統計Ⅱ（医療統計Ⅱ）
第3回　検定の基礎

授業担当：徳永伸一

東京医科歯科大学教養部　数学講座

あらためて：第13章 推定　の内容

Ⅰ．母集団と標本

Ⅱ．点推定

- 不偏性，不偏推定量

Ⅲ．区間推定

Ⅳ．母平均の区間推定

- 母分散が既知のとき

＊＊＊前期試験範囲ここまで＊＊＊

- 母分散が未知のとき

Ⅴ．母分散の区間推定

Ⅵ．母比率の区間推定

第14章 検定　の内容

Ⅰ．検定
Ⅱ．母平均の検定
Ⅲ．母分散の検定
Ⅳ．平均値の差の検定
Ⅴ．等分散の検定
Ⅵ．比率の検定
Ⅶ．適合度の検定
Ⅷ．独立性の検定

ではあらためて
第13章 の復習から

［復習］Ⅳ．母平均の区間推定

「母分散既知の場合」のポイントをまとめると

- 母分散σ^2を用いて標本平均の分布が表せる．
- 母集団分布が正規分布なら標本平均の分布も正規分布．
- 正規母集団を仮定せずとも，標本サイズnが十分大きければ標本平均の分布は正規分布で近似でき，いずれにしても正規分布の問題に帰着できる．
- だがその（標本平均が従う）正規分布$N(\mu, \sigma^2/n)$は，母分散σ^2を用いて表されているのだから，σ^2の値がわからないと推測できない．
- 現実にはσ^2は未知のケースが多い！

一方･･･

[復習] Ⅳ．母平均の区間推定

標本の不偏分散 $U^2 := \{ \Sigma (x_i - \bar{x})^2 \} / (n-1)$

は, 常にわかる (標本データから計算できる) .

そこで**母分散未知**の場合には:

- 方針1: U^2を σ^2の近似値として利用.
  - U^2は σ^2の (不偏) 推定量ですからね.
  - 「近似」を認めてしまえば, 結局またしても正規分布の問題に帰着. 推定の方法はほとんど同じ.

    (Uの値を σ のところに代入するだけ. 同じ公式が使える)
  - 標本サイズが大きければ, 良い (誤差の少ない) 近似値であることが期待できる.
  - だが小標本の場合は誤差が無視できないはず.

    →方針2へ

[復習] Ⅳ．母平均の区間推定（母分散未知）

- 方針2：σ^2を用いずに表せる（U^2を含む）統計量を導入する．

→第11章Ⅴ

3-その他の重要な標本分布 [2]　t分布

（★p.99のt分布の説明冒頭に「XとYが互いに独立な確率変数で、」と追加）

以下の定理を利用．

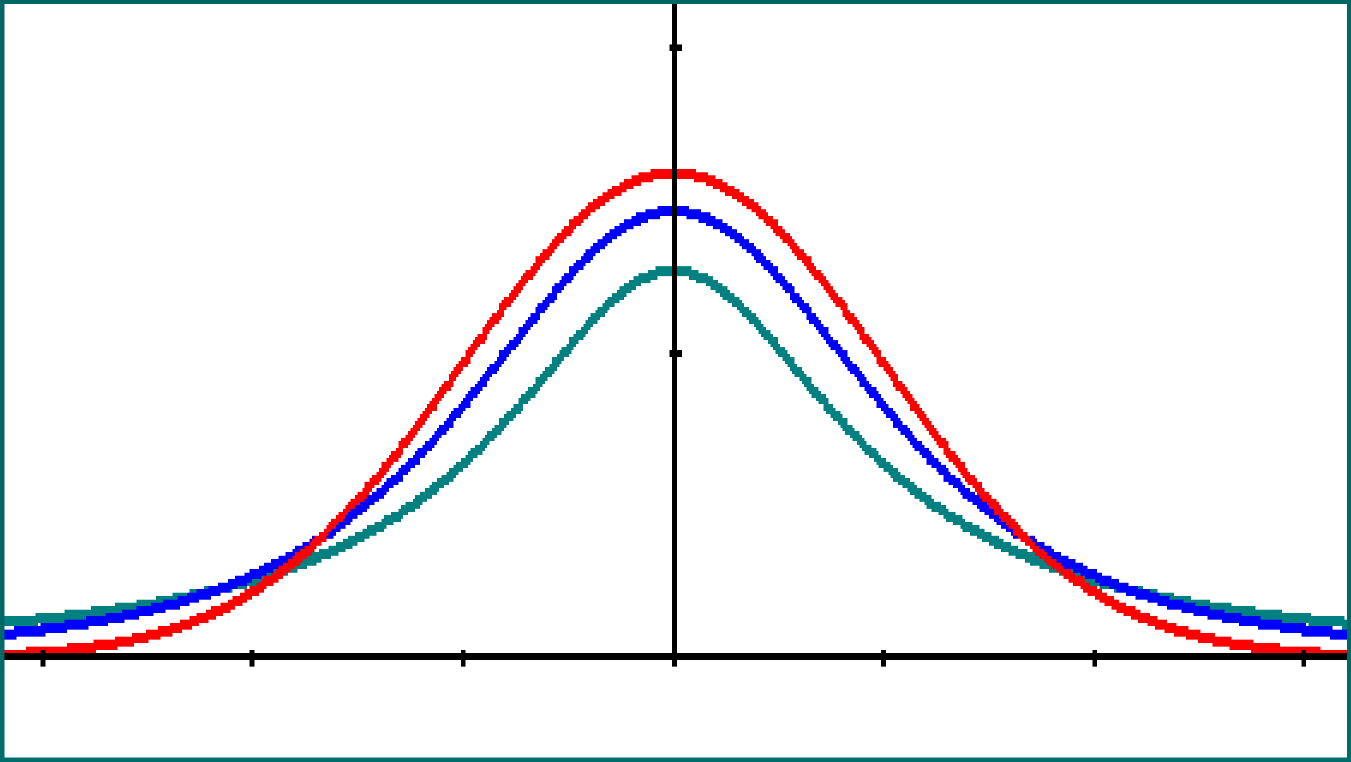
定理（教科書p.100）

正規母集団$N(\mu, \sigma^2)$からの，大きさnの無作為標本$\{X_1, X_2, \cdots, X_n\}$について，標本平均X'，不偏分散を$U^2$とする．このとき，確率変数

$$T := (X' - \mu) / (U / \sqrt{n})$$

は自由度n-1のt分布（と呼ばれる分布）に従う．

t分布の密度関数のグラフ
(明星大学 船津好明先生のウェブサイトより転載)



青:自由度3, 暗緑:自由度1 (赤:標準正規分布)

[復習] Ⅳ. 母平均の区間推定（母分散未知）

t分布について

- 教科書p.99 (第1刷) の定義に「XとYが独立」という条件を追加.
- 「スチューデントのt分布」とも呼ばれる.
  - 「スチューデントstudent」は発案者のペンネーム.
- 厳密に定義するのは少々やっかいなので省略. ともかく理論的に「わかっている」分布として (巻末の数表などを用いて) 利用する.
- 密度関数の概形は正規分布と似ている. 対称性のある釣鐘型.
- 自由度と呼ばれるパラメータを持つ.
  - サイズnの標本から求めたスチューデント比Tにおいては, 自由度ν (ギリシャ文字の「ニュー」) ＝n－1.
  - 自由度とは大雑把に言えば「独立なパラメータの個数」. あまりピンとこなくても,当面は「そういうもの」と割り切って利用すればよいです.
  - ν大きいほど標準正規分布N (0,1) に近づき, ν＝∞で完全に一致.
  - 自由度νのt分布の上側100α%点を$t_\nu(\alpha)$またはt(ν, α)で表す.
  - $t_\nu(\alpha)$の値は教科書p.137の表から.

[復習] Ⅳ．母平均の区間推定（母分散未知）

注意と補足

- $T = (X' - \mu) / (U / \sqrt{n})$ は**スチューデント比**と呼ばれる.
- 「**正規母集団**」の仮定はt分布を利用する上で本質的に重要.
- 標本平均X'の標準化変数:

  $Z = (X' - \mu) / (\sigma / \sqrt{n})$ との類似に注意.
  - つまりnが大きければ, T　Zであるということ.
    (ただし, 今考えているのはnが小さいケースでした)

[復習] Ⅳ．母平均の区間推定（母分散未知）

t分布を利用したμの推定

- 母分散σ^2が未知の正規母集団N（μ，σ^2）からの, 大きさnの無作為標本{x_1, x_2, ・・・,x_n}について, 標本平均x', 不偏分散をU^2とする.

このとき

$$t = (x' - \mu) / (U/\sqrt{n}) \sim (\text{自由度}n\text{-}1\text{の}t\text{分布})$$

だから

- 母平均μの信頼度γ＝1－αの信頼区間は

$$\left(x' - t_{n-1}(\alpha/2)\, U/\sqrt{n},\ x' + t_{n-1}(\alpha/2)\, U/\sqrt{n} \right)$$

【注意】やってることは母分散既知の場合と大差ない.
自由度という新たなパラメータが出てきただけ.

［復習］ VI. 母比率の区間推定

「母集団比率に関する推測」とは

- ベルヌーイ母集団に関する推測.
  - 2項分布の応用.
  - ある程度大きな標本を扱うケースがほとんどなので, たいていは「2項分布の正規近似」を利用する.
- 「世論調査の類」. 支持率調査など, 身近に興味深い例が多い.
  - 「統計的な理解を深めるよいチャンス」.

[復習] Ⅵ. 母比率の区間推定

2項分布の正規近似RECALL

中心極限定理により，nが十分大きいとき，
B (n,p) はN (np, np (1-p)) で近似できる．

∵ B (n,p) に従う確率変数は，B (1,p) (という同一の分布) に従う独立なn個の確率変数の和と見なせるから．

- 従って，標本比率P＝X/nの分布も，
  正規分布 N (p, p (1-p) /n) で近似できる．
- さらにPの標準化変数：

$$Z = (P-p) / \sqrt{p(1-p)/n}$$

は近似的に標準正規分布に従う．

- 分布が決まれば，あとはこれまでと同じ考え方で進めればよい．．．

がしかし！

［復習］ VI．母比率の区間推定

（以下母比率をp，標本比率$P＝X/n$ の実現値をP_0とする）

$$Z＝(P-p)/\sqrt{p(1-p)/n}$$

を元に$P(Z\in I)＝1-\alpha$となる区間Iを構成すると

$$\left(P_0-z(\alpha/2)\sqrt{p(1-p)/n},\ P_0+z(\alpha/2)\sqrt{p(1-p)/n}\right)$$

となり未知数pが残ってしまう．そこで

推定を行う際は，

$Z=(P-p)/\sqrt{p(1-p)/n}$ の分母（Pの標準偏差）のpを近似値（推定値）P_0で置き換えて計算．すなわち信頼度$\gamma＝1－\alpha$の信頼区間は以下のようになる：

$$\left(P_0-z(\alpha/2)\sqrt{P_0(1-P_0)/n},\ P_0+z(\alpha/2)\sqrt{P_0(1-P_0)/n}\right)$$

★ただし，誤差の最大値を見積もりたいときは$p=0.5$ を採用．

- $p(1-p)$ は$p=0.5$のとき最大値0.25を取ることに注意（2次関数のグラフを思い出せ！）．

[復習] Estimation of error (誤差の評価)

標準誤差SE = $\sqrt{p(1-p)/n}$ depends only on p and n,

i.e., it does not depend on the size of population!

(母集団の大きさは誤差に影響しない！)

- In many case, the estimation by a sample with n=1000 is reliable.

  (ただし無作為抽出がうまくいっていれば)

- On the other hand, we assume random sampling in the above estimation, which can be suspected.

［復習］平成14年度試験問題より

「2002年10月1日付けの朝日新聞記事によれば、最近5年間の日本棋院のプロ公式戦15000局において、黒盤勝率は51.86%であった。」

(2) 日本棋院は現行ルールでは先手が有利であるとしてルールの改正方針を決定した。この判断についてどう思うか。仮説検定の考えを用いて考察せよ。

【参考】朝日新聞の解説（ちなみに東京版夕刊の1面）：

「厳しい勝負の世界ではわずか2%の違いも無視できない」

↑「そりゃ違うだろ！」と突っ込むところ.

- 標本サイズが大きいから誤差は小さく、わずかな差も無視できない（14章で学ぶ仮説検定の用語を用いれば「有意差がある」）のであって、「勝負の厳しさ」とは無関係.

教科書p.111訂正

［例題5］

誤「回答者330人のうち，<u>18</u>人が」

正「回答者330人のうち，<u>23</u>人が」

（よって標本比率p’＝ 23/330　0.07となり，
解答はそのまま）

では14章 へ．

推定から検定へ

- もう数学的に重要な部分の大半は出尽くした．区間推定がわかっていれば検定も簡単！（なはず）
- 基本的なアイディアを正しく理解することが重要．
  - 計算だけなら中学生でもできるしパソコンがやってくれる
- 推定と検定を混同しないこと！
  - 単なる手法の違いというよりは，発想が根本的に異なる．

第14章 Ⅰ. 検定

仮説検定とは

- （帰無）仮説：
  「事前の情報や経験から抱いている既存のイメージ」
- その（帰無）仮説を疑いたくなるような状況が生じたとする.
  - その「疑い」に基づく（帰無仮説を否定する）仮説を対立仮説という.
    （ただし単に「仮説」という場合は帰無仮説の方を指します）
- まず（帰無）仮説が正しいとして（本当は疑いつつ）, 標本データを見ずに推論を行う. 一種の背理法的思考.
- 標本データから得た情報と（帰無）仮説に基づく推論の結果を照らし合わせ, 仮説の真偽を判断する（仮説を棄却または採択する）.

第14章 Ⅰ. 検定

1-検定の原理

- 例1：「甲状腺疾患患者の10人の血清アルブミンを測定したら健常人の60%しかなかったが，この疾患では血清アルブミンが低値になると断定してよいだろうか」
  - 「減ることはない」と仮説（帰無仮説）を立て，この仮説が否定できるかどうかを調べる．
  - 標本の実現値と理論値（期待値）は一致しないのが普通だが，その差異が誤差の範囲内なのか，意味のある差異なのか，が問題．
  - 確率の理論から言って誤差の範囲内とは考えられない差異があると判断できるとき，仮説は否定され，「有意である」「有意差がある」という．

- 例2：「サイコロを60回投げたら1の目が25回も出た．理論的には（すなわち期待値としては）1の目は10回（程度）出るはずである．この差異は大きすぎないだろうか」

  →「サイコロは正常」という（帰無）仮説を立て，「25回」が誤差の範囲内かどうかを調べる．

第14章 Ⅰ. 検定

2-帰無仮説, 対立仮説と2種類の過誤

- 最初に設定した仮説：**帰無仮説**. 通常H_0で表す.
- 帰無仮説を否定したとき受け入れる仮説：**対立仮説**. 通常H_1で表す.
- 帰無仮説H_0を否定することを, H_0を**棄却する**という.
- H_0が棄却されないとき, H_0を採択するという.
- さきほどの例1だと
  - H_0：「甲状腺疾患にかかっても血清アルブミンに変化はない」
  - H_1：「変化する」（もしくは「増える」,「減る」）
    - 医学的見地から増えることはあり得ない（らしい）ので, H_1は「減る」としてもよい（→後述の「片側検定」の例）.

第14章 Ⅰ. 検定

2-帰無仮説, 対立仮説と2種類の過誤 (続き)

- **第1種の過誤**：本当はH_0が正しいのにH_0を棄却してH_1を採択してしまう.
  - 第1種の過誤を犯す確率を**危険率**あるいは**有意水準**といい, 通常 α で表す.
  - 習慣的に, α ＝0.05, 0.01とすることが多い.
- **第2種の過誤**：本当はH_0が間違っているのにH_0を採択してしまう.
  - 第2種の過誤を犯す確率を通常 β で表し, 1－β を**検出力**という.

第14章 Ⅰ．検定

注意と補足

- （あたりまえだが）仮説検定による判断は実際の真偽と一致するとは限らない．あくまで客観的･合理的な判断をするための1つの手法．
- 必然的に確率の評価を含んだ判断となる．
  - 判断の基準となる確率が**有意水準**
- 典型的な例･･･新薬や新しい治療法のテスト
  - まず「効果がない」という帰無仮説を立てて推論

  →臨床試験のデータを分析した結果，帰無仮説を否定できれば，「効果がある」と判断できる．

第14章 Ⅰ．検定

3-片側検定と両側検定

棄却域：検定に用いる統計量 (**検定統計量**) において，棄却することが決定される範囲．棄却域の補集合が**採択域**．

母数θに対し，帰無仮説をH_0：$\theta = \theta_0$とするとき，対立仮説H_1の立て方により以下のように分類：

- **両側検定**：H_1はH_0の単純な否定，すなわち

$$H_1 : \theta \neq \theta_0$$

- **片側検定**：H_1は大小関係を考慮したH_0の否定，すなわち

$$H_1 : \theta > \theta_0 \quad \text{または} \quad H_1 : \theta < \theta_0$$

  - 異なっているとすれば確実に大きい，あるいは小さいことが予測されるときにはこちら．

第14章 Ⅰ. 検定

4-検定の手順

① 帰無仮説H_0, 対立仮説H_1, 有意水準 α を定める.

② 標本から統計量 (検定統計量) を求める.

③ 検定統計量 (の実現値) が棄却域に入るかどうか調べる.

④ 棄却域に入ればH_0を棄却してH_1を採択, 入らなければH_0を採択.

【注意】 H_0が棄却できなかったとき便宜上「H_0を採択」というが, 「棄却できない」ことは「H_0の正しさ」を積極的に主張するものではない！

(H_0が間違っていても, 検定統計量が”たまたま”採択粋に入ることは十分あり得る)