医学研究データを読み解くための早わかり講座

Contents

1. 統計解析は標本（得られたデータ）から母集団の特徴を推定する方法論
2. 標準偏差（SD）と標準誤差（SE）
3. P値と95%信頼区間
4. サンプルサイズについて
5. まとめ

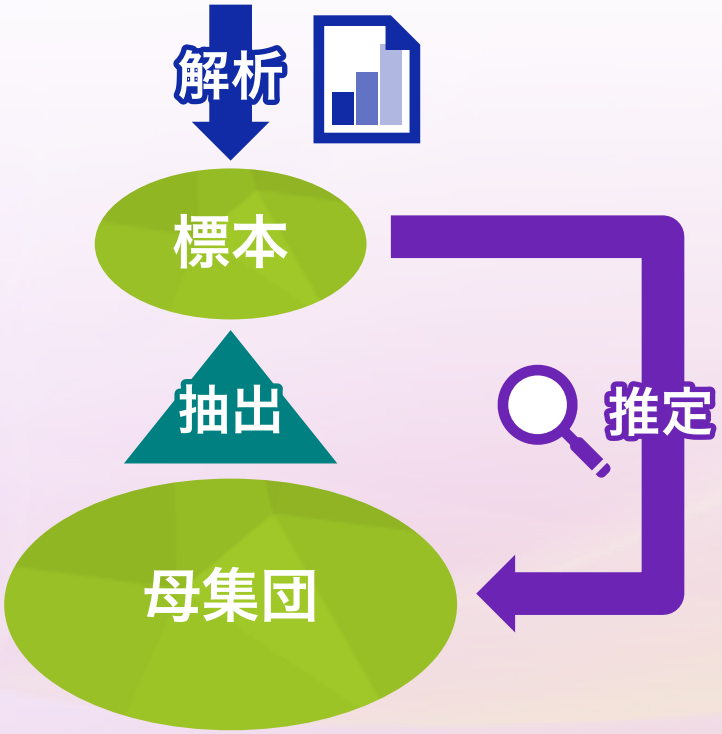
1 統計解析は標本(得られたデータ)から母集団の特徴を推定する方法論

臨床研究は一部の人に薬剤を投与し、そこで得られたデータをより多くの人にも通用するかを検証します。しかしその結果をみると効果にばらつきがあり、本当の効果なのか偶然なのかという判断が難しいことがあります。
そこで役立つツールが統計解析です。

たとえばある治療法の効果を調べたい場合、一部の対象者(標本)について調査した結果が、
ある疾患を持った患者全体(母集団)についても同様のことがいえるのかを統計解析によって判断できます。

つまり統計解析とは、
母集団の一部分である標本を解析し、母集団の特徴を推定したり予測したりする方法論といえます。

統計解析は医学研究において、なくてはならない存在です。
本コンテンツでは医学論文などで目にする医学研究データを正しく評価するためのポイントをわかりやすく解説していきます。



② 標準偏差（SD）と標準誤差（SE）

I. SDとSEとは

標準偏差（Standard Deviation: SD）

データが平均値を中心にどのくらいばらつきがあるのかを示す指標です。
標準偏差の最小値はゼロです。データの「ばらつきの程度」が大きければ値が大きくなり、
小さければ小さな値となります。
ばらつきの程度を数値化できれば、その値を比較することによって効果のばらつきを評価できます。

標準誤差（Standard Error：SE）

標準偏差を√nで割った値です。
標本から得られた値の精度を示す指標で、
得られた値が母数（母集団の値）にどれだけ近いかを示しています。
平均値の場合、標本から求めた平均値が母数の平均にどれだけ近いかを示します。
数値が大きければ母数の推定精度が低く、
小さければ母数の推定精度が高いことになります。
データ数が多くなると値は小さくなり、母数の推定精度が高くなります。

標準偏差

$$\mathrm{SD}=\sqrt{\frac{\sum_i (x_i-\bar{x})^2}{n-1}}$$

標準誤差

$$\mathrm{SE}=\sqrt{\frac{\sum_i (x_i-\bar{x})^2}{n(n-1)}}=\frac{\mathrm{SD}}{\sqrt{n}}$$

x_i は i 番目の値
$\bar{x}$ は平均値、nはデータ値

II. 論文執筆時に「SE」と「SD」をどちらを使用すべきか

標準偏差と標準誤差の使い分け

ばらつきを示したいのであれば「標準偏差」
母数の推定幅を示したいのであれば「標準誤差」

標準偏差を使用するケース

年齢や臨床検査値などの患者背景を示したい

- 平均値とそのばらつきを示したい場合は標準偏差を用います。
- → データが正規分布に従うときにはばらつきのよい指標となります。

平均値±SD
患者背景
年齢（年）　59.3±9.4
LDL (mg/dL)　121.8±18.1

標準誤差を使用するケース

変化量やリスクなど、“どれくらい”変わったか（変化量やリスクを母数と考える）を示したい

- 母数の値がどのあたりにあるかを推定精度とともに示したい場合は標準誤差を用います。
- → 最近では標準誤差の代わりに95%信頼区間(95%CI)で示されることも多いようです。

平均値±SE
治療前後のLDL変化(mg/dL)　-17.3±3.5
イベント発症相対リスク　2.5±0.2

③ P値と95%信頼区間

I. P値と95%信頼区間とは

p値は仮説検定において、標本から得られたデータが偶然得られる確率を数値で表したものです。
※「p」とはprobability（確率）の頭文字です。

- pの値は0 ～1の間の値であり、p値が事前に定めた有意水準（よく用いられる有意水準は0.05）よりも小さければ、得られたデータは偶然得られたものとは考えにくいと考え、有意な差があると判断します。

95%信頼区間（confidence interval：CI）は、どの程度差があるかを区間推定し、母数が存在しそうな範囲を数値（下限値と上限値）で表したものです。

- 100回標本を作製した場合、95回はこの範囲に母平均があります。
- 95%信頼区間が狭ければ、母数は標本から得られた値の近くにあると考えられます。

検定結果を示すp値

推定結果を示す
95%信頼区間

II. データから有意差を読みとるポイント

有意水準5%で検定した結果と95%CIで推定した結果は一致します。

- p値は差が大きくても、差の標準誤差が小さくても大きくなります（差の大きさはわかりません）。
- 95%CIはどのくらいの差があるかを推定幅で示しているので情報量が多いといえます。

差の検定（たとえばt検定）のp値と95%CI

- p<0.05ならば、差の95%CIに0は含まれません。
- p≧0.05ならば、差の95%CIに0が含まれます。

イベント発生リスクの低下（リスク差）

31%（95%CI：17~45%）　有意差あり

31%（95%CI：-5~66%）　有意差なし

比の検定（たとえばオッズ比）のp値と95%CI

- p<0.05ならば、比の95%CIに1は含みません。
- p≧0.05ならば、比の95%CIに1が含まれます。

イベント発生オッズ比

2.15（95% CI: 1.19-3.87）有意差あり

2.15（95% CI: 0.81-5.72）有意差なし

I. サンプルサイズとは

なぜサンプルサイズが重要なのか？

臨床試験は被験者の負担や費用を少なくするため、できるだけ少ない被験者数で結果を出したいという事情があります。

- しかし被験者が少なすぎると、母数の推定幅が広くなり（標準誤差が大きくなる）、効果があるのに効果を検出できないことがあります。

- また被験者が多すぎると母数の推定幅は狭くなり、医学的に意味のない差を検出することがあります。

そこで医学的に意味のある差をできるだけ少ない被験者数で示すためにサンプルサイズの設計が重要となります。

〈降圧薬の開発時〉

被験者数が少ない時の問題点

血圧低下幅：平均値20mmHg　95%信頼区間（-4~24mmHg）　n=5
効果がないと判断してよいのか？？

被験者が多すぎる場合の問題点

血圧差：平均値0.5mmHg　95%信頼区間（0.2~0.8mmHg）　n=50,000
効果があると判断してよいのか？？

II. サンプルサイズと有意差の検出の関係

検定結果は二つの過ちを犯す可能性があります

◆ 本当は差がないのに差を検出してしまう過誤（第一種の誤差）

第一種の過誤の大きさ：αエラー（偽陽性）
「効果のない薬を効果があると判断する」
「無罪の人を誤って逮捕する」

◆ 本当は差があるのに差を検出しない過誤（第二種の過誤）

第二種の過誤の大きさ：βエラー（偽陰性）
「効果のある薬を効果がないと判断する」
「真犯人を見逃してしまう」

この二つの過誤を一定の範囲に収めて差を検出するために必要なサンプル数を計算します。
一般的に、αエラーを5%、検出力※を80や90%に設定してサンプルサイズを計算することが多いといえます。

※（1-β）を検出力という

III. サンプルサイズの求め方

サンプルサイズを計算するのに必要な値

臨床的に効果があると判断される差（Δ）

データのばらつき（標準偏差）

- 検出すべき差（Δ）が大きいほど、必要なサンプル数は少なくてよいとされています。
- 小さい差を検出するには多くのサンプル数が必要となります。
- 検出すべき差（Δ）が半分になれば、必要なサンプル数は4倍となります。

- データのばらつきが小さければ、サンプル数は少なくてすみます。
- データのばらつきが大きければ多くのサンプル数が必要となります。
- 標準偏差が2倍になれば必要なサンプル数は4倍となります。

サンプルサイズの計算式

$$N = \frac{2\left(Z_{\alpha/2} + Z_{\beta}\right)^2 \mathrm{SD}^2}{\Delta^2}$$

N ：１グループあたりの人数
標準正規分布の上側 $\alpha/2$%点
標準正規分布の上側 β%点
SD：標準偏差
Δ ：検出すべき差

III. サンプルサイズの求め方

α=5%、検出力80%とした場合、便利な関係式（近似式）があります。

差の場合：16倍の法則

- 標準偏差をΔで割った値の二乗に16をかけるとおよそ1群の必要数となります。
- たとえば130±16mmHgの患者群で、8mmHg血圧降下を検出するには、1群$16\times16^2/8^2$=64例必要になります。

比の場合：3倍の法則

- 1/nのリスクを検出するには約3n例が必要となります。
- 1/1,000で発現するまれな副作用を検出するには3,000例を調べる必要があります。

α=5%、1-β=80%とすると
16倍の法則

$$N=\frac{2(1.96+0.84)^2SD^2}{\Delta^2}$$

$$\approx 16\frac{SD^2}{\Delta^2}$$

3倍の法則
1/nのリスクを検出するには3n例

5 まとめ

統計学は標本（得られたデータ）から母集団の特徴を推定する方法論です

標準偏差（SD）はデータのばらつきを示します

標準誤差（SE）は標本から得られた値が母数（母集団の値）にどれだけ近いかを示します

検定結果を「p値」、推定結果は「95%信頼区間」で表します

有意水準5%での検定結果と、95%信頼区間の推定結果は一致します

できるだけ少数で統計学的な差を検出するためにサンプルサイズを設計します

サンプルサイズの設計には差の16倍の法則、比の3倍の法則があります

〈監修〉
山村 重雄
城西国際大学 薬学部 医療薬学科
臨床統計学研究室 教授　薬学博士