

# 11 Deskriptive und funktionale Modelle in der statistischen Sozialforschung

Götz Rohwer

Bochum

**Zusammenfassung.** Der Beitrag beschäftigt sich mit zwei Arten von Modellen, die in der statistischen Sozialforschung verwendet werden: Deskriptive Modelle, die dazu dienen, tatsächlich beobachtete oder aus Beobachtungen ableitbare Sachverhalte zu repräsentieren; und analytische Modelle, die dazu dienen, in der Form von Regeln formulierbare Hypothesen über Zusammenhänge zwischen Variablen darzustellen und mit Hilfe von Daten numerisch bestimmbar zu machen. Korrespondierend zu dieser Unterscheidung können zwei Arten statistischer Generalisierungen unterschieden werden: Einerseits deskriptive Generalisierungen, die Sachverhalte, die mit den Daten einer Stichprobe ermittelt wurden, für eine korrespondierende Grundgesamtheit verallgemeinern; andererseits modale Generalisierungen, die darauf zielen, Regeln zu begründen, die sich für Voraussagen und Erklärungen eignen. Diese Unterscheidung wird im ersten Abschnitt näher ausgeführt. Der zweite Abschnitt beschäftigt sich kurz mit deskriptiven Modellen, der dritte Abschnitt dann ausführlicher mit analytischen Modellen. Schließlich wird im vierten Abschnitt ausgeführt, dass es bei der Konstruktion statistischer Modelle wichtig ist, auch Zusammenhänge zwischen den jeweils verwendeten erklärenden Variablen explizit zu repräsentieren.

## 1 Zwei Arten der Generalisierung

In diesem Abschnitt beschäftige ich mich mit einer Unterscheidung zwischen deskriptiven und modalen Generalisierungen im Kontext statistischer Sozialforschung (Rohwer 2010, 2011b, 2012a). Einerseits geht es darum, ausgehend von deskriptiven Aussagen über eine Stichprobe wiederum deskriptive Aussagen über eine korrespondierende Grundgesamtheit zu machen. Andererseits sollen Regeln begründet werden, die sich für Voraussagen und Erklärungen eignen.

*S. 309–329 in: Norman Braun & Nicole J. Saam, Hg. (2014). Handbuch Modellbildung und Simulation in den Sozialwissenschaften. Wiesbaden: Springer VS*

### 1.1 Deskriptive statistische Aussagen

Ich verwende folgende Definition: Deskriptive statistische Aussagen sind Aussagen über Häufigkeitsverteilungen von Merkmalen (oder daraus abgeleitete Aussagen), die für eine bestimmte Menge von Einheiten definiert sind. Als formalen Rahmen verwende ich statistische Variablen (Rohwer & Pötter 2001; Rohwer 2010). Die symbolische Notation ist:  $X : \Omega \rightarrow \mathcal{X}$ .  $X$  ist der Name der Variablen,  $\Omega$  ist die Referenzmenge, die aus endlich vielen tatsächlich beobachteten oder als existierend angenommenen Einheiten besteht, und  $\mathcal{X}$  ist ein Merkmalsraum, d.h. eine Menge von Merkmalswerten, durch die die Einheiten charakterisiert werden können. Statistische Variablen sind also Funktionen: Jeder Einheit  $\omega \in \Omega$  wird durch die Variable  $X$  ein Merkmalswert  $X(\omega) \in \mathcal{X}$ , der sie charakterisiert, zugeordnet.

Statistische Variablen können aus beliebig vielen Komponenten bestehen, formal:  $X = (X_1, \dots, X_m)$ . Wichtig ist: Damit man von deskriptiven Aussagen sprechen kann, muss man annehmen können, dass es die Elemente der Referenzmenge tatsächlich gibt oder gegeben hat. Zum Beispiel: Die Gesamtheit der Kinder, die in Deutschland im Jahr 2010 einen Kindergarten besucht haben.

### 1.2 Deskriptive Generalisierungen

Statistische Variablen eignen sich auch zur Charakterisierung deskriptiver Generalisierungen. Ausgangspunkt ist eine statistische Variable  $X : \Omega \rightarrow \mathcal{X}$ , die die Beobachtungen repräsentiert. Die Menge der beobachteten Einheiten,  $\Omega$ , wird dann als Teilmenge einer Menge  $\Omega^*$  aufgefasst, für die eine analog definierte statistische Variable  $X^* : \Omega^* \rightarrow \mathcal{X}$  mit dem gleichen Merkmalsraum  $\mathcal{X}$  angenommen werden kann. Somit kann man definieren: Eine *deskriptive Generalisierung* besteht darin, mit Hilfe der beobachteten Werte von  $X$  deskriptive Aussagen über die Verteilung von  $X^*$  in  $\Omega^*$  zu machen.

Überlegungen zur Rechtfertigung solcher Generalisierungen müssen auf den datengenerierenden Prozess Bezug nehmen, durch den die Beobachtungen (insbesondere die Auswahl der Elemente von  $\Omega$ ) zustande gekommen sind. Darauf kann in diesem Beitrag nicht näher eingegangen werden. Es sei aber hier bereits folgende terminologische Unterscheidung erwähnt: Ich verwende den Ausdruck *datengenerierender Prozess* zum Verweis auf einen Prozess, durch den Daten, d.h. Informationen über bereits existierende Sachverhalte, erzeugt werden. Dagegen spreche ich von einem *faktengenerierenden Prozess*, wenn durch den Prozess ein neuer Sachverhalt entsteht, den es zuvor noch nicht gab. Als Beispiel kann man an einen Lernprozess denken, durch den Schüler bestimmte Fähigkeiten entwickeln. Das ist ein faktengenerierender Prozess. Erst wenn dieser Prozess stattgefunden hat, kann ein datengenerierender Prozess einsetzen, durch den ermittelt (und als Werte einer Variablen aufgeschrieben) wird, was die Schüler gelernt haben.

Eine wichtige Frage betrifft die Grundgesamtheit, für die eine deskriptive Generalisierung vorgenommen werden kann. Insofern das Ziel darin besteht, Fakten zu ermitteln, gibt es enge Grenzen, insbesondere wenn sich die zu erforschenden Sachverhalte historisch schnell verändern. Die Grenzen werden offensichtlich, wenn zur

Rechtfertigung deskriptiver Generalisierungen mit zufällig ausgewählten Stichproben argumentiert wird: Die Grundgesamtheit,  $\Omega^*$ , kann dann nur Elemente umfassen, für die es *zum Zeitpunkt der Stichprobenziehung* eine positive Auswahlwahrscheinlichkeit gab.

### 1.3 Modale Generalisierungen mit Regeln

Jetzt betrachte ich Generalisierungen, die die Form prediktiver Regeln haben und nicht als deskriptive Aussagen formuliert werden können. Das Wort „Regel“ verwende ich zunächst in einer allgemeinen Bedeutung für Aussagen der Form

Wenn . . . , dann . . .

Unterschiedliche Arten von Regeln ergeben sich durch Modalitäten für die Formulierung des *dann*-Teils, zum Beispiel: Wenn . . . , dann ist . . . möglich, oder wahrscheinlich, oder notwendig, oder normativ gefordert.

In der statistischen Sozialforschung interessiert man sich in erster Linie für prediktive Regeln. Zum Beispiel, wobei  $\omega$  der Name *irgendeiner* Person ist, die einen Schulabschluss in Deutschland erworben hat: Wenn mindestens ein Elternteil von  $\omega$  ein Abitur hat, dann ist es sehr wahrscheinlich, dass auch  $\omega$  die Schule mit einem Abitur abgeschlossen hat. Man beachte, dass es sich um eine *generische* Regel handelt:  $\omega$  ist kein Name für eine bestimmte Person, sondern wird nur durch Werte von Variablen definiert (in diesem Beispiel: eine Person, die einen Schulabschluss in Deutschland erworben hat).

Dieses Beispiel illustriert *statische prediktive Regeln*. Sie haben die Form: Wenn  $\omega$  (ein generisches Objekt) die Eigenschaft  $x$  hat, dann hat  $\omega$  (wahrscheinlich) auch die Eigenschaft  $y$ . Hiervon zu unterscheiden sind *dynamische prediktive Regeln*, die sich auf einen faktengenerierenden Prozess beziehen, durch den ein Sachverhalt entstehen kann. Zum Beispiel: Wenn  $\omega$  (eine generisch spezifizierte Person) regelmäßig am Unterricht teilnimmt, wird sie wahrscheinlich die Abschlussprüfung erfolgreich bestehen.

### 1.4 Formulierungen prediktiver Regeln mit Variablen

Im Folgenden betrachte ich nur probabilistische Regeln, bei denen es für den *dann*-Teil eine probabilistische Qualifizierung gibt.<sup>1</sup> Es ist üblich, solche Regeln mit Hilfe von Variablen zu formulieren. Probabilistische Qualifizierungen können dann auf zwei unterschiedliche Arten ausgedrückt werden. Qualitativ: Wenn  $X = x$ , dann ist  $Y = y$  wahrscheinlich (wobei weitere Qualifizierungen hinzugefügt werden können); oder quantitativ: Wenn  $X = x$ , dann ist  $\Pr(Y = y) = p$  (wobei  $p$  eine bestimmte Zahl ist).

In der statistischen Sozialforschung werden meistens quantitative Formulierungen verwendet. Wenn dies angenommen wird, können prediktive Regeln durch mathematische Funktionen dargestellt werden; in einer allgemeinen Formulierung:

<sup>1</sup> Deterministische Regeln werden in der statistischen Sozialforschung normalerweise nicht verwendet, sie bilden jedoch den theoretischen Kern der von Charles Ragin vorgeschlagenen „Qualitative Comparative Analysis“; vgl. Rohwer 2011a, 2012a.

$$x \mapsto \Pr[Y|X = x] \quad (1)$$

zu lesen als eine mathematische Funktion, die jedem Wert  $x$  im Wertebereich von  $X$  eine durch  $X = x$  bedingte Wahrscheinlichkeitsverteilung von  $Y$  zuordnet. Wenn  $Y$  eine diskrete Variable ist, kann man auch Funktionen  $x \mapsto \Pr(Y = y|X = x)$  verwenden (für jeden möglichen Wert von  $Y$  gibt es dann eine solche Funktion). Oft werden auch bedingte Erwartungswerte verwendet, und die Regeln haben dann die Form  $x \mapsto E(Y|X = x)$ .

Ausgehend von solchen allgemeinen Formulierungen kann man bestimmte parametrische Funktionsformen festlegen, z.B. ein Logitmodell, wenn  $Y$  eine binäre Variable ist. Aber auch dann hat man zunächst nur eine Form für den angenommenen funktionalen Zusammenhang. Erst wenn die Funktion auch numerisch spezifiziert (mit Hilfe von Daten „geschätzt“) worden ist, erhält man schließlich eine Regel, die für Voraussagen verwendet werden kann.

### 1.5 *Prediktive Regeln versus deskriptive Aussagen*

Prediktive Regeln müssen von deskriptiven Aussagen unterschieden werden. Ein Unterschied besteht zunächst darin, dass sich deskriptive Aussagen auf eine jeweils bestimmte Menge von Einheiten beziehen, prediktive Regeln dagegen auf eine generische Einheit, die nur durch Werte von Variablen bestimmt ist. Korrespondierend gibt es eine begriffliche Differenz zwischen relativen Häufigkeiten (Anteilswerten),  $P(Y = y|X = x)$ , die eine endliche Referenzmenge voraussetzen, und Wahrscheinlichkeiten  $\Pr(Y = y|X = x)$ , die eine generische Einheit betreffen (Rohwer & Pötter 2002). Ich verwende deshalb unterschiedliche Symbole:  $P$  für relative Häufigkeiten,  $\Pr$  für Wahrscheinlichkeiten.

Ein Zufallsgenerator kann dazu dienen, die Unterscheidung zu verdeutlichen. Als Beispiel verwende ich das Werfen eines Würfels. Der Zufallsgenerator kann durch folgende Regel definiert werden: Wenn der Würfel geworfen wird, gibt es sechs mögliche Ergebnisse, von denen jedes mit der gleichen Wahrscheinlichkeit ( $1/6$ ) eintreten kann. Diese Regel muss von einer deskriptiven Aussage über eine Menge tatsächlich realisierter Ergebnisse unterschieden werden.

Angenommen, man wirft den Würfel 100 Mal. Die Ergebnisse können durch eine statistische Variable  $Z : \Omega \rightarrow \mathcal{Z} := \{1, \dots, 6\}$  repräsentiert werden. Die Häufigkeitsverteilung  $P[Z]$  gibt an, mit welchen Häufigkeiten die sechs möglichen Ergebnisse bei den 100 realisierten Würfeln tatsächlich aufgetreten sind. Sie muss offenbar von der Wahrscheinlichkeitsverteilung, die zur Charakterisierung des Zufallsgenerators verwendet wurde, unterschieden werden.

Es gibt eine wichtige Konsequenz. Da sich prediktive Regeln von deskriptiven Aussagen (und natürlich auch von analytischen Wahrheiten) unterscheiden, können sie nicht wahr oder falsch sein. Sie können nur pragmatisch gerechtfertigt werden, d.h. mit Argumenten, die zeigen, dass und wie sie von Menschen in ihren Aktivitäten genutzt werden können.

### 1.6 Statistische und modale Variablen

Die konzeptionelle Unterscheidung zwischen deskriptiven Aussagen und prediktiven Regeln motiviert eine korrespondierende Unterscheidung von Arten von Variablen. Wie schon ausgeführt wurde, betreffen deskriptive statistische Aussagen die Häufigkeitsverteilungen statistischer Variablen, d.h. tatsächlich realisierte Merkmalsverteilungen in bestimmten Referenzmengen, deren Elemente tatsächlich existieren oder existiert haben. Dies gilt natürlich auch für deskriptive Aussagen über bedingte Häufigkeiten; z.B. setzt  $P(Y = y|X = x)$  eine zweidimensionale statistische Variable  $(X, Y)$  voraus.

Bezieht man sich stattdessen auf bedingte Wahrscheinlichkeiten, gibt es zwei unterschiedliche theoretische Kontexte.

- a) Man kann annehmen, dass es eine gemeinsame Wahrscheinlichkeitsverteilung aller jeweils betrachteten Variablen gibt, von der dann bedingte und marginale Verteilungen abgeleitet werden können.
- b) Diese Annahme ist jedoch nicht ohne Weiteres sinnvoll, wenn bedingte Wahrscheinlichkeiten zur Formulierung prediktiver Regeln verwendet werden. Man betrachte z.B. die Regel  $x \mapsto \Pr[Y|X = x]$ . Hier dient die Variable  $X$  nur dazu, die *wenn*-Bedingung der Regel zu formulieren, und es bleibt vollständig offen, ob bzw. wie sie tatsächlich bestimmte Werte annimmt.  $X$  ist in diesem Kontext also weder eine statistische Variable, für die eine bestimmte Häufigkeitsverteilung, noch eine Zufallsvariable, für die eine bestimmte Wahrscheinlichkeitsverteilung angenommen werden kann. Und folglich ist auch  $Y$  keine gewöhnliche Zufallsvariable, für die eine unbedingte Wahrscheinlichkeitsverteilung angenommen werden kann.

Für die Überlegungen in diesem Beitrag ist der zweite Kontext relevant. Um daran zu erinnern, bezeichne ich Variablen, die zur Formulierung prediktiver Regeln verwendet werden, als *modale Variablen* und verwende eine spezielle Notation. Modale Variablen, für die keine Verteilungen angenommen werden können, werden durch zwei Punkte kenntlich gemacht, und modale Variablen, für die nur bedingte Verteilungen angenommen werden können, werden durch einen Punkt kenntlich gemacht. In unserem Beispiel:  $\ddot{X}$  anstelle von  $X$  und  $\dot{Y}$  anstelle von  $Y$ , so dass die Regel die Form  $x \mapsto \Pr[\dot{Y}|\ddot{X} = x]$  annimmt.

### 1.7 Statistische Erklärungen

Wofür sind prediktive Regeln nützlich? Ein Teil der Antwort ist offensichtlich: Sie sind erforderlich, um Beobachtungen für Voraussagen nutzen zu können. Aber können prediktive Regeln auch für Erklärungen verwendet werden?

Durch die Konzeption sog. induktiv-statistischer Erklärungen (Hempel 1965) wurde vorgeschlagen, dass probabilistische Regeln auch zur Erklärung *individueller* Ereignisse verwendet werden können. Diese Konzeption wurde jedoch kritisiert. Einer der Kritikpunkte ist leicht verständlich: dass A ein Ereignis B zu einem gewissen Grad wahrscheinlich macht, zeigt nicht, warum B eingetreten ist.

Diese Diskussion betrifft Unterschiede zwischen „erklären“ und „voraussagen“ und kann ignoriert werden, wenn man nur an Voraussagen interessiert ist. Ist man jedoch auch an Erklärungen interessiert, sollte man nicht nur zwischen verschiedenen Arten von Explananda, sondern auch zwischen Warum- und Wie-Fragen unterscheiden (Cross 1991; Faye 1999). Hier beziehe ich mich auf statistische Erklärungen, durch die statistische Sachverhalte (statistische Verteilungen oder daraus ableitbare Größen) erklärt werden sollen.

Ich nehme an, dass die Verteilung  $P[Y]$  einer statistischen Variablen  $Y : \Omega \rightarrow \mathcal{Y}$  erklärt werden soll. Eine *statistische Erklärung*, so wie ich diesen Ausdruck hier verstehen möchte, verwendet zwei Prämissen:

- die Verteilung einer Variablen  $X$ , die für die gleiche Referenzmenge  $\Omega$  definiert ist, und
- eine probabilistische Regel  $x \mapsto \Pr[\dot{Y}|\ddot{X}=x]$ , wobei  $\ddot{X}$  und  $\dot{Y}$  in ihrer inhaltlichen Bedeutung den Variablen  $X$  und  $Y$  entsprechen.

Der formale Teil der statistischen Erklärung besteht dann darin, mit Hilfe von

$$\Pr(\dot{Y}=y) := \sum_x \Pr(\dot{Y}=y|\ddot{X}=x)P(X=x) \quad (2)$$

eine unbedingte Verteilung  $\Pr[\dot{Y}]$  von  $\dot{Y}$  abzuleiten (diese Formulierung unterstellt, dass alle Variablen diskret sind). Der prediktive Anspruch besteht dann darin, dass  $\Pr[\dot{Y}]$  näherungsweise der zu erklärenden Verteilung  $P[Y]$  entspricht. Dieser Anspruch ist allerdings trivialerweise erfüllt, wenn die verwendete prediktive Regel aus der gemeinsamen Verteilung von  $X$  und  $Y$  abgeleitet worden ist. Die Idee, dass die prediktive Regel eine Verallgemeinerung erlaubt, die über die beobachteten Daten hinausgeht, ist deshalb für den Erklärungsanspruch wesentlich.

Für den Erklärungsanspruch sind jedoch auch noch Überlegungen relevant, die weder die formale Ableitung noch den Allgemeinheitsgrad der prediktiven Regel betreffen. Insbesondere sind folgende Fragen wichtig:

- a) Ob sich die erklärenden Variablen auf Bedingungen eines faktengenerierenden Prozesses für die Sachverhalte, die durch  $Y$  repräsentiert werden, beziehen.
- b) Welche Zusammenhänge es zwischen den erklärenden Variablen gibt. (Diese Frage wird in Abschnitt 4 weiter verfolgt.)
- c) Ob es relevante erklärende Variablen gibt, die in dem funktionalen Modell (das die prediktive Regel liefert) nicht berücksichtigt worden sind.
- d) Ob bzw. in welchem Ausmaß die prediktive Regel, die für die Erklärung verwendet wird, von den statistischen Verteilungen der erklärenden Variablen abhängt.<sup>2</sup>
- e) Ob bzw. in welchem Ausmaß die prediktive Regel, die für die Erklärung verwendet wird, vom historischen Kontext abhängt.

<sup>2</sup> Man kann dann von „verteilungsabhängiger Kausalität“ sprechen, vgl. Rohwer 2010.

## 2 Deskriptive Modelle

Korrespondierend zur Unterscheidung zwischen deskriptiven Aussagen und Regeln kann man zwei Arten von Modellen unterscheiden. Ich beginne mit einer kurzen Charakterisierung deskriptiver Modelle.

### 2.1 Deskriptive Modelle mit statistischen Variablen

Deskriptive Modelle können als Hilfsmittel zur Beschreibung von Häufigkeitsverteilungen statistischer Variablen definiert werden. Ausgangspunkt ist eine statistische Variable  $X : \Omega \rightarrow \mathcal{X}$ , die oft aus mehreren Komponenten besteht. Ein deskriptives Modell dient der Aufgabe, die Verteilung von  $X$ , symbolisch:  $P[X]$ , oder Aspekte dieser Verteilung mit Hilfe einer relativ einfacheren mathematischen Form zu beschreiben.

Als ein Beispiel, bei dem  $X$  nur aus einer Komponente besteht, kann man an die Beschreibung einer Einkommensverteilung durch eine Lognormal-Verteilung denken.

### 2.2 Regressionsmodelle mit statistischen Variablen

Wenn  $X$  aus zwei oder mehr Komponenten besteht, ist man oft an Beschreibungen bedingter Verteilungen interessiert. Diesem Zweck dienen Regressionsfunktionen und -modelle. Ausgangspunkt ist eine zweidimensionale statistische Variable, etwa  $(X, Y) : \Omega \rightarrow \mathcal{X} \times \mathcal{Y}$ . Eine allgemeine Regressionsfunktion ist eine Funktion  $x \mapsto P[Y|X = x]$ , die jedem Wert  $x \in \mathcal{X}$  die durch  $X = x$  bedingte Häufigkeitsverteilung von  $Y$  zuordnet. Ein deskriptives Regressionsmodell liefert eine relativ einfachere mathematische Repräsentation dieser Funktion. Meistens wird eine parametrische Funktion  $g(x; \theta)$  verwendet und der Parametervektor  $\theta$  so bestimmt, dass näherungsweise gilt:  $g(x; \theta) \approx P[Y|X = x]$ . Als Regressionsmodell erhält man dann die Funktion  $x \mapsto g(x; \theta)$ .

Oft werden spezielle Regressionsfunktionen verwendet, um bestimmte Aspekte einer allgemeinen Regressionsfunktion darzustellen. Sehr oft bezieht man sich auf bedingte Mittelwerte  $M(Y|X = x)$ ; als Beispiel kann man an ein lineares Regressionsmodell  $M(Y|X = x) \approx \alpha + x\beta$  denken. Hier wird der durch  $X = x$  bedingte Mittelwert von  $Y$  durch eine lineare Funktion von  $x$  approximiert.

### 2.3 Deskriptive Modelle und deskriptive Generalisierungen

Den Sinn deskriptiver Modelle kann man zunächst darin sehen, dass sie helfen sollen, den Informationsgehalt umfangreicher oder komplexer Datenmengen zugänglich zu machen. Der berühmte Statistiker R. A. Fisher hat das einmal so gesagt:

„Briefly, and in its most concrete form, the object of statistical methods is the reduction of data. A quantity of data, which usually by its mere bulk is incapable of entering the mind, is to be replaced by relatively few quantities which shall adequately represent the whole, or which, in other words, shall

contain as much as possible, ideally the whole, of the relevant information contained in the original data.“ (Fisher 1922: 311)

Deskriptive Modelle sind auch nützliche Hilfsmittel für deskriptive Generalisierungen. Da sich diese Modelle auf statistische Variablen beziehen, macht es keinen wesentlichen Unterschied, ob es sich bei der Referenzmenge um eine Stichprobe oder eine korrespondierende Grundgesamtheit handelt. Ein deskriptives Modell kann ohne weiteres auch für eine Grundgesamtheit formuliert werden, und die Parameter des Modells können dann mit den Daten einer Stichprobe geschätzt werden.

Es ist bemerkenswert, dass der Ausdruck „schätzen“ in diesem Zusammenhang eine klare Bedeutung hat: Man möchte näherungsweise diejenigen Werte der Modellparameter ermitteln, die durch ihre hypothetische Berechnung mit den vollständigen Daten für die Grundgesamtheit definiert sind. Daraus folgt auch, dass bereits die Definition der zu schätzenden Größen nicht nur von der Spezifikation eines Modells, sondern auch von der Festlegung einer bestimmten Schätzmethode abhängt.

### 3 Analytische Modelle

Jetzt bespreche ich Modelle, die dem Zweck dienen, Beziehungen zwischen modalen Variablen zu formulieren. Zur Unterscheidung von deskriptiven Modellen nenne ich sie im Folgenden *analytische Modelle*. Das wesentliche Hilfsmittel sind Funktionen (im mathematischen Sinn), und ich spreche deshalb gleichbedeutend auch von *funktionalen Modellen* (Rohwer 2010).

#### 3.1 Beziehungen zwischen Variablen

Man betrachte zwei Variablen:  $X$  mit dem Wertebereich  $\mathcal{X}$  und  $Y$  mit dem Wertebereich  $\mathcal{Y}$ . Zwei Arten funktionaler Beziehungen müssen unterschieden werden.

- Eine deterministische funktionale Beziehung besteht aus einer Funktion

$$x \mapsto y = f(x) \tag{3}$$

die jedem Wert  $x \in \mathcal{X}$  genau einen Wert  $f(x) \in \mathcal{Y}$  zuordnet.

- Eine probabilistische funktionale Beziehung besteht aus einer Funktion

$$x \mapsto \Pr[Y|X = x] \tag{4}$$

die jedem Wert  $x \in \mathcal{X}$  eine durch  $X = x$  bedingte Wahrscheinlichkeitsverteilung von  $Y$  zuordnet. Man beachte, dass bereits  $\Pr[Y|X = x]$  eine Funktion ist. Wenn  $Y$  eine diskrete Variable ist, kann sie in der Form  $y \mapsto \Pr(Y = y|X = x)$  dargestellt werden.



### 3.2 Eine allgemeine Definition funktionaler Modelle

Funktionale Modelle können mit deterministischen oder probabilistischen Funktionen gebildet werden. In diesem Beitrag betrachte ich nur funktionale Modelle, die Funktionen der Form (4) verwenden.<sup>3</sup> Diese Funktionen entsprechen den probabilistischen prädiktiven Regeln, die in Abschnitt 1 definiert wurden, und somit können diese Modelle auch als Hilfsmittel zur Formulierung solcher Regeln verstanden werden. Ich verwende folgende allgemeine Definition:

- Die Struktur des Modells wird durch einen gerichteten azyklischen Graphen gebildet.
- Jedem Knoten des Graphen entspricht eine modale Variable. Variablen mit dem Eingangsgrad 0 werden exogene Variablen genannt und durch zwei Punkte gekennzeichnet. Alle anderen Variablen werden endogene Variablen genannt und durch einen Punkt gekennzeichnet.
- Für jede endogene Variable gibt es eine probabilistische Funktion, die angibt, wie ihre Wahrscheinlichkeitsverteilung von Werten der ihr unmittelbar vorangehenden Variable(n) abhängt.
- Wenn keine weiteren Annahmen getroffen werden, gibt es für die exogenen Variablen keine Wahrscheinlichkeitsverteilungen.

### 3.3 Ein einfaches Beispiel

Um den Begriff eines funktionalen Modells zu illustrieren, verwende ich ein einfaches Beispiel, das sich auf den schulischen Erfolg eines Schulkindes bezieht. Es hat die folgende Form (die Pfeile mit Doppelspitzen sollen probabilistische Beziehungen symbolisieren):

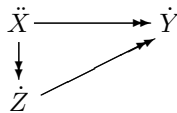


Abb. 1: Modell 1

Das Modell enthält drei Variablen, die ich zur Vereinfachung nur als binäre Variablen definiere:  $\dot{Y}$  = schulischer Erfolg (1 erfolgreich, 0 andernfalls),  $\ddot{X}$  = Bildungsniveau der Eltern (1 hoch, 0 niedrig),  $\dot{Z}$  = Schultyp (0 oder 1). Außerdem gibt es zwei probabilistische Funktionen: Eine Funktion  $x \mapsto \Pr[\dot{Z} | \ddot{X} = x]$ , die angibt, wie die Wahrscheinlichkeit der Schultypen vom Bildungsniveau der Eltern abhängt; und eine Funktion  $(x, z) \mapsto \Pr[\dot{Y} | \ddot{X} = x, \dot{Z} = z]$ , die angibt, wie die Wahrscheinlichkeit des Schulerfolgs sowohl vom Bildungsniveau der Eltern als auch vom Schultyp abhängt.

<sup>3</sup> Es ist jedoch möglich, dass ein solches Modell außerdem auch deterministische Funktionen enthält.

In diesem Beispiel genügt es natürlich, die Funktionen  $x \mapsto \Pr(\dot{Z} = 1 | \dot{X} = x)$  und  $(x, z) \mapsto \Pr(\dot{Y} = 1 | \dot{X} = x, \dot{Z} = z)$  zu betrachten.

Man beachte, dass nicht zu jedem Pfeil des Graphen eine separate Funktion gehört, sondern zu einer Funktion gehören jeweils alle Pfeile, die zu der gleichen endogenen Variablen führen.

### 3.4 Funktionale und kausale Beziehungen

Man beachte, dass die in einem funktionalen Modell verwendeten Funktionen nicht unmittelbar eine inhaltlich bestimmte Bedeutung haben. Damit sie eine Bedeutung gewinnen, bedarf es einer Interpretation. Natürlich wird man sich oft bemühen, die Funktionen so zu bestimmen, dass sie kausal interpretierbaren Abhängigkeitsbeziehungen entsprechen. Eine solche Interpretation kann aber nicht allein mit statistischen Überlegungen begründet werden (Rohwer 2012b).

### 3.5 Verteilungen für exogene Variablen?

Für die exogenen Variablen eines funktionalen Modells gibt es nicht bereits qua Modelldefinition Verteilungen. Es gibt aber manchmal Gründe, Annahmen über solche Verteilungen explizit einzuführen.

- a) Wenn man das Modell für eine statistische Erklärung verwenden möchte (s.o.).
- b) Wenn das Modell dazu dienen soll, um Voraussagen für eine bestimmte Einheit zu machen, für die Werte einiger exogener Variablen nicht bekannt sind. Man verwendet dann ein reduziertes Modell, bei dem über angenommene Verteilungen der fehlenden exogenen Variablen integriert wird. Zur Illustration sei angenommen, dass man mit dem Modell 1 den Schulerfolg eines Kindes voraussagen möchte, für das das Bildungsniveau seiner Eltern nicht bekannt ist. Wenn man dann die exogene Variable  $\dot{X}$  durch eine Variable  $\dot{X}$  substituiert und dafür eine bedingte Verteilung  $\Pr[\dot{X} | \dot{Z} = z]$  annimmt, kann man ein reduziertes Modell

$$\Pr(\dot{Y} = 1 | \dot{Z} = z) = \sum_x \Pr(\dot{Y} = 1 | \dot{Z} = z, \dot{X} = x) \Pr(\dot{X} = x | \dot{Z} = z) \quad (5)$$

verwenden, für das man nur den Schultyp des Kindes benötigt.

- c) Wenn man das Modell verwenden möchte, um ausgehend von Werten endogener Variablen Werte exogener Variablen vorauszusagen. Damit diese Form einer Bayesianischen Inferenz möglich wird, muss man zunächst mit Prior-Verteilungen für die exogenen Variablen beginnen.

### 3.6 Effekte erklärender Variablen

Es sei angenommen, dass eine endogene Variable  $\dot{Y}$  von einer exogenen (oder wiederum endogenen) Variablen  $\dot{X}$  abhängt. Die Vorstellung eines „Effekts“ von  $\dot{X}$  für  $\dot{Y}$  setzt einen Vergleich von

$$\Pr[\dot{Y}|\ddot{X}=x'] \quad \text{und} \quad \Pr[\dot{Y}|\ddot{X}=x''] \quad (6)$$

für (mindestens) zwei Werte  $x'$  und  $x''$  von  $\ddot{X}$  voraus. Dieser Vergleich betrifft bedingte Verteilungen und kann deshalb im Allgemeinen nicht durch eine einzige Zahl zusammengefasst werden. Deshalb werden oft vereinfachende Effektdefinitionen verwendet, die nur bedingte Erwartungswerte vergleichen:

$$E(\dot{Y}|\ddot{X}=x'') - E(\dot{Y}|\ddot{X}=x') \quad (7)$$

Es muss berücksichtigt werden, dass  $\dot{Y}$  meistens auch noch von anderen Variablen abhängt. Effekte können dann nicht mehr nur einer Variablen zugerechnet werden, sondern müssen als *kontextabhängige Effekte* konzipiert werden. Zur Verdeutlichung sei angenommen, dass  $\dot{Y}$  (wie oben im Modell 1) nicht nur von  $\ddot{X}$ , sondern auch von  $\ddot{Z}$  abhängt. Der Effekt einer Differenz in  $\ddot{X}$  muss dann in der Form

$$E(\dot{Y}|\ddot{X}=x'', \ddot{Z}=z) - E(\dot{Y}|\ddot{X}=x', \ddot{Z}=z) \quad (8)$$

geschrieben werden und hängt im Allgemeinen von dem Kovariablenkontext ab, auf den durch  $\ddot{Z}=z$  Bezug genommen wird.

### 3.7 Erklärte Varianz und statistische Erklärungen

Wenn über Schätzergebnisse von Regressionsfunktionen berichtet wird, werden oft Angaben zur „erklärten Varianz“ gemacht. Es ist bemerkenswert, dass sich dieser Begriff für funktionale Modelle nicht eignet. Zur Verdeutlichung betrachte man ein einfaches Modell:  $\ddot{X} \longrightarrow \dot{Y}$ . Da es für  $\ddot{X}$  keine Verteilung und somit auch keine Varianz gibt, kann die Vorstellung einer „erklärten Varianz“ nicht angewendet werden.<sup>4</sup>

Das Konzept einer „erklärten Varianz“ kann jedoch definiert werden, wenn man ein funktionales Modell für statistische Erklärungen (wie sie in Abschnitt 1 charakterisiert wurden) verwendet. Man beginnt dann mit einer zweidimensionalen statistischen Variablen:  $(X, Y) : \Omega \longrightarrow \mathcal{X} \times \mathcal{Y}$  und verwendet  $P[X]$ , die Verteilung von  $X$ , und ein funktionales Modell  $\ddot{X} \longrightarrow \dot{Y}$ , um eine *statistische Variable*, im Folgenden  $\hat{Y}$  genannt, zu konstruieren, deren Verteilung die durch die Daten gegebene Verteilung von  $Y$  approximiert. Zwei Konstruktionsmethoden sind möglich.

- a) Man kann eine Regel zur Voraussage individueller Werte von  $\hat{Y}$  verwenden, zum Beispiel  $\hat{Y}(\omega) := E(\dot{Y}|\ddot{X} = X(\omega))$ .
- b) Man kann unmittelbar eine Verteilung von  $\hat{Y}$  konstruieren.

Da es meistens, insbesondere wenn  $Y$  eine qualitative Variable ist, unterschiedliche Möglichkeiten gibt, individuelle Werte vorauszusagen, folge ich hier (zur Illustration) der zweiten Methode. Die Konstruktion einer Verteilung von  $\hat{Y}$  beginnt dann mit einer bedingten Verteilung:

<sup>4</sup> Man kann zwar bedingte Varianzen  $V(\dot{Y}|\ddot{X} = x)$  betrachten, um die Unsicherheit von Voraussagen zu quantifizieren; sie haben aber mit der Idee einer „erklärten Varianz“ nichts zu tun.

$$P(\hat{Y}=y|X=x) := P(\dot{Y} = y|\ddot{X}=x) \quad (9)$$

wobei ich hier annehme, dass  $\dot{Y}$  und folglich  $\hat{Y}$  diskrete Variablen sind. Dann kann die bekannte Verteilung von  $X$  verwendet werden, um eine Verteilung von  $\hat{Y}$  abzuleiten:

$$P(\hat{Y}=y) = \sum_x P(\hat{Y}=y|X=x)P(X=x) \quad (10)$$

Schließlich gibt es zwei Überlegungen, um das Ergebnis zu beurteilen:

- Man kann die Verteilung  $P[\hat{Y}]$  mit der durch die Daten gegebenen Verteilung  $P[Y]$  vergleichen und als „Verteilungsanpassung“ („goodness of distributional fit“) interpretieren.
- Man kann denjenigen Teil der Varianz von  $\hat{Y}$ , der der Variation von  $X$  zurechenbar ist, berechnen und als „erklärte Varianz“ interpretieren.

### 3.8 Numerische Illustration

Zur Illustration verwende ich folgende fiktive Daten für das Modell 1:

Tab. 1: Fiktive Daten für Modell 1

$X$	$Z$	$Y = 0$	$Y = 1$
0	0	300	300
0	1	80	320
1	0	40	160
1	1	80	720

Die Verteilungsanpassung hängt von dem parametrischen Modell ab, das zur Approximation des funktionalen Modells verwendet wird. Wenn man ein saturiertes Modell verwendet, ist die Verteilungsanpassung vollständig; in diesem Beispiel:

$$\Pr(\dot{Y}=1|\ddot{X}=x, \dot{Z}=z) = P(Y=1|X=x, Z=z) \implies P[\hat{Y}] = P[Y] \quad (11)$$

Die Verteilungsanpassung wäre jedoch nicht perfekt, wenn man zum Beispiel ein Logitmodell ohne einen Interaktionsterm verwendet hätte. Und natürlich gibt es auch dann im Allgemeinen keine perfekte Anpassung, wenn die statistische Erklärung neue Daten betrifft, die sich von den für die Modellschätzung verwendeten Daten unterscheiden.

Schließlich kann man die „erklärte Varianz“ berechnen, d.h. den Anteil der Varianz von  $\hat{Y}$  (nicht von  $Y$ ), der – vermittelt über das funktionale Modell – der Variation von  $X$  zugerechnet werden kann. Da die gemeinsame Verteilung von  $X$  und  $\hat{Y}$  bekannt ist, kann man eine gewöhnliche Varianzzerlegung vornehmen (Rohwer & Pötter 2001: 103 f.):

$$V(\hat{Y}) = V[M(\hat{Y}|X)] + M[V(\hat{Y}|X)] \tag{12}$$

Der erste Teil kann als „erklärte Varianz“

$$V[M(\hat{Y}|X)] = \sum_x [M(\hat{Y}|X=x) - M(\hat{Y})]^2 P(X=x) \tag{13}$$

interpretiert werden; der zweite Teil ist die restliche Variation:

$$M[V(\hat{Y}|X)] = \sum_x V(\hat{Y}|X=x)P(X=x) \tag{14}$$

Wenn man in unserem Beispiel ein saturiertes Modell verwendet, findet man zunächst  $M(\hat{Y}) = M(Y) = 0.75$  und  $V(\hat{Y}) = V(Y) = 0.1875$ , und schließlich für die „erklärte Varianz“:  $V[M(\hat{Y}|X, Z)] = 0.0285$ , für die restliche Variation:  $M[V(\hat{Y}|X, Z)] = 0.1590$ , und als Anteil der „erklärten Varianz“:  $0.0285/0.1875 \approx 15\%$ .

#### 4 Beziehungen zwischen erklärenden Variablen

Wenn man Werte endogener Variablen voraussagen möchte, sind meistens mehrere für die Voraussage relevante Variablen – im Folgenden „erklärende Variablen“ genannt – zu verwenden. Es ist dann wichtig, auch Beziehungen zwischen diesen Variablen zu berücksichtigen. Denn von diesen Beziehungen hängt ab, *wie* die erklärenden Variablen *gemeinsam* zum Zustandekommen der Werte der endogenen Variablen beitragen. Als ein formaler Rahmen, um Beziehungen zwischen erklärenden Variablen zu beschreiben, eignen sich wiederum funktionale Modelle. Dabei muss zwischen Interaktionen und funktionalen Beziehungen unterschieden werden.

##### 4.1 Interaktionen

###### *Eine allgemeine Definition*

Ich verwende folgende allgemeine Definition: Zwei Variablen  $\ddot{X}$  und  $\ddot{Z}$  sind *interaktive Bedingungen* für eine von ihnen abhängige Variable  $\dot{Y}$ , wenn der Effekt einer Veränderung in  $\ddot{X}$  [ $\ddot{Z}$ ] von Werten von  $\ddot{Z}$  [ $\ddot{X}$ ] abhängt. Die Formulierung zeigt, dass das Vorhandensein von Interaktion auch davon abhängt, wie „Effekte“ definiert werden. Zur Illustration betrachte ich Modell 2

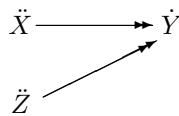


Abb. 2: Modell 2

wobei  $\dot{Y}$  ein Indikator für den schulischen Erfolg eines Kindes ist,  $\ddot{X}$  das Bildungsniveau seiner Eltern erfasst (0 niedrig, 1 hoch) und  $\ddot{Z}$  den Schultyp repräsentiert (0

oder 1). Anders als im Modell 1 gibt es in diesem Modell keine funktionale Beziehung zwischen  $\ddot{X}$  und  $\ddot{Z}$ , beide sind exogen.

Um den Effekt einer Differenz zwischen  $\ddot{X} = x'$  und  $\ddot{X} = x''$  zu erfassen, eignet sich folgende Definition:

$$\Delta^s(\dot{Y}; \ddot{X}[x', x''], \ddot{Z} = z) := E(\dot{Y} | \ddot{X} = x'', \ddot{Z} = z) - E(\dot{Y} | \ddot{X} = x', \ddot{Z} = z) \quad (15)$$

Verwendet man diese Definition, kann man leicht Beispiele sowohl mit als auch ohne eine Interaktion der beiden erklärenden Variablen angeben, wie etwa in Tabelle 2.

Tab. 2: Beispieldaten mit und ohne Interaktion

mit Interaktion			ohne Interaktion		
$x$	$z$	$E(\dot{Y}   \ddot{X} = x, \ddot{Z} = z)$	$x$	$z$	$E(\dot{Y}   \ddot{X} = x, \ddot{Z} = z)$
0	0	0.5	0	0	0.5
0	1	0.7	0	1	0.7
1	0	0.8	1	0	0.7
1	1	0.9	1	1	0.9

*Interaktion in parametrischen Modellen*

Die angegebene Definition einer Interaktion zwischen erklärenden Variablen ist unabhängig von der parametrischen Form, die zur Spezifikation eines funktionalen Modells gewählt wird. Sobald man parametrische Modelle verwendet, hängt es auch von deren Form ab, ob bzw. wie Interaktionen sichtbar gemacht werden können.

Lineare Modelle für Erwartungswerte erfordern eine explizite Formulierung von Interaktionseffekten (die dann meistens durch Produkte von Variablen definiert werden). Im Unterschied dazu implizieren die meisten nichtlinearen Modelle Interaktionen bereits durch ihre mathematische Form. Ein Logitmodell kann zur Verdeutlichung dienen. Für das Modell 2 kann ein Logitmodell beispielsweise so geschrieben werden:

$$\Pr(\dot{Y} = 1 | \ddot{X} = x, \ddot{Z} = z) \approx \frac{\exp(\alpha + x\beta_x + z\beta_z)}{1 + \exp(\alpha + x\beta_x + z\beta_z)} \quad (16)$$

In dieser Formulierung impliziert das Modell einen Interaktionseffekt, wenn man die Effektdefinition (15) verwendet, jedoch nicht, wenn man sich stattdessen auf sog. Odds Ratios

$$\frac{\Pr(\dot{Y} = 1 | \ddot{X} = x'', \ddot{Z} = z) / \Pr(\dot{Y} = 0 | \ddot{X} = x'', \ddot{Z} = z)}{\Pr(\dot{Y} = 1 | \ddot{X} = x', \ddot{Z} = z) / \Pr(\dot{Y} = 0 | \ddot{X} = x', \ddot{Z} = z)} \approx \exp((x'' - x')\beta_x) \quad (17)$$

bezieht. Natürlich kann es auch bei der Verwendung eines Logitmodells sinnvoll sein, zusätzlich einen expliziten Interaktionsterm aufzunehmen:

$$\Pr(\dot{Y} = 1 | \ddot{X} = x, \ddot{Z} = z) \approx \frac{\exp(\alpha + x\beta_x + z\beta_z + xz\beta_{xz})}{1 + \exp(\alpha + x\beta_x + z\beta_z + xz\beta_{xz})} \quad (18)$$

*Konsequenzen für das Verständnis von Effekten*

Wenn es bei zwei erklärenden Variablen eine Interaktion gibt, kann keiner von ihnen ein eindeutig bestimmter Effekt zugerechnet werden. Stattdessen muss man von *kontextabhängigen Effekten* sprechen. Dieses Konzept ist symmetrisch: Jede der beiden Variablen liefert einen Kontext für Effekte der jeweils anderen.

Zur Illustration verwende ich das Modell 2. Mit den Daten in der linken Hälfte von Tabelle 2 erhält man

$$\Delta^s(\dot{Y}; \ddot{X}[0, 1], \ddot{Z}=0) = 0.8 - 0.5 = 0.3 \quad (19)$$

$$\Delta^s(\dot{Y}; \ddot{X}[0, 1], \ddot{Z}=1) = 0.9 - 0.7 = 0.2 \quad (20)$$

Umgekehrt hängt der Effekt des Schultyps vom Bildungsniveau der Eltern ab.

*4.2 Funktionale Beziehungen*

Interaktionen betreffen die probabilistische Funktion, die zwei oder mehr erklärende Variablen mit einer von ihnen abhängigen Variablen verknüpft. Eine davon unabhängige Frage betrifft funktionale Beziehungen zwischen erklärenden Variablen.

Zur Illustration betrachte ich das Modell 1, in dem der Schultyp funktional vom Bildungsniveau der Eltern abhängt. Ob es in Bezug auf den schulischen Erfolg eine Interaktion des Schultyps und des Bildungsniveaus der Eltern gibt, ist vom Vorhandensein einer funktionalen Beziehung unabhängig.

*Mediator- und Moderatorvariablen*

In der Literatur wird oft zwischen Mediator- und Moderatorvariablen unterschieden (Baron & Kenny 1986; MacKinnon 2008). Im Kontext funktionaler Modelle können folgende Definitionen verwendet werden:

$\dot{Z}$  ist eine *Mediatorvariable* für  $\ddot{X}$  [oder  $\dot{X}$ ] bezüglich einer anderen Variablen  $\dot{Y}$ , wenn  $\dot{Z}$  auf einem gerichteten Weg liegt, der von  $\ddot{X}$  [oder  $\dot{X}$ ] zu  $\dot{Y}$  führt.

$\dot{Z}$  [oder  $\ddot{Z}$ ] ist eine *Moderatorvariable* bezüglich einer Beziehung zwischen  $\ddot{X}$  [oder  $\dot{X}$ ] und  $\dot{Y}$ , wenn der Effekt von  $\ddot{X}$  [oder  $\dot{X}$ ] auf  $\dot{Y}$  von Werten von  $\dot{Z}$  [oder  $\ddot{Z}$ ] abhängt.

Zum Beispiel ist im Modell 1  $\dot{Z}$  eine Mediatorvariable, und wenn es eine Interaktion von  $\dot{Z}$  und  $\ddot{X}$  gibt, ist  $\dot{Z}$  auch eine Moderatorvariable. Dagegen ist  $\ddot{Z}$  im Modell 2 keine Mediatorvariable, aber ggf. eine Moderatorvariable, wenn es eine Interaktion mit  $\ddot{X}$  gibt. Für die weiteren Überlegungen beziehe ich mich auf folgende Modellvarianten: Die leitende Frage ist, wie man in diesen Modellen von Effekten erklärender Variablen sprechen kann.

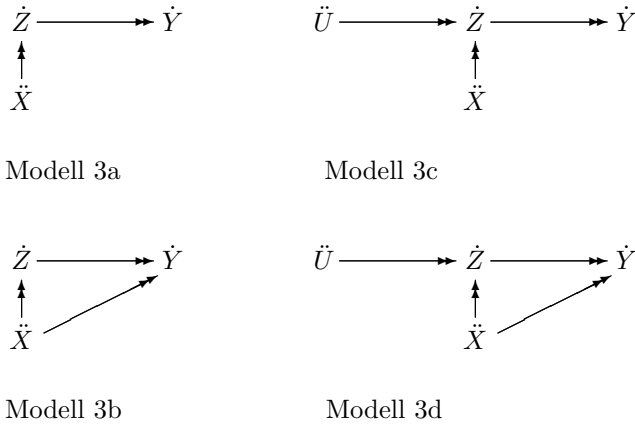


Abb. 3: Varianten zu Modell 3

*Effekte exogener Variablen*

In den Modellen 3a und 3b kann man die Mediatorvariable  $\dot{Z}$  eliminieren und einen Gesamteffekt der exogenen Variablen  $\ddot{X}$  berechnen. Ausgehend von

$$E(\dot{Y}|\ddot{X} = x) = \sum_z E(\dot{Y}|\ddot{X} = x, \dot{Z} = z) \Pr(\dot{Z} = z|\ddot{X} = x) \tag{21}$$

erhält man für Modell 3a den Gesamteffekt

$$\Delta^s(\dot{Y}; \ddot{X}[x', x'']) = \sum_z E(\dot{Y}|\dot{Z} = z) (\Pr(\dot{Z} = z|\ddot{X} = x'') - \Pr(\dot{Z} = z|\ddot{X} = x')) \tag{22}$$

und für Modell 3b den Gesamteffekt

$$\Delta^s(\dot{Y}; \ddot{X}[x', x'']) = \sum_z \left( E(\dot{Y}|\ddot{X} = x'', \dot{Z} = z) \Pr(\dot{Z} = z|\ddot{X} = x'') - E(\dot{Y}|\ddot{X} = x', \dot{Z} = z) \Pr(\dot{Z} = z|\ddot{X} = x') \right) \tag{23}$$

Um die Berechnung dieses Gesamteffekts zu illustrieren, verwende ich die Daten aus Tabelle 2 und nehme folgende Werte für die Abhängigkeit des Schultyps vom Bildungsniveau der Eltern an:

$$\Pr(\dot{Z} = 1|\ddot{X} = 0) = 0.4 \quad \text{und} \quad \Pr(\dot{Z} = 1|\ddot{X} = 1) = 0.8 \tag{24}$$

Mit den Daten aus der linken Hälfte der Tabelle 2 erhält man

$$\begin{aligned} E(\dot{Y}|\ddot{X} = 1) &= 0.8 \cdot 0.2 + 0.9 \cdot 0.8 = 0.88 \\ E(\dot{Y}|\ddot{X} = 0) &= 0.5 \cdot 0.6 + 0.7 \cdot 0.4 = 0.58 \\ \text{Gesamteffekt} &= 0.88 - 0.58 = 0.3 \end{aligned}$$



Mit den Daten aus der rechten Hälfte der Tabelle erhält man

$$\begin{aligned} E(\dot{Y}|\ddot{X}=1) &= 0.7 \cdot 0.2 + 0.9 \cdot 0.8 = 0.86 \\ E(\dot{Y}|\ddot{X}=0) &= 0.5 \cdot 0.6 + 0.7 \cdot 0.4 = 0.58 \\ \text{Gesamteffekt} &= 0.86 - 0.58 = 0.28 \end{aligned}$$

*Effekte endogener Variablen*

Wenn man Effekte endogener Variablen bestimmen möchte, tritt zunächst die Schwierigkeit auf, dass Verteilungen der Werte dieser Variablen bereits innerhalb des Modells von anderen Variablen abhängen. Wenn man annimmt, dass man gleichwohl Werte endogener Variablen hypothetisch fixieren kann, können auch Effekte endogener Variablen einfach definiert werden.

Zur Illustration beziehe ich mich auf die endogene Variable  $\dot{Z}$ . In den Modellen 3a und 3c kann man die Definition

$$\Delta^s(\dot{Y}; \dot{Z}[z', z'']) = E(\dot{Y}|\dot{Z}=z'') - E(\dot{Y}|\dot{Z}=z') \tag{25}$$

verwenden. In den Modellen 3b und 3d sind die Effekte kontextabhängig von Werten von  $\ddot{X}$ , so dass die Definition

$$\Delta^s(\dot{Y}; \dot{Z}[z', z''], \ddot{X}=x) = E(\dot{Y}|\dot{Z}=z'', \ddot{X}=x) - E(\dot{Y}|\dot{Z}=z', \ddot{X}=x) \tag{26}$$

zu verwenden ist.

*Direkte und indirekte Effekte*

Das Modell 3b (das dem Modell 1 entspricht) motiviert die weitere Frage, ob man den Gesamteffekt von  $\ddot{X}$  in einen direkten und einen indirekten Effekt aufspalten kann. Eine positive Antwort erfordert jedenfalls, dass man Werte von  $\dot{Z}$  hypothetisch konstant halten kann, obwohl ihre Wahrscheinlichkeitsverteilung von den Werten von  $\ddot{X}$ , die für die Effektdefinition verändert werden müssen, abhängt.

Aber selbst wenn man diese Annahme voraussetzt, kann man einen direkten Effekt von  $\ddot{X}$  nur definieren, wenn es keine Interaktion zwischen  $\ddot{X}$  und  $\dot{Z}$  gibt. In diesem Fall ist der Effekt

$$\Delta^s(\dot{Y}; \ddot{X}[x', x''], \dot{Z}=z) \tag{27}$$

unabhängig von  $z$  und kann deshalb sinnvoll als der direkte Effekt von  $\ddot{X}$  auf  $\dot{Y}$  interpretiert werden. Das wird durch die Daten in der rechten Hälfte der Tabelle 2 illustriert:

$$\Delta^s(\dot{Y}; \ddot{X}[0, 1], \dot{Z}=0) = \Delta^s(\dot{Y}; \ddot{X}[0, 1], \dot{Z}=1) = 0.2 \tag{28}$$

Und schließlich kann dann auch ein indirekter Effekt als Differenz zwischen dem Gesamteffekt und dem direkten Effekt definiert werden. Geht man von (23) aus, erhält man für den indirekten Effekt die Formulierung

$$\sum_z E(\dot{Y}|\ddot{X}=x', \dot{Z}=z) (\Pr(\dot{Z}=z|\ddot{X}=x'') - \Pr(\dot{Z}=z|\ddot{X}=x')) \tag{29}$$

Mit den Daten der rechten Hälfte der Tabelle 2 erhält man den indirekten Effekt 0.08, und der Gesamteffekt ist  $0.2 + 0.08 = 0.28$ .

Wenn es jedoch eine Interaktion zwischen  $\ddot{X}$  und  $\dot{Z}$  gibt, ist es nicht möglich, einen direkten Effekt eindeutig zu definieren, selbst wenn man annimmt, dass Werte von  $\dot{Z}$  konstant gehalten werden können. Dies wird durch die Daten in der linken Hälfte von Tabelle 2 illustriert:

$$\Delta^s(\dot{Y}; \ddot{X}[0, 1], \dot{Z}=0) = 0.3 \quad \text{und} \quad \Delta^s(\dot{Y}; \ddot{X}[0, 1], \dot{Z}=1) = 0.2 \quad (30)$$

Folglich gibt es dann auch keine eindeutige Zerlegung des Gesamteffekts in einen direkten und einen indirekten Anteil.

#### *Kontrafaktische Effektzerlegungen*

Wenn es eine Interaktion zwischen  $\ddot{X}$  und einer Mediatorvariablen  $\dot{Z}$  gibt, können direkte Effekte von  $\ddot{X}$  zunächst nur jeweils gesondert für bestimmte Werte von  $\dot{Z}$  definiert werden. Allerdings kann man davon ausgehend unterschiedliche Varianten durchschnittlicher direkter Effekte definieren. Dementsprechend haben einige Autoren einen sog. „natürlichen direkten Effekt“ (Pearl 2001; Petersen et al. 2006) vorgeschlagen. Im Kontext funktionaler Modelle kann folgende Formulierung verwendet werden:

$$\sum_z (\mathbb{E}(\dot{Y}|\ddot{X}=x'', \dot{Z}=z) - \mathbb{E}(\dot{Y}|\ddot{X}=x', \dot{Z}=z)) \Pr(\dot{Z}=z|\ddot{X}=x') \quad (31)$$

Die Idee ist, sich als Referenz auf die anfängliche Verteilung von  $\dot{Z}$ , die der Situation  $\ddot{X} = x'$  entspricht, zu beziehen und „kontrafaktisch“ anzunehmen, dass sich diese Verteilung trotz der Veränderung des Wertes von  $\ddot{X}$  nicht verändert.

Natürlich kann man auch andere Verteilungen von  $\dot{Z}$  hypothetisch voraussetzen. In jedem Fall kann dann der Gesamteffekt von  $\ddot{X}$  in einen mittleren direkten und einen mittleren indirekten Teil zerlegt werden. Ausgehend von (31) erhält man beispielsweise folgende Zerlegung

$$\begin{aligned} & \mathbb{E}(\dot{Y}|\ddot{X}=x'') - \mathbb{E}(\dot{Y}|\ddot{X}=x') = \\ & \sum_z [\mathbb{E}(\dot{Y}|\ddot{X}=x'', \dot{Z}=z) - \mathbb{E}(\dot{Y}|\ddot{X}=x', \dot{Z}=z)] \Pr(\dot{Z}=z|\ddot{X}=x') + \\ & \sum_z \mathbb{E}(\dot{Y}|\ddot{X}=x'', \dot{Z}=z) [\Pr(\dot{Z}=z|\ddot{X}=x'') - \Pr(\dot{Z}=z|\ddot{X}=x')] \end{aligned} \quad (32)$$

und kann dann den zweiten Term auf der rechten Seite als mittleren indirekten Effekt betrachten.

Um die Zerlegung zu illustrieren, verwende ich die Daten aus der linken Hälfte von Tabelle 2 zusammen mit (24). Der mittlere direkte Effekt ist dann  $(0.8 - 0.5)0.6 + (0.9 - 0.7)0.4 = 0.26$ , der mittlere indirekte Effekt ist  $0.8(0.2 - 0.6) + 0.9(0.8 - 0.4) = 0.04$ , und der Gesamteffekt ist  $0.26 + 0.04 = 0.3$ .

#### *Konfundierende Variablen*

In der Literatur wird in unterschiedlichen Bedeutungen von „konfundierenden Variablen“ gesprochen (Weinberg 1993). Für funktionale Modelle eignet sich folgende Definition:

$\dot{X}$  [oder  $\ddot{X}$ ] ist eine konfundierende Variable bezüglich einer Funktion, durch die  $\dot{Y}$  von  $\dot{Z}$  abhängt, wenn es einen gerichteten Pfad von  $\dot{X}$  [ $\ddot{X}$ ] zu  $\dot{Y}$  gibt und wenn es (a) einen gerichteten Pfad von  $\dot{X}$  [ $\ddot{X}$ ] zu  $\dot{Z}$  gibt, oder wenn es (b) eine weitere Variable gibt, etwa  $\dot{U}$  [oder  $\ddot{U}$ ], und gerichtete Pfade von  $\dot{U}$  [ $\ddot{U}$ ] sowohl zu  $\dot{Z}$  als auch zu  $\dot{X}$  führen.

Zum Beispiel ist  $\ddot{X}$  im Modell 3b eine konfundierende Variable bezüglich der Beziehung zwischen  $\dot{Z}$  und  $\dot{Y}$ . Andererseits ist in diesem Modell  $\dot{Z}$  keine konfundierende Variable, sondern eine Mediatorvariable bzgl. der Beziehung zwischen  $\ddot{X}$  und  $\dot{Y}$ . Es sei angemerkt, dass die vorgeschlagene Definition ein funktionales Modell mit funktionalen (gerichteten) Abhängigkeitsbeziehungen voraussetzt und nicht durch Verweise auf „Korrelationen“ ersetzt werden kann.

Durch die Definition werden konfundierende Variablen von *unabhängigen Kontextvariablen* unterschieden. Gäbe es zum Beispiel im Modell 3b keinen Pfeil von  $\ddot{X}$  zu  $\dot{Z}$ , wäre  $\ddot{X}$  für die Beziehung zwischen  $\dot{Z}$  und  $\dot{Y}$  keine konfundierende, sondern eine unabhängige Kontextvariable.

Die Unterscheidung ist wichtig für die Möglichkeiten, Effekte zu definieren. Ein wesentlicher Unterschied zeigt sich insbesondere dann, wenn es sich um nicht beobachtete Variablen handelt. Zur Illustration beziehe ich mich auf das Modell 2, in dem  $\ddot{X}$  eine unabhängige Kontextvariable für Effekte von  $\dot{Z}$  auf  $\dot{Y}$  ist:

$$\Delta^s(\dot{Y}; \ddot{Z}[z', z''], \ddot{X} = x) = E(\dot{Y} | \ddot{Z} = z'', \ddot{X} = x) - E(\dot{Y} | \ddot{Z} = z', \ddot{X} = x) \quad (33)$$

Wenn Werte von  $\ddot{X}$  nicht beobachtet werden können, kann man stattdessen eine Variable  $\dot{X}$  mit einer unbekanntem Verteilung annehmen. Da diese Verteilung qua Modelldefinition nicht von  $\dot{Z}$  abhängt, kann der beobachtbare Effekt von  $\dot{Z}$  jedenfalls als ein durchschnittlicher Effekt bzgl. der unbekanntem Verteilung von  $\dot{X}$  interpretiert werden:

$$\Delta^s(\dot{Y}; \ddot{Z}[z', z''], \dot{X}) = \sum_x \Delta^s(\dot{Y}; \ddot{Z}[z', z''], \dot{X} = x) \Pr(\dot{X} = x) \quad (34)$$

Und wenn es keine Interaktion zwischen  $\ddot{Z}$  und  $\dot{X}$  gibt, kann man diesen durchschnittlichen Effekt auch eindeutig der Variablen  $\dot{Z}$  zurechnen.

Eine ähnliche Überlegung ist jedoch nicht möglich, wenn es sich um eine konfundierende Variable handelt. Als Beispiel kann  $\dot{X}$  im Modell 3b dienen. Wenn diese Variable nicht beobachtet werden kann, kann man wiederum stattdessen eine Variable  $\dot{X}$  mit einer unbekanntem Verteilung annehmen, und ein Effekt von  $\dot{Z}$  auf  $\dot{Y}$  kann dann so formuliert werden:

$$E(\dot{Y} | \dot{Z} = z'') - E(\dot{Y} | \dot{Z} = z') = \sum_x \left( E(\dot{Y} | \dot{Z} = z'', \dot{X} = x) \Pr(\dot{X} = x | \dot{Z} = z'') - E(\dot{Y} | \dot{Z} = z', \dot{X} = x) \Pr(\dot{X} = x | \dot{Z} = z') \right) \quad (35)$$

Dieser Effekt resultiert jedoch nicht nur aus unterschiedlichen Werten von  $\dot{Z}$ , sondern außerdem aus den damit verbundenen unterschiedlich verteilten Werten von  $\dot{X}$ . Der Effekt kann deshalb nicht als ein durchschnittlicher Effekt interpretiert werden.

*Effekte von Mediatorvariablen*

Es ist bemerkenswert, dass die angeführten Schwierigkeiten bei Mediatorvariablen nicht auftreten. Anstelle von Modell 3b betrachte man Modell 4 (Abbildung 4), in dem es zusätzlich die Mediatorvariable  $\dot{V}$  gibt. Wie zuvor gilt: Wenn die konfundierende Variable  $\ddot{X}$  nicht beobachtet werden kann, kann die beobachtbare Beziehung zwischen  $\dot{Z}$  und  $\dot{Y}$  nur auf problematische Weise als ein Effekt interpretiert werden. Anders verhält es sich jedoch beim Effekt von  $\dot{V}$  auf  $\dot{Y}$ .

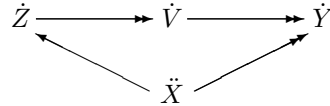


Abb. 4: Modell 4

Substituiert man  $\ddot{X}$  durch eine Variable  $\dot{X}$  mit einer unbekannten Verteilung, impliziert das Modell die Beziehungen

$$\Pr[\dot{Y}|\dot{X} = x, \dot{V} = v, \dot{Z} = z] = \Pr[\dot{Y}|\dot{X} = x, \dot{V} = v] \tag{36}$$

$$\Pr[\dot{V}|\dot{Z} = z, \dot{X} = x] = \Pr[\dot{V}|\dot{Z} = z] \tag{37}$$

und somit auch<sup>5</sup>

$$\Pr[\dot{X}|\dot{V} = v, \dot{Z} = z] = \Pr[\dot{X}|\dot{Z} = z] \tag{39}$$

Mit Hilfe dieser Beziehungen findet man

$$\begin{aligned} E(\dot{Y}|\dot{V} = v, \dot{Z} = z) &= \sum_x E(\dot{Y}|\dot{V} = v, \dot{Z} = z, \dot{X} = x) \Pr(\dot{X} = x|\dot{V} = v, \dot{Z} = z) \\ &= \sum_x E(\dot{Y}|\dot{V} = v, \dot{X} = x) \Pr(\dot{X} = x|\dot{Z} = z) \end{aligned} \tag{40}$$

und kann daraus

$$\begin{aligned} &E(\dot{Y}|\dot{V} = v'', \dot{Z} = z) - E(\dot{Y}|\dot{V} = v', \dot{Z} = z) \\ &= \sum_x [E(\dot{Y}|\dot{V} = v'', \dot{X} = x) - E(\dot{Y}|\dot{V} = v', \dot{X} = x)] \Pr(\dot{X} = x|\dot{Z} = z) \end{aligned} \tag{41}$$

ableiten. Diese Formulierung zeigt schließlich, dass sich Effekte von  $\dot{V}$  auf  $\dot{Y}$ , wenn man dafür Werte von  $\dot{Z}$  als konstant voraussetzt, durchaus als mittlere Effekte bzgl. einer unbekanntem Verteilung der nicht beobachteten Variablen  $\dot{X}$  interpretieren lassen. Und wiederum gilt auch, dass sich diese Effekte eindeutig der Variablen  $\dot{V}$  zurechnen lassen, wenn es keine Interaktion zwischen  $\dot{V}$  und  $\ddot{X}$  gibt.

<sup>5</sup> Geht man von

$$\Pr[\dot{X}, \dot{V}|\dot{Z}] = \Pr[\dot{X}|\dot{V}, \dot{Z}] \Pr[\dot{V}|\dot{Z}] = \Pr[\dot{V}|\dot{X}, \dot{Z}] \Pr[\dot{X}|\dot{Z}] \tag{38}$$

aus und verwendet (37), folgt sofort (39).

## 5 Grundlegende Literatur

Die in diesem Beitrag diskutierten funktionalen Modelle bilden eine Variante von Modellen, die von gerichteten azyklischen Graphen (DAGs) ausgehen. Eine grundlegende Monographie zu Modellen dieses Typs ist Pearl (2009). Dort wird auch ausführlich besprochen, wie diese Modelle für kausale Interpretationen verwendet werden können. Einen dafür sehr instruktiven Sammelband haben McKim & Turner (1997) herausgegeben; Beziehungen zur sozialwissenschaftlichen Forschung werden von Steel (2011) besprochen. Das für die Untersuchung von „Mechanismen“ wichtige Thema der Mediatorvariablen wird ausführlich von MacKinnon (2008) behandelt.

### Literaturverzeichnis

- BARON, R. M. UND D. A. KENNY (1986) „The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations.“ *Journal of Personality and Social Psychology* 51: 1173–1182.
- CROSS, C. B. (1991) „Explanation and the Theory of Questions.“ *Erkenntnis* 34: 237–260.
- FAYE, J. (1999) „Explanation Explained.“ *Synthese* 120: 61–75.
- FISHER, R. A. (1922) „On the Mathematical Foundations of Theoretical Statistics.“ *Philosophical Transactions of the Royal Society of London*. 222: 309–368.
- HEMPEL, C. G. (1965) *Scientific Explanation. Essays in the Philosophy of Science*. New York: Free Press.
- MACKINNON, D. P. (2008) *Introduction to Statistical Mediation Analysis*. New York: Lawrence Erlbaum.
- PEARL, J. (2001) „Direct and Indirect Effects.“ S. 411–420 in: J. BREESE UND D. KOLLER (HG.) *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence* San Francisco, CA: Morgan Kaufmann.
- PETERSEN, M. L., S. E. SINISI UND M. VAN DER LAAN (2006) „Estimation of Direct Causal Effects.“ *Epidemiology* 17: 276–284.
- ROHWER, G. (2010) *Models in Statistical Social Research*. London: Routledge.
- ROHWER, G. (2011a) „Qualitative Comparative Analysis. A Discussion of Interpretations.“ *European Sociological Review* 27: 728–740.
- ROHWER, G. (2011b) „Probleme der Generalisierung statistischer Aussagen.“ S. 135–145 in: D. FISCHER, W. BONSS, T. AUGUSTIN, F. BADER, M. PICHLBAUER UND D. VOGL (HG.) *Uneindeutigkeit als Herausforderung – Risikokalkulation, Amtliche Statistik und die Modellierung des Sozialen*. Neubiberg: Universität der Bundeswehr München.
- ROHWER, G. (2012a) „Factual and Modal Notions in Sozial Research.“ *Quality & Quantity*. doi: 10.1007/s11135-012-9786-9.
- ROHWER, G. (2012b) „Functional Models and Causal Interpretations.“ *NEPS Working Paper No. 9*. Bamberg: Otto-Friedrich Universität, Nationales Bildungspanel.
- ROHWER, G. UND U. PÖTTER (2001) *Grundzüge der sozialwissenschaftlichen Statistik*. Weinheim: Juventa.
- ROHWER, G. UND U. PÖTTER (2002) *Wahrscheinlichkeit. Begriff und Rhetorik in der Sozialforschung*. Weinheim: Juventa.
- WEINBERG, C. (1993) „Toward a Clearer Definition of Confounding.“ *American Journal of Epidemiology* 137: 1–8.

