§3.1　2 項分布

2 項定理を思い起こして下さい：自然数 n 及び数 a と b とに対して，

$$
\begin{aligned}
(a+b)^n &= {}_n\mathrm{C}_0\,a^n + {}_n\mathrm{C}_1\,a^{n-1}b + {}_n\mathrm{C}_2\,a^{n-2}b^2 + {}_n\mathrm{C}_3\,a^{n-3}b^3 + \cdots + \\
&\qquad {}_n\mathrm{C}_{n-3}\,a^3b^{n-3} + {}_n\mathrm{C}_{n-2}\,a^2b^{n-2} + {}_n\mathrm{C}_{n-1}\,ab^{n-1} + {}_n\mathrm{C}_n\,b^n \\
&= \sum_{k=0}^{n}({}_n\mathrm{C}_k\,a^{n-k}b^k) \ .
\end{aligned}
$$

実数 p に対して，$a=1-p \quad b=p$ とします.

$$(1-p+p)^n = \sum_{k=0}^{n}\{{}_n\mathrm{C}_k\,p^k(1-p)^{n-k}\} \ .$$

この等式の左辺は $(1-p+p)^n = 1^n = 1$ なので，

$$\sum_{k=0}^{n}\{{}_n\mathrm{C}_k\,p^k(1-p)^{n-k}\} = 1 \ .$$

このことより，次のような確率変数 X を考えることができます：X のとり得る値は 0 以上 n 以下の自然数 $k=0,1,2,3,\ldots,n-2,n-1,n$ であり，$\mathrm{P}(\,X=k\,) = {}_n\mathrm{C}_k\,p^k(1-p)^{n-k}$ ．このとき，

$$\sum_{k=0}^{n}\mathrm{P}(\,X=k\,) = \sum_{k=0}^{n}\{{}_n\mathrm{C}_k\,p^k(1-p)^{n-k}\} = 1 \ .$$

従って変数 X は離散型確率変数です.

定義　定数 n は自然数で，定数 p は実数で $0\le p\le 1$ とする．確率変数 X が 2 項分布 (binomial distribution) $\mathrm{B}(n,p)$ に従うとは，X がとり得る値が n 以下の自然数 $0,1,2,3,\ldots,n$ であり，n 以下の各自然数 $k=0,1,2,3,\ldots,n$ に対して

$$\mathrm{P}(\,X=k\,) = {}_n\mathrm{C}_k\,p^k(1-p)^{n-k}$$

となることである.

2 項分布のことをベルヌイ分布ということもあります.

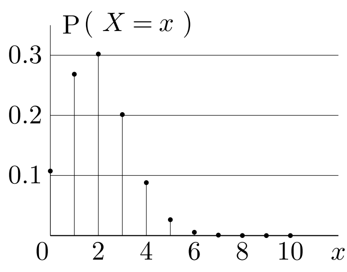
試行の反復について次の定理が成り立ちました.

試行 T において事象 A が起きる確率が p であるとする．自然数 n,k について $k\le n$ とする．試行 T を独立に n 回行うという試行において，事象 A が起きる回数が丁度 k になる確率は ${}_n\mathrm{C}_k\,p^k(1-p)^{n-k}$ である.

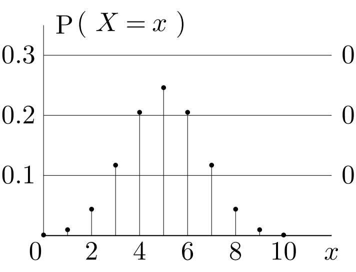
この定理より直ぐに次のことが分かります.

定理　定数 n は自然数で，定数 p は実数で $0\le p\le 1$ とする．試行 T において事象 A が起きる確率が p であるとする．この試行 T を独立に n 回行うとき，T の n 回の試行のうち事象 A が起きる回数を表す変数を X とおくと，X は確率変数であり，2 項分布 $\mathrm{B}(n,p)$ に従う.

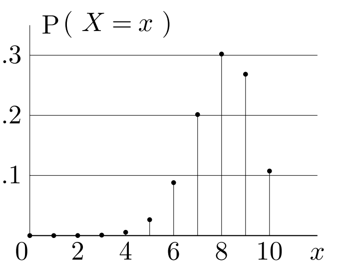
2 項分布に従う確率変数 X について，実数 x に対する確率 $\mathrm{P}(\,X=x\,)$ のグラフは次のようになります.



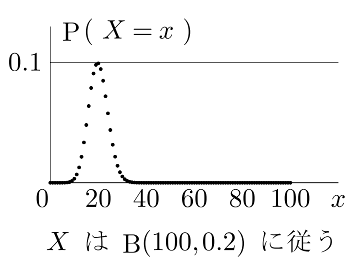
X は $\mathrm{B}(10,0.2)$ に従う



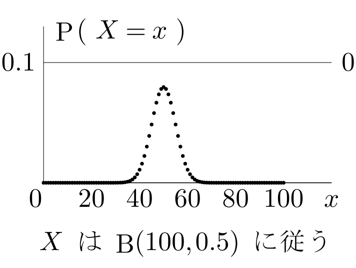
X は $\mathrm{B}(10,0.5)$ に従う



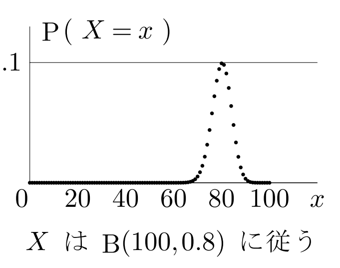
X は $\mathrm{B}(10,0.8)$ に従う



X は $\mathrm{B}(100,0.2)$ に従う



X は $\mathrm{B}(100,0.5)$ に従う



X は $\mathrm{B}(100,0.8)$ に従う

2 項分布の平均値と分散とは次のようになります．その証明は後にします.

定理 3.1　定数 n は自然数で，定数 p は実数で $0\le p\le 1$ とする．確率変数 X が 2 項分布 $\mathrm{B}(n,p)$ に従うならば，X は離散型確率変数であり，X の平均値 $\mathrm{E}[X]$ と分散 $\mathrm{V}[X]$ とは，

$$\mathrm{E}[X] = np \ , \qquad \mathrm{V}[X] = np(1-p) \ .$$

$\boxed{\text{例題 3.1}}$　試行 T において事象 A が起きる確率が $\frac{1}{3}$ であるとする．この試行を独立に 4 回行うとき，この 4 回の試行のうち事象 A が起きる回数を表す変数を X とおく．この確率変数 X の平均値 $\mathrm{E}[X]$ と分散 $\mathrm{V}[X]$ とを求める．また，確率 $\mathrm{P}\left(\,X\ge\frac{5}{2}\,\right)$ と $\mathrm{P}\left(\,1\le X\le\sqrt{11}\,\right)$ とを求める.

確率変数 X は 2 項分布 $\mathrm{B}\left(4,\frac{1}{3}\right)$ に従うので，

$$\mathrm{E}[X] = 4\cdot\frac{1}{3} = \frac{4}{3} \ , \qquad \mathrm{V}[X] = 4\cdot\frac{1}{3}\cdot\frac{2}{3} = \frac{8}{9} \ .$$

X が取りうる値 $k=0,1,2,3,4$ に対して $\mathrm{P}(\,X=k\,) = {}_4\mathrm{C}_k\left(\frac{1}{3}\right)^k\left(\frac{2}{3}\right)^{4-k}$ なので，

$$
\begin{aligned}
\mathrm{P}\left(\,X\ge\frac{5}{2}\,\right) &= \mathrm{P}(\,X=3\,)+\mathrm{P}(\,X=4\,) = {}_4\mathrm{C}_3\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^1 + {}_4\mathrm{C}_4\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^0 \\
&= 4\cdot\frac{2}{3^4}+\frac{1}{3^4} = \frac{8+1}{3^4} \\
&= \frac{1}{9} \ ,
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{P}\left(\,1\le X\le\sqrt{11}\,\right) &= \mathrm{P}(\,X=1\,)+\mathrm{P}(\,X=2\,)+\mathrm{P}(\,X=3\,) \\
&= {}_4\mathrm{C}_1\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^3 + {}_4\mathrm{C}_2\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^2 + {}_4\mathrm{C}_3\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^1 \\
&= 4\cdot\frac{2^3}{3^4}+6\cdot\frac{2^2}{3^4}+4\cdot\frac{2}{3^5} = \frac{32+24+8}{3^4} \\
&= \frac{64}{81} \ .
\end{aligned}
$$

$\boxed{\text{終}}$

$\boxed{\text{問題 3.1}}$　試行 T において事象 A が起きる確率が $\frac{2}{3}$ であるとします．この試行を独立に 5 回行うとき，この 5 回の試行のうち事象 A が起きる回数を表す変数を X とおきます．この確率変数 X の平均値 $\mathrm{E}[X]$ と分散 $\mathrm{V}[X]$ とを求めなさい．また，確率 $\mathrm{P}\left(\,X\le\frac{9}{4}\,\right)$ と $\mathrm{P}\left(\,\sqrt{7}\le X\le 4\,\right)$ とを求めなさい.

――――― 2 項分布に関する定理の証明

定理 3.1　定数 n は自然数で，定数 p は実数で $0\le p\le 1$ とする．確率変数 X が 2 項分布 $\mathrm{B}(n,p)$ に従うとき，$\mathrm{E}[X]=np$ ，$\mathrm{V}[X]=np(1-p)$ ．

証明　確率変数 X が 2 項分布 $\mathrm{B}(n,p)$ に従うとする．　$q=1-p$ とおく．　X が取り得る値は 0 以上 n 以下の自然数 $k=0,1,2,3,\ldots,n$ であり，$\mathrm{P}(\,X=k\,) = {}_n\mathrm{C}_k\,p^k(1-p)^{n-k} = \dfrac{n!}{k!(n-k)!}p^kq^{n-k}$ ．

$$
\begin{aligned}
\mathrm{E}[X] &= \sum_{k=0}^{n}\{k\,\mathrm{P}(\,X=k\,)\} = 0+\sum_{k=1}^{n}\{k\,\mathrm{P}(\,X=k\,)\} = \sum_{k=1}^{n}\{k\,\mathrm{P}(\,X=k\,)\} \\
&= \sum_{k=1}^{n}\left\{k\frac{n!}{k!(n-k)!}p^kq^{n-k}\right\} = \sum_{k=1}^{n}\left\{\frac{n!}{(k-1)!(n-k)!}p^kq^{n-k}\right\} \\
&= \sum_{k=0}^{n-1}\left\{\frac{n!}{k!\{n-(k+1)\}!}p^{k+1}q^{n-(k+1)}\right\} = \sum_{k=0}^{n-1}\left\{\frac{(n-1)!n}{k!(n-1-k)!}pp^kq^{n-1-k}\right\} \\
&= np\sum_{k=0}^{n-1}\left\{\frac{(n-1)!}{k!(n-1-k)!}p^kq^{n-1-k}\right\} \ .
\end{aligned}
$$

二項定理より，

$$\sum_{k=0}^{n-1}\left\{\frac{(n-1)!}{k!(n-1-k)!}p^kq^{n-1-k}\right\} = (p+q)^{n-1} = 1^{n-1} = 1 \ .$$

従って $\mathrm{E}[X]=np$ ．

$$
\begin{aligned}
\mathrm{E}[X(X-1)] &= \sum_{k=0}^{n}\{k(k-1)\,\mathrm{P}(\,X=k\,)\} = 0+0+\sum_{k=2}^{n}\{k(k-1)\,\mathrm{P}(\,X=k\,)\} \\
&= \sum_{k=2}^{n}\{k(k-1)\,\mathrm{P}(\,X=k\,)\} = \sum_{k=2}^{n}\left\{k(k-1)\frac{n!}{k!(n-k)!}p^kq^{n-k}\right\} \\
&= \sum_{k=2}^{n}\left\{\frac{n!}{(k-2)!(n-k)!}p^kq^{n-k}\right\} \\
&= \sum_{k=0}^{n-2}\left\{\frac{n!}{k!\{n-(k+2)\}!}p^{k+2}q^{n-(k+2)}\right\} \\
&= \sum_{k=0}^{n-2}\left\{\frac{(n-2)!(n-1)n}{k!(n-2-k)!}p^2p^kq^{n-2-k}\right\} \\
&= n(n-1)p^2\sum_{k=0}^{n-2}\left\{\frac{(n-2)!}{k!\{n-2-k\}!}p^kq^{n-2-k}\right\} \ .
\end{aligned}
$$

二項定理より，

$$\sum_{k=0}^{n-2}\left\{\frac{(n-2)!}{k!(n-2-k)!}p^kq^{n-2-k}\right\} = (p+q)^{n-2} = 1^{n-2} = 1 \ .$$

$$\mathrm{E}[X(X-1)] = n(n-1)p^2 \ .$$

$\mathrm{E}[X(X-1)] = \mathrm{E}[X^2-X] = \mathrm{E}[X^2]-\mathrm{E}[X] = \mathrm{E}[X^2]-np$ なので $\mathrm{E}[X^2]-np = n(n-1)p^2$ ，よって

$$\mathrm{E}[X^2] = n^2p^2-np^2+np \ .$$

定理 2.4.4 より

$$\mathrm{V}[X] = \mathrm{E}[X^2]-\mathrm{E}[X]^2 = n^2p^2-np^2+np-(np)^2 = -np^2+np = np(1-p) \ .$$

（証明終り）