統計学I

早稲田大学政治経済学術院
西郷 浩

本日の目標

- 確率変数
  - 確率変数とは
  - 確率変数の例
- 確率変数の性質
  - 確率変数の確率分布
  - 確率変数の期待値
  - 確率変数の分散

確率変数とは

- 確率変数 : X
  - 試行 : 偶然によって結果が定まる。
    - 例 : サイコロを1つ投げる。
      - 起こりうる結果 :
        1の目が出る、2の目が出る、…、6の目が出る。
  - 確率変数 : 試行の結果によって値が定まる。
    - 例 : サイコロを1つ投げるとき、
      - $X = 1$（もし出目が偶数なら）; 0（それ以外のとき）
      - X = （出目の2倍）

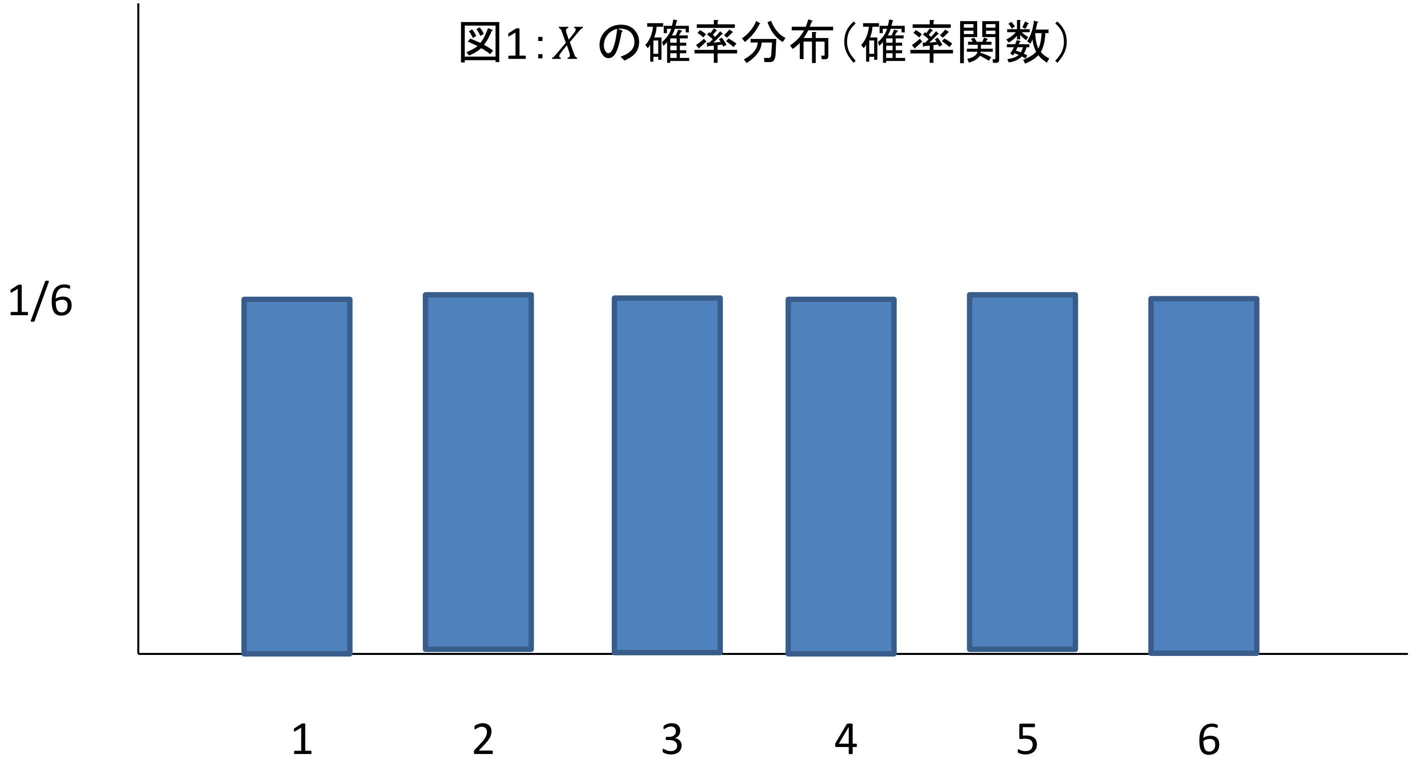
確率変数の例

- 確率変数の分類
  - 離散型確率変数
    - $X=$（サイコロの出目）
    - $X=$（一定時間内にかかってくる電話の回数）
    - $X=$（最初に表が出るまでのコイントスの回数）
  - 連続型確率変数
    - $X=$（100メートル走のタイム）
    - $X=$（ソフトボールの遠投の飛距離）

確率変数の確率分布(1)

- 離散型確率変数の確率関数
  - $p_X(x) = P(X = x)$
    - 例1: X =（1つのサイコロを投げた時の出目）
      - $p_X(1) = P(X = 1) = \frac{1}{6}$
      - $p_X(2) = P(X = 2) = \frac{1}{6}$
      - …
      - $p_X(6) = P(X = 6) = \frac{1}{6}$

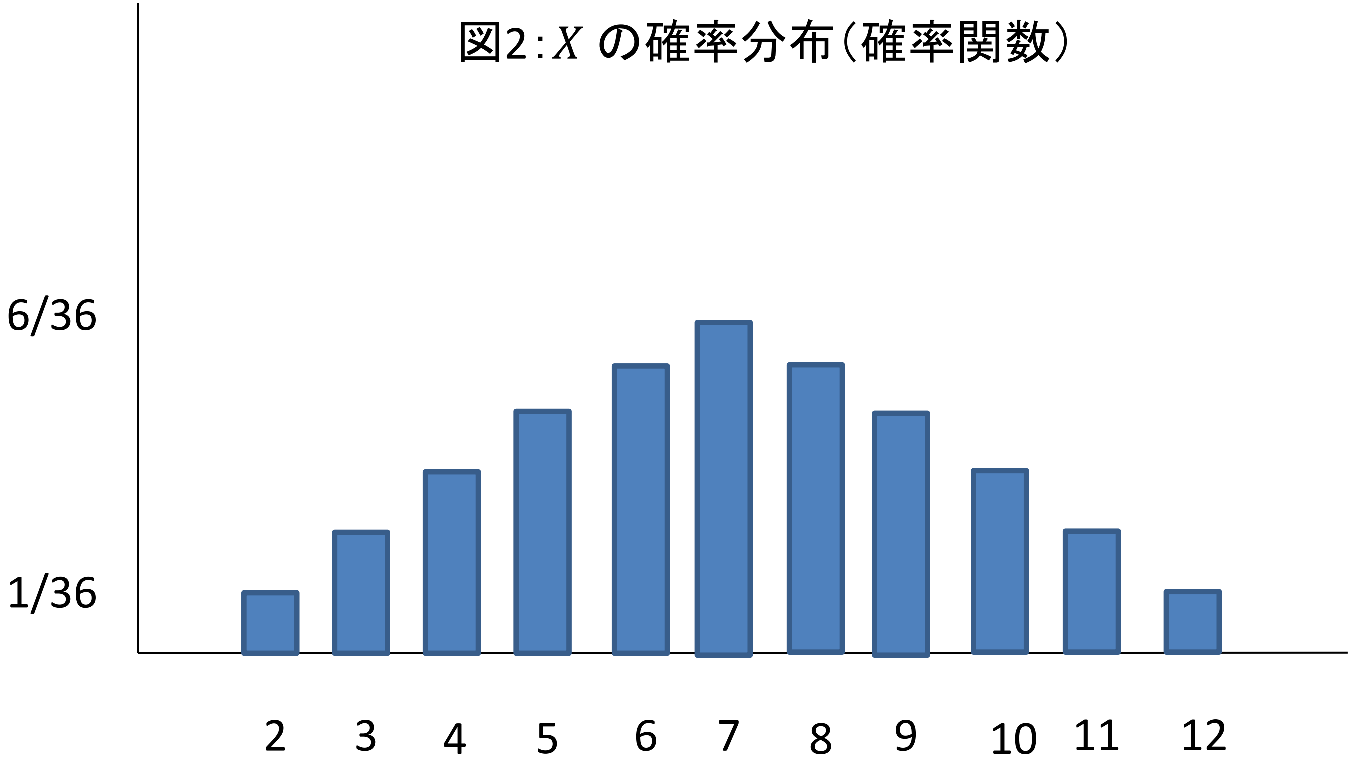
確率変数の確率分布(2)



確率変数の確率分布(3)

- 例2 : X = （サイコロを2つ投げたときの出目の和）
  - $p_X(2) = P(X = 2) = \frac{1}{36}$
  - $p_X(3) = P(X = 3) = \frac{2}{36}$
  - …
  - $p_X(7) = P(X = 7) = \frac{6}{36}$
  - $p_X(8) = P(X = 8) = \frac{5}{36}$
  - …
  - $p_X(12) = P(X = 12) = \frac{1}{36}$

確率変数の確率分布(4)

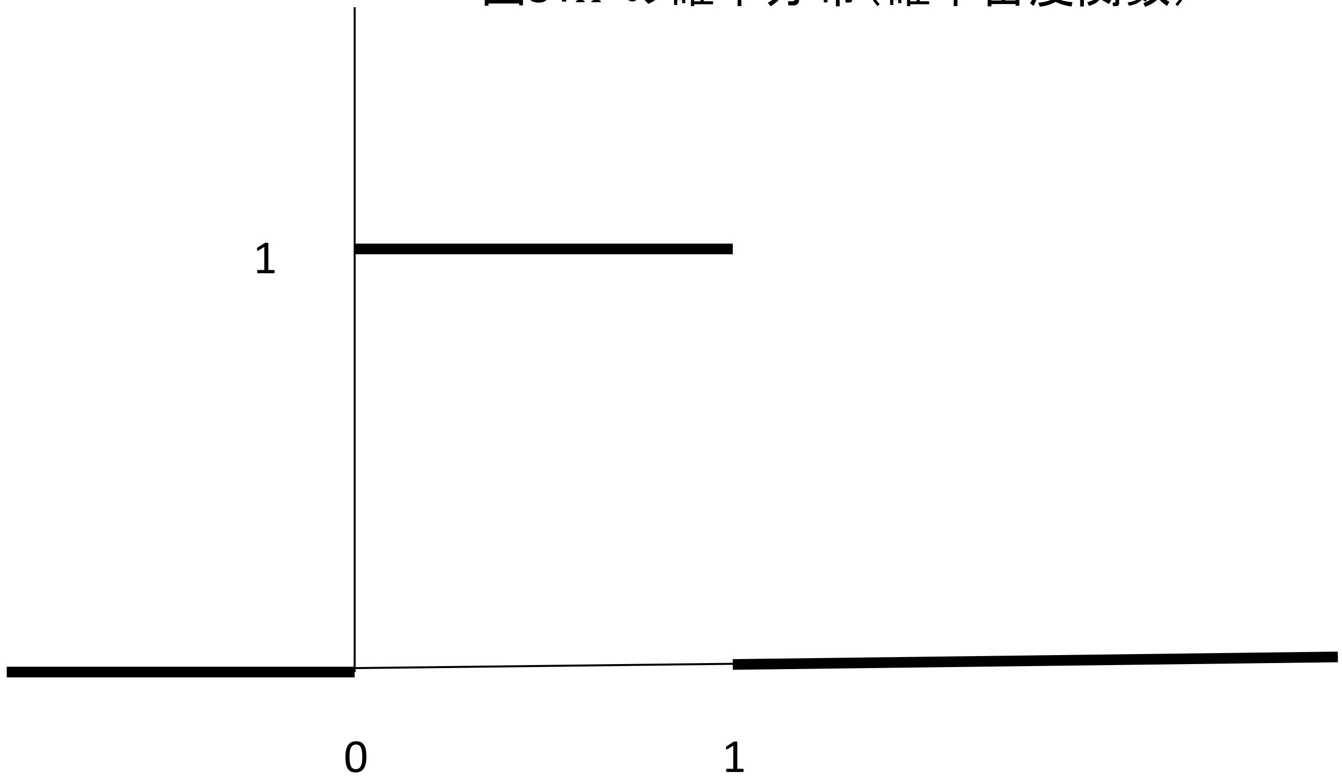


確率変数の確率分布(5)

- 連続型確率変数の確率密度関数
  - $f_X(x)$: つぎのような性質をもつ関数。
    $P(a \leq X \leq b) = \int_a^b f_X(x)dx$
    - 例 : X が、0と1の間で一様に出やすい確率変数であるときの確率密度関数
      - $f_X(x) = \begin{cases} 0 & x < 0 \\ 1 & 0 \leq x \leq 1 \\ 0 & 1 < x \end{cases}$

確率変数の確率分布(6)

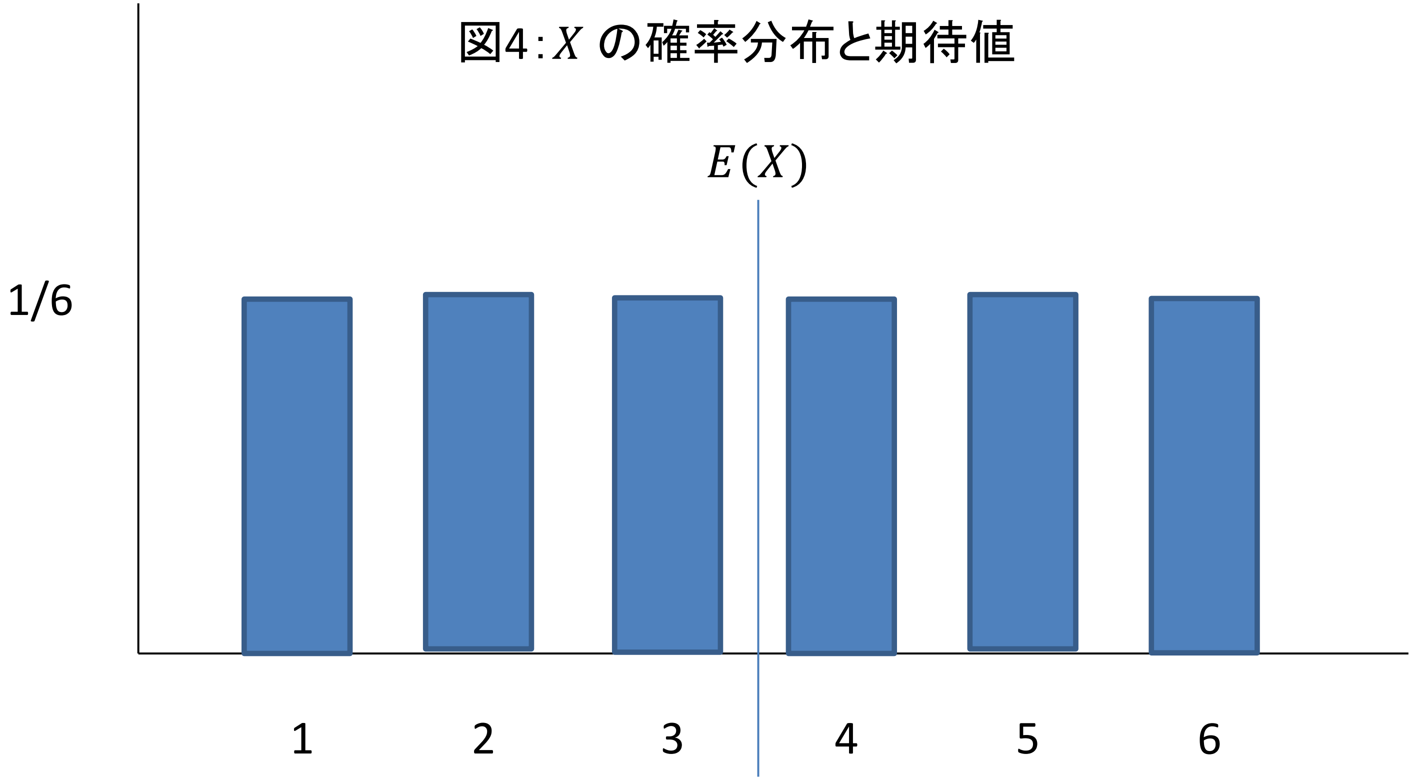
図3：X の確率分布（確率密度関数）



確率変数の期待値(1)

- 期待値：確率変数の平均的な値
  - 離散型確率変数：$E(X) = p_1 x_1 + p_2 x_2 + \cdots$
    - 例：$X =$（1つのサイコロを投げた時の出目）
      - $E(X) = (1/6) \times 1 + (1/6) \times 2 + (1/6) \times 3$
        $+(1/6) \times 4 + (1/6) \times 5 + (1/6) \times 6$
        $= 7/2$
      - 1つのサイコロを60,000回投げたとする。およそ10,000回1がでて、およそ10,000回2が出て、…、およそ10,000回6が出るであろう。したがって、出目の算術平均が期待値におおよそ等しくなる。

確率変数の期待値(2)

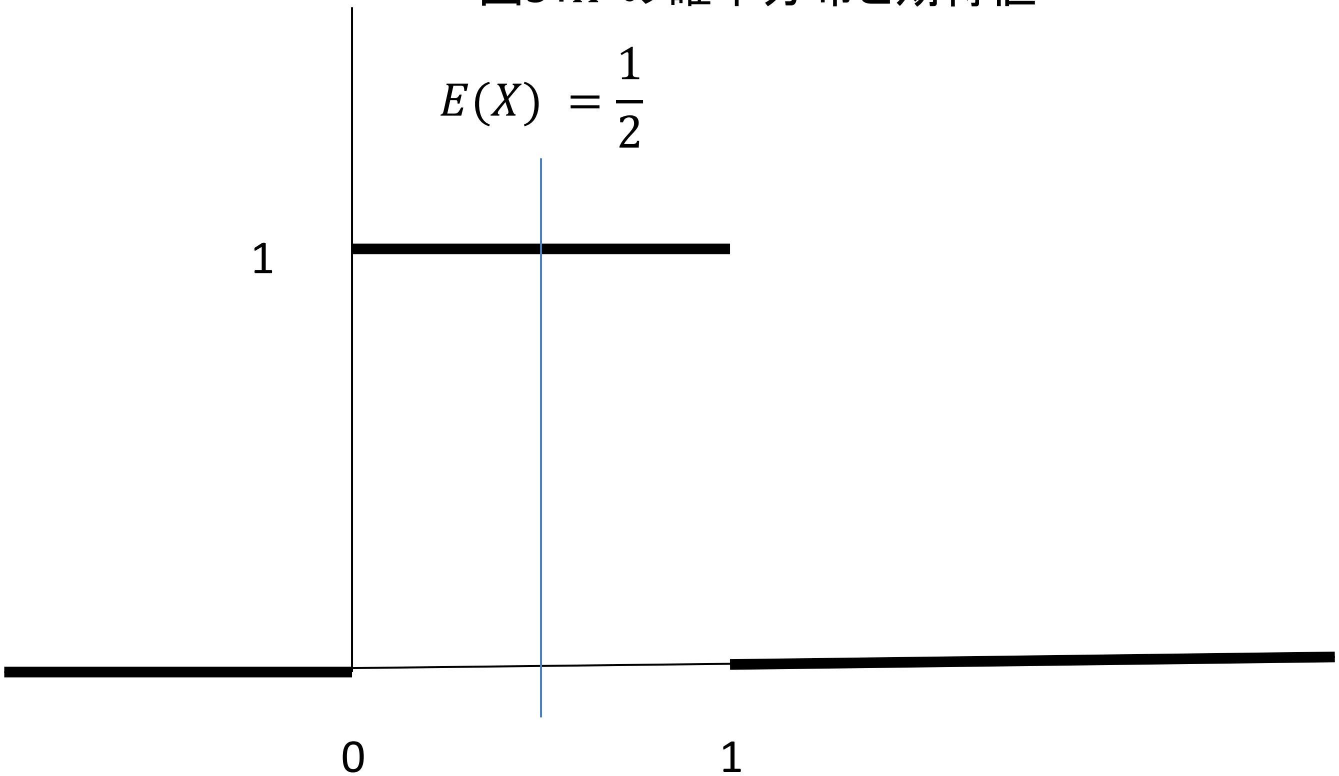


確率変数の期待値(3)

- 連続型確率変数:$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx$
  - 例:X が、0と1の間で一様に出やすい確率変数
    - $E(X) = \int_0^1 x dx = \frac{1}{2}$
    - 0から1までの値が一様に発生する。したがって、実際にX を10,000回発生させると、0から1までの値がまんべんなく発生するだろう。その算術平均は1/2におおよそ等しくなる。

確率変数の期待値(4)

図5 : X の確率分布と期待値

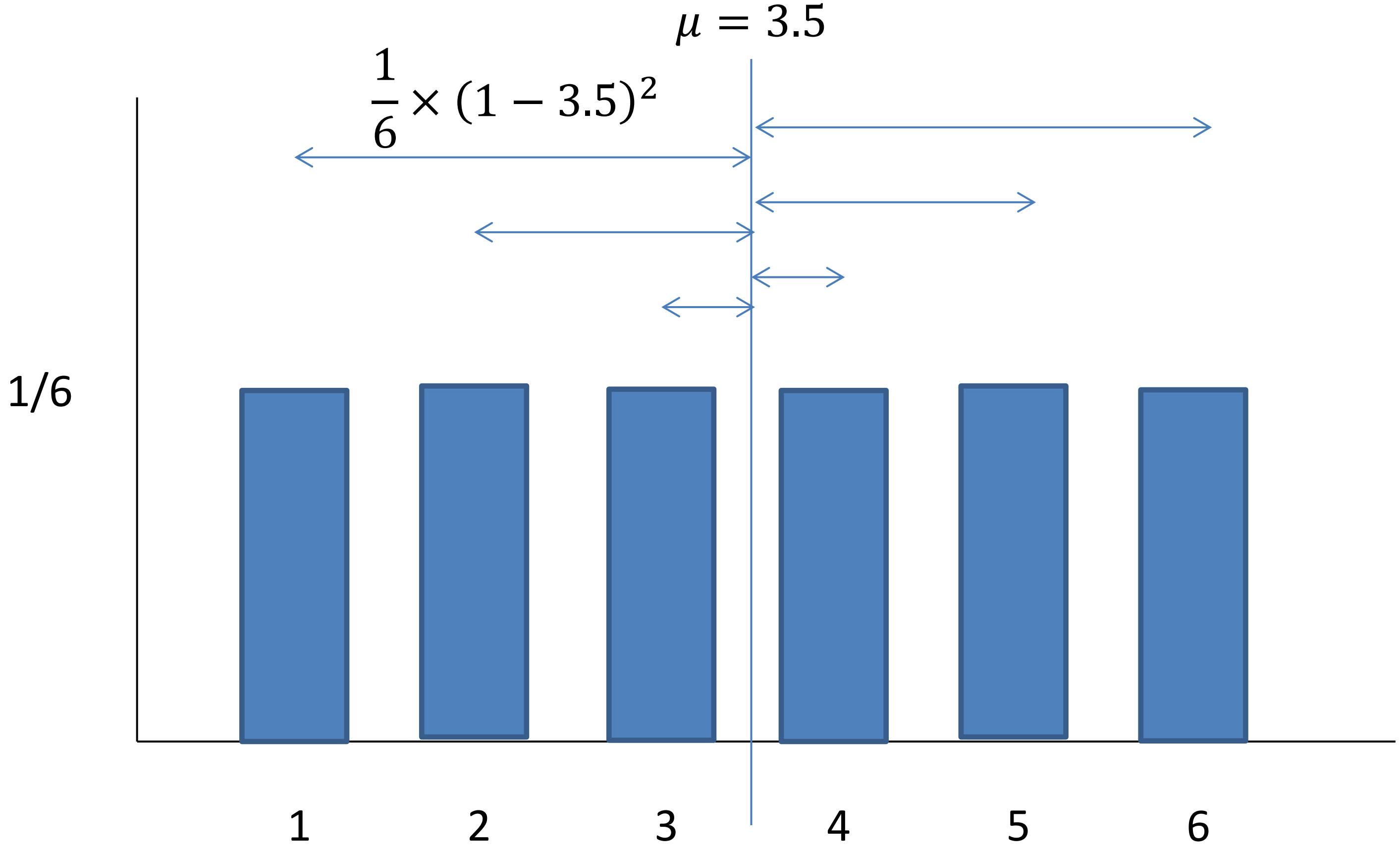


確率変数の分散(1)

- 分散:確率変数の平均的な散らばり
  - 離散型確率変数:
    - $V(X) = p_1(x_1 - \mu)^2 + p_2(x_2 - \mu)^2 + \cdots$
      - ただし、$\mu = E(X)$
    - 例: $X =$(1つのサイコロを投げた時の出目)
      - $V(X) = (1/6) \times (1 - 7/2)^2 + (1/6) \times (2 - 7/2)^2$
        $+(1/6) \times (3 - 7/2)^2 + (1/6) \times (4 - 7/2)^2$
        $+(1/6) \times (5 - 7/2)^2 + (1/6) \times (6 - 7/2)^2$
        $= 35/12$
    - 分散をあらわす記号:$\sigma^2 = V(X)$

確率変数の分散(2)

図5：X の確率分布と分散



確率変数の分散(3)

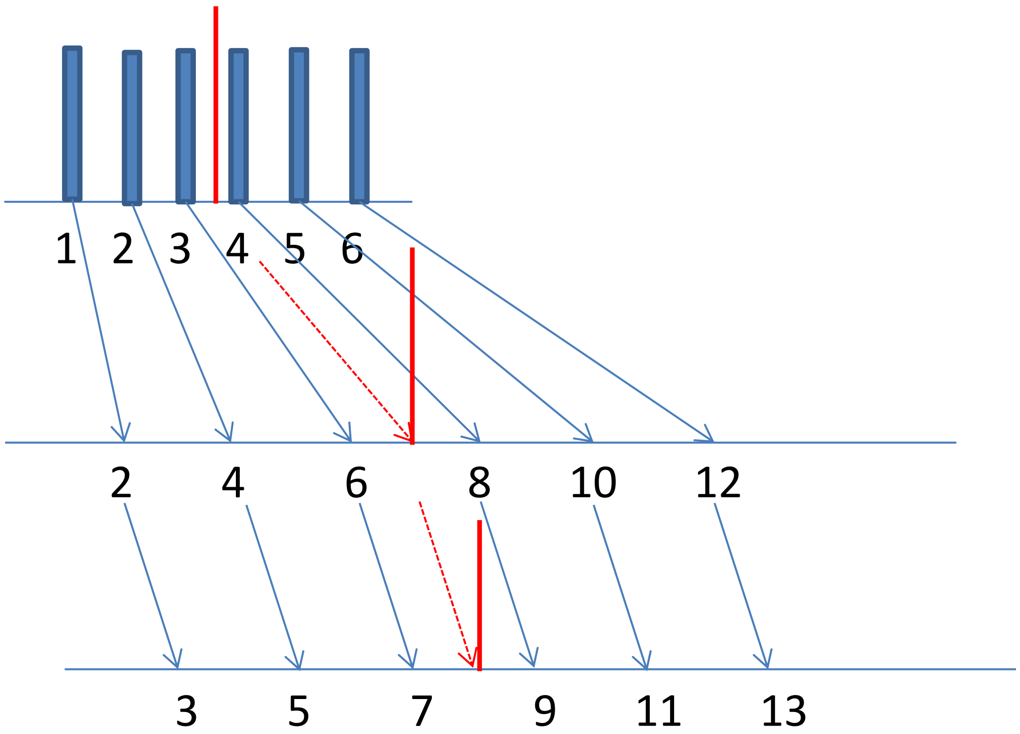
- 連続型確率変数:
  - $V(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f_X(x)\, dx$
    - ただし、$\mu = E(X)$
  - 例: X が、0と1の間で一様に出やすい確率変数
    - $V(X) = \int_0^1 (x - 1/2)^2 dx = 1/12$

期待値の性質(1)

- 直線による新しい確率変数
  - $Y = a + bX$　ただし、a と b は定数。
    - 例: $X =$（1つのサイコロを投げた時の出目）
      - $Y = 1 + 2X$
  - $E(Y) = a + bE(X)$
    - 例: $E(Y) = 1 + 2 \times (7/2) = 8$
      - Y の確率分布を求めて定義式どおりに計算しなくても、Xの期待値から Y の期待値を直接計算できる。

期待値の性質(2)

– 図による説明
  - $X \rightarrow 2X \rightarrow 1+2X$
  - (1)
  - (2)
  - (3)

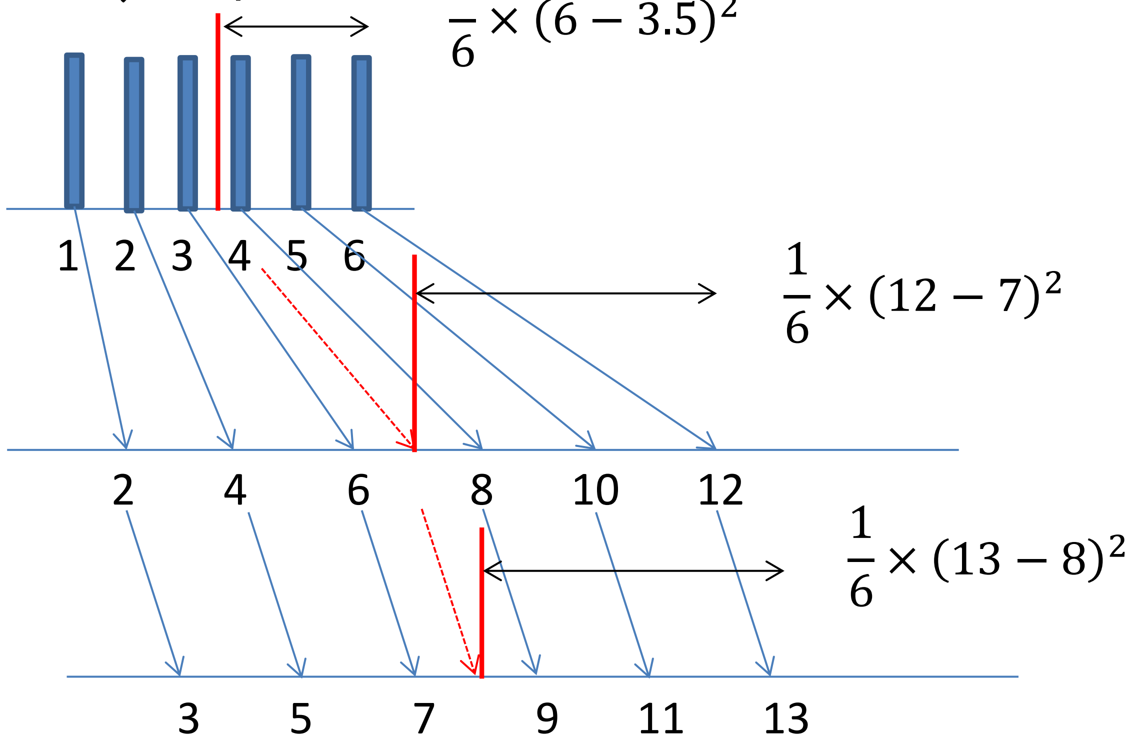


分散の性質(1)

- 直線による新しい確率変数
  - $Y = a + bX$　ただし、a と b は定数。
    - 例：X = （1つのサイコロを投げた時の出目）
      - $Y = 1 + 2X$
  - $V(Y) = b^2 V(X)$
    - 例：$V(Y) = 2^2 \times (35/12) = 35/3$
      - Y の確率分布を求めて定義式どおりに計算しなくても、Xの分散から Y の分散を直接計算できる。

分散の性質(2)

– 図による説明
  - $X \to 2X \to 1 + 2X$
  - (1)
  - (2)
  - (3)



確率変数の標準化

- 直線による変換によって、確率変数の期待値と分散を、それぞれ、0, 1に変更する。
  - $Z = \frac{X-\mu}{\sqrt{\sigma^2}} = \frac{X-\mu}{\sigma}$
    - 標準偏差:$\sigma = \sqrt{\sigma^2}$
  - 例:$X =$（1つのサイコロを投げた時の出目）
    - $Z = \frac{X-7/2}{\sqrt{35/12}}$
      - Z の期待値と分散が、それぞれ、0, 1 になることを確かめよ。

練習問題

- 問題1：
  - コインを1つ投げる。X を表が出たら1、そうでないとき0となる確率変数とする。X の確率関数と期待値、分散を求めなさい。
- 問題2：
  - X ＝（1つのサイコロを投げた時の出目）とする。このとき、$Y = 1 + 2X$の確率関数を求めなさい。確率関数からY の期待値と分散を求めなさい。

解答:問題1

- $E(X) = \frac{1}{2} \times 0 + \frac{1}{2} \times 1 = \frac{1}{2}$
- $V(X) = \frac{1}{2} \times \left(0 - \frac{1}{2}\right)^2 + \frac{1}{2} \times \left(1 - \frac{1}{2}\right)^2 = \frac{1}{4}$

解答:問題2

表1:XとYの確率分布

X	1	2	3	4	5	6
確率	1/6	1/6	1/6	1/6	1/6	1/6
Y	3	5	7	9	11	13

$$E(Y)=\frac{1}{6}\times 3+\frac{1}{6}\times 5+\frac{1}{6}\times 7+\frac{1}{6}\times 9+\frac{1}{6}\times 11+\frac{1}{6}\times 13=8=1+2E(X)$$

$$\begin{aligned}V(Y)=&\frac{1}{6}\times(3-8)^2+\frac{1}{6}\times(5-8)^2+\frac{1}{6}\times(7-8)^2+\frac{1}{6}\times(9-8)^2\\&+\frac{1}{6}\times(11-8)^2+\frac{1}{6}\times(13-8)^2=\frac{35}{3}=2^2V(X)\end{aligned}$$

確率変数の期待値と（確率変数の）観察値の算術平均

- 観察値の算術平均→確率変数の期待値
  - 大数の法則
    - サイコロの出目の期待値=3.5
    - サイコロを多数回投げて計算した算術平均→3.5

確率変数の分散と（確率変数の）観察値の分散

- 観察値の分散→確率変数の分散
  - 大数の法則
    - サイコロの出目の分散=35/12
    - サイコロを多数回投げて計算した分散→35/12