独立な確率変数に関する期待値，分散，共分散

独立な複数のr.v.について

Th

X, Y : r.v., independent

$\Rightarrow {}^{\forall} \underbrace{h_X(x)}_{x\text{だけの関数}}, \underbrace{h_Y(y)}_{y\text{だけの関数}},$

$$E[h_X(X) \cdot h_Y(Y)] = E[h_X(X)]\, E[h_Y(Y)]$$

Proof

(i) Suppose that X and Y are discrete r.v. .

It follows that

$$E[h_X(X) \cdot h_Y(Y)]$$

$$= \sum_i \sum_j h_X(x_i)\, h_Y(y_j)\, P_{ij}$$

X, Y : independent

$$= \sum_i \sum_j \underbrace{h_X(x_i)}_{j\text{と無関係}}\, h_Y(y_j)\, \underbrace{P_{i\circ}}_{=P(X=x_i)} P_{\circ j}$$

$$= \sum_i h_X(x_i)\, P_{i\circ} \cdot \sum_j h_Y(y_j)\, P_{\circ j}$$

$$= E[h_X(X)] \cdot E[h_Y(Y)] .$$ ┘

(ii) Suppose that X and Y are continuous r.v.

It holds that

$$E[h_X(X)h_Y(Y)]$$

$$= \int_{x=-\infty}^{\infty}\int_{y=-\infty}^{\infty} h_X(x)h_Y(y)f(x,y)\,dx\,dy$$

$$= \int_{x=-\infty}^{\infty}\int_{y=-\infty}^{\infty} h_X(x)h_Y(y)f_X(x)f_Y(y)\,dx\,dy$$

$$= \int_{x=-\infty}^{\infty} h_X(x)f_X(x)\,dx \cdot \int_{y=-\infty}^{\infty} h_Y(y)f_Y(y)\,dy$$

$$= E[h_X(X)]\cdot E[h_Y(Y)].$$

//

Th

$D=\{(x,y)\in\mathbb{R}^2 \mid a\leqq x\leqq b,\ c\leqq y\leqq d\}$

$f(x,y)$: D上で連続

$f(x,y)=g(x)h(y)$

$$\Rightarrow \iint_D f(x,y)\,dx\,dy = \int_a^b g(x)\,dx\cdot\int_c^d h(y)\,dy$$

この二重積分に関する定理を使っている。

Cor

X, Y : r.v., independent

$\Longrightarrow$ (1) $E[XY] = E[X]E[Y]$

(2) $\mathrm{Cov}[X, Y] = 0$

Proof

(1) Letting $h_X(x) = x$ and $h_Y(y)$

in the previous theorem, we obtain

$$E[XY] = E[X]E[Y].$$

(2) We can prove the desired result as follows:

$$\begin{aligned} \mathrm{Cov}[X, Y] &= E[XY] - E[X]E[Y] \\ &= E[X]E[Y] - E[X]E[Y] \\ &= 0. \end{aligned}$$

X, Y : independent

//

Cor

X, Y : r.v., independent

$\Rightarrow V[aX+bY] = a^2 V[X] + b^2 V[Y]$

$V[X-Y] = V[X] + V[Y]$

注意!

$V[X-Y] = 1^2 V[X] + 2\cdot 1 \cdot (-1) \cdot \underbrace{Cov[X, Y]}_{=0} + (-1)^2 V[Y]$

$= V[X] + V[Y]$.

r.v. X に対してそのモーメント母関数を

$$M_X(t) = E[e^{tX}]$$

と書くことにする。

Cor

X, Y : r.v., independent

$$\Rightarrow M_{X+Y}(t) = M_X(t) M_Y(t)$$

Proof

We have

$$\begin{aligned} M_{X+Y}(t) &= E[e^{t(X+Y)}] \\ &= E[e^{tX} \cdot e^{tY}] \\ &= E[e^{tX}] E[e^{tY}] \quad \text{) } X, Y : \text{independent} \\ &= M_X(t) \cdot M_Y(t) . \quad // \end{aligned}$$

※これは分布の再生性の証明に用いる。

cf. (参考)

X : r.v.

$M_X(t) = E[e^{tX}]$

$a \in \mathbb{R}$

$M_{(aX)}(t) = E[e^{t\cdot(aX)}]$

the moment generating function for r.v. aX

$\Longrightarrow M_{(aX)}(t) = M_X(at)$

Proof

$$M_{(aX)}(t) = E[e^{t\cdot(aX)}]$$
$$= E[e^{(at)X}]$$
$$= M_X(at)$$

//

例.

(X, Y) の同時確率分布

$X \backslash Y$	0	1	計
0	$5/12$	$5/12$	$5/6$
1	$1/12$	$1/12$	$1/6$
計	$1/2$	$1/2$	1

- X, Y : independent

i.e. $P(X=x_i, Y=y_j) = P(X=x_i)P(Y=y_j)$

- $E[XY] = E[X]E[Y]$

$$E[X] = 0 \cdot \frac{5}{6} + 1 \cdot \frac{1}{6} = \frac{1}{6}$$

$$E[Y] = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{1}{2}$$

$$E[XY] = 0 \cdot 0 \cdot \frac{5}{12} + 0 \cdot 1 \cdot \frac{5}{12} + 1 \cdot 0 \cdot \frac{1}{12} + 1 \cdot 1 \cdot \frac{1}{12} = \frac{1}{12}$$

$\therefore E[XY] = E[X]E[Y]$.

//

例 (共分散は0だが独立ではないr.v.の例)

同時確率分布

$X \backslash Y$	2	1	-1	-2	計
1	1/4	0	0	1/4	1/2
-1	0	1/4	1/4	0	1/2
計	1/4	1/4	1/4	1/4	1

- X, Y are not independent.

- $Cov[X, Y] = 0$

$$E[X] = E[Y] = 0$$

$$Cov[X, Y] = E[XY] - E[X]E[Y] = E[XY]$$

$$= 1 \cdot 2 \cdot \frac{1}{4} + 1 \cdot (-2) \cdot \frac{1}{4} + (-1) \cdot 1 \cdot \frac{1}{4} + (-1)(-1)\frac{1}{4}$$

$$= 0.$$

※ $\boxed{X, Y : \text{independent} \Rightarrow Cov[X, Y] = 0}$

だが、逆は必ずしも真ならず!

標本平均の期待値と分散

$X_1, X_2, \cdots, X_n$: r.v.

$E[X_1] = E[X_2] = \cdots = E[X_n] = \mu$

$\bar{X}_n \equiv \frac{1}{n} \sum_{i=1}^{n} X_i$ ← 標本平均

$\Longrightarrow E[\bar{X}_n] = \mu$

Proof.

We have

$$E[\bar{X}_n] = E\left[\frac{1}{n}(X_1 + X_2 + \cdots + X_n)\right]$$
$$= \frac{1}{n}\left(E[X_1] + \cdots + E[X_n]\right)$$
$$= \frac{1}{n} \cdot n\mu = \mu.$$

//

※ $X_1, \cdots, X_n$ は独立でなくてもいいし、
同一分布に従っていなくてもよい。

※ 分散はそれぞれ異なってもいいし、
存在しなくてもかまわない。

$X_1, \cdots, X_n$: r.v., independent

$V[X_1] = \cdots = V[X_n] = \sigma^2$

$\bar{X}_n \equiv \frac{1}{n}\sum_{i=1}^{n} X_i$

$\Rightarrow V[\bar{X}] = \frac{\sigma^2}{n}$

Proof.

It holds that

$$
\begin{aligned}
V[\bar{X}_n] &= V\left[\frac{1}{n}(X_1 + \cdots + X_n)\right] \\
&= \frac{1}{n^2} V[X_1 + \cdots + X_n] \\
&= \frac{1}{n^2}\left(V[X_1] + V[X_2] + \cdots + V[X_n]\right) \\
&= \frac{1}{n^2} n \cdot \sigma^2 = \frac{\sigma^2}{n}.
\end{aligned}
$$

//

※ $X_1, \cdots, X_n$ は同一分布に従っていなくてもよいし、期待値は異なっていてもかまわない。

実際によく使うのは次の形

Th

$X_1, \cdots, X_n \sim iid(\mu, \sigma^2)$

$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$

$\Rightarrow E[\bar{X}_n] = \mu$

$V[\bar{X}_n] = \frac{\sigma^2}{n}$

- iid とは、independent and identically distributed
  独立に同一の分布に従っているということ。

- $iid(\mu, \sigma^2)$ と書くと、その同一分布の
  平均が μ、分散が σ^2 であることを意味する。

- $iid\ N(\mu, \sigma^2)$ を使うこともある。
  その同一分布が $N(\mu, \sigma^2)$ であること。

cf.

Th.

$X_1, \cdots, X_n \sim \text{iid } N(\mu, \sigma^2)$

$\bar{X}_n \equiv \frac{1}{n} \sum_{i=1}^{n} X_i$

$\Rightarrow \bar{X}_n \sim N(\mu, \frac{\sigma^2}{n})$

- 正規分布の再生性の応用として示すことができる。
- 前ページの定理とくらべると、仮定と結論の両方が強い。

Th. (CLT)

$X_1, \cdots, X_n \sim \text{iid } (\mu, \sigma^2)$

$\bar{X}_n \equiv \frac{1}{n} \sum_{i=1}^{n} X_i$

$\Rightarrow \bar{X}_n \underset{\cdot}{\sim} N(\mu, \frac{\sigma^2}{n})$

(n: sufficiently large)

中心極限定理

(central limit theorem ; CLT).

例.

国民の体重
(μ, σ^2)

ある船の乗船客の体重

$X_i \quad (n=1, 2, \cdots, n)$

independent

$E[X_i] = \mu$

$V[X_i] = \sigma^2 (>0)$

これらは既知とする。

$S_n = \sum_{i=1}^{n} X_i$ n人乗ったときの総体重, r.v.

$\frac{S_n}{n} = \frac{1}{n} \sum_{i=1}^{n} X_i$ 〃 平均体重, r.v.

○これらの r.v. の期待値は？

$$E[S_n] = E[X_1 + \cdots + X_n]$$
$$= E[X_1] + \cdots + E[X_n]$$
$$= n\mu$$

$$E\left[\frac{S_n}{n}\right] = \frac{1}{n} E[S_n] = \mu$$

※ 期待値 $E[S_n]$, $E\left[\frac{S_n}{n}\right]$ を求める段階では

$X_1, \cdots, X_n$: independent

という仮定は必要ない。

・ S_n と $\dfrac{S_n}{n}$ の分散を求めよ。

$$V[S_n] = V[X_1 + \cdots + X_n]$$

$X_1, \cdots, X_n$: independent

$$= V[X_1] + \cdots + V[X_n]$$

$$= n\sigma^2$$

$$V\left[\frac{S_n}{n}\right] = \frac{1}{n^2} V[S_n]$$

$$= \frac{\sigma^2}{n}$$

※ n が大きくなると

総体重 S_n の分散 $V[S_n]$ は大きくなるが、

平均体重 $\dfrac{S_n}{n}$ の分散 $V\left[\dfrac{S_n}{n}\right]$ は小さくなる。

大きい船の方が、ある意味では"安定する"。

重要!

復習

X, Y : r.v., independent

$$\boxed{\begin{aligned}&\forall h_X(x),\ h_Y(y),\\ &\quad E[h_X(X)\,h_Y(Y)] = E[h_X(X)]\,E[h_Y(Y)]\end{aligned}}$$

$$\downarrow$$

$$\boxed{E[XY] = E[X]\,E[Y]}$$

$$\downarrow$$

$$\boxed{\mathrm{Cov}(X, Y) = E[XY] - E[X]E[Y] = 0}$$

$$\boxed{\begin{aligned}&V[aX+bY]\\ &= a^2V[X] + 2ab\,\mathrm{Cov}(X,Y) + b^2V[Y]\end{aligned}}$$

$$\Big\downarrow{\scriptstyle b=0}$$

$$\boxed{V[aX] = a^2V[X]}$$

X, Y : independent

$$\boxed{V[X+Y] = V[X] + V[Y]}$$

$$\boxed{\begin{aligned}&V\left[\frac{1}{n}(X_1+\cdots+X_n)\right]\\ &= \frac{1}{n^2}V[X_1+\cdots+X_n]\\ &= \frac{1}{n^2}(V[X_1]+\cdots+V[X_n])\\ &= \frac{1}{n^2}\cdot n\sigma^2 = \frac{\sigma^2}{n}\end{aligned}}$$

例

保険契約 (μ, σ^2)

$$X_i = \begin{cases} 1 & \text{if 事故発生} \\ 0 & \text{otherwise} \end{cases}$$

(第 i 番目の契約案件について)

$X_1, \cdots, X_n$: r.v., independent

- $S_n = \sum_{i=1}^{n} X_i$ n 件契約したときの事故件数, r.v.
- $\frac{S_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i$ 事故契約の総契約数に対する比率, r.v.

保険会社にとっては、これらの r.v. に関心がある。

- $E[S_n] = E[X_1 + \cdots + X_n]$
  $= E[X_1] + \cdots + E[X_n] = n\mu$

期待値の線型性

- $E\left[\frac{S_n}{n}\right] = \frac{1}{n}E[S_n] = \mu$

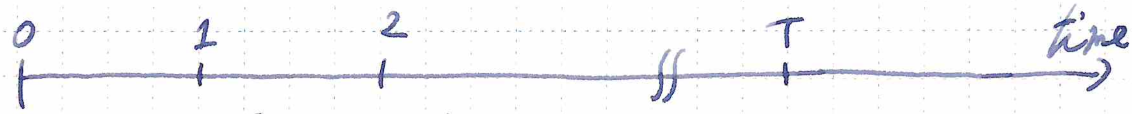
- $V[S_n] = V[X_1 + \cdots + X_n]$
  $= V[X_1] + \cdots + V[X_n]$
  $= n\sigma^2$

$X_1, \cdots, X_n$: independent

- $V\left[\frac{S_n}{n}\right] = \frac{1}{n^2}V[S_n] = \frac{\sigma^2}{n}$

※ 契約件数を大きくすると、
事故契約の総契約数に対する比率
は安定する。

例.



S_0　　$S_0 e^{R_1} = S_1$　　$S_1 e^{R_2} = S_0 e^{R_1+R_2} = S_2$　　$S_{T-1} e^{R_T} = S_0 e^{R_1+\cdots+R_T} = S_T$

S_0：時点における株価

R_t：時点 $t-1$ から t にかけての期待収益率，r.v.

$R_1, R_2, \cdots, R_t, \cdots$：independent

$$\left.\begin{array}{l} E[R_t]=\mu \\ V[R_t]=\sigma^2 \end{array}\right)$$ とする。

このとき．

$$E[R_1+\cdots+R_T] = T\mu$$

$$E\left[\frac{1}{T}\sum_{t=1}^{T} R_t\right] = \mu$$

$$V[R_1+\cdots+R_T] = T\sigma^2$$

$$V\left[\frac{1}{T}\sum_{t=1}^{T} R_t\right] = \frac{\sigma^2}{T}$$

※ 株を長く保有しつづけると収益率の平均は．

この意味で安定する。

おまけ

同時モーメント母関数

(同時)モーメント母関数

Def. ((同時)モーメント母関数)

$$M(s,t) = E\left[e^{sX+tY}\right]$$

$$= \begin{cases} \sum_i \sum_j p_{ij} e^{sx_i + ty_j} & (\text{discrete}) \\ \int_{x=-\infty}^{\infty} \int_{y=-\infty}^{\infty} f(x,y) e^{sx+ty} dx dy & (\text{conti.}) \end{cases}$$

Th

$$E[X] = M_1(0,0)$$

$$E[Y] = M_2(0,0)$$

$$V[X] = M_{11}(0,0) - \left(M_1(0,0)\right)^2$$

$$V[Y] = M_{22}(0,0) - \left(M_2(0,0)\right)^2$$

$$Cov[X,Y] = M_{12}(0,0) - M_1(0,0) M_2(0,0)$$

ここで $M_1(s,t) = \frac{\partial}{\partial s} M(s,t)$

$$M_2(s,t) = \frac{\partial}{\partial t} M(s,t)$$

$$M_{12}(s,t) = \frac{\partial}{\partial t}\left(\frac{\partial M(s,t)}{\partial s}\right)$$

$$= \frac{\partial}{\partial s}\left(\frac{\partial M(s,t)}{\partial t}\right) = M_{21}(s,t)$$ など

$$\boxed{\begin{aligned} E[X] &= M_1(0,0) \\ E[Y] &= M_2(0,0) \end{aligned}}$$

<u>Proof</u> (discrete)

$$M(s,t) = \sum_i \sum_j P_{ij}\, e^{s x_i + t y_j}$$

$$M_1(s,t) = \sum_i \sum_j P_{ij}\, x_i\, e^{s x_i + t y_j}$$

$$\therefore M_1(0,0) = \sum_i \sum_j P_{ij}\, x_i = E[X].$$

//

$$\boxed{\begin{aligned} E[X^2] &= M_{11}(0,0) \\ E[Y^2] &= M_{22}(0,0) \\ E[XY] &= M_{12}(0,0) \end{aligned}}$$

<u>Proof</u>

$$M_{11}(s,t) = \sum_i \sum_j P_{ij}\, x_i^2\, e^{s x_i + t y_j}$$

$$\therefore M_{11}(0,0) = \sum_i \sum_j P_{ij}\, x_i^2 = E[X^2]$$

Similarly, we can obtain $M_{22}(0,0) = E[Y^2]$.

$$M_{12}(s,t) = \sum_i \sum_j P_{ij}\, x_i y_j\, e^{s x_i + t y_j}$$

$$\therefore M_{12}(0,0) = \sum_i \sum_j P_{ij}\, x_i y_j = E[XY]$$

//

以上で IA が証明された。

Th (再掲)

$M(s,t) = E[e^{sX+tY}]$

$\Rightarrow$ $E[X] = M_1(0,0)$

$E[Y] = M_2(0,0)$

$V[X] = M_{11}(0,0) - (M_1(0,0))^2$

$V[Y] = M_{22}(0,0) - (M_2(0,0))^2$

$Cov[X,Y] = M_{12}(0,0) - M_1(0,0)M_2(0,0)$

ex

X \ Y	0	1	2
0	1/8	1/8	0
1	1/8	2/8	1/8
2	0	1/8	1/8

同時モーメント母関数 $M(s,t)$を導出し、それを使って
$E[X], E[Y]$,
$V[X], V[Y], Cov[X,Y]$
を計算する。

$$M(s,t) \equiv E[e^{sX+tY}]$$

$$= \frac{1}{8}e^{s\cdot0+t\cdot0} + \frac{1}{8}e^{s\cdot0+t\cdot1} + 0\cdot e^{s\cdot0+t\cdot2}$$

$$+ \frac{1}{8}e^{s\cdot1+t\cdot0} + \frac{2}{8}e^{s\cdot1+t\cdot1} + \frac{1}{8}e^{s\cdot1+t\cdot2}$$

$$+ 0\cdot e^{s\cdot2+t\cdot0} + \frac{1}{8}e^{s\cdot2+t\cdot1} + \frac{1}{8}e^{s\cdot2+t\cdot2}$$

$$\therefore M(s,t) = \frac{1}{8} + \frac{1}{8}e^t + \frac{1}{8}e^s + \frac{2}{8}e^{s+t}$$

$$+ \frac{1}{8}e^{s+2t} + \frac{1}{8}e^{2s+t} + \frac{1}{8}e^{2s+2t}$$

$$M_1(s,t) = \frac{1}{8}e^{s} + \frac{2}{8}e^{s+t} + \frac{1}{8}e^{s+2t} + \frac{2}{8}e^{2s+t} + \frac{2}{8}e^{2s+2t}$$

$$\therefore M_1(0,0) = \frac{1}{8}(1+2+1+2+2) = 1$$

$$\therefore E[X] = M_1(0,0) = 1.$$

$$M_2(s,t) = \frac{1}{8}e^{t} + \frac{2}{8}e^{s+t} + \frac{2}{8}e^{s+2t} + \frac{1}{8}e^{2s+t} + \frac{2}{8}e^{2s+2t}$$

$$\therefore M_2(0,0) = \frac{1}{8}(1+2+2+1+2) = 1$$

$$\therefore E[Y] = M_2(0,0) = 1$$

$$M_{11}(s,t) = \frac{1}{8}e^{s} + \frac{2}{8}e^{s+t} + \frac{1}{8}e^{s+2t} + \frac{4}{8}e^{2s+t} + \frac{4}{8}e^{2s+2t}$$

$$\therefore M_{11}(0,0) = \frac{1}{8}(1+2+1+4+4) = \frac{3}{2}$$

$$\therefore V[X] = M_{11}(0,0) - (M_1(0,0))^2$$

$$= \frac{3}{2} - 1^2 = \frac{1}{2}$$

$$M_{22}(s,t) = \frac{1}{8}e^{t} + \frac{2}{8}e^{s+t} + \frac{4}{8}e^{s+2t} + \frac{1}{8}e^{2s+t} + \frac{4}{8}e^{2s+2t}$$

$$\therefore M_{22}(0,0) = \frac{1}{8}(1+2+4+1+4) = \frac{3}{2}$$

$$\therefore V[Y] = M_{22}(0,0) - (M_2(0,0))^2$$

$$= \frac{3}{2} - 1^2 = \frac{1}{2}$$

$$M_{12}(s,t) = \frac{2}{8}e^{s+t} + \frac{2}{8}e^{s+2t} + \frac{2}{8}e^{2s+t} + \frac{4}{8}e^{2s+2t}$$

$$\therefore M_{12}(0,0) = \frac{1}{8}(2+2+2+4) = \frac{5}{4}$$

$$\therefore Cov[X,Y] = M_{12}(0,0) - M_1(0,0)\,M_2(0,0)$$

$$= \frac{5}{4} - 1 \cdot 1 = \frac{1}{4}$$

独立な確率変数に関する期待値、分散、共分散

問題1. 二つの確率変数X, Yが独立ならば、$E[XY] = E[X]E[Y]$が成り立つ。このことを証明しなさい。

問題2. 二つの確率変数X, Yが独立ならば、以下が成り立つ。このことを問題1の結果を用いて証明しなさい。

(1) $Cov[X, Y] = 0$

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

(3) $V[X - Y] = V[X] + V[Y]$

問題3. 確率変数XとYの同時確率分布が次の表で与えられているとする。

X \ Y	2	4	計
1	1/4	1/4	
3	1/4	1/4	
計			

XとYが独立であることと両者の共分散がゼロになることを確認しなさい。

問題4. 確率変数$X_1, X_2, \cdots, X_n$は同じ分布に従い、期待値はμ、分散はσ^2とする。また、$\overline{X}_n \equiv \frac{1}{n}\sum_{i=1}^{n} X_i$とおく。このとき、

$$E[\overline{X}_n] = \mu,$$

$$V[\overline{X}_n] = \frac{\sigma^2}{n}$$

である。このことを証明しなさい。その際、$X_1, X_2, \cdots, X_n$が独立であるという仮定をどこで使っているか確認しなさい。

問題5. (航空会社)

あなたは航空会社に勤めている。定員n人の飛行機が満席になると仮定する。乗客の体重を表す確率変数Xは、平均60kg、分散$10^2[kg^2]$の分布をする。$S_n = \sum_{i=1}^{n} X_i$とおくと、これは乗客数がn人のときの総体重、$S_n/n = \frac{1}{n}\sum_{i=1}^{n} X_i$は平均体重である。

(1) 乗客の総体重S_nと平均体重S_n/nの期待値を求めなさい。

(2) 乗客の総体重S_nと平均体重S_n/nの分散を求めなさい。

(3) 定員が20名の飛行機と100名の飛行機では、乗客の平均体重の標準偏差はどれだけ違うか？