Testtheorie und Testkonstruktion

MSc Klinische Psychologie und Psychotherapie

SoSe 2024

Prof. Dr. Dirk Ostwald

(5) Reliabilität

Überblick

Reliabilität einer Testmessung

Paralleltestreliabilität

Selbstkontrollfragen

Überblick

Reliabilität einer Testmessung

Paralleltestreliabilität

Selbstkontrollfragen

Überblick

Testgütekriterium Reliabilität

"Ein Test ist dann *reliabel*, wenn er das Merkmal, das er misst, exakt, d.h. ohne Messfehler misst. Vor dem Hintergrund der Klassischen Testtheorie unterscheidet man die *Retest-Reliabilität*, die *Paralleltest-Reliabilität*, die *Testhalbierungs-Reliabilität (Split-Half-Reliabilität)* und die *Innere Konsistenz* eines Tests."

Moosbrugger und Kelava (2012)

- Die Reliabilität einer Testmessung als quadrierte Korrelation von Observed- und True-Score definiert.
- Anhand der Eigenschaften des Modells multipler Testmessungen ergeben sich verschiedene Interpretationen.
- Die Interpretation nach Moosbrugger und Kelava (2012) findet sich auch.
- Messfehlerfreie Testmessungen sind allerdings sicherlich nur ein nicht zu erreichendes Ideal.
- Wir betrachten zunächst die Reliabilität von Testmessungen im Modell multipler Testmessungen.
- Allerdings ergibt sich dabei keine Möglichkeit, die Reliabilität von Testmessungen empirisch zu messen.
- Zentral ist dann die Definition der *Paralleltestreliabilität*. Diese ist empirisch messbar.
- Bei der Retest-Reliabilität handelt es sich um eine spezielle Paralleltestreliabilität.
- Split-Half-Reliabilität und Innere Konsistenz betrachten wir in Einheit (6) Interne Konsistenz.

Überblick

Reliabilität einer Testmessung

Paralleltestreliabilität

Selbstkontrollfragen

Definition (Reliabilität einer Testmessung)

Gegeben sei das Modell multipler Testmessungen für eine beliebige Testmessung j mit $1 \leq j \leq m$,

$$\mathbb{P}(\tau_{1j}, \upsilon_{1j} ..., \tau_{nj}, \upsilon_{nj}) := \prod_{i=1}^{n} \mathbb{P}(\tau_{ij}, \upsilon_{ij}) := \prod_{i=1}^{n} \mathbb{P}(\upsilon_{ij} | \tau_{ij}) \mathbb{P}(\tau_{ij}) \tag{1}$$

wobei nach Definition des Modells gilt, dass

$$\mathbb{P}(\tau_{1j}, \upsilon_{1j}) = \cdots = \mathbb{P}(\tau_{nj}, \upsilon_{nj}). \tag{2}$$

Die *Reliabilität der Testmessung* j ist dann definiert als

$$\mathsf{R}_j := \rho(\upsilon_{ij}, \tau_{ij})^2 \text{ für ein beliebiges } 1 \leq i \leq n. \tag{3}$$

Bemerkungen

- Vor dem Hintergrund des einfachen Modells der Klassischen Testtheorie schreibt man auch $\mathsf{R} = \rho(\upsilon, \tau)^2$

Reliabilität einer Testmessung

Bemerkungen (fortgesetzt)

- Per Definition ist die Reliabilität eine Testmessung die quadrierte Korrelation von Observed- und True-Score.
- Mit $-1 \leq \rho(\upsilon_{ij}, \tau_{ij}) \leq 1$ folgt direkt, dass $0 \leq \mathrm{R}_j \leq 1$.
- $\mathrm{R}_j = 0$ impliziert $\rho(\upsilon_{ij}, \tau_{ij}) = 0$, also die linear-affine Unabhängigkeit von Observed-Score und True-Score.
- $\mathrm{R}_j = 0$ impliziert also, dass der Observed-Score hinsichtlich des True- Scores nicht aussagekräftig ist
- $\mathrm{R}_j = 1$ impliziert $\rho(\upsilon_{ij}, \tau_{ij}) = \pm 1$, also die linear-affine Abhängigkeit von Observed-Score und True-Score.
- $\mathrm{R}_j = 1$ impliziert also, dass der Observed-Score hinsichtlich des True-Scores aussagekräftig ist.
- Weil True-Scores nur indirekt beobachtbar sind, wird eine alternative Form zu *Reliabilitätsschätzung* genutzt.

Theorem (Eigenschaften der Reliabilität einer Testmessung)

Gegeben sei das Modell multipler Testmessungen für eine beliebige Testmessung j mit $1 \leq j \leq m$,

$$\mathbb{P}(\tau_{1j}, \upsilon_{nj} ..., \tau_{nj}, \upsilon_{nj}) := \prod_{i=1}^{n} \mathbb{P}(\upsilon_{ij}|\tau_{ij})\mathbb{P}(\tau_{ij}). \tag{4}$$

Dann gelten für die Reliabilät R_j der Testmessung j, dass

(1) $\mathsf{R}_j = \frac{\mathbb{V}(\tau_{ij})}{\mathbb{V}(\upsilon_{ij})}$.

(2) $\mathsf{R}_j = 1 - \frac{\mathbb{V}(\varepsilon_{ij})}{\mathbb{V}(\upsilon_{ij})}$.

Beweis

(1) Mit Aussage (5) des Theorems zu den Eigenschaften des Modells multipler Testmessungen gilt

$$\mathsf{R}_j = \rho(\upsilon_{ij}, \tau_{ij})^2 = \left(\frac{\mathbb{C}(\upsilon_{ij}, \tau_{ij})}{\mathbb{S}(\upsilon_{ij})\mathbb{S}(\tau_{ij})} \right)^2 = \frac{\mathbb{C}(\upsilon_{ij}, \tau_{ij})^2}{\mathbb{V}(\upsilon_{ij})\mathbb{V}(\tau_{ij})} = \frac{\mathbb{V}(\tau_{ij})^2}{\mathbb{V}(\upsilon_{ij})\mathbb{V}(\tau_{ij})} = \frac{\mathbb{V}(\tau_{ij})}{\mathbb{V}(\upsilon_{ij})}. \tag{5}$$

(2) Mit Aussage (4) des Theorems zu den Eigenschaften des Modells multipler Testmessungen gilt dann weiter

$$\mathsf{R}_j = \frac{\mathbb{V}(\tau_{ij})}{\mathbb{V}(\upsilon_{ij})} = \frac{\mathbb{V}(\upsilon_{ij}) - \mathbb{V}(\varepsilon_{ij})}{\mathbb{V}(\upsilon_{ij})} = \frac{\mathbb{V}(\upsilon_{ij})}{\mathbb{V}(\upsilon_{ij})} - \frac{\mathbb{V}(\varepsilon_{ij})}{\mathbb{V}(\upsilon_{ij})} = 1 - \frac{\mathbb{V}(\varepsilon_{ij})}{\mathbb{V}(\upsilon_{ij})}. \tag{6}$$

Reliabilität einer Testmessung

Bemerkungen

- Da τ_{ij} und ε_{ij} nicht direkt beobachtbar sind, sind (1) und (2) nur von theoretischem Interesse.
- Die Reliabilität einer Testmessung ist der Anteil von True-Score Varianz an Observed-Score Varianz,

$$\mathsf{R}_j = \frac{\mathbb{V}(\tau_{ij})}{\mathbb{V}(\upsilon_{ij})} = \frac{\mathbb{V}(\tau_{ij})}{\mathbb{V}(\tau_{ij}) + \mathbb{V}(\varepsilon_{ij})}. \tag{7}$$

- Ist $\mathbb{V}(\tau_{ij}) = 0$, so ist auch $\mathsf{R}_j = 0$, ist $\mathbb{V}(\varepsilon_{ij}) = 0$ so ist $\mathsf{R}_j = 1$.
- $\mathsf{R}_j > 0$ impliziert also True-Score Varianz.
- Die Reliabilität einer Testmessung ist 1 minus dem Anteil von Error-Score Varianz an Observed-Score Varianz.

Überblick

Reliabilität einer Testmessung

Paralleltestreliabilität

Selbstkontrollfragen

Definition (Reliabilität einer Paralleltestmessung)

Gegeben sei das Modell paralleler Testmessungen für eine beliebige Testmessung j mit $1 \leq j \leq m$,

$$\mathbb{P}(\tau_1, \upsilon_{1j} ..., \tau_n, \upsilon_{nj}) := \prod_{i=1}^{n} \mathbb{P}(\tau_i, \upsilon_{ij}) = \prod_{i=1}^{n} \mathbb{P}(\tau_i) \mathbb{P}(\upsilon_{ij} | \tau_i) \tag{8}$$

wobei nach Definition des Modells gilt, dass

$$\mathbb{P}(\tau_1, \upsilon_{1j}) = \cdots = \mathbb{P}(\tau_n, \upsilon_{nj}). \tag{9}$$

Die *Reliabilität der Paralleltestmessung* j ist dann definiert als

$$\mathsf{R}_j := \rho(\upsilon_{ij}, \tau_i)^2 \text{ für ein beliebiges } 1 \leq i \leq n. \tag{10}$$

Bemerkungen

- Vor dem Hintergrund des einfachen Modells der Klassischen Testtheorie schreibt man auch hier $\mathsf{R} = \rho(\upsilon, \tau)^2$

Theorem (Paralleltestreliabilität)

Gegeben sei das Modell der paralleler Testmessungen

$$\mathbb{P}(\tau_1, \upsilon_{1j} ..., \tau_n, \upsilon_{nj}) := \prod_{i=1}^{n} \mathbb{P}(\tau_i, \upsilon_{ij}) = \prod_{i=1}^{n} \mathbb{P}(\tau_i)\mathbb{P}(\upsilon_{ij}|\tau_i) \tag{11}$$

Dann gelten

(1) $\mathsf{R}_j = \frac{\mathbb{V}(\tau_i)}{\mathbb{V}(\upsilon_{ij})}$ für alle $1 \leq j \leq m$.

(2) $\mathsf{R}_j = 1 - \frac{\mathbb{V}(\varepsilon_{ij})}{\mathbb{V}(\upsilon_{ij})}$ für alle $1 \leq j \leq m$.

(3) $\mathsf{R}_j = \rho(\upsilon_{ij}, \upsilon_{ik}) = \mathsf{R}_k$ für alle $1 \leq j, k \leq m$.

Bemerkungen

- Aussagen (1) und (2) sind analog zum Theorem der Reliabilität multipler Testmessungen.
- Aussage (3) des Theorems begründet die Paralleltest- und Retestreliabilitätsschätzung.
- Man sagt dazu vereinfachend, dass “die Korrelation paralleler Testmessungen gleich ihrer Reliabilität ist”.
- Zur Messung der Reliabilität nutzt man eine Schätzung der Korrelation zweier paralleler Testmessungen.
- Die Klassische Testtheorie begründet dieses Vorgehen vor der Annahme von von True- und Error-Scores.

Paralleltestreliabilität

Beweis

(1) Mit Aussage (5) des Theorems zu den Eigenschaften des Modells paralleler Testmessungen gilt

$$\mathsf{R}_j = \rho(\upsilon_{ij}, \tau_i)^2 = \left(\frac{\mathbb{C}(\upsilon_{ij}, \tau_i)}{\mathbb{S}(\upsilon_{ij})\mathbb{S}(\tau_i)} \right)^2 = \frac{\mathbb{C}(\upsilon_{ij}, \tau_i)^2}{\mathbb{V}(\upsilon_{ij})\mathbb{V}(\tau_i)} = \frac{\mathbb{V}(\tau_i)^2}{\mathbb{V}(\upsilon_{ij})\mathbb{V}(\tau_i)} = \frac{\mathbb{V}(\tau_i)}{\mathbb{V}(\upsilon_{ij})}. \tag{12}$$

(2) Mit Aussage (4) des Theorems zu den Eigenschaften des Modells multipler Testmessungen gilt dann weiter

$$\mathsf{R}_j = \frac{\mathbb{V}(\tau_i)}{\mathbb{V}(\upsilon_{ij})} = \frac{\mathbb{V}(\upsilon_{ij}) - \mathbb{V}(\varepsilon_{ij})}{\mathbb{V}(\upsilon_{ij})} = \frac{\mathbb{V}(\upsilon_{ij})}{\mathbb{V}(\upsilon_{ij})} - \frac{\mathbb{V}(\varepsilon_{ij})}{\mathbb{V}(\upsilon_{ij})} = 1 - \frac{\mathbb{V}(\varepsilon_{ij})}{\mathbb{V}(\upsilon_{ij})}. \tag{13}$$

(3) Mit Aussagen (2) und (5) des Theorems zu Eigenschaften des Modells paralleler Testmessungen gilt dann weiter

$$\mathsf{R}_j = \frac{\mathbb{V}(\tau_i)}{\mathbb{V}(\upsilon_{ij})} = \frac{\mathbb{C}(\upsilon_{ij}, \upsilon_{ik})}{\sqrt{\mathbb{V}(\upsilon_{ij})}\sqrt{\mathbb{V}(\upsilon_{ij})}} = \frac{\mathbb{C}(\upsilon_{ij}, \upsilon_{ik})}{\sqrt{\mathbb{V}(\upsilon_{ij})}\sqrt{\mathbb{V}(\upsilon_{ik})}} = \rho(\upsilon_{ij}, \upsilon_{ik}) \tag{14}$$

und dass ebenso

$$\mathsf{R}_k = \frac{\mathbb{V}(\tau_i)}{\mathbb{V}(\upsilon_{ik})} = \frac{\mathbb{C}(\upsilon_{ij}, \upsilon_{ik})}{\sqrt{\mathbb{V}(\upsilon_{ik})}\sqrt{\mathbb{V}(\upsilon_{ik})}} = \frac{\mathbb{C}(\upsilon_{ij}, \upsilon_{ik})}{\sqrt{\mathbb{V}(\upsilon_{ij})}\sqrt{\mathbb{V}(\upsilon_{ik})}} = \rho(\upsilon_{ij}, \upsilon_{ik}). \tag{15}$$

Beispiel

Wir betrachten den Fall zweier Testmessungen $j = 1, 2$ im Modell paralleler Testmessungen. Für alle $i = 1, \ldots, n$ seien in WDF Form, wobei wir der notationellen Einfachheit halber auf das i Subskript verzichten wollen,

$$p(t) = N(t; 0, \sigma_\tau^2) \text{ und } p(y_1|t) := N(y_1; t, \sigma_\varepsilon^2) \text{ und } p(y_2|t) := N(y_2; t, \sigma_\varepsilon^2) \tag{16}$$

Für Person i gibt es also nur eine True-Score Zufallsvariable für alle Testmessungen und die Propensitätsverteilungen unterscheiden sich zwischen Testmessungen nicht.

Dann gilt zunächst mit dem Theorem zu gemeinsamen Normalverteilungen mit $A := 1$ und $b := 0$

$$p(t)p(y_1|t) = p(t, y_1) = N\left(\begin{pmatrix} t \\ y_1 \end{pmatrix}; \begin{pmatrix} \mu_\tau \\ \mu_\tau \end{pmatrix}, \begin{pmatrix} \sigma_\tau^2 & \sigma_\tau^2 \\ \sigma_\tau^2 & \sigma_\tau^2 + \sigma_\varepsilon^2 \end{pmatrix}\right) \tag{17}$$

und weiterhin mit $A := \begin{pmatrix} 1 & 0 \end{pmatrix}$ und $b := 0$

$$p(t, y_1)p(y_2|t) = p(t, y_1, y_2) = N\left(\begin{pmatrix} t \\ y_1 \\ y_2 \end{pmatrix}; \begin{pmatrix} \mu_\tau \\ \mu_\tau \\ \mu_\tau \end{pmatrix}, \begin{pmatrix} \sigma_\tau^2 & \sigma_\tau^2 & \sigma_\tau^2 \\ \sigma_\tau^2 & \sigma_\tau^2 + \sigma_\varepsilon^2 & \sigma_\tau^2 \\ \sigma_\tau^2 & \sigma_\tau^2 & \sigma_\tau^2 + \sigma_\varepsilon^2 \end{pmatrix}\right) \tag{18}$$

Beispiel (fortgeführt)

Damit gilt dann durch Ablesen am Kovarianzmatrixparameter von $p(t, y_1, y_2)$, dass

$$\rho(\upsilon_1, \upsilon_2) = \frac{\mathbb{C}(\upsilon_1, \upsilon_2)}{\sqrt{\mathbb{V}(\upsilon_1)}\sqrt{\mathbb{V}(\upsilon_2)}} = \frac{\sigma_\tau^2}{\sqrt{\sigma_\tau^2 + \sigma_\varepsilon^2}\sqrt{\sigma_\tau^2 + \sigma_\varepsilon^2}} = \frac{\sigma_\tau^2}{\sigma_\tau^2 + \sigma_\varepsilon^2} \tag{19}$$

Insbesondere gilt also auch für $j = 1, 2$

$$\mathsf{R}_j = \rho(\upsilon_j, \tau)^2 = \left(\frac{\mathbb{C}(\upsilon_j, \tau)}{\mathbb{S}(\upsilon_j)\mathbb{S}(\tau)}\right)^2 = \frac{\mathbb{C}(\upsilon_j, \tau)^2}{\mathbb{V}(\upsilon_j)\mathbb{V}(\tau)} = \frac{\left(\sigma_\tau^2\right)^2}{(\sigma_\tau^2 + \sigma_\varepsilon^2)\sigma_\tau^2} = \frac{\sigma_\tau^2}{\sigma_\tau^2 + \sigma_\varepsilon^2} = \rho(\upsilon_1, \upsilon_2). \tag{20}$$

Gilt also beispielsweise $\sigma_\tau^2 := 1.0$ und $\sigma_\varepsilon^2 := 0.2$, so ergibt sich für die Paralleltestreliabilität

$$\mathsf{R}_j = \rho(\upsilon_j, \tau)^2 = \rho(\upsilon_1, \upsilon_2) = \frac{1.0}{1.0 + 0.2} \approx 0.83. \tag{21}$$

Definition (Paralleltestreliabilitätsschätzer)

Gegeben sei das Modell paralleler Testmessungen für n Personen und zwei Testmessungen $j = 1, 2$,

$$\mathbb{P}(\tau_1, \upsilon_{11}, \upsilon_{12}, ..., \tau_n, \upsilon_{n1}, \upsilon_{n2}) = \prod_{i=1}^{n} \mathbb{P}(\tau_i)\mathbb{P}(\upsilon_{i1}|\tau_i)\mathbb{P}(\upsilon_{i2}|\tau_i) \tag{22}$$

Dann wird der mit den Stichprobenmitteln

$$\bar{\upsilon}_1 := \frac{1}{n}\sum_{i=1}^{n} \upsilon_{i1} \text{ und } \bar{\upsilon}_2 := \frac{1}{n}\sum_{i=1}^{n} \upsilon_{i2} \tag{23}$$

definierte Stichprobenkorrelationskoeffizient

$$r_{12} := \frac{\frac{1}{n-1}\sum_{i=1}^{n}(\upsilon_{i1} - \bar{\upsilon}_1)(\upsilon_{i2} - \bar{\upsilon}_2)}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\upsilon_{i1} - \bar{\upsilon}_1)^2}\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\upsilon_{i2} - \bar{\upsilon}_2)^2}} \tag{24}$$

Paralleltestreliabilitätsschätzer genannt.

Bemerkungen

- **R** bietet mit der `cor()` Funktion eine Möglichkeit, Stichprobenkorrelationskoeffizienten zu berechnen.

Paralleltestreliabilität

Simulationsbeispiel

$p(t_i) = N(t_i; 0, \sigma^2_\tau)$ und $p(y_{ij}|t_i) := N(y_{ij}; t_i, \sigma^2_\varepsilon)$ für $j = 1, 2$ mit $\sigma^2_\tau := 1.0$ und $\sigma^2_\varepsilon := 0.2$ für $i = 1, \ldots, 30$.

```
n           = 30                                    # Personenanzahl
m           = 2                                     # Testmessungsanzahl
sigsqr_tau  = 1                                     # True-Score Varianz
sigsqr_eps  = .2                                    # Observed-Score-Varianz
R_12        = sigsqr_tau/(sigsqr_tau+sigsqr_eps)    # Paralletestreliabilität
T           = matrix(rep(NaN, n)  , nrow = n)       # True-Score Array
Y           = matrix(rep(NaN, n*m), nrow = n)       # Observed-Score Array
E           = matrix(rep(NaN, n*m), nrow = n)       # Error-Score Array
for(i in 1:n){                                      # Personeniterationen
    T[i]  = rnorm(1,1,sqrt(sigsqr_tau))             # True-Score Realisierung für j = 1,2
    for(j in 1:m){                                  # Testmessungsiterationen
        Y[i,j]  = rnorm(1,T[i],sqrt(sigsqr_eps))    # Observed-Score Realisierung f
        E[i,j]  = Y[i,j] - T[i]}}                   # Error-Score Realisierung
r_12      = cor(Y[,1],Y[,2])                        # Paralleltestreliabilitätsschätzer
```

```
Paralleltestreliabilität R_12           :  0.8333
Paralleltestrelibabilitätsschätzer  r_12 :  0.9105
```

Theorem (Fishertransformation des Paralleltestreliabilitätsschätzers)

Gegeben sei das Modell paralleler Testmessungen für n Personen mit Paralleltestreliabilität R_j. Für zwei Testmessungen $j = 1, 2$ sei r_{12} der Paralleltestreliabilitätsschätzer. Dann gilt, dass

$$\tilde{r}_{12} := \frac{1}{2} \ln \left(\frac{1 + r_{12}}{1 - r_{12}} \right) \tag{25}$$

asymptotisch normalverteilt ist nach

$$\tilde{r}_{12} \overset{a}{\sim} N \left(\mathsf{R}_j, (n-3)^{-1} \right) \tag{26}$$

Bemerkungen

- Die Transformation eines Stichprobenkorrelationskoeffizienten anhand von

$$\tilde{r} := \frac{1}{2} \ln \left(\frac{1 + r}{1 - r} \right) \tag{27}$$

  wird nach Fisher (1925) als *Fisher-Transformation* bezeichnet.
- Für eine ausführliche Darstellung, siehe Johnson, Kotz, und Balakrishnan (1994) Kapitel 32.
- Die approximative Verteilung von $\tilde{r}_{12}$ kann als Grundlage für Konfidenzintervalle dienen.

Paralleltestreliabilität

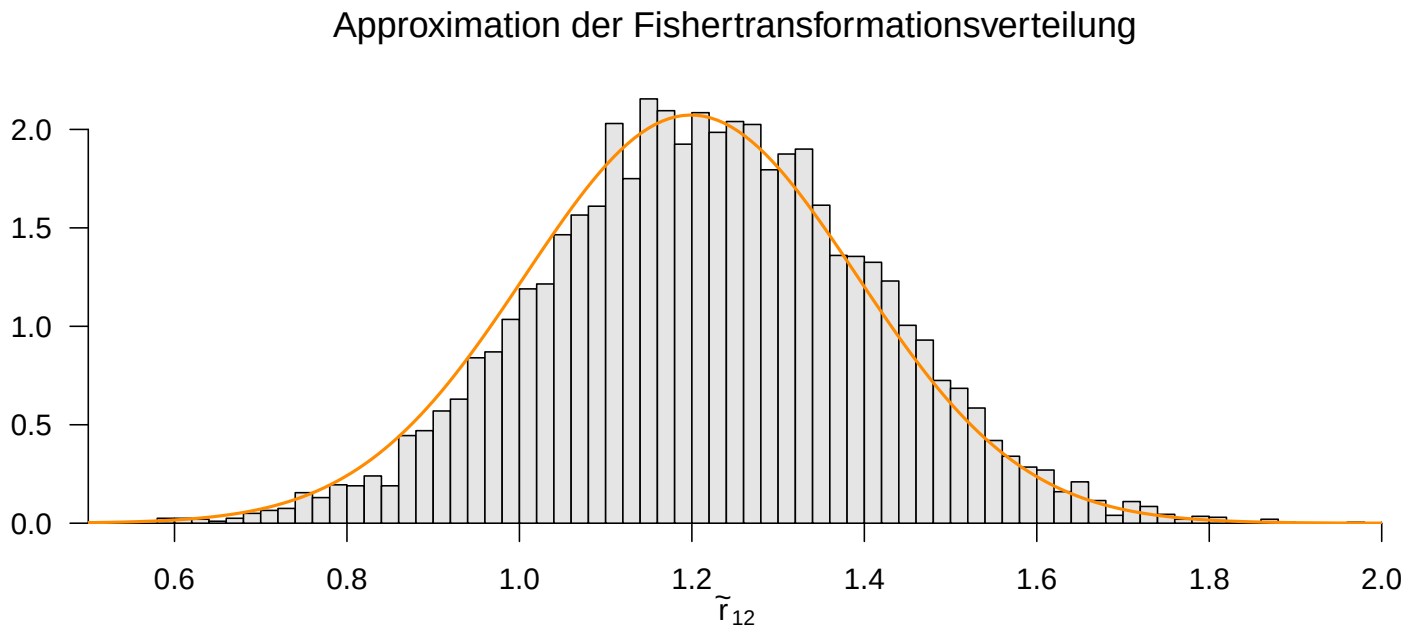
Simulationsbeispiel

$p(t_i) = N(t_i; 0, \sigma^2_\tau)$ und $p(y_{ij}|t_i) := N(y_{ij}; t_i, \sigma^2_\varepsilon)$ für $j = 1, 2$ mit $\sigma^2_\tau := 1.0$ und $\sigma^2_\varepsilon := 0.2$ für $i = 1, \ldots, 30$.

```
nsim               = 1e4                                  # Realisierungsanzahl
n                  = 30                                   # Personenanzahl
m                  = 2                                    # Testmessungsanzahl
sigsqr_tau         = 1                                    # True-Score Varianz
sigsqr_eps         = .2                                   # Observed-Score-Varianz
R_12               = sigsqr_tau/(sigsqr_tau+sigsqr_eps)   # Paralletestreliabilität
tilde_r_12         = rep(NaN, nsim)                       # Fisher Transformation von r_12 Array
for(s in 1:nsim){                                         # Simulationsiterationen
    T              = matrix(rep(NaN, n)  , nrow = n)      # True-Score Array
    Y              = matrix(rep(NaN, n*m), nrow = n)      # Observed-Score Array
    E              = matrix(rep(NaN, n*m), nrow = n)      # Error-Score Array
    for(i in 1:n){                                        # Personeniterationen
        T[i]       = rnorm(1,1,sqrt(sigsqr_tau))          # True-Score Realisierung für j = 1,2
        for(j in 1:m){                                    # Testmessungsiterationen
            Y[i,j] = rnorm(1,T[i],sqrt(sigsqr_eps))       # Observed-Score Realisierung f
            E[i,j] = Y[i,j] - T[i]}}                      # Error-Score Realisierung
    r_12           = cor(Y[,1],Y[,2])                     # Paralleltestreliabilitätsschätzer
    tilde_r_12[s]  = 1/2*log((1+r_12)/(1-r_12))           # Fisher Transformation von r_12
}
```

Simulationsbeispiel

$p(t_i) = N(t_i; 0, \sigma^2_\tau)$ und $p(y_{ij}|t_i) := N(y_{ij}; t_i, \sigma^2_\varepsilon)$ für $j = 1, 2$ mit $\sigma^2_\tau := 1.0$ und $\sigma^2_\varepsilon := 0.2$ für $i = 1, ..., 30$.



Überblick

Reliabilität einer Testmessung

Paralleltestreliabilität

Selbstkontrollfragen

Selbstkontrollfragen

1. Geben Sie die Definition der Reliabilität einer Testmessung wieder.
2. Geben Sie das Theorem zu den Eigenschaften der Reliabiltät einer Testmessung wieder.
3. Geben Sie die Definition der Reliabilität einer Paralleltestmessung wieder.
4. Geben Sie das Theorem zur Paralleltestreliabilität wieder.
5. Geben Sie die Definition des Paralleltestreliabilätsschätzers wieder.

Referenzen

Fisher, R. A. 1925. „Theory of Statistical Estimation“. *Mathematical Proceedings of the Cambridge Philosophical Society* 22 (5): 700–725. https://doi.org/10.1017/S0305004100009580.

Johnson, John, Kotz, und N. Balakrishnan. 1994. *Continuous Univariate Distributions, Volume 1*.

Moosbrugger, Helfried, und Augustin Kelava, Hrsg. 2012. *Testtheorie und Fragebogenkonstruktion: mit 66 Abbildungen und 41 Tabellen*. 2., aktualisierte und überarbeitete Auflage. Springer-Lehrbuch. Berlin Heidelberg: Springer.