2021/4/26　臨床データ解析演習③

2群の比較と分散分析

北海道大学　医学統計学
横田　勲

1

サンプルデータを利用

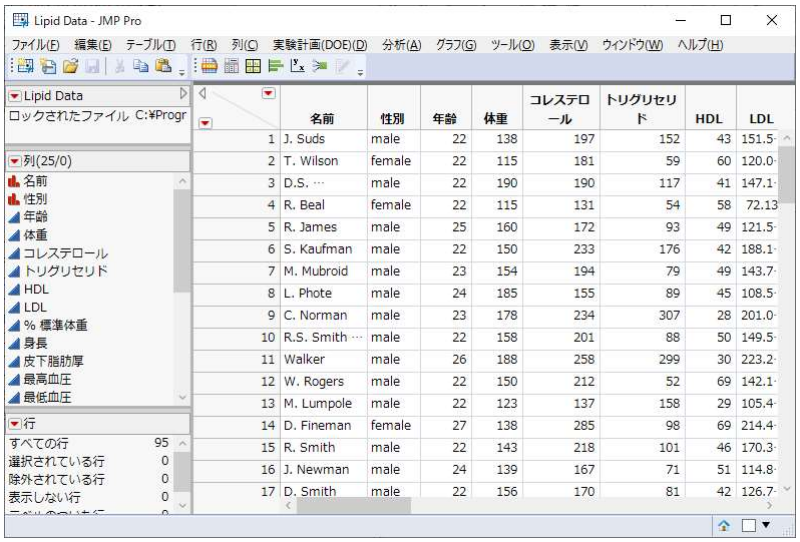
2

- ヘルプ＞サンプルデータ
- 医学研究のLipid data

2

データテーブル

3

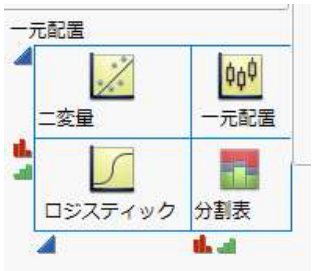


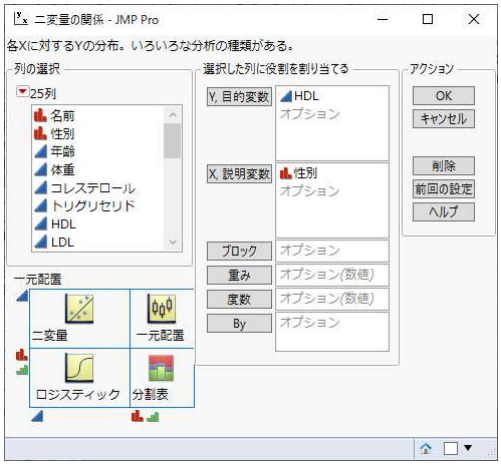
3

t検定、分散分析

4

- 分析> 二変量の関係
- YにHDL、Xに性別



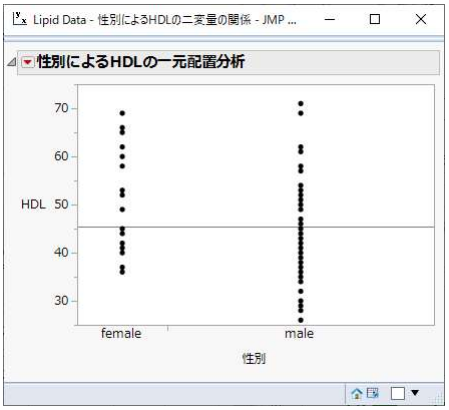


4

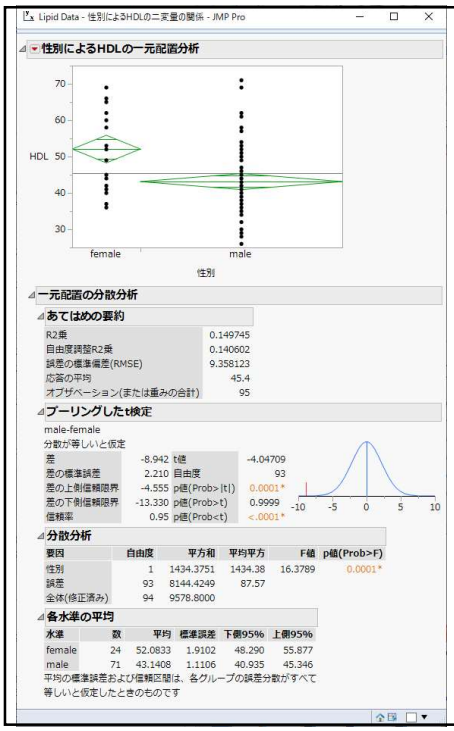
t検定

5

- ▼から
  「平均/ANOVA/プーリングしたt検定」



5



6

male-female
分散が等しいと仮定

差	-8.942	t値	-4.04709
差の標準誤差	2.210	自由度	93
差の上側信頼限界	-4.555	p値(Prob>|t|)	0.0001*
差の下側信頼限界	-13.330	p値(Prob>t)	0.9999
信頼率	0.95	p値(Prob<t)	<.0001*

6

Welchのt検定

7

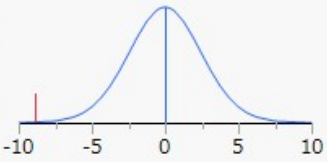
- ▼から「個々の分散を用いたt検定」

t検定

male-female

分散が等しくないと仮定

差	-8.942	t値	-3.60668
差の標準誤差	2.479	自由度	33.29477
差の上側信頼限界	-3.900	p値(Prob>|t|)	0.0010*
差の下側信頼限界	-13.985	p値(Prob>t)	0.9995
信頼率	0.95	p値(Prob<t)	0.0005*



7

Wilcoxonの順位和検定

8

- ノンパラメトリック＞Wilcoxon

Wilcoxon/Kruskal-Wallisの検定(順位和)

水準	度数	スコア和	スコアの期待値	スコア平均	(平均-平均0)/標準偏差0
female	24	1541.00	1152.00	64.2083	3.332
male	71	3019.00	3408.00	42.5211	-3.332

2標本検定(正規近似)

S	Z	p値(Prob>|Z|)
1541	3.33181	0.0009*

一元配置検定(カイ2乗近似)

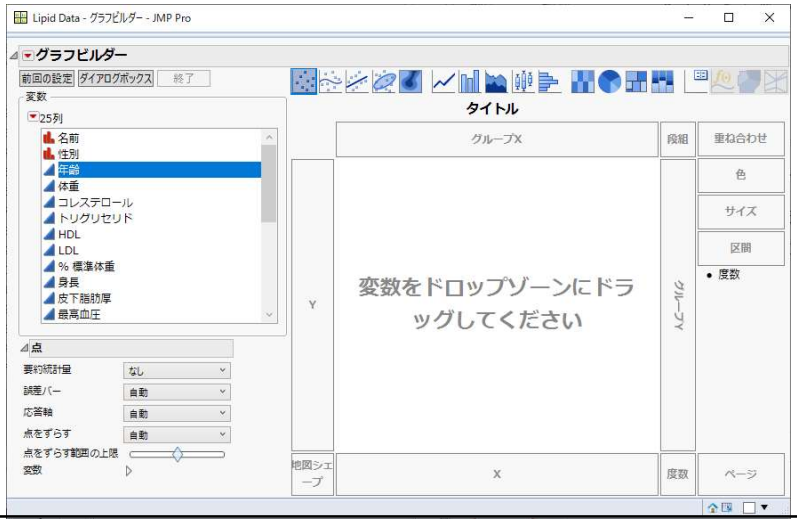
カイ2乗	自由度	p値(Prob>ChiSq)
11.1296	1	0.0008*

8

グラフ＞グラフビルダー

9

- 柔軟に様々なグラフを描くことが可能



9

練習①

10

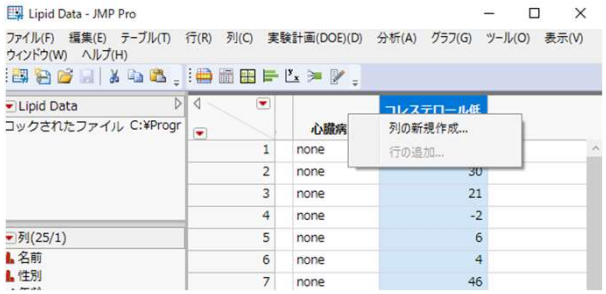
- コーヒー摂取量をアウトカムとして、
  喫煙歴による違いがみられるか
  分散分析を実行せよ
- 喫煙歴Noをコントロールとして、
  Dunnett法により対比較を実施せよ
- 喫煙歴グループごとに
  箱ひげ図を描いてみよ

10

対応のあるデータ

11

- データテーブルの何もない部分で
  - 右クリック> 列の新規作成　もしくは
  - ダブルクリック

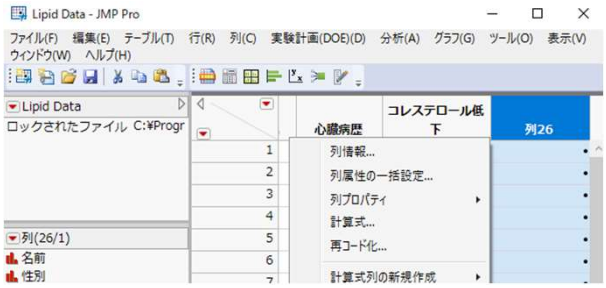


11

列情報の表示

12

- 新規作成された列(ここでは「列26」)を
  - 右クリック> 列情報　もしくは
  - ダブルクリック

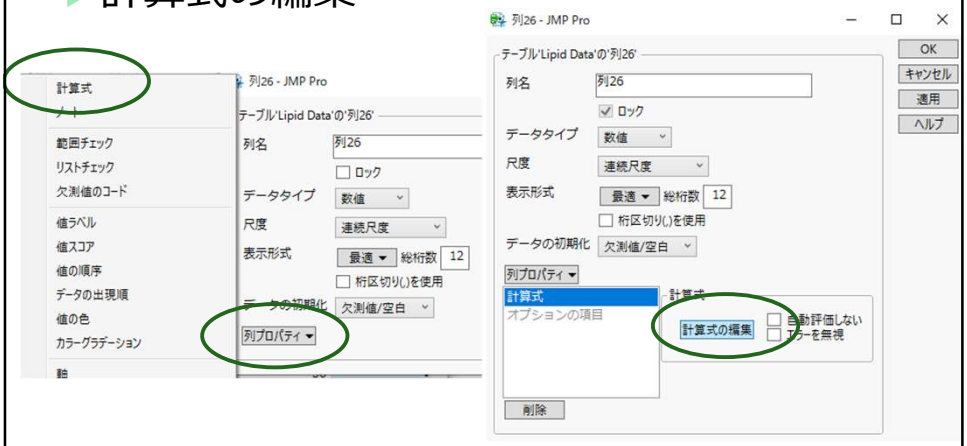


12

計算式による変化量の計算

13

- 列プロパティ＞計算式
- 計算式の編集

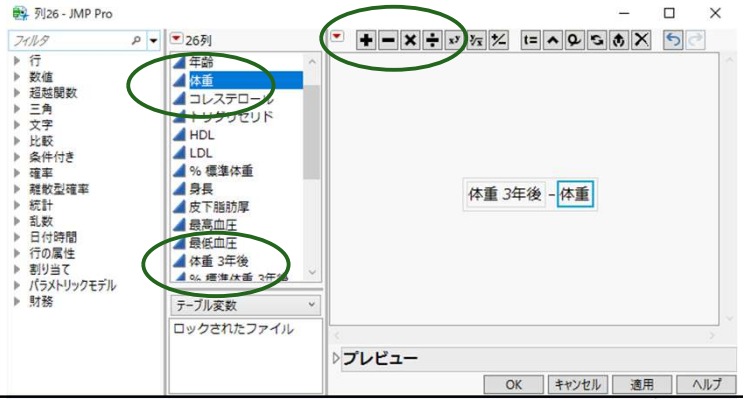


13

体重3年後 ー 体重を作成

14

- 体重 3年後をクリック
- ー　（＋－×÷の－）をクリック
- 体重をクリック

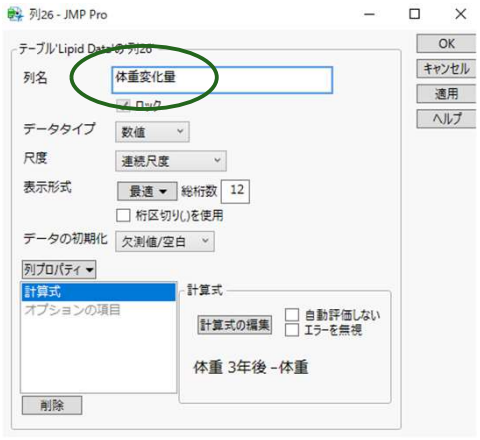


14

列名も変えてみる

15

- 列名に変数名を入力

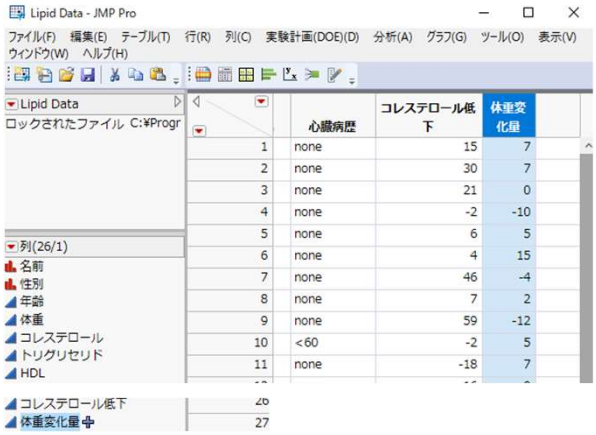


15

新たな列が追加

16

- 計算式を使用した場合、
  変数リストに＋が表示される



16

対応のあるデータの検定

17

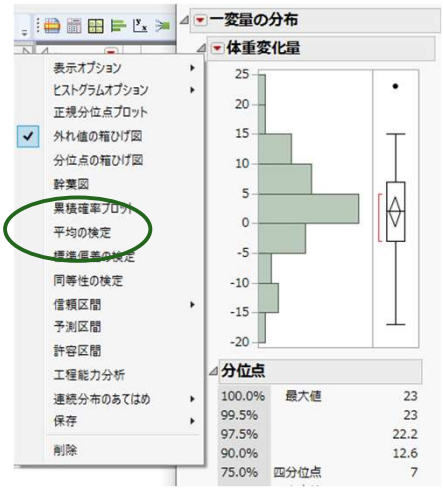
- 分析> 一変量の分布
- 作成した変化量データをYに

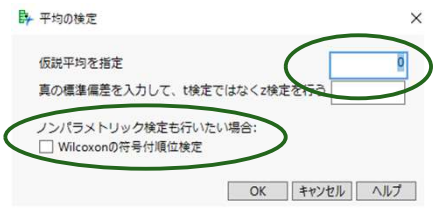
17

変数名の左にある▼

18

- 平均の検定をクリック





- 仮説平均
  （帰無仮説の値）
  を指定

18

対応のあるt検定と Wilcoxonの符号付順位検定

19

平均の検定

仮説値	0	
実際の推定値	1.90698	
自由度	42	
標準偏差	8.02619	
	t検定	Wilcoxonの符号付順位検定
検定統計量	1.5580	143.0000
p値(Prob>|t|)	0.1267	0.0837
p値(Prob>t)	0.0634	0.0419*
p値(Prob<t)	0.9366	0.9581



19

練習②

20

- コレステロール低下（既作成）について、対応のあるt検定を実行せよ

20