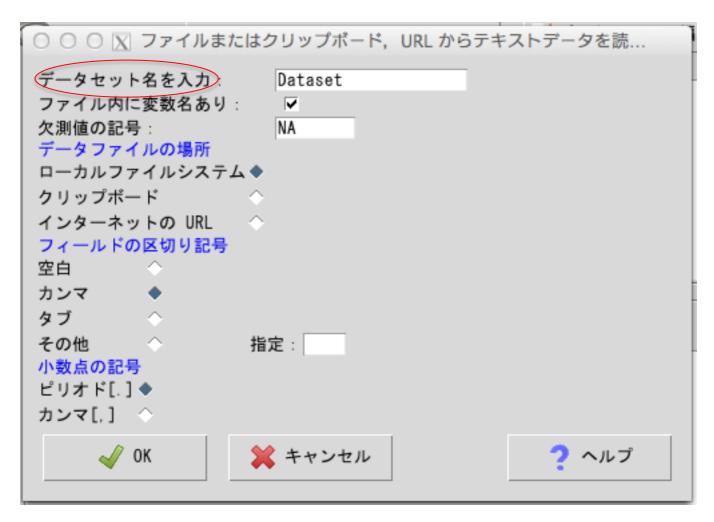
Rによる t検定・一元配置分散分析

- 1. 対応のある t 検定
 - 「データ」→「データのインポート」→「テキストファイルまたはクリップボード、URL から…」をクリック(SPSS などのセットがある場合は他の選択肢を選択)。





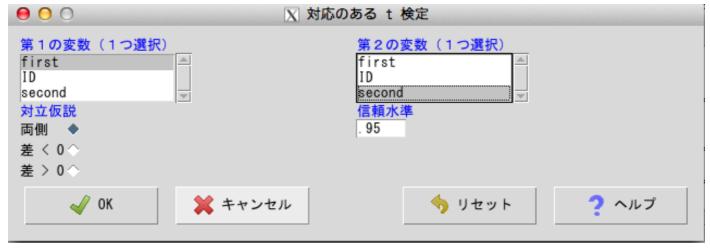
- 最上段の「データセット名を入力」の欄に適当な名前をつけ、
- ightharpoonup csv ファイルを使用する場合、中段の「フィールドの区切り記号」を「空白」ightharpoonup に変更する。ightharpoonup 「OK」をクリックする。
- ▶ xlsファイルを使用する場合は、「フィールドの区切り記号」を「タブ」に設定する。
- 対応するファイルを選ぶ画面になるので、使用するファイルを選択し、「Open」をクリックする。



■ 「データセットの編集」をクリックすると、上記のような画面が現れる。直接数値を変更 することは出来ないが、行や列の削除程度ならできる。

⊖ ⊝ ⊝			Rデータエディタ
Total Control			
ID first second			
1			
	18	30	
2	18	29	
3	17	30	
4	15	27	
5	14	28	
6	14	27	
7	13	24	
8	13	25	
9	13	26	
10	12	15	
11	12	17	
12	12	18	
13	12	16	
14	12	17	
15	12	18	
16	11	15	
17	11	14	
18	11	13	

■ 「統計量」 \rightarrow 「対応のある t 検定」を選択すると下記の画面になる。ここで、第一変数と第二変数をそれぞれ選択し、「OK」をクリックする。



すると…

```
### showData(Paired_t_test, placement='-20+200', font=getRcmdr('logFont'), maxwidth=80, maxheight=30)

> fix(Paired_t_test)

> showData(Paired_t_test, placement='-20+200', font=getRcmdr('logFont'), maxwidth=80, maxheight=30)

> t.test(Paired_t_test$first, Paired_t_test$second, alternative='two.sided', conf.level=.95, paired=TRUE)

Paired_t_test$first and Paired_t_test$second

t = -6.3087, df = 29, p-value = 6.853e-07

alternative hypothesis: true difference in means is not equal to 0

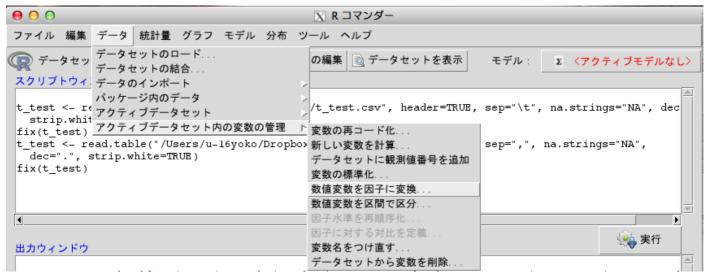
95 percent confidence interval:
-7.371322 -3.762011

sample estimates:
mean of the differences
-5.566667
```

「出力ウィンドウ」で結果が提示されるので、その内容と、先ほど R でスクリプトを打ってやった結果、または SPSS で出力した結果と比較してみよう。

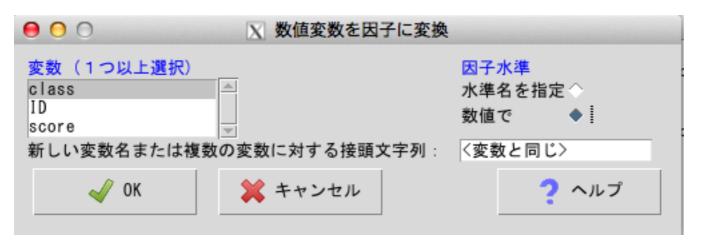
2. 対応のない t 検定

- 1.と同様にデータのインポートからデータを読み込む。
- 「データ」→「アクティブデータ内の変数の管理」→「数値変数を因子に変換」をクリックし、今回は「class」のデータを変換する。

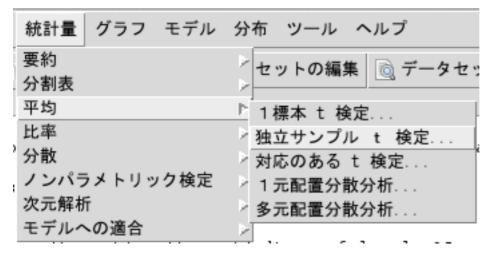


■ 変数の中から、因子にしたいもの(今回の場合は class)を一つ選択し、右側の「因子水準」 の部分「数値で」にチェックを入れる。

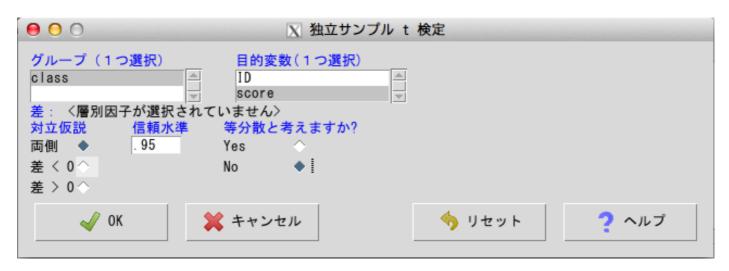
(水準名を指定しても良いのだが、手間がかかるため、このまま数値で行うのが無難)



- 『変数 class がすでに存在します。変数に上書きしますか?』と警告がでるので、「はい」を 選択し、次へ進む。
- 「統計量」→「平均」→「独立サンプル t 検定」を選択し、変数情報を入力する。

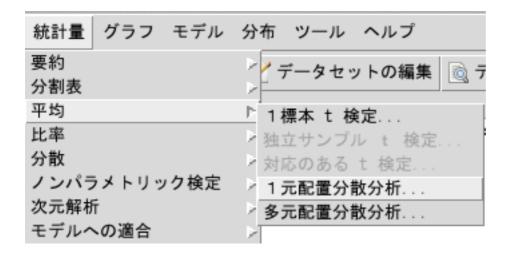


■ グループは class, 目的変数(説明変数)を score とし、対立仮説、信頼水準、等分散の仮定、の項目を確認し、「OK」をクリックする。



■ すると、今回の分析の結果が現れる。

- 3. 一元配置分散分析(対応なし)の場合
- データの入力方法は t 検定の時と同じなので割愛。
- 数値変数もt検定と同様に変換する。
- 「統計量」→「平均」→「1元配置分散分析」を選択し、入力画面で情報を確認する。



■ モデル名は自動で入力されたものをそのまま使用すればよい。グループ変数、目的変数が正しく選択されているか確認し、多重比較を行うようチェックを入れる。



■ これで結果が見られるようになる。

```
> summary(AnovaModel.1)
            Df Sum Sq Mean Sq F value Pr(>F)
                                  7.7 0.000834 ***
                  951 475.5
class
Residuals
            87
                 5372
                          61.7
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> numSummary(one_way_anova$score , groups=one_way_anova$class, statistics=c("mean",
"sd"))
                  sd data:n
      mean
1 33.76667 6.574157
                        30
2 27.46667 8.544945
                        30
3 26.40000 8.307039
                        30
> .Pairs <- glht(AnovaModel.1, linfct = mcp(class = "Tukey"))
> summary(.Pairs) # pairwise tests
         Simultaneous Tests for General Linear Hypotheses
Multiple Comparisons of Means: Tukey Contrasts
Fit: aov(formula = score ~ class, data = one_way_anova)
Linear Hypotheses:
           Estimate Std. Error t value Pr(>|t|)
2 - 1 == 0
                        2.029 -3.105 0.00717 **
          -6.300
                        2.029 -3.631 0.00139 **
3 - 1 == 0 -7.367
3 - 2 == 0 -1.067
                        2.029 -0.526 0.85897
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- single-step method)
```

> confint(.Pairs) # confidence intervals

Simultaneous Confidence Intervals

Multiple Comparisons of Means: Tukey Contrasts

Fit: aov(formula = score ~ class, data = one_way_anova)

Quantile = 2.3845

95% family-wise confidence level

Linear Hypotheses:

> remove(.Pairs)

```
Estimate lwr upr

2 - 1 == 0  -6.3000 -11.1380  -1.4620

3 - 1 == 0  -7.3667 -12.2047  -2.5287

3 - 2 == 0  -1.0667  -5.9047  3.7713

> cld(.Pairs) # compact letter display
2   3   1

"a" "a" "b"

> old.oma <- par(oma=c(0,5,0,0))

> plot(confint(.Pairs))

> par(old.oma)
```