

確率統計 A

第 3 章 平均と特性量 (Part 3)

(2012 年 6 月 26 日)

- 講師: 若木宏文
- E-mail: wakaki@math.sci.hiroshima-u.ac.jp
- Web: <http://home.hiroshima-u.ac.jp/~wakaki/lecture/probstatA12/>

確率統計 A, 2012年6月26日: P.1

平均ベクトル&共分散行列

定義 3.6. p 次元確率ベクトル $\mathbf{X}=(X_1, \dots, X_p)'$ について

$E(X_i)=\mu_i, E\{(X_i-\mu_i)(X_j-\mu_j)\}=\sigma_{ij} (i, j=1, \dots, p),$
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, 共分散行列と呼ぶ。

確率統計 A, 2012年6月26日: P.2

Example 2

Example 3.6.

1. 3項分布, $\mathbf{X} \sim M_3(n, \mathbf{p})$ の場合:

$$E(\mathbf{X}) = n\mathbf{p} = \begin{pmatrix} np_1 \\ np_2 \end{pmatrix}, \quad \text{Cov}(\mathbf{X}) = \begin{pmatrix} np_1(1-p_1) & -np_1p_2 \\ -np_1p_2 & np_2(1-p_2) \end{pmatrix}$$

2. 2次元正規分布, $\mathbf{X} \sim N_2(\boldsymbol{\mu}, \Sigma)$ の場合:

$$E(\mathbf{X}) = \boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \text{Cov}(\mathbf{X}) = \Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}$$

確率統計 A, 2012年6月26日: P.3

定理3.11

2次元確率ベクトル $\mathbf{X}=(X_1, X_2)'$ の平均ベクトルと共分散行列を

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

とする。ただし $\sigma_1 > 0, \sigma_2 > 0$ とする。このとき,

$$\min_{\alpha, \beta} E\{(X_2 - \alpha - \beta X_1)^2\} = E\{(X_2 - \hat{\alpha} - \hat{\beta} X_1)^2\} \\ = \sigma_2^2(1 - \rho^2)$$

となる。ここに $\hat{\alpha} = \mu_2 - \hat{\beta}\mu_1, \hat{\beta} = \rho\sigma_2 / \sigma_1$.

確率統計 A, 2012年6月26日: P.4

定理3.11の証明

確率統計 A, 2012年6月26日: P5

注 3.9

確率変数 X_2 を X_1 の1次結合 $\alpha + \beta X_1$ で近似するときの誤差を $E[\{X_2 - (\alpha + \beta X_1)\}^2]$ で測るとき、これを最小にする最良近似

$$\hat{\alpha} + \hat{\beta} X_2 = \mu_2 + \rho(\sigma_2 / \sigma_1)(X_1 - \mu_1)$$

の誤差は $\sigma_2^2(1 - \rho^2)$ である。 $|\rho|$ が大きくなると誤差は小さくなる。

確率統計 A, 2012年6月26日: P6

3.4. 条件付き分布と平均

確率統計 A, 2012年6月26日: P7

条件付き分布

X を確率変数とし、 $B \in \mathcal{B}$ は $P(B) > 0$ であるとする。
 $A \in \mathcal{B}$ に対して、事象 B を与えたときの事象
“ $X \in A$ ” の条件付き確率は、

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

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 $(\mathbb{R}, \mathcal{B}_1)$ 上の確率測度になる。これを **B を与えたときの X の条件付き分布** と呼ぶ。

確率統計 A, 2012年6月26日: P8

離散型の場合

(X, Y) を2次元の離散型確率変数とし, 同時確率関数を $f_{X,Y}(x,y)$, Y の周辺確率関数を $f_Y(y)$ とする.

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

を $Y=y$ を与えたときの X の条件付き確率関数と呼ぶ(条件付き確率関数で決まる確率分布を**条件付き分布**と呼ぶ). 条件付き確率関数は以下の性質を満たす.

[1]. $f_{X|Y}(x_j | y) \geq 0, j = 1, 2, \dots$ [2]. $\sum_{j \geq 1} f_{X|Y}(x_j | y) = 1$

確率統計 A, 2012年6月26日: P.9

連続型の場合

(X, Y) を2次元の連続型確率変数とし, 同時確率密度関数を $f_{X,Y}(x,y)$, Y の周辺確率密度関数を $f_Y(y)$ とする.

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

を $Y=y$ を与えたときの X の条件付き確率密度関数と呼ぶ(条件付き確率密度関数で決まる確率分布を**条件付き分布**と呼ぶ). 条件付き確率密度関数は以下の性質を満たす.

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

確率統計 A, 2012年6月26日: P.10

条件付き平均

条件付き分布での平均を**条件付き平均**と呼ぶ.
 $Y=y$ を与えたときの X の条件付き平均は

$$E(X | y) = \begin{cases} \sum_{i=1}^{\infty} x_i f_{X|Y}(x_i | y) & (\text{離散型の場合}) \\ \int_{-\infty}^{\infty} x f_{X|Y}(x | y) dx & (\text{連続型の場合}) \end{cases}$$

で定義される. また通常の期待値と同様に, 離散型の場合, 右辺の級数が絶対収束する, 連続型の場合, 右辺が絶対積分可能であるとき, 条件付き平均は存在するという.

確率統計 A, 2012年6月26日: P.11

性質 (1/2)

定理3.12. (X, Y) の同時確率関数(または確率密度関数)を $f_{X,Y}(x,y)$ とし, $Y=y$ を与えたときの X の条件付き確率関数(または確率密度関数)を $f_{X|Y}(x|y)$ とする. このとき次が成り立つ.

$$E\{g(X) | y\} = \begin{cases} \sum_{i=1}^{\infty} g(x_i) f_{X|Y}(x_i | y) & (\text{離散型の場合}) \\ \int_{-\infty}^{\infty} g(x) f_{X|Y}(x | y) dx & (\text{連続型の場合}) \end{cases}$$

注) 証明は定理3.1と同様して得られる.

確率統計 A, 2012年6月26日: P.12

性質 (2/2)

定理3.13. (X, Y) は同時確率関数(または確率密度関数) $f_{X,Y}(x,y)$ を持つとする. このとき,

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

連続型での証明

$$\begin{aligned} E\{E(X | Y)\} &= \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} x f_{X|Y}(x, y) dx \right\} f_Y(y) dy \\ &= \int_{-\infty}^{\infty} x \left\{ \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy \right\} dx \\ &= \int_{-\infty}^{\infty} x f_X(x) dx = E(X) \end{aligned}$$

確率統計 A, 2012年6月26日: P.13

Example

(X, Y) が2次元正規分布 $N_2(\mu, \Sigma)$ に従うとする. このとき,

$$\begin{aligned} f_{X,Y}(x, y) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \\ &\times \exp \left[-\frac{1}{2(1-\rho^2)} \left\{ \left(\frac{x-\mu_1}{\sigma_1} \right)^2 - 2\rho \left(\frac{x-\mu_1}{\sigma_1} \right) \left(\frac{y-\mu_2}{\sigma_2} \right) + \left(\frac{y-\mu_2}{\sigma_2} \right)^2 \right\} \right] \\ &= f_{X|Y}(x | y) \times f_Y(y) \end{aligned}$$

と表される.

確率統計 A, 2012年6月26日: P.14

Example (続き)

ここで,

$$\begin{aligned} f_{X|Y}(x | y) &= \frac{1}{\sqrt{2\pi(1-\rho^2)\sigma_1^2}} \exp \left[-\frac{1}{2(1-\rho^2)\sigma_1^2} \left\{ x - \mu_1 - \rho \frac{\sigma_1}{\sigma_2} (y - \mu_2) \right\}^2 \right] \\ f_Y(y) &= \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp \left[-\frac{1}{2\sigma_2^2} (y - \mu_2)^2 \right] \end{aligned}$$

したがって $Y=y$ が与えられたときの X の条件付き分布は

$$N(\mu_1 - \rho\sigma_1(y - \mu_2) / \sigma_2, (1 - \rho^2)\sigma_1^2)$$

$$E\{E(X | Y)\} = E\{\mu_1 - \rho\sigma_1(Y - \mu_2)\} = \mu_1 = E(X)$$

確率統計 A, 2012年6月26日: P.15