疫学に必要な統計：  
中村好一『基礎から学ぶ楽しい疫学　第3版』

2018年1月17日

小野裕介

[ 本日の内容 ]

＊シミュレーション：別紙参照

＊**標本抽出**

＊**偶然誤差**と**系統誤差**

#oy 「検定と推定」と「標本サイズ」の議論の前提となる「誤差」に関する議論を復習する.

＊**検定**と**推定**

＊**標本サイズ**

#oy 標本サイズには様々な議論が関わるため最後にした.

**[ 標本抽出 ]** (pp.155-160)(p.89)

＊すべての研究において, 「解くべき課題」(何を明らかにするのか)が存在する.

＊「解くべき課題」が決まれば, その対象である「母集団」も決まる.

・「母集団」とは, 「解くべき課題」が対象とする集団である. 「標的集団」ともいう(p.157 図10-1).

・「標的集団」の外部妥当性は問題になりうるが, 自明のこととしてしまうことが多いのが現実(p.157 注5).

＊実際の研究で対象とする集団が「観察対象集団」である(p.157 図10-1).

・「全数調査」では, 「標的集団」の全員を研究対象とする. よって, 「観察対象集団」は「標的集団」と等しい. (p.157)

・「標本調査」では, 「標的集団」の一部を抽出して「観察対象集団」とする. 「無作為抽出」するのが原則. (p.157 図10-1)

＊「無作為抽出」とは, すべての標的集団(母集団)の構成メンバーにおいて, 観察対象集団(標本)となる確率が等しい

抽出方法(p.159).

＊「無作為抽出」の方法にはいくつかある. #oy テキスト(p.159)参照

・単純無作為抽出

・系統抽出法

・層化抽出法

・多段抽出法

＊「観察対象集団」から参加拒否者を除いた, 実際に研究に参加した集団を「観察集団」という(p.157 図10-1).

・参加拒否者の存在は「選択の偏り」につながる. (p.158)

・参加拒否者をできるだけ少なくするためにいろいろな方法を講じる. (p.158)(p.92)

#oy もちろん, 「いろいろな方法」は, 数学的な手法である統計学的方法の範疇ではない.

**[ 偶然誤差** と **系統誤差 ]** (pp. 84-85)(p.87)

＊誤差：「真の姿」と「観察結果」の差 (p.84)

#oy 統計学, 疫学では, 「観察結果」が真実なのではない. 決して観察することのできない「真の姿」というものが

想定されている. 「観察結果」から, 「真の姿」というものを(数学的に)推し量ることが, 「統計学的推論」であるともいえる.

＊誤差の分類 (p.84 図5-1)

! ! !

+ - $

＊誤差

・偶然誤差 ( ! ! ! )：「標的集団」から「観察対象集団」を「無作為に抽出」する際に, 「観察結果」は標本ごとに異なる.

そうした「観察結果」は「真の姿」とは異なる. (p.84)

・系統誤差 ( + - $ )：系統的な, 一定の方向性をもった誤差. (p.85)

＊系統誤差 ( + - $ )

・狭義の偏り ( + - )

・交絡 ( $ )

＊狭義の偏り ( + - )

・選択の偏り ( + )：「標的集団」から「観察対象集団」を抽出する場合に, 偏った抽出方法を行った結果生じる偏り. つまり, 無作為抽出ではない場合に生じる偏り. (p.89)

あるいは, 観察対象集団と観察集団が異なる場合(参加協力拒否者が存在する場合)に生じる偏り. (p.89)

・情報の偏り ( - )：研究で得られた情報が事実とは異なる場合に生じる偏り. (p.92)

＊統計学的「推定」や「検定」などの「統計学的推論」は, 確率論を用いて「偶然誤差」の大きさを評価する. (p.84)

＊「偶然誤差」の小さな研究を行うためには, どうしたらよいか？答えは簡単で, 「標本サイズを大きくする. 」これしかない. (p.87)(p.160)

#oy 確認：以降の「検定と推定」と「標本サイズ」の議論は, 「偶然誤差」に関わるもの.

#oy 「偶然誤差」に関わる統計学的な議論は, 完全に数学的な議論である. 一方, 「狭義の偏り」は,

数学的には予測できない誤差である.

#oy 「狭義の偏り」は, 標本数が増えても必ずしも減少しない. それどころか, 増えることすらある. この点で, 「偶然誤差」とは

決定的に異なる. 中村がp.160で「標本サイズ」と「集める情報の質」について言及しているのは, こうしたことが背景にある.

[ **検定** と **推定** ] (pp.164-172)(p.88)

＊疫学と統計学は異なるが, 重なる部分も多い. (p.164)

・重なる：疫学は経験論的法則を処理する. #oy 経験論的法則を処理する技法が統計学.

・異なる #oy それぞれに固有の領域がある.

＊筆者は統計学を, (1)記述統計学と, (2)分析統計学に大別している.

(1) 記述統計学：集めたデータをわかりやすく示すために加工する作業. (p.165)

(例) 平均, 標準偏差. #oy 図表による可視化の技法も含むだろう. ヒストグラムなど.

・記述統計学は分析統計学よりも断然, 重要. (p.165)

(2) 分析統計学：推定や検定などの統計学的推論. 確率論を使いながら, 標本から得られた結果から母集団の状況を

推し量る. (p.165)(p166 図10-2)

#oy 「真の値」である母集団の平均と, 「観察結果」である標本平均の値とは異なる. ただし, 標本平均に関する規則性が, 「確率論」という数学の一分野の結果からある程度はわかっている(どういう分布をするのかということがわかっている). そのため,

かなりのことを確率的に言うことができる. それが, 以降で述べる「検定」や「推定」である.

＊統計学的検定：表10-6　コホート研究における罹患率比(相対危険)の検討)

・暴露と疫病発生が無関係な場合には相対危険は1.0(暴露群と非暴露群で罹患率が等しい).

相対危険が1.0よりも大きい場合には, 暴露に意味があるということ.

・帰無仮説：暴露と疫病発生が無関係, すなわち相対危険は1.0

#oy ただし, 偶然誤差があるため, 暴露と疫病発生が無関係な場合(帰無仮説が正しい場合)にも,

1.0よりも大きい相対危険を観察することはある.

・第1種の過誤(α) #oy 本当は, 暴露と疫病発生が無関係なのに, 無関係ではないと判断してしまう(帰無仮説を棄却)

間違い. その確率をαとする.

・第2種の過誤(β) #oy 暴露と疫病発生の間に関係があるのにもかかわらず, 無関係と判断してしまう

(帰無仮説を棄却しない)間違い. その確率をβとする. 逆に, 関係がある場合に無関係ではないと正しく判断する

(帰無仮説を棄却する)確率である1-βを「検出力」という.

#oy ちなみに, 検定における, 第1種の過誤, 第2種の過誤は, スクリーニングにおける, 「擬陽性」, 「偽陰性」に似ている

ともいえる. その場合, 「検出力」は「感度」に相当するともいえる.

・有意水準5%の検定は, αを5%にするような検定である.

#oy 仮説検定のもとでは, 確率5%でしか生じないような領域を棄却域とし, 観察結果が棄却域のときに帰無仮説を棄却

するようにする. 帰無仮説が正しいのにも関わらず棄却してしまう(第1種の過誤)可能性は, その確率(α)が低い(5%)とはいえ, 存在する.

・判断の結果が正しいか, 誤っているかは神のみぞ知る世界(p.169)

#oy 帰無仮説が棄却されない場合には, 第2種の過誤(関係あるのに関係ないとしてしまう間違い)が問題になる.

・通常の検定ではもっぱら第1種の過誤の確率αのみ注目しており, 第2種の過誤の確率βには無頓着である. したがって,

帰無仮説が棄却されない場合でも, 「母集団の罹患率比は1.0である」ということの保証にはならない.

せいぜい「有意な結果ではなかった」というような表現にとどめておきたい. (p.169)

#oy スクリーニング(陽性が有意)でいえば, 検定は, 陰性だからといって安心できない検査のようなものか.

＊観察項目と用いる検定方法 (表10-7)：詳細を知りたい場合には, 統計学の教科書を参照(p.171)

＊統計学的推定：95%信頼区間

・症例対照研究における区間推定の例(p.167 表10-5)

#oy 観察されたデータから, 「真の値」を区間として推定する. 「真の平均」が観察されたデータの平均の前後と考えるのは自然. 問題は, 区間の幅をどの程度にするのかである.

#oy 区間の幅が狭いほうが精度が高く, 情報量は多いといえるが, 間違いやすくなる(真の値が区間に入っていない可能性が

高まる). 5%という値はその間違いやすさ(信頼の程度)のさじ加減を決めている. 例えば, 99%信頼区間の幅は,

95%信頼区間の幅よりも広くなる(広いほうが間違いにくい).

#oy 間違いやすさのさじ加減を決めているという意味で, 5%というのは検定における有意水準に相当する.

#oy 観察している値が極端なものである可能性は, 低いとはいえ, ある(はたして自分が観察しているデータが極端なものなのかどうかも知ることはできない). 非常に極端な値を観察している場合には, 「真の値」は推定される区間には入らないことになる.

どれぐらい極端な値を観察すると, 区間に入らなくなってしまうのか, という程度を決めているのが, 5%という値であるともいえる.

ニュアンスとしては, 確率5%でしか生じないぐらいはしっこの極端な値を観察してしまう場合には外れる, という感じ.

#oy 真の値は観察不可能なので, 実際に計算した区間の中に入っているかどうかを知ることはできない.

#oy nが大きくなると, 区間の幅が小さくなることにも注意. 推定の精度が上がると考えるとよい.

・95%信頼区間のつくり方：標本で得られた値に標準誤差の1.96倍の数値を加えたものと減じたものを,

95%信頼区間の上限と下限として表示する(p.168)(p.166 図10-2)

#oy 標準誤差とは, 「偶然誤差」によって生じるばらつきだと思えばよい.

サンプルサイズ(n)が大きければ, 標準誤差は小さくなる.

#oy 標準正規分布(標準偏差1かつ平均0の正規分布)においては, 1.96から-1.96の値が全体の95%を占める.

模擬試験でいえば, 受験者の成績が正規分布する場合, 偏差値30.4から69.6の間に全受験者の95%が含まれる.

#oy まとめ：区間の中心は, 標本の平均値によって決まる.

#oy まとめ：区間の幅は(1) 95%信頼ということ, (2) サンプルサイズ(n)によって決まる.

＊**推定** と **検定**の関係

・時代の流れは検定よりも推定(p.170)

・推定のほうがより数量的に偶然誤差の大きさを提示する(p.88)

#oy 「より数量的」？推定のほうが「より直接的かつ直観的」にという感じだと思う.

・推定の結果から検定結果は判断できる. (p.170)(p.167 表10-5)

#oy 95％信頼区間に含まれる値は, 平均がその値であるという帰無仮説を有意水準5％では棄却することができないような値.   
#oy 検定と推定の違いは表現の仕方の違いといえる. 実質的にはやっていることは同じである(表裏).

だからこそ, 推定の結果から検定結果を判断できる.

・検定結果だけでは「有意ではない」という１つの情報だけしか入手できないが, 推定の結果はより多くの情報を提供している. (p.170)　#oy やや語弊のある表現か. 検定の場合はp値に情報が込められている(p値のみに情報がつめ込まれ過ぎかも).

＊注意点

・母集団の観察事項が小さくても, 標本サイズが大きければ検定結果は有意になる.

#oy 標本サイズが大きければ, 検定の精度が高まるため, ほんの少しの違いであっても検出できるようになるというイメージ.

・「母集団の観察事項の大きさ」が, 意味がある程度なのかどうかについては, 統計学的検定は何も情報提供してくれない.

意味があるかどうかの判断は, 既存の知識や経験などを発揮する必要. (p.172)

#oy 臨床的に意味がある差なのかどうか.

**[ 標本サイズ ]** (pp.160-163)(p.88)

＊標本サイズ

「標本サイズは大きければ大きいほど, 精度の高い結果が得られる」(p.160)

#oy 偶然誤差の問題

ただし, 「集める情報の質が落ちなければ」(p.160)

#oy バイアスの問題

＊「2群間の平均の差の検定」の場合の標本サイズ：表10-3.

・(例) 母集団の血圧の平均の差を検出したい.

・以下が決まれば, 必要な標本サイズの値が決まる(B6).

#oy ある程度の信頼性(第1種の過誤)と検出力(第2種の過誤)で, 2群間の平均血圧の違いを見出したい場合に,

どれぐらいの標本サイズがあればよいか.

#oy 「有意な結果が得られるための標本サイズ」(p.161)という表現があるが, やや誤解を招くか.

(1) 両群の平均の差(B1)：母集団の平均の差10.0mmHgを検出したい.

(2) 一方の群の標準偏差(B2)：15.6mmHgと仮定する.

(3) 他方の群の標準偏差(B3)：18.2mmHgと仮定する.

(4) 第1種の過誤(B4に影響)：0.05とする.

(5) 第2種の過誤(B5に影響)：0.2とする(検出力0.8).

#oy 見出したい差(1)が小さいほど, 必要な標本サイズは大きくなる.

#oy 高い信頼性(4)が欲しければ, 必要な標本サイズは大きくなる.

#oy 高い検出力(5)が欲しければ, 必要な標本サイズは大きくなる.

・標本サイズの計算はさまざまな前提のうえに成り立っている. その最たるものは研究を実施する前から

「母集団の状況はこの程度のものであろう」という前提を立てる必要があることである. (p.162)

#oy 上の例だと, 母集団の標準偏差がわかっている必要がある.

#oy 添付資料：10.5 症例数設計(『日本統計学会公式認定　統計検定1級対応　統計学』東京図書)

＊「母集団の平均の95%の信頼区間推定」の場合の標本サイズ：表10-1.

・(例) 母集団の血圧の平均を, 95%信頼区間で推定する際に, 区間を一定の間隔におさめたい.

・以下が決まれば, 必要な標本サイズの値が決まる(B3)

(1) 母集団の血圧の標準偏差(B1)：16.8mmHgと仮定する.

(2) 95%信頼区間の幅(B2)：5.0mmHg

#oy B3の計算式にある3.92という値は1.96を2倍した値である.