



Glossar für statistische Analysetechniken

13

Christoph Glatz, Dimitri Prandner und Wolfgang Aschauer

Zusammenfassung

Um den Zugang zu diesem Buch auch für Personen zu erleichtern, die wenig Erfahrung mit quantitativer Sozialforschung haben, wird in diesem Glossar ein grundlegender Einblick in die Praxis der Umfrageforschung sowie in statistische Analysetechniken gegeben. Wir geben Hinweise, wie man soziale und politische Einstellungen in Umfragen erhebt und welche Möglichkeiten der statistischen Analyse vorliegen. Insbesondere komplexere Verfahren wie Regressionsanalyse, Faktorenanalysen, Reliabilitätsanalysen und Kontrastgruppenanalysen werden näher erklärt.

C. Glatz (✉)

Institut für Soziologie, Karl Franzens Universität Graz, Graz, Österreich

E-Mail: christoph.glatz@edu.uni-graz.at

D. Prandner

Institut für Soziologie, Johannes Kepler Universität Linz, Linz, Österreich

E-Mail: Dimitri.Prandner@jku.at

W. Aschauer

Fachbereich Politikwissenschaft und Soziologie, Paris-Lodron-Universität Salzburg, Salzburg, Österreich

E-Mail: wolfgang.aschauer@sbg.ac.at

© Der/die Autor(en) 2022

W. Aschauer et al. (Hrsg.), *Die Österreichische Gesellschaft während der Corona-Pandemie*, https://doi.org/10.1007/978-3-658-34491-7_13

349

13.1 Einleitung zur statistischen Auswertung

In diesem Sammelband wird der Einfluss der Corona-Krise auf Werte, Einstellungen, Befindlichkeiten und Verhaltensweisen der österreichischen Gesellschaft behandelt. Die Erkenntnisse basieren dabei vor allem auf Umfragedaten und wurden mittels quantitativer statistischer Analysen gewonnen. Umfragedaten beinhalten dabei die Antworten der Teilnehmer*innen auf bestimmte Fragen (oder Items), wobei man von der Gesamtzahl der Teilnehmer*innen (der Stichprobe) durch statistische Verfahren auf die österreichische Grundgesamtheit schließen kann (bei statistischen Analysen üblicherweise mittels Signifikanzprüfung, siehe Tab. 13.1 „Signifikanz“). Bezüglich der Aussagekraft der Ergebnisse sind wir jedoch stets auch mit potenziellen Verzerrungen konfrontiert, die in den Schlussfolgerungen entsprechend zu berücksichtigen sind (siehe Kap. 12 (Prandner 2022))

Tab. 13.1 Grundbegriffe der statistischen Auswertung von Umfrageprogrammen

Antwortcodes	Numerischer Ausdruck der Antwortkategorien (z. B. „1“, „2“, „3“, „4“)
Antwortkategorien	Die Antwortmöglichkeiten (z. B. „Sehr schlecht“, „Schlecht“, „Gut“, „Sehr gut“)
Gewichtung	Um die Grundpopulation (hier die österreichische Bevölkerung) so gut wie möglich abzubilden, gibt es Gewichtungsvariablen im Datensatz, sodass die Stichprobe in bestimmten Merkmalen (üblicherweise Alter, Geschlecht, Bildung und Bundesland) der österreichischen Grundpopulation entspricht
Item	Fragen in Umfrageprogrammen an die Teilnehmer*innen (Bsp. „Wie geht es Ihnen?“)
Signifikanz (p-Wert < ,05)	Die Signifikanz wird mittels p-Wert überprüft. Dieser beschreibt, nach dem Prinzip der Falsifizierbarkeit, die Wahrscheinlichkeit, dass es <i>keinen</i> Effekt bzw. Einfluss der jeweiligen Variable gibt, sprich, dass es in der Grundpopulation <i>keinen</i> Unterschied gibt für die jeweilige Forschungsfrage. Je geringer der p-Wert desto wahrscheinlicher ist daher die Annahme, dass es tatsächlich einen Effekt gibt (z. B. Unterschiede in der Impfbereitschaft zwischen älteren und jüngeren Personen, etc.). Üblicherweise wird von einem signifikanten Effekt gesprochen, wenn die Wahrscheinlichkeit, dass sich ein Effekt aus der Grundgesamtheit <i>nicht</i> in der Stichprobe finden lässt, kleiner als 5 % ist ($p < ,05$)
Skalierung	Bandbreite der Antwortkategorien, in dem oben genannten Beispiel 1–4
Variable	Abgeleitete und kodierte Information aus dem Item für die weiterführende Analyse (z. B. „aktuelles Wohlbefinden“)

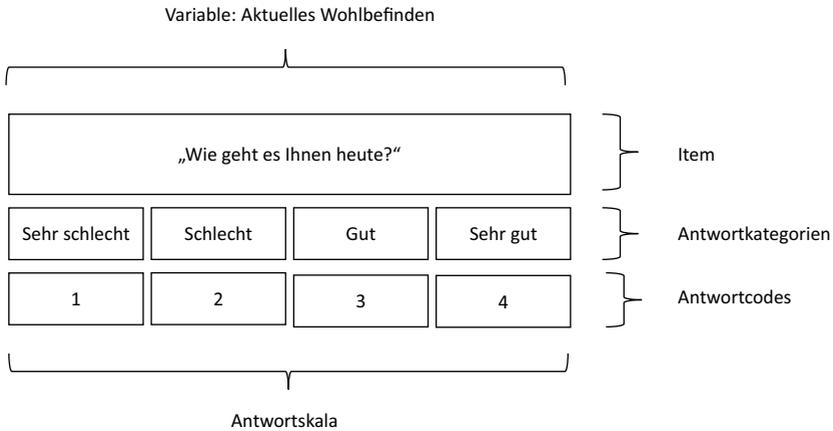


Abb. 13.1 Grundbegriffe der statistischen Auswertung von Umfragedaten anhand eines Beispielimens

in diesem Band). Die Antworten der Befragten werden anschließend in Zahlen übersetzt („codiert“), womit sie zu Variablen werden, welche für die statistische Analyse herangezogen werden. Wir sprechen bei den meisten Fragen bei Umfragedaten von geschlossenen Fragen, weil bereits vorgefertigte Antwortkategorien bestehen. Aus der Frage „Wie geht es Ihnen heute?“ lässt sich z. B. die Variable „aktuelles Wohlbefinden“ ableiten, mit den vorgefertigten Antwortkategorien „1“ (sehr schlecht) bis „4“ (sehr gut). Diese Kategorisierung wird als *Skalierung* bezeichnet und beträgt im dargestellten Beispiel (Abb. 13.1) 1–4.

13.2 Begrifflichkeiten der univariaten und bivariaten Analyse

Wenn die Variablen in einer Datenmatrix entsprechend aufbereitet wurden, erfolgt die quantitative Auswertung auf drei Wegen (für die **multivariate Analyse** siehe nächsten Abschnitt). **Eine univariate Analyse** betrachtet ausschließlich die Verteilung einer Variable. Ein Beispiel wäre die Auswertung der Frage, wie viele Österreicher*innen eine Covid-19-Impfung erhalten wollen sobald diese zugänglich ist. Eine **bivariate Analyse** betrachtet zwei Variablen gemeinsam, entweder mittels Zusammenhangs- oder Unterschiedsanalysen. Zusammenhangsanalysen befassen sich mit der Verbindung von zwei Variablen, beispielsweise ob ältere Personen

eine höhere Akzeptanz einer Covid-19-Impfung berichten oder nicht. Wird dagegen eine Unterschiedshypothese überprüft, so analysiert man, ob sich bestimmte Gruppen voneinander in einem Merkmal unterscheiden. Hier sind häufig Lageparameter (wie Mittelwertsunterschiede zwischen Gruppen) relevant. Ein einfaches Beispiel wäre die Frage: „Gibt es Unterschiede hinsichtlich der Impfbereitschaft zwischen Frauen und Männern?“. Zur Beantwortung der Frage würden die Mittelwerte in der Befürwortung der Impfung zwischen Männern und Frauen verglichen sowie die Signifikanz überprüft werden (siehe Tab. 13.1). Zentrale Begriffe und Tests der univariaten und bivariaten Analyse sind in der Tab. 13.2 dargestellt.

Tab. 13.2 Grundbegriffe und Kennwerte der univariaten und bivariaten Analyse

Univariate Analyse/Lageparameter	
Median	Der Wert, der eine Stichprobe in zwei gleich große Gruppe teilt. Wird z. B. die Impfbereitschaft von 0 (niedrig) bis 10 (hoch) erhoben, dann läge die Hälfte der Stichprobe bei einem Median von 5 unter diesem Wert, und die andere Hälfte darüber
Mittelwert	Durchschnittliche Ausprägung einer Variable in der jeweiligen Gruppe
Mittelwertindex	Gebildeter Index aus dem Mittelwert mehrerer Variablen
Schiefe-Wert	Erfasst, wie gleichmäßig sich die Stichprobe verteilt. Wenn beim oben genannten Beispiel (siehe Median) die meisten Teilnehmer*innen hohe Werte (z. B. 8, 9, 10) berichten, dann wäre diese Variable rechtssteil verteilt. Berichten die meisten Teilnehmer*innen dagegen geringe Werte wäre die Variable linkssteil. Ist die Verteilung ausgeglichen, so entspricht der Median ungefähr dem Mittelwert
Standardabweichung	Durchschnittliche Abweichung vom Mittelwert in der jeweiligen Gruppe
Z-Standardisierung	Transformierung von Variablen, sodass der Mittelwert 0 beträgt und die Standardabweichung 1. Diese Standardisierung ist sinnvoll, wenn Variablen mit unterschiedlichen Skalierungen miteinander verglichen werden

Bivariate Analyse

Kennwert zur Messung von Zusammenhangsanalysen

(Fortsetzung)

Tab. 13.2 (Fortsetzung)

Korrelationskoeffizient	Die Produkt-Moment-Korrelation (Pearsons´ s r) ist der zentrale Koeffizient zur Messung des Zusammenhangs zweier Variablen. Er nimmt Ausprägungen zwischen -1 und $+1$ an, je nachdem ob es sich um einen indirekt proportionalen (negativen) Zusammenhang (je höher x desto niedriger y) oder um einen direkt proportionalen (positiven) Zusammenhang (je höher x desto höher y) handelt. Je weiter der Kennwert von 0 abweicht desto stärker ist der Zusammenhang zweier Variablen. Je nach Skalenniveau gibt es weitere Zusammenhangsmaße, die angewendet werden können und ähnlich interpretiert werden
<i>Tests zur Durchführung von Unterschiedsanalysen</i>	
Chi ² -Test	Test für Unterschiedshypothesen bei Variablen mit nominalen Skalenniveaus (= Antwortmöglichkeiten, die qualitativ unterschiedlich sind; z. B. Unterschiede nach Bundesland)
F-Test	Überbegriff mehrerer statistischer Tests, welche für die Signifikanzprüfung auf einer speziellen Wahrscheinlichkeitsverteilung, der F-Verteilung, beruhen. Ein klassisches Verfahren hierzu wäre die Varianzanalyse (siehe unten)
T-Test	Überbegriff mehrerer statistischer Tests, welche für die Signifikanzprüfung auf der t-Verteilung beruhen
T-Test bei unabhängigen Stichproben	Überprüfung der Signifikanz zwischen zwei unabhängigen Gruppen mittels T-Test
T-Test bei verbundenen Stichproben	Überprüfung der Signifikanz zwischen zwei verbundenen Gruppen (z. B. dieselben Personen über zwei unterschiedliche Messzeitpunkte) mittels T-Test
Varianzanalyse (einfaktoriell)	Varianzanalyse mit einer unabhängigen Variable. Z. B. Mittelwertsvergleiche zwischen verschiedenen Altersgruppen
Varianzanalyse (mehrfaktoriell)	Varianzanalyse mit mehreren unabhängigen Variablen. Z. B. Mittelwertsvergleiche zwischen verschiedenen Alters- und Geschlechtergruppen
Varianzanalyse (Messwiederholung)	Varianzanalyse bei Längsschnittdaten. Überprüft signifikante Änderungen in der jeweiligen Gruppe über mehrere Messzeitpunkte
Welch-Test	Wird als Prüfinstrument eingesetzt, wenn sich die Standardabweichungen der Gruppen signifikant unterscheiden (robuster bzw. konservativer als T-Test)
<i>Deskriptive (=beschreibende) Merkmale einer Stichprobe</i>	
Effektstärke	Standardisierter Kennwert bei Unterschiedsanalysen mit einer Ausprägung von 0 bis 1, die Interpretation der Stärke des Effekts erfolgt in ähnlicher Form wie beim Korrelationskoeffizienten

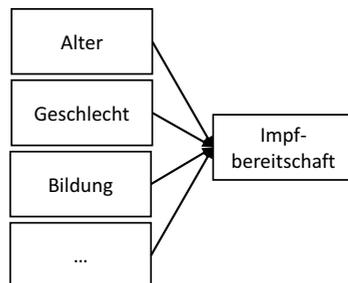
13.3 Erklärung und Begrifflichkeiten der multivariaten Analyse

Wir trennen in unseren sozialwissenschaftlichen Erklärungsmodellen oft zwischen unabhängigen Variablen (jene Merkmale, die wir als Einflussfaktoren bzw. Prädiktoren bewerten) und abhängigen Variablen (die wir möglichst gut erklären möchten). Da zur Erklärung einer abhängigen Variable meist mehrere Einflussfaktoren herangezogen werden müssen, werden in den (Sozial-)Wissenschaften häufig **multivariate Analysen** angewendet, um mehrere unabhängige Variablen in ein Analysemodell zu integrieren¹. So ist beispielsweise realistisch, dass nicht nur das Alter die Impfbereitschaft erklärt, sondern auch andere Variablen wie beispielsweise das Geschlecht, die Bildung, und vieles mehr (siehe Abb. 13.2). Demnach wäre es denkbar, dass niedriger gebildete Personen den Fortschritten der Wissenschaft und den allgemeinen Corona-Maßnahmen skeptischer gegenüberstehen und deshalb impfkritischer sind. Auch Frauen könnten im Vergleich zu Männern eine größere Impfskepsis aufweisen.

13.3.1 Regressionsanalyse

Ein klassisches Verfahren der multivariaten Analyse stellt die multiple Regressionsanalyse dar. Diese Analyse erlaubt es, mehrere erklärende (oder unabhängige) Variablen in ein Modell aufzunehmen, um eine zu erklärende (oder abhängige) Variable zu schätzen. Durch diese Art der Analyse kann man erkennen, wie sich die jeweilige unabhängige Variable auf die abhängige Variable

Abb. 13.2 Multivariate Analyse (Beispiel mehrere Einflussfaktoren/additive Effekte)



¹ Multivariate Analysen bieten außerdem die Möglichkeit Interaktionen zwischen Variablen zu entdecken bzw. zu überprüfen. Auf diese Interaktionseffekte wird in diesem Glossar allerdings nicht weiter eingegangen, weil dies den Rahmen des Kapitels sprengen würde.

auswirkt und man kann einzelne Effekte durch die Konstanthaltung der restlichen (Kontroll-)Variablen unverzerrt ausweisen. Im Beispiel der Abb. 13.2 sieht man jeweils den Effekt des Alters, des Geschlechts, und der Bildung auf die Impfbereitschaft unter Kontrolle der restlichen Variablen. Zentrale Kennwerte der multiplen Regressionsanalyse sind der Tab. 13.3 zu entnehmen.

Es gibt mehrere verschiedene Arten der Regressionsanalyse, welche sich hauptsächlich aufgrund der Skalenniveaus bzw. der Datenstruktur (Querschnitt vs. Längsschnitt) unterscheiden. Die klassische bzw. einfachste Form der Regressionsanalyse stellt die Ordinary Least Squares (OLS) Regressionsanalyse dar (oder lineare Regressionsanalyse). Darüber hinaus kommen in diesem Sammelband auch sequenzielle, ordinale und fixed-effects Regressionsanalysen zum Einsatz. Die zentralen Charakteristika der Verfahren sind jedoch ähnlich, auch wenn die Interpretation der Ergebnisse oftmals leicht unterschiedlich ist.

13.3.2 Explorative Faktorenanalyse

Im Gegensatz zur Regressionsanalyse besteht das Ziel der (explorativen) Faktorenanalyse nicht in der Erklärung einer abhängigen Variable, sondern darin, mehrere direkt erhobene Variablen zu einem oder mehreren Faktoren zu bündeln. Diese Faktoren stellen latente (und damit sehr messgenaue) Konstrukte dar, die mittels der direkt erhobenen Variablen geschätzt werden können. Diese geschätzten Faktorscores² können anschließend für weiterführende Analysen verwendet werden, beispielsweise als abhängige oder unabhängige Variablen in einer Regressionsanalyse.

Die explorative Faktorenanalyse basiert auf den Korrelationen der Variablen und der Annahme, dass diese Korrelation durch einen gemeinsamen latenten Faktor zustande kommt. Je besser eine Variable durch den jeweiligen Faktor erklärt werden kann, desto höher ist die Faktorladung dieser Variable auf den Faktor. Dieser Vorgang ist vergleichbar mit einer Regressionsanalyse bei der die Variablen die unabhängigen Variablen darstellen und die Faktorladung den Beta-Koeffizient, während der Faktor die abhängige Variable bildet. Demnach gibt es auch hier eine Varianzaufklärung des Faktors durch die Variablen. Die Varianzaufklärung gibt an, wie gut dieser Faktor durch die auf ihn gebündelten Items erklärt werden können. In der Regel werden Faktoren extrahiert, die mehr Information enthalten als eine der ursprünglichen Variablen; also einen

² Anstelle der Faktorscores ist es auch möglich, den Mittelwert der dem Faktor zugehörigen Items zu bilden.

Tab. 13.3 Zentrale Kennwerte der Regressionsanalyse

(Regressions-) Konstante	Ausprägung der abhängigen Variable, wenn alle unabhängigen Variablen die Ausprägung „0“ aufweisen
B-Wert	Die Veränderung der abhängigen Variable, wenn sich eine unabhängige Variable um eine Einheit erhöht während die restlichen unabhängigen Variablen kontrolliert werden. Bei positiven B-Werten zeigt sich ein positiver Zusammenhang der Variablen (je höher die Ausprägung der Variable x desto höher die Ausprägung der Variable y). Bei negativen B-Werten zeigt sich ein umgekehrter Zusammenhang
Beta-Wert	Standardisierter B-Wert zwischen -1 und 1 , sodass auch Variablen mit unterschiedlichen Skalierungen (z. B. $1-5$ vs. $1-10$) miteinander verglichen werden können. Die Interpretation erfolgt deckungsgleich zu den Korrelationskoeffizienten bzw. Effektstärken
Dummy Variablen	Spezielle Form der unabhängigen Variablen die mit „1“ und „0“ codiert (skaliert) sind. Typischerweise wird eine nominale Variable mit mehr als 2 Ausprägungen oder eine ordinale Messung für Regressionen in mehrere Dummy-Variablen aufgespalten. So können Spezifikationsfehler – Regressionen setzen metrisches Skalenniveau voraus – minimiert werden und die Variablen können dennoch differenziert analysiert werden
Multikollinearität	Eine Fehlerquelle in der Regressionsanalyse, die auftritt, wenn die Zusammenhänge zwischen den unabhängigen Variablen zu groß sind und dadurch eine wechselseitige Beeinflussung gegeben ist. Mittels verschiedener Koeffizienten kann Multikollinearität geprüft und somit auch kontrolliert werden
R^2	Die gesamte Erklärungskraft (Varianzaufklärung) der unabhängigen Variablen für die abhängige Variable. Variiert zwischen 0 und 1 (bzw. 0 % und 100 %) wobei der Wert aussagt, welcher Anteil der abhängigen Variable durch die unabhängige(n) Variable(n) erklärt werden kann
R^2 (korrigiertes)	Bei einem Regressionsmodell mit mehreren Variablen kommt es zu einer Verzerrung des R^2 Wertes da sich dieser Wert durch das Hinzufügen weiterer unabhängigen Variablen nur vergrößern kann, nicht aber verkleinern. Das korrigierte R^2 berücksichtigt dieses Problem und sollte in der Regel bei Regressionsanalysen mit mehreren unabhängigen Variablen berichtet werden
R^2 (Nagelkerke)	Alternativer (Pseudo-) R^2 Kennwert bei ordinalen, logistischen, binomialen und multinomialen Regressionsanalysen, da bei diesen Regressionsanalysen keine Varianzaufklärung berechnet werden kann. Der Wert zeigt an um wie viel Prozent sich das Erklärungsmodell (mit unabhängigen Variablen) im Vergleich zum Nullmodell (ohne unabhängige Variablen) verbessert

(Fortsetzung)

Tab. 13.3 (Fortsetzung)

Odds-ratio (OR)	Wird u. a. bei ordinalen oder logistischen Regressionsanalysen anstatt des Beta- und B-Wertes berichtet. Erklärt wird bei diesen Regressionsanalysen, im Vergleich zur klassischen (linearen) Regressionsanalyse, nicht eine kontinuierliche Änderung der abhängigen Variable (z. B. ein Zuwachs an Impfkzeptanz), sondern ein sprunghaftes Event (z. B. Ich lasse mich impfen oder nicht). Die Odds-Ratio schätzt dabei die Wahrscheinlichkeit des Eintretens dieses Events für die unabhängigen Variablen (UV). Bei der UV „Geschlecht“ (Mann = 0; Frau = 1) bedeutet eine OR von 1,10 beispielsweise, dass die Wahrscheinlichkeit, dass sich Frauen impfen lassen, um 10% erhöht ist (fiktives Beispiel)
-----------------	---

sogenannten Eigenwert > 1 aufweisen. Für den Fall, dass sich dabei zwei oder mehrere Faktoren bilden, wird in der Regel eine (orthogonale oder schiefwinklige) *Faktorrotation* durchgeführt, welche eine klare Zuordnung der Items zu den jeweiligen Faktoren erlaubt.

13.3.3 Reliabilitätsanalyse

Explorative Faktorenanalysen werden üblicherweise dann durchgeführt, wenn im vornherein nicht klar ist, wie viele verschiedene Dimensionen hinter einer Anzahl von Items verborgen sind. Wenn einzelne Itembatterien (das ist eine größere Anzahl an Items) nur auf einen speziellen Faktor zurückgeführt werden können, dann kann die Messgenauigkeit dieses Faktors (= dieser Skala) geprüft werden. Sind die Items metrisch skaliert, das heißt, mit drei oder mehr Antwortmöglichkeiten, wobei die Abstände zwischen den Antwortmöglichkeiten gleich groß sind, wird die Reliabilität der Skala in der Regel mit dem Koeffizienten Cronbach's Alpha (α) gemessen. Der Wert bildet die sogenannte interne Konsistenz der Skala ab. Sind die Items dagegen dichotom skaliert, das heißt, mit nur zwei Antwortkategorien (z. B. „Ja“ und „Nein“), so wird üblicherweise der Reproduktionskoeffizient zur Prüfung der internen Konsistenz herangezogen (vgl. Tab. 13.4). In der klassischen Literatur zur sozialwissenschaftlichen Statistik wird üblicherweise bei einem Wert $> 0,7$ von einer guten internen Konsistenz gesprochen, sprich, die Items messen das dahinterliegende Konstrukt mit ausreichender Präzision.

Tab. 13.4 Zentrale Kennwerte der Reliabilitätsanalyse und der Kontrastgruppenanalyse

<i>Reliabilitätsanalyse</i>	
Cronbach's Alpha	Überprüfung der internen Konsistenz einer Skala, der Koeffizient zeigt die Zuverlässigkeit der Messung einer latenten Variable (eines zugrunde liegenden Faktors) auf. Anwendung bei metrischen Items bzw. auch identisch skalierten ordinalen Items (siehe Fließtext)
Guttman Skala	Eine Guttman-Skala beruht auf dem Konzept einer Stufenleiter. Die Wahl von einer Antwortkategorie bzw. Ausprägung ist dabei immer in Abhängigkeit von nachfolgenden zu sehen. Z. B.: Wird die Aussage „Die Situation in allen EU-Ländern wird in den nächsten Jahren schlechter“ bejaht, muss logischerweise die Antwort auf Aussage „Die Situation in Österreich wird in den nächsten Jahren schlechter“ ebenso bejaht werden, da Österreich ein EU-Land ist. Werden diese logischen Ketten im Antwortverhalten bei zusammenhängenden Fragen eingehalten, kann man von einer Guttman-Skalierung ausgehen
Reproduktions Koeffizient	Die Beurteilung der internen Konsistenz erfolgt bei der Guttman-Skala auf der Grundlage des Reproduktionskoeffizienten, der das Verhältnis zwischen erwarteter logisch bedingter Antwort und tatsächlichem Antwortverhalten ausdrückt. Kommt es zu unlogischen Antworten und der Reproduktionskoeffizient liegt unter 0,9, kann nicht von einer Guttman-Skalierung ausgegangen werden
<i>Kontrastgruppenanalyse</i>	
CHAID-TREE-Analyse	Die TREE-Analyse mit Entscheidungsbäumen dient dazu, Gruppen zu identifizieren und Beziehungen zwischen den Gruppen zu entdecken. Das Modell erstellt mittels Kontrastgruppenanalyse Prognosen über Gruppenzugehörigkeit auf Grundlage von kategorialen Variablen und erstellt eine visuelle Klassifizierung. Ausgehend von der Wahrscheinlichkeit, dass eine bestimmte Antwortoption bei einer kategorialen Variablen gewählt wird, wird die Wahrscheinlichkeit für Segmentierung oder Stratifizierung nachvollziehbar gemacht. Die visuelle Darstellung der Ergebnisse erfolgt mittels Baumdiagrammen

13.3.4 Kontrastgruppenanalyse

Bei der Kontrastgruppenanalyse handelt es sich um ein Verfahren, welches besonders zur Aufdeckung von Interaktionseffekten geeignet ist (vgl. Tab. 13.4).

Im Unterschied zur näher besprochenen (OLS) Regressionsanalyse werden ausschließlich dichotome Variablen mit zwei Ausprägungen (z. B. Geschlecht mit den Ausprägungen männlich/weiblich) als unabhängige (erklärende) Variablen herangezogen. Diese werden im Gegensatz zur Regressionsanalyse außerdem nacheinander ins Modell inkludiert, um die abhängige Variable zu erklären. Dabei wird diejenige unabhängige Variable zuerst ausgewählt, welche die höchste Erklärungskraft bietet. Die Ergebnisse der Kontrastgruppenanalyse lassen sich dabei am einfachsten mittels Baumdiagrammen (wie man sie von Stammbäumen kennt) darstellen. Eine solche Darstellung ist im Beitrag „Arbeit und Familie im Covid-19-Alltag“ von (*Beham-Rabanser et al. 2021*) im Anhang enthalten.

Literatur

Prandner, D. (2022). Zu Datengrundlage und Datenqualität: Methodische Reflexion zur quantitativen Erhebung während der Corona-Krise. In Aschauer, W., Glatz, C., Prandner, D. (Hrsg.), *Die österreichische Gesellschaft während der Corona-Pandemie* (S. 327–349). Springer, V.S., Wiesbaden.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

