トリム平均を利用した統計解析 ブートストラップ法による 線型モデルの適用

堺伸也*, 小山暢之**, 井上貴博*,
幸坂美樹*, 松山千恵*, 山田剛久*

* イーピーエス株式会社 統計解析部
** 三共株式会社 臨床解析部

発表までの経緯(1)

- ある臨床試験について解析計画を検討
  - 3群 - それぞれの群で異なる薬剤を投与
  - 反応変数;　あるデータの投与前後の変化差(連続データ)
  - 共変量;　性別, 年齢, 反応変数の投与前の値

  ⇒ (通常であれば,) 線型モデルにより, 共変量を調整した群間差を算出し評価する.

- 外れ値
  - 外れ値がいくつか発生することが事前に分かっていた.
  - 解析上の厄介な問題の一つ.

- 対処方法
  - 「不適切な値」である正当な理由が付けば除外.
  - しかし, 臨床試験データでは「不適切な値」である理由が容易に付かないことが多い.
  - 今回の場合は, もともとのデータの分布に外れ値が含まれる.

  ⇒ 外れ値の影響を受け難いトリム平均を用いたい.

発表までの経緯(2)

外れ値のある臨床試験データに対して，
トリム平均を用いた評価をし，更に
共変量を調整するために線型モデルを利用して
群間比較を行いたい．

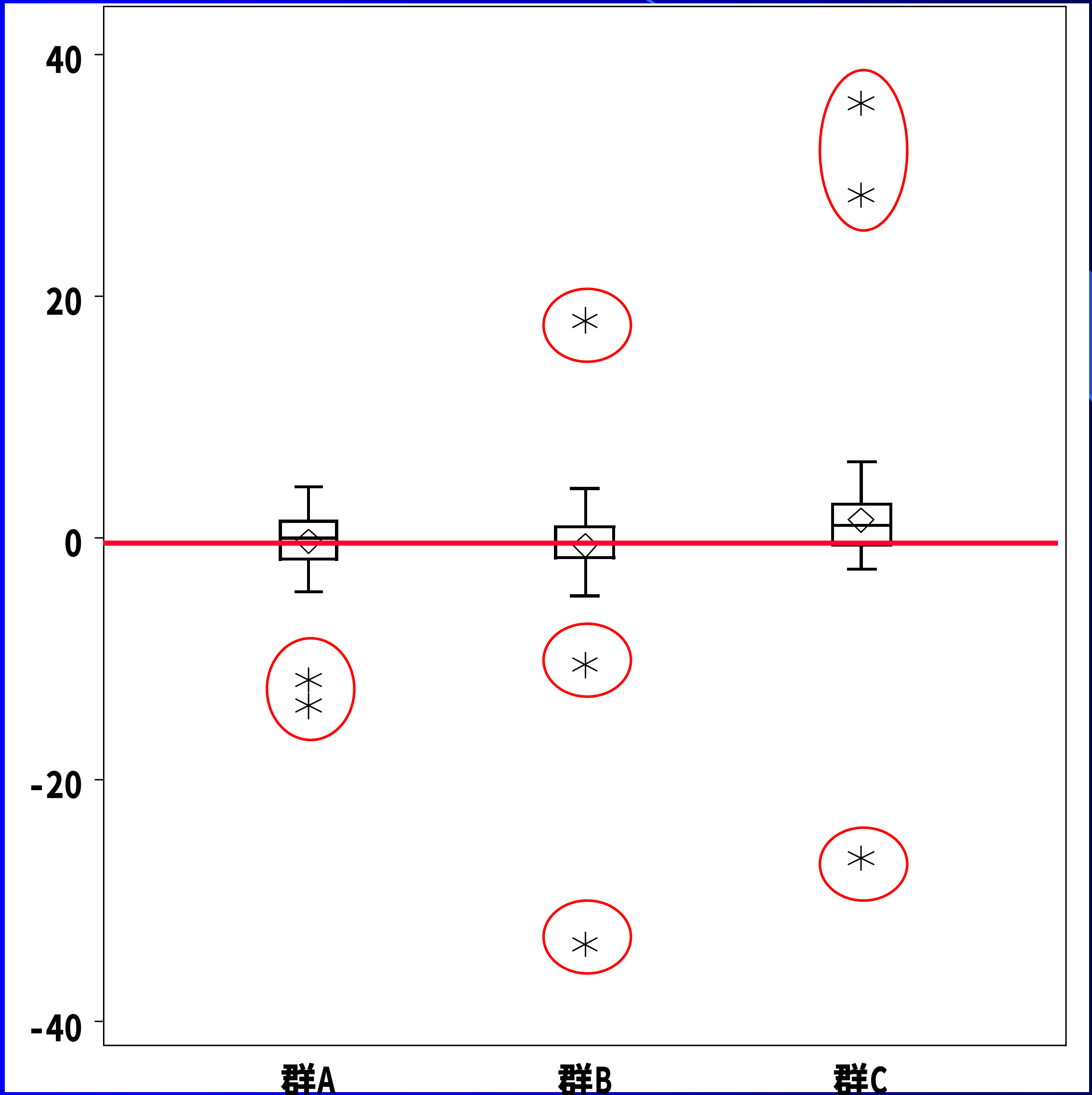
ブートストラップ法を用いてトリム平均に対応した
線型モデルのパラメータ推定(95％信頼区間)
を行うSASマクロを作成した．

次頁; 例を挙げつつ，経緯を確認

外れ値のあるデータセット

※仮想
データ
(1群100例)

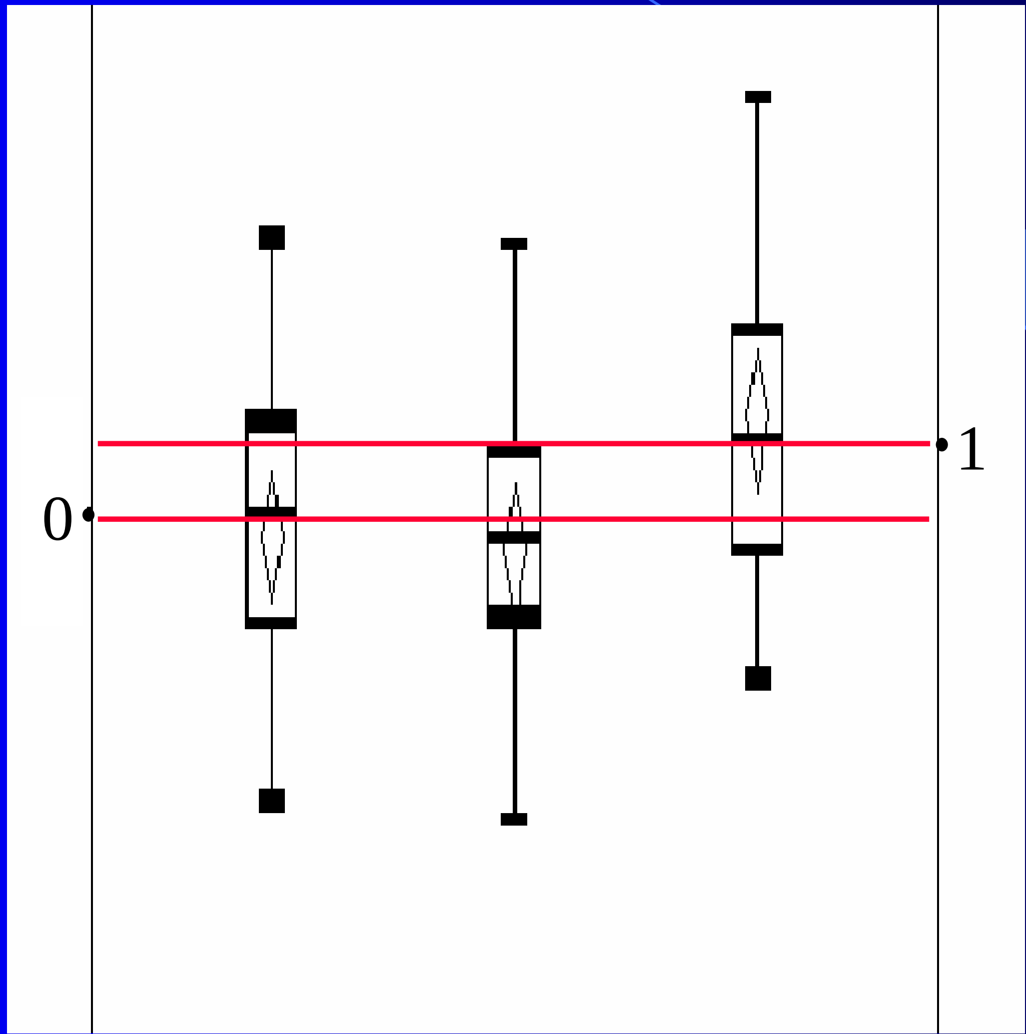
あるデータの
投与前後の
変化差の分布



データの
固まっている
-5〜5 の
部分を拡大

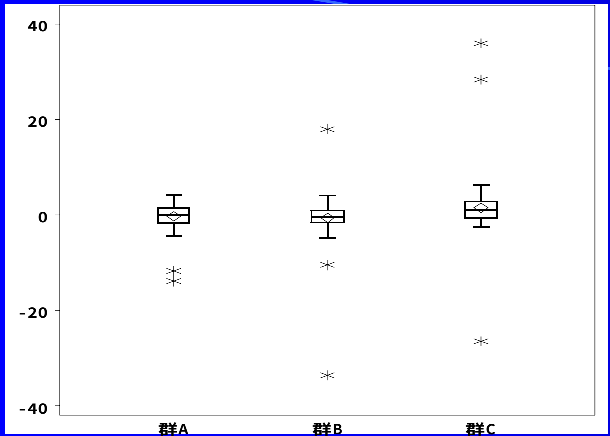
外れ値のあるデータセット

※仮想データ (1群100例)

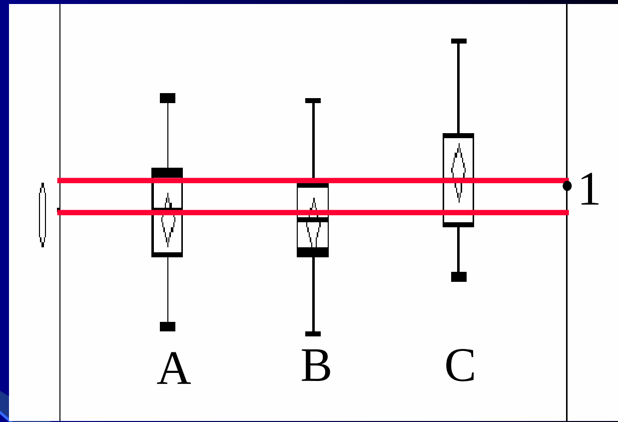


群A 群B 群C

この部分に注目すれば 1.0～1.5前後 群Cの値が大きい



要約指標



群Bと群C の差

群B-群C	95%信頼区間	幅
-2.1	(-3.5～-0.7)	2.8

群B-群C	95%信頼区間	幅
-1.6	(-2.1～-1.0)	1.1

特殊な手法は用いずに、
トリム後に95%信頼区間を求めた

- 平均値
  - 外れ値の影響を受け易い
  - さまざまな統計手法が開発されている
- トリム平均　それぞれの群でトリム
  - 外れ値の影響を受け難い
  - 統計手法が開発されていない
- 中央値%トリム平均 (100例のときは上下2例ずつ除く)
  - 外れ値の影響を受け難い -4.4, ・・・, 3.4, 3.5, ~~3.7, 4.2~~
  - 情報のロスが大きい
  (中央値は49.9%トリム平均であると考えることも出来る)

信頼区間の幅は狭いが,これでよいのか疑問

の平均値

線型モデルでの解析結果 (群B-群C)

- 更に, 共変量を調整した解析を行いたい
- 群, 性別, 年齢, 反応変数の投与前 を説明変数とした線型モデルでの解析

差が大きくなっている

外れ値の影響で
信頼区間の幅が広い

		群間差	95%信頼区間		幅
平均値 ※ [最小二乗平均]	群B-群C	-2.2	-3.4	-1.0	2.4
2%トリム平均 ※ [各群で2%トリムした後の最小二乗平均]	群B-群C	-1.6			

※ 線型モデルで調整した値

信頼区間の幅は狭いが, これでよいのか疑問

ほとんどの症例の状態を反映している

外れ値について

- トリム平均
  - 単純に外れ値を除いて解析すれば，95%信頼区間は不適切．
  - 外れ値を考慮した95%信頼区間を求めたい

⇨ ブートストラップ法

数値実験により評価

平均値

発表の流れ

1. ブートストラップ法 トリム平均の95%信頼区間
2. 作成した解析プログラム
3. 数値実験 ― シミュレーション条件
4. 数値実験 ― 結果

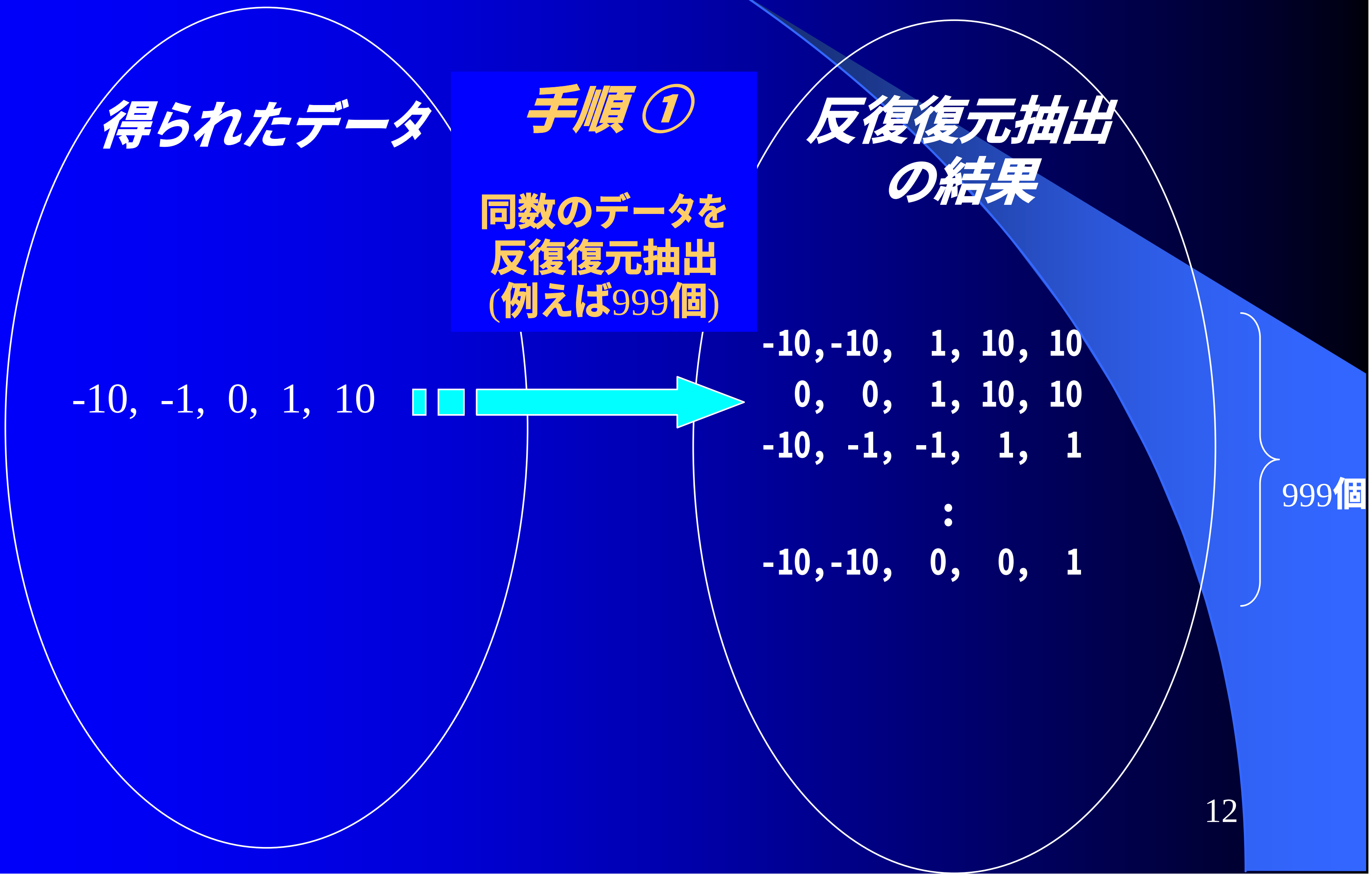
ブートストラップ法
トリム平均の95%信頼区間

外れ値の存在が
どのように考慮されているか.

ブートストラップ法

- ブートストラップ法は,
  得られたデータに基づき反復復元抽出することで
  推定量の分布を得て, 解析する方法
  (Efron , 1970年代)
- 次頁; トリム平均の95%信頼区間の算出例
  - ノンパラメトリック・ブートストラップ法
  - パーセンタイル法による95%信頼区間

ブートストラップ法
トリム平均の95%信頼区間(1)



(2)

得られたデータ

~~-10~~, -1, 0, 1, ~~10~~

手順②

トリムする

反復復元抽出の結果

~~-10~~, -10, 1, 10, ~~10~~
~~0~~, 0, 1, 10, ~~10~~
~~-10~~, -1, -1, 1, ~~1~~
:
~~-10~~, -10, 0, 0, ~~1~~

999個

(3)

手順③

トリム平均を求める

得られたデータ

~~-10~~, -1, 0, 1, ~~10~~ → 0

反復復元抽出の結果

~~-10~~, -10, 1, 10, ~~10~~ → 0.3
~~-1~~0, 0, 1, 10, ~~10~~ → 3.7
~~-10~~, -1, -1, 1, ~~1~~ → -0.3
:
~~-10~~, -10, 0, 0, ~~1~~ → -3.3

999個

(4)

手順④

95%信頼区間を求める

得られたデータ

~~-10~~, -1, 0, 1, ~~10~~ → 0

トリム平均の点推定値0

反復復元抽出の結果

外れ値が利用されている

~~-10~~, -10, 1, 10, ~~10~~ → 0.3
~~0~~, 0, 1, 10, ~~10~~ → 3.7
~~-10~~, -1, -1, 1, ~~1~~ → -0.3
:
~~-10~~, -10, 0, 0, ~~1~~ → -3.3

999個

小さい順に並べて
25番目と975番目の数値を抽出
95%信頼区間 (-3.7 , 3.7)

ブートストラップ法
トリム平均の95%信頼区間(5)

- トリム平均 (95%信頼区間)　　0.0 (-3.7 〜 3.7)
- 反復復元抽出を行った後に, トリムすることが重要.
  - 反復復元抽出のときには「外れ値」を含めておく.
- 先にトリムすると, 完全に解析から「外れ値」を除外したことになり, 推定誤差を過小評価することになってしまう.

作成した解析プログラム

信頼区間の求め方

トリム平均の95%信頼区間

- 反復復元抽出（999回）を行い，トリムの処理を行う．
- 平均値を求める．
- 平均値を小さい順に並べて25番目の数値と975番目の数値を95%信頼区間の限界値とする．

トリム後に線型モデルを当てはめる場合の群間差の95%信頼区間

- 群毎に反復復元抽出（999回）を行い，群毎にトリムの処理を行う．
- 最小二乗法で群間差の点推定値を求める．
- 点推定値を小さい順に並べて25番目の数値と975番目の数値を95%信頼区間の限界値とする．

作成した解析プログラム

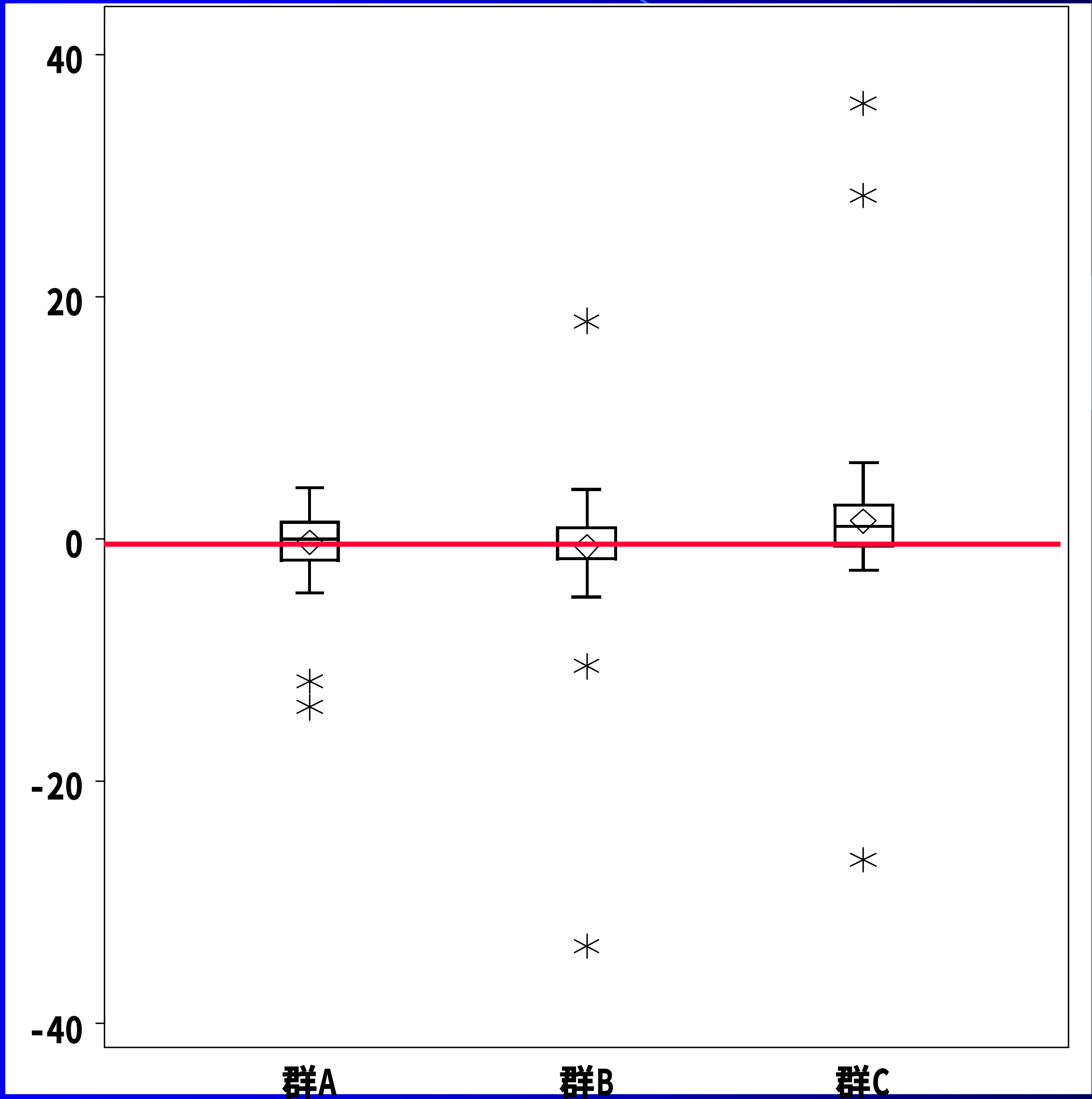
- 3つの群 (各群で反復復元抽出)
- SAS/STAT SURVEYSELECTプロシジャ (METHOD=URSオプション) を利用
- 任意の2つ群の群間差を推定

・3群での非劣性試験をイメージしてプログラム作成しています.

・3つの比較対がありますが, 多重性の調整はしていません.

外れ値のあるデータセット

※仮想データ
(1群100例)



仮想データに対する解析結果

- 群, 性別, 年齢, 反応変数の投与前を説明変数とした解析

		群間差	95%信頼区間		幅
平均値	群B-群C	-2.2	-3.4	-1.0	2.4
2%トリム平均 [ブートストラップ法]	群B-群C	-1.6	-3.0	-0.9	2.1
2%トリム平均 [完全に上下2%を除外]	群B-群C	-1.6	-2.1	-1.0	1.1

線型モデルで調整した値
ブートストラップ法のリサンプリング回数は1999回

数値実験 – シミュレーション条件

外れ値のあるデータに対して
今回の方法がどの程度，有効か？

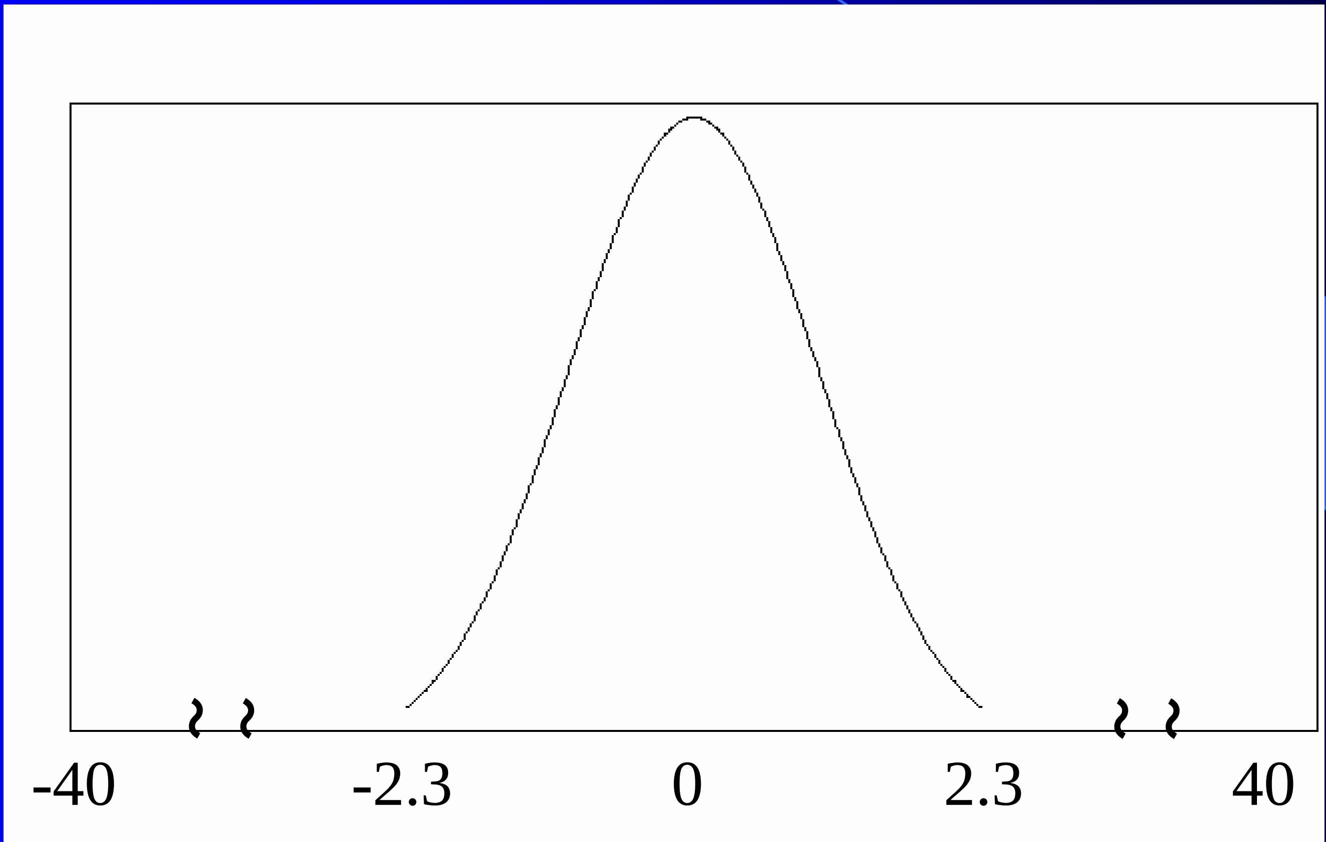
数値実験用のデータセット

- 1群100例　合計300例
- 乱数により以下のようなデータセットを作成

反応変数	投与群	性別	年齢	反応変数の投与前の値
0.256	A群	女	45	5
-0.418	A群	男	66	6
-1.941	B群	男	53	15
-0.589	C群	男	50	24
・	・	・	・	・

- モンテカルロ･シミュレーション
  - このようなデータセットを 10,000回 作り直してそれぞれ解析した
  - 反応変数 ＝ $\beta_0 + \beta_1$(投与群1)+ β_2(投与群2)+ β_3(性別) + β_4(年齢)+ β_5(投与前の値) + 誤差
  - ブートストラップ法の反復復元抽出は1999回

反応変数の誤差の分布



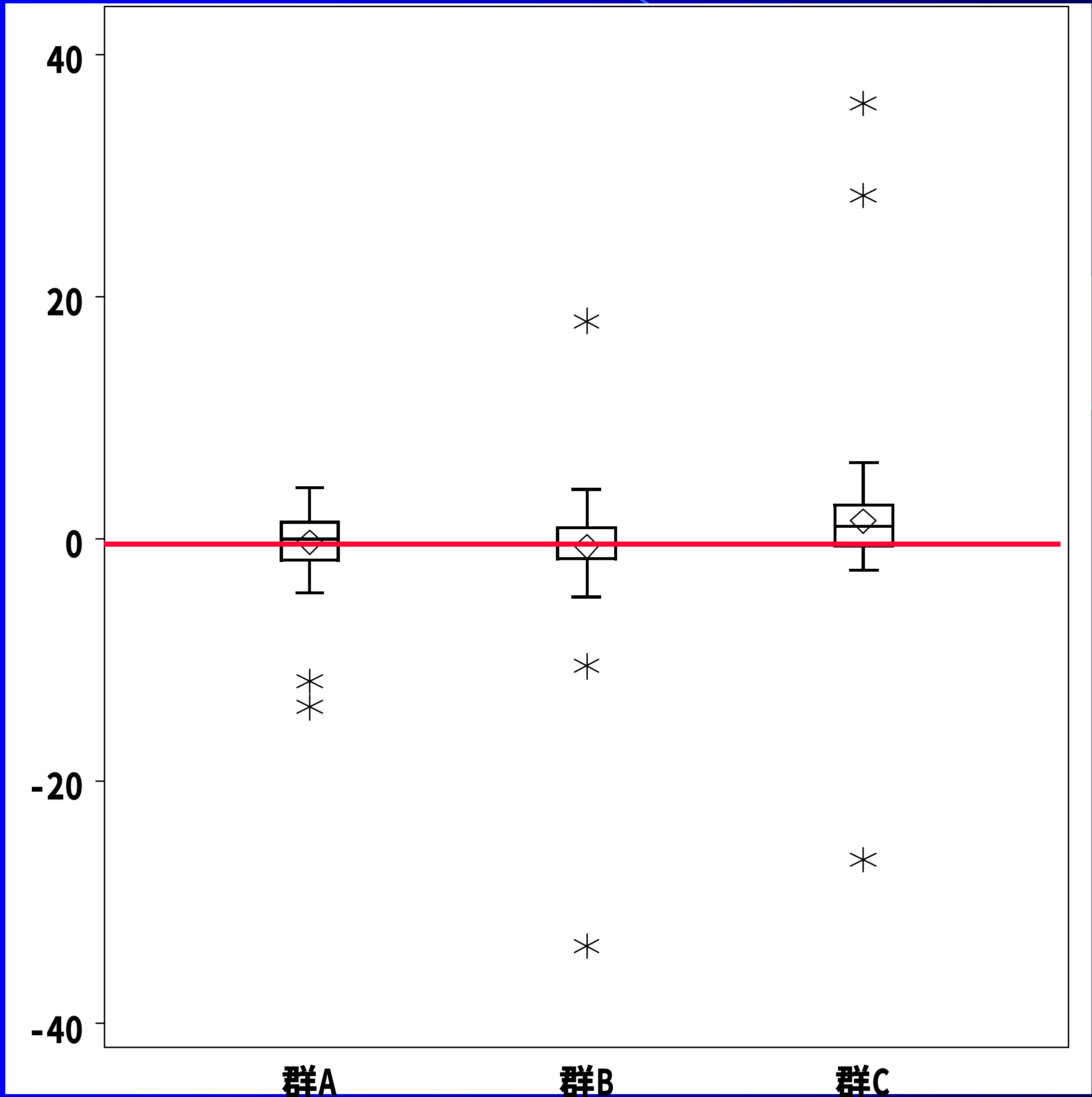
1% 外れ値

98% 正規分布(分散4)

1% 外れ値

反応変数の分布例

※仮想
データ
(1群100例)



評価方法

- 設定した各群の効果
  - 群A　0
  - 群B　0
  - 群C　+1
- 評価する比較対
  - 群A-群B　0 (帰無仮説)
  - 群B-群C　-1 (対立仮説)
- 次の数値を示した
  - 第1種の過誤の確率
  - 95%信頼区間の被覆確率
  - 検出力

数値実験 – 結果

群A-群B(=0 *帰無仮説*)の結果

- 第1種の過誤の確率 (95%信頼区間が真値 0 を含まない割合)
- 95%信頼区間の被覆確率 (95%信頼区間が真値 0 を含む割合)

	解析方法	第1種の過誤の確率	被覆確率 (1 - 第1種の過誤の確率)
	平均値の差	5.0%	95.0% *95%に近い*
	2%トリム平均の差 [ブートストラップ法]	3.6%	96.4% *95%よりやや大きい*
誤り	2%トリム平均の差 [完全に上下2%を除外]	7.3% *利用は難しい*	92.7%

線型モデルで調整した値　　シミュレーション回数10000回

群B-群C(=-1 *対立仮説*)の結果(1)

- **検出力** (95%信頼区間が0より小さい割合)
- 95%**信頼区間の被覆確率** (95%信頼区間が真値 -1 を含む割合)

解析方法	検出力	被覆確率 *帰無仮説と同様の結果*
平均値の差	46.6%	95.4%
2%トリム平均の差 [ブートストラップ法]	71.6% *検出力が高い*	96.6%
誤り 2%トリム平均の差 [完全に上下2%を除外]	91.7% *非常に検出力が高い*	92.1%

第1種の過誤の確率が保たれていない

線型モデルで調整した値　シミュレーション回数10000回

群B-群C(=-1 *対立仮説*)の結果(2)

- 推定値と95%信頼区間の幅の分布

mean±SD

解析方法	推定値 (真値 -1)	95%信頼区間の幅
平均値の差	-0.995±0.555	2.162±0.437
2%トリム平均の差 [ブートストラップ法]	-0.999±0.319 *バラツキが小さい*	1.579±0.374 *信頼区間の幅は狭い*
誤り 2%トリム平均の差 [完全に上下2%を除外]	-0.999±0.319	1.093±0.154

線型モデルで調整した値　シミュレーション回数10000回

数値実験の結果

- 平均値の差
  - 第一種の過誤の確率　保てている
  - 95%信頼区間の被覆確率　95%
  - 検出力　低い
- 2%トリム平均の差 [ブートストラップ法]
  - 第一種の過誤の確率　保てている
  - 95%信頼区間の被覆確率　ほぼ95%
  - 検出力　高い(71.6% ← 46.6%)
- 2%トリム平均の差 [完全に上下2%を除外]
  - 第一種の過誤の確率　保てていない

※ 線型モデルで調整した値

おわりに

おわりに

- トリム平均は, 外れ値に対して, 平均値よりもロバストな推定量である.

- トリム平均について, ブートストラップ法を用いることで, 線形モデルを適用したパラメータ推定(95%信頼区間)のが可能.

- 臨床データの解析において, トリム平均を利用することは選択肢の一つになりうる.

作成したプログラムについて補足

- プログラムは次の指定を変えて実行することが可能
  - 線型モデルのモデル式
  - 乱数のシード
  - 復元無作為抽出の回数
  - 信頼区間の信頼係数
  - トリムするデータの割合を指定

back up

トリムする割合について

- トリムする割合について
  - 恣意的にある群に有利になるように定めない
  - 事前に定義
  - 二重盲検した試験であれば，盲検下レビューを利用

  - 例; 正規分布に近づける
    ←→ できるだけ多くのデータを解析に用いる