
$$V(X) = E(X^2) - (E(X))^2 \tag{1}$$
$$V(Y) = E(Y^2) - (E(Y))^2 \tag{2}$$
$$Cov(X,Y) = E((X - E(X)(Y - E(Y))) \tag{3}$$
$$= E(XY) - E(XE(Y)) - E(E(X)Y) - E(X)E(Y) \tag{4}$$
$$= E(XY) - E(X)E(Y) \tag{5}$$

$$r = \frac{s_{xy}}{s_x s_y} \tag{6}$$
$$= \frac{Cov(X,Y)}{\sqrt{V(X)}\sqrt{V(Y)}} \tag{7}$$
$$= \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{y})^2}} \tag{8}$$

X, Y が独立である場合

$$E(XY) = E(X)E(Y), \quad Cov(XY) = 0 \tag{9}$$
$$V(X+Y) = V(X) + 2Cov(XY) + V(Y) = V(X) + V(Y) \tag{10}$$

---

1. $E(X_k) = 0,\ E(Y_k) = 0\ (k = 1, \ldots, n)$ とする。
   n 個の確率ベクトル (X_k, Y_k) が共分散 $\sigma_{X,Y} = E(X_kY_k)$ を持つ (2 変量の) 分布からの無作為標本とする。この時、

   $$\hat{\sigma}_{X,Y} = \frac{1}{n-1}\sum_{k=1}^{n}(X_k - \bar{X})(Y_k - \bar{Y}) \quad (\text{ただし、}\bar{X} = \frac{1}{n}\sum_{k=1}^{n}X_k, \bar{Y} = \frac{1}{n}\sum_{k=1}^{n}Y_k) \tag{11}$$

   が $\sigma_{X,Y}$ の不偏推定量であるか調べよ。不偏推定量でない場合は不偏推定量になるように修正せよ。
   (HINT:$k \neq k'$ ならば X_k と $Y_{k'}$ は独立なので $E(X_kY_{k'}) = 0$ となる)

   ....

   条件をまとめると次の式となる。

   $$E(X_1) = \cdots = E(X_n) = 0, \quad E(Y_1) = \cdots = E(Y_n) = 0 \tag{12}$$
   $$\sigma_{X,Y} = E(X_1Y_1) = \cdots = E(X_nY_n) \tag{13}$$
   $$k \neq k' \Rightarrow E(X_kY_{k'}) = 0 \tag{14}$$

$E(\hat{\sigma}_{X,Y})$ を計算する。

$$E(\hat{\sigma}_{X,Y}) = E\left(\frac{1}{n-1}\sum_{k=1}^{n}(X_k-\bar{X})(Y_k-\bar{Y})\right) \tag{15}$$

$$= \frac{1}{n-1}\sum_{k=1}^{n}(E(X_kY_k)-E(\bar{X}Y_k)-E(X_k\bar{Y})+E(\bar{X}\bar{Y})) \tag{16}$$

$E(\bar{X}Y_k)$ は次のように計算できる。

$$E(\bar{X}Y_k) = E\left(\frac{1}{n}\sum_{i=1}^{n}X_iY_k\right) = \frac{1}{n}\sum_{i=1}^{n}E(X_iY_k) = \frac{1}{n}E(X_kY_k) \tag{17}$$

同様に、$E(X_k\bar{Y}) = \frac{1}{n}E(X_kY_k)$ である。
$E(\bar{X}\bar{Y})$ を計算する。

$$E(\bar{X}\bar{Y}) = E\left(\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)\left(\frac{1}{n}\sum_{j=1}^{n}Y_j\right)\right) = \frac{1}{n^2}E\left(\sum_{i=1}^{n}\sum_{j=1}^{n}X_iY_j\right) \tag{18}$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}E(X_iY_j) = \frac{1}{n^2}\sum_{i=1}^{n}E(X_iY_i) \tag{19}$$

$$E(\hat{\sigma}_{X,Y}) = \frac{1}{n-1}\sum_{k=1}^{n}(E(X_kY_k)-E(\bar{X}Y_k)-E(X_k\bar{Y})+E(\bar{X}\bar{Y})) \tag{20}$$

$$= \frac{1}{n-1}\sum_{k=1}^{n}(E(X_kY_k)-\frac{1}{n}E(X_kY_k)-\frac{1}{n}E(X_kY_k)+\frac{1}{n^2}\sum_{i=1}^{n}E(X_iY_i)) \tag{21}$$

$$= \frac{1}{n-1}\left(\sum_{k=1}^{n}\frac{n-2}{n}E(X_kY_k)+\frac{n}{n^2}\sum_{i=1}^{n}E(X_iY_i)\right) \tag{22}$$

$$= \frac{1}{n-1}\left(\frac{n-2}{n}+\frac{1}{n}\right)\sum_{k=1}^{n}E(X_kY_k) \tag{23}$$

$$= \frac{1}{n}\sum_{k=1}^{n}E(X_kY_k) = \sigma_{X,Y} \tag{24}$$

よって、$\hat{\sigma}_{X,Y}$ は $\sigma_{X,Y}$ の不偏推定量である。

---

2. $X \sim Po(\lambda_1)$, $Y \sim Po(\lambda_2)$ で X と Y が独立とする $(\lambda_1, \lambda_2 > 0)$。その時、$U = X + Y$ の従う分布を求める。

…………………………………………………………………………

$$E(X) = V(X) = \lambda_1, \quad E(Y) = V(Y) = \lambda_2 \tag{25}$$

$$f_X(x) = P(X = x) = \frac{\lambda_1^x}{x!} \cdot e^{-\lambda_1}, \quad f_Y(y) = P(Y = y) = \frac{\lambda_2^y}{y!} \cdot e^{-\lambda_2} \tag{26}$$

$$\sum_{x=0}^{\infty} f_X(x) = 1, \quad \sum_{y=0}^{\infty} f_Y(y) = 1 \tag{27}$$

$$M_X(t) = E(e^{tX}) = \sum_{k=0}^{\infty} e^{tk} \frac{\lambda_1^k e^{-\lambda_1}}{k!} = e^{\lambda_1(e^t-1)} \tag{28}$$

$$M_Y(t) = E(e^{tY}) = \sum_{k=0}^{\infty} e^{tk} \frac{\lambda_2^k e^{-\lambda_2}}{k!} = e^{\lambda_2(e^t-1)} \tag{29}$$

(a) 積率母関数を用いて求めよ。

. . . .

$$\begin{aligned}
M_U(t) = E(e^{tU}) &= \sum_{k=0}^{\infty} \frac{t^k}{k!} E(U^k) = \sum_{k=0}^{\infty} \frac{t^k}{k!} E((X+Y)^k) &(30)\\
&= \sum_{k=0}^{\infty} \frac{t^k}{k!} E\left(\sum_{i=0}^{k} X^i Y^{k-i}\right) = \sum_{k=0}^{\infty}\sum_{i=0}^{k} \frac{t^k}{k!} E(X^i Y^{k-i}) &(31)\\
&= \left(\sum_{k=0}^{\infty} \frac{t^k}{k!} E(X^k)\right)\left(\sum_{l=0}^{\infty} \frac{t^l}{l!} E(Y^l)\right) &(32)\\
&= M_X(t) M_Y(t) = e^{\lambda_1(e^t-1)} e^{\lambda_2(e^t-1)} &(33)\\
&= e^{(\lambda_1+\lambda_2)(e^t-1)} &(34)
\end{aligned}$$

よって、$U = X + Y \sim Po(\lambda_1 + \lambda_2)$ となり、ポアソン分布に従うことが分かる。

(b) 畳み込みを用いて求めよ。

(HINT:$f_U(u) = \sum_{x=0}^{u} f_X(x) f_Y(u-x)$ を計算。f_X, f_Y は X, Y のそれぞれの確率質量関数)

. . . .

$$f_U(u) = \sum_{x=0}^{u} f_X(x) f_Y(u-x) = \sum_{x=0}^{u} \frac{\lambda_1^x e^{-\lambda_1}}{x!} \frac{\lambda_2^{u-x} e^{-\lambda_2}}{(u-x)!} \tag{35}$$

$$= e^{-\lambda_1-\lambda_2} \sum_{x=0}^{u} \frac{\lambda_1^x \lambda_2^{u-x}}{x!(u-x)!} = \frac{e^{-\lambda_1-\lambda_2}}{u!} \sum_{x=0}^{u} \frac{u!}{x!(u-x)!} \lambda_1^x \lambda_2^{u-x} \tag{36}$$

$$= \frac{e^{-(\lambda_1+\lambda_2)}}{u!} \sum_{x=0}^{u} {}_uC_x \lambda_1^x \lambda_2^{u-x} = \frac{e^{-(\lambda_1+\lambda_2)}}{u!} (\lambda_1+\lambda_2)^u \tag{37}$$

$$= \frac{(\lambda_1+\lambda_2)^u}{u!} e^{-(\lambda_1+\lambda_2)} \tag{38}$$

これにより、$U = X + Y \sim Po(\lambda_1 + \lambda_2)$ であることが分かる。

---

3. 確率変数 $X_1, \ldots, X_n$ が互いに独立にそれぞれ $N(\mu, \sigma^2)$ に従うとする。

(a) $Y_i = X_i - \bar{X}$ $(i = 1, \ldots, n)$ が従う分布を求めよ。
(HINT: 正規分布の再生性の性質を使う)

……………………………………………………………………

各 i について、$X_i \sim N(\mu, \sigma^2)$ である。
正規分布の再生性により $X_i + X_j$ も正規分布である。よって、$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ も正規分布である。
つまり、$Y_i = X_i - \bar{X}$ も正規分布である。

(b) Y_i と $\bar{X}$ の相関係数を求め Y_i と $\bar{X}$ の関係を考察せよ。

……………………………………………………………………

$$E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \mu \tag{39}$$

$$E(Y_i) = E(X_i - \bar{X}) = E(X_i) - E(\bar{X}) = 0 \tag{40}$$

$$V(\bar{X}) = E((\bar{X})^2) - (E(\bar{X}))^2 \tag{41}$$

$$V(Y_i) = E(Y_i^2) - (E(Y_i))^2 = E((X_i - \bar{X})^2) \tag{42}$$

$$= E(X_i^2 - 2X_i\bar{X} + \bar{X}^2) \tag{43}$$

$$= E(X_i^2) - 2E(X_i\bar{X}) + E(\bar{X}^2) \tag{44}$$

$V(X_i) = E(X_i^2) - (E(X_i))^2$ より $E(X_i^2) = \sigma^2 + \mu^2$ である。また、$i \neq j$ において X_i と X_j は独立であるので、$E(X_iX_j) = E(X_i)E(X_j)$ である。これ

らを用いて $E(X_i\bar{X})$ と $E(\bar{X}^2)$ を計算する。

$$
\begin{aligned}
E(X_i\bar{X}) &= E\left(X_i\frac{1}{n}\sum_{j=1}^{n}X_j\right) = \frac{1}{n}\sum_{j=1}^{n}E\left(X_iX_j\right) && (45)\\
&= (n-1)\mu^2+\sigma^2+\mu^2 = n\mu^2+\sigma^2 && (46)\\
E(\bar{X}^2) &= E\left(\left(\sum_{j=1}^{n}X_j\right)\left(\sum_{i=1}^{n}X_i\right)\right) = \sum_{j=1}^{n}\sum_{i=1}^{n}E(X_jX_i) && (47)\\
&= (n^2-n)\mu^2+n(\mu^2+\sigma^2) = n^2\mu^2+n\sigma^2 && (48)
\end{aligned}
$$

これらを用いて $V(Y_i), V(\bar{X})$ を求める。

$$
\begin{aligned}
V(Y_i) &= E(X_i^2)-2E(X_i\bar{X})+E(\bar{X}^2) && (49)\\
&= (\sigma^2+\mu^2)-2(n\mu^2+\sigma^2)+(n^2\mu^2+n\sigma^2) && (50)\\
&= (n-1)\sigma^2+(n-1)^2\mu^2 && (51)\\
V(\bar{X}) &= E((\bar{X})^2)-(E(\bar{X}))^2 = (n^2\mu^2+n\sigma^2)-\mu^2 && (52)
\end{aligned}
$$

共分散を求める。

$$
Cov(Y_i,\bar{X}) = E(Y_i\bar{X})-E(Y_i)E(\bar{X}) \qquad (53)
$$

このため、期待値 $E(Y_i\bar{X})$ を計算する。

$$
\begin{aligned}
E(Y_i\bar{X}) &= E\left((X_i-\bar{X})\bar{X}\right) = E(X_i\bar{X})-E(\bar{X}^2) && (54)\\
&= (n\mu^2+\sigma^2)-(n^2\mu^2+n\sigma^2) = n(1-n)\mu^2+(1-n)\sigma^2 && (55)
\end{aligned}
$$

よって、共分散は次のようになる。

$$
Cov(Y_i,\bar{X}) = n(1-n)\mu^2+(1-n)\sigma^2 \qquad (56)
$$

ここから相関係数を求める。

$$
\begin{aligned}
&\frac{Cov(Y_i,\bar{X})}{\sqrt{V(Y_i)}\sqrt{V(\bar{X})}} && (57)\\
&= \frac{n(1-n)\mu^2+(1-n)\sigma^2}{\sqrt{(n-1)\sigma^2+(n-1)^2\mu^2}\sqrt{(n^2-1)\mu^2+n\sigma^2}} && (58)\\
&= \frac{(1/n-1)\mu^2+(1/n^2-1/n)\sigma^2}{\sqrt{(1/n-1/n^2)\sigma^2+(1-1/n)^2\mu^2}\sqrt{(1-1/n^2)\mu^2+(1/n)\sigma^2}} && (59)\\
&\to 1(n\to\infty) && (60)
\end{aligned}
$$

このため、確率変数を多く取ると Y_i と $\bar{X}$ は正比例する。

4. 平均 μ_1、分散 $\sigma_1^2(>0)$ の分布に従う確率変数 X と、平均 μ_2、分散 $\sigma_2^2(>1)$ の分布に従う確率変数 Y とおく。

(a) $Y = (\sigma_2/\sigma_1)(X-\mu_1)+\mu_2$ とかける時、X と Y の相関を求めよ。($\sigma_i = \sqrt{\sigma_i^2} \quad i=1,2$)

………………………………………………………………………

$$\sigma_1^2 = E(X^2)-\mu_1^2, \quad \sigma_2^2 = E(Y^2)-\mu_2^2 \tag{61}$$
$$Cov(X,Y) = E(XY)-\mu_1\mu_2 \tag{62}$$

相関係数

$$\frac{Cov(X,Y)}{\sqrt{V(X)}\sqrt{V(Y)}} = \frac{E(XY)-\mu_1\mu_2}{\sigma_1\sigma_2} \tag{63}$$

$$E(XY) = E\left(X\left(\frac{\sigma_2}{\sigma_1}(X-\mu_1)+\mu_2\right)\right) \tag{64}$$
$$= E\left(\frac{\sigma_2}{\sigma_1}X^2 - \frac{\sigma_2}{\sigma_1}\mu_1 X + \mu_2 X\right) \tag{65}$$
$$= \frac{\sigma_2}{\sigma_1}E(X^2) - \frac{\sigma_2}{\sigma_1}\mu_1 E(X) + \mu_2 E(X) \tag{66}$$
$$= \frac{\sigma_2}{\sigma_1}(\sigma_1^2+\mu_1^2) - \frac{\sigma_2}{\sigma_1}\mu_1^2 + \mu_1\mu_2 \tag{67}$$
$$= \sigma_1\sigma_2 + \mu_1\mu_2 \tag{68}$$

$$\frac{Cov(X,Y)}{\sqrt{V(X)}\sqrt{V(Y)}} = 1 \tag{69}$$

これにより相関係数は 1 であることがわかる。
つまり、X と Y は正比例する。

(b) X と独立な平均 0、分散 1 の確率変数 Z を用いて、$Y = a(X-\mu_1)+\mu_2+Z$ と書ける時、X と Y の相関と $a(\geq 0)$ の値を求めよ。

………………………………………………………………………

$$
\begin{aligned}
E(XY) &= E\left(X(a(X-\mu_1)+\mu_2+Z)\right) && (70)\\
&= E\left(aX^2 - a\mu_1 X + \mu_2 X + XZ\right) && (71)\\
&= E(aX^2) - E(a\mu_1 X) + E(\mu_2 X) + E(XZ) && (72)\\
&= aE(X^2) - a\mu_1^2 + \mu_1\mu_2 + E(X)E(Z) && (73)\\
&= a(\sigma_1^2+\mu_1^2) - a\mu_1^2 + \mu_1\mu_2 && (74)\\
&= a\sigma_1^2 + \mu_1\mu_2 && (75)
\end{aligned}
$$

$$
\begin{aligned}
\frac{Cov(X,Y)}{\sqrt{V(X)}\sqrt{V(Y)}} &= \frac{E(XY)-\mu_1\mu_2}{\sigma_1\sigma_2} && (76)\\
&= \frac{a\sigma_1^2+\mu_1\mu_2-\mu_1\mu_2}{\sigma_1\sigma_2} && (77)\\
&= \frac{\sigma_1}{\sigma_2}a && (78)
\end{aligned}
$$

相関係数は -1 から 1 までの値をとるので、$-1 \leq \frac{\sigma_1}{\sigma_2}a \leq 1$ である。ここから、$-\frac{\sigma_2}{\sigma_1} \leq a \leq \frac{\sigma_2}{\sigma_1}$ がわかる。
$a \geq 0$ より、$0 \leq a \leq \frac{\sigma_2}{\sigma_1}$ である。
a が 0 に近くなれば X と Y は無相関となり、a が $\frac{\sigma_2}{\sigma_1}$ に近くなれば正の相関となる。