

2つの問題からなるテストの正答率に関する解析について

藤井良宜

Analysis of the Rates of Correct Answers for Two Problems

Yoshinori Fujii

ABSTRACT

When we compare teaching methods or instructional materials or investigate the understanding process and process of problem solving, we give students a problem in several condition and compare the rates of correct answers. In this case the rates of correct answers are considered to be influenced by the difficulty of the problem and the heterogeneity of the scholarship. Such a influence is solved to some extent by testing several problems. In this paper we concentrate the interest to the test which is consist of two problems in several conditions respectively. We construct the model with association of two problems and propose to estimate the parameter by maximam likelihood method useing EM algorithm. The data in Soeda and Fujii (1992) is analyzed. To investigate the influence of the association of two problems, we compare the proposed method with the method that we analyze the problems distinctively.

1. はじめに

新しい教授法や教具のよさを比較したり、生徒の理解や問題解決の過程を調べる場合には、異なる条件のもとで簡単なテストを実施し、その正答率を調べる方法がよく用いられる。このようなテストでの正答率は、問題の困難度や学力の個人差によって大きな影響を受けることが考えられるが、これらの影響をうまく制御して解析を行うことは困難である。しかし、この問題点は、単一の問題ではなく複数の問題を同時に実施することである程度解消できるであろう。例えば、添田、藤井(1992)¹⁾は、図形問題での添付された図による影響を調べる調査として、2つの文章題においてそれぞれ3通りの図を添付してその影響を調べている。複数の項目の調

査を行うことの利点は、困難度の異なる問題を同時に実施することで困難度の差の影響を調べることができる点と、各個人に対して複数の問題を実施することで学力の個人差を問題間の相関として考慮に入れることができる点にある。

この論文では、2つの問題に対し、それぞれ複数の条件設定をする場合の正答率の解析に絞って考えてゆくことにする。まず、問題間の相関を考慮したモデルを構成し、EMアルゴリズムを用いて各条件の影響を示すパラメータを推定する方法を提案する。また、その方法を添田、藤井（1992）のデータを用いて例証する。さらに、問題間の相関を考慮した場合の影響を見るために、提案した方法と各問題をそれぞれ別個に解析する方法との比較を行う。

2. モデル

次のような調査を考える。2つの問題に対して、それぞれ k_1, k_2 個の異なる条件をランダムに付け、 $k_1 k_2$ 通りのテストを準備する。このテストを n 人の生徒にランダムに配布し解答してもらおう。このとき、1問目が i 番目の条件で2問目が j 番目の条件（今後、条件 (i, j) と表す）の問題を与えられた生徒の結果を表1のようにまとめる。このようなデータは、問題設定で層別した分割表データとして取り扱うことができ、各条件のもとでデータは表2のセル確率をもつ多項分布にしたがうと仮定する。

		問題 1	
		正答	誤答
問題 2	正答	a_{ij}	c_{ij}
	誤答	b_{ij}	d_{ij}

— 表 1 —

		問 題 1	
		正 答	誤 答
問題 2	正答	$P_{11:i,j}$	$P_{12:i,j}$
	誤答	$P_{21:i,j}$	$P_{22:i,j}$

— 表 2 —

分割表解析でよく用いられる対数線形モデルを用いて、各セル確率を

$$P_{11:i,j} = \frac{\phi_{ij} \alpha_{ij} \beta_{ij}}{1 + \alpha_{ij} + \beta_{ij} + \phi_{ij} \alpha_{ij} \beta_{ij}}, \quad P_{12:i,j} = \frac{\beta_{ij}}{1 + \alpha_{ij} + \beta_{ij} + \phi_{ij} \alpha_{ij} \beta_{ij}}$$

$$P_{21:i,j} = \frac{\alpha_{ij}}{1 + \alpha_{ij} + \beta_{ij} + \phi_{ij} \alpha_{ij} \beta_{ij}}, \quad P_{22:i,j} = \frac{1}{1 + \alpha_{ij} + \beta_{ij} + \phi_{ij} \alpha_{ij} \beta_{ij}}$$

と表現することで、解析を容易にすることができる。このモデルは、一つの項目の正答率に関するロジスティックモデル（芝・渡辺（1988）²⁾、p155）の拡張でもある。ここで、 α_{ij} 、 β_{ij} はそれぞれ条件 (i, j) での問題1および問題2の正解率 / 誤答率に対応し、 ϕ_{ij} は問題1と問題2の間に関連性（dependency）に対応する指標である。2問題間に関連性は、学力の個人差によって生じると考えられる。各生徒は2つの問題を別個に解答するが、学力の高い生徒は両方の問題を正解するであろうし、学力の低い生徒は両方とも不正解となる傾向がある。

実際のデータは学力の異なる生徒を対象としているため、結果的に2問題間に関連が起こるのであろう。そのため、2問題間の関連性は学力の個人差の度合をある程度表しているとも言える。

一方、ここでの設定では、問題2の条件は問題1での正答率には直接的には余り影響しないと考えられる。そこで、 α_{ij} はjに関係しないと仮定し、単に α_i と表す。また、 β_{ij} もiに依存しないと仮定し、単に β_j と表すことにする。さらに、2問題間の関連性は問題の条件が変わっても余り変化しない場合が多く、一定であると仮定する方が解析結果を総合的に解釈できるので、 ϕ_{ij} は設定によらず一定であると仮定し、単に ϕ で表すことにする。これらの仮定については根拠が乏しく、このモデルを用いて解析を行う場合には、検定などを用いてモデルの適合度を何らかの形でチェックすることが必要となる。

3. パラメータの推定

3. 1 最尤推定法

最尤法を用いて α_i 、 β_j 、 ϕ の推定を考えよう。2節のモデルから対数尤度Lは

$$L = \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} b_{ij} \log \alpha_i + c_{ij} \log \beta_j + a_{ij} (\log \alpha_i + \log \beta_j + \log \phi) \\ + n_{ij} \log(1 + \alpha_i + \beta_j + \phi \alpha_i \beta_j) \quad + \text{定数}$$

となり、これを各パラメータで偏微分すると

$$\frac{\partial L}{\partial \alpha_i} = \frac{1}{\alpha_i} \sum_{j=1}^{k_2} (a_{ij} + b_{ij}) - \sum_{j=1}^{k_2} \frac{n_{ij}(1 + \phi \beta_j)}{1 + \alpha_i + \beta_j + \phi \alpha_i \beta_j} \quad (i=1, 2, \dots, k_1)$$

$$\frac{\partial L}{\partial \beta_j} = \frac{1}{\beta_j} \sum_{i=1}^{k_1} (a_{ij} + c_{ij}) - \sum_{i=1}^{k_1} \frac{n_{ij}(1 + \phi \alpha_i)}{1 + \alpha_i + \beta_j + \phi \alpha_i \beta_j} \quad (j=1, 2, \dots, k_2)$$

$$\frac{\partial L}{\partial \phi} = \frac{1}{\phi} \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} a_{ij} - \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} \frac{n_{ij} \phi \alpha_i \beta_j}{1 + \alpha_i + \beta_j + \phi \alpha_i \beta_j}$$

となる。しかし、第2項の形はロジスティックモデルと似た形をしているため、これらの式の値をすべて0にするようなパラメータを数値的に求めることは、それほど容易ではないことはよく知られている。実際、ニュートン・ラフソン法を用いても、かなりうまく初期値を設定しなければ収束しない。そこで、Dempster et al. (1977)³¹⁾で提案されたEMアルゴリズムを用いた反復法を考える。この方法はニュートン・ラフソン法ほど初期値の選択に依存しないうえ、計算プログラムも容易に作成できる点ですぐれている。

まず、適当な初期推定値 α_i^0 、 β_j^0 、 ϕ^0 を設定し、その推定値に対して、次の推定値 α_i^1 、 β_j^1 、 ϕ^1 を次のように構成する。

$$\alpha_i^1 = \frac{\sum_j (a_{ij} + b_{ij})}{\sum_j \frac{n_{ij}(1 + \phi^0 \beta_j^0)}{1 + \alpha_i^0 + \beta_j^0 + \phi^0 \alpha_i^0 \beta_j^0}}, \quad \beta_j^1 = \frac{\sum_i (a_{ij} + c_{ij})}{\sum_i \frac{n_{ij}(1 + \phi^0 \alpha_i^0)}{1 + \alpha_i^0 + \beta_j^0 + \phi^0 \alpha_i^0 \beta_j^0}}$$

$$\phi^1 = \frac{\sum_{ij} a_{ij}}{\sum_{ij} \frac{n_{ij} \phi^0 \alpha_i^0 \beta_j^0}{1 + \alpha_i^0 + \beta_j^0 + \phi^0 \alpha_i^0 \beta_j^0}}$$

さらに、この操作を推定値が収束するまで繰り返すことで最尤推定量を求めることができる。

3. 2 初期推定値

ニュートン・ラフソン法やEMアルゴリズムを用いた場合には、初期推定値をどう選ぶかによってかなり収束の速さや収束する可能性が変化する。そのため、簡単に求められて、さらにある程度パラメータに近い初期推定量を見つけることが重要になる。そこで、推定関数の考え方をを用いて、次のような初期推定量を提案する。

$$\alpha_i^0 = \sum_j b_{ij} / \sum_j d_{ij}, \quad \beta_j^0 = \sum_i c_{ij} / \sum_i d_{ij},$$

$$\phi^0 = \sum_{ij} \frac{a_{ij} d_{ij}}{n_{ij}} / \sum_{ij} \frac{b_{ij} c_{ij}}{n_{ij}}.$$

ϕ^0 は層別された分割表における共通オッズ比の推定によく用いられるマンテル・ヘンセル推定量で、その一致性はGart (1962)⁴⁾ですでに示されている。また、次の定理によって α_i^0 、 β_j^0 の一致性も証明することができる。

定理 1

α_i^0 、 β_j^0 は一致推定量である。

(証明)

α_i^0 は次の推定関数 $g(\alpha_i)$ を 0 とする α_i を求めることによって得られる。

$$g_i(\alpha_i) = \sum_{j=1}^{k_2} (d_{ij} \alpha_i - b_{ij})$$

この推定関数は不偏で α_i に関して単調増加であるから α_i^0 は一致推定量となる。 β_j^0 に関しても同様に証明できる。(証明終)

定理 1 より、この初期推定量は、簡単な推定量であるが、データ数を多くするとパラメータに収束することが分かる。しかし、 α_i や β_j の推定においては a_{ij} の情報を余り利用してないためそれほど効率の高い推定量ではない。

4. 解析例

添田・藤井 (1992) のデータを用いて、実際に解析してみよう。このデータは図形問題における添付された図の影響を調べるために、大学生 119 名に 2 つの図形問題を解いてもらった結果である。各問題に対してそれぞれ 3 通りの図を用意し、その組み合わせで 9 通りの問題を大学生にランダムに配っている。実際のデータを表 3 に示す。ここでは、簡単のため問題 1 では一般四角形、平行四辺形、長方形をそれぞれ条件 1、2、3 とし、問題 2 では一般四角形、台

形、平行四辺形をそれぞれ条件1、2、3と表すことにする。

まず、初期推定量を求めてみると $(\alpha_1^0, \alpha_2^0, \alpha_3^0) = (1.375, 0.857, 0.500)$ 、 $(\beta_1^0, \beta_2^0, \beta_3^0) = (0.500, 2.143, 0.118)$ 、 $\phi^0 = 2.436$ となる。さらに、この初期推定量を用いてEMアルゴリズムを用いて最尤推定量を求めると $(\alpha_1, \alpha_2, \alpha_3) = (1.255, 0.957, 0.504)$ 、 $(\beta_1, \beta_2, \beta_3) = (0.530, 2.098, 0.205)$ 、 $\phi = 1.565$ となる。この結果から問題2については適切な図を添付することによってかなり正答率が増加し、問題1については一般の四角形を用いた方が正答率が増加することが分かる。一方、初期推定値を見ても、最尤推定量と同じような傾向を示していることが分かるが、最尤推定量の方が初期推定量に比べて保守的(1に近い推定値)となっている。これは、初期推定量が他の問題で不正解であった者の情報だけを用いているために起こるのではないだろうか。また、問題間の相関を考慮せずに各問題で別個に解析をしても同じような傾向を示している。この点については5節でもう少し詳しく述べることにする。

問題1 問題2		一般四角形		平行四辺形		長方形		合計	
		正答	誤答	正答	誤答	正答	誤答		
一般四角形	正答	4	2	4	1	2	3	16	40
	誤答	5	2	6	2	1	8	24	
台形	正答	6	3	4	3	4	9	29	40
	誤答	2	2	2	4	0	1	11	
平行四辺形	正答	2	2	1	0	3	0	8	39
	誤答	4	4	4	8	6	5	31	
合計		23	15	21	18	16	26	119	
		38		39		42			

-表3-

また、2節のモデルではパラメータに対してかなり制限を加えているので、パラメータの値によって解釈するには、モデルの適合度を検定することが必要となる。そこで、最尤推定量を用いて、ピアソントイプのカイ2乗統計量を構成し、モデルの適合度を調べる。カイ2乗統計量は25.25となり、有意水準5%で棄却できないため適合度はそれほど悪くはないことが分かる。よって、このモデルによる解析は意味のあるものであるといえる。

5. 問題間の関連性を考慮しない場合の解析

複数の問題を同時に生徒に解かせた場合に、問題間に相関がみられるのが普通である。各生徒の学力は複数の問題に共通に反映するため、ある問題が正答であった生徒と誤答であった生徒では他の問題での正答率に違いがあるのは当然と考えられる。しかし、実際の解析において

は問題間の相関を考慮せず問題ごとに解析が行われる場合が多い。これは、問題数が多くなると条件パターンが急激に増加し、それをうまく解析するにはかなり多くのデータ数が必要となるためである。そのため、各問題ごとの解析のみが行われるのも、ある程度は仕方ない面もある。そこで、問題間の相関を無視して項目ごとに解析した場合の影響について明らかにしておくことは重要である。ここでは問題数が2つの場合にしばって考えていくことにする。ただし、簡単のため条件(i,j)での被調査者数はjに依存しないことを仮定しておく。この仮定は多くの調査では各設定の被調査者数はほぼ等しく取られているであろうから、それほど厳しい仮定ではない。

問題2の結果を考慮せず、問題1のみで解析をする場合にはデータは表4の形の $2 \times k_1$ の分割表としてみる事ができる。このとき、条件と正答率の関係を見る場合には、オッズ比とい

	条件 1	条件 2	...	条件 k_1
誤答	X_{11}	X_{12}	...	X_{1k_1}
正答	X_{21}	X_{22}	...	X_{2k_1}

— 表 4 —

う指標が用いられる。オッズ比は条件1を基準として各条件での正答率を比較する指標で、

$$\phi_i = \frac{E(X_{1i})E(X_{21})}{E(X_{11})E(X_{2i})}$$

と表す。条件iでの正答率が条件1での正答率よりも高ければ、 ϕ_i は1より大きくなり、条件1の正答率の方が高ければ、 ϕ_i は1より小さくなる。また、条件iと条件hを比較する場合には、 ϕ_h / ϕ_i を考えればよい。上のモデルのパラメータ α 、 β 、 ϕ と ϕ_i との関係を明らかにするために次の定理を示しておく。

定理2

$b_j \geq 1, \pi_j \geq 0$ ($i=1, 2, \dots, k_2$)、ただし $\sum \pi_j = 1$ とする。さらに、 $x > 0$ に対して $\Delta(x)$ を次の方程式を満足するようにとる。

$$\sum_{j=1}^{k_2} \frac{\pi_j}{1+xb_j} = \frac{1}{1+x\Delta(x)}$$

このとき、 $\Delta(x)$ はxの単調減少関数となる。特に、 $b_1=b_2=\dots=b_{k_2}$ のときのみ $\Delta(x)$ は定数関数となる。

(証明)

上の方程式を変形することで次式を得る。

$$\log \Delta(x) = -\log x + \log \left(1 - \sum_j \frac{\pi_j}{1+xb_j} \right) - \log \left(\sum_j \frac{\pi_j}{1+xb_j} \right)$$

両辺をxで微分すると

$$\frac{\partial}{\partial x} \log \Delta(x) = -\frac{1}{x} + \frac{\sum_j \frac{b_j \pi_j}{(1+xb_j)^2}}{1 - \sum_j \frac{\pi_j}{1+xb_j}} + \frac{\sum_j \frac{b_j \pi_j}{(1+xb_j)^2}}{\sum_j \frac{\pi_j}{1+xb_j}}$$

$$\begin{aligned}
 &= -\frac{1}{x} \left[\frac{\left(\sum_j \frac{\pi_j}{1+xb_j} \right) \left(1 - \sum_j \frac{\pi_j}{1+xb_j} \right) - x \sum_j \frac{b_j \pi_j}{(1+xb_j)^2}}{\left(\sum_j \frac{\pi_j}{1+xb_j} \right) \left(1 - \sum_j \frac{\pi_j}{1+xb_j} \right)} \right] \\
 &= -\frac{1}{x} \frac{\sum_j \left(\frac{1}{1+xb_j} - \sum_h \frac{\pi_h}{1+xb_h} \right)^2 \pi_j}{\left(\sum_j \frac{\pi_j}{1+xb_j} \right) \left(1 - \sum_j \frac{\pi_j}{1+xb_j} \right)} \leq 0
 \end{aligned}$$

等号成立は、分子が0となるときのときであるから、 b_j がすべて等しくなるときのときとなる。

(証明終)

つぎに、対数オッズ比 $\log \phi_i$ と α_i 、 α_i の間には次の不等式が成り立つことが示す。

定理3

$$0 \leq \log \phi_i \leq \log \alpha_i - \log \alpha_1 \quad (\alpha_1 \leq \alpha_i \text{ のとき})$$

$$0 \geq \log \phi_i \geq \log \alpha_i - \log \alpha_1 \quad (\alpha_1 \geq \alpha_i \text{ のとき})$$

(証明)

$$\begin{aligned}
 E(X_{1i}) &= \sum_{j=1}^{k_2} E(c_{ij} + d_{ij}) = \sum_{j=1}^{k_2} n_{ij} \left(\frac{1 + \beta_j}{1 + \alpha_i + \beta_j + \phi \alpha_i \beta_j} \right) \\
 &= \sum_{j=1}^{k_2} n_{ij} \left(\frac{1}{1 + \alpha_i \left(\frac{1 + \phi \beta_j}{1 + \beta_j} \right)} \right)
 \end{aligned}$$

ここで、定理2において

$$b_j = \frac{1 + \phi \beta_j}{1 + \beta_j}, \quad \pi_j = \frac{n_{ij}}{N}, \quad \text{ただし} \quad N = \sum_{j=1}^{k_2} n_{ij}$$

とすると、

$$E(X_{1i}) = N \left(\frac{1}{1 + \alpha_i \Delta(\alpha_i)} \right)$$

となる。同様にして、

$$E(X_{2i}) = N \left(\frac{1}{1 + \alpha_i \Delta(\alpha_i)} \right)$$

をえる。また、 $X_{1i2} + X_{2i} = X_{1i} + X_{2i} = N$ より

$$E(X_{2i}) = N \left(\frac{\alpha_i \Delta(\alpha_i)}{1 + \alpha_i \Delta(\alpha_i)} \right), \quad E(X_{21}) = N \left(\frac{\alpha_1 \Delta(\alpha_1)}{1 + \alpha_1 \Delta(\alpha_1)} \right)。$$

対数オッズ比 $\log \phi_i$ は

$$\log \phi_i = \log \alpha_i - \log \alpha_1 + \log \Delta(\alpha_i) - \log \Delta(\alpha_1)$$

定理2より Δ は単調減少関数であるから定理は成立する。

(証明終)

定理3より問題間の相関を考慮しない場合には、各設定の影響を過小評価する傾向があること

が分かる。また、 $\phi = 1$ の場合には b_i は i に依存しないため、 Δ は定数関数となり $\log \phi_i$ と $\log \alpha_i - \log \alpha_1$ は一致する。そのため、項目間に相関がない場合にはこの2つの解析は同じ結果を導くことが示される。

6. 考察

この論文では対数線形モデルを用いてモデル構成を行っているが、もう少し学力の個人差を考慮するならば、生存データの異質性モデル (Hougaard, 1986)⁵⁾ や比率に関するオーバードispersションモデル (Anderson, 1988)⁶⁾ と同様に異質性の確率変数 Z を導入して、 Z を与えたとき独立に各問題に正答するというモデルが考えられる。各問題の正答率にはロジスティックモデルを考えると、各セルの確率は

$$P_{11 \dots ij} = E \left(\frac{\alpha Z}{1 + \alpha Z} \frac{\beta Z}{1 + \beta Z} \right)$$

などと表せる。ここで、期待値は確率変数 Z に関するものである。しかし、このようなモデルを考える場合に、 Z の分布に何を用いるかによって結果が異なるうえ、期待値の部分を書き表すことができないため、解析が困難である。さらに、特定の分布 (例えば、正規分布) を用いた場合でも解析が複雑となるため余り実用的ではない。一方、2節で提案したモデルの場合にも相関のパラメータ ϕ を導入しているので、2元配置の分散分析の場合と同様に ϕ が大きい場合には α や β の解釈が難しくなるので注意が必要である。

5節で述べたように、相関を考慮せずに解析を行った場合には各設定の影響を過小評価する傾向が示される。しかし、問題間の相関 ϕ がそれほど大きくない場合にはその影響は余り大きくない。表3のデータの設定で、問題1の一般四角形を用いた場合と長方形の場合の比較における ϕ の影響を見てみよう。表3では、問題1で一般の四角形、問題2で平行四辺形を与えている被調査者が他と比べて1名少ないので、そこを他と同じ13名にとり、 α 、 β は4節で求めた最尤推定量とする。

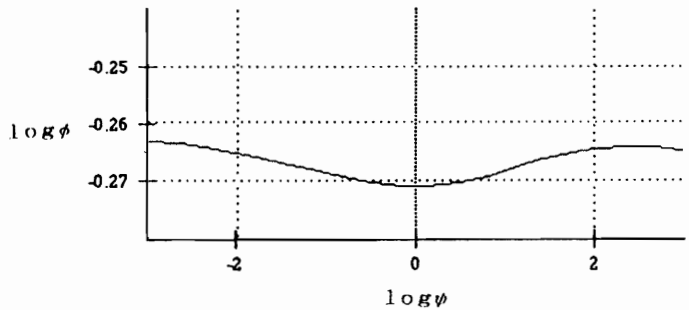


図 1

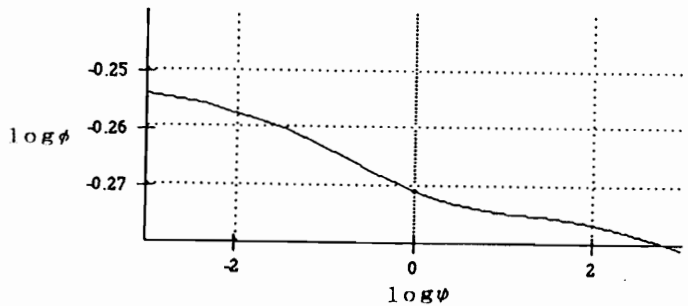


図 2

このとき、 ϕ を動かしたときの $\log \phi$ の変化を図1に表す。ただし、図1では ϕ も対数変換を

施している。この図では $\log \phi = 0$ のとき、 $\log \phi$ は最小になっており、定理3の結果と一致するが、その変化の幅はそれほど大きくはない。一方、条件ごとの被調査者数が j に依存する場合には定理3は成立しないため、逆に過大評価する場合も考えられる。実際、図1において被調査者数を表3の設定通りにしておいた場合の $\log \phi$ の変化を調べると、図2のようになる。ここでは、 $\log \phi$ はほとんど単調減少となって、 $\log \phi$ が負の場合には過大評価になることが分かるであろう。この点については、今後パラメータのバイアスばかりでなく推定量の分布の比較などを数値実験によって評価する必要があると思われる。

まとめとして、問題間の相関がそれほど大きくない場合には過小評価になる点を考慮しながら各問題ごとに解析をすればよいのではないか。また、実際に調査を行う場合には解析結果の解釈を容易にする点からも、問題ごとに複数の設定がある場合にはできるだけ均等に割り当てる方がよい。

参考文献

- 1) 添田、藤井 (1992) . 図的表記の思考への影響について — 大学生と中学生との比較を中心に — 宮崎大学教育学部紀要, 教育科学, 第72号, 65-73
- 2) 芝、渡辺 (1988) . 入試データの解析 新曜社
- 3) Gart, J.J. (1962) . On the combination of the relative risks. *Biometrics*, 18, 594-600.
- 4) Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977) . Maximum likelihood from incomplete data via the EM algorithm (with discussion). *J.R.Statist.Soc.*, B39, 1-38.
- 5) Hougaard, P. (1986) . Survival models for heterogeneous populations derived from stable distributions. *Biometrika*, 73, 387-396.
- 6) Anderson, D. A. (1988) . Some models for overdispersed binomial data. *Australian Journal of Statistics*, 30, 125-148.

(1992年9月30日受理)