

母集団と標本

資料：Rによるやさしい統計学・第4章

ISBN978-4-274-06710-5 オーム出版社

著者：山田剛史 杉澤武俊 村井潤一郎

概要

- ▶ 母集団

- ▶ 対象のデータ全体

- ▶ 標本

- ▶ 一部を取り出したデータ

推測統計

▶ 推測統計

▶ 推定

▶ 点推定

▶ 区間推定

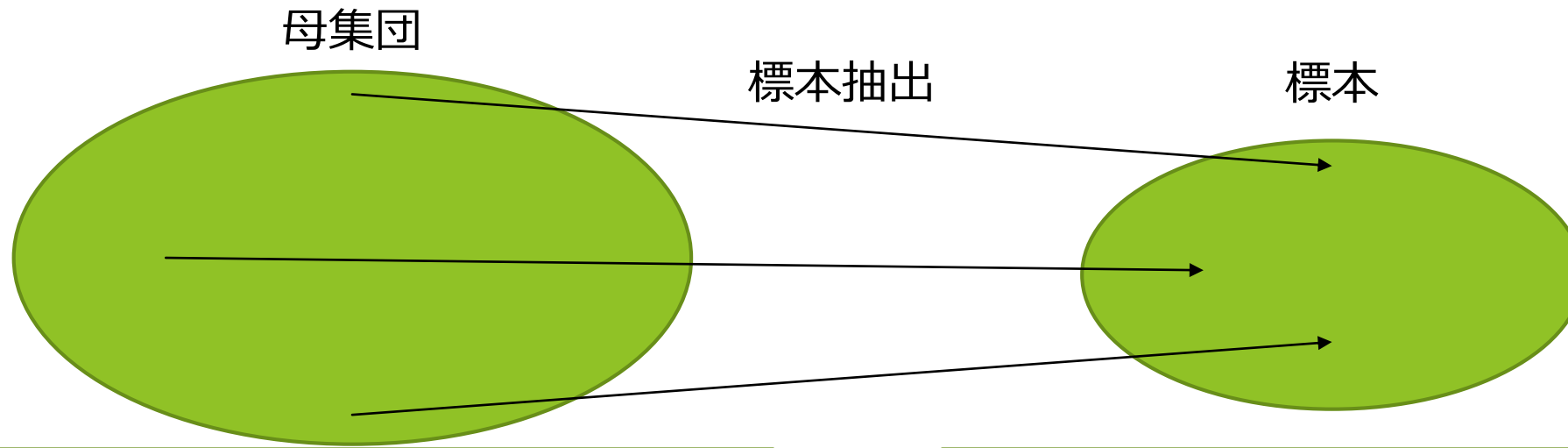
▶ 検定

▶ 第5章で

推定

- ▶ 標本から母数の値の結論を導く
- ▶ 点推定
 - ▶ 「60点くらいだろう」
- ▶ 区間推定
 - ▶ 「50～70点くらいだろう」

推定量と推定値



母数

- ▶ 母平均
- ▶ 母分散
- ▶ 母標準偏差
- ▶ 母相関係数
- ▶ 母比率

標本統計量

- ▶ 標本平均
- ▶ 標本分散 , 不偏分散
- ▶ 標本標準偏差
- ▶ 標本相関係数
- ▶ 標本比率

標本抽出に伴う誤差・準備

＊ 推定値がどれくらいあてになるのか調べる

- ▶ 単純無作為抽出
- ▶ 確率変数
- ▶ 確立分布
- ▶ 母集団分布
- ▶ 正規分布

単純無作為抽出

▶ 母集団の中のどのデータも平等に選ばれる

▶ 確立変数

▶ 実際に結果が得られるまでわからない

▶ 結果→実現値

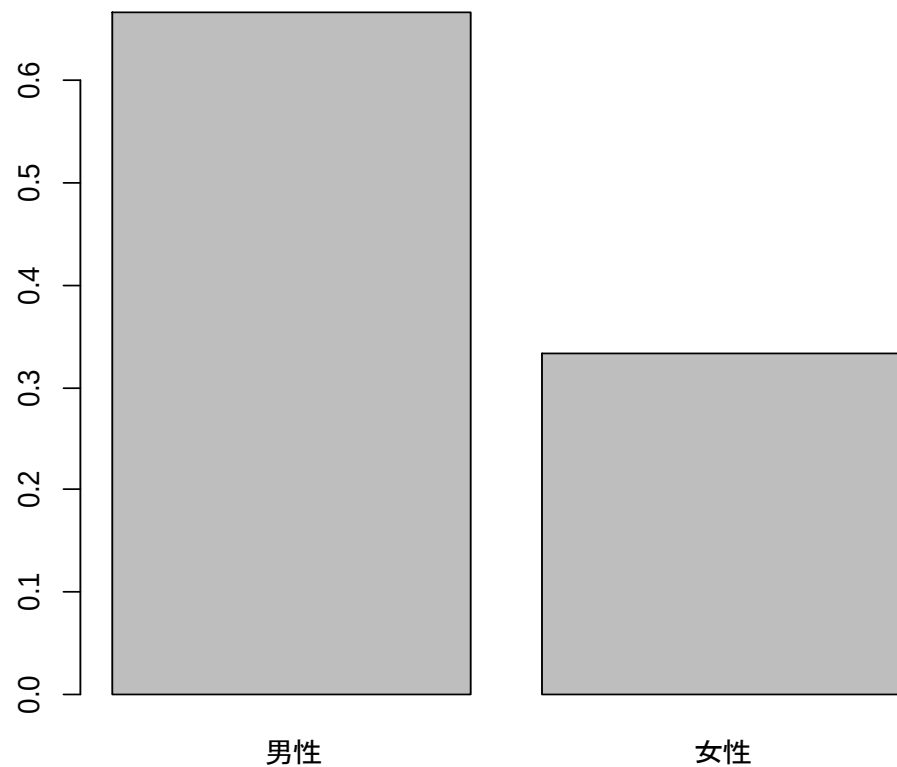
▶ 確率分布

サイコロの目	1	2	3	4	5	6
確立	1/6	1/6	1/6	1/6	1/6	1/6

Ex)サイコロの出る目の確率分布

母集団分布

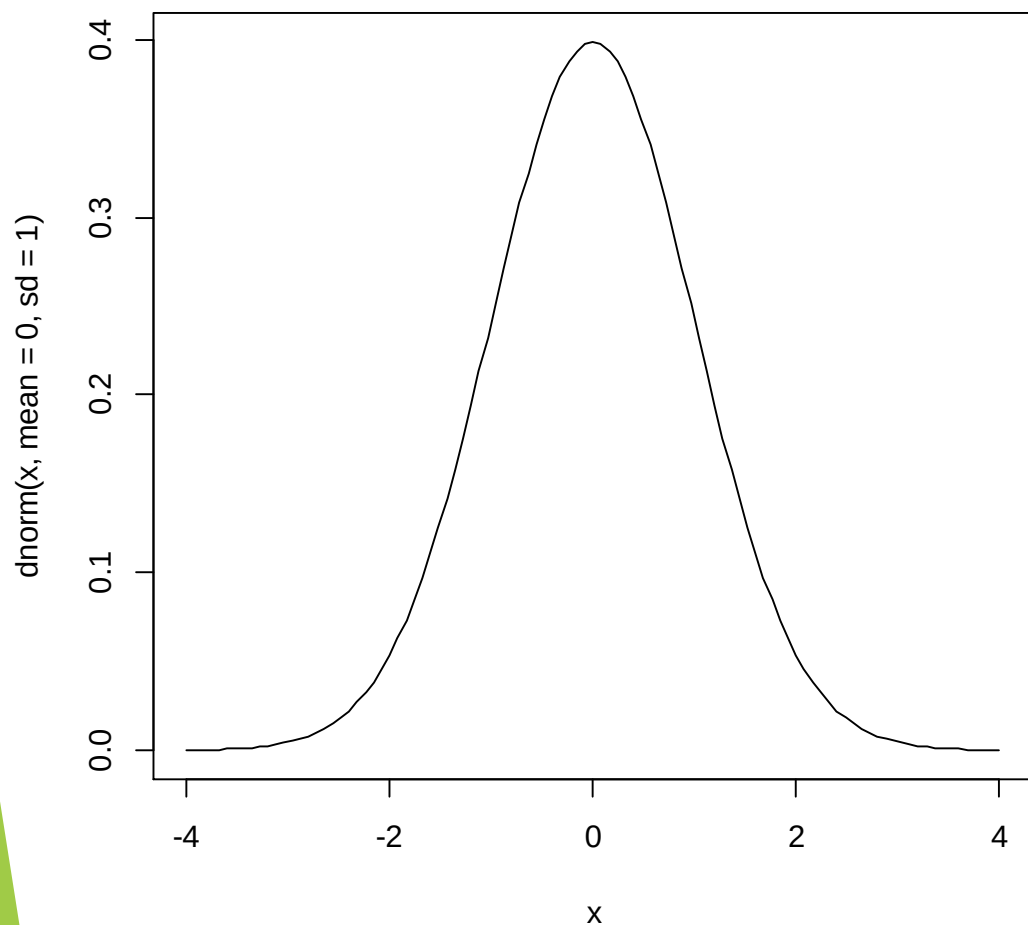
▶ ある変数の母集団における分布



Ex)男女比が2:1の分布

正規分布

▶ 平均 μ , 分散 σ^2 の時, $N(\mu, \sigma^2)$ と表す

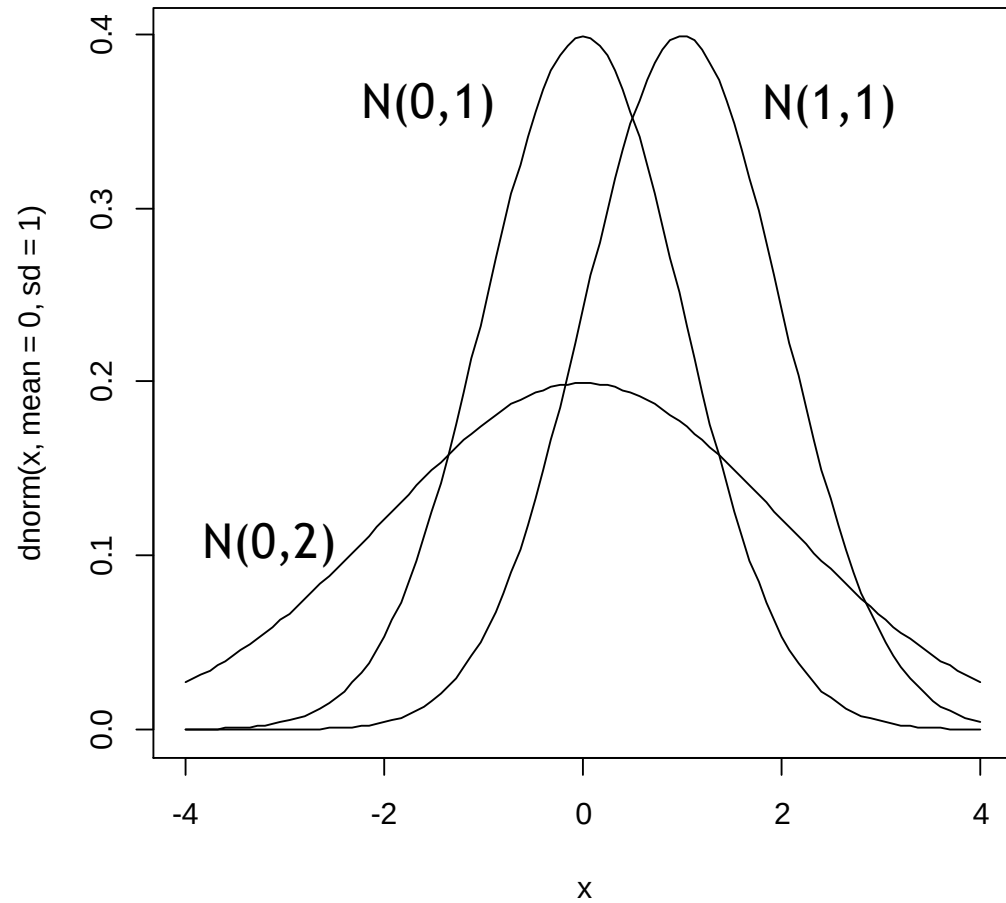


$N(0, 1)$

変数 X が $N(\mu, \sigma^2)$ にしたがうとき
 $X \sim N(\mu, \sigma^2)$ と表す

正規分布 2

- ▶ $N(1,1)$ と $N(0,2)$ を追加



変数の値

確率密度

標本分布

- ▶ 母数の推定値がどれくらいあてになるかを表す
- ▶ 母集団分布, 標本統計量の種類, サンプルサイズ n から数学的に求める
- ▶ $N(\mu, \sigma^2)$ の正規母集団の場合, 標本平均の標本分布は $N(\mu, \sigma^2/n)$ となる

不偏性と標準誤差

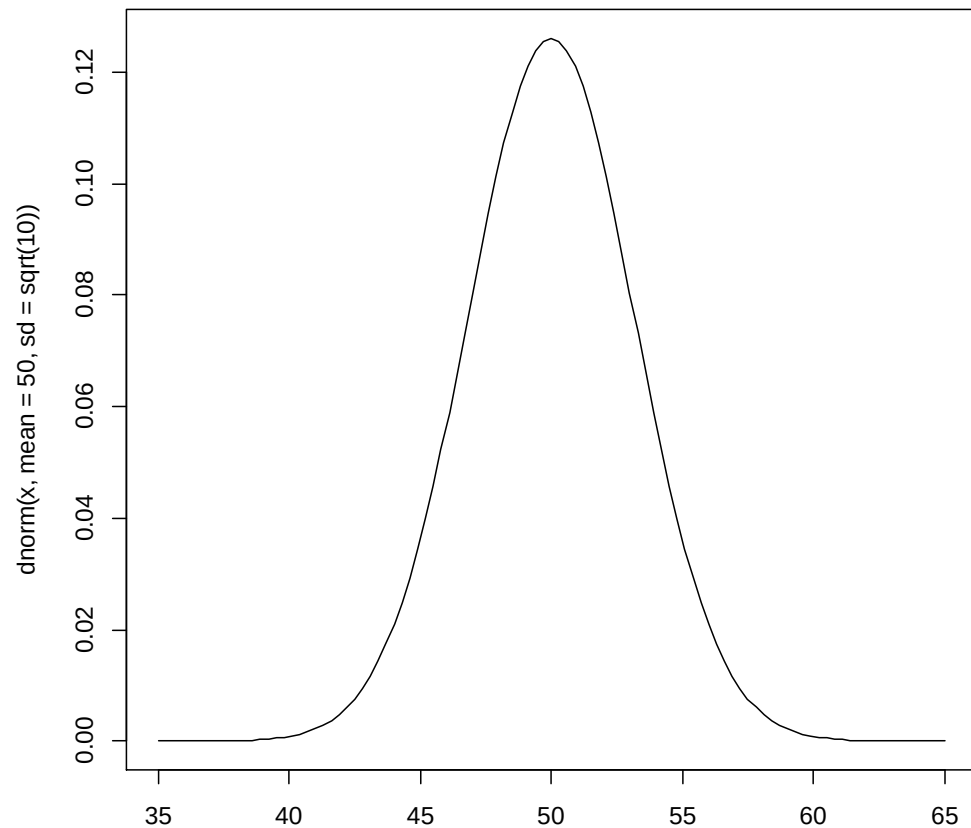
▶ 不偏性

- ▶ 無限に推定を繰り返した時,標本分布の平均と母数が一致する
- ▶ 不偏性のある推定量は不偏推定量と呼ばれる

▶ 標準誤差

- ▶ 標本分布の平均からの散らばり度合い
- ▶ 標本分布の標準偏差で表される

標本平均の標本分布



N(50,100)の正規母集団の場合
サンプルサイズ $n=10$

標準誤差 $= \sqrt{10}$

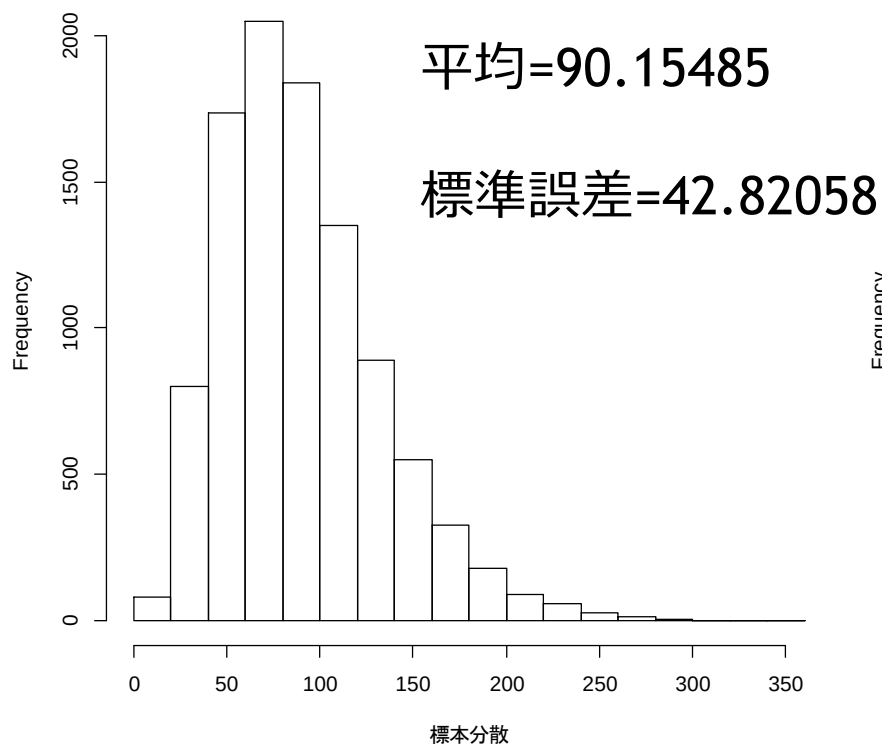
標本平均は母平均の不偏推定量

- ▶ 母集団が正規分布でなくとも, n が大きい時,標本平均の標本分布はほぼ正規分布になる

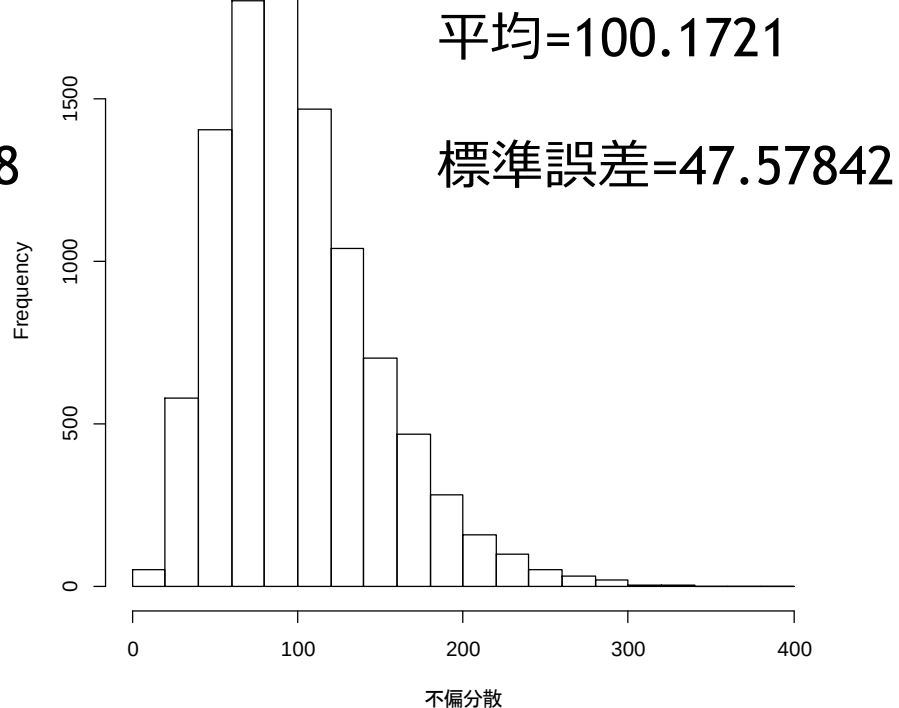
→ 中心極限定理

標本分散と不偏分散

Histogram of 標本分散



Histogram of 不偏分散



- ▶ 正規母集団 $N(50, 100)$ から10000回ずつ抽出
- ▶ 不偏分散には不偏性があるが, 誤差は大きい

まとめ

▶ 母集団

- ▶ 対象のデータ全体

▶ 標本

- ▶ 一部を取り出したデータ

▶ 単純無作為抽出

- ▶ 母集団の中のどのデータも平等に選ばれるような抽出

▶ 正規分布

- ▶ 統計で扱われる代表的な分布の形

▶ 標本分布

- ▶ 母数の推定値がどれくらいあてになるかを表す分布