

統計手法の選択

B3 久木田 雄亮

資料

医薬研究者の視点からみた道具としての統計学
奥田千恵子著

第15章 統計手法の選択



統計手法は実験計画時に選ぶ

- 実験の計画段階で統計手法を選ぶことが原則
- 誤った統計手法を用いたために誤った結論を出してしまった実験を解析し直すことはできる
- 実験デザインが悪い研究から統計学を用いて意味のある情報を引き出すことは難しい

解析手法を選択するための観点

- 統計手法はさまざまな観点から選ぶことが必要
1. データは連続量か離散量か？
 2. パラメトリック検定か，ノンパラメトリック検定か？
 3. データは独立しているか，対応しているか？
 4. 群の数は2つか，3つ以上か？
 5. 1変数の解析か，2変数の関係か？

1. 連続量か，離散量か

連続量の検定法

	検定法
正規分布に従うと仮定できる	パラメトリック検定
正規分布に従うと仮定できない	ノンパラメトリック検定

離散量の検定法

	例	検定法
大きな整数値データ	年齢，テストの点数	連続量と同じ
順序カテゴリデータ	薬効判定（著効・有効・不変・悪化）	ノンパラメトリック検定 累積 χ^2 検定
カテゴリデータ	A型・B型・O型・AB型	χ^2 検定
2値変数	生存・死亡	

2. パラメトリックか, ノンパラメトリックか

	パラメトリック検定	ノンパラメトリック検定
対応のある2群の比較	対応のあるt検定	マン・ホイットニーの検定
対応のない2群の比較	対応のないt検定	ウィルコクソンの符号付順位検定
対応のある3群以上の比較	分散分析	クラスカル・ウォリスの検定
対応のない3群以上の比較	反復測定分散分析	フリードマンの検定
相関関係	ピアソンの相関	スピアマンの相関

- ノンパラメトリック検定は正規分布に従うか否かに左右されないが検出力は弱くなる
- データが正規分布に近似することができるのならばパラメトリック検定を用いた方が検出力は高くなる

3. 独立データか，対応データか

対応のあるデータの検定法

	連続量(正規分布)	連続量(非正規分布)， 順序カテゴリデータ	2値変数
2群間の比較	対応のあるt検定	ウィルコクソンの 符号付き順位検定	マクニマーの検定
3群間以上の比較	反復測定分散分析	フリードマンの検定	コ克兰のQ検定

- データの対応づけは事前に行う
- 対応のあるデータ
 - 同一個体から複数データを得た場合
 - ある個体に対して，無作為に選んだ別の個体よりも無作為度が低いデータ

4. 2群か3群以上か

3群以上の実験系で用いる検定法

	実験系全体としての検定	群と群の個別比較
パラメトリック	分散分析 ・用量依存性がある場合 回帰分析	ボンフェローニの補正などの多重比較
パラメトリック, 反復測定	反復測定分散分析	コントラスト変換 (多重比較)
ノンパラメトリック	クラスカル・ウォリスの検定	順位によるノンパラメトリックな多重比較
ノンパラメトリック, 反復測定	フリードマンの検定	

- ・「全ての群の母平均値は等しい」という帰無仮説の検定
- ・「どの群とどの群の間に差があるのか」という個別の比較の場合は多重比較

5. 2変数の関係を求める

変数x	変数y	関係	係数, 検定法
連続量(正規分布)	連続量(正規分布)	直線関係 x,yは入れ替え可能	ピアソンの相関係数
連続量(非正規分布) 順序カテゴリー	連続量(非正規分布) 順序カテゴリー	順位が直線関係 x,yは入れ替え可能	スピアマンの順位相関係数
連続量(正規分布・研究者が任意に指定したり, 制御して一定条件で固定できる変数)	連続量(正規分布)	直線関係 $y=a+bx$	単回帰
同上	同上	曲線関係	非線形回帰
連続量, 離散量	2値変数		ロジスティック回帰
時間	生存率など, 観察期間中1度だけ生じる事象		カプラン・マイヤーの検定 コックスの比例ハザード回帰

さらに進んだ統計手法

1. 多変量解析
2. 共分散分析
3. メタアナリシス

多変量解析

- 変数が3つ以上の場合に多変量解析を行う
- 重回帰, ロジスティック回帰
 - 独立変数, $x_1, x_2, \dots$ から従属変数 y を求める
- 変数を減らす
 - 主成分分析・因子分析
- グループにまとめる, グループを調べる
 - 類似分析・判別分析

共分散分析

- 2つ, それ以上の回帰直線の比較
- 分散分析と回帰分析を組み合わせた解析方法
- 水準間の差の検定が目的(分散分析と同じ)
- 共変量を使って, 誤差の部分を小さくし, 検定力を高めようとする

メタアナリシス

- 独立して行われた2つ以上の研究結果を統合するための手法
- 過去の似た研究論文から将来の治療の示唆を得るために行う
- 標本数が少ない研究結果が有意でなかった同質の研究で行うと検出力が上がり有意差が出ることもある