

## 統計学 01

早稲田大学政治経済学部  
第4回  
西郷 浩

1

## 本日の目標

- 第6章 確率分布
  - 離散型確率変数の確率分布の代表例
    - 二項分布
    - ポアソン分布
  - 連続型確率変数の確率分布の代表例
    - 正規分布
- 学習に当たっての注意
  - ユーザーの立場からすれば、
    - 数学的な導出の過程はそれほど重視しなくて良い。
    - 「便利な性質が知られている」という事実が重要である。

2

## 二項分布(1)

- ベルヌーイ試行
  - 試行の結果: (S, F) 2値
  - 発生確率: 一定
    - $P(\{S\text{が発生する}\})=p$ ;  $P(\{F\text{が発生する}\})=1-p$
  - 例: ひとつのサイコロを振る:
    - $S=\{1\text{の目が出る}\}$ ,  $F=\{\text{それ以外の目が出る}\}$
- 独立試行
  - 複数の試行において、試行の結果が相互に独立。
    - 例:  $P(\{S@試行2\}|\{S@試行1\}) = P(\{S@試行2\})$

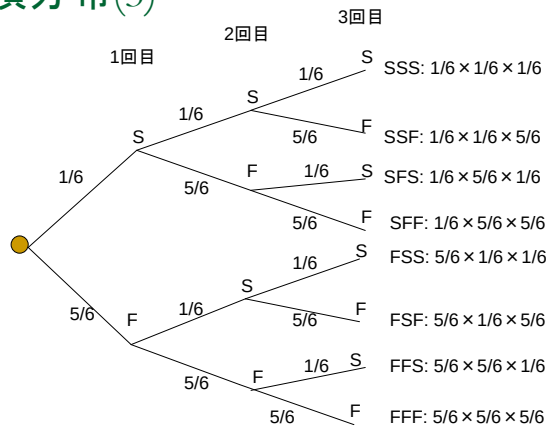
3

## 二項分布(2)

- 二項分布
  - $X = (n \text{ 回の独立ベルヌーイ試行において、} S \text{ が発生する回数})$
  - 確率変数  $X$  の確率分布は？
- 考察を具体化するために
  - $X = (3 \text{ 回サイコロを振って、} 1 \text{ の目が出た回数})$  として、 $P(X=1)=f(1)$  を求めてみる。
  - 考察の方針
    - まず、樹形図で確率を計算し、それを眺めて体系的な計算方法を考える。

4

## 二項分布(3)



5

## 二項分布(4)

- $f(1) = P(X=1)$   
 $= P(\{\text{SFFが発生}\} \cup \{\text{FSFが発生}\} \cup \{\text{FFSが発生}\})$   
 $= P(\{\text{SFFが発生}\}) + P(\{\text{FSFが発生}\}) + P(\{\text{FFSが発生}\})$   
 $= 1/6 \times 5/6 \times 5/6 + 5/6 \times 1/6 \times 5/6 + 5/6 \times 5/6 \times 1/6$   
 $= 3 \times (1/6 \times 5/6 \times 5/6)$   
 $= (\text{3回の試行のうちSが1回生じるパターンの数})$   
 $\quad \times (\text{おのおののパターンが生じる確率})$   
 $= (\text{3つの番号札から1枚取ってくる組み合わせの数})$   
 $\quad \times (\text{おのおののパターンが生じる確率})$
- この考え方は一般化できる。

6

## 二項分布(5)

### ■ 一般の場合

- $X = (n \text{ 回の独立ベルヌーイ試行において、Sが発生する回数})$
- $f(x) = P(X=x)$  (確率変数  $X$  が、ある値  $x$  をとる確率)  
 $= P(n \text{ 回の独立ベルヌーイ試行において、Sが} x \text{ 回発生する})$   
 $= (n \text{ の試行のうち } x \text{ 回のSが生じるパターンの数})$   
 $\quad \times (\text{おのおののパターンの発生確率})$   
 $= (n \text{ 枚の番号札から } x \text{ 枚とってくる組み合わせの数})$   
 $\quad \times (\text{おのおののパターンの発生確率})$   
 $= {}_n C_x p^x (1-p)^{n-x}$

7

## 二項分布(6)

- ${}_n C_x = n! / \{x!(n-x)!\}$  ただし、 $n! = n(n-1)\dots 2 \times 1$ 
  - 公式そのものではなくて、「組み合わせの数は計算できる」という事実が重要である。

### ■ 二項分布にしたがう確率変数 $X$ の性質

$$\sum_{x=0}^n f(x) = \sum_{x=0}^n {}_n C_x p^x (1-p)^{n-x} = \{p + (1-p)\}^n = 1$$

$$E(X) = np$$

$$V(X) = np(1-p)$$

8

## 二項分布(7)

### ■ 例

- 不良品率が 0.005 であるような製造工程で、1日に2000個の製品を作っているとする。
  - 1日の不良品数の期待値は？
  - 1日の不良品数が2以下である確率は？

### □ 解答

- S(成功!)が発生する確率を 0.005 とする独立ベルヌーイ試行が2000回おこなわれると考える。X を不良品数とする。
- $E(X) = 2000 \times 0.005 = 10$
- $P(X \leq 2) = P(X=0) + P(X=1) + P(X=2)$   
 $= 1 \times (1-0.005)^{2000} + 2000 \times 0.005 \times (1-0.005)^{1999}$   
 $+ 1999000 \times 0.005^2 \times (1-0.005)^{1998} = 0.002724205$

9

## ポアソン分布(1)

### ■ 独立なベルヌーイ試行において

- $n$  が非常に大きい。
- Sの発生確率  $p$  が非常に小さい。
- $np \rightarrow \lambda$  ( $n \rightarrow \infty$  のとき)
- このとき、Sの発生回数  $X$  の確率分布  $f(x)$  はポアソン分布にしたがう。
  - $f(x) = e^{-\lambda} \lambda^x / x!$
- 例: 事故の発生件数

10

## ポアソン分布(2)

### ■ ポアソン分布にしたがう確率変数 $X$ の性質

$$\sum_{x=0}^{\infty} f(x) = \sum_{x=0}^{\infty} e^{-\lambda} \lambda^x / x! = e^{-\lambda} \sum_{x=0}^{\infty} \lambda^x / x! = e^{-\lambda} e^{\lambda} = 1$$

$$E(X) = \sum_{x=0}^{\infty} x f(x) = \sum_{x=1}^{\infty} x f(x) = \lambda \sum_{x=1}^{\infty} e^{-\lambda} \lambda^{x-1} / (x-1)! = \lambda$$

$$E\{X(X-1)\} = \sum_{x=0}^{\infty} x(x-1) f(x) = \sum_{x=2}^{\infty} x(x-1) f(x) = \lambda^2$$

$$V(X) = E(X^2) - \{E(X)\}^2 = E\{X(X-1)\} + E(X) - \{E(X)\}^2 = \lambda$$

### ■ ポアソン分布の適用例

- 教科書 表6.2-6.5(p. 116)

11

## 正規分布(1)

### ■ 正規分布にしたがう確率変数 $X$ の確率分布

$$P(X \leq x) = F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(u-\mu)^2}{2\sigma^2}} du = \int_{-\infty}^x f(u) du$$

### ■ 正規分布の密度関数 $f(x)$ の形状

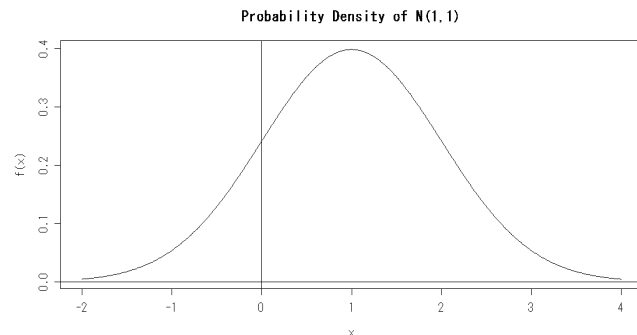
- $x = \mu$  を中心とした左右対称の釣鐘型の分布

### ■ 正規分布の表記

- 2つのパラメータ  $\mu, \sigma^2$  が決まれば一意に決まる。
  - $X \sim N(\mu, \sigma^2)$  と表記することが多い。

12

## 正規分布(2)



13

## 正規分布(3)

- $X \sim N(\mu, \sigma^2)$  である確率変数  $X$  の性質

$$\int_{-\infty}^{\infty} f(u) du = 1$$

$$E(X) = \mu \quad (\mu \text{ は期待値に等しい})$$

$$V(X) = \sigma^2 \quad (\sigma^2 \text{ は分散に等しい})$$

$$Y = aX + b \rightarrow Y \sim N(a\mu + b, a^2\sigma^2)$$

$$\text{とくに } Z = \frac{X - \mu}{\sigma} \sim N(0, 1) \quad (\text{標準正規分布})$$

14

## 正規分布(4)

- $Z \sim N(0, 1)$  の(累積)分布関数

$$\Phi(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$$

- 教科書(p.280) は上側確率  $1 - \Phi(z)$  を掲載している。

- 覚えておくと便利(教科書の pp. 121-122 も参照)

$X \sim N(\mu, \sigma^2)$  のとき、

$$P(\mu - 1.96\sigma \leq X \leq \mu + 1.96\sigma) = 0.95$$

$$P(\mu - 1.645\sigma \leq X) = P(X \leq \mu + 1.645\sigma) = 0.95$$

15

## 正規分布(5)

- 例:

- Aさんの通学時間(分)は期待値60、分散100の正規分布にしたがうという。最初の講義の開始時間までに間に合う確率を0.95にするためには、講義開始時間の何分前に家を出る必要があるか。

- 考え方の整理

- Aさんは  $a$  分前に家を出たとする。
- 実際の通学時間を  $X$  とすれば、 $X \sim N(60, 100)$
- $X \leq a$  であれば間に合うことになる。
- $P(X \leq a) = 0.95$  となる  $a$  を求めればよい。

16

## 正規分布(6)

$$P(X \leq a) = P\left(\frac{X - 60}{10} \leq \frac{a - 60}{10}\right)$$

$$\frac{X - 60}{10} \sim N(0,1) \text{なので}$$

$$P\left(\frac{X - 60}{10} \leq 1.645\right) = 0.95$$

したがって、

$$\frac{a - 60}{10} = 1.645 \Leftrightarrow a = 60 + 1.645 \times 10 = 76.45$$

17

## 練習問題 (1)

### ■ 二項分布(2番はベイズの定理を使う)

- 2つの壺(外見では区別できない)がある。
  - 壺Aには赤玉70個と白球30個
  - 壺Bには赤玉30個と白球70個
- 最初に壺をひとつ選ぶ(どちらかはわからない)。
- 「壺からひとつ玉を取り出す→色を記録する→玉をもどす」という実験を12回繰り返した。
- 結果は、赤8回・白4回であった。
- 問題
  1. 「最初に選んだ壺がAである」という条件のもとで「赤8回・白4回観察される」という事象が発生する確率を求めよ。
  2. 「赤8回・白4回観察される」という条件のもとで、取り出した壺がAである確率を求めよ。

18

## 練習問題 (2)

### ■ 正規分布

- 次の講義の教室の移動時間(分)  $X$  が、平均8、分散1にしたがうとする。休み時間を10分とする。
  1. 前の講義が時間通りに終わった場合、次の講義の開始時間までに教室に到着できる確率を求めよ。
  2. 前の講義が4分延長された場合、次の講義に遅刻する確率を求めよ。ただし、次の講義は刻限に始まるとする。
  3. 次の講義の開始時間に間に合う確率を0.99とするためには、前の講義の延長は何分までにすべきか。
  4. 全学的な規則の変更で休み時間が15分になったとする。上の1の答えはどのように変わるか。

19