2000/5/1
第　3　講

比較統計システム論　講義録（2000 年度）

推定に有効な分布

　実証研究では、母集団の分数が分かっていることはまれであるし、標本数も限定されている。この場合には正規分布以外の分布を適用する必要がある。

t分布

　アイルランドのダブリンにあるビール会社ギネスの技師ウィリアム・ゴセットが発見したのがt分布である。彼はビールの品質を管理するうえで、ビールを選んで平均$\bar{x}$や分散s^2を計算していたが、標本数が少ないときにs^2の値がおかしくなることに気づき、品質の平均$\bar{x}$の信頼性に $s/\sqrt{n}$ （標準偏差）を用いてよいかどうかに疑問を持ち始めた。その結果、彼は「平均が有する可能な誤差」（1908）でt分布を発表した。
　ゴセットはギネスの立場を重視して、修正一遍の例外を除き、すべての論文を「student」という名で発表した。t分布が student の t 分布と呼ばれることがあるのはそのためである。

標本数が少ない場合の標本平均は、正規分布ではなく、σの代わりに $\hat{s}$を代入して

$$t=\frac{(\bar{x}-\mu)\sqrt{n}}{\hat{s}}=\frac{(\bar{x}-\mu)\sqrt{n-1}}{s}$$

とおくと、この確率分布は平均δの左右対称な分布になることが知られており、t 分布と呼ばれる。
t分布を変形すると、

$$t=\frac{\bar{x}-\mu}{\sqrt{\sigma^2/n}}\Bigg/\sqrt{\frac{s^2}{\sigma^2}}=\frac{\bar{x}-\mu}{\sqrt{\sigma^2/n}}\Bigg/\sqrt{\frac{(n-1)s^2}{\sigma^2}\Big/(n-1)}$$

分子 $(\bar{x}-\mu)\big/\sqrt{\sigma^2/n}$ は、標準正規分布N（0.1）に、分母（n－1）s^2／σ^2は自由度n－1のX^2分布X^2（n－1）に従う。

$\bar{x}$とs^2は独立であり、N（0.1）とX^2（n－1）の組み合わせからtの密度関数が求まる。

t分布　 2 つの確率変数YとZが次の条件を満たす。
（a）Zは標準正規分布N（0.1）に従う。
（b）Yは自由度kのX^{2}分布X^2（k）に従う。
（c）ZとYは独立である。

確率変数 $t = \dfrac{Z}{\sqrt{Y/k}}$, と定義すると、tが従う確率分布を自由度kのt分布という。

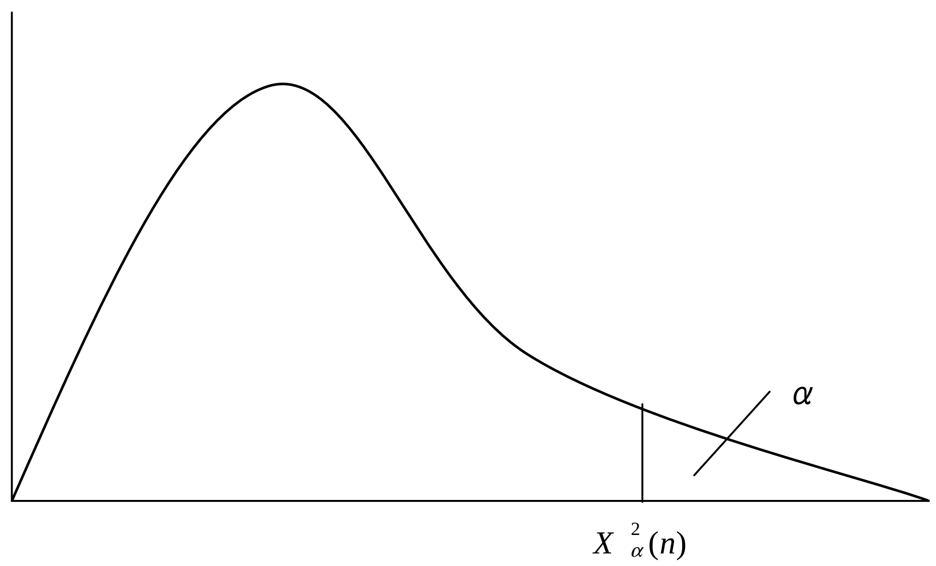
$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

で定義されたt統計量は、自由度n－１のt分布に従う。$\bar{x}$の標準偏差 $s/\sqrt{n}$ を標本平均の標準誤差という。

　t分布は正規分布よりばらつきが大きく、nの値が小さくなるほどばらつきは大きくなる。nが大きくなるにしたがって、その形は正規分布に近くなる。

カイ2乗分布

平均0、分散1の標準正規分布の母集団から、互いに独立に選ばれたn個の標本を
Z_1, Z_2, …Z_nとすると、
$X^2 = Z_1^2 + Z_2^2 + Z_3^2 + \cdots + Z_n^2$
の分布をカイ2乗分布という。下図のような形状を示す。nは自由度。



$X^2_\alpha(n)$

E（Z_i^2）＝１であるから、E（X^2）＝nとなる。nを自由度という。

　X^{2}分布を用いると、正規母集団からの標本Z_1, Z_2, …Z_nに基づく標本分散S^{2}の標本分布は次のように表わせる。

$$S^2 = \frac{1}{n-1}\left\{(Z_1 - \bar{Z})^2 + (Z_2 - \bar{Z})^2 + \ldots + (Z_n - \bar{Z})^2\right\}$$

標本分散

総計量　$X^2 = \dfrac{(n-1)S^2}{\sigma}$

は自由度n—1のX^{2}分布X^2（n－1）に従う。

S^{2}の定義式を と変化すると、　$(Z_1-\mu)^2+(Z_2-\mu)^2+\ldots+(Z_n-\mu)^2$

$$F = \dfrac{\dfrac{u_1}{n_1}}{\dfrac{u_2}{n_2}}$$

に類似している。標本Z$_1$，Z$_2$，…Z$_n$を標準化すると、

$\dfrac{Z_1-\mu}{\sigma},\dfrac{Z_2-\mu}{\sigma}\ldots\dfrac{Z_n-\mu}{\sigma}$ は標準正規分布 N（0．１）に従い、互いに独立であるから、

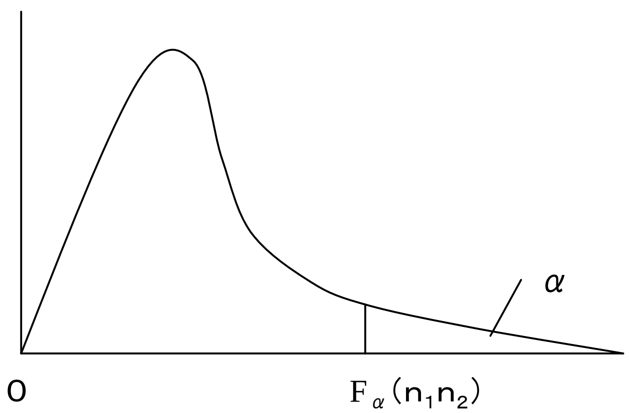
$\left(\dfrac{Z_1-\mu}{\sigma}\right)^2+\left(\dfrac{Z_2-\mu}{\sigma}\right)^2+\ldots+\left(\dfrac{Z_n-\mu}{\sigma}\right)^2$ は自由度nの X^{2}分布に従う。これを先の（n－1）S^{2}の形に近づけると、

$$\frac{(n-1)S^2}{\sigma} = \left(\frac{Z_1-\overline{Z}}{\sigma}\right)^2+\left(\frac{Z_2-\overline{Z}}{\sigma}\right)^2+\ldots+\left(\frac{Z_n-\overline{Z}}{\sigma}\right)^2$$

となり、自由度n－1の X^{2}分布に従うことがわかる。

F分布

お互いに独立な2つの確率変数U$_1$，U$_2$がそれぞれ自由度n$_1$，n$_2$のカイ2乗分布をするとき、次の式で定義される確率変数
の分布を自由度n$_1$，n$_2$のF分布という。分布の形状は非対称である。



母集団が分散σ_1^2の正規分布で、S_1^2を大きさn_1の標本分散、母集団が分散σ_2^2の正規分布で、S_2^2を大きさn_2の標本分散とすると、

$$\frac{n_1 - S_1^2}{\sigma_1^2}, \frac{n_2 - S_2^2}{\sigma_2^2}$$ はそれぞれ自由度n_1-1、n_2-1のカイ2乗分布に従う。そこで

$$F = \frac{\dfrac{n_1 S_1^2}{(n_1-1)\sigma_1^2}}{\dfrac{n_2 S_2^2}{(n_2-1)\sigma_2^2}}$$

は下分布に従う。

分散分析

標本グループ（級あるいは層）が複数あるときに、各級が同一集団のものかどうかを検定する手法として、母集団の分散を級（層）間分散（Between variance）で説明できる部分と、級（層）間分散で説明できる部分が相対的に小さければ同一母集団の標本であるとみなすものを分散分析という。

X_{ij}（$i=1, 2, \cdots, m$, $j=1, 2, \cdots, n_i$）を正規分布に従う確率変数で大きさ $N=\sum_{i=1}^{m} n_i$ の母集団の一つの標本と考える。

$$\frac{1}{N}\sum_{i=1}^{m}\sum_{i=1}^{n_i} n_i (x_{ij} - \bar{x})^2$$

$$= \frac{1}{N}\sum_{i=1}^{m}\sum_{i=1}^{n_i} (x_{ij} - \bar{x})^2 + \frac{1}{N}\sum_{i=1}^{m} n_i (\bar{x}_i - \bar{x})^2$$

左辺は全体の分散（全分散）を表わし、右辺第１項は同一級（層）内の分散＜級（層）内分散＞、第2項は級（層）間の分散＜級（層）間分散＞を表わしている。

　級（層）内分散は級（層）ごとの内部の変化であり、各級（層）を同一母集団から抽出したとする仮説が正しいか否かに関係なく、確率的な分布をしていると考える。これに対して級（層）間分散は、もし各級（層）を同一母集団から取り出したのであれば確率的な変化が主因となるが、級（層）ごとに母集団が違うのであれば、母集団の差が主因と考えられる。

　級内分散に比べ、確率で説明できる以上に級間分散が大きいならば、級間分散は有意、すなわち各級の平均の差は有意ということになる。

　X_{ij}は正規分布に従う変数で、その分散はσ^2とすると、

$$\frac{Nx(\text{全分散})}{\sigma^2}$$ は自由度$N-1$のカイ2乗分布に従う。

同様に、

$$\frac{Nx(\text{級内分散})}{\sigma^2}$$ は自由度$N-m$のカイ2乗分布に従う。

$$\frac{Nx(\text{級間分散})}{\sigma^2}$$

は自由度m－1のカイ2乗分布に従う。

カイ2乗分布の期待値は、その自由度に等しくなる＜E(X^2)＝n n＝自由度＞ので、

$$E\left[\frac{N(\text{全分散})}{N-1}\right]=E\left[\frac{N(\text{級内分散})}{N-m}\right]=E\left[\frac{N(\text{級間分散})}{m-1}\right]=\sigma^2$$

級内分散と級間分散は互いに独立なので、母集団が1つであるならば、分散比

$$F=\frac{\dfrac{\text{級間分散}}{m-1}}{\dfrac{\text{級間分散}}{N-m}}$$

は自由度(m－1、N－m)のF分布に従う。

個々の級が同一母集団から抽出されたという仮説をF分布に基づいて検定することができる(Fが有意に大きければ同一母集団から抽出されたという仮説は棄却される。逆にFが有意に小さければ、同一母集団から抽出されたという仮説は棄却できない)。

実験計画法と分散分析

実験計画法とは、⑴ いくつかの原因系について複数個の処理条件を人為的に設定し、それらが特性値に与える効果の比較を目的とするもの、⑵ 実験結果を示すデータは、同じ条件下で繰り返しても一定にはならず、バラツキを伴う、という2つの条件下で得られるべき情報量を最大化するための手法を指す。

これは一定の情報量を獲得するのに最少の実験費用で行う手法と言い換えてもよい。

これらの方法は、1920 年代に R. A. Fischer によってその基礎が築かれたものであるが、経済学の研究では無視されることが多い。

実験計画の仕方には2つの側面がある。

(1) 処理条件の選択――因子と水準の選び方

(2) 誤差の制御――実験配置の仕方

Ⅰ. 因子と水準の選び方

因子とは、結果の特性値に影響を及ぼすと考えられる種々の原因系のうち、その実験に取り上げて比較されるもの(経済学の説明変数)。

水準とは、因子の取る種々の条件で、具体的には数量、%、価格、男女、国々などで表わされる。

因子の選び方

二つの因子A、Bの最適条件を知るために24回実験するとしよう。次の三つの実験計画が考えられる。

（1） A、Bの因子について別々に「1因子実験」を行う。
（2） A、Bの両因子について「2因子実験」を行う。
（3） A、Bの両因子の水準の組み合わせのうち、技術的に判断して適当なものだけを実施する。

一般に「1因子実験をいくつか平行して実施するよりは、それらの水準を組み合わせて多因子実験を構成する方がよい」とされている。

経済学では、理論モデルで用いられる説明変数によって因子が決められる。

<u>水準の選び方</u>

水準数は質的因子の場合、あらかじめ決まっていることが多い。たとえば、原料を3者から購入しているとすれば、この因子の水準は3である。比較される品種、銘柄が 10 個あれば、その水準は 10 である。

これに対して、量的因子はどのようにも決まる。1 つの実験になるべく多くの因子を同時に取り上げる方が望ましく、かつ実験規模を大きくしないためには、各因子の水準数はできるだけ少ない方が望ましい。具体的には、量的因子の水準数は 2 または 3 とすることが多い(データの利用可能性に制約される)。量的因子の水準幅は、技術的に可能と考えられる限り、初めはなるべく広くとるべきである。

Ⅱ. 統計的判定の考え方

実験をいかに精密に行っても、管理しきれない誤差（error）は残る。しかし、この誤差が一定の統計的分布法則（正規分布等）に従うとき、この状態は<u>統計的管理状態</u>にあるという。

経済学でモデルの誤差項がホワイト・ノイズでなければならないというのは、モデルが統計的管理を十分やっているということを意味する。逆に、誤差項にゆがみがあれば、それは統計的管理が不十分であるということで、モデルの設定を考え直す必要があることを意味している。

Ⅲ. Fisher の3原則

統計的判定方式を採用するためには、その実験の精度を表わす誤差の値を、その実験で得られるデータ自身から評価しなければならない。そのためには実験の場を3つの原則に基づいて管理すべきことを、Ronald Fisher は提唱した。

（1） 反復（replication）→誤差分散の評価
（2） 無作為化（randomization）→系統誤差の偶然誤差化への転化
（3） 局所管理（local control）→系統誤差の除去

（1） 反復

誤差の大きさを評価するためには「同じ条件下での繰り返し（repetition）」が必要であるが、これを局所管理（属性ブロック毎に実験）を組み合わせると、「処理の一揃いを2回以上反復（replication）すべし」という原則になる。

（2） 無作為化

実験の場で生起する誤差は、偶然的に起こって統計的分布に従う偶然誤差（random error）だけではなく、ある実験単位（時間や空間の差などに従う）に伴う誤差もあり、これを系統誤差（systematic error）という。Fisher は、処理の順序をランダム化して、系統的誤差を偶然誤差に転化することを提唱したのである。

（3） 局所管理

実験の場全体の無作為化や均一化が困難である場合に実験の場を適当に区切って、局所的に高度の均一性を持つように管理することである。たとえば、業種別区分、男女別、年齢別、等々。

「系統誤差の大きい部分をブロック間差異として、誤差の評価から除去し、各ブロック内はできるだけ均一になるように管理する」のが「局所管理」の原則であり、「ブロック構成原理」とも呼ばれる。

ひとたび、この原則が採用されてブロックが構成されると、初めに意図した原因系（系統誤差を生ぜしめるもの）だけでなく、その他均一にはできない原因系は全てこれにかぶせればよい。

Ⅳ．基本設計

Fisher の３原則のうち、反復と無作為化の２つを満足する設計を完全無作為化法（completely randomized design）という。

局所管理を導入した設計を乱塊法（randomized block design）という。この時には、ブロックを構成し、各ブロック内を局所管理するとともに、処理の割付はブロック毎にランダムに決める。

３つの処理方法A、B、Cを３日間にわたって各３回ずつ実施する実験の種々の配置法を考えよう。

（a） 規則的配置

第１日	第2日	第3日
Ａ Ａ Ａ	Ｂ Ｂ Ｂ	Ｃ Ｃ Ｃ

実験装置を置き換えることの面倒を考えれば、Aを第１日目、Bを第2日目、Cを第3日目にやることはありうる。しかし「日」による系統誤差が大きいとき（実験に対する慣れ等）には、その「日」の効果と「処理方法」の差がこん然として区別がつかない（「日」と「処理」が交絡して＜confound＞いるという）。

(b) 完全無作為化法による配置

第１日	第２日	第３日
Ｂ Ｃ Ａ	Ｃ Ｂ Ｂ	Ａ Ｃ Ａ

系統誤差との交絡を避ける方法が無作為化である。この方法の下でくり返し数を増やせば、各処理に伴う誤差は打ち消しあって平均として公平な比較に近づくことが期待される。「日」による系統誤差が大きい場合には、処理の配置はバランスをとった方がよい。

(c) 乱塊法による配置

ブロック1	ブロック2	ブロック3
Ｂ Ｃ Ａ	Ｃ Ａ Ｂ	Ａ Ｃ Ｂ

(b) の問題を避けるために、ランダムな配置を各日ごとに行う乱塊法を採用することも考えられる。完全無作為化法における誤差要因としては、「日間変動」と「日内変動」の双方を含み、反復数を増やすとこの「日間変動」が一層大きくなる恐れがある。これに反して乱塊法における誤差は、「日内変動」のみしか含まないので、日数を増やしてもブロック間差違を大きくするだけで「誤差」が大きくなる心配はない。

(d) ラテン方格法による配置

ブロック1	ブロック2	ブロック3
Ｂ Ｃ Ａ	Ｃ Ａ Ｂ	Ａ Ｂ Ｃ

１日の中での実験順序も系統誤差を生じさせるのであれば、「実験順序」も「日」と同様にブロック因子として管理すべきかもしれない。このように、２種のブロックを構成する計画を、ラテン方格法（Latin square method）という。

以上の例では、どのブロックでも、比較したい処理の一揃いがいつでも実験されていた。この条件を満足する計画を、<u>完備型計画</u>という。

これに対して、各ブロックには比較したい処理の一部しか配置できない計画を総称して<u>不完備型計画</u>という。

以上のような実験計画法を明示的に取り込んだ統計的手法を、分散分析（Analysis of Variance, ANOVA）という。

基本モデル（一元配置法）

因子水準		A_1	A_2	A_3	…	A_K	
繰	1	X_{11}	X_{21}	X_{31}	…	X_{K1}	
り	2	X_{12}	X_{22}	X_{32}	…	X_{K2}	
返	3	X_{13}	X_{23}	X_{33}	…	X_{K3}	
し	・	・	・	・		・	
数	・	・	・	・		・	
	・	・	・	・		・	
	n	X_{1n}	X_{2n}	X_{3n}		X_{Kn}	
列の合計		T_1	T_2	T_3		T_K	総計 T
列の平均		$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_3$		$\bar{X}_K$	総平均 $\bar{X}$
列の効果		$\bar{X}_1-\bar{X}$	$\bar{X}_2-\bar{X}$	$\bar{X}_3-\bar{X}$		$\bar{X}_K-\bar{X}$	

各データから列の平均を引くと、誤差が求められる。

因子水準		A_1	A_2	A_3	…	A_K
繰	1	$X_{11}-\bar{X}_1$	$X_{21}-\bar{X}_2$	$X_{31}-\bar{X}_3$	…	$X_{K1}-\bar{X}_K$
り	2	$X_{12}-\bar{X}_1$	$X_{22}-\bar{X}_2$	$X_{32}-\bar{X}_3$	…	$X_{K2}-\bar{X}_K$
返	・	・	・	・		・
し	・	・	・	・		・
数	・	・	・	・		・
	n	$X_{1n}-\bar{X}_1$	$X_{2n}-\bar{X}_2$	$X_{3n}-\bar{X}_3$		$X_{Kn}-\bar{X}_K$
誤差の和		0	0	0	…	0

「列の効果」が「誤差」に比べて有意かどうかは F 検定で判定する。

$$\frac{\text{列の効果の不偏分散}}{\text{誤差の不偏分散}} = \frac{V_1}{V_2} = F\langle K-1,\ K(n-1)\rangle$$

V_1 の自由度 ϕ_1 は K-1 である。

$$V_1 = \frac{n(\bar{x}_1-\bar{x})^2 + n(\bar{x}_2-\bar{x})^2 + \cdots + n(\bar{x}_K-\bar{x})^2}{K-1}$$

V_2 の自由度 ϕ_1 は K(n-1)である。

$$V_2 = \frac{1}{K(n-1)}\left\{(x_{11}-\bar{x}_1)^2 + (x_{12}-\bar{x}_1)^2 + \cdots + (x_{1n}-\bar{x}_1)^2 + \cdots + (x_{21}-\bar{x}_2)^2 + (x_{22}-\bar{x}_2)^2 + \cdots + (x_{2n}-\bar{x}_2)^2 + (x_{Kn}-\bar{x}_K)^2\right\}$$

これより F を求めて F-table の5%、1%有意水準と比較して、列の効果が有意かどうかを

判断すればよい。

Note:
また、このモデルでは、次の関係式が成立する(平方和の分解)。

$$\sum_i \sum_j (x_{ij} - \bar{x})^2 = \sum n_i (\bar{x}_i - \bar{x})^2 + \sum_i \sum_j (x_{ij} - \bar{x}_i)^2$$

全変動 S_T＝水循環変動 S_A＋水準内変動 S_E
（Between）　(Within)

このことから、水準間に差があるかどうかは水準間変動と水準内変動の自由度調整をした比で判別したもの F でテストできると考えられるのである。

参考文献

石村貞夫(1992)『分散分析のはなし』、東京図書。

奥野忠一、芳賀敏郎(1969)『実験計画法』、培風館。

大村　平(1992)『実験計画と分散分析のはなし』、日科技？出版。

広津千尋(1992)『実験データの解析――分散分析を超えて――』、共立出版。

豊田秀樹(1994)『違いを見抜く統計学　実験計画と分散分析入門』、講談社ブルーバックス

豊田秀樹、前田忠彦、柳井晴夫(1992)『原因をさぐる統計学　共分散構造分析入門』、講談社ブルーバックス

Cochran, W.G. and Cox, G.M. (1957) Experimental Designs, 2nd ed., New York: John Wiley & Sons.

Cox, D.R. (1958) Planning of Experiments, New York: John Wiley & Sons.