

確率・統計の初歩(3)

別所俊一郎



1

標本統計量の大標本近似



- 標本統計量は確率変数になるから、その分布を
考えることができる
 - 手許にある標本から母集団の特性を統計的に推測
するとき、得られた標本統計量を持つはずの性質・分
布を知っておくことは有益
 - 一般には、標本統計量の分布は、母集団の分布、標
本の大きさ、サンプリングの方法に依存する
 - しかし、ランダムサンプリングで標本の大きさが十分
に大きければ、標本統計量の分布はあるいいてど推
測できる
- 大数の法則と中心極限定理

2

標本分布の扱い方



- Exactアプローチ
 - サンプルサイズがどのようなものであっても一般的に成り立つような公式を求める
 - exact distribution, finite-sample distribution (小標本分布)と呼ばれる
 - 求めにくいことがおおい: 母集団の分布が正規分布なら標本平均の分布は正規分布に従う

3

標本分布の扱い方



- Approximateアプローチ
 - サンプルサイズが十分に大きい ($n \rightarrow \infty$) のときの標本統計量の分布を近似的に用いる
 - Asymptotic distribution (漸近分布) と呼ぶ
 - 「十分に大きい」とは? : $n = 30$ くらいでもよい。
 - 「大数の法則」と「中心極限定理」を用いる
 - Exactアプローチよりシンプルで、母集団の分布に依存しないので使いやすい

4

大数の法則 Law of Large Numbers



- かなり一般的な条件の下で、サンプルサイズ n が十分に大きく、観測値が i.i.d. であれば、標本平均はかなり高い確率で母集団平均の近くにある
 - 平均が同じ確率変数を多く集めてその平均を取ると、ちらばりが相殺されて共通の平均値へ近づく
 - $n \rightarrow \infty$ のとき分布が「真の平均」の周りに集まってくる
 - 「確率収束する convergence in probability」と呼び、標本平均が真の平均に確率収束するとき、その標本平均は「一致性 consistency」を持つ、という
- Fig 2.6: ベルヌーイ分布の標本平均の分布

5

中心極限定理 Central Limit Theorem



- かなり一般的な条件の下で、サンプルサイズ n が十分に大きく、観測値が i.i.d. であれば、標本平均の分布は正規分布で近似される
 - 「分布収束 convergence in distribution」とよび、以下の記法を用いる
- 標準化された標本分布は「漸近的に」標準正規分布に従う
- 母集団の分布が正規分布なら、exactに正規分布になる
- そうでないばあいには、 $n = 100$ くらいにならないと近似が不十分 (Fig. 2.7, 2.8)

6

Key Concepts

- 大数の法則

- 中心極限定理

