統計Ｉ
第6回　推定･検定の基礎 ⑵

授業担当：徳永伸一

東京医科歯科大学教養部　数学講座

今日 (第6回) の授業の概要:

- 第5章前半 (推定) の復習と補足
- 第5章後半 (検定) へ

第5章前半の復習 (その1)

教科書p.58 (既出の用語の復習といくつかの新たな用語)

KEYWORDS：

– 母集団, 無作為標本
– 母数parameter ･･･母平均, 母分散, 母比率
– 統計量, 統計値 (統計量の実現値)
  ･･･標本平均, 標本分散, 標本比率
– 推定量 (母数の推定に用いる統計量) ，
  推定値 (推定量の実現値)

– 点推定⇔区間推定

前回はここまで.

– 検定 (統計的仮説検定)

第5章前半の復習 (その2～不偏性)

推定量が「**不偏**である (偏りがない)」とは:

- 「対応する母数より大きい (or小さい) 値が得られやすい」といった傾向がない.
- その推定量を繰り返し実測し、得られた値 (推定値) の平均値は, 繰り返しの回数を増やすほど対応する母数に近づく.

厳密には:

- **[定義]** 母数θの推定量θ'に対し,

$$E(\theta') = \theta$$

のときθ'をθの**不偏推定量**であるという.

第5章前半の復習 (その3～不偏性続き)

- **標本分散** $s^2 := \{ \Sigma (x_i - \bar{x})^2 \} / (n-1)$ は, 母分散 σ^2 の不偏推定量.
- すなわち, (定義より) $E(s^2) = \sigma^2$ である (証明は割愛).

ということは

- $s'^2 := \{ \Sigma (x_i - \bar{x})^2 \} / n$ とおけば,

  $E(s'^2) = E(((n-1)/n) s^2) = ((n-1)/n)\sigma^2$
- つまり s'^2 の実測値は, 母分散より小さめの値をとる傾向にある (すなわち不偏でない).

不偏性に関する補足

- **標本平均** $\bar{x} := \Sigma x_i / n$も母平均$\mu$の不偏推定量.

  ∵$E(\bar{x}) = E(\Sigma x_i / n) = \Sigma E(x_i) / n = \mu$
- たとえばn＝3のとき$x' = (x_1 + x_2 + 2x_3)/4$ とおいてもx'はμの不偏推定量（確認せよ）.
- 標準偏差つまり$s = \sqrt{(s^2)}$は, 不偏推定量ではない！

  ∵$E(s) = E(\sqrt{(s^2)}) \neq \sqrt{(E(s^2))} = \sigma$

第5章前半の復習（その4～区間推定）

5.2 母数の区間推定

- 「母数θが区間Ｉに含まれる確率が○%」といった形の推定を行なう.
- ↑における
  - 区間・・・（○%）信頼区間
  - 確率・・・信頼度（または信頼計数）
- 推定する母数は，この授業（統計Ｉ）では母平均（試験ではひょっとすると母比率も）．
  - （必要なら近似を導入し）正規分布の問題に帰着できるケース.

第5章前半の復習 (その5～区間推定)

母平均μの区間推定について整理

問題設定:

- 標本サイズn, 標本平均 $\bar{X}$ の実測値が与えられており, 母分散 σ^2 は**既知**とする.
- 母集団分布・・・正規分布なら好都合だが, 任意の分布でもnが大きければ, 標本平均の分布は中心極限定理により正規分布で近似可能.

以上の条件のもとで,

「μの 100γ%信頼区間」

を求める (γは信頼度＝信頼係数で, 具体的には 0.95, 0.99, 0.90など).

第5章前半の復習（その6～区間推定の続き）

母平均μの区間推定について（続き）

- $\bar{X} \sim N(\mu, \sigma^2/n)$ と近似できる．
  - 注意：正規母集団を仮定すれば厳密．
- $Z = (\bar{X} - \mu)/(\sigma/\sqrt{n})$ と標準化すると $Z \sim N(0, 1)$
- $P(-z(\alpha/2) \leqq Z \leqq z(\alpha/2)) = \gamma$
  - ただし $\gamma = 1 - \alpha$．$z(\alpha)$ は $P(Z \geqq z(\alpha))$ を満たす値．
  - たとえば $\gamma = 0.95$ のとき $\alpha/2 = 0.025$， $z(\alpha/2) = z(0.0$
  - $25) = 1.96$
  - $z(\alpha/2)$ は「上側 $\alpha/2$ 点」と呼ばれる．
- ↑を同値変形すると

$$P\left(\bar{X} - z(\alpha/2)\ \sigma/\sqrt{n} \leqq \mu \leqq \bar{X} + z(\alpha/2)\ \sigma/\sqrt{n}\right) = \gamma$$

よって「μの（100×γ）％信頼区間」は

$$\left(\bar{X} - z(\alpha/2)\ \sigma/\sqrt{n}\ ,\ \bar{X} + z(\alpha/2)\ \sigma/\sqrt{n}\right)$$

第5章前半の復習（その7～区間推定についての補足）

「μの（100×γ）％信頼区間」：

$$\left(\bar{X} - z(\alpha/2)\ \sigma/\sqrt{n}\ ,\ \bar{X} + z(\alpha/2)\ \sigma/\sqrt{n} \right)$$

について：

- 標本平均（の実測値）を中心とする区間である．
  - 問5.1・・・中心位置がずれるだけ．
- $\sigma/\sqrt{n}$ は標準誤差と呼ばれる．
- γを大きくすると・・・
  - $\alpha = 1 - \gamma$ は小さくなる → $z(\alpha/2)$ は大きくなる．
    →信頼区間の幅が大きくなる．
    （外れる確率を減らすのだから、幅を大きく取る必要があるのは当然）
- nを大きくすると → 信頼区間の幅が小さくなる．
  - 情報量が増えるのだから，誤差が減るのは当然

推定から検定へ

- 以上で第5章前半 (推定) はおしまい.

次に検定 (仮説検定) へ

- もう数学的な内容は出尽くした. 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.

5.3 仮説検定の考え方 (その1)

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) .

5.3 仮説検定の考え方 (その2)

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・・・**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.

5.3 仮説検定の考え方（その3）

例題5.2

- 「あるコインを10回投げたところ，表が2回，裏が8回出た」
- 曲がっていないコインなら，表裏の出る確率は等しく，ほぼ半々の割合で出るはず

  →「曲がっているかも」と疑いたくなる．
- 上の結果をもとに，「コインは曲がっていない」という（帰無）仮説の当否を判定（検定）せよ，という問題．
  - 対立仮説：「コインは曲がっている」．

5.3 仮説検定の考え方（その4）

例題5.2の解答

手順1：帰無仮説，対立仮説の明示・定式化.

表が出る確率をpとおく.

- 帰無仮説H_0：p＝0.5
- 対立仮説H_1：p≠0.5

［注意］

- 対立仮説はここでは単純に「H_0の否定」と考えて「p≠0.5」を採用する.
- 「p＜0.5」（裏が出やすい）などとする考え方もある（→5.4片側検定）

5.3 仮説検定の考え方（その5）

例題5.2の解答（続き）

手順2：帰無仮説H_0にもとづく推論.

- X：コインを10回投げるとき，表の出る回数.
- H_0のもとで，X～B（10,0.5）であり，

  $P(X=x)={}_{10}C_X \cdot 0.5^{10}$

  また$\mu=E(X)=10\times 0.5=5$
- H_0のもとで，「（Xが）x＝2と同等かそれ以上に，μ＝5から離れた値をとる確率」を計算.
  - – Xが特定の値（ここでは2）をとる確率だけでなく，それと同等かそれ以上に期待値から離れた値をとる確率の和を考えるところがポイント（教科書p.64-65）

5.3 仮説検定の考え方（その6）

例題5.2の解答（続き）

手順2の続き：

- ［Xが2と同等かそれ以上に，μ＝5から離れた値をとる確率］
  ＝P（｜X－5｜≧｜2－5｜）
  ＝P（X＝0,1,2,8,9,10）＝（中略）＝0.1094
- あとは「確率0.1094をどう評価するか」という問題．
  （「ありそうもないこと」か，あるいは「十分可能性がある」とみるか）
  → 有意水準の設定次第．

手順3：棄却／採択の決定

- とりあえず有意水準α＝0.05とする
  ・・・「手順2」で求めた確率はこれより大きいので
  H_0は棄却できない（＝採択される）
- すなわち検定結果は有意でなく，「コインは曲がっているとはいえない」と判断された（仮説検定の終了）．

5.3 仮説検定の考え方（その7）

例題5.2の結果についての考察と注意

- 「曲がっているとは言えない」という表現について：
  - 【重要】仮説検定においては，「棄却」によって主張するのが基本．棄却されなかった場合，「帰無仮説が正しい」と積極的に主張するのは言い過ぎ．
  - 例題5.2ではたまたまこ「コインは曲がっている」ことが立証されなかっただけなので、「コインは曲がっていない」とまで言うべきではない．
- 有意水準の値によって検定結果は変わり得るが，今の問題では有意水準 $\alpha=0.10$ としても結論は同じ（確認せよ）．
- では試行の結果が「表1回，裏9回」の場合は？
  - 手順2で求めた確率（「Xが1と同等かそれ以上に，$\mu=5$ から離れた値をとる確率」）$=0.0216$
  - $\alpha=0.05$ なら棄却，$\alpha=0.01$ なら棄却されない（採択）．
- 「帰無仮説が棄却されるような，統計量の値の範囲」を，その統計量の**棄却域**という．
  - 今の問題では $X=0,1,9,10$ が，有意水準 $\alpha=0.05$ における X の棄却域．
- 今の問題では試行回数が少なかったので確率の厳密な値を計算したが，試行回数が多い場合は「2項分布の正規近似」を用いる（→第6章6.5）．

5.3 仮説検定の考え方（その8）

母平均μに関する検定.

例題5.3

- 検定の対象：母集団の平均身長μ
- 母集団・・・ある都市の高校3年生男子
- 標本サイズn＝81（人）
- 標本平均 $\bar{x}$ ＝171.3（cm）
- 母分散σ^2＝30（cm^2）は不変.

（ここまで例題5.1とほぼ同じ）

- 10年前の平均身長μ_0＝170.0

以上の条件のもとで,

「μが変化したかどうかを有意水準5％で検定せよ」

という問題.

5.3 仮説検定の考え方 (その9)

例題5.3の解答

手順1：帰無仮説, 対立仮説の明示・定式化.

- 帰無仮説H_0：$\mu = \mu_0 = 170.0$
- 対立仮説H_1：$\mu \neq 170.0$

［注意］ここでも対立仮説は単純に「H_0の否定」と考えた.

– 「食生活の変化」等を考慮し, $\mu > 170.0$を対立仮説とする立場もあり得る (→片側検定)

5.3 仮説検定の考え方 (その10)

例題5.3の解答 (続き)

手順2：帰無仮説H_0にもとづく推論.

- $\bar{X} \sim N(\mu, \sigma^2/n)$ と近似できる.
  - 注意：正規母集団を仮定すれば厳密.
- $Z = (\bar{X} - \mu) / (\sigma / \sqrt{n})$ と標準化するとZ $\sim N(0, 1)$
- $P(-z(\alpha/2) \leqq Z \leqq z(\alpha/2)) = \gamma$
  - ただし$\gamma = 1 - \alpha$. $z(\alpha)$ は$P(Z \geqq z(\alpha)) = \alpha$を満たす値.
  - ここまでは例題5.1とほぼ同じ (次の変形に注意).
- ↑を同値変形すると

$$P\left(\mu - z(\alpha/2)\ \sigma/\sqrt{n} \leqq \bar{X} \leqq \mu + z(\alpha/2)\ \sigma/\sqrt{n}\right) = \gamma$$

[注意] 推定のときは以下のように変形した.

$$P\left(\bar{X} - z(\alpha/2)\ \sigma/\sqrt{n} \leqq \mu \leqq \bar{X} + z(\alpha/2)\ \sigma/\sqrt{n}\right) = \gamma$$

5.3 仮説検定の考え方（その11）

例題5.3の解答（続き）

［手順2の続き］

よって標本平均 $\bar{X}$ の値が

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

の範囲から外れるとき，H_0は有意水準αで棄却される．

すなわちこの範囲が**採択域**で、この外が**棄却域**．

μ＝170.0，α＝0.05を入れて計算すると

（168.81，171.19）

手順3：棄却／採択の決定

標本平均＝171.3は上の採択域に入らない（棄却域に落ちる）

　ので、H_0は有意水準5％で棄却された．

すなわち「平均身長は伸びた」と判断される．

5.3 仮説検定の考え方（その12）

例題5.3の結果についての注意と補足

- 例題5.2の場合と違い，帰無仮説H_0が棄却されたので「H_0が間違っていた」ことをより積極的に主張できる．
- $\bar{X}$の採択域（168.81, 171.19）は，例題5.1（区間推定）におけるμの95％信頼区間（170.11, 172.49）を並行移動したもの．
- このように，母平均に関しては区間推定と**両側検定**が密接に関係している（信頼区間を見れば検定結果がわかる）．
  - 具体的には，帰無仮説H_0：$\mu=\mu_0$，対立仮説H_0：$\mu\neq\mu_0$ に対し「μ_0がμの100（1－α）％信頼区間に入る」こと と「$\bar{X}$の実現値が有意水準αでの採択域に入る」ことが同値．
  - 例題5.3では$\mu_0=170.0$が例題5.1で求めた95％信頼区間に入らないことから，実は計算せずとも有意水準5％で棄却されることはわかっていた．
- 「**片側検定**の場合はどうなるか？→5.4参照
- 用語：「第1種の誤り」：本当はH_0が正しいのに棄却すること．
  「第2種の誤り」：本当はH_0が正しくないのに採択すること．

5.4 片側検定（その1）

母比率pに関する検定.

<u>例題5.4</u>

- 検定の対象:「ある症状」の3日後の治癒率p
- 母集団・・・「ある症状」の患者
- 投薬なしで治癒する確率p_0＝0.2
- 標本サイズn＝20（人）
- 標本において投薬の結果治癒した人数 x ＝9（人）

以上の条件のもとで,

「投薬により治癒率pが増加した（薬の効き目があった）かどうかを有意水準5％で検定せよ」

という問題.

5.4 片側検定（その2）

例題5.4の解答

手順1：帰無仮説，対立仮説の明示・定式化．

p：投薬した場合の治癒率 として，

- 帰無仮説H_0：$p = p_0 = 0.2$
- 対立仮説H_1：$p > 0.2$

［注意］投薬により症状が悪化することがないならば，「$p < 0.2$」の可能性は最初から考える必要がない．

5.4 片側検定（その3）

例題5.4の解答（続き）

手順2：帰無仮説H_0にもとづく推論.

- X：20人が薬を飲み3日間静養した場合の，治癒した人数.
- H_0のもとで，X～B（20,0.2）であり，

  $P(X=x) = {}_{20}C_X \cdot 0.2^{20} \cdot 0.8^{20-x}$

  （期待値$E(X) = 20 \times 0.2 = 4$）
- H_0のもとで，「X≧9となる確率」を計算.
  - もし両側検定であれば「（Xが）x＝9と同等かそれ以上に，μ＝4から離れた値をとる確率」を計算する（例題5.2参照）.
- 表5.2より$P(X \geqq 9) = 0.010$

5.4 片側検定（その4）

例題5.4の解答（続き）

手順3：棄却／採択の決定

- P（X≧9）＝0.010＜0.05 （＝有意水準α）

  →よってH_0は棄却される．

- すなわち検定結果は**有意**で，「薬は効く」と判定された（仮説検定の終了）．

5.4 片側検定（その5）

例題5.4の注意と補足

- Xの有意水準5%での棄却域は

$$\{8, 9, \cdots, 20\}$$

- 用語：
  - 片側検定・片側棄却域

  （例題5.4では右片側検定・右片側棄却域）
  - 両側検定・両側棄却域

5章全体についての注意と補足

- 例題5.1と例題5.3のように, 同じ問題設定で推定と検定を考える場合について, 両者の関係をしっかり把握しておきましょう.
- 両側検定と片側検定の違いや使い分け方についてよく理解しておきましょう.
  - 両者を使い分ける基準は必ずしも明確ではないが, 事前情報や検定の意図により「明らかに片側検定すべき」状況というのはある.
  - 逆に「片側検定すべき」理由がなければ両側検定を行なえばよいだろう.
  - 同じ問題設定で片側検定と両側検定の2通りを行なった場合, 棄却域 (片側棄却域と両側棄却域) がどのように変化するかを理解しておこう.
  - 一般に, 片側検定の方が (その「片側」だけに注目すれば) 棄却域が広いので, 両側検定より棄却されやすい.
- 信頼度が変化したときの信頼区間の変化, 有意水準が変化したときの棄却域の変化について理解しておきましょう.

試験に向けての注意と補足

- 教科書P.70の演習問題は, いずれも推定･検定を理解する上で重要なので一通りやっておきましょう.
- pp.133-134の表はコピーして試験問題と一緒に配布するので, これを活用できるようにしておくこと.
- 試験は例年通り, 電卓持込可です. ただし (多少不利になるとはいえ) 手計算で合格点を取ることも可能なので, 当日忘れた場合はあきらめてください. 試験中の貸し借りは禁止.
- パソコン, ポケコン, PDA, 高機能関数電卓等, 公式を記憶可能な機器, および携帯電話等通信機能を持った機器の使用は不正行為と紛らわしいので禁止します.