

# 効果量・検出力・サンプルサイズ

M1 布井雅人

# 本日の発表内容

- 効果量について
- 検定力・検出力について
- 統計ソフトG \* Powerについて

# 統計的検定では

- 設定された有意水準よりもp値が低いと“有意である”と判断
- この場合の“有意である”は**実質的**なものと言えるのか？

- p値: サンプルサイズによって決まる
  - … サンプルサイズが大きくなると有意と判断されやすくなる
- 決して**効果の大きさを表すものではない**
  - 10人では非有意だが、1000人集めれば有意になることがある

# サンプルサイズの影響例

- 2変量A・Bの相関係数  $r=0.2$

サンプルサイズ=10  
…  $p=.58$  } 非有意

サンプルサイズ=1000  
…  $p<.000$  } 有意

サンプルサイズ=98  
…  $p = .048$  }

相関係数そのものに変化はない  
 $r=0.2$ は大きくないはず!!

- 統計的検定で見ているもの
  - 効果(差)があるかどうか
  - × 効果(差)の大きさ
- サンプルサイズに影響されない効果の大きさを示す必要あり

**効果量**

# 効果量d

- 帰無仮説からのずれを標準化したもの
  - 平均値の差を標準化したもの
  - 平均値と標準偏差を用いて算出する
- 近年論文で、p値だけでなく、効果量を報告することが推奨されている

# 効果量 t検定

– 1母集団のt検定の場合  $H_0: \mu = \mu_0$

$$d = \frac{\mu - \mu_0}{\sigma}$$

– 2母集団のt検定の場合  $H_0: \mu_1 = \mu_2$

$$d = \frac{\mu_2 - \mu_1}{\sigma}$$

$\frac{\text{2条件の平均値の差}}{\text{2条件の標準偏差の平均値}}$

# 効果量の目安

- Cohen(1988)が提案
  - 行動科学における慣習的な数値

Table. 効果量の主な指標とその大きさの目安(Cohen, 1988 から抜粋).

	指標	効果量の目安			説明できる変動の割合 (効果/[効果+誤差])		
		小	中	大	小	中	大
2変量の積率相関	$r$	0.10	0.30	0.50	1%	9%	25%
t検定	$d$	0.20	0.50	0.80	1%	6%	14%
分散分析	$f$	0.10	0.25	0.40	1%	6%	14%
多変量分散分析	$f^2$	0.02	0.15	0.35	2%	13%	26%

注意:  $r$  (の絶対値)は1を超えないが,  $d$  や  $f$ ,  $f^2$  は1を超えることもある.

$$f^2 = \frac{\text{効果による変動}}{\text{誤差による変動}}$$

# 効果量 分散分析

- 分散分析の効果量の表し方は様々
- 関連度(strength of association)で表すことが多い
  - ある要因の効果がデータの変動のどれだけの割合を説明できるかを示したもの

# 関連度

- $\eta^2$  : 特定のサンプルについて求めたもの
- 偏  $\eta^2$  : 参加者の要因やその他の主効果・交互作用によって説明できる変動を除いて求めたもの

$$\text{偏 } \eta^2 = \frac{\text{効果の自由度} \times \text{効果のF値}}{\text{効果の自由度} \times \text{効果のF値} + \text{効果の検定に用いる誤差の自由度}}$$

- $\varepsilon^2$   $\omega^2$  : 母集団推定値
  - …他の研究データと比較する際はこっちがよい

# 効果量の用途

- 結果を示す(効果の大きさ)
- メタ分析に用いる
  - メタ分析の回を参照
- 検出力分析に用いる
  - サンプルサイズを決定する
  - 第二種の過誤が起きてないかを確認するため

# 検出力

		実際に成り立っているのは	
		帰無仮説: $H_0$	対立仮説: $H_1$
検定結果	$H_0$	正しい ( $1-\alpha$ )	第二種の過誤 ( $\beta$ )
	$H_1$	第一種の過誤 ( $\alpha$ )	正しい ( $1-\beta$ )

検出力

# 検出力

- 第二種の過誤を犯さない確率のことを検出力と言う
  - 帰無仮説が棄却されるはずの場合に棄却する確率
    - 検出力: 0.8
      - 100回仮説検証を行った場合に80回は検出できる
    - 検出力: 0.5
      - 2回に1回しか仮説は支持されない

# 第一種の過誤と第二種の過誤

- **第一種の過誤を犯す確率を低くする**  
(=誤って帰無仮説を棄却しないようにする)  
**第二種の過誤を犯す確率が高くなる**
- **第二種の過誤を犯す確率を低くする**  
(=帰無仮説を棄却する確率を上げる)  
**第一種の過誤を犯す確率が高くなる**

第一種の過誤と第二種の過誤はトレードオフの関係

# サンプルサイズ

- サンプルサイズを増やしていくと、検出力は向上する
    - サンプルが母集団に近づいていく
- サンプルサイズは大きければ大きいほど良い  
but コストなどを考慮する必要あり

第一種の過誤も第二種の過誤もなるべく犯さない  
ような、適切なサンプルサイズを設定する

# サンプルサイズを決定する

- サンプルサイズを決めるには…
  - 検出力
  - 効果量 を決定すればよい
  - 検出力はCohenの示唆によって0.8に設定することが多い
  - 効果量は研究分野によって異なる
    - 社会科学では中～小程度の効果量が多い
    - わずかな結果でも見逃したくない→小
    - 大きな効果があれば検出したい→大

# G \* Power

- 検定力分析のためのプログラム
- 東京大学栗田さんによる解説
  - <http://www.nuis.ac.jp/pda-j/doc/00008/gpowermanual.html>
- インストール方法など
  - <http://ningen.human.tsukuba.ac.jp/~p-950780/gpower/>

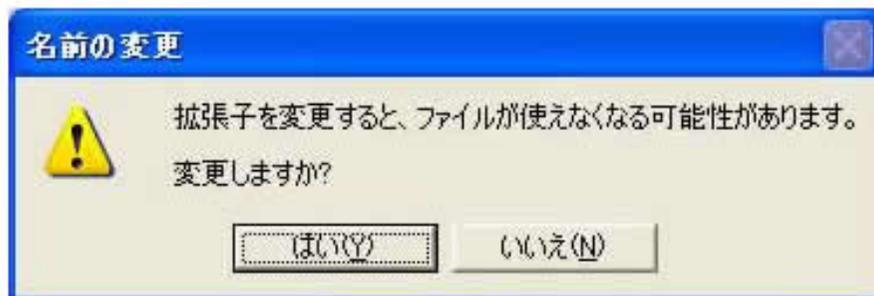
# インストール

- G \* PowerはWindows上では使いにくいようです。
- 少し変わったことをするといいみたいです。
  - <http://www.psych.uni-duesseldorf.de/aap/projects/gpower/>  
にアクセス
    - 筑波大学三好先生のホームページ
  - 下のほうのG \* Power2がフリーソフトになります
  - [GP-Setup.exe](#) (800 K)をクリック
  - インストーラーを起動させインストールしてください
    - サテライトのPCではローカルディスクに保存できないので、マイドキュメントに保存してください

# バッチファイルの作成

- G \* Powerを一発で起動するためのファイルらしい・・・
  - txtファイル(メモ帳)に以下の文字をコピーして、“GPower”と名前をつけデスクトップに保存
    - @echo off
    - cd m:¥Gpower
    - call us.bat
    - GPOWER
  - “コントロールパネル” → “デスクトップの表示とテーマ” → “フォルダオプション” → “表示” で “登録されている拡張子は表示しない” のチェックを外す
    - 下から3番目くらい
    - サテライトのPCは始めから外れているようです

- 保存したテキストファイルの拡張子を“txt”から“bat”に変更する

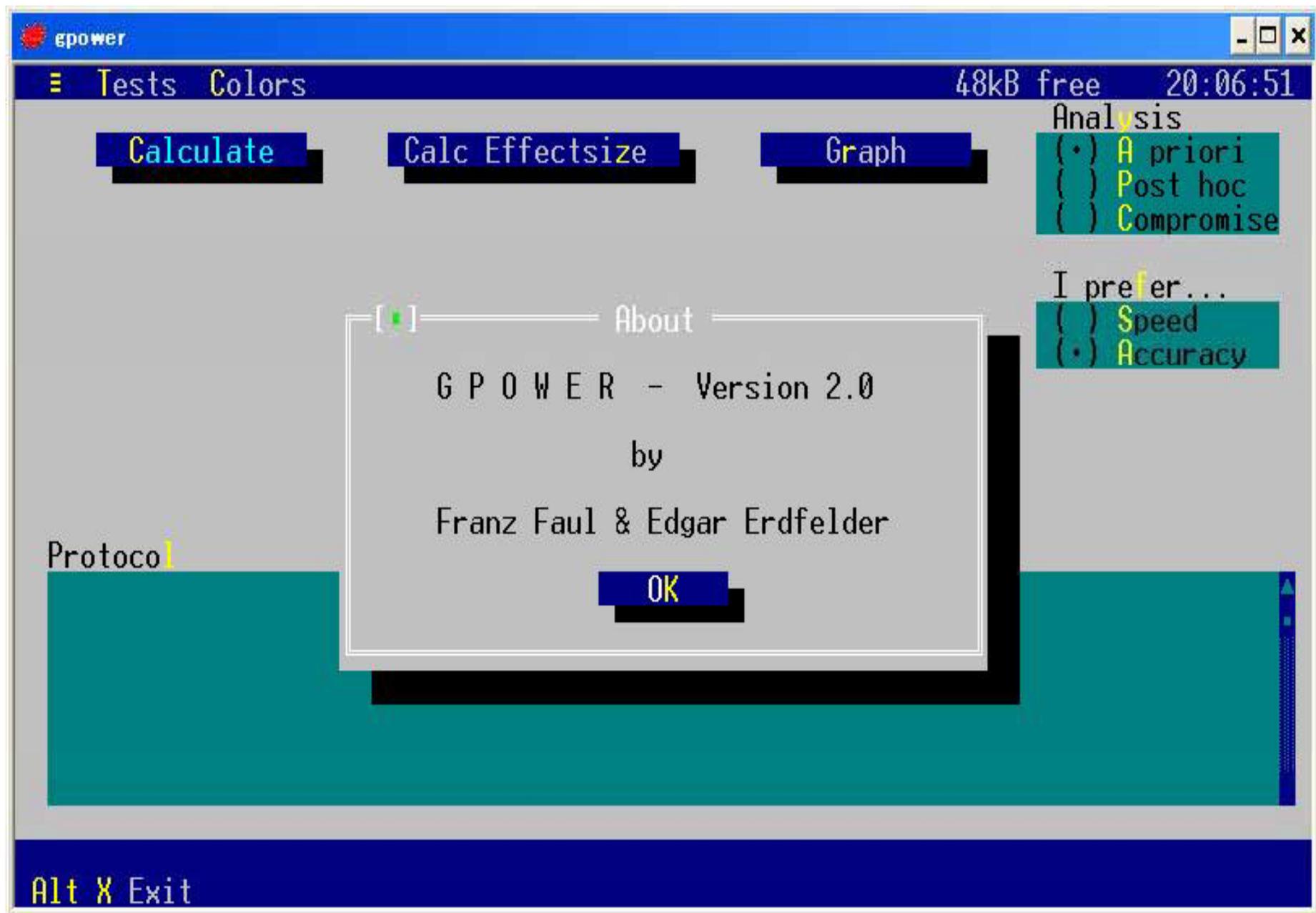


- これが出てくるが変更する



デスクトップ上にこれが出てきたらバッチファイルの作成は完了

- 作成したバッチファイルをG \* Powerを保存した同じフォルダに移動させる
  - “マイドキュメント” → “G \* Power”
- 移動が完了したら、そこでバッチファイルをダブルクリックするとG \* Powerが起動します
  
- あとはショートカットをデスクトップ上に作ると便利



# メイン画面

- 3つのメニューからなる
  - Calculate (メインウィンドウ)
    - 検定方法の選択, 計算精度の選択, 検定力分析の方法の選択, 各種値の設定など
  - Calc Effectsize
    - 検定力分析に用いる効果の大きさを計算によって求めたいときに選択
  - Graph
    - メインウィンドウでの計算結果をグラフ表示する際選択

# Calculate 画面右側

- Analysis
  - A priori
    - 効果量・有意水準・検定力からサンプルサイズを求める
  - Post hoc
    - 効果量・有意水準・サンプルサイズから検定力を求める
  - Compromise
    - 効果量・ $\beta/\alpha$ 比・サンプルサイズから有意水準・検定力を求める

# 続き

- I prefer...
  - Speed
    - 速度優先
  - Accuracy
    - 精度優先
- Test is...
  - One tailed
    - 片側検定
  - Two tailed
    - 両側検定
- 以上の条件を指定して、選択します。

# Calculate ~ A prioriを選択した場合 ~

## 画面左側

- Effect size d
  - 効果の大きさを指定
  - 画面下部にCohenの基準が書いてある
- Alpha
  - 有意水準を指定
- Power
  - 検定力を指定

## 画面中央部

- …結果の表示
- Delta
  - 非心t分布の非心度
- Critical t
  - 棄却の限界
- Total sample size
  - 必要なサンプルサイズ

# Calculate ~ Post hoc を選択した場合 ~

## 画面左側

- Effect size d
  - 効果の大きさを指定
  - 画面下部にCohenの基準が書いてある
- Alpha
  - 有意水準を指定
- Sample size n1/n2
  - 各集団のサンプルサイズを指定

## 画面中央部

- Delta
  - 非心t分布の非心度
- Critical t
  - 棄却の限界
- Power
  - 算出された検定力

Compromiseを選択した場合は割愛

# 検定方法

メインウィンドウで検定方法の選択を行うことが可能

- F2 : t-test for means
  - 平均値差に関するt検定(2グループ)
- F3 : t-test for correlations
  - 相関係数に関するt 検定
- F4 : Other t-tests (対応がある場合など)
- F5 : F-test on means in the ANOVA
  - 分散分析における F検定
- F6 : F-test in multiple regression
  - 重回帰分析における F検定
- F7 : Other F-tests
- F8 :  $\chi^2$ -test

# Calc Effectsize

- Mean (Group1/Group2)
  - 2群の母平均を入力
- Sigma (within each group)
  - 母標準偏差を入力

効果量が算出される

- Calc & Copyでメインウィンドウに代入可能

# Graph

## 画面左側

- 1行目
  - X軸に指定されたものを表示
  - 値の範囲選択を行う
- 2,3行目
  - X軸・Y軸に指定されなかった残りの2つが表示
  - 数値の指定を行う
- Number of steps
  - プロットする点の数を指定

## 画面右側

- I want to see...
  - Y軸にくる変数の指定
- as a function of...
  - X軸にくる変数の指定
- Option
  - Grid: グリッド線の表示
  - Plot values: 数値の表示
  - Plot lines: 点の連結
  - Protocol: グラフに用いた数値をプロトコルに保存
  - Speed mode: グラフ作成時間を早く

## Ex.) Bayliss et. al. (2006)

- 視線方向と同じ方向にある対象と反対の方向にある対象の好ましさの違い
- 被験者内要因
- 被験者 24名
- Congruency :  $M=4.74$ ,  $SD=0.73$
- Incongruency :  $M=4.59$ ,  $SD=0.78$
- $t(23)=4.64$ ,  $p<.001$ ;  $r=.49$
- この際の効果量の求め方は以下の通り

$$f = \frac{M_1 - M_2}{\sqrt{\sigma_1^2 + \sigma_2^2 - 2r\sigma_1\sigma_2}}$$

- 対応のあるt検定はF4
- それぞれの数値を入力
  - $df(\text{自由度}) = \text{被験者数} - 1$
- $ES = 0.257$
- $\text{Power} = 0.2265$

# ファイルの保存と終了

- 保存は終了の際にのみ可能
- メインウィンドウ上で“Alt+X”を押すと終了
  - “OK”で保存
  - “Discard”で保存せず終了
- 詳しくは栗田さんのHPを
  - <http://www.nuis.ac.jp/pda-j/doc/00008/gpowermanual.html>

# Rで

- Rでも出来るようですが、時間の都合上割愛します
  - <http://aoki2.si.gunma-u.ac.jp/R/>
  - パワーアナリシス参照

# 参考文献

- <http://www.geocities.co.jp/WallStreet/4716/poweranalysis.htm>
- <http://www.geocities.co.jp/WallStreet/4716/poweranalysis2.htm>
- <http://home.hiroshima-u.ac.jp/nittono/QA.html#Stat>
- <http://aoki2.si.gunma-u.ac.jp/R/>
- <http://www.nuis.ac.jp/pda-j/doc/00008/gpowermanual.html>
- <http://www.psych.uni-duesseldorf.de/aap/projects/gpower/>
- Bayliss, A. P., Paul, M. A., Cannon, P. R., & Tipper, S. P. (2006). Gaze cuing and affective judgment of objects: I like what you look at. *Psychonomic Bulletin & Review*, 13(6), 1061-1066
- 永田靖 (2003) サンプルサイズの決め方 朝倉書店