



第15回 6月16日の授業内容

□ § 5. 標本分布

- § 5.3 統計量の分布の近似
 - § 5.3.2 中心極限定理
 - § 5.3.3 大数の法則と中心極限定理: 再論

6/16/09

1



§ 5.3.2 中心極限定理

□ 大数の法則

- n のとき、標本平均が母集団平均 μ の一点分布になる。
- $\text{Var}(\bar{X}_n) = \sigma^2/n \xrightarrow{P} 0$ であるし、また $\text{Var}(n\bar{X}_n) = n\sigma^2 \xrightarrow{P} \infty$
- そこで、 $\text{Var}(\delta\bar{X}_n)$ が0にも収束しないし、発散もしない定数を探そう。

➡ $\delta = \sqrt{n}$ がその解。

6/16/09

2



§ 5.3.2 中心極限定理(2)

□ $\sqrt{n}\bar{X}_n$ の確率分布

- モーメント

$$E(\sqrt{n}\bar{X}_n) = \sqrt{n}\mu \quad \text{Var}(\sqrt{n}\bar{X}_n) = \sigma^2$$

□ 中心極限定理 Central Limit Theorem

- Lindeberg=Levyのバージョン
- X_1, X_2, \dots, X_n を互いに独立な平均 μ 、分散 σ^2 の確率変数とする。このとき、 $n \rightarrow \infty$ であれば、

$$\frac{\bar{X}_n - E(\bar{X}_n)}{\sqrt{\text{Var}(\bar{X}_n)}} \xrightarrow{d} N(0,1) \iff \frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} N(0,1)$$

6/16/09

3



§ 5.3.2 中心極限定理(3)

□ 中心極限定理が何故重要か

- 標本のデータから計算した統計量に基づいて推論を行う時の精度が評価できる(わかる)。
- X_i の確率分布を知らなくても、平均と分散が同一であることだけわかっていれば、標本平均の確率分布を正規分布で十分に近似できる。

6/16/09

4



§ 5.3.2 中心極限定理(4)

□ 二項分布の正規近似

- $X \sim \text{Bin}(n, p)$ とする。
- 二項分布の再生性から、ベルヌイ分布 $W_i \sim \text{Bin}(1, p)$ をもちいて、 $X = W_1 + \dots + W_n$ で表すことができる。

$\bar{W} = X/n$ とおくと、 $E(\bar{W}) = p$ 、 $\text{Var}(\bar{W}) = p(1-p)/n$ 。

中心極限定理から、

$$\frac{\bar{W} - p}{\sqrt{p(1-p)/n}} \xrightarrow{d} N(0,1)$$

よって十分に n が大ならば、 X を $N(np, np(1-p))$ で近似可能。

6/16/09

5



§ 5.3.2 中心極限定理(5)

□ 二項分布の正規近似

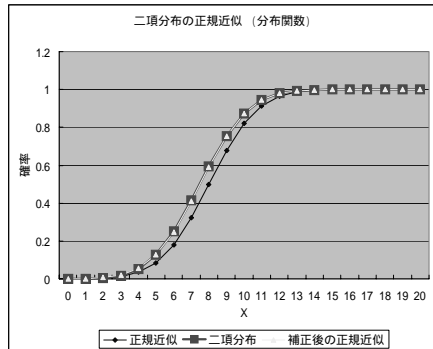
- 近似例: $X \sim \text{Bin}(20, 0.4)$ の場合 $E(X) = 8$ 、 $\text{Var}(X) = 4.8$
- $P(X = 5) = 0.1256$
- 正規分布で近似 $P(X \leq 5) = P\left(\frac{X-8}{\sqrt{4.8}} \leq \frac{5-8}{\sqrt{4.8}}\right) = \Phi\left(-\frac{3}{\sqrt{4.8}}\right) = \Phi(-1.369) = 0.0855$
- 連続補正 $P(X \leq 5) = P\left(\frac{X+0.5-8}{\sqrt{4.8}} \leq \frac{5+0.5-8}{\sqrt{4.8}}\right) = \Phi\left(-\frac{2.5}{\sqrt{4.8}}\right) = \Phi(-1.141) = 0.1269$

6/16/09

6



§ 5.3.2 中心極限定理(6)



6/16/09

7



§ 5.3.3 大数の法則と中心極限定理:再論

6/16授業後訂正

□ 中心極限定理を確認するための実験

- $X \sim U(0,8)$ にしたがう確率変数を n 個抽出した標本を 150ヶ作成。
- 150ヶの標本それぞれについて、標本平均(算術平均)を求め、標本平均の累積相対頻度分布(経験分布関数)を作成。
- n として、20, 400, 4000を採用

$$E(\bar{X}_n) = \frac{8+0}{2} = 4, \quad \text{Var}(\bar{X}_n) = \frac{(8-0)^2}{12n} = \frac{16}{3n}$$

$$Y_n = \frac{\bar{X}_n - 4}{\sqrt{16/3n}}$$

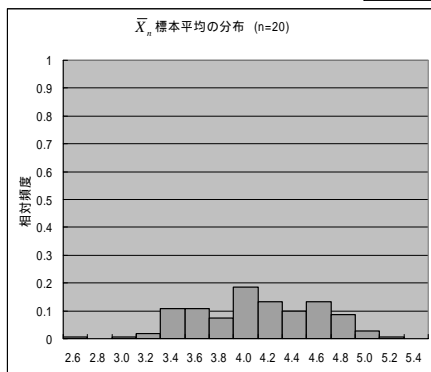
6/16/09

8



§ 5.3.3 大数の法則と中心極限定理:再論(2)

6/16授業後訂正



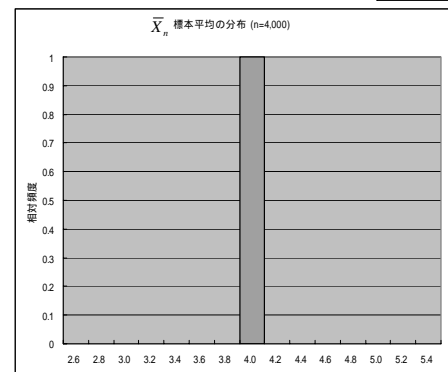
6/16/09

9



§ 5.3.3 大数の法則と中心極限定理:再論(4)

6/16授業後訂正

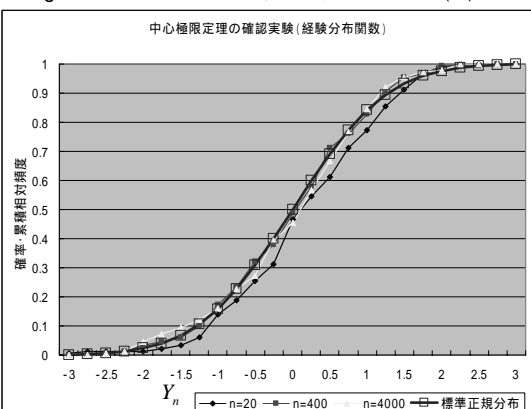


6/16/09

11



§ 5.3.3 大数の法則と中心極限定理:再論(5)



6/16/09

12



§ 5.3.3 大数の法則と中心極限定理:再論(6)

□ 大数の法則も中心極限定理も確率変数の収束

□ 確率変数列の収束概念

- 実数列の収束概念と異なる点
 - 収束先が確率変数
- 確率変数の同一性
 - 分布の同一
 - 変数自身の同一

6/16/09

13