

2020/4/20 JMP実習①

データの要約と t検定

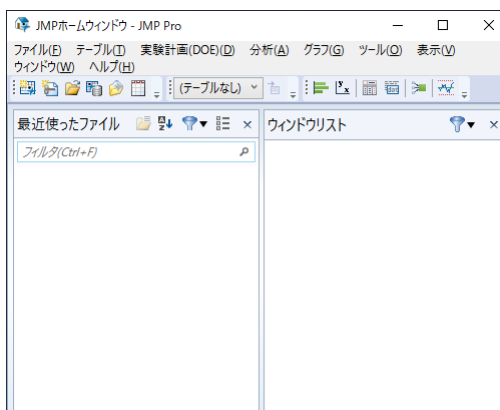


北海道大学 医学統計学
横田 勲

ホームウィンドウ

2

- ▶ ファイル> 新規作成> データテーブル
 - ▶ JMPに直接データを作成
- ▶ ファイル> 開く
 - ▶ CsvやExcelを読み込み



サンプルデータを利用

3

- ▶ ヘルプ> サンプルデータ
- ▶ 医学研究のLipid data

データテーブル




4

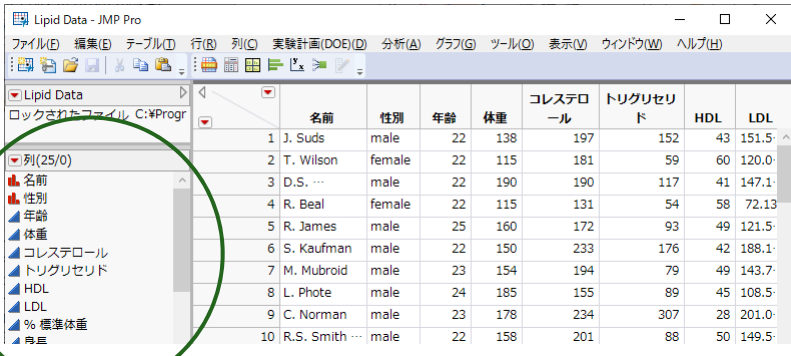
The screenshot shows the JMP Pro interface with a data table titled 'Lipid Data'. The table contains 17 rows of data with columns for Name, Sex, Age, Weight, Cholesterol, Triglycerides, HDL, and LDL. The left sidebar shows a tree view of the data table with various columns expanded.

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

尺度の種類

5

- ▶  連続：年齢、BMI、血圧
- ▶  順序：改善/不変/悪化のような順序性
- ▶  名義：あり/なし、術式A/B/C



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

列情報

6

- ▶ 変数を選んで右クリック
- ▶ データタイプと尺度で確認可
- ▶ データタイプが文字だと連続尺度を選べない



名前	性別	年齢	体重
J. Suds	male	22	138
T. Wilson	female	22	115
D.S. ...	male	22	190
R. Beal	female	22	115
R. James	male	25	160
S. Kaufman	male	22	150
M. Mubroid	male	23	154
L. Phote	male	24	185
C. Norman	male	23	178
R.S. Smith ...	male	22	158
Walker	male	26	188
W. Rogers	male	22	150
M. Lumpole	male	22	123
D. Fineman	female	27	138

名前 - JMP Pro

テーブル'Lipid Data'の'名前'

列名: 名前

ロック

データタイプ: 文字

尺度: 名義尺度

列プロパティ

変数の要約①

7

▶ 分析 > 一変量の関係

The screenshot shows the JMP Pro interface. The 'Analyze' menu is open, and the path 'Analyze > Univariate' is highlighted. The data table below shows the following columns: 年齢 (Age), 体重 (Weight), コレステロール (Cholesterol), トリグリセリド (Triglycerides), HDL, and LDL. The data rows are as follows:

年齢	体重	コレステロール	トリグリセリド	HDL	LDL			
22	138	197	152	43	151.5			
22	115	181	59	60	120.0			
22	190	190	117	41	147.1			
22	115	131	54	58	72.13			
25	160	172	93	49	121.5			
22	150	233	176	42	188.1			
23	154	194	79	49	143.7			
24	185	155	89	45	108.5			
23	178	234	307	28	201.0			
22	158	201	88	50	149.5			
26	188	258	299	30	223.2			
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

変数の要約②

8

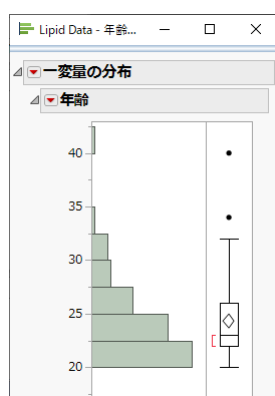
- ▶ 列の選択で変数をクリック
 - ▶ 複数指定可
- ▶ 「Y, 列」をクリック
- ▶ OKをクリック

The screenshot shows the 'Univariate' dialog box in JMP Pro. The 'Y, 列' (Y, Column) field is selected, and the '必須' (Required) option is chosen. The '重み' (Weight) field is set to 'オプション(数値)' (Option (Numeric)), and the '度数' (Degree) field is set to 'オプション(数値)' (Option (Numeric)). The 'By' field is set to 'オプション' (Option). The 'アクション' (Action) buttons include 'OK', 'キャンセル' (Cancel), '削除' (Delete), and '前回の設定' (Previous Settings). The 'ヘルプ' (Help) button is also visible.

連続変数の要約

9

- ▶ ヒストグラム
- ▶ 箱ひげ図
- ▶ 平均、標準偏差
- ▶ 分位点



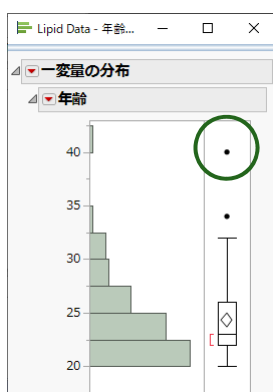
分位点		
100.0%	最大値	40
99.5%		40
97.5%		33.2
90.0%		29
75.0%	四分位点	26
50.0%	中央値	23
25.0%	四分位点	22
10.0%		22
2.5%		21.4
0.5%		20
0.0%	最小値	20

要約統計量	
平均	24.315789
標準偏差	3.2690136
平均の標準誤差	0.3353937
平均の上側95%	24.981722
平均の下側95%	23.649857
N	95

箱ひげ図で外れ値をクリック

10

- ▶ データテーブルの34行目が指定される
- ▶ どのデータであったか一目瞭然！

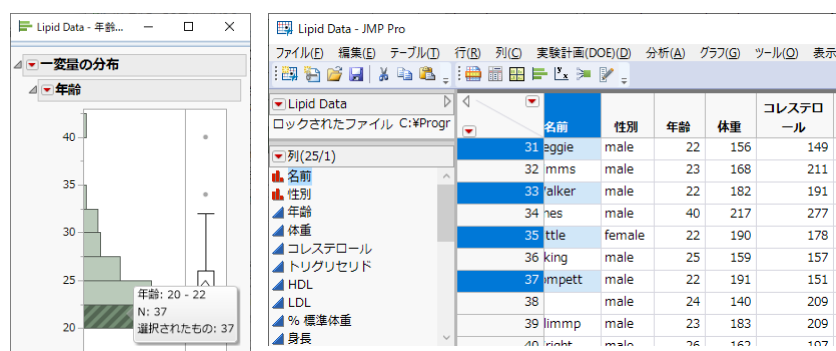


Lipid Data - JMP Pro			
ロックされたファイル C:\Progr			
	名前	性別	年齢
列(25/1)	31 Q. Reggie	male	22
	32 H. Timms	male	23
	33 H. Walker	male	22
	34 J. Jones	male	40
	35 M. Little	female	22
	36 G. Oking	male	25
	37 L. Lompett	male	22
	38 R. ...	male	24
	39 G. Glimmp	male	23

ヒストグラムをクリック

11

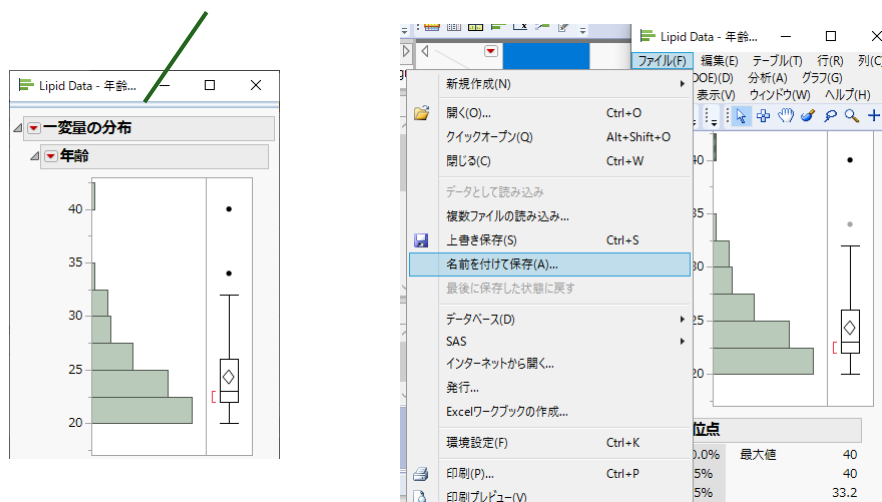
- ▶ 当該幅に含まれるデータの詳細が表示
- ▶ 当該幅に含まれるデータが選択



結果の保存

12

- ▶ この辺にマウスマウスカーソルをやる



他の要約統計量を出力

13

- ▶ 要約統計量の左にある▼をクリック
- ▶ 要約統計量のカスタマイズ

分位点		
100.0%	最大値	40
99.5%		40
97.5%		33.2
90.0%		29
75.0%	四分位点	26
50.0%	中央値	23
25.0%	四分位点	22
10.0%		22
2.5%		21.4
0.5%		20
0.0%	最小値	20

要約統計量	
平均	24.315789
標準偏差	3.2690136
平均の標準誤差	0.3353937
平均の上側95%	24.981722
平均の下側95%	23.649857
N	95

練習①

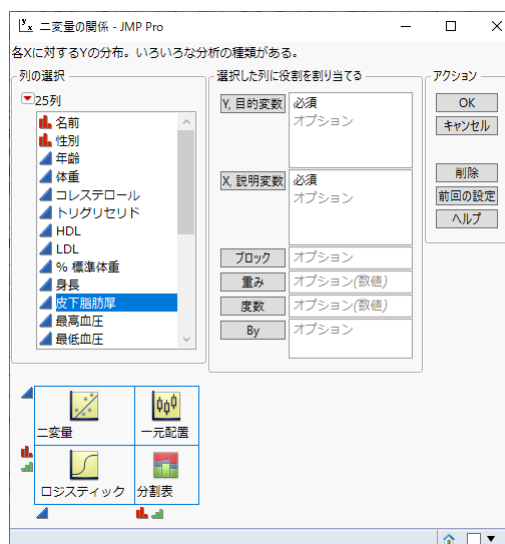
14

- ▶ LipidデータのHDLについて、
- ▶ 箱ひげ図を描け
- ▶ 中央値、四分位範囲、範囲、平均、標準偏差、歪度、尖度を確認せよ
- ▶ 正規分位点プロットを描け
 - ▶ 色々な▼をクリックして探してみよう

分析＞二変量の関係

15

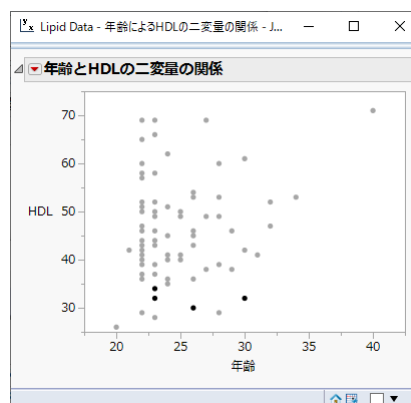
- ▶ Y, 目的変数
- ▶ X, 説明変数
にそれぞれ
変数を指定
- ▶ 例えば、
YにHDL
Xに年齢



まずは散布図

16

- ▶ ▼から様々な解析内容の表示を追加



練習②

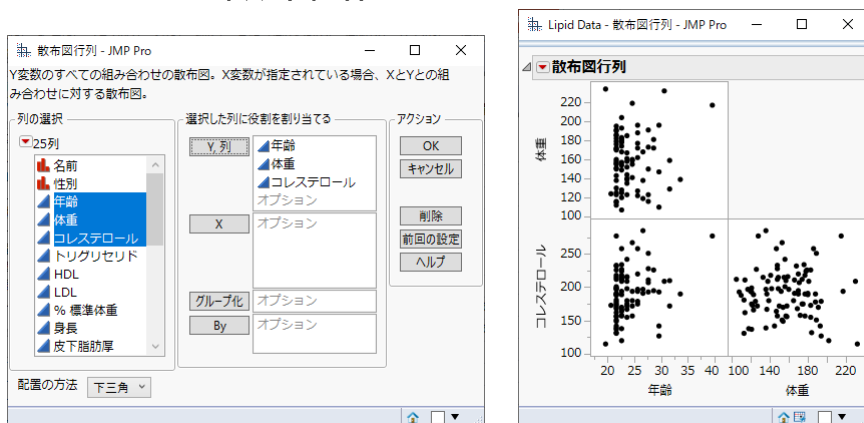
17

- ▶ HDLと年齢の関係について、
- ▶ 散布図を描き、95%確率楕円を重ねよ
- ▶ 相関係数を求めよ

(補) 散布図行列を描きたい

18

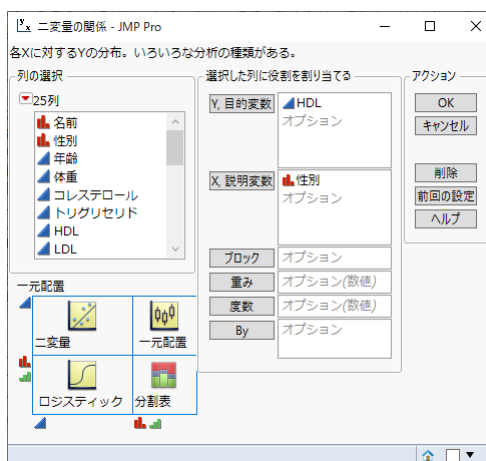
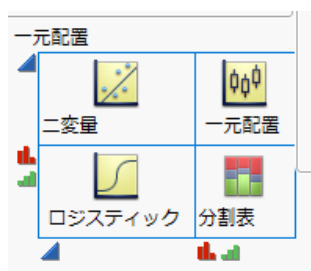
- ▶ 3つ以上の変数間の散布図
- ▶ グラフ> 散布図行列



t検定、分散分析

19

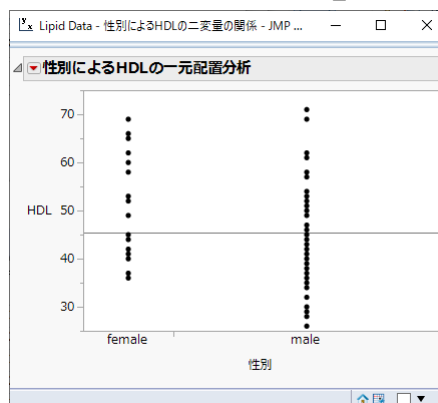
- ▶ 分析> 二変量の関係
- ▶ YにHDL、Xに性別

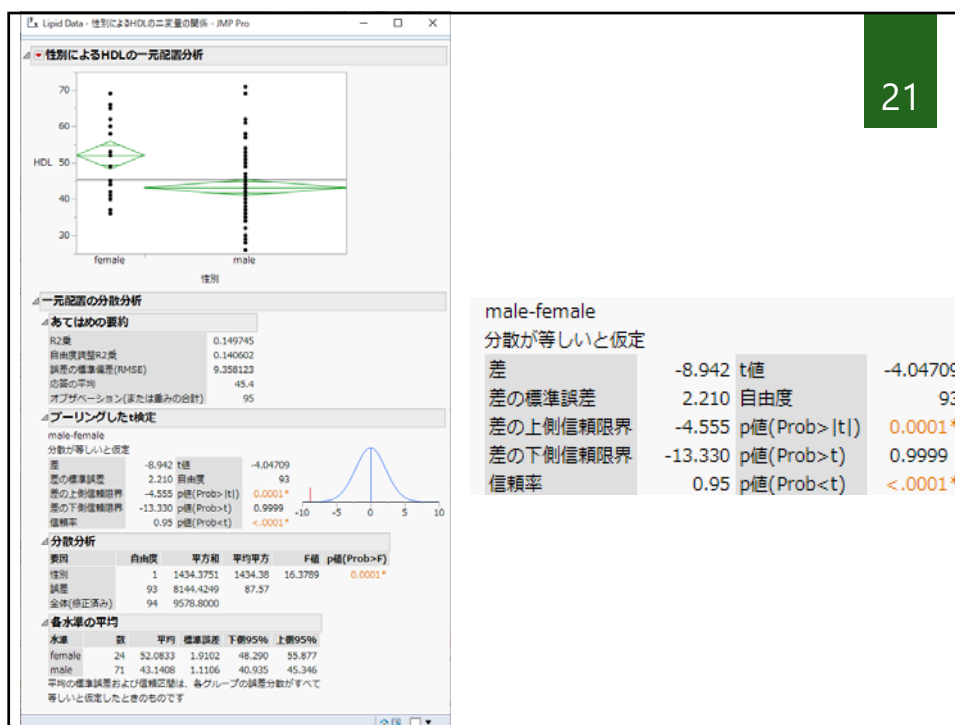


t検定

20

- ▶ ▼から
「平均/ANOVA/プーリングしたt検定」





21

Welchのt検定

22

▶ ▼ から「個々の分散を用いたt検定」



Wilcoxonの順位和検定

23

▶ 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*				

グラフ> グラフビルダー

24

▶ 柔軟に様々なグラフを描くことが可能

練習③

25

- ▶ 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