文系学生に対する2項分布とポアソン分布の数理の教え方

青山学院大学社会情報学部 寺尾 敦

本稿では、2項分布とポアソン分布の確率関数、平均、分散の式導出について、文系学生への統計学教育でどのように教えることが適切かを吟味する。いくつかの方法が考えられる場合には、それらを比較検討する。確率分布の平均と分散の計算式は、積率母関数を用いれば統一的な方法で求めることができるが、文系学生にこれを教えることは難しいと思われるので、本稿では述べていない。

最終的な目的は、統計学を教える教員が利用できる事典あるいはハンドブック（たとえば、遠山、1965）の作成である。一般のテキストのように特定の導出方法だけを示すのではなく、可能な方法をいくつか示し、教員が教え方を選択するときに役立つようにしたい。

2項分布

2項分布の確率関数

2項分布の確率関数 $P(X=x) = {}_n C_x p^x (1-p)^{n-x}$ は、試行数が n 、成功確率が p のベルヌーイ試行で x 回成功する確率を考えることで、容易に導入できる。筆者は所属する学部での1年次必修科目「統計入門」でこれを教えている。受講者の半数以上は文系学生だが、大きな困難を感じたことはない。

2項分布の平均と分散

2項分布の平均と分散は、ベルヌーイ分布（試行数が1回の2項分布）の平均と分散を定義にしたがって計算し、確率変数の和の平均の公式 $E[X_1 + X_2] = E[X_1] + E[X_2]$ と、独立な確率変数の和の分散の公式 $V[X_1 + X_2] = V[X_1] + V[X_2]$ を使うことで、簡単に求められる（篠原、1989；野田・宮岡、1990）。第 i 試行 ($i=1,2,\dots,n$) での結果を確率変数 X_i で表し、結果が成功の場合は $X_i=1$ 、失敗の場合は $X_i=0$ とする。この確率変数 X_i はベルヌーイ分布に従い、その平均は

$E[X_i] = 1 \times p + 0 \times (1-p) = p$ 、分散は

$$V[X_i] = (1-p)^2 \times p + (0-p)^2 \times (1-p) = (1-p)\{(1-p)p + p^2\} = p(1-p)$$

となる。確率変数 X を、全部で n 個の確率変数の和 $X = \sum_{i=1}^n X_i$ で定義すると、 X は n 回のベルヌーイ試行での成功回数を表す。したがって、確率変数 X は2項分布に従い、平均は

$$E[X] = E\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n E[X_i] = np, \text{ 分散は } V[X] = V\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n V[X_i] = np(1-p) \text{ となる.}$$

2 項分布の確率関数を用いて、平均と分散を定義に従って計算することは少し厄介である（東京大学教養部統計学教室，1991：ホーエル，1978）．平均は、

$$E[X] = \sum_{x=0}^n x \cdot {}_n C_x p^x (1-p)^{n-x} = \sum_{x=1}^n x \cdot \frac{n}{x} \cdot {}_{n-1} C_{x-1} (p \cdot p^{x-1}) (1-p)^{n-x} = np \sum_{x=1}^n {}_{n-1} C_{x-1} p^{x-1} (1-p)^{(n-1)-(x-1)}$$

と式変形して、 $y = x-1$ とおくと、 $E[X] = np \sum_{y=0}^{n-1} {}_{n-1} C_y p^y (1-p)^{(n-1)-y} = np$ となる．分散は、

$$\begin{aligned} E[X(X-1)] &= \sum_{x=0}^n x(x-1) {}_n C_x p^x (1-p)^{n-x} = \sum_{x=2}^n x(x-1) \cdot \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\ &= \sum_{x=2}^n \frac{n!}{(x-2)!(n-x)!} p^x (1-p)^{n-x} = n(n-1)p^2 \sum_{x=2}^n \frac{(n-2)!}{(x-2)!(n-x)!} p^{x-2} (1-p)^{(n-2)-(x-2)} \end{aligned}$$

と式変形して、 $y = x-2$ とおくと、 $E[X(X-1)] = n(n-1)p^2 \sum_{y=0}^{n-2} {}_{n-2} C_y p^{n-2} (1-p)^{(n-2)-y} = n(n-1)p^2$

を得る．あとは、 $V[X] = E[X(X-1)] + E[X] - \{E[X]\}^2 = np(1-p)$ として分散が求められる．

小針（1973）は、2 項分布の平均と分散を一続きの論理で示すエレガントな方法を示している．等

式 $(p+q)^n = \sum_{x=0}^n {}_n C_x p^x q^{n-x}$ の両辺を p について偏微分すると、 $n(p+q)^{n-1} = \sum_{x=0}^n x \cdot {}_n C_x p^{x-1} q^{n-x}$ とな

る．この両辺に p をかけると、 $np(p+q)^{n-1} = \sum_{x=0}^n x \cdot {}_n C_x p^x q^{n-x}$ を得る．ベルヌーイ試行での成功確

率を p 、失敗確率を q で表すと、この式の右辺は $E[X]$ であり．左辺は $p+q=1$ に注意して np

となる．小針（1973）も注意しているように、 p での偏微分で $q=p-1$ を無視していることが気にか

かかるが、偏微分する等式は任意の p と q に対して成立するので問題ない．分散については、平均の

導出過程で得られた等式 $np(p+q)^{n-1} = \sum_{x=0}^n x \cdot {}_n C_x p^x q^{n-x}$ の両辺を p についてもう一度偏微分する

と $n(n-1)p(p+q)^{n-2} + n(p+q)^{n-1} = \sum_{x=0}^n x^2 \cdot {}_n C_x p^{x-1} q^{n-x}$ となる．この等式の両辺に p をかけると

$n(n-1)p^2(p+q)^{n-2} + np(p+q)^{n-1} = \sum_{x=0}^n x^2 \cdot {}_n C_x p^x q^{n-x}$ が得られる．ここで p と q をそれぞれベルヌーイ試行での成功確率と失敗確率であるとすると、この等式の右辺は $E[X^2]$ である．左辺で

$p+q=1$ に注意して、 $V[X] = E[X^2] - \{E[X]\}^2 = n(n-1)p^2 + np - (np)^2 = np(1-p)$ が得られる．

ポアソン分布

ポアソン分布の確率関数

ポアソン分布の確率関数 $P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$ の導出は、テキストによって細部の書き方は異なるものの、実質的にはどれも同じである。2 項分布の確率関数において、2 つのパラメータの積 np を一定の正数 λ に保って、 $n \rightarrow \infty$, $p \rightarrow 0$ の極限をとる。いくつかのテキストでは、 np を一定の値とするのではなく、 $np \rightarrow \lambda$ としている（たとえば、稲垣, 2003; 尾畑, 2014; 東京大学教養部統計学教室, 1991）。

文系学生にポアソン分布の導出を教えるときの難しさは、極限 $\lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda}$ が出てくるところにある。こういうものだと思って受け入れてもらうか、高校数学で学習するネイピア数の定義のひとつ $\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e$ で $n \rightarrow -\infty$ としても e に収束することを示してこれを利用する。

ポアソン分布の確率関数を覚えておくためには、2 項分布からの導出を覚えておくか、関数 e^λ の $\lambda = 0$ におけるテイラー展開 $e^\lambda = 1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \dots$ の両辺に $e^{-\lambda}$ をかけると右辺が $P(X = 0) + P(X = 1) + P(X = 2) + \dots$ となることを覚えておく。これはもともと、ポアソン分布の確率関数において $\sum_{k=0}^{\infty} P(X = k) = 1$ を確認する計算である。テイラー展開の理論は文系学生にとって難しいが、展開式は受け入れてもらえるようである。

ポアソン分布の平均と分散

ポアソン分布の平均と分散は、2 項分布の平均 $E[X] = np$ と分散 $V[X] = np(1 - p)$ の極限を考えることで容易にわかる（ホーエル, 1978）。文系学生への教育では、この方法がもっともよいだろう。平均は定義に従って計算しても比較的容易に得られる（東京大学教養部統計学教室, 1991）。

定義に従って分散を計算することは少し厄介である。よく知られた方法は、 $E[X(X - 1)] = \sum_{x=0}^{\infty} x(x-1) \frac{\lambda^x e^{-\lambda}}{x!} = \sum_{x=2}^{\infty} x(x-1) \frac{\lambda^x e^{-\lambda}}{x!} = \lambda^2 e^{-\lambda} \sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{(x-2)!} = \lambda^2 e^{-\lambda} e^\lambda = \lambda^2$ を利用して、 $V[X] = E[X(X - 1)] + E[X] - \{E[X]\}^2 = \lambda$ と計算するものである（篠原, 1989; 白石, 2012）。ここで用いた等式 $E[X^2] = E[X(X - 1)] + E[X]$ は 2 項分布の分散の計算でも用いた。

小針(1973)は、ポアソン分布の平均と分散を一続きの論理で示すエレガントな方法を示している。

関数 e^λ の $\lambda=0$ におけるテイラー展開 $e^\lambda = \sum_{x=0}^{\infty} \frac{\lambda^x}{x!}$ の両辺を λ について微分すると、右辺は λ

のべき級数だから項別微分ができてその結果は左辺の微分に一致し、 $e^\lambda = \sum_{x=0}^{\infty} \frac{x\lambda^{x-1}}{x!}$ となる。両辺に

$\lambda e^{-\lambda}$ をかけると、 $\lambda e^{-\lambda} e^\lambda = \sum_{x=0}^{\infty} x \frac{\lambda^x e^{-\lambda}}{x!}$ となる。この左辺の値は λ 、右辺は $E[X]$ である。分散を

求めるには、まず $\lambda e^\lambda = \sum_{x=0}^{\infty} \frac{x\lambda^x}{x!}$ の両辺を λ について微分して、 $\lambda e^\lambda + e^\lambda = \sum_{x=0}^{\infty} \frac{x^2 \lambda^{x-1}}{x!}$ を得る。この

両辺に $\lambda e^{-\lambda}$ をかけると、 $\lambda^2 + \lambda = \sum_{x=0}^{\infty} x^2 \frac{\lambda^x e^{-\lambda}}{x!} = E[X^2]$ を得る。あとは、 $V[X] = E[X^2] - \{E[X]\}^2$

より、分散を求めることができる。

小針(1973)の方法はエレガントで、微分操作そのものは簡単だから文系学生にも理解可能だろう。

微分を利用して級数の和を求める問題は大学入試でも出題されているから (河田, 2017, p.67-68),

こうした解法を学習した経験のある学生もいるかもしれない。ただし、この解法はべき級数の項別微

分の定理を前提としていることに注意が必要である。項別微分の定理を文系学生に教えることはしな

くても、こうした操作が常に許されるわけではないということは教えておくべきだろう。高校生向け

の受験参考書である清 (2005, p.163) はこの注意を与えている。

引用文献

- 稲垣宣生 (2003) 数理統計学 (改訂版) 裳華房
尾畑伸明 (2014) 数理統計学の基礎 共立出版
小針暁宏 (1973) 確率・統計入門 岩波書店
河田直樹 (2017) 大学入試 無限級数の解法研究 聖文新社
篠原昌彦 (1989) 確率・統計 朝倉書店
白石高章 (2012) 統計科学の基礎 日本評論社
清史弘 (2005) 分野別 受験数学の理論 6. 数列 駿台文庫
東京大学教養部統計学教室 (編) (1991) 統計学入門 東京大学出版会
遠山啓 (編) (1965) 現代数学教育事典 明治図書
野田一雄・宮岡悦良 (1991) 入門・演習 数理統計 共立出版
ホーエル, P. G. (1978) 入門数理統計学 培風館