■復習

- $\boldsymbol{X} = (X_1, \cdots, X_p),\ Y = g(\boldsymbol{X})$

  $$\mathrm{E}[g(\boldsymbol{X})] = \mathrm{E}[Y] = \begin{cases} \int y f_Y(y) dy \\ \sum y_j f_Y(y_j) \end{cases}$$

  $$= \begin{cases} \int g(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) dx_1 \cdots dx_n \\ \sum g(\boldsymbol{x}_i) f_{\boldsymbol{X}}(\boldsymbol{x}_i) \end{cases}$$

- $g(\boldsymbol{X})$ の確率密度関数の導出定理 3.23
- k 次モーメント：$\mu_k = E[X^k]$

  k 次絶対モーメント：$\nu_k = E[|X|^k]$

  平均の周りの k 次モーメント：$\alpha_k = E[(X - \mu_1)^k]$
- 分散、歪度、尖度

■今日の内容

-
  - 不等式（Markov, Chevyshev, Schwarz）
  - 平均の線形性
  - 積の期待値と独立性の関係
  - 共分散とその性質
  - 相関係数とその性質、シュワルツの不等式
  - 独立和の平均と分散
- 3.4 条件付き分布の平均
  - 事象が与えられたときの条件付き分布
  - $Y = y$ が与えられたときの X の条件付き分布
  - 条件付き平均と性質

■不等式

- $X \geq 0$ のとき $\mathrm{E}[X] \geq 0$

  $\because \mathrm{E}[X] = \int x f_X(x) dx \geq 0$

- $X \geq Y$ のとき $\mathrm{E}[X] \geq \mathrm{E}[Y]$

$$\begin{aligned}
\because \mathrm{E}[X] - \mathrm{E}[Y] &= \int x f_X(x) dx - \int y f_Y(y) dx \\
&= \iint x f_{X,Y}(x,y) dy dx - \iint y f_{X,Y}(x,y) dx dy \\
&= \iint (x-y) f_{X,Y}(x,y) dx dy \geq 0
\end{aligned}$$

定理 3.30

(1) X: 非負の値をとる確率変数, $a > 0$: 定数

$$\Rightarrow \quad \mathrm{P}(X \geq a) \leq \mathrm{E}[X]/a$$

$\because$ $Y = \begin{cases} 0 & (X < a) \\ a & (X \geq a) \end{cases}$ とすると $\mathrm{E}[X] \geq \mathrm{E}[Y] = a\mathrm{P}(X \geq a)$

(2) $\mu = \mathrm{E}[X],\ \sigma^2 = \mathrm{Var}(X) < \infty$ とすると

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

$\because |X - \mu|^2$ に対して (1) を用いる。

■平均の線形性　X, Y ： 確率変数、a, b ： 定数

$$
\begin{aligned}
E[aX + bY] &= \iint (ax + by) f_{X,Y}(x, y) dxdy \\
&= a \iint x f_{X,Y}(x, y) dxdy + b \iint y f_{X,Y}(x, y) dxdy \\
&= a \int x f_X(x) dx + b \int y f_Y(y) dy
\end{aligned}
$$

■積の期待値と独立性　X, Y ： 独立な確率変数

$$
\begin{aligned}
\mathrm{E}[g(X)h(Y)] &= \iint g(x)h(y) f_{X,Y}(x, y) dxdy \\
&= \iint g(x)h(y) f_X(x) f_Y(y) dxdy \\
&= \int g(x) f_X(x) dx \int {\color{red}h}(y) f_Y(y) dy = \mathrm{E}[g(X)]\mathrm{E}[h(Y)]
\end{aligned}
$$

■共分散とその性質

定義 3.13 X, Y : 確率変数

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

$$\text{相関係数：}\ \rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$$

性質

- 平均の線形性から

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

- X, Y が独立のとき、

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

- シュワルツの不等式

$$\mathrm{E}[X^2]\mathrm{E}[Y^2] \leq (\mathrm{E}[XY])^2$$

$\because$ X, Y を有限な2次モーメントを持つ確率変数、k を定数とする。

$0 \leq \mathrm{E}[(kX+Y)^2] = k^2\mathrm{E}[X^2] + 2k\mathrm{E}[XY] + \mathrm{E}[Y^2]$

k の2次式が常に非負だから判別式：

$(\mathrm{E}[XY])^2 - \mathrm{E}[X^2]\mathrm{E}[Y^2] \leq 0$

X, Y を $(X - \mathrm{E}[X])$, $(Y - \mathrm{E}[Y])$ で置き換えると

$$-1 \leq \rho(X, Y) \leq 1$$

を得る

■独立和の平均と分散

$X_1, \cdots, X_n$：独立な確率変数, $a_1, \cdots, a_n$：定数

(1) $$\mathrm{E}\Big[\sum_{j=1}^{n} a_j X_j\Big] = \sum_{j=1}^{n} a_j \mathrm{E}[X_j]$$

(2) $$\mathrm{Var}\Big[\sum_{j=1}^{n} a_j X_j\Big] = \sum_{j=1}^{n} a_j^2 \mathrm{Var}[X_j]$$

$$\begin{aligned}
\mathrm{Var}\Big[\sum_{j=1}^{n} a_j X_j\Big] &= \mathrm{E}\Big[\Big(\sum_{j=1}^{n} a_j (X_j - \mathrm{E}[X_j])\Big)^2\Big] \\
&= \sum_{j=1}^{n}\sum_{k=1}^{n} \mathrm{E}[(X_j - \mathrm{E}[X_j])(X_k - \mathrm{E}[X_k])]
\end{aligned}$$

条件付き分布と平均

■事象が与えられたときの条件付き分布

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

B が与えられたときの X の条件付き分布

$$\mathrm{P}_{X|B}(A) = \mathrm{P}(X^{-1}(A)|B)$$

$P_{X|B}(*)$ は $(\boldsymbol{R}^1, \mathcal{B}_1)$ 上の確率測度

■$Y = y$ が与えられたときの X の条件付き分布

離散型の場合

$f_{X,Y}(x, y)$：確率密度関数

$D_X = \{x_1, x_2, \cdots\}$：$X$ のとり得る値

$D_Y = \{y_1, y_2, \cdots\}$：$Y$ のとり得る値

$y \in D_Y$ のとき、事象 $Y = y$ が与えられたときの

$X \in A$ の条件付き確率は

$$\mathrm{P}(X \in A | Y = y) = \frac{\mathrm{P}(X \in A, Y = y)}{\mathrm{P}(Y = y)} = \frac{\sum_{x_i \in A} f_{X,Y}(x_i, y)}{f_Y(y)}$$

$Y = y$ が与えられたときの X の条件付き確率 (密度) 関数は

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

連続型の場合

$f_{X,Y}(x,y)$：確率密度関数

$Y=y$ が与えられたときの X 条件付き分布関数を

$$\lim_{h\to 0+0} \mathrm{P}(X \le x | y < Y \le y+h)$$

$$= \lim_{h\to 0+} \frac{\int_{-\infty}^{x} \left\{ \int_{y}^{y+h} f_{X,Y}(u,v)dv \right\} du}{\int_{y}^{y+h} f_Y(v)dv}$$

$$= \lim_{h\to 0+} \frac{\int_{-\infty}^{x} f_{X,Y}(u, y+\theta_1 h) h du}{f_Y(y+\theta_2 h)h} = \int_{-\infty}^{x} f_{X|Y}(u|y)du,$$

$f_{X|Y}$ は $Y=y$ が与えられたときの X の条件付き確率密度関数：

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

■条件付き平均と性質

離散型、連続型のいずれの場合でも

条件付き確率密度関数 $f_{X|Y}$ は確率密度関数の性質を満たす：

(1) $f_{X|Y}(x, y) \geq 0$,

(2) $\int f_{X|Y}(x|y)dx = 1$ または $\sum_{i=1}^{\infty} f_{X|Y}(x_i|y) = 1$

条件付き平均

$Y = y$ が与えられたときの $g(X)$ の条件付き平均

$$\mathrm{E}[g(X)|Y = y] = \begin{cases} \int g(x) f_{X|Y}(x, y)dx & \text{(連続型)} \\ \sum_{i=1}^{\infty} g(x_i) f_{X|Y}(x_i|y) & \text{(離散型)} \end{cases}$$

性質

$h(y) = \mathrm{E}[g(X)|Y = y]$ とし、$\mathrm{E}[g(X)|Y]$ は、$h(Y)$ を表すとする。

定理 3.35

$$\mathrm{E}_Y[\mathrm{E}[g(X)|Y]] = \mathrm{E}_{X,Y}[g(X)]$$

証明（連続型の場合）

$$\begin{aligned}
\mathrm{E}_Y[\mathrm{E}(g(X)|Y)] &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} g(x) f_{X|Y}(x|y)dy\right] f_y(y)dy \\
&= \int_{-\infty}^{\infty} g(x) \left[\int_{-\infty}^{\infty} f_{X,Y}(x,y)dy\right] dx \\
&= \int_{-\infty}^{\infty} g(x) f_X(x)dx = \mathrm{E}(g(X)).
\end{aligned}$$