独立性・二項分布・ベルヌーイ分布・大数の法則

樋口さぶろお

龍谷大学理工学部数理情報学科

確率統計☆演習 I　L08(2016-11-24 Thu)

最終更新: Time-stamp: "2016-11-24 Thu 06:49 JST hig"

今日の目標

- 確率変数の独立性が判定できる・利用できる 塚田確率統計 2.4
- 二項分布の母期待値が計算できる 塚田確率統計 4.2 塚田確率統計 4.1
- 大数の法則の意味が説明できる 塚田確率統計 5.2

http://hig3.net

L07-Q1
Quiz 解答:多次元の確率変数の期待値

1. $\mathrm{E}[X+2Y] = 0\cdot(1+2\cdot 0) + \frac{2}{12}(2+2\cdot 0) + \frac{1}{12}(3+2\cdot 0) + \frac{4}{12}(1+2\cdot 2) + 0(2+2\cdot 2) + \frac{5}{12}(3+2\cdot 2) = \frac{62}{12}$.
2. $\mathrm{E}[\mathbf{1}_{[Y\geq 1]}(X,Y)] = 0\cdot 0 + \frac{2}{12}\cdot 0 + \frac{1}{12}\cdot 0 + \frac{4}{12}\cdot 1 + 0\cdot 1 + \frac{5}{12}\cdot 1 = \frac{9}{12}$.
3. $$f_{\mathrm{X}}(x) = \begin{cases} 4/12 & (x=1) \\ 2/12 & (x=2) \\ 6/12 & (x=3) \\ 0 & (\text{他}) \end{cases} \qquad f_{\mathrm{Y}}(y) = \begin{cases} 3/12 & (y=0) \\ 9/12 & (y=2) \\ 0 & (\text{他}) \end{cases}$$
4. (1 の別解)

L07-Q4 Quiz 解答:離散型確率変数の独立性

1. $f_X(2) = f_X(4) = f_Y(2) = f_Y(4) = \frac{1}{2}$ であり，
   $f_X(x) \cdot f_Y(y) \neq f_{XY}(x, y)$ であり独立でない．
2. $\mathrm{E}[X] = \frac{1}{2} \cdot 2 + \frac{1}{2} \cdot 4 = 3, \mathrm{E}[Y] = \frac{1}{2} \cdot 2 + \frac{1}{2} \cdot 4 = 3, \mathrm{E}[XY] = \frac{1}{2} \cdot 4 + 0 \cdot 8 + 0 \cdot 8 + 16 \cdot \frac{1}{2} = 10, \mathrm{E}[X+Y] = 3+3 = 6, \mathrm{Cov}[X, Y] = 10 - 3 \times 3 = 1.$
   $\mathrm{Cov}[X, Y] = \mathrm{E}[XY] = \mathrm{E}[X]\mathrm{E}[Y] = 10 - 3 \cdot 3 = 1.$ $\mathrm{Cov}[X, Y] \neq 0$ からも独立でないことがわかる．
   $\mathrm{V}[X] = \mathrm{V}[Y] = 1$ より，$\rho[X, Y] = \frac{1}{\sqrt{1}\sqrt{1}} = 1.$

L07-Q6
Quiz 解答:独立と限らない確率変数の母期待値

1. $\mathrm{E}[-2X + 3Y] = -2\mathrm{E}[X] + 3\mathrm{E}[Y] = 5.$
2. $\mathrm{V}[-2X + 3Y] = \mathrm{E}[(-2X + 3Y)^2] - \mathrm{E}[-2X + 3Y]^2 = (-2)^2\mathrm{V}[X] + 2(-2)(3)\mathrm{Cov}[X, Y] + 3^2\mathrm{V}[X] = 20 - 84 + 99 = 35.$

ここまで来たよ

1. 確率不等式・2 変数の同時確率分布
   - 独立性
2. 独立性・二項分布・ベルヌーイ分布・大数の法則
   - 二項分布
   - ベルヌーイ分布
   - 独立同分布に従う確率変数の和と大数の法則

独立性 塚田確率統計 §2.4,§3.6 高校 数学 B

独立性

確率変数 X, Y が同時分布 $f_{\mathrm{XY}}(x, y)$ を持つとする.
X, Y が独立とは,

$$f_{\mathrm{XY}}(x, y) = f_{\mathrm{X}}(x) \times f_{\mathrm{Y}}(y)$$

が成立することをいう (世の中には, 同値な定義が多数).

独立とは, X,Y が互いに $\boxed{}$
事象 A, B が独立 $\Leftarrow P(A \text{かつ} B) = P(A) \times P(B)$ 塚田確率統計 2.4 の特別な場合.

独立性と母共分散 塚田確率統計定理 3.6.3

X, Y が独立なとき, 母共分散 $\mathrm{Cov}[X, Y] = 0$.

すぐ後で証明.

母共分散 $\mathrm{Cov}[X, Y] = 0$ は, X, Y が独立であるための $\boxed{}$ 条件.

L08-Q1

Quiz(離散型確率変数の独立性)

2 次元の離散型確率変数 (X, Y) を考える．同時分布 $f_{\mathrm{XY}}(x, y)$ は次の表で与えられる (現れない X, Y の確率は zero である).

$y \backslash x$	2	3
3	2/12	1/12
7	A	B

X, Y が独立になるように, 実数 A, B を定めよう.

X, Y が独立であるとき ‘だけ’ 成立する性質

$$
\begin{aligned}
\mathrm{E}[\phi_1(X) \times \phi_2(Y)] =& \mathrm{E}[\phi_1(X)] \times \mathrm{E}[\phi_2(Y)] \\
\text{特に } \mathrm{Cov}[X, Y] =& (\mathrm{E}[XY] - \mathrm{E}[X] \times \mathrm{E}[Y] =)0 \\
\mathrm{V}[X+Y] =& \mathrm{V}[X] + \mathrm{V}[Y]
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}[XY] =& \sum_x \sum_y f_{\mathrm{XY}}(x, y) \cdot x \cdot y \\
=& \sum_x \sum_y f_{\mathrm{X}}(x) \times f_{\mathrm{Y}}(y) \times x \times y \\
=& \sum_x f_{\mathrm{X}}(x) \cdot x \times \sum_y f_{\mathrm{Y}}(y) \cdot y = \mathrm{E}[X] \times \mathrm{E}[Y] \\
\mathrm{V}[X+Y] =& \mathrm{E}[(X+Y)^2] - \mathrm{E}[X+Y]^2 \\
=& \mathrm{E}[X^2] + 2\mathrm{E}[XY] + \mathrm{E}[Y^2] - (\mathrm{E}[X]^2 + 2\mathrm{E}[X]\mathrm{E}[Y] + \mathrm{E}[Y]^2) \\
=& \mathrm{V}[X] + 2\mathrm{Cov}[X, Y] + \mathrm{V}[Y]
\end{aligned}
$$

L08-Q2

Quiz(独立な確率変数の期待値)

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. $\mathrm{E}[(-2X + 3Y)(X + 5Y)]$ を求めよう.
2. $\mathrm{V}[-2X + 3Y]$ を求めよう.

ここまで来たよ

1. 確率不等式・2 変数の同時確率分布
   - 独立性
2. 独立性・二項分布・ベルヌーイ分布・大数の法則
   - 二項分布
   - ベルヌーイ分布
   - 独立同分布に従う確率変数の和と大数の法則

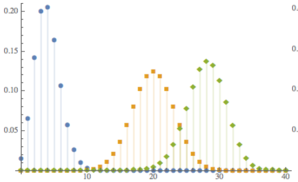
二項分布 塚田確率統計 §4.2 高校 数学 B

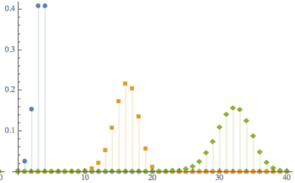
二項分布

離散型確率変数 X が次の確率分布を持つとき, X は二項分布 $\mathrm{B}(n,p)$ に従うという.

$$P(X=x)=f(x)=\begin{cases} {}_n\mathrm{C}_x\, p^x(1-p)^{n-x} & (x=0,1,2,3,\ldots,n) \\ 0 & (\text{他}) \end{cases}$$

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.





$\mathrm{B}(40,0.1),\mathrm{B}(40,0.5),\mathrm{B}(40,0.7)$, $\mathrm{B}(4,0.8),\mathrm{B}(20,0.8),\mathrm{B}(40,0.8)$

二項分布の母平均値と母分散

$$\mathrm{E}[X] = \boxed{}, \mathrm{V}[X] = \boxed{}$$

二項定理高校 数学 A

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

ここまで来たよ

1. 確率不等式・2 変数の同時確率分布
   - 独立性
2. 独立性・二項分布・ベルヌーイ分布・大数の法則
   - 二項分布
   - ベルヌーイ分布
   - 独立同分布に従う確率変数の和と大数の法則

ベルヌーイ分布 塚田確率統計 §4.1

ベルヌーイ分布

$n = 1$ の二項分布 $\mathrm{B}(1, p)$ のこと

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

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.

ベルヌーイ分布の母平均値と母分散

$$\mathrm{E}[X] = \boxed{}, \qquad \mathrm{V}[X] = \boxed{}$$

ベルヌーイ分布と二項分布

$X_1, X_2, \ldots, X_n$ が独立で $X_i \sim \mathrm{B}(1,p)$ のとき,
$U_n = X_1 + \cdots + X_n$ は $U_n \sim \mathrm{B}(n,p)$.

なぜなら

L08-Q3
L08-Q4

塚田確率統計章末問題 4.13.1

Quiz(二項分布)

確率 $p = \frac{2}{3}$ で表のでるいかさまコインがある．100 回投げる．

1. 表が 50 回でる確率を求めよう．
2. 表がでる回数の母平均値を求めよう．
3. 表がでる回数の母分散を求めよう．

L08-Q5

Quiz(ベルヌーイ分布)

ある宝くじは, あたりとはずれの 2 種類の結果だけがある. あたりの確率は 0.05 である. あたりの賞金は 1000 円, はずれの賞金は 0 円である. 賞金を確率変数 Y とする.

1. Y と, ベルヌーイ分布 $\mathrm{B}(1, p)$ に従う確率変数 X との関係を書こう.
2. Y の母平均値と母分散を求めよう.

ここまで来たよ

1. 確率不等式・2 変数の同時確率分布
   - 独立性
2. 独立性・二項分布・ベルヌーイ分布・大数の法則
   - 二項分布
   - ベルヌーイ分布
   - 独立同分布に従う確率変数の和と大数の法則

復習

チェビシェフの不等式 Chebyshev's inequality 塚田確率統計 3.5

X: 離散型または連続型確率変数
$\mu = \mathrm{E}[X]$: 母平均値
$\sigma^2 = \mathrm{V}[X]$: 母分散
$a > 0$: 任意の正の実数
のとき次が成立する.

$$P(|X - \mu| \geq a\sigma) \leq \frac{1}{a^2}$$

要するに

独立同分布の性質 塚田確率統計 5.1

独立同分布 (i.i.d.)

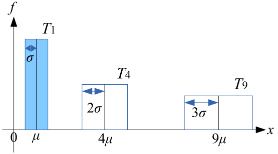
離散型/連続型確率変数 $X_1, X_2, \ldots, X_n$ が, たがいに独立で, すべて同じ確率分布に従う (同じ確率関数 $f(x)$) とする.
これを $X_1, \ldots, X_n$ は独立同分布に従う (i.i.d.=independent and identically-distributed) という.

新しい確率変数: $U_n = X_1 + \cdots + X_n$
母平均値 $\mathrm{E}[X_i] = \mu$, 母分散 $\mathrm{V}[X_i] = \sigma^2$ としたとき,

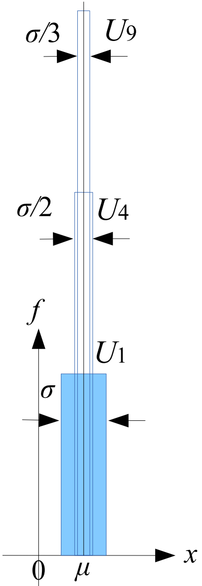
$$\mathrm{E}[U_n] = \sum_{i=1}^{n} \mathrm{E}[X_i] = n \times \mu.$$

$$\mathrm{V}[U_n] = \sum_{i=1}^{n} \mathrm{V}[X_i] = n \times \sigma^2.$$

U_n の確率密度関数はこんな感じ?



W_n の確率密度関数はこんな感じ?



新しい確率変数: $W_n = \frac{1}{n}U_n = \frac{1}{n}(X_1 + \cdots + X_n)$

$$\mathrm{E}[W_n] = \mathrm{E}\left[\tfrac{1}{n}U_n\right] = \frac{1}{n} \times n \times \mu.$$

$$\mathrm{V}[W_n] = \mathrm{V}\left[\tfrac{1}{n}U_n\right] = \left(\frac{1}{n}\right)^2 \times n \times \sigma^2.$$

L08-Q6

Quiz(独立同分布にしたがう変数の和)

確率変数 $X_1, \ldots, X_{100}$ は $\mathrm{E}[X_i] = 3, \mathrm{V}[X_i] = 7$ の独立同分布に従う.
次の確率変数の母平均値と母分散を求めよう.

1. 確率変数 $A = \frac{1}{100}(X_1 + X_2 + X_3 + \cdots + X_{100})$
2. 確率変数 $B = \frac{1}{10}(X_1 + X_2 + X_3 + \cdots + X_{100} - 100 \cdot 3)$
3. 確率変数 $C = \frac{1}{10\sqrt{7}}(X_1 + X_2 + X_3 + \cdots + X_{100} - 100 \cdot 3)$

大数の (弱) 法則 塚田確率統計 5.2

$W_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ は $\forall \epsilon > 0 \lim_{n\to+\infty} P(|W_n - \mu_X| \geq \epsilon) = 0$ を満たす.
つまり, n 大で W_n は母平均値 $\mu_X = \mathrm{E}[X]$ に「必ず近い」(確率収束)

証明　$\mu_{W_n} = \mu_X, \sigma^2_{W_n} = \sigma^2_X/n$. W_n に対するチェビシェフの不等式より,

$$P(|W_n - \mu_X| \geq a \times \frac{\sigma_X}{\sqrt{n}}) \leq \frac{1}{a^2}$$

$a = \frac{\epsilon}{\frac{\sigma_X}{\sqrt{n}}}$ とすると, $n \to +\infty$ で

$$P(|W_n - \mu_X| \geq \epsilon) \leq \frac{\sigma^2_X}{\epsilon^2 n} \to 0$$

これが母期待値の直観的意味. 要するに,

連絡

- 予習問題は, $\boxed{\text{次々回の授業直前}}$ を締切 (そこまでの最高点を記録) とします. でも, Trial までにやったほうが効率いいと思う.
  前からそうだけど, 予習問題が満点だと, Trial の満点の 1/3 まで保証されます.
- 配布資料は 1-503 向かいの引出, `http://hig3.net` で再配布.
- 加減乗除と平方根 (ルート) の使える電卓持ってきてね. 関数電卓でなくてもいいです. 携帯電話の機能・アプリでもかまいません.
- 樋口オフィスアワー木 6 金昼 (1-502), Math ラウンジ月-木昼 (1-614)
- 次回は 塚田確率統計 4.7 塚田確率統計 5.3

`https://manaba.ryukoku.ac.jp`