July, 2019

情報領域演習第二 第１４回 Ｋ演習 (確率論)

演習問題１：二項分布

互いに独立に n 回のベルヌーイ試行を繰り返したときの成功回数 X が従う確率分布を二項分布と呼ぶ。その確率関数を

$$p(x; n, p) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \tag{1}$$

と表すとして、以下の問いに答えよ。

1. この確率分布の標本空間と母数空間を答えよ。
2. この確率分布の平均 (X の期待値) を求めよ。
3. この確率分布の分散 ($(X - \mathrm{E}[X])^2$ の期待値) を求めよ。
4. この確率分布のモーメント母関数 ($\exp(tX)$ の期待値) を求めよ。
5. 前問のモーメント母関数から X の期待値と X^2 の期待値を求めよ。
6. X_1 と X_2 が互いに独立に p が共通の二項分布 $Bin(n_1, p)$ と $Bin(n_2, p)$ に従うとき、X_1 と X_2 の和 $Y = X_1 + X_2$ の確率関数を求めよ。

ヒント：平均と分散とモーメント母関数の計算には、ベルヌーイ試行の独立な繰り返しであることを用いる手順と、それぞれの期待値の計算に正面から取り組む手順とがある。演習の時間では後者を解説する。和の分布は、原則として畳み込みの計算 $f_c(y) = \sum_{y-x_1 \in \mathcal{X}} p_{X_1}(x_1) p_{X_2}(y - x_1)$ で求める。ただしモーメント母関数から和の分布が分かる場合もある。

解説：二項分布の期待値の計算には、ベルヌーイ分布に従う n 個の互いに独立な確率変数である事実を用いることができる。ベルヌーイ試行の期待値を求めてから、その事実に基づいて

$$E[X] = np$$
$$V[X] = np(1-p)$$
$$M_X(t) = \left(e^t + 1 - p\right)^n$$

と答えるのが、もっとも誤りにくい。しかし正面から取り組む計算も可能であり、今回はそちらを解説する。

まず二項分布の名前は、次のような<u>二項定理</u>に由来する。二項分布のモーメントの計算には必ず二項定理が現れる。

$$(a+b)^n = {}_nC_0 a^0 b^n + {}_nC_1 a^1 b^{n-1} + {}_nC_2 a^2 b^{n-2} + \cdots + {}_nC_{n-1} a^{n-1} b^1 + {}_nC_n a^n b^0 \tag{2}$$

$a = p$、$b = 1 - p$ と置けば直ちに

$$\begin{aligned}(p + (1-p))^n &= {}_nC_0 p^0 (1-p)^{n-0} + {}_nC_1 p^1 (1-p)^{n-1} + {}_nC_2 p^2 (1-p)^{n-2} + \cdots \\ &\quad + {}_nC_{n-1} p^{n-1} (1-p)^1 + {}_nC_n p^n (1-p)^0 & (3) \\ &= 1 & (4)\end{aligned}$$

の関係が得られる。1 行目は $X = k$ の確率が ${}_nC_k p^k (1-p)^{n-k}$ であることを意味し、2 行目は全確率が 1 であることを意味する。

X の期待値 (二項分布の平均、二項分布の期待値) は、次の 3 通りのいずれも、それほど複雑ではない。

1. 互いに独立な同一のベルヌーイ試行の和である事実を用いる計算
2. 期待値の定義通りの計算
3. モーメント母関数を一階偏微分して得る計算

一つめの計算方法は、各ベルヌーイ試行を $X_1, \ldots, X_n$ で表し、その和を Y とする。確率変数の和の期待値はそれぞれの期待値の和であることを用いて

$$E[Y] = E[X_1 + \cdots + X_n] = E[X_1] + \cdots + E[X_n] = p + \cdots + p = np \tag{5}$$

と Y の期待値を得る。

二つめは定義通りである。

$$\begin{aligned}
E[X] &= \sum_{k=0}^{n} k {}_nC_k p^k (1-p)^{n-k} \\
&= \sum_{k=0}^{n} k \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} k \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} \frac{n!}{(k-1)!\,(n-k)!} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} n \frac{(n-1)!}{(k-1)!\,((n-1)-(k-1))!} p \cdot p^{k-1} (1-p)^{(n-1)-(k-1)} \\
&= np \sum_{l=0}^{n-1} \frac{(n-1)!}{l!\,((n-1)-l)!} p^l (1-p)^{(n-1)-l} \\
&= np \{p + (1-p)\}^{n-1} \\
&= np
\end{aligned} \tag{6}$$

と項をくくりだして、二項展開に帰着させる。この計算は、二項分布の確率関数に直に触れる。

三つめはモーメント母関数さえ導ければ

$$\left.\frac{\partial (pe^t + 1 - p)^n}{\partial t}\right|_{t=0} = n\left(pe^t + 1 - p\right)^{n-1} pe^t\Big|_{t=0} = n(p + 1 - p)^{n-1} p = np \tag{7}$$

となる。

二項分布の分散は、X^2 の期待値 (= X^2 の平均)

$$E\left[X^2\right] = \sum_{k=0}^{n} k^2 {}_nC_k p^k (1-p)^{n-k} = \sum_{k=0}^{n} k^2 \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k} \tag{8}$$

あるいは X の分散 (= X^2 の平均 - X の平均の二乗)

$$V[X] = E\left[(X - E[X])^2\right] = E\left[X^2\right] - (E[X])^2 \tag{9}$$

のいずれかで計算を進めると遠回りで、少し巧妙に

$$V[X] = E[X(X-1)] + E[X] - (E[X])^2 \tag{10}$$

を用いて、計算を簡単にする。確率関数の分母に階乗がある場合は、X^2 の期待値より $X(X-1)$ の期待値の方が簡単に計算できる。$k^2 p(k)$ の総和より、$k(k-1)p(k)$ の総和の方が、間違えにくく、二項定理に帰着させやすい。

$E[X(X-1)]$ を計算すると

$$
\begin{aligned}
E[X(X-1)] &= \sum_{k=0}^{n} k(k-1)\,{}_nC_k p^k (1-p)^{n-k} \\
&= \sum_{k=0}^{n} k(k-1)\frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k} \\
&= \sum_{k=2}^{n} k(k-1)\frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k} \\
&= \sum_{k=2}^{n} n(n-1)\frac{(n-2)!}{(k-2)!\,((n-2)-(k-2))!} p^2 \cdot p^{k-2} (1-p)^{(n-2)-(k-2)} \\
&= n(n-1)p^2 \sum_{l=0}^{n-2} \frac{(n-2)!}{l!\,((n-2)-l)!} p^l (1-p)^{(n-2)-l} \\
&= n(n-1)p^2 \{p+(1-p)\}^{n-2} \\
&= n(n-1)p^2
\end{aligned}
\tag{11}
$$

となる。これと (10) 式より、

$$
V[X] = n(n-1)p^2 + np - (np)^2 = np(1-p) \tag{12}
$$

と求まる。

二項分布はモーメント母関数も比較的簡単に導ける。二項分布の確率変数が、互いに独立なベルヌーイ試行の確率変数の和で表されることを利用するなら、成功確率 p のベルヌーイ試行のモーメント母関数が

$$
\begin{aligned}
M_X(t) &= E\left[e^{tX}\right] = e^{t\cdot 0}p(0) + e^{t\cdot 1}p(1) \\
&= 1\cdot(1-p) + e^t \cdot p = pe^t + 1 - p
\end{aligned}
\tag{13}
$$

であったことを思い出す。$X_1,\ X_2,\ \ldots$ を互いに独立なベルヌーイ試行変数の列とすると、それぞれのベルヌーイ試行の確率分布のモーメント母関数はすべて共通の

$$
M_{X_i}(t) = 1\cdot p + e^t \cdot p,\ \ i = 1, 2, \ldots \tag{14}
$$

となる。今 $Y = X_1 + X_2 + \cdots + X_n$ とすると、Y は n 回の独立なベルヌーイ試行での成功回数であり、Y の確率分布は二項分布となるはずである。独立な確率変数の和の分布のモーメント母関数が、それぞれの確率変数のモーメント母関数の積で得られる、との関係を思い出せば、Y の確率分布のモーメント母関数は

$$
\begin{aligned}
M_Y(t) &= \prod_{i=1}^{n} M_{X_i}(t) \\
&= \left(pe^t + 1 - p\right)^n
\end{aligned}
\tag{15}
$$

と得られる。

また、直接に $\mathrm{E}\left[e^{tX}\right]$ を計算してもよい。$n=1$ の場合、

$$M_X\left(t\right) = pe^t + 1 - p \tag{16}$$

$n=2$ の場合、

$$\begin{aligned} M_X\left(t\right) &= p^2e^{2t} + 2pe^t\left(1-p\right) + \left(1-p\right)^2 \\ &= \left(pe^t + 1 - p\right)^2 \end{aligned}$$

$n=3$ の場合も、

$$\begin{aligned} M_X\left(t\right) &= p^3e^{3t} + 3p^2e^{2t}\left(1-p\right) + 3pe^t\left(1-p\right)^2 + \left(1-p\right)^2 \\ &= \left(pe^t + 1 - p\right)^3 \end{aligned}$$

となる。$n=3$ までなら、二項定理を思い出さずとも計算できるが、一般の n については e^{tX} が $(e^t)^X$ であることに気づけば、

$$\begin{aligned} M_X\left(t\right) &= \sum_{k=0}^{n} e^{tk}\frac{n!}{k!\left(n-k\right)!}p^k\left(1-p\right)^{n-k} \\ &= \sum_{k=0}^{n} \left(e^t\right)^k\frac{n!}{k!\left(n-k\right)!}p^k\left(1-p\right)^{n-k} \\ &= \sum_{k=0}^{n} \frac{n!}{k!\left(n-k\right)!}\left(pe^t\right)^k\left(1-p\right)^{n-k} \\ &= \left(pe^t + 1 - p\right)^n \end{aligned} \tag{17}$$

とモーメント母関数の計算にも、二項定理を用いる。

演習問題２：指数分布

ある繰り返し起こる事象の単位時間あたりの発生回数がポアソン分布に従うとき、個々の事象の発生間隔 X は指数分布に従うことが分かっている。その確率密度関数を

$$f(x;\lambda) = \lambda e^{-\lambda x} \tag{18}$$

と表すとして、以下の問いに答えよ。

1. この確率分布の標本空間と母数空間を答えよ。
2. この確率分布の平均 (X の期待値) を求めよ。
3. この確率分布の分散 ($(X - \mathrm{E}[X])^2$ の期待値) を求めよ。
4. この確率分布のモーメント母関数 ($\exp(tX)$ の期待値) を求めよ。
5. 前問のモーメント母関数から X の期待値と X^2 の期待値を求めよ。
6. X_1 と X_2 が互いに独立に同一の指数分布 $Exp(\lambda)$ に従うとき、X_1 と X_2 の和 $Y = X_1 + X_2$ の確率密度関数を求めよ。

ヒント：部分積分を用いる。また和の分布は、原則として畳み込みの計算 $f_Y(y) = \int_{x_1=0}^{y} f(x_1) f(y - x_1) dx_1$ で求める。ただしモーメント母関数から和の分布が分かる場合もある。

解説：指数分布は、ランダムな到着間隔の確率分布、あるいは待ち時間の分布、として用いられることが多い。オペレーションズリサーチで待ち行列というトピックを学ぶ際にも、M/M/1 という基本的なモデルの到着間隔と処理時間の確率分布として紹介される。

累積分布関数が

$$F(k) = Pr[X \leq k] = 1 - \exp(-\lambda x) \tag{19}$$

であり、確率密度関数はこれを微分して

$$f(x) = \frac{d}{dx}(1 - \exp(-\lambda x)) = \lambda \exp(-\lambda x) \tag{20}$$

となる。

指数分布の期待値の計算は、指数関数の積分

$$\int e^a x dx = e^{ax}/a + C \tag{21}$$

と、部分積分

$$\int_{x\in\mathcal{X}} g(x) e^x dx = [e^x g(x)]_{x\in\mathcal{X}} - \int_{x\in\mathcal{X}} g'(x) e^x dx \tag{22}$$

の組み合わせ、あるいは繰り返しで乗り切れる。例えば X の期待値の計算は、

$$
\begin{aligned}
E\left[X\right] &= \int_0^\infty x\lambda e^{-\lambda x} dx \\
&= \left[x\left(-e^{-\lambda x}\right)\right]_{x=0}^\infty - \int_0^\infty \left(-e^{-\lambda x}\right) dx \\
&= 0 - 0 + \int_0^\infty e^{-\lambda x} dx \\
&= \left[-\frac{1}{\lambda} e^{-\lambda x}\right]_{x=0}^\infty \\
&= 0 - \left(-\frac{1}{\lambda}\right) = \frac{1}{\lambda}
\end{aligned} \tag{23}
$$

となる。ここで $\frac{-1}{\lambda} e^{-\lambda x}$ の $x \to \infty$ の極限は、$x \to \infty$ よりも速く $e^{-\lambda x} \to 0$ となることから 0 になるのは、既知とする。

2 乗の期待値は、途中で X の期待値の $2/\lambda$ 倍になったことに気づけば、そこで計算はおしまい。

$$
\begin{aligned}
E\left[X^2\right] &= \int_0^\infty x^2 \lambda e^{-\lambda x} dx \\
&= \left[x^2\left(-e^{-\lambda} x\right)\right]_{x=0}^\infty - \int_0^\infty \left(-2xe^{-\lambda x}\right) dx \\
&= \int_0^\infty 2xe^{-\lambda x} dx \\
&= \frac{2}{\lambda} \int_0^\infty \lambda x e^{-\lambda x} dx \\
&= \frac{2}{\lambda} E\left[X\right] = \frac{2}{\lambda^2}
\end{aligned} \tag{24}
$$

気づかなければ、3 行目以降、部分積分をもう一度繰り返す。

$$
\begin{aligned}
E\left[X^2\right] &= \int_0^\infty x^2 \lambda e^{-\lambda x} dx \\
&= \left[x^2\left(-e^{-\lambda} x\right)\right]_{x=0}^\infty - \int_0^\infty \left(-2xe^{-\lambda x}\right) dx \\
&= 0 - 0 + \int_0^\infty 2xe^{-\lambda x} dx \\
&= \left[2x\frac{-1}{\lambda} e^{-\lambda x}\right]_{x=0}^\infty - \int_0^\infty 2\frac{-1}{\lambda} e^{-\lambda x} dx \\
&= 0 - 0 + \int_0^\infty 2\frac{1}{\lambda} e^{-\lambda x} dx \\
&= \left[\frac{2}{\lambda}\frac{-1}{\lambda} e^{-\lambda x}\right]_{x=0}^\infty \\
&= 0 - \frac{-2}{\lambda^2} = \frac{2}{\lambda^2}
\end{aligned} \tag{25}
$$

以上から分散は次のように求まる。

$$
V\left[X\right] = E\left[X^2\right] - \left(E\left[X\right]\right)^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2} \tag{26}
$$

指数分布のモーメント母関数は

$$
\begin{aligned}
M_X(t) = E\left[e^{tx}\right] &= \int_0^\infty e^{tx}\lambda e^{-\lambda x}dx \\
&= \lambda\int_0^\infty e^{(t-\lambda)x}dx \\
&= \lambda\left[\frac{1}{t-\lambda}e^{(t-\lambda)x}\right]_0^\infty \\
&= \lambda\left[\frac{1}{t-\lambda}\lim_{x\to\infty}e^{(t-\lambda)x} - \frac{1}{t-\lambda}e^{(t-\lambda)0}\right] \\
&= \lambda\left[\frac{1}{t-\lambda}\lim_{x\to\infty}e^{(t-\lambda)x} - \frac{1}{t-\lambda}\right]
\end{aligned} \tag{27}
$$

となる。ここで $t-\lambda=0$ のとき分母が 0 、$t-\lambda>0$ ならば極限が ∞ に発散となるので、$t\geq\lambda$ の範囲ではモーメント母関数は存在しない。$t-\lambda<0$ のとき、$\exp(t-\lambda)x$ で $x\to\infty$ とすると 0 に収束するので、

$$
M_X(t) = \frac{\lambda}{t-\lambda}(0-1) = \frac{\lambda}{\lambda-t} \tag{28}
$$

を得る。このモーメント母関数の表現から、和の分布のモーメント母関数は

$$
\left(\frac{\lambda}{\lambda-t}\right)^n \tag{29}
$$

となり、これは $n=1$ 以外では指数分布にはならないことが証明される。

そこで、和の分布を畳み込み計算で求める。$Y=y$ のときの $Y=X_1+X_2$ の確率密度関数は、$0\leq x_1\leq y$ かつ $0\leq x_2\leq y$ かつ $x_1+x_2=y$ を満たす範囲で同時密度関数 $f_{X_1}(x_1)f_{X_2}(x_2)$ を積分して得る。

$$
f_Y(y) = \int\int_{(x_1,x_2):0\leq x_1\leq y,\,0\leq x_2\leq y,\,x_1+x_2=y} f_{X_1}(x_1)f_{X_2}(x_2)\,dx_1dx_2 \tag{30}
$$

この定積分の範囲は、平面上の直線 $x_2=y-x_1$ の第一象限内の線分に他ならない。

$$
\{(x_1,x_2):0\leq x_1\leq y,\,0\leq x_2\leq y,\,x_1+x_2=y\} = \{(x,y-x):0\leq x\leq y\} \tag{31}
$$

である。

よって改めて、後者で定積分を記すと

$$
\begin{aligned}
f_Y(y) &= \int_0^y f_{X_1}(x)f_{X_2}(y-x)\,dx \\
&= \int_0^y \lambda e^{-\lambda x}\lambda e^{-\lambda(y-x)}dx \\
&= \int_0^y \lambda^2 e^{-\lambda y}dx \\
&= \lambda^2 e^{-\lambda y}\int_0^y dx \\
&= \lambda^2 y e^{-\lambda y}
\end{aligned}
$$

となる。確かに指数分布とは異なる確率密度関数を得る。また、この計算が畳み込みである。

補足：指数分布の確率密度関数には

$$
f(x;\mu) = \frac{1}{\mu}e^{-x/\mu} \tag{32}
$$

という表現もある。$\mu = 1/\lambda$ である。

参考：確率変数 $X_1,\ X_2,\ \ldots,\ X_n$ が互いに独立に同一の指数分布 $Exp\left(\lambda\right)$ に従うとき、これらの和 $Y = X_1 + X_2 + \cdots + X_n$ は形状母数が n、尺度母数が $1/\lambda$ のガンマ分布に従う。密度関数は

$$\begin{aligned} f\left(y\right) &= \frac{\lambda^n y^{n-1} e^{-\lambda y}}{(n-1)!} \\ &= \frac{\lambda^n y^{n-1} e^{-\lambda y}}{\Gamma\left(n\right)} \end{aligned}$$

となる。これはアーラン分布とも呼ばれる。