連続型確率変数

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 L02(2025-04-21 Mon)

最終更新: Time-stamp: "2025-04-21 Mon 11:47 JST hig"

今日の目標

- 久保川 統計学入門 §5.1 連続型確率変数とは何か説明できる
- 久保川 統計学入門 §5.2 連続型確率変数の母平均値, 母分散, 母期待値が計算できる

L01-Q1

Quiz 解答: 離散的な確率変数の母平均・母分散・母標準偏差

1. 母期待値 $\mathrm{E}[3+12\mathrm{e}^X] =$
$\frac{4}{12}\cdot(3+12\mathrm{e}^{-1}) + \frac{5}{12}\cdot(3+12\mathrm{e}^{0}) + \frac{3}{12}\cdot(3+12\mathrm{e}^{2}) = 8+4\mathrm{e}^{-1}+3\mathrm{e}^{2}$.
2. 母平均値 $\mathrm{E}[X] = \frac{4}{12}\cdot(-1) + \frac{5}{12}\cdot 0 + \frac{3}{12}\cdot 2 = \frac{1}{6}(=\mu)$.
3. 母分散
$\mathrm{Var}(X) = \mathrm{E}[(X-\mu)^2] = \frac{4}{12}\cdot(-1-\frac{1}{6})^2 + \frac{5}{12}\cdot(0-\frac{1}{6})^2 + \frac{3}{12}(2-\frac{1}{6})^2 = \frac{47}{36}$.

計算を楽にする母期待値の性質

命題 (母期待値の性質高校 数学 B 久保川 統計学入門 公式 4.4(p.85)**)**

確率変数 X，$a, b \in \mathbb{R}$: 定数，関数 g, h に対して，

$$\begin{aligned} \mathrm{E}[ag(X) + bh(X)] &= \sum_x (ag(x) + bh(x)) \times p(x) = a\mathrm{E}[g(X)] + b\mathrm{E}[h(X)] \\ \mathrm{E}[a] &= a, \\ \mathrm{E}[1] &= 1, \\ \mathrm{E}[aX + b] &= a\mathrm{E}[X] + b \end{aligned}$$

定義 (k 次のモーメント久保川 統計学入門 p.121**)**

$\mathrm{E}[X^k]$ を X の k 次のモーメントという $(k = 0, 1, 2, \dots)$.

k 次のモーメントを部品として計算しておき, 他の母期待値は性質から導くのが楽.
$\mathrm{E}[X^2] \neq \mathrm{E}[X]^2$

母分散の性質 久保川 統計学入門 公式 4.5(p.86)

分散の計算方法（分散公式） 高校 数学 B 久保川 統計学入門 公式 4.5(p.86),p.106

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

$$\mathrm{Var}(X) = \mathrm{E}[X^2] - (\mathrm{E}[X])^2$$
$$\mathrm{Var}(aX + b) = a^2 \mathrm{Var}(X).$$

証明 $\mu = \mathrm{E}[X]$ とする.
$\mathrm{Var}(X) = \mathrm{E}[(X-\mu)^2] = \mathrm{E}[X^2 - 2\mu X + \mu^2] =$
$\mathrm{E}[X^2] - 2\mu\mathrm{E}[X] + \mu^2\mathrm{E}[1] = \mathrm{E}[X^2] - 2\mu^2 + \mu^2.$
証明 $\mu = \mathrm{E}[X]$ と, $\mathrm{E}[Y] = \mu_{\mathrm{Y}} = a\mu + b$ とする.

$$\begin{aligned}\mathrm{Var}(Y) =& \mathrm{E}[(Y - \mu_{\mathrm{Y}})^2]\\ =& \mathrm{E}[((aX+b) - (a\mu + b))^2]\\ =& a^2\mathrm{E}[(X-\mu)^2] = a^2\mathrm{Var}(X).\end{aligned}$$

ここまで来たよ

1. 離散型確率変数
   - 標本抽出と Python の scipy による確率変数の扱い
2. 連続型確率変数
   - 連続型確率変数

確率変数の標本抽出

久保川 統計学入門 §7 を先取り

定義 (標本, 標本抽出)

確率変数 X の値を得る試行を n 回行って得られる実現値 x の長さ n の列を, サイズ (size)n の標本 (sample) という.
標本を得る操作を標本抽出 (sampling) という.

標本は‘毎回違った結果’になる.

データ分析 では標本を扱っていた.

scipy による確率変数の標本抽出

https://colab.research.google.com

```
import numpy as np # ライブラリの読込
from scipy import stats

xk = (0,   1 ,  2  , 3  ,4   , 5  ,6)    # x_1, x_2,..,x_k
pk = (0.1, 0.2, 0.3, 0.1, 0.1, 0.0, 0.2) # p_1, p_2,..,p_k
rvx = stats.rv_discrete(values=(xk, pk)) # 離散型確率変数 X を定義
sample=rvx.rvs(size=30) # サイズ30の標本抽出(毎回違った結果になる)

# 母ナントカ. 答はいつも同じ
rvx.mean()# 母平均値
rvx.var() # 母分散
rvx.std() # 母標準偏差

# データ分析 でやってた標本平均値
## 標本抽出のたびに結果が変わる
sample.mean() # 標本平均値
sample.var()  # 不偏標本平均値
sample.std()  # 標本標準偏差
```

ここまで来たよ

1. 離散型確率変数
   - 標本抽出と Python の scipy による確率変数の扱い
2. 連続型確率変数
   - 連続型確率変数

場合分けとそうでない確率変数

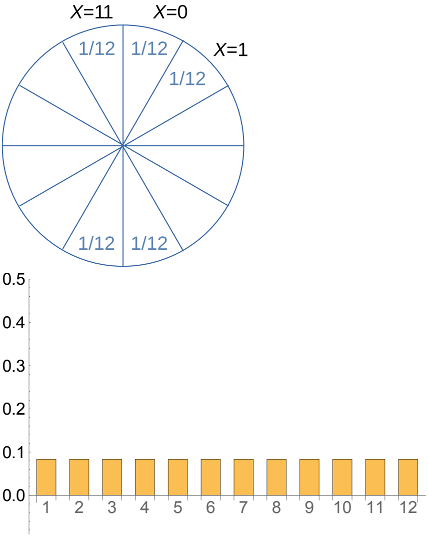
整数に値をとる離散型確率変数 X が次の確率関数 $p(x)$ を持つ．
(同じこと)

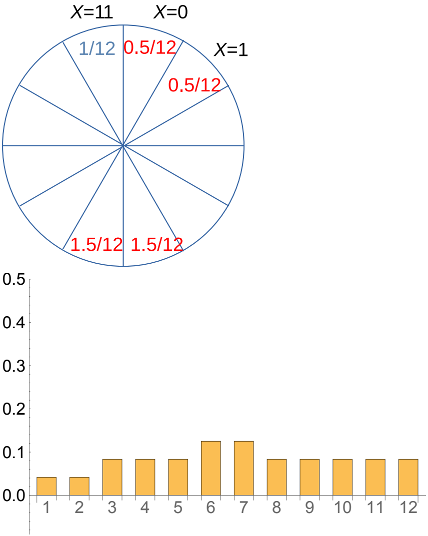
$$p(x) = \begin{cases} 0.1 & (x = 6) \\ 0.2 & (x = 7) \\ 0.3 & (x = 8) \\ 0.4 & (x = 9) \\ 0 & (\text{他}) \end{cases}$$

$$p(x) = \begin{cases} (x-5)/10 & (6 \le x \le 9) \\ 0 & (\text{他}) \end{cases}$$

離散型確率変数の例：離散ルーレット

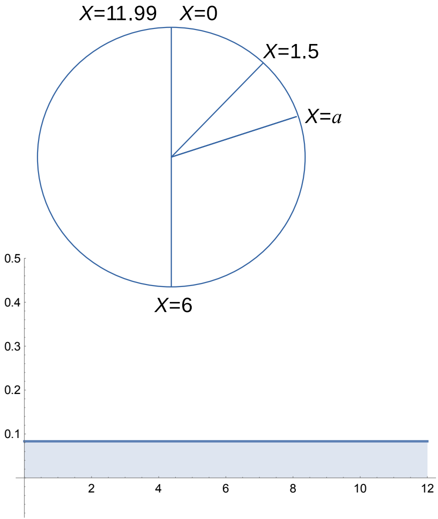
$P(X=1)=1/12, P(0\leq X\leq a)=1,\dots,a$ の棒の高さの和

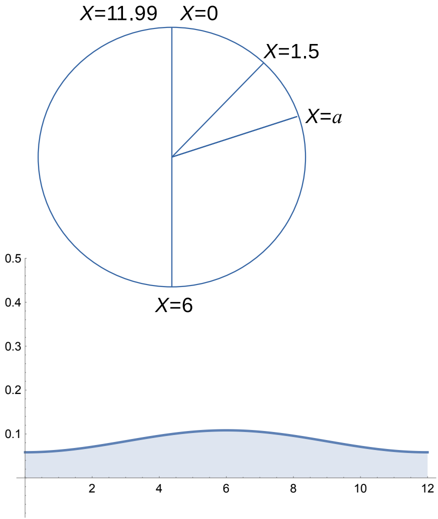




連続型確率変数の例：連続型ルーレット

$$P(X=1)=0, P(0\le X\le a)=a/12=(0 \text{ から } a \text{ までのグラフの面積})$$





連続型確率変数 久保川 統計学入門 §5.1

定義 (連続型確率変数 久保川 統計学入門 p.104)

連続型確率変数 X とは, 実数のように連続値をとる確率変数のこと.

$\rightsquigarrow$ X の実現値の出やすさは x のある関数 $f(x)$ で表される!
連続型確率変数 連続型確率分布
連続型 確率密度関数 $f(x)$ (x は実数)
離散型 確率関数 $p(x)$ (x は整数またはとびとびの値)

定義 (確率密度関数 $\boxed{\text{久保川 統計学入門 p.104}}$)

次の性質を持つ関数 $f(x)$ を，連続型確率変数の確率密度関数という．X の確率分布を指定できる．

$$
\begin{aligned}
f(x) &\geq 0 \\
\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x &= 1
\end{aligned}
$$

$f(x) \leq 1$ とはいえない！

離散的

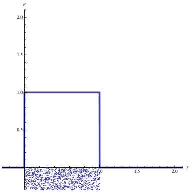
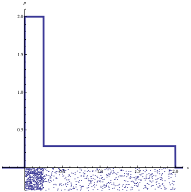
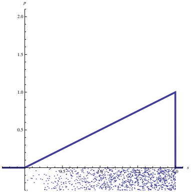
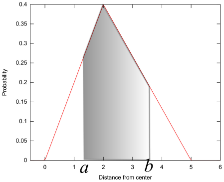
得点 x	確率 $p(x)$
0	0.1
1	0.3
$\vdots$	
x	$p(x)$

連続的

- $f(x)$ が大きいほど, その値 x がでやすい

物理・工学系では $p(x)$ と書いたら確率密度関数 $f(x)$ を意味することも

確率密度関数の例

横軸下の細かい点が, 標本 (縦方向の位置はランダムで意味なし)

命題 (確率密度関数と確率 久保川 統計学入門 p.104)

$$P(c < X \leq d) = \int_c^d f(x)\,\mathrm{d}x \quad \text{(下側面積)}$$

連続型確率変数の母期待値

定義 (母期待値 久保川 統計学入門 定義 5.3, 5.4)

$$\text{離散型確率変数} \quad \mathrm{E}[g(X)] = \sum_{x} g(x) \cdot p(x)$$

$$\text{連続型確率変数} \quad \mathrm{E}[g(X)] = \int_{-\infty}^{+\infty} g(x) \cdot f(x)\,\mathrm{d}x$$

- 離散型と同じ定義，同じ公式が成立．
- 母平均値 $\mu = \mathrm{E}[X]$, 母分散 $\mathrm{Var}(X) = \mathrm{E}[(X-\mu)^2]$.

定義 (k 次のモーメント ($k = 0, 1, 2, 3, \cdots$) 久保川 統計学入門 p.121)

$$\mathrm{E}[X^k] = \int_{-\infty}^{+\infty} x^k \cdot f(x)\,\mathrm{d}x$$

L02-Q1

Quiz(連続型確率変数)

次の確率密度関数 $f(x)$ を持つ連続型確率変数 X を考える．次を求めよう．

$$f(x) = \begin{cases} 8x & (0 \leq x < \frac{1}{2}) \\ 0 & (\text{それ以外}) \end{cases}$$

1. 母期待値 $\mathrm{E}[X^k]$ $(k = 0, 1, 2, \ldots)$
2. 母平均値 $\mathrm{E}[X]$
3. 母分散 $\mathrm{Var}(X)$
4. 母期待値 $\mathrm{E}[(2X+3)^2]$
5. 母分散 $\mathrm{Var}(2X+3)$
6. 確率 $P(-\frac{1}{4} < X \leq \frac{1}{4})$

久保川 統計学入門 例題 5.6,5.7(p.107),§5 基本問題問 2,3(p,124)