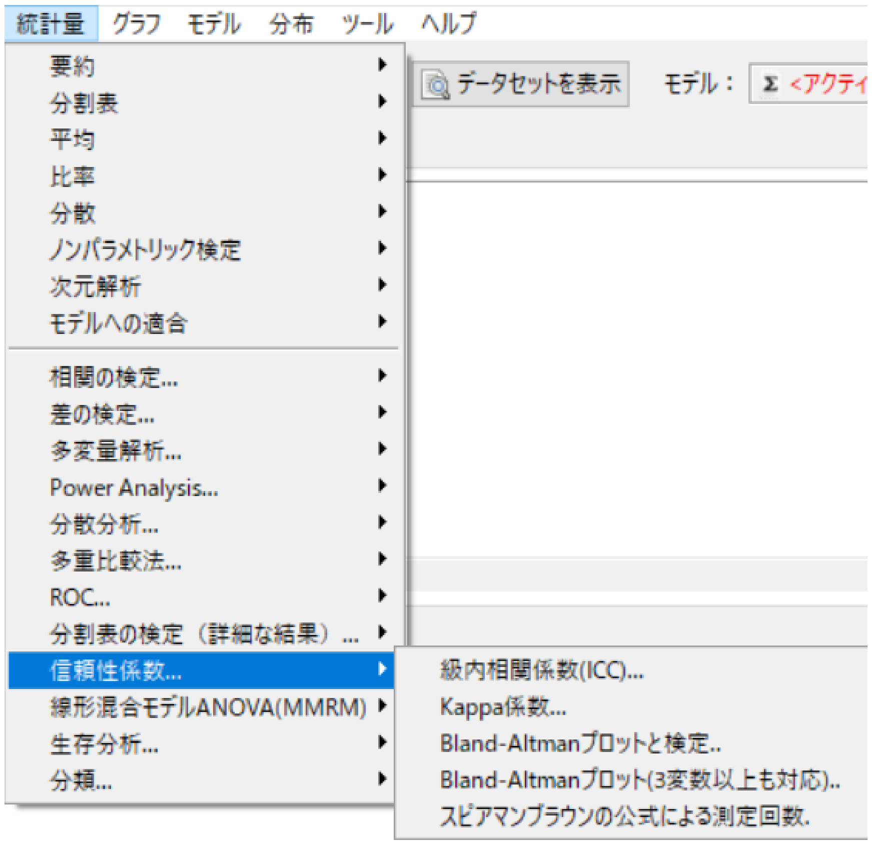
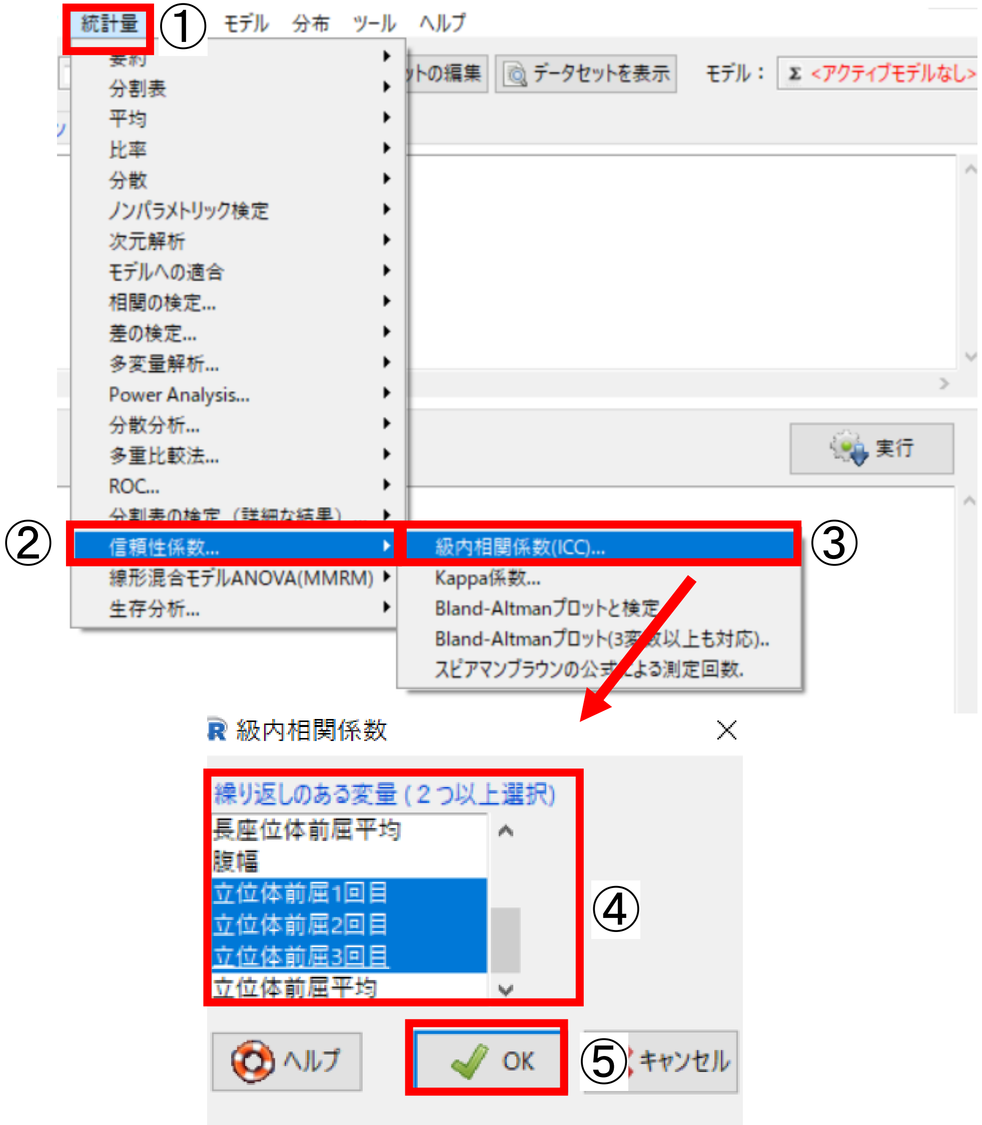
統計量：信頼性係数



- ■ 級内相関係数（ICC）
  - ✓ 正規分布に従うデータ
  - ✓ 2回以上の測定の再現性
  - ✓ 2名以上の測定者の再現性
- ■ Kappa係数
  - ✓ 段階的に評価されたデータ
  - ✓ 2回以上の測定の再現性
  - ✓ 2名以上の測定者の再現性
- ■ Bland-Altmanプロットと検定
- ■ Bland-Altmanプロット(3変数以上も対応)
  - ✓ 正当なBland-Altmanプロットではありません
- ■ スピアマンブラウンの公式による測定回数

作製担当：松田　陽子，石田　水里

統計量：信頼性係数
級内相関係数（ICC）



- 検者内または検者間信頼性（再現性）を調べます
- 比率尺度か間隔尺度，段階数の多い順序尺度のデータに適用します
- 級内相関係数の3つのタイプを出力します
  - ✓ Shrout PE, Fleiss JL. : Intraclass correlations: uses in assessing rater reliability. Psychol Bull. 1979 Mar;86(2):420-8.
  - ✓ https://www. aliquote. org/cours/2012_biomed/biblio/Shrout1979. pdf#search=%27shrout+fleiss%27 の論文に基づいて作製者がプログラムしています
  - ✓ SEMは Stratford PW,et al. :Use of the Standard Error as a Reliability Index of Interest: An Applied Example Using Elbow Flexor Strength Data. Physical Therapy 77(7):745-750,1997を参照に作製者がプログラムしています
- 立位体前屈データを選びます
- ①[統計量]－②[信頼性係数]－③[級内相関係数（ICC）]を選びます
- ダイアログボックスの④[繰り返しのある変量（2つ以上選択）]から，変数をクリックして選びます
  - ✓ 再現性を求めたい変数を選びます
- ⑤[OK]をクリックします

統計量：信頼性係数
級内相関係数（ICC） 結果

```
> ICC.CI(.Responses,95)
+ #ICC：3タイプ同時出力されています．（1，1）は検者内，（2，1）（3，1）は検者間信頼性です．
              estimate lower bound-95% upper bound-95%
ICC(1,1)     0.9735976       0.9548406       0.9855082
ICC(1, 3 )   0.9910415       0.9844796       0.9951223
ICC(2,1)     0.9736947       0.9184975       0.9890436
ICC(2, 3 )   0.9910751       0.9712715       0.9963210
ICC(3,1)     0.9845578       0.9733552       0.9915687
ICC(3, 3 )   0.9947991       0.9909578       0.9971737
```

①

②

```
> sem.2(.Responses)
+ #SEM：他のICC値と比較するときに使用します．SEMに差がないときにICCどうしを比較可能です．
       SEM lower bound-95% upper bound-95% lower bound-99% upper bound-99%
 1.1598440       0.9954829       1.3897209       0.9505691       1.4751261
```

③

```
> SB(.Responses)
+ #ICCの推定値に対してスピアマンブラウンの公式を用いて計算しています．厳しくするなら下限値を用いた方が良いかもしれません
$検者内信頼性ρ以上を満たすための測定回数

1 "ρ ≧ 0.7 を満たすためには 1 回以上測定した平均値を用います"
2 "ρ ≧ 0.75 を満たすためには 1 回以上測定した平均値を用います"
3 "ρ ≧ 0.8 を満たすためには 1 回以上測定した平均値を用います"
4 "ρ ≧ 0.85 を満たすためには 1 回以上測定した平均値を用います"
5 "ρ ≧ 0.9 を満たすためには 1 回以上測定した平均値を用います"
6 "ρ ≧ 0.95 を満たすためには 1 回以上測定した平均値を用います"

$検者間信頼性ρ以上を満たすための測定回数

1 "ρ ≧ 0.7 を満たすためには 1 回以上測定した平均値を用います"
2 "ρ ≧ 0.75 を満たすためには 1 回以上測定した平均値を用います"
3 "ρ ≧ 0.8 を満たすためには 1 回以上測定した平均値を用います"
4 "ρ ≧ 0.85 を満たすためには 1 回以上測定した平均値を用います"
5 "ρ ≧ 0.9 を満たすためには 1 回以上測定した平均値を用います"
6 "ρ ≧ 0.95 を満たすためには 1 回以上測定した平均値を用います"
```

- ①ICCは，3タイプ6種類出力されます
- 0.7以上が望ましいといわれます
  - ✓ 検者内信頼性を見るときはICC(1,1)，検者間信頼性をみるときはICC(2,1)を見ます
  - ✓ ICC(3,1)は，滅多に使いません
  - ✓ ICC（1,k），(2,k）は最初は使いません
  - ✓ ｋ回測定の平均値を使ったら，ICC（1,k），(2,k）の値になります，という意味となります
    - ➢ k=1,2,3,...とデータ列に応じて変わります
- 他のICCと大小を比較したいときは，②SEMに差がないことが条件です
  - ✓ 2つのICCの95%信頼区間が重複している（＝差があるとはいえない）ときに比較できます
  - ✓ この比較方法は理論的に正しいとはいえません
- 測定回数は③に出力されます．スピアマンブラウンの公式を利用しています
  - ✓ ρ≧0.7を満たす回数が理想です
  - ✓ 厳しくするなら下限値を用いた方がよいかもしれません

統計量：信頼性係数
Kappa係数

- カッパ係数を出力します
- 名義尺度または順序尺度のデータに適用します
- 3名以上の一致度についても計算可能です
  - ✓ Concordパッケージを使用しています
- X線頸椎アライメント分類データを選びます
- 数値変数は選択できませんので，必要に応じてあらかじめ因子変数への変換を行ってください
  - ✓ ここではすべて因子変数へ変換しています
- ①[統計量]ー②[信頼性係数]ー③[Kappa係数]を選びます
- ダイアログボックスの④[繰り返しのある変量（2つ以上選択）]から，変数をクリックして選びます
  - ✓ 再現性を求めたい変数を選びます
  - ✓ アライメント分類1回目とアライメント分類2回目とアライメント分類3回目を選んでいます
- ⑤[OK]をクリックします

統計量:信頼性係数
Kappa係数　結果

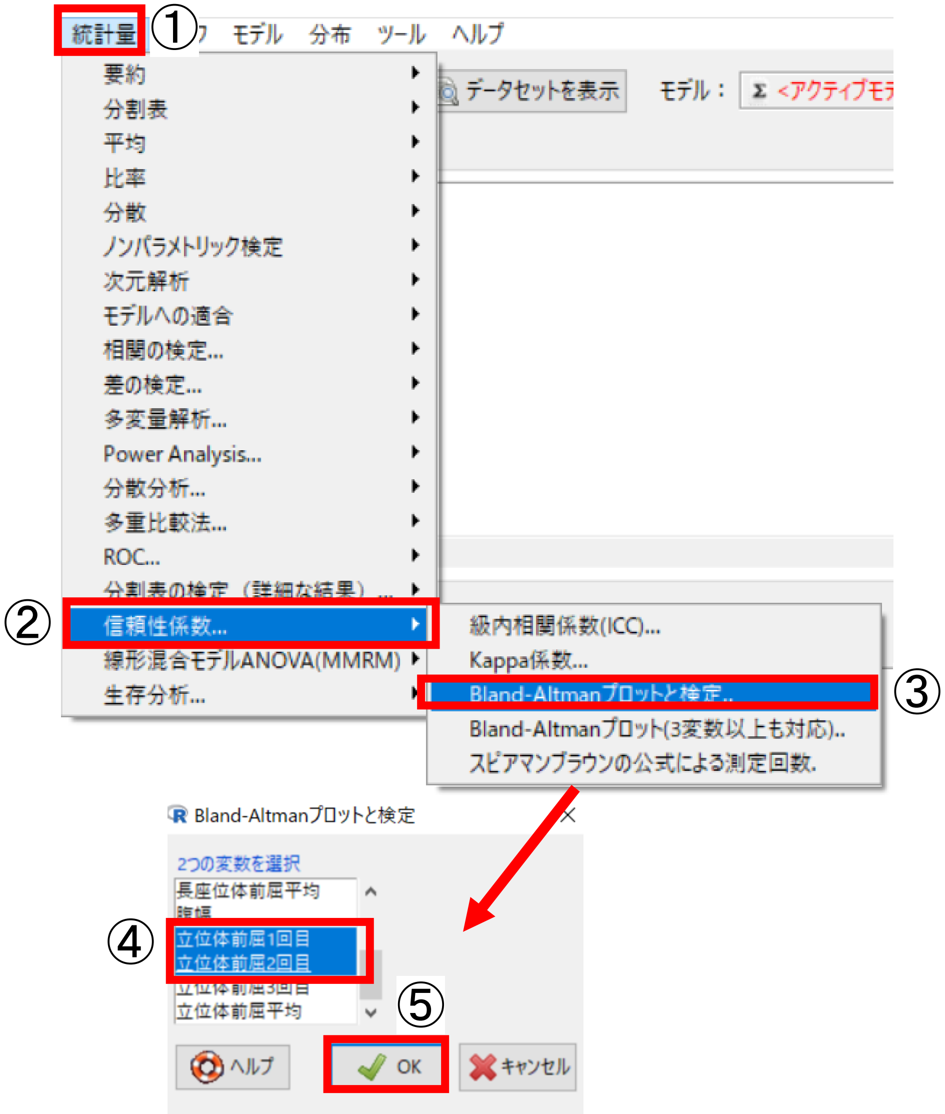
```
> cohen.kappa(.Responses)
+ #一般には，SiegelのKappa係数（=Fleissのκ係数）を見ます
                   coef.        Z p
Kappa(Cohen)  0.9265381 12.97869 0
Kappa(Siegel) 0.9061033 10.55970 0

> SBK(.Responses)
$`検者内（間）信頼性ρ以上を満たすための測定回数`

1   ρ≧ 0.7 を満たすためには 1 回以上測定した平均値を用います
2  ρ≧ 0.75 を満たすためには 1 回以上測定した平均値を用います
3   ρ≧ 0.8 を満たすためには 1 回以上測定した平均値を用います
4  ρ≧ 0.85 を満たすためには 1 回以上測定した平均値を用います
5   ρ≧ 0.9 を満たすためには 1 回以上測定した平均値を用います
6  ρ≧ 0.95 を満たすためには 2 回以上測定した平均値を用います
```

①

②

- カッパ係数は①を見ます
- ICCと同様に0.7以上が理想です
  - ✓ これは，2人の検者，2回の繰り返し一致度を求める本来のカッパ係数とは異なり，それを拡張したSiegelのカッパ係数（Fleissの係数）となっています
  - ✓ Siegel,S.　and Castellan,N.　J.　Jr.　：Nonparametric statistics for the behavioral sciences (2nd ed.　). 　McGraw-Hill,　1988
- 測定回数は②に出力されます．スピアマンブラウンの公式を利用しています
  - ✓ ρ≧0.7を満たす回数が理想です
  - ✓ この例では，1回以上測定した平均値を使用すれば0.7以上，さらには0.9以上までも満たせそうです

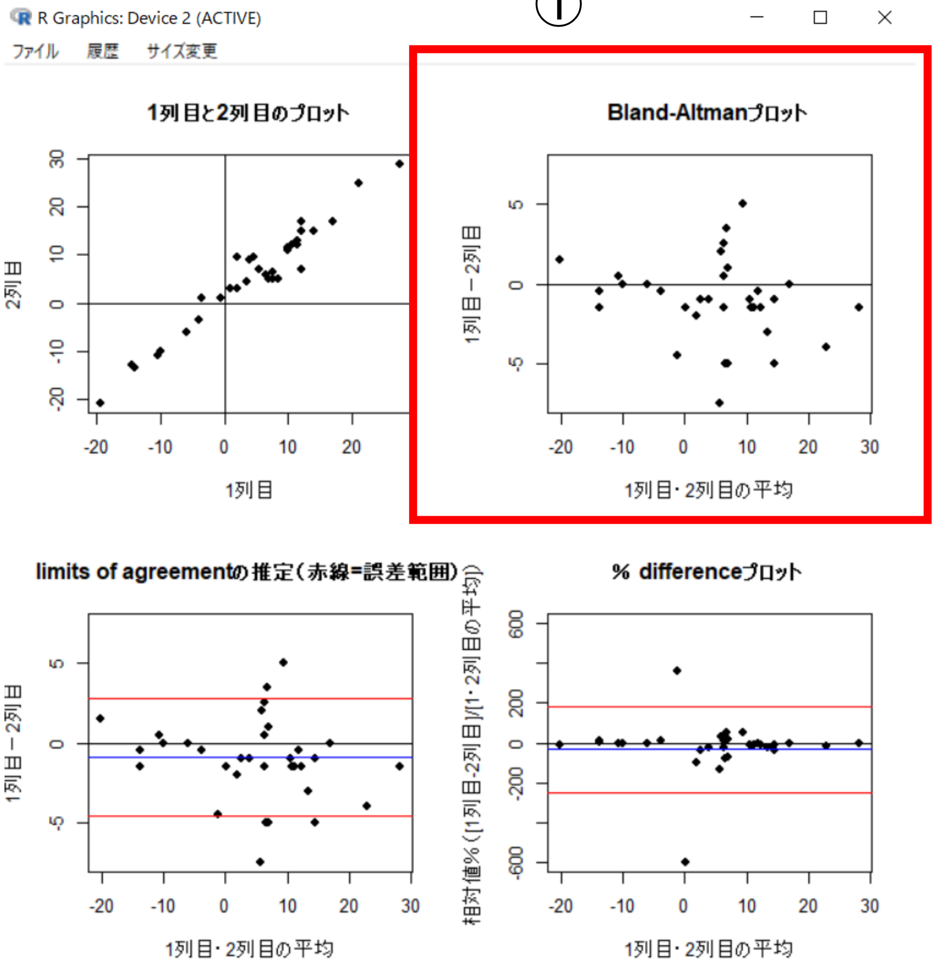
統計量：信頼性係数
Bland-Altmanプロットと検定



- Bland-Altman分析は測定値に系統誤差が混入しているかを確認する手法の1つです
- 一対2つの測定値の差をy軸，2つの測定値の平均値をx軸にプロットした散布図（Bland-Altman plot）を作成して，測定値に含まれる系統誤差の有無・程度を可視的に確認します
- このほか，固定誤差や比例誤差に関する検定と誤差の程度も出力します
- 立位体前屈データを選びます
- ①[統計量]ー②[信頼性係数]ー③[Bland-Altmanプロットと検定]を選びます
- ダイアログボックスの④[2つの変数を選択]から，変数をクリックして選びます
  - ✓ 立位体前屈1回目と立位体前屈2回目を選んでいます
- ⑤[OK]をクリックします

統計量：信頼性係数
Bland-Altmanプロットと検定　結果

①



- ①のグラフがBland-Altman plotです．2つの測定値について，x軸は平均，y軸は差の値を表しています．つまり，被検者ごとに求めた平均の全被検者の中でのバラツキがx軸に，各測定値間のバラツキがy軸に反映されます
- 系統誤差が存在せず，偶然誤差のみが存在する場合，y軸の0付近を中心に正負の方向にばらついた分布を示します．被検者によって正負の様々な方向にばらつくという意味です
- 系統誤差のうち固定誤差（真の値に関わらず，特定方向に一定の幅で生じる乖離）が存在する場合，y軸の0付近から正あるいは負の特定方向に偏った分布を示します
- 比例誤差（真の値に比例して増減する，特定方向に生じる乖離）が存在する場合，x軸の値が増加するにつれて，2つの測定値の差の増加がみられる扇型の分布を示します

✓ 今回の立位体前屈データでは，y軸0から負の方向に偏った分布を示しており，系統誤差のうち，固定誤差が存在しています

統計量：信頼性係数
Bland-Altmanプロットと検定　結果

出力　実行

①

```
> Bland.Altman(.Responses)
$`【固定誤差 fixed error】1列目と2列目データの差の検定（有意→差の平均が0ではない固定誤差あり）`

        One Sample t-test

data:  diff
t = -3.9849, df = 35, p-value = 0.0003261
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -1.7610214 -0.5723119
sample estimates:
mean of x
-1.166667


$`差の値の標準誤差(SE)`
[1] 0.29277

$`固定誤差の程度：limits of agreementの推定（下限，平均，上限の順）→同タイトルの図で判断`
[1] -3.580190 -1.166667  1.246856
```

②

```
$`【比例誤差 proportional error】plotの相関係数
       cor
-0.1948034

$`相関係数の検定p値（有意な時は比例誤差あり）`
[1] 0.2549036

$`plotの回帰分析（Pr下段が有意な時は比例誤差あり）`
              Estimate Std. Error   t value   Pr(>|t|)
(Intercept) -0.9353786 0.35323369 -2.648045 0.01218557
average     -0.0370061 0.03195482 -1.158076 0.25490357

$`比例誤差の程度：%differenceの推定（下限，平均，上限）→同タイトルの図で判断`
[1] -112.48892  -22.16654   68.15584
```

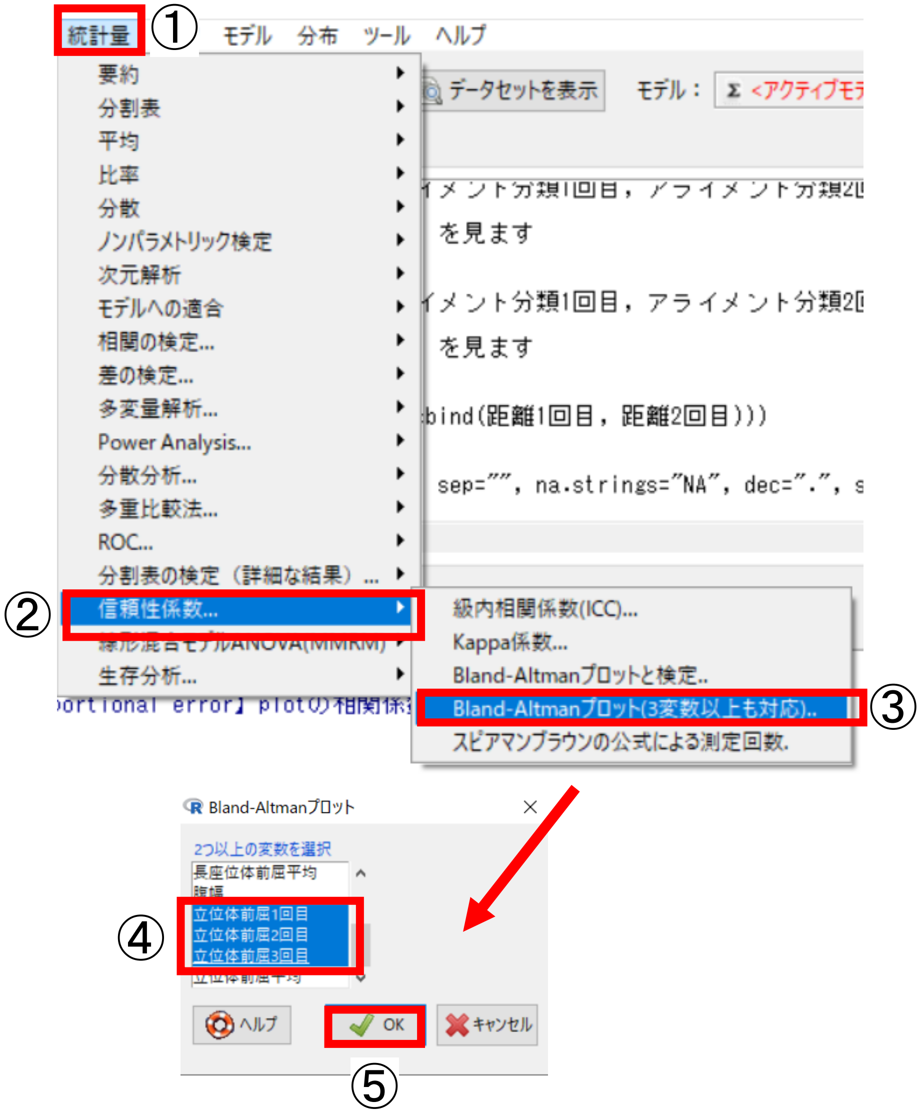
③

```
$`偶然誤差のみのとき（上記検定で全て有意ではないとき）：最小可検誤差(minimal detectable change)
[1] 3.442975
```

- 出力には固定誤差や比例誤差に関する検定と誤差の程度が表されます
- ①は固定誤差に関する結果です．下線部は2つの測定値間の差の検定結果で，p＜0.05で有意のとき固定誤差ありと判断します．固定誤差の程度（下限，平均，上限）も4つのグラフのうち左下「limits of agreementの推定」タイトルのグラフと照らし合わせて確認できます．グラフ内赤線が下限と上限，青線が平均を表しています
- ②は比例誤差に関する結果です．下線部（相関係数の検定p値およびplotの回帰分析のpr下段）がp＜0.05で有意なときに比例誤差ありと判断します．比例誤差の程度も右下「%differenceプロット」タイトルのグラフに対応しており，赤線が下限と上限，青線が平均を表します
- 固定誤差と比例誤差の検定ともに有意でなく，偶然誤差のみが存在する場合，最小可検誤差（MDC）の値を確認できます．MDCは偶然誤差のうち測定誤差範囲の指標となります

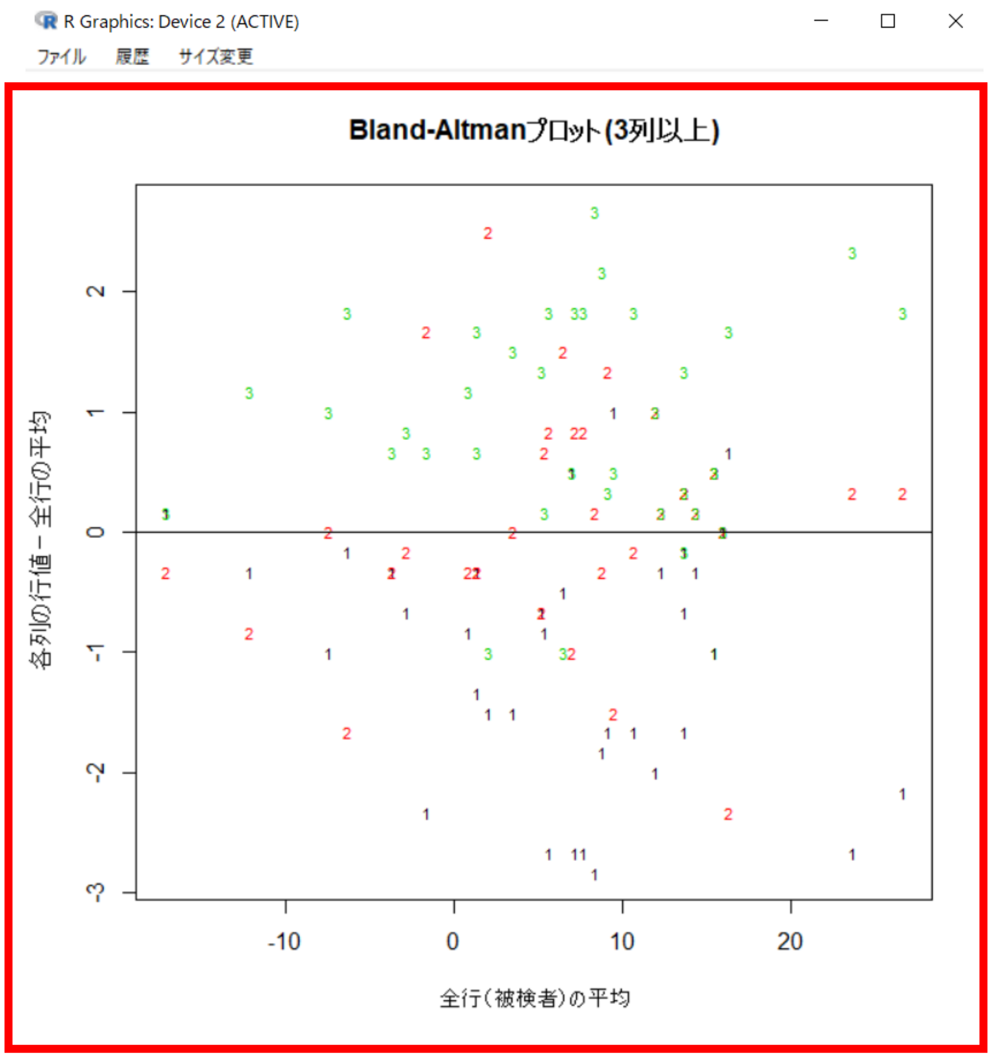
✓ ここでの結果は固定誤差の検定結果で有意となり，Bland-Altman plotの可視的判断とも一致しています

統計量：信頼性係数
Bland-Altmanプロットと検定（3変数以上も対応）



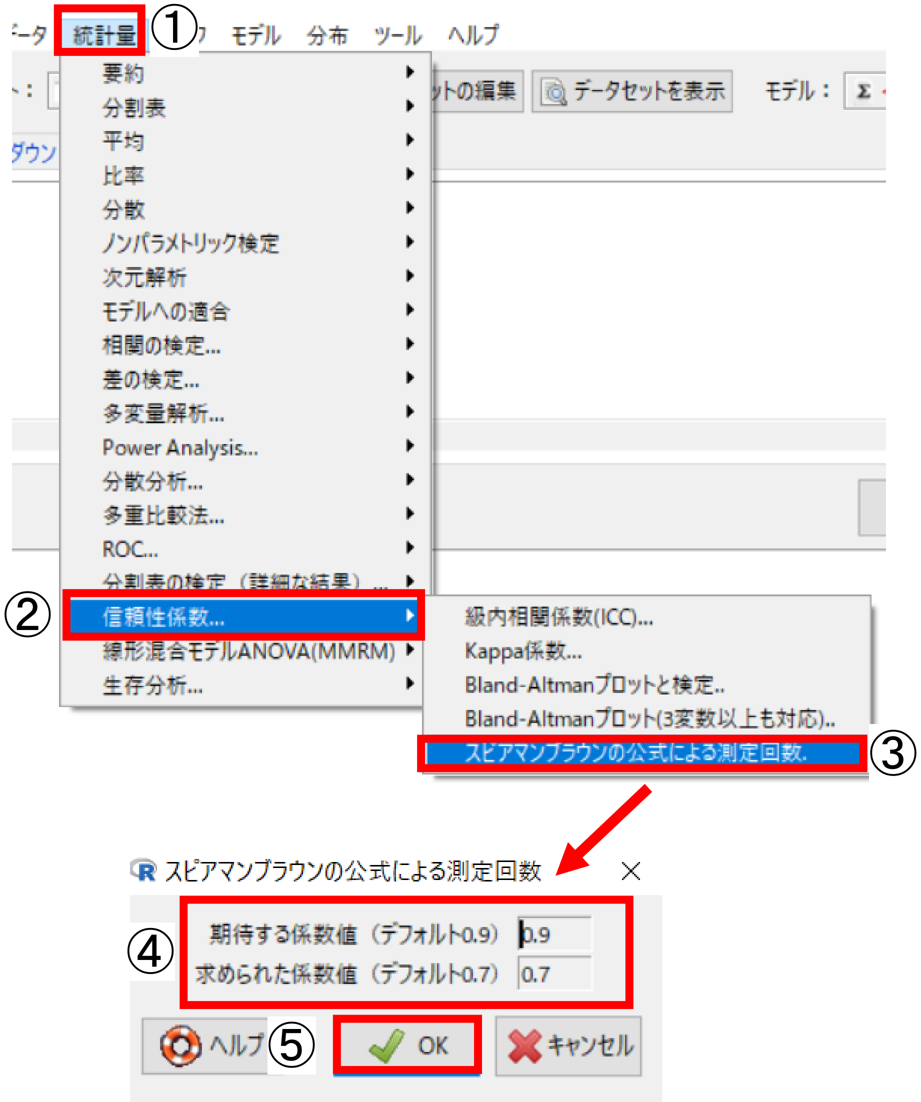
- 3変数以上に関する誤差の状態を確認するための参照とします
- 各行の平均をx軸，各測定値と各行の平均の差をy軸にプロットした散布図を出力しますが，正当なBland-Altman plotとは異なります
- 立位体前屈データを選びます
- ①[統計量]ー②[信頼性係数]ー③[Bland-Altmanプロットと検定（3変数以上も対応）]を選びます
- ダイアログボックスの④[2つ以上の変数を選択]から，変数をクリックして選びます
- 立位体前屈1回目と立位体前屈2回目と立位体前屈3回目を選んでいます
- ⑤[OK]をクリックします

統計量：信頼性係数
Bland-Altmanプロットと検定（3変数以上も対応）結果



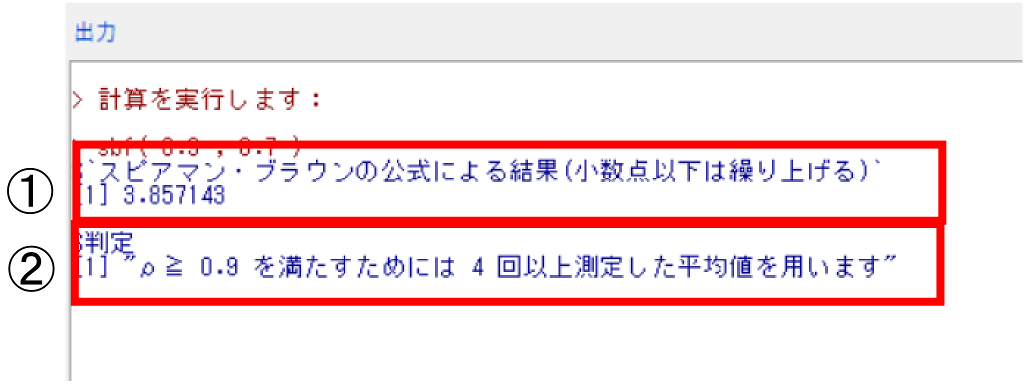
- グラフのx軸は各行の平均で，被検者ごとに求めた平均の全被検者の中でのバラツキが反映されます．y軸は各測定値と各行の平均の差で，各測定値の平均からのバラツキが反映されます
  - ✓ ここでは距離1回目を1，距離2回目を2，距離3回目を3としてグラフ上にプロットしています
- 通常のBland-Altman plotに準じて分布の偏り・形状から固定誤差や比例誤差の存在や，各測定値の特徴（特定の測定値で偏りが大きいなど）を確認します
  - ✓ このデータでは，全被検者で1回目が負方向，3回目が正方向へ偏る系統誤差の存在が示されます

統計量：信頼性係数
スピアマンブラウンの公式による測定回数



- 同一の特性について測定するいくつかの平行テストがあるとき，全平行テストによる合計得点の信頼性係数を用いて計算することができます
- 立位体前屈データを選びます
- ①[統計量]ー②[信頼性係数]ー③[スピアマンブラウンの公式による測定回数]を選びます
- ダイアログボックスの④[期待する係数値（デフォルト0.9）]と [求められた係数値（デフォルト0.7）]へは，それぞれ該当する係数値を記入します
  - ✓ 求められた係数値とは実際のデータで計算されたICCやKappa係数の値となります
- ⑤[OK]をクリックします

統計量:信頼性係数
スピアマンブラウンの公式による測定回数　結果



- ①には「期待する係数値」を得るために必要な測定回数が出力されています
  - ✓ 小数点以下は繰り上げますので，この例では4回と判断します
- ②には判定として結果がまとめられています
  - ✓ 4回以上測定した平均を使用すれば0.9以上を満たせそうです