27 juin 2013
PROB #11

確率論 (*Probability Theory*) 第 11 週 ベルヌーイ試行に関わる分布たち

w.yamamoto

1 二項分布

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。これをベルヌーイ試行と呼び、その確率関数は

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

となる。毎回の試行を独立[1]とすると、この試行を 2 回繰り返したとき、成功する回数は $\{0, 1, 2\}$ の何れかとなる。それぞれの値を取る確率は

$$\begin{aligned} p(0) &= Pr[X_1 = 0 \; and \; X_2 = 0] \\ &= Pr[X_1 = 0]\, Pr[X_2 = 0] = (1-p)^2 \qquad (2) \\ p(1) &= Pr[\{X_1 = 0 \; and \; X_2 = 1\} \; or \, \{X_1 = 1 \; and \; X_2 = 0\}] \\ &= Pr[X_1 = 1]\, Pr[X_2 = 0] + Pr[X_1 = 0]\, Pr[X_2 = 1] \\ &= p(1-p) + (1-p)\, p = 2p(1-p) \\ &= {}_2C_1 p(1-p) \qquad (3) \\ p(2) &= Pr[X_1 = 1 \; and \; X_2 = 1] \\ &= Pr[X_1 = 1]\, Pr[X_2 = 1] = p^2 \qquad (4) \end{aligned}$$

n 回の互いに独立な試行で k 回の成功を観測する確率は、

$$p(k) = Pr[X = k] = {}_nC_k p^k (1-p)^{n-k} = \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k} \tag{5}$$

となる。この確率関数を持つ確率分布は <u>二項分布</u> と呼ばれる。この確率分布の名前の由来は、次のような <u>二項展開</u> に依る。

$$\begin{aligned} (p + (1-p))^n &= 1 \\ &= {}_nC_0 p^0 (1-p)^{n-0} + {}_nC_1 p^1 (1-p)^{n-1} + {}_nC_2 p^2 (1-p)^{n-2} + \cdots \\ &\quad + {}_nC_{n-1} p^{n-1} (1-p)^1 + {}_nC_n p^n (1-p)^0 \qquad (6) \end{aligned}$$

この式は、全確率が 1 であることも意味している。

確率関数

$$p(k) = p(k; n, p) = Pr[X = k] = {}_nC_k p^k (1-p)^{n-k} \tag{7}$$

累積分布関数

$$F(k) = Pr[X \le k] = \sum_{l=0}^{k} {}_nC_l p^l (1-p)^{n-l} \tag{8}$$

ここまでをグラフに表すと、次の通り。まずは確率関数から。

[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。

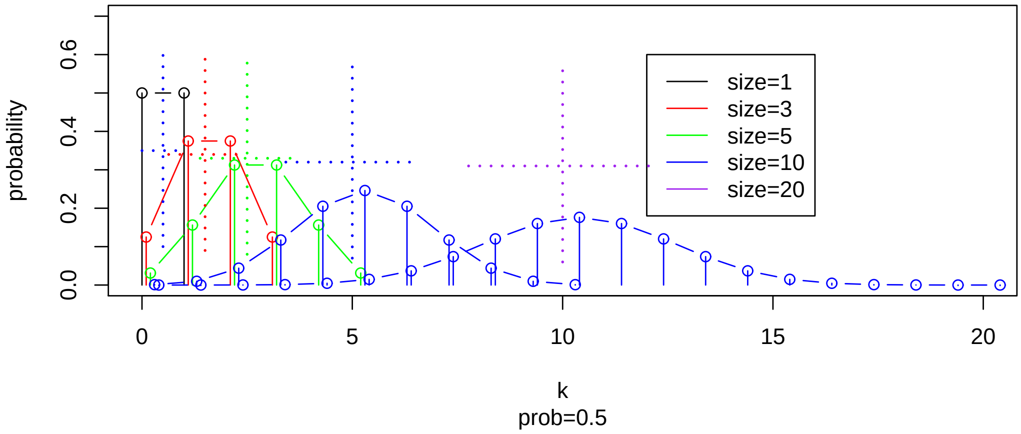

図 4 $p = 0.5$ の二項分布の確率関数 ($n = 1, 3, 5, 10, 20$)

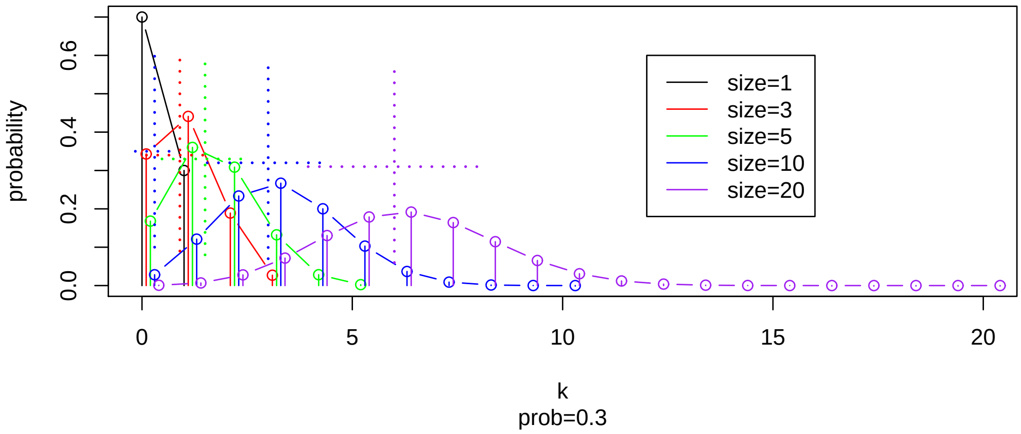

図 5 $p = 0.3$ の二項分布の確率関数 ($n = 1, 3, 5, 10, 20$)

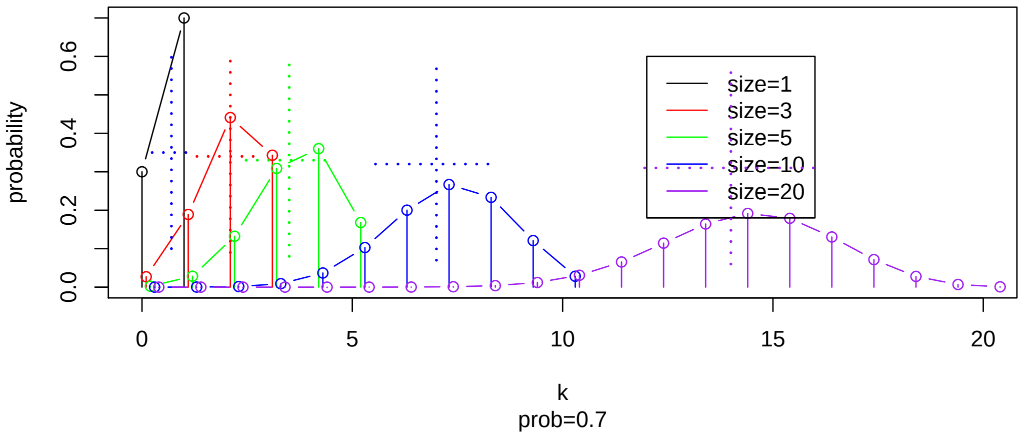

図 6 $p = 0.7$ の二項分布の確率関数 ($n = 1, 3, 5, 10, 20$)

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。また、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。

以下は、確率関数と累積分布関数の関係の例である。総試行回数が 2 の場合は、

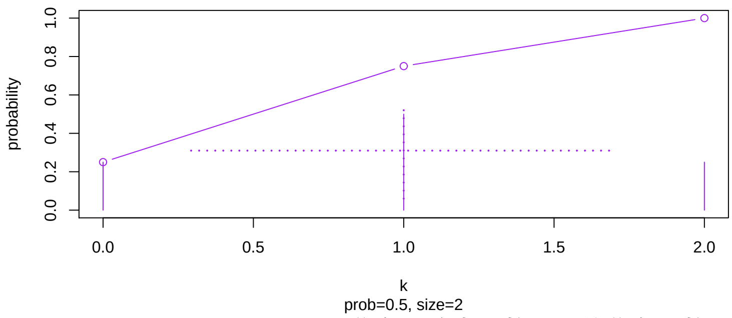

図 7 $n = 2, p = 0.5$ の二項分布の確率関数と累積分布関数

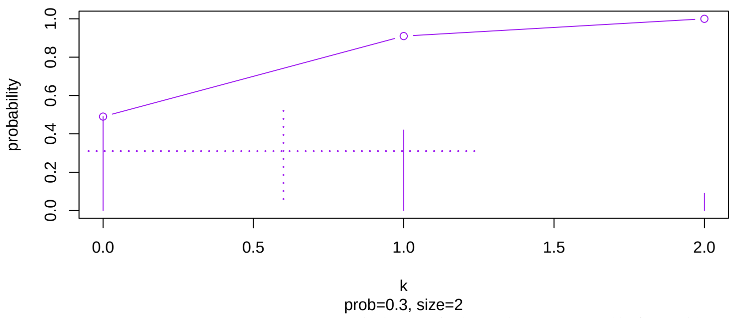

図 8 $n = 2, p = 0.3$ の二項分布の確率関数と累積分布関数

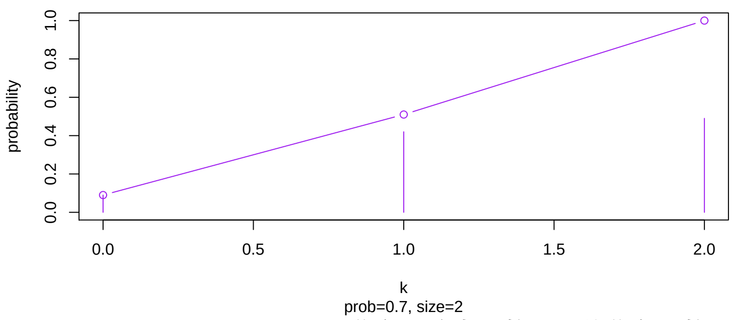

図 9 $n = 2, p = 0.7$ の二項分布の確率関数と累積分布関数

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、

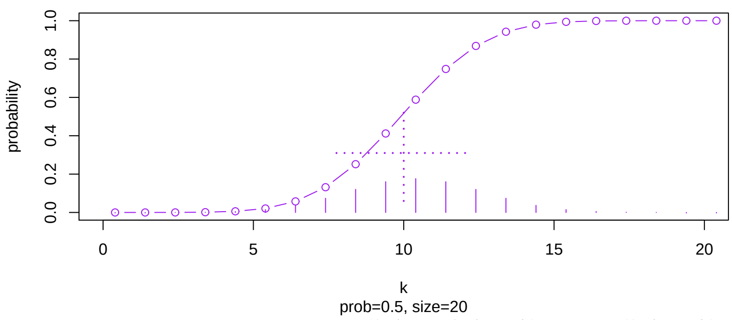

図 10 $n = 20, p = 0.5$ の二項分布の確率関数と累積分布関数

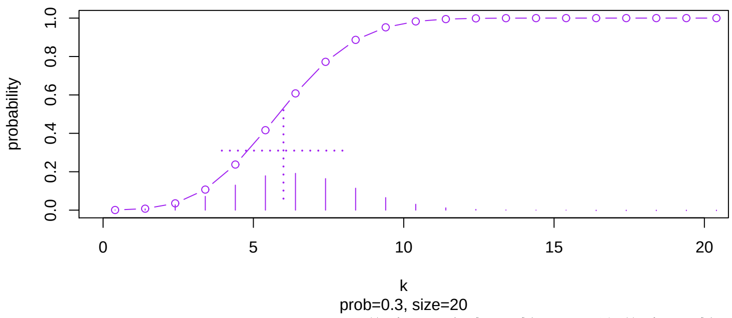

図 11 $n = 20,\ p = 0.3$ の二項分布の確率関数と累積分布関数

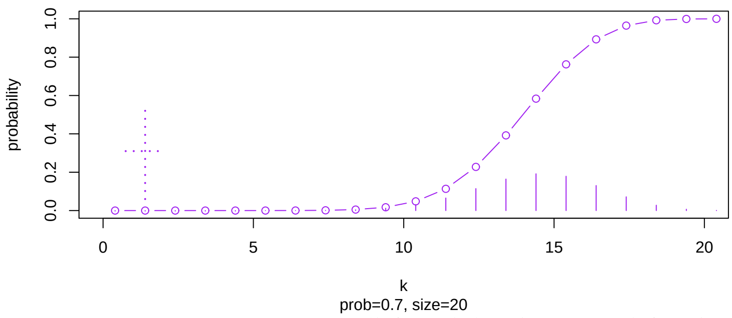

図 12 $n = 20,\ p = 0.7$ の二項分布の確率関数と累積分布関数

となる。

X の期待値 (= X の平均)

$$E[X] = \sum_{k=0}^{n} k_n C_k p^k (1-p)^{n-k} = \sum_{k=0}^{n} k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = ? \tag{9}$$

X^2 の期待値 (= X^2 の平均)

$$E\left[X^2\right] = \sum_{k=0}^{n} k^2{}_n C_k p^k (1-p)^{n-k} = \sum_{k=0}^{n} k^2 \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = ? \tag{10}$$

X の分散 (= X^2 の平均 - X の平均の二乗)

$$V[X] = E\left[(X - E[X])^2\right] = E\left[X^2\right] - (E[X])^2 = ? \tag{11}$$

1.1 二項分布にまつわる計算

組み合わせの数

$${}_nC_k = \frac{n!}{k!(n-k)!} = \begin{pmatrix} n \\ k \end{pmatrix} \tag{12}$$

二項定理

$$(a+b)^n = \sum_{k=0}^{n} {}_nC_k a^k b^{n-k} = \sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} a^k b^{n-k} \tag{13}$$

これらを、二項分布では期待値や分散の計算に用いる。期待値では

$$
\begin{aligned}
E[X] &= \sum_{k=0}^{n} k_n C_k p^k (1-p)^{n-k} \\
&= \sum_{k=0}^{n} k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} n \frac{(n-1)!}{(k-1)!((n-1)-(k-1))!} p \cdot p^{k-1} (1-p)^{(n-1)-(k-1)} \\
&= np \sum_{l=0}^{n-1} \frac{(n-1)!}{l!((n-1)-l)!} p^l (1-p)^{(n-1)-l} \\
&= np \{p + (1-p)\}^{n-1} \\
&= np
\end{aligned} \tag{14}
$$

分散では、

$$
V[X] = E\left[(X-\mu)^2\right] = E\left[X^2\right] - (E[X])^2 \tag{15}
$$

のどちらでもなく、少し巧妙に、

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

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。

$$
\begin{aligned}
E[X(X-1)] &= \sum_{k=0}^{n} k(k-1)\,{}_nC_k p^k (1-p)^{n-k} \\
&= \sum_{k=0}^{n} k(k-1) \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\
&= \sum_{k=2}^{n} k(k-1) \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\
&= \sum_{k=2}^{n} n(n-1) \frac{(n-2)!}{(k-2)!((n-2)-(k-2))!} p^2 \cdot p^{k-2} (1-p)^{(n-2)-(k-2)} \\
&= n(n-1)p^2 \sum_{l=0}^{n-2} \frac{(n-2)!}{l!((n-2)-l)!} p^l (1-p)^{(n-2)-l} \\
&= n(n-1)p^2 \{p + (1-p)\}^{n-2} \\
&= n(n-1)p^2
\end{aligned} \tag{17}
$$

より、

$$
V[X] = n(n-1)p^2 + np - (np)^2 = np(1-p) \tag{18}
$$

1.2 二項分布のモーメント母関数

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。

成功確率 p のベルヌーイ試行の確率関数は

$$
p(k) = p^k (1-p)^{1-k}, \; k = 0, 1. \tag{19}
$$

この確率関数を持つ確率分布のモーメント母関数は

$$\begin{aligned} M_X(t) &= E\left[e^{tX}\right] = e^{t\cdot 0}p(0) + e^{t\cdot 1}p(1) \\ &= 1\cdot(1-p) + e^t \cdot p = pe^t + 1 - p \end{aligned} \tag{20}$$

$X_1, X_2, \ldots$ 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上と同じく

$$M_{X_i}(t) = 1\cdot p + e^t \cdot p, \ \ i = 1, 2, \ldots \tag{21}$$

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となるから、Y の確率分布は二項分布となる。このとき、Y の確率分布のモーメント母関数は

$$\begin{aligned} M_Y(t) &= \prod_{i=1}^{n} M_{X_i}(t) \\ &= \left(pe^t + 1 - p\right)^n \end{aligned} \tag{22}$$

別解、直接計算してみると、二項分布のモーメント母関数は、$n = 3$ ぐらいまでならなんとか計算できる。$n = 1$ の場合、

$$M_X(t) = pe^t + 1 - p \tag{23}$$

$n = 2$ の場合、

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

$n = 3$ の場合も、

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

実は一般の n についても

$$\begin{aligned} M_X(t) &= \sum_{k=0}^{n} e^{kt}\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} \\ &= \sum_{k=0}^{n} \left(e^t\right)^k \frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} \\ &= \sum_{k=0}^{n} \frac{n!}{k!(n-k)!}\left(pe^t\right)^k(1-p)^{n-k} \\ &= \left(pe^t + 1 - p\right)^n \end{aligned} \tag{24}$$

と求めることができる。最後の式は二項定理による。

2 幾何分布と負の二項分布

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 Y 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

$$p(k) = p^k(1-p)^{1-k}, \ k = 0, 1. \tag{25}$$

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

$$\begin{aligned} Pr[Y=k] &= p_Y(k) \\ &= Pr[X_1=0 \text{ and } X_2 1=0 \text{ and } \ldots \text{ and } X_k=0 \text{ and } X_{k+1}=1] \\ &= Pr[X_1=0]\,Pr[X_2=0]\cdots Pr[X_k=0]\,Pr[X_{k+1}=1] \\ &= (1-p)^k p \end{aligned} \tag{26}$$

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 Y 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

$$\begin{aligned} E_Y[Y] &= \sum_{k=0}^{\infty} k p_Y(k) = \sum_{k=0}^{\infty} kp(1-p)^k \\ &= p(1-p)\sum_{k=0}^{\infty} k(1-p)^{k-1} \\ &= p(1-p)\sum_{k=0}^{\infty} \frac{\partial (1-p)^k}{\partial (1-p)} \\ &= p(1-p)\frac{\partial \sum_{k=0}^{\infty}(1-p)^k}{\partial (1-p)} \\ &= p(1-p)\frac{\partial (1-(1-p))^{-1}}{\partial (1-p)} \\ &= \frac{p(1-p)}{p^2} = \frac{1-p}{p} \end{aligned} \tag{27}$$

モーメント母関数も $t < \log(1-p)^{-1}$ の範囲で

$$M_Y(t) = \frac{p}{1-(1-p)e^t} \tag{28}$$

と求まる。分散の計算を飛ばしたのは、参考書によれば、モーメント母関数から求める方が簡単なため。(今日の課題 8-1 および 8-2)

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 $ZY = Y_1 + \cdots + Y_r$ 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、

$$Pr[Z=k] = Pr[X_1+X_2+\cdots+X_{k+r-1}=r-1 \text{ and } X_{k+r}=1] \tag{29}$$

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

$$\begin{aligned} Pr[Z=k] &= Pr[X_1+X_2+\cdots+X_{k+r-1}=r-1]\,Pr[X_{k+r}=1] \\ &= \left[{}_{k+r-1}C_{r-1}p^{r-1}(1-p)^k\right]\left[p^1(1-p)^0\right] \\ &= {}_{k+r-1}C_{r-1}p^r(1-p)^k \end{aligned} \tag{30}$$

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

$$E[Z] = \frac{r(1-p)}{p} \tag{31}$$

$$V[Z] = \frac{r(1-p)}{p^2} \tag{32}$$

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

$$M_Z(t) = \left\{\frac{p}{1-(1-p)e^t}\right\}^r \tag{33}$$

となる。

3　レポート課題

確率分布 F のモーメント母関数は、

$$E\left[e^{tX}\right] = \sum_{k\in\mathcal{X}} \exp(tk)\,p(k) \tag{34}$$

あるいは

$$E\left[e^{tX}\right] = \int_{x\in\mathcal{X}} \exp(tk)\,f(k)\,dx \tag{35}$$

で定義される。

#6-1　二項分布と幾何分布の平均、分散、モーメント母関数を計算できるようにしておくこと。

#6-2　二項分布と幾何分布のモーメント母関数を、$t=0$ の回りでマクローリン展開し、平均と分散を計算できるようにしておくこと。

3.1　マクローリン展開とテイラー展開

無限回微分可能な関数 $g(x)$ の k 階の導関数を $g^{(k)}$ とおく。まず、テイラーの公式

$$g(x) = \sum_{k=0}^{n-1} \frac{g^{(k)}(a)}{k!} + R_n(x)(x-a)^k \tag{36}$$

を既知とする。R_n はこの式の等号が成り立つように定められる剰余項である。ロルの定理から x に依存して決まり、区間 (a,b) の値を取る関数 $c(x)$ を用いて

$$R_n(x) = \frac{g^{(n)}(c(x))}{n!}(x-a)^n \tag{37}$$

と表せる。

テイラーの公式で $n\to\infty$ とせず、有限の n のまま、$x=a$ の近傍で

$$g(x) \simeq \sum_{k=0}^{n} \frac{g^{(k)}(a)}{k!}(x-a) \tag{38}$$

とすると、$g(x)$ の $x=a$ の近傍での多項式近似が得られる。ただし、この近似が良い精度を持つためには、剰余項が

$$R_n(x) = o\left((x-a)^{n-1}\right) \tag{39}$$

を満たさなければならない。

$g(x)$ のマクローリン展開は、テイラーの公式で $a-0$ と置き、さらに $n\to\infty$ の極限

$$\begin{aligned} g(x) &= \sum_{k=0}^{\infty} \frac{g^{(k)}(0)}{k!}x^k \\ &= g(0) + \sum_{k=1}^{\infty} \frac{g^{(k)}(0)}{k!}x^k \\ &= g(0) + g'(0)\,x + \frac{g''(0)}{2}x^2 + \frac{g'''(0)}{6}x^3 + \cdots \end{aligned} \tag{40}$$

として与えられる。マクローリン展開の展開の中心を $x = 0$ ではなく $x = a$ とすると、テイラー展開と言われる。

$$\begin{aligned} g(x) &= \sum_{k=0}^{\infty} \frac{g^{(k)}(a)}{k!}(x-a)^k \\ &= g(a) + \sum_{k=1}^{\infty} \frac{g^{(k)}(a)}{k!}(x-a)^k \\ &= g(a) + g'(a)(x-a) + \frac{g''(a)}{2}(x-a)^2 + \frac{g'''(a)}{6}(x-a)^3 + \cdots \end{aligned} \tag{41}$$

テイラー展開やマクローリン展開は、テイラーの公式の極限であり、これら存在する条件は、テイラーの公式の剰余項が

$$\lim_{n\to\infty} R_n(x) = 0 \tag{42}$$

を満たすことである。

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)