.

Wachstumsprozesse, interagierende Populationen, Epidemien, und die Messung psychischer Merkmale.

Die logistische Funktion

U. Mortensen

Fassung 03. 11. 2017[1]

[1]Korrektur 10. 05. 2020

Inhaltsverzeichnis

1 Wachstumsprozesse **3**

1.1 Einführung 3

1.2 Exponentielles Wachstum 4

1.3 Der diskrete Fall 7

1.4 Logistisches Wachstum 10

1.5 Interagierende Populationen 14

2 Das klassische epidemische Modell **17**

2.1 Das ”triviale” Modell 17

2.2 Das klassische epidemiologische Modell 18

3 Die logistische Verteilung **22**

3.1 Die Definition der logistischen Verteilung 22

3.2 Logistische versus Gauß-Verteilung 24

3.3 Spezielle Eigenschaften der logistischen Verteilung 25

4 Regression und Klassifikation **27**

5 Diskussion **30**

6 Anhang **32**

1 Wachstumsprozesse

1.1 Einführung

Wenn eine Größe im Laufe der Zeit zunimmt, kann man von einem Wachstumsprozess sprechen. Wächst die Bevölkerung in einem Land, vermehren sich Bakterien in einer Petrischale, vermehrt sich ein Vermögen durch Verzinsung, nimmt die Verkehrsdichte auf einer Autobahn zu, so hat man es mit Wachstumsprozessen zu tun. Mathematisch kann man einen Wachstumsprozess durch eine Funktion der Zeit abbilden. Der Wachstumsprozess kann zufällige Elemente enthalten, so dass der Verlauf der beschreibenden Funktion der Trajektorie $x(t)$ mit $T_1 \leq t < T_2$ eines Zufallsprozesses gleicht. Dabei kann es gelingen, $x(t)$ in der Form $x(t) = f(t) + \varepsilon(t)$ darzustellen, wobei f eine deterministische Funktion, d.h. eine Funktion ohne zufällige Komponenten ist, und $\varepsilon(t)$ ist eine Funktion, die die zufälligen Aspekte in der betrachteten Größe ist; der Kurs einer Aktie kann u.U. in dieser Weise beschrieben werden. f wird dann auch als Trend bezeichnet. Im Folgenden werden zunächst nur Trendfunktionen betrachtet. Dabei wird angenommen, dass f eine differenzierbare Funktion der Zeit ist, d.h. es wird angenommen, dass für alle $t \in [T_1, T_2)$ der Limes[2]

$$\lim_{\Delta t \to 0} \frac{f(t + \Delta t) - f(t)}{\Delta t} = \frac{df(t)}{dt} = f'(t) \tag{1}$$

existiert.

Beispiel 1.1 Ein einfaches Beispiel ist das *lineare Modell*. Hier ist die Funktion f durch

$$f(t) = a \cdot t + b \tag{2}$$

gegeben. Für (1) erhält man

$$\lim_{\Delta t \to 0} \frac{a \cdot (t + \Delta t) + b - a \cdot t - b)}{\Delta t} = \lim_{\Delta t \to 0} \frac{a \Delta t}{\Delta t} = a, \tag{3}$$

d.h. die Bildung des Grenzwerts für $\Delta t \to 0$ ist trivial, da sich Δt herauskürzt. Der Zuwachs ist unabhängig von t stets gleich a. Man denke an eine Badewanne, die bei gleichbleibendem Zulauf von Wasser gefüllt wird. □

Häufig ist der Zuwachs aber nicht unabhängig von t; so kann der Zuwachs im Zeitpunkt t proportional zum Wert von $f(t)$ sein, so dass[3]

$$f(t + \Delta t) = f(t) + pf(t), \quad 0 \leq t < \infty \tag{4}$$

gilt. Hier ist p eine Zahl, von der sich zeigen läßt, dass sie im Allgemeinen von Δt abhängt, und $pf(t)$ ist ein Anteil von dem Wert, den f zum Zeitpunkt t annimmt. Dieser Ansatz wird im folgenden Abschnitt 1.1 diskutiert.

f kann unbegrenzt oder begrenzt wachsen; begrenztes Wachstum wird von Abschnitt 1.4 an diskutiert. Die Bevölkerung eines Landes wächst im Allgemeinen nicht unbegrenzt;

[2] Die Bezeichnung $[T_1, T_2)$ bedeutet $t_1 \leq t < T_2$, d.h. t kann den Wert T_2 nicht annehmen. Das ist z.B. der Fall, wenn $T_2 = \infty$ ist.

[3] Δ ist das große Delta, im Unterschied zum kleinen Delta, δ. Das große Delta Δ wird üblicherweise zur Bezeichnung eines Unterschieds bzw. einer Diffferenz verwendet.

von einem bestimmten Zeitpunkt an wächst sie immer langsamer und strebt asymptotisch gegen einen maximalen Wert $f_{\max} < \infty$. Unter Umständen wird sie auch kleiner, nachdem sie einen maximalen Wert erreicht hat. Wächst die Größe unbegrenzt über dem betrachteten Zeitintervall, so gilt $f'(t) > 0$ für alle $t \in [T_1, T_2)$. Wächst sie asymptotisch gegen einen maximalen Wert, so gilt $f'(t) \geq 0$; $f'(t) = 0$, wenn zum Zeitpunkt t das Maximum bereits erreicht wurde. Ein im Prinzip unbegrenztes Wachstum kann man bei einer Bakterienpopulation (Zellteilung) beobachten: die Population würde wachsen, wenn die Petrischale unbegrenzt wäre. Eine analoge Situation hat man, wenn man sagt, der Verkehr nähme jedes Jahr um einen bestimmen Prozentsatz zu. In der Realität wird es aber kein unbegrenztes Wachstum geben, weil irgendwann alle Autobahnen komplett mit Autos gefüllt sein werden und jede nicht schon von Häusern bedeckte Fläche von Autobahnen und Straßen überbaut sein wird. Die von den Autofabriken täglich produzierten Fahrzeuge müssen, weil es sonst keine freie Fläche mehr gibt, in immer höheren Türmen gestapelt werden, deren Höhenwachstum anfänglich durch ein lineares Modell der Form (2) modelliert werden kann, wobei aber letzlich von einem begrenzten Wachstum ausgegangen werden muß, da wegen der allgemeinen Verstopfung kein Material für Autos und Türme mehr angeliefert werden kann[4]. Die Population der durch einen Virus infizierten Personen wird spätestens dann nicht mehr wachsen, wenn alle suszeptiblen (ansteckbaren) Personen infiziert sind. Die Charakterisierung eines Wachstumsprozesses als unbegrenzt, wie sie im folgenden Abschnitt vorgestellt wird, wird also im Allgemeinen nur eine Approximation des Prozesses für einen ersten Teilabschnitt des Intervalles $[T_1, T_2)$ sein.

1.2 Exponentielles Wachstum

Vorbemerkung: Es sei zunächst daran erinnert, dass für[5] $g(t) = ce^{\lambda t}$, $c, \lambda \neq 0$ und $0 \leq t < \infty$

$$\frac{dg(t)}{dt} = g'(t) = \lambda c e^{\lambda t} = \lambda g(t) \tag{5}$$

gilt. Umgekehrt folgt aus der Differentialgleichung $g'(t) = \lambda g(t)$, dass $g(t) = ce^{\lambda t}$. Wie in der Theorie der Differentialgleichungen gezeigt wird, ist die Lösung für f eindeutig (Satz von Picard-Lindelöf).

Satz 1.1 *Die Funktion f repräsentiere eine Größe, die mit der Zeit t wachse, so dass (i) $f(t) \geq 0$, (ii) $f'(t) = df(t)/dt > 0$ für $t \geq 0$; insbesondere gelte*

$$f(t + \Delta t) = f(t) + p(\Delta t) f(t), \quad 0 \leq \Delta t < \infty \tag{6}$$

mit $p(\Delta t) > 0$ für $\Delta t > 0$ und p sei mindestens einmal differenzierbar bezüglich Δt. Dann gilt

$$f(t) = ce^{\lambda t}, \quad c > 0. \tag{7}$$

[4] Anhänger der Theorie, dass nur ökonomisches Wachstum unseren Wohlstand sichert, können allerdings auf die Möglichkeit verweisen, ein zweites, nun aber unterirdisches Autobahnsystem zu bauen, und/oder ein Autobahnsystem auf Stelzen über dem alten zu errichten, etc. Dieser Ansatz kann hier nicht weiter vertieft werden.

[5] λ = lambda, griechischer Buchstabe, entspricht dem lateinischen ℓ.

Anmerkung: Der Proportionalitätsfaktor p muß notwendig als eine Funktion von Δt angenommen werden, soll nicht $f(t) = f_0$, f_0 eine Konstante gelten. Denn der Ansatz (6) impliziert $f(t + \Delta t) = (1 + p)f(t)$, und p =konstant bedeutet $df/d\Delta t = 0$ für alle Δt, d.h. f verändert sich nicht mit Δt und damit auch auch nicht mit t, entgegen der Annahme, dass f einen Wachstumsprozess ($df(t)/dt > 0$) beschreiben soll. □

Beweis: Zunächst wird gezeigt, dass $f(0) \neq 0$. Denn sei $f(0) = 0$. Dann folgt $f(\Delta t) = f(0) + p(\Delta t)f(0) = 0$ für alle $\Delta t > 0$ im Widerspruch zu $f'(t) > 0$; also folgt $f(0) > 0$ wegen (i). Dann folgt aber $p(0) = 0$. Denn es sei $\Delta t = 0$, so dass $f(t) = f(t) + p(0)f(t)$, also $0 = p(0)f(t)$, und wegen $f(t) > 0$ für $0 \leq t < \infty$ folgt $p(0) = 0$.

Die Funktion $p(\Delta t)$ ist zunächst nicht bekannt, aber man kann eine Reihenentwicklung (Taylor-Reihe) für p ansetzen[6], d.h. es möge

$$p(\Delta t) = p(0)+\Delta t p'(0)+\frac{\Delta t^2}{2!}p''(0)+\frac{\Delta t^3}{3!}p'''(0)+\cdots = \Delta t p'(0)+\frac{\Delta t^2}{2!}p''(0)+\frac{\Delta t^3}{3!}p'''(0)+\cdots$$

gelten, wobei p', p'', p''' etc die erste, die zweite, die dritte etc. Ableitung von p nach Δt sind. Für hinreichend kleine Werte von Δt gilt dann die Approximation $p(\Delta t) \approx \Delta t p'(0)$, woraus wegen $p(\Delta t) \neq 0$ auch $p'(0) \neq 0$ folgt. Aus (6) ergibt sich mithin nach Division durch $\Delta t > 0$ die Diffferenzengleichung

$$\frac{f(t + \Delta t) - f(t)}{\Delta t} = \frac{\Delta t p'(0)}{\Delta t} f(t) = p'(0)f(t),$$

woraus sich durch den Grenzübergang $\Delta t \to 0$ der Differentialquotient

$$f'(t) = \lim_{\Delta t \to 0} \frac{f(t + \Delta t) - f(t)}{\Delta t} = \lambda f(t), \quad \lambda = p'(0)$$

ergibt, und diese Differentialgleichung hat bekanntlich die Lösung $f(t) = ce^{\lambda t}$ (vergl. (5)), also (7) mit $f(0) = c > 0$. □

Anmerkungen:
1. **Bedeutung des Parameters λ I:** Offenbar gilt

$$p'(0) = \frac{f'(t)}{f(t)} = \frac{\lambda c e^{\lambda t}}{c e^{\lambda t}} = \lambda \,\text{für alle } t \geq 0, \tag{8}$$

d.h. λ ist gleich der *Veränderung* von f an der Stelle t relativ zum Wert von f an der Stelle t. Gleichzeitig ist $\lambda = p'(0)$. Während $p(\Delta t) = 0$ ist für $\Delta t = 0$, ist die *Veränderung* von p für $\Delta t = 0$ ungleich Null und unabhängig von t.
2. **Die Form der Funktion $p(\Delta t)$:** Die Funktion $p(\Delta t)$ kann nun explizit angegeben werden: aus (6) folgt insbesondere für $t = 0$[7]

$$p(\Delta t) = \frac{f(\Delta t) - f(0)}{f(0)} = \frac{ce^{\lambda \Delta t} - f(0)}{f(0)} = .\frac{c(e^{\lambda \Delta t} - 1)}{c} = e^{\lambda \Delta t} - 1. \tag{9}$$

□

[6] Wenn man nicht mehr weiter kann, schaut man sich die Reihe an!

[7] Man kann den Ausdruck für beliebigen t-Wert anschreiben; für den Fall $t = 0$ werden die Ausdrücke einfacher, weil $f(0) = c$.

Verdoppelungszeit: Natürlich kann man (6) für feste Zeitabschnitte Δt betrachten, etwa für Δt eine Konstante. Insbesondere kann man fragen, für welchen Wert von $\Delta t = \Delta t_d$ sich $f(t)$ verdoppelt, so dass $f(t + \Delta t_d) = 2f(t)$. Unter Berücksichtigung von (7) findet man

$$f(t + \Delta t_d) = ce^{\lambda(t+\Delta t_d)} = ce^{\lambda t}e^{\lambda \Delta t_d} = 2ce^{\lambda t},$$

d.h.

$$e^{\lambda \Delta t_d} = 2,$$

so dass

$$\Delta t_d = \frac{\log_e 2}{\lambda}. \tag{10}$$

Δt_d ist demnach unabhängig von t und damit auch unabhängig vom Startwert $f(t_0)$. Δt_d heißt *Verdopplungszeit* (doubling time).

Es ist $\log 2 = .693147 \approx .7$, und man erhält aus (10)

$$\Delta t_d \approx \frac{.7}{\lambda}. \tag{11}$$

So erhält man für $\lambda = .05$ den Wert $\Delta t_d = 14.00$. Hat man z.B. ein Kapital von $f(t_0)$ und bekommt man pro Jahr 5% Zinsen darauf, so braucht es

$$\Delta t_d \approx \frac{.7 \times 100}{.05 \times 100} = \frac{70}{5} = 14$$

Jahre, bis man sein Kapital verdoppelt hat. Nach weiteren 14 Jahren hat man es um weiteres Mal verdoppelt und besitzt nun $4f(t_0)$, nach weiteren 14 Jahren besitzt man $8f(t_0)$, etc. Allgemein hat man nach k Verdoppelungszeiten sein Kapital um den Faktor 2^k vervielfacht.

Natürlich kann (11) verwendet werden, um den Wert von λ aus beobachteten Verdopplungszeiten zu schätzen:

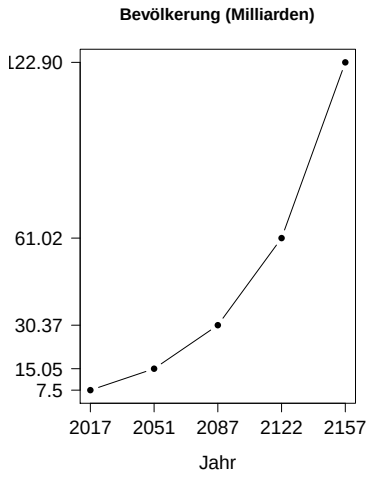
$$\hat{\lambda} \approx \frac{.7}{\Delta t_d}. \tag{12}$$

Beispiel 1.2 Verbrauch endlicher Resourcen: Analoge Betrachtungen gelten für den wachsenden Verbrauch von Ressourcen. Im Jahre t_0 habe man $x_0 = f(t_0)$ Barrel Öl verbraucht. Jedes Jahr wachse der Verbrauch um, sagen wir 1.5%, d.h. $\lambda = .015$. Die Verdoppelungszeit ist die Zeit, bis zu der sich der Verbrauch verdoppelt hat:

$$\Delta t_d \approx \frac{70}{1.5} \approx 47 \text{ Jahre}$$

In 47 Jahren hat die Welt dann soviel Öl verbraucht wie *in der gesamten Zeit* bis t_0, bis zu dem Zeitpunkt also, bis zu dem x_0 verbraucht wurde. Nach weiteren 47 hat man die Menge $4x_0$ verbraucht, d.h. man hat nur in diesem Zeitabschnitt so viel Öl verbraucht, wie insgesamt bis zum Zeitpunkt $2x_0$ konsumiert wurde. Dieses exponentielle Wachstum hat seine Tücken. Denn angenommen, man habe nach k Verdoppelungszeiten gerade die Hälfte der verfügbaren Ressourcen verbraucht. Intuitiv wird oft angenommen, dass man noch einmal den gesamten Zeitraum, der seit Beginn des Verbrauchs der Ressourcen verflossen ist, um die zweite Hälfte zu verbrauchen. Aber dies ist ein Trugschluß: man hat

Abbildung 1: Bevölkerungszuwachs in Millliarden; $\lambda = .02$, die Verdoppelungszeit beträgt 35 Jahre



nur nocht *eine* Verdoppelungszeit. Also: Benötigt man für den Verbrauch einer Ressource n Verdoppelungszeiten, so hat man nach $n-1$ Zeiten die Hälfte verbraucht und in der letzten, n-ten Δt_d-Zeit verbraucht man den Rest! □

Beispiel 1.3 Populationswachstum In Abb. 2, Seite 11, wird das Malthussche Gesetz illustiert, demzufolge die Weltbevölkerun exponentiell wächst. Abb. 1 illustriert das Wachstum einer Population als Funktion der Verdoppelungszeiten. Zum Zeitunkt $t = 2017$ gibt es ca 76.5 Mi8lliarden MEnschen auf der Erde. Schätzungen zufolge beträgt dier Verdoppelungszeit ca. 35 Jahre, was einem $\lambda = .02$ entspricht. Nach vier Verdoppelungszeiten , also nach 140 Jahren, ist die Bevölkerung bei nahezu 123 Milliarden. Berachtet man dazu den erwarteten Verlust an Landfläche durch den durch den Klimawandel bedingten Anstieg des Meeresspiegels, so kann man sagen, dass es eng wird, – es sei denn, die Wachstumsrate wird drastisch reduziert bzw. andere Wachstumsmodelle wie das in Abschnitt 1.4 vorgestellte logistische Modell beschreiben das Wachstum.

Es liegt nahe, zu vermuten, dass in den kommenden Jahrzehnten die Bedingungen für ein exponentielles Wachstumm nicht mehr gegeben sein werden. □

1.3 Der diskrete Fall

Oft wird das Wachstum einer Größe nach festen Zeitintervallen bestimmt, etwa das Wirtschaftswachstum, das nach jeweils einem Jahr bestimmt wird. Es sei α das durchschnittliche Wachstum nach einer Zeiteinheit; es gilt dann

$$x(t+1) = x(t) + \alpha x(t) = (1+\alpha)x(t). \tag{13}$$

Beginnt man zum Zeitpunkt $t = t_0$, so hat man

$$x(t_0 + 1) = (1 + \alpha)x(t_0),\ x(t_0 + 2) = (1 + \alpha)x(t_0 + 1) = (1 + \alpha)^2 x(t_0),$$

so dass man nach n Zeitintervallen

$$x(t_0 + n) = (1 + \alpha)^n x(t_0) \tag{14}$$

erhält.

Einen auf Jakob Bernoulli (1655 – 1701) zurückgehenden Spezialfall erhält man für $x(t_0) = 1$ und $\alpha = 1$; $x(t_0)$ entspreche einem Euro und $\alpha = 1$ sei die Verzinsung, die nach einem Jahr ausgezahlt wird. Dann ist $x(t_1) = 2$. Bernoulli betrachtete die Auszahlung für den Fall einer halbjährlichen Auszahlung, für die dann $\alpha = 1/2$ ist; man erhält

$$1 \cdot \left(1 + \frac{1}{2}\right)^2 = 1.5^2 = 2.25.$$

Bei einer monatlichen Auszahlung ergibt sich am Ende des Jahres

$$1 \cdot \left(1 + \frac{1}{12}\right)^{12} = 2.613035...$$

Wird halbmonatlich ausgezahlt, so erhält man

$$1 \cdot \left(1 + \frac{1}{24}\right)^{24} = 2.663731...$$

etc, und für $n = 1000$ findet man

$$1 \cdot \left(1 + \frac{1}{1000}\right)^{1000} = 2.716924...$$

Es läßt sich zeigen, dass

$$\lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = e = 2.71828\ldots \tag{15}$$

e ist eine irrationale[8] Zahl und gehört zu den fundamentalen Konstanten in der Mathematik. Mit $h = 1/n$ hat man

$$\lim_{h\to 0} (1 + h)^{1/h} = e$$

schreiben, und da auch $xh \to 0$ für $h \to 0$, erhält man für $\alpha = x$

$$\lim_{h\to 0} \left((1 + xh)^{1/xh}\right)^x = e^x. \tag{16}$$

Für e^x wird auch $\exp(x)$ geschrieben.

[8]Eine Zahl x heißt *rational* (von lat. ratio = Bruch oder Quotient), wenn sie als Quotient p/q zweier natürlichen Zahlen p und q drgestellt werden kann (natürliche Zahlen: $1, 2, 3, \ldots$). Sie heißt *irrational*, wenn sie nicht durch einen Quotienten der Form p/q darstellbar ist. So sind $\sqrt{2}$ und π irrationale Zahlen. Irrationale Zahlen haben unendlich viele Dezimalstellen ohne Periodizität.

Eine alternative Herleitung von (16) ergibt sich aus (14); es ist

$$\frac{x(t_0+n)}{x(t_0)} = (1+\alpha)^n. \tag{17}$$

Der Quotient $x(t_0+n)/x(t_0)$ ist unabhängig von x bzw. von t_0 und hängt nur vom Zuwachs ab. Nun sei also x eine stetige Funktion t, und es soll

$$\frac{x(t+h)}{x(t)} = \phi(h), \quad \text{für alle } h > 0$$

gelten, wobei ϕ eine noch unbekannte Funktion von h, aber nicht von x ist; sie ersetzt $(1+\alpha)^n$. Dann folgt zunächst

$$x(t+h) = \phi(h)x(t).$$

und subtrahiert man $x(t)$ von beiden Seiten und dividiert durch h, so erhält man

$$\frac{x(t+h)-x(t)}{h} = \frac{\phi(h)x(t)-x(t)}{h} = x(t)\frac{\phi(h)-1}{h}.$$

Für $h \to 0$ geht die linke Seite in

$$\lim_{h\to 0} \frac{x(t+h)-x(t)}{h} = \frac{dx(t)}{dt} = x'(t)$$

über, und wenn man voraussetzt, dass der Limes

$$\lim_{h\to 0} \frac{\phi(h)-1}{h} = a$$

existiert, erhält man wieder

$$x'(t) = ax(t), \quad \text{d.h. } x(t) = ce^{at}. \tag{18}$$

Man kann wieder nach der Anzahl n_d von Zeiteinheiten fragen, nach denen sich $x(t_0)$ verdoppelt hat. Es ist

$$2x(t_0) = (1+\alpha)^{n_d}x(t_0),$$

so dass

$$n_d = \frac{\log 2}{\log(1+\alpha)}. \tag{19}$$

Zur Bedeutung des Parameters λ II: Bakterien teilen sich, und dieser Prozess impliziert exponentielles Wachstum. Die meisten Tiere einschließlich der Menschen teilen sich nicht bei der Vermehrung, sondern erzeugen weitere Mitglieder der Population und sterben. Gestorbene sind aber keine Mitglieder der Population mehr, und dieser Sachverhalt muß bei der Beschreibung des Wachstums einer Population berücksichtigt werden. Außerdem kann eine Population durch Zuwanderung wachsen bzw. sich durch Abwanderung verringern.

Es sei $N(t)$ der Wert einer Population (genauer: die Anzahl der Mitglieder einer Population) zur Zeit t. Die Rate der Veränderung von N zur Zeit t ist durch

$$\frac{dN(t)}{dt} = \text{Geburten - Todesfälle + Migration} \tag{20}$$

gegeben; Migration repräsentiert die Ab- und Zuwanderungen. (20) heißt auch *Konservierungsgesetz* für die Population: offenbar ist $dN/dt = 0$ wenn

$$\text{Geburten} = \text{Todesfälle - Migration} \tag{21}$$

gilt; gilt diese Beziehung, verändert sich die Größe N der Population nicht.

Um den tatsächlichen Verlauf von $N(t)$ zu bestimmen, müssen weitere Annahmen über die rechte Seite von (20) gemacht werden. Ein erster Ansatz besteht darin, die Migration zu vernachlässigen und die Zahl der Geburten und der Todesfälle jeweils proportional zu N anzunehmen. Man erhält dann

$$\frac{dN(t)}{dt} = aN(t) - bN(t) = (a-b)N(t) = \lambda N(t) \tag{22}$$

wobei a der Proportionalitätsfaktor für die Geburten und b der Proportionalitätsfaktor für die Todesfälle ist. $a - b$ gibt dann den "Nettoeffekt" an: für $a - b > 0$ überwiegen die Geburten und die Population wächst, für $a - b < 0$ überwiegen die Todesfälle und die Population stirbt aus. Mit $\lambda = a - b$ hat man also die Gleichung $dN/dt = N_0\lambda N(t)$, $N_0 = N(0)$. Die Gleichung hat die Lösung (vergl. Bemerkung zum Parameter λ auf Seite 5)

$$N(t) = N_0 e^{\lambda t}. \tag{23}$$

Sie wurde zunächst durch Malthus[9] bekannt, ist aber wohl vorher schon von Euler diskutiert worden. In Abb. 2 wird das Bevölkerungswachstum von von 10 000 vChr bis heute angegeben; die Daten legen ein bisher exponentielles Wachstum nahe.

1.4 Logistisches Wachstum

Den gezeigten Daten nach ist scheint der Malthussche Ansatz nicht gar so falsch zu sein. Man kann aber davon ausgehen, daß das Wachstum irgendwann langsamer wird. Daten für einzelne Nationen bestätigen diese Vermutung: in den meisten untersuchten Ländern - außer den europäischen Ländern sind auch Länder wie Japan, Malawi, Ägypten etc untersucht worden - zeigen die Wachstumskurven Sättigungen, die in den globalen Daten der Abbildung 2 untergehen.

Das Wachstum ist exponentiell, wenn die Wachstumsrate λ eine Konstante ist. Soll das Wachstum nach einer gewissen Zeit gegen Null gehen, so muß die Wachstumsrate eine Funktion der Zeit sein, also $\lambda = \lambda(t)$, und $\lambda(t) \to 0$ für größer werdende Zeit t, denn für $\lambda = a = 0$ ist ja $dx(t)/dt = 0$ in (7). Der belgische Mathematiker Verhulst[10] wurde (1838) von der französischen Regierung beauftragt, das Wachstum der Bevölkerung von Paris abzuschätzen, da die Anzahl der Wohnungen (französisch *logis*) in Paris entsprechend

[9] Malthus, T.R. (1798) An Essay on the Principle of Population. Nachdruck bei Penguin (1979)
[10] Pierre-François Verhulst (1804 – 1849)

Abbildung 2: Wachstum der Weltbevölkerung, (a) McEvedy and Jones 1978, UN 1994, (b) Daten aus Murray (1989)

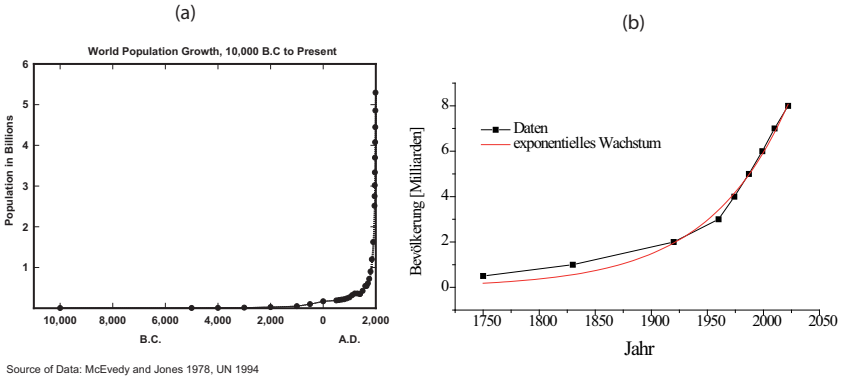


Abbildung 3: Pierre Verhulst (1804 - 1849)

wachsen mußte. Verhulsts Ansatz ist dementsprechend als der *logistische Modell* bekannt. Verhulst nahm an, dass die Bevölkerung nicht größer als ein bestimmter, kritischer Wert K sein könne, da die Bevölkerung ja aus der Umgebung mit Nahrung und Wasser versorgt werden muß; die Umgebung kann demnach eine Maximalzahl K gerade noch "tragen", weshalb K auch als *Trägerkonstante* bekannt ist. Ist $N(t)$ die Größe der Population zum Zeitpunkt t, so soll demnach $N(t)/K \leq 1$ sein. Also kann man (Verhulst)

$$\lambda(t) = r\left(1 - \frac{N(t)}{K}\right) \tag{24}$$

setzen, wobei r eine geeignet gewählte Konstante ist. Mit $N(t) \to K$ strebt $\lambda(t)$ gegen 0. Man macht nun wieder den Ansatz (7), ersetzt aber die Konstante a durch $\lambda(t)$ und

erhält nun die Gleichung

$$\frac{dN(t)}{dt} = rN(t)\left(1 - \frac{N(t)}{K}\right), \quad r > 0, \quad K > 0. \tag{25}$$

In diesem Modell ist die Geburtsrate pro Kopf durch $r(1 - N/K)$ gegeben; sie ist also nicht konstant, wie λ in (22) und (23), sondern hängt von N ab. (25) charakterisiert das *logistische Wachstum*.

Natürlich kann man fragen, warum man gerade die Funktion $\lambda(t)$ in (24) als Wachstumsrate annehmen soll. Eine Plausibilitätsbetrachtung zu (25) liefert eine erste Motivation: Multipliziert man die Gleichung aus, erhält man

$$\frac{dN(t)}{dt} = rN(t) - \frac{rN(t)N(t)}{K}.$$

Der erste Term auf der rechten Seite, $rN(t)$, impliziert exponentielles Wachstum. Der zweite Term – $rN^2(t)/K$ – impliziert, dass die Veränderung, repräsentiert durch dN/dt, proportional zur momentanen Größe $N(t)$ der Bevölkerung reduziert wird, wobei der Propotionalitätsfaktor durch $rN(t)/K$ gegeben sein soll. Die Veränderung wird also proportional zur momentan gegebenen Bevölkerungsgröße $N(t)$ reduziert. Eine weitergehende Begründung erhält man, wenn man spezifische Merkmale der Bevölkerung betrachtet, vergl. den Abschnitt 1.5.

Systeme wie (22) und (25) haben verschiedene Gleichgewichtslagen. Dies sind Werte für N, bei denen sich N nicht mit der Zeit ändert, für die also $dN(t)/dt = 0$ gilt. Für das exponentielle Wachstum (22) gibt es nur eine, nämlich $N(0) = N_0 = 0$. Wird das System, d.h. die Population aus dieser Gleichgewichtslage gebracht, wächst es unbeschränkt für $\lambda > 0$. Man sagt, daß System sei *instabil* für $\lambda > 0$. Für $\lambda < 0$ kehrt es in die Gleichgewichtslage zurück; das System ist dann *stabil*.

Für (25) ist die Bedingung $dN(t)/dt = 0$ für zwei Werte von N erfüllt: (i) $N(t) = 0$, oder (ii) $N(t) = N_\infty$. Im Falle (i) hat die Population gar kein Mitglied mehr, sie ist entweder ausgestorben oder hat nie existiert. Ist andererseits $N(t) > 0$, so impliziert (25), dass $dN(t)/dt > 0$ so lange $1 - N(t)/K > 0$ (und natürlich $r > 0$), also $1 > N(t)/K$ bzw. $K > N(t)$. Für $N(t) \to K$ folgt $1 - N(t)/K \to 0$ und damit $dN(t)/dt \to 0$. Für

$$N(t) = N_\infty = K. \tag{26}$$

ist $dN(t)/dt = 0$ und die Bevölkerung bleibt konstant. K heißt dementsprechend die *Trägerkapazität* (*carrier capacity*) und reflektiert die Ressourcen, von denen die Population lebt. Die Lage (26) ist stabil: weicht N nach unten von K ab, kehrt der Wert von N wieder zu K zurück. Die andere Möglichkeit für einen Gleichgewichtzustand ist $N(t) = 0$. Diese Lage ist instabil: die kleinste Auslenkung zu $N \neq 0$ impliziert das Wachstum gegen K. Abb. 4 zeigt das Wachstum der Lebenserwartung in Norwegen zwischen den Jahren 1840 und 1990: auch diese Variable folgt sehr gut dem logistischen Modell. Die Werte der Parameter r und K repräsentieren die Lebensbedingungen der Population. Natürlich kann man auch den Fall $r < 0$betrachten; für $1 - N(t)/K > 0$ ist dann $dN(t)/dt < 0$, d.h. $N(t)$ wird *kleiner* mit wachsendem t, bis schließlich $N(t) = 0$, d.h. bis die Population ausgestorben ist.

Abbildung 4: Wachstum der Lebenserwartung in Norwegen zwischen 1840 und 1990

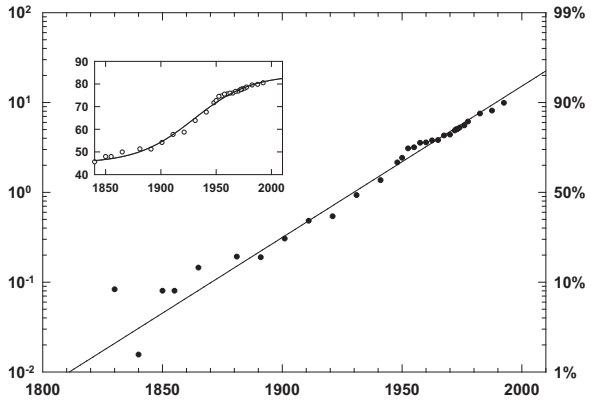
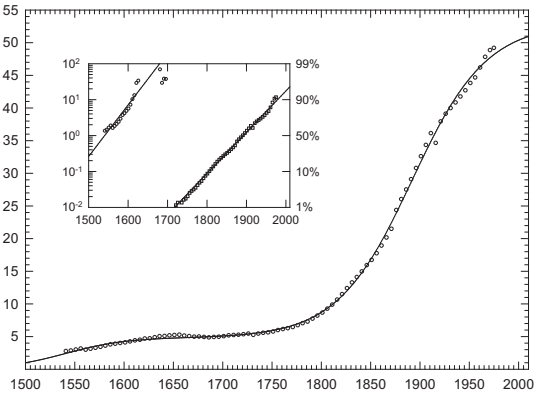


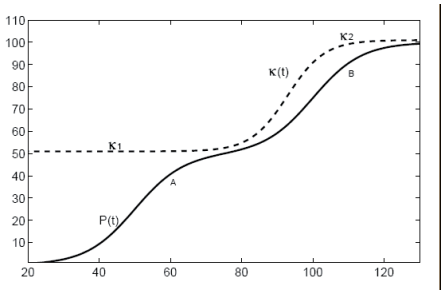
Abbildung 5: Die Population von England: Bi-logistisches Wachstumsmodell, 1541-1975



Verändern sich die Lebensbedingungen, so kann ein neuer Wachstumsprozess beginnen, der über N_∞ hinausführt. Das bi-logistische Modell beschreibt einen solchen fortgesetzten Wachstumsprozess:

Bi-logistisches Wachstumsmodell: Ein Beispiel ist das Bevölkerungswachstum von England zwischen den Jahren 1541 und 1975, vergl. Abb. 5. Das Interessante an diesen Daten ist, dass ein erster Wachstumsprozess gegen 1650 abgeschlossen war, und erst gegen 1750 ein neuer Prozess, diesmal aber mit anderen Parametern startete. Abbildung 6 zeigt eine logistsche Wachsturmskurve bei zeitabhängigem Wert von K.

Abbildung 6: Bi-logistisches Wachstumsmodell: K als Funktion der Zeit. Nach Meyer & Ausubel (1999).



1.5 Interagierende Populationen

Modifikationen von (25) erlauben, eine Reihe von dynamischen Prozessen zu interpretieren (d.h. zu verstehen), einschließlich physiologischer Prozesse. Eine interessante Erweiterung besteht in der Betrachtung *interagierender Populationen*, etwa B und R. B kann z.B. eine Population von Beutefischen sein, und R eine Population von Raubfischen; Volterra (1926) hat insbesondere diese Wechselwirkungen zwischen dieser Art von Populationen in der Adria betrachtet. Man spricht dementsprechend von Räuber-Beute-Systemen. Das von Volterra vorgeschlagene Modell ist der Ausgangspunkt für eine Klasse von Modellen, mit denen sich u.a. die Interaktionen von Neuronenpopulationen, oder die Interaktionen von Pro- und Opponenten in bezug auf eine bestimmte Meinung, beschreiben lassen. Hier soll nur der Ansatz vorgestellt werden. Der besteht in den Gleichungen

$$\frac{dN(t)}{dt} = N(t)(a - bP(t)) \tag{27}$$

$$\frac{dP(t)}{dt} = P(t)(cN(t) - d), \tag{28}$$

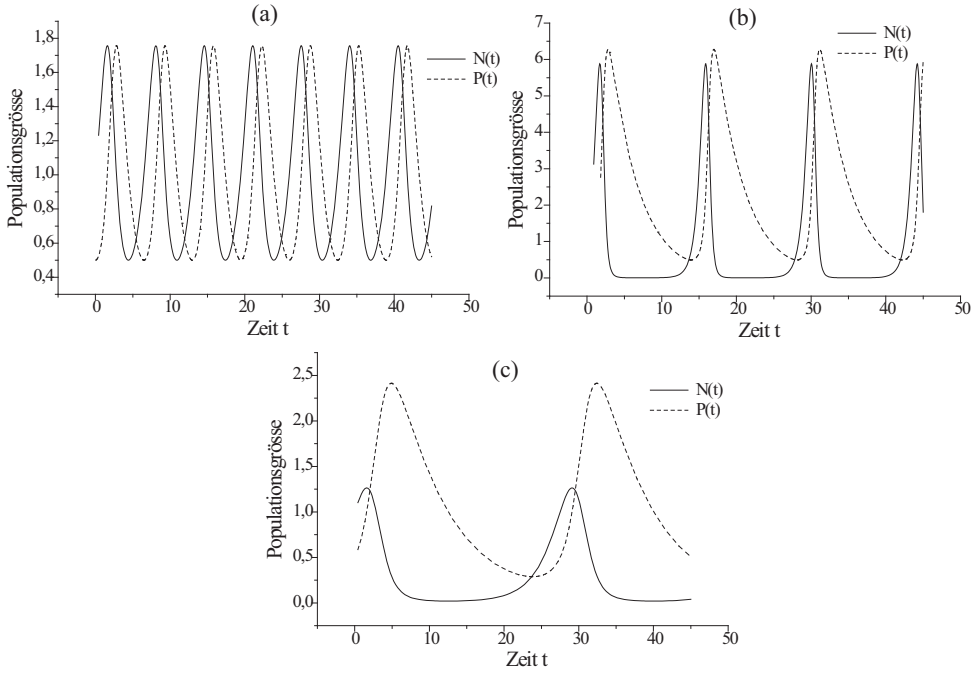
wobei N die Anzahl der Elemente in B und P die Anzahl der Elemente in R ist. Formal besteht eine Ähnlichkeit zur logistischen Gleichung (25) (man ersetze in (27) $P(t)$ durch $N(t)$). Ausgeschrieben kann (27) in der Form

$$N'(t) = aN(t) - abN(t)P(t) \tag{29}$$

geschrieben werden. Für $b = 0$ erhält man $N'(t) = aN(t)$, also exponentielles Wachstum für N. Für $b \neq 0$ wird aber die Veränderung N' durch den Term $abN(t)P(t)$ reduziert. $N(t)P(t)$ ist die Anzahl der möglichen Paare von Raub- und Beutefischen zur Zeit t. Sind N und P groß, so gibt es viele tödliche Treffen zwischen den beiden Arten und das Wachstum wird verringert (es gibt hier auch negatives Wachstum, d.h. N schrumpft). Für NP klein wird N' nur wenig reduziert und man erhält ein größeres Wachstum von N.

Die Gleichung (28) beschreibt die Veränderung der Population P; man erhält $P' = cN(t)P(t) - dP(t)$. (d ist hier nicht das Differential.) Für große Werte von P wird das Wachstum von P reduziert. Für große Werte von N verändert sich P schnell, d.h. die

Abbildung 7: Interaktion zwischen zwei Populationen: das Lotka-Volterra Modell. (a) $a = b = c = d = 1$, (b) $a = 1.7$, $b = .75$, $c = .35$, $c = .255$; (c) $a = .5$, $b = .5$, $c = .5$, $d = .15$



Population wächst: es gibt viel Nahrung. Für kleine Werte von N wächst P nur schwach oder wird, wenn $cN(t)P(t) - dP(t) < 0$, d.h. $cN(t)P(t) < dP(t)$, sogar kleiner. Für $N = 0$ stirbt die Population der Räuber exponentiell aus. Abbildung 7 zeigt die Form der Interaktion für verschiedene Kombinationen von Parameterwerten.

Das System von Gleichungen (27) und (28) ist auch als Lotka-Volterra System bekannt. Es ist das einfachste Modell für interagierende Populationen und liefert in guter Näherung z.B. die Beziehung nicht nur zwischen Fischpopulationen, sondern auch für die Wechselwirkung zwischen Hase und Fuchs, etc. Interessant sind die Verallgemeinerungen z.B. für die Interaktion von Enzymen, Neuronen etc. Eine Übersicht findet man in Murray, (1989).

Modelle wie das Lotka-Volterra System sind zumindest dem Ansatz nach geeignet, die Ausbreitung bestimmter Ansichten quantitativ zu erklären. Ansichten sind hierbei Aussagen, deren Wahrheit oder Falschheit nicht eindeutig bewiesen werden kann. Eine Ansicht wie die "Die Arbeitslosigkeit ist in Deutschland so hoch, weil die Löhne zu hoch sind" hat für eine Teilmenge der Einwohner Deutschlands Evidenzcharakter, für eine andere Teilmenge aber nicht; sie kann jedenfalls nicht aus allgemein anerkannten Axiomen logisch hergeleitet werden. Gleichwohl gibt es Zeiten, in denen Ansichten wie diese von mehr und mehr Personen adoptiert werden. Mit einer gewissen Verzögerung beginnen

Abbildung 8: Zustimmung zur Todesstrafe (I). Quelle: US Department of Justice. Die Kurve wurde durch Anpassen des Klassischen epidemischen Modells bestimmt.

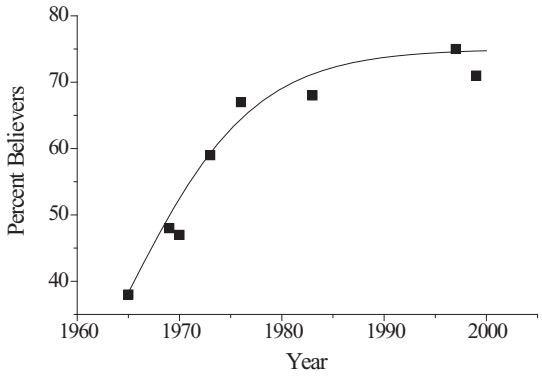
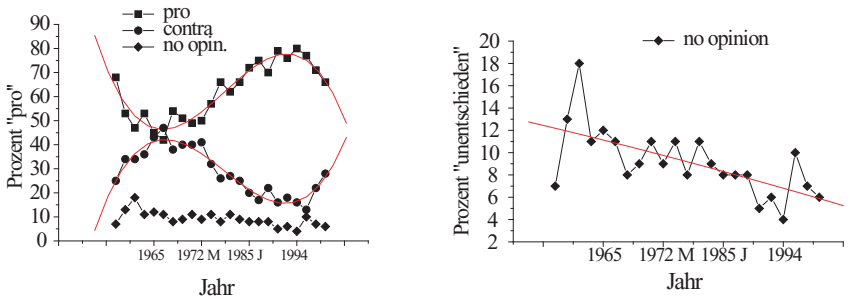


Abbildung 9: Zustimmung zur Todesstrafe (II)



dann diejenigen, die nicht dieser Ansicht sind, ihrerseits ihre Meinung zu verkünden und es kommt zu einer gewissen Inhibierung der Ausbreitung der ersten Ansicht. Ein typisches Beispiel für eine Ansicht, deren Wahrheit nicht bewiesen werden kann (und gegen die die empirische Evidenz spricht) ist die, dass die Todesstrafe eine ebenso sinnvolle wie zu rechtfertigende Strafe sei. In Abb. 8 werden Daten gezeigt, die die Zustimmung sowie die Ablehnung der Aussage zur Todesstrafe zeigen. Die Daten zeigen nur den oberen Zweig des Datenverlaufs. Offenbar ist die Zahl der Anhänger gewachsen, nähert sich aber einer Asymptote, dh einem Sättigungswert; aus der Menge der Indifferenten oder der Gegner der Todesstrafe können keine Personen mehr mit der Ansicht infiziert werden. Abb. 9 zeigt einen ausführlicheren Datensatz, demzufolge die Zahl der Anhänger der Todesstrafe in den USA zunächst schrumpfte, um dann wieder anzusteigen und erneut zu schrumpfen. Die Kurve für die Anzahl der Gegner der Todesstrafe ist spiegelbildlich zu der der Proponenten dieser Strafart. Die Anzahl der Indifferenten wird im Mittel gleichzeitig immer geringer. Über längere Zeiträume wird man die Anzahl der Proponenten wie die der Opponenten besser mit Systemen der Art des Lotka-Volterra Systems beschreiben.

2 Das klassische epidemische Modell

Dieses Modell ist eine Anwendung des von Pierre Verhulst entwickelten Modellszum Wachstum von Populationen zu modellieren. Erste Modelle gingen von einem exponentiellen Wachstum aus, aber Verhulst hatte den Verdacht, dass beim Bevölkerungswachstum Sättungsphänomene auftreten, die das Wachstum begrenzen.

Bei Ansteckungsprozessen kann diese Annahme ebenfalls gemacht werden: je mehr Personen bereits mit einem Grippevirus infiziert wurden, desto weniger können noch angesteckt werden, umso langsamer breitet sich die Krankheit aus, bis es schließlich keine "suszeptiblen" Personen mehr gibt. Verhulst konnte 1838 mit seinem Modell das Wachstum der Bevölkerung von Paris und damit die Anzahl der benötigten Wohnungen (*logis*) voraussagen; deshalb wird in ähnlichen Zusammenhängen vom *logistischen Modell* gesprochen.

Einfache Modelle aus der Epidemiologie (Frauenthal (1980), p. 25) illustrieren die Konstruktion des Modells. Diese Modelle sind für die Psychologie deswegen interessant, weil sie ebenso einfache Modelle zur Ausbreitung von Gerüchten, Vorurteilen, Stereotypen, etc. darstellen. Dazu werden die folgenden Annahmen gemacht:

A1: Gegeben sei eine Population mit konstantem Umfang N, d.h. es gibt keine Geburten, Todesfälle, Zu- oder Abwanderungen; im allgemeinen wird dies die Teilpopulation der überhaupt Infizierbaren einer größeren Population sein.

A2: Eine Person (allgemein: ein Element) aus der Population hat ein Merkmal A (ist mit dem Gerücht oder Vorurteil infiziert), oder $\bar{A}$ (ist nicht infiziert)

A3: A wird durch Kontakt weitergegeben, wobei es keine latente Phase gibt

A4: Alle Personen (Elemente) der Population sind gleich infizierbar, und alle Infizierten sind gleich ansteckend.

2.1 Das "triviale" Modell

Hier wird $N = \infty$ angenommen. Das Modell ist nur deswegen interessant, weil es die grundsätzliche Vorgehensweise bei der Konstruktion solcher Modelle illustriert und den Übergang zu einem realistischeren Modell mit $N < \infty$ erleichtert.

Es sei $X(t)$ die Anzahl der Infizierten zur Zeit t, und es sei β die durchschnittliche Anzahl von Kontakten noch Infizierbarer, die zu einem neuen Infizierten pro Zeiteinheit und pro bereits Infiziertem in der Poplation führen. $X(t+h)$ sei die Anzahl Infizierter zur Zeit $t+h$. Dann gilt

$$X(t+h) = X(t) + \beta h X(t). \tag{30}$$

$X(t+h)$ ergibt sich demnach aus der Anzahl $X(t)$ der zur Zeit t Infizierten, plus einem Anteil $\beta h X(t)$ Neuinfizierter. Dieser Anteil ist proportional zu $X(t)$, denn es werden ja um so mehr infiziert, als es bereits Infizierte gibt. Je größer das Zeitintervall h ist, desto mehr Neuinfizierte wird es geben, so daß h ein Teil des Proportionalitätsfaktors ist. β ist der andere Teil des Proportionalitätsfaktors. Auf einer Party wird der Wert von β groß sein und X wird schnell wachsen, da im Prinzip jeder mit jedem Kontakt macht. Besteht

die Population aus monadischen Individuen, so wird β nahe bei Null sein, und X wird sehr langsam wachsen.

Analog zu dem Vorgehen in Abschnitt 1.2 ergibt sich aus (30) sofort die Gleichung

$$\frac{X(t+h) - X(t)}{h} = \beta X(t), \tag{31}$$

woraus sich für $h \to 0$ die *Differentialgleichung*

$$\frac{dX(t)}{dt} = \beta X(t), \tag{32}$$

ergibt. Sie hat, wie ja schon gezeigt wurde, die Lösung

$$X(t) = x_0 e^{\beta t}, \tag{33}$$

Für $t = 0$ folgt $X(0) = x_0$; x_0 ist die Anzahl der Infizierten zu Beginn des Prozesses. Für $x_0 = 0$ folgt $x(t) = 0$ für alle t, d.h. es muß mindestens eine(n) Infizierte(n) geben, damit überhaupt ein Ansteckungsprozess in Gang kommt. Das Modell ist sicherlich unvollständig, weil es nicht angibt, woher die erste Infektion kommt, da diese den Annahmen entsprechend ja nur über die Ansteckung verbreitet wird. β entspricht dem Parameter λ in Gleichung (7), vergl. die Bemerkung auf Seite 5.

Damit die Anzahl der Infizierten wächst, muß $\beta > 0$ gelten; β ist die *Wachstumsrate* für die (Teil-)Population der Infizierten. Für $t \to \infty$ folgt dann $X(t) \to \infty$, d.h. am Ende sind alle infiziert.

Populationen sind im Allgemeinen nicht unendlich groß, so daß man $N < \infty$ annehmen muß. Die Gleichung (33) kann als erste Näherung benutzt werden, wenn t noch klein ist und N als groß angenommen werden kann. Eine allgemeine Lösung für den Fall $N < \infty$ wird im folgenden Abschnitt hergeleitet.

2.2 Das klassische epidemiologische Modell

Es sei $N < \infty$. Wenn $X(t)$ die Anzahl der Angesteckten zur Zeit t ist, so gibt es $S(t) = N - X(t)$ Personen, die zur Zeit t noch nicht angesteckt sind; es wird S für diese Anzahl geschrieben, um anzuzeigen, daß es sich um die "Suszeptiblen", also um die überhaupt Infizierbaren handelt. Jedenfalls gilt dann

$$N = X(t) + S(t) = \text{konstant}. \tag{34}$$

Man kann den Ansatz (30) nun verallgemeinern. $X(t)$ wird von der Anzahl der möglichen Kontakte zwischen Infizierten und Nichtinfizierten abhängen. Da jede noch infizierbare Person mit jeder bereits infizierten Person in Kontakt treten kann, ist die Anzahl der *möglichen* Kontakte gerade $S(t)X(t)$, so daß man

$$X(t+h) = X(t) + \beta h X(t) S(t) \tag{35}$$

ansetzen kann. Man erhält daraus

$$\frac{X(t+h) - X(t)}{h} = \beta X(t) S(t), \tag{36}$$

und für $h \to 0$ ergibt sich die Differentialgleichung

$$\frac{dX(t)}{dt} = \beta n(t), \quad n(t) = S(t)X(t). \tag{37}$$

Anmerkung (Bedeutung des Parameters β): Es kann hilfreich sein, sich die Bedeutung von β in Gleichung (37) etwas expliziter klar zu machen: das Produkt $n(t) = S(t)X(t)$ ist die Anzahl der *möglichen Kontakte* zwischen Suszeptiblen (Infizierbaren, $S(t)$) und bereits Infizierten ($X(t)$) zum Zeitpunkt t. Da jeder Kontakt einen Suszeptiblen in eine Infizierten verwandelt, ist $S(t)X(t)$ die mögliche Anzahl Infizierter zur Zeit t Da ein Neuinfizierter einem Kontakt zwischen einem Infizierten und einem Suszetiblen bedeutet, entspricht $dX(t)$ der Anzahl der Kontakte während des Zeitintervalls der Dauer dt. Betrachtet man

$$\frac{1}{n(t)dt}$$

als Wahrscheinlichkeit eines Kontaktes in einem Intervall der Dauer dt, so ist

$$\beta = \frac{dX(t)}{n(t)dt} = \frac{\dot{X}(t)}{n(t)} \tag{38}$$

die durchschhnittliche Anzahl von Kontakten[11], d.h. Neuinfizierten während des Zeitintervalls $[t, t + dt)$. $\dot{X}(t)$ ist die newtonsche Schreibweise für die Derivierte $dX(t)/dt$ von $X(t)$. Etwas anschaulicher ist vielleicht der Bezug auf die Gleichung (36), wo β zum Quotienten $(X(t+h) - X(t)/h$ in Beziehung gesetzt wird; β ist die durchschnittliche Anzahl von Kontakten pro Zeitintervall von "hinreichend kleiner" Dauer h.

Je größer das Produkt $S(t)X(t)$ zur Zeit t ist, je größer also die Anzahl der möglichen Kontakte zur Zeit t ist, desto größer wird die *Veränderung* von $X(t)$ zur Zeit t sein. Aus (34) folgt

$$S(t) = N - X(t);$$

setzt man diesen Ausdruck für S in (37) ein, so erhält man

$$\frac{dX(t)}{dt} = \beta X(t)(N - X(t)). \tag{39}$$

Dies ist die *logistische Gleichung*, die die Lösung

$$X(t) = \frac{N}{1 + Ae^{-\alpha t}}, \quad A = \frac{N - X(0)}{X(0)}, \ \alpha = N\beta > 0 \tag{40}$$

hat; die Herleitung von (40) findet man im Anhang, Seite 32, ihre Details sind für das Folgende nicht wesentlich. Damit der Ansteckungsprozess überhaupt in Gang kommt, muß $X(0) \neq 0$ sein.

Für $t \to \infty$ folgt $X(t) \to N$, und für $t \to 0$ folgt $X(t) \to X(0)$, so dass $X(0) \leq X(t) \leq N$. Das (deterministische) Modell impliziert also, daß am Ende alle Elemente

[11] Führt man n Bernoulli-Experimente mit dem Parameter p aus, d.h. hat man bei einem Experiment entweder einen "Erfolg" (mit der Wahrsccheinlichkeit p) oder einen "Misserfolg" (mit der Wahrschheinlichkeit $1 - p$, und sind die Experimente voneinander unabhängig, so ist bekanntlich np die "erwartete", d.h. durchschnittliche Anzahl von Erfolgen.

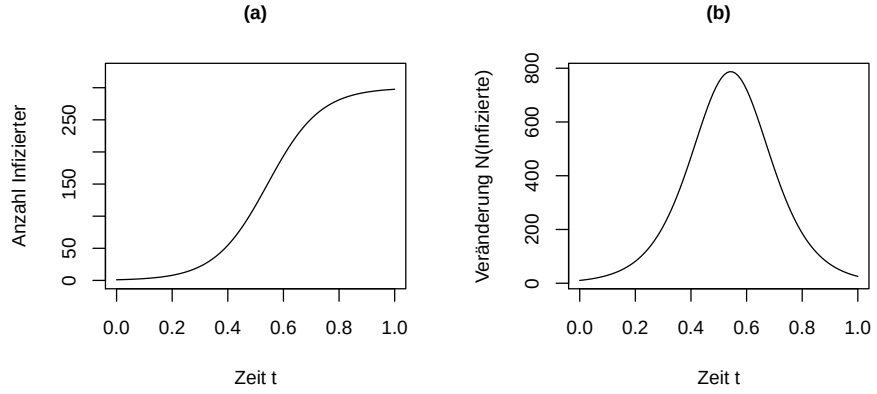
der Population bzw. der Teilpopulation infiziert sind, sofern alle Elemente überhaupt infizierbar sind.

Die Gleichung (40) erlaubt es, $\dot{X}(t)$ in alternativer Form darzustellen. Differenziert man die rechte Seite von (40) nach t, so erhält man wegen $S(t) = N - X(t)$

$$\dot{X}(t) = \frac{dX(t)}{dt} = \frac{\alpha N A e^{-\alpha t}}{(1 + Ae^{-\alpha t})^2} =: W(t) \tag{41}$$

Der Ausdruck auf der rechten Seite (also $W(t) := \dot{X}(t)$) ist auch als *epidemische Kurve* bekannt.

Abbildung 10: Logistisches Wachstum der Anzahl Infizierter, (a) Anzahl der Infizierten nach der Zeit t, Gl (40), (b) Veränderung der Anzahl Infizierter zur Zeit t (epidemisches Wachstum $W(t)$), Gl. (41).



Beispiel 2.1 Es wird eine große Party mit $N = 299$ suszeptiblen Personen und einer bereits mit einer ansteckenden Krankheit infizierten Person gefeiert, d.h. man habe $N = 300$ Teilnehmer. Der Parameter β ist so gewählt worden, dass das Wachstum innerhalb eines Zeitintervalls der Länge 1 (eine Stunde, ein Tag, eine Woche, etc) stattfindet: dazu muß $\beta = .035$ gesetzt werden. Setzt man die 1 an der t-Achse für 1 Stunde, so sind nach einer Stunde alle infiziert; auf einer ziemlich großen Party kann so etwas passieren. Nach Definition von β finden also pro "Augenblick" im Durchschnitt .035 Kontakte statt (vergl. Anmerkung zur Bedeutung on β). Abb. 10 illustriert den Verlauf der Infektionen. Abbildung 2.1 illustiert das Wachstum der Anzahl Infizierter und die Veränderung der Anzahl Infizierter mit der Zeit t[12]. □

In analoger Weise kann man die Veränderung von $S(t)$, der Anzahl der zur Zeit t noch nicht Infizierten, beschreiben. Je größer die Anzahl der möglichen Kontakte zwischen

[12]Mein Dank gilt Dr. G. Frank für seinen Hinweis auf eine unklare Formulierung des Beispiels in der vorangegangenen Fassung dieses Skripts

Infizierten und Nichtinfizierten ist, desto schneller wird die Anzahl der noch nicht Infizierten abnehmen, denn jeder Kontakt bedeutet Infektion. Dementsprechend wird man die (Differential-)Gleichung

$$\frac{dS(t)}{dt} = -\beta S(t)X(t) \tag{42}$$

ansetzen. Das Vorzeichen von $\beta S(t)X(t)$ ist hier negativ, da ja S in Abhängigkeit von der Anzahl der Kontakte kleiner wird. Aus (34) folgt $X(t) = N - S(t)$, so daß man aus (42) die Differentialgleichung (DGL)

$$\frac{dS(t)}{dt} = -\beta S(t)(N - S(t)) \tag{43}$$

erhält. Sie entspricht der DGL (39); das Vorzeichen von β ist hier negativ, weil ja S mit der Zeit kleiner wird. Wegen $S(t) = N - X(t)$ ergibt sich die Lösung für S sofort aus der für X.

Zur Interpretation von β: Zunächst sei festgestellt, daß β von der Einheit abhängt, in der die Zeit t gemessen wird. Denn der Übergang von einer Einheit zur anderen (etwa Sekunden zu Minuten, oder umgekehrt) bedeutet, dass man von t zu $\Delta t = at$, $a > 0$ übergeht. Wird die Kettenregel auf (39) angewendet hat man

$$\frac{dX(\Delta t)}{dt} = a\frac{adX(\Delta t)}{d\Delta t},$$

d.h.

$$a\frac{dX(\Delta t)}{d\Delta t} = \beta X(\Delta t)(N - X(\Delta t)),$$

d.h. aber, wenn man wieder t statt Δt schreibt,

$$\frac{dX(t)}{dt} = \frac{\beta}{a}X(t)(N - X(t)).$$

Der Proportionalitätsfaktor ist jetzt β/a. Man kann nun wieder β für β/a schreiben; man sieht aber, daß die neue Einheit den Wert des Faktors mit bestimmt.

Aus (36) folgt

$$\frac{1}{h}\left(\frac{X(t+h) - X(t)}{X(t)S(t)}\right) = \beta. \tag{44}$$

Demnach ist β gleich dem Anteil der Anzahl Infizierter, um die $X(t)$ nach der Zeit h zunimmt, an der Gesamtzahl möglicher Kontakte $X(t)S(t)$ zur Zeit t, geteilt durch h. Man kann dann nach der Zeit $h_1(t)$ fragen, die nötig ist, damit gerade $X(t+h)-X(t) = 1$ ist. Wegen $S(t) = N - X(t)$ erhält man

$$\frac{1}{h_1(t)}\left(\frac{1}{X(t)(N - X(t))}\right) = \beta. \tag{45}$$

$1/(X(t)(N - X(t))) =: p(1)$ kann als Wahrscheinlichkeit für einen bestimmten Kontakt interpretiert werden, wobei idealisierend vorausgesetzt wurde, dass alle möglichen Kontakte die gleiche Wahrscheinlichkeit haben. (45) kann dann in der Form

$$\beta = \frac{p(1)}{h_1}, \tag{46}$$

geschrieben werden, d.h. β ist das Verhältnis der Wahrscheinlichkeit eines Kontakts zur Dauer, die für das Herstellen des Kontakts benötigt wird. Dem Ansatz zufolge hat dieses Verhältnis für alle t den gleichen Wert. Gibt es viele Möglichkeiten für einen Kontakt, so wird die Wahrscheinlichkeit für einen bestimmten Kontakt klein; dafür wird die Zeitdauer, die bis zum Abschluß (Infektion) eines Kontakt benötigt wird, entsprechend kurz. Gibt es wenige Möglichkeiten, so wächst die Wahrscheinlichkeit $p(1)$ für einen der möglichen Kontakte, aber dafür wird die Wartezeit h_1 entsprechend lang.

3 Die logistische Verteilung

3.1 Die Definition der logistischen Verteilung

Verteilungsfunktionen, d.h. Funktionen vom Typ $F(x) = P(X \leq x)$, haben oft einen S-förmigen Verlauf, so etwa die Gauß-Verteilung (Normalverteilung). Die Wachstumskurve in Abbildung (10) zeigt einen ähnlichen Verlauf, und es stellt sich die Frage, ob man die Funktion (40) (Seite 19), nämlich

$$X(t) = \frac{N}{1 + Ae^{-\alpha t}},$$

die der Abbildung (10) unterliegt, nicht durch geeignete Normierung in eine Verteilungsfunktion verwandeln kann, – für eine Verteilungsfunktion muß ja $0 \leq F(x) \leq 1$ für alle x aus dem Definitionsbereich von F gelten. Der praktische Reiz einer solchen Übung liegt darin, dass die Funktion (40) leicht zu berechnen ist, zumindest im Vergleich zur Gauß-Funktion, die durch ein Integral definiert ist, das nicht in geschlossener Form darstellbar und deshalb numerisch approximiert werden muß. Tatsächlich liefert die Normierung von (40) eine vorzügliche Approximation für eine Gauß-Funktion. xxx

Es zeigt sich aber darüber hinaus, dass gewisse theoretische Überlegungen auf eine Adaptation von (40) führen.

Die logistische Verteilung ist durch

$$P(X \leq x) = F(x) = \frac{1}{1 + \exp\left(-\frac{(x-\mu)\alpha}{\sigma}\right)}, \quad \alpha = \pi/\sqrt{3} \tag{47}$$

definiert. Mit $\beta = \alpha/\sigma$ hat man

$$P(X \leq x) = \frac{1}{1 + \exp\left(-(x-\mu)\beta\right)} = \frac{1}{1 + e^{-x\beta}e^{\mu\beta}},$$

und für $x \to -\infty$ folgt offenbar $F(x) \to 0$, und für $x \to +\infty$ folgt $F(x) \to 1$, so dass F auf $(-\infty, +\infty)$ definiert ist. Der Erwartungswert $\mathbb{E}(X)$ und die Varianz $\mathbb{V}(X)$ sind durch

$$\mathbb{E}(X) = \mu, \quad \mathbb{V}(X) = \sigma^2 \tag{48}$$

gegeben. Die zugehörige Dichtefunktion ergibt sich durch Differentiation:

$$f(x) = \frac{dF(x)}{dx} = \frac{\alpha}{\sigma}\left(\frac{\exp(-(x-\mu)\alpha/\sigma))}{(1 + \exp(-(x-\mu)\alpha/\sigma)))^2}\right) \tag{49}$$

Die logistische Verteilung ähnelt sehr der Gaußschen Verteilung (vergl. Abb.11, Seite 24).

Es ist

$$1 - F(x) = \frac{\exp(-(x-\mu)\alpha/\sigma))}{1+\exp(-(x-\mu)\alpha/\sigma))},$$

so daß (49) in der Form

$$f(x) = \frac{dF(x)}{dx} = \frac{\alpha}{\sigma}F(x)(1-F(x)) = \beta F(x)(1-F(x)) \tag{50}$$

geschrieben werden kann, mit $\beta = \alpha/\sigma$. Dies ist eine Differentialgleichung, die der Gleichung (39) entspricht, allerdings ohne dass x von der Zeit t abhängt.

Es stellt sich die Frage, welche Eigenschaften der zufälligen Veränderlichen X implizieren, dass die Verteilungsfunktion und damit die Dichte von X durch die logistische Verteilung gegeben sind. Es liegt nahe, X in einen Zusammenhang mit Wachstumsprozessen zu bringen, wie sie zum Beispiel im klassischen epidemiologischen Modell betrachtet wurden. Teilt man die Gleichung (39) durch N, so erhält man

$$\frac{1}{N}\frac{dX(t)}{dt} = \beta X(t)\frac{N-X(t)}{N} = N\beta\frac{X(t)}{N}\left(\frac{N-X(t)}{N}\right),$$

und setzt man

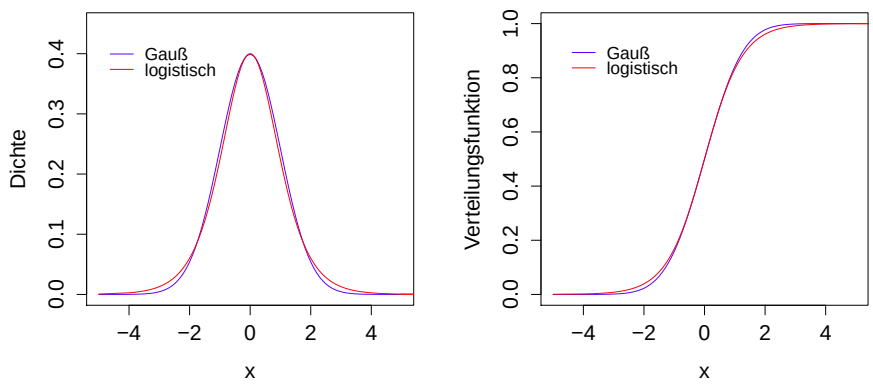
$$F_x(t) = \frac{X(t)}{N}, \quad \beta^* = N\beta, \tag{51}$$

so erhält man

$$\frac{d(X(t)/N)}{dt} = \frac{dF_x(t)}{dt} = \beta^* F_x(t)(1-F_x(t)). \tag{52}$$

$F_x(t)$ ist der Anteil der Infizierten an der Gesamtpopulation der überhaupt Infizierbaren, zur Zeit t; man kann $F_x(t)$ als die Wahrscheinlichkeit auffassen, zur Zeit t auf einen Infizierten zu treffen, wenn man zufällig eine Person aus der entsprechenden Person wählt, und wenn der Prozess seit t Zeiteinheiten läuft. Insofern entsprechen sich die Bedeutungen der Gleichungen (50) und (52). Andererseits muß man sehen, dass die Anzahl $X(t)$ von Infizierten implizit als durchschnittliche Anzahl von zur Zeit t Infizierten angesehen werden muß, denn grundsätzlich ist der Ansteckungsprozess ja stochastisch definiert: nicht Infizierte und Infizierte treffen zufällig aufeinander, und läßt man den Prozeß unter identischen Bedingungen mehrfach ablaufen, so wird die Anzahl $\xi(t)$ der zur Zeit t tatsächlich Infizierten bei jedem Durchlauf anders ausfallen, so dass $X(t)$ als deterministisch definierte Größe nur die mittlere oder erwartete Anzahl sein kann. Somit läge es nahe, $X(t)/N$ mit dem Erwartungswert μ in (50) in Verbindung zu bringen, wobei man jetzt μ als Funktion der Zeit auffasst, $\mu = \mu(t)$. Dann ergibt sich aber die Frage, was die zufällige Veränderliche X repräsentiert, schließlich muß für die logistische Verteilung ja $-\infty < X < \infty$ gelten, und darüber hinaus wird impliziert, dass die logistische Verteilung durch eine Funktion $\mu(t)$ definiert ist, die ihrerseits einer logistischen Wachstumskurve entspricht. Wie es scheint, ist eine direkte Beziehung zwischen der logistischen Verteilung und der logistischen Funktion nicht leicht herzustellen. Es genügt fürs Erste, die logistische Verteilung nur analog zur logistischen Funktion zu sehen: $F(x) = P(X \leq x)$ wächst für kleine Werte von F zunächst angenähert exponentiell mit x, dann approximativ linear und geht chließlich asymptotisch wie $1 - \exp(-\alpha t)$ gegen 1. Meistens wird die logistische Verteilung einfach als eine relativ gute Approximation an die Normalverteilung gewählt,

Abbildung 11: (a) Gauß-Dichte ($\mu = 0, \sigma^2 = 1$) und logistische Dichte ($\mu = 0, \sigma^2 = .625$), (b) korrespondierende Verteilungsfunktionen.



die den Vorteil hat, eine geschlossene Formel für $F(x)$ zu haben. In anderen Anwendungen, etwa in der Theorie psychometrischer Tests, wird die logistische Verteilung gewählt, weil sie bestimmte wünschenswerte Eigenschaften hat. Darauf wird in Abschnitt 3.3 näher eingegangen. In den Abschnitten 3.3 und 4 werden weitere Beziehungen zwischen logistischer Funktion und Wahrscheinlichkeiten aufgezeigt.

3.2 Logistische versus Gauß-Verteilung

Die Verteilungsfunktionen für die Gauß-Verteilung einerseits und die der logistische Verteilung andererseits sind außerordentlich ähnlich (bei geeigneter Wahl der Parameterwerte), während bei den Dichtefunktionen gewisse Unterschiede deutlich werden: die logistische Dichte ist etwas schmaller, nimmt dafür im Zentrum höhere Werte an; weitere Unterschiede finden sich an den Enden der Dichtefunktionen (vergl. Abbildung 11). Die logistische Dichte geht wie eine Exponentialfunktion $\exp(-\alpha x)$ und damit langsamer als die Normalverteilung gegen Null, die wie $exp(-x^2/2\sigma^2)$ gegen Null geht.

Die logistische Verteilung (47) ist im Vergleich zur Gauß-Verteilung leicht zu berechnen, da für die Gauß-Verteilung kein geschlossener Ausdruck für das Integral existiert. Dies legt nahe, Die logistische Verteilung als Approximation für die Gauß-Verteilung zu benützen. Diese Möglichkeit erweist sich als besonders nützlich, wenn es darum geht, den Effekt unabhängiger Variablen auf die Häufigkeit des Eintretens bestimmter zufälliger Ereignisse zu bestimmen. So sei η eine zufällige Veränderliche und A ein zufälliges Ereignis, das dann eintritt, wenn $\eta > \eta_0$ ist, wobei man η_0 als einen Schwellenwert interpretieren kann. η_0 sei $N(\mu, \sigma^2)$-verteilt. Ein Beispiel ist die zufällig in der Zeit fluktuierende Fähigkeit, Aufgaben eines bestimmten Typs (etwa: $17 \times 13 = ?$) innerhalb eines bestimmten Zeitintervalls im Kopf zu lösen. A ist das Ereignis, dass eine Aufgabe gelöst wird, und

die Aufgabe wird gelöst, wenn $\eta > \eta_0$ ist. Es werde angenommen, dass μ von bestimmten Variablen X_j abhängt: es könnte der Sauerstoffgehalt der Luft zum Zeitpunkt des Testens, die Anzahl der Stunden Schlaf in der vorausgegangenen Nacht, die Anzahl der vor dem Test genossenen Tassen Kaffees etc. bedeuten. In jedem Fall kann man von dem Ansatz

$$\mu = \mu(X_1, \ldots, X_n) \tag{53}$$

ausgehen. Setzt man $\eta = \mu + \xi$, so hat man

$$P(A|\mu) = P(\eta > \eta_0) = P(\mu + \xi > \eta_0) = P(\xi > \eta_0 - \mu).$$

Bei geeigneter Normierung hat man $\xi \sim N(0, 1)$, so dass man insbesondere

$$P(A|\mu(X)) = 1 - \Phi(\eta_0 - \mu) = \Phi(\mu - \eta_0) \tag{54}$$

hat. Die nähere Bestimmung der Beziehung zwischen μ und den $X_1, \ldots, X_n$ ist die Probit-Analyse. Die Evaluierung des Integrals Φ kann man sich ersparen, wenn man statt der Gauß- die logistische Verteilung betrachtet, aus der sich leicht die Beziehung

$$\log \frac{p}{1-p} = \mu(X_1, \ldots, X_n) \tag{55}$$

herleiten läßt, wobei $p = P(A|\mu)$ ist. Die freien Parameter, die für μ in (54) einerseits und (55) andererseits geschätzt werden müssen, können aufeinander bezogen werden.

Eine ganz andere Beziehung zwischen logistischer Funktion und Gauß-Verteilung wird in Abschnitt 4 angesprochen.

3.3 Spezielle Eigenschaften der logistischen Verteilung

Die logistische Verteilung ist u. a. durch den Erwartungswert μ definiert. Setzt man $\beta = \alpha/(\sigma\sqrt{3})$, so läßt sich die Verteilung in der Form

$$P(X \leq x|\mu) = F(x; \mu) = \frac{1}{1 + \exp(-(x - \mu)\beta)} \tag{56}$$

schreiben. Man kann nun $P(X \leq x|\mu)$ als Funktion von μ betrachten. Man findet

$$\frac{dF(x;\mu)}{d\mu} = -\frac{\beta \exp(-(x-\mu)\beta)}{(1 + \exp(-(x-\mu)\beta))^2}, \tag{57}$$

d.h. für fixen Wert von x und wachsendem Wert von μ wird $F(x; \mu)$ kleiner, und für kleiner werdenden Wert von μ wird $F(x; \mu)$ größer. Es sei nun μ eine Funktion eines Parameters θ, etwa

$$\mu(\theta) = \theta. \tag{58}$$

Für steigenden Wert von θ wird nun $F(x; \mu)$ kleiner, und

$$P(X > x|\mu) = P(X > x|\theta) = 1 - F(x; \theta) \tag{59}$$

wird größer. Es sei $G(\xi) = 1/(1 + e^{-\xi})$. Dann ist

$$1 - G(\xi) = 1 - \frac{1}{1 + e^{-\xi}} = \frac{1 + e^{-\xi} - 1}{1 + e^{-\xi}} = \frac{e^{-\xi}}{1 + e^{-\xi}} = \frac{1}{1 + e^{\xi}},$$

so dass man mit $\xi = (x - \mu)\beta$

$$P(X > x|\theta) = \frac{1}{1 + \exp(-(\theta - x)\beta)} \tag{60}$$

hat. Als Funktion von θ erhält man, für fixen Wert von x, den Ausdruck

$$P(X > x|\theta) = \pi(\theta) = \frac{1}{1 + \exp(a_0 + a_1\theta)}, \tag{61}$$

mit $a_0 = x\beta$, $a_1 = -\beta$. (61) ist eine reparametrisierte Form von (60) und repräsentiert einen Spezialfall der logistischen Regression, mit θ als unabhängiger Variablen. Wegen

$$\frac{d\pi(\theta)}{d\theta} = \frac{a_1 e^{a_0 + a_1\theta}}{(1 + \exp(a_0 + a_1\theta))^2} > 0 \text{ für } a_1 > 0$$

entspricht (61) einer logistischen Wachstumsfunktion (von θ).

Es zeigt sich, dass die logistische Verteilung Eigenschaften hat, die der Gauß-Verteilung *nicht* zukommen. Diese Eigenschaften zeigen sich u. A. in der Theorie psychometrischer Messungen. Dabei werden Probanden Items I_g, $g = 1, \ldots, n$ vorgelegt; der Einfachheit halber werden hier nur dichotome Items betrachtet, d.h. Aufgaben, die entweder gelöst oder nicht gelöst werden, oder Aussagen, denen entweder zugestimmt wird oder nicht. Es sei X_{ga} eine Indikatorvariable, d.h. $X_{ga} = 1$ wenn die Person a dem Item I_g zugestimmte hat bzw. die Aufgabe gelöst hat, und $X_{ga} = 0$ sonst. Es wird angenommen, dass die Wahrscheinlichkeit für $X_{ga} = 1$ vom Wert eines für die Person charakteristischen Parameters θ_a und von der Schwierigkeit κ_g des g-ten Items abhängt. Die Wahrscheinlichkeit wird durch eine *Item-Response*-Funktion oder kurz *Itemfunktion* definiert:

$$P(X_g = 1|\theta) = F_g(\theta); \tag{62}$$

hier wurde der Index a unterdrückt, weil F_g für eine beliebige Person gelten soll und F_g eben als Funktion von θ erklärt wird. Man kann $A = \{X_g = 1\}$ setzen und dann insbesondere den Ansatz (54) machen

$$P(X_g = 1|\theta) = F_g(\theta) = \Phi(\theta - \kappa_g) \tag{63}$$

setzen, d.h. man hat damit θ mit dem Erwartungswert μ einer 'latenten' Variablen in Beziehung gebracht, und denSchwellenwert η_0 mit der Schwierigkeit κ_g des Items gleichgesetzt. Alternativ, aber analog dazu, kann man von der Logit-Funktion (55) ausgehen und

$$\log \frac{P(X_g = 1|\theta)}{1 - P(X_g = 1|\theta)} = \theta - \kappa_g \tag{64}$$

ausgehen. Löst man die Gleichung nach $P(X_g = 1|\theta)$ auf, so erhält man

$$P(X_g = 1|\theta) = \frac{e^{\theta - \kappa_g}}{1 + e^{\theta - \kappa_g}}; \tag{65}$$

als Funktion von θ spricht man von der logistischen Itemfunktion (sie ist ein Spezialfall von (61)). Dieses Modell wurde zuerst von dem dänischen Statistiker Georg Rasch[13] in

[13] 1901 – 1980

Bezug auf die im Folgenden (s. u.) betrachtete Eigenschaft der *spezifischen Objektivität* diskutiert; man spricht dementsprechend auch vom Rasch-Modell. Natürlich kann auch hier der Ansatz der 'latenten' Variablen η gemacht werden, indem man postuliert, dass ein Item beantwortet wird, wenn $\eta > \eta_0 = \kappa_g$ ist, und dass η logistisch verteilt ist mit dem Erwartungswert $\mu = \theta$. Diese Interpretation entspricht dem Modell des *stochastischen Subjekts*, wie es von Holland (1990) diskutiert wurde. Der alternative Ansatz einer Interpretation verzichtet auf die Annahme eines stochastischen Antwortprozesses innerhalb der Probanden und nimmt statt dessen an, dass für jede Person der Antwortprozess deterministisch verläuft und die Wahrscheinlichkeit $P(X_g = 1|\theta)$ nur deswegen zwischen 0 und 1 (also *nicht* gleich 1) ist, weil θ insofern zufällig ist, als die Wahl einer Person mit einem bestimmten θ-Wert zufällig ist. Viele Psychometriker ziehen diese Interpretation der des stochastischen Subjekts vor, während Psychologen, die sich eher mit Prozessmodellen psychischer Vorgänge beschäftigen, eine größere Neigung haben, der Interpretation des stochastischen Subjekts zuzuneigen. Für das Folgende ist die Diskussion der beiden Standpunkte aber nicht wichtig.

Eine wichtige Eigenschaft des logistischen Ansatzes (65) ist die spezifische Objektivität, das ist die Unabhängigkeit des Vergleichs verschiedener Items von den Personparametern, und die Unabhängigkeit des Vergleichs verschiedener Personen von den Itemparametern κ_g. Zur Vereinfachung werde $p_g = P(X_g = 1|\theta)$ geschrieben. (64) erscheint dann in der Form

$$\log \frac{p_g}{1 - p_g} = \theta - \kappa_g.$$

Bildet man nun die Differenz zwischen den Logits für zwei verschiedene Items I_g und $I_{g'}$, so erhält man

$$\log \frac{p_g(\theta)}{1 - p_g(\theta)} - \log \frac{p_{g'(\theta)}}{1 - p_{g'(\theta)}} = \theta - \kappa_g - \theta + \kappa_{g'} = \kappa_{g'} - \kappa_g, \tag{66}$$

der Parameter θ fällt also heraus. In gleicher Weise fällt κ_g heraus, wenn man zwei verschiedene Personen mit den Parametern θ und θ' betrachtet:

$$\log \frac{p_g(\theta)}{1 - p_g(\theta)} - \log \frac{p_g(\theta')}{1 - p_g(\theta')} = \theta - \kappa_g - \theta' + \kappa_g = \theta - \theta'. \tag{67}$$

Gelten die Gleichungen (66) und (67), so erfüllt die Itemfunktion die Bedingungen der spezifischen Objektivität. Einige Psychometriker postulieren, dass spezifische Objektivität eine grundlegende Eigenschaft zumindest fundamentaler wissenschaftlicher Messungen ist. Akzeptiert man dieses Postulat, so ist man mit der Wahl der Itemfunktion festgelegt: Rasch (1967), Andersen (1973), Irtl (1987) haben gezeigt, dass die Eigenschaft der spezifischen Objektivität impliziert, dass die Itemfunktion durch die logistische Funktion (65) definiert ist.

4 Regression und Klassifikation

Oft müssen Entscheidungen unter Unsicherheit getroffen werden, d.h. die Daten $\mathbf{X}$, die als Basis für die Entscheidung dienen, lassen keine eindeutige Entscheidung zu. Man denke etwa an Entscheidungen über das zukünftige Verhalten von Menschen: wird ein

Sexualstraftäter nach erfolgter Therapie rückfällig werden oder nicht, wird ein Bewerber um einen Studienplatz in der Medizin das Studium erfolgreich abschließen oder nicht, etc. Auch diagnostische Entscheidungen sind oft mit Unsicherheiten verbunden: ob ein Röntgenfoto einen Tumor oder eine harmlose Zyste zeigt, ist anhand des Fotos nicht immer eindeutig zu entscheiden.

Im einfachen Fall muß man sich in all diesen Situationen zwischen zwei Alternativen $\mathcal{C}_1$ und $\mathcal{C}_2$ entscheiden. Die Daten mögen als Vektor $\mathbf{x} = (x_1, \ldots, x_p)'$ vorliegen, d.h. es werden p Variablen gemessen (die Messungen dürfen auf kategorial sein). Nach dem Satz von Bayes gilt nun

$$P(\mathcal{C}_k|\mathbf{x}) = \frac{P(\mathbf{x}|\mathcal{C}_k)P(\mathcal{C}_k)}{P(\mathbf{x})}, \quad k = 1, 2 \tag{68}$$

und mit

$$P(\mathbf{x}) = P(\mathbf{x}|\mathcal{C}_1)P(\mathcal{C}_1) + P(\mathbf{x}|\mathcal{C}_2)P(\mathcal{C}2) \tag{69}$$

folgt

$$P(\mathcal{C}_k|\mathbf{x}) = \frac{P(\mathbf{x}|\mathcal{C}_k)P(\mathcal{C}_k)}{P(\mathbf{x}|\mathcal{C}_1)P(\mathcal{C}_1) + P(\mathbf{x}|\mathcal{C}_2)P(\mathcal{C}2)} \tag{70}$$

Es sei $k = 1$. Dividiert man auf der rechten Seite Zähler und Nenner durch $P(\mathbf{x}|\mathcal{C}_k)P(\mathcal{C}_k)$, so wird aus (70)

$$P(\mathcal{C}_1|\mathbf{x}) = \frac{1}{1 + \frac{P(\mathbf{x}|\mathcal{C}_2)P(\mathcal{C}2)}{P(\mathbf{x}|\mathcal{C}_1)P(\mathcal{C}_1)}} \tag{71}$$

Es sei angemerkt, dass

$$\lambda_{21}(\mathbf{x}) = \frac{P(\mathbf{x}|\mathcal{C}_2)}{P(\mathbf{x}|\mathcal{C}_1)} \tag{72}$$

ein Likelihood-Quotient ist. Es sei nun

$$a(\mathbf{x}) = -\log\left[\frac{P(\mathbf{x}|\mathcal{C}_2)P(\mathcal{C}2)}{P(\mathbf{x}|\mathcal{C}_1)P(\mathcal{C}_1)}\right], \tag{73}$$

d.h.

$$e^{-a(\mathbf{x})} = \frac{P(\mathbf{x}|\mathcal{C}_2)P(\mathcal{C}2)}{P(\mathbf{x}|\mathcal{C}_1)P(\mathcal{C}_1)}, \tag{74}$$

und (71) kann in der Form

$$P(\mathcal{C}_1|\mathbf{x}) = \frac{1}{1 + e^{-a(\mathbf{x})}} \tag{75}$$

geschrieben werden; für $k = 2$ kann man natürlich völlig analog verfahren.

Nach (75) kann $P(\mathcal{C}_1|\mathbf{x})$ durch eine logistische Funktion ausgedrückt werden. Eine Annahme über die Verteilung von $\mathbf{x}$ wurde dabei nicht gemacht, weshalb die Funktion $a(\mathbf{x})$ nicht weiter spezifiziert ist. Insbesondere kann man nun annehmen, dass $\mathbf{x}$ multivariat normalverteilt ist, so dass

$$f_k(x) = \frac{1}{(2\pi)^{p_2}|\Sigma_k|^{1/2}} \exp\left[-\frac{1}{2}(x - \mu_k)\Sigma_k^{-1}(x - \mu_k)\right], \tag{76}$$

wobei die Kovarianzmatrizen Σ_k und die Vektoren der Erwartungswerte μ_k klassenspezifisch sein können. Setzt man zur Abkürzung

$$p_{12} = \frac{P(\mathcal{C}_1)}{P(\mathcal{C}_2)}, \tag{77}$$

so hat man

$$\begin{aligned} a(\mathbf{x}) &= \log\left[\frac{P(\mathbf{x}|\mathcal{C}_1)}{P(\mathbf{x}|\mathcal{C}_2)}\right] + \log p_{12} \\ &= \log\left[\frac{f_1(\mathbf{x})}{f_2(\mathbf{x})}\right] + \log p_{12} \end{aligned}$$

Hierin ist

$$\log\left[\frac{f_1(\mathbf{x})}{f_2(\mathbf{x})}\right] = \frac{1}{2}((\mathbf{x}-\mu_2)\Sigma_2^{-1}(\mathbf{x}-\mu_2) - (\mathbf{x}-\mu_1)\Sigma_1^{-1}(\mathbf{x}-\mu_1)) + \frac{1}{2}\log\left[\frac{|\Sigma_2|}{|\Sigma_1|}\right] \tag{78}$$

Multipliziert man die $(\mathbf{x}-\mu_k)\Sigma_k^{-1}(\mathbf{x}-\mu_k)$ aus, so sieht man, dass die Komponenten von $\mathbf{x}$ quadratisch in die logistische Funktion eingehen. Für $\Sigma_1 = \Sigma_2$ vereinfacht sich der Ausdruck und nimmt eine Form an, die mit üblichen Darstellung der logistischen Funktion übereinstimmt. Denn dann ist

$$(\mathbf{x}-\mu_k)'\Sigma^{-1}(\mathbf{x}-\mu_k) = \mathbf{x}'\Sigma^{-1} - 2\mathbf{x}'\Sigma^{-1}\mu_k + \mu_k'\Sigma^1\mu_k,$$

und

$$\log\left[\frac{f_1(\mathbf{x})}{f_2(\mathbf{x})}\right] = \mathbf{x}'\Sigma^{-1}(\mu_1 - \mu 2) + \mu_2'\Sigma^{-1}\mu_2 - \mu_1'\Sigma^{-1}\mu_1.$$

Setzt man nun zur Abkürzung

$$\mathbf{w} = \Sigma^{-1}(\mu_1 - \mu 2), \quad w_0 = \mu_2'\Sigma^{-1}\mu_2 - \mu_1'\Sigma^{-1}\mu_1 + \log p_{12}, \tag{79}$$

so erhält man

$$P(\mathcal{C}_1|\mathbf{x}) = \frac{1}{1+\exp(-(\mathbf{x}'w + w_0))} \tag{80}$$

d.h. die posteriori-Wahrscheinlichkeit für $\mathcal{C}_1$ ist durch die logistische Funktion gegeben. Da $P(\mathcal{C}_2|\mathbf{x}) = 1 - P(\mathcal{C}_1|\mathbf{x})$ hat man

$$P(\mathcal{C}_2|\mathbf{x}) = 1 - \frac{1}{1+\exp(-(\mathbf{x}'w + w_0))} = \frac{1}{1+\exp(\mathbf{x}'w + w_0)}. \tag{81}$$

Das Bradley-Terry-Luce Modell: Es seien a und b zwei Stimuli und $v(a)$ und $v(b)$ seien Skalenwerte, die die "Antwortstärken" (response strength) für a und b repräsentieren – was auch immer diese Antwortstärken sein mögen. In einem Wahrnehmungsexperiment kann v die Helligkeit zweier Stimuli, oder die Lautheit zweier Töne repräsentieren, in einer Marktforschungsuntersuchung die Luxuriosität zweier Armbanduhren, in einer Untersuchung zum Wahlverhalten die Attraktivität verschiedener Politiker. p_{ab} sei die Wahrscheinlichkeit, dass a vor b "präferiert" wird; im Wahrnehmungsexperiment soll dies

heißen, dass a als heller oder lauter als b wahrgenommen wird, etc. Nach dem Bradley-Terry-Luce Modell (Suppes & Zinnes, 1963, p. 48) soll dann

$$p_{ab} = \frac{v(a)}{v(a) + v(b)} \tag{82}$$

gelten. Nun ist sicherlich

$$\frac{v(a)}{v(a) + v(b)} = \frac{1}{1 + \frac{v(b)}{v(a)}} = \frac{1}{1 + e^{-q(a,b)}}, \tag{83}$$

mit

$$q(a, b) = -\log\left[\frac{v(b)}{v(a)}\right] = \log v(b) - \log v(a). \tag{84}$$

(83) beschreibt offenbar eine logistische Funktion. Man kann nun versuchen, $q(a, b)$ und damit p_{ab} direkt über messbare Stimuluseigenschaften auszudrücken, etwa indem man Vektoren $\mathbf{x}_a = (x_{a1}, x_{a2}, \ldots, x_{an})'$ und analog $\mathbf{x}_b$ betrachtet; die Komponenten x_{aj} und x_{bj} dieser Vektoren repräsentieren Ausprägungen bestimmter Aspekte von a und b. Im Bereich der Psychophysik könnte insbesondere $n = 1$ und

$$v(x) = \alpha(x + \gamma)^\beta$$

gelten, x die Stimulusstärke (etwa: Intensität). Die Elaboration solcher Ansätze geht aber über das Ziel dieser Anmerkungen hinaus.

5 Diskussion

Es sind zwei Herleitungen der logistischen Funktion betrachtet worden. Die erste beruht auf vereinfachenden Annahmen über Wachstumsprozesse. Auch die Ausbreitung einer Epidemie oder von Gerüchten kann. als Wachstumsprozess betrachtet werden. Die Annahmen sind (i), dass der Wachstumsprozess durch endliche Ressourcen begrenzt ist, und dass (ii) die Veränderung der betrachteten Größe X einerseits proportional zu X und andererseits proportional zum noch verbleibenden Rest $N - X$ ist, wobei $N < \infty$ die ein Maß der verfügbaren Ressource ist. Diese Annahmen werden in der Differentialgleichung

$$\frac{dX(t)}{dt} = \beta X(t)(N - X(t))$$

ausgedrückt, deren Lösung eben die logistische Funktion ist. Durch geeignete Normierung läßt sich daraus die Verteilungsfunktion einer logistisch verteilten zufälligen Veränderlichen herleiten. Der Ansatz läßt sich auf das Wachstum von interagierenden Populationen verallgemeinern.

Die zweite Herleitung ergibt sich aus Forderungen an Messungen, wie sie etwa beim Testen von Probanden gemacht werden. Spezifische Objektivität wird von einigen Psychometrikern als eine für wissenschaftliche Messungen notwendige Eigenschaft von Messungen betrachtet, womit sie sich Überlegungen von Rasch (1967) anschließen. Rasch selbst illustrierte dieses Prinzip anhand des zweiten newtonschen Axioms, demzufolge

die Kraft F als Produkt der Masse m eines Körpers und seiner Beschleunigung b erklärt ist, $F = m \times b$. Hat man nämlich zwei Körper mit den Massen m_1 und m_2, auf die die gleiche Kraft F einwirkt, so muß also $F = m_1 b_1 = m_2 b_2$ gelten, woraus wiederum

$$\frac{m_1}{m_2} = \frac{b_2}{b_1} \tag{85}$$

folgt. Äquivalent dazu ist die Gleichung

$$\log m_1 - \log m_2 = \log b_2 - \log b_1, \tag{86}$$

die einer Reparametrisierung der Gleichung (85) entspricht. Diese Gleichungen erlauben es, zwei Massen miteinander zu vergleichen, ohne dass die Kraft explizit bestimmt werden muß, – man muß nur die Beschleunigungen messen. Die Beschleunigungen wiederum können bestimmt werden, ohne dass die Massen oder Kräfte bekannt sein müßten.

Dieses Prinzip der Unabhängigkeit von Messungen einer Größe von anderen Größen soll, so Rasch, auch auf psychometrische Messungen übertragen werden. Die Vorteil liegt zunächst auf der Hand: die Schwierigkeiten der Items – die Itemparameter κ_g – erscheinen als unabhängig von den Fähigkeitsparametern θ der Probanden. Das Interessante an diesem Postulat ist nun, dass ein eher wissenschaftstheoretisch motiviertes Postulat eine bestimmte – eben die logistische – Funktion F_g festlegt. Die Beweise sind länglich und werden, bis auf Irtls Beweis, in Fischer (1995) präsentiert, neben anderen Beweisen, die ebenfalls auf die logistische Funktion führen. Die logistische Funktion erhält dadurch einen ganz anderen Status als nur den der Approximation an die "eigentlich" zu betrachtende Gauß-Funktion, nämlich den Status der "wahren" Itemfunktion, – sofern man das Prinzip der spezifischen Objektivität und anderer, von Fischer (1995) diskutierten Eigenschaften (z.B. die Suffizienz der Schätzungen für die Parameter θ und κ_g) als "wahr" akzeptiert.

Ob man aber etwa das Prinzip der spezifischen Objektivität als notwendig für Modelle, die Verhalten beschreiben, akzeptieren muß, ist fraglich. Denn letzlich verbirgt sich hinter der Forderung nach spezifisch objektiver Messung ein ontologisches Postulat: zumindest die grundlegenden Gesetzmäßigkeiten, die erfaßt werden sollen, müssen so strukturiert sein, dass sie spezifisch objektive Messungen überhaupt ermöglichen. Die newtonsche Formel $F = mb$ hat diese Eigenschaft – zunächst. Aber es ist, wie wohl alle einfachen Gesetze, ein *ceteri paribus*-Gesetz: es gilt z.B. nur dann, wenn die Geschwindigkeiten v der Körper nicht zu groß sind, wie man seit Einstein (1905) weiß. b, die Beschleunigung, repräsentiert die *Veränderung* der Geschwindigkeit, und die kann für verschiedene Geschwindigkeiten gleich groß sein. In (85) geht die Geschwindigkeit der Körper nicht ein. Damit Aussagen über Massen, die man aus (85) ableitet, sinnvoll sind, müssen Randbedingungen bezüglich der Geschwindigkeit der Körper festgelegt werden. Bei den "normalen" Geschwindigkeiten, mit denen man es zu tun hat, ist der Effekt der Geschwindigkeit sicherlich vernachlässigbar, bei sehr hohen nicht. So suggestiv der Hinweis auf elementare physikalische Gesetze sein mag, so wenig kann man postulieren, dass die Struktur dieser Gesetze denen auch elementarer psychologischer Gesetzmäßigkeiten entspricht. Die *ceteris paribus*-Bedingungen können wesentlich einschränkender sein als in der Physik. Ob man also spezifische Objektivität als generelles Prinzip akzeptieren kann, hängt davon ab, ob man sicherstellen kann, dass die zu erklärenden Zusammenhänge dieses Prinzip überhaupt zulassen. Für die dafür notwendigen Messungen kann spezifische Objektivität noch nicht vorausgesetzt werden.

Ein konkretes Beispiel ist der Ansatz Mickos (1969), Reaktionszeiten spezifisch objektiv zu modellieren. Dazu sollte das Rasch-Modell (65) implementiert werden. Dieses Modell kann reparametrisiert werden, denn

$$e^{\theta-\kappa_g} = e^{\theta}e^{\kappa_g} = \delta\sigma_g,$$

mit eben $\delta = \delta(\theta) = e^{\theta}$ und $\sigma_g = \sigma(\kappa_g) = e^{\kappa_g}$. (65) nimmt dann die Form

$$P(X_g = 1|\theta) = \frac{\delta\sigma_g}{1+\delta\sigma_g}$$

an. Micko postuliert, dass die Verteilung der Reaktionszeit τ bei einer Person mit dem Fähigkeitsparameter θ dementsprechend in der Form

$$P(\tau \leq t|\theta) = \frac{\delta(\theta)\sigma(t)}{1+\delta(\theta)\sigma(t)} \tag{87}$$

angeschrieben werden kann, wobei $\sigma(t)$ eine aufgabenspezifische Funktion der Zeit ist, die für alle Personen gleich ist. $\delta(\theta)$ wiederum charakterisiert die Person, unabhängig von der zeitlichen Struktur der Aufgabe. Weitere Einschränkungen über die Form von $\sigma(t)$ werden nicht gemacht, und insofern erscheint der Ansatz (87) sehr allgemein. Vorberg & Schwarz (1990) haben gezeigt, dass schon eine Reihe von einfachen Modellen für Reaktionszeiten der Forderung nach spezifischer Objektivität, also der Formel (87), nicht genügt, und dass nur unrealistische Modelle für Reaktionszeiten mit diesem Ansatz kompatibel sind (es handelt sich um Modelle der seriellen oder parallelen Informationsverarbeitung). Als Vertreter der Philosophie spezifisch objektiver Messung könnte man nun sagen, dass mit den genannten Modelle ja nur falsche Modelle ausgeschlossen würden, das Prinzip spezifisch objektiver Messung also nach wie vor gelte. Leider können aber die Modelle, die mit diesem Prinzip kompatibel sind, wegen mangelnder Kompatibilität mit empirischen Daten nicht akzeptiert werden. Spezifische Objektivität ist offenbar nicht in dem Ausmaß ontologisch zu rechtfertigen, wie sie im ersten Augenblick plausibel zu sein scheint. Die logistische Funktion leistet gute Dienste als Approximation, aber sie ist nicht notwendig die "wahre" Funktion. Mehr kann man kaum verlangen.

6 Anhang

Herleitung von (40): Die Gleichung läßt sich mit der Methode der Trennung der Variablen lösen. Dazu schreibt man Gleichung (39) zunächst in der Form

$$\frac{dX(t)}{X(t)(N-X(t))} = \beta dt. \tag{88}$$

Weiter ist

$$\frac{1}{X(N-X)} = \frac{1}{N}\left(\frac{1}{X} + \frac{1}{N-X}\right),$$

wie man leicht nachrechnet. Dann kann (88) in der Form

$$dX(t)\left(\frac{1}{X(t)} + \frac{1}{N-X(t)}\right) = N\beta dt$$

geschrieben werden, woraus

$$\int_{X(0)}^{X(t)} \frac{1}{U} dU + \int_{X(0)}^{X(t)} \frac{1}{N-U} dU = \int_0^t \alpha dt, \quad \alpha = N\beta$$

folgt. Bekanntlich ist

$$\int \frac{1}{x} dx = \log x + k_0, \quad \int \frac{1}{N-x} dx = -\log(N-x) + k_1,$$

k_0 und k_1 Integrationskonstanten, die aber beim Übergang zu bestimmten Integralen wegfallen. Demnach hat man

$$\log X(t) - \log X(0) - \log(N - X(t)) + \log(N - X(0)) = \alpha t.$$

Mit $x_0 = X(0)$ erhält man daraus

$$\log\left(\frac{X(t)}{x_0}\frac{N-x_0}{N-X(t)}\right) = \log\left(\frac{AX(t)}{N-X(t)}\right) = \alpha t, \quad A = \frac{N-x_0}{x_0}$$

und daraus

$$\frac{A\,X(t)}{N-X(t)} = e^{\alpha t},$$

Löst man diese Gleichung nach $X(t)$ auf, erhält man

$$X(t) = \frac{(N/A)e^{\alpha t}}{1 + e^{\alpha t}/A} = \frac{N}{A\left(e^{-\alpha t} + \frac{1}{A}\right)} = \frac{N}{1 + Ae^{-\alpha t}},$$

und dies ist die Gleichung (40).

□

Index

Bradley-Terry-Luce Modell, 29

doubling time, 6

epidemiscche Kurve, 20

Itemfunktion, 26

Likelihood-Quotient, 28
logistisches Modell, 11
logistisches Wachstum, 12
Lotka-Volterra System, 15

Objektivität, spezifische, 27, 31

psychometrische Messungen, 26

stochastisches Subjekt, 27
Suffizienz, 31

Trägerkonstante, 11

Verdopplungszeit, 6
Verhulst, 10

Literatur

[1] Andersen, E. B. (1973) Conditional inference for multiple choice questionaires. *British Journal of Mathematical and Statistical Psychology* ,26, 31–44

[2] Einstein, A. (1905) Zur Elektrodynamik bewegter Körper. *Annalen der Physik*, 4. Serie 17, 891–921

[3] Frauenthal, J.C.: Mathematical Modelling in Epidemiology. Springer-Verlag, Berlin, Heidelberg, New York, 1980

[4] Holland, P. W. (1990) On the sampling theory foundations of item response theory models. *Psychometrika*, 55, 577–601

[5] Irtl, H.: On specific objectivity as a concept in measurement. In: Roskam, E., Suck, R.: Progress in Mathematical Psychology - I, Amsterdam 1987

[6] Meyer, P.S., Ausubel, J. H. (2006) Carrying cpacity: a model with logistically varying limits. *Technological Forecasting and Social Change*, 61(3), 209–214

[7] Micko, H. C. (1969) A psychological scale for reaction time measurement. *Acta Psychologica*, 30 Attention and Performance II(W.G. Koster, ed.) 324–335

[8] Murray, J.D. Mathematical Biology. Springer-Verlag, Berlin 1989

[9] Rasch, G. (1977) On specific objecitivty: An attempt at formalizing the request for generality and validity of scientific statements. *Danish Yearbook of Philosophy*, 14, 58–94

[10] Verhulst, P. F. (1838) Notice sur la loi que la population suit dans son accroissement . Corr. Math. et Phys., 10, 113-121

[11] Vorberg, D., Schwarz, W. (1990) Rasch-representable reaction time distributions. *Psychometrika*, 55, 617-632

[12] Suppes, P.,Zinnes, J.L.: Basic Measurement Theory. In Luce, R.D., Bush, R. R., Galanter, E. (eds) Handbook of Mathematical Psychology, Vol. I,New York & London 1963