



Die Auswahl der passenden statistischen Methode

von Daniela Keller

Das häufigste und wichtigste Thema in meinen Beratungen und meinen Kursen ist die Frage: Was ist die passende statistische Analysemethode für mein Projekt? Und dies ist zu Recht eine häufig gestellte Frage, denn sie ist nicht pauschal und meist auch nicht kurz und einfach zu beantworten. Die Antwort hängt vom Studiendesign, den Variablentypen, der Stichprobengröße und den Fragestellungen ab und klärt sich häufig erst im Verlauf der Datenanalyse, weil z.B. durch Voranalysen neue Erkenntnisse zu den Forschungsfragen oder auch einfach nur zu den untersuchten Daten gewonnen werden.

Leider kann ich Ihnen diese individuelle Frage also nicht in diesem Artikel beantworten. Ich will Ihnen aber eine Vorgehensweise an die Hand geben, mit der Sie selbst für Ihr Projekt – bzw. für jede Hypothese Ihres Projekts – die passende Methode herausfinden können. Gehen Sie dazu einfach die folgenden Punkte durch:

1. Von der Forschungsfrage zu den statistischen Hypothesen

Starten Sie für die Auswahl der passenden statistischen Methode noch einmal bei der Formulierung Ihrer For-

schungsfrage. Das klingt wie ein Schritt in die falsche Richtung, ist aber wichtig. Denn Sie brauchen richtig formulierte Hypothesen und dazu ist es notwendig, nochmal genau zu überlegen, was in der Forschungsfrage untersucht werden soll.

Aus der Forschungsfrage heraus formulieren Sie sich die Hypothesen. Hier sollten Sie darauf achten, dass Sie die Hypothesen präzise und genau formulieren, dass sie nur eine Idee pro Hypothesenpaar verarbeiten und dass die Hypothesen auch messbar sind. Dass Sie also alle in den Hypothesen verwendeten Parameter auch erhoben haben.

Das statistische Hypothesenpaar besteht aus Null- und Alternativhypothese. Die Alternativhypothese beschreibt das, was Sie nachweisen wollen, z.B. „Es gibt einen Zusammenhang“. Die Nullhypothese ist das Gegenteil der Alternativhypothese, z.B. formuliert „Es gibt keinen Zusammenhang.. Der Signifikanztest später versucht die Nullhypothese abzulehnen und damit die Alternativhypothese zu bestätigen.

2. Das Studiendesign ergibt die Variablen und Variablentypen

Wenn Sie Ihre Hypothesen fix haben, werden Sie sich klar über Ihr verwendetes Studiendesign: Haben Sie mehrere Messzeitpunkte? Wenn ja, welche und wie viele? Werden verschiedene Gruppen verwendet? Wie viele und welche? Welche Parameter werden erhoben?

Sollten Sie Ihre Daten schon vorliegen haben, ist das längst alles schon fix. Trotzdem ist es sinnvoll, noch einmal darüber nachzudenken. Dadurch haben Sie Ihr Studiendesign dann glasklar vor sich liegen, wenn Sie im nächsten Schritt über die passende statistische Methode entscheiden.

Die Beantwortung der oben genannten Fragen führt dazu, dass Sie wissen, welche Variablen Sie mit welchem Messniveau erhoben haben. Außerdem wird klar, ob es bei den vorher formulierten Hypothesen um Gruppenvergleiche oder um Analysen von Messwiederholungen geht. Beides ist entscheidend für die Auswahl der passenden statistischen Methode.

3. Voraussetzungen prüfen

Wenn Sie nun sowohl Ihre Hypothese als auch die dazugehörigen Variablen und das Studiendesign klar vor Augen haben, ergibt sich damit eine Auswahl an möglichen statistischen Methoden. Das sind einerseits deskriptive Methoden, die Sie zum Beschreiben der Daten, der Unterschiede oder Zusammenhänge verwenden. Außerdem Grafiken, mit denen Sie diese Unterschiede oder Zusammenhänge veranschaulichen können.

Auch haben Sie jetzt eine Idee, welche Signifikanztests oder auch welche komplexeren Modelle für Ihre Auswertung in Frage kommen. Beides sind Methoden der schließenden Statistik, mit denen Sie die von Ihnen formulierten Hypothesen auf Signifikanz prüfen. Sie haben nach den oben durchgeführten Schritten nun bereits eine oder mehrere mögliche Methoden ausgewählt. Zusätzlich ist aber wichtig, dass viele dieser Tests und Modelle Voraussetzungen an z.B. die Verteilung der Daten stellen.

Diese Voraussetzungen müssen nun geprüft werden, damit Sie entscheiden können, ob sie die angedachte

Methode einsetzen dürfen. Sie können z.B. die Normalverteilung der Variablen oder auch der Residuen betreffen, eine bestimmte Form des Zusammenhangs (Linearität) zwischen den Parametern oder auch eine genügend große Zellbesetzung bei der Analyse von Kreuztabellen beinhalten. Abhängig vom Ergebnis dieser Prüfung entscheiden Sie dann, ob die bisher ausgewählte Methode möglich ist. Wenn ja, wenden Sie sie an. Wenn nein, brauchen Sie ein alternatives Vorgehen.

4. Alternativen finden

Alternativen können darin bestehen, Datentransformationen durchzuführen, z.B. um eine Normalverteilung in der metrischen Variablen zu erzeugen oder indem Sie Kategorien sinnvoll zusammenfassen, um eine genügend große Zellbesetzung zu erreichen. Für einige Methoden stehen als Alternative auch nichtparametrische Verfahren zur Verfügung, die keine oder weniger Anforderungen an die Verteilung der Daten stellen. Oder Sie argumentieren über die Robustheit der ausgewählten Methoden oder setzen Bootstrapping zur Absicherung Ihrer Ergebnisse ein.

Welcher Weg der Beste ist, hängt von der individuellen Situation ab, von der Stichprobengröße, der Art der Hypothese, den zur Verfügung stehenden Methoden (auch softwareabhängig) und dem Ziel Ihrer Arbeit.

Mit dieser Schritt-für-Schritt-Vorgehensweise schaffen Sie es, die passende Analysemethode für jede Ihrer Hypothesen herauszufinden, ohne sich dabei zu verlaufen und ohne dabei etwas Wichtiges zu vergessen. Je nach Datensituation und Komplexität der Hypothese kommen Sie durch diesen Prozess sehr schnell und einfach hindurch oder müssen auch mal ein paar Schleifen drehen und – wie im letzten Abschnitt beschrieben – nach alternativen Methoden suchen.

Was Sie in jedem Fall brauchen ist ein Grundverständnis der statistischen Konzepte und das Wissen um die Eigenschaften und Voraussetzungen der einzelnen statistischen Methoden. Beides möchte ich Ihnen in dieser Artikelserie vermitteln.

Die Autorin

Daniela Keller ist leidenschaftliche Statistik-Expertin und berät Studierende und Wissenschaftler*innen zu allen Themen der statistischen Datenanalyse. Während ihres Studiums der Diplom-Mathematik gründete sie mit Kommilitonen eine studentische statistische Beratung und arbeitete anschließend selbständig in diesem Feld. Neben Einzelberatungen und Workshops unterstützt sie Ihre Kund*innen seit 2019 mit der Statistik-Akademie, ihrem Online-Mitgliederbereich für alle, die Statistik verstehen und selbständig anwenden wollen. Ihr Blog (www.statistik-und-beratung.de/blog) und ihr YouTube-Kanal sind Fundgruben für leicht verständlich aufbereitetes Statistikwissen für die Praxis.



© privat