条件付確率と独立性

確率・統計 - 第 5 講

村田 昇

前回のおさらい

- 確率変数
  - 確率空間上の関数としての確率変数
  - 確率変数の確率空間
  - 複数の確率変数の同時確率 (結合確率)
  - 確率変数の平均，分散，モーメント
  - **確率変数の特性関数**

特性関数

- 定義

  特性関数は指数関数を用いて変数変換された確率変数の平均として

  $$\phi(s) = \mathbb{E}[e^{isX}]$$

  で定義される．
  - 特性関数は実数値ではなく複素数値をとることに注意

特性関数の性質

- 逆変換の存在

  特性関数と確率測度 (確率密度) は一対一に対応している
  (Levy-Haviland の反転公式)

- 分布の同一性

  2 つの分布の特性関数が同じであれば，2 つの分布が同じであることが言える．

演習

練習問題

- 以下の確率測度の特性関数を求めよ．
  1. 標準正規分布
  2. 平均 μ 分散 σ^2 の正規分布

解答例

- X: 標準正規分布に従う確率変数
- 定義に従って計算

$$
\begin{aligned}
\phi_X(s) &= \mathbb{E}[e^{isX}] \\
&= \int_{-\infty}^{\infty} e^{isx} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}+isx} dx
\end{aligned}
$$

- 2 次式を整理

$$
\begin{aligned}
-\frac{x^2}{2} + isx &= -\frac{x^2 - 2isx}{2} \\
&= -\frac{x^2 - 2isx + (is)^2 - (is)^2}{2} \\
&= -\frac{(x - is)^2}{2} - \frac{s^2}{2}
\end{aligned}
$$

- 正規分布の密度関数の積分 (=1) に帰着

$$
\begin{aligned}
\phi_X(s) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-is)^2}{2}-\frac{s^2}{2}} dx \\
&= e^{-\frac{s^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-is)^2}{2}} dx \\
&= e^{-\frac{s^2}{2}}
\end{aligned}
$$

  - 厳密には複素積分の基本的性質 (Cauchy の積分定理) を使う
- Y: 平均 μ 分散 σ^2 の正規分布に従う確率変数
- 前問と同様に定義に従って計算

$$
\begin{aligned}
\phi_Y(s) &= \mathbb{E}[e^{isY}] \\
&= \int_{-\infty}^{\infty} e^{isy} \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy
\end{aligned}
$$

- $Y = \sigma X + \mu$ の関係を利用

$$
\begin{aligned}
\phi_Y(s) &= \mathbb{E}[e^{isY}] = \mathbb{E}[e^{is(\sigma X+\mu)}] \\
&= e^{\mu is} \cdot \mathbb{E}[e^{i(\sigma s)X)}] = e^{\mu is} \cdot \phi_X(\sigma s) \\
&= e^{\mu is} e^{-\frac{\sigma^2 s^2}{2}} = e^{\mu is-\frac{\sigma^2 s^2}{2}}
\end{aligned}
$$

条件付確率

条件付確率

- 定義

  確率空間 $(\Omega, \mathcal{F}, P)$ において 2 つの事象 $A, B \subset \Omega$ を考える．事象 A のもとにおける事象 B の**条件付確率** (conditional probability) を

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

  で定義する．

同時確率と条件付確率

- **同時確率** $P(A \cap B)$

  “事象 A と事象 B が同時に起きている” 確率

- **条件付確率** $P(B|A)$

  “事象 A が起きたときに事象 B が起きている” 確率

条件付確率に関する注意

- 事象 A は空でない $A \neq \emptyset$ とする
- $P(A) = 0$ のときは $P(B|A)$ は意味がないが，適当に一点 $\omega_0 \in A$ を選び

  $$P(B|A) = \begin{cases} 1, & \omega_0 \in B \\ 0, & \omega_0 \notin B \end{cases}$$

  とすることがある

条件付確率の例

- 骰子を一回振る試行

  以下の事象を定義する．

  $$A = \{\text{偶数の目が出る}\}$$
  $$B = \{\text{素数の目が出る}\}$$

  このとき以下の確率は条件付確率で表される．
  - “偶数の目が出たとき，それが素数である” 確率
  - “偶数の目が出たとき，それが素数でない” 確率

- “偶数の目が出たとき，それが素数である” 確率

  $$\begin{aligned} P(B|A) &= \frac{P(A \cap B)}{P(A)} \\ &= \frac{P(\text{偶数かつ素数の目が出る})}{P(\text{偶数の目が出る})} \\ &= \frac{P\{2\}}{P(\{2,4,6\})} = \frac{1/6}{1/2} = \frac{1}{3} \end{aligned}$$

- “偶数の目が出たとき，それが素数でない” 確率

$$
\begin{aligned}
P(B^c|A) &= \frac{P(A \cap B^c)}{P(A)} \\
&= \frac{P(\text{偶数かつ素数でない目が出る})}{P(\text{偶数の目が出る})} \\
&= \frac{2/6}{1/2} = \frac{2}{3} = 1 - P(B|A)
\end{aligned}
$$

確率変数による条件の表現

- 確率変数が一つの値を取る場合

$$
\begin{aligned}
A &= \{\omega | X(\omega) = x\} \Leftrightarrow A = X^{-1}(x)\ (x \in \Omega^X) \\
B &= \{\omega | Y(\omega) = y\} \Leftrightarrow B = Y^{-1}(y)\ (y \in \Omega^Y)
\end{aligned}
$$

 のようにある事象に対応づけられるので,

$$P(Y{=}y|X{=}x) \quad \text{または} \quad P_{X=x}(Y{=}y)$$

 という条件付確率を考えることができる.

- 確率変数がある集合に含まれる場合
 同様に

$$
\begin{aligned}
A &= \{\omega | X(\omega) \in E\} \Leftrightarrow A = X^{-1}(E)\ (E \subset \Omega^X) \\
B &= \{\omega | Y(\omega) \in F\} \Leftrightarrow B = Y^{-1}(F)\ (F \subset \Omega^Y)
\end{aligned}
$$

 と事象と対応づけられるので,

$$P(Y \in F|X \in E) \quad \text{または} \quad P_{X\in E}(Y \in F)$$

 のように表した条件付確率を考えることができる.

- 同時確率の表現
 単に変数を並べて

$$
\begin{aligned}
P(X{=}x, Y{=}y) &= P(\{\omega | X(\omega) = x\} \cap \{\omega | Y(\omega) = y\}) \\
P(X \in E, Y \in F) &= P(\{\omega | X(\omega) \in E\} \cap \{\omega | Y(\omega) \in F\})
\end{aligned}
$$

 と表すことが多い.

- 確率法則 (確率分布) の表現
 確率変数 X,Y のいろいろな値に対する同時確率や条件付確率をまとめて表現する場合

$$P(X, Y), \quad P(Y|X)$$

 といった記述を行うことが多い.

条件付確率の性質

- 定理
 条件付確率には次の性質がある.
 1. $P(A \cap B) = P(A)P(B|A)$
 2. $P(A|A) = 1$
 3. $P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2|A_1)\cdots P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})$
 4. $P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$
 - 証明は定義から明らか

条件付確率分布

- 定義

  定理の 2. より $P(\cdot|A)$ は A を標本空間とする確率法則になっている．これを A のもとにおける**条件付確率分布** (conditional probability distribution) という．

条件付平均値

- 定義

  条件付確率分布のもとでの平均値

  $$\mathbb{E}[Y|A] = \sum_{\omega \in A} Y(\omega) P(\omega|A) \quad \text{(離散分布の場合)}$$
  $$\mathbb{E}[Y|A] = \int_A Y(\omega) P(d\omega|A) \quad \text{(連続分布の場合)}$$

  を**条件付平均値** (conditional expectation) と呼ぶ．

Bayes の定理

Bayes の定理 (基本形)

- 定理

  条件付確率では次の等式が成り立つ．

  $$P(A|B) = \frac{P(A)P(B|A)}{P(B)}.$$

  - 前出の定理の 4. より導かれる
  - 左辺と右辺で事象 A, B の役割が異なる

Bayes の定理 (一般形)

- 定理

  $\Omega = A_1 + A_2 + \cdots + A_n$ のとき

  $$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{k=1}^n P(A_k)P(B|A_k)}$$

  が成り立つ．

演習

練習問題

- Bayes の定理 (一般形) を証明せよ．

解答例

- 証明

  事象 B と A_i との関係は全事象を利用すると

  $$P(B) = P(\Omega \cap B) = P(\sum_{k=1}^n A_k \cap B) = \sum_{k=1}^n P(A_k \cap B)$$

  と分解できる．すなわち $\Omega = A_1 + A_2 + \cdots + A_n$ のとき $A_k \cap B$ は互いに排反な事象となる．これと定理を用いれば分母の表現が得られる．

独立性

確率変数の独立性

- 定義 (有限試行の場合)

  二つの確率変数 X, Y が互いに**独立** (independent) であるとは，任意の x, y に対して

  $$P(X=x, Y=y) = P(X=x)P(Y=y)$$

  となることである．

- 条件付確率による意味付け

  条件付確率を用いて同時確率を書き直すと

  $$P(X=x, Y=y) = P(X=x)P(Y=y|X=x)$$

  であるが，独立性の定義から

  $$P(Y=y|X=x) = P(Y=y)$$

  となり，Y の確率は X に左右されないことを示す．

- 定義 (無限試行の場合)

  任意の事象 (集合) $E \subset \Omega^X, F \subset \Omega^Y$ に対して

  $$P(X \in E, Y \in F) = P(X \in E)P(Y \in F)$$

  が成り立つことである．

- 定義 (確率密度を持つ場合)

  X, Y の同時確率分布の密度関数 (**同時確率密度**; joint probability density) を $f(x, y)$，X, Y のそれぞれの確率分布の密度関数 (**周辺確率密度**; marginal probability density) を $g(x), h(y)$ とすると，確率変数 X, Y が互いに独立とは任意の x, y に対して

  $$f(x, y) = g(x)h(y)$$

  となることである．

- 特性関数による表現 (Kac の定理)

  確率変数 X, Y が互いに独立とは，特性関数を用いて

  $$\mathbb{E}[e^{isX+itY}] = \mathbb{E}[e^{isX}]\mathbb{E}[e^{itY}]$$

  が成り立つことと同値である．

独立性の性質

- 定理

  独立性に関しては以下の性質が成り立つ．

  - 確率変数 X, Y が独立 $\Leftrightarrow P(Y|X) = P(Y)$ a.s.
    (周辺分布 $P(Y)$ と条件付分布 $P(Y|X)$ が確率 1 で等しい)
  - 確率変数 $X_1, X_2, \ldots, X_n$ が独立 $\Rightarrow$ 変数変換した確率変数

    $$Y_1 = \phi_1(X_1), Y_2 = \phi_2(X_2), \ldots, Y_n = \phi_n(X_n)$$

    は独立 (逆が成り立つとは限らないことに注意)

- (定理のつづき)
    - 確率変数 $X_1, X_2, \ldots, X_n$ が独立 $\Rightarrow$

      $$\mathbb{E}[X_1 X_2 \cdots X_n] = \mathbb{E}[X_1]\mathbb{E}[X_2] \cdots \mathbb{E}[X_n]$$

      (平均値の乗法性)
    - 確率変数 $X_1, X_2, \ldots, X_n$ の任意の 2 変数が独立 $\Rightarrow$

      $$\mathrm{Var}(X_1 + X_2 + \cdots + X_n) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \cdots + \mathrm{Var}(X_n)$$

      (分散の加法性)
  - 平均値の加法性は独立とは無関係に成り立つ
  - 平均値の乗法性は独立性の定義から明らか
  - 分散の加法性は展開して平均値の乗法性を用いれば明らか

事象の独立性

- 指示関数

  集合 A の **指示関数 (定義関数)** を

  $$1_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases}$$

  で定義する．
  - 指示関数は実数値関数
  - $Y = 1_A(\omega)$ 確率変数
    (標本点 ω が事象 A に含まれたかどうかを表す)
- 定義

  事象 $A_1, A_2, \ldots, A_n$ のそれぞれの指示関数で定義される確率変数 $Y_1, Y_2, \ldots, Y_n$ が互いに独立である場合，事象 $A_1, A_2, \ldots, A_n$ は独立であるという．
  - 確率変数の独立性の定義に帰着することに注意

事象の独立性の同値条件

- 定理

  以下の 3 つは同値である．
  1. 事象 $A_1, A_2, \ldots, A_n$ が独立
  2. $B_i = A_i$ または $B_i = A_i^c$ としたとき

     $$P(B_1 \cap B_2 \cap \cdots \cap B_n) = P(B_1)P(B_2) \cdots P(B_n)$$

  3. $\forall k = 2, 3, \ldots, n,\ 1 \leq \forall i_1 < \forall i_2 < \cdots \forall i_k \leq n$

     $$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_k})$$

演習

練習問題

- 以下の問に答えよ．

  高田馬場のラーメン屋 A の売上は，前の週の口コミサイト B と C における評価に強く影響される．詳しく調べた結果以下のことが判った．

  - 売上，および評価は上がるか下がるかの二通りしかない
  - B,C 共に評価が上がったときに売上が上がる確率は 0.9 である
  - B,C の一方のみ評価が上がったときに売上が上がる確率は 0.6 である
  - B,C 共に評価が下がったときに売上が下がる確率は 0.8 である
  - B の評価が上がる確率は 0.5 である
  - C の評価が上がる確率は 0.6 である
  - B の評価が上がるとき $X = 1$, 下がるとき $X = 0$, C の評価が上がるとき $Y = 1$, 下がるとき $Y = 0$ となる確率変数 X, Y を考えると，X, Y の共分散は 0.1 である

- このとき以下の確率を求めよ．
  - B,C 共に評価が上がる確率はいくつか？
  - A の売上が下がったとき，B,C 共に評価が下がった確率はいくつか？
  - A の売上が上がったとき，B,C の一方のみ評価が上がった確率はいくつか？

ヒント

- 扱う事象を整理する

$$
\begin{aligned}
A &= \{\text{A の売上が上がる}\} & A^c &= \{\text{A の売上が下がる}\}\\
B &= \{\text{B の評価が上がる}\} & B^c &= \{\text{B の評価が下がる}\}\\
C &= \{\text{C の評価が上がる}\} & C^c &= \{\text{C の評価が下がる}\}
\end{aligned}
$$

- 前提条件を書き下す

$$
\begin{aligned}
&P(A|B \cap C) = 0.9 & &P(A^c|B \cap C) = 0.1\\
&P(A|B \cap C^c) = P(A|B^c \cap C) = 0.6 & &P(A^c|\cdot) = 0.4\\
&P(A|B^c \cap C^c) = 0.2 & &P(A^c|B^c \cap C^c) = 0.8\\
&P(B) = 0.5 & &P(B^c) = 0.5\\
&P(C) = 0.6 & &P(C^c) = 0.4\\
&\mathrm{Cov}(X, Y) = 0.1
\end{aligned}
$$

- 求める確率は以下のとおり

$$
\begin{aligned}
&P(B \cap C)\\
&P(B^c \cap C^c|A^c) = P(A^c \cap B^c \cap C^c)/P(A^c)\\
&P(B \cap C^c|A) + P(B^c \cap C|A)\\
&\qquad = P(A \cap B \cap C^c)/P(A) + P(A \cap B^c \cap C)/P(A)
\end{aligned}
$$

  - 第 1 問は共分散の計算を利用すれば求められる

$$
\begin{aligned}
\mathrm{Cov}(X, Y) &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \\
&= P(X = 1, Y = 1) - P(X = 1)P(Y = 1)
\end{aligned}
$$

- 第 1 問の結果を利用すると前提条件から $P(B \cap C^c), P(B^c \cap C), P(B^c \cap C^c)$ も求められる
- 第 2,3 問は必要な確率を整理して Bayes の定理を用いればよい

解答例

- 第 1 問

$$
\begin{aligned}
\mathrm{Var}(X, Y) &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \\
&= P(B \cap C) - P(B)P(C) \\
&= P(B \cap C) - 0.5 \times 0.6 = 0.1 \\
P(B \cap C) &= 0.4
\end{aligned}
$$

これより

$$
\begin{aligned}
P(B \cap C^c) &= P(B) - P(B \cap C) = 0.1 \\
P(B^c \cap C) &= P(C) - P(B \cap C) = 0.2 \\
P(B^c \cap C^c) &= 0.3
\end{aligned}
$$

- 第 2 問

$$
\begin{aligned}
P(B^c \cap C^c | A^c) &= P(A^c \cap B^c \cap C^c)/P(A^c) \\
&= \frac{P(A^c | B^c \cap C^c) P(B^c \cap C^c)}{\sum_\lambda P(A^c | B^\lambda \cap C^\lambda) P(B^\lambda \cap C^\lambda)} \\
&= 0.6
\end{aligned}
$$

- 第 3 問

$$
\begin{aligned}
&P(B \cap C^c | A) + P(B^c \cap C | A) \\
&\quad = P(A \cap B \cap C^c)/P(A) + P(A \cap B^c \cap C)/P(A) \\
&\quad = 0.3
\end{aligned}
$$

今回のまとめ

- 条件付確率と独立性
  - 同時確率，周辺確率，条件付確率
  - 条件付確率分布，条件付確率密度，条件付平均値
  - 条件付確率と Beyes の定理
  - 確率変数と事象の独立性