

独立同分布の和・二項分布

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(2024-06-10 Mon)

最終更新: Time-stamp: "2024-06-09 Sun 08:08 JST hig"

今日の目標

- 独立同分布にしたがう確率変数の和の母平均値, 母分散を求められる 岩薩林 確率・統計 §4.4
- 二項分布の母期待値, 母分散, 確率を求められる

岩薩林 確率・統計 §3.4



L07-Q1

Quiz 解答: ベイズの定理

甘い品種 1: $x = 1$, 渋い品種 2: $x = 2$, 赤: $y = 10$, 黄: $y = 20$ と表現する.

①

$$p_{Y|X}(y|1) = \begin{cases} 0.95 & (y = 10) \\ 0.05 & (y = 20) \end{cases}$$

$$p_{Y|X}(y|2) = \begin{cases} 0.125 & (y = 10) \\ 0.875 & (y = 20) \end{cases}$$

L07-Q2

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

$$p_X(100) = 0.005, p_{Y|X}(80|100) = 0.99, p_{Y|X}(80|0) = 0.02$$

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

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

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

L07-Q3

Quiz 解答:2次元の独立でない確率変数の母期待値

- ① $E[(X + 2)(Y + 3)] = E[XY + 3X + 2Y + 6] =$
 $E[XY] + 3E[X] + 2E[Y] + 6E[1] = 25 + 9 + 20 + 6 = 60.$
- ② $\text{Cov}[X, Y] = E[XY] - E[X]E[Y] = 25 - 3 \cdot 10 = -5.$
 $\text{Cov}[X, Y] = E[(X - \mu_X)(Y - \mu_Y)]$ を展開しても同じ結果.

L07-Q4

Quiz 解答: 独立な確率変数の母期待値

- ① X, Y は独立なので $E[XY] \stackrel{IE1}{=} E[X]E[Y] = 2 \cdot 3.$
- ② X, Y は独立なので $\text{Cov}[X, Y] \stackrel{IC1}{=} 0.$
- ③ $E[(-2X + 3Y)(X + 5Y)] \stackrel{E5}{=} E[-2X^2] + E[-7XY] + E[15Y^2] \stackrel{V1, IE1}{=} -2(V[X] + E[X]^2) - 7E[X]E[Y] + 15(V[Y] + E[Y]^2) = 240.$
- ④ X, Y は独立なので $V[g_1(X) + g_2(Y)] \stackrel{IC2}{=} V[g_1(X)] + V[g_2(Y)]$ であることに注意して,
 $V[-2X + 3Y] \stackrel{IC2}{=} V[-2X] + V[3Y] \stackrel{V2}{=} (-2)^2V[X] + 3^2V[Y] = 119.$

ここまで来たよ

7 バイズ推定・独立性・独立同分布

8 独立同分布の和・二項分布

- 独立同分布
- 二項分布
- ベルヌーイ分布と二項分布・独立同分布の性質の応用
- 独立同分布にしたがう確率変数の和

復習: 独立, 非独立な確率変数の母期待値と母分散

X_1, X_2 が独立でも独立でなくても成立すること

$$E[aX_1 + bX_2 + c] = aE[X_1] + bE[X_2] + c,$$

$$V[aX_1 + c] = a^2V[X_1],$$

$$V[aX_1 + bX_2 + c] = a^2V[X_1] + 2ab\text{Cov}[X_1, X_2] + b^2V[X_2],$$

$$\text{Cov}[X_1, X_2] = E[X_1X_2] - E[X_1]E[X_2].$$

X_1, X_2 が独立なとき成立すること

$$E[g_1(X_1) \times g_2(X_2)] = E[g_1(X_1)] \times E[g_2(X_2)],$$

$$\text{Cov}[X_1, X_2] = 0,$$

$$V[aX_1 + bX_2 + c] = a^2V[X_1] + b^2V[X_2].$$

独立同分布

$X = X_1, Y = X_2$ は独立同分布 定義: 独立かつ同分布

例: 形の同じ青色コイン X_1 と赤色コイン X_2 .
表がでる (1) の確率 p

$x_2 \backslash x_1$	0	1
0	$(1-p)^2$	$p(1-p)$
1	$p(1-p)$	p^2

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, \dots, X_n が、たがいに独立で、すべて同分布に従う (同じ周辺分布 $p_{X_i}(x_i) = p(x)$ を持つ) とする。

このとき X_1, \dots, X_n は独立同分布に従う (i.i.d.=independent and identically-distributed) という。

母期待値も等しい. $E[X_i] = \mu, V[X_i] = \sigma^2$.

例: 同じ形の n 色のサイコロ. X_1 =青色, X_2 =緑色, \dots .

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
 $E[g(X)] = E[g(Y)]$.

$y \backslash x$	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 $p_X(a) = p_Y(a)$.

- 「同じ」とはかぎらない
- $E[g(X)] = E[g(Y)]$.

$y \backslash x$	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

例: 形の同じ青色コイン X (表 p) と赤色コイン Y (表 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

- 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 $p_{Y|X}(y|x) = p_Y(y)$. x の条件によらない.
 - ▶ 条件付き分布 $p_{X|Y}(x|y) = p_X(x)$. y の条件によらない.

X, Y は独立同分布定義: 独立 (上の形) かつ同分布 ($p = q$)

L08-Q1

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

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

- 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 $E[X_1 + X_2]$ を求めよう。
 - ③ 母分散 $V[X_1 + X_2]$ を求めよう。
 - ④ 母共分散 $\text{Cov}[X_1, X_2]$ 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 $E[X_1 + X_2]$ を求めよう。
 - ③ 母分散 $V[X_1 + X_2]$ を求めよう。
 - ④ 母共分散 $\text{Cov}[X_1, X_2]$ を求めよう。

ここまで来たよ

7 バイズ推定・独立性・独立同分布

8 独立同分布の和・二項分布

- 独立同分布
- 二項分布
- ベルヌーイ分布と二項分布・独立同分布の性質の応用
- 独立同分布にしたがう確率変数の和

復習+ちょっと (x で書かれた確率関数) !

L08-Q2

Quiz(離散的な確率変数の母平均値・母分散・母標準偏差・確率)

整数に値をとる離散型確率変数 X は次の確率分布に従う.

$$p(x) = \begin{cases} \frac{x}{55} & (0 \leq x \leq 10) \\ 0 & (\text{他}) \end{cases}$$

- 1 確率 $P(X \leq 5)$ を求めよう.
- 2 母平均値 $E[X]$ を求めよう.
- 3 母分散 $V[X]$ を求めよう.

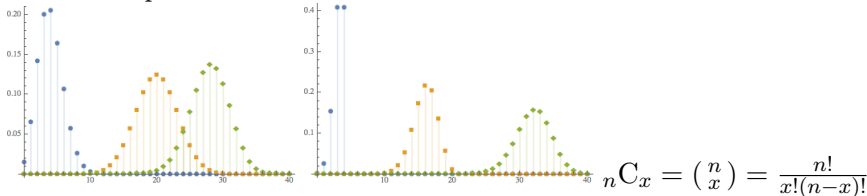
二項分布 高校 数学 B 岩薩林 確率・統計 §3.4

定義 (二項分布 岩薩林 確率・統計 (3.24)p.66)

離散型確率変数 X が次の確率分布を持つとき、 X はパラメタ n, p の二項分布 $B(n, p)$ にしたがうという。

$$p(x) = \begin{cases} \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} & (x = 0, 1, 2, 3, \dots, n) \\ 0 & (\text{他}) \end{cases}$$

意味: 確率 p で表の出るコインを n 回投げたとき、 x 回表が出る確率。



$B(40, 0.1), B(40, 0.5), B(40, 0.7), B(4, 0.8), B(20, 0.8), B(40, 0.8)$

二項分布 $B(n, p)$ の Python での扱い

Google Colab on Moodle

```
1 from scipy import stats
2
3 # 二項分布 = binomial distribution
4 # scipy.stats.binom(n=10,p=0.3) 確率分布名=binom(n,p)
5 rvx=stats.binom(n=10,p=0.3)
6
7 # 確率(質量)関数 probability mass function 離散型に使う用語
8 rvx.pmf(x) #  $=p(x)$ 
9
10 # 累積分布関数
11 rvx.cdf(x)
12
13 # 母平均値
14 rvx.mean()
```

命題 (二項分布の母平均値と母分散 岩薩林 確率・統計 定理 3.1(p.70))

$$E[X] = np, V[X] = np(1 - p)$$

(証明延期)

$$E[1] = \text{岩薩林 確率・統計 式 (3.26)-(3.27)}$$

定理 (二項定理 高校 数学 A 岩薩林 確率・統計 (3.25)p.68)

$$(a + b)^n = \sum_{x=0}^n {}_n C_x a^x b^{n-x}$$

L08-Q3

Quiz(二項分布)

確率 $\frac{2}{3}$ で表がでるコインを 100 回投げる.

- 1 表がでる回数 X は確率変数である. X のしたがう確率分布を, パラメタ $\circ\circ$ の...分布のように, または記号で答えよう.
- 2 表が 40 回でる確率を求めよう. 階乗 $n!$ とべき乗 a^b と分数 $\frac{a}{b}$ は簡単化・約分しなくてよい. $P, C, \binom{n}{k}$ などの記号は使わないで答えること.
- 3 X の母平均値, 母分散を求めよう.

L08-Q4

Quiz(二項分布または独立同分布)

あるスーパーのおでんセット 1 パックには、確率 $\frac{7}{10}$ で卵が 2 個、確率 $\frac{3}{10}$ で卵が 0 個 (!) 入っている。おでんセットを 10 パックを買ったときに得られる卵の合計の個数を確率変数 Y とする。各パックの卵の個数は独立とする。

- 1 10 パック中 X パックが卵 2 個入りとする。 Y を X で表そう。 X のしたがう確率分布を答えよう。
- 2 母平均値 $E[Y]$ を求めよう。
- 3 母分散 $V[Y]$ を求めよう。
- 4 確率 $P(Y = 16)$ を求めよう。

ここまで来たよ

7 バイズ推定・独立性・独立同分布

8 独立同分布の和・二項分布

- 独立同分布
- 二項分布
- ベルヌーイ分布と二項分布・独立同分布の性質の応用
- 独立同分布にしたがう確率変数の和

ベルヌーイ分布

定義 (ベルヌーイ分布 岩薩林 確率・統計 名前が出てこない)

$n = 1$ の二項分布 $B(1, p)$ をパラメタ p のベルヌーイ分布という.

$$P(X = x) = p(x) = p^x \cdot (1 - p)^{1-x} = \begin{cases} 1 - p & (x = 0) \\ p & (x = 1) \\ 0 & (\text{他}) \end{cases}$$

意味: **ベルヌーイ試行**=(不公平な) コイン投げ. 表がでる確率 p . 表 $x = 1$.

命題 (ベルヌーイ分布のモーメントと母平均値と母分散)

$$E[X^k] = p \quad (k = 1, 2, \dots), \quad V[X] = p(1 - p).$$

命題 (ベルヌーイ分布と二項分布のもうひとつの関係)

岩薩林 確率・統計 例題 4.6 の最初の 1 文

X_1, X_2, \dots, X_n が独立同分布 $B(1, p)$ にしたがうとき,
 $S_n = X_1 + \dots + X_n$ は $S_n \sim B(n, p)$.

なぜなら

X_i は 1 枚のコインのうち表が出た枚数. S_n は n 枚のコインのうち表が出た枚数.

ここまで来たよ

7 バイズ推定・独立性・独立同分布

8 独立同分布の和・二項分布

- 独立同分布
- 二項分布
- ベルヌーイ分布と二項分布・独立同分布の性質の応用
- 独立同分布にしたがう確率変数の和

独立同分布にしたがう確率変数の和の性質

岩薩林 確率・統計 例題 4.6

命題 (i.i.d にしたがう確率変数の和)

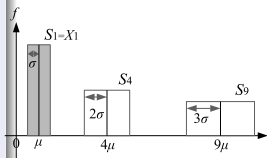
X_1, \dots, X_n : i.i.d. 母平均値 $E[X_i] = \mu$, 母分散 $V[X_i] = \sigma^2$.

和の確率変数 $S_n = X_1 + \dots + X_n$.

$$E[S_n] = \sum_{i=1}^n E[X_i] \stackrel{\text{同分布}}{=} n \times \mu.$$

$$V[S_n] \stackrel{\text{独立}}{=} \sum_{i=1}^n V[X_i] \stackrel{\text{同分布}}{=} n \times \sigma^2$$

S_n の確率 (密度) 関数はこんな感じ?



L08-Q5

Quiz(独立同分布にしたがう確率変数の和)

ダーツで、1回投げごとに点数が得られ、 n 回投げた合計点で競うルールでプレイしている。

あるプレイヤーの i 回目の点数を確率変数 X_i とすると、 X_i は独立同分布にしたがい、 $E[X_i] = 80$ 、 $V[X_i] = 200$ だという。

このプレイヤーが n 回投げた合計点を確率変数 $S_n = \sum_{i=1}^n X_i$ とする。

- 1 $E[S_{10}]$, $V[S_{10}]$ を求めよう。
- 2 $E[\frac{1}{10}S_{10}]$, $V[\frac{1}{10}S_{10}]$ を求めよう。
- 3 $E[\frac{1}{100}S_{100}]$, $V[\frac{1}{100}S_{100}]$ を求めよう。

二項分布の母平均値・母分散の導出

$X_1, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} B(1, p)$ とすると, $E[X_i] = p, V[X_i] = p(1 - p)$.

このとき, 独立同分布の和 確率統計 I(2022)L09 $S_n = \sum_{i=1}^n X_i \sim B(n, p)$ に対して,

$$E[S_n] = E[X_1] + \dots + E[X_n] \stackrel{\text{同分布}}{=} n \times p$$

$$V[S_n] \stackrel{\text{独立}}{=} V[X_1] + \dots + V[X_n] \stackrel{\text{同分布}}{=} n \times p(1 - p).$$