確率統計 A

第 3 章 平均と特性量 (Part 3)

(2012 年 6 月 12 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/

3.3. 特性量

モーメントの定義

定義 3.3.

1. 原点のまわりの k 次モーメント: $\mu_k = \mathrm{E}(X^k)$.
2. 原点のまわりの k 次絶対モーメント: $\nu_k = \mathrm{E}(|X|^k)$.
3. 平均値のまわりの k 次モーメント:
$$\alpha_k = \mathrm{E}[\{X - \mathrm{E}(X)\}^k] = \mathrm{E}\{(X-\mu)^k\} \ (\mu = \mu_1).$$

モーメントの存在

定理 3.6. 確率変数 X は有限な n 次モーメントをもつとする. このとき, X の n 次以下のモーメントも有限である. すなわち
$$\mathrm{E}(|X|^n) < \infty \Rightarrow \mathrm{E}(|X|^m) < \infty \ (m = 0, 1, \ldots, n).$$

(Proof) 定理 3.4 (3) および不等式 $|X|^m \le |X|^n + 1$ より得られる.

分散, 歪度, 尖度

定義 3.4.

1. **分散 (Variance)**: 分布のばらつきを表す尺度.
   $\sigma^2 = \mathrm{Var}(X) = \mathrm{E}\{(X-\mu)^2\} = \alpha_2$.
2. 歪度 (Skewness): 分布の歪みを表す尺度.
   $\beta_1 = \alpha_3/\sigma^3 = \mathrm{E}[\{(X-\mu)/\sigma\}^3]$.
   特に正規分布の場合は $\beta_1 = 0$.
3. 尖度 (Kurtosis): 分布の裾の広さを表す尺度.
   $\beta_2 = \alpha_4/\sigma^4 = \mathrm{E}[\{(X-\mu)/\sigma\}^4]$.
   特に正規分布の場合は $\beta_2 = 3$. そのため, 尖度 $\beta_2 = \alpha_4/\sigma^4 - 3$ と定義することもある.

Example 1

Example 3.5.

1. 正規分布, $X \sim N(\mu, \sigma^2)$ のとき:
   $\mathrm{E}(X) = \mu$, $\mathrm{Var}(X) = \sigma^2$, $\beta_1 = 0$, $\beta_2 = 3$.
2. 2 項分布, $X \sim B(n, p)$ のとき:
   $\mathrm{E}(X) = np$, $\mathrm{Var}(X) = npq$, ($q = 1 - p$),
   $\beta_1 = -(p-q)/(npq)^{1/2}$,
   $\beta_2 = (1 - 6pq + 3npq)/npq$.

分散に関する定理

定理 3.7. 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. $\mathrm{Var}(X+a)=\mathrm{Var}(X)$,
2. $\mathrm{Var}(aX)=a^2\mathrm{Var}(X)$,
3. $\mathrm{Var}(X)=\mathrm{E}(X^2)-\mathrm{E}(X)^2$.

定理3.7の証明

モーメントに関する不等式

定理 3.8.

1. (**Markov**) X を非負値確率変数とする. 任意の正数 a に対して

$$\mathrm{P}(X\geq a)\leq \mathrm{E}(X)/a.$$

注 3.6) $\mathrm{E}(X)=0$ のとき, $\mathrm{P}(X=0)=1$.

2. (**Chevyshev**) 確率変数 X の平均, 分散を

$$\mathrm{E}(X)=\mu,\ \mathrm{Var}(X)=\sigma^2<\infty$$

とする. 任意の正数 k に対して

$$\mathrm{P}(|X-\mu|\geq k)\leq \sigma^2/k^2.$$

定理3.8の証明

共分散&相関係数

定義 3.5. 確率変数X, Yに対して

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

をXとYの**共分散** (**Covariance**) と呼ぶ. また

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

をXとYの**相関係数** (**Correlation Coefficient**) と呼ぶ. $\rho(X,Y)=0$のとき無相関, また$\rho(X,Y)$が正のとき**正の相関**があると言い, 負のときは**負の相関**があると言う.

注 3.7

定理3.10で示されるように, $|\rho(X,Y)|\leq 1$である. また定理3.11より$|\rho(X,Y)|$の大きさはXとYの線形関係の強さを表している. そのため$|\rho(X,Y)|$が1に近いときは**強い相関**があると言い, 0に近いときは**弱い相関**があると言う. 確率変数XとYの相関係数$\rho(X,Y)$は, それらの関連の強さを測る尺度であるが, 直線的関連の度合いを測る尺度であることに注意したい.

共分散&相関係数に関する定理 (1/2)

定理 3.9. X, Y, Zを確率変数, aを定数とする. このとき以下が成り立つ.

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

定理3.9の証明

共分散&相関係数に関する定理 (2/2)

定理 3.10. 確率変数X, Yは有限な分散を持つとする. このとき

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

が成り立つ. したがって,

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

等号が成り立つのは以下の1, 2, 3のいずれかである.

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

定理3.10の証明

平均ベクトル&共分散行列

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

$$\mathrm{E}(X_i)=\mu_i,\ \mathrm{E}\{(X_i-\mu_i)(X_j-\mu_j)\}=\sigma_{ij}\ (i,j=1,\ldots,n),$$

とする. このとき

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

をそれぞれXの平均ベクトル, 共分散行列と呼ぶ.

注 3.8

$\mathrm{E}(\boldsymbol{X})=\boldsymbol{\mu}$とおくと, 共分散行列は

$$\mathrm{Cov}(\boldsymbol{X})=\Sigma=\mathrm{E}\{(\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{X}-\boldsymbol{\mu})'\}$$

と表せる. このことより, 共分散行列は半正定値行列であることがわかる.

Example 2

Example 3.6.

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

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

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

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

定理3.11

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

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

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

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

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

定理3.11の証明

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$\mathrm{E}[\{X_2-(\alpha+\beta X_1)\}^2]$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.