

Übung zur Vorlesung - Theorien Psychometrischer Tests II

N. Rose

2. Übung (20.11.2008)



- **Agenda**
 - Item-, Testinformationsfunktion und Standardfehler der Personenparameter
 - Berechnung der „*Marginal Error Variance*“ und der Andrich's Reliabilität aus dem WINMIRA-Output
 - Rückblick „Spezifische Objektivität“

Standardfehler der Personenparameter

- Erläutern Sie den Zusammenhang zwischen Iteminformationsfunktion und Testinformationsfunktion.

$$I(T | \xi = \xi(u)) = \sum_{i=1}^k I(Y_i | \xi = \xi(u))$$

- Erläutern Sie den Zusammenhang zwischen Testinformationsfunktion und den Standardfehlern der Personenparameterschätzungen.

$$SE(\hat{\xi} | \xi = \xi(u)) = \frac{1}{\sqrt{\sum_{i=1}^k I(Y_i | \xi = \xi(u))}}$$

Reliabilität - WINMIRA

score frequency | person parameters and standard errors:

Raw-score	Expected freq.	MLE-estimate	std. error MLE	WLE-estimate	std. error WLE
0	0.02	*****	*****	-6.440	1.861
1	0.07	-4.995	1.260	-4.708	1.190
2	0.22	-3.753	0.999	-3.578	0.968
3	0.61	-2.899	0.859	-2.779	0.841
4	1.55	-2.238	0.774	-2.157	0.765
5	3.49	-1.681	0.722	-1.626	0.717
6	6.98	-1.187	0.687	-1.147	0.685
7	12.40	-0.731	0.664	-0.703	0.663
8	19.58	-0.301	0.649	-0.284	0.648
9	27.49	0.115	0.642	0.120	0.642
10	34.32	0.526	0.643	0.519	0.642
11	38.10	0.944	0.652	0.924	0.651
12	37.60	1.380	0.669	1.349	0.667
13	32.99	1.842	0.693	1.803	0.691
14	25.74	2.344	0.726	2.292	0.722
15	17.85	2.905	0.776	2.825	0.768
16	11.01	3.575	0.872	3.432	0.847
17	6.04	4.525	1.119	4.206	1.018
18	2.94	*****	*****	5.562	1.638

$$s^2(\hat{\xi})$$

$$s(\hat{\xi})$$

$$\frac{1}{N} \sum_{u=1}^N SE[\hat{\xi}(u)]^2$$

$$\sqrt{\frac{1}{N} \sum_{u=1}^N SE[\hat{\xi}(u)]^2}$$

WLE estimates : Mean = 1.151 Var = 1.778 stdev = 1.334
 marginal error variance = 0.504 stdev = 0.710
 anova reliability = 0.779
 Andrichs reliability = 0.717



Kennwerte anhand des Outputs berechnen

- Mittelwert des Schätzers für ξ : $\bar{\hat{\xi}} = \frac{1}{N} \sum_{u=1}^N \hat{\xi}(u) = \frac{1}{N} \sum_{j=0}^k n_{S=j} \cdot \hat{\xi}_{S=j}$

→ die Gewichte $n_{s=j}$ werden in der Spalte „Expected Freq.“ angegeben.

- Varianz des Schätzers für ξ :

$$s^2(\hat{\xi}) = \frac{1}{N} \sum_{u=1}^N \left(\hat{\xi}(u) - \bar{\hat{\xi}} \right)^2 = \frac{1}{N} \sum_{j=0}^k n_{S=j} \cdot \left(\hat{\xi}_{S=j} - \bar{\hat{\xi}} \right)^2$$

- Standardabweichung des Schätzers für ξ :

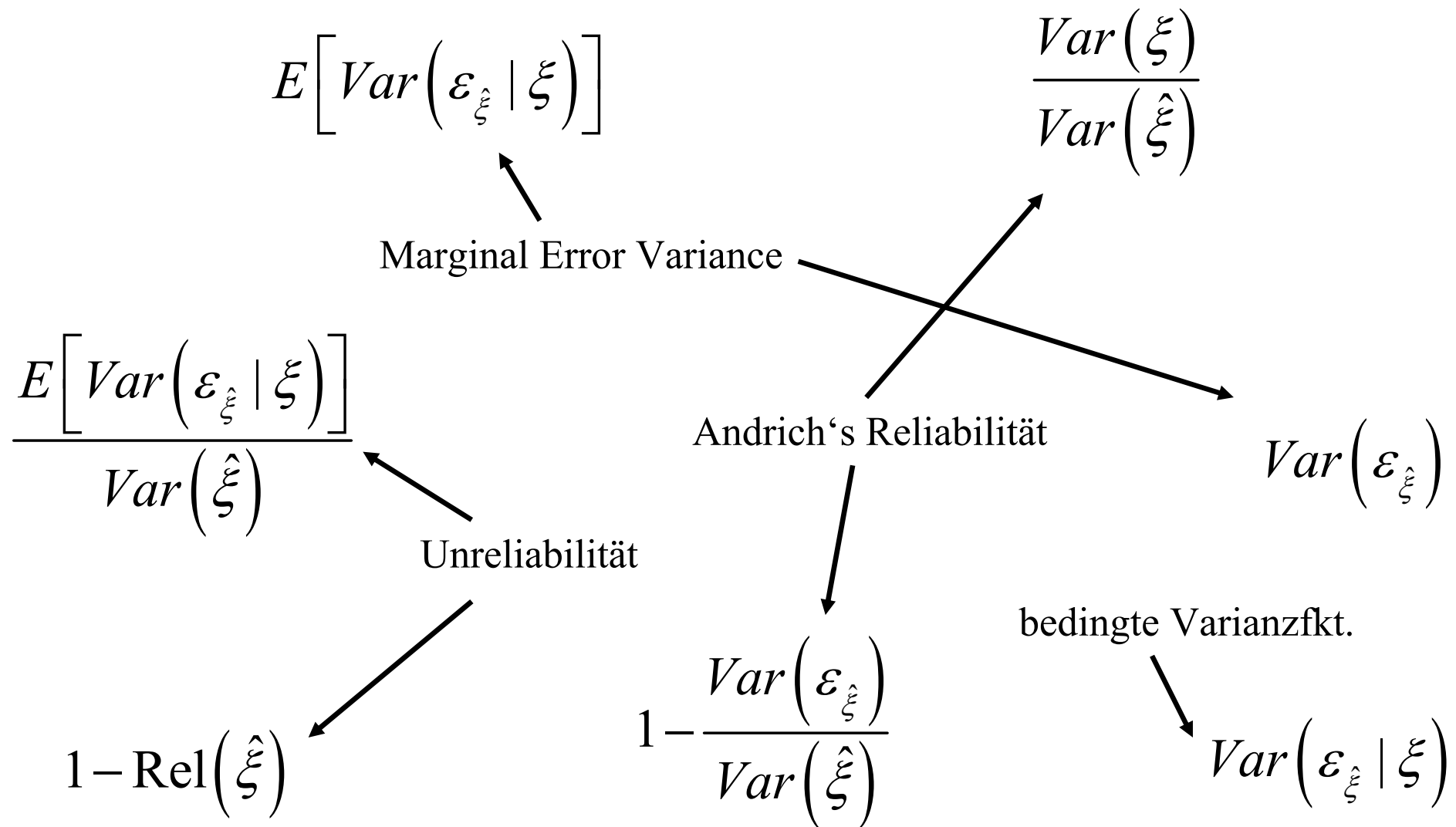
$$s(\hat{\xi}) = \sqrt{s^2(\hat{\xi})}$$

Kennwerte anhand des Outputs berechnen

- Berechnen der Andrich Reliabilität in 2 Schritten:
 1. Berechnen der *Marginal Error Variance* (Randfehlervarianz)
 2. Berechnen der Andrich Reliabilität

➔ Aber vorher: „Munteres Begriffe raten!!“

Begriffe?!



Kennwerte anhand des Outputs berechnen

1. Berechnen der Marginal Error Variance:

→ bei abzählbaren Ausprägungen (nur als Merkhilfe! ... eigentlich ist ξ eine metrische Variable!)

$$E \left[\text{Var} \left(\varepsilon_{\hat{\xi}} \mid \xi \right) \right] = \sum_{v=1} \text{Var} \left(\varepsilon_{\hat{\xi}} \mid \xi = \xi(v) \right) \cdot P \left(\xi = \xi(v) \right)$$

→ die Marginal Error Variance ist die über die Verteilung von ξ gemittelte Fehlervarianz bzw. der Erwartungswert der bedingten Varianzfunktion des Fehlers gegeben ξ .

→ wichtig:
$$\text{Var} \left(\hat{\xi} \mid \xi \right) = \text{Var} \left(\xi + \varepsilon_{\hat{\xi}} \mid \xi \right) = \text{Var} \left(\xi \mid \xi \right) + \text{Var} \left(\varepsilon_{\hat{\xi}} \mid \xi \right)$$
$$= \text{Var} \left(\varepsilon_{\hat{\xi}} \mid \xi \right)$$

$$\text{Var} \left(\varepsilon_{\hat{\xi}} \mid \xi \right) = \text{SE} \left(\hat{\xi} \mid \xi \right)^2$$

Kennwerte anhand des Outputs berechnen

1. Berechnen der Marginal Error Variance:

- in der Stichprobe kann die Marginal Error Variance wie folgt geschätzt werden:

$$\begin{aligned} \text{Var}\left(\varepsilon_{\hat{\xi}}\right) &= \frac{1}{N} \sum_{u=1}^N \text{SE}\left(\hat{\xi} \mid \xi = \xi(u)\right)^2 \\ &= \frac{1}{N} \sum_{u=1}^N \text{SE}\left(\hat{\xi} \mid U = u\right)^2 \end{aligned}$$

- da jedoch alle Personen mit dem selben Summenscore die gleiche Fähigkeitsschätzung erhalten, vereinfacht sich die Berechnung:

$$\begin{aligned} \text{Var}\left(\varepsilon_{\hat{\xi}}\right) &= \frac{1}{N} \sum_{j=0}^k n_{S=j} \cdot \text{SE}\left(\hat{\xi} \mid S = j\right)^2 \\ &= \frac{1}{N} \sum_{j=0}^k n_{S=j} \cdot \text{Var}\left(\varepsilon_{\hat{\xi}} \mid S = j\right)^2 \end{aligned}$$

Kennwerte anhand des Outputs berechnen

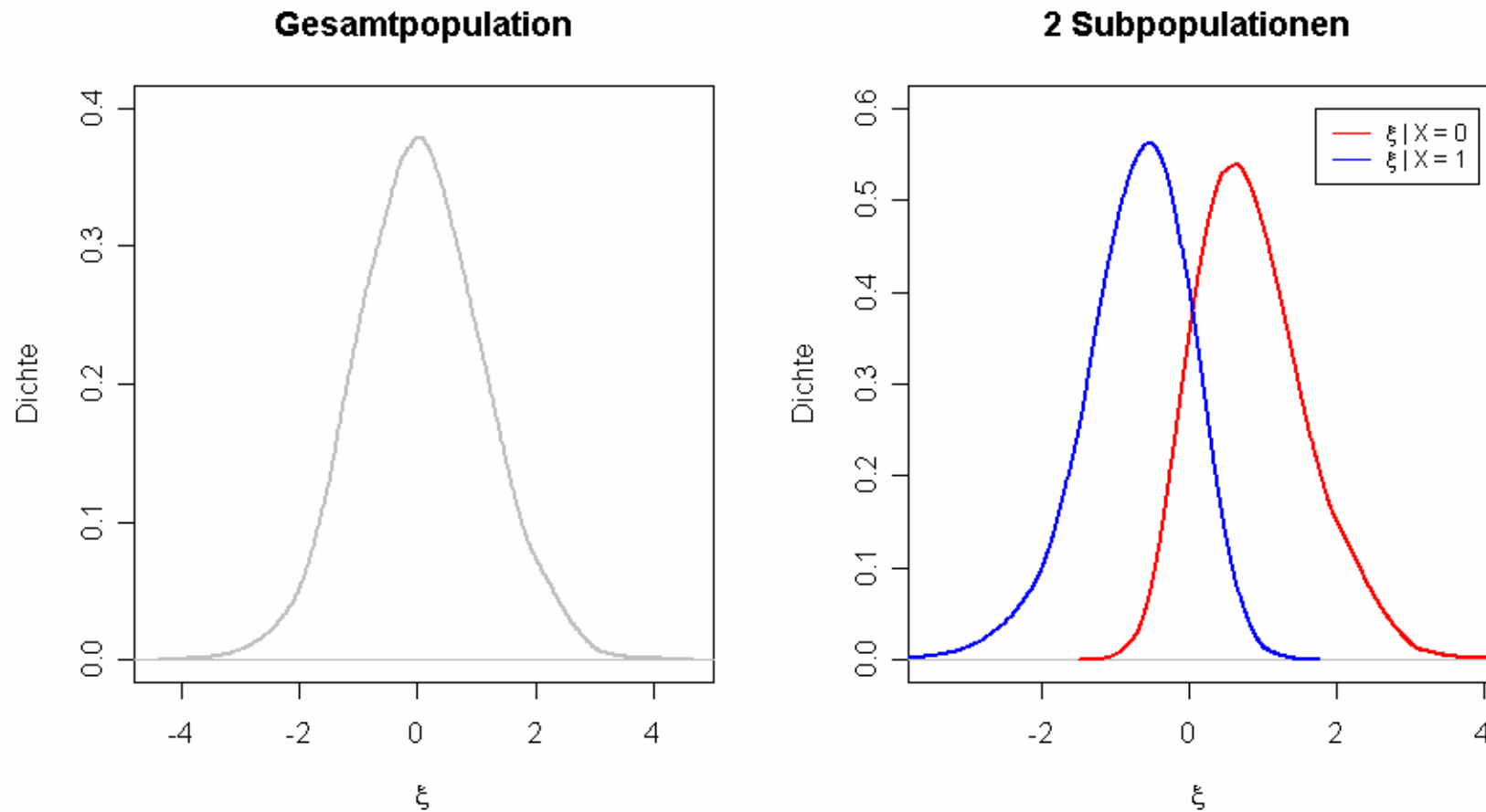
2. Berechnen der Andrich Reliabilität:

$$\text{Rel}(\hat{\xi}) = 1 - \frac{\text{Var}(\varepsilon_{\hat{\xi}})}{\text{Var}(\hat{\xi})}$$

$$\text{Rel}(\hat{\xi}) = 1 - \frac{\frac{1}{N} \sum_{j=0}^k n_{S=j} \cdot SE(\hat{\xi} | S = j)^2}{\text{Var}(\hat{\xi})}$$

Itemparameter in verschiedenen Subpopulationen

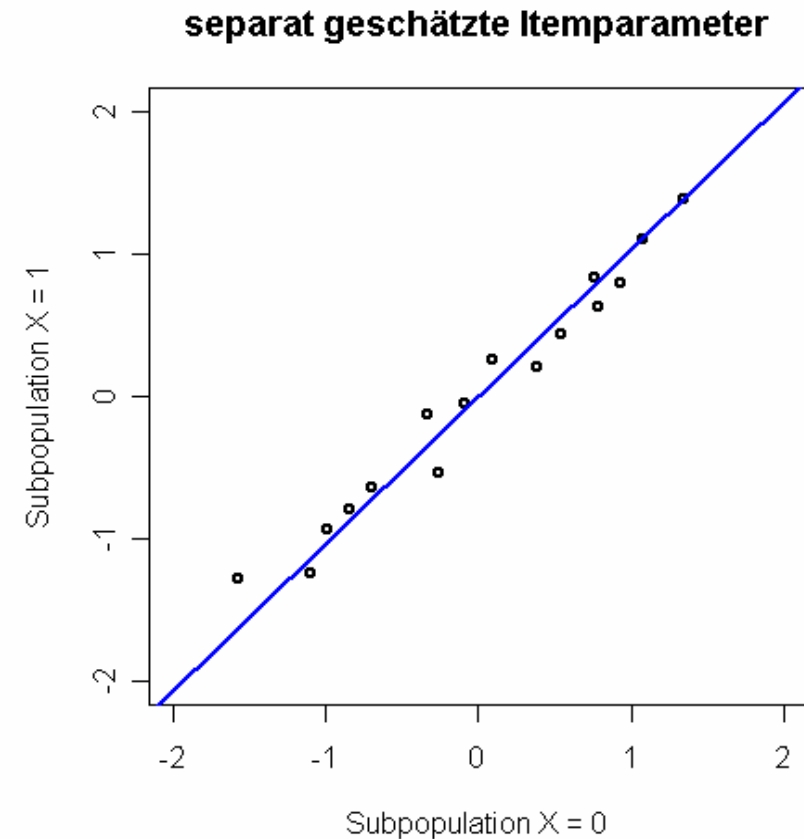
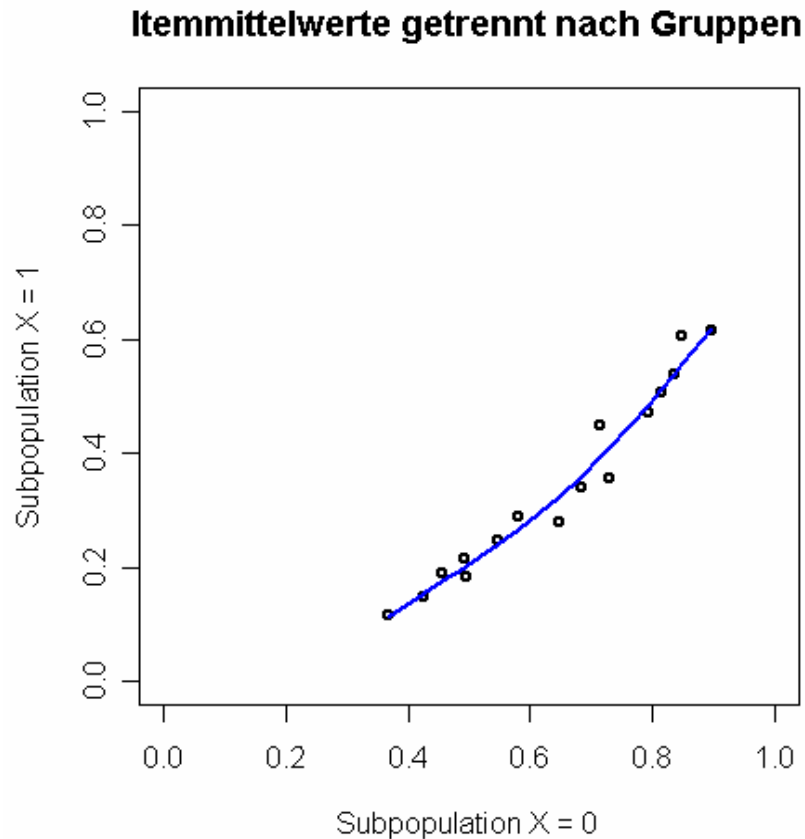
- **Datenbsp.:** „bsp_V3_1.sav“ bzw. „bsp_V3_1x0.sav“, „bsp_V3_1x1.sav“



- 2 Subpopulationen in denen das Rasch-Modell gilt!

Itemparameter in verschiedenen Subpopulationen

- Erläutern Sie an den beiden Graphiken die spezifische Objektivität.



Antwort: Die Differenzen der Itemschwierigkeiten sind im Rasch-Modell bis auf zufällige Stichprobenschwankungen gleich in beiden Gruppen (Spezifische Objektivität!). Die Differenzen der Itemmittelwerte (als Schätzer der unbedingten Lösungswahrscheinlichkeiten) sind nicht gleich über die Gruppen.

Reliabilität vs. Standardfehler

Expected Score Frequencies and Personparameters:

Raw- score	Expected freq.	MLE- estimate	std. error MLE	WLE- estimate	std. error WLE
0	72.28	*****	*****	-3.333	1.564
1	127.58	-2.414	1.079	-2.079	0.961
2	195.87	-1.549	0.825	-1.398	0.795
3	261.56	-0.954	0.729	-0.877	0.721
4	303.81	-0.457	0.687	-0.423	0.685
5	306.92	0.004	0.674	0.005	0.674
6	269.70	0.464	0.686	0.431	0.684
7	206.14	0.958	0.727	0.882	0.718
8	137.04	1.548	0.821	1.398	0.792
9	79.24	2.407	1.076	2.071	0.957
10	39.85	*****	*****	3.317	1.558

WLE estimates : Mean = -0.175 Var = 1.594 stdev = 1.262
 marginal error variance = 0.662 stdev = 0.814
 anova reliability = 0.706
 Andrichs reliability = 0.584

Expected Score Frequencies and Personparameters:

score frequency | person parameters and standard errors

Raw- score	Expected freq.	MLE- estimate	std. error MLE	WLE- estimate	std. error WLE
0	0.50	*****	*****	-3.320	1.566
1	1.93	-2.399	1.080	-2.062	0.961
2	6.43	-1.534	0.824	-1.381	0.793
3	18.44	-0.942	0.727	-0.863	0.718
4	45.69	-0.449	0.684	-0.413	0.682
5	97.80	0.006	0.670	0.008	0.670
6	180.80	0.460	0.681	0.428	0.680
7	288.67	0.949	0.723	0.872	0.714
8	398.06	1.533	0.818	1.381	0.788
9	474.06	2.387	1.074	2.049	0.953
10	487.61	*****	*****	3.290	1.554

WLE estimates : Mean = 1.703 Var = 1.264 stdev = 1.124
 marginal error variance = 1.084 stdev = 1.041
 anova reliability = 0.508
 Andrichs reliability = 0.142

Daten von zwei Subpopulationen (Stichproben je 1000 Fälle) die sich hinsichtlich der Verteilung der latenten Variable unterscheiden. Beide Datensätze wurden unter Gültigkeit des Rasch-Modells erzeugt (10 Items).