

確率・統計 A

若木宏文

wakaki@hiroshima-u.ac.jp

<http://home.hiroshima-u.ac.jp/wakaki/lecture/index.shtml>

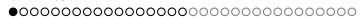
2018.7.17

目次

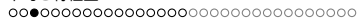
平均と特性量

特性量

条件付き分布と平均



3.3. 特性量



メモ用紙

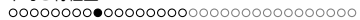


定理 3.7 (分散の性質)

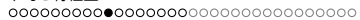
確率変数 X の分散に対して, 次が成り立つ. ただし, a は定数である.

- (1) $\text{Var}(X + a) = \text{Var}(X)$.
- (2) $\text{Var}(aX) = a^2\text{Var}(X)$.
- (3) $\text{Var}(X) = E(X^2) - E(X)^2$.

証明: 分散の定義と平均の線形性を用いる.



メモ用紙



メモ用紙

共分散と相関係数

定義 3.5

確率変数 X, Y に対して

$$\text{Cov}(X, Y) = E[\{X - E(X)\}\{Y - E(Y)\}]$$

を X と Y との**共分散**とよぶ。また

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}}$$

を X と Y との**相関係数**とよぶ。

$\rho(X, Y) = 0$ であるとき**無相関**という。

$\rho(X, Y)$ が正であるとき, X と Y の間には**正の相関**があるといい,

負であるときは**負の相関**があるという。

分散・共分散の基本性質

定理 3.9

X, Y, Z を確率変数, a を定数とする. このとき, 次が成り立つ.

- (1) $\text{Cov}(X, X) = \text{Var}(X)$.
- (2) $\text{Cov}(X, Y) = \text{Cov}(Y, X)$.
- (3) $\text{Cov}(X + Y, Z) = \text{Cov}(X, Z) + \text{Cov}(Y, Z)$.
- (4) $\text{Cov}(aX, Y) = a\text{Cov}(X, Y)$.
- (5) $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$.
- (6) X と Y が独立 $\Rightarrow \text{Cov}(X, Y) = 0$.
- (7) X と Y が有限な分散をもつとする. このとき

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y).$$

証明：分散の定義と期待値の基本性質から容易

シュワルツの不等式

定理 3.10

確率変数 X, Y は有限な分散をもつとする。このとき

$$\{\text{Cov}(X, Y)\}^2 \leq \text{Var}(X)\text{Var}(Y)$$

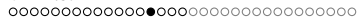
が成り立つ。したがって

$$|\rho(X, Y)| \leq 1.$$

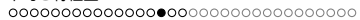
等号が成り立つのは、次の (1), (2), (3) のいずれかの場合である:

- (1) $P(X = E(X)) = 1.$
- (2) $P(Y = E(Y)) = 1.$
- (3) 定数 a, b が存在して, $P(Y = a + bX) = 1.$

(証明は板書で)



メモ用紙



メモ用紙

定義 3.6 (平均ベクトル, 共分散行列)

p 次元確率ベクトル $\mathbf{X} = (X_1, \dots, X_p)'$ について

$$E(X_i) = \mu_i, \quad E\{(X_i - \mu_i)(X_j - \mu_j)\} = \sigma_{ij}, \quad i, j = 1, \dots, p$$

とする. このとき

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_p \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1p} \\ \vdots & \ddots & \vdots \\ \sigma_{p1} & \cdots & \sigma_{pp} \end{pmatrix}$$

をそれぞれ \mathbf{X} の平均ベクトル, 共分散行列とよぶ.

共分散行列の性質

注 3.8

確率行列 $X = (X_{ij})$ の平均を $E(X) = (E(X_{ij}))$ と定義する. このとき

$$E(\mathbf{X}) = \boldsymbol{\mu}, \quad \text{Cov}(\mathbf{X}) = \Sigma = E\{(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})'\}$$

と表せる. また, 任意の実ベクトル $\mathbf{a} = (a_1, \dots, a_p)'$ に対して

$$\sum_{i,j=1}^p \sigma_{ij} a_i a_j = \text{Var}(a_1 X_1 + \dots + a_p X_p) \geq 0$$

となる. したがって, 共分散行列 Σ は半正定値行列である.



3.4. 条件付き分布と平均

事象を与えたときの条件付き分布

$X : (\Omega, \mathcal{B}, P)$ 上の確率変数, $B \in \mathcal{B}$, $P(B) > 0$

$A \in \mathbb{B}_1$

事象 B を与えたときの, 事象 “ $X \in A$ ” の条件付き確率 :

$$P(X \in A | B) = \frac{P("X \in A" \cap B)}{P(B)}$$

B を固定して, $P_{X|B}(A) = P(X \in A | B)$ を A の関数とみると,

$P_{X|B}$ は $(\mathbb{R}^1, \mathbb{B}_1)$ 上の確率となる.

$P_{X|B}$ を B を与えたときの X の条件付き分布とよぶ.

離散型分布の場合

(X, Y) : 2次元離散型確率変数,

$D = \{(x_i, y_j); i, j = 1, 2, \dots\}$: 取り得る値の集合

$f_{X,Y}(x, y)$: (X, Y) の確率関数, $f_Y(y)$: Y の周辺確率関数

事象 " $Y = y_j$ " を与えたときの, 事象 " $X \in A$ " の条件付き確率は

$$\begin{aligned} P_{X|Y=y_j}(A) &= P(X \in A | Y = y_j) \\ &= \frac{P(X \in A, Y = y_j)}{P(Y = y_j)} = \sum_{x_i \in A} \frac{f_{X,Y}(x_i, y_j)}{f_Y(y_j)} \end{aligned}$$

条件つき分布（離散型の場合）

$P_{X|Y=y_j}$ は $(\mathbb{R}^1, \mathbb{B}_1)$ 上の離散型分布であり、その確率関数は

$$f_{X|Y}(x|y_j) \equiv \frac{f_{X,Y}(x_i, y_j)}{f_Y(y_j)}$$

によって与えられる.

$P_{X|Y=y_j}$ を $Y = y_j$ を与えたときの X の条件付き分布とよぶ.

$f_{X|Y}(x|y_j)$ を $Y = y_j$ を与えたときの X の条件付き確率関数とよぶ.

条件付き期待値 (離散型の場合)

(X, Y) は離散型で, $f_{X|Y}(x_i|y_j)$ を $Y = y_j$ を与えたときの X の条件付き確率関数とする.

$$\sum_{i=1}^{\infty} |x_i| f_{X|Y}(x_i|y_j) < \infty$$

であるとき, $Y = y_j$ を与えたときの X の条件付き平均は存在するといふ,

$$E(X | Y = y_j) \equiv \sum_{i=1}^{\infty} x_i f_{X|Y}(x_i|y_j)$$

を $Y = y_j$ を与えたときの X の条件付き平均, あるいは条件付き期待値とよぶ.

例 3.10 (3 項分布)

$(X, Y) \sim M_3(n, (p, q))$ (3 項分布), $p > 0, q > 0, 1 - p - q > 0$.

$$\begin{aligned} f_{X|Y}(i|j) &\equiv \mathbf{P}(X = i|Y = j) \\ &= \frac{\frac{n!}{i!j!(n-i-j)!} p^i q^j (1-q-p)^{n-j-i}}{\frac{n!}{j!(n-j)!} q^j (1-q)^{n-j}} \\ &= \frac{(n-j)!}{i!(n-j-i)!} \left(\frac{p}{1-q}\right)^i \left(1 - \frac{p}{1-q}\right)^{n-j-i}. \end{aligned}$$

$Y = j$ を与えたときの X の条件付き分布は
2 項分布 $B(n-j, p/(1-q))$,

$$\mathbf{E}(X|Y = j) = (n-j) \frac{p}{1-q}.$$

連続型の場合

(X, Y) : 2次元連続型確率変数

$f_{X,Y}(x, y)$: (X, Y) の確率密度関数, $f_Y(y)$: Y の周辺確率密度関数

$f_Y(y) > 0$ である y に対して

$$f_{X|Y}(x|y) \equiv \frac{f_{X,Y}(x, y)}{f_Y(y)}$$

は、確率密度関数の性質

$$(1) f_{X,Y}(x|y) \geq 0 (\forall x \in \mathbb{R}), (2) \int_{-\infty}^{\infty} f_{X,Y}(x|y) dx = 1$$

を満たす.

$f_{X|Y}(x|y)$ を $Y = y$ を与えたときの X の条件付き確率密度関数,

$f_{X|Y}(x|y)$ によって定められる連続型分布を

$f_{X|Y}(x|y_j)$ を $Y = y_j$ を与えたときの X の条件付き分布とよぶ.

注 3.11

$f_{X,Y}(x, y)$, $f_Y(y)$ が連続であれば

$$\begin{aligned} \lim_{h \rightarrow 0+} \mathbf{P}(X \leq x | y < Y \leq y + h) &= \lim_{h \rightarrow 0+} \frac{\int_{-\infty}^x \left\{ \int_y^{y+h} f_{X,Y}(u, v) dv \right\} du}{\int_y^{y+h} f_Y(v) dv} \\ &= \lim_{h \rightarrow 0+} \frac{\int_{-\infty}^x f_{X,Y}(u, y + \theta_1 h) h du}{f_Y(y + \theta_2 h) h} = \int_{-\infty}^x f_{X|Y}(u|y) du \end{aligned}$$

が成り立つ。ただし, $0 \leq \theta_i \leq 1$, $i = 1, 2$.

条件付き期待値 (連続型の場合)

(X, Y) は連続型で, $f_{X|Y}(x|y)$ を $Y = y$ を与えたときの X の条件付き確率密度関数とする.

$$\int_{-\infty}^{\infty} |x| f_{X|Y}(x|y) dx < \infty$$

であるとき, $Y = y$ を与えたときの X の条件付き平均は存在する
といい,

$$E(X | Y = y) \equiv \int_{-\infty}^{\infty} x f_{X|Y}(x|y) dx$$

を $Y = y$ を与えたときの X の条件付き平均, あるいは条件付き期待値とよぶ.

例 3.11 (2次元正規分布)

 $(X_1, X_2)' \sim N_2(\boldsymbol{\mu}, \Sigma)$ (2次元正規分布),

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

 $X_2 = x_2$ を与えたときの X_1 の条件付き分布は

$$N\left(\mu_1 + \rho \frac{\sigma_1}{\sigma_2}(x_2 - \mu_2), (1 - \rho^2)\sigma_1^2\right)$$

条件付き平均は

$$E(X_1 | X_2 = x_2) = \mu_1 + \rho \frac{\sigma_1}{\sigma_2}(x_2 - \mu_2).$$

(証明は板書で)



メモ用紙



メモ用紙



メモ用紙

条件付き平均の性質

定理 3.12

$f_{X,Y}(x,y)$: (X,Y) の同時確率 (密度) 関数

$f_{X|Y}(x|y)$: $Y = y$ を与えたときの X の条件付き確率 (密度) 関数

$$E(g(X)|Y = y) = \begin{cases} \sum_{j=1}^{\infty} g(x_j) f_{X|Y}(x_j|y), & \text{離散型のとき} \\ \int_{-\infty}^{\infty} g(x) f_{X|Y}(x|y) dx, & \text{連続型のとき} \end{cases}$$

次が成り立つ.

$g(y) = E(X|Y = y)$ とするとき, $g(Y)$ は確率変数となるが, これを $E(X|Y)$ と表す.

定理 3.13

$f_{X,Y}(x, y) : (X, Y)$ の同時確率密度関数 (確率関数) このとき

$$E\{E(X|Y)\} = E(X).$$

が成り立つ.

(証明は板書で)



メモ用紙