

# 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

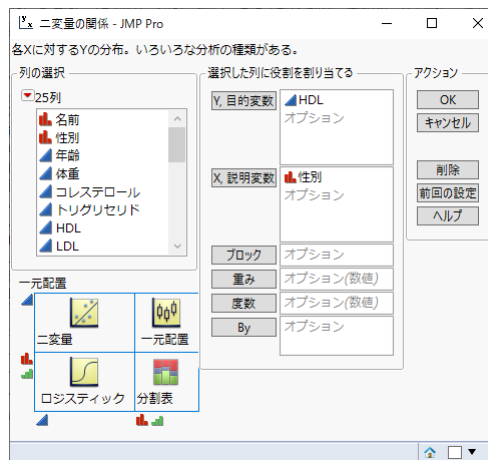
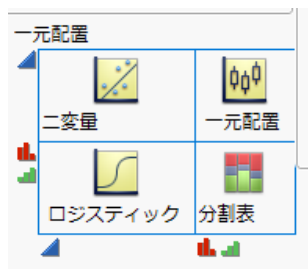
名前	性別	年齢	体重	コレステロール	トリグリセリド	HDL	LDL
1 J. Suds	male	22	138	197	152	43	151.5
2 T. Wilson	female	22	115	181	59	60	120.0
3 D.S. ...	male	22	190	190	117	41	147.1
4 R. Beal	female	22	115	131	54	58	72.13
5 R. James	male	25	160	172	93	49	121.5
6 S. Kaufman	male	22	150	233	176	42	188.1
7 M. Mubroid	male	23	154	194	79	49	143.7
8 L. Phote	male	24	185	155	89	45	108.5
9 C. Norman	male	23	178	234	307	28	201.0
10 R.S. Smith ...	male	22	158	201	88	50	149.5
11 Walker	male	26	188	258	299	30	223.2
12 W. Rogers	male	22	150	212	52	69	142.1
13 M. Lumpole	male	22	123	137	158	29	105.4
14 D. Fineman	female	27	138	285	98	69	214.4
15 R. Smith	male	22	143	218	101	46	170.3
16 J. Newman	male	24	139	167	71	51	114.8
17 D. Smith	male	22	156	170	81	42	126.7

3

# t検定、分散分析

4

- ▶ 分析 > 二変量の関係
- ▶ YにHDL、Xに性別

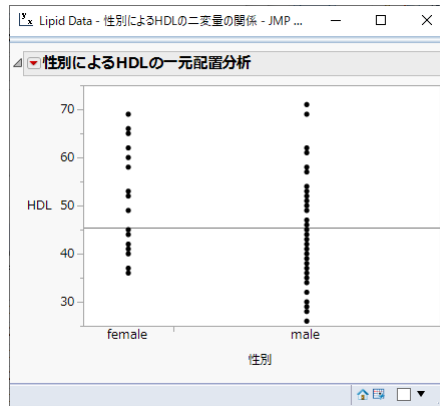


4

# t検定

5

▶ ▼ から  
「平均/ANOVA/プーリングしたt検定」



5

性別によるHDLの一元配置分析

一元配置の分散分析

あてはめの要約

R2乗	0.149745
自由度調整R2乗	0.140602
誤差の標準偏差(RMSE)	9.358123
応答の平均	45.4
オブザベーション(または重みの合計)	95

プーリングしたt検定

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*

分散分析

要因	自由度	平方和	平均平方	F値	p値(Prob>F)
性別	1	1434.3751	1434.38	16.3789	0.0001*
誤差	93	8144.4249	87.57		
全変(修正済み)	94	9578.8000			

各水準の平均

水準	数	平均	標準誤差	下側95%	上側95%
female	24	52.0833	1.9102	48.290	55.877
male	71	43.1408	1.1106	40.935	45.346

平均の標準誤差および信頼区間は、各グループの誤差分散がすべて等しいと仮定したときのものです

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」



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)

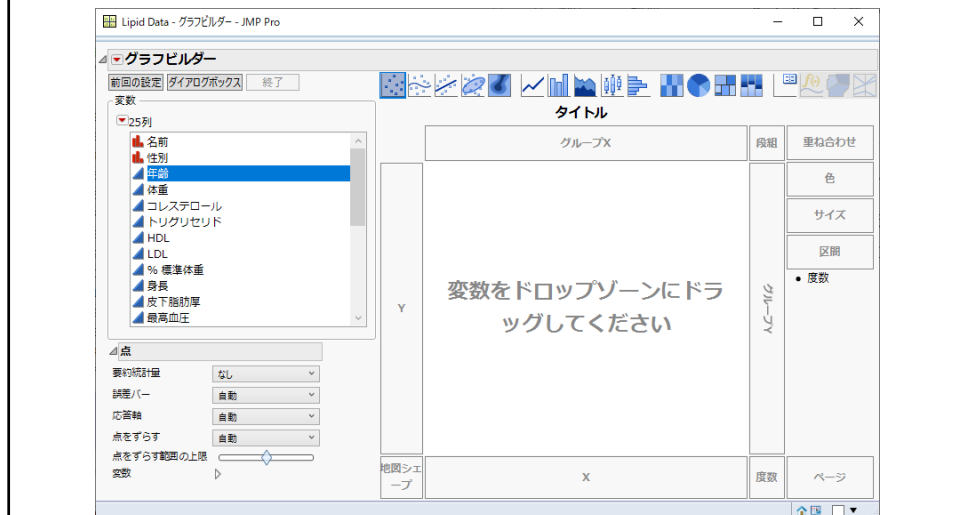
カイ2乗	自由度	p値(Prob>ChiSq)
11.1296	1	0.0008*

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

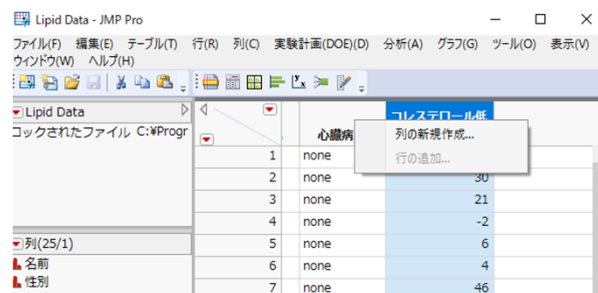
- ▶ コーヒー摂取量をアウトカムとして、喫煙歴による違いがみられるか分散分析を実行せよ
- ▶ 喫煙歴Noをコントロールとして、Dunnett法により対比較を実施せよ
- ▶ 喫煙歴グループごとに箱ひげ図を描いてみよ

10

## 対応のあるデータ

11

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

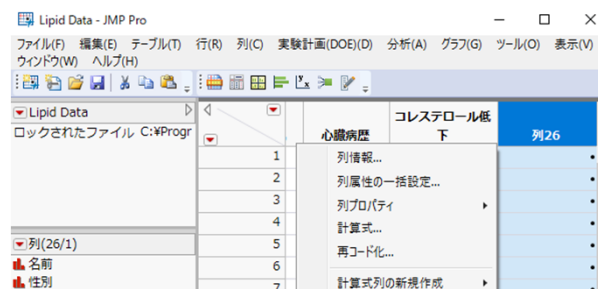


11

## 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

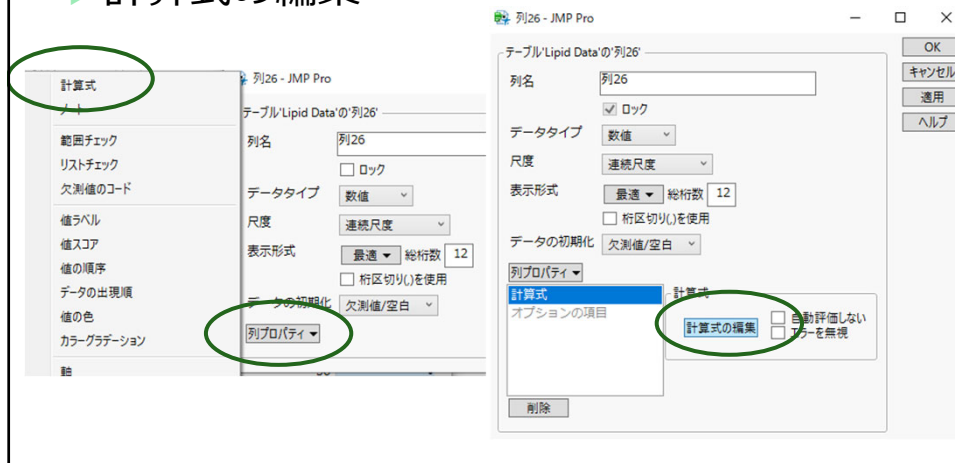


12

## 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

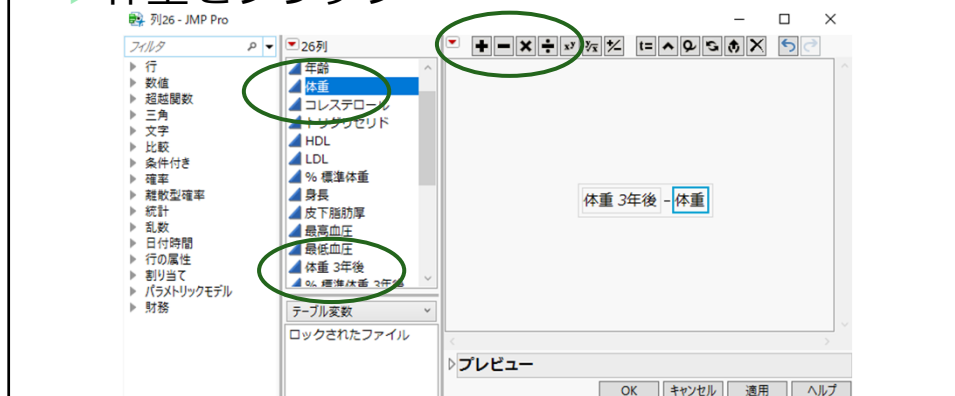


13

## 体重3年後 — 体重を作成

14

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

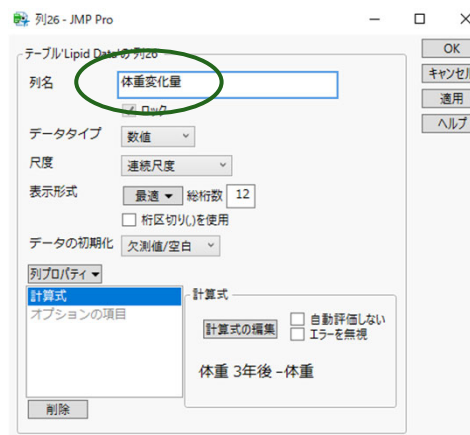


14

## 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

	心臓病歴	コレステロール低下	体重変化量
1	none	15	7
2	none	30	7
3	none	21	0
4	none	-2	-10
5	none	6	5
6	none	4	15
7	none	46	-4
8	none	7	2
9	none	59	-12
10	<60	-2	5
11	none	-18	7

16



## 対応のあるデータの検定

17

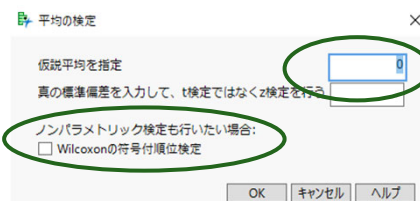
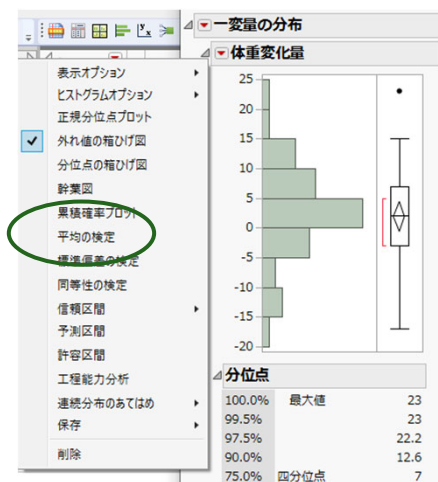
- ▶ 分析> 一変量の分布
- ▶ 作成した変化量データをYに

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



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