Hand out

確率論 (*Probability*) 第 11 週

due	19 juin 2014
cur	26 juin 2014
ver	1
rev	1

w.yamamoto

1　ベルヌーイ試行とその仲間たち

再現性は、ある法則が科学的に正しいと認定されるための条件の一つである。例えば「条件 A の下で事象 B が発生する」ことを法則として訴えるとしよう。条件を揃えても事象 B の発生が再現できなければ、「その条件の下で事象が必ず起こる」との主張を受け入れてもらえない。事象自体は、実験条件を A に整えれば B という成果が得られるという決定論的な事象でも、あるいは実験条件 A の下で B が得られる確率は 60%などの確率的な事象でも、どちらでもよい。が、ともかく再現しなければならない。

ベルヌーイ試行はそのような科学的な探求の最も単純化されたモデルである。まず 1 回の試行の成功と失敗を、ベルヌーイ試行でモデル化する。そして試行を複数回繰り返す時の、成功回数の確率モデルを順に導出していく。

1.1　ベルヌーイ試行

1 回の試行で、成功する確率を p 、失敗する確率を $1-p$ とする。この 1 回の試行を、ベルヌーイ試行と呼ぶ。

ベルヌーイ試行を確率分布でモデル化すると、その標本空間は $\{0,1\}$ 、加法族は $\{\{\phi\},\{0\},\{1\},\{0,1\}\}$ 、そして確率関数は

$$p(0) = Pr[X=0] = 1-p,\ p(1) = Pr[X=1] = p \tag{1}$$

となる。これはとても単純な分布であり、平均や分散などの計算は、やる気さえ出せば、極めて容易である。

$$\begin{aligned}
\mathrm{E}[X] &= \sum_{x=-\infty}^{x=\infty} xp(x) = \sum_{x=0}^{x=1} xp(x) = 0\cdot(1-p) + 1\cdot p = p \\
\mathrm{E}\left[X^2\right] &= \sum_{x=-\infty}^{x=\infty} x^2 p(x) = \sum_{x=0}^{x=1} x^2 p(x) = 0\cdot(1-p) + 1\cdot p = p \\
\mathrm{Var}[X] &= \sum_{x=-\infty}^{x=\infty} (x - E[X])^2 p(x) = \sum_{x=0}^{x=1} (x-p)^2 p(x) = p^2\cdot(1-p) + (1-p)^2\cdot p = p(1-p)
\end{aligned}$$

モーメント母関数も

$$\mathrm{E}\left[e^{tX}\right] = \sum_{x=-\infty}^{x=\infty} e^{tx} p(x) = \sum_{x=0}^{x=1} e^{tx} p(x) = e^{t\cdot 0}\cdot(1-p) + e^{t\cdot 1}\cdot p = 1 - p + pe^t$$

と簡単に得られる。

この分布の平均と分散の間には、

$$Var[X] = E[X](1 - E[X]) \tag{2}$$

の関係があり、$p = 0.5$ のときに分散 (ばらつき、不確実性) が最大となる。また平均情報量 (エントロピー、シャノン情報量、情報エントロピー) も

$$H[X] = -(1-p)\log(1-p) - p\log p \tag{3}$$

であり、これも同様に $p = 0.5$ のときに最大となる。双方とも最小になるのは、$p = 0$ もしくは $p = 1$ と結果が明らかな状況であり、最大になるのは $p = 0.5$ と結果が最も不確かな状況である。

以下では、このベルヌーイ試行を独立に繰り返すことを考える。繰り返すので、p は共通である。

1.2　二項分布

同じベルヌーイ試行を独立[1]に繰り返す。例えば総試行回数を 2 回と決めて繰り返すと、全試行回数中の成功する回数は $\{0, 1, 2\}$ の何れかとなる。これが標本空間である。それぞれの値を取る確率は

$$\begin{aligned} p(0) &= Pr[X_1 = 0 \ and \ X_2 = 0] \\ &= Pr[X_1 = 0] Pr[X_2 = 0] = (1-p)^2 \quad (4) \\ p(1) &= Pr[\{X_1 = 0 \ and \ X_2 = 1\} \ or \ \{X_1 = 1 \ and \ X_2 = 0\}] \\ &= Pr[X_1 = 1] Pr[X_2 = 0] + Pr[X_1 = 0] Pr[X_2 = 1] \\ &= p(1-p) + (1-p)p = 2p(1-p) \\ &= {}_2C_1 p(1-p) \quad (5) \\ p(2) &= Pr[X_1 = 1 \ and \ X_2 = 1] \\ &= Pr[X_1 = 1] Pr[X_2 = 1] = p^2 \quad (6) \end{aligned}$$

と求まる。

同様の計算から、n 回の互いに独立な試行の繰り返しの中で成功が k 回となる確率は、

$$p(k) = Pr[X = k] = {}_nC_k p^k (1-p)^{n-k} = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \quad (7)$$

と得られる。この確率関数を持つ確率分布を 二項分布 と呼ばれる。

二項分布の確率関数

$$p(k) = p(k; n, p) = Pr[X = k] = {}_nC_k p^k (1-p)^{n-k}$$

二項分布の標本空間は、すべて失敗する $X = 0$ からすべて成功する $X = n$ までの整数であり、 $\{0, 1, \ldots, n\}$ となる。

二項分布の名前は、次のような 二項定理 に由来する。

$$(a+b)^n = {}_nC_0 a^0 b^n + {}_nC_1 a^1 b^{n-1} + {}_nC_2 a^2 b^{n-2} + \cdots + {}_nC_{n-1} a^{n-1} b^1 + {}_nC_n a^n b^0 \quad (8)$$

$a = p$、$b = 1 - p$ と置けば直ちに

$$\begin{aligned} (p + (1-p))^n &= {}_nC_0 p^0 (1-p)^{n-0} + {}_nC_1 p^1 (1-p)^{n-1} + {}_nC_2 p^2 (1-p)^{n-2} + \cdots \\ &\quad + {}_nC_{n-1} p^{n-1} (1-p)^1 + {}_nC_n p^n (1-p)^0 \quad (9) \\ &= 1 \quad (10) \end{aligned}$$

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。

二項分布の累積分布関数は、地道に $p(k)$ の累積和を計算するしかない。

二項分布の累積分布関数

$$F(k) = Pr[X \le k] = \sum_{l=0}^{k} {}_nC_l p^l (1-p)^{n-l}$$

図 4~6 に確率関数を図示してみる。

[1]独立ならば同時に起こる確率はそれぞれが単独で起こる確率の積となる。$Pr[X_1 \le a \ and \ X_2 \le b] = Pr[X_1 \le a] Pr[X_2 \le b]$ も、$Pr[X_1 = a \ and \ X_2 = b] = Pr[X_1 \le a] Pr[X_2 \le b]$ も成り立つ。

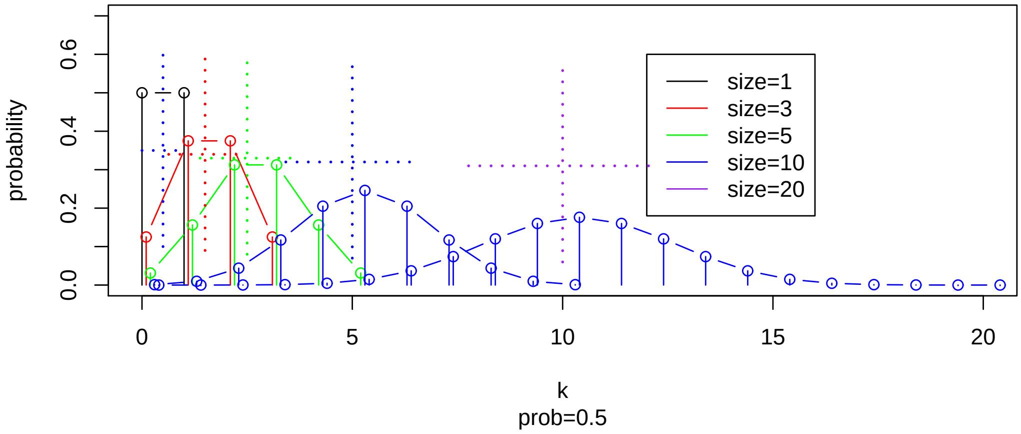

図 4 $p = 0.5$ の二項分布の確率関数 $(n = 1, 3, 5, 10, 20)$

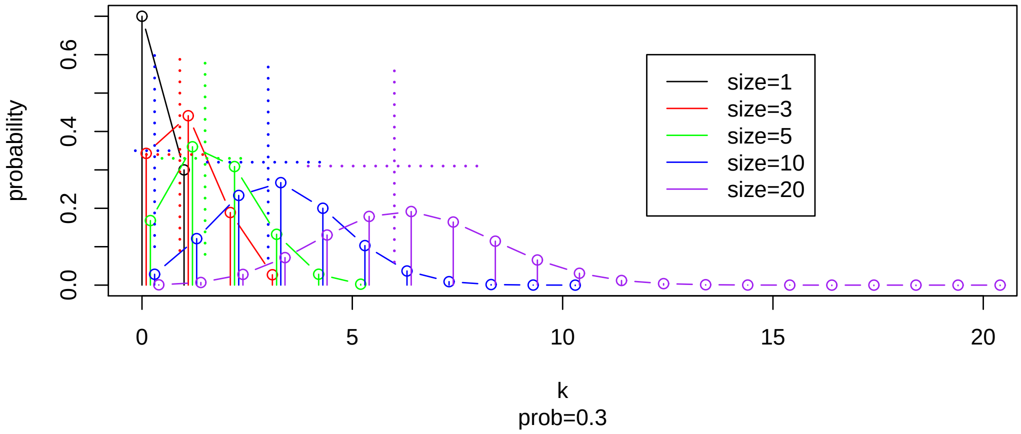

図 5 $p = 0.3$ の二項分布の確率関数 $(n = 1, 3, 5, 10, 20)$

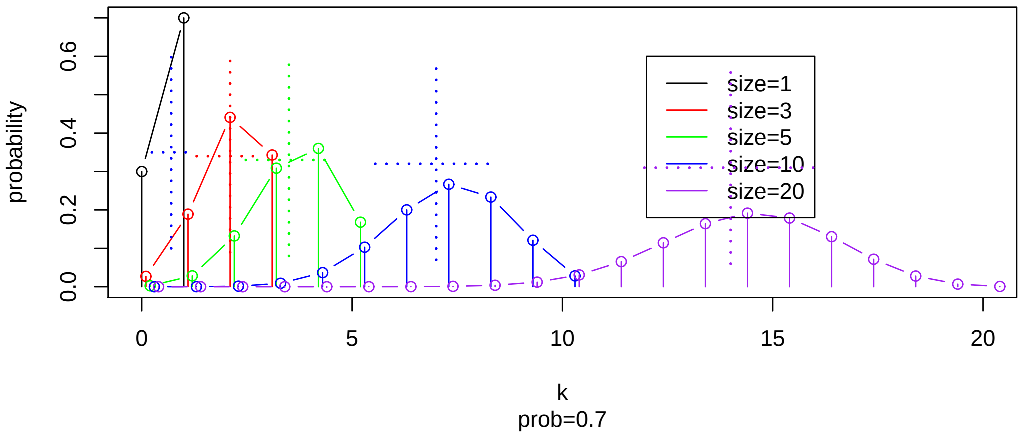

図 6 $p = 0.7$ の二項分布の確率関数 $(n = 1, 3, 5, 10, 20)$

確率関数は $p = 0.5$ の時に対象、また $p < 0.5$ では右に偏り、$p > 0.5$ では左に偏る。n 回の試行を繰り返すと np に近い整数の方が、np より遠い整数よりも、観測されやすい。

次に、確率関数と累積分布関数の関係を図示してみる。総試行回数が 2 の場合は、

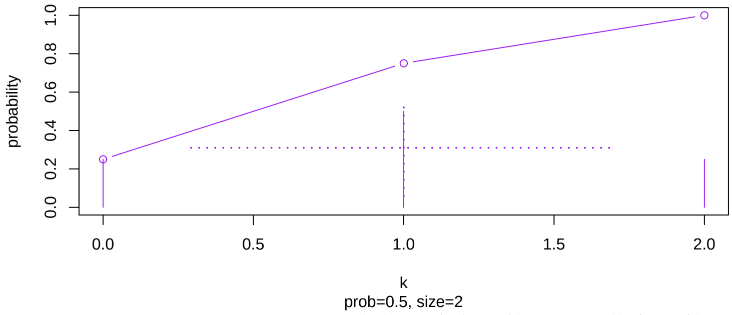

図 7 $n = 2, p = 0.5$ の二項分布の確率関数と累積分布関数

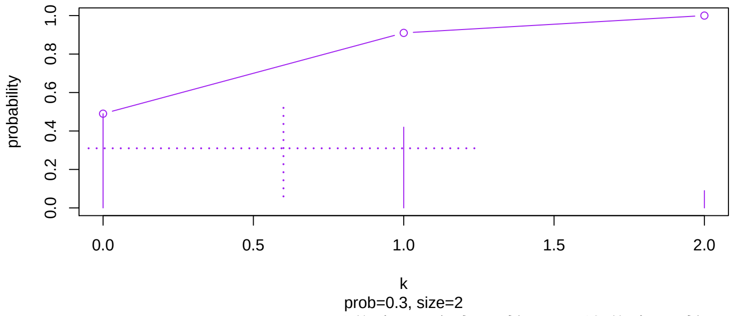

図 8 $n = 2, p = 0.3$ の二項分布の確率関数と累積分布関数

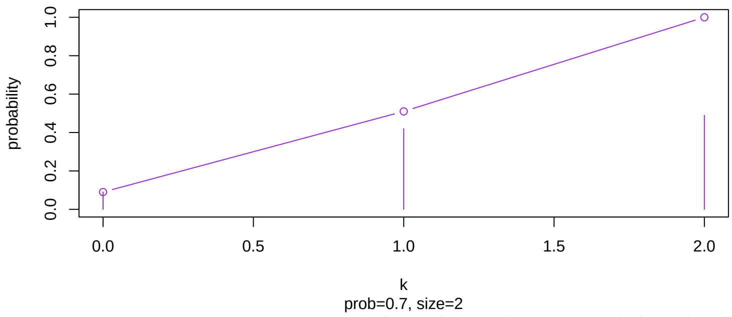

図 9 $n = 2, p = 0.7$ の二項分布の確率関数と累積分布関数

のように、自分で計算して描くことも簡単だろう。総回数が 20 の場合は、

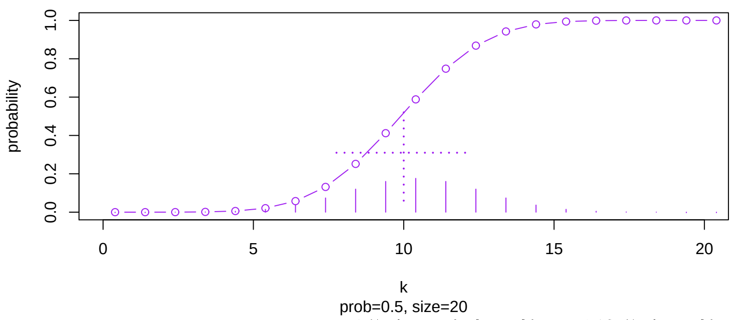

図 10 $n = 20, p = 0.5$ の二項分布の確率関数と累積分布関数

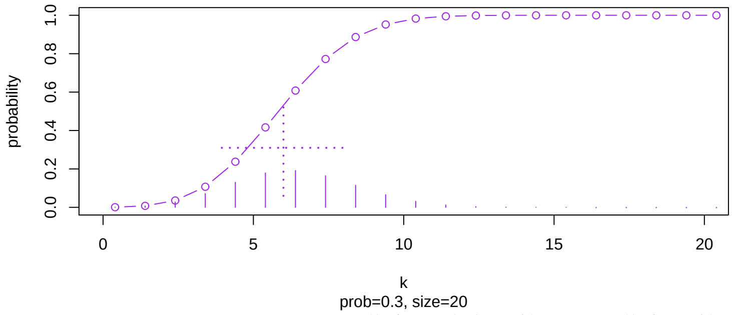

図 11 $n = 20, p = 0.3$ の二項分布の確率関数と累積分布関数

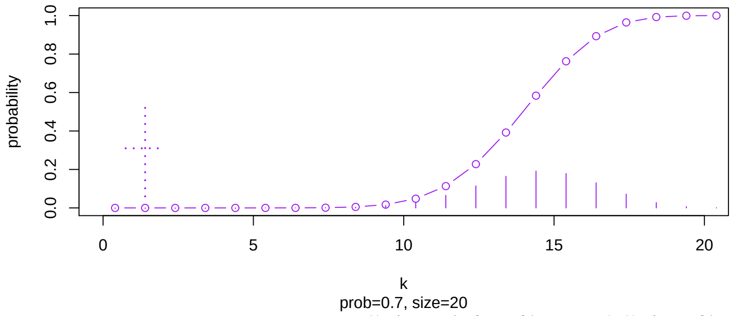

図 12 $n = 20, p = 0.7$ の二項分布の確率関数と累積分布関数

となる。

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。X の期待値 (二項分布の平均、二項分布の期待値) は

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

と項をくくりだして、二項展開に帰着させる。

この分布の分散については、X^2 の期待値 (= X^2 の平均)

$$E\left[X^2\right] = \sum_{k=0}^{n} k^2 {}_nC_k p^k (1-p)^{n-k} = \sum_{k=0}^{n} k^2 \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \tag{12}$$

あるいは X の分散 (= X^2 の平均 - X の平均の二乗)

$$V[X] = E\left[(X - E[X])^2\right] = E\left[X^2\right] - (E[X])^2 \tag{13}$$

のいずれかで計算を進めると遠回りになる。

二項定理に馴染む計算をするために、少し巧妙に、

$$V[X] = E[X(X-1)] + E[X] - (E[X])^2 \tag{14}$$

を用いると、計算が簡単になる。k^2 をかけて総和を求めるより、$k(k-1)$ をかけた総和を求める方が、二項定理に馴染みやすいのである。実際、

$$\begin{aligned}
E[X(X-1)] &= \sum_{k=0}^{n} k(k-1)\,{}_nC_k p^k (1-p)^{n-k} \\
&= \sum_{k=0}^{n} k(k-1)\frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\
&= \sum_{k=2}^{n} k(k-1)\frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\
&= \sum_{k=2}^{n} n(n-1)\frac{(n-2)!}{(k-2)!((n-2)-(k-2))!} p^2 \cdot p^{k-2} (1-p)^{(n-2)-(k-2)} \\
&= n(n-1)p^2 \sum_{l=0}^{n-2} \frac{(n-2)!}{l!((n-2)-l)!} p^l (1-p)^{(n-2)-l} \\
&= n(n-1)p^2 \{p + (1-p)\}^{n-2} \\
&= n(n-1)p^2
\end{aligned} \tag{15}$$

と (14) 式より、

$$V[X] = n(n-1)p^2 + np - (np)^2 = np(1-p) \tag{16}$$

と求まる。

二項分布のモーメント母関数は、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して求める。

まず成功確率 p のベルヌーイ試行のモーメント母関数が

$$\begin{aligned}
M_X(t) &= E\left[e^{tX}\right] = e^{t\cdot 0} p(0) + e^{t\cdot 1} p(1) \\
&= 1\cdot(1-p) + e^t \cdot p = pe^t + 1 - p
\end{aligned} \tag{17}$$

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

$$M_{X_i}(t) = 1\cdot p + e^t \cdot p,\ \ i = 1, 2, \ldots \tag{18}$$

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

$$\begin{aligned}
M_Y(t) &= \prod_{i=1}^{n} M_{X_i}(t) \\
&= \left(pe^t + 1 - p\right)^n
\end{aligned} \tag{19}$$

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、

$$M_X(t) = pe^t + 1 - p \tag{20}$$

$n = 2$ の場合、

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

$n=3$ の場合も、

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

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

$$\begin{aligned} M_X(t) &= \sum_{k=0}^{n} e^{kt}\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} \\ &= \sum_{k=0}^{n}\left(e^t\right)^k\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} \\ &= \sum_{k=0}^{n}\frac{n!}{k!(n-k)!}\left(pe^t\right)^k(1-p)^{n-k} \\ &= \left(pe^t+1-p\right)^n \end{aligned} \tag{21}$$

と、二項定理に合わせる必要がある。

1.3　幾何分布

二項分布は n 回の試行の中での成功回数 X の確率分布であった。今度は、1 回成功するまでに連続して失敗する回数 X の確率分布を考える。

再び、ベルヌーイ試行を繰り返すことを考える。成功確率 p のベルヌーイ試行の確率関数は

$$p(k)=p^k(1-p)^{1-k},\ k=0,1. \tag{22}$$

であった。k 回連続して失敗した後に成功する確率は、ベルヌーイ試行の確率関数と同様に

$$\begin{aligned} Pr[X=k] &= p_X(k) \\ &= Pr[X_1=0\ \text{ and }\ X_21=0\ \text{ and }\ \ldots\text{ and }\ X_k=0\ \text{ and }\ X_{k+1}=1] \\ &= Pr[X_1=0]\,Pr[X_2=0]\cdots Pr[X_k=0]\,Pr[X_{k+1}=1] \\ &= (1-p)^k p \end{aligned} \tag{23}$$

と求まる。この確率関数を持つ 確率分布 を幾何分布と言う。

幾何分布の確率関数

$$p(k)=(1-p)^k p$$

幾何分布の定義は前述の通り、「1 回成功するまでに連続して失敗する回数 X の確率分布」である。この確率関数の形状は、参考書 p.63 の図 5.3 および 5.4 を参照のこと。

幾何分布の期待値は

$$\begin{aligned} E_X[X] &= \sum_{k=0}^{\infty}kp_X(k)=\sum_{k=0}^{\infty}kp(1-p)^k \\ &= p(1-p)\sum_{k=0}^{\infty}k(1-p)^{k-1} \\ &= p(1-p)\sum_{k=0}^{\infty}\frac{\partial(1-p)^k}{\partial(1-p)} \\ &= p(1-p)\frac{\partial}{\partial(1-p)}\sum_{k=0}^{\infty}(1-p)^k \\ &= p(1-p)\frac{\partial}{\partial(1-p)}(1-(1-p))^{-1} \\ &= \frac{p(1-p)}{p^2}=\frac{1-p}{p} \end{aligned} \tag{24}$$

幾何分布の分散は、二項分布と同じく

$$Var[X] = E[X(X-1)] + E[X] - (E[X])^2$$

の関係から求めるのが筋がいい。

$$\begin{aligned}
E_X[X(X-1)] &= \sum_{k=0}^{\infty} k(k-1)\,p_X(k) = \sum_{k=2}^{\infty} k(k-1)\,p(1-p)^k \\
&= p(1-p)^2 \sum_{k=2}^{\infty} k(k-1)(1-p)^{k-2} \\
&= p(1-p)^2 \sum_{k=2}^{\infty} \frac{\partial^2 (1-p)^k}{\partial (1-p)^2} \\
&= p(1-p)^2 \frac{\partial \sum_{k=2}^{\infty} (1-p)^k}{\partial (1-p)} \\
&= p(1-p)^2 \frac{2}{(1-(1-p))^3} \\
&= \frac{2(1-p)^2}{p^2}
\end{aligned}$$

より、

$$\begin{aligned}
Var[X] &= \frac{2(1-p)}{p^2} + \frac{1-p}{p} - \frac{(1-p)^2}{p^2} \\
&= \frac{1-p}{p^2}
\end{aligned}$$

モーメント母関数も $t < \log(1-p)^{-1}$ の範囲で

$$\begin{aligned}
M_X(t) = E\left[e^{tX}\right] &= \sum_{k=0}^{\infty} e^{tk}(1-p)^k p \\
&= p\sum_{k=0}^{\infty} \left\{e^t(1-p)\right\}^k \\
&= \frac{p}{1-e^t(1-p)}
\end{aligned}$$

と求まる。参考書によると、分散はモーメント母関数の 2 階から求める方が簡単かもしれない。

1.4 負の二項分布

次に幾何分布を少し一般化する。独立なベルヌーイ試行の繰り返し $X_1, X_2, \ldots$ の中で、通算で r 回の成功が起こるまでに失敗した回数を Z とする。幾何分布は $r = 1$ の場合に相当するし、実は幾何分布に従う独立な確率変数を $Y_1, \ldots, Y_r$ とすると、負の二項分布は $Z = Y_1 + \cdots + Y_r$ 個の従う確率分布とも言える。

「$Z = k$」という事象は、「$k + r - 1$ 回の試行を終えた時点で $r - 1$ 回の成功が起きていて、$k + r$ 回目の試行で r 回目の試行が起こる」事象とも言える。$X_1, X_2, \ldots$ は各試行回で成功していれば 1、失敗していれば 0 であることを思い出せば、

$$Pr[Z = k] = Pr[X_1 + X_2 + \cdots + X_{k+r-1} = r-1 \ \text{ and } \ X_{k+r} = 1] \tag{25}$$

となる。これは $X_i, i = 1, 2, \ldots$ が互いに独立なことから、

$$\begin{aligned}
Pr[Z = k] &= Pr[X_1 + X_2 + \cdots + X_{k+r-1} = r-1]\,Pr[X_{k+r} = 1] \\
&= \left[{}_{k+r-1}C_{r-1}\,p^{r-1}(1-p)^k\right]\left[p^1(1-p)^0\right] \\
&= {}_{k+r-1}C_{r-1}\,p^r(1-p)^k
\end{aligned} \tag{26}$$

この確率関数を持つ確率分布を 負の二項分布 という。

負の二項分布の確率関数

$$p(k) = {}_{k+r-1}C_{r-1} p^r (1-p)^k$$

負の二項分布の確率関数の例は、参考書 p.65 の図 5.5 および 5.6 を参照のこと。

負の二項分布の平均と分散はそれぞれ

$$E[Z] = \frac{r(1-p)}{p} \tag{27}$$

$$V[Z] = \frac{r(1-p)}{p^2} \tag{28}$$

となる。モーメント母関数は、負の二項分布が幾何分布の和の分布であることを利用して、$t < \log(1-p)^{-1}$ の範囲で

$$M_Z(t) = \left\{ \frac{p}{1-(1-p)e^t} \right\}^r \tag{29}$$

となる。

参考文献

[1] 久保木久孝 (2007)「確率・統計解析の基礎」, 朝倉書店. (教科書)

[2] 宮川雅巳 (1998)「統計技法」工系数学講座 14, 共立出版. (教科書)

[3] 永田靖 (2005)「統計学のための数学入門 30 講」科学のことばとしての数学, 朝倉書店. (講義・教科書などで、計算が追えないところが見つかった時、それを尋ねるつもりで開くと助けてくれる本)

[4] 藤田岳彦 (2010)「弱点克服 大学生の確率・統計」東京図書. (講義・教科書などで、計算を追えないところが見つかった時、それを尋ねるつもりで開くと助けてくれる本)

[5] 東京大学教養学部統計学教室・編 (1991)「統計学入門」基礎統計学 I, 東京大学出版会. (教科書)

[6] 薩摩純吉 (1989)「確率・統計」理工系の数学入門コース 7, 岩波書店. (教科書)