

10 typische Fehler bei der Datenanalyse / Datenauswertung:

1. Schon am Anfang der Datenerhebung, z.B. bei der Konstruktion von Fragebögen, wird der Möglichkeitsraum des Interviewten nicht angemessen berücksichtigt (beispielsweise sind die Antwortmöglichkeiten "weiß ich nicht", "keine Angabe" vorzusehen, nicht nur "ja" und "nein").

2. Bei sogenannten fehlenden Werten (*Missings*) ist zu unterscheiden zwischen nichtinformativen *Missings* (non Response) und informativen *Missings*. Letztere liegen in multivariaten Datensätzen vor, wenn einige Variablen nicht für alle Untersuchungseinheiten zutreffend bzw. existent sind. Ein Beispiel ist die Variable „Schwangerschaftsdauer“, die für männliche Personen bzw. Kleinkinder informative *Missings* aufweist.

3. Berechnen von Statistiken, die nicht dem Skalenniveau der Variablen entsprechen

- Beispiel 1:
Berechnen von Mittelwerten auf Basis von ordinalskalierten Daten mit wenigen Merkmalsausprägungen in den Daten führt zu Verzerrungseffekten und Scheinpräzision
- Beispiel 2:
Berechnen einer Kategorienskala aus metrisch skalierten Daten führt zu Informationsverlust.

4. Errechnen von Pearson'schen Korrelationskoeffizienten an nicht approximativ normalverteilten Daten, insbesondere Daten, die keine eingipflige, sondern zwei- oder mehrgipflige Verteilung aufweisen. Gefahr: Anschein von starken Zusammenhängen beispielsweise aufgrund eines einzigen Ausreißers.

5. Anwendung von inadäquaten Testverfahren in der Datenanalyse - zwei einfache Beispiele:

- Anwenden von parametrischen Testverfahren an Daten, bei welchen die Voraussetzungen für deren Durchführung (Verteilungsform) nicht berücksichtigt wird. (Beispiel: t-Test bei nicht normalverteilten Daten)
- Anwenden von nicht parametrischen Testverfahren an Daten, bei welchen die Voraussetzungen für Durchführung von parametrischen Verfahren nicht berücksichtigt wird. (z.B. Mann-Whitney-U- Test bei normalverteilten Daten). Dies führt zu Informationsverlust.

6. Multiples Testen: Wenn Untersuchungsgruppen hinsichtlich mehr als einer Zielgröße verglichen werden sollen, darf das konventionelle Signifikanzniveau nicht beibehalten werden; es ist zu adjustieren.

7. Anwenden statistischen Methoden, die für prospektive Studien entwickelt wurden, bei retrospektiv erhobenen Daten. Beispiel: Errechnen relativer Risiken aus Fall-Kontrollstudien.

8. Fehlinterpretation des Signifikanztests bzw. der Bedeutung des p-Werts: Dieser gibt nicht Auskunft über die Wahrscheinlichkeit, mit der die Forschungshypothese (H_A oder H_1) zutrifft. Er vermittelt also keine Sicherheit über deren Zutreffen. Er gibt Auskunft über die Wahrscheinlichkeit, sich bei der Ablehnung der entgegengesetzten Nullhypothese zu

täuschen. Dazu kommt oft eine Überbewertung der Bedeutung des Signifikanztests. Dieser gibt inhaltlich Auskunft darüber, ob ein bestimmtes statistisches Ergebnis mit einer bestimmten Wahrscheinlichkeit zufällig zustande gekommen ist, er gibt keine Auskunft über die praktische Bedeutsamkeit der untersuchten Differenz (Effektgröße), die Qualität des Forschungsdesigns, die Validität und Reliabilität der verwendeten Messinstrumente, ein korrekt durchgeführtes Interventionsprogramm oder die Reproduzierbarkeit der Ergebnisse. Ein Signifikanztest ist also nur eines von vielen Kriterien, die bei der Gütebewertung der statistischen Datenanalyse anzulegen sind.¹

9. Fehler und Anomalien in den Daten werden am besten erkannt, wenn man per Hand vorgeht - sie geraten bei automatisierten Auswertungsroutinen gerne aus dem Blick. Daher sind vor Durchführung multivariater statistischer Methoden vorausgehende Analysen absolut notwendig.²

10. Anwendung von statistischen Modellen ohne vorherige Prüfung der wichtigsten Modellvoraussetzungen. Das führt dann oft zu dem Versäumnis, die statistische Aussagekraft einzuschränken, wenn wichtige Voraussetzungen nicht oder nur partiell erfüllt sind.

Literatur: Tabachnik/Fidel (2013): Using multivariate statistics

¹ 3.1.4 Controversy around Significance Testing

² 1.1.4 Garbage in, Roses out