二項分布のポアソン分布による近似について

高知高専
高木和久

ポアソン分布と二項分布

ポアソン分布

$$\mathrm{Poi}(k,\lambda)=\frac{\lambda^k}{k!}\cdot e^{-\lambda}$$

二項分布

$$\mathrm{Bin}(k,n,p)=\frac{n!}{(n-k)!k!}p^k(1-p)^{n-k}$$

表 1.　2つの分布の比較

k	n	p	Poi(k,np)	Bin(k,n,p)	比
30	100	0.5	0.00068	0.00002	34.0
3	10	0.6	0.08924	0.04247	2.1
3	10	0.7	0.05213	0.00900	5.8
3	10	0.8	0.02863	0.00079	36.2

$n\to\infty$ のとき$\mathrm{Bin}(k,n,p)\to$ $\mathrm{Poi}(k,np)$であることの証明

$\lambda=np$ とおくと $p=\frac{\lambda}{n}$ だから

$$\mathrm{Bin}(k,n,p)=\frac{n!}{(n-k)!\,k!}p^k(1-p)^{n-k}$$

$$=\frac{n!}{(n-k)!\,k!}\,\frac{\lambda^k}{n^k}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}$$

$$=\frac{n!}{(n-k)!\,n^k}\,\frac{\lambda^k}{k!}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}$$

$$\mathrm{Bin}(k,n,p) = \frac{n!}{(n-k)!\,n^k}\,\frac{\lambda^k}{k!}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}$$

$n\to\infty$ のとき $\left(1-\frac{\lambda}{n}\right)^{-k}\to 1$ である。

また $\left(1+\frac{a}{n}\right)^n \to e^a$ であるから

$$\frac{\lambda^k}{k!}\left(1-\frac{\lambda}{n}\right)^n \to \frac{\lambda^k}{k!}e^{-\lambda} = \mathrm{Poi}(k,\lambda)$$

そして

$$\frac{n!}{(n-k)!\,n^k} = \frac{n}{n}\cdot\frac{n-1}{n}\cdot\frac{n-2}{n}\cdot\cdots\cdot\frac{n-k+1}{n} \to 1$$

であるから $n\to\infty$ のとき $\mathrm{Bin}(k,n,p)\to\ \mathrm{Poi}(k,np)$ が成り立つ。（証明終）

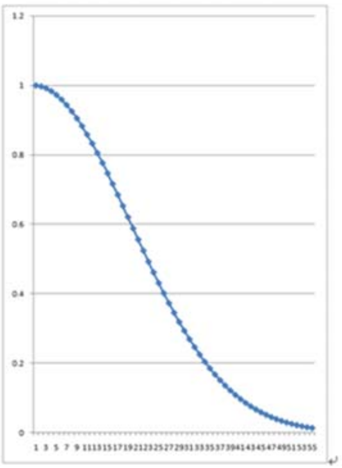
誕生日の確率

k 人の集団がある。この中のどの 2 人の誕生日も同じでない確率はどの位であろうか。
k の値が与えられた時、この確率は次の式で計算できる。

$$\frac{365}{365}\cdot\frac{364}{365}\cdot\frac{363}{365}\cdot\frac{362}{365}\cdot\cdots\cdot\frac{365-k+1}{365}$$

$$=\prod_{j=1}^{k-1}\left(1-\frac{j}{365}\right)$$

誕生日の確率



誕生日の数 365 を n で置き換えて得られる
確率を $\mathrm{Bir}(k,n)$ で表わすと、

$$\mathrm{Bir}(k,n) = \frac{n}{n}\cdot\frac{n-1}{n}\cdot\frac{n-2}{n}\cdot\cdots\cdot\frac{n-k+1}{n}$$

$$=\prod_{j=1}^{k-1}\left(1-\frac{j}{n}\right)$$

となる。

$$\mathrm{Bin}(k,n,p) = \frac{n!}{(n-k)!\,n^k}\ \frac{\lambda^k}{k!}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}$$

$$\therefore\ \mathrm{Bin}(k,n,p) = \mathrm{Bir}(k,n)\ \frac{\lambda^k}{k!}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}$$

近似式 $e^x \sim 1+x$ で $x=\frac{a}{n}$ とおくと

$$e^{\frac{a}{n}} \sim 1+\frac{a}{n}$$

が成り立つから

$$\left(1+\frac{a}{n}\right)^n \sim e^a$$

を得る。

$$\frac{\lambda^k}{k!}\left(1-\frac{\lambda}{n}\right)^n \sim \frac{\lambda^k}{k!}\,e^{-\lambda} = \mathrm{Poi}(k,\lambda)$$

$$\mathrm{Bin}(k,n,p) = \frac{n!}{(n-k)!\,n^k}\ \frac{\lambda^k}{k!}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}$$

$$\therefore\ \mathrm{Bin}(k,n,p) = \mathrm{Bir}(k,n)\ \frac{\lambda^k}{k!}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}$$

ここで $\lambda = np$ とおくと

$$\frac{\lambda^k}{k!}\left(1-\frac{\lambda}{n}\right)^n \sim \frac{\lambda^k}{k!}\,e^{-\lambda} = \mathrm{Poi}(k,\lambda)$$

より

$$\therefore\ \mathrm{Bin}(k,n,p) \sim \mathrm{Bir}(k,n)\ \mathrm{Poi}(k,\lambda)\left(1-\frac{\lambda}{n}\right)^{-k}$$

対数関数の新しい近似式を作る

$$\ln x \sim 2 \cdot \frac{x-1}{x+1} \cdots ①$$

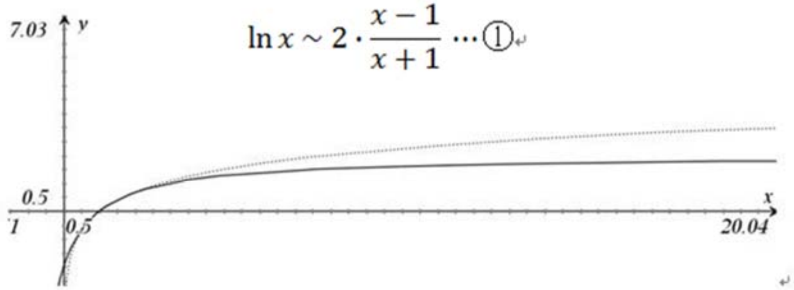


図 1. $y = \ln x$ のグラフ（点線）と $y = 2 \cdot \frac{x-1}{x+1}$ のグラフ（実線）

$$\ln x \sim \frac{x-1}{\sqrt{x}} \cdots ②$$

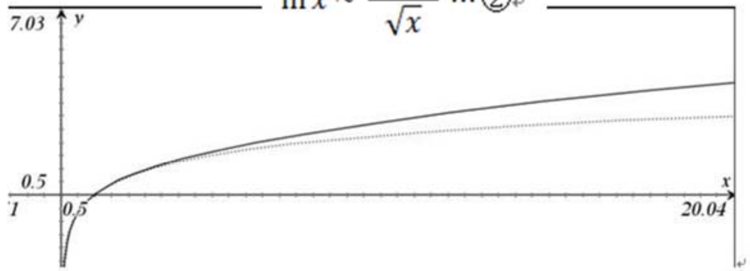


図 2. $y = \ln x$ のグラフ（点線）と $y = \frac{x-1}{\sqrt{x}}$ のグラフ（実線）

$a < b$ のとき不等式

$$\frac{b-a}{\frac{a+b}{2}} < \ln\frac{b}{a} < \frac{b-a}{\sqrt{ab}}$$

が成り立っている.そこで,

$$\sqrt{ab} < \mathrm{M}(a,b) < \frac{a+b}{2}$$

が成り立つような $\mathrm{M}(a,b)$を用いて

$$\ln\frac{b}{a} \sim \frac{b-a}{\mathrm{M}(a,b)}$$

という形で $\ln\frac{b}{a}$ を近似することにする.

$$\mathrm{M}(a,b) = \frac{\frac{a+b}{2} + 2\sqrt{ab}}{3} = \frac{a+b+4\sqrt{ab}}{6}$$

このとき

$$\ln\frac{b}{a} \sim \frac{b-a}{\mathrm{M}(a,b)} = \frac{6(b-a)}{a+b+4\sqrt{ab}}$$

すなわち

$$\ln x \sim \frac{6(x-1)}{x+1+4\sqrt{x}} \cdots ③$$

$$\ln x \sim \frac{6(x-1)}{x+1+4\sqrt{x}} \quad \cdots ③$$

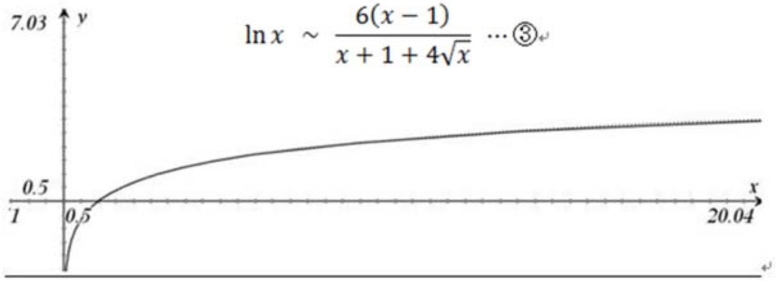


図 3. $y = \ln x$ のグラフ（点線）と $y = \frac{6(x-1)}{x+1+4\sqrt{x}}$ のグラフ（実線）

マクローリン展開による精度の検証

$$\ln(x+1) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots$$

$$\sqrt{1+x} = 1 + \frac{x}{2} - \frac{x^2}{8} + \frac{x^3}{16} + \cdots$$

$$\ln(x+1)\cdot\left(x+2+4\sqrt{x+1}\right) = 6x + \frac{x^5}{480} - \frac{x^6}{240} + \cdots$$

$$\ln(x+1)\cdot\left(x+2+4\sqrt{x+1}\right) \sim 6x$$

この近似式は x^2, x^3, x^4 の項が消えていて,かなり精度が高い.

x を $x-1$ で置き換えて

$$\ln x \cdot \left(x+1+4\sqrt{x}\right) \sim 6(x-1)$$

これより近似式

$$\ln x \sim \frac{6(x-1)}{x+1+4\sqrt{x}}$$

を得ることができる.

平均Mの満たす性質

1. $\mathrm{M}(b,a) = \mathrm{M}(a,b)$
2. $a < b$ のとき $a < \mathrm{M}(a,b) < b$
3. $k > 0$ のとき $\mathrm{M}(ka,kb) = k\cdot\mathrm{M}(a,b)$
4. $\sqrt{ab} \le \mathrm{M}(a,b) \le \frac{a+b}{2}$
5. $\mathrm{M}(1,2), \mathrm{M}(2,3), \cdots, \mathrm{M}(n-1,n)$ の調和平均は近似的に $\mathrm{M}(1,n)$ に等しい。即ち、
$$\frac{n-1}{\mathrm{M}(1,n)} \sim \frac{1}{\mathrm{M}(1,2)} + \frac{1}{\mathrm{M}(2,3)} + \cdots + \frac{1}{\mathrm{M}(n-1,n)}$$
6. $\frac{ab-1}{\mathrm{M}(ab,1)} \sim \frac{a-1}{\mathrm{M}(a,1)} + \frac{b-1}{\mathrm{M}(b,1)}$
7. $\mathrm{M}(a^2,b^2) \sim \frac{a+b}{2}\cdot\mathrm{M}(a,b)$

近似式

$$\ln x \sim \frac{6(x-1)}{x+1+4\sqrt{x}} \cdots ③$$

で x に $1-p$ を代入すると

$$\ln(1-p) \sim \frac{-p}{\mathrm{M}(1-p,1)}$$

よって任意の正の整数 n に対し

$$n\cdot\ln(1-p) \sim \frac{-np}{\mathrm{M}(1-p,1)}$$

が成り立つ。

λ を次のように定める。

$$\lambda = \frac{np}{\mathrm{M}(1-p,1)}$$

すると

$$\ln(1-p)^n = n\cdot\ln(1-p) \sim -\lambda$$

$$\therefore (1-p)^n \sim e^{-\lambda}$$

となる。 $\frac{np}{\lambda} = \mathrm{M}(1-p,1)$ であるから

$$\frac{np}{(1-p)\lambda} = \frac{1}{1-p}\cdot\mathrm{M}(1-p,1) = \mathrm{M}\left(1,\frac{1}{1-p}\right)$$

となる。

$$\mathrm{Bin}(k,n,p) = \frac{n!}{(n-k)!k!}p^k(1-p)^{n-k} = \frac{n!}{(n-k)!k!}p^k(1-p)^{-k}(1-p)^n$$

$$= \frac{n!}{(n-k)!n^k}\cdot\frac{p^k n^k}{(1-p)^k}\cdot\frac{1}{k!}(1-p)^n = \frac{n!}{(n-k)!n^k}\cdot\frac{p^k n^k}{(1-p)^k\lambda^k}\cdot\frac{\lambda^k}{k!}(1-p)^n$$

$$= \mathrm{Bir}(k,n)\cdot\left\{\frac{pn}{(1-p)\lambda}\right\}^k\cdot\frac{\lambda^k}{k!}(1-p)^n = \mathrm{Bir}(k,n)\cdot\left\{\mathrm{M}\left(1,\frac{1}{1-p}\right)\right\}^k\cdot\frac{\lambda^k}{k!}(1-p)^n$$

$$\sim \mathrm{Bir}(k,n)\cdot\left\{\mathrm{M}\left(1,\frac{1}{1-p}\right)\right\}^k\cdot\frac{\lambda^k}{k!}e^{-\lambda}$$

$$\mathrm{Bin}(k,n,p) \sim \mathrm{Bir}(k,n)\cdot\left\{\mathrm{M}\left(1,\frac{1}{1-p}\right)\right\}^k\cdot\mathrm{Poi}(k,\lambda)$$

が成り立つ。$\lambda = \frac{np}{\mathrm{M}(1-p,1)}$ であったから、

近似式

$$\mathrm{Bin}(k,n,p) \sim \mathrm{Bir}(k,n)\cdot\mathrm{Poi}\left(k,\frac{np}{\mathrm{M}(1-p,1)}\right)\cdot\left\{\mathrm{M}\left(1,\frac{1}{1-p}\right)\right\}^k \cdots ④$$

を得る。

k	n	p	$\mathrm{Poi}(k,np)$	$\mathrm{Bin}(k,n,p)$	$\mathrm{Poi}(k,np)$ /$\mathrm{Bin}(k,n,p)$	近似式④ /$\mathrm{Bin}(k,n,p)$
50	200	0.2	0.0177	0.0149	1.2	1.0000
50	100	0.7	0.0024	0.0000	180.5	1.0878
30	100	0.3	0.0726	0.0868	0.8	1.0002
30	100	0.5	0.0007	0.0000	29.2	1.0055
30	50	0.8	0.0185	0.0006	30.2	1.1895
25	50	0.5	0.0795	0.1123	0.7	1.0027
15	20	0.8	0.0992	0.1746	0.6	1.0719
10	20	0.5	0.1251	0.1762	0.7	1.0011
5	15	0.5	0.1094	0.0916	1.2	1.0008
3	10	0.5	0.1404	0.1172	1.2	1.0005
3	10	0.6	0.0892	0.0425	2.1	1.0022
3	10	0.7	0.0521	0.0090	5.8	1.0084
3	10	0.8	0.0286	0.0008	36.4	1.0353
3	10	0.9	0.0150	0.0000	1714.0	1.2117

従来の近似が使える場合とは

$p\sim 0$ の時 $1-p\sim 1$ だから $\mathrm{M}(1-p,1)\sim 1$ であり、更に

$$\mathrm{M}\left(1,\frac{1}{1-p}\right) = \frac{1}{1-p}\,\mathrm{M}(1-p,1)\sim 1$$

が成り立つから近似式④は次の形になる。

$$\mathrm{Bin}(k,n,p)\sim \mathrm{Bir}(k,n)\cdot \mathrm{Poi}(k,np)\ \cdots ⑤$$

定義より

$$\mathrm{Poi}(n-k,n) = \frac{n^{n-k}}{(n-k)!}\cdot e^{-n}$$

だから、$k\geq 2$ のとき

$$\mathrm{Poi}(n-k,n) = \frac{n^n}{n^k e^n (n-k)!} = \frac{n^n\cdot(n-k+1)(n-k+2)\cdots(n-1)n}{e^n(n-k)!\cdot(n-k+1)(n-k+2)\cdots(n-1)n\cdot n^k}$$

$$= \frac{n^n\cdot n(n-1)(n-2)\cdots(n-k+1)}{e^n\cdot n!\cdot n^k} = \mathrm{Poi}(n,n)\cdot\prod_{j=1}^{k-1}\left(1-\frac{j}{n}\right) = \mathrm{Poi}(n,n)\cdot \mathrm{Bir}(k,n)$$

$$\therefore \mathrm{Bir}(k,n) = \frac{\mathrm{Poi}(n-k,n)}{\mathrm{Poi}(n,\ n)}$$

kに対してnが十分大きい時は

$$\mathrm{Bir}(k,n) = \frac{\mathrm{Poi}(n-k,n)}{\mathrm{Poi}(n,\ n)}\sim 1$$

が成り立つから、

$$\mathrm{Bin}(k,n,p)\sim \mathrm{Bir}(k,n)\cdot \mathrm{Poi}(k,np)\ \cdots ⑤$$

より

$$\mathrm{Bin}(k,n,p)\ \sim\ \mathrm{Poi}(k,np)$$

を得る。

従来の近似が使える場合は

$$\mathrm{Bin}(k, n, p) \sim \mathrm{Poi}(k, np)$$

- pの値が小さく
- そしてnがkに対して十分大きい場合

- ご清聴ありがとうございました