

比較現代日本論研究演習/現代日本論演習 II (田中重人)

第2回「無作為抽出と区間推定」(2003.10.09)

1. 中心極限定理
2. 平均値の区間推定

【標本誤差の推定】

「標本誤差」(sampling error)

= 無作為抽出による誤差

- ★ 方向性をもたない
 - ★ 確率的に決まる
 - ★ 標本数が大きいほど誤差の範囲が小さい
- ➡ 「統計的推測」によって範囲を推定できる

【中心極限定理】

central limit theorem

- ★ 等確率標本の平均値は、母集団の平均値より高くなったり低くなったりする。
- ★ しかし平均的にみれば母集団の平均値に一致すると期待できる (点推定)
- ★ 標本サイズが大きいくほど、母集団の平均とのずれが小さくなる

別紙の乱数表から、1桁の数字を
10個と20個抜き出して、
それぞれ平均値を求めてみよう

【平均値の信頼区間】

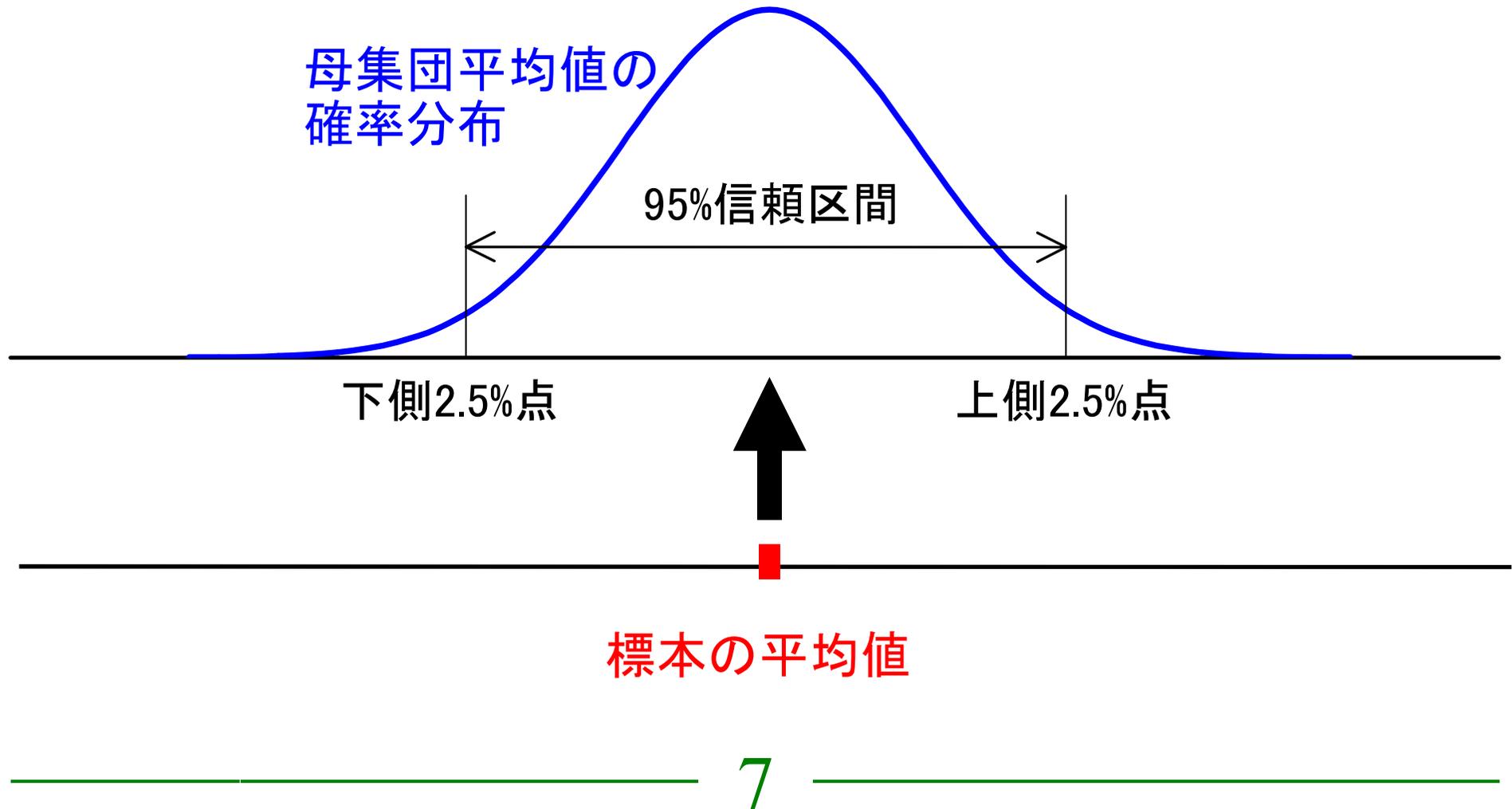
※「母集団では正規分布」の仮定が必要

- ★ 標本の平均値が母集団平均値からはずれる確率は正規分布にしたがう
 - ▶ 標本平均値から逆算すれば、母集団の平均値の確率分布 (t 分布) がわかる

★ 母集団の平均値の確率分布から両端を α %分だけ切り落としてえられる区間を $(100 - \alpha)$ %の「信頼区間」という。

α を「危険率」、 $(100 - \alpha)$ を「信頼率」という。
この値は自由に決めていいのだが、通常は $\alpha = 5\%$ として、95%信頼区間を求める。

信頼区間のもともめかた



【無限母集団の仮定】

母集団がある程度大きければ、統計的推測のうえでは、母集団は無限大とみなしてよい。

厳密にいうと、 $\frac{N-n}{(N-1)n} \approx \frac{1}{n}$ の場合

➡ 無限大の母集団から n 個の標本を無作為に選んだ場合について考える

★ 無限母集団からの標本の場合の
平均値の信頼区間のおおよその値：

$$\underbrace{\bar{m}}_{\text{標本平均}} \pm \underbrace{1.96}_{t \text{ 臨界値}} \times \underbrace{\frac{SD}{\sqrt{n}}}_{\text{標準誤差}}$$