離散型確率変数

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　L06(2014-11-07 Fri)

今日の目標

- 離散確率分布が与えられたときに, 確率変数の母期待値・母平均値・母分散・母標準偏差が計算できる
- 離散確率分布が与えられたときに, 事象の確率が計算できる

http://hig3.net

L05-S1

Quiz 解答:共分散と相関係数

共分散 $C_{xy} = \frac{1}{5}[(2-4)(4-9)+\cdots] = 6.8.$
相関係数 $r = \frac{6.8}{1.90\times 3.85} = 0.93.$
回帰係数 $b = \frac{0.93\times 3.85}{1.90} = 1.88.$
よって回帰直線は, $y = 1.88(x-4)+9.$

ここまで来たよ

1. 略解:回帰分析
2. 離散型確率変数
   - 確率分布
   - 母期待値・母平均値・母分散・母標準偏差

データ 対 確率

データ＝これまでの状況:有限母集団
78 人のアイドル集団の身長データが Excel のデータで与えられました.
確率＝今日から考える少し不自然な状況:無限母集団
アイドル工場で, サイコロをふって 3 種類の身長から選んで, 何人でも生産される.

目	x	確率 $P(X=x)$
1,2,3	158	$\frac{3}{6}$
4,5	160	$\frac{2}{6}$
6	165	$\frac{1}{6}$

,　$P(X=158)=\frac{3}{6}$

生産結果でなく, ‘本当の’ 比率 (確率モデル) のほうを気にしよう.
確率変数: 変数 x と確率分布を組で考えたもの.　いいかげんな定義
「確率変数 X は, …という確率分布にしたがう」
X は確率変数, x は具体的な値.

確率=今日から考える少し不自然な状況 2:無限母集団

自然数すべての, 3 で割った余りを考えよう.

自然数 x	余り
1	1
2	2
3	0
4	1
$\vdots$	

先週までののりでの平均値 $\bar{x} = \frac{1}{+\infty}(1+2+0+1+2+\cdots) = \frac{+\infty}{+\infty}$?

データを生成する仕組み (確率モデル) のほうを気にする.

x	確率 $P(X=x)$
0	$\frac{1}{3}$
1	$\frac{1}{3}$
2	$\frac{1}{3}$

データと確率の関係

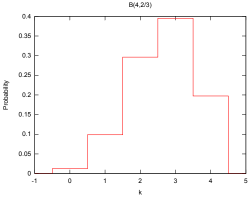
確率分布の表ってちょっと $\boxed{}$ に似てる…

x	確率
158	$\frac{3}{6}$
160	$\frac{2}{6}$
165	$\frac{1}{6}$
合計	1

$\leftrightarrow$

x	度数	相対度数
158	39	$\frac{39}{78}$
160	26	$\frac{26}{78}$
165	13	$\frac{13}{78}$
合計	78	1

$P(X = x)$ のグラフ, データの $\boxed{}$ に似てる…



事象と確率

X:確率変数

ここでは, (量的) 確率変数 1 個という限られた範囲で確率論を展開しています.

本来は, 事象が基本で, そこを定義域とする関数として確率変数を後から考えます.

事象 起きたかどうか判定できる物事. $\rightsquigarrow$ 「 $X \in A$ である」$A \subset \mathbb{R}$

例: 「$X = 3$」 「$X \leq 2$」, 「X は素数」,「$X = 1$ or $X = 3$」

全事象 $U = \mathbb{R}$. 事象はこの部分集合 $A \subset \mathbb{U}$.

空事象 $\emptyset$

基本事象 $A = \{a\}$. それ以上分けられない

以下は当面高校の知識で

補事象 $A^{\mathrm{c}} = U \setminus A$. A が起きなかったという事象.

和事象 $A \cup B$ または,

積事象 $A \cap B$ かつ,

排反事象 「A, B が排反事象」 $\Leftrightarrow A \cap B = \emptyset$. 同時に起きない

確率分布の記号

事象の確率 P(事象).
基本事象の確率 $P(X = x) = f(x)$　　$f(x)$:確率関数, (離散) 確率分布.

x_k	確率 $P(X = x_k)$
$x_1 = 158$	$\frac{3}{6}$
$x_2 = 160$	$\frac{2}{6}$
$x_3 = 165$	$\frac{1}{6}$
合計	1

$k = 1, 2, 3 = m.$

$$\to f(x) = \begin{cases} \frac{3}{6} & (x = 158) \\ \frac{2}{6} & (x = 160) \\ \frac{1}{6} & (x = 165) \\ 0 & (\text{他}) \end{cases}$$

確率の性質

$$P(U) = 1, \quad P(\emptyset) = 0, \quad 0 \leq P(\text{事象 } A) \leq 1$$

$$A,B \text{ が排反事象のとき} \qquad P(A \cup B) = P(A) + P(B)$$

確率分布の性質

$$\sum_{k=1}^{m} P(X = x_k) = 1, \quad \text{対応物無}, \quad 0 \leq P(X = x_k) \leq 1 \qquad (k = 1, \ldots, m)$$

m は $P \neq 0$ であるような値の個数.

ここまで来たよ

1. 略解:回帰分析
2. 離散型確率変数
   - 確率分布
   - 母期待値・母平均値・母分散・母標準偏差

母期待値

X:確率変数, $x_1, \ldots, x_m$:X のとる値, $\phi(x)$:ふつうの関数.

関数 $\phi(x)$ の母期待値

$$\mathrm{E}[\phi(X)] = \sum_{k=1}^{m} P(X = x_k) \times \phi(x_k)$$

例:$\phi(x) = 2x, x^2, \mathrm{e}^x$, ($x$ の場合分けで書かれた関数), $\cdots$.

性質

$E[1] = 1$. ($\phi(x) = 1$ と $P(U) = 1$ から示せる)

定義

- 母平均値 $E[X] = m$.　　$\phi(x) = x$.
- 母分散 $= V[X] = E[(X - m)^2]$.　　$\phi(x) = (x - m)^2$.
- 母標準偏差 $= \sqrt{V[X]}$

L06-Q1

Quiz(離散的な確率変数の母平均・母分散・母標準偏差)

確率変数 X は

- 値 $x_1 = -1$ を確率 $\frac{1}{3}$ で
- 値 $x_2 = 0$ を確率 $\frac{5}{12}$ で
- 値 $x_3 = +2$ を確率 $\frac{1}{4}$ で

とる.

1. 期待値 $\mathrm{E}(\mathrm{e}^X)$ を求めよう.
2. X の母平均値を求めよう.
3. X の母分散を求めよう.
4. X の母標準偏差を求めよう.
5. 事象 $X \leq 1$ の確率を求めよう.

ベルヌーイ分布 B(1,p)

$$P(X=1)=1-p,\quad P(X=0)=p,\quad (0\le p\le 1)$$

「確率 p で表のでるコイン」

L06-Q2

Quiz(離散的な確率変数の母平均・母分散・母標準偏差)

ベルヌーイ分布に従う確率変数 X を考える．つまり X は確率 p で値 $x_1=1$ を, 確率 $1-p$ で 値 $x_2=0$ をとる．

1. 母期待値 $\mathrm{E}[\sin(\frac{\pi}{2}X)]$ を求めよう．
2. X の母平均値を求めよう．
3. X の母分散を求めよう．
4. X の母標準偏差を求めよう．
5. 事象 $X\le\frac{1}{2}$ の確率を求めよう．
6. 母期待値 $\mathrm{E}[2\cos(\frac{\pi}{2}X)]$ を求めよう．

事象の確率

事象 A が起きる $\Leftrightarrow$ X の条件 $X \in A$ が成立

特徴関数

$$\mathbf{1}_{[\text{事象}]}(x) = \begin{cases} 1 & (X = x \text{ のとき事象が起きる}) \\ 0 & (X = x \text{ のとき事象が起きない}) \end{cases}$$

とすると，

$$P(A) = \mathrm{E}[\mathbf{1}_{[A]}(X)]$$

事象のかわりに，「条件」と思うと考えやすいかも．

例

平均値, 分散の性質

母平均値の性質

X: 確率変数, $a, b \in \mathbb{R}$:定数 のとき,

$$\begin{aligned}
\mathrm{E}[aX+b] &= \sum_{k=1}^{m} P(X=x_k) \times (ax_k+b) \\
&= \left(a \sum_{k=1}^{m} P(X=x_k) x_k\right) + b = a\mathrm{E}[X] + b.
\end{aligned}$$

$$\begin{aligned}
\mathrm{E}[\phi_1(X)+\phi_2(X)] &= \sum_{k=1}^{m} P(X=x_k) \times (\phi_1(X)+\phi_2(X)) \\
&= \mathrm{E}[\phi_1(X)] + \mathrm{E}[\phi_2(X)].
\end{aligned}$$

もちろん一般には $\mathrm{E}(\phi(X)) \neq \phi(\mathrm{E}(X))$, $\mathrm{E}(X^2) \neq (\mathrm{E}(X))^2$.
これ, $\sin(x^2) \neq (\sin(x))^2$ と同じくらいだいじ.

母分散の性質

X: 確率変数, $a, b \in \mathbb{R}$:定数 のとき,

$$\mathrm{V}[aX + b] = a^2 \mathrm{V}[X].$$

母分散の性質

$$\mathrm{V}[X] = \mathrm{E}[X^2] - (\mathrm{E}[X])^2$$

2 項分布 $B(n,p)$

$$P(X=k)={}_n\mathrm{C}_k\cdot p^k(1-p)^{n-k}\qquad (k(=x_k)=0,1,2,\dots,n)$$

- ベルヌーイ分布 $B(1,p)$ に従う X を n 回繰りかえして試行したときに, $X=1$ となる回数の分布
- 表が確率 p で出るコインを n 回投げたとき, 表が k 回出る確率

$x_k=k$	確率
0	$1\cdot p^0(1-p)^n$
1	$n\cdot p^1(1-p)^{n-1}$
2	$\frac{n(n-1)}{2}\cdot p^2(1-p)^{n-2}$
$\vdots$	
k	${}_n\mathrm{C}_k p^k(1-p)^{n-k}$
$\vdots$	
n	$1\cdot p^n(1-p)^0$

正攻法では計算が難しい.
L06-Q3

Quiz(2 項分布)

2 項分布 $B(n, p)$ に従う確率変数 X を考える.

1. 期待値 $\mathrm{E}(1) = 1$ を確かめよう.
2. 事象 $X \geq 2$ の確率を求めよう.
3. X の母平均値を求めよう.
4. X の母分散を求めよう.
5. X の母標準偏差を求めよう.

幾何分布 $G(n,p)$

$$P(X=k)=p(1-p)^{k-1} \qquad (k(=x_k)=1,2,\ldots)$$

- ベルヌーイ分布 $B(1,p)$ に従う X を 繰りかえして試行したときに, k 回目で初めて $X=1$ となる確率

L06-Q4

Quiz(幾何分布)

確率変数 X が幾何分布
$P(X=k)=p(1-p)^{k-1} \qquad (k(=x_k)=1,2,\ldots)$ に従うとする.

1. $P(X\in\mathbb{R})=1$ を確かめよう.
2. 確率 $P(X\le k)$ を求めよう.
3. X の母平均値を求めよう.
4. X の母分散を求めよう.
5. X の母標準偏差を求めよう.

連絡

- 予習問題ポリシー変更: 点数:最終受験→最高点, 締切:水 9:20 →金 9:00, 正誤表示:締切後→受験後.
- 2014-11-07 金 23:55 レポート課題締切 提出場所: RaMMoodle
- 2014-11-12 水 5, 12-03 水 4 数理情報学科特別講義
- 2014-11-15 土 3 数学検定勉強会 1-537.
- 2014-11-17 から チューターは月火水木昼 (1-614).
- 2014-11-21 金 2 プチテスト. 非参照. 30 ピーナッツ.
- 2014-11-21 金 3 特別研究履修説明会 (3 年生向け) 英語と重なっちゃう…

プチテスト計画!

- 2014-11-21 金 2, 75 分 (御生誕法要だから), 30 ピーナッツ, 参照相談なし. 紙のテスト.
- まず授業でやらなかったページに×つけましょう.
- 過去問ありません. 下の出題計画, 非参照 Quiz, 予習問題をやり直すことをお奨めします.
- 出題計画 (2014-11-14 金ごろ修正, 確定します). Excel 関係のものはありません.
  - データから平均値, 分散, 標準偏差を求める
  - データから箱ひげ図を描く
  - データから共分散, 相関係数を求める
  - データから回帰係数, 回帰直線を求める
  - 離散型確率変数について, 確率, 期待値, 平均値, 分散, 標準偏差を求める
  - 連続型確率変数について, 確率, 期待値, 平均値, 分散, 標準偏差を求める
  - 選択肢的な問