数学B　統計的な推測「確率分布」

確率変数／ 確率変数と確率の表記／ 確率分布／ 確率変数の期待値（平均）／
1 次式の確率変数の期待値（平均）／ 偏差／ 確率変数の分散／
確率変数の標準偏差／ 分散の計算／ 1 次式の確率変数の分散・標準偏差／
2 個以上の確率変数と確率の表記／ 同時分布／
2 個の確率変数の和の期待値（平均）／ 3 個の確率変数の和の期待値（平均）／
n 個の確率変数の和の期待値（平均）／ ☆カウンター／
1 次式の確率変数の和の期待値（平均）／ 確率変数の独立／
3 個の確率変数の独立／ 条件付き確率／ 確率の乗法定理／
事象の独立と従属／ 2 つの事象が独立であることの条件／
2 個の独立な確率変数の積の期待値（平均）／
3 個の独立な確率変数の積の期待値（平均）／
n 個の独立な確率変数の積の期待値（平均）／
2 個の独立な確率変数の積の分散・標準偏差／
3 個の独立な確率変数の積の分散・標準偏差／
n 個の独立な確率変数の積の分散・標準偏差／
1 次式の確率変数の和の分散／ 二項分布／ 二項分布の平均と分散／
連続型確率変数・確率密度関数・分布曲線／ 連続分布／ 離散型確率変数／
連続型確率変数の期待値（平均）・分散・標準偏差／ 正規分布／
正規分布の標準化／ 標準正規分布／ 正規分布表／
二項分布の正規分布による近似

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 の確率分布が右になるから

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

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

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

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

$E(X)=\sum_{k=1}^{n} x_kp_k$，$\sum_{k=1}^{n} p_k=1$ 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），あるいは単に 偏差 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 $V(X)$ または $\sigma^2(X)$ 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

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$)^2$

㊂ 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

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

$m=E(X)$ として

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

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.

$X \backslash Y$	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

すべての k と ℓ の mn 個の組み合せについての (x_k, y_ℓ) 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 $P(X=x_k)=\sum_{\ell=1}^{n} p_{k\ell}=p_{k1}+p_{k2}+\cdots+p_{kn}=p_k$

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

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

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

例 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

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

補 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

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

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

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

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

$X \backslash Y$	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

$$\begin{aligned}E(aX+bY)=&\ (ax_1+by_1)p_{11}+(ax_1+by_2)p_{12}+(ax_1+by_3)p_{13}\\&+(ax_2+by_1)p_{21}+(ax_2+by_2)p_{22}+(ax_2+by_3)p_{23}\\=&\ ax_1(p_{11}+p_{12}+p_{13})+ax_2(p_{21}+p_{22}+p_{23})\\&+by_1(p_{11}+p_{21})+by_2(p_{12}+p_{22})+by_3(p_{13}+p_{23})\\=&\ ax_1p_1+ax_2p_2+by_1q_1+by_2q_2+by_3q_3\\=&\ a(x_1p_1+x_2p_2)+b(y_1q_1+y_2q_2+y_3q_3)\\=&\ aE(X)+bE(Y)\end{aligned}$$

㊉ 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 $p_{k\ell}=p_k q_\ell$ 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 \ 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 \ 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

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

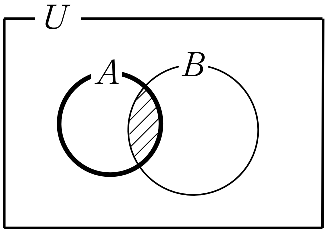
が成り立つとき 確率変数 X と Y と Z は 独立である という．

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**（じょうけんつきかくりつ）といい

$P_A(B)$ で表す.



この確率は

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

ただし　$P(A)\neq 0$

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）

2つの事象が独立であることの条件

2つの事象 A，B について

A と B 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

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

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

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

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

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・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 $\sigma(X+Y)=\sqrt{V(X)+V(Y)}$

㊞ 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・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・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

1. X の期待値は $E(X)=np$
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

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

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

別 $\boxed{\text{二項分布の平均と分散}}$ 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

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

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

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

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

連続型確率変数・確率密度関数・分布曲線

実数のある区間全体に値をとる確率変数 X に対して，

1 つの関数 $y=f(x)$ が対応して次の性質をもつとする．

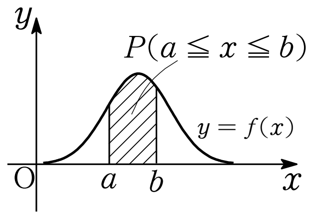
$\boxed{1}$ $f(x) \geqq 0$

$\boxed{2}$ X が $a \leqq x \leqq b$ の範囲の値をとる確率 $P(a \leqq X \leqq b)$ は

曲線 $y=f(x)$ と x 軸および 2 直線 $x=a$, $x=b$ で

囲まれた部分の面積に等しい．

つまり $P(a \leqq X \leqq b)=\int_a^b f(x)\,dx$



$\boxed{3}$ 曲線 $y=f(x)$ と x 軸の間の面積は 1 である．

つまり $\int_{-\infty}^{\infty} f(x)\,dx=1$

このとき

X を **連続型確率変数**（れんぞくがたかくりつへんすう）といい，関数 $f(x)$ を X の **確率密度関数**（かくりつみつどかんすう）という．

$y=f(x)$ のグラフをその **分布曲線**（ぶんぷきょくせん）という．

連続分布

確率密度関数によって確率分布が定められているとき，

その分布を **連続分布**（れんぞくぶんぷ）という．

離散型確率変数

連続型確率変数に対して，とびとびの値をとる確率変数を

離散型確率変数（りさんがたかくりつへんすう）という．

(補) 大雑把に説明すると，連続型は $\int$，離散型は Σ

連続型確率変数の期待値（平均）・分散・標準偏差

連続型確率変数 X のとり得る値の範囲が $a \leqq X \leqq b$ であり，

その確率密度関数を $f(x)$ とするとき

$\boxed{1}$ 確率変数 X の期待値は $E(X) = \int_a^b xf(x)\,dx$

$\boxed{2}$ $E(X) = m$ とする．

確率変数 X の分散は $V(X) = \int_a^b (x-m)^2 f(x)\,dx$

$$= \int_a^b x^2 f(x)\,dx - m^2$$

$\boxed{3}$ 確率変数 X の標準偏差は $\sigma(X) = \sqrt{V(X)}$

補 離散型確率変数の Σ を $\int$ で置き換えている．

要

確率変数 X が離散型，連続型かで次のようになる．

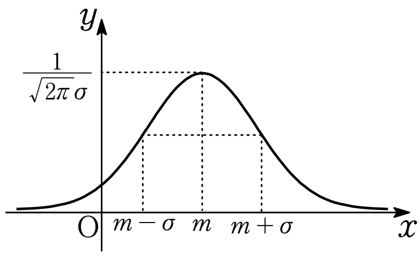
	離散型	連続型
期待値 $E(X)$	$\sum_{k=1}^{n} x_k P(X=k)$	$\int_a^b xf(x)\,dx$
分散 $V(X)$	$\sum_{k=1}^{n} (x_k - m)^2 P(X=k)$	$\int_a^b (x-m)^2 f(x)\,dx$
分散 $V(X)$	$\sum_{k=1}^{n} {x_k}^2 P(X=k) - \{E(X)\}^2$	$\int_a^b x^2 f(x)\,dx - m^2$

正規分布

連続型確率変数 X の確率密度関数 $f(x)$ が

$$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}\,e^{-\frac{(x-m)^2}{2\sigma^2}}$$

で与えられているとき



X は **正規分布**（せいきぶんぷ） $N(m,\sigma^2)$ **に従う** という．

ここで m は X の平均，σ^2 は X の分散である．なお，σ は X の標準偏差である．

$y=f(x)$ のグラフを **正規分布曲線**（せいきぶんぷきょくせん）といい，次の性質がある．

1. 直線 $x=m$ に関して対称になる．
2. $f(x)$ の値は $x=m$ で最大値をとる．
3. x 軸を漸近線とする．
4. 曲線の山は，σ が大きくなるほど低くなって横に広がり，
   σ が小さくなるほど高くなって対称軸 $x=m$ のまわりに集まる．

㊗補 正規分布を意味する英語は Normal distribution

補 $f(x)$ が確率密度関数になることは知られているとする．

考 $\boxed{4}$ σ が大きいとばらつきも大きくなり，小さいとばらつきが小さくなる．

正規分布の標準化

確率変数 X が正規分布 $N(m,\,\sigma^2)$ に従うとき

$$Z=\frac{X-m}{\sigma}\quad \text{すなわち}\quad Z=\frac{X-(X\text{の平均})}{(X\text{の標準偏差})}$$

とおくと 確率変数 Z は $N(0,\,1)$ に従う.

この Z を **標準化** した確率変数という.

㊂ X が $N(m,\,\sigma^2)$ に従うとき $E(X)=m,\ V(X)=\sigma^2$

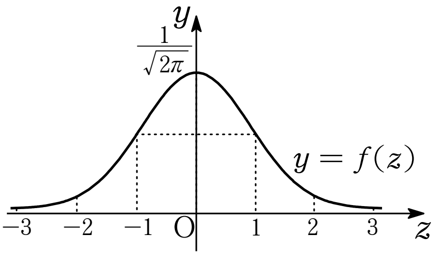
確率変数 Z の平均は

$$E(Z)=E\left(\frac{X-m}{\sigma}\right)=E\left(\frac{1}{\sigma}X-\frac{m}{\sigma}\right)=\frac{1}{\sigma}E(X)-\frac{m}{\sigma}=\frac{1}{\sigma}\cdot m-\frac{m}{\sigma}=0$$

確率変数 Z の分散は

$$V(Z)=V\left(\frac{X-m}{\sigma}\right)=V\left(\frac{1}{\sigma}X-\frac{m}{\sigma}\right)=\left(\frac{1}{\sigma}\right)^2V(X)=\frac{1}{\sigma^2}\cdot\sigma^2=1$$

標準正規分布



ひょうじゅんせいきぶんぷ
正規分布 $N(0,\,1)$ を **標準正規分布** という．

標準正規分布の確率密度関数は

$$f(z)=\frac{1}{\sqrt{2\pi}}\,e^{-\frac{z^2}{2}}$$

$y=f(z)$ のグラフを **標準正規分布曲線** といい，次の性質がある．

には次の性質がある．

$\boxed{1}$ $f(z)$ の値は $z=0$ で最大値をとる．

$\boxed{2}$ z 軸を漸近線とする．

$\boxed{3}$ y 軸に関して対称になるので

$$\int_0^{\infty} f(z)\,dz=\int_{-\infty}^{0} f(z)\,dz=0.5$$

$\boxed{4}$ 確率変数 Z が正規分布 $N(0,\,1)$ に従うとき

$$P(0\leqq Z\leqq u)=\int_0^{u} f(z)\,dz$$

この値は正規分布表から求まる．

要

確率変数 X が正規分布 $N(m,\,\sigma^2)$ に従うとき

$$Z=\frac{X-m}{\sigma}\quad \text{すなわち}\quad Z=\frac{X-(X\text{の平均})}{(X\text{の標準偏差})}$$

と標準化すると，確率変数 Z は標準正規分布 $N(0,\,1)$ に従う．

標準正規分布曲線 $y=f(z)$ において z 座標 1 が σ に対応する．

例えば

$$m-\sigma\leqq x\leqq m+\sigma \iff -1\leqq z\leqq 1$$

すなわち

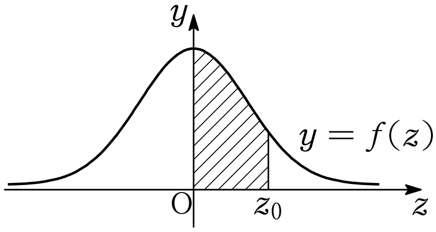
$$P(m-\sigma\leqq X\leqq m+\sigma)=P(-1\leqq Z\leqq 1)$$

正規分布表

標準正規分布曲線 $y=f(z)$ に対し

右図斜線部の面積

$$\int_0^{z_0} f(z)\,dz$$



の近似値を次のように表にしたものを **正規分布表** という．

z_0	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990

上表は小数第５位を四捨五入して小数第４位までの値を表わしている．

確率変数 Z が標準正規分布 $N(0,\,1)$ に従うとすると，次の確率が求まる．

$$P(0 \leqq Z \leqq z_0)=\int_0^{z_0} f(z)\,dz$$

例 確率変数 Z が $N(0,\,1)$ に従うとして $P(0 \leqq Z \leqq 1.96)=\int_0^{1.96} f(z)\,dz=0.4750$

二項分布の正規分布による近似

二項分布 $B(n,\ p)$ に従う確率変数 X は n が十分大きいとき

近似的に正規分布 $N\big(np,\ np(1-p)\big)$ に従う．

このことから

$$Z=\frac{X-np}{\sqrt{np(1-p)}}$$

とおくと 確率 Z は標準正規分布 $N(0,\ 1)$ に従う．

㊐ X が $B\left(180,\ \dfrac{1}{6}\right)$ に従うとき，180 が十分大きいので $N(30,\ 25)$ に従う．

$$Z=\frac{X-30}{\sqrt{25}}=\frac{X-30}{5}$$

とおくと Z は $N(0,\ 1)$ に従う．

$P(30 \leqq X \leqq 40)=P\left(0 \leqq \dfrac{X-30}{5} \leqq 2\right)=P(0 \leqq Z \leqq 2)=0.4772$