

# 目次

第 1 章	Ronald Fisher と農事試験, Pearson と生物測定学	2
1.1	概観	2
1.1.1	3つの学派	2
1.1.2	Fisher の業績	2
1.2	実験計画法	3
1.2.1	農事実験	3
1.2.2	分散分析と有意性検定	3
1.3	有意性検定	4
1.3.1	Fisher 以前の検定論	4
1.3.2	Fisher の有意性検定	5
1.4	Pearson の統計理論	5
1.4.1	記述統計学	5
1.4.2	モーメント法とピアソン系	5
第 2 章	Pearson and Neyman	6
2.1	統計的品質管理の父	6
2.2	Neyman-Pearson の統計的仮説検定理論	6
第 3 章	参考文献	7
参考文献		8

豊かな統計理論の展開の歴史は、豊かな応用とともにある。

- (1) Fisher の統計は Rothamsted 農業研究所におけるデータ解析問題に深く根ざしている。
- (2) 1940s の Neyman-Pearson の統計理論は統計的品質管理の問題があった。

それぞれの大家をまとめる。

- (1) Ronald Fisher 90-62：近視だったために夜は音で数学を学び、独自の数学への適性を身に付け、奨学金を得て Cambridge 大学で数学と物理学を学んだが、その後は第一次世界大戦もあってうやむやになる（修士はとっていないのでは？）。Pearson への反感から大学教員ではなく農事試験場のしごとを選び、36 歳にはそこで実験計画法を確立する。その後に Galton 教授職へ戻り、そこでは溜まりに溜まっていた優生学への興味を爆発させる。
- (2) Karl Pearson 57-36：本当は Carl. 菊池大麓と Cambridge で学び、ドイツ留学経験も積んで（ここでマルクスに傾倒）中世ドイツ文学や法学にも打ち込んだが、最終的には数学に戻った真のアカデミアの人。Galton と仲良くなって優生学に親しみ、優生学部の初代教授となる。生物測定学という全く独立の Darwin 以来の流れから、記述統計学を作る。
- (3) Egon Pearson 95-80：もともとは Cambridge で天体物理学を専攻してから転向。25 年に Neyman と会い、31 年に Shewhart と会う（ここで、統計的品質管理の問題意識と出会う）。父の退官後、Fisher は優生学教授で、Egon は応用統計学部教授となる。
- (4) Jerzy Neyman 94-81：父がロシアで仕事をしていた法律家で、ウクライナの大学で Bernstein に数学を学び、祖国ポーランドへ還る。ワルシャワ大学で仕事を持ちながら、ピアソンの下で留学したときに、息子と意気投合。

## 第 1 章

# Ronald Fisher と農事試験，Pearson と生物測定学

Sir Ronald Aylmer Fisher 90-62 は一時期ロザムステッド農事試験場 (Rothamsted Experimental Station) の統計研究員だったが、最終的には UCL の Pearson の座 (Galton 教授職) を継いでいる。

膨大かつ広範囲 (数理統計学から遺伝学・優生学)。

### 1.1 概観

Neyman-Pearson 学派が巨視的で、Bayes 学派が微視的であり、Fisher 学派はその中間点で、あくまで物理学実験計画法を主軸とする。巨視的には確率は相対頻度であり、微視的には確からしさである。

#### 1.1.1 3つの学派

**歴史 1.1.1.** Pearson が大標本論・記述統計学と説明され、Fisher の 20s のしごとは精密標本論・推測統計学と説明される。社会からの統計学への要請で最も大きなものは検定論である (この点がのちのち Fisher が排撃した Bayes を蘇らせることとなる)。Fisher の有意性検定は農事試験のデータ解析であったが、これを 1950s の Neyman-Pearson の仕事は、大量生産を背景とした統計的品質管理が背景にあった。ここからさらに、Wald や Lehmann によって数学的に体系化されていく。この「最適性」を指導原理とした数学的体系化、Neyman-Pearson や Bayes などの社会情勢に応えた派生は、原点である Fisher の理論とは違う色を帯びていくことになる。この3派閥は、現在にもあとを引く。

- (1) Frequentist は頻度派ともいうが、Neyman-Pearson 学派のことをいう。今日では数理統計学の主流となっている。
- (2) Bayesian はベイズ派と呼ばれる。産業応用派・意思決定理論への応用である。
- (3) Fisherian は Fisher 学派のことをいう。実験家の統計学である。

**議論 1.1.2** (Frequentist vs. Fisherian). Fisher にとっては、統計的推測理論はあくまで物理学実験手法の延長であり、科学的な帰納的推論の一環である。一方で、Frequentist は、現在の数理統計学同様、「最適性」を指導原理とする。

**議論 1.1.3** (Frequentist vs. Bayesian). Bayes の「主観確率」の理論は一度は Fisher によって排斥されたが、1950s に再構築された。

#### 1.1.2 Fisher の業績

- (1) 統計量の精密な標本分布について、Fisher が関わっていないものを見つける方が難しい。このとき母集団が正規分布に従うことを仮定していたが、これは主に生物データを扱っていた Fisher にとっては現実的である。
- (2) さらにその間の関係まで整理し、 $\chi^2$  分布、 $t$  分布、正規分布、標本相関係数  $r$  の分布は、すべて  $F$ -分布の特別な場合であることを理解した。
- (3) 大標本理論に関する論文 Fisher 22[6] にて、逆確率法を排撃し、Pearson のモーメント法に代わって最尤法を提示。<sup>t1</sup>十分

<sup>t1</sup> 考え方自体は、Edgeworth の "genuine inverse" や Pearson の段階ですでにあった。

性の概念も Laplace がすでに用いていたが、有効性、一致性などの観点から検討し、一般的な方法として提示した。母集団の母数と標本から得られる統計量を明確に区別。実際、農事試験データに Bayes の方法はフィットしない、純粋な科学的帰納法のみが要請される現場であるからだ。また、この論文では、統計学の目的は、次の3ステップから成る「データの縮約 (reduction of data)」であるとした。

- (i) 特定化 (specification) の問題：モデル選択。
  - (ii) 推定の問題：選択したモデルの母数を最適に推定する。
  - (iii) 分布の問題：構成した統計量の分布を求める。ここで漸近論を採用せず、精密な分布を求めたのが Fisher の特徴である。
- (4) あくまで関心は小標本にあり、情報量・十分性・尤度の概念をそちらに適用しようとした。
- (5) 実験計画法とは実験化のお家芸であったが、これが統計家の領域に引き込まれた。そして因果推論の基本となる分散分析 (ANOVA) という検定法を開発。

**歴史 1.1.4** (日本での受容). Pearson までの統計学で、母集団のパラメータと標本統計量が区別されなかったように、全数調査じゃなくてランダムサンプリングを使うという数学的トリックが生じたのが Fisher からである。そこで日本で受容する際も論争が起こり、数理統計学者が説得する形となった。

## 1.2 実験計画法

農場実験での実験計画法を確立するために、精密分布を計算するなどの仕事は必要であった。なぜならば、まず暗黙知であった有意性検定をするためには、誤差のコントロールが必要になった (分散分析の開発)。特に、肥沃土の不均一性が一番大きな系統的誤差原因となるので、必然的に実験の規模は限られ、必然的に小標本理論になる。

### 1.2.1 農事実験

**歴史 1.2.1** (高度農業). 1846 年の穀物法撤廃を契機に、英国で「高度農業」が始まり、ここから農事実験は「品種改良」「化学肥料の効果比較」が中心となった。Rothamsted を設立する Lawes J. B. は 1842 年にはじめて化学肥料を開発した。このとき農事試験を科学化しなければいけないという要請が高まり、その最初の研究が Johnston, J. F. W. (1849). *Experimental Agriculture, Being the Results of Past and Practical Agriculture* である。これはすごく自然言語で書かれた実験計画法の萌芽でもある。

**歴史 1.2.2.** Rothamsted では 1919 年時点で 75 年ほどの実験データが溜まっており、この分析法を考えあぐねていた。

### 1.2.2 分散分析と有意性検定

農事実験では統制すべき変数が多いので、そのそれぞれについて、変動への寄与を分離することが必要だった。当時 Fisher は変動分析 (analysis of variation) と呼んでいた。そしてその後、処置による変動に対して、有意性を判定することが必要。特に、精密標本分布を用いると、標本の大きさに拘わらず、正確な有意性検定を構成できる。

特に、肥沃土の不均一性が一番大きな系統的誤差原因となるので、必然的に実験の規模は限られ、必然的に小標本理論になる。

**歴史 1.2.3.** Fisher 18 [5] で分散 (variance) を命名し、遺伝学研究のときに発明され、翌年から就職した Rothamsted で精緻化される。ここで  $F$  分布を、当時は  $z$  分布と命名した。唯一の先行研究は、Gosset=Student 08 [4] の  $t$  分布が、当時の大標本理論に頼ることなく、標本統計量のみに基づいて検定を構成したものである。

そしてブロックデザインが大事になる。当時の実践家は誤差を最小にしようとしていたが、Fisher 以後は誤差が統計的にコントロール可能であればよい。

次に、実践家のお家芸であった有意性判定を、有意性検定で置き換えた。

## 1.3 有意性検定

Fisher の有意性検定以前にも, Pearson の  $\chi^2$ -適合度検定, Gosset の  $t$  検定 (当時は  $z$  検定) があったが, 普及したのは Fisher のものが初である. その理由は, 小標本にも適用可能であること (これは  $t$ -検定もそう), 社会的要請の 2 点がある. Gosset は観測値の精度評価を主な目的とした誤差論的な発想であったのに対し, 誤差が系統的に把握できている時点で OK であり, 有意性の検定に使える, という発想の転換がある. これはある意味で用途の制限であるが, これが爆発的な応用を生んだ.

### 1.3.1 Fisher 以前の検定論

**歴史 1.3.1** (確率誤差検定 (Wood and Stratton 1910)). 基本的な問題設定は, ある正規分布を観測しており, その母平均を推定する問題である. そのために, 誤差の発生要因を統制する必要がある (これと付随する分散分析の技法をあわせて実験計画という).

標本  $X_1, \dots, X_n \sim N(\mu, \sigma^2)$  から定まる**確率誤差 (probable error)** とは,  $0.67\sigma/\sqrt{n}$  とする. すると,  $\mu \pm 0.67\sigma/\sqrt{n}$  は確率 0.5 でこの範囲に入る. これを用いて検定していたので, 「帰無仮説」という概念はない. Fisher を待つ必要がある.

Edgeworth, F. Y. は経済学の文脈で, Pearson, K. は優生学の分野で検定を構成した. 前者は *Economic Journal* の創刊から編集者で, 後者は *Biometrika* を主宰.

**歴史 1.3.2** (Francis Ysidro Edgeworth 1845-1926 アイルランド と 数理心理学). 発明家の祖父を持つ名家の出身. 父はユグノーの子孫で, ドイツに向かう途中, 英国博物館の階段であったスペインの難民と駆け落ち中に生んだ子供. 基本的にあらゆる学問をやっていたが (弁護士資格もある), King's College London で経済学の職に就く. 新古典派の主要人物となり, 数学的形式を個人の意思決定へと応用した最初の人物となった. 効用理論と無差別曲線 (indifference curve) でミクロ経済学, Edgeworth 展開で数理統計学に名前を残す. 81 [2] は数学も, フランス語・ラテン語・古典ギリシャ語にまたがる文章も, アホみたいに読みにくいらしい.

マーシャルとはともに数学と倫理学を通じて経済学に達したという類似点がある. エッジワースは社会科学に数学の手法を適用した先駆者の一人である. 彼自身はその手法を「数理心理学」と名づけていた.

**歴史 1.3.3** (Karl Pearson の適合度検定). Galton が確立した相関・回帰の概念から出発して, ピアソン系, モーメント法, 確率誤差論,  $\chi^2$ -適合度検定を開発した. これらは記述統計学と呼ばれたが, Fisher のように標本の大小に拘らない枠組みを作ることはなかった.

$X_1, \dots, X_n$  に関する理論値と観測値の間の乖離の尺度を

$$\chi^2 := \sum_{i=1}^k \frac{(m'_i - m_i)^2}{m_i}$$

とする. ただし,  $k \leq n$  について  $m_i$  ( $i \in [k]$ ) はそれぞれの値に対する頻度である. すると,  $\chi^2$  なる統計量の漸近分布は, 自由度  $k-1$  の  $\chi^2(k-1)$  に従うことを示した [3].

**歴史 1.3.4** (William Sealy Gosset 76-37 とギネスビール). Oxford で化学と数学を学び, 学士のままギネスビール社のダブリン醸造所に就職した (農事試験同様, 数年前から統計家を積極採用していた). その後あまりの統計的困難さに, 30 歳前後で Pearson の研究室に赴いて, [4] を出した. ギネスビール社は機密保護の観点から社員が論文を出すことを禁止していたので, Student というペンネームで発表していた. もともと  $z = \frac{t}{\sqrt{n-1}}$  を用いていたが, Fisher が自身の自由度理論に併せるために  $t$  に変えた. Fisher と Pearson の仲を取り持つ立ち回りをした.

醸造過程の共変量は互いに独立でなく, データも少ない. 誤差の正規性も仮定できないので, 当時の誤差論は適用不可能であった. [4] のアブストでは

繰り返すことが困難な実験は少なくない. このような場合, データ数は少なく, いくつかの化学実験, 多くの生物学的実験, そして殆どの農事試験・大規模実験が当てはまり, 従来これらはほとんど統計学の範囲外であった.

また, 正規曲線の方法は大標本のみで適用可能であるが, いつ適用不可能になるかの研究がない.

そこで、標本平均を  $\bar{x}$ 、標本不偏標準偏差を  $s$  として

$$z = \frac{\bar{x} - \mu}{s}$$

とし、この分布と確率積分表を作った。モデルを正規  $N(\mu, \sigma)$  として  $\mu$  を推定する際、2つの平均値の統計的有意性を検定できる。

しかし、Gosset はこれを小標本の場合の正規近似に用いただけで、検定を構成はしなかった。ここでも確率誤差検定同様、棄却域を設定せず、検定に2つ以上の役割を持たせており、Wald の意味で「検定」とは言えない。

### 1.3.2 Fisher の有意性検定

有意性検定に用いるには、Gosset の研究はほとんどそのまま使えるが、唯一、有意性査定のために特殊化する必要があった。これは、Gosset は観測値の精度評価を主な目的とした誤差論的な発想であったのに対し、誤差が系統的に把握できている時点で OK であり、有意性の検定に使える、という発想の転換がある。これはある意味で用途の制限であるが、これが爆発的な応用を生んだ。

また、Pearson の適合度検定は、事前に設定した棄却域の下で棄却するのではなく、単純に  $P[\chi^2 \geq c]$  を見て判断する。

さらに、母集団と標本との確率論的な関係を厳密に扱うために自由度の概念を導入し、 $t$  検定として作り直した。

## 1.4 Pearson の統計理論

標本の理解と、特定の分布への当てはめ (fitting) を通じて、統計的手法を開発した。もちろん推測への萌芽は含まれているが、あくまで、大数観察とそこから科学的な知見を引き出そうとすることが学問的興味であった。

Pearson の検定は誤差が小さいことの確認であり（その上でマクロ生物理論を立てたいので）、Fisher の検定は系統的な誤差の中に有意なものが混じっていないかの検出である。

### 1.4.1 記述統計学

標本が十分に大きいときに、これ自体を母集団とみなすならば、背後に分布を仮定することも、標本と母平均との確率的関係に頭を悩ませることもない。この特権を享受した下での統計的手法を、記述統計学という。これは、全体を見渡して知恵を得たい、複雑な大規模データを要約して理解したいという傾向のある学問（マクロ生物学など）で取られる手法となる。一方で Gosset や Fisher は推論が必要になり、Edgeworth は意思決定が必要であった。

### 1.4.2 モーメント法とピアソン系

しかし、Pearson も非対称分布を持つ標本に出会い、これを正規分布の混合として説明しようと試みた。そのための方法としてモーメント法を開発した。4次までの積率に注目し、正規分布の歪度  $\gamma_1 = 0$  と尖度  $\gamma_2 = 3$  と比較することで近さを測れる。母モーメントを知っていれば、標本モーメントはそれに確率収束するだろうから、必要な次数までモーメント方程式を用意すれば、その解そして一致推定量が得られる、という手法である。このとき、母分布にパラメトリックな仮定をおくことになるが、これをピアソン系と呼ぶ。左右非対称性や尖り具合によって 12 に分類されている。

こうして最後に、ピアソン系の適合度が知りたくなる。そこで適合度検定が要請された。



## 第 2 章

# Pearson and Neyman

統計的推測理論が次の発展をするのは、工業化が契機になる。農事試験などの精密なものではなく、大量生産とその誤差管理が問題になる。そこで、最適な有意性検定 (optimal test of significance) の理論的枠組が生まれた。

### 2.1 統計的品質管理の父

**歴史 2.1.1** (Walter A. Shewhart 91-67 米). California Berkeley で物理学博士を取ったのちに、AT&T の前身である Bell Telephone で通信システムの信頼性向上の仕事をした際に残した知的遺産により、統計的品質管理の父と呼ばれる。プラグマティズム哲学者 C・I・ルイス の著作に影響されて操作主義的姿勢を鮮明に表し、それが彼の統計処理に影響していた。

シューハートの上司であった George D Edwards は「シューハート博士はほんの 1 ページのメモを書いた。その 3 分の 1 は、我々が今日概略の管理図と呼ぶような単純な図だった。その図と前後の文章には、今日の我々がプロセス品質管理として知っている基本原則と考慮すべきことが全て記述されていた」と回想している。

製造工程のばらつきが正規分布の形になるならば、「特殊要因」がなくなり、「偶然要因」のみからなることになり、これが理想的な管理状態とした。管理図とは、この 2 状態の区別をするツールである。

彼はアカデミアにいる William E. Deming 00-93 米にも影響を及ぼした。デミングはシューハートの考え方に基づいて科学的推論に関して研究を展開し、PDCA サイクルも生み出した。デミングは Yale で数学と物理学の博士を取ってから Bell でインターンをしたのち、ダグラス・マッカーサー将軍の下で日本政府の国勢調査コンサルタントを務め、統計的プロセス管理手法を日本の企業経営者に教えた。その後も何度も日本に赴き、1950 年から日本の企業経営者に、Bell 研究所で Shewhart から学んだ設計/製品品質/製品検査/販売などを強化する方法を伝授していった。彼が伝授した方法は、分散分析や仮説検定といった統計学的手法の応用などである。デミングは、日本の製造業やビジネスに最も影響を与えた外国人であった。このため、以前から英雄的な捉え方をされていたが、アメリカでの認知は彼が死去したころやっと広まり始めたところであった。

### 2.2 Neyman-Pearson の統計的仮説検定理論

## 第 3 章

## 参考文献

## 参考文献

- [1] 芝村良. (2004). 『R. A. フィッシャーの統計理論』. 九州大学出版会.
- [2] Edgeworth, F. Y. (1881). *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Science*.
- [3] Pearson, K. (1900). On the Criterion that a Given System of Deviations from the Probable in the Case of a Correlated System of Variables is Such that it can be Reasonably Supposed to Have Arisen from Random Sampling. *Philosophical Magazine*. 5th Series, Vol. L, pp. 157 – 175.
- [4] Gosset, W. S. (1908). The Probability Error of a Mean.
- [5] Fisher, R. A. (1918). The Correlation between Relatives on the Supposition of Mendelian Inheritance.
- [6] Fisher, R. A. (1922). On the Mathematical Foundations of Theoretical Statistics.
- [7] Fisher, R. A. (1925). Theory of Statistical Estimation.