

「確率とその応用」ノート、その肆

逆瀬川浩孝

4 期待値

例題 4.1 (確率関数と期待値) 確率関数が次のように与えられている確率変数 X の期待値を計算しなさい

$$P(X = -1) = 0.4, P(X = 0) = 0.3, P(X = 1) = 0.2, P(X = 2) = 0.1$$

こたえ 「期待値は値 × 確率の和」を忠実に実行して

$$E(X) = -1 \times 0.4 + 0 \times 0.3 + 1 \times 0.2 + 2 \times 0.1 = 0$$

例題 4.2 (確率変数の関数の期待値) 確率変数 X の確率関数は次のように与えられている。このとき、 $Y = X^2$ の期待値を計算しなさい。

$$P(X = -2) = P(X = -1) = P(X = 0) = P(X = 1) = P(X = 2) = 0.2$$

こたえ Y の確率関数は次のように計算できる。

$$\begin{aligned} P(Y = 0) &= P(X = 0) = 0.2, \\ P(Y = 1) &= P(X = 1) + P(X = -1) = 0.4 \\ P(Y = 4) &= P(X = 2) + P(X = -2) = 0.4 \end{aligned}$$

したがって、

$$E(Y) = 0.4 + 1.6 = 2$$

計算の過程を分解すると、

$$\begin{aligned} E(Y) &= 0 \times P(Y = 0) + 1 \times P(Y = 1) + 4 \times P(Y = 4) \\ &= 1^2 \times P(X = 1) + (-1)^2 \times P(X = -1) \\ &\quad + 2^2 \times P(X = 2) + (-2)^2 \times P(X = -2) \\ &= \sum_{k=-2}^2 k^2 P(X = k) \end{aligned} \tag{4.1}$$

と書ける。定義ではないが、 $Y = g(X)$ という確率変数の期待値は

$$E(Y) = \sum_k g(k) P(X = k) \tag{4.2}$$

としてもよい。

$E(X^2) \neq (E(X))^2$ であることに注意。

練習問題 4.1 確率関数が下のように入えられている確率変数 X に対して、 $Z = \max(X - 1, 0)$ と定義します。(1) Z の取りうる値はいくつですか。(2) Z の確率関数を求めなさい。(3) Z の期待値を計算しなさい。

$$f_X(0) = 0.2, f_X(1) = 0.3, f_X(2) = 0.2, f_X(3) = 0.15, f_X(4) = 0.1, f_X(5) = 0.05$$

練習問題 4.2 (定義関数) 事象 A の起きる確率を p とします。確率変数 X は、事象 A が起きたら 1、さもなければ 0 という値を取るものとします。このとき、 X の期待値 (平均) と分散を計算しなさい (X は A の定義関数と呼ばれ、 1_A と書かれることが多い)。

練習問題 4.3 X はパラメータ p のベルヌイ分布に従う確率変数としたとき、 $|X - p|$ の期待値を計算して、 X の標準偏差 $S(X) = \sqrt{E((X - p)^2)}$ と比較しなさい。普通の数ならば $|X| = \sqrt{X^2}$ ですが、 $|E(X)| = \sqrt{E(X)^2}$ は成り立ちますか？

ヒント: 「比較しなさい」？ p がいくつの時どっちが大きいとか、その差はどれくらいとか...
まずは両者を p の関数としてグラフ表示するところから始めなさい。

例題 4.3 (幾何分布) p は $0 < p < 1$ を満たす定数としたとき、次の関数が確率関数になることを示し、確率変数 X の期待値を計算しなさい (X はパラメータ p の**幾何分布**に従うと言われます)

$$f(i) = P(X = i) = p(1 - p)^{i-1}, \quad i = 1, 2, \dots$$

こたえ $1 + r + r^2 + r^3 + \dots = (1 - r)^{-1}$ という公式を使えば、 $f(1) + f(2) + \dots = 1$ はすぐに分かるでしょう。「期待値は値 $\times$ 確率の和」を計算して

$$\begin{aligned} E(X) &= \sum_{i=1}^{\infty} i \times p(1 - p)^{i-1} = \sum_{i=1}^{\infty} \sum_{j=1}^i p(1 - p)^{i-1} \\ &= \sum_{j=1}^{\infty} p(1 - p)^{j-1} \sum_{i=j}^{\infty} (1 - p)^{i-j} = \sum_{j=1}^{\infty} (1 - p)^{j-1} = \frac{1}{p} \end{aligned}$$

練習問題 4.4 (2項分布) p は $0 < p < 1$ を満たす定数、 n は自然数としたとき、次の関数が確率関数になることを示し、確率変数 X の期待値を計算しなさい (X はパラメータ n, p の**2項分布**に従うと言われます)

$$f(i) = P(X = i) = \binom{n}{i} p^i (1 - p)^{n-i}, \quad i = 0, 1, 2, \dots, n$$

練習問題 4.5 (ポワソン分布) a を正の定数としたとき、次の関数が確率関数になることを示し、確率変数 X の期待値を計算しなさい (X はパラメータ a の**ポワソン分布**に従うと言われます)

$$f(i) = P(X = i) = \frac{a^i}{i!} e^{-a}, \quad i = 0, 1, 2, \dots$$

テキスト 121 ページの説明にあるように、ポワソン分布は (航空機事故のように) 滅多に起きない事象の起きる回数を説明するためによく使われます。1 秒ごとに事故が起きたら 1、さもなければ 0 という実験を繰り返すと、1 年間 $3600 \times 24 \times 365 = 31536000$ 秒のなかで 1 の回数が年間の事故数です。これはちょうど、内閣支持率調査の考え方と同じですから、1 秒間で事故を起こす確率 (これが「支持する」と答える確率にあたる) を p とすると、年間の事故数を表す確率変数を X とすると、 X はパラメータ $31536000, p$ の 2 項分布に従います。2 項分布といっても、 X の値の範囲はせいぜい 100 止まりでしょう。そうすると、一々丁寧に 2 項係数を使って 2 項分布の確率関数を計算するよりは近似計算した方が簡単です。それがポワソン分布なのです。

本のミスプリントも事故と同じように考えることが出来ます。「確率とその応用ノート」の文字数は全部で 5 万くらいです。ミスプリントが 20 箇所位あるとすれば、1 文字を打つときにミスの確率は 2500 分の 1 ($= p$) くらいのもので、1 ページの文字数は 1000 字くらいですから、

各ページに含まれるミスプリントの数は、パラメータ 1000, p の 2 項分布にしたがいます。それを上のように考えると、パラメータ 0.4 のポワソン分布に従うといっても良い、ということになります。

例題 4.4 (密度関数と期待値) 確率変数 X の密度関数が次の式で与えられているとき、定数 c の値を求め、期待値を計算しなさい。

$$f_1(x) = \begin{cases} c(1-x^2), & -1 \leq x \leq 1 \\ 0, & \text{それ以外} \end{cases}$$

$$f_2(x) = \begin{cases} c(2-|x|), & -2 \leq x \leq 2 \\ 0, & \text{それ以外} \end{cases}$$

こたえ 全域での積分が 1 になるという制約条件によって c が決まる。あとは定義通り計算するだけ。

● $f_1(x)$ の場合

$$c \int_{-1}^1 (1-x^2) dx = c \left(2 - \frac{2}{3} \right) = 1 \Leftrightarrow c = \frac{3}{4}$$

したがって、期待値は

$$E(X) = \frac{3}{4} \int_{-1}^1 x(1-x^2) dx = 0$$

● $f_2(x)$ の場合

$$\begin{aligned} c \int_{-2}^2 (2-|x|) dx &= c \int_0^2 (2-x) dx + c \int_{-2}^0 (2+x) dx \\ &= c(4-2) + c(4-2) = 4c \Leftrightarrow c = \frac{1}{4} \end{aligned}$$

したがって、期待値は

$$E(X) = \frac{1}{4} \int_{-2}^2 x(2-|x|) dx = \frac{1}{4} \int_{-2}^2 x|x| dx = 0$$

いずれも偶関数なので、期待値は 0、という解答でもよい。

例題 4.5 (累積分布関数と期待値) 確率変数 X の累積分布関数が下のようになら与えられているとき、 X の期待値 $E(X)$ を計算しなさい。

$$F_1(x) = \begin{cases} 0, & x < 0 \\ 2x - x^2, & 0 \leq x \leq 1 \\ 1, & x > 1 \end{cases}$$

$$F_2(x) = \begin{cases} 0, & x < 0 \\ \sqrt{x}, & 0 \leq x \leq 1 \\ 1, & x > 1 \end{cases}$$

こたえ まず、密度関数を計算してから、定義にしたがって期待値を計算する。定義域はいずれも $[0, 1]$ 。

$$f_1(x) = 2 - 2x \tag{4.3}$$

より、

$$\int_0^1 x f_1(x) dx = 2 - 1 = 1 \tag{4.4}$$

$F_2(x)$ については、

$$f_2(x) = \frac{1}{2\sqrt{x}} \quad (4.5)$$

より、

$$\int_0^1 x f_2(x) dx = \frac{1}{2} \int_0^1 \sqrt{x} dx = \frac{1}{3} \quad (4.6)$$

例題 4.6 (指数分布) 累積分布関数が次のように与えられている確率変数 X の期待値を計算しなさい。

$$F(x) = \begin{cases} 1 - e^{-2x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

こたえ 期待値を計算するために、まず密度関数を求める。

$$f(x) = \frac{d}{dx} F(x) = \begin{cases} 2e^{-2x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

あとは「期待値は値 × 確率の積分」を忠実に計算するだけ

$$E(X) = \int_0^{\infty} x \times 2e^{-2x} dx = [-xe^{-2x}]_0^{\infty} + \int_0^{\infty} e^{-2x} dx = \frac{1}{2}$$

コメント 確率変数が正値 ($X \geq 0$) の場合の期待値は

$$\begin{aligned} E(X) &= \int_0^{\infty} x f(x) dx = \int_0^{\infty} \left(\int_0^x du \right) f(x) dx = \int_0^{\infty} \left(\int_u^{\infty} f(x) dx \right) du \\ &= \int_0^{\infty} (1 - F(u)) du \end{aligned}$$

によって「も」計算できる。したがって、

$$E(X) = \int_0^{\infty} e^{-2x} dx = \frac{1}{2}$$

例題 4.7 (正規分布) X はパラメータ μ, σ の正規分布に従う確率変数としたとき、その期待値と分散を計算しなさい。

こたえ 定義通りに計算する。

$$E(X) = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)} dx$$

$z = \frac{x-\mu}{\sigma}$ と変数変換すると、

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} (\sigma z + \mu) \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \\ &= \sigma \int_{-\infty}^{\infty} z \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz + \mu \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \end{aligned}$$

第一項の積分は $ze^{-z^2/2}$ が奇関数なので 0。第二項の被積分関数は、平均 0、分散 1 の正規分布の密度関数なので、その定積分は 1。したがって、 $E(X) = \mu$ 。

分散も、同じ変数変換を利用して、定義通りに計算できる。

$$\begin{aligned}
 V(X) &= E((X - \mu)^2) \\
 &= \int_{-\infty}^{\infty} (x - \mu)^2 \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)} dx, \quad \left(z = \frac{x - \mu}{\sigma}\right) \\
 &= \sigma^2 \int_{-\infty}^{\infty} z^2 \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \\
 &= \sigma^2 \int_0^{\infty} z^2 \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz + \sigma^2 \int_{-\infty}^0 z^2 \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \\
 &= 2\sigma^2 \int_0^{\infty} z^2 \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz
 \end{aligned}$$

この定積分は部分積分を使って計算する。

$$\int_0^{\infty} z^2 \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = \frac{1}{\sqrt{2\pi}} \left[-ze^{-z^2/2} \right]_0^{\infty} + \frac{1}{\sqrt{2\pi}} \int_0^{\infty} e^{-z^2/2} dz = \frac{1}{2}$$

したがって、 $V(X) = \sigma^2$ 。

例題 4.8 (正規分布) X はパラメータ μ, σ の正規分布に従う確率変数としたとき、 $Z = \frac{X - \mu}{\sigma}$ と定義すると、 Z は標準正規分布に従うことを示しなさい。

こたえ Z の累積分布関数は

$$\begin{aligned}
 P(Z \leq x) &= P\left(\frac{X - \mu}{\sigma} \leq x\right) = P(x \leq \sigma x + \mu) \\
 &= \int_{-\infty}^{\sigma x + \mu} \frac{1}{\sqrt{2\pi}\sigma} e^{-(u-\mu)^2/(2\sigma^2)} du
 \end{aligned} \tag{4.7}$$

と計算されるので、その密度関数は微分によって求めることができる。

$$\frac{d}{dx} P(Z \leq x) = \sigma \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2} = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \tag{4.8}$$

これは平均 0、分散 1 の正規分布の密度関数に他ならない。

ある確率変数 X に対して、 $(X - E(X))/\sqrt{S(X)}$ という確率変数は、確率変数の標準化、あるいは基準化という。基準化された確率変数の平均は 0、分散は 1 である。

例題 4.9 (確率変数の関数の期待値) パラメータ 0, 1 の正規分布、すなわち標準正規分布に従う確率変数 X に対して、 $Z = \max\{X, 0\}$ と定義したとき、確率変数 Z の期待値を計算しなさい。

こたえ 「期待値は値 × 確率の積分」を忠実に計算するだけ

$$\int_0^{\infty} x \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \frac{1}{\sqrt{2\pi}} \left[-e^{-x^2/2} \right]_0^{\infty} = \frac{1}{\sqrt{2\pi}}$$

練習問題 4.6 (三角分布) 確率変数 X の密度関数は、区間 $[a, b]$ の外では 0、点 $(a, 0), (c, d), (b, 0)$ を結んだ三角形状をしています (ただし、 $a < c < b, d > 0$ とします)。このとき、 X の期待値が $(a + b + c)/3$ になることを示しなさい。

例題 4.10 (一様分布、確率変数の関数の密度関数) 確率変数 Y は、区間 $[-1, 1]$ 上で一様分布するものとします。また、 $Z = 2Y + 2$ と定義します。(1) Y の密度関数 $f_Y(x)$ と、 Z の密度関数 $f_Z(x)$ の略図を描きなさい (式に表せるように、主要な点の座標を描くこと)。(2) Y の期待値を計算しなさい。(3) Z の期待値を計算しなさい。

こたえ (1) 一様分布の密度関数は定義された区間で同じ値を取る。積分値が 1 になることから、その値は区間の幅の逆数に等しい。したがって、

$$f_Y(x) = \begin{cases} 0.5, & -1 \leq x \leq 1 \\ 0, & \text{それ以外} \end{cases}$$

Z は Y が -1 から 1 へ変化するにつれて 0 から 4 へ動く。 Y の値を 2 倍にして 2 だけずらしたもののだから、一様分布であることには変わりがない。したがって、

$$f_Z(x) = \begin{cases} 0.25, & 0 \leq x \leq 4 \\ 0, & \text{それ以外} \end{cases}$$

式で導くには、累積分布関数を計算すればよい

$$P(Z \leq x) = P(2Y + 2 \leq x) = P\left(Y \leq \frac{x}{2} - 1\right)$$

密度関数を求めるには x で微分すればよいので、

$$f_Z(x) = \frac{d}{dx} P(Z \leq x) = \frac{d}{dx} P\left(Y \leq \frac{x}{2} - 1\right) = \frac{1}{2} f_Y\left(\frac{x}{2} - 1\right) = \frac{1}{4}$$

(2)(3) 期待値は定義通り

$$\begin{aligned} E(Y) &= \int_{-1}^1 x \frac{dx}{2} = 0 \\ E(Z) &= \int_0^4 x \frac{dx}{4} = \frac{16}{8} = 2 \end{aligned}$$

コメント $Z = 2Y + 2$ ならば $f_Z(x) = 2f_Y(x) + 2$ 、という「思い込み」をする人が多い。そのまま計算すると、 Z の密度関数は区間 $[-1, 1]$ で高さ 3 に、それ以外で 2 なる！！ その「密度関数」は積分したら ∞ になってしまう、この辺りで気がついてほしいのだが。

練習問題 4.7 確率変数 U は区間 $[0, 1]$ で一様分布する確率変数、 X は $X = aU + b$ によって定義された確率変数としたとき、 X の期待値と分散を計算しなさい

練習問題 4.8 X は区間 $[a, b]$ で一様分布する確率変数としたとき、その期待値と分散を計算しなさい。

練習問題 4.9 悪天候で工事が中止になると A の損害が発生する。工事が中止にならないければ、もちろん損害額は 0 。明日、工事が中止になるほどの悪天候になる確率は p だという。さて、明日の損害額の期待値を計算しなさい。

練習問題 4.10 (続き) 工事が中止になったらその損害額 A を補償しましょう、その保険料として B 払ってください、という保険に入ったとする。(1) 出費合計(保険料マイナス補償額(=損害額))の期待値を求めなさい。(2) 保険料がいくらだったら保険を利用しますか。

練習問題 4.11 「あたり」と書いた 1 枚の紙と 9 枚の白紙とを箱に入れ、 10 人に順番に取り出してもらい、という抽選を行うと考えてください(プロ野球のドラフトのようなもの)。引いた人はその場で当たりはずれを確かめるものとし、当たりくじが出たら、そこでくじ引きは終わるものとします。このとき、(1) くじを引く人数の期待値を計算しなさい。(2) ある人のくじに当たる確率を計算しなさい。(3) くじを引く人数と、ある人が当たりくじを引く確率の関係を説明しなさい。

例題 4.11 (年金のモデル) ある年金プランでは、 m 才まで毎年 a ずつ払い続け、 $m+1$ 才から b ずつ受け取るという契約になっています。支払い、受け取りは誕生日に生きていれば決済されるものとします。現在 k 才の人の余命を確率変数と考え X と置き、その確率関数を $f(x)$ とします。ただし、余命は 1 年未満の日数を切り捨てて数えることにします。つまり、35 年と 1 日生きる人も、35 年と 364 日生きる人も余命は同じ 35 年、とします。以下の問いに答えなさい。ただし、誕生日にだけ加入でき、加入すると直ちに 1 回目の支払いが発生するものとします。(1) 現在 k 才の人が一生に払う年金積立額を X, a, m, k で表しなさい。(2) 現在 k 才の人が一生に受け取る年金額を X, b, m, k で表しなさい。(3) この年金プランに k 才で契約したとき、生涯の収支 (受取年金総額から年金積立額を引いたもの) を Z としたとき、 Z を X, a, b, m, k で表しなさい。(4) Z の期待値を $f(x)$ を使って表しなさい。(5) 人の寿命分布を $g(x)$ としたとき、 Z の期待値を $g(x)$ を使って表しなさい。

こたえ キャッシュフロー (お金の流れ) を描いてみればよい。長生きする場合、 $k, k+1, \dots, m$ で $-a, m+1, m+2, \dots, k+X$ で $+b$ という流れがある。 $m+1$ 才になる前に死ねば、キャッシュフローは $-a$ だけ、そうなるのは $X \leq m-k$ の場合。ということを理解しておけば、分かるはず。

(1) 積立額は X と $m-k$ の小さい方に a を掛けた額プラス a (契約時に払うお金)、すなわち $a(\min\{X, m-k\} + 1)$ 。(2) $X \geq m-k+1$ の場合だけ、 $b(k+X-m)$ だけのお金を受け取ることができる (k も必要です)、したがって $b \max\{X+k-m, 0\}$ 。(3)

$$Z = b \max\{X+k-m, 0\} - a(\min\{X, m-k\} + 1)$$

(4) 期待値の定義式を使って

$$E(Z) = b \sum_{i \geq m-k} (i+k-m)f(i) - a \sum_{i=0}^{m-k} if(i) - a - a(m-k) \sum_{i > m-k} f(i)$$

(5) 年齢を A 、寿命を Y とすると、 Y と $A=k$ という条件の下での余命 X の関係は

$$P(X=i | A=k) = \frac{P(Y=k+i)}{P(Y \geq k)} = \frac{g(k+i)}{1-G(k-1)} = f(i)$$

と表される。ただし、 $G(k)$ は寿命の累積分布関数 ($= g(0) + g(1) + \dots + g(k)$) を表す。(4) で得られた結果にこの式を代入するだけ。

コメント こういう問題を考える場合は、具体的な数を使って、いろいろな場合を想定しながらシミュレーションしてみる必要があります。 $m=64$ として、 $k=63$ 、 $X=1$ としたら、2 回支払って終わり、 $X=2$ としたら支払いは 2 回、その後 1 回受け取って終わり、 $X=3$ としたら支払いは $X=2$ の場合と同じ 2 回、その後 2 回受け取って終わり、...、「ああそうか」と分かるまでいろいろなケースを試算し、支払いは $\min\{X, m-k\} + 1 (= A)$ 回、受け取りは $\max\{X+k-m, 0\} (= B)$ という式が得られたら、もう一度最初に戻って、その式が本当に正しいか、検算してみる、という手順が必要です。 $m=64, k=63, X=1$ ならば確かに $A=2, B=0$ 、 $X=2$ ならば $A=2, B=1$ 、「あああっているな」。

余命というのは (現在まで生きながらえた、という条件付きの死ぬまでの年数ですから死ぬときの年齢ではありません。ここでは年齢が k 才の人の余命しか考えていませんから確率変数は一つで良かったのですが、一般に余命という場合は年齢を変数として取り入れて、年齢が k 才の人の余命を X_k と記すことにしましょう。年齢を Y とすると、 $X_k = n$ ということは $Y \geq k$ という条件の下で $Y = n+k$ ということと同じです。したがって、上の解答のような関係式を使う

必要があるのです。 $g(i) = f(i - k)$ とすると、余命の確率を全部足しても 1 にならない（不老長寿！） ちなみに、日本人の平均寿命というのは、0 歳児の平均余命と（正式に）定義されています。

練習問題 4.12（コールオプションの価値） X をパラメータ 1000, 0.5 の 2 項分布に従う確率変数として、 $Y = 14000 + (X - 500)$ と定義します（ Y は 1 週間後の株価を表し、 Y は最高でも 14,500 円、最低でも 13,500 円になる、としたものです）。 $\max\{Y - m, 0\}$ の期待値を m の関数とみなして $g(m)$ と書くことにします。このとき、Excel を使って、 $g(m)$ を $13980 \leq m \leq 14020$ の範囲で計算し、そのグラフを描きなさい。ただし、次の手順に従って計算しなさい。(1) Y の平均 μ と標準偏差 s を計算し、 $\mu - 3s, \mu + 3s$ に近い切りの良い整数をそれぞれ k_0, k_1 としなさい。(2) Excel で、「=binomdist(x,n,p,false)」という数式を入力すると、パラメータ n, p の 2 項分布の確率関数を計算してくれるので、それを利用して $P(Y = i)$ を $k_0 \leq i \leq k_1$ の範囲で計算しなさい。(3) 次の式を使って $g(m)$ を計算しなさい（半分相対参照を使うと、いっぺんに計算が出来るはず）。

$$g(m) = E(\max\{Y - m, 0\}) \cong \sum_{i=k_0}^{k_1} \max\{i - m, 0\} P(Y = i)$$

$g(m)$ はコールオプションの価値と呼ばれます。

練習問題 4.13（セントペテルスブルグのパラドックス）コインを表が出るまで投げて、投げた回数に応じて賞金を上げます。賞金は最初 1 万円、裏が出るたびに倍にしていきます。つまり、最初に表が出たら賞金は 1 万円、2 回目に初めて表が出たら 2 万円、3 回目に初めて表が出たら 4 万円、 k 回目に初めて表が出たら賞金は 2^{k-1} 万円です、賞金額に上限はありません、というオファーがあったとしましょう。(1) この賭をやったときの賞金を X としたとき、 X の確率関数を計算しなさい。(2) X の期待値が無限大になることを計算で確かめなさい。(3) Excel を使って、この賭けを繰り返し行って、平均賞金額を計算し、期待値 ∞ と比較しなさい。(4) 参加費が 10 万円だったとした場合、あなたはこの賭けに参加しますか。

数学者ダニエル・ベルヌイがセントペテルスブルグに滞在中に書いた論文で紹介されていることから、その地名でよばれるようになりました。ちなみに、ベルヌイは有名な数学者物理学者を多数輩出した家系で、このベルヌイのおじさんに、ベルヌイ分布のもとになったヤコブ・ベルヌイがいます。何が「パラドックス」？

練習問題 4.14（天候デリバティブ）ある「海の家」の収益はシーズンの平均気温に左右され、平均気温が x のとき、収益は $500 + 100(x - 25)$ で与えられるものとします。今夏の平均気温の確率分布は次の表のように予想されているとしたとき、冷夏の収益の落ち込みをカバーするために、平均気温が 25 度ならば 200、26 度ならば 100 を受け取る保険契約を結ぶことを考えています。27 度以上の場合には保証はありません。この保険料はいくらくらいが妥当ですか、試算しなさい。

平均気温	25	26	27	28	29	30
確率	0.05	0.1	0.15	0.3	0.25	0.15

練習問題 4.15（在庫管理、テキスト 4.2.1 を理解してから解いてください。テキストの問題を一般化し、需要を連続と考えた場合の最適化問題です） X は累積分布関数 $F(x)$ 、密度関数 $f(x)$

を持つ $[0, \infty)$ で定義された連続確率変数とし、 m をある正の数とします。 m によって異なる確率変数 $C(m)$ を次のように定義します。

$$C(m) \equiv 70 \min\{X, m\} - 50 \max\{m - X, 0\}$$

(1) $C(m)$ の期待値を $g(m)$ としたとき、 $g(m)$ を $f(x)$ を使って表しなさい。(2) $g(m)$ を m で微分することにより、 $g(m)$ の極値を計算しなさい。(3) $g(m)$ を最大とする m の値を求めなさい。

例題 4.12 (視聴率調査) 視聴率調査で、ある時刻にある番組を見ている世帯の数を X としたとき、 X はパラメータ $n = 600, p$ の 2 項分布に従っているとします。(1) その番組の視聴率 X/n の期待値と標準偏差を計算しなさい。(2) $p = 0.1$ としたとき、相対誤差 $|X/n - p|/p$ が 10% 以内の確率を計算しなさい ($n = 600$ は視聴率調査機器の設置されている世帯の数です)。

こたえ (1) X がパラメータ n, p の 2 項分布に従うならば、その期待値は np 、分散は $np(1-p)$ 、したがって

$$E\left(\frac{X}{n}\right) = p, V\left(\frac{X}{n}\right) = \frac{p(1-p)}{n}$$

(2) X が 2 項分布に従うならば、相対誤差が 10% 以内という事象は

$$\frac{|X/600 - 0.1|}{0.1} \leq 0.1 \Leftrightarrow 0.09 \leq \frac{X}{600} \leq 0.11 \Leftrightarrow 54 \leq X \leq 66$$

となるので、2 項分布の 54 以上 66 以下になる確率を計算すればよい。一般的に書けば

$$P\left(\frac{|X/n - p|}{p} \leq 0.1\right) = P\left(0.9p \leq \frac{X}{n} \leq 1.1p\right) = P(0.9np \leq X \leq 1.1np)$$

2 項分布の累積分布関数を Excel で計算する関数は「BINOMDIST(..., TRUE)」なので、「=BINOMDIST(66, 600, 0.1, TRUE) - BINOMDIST(53, 600, 0.1, TRUE)」を計算すればよい。答えは 0.624、つまり、視聴率が 9% から 11% である確率は 0.62。相対誤差 20% (8% 以上 12% 以下) で良ければ 0.1 の代わりに 0.2 を使えば良いので、「66, 53」を「72, 47」として計算すればよい。答えは 0.911。

練習問題 4.16 (内閣支持率調査) 内閣支持率調査で、大きさ n の無作為抽出標本の中に含まれる「支持」回答数を X とすると、 X はパラメータ n, p の 2 項分布に従うという。調査結果を n で割ったもの $Z = X/n$ が支持率の推定値となるが、 Z の期待値と標準偏差を計算しなさい。また、 $n = 1000$ としたとき、 $p = 0.1, 0.2, 0.3$ のそれぞれについて、 $|Z - p| > 0.03$ となる確率を計算しなさい。このことから何が分かりますか。

4.1 確率母関数とモーメント母関数

数列 $\{a_0, a_1, \dots\}$ の母関数は

$$G(z) = \sum_{i=0}^{\infty} a_i z^i \quad (4.9)$$

によって定義されます。 $G(0) = a_0$ です。1 回微分して $z = 0$ を代入すると $G'(0) = a_1$ が得られます。2 回微分して $z = 0$ を代入すると $G''(0) = 2a_2$ が得られます。一般に、これを n 回微分して $z = 0$ を代入すると、

$$G^{(n)}(0) = n!a_n \Rightarrow a_n = \frac{G^{(n)}(0)}{n!}, \quad n = 0, 1, 2, \dots \quad (4.10)$$

のように数列の各項を計算（生成）することができます。これが「母」関数 generating function と呼ばれる所以です。

非負の整数値を取る離散確率変数 X の確率関数 $\{P(X = i), i = 0, 1, 2, \dots\}$ を数列と考えれば、その母関数が定義できますが、それを**確率母関数**といいます。

$$G_X(z) = \sum_{i=0}^{\infty} P(X = i) \times z^i = E(z^X)$$

式の形から、値 $(z^i) \times$ 確率 $(P(X = i))$ の和、の形をしているので、これは確率変数 z^X の期待値にほかなりません。

例題 4.13 パラメータ p のベルヌイ分布の確率母関数を求め、それを利用して、確率関数を計算しなさい。

こたえ 定義に従い

$$G(z) = z^0 \times (1 - p) + z^1 \times p = 1 - p + pz \quad (4.11)$$

が得られる。微分するまでもなく、定数項、 z の係数を見れば

$$P(X = 0) = 1 - p, P(X = 1) = p \quad (4.12)$$

という良く見慣れたベルヌイ分布の確率関数が得られる。

例題 4.14 パラメータ n, p の2項分布の確率母関数を求め、それを微分することにより、確率関数を計算しなさい。

こたえ 確率母関数は、定義に従い

$$G(z) = \sum_{i=0}^{\infty} P(X = i) \times z^i = \sum_{i=0}^n \binom{n}{i} p^i (1 - p)^{n-i} \times z^i \quad (4.13)$$

$$= (pz + 1 - p)^n \quad (4.14)$$

となる。これを微分すると、

$$G'(z) = np((pz + 1 - p)^{n-1} \Rightarrow P(X = 1) = np(1 - p)^{n-1} \quad (4.15)$$

$$G''(z) = n(n-1)p^2(pz + 1 - p)^{n-2} \Rightarrow P(X = 2) = \frac{n(n-1)}{2} p^2 (1 - p)^{n-2} \quad (4.16)$$

$$G'''(z) = \dots \quad (4.17)$$

などが得られる。

確率母関数の威力は複雑な確率計算の際に発揮されます。確率変数 X の確率関数を $\{p_i\}$ 、確率変数 Y の確率関数を $\{q_i\}$ としたとき、すべての i に対して $p_i = q_i$ が成り立つならば、 X と Y は全く同じ性質を持つ確率変数といってよいでしょう。このことから、もし X の確率母関数 $G_X(z)$ と Y の確率母関数 $G_Y(z)$ が一致するならば、 z の級数に展開したときすべての z^i の係数は等しいので、確率変数 X, Y の分布は等しい、という性質が導かれます。このことから、複雑な確率変数 Z の分布を計算するために、その確率母関数を計算して、それが性質の分かっている確率変数 X の確率母関数に一致することを確認するというやり方で、複雑な計算を容易に実行することが可能です。

例題 4.15 X, Y は互いに独立にパラメータ p のベルヌイ分布に従うとき、 $X + Y$ はパラメータ $2, p$ の 2 項分布に従うことを示しなさい。

こたえ パラメータ p のベルヌイ分布の確率母関数は $1 - p + pz$ であった。したがって、 $X + Y$ の確率母関数は

$$E(z^{X+Y}) = E(z^X z^Y) = E(z^X)E(z^Y) \quad (4.18)$$

$$= (1 - p + pz)^2 = (1 - p)^2 + 2p(1 - p)z + p^2 z^2 \quad (4.19)$$

となる。2 番目の等式は X, Y が独立であることから言える。 z^i の各係数 $((1-p)^2, 2p(1-p), p^2)$ を確かめてみると、これらはパラメータ $2, p$ の 2 項分布に確率関数に他ならない。したがって、 $X + Y$ の確率母関数はパラメータ $2, p$ の 2 項分布の確率母関数と一致するため、それはパラメータ $2, p$ の 2 項分布に従うといえる。

例題 4.16 X はパラメータ n, p の 2 項分布に従い、 Y はパラメータ m, p の 2 項分布に従い、両者は互いに独立とする。このとき、 $X + Y$ がパラメータ $n + m, p$ の 2 項分布に従うことを、確率母関数を使って示しなさい。

こたえ パラメータ n, p の 2 項分布の確率母関数は $(pz + 1 - p)^n$ であった。したがって、 $X + Y$ の確率母関数を $G_{X+Y}(z)$ と置くと、 X, Y が独立であることを利用して

$$\begin{aligned} G_{X+Y}(z) &= E(z^{X+Y}) = E(z^X z^Y) = E(z^X)E(z^Y) = G_X(z)G_Y(z) \\ &= (pz + 1 - p)^n (pz + 1 - p)^m = (pz + 1 - p)^{n+m} \end{aligned}$$

と表される。これはパラメータ $n + m, p$ の 2 項分布の確率母関数に他ならない。確率母関数と確率分布は一対一に対応しているので、このことから $X + Y$ が $n + m, p$ の 2 項分布に従っていることが分かる。

コメント 確率母関数を知らない場合、 $X + Y$ の確率関数を計算するには合成積を計算します。全確率の公式を適用すると、

$$\begin{aligned} P(X + Y = k) &= \sum_i P(X + Y = k \mid Y = i)P(Y = i) \\ &= \sum_{i=\max\{0, k-n\}}^{\min\{k, m\}} P(X = k - i)P(Y = i) \\ &= \sum_{i=\max\{0, k-n\}}^{\min\{k, m\}} \binom{n}{k-i} p^{k-i} (1-p)^{n-k+i} \binom{m}{i} p^i (1-p)^{m-i} \\ &= p^k (1-p)^{n+m-k} \sum_{i=\max\{0, k-n\}}^{\min\{k, m\}} \frac{n!m!}{(k-i)!(n-k+i)!i!(m-i)!} \end{aligned}$$

最後の式のシグマを計算すると $\binom{n+m}{k}$ となって、確かに 2 項分布になりますが、その計算は大変。

確率母関数のもう一つの効用は、**期待値の計算が微分で計算できる**ようになることです。実際、確率母関数を微分すると、

$$\frac{d}{dz} G_X(z) = \frac{d}{dz} E(z^X) = E\left(\frac{d}{dz} z^X\right) = E(Xz^{X-1}) \quad (4.20)$$

となるので、 $z = 1$ を代入する $E(X)$ が得られます。もう一回微分して $z = 1$ を代入すると $E(X(X-1))$ が得られます。以下同様にして、

$$\left. \frac{d^k}{dz^k} G_X(z) \right|_{z=1} = E(X(X-1)\cdots(X-k+1)) \quad (4.21)$$

が得られることが分かります。この期待値は k 次の階乗モーメントと呼ばれます。階乗モーメントを組み合わせると、通常モーメントを計算することができるので、各種の期待値を微分で計算できるのです。無限級数の計算はテクニックを要することが多いので、これは大きなメリットです。

例題 4.17 X はパラメータ p の幾何分布に従うとき、その平均と分散を求めなさい。

こたえ パラメータ p の幾何分布の確率母関数は

$$G(z) = \sum_{i=1}^{\infty} p(1-p)^{i-1} z^i = \frac{pz}{1-(1-p)z} \quad (4.22)$$

なので、それを微分することにより、

$$G'(z) = \frac{p(1-(1-p)z) + (1-p)pz}{(1-(1-p)z)^2} = \frac{p}{(1-(1-p)z)^2} \quad (4.23)$$

従って、平均は

$$E(X) = G'(1) = \frac{1}{p} \quad (4.24)$$

もう一回微分すると、

$$G''(z) = \frac{2p(1-p)}{(1-(1-p)z)^3} \quad (4.25)$$

したがって、2 次の階乗モーメントは

$$E(X(X-1)) = G''(1) = \frac{2(1-p)}{p^2} \quad (4.26)$$

したがって、分散は

$$V(X) = E(X(X-1)) + E(X) - (E(X))^2 = \frac{2(1-p)}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{1-p}{p^2} \quad (4.27)$$

コメント 確率母関数を使わないで計算する場合は

$$\sum_{i=1}^{\infty} ix^{i-1} = \sum_{i=1}^{\infty} \sum_{j=1}^i x^{i-1} = \sum_{j=1}^{\infty} \sum_{i=j}^{\infty} x^{i-1} = \sum_{j=1}^{\infty} \frac{x^{j-1}}{1-x} = \frac{1}{(1-x)^2} \quad (4.28)$$

$$\sum_{i=1}^{\infty} i(i+1)x^{i-1} = 2 \sum_{i=1}^{\infty} \sum_{j=1}^i jx^{i-1} = 2 \sum_{j=1}^{\infty} j \sum_{i=j}^{\infty} x^{i-1} = 2 \sum_{j=1}^{\infty} j \frac{x^{j-1}}{1-x} = \frac{2}{(1-x)^3} \quad (4.29)$$

という関係を使って

$$E(X) = \sum_{i=1}^{\infty} ip(1-p)^{i-1} = \frac{1}{p} \quad (4.30)$$

$$E(X^2) = \sum_{i=1}^{\infty} i(i+1)p(1-p)^{i-1} - \sum_{i=1}^{\infty} ip(1-p)^{i-1} = \frac{2}{p^2} - \frac{1}{p} \quad (4.31)$$

$$V(X) = \frac{2}{p^2} - \frac{1}{p} - \frac{1}{p^2} = \frac{1-p}{p^2} \quad (4.32)$$

のように計算できますが、ちょっとした工夫が必要。

確率母関数はこのように優れたものの道具ですが、残念ながら非負整数値を取る離散型確率変数に対してしか適用できません。これを一般の確率変数に適用できるようにしたのが、モーメント母関数です。一般の確率変数 X のモーメント母関数は、確率変数 $e^{\theta X}$ の期待値として定義され、 $M_X(\theta)$ と書くことにします。

$$M_X(\theta) = E(e^{\theta X}) = \begin{cases} \sum_i e^{\theta a_i} P(X = a_i), & X \text{ が離散の場合} \\ \int_{-\infty}^{\infty} e^{\theta x} f_X(x) dx, & X \text{ が連続の場合} \end{cases} \quad (4.33)$$

確率変数が非負整数値を取る離散型確率変数の場合は、 $e^{\theta} = z$ と置くと、モーメント母関数は確率母関数に一致することに注意してください。

モーメント母関数の場合は確率母関数の場合よりもっと簡単に、微分とモーメントが結びついています。実際、次の式が成り立ちます。

$$\left. \frac{d^k}{d\theta^k} M_X(\theta) \right|_{\theta=0} = E(X^k) \quad (4.34)$$

すべての次数のモーメントが一致すれば、確率変数として一致する、という性質があるため、同じモーメント母関数を持つ二つの確率変数は同じ分布に従います。そこで、分布が知られていないある確率変数 X のモーメント母関数を何らかの方法で計算できたとして、それが既知の分布 $F(x)$ のモーメント母関数に一致するということを確かめることが出来たとしたら、 X は $F(x)$ に従う確率変数であることが確信を持って言えることになります。その例は下の例題に出きます。

例題 4.18 標準正規分布に従う確率変数 X のモーメント母関数を計算し、それを微分することにより、その平均が 0、分散が 1 であることを確かめなさい。

こたえ 定義通りに計算すると、

$$E(e^{\theta X}) = \int_{-\infty}^{\infty} e^{\theta x} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = e^{\theta^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-(x-\theta)^2/2} dx = e^{\theta^2/2} \quad (4.35)$$

最後の等式は、定積分の被積分関数が平均 θ 、分散 1 の正規分布の密度関数になっていることから導かれる。これを微分すると、

$$M'_X(\theta) = \theta e^{\theta^2/2} \quad (4.36)$$

$$M''_X(\theta) = e^{\theta^2/2} + \theta^2 e^{\theta^2/2} \quad (4.37)$$

したがって、

$$E(X) = M'_X(0) = 0 \quad (4.38)$$

$$V(X) = M''_X(0) - (M'_X(0))^2 = 1 \quad (4.39)$$

例題 4.19 パラメータ μ, σ^2 の正規分布に従う確率変数 X のモーメント母関数を計算し、それを微分することにより、その平均が μ 、分散が σ^2 であることを確かめなさい。

こたえ X を標準化する（平均を引いて標準偏差で割る）と標準正規分布に従う。標準正規分布のモーメント母関数は上の例題で求められているので、それを使って変形すればよい。

標準正規分布にしたがう確率変数を Z とすると、 $X = \sigma Z + \mu$ 。したがって、 X のモーメント母関数は

$$M_X(\theta) = E(e^{\theta(\sigma Z + \mu)}) = E(e^{\theta\sigma Z})E(e^{\theta\mu}) = e^{\theta\mu}M_Z(\theta\sigma) = \exp\left(\theta\mu + \frac{\sigma^2\theta^2}{2}\right) \quad (4.40)$$

これを微分すると、

$$M'_X(\theta) = (\mu + \sigma^2\theta) \exp\left(\theta\mu + \frac{\sigma^2\theta^2}{2}\right) \quad (4.41)$$

$$M''_X(\theta) = (\sigma^2 + (\mu + \sigma^2\theta)^2) \exp\left(\theta\mu + \frac{\sigma^2\theta^2}{2}\right) \quad (4.42)$$

したがって、

$$E(X) = M'_X(0) = \mu \quad (4.43)$$

$$V(X) = M''_X(0) - M'_X(0)^2 = \sigma^2 \quad (4.44)$$

例題 4.20 X はパラメータ μ, σ^2 の正規分布に従い、 Y はパラメータ ν, τ^2 の正規分布に従い、両者は互いに独立とする。このとき、 $X + Y$ がパラメータ $\mu + \nu, \sigma^2 + \tau^2$ の正規分布に従うことを、モーメント母関数を使って示しなさい。

こたえ X, Y のモーメント母関数は

$$M_X(\theta) = \exp\left(\theta\mu + \frac{\sigma^2\theta^2}{2}\right) \quad (4.45)$$

$$M_Y(\theta) = \exp\left(\theta\nu + \frac{\tau^2\theta^2}{2}\right) \quad (4.46)$$

であるが、それらが互いに独立なので、 $e^{\theta X}, e^{\theta Y}$ も互いに独立。したがって、 $X + Y$ のモーメント母関数は

$$M_{X+Y}(\theta) = M_X(\theta)M_Y(\theta) = \exp\left(\theta\mu + \frac{\sigma^2\theta^2}{2}\right) \exp\left(\theta\nu + \frac{\tau^2\theta^2}{2}\right) \quad (4.47)$$

$$= \exp\left(\theta(\mu + \nu) + \frac{(\sigma^2 + \tau^2)\theta^2}{2}\right) \quad (4.48)$$

これはパラメータ $\mu + \nu, \sigma^2 + \tau^2$ の正規分布のモーメント母関数にほかならない。ということは $X + Y$ がパラメータ $\mu + \nu, \sigma^2 + \tau^2$ の正規分布にしたがうということになる。

モーメント母関数を知らない場合は、合成積を使って定積分を計算しなければいけない。ちょっと大変。

練習問題 4.17 平均 0、分散 σ^2 の正規分布にしたがう確率変数 X について、 X^2 の平均は σ^2 、分散は $2\sigma^4$ になることを示せ（モーメント母関数を使う）