統計的な推測【二項分布】　p.70～73

1 学習内容の説明 ⇒ 2 問題演習 ⇒ 3 振り返り（確認テスト・相互採点・リフレクションの記入）

【態度目標】しゃべる、質問する、説明する、動く、協力する、貢献する

【内容目標】二項分布の考え方を理解しよう

□二項分布

同じ条件のもとでの試行の繰り返しを反復試行という。ここでは，反復試行において，ある事象の起こる回数 X の確率分布について考えよう。

一般に，反復試行の確率について，次のことが成り立つ。

> 1 回の試行で事象 A の起こる確率を p とする。この試行を n 回行う反復試行において，A がちょうど r 回起こる確率は
>
> ${}_n\mathrm{C}_r p^r q^{n-r}$　　ただし，$q=1-p$

数学Aのおさらい

このような n 回の反復試行において，事象 A の起こる回数を X とすると，X は確率変数で，その確率分布は次の表のようになる。

X	0	1	……	r	……	n	計
P	${}_n\mathrm{C}_0 q^n$	${}_n\mathrm{C}_1 pq^{n-1}$	……	${}_n\mathrm{C}_r p^r q^{n-r}$	……	${}_n\mathrm{C}_n p^n$	1

それぞれの確率が，二項定理による $(q+p)^n$ の展開式の各項と同じなので，二項分布と呼ばれている。

$(q+p)^n=\sum_{r=0}^{n} {}_n\mathrm{C}_r p^r q^{n-r}$ である。

この表で与えられる確率分布を **二項分布** といい，

$B(n,\ p)$ で表す。

$B(n,\ p)$ の B は，二項分布を意味する英語 binomial distribution の頭文字である。

B（**何回行うか，確率 p の試行を**）

また，確率変数 X は二項分布 $B(n,\ p)$ に従うという。

離散的変量の確率分布で二項分布以外に重要なものとしてポアソン分布というものもある

例１５） 1 枚の硬貨を 10 回投げて，表の出る回数を X とする。

1 回だけ投げるとき，表の出る確率は $\frac{1}{2}$ であるから，X は二項分布 $B\left(10,\ \frac{1}{2}\right)$ に従う確率変数である。

$B\left(10\text{回行う，確率 } \frac{1}{2} \text{ の試行を}\right)$

また，たとえば $P(X=3)$ は，次のように計算される。

$$P(X=3)={}_{10}\mathrm{C}_3\left(\frac{1}{2}\right)^3\left(1-\frac{1}{2}\right)^7=\frac{120}{2^{10}}=\frac{15}{128}$$ 終

10回中 3 回表が出る反復試行の計算

□二項分布に従う確率変数の期待値と分散

二項分布に従う確率変数の期待値と分散を調べてみよう。

1 回の試行において，事象 A の起こる確率を p とし，この試行を 3 回行うとする。$k=1,\ 2,\ 3$ に対して，k 回目の試行で A が起これば 1，A が起こらなければ 0 の値をとる確率変数 X_k を考える。この反復試行において，A が起こる回数を X とすると，$X=X_1+X_2+X_3$ と表され，

X は二項分布 $B(3,\ p)$ に従う。

X	X_1	X_2	X_3
0	0	0	0
1	1	0	0
1	0	1	0
1	0	0	1
2	1	1	0
2	1	0	1
2	0	1	1
3	1	1	1

すべての k について，X_k の確率分布は，右の表のようになる。

ただし，$q=1-p$ である。

X_k	0	1	計
P	q	p	1

よって　$E(X_k)=0\cdot q+1\cdot p=p$

$E(X_k{}^2)=0^2\cdot q+1^2\cdot p=p$

$V(X_k)=E(X_k{}^2)-\{E(X_k)\}^2=p-p^2=p(1-p)=pq$

$X=X_1+X_2+X_3$ の期待値は，次のようになる。

$$E(X)=E(X_1)+E(X_2)+E(X_3)=3p$$

また，X_1，X_2，X_3 は互いに独立であるから，$X=X_1+X_2+X_3$ の分散は，次のようになる。

$$V(X)=V(X_1)+V(X_2)+V(X_3)=3pq$$

深める　上の確率変数 X の期待値と分散を，
それぞれ定義にもとづいて計算して求めてみよう。

$q=1-p$ とする。

$E(X)=0\cdot q^3+1\cdot 3pq^2+2\cdot 3p^2q+3\cdot p^3$

$=3p(q^2+2pq+p^2)=3p(p+q)^2=3p$

$p+q=1$ なので、上と同じになる

また，$V(x)=(0-3p)^2\cdot q^3+(1-3p)^2\cdot 3pq^2+(2-3p)^2\cdot 3p^2q+(3-3p)^2\cdot p^3$

$q=1-p$

$=9p^2q^3+(1-6p+9p^2)\cdot 3pq^2+(4-12p+9p^2)\cdot 3p^2q+\{3(1-p)\}^2\cdot p^3$

$=9p^2q^3+(1-6p+9p^2)\cdot 3pq^2+(4-12p+9p^2)\cdot 3p^2q+9q^2\cdot p^3$

$=3pq\{3pq^2+(1-6p+9p^2)\cdot q+(4-12+9p^2)\cdot p+3p^2q\}$

$=3pq\{9p^3+(12q-12)p^2+(3q^2-6q+4)p\}+q$

$=3pq\{9p^3+(12-12p-12)p^2+\{3(1-p)^2-6(1-p)+4\}\cdot p+1-p\}$

$q=1-p$

$=3pq\{9p^3-12p^2+(3p^2+1)\cdot p+1-p\}$

$=3pq$

n 回の反復試行では，事象 A の回数 X は，$X=X_1+X_2+X_3+\cdots\cdots+X_n$ になる。

また，X_1，X_2，……，X_n は独立であるから

$E(X)=E(X_1+X_2+X_3+\cdots\cdots+X_n)$

$=E(X_1)+E(X_2)+E(X_3)+\cdots\cdots+E(X_n)$

$=p+p+p+\cdots\cdots+p$

$=np$

$V(X)=V(X_1+X_2+X_3+\cdots\cdots+X_n)$

$i\neq j$ のとき，X_i と X_j は独立だから

$V(X)=V(X_1)+V(X_2)+V(X_3)+\cdots\cdots+V(X_n)$

$=pq+pq+pq+\cdots\cdots+pq$

$=npq$

一般に，次のことが成り立つ。

二項分布に従う確率変数の期待値，分散，標準偏差

確率変数 X が二項分布 $B(n,\ p)$ に従うとき

期待値は　$E(X)=np$

分散は　$V(X)=npq$　ただし，$q=1-p$

標準偏差は　$\sigma(X)=\sqrt{npq}$

統計的な推測【二項分布】　p.70〜73

例題2） 1個のさいころを90回投げて，2以下の目が出る回数を X とする。
X の期待値と分散および標準偏差を求めよ。

解答　さいころを1回投げて，2以下の目が出る確率 p は　$p=\frac{2}{6}=\frac{1}{3}$

よって，X は二項分布 $B\left(90,\ \frac{1}{3}\right)$ に従うから　X の期待値は　$E(X)=90\cdot\frac{1}{3}=30$

X の分散は　$V(X)=90\cdot\frac{1}{3}\cdot\left(1-\frac{1}{3}\right)=20$

X の標準偏差は　$\sigma(X)=\sqrt{V(X)}=\sqrt{20}=2\sqrt{5}$

コインを投げるときの「表」と「裏」のように、二種類の結果を取りうる試みがあったときに、このような試行のことを ベルヌーイ試行 と呼びます。二つの事象を $X=1$，$X=0$ とすると、$P(X=1)=p$，$P(X=0)=1-p$ であり、$P(X=k)=p^k(1-p)^{1-k}\ (k=0,1)$ とも表せます。これらの分布のことを ベルヌーイ分布 と呼びます。ベルヌーイ分布は二項分布の試行回数が $n=1$ の場合の特殊な例と考えられます。


ヤコブ・ベルヌーイ
Jakob Bernoulli
1654/12/27−1705/8/16

研究　二項分布のグラフ

1個のさいころを n 回投げるとき，
1の目が出る回数を X とすると，
X は二項分布 $B\left(n,\ \frac{1}{6}\right)$ に従う。

二項分布 $B\left(n,\ \frac{1}{6}\right)$ について

$$P_r=P(X=r)$$
$$={}_n\mathrm{C}_r\left(\frac{1}{6}\right)^r\left(\frac{5}{6}\right)^{n-r}$$

とおく。

$n=10,\ 20,\ 30,\ 50$

の各場合について，確率 P_r は右の表のようになり，
折れ線グラフをかくと下の図のようになる。
図から予想されるように，二項分布のグラフは，
n が大きくなるにつれて，左右対称なつりがね状の
曲線に近づく。

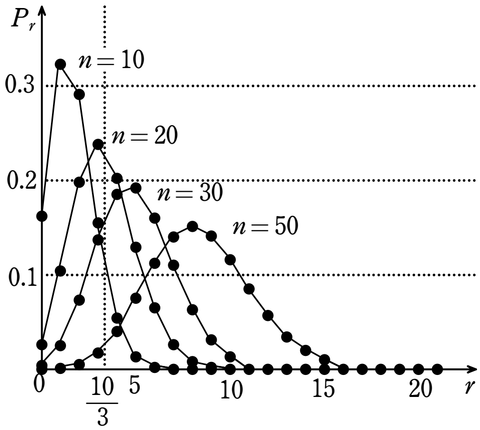
P_r \ n	10	20	30	50
P_0	0.162	0.026	0.004	0.000
P_1	0.323	0.104	0.025	0.001
P_2	0.291	0.198	0.073	0.005
P_3	0.155	0.238	0.137	0.017
P_4	0.054	0.202	0.185	0.040
P_5	0.013	0.129	0.192	0.075
P_6	0.002	0.065	0.160	0.112
P_7	0.000	0.026	0.110	0.140
P_8	：	0.008	0.063	0.151
P_9	：	0.002	0.031	0.141
P_{10}	：	0.000	0.013	0.116
：		：	：	：

なお，n が大きいときは，P_r の値が小さくなるので，P_r の目盛りの幅を大きくとると見やすくなる。

　この図から，たとえば $n=20$ の場合の 1 の目が出る回数について確率を考えてみると，期待値

$$20\cdot\frac{1}{6}=\frac{10}{3}$$

に近い回数ほど出る確率が大きいことがわかる。



一般に，二項分布 $B(n,\ p)$ において，確率 P_r を最大にする r の値を求めるには，次のようにする。

$$P_{r+1}={}_n\mathrm{C}_{r+1}\cdot p^{r+1}\cdot q^{n-r-1}\ ,\quad P_r={}_n\mathrm{C}_r\cdot p^r\cdot q^{n-r}$$

より　$$\frac{P_{r+1}}{P_r}=\frac{{}_n\mathrm{C}_{r+1}\cdot p^{r+1}\cdot q^{n-r-1}}{{}_n\mathrm{C}_r\cdot p^r\cdot q^{n-r}}=\frac{{}_n\mathrm{P}_{r+1}}{(r+1)!}\cdot\frac{r!}{{}_n\mathrm{P}_r}\cdot\frac{p}{q}=\frac{n-r}{r+1}\cdot\frac{p}{q}$$

$\dfrac{P_{r+1}}{P_r}>1$ とすると　$(n-r)p>(r+1)q$

よって　$np-q>(p+q)r$　$\therefore\ np-q>r$

$r<np-q$ のとき　$P_r<P_{r+1}$

$r>np-q$ のとき　$P_r>P_{r+1}$

したがって，n が大きいとき，P_r を最大にする r の値は np に近いことがわかる。

ド・モアブル－ラプラスの定理

確率変数 X_n の分布が二項定理 $B(n,\ p)$ であるとき X_n を標準化した確率変数を

$$Z_n=\frac{X_n-np}{\sqrt{npq}}$$

とおくと，n が十分に大きいときは，Z_n はほぼ（近似的に）標準正規分布に従う。

（→p.81 以降で扱っています）