Multiple Regressionsanalyse – Verzerrungen

Dr. Marcel Miché

21. Mai 2025

Info

In der klinischen Psychologie werden fast ausschliesslich "Konstrukte" in den empirischen Analysen verwendet. "Tolerance of uncertainty" oder "trust in institutions" sind nur zwei von sehr vielen Beispielen psychologischer Konstrukte. Ausnahmen sind Variablen wie Alter, biologisches Geschlecht oder andere Repräsentanten aus Biologie und Neurologie, die zweifelsfrei (annähernd) perfekte Messgenauigkeit haben. Alter und biol. Geschlecht haben darüber hinaus sogar perfekte Messgültigkeit (Validität). Andere "Konstrukte" sind hingegen nie perfekt messgenau (= reliabel, häufig gemessen mit Cronbach's alpha). Validität ist eine eigenständige, zusätzliche Herausforderung. Beides zusammen könnte als 'Datenqualität' gelten (z.B. Ghasemaghaei & Calic, 2019 und Cichy & Rass, 2019), jedenfalls für strukturierte Daten (z.B. Cheng et al. , 2024), besser noch 'Datenbedeutsamkeit' (Pedersen et al., 2025).

In empirischen Analysen, in denen Konstrukte als unabhängige sowie abhängige Variablen verwendet werden, könnte man also die fehlende Messgenauigkeit der Konstrukte innerhalb der Datenanalyse berücksichtigen. Strukturgleichungsmodelle bieten diese Möglichkeit, indem man das Messmodel (aus welchen Einzelitems werden die Konstrukte auf welche Weise erzeugt?) simultan mit dem Strukturmodel (die eigentliche multiple Regressionsgleichung) einem empirischen Schätzmechanismus übergibt (siehe Codebeispiel in Abbildung 1).

Abbildung 1: Strukturgleichungsmodel (SEM).

```
m6a <- '
# measurement model
adjust =~ motiv + harm + stabi
risk =~ verbal + ppsych + ses
achieve =~ read + arith + spell
# structual model: multiple linear regression
achieve ~ adjust + risk'
```

In Codebeispiel 2 (siehe Abbildung 2) ist zu sehen wie das SEM auf die Daten angewendet wird, wie das Ergebnis (mit dem summary Befehl) ausgegeben wird, sowie den relevanten Teil des Ergebnisses.

Abbildung 2: Model fitting und Ergebnis.

```
# Fit the model.
fit6a <- sem(m6a, data=dat)
# Show results.
summary(fit6a, standardized=TRUE, fit.measures=FALSE,
rsquare=TRUE)
# Output in R console
# Regressions:
#                  Estimate  Std.Err  z-value  P(>|z|)  Std.all
# achieve ~
# adjust              0.375    0.046    8.085    0.000    0.372
# risk                0.724    0.078    9.253    0.000    0.564
# R-square         Estimate
# achieve             0.653
```

Multiple lineare Regression (in Form eines SEM)

Abbildung 3: Model fitting und Ergebnis.

```
# Fit the model.
fit6aNew <- sem(model='achieve ~ adjust + risk', data=datlm)
# Show results.
summary(fit6aNew, standardized=TRUE, fit.measures=FALSE,
rsquare=TRUE)
# Output in R console
# Regressions:
#                  Estimate  Std.Err  z-value  P(>|z|)  Std.all
# achieve ~
# adjust              0.529    0.037   14.358    0.000    0.502
# risk                0.583    0.066    8.859    0.000    0.310
# R-Square:        Estimate
# achieve             0.419
```

Multiple lineare Regression (konventionell)

Mit 'konventionell' ist der R-Befehl lm (linear model) gemeint. Das lineare Model nutzt statt der Standardnormalverteilung die t-Verteilung für die statistischen Tests jedes geschätzten Modelparameters.

Abbildung 4: Model fitting und Ergebnis.

```
lm.mod <- lm(achieve ~ adjust + risk,
data=data.frame(scale(datlm))); summary(lm.mod)
# Output in R console
#             Estimate Std. Error t value Pr(>|t|)
# adjust        0.5024    0.00351  14.315   <2e-16
# risk          0.3100    0.00351   8.832   <2e-16
# Multiple R-squared:  0.4189,  Adjusted R-squared:  0.4165
```

Fazit 1: Die aufgeklärte Varianz in Abbildungen 3 und 4 ist identisch (0.419).

Fazit 2: Die z-Werte im Output sind approximativ dieselben wie die t-Werte (in Abbildungen 3 und 4), weil bei grossen Stichproben (hier: $N = 500$) diese beiden Verteilungsformen sich einander annähern.

Fazit 3: Die z-Werte in Abbildung 3 sind einmal stärker ($14.358 > 8.085$), einmal schwächer ausgeprägt ($8.859 < 9.253$) als die z-Werte in Abbildung 2.

Wichtig: Es können alle möglichen Mischformen verzerrter Ausprägungen auftreten, wichtig ist also zu wissen, **dass** Verzerrungen (biases) auftreten.

Gesamtfazit: Die SEM Variante (siehe Abbildungen 1 und 2) mit der konventionellen Variante abzugleichen, ist konsistent mit der häufig betonten Zielsetzung explanativer Forschung, Verzerrungen der Regressionsgewichte ('Bias') so gut wie möglich vermeiden zu wollen. Diesen Vergleich (und die dabei auftretenden Verzerrungen) transparent zu machen, würde eine hintergründig stets vorhandene Quelle der Unsicherheit in den Vordergrund rücken. Diese Unsicherheit lautet: Welche von beiden Varianten zeigt das richtige(re) Ergebnis, und vor allem, warum? Solche und ähnliche Unsicherheiten stehen im Konflikt mit der Vorstellung davon, dass ein/e Expert/in Sicherheit liefern sollte. Die Nebenwirkung solcher Vorstellungen und deren Umsetzung in die Realität scheint leider dafür (mit-)verantwortlich zu sein, dass **volle Transparenz** (= **open** science) nur schwer erreicht werden kann, weil hierfür jene (unsinnigen) Vorstellungen über Expert/inn/en korrigiert werden müssen.

Datenherkunft dieses Handouts. Siehe zudem das R-Skript `SEMversusLM1.R`. Sowie: Weiteres R-Skript `SEMversusLM2.R`, dessen Datenherkunft, siehe Skriptzeile 1.

Ein Beispielpaper aus dem Gesundheitsbereich, in dem SEM die Hauptrolle spielt (auf scholar google bisher 601 Mal zitiert; Heute: 21. Mai 2025).