Psychologische Testtheorie

Sitzung 14

Einzelfalldiagnostik II

Organisatorisches zur Modulprüfung I

- Fragestunde: **Mittwoch, 05.02., 10:15-11:45** ➛ Zoom-Link im Moodle!
- Modulprüfung: **Mittwoch, 12.02., 16:00-17:30**
  - Seien Sie bitte pünktlich!
  - Prüfungsräume und Raumaufteilung: Zwei Räume im Hauptgebäude am Geschwister-Scholl-Platz 1
    - **N120**: Nachnamen mit Anfangsbuchstaben A bis M
    - **A140**: Nachnahmen mit Anfangsbuchstaben N bis Z
  - Erlaubtes Hilfsmittel: Taschenrechner
  - Achtung: Smartwatches sind während der Klausur nicht erlaubt!
  - Achtung: Verwenden Sie bitte dokumentenechte Stifte (d.h., kein Bleistift)!

Organisatorisches zur Modulprüfung II

- Was ist klausurrelevant? → Siehe Syllabus!
  - „Alle Folien und ~~Videos~~ Tutorials aus Vorlesung und UK sind klausurrelevant, **außer diejenigen mit einem roten Punkt** im oberen rechten Eck. Alle in den Folien angegebenen Formeln müssen auswendig gekonnt werden, außer diejenigen mit einem grünen Punkt, denn diese werden in der Klausur angegeben."
- Wie sehen die Klausuraufgaben aus? → Siehe Syllabus!
  - „Die Klausur wird ungefähr 30 Aussagen enthalten, welche als ‚wahr' oder ‚falsch' beurteilt werden müssen. Für jede richtige Einschätzung erhalten Sie einen Punkt, sodass bei 30 Aussagen maximal 30 Punkte (= alles richtig) und minimal 0 Punkte (= alles falsch) erzielt werden können."
  - „Die Aussagen werden unterschiedliche Anforderungsbereiche von reiner Reproduktion über Transfer bis hin zur Anwendung abdecken. Sie können sich beispielsweise auch auf die Interpretation vorgegebener R Outputs oder die Berechnung relevanter Parameter beziehen."

Organisatorisches zur Modulprüfung III

- Wie ist die Klausur aufgebaut? → Siehe Syllabus und Beispiel im Moodle!
  - „Die Klausur besteht aus a) einem Aufgabenkatalog inkl. Deckblatt (ca. 5-8 Seiten) und b) einem Antwortbogen (1 Seite). Es muss nur der Antwortbogen abgegeben werden - alle anderen Blätter der Klausur dürfen Sie mitnehmen. Auf dem Antwortbogen müssen der volle Name und die Matrikelnummer gut leserlich eingetragen werden.“ → Jede Klausur enthält 2 Antwortbögen, von denen einer abgegeben werden muss (der andere dient als Backup, falls Ihnen beim Ausfüllen ein Fehler unterläuft). Die Antwortbögen sind an den Aufgabenkatalog getackert und müssen von Ihnen vor der Abgabe selbstständig herausgelöst werden!
  - „Bitte beachten Sie, dass die Wahr-Falsch-Aussagen nur im Aufgabenkatalog stehen und nicht auf dem Antwortbogen. D.h., Sie müssen Sie müssen beide Bögen bei der Bearbeitung der Klausur nebeneinanderlegen um Ihre Antworten direkt bzw. am Ende der Klausur in den Antwortbogen zu übertragen.“→ Sie sind selbst dafür verantwortlich, dass Sie ihre Antworten vor Ende der Bearbeitungszeit in den Antwortbogen übertragen. Rechnen Sie also genügend Zeit dafür ein!

Organisatorisches zur Modulprüfung IV

- Wie ist die Klausur aufgebaut? → Siehe Syllabus und Beispiel im Moodle!
  - Jede Aufgabe besteht aus einem **Block von zwei bis fünf Aussagen**, die sich entweder auf ein gemeinsames Thema aus der Lehrveranstaltung oder auf eine gemeinsame Aufgabenstellung (ggf. mit R-Output) beziehen.
  - Auf dem Antwortbogen setzen Sie bei jeder Aufgabe **für jede richtige Aussage ein Kreuz** in das entsprechende Kästchen. Für jede falsche Aussage lassen Sie das entsprechende Kästchen frei. Es können keine bis alle Aussagen pro Aufgabe richtig sein.
  - Achten Sie darauf, dass **eindeutig** erkennbar ist, in welchen Kästchen Sie Ihre Kreuze gesetzt haben!
  - Achtung: Jede einzelne Aussage wird immer mit entweder 1 Punkt (Aussage korrekt beurteilt) oder mit 0 Punkten (Aussage nicht korrekt beurteilt) bewertet. Es gibt keinen Punktabzug für falsche Antworten innerhalb einer Aufgabe.

Wahr-oder-Falsch Aussagen zur letzten Sitzung

1. Um den latenten Variablenwert $\hat{\theta}_{Person}$ einer festen Person schätzen zu können müssen wir das Messmodell des verwendeten Tests kennen.
2. Die Breite des modellbasierten Konfidenzintervalls um $\hat{\theta}_{Person}$ hängt der Reliabilität des Tests ab.

Wahr-oder-Falsch Aussagen zur letzten Sitzung

3. Eine Person hat vier Reaktionszeititems bearbeitet und dabei folgende Werte erzielt: x_1 = 287, x_2 = 258, x_3 = 272, x_4 = 280. Für den verwendeten Reaktionszeittests gilt das essentiell parallele Modell und in dessen Handbuch finden Sie folgenden Output:

Gegeben der drei Formeln unten ergibt sich für diese Person ein modell-basiertes 95%-Konfidenzintervall von [-39.22; 31.54].

```
Intercepts:
                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
   .almd1            288.691    5.695   50.694    0.000  288.691    4.347
   .almd2            269.739    5.695   47.366    0.000  269.739    4.062
   .almd3            266.092    5.695   46.725    0.000  266.092    4.007
   .almd4            287.835    5.695   50.543    0.000  287.835    4.334
    alertness          0.000                               0.000    0.000

Variances:
                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
   .almd1     (var) 1303.854   91.288   14.283    0.000 1303.854    0.296
   .almd2     (var) 1303.854   91.288   14.283    0.000 1303.854    0.296
   .almd3     (var) 1303.854   91.288   14.283    0.000 1303.854    0.296
   .almd4     (var) 1303.854   91.288   14.283    0.000 1303.854    0.296
    alertnss        3106.731  416.900    7.452    0.000    1.000    1.000
```

Formelsammlung:

$$\mathrm{SE}\left(\hat{\theta}_{Person}\right) = \sqrt{\frac{VAR(\varepsilon_i)}{k}}$$

$$\hat{\theta}_{Person} = \frac{1}{k}\sum_{i=1}^{k} X_{iPerson} - \frac{1}{k}\sum_{i=1}^{k}\sigma_i$$

$$I_{Person} = \left[\hat{\theta}_{Person} \pm 1.96 \cdot SE\left(\hat{\theta}_{Person}\right)\right]$$

Überblick

Sitzung	Datum	Thema	Themenblock
1	14.10.24	Einführung	Begriffe, Modellierung von Antwortverhalten durch Zufallsvariablen & mathematische Grundlagen der Testtheorie
2	21.10.24	Wahrscheinlichkeitstheoret. Grundlagen	
3	28.10.24	Testtheoretische Modelle I	Testtheoretische Modelle
4	04.11.24	Testtheoretische Modelle II	
5	11.11.24	Testtheoretische Modelle III	
6	18.11.24	Skalierung I	Gütekriterien psychologischer Tests
7	25.11.24	Skalierung II	
8	02.12.24	Faktorenanalyse I	
9	09.12.24	Faktorenanalyse II	
10	16.12.24	Reliabilität I	
11	23.12.24	Sitzung entfällt wegen Weihnachten!	
	30.12.24	Offizielle Winterpause	
	06.01.25		
12	13.01.25	Reliabilität II	Gütekriterien psychologischer Tests
13	20.01.25	Validität	
14	27.01.25	Einzelfalldiagnostik I	Methoden der Einzelfalldiagnostik
15	03.02.25	Einzelfalldiagnostik II	

2.5. τ –kongenerisches Modell

Schätzfunktion I

- Eine feste Person antwortet auf k Items, von denen wir wissen, dass Sie einem τ–kongenerischen Modell folgen
- Die Itemantworten der Person können als Zufallsvariablen $X_{iPerson}$ aufgefasst werden
- Im τ–kongenerischen Modell gilt auf der Ebene der festen Person:

$$X_{iPerson} = \sigma_i + \beta_i \cdot \theta_{Person} + \varepsilon_{iPerson}$$
$$COV\left(\varepsilon_{iPerson}, \varepsilon_{jPerson}\right) = 0$$

- Gesucht wird ein Konfidenzintervall für den Parameter θ_{Person}

Schätzfunktion II

- Im τ–kongenerischen Modell vereinfacht sich die Schätzfunktion, wenn wir statt der normalen Itemantworten $X_{iPerson}$ die **z-standardisierten Itemantworten** $Z_{iPerson}$ betrachten:

$$Z_{iPerson} = \frac{X_{iPerson} - E(X_i)}{\sqrt{VAR(X_i)}} = \frac{X_{iPerson} - \sigma_i}{\sqrt{VAR(X_i)}}$$

- Schätzwerte für σ_i und $VAR(X_i)$ erhält man aus der Normstichprobe

Schätzfunktion III

- Als erwartungstreue und effiziente Schätzfunktion für θ_{Person} ergibt sich im essentiell τ–kongenerischen Modell dann:

$$\hat{\theta}_{Person} = \frac{\sum_{i=1}^{k} \frac{\beta_{zi} \cdot Z_{iPerson}}{1 - \beta_{zi}^2}}{\sum_{i=1}^{k} \frac{\beta_{zi}^2}{1 - \beta_{zi}^2}}$$

- Der Zähler der Schätzfunktion eine mit den standardisierten Steigungsparametern gewichtete Summe der z-standardisierten Itemantworten der Person: Je geringer der standardisierte Steigungsparameter eines Items, desto weniger geht die Itemantwort der Person auf diesem Item in die Schätzung mit ein
- Der Nenner sorgt wieder dafür, dass die Schätzfunktion erwartungstreu ist
- Schätzwerte für β_{zi} erhalten wir aus der Normstichprobe

Standardfehler

- Als Standardmessfehler ergibt sich im τ–kongenerischen Modell:

$$SE\left(\hat{\theta}_{Person}\right) = \frac{1}{\sqrt{\sum_{i=1}^{k} \frac{\beta_{zi}^2}{1 - \beta_{zi}^2}}}$$

- Schätzwerte für die standardisierten Steigungsparameter β_{zi} aller Items erhalten wir aus der Normstichprobe

Konfidenzintervall

- Als zweiseitiges Konfidenzintervall für θ_{Person} ergibt sich somit im τ–kongenerischen Modell:

$$I_{Person} = \left[\hat{\theta}_{Person} \pm z_{1-\frac{\alpha}{2}} \cdot SE\left(\hat{\theta}_{Person}\right)\right]$$

$$= \left[\frac{\sum_{i=1}^{k} \frac{\beta_{zi} \cdot Z_{iPerson}}{1-\beta_{zi}^2}}{\sum_{i=1}^{k} \frac{\beta_{zi}^2}{1-\beta_{zi}^2}} \pm z_{1-\frac{\alpha}{2}} \cdot \frac{1}{\sqrt{\sum_{i=1}^{k} \frac{\beta_{zi}^2}{1-\beta_{zi}^2}}}\right]$$

Beispiel I

- Vier τ–kongenerische Items ($k = 4$)
- Schätzwert β_{zi} aus der Normstichprobe:

```
Latent Variables:
                   Estimate   Std.Err    z-value  p(>IzI)   Std.lv   Std.all
      .item1          2.056     0.046     44.246    0.000    2.056     0.998
      .item2          0.556     0.025     21.866    0.000    0.556     0.622
      .item3          1.058     0.039     26.935    0.000    1.058     0.731
      .item4          1.529     0.036     42.747    0.000    1.529     0.980
```

Beispiel II

- Damit können wir den Schätzwert für den Standardmessfehler $SE(\hat{\theta}_{Person})$ berechnen:

$$\widehat{SE}(\hat{\theta}_{Person}) = \frac{1}{\sqrt{\sum_{i=1}^{k} \frac{\beta_{zi}^2}{1-\beta_{zi}^2}}}$$

$$= \frac{1}{\sqrt{\frac{0.998^2}{1-0.998^2} + \frac{0.622^2}{1-0.622^2} + \frac{0.731^2}{1-0.731^2} + \frac{0.980^2}{1-0.980^2}}} \approx 0.06$$

Beispiel III

- Die Person hat auf die vier Items folgende z-standardisierte Itemantworten gegeben:

$$z_{1Person} = 1.3$$
$$z_{2Person} = 2.0$$
$$z_{3Person} = 1.1$$
$$z_{4Person} = 1.5$$

- Der Schätzwert für ihren Wert auf der latenten Variable ist somit:

$$\hat{\theta}_{Person} = \frac{\sum_{i=1}^{k} \frac{\beta_{zi} \cdot z_{iPerson}}{1-\beta_{zi}^2}}{\sum_{i=1}^{k} \frac{\beta_{zi}^2}{1-\beta_{zi}^2}} = \frac{\frac{0.998 \cdot 1.3}{1-0.998^2} + \frac{0.622 \cdot 2}{1-0.622^2} + \frac{0.731 \cdot 1.1}{1-0.731^2} + \frac{0.980 \cdot 1.5}{1-0.980^2}}{\frac{0.998^2}{1-0.998^2} + \frac{0.622^2}{1-0.622^2} + \frac{0.731^2}{1-0.731^2} + \frac{0.980^2}{1-0.980^2}} \approx 1.33$$

Beispiel IV

- Als zweiseitiges 95 Prozent Konfidenzintervall für ihren latenten Variablenwert θ_{Person} ergibt sich

$$I_{Person} = \left[\hat{\theta}_{Person} \pm z_{1-\frac{\alpha}{2}} \cdot \widehat{SE}\left(\hat{\theta}_{Person}\right)\right]$$

$$= [1.33 \ \pm 1.96 \cdot 0.06] =$$

$$= [1.212; 1.448 \ ]$$

- Die plausiblen Werte für den latenten Variablenwert θ_{Person} liegen also zwischen 1.212 und 1.488

2.6. Mehrdimensionales τ –kongenerisches Modell

Mehrdimensionales τ –kongenerisches Modell

- Besprechen wir in dieser Vorlesung nicht, da sehr kompliziert
- Falls eine Einfachstruktur vorliegt, sollte der Test auf der Basis dieser Einfachstruktur in mehrere eindimensionale τ–kongenerische Modelle aufgeteilt werden und die Einzelfalldiagnostik auf dieser Basis durchgeführt werden

2.7. Übersicht

Übersicht

Modell	$\hat{\theta}_{Person}$	$SE(\hat{\theta}_{Person})$
parallel	$\frac{1}{k}\sum_{i=1}^{k} X_{iPerson}$	$\sqrt{\frac{VAR(\varepsilon_i)}{k}}$
essentiell parallel	$\frac{1}{k}\sum_{i=1}^{k} X_{iPerson} - \frac{1}{k}\sum_{i=1}^{k} \sigma_i$	$\sqrt{\frac{VAR(\varepsilon_i)}{k}}$
τ–äquivalent	$\frac{\sum_{i=1}^{k} \frac{X_{iPerson}}{VAR(\varepsilon_i)}}{\sum_{i=1}^{k} \frac{1}{VAR(\varepsilon_i)}}$	$\frac{1}{\sqrt{\sum_{i=1}^{k} \frac{1}{VAR(\varepsilon_i)}}}$
essentiell τ–äquivalent	$\frac{\sum_{i=1}^{k} \frac{X_{iPerson} - \sigma_i}{VAR(\varepsilon_i)}}{\sum_{i=1}^{k} \frac{1}{VAR(\varepsilon_i)}}$	$\frac{1}{\sqrt{\sum_{i=1}^{k} \frac{1}{VAR(\varepsilon_i)}}}$
τ–kongenerisch	$\frac{\sum_{i=1}^{k} \frac{\beta_{zi} \cdot Z_{iPerson}}{1-\beta_{zi}^2}}{\sum_{i=1}^{k} \frac{\beta_{zi}^2}{1-\beta_{zi}^2}}$	$\frac{1}{\sqrt{\sum_{i=1}^{k} \frac{\beta_{zi}^2}{1-\beta_{zi}^2}}}$

2.8. Approximative Konfidenzintervalle

Approximative Konfidenzintervalle I

- In vielen Testmanualen werden weder Schätzwerte für den Standardmessfehler noch für die Modell-Parameter aus denen man ihn berechnen könnte angegeben. Die bisher vorgestellte modellbasierte Konstruktion eines Konfidenzintervalls für den latenten Variablenwert einer Person ist dann **nicht möglich!**
- Zudem wird in den meisten Tests ohne Berücksichtigung der testtheoretischen Modelle als Schätzfunktion für den latenten Variablenwert einfach der Summenwert oder der Mittelwert der Items verwendet
- In diesem Fall kann ein Konfidenzintervall auf der Basis folgender approximativer Formel für den Standardmessfehler konstruiert werden:

$$SE\left(\hat{\theta}_{Person}\right) = \sqrt{VAR(X) \cdot \left(1 - Rel(X)\right)}$$

- Hierbei ist X entweder die Itemsumme oder der Itemmittelwert oder der Normwert und für $Rel(X)$ kann dementsprechend ein Schätzwert für die Reliabilität dieser Größen verwendet werden (z.B. Schätzwert für Cronbach α)

Approximative Konfidenzintervalle II

- Wichtig: Die approximativen Konfidenzintervalle stellen eine **Notlösung** dar, falls die modellbasierten Konfidenzintervalle nicht berechnet werden können
- In vielen Fällen sind sie deutlich breiter als die modellbasierten Intervalle

Vergleich: Reliabilität des Summenwerts vs. Standardmessfehler

- **Für die Einzelfalldiagnostik ist von entscheidender praktischer Bedeutung, mit welcher Genauigkeit der Wert einer Person auf der latenten Variable mithilfe eines vorliegenden psychologischen Tests geschätzt wird!**
  - Der Standardmessfehlerfehler ist hierfür ein **direktes Maß**, wie man an den Formeln der Konfidenzintervalle für den Parameter θ_{Person} sehen kann
  - Die Reliabilität des Summenwerts (oder Mittelwerts) ist hierfür nur ein **indirektes Maß**, was sich darin äußert, dass die mithilfe der Reliabilität des Summenwerts berechneten Konfidenzintervalle **nur approximativ** gelten

3. Interpretation und Normwerte

Normorientierte Interpretation des Konfidenzintervalls I

- Im Beispiel zum τ–kongenerischen Modell hatten wir als Konfidenzintervall für den latenten Variablenwert θ_{Person} das Intervall $[1.212; 1.448]$ berechnet
- Wir können die Werte im Intervall nun in Beziehung zu den latenten Variablenwerten in der Population setzen (= normorientierte Auswertung)
- Im τ–*kongenerischen Modell* wissen wir, dass aufgrund der Normierung $E(\theta) = 0$ und $VAR(\theta) = \sqrt{VAR(\theta)} = 1$ gilt. Der Mittelwert der latenten Variablenwerte in der Population ist also 0 und die Standardabweichung 1
- In den *übrigen Modellen* ergibt sich $VAR(\theta)$ (und in den nicht-essentiellen Modellen auch $E(\theta)$) nicht aus der Normierung, kann aber aus der Normstichprobe geschätzt werden

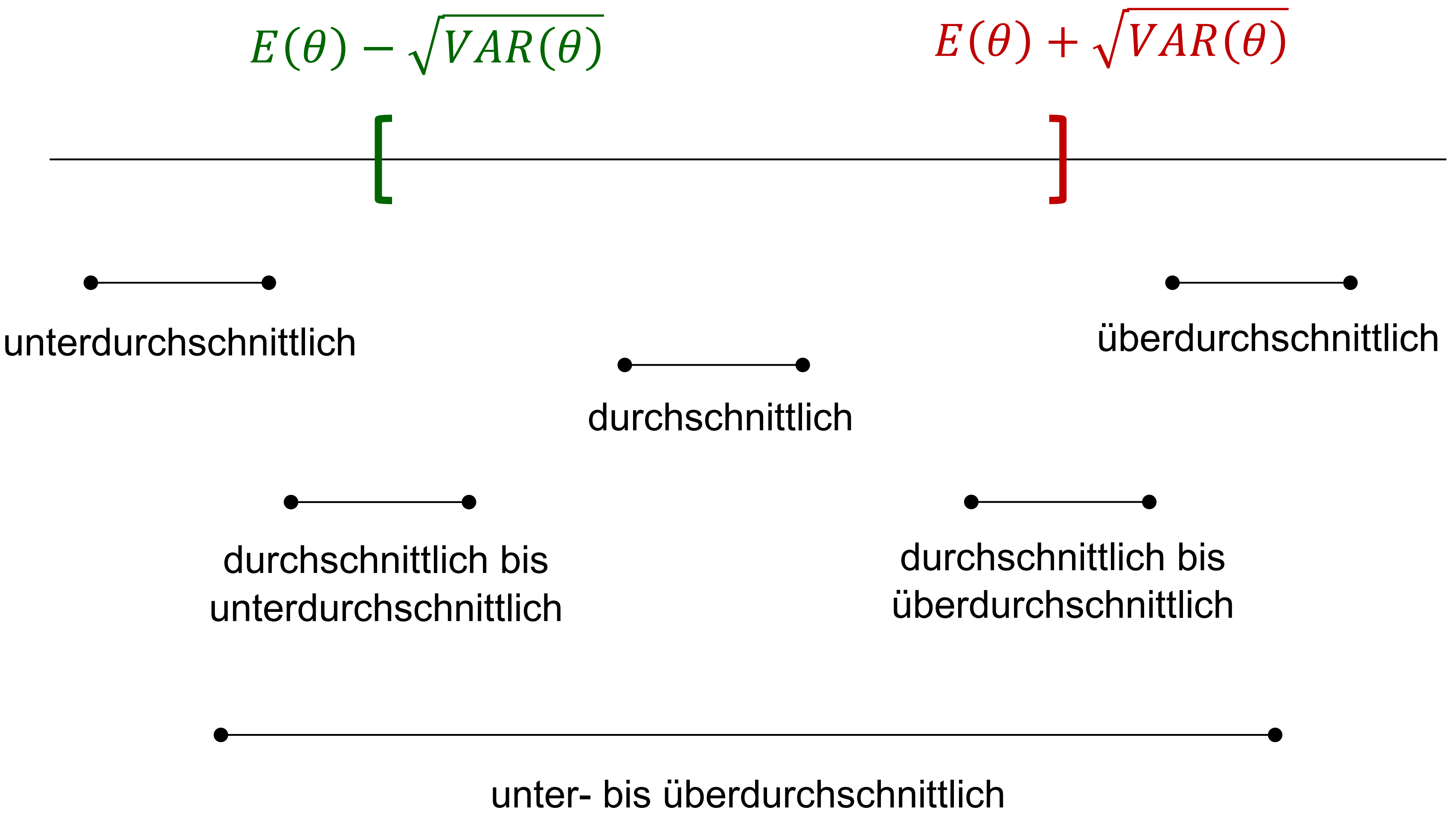
Normorientierte Interpretation des Konfidenzintervalls II

- Wir können nun das Konfidenzintervall um den latenten Variablenwert einer Person zu diesen Populations-Werten (Mittelwert und Standardabweichung der latenten Variable) in Beziehung setzten:
  - Liegt das gesamte Intervall mehr als eine Standardabweichung unter $E(\theta)$ (dem Mittelwert in der Population), gehen wir davon aus, dass die Person einen **unterdurchschnittlichen** latenten Variablenwert aufweist
  - Liegt das Intervall zwischen einer Standardabweichung unter $E(\theta)$ und einer Standardabweichung über $E(\theta)$, gehen wir davon aus, dass die Person einen **durchschnittlichen** latenten Variablenwert aufweist
  - Liegt das gesamte Intervall mehr als eine Standardabweichung über $E(\theta)$, gehen wir davon aus, dass die Person einen **überdurchschnittlichen** latenten Variablenwert aufweist

Normorientierte Interpretation des Konfidenzintervalls III

- Wir können nun das Konfidenzintervall um den latenten Variablenwert einer Person zu diesen Populations-Werten (Mittelwert und Standardabweichung der latenten Variable) in Beziehung setzten:
  - Liegt das Intervall nicht genau in einem dieser drei Bereiche, sondern über einer der durch die Standardabweichungen definierten Grenzen, gehen wir davon aus, dass die Person einen **durchschnittlichen bis unterdurchschnittlichen** bzw. einen **durchschnittlichen bis überdurchschnittlichen** latenten Variablenwert aufweist (je nachdem auf welcher Grenze das Konfidenzintervall liegt)

Normorientierte Interpretation des Konfidenzintervalls IV



Normorientierte Interpretation des Konfidenzintervalls V

- Achtung: Ist im Intervall der **Mittelwert** der Normstichprobe enthalten, ist der (geschätzte) Populationsmittelwert ein plausibler Wert für diese Person. D.h., der entsprechende Hypothesentest (siehe Kapitel 4) würde sich für eine ungerichtete Nullhypothese entscheiden, die besagt, dass der Wert der Person genau dem Mittelwert der Normstichprobe entspricht.

Beispiel

- Im Beispiel zum τ–kongenerischen Modell hatten wir als Konfidenzintervall für den latenten Variablenwert θ_{Person} das Intervall $[1.212; 1.448]$ berechnet
- Der Wert eine Standardabweichung über dem Mittelwert in der Population ist aufgrund der Normierung $E(\theta) + \sqrt{VAR(\theta)} = 0 + 1 = 1$
- Das Intervall in unserem Beispiel liegt über diesem Wert
- Wir gehen somit davon aus, dass die Person einen überdurchschnittlichen latenten Variablenwert aufweist

Normwerte I

- Um nicht für jeden psychologischen Test andere Referenzwerte $E(\theta)$ und $VAR(\theta)$ bzw. $\sqrt{VAR(\theta)}$ verwenden zu müssen für die Interpretation, werden die Schätzwerte für die latenten Variablenwerte oft in sogenannte **Normwerte** umgerechnet
- Beispielsweise können wir die Schätzwerte in IQ-Werte transformieren, so dass wir $E(\theta_{Norm}) = 100$ und $\sqrt{VAR(\theta_{Norm})} = 15$ erhalten
- Vorsicht: Die Umrechnung der Schätzwerte in Normwerte entspricht einer **Transformation** der Schätzfunktion für die latenten Variablenwerte und **ändert auch den Standardmessfehler** zur Konstruktion des Konfidenzintervalls
- Am einfachsten ist es daher, zunächst ein Konfidenzintervall auf der Basis der normalen Schätzwerte und des regulären Standardmessfehlers zu konstruieren und dann die Grenzen des Intervalls in die Normwerte zu transformieren (siehe nächste Folie)

Normwerte II

- Bei der Umrechnung in die Normwerte werden die Schätzwerte zunächst z-standardisiert, sodass danach $E(\theta_z) = 0$ und $VAR(\theta_z) = 1$ gilt (im τ–kongenerischen Modell ist dies durch die Normierung schon gegeben):

$$\hat{\theta}_{Person,z-Wert} = \frac{\hat{\theta}_{Person} - E(\theta)}{\sqrt{VAR(\theta)}}$$

- Werte für $E(\theta)$ und $VAR(\theta)$ erhalten wir entweder aus der Normierung oder als Schätzwerte aus der Normstichprobe

Normwerte III

- Diese standardisierten Schätzwerte können wir dann wie folgt in Normwerte umrechnen:

$$\hat{\theta}_{Person,Normwert} = E(\theta_{Norm}) + STD(\theta_{Norm}) \cdot \hat{\theta}_{Person,z-Wert}$$

- Es gibt verschiedene Konventionen für die Wahl von $E(\theta_{Norm})$ und $STD(\theta_{Norm})$:

	$E(\theta_{Norm})$	$STD(\theta_{Norm})$
Z-Werte	0	1
IQ-Werte	100	15
Standardwerte	100	10
T-Werte	50	10
Stanine	5	2

Beispiel I

- Im Beispiel zum τ–kongenerischen Modell hatten wir als Konfidenzintervall für den latenten Variablenwert θ_{Person} das Intervall $[1.212; 1.448]$ berechnen

→ **Wir transformieren nun die Konfidenzintervallgrenzen in Normwerte!**

- Schritt 1: Umrechnung der beiden Intervallgrenzen in **z-standardisierte Werte**
  - Wir wissen, dass im τ–kongenerischen Modell wegen der Normierung schon $E(\theta) = 0$ und $VAR(\theta) = \sqrt{VAR(\theta)} = 1$ gilt
  - Eine Umrechnung der beiden Intervallgrenzen in z-standardisierte Werte ist deswegen nicht nötig (bzw. würde sich die Werte nicht ändern, da in diesem Fall $\hat{\theta}_{Person} = \hat{\theta}_{Person,z-Wert}$ gilt)

Beispiel II

- Im Beispiel zum τ–kongenerischen Modell hatten wir als Konfidenzintervall für den latenten Variablenwert θ_{Person} das Intervall $[1.212; 1.448]$ berechnen

→ **Wir transformieren nun die Konfidenzintervallgrenzen in Normwerte!**

- Schritt 2: Umrechnung der beiden Intervallgrenzen in **Normwerte (hier: IQ-Werte)**
  - untere Grenze:

$$E(\theta_{Norm}) + STD(\theta_{Norm}) \cdot 1.212 = 100 + 15 \cdot 1.212 = 118.18$$

  - obere Grenze:

$$E(\theta_{Norm}) + STD(\theta_{Norm}) \cdot 1.448 = 100 + 15 \cdot 1.448 = 121.72$$

Beispiel III

- Im Beispiel zum τ–kongenerischen Modell hatten wir als Konfidenzintervall für den latenten Variablenwert θ_{Person} das Intervall $[1.212; 1.448]$ berechnen

→ **Wir transformieren nun die Konfidenzintervallgrenzen in Normwerte!**

- Schritt 3: **Interpretation** des Konfidenzintervalls
  - Das Konfidenzintervall auf der Basis der IQ-Werte ist somit $[118.18; 121.72]$
  - Bei der Interpretation dieses Konfidenzintervalls müssen natürlich die neuen Referenzgrößen $E(\theta_{Norm}) = 100$ und $\sqrt{Var(\theta_{Norm})} = STD(\theta_{Norm}) = 15$ berücksichtigt werden
  - **Wichtig:** Die Umrechnung in Normwerte kann die Interpretation nicht ändern. Eine Person kann z.B. nicht vor der Umrechnung unterdurchschnittlich und danach durchschnittlich sein. Die Umrechnung in Normwerte erleichtert lediglich die Interpretation

Prozentränge

- Für die Konfidenzintervalle haben wir bereits angenommen, dass die Itemantworten $X_{iPerson}$ der festen Person unabhängig und normalverteilt sind. Wenn wir auch annehmen, dass die Normwerte in der Population normalverteilt sind, können wir Normwerte in Prozentränge umrechnen.

- Der Prozentrang einer Person mit Normwert $\hat{\theta}_{Person,Normwert}$ entspricht $P(\theta_{Norm} \leq \hat{\theta}_{Person,Normwert}) = F(\hat{\theta}_{Person,Normwert})$ also dem Wert der Verteilungsfunktion einer Normalverteilung mit Erwartungswert $E(\theta_{Norm})$ und Standardabweichung $STD(\theta_{Norm})$.

- Beispiel in R: Für einen IQ-Wert von 110 ergibt sich ein Prozentrang von ca. 75

```
> pnorm(110, mean = 100, sd = 15)
[1] 0.7475075
```

- Interpretation: „Ungefähr 75% der Personen in der Population haben den gleichen oder einen niedrigeren Wert auf der latenten Variable“

4. Hypothesentests

Fragestellung

- Neben der Berechnung von Konfidenzintervallen für den latenten Variablenwert θ_{Person} von Personen ist es auch möglich, verschiedene Hypothesen bezogen auf den latenten Variablenwert θ_{Person} von Personen mithilfe von statistischen Hypothesen-Tests zu überprüfen
- Beispiele für Fragestellungen:
  - Unterscheidet sich der latente Variablenwert einer Person von einem bestimmten, vorgegebenen Wert?
  - Unterscheiden sich zwei Personen in ihren latenten Variablenwerten?
  - Verändert sich der latente Variablenwert einer Person über die Zeit?

Statistische Hypothesen I

1. Unterscheidet sich der latente Variablenwert einer Person von einem bestimmten, vorgegebenen Wert?

- Statistische Hypothesen im Falle einer ungerichteten Fragestellung:

$$H_0: \theta_{Person} = \theta_0$$

$$H_1: \theta_{Person} \neq \theta_0$$

- Statistische Hypothesen im Falle einer gerichteten Fragestellung (hier z.B. rechtsseitig):

$$H_0: \theta_{Person} \leq \theta_0$$

$$H_1: \theta_{Person} > \theta_0$$

- θ_0 stellt hierbei einen aus inhaltlichen Gründen interessanten Wert dar

Statistische Hypothesen II

2. Unterscheiden sich zwei Personen in ihren latenten Variablenwerten?

- Statistische Hypothesen im Falle einer ungerichteten Fragestellung:

$$H_0: \theta_{Person1} = \theta_{Person2}$$

$$H_1: \theta_{Person1} \neq \theta_{Person2}$$

- Statistische Hypothesen im Falle einer gerichteten Fragestellung (hier z.B. rechtsseitig):

$$H_0: \theta_{Person1} \leq \theta_{Person2}$$

$$H_1: \theta_{Person1} > \theta_{Person2}$$

Statistische Hypothesen III

3. Verändert sich der latente Variablenwert einer Person über die Zeit?

- Statistische Hypothesen im Falle einer ungerichteten Fragestellung:

$$H_0 : \theta_{PersonZeitpunkt1} = \theta_{PersonZeitpunkt2}$$

$$H_1 : \theta_{PersonZeitpunkt1} \neq \theta_{PersonZeitpunkt2}$$

- Statistische Hypothesen im Falle einer gerichteten Fragestellung (hier z.B. rechtsseitig):

$$H_0 : \theta_{PersonZeitpunkt1} \leq \theta_{PersonZeitpunkt2}$$

$$H_1 : \theta_{PersonZeitpunkt1} > \theta_{PersonZeitpunkt2}$$

Teststatistiken I

- Für die Überprüfung der statistischen Hypothesen brauchen wir Teststatistiken
  - Teststatistik für die erste Fragestellung:

$$Z = \frac{\hat{\theta}_{Person} - \theta_0}{SE\left(\hat{\theta}_{Person}\right)}$$

  - Teststatistik für die zweite Fragestellung:

$$Z = \frac{\hat{\theta}_{Person1} - \hat{\theta}_{Person2}}{\sqrt{2} \cdot SE\left(\hat{\theta}_{Person}\right)}$$

  - Teststatistik für die dritte Fragestellung:

$$Z = \frac{\hat{\theta}_{PersonZeitpunkt1} - \hat{\theta}_{PersonZeitpunkt2}}{\sqrt{2} \cdot SE\left(\hat{\theta}_{Person}\right)}$$

Teststatistiken II

- Alle drei Teststatistiken folgen unter der Bedingung, dass die Nullhypothese gilt, einer Standardnormalverteilung, d.h. in allen drei Fällen gilt:

$$Z \sim N(0,1)$$

- Wir können also kritische Bereiche oder p-Werte auf der Basis der Standardnormalverteilung berechnen (siehe z-Test aus Statistik I)
- Die genaue Form von $\hat{\theta}_{Person}$ und $SE(\hat{\theta}_{Person})$ ist wie bei den Konfidenzintervallen abhängig davon, welches testtheoretische Modell gilt

Beispiel I

- **Wir wollen wissen, ob eine Person hochbegabt ist, also ihre Intelligenz höher als zwei Standardabweichungen über dem Populationsdurchschnitt liegt**
  - Wir betrachten einen IQ-Test, der aus vier τ–kongenerischen Items besteht
  - Aufgrund der Normierung wissen wir, dass $E(\theta) = 0$ und $VAR(\theta) = 1$
  - Der Wert zwei Standardabweichungen über dem Populationsdurchschnitt ist somit $E(\theta) + 2 \cdot \sqrt{VAR(\theta)} = 0 + 2 \cdot 1 = 2$

→ Unsere statistischen Hypothesen lauten also:

$$H_0: \theta_{Person} \leq 2$$

$$H_1: \theta_{Person} > 2$$

Beispiel II

- Die Person bearbeitet unseren IQ-Test
- Als Schätzwert für ihren latenten Variablenwert ergibt sich $\hat{\theta}_{Person} = 2.3$
- Der Schätzwert für den Standardmessfehler sei $\widehat{SE}\left(\hat{\theta}_{Person}\right) = 0.3$
- Die Teststatistik für unsere Fragestellung ist

$$Z = \frac{\hat{\theta}_{Person} - \theta_0}{SE\left(\hat{\theta}_{Person}\right)} = \frac{\hat{\theta}_{Person} - 2}{SE\left(\hat{\theta}_{Person}\right)}$$

- Die Realisation dieser Teststatistik ist

$$z = \frac{\hat{\theta}_{Person} - 2}{\widehat{SE}\left(\hat{\theta}_{Person}\right)} = \frac{2.3 - 2}{0.3} = 1$$

Beispiel III

- Der realisierte Wert $z = 1$ liegt bei einem α-Niveau von 0.05 nicht im rechtsseitigen kritischen Bereich $K_z = [1.64; \infty[$ des z-Tests
- Wir entscheiden uns also für die $H_0: \theta_{Person} \leq 2$ und somit gegen die Hochbegabung

Kritische Differenzen

- In Testmanualen sind häufig sogenannte **kritische Differenzen** angegeben
- Prinzip: Ist z.B. die Differenz zwischen den Schätzwerten der latenten Variablenwerte zweier Personen größer als die angegebene kritische Differenz, gehen wir davon aus, dass sich die latenten Variablenwerte der beiden Personen unterscheiden
- Die kritischen Differenzen stellen eine (schlechte) Alternative zu den hier vorgestellten Hypothesentests dar, da sie häufig auf der Basis der oben angesprochenen approximativen Formel für den Standardmessfehler berechnet werden

Das war‘s schon fast für heute…

- *Ausblick:* Im Sommersemester befassen wir uns mit weiteren, eher praktischen Überlegungen zur psychologischen Diagnostik.

- *Aber zuerst:*
  - **Gibt es offene Fragen zur heutigen Vorlesung?**
  - Zur Vertiefung: Übungsblatt 9 im Moodle (Achtung: Aufgaben 5f bis 5h beziehen sich auf den von der Klausurrelevanz ausgenommenen Hypothesentest)