
Jürgen Rost

Lehrbuch Testtheorie Testkonstruktion

Verlag Hans Huber
Bern · Göttingen · Toronto · Seattle

Inhaltsverzeichnis

	ÜBER DIESES BUCH	9
	LESEHINWEISE	14
1.	WAS IST TESTTHEORIE?	17
1.1	Der Gegenstand der Testtheorie	17
✗ 1.1.1	Was ist ein Test?	17
1.1.2	Warum eine Theorie über Tests?	20
1.2	Die wissenschaftstheoretischen Grundlagen der Testtheorie	22
1.2.1	Was sind Theorien?	22
1.2.2	Was ist ein formales Modell?	24
✗ 1.2.3	Was sind Testmodelle?	27
✗ 1.2.4	Was erklären Theorien über das Testverhalten?	28
2.	TESTKONSTRUKTION	31
✗ 2.1	Gütekriterien für Tests	31
2.1.1	Validität	32
2.1.2	Reliabilität und Meßgenauigkeit	34
2.1.3	Objektivität	37
2.1.4	Logische Beziehungen zwischen den drei Gütekriterien	39
2.1.5	Normierung	40
✗ 2.2	Schritte der Testentwicklung	42
2.2.1	Arten von latenten Variablen	42
✗ 2.2.2	Arten von Tests	44
	2.2.2.1 Leistungstests	44
	2.2.2.2 Persönlichkeitsfragebögen	46
	2.2.2.3 Objektive Persönlichkeitstests	47
	2.2.2.4 Projektive Tests	48
	2.2.2.5 Situationsfragebögen	50
	2.2.2.6 Einstellungstests	50
	2.2.2.7 Motivations- und Interessensfragebögen	53
	2.2.2.8 Verhaltensfragebögen	54
2.2.3	Definition des Itemuniversums	56
2.2.4	Ziehung einer Itemstichprobe	57
2.2.5	Auswahl eines geeigneten Testmodells	58

2.3	Itemkonstruktion	60
2.3.1	Arten von Antwortformaten	61
2.3.1.1	Freie Antwortformate	61
2.3.1.2	Gebundene Antwortformate	63
2.3.1.3	Ratingformate	66
2.3.2	Die sprachliche Formulierung der Items	71
2.3.3	Die Zusammenstellung des Tests	73
2.4	Datenerhebung	77
2.4.1	Stichprobenprobleme	77
2.4.2	Durchführungsprobleme	80
2.5	Kodierung der Antworten	83
2.5.1	Die Signierung freier Antworten	84
2.5.2	Die Kodierung von Antwortkategorien	88
3	TESTMODELLE	94
3.1	Modelle für dichotome Itemantworten	94
3.1.1	Modelle mit quantitativer Personenvariable	100
	3.1.1.1 Stufenförmige Itemfunktionen	104
X	3.1.1.1.1 Die Guttman-Skala: der Sprung von Null auf Eins	104
	3.1.1.1.2 Antwortfehlermodelle: Irrtum und Raten	109
	3.1.1.2 Kontinuierlich ansteigende Itemfunktionen	112
	3.1.1.2.1 Das Binomialmodell: eine Gerade als Itemfunktion	113
X	3.1.1.2.2 Das Rasch-Modell: parallele Itemfunktionen	120
X	3.1.1.2.3 Item Response Theorie (IRT): Rate- und Trennschärfeparameter	134
	3.1.1.2.4 Die Mokken-Analyse: unbekannte Itemfunktionen	136
	3.1.1.3 Nichtmonotone eingipflige Itemfunktionen	138
	3.1.1.3.1 Das Parallelogramm-Modell: kastenförmige Itemfunktionen	140
	3.1.1.3.2 Kontinuierliche, eingipflige Itemfunktionen	143
3.1.2	Modelle mit qualitativer Personenvariable	149
	3.1.2.1 Deterministische Klassen: verbotene Antwortmuster	151
X	3.1.2.2 Die Analyse latenter Klassen: wahrscheinliche Antwortmuster	153
	3.1.2.3 Das Fixieren und Gleichsetzen von Parametern	159
	3.1.2.4 Lokalisierte Klassen: Punkte auf einem Kontinuum	165
	3.1.3 Das mixed Rasch-Modell: klassifizieren und quantifizieren zugleich	169

3.2	Modelle für nominale Itemantworten	178
3.2.1	Klassenanalyse nominaler Daten	180
3.2.2	Das mehrdimensionale Rasch-Modell	184
3.3	Modelle für ordinale Itemantworten	194
✕ 3.3.1	Das ordinale Rasch-Modell	196
3.3.2	Modelle für Ratingskalen	209
3.3.3	Klassenanalyse ordinaler Daten	219
3.3.4	Klassenmodelle für Ratingskalen	228
3.3.5	Mixed Rasch-Modelle für ordinale Daten	236
3.4	Itemkomponentenmodelle: Modelle für systematisch konstruierte Items	245
3.4.1	Linear-logistische Testmodelle: Komponenten der Aufgabenschwierigkeit	246
3.4.2	Mehrdimensionale Komponentenmodelle	252
3.4.3	Linear-logistische Klassenanalyse	255
3.5.	Modelle der Veränderungsmessung	259
3.5.1	Klassische Probleme der Veränderungsmessung	260
3.5.1.1	Die Reliabilität von Differenzwerten	261
3.5.1.2	Die Korrelation von Anfangswert und Differenzwert	264
3.5.1.3	Messen Vor- und Nachtest dasselbe?	267
3.5.2	Dreifaktorielle Testmodelle: Personen, Items und Zeitpunkte	270
3.5.3	Dynamische Modelle: Lernen während der Testbearbeitung	277
3.5.3.1	Personenspezifisches Lernen	277
3.5.3.2	Itemspezifisches Lernen	279
3.5.3.3	Globales reaktionskontingentes Lernen	281
3.5.4	Die Messung der Wirksamkeit von Maßnahmen	284
4.	PARAMETERSCHÄTZUNG	292
✕ 4.1	Die Likelihoodfunktion	294
✕ 4.2	Die Suche nach dem Maximum	298
✕ 4.2.1	Parameterschätzung für das dichotome Rasch-Modell	300
✕ 4.2.2	Parameterschätzung für die dichotome Klassenanalyse	309
✕ 4.3	Die Eindeutigkeit der Parameterschätzungen	315
✕ 4.4	Die Genauigkeit der Parameterschätzungen	320

✕ (5.)	MODELLGELTUNGSTESTS	324
5.1	Modellvergleiche anhand der Likelihood	325
5.1.1	Informationstheoretische Maße	328
5.1.2	Likelihoodquotiententests	330
5.2	Reproduzierbarkeit der Patternhäufigkeiten	335
✕ 5.3	Die Prüfung einzelner Modellannahmen	340
✕ 5.3.1	Prüfung der Personenhomogenität	340
✕ 5.3.2	Prüfung der Itemhomogenität	345
6.	TESTOPTIMIERUNG	349
6.1	Optimierung der Meßgenauigkeit eines Tests	350
6.1.1	Meßgenauigkeit der Personenmeßwerte	351
6.1.2	Reliabilitätssteigerung durch Testverlängerung	355
6.1.3	Berechnung von Vertrauensintervallen	357
6.1.4	Erhöhung der Zuordnungssicherheit	361
6.2	Optimierung durch Itemselektion	363
6.2.1	Itemselektion bei quantitativen Modellen	363
6.2.2	Itemselektion bei klassifizierenden Modellen	373
6.2.3	Die Identifizierung eindimensionaler Itemgruppen	376
6.3	Optimierung durch Personenselektion	381
6.3.1	Abweichende Antwortmuster	382
6.3.2	Unskalierbare Personengruppen	386
6.4	Optimierung der externen Validität	390
6.4.1	Die Berechnung der externen Validität	390
6.4.2	Maximal erreichbare Validitäten	394
6.4.3	Das Reliabilitäts-Validitäts-Dilemma	397
6.5	Die Normierung von Tests	401
7.	LITERATURVERZEICHNIS	405
	ANHANG:	
	Lösungen der Übungsaufgaben	416
	Chi-quadrat Tabelle	422
	Notationstabelle	423
	Stichwortverzeichnis	426
	Anforderungsschein für PC-Programm	