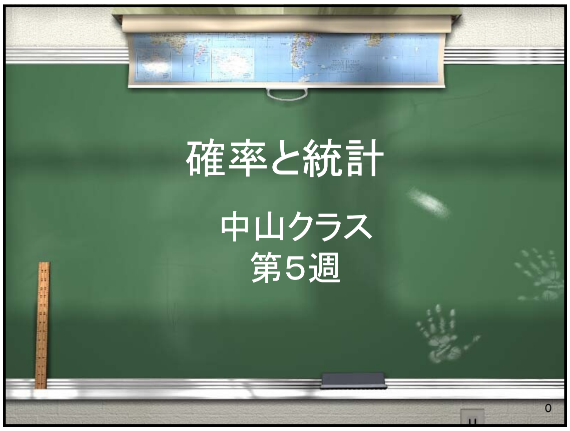


今日の内容

◆小テストについて

◆予想問題の解説

◆第4章　母集団と標本
4.1　母集団と標本
4.2　推定統計の分類
4.3　点推定

◆自習，コンピュータ演習，質問

1

小テストについて

11月6日(木)3限

試験時間：60分

試験範囲：教科書第1章～3章

持ち込み：筆記用具，時計，電卓(専用機)，学生証

試験の準備：
予想問題と過去問が完全に解けるように！

2

予想問題の解説

8

第4章　母集団と標本

3章まで：手元にあるデータの様子を記述する方法

4章から：大きな集団から一部を取り出した少数のデータの情報を使って，元の集団の性質について推測→推測統計

9

4.1　母集団と標本

非常に大規模なデータ全体(→母集団)の
(日本国民全体や工場で生産される製品全体に関するデータ)
統計的な性質(→母数)を
（度数の比率，平均，分散，相関など）
対象の一部を取り出したデータ(→標本)から推測

標本抽出：　母集団から標本を取り出すこと.
母数：母集団の性質を表す統計的指標(比率，平均，分散，相関係数など)

10

4.2　推測統計の分類

推測統計 ― 推定 ― 点推定 / 区間推定
推測統計 ― 検定 ・・・第5章以降

推定：具体的な値を用いて「母数は〇〇くらいだろう」
　日本の中学生全体の平均点は60点位だろう
　→点推定
　日本の中学生全体の平均は50点から70点くらい
　→区間推定

検定：母集団について述べた異なる立場の2つの主張（仮説）のうち，どちらを採択するかを決めるもの．
　「平均的な学力は5年前から変化していない」　仮説A
　「平均的な学力は5年前から変化した」　仮説B

4.3　点推定

4.3.1　点推定の手順

1. 17才の日本人男性全体の平均身長を推定
2. 10人を標本として抽出
　→サンプルサイズ，標本の大きさ=10
3. 10人の身長データ（165.2～171.3cm）を用いて平均を計算
4. 平均=169.36cm
5. 17才の日本人男性全体の平均身長=169.36cm と推定．

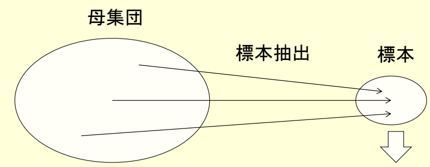
4.3.2　推定量と推定値

標本統計量：　標本データから計算される統計量
　標本平均，標本分散，標本相関係数など

母集団の統計量：
　母平均，母標準偏差，母相関係数など

母数の推定量：
　母数を推定するために用いられる標本統計量
　計算式（関数）や名前（平均など）

推定値：　標本データを用いて計算された推定量の値（数値）



＜母数＞
（本当に知りたいもの）
母平均
母分散
母標準偏差
母相関係数
母比率
など

推定

＜標本統計量＞
（標本から計算できる）
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
など

母数と推定量

母数	推定量	推定値
母平均	標本平均	標本データから計算
母分散	不偏分散	標本データから計算
母標準偏差	不偏分散の正の平方根	標本データから計算
母相関係数	標本相関係数	標本データから計算
母比率	標本比率	標本データから計算

母比率／標本比率

比率：　ある特徴を持った人や物が全体の中に含まれる割合

（例）
- ある学校の生徒にしめる男子生徒の割合
- 日本全体の有権者全体の中で現内閣を支持する人の割合
- あるテレビ番組を見ている世帯の割合（視聴率）

内閣支持率やテレビの視聴率などは，標本抽出に基づいた母数の推定が行われている．

4.3.3 標本抽出に伴う誤差

実際の母数の値にどの程度近い推定値が得られるか？
推定の結果はどのくらい信用できるか？

標本誤差：抽出された標本データのみを用いることで生じる誤差
(例)母集団=[1, 2, 3, 4, 5]　母平均=3
　標本データ=[1,4]　　標本平均=2.5
　標本データ=[2,5]　　標本平均=3.5

推定の結果「この程度の誤差が生じる」ことを知ることが重要. →標本分布　･･･　4.4節以降で取り上げる

17

不偏分散について

nサンプルの標本データ

$$X=[x_1, x_2, \dots, x_n]$$

標本平均

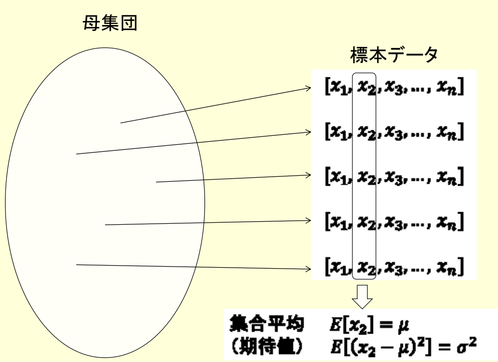
$$\bar{x}=\frac{1}{n}(x_1+x_2+\cdots+x_n)$$

標本分散

$$s^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

nサンプルの標本データを多数回抽出(独立試行)する
→x_iに対して多くのデータを得る
→集合平均(期待値)$E[x_i]$を計算→母集団の平均(μ)
同様に, $E[(x_i-\mu)^2]$を計算→母集団の分散(σ^2)

18



集合平均　$E[x_2]=\mu$
(期待値)　$E[(x_2-\mu)^2]=\sigma^2$

19

標本平均の期待値, μは母集団の平均

$$E[\bar{x}]=E[\frac{1}{n}(x_1+x_2+\cdots+x_n)]$$
$$=\frac{1}{n}\{E[x_1]+E[x_2]+\cdots+E[x_n]\}$$
$$=\frac{1}{n}(\mu+\mu+\cdots+\mu)=\mu$$

標本平均の期待値＝母集団の平均

標本平均の散らばり, σ^2は母集団の分散

$$E[(\bar{x}-\mu)^2]=E[\{\frac{1}{n}(x_1+x_2+\cdots+x_n-n\mu)\}^2]$$
$$=E\left[\left\{\frac{1}{n}((x_1-\mu)+(x_2-\mu)+\cdots+(x_n-\mu))\right\}^2\right]$$
$$=\frac{1}{n^2}\sum_{i=1}^{n}E[(x_i-\mu)^2]=\frac{1}{n^2}(\sigma^2+\sigma^2+\cdots+\sigma^2)=\frac{1}{n}\sigma^2$$

20

標本分散の変形

$$s^2=\frac{1}{n}\sum_{i=1}^{n}\{(x_i-\mu)-(\bar{x}-\mu)\}^2$$
$$=\frac{1}{n}\sum_{i=1}^{n}\{(x_i-\mu)^2-2(x_i-\mu)(\bar{x}-\mu)+(\bar{x}-\mu)^2\}$$

第2項の変形

$$-2\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)(\bar{x}-\mu)=-2(\bar{x}-\mu)(\bar{x}-\mu)=-2(\bar{x}-\mu)^2$$

21

標本分散の期待値

$$E[s^2]=\frac{1}{n}\sum_{i=1}^{n}E[(x_i-\bar{x})^2]-E[(\bar{x}-\mu)^2]$$
$$=\sigma^2-E[(\bar{x}-\mu)^2]=\sigma^2-\frac{1}{n}\sigma^2=\frac{n-1}{n}\sigma^2$$

母集団の分散σ^2と標本分散の期待値$E[s^2]$の関係

$$\sigma^2=\frac{n}{(n-1)}E[s^2]$$

母集団の分散は不偏分散となる.

22