

目次

第 1 章	統計の誕生	3
1.1	オランダ：英国政治算術とフランス確率論の衝突	3
1.1.1	英国政治算術	3
1.1.2	伊・仏の確率論	3
1.2	オランダでの人口統計と確率論	3
1.3	フランスでの最小二乗法の誕生	4
1.3.1	天文学の大衆化	4
1.3.2	Lagrange と確率変数の和	4
1.3.3	Bernoulli と最尤法	4
1.3.4	Legendre と最小二乗法	4
1.4	Quetelet の社会物理学	4
1.5	ドイツでの社会統計学	5
1.6	現代誤差論	5
1.6.1	田口玄一	5
第 2 章	Ronald Fisher と農事試験, Pearson と生物測定学	6
2.1	概観	6
2.1.1	3つの学派	6
2.1.2	Fisher の業績	6
2.2	Galton までの優生学	7
2.2.1	Mendel の遺伝理論	7
2.3	最尤法	7
2.4	実験計画法	8
2.4.1	農事実験	8
2.4.2	分散分析と有意性検定	8
2.5	有意性検定	9
2.5.1	Fisher 以前の検定論	9
2.5.2	Fisher の有意性検定	10
2.6	Pearson の統計理論	10
2.6.1	記述統計学	10
2.6.2	モーメント法とピアソン系	10
第 3 章	Pearson and Neyman	11
3.1	統計的品質管理の父	11
3.2	Neyman-Pearson の統計的仮説検定理論	11
第 4 章	参考文献	13
参考文献		14

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着け、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。

- (1) Fisher と Gosset は大量生産できない農業・ビール醸造なる対象に対して、比較実験を通じて科学的結論を導く必要があった。このため、誤差の統制（分散分析）と精密分布の決定と有意性検定を作った。
- (2) Karl Pearson と Galton は大量のデータから優生学的知見を引き出したかった、どちらかといえば社会科学的営みであった。このため、誤差を最小にする fitting の技法を開発した。
- (3) Neyman と Egon Pearson は、企業意思決定に必要なサンプリング問題から発展して、契約上の品質に関する合意形成手段としての統一検定方式を開発する必要があった。そこで、決定理論の枠組みと最適性が問題になった。^{†1}

こう見ると、検定論は、最もコミュニケーションの道具、人類生態系を別の論理で飛び回るミームであり、言語の権化であるようだ。品種・肥料の有効性について有意性を下す判定が有力者の権威から統計技法の権威へと移り、また品質管理についての権威が統計技法に移った。デジタルかもその延長であるとみれる。

人々の行動指針を与えるようなミーム。人生が模倣するような権威を帯びた形式として、芸術とは異なる作品。このようなものを作るのが夢であった。これはむしろ社会彫刻のようなものであって、統計技法だけではなく、周りの言論と社会実践が肝要になってくる。端的に言えば、「立場が違ふ人を結ぶ形式」が好きなら、それはシミュレーションや統計の手法になるわけだ。僕はコミュニケーションのプロでなくてはならない。

^{†1} これはプログラムに関する法律などでも、採択されることになるだろう。

第 1 章

統計の誕生

人類が最初に直面した統計的問題は、次の 3 つである：

観測の誤差と確率論の問題 同一観測を繰り返しても値がばらつくが、その中でどれを「最良値」「代表値」として採用すれば良いかの問題である。答えは最小二乗法が与えるが、なぜこれが「良い」のか？この問題設定は実は線形回帰問題と関連し、誤差分布が正規性の仮定を満たすとき（根源誤差の集まりとみなせるとき）、最小二乗推定量が最尤推定量となる。ここまでの問題に対する理論を誤差論という。

政治算術 William Petty 1623-1687 は political arithmetic を著し、労働価値説を唱えて、古典派経済学と統計学の祖と呼ばれる。

国状論

この 3 つを統合したのが、Quetelet の社会物理学である。

20 世紀のはじめに、英米の生物測定学と、ドイツの社会統計学とに分岐し、現代に至る。英国統計学は特に数学との結びつきが強く、記述論から推測論・検定論を作り、数理統計学の系譜となる。政治算術は、確率論の土壌を生かした Quetelet によって社会物理学となり、刺激を受けたドイツで国状論・社会統計学となり、計量経済学の系譜となる。

Leonard Savage 1917-1971 の『統計学の基礎』で Bayes 学派が復活し、そもそも母集団の前提がない際にも応用された。現在実社会で最も中心的なのが Bayes 学派である。

1.1 オランダ：英国政治算術とフランス確率論の衝突

イギリス経験論と大陸合理主義との二項対立がある。

英国的な統計的経験的確率と、伊・仏的な合理的先験的確率との、認識論的方法論の対立は最初からあった。「確率統計学」というときの両極端から歩み寄った。

1.1.1 英国政治算術

英国で互いに友人である 2 人、William Petty の『政治算術』(1690) と John Graunt 『死亡表に関する自然のおよび政治的諸観察』(1660) が人口統計学の源流となった。

1.1.2 伊・仏の確率論

ガリレイに源流を持ちつつ、Pascal と Fermat が完成させたのがこの確率論的な見方である。

1.2 オランダでの人口統計と確率論

ちょうど 2 つの接点に居たのがオランダで、ここで真に豊かな理論が生まれた。

Huygens の『チャンスの価格』は、完全に Pascal の「勝負の値」の系譜にある。実はデカルトの祖国以上にデカルト学の拠点であったのがオランダであり、チャンスの値の算出方法が体系化された。スピノザも、これに関して「確率書簡」などの記録があ

る。しかし Huygens は、Graunt の『諸観察』を読んで、人口統計との融合を試みた。

Johan de Witt 1625-1672 は数学と法律学を学び、28 にしてホラント州法律顧問（事実上の最高指導者）になり、財政再建を進めた。特に「償還年金と比べた終身年金の価格」を国会に提出した。これは独自の生命表と、Huygens の『チャンスの価格』をもとにしていた。

1.3 フランスでの最小二乗法の誕生

母関数の方法 (Lagrange) も、特性関数 (Laplace) も、最小二乗法の発展 (Legendre) も、フランスで起こった。確率論と結びつけて最適性を初めて指摘したのが Gauss、最尤法の意味論を読み取り、推定論を初めて展開したのが Laplace、数理統計学の応用先として推定論の道をより広くしたのが Quetelet である。

最尤法や十分統計量の命名は Fisher だが本質的には Laplace が開拓していた、Laplace の数学史の講義を拓いたのが Pearson。

19 世紀に、最小二乗法が工学者の常識となった中で、初めて根源誤差の仮定に辿り着いたのが Young。

1.3.1 天文学の大衆化

Tycho Brahe 1546-1601 は助手たちに反復同一観測させ、互いに吟味させた。Lippershey 1570-1619 が望遠鏡を発明したことより、天体観測は大衆化された。ガリレオは『天文対話』で、自然言語により誤差の扱いを議論した。

1.3.2 Lagrange と確率変数の和

Thomas Simpson 1710-1761 (英) は Kepler 則として発見されていて、暗黙知としてもよく使われていた公式を、定積分の近似式である Simpson 式として名を残す。A Letter to the Right Honourable George Earl of Macclesfield, President of the Royal Society, on the Advantage of taking the Mean of a Number of Observations, in practical Astronomy (1775)。にて、知識人でも「注意深くなされた 1 回の測定」神話を信じていた。そこで、Lagrange らと共に、算術平均の統計量としての分布を調べた。L'utilité de la Méthode de Prendre le Milieu. Miscellanea Taurinensia にて、母関数の方法を用いて後継した。

1.3.3 Bernoulli と最尤法

Lambert, J. H. 1728-1777 (独) は円周率の無理性を証明もした。それと Bernoulli は最尤推定の考え方を育てていた。

一方で、誤差に正規性の仮定がおけると、その分布に対応した最尤推定法が最小二乗法であることを初めて発見したのが Laplace である。

1.3.4 Legendre と最小二乗法

Nouvelles Méthodes: Pour la Détermination (1805) にて、normal equation を導く。しかし、確率論的な発想がなかった。

Robert Adrain 1775-1843 が最初に証明をしようとする。まずは誤差の従う分布がなにかを直接考察した。

Karl Gauss 1777-1855 は『天体運動論』で、正規性の仮定をおくと、算術平均が最確値であることを導く。また、先取権を巡って筆を執っている。一方でフランス学会は極めて鷹揚な態度を取る。

Laplace は著書『確率の解析的理論』の「試論」の中で、Legendre と Gauss を同等に扱っている。

1.4 Quetelet の社会物理学

応用統計学の祖と呼ばれるのが Quetelet である。フランスには、Laplace が統計的推論の手法を得ていた土壌があった。

『人間について』(1835) では「平均人の理論と社会体制の組織に関する」思想を社会物理学と呼んだ。その内実は、肉体的・知的、並びに道徳的諸性質を支配する諸法則を統計的に研究する学問であった。したがって、事実を確定する手法である統計学から

派生する理論である（ちょうど生物測定学と同じ）。そして、官庁統計・人口学とも違う。

なお, Quetelet による「統計学」とは, Gottfried Achenwall(アッヘンヴァル) 1719-1772 独の言葉を借りて, 「事実・状態を説明するものであって, 政治社会に影響を与えるすべての要素を説明するものである」という。または歴史学者 Schläözer(シュレーツァー)

1.5 ドイツでの社会統計学

Adolf Wagner 1835-1917, Ernst Engel 1821-1896, Georg Mayr ら。

1.6 現代誤差論

1.6.1 田口玄一

海軍水路部天文部から統計数理研究所へ。田口メソッドは品質管理手法を開発・設計工程に取り入れた。

第2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。Fisher が導入した精密標本論をまとめると、次の Rao の言葉になる。

特定の標本から一般化されて得た知識は不確実なものであるが、一度その不確実性を数量化できれば、種類は異なれど、それは確かな知識となる。

また、Fisher 以降、確率モデルを想定した上でパラメータを推定する「理論モデル主導的」な枠組みが支配したが、計算資源が豊かになったことで、「データ主導的方法」が興隆した。

2.1 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。

2.1.1 3つの学派

歴史 2.1.1. 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 2.1.2 (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 2.1.3 (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。

2.1.2 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、 χ^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[8] 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。^{†1} 十分性の概念も 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 2.1.4 (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り、数理統計学者が説得する形となった。

2.2 Galton までの優生学

2.2.1 Mendel の遺伝理論

歴史 2.2.1 (Gregor Johann Mendel 1822-84). 修道会に入会し、所属した修道院は学術研究・教育が盛んだった。司祭になってから 51 年から 2 年間 Doppler の下で数学・物理学、ウンガーから植物生理学を学び、戻ってからえんどう豆の交配実験を行った。論文 [4] は数学的で抽象的な議論が理解されず、「反生物的」とさえ評された。後に井戸の水位や太陽の黒点などの気象との関係を研究し、没した時点ではむしろ気象学者としての評価のほうが高かった。

理論 2.2.2 (Hardy – Weinberg principle). すべての生物の因子型は、2 値変数 A, a の非順序対 $AA, Aa, aa \in [\{A, a\}]^2$ によって表現される。 A の遺伝子頻度を p として測られる。すると、3 元状態空間 $\{AA, Aa, aa\}$ 上の確率分布の世代毎の変化・発展が考え得る。このとき、十分大きな集団においては、親の母集団の因子型の組成がどうであろうとも、ランダムな交配は 1 世代以内に、「近似的に安定な因子型分布」を持つ [2]。

要諦 2.2.3. したがって、進化＝遺伝子頻度の時間発展の研究は、どのような要因で Hardy-Weinberg の法則が破られているかに関する研究となる。

理論 2.2.4 (性染色体). 染色体のうち、性決定に関与する、雄雌異体の生物で唯一異なる染色体。男性が XY であるから、性に関係する性質は父から息子へは遺伝し得ない。

色盲などのような性に伴う遺伝因子の多くは劣性で、 a の男性すべてと aa の女性すべてに現れる。したがって、起こる確率が男性で α ならば、女性では α^2 になる。

2.3 最尤法

Pearson のモーメント法は静的な解析であったが、最尤法は攻撃で動的で推論的である！ 22 年 [8] に系統的に提示する。

歴史 2.3.1 (赤池さんは Fisher の超克を意識していた). [12] において、

MLE は Fisher によって初めて詳細に論じられてから極めて広く用いられているが、その有効性の根拠が何にあるのかは明

^{†1} 考え方自体は、Edgeworth の "genuine inverse" や Pearson の段階ですでにあった。

らかではなく、その合理性の説明は通常直感に訴えてなされてきた。前項で論じた因子分析法には最尤法が用いられている。そこで、最尤法が何を最適化しようとしているのかが明らかになりさえすれば、先の問題は解決されることになる。

この後すぐに、漸近論的な立場から、 M -推定量のような発想「最尤法は $E[\log_e f(Y|\theta)]$ を最大にする θ の推定量である」さらに言い換えると、「KL-情報量を規準として、真の分布を最もよく近似する $f(y|\theta)$ を求めようとしている」と捉え直している。

AIC の導入は、平均対数尤度 $l(\theta)/N = \mathbb{P}_n[\log f(Y|\theta)]$ が、 $E[\log_e f(Y|\theta)]$ の推定値であるという着想の下に成立した。この見方に従えば、 $\{f(y|\theta)\}_{\theta \in \Theta}$ の族が Y の真の分布を与える $g(y)$ を含まない場合においても、最尤法が統計的モデルのパラメータ決定に有効なものでありうるということが容易にわかる。直交射影ではないが、集合 $\{f(y|\theta)\}_{\theta \in \Theta}$ の中で g に最も KL-距離の意味で近い点を選出する算譜なのだろう。つまり、赤池さんの言葉で言えば、最尤推定量 $\hat{\theta}$ はエントロピーに関して g を最もよく近似する $f(y|\theta)$ を与える θ の、観測値 y_1, \dots, y_N に基づく推定値を与える。ただし、適用条件は、平均対数尤度が $E[\log_e f(Y|\theta)]$ の推定値として有効な限りであり、またモデル $\{f(Y|\theta)\}$ も人間が勝手に自分の責任において取捨選択するものである。「Fisher が明らかにし得なかった尤度の本質的に実験的帰納的な性格と、統計理論における尤度利用の必然性とは、ここに客観的に描き出されている」。

2.4 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。

2.4.1 農事実験

歴史 2.4.1 (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 2.4.2. 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。

2.4.2 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 2.4.3. Fisher 18 [7] 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 [6] 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。

2.5 有意性検定

Fisher の有意性検定以前にも, Pearson の χ^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.

2.5.1 Fisher 以前の検定論

歴史 2.5.1 (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 2.5.2 (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 [4] 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 2.5.3 (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, χ^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, χ^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 [5].

歴史 2.5.4 (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, [6] 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. [6] 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 μ 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。

2.5.2 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。

2.6 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。

2.6.1 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。

2.6.2 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 3 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。

3.1 統計的品質管理の父

歴史 3.1.1 (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、統計の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。

3.2 Neyman-Pearson の統計的仮説検定理論

28, 33 にて、対立仮説、第 I,II 種の過誤、検出力の枠組みから Fisher の有意性検定を捉え直した。そして検定の「最適性」の枠組みをこしらえた、このメタ視点が達成点である。意思決定理論的だね。2 種類の過誤はトレードオフの形になっていることに注目すると、過誤確率の最適制御の問題になる。これについて、まず単純検定仮説・対立仮説なる対象に絞って、Neyman-Pearson の補題を立てた。

歴史 3.2.1 (Fisher の最尤法からアイデアを得て、検定論でも「最適性」を追求した。). 28[10] では、尤度比検定が導入された。33[11] では、最強力検定 (most efficient test) が定義された。複合仮説の場合については一様最強力検定が最適だろうが、これが構成できる状況は稀である。33 年以降の研究は、次善策の提案が続き、38 年まで共同研究を続ける。不偏検定、一様不偏検定、相似検定 (similar test) など。

Karl Pearson と Fisher の一番の違いは、誤差への関心の高さと、最適性への関心の高さの別である。この Fisher の立場は Neyman-Pearson に確かに受け継がれた。

統計的仮設検定の問題とは、棄却域を選択する問題である。

しかし、この立場に対する Fisher の反論としては、「有意性が棄却されるか」しかないのであって、その逆は「採択する」ではないはずである。これは帰納的推論の枠組みであり、この立場からは「形式主義すぎる」ということになる。一方で Neyman-Pearson は帰納的行動の枠組みと言える。いずれにしろ、検定の用途の違いである。

背景 3.2.2 (sampling inspection). 1924 年から Bell 研究所にて、Dodge, H. F. and Romig, H. G. が抜き取り検査の研究が始められた。誤って合否を下すことは、producer's risk と consumer's risk の狭間にある。つまり、Neyman-Pearson の応用先的には、どちらかに意思決定をする必要があったのだ！

最終的に、Stanford で品質管理の集中講義が行われ、1945 年夏の時点では各企業から派遣されて出席者は 1 万人を超えたという。

ここにおいて、Neyman-Pearson の検定技法は、契約上の品質に関する合意形成手段として用いられたのである！

このように、[3] の観点から見ると、統計手法に関する論争は全て応用先の要請の違いから起因したものであり、その解決は、「立場の違い自体をモデル化し、決定理論を作る」ことでアルゴリズム的に解決されてきた。

第 4 章

参考文献

参考文献

- [1] Mendel, G. J. (1866). “Versuche über Pflanzen-Hybriden” (“Experiments on Plant Hybridization”). *Proceedings of the Natural History Society of Brunn*.
- [2] Hardy, G. H. (1908). Mendelian Proportions in a Mixed Population, Letter to the Editor, *Science*, N. S., 28: 49-50.
- [3] 芝村良. (2004). 『R. A. フィッシャーの統計理論』. 九州大学出版会.
- [4] Edgeworth, F. Y. (1881). *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Science*.
- [5] 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.
- [6] Gosset, W. S. (1908). The Probability Error of a Mean.
- [7] Fisher, R. A. (1918). The Correlation between Relatives on the Supposition of Mendelian Inheritance.
- [8] Fisher, R. A. (1922). On the Mathematical Foundations of Theoretical Statistics.
- [9] Fisher, R. A. (1925). Theory of Statistical Estimation.
- [10] Neyman, J., and Pearson, E. S. (1928). On the Use and Interpretation of Certain Test Criteria for Purposes of Statistical Inference. *Biometrika*, 20A, 175-240.
- [11] Neyman, J., and Pearson, E.S. (1933). On the problem of the most efficient tests of statistical hypotheses. *Phil. Trans. R. Soc., Ser. A*, 231, 289 – 337.
- [12] 赤池広次 (1976) 「情報量規準 AIC とは何か」 数理科学 153 号, 5-11 ページ.
- [13] 安藤洋美 (1995) 『最小二乗法の歴史』. 現代数学社.
- [14] 長屋政勝, 金子治平, 上藤一郎 (1999) 『統計と統計理論の社会的形成』 (北海道大学図書刊行会, 統計と社会経済分析 I).