

# Übereinstimmungsanalyse zweier stetiger Merkmale umgesetzt in einem SAS-Macro

Gisela Büchele  
Institut für Epidemiologie  
Universität Ulm  
Helmholtzstrasse 22  
89081 Ulm  
gisela.buechele@uni-ulm.de

## Zusammenfassung

### Hintergrund

In einer Vielzahl medizinischer Anwendungsbereiche besteht der Bedarf für die Übereinstimmungsüberprüfung zweier stetiger Merkmale.

### Statistische Methoden

Die Berechnung eines Korrelationskoeffizienten nach Pearson birgt das Problem, dass der Korrelationskoeffizient bei Abweichungen von der Winkelhalbierenden, die dennoch mehr oder minder auf einer Geraden liegen, unempfindlich ist. [1,2] Der Korrelationskoeffizient nach Spearman ist darüber hinaus noch robust gegen eine Verletzung der Linearitätsannahme eines Zusammenhangs.

Verschiedene Lösungsvorschläge wurden genannt, u.a. von Lin der Konkordanz-Korrelationskoeffizient [3] oder von anderen Autoren ein ICC(intra class correlation)-basiertes Konzept [4]. Der Konkordanz-Korrelationskoeffizient nach Lin ( $CCC_{Lin}$ ) gewichtet den Pearsonschen Korrelationskoeffizienten  $r$  mit einem Korrekturterm entsprechend der Abweichungen von der Winkelhalbierenden.

### Grafische Darstellung

Zur grafischen Ergebnispräsentation wurde in Ergänzung zur einfachen Punktwolke (mit Winkelhalbierender und Regressionsgerade) der Residuenplot von Bland und Altman empfohlen [2]. Diese Darstellung ermöglicht die qualitative Beurteilung einer systematischen Abweichung sowie der Streuung und trägt zusätzlich der speziellen Situation Rechnung, wenn keine der beiden Messungen als Referenz angesehen werden kann.

### Umsetzung als SAS-Macro

Da weder der  $CCC_{Lin}$  noch der Residuenplot nach Bland und Altman in SAS standardmäßig implementiert sind, wurde ein SAS-Macro erstellt, das alle notwendigen Berechnungen durchführt und grafisch ausgibt.

**Schlüsselwörter:** Übereinstimmungsanalyse, Konkordanz-Korrelationskoeffizient nach Lin, Residuenplot nach Bland und Altman

## 1 Hintergrund

In einer Vielzahl von medizinischen Anwendungsbereichen besteht der Bedarf für die Übereinstimmungsüberprüfung zweier stetiger Merkmale. Anwendungsbeispiele wären der Vergleich von Messwerten verschiedener Untersucher (*inter observer-Reliabilität*), verschiedener Messgeräte oder Labore sowie der Vergleich von Messwerten, die vom gleichen Untersucher in einem zeitlichen Abstand am gleichen Individuum erhoben wurden (*intra observer-Reliabilität*).

## 2 Statistische Methoden

Die Berechnung eines Korrelationskoeffizienten nach Pearson birgt das Problem, dass der Korrelationskoeffizient bei Abweichungen von der Winkelhalbierenden, die dennoch mehr oder minder auf einer Geraden liegen, unempfindlich ist. [1,2] Der Korrelationskoeffizient nach Spearman ist darüber hinaus noch robust gegen eine Verletzung der Linearitätsannahme eines Zusammenhangs. Verschiedene Lösungsvorschläge wurden genannt, u.a. von Lin der Konkordanz-Korrelationskoeffizient [3] oder von anderen Autoren ein ICC(intra class correlation)-basiertes Konzept [4].

### 2.1 Konkordanz-Korrelationskoeffizient

Der Korrelationskoeffizient nach Lin ( $CCC_{Lin}$ ) gewichtet den Pearsonschen Korrelationskoeffizienten  $r$  mit einem Korrekturterm entsprechend der Abweichungen von der Winkelhalbierenden. Die folgende Formel für  $CCC_{Lin}$  wurde nach Guggenmoos-Holzmann [1] wie folgt zitiert:

$$CCC_{Lin} = r_{Pearson} \cdot \frac{2}{\frac{S_A}{S_B} + \frac{S_B}{S_A} + \frac{(\bar{x}_A - \bar{x}_B)^2}{S_A \cdot S_B}}$$

wobei A und B die beiden Untersucher bzw. zu vergleichenden Untersuchungen darstellen und  $\bar{x}$  die Mittelwerte und S die Standardabweichungen der jeweiligen Messreihen sind.

### 2.2 Grafische Darstellung

Zur grafischen Ergebnispräsentation wurde in Ergänzung zur einfachen Punktwolke (mit Winkelhalbierender und Regressionsgerade) von Bland und Altman der Residuenplot empfohlen [2]. Diese Darstellung ermöglicht die qualitative Beurteilung einer systematischen Abweichung sowie der Streuung und trägt zusätzlich der speziellen Situation Rechnung, wenn keine der beiden Messungen als Referenz angesehen werden kann.

### 3 Umsetzung als SAS-Macro

Da weder der  $CCC_{Lin}$  noch der Residuenplot nach Bland und Altman in SAS standardmäßig implementiert sind, wurde ein SAS-Macro erstellt, das alle notwendigen Berechnungen durchführt und grafisch ausgibt.

#### 3.1 Macro-Definition und Aufruf

Definition: `%MACRO agreement (data, mw1, mw2, lab1, lab2);`

Aufruf: `%agreement (gesamt, wert1, wert2, Observer1, Observer2);`

#### 3.2 Berechnung des $CCC_{Lin}$

```
DATA Lab12;
  SET &data;
  Differences=&mw1-&mw2;
  Means=(&mw1+&mw2)/2;
  IF (&mw1 NE . AND &mw2 NE .);
RUN;

*Korrelationskoeffizient nach Pearson;
PROC CORR DATA=lab12 NOPRINT OUTP=rp;
  VAR &mw1 &mw2; RUN;
DATA _NULL_; SET rp;
  IF _N_=3 THEN CALL SYMPUT ('N', TRIM(LEFT(&mw1)));
  IF _N_=5 THEN CALL SYMPUT ('r_pearson',
    TRIM(LEFT(ROUND(&mw1, 0.01))));
RUN;

*Berechnung der Mittelwerte und Standardabweichungen;
PROC MEANS NOPRINT DATA=lab12;
  VAR differences &mw1 &mw2;
  OUTPUT OUT=ccc MEAN=meandiff mean1 mean2
    STD=stddiff std1 std2
    MAX=maxdiff max1 max2; RUN;

*Anlegen von Macro-Variablen;
DATA _NULL_;
  SET ccc;
  CALL SYMPUT ('mean', TRIM(LEFT(ROUND(meandiff, 0.01))));
  CALL SYMPUT ('std', TRIM(LEFT(ROUND(stddiff, 0.01))));
  CALL SYMPUT ('lowstd', TRIM(LEFT(ROUND(meandiff-2*stddiff, 0.01))));
  CALL SYMPUT ('uppstd', TRIM(LEFT(ROUND(meandiff+2*stddiff, 0.01))));
  CALL SYMPUT ('max', MAX(max1, max2));

  CCC_Lin = &r_pearson * ( 2 / (std1/std2 + std2/std1 +
    ((mean1-mean2)**2 / (std1*std2))));
  CALL SYMPUT ('CCC', TRIM(LEFT(ROUND(CCC_Lin, 0.01))));
RUN;
```

```
*zur Erstellung der Winkelhalbierenden;
DATA winkel;
  x=0; y=0; OUTPUT;
  x=&max; y=&max; OUTPUT;
RUN;
*Erstellung der SAS-Datei für die grafische Ausgabe;
DATA graf;
  SET lab12 winkel;
RUN;
```

## 3.2 Grafische Ergebnisausgabe

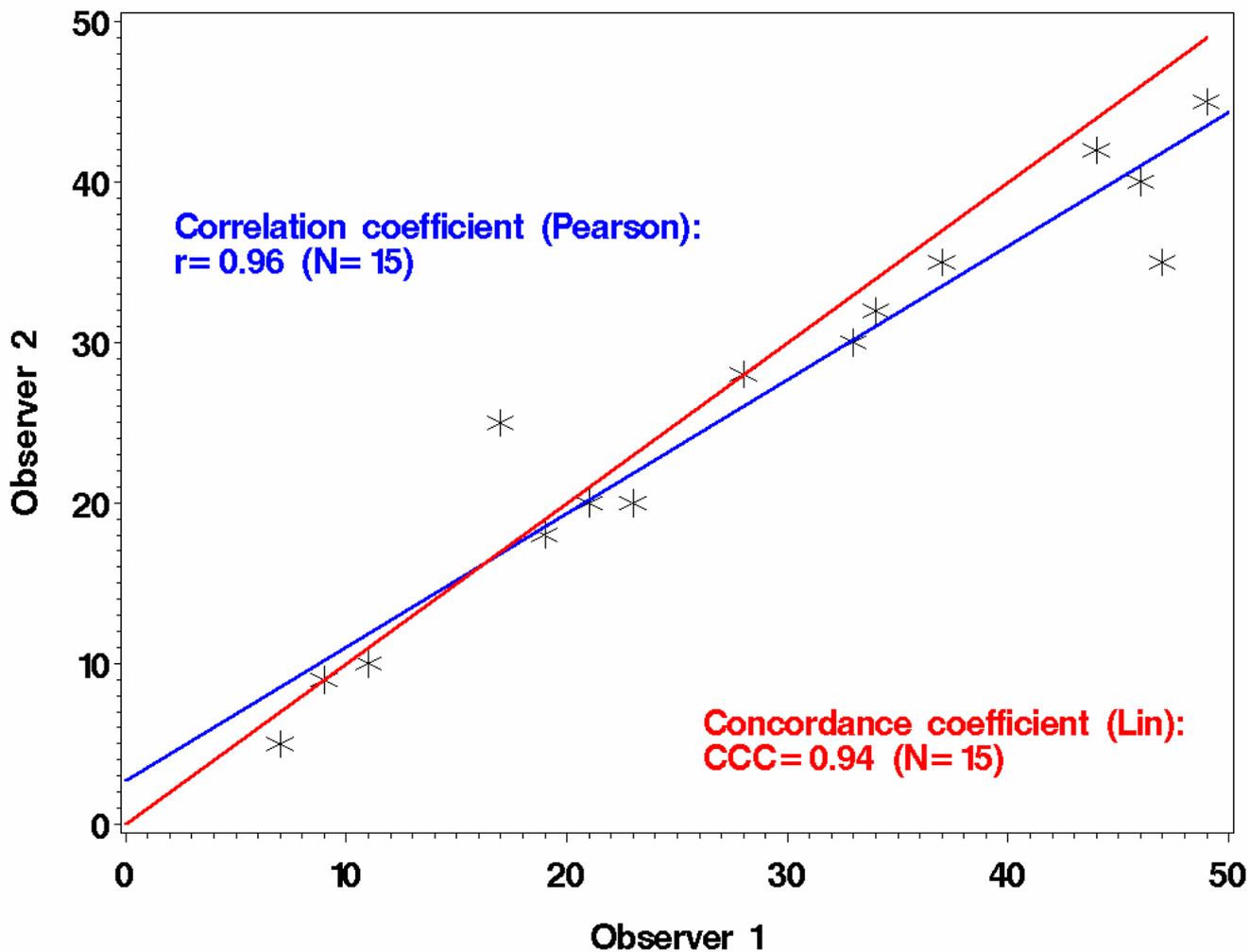
```
* Punktwolke mit Regressionsgeraden (dunkle Linie, "blue") und
  Winkelhalbierenden (helle Linie, "red");
```

```
SYMBOL1 V=star CV=black H=3 I=R W=2 C=blue;
SYMBOL2 V=none I=join H=1 W=2 C=red;
PROC GPLOT DATA=graf;
  PLOT &mw2*&mw1 y*x /VAXIS=axis1 HAXIS=axis2 OVERLAY;
  AXIS1 LABEL=(A=90 "&lab2");
  AXIS2 LABEL=("&lab1");
  NOTE H=2 C=blue M=(18,46) 'Correlation coefficient (Pearson):'
      M=(18,44) "r=&r_pearson (N=&N)";
  NOTE H=2 C=red M=(70,17) 'Concordance coefficient (Lin):'
      M=(70,15) "CCC=&CCC (N=&N)";
RUN; QUIT;
```

```
* Residuenplot;
```

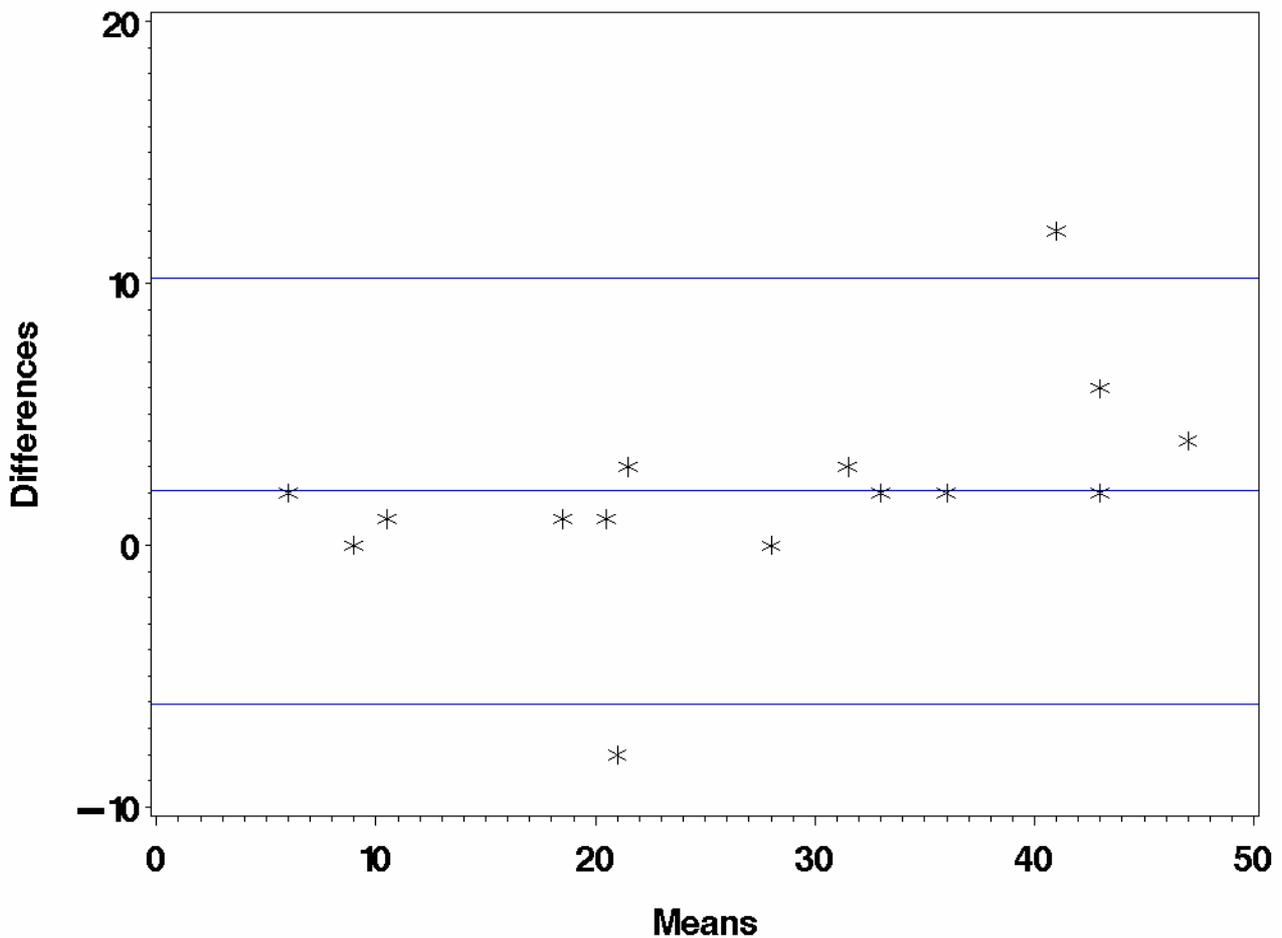
```
SYMBOL1 V=star CV=black H=2 I=none W=1 C=blue;
PROC GPLOT DATA=lab12;
  PLOT differences * means
    / VAXIS=axis1 CVREF=blue VREF=&mean &lowstd &uppstd ;
  AXIS1 LABEL=(A=90);
  TITLE1 H=1.8 "Residual plot: differences "
        H=1.5 "(&lab1 - &lab2) "
        H=1.8 " against means of values";
  FOOTNOTE1 J=C H=1.5 "Lines represent the mean +/- 2 SD"
              " (mean=&mean, SD=&std)";
RUN; QUIT;
```

## 4 Grafiken



**Abbildung 1:** Punktwolke mit Korrelationskoeffizienten nach Pearson und Regressionsgeraden (dunkle Linie) sowie mit  $CCC_{Lin}$  und Winkelhalbierenden (helle Linie)

**Residual plot: differences (Observer 1 – Observer 2) against means of values**



Lines represent the mean  $\pm$  2 SD (mean= 2.07, SD= 4.08)

**Abbildung 2:** Residuenplot nach Bland und Altman. Werte außerhalb des Bereiches Mittelwert  $\pm$ 2 Standardabweichungen gelten als starke Abweichungen.

## 5 Diskussion

Eine Vielzahl von Anwendungen nicht nur in der medizinischen Forschung bedürfen des Einsatzes von Methoden der Übereinstimmungsanalyse. Der von Lin vorgeschlagene Konkordanz-Korrelationskoeffizient dient der Berechnung einer Maßzahl zur Übereinstimmungsbeurteilung, die auch für Anwender aus nicht-statistischen Bereichen einfach bestimmbar und relativ leicht verständlich ist.

Der Residuenplot nach Bland und Altman ermöglicht eine informative grafische Darstellung, die als Ergänzung zur Punktwolke systematische und/oder große Abweichungen deutlich zeigt.

Die Umsetzung in ein SAS-Macro dient zur Reduktion des Programmieraufwandes bei wiederkehrenden Anwendungen und kann auch von SAS-ungeübten Benutzern problemlos angewandt werden. Die Ausgabe der Ergebnisse erfolgt direkt in Präsentationsform.

## **Literatur**

- [1] Guggenmoos-Holzmann I, Wernecke D. Medizinische Statistik. Berlin; Blackwell Wissenschaftsverlag. 1995
- [2] Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986; 307-10
- [3] Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics*. 1989; 45(1): 255-68
- [4] Shrout PE, Fleiss JL. Intra class correlations: Uses in assessing rater reliability. *Psychological Bulltin*. 1979 (86): 420-8