

統計学 01

早稲田大学政治経済学部
第3回
西郷 浩

1

本日の目標

- 第5章 確率変数
 - 5.1 確率変数と確率分布
 - 5.2 確率変数の期待値と分散
 - 5.3 モーメント(とモーメント母関数)
 - 5.4 チェビシェフの不等式

2

確率変数と確率分布(1)

- 確率変数 X
 - 偶然によって値が決まる変数
 - 例: $X =$ (コイントスの結果)
- 確率変数の種別
 - 離散型確率変数: 確率変数の取りうる値が可算
 - $X =$ (サイコロの出目)
 - $X =$ (コイントスで、初めて表が出るまでの回数)
 - 連続型確率変数: 確率変数の取りうる値が連続的
 - $X =$ (ある人の 100m 走 のタイム)
 - $X =$ (ある人の 遠投距離)

3

確率変数と確率分布(2)

- 確率変数の確率分布
 - 確率変数の出方(確率的な挙動)をあらわしたもの。
- 離散型確率変数 X の確率分布
 - $P(X = x_k) = f(x_k)$ ($k=1,2,\dots$) (f を確率関数と呼ぶ)
 - 例: $X =$ (サイコロの出目)
 - $P(X=1)=f(1)=1/6, P(X=2)=f(2)=1/6 \dots, P(X=6)=f(6)=1/6$
- 連続型確率変数 X の確率分布
 - $P(a \leq X \leq b) = \int_{[a,b]} f(x) dx$ (f を確率密度関数と呼ぶ)
 - 例: $X =$ (電球の寿命までの期間)
 - $P(a \leq X \leq b) = \int_{[a,b]} \lambda e^{-\lambda x} dx$ (指数分布)

4

確率変数と確率分布(3)

■ (累積)分布関数:

- $F(x) = P(X \leq x)$
 - 確率変数 X が、ある値 x 以下の値を取る確率
- 分布関数の形状 (図5.6-5.9、教科書 pp. 92-94)
 - 離散型確率変数: 階段関数
 - 連続型確率変数: 連続関数

$$F(x) = \begin{cases} \sum_{u \leq x} f(u) & (\text{離散型確率変数}) \\ \int_{-\infty}^x f(u) du & (\text{連続型確率変数}) \end{cases}$$

5

確率変数と確率分布(4)

□ 分布関数の性質

- 単調性: $x_1 < x_2$ ならば $F(x_1) \leq F(x_2)$
- $F(-\infty) = 0, F(\infty) = 1$
- 右連続

■ モードとメディアン

- モード x_0 :
 - $f(x)$ が最大になるところに対応する x
- メディアン x_m :
 - $P(X \leq x) = F(x) \geq 1/2$ が成立する x の下限

6

確率変数の期待値と分散(1)

■ 確率変数 X の期待値 (平均)

$$E(X) = \begin{cases} \sum_k x_k f(x_k) & (\text{離散型確率変数}) \\ \int_{-\infty}^{\infty} x f(x) dx & (\text{連続型確率変数}) \end{cases}$$

■ 確率変数の関数 $\phi(X)$ の期待値

$$E(\phi(X)) = \begin{cases} \sum_k \phi(x_k) f(x_k) & (\text{離散型確率変数}) \\ \int_{-\infty}^{\infty} \phi(x) f(x) dx & (\text{連続型確率変数}) \end{cases}$$

7

確率変数の期待値と分散(2)

■ 確率変数 X の期待値の意味

- 確率変数 X が取る平均的な値
 - $X = (\text{サイコロの出目})$ とし、 X を多数(n)回観察する。
 - $X = 1$ となる回数: およそ $n/6$ 回
 - $X = 2$ となる回数: およそ $n/6$ 回
 - ...
 - $X = 6$ となる回数: およそ $n/6$ 回
 - X の観察値の平均値は、およそ
 - $(1 \times n/6 + 2 \times n/6 + \dots + 6 \times n/6) / n$
 $= 1 \times 1/6 + 2 \times 1/6 + \dots + 6 \times 1/6$
 $= E(X)$

8

確率変数の期待値と分散(3)

■ 確率変数 X の関数 $\phi(X)$ の期待値の意味

- 確率変数 X の関数 $\phi(X)$ が取る平均的な値
 - $X = (\text{サイコロの出目})$, $\phi(x) = x^2$ とし、 $\phi(X)$ を多数(n)回観察する。
 - $X^2 = 1$ となる回数: およそ $n/6$ 回
 - $X^2 = 4$ となる回数: およそ $n/6$ 回
 - $X^2 = 9$ となる回数: およそ $n/6$ 回
 - ...
 - $X^2 = 36$ となる回数: およそ $n/6$ 回
 - $\phi(X)$ の観察値の平均値は、およそ
 - $(1 \times n/6 + 4 \times n/6 + \dots + 36 \times n/6) / n = E(\phi(X))$

9

確率変数の期待値と分散(4)

■ 確率変数 X の期待値:

- X の確率分布の中心の位置を示す指標として用いられることが多い。

■ $\mu = E(X)$ と定義する。(μ : ミュー)

■ 期待値の性質

- $E(c) = c$ (定数 c の期待値)
- $E(X + c) = E(X) + c$
- $E(cX) = c E(X)$
- $E(X + Y) = E(X) + E(Y)$

これは、平均値(平均点)を想起すると覚えやすい。

10

確率変数の期待値と分散(5)

■ 確率変数 X の分散と標準偏差

$$V(X) = E\{(X - \mu)^2\} = \begin{cases} \sum_k (x_k - \mu)^2 f(x_k) & (\text{離散型確率変数}) \\ \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx & (\text{連続型確率変数}) \end{cases}$$

□ 便利な公式

$$\begin{aligned} V(X) &= E\{(X - \mu)^2\} = E\{X^2 - 2\mu X + \mu^2\} \\ &= E(X^2) - 2\mu E(X) + E(\mu^2) \\ &= E(X^2) - 2\mu^2 + \mu^2 \\ &= E(X^2) - \mu^2 \end{aligned}$$

11

確率変数の期待値と分散(6)

■ 確率変数 X の標準偏差

- $D(X) = \{V(X)\}^{1/2}$
 - 確率変数 X と同じ測定単位をもつ。

■ 確率変数 X の分散(標準偏差)の意味

- 中心の位置 μ からの平均的なズレ具合

■ 今後は、 $\sigma^2 = V(X)$ 、 $\sigma = D(X)$ と定義する。

■ 分散の性質

- $V(c) = 0$
- $V(X + c) = V(X)$
- $V(cX) = c^2 V(X)$

12

確率変数の期待値と分散(7)

■ 確率変数の標準化

$$Z = \frac{X - E(X)}{\sqrt{V(X)}} = \frac{X - \mu}{\sigma}$$

■ 標準化した確率変数 Z の性質

$$E(Z) = 0, \quad V(Z) = 1, \quad D(Z) = 1$$

- 「平均0、分散1」となるように変換しておくと、(とくに正規分布との関連で)あつかいやすい。

13

モーメント

■ 確率変数 X のモーメント(積率)

原点のまわりの r 次のモーメント $\mu_r = E(X^r)$

期待値のまわりの r 次のモーメント $\mu'_r = E\{(X - \mu)^r\}$

- X の期待値 μ : 原点のまわりの1次のモーメント
- X の分散 σ^2 : 期待値のまわりの2次のモーメント

■ 高次のモーメントを使った指標

歪度 (歪みの尺度) $\beta_3 = \alpha_3 = \mu'_3 / \sigma^3$

尖度 (尖りの尺度) $\beta_4 = \alpha_4 - 3 = \mu'_4 / \sigma^4 - 3$

14

チェビシェフの不等式

■ チェビシェフの不等式

$$P(|X - \mu| \geq k\sigma) \leq 1/k^2 \Leftrightarrow P(|X - \mu| < k\sigma) > 1 - 1/k^2$$

- 確率変数 X がどんな確率分布に従っていたとしても、 X の期待値 μ から標準偏差 σ の k 倍以上はなれたところに出現する確率は $1/k^2$ を絶対に超えない。
 - 証明:教科書 p.105.

15

練習問題

■ 確率変数

$X = (\text{サイコロの出目})$

- 期待値は?
- 分散は?

■ 確率変数

$Y = (\text{コインの結果: 表} \rightarrow 1; \text{裏} \rightarrow 0)$

- 期待値は?
- 分散は?

16