数学Ｂ　統計的な推測「確率分布」

~高校数学のまとめ~

確率変数

試行の結果によってその値が定まる変数を **確率変数**（かくりつへんすう）という.

確率変数と確率の表記

確率変数は X, Y, Z などの大文字で表すことが多い.

$X = a$ の確率は $P(X = a)$

$a \leqq X \leqq b$ の確率は $P(a \leqq X \leqq b)$

のように表す.

また 確率を P とかくことがある.

例 さいころを 1 回振る試行によって，出た目を X とする.

$$P(X = 1) = \frac{1}{6},\ P(1 \leqq X \leqq 3) = \frac{3}{6} = \frac{1}{2}$$

確率分布

確率変数 X のとり得る値が x_1，x_2，$\cdots$，x_n であり，

それぞれの値をとる確率が p_1，p_2，$\cdots$，p_n とする．

すなわち $P(X=x_k)=p_k\ (k=1, 2, \cdots, n)$ とすると，次のことが成り立つ．

1. $\displaystyle\sum_{k=1}^{n} p_k = p_1+p_2+\cdots+p_n=1$
2. $p_1 \geqq 0$，$p_2 \geqq 0$，$\cdots$，$p_n \geqq 0$

また 確率変数 X と確率の対応は下の表になる．

X	x_1	x_2	$\cdots$	x_n	計
P	p_1	p_2	$\cdots$	p_n	1

このように確率変数 X のとり得る値と，その値をとる確率との対応を示したものを

その確率変数の **確率分布**（かくりつぶんぷ） または 単に **分布**（ぶんぷ） といい

確率変数 X はこの分布に従（したが）うという．

例 さいころを1回振る試行で出る目を X とすると，その確率分布は

X	1	2	3	4	5	6	計
P	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	1

確率変数の期待値（平均）

確率変数 X の各々の値と確率をかけて，すべてたした値を

確率変数 X の 期待値（きたいち） または 平均（へいきん） といい $E(X)$ または m などと表す．

すなわち 確率変数 X の確率分布を次として

X	x_1	x_2	$\cdots$	x_n	計
P	p_1	p_2	$\cdots$	p_n	1

$$E(X)=\sum_{k=1}^{n} x_k p_k = x_1p_1+x_2p_2+\cdots+x_np_n$$

(補) 期待値を意味する英語は Expected value，平均を意味する英語は mean

(補)「データの分析」（数学 Ⅰ）の平均値は $\overline{X}$ と表す．

(例) あるテストで 50 点が 2 人，80 点が 1 人のとき

平均点は $\dfrac{50\times2+80}{3}=60$（点）

これは得点 X の確率分布として右になるから

$E(X)=50\cdot\dfrac{2}{3}+80\cdot\dfrac{1}{3}=60$

X	50	80	計
P	$\dfrac{2}{3}$	$\dfrac{1}{3}$	1

1 次式の確率変数の期待値（平均）

確率変数 X の期待値を $E(X)$ とする．

a，b を定数，$a\neq0$ として 1 次式の確率変数 $aX+b$ の期待値は

$$E(aX+b)=aE(X)+b$$

(考) 確率変数 X の確率分布を次とする．

X	x_1	x_2	$\cdots$	x_n	計
$aX+b$	ax_1+b	ax_2+b	$\cdots$	ax_n+b	
P	p_1	p_2	$\cdots$	p_n	1

$E(X)=\sum\limits_{k=1}^{n} x_kp_k$，$\sum\limits_{k=1}^{n} p_k=1$ に注意して

$$E(aX+b)=\sum_{k=1}^{n}(ax_k+b)p_k=a\sum_{k=1}^{n}x_kp_k+b\sum_{k=1}^{n}p_k=aE(X)+b$$

偏差

確率変数 X の平均を m とするとき

確率変数 $X-m$ を X の平均からの 偏差（へんさ） という.

㊓ 偏差の平均は０になる．つまり $E(X-m)=0$

確率変数の分散

確率変数 X の平均を m とする.

確率変数 X の偏差の２乗 $(X-m)^2$ の期待値を

確率変数 X の 分散（ぶんさん） といい $V(X)$ などと表す.

すなわち 確率変数 X の確率分布を次として

X	x_1	x_2	$\cdots$	x_n	計
$(X-m)^2$	$(x_1-m)^2$	$(x_2-m)^2$	$\cdots$	$(x_n-m)^2$	
P	p_1	p_2	$\cdots$	p_n	1

$$V(X)=E\big((X-m)^2\big)=\sum_{k=1}^{n}(x_k-m)^2p_k$$
$$=(x_1-m)^2p_1+(x_2-m)^2p_2+\cdots+(x_n-m)^2p_n$$

分散は 平均からの散らばりの具合いを表す数値であり,

確率変数の値が平均から離れるほど大きな値をとる.

㊓ 分散を意味する英語は Variance

㊓「データの分析」（数学 Ｉ）の分散は $s_X{}^2$ と表す.

確率変数の標準偏差

確率変数 X の分散の正の平方根 を

確率変数 X の 標準偏差（ひょうじゅんへんさ） といい σ（シグマ）(X) または $s(X)$ などと表す.

すなわち 確率変数 X の分散を $V(X)$ とすると

$$\sigma(X)=\sqrt{V(X)} \quad \text{または} \quad s(X)=\sqrt{V(X)}$$

㊓ 標準偏差を意味する英語は standard deviation

㊓ σ はギリシャ文字

分散の計算

確率変数 X の期待値を $E(X)$，分散を $V(X)$ とすると

$$V(X)=E(X^2)-\{E(X)\}^2$$

つまり $(X$ の分散$)=(X^2$の期待値$)-(X$ の期待値$)^2$

㊟ 確率変数 X の確率分布を次とする．

X	x_1	x_2	$\cdots$	x_n	計
X^2	${x_1}^2$	${x_2}^2$	$\cdots$	${x_n}^2$	
P	p_1	p_2	$\cdots$	p_n	1

$E(X)=\sum\limits_{k=1}^{n}x_kp_k$，$E(X^2)=\sum\limits_{k=1}^{n}{x_k}^2p_k$，$\sum\limits_{k=1}^{n}p_k=1$ に注意する．

$m=E(X)$ として

$$\begin{aligned}V(X)&=E\big((X-m)^2\big)\\&=\sum_{k=1}^{n}(x_k-m)^2p_k\\&=\sum_{k=1}^{n}\left({x_k}^2-2mx_k+m^2\right)p_k\\&=\sum_{k=1}^{n}{x_k}^2p_k-2m\sum_{k=1}^{n}x_kp_k+m^2\sum_{k=1}^{n}p_k\\&=E(X^2)-2m\cdot m+m^2\cdot 1\\&=E(X^2)-m^2\\&=E(X^2)-\{E(X)\}^2\end{aligned}$$

1次式の確率変数の分散・標準偏差

確率変数 X の分散を $V(X)$，標準偏差を $\sigma(X)$ とする．

a，b を定数，$a \neq 0$ として

$\boxed{1}$ 1次式の確率変数 $aX+b$ の分散は $V(aX+b)=a^2V(X)$

$\boxed{2}$ 1次式の確率変数 $aX+b$ の標準偏差は $\sigma(aX+b)=|a|\sigma(X)$

考 確率変数 X の平均を m とする．

1次式の確率変数 $aX+b$ の偏差は

$$aX+b-E(aX+b)=aX+b-\{aE(X)+b\}=a\{X-E(X)\}=a(X-m)$$

$\boxed{1}$ $V(aX+b)=E\big(a^2(X-m)^2\big)=a^2E\big((X-m)^2\big)=a^2V(X)$

$\boxed{2}$ $\sigma(aX+b)=\sqrt{V(aX+b)}=\sqrt{a^2V(X)}=|a|\sqrt{V(X)}=|a|\sigma(X)$

2個以上の確率変数と確率の表記

確率変数 X, Y, Z について

$X=a$ かつ $Y=b$ の確率は $P(X=a,\, Y=b)$

$X=a$ かつ $Y=b$ かつ $Z=c$ の確率は $P(X=a,\, Y=b,\, Z=c)$

のように表す.

同時分布

2個の確率変数 X, Y について

X のとる値が $x_1,\, x_2,\, \cdots,\, x_m$

Y のとる値が $y_1,\, y_2,\, \cdots,\, y_n$

とし $P(X=x_k,\, Y=y_\ell)=p_{k\ell}$ とおくと対応は下の表になる.

X \ Y	y_1	y_2	$\cdots\cdots$	y_n	計
x_1	p_{11}	p_{12}	$\cdots\cdots$	p_{1n}	p_1
x_2	p_{21}	p_{22}	$\cdots\cdots$	p_{2n}	p_2
$\vdots$		$\cdots\cdots\cdots\cdots$			$\vdots$
$\vdots$		$\cdots\cdots\cdots\cdots$			$\vdots$
x_m	p_{m1}	p_{m2}	$\cdots\cdots$	p_{mn}	p_m
計	q_1	q_2	$\cdots\cdots$	q_n	1

すべての k と ℓ の mn 個の組み合せについての $(x_k,\, y_\ell)$ と $p_{k\ell}$ の対応を

X と Y の **同時分布**（どうじぶんぷ） という.

上の表から

各 k について $P(X=x_k)=\displaystyle\sum_{\ell=1}^{n} p_{k\ell}=p_{k1}+p_{k2}+\cdots+p_{kn}=p_k$

各 ℓ について $P(Y=y_\ell)=\displaystyle\sum_{k=1}^{m} p_{k\ell}=p_{1\ell}+p_{2\ell}+\cdots+p_{m\ell}=q_\ell$

X と Y はそれぞれ次の分布に従う.

X	x_1	x_2	$\cdots$	x_m	計
P	p_1	p_2	$\cdots$	p_m	1

Y	y_1	y_2	$\cdots$	y_n	計
P	q_1	q_2	$\cdots$	q_n	1

2 個の確率変数の和の期待値（平均）

確率変数 X，Y の期待値をそれぞれ $E(X)$，$E(Y)$ とする．

確率変数の和 $X+Y$ の期待値は

$$E(X+Y)=E(X)+E(Y)$$

㊀ 2 つの確率変数 X，Y の確率分布を次とする．

X	x_1	x_2	計
P	p_1	p_2	1

Y	y_1	y_2	y_3	計
P	q_1	q_2	q_3	1

X＼Y	y_1	y_2	y_3	計
x_1	p_{11}	p_{12}	p_{13}	p_1
x_2	p_{21}	p_{22}	p_{23}	p_2
計	q_1	q_2	q_3	1

右のように $P(X=x_k,\ Y=y_\ell)=p_{k\ell}$ とすると

$X+Y$ の確率分布は次になる．

$X+Y$	x_1+y_1	x_1+y_2	x_1+y_3	x_2+y_1	x_2+y_2	x_2+y_3	計
P	p_{11}	p_{12}	p_{13}	p_{21}	p_{22}	p_{23}	1

$$\begin{aligned}
E(X+Y)=&\ (x_1+y_1)p_{11}+(x_1+y_2)p_{12}+(x_1+y_3)p_{13}\\
&+(x_2+y_1)p_{21}+(x_2+y_2)p_{22}+(x_2+y_3)p_{23}\\
=&\ x_1(p_{11}+p_{12}+p_{13})+x_2(p_{21}+p_{22}+p_{23})\\
&+y_1(p_{11}+p_{21})+y_2(p_{12}+p_{22})+y_3(p_{13}+p_{23})\\
=&\ x_1p_1+x_2p_2+y_1q_1+y_2q_2+y_3q_3\\
=&\ E(X)+E(Y)
\end{aligned}$$

㊟ X，Y の取り得る値がいくつになっても同じ計算で成り立つ．

3個の確率変数の和の期待値（平均）

確率変数 X，Y，Z の期待値をそれぞれ $E(X)$，$E(Y)$，$E(Z)$ とする．

確率変数の和 $X+Y+Z$ の期待値は

$$E(X+Y+Z)=E(X)+E(Y)+E(Z)$$

考 $E(X+Y+Z)=E\big((X+Y)+Z\big)=E(X+Y)+E(Z)=E(X)+E(Y)+E(Z)$

n 個の確率変数の和の期待値（平均）

確率変数 X_k $(k = 1, 2, \cdots, n)$ の期待値を $E(X_k)$ とする.

確率変数の和 $X_1 + X_2 + \cdots + X_n$ の期待値は

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

☆カウンター

確率変数 X_k $(k = 1, 2, \cdots, n)$ の分布が

$$X_k = \begin{cases} 1 & （確率\ p） \\ 0 & （確率\ 1 - p） \end{cases}$$

X_k	1	0	計
確率	p	$1-p$	1

となるとき，X_k のの期待値は

$$E(X_k) = 1 \cdot p + 0 \cdot (1 - p) = p$$

このとき，確率変数の和 $X_1 + X_2 + \cdots + X_n$ の期待値は

$$\begin{aligned} E(X_1 + X_2 + \cdots + X_n) &= E(X_1) + E(X_2) + \cdots + E(X_n) \\ &= \underbrace{p + p + \cdots + p}_{n\text{個}} \\ &= pn \end{aligned}$$

1 次式の確率変数の和の期待値（平均）

確率変数 X，Y の期待値をそれぞれ $E(X)$，$E(Y)$ とする．

a, b をともに 0 以外の定数 として，確率変数の実数倍の和 $aX+bY$ の期待値は

$$E(aX+bY)=a\,E(X)+b\,E(Y)$$

㊎ 2 つの確率変数 X，Y の確率分布を次とする．

X	x_1	x_2	計
P	p_1	p_2	1

Y	y_1	y_2	y_3	計
P	q_1	q_2	q_3	1

X＼Y	y_1	y_2	y_3	計
x_1	p_{11}	p_{12}	p_{13}	p_1
x_2	p_{21}	p_{22}	p_{23}	p_2
計	q_1	q_2	q_3	1

右のように $P(X=x_k,\, Y=y_\ell)=p_{k\ell}$ とすると
$aX+bY$ の確率分布は次になる．

$aX+bY$	ax_1+by_1	ax_1+by_2	ax_1+by_3	ax_2+by_1	ax_2+by_2	ax_2+by_3	計
P	p_{11}	p_{12}	p_{13}	p_{21}	p_{22}	p_{23}	1

$$\begin{aligned}E(aX+bY)=&\ (ax_1+by_1)p_{11}+(ax_1+by_2)p_{12}+(ax_1+by_3)p_{13}\\&+(ax_2+by_1)p_{21}+(ax_2+by_2)p_{22}+(ax_2+by_3)p_{23}\\=&\ ax_1(p_{11}+p_{12}+p_{13})+ax_2(p_{21}+p_{22}+p_{23})\\&+by_1(p_{11}+p_{21})+by_2(p_{12}+p_{22})+by_3(p_{13}+p_{23})\\=&\ ax_1p_1+ax_2p_2+by_1q_1+by_2q_2+by_3q_3\\=&\ a(x_1p_1x_2p_2)+b(y_1q_1+y_2q_2+y_3q_3)\\=&\ aE(X)+bE(Y)\end{aligned}$$

㊓ X，Y の取り得る値がいくつになっても同じ計算で成り立つ．

確率変数の独立

2個の確率変数 X，Y があり，

X のとる任意の値 x_k と Y のとる任意の値 y_ℓ について

$$P(X=x_k,\, Y=y_\ell)=P(X=x_k)\cdot P(Y=y_\ell)$$

が成り立つとき 確率変数 X と Y は 独立（どくりつ）であるという.

すなわち 確率変数 X，Y について

X のとる値が $x_1,\, x_2,\, \cdots,\, x_m$

Y のとる値が $y_1,\, y_2,\, \cdots,\, y_n$

とし

$$P(X=x_k,\, Y=y_\ell)=p_{k\ell},\ P(X=x_k)=p_k,\ P(Y=y_\ell)=q_\ell$$

とすると mn 個のすべての組 $(k,\, \ell)$ で $p_{k\ell}=p_kq_\ell$ が成り立つとき

確率変数 X と Y は 独立であるという.

このとき，下の2つの表が対応する.

$X \backslash Y$	y_1	y_2	$\cdots\cdots$	y_n	計
x_1	p_{11}	p_{12}	$\cdots\cdots$	p_{1n}	p_1
x_2	p_{21}	p_{22}	$\cdots\cdots$	p_{2n}	p_2
$\vdots$		$\cdots\cdots\cdots\cdots$			$\vdots$
$\vdots$		$\cdots\cdots\cdots\cdots$			$\vdots$
x_m	p_{m1}	p_{m2}	$\cdots\cdots$	p_{mn}	p_m
計	q_1	q_2	$\cdots\cdots$	q_n	1

$X \backslash Y$	y_1	y_2	$\cdots\cdots$	y_n	計
x_1	p_1q_1	p_1q_2	$\cdots\cdots$	p_1q_n	p_1
x_2	p_2q_1	p_2q_2	$\cdots\cdots$	p_2q_n	p_2
$\vdots$		$\cdots\cdots\cdots\cdots$			$\vdots$
$\vdots$		$\cdots\cdots\cdots\cdots$			$\vdots$
x_m	p_mq_1	p_mq_2	$\cdots\cdots$	p_mq_n	p_m
計	q_1	q_2	$\cdots\cdots$	q_n	1

3 個の確率変数の独立

3 個の確率変数 X，Y，Z について

X のとる任意の値 a，Y のとる任意の値 b，Z のとる任意の値 c について

$$P(X=a, Y=b, Z=c)=P(X=a)\cdot P(Y=b)\cdot P(Z=c)$$

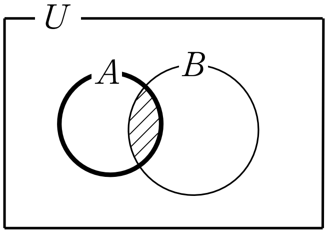
が成り立つとき 確率変数 X と Y と Z は **独立である** という．

条件付き確率

事象 A が起こったときに事象 B の起こる確率を

事象 A が起こったときの事象 B の起こる **条件付き確率**（じょうけんつき）といい

$P_A(B)$ で表す．



この確率は

$$P_A(B)=\frac{n(A\cap B)}{n(A)}=\frac{P(A\cap B)}{P(A)}$$

ただし　$n(A)\neq 0$，$P(A)\neq 0$

確率の乗法定理

2 つの事象 A，B がともに起こる確率 $P(A\cap B)$ は

$$P(A\cap B)=P(A)P_A(B)$$

ただし　A が空事象のときは $P(A\cap B)=0$

事象の独立と従属

2 つの事象 A，B があって

$$P_A(B)=P(B) \text{ かつ } P_B(A)=P(A)$$

が成り立つとき　A と B は**独立**（どくりつ）であるという．

A と B が独立でないとき A と B は**従属**（じゅうぞく）であるという．

㊟ $P_A(B)=P(B)$ と $P_B(A)=P(A)$ のうち一方が成り立てば他方の式も成り立つ．
（乗法定理からわかる）

独立な事象と乗法定理

2 つの事象 A，B について

$$A \text{ と } B \text{ が独立} \iff P(A\cap B)=P(A)\cdot P(B)$$

㊎ 事象 A と B が独立であることと，対応する確率変数 X と Y が独立であることは同値

2 個の独立な確率変数の積の期待値（平均）

確率変数 X，Y の期待値をそれぞれ $E(X)$，$E(Y)$ とする．

確率変数 X，Y が **独立** であるとき 確率変数の積 XY の期待値は

$$E(XY) = E(X)E(Y)$$

㊐ 2 つの確率変数 X，Y の確率分布を次とする．

X	x_1	x_2	計
P	p_1	p_2	1

Y	y_1	y_2	y_3	計
P	q_1	q_2	q_3	1

X \ Y	y_1	y_2	y_3	計
x_1	p_1q_1	p_1q_2	p_1q_3	p_1
x_2	p_2q_1	p_2q_2	p_2q_3	p_2
計	q_1	q_2	q_3	1

X，Y は独立なので右のように

$$P(X = x_k,\ Y = y_\ell) = P(X = x_k)\cdot Y(y = y_\ell) = p_k q_\ell$$

XY の確率分布は次になる．

XY	x_1y_1	x_1y_2	x_1y_3	x_2y_1	x_2y_2	x_2y_3	計
P	p_1q_1	p_1q_2	p_1q_3	p_2q_1	p_2q_2	p_2q_3	1

$$\begin{aligned} E(XY) &= x_1y_1\cdot p_1q_1 + x_1y_2\cdot p_1q_2 + x_1y_3\cdot p_1q_3 + x_2y_1\cdot p_2q_1 + x_2y_2\cdot p_2q_2 + x_2y_3\cdot p_2q_3 \\ &= (x_1p_1 + x_2p_2)(y_1q_1 + y_2q_2 + y_3q_3) \\ &= E(X)E(Y) \end{aligned}$$

㊟ X，Y の取り得る値がいくつになっても同じ計算で成り立つ．

3 個の独立な確率変数の積の期待値（平均）

確率変数 X，Y，Z の期待値をそれぞれ $E(X)$，$E(Y)$，$E(Z)$ とする.

確率変数 X，Y，Z が **独立** であるとき 確率変数の積 XYZ の期待値は

$$E(XYZ)=E(X)E(Y)E(Z)$$

考 $E(XYZ)=E(XY\cdot Z)=E(XY)E(Z)=E(X)E(Y)E(Z)$

n 個の独立な確率変数の積の期待値（平均）

確率変数 X_k の期待値を $E(X_k)$ とする.

n 個の確率変数 X_1，X_2，$\cdots$，X_n が **独立** であるとき

n 個の確率変数の積 $X_1X_2\cdots X_n$ の期待値は

$$E(X_1X_2\cdots X_n)=E(X_1)\cdot E(X_2)\cdot\cdots\cdot E(X_n)$$

2個の独立な確率変数の積の分散と標準偏差

確率変数 X，Y の分散をそれぞれ $V(X)$，$V(Y)$ とする．

確率変数 X，Y が 独立 であるとき

$\boxed{1}$ 確率変数の和 $X+Y$ の分散は $V(X+Y)=V(X)+V(Y)$

$\boxed{2}$ 確率変数の和 $X+Y$ の標準偏差は $\sigma(X+Y)=\sqrt{V(X)+V(Y)}$

例 確率変数 X，Y は独立であることから

$\boxed{1}$
$$\begin{aligned} V(X+Y) &= E\big((X+Y)^2\big)-\{E(X+Y)\}^2 \\ &= E(X^2+2XY+Y^2)-\{E(X)+E(Y)\}^2 \\ &= E(X^2)+2E(XY)+E(Y^2)-\{E(X)\}^2-2E(X)E(Y)-\{E(Y)\}^2 \\ &= E(X^2)-\{E(X)\}^2+E(Y^2)-\{E(Y)\}^2 \ \big(\because E(XY)=E(X)E(Y)\big) \\ &= V(X)+V(Y) \end{aligned}$$

$\boxed{2}$ $\sigma(X+Y)=\sqrt{V(X+Y)}=\sqrt{V(X)+V(Y)}$

3個の独立な確率変数の積の分散と標準偏差

確率変数 X，Y，Z の分散をそれぞれ $V(X)$，$V(Y)$，$V(Z)$ とする．

確率変数 X，Y，Z が 独立 であるとき

$\boxed{1}$ 確率変数の和 $X+Y+Z$ の分散は

$$V(X+Y+Z)=V(X)+V(Y)+V(Z)$$

$\boxed{2}$ 確率変数の和 $X+Y+Z$ の標準偏差は

$$\sigma(X+Y)=\sqrt{V(X)+V(Y)+V(Z)}$$

n 個の独立な確率変数の和の分散と標準偏差

確率変数 X_k の分散を $V(X_k)$ とする．

n 個の確率変数 X_1, X_2, $\cdots$, X_n が 独立 であるとき

$\boxed{1}$ n 個の確率変数の和 $X_1+X_2+\cdots+X_n$ の分散は

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

$\boxed{2}$ n 個の確率変数の和 $X_1+X_2+\cdots+X_n$ の標準偏差は

$$\sigma(X_1+X_2+\cdots+X_n)=\sqrt{V(X_1)+V(X_2)+\cdots+V(X_n)}$$

1次式の確率変数の和の分散

確率変数 X，Y の分散をそれぞれ $V(X)$，$V(Y)$ とする.

確率変数 X，Y が **独立** であるとき

a, b をともに0以外の定数として，確率変数の実数倍の和 $aX+bY$ の分散は

$$V(aX+bY)=a^2V(X)+b^2V(Y)$$

考 確率変数 X，Y が独立であるとき，aX，bY も独立なので

$$V(aX+bY)=V(aX)+V(bY)=a^2V(X)+b^2V(Y)$$

二項分布

ある試行で事象 A が起こる確率を p，起こらない確率を $q(=1-p)$ とおく．

この試行を n 回繰り返す反復試行において，事象 A が起こる回数を X とすると

X は確率変数であり $X=0,\ 1,\ 2,\ \cdots,\ n$ の値をとる．

$X=k$ となる確率は　$P(X=k)={}_n\mathrm{C}_k p^k q^{n-k}$

つまり　確率変数 X の確率分布は次である．

X	0	1	$\cdots$	k	$\cdots$	n	計
P	${}_n\mathrm{C}_0 q^n$	${}_n\mathrm{C}_1 pq^{n-1}$	$\cdots$	${}_n\mathrm{C}_k p^k q^{n-k}$	$\cdots$	${}_n\mathrm{C}_n p^n$	1

この分布を 確率 p に対する次数 n の **二項分布**（にこうぶんぷ） といい　$B(n,\ p)$　で表す．

補　二項分布を意味する英語は Binomial distribution

例　1 つのさいころを振ることを 180 回繰り返すとき，1 の目が出る回数を X とすると
$X=0,\ 1,\ 2,\ \cdots,\ 180$ の値をとる．
$X=k$ となる確率は　$P(X=k)={}_{180}\mathrm{C}_k\left(\dfrac{1}{6}\right)^k\left(1-\dfrac{1}{6}\right)^{180-k}$ $(k=0,\ 1,\ 2,\ \cdots,\ 180)$
X は二項分布に従い，$B\left(180,\ \dfrac{1}{6}\right)$ で表す．

二項分布の平均と分散

確率変数 X が二項分布 $B(n,\ p)$ に従うとき

$\boxed{1}$ X の期待値は $E(X)=np$

$\boxed{2}$ X の分散は $V(X)=np(1-p)$

(考) 確率変数 X が $B(n,\ p)$ に従うとすると，$q=1-p$ として次になる．

X	0	1	$\cdots$	k	$\cdots$	n	計
P	${}_n\mathrm{C}_0q^n$	${}_n\mathrm{C}_1pq^{n-1}$	$\cdots$	${}_n\mathrm{C}_kp^kq^{n-k}$	$\cdots$	${}_n\mathrm{C}_np^n$	1

$\boxed{1}$
$$\begin{aligned}
E(X)&=\sum_{k=0}^{n}kP(X=k)\\
&=\sum_{k=0}^{n}k{}_n\mathrm{C}_kp^kq^{n-k}\\
&=\sum_{k=0}^{n}n{}_{n-1}\mathrm{C}_{k-1}p^kq^{n-k}\quad(\because k{}_n\mathrm{C}_k=n{}_{n-1}\mathrm{C}_{k-1})\\
&=np\sum_{k=0}^{n}{}_{n-1}\mathrm{C}_{k-1}p^{k-1}q^{n-k}\\
&=np(p+q)^{n-1}\quad(\because \text{二項定理})\\
&=np\quad(\because p+q=1)
\end{aligned}$$

$\boxed{2}$
$$\begin{aligned}
E(X^2)&=\sum_{k=1}^{n}k^2P(X=k)\\
&=\sum_{k=0}^{n}\{k(k-1)+k\}{}_n\mathrm{C}_kp^kq^{n-k}\\
&=\sum_{k=0}^{n}k(k-1){}_n\mathrm{C}_kp^kq^{n-k}+\sum_{k=0}^{n}k{}_n\mathrm{C}_kp^kq^{n-k}\\
&=\sum_{k=2}^{n}n(n-1){}_{n-2}\mathrm{C}_{k-2}p^kq^{n-k}+E(X)\quad(\because k(k-1){}_n\mathrm{C}_k=n(n-1){}_{n-2}\mathrm{C}_{k-2})\\
&=n(n-1)p^2\sum_{k=2}^{n}{}_{n-2}\mathrm{C}_{k-2}p^{k-2}q^{n-k}+np\\
&=n(n-1)p^2\sum_{k=0}^{n-2}{}_{n-2}\mathrm{C}_kp^kq^{n-k}+np\\
&=n(n-1)p^2(p+q)^{n-2}+np\quad(\because \text{二項定理})\\
&=n(n-1)p^2+np\quad(\because p+q=1)\\
&=n^2p^2+np(1-p)
\end{aligned}$$

よって $V(X)=E(X^2)-\{E(X)\}^2=n^2p^2+np(1-p)-n^2p^2=np(1-p)$

(例) 1 つのさいころを振ることを 180 回繰り返すとき，1 の目が出る回数 X について，
X は二項分布 $B\left(180,\ \dfrac{1}{6}\right)$ に従い

平均は $E(X)=180\cdot\dfrac{1}{6}=30$

分散は $V(X)=180\cdot\dfrac{1}{6}\left(1-\dfrac{1}{6}\right)=25$

㊊ $\boxed{\text{二項分布の平均と分散}}$ はカウンターが有効である．

確率変数 X_k $(k=1,\ 2,\ \cdots,\ n)$ の分布を

$$X_k=\begin{cases}1 & (\text{確率 } p)\\ 0 & (\text{確率 } 1-p)\end{cases}$$

X_k	1	0	計
確率	p	$1-p$	1

として，期待値と分散は

$$E(X_k)=1\cdot p+0\cdot(1-p)=p$$

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

${X_k}^2$	1	0	計
確率	p	$1-p$	1

ここで

$$X=X_1+X_2+\cdots+X_n$$

として，$X=k$ となる確率は ${}_n\mathrm{C}_k p^k(1-p)^{n-k}$ $(k=0,\ 1,\ 2,\ \cdots,\ n)$

これより，確率変数 X は二項分布 $B(n,\ p)$ に従う．

よって，二項分布の平均と分散は

$$\begin{aligned}E(X)=E(X_1+X_2+\cdots+X_n)&=E(X_1)+E(X_2)+\cdots+E(X_n)\\&=\underbrace{p+p+\cdots+p}_{n\text{ 個}}\\&=np\end{aligned}$$

$$\begin{aligned}V(X)=V(X_1+X_2+\cdots+X_n)&=V(X_1)+V(X_2)+\cdots+V(X_n)\\&=\underbrace{p(1-p)+p(1-p)+\cdots+p(1-p)}_{n\text{ 個}}\\&=np(1-p)\end{aligned}$$