

159. 医学統計

From MY point of view

- 記述統計と推測統計
- カテゴリーデータと数量データ
- カテゴリーデータの解析
- 数量データの解析(2 群比較): Paired? Unpaired? パラメトリック? ノンパラメトリック?
- 数量データの解析(3 群以上): ANOVA? post hoc test?

1. 記述統計と推測統計

記述統計: 母集団(本当に知りたい集団)=対象集団: 国民調査、当院における麻酔件数、鎮痛薬 A の内服状況 etc.

推測統計: 対象集団の情報から母集団(人類)の状況を推測する→医学研究はほぼコレ

できるだけ正確に母集団(人類)の情報を推測するために... サンプル数、多施設、多人種、傾向マッチング、Meta-analysis

2. 数量データとカテゴリーデータ

数量データ: 測定できる=数値が等間隔である (身長 152cm の次は 153cm, 154cm 155cm で間1cm は等間隔)

カテゴリーデータ: 測定できない=数値を記号に置き換えられる (クチコミの★★★★☆=ABCDE ランクで置き換えられる)

患者名	血液型	出身地	将棋の段位	ASA-PS	NRS	身長	体重	EF(%)	虫歯の数
Aさん	A	富山	5級	2	0	165	52.3	67.1	0
Bさん	B	新潟	4級	1	1	183	70.1	65.3	1
Cさん	AB	静岡	1級	3	2	169	57.5	62.1	0
Dさん	O	栃木	初段	4	5	173	65.4	70.3	2
Eさん	A	愛知	3級	5	10	172	62.4	23.1	8

カテゴリーデータ: 血液型, 出身地, 将棋の段位
数量データ: ASA-PS, NRS, 身長, 体重, EF(%), 虫歯の数

名義データ: 血液型, 出身地
順序データ: 将棋の段位
連続データ: 身長, 体重, EF(%), 虫歯の数
離散データ: ASA-PS, NRS

3. カテゴリーデータの解析

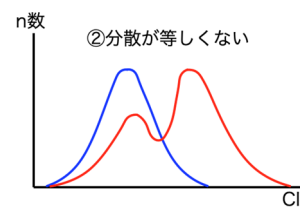
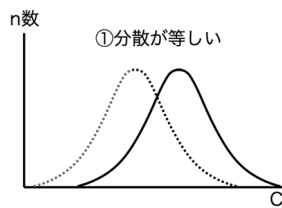
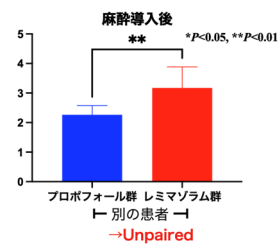
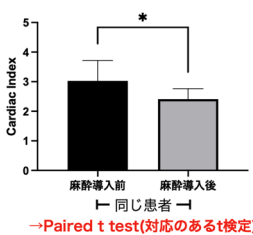
サンプル数が大きい→カイニ乗検定(Chi-squared test)

サンプル数が小さい(n<10)→フィッシャーの正確確率検定(Fisher's exact test)

4. 数量データの解析(2 サンプル比較)

Paired? Unpaired?

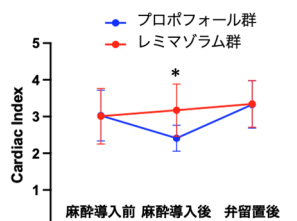
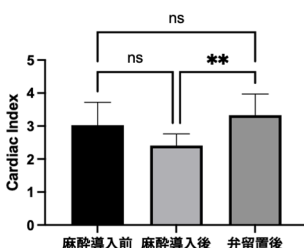
Unpaired の場合



①分散が等しい(F 検定で $P \geq 0.05$)→unpaired t test (パラメトリック) 数値は平均値±標準偏差 or 標準誤差

②分散が等しくない(F 検定で $P < 0.05$)→Mann-Whitney U test (ノンパラメトリック) 数値は中央値(最大値, 最小値)

5. 数量データの解析(3 サンプル以上)



3 サンプル以上: One-way ANOVA(左)

2 群以上 x2サンプル以上: Two-way ANOVA(右)

ANOVA はサンプル間に差があるかどうかを調べる検定。各ポイントの有意差を調べるためには Post hoc test (Bonferroni, Tukey, Sidak など) が必要 →Bonferroni が最もよく使われている(慣習)

2021.6.10.

Hisakatsu Ito

第 1 版(最終更新:2021.6.10.)

コラム P 値とは？ :P 値とは比較している対象が「同じもの」である確率(probability)である。A 群と B 群を比較して $P < 0.01$ ということは A 群と B 群が同じものである確率が 1%以下→たぶん違うもの→「有意差がある」と判定する。逆に $P = 0.12$ だったとき A 群と B 群が同じものである確率はたった 12%である。P 値の意味を知ると「有意差がない」=「同じもの」ではないことがわかる。ちなみに $P < 0.05$ を有意差ありとするという定義は慣習によって決まったもので必ずしも統計学的には正しくない。