二項分布の補足

経営統計Iの補足の補足資料

2019年12月16日
金沢学院大学経営情報学部
藤本祥二

階乗（factorial）

階段状に減っていく数値の積

$$n! = n \cdot (n-1) \cdot \cdots \cdot 2 \cdot 1$$

n個の物を並べ替える場合の数は$n!$通り

A─B─C─D
　　　└D─C
　└C─B─D
　　　└D─B
　└D─B─C
　　　└C─B

B─A─C─D
　　　└D─C
　└C─A─D
　　　└D─A
　└D─A─C
　　　└C─A

C─A─B─D
　　　└D─B
　└B─A─D
　　　└D─A
　└D─A─B
　　　└B─A

D─A─B─C
　　　└C─B
　└B─A─C
　　　└C─A
　└C─A─B
　　　└B─A

青で4つに分岐して
赤で3つに分岐して
緑で2つに分岐して
紫で1つに分岐する

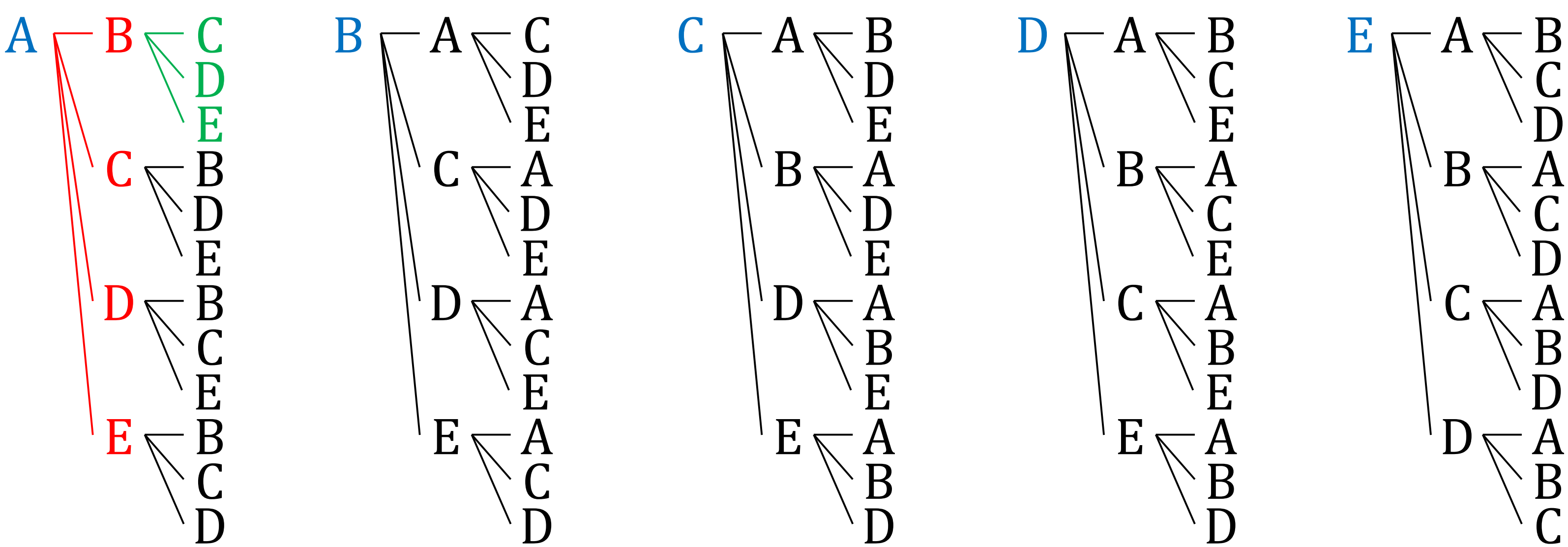
{A,B,C,D}の4つの物を並べ替えるやり方は24通り

$$4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$$

順列（permutation）

n個の物からr個取り出して並べ替える場合の数

$$
{}_nP_r = n \cdot (n-1) \cdot \cdots \cdot (n-r+1) = \frac{n!}{(n-r)!}
$$



青で5つに分岐して
赤で4つに分岐して
緑で3つに分岐する

{A,B,C,D,E}の5つの物から3つを取り出して
並べ替えるやり方は60通り

$$
{}_5P_3 = \frac{5!}{2!} = \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{2 \cdot 1} = 5 \cdot 4 \cdot 3 = 60
$$

$0! = 1$

n個の物からr個取り出して並べ替える場合の数

$${}_nP_r = n \cdot (n-1) \cdot \cdots \cdot (n-r+1) = \frac{n!}{(n-r)!}$$

$r = n$でこの2つ目の等式が成り立つためには
$0! = 1$としないと辻褄が合わなくなってしまう

$${}_nP_r = n \cdot (n-1) \cdot \cdots \cdot (n-r+1)$$

上式で　$r = n$　とすると

$${}_nP_n = n \cdot (n-1) \cdot \cdots \cdot (n-n+1) = n \cdot (n-1) \cdot \cdots \cdot 1 = n!$$

$${}_nP_r = \frac{n!}{(n-r)!}$$

上式で　$r = n$　とすると

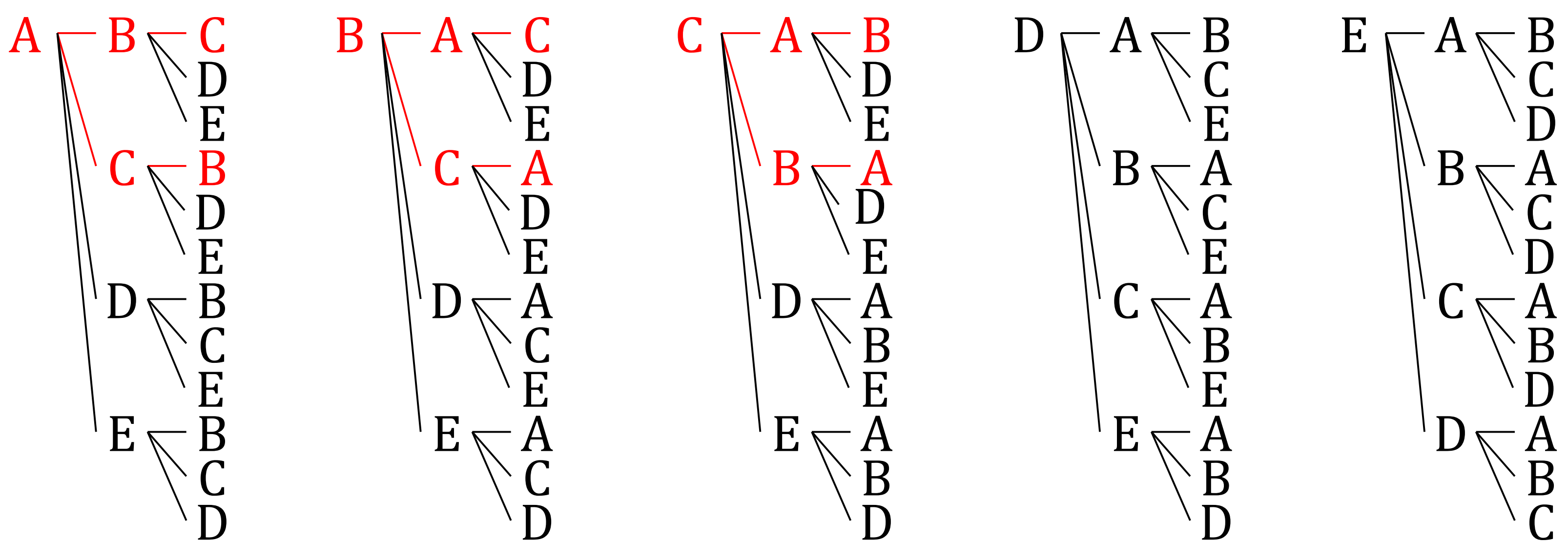
$${}_nP_n = \frac{n!}{0!}$$

$0! = 1$と定義することで矛盾がなくなる

組み合わせ（combination）

n個の物からr個取り出す組み合わせの場合の数

$$
{}_{n}C_{r} = \frac{{}_{n}P_{r}}{r!} = \frac{n!}{(n-r)!\,r!}
$$



順列において順番が違うだけの場合を同一視
3個を取り出す場合は赤い3!通りのパターンを同一視
ABC,ACB,BAC,BCA,CAB,CBAをすべてABCと同じとみなす
{ABC,ABD,ABE,ACD,ACE,ADE,BCD,BCE,BDE,CDE}
の10通り

{A,B,C,D,E}の5つの物から3つを取り出す
組み合わせは10通り

$$
{}_{5}C_{3} = \frac{{}_{5}P_{3}}{3!} = \frac{5!}{2!\,3!} = \frac{60}{6} = 10
$$

{青,赤,緑,紫}の4個からx個取り出す組み合わせの数を考える

- 0個取り出す組み合わせの数
  {}　何も取り出さないという1通り
$$ {}_4C_0 = \frac{4!}{(4-4)!\,4!} = \frac{4!}{0!\,4!} = \frac{4\cdot3\cdot2\cdot1}{1\cdot4\cdot3\cdot2\cdot1} = 1 $$
- 1個取り出す組み合わせの数
  {青，赤，緑，紫}　の4通り
$$ {}_4C_1 = \frac{4!}{(4-1)!\,1!} = \frac{4!}{3!\,1!} = \frac{4\cdot3\cdot2\cdot1}{3\cdot2\cdot1\cdot1} = 4 $$
- 2個取り出す組み合わせの数
  {青赤，青緑，青紫，赤緑，赤紫，緑紫}　の6通り
$$ {}_4C_2 = \frac{4!}{(4-2)!\,2!} = \frac{4!}{2!\,2!} = \frac{4\cdot3\cdot2\cdot1}{2\cdot1\cdot2\cdot1} = 6 $$
- 3個取り出す組み合わせの数
  {青赤緑，青赤紫，青緑紫，赤緑紫}　の4通り
$$ {}_4C_3 = \frac{4!}{(4-3)!\,3!} = \frac{4!}{1!\,3!} = \frac{4\cdot3\cdot2\cdot1}{1\cdot3\cdot2\cdot1} = 4 $$
- 4個取り出す組み合わせの数
  {青赤緑紫}　の1通り
$$ {}_4C_4 = \frac{4!}{(4-4)!\,4!} = \frac{4!}{0!\,4!} = \frac{4\cdot3\cdot2\cdot1}{1\cdot4\cdot3\cdot2\cdot1} = 1 $$

- {青,赤,緑,紫}の4個から2個取り出す

青赤①, 赤青①, 緑青②, 紫青③,
青緑②, 赤緑④, 緑赤④, 紫赤⑤,
青紫③, 赤紫⑤, 緑紫⑥, 紫緑⑥

順列だと12通り

$${}_4P_2 = 12$$

青赤↔赤青の①のように順番変えると同じになるものを同一視すると6通り

$${}_4C_2 = 6$$

- {青,赤,緑,紫}の4個から3個取り出す

青赤緑①, 赤青緑①, 緑青赤①, 紫青赤②,
青赤紫②, 赤青紫②, 緑青紫③, 紫青緑③,
青緑赤①, 赤緑青①, 緑赤青①, 紫赤青②,
青緑紫③, 赤緑紫④, 緑赤紫④, 紫赤緑④,
青紫赤②, 赤紫青②, 緑紫青③, 紫緑青③,
青紫緑③, 赤紫緑④, 緑紫赤④, 紫緑赤④

順列だと24通り

$${}_4P_3 = 24$$

順番変えると
同じになるものを同一視すると4通り

$${}_4C_3 = 4$$

- {青,赤,緑,紫}の4個から3個取り出す

青赤緑紫, 赤青緑紫, 緑青赤紫, 紫青赤緑,
青赤紫緑, 赤青紫緑, 緑青紫赤, 紫青緑赤,
青緑赤紫, 赤緑青紫, 緑赤青紫, 紫赤青緑,
青緑紫赤, 赤緑紫青, 緑赤紫青, 紫赤緑青,
青紫赤緑, 赤紫青緑, 緑紫青赤, 紫緑青赤,
青紫緑赤, 赤紫緑青, 緑紫赤青, 紫緑赤青

順列だと24通り

$${}_4P_4 = 24$$

順番変えると
同じになるものを同一視すると1通り

$${}_4C_4 = 1$$

組み合わせの数と二項係数

$(q+p)^2 = (q+p)(q+p) = q^2 + 2qp + p^2$

組み合わせに注目して
この計算を行う

$= qq$

青赤の2個の括弧から青を選んでpを取り出した

青赤の2個の括弧から赤を選んでpを取り出した

$+pq + qp$

$+pp$

青赤の2個の括弧から青赤2つを選んでppを取り出した

$= 1q^{2-0}p^0 + 2q^{2-1}p^1 + 1q^{2-2}p^2$

2個の括弧から0個の括弧を選ぶ組み合わせの数 ${}_2C_0 = 1$

2個の括弧から1個の括弧を選ぶ組み合わせの数 ${}_2C_1 = 2$

2個の括弧から2個の括弧を選ぶ組み合わせの数 ${}_2C_2 = 1$

$$(q+p)^3=(q+p)(q+p)(q+p)\boxed{=q^3+3q^2p+3qp^2+p^3}$$

$$=qqq$$

青赤緑の3個の括弧から青を選んでpを取り出した

青赤緑の3個の括弧から赤を選んでpを取り出した

青赤緑の3個の括弧から緑を選んでpを取り出した

$$+pqq+qpq+qqp$$

青赤緑の3個の括弧から青赤を選んでppを取り出した

青赤緑の3個の括弧から青緑を選んでppを取り出した

青赤緑の3個の括弧から赤緑を選んでppを取り出した

$$+ppq+pqp+qpp$$

$$+ppp$$

青赤緑の3個の括弧から青赤緑を選んでpppを取り出した

$$=1q^{3-0}p^0+3q^{3-1}p^1+3q^{3-2}p^2+1q^{3-3}p^3$$

3個の括弧から0個の括弧を選ぶ組み合わせの数 ${}_3C_0=1$

3個の括弧から1個の括弧を選ぶ組み合わせの数 ${}_3C_1=3$

3個の括弧から2個の括弧を選ぶ組み合わせの数 ${}_3C_2=3$

3個の括弧から3個の括弧を選ぶ組み合わせの数 ${}_3C_3=1$

$$(q+p)^4 = (q+p)(q+p)(q+p)(q+p)$$

$$\boxed{= q^4 + 4q^3p + 6q^2p^2 + 4qp^3 + p^4}$$

$$
\begin{aligned}
&= qqqq \\
&\quad + pqqq + qpqq + qqpq + qqqp \\
&\quad + ppqq + pqpq + pqqp + qppq + qpqp + qqpp \\
&\quad + pppq + ppqp + pqpp + qppp \\
&\quad + pppp
\end{aligned}
$$

$$= 1q^{4-0}p^0 + 4q^{4-1}p^1 + 6q^{4-2}p^2 + 4q^{4-3}p^3 + 1q^{4-4}p^4$$

4個の括弧から0個の括弧を選ぶ組み合わせの数 ${}_4C_0 = 1$

4個の括弧から1個の括弧を選ぶ組み合わせの数 ${}_4C_1 = 4$

4個の括弧から2個の括弧を選ぶ組み合わせの数 ${}_4C_2 = 6$

4個の括弧から3個の括弧を選ぶ組み合わせの数 ${}_4C_3 = 4$

4個の括弧から4個の括弧を選ぶ組み合わせの数 ${}_4C_4 = 1$

一般の二項展開

$$(q+p)^n = \sum_{x=0}^{n} {}_nC_x q^{n-x} p^x$$

$$= {}_nC_0 q^{n-0} p^0 + {}_nC_1 q^{n-1} p^1 + \cdots + {}_nC_x\, q^{n-x} p^x + \cdots + {}_nC_{n-1} q^1 p^{n-1} + {}_nC_n q^0 p^n$$

n個の括弧から0個の括弧を選ぶ組み合わせの数 ${}_nC_0 = 1$

n個の括弧から1個の括弧を選ぶ組み合わせの数 ${}_nC_1 = n$

n個の括弧からx個の括弧を選ぶ組み合わせの数 ${}_nC_x = \frac{n!}{(n-x)!x!}$

n個の括弧から$n-1$個の括弧を選ぶ組み合わせの数 ${}_nC_{n-1} = n$

n個の括弧からn個の括弧を選ぶ組み合わせの数 ${}_nC_n = 1$

$${}_nC_x = \frac{n!}{(n-x)!\,x!}$$

は二項係数とも呼ばれる

$$(q+p)^n = q^n + nq^{n-1}p + \cdots + \frac{n!}{(n-x)!\,x!} q^{n-x} p^x + \cdots + nqp^{n-1} + p^n$$

二項係数 ${}_nC_x$ の性質

- パスカルの三角形の左右対称性

$$ {}_nC_x = {}_nC_{n-x} $$

- パスカルの三角形の端の値は1

$$ {}_nC_0 = 1, \qquad {}_nC_n = 1 $$

- パスカルの三角形の足し上げの原理

$$ {}_nC_x + {}_nC_{x-1} = {}_{n+1}C_x $$

二項係数 ${}_nC_x$ の性質の証明

- ${}_nC_x = {}_nC_{n-x}$ の証明

$$ {}_nC_x = \frac{n!}{(n-x)!\,x!} = \frac{n!}{x!\,(n-x)!} = {}_nC_{n-x} $$

- ${}_nC_0 = 1$，${}_nC_n = 1$ の証明

$$ {}_nC_0 = \frac{n!}{(n-0)!\,0!} = \frac{n!}{n!\,0!} = 1, \qquad {}_nC_n = \frac{n!}{(n-n)!\,n!} = \frac{n!}{0!\,n!} = 1 $$

- ${}_nC_x + {}_nC_{x-1} = {}_{n+1}C_x$ の証明

$$ \begin{aligned} {}_nC_x + {}_nC_{x-1} &= \frac{n!}{(n-x)!\,x!} + \frac{n!}{(n-x+1)!\,(x-1)!} \\ &= \frac{n!}{(n-x)!\,x(x-1)!} + \frac{n!}{(n-x+1)(n-x)!\,x!} \\ &= \frac{n!}{(n-x)!\,(x-1)!}\left\{\frac{1}{x} + \frac{1}{n-x+1}\right\} \\ &= \frac{n!}{(n-x)!\,(x-1)!}\left\{\frac{n+1}{x(n-x+1)}\right\} \\ &= \frac{(n+1)n!}{(n-x+1)(n-x)!\,x(x-1)!} = \frac{(n+1)!}{(n+1-x)!\,x!} = {}_{n+1}C_x \end{aligned} $$

パスカルの三角形と二項係数

$${}_{n}C_{x-1} + {}_{n}C_{x} = {}_{n+1}C_{x}$$

この式を次の図に置き換える

$${}_{n}C_{x-1} \quad + \quad {}_{n}C_{x}$$

$${}_{n+1}C_{x}$$

足し上げの原理

$$1\,{}_{0}C_{0}$$

$$1\,{}_{1}C_{0} \quad 1\,{}_{1}C_{1}$$

$$1\,{}_{2}C_{0} \quad 2\,{}_{2}C_{1} \quad 1\,{}_{2}C_{2}$$

$$1\,{}_{3}C_{0} \quad 3\,{}_{3}C_{1} \quad 3\,{}_{3}C_{2} \quad 1\,{}_{3}C_{3}$$

$$1\,{}_{4}C_{0} \quad 4\,{}_{4}C_{1} \quad 6\,{}_{4}C_{2} \quad 4\,{}_{4}C_{3} \quad 1\,{}_{4}C_{4}$$

$$1\,{}_{5}C_{0} \quad 5\,{}_{5}C_{1} \quad 10\,{}_{5}C_{2} \quad 10\,{}_{5}C_{3} \quad 5\,{}_{5}C_{4} \quad 1\,{}_{5}C_{5}$$

$$1\,{}_{6}C_{0} \quad 6\,{}_{6}C_{1} \quad 15\,{}_{6}C_{2} \quad 20\,{}_{6}C_{3} \quad 15\,{}_{6}C_{4} \quad 6\,{}_{6}C_{5} \quad 1\,{}_{6}C_{6}$$

$${}_{n}C_{0} = {}_{n}C_{n} = 1$$

端っこは1

$${}_{n}C_{x-1} = {}_{n}C_{x}$$

左右対称

$n = 3$の二項分布の平均と分散

事象	X	$P(X)$	$XP(X)$	$(X-\mu)^2$	$(X-\mu)^2 P(X)$
$\{\omega_{\mathrm{FFF}}\}$	0	q^3	0	$(3p)^2$	$9p^2q^3$
$\{\omega_{\mathrm{FFS}}, \omega_{\mathrm{FSF}}, \omega_{\mathrm{SFF}}\}$	1	$3q^2p$	$3q^2p$	$(1-3p)^2$	$3(1-3p)^2q^2p$
$\{\omega_{\mathrm{FSS}}, \omega_{\mathrm{SFS}}, \omega_{\mathrm{SSF}}\}$	2	$3qp^2$	$6qp^2$	$(2-3p)^2$	$3(2-3p)^2qp^2$
$\{\omega_{\mathrm{SSS}}\}$	3	p^3	$3p^3$	$(3-3p)^2$	$9(1-p)^2p^3$
計	—	1	$3p$	—	$3pq$

- 全確率の合計は1

$$\sum_{X=0}^{3} P(X) = q^3 + 3q^2p + 3qp^2 + q^3 = (q+p)^3 = 1^3 = 1$$

- 平均μは$3p$

$$\mu = \sum_{X=0}^{3} XP(X) = 3p$$

- 分散σ^2は$3pq$

$$\sigma^2 = \sum_{X=0}^{3} (X-\mu)^2 P(X) = 3pq$$

平均と分散の計算の詳細を示す

- 全確率の合計は1

$$\sum P(X) = q^3 + 3q^2p + 3qp^2 + p^3 = (q+p)^3 = 1$$

- 平均μは$3p$

$$\mu = \sum XP(X) = 0 + 3q^2p + 6qp^2 + 3q^3 = 3p(q^2 + 2qp + p^2)$$

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

- 分散σ^2は$3pq$

$$\sigma^2 = \sum (X-\mu)^2 P(X)$$

$$\begin{aligned}
&= 9p^2q^3 + 3(1-3p)^2q^2p + 3(2-3p)^2qp^2 + 9(1-p)^2p^3 \\
&= 3\{3p^2q^3 + (1-p-2p)^2q^2p + (2-2p-p)^2qp^2 + 3(1-p)^2p^3\} \\
&= 3\{3p^2q^3 + (q-2p)^2q^2p + (2q-p)^2qp^2 + 3q^2p^3\} \\
&= 3pq\{3pq^2 + (q^2-4pq+4p^2)q + (4q^2-4pq+p^2)p + 3qp^2\} \\
&= 3pq(3pq^2 + q^3 - 4q^2p + 4qp^2 + 4q^2p - 4p^2q + p^3 + 3qp^2) \\
&= 3pq(q^3 + 3pq^2 + 3qp^2 + p^3) \\
&= 3pq(q+p)^3 \\
&= 3qp
\end{aligned}$$

$q + p = 1$が使えるように変形して式を簡単にしていく

一般の二項分布

- 成功確率 p，失敗確率 $q=1-p$ の独立試行をn回反復した試行
- n回中で何回成功したかを確率変数Xにした時の確率分布が二項分布
- 二項分布の確率分布は二項展開後の各項を計算すれば分かる

$$(q+p)^n = q^n + nq^{n-1}p + \cdots + {}_nC_x\, q^{n-x}p^x + \cdots + nqp^{n-1} + p^n$$

失敗n，成功0，の確率 $P(X=0)$	失敗$n-1$，成功1，の確率 $P(X=1)$	失敗$n-x$，成功x，の確率 $P(X=x)$	失敗1，成功$n-1$，の確率 $P(X=n-1)$	失敗0，成功n，の確率 $P(X=n)$

- n回の反復試行でX回成功する確率$P(X)$

$$P(X) = {}_nC_X q^{n-X}p^X$$

${}_nC_X$はn個の物から
X個を取り出す組み合わせの数
パスカルの三角形を使えば
比較的簡単に計算できる

$${}_nC_X = \frac{n!}{(n-X)!\,X!}$$

「!」は階乗記号
（階段状に小さくなる数の乗積）

$0! = 1$
$1! = 1$
$2! = 2\cdot 1 = 2$
$3! = 3\cdot 2\cdot 1 = 6$
$4! = 4\cdot 3\cdot 2\cdot 1 = 24$
$5! = 5\cdot 4\cdot 3\cdot 2\cdot 1 = 120$

- 二項分布の期待値μと分散σ^2の公式

$$\mu = np, \qquad \sigma^2 = npq$$

一般の二項分布の平均と分散の証明

$$\sum_{X=0}^{n} P(X) = \sum_{X=0}^{n} {}_nC_X q^{n-X} p^X = (q+p)^n = 1^n = 1$$

二項分布の全確率の合計は1

$\dfrac{X}{X!} = \dfrac{1}{(X-1)!}$
を使う

$$\mu = \sum_{X=0}^{n} XP(X) = \sum_{X=1}^{n} XP(X) = \sum_{X=1}^{n} X\, {}_nC_X q^{n-X} p^X = \sum_{X=1}^{n} X \frac{n!}{(n-X)!\,X!} q^{n-X} p^X$$

$$= \sum_{X=1}^{n} \frac{n!}{(n-X)!\,(X-1)!} q^{n-X} p^X$$

$\cdots(*)$とする，次ページで使う

$$= \sum_{X=1}^{n} \frac{n\cdot(n-1)!}{\{(n-1)-(X-1)\}!\,(X-1)!} q^{\{(n-1)-(n-X)\}} p \cdot p^{X-1}$$

$X-1=Y$
とする

$$= np \sum_{Y=0}^{n-1} \frac{(n-1)!}{\{(n-1)-Y\}!\,Y!} q^{(n-1)-Y} p^Y = np \sum_{Y=0}^{n-1} {}_{n-1}C_Y q^{(n-1)-Y} p^Y$$

$$= np(q+p)^{n-1} = np$$

二項分布の平均（期待値）はnp

$$
\begin{aligned}
\sum_{X=0}^{n} X^2 P(X) &= \sum_{X=1}^{n} X^2 P(X) = \sum_{X=1}^{n} X^2 {}_nC_X q^{n-X} p^X = \sum_{X=1}^{n} X^2 \frac{n!}{(n-X)!\,X!} q^{n-X} p^X \\
&= \sum_{X=1}^{n} X \frac{n!}{(n-X)!\,(X-1)!} q^{n-X} p^X = \sum_{X=1}^{n} (X-1+1) \frac{n!}{(n-X)!\,(X-1)!} q^{n-X} p^X \\
&= \sum_{X=1}^{n} (X-1) \frac{n!}{(n-X)!\,(X-1)!} q^{n-X} p^X + \sum_{X=1}^{n} \frac{n!}{(n-X)!\,(X-1)!} q^{n-X} p^X \\
&= \sum_{X=1}^{n} (X-1) \frac{n \cdot (n-1)!}{\{(n-1)-(X-1)\}!\,(X-1)!} q^{\{(n-1)-(X-1)\}} p \cdot p^{X-1} + np \\
&= np \sum_{Y=0}^{n-1} Y \frac{(n-1)!}{\{(n-1)-Y\}!\,Y!} q^{(n-1)-Y} p^Y + np \\
&= np(n-1)p + np
\end{aligned}
$$

前ページ(*)式と同じ

$n \to n-1$
としたときの二項分布の平均
$(n-1)p$

$$
\begin{aligned}
\sigma^2 &= \sum_{X=0}^{n} X^2 P(X) - \mu^2 \\
&= np(n-1)p + np - (np)^2 \\
&= n^2p^2 - np^2 + np - n^2p^2 = np - np^2 = np(1-p) \\
&= npq
\end{aligned}
$$

二項分布の分散はnpq