確率変数の独立, 独立同分布

樋口さぶろお https://hig3.net

龍谷大学 先端理工学部 数理・情報科学課程

確率統計 I L08(2023-05-29 Mon)

最終更新: Time-stamp: "2023-05-29 Mon 11:26 JST hig"

今日の目標

- 確率変数の独立性を判定し利用できる 岩薩林 確率・統計 §3.3
- 独立同分布の和の母平均値/母分散を計算できる

岩薩林 確率・統計 p.113, 定理 5.2

L07-Q1

Quiz 解答: 条件付き分布

1. $p_{\mathrm{Y}|\mathrm{X}}(0|9) = \dfrac{p(9,0)}{p_{\mathrm{X}}(9)} = \dfrac{\frac{1}{6}}{\frac{1}{6}+\frac{1}{3}} = \dfrac{1}{3}.$
2. $p_{\mathrm{X}|\mathrm{Y}}(9|0) = \dfrac{p(9,0)}{p_{\mathrm{Y}}(0)} = \dfrac{\frac{1}{6}}{\frac{1}{6}+\frac{1}{3}+\frac{1}{6}} = \dfrac{1}{4}.$

L07-Q2,Q4

Quiz 解答: 同時分布・条件付き分布・周辺分布

$y \backslash x$	2	3	
20	3/12	1/12	4/12
30	4/12	4/12	8/12
	7/12	5/12	1

L07-Q3

L07-Q5

Quiz 解答: ベイズ推定

Y を色, X を当たり外れとすると,

$$
\begin{aligned}
&P(X=\text{当}|Y=\text{赤})\\
=&\frac{P(Y=\text{赤}|X=\text{当})P(X=\text{当})}{P(Y=\text{赤}|X=\text{当})P(X=\text{当})+P(Y=\text{赤}|X=\text{外})P(X=\text{外})}\\
=&\frac{\frac{1}{10}\frac{2}{10}}{\frac{1}{10}\frac{2}{10}+\frac{7}{10}\frac{8}{10}}=\frac{2}{58}.
\end{aligned}
$$

$Y\backslash X$	当たり	外れ
赤	$\frac{1}{10}\cdot\frac{2}{10}$	$\frac{7}{10}\cdot\frac{8}{10}$
白	$\frac{9}{10}\cdot\frac{2}{10}$	$\frac{3}{10}\cdot\frac{8}{10}$
合計	$\frac{2}{10}$	$\frac{8}{10}$

L07-Q6

Quiz 解答: ベイズ推定

検査 $Y=\begin{cases}80 & (\text{陽性})\\ 20 & (\text{陰性})\end{cases}$, 病状 $X=\begin{cases}100 & (\text{病気})\\ 0 & (\text{病気でない})\end{cases}$ で表す.

1. $$p_{\mathrm{X}}(100) = 0.005, p_{\mathrm{Y}|\mathrm{X}}(80|100) = 0.99, p_{\mathrm{Y}|\mathrm{X}}(80|0) = 0.02$$

同時確率分布

$x \backslash y$	80	20	計
100	真陽性 0.99×0.005	偽陰性 0.01×0.005	0.005
0	偽陽性 0.02×0.995	真陰性 0.98×0.995	0.995

$$\begin{aligned} p_{\mathrm{X}|\mathrm{Y}}(100|80) &= \frac{p_{\mathrm{Y}|\mathrm{X}}(80|100) \times p_{\mathrm{X}}(100)}{p_{\mathrm{Y}|\mathrm{X}}(80|100) \times p_{\mathrm{X}}(100) + p_{\mathrm{Y}|\mathrm{X}}(80|0) \times p_{\mathrm{X}}(0)} \\ &= \frac{0.99 \times 0.005}{0.99 \times 0.005 + 0.02 \times 0.995} = 0.199185. \end{aligned}$$

2. $0.99^2 \cdot 0.005/(0.99^2 \cdot 0.005 + 0.02^2 \cdot 0.995) = 0.924884.$

ここまで来たよ

7 条件付き確率とベイズの公式

8 確率変数の独立, 独立同分布
- 確率変数の独立性
- 独立同分布

確率変数が「同じ」対「同分布」

2 つの確率変数 X, Y.
X, Y は同じ　定義:$X = Y$, 各試行で同じ値.

- X, Y が同じ, ならば, 同分布 (次)
- 例: コイン 1 枚 (表 (1) の確率 $0 < p < 1$) を 1 回投げた結果に, 2 つの変数名 X, Y をつけた
- もちろん周辺分布は等しい
  $\mathrm{E}[g(X)] = \mathrm{E}[g(Y)]$.

$y \backslash x$	0	1
0	$1-p$	0
1	0	p

X, Y は同分布　定義: 周辺分布が等しい $p_{\mathrm{X}}(a) = p_{\mathrm{Y}}(a)$.

- 「同じ」とはかぎらない
- $\mathrm{E}[g(X)] = \mathrm{E}[g(Y)]$.

$y \backslash x$	0	1
0	$1-\frac{3}{2}p$	$\frac{1}{2}p$
1	$\frac{1}{2}p$	$\frac{1}{2}p$

X, Y は独立同分布

例: 形の同じ青色コイン X(表 p) と赤色コイン Y(表 q). $p = q$

$y \backslash x$	0	1	
0	$(1-p)(1-q)$	$p(1-q)$	$1-q$
1	$(1-p)q$	pq	q
	$1-p$	p	1

例

$y \backslash x$	0	1	
0	4/9	2/9	2/3
1	2/9	1/9	1/3
	2/3	1/3	1

独立: この形 ($p \neq q$ も可) $\rightsquigarrow$ カモ

同分布: $p = q$

- X の確率が Y によらない. Y の確率が X によらない.
  - 条件付き分布 $p_{\mathrm{Y}|\mathrm{X}}(y|x) = p_{\mathrm{Y}}(y)$. x の条件によらない.
  - 条件付き分布 $p_{\mathrm{X}|\mathrm{Y}}(x|y) = p_{\mathrm{X}}(x)$. y の条件によらない.

確率変数の独立性 高校 数学 B

定義 (事象の独立性 岩薩林 確率・統計 (2.14)p.40)

事象 A, B が独立 (independent) とは次が成立すること.

$$P(A \cap B) = P(A) \times P(B)$$

定義 (確率変数の独立性 岩薩林 確率・統計 (3.14)p.58)

確率変数 X, Y が独立 (independent) とは次が成立すること.

$$p(x, y) = p_{\mathrm{X}}(x) \times p_{\mathrm{Y}}(y)$$

X, Y が独立とは

- すべての行 (周辺分布も) がベクトルとして平行. すべての列 (周辺分布も) がベクトルとして平行
- 条件付き確率が周辺分布に等しい
  $p_{\mathrm{X|Y}}(x|\text{何でも}) = p_{\mathrm{X}}(x)$, $p_{\mathrm{Y|X}}(y|\text{何でも}) = p_{\mathrm{Y}}(y)$.
- ベイズ推定は役立たない. 事前確率=事後確率.
- X, Y が互いに「無関係」であること
- 同時分布が, 周辺分布だけから (積で) 決まっちゃうこと

独立である例

例

$y\backslash x$	0	1
10	$4/9$	$2/9$
11	$2/9$	$1/9$

例

赤白のサイコロを各 1 個振って, $X=$ 赤の目, $Y=$ 白の目.

例

下のカードから引いて, $X=$ 色, $Y=$ 数.

$\heartsuit 8 \quad \diamondsuit 8 \quad \diamondsuit 9$

$\spadesuit 8 \quad \clubsuit 8 \quad \clubsuit 9$

独立でない例

サイコロで, 目が 1–4 なら $X = 0$, 目が 5,6 なら $X = 1$.

例 1 $X = Y$. 確率変数 X, Y は同じ.

$y \backslash x$	0	1
0	2/3	0
1	0	1/3

例 2 例 1 の X, 同じサイコロの目で，1–2 なら $Y = 1$, 3–6 なら $Y = 0$.

$y \backslash x$	0	1
0	1/3	1/3
1	1/3	0

命題 (独立と限らない 2 次元の確率分布の母期待値の性質)

高校 数学 B | 岩薩林 確率・統計 (3.17)p.61 | 確率統計 I(2023)L06

$$\mathrm{E}[g_1(X,Y)+g_2(X,Y)] = \mathrm{E}[g_1(X,Y)]+\mathrm{E}[g_2(X,Y)] \quad \text{(E5)}$$

$$\text{特に}\mathrm{E}[X+Y] = \mathrm{E}[X]+\mathrm{E}[Y] \quad \text{(E6,(3.17))}$$

$$\mathrm{E}[g_1(X)\times g_2(Y)] \neq \mathrm{E}[g_1(X)]\times\mathrm{E}[g_2(Y)] \quad \text{(8.1)}$$

$$\text{母共分散}\mathrm{Cov}[X,Y] \overset{\text{定義}}{=} \mathrm{E}[(X-\mu_{\mathrm{X}})(Y-\mu_{\mathrm{Y}})] \quad \text{(3.19)}$$

$$= \mathrm{E}[XY]-\mathrm{E}[X]\times\mathrm{E}[Y]. \quad \text{(C1,(3.20))}$$

L08-Q1

Quiz(2 次元の独立でない確率変数の母期待値)

確率変数 X, Y について, $\mathrm{E}[X] = 3, \mathrm{E}[X^2] = 16, \mathrm{E}[Y] = 10, \mathrm{E}[Y^2] = 102, \mathrm{E}[XY] = 25$ が成立する. 次の量を求めよう.

1. $\mathrm{E}[(X+2)(Y+3)]$
2. $\mathrm{Cov}[X, Y]$

命題 (X, Y が独立のとき成立するいい性質)

岩薩林 確率・統計 (3.18)p.61

$$\mathrm{E}[g_1(X) \times g_2(Y)] = \mathrm{E}[g_1(X)] \times \mathrm{E}[g_2(Y)] \quad \text{(IE1)}$$

$$\text{特に}\mathrm{E}[XY] = \mathrm{E}[X] \times \mathrm{E}[Y] \quad \text{(IE2,(3.18))}$$

$$0 = \mathrm{Cov}[X, Y] = \mathrm{E}[XY] - \mathrm{E}[X] \times \mathrm{E}[Y], \quad \text{(IC1,(3.21))}$$

$$\mathrm{V}[g_1(X) + g_2(Y)] = \mathrm{V}[g_1(X)] + \mathrm{V}[g_2(Y)] \quad \text{(IC2)}$$

岩薩林 確率・統計 (3.21)p.63

$$\text{特に}\mathrm{V}[X + Y] = \mathrm{V}[X] + \mathrm{V}[Y] \quad \text{(IC3,(3.21))}$$

岩薩林 確率・統計 (3.21)p.63

$$\mathrm{V}[aX + bY + c] = a^2\mathrm{V}[X] + b^2\mathrm{V}[Y] \quad \text{(IC4)}$$

Cov は「独立でない度」, ただし,
IC1, IC3,IC4 は, 独立でなくても $\mathrm{Cov}[X, Y] = 0$ だけで成り立つ.
$\mathrm{Cov}[X, Y] = 0$ は独立の 必要条件 (だが 十分条件 ではない)

一部の証明連続型で書くけど, $\int \to \sum$ で離散型でも同様.

$$\begin{aligned}
\mathrm{E}[g_1(X)\cdot g_2(Y)] &= \iint g_1(x)g_2(y)\cdot f(x,y)\mathrm{d}y\,\mathrm{d}x \\
&\overset{\text{独立}}{=} \iint g_1(x)g_2(y)\cdot f_{\mathrm{X}}(x)\times f_{\mathrm{Y}}(y)\mathrm{d}y\,\mathrm{d}x \\
&= \int g_1(x)f_{\mathrm{X}}(x)\left(\int g_2(y)\cdot f_{\mathrm{Y}}(y)\mathrm{d}y\right)\mathrm{d}x \\
&= \mathrm{E}[g_1(X)]\times \mathrm{E}[g_2(Y)]
\end{aligned}$$

$$\begin{aligned}
\mathrm{V}[X+Y] &\overset{\text{V1}}{=} \mathrm{E}[(X+Y)^2] - \mathrm{E}[X+Y]^2 \\
&\overset{\text{E6}}{=} \mathrm{E}[X^2] + 2\mathrm{E}[XY] + \mathrm{E}[Y^2] - (\mathrm{E}[X]^2 + 2\mathrm{E}[X]\mathrm{E}[Y] + \mathrm{E}[Y]^2) \\
&\overset{\text{V1,C1}}{=} \mathrm{V}[X] + 2\mathrm{Cov}[X,Y] + \mathrm{V}[Y] \\
&\overset{\text{IC1}}{=} \mathrm{V}[X] + 0 + \mathrm{V}[Y]
\end{aligned}$$

L08-Q2

Quiz(独立な確率変数の母期待値)

独立な確率変数 X, Y を考える.
$\mathrm{E}[X]=2, \mathrm{E}[Y]=3, \mathrm{V}[X]=5, \mathrm{V}[Y]=11$ である.

1. 母期待値 $\mathrm{E}[XY]$ を求めよう.
2. 母共分散 $\mathrm{Cov}[X, Y]$ を求めよう.
3. $\mathrm{E}[(-2X+3Y)(X+5Y)]$ を求めよう.
4. $\mathrm{V}[-2X+3Y]$ を求めよう.

岩薩林 確率・統計 問題 6(p.64)
岩薩林 確率・統計 第 3 章練習問題 2(p.73)
岩薩林 確率・統計 第 3 章練習問題 5(p.73)

ここまで来たよ

7 条件付き確率とベイズの公式

8 確率変数の独立, 独立同分布
- 確率変数の独立性
- 独立同分布

独立同分布

定義 (独立同分布 (i.i.d.) 岩薩林 確率・統計 定理 4.2(p.87) の仮定,p.113)

離散型/連続型確率変数 $X_1, X_2, \ldots, X_n$ が, たがいに独立で, 同分布 (同じ周辺分布 $p_{\mathrm{X}_i}(x_i) = p(x)$ を持つとする.
このとき $X_1, \ldots, X_n$ は独立同分布に従う (i.i.d.=independent and identically-distributed) という.

母期待値も共通になる. $\mathrm{E}[X_i] = \mu, \mathrm{V}[X_i] = \sigma^2$.
例: 同じ形の n 色のサイコロ. X_1=青色, X_2=緑色, $\cdots$.

L08-Q3

Quiz(同分布の確率母期待値母分散)

X_1, X_2 は同分布にしたがう確率変数で，コインの表 (1) 裏 (0) を表す．確率関数は，

$$p_{X_i} x_i = \begin{cases} 1-p & (x=0) \\ p & (x=1) \end{cases},$$

母平均値・母分散は

$$\mathrm{E}[X_i^k] = 0^k \cdot (1-p) + 1^k \cdot p = p, \mathrm{V}[X_i] = \mathrm{E}[X_i^2] - \mathrm{E}[X_i]^2 = p - p^2.$$

1. $X_1 = X_2$ であるとき (1 枚のコインに 2 個の名前をつけたとき),
   1. 確率 $P(X_1 = X_2 = 1)$ を求めよう．
   2. 母平均値 $\mathrm{E}[X_1 + X_2]$ を求めよう．
   3. 母分散 $\mathrm{V}[X_1 + X_2]$ を求めよう．
   4. 母共分散 $\mathrm{Cov}[X_1 + X_2]$ を求めよう．
2. X_1, X_2 が独立であるとき (形は同じだが色の違う別々のコインであるとき),
   1. 確率 $P(X_1 = X_2 = 1)$ を求めよう．
   2. 母平均値 $\mathrm{E}[X_1 + X_2]$ を求めよう．
   3. 母分散 $\mathrm{V}[X_1 + X_2]$ を求めよう．
   4. 母共分散 $\mathrm{Cov}[X_1 + X_2]$ を求めよう．

L08-Q4

Quiz(独立同分布の母期待値母分散)

確率変数 X_1, X_2, X_3 は独立同分布に従い, $\mathrm{E}[X_i] = \mu, \mathrm{V}[X_i] = \sigma^2$ とする. 次の母期待値, 母分散を μ と σ^2 で表そう.

1. $\mathrm{E}[X_1 + X_1 + X_1]$
2. $\mathrm{E}[X_1 + X_2 + X_3]$
3. $\mathrm{E}[X_1 + X_1 + X_2]$
4. $\mathrm{V}[X_1 + X_1 + X_1]$
5. $\mathrm{V}[X_1 + X_2 + X_3]$
6. $\mathrm{V}[X_1 + X_1 + X_2]$
7. $\mathrm{E}[(X_1 + X_2 + X_3)^2]$