

統計講習会テキスト附録

経営	47
測定	13

標本調査法および問題集

昭和31年12月

農林省林業試験場経営部

附 録 1 標 本 調 査 法

1. 標 本 調 査

社会調査、その他現存する集団の調査研究においては、対象集団 $\mathcal{U}$ は有限個の要素から構成されている。このような対象集団からとられる標本について反省してみよう。

いままでの議論では、対象集団 $\mathcal{U}$ からランダムに一要素を抽出する操作を考え、抽出された要素の数標識を表わす確率変数を X とし、これを母確率変数として母集団 $(\mathcal{U}, X)$ を設定した。そして、母確率変数 X と同じ確率分布をする n 個の互に独立な確率変数 $X_1, \dots, X_n$ の実現値であると考えることのできる n 個の値の組 $x_1, \dots, x_n$ を大きさ n の標本と定義したのである。そして、そのような標本を得るためには次のようにしなければならなかった。すなわち、対象集団 $\mathcal{U}$ からランダムに一要素を抽出し、その数標識を調べてこれを x_1 とする。次に、抽出された要素をもとの集団にもどして、再び始めと同じ状態の集団にしておいて、そこからランダムに一要素を抽出し、その数標識を調べてこれを x_2 とする。このようにして、要素を抽出して数標識を調べたならば、それをもとの集団にもどして、毎回同一対象集団に保持しながら、ランダムに一要素を抽出して数標識を調べる、という操作を n 回繰り返して n 個の数標識の値の組 $x_1, x_2, \dots, x_n$ を求めるのである。しかるに、この $x_1, x_2, \dots, x_n$ がわれわれの定義した意味での標本である。上のように、抽出された要素はもとにもどして次の抽出をするという抽出法を元に戻す抽出法drawing with replacement とよんでいる。

しかるに、実際問題として、元に戻す抽出はあまりにも煩雑である。そこでまず、対象集団からランダムに一要素を抽出し、その数標識を調べて x_1 を得たならば、次にはこの抽出された要素をもとにもどすことなく、残りの集団からランダムに一要素を抽出し、その数標識を調べて x_2 とする。というようにして進むことを考える。このようにして、一度抽出された要素をもとにもどすことなく、残りの集団からランダムに一要素を抽出するという操作を元に戻さない抽出法drawing without replacement という。すなわち、 n 回の元に戻さない抽出法で得られる要素の数標識 $x_1, x_2, \dots, x_n$ を求めて、これをもつて推論の足場とするならば、抽出操作が元に戻す抽出法よりもはるかに便利になる。

しかるに、元に戻さない抽出法で得られる n 個の数標識の組 $x_1, x_2, \dots, x_n$ が、われわれの定義した意味での標本となつていようか。

もし対象集団 $\mathcal{U}$ が無限集団であるならば、元に戻さない抽出法で得られた $x_1, x_2, \dots, x_n$ は、やはり標本であると考えることが許されるであろう。

しかるに、もし対象集団 $\mathcal{U}$ が有限集団であるならば、このような元に戻さない抽出法で得られた $x_1, x_2, \dots, x_n$ は、われわれの意味での標本とはなっていないのである。その理由は、第1回に抽出される要素の数標識を表わす確率変数を X_1 、第2回に抽出される要素の数標

識を表わす確率変数を X_1 とすれば、第1回に、どのような要素が抽出されるかに依存して、 X_1 の確率分布が定まることになり、 X_1 と X_2 とは独立になつていないからである。

したがつて、現存する対象集団が有限集団であるということ、そして実際問題としては、元に戻さない抽出をしなければならないという2つの制約を表面に押し出すとき、いままで論じた推論は、そのままの形では適用できないことになる。そこで上の2つの制約条件を考慮に入れた上で、対象集団の集団特性を推論するための理論をどのように構成するか、ということが問題になつてくる。こうした理論を述べるのが本章の標本調査論 theory of sample survey である。

さて、標本調査は次の諸段階をふんで実行される。

1) 調査目的の確立 まず、目的を明確に述べるのが大事である。そうでないと調査の細部にばかり心を奪われて、大局を見失うことが多い。

2) 標本を抽出すべき対象集団の確立 ある工場で、一日に生産された製品の中に、何ほどの不合格品があるか、というような場合には対象集団は明確である。しかし、例えば、農家調査をする、というような場合には、何をもちて農家と定義するか、また、その如何なる集団をもつて対象集団とするか、を一点の疑いもなく、明確に述べることは相当困難な問題である。それにしても、調査の出発にあつては、対象集団を明確に述べておかなければならない。

3) 調査すべき数標識の決定、われわれは対象集団に関する、ある統計的意識を求めているのであるから、対象集団の各要素について、如何なる数標識を注目するかを、調査の目的に応じて決定しなければならない。

4) 測定法の確立 調査すべき数標識が定まつたならば、それを如何にして測定するかを決定しなければならない。例えば、人口の年齢構成を知りたいという場合に、ある人の年齢を如何にして定めるか。調査日における満幾歳幾ヶ月までも測るものか、それとも調査には最も近い誕生日の年齢をとるものかを明かにしておかなければならない。さらに、記入申告させるものか、それとも調査員が面接して記入するものかも決定しておかなければならない。

5) 抽出単位の選定 標本抽出の準備として、対象集団を抽出単位とよばれる単位に分割しておかなければならない。それは標本として抽出されるべき単位であり、その総和は、対象集団を尽くし、かつ互に重なり合つてはいけぬ。すなわち、対象集団を構成する各要素は、一つそして唯一つの抽出単位に属しなければならない。例えば、電球の一山が対象集団であるときには、その個々の電球自身が抽出単位にとられるであろう。しかし、ある都市の人口調査において、各人が対象集団の構成要素であり、それが抽出単位にとられるときもあるであろうし、あるいは個人の集団である各世帯が抽出単位にとられることもあるであろう。そのいずれを抽出単位にとるかあらかじめ決定しておかなければならない。

抽出単位の完全なリストを枠 frame という。こうした枠を構成するという仕事は、實際上、一つの大事な問題になる。

6) 標本の選定 標本の抽出の仕方については種々の方法が考えられ、節を追つて説明するが、それらの抽出法のうちいずれを選ぶかを定めなければならない。この問題の中には、標本の大きさを決定することも含まれるであろう。与えられた予算と時間内で調査が実行できるように標本を選定しなければならない。

7) 現地作業の組織 大きい調査の遂行のためには、事務組織の問題が重要になつてくる。所定の測定法に従つて、データが円滑に蒐集されるように作業体系を組織化しなければならない。

8) データの蒐集と分析 どのような質問によつて調査するかということも慎重に考えなければならない。そして所定の方式に従つて、実際にデータを集め、それを所定の方法によつて分析し、推定するのである。

9) 将来の標本調査のために役立つ知識の獲得 対象集団に関する情報を、あらかじめ多くもてばもつほど、よりよい標本調査法が計画できるであろう。現実の標本調査から得られる対象集団に関する知識、および調査運営上の経験は、将来の調査において有効に活用しようように配慮すべきである。

以上、標本調査の全過程の、いずれの段階に欠陥があつても、調査は全体として無意味なものになつてしまうであろう。しかし、その全段階については、ここに述べることは出来ないもので、8)の段階について、最低の費用で所期の精度をもつ推定をなし得ることを目標として、標本を得るための方法ならびにその推定の方法について説明することにする。かくして、今後本書では、標本調査法というときには、標本を得るための抽出法と、得られた標本から対象集団の集団的特性を推定するための方法の2つを内容にもつことになる。

ここで、標本調査法の精度について説明しておこう。同一対象集団に対して、何回も何回もその標本調査法を適用して推定をしたとすれば、その推定値は、一定の確率法則に従つて、確率分布をするであろう。しかも実際にとられるような標本の大きさにおいては、中心極限定理により、推定値が近似的に正規分布をする、とみなし得る場合が大部分である。したがつて、推定値の標本分布の分散が小さいほど、より精度の高い標本調査法ということが出来るであろう。そこで、推定値の標本分布の分散の逆数をもつて、その標本調査法の、あるいは抽出法を一定にしておく場合であれば推定法の精度 precision と定義するのである。

したがつて、標本調査論においては、各種の方法で得られる推定値の分散を表わす公式を求めることが大切な問題になつてくる。それを通して、各種の標本調査法の優劣を判断していくわけである。

2. 単純ランダム抽出法

対象集団を構成する要素の数を N とし、その数標識を、それぞれ、 $y_1, y_2, \dots, y_N$ で表わす。この集団からとられる $n (\leq N)$ 個の要素の組は全部で

$${}^N C_n = \frac{N!}{n!(N-n)!}$$

たす。そしてこれらの各組の大きさ n の標本とよぶことにし、標本を構成する要素の数標識を一般的に、 $y_1, \dots, y_n$ で表わすことにする。これは必ずしも対象集団の始めの n 個の要素の数標識という意味ではないことを注意しておく。以下においても同様である。

対象集団 π から、大きさ n の標本がすべて等しい確率、すなわち、標本 $(y_1, \dots, y_n)$ が確率 $\frac{1}{{}^N C_n}$ で抽出されるような仕方抽出することを、単純ランダム抽出法 simple random sampling という。実際に単純ランダム抽出を行うためには次のようにすればよいであろう。まず、対象集団の各要素に 1 番から N 番までの番号を付けておく。別に 1 番から N 番までの番号を書いた N 本のくじを用意し、そこからランダムに 1 本抽出して、その番号を読む。次にそのくじを元に戻すことなく、残りの $(N-1)$ 本のくじの中からランダムに 1 本抽出してその番号を読む。次にそのくじも元に戻さずに残りの $(N-2)$ 本のくじの中からランダムに 1 本を抽出する、というようにして、合計 n 本のくじを引く。そして引かれた n 本のくじの番号をもつ対象集団の要素についてその数標識 $y_1, \dots, y_n$ を読めばよいであろう。

(さらに実際的には、乱数表という数字の配列が出来ていて、そこから、ランダムに数字をとれば、上の、くじ引きと同等の抽出が出来るとなっている)

この抽出法においては、抽出のいずれの段階においても、まだ抽出されていない要素は、すべて同一の抽出される確率をもつことになる。そして、単純ランダム抽出法によつて得られた標本 $(y_1, \dots, y_n)$ は、対象集団の要素の数標識 $y_1, \dots, y_N$ から定まるところの一定の確率分布をもつた離散的 n 次元確率変数 $(y_1, \dots, y_n)$ の実現値であると解釈される。

われわれは、対象集団 π と、単純ランダム抽出法を通して定まるこの確率変数 $(y_1, \dots, y_n)$ とを組にして $(\pi, (y_1, \dots, y_n))$ を母集団、 $(y_1, \dots, y_n)$ を単純ランダム抽出における母確率変数とよぶことにする。

したがって、母確率変数の実現値 $(y_1, \dots, y_n)$ が単純ランダム抽出法における大きさ n の標本である。 n/N を抽出比 sampling ratio, その逆数 N/n を拡大乗数 expansion factor という。

われわれの目的は、この標本 $(y_1, \dots, y_n)$ を通して、対象集団 π のすべての要素の数標識の

$$\text{総和(母総和)}: Y = y_1 + y_2 + \dots + y_N,$$

$$\text{平均値(母平均)}: \bar{Y} = \frac{1}{N}(y_1 + y_2 + \dots + y_N) = \frac{Y}{N}$$

を推定することである。

標本 $(y_1, \dots, y_n)$ については次の記号を用いる:

$$\text{標本和}: y = y_1 + \dots + y_n,$$

$$\text{標本平均}: \bar{y} = \frac{1}{n}(y_1 + \dots + y_n) = \frac{y}{n}$$

このとき、次の定理が成り立つことを証明しよう。

定理 1. 標本平均 $\bar{y}$ は、母平均 $\bar{Y}$ の不偏推定量である。この意味で、母平均 $\bar{Y}$ の推定量として、標本平均 $\bar{y}$ をとることにする。

証明 標本平均 $\bar{y}$ が、どのような確率分布をするかを調べてみよう。 $\bar{y}$ は、各標本毎に、 $\bar{y} = \frac{1}{n}(y_1 + \dots + y_n)$ という値をとるのであるが、その値をとる確率は、標本 $(y_1, \dots, y_n)$ が得られる確率 $\frac{1}{{}^N C_n}$ に等しい。よつて

(11.1)

$$\Sigma\{\bar{y}\} = \Sigma\left(\bar{y} \times \frac{1}{{}^N C_n}\right) = \frac{n!(N-n)!}{n \times N!} \Sigma(y_1 + \dots + y_n),$$

ここに Σ は、 ${}^N C_n$ 個の標本のすべてにわたつて加えることを示す。

この和を求めるために、任意の一つの値 y_i が、何個の標本に含まれているかを調べてみよう。 y_i を含む標本を構成するために用い得る要素は、 y_i を除いた残りの $(N-1)$ 個である。そこから $(n-1)$ 個の要素をとつて、 y_i と併せた n 個の要素が y_i を含む標本のすべてであり、したがつてその個数は

(11.2)

$${}^{N-1} C_{n-1} = \frac{(N-1)!}{(n-1)!(N-n)!} \quad (i=1, \dots, N)$$

である。したがつて

(11.3)

$$\Sigma(y_1 + \dots + y_n) = \frac{(N-1)!}{(n-1)!(N-n)!} (y_1 + y_2 + \dots + y_N)$$

よつて (11.1), (11.3) から

(11.4)

$$\begin{aligned} \Sigma\{\bar{y}\} &= \frac{(N-1)!}{(n-1)!(N-n)!} \cdot \frac{n!(N-n)!}{n \times N!} (y_1 + y_2 + \dots + y_N) \\ &= \frac{1}{N} (y_1 + y_2 + \dots + y_N) = \bar{Y} \end{aligned}$$

よつて証明された。(終)

系 1. $\hat{Y} = N\bar{y}$ は、母総和 Y の不偏推定量である。

証明 期待値をとる演算がもつ性質から明かなように

$$(11.5) \quad E\{\hat{Y}\} = E\{Ny\} = Ne\{\bar{y}\} = N\bar{Y} = Y. \quad (\text{終})$$

次に、これら推定量の分散(その正の平方根を推定の標準誤差 standard error of estimate という)を求めてみよう。

対象集団 π を構成する要素の数標識 $y_1, y_2, \dots, y_N$ の分散 σ^2 は

$$(11.6) \quad \sigma^2 = \frac{1}{N} \sum_{i=1}^N (y_i - \bar{Y})^2$$

で定義するのが普通であるが、今後計算の便宜上、次式で与えられる S^2 をもつて(有限集団の)母分散と定義することにする:

$$(11.7) \quad S^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{Y})^2$$

定理 11.2 単純ランダム抽出における標本平均 $\bar{y}$ の分散 $V(\bar{y})$ は、次式で与えられる:

$$(11.8) \quad V(\bar{y}) = E\{(\bar{y} - \bar{Y})^2\} = \frac{s^2}{n} \cdot \frac{N-n}{N}.$$

証明 11.3 を導いたと同じ考えで、次の二つの式の成立が証明される:

$$E\{(y_1 - \bar{Y})^2 + \dots + (y_n - \bar{Y})^2\} = \frac{n}{N} \{(y_1 - \bar{Y})^2 + \dots + (y_N - \bar{Y})^2\} \quad (9)$$

$$E\{(y_1 - \bar{Y})(y_2 - \bar{Y}) + (y_1 - \bar{Y})(y_3 - \bar{Y}) + \dots + (y_{n-1} - \bar{Y})(y_n - \bar{Y})\}$$

$$= \frac{n(n-1)}{N(N-1)} \{(y_1 - \bar{Y})(y_2 - \bar{Y}) + (y_1 - \bar{Y})(y_3 - \bar{Y}) + \dots + (y_{N-1} - \bar{Y})(y_N - \bar{Y})\}$$

次に、期待値をとるために等式

$$n(\bar{y} - \bar{Y}) = (y_1 - \bar{Y}) + (y_2 - \bar{Y}) + \dots + (y_n - \bar{Y}) \quad (11)$$

の両辺を平方して、(11.9)(11.10)を用いれば、次の結果が得られる:

$$\begin{aligned} n^2 E\{(\bar{y} - \bar{Y})^2\} &= \frac{n}{N} \{(y_1 - \bar{Y})^2 + \dots + (y_N - \bar{Y})^2\} \\ &+ \frac{2n(n-1)}{N(N-1)} \{(y_1 - \bar{Y})(y_2 - \bar{Y}) + \dots + (y_{N-1} - \bar{Y})(y_N - \bar{Y})\} \\ &= \frac{n}{N} \{(y_1 - \bar{Y})^2 + \dots + (y_N - \bar{Y})^2\} \end{aligned}$$

$$+ 2 \frac{n-1}{N-1} \{(y_1 - \bar{Y})(y_2 - \bar{Y}) + \dots + (y_{N-1} - \bar{Y})(y_N - \bar{Y})\} \quad (12)$$

しかし

$$\begin{aligned} \{(y_1 - \bar{Y}) + \dots + (y_N - \bar{Y})\}^2 &= (y_1 - \bar{Y})^2 + \dots + (y_N - \bar{Y})^2 \\ &+ 2\{(y_1 - \bar{Y})(y_2 - \bar{Y}) + \dots + (y_{N-1} - \bar{Y})(y_N - \bar{Y})\} \quad (13) \end{aligned}$$

(11.12) の最後の式の大括弧内の、最初の中括弧内の式を A, 第 2 の中括弧内の式を B,

(11.13) の左辺を C とおけば

$$ZB = C - A$$

であることを用いて 11.12 の右辺を変形すれば次式が得られる:

$$\begin{aligned} n^2 E\{(\bar{y} - \bar{Y})^2\} &= \frac{n}{N} \left\{ \left(1 - \frac{n-1}{N-1}\right) \{(y_1 - \bar{Y})^2 + \dots + (y_N - \bar{Y})^2\} \right. \\ &\quad \left. + \frac{n-1}{N-1} \{(y_1 - \bar{Y}) + \dots + (y_N - \bar{Y})\}^2 \right\} \quad (14) \end{aligned}$$

しかし、この右辺の第 2 の中括弧内の和は 0 になるから、(11.14) の両辺を n^2 で割れば、求める (11.8) が得られる。(終)

系 1. 標本平均 $\bar{y}$ の標準誤差 $\sigma_{\bar{y}}$ は次式で与えられる:

$$\sigma_{\bar{y}} = \frac{S}{\sqrt{n}} \sqrt{\frac{N-n}{N}} \quad (15)$$

系 2. 母総和 Y の推定量 $\hat{Y} = N\bar{y}$ の分散 $V(\hat{Y})$ および標準誤差 $\sigma_{\hat{Y}}$ はそれぞれ次式で与えられる:

$$V(\hat{Y}) = E\{(\hat{Y} - Y)^2\} = \frac{N^2 S^2}{n} \cdot \frac{N-n}{N}, \quad (16)$$

$$\sigma_{\hat{Y}} = \frac{Ns}{\sqrt{n}} \sqrt{\frac{N-n}{N}} \quad (17)$$

われわれは前に、元に戻す抽出、あるいは無限集団からの抽出においては、母分散を σ^2 とすれば、大きさ n の(ランダム)標本の平均 $\bar{y}$ の分散は $\frac{\sigma^2}{n}$ で与えられることを知っている。しかし(11.8)の結果によつて有限母集団からの、元に戻さない単純ランダム抽出においては、大きさ n の標本の平均 $\bar{y}$ の分散には、さらに、 $\frac{(N-n)}{N}$ という因数が伴ってくることをみる。これを有限修正項 finite correction とよんでいる。しかし

$$\frac{N-n}{N} = 1 - \frac{n}{N}$$

であるから、有限修正項は、抽出比 n/N が十分小であれば、ほとんど 1 に等しいことを知る。したがって、このような場合には、対象集団の大きさが、標本平均の標準誤差に対してなんら直接の影響を与えず、標本の大きさが n であれば、それは常に $\frac{s}{\sqrt{n}}$ に十分近いことが知られる。

例えば、2つの対象集団 π_1 (大きさ 250,000) および π_2 (大きさ 10,000) において、 S の値が等しいならば、 π_1 からの大きさ 100 の標本も、 π_2 からの大きさ 100 の標本も、母平均の推定に対してほとんど等しい精度の推定値を与えることを注意すべきである。常識的には、 π_1 からの大きさ 100 の標本による母平均の推定と、同程度の精度をもつ推定を、 π_2 からの標本で得るためには、 π_2 から大きさ 100×25 の標本をとらなければならないだろうと考えやすい。しかし、そうではないことを教えるのが上の諸定理である。

これら母平均および母総和の推定量の標準誤差は、次の 3つの目的のために用いられる：

1) 単純ランダム抽出法によつて得られる推定の精度と他の方法で得られる推定の精度とを比較するために、

2) 計画している調査に必要な標本の大きさを決定するために、

3) 完了した調査において、実際に達成された推定の精度を推定するために、

しかし、以上の推定量の分散に関する公式は、母分散 S^2 を含み、実際上これは未知である。したがって、それを標本から推定しなければならない。

そこで次に、いかにして推定量の標準誤差を推定するかを述べよう。

定理 11.3 単純ランダム抽出法において、次式で定義される標本分散 s^2

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2 \quad (18)$$

は、母分散 S^2 の不偏推定量である。

証明 (11.18) で定義される s^2 を、次のように書いておく：

$$\begin{aligned} s^2 &= \frac{1}{n-1} \sum_{i=1}^n \{ (y_i - \bar{Y}) - (\bar{y} - \bar{Y}) \}^2 \\ &= \frac{1}{n-1} \left[\sum_{i=1}^n (y_i - \bar{Y})^2 - n(\bar{y} - \bar{Y})^2 \right] \end{aligned} \quad (19)$$

しかして、次のことが計算される：

$$E \left\{ \sum_{i=1}^n (y_i - \bar{Y})^2 \right\} = \frac{n}{N} \sum_{i=1}^N (y_i - \bar{Y})^2 = \frac{n(N-1)}{N} S^2 \quad (20)$$

さらに、定理 11.2 によつて

$$E \{ n(\bar{y} - \bar{Y})^2 \} = \frac{N-n}{N} S^2$$

よつて、(19)、(20) および (21) から

$$E \{ s^2 \} = \frac{S^2}{(n-1)N} \{ n(N-1) - N - n \} = S^2 \quad (\text{終}) \quad (22)$$

系 1. $\bar{y}$ および $\hat{Y} \equiv N\bar{y}$ の分散 $V(\bar{y}) \equiv \sigma_{\bar{y}}^2$, および $V(\hat{Y}) \equiv \sigma_{\hat{Y}}^2$ の不偏推定量は、それぞれ、次式で定義される $v(\bar{y}) \equiv s_{\bar{y}}^2$, および $v(\hat{Y}) \equiv s_{\hat{Y}}^2$ で与えられる：

$$v(\bar{y}) \equiv s_{\bar{y}}^2 = \frac{s^2}{n} \cdot \frac{N-n}{N}, \quad (23)$$

$$v(\hat{Y}) \equiv s_{\hat{Y}}^2 = \frac{N^2 s^2}{n} \cdot \frac{N-n}{N}. \quad (24)$$

なお標準誤差の推定量としては、それぞれ、次式で与えられる $s_{\bar{y}}$ および $s_{\hat{Y}}$ を用いる

$$s_{\bar{y}} = \frac{s}{\sqrt{n}} \sqrt{\frac{N-n}{N}}, \quad (25)$$

$$s_{\hat{Y}} = \frac{Ns}{\sqrt{n}} \sqrt{\frac{N-n}{N}}. \quad (26)$$

これら標準誤差の推定量 $s_{\bar{y}}$ および $s_{\hat{Y}}$ は、少し偏つた推定量であるが、その偏りは、実際問題としてはほとんど影響を与えないくらいのものである。

実際問題としては、標本の大きさ n が、かなり大きいので、推定量 $\bar{y}$ および $\hat{Y}$ は、それぞれ対応する母集団値 $\bar{Y}$ および Y の周りに正規分布をする、と仮定できるのが普通である。もしこの仮定が許されるならば、母平均 $\bar{Y}$ および母総和 Y の信頼係数 $(1-\alpha)$ の信頼区間は、それぞれ、次の (27) および (28) で与えられると考えてよいであろう：

$$\left(\bar{y} + Z_{\frac{1}{2}\alpha} \frac{s}{\sqrt{n}} \sqrt{\frac{N-n}{N}}, \bar{y} + Z_{1-\frac{1}{2}\alpha} \frac{s}{\sqrt{n}} \sqrt{\frac{N-n}{N}} \right) \quad (27)$$

$$\left(N\bar{y} + Z_{\frac{1}{2}\alpha} \frac{Ns}{\sqrt{n}} \sqrt{\frac{N-n}{N}}, N\bar{y} + Z_{1-\frac{1}{2}\alpha} \frac{Ns}{\sqrt{n}} \sqrt{\frac{N-n}{N}} \right) \quad (28)$$

ここに $Z_{\frac{1}{2}\alpha}$ ($= -Z_{1-\frac{1}{2}\alpha}$) は標準正規分布 $N(0, 1)$ の $\frac{1}{2}\alpha \times 100\%$ 点である。

もし、標本の大きさ n が 30 以下であるならば、上式で $Z_{\frac{1}{2}\alpha}$ ($= -Z_{1-\frac{1}{2}\alpha}$) の代りに、自由度 $(n-1)$ の t -分布における $\frac{1}{2}\alpha \times 100\%$ 点 ($t_{\frac{1}{2}\alpha} = -t_{1-\frac{1}{2}\alpha}$) を用いるべきである。事実、もし観測値 y_i ($i=1, \dots, n$) が、互に独立に、平均値 $\bar{Y}$ をもつて、同一の正規分布をするならば、

$$\frac{\bar{y} - \bar{Y}}{\frac{s}{\sqrt{n}}}$$

は、正確に、自由度 $(n-1)$ の t -分布をすることを既に 6.6 で知っている。

3. 単純ランダム抽出法による比率の推定

対象集団のいずれの要素も、2つのクラス C および $\bar{C}$ のいずれか一つに分類されるものとし、対象集団 $\mathcal{U}$ の総数 N 個の要素のうち、クラス C に入るものの比率（母集団比率） p を、単純ランダム抽出によつて推定したいものとしよう。

この問題においては、 N 個の要素の教標識 $y_1, y_2, \dots, y_N$ を、クラス C に入る要素については $y_i = 1$ 、クラス $\bar{C}$ に入る要素については、 $y_i = 0$ と定義することにすれば、

$$p = \frac{1}{N} (y_1 + y_2 + \dots + y_N) = \bar{Y}$$

である。したがつて、対象集団の要素の教標識 y_i を、上のように定義することによつて、母集団比率 p を推定する問題は、母平均 $\bar{Y}$ を推定する問題に還元される。なお、対象集団 $\mathcal{U}$ で、クラス C に入る要素の個数を A で表わすことにすれば

$$A = y_1 + y_2 + \dots + y_N = Y$$

であり、 A に関する推定の問題は、母総和 Y に関する推定の問題に還元される。したがつて、 p および A を推定するためには、前に得られた諸結果をそのまま用いることが出来る。大きさ n の単純ランダム標本 $y_1, \dots, y_n$ において、クラス C に入る要素の比率、すなわち標本比率を $\hat{p}$ で表わすことにすれば

$$\hat{p} = \frac{1}{n} (y_1 + \dots + y_n) = \bar{y}$$

である。前の諸結果を用いるために、 S^2 および s^2 を、 P および p を用いて表現しておこう。そのために、次の関係式が成り立つことを注意しておく：

$$\sum_{i=1}^N y_i^2 = A = NP, \quad \sum_{i=1}^N y_i^2 = np$$

したがつて、(29) を用いて

$$s^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1} = \frac{\sum_{i=1}^n y_i^2 - N\bar{y}^2}{n-1} = \frac{1}{n-1} (NP - N\bar{y}^2),$$

すなわち、

$$s^2 = \frac{N}{n-1} P Q, \quad \text{ここに } Q = 1 - P. \quad (30)$$

同様にして、次の結果が得られる：

$$s^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1} = \frac{n}{n-1} P q, \quad q = 1 - p \quad (31)$$

したがつて、定理 11.1 および定理 11.2 を、今の場合に適用すれば、次の諸結果が得られる：

定理 4 標本比率 $\hat{p}$ は、母集団比率 P の不偏推定量である。

定理 5 推定量 $\hat{p}$ の分散 $V(\hat{p})$ は次式で与えられる：

$$V(\hat{p}) = E\{(\hat{p} - P)^2\} = \frac{S^2}{n} \cdot \frac{N-n}{N} = \frac{P Q}{n} \left(\frac{N-n}{N-1} \right)$$

系 1. 標本および対象集団で、クラス C に入る要素の百分率を、それぞれ、 p および P で表わしたものとすると、上の定理 11.4 および定理 11.5 は、そのままの形で成立する。

定理 11.6 対象集団 $\mathcal{U}$ において、クラス C に入る要素の数 A の不偏推定量 $\hat{A}$ は次式で与えられる：

$$\hat{A} = N \hat{p} \quad (33)$$

また、その分散 $V(\hat{A})$ は次式で与えられる：

$$V(\hat{A}) = \frac{N^2 P Q}{n} \left(\frac{N-n}{N-1} \right). \quad (34)$$

定理 11.7 推定量 $\hat{p}$ の分散 $V(\hat{p})$ の不偏推定量 $v(\hat{p})$ は次式で与えられる：

$$v(\hat{p}) \equiv S p^2 = \frac{N-n}{(n-1)N} p q. \quad (35)$$

証明 標本平均 $\bar{y}$ の分散の不偏推定量は、(23) で与えられることを知っている。(23) の s^2 の代りに (31) を代入すれば $V(\hat{p})$ の不偏推定量として次式が得られる：

$$v(\hat{p}) = \frac{s^2}{n} \cdot \frac{N-n}{N} = \frac{1}{n} \cdot \frac{n}{n-1} \cdot p q \cdot \frac{N-n}{N} = \frac{N-n}{(n-1)N} p q. \quad (\text{終})$$

もし、 n に比して N が十分大きく、有限修正項が無視できるならば、 p の分散の不偏推定量は (35) により次式で与えられる：

$$v(\hat{p}) = \frac{p q}{n-1}. \quad (36)$$

系 1. 対象集団において、クラス C に入る要素の数を A とする。しかるとき、 A の推定量 $\hat{A} = N \hat{p}$ の分散の不偏推定量は次式で与えられる：

$$V(\hat{A}) \equiv s_{Np}^2 = \frac{N(N-n)}{n-1} p q. \quad (37)$$

次に、母集団百分率 P の信頼区間を求めよう。

そのために、やはり、標本の大きさが大であるときには、標本比率 p は、 P を平均値とし、分散 $V(p)$ をもって正規分布をするとみなしてよいことを用いる。(32)により、 p の標準誤差 $Q_p = \sqrt{VP}$ は次式で与えられる：

$$Q_p = \sqrt{\frac{N-n}{N-1}} \sqrt{\frac{PQ}{n}}$$

この式の P を p で推定して、 p が標準誤差

$$Sp' = \sqrt{\frac{N-n}{N-1}} \sqrt{\frac{pq}{n}}$$

もって正規分布をすると仮定すれば、母集団比率 P の信頼係数 $(1 - \alpha)$ の信頼区間は次式で与えられる：

$$\left(P + Z_{1-\frac{\alpha}{2}} \sqrt{\frac{N-n}{N-1}} \sqrt{\frac{pq}{n}}, P - Z_{1-\frac{\alpha}{2}} \sqrt{\frac{N-n}{N-1}} \sqrt{\frac{pq}{n}} \right)$$

一方われわれは、定理 11.7 により、 $V(p)$ の不偏推定量が (35) で与えられることを知っているから、 p の標準誤差を

$$S_p = \sqrt{\frac{N-n}{N}} \sqrt{\frac{pq}{n-1}}$$

で定義するならば、(39)において、(38)で与えられる Sp' の代りに、(40)で与えられる S_p を用いるべきである。

しかし、実際問題としては、 S_p を用いても Sp' を用いてもほとんど相等しい程度の良い近似が得られるであろう。

4. 標本の大きさの決定

対象集団 π の母平均 $\bar{Y}$ を単純ランダム抽出法で得られた標本にもとづいて推定したいものとする。このとき、われわれは、単純ランダム標本の標本平均 $\bar{y}$ が $\bar{Y}$ の不偏推定量であることは知っている。しかしこのことは、各標本毎に計算される $\bar{y}$ の値が、いつも $\bar{Y}$ に等しいということを保証しているのではなく、一般に、標本から計算される $\bar{y}$ の値は $\bar{Y}$ と異なることであろう。実際上の問題として、推定の誤差の絶対値 $|\bar{y} - \bar{Y}|$ が、あらかじめ指定された値 d 以上になることは望ましいことではなく、そのようなことの起る確率が、あらかじめ指定された小さい値 α に等しくなることが要求されているものとする。すなわち

$$\Pr \{ |\bar{y} - \bar{Y}| \geq d \} = \alpha$$

であることが要求されているものとする。そして問題は、“この要求(41)を満たすためには、単純ランダム抽出法における標本の大きさ n を幾らにとるべきか”ということであるとしよう。

標本平均 $\bar{y}$ は正規分布をすると仮定できるものとする。その平均値は $\bar{Y}$ であり、分散 $V(\bar{y})$ は (3) で与えられる。したがって、このとき

$$z = \frac{\bar{y} - \bar{Y}}{\sqrt{\frac{N-n}{N}} \frac{s}{\sqrt{n}}}$$

は標準正規分布 $N(0, 1)$ をする。しかも不等式 $|\bar{y} - \bar{Y}| \geq d$ は、不等式

$$|z| \geq \frac{d}{\sqrt{\frac{N-n}{N}} \frac{s}{\sqrt{n}}} \quad (43)$$

と同値である。したがって(41)を成立せしめるためには

$$\frac{d}{\sqrt{\frac{N-n}{N}} \frac{s}{\sqrt{n}}} = Z_{1-\frac{\alpha}{2}} \quad (=z^* \text{ とおく}) \quad (44)$$

であるように標本の大きさ n をとればよい。(44)を n について解いて次の結果が得られる：

$$n = \frac{\left(\frac{z^*}{d}\right)^2 S^2}{1 + \frac{1}{N} \left(\frac{z^*}{d}\right)^2 S^2} \quad (45)$$

実際には、 S^2 の値が未知であるから、予備調査とか、過去の調査によつて、あらかじめ S^2 の値を推定し、その推定値 s^2 を (45) の S^2 の代りに代入して n の値を決定しなければならない。

N の値が相当大きい場合には、(45) の第一近似として

$$n_0 = \left(\frac{z^*}{d}\right)^2 S^2$$

をとることが出来る。この公式は、 n_0/N が十分小さいときには、適当なものである。しからざるときには

$$n = \frac{n_0}{1 + \frac{n_0}{N}} \quad (47)$$

をもつて標本の大きさとする。

次に、対象集団 π のすべての要素が 2 つのクラス C および C' のいずれか一方に分類される場

合に、クラスCに入るものの比率Pを推定したいものとする。このとき、標本比率pが、母集団比率Pの不偏推定量であることは知っている。

問題は“dおよびαをあらかじめ与えられた正数として、条件

$$\Pr\{|p-P| \leq d\} = \alpha \quad (48)$$

を満たすように、標本の大きさnを決定すること”であるとしよう。

標本比率pが正規分布をすると仮定できるくらいに、標本の大きさnが大であるとすれば、その平均値はp、分散V(p)は(32)で与えられる。したがって、上に求めたと同様にして、条件(48)を満たすためには

$$d = z^* \sqrt{\frac{N-n}{N-1}} \sqrt{\frac{PQ}{n}} \quad (49)$$

を満足するようにnの値を定めればよい。(49)をnについて解いて、次の結果が得られる：

$$n = \frac{\left(\frac{z^*}{d}\right)^2 PQ}{1 + \frac{1}{N} \left\{ \left(\frac{z^*}{d}\right)^2 PQ - 1 \right\}} \quad (50)$$

しかし実際には、pの値は未知であるから、予備調査等によつてPの値を推定し、その推定値pを(50)のPの代りに代入して、nの値を決定すべきである。もしNの値が十分大であるならば、(50)の第一近似として

$$n_0 = \left(\frac{z^*}{d}\right)^2 pq \quad (51)$$

を標本の大きさにとることが出来る。

もし、 n_0/N が十分小であるならば、(51)の n_0 が妥当な標本の大きさである。もし、 n_0/N が無視できないならば、

$$n = \frac{n_0}{1 + \frac{n_0 - 1}{N}} \doteq \frac{n_0}{1 + \frac{n_0}{N}}$$

をもつて、求める標本の大きさとする。

5. 層化抽出法

対象集団πを構成するN個の要素の数標識の平均値(母平均) $\bar{Y}$ を推定することが目的であるとする。そのために対象集団πを、互に重なり合うことがなく、またその和がπを尽くすような仕方ではL個の部分集団 $\pi_1, \dots, \pi_L$ に分割する。すなわち

$$\pi = \pi_1 + \pi_2 + \dots + \pi_L.$$

したがって部分集団 π_h を構成する要素の数を N_h ($h=1, \dots, L$)とすれば

$$N = N_1 + N_2 + \dots + N_L$$

である。このとき、各部分集団 π_h を層 stratumという、($h=1, \dots, L$)。

いま、各層 π_h において、互に独立に単純ランダム抽出を行うとき、この全手続きを層化抽出法 stratified random sampling という。そして、各層 π_h からの単純ランダム抽出における標本の大きさを n_h とする。

何故こうした層化抽出法を考えるかという、それは次のような考えにもとづいている。すなわち、もし各層 π_h の内部が、非常に均一であるように層化できるならば、各層からの小さな標本によつて、層平均のよい推定値が得られるであろう。そうすれば、これらをうまく結合することによつて、小さな標本で、しかも精度よく母平均を推定できることが考えられるからである。

例えば、幼稚園から大学まである学校で、全学生生徒の平均身長を推定したいために、200人の人を抽出するものとしよう。このとき、幼稚園から大学までの学生生徒全体の集団がπである。いまその学校で、幼稚園の児童全体の集団を π_1 、小学生全体の集団を π_2 、中学生全体の集団を π_3 、高等学生全体の集団を π_4 、大学生全体の集団を π_5 とすれば、全集団πが $\pi_1, \pi_2, \pi_3, \pi_4, \pi_5$ の5つの層に分割されることになる。このとき常識的に考えても、全集団πからランダムに200人抽出するよりも、各層 π_i から幾人かずつ、計200人抽出して、各層の平均身長を推定し、それらの推定値を適合に結合して全学生生徒の平均身長の推定値を求める方がよいのではないかと考えられる。そうした考えに対する理論とその結論を述べようとするのが本節の目的である。

このような層化抽出法においては、各層別抽出からの推定値をどのように結合して母平均の推定値を求めるか、また、より良い精度の推定を得るためには、各層からの標本の大きさ n_h をどのようにとるか、ということが問題になる。

(もちろん、層化抽出法においては、その出発点として、層をどのように構成するか、ということが非常に大事な問題である。ここでは、層化が既に行われているものとして、そこから議論を進めることとする。)

本節において、添数hは層を、iは層内の要素を表わすものとして、次の記号を用いる：

N_h = 層 π_h に含まれる要素の総数、

n_h = 層 π_h からの標本の大きさ、

Y_{hi} = 層 π_h の第i要素の数標識

$$\bar{Y}_h = \frac{1}{N_h} \sum_{i=1}^{N_h} Y_{hi} = \text{層}\pi_h\text{の平均値,}$$

$$\bar{y}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} y_{hi} = \text{層}\pi_h\text{からの標本の平均値,}$$

$$S_h^2 = \frac{1}{N_h - 1} \sum_{i=1}^{N_h} (y_{hi} - \bar{Y}_h)^2 = \text{層}\pi_h \text{の分散}$$

各層 h から単純ランダム抽出法によつて、標本 $(y_{h1}, \dots, y_{hN_h})$ を求めるのであるから、ここに、11.2 におけると同じ意味で、一つの母集団 $(\pi_h, (y_{h1}, \dots, y_{hN_h}))$ が構成される ($h=1, \dots, L$)。これらをすべて結合したものを

$$(\pi = \pi_1 + \dots + \pi_L, (y_{11}, \dots, y_{1n_1}, \dots, (y_{L1}, \dots, y_{1n_L})))$$

で表わし、 $(y_{11}, \dots, y_{1n_1}, \dots, y_{L1}, \dots, y_{Ln_L})$ を母確率変数として、ここに層化抽出法に伴う母集団が設定される。この母確率変数の実現値が層化抽出法における標本である。

層 π_h からの標本平均 $\bar{y}_h$ ($h=1, \dots, L$) を結合して、次の統計量 $\bar{y}_{st}$ を作る：

$$\bar{y}_{st} = \frac{\sum_{h=1}^L N_h \bar{y}_h}{N} \quad (54)$$

もし標本 $(y_{11}, \dots, y_{Ln_L})$ を単純に平均するならば、その平均値 $\bar{y}$ は次のように書くことが出来る：

$$\bar{y} = \frac{\sum_{h=1}^L n_h \bar{y}_h}{n}, \quad \text{ここに} \quad n = n_1 + \dots + n_L \quad (55)$$

もし

$$\frac{n_h}{n} = \frac{N_h}{N}, \quad \text{すなわち} \quad \frac{n_h}{N_h} = \frac{n}{N} = \text{一定}, \quad h=1, \dots, L,$$

であるならば $\bar{y} = \bar{y}_{st}$ である。しからざる場合には、一般には $\bar{y}_{st}$ と $\bar{y}$ とは異なるであろう。

特に、各層の標本の大きさ n_h が (56) を満足するようにとられているときに、これを比例割当による層化抽出法という。

定理 11.8 層化抽出法において、(54) で定義される推定量 $\bar{y}_{st}$ は、母平均 $\bar{Y}$ の不偏推定量である。

証明 各層 π_h で単純ランダム抽出が行われているのであるから、定理 11.1 によつて

$$E\{\bar{y}_h\} = \bar{Y}_h$$

である。このことを用いて次式が計算される。

$$E\{\bar{y}_{st}\} = E\left\{\frac{\sum_{h=1}^L N_h \bar{y}_h}{N}\right\} = \frac{\sum_{h=1}^L N_h \bar{Y}_h}{N} \quad (57)$$

しかるに、母平均 $\bar{Y}$ は次のように書ける：

$$\bar{Y} = \frac{1}{N} \sum_{h=1}^L \sum_{i=1}^{N_h} y_{hi} = \frac{1}{N} \sum_{h=1}^L N_h \bar{Y}_h. \quad (58)$$

よつて、(57), (58) から

$$E\{\bar{y}_{st}\} = \bar{Y}$$

が成立する。(終)

定理 11.9 $\bar{y}_{st}$ の分散 $V(\bar{y}_{st})$ は次式で与えられる：

$$V(\bar{y}_{st}) = \frac{\sum_{h=1}^L N_h^2 V(\bar{y}_h)}{N^2}, \quad (59)$$

ここに

$$V(\bar{y}_h) = E(\bar{y}_h - \bar{Y}_h)^2.$$

証明 次のような計算をしておく：

$$\bar{y}_{st} - \bar{Y} = \frac{\sum N_h \bar{y}_h}{N} - \frac{\sum N_h \bar{Y}_h}{N} = \frac{\sum N_h (\bar{y}_h - \bar{Y}_h)}{N},$$

ここに $\sum$ は、すべての層にわたつて加えることを示す。したがつて

$$(\bar{y}_{st} - \bar{Y})^2 = \frac{1}{N^2} \sum N_h^2 (\bar{y}_h - \bar{Y}_h)^2 + \frac{2}{N^2} \sum N_h N_j (\bar{y}_h - \bar{Y}_h)(\bar{y}_j - \bar{Y}_j), \quad (62)$$

ここに、右辺の第 2 項における $\sum$ は、層のすべての相異なる 2 つの組にわたつて加えるものである。

(62) の両辺の期待値をとるのに、各層 π_h から互に独立に、単純ランダム抽出が行われているのであるから、右辺の第 2 項の期待値は 0 に等しい。したがつて

$$V(\bar{y}_{st}) = E\{(\bar{y}_{st} - \bar{Y})^2\} = \frac{1}{N^2} \sum N_h^2 E(\bar{y}_h - \bar{Y}_h)^2 = \frac{1}{N^2} \sum N_h^2 V(\bar{y}_h)$$

よつて、(59) の成立が証明された。(終)

この定理 11.9 の教える大切なことは、推定量 $\bar{y}_{st}$ の分散が、各層 π_h の層平均 $\bar{Y}_h$ の推定量 $\bar{y}_h$ の分散 $V(\bar{y}_h)$ にだけ依存しているということである。したがつて、対象集団 π 全体としては変動が大きくても、もし各層 π_h 内では非常に斉一であるように層化されているものであるならば、母平均 $\bar{Y}$ は $\bar{y}_{st}$ によつて精度よく推定されることになる。

定理 11.2 を各層に用いて

$$V(\bar{y}_h) = \frac{S_h^2}{n_h} \cdot \frac{N_h - n_h}{N_h}$$

が得られ、これを (59) に代入すれば次の結果が得られる：

定理 11.10 層化抽出法における推定量 $\bar{y}_{st}$ の分散 $V(\bar{y}_{st})$ は次のように書ける:

$$V(\bar{y}_{st}) = \frac{1}{N^2} \sum_{h=1}^L N_h (N_h - n_h) \frac{S_h^2}{n_h} \quad (63)$$

系 1. もし各層で抽出比 $\frac{n_h}{N_h}$ が無視できるならば

$$V(\bar{y}_{st}) = \frac{1}{N^2} \sum_{h=1}^L N_h^2 \frac{S_h^2}{n_h} = \sum_{h=1}^L \frac{W_h^2 S_h^2}{n_h}, \quad \text{ここに } W_h = \frac{N_h}{N} \quad (64)$$

系 2 比例割当による層化抽出法では

$$n_h = n \frac{N_h}{N}$$

を(63)に代入して次の結果が得られる:

$$V(\bar{y}_{st}) = \sum_{h=1}^L \frac{N_h}{N} \cdot \frac{S_h^2}{n} \left(\frac{N - n}{N} \right) = \frac{N - n}{hN} \sum_{h=1}^L W_h S_h^2 \quad (65)$$

母総和 Y の推定のためには、定理 11.1 の系 1, および定理 11.10 から、容易に次の結果が得られる:

定理 11.11 $\hat{Y}_{st} \equiv N \bar{y}_{st}$ は母総和 Y の不偏推定量であり、その分散 $V(\hat{Y}_{st})$ は次式で与えられる:

$$V(\hat{Y}_{st}) = \sum_{h=1}^L N_h (N_h - n_h) \frac{S_h^2}{n_h}$$

以上の諸公式によつて推定の精度を知ろうとしても、それは未知の分散 S_h^2 を含んでいるからできない。しかるに定理 11.3 により、分散 S_h^2 の不偏推定量が

$$s_h^2 = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2$$

で与えられるから次の結果が得られる:

定理 11.12 層化抽出法における推定量 $\bar{y}_{st}$ の分散の不偏推定量 $v(\bar{y}_{st})$ は次式で与えられる:

$$v(\bar{y}_{st}) \equiv S^2(\bar{y}_{st}) = \frac{1}{N^2} \sum_{h=1}^L N_h (N_h - n_h) \frac{S_h^2}{n_h} \quad (68)$$

もし、 $\bar{y}_{st}$ の確率分布が正規分布であると仮定できるならば、(68) の $s(\bar{y}_{st})$ を用いて信頼系数 $1 - \alpha$ の信頼区間は、近似的に次のように与えられるであろう:

母平均について: $(\bar{y}_{st} + z_{1-\frac{\alpha}{2}} s(\bar{y}_{st}), \bar{y}_{st} + z_{1-\frac{\alpha}{2}} s(\bar{y}_{st}))$,

母総和について: $(N \bar{y}_{st} + z_{1-\frac{\alpha}{2}} N s(\bar{y}_{st}), N \bar{y}_{st} + z_{1-\frac{\alpha}{2}} N s(\bar{y}_{st}))$.

ここに $z_{1-\frac{\alpha}{2}}$ は、標準正規分布の $\frac{1}{2} \times 100\%$ 点である。

6. 標本の大きさ n_h の最適割当

各層における標本の大きさ n_h をどのように選定するかに関して、ネイマン Neyman の割当法とよばれる次の定理を証明しよう。

定理 11.13 層化抽出法において、標本の全体としての大きさ n が一定値として与えられているときに、推定量 $\bar{y}_{st}$ の分散を最小にするという意味での最適な割当は、 n_h が $N_h S_h$ に比例するとき、すなわち

$$n_h \propto N_h S_h, \quad h=1, \dots, L \quad (69)$$

であるときである。さらに詳しくは、 n_h が次式(70)で与えられるときである。

$$n_h = \frac{n}{\sum N_h S_h} N_h S_h, \quad h=1, \dots, L$$

証明 問題は、(63) で与えられる $\bar{y}_{st}$ の分散 $V(\bar{y}_{st})$ を、制約条件

$$n_1 + \dots + n_L = n \quad (71)$$

のもとで、 $n_1, \dots, n_L$ に関して最小にする $n_1, \dots, n_L$ の値を求めることである。そのために、一応 n_h を連続変数と考えて、ラグランジュ Lagrange の乗数法を用いて次式を作る:

$$\begin{aligned} F &= V(\bar{y}_{st}) + (n_1 + \dots + n_L - n) \\ &= \frac{1}{N^2} \sum_{h=1}^L \left(\frac{N_h^2}{n_h} - N_h \right) S_h^2 + (n_1 + \dots + n_L - n) \end{aligned}$$

$\frac{\partial F}{\partial n_h} = 0$, ($h=1, \dots, L$) を計算して次の方程式が得られる:

$$-\frac{N_h^2}{N^2} \frac{S_h^2}{n_h^2} + \lambda = 0, \quad h=1, \dots, L \quad (73)$$

これから次の結果が得られる:

$$n_h = \frac{N_h S_h}{N \sqrt{\lambda}}, \quad \text{すなわち, } n_h \propto N_h S_h, (h=1, \dots, L). \quad (74)$$

n_h の値を具体的に求めるためには、(74) をすべての層の上で加えて

$$\sum n_h = n = \frac{\sum N_h S_h}{N \sqrt{\lambda}} \quad (75)$$

この式から $\sqrt{\lambda}$ の値を求めて、(74) に代入すれば、 n_h の値として(70)の結果が得られる。

なおこのとき、公式(63)の nh に(70)の値を代入することによつて、分散の最小値 $V_{\min}(\bar{y}_{st})$ は次式で与えられる:

$$V_{\min}(\bar{y}_{st}) = \frac{1}{N^2} \frac{(\sum N_h S_h)^2}{n} - \frac{1}{N^2} \sum N_h S_h^2 \quad (76)$$

定理11.13 は、層の大きさが大きいほど、また層内変動の大きさが大きいほど、その層から大きい標本をとるべきであることを教えている。

7. 層化抽出法と単純ランダム抽出法との相対精度

対象集団からとるべき標本の、全体としての大きさ n は一定であるものとして、層化抽出法による場合と単純ランダム抽出法による場合とで、母平均の推定量の分散の大小が(したがって精度の良し悪しが)どうなっているかを調べてみよう。ほとんど大抵の場合に、層化抽出法による推定の精度は単純ランダム抽出法による推定の精度よりもすぐれている。ただ、層のとり方、標本の割当如何によつて、ある極端な場合には、層化抽出標本の精度が、単純ランダム標本のそれよりも落ちる場合があることが示されるであろう。

まず、標本の全体としての大きさ n が一定という条件のもとで、次の三つの抽出法をとる場合に、母平均の推定量の分散 V をそれぞれ次のように表わすこととする。

V_{ran} : 単純ランダム抽出の場合、

V_{prop} : 比例割当をした層化抽出の場合、

V_{opt} : 最適割当(すなわち $nh \propto N_h S_h$ のとき)をした層化抽出の場合、

しかるとき次の定理が得られることを証明しよう。

定理11.14 項 $\frac{nh}{N_h}$ を無視し、したがって有限修正項を無視するならば、次の関係が成り立つ:

$$V_{opt} \leq V_{prop} \leq V_{ran} \quad (77)$$

もし、有限修正項を無視することができないならば、次の等式が成り立つ:

$$V_{ran} = V_{prop} + \frac{(N-n)}{nN(N-1)} \left\{ \sum N_h (Y_h - \bar{Y})^2 - \frac{1}{N} \sum (N - N_h) S_h^2 \right\}.$$

証明 有限修正項を無視すれば

$$V_{ran} = \frac{S^2}{n}, \quad ((8) \text{により}), \quad (79)$$

$$V_{prop} = \frac{1}{nN} \sum N_h S_h^2, \quad ((65) \text{により}), \quad (80)$$

$$V_{opt} = \frac{1}{nN^2} (\sum N_h S_h)^2, \quad ((96) \text{により}). \quad (81)$$

一方、分散分析における恒等式(7.1)とまったく同様にして、次の等式が成り立つ:

$$\begin{aligned} (N-1) S^2 &= \sum_h \sum_i (y_{hi} - \bar{Y})^2 \\ &= \sum_h \sum_i (y_{hi} - \bar{Y}_h)^2 + \sum_h N_h (\bar{Y}_h - \bar{Y})^2 \\ &= \sum_h (N_h - 1) S_h^2 + \sum_h N_h (\bar{Y}_h - \bar{Y})^2. \end{aligned} \quad (82)$$

しかるに、 $1/N_h$ のような項は無視できるものと仮定しているから、(82)は次のように書ける:

$$NS^2 = \sum_h N_h S_h^2 + \sum_h N_h (\bar{Y}_h - \bar{Y})^2 \quad (83)$$

よつて、(79), (83)により

$$V_{ran} = \frac{S^2}{n} = \frac{\sum N_h S_h^2}{nN} + \frac{\sum N_h (\bar{Y}_h - \bar{Y})^2}{nN}.$$

ここで、(80)を用いて次式が得られる:

$$V_{ran} = V_{prop} + \frac{\sum N_h (\bar{Y}_h - \bar{Y})^2}{nN}, \quad (84)$$

また、 V_{opt} の定義から、当然、 $V_{opt} \leq V_{prop}$ であるから、(84)により(77)が成り立つことがわかる。同様な解析で(78)の成立も証明される。(終)

(78)の結果から次のことが知られる。すなわち、もし

$$\sum N_h (\bar{Y}_h - \bar{Y})^2 < \frac{1}{N} \sum (N - N_h) S_h^2 \quad (85)$$

であるならば $V_{prop} > V_{ran}$ 、しからざるときには $V_{prop} \leq V_{ran}$ となる。それでは(85)が成立するのはどういう場合かという、 $\bar{Y}_h$ がすべて相等しいような場合である。こうした極端な場合を除けば、右辺に除数 N があるから、この不等式(85)は成立しないであろう。したがって一般には $V_{prop} \leq V_{ran}$ である。

8. 系統抽出法

N 個の要素から成る対象集団 π から、大きさ n の標本をとるのに次の方法によるとき、その方法を系統抽出法systematic samplingという。

すなわち、対象集団の N 個の要素に1番から N 番までの番号を付し、初めの1番から K 番までの番号の中から、ランダムに一つの番号を選び、それから以後、 K 番目毎の要素を次々にとつて、結局大きさ n の標本を抽出するのである。例えば、 K が20で、最初に抽出された番号が7であれば、7, 27, 47, ... の番号の要素を抽出するのである。この抽出法は、単純ランダム抽出法よりも、実行が非常に便利である。この際もしも、 $N = nK$ であるならば、 K 個の系統的標本が可能なので、これを表示すると次の表のようになる。

標 本 番 号	1..... i..... K
標 本	$y_1 \dots y_i \dots y_k$ $y_{k+1} \dots y_{k+1} \dots y_{k+1}$ $\vdots$ $y_{(n-1)k+1} \dots y_{(n-1)k+1} \dots y_{nk}$

しかし、一般にはNはKの整数倍になつていないであろうから、その場合には、同一対象集団からの系統的標本で、標本の大きさが1だけ違うものが出てくる。例えば、N=22, K=4とすれば、4つの系統的標本の要素の番号は、次の表のようになっている：

標 本 番 号	I	II	III	III
	1	2	3	4
標 本 の 要 素	5	6	7	8
の 番 号	9	10	11	12
	13	14	15	16
	17	18	19	20
	21	22		

すなわち、初めの2つの系統的標本の大きさは6であるが、後の2つの系統的標本の大きさは5である。しかし、この差はnが50以上ならばほとんど無視してもよいことが知られているので、今後、われわれは、N=Knであると仮定して議論をすすめることにする。このとき、系統抽出法に伴つて $(y_i, y_{k+i}, \dots, y_{(n-1)K+i})$ を母確率変数として、母集団 $(y_i, y_{k+i}, \dots, y_{(n-1)K+i})$ が設定されることになり、この母確率変数の実現値が、大きさnの系統的標本となるのである。

今後系統的標本の平均値を一般に $\bar{y}_{sy}$ で、また、第i系統的標本の第j要素を $\bar{y}_{ij}$ で表わすことにする。したがつて、 $i=1, \dots, K$ $j=1, \dots, n$ である。また、第i標本の平均値を $\bar{y}_i$ で表わす。すなわち $\bar{y}_i = \frac{1}{n} \sum_j y_{ij}$ 。

系統的標本抽出は、他面、次のように考えることも出来る。すなわち、対象集団をn個の要素 $(y_i, y_{k+i}, \dots, y_{(n-1)K+i})$, $i=1, 2, \dots, K$ から成るK個の部分集団に分割し、改めてこれら部分集団を抽出単位にとり、そして、これら抽出単位の一つをランダムに抽出して、これを大きさnの標本と考えれば、これは上の系統的標本にはかならない。

(これは後に述べる集落抽出法で、唯一つの集落を抽出する、という特別の場合に当たっている。)

定理11.15 もし、N=Knならば、大きさnの系統的標本の標本平均 $\bar{y}_{sy}$ は、母平均 $\bar{Y}$

の不偏推定量である。

証明 系統抽出法を、上述のように集落抽出法の特別の場合と考えれば、 $\bar{y}_{sy}$ はK個の値 $\bar{y}_1, \dots, \bar{y}_K$ をそれぞれ確率 $1/K$ でとる確率変数である。したがつて

$$E\{\bar{y}_{sy}\} = \frac{1}{K} (\bar{y}_1 + \dots + \bar{y}_K) = \frac{1}{nk} \sum_i \sum_j y_{ij} = \bar{Y}$$

が成り立ち、よつて証明された。(終)

定理11.16 系統的標本の平均 $\bar{y}_{sy}$ の分散 $V(\bar{y}_{sy})$ は次式で与えられる：

$$V(\bar{y}_{sy}) = \frac{N-1}{N} S^2 - \frac{K(n-1)}{N} S_{wsy}^2$$

ここに

$$S_{wsy}^2 = \frac{1}{K(n-1)} \sum_{i=1}^K \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2$$

は、系統的標本内分散とよばれるものであり、分母の $K(n-1)$ は分子の平方和の自由度である。

証明 分散分析における恒等式から、次の等式が成り立つ：

$$(N-1)S^2 = \sum_i \sum_j (y_{ij} - \bar{Y})^2 = n \sum_i (y_i - \bar{Y})^2 + \sum_i \sum_j (y_{ij} - \bar{y}_i)^2$$

しかるに、 $\bar{y}_{sy}$ の分散は、その定義から次式で与えられる：

$$V(\bar{y}_{sy}) = E\{(\bar{y}_{sy} - \bar{Y})^2\} = \frac{1}{K} \sum_i (\bar{y}_i - \bar{Y})^2 \quad (89)$$

したがつて、(88)、(89)から次式が成り立つ：

$$(N-1)S^2 = nK V(\bar{y}_{sy}) + K(n-1)S_{wsy}^2$$

したがつて、(86)の成立することがわかる。(終)

定理11.17 系統的標本の平均 $\bar{y}_{sy}$ の分散が、単純ランダム標本の平均 $\bar{y}$ の分散より小となるのは

$$S_{wsy}^2 > s^2 \quad (91)$$

の成立するときであり、またそのときに限る。(ただし標本の大きさはいずれもnとする。)

証明(11.8)および(11.86)の結果から $V(\bar{y}_{sy}) < V(\bar{y})$ となるのは

$$\frac{N-1}{N} S^2 - \frac{K(n-1)}{N} S_{wsy}^2 < \frac{N-n}{N} \cdot \frac{s^2}{n}$$

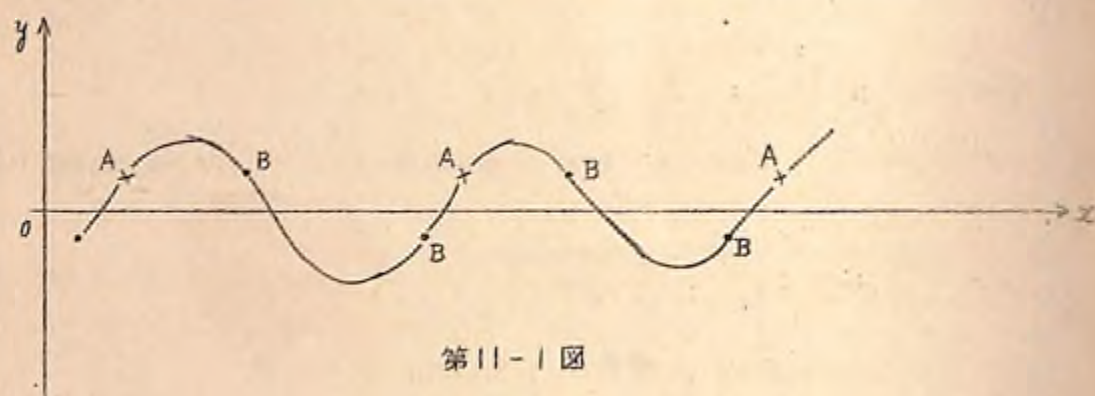
すなわち

$$K(n-1)S_{wsy}^2 > (N-1 - \frac{N-n}{n}) s^2 = K(n-1)S^2$$

の成立するときであり、またそのときに限る。よつて(91)の結果が証明された。(終)

この定理 11.17 から次のことがわかる。すなわち、系統抽出法が、単純ランダム抽出法よりも精度の良い母平均の推定を与えるのは、系統的標本内分散が、母分散よりも大きい場合においてである。換言すれば、系統抽出法は、同一標本内の要素が不斉一であるときに精度が高く、それらが斉一であるときに精度が低い、ということになる。この結果は直観的にも明かなことである。

もし対象集団の要素の数標識が、周期性をもつ場合には、系統抽出法はよほど注意して適用されなければならない。例えば、対象集団の要素の数標識が正弦曲線を描いて表われてくるものとしよう。しかるとき、系統抽出法の効力は K の値に依存する。



第 11 - 1 図において、曲線上の点の縦座標が数標識 y である。点 A で表わされる系統的標本は、最も不都合な場合である。すなわち、この場合は K が正弦曲線の周期 2π の整数倍になつており、系統的標本における各観測値が、すべて同一値をとり、大きさ n の標本であつても、それは唯一つの観測値だけの効果しかない。これに対して、点 B で表わされる系統的標本は、最も好都合な場合である。これは、 K が半周期 π の奇数倍になつているときには、どの系統的標本の平均も、母平均に等しくなつていゝ。しかしてこれらの両極端の中間に位する場合には、 K と周期との関係に依存して、各種の効力をもつ系統的標本が得られることになる。

9. 集落抽出法

対象集団 π を、各種の仕方でも部分に分割し、その各部分集団を抽出単位とする場合がある。例えば、全国の世帯について所得を研究するという問題においては、各世帯を要素として、全国にわたる要素（二世帯）の集団が対象集団 π である。ここから標本をとるのに、全国を郡に分割して、郡を抽出単位にとり、そこから幾つかの郡を抽出して、抽出された郡についてはす

べての世帯を調査するというような場合である。

このように、対象集団 π を幾つかの部分集団に分割し、これら部分集団そのものを抽出単位にとるとき、これら各部分集団を集落（集落 cluster という。（あるいは、単位ともよぶことにする）その意味で、こうした抽出法を集落抽出法 cluster sampling という。

層化抽出法においては、どのようにして対象集団を層化するかが問題であるように、集落抽出法においても、対象集団をどのように集落に分割するかが重要な問題である。

ここでは簡単のために、対象集団 π が、すべて M 個の要素から成る単位（二集落）に分割されたものとし、単位の総数を N とする。したがつて対象集団 π の要素の総数は MN である。次の記号を用いる：

y_{ij} = 第 i 単位内に含まれる第 j 要素の数標識、

y_i = 第 i 単位に含まれる要素の数標識の総和

$$\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i = \text{単位当り平均値、}$$

$$\bar{Y} = \frac{1}{N} \sum_{i=1}^N y_i = \frac{\bar{y}}{M} = \text{要素当り平均値、}$$

われわれの目的は、対象集団 π における要素当り平均値 $\bar{Y}$ を推定することであるとする。そのために、 N 個の単位の中から、単純ランダム抽出法で n 個の単位を抽出して、それらの単位に含まれるすべての要素の数標識を調査するという、集落抽出法を行うものとする。

まず、抽出された n 個の単位について、各単位に含まれる要素の数標識の総和 y_i ($i = 1, \dots, n$) を求め、

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = \text{標本における要素当り平均値}$$

をもつて $\bar{Y}$ の推定値にとることとする。このとき、次の定理が成り立つ：

定理 11.18 $\bar{y}$ は $\bar{Y}$ の不偏推定量である。

証明 いま、 $\bar{Y}$ および $\bar{y}$ を次のように書きかえてみる：

$$\bar{Y} = \frac{1}{N} \sum_{i=1}^N \left(\frac{y_i}{M} \right), \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n \left(\frac{y_i}{M} \right).$$

したがつて、今の問題は、あらためて数標識 $y_i^* = \frac{y_i}{M}$ ($i = 1, \dots, N$) をもつ N 個の要素から成る対象集団 π^* を考えて、そこから単純ランダム抽出法で n 個の要素を抽出することにより、母集団 $(\pi^*, (y_1^*, \dots, y_N^*))$ を設定し、その母平均 $\bar{Y}$ を、標本平均 $\bar{y}$ によつて推定していることになる。したがつて、定理 11.1 により、 $\bar{y}$ は $\bar{Y}$ の不偏推定量である。（終）

対象集団 π における要素間の分散 S^2 は、定義から次式で与えられる：

$$S^2 = \frac{\sum_{i=1}^N \sum_{j=1}^M (y_{ij} - \bar{Y})^2}{MN - 1} \quad (93)$$

なお、集落間相関係数 p を次式で定義する：

$$p = \frac{\sum_j \sum_{j < k} (y_{ij} - \bar{Y})(y_{ik} - \bar{Y})}{\frac{MN(M-1)S^2}{2}} \quad (94)$$

つぎに、推定量 $\bar{y}$ の分散 $V(\bar{y})$ を求めてみよう。そのために

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = \text{標本における単位当り平均値とおけば,}$$

$$\bar{\bar{y}} = \frac{1}{M} \bar{y}$$

である。いま、あらためて数標識 y_i , ($i=1, \dots, N$) をもつ N 個の要素から成る対象集団 π^{**} を考え、そこから単純ランダム抽出法で n 個の要素を抽出することによって、母集団 $(\pi^{**}, (y_1, \dots, y_n))$ を想定すれば、定理 11.2 より、 y の分散 $V(\bar{y})$ は次式で与えられる：

$$V(\bar{y}) = \frac{N-n}{Nn} \cdot \frac{\sum_{i=1}^N (y_i - \bar{Y})^2}{N-1} \quad (95)$$

したがって

$$V(\bar{\bar{y}}) = \frac{1}{M^2} V(\bar{y}) = \frac{(N-n)}{NnM^2} \cdot \frac{\sum_{i=1}^N (y_i - \bar{Y})^2}{N-1}$$

しかし

$$y_i - \bar{Y} = (y_{i1} - \bar{Y}) + (y_{i2} - \bar{Y}) + \dots + (y_{iM} - \bar{Y})$$

と書けることから、両辺を平方して、 N 個の単位 (集落) のすべてにわたって加えれば、次式が得られる：

$$\begin{aligned} \sum_{i=1}^N (y_i - \bar{Y})^2 &= \sum_{i=1}^N \sum_{j=1}^M (y_{ij} - \bar{Y})^2 + 2 \sum_{i=1}^N \sum_{j < k} (y_{ij} - \bar{Y})(y_{ik} - \bar{Y}) \\ &= (MN-1)S^2 + NM(M-1)pS^2. \end{aligned} \quad (97)$$

(97) を (96) に代入することによって、次の結果が得られる：

定理 11.19 推定量 y の分散 $V(\bar{y})$ は次式で与えられる：

$$V(\bar{y}) = \frac{N-n}{N} \cdot \frac{1}{nM} \cdot S^2 \left\{ \frac{(MN-1)}{(N-1)M} + \frac{N(M-1)}{N-1} p \right\}.$$

N が十分大きいときには、近似的に次のように書ける：

$$V(\bar{y}) = \frac{N-n}{N} \cdot \frac{1}{nM} S^2 \{ 1 + (M-1)p \}. \quad (99)$$

もし対象集団 π から、単純ランダム抽出法で nM 個の要素を抽出して $\bar{y}$ を計算するのであれば、

そのときの推定量 $\bar{y}$ の分散 $V(\text{ran} \bar{y})$ は定理 11.2 より、次式で与えられる。

$$V(\text{ran} \bar{y}) = \frac{N-n}{N} \cdot \frac{1}{nM} S^2.$$

したがって (99) および (100) により

$$\frac{V(\bar{y})}{V(\text{ran} \bar{y})} = 1 + (M-1)p. \quad (101)$$

したがって、同じ個数の要素を調査する場合、

$p > 0$ であれば、集落抽出法が単純ランダム抽出法よりも精度が悪く、

$p < 0$ であれば、集落抽出法が単純ランダム抽出法よりも精度が良い。

10. 2 段抽出法

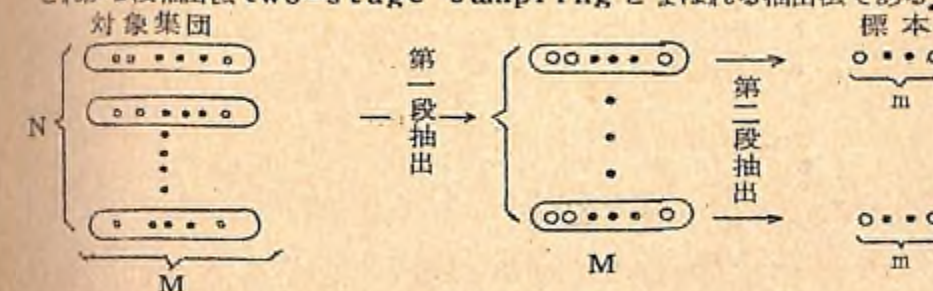
集落抽出法においては、対象集団 π を集落とよばれる部分集団に分割して、これを抽出単位にとり、そこから幾つかの単位を抽出し、抽出された単位については、そのすべての要素を調査するという方法であった。しかし、もし単位の内部において、要素の数標識が割合に均一であるならば、そのすべての要素について調査することは不経済なことである。そのようなときには抽出する単位の数を多くして、その代りに抽出された単位からさらにその一部分の要素を抽出して、それについて、数標識を調査することが望ましいであろう。こうした考えに立つ抽出法が 2 段抽出法である。これはいろいろの形において用いられ、非常に応用の広い抽出法である。

ここでは、対象集団 π が、すべて M 個の要素から成る部分集団に分割され、これが最初の抽出における抽出単位にとられているものとするこの部分集団を第 1 次抽出単位 (あるいは、簡単に単位) という。第 1 次抽出単位の総数は N 個であるとする。

第 1 段階として、 N 個の第 1 次抽出単位の集団の中から、単純ランダム抽出法で n 個の単位を抽出する。つぎに

第 2 段階として、抽出された n 個の単位のおのおのから、単純ランダム抽出法で m 個の要素を抽出し、これら合計 nm 個の要素について、その数標識を調査して、これをもって標本とする。

これが 2 段抽出法 two-stage sampling とよばれる抽出法である。



第 11 - 2 図

次の記号を用いる：

y_{ij} = 第 i 抽出単位における第 j 要素の数標識，

y = 標本における要素の数標識和，

y_i = 第 i 抽出単位から標本における数標識の総和

$$\bar{y}_i = \frac{1}{m} \sum_{j=1}^m y_{ij} = \frac{y_i}{m} = \text{第 } i \text{ 抽出単位からの標本における要素当り平均値，}$$

$$\bar{\bar{y}} = \frac{1}{n} \sum_{i=1}^n \bar{y}_i = \frac{y}{nm} = \text{標本全体における要素当り平均値，}$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n \frac{y_i}{m} = \frac{y}{nm} = \text{標本における抽出単位当りの平均値}$$

以上は、標本についての記号であるが、対象集団 π についても、 $Y, Y_i, \bar{Y}_i, \bar{\bar{Y}}$ 、および $\bar{Y}$ を上と同様に定義する。そして、われわれの目的は、対象集団 π における要素当り平均値 $\bar{Y}$ を推定することであるとする。

いま、第 i 抽出単位の第 i 要素に対する観測値 y_{ij} は、次の構造によつて規定されるものと仮定する：

$$y_{ij} = \bar{Y} + U_i + w_{ij}$$

ここに U_i, W_{ij} はいずれも確率変数で、互に独立であるものとし

$$E\{U_i\} = 0, \quad E\{W_{ij}\} = 0$$

とする。なお U_i の分散を S_u^2 、 W_{ij} の分散を S_w^2 で表わす。

U_i は第 1 次抽出単位からの寄与を表わす成分で、それは第 i 抽出単位内では一定の値をとるものであり、 W_{ij} は第 i 抽出単位内で、第 j 要素からの寄与を表わす成分である。

N および M は、いずれも無限に大きいと仮定し、したがつて第 1 段階の抽出も、第 2 段階の抽出もいずれも無限集団からのランダム抽出であると仮定する。

しかるとき、 $\bar{Y}$ の推定量として $\bar{y}$ をとるとき、それが次の性質をもつことを証明しよう。

定理 11.20 構造 (102) を仮定するとき、 $\bar{y}$ は $\bar{Y}$ の不偏推定量である。

証明 (102) を用いて、 $\bar{y}$ を次のように表わしておく：

$$\begin{aligned} \bar{y} &= \frac{1}{n} \sum_{i=1}^n \bar{y}_i = \frac{1}{n} \sum_{i=1}^n \frac{1}{m} \sum_{j=1}^m y_{ij} \\ &= \frac{1}{n} \sum_{i=1}^n \frac{1}{m} \sum_{j=1}^m (\bar{Y} + U_i + w_{ij}). \end{aligned} \quad (104)$$

すなわち

$$\bar{y} = \bar{Y} + \frac{1}{n} \sum_{i=1}^n U_i + \frac{1}{nm} \sum_{i=1}^n \sum_{j=1}^m w_{ij} \quad (105)$$

ここで両辺の期待値をとれば (103) を用いて

$$E\{\bar{y}\} = \bar{Y}$$

が成り立つ。(終)

次に、 $\bar{y}$ の分散を与える公式を示そう。

定理 11.21 構造 (102) を仮定するとき、推定量 $\bar{y}$ の分散 $V(\bar{y})$ は次式で与えられる：

$$V(\bar{y}) = \frac{Su^2}{n} + \frac{Sw^2}{nm} \quad (106)$$

証明 (102) を用いて、 $\bar{y}$ が (105) のように書けることを知っている。したがつて、分散に関する定理 3.3 によつて (106) が成り立つ。(終)

この定理 11.21 から次のことが知られる。すなわち、 m を大きくするときは、 $\bar{y}$ の分散に対する Sm^2 からの影響だけを減少し、 n を大きくするときは、 Su^2 からの影響も、 Sw^2 からの影響もともに減少する。

定理 11.22 分散 $V(\bar{y})$ の不偏推定量 $v(\bar{y})$ は次式で与えられる：

$$v(\bar{y}) = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{m^2 n (n-1)} = \frac{\sum_{i=1}^n (\bar{y}_i - \bar{\bar{y}})^2}{n(n-1)}. \quad (107)$$

証明 次の恒等式が成り立つ：

$$\sum_{i=1}^n (\bar{y}_i - \bar{\bar{y}})^2 = \sum_{i=1}^n (\bar{y}_i - \bar{Y})^2 - n(\bar{\bar{y}} - \bar{Y})^2 \quad (108)$$

しかして (102) を用いて、次の関係が計算される：

$$\bar{y}_i - \bar{\bar{y}} = u_i + \frac{1}{m} \sum_{j=1}^m w_{ij} \quad (109)$$

したがつて

$$E\{(\bar{y}_i - \bar{\bar{y}})^2\} = Su^2 + \frac{Sw^2}{m}. \quad (110)$$

よつて

$$E\left\{\sum_{i=1}^n (\bar{y}_i - \bar{\bar{y}})^2\right\} = nSu^2 + \frac{n}{m} Sw^2.$$

また定理 11.21 により

$$nE\{(\bar{y} - \bar{\bar{y}})^2\} = nV(\bar{y}) = Su^2 + \frac{1}{m} Sw^2. \quad (112)$$

したがつて (108), (111) および (112) により

$$E\left\{\sum_{i=1}^n (\bar{y}_i - \bar{\bar{y}})^2\right\} = (n-1) \left(Su^2 + \frac{Sw^2}{m}\right) = n(n-1)V(\bar{y}) \quad (113)$$

すなわち

附 録 2 問 題 集

第 1 章

1. 単一測定では有用な情報が非常に少ないという例を三つあげよ
2. 次の確率を計算せよ
 - (a) 貨幣を一回投げて表が出る確率
 - (b) 2枚の貨幣を投げて二つとも表が出る確率
3. 表と裏がでるということに関して、貨幣の『真実性』ということの考えを得るためにはどのように実験を行えばよいかを示せ。
4. 全部で52枚のトランプの中から3枚ぬいて次のカードを得る確率を計算せよ。
 - (a) エース、ジャック、ジョース(2の札)
 - (b) エース、ジャック、スペードの2。
5. 1組のトランプから5枚のカードをぬいて、1そろいのエース、キング、クイーン、ジャックおよび10をぬく確率はどうか?
6. 次の順列および組合せを計算せよ。
 - (a) 15のものから同時に3つをとる順列。
 - (b) 6つのものから同時に6つをとる順列。
 - (c) 15のものから同時に3つをとる組合せ。
 - (d) 5つのものから同時に5つをとる組合せ。
7. 16人のグループから5人の委員をえらぶには幾通りの方法があるか?
8. 一度に8枚の貨幣を投げた場合、表になる枚数は0, 1, 2, …………… 8である。0, 1, 2, または3枚の表が出る確率を決定せよ。

第 2 章

1. 2-2表の級の値に自然数1, 2, 3, …………… 16を代入して、平均値および標準偏差を計算せよ。公式(14)および(15)を用いて計算したものを実際の値に変換せよ
 $\bar{X}' = 7.852, \quad \bar{X} = 7.902 \quad S' = 2.578, \quad S = 2.578$
2. 2-3表には大麦の400平方ヤードのプロットのグラムで表わした数量が与えてある。これらの収量に対して度数表およびヒストグラムを作れ。クラス間隔は11とし、最初のクラスは14~24とせよ。
 作った度数表から平均値および標準偏差を計算せよ。
 $\bar{X}' = 13.06 \quad \bar{X} = 15.169 \quad S' = 2.862 \quad S = 3.149$

$$E \left\{ \frac{\sum_{i=1}^n (\bar{y}_i - \bar{\bar{y}})^2}{n(n-1)} \right\} = V(\bar{\bar{y}})$$

よつて(107)で与えられる $v(\bar{\bar{y}})$ は、分散 $V(\bar{\bar{y}})$ の不偏推定量である。実際計算のためには、(107)の第1式が便利である。(終)

構造(102)を仮定しない場合にも、定理11.21とほとんど類似の結果が得られるのであるが、それについては省略する。

標本調査論においては、一定経費のもとで最も良い精度の推定を得るために、あるいは、一定の精度で推定するのに、最小の費用ですむようにするために、ここで述べた抽出法以外にも各種の抽出法が研究されている。

3. 次の等式を証明せよ

$$\sum(x^2) = \sum(X^2) - \frac{(\sum X)^2}{N}$$

2-3表

大麦の400平方ヤードプロットのグラムで表わした収量

185	162	136	157	141	130	129	176	171	190	157	147	176	126	175	134	169	189	180	128
169	205	129	117	144	125	165	170	153	186	164	123	165	203	156	182	164	176	176	150
216	154	184	203	166	155	215	190	164	204	194	148	162	146	174	185	171	181	158	147
165	157	180	165	127	186	133	170	134	177	109	169	128	152	165	139	146	144	178	188
133	128	161	160	167	156	125	162	128	103	116	87	123	143	130	119	141	174	157	168
195	180	158	139	139	168	145	166	118	171	143	132	126	171	176	115	165	147	186	157
187	174	172	191	155	169	139	144	130	146	159	164	160	122	175	156	119	135	116	134
157	182	209	136	153	160	142	179	125	149	171	186	196	175	189	214	169	166	164	195
189	108	118	149	178	171	151	192	127	148	158	174	191	134	188	248	164	206	185	192
147	178	189	141	173	187	167	128	139	152	167	131	203	231	214	177	161	194	141	161
124	130	112	122	192	155	196	179	166	156	131	179	201	122	207	189	164	131	211	172
170	140	156	199	181	181	150	184	154	200	187	169	155	107	143	145	190	176	162	123
189	194	146	22	160	107	70	84	112	162	124	156	138	101	138	141	143	135	163	183
99	118	150	151	83	136	171	191	155	164	98	136	115	168	130	111	136	129	122	120
179	172	192	171	151	142	193	174	146	180	140	137	138	194	109	120	124	126	126	147
115	148	195	154	149	139	163	118	126	127	139	174	167	175	179	172	174	167	142	169
122	163	144	147	123	160	137	161	122	101	158	103	119	164	112	57	94	106	132	122
164	142	155	147	115	143	68	184	183	167	160	138	191	133	160	156	122	111	153	148
103	131	180	142	191	175	146	181	111	110	154	176	168	175	175	146	148	167	106	123
121	154	148	91	93	74	113	79	132	119	96	80	97	98	106	107	69	86	94	129

第 3 章

1. 第2章の問題(2)の度数分布に対して縦軸(Y)および理論正規度数を計算せよ。理論度数を加えると計算のチェックができるであろう。

2. 問題1の実際度数に対してヒストグラムを作り、これをプロットして縦軸の端点を結び平滑な曲線を描け。

3. 3節の方程式 (1) を

$$Y = C e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$$

に等しいとき、曲線の型にσがどのように影響するかを示せ。

4. 母平均 21.65, σ=3.21であれば、無作為にとられた変数が28.55よりも大きく14.75より小さくなる確率を見出せ。 P=0.0316

5. 問題4の母集団で、400個の変量の標本平均の標準偏差がσ/√400であれば、無作為にとられた標本平均が21.33と21.97の限界の外におちる確率を見出せ。 P=0.045

6. 正規性からのかけはなれに対して次の分布を検定せよ。

X	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
f	1	57	185	217	177	126	87	54	30	20	14	11	13	5	1	2

X	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
f	2	3	4	5	7	11	17	30	50	34	21	10	7	5	2	2

X	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
f	1	7	13	19	23	26	27	28	26	24	22	17	14	9	4	1

(a) $g_1 = 1.360, g_2 = 2.143$ (b) $g_1 = -0.327, g_2 = 0.939$
(c) $g_1 = 0.107, g_2 = -0.766$

7. 20%が黒で、残りが白であるような豆の袋があると仮定しよう。8個の豆の単一標本を抽出して6個以上が白である確率を決定せよ。

P = 0.7969

8. 1000粒の小麦の種子の標本を種子供給所から抽出し落穂型植物の割合を調べるために栽培した。供給所が1000植物当たり正確に3つの落穂型植物を含んでいるとき、3つの落穂型がこの試験において見出される確率を決定せよ。

9. 3-3表は10個の無作為な数字を加えた1000の総計に対する度数表である。45を仮想平均として平均値および標準偏差を計算せよ。

3 - 3 表

10の数字の1000総計に対する度数表

級 間 隔	度 数	級 間 隔	度 数	級 間 隔	度 数
14~16	2	35~37	77	56~58	64
17~19	1	38~40	86	59~61	40
20~22	3	41~43	130	62~64	18
23~25	9	44~46	137	65~67	9
26~28	16	47~49	119	68~70	5
29~31	37	50~52	104	71~73	1
32~34	56	53~55	86		

1. 下の数字は2つの地区で生育した同じ小麦の品種の蛋白質の試験結果である。地区1の平均は12.74で地区2の平均は13.03である。2平均値間の差の有意性を検定せよ。この例で5節および6節で述べたもっと正確な検定を行なう必要があると考えるか?

蛋白質の結果

地区1	12.6	13.4	11.9	12.8	13.0		
地区2	13.1	13.4	12.8	13.5	13.3	12.7	12.4

$$t = 1.04, P = 0.3 \text{ (近似的)}$$

2. Mitchellは豚について骨格形成に対する石灰岩と骨粉との対にされた、資料実験を行った。その結果は下の4-2表に与えてある。

4 - 2 表

石灰岩と骨粉で飼育した豚の対のけんこう骨の百分率で表わした灰含有量

対	石灰岩	骨粉
1	49.2	51.5
2	53.3	54.9
3	50.6	52.2
4	52.0	53.3
5	46.8	51.6
6	50.5	54.1
7	52.1	54.2
8	53.0	53.3
平均	50.90	53.14

次の二通りの方法で平均値間の差の有意性を決定せよ。

(1) 値が対であると仮定した場合 (2) 値が対でないと仮定した場合。その結果を基として対にする効果を論ぜよ。

(1) 対の場合 : $t = 4.42, p = 0.01$ 以下

(2) 対でない場合 : $t = 2.48, p =$ 近似的に 0.02

3. 小麦の品種試験において、2つの品種間の平均差はエーカー当り4.5ブッシェルであることがわかった。標準誤差(Smd)はエーカー当り1.5ブッシェルであった。平均差に対する5%点での信頼限界を決定せよ。

注意。非常に多くの試験を行ったので、 t は1.96とすることができる。

信頼限界。 1.56~7.44

1. CramptonおよびHopkinsの資料からとった5-4表は比較できる資料をやった豚の体重の増加量を示す。豚の5ロットは5つの異った処理を表わす。そして各ロットには10匹の豚がいる。資料の分散分析を行って、処理の差の有意性を検定せよ。

5 - 4 表

比較できる飼料をやった場合の豚の体重の増加量

反復	ロットI	ロットII	ロットIII	ロットIV	ロットV
1	165	168	164	185	201
2	156	180	156	195	189
3	159	180	189	186	173
4	167	166	138	201	193
5	170	170	153	165	164
6	146	161	190	175	160
7	130	171	160	187	200
8	151	169	172	177	142
9	164	179	142	166	184
10	158	191	155	165	149

この実験での誤差の平均平方は計算すると243.6である。

2. 穀類の品種の農耕試験においては、考えている地域の色々な場所で試行を行い、2年以上の期間に対してそれを実行することが望ましい。Inner等は2年以上にわたってミネソタ州の二三の場処で大麦についての資料をとった。5-5表は6品種に対する2年間の3つの場処の収量を示す。この結果を分析せよ。

ブロックは1, 2, 3と番号がつけてあるが、University農場のブロック1はWasecaや他の場処のブロック1と関係があるという意味ではないことを注意しておく。したがって、ブロックに対する平方和および自由度は場処毎に計算して、最終分析で一緒にまとめればよい。この種の試験で、初心者が自由度を分ける際におちいる共通の誤まりはブロックを3水準であられる因子として考えることである。したがって彼等は分析においては次のような表わし方をする。

ブロック×場処

ブロック×年

ブロック×場処×年

等

これらの表わし方は、ブロックの数がすべての場処で一様であるような一定の水準を表わさないから意味はない。したがって正しい方法は冬実験に対してブロックの平方和を計算し、最終

の分析を示すためにすべての平方和を一緒に加えることである。

5 - 5 表

2年の各々で3つの場所で生育した大麦の6品種のエーカー当り
ブッシェルで表わした数量

プロット	Manchuria	Glabron	Svansota	Velvet	Tvebi	Peatland	場所	年
1	29.2	44.6	33.9	36.7	41.2	38.5	University	
2	25.0	39.1	39.4	41.0	31.9	29.6	農場	1931
3	26.8	45.5	32.1	42.0	36.6	30.2		
1	19.7	28.6	20.1	20.3	19.3	22.3	University	
2	31.4	38.3	30.8	27.5	22.4	30.8	農場	1932
3	29.6	43.5	31.4	32.6	44.5	31.1		
1	47.5	55.4	44.5	56.9	63.9	41.2		
2	52.2	53.4	46.0	40.6	63.8	51.5	Waseca	1931
3	46.9	56.8	51.5	53.2	63.8	53.0		
1	40.8	44.4	41.0	44.6	53.5	39.8		
2	29.4	34.9	41.1	41.4	44.2	39.2	Waseca	1932
3	30.2	33.9	33.4	26.2	50.0	29.1		
1	24.0	27.5	26.5	27.2	42.1	24.7		
2	24.7	25.5	21.5	28.0	42.5	29.5	Morris	1931
3	33.6	33.3	29.3	23.2	46.7	35.4		
1	29.6	36.6	27.1	35.9	40.4	35.7		
2	34.1	34.3	35.7	33.9	46.9	41.9	Morris	1932
3	39.4	34.5	42.3	46.7	53.0	52.0		

次の平方和の値は計算の際のチェックとなる。

全 体	11,504.61
品 種	1,566.58
品種×場所×年	230.52

3. $\sum_1^n (x^2) = \sum_1^n (X^2) - \frac{(\sum_1^n X)^2}{n}$ なることを説明せよ。

n_1 および n_2 の次の値に対してFの5%点を見出せ。

n_1	n_2	n_1	n_2
3	51	16	39
6	43	18	215
4	92	17	19
12	195	36	28
7	36	28	154
11	64	53	42

4. 5-6表はJonesがはりがね虫を数えた資料の一部である。資料を $\sqrt{1/2 + X}$ に変換し、分散分析を行え。処理に対する平方値を計算してから、数に再び変換せよ。平方根は小数以下2位まで正確に用いよ。

5 - 6 表

3反復の実験での5処理に対するはりがねむしの数

P	O	N	K	M
3	2	5	1	4
M	K	O	N	P
6	0	6	4	4
O	M	K	P	N
4	9	1	6	5

処理の平均平方 = 1.0133

第 6 章

1. 6-3表は種子のまきつけから蒔つみまでの日数およびまきつけから成熟までの日数について小麦の25品種の実験で得た結果である。

成熟までの日数の蒔つみまでの日数への回帰の回帰係数を計算し、回帰係数の有意性を検定せよ。計算を始める前に蒔つみまでの回数から50を、成熟までの日数から85を差引いて資料を簡略化せよ。回帰係数の5%点での信頼限界を求め、この例に与えた資料に基づいて、成熟までの日数の代りに蒔つみまでの日数を代用してよいかどうかを決定せよ。

$$\text{回帰係数} = \frac{105.23}{125.68} = 0.8373 \text{ (簡略資料)}$$

6 - 3 表

25の小麦品種の蒔きまでの日数と成熟までの日数に対する資料

品 種	蒔きまでの日数	成熟までの日数	品 種	蒔きまでの日数	成熟までの日数
1	60.0	94.4	14	58.2	92.4
2	53.6	89.0	15	58.0	91.6
3	59.0	94.0	16	59.4	94.0
4	61.8	95.4	17	55.4	90.8
5	53.8	88.2	18	61.6	95.2
6	57.8	93.4	19	63.0	97.2
7	57.8	93.6	20	60.2	94.6
8	58.4	92.0	21	61.6	96.0
9	57.8	92.8	22	57.6	92.6
10	59.0	93.4	23	60.8	95.4
11	59.2	93.8	24	61.2	94.4
12	59.0	92.8	25	58.2	94.0
13	58.6	94.2			

2. 6-4表は139の小麦品種に対して二つの方法で決定したカロチン含有量の資料である。一つの方法は、小麦全体についてカロチンを決定し、もう一つの方法では花粉についてカロチンを決定したものである。小麦のカロチンの数字は粉のカロチンよりも小さい。これはもちろん実際の状態とは逆である。これは小麦全体について小さいけれども適切な結果を与える別の抽出法を行ったことによる。

値の各対に対して一枚のカードを作り花粉のカロチンを従属変数Yとして、回帰表を準備せよ。計算労力を減ずるため、かなり大きな級をつくる。例えばXに対しての第1の級は0.85 ~ 0.95、Yに対しての第1の級は1.33 ~ 1.49とせよ。また計算を進める前に、1から始まる自然数に実際の級の値をおきかえることを忘れてはならない。回帰方程式を決定し、回帰のグラフを描け。

$$b_{yx} = 438.39 / 665.96 = 0.6583 \quad (\text{簡略資料})$$

3. 次を証明せよ

$$(a) \quad \sum xy = \sum XY - \frac{\sum X \sum Y}{n}$$

$$(b) \quad \sum (Y - Y_e)^2 = \sum y^2 - b_y^2 \sum x^2 \\ = \sum y^2 - b \sum xy$$

6 - 4 表
139品種に対する花粉および全体の小麦のカロチン含有量

品種	花粉のカロチン	小麦のカロチン	品種	花粉のカロチン	小麦のカロチン	品種	花粉のカロチン	小麦のカロチン
1	2.39	1.18	48	1.71	1.16	95	1.97	1.33
2	3.11	2.13	49	1.93	1.14	96	1.83	1.14
3	2.15	1.41	50	1.81	1.30	97	2.00	1.51
4	1.96	1.42	51	1.89	1.32	98	1.96	1.28
5	2.02	1.50	52	1.65	1.32	99	2.00	1.33
6	1.76	1.25	53	1.93	1.28	100	2.02	1.32
7	2.10	1.65	54	2.12	1.48	101	1.78	1.17
8	2.12	1.24	55	2.25	1.50	102	1.83	1.10
9	2.28	1.48	56	1.92	1.42	103	1.93	1.22
10	1.86	1.35	57	2.25	1.66	104	2.14	1.44
11	2.60	1.58	58	2.25	1.63	105	2.15	1.54
12	2.11	1.45	59	1.65	1.18	106	2.13	1.46
13	2.30	1.74	60	1.63	1.14	107	1.97	1.40
14	1.80	1.42	61	1.70	1.22	108	1.83	1.11
15	2.00	1.45	62	1.61	1.20	109	2.10	1.40
16	2.05	1.87	63	1.83	1.33	110	1.84	1.19
17	2.09	2.00	64	1.60	1.13	111	1.98	1.39
18	2.33	1.65	65	1.37	0.92	112	2.31	1.60
19	2.29	1.64	66	1.96	1.20	113	2.29	1.53
20	2.30	1.62	67	1.88	1.26	114	2.15	1.45
21	1.97	1.55	68	1.92	1.34	115	1.96	1.44
22	2.36	1.68	69	1.89	1.04	116	1.98	1.40
23	1.73	1.32	70	1.99	1.26	117	1.89	1.30
24	1.72	1.47	71	1.82	0.98	118	2.08	1.33
25	1.70	1.53	72	2.12	1.31	119	2.00	1.42
26	1.63	1.50	73	2.16	1.16	120	2.06	1.44
27	1.93	1.48	74	2.14	1.04	121	1.96	1.36
28	1.50	1.25	75	1.63	0.88	122	2.07	1.38
29	1.77	1.33	76	2.76	1.91	123	2.24	1.51
30	1.60	1.40	77	2.07	1.20	124	2.15	1.38
31	2.31	1.49	78	1.67	1.07	125	1.83	1.18
32	2.17	1.42	79	2.78	1.80	126	1.84	1.20
33	2.10	1.35	80	3.40	2.02	127	2.03	1.45
34	2.90	1.58	81	3.67	2.10	128	1.87	1.05
35	2.17	1.50	82	2.41	1.61	129	2.24	1.44
36	2.15	1.40	83	2.23	1.38	130	2.14	1.06
37	2.01	1.40	84	3.07	1.93	131	2.13	1.10
38	2.35	1.67	85	2.22	1.44	132	2.03	0.98
39	2.34	1.62	86	2.55	1.58	133	2.25	1.31
40	2.00	1.47	87	2.12	1.39	134	2.33	1.08
41	2.18	1.55	88	1.94	1.27	135	2.01	1.14
42	2.47	1.73	89	1.95	1.41	136	1.89	1.41
43	2.25	1.62	90	1.59	1.08	137	3.00	2.20
44	1.77	1.39	91	2.00	1.30	138	2.16	1.73
45	1.68	1.34	92	1.77	1.22	139	2.29	1.61
46	2.46	1.29	93	1.98	1.26			
47	1.86	1.28	94	1.97	1.30			

第 7 章

1. 7-2表の数字は, Manitoba大学の家庭経済学部(Home Economics)の学生の物理と英語の点数である。二つの学科の点数間の関係に対して相関係数を決定せよ。直接法で計算し, 係数の有意性を検定せよ。

$$r = +0.634 \quad t = 5.67$$

7 - 2 表

Manitoba 大学の家庭経済学部の50人の学生の物理と英語の点数

学生	物理	英語	学生	物理	英語	学生	物理	英語
1	20	21	18	26	29	35	23	26
2	25	26	19	24	27	36	26	27
3	24	27	20	19	26	37	22	21
4	22	24	21	25	25	38	26	25
5	27	27	22	18	20	39	23	21
6	26	28	23	20	24	40	29	26
7	26	24	24	23	24	41	23	19
8	32	26	25	22	20	42	20	19
9	22	24	26	23	26	43	23	30
10	27	26	27	31	27	44	33	32
11	22	23	28	24	25	45	21	19
12	22	25	29	28	30	46	28	30
13	22	24	30	25	28	47	24	21
14	29	30	31	26	32	48	26	28
15	26	30	32	24	25	49	24	22
16	24	28	33	27	30	50	30	25
17	25	29	34	28	30			

2. 7-3表の資料に対して, 25の小麦品種の蒔きまでの日数と成熟までの日数に対する相関係数を決定せよ。この係数に対して5%点で信頼限界を見出せ。

$$r = +0.946, \quad l_1 = 0.916 \quad l_2 = 0.972$$

7 - 3 表

小麦の25品種の蒔きまでの日数と成熟までの日数

品種	蒔きまでの日数	成熟までの日数	品種	蒔きまでの日数	成熟までの日数
1	60.0	94.4	14	58.2	92.4
2	53.6	89.0	15	58.0	91.6
3	59.0	94.0	16	59.4	94.0
4	61.8	95.4	17	55.4	90.8
5	53.8	88.2	18	61.6	95.2
6	57.8	93.4	19	63.0	97.2
7	57.8	93.6	20	60.2	94.6
8	58.4	92.0	21	61.6	96.0
9	57.8	92.8	22	57.6	92.6
10	59.0	93.4	23	60.8	95.4
11	59.2	93.8	24	61.2	94.4
12	59.0	92.8	25	58.2	94.0
13	58.6	94.2			

第 8 章

1. R. A. Fisherは稈度(X_1), 稈度(X_2)および高度(X_3)の, 収量 Y に対する影響についての次の資料を与えている。

$$\sum y^2 = 1786.6$$

$$\sum(xy) = 1137.4 \quad \sum(x_1)^2 = 1934.1 \quad \sum(x_1x_2) = -772.2$$

$$\sum(xy) = 592.9 \quad \sum(x_2)^2 = 2889.5 \quad \sum(x_1x_3) = 924.1$$

$$\sum(xy) = 891.8 \quad \sum(x_3)^2 = 1750.8 \quad \sum(x_2x_3) = 119.6$$

$$b_{y1.23} = 0.3962 \quad b_{y2.13} = -0.1120 \quad b_{y3.11} = 0.3079$$

これらの資料に対して, (a) 例8-1の直接法および(b) 例8-2の乗数のマトリックスを求めることによって偏回帰係数を計算せよ。

2. 公式(10)を用いて, 問題1で得た偏回帰係数の有意性を検定せよ。 b_3 の信頼限界を決定する方法を工夫せよ。

3. 例8-2の b_1 および b_2 間の差の有意性を検定せよ。

4. 問題1の資料から X_1 に Y を, X_2 に X_1 を, X_3 に X_2 を, X_4 に X_3 を代入して, 相関係数 r_{12} , r_{13} , r_{14} , r_{23} , r_{24} , r_{34} を得よ。それから, 例8-4のようにして偏相関係数について

解け。

5. 問題4から重相関係数 $r_{1.234}$ を計算し、その有意性を検定せよ。
6. $n' = 40$ に対して、丁度有意である重相関 $r_{1.234}$ を決定せよ。
7. 下に与えた例で、重相関の計算によって情報の増加の有意性を検定せよ。各比較に対して、言葉で諸君の結論を述べよ。

$n' = 40$	$r_{12} = 0.7643$	$r_{1.234} = 0.8031$
$n' = 62$	$r_{12} = 0.8744$	$r_{1.2345} = 0.9664$
$n' = 20$	$r_{12} = 0.7621$	$r_{1.23} = 0.7635$
$n' = 20$	$r_{12} = 0.7316$	$r_{1.23456} = 0.7329$

第 9 章

1. 9-5表の資料はEden と Fisherによって与えられた8つの肥料処理および各8反復に対する穀粒とわらの収量である。共分散分析で附属変量としてわらの重さを取り、補正した穀粒収量の有意性を検定せよ。得た結果および資料の他の特徴から、穀粒のわら収量による補正の適合性を決定せよ。

9 - 5 表

8つの異なる肥料処理と各8反復の64プロットに対する平均穀粒とわらの収量

処 理	ブロック1		ブロック2		ブロック3		ブロック4	
	わ	ら	わ	ら	わ	ら	わ	ら
A	242	620	321	646	261	681	317	644
B	267	644	382	745	201	542	316	711
C	215	523	330	713	298	686	381	688
D	212	601	292	693	265	685	255	714
E	322	664	370	693	284	666	323	516
F	200	514	261	637	259	697	361	710
G	260	550	318	708	266	663	340	673
H	203	521	275	661	207	594	331	730

処 理	ブロック5		ブロック6		ブロック7		ブロック8	
	わ	ら	わ	ら	わ	ら	わ	ら
A	255	706	331	615	216	552	295	726
B	280	705	285	637	200	543	309	646
C	300	692	294	612	256	635	284	748
D	238	699	309	697	283	701	324	746
E	232	656	393	663	351	657	363	683
F	234	633	258	595	306	697	376	712
G	362	671	400	626	276	655	385	671
H	229	625	266	644	276	745	328	747

第 10 章

1. 適合の各段階で検査することによって最良の適合曲線を決定しながら例10-3の方法で10-11表の資料に直交多項式を適合せよ。同じ表で多項値を計算し、実際値と共に図表を明け。

2. 10-12表はJulia Bellの研究で、遺伝的な筋肉のdystrophyについて得られたものである。各対に対して2つの記入事項があるから、これは遺伝研究でよく用いられる対称表である。表のようにしてYおよびXをとり、Y列の平均を計算し、図に描け。列10-1のようにして、回帰の一次、二次および三次の成分を検定する段階にまで三次の多項式の適合を進めよ。

10 - 11 表

2つの型の食物で飼育したChoristoneura fomiferanae(Clem)

(もみの害虫)の幼虫の生長割合の差

Instar	2	3	4	5	6	7
差	6	20	63	75	64	39

10 - 12 表

実のいとこの対の筋肉 dystrophy の関係年令 (Bell)

	10	68	28	13	7	4	4	134
60~69		1		2	1	1		4
50~59		1	1	1			1	4
40~49			3	3			1	7
30~39			3	4	3	1	2	13
20~29		9	12	3	3	1		28
10~19	2	56	9			1		68
0~9	8	2						10

0~9 10~19 20~29 30~39 40~49 50~59 60~69

X の 年 齢

平均平方: 二次 = 13.653

三次 = 0.203

残差 = 0.930

附 録 1.

1. 13,266世帯から成る某市の人口を推定したいために、30世帯をランダムに抽出し、その各世帯の人員数を調べて次の結果が得られた。

5, 6, 3, 3, 2, 3, 3, 4, 4, 2, 6, 4, 3, 1, 5,
6, 2, 4, 3, 1, 2, 8, 4, 3, 5, 6, 1, 4, 2, 5,

この結果からその市の人口の推定値を求めよ。

2. 大学入試の合格者500人の中で、幾パーセントが浪人生活の後合格したものであるかを知りたいために、100人をランダムに抽出して調査した。その結果37人が浪人生活の後合格していることが分った。この調査結果から、95%の信頼係数をもって、浪人生活の後合格したものの比率を推定せよ。

3. 層化抽出法において、層から大きさ nh , ($h=1, 2, \dots, L$)の標本をとるとき、総費用 C が次式で与えられるとする：

$$C = a + \sum c_h n_h$$

しかるとき、与えられた総費用 C のもとで、母平均 $\bar{Y}$ を推定するのに、推定の精度を最良にするためには、各層からとるべき標本の大きさ n_h は

$$\frac{N_h S_h}{\sqrt{c_h}}$$

に比例した大きさにとるべきであることを証明せよ。

この結果から次のことが知られる。すなわち与えられた層において

(1) 層が大きいほど、(2) 層内の変動が大きいほど、(3) 層から抽出の費用が安いほど
その層から大きい標本をとるべきである。

附 録 3 数 表

1. 偏差によって表わした正規曲線の面積 $\frac{1}{2} (1 + \alpha)$

	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8820
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998
3.6	.9998	.9998	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999
3.7	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999
3.8	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	1.0000

2. 偏差によって表わした正規曲線の高さ

	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.3989	.3989	.3989	.3988	.3986	.3984	.3982	.3980	.3977	.3973
0.1	.3970	.3965	.3961	.3956	.3951	.3945	.3939	.3932	.3925	.3918
0.2	.3910	.3902	.3894	.3885	.3876	.3867	.3857	.3847	.3836	.3825
0.3	.3814	.3802	.3790	.3778	.3765	.3752	.3739	.3725	.3712	.3697
0.4	.3683	.3668	.3653	.3637	.3621	.3605	.3589	.3572	.3555	.3538
0.5	.3521	.3503	.3485	.3467	.3448	.3429	.3410	.3391	.3372	.3352
0.6	.3332	.3312	.3292	.3271	.3251	.3230	.3209	.3187	.3166	.3144
0.7	.3123	.3101	.3079	.3056	.3034	.3011	.2989	.2966	.2943	.2920
0.8	.2897	.2874	.2850	.2827	.2803	.2780	.2756	.2732	.2709	.2685
0.9	.2661	.2637	.2613	.2589	.2565	.2541	.2516	.2492	.2468	.2444
1.0	.2420	.2396	.2371	.2347	.2323	.2299	.2275	.2251	.2227	.2203
1.1	.2179	.2155	.2131	.2107	.2083	.2059	.2036	.2012	.1989	.1965
1.2	.1942	.1919	.1895	.1872	.1849	.1826	.1804	.1781	.1758	.1736
1.3	.1714	.1691	.1669	.1647	.1626	.1604	.1582	.1561	.1539	.1518
1.4	.1497	.1476	.1456	.1435	.1415	.1394	.1374	.1354	.1334	.1315
1.5	.1295	.1276	.1257	.1238	.1219	.1200	.1182	.1163	.1145	.1127
1.6	.1109	.1092	.1074	.1057	.1040	.1023	.1006	.0989	.0973	.0957
1.7	.0940	.0925	.0909	.0893	.0878	.0863	.0848	.0833	.0818	.0804
1.8	.0790	.0775	.0761	.0748	.0734	.0721	.0707	.0694	.0681	.0669
1.9	.0656	.0644	.0632	.0620	.0608	.0596	.0584	.0573	.0562	.0551
2.0	.0540	.0529	.0519	.0508	.0498	.0488	.0478	.0468	.0459	.0449
2.1	.0440	.0431	.0422	.0413	.0404	.0400	.0387	.0379	.0371	.0363
2.2	.0355	.0347	.0339	.0332	.0325	.0317	.0310	.0303	.0297	.0290
2.3	.0283	.0277	.0270	.0264	.0258	.0252	.0246	.0241	.0235	.0229
2.4	.0244	.0219	.0213	.0208	.0203	.0198	.0194	.0189	.0184	.0180
2.5	.0175	.0171	.0167	.0163	.0158	.0154	.0151	.0147	.0143	.0139
2.6	.0135	.0132	.0129	.0126	.0122	.0119	.0116	.0113	.0110	.0107
2.7	.0104	.0101	.0099	.0096	.0093	.0091	.0088	.0086	.0084	.0081
2.8	.0079	.0077	.0075	.0073	.0071	.0069	.0067	.0065	.0063	.0061
2.9	.0060	.0058	.0056	.0055	.0053	.0051	.0050	.0048	.0047	.0046
3.0	.0044	.0043	.0042	.0040	.0039	.0038	.0037	.0036	.0035	.0034
3.1	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026	.0025	.0025
3.2	.0024	.0023	.0021	.0022	.0022	.0020	.0020	.0019	.0018	.0018
3.3	.0017	.0017	.0016	.0016	.0015	.0015	.0014	.0014	.0013	.0013
3.4	.0012	.0012	.0012	.0011	.0011	.0010	.0010	.0010	.0009	.0009
3.5	.0009	.0008	.0008	.0008	.0008	.0007	.0007	.0007	.0007	.0006
3.6	.0006	.0006	.0006	.0005	.0005	.0005	.0005	.0005	.0005	.0004
3.7	.0004	.0004	.0004	.0004	.0004	.0004	.0003	.0003	.0003	.0003
3.8	.0003	.0003	.0003	.0003	.0003	.0002	.0002	.0002	.0002	.0002
3.9	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0001	.0001
4.0	.0001									

3. t の 表

DF	0.50	0.10	0.05	0.02	0.01	0.001
1	1.000	6.31	12.71	31.82	63.66	63.7
2	0.816	2.92	4.30	6.96	9.92	31.6
3	.765	2.35	3.18	4.54	5.84	12.9
4	.741	2.13	2.78	3.75	4.60	8.61
5	.727	2.02	2.57	3.36	4.03	6.86
6	.718	1.94	2.45	3.14	3.71	5.96
7	.711	1.90	2.36	3.00	3.50	5.40
8	.706	1.86	2.31	2.90	3.36	5.04
9	.703	1.83	2.26	2.82	3.25	4.78
10	.700	1.81	2.23	2.76	3.17	4.59
11	.697	1.80	2.20	2.72	3.11	4.44
12	.695	1.78	2.18	2.68	3.06	4.32
13	.694	1.77	2.16	2.65	3.01	4.22
14	.692	1.76	2.14	2.62	2.98	4.14
15	.691	1.75	2.13	2.60	2.95	4.07
16	.690	1.75	2.12	2.58	2.92	4.02
17	.689	1.74	2.11	2.57	2.90	3.96
18	.688	1.73	2.10	2.55	2.88	3.92
19	.688	1.73	2.09	2.54	2.86	3.88
20	.687	1.72	2.09	2.53	2.84	3.85
21	.686	1.72	2.08	2.52	2.83	3.82
22	.686	1.72	2.07	2.51	2.82	3.79
23	.685	1.71	2.07	2.50	2.81	3.77
24	.685	1.71	2.06	2.49	2.80	3.74
25	.684	1.71	2.06	2.48	2.79	3.72
26	.684	1.71	2.06	2.48	2.78	3.71
27	.684	1.70	2.05	2.47	2.77	3.69
28	.683	1.70	2.05	2.47	2.76	3.67
29	.683	1.70	2.04	2.46	2.76	3.66
30	.683	1.70	2.04	2.46	2.75	3.65
35	.682	1.69	2.03	2.44	2.72	3.59
40	.681	1.68	2.02	2.42	2.71	3.55
45	.680	1.68	2.02	2.41	2.69	3.52
50	.679	1.68	2.01	2.40	2.68	3.50
60	.678	1.67	2.00	2.39	2.66	3.46
70	.678	1.67	2.00	2.38	2.65	3.44
80	.677	1.66	1.99	2.38	2.64	3.42
90	.677	1.66	1.99	2.37	2.63	3.40
100	.677	1.66	1.98	2.36	2.63	3.39
120	.676	1.66	1.98	2.36	2.62	3.37
150	.676	1.66	1.98	2.35	2.61	3.36
200	.675	1.65	1.97	2.35	2.60	3.34
300	.675	1.65	1.97	2.34	2.59	3.32
400	.675	1.65	1.97	2.34	2.59	3.32
500	.674	1.65	1.96	2.33	2.59	3.31
1000	.674	1.65	1.96	2.33	2.58	3.30
∞	.674	1.64	1.96	2.33	2.58	3.29

6. 乱 数 表

22	17	68	65	84	68	95	23	92	35	87	02	22	57	51	61	09	43	95	06	58	24	82	03	47
19	36	27	59	46	13	79	93	37	55	39	77	32	77	09	85	52	05	30	62	47	83	51	62	74
16	77	23	02	77	09	61	87	25	21	28	06	24	25	93	16	71	13	59	78	23	05	47	47	25
78	43	76	71	61	20	44	90	32	64	97	67	63	99	61	46	38	03	93	22	69	81	21	99	21
03	28	28	26	08	73	37	32	04	05	69	30	16	09	05	88	69	58	28	99	35	07	44	75	47
93	22	53	64	39	07	10	63	76	35	87	03	04	79	88	08	13	13	85	51	55	34	57	72	69
78	76	58	54	74	92	38	70	96	92	52	06	79	79	45	82	63	18	27	44	69	66	92	19	09
23	68	35	26	00	99	53	93	61	28	52	70	05	48	34	56	65	05	61	86	90	92	10	70	80
15	39	25	70	99	93	86	52	77	65	15	33	59	05	28	22	87	26	07	47	86	96	98	29	06
58	71	96	30	24	18	46	23	34	27	85	13	99	24	44	49	18	09	79	49	74	16	32	23	02
57	35	27	33	72	24	53	63	94	09	41	10	76	47	91	44	04	95	49	66	39	60	04	59	81
48	50	86	54	48	22	06	34	72	52	82	21	15	65	20	33	29	94	71	11	15	91	29	12	03
61	96	48	95	03	07	16	39	33	66	98	56	10	56	79	77	21	30	27	12	90	49	22	23	62
36	93	89	41	26	29	70	83	63	51	99	74	20	52	36	87	09	41	15	09	98	60	16	03	03
18	87	00	42	31	57	90	12	02	07	23	47	37	17	31	54	08	01	88	63	39	41	88	92	10
22	85	61	68	90	49	64	92	85	44	16	40	12	89	88	50	14	49	81	06	01	82	77	45	12
67	80	43	79	33	12	83	11	41	16	25	58	19	68	70	77	02	54	00	52	53	43	37	15	26
27	62	50	96	72	79	44	61	40	15	14	53	40	65	39	27	31	58	50	28	11	39	03	34	25
33	78	80	87	15	38	30	06	38	21	14	47	47	07	26	54	96	87	53	32	40	36	40	96	76
13	13	92	66	99	47	24	49	57	74	32	25	43	62	17	10	97	11	69	84	99	63	22	32	98

6. 乱 数 表

22	17	68	65	84	68	95	23	92	35	87	02	22	57	51	61	09	43	95	06	58	24	82	03	47
19	36	27	59	46	13	79	93	37	55	39	77	32	77	09	85	52	05	30	62	47	83	51	62	74
16	77	23	02	77	09	61	87	25	21	28	06	24	25	93	16	71	13	59	78	23	05	47	47	25
78	43	76	71	61	20	44	90	32	64	97	67	63	99	61	46	38	03	93	22	69	81	21	99	21
03	28	28	26	08	73	37	32	04	05	69	30	16	09	05	88	69	58	28	99	35	07	44	75	47
93	22	53	64	39	07	10	63	76	35	87	03	04	79	88	08	13	13	85	51	55	34	57	72	69
78	76	58	54	74	92	38	70	96	92	52	06	79	79	45	82	63	18	27	44	69	66	92	19	09
23	68	35	26	00	99	53	93	61	28	52	70	05	48	34	56	65	05	61	86	90	92	10	70	80
15	39	25	70	99	93	86	52	77	65	15	33	59	05	28	22	87	26	07	47	86	96	98	29	06
58	71	96	30	24	18	46	23	34	27	85	13	99	24	44	49	18	09	79	49	74	16	32	23	02
57	35	27	33	72	24	53	63	94	09	41	10	76	47	91	44	04	95	49	66	39	60	04	59	81
48	50	86	54	48	22	06	34	72	52	82	21	15	65	20	33	29	94	71	11	15	91	29	12	03
61	96	48	95	03	07	16	39	33	66	98	56	10	56	79	77	21	30	27	12	90	49	22	23	62
36	93	89	41	26	29	70	83	63	51	99	74	20	52	36	87	09	41	15	09	98	60	16	03	03
18	87	00	42	31	57	90	12	02	07	23	47	37	17	31	54	08	01	88	63	39	41	88	92	10
88	56	53	27	59	33	35	72	67	47	77	34	55	45	70	08	18	27	38	90	16	95	86	70	75
09	72	95	84	29	49	41	31	06	70	42	38	06	45	18	64	84	73	31	65	52	53	37	97	15
12	96	88	17	31	65	19	69	02	83	60	75	86	90	68	24	64	19	35	51	56	61	87	39	12
85	94	57	24	16	92	09	84	38	76	22	00	27	69	85	29	81	94	78	70	21	94	47	90	12
38	64	43	59	98	98	77	87	68	07	91	51	67	62	44	40	98	05	93	78	23	32	65	41	18
53	44	09	42	72	00	41	86	79	79	68	47	22	00	20	35	55	31	51	51	00	83	63	22	55
40	76	66	26	84	57	99	99	90	37	36	63	32	08	58	37	40	13	68	97	87	64	81	07	83
02	17	79	18	05	12	59	52	57	02	22	07	90	47	03	28	14	11	30	79	20	69	22	40	98
95	17	82	06	53	31	51	10	96	46	92	06	88	07	77	56	11	50	81	69	40	23	72	51	39
35	29	22	42	92	96	11	83	44	80	34	68	35	48	77	33	42	40	90	60	73	96	53	97	86
26	29	13	56	41	85	47	04	66	08	34	72	57	59	13	82	43	80	46	15	38	26	61	70	04
77	80	20	75	82	92	82	32	99	90	63	95	73	76	63	89	73	44	99	05	48	67	26	43	18
46	40	66	44	52	91	36	74	43	53	30	82	13	54	00	78	45	63	98	35	55	03	36	67	68
37	56	08	18	09	77	53	84	46	47	31	91	18	95	58	24	16	74	11	53	44	10	13	85	57
61	65	61	68	66	37	27	47	39	19	84	83	70	07	48	53	21	40	06	71	95	06	79	88	54
93	43	69	64	07	34	18	04	52	35	56	27	09	24	86	61	85	53	83	45	19	90	70	99	00
21	96	60	12	99	11	20	99	45	18	48	13	93	55	34	18	37	79	49	90	65	97	38	20	46
95	20	47	97	97	27	37	83	28	71	00	06	41	41	74	45	89	09	39	84	51	67	11	52	49
97	86	21	78	73	10	65	81	92	59	58	76	17	14	97	04	76	62	16	17	17	95	70	45	80
69	92	06	34	13	59	71	74	17	32	27	55	10	24	19	23	71	82	13	74	63	52	52	01	41
04	31	17	21	56	33	73	99	19	87	26	72	39	27	67	53	77	57	68	93	60	61	97	22	61
61	06	98	03	91	87	14	77	43	96	43	00	65	98	50	45	60	33	01	07	98	99	46	50	47
85	93	85	86	88	72	87	08	62	40	16	06	10	89	20	23	21	34	74	97	76	38	03	29	63
21	74	32	47	45	73	96	07	94	52	09	65	90	77	47	25	76	16	19	33	53	05	70	53	30
15	69	53	82	80	79	96	23	53	10	65	39	07	16	29	45	33	02	43	70	02	87	40	41	45
02	89	08	04	49	20	21	14	68	86	87	63	93	95	17	11	29	01	95	80	35	14	97	35	33
87	18	15	89	79	85	43	01	72	73	08	61	74	51	69	89	74	39	82	15	94	51	33	41	67
98	83	71	94	22	59	97	50	99	52	08	52	85	08	40	87	80	61	65	31	91	51	80	32	44
10	08	58	21	66	72	68	49	29	31	89	85	84	46	06	59	73	19	85	23	65	09	29	75	63
47	90	56	10	08	88	02	84	27	83	42	29	72	23	19	66	56	45	65	79	20	71	53	20	25
22	85	61	68	90	49	64	92	85	44	16	40	12	89	88	50	14	49	81	06	01	82	77	45	12
67	80	43	79	33	12	83	11	41	16	25	58	19	68	70	77	02	54	00	52	53	43	37	15	26
27	62	50	96	72	79	44	61	40	15	14	53	40	65	39	27	31	58	50	28	11	39	03	34	25
33	78	80	87	15	38	30	06	38	21	14	47	47	07	26	54	96	87	53	32	40	36	40	96	76
13	13	92	66	99	47	24	49	57	74	32	25	43	62	17	10	97	11	69	84	99	63	22	32	98

4. χ^2 の 表										
DF	0.99	0.95	0.50	0.30	0.20	0.10	0.05	0.02	0.01	0.001
1	0.0002	0.004	0.46	1.07	1.64	2.71	3.84	5.41	6.64	10.83
2	0.020	0.103	1.39	2.41	3.22	4.60	5.99	7.82	9.21	13.82
3	0.115	0.35	2.37	3.66	4.64	6.25	7.82	9.84	11.34	16.27
4	0.30	0.71	3.36	4.88	5.99	7.78	9.49	11.67	13.28	18.46
5	0.55	1.14	4.35	6.06	7.29	9.24	11.07	13.39	15.09	20.52
6	0.87	1.64	5.35	7.23	8.56	10.64	12.59	15.03	16.81	22.46
7	1.24	2.17	6.35	8.38	9.80	12.02	14.07	16.62	18.48	24.32
8	1.65	2.73	7.34	9.52	11.03	13.36	15.51	18.17	20.09	26.12
9	2.09	3.32	8.34	10.66	12.24	14.68	16.92	19.68	21.67	27.88
10	2.56	3.94	9.34	11.78	13.44	15.99	18.31	21.16	23.21	29.59
11	3.05	4.58	10.34	12.90	14.63	17.28	19.68	22.62	24.72	31.26
12	3.57	5.23	11.34	14.01	15.81	18.55	21.03	24.05	26.22	32.91
13	4.11	5.89	12.34	15.12	16.98	19.81	22.36	25.47	27.69	34.53
14	4.66	6.57	13.34	16.22	18.15	21.06	23.68	26.87	29.14	36.12
15	5.23	7.26	14.34	17.32	19.31	22.31	25.00	28.26	30.58	37.70
16	5.81	7.96	15.34	18.42	20.46	23.54	26.30	29.63	32.00	39.25
17	6.41	8.67	16.34	19.51	21.62	24.77	27.59	31.00	33.41	40.79
18	7.02	9.39	17.34	20.60	22.76	25.99	28.87	32.35	34.80	42.31
19	7.63	10.12	18.34	21.69	23.90	27.20	30.14	33.69	36.19	43.82
20	8.26	10.85	19.34	22.78	25.04	28.41	31.41	35.02	37.57	45.32
21	8.90	11.59	20.34	23.86	26.17	29.62	32.67	36.34	38.93	46.80
22	9.54	12.34	21.34	24.94	27.30	30.81	33.92	37.66	40.29	48.27
23	10.20	13.09	22.34	26.02	28.43	32.01	35.17	38.97	41.64	49.73
24	10.86	13.85	23.34	27.10	29.55	33.20	36.42	40.27	42.98	51.18
25	11.52	14.61	24.34	28.17	30.68	34.38	37.65	41.57	44.31	52.62
26	12.20	15.38	25.34	29.25	31.80	35.56	38.88	42.86	45.64	54.05
27	12.88	16.15	26.34	30.32	32.91	36.74	40.11	44.14	46.96	55.48
28	13.56	16.93	27.34	31.39	34.03	37.92	41.34	45.42	48.28	56.89
29	14.26	17.71	28.34	32.46	35.14	39.09	42.56	46.69	49.59	58.30
30	14.95	18.49	29.34	33.53	36.25	40.26	43.77	47.96	50.89	59.70

B F分布の5%点(上)と1%点(下)

大きい方の自由度に対する自由度

	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞
1	161.4	199.5	215.7	227.2	235.0	241.9	248.0	253.3	258.0	262.2	266.0	269.5	272.8	275.9	278.9	281.8	284.6	287.3	289.9	292.4	294.8	297.1	299.3	301.4
2	19.0	19.0	19.1	19.2	19.3	19.4	19.5	19.6	19.7	19.8	19.9	20.0	20.1	20.2	20.3	20.4	20.5	20.6	20.7	20.8	20.9	21.0	21.1	21.2
3	16.0	16.0	16.1	16.2	16.3	16.4	16.5	16.6	16.7	16.8	16.9	17.0	17.1	17.2	17.3	17.4	17.5	17.6	17.7	17.8	17.9	18.0	18.1	18.2
4	14.5	14.5	14.6	14.7	14.8	14.9	15.0	15.1	15.2	15.3	15.4	15.5	15.6	15.7	15.8	15.9	16.0	16.1	16.2	16.3	16.4	16.5	16.6	16.7
5	13.7	13.7	13.8	13.9	14.0	14.1	14.2	14.3	14.4	14.5	14.6	14.7	14.8	14.9	15.0	15.1	15.2	15.3	15.4	15.5	15.6	15.7	15.8	15.9
6	13.3	13.3	13.4	13.5	13.6	13.7	13.8	13.9	14.0	14.1	14.2	14.3	14.4	14.5	14.6	14.7	14.8	14.9	15.0	15.1	15.2	15.3	15.4	15.5
7	13.0	13.0	13.1	13.2	13.3	13.4	13.5	13.6	13.7	13.8	13.9	14.0	14.1	14.2	14.3	14.4	14.5	14.6	14.7	14.8	14.9	15.0	15.1	15.2
8	12.8	12.8	12.9	13.0	13.1	13.2	13.3	13.4	13.5	13.6	13.7	13.8	13.9	14.0	14.1	14.2	14.3	14.4	14.5	14.6	14.7	14.8	14.9	15.0
9	12.6	12.6	12.7	12.8	12.9	13.0	13.1	13.2	13.3	13.4	13.5	13.6	13.7	13.8	13.9	14.0	14.1	14.2	14.3	14.4	14.5	14.6	14.7	14.8
10	12.5	12.5	12.6	12.7	12.8	12.9	13.0	13.1	13.2	13.3	13.4	13.5	13.6	13.7	13.8	13.9	14.0	14.1	14.2	14.3	14.4	14.5	14.6	14.7
11	12.4	12.4	12.5	12.6	12.7	12.8	12.9	13.0	13.1	13.2	13.3	13.4	13.5	13.6	13.7	13.8	13.9	14.0	14.1	14.2	14.3	14.4	14.5	14.6
12	12.3	12.3	12.4	12.5	12.6	12.7	12.8	12.9	13.0	13.1	13.2	13.3	13.4	13.5	13.6	13.7	13.8	13.9	14.0	14.1	14.2	14.3	14.4	14.5
13	12.2	12.2	12.3	12.4	12.5	12.6	12.7	12.8	12.9	13.0	13.1	13.2	13.3	13.4	13.5	13.6	13.7	13.8	13.9	14.0	14.1	14.2	14.3	14.4
14	12.1	12.1	12.2	12.3	12.4	12.5	12.6	12.7	12.8	12.9	13.0	13.1	13.2	13.3	13.4	13.5	13.6	13.7	13.8	13.9	14.0	14.1	14.2	14.3
15	12.0	12.0	12.1	12.2	12.3	12.4	12.5	12.6	12.7	12.8	12.9	13.0	13.1	13.2	13.3	13.4	13.5	13.6	13.7	13.8	13.9	14.0	14.1	14.2
16	11.9	11.9	12.0	12.1	12.2	12.3	12.4	12.5	12.6	12.7	12.8	12.9	13.0	13.1	13.2	13.3	13.4	13.5	13.6	13.7	13.8	13.9	14.0	14.1
17	11.8	11.8	11.9	12.0	12.1	12.2	12.3	12.4	12.5	12.6	12.7	12.8	12.9	13.0	13.1	13.2	13.3	13.4	13.5	13.6	13.7	13.8	13.9	14.0
18	11.7	11.7	11.8	11.9	12.0	12.1	12.2	12.3	12.4	12.5	12.6	12.7	12.8	12.9	13.0	13.1	13.2	13.3	13.4	13.5	13.6	13.7	13.8	13.9
19	11.6	11.6	11.7	11.8	11.9	12.0	12.1	12.2	12.3	12.4	12.5	12.6	12.7	12.8	12.9	13.0	13.1	13.2	13.3	13.4	13.5	13.6	13.7	13.8
20	11.5	11.5	11.6	11.7	11.8	11.9	12.0	12.1	12.2	12.3	12.4	12.5	12.6	12.7	12.8	12.9	13.0	13.1	13.2	13.3	13.4	13.5	13.6	13.7
21	11.4	11.4	11.5	11.6	11.7	11.8	11.9	12.0	12.1	12.2	12.3	12.4	12.5	12.6	12.7	12.8	12.9	13.0	13.1	13.2	13.3	13.4	13.5	13.6
22	11.3	11.3	11.4	11.5	11.6	11.7	11.8	11.9	12.0	12.1	12.2	12.3	12.4	12.5	12.6	12.7	12.8	12.9	13.0	13.1	13.2	13.3	13.4	13.5
23	11.2	11.2	11.3	11.4	11.5	11.6	11.7	11.8	11.9	12.0	12.1	12.2	12.3	12.4	12.5	12.6	12.7	12.8	12.9	13.0	13.1	13.2	13.3	13.4
24	11.1	11.1	11.2	11.3	11.4	11.5	11.6	11.7	11.8	11.9	12.0	12.1	12.2	12.3	12.4	12.5	12.6	12.7	12.8	12.9	13.0	13.1	13.2	13.3
25	11.0	11.0	11.1	11.2	11.3	11.4	11.5	11.6	11.7	11.8	11.9	12.0	12.1	12.2	12.3	12.4	12.5	12.6	12.7	12.8	12.9	13.0	13.1	13.2
26	10.9	10.9	11.0	11.1	11.2	11.3	11.4	11.5	11.6	11.7	11.8	11.9	12.0	12.1	12.2	12.3	12.4	12.5	12.6	12.7	12.8	12.9	13.0	13.1
27	10.8	10.8	10.9	11.0	11.1	11.2	11.3	11.4	11.5	11.6	11.7	11.8	11.9	12.0	12.1	12.2	12.3	12.4	12.5	12.6	12.7	12.8	12.9	13.0
28	10.7	10.7	10.8	10.9	11.0	11.1	11.2	11.3	11.4	11.5	11.6	11.7	11.8	11.9	12.0	12.1	12.2	12.3	12.4	12.5	12.6	12.7	12.8	12.9
29	10.6	10.6	10.7	10.8	10.9	11.0	11.1	11.2	11.3	11.4	11.5	11.6	11.7	11.8	11.9	12.0	12.1	12.2	12.3	12.4	12.5	12.6	12.7	12.8
30	10.5	10.5	10.6	10.7	10.8	10.9	11.0	11.1	11.2	11.3	11.4	11.5	11.6	11.7	11.8	11.9	12.0	12.1	12.2	12.3	12.4	12.5	12.6	12.7
31	10.4	10.4	10.5	10.6	10.7	10.8	10.9	11.0	11.1	11.2	11.3	11.4	11.5	11.6	11.7	11.8	11.9	12.0	12.1	12.2	12.3	12.4	12.5	12.6
32	10.3	10.3	10.4	10.5	10.6	10.7	10.8	10.9	11.0	11.1	11.2	11.3	11.4	11.5	11.6	11.7	11.8	11.9	12.0	12.1	12.2	12.3	12.4	12.5
33	10.2	10.2	10.3	10.4	10.5	10.6	10.7	10.8	10.9	11.0	11.1	11.2	11.3	11.4	11.5	11.6	11.7	11.8	11.9	12.0	12.1	12.2	12.3	12.4
34	10.1	10.1	10.2	10.3	10.4	10.5	10.6	10.7	10.8	10.9	11.0	11.1	11.2	11.3	11.4	11.5	11.6	11.7	11.8	11.9	12.0	12.1	12.2	12.3
35	10.0	10.0	10.1	10.2	10.3	10.4	10.5	10.6	10.7	10.8	10.9	11.0	11.1	11.2	11.3	11.4	11.5	11.6	11.7	11.8	11.9	12.0	12.1	12.2
36	9.9	9.9	10.0	10.1	10.2	10.3	10.4	10.5	10.6	10.7	10.8	10.9	11.0	11.1	11.2	11.3	11.4	11.5	11.6	11.7	11.8	11.9	12.0	12.1
37	9.8	9.8	9.9	10.0	10.1	10.2	10.3	10.4	10.5	10.6	10.7	10.8	10.9	11.0	11.1	11.2	11.3	11.4	11.5	11.6	11.7	11.8	11.9	12.0
38	9.7	9.7	9.8	9.9	10.0	10.1	10.2	10.3	10.4	10.5	10.6	10.7	10.8	10.9	11.0	11.1	11.2	11.3	11.4	11.5	11.6	11.7	11.8	11.9
39	9.6	9.6	9.7	9.8	9.9	10.0	10.1	10.2	10.3	10.4	10.5	10.6	10.7	10.8	10.9	11.0	11.1	11.2	11.3	11.4	11.5	11.6	11.7	11.8
40	9.5	9.5	9.6	9.7	9.8	9.9	10.0	10.1	10.2	10.3	10.4	10.5	10.6	10.7	10.8	10.9	11.0	11.1	11.2	11.3	11.4	11.5	11.6	11.7
41	9.4	9.4	9.5	9.6	9.7	9.8	9.9	10.0	10.1	10.2	10.3	10.4	10.5	10.6	10.7	10.8	10.9	11.0	11.1	11.2	11.3	11.4	11.5	11.6
42	9.3	9.3	9.4	9.5	9.6	9.7	9.8	9.9	10.0	10.1	10.2	10.3	10.4	10.5	10.6	10.7	10.8	10.9	11.0	11.1	11.2	11.3	11.4	11.5
43	9.2	9.2	9.3	9.4	9.5	9.6	9.7	9.8	9.9	10.0	10.1	10.2	10.3	10.4	10.5	10.6	10.7	10.8	10.9	11.0	11.1	11.2	11.3	11.4
44	9.1	9.1	9.2	9.3	9.4	9.5	9.6	9.7	9.8	9.9	10.0	10.1	10.2	10.3	10.4	10.5	10.6	10.7	10.8	10.9	11.0	11.1	11.2	11.3
45	9.0	9.0	9.1	9.2	9.3	9.4	9.5	9.6	9.7	9.8	9.9	10.0	10.1	10.2	10.3	10.4	10.5	10.6	10.7	10.8	10.9	11.0	11.1	11.2
46	8.9	8.9	9.0	9.1	9.2	9.3	9.4	9.5	9.6	9.7	9.8	9.9	10.0	10.1	10.2	10.3	10.4	10.5	10.6	10.7	10.8	10.9	11.0	11.1
47	8.8	8.8	8.9	9.0	9.1	9.2	9.3	9.4	9.5	9.6	9.7	9.8	9.9	10.0	10.1	10.2	10.3	10.4	10.5	10.6	10.7	10.8	10.9	11.0
48	8.7	8.7	8.8	8.9	9.0	9.1	9.2	9.3	9.4	9.5	9.6	9.7	9.8	9.9	10.0	10.1	10.2	10.3	10.4	10.5	10.6	10.7	10.8	10.9
49	8.6	8.6	8.7	8.8	8.9	9.0	9.1	9.2	9.3	9.4	9.5	9.6	9.7	9.8	9.9	10.0	10.1	10.2	10.3	10.4	10.5	10.6	10.7	10.8
50	8.5	8.5	8.6	8.7	8.8	8.9	9.0	9.1	9.2	9.3	9.4	9.5	9.6	9.7	9.8	9.9	10.0	10.1	10.2	10.3	10.4	10.5	10.6	10.7
51	8.4	8.4	8.5	8.6	8.7	8.8	8.9	9.0	9.1	9.2	9.3	9.4	9.5	9.6	9.7	9.8	9.9	10.0	10.1	10.2	10.3	10.4	10.5	10.6
52	8.3	8.3	8.4	8.5	8.6	8.7	8.8	8.9	9.0	9.1	9.2	9.3	9.4	9.5	9.6	9.7	9.8	9.9	10.0	10.1	10.2	10.3	10.4	10.5
53	8.2	8.2	8.3	8.4	8.5	8.6	8.7	8.8	8.9	9.0	9.1	9.2	9.3	9.4	9.5	9.6	9.7	9.8	9.9	10.0	10.1	10.2	10.3	10.4
54	8.1	8.1	8.2	8.3	8.4	8.5	8.6	8.7	8.8	8.9	9.0	9.1	9.2	9.3	9.4	9.5	9.6	9.7	9.8	9.9	10.0	10.1	10.2	

附 録 4 マトリックスの計算法

1 行列式の値

a 非対角の場合

行列式は、下の左に与えられたようなものとする。この計算の大部分は、補助マトリックスを右のように作ることである。

$$\begin{array}{cccc} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{array}$$

与えられたマトリックス

$$\begin{array}{cccc} a_{11} & x_{12} & x_{13} & x_{14} \\ a_{21} & x_{22} & x_{23} & x_{24} \\ a_{31} & x_{32} & x_{33} & x_{34} \\ a_{41} & x_{42} & x_{43} & x_{44} \end{array}$$

(1) (2) (3) (4)

補助マトリックス

(1) 与えられたマトリックスから第1列を写して

$$x_{12} = a_{12}/a_{11} \quad x_{13} = a_{13}/a_{11} \quad x_{14} = a_{14}/a_{11}$$

とする。

(2)

$$\begin{aligned} x_{22} &= a_{22} - a_{21}x_{12} \\ x_{32} &= a_{32} - a_{31}x_{12} \\ x_{42} &= a_{42} - a_{41}x_{12} \end{aligned}$$

(3)

$$\begin{aligned} x_{23} &= (a_{23} - a_{21}x_{13})/x_{22} \\ x_{33} &= a_{33} - a_{31}x_{13} - x_{32}x_{23} \\ x_{43} &= a_{43} - a_{41}x_{13} - x_{42}x_{23} \end{aligned}$$

(4)

$$\begin{aligned} x_{24} &= (a_{24} - a_{21}x_{14})/x_{22} \\ x_{34} &= (a_{34} - a_{31}x_{14} - x_{32}x_{24})/x_{23} \\ x_{44} &= a_{44} - a_{41}x_{14} - x_{42}x_{24} - x_{43}x_{34} \end{aligned}$$

$$D = a_{11} \times x_{22} \times x_{33} \times x_{44}$$

例 行列式の値

与えられたマトリックス

27.948	-6.2992	-5.2125	8.1643
-84.522	174.46	19.412	-9.8742
-6.3428	4.5542	3.9844	12.748
7.1932	-10.764	20.398	115.35

補助マトリックス

27.948	-0.22539	-0.18651	0.29212
-84.522	155.41	0.023472	0.095337
-6.3428	3.1246	2.7281	5.2428
7.1932	-9.1427	21.954	-0.98007

$$D = 27.948 \times 155.41 \times 2.7281 \times -0.98007 = -11613$$

b. 対称の場合

行列式及び補助マトリックスが

与えられたマトリックス

a_{11}	a_{12}	a_{13}	a_{14}
a_{12}	a_{22}	a_{23}	a_{24}
a_{13}	a_{23}	a_{33}	a_{34}
a_{14}	a_{24}	a_{34}	a_{44}

補助マトリックス

a_{11}	x_{12}	x_{13}	x_{14}
a_{12}	x_{22}	x_{23}	x_{24}
a_{13}	x_{32}	x_{33}	x_{34}
a_{14}	x_{42}	x_{43}	x_{44}
(1)	(2)	(3)	

であるとする。このやり方は

(1) 非対称マトリックスの場合と同じ

$$\begin{aligned} (2) \quad x_{22} &= a_{22} - a_{12}x_{12} \\ x_{32} &= a_{23} - a_{13}x_{12} & x_{23} &= x_{32}/x_{22} \\ x_{42} &= a_{24} - a_{14}x_{12} & x_{24} &= x_{42}/x_{22} \end{aligned}$$

$$\begin{aligned} (3) \quad x_{33} &= a_{33} - a_{13}x_{13} - x_{32}x_{23} \\ x_{43} &= a_{34} - a_{14}x_{13} - x_{42}x_{23} & x_{34} &= x_{43}/x_{33} \end{aligned}$$

$$\begin{aligned} (4) \quad x_{44} &= a_{44} - a_{14}x_{14} - x_{42}x_{24} - x_{43}x_{34} \\ D &= a_{11} \times x_{22} \times x_{33} \times x_{44} \end{aligned}$$

例 行列式の値

与えられたマトリックス

1.2643	0.98724	-0.40312	1.5734
0.98724	1.4396	2.0937	-0.9682
-0.40312	2.0937	0.87326	2.0431
1.5734	-0.96820	2.0431	2.6114

補助マトリックス

1.2643	0.78086	-0.31885	1.2445
0.98724	0.66870	3.6017	-3.2855
-0.40312	2.4085	-7.9300	-1.3188
1.5734	-2.1970	10.458	7.2271

$$D = -48.453$$

2. 連立方程式

方程式は

$$\begin{aligned} a_{11}X_1 + a_{12}X_2 + a_{13}X_3 &= a_{14} \\ a_{12}X_1 + a_{22}X_2 + a_{23}X_3 &= a_{24} \\ a_{13}X_1 + a_{23}X_2 + a_{33}X_3 &= a_{34} \end{aligned}$$

であるとする。与えられたマトリックスを書き、補助マトリックスを計算する。

与えられたマトリックス

a_{11}	a_{12}	a_{13}	a_{14}
a_{12}	a_{22}	a_{23}	a_{24}
a_{13}	a_{23}	a_{33}	a_{34}

補助マトリックス

a_{11}	x_{12}	x_{13}	x_{14}
a_{12}	x_{22}	x_{23}	x_{24}
a_{13}	x_{32}	x_{33}	x_{34}
(1)	(2)	(3)	(4)

(1) 行列式の計算と同様

$$\begin{aligned} (2) \quad x_{22} &= a_{22} - a_{12}x_{12} \\ x_{32} &= a_{23} - a_{13}x_{12} & x_{23} &= x_{32}/x_{22} \\ & & x_{24} &= (a_{24} - a_{12}x_{14})/x_{22} \end{aligned}$$

$$(3) \quad x_{33} = a_{33} - a_{13}x_{13} - x_{32}x_{23}$$

$$(4) \quad x_{34} = (a_{34} - a_{13}x_{14} - x_{32}x_{24})/x_{33}$$

そうすると

$$\begin{aligned} X_3 &= x_{34} \\ X_2 &= x_{24} - x_{23} X_3 \\ X_1 &= x_{14} - x_{12} X_2 - x_{13} X_3 \end{aligned}$$

例 連立方程式の解

与えられたマトリックス

947.17	-20834	272.76	694.96
-20834	414.23	-354.00	79.648
272.76	-354.00	630.00	407.00

補助マトリックス

947.17	-0.021996	0.28798	0.73372
-20834	413.77	-0.84105	0.22944
272.76	-348.00	258.77	1.1080

$$X_3 = 1.1080$$

$$X_2 = 0.22944 - (-0.84105 \times 1.1080) = 1.1613$$

$$X_1 = 0.73372 - (-0.021996 \times 1.1613) - (0.28797 \times 1.1080) = 0.44019$$

3. ガウスの常数と偏相関および相関係数の相関マトリックスからの計算

直交マトリックス

a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	1	0	0	0	0	0
a_{12}	a_{22}	a_{23}	a_{24}	a_{25}	a_{26}	0	1	0	0	0	0
a_{13}	a_{23}	a_{33}	a_{34}	a_{35}	a_{36}	0	0	1	0	0	0
a_{14}	a_{24}	a_{34}	a_{44}	a_{45}	a_{46}	0	0	0	1	0	0
a_{15}	a_{25}	a_{35}	a_{45}	a_{55}	a_{56}	0	0	0	0	1	0
a_{16}	a_{26}	a_{36}	a_{46}	a_{56}	a_{66}	0	0	0	0	0	1

補助マトリックス

a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	1					
a_{12}	x_{22}	x_{23}	x_{24}	x_{25}	x_{26}	y_{21}	y_{22}				
a_{13}	x_{32}	a_{33}	x_{34}	x_{35}	x_{36}	y_{31}	y_{32}	y_{33}			
a_{14}	x_{42}	a_{43}	x_{44}	x_{45}	x_{46}	y_{41}	y_{42}	y_{43}	y_{44}		
a_{15}	x_{52}	x_{53}	x_{54}	x_{55}	x_{56}	y_{51}	y_{52}	y_{53}	y_{54}	y_{55}	
a_{16}	x_{62}	x_{63}	x_{64}	x_{65}	x_{66}	y_{61}	y_{62}	y_{63}	y_{64}	y_{65}	y_{66}

手順 (1) (2) (3) (4) (5) (6)

b マトリックス

b_{61}	b_{62}	b_{63}	b_{64}	b_{65}	
b_{51}	b_{52}	b_{53}	b_{54}		b_{56}
b_{41}	b_{42}	b_{43}		b_{45}	b_{46}
b_{31}	b_{32}		b_{34}	b_{35}	b_{36}
b_{21}		b_{23}	b_{24}	b_{25}	b_{26}
	b_{12}	b_{13}	b_{14}	b_{15}	b_{16}

c マトリックス

c_{11}	c_{12}	c_{13}	c_{14}	c_{15}	c_{16}
	c_{22}	c_{23}	c_{24}	c_{25}	c_{26}
		c_{33}	c_{34}	c_{35}	c_{36}
			c_{44}	c_{45}	c_{46}
				c_{55}	c_{56}
					c_{66}

γ マトリックス

γ_{12}	γ_{13}	γ_{14}	γ_{15}	γ_{16}
	γ_{23}	γ_{24}	γ_{25}	γ_{26}
		γ_{34}	γ_{35}	γ_{36}
			γ_{45}	γ_{46}
				γ_{56}

$$\gamma_{123456}^2 = a_{16}^2 + x_{62}x_{26} + x_{63}x_{36} + x_{64}x_{46} + x_{65}x_{56}$$

補助マトリックスの計算

(1) もとのマトリックスから第1行と第2行をうつす.

$$\begin{aligned} x_{22} &= a_{22} - a_{12}a_{12} \\ x_{32} &= a_{23} - a_{13}a_{12} \\ x_{42} &= a_{24} - a_{14}a_{12} \\ x_{52} &= a_{25} - a_{15}a_{12} \\ x_{62} &= a_{26} - a_{16}a_{12} \end{aligned}$$

$$\begin{aligned} x_{23} &= x_{32}/x_{22} \\ x_{24} &= x_{42}/x_{22} \\ x_{25} &= x_{52}/x_{22} \\ x_{26} &= x_{62}/x_{22} \\ y_{21} &= -a_{12}/x_{22} \\ y_{22} &= 1/x_{22} \end{aligned}$$

$$\begin{aligned} x_{33} &= a_{33} - a_{13}a_{13} - x_{32}x_{23} \\ x_{43} &= a_{34} - a_{14}a_{13} - x_{42}x_{23} \\ x_{53} &= a_{35} - a_{15}a_{13} - x_{52}x_{23} \\ x_{63} &= a_{36} - a_{16}a_{13} - x_{62}x_{23} \end{aligned}$$

$$\begin{aligned} x_{34} &= x_{43}/x_{33} \\ x_{35} &= x_{53}/x_{33} \\ x_{36} &= x_{63}/x_{33} \\ y_{31} &= (-a_{13} - x_{32}y_{21})/x_{33} \\ y_{32} &= (-x_{32}y_{22})/x_{33} \\ y_{33} &= 1/x_{33} \end{aligned}$$

$$\begin{aligned}
 (4) \quad x_{44} &= a_{44} - a_{14} a_{14} - x_{42} x_{24} - x_{43} x_{34} \\
 x_{54} &= a_{45} - a_{15} a_{14} - x_{52} x_{24} - x_{53} x_{34} \\
 x_{64} &= a_{46} - a_{16} a_{14} - x_{62} x_{24} - x_{63} x_{34} \\
 x_{45} &= x_{54} / x_{44} \\
 x_{46} &= x_{64} / x_{44} \\
 y_{41} &= (-a_{14} - x_{42} y_{21} - x_{43} y_{31}) / x_{44} \\
 y_{42} &= (-x_{42} y_{22} - x_{43} y_{32}) / x_{44} \\
 y_{43} &= (-x_{43} y_{33}) / x_{44} \\
 y_{44} &= 1 / x_{44}
 \end{aligned}$$

$$\begin{aligned}
 (5) \quad x_{55} &= a_{55} - a_{15} a_{15} - x_{52} x_{25} - x_{53} x_{35} - x_{54} x_{45} \\
 x_{65} &= a_{56} - a_{16} a_{15} - x_{62} x_{25} - x_{63} x_{35} - x_{64} x_{45} \\
 x_{56} &= x_{65} / x_{55} \\
 y_{51} &= (-a_{15} - x_{52} y_{21} - x_{53} y_{31} - x_{54} y_{41}) / x_{55} \\
 y_{52} &= (-x_{52} y_{22} - x_{53} y_{32} - x_{54} y_{42}) / x_{55} \\
 y_{53} &= (-x_{53} y_{33} - x_{54} y_{43}) / x_{55} \\
 y_{54} &= (-x_{54} y_{44}) / x_{55} \\
 y_{55} &= 1 / x_{55}
 \end{aligned}$$

等々

c マトリックスの計算

$$\begin{aligned}
 (1) \quad c_{16} &= y_{61} \\
 c_{15} &= y_{51} - x_{56} c_{16} \\
 c_{14} &= y_{41} - x_{45} c_{15} - x_{46} c_{16} \\
 c_{13} &= y_{31} - x_{34} c_{14} - x_{35} c_{15} - x_{36} c_{16} \\
 c_{12} &= y_{21} - x_{23} c_{13} - x_{24} c_{14} - x_{25} c_{15} - x_{26} c_{16} \\
 c_{11} &= 1 - a_{12} c_{12} - a_{13} c_{13} - a_{14} c_{14} - a_{15} c_{15} - a_{16} c_{16}
 \end{aligned}$$

$$\begin{aligned}
 (2) \quad c_{26} &= y_{62} \\
 c_{25} &= y_{52} - y_{56} c_{26} \\
 c_{24} &= y_{42} - x_{45} c_{25} - x_{46} c_{26} \\
 c_{23} &= y_{32} - x_{34} c_{24} - x_{35} c_{25} - x_{36} c_{26} \\
 c_{22} &= y_{22} - x_{23} c_{23} - x_{24} c_{24} - x_{25} c_{25} - x_{26} c_{26}
 \end{aligned}$$

$$\begin{aligned}
 (3) \quad c_{36} &= y_{63} \\
 c_{35} &= y_{53} - x_{56} c_{36} \\
 c_{34} &= y_{43} - x_{45} c_{35} - x_{46} c_{36} \quad \text{等々}
 \end{aligned}$$

b マトリックスの計算

一般公式は

$$b_{ij} = \frac{c_{ij}}{c_{ii}}$$

である。

(1) $d_{11}, d_{22}, \dots, d_{66}$ を与える $c_{11}, c_{22}, \dots, c_{66}$ の逆数を計算する

$$(2) \quad b_{61} = d_{66} c_{16} \quad b_{62} = d_{66} c_{26} \quad b_{63} = d_{66} c_{36} \quad b_{64} = d_{66} c_{46} \quad b_{65} = d_{66} c_{56}$$

$$(3) \quad b_{51} = d_{55} c_{15} \quad b_{52} = d_{55} c_{25} \quad b_{53} = d_{55} c_{35} \quad b_{54} = d_{55} c_{45} \quad b_{56} = d_{55} c_{56}$$

等々

γ マトリックスの計算

一般公式は

$$\gamma_{ij} = \frac{c_{ij}}{\sqrt{c_{ii} c_{jj}}}$$

$$(1) \quad \rho_{11} = \frac{1}{\sqrt{c_{11}}}, \rho_{22} = \frac{1}{\sqrt{c_{22}}}, \dots, \rho_{66} = \frac{1}{\sqrt{c_{66}}} \quad \text{を計算する}$$

$$(2) \quad \gamma_{12} = c_{12} \rho_{11} \rho_{22}, \quad \gamma_{13} = c_{13} \rho_{11} \rho_{33} \quad \text{等}$$

例 計算係數と重回歸係數の計算

[illegible]

$$\gamma_{1,23}^2 \dots \gamma_b^2 = 0.6412^2 + (0.03135 \times 0.02475) + (0.1243 \times 0.08467) + (0.1935 \times 0.1368) + (0.1094 \times 0.09486) = 0.4593$$

$$r_{1.23\dots 6} = 0.678$$