



連載TOPへ

【書籍化決定！】本コンテンツの公開は2015年5月7日までとなります。

[SHARE] [Tweet](#) [シェア](#) [BI](#)

第7回 正規分布という王様が誕生する

実験医学2014年11月号

はじめに

前回の 記事では、パラメトリック統計学に歩み入るための“お守り”として「確率変数」と「確率分布」を読者の皆さんにお渡ししました。観察されたデータという限られた情報源から、いかにして未知の母集団の属性を探ることができるのか。ある確率分布に従う確率変数という考え方は、母集団のふるまいをモデル化するためにパラメトリック統計学が提唱する基本方針といえます。

18世紀はじめにジャック・ベルヌーイによって打ち立てられた近代確率論は、偶然性に支配されたできごと（事象）を数学によって記述する という選択肢を研究者に選ばせました。それは同時に、得られた知見を数値化することによって客観性と普遍性をもたせるといふ知の歴史の大きな流れにも合致 していたに違いありません。

前回導入した確率分布という概念にはもっと説明すべきことがたくさん残されています。コインやサイコロを投げることだけが確率分布が 扱える問題ではありません。もっと普遍的に使えるツールとして確率分布の威力を発現するにはどうすればいいのでしょうか。今回はそれについて考えてみましょう。

§ 研究現場から統計学のリクツを見直す

初学者が統計学を学ぶとき、確率変数や確率分布はカリキュラムの最初の方で教えられるのが今でも普通でしょう。しかし、それは必ずしも効果的 な順序ではないかもしれません。確率論と統計学は、単に純粋学問的な興味から発展してきたのではなく、例えばサイコロの賭けでいかに効率を上げるか、の ような、むしろ具体的な個別問題を契機として積み上げられてきました。そういう個々の問題状況を理解することが統計的思考を身につけるうえでもっとも効果的でしょう。しかし、残念ながら、きれいに磨き上げられて干からびてしまった“数理統計学”からはそういうリアルな現場感覚の痕跡はとどめられていません。

確率統計はいくつ知っていればいいのですか？

後の回であらためて説明しますが、統計学の世界では100を超えるさまざまな確率分布が用いられています。そのなかには、次に説明する正規分布のように、きわめて頻繁に用いられるものもあれば、特殊な用途にのみ使われる確率分布もあります。どの分布を使うにしても私たち統計ユーザーにとって理解しなければならないのは、理論的な確率分布と現実世界との結びつきです。

パラメトリック統計学はたしかに現在では厳密な数学の理論体系として構築されています。しかし、統計ユーザーである私たちに求められているのは、干からびた統計理論をよくわからないまま鵜呑みにすることではなく、むしろそういう理屈がどのような現実的状況のもとで生まれ、発展していったのかという歴 史的経緯の理解でしょう。

羊土社HP会員

English page

ログインしていません

羊土社HP会員とは？ ログイン

書籍検索

実験医学の定期購読

最新号がWEBでも読める！

国内送料無料

月刊実験医学新刊

実験医学

次号予告

バックナンバー

連載一覧

掲載広告一覧

定期購読案内

詳細をみる

カートに入れる

実験医学増刊号新刊

実験医学

先制医療

次号予告

バックナンバー

掲載広告一覧

定期購読案内

詳細をみる

カートに入れる

実験医学 電子バックナンバー発売中

DIGITAL ARCHIVE

新着情報 人材・セミナー案内

東京大学大学院 医学系研究科 疾患生命工学センター 分子病態医科学部 門

平成28年度 大学院生（修士・博士）募集および説明会

詳細や他の情報はINFORMATIONコーナーをご覧ください

羊土社新刊・近刊

骨ベテア

サイトカイン増殖因子キーワード

Dr.北野の0から始めるシステムバイオロジー

詳細 購入

詳細 購入

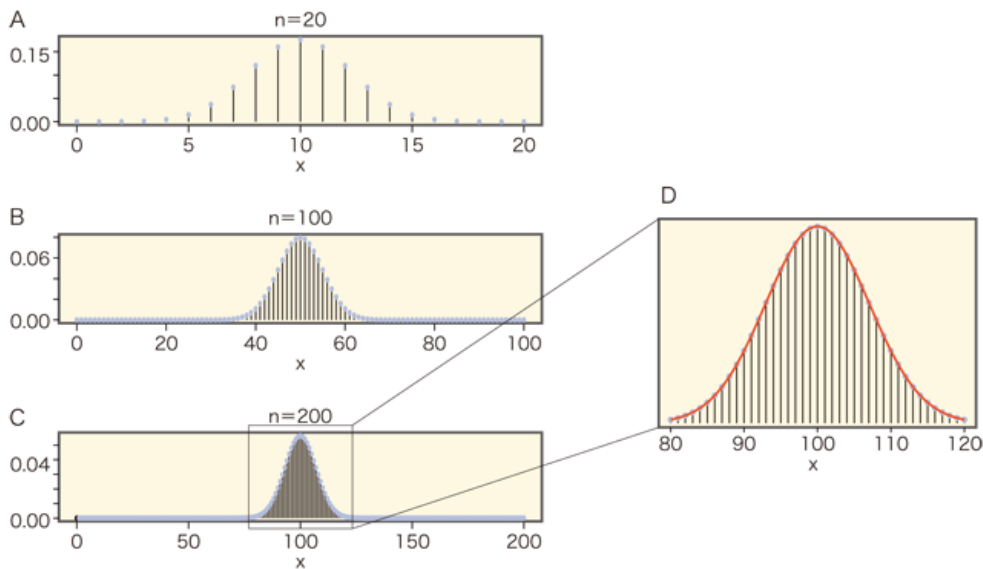
詳細 購入

>>新刊一覧へ

統計ツールのうっかり誤用を避けるためにも、歴史の理解は不可欠だと私は考えます^{1) 2)}。

§ 正規分布のひそやかな誕生

一つの例として、前回取り上げた二項分布の話の続けましょう。すでに説明したように、二項分布は、たとえばコイン投げ試行に代表される具体的な状況を確認の観点からモデル化したものです。整数値をとる二項分布の確率変数は、コインを投げる回数 n と表の出る確率 p によって厳密に決定されます。



☒ 二項分布から正規分布へ
A～C) 表と裏の出る確率がそれぞれ0.5であるコインを n 回 ($n=20$ (A), 100 (B), 200 (C)) 投げたときに表が出る回数 x の確率を明示したグラフ。D) Cの一部を切り抜き、正規分布の曲線グラフを重ね描きした(赤)。その関数は $1/\sqrt{100\pi} \exp\{-(x-100)^2/100\}$ となります。

では、ある p のもとで投げる回数 n をどんどん増やしていったら、この二項分布はどのようなになるでしょうか。☒A～Cは $n=20, 100, 200$ と増やしたときの二項分布のグラフです。最初はばらばらの棒グラフの集まりにしかみえなかった二項分布が、回数 n が増えるとともに、隣接する棒グラフがつながってしだいになめらかな連続するグラフにみえてきます。

ベルヌーイの著書『推測法』(1713)を読んだ数学者アブラハム・ド・モアブル (Abraham De Moivre: 1667-1754) は、ベルヌーイの提唱する二項分布の回数 n に関する極限分布について考察を重ね、1738年に出版された著書『偶然論』の中でその結果を公表しました。ド・モアブルは二項分布をする確率変数 x に関して n が無限大に発散したとき、 x の極限分布は

$$\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$$

という指数関数によって与えられることを証明しました。この指数関数こそ後に正規分布 (normal distribution) とよばれることになる確率分布の誕生です。

n の増大とともに二項分布が正規分布に近づいていく様子を見るには、 n を無限大まで発散させる必要はありません。☒Dを見ると $n=200$ と設定した場合でも二項分布は正規分布によって無理なく近似できることがわかるでしょう。

ド・モアブルはあくまでも二項分布の極限として正規分布の関数を導き出しました。しかし、正規分布の威力は実はもっと強力で、しかももっと広範囲に及ぶことが後の研究によって明らかになりました。その一つは、離散的な数値をとる確率変数から連続的な数値をとる確率変数への確率分布の一般化です。二項分布は整数値の確率変数に

学会売行き良好書情報

- 第37回 日本分子生物学会 年会 (14/12/02)

>>過去の売行き情報はこちら

実験医学 550号 突破!

アンケートに答えて

ポイントカードを

多数のご回答ありがとうございました

実験医学 @Yodosha_EM on Twitter

実験医学 jikkenigaku on Facebook

教科書・サブテキスト

をお探しの方へ

臨床医学系書籍 TOPページ(総合)

プライマリケアと救急を中心とした総合誌

レジデントノート

月刊 増刊

限定された確率分布でしたが、正規分布は一般の連続実数値の確率変数にも適用できます。この一般化はさらに一世紀下った19世紀初頭になって確立されました。

§ 正規分布のすこやかな成長

今から15年前の前世紀末1999年のこと、私はドイツの大学都市ゲッティンゲンに仕事で滞在したことがありました。中世の街の面影を残す石 畳の旧市街区（アルトシュタット）は、200年前の19世紀はじめに有名な数学者カール・フリードリッヒ・ガウス（Carl Friedrich Gauss：1777-1855）が活躍しました。しかし、ユーロ通貨に切り替わる前のドイツでは、ガウスは単に歴史的人物というだけではありませんでした。というのも、旧ドイツのマルクが流通していたころ、10マルク紙幣にはガウスの肖像とともに、彼が発見した正規分布曲線が描かれていたからです。

すでに説明したように、指数関数の一つである正規分布の関数そのものはド・モアブルによって18世紀前半に導かれていました。彼はあくまでも 二項分布の極限形としてそれを導出したのですが、そのままでは普遍的な利用は望めません。これに対して、ガウスが1809年にラテン語で出版した著書『太陽の周りを楕円軌道で公転する天体の運行に関する理論』において提出された正規分布関数は、もっと一般的に、観測値の誤差のふるまいを記述する数式としての役割を担っていました。

ガウスによる正規分布の理論に衝撃を受けたのは、彼と同時代のピエール-シモン・ラプラス（Pierre-Simon Laplace：1749-1827）でした。ラプラスは、ガウスの正規分布関数を用いることにより、データの総和や平均はデータ数が無限大になれば必ず 正規分布をするという定理、すなわち中心極限定理（central limit theorem）を証明しました。さらに、ラプラスは観測データから近似式を計算する最小二乗法（least square method）の前提として正規分布が必要であることも認識していました。

§ 王様の出自は実は庶民だった

このように、18世紀前半から19世紀前半の一世紀の間に、ド・モアブル、ガウス、そしてラプラスらの研究を通じて、その後の理論統計学を制 覇することになる正規分布の初期理論はすべて構築されました。賭け事の数理にはじまる確率論と統計学の歴史は、発展しつつあった数学の力を存分に駆使して、現実世界の不確定な現象を数理の観点からアプローチするという新たな展開を見せることになります。そして、誤差関数としての正規分布関数の評価は急速 上がっていきま



日常生活に密着していたマルク紙幣に正規分布曲線が描かれていたという事実は、ドイツの国民性が厳密な論理を身近に感じていたという点だけではなく、理論統計学のルーツがそもそも日常生活空間のなかにあったのだということを現代の私たちに再認識させてくれます。

では、統計学の理論的基盤として運命づけられた正規分布は、その後いったいどのような発展を遂げることになるのでしょうか。正規分布がもついくつかの強力な性質のおかげで、パラメトリック統計学の理論が数学的体系として構築できたという点は強調すべきでしょう。次回ではこのあたりのこととお話 することにしましょう。

文献

- 1) Theodore M Porter：The Rise of Statistical Thinking, 1820-1900. Princeton University Press, 1988
- 2) Stephen M Stigler：The History of Statistics: The Measurement of Uncertainty before 1900. Harvard University Press, 1986

[SHARE] Tweet シェア

「第8回 ピアソンが築いたパラメトリック統計学の礎石」は、本誌2014年12月号を御覧ください

本記事の掲載号



実験医学 2014年11月号 Vol.32 No.18
化学感覚と脳

東原和成／企画
定価 2,000円＋税、2014年10月発行
[▶詳細](#) [▶購入](#)

本連載に関する質問・感想、統計に関する具体的な悩みを編集部までお寄せください！

- 下記画像中の英数字をご入力ください



[画像を変更する](#)

おすすめ書籍



[▶詳細](#) [▶購入](#)



[▶詳細](#) [▶購入](#)



[▶詳細](#) [▶購入](#)



[▶詳細](#) [▶購入](#)



[▶詳細](#) [▶購入](#)

[会社案内](#) | [採用情報](#) | [個人情報取扱い](#) | [お問い合わせ](#) | [広告掲載について](#)

(C)2014 YODOSHA CO., LTD. All Rights Reserved.