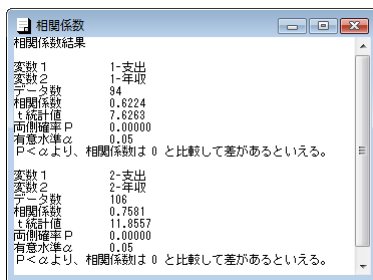


図 4.2 相関係数結果

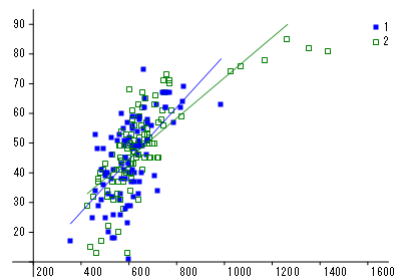


図 4.3 多重散布図

群分けの回帰分析も同様の設定で実行できる。図 4.1 のメニューで「回帰分析」ボタンをクリックすると図 4.4 のような結果が表示される。これは 1 つの群についてだけで、同じものが群の数だけ表示される。また「予測値と残差」ボタンをクリックすると、各群別の回帰式を使った予測値と残差が図 4.5 のように示される。

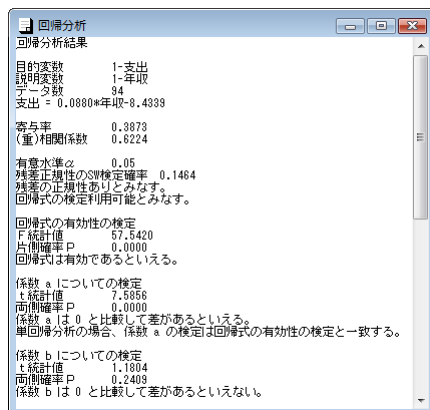


図 4.4 回帰分析結果

地域	実測値	予測値	残差
1	49	42.892	6.108
2	33	41.308	-8.308
3	32	38.574	-6.574
4	31	42.449	-11.449
5	57	43.861	13.139
6	47	46.460	0.540
7	48	45.886	2.114
8	53	31.888	21.112

図 4.5 予測値と残差結果

予測値と実測値の関係を散布図に表示するには、「予測実測散布図」ボタンをクリックする。図 4.5 で与えられた群別の予測値が図 4.6 のように横軸が予測値、縦軸が実測値で表示される。

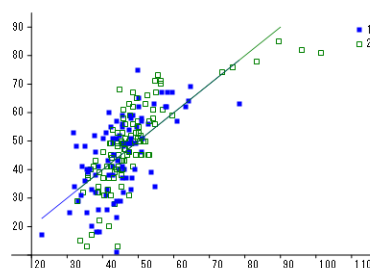


図 4.6 予測実測散布図結果

## 5.5 層別

層別の集計や層別の比較検定についてはこれまで基本統計の枠組みの中でプログラムを作成してきた。これまでの論文で分析メニューについても説明してきたので、ここでは分析用のガイドメニューだけを図 5.1 と図 5.2 に表示するに留める。

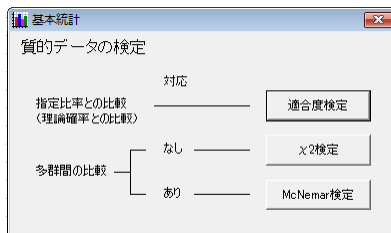


図 5.1 質的データの検定

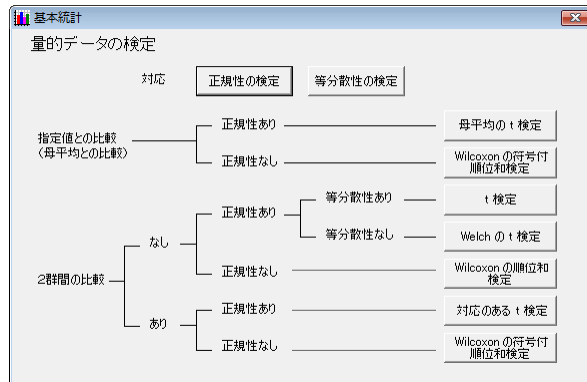


図 5.2 量的データの検定

## 5.6 特性要因図

問題の原因を探究するツールとして特性要因図（フィッシュボーン図）は重要である。まず図 1 のメインメニューで「特性要因図」ボタンをクリックしてグラフィックエディタを表示する<sup>3)</sup>。これは、特性要因図描画用に描画ボタンが特殊化されたものであり、まず「特性」ボタンをクリックしてピクチャーボックス上をクリックし、背骨の部分を描画する。その後、「要因」ボタンをクリックして要因を配置し、「矢印」ボタンで要因を結ぶようにして矢印を描いて行く。簡単な例を図 6.1 に示す。

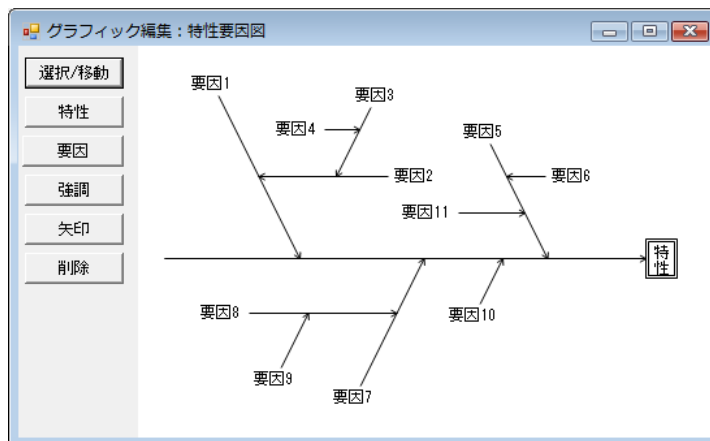


図 6.1 簡単な特性要因図

実際の特性要因図は多くの要因を含むため、画面が大きくなる。ここでは参考文献 1) で与えられたサンプルについて作成した図を図 6.2 に表示する。

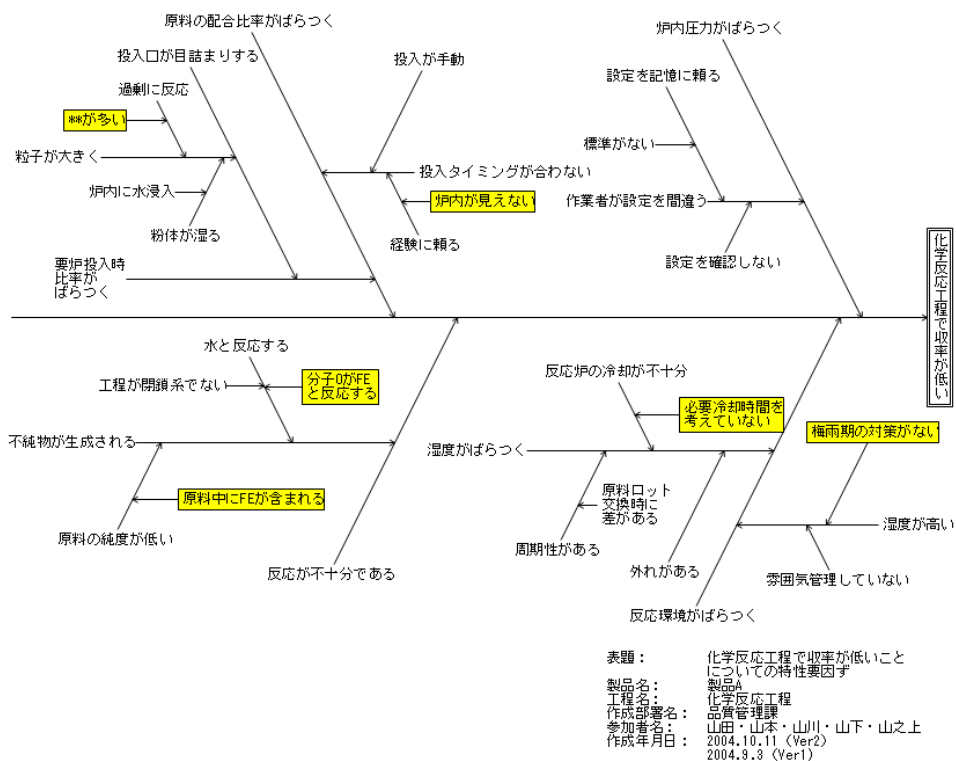


図 6.2 特性要因図サンプル

特性要因図もグラフィックエディタのメニューで、グリッドエディタのデータに変換して保存することができる。グラフィックエディタの使用法やデータ形式については参考文献3) に詳しい。

## 6. 在庫管理シミュレータ

### 6.1 在庫管理の考え方

定量発注方式と定期発注方式の在庫管理に必要なデータは、単位期間（以後 1 日を単位期間とする）の出庫の量の分布と発注から入庫までの期間を与えるリードタイムである。多くの場合、1 日当たりの出庫量  $X$  は正規分布  $N(\mu, \sigma^2)$  で近似し、リードタイムは固定の場合と変動する場合とを考える。リードタイムが変動する場合は、正規分布を仮定し、リードタイム  $L$  の上側確率が指定された遅れの確率になるようにする。

問題点が指摘され、あまり使われないことがないようであるが、1 日 1 品目当たりの在庫費用を  $h$ 、発注 1 回当たりの発注費用を  $K$  として、定量発注方式の経済的発注量  $\sqrt{2K\mu/h}$  や定期発注方式の経済的発注間隔  $\sqrt{2K/\mu h}$  を考える場合もある。

#### 定量発注方式

1 回の発注量  $Q$  の定量発注方式の場合、発注から入庫までの期間  $L$  の出庫量が、 $N(L\mu, L\sigma^2)$  に従うことを用いて、安全在庫  $\lambda\sqrt{L}\sigma$  を考える。ここに  $\lambda$  は安全係数と呼ばれ、欠品確率を  $\alpha$  とする場合に、標準正規分布の確率分布を用いて  $P(X > \lambda) = \alpha$  で与えられる。サイクル在庫は発注量の半分で  $Q/2$  となり、サイクル在庫と安全在庫を加えた理論在庫は  $Q/2 + \lambda\sqrt{L}\sigma$  である。発注点は在庫量が  $I = L\mu + \lambda\sqrt{L}\sigma$  になった時点である。図 1.1 は定量発注方式の在庫量の推移を表すイメージである。

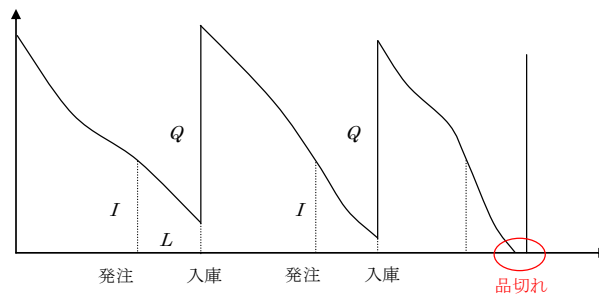
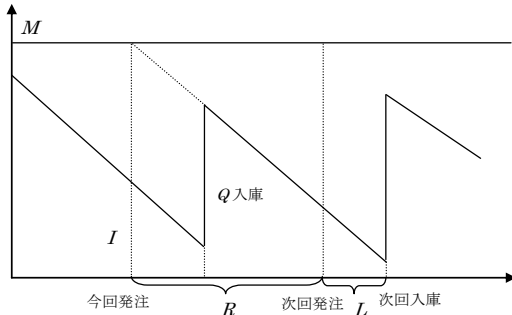
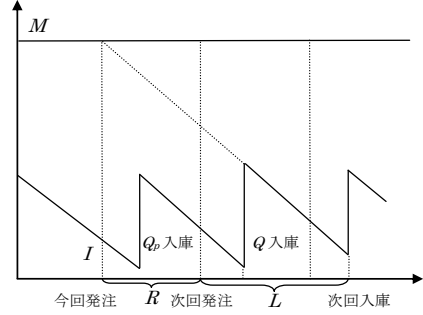


図 1.1 定量発注方式のイメージ

#### 定期発注方式

発注間隔  $R$  の定期発注方式の場合、一度発注すると次回発注分の入庫までの  $L + R$  の間、入庫量を調節できないので、その間の出庫量の分布  $N((L + R)\mu, (L + R)\sigma^2)$  から、安全在庫  $\lambda\sqrt{L + R}\sigma$  が求められる。定量発注方式の場合と同様、 $\lambda$  は安全係数である。この場合サイクル在庫は  $(L + R)\mu/2$  であり、理論在庫は  $(L + R)\mu/2 + \lambda\sqrt{L + R}\sigma$  である。発注量は、最大在庫を  $(L + R)\mu + \lambda\sqrt{L + R}\sigma$  として、最大在庫－現在の在庫量－現在の発注残量（入庫待ち量）で与えられる。定期発注方式の在庫量の推移を表すイメージを図 1.2a と図 1.2b に示す。

図 1.2a 定期発注方式イメージ ( $L < R$ )図 1.2b 定期発注方式イメージ ( $L > R$ )

## 6.2 在庫管理シミュレータの動作

メニュー「分析－数学・OR－在庫管理」を選択すると図 2.1 に示す在庫管理シミュレータのメニュー画面が表示される。

図 2.1 在庫管理シミュレータメニュー画面

まず、発注方式グループボックスから、「定量（発注点）」発注方式か、「定期」発注方式かを選択する。在庫管理のためのデータは、データグループボックスによって「メニューから」と「ファイルから」のどちらかを選択できる。メニューから入力する場合は、左側のテキストボックスにデータを記入する。\*の付いているところは必須項目である。左上の [ ] の付いた保管費用と発注費用については、右上の「経済的極値を利用」チェックボックスにチェックがある場合のみ有効で、発注量や発注間隔の代わりに経済的発注量または経済的発注間隔を利用することができる。

リードタイムは「調達平均\*」テキストボックスに記入する。リードタイムが確定的な場合は、「調達偏差」テキストボックスを 0 のままにしておく、変動する場合は、リードタイムの標準偏差の値を入れ、調達遅れの危険率を「調達危険率」テキストボックスで指定する。出庫量の平均と出庫量の標

標準偏差は、何日分かまとめた数値を入れることもあると考え、左下に「出庫記載単位」テキストボックスを設けている。これは1日当たりの標準偏差の値が単純にまとめた日数で割った値でなく、まとめた日数の平方根で割った値になることから、学生の間違いを減らすねらいもある。

左のテキストボックスにデータを入力終わったら、「在庫データ出力」ボタンをクリックすると、図 2.2a と図 2.2b のような結果が出力される。

在庫基礎データの出力	
定量発注(発注点)方式	
集計単位	1
発注間隔は発注点以下になる日で決定	
サイクル在庫	: 100
安全在庫	: 22
理論在庫	: 122
発注点	: 182
-----	
出庫平均/日	: 20.00
出庫標準偏差/日	: 5.00
-----	
発注量	: 200
欠品危険率(%)	: 5.00
安全係数	: 1.64
調達日数平均	: 7
調達日数標準偏差	: 0
調達危険率(%)	: 5.00

図 2.2a 定量発注方式の理論値

在庫基礎データの出力	
定期発注方式	
集計単位	1
発注量は「最大在庫-現在在庫-発注残」で決定	
サイクル在庫	: 100
安全在庫	: 34
理論在庫	: 134
最大在庫	: 374
-----	
出庫平均/日	: 20.00
出庫標準偏差/日	: 5.00
-----	
発注間隔	: 10
欠品危険率(%)	: 5.00
安全係数	: 1.64
調達日数平均	: 7
調達日数標準偏差	: 0
調達危険率(%)	: 5.00

図 2.2b 定期発注方式の理論値

前者は定量発注方式の理論値の結果であり、後者は定期発注方式の理論値の結果である。

データをファイルから読み込む場合は、図 2.3a と図 2.3b のような形式のファイルを用いる。

	A	B	C
発注量(定量)	200	300	100
発注間隔(定期)	3	3	3
欠品危険率(%)	2	2	2
調達日数平均	7	5	7
調達日数標準偏差	0	0	3
調達危険率(%)	5	5	5

図 2.3a 在庫基礎データ

	A	B	C	
1	18	34	10	
2	17	25	13	
3	14	30	14	
4	17	24	12	
5	19	41	16	
6	21	28	4	
7	21	30	13	

図 2.3b 出庫データ

このファイルには3つの品目のデータが並んでいる。前者は計算用の基礎データで、後者は実際の出庫データである。このファイルを利用する場合、出庫の平均と標準偏差はこのデータから求められる。計算時には、前者を前面に出し、後者のページを「出庫データ」テキストボックスで指定しておく必要がある。「ファイルから」ラジオボタンを選択して、「変数選択」ボタンで品目を選んで実行する。実行結果は、複数の品目があることから、図 2.4a と図 2.4b のようにグリッド形式で表示される。



定量発注（発注点）方式			
	A	B	C
サイクル在庫	100	150	50
安全在庫	30	38	27
理論在庫	130	188	77
発注点	173	185	147
データ数	30	30	30
出庫最小／日	10	12	2
出庫最大／日	33	50	19
出庫平均／日	20.37	29.33	10.07
出庫標準偏差／日	5.57	8.36	3.74
発注量	200	300	100
欠品危険率(%)	2	2	2
安全係数	2.05	2.05	2.05
調達日数平均	7	5	7
調達日数標準偏差	0	0	3
調達危険率(%)	5	5	5

定期発注方式			
	A	B	C
サイクル在庫	31	44	15
安全在庫	36	49	30
理論在庫	67	93	45
最大在庫	240	283	180
データ数	30	30	30
出庫最小／日	10	12	2
出庫最大／日	33	50	19
出庫平均／日	20.37	29.33	10.07
出庫標準偏差／日	5.57	8.36	3.74
発注間隔	3	3	3
欠品危険率(%)	2	2	2
安全係数	2.05	2.05	2.05
調達日数平均	7	5	7
調達日数標準偏差	0	0	3
調達危険率(%)	5	5	5

図 2.4a 定量発注方式の理論値（ファイルから） 図 2.4b 定期発注方式の理論値（ファイルから）  
データのファイル形式は、メニューの「入力用画面作成」ボタンをクリックして、グリッドエディタに枠組みを作成することもできる。

在庫シミュレーションは、出庫量に正規乱数を仮定して実行することができる。出庫量の平均値と標準偏差を指定し、シミュレーショングループボックスの「実行回数」、「実行期間」、「初期在庫」、定期発注方式の場合は「初期発注（定期）」テキストボックスを指定して、「実行」ボタンをクリックと、図 2.5a のようなシミュレーション結果の折れ線グラフと図 2.5b のようなテキストによる結果が表示される。

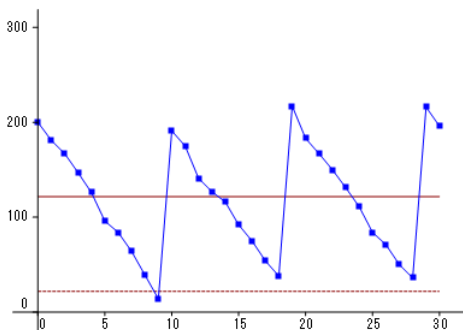


図 2.5a シミュレーション結果（乱数）

シミュレーション結果	
結果はシミュレーション回数 1 回の平均です。	
実行期間	30
総出庫	604.00
出庫平均	20.13
出庫標準偏差	5.84
-----	
発注回数	3.00
総発注量	600.00
発注量平均	200.00
-----	
入庫回数	3.00
総入庫量	600.00
入庫量平均	200.00
-----	
平均在庫量	120.81
欠品回数	0.000
欠品率	0.00%

図 2.5b シミュレーション結果まとめ

ここに欠品率は「入庫前日の欠品回数÷入庫回数」で定義している。

これは定量発注方式の結果であるが、シミュレーションは 1 回だけの結果を表示することもできるし、多数回実行してその平均を表示することもできる。人に説明を行う場合は、1 回のシミュレーションを乱数を変えて何回か行い、これを繰り返した結果は、というように多数回繰り返した結果を表示するのが効果的であろう。定量発注方式の場合、多数回繰り返した場合の結果は図 2.6 のように時間の経過とともに振幅が小さくなっていく。これは平均を取るという性質上、当然の結果である。

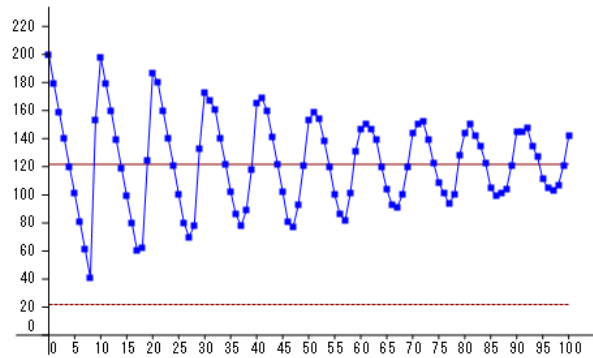


図 2.6 定量発注方式 100 回のシミュレーション結果（平均）

定期発注方式のシミュレーション結果については省略する。

シミュレーションは乱数を用いたもの以外にファイルに保存された実際のデータを用いて実行することもできる。その方法は理論的な在庫データ出力の場合と同様であるので、ここでは省略し、結果のみを図 3.2.7a と図 3.2.7b に示す。

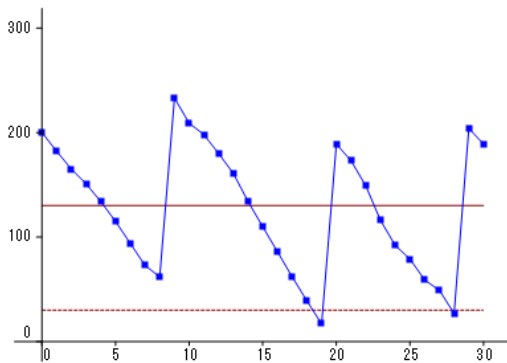


図 2.7a シミュレーション結果（ファイルから）

シミュレーション結果	
結果はシミュレーション回数 1 回の平均です。	
実行期間	30
総出庫	611.00
出庫平均	20.37
出庫標準偏差	5.57
-----	
発注回数	3.00
総発注量	600.00
発注量平均	200.00
-----	
入庫回数	3.00
総入庫量	600.00
入庫量平均	200.00
-----	
平均在庫量	126.81
欠品回数	0.000
欠品率	0.00%

図 2.7b シミュレーション結果まとめ

### 6.3 間欠出庫に対する処理

これまで述べてきた方法は、出庫が毎日あり、正規分布に従う場合に適用できる。現実には多少正規分布からずれても、中心極限定理により、理論は有効である。しかし、出庫が時々しかないような場合にはこれまでの理論は利用できなくなる。このような出庫の形態は間欠出庫と呼ばれる（間欠需要ともいう）。我々はまずこの間欠出庫に対して、出庫の有無と出庫があった場合の分布は独立と考える。出庫の有無に対しては出庫率を $p$ とする 2 項分布、出庫があった場合の分布は平均 $\mu$ 、分散 $\sigma^2$ の正規分布とする。但し、 $\mu \gg \sigma$ を仮定する。

この間欠出庫 1 回の密度関数は図 3.1 のように考えられる。

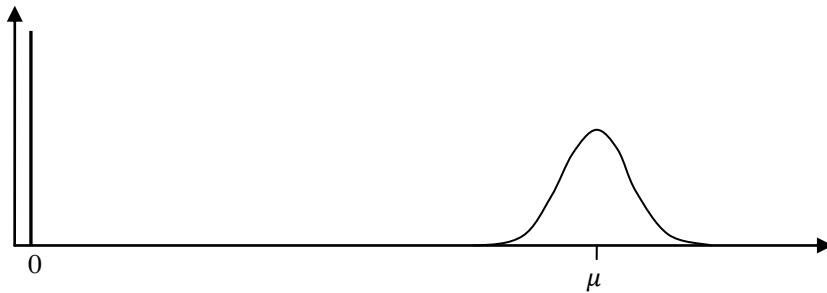


図 3.1 1 回の間欠出庫の密度関数

ここで横軸は出庫の量、0 の位置の棒はデルタ関数を表すものとする。棒または山の面積はそれぞれ  $1-p, p$  で与えられる。

対象となる期間を  $L$  日として、この間に何回出庫があるか考える。出庫の発生回数は 2 項分布に従うと仮定しているので、 $n$  回の発生の確率  $P_n$  は以下で与えられる。

$$P_n = {}_L C_n p^n (1-p)^{L-n}$$

今欠品危険率を  $\alpha$  とした場合、

$$\sum_{i=n}^L P_i > \alpha$$

となる、最大の  $n$  を  $n_{max}$  とすると、以下の関係となる。

$$\sum_{i=n_{max}}^L P_i > \alpha > \sum_{i=n_{max}+1}^L P_i \equiv P'$$

この関係は図 3.2 の密度関数で考えると分かり易い。

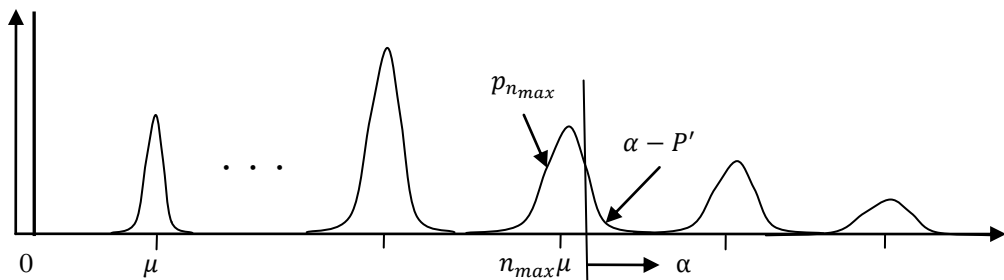


図 3.2 複数回の間欠出庫の密度関数

ここに横軸は出庫の大きさ、それぞれの山の面積は 2 項分布の確率  $P_n$  で与えられる。

仮定  $\mu \gg \sigma$  から、期間開始時の在庫の量を  $(n_{max} + 1)\mu$  とすると、出庫変動に関わらず欠品確率は  $\alpha$  より小さくなり、 $(n_{max} - 1)\mu$  にすると、欠品確率は  $\alpha$  より大きくなる。(これを仮定と考えてもよい)。これより、我々が出庫の確率分布を考えるのは、在庫の量が  $n_{max}\mu$  で、出庫の発生が  $n_{max}$  回あった場合である。

出庫の確率分布によって欠品確率を $\alpha$ にしようとする、出庫が $n_{max}$ 回発生した事後確率として求めればよい。まず、出庫が $(n_{max} + 1)$ 回以上起きる場合は考えないので（在庫の量が $n_{max}\mu$ の場合は必ず欠品）、その確率を $P'$ として欠品確率から引いておく。それを出庫が $n_{max}$ 回起きる確率 $P_{n_{max}}$ で割ると事後確率を求められるので、標準正規分布の分布関数を $F(x)$ として、以下のように安全係数 $\lambda$ を与えることができる。

$$F(\lambda) = 1 - \frac{\alpha - P'}{P_{n_{max}}}, \quad P' = \sum_{i=n_{max}+1}^L P_i$$

これから期間開始時の在庫の量 $I$ を以下のように定める。

$$I = n_{max}\mu + \lambda\sqrt{n_{max}}\sigma$$

現実の計算では、 $\lambda\sqrt{n_{max}}\sigma > \mu/2$ または、 $\lambda\sqrt{n_{max}}\sigma < -\mu/2$ の場合、それぞれを $\mu/2$ または $-\mu/2$ で置き変えてもよいであろう。

### 定量発注方式

定量発注方式の場合、発注点を求める。在庫が発注点を下回ったときに発注して、欠品危険率が $P_{n_{max}}$ となる在庫量は、発注点を下回る量の平均が $\mu/2$ であることから、 $(n_{max} + 1/2)\mu$ とすればよい。これに出庫の確率変動部分を加えて、発注点は以下となる。

$$I = (n_{max} + 1/2)\mu + \lambda\sqrt{n_{max} + 1/4}\sigma$$

### 定期発注方式

定期発注方式では、発注間隔を $R$ として、入庫の調節ができない期間が $L + R$ となる。それゆえ、 $n$ 回の入庫の発生確率は以下となる。

$$P'_n = {}_{L+R}C_n p^n (1-p)^{L+R-n}$$

この $P'_n$ を用いて新たに $n'_{max}$ を計算し、信頼係数 $\lambda'$ を求めて、最大在庫量 $M$ を以下のように定める。

$$M = n'_{max}\mu + \lambda'\sqrt{n'_{max}}\sigma$$

この理論をシミュレーションに適用してみる。

図 3.2 間欠出庫のシミュレーション

図 3.2 の設定でシミュレーションを実行すると以下のような結果となった。

シミュレーション回数	1000 回
実行期間	300
平均在庫量	91.88
入庫回数	5.92
欠品回数	0.231
欠品率	3.90%

この他、仮定の範囲内では、欠品危険率を変動させても、出庫率を変動させても、欠品危険率で指定した値とあまり変わらない欠品率が得られた。

## 7. PERT

スケジュール管理に使われる基本的な手法である PERT を学ぶ際には、アローダイアグラムの作成が一番難しい。現在開発中のプログラムのメニュー画面とそのデータをそれぞれ図 1、図 2 に示す。図 2 の中では作業名、先行作業（ / または , 区切りで入力）、所要日数が利用され、分析メニューの「アローダイアグラム」ボタンで結果が表示されるが、表示のアルゴリズムの検討が不十分なため、ISM のときと同様の手直しが必要である。

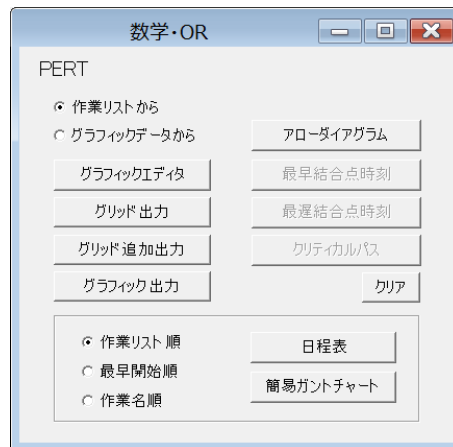


図 1 PERT 分析メニュー

データ編集 PERT1.txt

	作業名	先行作業	作業時間	作業内容
▶ 1	A		7	設計
2	B	A	3	地盤工事
3	C	B	5	基礎工事
4	D	A	6	資材調達
5	E	C	3	屋根工事
6	F	C/D	6	外壁・防水...
7	G	E	4	床面工事
8	H	F/G	5	内壁工事
9	I	E	3	ガス・水道工...
10	J	H/I	2	電気工事
11	K	F/G	10	仕上工事

1/2 (1,1)      分析:      備考:

図 2 PERT 用データ

「アローダイアグラム」ボタンをクリックして、自動出力された図のノード（円形のボックス）の位置を動かし、手直しをした画面が図 3 である。

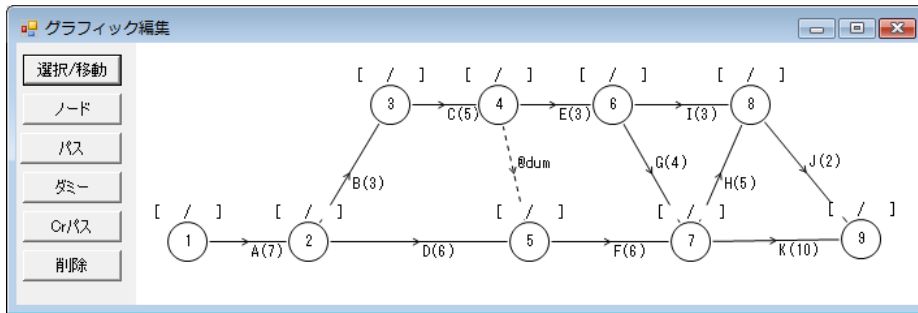


図 3 PERT 図結果画面

矢印の名前の括弧内の値はラインの Value である。

このデータは、必要があれば、分析メニューの「グリッド出力」か「グリッド追加出力」で、グリッドエディタに移し、保存しておく。以後はグラフィック画面を見ながら分析を進めるので、「グラフィックデータから」ラジオボタンを選択する。その際、メニューは図 4 のように変わる。

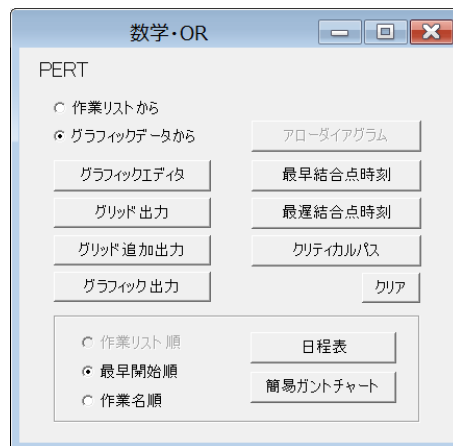


図 4 分析メニュー 2

分析メニューの「最早結合点時刻」をクリックすると、図 5 のように表示される。

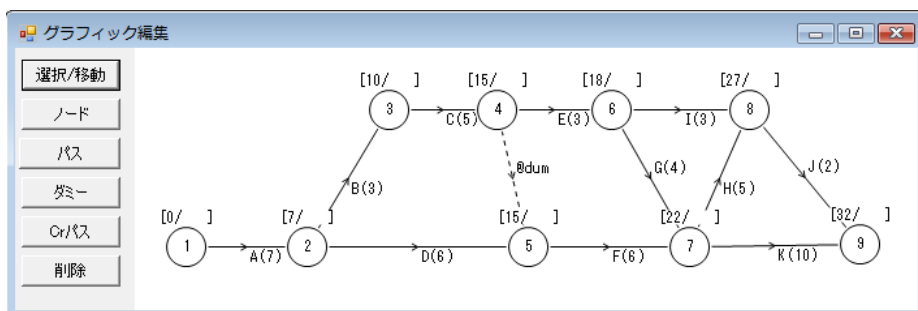


図 5 最早結合点時刻

さらに、「最遅結合点時刻」をクリックすると、図 6 のようになる。

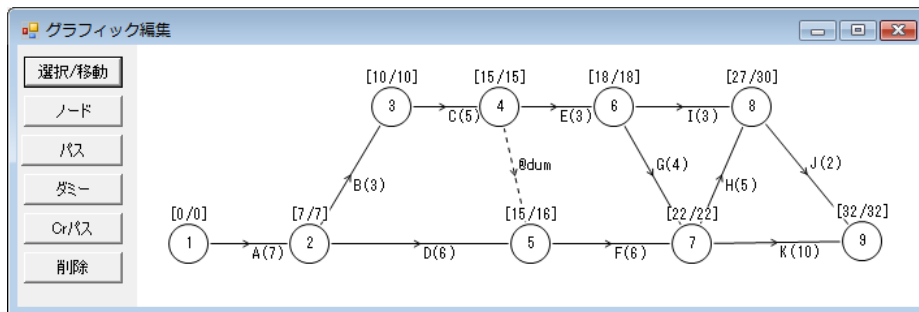


図 6 最遅結合点時刻

ここで、ノードの上にある分数表示は、分子のところが最早結合点時刻と呼ばれ、次の仕事を始められる最早の時刻で、分母のところが最遅結合点時刻と呼ばれ、いつまで次の仕事を待てるかを表す時刻である。この分数形式もノードの Value である。Value の型を文字列型にしたのはこのような場合に対応させるためである。

分析メニューの「クリティカルパス」ボタンをクリックすると、図 7 のように赤い線でクリティカルパスが表示される。

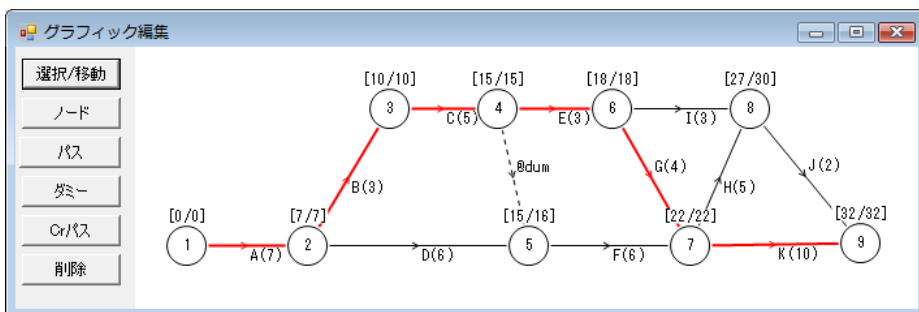


図 7 クリティカルパス

さらに「日程表」ボタンをクリックすると、図 8 のように簡単な日程表が表示される。

作業時間	最早開始	最早終了	最遅開始	最遅終了	Free Float	Total Float	Critical Path
▶ A	7	0	7	0	7	0	1
B	3	7	10	7	10	0	1
D	6	7	13	10	16	2	0
C	5	10	15	10	15	0	1
E	3	15	18	15	18	0	1
F	6	15	21	16	22	1	0
G	4	18	22	18	22	0	1
I	3	18	21	27	30	6	0
H	5	22	27	25	30	0	0
K	10	22	32	22	32	0	1
J	2	27	29	30	32	3	0

図 8 日程表



ここに、作業の並び順は分析メニューで指定できる。各データの意味は以下の通りである。

最早開始時刻

その作業を最も早く開始できる時刻

最早終了時刻

その作業を最も早く終了できる時刻

最遅開始時刻

その作業を遅くとも開始しなければならない時刻

最遅終了時刻

その作業を遅くとも終了しなければならない時刻

フリー・フロート

後続の作業に影響を与えないゆとり

トータル・フロート

最大限許されるゆとり（フリー・フロート $\leq$ トータル・フロート）

最後に「簡易ガントチャート」ボタンをクリックすると、図 9 のように、グリッドエディットでガントチャートが表示される。

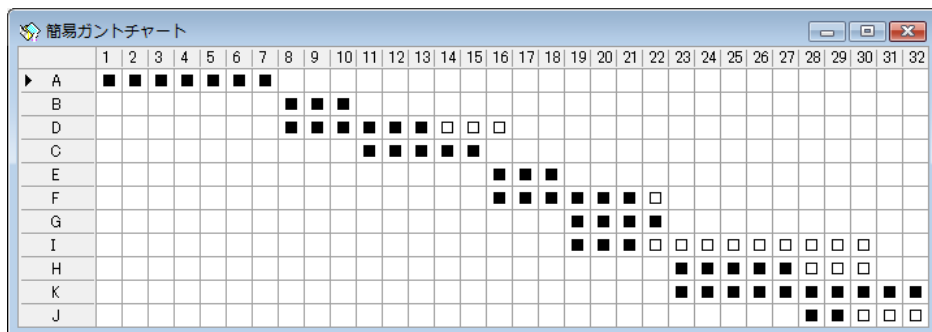


図 9 ガントチャート

## 8. システムダイナミクス

システムダイナミクスは 1950 年代から発展してきたシミュレーション手法で、有益性は広く認識されている。ソフトウェアも STELLA/iThink（以後 STELLA と呼ぶ）を代表として、多く提供されており、社会システム分析には欠かせない 1 分野を形成している。元々この手法は、シミュレーション記述言語によって分析を行っていたが、コンピュータの発達によって、グラフィックによる記述が一般的になり、見た目にも分かり易い表現となった。我々は、このグラフィック表現をグラフィックエディタによって描画することを考え、College Analysis への導入を検討した。プログラム作成に当っては参考文献 1) の STELLA の描画を参考にしたが、これまでのグラフィックエディタの利用法との相違がないようにした。また、外界を表すクラウドについては、開発当初は導入していたが、グラフィック画面の簡単化のために省略している。

グラフィックエディタにおける画面の構成要素は、ボックスとラインである。そのため、フロー（バルブを含む）のように一体化した考え方はない。また、グラフィックエディタでは、ボックス間を結ぶラインは、方向が同じものは 1 本だけしか定義されない。この点が STELLA と最も大きな違いであるが、これはライン間に節点と呼ばれるボックスを 1 つ挟むことで解決した。この節点は、KJ 法の描画の際に導入したもので、コピーや印刷の際には表示されないので、うまく工夫すると描画から完全に消し去ることもできる。

システムダイナミクスを取り扱う際の、グラフィックエディタでのボックスは、ストック、バルブ、コンバータであり、ラインは、フロー（直線矢印）とコントロール（左右円弧矢印）である。ボックスにおけるデータは、数値や数式の形で、ボックスの value の中に保存される。数式の表示では、ファイルデータや  $n$  期前のデータなどの参照が可能である。

メニュー [分析－OR－システムダイナミクス] を選択すると、図 1 のようなシステムダイナミクス分析メニューが表示される。

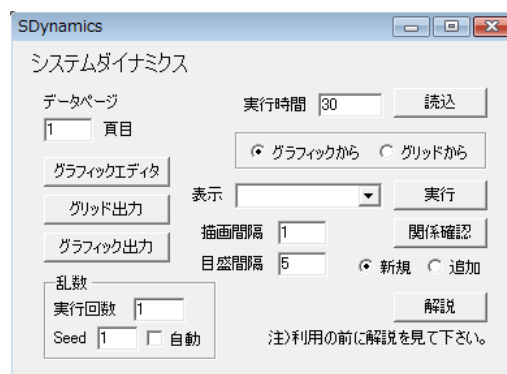


図 1 システムダイナミクス分析メニュー

最初にこのメニューで「グラフィックエディタ」ボタンをクリックして、図 2 のグラフィックエディタ画面を表示する。

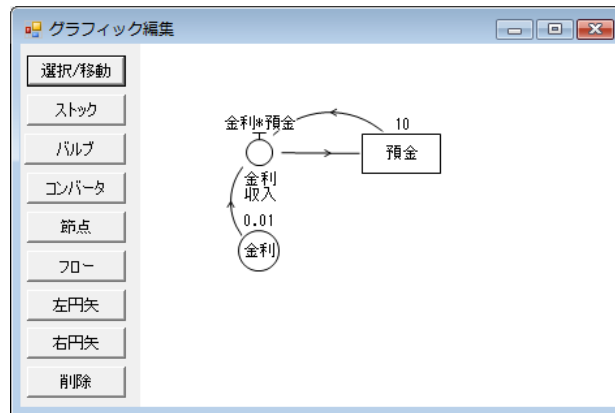


図 2 グラフィックエディタ画面

描かれた図は金利収入による預金の増加モデルである。預金は「ストック」と呼ばれる貯蔵庫、金利は「コンバータ」と呼ばれる影響を表す記号、金利収入は「バルブ」と呼ばれる制御機能を表す記号で表示されている。金利収入から預金への直線は「フロー」と呼ばれ、ものの流れ（この場合は金銭の流れ）を表している。バルブはこのフローを制御している。金利から金利収入、または預金から金利収入への円弧はコントロールと呼ばれ、コンバータやストックから金利収入への影響を表している。図の中では、金利から金利収入への円弧は「左円矢」、預金から金利収入へは「右円矢」を利用している。二つの円形のラインは進行方向左側の円弧か、右側の円弧かの違いだけで同じものであり、見た目にバランスの取れた方を利用すればよい。

ボックス名の入力にはボックスをダブルクリックし、数値や数式を表す **value** の設定は右クリックで入力する。図 2 では、金利の **value** は「0.01」、金利収入の **value** は「金利\*預金」、預金の **value** は初期値の 10（万円）に設定している。

グラフィックエディタメニュー「表示-ボックス値表示[ON/OFF]」で表示のモードを切り替えると、図 3 のようにボックスの上の **value** 値を非表示にすることができる。

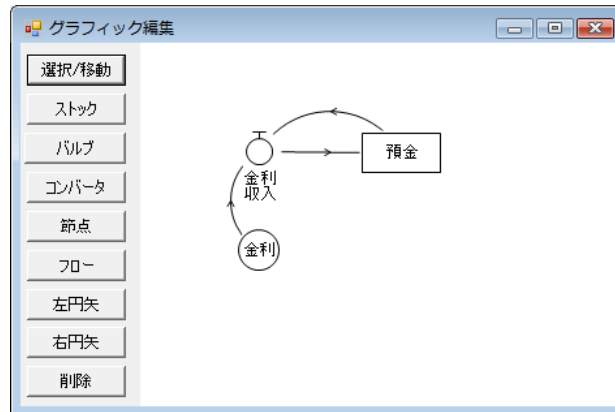


図3 ボックス値非表示モード

できたデータは、「グリッド出力」ボタンをクリックしてグリッドエディタに保存することができる。分析の実行は、「グラフィックから」、「グリッドから」ラジオボタンで、グラフィックデータを直接利用するか、保存したグリッドデータを利用するか選択する。通常は「グラフィックから」にしておき、必要に応じてグリッドエディタに保存する。新規の場合やデータの構造を変えた場合は、「読込」ボタンでプログラム中にデータを読み込み、「表示」コンボボックスで、出力するデータを選択し、「実行」ボタンをクリックする。データの構造を変えず、**value** のみ変更した場合は、そのまま「実行」ボタンをクリックしてもよい。「表示」コンボボックスで預金を選んだ実行結果を図4に示す。

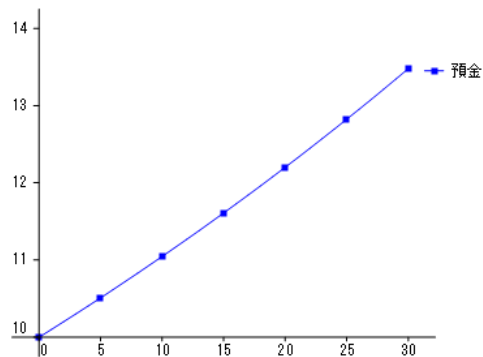


図4 預金表示結果

数式や数値の評価で、コンバータとバルブはその時点で計算を行うが、ストックに関しては、常に1期前のデータを参照している。

金利を、0.01, 0.02, 0.03 と変えた場合の比較は、**value** を変更しながら、メニュー下の「追加」ラジオボタンを選択してグラフを描画する。その結果を図5に示す。ここで凡例は変更してある。

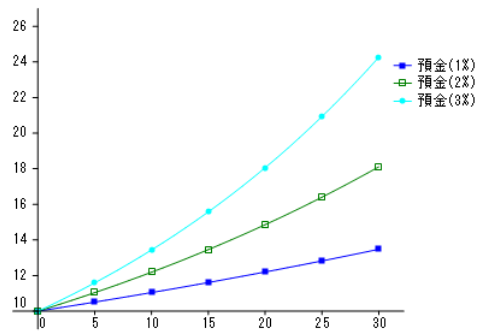


図 5 金利を変化させた試行

「関係確認」ボタンをクリックすると、図 6 のようなボックス間の関係がまとめて表示される。資料などに使用する際に利用する。

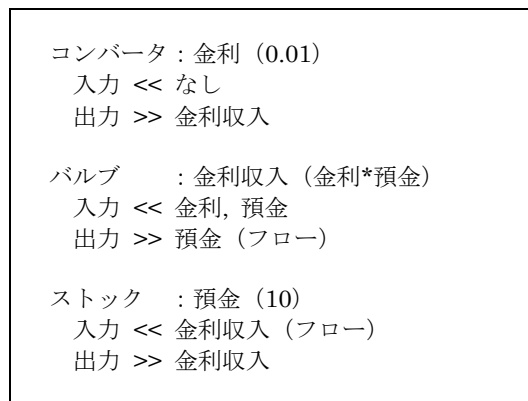


図 6 ボックス間の関係

次にアルコールの吸収・排出問題の例である<sup>1)</sup>。図 7a にグラフィックエディタの画面、図 7b にそれをコピーしたものを示す。

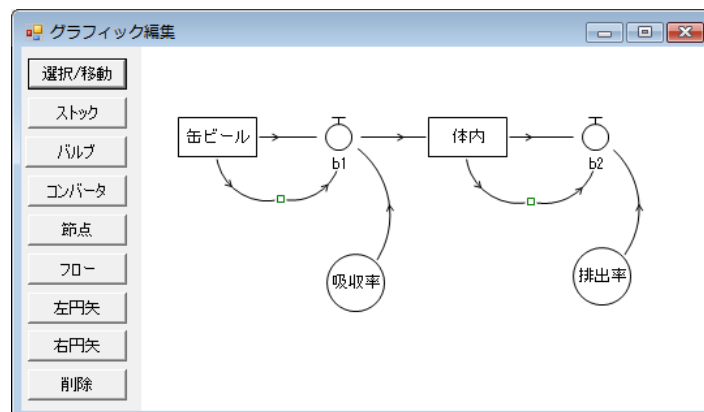


図 7a グラフィックエディタ画面

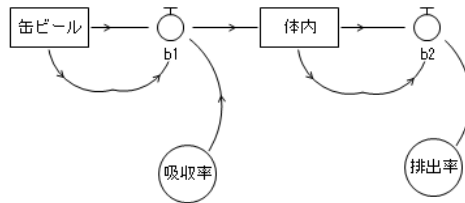


図 7b グラフィックデータのコピー

これは缶ビールからバルブ b1 に、体内からバルブ b2 に、フローとコントロールの 2 本のラインが入る例であるが、グラフィックエディタの仕様から、同じボックス間に同方向の複数のラインを引くことができず、間にグラフィックエディタの左側にボタンのある「節点」を入れている。節点はコピー画像では消えている。しかし、この節点が気になる場合は、図 8a のようにボックスの内部に節点を持って行く。コピーすると図 8b のようにあたかも節点がなかったかのように見える。

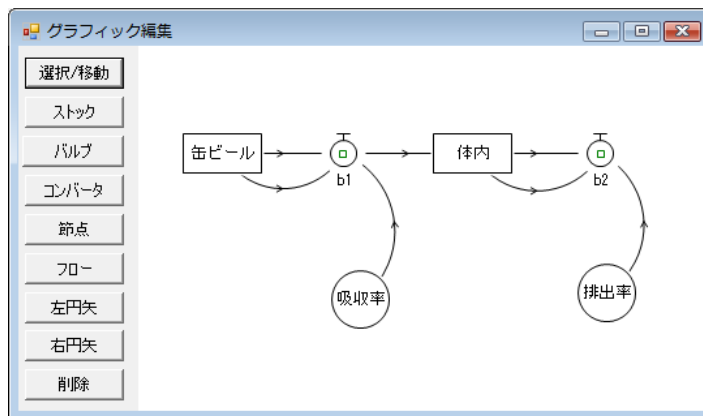


図 8a グラフィックエディタ画面

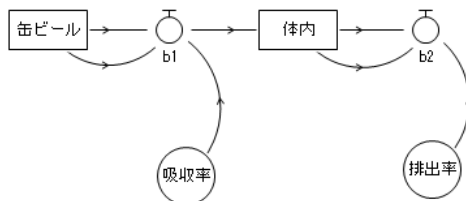


図 8b グラフィックデータのコピー

図 9 に value 値を表示したグラフィックデータ、図 10 に「実行時間」を 100、「目盛間隔」を 10

とした実行結果を示す（見易いように、排出率の値は参考文献と変えてある）。節点があっても、それについての設定は不要である。

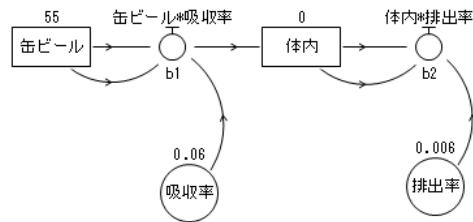


図 9 value 値付きグラフィックデータ

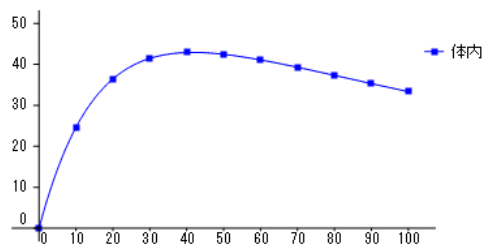


図 10 実行結果

次にファイルのデータを利用する在庫管理問題の例を考える。図 11 がそのモデル画面である。ファイルのデータは出庫データで、このモデルのデータが入っているファイルの 1 ページ目に図 12 のように含まれている。

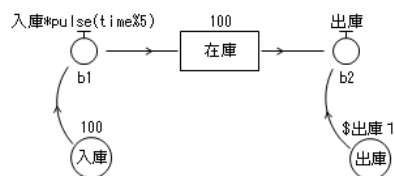


図 11 ファイルデータを利用した在庫の変化

データ編集 システムダイナミクス3 (在庫管理) .txt

	出庫1	出庫2	出庫3
1	18	34	10
2	17	25	13
3	14	30	14
4	17	24	12
5	19	41	16
6	21	30	4

1/5 (1,1)      分析:      備考:

図 12 出庫データ

入庫量は常に 100、入庫は 5 日ごとにする。バルブ b1 の式「 $\text{入庫} * \text{pulse}(\text{time} \% 5)$ 」の time 変数は経過時間で 1 から始まる。「 $\text{time} \% 5$ 」は time を 5 で割った余りを表す。pulse(x) 関数は、 $x=0$  のとき 1、その他のとき 0 の値を持つ。これを文章で表すと、「時間を 5 で割った余りが 0 のときだけ、入庫(100)で表わされる値を在庫ストックに送り込む」ということになる。出庫コンバータの式「\$出庫 1」で、変数名の先頭に付いた\$マークは、分析メニューの「データページ」テキストボックスで与えたページにあるデータを参照することを意味する。変数名は「出庫 1」であり、経過時間  $t$  に  $t$  行目のデータを参照する。在庫量のシミュレーション結果を図 13 に示す。

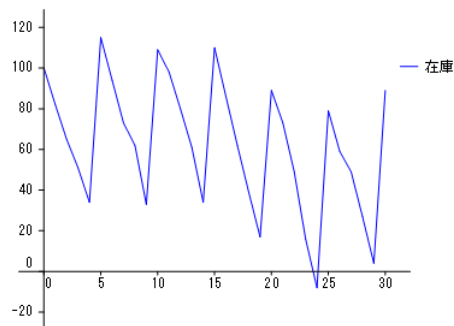


図 13 在庫量シミュレーション結果

グラフ画面のメニュー [設定-データポイント[ON/OFF]] で、データ点は消してある。

次は平均 20、標準偏差 5 の正規出庫の場合の、定期発注方式のシミュレーションである。図 14 にモデルを示す。

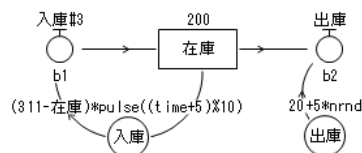


図 14 乱数出庫の定期発注方式



出庫は「 $20+5*\text{nrnd}$ 」で表わされるが、「nrnd」は標準正規乱数であるので、出庫は平均 20、標準偏差 5 の正規乱数データとなる。入庫は 10 日ごとに発生するように  $\text{pulse}()$  関数を用いて設定されているが、その大きさは、最大在庫 311 から現在の在庫の量を引いたものである。但し、在庫納入はバルブ b1 によって、入庫発生から 3 日後となる。b1 の式「入庫#3」は 3 日前の入庫データを参照することを意味する。ストックはその 1 日前の値を用いるので、この式の在庫は 4 日前の値である。そのため、入庫のリードタイムは 4 日と解釈される。在庫量のシミュレーション結果を図 15 に示す。

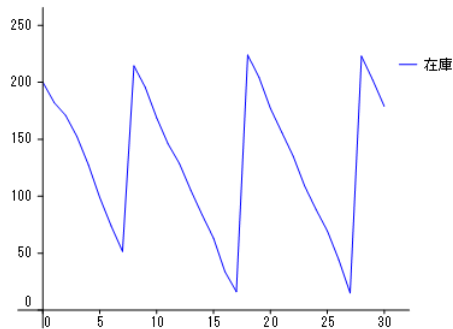


図 15 在庫量シミュレーション結果

乱数は、図 1 のメニューにあるように、Seed を決めて、同じ乱数系列を発生させることも、自動的に Seed を決めて毎回違う乱数系列にすることもできる。メニューの実行回数は、今後拡張するモンテカルロシミュレーション用に作った部分であるが、現在はまだ使用していない。

定量発注方式の表現は、定期発注方式より少し複雑である。在庫が発注点在庫 (96) より少なくなった時点で発注するので、図 16 の入庫コンバータのような複雑な式になる。 $\text{theta}(x)$  関数は  $x \geq 0$  のとき 1、 $x < 0$  のとき 0 となる関数である。 $\text{pulse}()$  関数と合わせて使うと、条件文などが表現できるようになる。図 17 に定量発注方式の在庫量シミュレーション結果を示す。

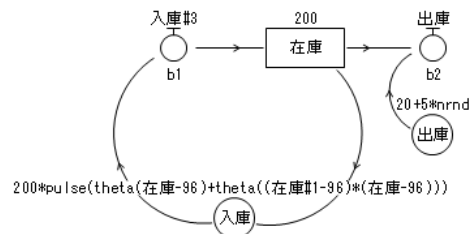


図 16 乱数出庫の定量発注方式

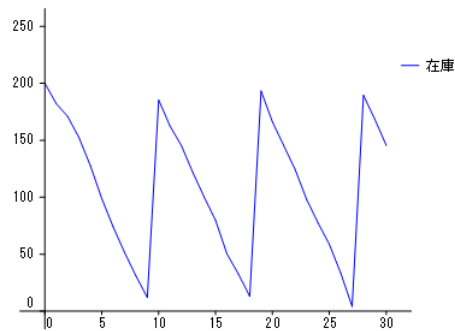


図 17 在庫量シミュレーション結果

これまではストックに上限や下限はなかったが、新しくこれを設定できるようにした。例えば、ストックの値として、100;0;200 のように、セミコロンで区切ると、初期値 100、下限 0、上限 200 となる。また、100;0; や 100;;200 のように、上限や下限を略すとその部分の制約は付かなくなる。この制約は現在のバージョンでは、単独のバルブを介する場合のストック間に有効で、1つのストックに複数のバルブが繋がった場合の制御はまだ完全ではない。図 18 と図 19 に最も簡単な制約付きストックの例を示す。

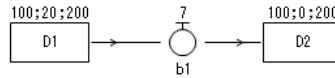


図 18 制約付きストック

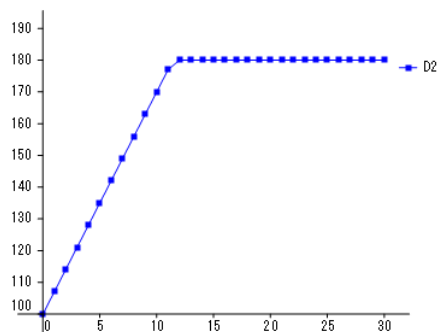


図 19 制約付きストックシミュレーション結果

## 参考文献

- 1) シミュレーションによるシステムダイナミクス入門, 土金達男, 東京電機大学出版局, 2005.

## 9. 不良品診断

ここでは不良品の発生に関する2つのプログラムを紹介する。1つは不良率診断についてのプログラムで、2項分布を使った正確な適合度検定を使い易くしたプログラムである。これは過去の実績から、新しい不良品の発生頻度の確率を求めるものである。もう一つは、抜き取り検査についてのプログラムである。抜き取り検査は多数のロットの中から、サンプルを取り出し、その中の不良品数でロットの合否を決めるものである。プログラムは、サンプル数、合格判定数、ロット中の不良品率、ロットの合格率のうちから、3つのデータを指定し、残りの1つを計算する機能とOC曲線を描く機能を持つ。

### 9.1 不良率診断

不良品の発生確率を過去の実績から求めることは生産ラインの異常を知る上で重要である。発生確率が小さい場合、発生は偶然とは考えられない。

例えばこれまで不良品の発生確率が  $p$  であったとする。これが今回の観測では  $N$  個中  $n$  個であったとする。この事象の起こる確率  $P$  は、通常の検定では自由度1の  $\chi^2$  分布の性質を使った以下の式を利用する。

$$\chi_1^2(P) = \frac{(n - pN)^2}{pN} + \frac{(n - pN)^2}{(1 - p)N} \quad (1)$$

しかし、これは  $p$  と  $n$  が、ある程度大きい場合の近似式で不良品の発生などの少ない発生件数では使えない。正確な計算を行うには以下の2項分布の式を利用する。

$$P = \sum_{i=n}^N {}_N C_i p^i (1 - p)^{N-i} \quad (2)$$

プログラムでは、(2)式を使って計算している。

メニュー「分析－OR－品質管理－不良品診断」を選択すると図1に示す分析プログラムが表示される。データの入力方法は「ファイルから一括処理」と「入力データから」を選択でき、データの種類は「過去発生数から」と「過去発生率から」を選択できる。図1の左がファイルからで、右が入力データからの処理である。

図 1 不良率診断の分析実行メニュー

例えば右の図のように、「データ」テキストボックスに、データを入力し、「分析実行」ボタンをクリックすると図 2 のような結果が表示される。

	過去売上数	過去不良数	過去不良率	売上数	不良数	不良率	検定確率
▶	1000	4	0.0040	100	2	0.0200	0.061226

図 2 出力結果

データファイルの形式は図 3 のように、過去の個数、過去の不良品数、今回の個数、今回の不良品数を商品ごとに並べて入力する。過去の不良品率の場合は、過去の個数と不良品数の代わりにそれだけがあればよい。

	過去個数	過去Claim	今年個数	今年Claim
a	1000	5	200	4
b	2000	7	300	5
▶ c	1000	4	135	2

1/1 (3,4)      分析:      備考:

図 3 データファイル形式

分析実行メニューで図 4 のように変数を選んで、「ファイルから一括処理」ラジオボタンを選択し、「分析実行」ボタンをクリックすると、図 5 のような結果が表示される。

図 4 分析実行メニュー（ファイルの変数選択）

	過去売上数	過去不良数	過去不良率	売上数	不良数	不良率	検定確率
a	1000	5	0.0050	200	4	0.0200	0.018679
b	2000	7	0.0035	300	5	0.0167	0.004393
c	1000	4	0.0040	135	2	0.0148	0.102275

図 5 ファイルから一括処理の出力結果

## 9.2 抜き取り検査

抜き取り検査は、大きさ  $N$  個のロットの中から  $n$  個のサンプルを取り出して、その中に含まれる不良品の数が  $c$  個以下ならそのロットは合格、それより多い場合はそのロットを不合格にする品質検査の手法である。今このロットの中に確率  $p$  で不良品が含まれているとき、サンプルの中から不良品が  $x$  個見つかる確率  $p(x)$  は以下の超幾何分布の確率で与えられる。

$$p(x) = \frac{{}^pN C_x \cdot (1-p)^N C_{n-x}}{{}^N C_n}$$

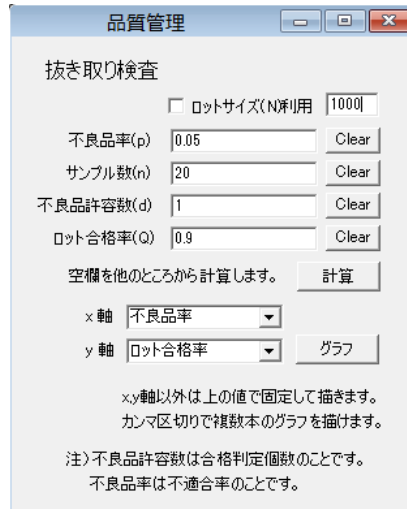
特に  $N \gg n$  で、ロットの中から取り出す個数の減少の影響を考慮しなくても良い場合、確率  $p(x)$  は以下の 2 項分布の確率で与えられる。

$$P(x) = {}_n C_x p^x (1-p)^{n-x}$$

これより、ロットが合格する確率  $P$  は以下となる。

$$P = \sum_{x=0}^c p(x)$$

メニュー「分析－OR－品質管理－抜き取り検査」を選択すると、図 6 のような分析実行画面が表示される。



品質管理

抜き取り検査

☐ ロットサイズ(N)利用 1000

不良品率(p) 0.05 Clear

サンプル数(n) 20 Clear

不良品許容数(d) 1 Clear

ロット合格率(Q) 0.9 Clear

空欄を他のところから計算します。 計算

x 軸 不良品率

y 軸 ロット合格率

グラフ

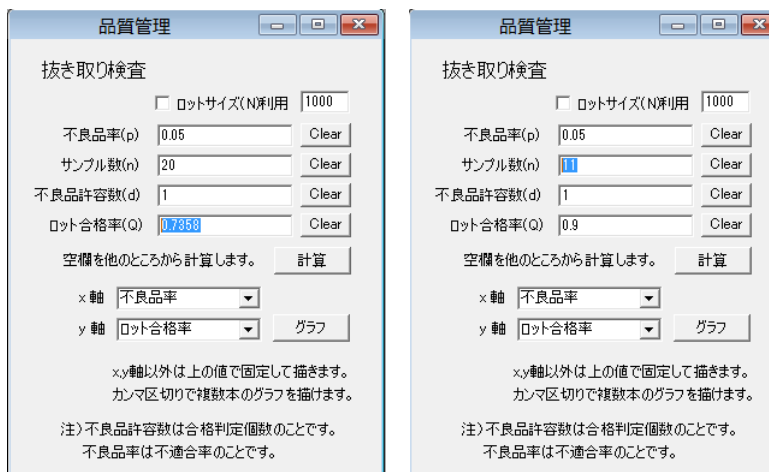
x,y軸以外は上の値で固定して描きます。  
カンマ区切りで複数本のグラフを描けます。

注>不良品許容数は合格判定個数のことです。  
不良品率は不適合率のことです。

図 6 抜き取り検査の分析実行メニュー

上の 4 つのテキストボックスは、そのうちのどれかを空欄にして、その下にある「計算」ボタンをクリックすると計算結果で空欄を埋めてくれる。例えばこの状態で、「ロット合格率」を空欄にして、「計算」ボタンをクリックすると、計算結果の 0.7358 でテキストボックスが埋まる。また、「ロット合格率」を 0.9 にして、「サンプル数」を空欄にして「計算」ボタンをクリックすると、「サンプル数」で 11 という答えが表示される。

この結果は図 7 に示す。



品質管理

抜き取り検査

☐ ロットサイズ(N)利用 1000

不良品率(p) 0.05 Clear

サンプル数(n) 20 Clear

不良品許容数(d) 1 Clear

ロット合格率(Q) 0.7358 Clear

空欄を他のところから計算します。 計算

x 軸 不良品率

y 軸 ロット合格率

グラフ

x,y軸以外は上の値で固定して描きます。  
カンマ区切りで複数本のグラフを描けます。

注>不良品許容数は合格判定個数のことです。  
不良品率は不適合率のことです。

品質管理

抜き取り検査

☐ ロットサイズ(N)利用 1000

不良品率(p) 0.05 Clear

サンプル数(n) 11 Clear

不良品許容数(d) 1 Clear

ロット合格率(Q) 0.9 Clear

空欄を他のところから計算します。 計算

x 軸 不良品率

y 軸 ロット合格率

グラフ

x,y軸以外は上の値で固定して描きます。  
カンマ区切りで複数本のグラフを描けます。

注>不良品許容数は合格判定個数のことです。  
不良品率は不適合率のことです。

図 7 3つの条件で他の条件を決める方法

メニューの条件で、不良品率を変化させた場合のロット合格率の変化を見たいときは、図 7 の設定で「グラフ」ボタンをクリックすると、図 8 のような結果が表示される。

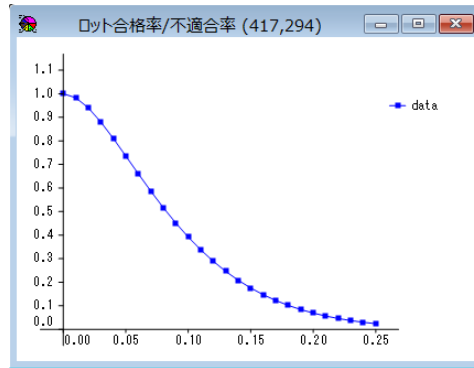


図 8 不良品率の変化によるロット合格率の変化

さらに不良品許容数を 0,1,2,3 と変化させたいときは、「不良品許容数」にカンマ区切りでこのまま入力し、「グラフ」ボタンをクリックする。結果は図 9 のように表示される。

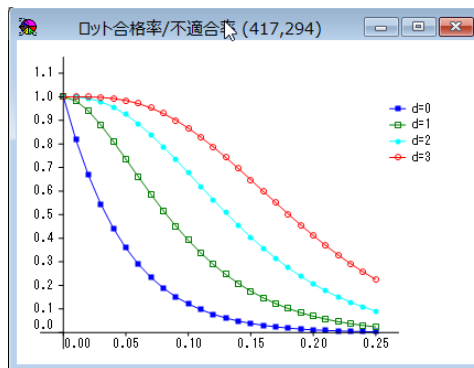


図 9 不良品許容数を変化させたロット合格率の変化

同じ設定で、ロット合格率と不良率を逆にすると図 10 のような結果になる。

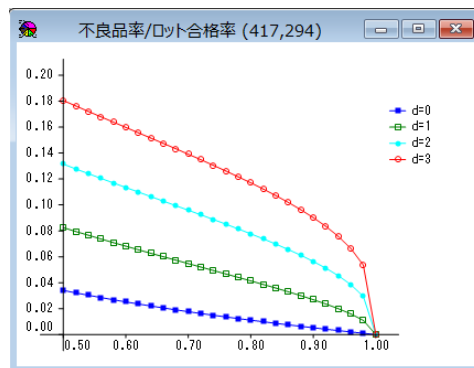


図 10 ロット合格率の変化による不良率の変化

## 10. パラメータ設計

### 10.1 パラメータ設計の理論と考え方

ここではまずゼロ点比例式の動特性パラメータ設計について考えるが、その前に SN 比と感度について、理論的な考察を加えておく。

1 つの実験では、信号水準  $M_j$  ( $j=1, \dots, p$ ) と誤差水準  $N_\alpha$  ( $\alpha=1, \dots, n$ ) によって、表 1 のように  $pn$  個のデータ  $y_{j\alpha}$  が得られる。誤差水準はできるだけ広く散らばるよう配慮されるものとする。

表 1 パラメータ設計におけるデータ

$M_1$			...	$M_p$		
$N_1$	...	$N_n$	...	$N_1$	...	$N_n$
$y_{11}$	...	$y_{1n}$	...	$y_{p1}$	...	$y_{pn}$

この実験についての、誤差水準  $\alpha$  のゼロ比例式回帰直線を考える。実測値  $y_{j\alpha}$  についての推定回帰式を  $Y_{j\alpha} = b_\alpha M_j$  とすると、実測値との差の 2 乗和は以下となる。

$$EV_\alpha = \sum_{j=1}^p (y_{j\alpha} - Y_{j\alpha})^2 = \sum_{j=1}^p (y_{j\alpha} - b_\alpha M_j)^2$$

これを最小とするには、

$$\frac{\partial}{\partial b_\alpha} EV_\alpha = -2 \sum_{j=1}^p M_j (y_{j\alpha} - b_\alpha M_j) = 0$$

として、以下を得る。

$$b_\alpha = \sum_{j=1}^p M_j y_{j\alpha} / \sum_{j=1}^p M_j^2 = \frac{L_\alpha}{r}, \quad \text{ここに、} L_\alpha = \sum_{j=1}^p M_j y_{j\alpha}, \quad r = \sum_{j=1}^p M_j^2$$

全体のゼロ比例式回帰直線については、推定回帰式を  $Y_j = b M_j$  とすると、実測値との差の 2 乗和は以下となる。

$$EV = \sum_{j=1}^p \sum_{\alpha=1}^n (y_{j\alpha} - Y_j)^2 = \sum_{j=1}^p \sum_{\alpha=1}^n (y_{j\alpha} - b M_j)^2$$

これを最小とするには、

$$\frac{\partial}{\partial b} EV = -2 \sum_{j=1}^p \sum_{\alpha=1}^n M_j (y_{j\alpha} - b M_j) = 0$$

として、以下を得る。



$$b = \sum_{j=1}^p \sum_{\alpha=1}^n M_j y_{j\alpha} / \sum_{j=1}^p \sum_{\alpha=1}^n M_j^2 = \frac{1}{nr} \sum_{j=1}^p \sum_{\alpha=1}^n M_j y_{j\alpha} = \frac{1}{nr} \sum_{\alpha=1}^n L_\alpha = \frac{1}{n} \sum_{\alpha=1}^n b_\alpha$$

次にデータの変動について考察する。まず、 $\mathbf{y} = \mathbf{0}$ からの全体の変動  $S_T$  は以下となる。

$$S_T = \sum_{j=1}^p \sum_{\alpha=1}^n y_{j\alpha}^2 \quad \text{自由度 } pn$$

また、 $\mathbf{y} = \mathbf{0}$ からの全体の回帰変動  $S_\beta$  は以下となる。

$$S_\beta = \sum_{j=1}^p \sum_{\alpha=1}^n (bM_j)^2 = nrb^2 = \frac{1}{nr} \left( \sum_{\alpha=1}^n L_\alpha \right)^2 \quad \text{自由度 } 1 \text{ ( } b \text{ のみ)}$$

これより、 $b^2 = S_\beta / nr$  となる。

誤差水準  $\alpha$  の回帰直線の全体の回帰直線からの変動  $S_{N\beta}$  は以下となる。

$$S_{N\beta} = \sum_{j=1}^p \sum_{\alpha=1}^n (b_\alpha M_j - bM_j)^2 = r \sum_{\alpha=1}^n (b_\alpha - b)^2 = r \sum_{\alpha=1}^n b_\alpha^2 - nrb^2 = \sum_{\alpha=1}^n L_\alpha^2 / r - S_\beta$$

自由度  $n-1$     束縛条件  $\sum_{\alpha=1}^n (b_\alpha - b) = 0$     1 個

各点の誤差水準  $\alpha$  の回帰直線からの変動  $S_e$  は以下となる。

$$\begin{aligned} S_e &= \sum_{j=1}^p \sum_{\alpha=1}^n (y_{j\alpha} - b_\alpha M_j)^2 = \sum_{j=1}^p \sum_{\alpha=1}^n [y_{j\alpha} - (b_\alpha - b)M_j - bM_j]^2 \\ &= \sum_{j=1}^p \sum_{\alpha=1}^n y_{j\alpha}^2 + \sum_{j=1}^p \sum_{\alpha=1}^n (b_\alpha - b)^2 M_j^2 + \sum_{j=1}^p \sum_{\alpha=1}^n b^2 M_j^2 - 2 \sum_{j=1}^p \sum_{\alpha=1}^n y_{j\alpha} b_\alpha M_j \\ &= S_T + S_{N\beta} + S_\beta - 2 \sum_{\alpha=1}^n L_\alpha^2 / r^2 = S_T + S_{N\beta} + S_\beta - 2(S_{N\beta} + S_\beta) = S_T - S_{N\beta} - S_\beta \end{aligned}$$

自由度  $pn - n$     束縛条件  $\sum_{j=1}^p M_j (y_{j\alpha} - b_\alpha M_j) = 0$      $n$  個

各点の誤差水準  $\alpha$  の回帰直線からの不偏分散  $V_e$  は、変動を自由度で割って以下となる。

$$V_e = S_e / (np - n)$$

これが不偏分散となることはこの節の最後に詳しく説明する。

ここまでの議論で、全変動  $S_T$  は、全体の回帰変動  $S_\beta$ 、全体の回帰直線からの誤差水準  $\alpha$  の回帰変動  $S_{N\beta}$ 、各点の誤差水準  $\alpha$  の回帰直線からの変動  $S_e$  の和で以下のように表されることが分かった。

$$S_T = S_\beta + S_{N\beta} + S_e$$

各点の全体の回帰直線からの変動  $S_N$  は以下となる。

$$\begin{aligned} S_N &= \sum_{j=1}^p \sum_{\alpha=1}^n (y_{j\alpha} - bM_j)^2 = \sum_{j=1}^p \sum_{\alpha=1}^n y_{j\alpha}^2 + nb^2 \sum_{j=1}^p M_j^2 - 2b \sum_{j=1}^p \sum_{\alpha=1}^n y_{j\alpha} M_j \\ &= S_T + nrb^2 - 2nrb^2 = S_T - S_\beta = S_{N\beta} + S_e \end{aligned}$$

$$\text{自由度 } pn-1 \quad \text{束縛条件} \quad \sum_{j=1}^p \sum_{\alpha=1}^n M_j (y_{j\alpha} - bM_j) = 0$$

各点の全体の回帰直線からの不偏分散  $V_N$  は、変動を自由度で割って以下となる。

$$V_N = S_N / (pn-1)$$

これらを使って、SN 比  $\eta$  と感度  $S$  を定義する。SN 比は、測定誤差の分散  $\sigma^2$  に対する有効な信号の変化の大きさ  $\beta^2$  の比を用いて、また感度  $S$  は  $\beta^2$  の値を用いて以下のように定義される。

$$\text{SN 比 : } \eta = 10 \log_{10} \frac{\beta^2}{\sigma^2}, \quad \text{感度 : } S = 10 \log_{10} \beta^2$$

実際の計算では  $\beta^2$  と  $\sigma^2$  の値は不明であるので、これらの不偏推定量を用いて置き換える。

$$\text{SN 比 : } \eta = 10 \log_{10} \left[ \frac{(S_\beta - V_e)/nr}{V_N} \right], \quad \text{感度 : } S = 10 \log_{10} \left[ (S_\beta - V_e)/nr \right]$$

一般に SN 比は大きな値ほど、有効な信号を誤差の中から拾いやすくなり、良好な結果である。また、感度は対象により、大きな値がよい場合、小さな値がよい場合、目標値がよい場合など様々であるが、感度があまり変化しない制御因子を用いて SN 比を上げることを考える場合もある。

最後に、 $\beta^2$  の不偏推定量を求めておく。

$$y_{j\alpha} = \beta_\alpha M_j + \varepsilon_{j\alpha}, \quad E[\varepsilon_{j\alpha}] = 0, \quad E[\varepsilon_{j\alpha} \varepsilon_{j'\alpha'}] = \delta_{jj'} \delta_{\alpha\alpha'} V[\varepsilon]$$

とすると、

$$b_\alpha = \frac{1}{r} \sum_{j=1}^p M_j y_{j\alpha} = \frac{1}{r} \sum_{j=1}^p M_j (\beta_\alpha M_j + \varepsilon_{j\alpha}) = \beta_\alpha + \frac{1}{r} \sum_{j=1}^p M_j \varepsilon_{j\alpha}$$

より

$$\begin{aligned}
 S_e &= \sum_{j=1}^p \sum_{\alpha=1}^n (\beta_\alpha M_j + \varepsilon_{j\alpha} - b_\alpha M_j)^2 = \sum_{j=1}^p \sum_{\alpha=1}^n [(\beta_\alpha - b_\alpha) M_j + \varepsilon_{j\alpha}]^2 \\
 &= \sum_{j=1}^p \sum_{\alpha=1}^n \left[ -\frac{1}{r} M_j \sum_{j'=1}^p M_{j'} \varepsilon_{j'\alpha} + \varepsilon_{j\alpha} \right]^2 \\
 &= \sum_{j=1}^p \sum_{\alpha=1}^n \left[ \frac{1}{r^2} M_j^2 \left( \sum_{j'=1}^p M_{j'} \varepsilon_{j'\alpha} \right)^2 - \frac{2}{r} M_j \varepsilon_{j\alpha} \sum_{j'=1}^p M_{j'} \varepsilon_{j'\alpha} + \varepsilon_{j\alpha}^2 \right] \\
 &= -\frac{1}{r} \sum_{\alpha=1}^n \sum_{j=1}^p \sum_{j'=1}^p M_j M_{j'} \varepsilon_{j\alpha} \varepsilon_{j'\alpha} + \sum_{j=1}^p \sum_{\alpha=1}^n \varepsilon_{j\alpha}^2
 \end{aligned}$$

となり、

$$E[S_e] = E \left[ -\frac{1}{r} \sum_{\alpha=1}^n \sum_{j=1}^p \sum_{j'=1}^p M_j M_{j'} \varepsilon_{j\alpha} \varepsilon_{j'\alpha} + \sum_{j=1}^p \sum_{\alpha=1}^n \varepsilon_{j\alpha}^2 \right] = (np - n)V(\varepsilon) \quad (\text{A1})$$

また、

$$b = \sum_{j=1}^p \sum_{\alpha=1}^n M_j y_{j\alpha} / \sum_{j=1}^p \sum_{\alpha=1}^n M_j^2 = \frac{1}{nr} \sum_{j=1}^p \sum_{\alpha=1}^n M_j y_{j\alpha}$$

より、

$$\begin{aligned}
 b^2 &= \frac{1}{n^2 r^2} \left( \sum_{j=1}^p \sum_{\alpha=1}^n M_j (\beta_\alpha M_j + \varepsilon_{j\alpha}) \right)^2 \\
 &= \frac{1}{n^2 r^2} \sum_{j=1}^p \sum_{\alpha=1}^n \sum_{j'=1}^p \sum_{\alpha'=1}^n M_j M_{j'} (\beta_\alpha \beta_{\alpha'} M_j M_{j'} + 2\beta_{\alpha'} \varepsilon_{j\alpha} M_{j'} + \varepsilon_{j\alpha} \varepsilon_{j'\alpha'}) \\
 &= \frac{1}{n^2 r^2} \left[ \beta^2 n^2 r^2 + 2\beta nr \sum_{j=1}^p \sum_{\alpha=1}^n \varepsilon_{j\alpha} + \sum_{j=1}^p \sum_{\alpha=1}^n \sum_{j'=1}^p \sum_{\alpha'=1}^n M_j M_{j'} \varepsilon_{j\alpha} \varepsilon_{j'\alpha'} \right]
 \end{aligned}$$

であるから、

$$\begin{aligned}
 E[b^2] &= \frac{1}{n^2 r^2} E \left[ \beta^2 n^2 r^2 + 2\beta nr \sum_{j=1}^p \sum_{\alpha=1}^n \varepsilon_{j\alpha} + \sum_{j=1}^p \sum_{\alpha=1}^n \sum_{j'=1}^p \sum_{\alpha'=1}^n M_j M_{j'} \varepsilon_{j\alpha} \varepsilon_{j'\alpha'} \right] \\
 &= \beta^2 + \frac{1}{n^2 r^2} \sum_{j=1}^p \sum_{\alpha=1}^n \sum_{j'=1}^p \sum_{\alpha'=1}^n M_j M_{j'} \delta_{jj'} \delta_{\alpha\alpha'} V(\varepsilon) \\
 &= \beta^2 + \frac{1}{nr} V(\varepsilon)
 \end{aligned}$$

よって、以下となる。

$$\beta^2 = E[b^2] - \frac{1}{nr} V(\varepsilon) \quad (\text{A2})$$

(A2)と(A1)、及び  $b^2 = S_\beta / nr$  の関係から、

$$\beta^2 = \frac{1}{nr} E[S_\beta] - \frac{1}{nr} E[S_e / (np - n)] = E[(S_\beta - V_e) / nr] \quad (A3)$$

すなわち、 $\beta^2$  の不偏推定量は  $(S_\beta - V_e) / nr$  である。

同様の考え方で  $\sigma^2$  の不偏推定量が  $V_N = S_N / (pn - 1)$  であることも示すことができる。

次に我々は SN 比を最大にする制御因子の最適設定について考える。制御因子 A, B, …について直交表を作ると、他の制御因子の影響をならした、1つの制御因子の影響を調べることができるようになる。表 2 に直交表を加えたデータを示す。

表 2 パラメータ設計におけるデータ

	A	B	…	$M_1$			…	$M_p$			SN 比	感度
				$N_1$	…	$N_n$		$N_1$	…	$N_n$		
1	1	1	…	$y_{111}$	…	$y_{11n}$	…	$y_{1p1}$	…	$y_{1pn}$	$\eta_1$	$S_1$
2	1	1	…	$y_{211}$	…	$y_{21n}$	…	$y_{2p1}$	…	$y_{2pn}$	$\eta_2$	$S_2$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
D	2	2	…	$y_{d11}$	…	$y_{d1n}$	…	$y_{dp1}$	…	$y_{dpn}$	$\eta_d$	$S_d$

ここに SN 比と感度は上で述べた方法で求めて加えてあるものとする。直交表は、各制御因子の同じ番号の行を見ると、他の制御因子について、すべての番号が同じ数だけ入っているという特徴を持つ。

例えば制御因子 A が  $k$  になる行について、SN 比及び感度の平均を取ったものをそれぞれ  $\bar{\eta}_{A=k}$ 、 $\bar{S}_{A=k}$  と書くにとすると、SN 比の補助表は表 3 のようになる。感度の補助表も同様である。

表 3 SN 比の補助表

制御因子	水準 1	…	水準 $r$
A	$\bar{\eta}_{A=1}$	…	$\bar{\eta}_{A=r}$
B	$\bar{\eta}_{B=1}$	⋮	$\bar{\eta}_{B=r}$
⋮	⋮	…	⋮

ここに水準の少ない制御因子の場合、その部分は空欄にしておく。

この補助表の SN 比の中で、制御因子ごとの水準値の最も大きな水準を並べたものを最適条件といい、例えば A1B2C1D3…などと表す。我々のプログラムでは制御因子名は省略して番号だけで表している。この最適な水準の SN 比を合計したものを SN 比の最適値という。感度についても SN 比の最適条件を用いて最適値を定義する。

これに対して現実の制御因子の設定を比較条件または現状条件という。この条件を用いて SN 比を合計したものを SN 比の比較値または現状値という。感度についても同様である。最適値と比較値の差は、今後の改善の可能性として検討すべき値である。

ここで述べた水準値は理論的な推測値である。この値が妥当なものかどうか、追実験をして検証しておかなければならない。また、現実的に考えて最適な制御条件が最良のものであるとは限らない。その際は、できるだけ SN 比の値を落とさず、感度で制御因子の調整を行うこともある。

## 10.2 プログラムの利用法

パラメータ設計のデータは図 1 のように、左の直交表の部分と右の実験結果の部分に分けられる。

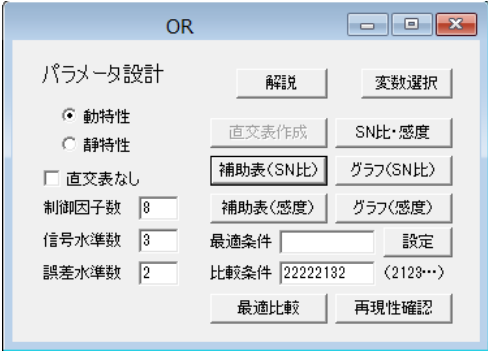


	E	F	G	H	0.025	0.025	0.1	0.1	0.5	0.5
1	1	1	1	1	0.046	0.040	0.068	0.056	0.235	0.276
2	2	2	2	2	0.052	0.036	0.085	0.078	0.293	0.311
3	3	3	3	3	0.071	0.079	0.111	0.109	0.327	0.321
4	1	2	2	3	0.043	0.037	0.072	0.050	0.224	0.246
5	2	3	3	1	0.058	0.038	0.083	0.082	0.283	0.252
6	3	1	1	2	0.076	0.074	0.104	0.131	0.371	0.404
7	2	1	3	2	0.030	0.027	0.062	0.065	0.278	0.298
8	3	2	1	3	0.044	0.036	0.108	0.086	0.354	0.361

図 1 パラメータ設計のデータ

ここでは、制御因子が A～H の 8 種類、信号因子が 3 種類、誤差因子が 2 種類である。信号因子と誤差因子の部分の変数名には信号因子の数値が与えられている。

メニュー [分析－OR－パラメータ設計] を選択すると図 2 のような分析メニューが表示される。



OR

パラメータ設計

☒ 動特性 ☐ 静特性

☐ 直交表なし

制御因子数: 8

信号水準数: 3

誤差水準数: 2

解説

変数選択

直交表作成

SN比・感度

補助表<SN比>

グラフ<SN比>

補助表<感度>

グラフ<感度>

最適条件

設定

比較条件: 22222132 <2123...>

最適比較

再現性確認

図 2 パラメータ設計分析メニュー

パラメータ設定には動特性と静特性の 2 種類あるが、今回は動特性のみについて紹介する。まずメニュー中にある、「制御因子数」、「信号水準数」、「誤差水準数」の値を入力する。この例題の場合、デフォルトの数値がそのまま利用できる。次に「変数選択」ボタンですべての変数を選択する。現在の例では直交表が付いているが、単純に SN 比と感度のみを求める場合には、直交表を省略したデータ

を用いることもできる。その際には「直交表なし」チェックボックスにチェックを入れておく。

「SN 比・感度」ボタンをクリックすると、図 3 の計算結果が表示される。

SN比・感度												
	A	B	C	D	E	F	G	H		SN比	感度	
▶ 1	1	1	1	1	1	1	1	1	1	26.534	-5.730	
2	1	1	2	2	2	2	2	2	2	28.258	-4.239	
3	1	1	3	3	3	3	3	3	3	23.149	-3.516	
4	1	2	1	1	2	2	3	3	3	26.656	-6.431	
5	1	2	2	2	3	3	1	1	1	25.083	-5.230	

図 3 各実験の SN 比・感度

ここでは各実験に対して、単純に SN 比と感度を求めて表示している。

直交表を使った SN 比の補助表は「補助表 (SN 比)」ボタンをクリックすることで図 4 のように与えられる。感度の補助表については「補助表 (感度)」ボタンをクリックして得られる。

補助表 (SN比)				
	水準1	水準2	水準3	MAX-MIN
▶ A	27.494	23.169		4.325
B	23.656	22.961	29.378	6.418
C	28.332	25.959	21.704	6.628
D	22.358	26.353	27.284	4.926
E	26.649	24.186	25.160	2.463
F	25.040	27.077	23.878	3.199
G	26.675	24.008	25.313	2.667
H	23.489	26.686	25.820	3.197

図 4 補助表 (SN 比)

ここで制御因子 A は 2 水準であるから、空白が 1 つできている。

補助表をグラフにした図は「グラフ (SN 比)」ボタンをクリックして表示される。描画結果を図 5 に示す。

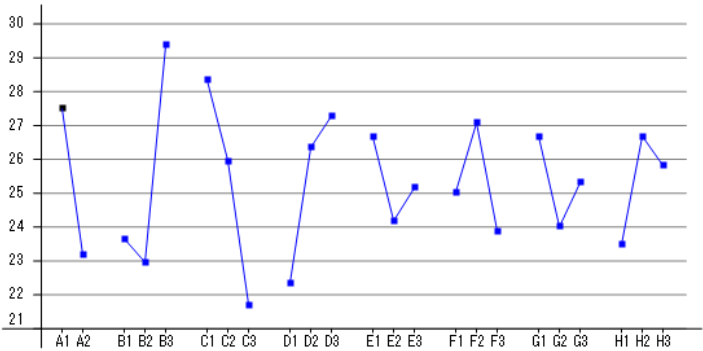


図 5 補助表のグラフ (SN 比)

図 4 と図 5 に対する感度の補助表とグラフは、それぞれ図 6 と図 7 で与えられる。

補助表（感度）				
	水準1	水準2	水準3	MAX-MIN
▶ A	-4.515	-3.975		0.540
B	-3.458	-5.041	-4.236	1.583
C	-4.337	-3.729	-4.669	0.940
D	-6.000	-4.117	-2.618	3.381
E	-3.963	-4.762	-4.011	0.799
F	-3.309	-4.239	-5.187	1.878
G	-4.045	-4.400	-4.290	0.355
H	-5.052	-3.968	-3.716	1.336

図 6 補助表（感度）

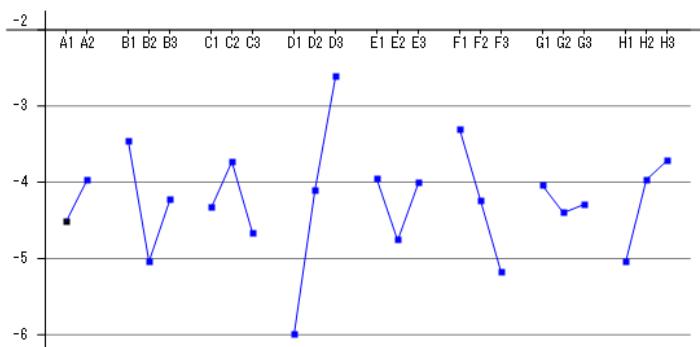


図 7 補助表のグラフ（感度）

SN 比の補助表やグラフを使った最適条件は「設定」ボタンをクリックすることでメニュー上の最適条件の部分に図 8 のように表示される。

パラメータ設計		OR	
<input checked="" type="radio"/> 動特性 <input type="radio"/> 静特性 <input type="checkbox"/> 直交表なし		解説    変数選択	
制御因子数 8 信号水準数 3 誤差水準数 2		直交表作成    SN比・感度 補助表<SN比>    グラフ<SN比> 補助表<感度>    グラフ<感度>	
		最適条件 13131212    設定	
		比較条件 22222132    (2123...)	
		最適比較    再現性確認	

図 8 SN 比の最適条件の設定

比較条件で現在の実験から得られるデータの値を求めることができるが、最適条件との比較も可能である。これらの数値は「最適比較」ボタンで得ることができる。表示結果を図 9 に示す。最適条件の制御因子の組み合わせを変えることで結果を手動で訂正することもできる。

最適条件
A1+B3+C1+D3+E1+F2+G1+H2-7T
SN比 = 42.254
感度 = -2.206
比較（現状）条件
A2+B2+C2+D2+E2+F1+G3+H2-7T
SN比 = 22.344
感度 = -3.476
利得（最適-比較）
SN比 = 19.910
感度 = 1.270

図 9 最適条件と比較条件

これらの最適条件と比較条件を実験で再現し、結果を得て、それをデータに追加する。当然その部分の直交表は空白になっているが、そのまま「再現性確認」ボタンをクリックすると図 10 に示す再現性確認表が得られる。

	SN比推定値	感度推定値	SN比実験値	感度実験値
▶ 最適条件	42.254	-2.206	44.369	-0.962
比較条件	22.344	-3.476	26.177	-4.059
利得	19.910	1.270	18.192	3.097

図 10 再現性確認表

参考文献

[1] 井上清和・中野恵司他, 入門パラメータ設計, 日科技連出版社, 2008.



## 11.1. オンライン品質工学

### 11.1 オンライン品質工学の考え方

工場の生産ラインでは、目標特性の製品を作るために、品質特性値を計測し、ラインを調整する。これらには時間と人手を要し、そこには費用が発生する。オンライン品質工学では、品質を金額ベースで表し、最適な計測間隔と調整間隔を求め、現行の値と比較し、どの程度の金額の削減につながるかを検討する。

#### 11.1.1 目的特性の規格値が $[-\Delta, \Delta]$ の場合

計算には以下の量を利用するが、規格値の中央値は0に設定している。特にどんな値でもずらして考えればよい。

不良品損失： $A$ ，計測コスト： $B$ ，調整コスト： $C$

計測間隔： $n$ ，調整間隔： $u$ ，調整限界： $D$ ，計測タイムラグ： $l$

製品1個当たりの計測コストは $B/n$ 、調整コストは $C/u$ である。また、製品1個当たりの品質損失（品質水準） $Q$ を以下のように与える<sup>[1]</sup>。

$$Q = \frac{A}{\Delta^2} \left[ \frac{D^2}{3} + \left( \frac{n+1}{2} + l \right) \frac{D^2}{u} \right]$$

ここに、最初の項は調整限界内のばらつき、次の項は問題があった計測時の調整限界を超えた製品の予測個数と計測タイムラグによる計測が遅れた製品の個数である。この品質損失の角括弧の中は誤差分散であり、その平方根を誤差標準偏差またはRMSと呼ぶ。

ここで与えた計測調整コストと品質損失を足して、製品1個当たりの総損失 $L$ は以下となる。

$$L = \frac{B}{n} + \frac{C}{u} + \frac{A}{\Delta^2} \left[ \frac{D^2}{3} + \left( \frac{n+1}{2} + l \right) \frac{D^2}{u} \right] \quad (1)$$

後の考察のために、品質損失（品質水準） $Q$ を、損失関数という考え方に基づいて、少し補足的説明も加えて考えてみる。まず損失関数を以下と仮定する。

$$Q(x) = \frac{A}{\Delta^2} x^2$$

ここに係数は $Q(\Delta) = A$ となるように決めている。

今、損失を調整限界内損失（限界内損失と略す） $Q_{in}$ と調整限界外損失（限界外損失と略す） $Q_{out}$ に分けて考える。

限界内損失のデータの確率分布を、範囲 $-D \leq x \leq D$ の一樣分布と考えると、確率密度関数は $1/2D$ となる。そのため損失関数の期待値は以下となる。

$$Q_{in} = \frac{1}{2D} \int_{-D}^D \frac{A}{\Delta^2} x^2 dx = \frac{A}{6D\Delta^2} [x^3]_{-D}^D = \frac{AD^2}{3\Delta^2}$$

次に限界外損失について考える。 $n$  個に 1 回の計測で管理外データが発見された場合、残りの  $n-1$  個については計測されていないのでエラーがあるかどうか分からない。そのため確率を  $1/2$  として期待値をとると、 $(n-1)/2$  となる。これと管理外データが発見された 1 個を足して、期待値は  $(n+1)/2$  となる。また、これにタイムラグ  $l$  による調整の遅れで管理外データが発生すると考えると、これらを合わせて、1 回の調整での管理外データの個数の期待値  $n_{out}$  は以下となる。

$$n_{out} = \frac{n+1}{2} + l$$

これが調整間隔  $u$  個に 1 回起こるので、商品 1 個当たりの管理外データの発生確率  $p_{out}$  は以下となる。

$$p_{out} = \frac{n_{out}}{u} = \left( \frac{n+1}{2} + l \right) \frac{1}{u} \quad (2)$$

管理外データの損失関数の値は  $x = D$  として、1 回につき  $AD^2/\Delta^2$  で与えられるので、商品 1 個当たりの限界外損失  $Q_{out}$  は以下で与えられる。

$$Q_{out} = \frac{A}{\Delta^2} \left( \frac{n+1}{2} + l \right) \frac{D^2}{u}$$

限界内損失と限界外損失を合わせて品質損失は以下となる。

$$Q = Q_{in} + Q_{out} = \frac{A}{\Delta^2} \left[ \frac{D^2}{3} + \left( \frac{n+1}{2} + l \right) \frac{D^2}{u} \right] \quad (3)$$

これは、分散という考えから求めた結果と同じである。これに計測費用と調整費用を加えて総費用は以下となる。

$$L = \frac{B}{n} + \frac{C}{u} + \frac{A}{\Delta^2} \left[ \frac{D^2}{3} + \left( \frac{n+1}{2} + l \right) \frac{D^2}{u} \right]$$

現行計測間隔を  $n_0$ ，現行調整限界を  $D_0$ ，現行調整間隔を  $u_0$  とすると、現行総損失  $L_0$  は以下のようになる。

$$L_0 = \frac{B}{n_0} + \frac{C}{u_0} + \frac{A}{\Delta^2} \left[ \frac{D_0^2}{3} + \left( \frac{n_0+1}{2} + l \right) \frac{D_0^2}{u_0} \right] \quad (4)$$

次に、(4)式に基づく最適な総損失を求めてみよう。そのために、我々は  $u = \lambda D^2$  という仮定を考える。パラメータ  $\lambda$  は、現行の値を使って  $\lambda = u_0/D_0^2$  で与えられる。この関係は限界調整間隔を小さく設定すると、平均調整間隔も小さくする必要があるということに基づく。しかし、これには計算

上の理由もある。(1)式にこの関係を代入すると以下となり、

$$L = \frac{B}{n} + \frac{C}{\lambda D^2} + \frac{A}{\Delta^2} \left[ \frac{D^2}{3} + \left( \frac{n+1}{2} + l \right) \frac{1}{\lambda} \right] \quad (5)$$

$n$  と  $D$  とは和の項として分離され、極小化条件を求めるために微分した後、単独に解を求めることが可能となる。他の関係では、解を求めるために、数値計算の必要が生じ、解析的には困難となる。

我々は(4)式を  $n$  と  $D$  とで微分して 0 とおき、最適計測間隔  $\hat{n}$  と最適調整限界  $\hat{D}$  を求める。これらは以下のように与えられる。

$$\hat{n} = \sqrt{\frac{2B\lambda\Delta^2}{A}} \quad , \quad \hat{D} = \left( \frac{3C\Delta^2}{A\lambda} \right)^{1/4} \quad (6)$$

これより、最適総損失  $\hat{L}$  は以下ようになる。

$$\hat{L} = \frac{B}{\hat{n}} + \frac{C}{\lambda \hat{D}^2} + \frac{A}{\Delta^2} \left[ \frac{\hat{D}^2}{3} + \left( \frac{\hat{n}+1}{2} + l \right) \frac{1}{\lambda} \right] \quad (7)$$

ここで、ハットの付いた量を含む式は  $u = \lambda D^2$  の仮定を入れていると考えてもらいたい。

次に  $u = \lambda D^2$  の条件を外してその他の指標をみてみよう。誤差分散  $\sigma^2$  は(1)式の[ ]の中で、以下のように与えられる。

$$\sigma^2 = \frac{D^2}{3} + \left( \frac{n+1}{2} + l \right) \frac{D^2}{u}$$

誤差の標準偏差を RMS と呼び、以下で与えられる。

$$RMS = \sigma = \sqrt{\frac{D^2}{3} + \left( \frac{n+1}{2} + l \right) \frac{D^2}{u}} \quad (8)$$

また、規格値の範囲  $\pm\Delta$  とこの標準偏差の  $\pm 3\sigma$  の範囲との比を工程能力指数  $C_p$  と呼び、精度評価の 1 つの基準としている。

$$C_p = \frac{2\Delta}{6\sqrt{\frac{D^2}{3} + \left( \frac{n+1}{2} + l \right) \frac{D^2}{u}}} \quad (9)$$

バッチ処理の場合には、 $A, n, u$  等はバッチ単位の数として総損失等は計算される。計測方法としてバッチ内のいくつかの製品についてサンプリング検査されるので、バッチ処理の場合には計測のバッチ内分散  $s_m^2$  も考慮しなくてはならない。そのため、品質損失（品質水準） $Q$  を以下のように考える。

$$Q = \frac{A}{\Delta^2} \left[ \frac{D^2}{3} + \left( \frac{n+1}{2} + l \right) \frac{D^2}{u} + s_m^2 \right] \quad (10)$$

これに関連して、総損失  $L$  や RMS、工程能力指数  $C_p$  も変更を受ける。プログラムではバッチ内分散の値が全体の分散の値に占める比率をバッチ内分散比率  $r_B$  として以下のように与えている。

$$r_B = \frac{s_B^2}{\frac{D^2}{3} + \left( \frac{n+1}{2} + l \right) \frac{D^2}{u} + s_m^2}$$

### 11.1.2 目的特性の規格値が $[0, \Delta]$ の場合

規格の下限值は 0 に設定している。特にどんな値でもずらして考えればよい。ここでは、区間が正に限られた場合（製品によって負になる場合も含む）の損失関数を以下と考える。

$$Q(x) = \frac{A}{\Delta^2} x^2$$

限界内損失のデータの確率分布を、範囲  $m-D \leq x \leq m+D$  の一様分布と考えると、確率密度関数は  $1/2D$  となる。そのため限界内損失関数の期待値は以下となる。

$$Q_{in} = \frac{1}{2D} \int_{m-D}^{m+D} \frac{A}{\Delta^2} x^2 dx = \frac{A}{6D\Delta^2} [x^3]_{m-D}^{m+D} = \frac{A}{3\Delta^2} (D^2 + 3m^2)$$

定数  $\Delta$  の考え方が異なるので注意を要するが、形式的には以前の結果に比べて  $3m^2$  だけ追加されている。

限界外損失の場合の損失関数の値は以下となる。)

$$\text{内側 (0 に近い側)} \quad Q(m-D) = A(m-D)^2 / \Delta^2$$

$$\text{外側 (0 から遠い側)} \quad Q(m+D) = A(m+D)^2 / \Delta^2$$

平均的に内側と外側で、確率  $1/2$  で管理外データが発生すると考えると、1 節の  $p_{out}$  を用いて限界外損失  $Q_{out}$  は以下ようになる。

$$\begin{aligned} Q_{out} &= \frac{1}{2} [Q(m-D) + Q(m+D)] p_{out} = \frac{A}{2\Delta^2} [(m+D)^2 + (m-D)^2] p_{out} \\ &= \frac{A}{\Delta^2} \left( \frac{n+1}{2} + l \right) \frac{D^2 + m^2}{u} \end{aligned}$$

ここで値が 0 未満になることも考えられるが、これは廃棄になることはないと考えて、損失関数がそのまま適用できると考える。廃棄になる場合は、それを見越して精度や中心の位置を調整する必要がある。

これより、品質損失は以下となる。

$$Q = Q_{in} + Q_{out} = \frac{A}{\Delta^2} \left[ \frac{D^2 + 3m^2}{3} + \left( \frac{n+1}{2} + l \right) \frac{D^2 + m^2}{u} \right] \quad (11)$$

これに計測費用と調整費用を加えて総費用は以下となる。

$$L = \frac{B}{n} + \frac{C}{u} + \frac{A}{\Delta^2} \left[ \frac{D^2 + 3m^2}{3} + \left( \frac{n+1}{2} + l \right) \frac{D^2 + m^2}{u} \right] \quad (12)$$

$m$  が定数の場合（例えば  $\Delta$  に比例するとしても）、1 節で述べた、 $n$  と  $D$  との和の項としての独立性は得られず、(4)式の最小化のためには連立方程式の数値計算が必要になる。しかし、 $m = \alpha D$  のように決めることができると、

$$L = \frac{B}{n} + \frac{C}{u} + \frac{A}{\Delta^2} \left[ \frac{(1+3\alpha^2)D^2}{3} + \left( \frac{n+1}{2} + l \right) \frac{(1+\alpha^2)D^2}{u} \right]$$

のようになり、前と同じ方法で計算が容易になる。

今、 $u = \lambda D^2$  を仮定すると

$$L = \frac{B}{n} + \frac{C}{\lambda D^2} + \frac{A}{\Delta^2} \left[ \frac{(1+3\alpha^2)D^2}{3} + \left( \frac{n+1}{2} + l \right) \frac{1+\alpha^2}{\lambda} \right] \quad (13)$$

我々は(4)式を  $n$  と  $D$  とで微分して 0 とおき、最適計測間隔  $\hat{n}$  と最適調整限界  $\hat{D}$  を求めると、以下のように与えられる。

$$\hat{n} = \sqrt{\frac{2B\lambda\Delta^2}{A(1+\alpha^2)}} \quad , \quad \hat{D} = \left( \frac{3C\Delta^2}{A\lambda(1+3\alpha^2)} \right)^{1/4} \quad (14)$$

誤差分散  $\sigma^2$  については、損失関数に関係なく、(1)式の[]の中で、以下のように与えられる。

$$\sigma^2 = \frac{D^2}{3} + \left( \frac{n+1}{2} + l \right) \frac{D^2}{u}$$

誤差の標準偏差を RMS も同様である。

$$RMS = \sigma = \sqrt{\frac{D^2}{3} + \left( \frac{n+1}{2} + l \right) \frac{D^2}{u}} \quad (15)$$

工程能力指数  $C_p$  については目的特性の範囲が異なるため、以下とする。

$$C_p = \frac{\Delta}{6\sqrt{\frac{D^2}{3} + \left( \frac{n+1}{2} + l \right) \frac{D^2}{u}}} \quad (16)$$

パッチ処理の場合には、前と同様に計測のパッチ内分散  $s_m^2$  も考慮して、品質損失（品質水準） $Q$

を以下のように考える。

$$Q = \frac{A}{\Delta^2} \left[ \frac{D^2 + 3m^2}{3} + \left( \frac{n+1}{2} + l \right) \frac{D^2 + m^2}{u} + s_m^2 \right] \quad (17)$$

$u = \lambda D^2$  と  $m = \alpha D$  の仮定の元に総費用は以下となる。

$$L = \frac{B}{n} + \frac{C}{\lambda D^2} + \frac{A}{\Delta^2} \left[ \frac{(1+3\alpha^2)D^2}{3} + \left( \frac{n+1}{2} + l \right) \frac{1+\alpha^2}{\lambda} + s_m^2 \right] \quad (18)$$

バッチ内分散の値は定数として与えるので、(14)式で与えられる最適計測間隔  $\hat{n}$  と最適調整限界  $\hat{D}$  には変更がない。

## 11.2 プログラムの利用法

メニュー「分析－OR－品質管理－オンライン品質工学」を選択すると図1に示す分析メニューが表示される。

図1 オンライン品質工学分析メニュー

このまま「実行」ボタンをクリックすると、画面上で与えられた数値を使って、図2に示す結果が表示される。

計算結果		
	現行	最適
▶ 計測間隔N	600.000	201.246
調整限界D	5.000	3.811
調整間隔U	1200.000	697.137
パラメータS	48.000	48.000
計測費/個	0.250	0.745
調整費/個	1.000	1.721
品質損失/個	5.196	2.478
総損失/個	6.446	4.944
工程能力指数Cp	1.308	1.894
RMS	3.823	2.640
バッチ内分散比率	0.000	0.000
累積計測回数	4.000	11.926
累積調整回数	2.000	3.443
累積計測費	600	1789
累積調整費	2400	4131
累積品質損失	12471	5947
累積総損失	15471	11867

図 2 画面からの実行結果

累積となっている部分は、メニューの「時間当たり生産数」と「時間」の値を使って全時間で発生する回数と金額を計算している。例えば 1 日 8 時間稼働とすると「時間」を 8 にして、「時間当たり生産数」を 1 時間当たりの生産数にする。単位時間を 1 日にして、年間の稼働日数を「時間」として設定してもよい。

現行のデータを複数与えて、最適解を求めるときには、ファイルからチェックボックスをチェックして、変数選択して「実行」ボタンを押す。しかし、このプログラムは最適な結果を出すだけでなく、徐々に最適な結果に近づけて行くときにも利用できる。この方法を参考文献に従って問題形式で考えて行く。

オンライン品質工学 1.txt のファイルを開くと、図 3 のように現行のデータが入力されている。

データ編集 オンライン品質工学1.txt				
製品 1				
▶ 規格値Δ	15			
不良品損失A(円)	80			
計測コストB(円)	150			
調整コストC(円)	1200			
現行計測間隔N0(個)	600			
現行調整限界D0(個)	5			
現行調整間隔U0(個)	1200			
タイムラグL(個)	1			
バッチ内標準偏差Sm	0			
時間当たり生産数	300			
1/3 (1.1)	分析:		備考:	

図 3 変更用ファイルデータ

データは 1 品目についてだけであるが、このデータを変更しながら最適な結果に近づけて行く。

質問 1 最適計測間隔と最適調整限界を求めて現行より、いくらの改善かを推定する。

この問題の解答の最後には簡単な金額計算をする必要があるので、メニュー画面には下に簡単な電卓機能を付けておいた。

図 3 のデータで変数選択を「All」にして、「ファイルから」チェックボックスにチェックを入れ、

「実行」ボタンをクリックすると、図4のような結果が出る。

計算結果								
	現行製品1	最適製品1	現行var2	最適var2	現行var3	最適var3	現行var4	最適var4
▶ 計測間隔N	600.000	201.246						
調整限界D	5.000	3.811						
調整間隔U	1200.000	697.137						
パラメータs	48.000	48.000						
計測費/個	0.250	0.745						
調整費/個	1.000	1.721						
品質損失/個	5.196	2.478						
総損失/個	6.446	4.944						
工程能力指数Cp	1.308	1.894						
RMS	3.823	2.640						
バッチ内分散比率	0.000	0.000						
累積計測回数	4.000	11.926						
累積調整回数	2.000	3.443						
累積計測費	600	1789						
累積調整費	2400	4131						
累積品質損失	12471	5947						
累積総損失	15471	11867						

図4 ファイルからの実行結果

変数が現行と最適合わせて4つ表示されるのはすべての変数を選んでいるためで、変数を1つずつ追加で選択して行く必要がないように、データの「規格Δ」の値が空欄のときは、関連する2列すべてを空欄になるように設定しているためである。もちろん他の列にデータがあるときには、その列も計算する。

最適な結果が出ているので、その値を参考にして、現行の値を変更することができる。グリッドエディタ上で、現行の値を次の列にコピーして必要な部分を変更した結果を図5に示す。

データ編集 オンライン品質工学1.txt				
製品1				
規格値Δ	15	15		
不良品損失A(円)	80	80		
計測コストB(円)	150	150		
調整コストC(円)	1200	1200		
現行計測間隔N0(個)	600	200		
現行調整限界D0(個)	5	4		
▶ 現行調整間隔U0(個)	1200	768		
タイムラグL(個)	1	1		
バッチ内標準偏差Sm	0	0		
時間当たり生産数	300	300		
1/3 (7.2)	分析:		備考:	

図5 現行値を変更した後のデータ

これを元に計算した結果が図6である。



計算結果								
	現行製品1	最適製品1	現行var2	最適var2	現行var3	最適var3	現行var4	最適var4
▶ 計測間隔N	600.000	201.246	200.000	201.246				
調整限界D	5.000	3.811	4.000	3.811				
調整間隔U	1200.000	697.137	768.000	697.137				
パラメータ	48.000	48.000	48.000	48.000				
計測費/個	0.250	0.745	0.750	0.745				
調整費/個	1.000	1.721	1.563	1.721				
品質損失/個	5.196	2.478	2.648	2.478				
総損失/個	6.446	4.944	4.961	4.944				
工程能力指数Cp	1.308	1.894	1.832	1.894				
RMS	3.823	2.640	2.729	2.640				
バッチ内分散比率	0.000	0.000	0.000	0.000				
累積計測回数	4.000	11.926	12.000	11.926				
累積調整回数	2.000	3.443	3.125	3.443				
累積計測費	600	1789	1800	1789				
累積調整費	2400	4131	3750	4131				
累積品質損失	12471	5947	6356	5947				
累積総損失	15471	11867	11906	11867				

図 6 現行値を変更した後の結果

ここでは現行の総損失の差をしてみる。1 個当たりの改善は 1.4850 円、1 時間 300 個、1 年 2000 時間として 89.1 万円となる。これはメニューの「時間」のところを「2000」に変えて出力すると、そのまま全時間の合計が計算されるので、差を見つけ易い。

工程能力指標  $C_p$  を見てみると現行 1.308、改善 1.832 となる。

最後に、計測に必要なマンパワーを求めてみよう。計測に 3 分、調整に平均 15 分かかるとする。計測回数は 1 日 8 時間として、1 日当たり 12 回、時間は  $12 \times 3 = 36$  分である。調整回数は 1 日当たり 3.443 回、時間は約  $3.443 \times 15 = 47.9$  分である。両方合わせて、約 82.9 分となる。そのため、1 日当たり  $82.9 / 480 = 0.173$  人必要である。

規格区間が  $[0, \Delta]$  となる場合については、図 1 の「 $[0, \Delta]$ 」ラジオボタンを選択し、その下の「中心」の位置を調整限界値の何倍に設定するかを決めて同じ処理を行うが、捜査は全く同じなので、説明は省略する。

#### 参考文献

[1] 矢野宏, 品質工学概論, 日本規格協会, 2009.

## 12. 異常検知

製造現場における検査過程では多くのデータが測定されるが、正常なデータと異常なデータの迅速な選別は品質管理の上で非常に重要である。ここではその主要な理論について説明し、それを実践するプログラムを紹介する。理論はデータが多変量正規分布に従うと仮定される場合とそうでない場合を扱う。データが多変量正規分布に従う場合、マハラノビス距離の2乗を元にしたホテリングの $t^2$ 統計量に基づく判定法を用いる。また、多変量正規分布に従わない場合は、確率的な解釈も可能な混合正規分布モデルを仮定する方法を用いている。また、1次元データについては、ガンマ分布による異常検知の方法も加えている。

### 12.1 複数変数の異常検知

#### 多変量正規分布に基づく異常検知

一般に  $p$  変数の多変量正規分布の密度関数は以下で与えられる。

$$f(\mathbf{x} | \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{\sqrt{(2\pi)^p |\boldsymbol{\Sigma}|}} \exp \left[ -\frac{1}{2} {}^t(\mathbf{x} - \boldsymbol{\mu}) \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right]$$

データ  $D = \{\mathbf{x}_1, \dots, \mathbf{x}_\lambda\}$  が与えられた場合の対数尤度関数は以下で与えられる。

$$L(\boldsymbol{\mu}, \boldsymbol{\Sigma} | D) = -\frac{pN}{2} \log(2\pi) - \frac{N}{2} \log |\boldsymbol{\Sigma}| - \frac{1}{2} \sum_{\lambda=1}^N {}^t(\mathbf{x}_\lambda - \boldsymbol{\mu}) \boldsymbol{\Sigma}^{-1} (\mathbf{x}_\lambda - \boldsymbol{\mu})$$

我々は最尤法を用いて(2)式を最大化するが、その解は以下となる。

$$\hat{\boldsymbol{\mu}} = \frac{1}{N} \sum_{\lambda=1}^N \mathbf{x}_\lambda, \quad \hat{\boldsymbol{\Sigma}} = \frac{1}{N} \sum_{\lambda=1}^N (\mathbf{x}_\lambda - \hat{\boldsymbol{\mu}}) {}^t(\mathbf{x}_\lambda - \hat{\boldsymbol{\mu}})$$

ここで、同じ正規分布の確率変数  $\mathbf{x}'$  に対する異常度  $a(\mathbf{x}')$  を  $-2 \log f(\mathbf{x}' | \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})$  を元に以下のように定義する。

$$a(\mathbf{x}') = {}^t(\mathbf{x}' - \hat{\boldsymbol{\mu}}) \hat{\boldsymbol{\Sigma}}^{-1} (\mathbf{x}' - \hat{\boldsymbol{\mu}})$$

ここで上式と  $-2 \log f(\mathbf{x}' | \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})$  の差は定数であるので、評価関数として本質的な差はない。また上式は1次元変数の場合の変数の標準化の一般形である。

異常度の式については、以下のように定数を掛けると、分布が自由度  $p, N-p$  のF分布に従うことが知られている。

$$T^2 \equiv \frac{N-p}{(N+1)p} {}^t(\mathbf{x}' - \hat{\boldsymbol{\mu}}) \hat{\boldsymbol{\Sigma}}^{-1} (\mathbf{x}' - \hat{\boldsymbol{\mu}}) \sim F_{p, N-p}$$

この  $T^2$  をホテリング統計量という。

異常検知には、この統計量を使って確率の値を指定するか、直接  $T^2$  値を指定して閾値とする。

### 混合多変量正規分布に基づく異常検知

$p$  変数、 $n$  群混合多変量正規分布の密度関数は、群  $\alpha$  の確率密度関数

$$f_{\alpha}(\mathbf{x} | \boldsymbol{\mu}_{\alpha}, \boldsymbol{\Sigma}_{\alpha}) = \frac{1}{\sqrt{(2\pi)^p |\boldsymbol{\Sigma}_{\alpha}|}} \exp \left[ -\frac{1}{2} {}^t(\mathbf{x} - \boldsymbol{\mu}_{\alpha}) \boldsymbol{\Sigma}_{\alpha}^{-1} (\mathbf{x} - \boldsymbol{\mu}_{\alpha}) \right] \quad (\alpha = 1, \dots, n)$$

を利用して以下で与えられる。

$$\begin{aligned} f(\mathbf{x} | \boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_{\alpha}, \boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_{\alpha}) &= \sum_{\alpha=1}^n \pi_{\alpha} f_{\alpha}(\mathbf{x} | \boldsymbol{\mu}_{\alpha}, \boldsymbol{\Sigma}_{\alpha}) \\ &= \sum_{\alpha=1}^n \frac{\pi_{\alpha}}{\sqrt{(2\pi)^p |\boldsymbol{\Sigma}_{\alpha}|}} \exp \left[ -\frac{1}{2} {}^t(\mathbf{x} - \boldsymbol{\mu}_{\alpha}) \boldsymbol{\Sigma}_{\alpha}^{-1} (\mathbf{x} - \boldsymbol{\mu}_{\alpha}) \right] \end{aligned}$$

ここに、 $\pi_{\alpha}$  は群  $\alpha$  の生起確率である。

この密度関数に従うデータ  $D = \{\mathbf{x}_1, \dots, \mathbf{x}_{\lambda}\}$  による対数尤度は以下である。

$$\begin{aligned} L(\pi_1, \dots, \pi_n, \boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_n, \boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_n | D) &= \sum_{\lambda=1}^N \log \left[ \sum_{\alpha=1}^n \pi_{\alpha} f_{\alpha}(\boldsymbol{\mu}_{\alpha}, \boldsymbol{\Sigma}_{\alpha} | \mathbf{x}_{\lambda}) \right] \\ &= \sum_{\lambda=1}^N \log \left[ \sum_{\alpha=1}^n \frac{\pi_{\alpha}}{\sqrt{(2\pi)^p |\boldsymbol{\Sigma}_{\alpha}|}} \exp \left\{ -\frac{1}{2} {}^t(\mathbf{x}_{\lambda} - \boldsymbol{\mu}_{\alpha}) \boldsymbol{\Sigma}_{\alpha}^{-1} (\mathbf{x}_{\lambda} - \boldsymbol{\mu}_{\alpha}) \right\} \right] \end{aligned}$$

最尤法を用いてこの対数尤度の最大値を求めるが、その際以下のアルゴリズムを利用する。

- 1) パラメータ  $\pi_{\alpha}, \boldsymbol{\mu}_{\alpha}, \boldsymbol{\Sigma}_{\alpha}$  に初期値  $\hat{\pi}_{\alpha}, \hat{\boldsymbol{\mu}}_{\alpha}, \hat{\boldsymbol{\Sigma}}_{\alpha}$  を与える。
- 2)  $\hat{\pi}_{\alpha}, \hat{\boldsymbol{\mu}}_{\alpha}, \hat{\boldsymbol{\Sigma}}_{\alpha}$  の値を用いて、各データの群  $\alpha$  への帰属度  $q_{\alpha}(\mathbf{x}_{\lambda})$  を以下で求める。

$$q_{\alpha}(\mathbf{x}_{\lambda}) = \frac{\hat{\pi}_{\alpha} f_{\alpha}(\mathbf{x}_{\lambda} | \hat{\boldsymbol{\mu}}_{\alpha}, \hat{\boldsymbol{\Sigma}}_{\alpha})}{\sum_{\beta=1}^n \hat{\pi}_{\beta} f_{\beta}(\mathbf{x}_{\lambda} | \hat{\boldsymbol{\mu}}_{\beta}, \hat{\boldsymbol{\Sigma}}_{\beta})}$$

- 3) この帰属度を使い、新しいパラメータを以下のように決定する。

$$\begin{aligned} \hat{\pi}_{\alpha} &= \frac{1}{N} \sum_{\lambda=1}^N q_{\alpha}(\mathbf{x}_{\lambda}), \quad \hat{\boldsymbol{\mu}}_{\alpha} = \sum_{\lambda=1}^N q_{\alpha}(\mathbf{x}_{\lambda}) \mathbf{x}_{\lambda} \bigg/ \sum_{\lambda=1}^N q_{\alpha}(\mathbf{x}_{\lambda}), \\ \hat{\boldsymbol{\Sigma}}_{\alpha} &= \sum_{\lambda=1}^N q_{\alpha}(\mathbf{x}_{\lambda}) (\mathbf{x}_{\lambda} - \hat{\boldsymbol{\mu}}_{\alpha}) {}^t(\mathbf{x}_{\lambda} - \hat{\boldsymbol{\mu}}_{\alpha}) \bigg/ \sum_{\lambda=1}^N q_{\alpha}(\mathbf{x}_{\lambda}) \end{aligned}$$

- 4) 新しいパラメータと元のパラメータを比較し、十分近ければ（プログラムではすべての成分が 0.001 未満）終了し、そうでなければ 2) へ戻る。

この方法によって求めたパラメータを使って、異常判定には以下の指標を用いる。

$$a(\mathbf{x}') = -\log \sum_{\alpha=1}^n \hat{\pi}_{\alpha} f_{\alpha}(\mathbf{x}' | \hat{\boldsymbol{\mu}}_{\alpha}, \hat{\boldsymbol{\Sigma}}_{\alpha})$$

判定基準はこの指標を小さい方から順番に並べ、分位点を閾値として決めるか、直接指標の閾値を指定する。

このモデルの適合度は赤池情報量基準 AIC、ベイズ情報量基準 BIC などを使って求める。今このモデルのパラメータ数を  $M_n$  とすると、AIC と BIC はそれぞれ以下のように表現される。

$$AIC = -2L(\Theta | D) + 2M_n$$

$$BIC = -2L(\Theta | D) + M_n \log N, \quad M_n = \frac{p}{2}(n+1)k +$$

ここに、 $L(\Theta | D)$  はパラメータを  $\Theta$  で代表させて書いた対数尤度である。具体的には(8)式に求めたパラメータの値を代入したものである。適合度を求めるために、交差検証を使う方法も考えられるが、プログラムでは使用していない。

## 12.2 時系列データの異常検知

時系列データの異常検知で、周波数変化や波形変化に対応した検知法に特異スペクトル変換法がある。これは、時系列データ  $x_{t-w}$  から  $x_{t-1}$  までの  $w$  個のデータを抽出し、それを 1 つのベクトル  ${}^t\mathbf{x}_{t-w} = (x_{t-w} \ x_{t-w+1} \cdots x_{t-1})$  とする。そのベクトルの開始時点をずらして  $k$  ( $k < w$ ) 本並べ、以下のような行列を作る。

$$\mathbf{X}_1^{(t)} = (\mathbf{x}_{t-k-w+1} \ \mathbf{x}_{t-k-w+2} \ \cdots \ \mathbf{x}_{t-w})$$

これに対して時間  $L$  だけ経過した時点から作ったデータを以下のように  $\mathbf{X}_2^{(t)}$  とする。

$$\mathbf{X}_2^{(t)} = (\mathbf{x}_{t-k-w+1+L} \ \mathbf{x}_{t-k-w+2+L} \ \cdots \ \mathbf{x}_{t-w+L})$$

この二つのデータの間で違いを見ることになる。

今、 ${}^t\mathbf{X}_1^{(t)}$ 、 ${}^t\mathbf{X}_2^{(t)}$  の各列ベクトルに係数をかけて足し合わせ、特徴的な量を求めるために、その大きさを以下のように最大化する。

$$\|{}^t\mathbf{X}_1^{(t)} \mathbf{u}^{(t)}\|^2 \rightarrow \text{最大化} \quad \text{但し、} {}^t\mathbf{u}^{(t)} \mathbf{u}^{(t)} = 1, \quad \mathbf{u}^{(t)} \text{ は } (w \times 1) \text{ ベクトル}$$

$$\|{}^t\mathbf{X}_2^{(t)} \mathbf{v}^{(t)}\|^2 \rightarrow \text{最大化} \quad \text{但し、} {}^t\mathbf{v}^{(t)} \mathbf{v}^{(t)} = 1, \quad \mathbf{v}^{(t)} \text{ は } (w \times 1) \text{ ベクトル}$$

このベクトルは、以下の固有方程式から求められ、

$$\mathbf{X}_1^{(t) \ t} \mathbf{X}_1^{(t)} \mathbf{u}^{(t)} = \lambda \mathbf{u}^{(t)}$$

$${}^t\mathbf{X}_2^{(t) \ t} \mathbf{X}_2^{(t)} \mathbf{v}^{(t)} = \mu \mathbf{v}^{(t)}$$

固有値の大きい順に固有ベクトルを  $m$  個並べて表した  $(w \times m)$  行列を以下のように定義する。

$$\mathbf{U}_m^{(t)} (w \times m) = (\mathbf{u}_1^{(t)} \ \mathbf{u}_2^{(t)} \ \cdots \ \mathbf{u}_m^{(t)})$$

$$\mathbf{V}_m^{(t)} (w \times m) = (\mathbf{v}_1^{(t)} \mathbf{v}_2^{(t)} \cdots \mathbf{v}_m^{(t)})$$

これら 2 つの行列を用いて、変化度  $a(t)$  は行列  ${}^t\mathbf{U}_m^{(t)}\mathbf{V}_m^{(t)} (m \times m)$  の 2 ノルムを用いて以下のよう  
に定義される。

$$a(t) = 1 - \left\| {}^t\mathbf{U}_m^{(t)}\mathbf{V}_m^{(t)} \right\|_2^2 = 1 - \left( {}^t\mathbf{U}_m^{(t)}\mathbf{V}_m^{(t)} \text{ の最大特異値} \right)^2$$

ここに、実行列  $\mathbf{A}$  の 2 ノルム  $\|\mathbf{A}\|_2$  は以下のように定義される。

$$\|\mathbf{A}\|_2 = \max_{\mathbf{u}} \sqrt{\frac{{}^t(\mathbf{A}\mathbf{u})\mathbf{A}\mathbf{u}}{{}^t\mathbf{u}\mathbf{u}}} = \sqrt{\lambda_{\max}}$$

注) 行列  $\mathbf{A}$  の特異値とは、行列  $\mathbf{A}$  と  $\mathbf{A}$  の随伴行列  $\mathbf{A}^*$  の積  $\mathbf{A}\mathbf{A}^*$  ( $\mathbf{A}^*\mathbf{A}$ ) の非負の固有値の平方根である。

### 12.3 入力と出力がある異常検知

入力と出力の関係で生じる異常の検知方法については、重回帰分析を用いる手法が考えられる。しかし、重回帰分析は、入力変数が多くその値が似通っている場合に、多重共線性の問題が発生する可能性があり、予測が不安定となる。これに対して改善方法と考えられている代表的な手法がリッジ回帰分析と PLS 回帰分析である。リッジ回帰分析は、多重共線性の元となる分散共分散行列に手を加える手法であり、PLS 回帰分析は多重共線性を与える変数間の自由度を制約する手法である。我々のプログラムは 3 者を比較するように作成しており、その違いを理解し易くなっている。

#### 重回帰分析

重回帰分析の目的変数を  $y_\lambda$  ( $\lambda = 1, 2, \dots, N$ )、説明変数を  $x_{i\lambda}$  ( $i = 1, 2, \dots, p$ ) とし、それらの関係を  $\varepsilon_\lambda$  を誤差項として以下とする。

$$y_\lambda = \sum_{i=1}^p b_i x_{i\lambda} + b_0 + \varepsilon_\lambda$$

最小 2 乗法としての重回帰分析では、以下の値  $D$  が最小になるように、パラメータ  $b_i, b_0$  を決定する。

$$D = \sum_{\lambda=1}^N \left( y_\lambda - \sum_{i=1}^p b_i x_{i\lambda} - b_0 \right)^2 = {}^t(\mathbf{y} - \mathbf{X}\mathbf{b})(\mathbf{y} - \mathbf{X}\mathbf{b})$$

ここに、

$$(\mathbf{X})_{\lambda i} = \tilde{x}_{i\lambda} = x_{i\lambda} - \bar{x}_i, \quad (\mathbf{y})_\lambda = \tilde{y}_\lambda = y_\lambda - \bar{y}, \quad \mathbf{b} = {}^t(b_1, b_2, \dots, b_p)$$

である。パラメータは以下で与えられる。

$$\mathbf{b} = ({}^t\mathbf{X}\mathbf{X})^{-1}{}^t\mathbf{X}\mathbf{y}, \quad b_0 = \bar{y} - \sum_{i=1}^p b_i \bar{x}_i$$

問題となる多重共線性は、行列 ${}^t\mathbf{X}\mathbf{X}$ の非正則性から生じる。

多重共線性について、 $i$  番目の説明変数を、他の説明変数で予測して重相関係数 $r_i$ を求め、以下の式で定義される VIF 指標を利用している。

$$VIF_i = 1/(1-r_i^2)$$

異常度について、通常重回帰分析では以下で定義する。

$$a(y', \mathbf{x}') = \frac{(y' - b_0 - \mathbf{b}\mathbf{x}')^2}{\sigma^2}$$

ここで分散 $\sigma^2$ については以下で推測する。

$$\sigma^2 = \frac{D}{N} = \frac{1}{N} {}^t(\mathbf{y} - \mathbf{X}\mathbf{b})(\mathbf{y} - \mathbf{X}\mathbf{b})$$

### リッジ回帰分析

リッジ回帰分析は重回帰分析の正則性の問題に対して、以下のように置くことによって正則性を確保しようとする手法である。

$$\mathbf{b}' = ({}^t\mathbf{X}\mathbf{X} + \eta\mathbf{I})^{-1}{}^t\mathbf{X}\mathbf{y}$$

これは、以下を最小化する解でもある。

$$D' = {}^t(\mathbf{y} - \mathbf{X}\mathbf{b}')(\mathbf{y} - \mathbf{X}\mathbf{b}') + \eta {}^t\mathbf{b}'\mathbf{b}'$$

ここでパラメータ $\eta$ の値は以下のようにして求められる。 $\alpha$  番目の個体を抜いた 1 個抜き交差検証のパラメータを $\mathbf{b}^{(-\alpha)}$ とすると、そのときの平均 2 乗誤差 $e(\eta)$ は以下で与えられることが知られている。

$$e(\eta) = \frac{1}{N} \sum_{\lambda=1}^N (\tilde{y}_\lambda - \sum_{i=1}^p \tilde{x}_{i\lambda} b'_i)^2$$

これは、以下のように書くこともできる。

$$e(\eta) = \frac{1}{N} {}^t\mathbf{A}\mathbf{A}$$

ここに、

$$\mathbf{A} = \text{diag}(\mathbf{I} - \mathbf{H})^{-1}(\mathbf{I} - \mathbf{H})\mathbf{y}, \quad \mathbf{H}(N \times N) = \mathbf{X}({}^t\mathbf{X}\mathbf{X} + \eta\mathbf{I})^{-1}{}^t\mathbf{X}$$

また、 $\text{diag}(\mathbf{I} - \mathbf{H})^{-1}$ は対角要素が $(\mathbf{I} - (\mathbf{H})_{ii})^{-1}$ となる対角行列である。運用上はパラメータ $\eta$ の値を変化させて、この $e(\eta)$ が最小になるようなパラメータ $\eta$ を選ぶ。

もう少し安全性を考えて、以下の一般化交差確認検証法と呼ばれる方法から与えられる誤差 $e_{GCV}(\eta)$ を最小化する場合もある。

$$e_{GCV}(\eta) = \frac{1}{N} {}^t \mathbf{A}' \mathbf{A}'$$

ここに、 $\mathbf{A}' = (\mathbf{I} - \mathbf{H})\mathbf{y} / [1 - \text{tr}\mathbf{H} / N]$  である。我々のプログラムでは前者の判定法を利用している。

多重共線性がある場合、重回帰分析の予測は、そのデータに対してだけは良い精度を与えるが、他の新しいデータを用いた場合、予測の精度が著しく低下する。そのため 1 個抜き交差検証は必須である。

異常度について、リッジ回帰分析では以下で定義する。

$$a(y', \mathbf{x}') = \frac{(y' - b'_0 - \mathbf{b}'\mathbf{x}')^2}{\sigma'^2}$$

ここで分散  $\sigma'^2$  については以下で推測する。

$$\sigma'^2 = \frac{D'}{N} = \frac{1}{N} [{}^t (\mathbf{y} - \mathbf{X}\mathbf{b})(\mathbf{y} - \mathbf{X}\mathbf{b}) + \eta {}^t \mathbf{b}' \mathbf{b}']$$

我々のプログラムでは、重相関係数と寄与率について、残差分散を以下で計算している。

$$\sigma^2 = \frac{1}{N} {}^t (\mathbf{y} - \mathbf{X}\mathbf{b}')(\mathbf{y} - \mathbf{X}\mathbf{b}')$$

そのため、出力される分散の値が他の分析（例えば以下の PLS 回帰）より大きいのに、重相関係数の値が大きく表示されるということもある。

## PLS 回帰分析

PLS 回帰分析ではまず、変数の線形結合を考える。

$$r_{i\lambda} = \sum_{j=1}^p u_{ij} \tilde{x}_{j\lambda} \quad (i=1, 2, \dots, r; r < p)$$

この式を行列記号を用いて書くと以下となる。

$$\mathbf{R} = \mathbf{X}\mathbf{U} \quad \mathbf{U} = (\mathbf{u}_1 \mathbf{u}_2 \cdots \mathbf{u}_r)$$

ここで、行列  $\mathbf{U}$  の各列ベクトルは直交し、順番に  $\mathbf{X}\mathbf{u}_i$  と  $\mathbf{y}$  との内積が最大化されるように選ばれる。

詳細は後に示す。

この新しい変数を用いて、目的変数を以下のように予測する。

$$\tilde{y}_\lambda = \sum_{j=1}^r \beta_j r_{j\lambda} + \varepsilon_\lambda$$

即ち、

$$\mathbf{y} = \mathbf{R}\boldsymbol{\beta} + \boldsymbol{\varepsilon} = \mathbf{X}\mathbf{U}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

最小 2 乗法を使うと、以下の量を最小化するようにパラメータを決定する。

$$D'' = {}^t (\mathbf{y} - \mathbf{R}\boldsymbol{\beta})(\mathbf{y} - \mathbf{R}\boldsymbol{\beta})$$

その解は次のように与えられる。

$$\boldsymbol{\beta} = ({}^t\mathbf{R}\mathbf{R})^{-1}{}^t\mathbf{R}\mathbf{y} = ({}^t\mathbf{U}'\mathbf{X}\mathbf{X}\mathbf{U})^{-1}{}^t\mathbf{U}'\mathbf{X}\mathbf{y}$$

これから、標準化偏回帰係数  $\tilde{\mathbf{b}}$  は以下となる。

$$\tilde{\mathbf{b}} = \mathbf{U}\boldsymbol{\beta}$$

また、通常の回帰係数は、以下で与えられる。

$$b_i'' = \tilde{b}_i s_y / s_i, \quad b_0'' = \bar{y} - \sum_{i=1}^p b_i'' \bar{x}_i$$

多重共線性の改善の程度については、変数を  $\mathbf{U}$  行列で変換した後の  $i$  番目の説明変数を、他の説明変数で予測して重相関係数  $r_i$  を求め、以下の式で定義される VIF 指標を利用している。

$$VIF_i = 1/(1-r_i^2)$$

異常度については以下で定義する。

$$a(\mathbf{y}', \mathbf{x}') = \frac{(\mathbf{y}' - b_0'' - \mathbf{b}''\mathbf{x}')^2}{\sigma^2}$$

ここで分散  $\sigma^2$  については以下で推測する。

$$\sigma^2 = \frac{D}{N} = \frac{1}{N} {}^t(\mathbf{y} - \mathbf{X}\mathbf{b}'')(\mathbf{y} - \mathbf{X}\mathbf{b}'')$$

最後に行列  $\mathbf{U} = (\mathbf{u}_1 \mathbf{u}_2 \cdots \mathbf{u}_r)$  の決定法について述べる。この行列の 1 列目  $\mathbf{u}_1$  は  $\mathbf{X}\mathbf{u}_1$  が最も  $\mathbf{y}$  の方向に向くように、以下のように求める。

$$L_1 = {}^t\mathbf{y}\mathbf{X}\mathbf{u}_1 - \mu_1({}^t\mathbf{u}_1\mathbf{u}_1 - 1) \rightarrow \text{最大化}$$

この解は以下で与えられる。

$$\mathbf{u}_1 = {}^t\mathbf{X}\mathbf{y} / \|{}^t\mathbf{X}\mathbf{y}\|$$

次の  $\mathbf{u}_2$  については、 $\mathbf{X}$  から  $\mathbf{u}_1$  方向の成分を取り除き、以下のように求める。

$$L_2 = {}^t\mathbf{y}(\mathbf{X} - \mathbf{d}_1 {}^t\mathbf{d}_1\mathbf{X})\mathbf{u}_2 - \mu_2({}^t\mathbf{u}_2\mathbf{u}_2 - 1) \rightarrow \text{最大化}$$

ここに、 $\mathbf{d}_1 = \mathbf{X}\mathbf{u}_1 / \|\mathbf{X}\mathbf{u}_1\|$  である。確かに  $\mathbf{X} - \mathbf{d}_1 {}^t\mathbf{d}_1\mathbf{X}$  は、以下のように  $\mathbf{u}_1$  方向の成分を取り除いている。

$$(\mathbf{X} - \mathbf{d}_1 {}^t\mathbf{d}_1\mathbf{X})\mathbf{u}_1 = \mathbf{X}\mathbf{u}_1 - \mathbf{X}\mathbf{u}_1 {}^t(\mathbf{X}\mathbf{u}_1)\mathbf{X}\mathbf{u}_1 / \|\mathbf{X}\mathbf{u}_1\|^2 = \mathbf{0}$$

この解は以下で与えられる。

$$\mathbf{u}_2 = {}^t(\mathbf{X} - \mathbf{d}_1 {}^t\mathbf{d}_1\mathbf{X})\mathbf{y} / \|{}^t(\mathbf{X} - \mathbf{d}_1 {}^t\mathbf{d}_1\mathbf{X})\mathbf{y}\|$$

このベクトル  $\mathbf{u}_2$  は以下のように  $\mathbf{u}_1$  と直交する。

$${}^t\mathbf{u}_1\mathbf{u}_2 \propto {}^t\mathbf{u}_1 {}^t(\mathbf{X} - \mathbf{d}_1 {}^t\mathbf{d}_1\mathbf{X})\mathbf{y} = {}^t(\mathbf{X}\mathbf{u}_1 - \mathbf{X}\mathbf{u}_1 {}^t(\mathbf{X}\mathbf{u}_1)\mathbf{X}\mathbf{u}_1 / \|\mathbf{X}\mathbf{u}_1\|^2)\mathbf{y} = 0$$



これを続けると、 $k$  番目の係数ベクトル  $\mathbf{u}_k$  は以下のように求められることが分かる。

$$L_k = {}^t\mathbf{y} \left( \mathbf{X} - \sum_{i=1}^{k-1} \mathbf{d}_i {}^t\mathbf{d}_i \mathbf{X} \right) \mathbf{u}_k - \mu_k ({}^t\mathbf{u}_k \mathbf{u}_k - 1) \rightarrow \text{最大化}$$

$$\mathbf{u}_k = \left( \mathbf{X} - \sum_{i=1}^{k-1} \mathbf{d}_i {}^t\mathbf{d}_i \mathbf{X} \right) \mathbf{y} / \left\| \left( \mathbf{X} - \sum_{i=1}^{k-1} \mathbf{d}_i {}^t\mathbf{d}_i \mathbf{X} \right) \mathbf{y} \right\|$$

どこまでの次元数を求めればよいかは、1つの方法として1個抜き交差検証法の重相関係数または残差分散の大きさを元にして決めればよい。我々のプログラムではこの方法を用いている。

#### 4. プログラムの利用法

メニュー「分析－OR－品質管理－異常検知」を選択すると、図1のような実行メニューが表示される。

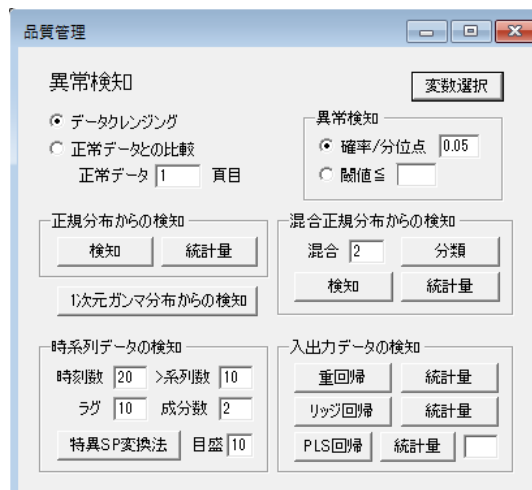


図1 異常検知実行メニュー

このプログラムでは、グリッドエディタに表示されているデータから、異常データを選別する「データクレンジング」機能と、別頁にある正常データを用いて、現在表示されているデータの異常データを選別する「正常データとの比較」機能がある。後者の場合は、正常データがどの頁にあるかを指定しなければならない。

分析には、(多変量)「正規分布からの検知」、(多変量)「混合正規分布からの検知」、「時系列データの検知」、「入出力データの検知」がある。最初の「正規分布からの検知」は1つの正規分布からのデータのずれを検知するもので、マハラノビス距離を基にした手法である。これは分布が限定されているが、多くのデータでほぼ正規分布の仮定が成り立つと考えられるので、利用範囲は広い。「混合正規分布からの検知」は、正規分布が仮定できない場合で、しかもどのような分布になっているか予想が困難な場合に適用が可能である。これは、分布を複数の正規分布の重ね合わせとして考えるモデ

ルで、いくつかの正規分布の重ね合わせて考えると効果的かという判断も可能である。「時系列データの検知」では、最も実用的と思われる「特異 SP 変換法」（特異スペクトル変換法）が利用できる。「入出力データの検知」では、「重回帰」分析を基本として、多重共線性のある場合の手法で「リッジ回帰」分析と「PLS 回帰」分析が利用可能である。

他に「1次元ガンマ分布からの検知」があるが、これは変数が1つの場合にしか適用できないので、現実には利用しにくいかも知れない。以後、まず2変数のデータを用いて、正規分布と混合席分布からの検知について、プログラムの説明をする。

図2に示すファイル「異常検知1(正規分布).txt」の3頁目は、2変数で異常値を含んだデータである。

	身長	体重
1	148	41
2	160	49
3	159	45
4	153	43
5	151	42
6	140	29
7	158	49
8	137	31
9	149	47
10	160	47
11	151	42

図2 異常値を含んだデータ

この中から異常値を検出するには、「データクレンジング」ラジオボタンを選び、「確率/分位点」ラジオボタンを選んで、異常値の確率値を指定する（この場合は確率値となる）。ここでは5%に設定している。その後、正規分布からの異常検知の「検知」ボタンをクリックすると図3のような出力結果を得る。

	Maha2乗	異常度(f値)	確率	異常
15	2.011	0.908	0.415	0
16	2.851	1.288	0.292	0
17	0.658	0.297	0.745	0
18	1.026	0.463	0.634	0
19	0.565	0.255	0.776	0
20	1.789	0.808	0.456	0
21	14.438	6.520	0.005	1
22	1.718	0.776	0.470	0
23	3.303	1.492	0.242	0
24	0.310	0.140	0.870	0

図3 データクレンジング検知結果

出力は、このデータから求めた多変量正規分布の平均からのマハラノビス距離の2乗、それを元にしたホテリング統計量（F分布のf値）、その検定確率、異常かどうかの判定である。判定は正常と異常でそれぞれ0または1で出力される。

次に、「統計量」ボタンをクリックすると、図 4 のように、平均や共分散等のパラメータ推定値等と共に、異常の判定に使われるホテリング統計量の閾値が出力される。利用者はこの値を参考にして、閾値としてホテリング統計量を用いてもよい。



	身長	体重
▶ 平均	149.000	38.700
分散共分散		
身長	51.733	39.433
体重	39.433	40.343
閾値確率	0.050	
自由度	2.28	
閾値(Ht2-f値)	3.340	

図 4 パラメータ推定値

次に「正常データとの比較」ラジオボタンをクリックし、同じファイルの 1 頁目を開く、正常データは 2 頁目に入っているものとして、「正常データ」テキストボックスの中に 2 を入力する。「検知」ボタンをクリックすると、図 5 のように 2 頁目の正常データから求められるパラメータを元にした、異常検知結果が出力される。

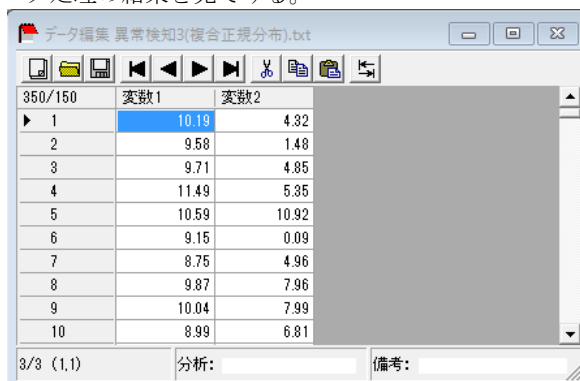


	Maha2乗	異常度(f値)	確率	異常
▶ 1	2.295	1.036	0.368	0
2	0.484	0.219	0.805	0
3	10.957	4.948	0.014	1
4	0.368	0.166	0.848	0
5	2.384	1.077	0.354	0
6	0.645	0.292	0.749	0
7	2.763	1.248	0.303	0
8	2.048	0.925	0.408	0
9	2.412	1.089	0.350	0
10	6.417	2.898	0.072	0

図 5 正常データとの比較検知結果

出力項目については図 3 と同様である。「統計量」ボタンをクリックした結果は、正常データを元にした結果であり、図 4 と同じ様式であるので省略する。

ファイル「異常検知 3(複合正規分布).txt」を図 6 のように読み込み、「データクレンジング」ラジオボタンを選択し、データ処理の結果を見てみる。



	変数1	変数2
▶ 1	10.19	4.32
2	9.58	1.48
3	9.71	4.85
4	11.49	5.35
5	10.59	10.92
6	9.15	0.09
7	8.75	4.96
8	9.87	7.96
9	10.04	7.99
10	8.99	6.81

3/3 (1,1)      分析:      備考:

図 6 非正規分布のデータ

このデータは、実際には 2 つの正規分布を合わせたものであるが、今の段階ではそれが分からないものとする。仮に「混合」テキストボックスを 2 とし、処理を進める。後にこの数字を変更して最も良いモデルを選択する。

「分類」ボタンをクリックすると、図 7 のように、2 つの群についてのデータの帰属度と分類結果が得られる。

	帰属度1	帰属度2	分類
343	0.999	0.001	1
344	1.000	0.000	1
345	0.999	0.001	1
346	0.984	0.016	1
347	0.999	0.001	1
348	0.999	0.001	1
349	0.999	0.001	1
350	1.000	0.000	1
351	0.000	1.000	2
352	0.000	1.000	2

図 7 レコード毎の帰属度と分類結果

この分類結果をコピーして、元のデータに図 8 のように貼り付け、

350/150	変数1	変数2	
1	10.19	4.32	1
2	9.58	1.48	1
3	9.71	4.85	1
4	11.49	5.35	1
5	10.59	10.92	1
6	9.15	0.09	1
7	8.75	4.96	1
8	9.87	7.96	1
9	10.04	7.99	1
10	8.94	6.81	1

図 8 分類項目の貼り付け

分析「相関と回帰分析」の「先頭列で群分け」ラジオボタンを選択して、図 9 のような散布図を描くことも可能である。

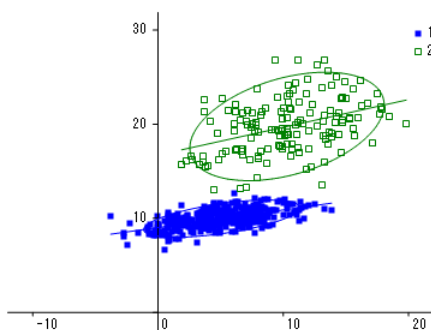


図 9 データの散布図

但し、このグラフには、グラフメニュー「設定－正規楕円半径」を用いて、 $2\sigma$ の確率楕円を加えてある。

次に「統計量」ボタンをクリックすると、図 10 のように 2 つの群の平均と共分散、群の生起確率の推測値などが表示される。この中で一番下の BIC はモデル判定によく利用されるベイズ情報量基準と呼ばれるもので、この値が小さいほど良いモデルとして評価される。

	変数1	変数2
群1		
生起確率	0.697	
平均	9.973	5.091
共分散		
変数1	0.992	1.845
変数2	1.845	9.744
群2		
生起確率	0.303	
平均	19.738	10.267
共分散		
変数1	8.540	4.666
変数2	4.666	15.193
閾値(a値)	7.305	
対数尤度	-2392.768	
AIC	4809.535	
BIC	4860.110	

図 10 2 群の場合のパラメータ推定値

現在の 2 群の場合は BIC=4860.110 であるが、3 群にすると 4888.488 となり、2 群の方が良いモデルであると判断される。これは、2 群を故意に作ったモデルであるので、当然の結果である。

また、「検知」ボタンをクリックすると、図 11 のように異常度の値と判別結果が表示される。

	異常度(a値)	異常
97	5.259	0
98	4.859	0
99	4.379	0
100	3.285	0
101	3.165	0
102	3.323	0
103	7.567	1
104	3.786	0
105	3.606	0
106	3.473	0

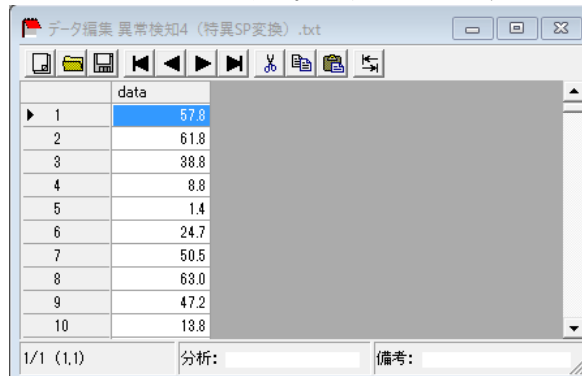
図 11 混合正規分布からの検知

判定には「確率/分位点」ラジオボックスの分位点を利用している。

1 次元ガンマ分布でも同様の結果の表示となるので、ここでは省略するが、ガンマ分布の場合は、2 つのパラメータが推測される。

時系列データからの異常検知では、正常データを指定しないので、図 1 の分析メニューでは「デー

「タクレンジング」ラジオボタンを選んで実行する。時系列データは、図 12 のように与えられる。



	data
1	57.8
2	61.8
3	38.8
4	8.8
5	1.4
6	24.7
7	50.5
8	63.0
9	47.2
10	13.8

図 12 時系列データ (異常検知 4(特異 SP 変換).txt)

異常検知の方法は簡単で、変数選択をして、「特異 SP 変換」ボタンをクリックする。その際、必要ならば「時刻数」、「系列数」、「ラグ」、「成分数」の値を変更する。これらの値はそれぞれ理論の説明の中の、 $w, k, L, m$ に相当する。「目盛」の値は、グラフの目盛間隔の値を与える。図 13 に実行結果を示す。

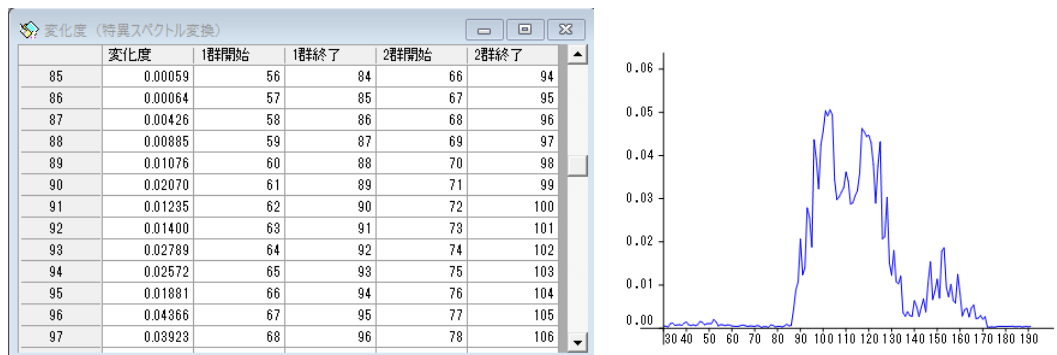


図 13 特異スペクトル変換による異常検知結果

図 13 の表の中で、左端が時刻であり、右 4 列は 2 つの群が利用したデータの範囲を示している。1 群は時刻の 1 つ前までのデータを利用し、2 群はそれから「ラグ」だけ遅れた時刻のデータを使っている。

入力と出力があるデータでは、重回帰分析が基本となる。通常の重回帰分析と同様に、目的変数、説明変数の順に変数を選んで、図 1 の分析メニューの「重回帰分析」ボタンをクリックすると図 14 のような実行結果が表示される。



	実測値	予測値	残差	異常度	判定
4	39	40.859	-1.859	0.160	0
5	81	84.237	-3.237	0.486	0
6	47	53.870	-6.870	2.192	0
7	92	91.659	0.341	0.005	0
8	75	76.076	-1.076	0.054	0
9	77	79.557	-2.557	0.304	0
10	64	61.139	2.861	0.380	0
11	76	82.910	-6.910	2.217	0
12	97	85.622	11.378	6.011	1
13	65	65.551	-0.551	0.014	0

図 14 重回帰分析実行結果（異常検知 5(入出力).txt 1 頁目）

また、隣の「統計量」ボタンをクリックすると、図 15 のような結果が表示される。

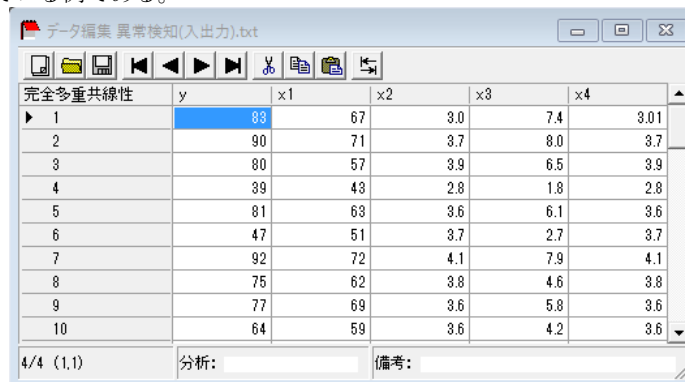


	偏回帰係数	標準化係数	VIF	残差分散	閾値	重相関R	寄与率R <sup>2</sup>
▶ x1	0.1490	0.0749	1.1796	21.5362	4.0705	0.943	0.889
x2	2.2233	0.0527	1.0554				
x3	2.7614	0.3522	3.9544				
x4	0.4314	0.5888	3.9670				
切片	6.8654	0.0000					

図 15 重回帰分析統計量（異常検知 5(入出力).txt 1 頁目）

これは正常に重回帰分析が行われた結果である。

データに多重共線性がある場合、例えば、図 16 は変数 x2 と x4 が殆ど同じで、1 番目のデータだけが、0.01 違っている例である。



完全多重共線性	y	x1	x2	x3	x4
▶ 1	83	67	3.0	7.4	3.01
2	90	71	3.7	8.0	3.7
3	80	57	3.9	6.5	3.9
4	39	43	2.8	1.8	2.8
5	81	63	3.6	6.1	3.6
6	47	51	3.7	2.7	3.7
7	92	72	4.1	7.9	4.1
8	75	62	3.8	4.6	3.8
9	77	69	3.6	5.8	3.6
10	64	59	3.6	4.2	3.6

4/4 (1.1)      分析:      備考:

図 16 強い多重共線性があるデータ（同 4 頁目）

このデータに対して重回帰分析の「統計量」ボタンをクリックすると、図 17 のような結果が得られる。



	偏回帰係数	標準化係数	残差分散	閾値	重相関R	寄与率R <sup>2</sup>
▶ x1	0.2251	0.1132	38.0310	3.3607	0.897	0.804
x2	521.1502	12.3456	共線性指標	0.0000		
x3	6.7407	0.8596				
x4	-519.5957	-12.2978				
切片	18.5670	0.0000				

図 17 強い多重共線性があるデータの重回帰分析統計量（同 4 頁目）

この結果ではほぼ同じ値である変数x2とx4の係数が極端に大きくなっている。これは予測において、x2とx4の値が少しずれると大きな差となって表れることを示しており、予測の頑健性において問題となる。

これを解決するための代表的な手法に、リッジ回帰分析とPLS回帰分析がある。理論のところで示したように、リッジ回帰分析は多重共線性の元となる説明変数の共分散行列の対角成分に、ある定数（ここでは、 $\eta \div N$ ）を加える方法である。この定数 $\eta$ は、1個抜き交差検証の残差分散が最小となるように決める。

図16のデータに対して、「リッジ回帰分析」とその「統計量」を与えた結果を図18と図19に示す。

	実測値	予測値	残差	異常度	判定
4	39	41.477	-2.477	0.267	0
5	81	84.407	-3.407	0.505	0
6	47	53.606	-6.606	1.899	0
7	92	90.447	1.553	0.105	0
8	75	76.297	-1.297	0.073	0
9	77	79.579	-2.579	0.290	0
10	64	60.639	3.361	0.492	0
11	76	82.966	-6.966	2.112	0
12	97	84.711	12.289	6.573	1
13	65	65.043	-0.043	0.000	0

図18 リッジ回帰分析による異常検知

	偏回帰係数	標準化係数	残差分散	閾値	重相関R	寄与率R^2
x1	0.1693	0.0851	22.9761	3.9367	0.942	0.887
x2	0.8338	0.0198			最良の	8.600
x3	2.2780	0.2905				
x4	0.4691	0.6402				
切片	9.9390	0.0000				

図19 リッジ回帰分析統計量

また、PLS回帰分析は多重共線性に対して、独立成分として与える変数の数を減らす方法で対応する。どれだけの変数を減らすかは、リッジ回帰分析と同様に1個抜き交差検証の残差分散が最小となるように決める。図20と図21にメニューの指定で独立成分を3にしたPLS回帰分析を実行した結果を示す。

	実測値	予測値	残差	異常度	判定
4	39	41.067	-2.067	0.196	0
5	81	83.798	-2.798	0.359	0
6	47	54.035	-7.035	2.268	0
7	92	92.300	-0.300	0.004	0
8	75	75.417	-0.417	0.008	0
9	77	79.388	-2.388	0.261	0
10	64	61.694	2.306	0.244	0
11	76	82.522	-6.522	1.949	0
12	97	85.634	11.366	5.921	1
13	65	66.501	-1.501	0.103	0



図 20 PLS 回帰分析による異常検知



	偏回帰係数	標準化係数	r-VIF	残差分散	関値	重相関R	寄与率R <sup>2</sup>
▶ x1	0.1312	0.0659	1.0160	21.8210	4.9222	0.942	0.888
x2	2.7116	0.0642	1.1392	交差検証R	0.9337	最良自由度	3
x3	3.3141	0.4227	1.1233				
x4	0.3820	0.5214					
切片	7.3511	0.0000					

図 21 PLS 回帰分析統計量

「最良自由度」は、目的変数の最適な自由度である。PLS 回帰分析の寄与率の値は、他に比べると少し落ちているが、リッジ回帰分析、PLS 回帰分析とも重回帰分析で問題となった変数  $x_2$  と  $x_4$  の係数が小さくなっている。これにより、これらの変数のずれの影響は小さく抑えられる。

## 参考文献

[1] 井出剛, 入門機械学習による異常検知, コロナ社, 2015.