二項分布・独立同一分布の和

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 L07(2025-06-02 Mon)

最終更新: Time-stamp: "2025-06-02 Mon 06:46 JST hig"

今日の目標

- 独立同一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 久保川 統計学入門 §6.5, p.152
- 二項分布の母期待値, 母分散, 確率を求められる 久保川 統計学入門 §4.3.1,§4.3.2

L06-Q1 L06-Q2 L06-Q3
Quiz 解答: 多変数の確率変数の期待値

1. $p(x,y) =$

$y \backslash x$	1	2	3
0	0	$\frac{2}{12}$	$\frac{1}{12}$
2	$\frac{4}{12}$	0	$\frac{5}{12}$

$g(x,y) = 2x^2 + \mathrm{e}^y =$

$y \backslash x$	1	2	3
0	$2 \cdot 1^2 + \mathrm{e}^0$	$2 \cdot 2^2 + \mathrm{e}^0$	$2 \cdot 3^2 + \mathrm{e}^0$
2	$2 \cdot 1^2 + \mathrm{e}^2$	$2 \cdot 2^2 + \mathrm{e}^2$	$2 \cdot 3^2 + \mathrm{e}^2$

$$\mathrm{E}[2X^2 + \mathrm{e}^Y] = (2 \cdot 1^2 + \mathrm{e}^0)0 + (2 \cdot 2^2 + \mathrm{e}^0)\frac{2}{12} + (2 \cdot 3^2 + \mathrm{e}^0)\frac{1}{12}$$
$$+(2 \cdot 1^2 + \mathrm{e}^2)\frac{4}{12} + (2 \cdot 2^2 + \mathrm{e}^2)0 + (2 \cdot 3^2 + \mathrm{e}^2)\frac{5}{12}.$$

2. $\frac{4}{12} + 0 + \frac{5}{12} = \frac{9}{12}.$

3.

$$p_X(x) = \begin{cases} 4/12 & (x=1) \\ 2/12 & (x=2) \\ 6/12 & (x=3) \\ 0 & (\text{他}) \end{cases} \qquad p_Y(y) = \begin{cases} 3/12 & (y=0) \\ 9/12 & (y=2) \\ 0 & (\text{他}) \end{cases}$$

4. $\mathrm{E}[2X^2 + \mathrm{e}^Y] = 2\mathrm{E}[X^2] + \mathrm{E}[\mathrm{e}^Y] \overset{\text{周辺分布}}{=} 2(1^2 \cdot \frac{4}{12} + 2^2 \cdot \frac{2}{12} + 3^2 \cdot \frac{6}{12}) + (\mathrm{e}^0 \cdot \frac{3}{12} + \mathrm{e}^2 \cdot \frac{9}{12})$.
5. $\mathrm{Var}(X) = \mathrm{E}[X^2] - \mathrm{E}[X]^2 \overset{\text{周辺分布}}{=} \frac{66}{12} - (\frac{13}{12})^2 = \frac{29}{36}$.
6. $\mathrm{E}[XY] \overset{\text{同時分布}}{=} 1 \cdot 2 \cdot \frac{4}{12} + 3 \cdot 2 \cdot \frac{5}{12} = \frac{19}{6}$.
7. $\mathrm{Cov}(X, Y) = \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y] = \frac{19}{6} - \frac{26}{12} \cdot \frac{18}{12} = -\frac{1}{12}$.

L06-Q4

Quiz 解答: 離散型確率変数の独立性

$y\backslash x$	3	4	計
13	$\frac{3}{10}\cdot\frac{2}{7}$	$\frac{7}{10}\cdot\frac{2}{7}$	$\frac{2}{7}$
14	$\frac{3}{10}\cdot\frac{5}{7}$	$\frac{7}{10}\cdot\frac{5}{7}$	$\frac{5}{7}$
計	$\frac{3}{10}$	$\frac{7}{10}$	1

L06-Q5

Quiz 解答:2 次元の独立でない確率変数の母期待値

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

L06-Q6

Quiz 解答: 独立な確率変数の母期待値

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

ここまで来たよ

6 多変数の確率変数

7 二項分布・独立同一分布の和
- 二項分布
- ベルヌーイ分布
- 独立同一分布
- 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) I

L07-Q1

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 $\mathrm{E}[X]$ を求めよう.
3. 母分散 $\mathrm{Var}(X)$ を求めよう.

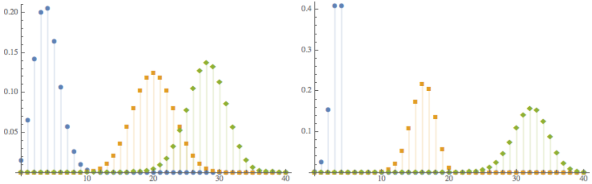
二項分布高校 数学 B 久保川 統計学入門 §4.3.2

定義 (二項分布 久保川 統計学入門 §4.3.2)

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 $Bin(n,p)$ 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.



二項係数 高校 数学 A

$${}_n\mathrm{C}_x = \binom{n}{x} = \frac{n!}{x!(n-x)!}$$

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

二項分布 $Bin(n,p)$ の Python での扱い

Google Colab on Moodle

$X \sim Bin(n,p) \leftrightarrow$ `rvx=scipy.stats.binom(n=`n`,p=`p`)`

```
from scipy import stats

# 二項分布 = binomial distribution
# scipy.stats.binom(n=10,p=0.3) 確率分布名=binom(n,p)
rvx=stats.binom(n=10,p=0.3)

# 確率(質量)関数 probability mass function 離散型に使う用語
rvx.pmf(x) # =p(x)

# 累積分布関数
rvx.cdf(x)

# 母平均値
rvx.mean()
```

命題 (二項分布の母平均値と母分散 久保川 統計学入門 公式 4.10)

$$\mathrm{E}[X] = \boxed{np}, \mathrm{Var}(X) = \boxed{np(1-p)}$$

(証明延期)

$\mathrm{E}[X^0] =$

定理 (二項定理 高校 数学 A 久保川 統計学入門 (4.4))

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

L07-Q2

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，パラメタ〇〇の…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.

久保川 統計学入門 例題 4.11

L07-Q3

Quiz(二項分布または独立同一分布)

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 $\mathrm{E}[Y]$ を求めよう.
3. 母分散 $\mathrm{Var}(Y)$ を求めよう.
4. 確率 $P(Y = 16)$ を求めよう.

ここまで来たよ

6 多変数の確率変数

7 二項分布・独立同一分布の和
- 二項分布
- ベルヌーイ分布
- 独立同一分布
- 独立同分布にしたがう確率変数の和

ベルヌーイ分布 久保川 統計学入門 §4.3.1

定義 (ベルヌーイ分布 久保川 統計学入門 §4.3.1)

$n=1$ の二項分布 $Bin(1,p)$ をパラメタ p のベルヌーイ分布 $Ber(p)$ 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と母分散 久保川 統計学入門 定理 4.8)

$$\mathrm{E}[X^k]=\boxed{p}\ (k=1,2,\ldots),\quad \mathrm{Var}(X)=\boxed{p(1-p)}.$$

公式に頼らず言える.

命題 (ベルヌーイ分布と二項分布の関係 久保川 統計学入門 例題 6.20)

$X_1, \ldots, X_n$ が独立同一分布 $Ber(p)$ にしたがうとき,
$S_n = X_1 + \cdots + X_n$ は $S_n \sim Bin(n, p)$.

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

6 多変数の確率変数

7 二項分布・独立同一分布の和
- 二項分布
- ベルヌーイ分布
- 独立同一分布
- 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

$$
\begin{aligned}
\mathrm{E}[aX_1 + bX_2 + c] =& a\mathrm{E}[X_1] + b\mathrm{E}[X_2] + c,\\
\mathrm{Var}(aX_1 + c) =& a^2\mathrm{Var}(X_1),\\
\mathrm{Var}(aX_1 + bX_2 + c) =& a^2\mathrm{Var}(X_1) + 2ab\mathrm{Cov}(X_1, X_2) + b^2\mathrm{Var}(X_2),\\
\mathrm{Cov}(X_1, X_2) =& \mathrm{E}[X_1X_2] - \mathrm{E}[X_1]\mathrm{E}[X_2].
\end{aligned}
$$

X_1, X_2 が独立なとき成立すること

$$
\begin{aligned}
\mathrm{E}[g_1(X_1) \times g_2(X_2)] =& \mathrm{E}[g_1(X_1)] \times \mathrm{E}[g_2(X_2)],\\
\mathrm{Cov}(X_1, X_2) =& 0,\\
\mathrm{Var}(aX_1 + bX_2 + c) =& a^2\mathrm{Var}(X_1) + b^2\mathrm{Var}(X_2).
\end{aligned}
$$

独立同一分布

定義 (独立同一分布 久保川 統計学入門 p.143,p.150)

確率変数 $X_1, \dots, X_n$ が互いに独立であり，同一の分布にしたがう（周辺分布の確率関数 p_{X_i} が同じである）とき，$X_1, \dots, X_N$ は独立同一分布 (i.i.d. independent and identically distributed) といい，次のように書く（たとえばその分布が $Ber(p)$ の場合）.

$$X_1, \dots, X_n, \text{i.i.d.} \sim Ber(p)$$

このとき，母期待値も等しい. $\mathrm{E}[X_i] = \mu, \mathrm{Var}(X_i) = \sigma^2$.

例: 形の同じ青色コイン X_1 と赤色コイン X_2.
表がでる $(X_i = 1)$ の確率 p
$X_1, X_2, \text{i.i.d.} \sim Ber(p)$.

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

ここまで来たよ

6 多変数の確率変数

7 二項分布・独立同一分布の和
- 二項分布
- ベルヌーイ分布
- 独立同一分布
- 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 久保川 統計学入門 §6.5

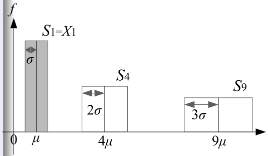
命題 (i.i.d にしたがう確率変数の和 久保川 統計学入門 公式 6.19)

$X_1, \dots, X_n$, i.i.d. 母平均値 $\mathrm{E}[X_i] = \mu$, 母分散 $\mathrm{Var}(X_i) = \sigma^2$.
和の確率変数 $S_n = X_1 + \cdots + X_n$.

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

$$\mathrm{Var}(S_n) \overset{\text{独立}}{=} \sum_{i=1}^{n} \mathrm{Var}(X_i) \overset{\text{同分布}}{=} n \times \sigma^2$$

S_n の確率 (密度) 関数はこんな感じ?



L07-Q4

Quiz(独立同一分布にしたがう確率変数の和)

ダーツで, 1 回投げる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, $\mathrm{E}[X_i] = 80, \mathrm{Var}(X_i) = 200$ 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. $\mathrm{E}[S_{10}]$, $\mathrm{Var}(S_{10})$ を求めよう.
2. $\mathrm{E}[\frac{1}{10}S_{10}]$, $\mathrm{Var}(\frac{1}{10}S_{10})$ を求めよう.
3. $\mathrm{E}[\frac{1}{100}S_{100}]$, $\mathrm{Var}(\frac{1}{100}S_{100})$ を求めよう.

L07-Q5

Quiz(独立同一分布にしたがう変数の和)

確率変数 $X_1, \ldots, X_n$ は $\mathrm{E}[X_i] = \mu, \mathrm{Var}(X_i) = \sigma^2$ の独立同分布に従う.

1. 確率変数 $S_n = X_1 + X_2 + X_3 + \cdots + X_n$ の母平均値と母分散を求めよう.
2. 確率変数 $U_n = \frac{1}{n}(X_1 + X_2 + X_3 + \cdots + X_n)$ の母平均値と母分散を求めよう.
3. 確率変数 U_n を標準化して, 確率変数 Y_n を, 母平均値 0, 母分散 1^2 になるように定義しよう.

二項分布の母平均値・母分散の導出

$X_1, \ldots, X_n$, i.i.d. $\sim Ber(p)$ とすると, $\mathrm{E}[X_i] = p, \mathrm{Var}(X_i) = p(1-p)$.

このとき, $S_n = \sum_{i=1}^{n} X_i \sim Bin(n, p)$ に対して,

$$\mathrm{E}[S_n] \overset{\text{いつでも}}{=} \mathrm{E}[X_1] + \cdots + \mathrm{E}[X_n] \overset{\text{同一}}{=} n \times p$$

$$\mathrm{Var}(S_n) \overset{\text{独立}}{=} \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n) \overset{\text{同一分布}}{=} n \times p(1-p).$$