

1984年10月27日  
JSC第14回年次大会  
研究発表番号 4-17

カイ2乗(2)プロット

日本エステル㈱ 松本哲夫\*  
 ㈱日本科学技術連盟 DE0部会

1. はじめに

最初の実験の時には、(1)実験の場の誤差はどれくらいあるのか、(2)影響力の大きい因子はどれくらいあるのか、(3)影響力の大きい因子は誤差に比べてどれくらい大ききであるのか、ということに対して事前情報はないことが多い。このような場合には、直交配列表を用いて実験を行い、誤差列から実験の場の誤差を推定することが多い。

しかし、この場合どこまでを誤差と考えるのかが釈然としないうし、全列に因子を割り付けてしまった時には分散分析そのものが不可能となる。このような場合にも適用できる図的分散分析法として、Daniel (1959) ①は2水準の直交配列表に対し、Half-Normal Plotを導入した。この方法の一部理論付けは、Birnbaum (1959) ②によってなされ、日本においても吉村③によって紹介されている。近年Zahn (1975) ④はその一部を修正して、問題点の抽出と検出力を実例やシミュレーションによって検討している。

本報においては、この考え方を3水準の直交配列表に拡張し、自由度2のカイ2乗プロットを導入した。

2. 理論

2.1 確率密度関数

3水準の直交配列表を用いて実験を行った時、結果を解析すれば表1の分散分析表が得られる。これらの平方和のうち、主効果の平方和の自由度は2である。一方、2因子間の交互作用の平方和の自由度は4であるが、Yates型の分解をすればそれぞれ自由度2の平方和に分けて考えられる。

表1. 分散分析表

S.V.	d.f.	S.S.
A	2	S(A)
B	2	S(B)
A × B (1)	2	S(A × B (1))
A × B (2)	2	S(A × B (2))
C	2	S(C)
⋮	⋮	⋮
⋮	⋮	⋮
X	2	S(X)

誤差の分散を $\sigma^2$ とすれば総ての要因効果が零であるとの帰無仮説の下では、 $S(i) / \sigma^2 \equiv \chi^2_i$ は自由度2のカイ2乗分布に従う。  
 $(i = A, B, A \times B (1), A \times B (2), C, \dots, X)$   
 さて、自由度2のカイ2乗分布は母数2の指数分布と一致しそのp.d.f.は①式で与えられる。

$$f(x) = \frac{1}{2} e^{-\frac{1}{2}x} \quad (x \geq 0)$$

$$= 0 \quad (x < 0) \quad \text{①}$$

また、 $x$ が0から $x$ までの値をとる確率、即ち $x$ の分布関数は②式で与えられる。

$$P = F(x) = \int_0^x f(x) dx$$

$$= 1 - e^{-\frac{1}{2}x} \quad (x \geq 0) \quad \text{②}$$

一方、 $x$ の期待値及び分散はそれぞれ③、④式で与えられる。

$$E(x) = \int_0^\infty x \cdot \frac{1}{2} e^{-\frac{1}{2}x} dx = 2 \quad \text{③}$$

$$V(x) = \int_0^\infty (x - E(x))^2 \cdot \frac{1}{2} e^{-\frac{1}{2}x} dx = 4 \quad \text{④}$$

②式を変形すると、

$x = -2 \ln(1 - F(x))$  となるので、任意の $S(i)$ を $-2 \ln(1 - F(x))$ に対してプロットした結果、原点を通る直線に旨く乗るようなら、その直線の傾きから誤差分散が推定できる。たとえ誤差項がない場合であっても各要因効果が何れも零であるとの帰無仮

説が成立するかどうかを、少なくとも直観的に判定することは可能となるであろう。

一方、現実の実現値をプロットする場合には、誤差分散が未知であるから、ある程度客観的な基準の下で検定、推定を行うことが望ましく、以下にその方法を考えることにする。

まず、 $X_1, X_2, \dots, X_n$  のサンプリル  $(X_i \sim \chi^2_{ID(2)})$  を大ききの順に並べ換え、 $X(1) \leq X(2) \leq X(3) \leq \dots \leq X(a) \leq \dots \leq X(n)$  とする。そこで、Zahnの考え方と同様に、横軸は  $x$  でなく  $Z(i, n)$ 、即ち  $X(i)$  の期待値とすることを考えよう。 $Z(i, n)$  は⑤式で与えられ、表2に  $Z(i, n)$  と  $-21n(1 - \pi_i)$  の計算結果を示した。ここで、 $\pi_i$  は⑥式で計算される。表より、 $i$  が大きくなれば、両者の差は殆どなくなる。

$$Z(i, n) = \frac{2 \cdot n!}{(i-1)! (n-i)!} \quad (i=1, 2, \dots, n) \quad (5)$$

$$\cdot \sum_{k=0}^{i-1} (-1)^k \binom{i-1}{k} \frac{1}{(n-i+k+1)^2} \quad (6)$$

$$\pi_i = \frac{1}{2} \left( i - \frac{1}{2} \right) \quad (7)$$

表2.  $Z(i, n)$  と  $-21n(1 - \pi_i)$  の比較

i	L9 n=4		L27 n=13	
	$Z(i, n)$ $-21n(1 - \pi_i)$	$Z(i, n)$ $-21n(1 - \pi_i)$	$Z(i, n)$ $-21n(1 - \pi_i)$	$Z(i, n)$ $-21n(1 - \pi_i)$
1	0.500	0.267	0.154	0.078
2	1.166	0.940	0.321	0.245
3	2.166	1.962	0.502	0.427
4	4.166	4.152	0.702	0.627
5			0.925	0.850
6			1.175	1.100
7			1.460	1.386
8			1.794	1.720
9			2.194	2.122
10			2.694	2.624
11			3.360	3.297
12			4.360	4.319
13			6.360	6.516

今、Danielと同様の考えの下に a 番目の順序統計量を基準にとり、その a は⑦式によって定めると良いであろう。即ち、 $\frac{1}{2} \int_0^x e^{-x} dx = 1 - 1/e = 0.632$  ⑧

$Pa = \Pr \{ \chi^2_{(2)} \leq 2 \}$   $\chi^2_{(2)}$  の定め方は⑦  
 $Pa = \frac{1}{2} \int_0^2 e^{-x/2} dx = 1 - 1/e = 0.632$  ⑧  
 故に、 $F[Z(i, n)] \approx Pa$  となるような  $i$  を a と定めると良い。⑦式において、

$\chi^2_{(2)} \leq 2$  とするのは、Half-Normal Plot の場合の値、即ち

$Pa' = \Pr \{ \text{Half-Normal ID} \leq \sigma \} = 0.683$  と概略等しくするためである。ここに、 $V(x) = 2^2$  であり、かつ  $E(x) = 2$  である。

## 2. 2 検定線の定め方

前記の如く、 $X(1) \leq X(2) \leq X(3) \leq \dots \leq X(a) \leq \dots \leq X(n)$  と大ききの順に並べ換えた時、 $X(a)$  と  $X(i)$  ( $i = h, \dots, n$ 、但し  $h > a$ ) の joint p.d.  $f$  は次式で与えられる。

$$f(xa, xi) = \frac{n!}{(a-1)! (i-a-1)! (n-i)!} \cdot \{ F(xa) \}^{a-1} \cdot \{ F(xi) - F(xa) \}^{i-a-1} \cdot \{ 1 - F(xi) \}^{n-i} \cdot f(xa) \cdot f(xi) \quad (9)$$

従って、 $ti = X(i, n) / X(a, n)$  に対し、 $ti > ti(\alpha)$  ならば有意と判定する時の  $\alpha\%$  の有意水準の値  $ti(\alpha)$  は⑩式で与えられる。

$$\alpha = \frac{n!}{(a-1)! (i-a-1)! (n-i)!} \cdot \sum_{k=0}^{a-1} \sum_{\ell=0}^{i-a-1} \binom{k+\ell}{k} \binom{a-1}{\ell} \binom{i-a-1}{\ell} \cdot (n-i+\ell+1)^{-1} \cdot \{ (i-a+k-\ell) + (n-i+\ell+1) ti(\alpha) \}^{-1} \quad (10)$$

L27 の場合についての数値計算の結果を表3と図1に示した。なお、Half-normal plot の場合にはこのように旨く行かない。従って、Danielはシミュレーションによってかかる有意水準の値を求めている。

表3. Guardrail for  $n=13$  and  $a=8$

h	$\alpha$			
	0.05	0.10	0.20	0.40
13	7.822	6.424	5.108	3.827
12	4.840	4.063	3.326	2.599
11	3.477	2.933	2.446	1.968
10	2.531	2.184	1.857	1.542
9	1.819	1.600	1.400	1.217

## 3. カイ2乗プロットの手順

### 3. 1 手順

3水準直交配列表に依って実施した実験結果は、次の手順によって解析する。

(ステップ1) 順序統計量を求める。

- ① 各平方和を大きさの順に並べて、小さい方から順に、 $X(1), X(2), \dots, X(n)$  とする。②  $Z(a, n) = 2$  となるような整数  $a$  を求める。例を表4に示した。

表4. $a$ の値	
実験計画	L27
平方和の数	13
$a$ の値	8

(ステップ2) 用意したカイ2乗プロット用に、 $t(i) (i = 1, 2, \dots, n)$  をプロットする。

(ステップ3) 適当な有意水準で検定する。

- ① 各  $t(i) (i = h, \dots, n, \text{但し } h > a)$  を検定する。

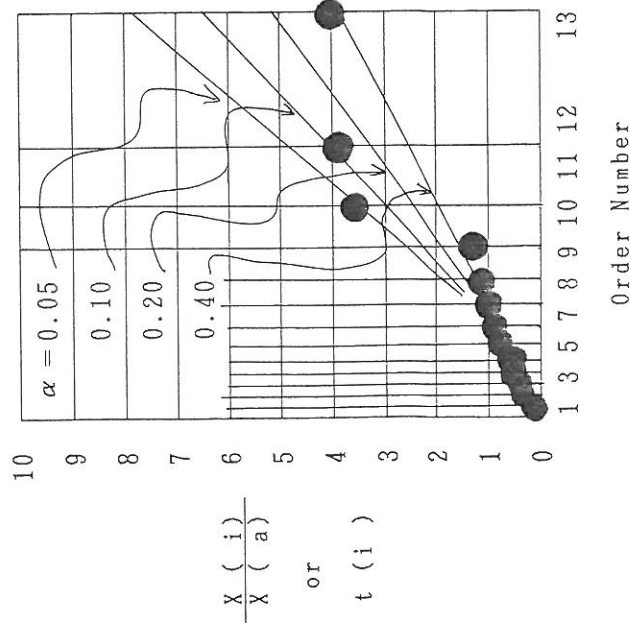
- ②  $t(h)$  より小さい  $t(i)$  については検定しない。

(ステップ4) 誤差分散の推定値を求める。

- ①  $Z(i, n) = 2$  の時の  $X(i)$  の値、即ち  $X(a)$  の値の  $\frac{1}{2}$  を第一近似の誤差分散の推定値とする。

図1. カイ2乗プロット用紙

図2. カイ2乗プロット



### 3. 2 数値例

実例により、カイ2乗プロットの使用法を説明する。L27型直交配列表を用いて、A, B, C, D, F, G, H, I, J, A×B, B×C の11要因を割付けて実験した。従来の分散分析法では、本来

解析不可能である。理論的に問題が残るが、小さい平方和を誤差と考えることにすれば、一応分散分析はでき、表5に示した通り、B, C, B×C が5%有意と判定されることになる。また、有意でない因子をプールした結果、 $\sigma^2$ の推定値は、157 となった。

表5. 分散分析表

S.V.	S.S.	d.f.	m.s.	F-test
A	493.3	2	246.4	1.79
B	1900.4	2	950.2	6.90*
A×B	432.5	4	108.1	0.78
C	1958.7	2	979.3	7.12*
B×C	2326.5	4	581.6	4.22*
D	523.1	2	261.5	1.90
ERROR	1375.4	10	137.5	
TOTAL	9010.2	26		

次にこれをカイ2乗プロットを用いて解析するため、まず平方和を大きさの順に並べる。

表6. 平方和の並べ換え

i	X(i)	t(i)	要因
1	37	0.075	G
2	210	0.426	A×B
3	223	0.452	A×B
4	269	0.546	F
5	281	0.570	H
6	343	0.696	J
7	446	0.905	I
8	493	1.000	A
9	523	1.061	D
10	573	1.162	B×C
11	1753	3.556	B×C
12	1900	3.854	B
13	1959	3.974	C

次に表4 から、 $Z(i, n) = 2$  となる  $a = 8$  を求める。そして  $X(i) (i = 1, 2, \dots, n)$  を  $X(8)$  で割って  $t(i)$  を計算し、これをカイ2乗プロット用紙にプロットする。その結果図2 が得られ、 $t(13), t(12), t(11)$  が有意となることが分かる。この時の  $\sigma^2$  の推定値は第一近似として、 $\frac{1}{2} \cdot 493$  即ち 247 と求められる。表5 の分散分析をこの結果と比較すると、誤差分散の推定値の値が若干小さく、有意と判定した因子は略同じである。

### 4. シミュレーション結果

日立製作所製 HITAC-20 型ミニコンピュータを用いて、必要な指数乱数を発生させ、帰無

仮説の下における有意水準 $\alpha$ と対立仮説の下における検出力 $1 - \beta$ とを、L27の場合についてシミュレーションにより検討した。

(1)帰無仮説の下における有意水準 $\alpha$ の検討  
50,000回のシミュレーションを実施し、表7に示す如くほぼ妥当な結果が得られた。

表7.有意水準の検討

$\alpha$	有意と判定した回数	観測した $\alpha$
0.05	2,505	0.050
0.10	5,099	0.102
0.20	10,219	0.204
0.40	20,265	0.405

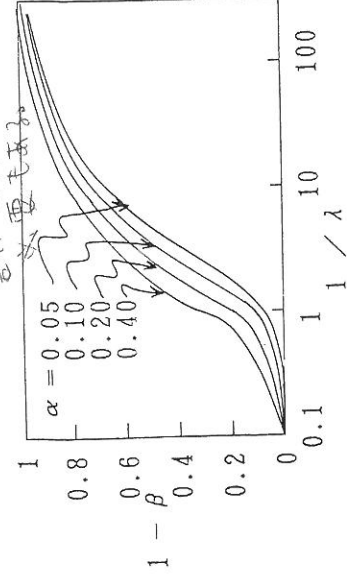
(2)対立仮説の下における検出力 $1 - \beta$ の検討  
対立仮説として1個だけ有意な平方和がある場合について、(1)と同様に10,000回のシミュレーションを実施し、表8に示すような結果が得られた。但し、検出力 $1 - \beta$ の検討には本来非心カイ2乗分布で計算すべきであるが、簡単な為此ここでは指数分布で近似した。

表8.注目したものが正しく有意と判定された回数

有意水準 $\alpha$				
$1/\lambda$	0.05	0.10	0.20	0.40
0.1	4	6	19	49
0.5	173	363	805	1,542
1	635	1,078	1,877	3,177
2	1,951	2,677	3,684	5,104
10	6,933	7,462	8,016	8,632
100	9,601	9,686	9,770	9,857

但し、 $\lambda$ は指数乱数における母数であり、 $\lambda = 2$ が帰無仮説に相当する。これを、検出力曲線にして図3に示す。

図3.検出力曲線



通常の分散分析法では、解析に問題がある為、比較はできないが恐らく検出力はそれほど高くないであろう。なお、帰無仮説の

下における有意水準 $\alpha$ は検出力 $1 - \beta$ とは一致しない。この理由は、本方法においてはその平方和を検定するのではなく、一部の平方和を検定するに過ぎないためであり、L27の場合は5/13の確率でしか検定される機会がないことは明白であろう。

## 5. 一般の分散分析法との比較

一般に用いられている分散分析法と比較すると、以下の特徴がある。

(利点) (1)図的解法であるから、簡便である。(2)異常値の発見が容易である。(3)誤差分散の推定値がなくとも分散分析できる。

(欠点) (1)実験の大きさと比較して有意な因子の数が多くと適用できない。(2)総の実験計画に適用出来るものではない。(3)他の方法と比較ができないが、若干検出力は低いであろう。

## 6. 結言

以上述べたように、カイ2乗プロットの有用性には見るべきものがあると思われる。

今後は、(1)検出力を厳密に把握する、(2)検定順序を考慮する、(3)有意なものを除外して残ったもので再度カイ2乗プロットしてさらに解析する、等について検討してゆく必要があると思われる。

本研究を行うに当たり、神戸商大の田坂教授を始め日科技連 DE0部会の諸氏ならびに日科技連大阪事務所の田中氏より有益な御助言を得た。ここに深甚なる謝意を表する次第である。

## 7. 参考文献

- Daniel, C., "Use of half-normal plot in interpreting factorial two-level experiments", Technometrics, 1, 4, 311 (1959)
- Birnbaum, A., "On the analysis of factorial experiments without replication", Technometrics, 1, 4, 343 (1959)
- 吉村 功、"正規確率紙による 2 型実験の解析"、品質管理、12, April, 273 (1961)
- Zahn, D. A., "Modification of and revised critical values for the half-normal plot", Technometrics, 17, 2, 189 (1975)