連続型確率変数

樋口さぶろお　https://hig3.net

龍谷大学理工学部数理情報学科

確率統計☆演習　L09(2020-11-23 Mon)

最終更新: Time-stamp: "2020-12-04 Fri 08:46 JST hig"

今日の目標

- 連続型確率変数とは何か説明できる　高校 数学 B
- 連続型確率変数の確率, 母ナントカを計算できる

  岩薩林 確率・統計 §4.1
- 一様分布を例に説明できる

L08-Q1

Quiz 解答:離散的な確率変数の母平均値・母分散・母標準偏差・確率

1. $\mathrm{E}[\mathrm{I}_{[X\le 5]}(X)] = \sum_{x=0}^{10} \frac{x}{55} \mathrm{I}_{[X\le 5]}(x) = \sum_{x=1}^{5} \frac{x}{55} = \frac{15}{55} = \frac{3}{11}.$
2. $\mathrm{E}[X] = \sum_{x=0}^{10} \frac{x}{55} \cdot x = \frac{\frac{1}{6}\cdot 10 \cdot (10+1)(2\cdot 10+1)}{55} = \frac{385}{55} = 7.$
3. $\mathrm{V}[X] = \mathrm{E}[X^2] - (\mathrm{E}[X])^2 = \sum_{x=0}^{10} \frac{x}{55} \cdot x^2 - 7^2 = 55 - 7^2 = 6.$

L08-Q2

Quiz 解答:二項分布

表がでる回数は二項分布 $\mathrm{B}(100, \frac{2}{3})$ に従う確率変数を X である.

1. $P(X=40)$ を求めればいいから, ${}_{100}\mathrm{C}_{40}p^{40}(1-p)^{100-40} = \frac{100!}{40!60!}(\frac{2}{3})^{40}(1-\frac{2}{3})^{60}$.
2. $\mathrm{E}[X] = n \times p = \frac{200}{3}$. $\mathrm{V}[X] = n \times p(1-p) = \frac{200}{9}$.

L08-Q3

Quiz 解答:二項分布または独立同分布

1. $X \sim \mathrm{B}(10, 0.3)$ とすると, $Y = 2X$.
2. $\mathrm{E}[Y] = \mathrm{E}[2X] = 2 \times 10 \cdot 0.7 = 14$ 個.
3. $\mathrm{V}[Y] = 2^2\mathrm{V}[X] = 2^2 \times 10 \cdot 0.7 \cdot 0.3 = 8.4$ 個 ^{2}.
4. $P(Y=16) = P(X=8) = \frac{10}{2!8!}0.7^8 0.3^2$

別解．10 パックそれぞれに入っている卵の個数を確率変数 B_i とすると，B_i は独立同分布にしたがう．$B_i = 2A_i$, $A_i \sim \mathrm{B}(1, 0.7)$ である．
$\mathrm{E}[B_i] = 1.4$, $\mathrm{V}[B_i] = 0.84$ なので,
$\mathrm{E}[Y] = 10\mathrm{E}[B_i] = 14, \mathrm{V}[Y] = 10\mathrm{V}[B_i] = 8.4.$

L08-Q4

Quiz 解答:独立同分布の母期待値母分散

1. 3μ.
2. 3μ.
3. 3μ
4. $9\sigma^2$.
5. $3\sigma^2$.
6. $5\sigma^2$.
7. $3(\sigma^2 + \mu^2) + 3 \times 2 \times \mu^2$.

L08-Q5

Quiz 解答:独立同分布にしたがう確率変数の和

1. $\mathrm{E}[T_{10}] = 800$, $\mathrm{V}[T_{10}] = 2000$.
2. $\mathrm{E}[\frac{1}{10}T_{10}] = 80$, $\mathrm{V}[\frac{1}{10}T_{10}] = 20$.
3. $\mathrm{E}[\frac{1}{100}T_{100}] = 80$, $\mathrm{V}[\frac{1}{100}T_{100}] = 2$.

Q6 は Q3 の別解参照.

ここまで来たよ

8 二項分布と独立同分布にしたがう確率変数の和

9 連続型確率変数
- 連続型確率変数
- 連続型一様分布
- 確率変数の 1 次式による変換と標準化

復習 (x で書かれた確率関数)

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 $\mathrm{B}(n, p)$ にしたがうという．

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

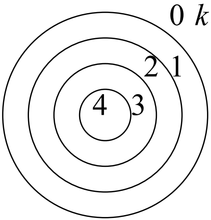
$n = 3$ のとき，

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

$$\mathrm{E}[g(X)] = \sum_{x=0}^{3} g(x) p(x)$$

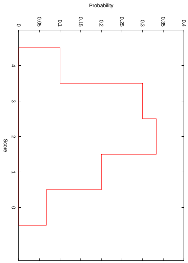
あるプレイヤーのダーツの得点確率

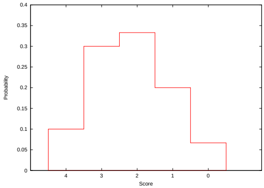
得点: 的の真ん中から順に $4, 3, 2, 1, 0$ 点



離散型確率分布

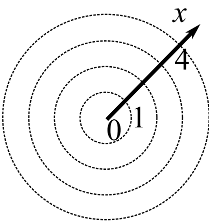
得点 s	確率関数 $f(s)$
4	0.1
3	0.3
2	0.3333
1	0.2
0	0.0667

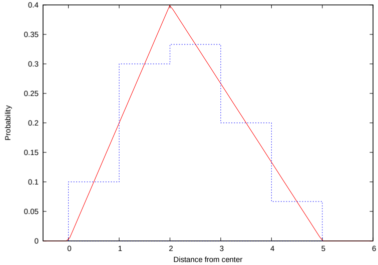




中心から x cm にあてる確率 岩薩林 確率・統計 §4.1

的の真ん中からの距離 x cm, 得点 $s = 4 - x$ 点 (実数).





$x = 0.5$cm と 0.9 cm への当たりやすさは違う. $x = 1.0$cm を境に急に変わるわけじゃない. これを表現したい.

$\rightsquigarrow$ 点数の出やすさは x のある関数 $p(x)$ 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)

連続型確率変数

連続型確率変数

連続型確率変数 X とは, 実数値をとり, 確率が確率密度関数 $f(x)$ で指定されるもの.

離散的

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

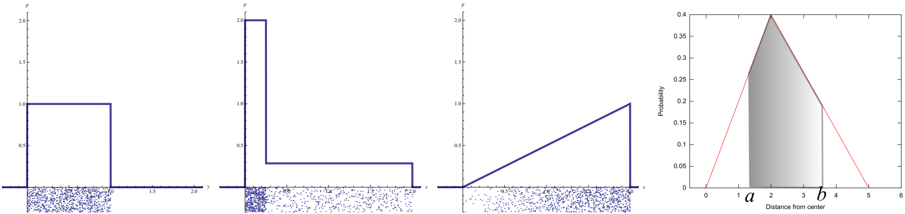
連続的

- $f(x)$ が大きいほど, その値 x がでやすい
- $0 \leq f(x)$ である. $f(x) \leq 1$ とは限らない.

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

計算科学☆演習 II

確率密度関数の例



横軸下の細かい点が, 試行の結果 (縦方向の位置はランダムで意味なし)

確率密度関数と確率 岩薩林 確率・統計 (4.1)

$$P(a \leq X < b) = (\text{あとで}) = \int_a^b f(x)\,\mathrm{d}x \quad (\text{下側面積})$$

連続型確率変数の母期待値 岩薩林 確率・統計 §4.2

母期待値の定義 岩薩林 確率・統計 (4.8)

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

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

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

k 次のモーメント $(k = 0, 1, 2, 3, \cdots)$ 岩薩林 確率・統計 (4.9)

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

L09-Q1

Quiz(連続型確率変数)

次の確率密度関数 $f(x)$ を持つ確率変数 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 $\mathrm{V}[X]$
4. 母期待値 $\mathrm{E}[(2X+3)^2]$
5. 母分散 $\mathrm{V}[2X+3]$
6. 確率 $P(|4X| \leq 1)$

岩薩林 確率・統計 例題 4.2,4.3

確率密度関数から事象の確率を求める

$$P(\text{事象}) = P(\text{条件}) = \mathrm{E}[\mathrm{I}_{[\text{条件}]}(X)]$$

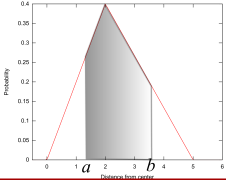
$$P(a \leq X < b) = \mathrm{E}[\mathrm{I}_{[a \leq X < b]}(X)]$$

$$= \int_{-\infty}^{+\infty} f(x)\mathrm{I}_{[a \leq X < b]}(x)\,\mathrm{d}x = \int_{a}^{b} f(x)\,\mathrm{d}x \quad \boxed{\text{面積}}$$

$$\text{全事象の確率} = 1 = \int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = \mathrm{E}[1]$$

岩薩林 確率・統計 第 4 章問題 1(p.79) 岩薩林 確率・統計 第 4 章練習問題 1

じゃあ，ちょうど距離 $x = a$ cm となる確率は? $\rightsquigarrow \boxed{0}$.

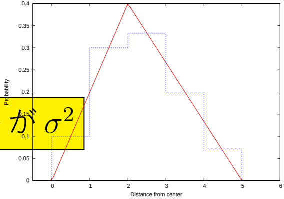


$$\mathrm{I}_{[X\,\text{の条件}]}(x) = \begin{cases} 1 & (x\,\text{が条件を満たす}) \\ 0 & (\text{それ以外}) \end{cases}$$

連続型確率変数の母平均値と母分散の直観的意味

$\mu = \mathrm{E}[X], \sigma^2 = \mathrm{V}[X]$

$f(x)$ のグラフ



その 1:確率密度関数のグラフから

$0 \le y \le f(x)$ 部分の横方向の重心の位置 μ

$0 \le y \le f(x)$ 部分の幅が σ, 慣性モーメントが σ^2

その 2:チェビシェフの不等式

母平均値から, (母標準偏差 σ を単位に測って) 離れた値がでる確率は小さい.

その 3:大数の弱法則

宝くじを何回も買うと, 1 回あたりの平均の賞金は,
E[賞金] に近い

母期待値・母分散・確率のチェックリスト

検算方法

- $\mathrm{E}[X^k] \geq 0.$ (k:偶数)　0 以上の関数の積分だから
- $\mathrm{V}[X] \geq 0, \mathrm{V}[aX + b] \geq 0.$　0 以上の関数の積分だから
- $\mathrm{E}[X^0] = 1.$　全確率. 確率密度関数の定義.
- $0 \leq P(\text{条件}) \leq 1.$　確率だから.

ぜんぶ, 積分の性質を使って不等号をつないでいって証明できます.

母平均値が重心, 母分散が幅 ^{2} であることから来る検算方法

$a < x < b$ でだけ, 確率密度関数 $f(x) > 0$ であるなら

- $a < \mathrm{E}[X] < b.$　板の外がつりあいポイントになることないでしょ
- $\mathrm{V}[X]^(1/2) \leq b - a.$　重心からの距離が板の幅 $b - a$ より大きくなることないでしょ

これも, 積分の性質を使って不等号をつないでいって証明できます.

多次元の連続型確率分布 岩薩林 確率・統計 §4.6

2 次元の同時確率密度関数 $f_{\mathrm{XY}}(x, y)$.

母期待値の定義

離散型確率変数　$\mathrm{E}[g(X, Y)] = \sum_{x} \sum_{y} g(x, y) \cdot p_{\mathrm{XY}}(x, y)$

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

周辺分布の確率密度関数 $f_{\mathrm{X}}(x) = \int_{-\infty}^{+\infty} f_{\mathrm{XY}}(x, y)\mathrm{d}y,\ f_{\mathrm{Y}}(y) = \cdots$.

X, Y が独立 $\Leftrightarrow f_{\mathrm{XY}}(x, y) = f_{\mathrm{X}}(x) f_{\mathrm{Y}}(y)$.

ここまで来たよ

8. 二項分布と独立同分布にしたがう確率変数の和

9. 連続型確率変数
   - 連続型確率変数
   - 連続型一様分布
   - 確率変数の 1 次式による変換と標準化

連続型一様分布 岩薩林 確率・統計 例題 4.1(p.78)

連続型一様分布 $\mathrm{U}(c,d)$

確率変数 X の確率密度関数が次で与えられるとき, X は区間 $[c,d)$ の連続型一様分布 $\mathrm{U}(c,d)$ に従う ($X \sim \mathrm{U}(c,d)$) という.

$$f(x) = \frac{1}{d-c}\mathrm{I}_{[c \le x < d]}(x) = \begin{cases} \frac{1}{d-c} & (c \le x < d) \\ 0 & (\text{他}) \end{cases}$$

L09-Q2
TA Prob and Sol:連続型一様分布

連続型確率変数 X が連続型一様分布 $\mathrm{U}(c,d)$ にしたがう.

1. $\mathrm{E}[X]$ を求めよう.
2. $\sqrt{\mathrm{V}[X]}$ を求めよう.

略解

$$\mathrm{E}[X^k] = \frac{1}{d-c}\int_c^d x^k \, \mathrm{d}x = \frac{1}{k+1}\frac{d^{k+1}-c^{k+1}}{d-c}.$$

1. $\mathrm{E}[X^1] = \frac{c+d}{2}$.
2. $\mathrm{V}[X] = \mathrm{E}[X^2] - \mathrm{E}[X^1]^2 = \frac{(d-c)^2}{12}$. $\sqrt{\mathrm{V}[X]} = \frac{d-c}{\sqrt{12}} \simeq \frac{d-c}{3.5}$.

$\mathrm{U}(c,d)$ に対するこの結果は, 公式のように記憶して使おう.

ここまで来たよ

8 二項分布と独立同分布にしたがう確率変数の和

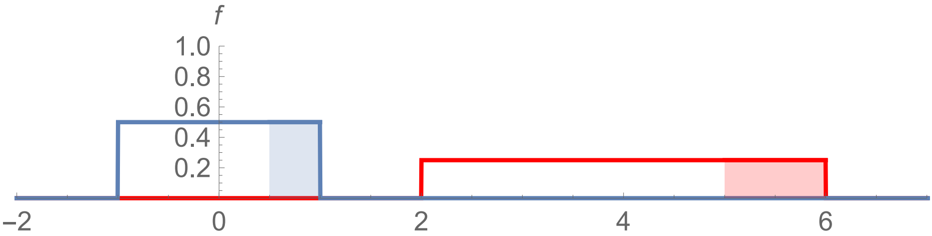
9 連続型確率変数
- 連続型確率変数
- 連続型一様分布
- 確率変数の 1 次式による変換と標準化

L09-Q3

Quiz(連続型一様分布の母期待値)

$X \sim \mathrm{U}(-1, 1)$ とする．次を求めよう．

1. $\mathrm{E}[2X + 4]$
2. $\mathrm{V}[2X + 4]$
3. $P(2X + 4 > 5)$



左から $X \sim \mathrm{U}(-1,1)$, $Y = aX + b = 2X + 4 \sim \mathrm{U}(2,6)$.

岩薩林 確率・統計 例題 4.4

$$f_{\mathrm{Y}}(y) = \frac{1}{a} \times f_{\mathrm{X}}\left(\tfrac{y-b}{a}\right) = \frac{1}{a} \times \begin{cases} \frac{1}{2} & \left(-1 \leq \frac{y-b}{a} < 1\right) \Leftrightarrow (b - a \leq y < b + a) \\ 0 & (\text{他}) \end{cases}$$

確率変数の標準化 岩薩林 確率・統計 例題 4.4(p.80)

任意の確率変数 X に対して, $\mu = \mathrm{E}[X], \sigma^2 = \mathrm{V}[X], \sigma > 0$ とする.
確率変数 Z を $Z = \frac{X-\mu}{\sigma}$ と定めると, $\mathrm{E}[Z] = 0$, $\mathrm{V}[Z] = 1$ となる.

Z は標準化された確率変数

$Y \sim \mathrm{U}(2,6)$ を標準化すると, $Z = \dfrac{Y-\frac{2+6}{2}}{\frac{6-2}{\sqrt{12}}} = \dfrac{Y-4}{4/\sqrt{12}} \sim \mathrm{U}(-\sqrt{3}, +\sqrt{3})$.

$X \sim \mathrm{U}(c,d)$ を標準化すると,$Z \sim \mathrm{U}(-\sqrt{3}, +\sqrt{3})$

L09-Q4

Quiz(連続型一様分布)

連続型確率変数 $Z \sim \mathrm{U}(-\sqrt{3}, \sqrt{3})$ に対して, $X = 2Z + 1$ を考える.

1. X の確率密度関数 $f_X(x)$ とそのグラフを答えよう.
2. $\mathrm{E}[X]$ を求めよう.
3. $\mathrm{V}[X]$ を求めよう.