Kapitel 9: Verhältniszahlen, Wachstumsraten und Aggregation

1. Arten von Verhältniszahlen 302
2. Gliederungs- und Beziehungszahlen 310
a) Gliederungszahlen 310
b) Beziehungszahlen.. 311
c) Partielle Assoziation 315
d) Allgemeine Interpretationsprobleme von Gliederungs- und Beziehungszahlen 316
3. Messzahlen 320
4. Wachstumsraten und Wachstumsfaktoren... 325
a) Wachstumsraten und Wachstumsfaktoren bei diskreter Zeit 325
b) Wachstumsraten und Wachstumsfaktoren bei stetiger Zeit 332
c) Weitere Bemerkungen zu Wachstumsraten.................................... 336
5. Aggregationsprobleme... 341
a) Begriff der Aggregation... 342
b) Aggregation von Verteilungen 343
c) Verteilung einer Linearkombination 344
d) Aggregation von Mittelwerten, Beziehungszahlen und Quoten, Struktureffekt und Standardisierung.. 346
e) Aggregation von Messzahlen und Wachstumsraten....................... 351

1. Arten von Verhältniszahlen

Absolute Größen, wie Summen oder Durchschnitte (Mittelwerte) usw. sind, isoliert betrachtet, meist wenig aussagefähig. Sie können oft erst im Vergleich mit anderen Größen richtig eingeschätzt werden und sie sind auch abhängig von Umfang und Struktur der Masse, auf die sie sich beziehen.

So mag es z.B. nicht überraschen, dass im Lande A die absolute Anzahl der Geburten größer ist als im Lande B, weil A auch mehr Einwohner hat als B. Vergleicht man A und B aber auf der Basis der Geburtenrate (Lebendgeborene/Wohnbevölkerung), so ist der unterschiedliche **Umfang** der Massen (d.h. der Länder A und B hinsichtlich der Einwohnerzahl) "ausgeschaltet". Noch aussagefähiger für die Darstellung "echter" Unterschiede hinsichtlich der Fruchtbarkeit ist ein Vergleich, der auch die unterschiedliche **Struktur** der beiden Länder hinsichtlich Alter und Geschlecht berücksichtigt. Entsprechend ist auch die absolute Zunahme des Sozialprodukts oft nur von geringem Interesse. Um die Entwicklung zu beurteilen ist es besser mit Wachstumsraten zu rechnen.

Verhältniszahlen dienen dem Vergleich von Massen, der Beschreibung von Strukturen (Verteilungen bzgl. "qualitativer" Merkmale) und der Charakterisierung der Dynamik einer Entwicklung. Sie sind neben Mittelwerten die am häufigsten auch von Nichtstatistikern benutzten statistischen Kennzahlen und in ihrer Aussage meist leicht verständlich. Sie werden aber auch sehr oft mißverstanden. Die Probleme mit ihnen sind mehr inhaltlicher als formaler Art.

Bei der Analyse nominalskalierter Variablen spielen Verhältniszahlen oft die Rolle, die Mittelwerte bei metrischen Variablen haben. Mit Messzahlen (eine spezielle Art von Maßzahlen) - aber auch mit den gerade bei Ökonomen sehr beliebten Wachstumsraten - werden Daten, die als Zeitreihe (vgl. Def. 11.1) vorliegen, auf sehr einfache Art beschrieben.

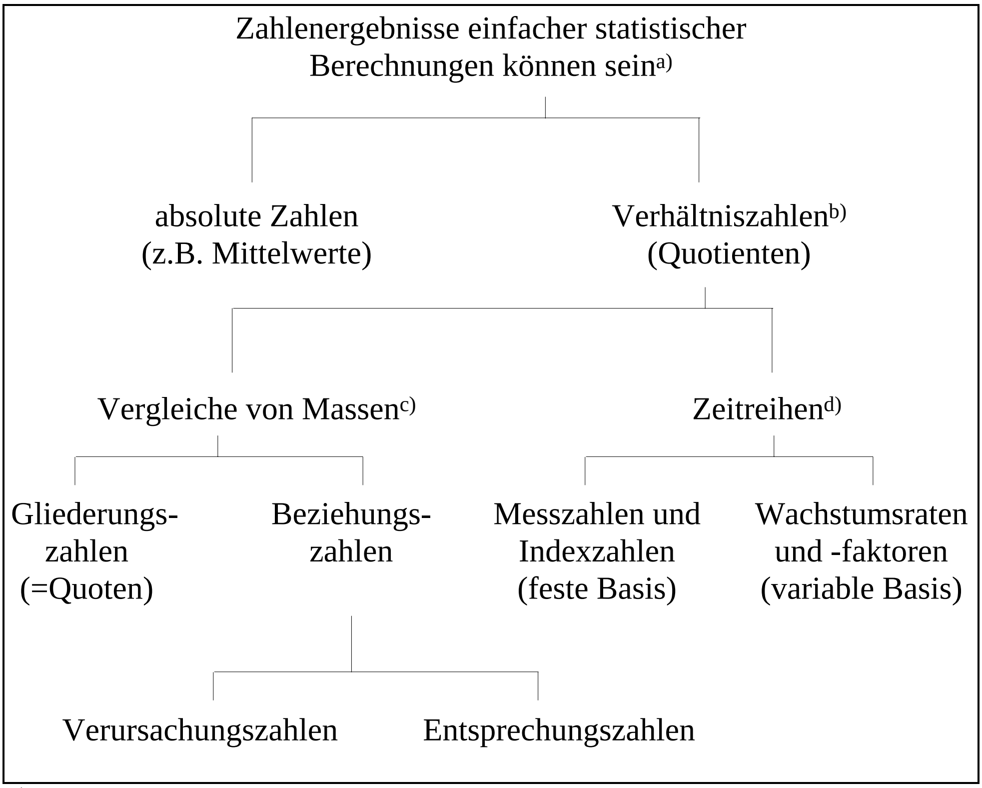
Def. 9.1: Verhältniszahlen

a) Kennzahlen, die als Quotient gebildet sind heißen **Verhältniszahlen**. Man unterscheidet zwischen Gliederungszahlen, Beziehungszahlen und Messzahlen, je nachdem, wie Zähler und Nenner des Quotienten definiert sind. Auch Wachstumsfaktoren und Wachstumsraten sind als Quotienten Verhältniszahlen im weiteren Sinne (vgl. Übers. 9.1).

b) Bei **Gliederungszahlen** G_i ist der Zähler eine Teilmenge des Nenners. Die Gesamtheit (Nennermenge) wird nach einem i.d.R. kategorialen (nominalskalierten) Merkmal in m Teilmassen zerlegt (vgl. Kap. 2 zum Begriff der Zerlegung). Mit dem Umfang n_i der i-ten Teilgesamtheit und $n = \Sigma n_i$, der Gesamtheit bzw. den Merkmalssummen S_i und $S = \Sigma S_i$ ist eine Gliederungszahl wie folgt definiert

$$(9.1) \quad G_i = \frac{n_i}{n} \quad \text{oder} \quad G_i = \frac{S_i}{S} \qquad (i = 1,2,...,n).$$

Übersicht 9.1: Arten von Verhältniszahlen



a) Zahlenergebnisse statistischer Berechnungen können auch Schätzwerte für die Parameter eines Modells sein, z.B. Regressionskoeffizienten.
b) Die englischen Begriffe sind ratios (Verhältniszahlen), rates (Beziehungszahlen), proportions (Gliederungszahlen) und relatives (Messzahlen).
c) ohne Zeitbezug (Querschnittsdaten).
d) Darstellung eines zeitlichen Ablaufs.

Eine Gliederungszahl (Quote, Anteilswert) G_i ist "dimensionslos" (genauer: G_i hat keine Maßeinheit). In der Praxis wird G_i mit 100 multipliziert und hat dann die Maßeinheit "Prozent".

c) Bei **Beziehungszahlen** sind Zähler und Nenner Umfänge oder Merkmalssummen von selbständigen Massen, die jedoch in sinnvoller Beziehung zueinander stehen sollten. Die Beziehungszahl ist deshalb auch i.d.R. nicht dimensionslos. Je nachdem, ob die Zählermasse als von der Nennermasse "verursacht" gelten kann oder nicht unterscheidet man zwischen Verursachungszahlen und Entsprechungszahlen.

Übersicht 9.2: Beispiele für Gliederungs- und Beziehungszahlen

Typ: E = Entsprechungszahl
G = Gliederungszahl
V = Verursachungszahl

Name	Typ	Zähler	Nenner
Erwerbsquote	G	Erwerbspersonen	Wohnbevölkerung
Arbeitslosenquote	G	Arbeitslose	Erwerbspersonen
Produktivität[a)b)]	V	output	input
Geburtenrate	V	Lebendgeborene	Wohnbevölkerung
Bevölkerungsdichte	E	Wohnbevölkerung	Fläche
Arealitätsziffer	E	Fläche	Wohnbevölkerung
Arztdichte	E	Anzahl der Ärzte	Wohnbevölkerung
Rentabilität	V	Gewinn	Kapital
Umschlagshäufigkeit[c)]	V	Umsatz	Bestand
durchschn. Steueraufk.	V	Steueraufkommen	Wohnbevölkerung
Hektarertrag[b)]	V	Ernteertrag	landw.Anbaufläche
Belastungsquote[d)]	E	$B(x<20,x>60)$	$B(20 \leq x \leq 60)$

a) Produktivität allgemein. Bei der Arbeitsproduktivität ist z.B. der Output zu dividieren durch den Arbeitseinsatz also speziell dem Input des Faktors Arbeit.

b) Von einigen Autoren auch als Entsprechungszahlen bezeichnet.

c) z.B. Umschlagshäufigkeit (-geschwindigkeit) eines Lagers: Lagerbewegungen/ Lagerbestand.

d) Es bedeuten $B(x < 20, x > 60)$ = Bevölkerung im Alter von unter 20 und über 60 Jahren und $B(20 \leq x \leq 60)$ = Bevölkerung im Alter von 20 bis 60 Jahren.

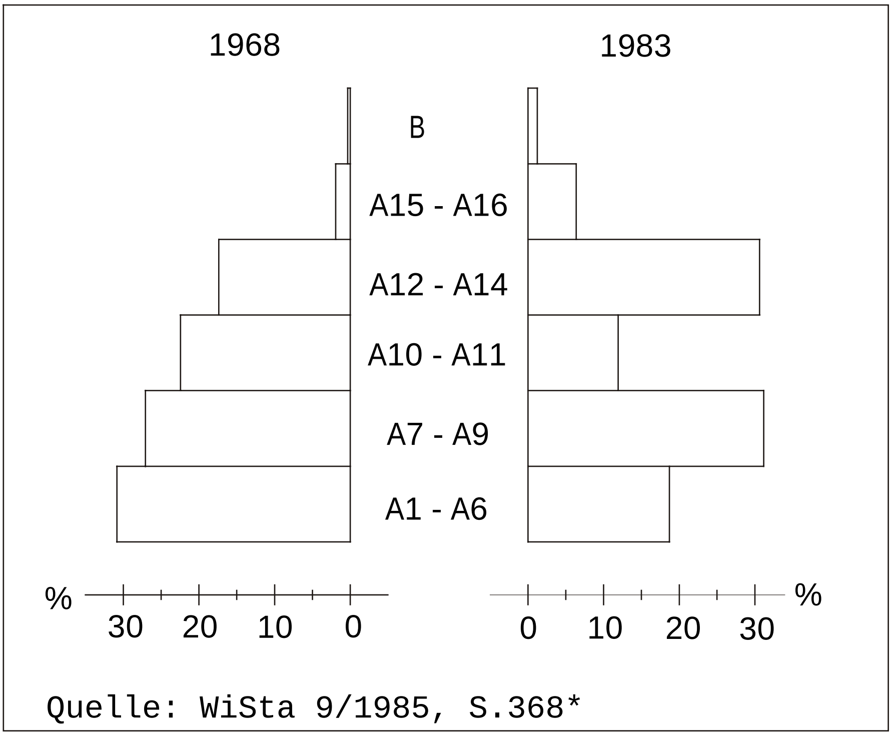
d) Eine **Messzahl** setzt einen (meist aktuellen) Wert y_t ins Verhältnis zum Basiswert y_0, wobei t die "Berichtsperiode" und 0 die (meist zurückliegende) "Basisperiode" (Referenzperiode) ist. Eine dem räumlichen Vergleich dienende Messzahl ist analog definiert. Auch Messzahlen sind wie Gliederungszahlen dimensionslos, weil Kenngrößen (Umfänge, Merkmalsbeträge) gleichartiger Massen ins Verhältnis gesetzt werden. Indexzahlen (Kap. 10) sind zusammengefaßte Messzahlen. Wachstumsraten und -faktoren werden in Def. 9.3 definiert.

Die Unterscheidung zwischen Gliederungs- und Beziehungszahlen (Übers. 9.1) wird auch mit Beispielen erläutert (Übers. 9.2).

Zum Sprachgebrauch:

Die Begriffe "Quote" und "Rate" werden in Politik, Wirtschaft und Journalismus oft völlig willkürlich verwendet. Ein krasser Fall von Verwirrung, die entstehen kann, ist Beispiel 9.2.
So wird z.B. oft von der "Arbeitslosenrate" gesprochen obgleich es sich dabei um eine Quote handelt. Oder die sog. "Belastungsquote" (vgl. Übers. 9.2) ist keine Quote, sondern eine Beziehungszahl.
Quoten sind stets Gliederungszahlen, wobei der Zähler eine echte Teilmenge des Nenners sein muss und "Raten" sind Beziehungszahlen. In diesem Sinne sind z.B. auch die Schuldenquote (Schuldenstand/Sozialprodukt) oder die Scheidungsquote (vgl. Bsp. 9.2) keine echten Quoten. Besonders inflationär wird der Begriff "Quote" bei betriebswirtschaftlichen Kennzahlen oder in der Finanzstatistik verwendet (z.B. Schuldendienstquote = Ausgaben für den Schuldendienst/Staatseinnahmen).

Abb. 9.1: Der "Stellenkegel" der Beamten 1968 und 1983



Beispiel / Lösung 9.1:

In den 70er Jahren vollzog sich im ganz besonderen Maße unter den Beamten der Bundesrepublik quasi ein "kollektiver Aufstieg", nicht nur durch eine allgemeine Erhöhung der Besoldung, sondern auch durch eine Veränderung der Struktur zugunsten der höheren Laufbahn- und Besoldungsgruppen[1*)], wie dies durch die folgende Tabelle gezeigt wird, für die geeignete Verhältniszahlen zu berechnen sind.

Als Verhältniszahlen eignen sich in diesem Fall besonders Meß- und Gliederungszahlen, die in der nachfolgenden Tabelle wiedergegeben sind.

Beamte und Richter nach Besoldungsgruppe[*)] am 2.10.1968 und am 30.6.1983 in Tausend[**)]

	Daten			Messzahl 1983	Quoten in vH	
	Besoldungs-gruppe	1968	1983	1968=100 (1)	1968 (2)	1983 (3)
1	B,R3-10,C4	4,3	19,3	448,8	0,34	1,23
2	A15-A16[a)]	24,4	100,2	410,7	1,93	6,38
3	A12-A14	220,1	481,1	218,6	17,37	30,61
4	A10-A11	284,5	187,5	65,9	22,45	11,93
5	A7-A9	334,1	489,9	142,8	27,07	31,17
6	A1-A6	390,9	293,7	75,1	30,85	18,70
	Summe	1267,3	1571,7	124,0[b)]	100[c)]	100[c)]

Quelle: Wirtschaft und Statistik Sept. 1985, S. 368[*]

*) Bekanntlich gliedert sich die Beamtenschaft in vier Laufbahngruppen (einfacher-, mittlerer-, gehobener- und höherer Dienst). Innerhalb jeder Laufbahngruppen gibt es verschiedene Besoldungsgruppen. A und B sind Besoldungsgruppen in der Verwaltung, R für Richter und C für Hochschullehrer. Die Höhe der Besoldung nimmt von der ersten bis zur sechsten Zeile ab.

**) ohne Personen in Ausbildung.

a) sowie R1/R2 und C2/C3 in der Richter- und Hochschullehrerbesoldung.

b) ein gewogenes arithmetisches Mittel der einzelnen Messzahlen, gewogen mit der Struktur gem. Spalte 2.

c) Die Summen können nur im Rahmen von Rundungsfehlern von 100 abweichen.

Die Berechnung von Gliederungszahlen (Spalten 2 und 3) und deren Darstellung in Abb. 9.1 zeigt deutlich die behauptete Strukturveränderung zugunsten der höheren Besoldungsgruppen. Das bestätigt sich auch durch die Messzahlen: man sieht dass bestimmte Gruppen weit mehr (z.B. um 348,8%) zunehmen als der Durchschnitt (+24%) und andere unterdurchschnittlich zunehmen bzw. sogar abnehmen (d.h. die Messzahl ist kleiner als 100%).

Die Abnahme der Anteile in Zeile 4 (Bes. Gr. A10/A11) und die Zunahme in Zeile 5 (Bes. Gr. A7-A9) scheinen der These vom Aufstieg durch "Strukturverbesserung" zu

---

1 vgl. Fußnoten zur folgenden Tabelle in der Lösung 9.1.

widersprechen. Dies zeigt sich aber aus der hier vorgenommenen Zusammenfassung von Besoldungsgruppen ohne Rücksicht auf Laufbahngruppen. Bei genauerer Betrachtung der Zahlen zeigt sich, dass die Strukturveränderungen weitgehend darin bestanden, dass sich 1985 sehr viele Beamte jeweils auf der für ihre Laufbahngruppe höchste Besoldungsgruppe befanden (insbesondere im gehobenen Dienst), während 1968 die Verhältnisse noch ausgeglichener waren:

einfacher Dienst

Bes. Gr.	1968	1983
A5"S"	33,2	51,5
A4	63,7	80,4
A3	74,5	17,3
A1/A2	30,5	2,4
S	201,9	151,6

mittlerer Dienst

Bes. Gr.	1968	1983
A9"S"	29,5	104,0
A7/A8	260,6	334,7
A5/A6	189,0	142,1
S	479,1	580,8

gehobener Dienst

Bes. Gr.	1968	1983
A13*)	5,0	101,9
A12	68,2	205,6
A11	197,2	99,0
A10	87,3	88,5
S	357,7	495,0

*) und A13"S" bis A15"S" (also Besoldungen für Beamte des gehobenen Dienstes die denen des höheren Dienstes entsprechen)

Die These vom kollektiven Aufstieg belegen auch Messzahlen nach den vier Laufbahn gruppen für 1983 (1968 = 100):

einfacher Dienst	75,1	unterdurchschnittliche
mittlerer Dienst	121,2	Zunahme (<24%)
gehobener Dienst	138,4	überdurchschnittliche
höherer Dienst	166,9	Zunahme (>24%)
insgesamt	124,0	

Das Beispiel zeigt auch, dass es mit einer oder einigen wenigen Verhältniszahlen oft nicht getan ist, um sich ein zutreffendes Bild zu machen und dass dieses Bild auch sehr davon abhängt, wie das Datenmaterial gegliedert ist.

Beispiel 9.2:

Der nebenstehende Text aus der Zeitung DER SPIEGEL (Nr. 37/1976) ist ein schönes Beispiel für einen Verwirr-**Text**, denn es ist sehr schwer, den wahren Sachverhalt aus dem Text zu "rekonstruieren":

a) Welche Größen werden miteinander verglichen?

b) In welchem Sinne wird das Wort "Quote" benutzt?

c) Was heißt "Unfallraten" mit Fahrzeugkilometer "korrelieren"?

d) Nennen Sie Fälle, in denen gleiche Größen (Berechnungen) mit unterschiedlichen Worten beschrieben werden (oder umgekehrt)!

Verwirr-Statistik

Einen neuerlichen Verfall der westdeutschen Verkehrssitten schienen die jüngsten Unfallziffern des Statistischen Bundesamtes in Wiesbaden zu belegen: Verglischen mit dem Juni des Vorjahres, sei im letzten Juni die Uahl der Verkehrstoten um 7,7 Prozent, die Zahl der Verletzten um 5,3 Prozent höher gewesen; im Vergleich der ersten Halbjahre 1975 und 1976 sei die Totenquote um 0,8 Prozent geringer, die Verletzenrate um 3,7 Prozent höher ausgefallen. Prompt fragt die „Frankfurter Allgemeine“: „Fahren wir wieder unvorsichtiger?“ Wohl nicht. Denn unberücksichitgt blieben bei den amtlichen Zahlen die im Vergleichszeitraum beträchtlich gestiegenen Kfz-Zulassungen und Fahrleistungen. Experten des ADAC machten sich die Mühe, das schiefe Bild aus Wiesbaden zurechtzurücken und die Unfallraten mit Fahrzeugkilometern zu korrelieren Im Juni-Vergleich ergab sich dananch, statt eines Anstiegs der Todesfälle um 7,7 Prozentpunkte, eine gegenüber 1975 um 0,4 Punkte niedrigere Zahl der tödlichen Unfälle; die absolut um 5,3 Prozent höhere Verletztenquote leigt bei korrekter Auswertung um 2,5 Prozent unter der des Juni 1975. Ähnliches gilt für den Halbjahresvergleich: 0,8 Prozent weniger Tote als im Vorjahr bedeuten in Wahrheit ein Absinken um vier Prozent; und die Verletzten-Rate lag, statt um 3,7 Prozent, nur um 0,4 Prozent höher.

DER SPIEGEL, Nr. 37/1976

Lösung 9.2:

a) Verglichen werden verschiedene Wachstumsraten und zwar

- im "Jahresvergleich" Juni 76 zu Juni 75 und im "Halbjahresvergleich" Januar bis Juni 76 zu Januar bis Juni 75,
- der Anzahl von Unfällen, Unfallverletzten und Unfalltoten mit und ohne Berücksichtigung der 1975 bzw. 1976 gefahrenen Entfernungen.

b) "Quote" ist ganz offensichtlich die relative Zunahme (also Wachstumsrate), wie z.B. bei "Toten-" oder "Verletztenquote", wobei letztere beim Halbjahresvergleich auch Verletztenrate genannt wird. Auch die nicht definierte Unfallrate ist wohl so zu verstehen. Im Text sind Quote und Rate ganz nach Belieben und immer falsch verwendet worden, so dass der Text verwirrend und fast völlig unverständlich ist.

c) Gemeint ist offenbar, dass durch die Anzahl der 1975 bzw. 1976 gefahrenen Fahrzeugkilometer (die den Zahlenangaben zufolge um etwa 8% zugenommen hat) dividiert wurde.

d) Gleich gesetzt werden nicht nur Quoten und Raten, sondern auch Prozent, Prozentpunkte und "Punkte", ferner Tote, Verkehrstote, Totenquote, Todesfälle und tödliche Unfälle.

Fazit: Es ist unschön, wenn jemand schulmeisterlich Statistiken kritisiert und dabei selbst die einfachsten Begriffe der Statistik durcheinanderwirft.

2. Gliederungs- und Beziehungszahlen

a) Gliederungszahlen

Eigenschaften von Gliederungszahlen

Es ergibt sich als unmittelbare Folgerung aus Gl. 9.1:

(9.2) $0 \leq G_i \leq 1$ und

(9.3) $\Sigma G_i = 1$ $(i = 1,2,...,m)$.

Zu den Gliederungszahlen gehört auch die relative Häufigkeit h_i der Merkmalsausprägung x_i bzw. der i-ten Größenklasse der Variablen X (Def. 3.1). Gliederungszahlen beziehen sich aber auch auf kategorial gebildete Teilgesamtheiten, nicht nur auf metrisch skalierte Variablen und sie können auch Quotienten von Merkmalssummen sein, nicht nur von Häufigkeiten (Umfänge von Massen und Teilmassen).

Interpretationsprobleme

Häufig werden aus Gliederungszahlen Schlüsse gezogen, insbesondere über Entwicklungen, die nur in Verbindung mit entsprechenden Beziehungszahlen und deren Veränderung zulässig sind. Hierzu zwei Beispiele:

1. Man schließt auf ein besonders hohes Unfallrisiko auf dem Beifahrersitz eines PKW, weil die Zahl der Unfallopfer, gegliedert nach den vier Sitzen eines PKW, auf denen die verunfallte Person jeweils gesessen hat, im Falle des Beifahrersitzes am größten ist. Der Fehlschluß rührt daher, dass nicht berücksichtigt wird, wie oft die

einzelnen Sitze besetzt sind. So kann das Unfallrisiko auf einem der hinteren Sitze größer sein, wenn man berücksichtigt, dass diese nicht so oft besetzt sind.

2. Ein zunehmender Anteil Krebstoter an der Gesamtzahl der Sterbefälle muss nicht ein höheres Krebsrisiko bedeuten. Mögliche Gründe: die allgemeine Sterblichkeit sinkt, die Krebssterblichkeit aber unterdurchschnittlich oder die Altersstruktur (auch wegen des Rückgangs der Sterblichkeit) hat sich verändert.

Auf ein weiteres Interpretationsproblem, nämlich die Strukturabhängigkeit einer Verhältniszahl, wird an späterer Stelle gesondert hingewiesen. Auch hierfür ein Beispiel:

Eine Klausur gilt als schwerer, weil die Durchfallquote höher ist. Tatsächlich ist aber die Struktur der Klausurteilnehmer anders: es gab z.B. in der "schwierigeren" Klausur mehr Wiederholer, deren Durchfallquote oft größer ist als bei Studenten in ihrem ersten Klausurversuch.

Gliederungszahlen und Wahrscheinlichkeiten

Der Anteil der Knabengeburten an der Gesamtzahl der Geburten (eine Gliederungszahl) ist eine Schätzung für die Wahrscheinlichkeit einer Knabengeburt. Die Wahrscheinlichkeit ist wie eine Gliederungszahl eine Zahl zwischen 0 und 1. Sie bezieht sich aber nicht auf eine empirisch beobachtete und daher notwendig endliche Gesamtheit, sondern auf Ergebnisse eines prinzipiell beliebig (unendlich) oft wiederholbaren Zufallsversuch.

Zieht man in einer Lotterie bei zehn Losen 9 Nieten (Gewinn**quote** 1/10 also, 10%) so ist das natürlich kein Widerspruch zu einer Gewinn**wahrscheinlichkeit** von 20% (ein anderer Teilnehmer hat mehr Glück und erzielt eine Quote von 30%). Auch geeignet gebildete Raten (also Beziehungszahlen, z.B. die Todesrate) stehen im Zusammenhang mit Wahrscheinlichkeiten (z.B. der Sterbewahrscheinlichkeit).

b) Beziehungszahlen

Allgemeine Eigenschaften

1. Dimension:
   Anders als Gliederungszahlen und Messzahlen setzen Beziehungszahlen verschiedenartige Massen ins Verhältnis. Sie sind deshalb auch nicht dimensionslos, sondern haben eine Maßeinheit.

2. Umkehrbarkeit:
   Beziehungszahlen sind grundsätzlich umkehrbar wie das Beispiel der Bevölkerungsdichte und der Arealitätsziffer (Übers. 9.2) zeigt. Einmal

mißt man die Anzahl der Personen je Fläche (Bevölkerungsdichte), im anderen Fall (Arealitätsziffer) die durchschnittliche Fläche um eine Person.

Ein ökonomisches Beispiel für die Umkehrbarkeit ist die

$$\boxed{\text{Kapitalproduktivität} = \frac{\text{Output (z.B. Inlandsprodukt)}}{\text{Kapitaleinsatz (Anlagevermögen)}}}$$

Der reziproke Wert ist der (durchschnittliche) Kapitalkoeffizient. Die Kapitalproduktivität zeigt, wieviel DM Output mit einer eingesetzten DM Kapital zu erwirtschaften ist und der Kapitalkoeffizient zeigt wie groß der Kapitalbedarf ist (wieviel investiert werden muss) um eine DM Output zu erwirtschaften.

3. Zusammenhang mit Mittelwerten:

   Beziehungszahlen sind nichts anderes als Mittelwerte, wenn eine Merkmalssumme (Zähler) zu einer entsprechend abgegrenzten Personengesamtheit (Nenner) ins Verhältnis gesetzt wird. Ein Beispiel ist das durchschnittliche Steueraufkommen (Übers. 9.2) oder die Beziehungszahl Bierkonsum/Einwohner, der durchschnittliche Bierverbrauch (der dann allerdings nicht aufgrund der individuellen Verbrauchsmengen errechnet ist).

Alle Verhältniszahlen sind ferner Mittelwerte durch Aggregation (vgl. Kap. 9, Abschn. 5), d.h. sie sind in dem Sinne Mittelwerte dass eine auf die Gesamtmasse bezogene Verhältniszahl ein Mittel der entsprechenden Verhältniszahlen der Teilmassen ist (so ist z.B. die rohe Todesrate ein Mittel der altersspezifischen Todesraten).

Verursachungszahlen

1. Der Begriff "Verursachungszahl" ist problematisch weil er eine monokausale Verursachung suggeriert und der Quotient ist meist nicht eindeutig, denn

   a) es ist oft nicht klar welche Nennergröße zu nehmen ist und
   b) die erwähnte Umkehrbarkeit ist zu beachten.

   zu a)

   Ein bekanntes Problem ist die zutreffende Beschreibung eines Unfallrisikos mit PKWs (oder entsprechend die Messung des Flugrisikos) durch eine geeignete Maßzahl. Soll man die Anzahl der Unfälle mit PKWs dividieren durch

   - die Einwohnerzahl,

- den Bestand an PKWs,
- die Anzahl der gefahrenen Kilometer oder
- das Produkt dieser Größe mit der Anzahl der beförderten Personen (also die "Personenkilometer")?

zu b)

Man mag die Kapitalproduktivität als Verursachungszahl bezeichnen, wohl kaum aber den reziproken Kapitalkoeffizienten. Im Zähler der Kapitalproduktivität steht übrigens der Output *aller* Produktionsfaktoren, nicht nur der des Kapitals, weshalb diese Beziehungszahl auch nicht kausal interpretiert werden darf.

2. Verursachungszahlen sind meistens Verhältnisse von Bestands- und Bewegungsmassen, wie z.B. die

$$\text{Fruchtbarkeitsrate}^{a]} = \frac{\text{Lebendgeborene eines Jahres}}{\text{gebärfähige Frauen}^{\,b]} \text{ in der Mitte}^{\,c]} \text{ dieses Jahres}}$$

a] früher war die Bezeichnung Fruchtbarkeits"ziffer" üblicher, entsprechend sprach man früher von der Sterbeziffer oder der Geburtenziffer (heute: Todesrate, Geburtenrate [besser: Geborenenrate]).

b] Frauen im Alter zwischen 15 und 45 Jahren.

c] oder im Durchschnitt des Jahres.

Begriff der "Rate"

Eine "Rate" bezieht die Häufigkeit eines Ereignisses in einem Zeitintervall auf die durchschnittliche Anzahl der Einheiten, die (zu Beginn des Intervalls) dem Risiko des Ereignisses ausgesetzt waren.
Sie steht in einem Zusammenhang mit

- der Wachstumsrate und
- der Wahrscheinlichkeit.

Die Differenz zwischen der Geburten- und der Todesrate ist die "natürliche" (durch Geburt und Tod) Wachstumsrate der Bevölkerung.
Die altersspezifische Sterberate (als Verursachungszahl) z.B. der 60jährigen Männer könnte eine Schätzung der Sterbewahrscheinlichkeit sein, wenn die Sterbefälle 60-jähriger Männer des Jahres t bezogen wären auf den Bestand an Lebenden *zu Beginn* des Jahres t (üblich ist aber: in der *Mitte*, oder im *Durchschnitt* des Jahres t) .
Gemeint ist dabei die *einjährige* Sterbewahrscheinlichkeit 60-jähriger Männer, d.h. die (bedingte) Wahrscheinlichkeit das Alter 61 nicht mehr zu erreichen, *wenn* man (Bedingung!) das Alter von 60 erreicht hat. Die *unbedingte* Sterbewahrscheinlichkeit, irgendwann einmal zu sterben, egal in welchem Alter, ist natürlich immer 1.

Beispiel 9.3:

Für die Bundesrepublik ("alte" Bundesländer) galten 1990 etwa (gerundet) die folgenden Zahlen:

Ehescheidungen	140 Tausend	Wohnbevölkerung	60 Millionen
Eheschließungen	400 Tausend	Bestand an Ehen	15 Millionen
Lebendgeborene	600 Tausend		

Man berechne die folgenden Beziehungszahlen

1. Lebendgeborene je
   a) 1000 Einwohner
   b) 100 bestehende Ehen
2. Ehescheidungen je
   a) 1000 bestehende Ehen
   b) 100 geschlossene Ehen

und interpretiere die Ergebnisse!

Lösung 9.3:

Die vier Beziehungszahlen sind leicht zu berechnen:

1a) 10 also 1% (das ist die sog. rohe [allgemeine] Geburtenrate),

1b) 4 also 4%.

2a) 140/15 = 9,333 also 0,93%. Diese Größe **und** 2b) sind auch bekannt als Scheidungs-"quote",

2b) (140/400)100 = 35 (also 35%).

Interpretation (verbreitete Fehlinterpretationen):

Bei der Interpretation der Geburtenrate [Geburtenziffer] und der Scheidungsquote (die keine Quote, sondern eine Beziehungszahl ist) wird oft vergessen, dass es sich um die Geburten bzw. Scheidungen **eines Jahres** handelt. Die Interpretation von 1b): "von 100 Ehen **haben** nur vier ein Kind" widerspricht ganz offensichtlich der Erfahrung. Richtig wäre: nur vier (von 100) Ehepaare haben **in diesem Jahr** ein Kind **bekommen**. Auf der gleichen Linie liegt die Fehlinterpretation von 2a): noch nicht einmal 1% der Ehen endet vor dem Scheidungsrichter (diese Interpretation wäre nur zulässig, wenn die durchschnittliche Verweildauer in der Ehe ein Jahr wäre). Man kann auch nicht sagen [im Sinne von 2b)]: etwa jede 3-te Ehe (35%) wird geschieden, denn es sind die Scheidungen eines Jahres (von Ehen, die in einem Zeitraum von vielen früheren Jahren geschlossen wurden), die verglichen werden mit den Eheschließungen dieses gleichen Jahres. Eine solche Interpretation wäre aber zulässig, wenn der Bestand an Ehen dem Modell der stationären Bevölkerung (=Sterbetafelbevölkerung) entspräche und diese Scheidungsquote (gem. 2b) jedes Jahr 100 % (statt 35%) betragen würde (vgl. Kap. 12).

c) Partielle Assoziation

Gegeben seien drei dichotome Merkmale X,Y und Z. Unter der partiellen Assoziation erster Ordnung versteht man die Assoziation zwischen zwei Merkmalen (etwa X und Y) in einer der beiden Teilgesamtheiten bezüglich des dritten Merkmals (also wenn $Z = z_1 = 1$ und wenn $Z = z_o = 0$ ist) . Von Interesse sind die Beziehungen zwischen diesen beiden partiellen Assoziationen und der Assoziation von X mit Y in der Gesamtheit (die man auch partielle Assoziation nullter Ordnung nennen könnte) und deren Interpretation. Mit dem Kreuzprodukt (Gl. 7.47) als Assoziationsmaß erhält man:

$$(9.4) \quad |xy| = p_z \cdot |xy;z_1| + q_z \cdot |xy;z_0| + \frac{|xz| \cdot |yz|}{p_z q_z}$$

Hierin sind $|xy;z_1|$ und $|xy;z_0|$ die beiden partiellen Assoziationen, p_z ist der Anteil der Merkmalsträger mit $Z = z_1 = 1$ und $q_z = 1 - p_z$ der Anteil der Merkmalsträger mit $Z = z_0 = 0$.
Gl. 9.4 hat folgende Interpretation:
Die Assoziation zwischen X und Y in der (inhomogenen) Gesamtheit ist eine Summe der "**internen Assoziation**" (AI)

$$(9.4a) \quad AI = p_z \cdot |xy;z_1| + q_z \cdot |xy;z_0|$$

und der **marginalen Assoziation** AM, d.h. der auf die eindimensionalen Randverteilungen bezogenen Beziehung zwischen X und Y mit Z (der Kontrollvariablen):

$$(9.4b) \quad AM = \frac{|xz| \cdot |yz|}{p_z q_z} = \frac{|xz| \cdot |yz|}{s^2_z},$$

so dass man erhält

$$(9.4c) \quad |xy| = AI + AM.$$

Man beachte, dass die Kreuzprodukte |xy|, |xz| und |yz| jeweils Kovarianzen darstellen und dass die Varianzen der dichotomen Merkmale $s_x^2 = p_x q_x$ und s_y^2 und s_z^2 entsprechend definiert sind.
Für die Interpretation von besonderem Interesse sind die folgenden beiden Spezialfälle:

1) Scheinkorrelation AI = 0 (die Assoziation |xy| ist ausschließlich auf die Assoziation von X und Y mit Z zurückzuführen) und
2) AM = 0 (die Assoziation |xy| ist das gewogene Mittel der partiellen Assoziationen), weil X und Y nicht mit Z korreliert sind.

Im ersten Fall gilt für die Vierfelderkorrelationen (Phi-Koeffizienten gem. Gl. 7.46), die hier mit r bezeichnet werden sollen,

(9.5) $r_{xy} = r_{xz}\ r_{yz}$.

Das ist die Analogie zum Verschwinden der partiellen Korrelation bei metrisch skalierten Merkmalen ($r_{xy.z} = 0$, was kennzeichnend ist für Scheinkorrelation zwischen X und Y).

d) Allgemeine Interpretationsprobleme von Gliederungs- und Beziehungszahlen

In diesem Abschnitt sollen die folgenden Probleme im Umgang mit Gliederungs- und Beziehungszahlen betrachtet werden:

1. die Konstruktionsprinzipien von Beziehungszahlen:
   - Bildung homogener Massen
   - Ausscheiden unbeteiligter Massen
2. das Problem der Strukturabhängigkeit, d.h.
   - der Unterschied zwischen Verhältniszahlen (insbesondere Gliederungs- und Beziehungszahlen) kann allein strukturbedingt (und nicht "echt") sein,
   - das Simpson-Paradoxon
3. die Scheinkorrelation.

Die Probleme Nr. 2 und 3 sind Folge der Abhängigkeit der untersuchten Beziehung von einer "dritten" Variablen.

1. Konstruktionsprinzipien

Der oberste Grundsatz der Konstruktion von Beziehungszahlen (oder allgemeiner Maßzahlen) ist:

Äquivalenten (in den relevanten Aspekten gleichen) Sachverhalten ist ein gleicher Zahlenwert der Maßzahl zuzuordnen, nicht-äquivalenten Sachverhalten dagegen ein verschiedener Zahlenwert.

Ist für den Ernteertrag allein die Größe der Anbaufläche relevant, weil z.B. der Ertrag der Fläche proportional ist, so ist der Hektarertrag (Übers. 9.2) eine sinnvolle Beziehungszahl. Andere Einflußfaktoren, wie z.B. Klima, Bodenart, Bearbeitungstechnik [z.B. Düngung], oder Art [Rechtsform usw.] des landwirtschaftlichen Betriebs sollten für die Größe des Ertrags irrelevant sein.

Wenn das nicht der Fall ist, dann sollten (am Beispiel des Hektarertrags) nur Erträge von Flächen bei *gleichem* Klima, *gleicher* Bodenart usw. in die Berechnung der entsprechenden Verhältniszahlen einbezogen werden.

Häufig sind deshalb Zähler und Nenner einer Verhältniszahl zu spezifizieren, um die verglichenen Massen homogener zu machen, denn:
Nur homogene Massen repräsentieren einen einheitlichen Ursachenkomplex; sind die Massen inhomogen, so nivellieren sich charakteristische Unterschiede, die herauszustellen gerade die Aufgabe einer Beziehungszahl sein kann. Die Homogenisierung geschieht durch:

- Ausscheiden unbeteiligter Teile einer Masse:
  Ein Land A kann einen geringeren Bierverbrauch "pro Kopf" der Bevölkerung haben als das Land B, weil die Altersstruktur anders ist. Es kann sinnvoller sein, Säuglinge und Kinder aus der Nennermasse auszuscheiden und nicht die Gesamtbevölkerung, sondern die Bevölkerung ab eines bestimmten Alters als Bezugsgröße zu wählen.

- Bildung spezifischer Verhältniszahlen:
  Es ist z.B. in der Bevölkerungsstatistik üblich, viele Verhältniszahlen nach Alter und Geschlecht zu differenzieren, etwa die Erwerbsquoten: der Verlauf der altersspezifischen Erwerbsquoten der Frauen zeigt u.a. deutlich den Einfluß der Familienbildung (in einem Altersintervall von ca. 25 bis 40 Jahre) auf die Erwerbsbeteiligung von Frauen.

2. Strukturabhängigkeit und Simpson-Paradoxon

Wie an späterer Stelle (Abschn. 5d) gezeigt wird, ist eine Verhältniszahl stets ein gewogenes Mittel von speziellen ("spezifischen", d.h. für Teilgesamtheiten gebildeten) Verhältniszahlen des gleichen Typs. Folgerungen:

- Deshalb können sich zwei Verhältniszahlen allein durch die Gewichtung unterscheiden (***Struktureffekt***; Beispiel 9.4);
- ***Standardisierung*** (vgl. Abschn. 5d) ist ein Verfahren, um beim Vergleich von Verhältniszahlen den Struktureffekt auszuschalten;
- Aufgrund unterschiedlicher Gewichtung (Struktur) kann im Extremfall das ***Simpson Paradoxon*** auftreten (Beispiel 9.5).

Def. 9.2: Simpson Paradoxon

Die Tatsache, dass ein Mittelwert oder eine Verhältniszahl (z.B. eine Quote, ein Anteilswert) für eine Gesamtheit A größer sein kann, als für eine andere Gesamtheit B, obgleich diese Größe (Mittelwert oder Verhältniszahl) in allen Teilgesamtheiten von A kleiner ist als in denen von B, ist bekannt als "Simpson-Paradoxon" (nach Th. Simpson 1710 - 1761).

Beispiel 9.4:

(Strukturabhängigkeit): Die Sterbeziffer [Todesrate] (= Zahl der Gestorbenen je 1.000 Lebende) von Geistlichen ist viel höher (0,55%) als die der Bergarbeiter (0,15%) [fiktive Zahlen]:

Altersklasse	Geistliche		Bergarbeiter	
	Lebende	Gestorbene	Lebende	Gestorbene
unter 50	100	10	900	90
über 50	900	540	100	60
insgesamt	1000	550	1000	150

Kann man aus den Angaben schließen, dass der Beruf des untertage arbeitenden Bergmanns "gesünder" ist, als der des Geistlichen?

Lösung 9.4:

Man kann natürlich nicht in der Weise schließen, wie dies die Fragestellung nahelegt. Der Grund ist die unterschiedliche Altersstruktur der Geistlichen und der Bergarbeiter. Der Anteil der unter 50-jährigen ist bei den Geistlichen 10% und bei den Bergarbeitern 90% (entsprechend sind die Anteile der über 50-jährigen 90% und 10%). Die altersspezifischen Todesraten (Sterbeziffern) sind für beide Berufe in beiden Altersklassen gleich, nämlich bei

den unter 50-jährigen : 10%

den über 50-jährigen : 60%.

Die rohen Todesraten sind (mit den Anteilen 10% und 90% für die Altersstruktur) gewogene arithmetische Mittelwerte:

für die Geistlichen: $0{,}55 = 0{,}1 \cdot 0{,}1 + 0{,}9 \cdot 0{,}6$

für die Bergarbeiter: $0{,}15 = 0{,}9 \cdot 0{,}1 + 0{,}1 \cdot 0{,}6$.

Beispiel 9.5:

(Simpson-Paradoxon): Das Beispiel 9.4 wird wie folgt modifiziert:

Alters-klasse	Geistliche			Bergarbeiter		
	Lebende	Gestorbene	Rate[a)]	Lebende	Gestorbene	Rate[a)]
< 50	100	10	0,10	600	80	0,13
≥ 50	900	540	0,60	400	280	0,70
S[b)]	1000	550	0,55	1000	360	0,36

a) Todesraten.
b) bzw. Durchschnitt.

Die altersspezifischen Todesraten der Bergarbeiter sind in allen (beiden) Altersklassen höher, als die der Geistlichen und trotzdem ist die rohe (gesamte) Todesrate der Bergarbeiter niedriger, als die der Geistlichen (Simpson-Paradoxon). Woran liegt das?

Lösung 9.5

Der Struktureffekt (Bergarbeiter sind jünger), der dahingehend wirkt, dass die rohe Todesrate der Bergarbeiter kleiner sein müßte, als die der Priester, wirkt dem echten Unterschied in der Sterblichkeit (Todesraten der Bergarbeiter in allen Altersklassen größer) entgegen.

3. Scheinkorrelation

Das Konzept der Scheinkorrelation ist in Def. 7.9 definiert (Korrelation zwischen zwei Variablen nur weil diese gemeinsam abhängig sind von einer dritten Variablen) und wird im folgenden Beispiel demonstriert.

Beispiel 9.6:

Einer (fiktiven) Statistik zufolge ergaben sich die folgenden Daten über die Unfallhäufigkeit von Männern und Frauen:

Autounfall	Männer	Frauen	Summe
wenigstens einmal	3.122	2.255	5.377
nie	3.958	4.695	8.653
Summe	7.080	6.950	14.030

Beispiel entnommen aus dem Lehrbuch von H. Zeisel, Say it with Figures (deutsche Übersetzung: Die Sprache der Zahlen, Köln, Berlin 1970, S. 126).

Kann man aufgrund dieser Zahlen schließen, dass Frauen bessere (sicherere) Autofahrer sind als Männer?

Lösung 9.6:

Eine Schlussfolgerung in der genannten Weise ist sehr verbreitet. Sie wird auch nahegelegt durch die Berechnung von Quoten. Betrachtet man die "Quote der mindestens einmal Verunfallten", so ist sie bei den Männern 3122/7080 = 0,441 (also 44,1%) und bei den Frauen 2255/6950 = 0,324, also nur 32,4% (für die Gesamtheit ist die Quote natürlich ein gewogenes Mittel der beiden Quoten; sie beträgt 5377/14030 = 0,383), so dass Frauen scheinbar die besseren Autofahrer sind als Männer.

Bei genauerem Hinsehen kann sich dies aber als Scheinkorrelation erweisen. H. Zeisel untersucht die "Fahrhäufigkeit" (>10.000 oder ≤ 10.000 Meilen im Jahr) als dritte, die Scheinkorrelation erzeugende Variable mit folgenden Zahlen:

	häufiges Fahren			seltenes Fahren		
Unfall	Männer	Frauen	Σ	Männer	Frauen	Σ
wen. Einmal	2.605	996	3.601	517	1.259	1.776
nie	2.405	919	3.324	1.553	3.776	5.329
Summe	5.010	1.915	6.925	2.070	5.035	7.105

Berechnet man jetzt die entsprechenden Quoten, so sind sie in den Teilgesamtheiten "häufiges-" und "seltenes Fahren" für Männer und Frauen jeweils gleich, was ein Zeichen dafür ist, dass die Korrelation zwischen Geschlecht und Fahrtüchtigkeit nur eine Scheinkorrelation ist.

Man kann den Nachweis der Scheinkorrelation auch mit Assoziationsmaßen führen. Der Zusammenhang von Gl. 9.4 stellt sich hier mit X: Geschlecht, Y: Unfallhäufigkeit und Z: Fahrhäufigkeit wie folgt dar:

(9.4) $|xy| = p_z \cdot |xy;z_1| + q_z \cdot |xy;z_0| + |xz| \cdot |yz| / p_z q_z$

$|xy| = 0{,}0291225$, $f = r_{xy} = 0{,}1198$ (Phi-Koeffizient)

$p_z = 6925/14030 = 0{,}4936$ und entsprechend $q_z = 0{,}5064$

$|xy;z_1| = -0{,}00002888$ ($f = -0{,}000129$ also » 0)

und $|xy;z_0| = -0{,}00006012$ ($f = -0{,}0003056$ also » 0).

Das Verschwinden der beiden partiellen Assoziationen ist ein Zeichen für Scheinkorrelation. Für die marginalen Beziehungen erhält man: $|xz| = 0{,}108013$ ($f = r_{xy} = -0{,}4321$) und $|yz| = 0{,}067498$ ($r_{yz} = 0{,}2777$), so dass $r_{xy} \approx r_{xz}\, r_{yz}$. Ferner ist AM = 0,029167 ≈ $|xy| = 0{,}0201225$ und AI ≈ 0.

3. Messzahlen

Die Messzahl m_{0t} (d.h. zur Basis 0, Berichtszeit t) einer Variablen Y ist nach Def. 9.1 die Größe:

$$(9.6) \quad m_{0t} = \frac{y_t}{y_0},$$

bei der diskreten Zeitvariable t = 0,1,2,...,T bzw. die mit 100 multiplizierten Größen

$$(9.6a) \quad m_{0t},^{*} = 100 m_{0t} = 100\frac{y_t}{y_0}.$$

Die Größe t kann, muss aber nicht "Zeit" bedeuten. Messzahlen können z.B. auch dem räumlichen Vergleich dienen, wenn 0 das Basisland und t das Vergleichsland ist.

Eigenschaften von Messzahlen:

1. Messzahlen haben die in Übers. 9.3 definierten und leicht zu verifizierenden Eigenschaften, die nicht unbedingt auch für Indexzahlen (Kapitel 10) gelten müssen.

2. Aus Gl. 9.6 folgt, dass die Bildung von Messzahlen eine Lineartransformation darstellt: y_t wird mit der Konstanten $(y_0)^{-1}$ multipliziert, d.h. die Entwicklung der Zeitreihe y_t (also der Folge $y_{1,}$ y_2,..., y_t wird in Einheiten [bzw. bei $m_{0t},^{*}$ in Prozent] des Basiswerts y_0 dargestellt.

3. Das wohl wichtigste methodische Problem ist die ***Wahl der Basisperiode***, z.B. des Basisjahres: Die Regel ist, ein "Normaljahr" (ohne auffallende Besonderheiten) zu wählen damit die nachfolgende Entwicklung nicht unter- oder überzeichnet wird. Man kann sich an Beispielen leicht klar machen, welchen Effekt die Wahl eines extremen Jahres als Basisjahr hätte.

Übersicht 9.3: Eigenschaften von Messzahlen

Eigenschaft	Inhalt der Forderung
Identität	$m_{00} = m_{tt} = 1$ ($m^*_{00} = m^*_{tt} = 100$) Identität von Basis und Berichtsperiode
Dimensionalität	$m_{(a)0t} = ay_t/ay_0 = m_{0t} = y_t/y_0$ Unabhängigkeit von der Maßeinheit der Messwerte
Zeitumkehrbarkeit (Reversibiltät)	$m_{t0} = m_{0t}^{-1}$ Vertauschung von Basis- und Berichtsperiode ($m_{t0}m_{0t} = 1$)
Zirkularität (Transitivität, Verkettbarkeit)	für je drei Perioden 0,s und t gilt $m_{0t} = m_{0s}m_{st}$ (= Verkettung; Folgerung: $m_{st} = m_{0t}/m_{0s}$ [= Umbasierung])
Faktorumkehrprobe	ist für alle Perioden die Größe W das Produkt aus P und Q so gilt für die entsprechenden Messzahlen $m^W_{0t} = m^P_{0t} \cdot m^Q_{0t}$ *)

*) z.B. eine Wertmesszahl m^W_{0t} ist das Produkt aus Preis- und Mengenmesszahl.

4. Zweck der Messzahlenbildung ist die Vergleichbarmachung von Wachstumsvorgängen, die sich auf einem unterschiedlichen absoluten Niveau abspielen, z.B. der Staatsverbrauch und der meist fast dreimal so große Private Verbrauch in Beispiel 9.7.

5. Die Messzahlen $m^{(x)}_{0t} = x_t/x_0$ und $m^{(y)}_{0t} = y_t/y_0$ der Zeitreihen X_t und Y_t haben für alle Perioden t = 1,2,...,T die gleichen Wachstumsraten wie die Zeitreihen X_t und Y_t selber.

6. Die Messzahl m^z_{0t} der ***Beziehungszahl*** $z_t = y_t/x_t$ ist das Verhältnis (die Beziehungszahl) der Messzahlen m^y_{0t} und m_{0t},^{x} also $m^z_{0t} = m^y_{0t}/m^x_{0t}$. Praktische Bedeutung: Berechnung der Messzahlen der Produktivität aufgrund der Messzahlen von Output und Input ohne Rückgriff auf die absoluten Größen von Output und Input.
   Ist Z ein ***Produkt***, etwa $z_t = y_t \cdot x_t$ so gilt (Faktorumkehrprobe!) $m^z_{0t} = m^x_{0t} \cdot m^y_{0t}$.

Beispiel / Lösung 9.7:

Gegeben sind die Werte (zu jeweiligen Preisen) für den Privaten Verbrauch (PV) und den Staatsverbrauch (SV) in Mrd. DM in der Bundesrepublik Deutschland, für die Messzahlen zur Basis 1980 zu bilden sind (bei der mit 100 multiplizierten Messzahl m_{0t},* schreibt man üblicherweise die an sich unsinnige "Gleichung" 1980 = 100):

Daten			Messzahlen m_{0t} zur Basis 1980		m_{0t},* (für SV)
Jahr	PV	SV	PV	SV	1980=100
1980	840,78	297,79	1,0	1,0	100
1981	887,85	318,16	1,0560	1,0684	106,84
1982	918,05	326,19	1,0919	1,0954	109,5
1983	964,16	336,21	1,1467	1,1290	112,90
1984	1003,57	350,23	1,1936	1,1761	117,61
1985	1038,34	365,66	1,2350	1,2279	122,79
1986	1068,61	382,72	1,2710	1,2852	128,52
1987	1112,68	396,97	1,3234	1,3331	133,31
1988	1156,81	411,46	1,3759	1,3817	138,17

Quelle: Jahresgutachten des Sachverständigenrats 1989/90, Tab 25*

Man erkennt unschwer, dass der Staatsverbrauch von 1980 bis 1988 nur geringfügig stärker angestiegen ist (nämlich um 38,2%, Prozent von 1980), als der Private Verbrauch (der um 37,6% gestiegen ist). Beim Vergleich der Ursprungszahlen 411,46 und 1156,81 für 1988 mit 297,79 und 840,78 für 1980 wäre das auf einen Blick kaum möglich zu erkennen.
Außerdem ist leicht zu sehen (vgl. Abb. 9.2), dass es auch Perioden gibt, in denen SV langsamer ansteigt, als PV, auch wenn über die ganze Zeitspanne (von 1980 bis 1988) SV stärker angestiegen ist, als PV.

Beispiel 9.8:

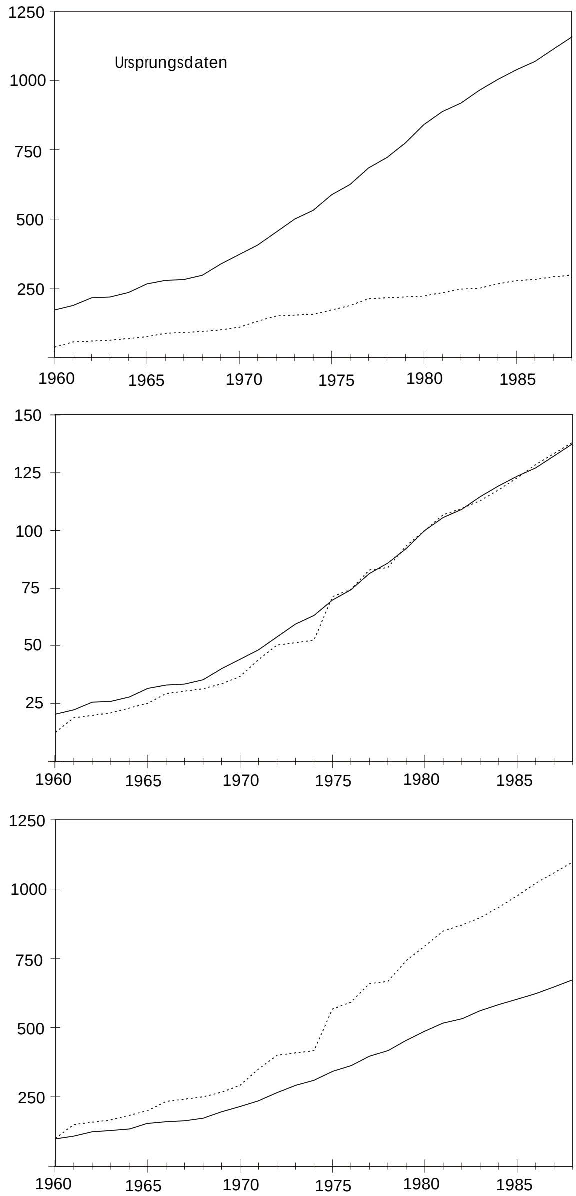
Die Messzahlen des Beispiels 9.7 zur Basis 1980 sind umzubasieren auf das "neue" Basisjahr 1985!

Lösung 9.8:

Eine Neuberechnung sämtlicher Messzahlen aus den Ursprungswerten ist nicht erforderlich. Zur Berechnung von Messzahlen $m_{85,t}$ aus den Messzahlen $m_{80,t}$ reicht es aus, alle Messzahlen $m_{80,t}$ durch $m_{80,85}$ zu dividieren, also bei PV alle Messzahlen durch 1,235 und bei SV durch 1,2279.
Für die mit 100 multiplizierten Messzahlen m^* gilt entsprechend die Formel: $m^*_{85,t} = (m^*_{80,t}/m^*_{80,85})\cdot 100$ für alle Jahre t = 80,81,...,88.

Abb. 9.2: Messzahlen und Ursprungsdaten von Privatem Verbrauch und Staatsverbrauch 1960 bis 1988

Privater Verbrauch: durchgezogene Linie, Staatsverbrauch: gestrichelte Linie
a) Ursprungsdaten b) Messzahlen 1980=100 c) Messzahlen 1960=100



Umbasierung und Verkettung:

Umbasierung (Basiswechsel) ist die Umkehrung der Verkettung. Mit den Perioden 0,s und t (etwa 1980, 1985 und 1990) bedeutet

Umbasierung: die bisherige Messzahl m_{0t} ist auf die neue Basis s umzustellen (um sie z.B. mit anderen Messzahlen der Basis s vergleichen zu können). Es ist also die Messzahl m_{0t} zu bestimmen.
Verkettung: zwei Messzahlenreihen zur Basis 0 und s sind zu einer langen Reihe zusammenzufügen (die Reihe mit der Basis 0 ist mindestens bis s geführt worden).

Lösung:

a) Messzahlen m_{0t},m_{st}:

Umbasierung:	$m_{st} = \frac{m_{0t}}{m_{0s}}$
Verkettung:	$m_{0t} = m_{0s} \cdot m_{st}$

(Bemerkung: dahinter steht der "Dreisatz" $m_{0t}/m_{0s} = m_{st}/m_{ss}$ wegen $m_{ss} = 1$)

b) Messzahlen m^*_{0t}, m^*_{st} (mit 100 multiplizierte Messzahlen):

Umbasierung:	$m^*_{st} = \frac{m_{0t}}{m_{0s}} \; 100$
Verkettung:	$m^*_{0t} = \frac{m_{0s} \cdot m_{st}}{100}$

4. Wachstumsraten und Wachstumsfaktoren

a) Wachstumsraten und Wachstumsfaktoren bei diskreter Zeit

Gerade in der Ökonomie ist es sehr verbreitet mit Wachstumsraten (relativen Zuwächsen) zu argumentieren. Die Rechenregeln für den Umgang mit Wachstumsraten (z.B. Berechnung einer mittleren Wachstumsrate, der Schluß von der Wachstumsrate des Preisniveaus P auf die Wachstumsrate der reziproken Kaufkraft K [K = 1/P], vgl. Beispiel 9.12) sind aber häufig nicht bekannt, so dass es angebracht erscheint, hierauf kurz einzugehen.

Jede Wachstumsrate bezieht sich auf ein Intervall bestimmter Länge. Es ist unmittelbar einsichtig, dass die Verzinsung eines Kapitals mit 5% per anum (also jährlich) einen geringeren Zinsertrag bedeutet als eine monatliche Verzinsung von 5%. Mit einer Grenzbetrachtung bei infinitesimal kleinen (kurzen) Intervallen wird die Zeit zu einer stetigen Variable und die Formeln haben dann z.T. eine andere Gestalt als bei diskreter Zeitvariable.

Def. 9.3: Wachstumsrate und Wachstumsfaktor bei diskreter Zeit t

a) Mit der diskreten Zeitvariable $t = 0,1,2,\dots,T$ erhält man für die Wachstumsrate und den Wachstumsfaktor (auch Gliedziffer oder Kettenindex genannt) der Zeitreihe y_t (d.h. der Zahlenfolge $y_0, y_1, \dots, y_t, \dots, y_T$) die folgenden Ausdrücke:

$$(9.7) \quad r_t = \frac{y_t - y_{t-1}}{y_{t-1}} = w_t - 1 \qquad (r_t\text{: Wachstumsrate}),$$

$$(9.8) \quad w_t = \frac{y_t}{y_{t-1}} = r_t + 1 \qquad (w_t\text{: Wachstumsfaktor}).$$

b) Für ein Wachstum mit konstanter Wachstumsrate [z.B. Verzinsung mit Zinseszins] r ($r_t = r$ für alle $t = 0,1,\dots,T$) gilt:

$$(9.9) \quad y_t = y_o \cdot w^t = y_o \cdot (1+r)^t \qquad (\text{Wachstum mit konstanter Rate r}).$$

Bei variierenden Wachstumsraten r_t lautet die Wachstumsgleichung:

$$(9.10) \quad y_T = y_0(1+r_1)(1+r_2)\dots(1+r_T) = y_0 \prod_{t=1}^{T}(1+r_t) = y_0 \prod_{t=1}^{T} w_t \ .$$

c) Als mittlere Wachstumsrate r soll diejenige konstante Wachstumsrate bezeichnet werden, die über den gleichen Zeitraum von 0 bis T zum gleichen Wachstum von y_0 zu y_T geführt hätte, wie die tatsächlichen (unterschiedlichen) Wachstumsraten $r_1, r_2, \dots, r_T$. Daraus folgt, dass r aus dem geometrischen Mittel der Wachstumsfaktoren w_t zu berechnen ist

$$(9.11) \quad r = (w_1 w_2 \dots w_T)^{1/T} - 1 = \left[\prod_{t=1}^{T} w_t\right]^{1/T} - 1 \ .$$

Bemerkungen zu Def. 9.3:

1. Drückt man die Wachstumsrate in Prozent aus, so ist der Wert r_t mit 100 zu multiplizieren (diese prozentuale Wachstumsrate sei p_t genannt: $p_t = 100 \cdot r_t$). Für das Wachstum mit konstanter Rate erhält man dann die vielen Kaufleuten bekannte Formel:
   $y_0(1 + p/100)^n$ für das Kapital y_n nach $t = n$ Perioden (Jahren).
2. Gl. 9.9 ist die Lösung der Differenzengleichung
   $y_t = w \cdot y_{t-1}$ $(t=1,2,...,T)$ mit dem Anfangswert y_0.
3. In der Ökonomie ist die ***halblogarithmische*** graphische ***Darstellung*** von Daten sehr beliebt (Ordinate: $\log(y_t)$ statt y_t und Abszisse t). Aus Gl. 9.9 folgt:

   $$\boxed{(9.9a)\quad \log(y_t) = \log(y_o) + w \cdot t,}$$

   d.h. bei Wachstum mit konstanter Rate ist die Zeitreihe y_t in halblogarithmischer Darstellung eine Gerade. Zwei Zeitreihen x_t und y_t, die in bestimmten Intervallen (etwa $a \leq t \leq b$) in halblogarithmischer Darstellung parallel verlaufen, haben in diesem Intervall gleiche Wachstumsraten.
4. Der Wachstumsfaktor w_t ist der Faktor, mit dem y_{t-1} zu multiplizieren ist, um y_t zu erhalten. Man kann w_t als Messzahl mit variabler Basis betrachten und w_t ist unabhängig von der Wahl einer Basisperiode; d.h. man kann Wachstumsfaktoren (und damit auch Wachstumsraten) auch aus entsprechenden Messzahlen mit beliebiger aber gleicher Basis errechnen, da

   $$w_t = \frac{m_{0t}}{m_{0,t-1}} = \frac{m_{st}}{m_{s,t-1}}.$$

5. Während die Wachstumsrate r_t positive und negative Werte annehmen kann, gilt i.d.R. für den Wachstumsfaktor $w_t > 0$ (denn $w_t < 0$ wäre eine Abnahme um mehr als 100% und y wird i.d.R. als nichtnegativ angenommen). Aus Def. 9.3 folgt, dass für r_t und w_t stets gilt:

   $r_t = w_t - 1$ und $w_t = r_t + 1$. Hierzu folgende Beispiele:

   $r_t = +0{,}2$ (20% Zunahme) $w_t = 1{,}2$
   $r_t = +0{,}06$ (6% Zunahme) $w_t = 1{,}06$
   $r_t = -0{,}05$ (5% Abnahme) $w_t = 0{,}95$
   $r_t = -0{,}32$ (32% Abnahme) $w_t = 0{,}68$.

6. Man beachte, dass die Zeit diskret ist, d.h. es wird mit "Einheitsintervallen" (z.B. ein Monat, ein Quartal, ein Jahr usw.) gerechnet. Wird dies nicht beachtet, so entstehen typische Fehlschlüsse (weil trotz diskreter Zeit so gerechnet wird, wie es nur mit stetiger Zeit zulässig ist), auf die in den Beispielen 9.10 bis 9.12 hingewiesen wird.
7. Auf dem gleichen Fehlschluß beruht die Berechnung eines **arithmetischen** Mittels wenn es gilt, eine mittlere Wachstumsrate r zu berechnen. Damit wird r meist überschätzt. Wie Teil c von Def. 9.3 zeigt, ist vom **geometrischen** Mittel der Wachstums**faktoren** auszugehen (vgl. auch Bsp. 9.11).
8. Die Wachstumsraten-Transformation ist eine **nichtlineare** Transformation der Zeitreihe y_t (während die Bildung von Messzahlen eine lineare Transformation darstellt), d.h. dass die Gestalt der Zeitreihe r_t von derjenigen der Zeitreihe y_t durchaus verschieden sein kann. Das gilt v.a. für die Lage von Maxima und Minima, die in der Ökonomie (z.B. bei der Konjunkturdiagnose) auch "Wendepunkte" genannt werden (vgl. Bsp. 9.15).

Beispiel 9.9:

(Ein Beispiel für Wachstum mit konstanter Wachstumsrate bei diskreter Zeit: Ausbreitung des Vampirismus nach dem Wiedererscheinen von Graf Dracula. Hinweis für Leser, die mit dem Vampirismus nicht vertraut sind: wenn ein Vampir V einen Nichtvampir beißt und ihm das Blut aussaugt, dann wird dieser ebenfalls zum Vampir, aber V stirbt durch den Biss nicht, sondern er ist im Gegenteil darauf angewiesen weiteren Menschen das Blut auszusaugen.)

Der häufig von skurrilen Vorstellungen geplagte Statistiker L wird nach dem Besuch einer einschlägigen Filmvorführung den Alptraum nicht los, dass Graf Dracula von den Toten auferstehen könnte. Er geht davon aus, dass der "Durchschnittsvampir" pro Monat zwei Menschen das Blut aussaugt. Wie lange wird es dauern, bis nach Draculas Wiedererscheinen eine Bevölkerung vom Umfang

- einer Großstadt mit 700.000 Menschen
- der Bundesrepublik Deutschland (70 Millionen Menschen)

vollständig vom Vampirismus befallen sein wird?

Lösung 9.9:

Die Anzahl y_t der Vampire entwickelt sich wie folgt: $y_0 = 1 = 3^0$ (Start mit Dracula) $y_1 = 3 = 3^1$, $y_2 = 3^2 = 9$ usw. Somit gilt $w = 3$ (Wachstumsrate r =

2 also 200% jeden Monat) sowie $y_0 = 1$ (Dracula). Nun ist t zu bestimmen aus der Gleichung

- $700.000 = 3^t$ ($t = \log(700000)/\log(3) = 12{,}25$) bzw.
- $70.000.000 = 3^{t^*}$ ($t^* = \log(70000000)/\log(3) = 16{,}44$),

d.h. die Großstadt ist nach einem Jahr (12 Monate) und einer Woche (0,25 Monate) und die Bundesrepublik nach 16,44 Monaten vollständig vom Vampirismus befallen.

Beispiel 9.10:

Einem verbreiteten Missverständnis zufolge entspricht einer monatlichen Wachstumsrate von 0,5% einer jährlichen Wachstumsrate von $12 \cdot 0{,}5 = 6\%$. Wie groß ist die tatsächliche jährliche Wachstumsrate bei einer monatlichen Wachstumsrate von

- 0,5%
- 5%
- 15%?

Lösung 9.10:

Die Rechnung: jährliche Wachstumsrate = 12·monatliche Wachstumsrate, die der gleichen Logik folgt, wie die Berechnung eines arithmetischen Mittels (Bsp. 9.11), ist nur bei kleinen Wachstumsraten annähernd richtig; je größer die monatliche Wachstumsrate, desto größer ist der Unterschied zur richtig berechneten jährlichen Wachstumsrate:

0,5%: $(1{,}005)^{12} - 1 = 0{,}06168$ also 6,17% statt 6%
5%: $(1{,}05)^{12} - 1 = 0{,}79586$ also 79,59% statt 60%
15%: $(1{,}15)^{12} - 1 = 4{,}35025$ also 435% statt 180% (=12·15%).

Folgerung aus Beispiel 9.10:

Der Grund für die Diskrepanz zwischen den beiden Rechenergebnissen ist in folgender Ungleichung zu finden:

$$\boxed{(9.12)\quad w^t - 1 \;^3\; (w-1)t = rt.}$$

Die Differenz zwischen w^t-1 und rt ist leicht zu bestimmen, wenn man berücksichtigt, dass w^t als binomiale Entwicklung

$$w^t = (1+r)^t = \sum_{i=0}^{t} \binom{t}{i} r^i = 1 + rt + \left[\binom{t}{2} r^2 + \binom{t}{3} r^3 + \ldots + tr^{t-1} + r^t\right]$$

darzustellen ist. Sie ist mithin der Ausdruck in der eckigen Klammer.

Beispiel 9.10 zeigt, dass die Exponentialfunktion w^t von der linearen Entwicklung rt + 1 umso mehr abweicht, je größer w und damit auch r^2, r^3,... ist. Die Differenz $D(t) = (1+r)^t - rt - 1$ wächst auch mit t, denn es gilt:
$D(0) = D(1) = 0$, $D(2) = r^2$, $D(3) = r^3 + 3r^2$, $D(4) = r^4 + 4r^3 + 6r^2$ usw.

Beispiel 9.11:

a) Man zeige: der Mittelwert der Wachstumsraten
   - 3, 5, 4 und 8% ist nicht 5% sondern nur 4,9835%.
   - 30, 50, 40 und 80% ist nicht 50% sondern nur 48,8877%.

b) Gegeben seien die Daten 200, 260, 273, 284 und 307. Wie groß ist die mittlere Wachstumsrate?

c) Wie unterscheiden sich die korrekt gerechnete mittlere Wachstumsrate aus zwei Wachstumsraten r_1, r_2 von dem fälschlich angewendeten arithmetischen Mittel von Wachstumsraten?

Lösung 9.11:

a) Die mittlere Wachstumsrate r ist nach Def. 9.3 nicht das arithmetische Mittel 5% (vgl. auch Bsp. 4.4), sondern über das geometrische Mittel der Wachstumsfaktoren zu berechnen denn wegen:

$y_4 = y_0 \cdot 1{,}03 \cdot 1{,}05 \cdot 1{,}04 \cdot 1{,}08 = y_0 \cdot (1+r)(1+r)(1+r)(1+r)$ ist
$r = (1{,}03 \cdot 1{,}05 \cdot 1{,}04 \cdot 1{,}08)^{1/4} - 1 = 0{,}049835$ also 4,98%

Man kann als Ergebnis festhalten:

Die mittlere Wachstumsrate ist nach Def. 9.3 aus dem **geometrischen** Mittel der Wachstums**faktoren** zu bestimmen, nicht aber als **arithmetisches** Mittel der Wachstums**raten**.

Entsprechend ist zu rechnen
$r = (1{,}3 \cdot 1{,}5 \cdot 1{,}4 \cdot 1{,}8)^{1/4} - 1 = 4{,}914^{1/4} - 1 = 0{,}4888$ also 48,89% und nicht $(30+50+40+80)/4 = 200/4 = 50\%$.

b) Es ist nicht nötig, die einzelnen Wachstumsraten auszurechnen. Für die Wachstumsfaktoren erhält man 260/200, 273/260, 284/273 und 307/284. Bildet man das Produkt dieser vier Brüche, so sieht man, dass sich einiges wegkürzt, so dass man einfach erhält:
$r = (307/200)^{1/4} - 1 = 0{,}113$ (also 11,3%)
oder allgemein: $(y_t / y_0)^{1/t} - 1$ (wenn y_0 der erste Wert und y_t der letzte Wert ist). Man beachte, dass der Bruch y_t/y_0 die Messzahl m_{0t} darstellt, so dass man für die mittlere Wachstumsrate r erhält

$$(9.11) \quad r = (y_t / y_0)^{1/t} - 1 = \sqrt[t]{m_{0t}} - 1 \qquad \text{(mittlere Wachstumsrate)},$$

bzw. in Prozent:

$$(9.11a) \quad r = [(y_t/y_0)^{1/t} - 1]100$$

c) Die korrekt aufgrund eines geometrischen Mittels der Wachstumsfaktoren berechnete mittlere Wachstumsrate aus den beiden Wachstumsraten r_1, r_2 soll $r = r_G$ und das arithmetische Mittel soll r_A genannt werden ($r_A = ½(r_1+r_2)$). Dann ist offensichtlich stets $r_A > r_G$, es sei denn $r_1 = r_2$ denn es gilt:

$$\boxed{(1 + r_A)^2 - ¼(r_1 - r_2)^2 = (1 + r_G)^2.}$$

Es kommt mithin auf die Unterschiedlichkeit von r_1 und r_2 an. Der Ausdruck $¼(r_1 - r_2)^2$ ist übrigens die Varianz der Wachstumsraten r_1 und r_2 also (s_r^2). Insbesondere bedeutet auch $r_A = 0$, nicht $r_G = 0$ denn es gilt: wenn $r_A = 0$ dann $r_1 = - r_2$ und $(1+r_G)^2 = 1 - ¼(2r_1)^2 = 1-r_1{}^2$, so erhält man z.B. bei zunehmender Varianz s_r^2:
$r_1 = 0{,}1$ und $r_A = 0$ dann $r_G = -0{,}0050126$
$r_1 = 0{,}2$ und $r_A = 0$ dann $r_G = -0{,}0202041$ oder
$r_1 = 0{,}3$ und $r_A = 0$ dann $r_G = -0{,}0460608$.

Beispiel 9.12:

Einer Zunahme des Preisniveaus von 10% entspricht eine Abnahme der Kaufkraft in Höhe von (Richtiges ankreuzen)

- 10%
- mehr als 10%
- weniger als 10%

Lösung 9.12:

Viele sind geneigt, 10% anzukreuzen. Dass dies falsch ist, erkennt man daran, dass bei dieser Logik eine Zunahme des Preisniveaus um 100% (also eine Verdoppelung der Preise, was als Inflationsrate durchaus vorkommen kann) ein Kaufkraftverlust von 100% bedeuten würde, was natürlich nicht möglich ist. Eine Verdoppelung der Preise bewirkt vielmehr eine Halbierung der Kaufkraft. Der Kaufkraftverlust beträgt also "nur" 50% und nicht 100%. Da die Kaufkraft (K) das reziproke Preisniveau (P) ist, d.h. es ist $K_t = 1/P_t$ (für jede Periode t), gilt für die Wachstumsfaktoren (bzw. Wachstumsraten) von P und K, w_P und w_K (bzw. r_P und r_K) bei diskreter Zeit (vgl. auch Übers. 9.5):

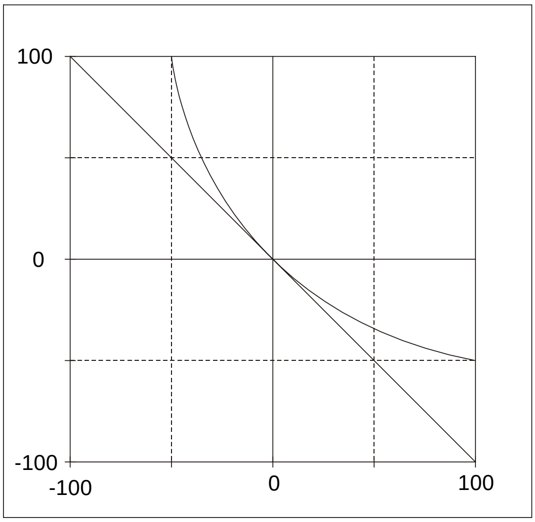
$$(9.13) \quad w_K = \frac{1}{w_P} \quad \text{und damit} \quad r_K = \frac{1}{1 + r_P} - 1 \qquad \text{wenn} \quad K = \frac{1}{P}$$

bzw. wenn die Wachstumsraten in Prozent ausgedrückt sind:

(9.13a) $r_K^* = 10000/(100+r_P^*) -100$.

Der Zusammenhang zwischen den Wachstumsraten von P und K ist also hyperbolisch wie die nebenstehende Abb. 9.3 zeigt. Beispiele für Gl. 9.13a: r_P^* = 25% dann ist r_K^* = -20% oder r_P^* = 33.3% (also 1/3) r_K^* = -25% (also 1/4) usw. Auch hier gilt: der Fehler, der dadurch entsteht, dass man bei diskreter Zeit so rechnet, wie mit stetiger Zeit (also $r_K^* = -r_P^*$, die eingezeichnete Gerade in Abb. 9.3) ist umso geringer, je kleiner die Wachstumsrate ist.

Abb. 9.3: Kaufkraft und Preisniveau



b) Wachstumsraten und Wachstumsfaktoren bei stetiger Zeit

Bei diskreter Zeit wurde die Zeitreihe und deren Wachstumsrate mit y_t und r_t (mit t=0,1,...,T als Periodenindex) bezeichnet. Wird die Zeit dagegen als stetige Variable ($t \in \mathbb{R}$) betrachtet, so ist eine Zeitreihen- und eine Wachstumsratenfunktion gegeben, die zur besseren Unterscheidbarkeit mit y(t) und r(t) bezeichnet werden sollen. Im Falle einer konstanten Wachstumsrate soll diese bei stetiger Zeit mit a bezeichnet werden (anstelle von r bei diskreter Zeit).

Der Übergang von $r_1,...,r_T$ der diskreten Folge $y_0,y_1,...,y_T$ gem. Def. 9.3 zur stetigen Wachstumsratenfunktion r(t) für die Zeitreihe y(t) erfolgt durch Verkürzung der Intervalle, auf die sich die Wachstumsraten beziehen. Mit

$$r(t,\Delta) = \frac{(y_t - y_{t-\Delta})/\Delta}{y_{t-\Delta}}$$

(man beachte jedoch, dass in diesem Fall t nicht mehr als diskret, sondern als stetig angenommen werden muss, weil D beliebig klein sein darf) ist die Wachstumsrate bezüglich des Intervalls der Länge D gegeben. Mit den Einheitsintervallen ($\Delta = 1$) erhält man Gl. 9.7 und der Grenzübergang

$$\lim_{\Delta \to 0} (r(t,\Delta)) = r(t)$$

liefert für die Funktion y(t) an jeder beliebigen Stelle t, etwa bei $t = t_o$ die Wachstumsrate r(t) von y(t). Sie ist nach Def. 9.4

$$r(t=t_o) = \left.\frac{y'(t)}{y(t)}\right|_{t_o} = \left.\frac{dy/dt}{y}\right|_{t_o}$$

Man kann y'(t) auch als absolutes - und r(t) = y'(t)/y(t) als relatives Wachstum bezeichnen.

Def. 9.4: Wachstumsrate bei stetiger Zeit

a) Die Wachstumsrate r(t) einer stetigen Funktion y = y(t) ist

$$\boxed{(9.14)\quad r(t) = \frac{y'(t)}{y(t)} = \frac{dy/dt}{y} = \frac{d\ln(y)}{dt}}$$

b) Bei konstanter Wachstumsrate r(t) = a (für jeden Wert von t) ist die stetige Zeitreihe y(t) gegeben mit

$$\boxed{(9.15)\quad y(t) = y(0)\cdot e^{at} = y(0)\cdot \exp(at).}$$

c) Bei variablen Wachstumsraten r(t) gilt entsprechend

$$\boxed{(9.16)\quad y(T) = y(0) \cdot \exp\left(\int_0^T r(t)dt\right)}$$

wenn r(t) eine im Intervall (0,T) stetige Funktion ist, so dass man für eine mittlere Wachstumsrate in Analogie zu Def. 9.3

$$\boxed{(9.17)\quad a = \frac{1}{T}\int_0^T r(t)\,dt = \frac{\ln[y(T)/y(0)]}{T} = \frac{1}{T}\{\ln[y(T)] - \ln[y(0)]\}}$$

erhält.

Bemerkungen zu Def. 9.4:

1. Im stetigen Fall hat der Begriff "Wachstumsfaktor" wenig Sinn (im Unterschied zur Wachstumsrate). Nach Gl. 9.14 erhält man die stetige Wachstumsrate indem man die erste Ableitung der Funktion y(t) bildet (y'(t) = dy/dt) und diese durch die Funktion y(t) teilt. Die Funktion r(t) ist die Wachstumsrate von y(t), d.h. für jeden Wert von t, etwa für $t=t_o$ stellt $r(t=t_o)$ die Wachstumsrate von $y(t=t_o)$ dar. Wegen

$$\frac{d\ln(y)}{dt} = \frac{1}{y}\frac{dy}{dt}$$

wird Gl. 9.14 auch logarithmische Ableitung von y(t) genannt.

2. Gl. 9.15 ist die Lösung der Differentialgleichung dy/dt = ay mit dem Anfangswert y(0). Nach Gl. 9.15 ist ferner ln[y(t)] = ln[y(0)] + at und die logarithmische Ableitung hiervon ist dln[y(t)]/dt = a. Der Übergang von Gl. 9.9 zu Gl. 9.15 ist auch leicht einsichtig zu machen, wenn man den sog. "Aufzinsungsfaktor" bei jährlicher, halbjährlicher, monatlicher Verzinsung betrachtet, denn es gilt dann:

$$(1+r)^t,\ (1+\frac{r}{2})^{2t},\ (1+\frac{r}{12})^{12t} \text{ usw., mit } \lim_{n\to\infty}(1+\frac{r}{n})^{nt} = e^{rt}\ .$$

3. Aus Gl. 9.15 und 9.9 folgt, dass zwischen der Wachstumsrate α (stetige Zeit) und r (diskrete Zeit) die folgende Beziehung besteht:

$$\boxed{(9.18)\quad e^{\alpha} = w = 1+r, \quad \text{so dass gilt} \quad (9.19)\quad a = \ln(1+r).}$$

Man erhält somit im Zusammenhang mit der Reihenentwicklung von e^a und ln(1+r) die folgenden Umrechnungen

$$\boxed{(9.20)\quad r = e^{\alpha} - 1 = a + \frac{a^2}{2!} + \frac{a^3}{3!} + \frac{a^4}{4!} + ...}$$

für die Umrechnung von a in r (so dass a < r) und

$$\boxed{(9.21)\quad \alpha = \ln(1+r) = r - \frac{r^2}{2!} + \frac{r^3}{3!} + \frac{r^4}{4!} + ...}$$

für die Umrechnung von r nach α.
Wie man sieht, gilt nur bei kleinen Wachstumsraten $r \approx \alpha$.

<u>Beispiel 9.13:</u>

Gegeben sei die Funktion $y(t) = \exp(a + bt + ct^2)$, die auch logarithmische Parabel genannt wird, da $\ln[y(t)] = a + bt + ct^2$.

a) Wie lautet die Funktion r(t) und welche Besonderheiten hat sie?

b) Wie lautet für die Funktion $\ln[y(t)] = 0{,}1 + 0{,}5t - 0{,}04t^2$ die
- Wachstumsrate r(3) (also r(t) für t=3)
- mittlere Wachstumsrate im Intervall $0 \leq t \leq 6$?

Lösung 9.13:

a) Man sieht an diesem Beispiel leicht, dass gilt

$$\frac{dy/dt}{y} = \frac{d\ln(y)}{dt}$$

denn die Ableitung dy/dt beträgt nach der Kettenregel mit $z = a + bt + ct^2$ $dy/dt = dy/dz \cdot dz/dt = e^z \cdot (b+2ct)$, so dass die Wachstumsrate $r(t) = b+2ct$ ist, also linear zu- oder abnimmt. Zum gleichen Ergebnis gelangt man durch die Ableitung von lny nach t.

b) Die Funktion y(t) ist in Abb. 9.4 (links oben y(t) und links unten r(t)) dargestellt. Sie steigt von $y(0) = e^{0,1} = 1,10517$ auf $y(6,25) = 5,2725$ und fällt dann auf $y(10) = e^{-2,9} = 0,05502$. Die Funktion $r(t) = 0,5 - 0,08t$ fällt linear von $r(0) = 0,5$ zu $r(6) = 0,02$. Sie nimmt bei $t = 3$ den Wert $r(3) = 0,26$ an und den Wert Null an der Stelle $t = 6,25$. An dieser Stelle durchläuft y(t) ein Maximum mit $y(t=6,25) = \exp(1,6625) = 5,2725$. Das Integral über r(t) in den Grenzen 0 bis 6 beträgt 1,56. Man beachte, dass gilt $y(0) = e^{0,1}$ und $y(6) = e^{1,66} = 5,2593$ und $y(0) \cdot e^{1,56} = y(6)$, so dass man nach Gl. 9.17 für die mittlere Wachstumsrate erhält

$$\begin{aligned}\alpha &= \ln[y(T)/y(0)]/T = T^{-1}\{\ln[y(T)] - \ln[y(0)]\}\\ &= \{\ln[e^{1,66}] - \ln[e^{0,1}]\}/6 = (1,66-0,1)/6 = 1,56/6 = 0,26.\end{aligned}$$

Die mittlere stetige Wachstumsrate beträgt also 26% (dem entspräche eine mittlere Wachstumsrate im diskreten Fall von 29,69%, denn $(e^{0,26})^6 = (1,2969)^6 = 4,7588 = 5,2593/1,10517$. Die Größe 4,7588 ist genau das Verhältnis $y(6)/y(0) = e^{1,66}/e^{0,1} = e^{1,56}$.

Beispiel 9.14:

In einem stetigen Wachstumsmodell werden konstante Wachstumsraten in Höhe von 0,5%, 5% und 50% unterstellt. Wie groß sind die entsprechenden diskreten Wachstumsraten?

Lösung 9.14:

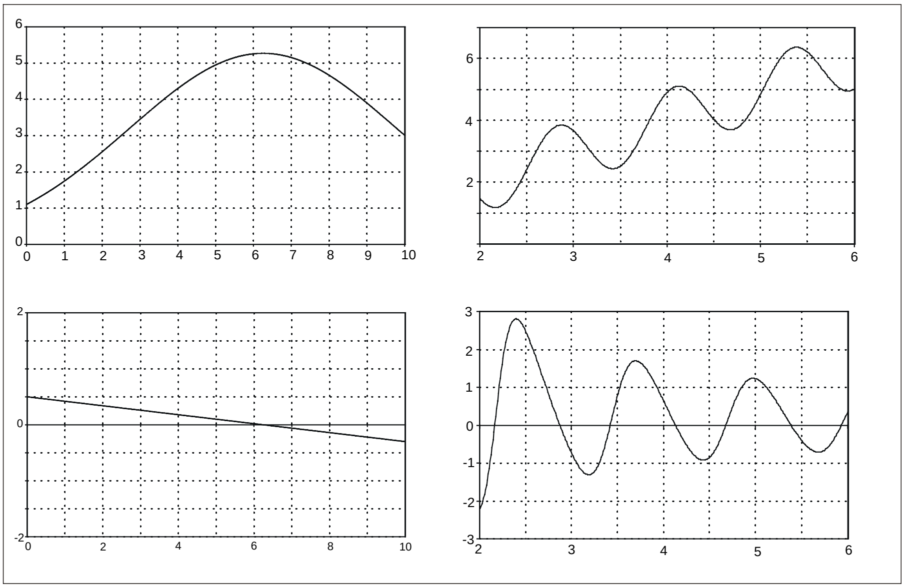
Wegen $r = e^a - 1$ gilt a = 0,005 entspricht r = 0,0050125 (der Unterschied zwischen r und a ist also gering), a = 0,05 entspricht r = 0,05127 und a = 0,5 entspricht r = 0,64872 (der Unterschied zwischen r und a ist beträchtlich).

Beispiel 9.15:

Man kann leicht zeigen, dass die Wachstumsratenfunktion andere Maxima und Minima besitzt, als die Zeitreihe y(t). Hierzu das Beispiel: y(t) = t + sin(5t). Wie lautet die Funktion der Wachstumsraten?

Abb. 9.4: Wachstumsraten für einige stetige Funktionen

Links: Logarithmische Parabel (Beispiel 9.13), Rechts: y(t) = t + sin(5t) (vgl. Beispiel 9.15), oberes Bild jeweils y(t) und unteres Bild r(t)



Lösung 9.15:

Die Ableitung von y(t) nach t lautet dy/dt = 1 + 5cos(5t), so dass r(t) = [1+5cos(5t)]/[t+sin(5t)]. Offensichtlich haben y(t) und r(t), wie Abb. 9.4 zeigt, unterschiedliche Extrema.

c) Weitere Bemerkungen zu Wachstumsraten

In diesem Exkurs soll noch einmal auf verbreitete Mißverständnisse im Umgang mit Wachstumsraten eingegangen werden und es werden einige Formeln zu bekannten Funktionen y(t) und deren Wachstumsraten r(t) zusammengestellt, die häufig als Modelle eines Wachstumsprozesses oder auch als Trendmodelle benutzt werden.

1. Die Wachstumsrate eines Produkts, Quotienten und Kehrwerts

Das Beispiel 9.12 hat gezeigt, dass oft falsche Vorstellungen über die Wachstumsrate einer Funktion von y(t) bestehen. Gerade in der Ökonomie werden oft Größen als Produkte oder Quotienten (oder z.B. als Kehrwert) definiert:

- So ist z.B. der Umsatz (z) das Produkt aus Menge (x) und Preis (y) und viele sind geneigt, z.B. wie folgt zu rechnen: eine Zunahme des Preises um 7% und eine Zunahme der Menge um 3% bedeuten eine Umsatzzunahme von 10%, was nur bei stetiger Zeit gilt (ist t diskret so ist die Umsatzzunahme 1,07·1,03 - 1 = 0,1021 also nicht 10% sondern 10,21%);
- Das Realeinkommen ist der Quotient Nominaleinkommen/Preisniveau und oft wird wie folgt gerechnet: eine Zunahme des Nominaleinkommens z.B. um 20% bedeutet bei einer Preissteigerung von 6% eine Zunahme des Realeinkommens um 20 - 6 = 14%. Bei diskreter Zeit beträgt jedoch die Zunahme des Realeinkommens nicht 14% sondern nur 13,2%, denn 1,2/1,06 - 1 = 0,13208.

In Übersicht 9.4 sind die relevanten Formeln zusammengestellt.

Übersicht 9.4: Wachstumsraten von Produkten, Quotienten und Kehrwerten

	diskrete Zeit	stetige Zeit
Produkt $z = xy$	$w_z = w_x w_y$	$r_z(t) = r_x(t) + r_y(t)$
Quotient $z = \frac{x}{y}$	$w_y = \frac{w_x}{w_y}$	$r_z(t) = r_x(t) - r_y(t)$
Kehrwert $z = \frac{1}{y}$	$w_z = \frac{1}{w_y}$	$r_z(t) = - r_y(t)$

2. Einige Wachstumsfunktionen insbesondere Kurven mit einem Sättigungsniveau und deren Wachstumsraten

In Übersicht 9.5 sind einige Funktionen y(t) mit stetiger Zeitvariable t zusammengestellt, die bei Ökonomen von gewissem Interesse sind. Dies gilt insbesondere für solche Kurven, die einer Sättigungsgrenze k entgegenstreben (Nr. 4 bis 9).

Übersicht 9.5.: Wachstumsraten ausgewählter Funktionen

	Funktion y(t)	Ableitung y = dy/dt	Wachstumsrate r(t) (1)
1	$(a+bt)^{\alpha}$ (2)	$ab(a+bt)^{\alpha-1}$	$ab(a+bt)^{-1}$
1a	$\alpha=1$: Gerade $y = a+bt$	b	$b/(a+bt) = r_G$
1b	$\alpha=-1$: $\frac{1}{a+bt}$	$-\frac{b}{(a+bt)^2}$	$-\frac{b}{a+bt} = -r_G$
1c	$\alpha=½$: $\sqrt{(a+bt)}$	$½b(a+bt)^{-1/2}$	$½r_G$
1d	Potenzfunktion bt^a	$b\alpha t^{\alpha-1}$	α/t (hyperbolisch)
2	Parabel $a+bt+ct^2$ (3)	$b+2ct$	$\frac{b+2ct}{a+bt+ct^2}$
3	$a\cdot\exp(bt^{\alpha})$	$y\cdot\alpha bt^{\alpha-1}$	$\alpha bt^{\alpha-1}$
3a	$\alpha=1$: ae^{bt} oder: ar^t mit $r = e^b$	yb $y\cdot\ln(r)$	b $b = \ln(r)$
3b	$\alpha=-1$: $ae^{b/t}$	$-yb/t^2$	$-b/t^2$
4	$k+be^{ct}$ oder $y = k+br^t$ mit $r=e^c$ (k: Sättigungsniveau)	cbe^{ct} strebt gegen 0, wenn r<1, c=lnr<0	$-c\,\frac{k-y}{y} = -cR$ speziell für: c = -1, ist r(t) = R; Abb. 9.5
5	$k+\frac{b}{c+t}$ (Hyperbel)	$-\frac{b}{(c+t)^2}$	$-\frac{b}{k(c+t)^2+b(c+t)}$
6	$\frac{k(t+a)}{t+b}$ (b>a)	$\frac{k(b-a)}{(t+b)^2}$	$\frac{b-a}{(t+a)(t+b)}$
7	$\exp(K+br^t)$ mit $r=e^c$ oder: $\ln(y) = K+be^{ct}$; $k=e^K$ Sättigungsniveau (b < 0)	$ybcr^t$ strebt gegen 0, wenn r<1 (c<0)	$br^t\ln(r) = bce^{ct}$ [mit ln(r) = c < 0] (vgl. Abb. 9.5)
8	$\frac{k}{1+e^{a-bt}}$ a,b,k>0 k Sättigungsniveau	$\frac{by(k-y)}{k}$	$b+ßy$ ($ß = -\frac{b}{k}$) (4) (vgl. Abb. 9.5)
9	$\ln(y) = K - \frac{a}{b+t}$ $k=e^K$ Sättigungsniveau	$\frac{ya}{(b+t)^2}$	$\frac{a}{(b+t)^2}$

(1) $r(t) = y'/y$

(2) In dieser allgemeinen Form liegt eine Polynomfunktion vom Grade a vor, wobei a Î Z, Z^+ ist. Bei a=½ liegt eine Wurzelfunktion und bei ganzzahligem a und a>0 eine Potenzfunktion vor (Fall 1d).

(3) Kann entsprechend verallgemeinert werden wie Funktion Nr.1.

(4) Kennzeichnend für die logistische Funktion: r(t) = f[y(t)] (f: linear).

Man beachte:

1. cy(t) und y(t) haben die gleiche Wachstumsrate r(t) [c: Konstante]
2. hat y(t) die Wachstumsrate r(t), so hat $[y(t)]^{-1}$ die Wachstumsrate -r(t)

Weitere Funktionen in den Beispielen 9.13 (logarithmische Parabel) und 9.15.

Sättigungsniveau:

Die Funktionen Nr. 4 ff haben Sättigungsniveaus (k), denen sich y(t) asymptotisch nähert und zwar [bei bestimmter Konstellation der Parameter] monoton (Nr. 4 bis 6) oder S-förmig (d.h. mit Wendepunkt, Nr. 7 - 9), weil die Ableitung dy/dt mit wachsendem t gegen 0 strebt.

Namen von Funktionen:

4: modifizierte Exponentialfunktion, 6: Törnquist-Funktion (für a=0 oder b=0 ergeben sich einfach linearisierbare Funktionen [Übers. 8.2], 7: Gompertz - Funktion, 8: logistische Funktion, 9: Johnson-Funktion (eine "logarithmische Hyperbel").

1. Die modifizierte Exponentialfunktion (Nr.4 in Übers. 9.5)

In Abb. 9.5 ist links oben die folgende Funktion dargestellt:
$y(t) = k + br^t$ mit k = 20, b= -16 und r = 0,95 (so dass c = ln(r) = -0,0513), also $y(t) = 20 - 16 \cdot 0{,}95^t$.

Die Funktion steigt von y(0) = k+b = 10+(-16) = 4 monoton zum Sättigungsniveau k=20 und zwar umso steiler, je kleiner r ist (für r=1 ist y(t) eine Parallele der Abszisse t). Sie

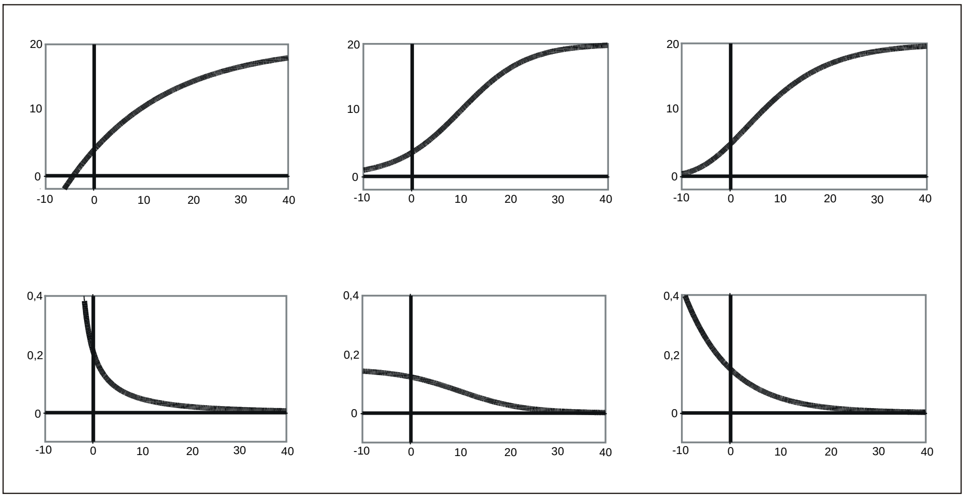
- nähert sich k monoton wenn r < 1 und zwar steigend wenn b < 0, fallend wenn b > 0
- entfernt sich monoton von k wenn r > 1, steigend wenn b > 0, fallend wenn b < 0.

Die Wachstumsrate (Abb. 9.5 links unten) ist $r(t) = (-\ln(r))[k-y(t)] / y(t) = 0{,}051294 \cdot 16 \cdot 0{,}95^t/[20+(-16 \cdot 0{,}95^t)]$und sie ist monoton fallend; $r(0) = 0{,}051294 \cdot (20-4)/4 = 0{,}20517$. Der Ausdruck

R = [k - y(t)]/y(t)

ist der relative Abstand der Funktion vom Sättigungsniveau k. Die modifizierte Exponentialfunktion ist also dadurch ausgezeichnet, dass die Wachstumsrate r dieser Größe R proportional ist (Proportionalitätsfaktor c). Die Logarithmen von R heißen **Logits**. Wie man leicht sieht erhält man mit k = 0 die einfache Exponentialfunktion. Dann ist R = -1 und die Wachstumsrate ist r(t) = (-c)R = c, was der Größe b in Zeile 3a der Übersicht 9.5 entspricht.

Abb. 9.5: Einige Funktionen mit einer Sättigungsgrenze und deren Wachstumsraten



2. Die logistische Funktion (Nr. 8 in Übersicht 9.5)

Sie hat, anders als die modifizierte Exponentialfunktion einen Wendepunkt. Als logistische Funktion bezeichnet man eine Funktion der folgenden Gestalt:

$$\boxed{y(t) = \frac{k}{1 + e^{f(t)}}}$$

wobei f(t) eine monoton fallende Funktion ist, z.B. die Gerade f(t) = a - bt (mit a,b,k > 0). In Abb. 9.5 (Mitte oben) wurde die folgende Funktion dargestellt

$$y(t) = \frac{20}{1 + \exp(1{,}5-0{,}15t)}$$

Es gilt: $y(0) = k/(1+e^a) = 20/(1+e^{1,5}) = 3{,}6485$.
Der Wendepunkt liegt bei t = a/b = 10 und dann ist y(t) = k/2 = 10. Die Kurve steigt umso steiler, je größer bei gegebenen k und a der Parameter b ist. Sie ist zentralsymmetrisch um den Wendepunkt.
Die logistische Funktion ergibt sich aus der Differentialgleichung dy/dt = by(k-y)/k.
Die Wachstumsrate (Abb. 9.5 Mitte unten) der logistischen Funktion ist linear abhängig vom erreichten Niveau y was die Schätzung der logistischen Funktion erleichtert:
r(t) = b - by/k = 0,15 - 0,0075y(t).
Sie fällt ab $r(0) = b - b/(1+e^a) = 0{,}15 - 0{,}15/(1+e^{1,5}) = 0{,}12264$ monoton. Die Logits L(t) haben einen linearen Trend: L(t) = ln(R) = ln{[k-y(t)]/y(t)} = a - bt.
Man erkennt an der Gestalt der Logit-Funktion übrigens auch die Symmetrie. So gilt z.B.
für y = k/2: L = ln(1) = 0, so dass t = a/b = 10
für y = k/4: L = ln(3) , so dass t = (a-ln3)/b = 2,676 und

für y = k/2: L = ln(1/3), so dass t = (a-ln3)/b = 17,324.

3. Die Gompertzfunktion (Nr. 7 in Übers. 9.5)

Auch sie hat einen Wendepunkt ist aber weniger bekannt als die logistische Funktion. In Abb. 9.5 (rechts oben) wurde dargestellt:

$y(t) = k \cdot \exp(b \cdot r^t) = 20 \cdot \exp(-1{,}4 \cdot 0{,}9^t)$.
k = 20, b = -1,4 und r = 0,9 so dass c = ln(r) = -0,10536.

Mit kleinerem Wert von r (etwa r=0,8) verläuft die Kurve steiler. Ist r > 1, so ist die Kurve fallend. Es gilt $y(0)= ke^{-b} = 20e^{-1,4} = 4{,}932$ (je größer betragsmäßig b ist, desto niedriger ist der Ordinatenabschnitt). Der Wendepunkt liegt bei $t_W = -\ln(|b|)/c = -\ln(1{,}4)/(-0{,}10536) = 3{,}1935$, wobei y stets $y(t_W) = k/e = 7{,}3576$ (mit k = 20) beträgt.
Die Wachstumsrate (Abb. 9.5 rechts unten) lautet $r(t) = b(\ln r)r^t = 0{,}147505 \cdot 0{,}9^t$, sie ist also ab r(0) = 0,1475 monoton fallend. Es gilt $\ln[k/y(t)] = br^t$, so dass c[ln(y) - ln(k)] = r(t), d.h. die Wachstumsrate ist proportional zur Differenz der Logarithmen von y und k (bzw. zu ln(y/k)).
Zusammenfassend kann festgestellt werden:

	Funktion	die Wachstumsrate ist
4	modifizierte Exponentialfunktion	proportional zu $R = \frac{k-y}{y}$ relativer Abstand zum Sättigunsniveau
7	Gompertzfunktion	proportional zu $\ln\left(\frac{y}{k}\right) = \ln\left(\frac{1}{1+R}\right)$
8	logistische Funktion	linear abhängig von y/k $r(t) = b - b\frac{y}{k}$ die Logits L(t) = ln(R) haben einen linearen Trend L(t) = a - bt

5. Aggregationsprobleme

In diesem Abschnitt werden einige Probleme behandelt, die nicht beschränkt sind auf Verhältniszahlen, aber andererseits auch nicht ein eigenes Kapitel rechtfertigen dürften. Um den Begriff der Aggregation deutlich zu machen, ist auch auf Daten zurückzugreifen, die als Häufigkeitsverteilung vorliegen.

a) Begriff der Aggregation

Im Rahmen der Deskriptiven Statistik wird unter Aggregation meist eine Summenbildung verstanden. Liegen Daten als Häufigkeitsverteilung vor, so wird der Begriff benutzt für unterschiedliche Fragestellungen:

a) Aussagen über eine Variable X für verschiedene Teilgesamtheiten und Zusammenhänge zwischen den Aussagen (z.B. den Mittelwerten), die für die Gesamtheit und jene, die für die Teilmassen gelten. Dabei können diese Teilgesamtheiten aufgrund aneinandergrenzender Intervalle auf der x-Achse (Klassenbildung) oder aufgrund eines anderen (meist qualitativen aber auch quantitativen[1]) Merkmals gebildet sein (Zerlegung, vgl. S. 24). In diesem Sinne interessieren auch Zusammenhänge zwischen einer aggregierten Beziehungszahl X/Y und den Teilbeziehungszahlen x_j/y_j, wenn bei j =1, 2, .., J Teilgesamtheiten gilt $X = \Sigma x_j$ und $Y = \Sigma y_j$.

b) Aussagen über eine Variable, die eine Summe darstellt (etwa Z = X + Y als ungewogene Linearkombination oder $Z = b_1 X_1 + b_2 X_2$ als [mit den Gewichten b_1 und b_2] gewogene Linearkombination).

Um den Unterschied deutlich zu machen, sollte man **im Fall a)** von **Aggregation von Verteilungen** (Summation über Einheiten bei einer Variablen) sprechen, und **im Fall b)** von einer **Linearkombination** (Summation über Variablen bei einer Gesamtheit). Im Fall a) wird über die Häufigkeiten n_i (auf der Ordinate) und im Fall b) über die Merkmalswerte (Abszisse) summiert.

Der Begriff Aggregation tritt auch in der Wirtschaftsstatistik auf, insbesondere in der Sozialproduktsrechnung. Gemeint ist damit die sog. "fundierte Schätzung" eines (inhomogenen) "Aggregats", die gerade nicht durch eine einfache Summenbildung möglich ist. Indexzahlen betrachten aggregierte Veränderungen. Die speziellen Probleme der Indizes (Kap. 10) im Unterschied zu einfachen Messzahlen (Def. 9.1 d) entstehen alle dadurch, dass Indizes Aggregate vergleichen.

Das Aggregationsproblem (im engeren Sinne) tritt auch in der Ökonometrie bei der Frage auf, ob und unter welchen Voraussetzungen eine funktionale Beziehung zwischen mikroökonomischen (auf das einzelne Wirtschaftssubjekt bezogene) Variablen, etwa X_v und Y_v (etwa eine Regressionsfunktion $\hat{y} = a + bx_v$) auch für aus diesen Größen durch

[1] Ein Beispiel für ein quantitatives Merkmal wäre: Zusammenhang zwischen der rohen und den altersspezifischen Todesraten.

Summation hervorgegangene Makrovariablen $X = \Sigma X_i$ und $Y = \Sigma Y_i$ gelten kann, und wie dann die entsprechenden Koeffizienten zu interpretieren sind.

b) Aggregation von Verteilungen

Bei der Aggregation von Verteilungen ist es im Unterschied zur Klassierung (Klassenbildung) nicht notwendig, dass die Teilgesamtheiten disjunkt sind. Die zulässigen Wertebereiche für die Variable X können sich auch überschneiden (vgl. Beispiel 9.16) oder gar identisch sein. Beispiele für das Auftreten dieses Problems sind die Schätzung einer "gemeinsamen" Einkommensverteilung auf der Grundlage der Lohn- und Einkommensteuerstatistik, die Verteilung der Arbeitsverdienste der Arbeitnehmer eines Unternehmens mit mehreren Betrieben oder die Schätzung von Bundesergebnissen auf der Basis der Bundesländer.

Über die Gestalt der so entstehenden "Mischverteilungen" lassen sich keine allgemeinen Aussagen machen. Durch Aggregation können z.B. mehrgipflige oder schiefe Verteilungen entstehen auch wenn alle zugrundeliegenden Verteilungen der Teilgesamtheiten unimodal und symmetrisch sind.
Allerdings gibt es für die Zusammenhänge zwischen den Verteilungsmaßzahlen der J Teilgesamtheiten (j = 1,2,...,J) und der entsprechenden Verteilungsmaßzahl der aggregierten Verteilung oft einfache Beziehungen (z.B. ein gewogenes Mittel).
Für Mittelwerte und Streuungsmaße sind sie in der Regel bei der Darstellung der entsprechenden Maßzahl erwähnt worden. So gilt z.B. für das arithmetische Mittel $\bar{x}$ und die Varianz s^2 der aggregierten Verteilung

(*) $\bar{x} = \Sigma g_j \bar{x}_j$ (Teilgesamtheit j = 1,2,...,J) und

(**) $s^2 = \Sigma(\bar{x}_j - \bar{x})^2 g_j + \Sigma s_{j'}^2 g_j$

mit $g_j = n_j / n$, so dass $0 \leq g_j \leq 1$ und $\Sigma g_j = 1$.

Man beachte, dass die Formeln mit denen übereinstimmen, die für die Klassierung gelten.

<u>Beispiel 9.16:</u>

Gegeben seien die Häufigkeitsverteilungen für das Merkmal X für die beiden Teilgesamtheiten A und B (mit den absoluten Häufigkeiten n_{Ai} und n_{Bi}).

x_i	n_{Ai}	n_{Bi}
2	10	-
3	20	10
4	10	40
5	-	10

$\Sigma x_i n_{Ai} = 120$
$\Sigma x_i n_{Bi} = 240$

Bestimmen Sie die aggregierte Verteilung, sowie deren arithmetisches Mittel, Varianz und Schiefe.

<u>Lösung 9.16:</u>

x_i	n_i
2	10
3	30
4	50
5	10

$\Sigma x_i n_i = 360$

Es gilt $\bar{x}_A = 3$, $\bar{x}_B = 4$, $n_A = 40$, $n_B = 60$, so dass man erhält:
die Gewichte $g_A = 40/100 = 0{,}4$ und $g_B = 0{,}6$, ferner $\bar{x} = 3{,}6$ und für die Varianzen $s_A^2 = 1/2$, $s_B^2 = 1/3$. Die innere Varianz $\Sigma s_j^2 g_j$ beträgt somit 0,4 und für die äußere Varianz $\Sigma(\bar{x}_j - \bar{x})^2 g_j$ erhält man 0,24, so dass die Gesamtvarianz s^2 der aggregierten Verteilung $0{,}4 + 0{,}24 = 0{,}64$ beträgt. Man sieht ferner, dass die Verteilungen in den Teilgesamtheiten A und B symmetrisch sind, die aggregierte Verteilung dagegen rechtssteil ist (Momentschiefe: $- 0{,}168/0{,}8^3 = -0{,}328$). Betrachtet man die Merkmalssummen, so gilt einfach $\Sigma x_i n_i = \Sigma x_i n_{iA} + \Sigma x_i n_{iB} = 120+240 = 360$.

c) Verteilung einer Linearkombination

Gegeben sei eine Häufigkeitsverteilung der Variablen X und eine der Variable Y und gefragt ist nach der Verteilung der Größe X + Y, wobei jedoch die gemeinsame Verteilung von X und Y gegeben sein muss. Das folgende Beispiel möge dies veranschaulichen.

<u>Beispiel 9.17:</u>

Die Variablen X und Y besitzen die folgenden Verteilungen

Randverteilungen

x_i	n_i	y_j	n_j
2	10	2	20
4	12	3	10
5	8		

$\bar{x} = 3{,}6$ $\bar{y} = 2{,}333 = 7/3$

$s_x^2 = 1{,}44$ $s_y^2 = 2/9$

gemeinsame Verteilung

	y=2	y=3
x=2	8	2
x=4	9	3
x=5	3	5

$s_{xy} = 0{,}1667$ (Kovarianz)

Die Variable $Z_1 = X + Y$ ist eine ungewogene Linearkombination und die Variable $Z_2 = 0{,}8X + 0{,}2Y$ ist eine gewogene Linearkombination mit den Gewichten $b_1 = 0{,}8$ und $b_2 = 0{,}2$. Bestimmen Sie die Häufigkeitsverteilungen, (arithmetischen) Mittelwerte und Varianzen der Linearkombinationen Z_1 und Z_2!

Lösung 9.17:

Für die Häufigkeitsverteilungen der Linearkombinationen Z_1 und Z_2 erhält man:

z_{1h}	n_h	z_{2h}	n_h
4	8	2	8
5	2	2,2	2
6	9	3,6	9
7	6	4,2	3
8	5	4,4	3
		4,6	5
	30		30

Die Anwendung obiger Formeln für Mittelwert und Varianz einer Linearkombination führt zu:

ungewogene Linearkombination

$$\bar{z}_1 = \bar{x} + \bar{y} = 5{,}933 = 178/30$$
$$s_{z1}^2 = s_x^2 + s_y^2 + 2s_{xy} = 1{,}9956$$

gewogene Linearkombination

$$\bar{z}_2 = 0{,}8\bar{x} + 0{,}2\bar{y} = 3{,}34667$$
$$s_{z2}^2 = 0{,}8^2 \cdot s_x^2 + 0{,}2^2 \cdot s_y^2 + 2 \cdot 0{,}8 \cdot 0{,}2 \cdot s_{xy} = 0{,}98.$$

Man beachte, dass insbesondere die Formeln für die Varianzen der Linearkombinationen völlig verschieden sind von denen einer Aggregation im engeren Sinne (Abschn. b). Betrachtet man dagegen einfach die Merkmalssummen, so gilt $Sx_i\, n_i = 108$, $Sy_j\, n_j = 70$ und (wie in Abschn. b) $Sz_{1h}n_h = 178 = 108 + 70$.

d) Aggregation von Mittelwerten, Beziehungszahlen und Quoten, Struktureffekt und Standardisierung

Angenommen, für jede Teilgesamtheit j sei die Relation $Q_j = x_j/y_j$ (j=1,2,...,J) eine sinnvolle Beziehungszahl. Die für die Gesamtmasse definierte Beziehungszahl Q der Summen $X = Sx_j$ und $Y = Sy_j$ steht in folgender Beziehung zu den Teilbeziehungszahlen Q_j.

$$(9.22) \quad Q = \frac{X}{Y} = \Sigma Q_j g_{yj} \quad \text{mit} \quad g_{yj} = \frac{y_j}{\Sigma y_j} = \frac{y_j}{Y}$$

danach ist Q das **arithmetische Mittel** der Q_j, gewogen mit den Größen g_j, die Ausdruck der Struktur der Nennergröße sind.
Man kann Q auch darstellen als **harmonisches Mittel** der Q_j mit den Gewichten $g_{xj} = x_j / \Sigma x_j = x_j/X$, die die Struktur der Zählermasse darstellen

$$(9.23) \quad \frac{1}{Q} = \frac{Y}{X} = \sum \frac{1}{Q_j} g_{xj} \quad \text{mit } g_{xj} = \frac{x_j}{\Sigma x_j} = \frac{x_j}{X} \quad .$$

Wie leicht zu sehen ist, handelt es sich bei g_{xj} und g_{yj} jeweils um $\overline{x}$ (auf die Summe 1) normierte Gewichte.

Gl. 9.22 ist der Ausgangspunkt für eine Betrachtung

1. des bereits behandelten Struktureffekts und Simpson-Paradoxons (vgl. Def. 9.2) und
2. der Standardisierung von Beziehungszahlen.

Beispiel 9.4 demonstrierte den Struktureffekt von Beziehungszahlen. Es zeigte sich, dass die rohe Todesrate (d.h. die Todesrate für die Gesamtheit) nach Gl. 9.22 ein gewogenes arithmetisches Mittel der altersspezifischen Todesraten (d.h. der Größen Q_j für die beiden Alters-Teilgesamtheiten) darstellt. Vergleicht man rohe Todesraten, die unter Verwendung der gleichen Gewichte (also von Standardgewichten) errechnet sind, so spricht man von standardisierten Todesraten.

Def. 9.5: Struktureffekt, Standardisierung

Nach Gl. 9.22 ist eine aggregierte (für die Gesamtmasse errechnete) Beziehungszahl Q = X/Y das gewogene arithmetische Mittel der Teil-Beziehungszahlen $Q_j = x_j / y_j$ (j=1,1,...,J)

$$(9.22) \quad Q = \Sigma Q_j g_{yj} \quad .$$

Daraus folgt: Zwei Beziehungszahlen Q_A und Q_B für Gesamtheiten A und B, die sich jeweils in J Teilmassen gliedern lassen, können sich unterscheiden aufgrund unterschiedlicher

a) Teil-Beziehungszahlen Q_{Aj}, Q_{Bj}
b) Gewichte der Nennermasse g_{Ayj}, g_{Byj} .

Die Unterschiedlichkeit aufgrund von a) gilt als "echter" Unterschied, diejenige aufgrund von b) wird als Struktureffekt gedeutet. Um die echten Unterschiede herauszuarbeiten, vergleicht man nicht Q_A mit Q_B, sondern

$$(9.24) \quad Q_A^* = \sum Q_{Aj} g_j^* \quad \text{mit} \quad Q_B^* = SQ_{Bj} g_j^* ,$$

d.h. man vergleicht Beziehungszahlen, die unter Zugrundelegung der gleichen Gewichte (Standardgewichte) g_j^* berechnet sind. Die Größen Q^* heißen dann standardisierte Beziehungszahlen.

Bemerkungen zur Def. 9.5:

1. **Standardisierte und erwartungsgemäße Verhältniszahlen**
   So wie Verhältniszahlen bei gleicher Gewichtung (gleicher Struktur) verglichen werden können (d.h. standardisierte Zahlen) können auch Verhältniszahlen bei gleicher Sachkomponente verglichen werden, also etwa Q_A^0 mit Q_B^0 gem. Übersicht 9.6. Man spricht auch von "erwartungsgemäßen" Verhältniszahlen.

2. **Zerlegung in den Struktureffekt und den echten Effekt**
   Der Struktureffekt kann wie folgt dargestellt werden:

   $Q^D = Q_A - Q_B$

   ist die Differenz der globalen (Gesamt)-Beziehungszahlen oder Quoten und

   $Q_j^D = Q_{Aj} - Q_{Bj}$

   ist die Unterschiedlichkeit der Teil-Beziehungszahlen $1 \leq j \leq J$ oder (**Q_j,^{D} = Quotendifferenzen**).

   $G_j^D = g_{Ayj} - g_{Byj}$

   gibt die Unterschiedlichkeit der Gewichte wieder (**G_j^D = Gewichtsdifferenzen**).
   Im folgenden soll der Einfachheit halber das Subskript y weggelassen werden, also $G_j^D = g_{Aj} - g_{Bj}$. Dann gilt:

(9.25) $Q^D = \sum Q_{Aj}G_j^D + \sum g_{Bj}Q_j^D = S + E$

Der erste Summand S ist Ausdruck der Unterschiedlichkeit der Gewichte (G_j^D). Er ist also der Struktureffekt, ausmultipliziert

(9.25a) $S = \sum Q_{Aj}g_{Aj} - \sum Q_{Aj}g_{Bj}$.

Der zweite Summand E ist Ausdruck der unterschiedlichen Teil-Beziehungszahlen oder Teilquoten (Q_j^D)bei gleicher Gewichtung mit g_{Bj}, also

(9.25b) $E = \sum Q_{Aj}g_{Bj} - \sum Q_{Bj}g_{Bj}$,

und somit ein Maß für den "echten" Unterschied.

Übersicht 9.6: Struktureffekt und echter Effekt bei Verhältniszahlen

(am Beispiel von Beziehungszahlen Q = X/Y)

	zu vergleichende Größen	Unterschiedlichkeit ist Ausdruck von
unbereinigter Vergleich	$Q_A = \sum_j Q_{Aj}\, g_{Aj}$ mit $Q_B = \sum_j Q_{Bj}\, g_{Bj}$ (j=1,2,..,n)	Struktureffekt **und** echter Effekt (Gl. 9.25)
standardisierte Verhältniszahlen	Q_A^* mit Q_B^* gem. Gl. 9.24; gleiche Strukturkomponente g_j^*; ungleiche Sachkomponenten Q_{Aj}, Q_{Bj}	einem echten Unterschied (bei $g_j^* = g_{Bj}$ ist die Differenz $Q_A^* - Q_B^* = E$ gem. Gl. 9.25b)
erwartungsgemäße Verhältniszahlen	$Q_A^0 = \sum Q_j^0 g_{Aj}$ mit $Q_B^0 = \sum Q_j^0 g_{Bj}$ gleiche Sachkomponente Q_A^0, ungleiche Strukturkomponenten g_{Aj}, g_{Bj}	einem Strukturunterschied (wenn $Q_j^0 = Q_{Aj}$ ist die Differenz $Q_A^0 - Q_B^0 = S$ gem. Gl. 9.25a).

3. **Standardisierung**

Der Unterschied zwischen Q und Q^*, der nichtstandardisierten und der standardisierten Beziehungszahl, ist eine mit den Teil-Beziehungszahlen gewogene Summe der Differenzen zwischen den empirischen Gewichten g_j und den Standardgewichten g_j^*

(9.26) $Q - Q^* = \sum Q_j(g_j - g_j^*)$,

wobei die Differenzen wie G_j^D in Gl. 9.25 Unterschiede im Wägungsschema wiedergeben, so dass man die Differenz zwischen einer nicht-standardisierten Beziehungszahl Q und der entsprechenden standardisierten Beziehungszahl Q^* als Ausdruck des Struktureffekts deuten kann.

4. **Simpson-Paradoxon**
Das in Def. 9.2 beschriebene Simpson-Paradoxon, wonach z.B. ein Mittelwert für eine Gesamtheit A größer sein kann, als für eine andere B, obgleich er in allen Teilgesamtheiten von A kleiner ist als in denen von B, ergibt sich aus Gl. 9.25a (in Analogie zu Gl. 9.25):

$$\boxed{(9.25a) \quad \bar{x}^D = \Sigma \bar{x}_{Aj} G_j^D + \Sigma g_{Byj} M_j^D \;.}$$

Danach kann sehr wohl $\bar{x}^D = \bar{x}_A - \bar{x}_B > 0$ sein, obgleich für alle j gilt $M_j^D = \bar{x}_j - \bar{x}_j < 0$ (die Mittelwertdifferenzen M_j^D treten in Gl. 9.25a an die Stelle der Quotendifferen-zen Q_j^D von Gl. 9.25).

Beispiel 9.18:

Bekanntlich ist die Lohnquote der Anteil der Bruttoeinkommen aus unselbständiger Arbeit (L_t) am Volkseinkommen (Y_t). Dass diese Quote im Zeitablauf zunimmt ist schon deshalb zu erwarten, weil der Anteil der unselbständig Beschäftigten an den Beschäftigten zunimmt. Neben anderen "Bereinigungen" wird deshalb häufig eine Standardisierung dergestalt vorgenommen, dass man die Lohnquote berechnet, wie sie wäre, wenn sich dieser Anteil gegenüber einer Basisperiode (z.B. 1960) nicht verändert hätte. Wie ist eine solche Berechnung durchzuführen und was ist hieran kritisch anzumerken?

Lösung 9.18:

a) Berechnung: Es sei $Q_t = L_t/Y_t$ die "Lohnquote" zur Zeit t. Dann ist bei einer Anzahl U_t von Unselbständigen und E_t von Erwerbstätigen $l_t = L_t/U_t$ eine Art "Durchschnittslohn" und entsprechend $y_t = Y_t/E_t$ eine Art gesamtes (Arbeits- <u>und</u> Profit-) "Durchschnittseinkommen" für die Periode t. Es gilt also, dass die Lohnquote das Produkt einer "Pro-Kopf-Lohnquote" $q_t = l_t / y_t$ und einer Unselbständigen- oder "Abhängigen"-quote $a_t = U_t / E_t$ ist, denn:

$$Q_t = \frac{L_t}{Y_t} = \frac{l_t}{y_t} \cdot \frac{U_t}{E_t} = q_t a_t \quad \text{(unbereinigte Lohnquote).}$$

Die "bereinigte" (mit der Beschäftigtenstruktur des Basisjahres 0 gewogene, d.h. standardisierte) Lohnquote ist dann:

$$Q_t^* = \frac{l_t U_0}{y_t E_0} = q_t a_0 \quad \text{mit } q_t = \frac{l_t}{y_t} = \frac{L_t}{U_t}\frac{Y}{E_t}. \quad \text{(bereinigte Lohnquote)}$$

Abb. 9.6 stellt die unbereinigte Lohnquote Q_t (durchgezogene Linie) und die mit der Struktur a_o von 1960 bereinigte Lohnquote $Q_t,^*$ (gestrichelte Linie) dar.

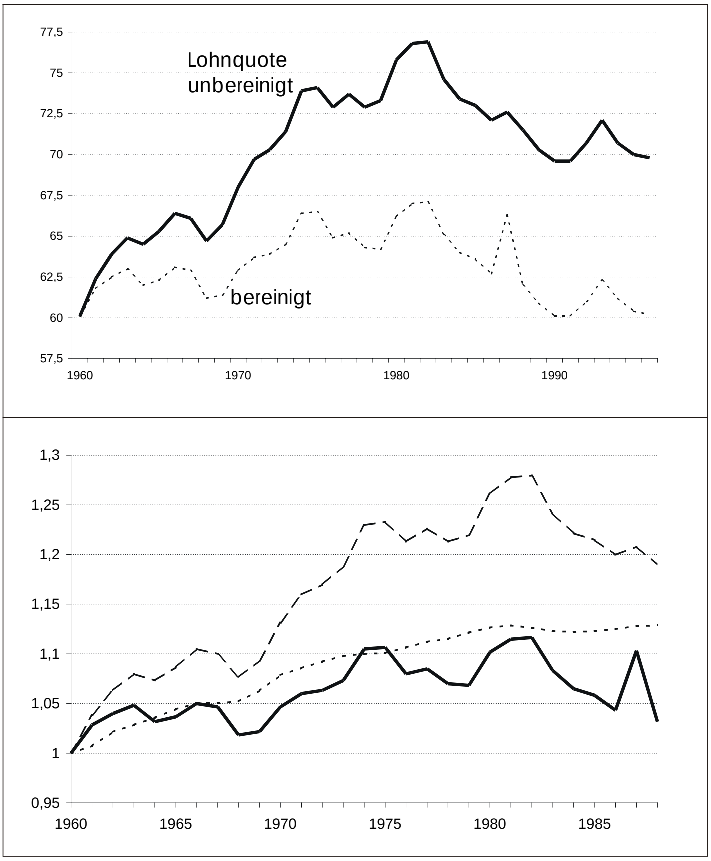
Offenbar ist

- $Q_t/Q_t^* = a_t/a_0$, d.h. die Veränderung (Messzahl) der Abhängigenquote (mittlere Linie in ***Abb. 9.7***) oder die Differenz $Q_t\text{-}Q_t^* = q_t(a_t\text{-}a_0)$ ist Ausdruck des Struktureffekts. Demgegenüber gilt
- $Q_t^*/Q_0 = q_t/q_0$ (untere durchgezogene Linie in Abb. 9.7) oder die Differenz $Q_t^*\text{-}Q_o = a_0(q_t\text{-}q_0)$ als echte Zunahme der Lohnquote (denn $Q_0^* = Q_0$).
- Der Gesamteffekt $Q_t /Q_0 = (Q_t,^* /Q_0)(Q_t /Q_t,^*) = (q_t /q_0)(a_t /a_0)$ ist dann das Produkt der beiden Effekte (obere gestrichelte Linie in Abb. 9.8), bzw. die Differenz $Q_t\text{-}Q_0 = (Q_t\text{-}Q_t,^*) + (Q_t,^*\text{-}Q_0) = q_t(a_t\text{-}a_0) + a_0(q_t\text{-}q_0) = S + E$ ist die Summe der beiden Effekte (die letzte Gleichung ist analog Gl. 9.25 gebildet).

Die Daten (Abb. 9.6,7) zeigen, dass die Abnahme der Lohnquote v.a. Ergebnis einer echten Veränderung der Verteilungsrelation ist. Das wird demonstriert am Vergleich von unbereinigter und bereinigter (standardisierter) Lohnquote (Abb. 9.6) und an der Abnahme von q_t/q_0 (untere Linie in Abb. 9.7). Die Abnahme der Lohnquote ist dagegen nicht auf einen Struktureffekt (a_t/a_0) zurückzuführen. Letzterer hätte eher im Gegenteil eine Zunahme der Lohnquote bewirken müssen.

b) Der kritische Punkt bei dieser Art von Betrachtung, die in der Statistik sehr verbreitet ist (so sind z.B. Preisindizes Verhältnisse von standardisierten Ausgaben), ist dass die Pro-Kopf-Lohnquote $q_t = l_t /y_t$ sich wahrscheinlich nicht so entwickelt hätte, wie sie sich tatsächlich entwickelt hat, wenn die Abhängigenquote konstant geblieben wäre. Es wird also eine Unabhängigkeit der Einkommensentwicklung von der Beschäftigtenstruktur unterstellt, die mit Sicherheit (was sich auch theoretisch begründen läßt) nicht gegeben ist. Außerdem ist die Beschäftigtenstruktur nicht der einzige Faktor, der die echte Lohnzunahme überlagert. Ein anderer Effekt ist z.B. die Arbeitszeitverkürzung. Danach wäre es auch gerechtfertigt, die Lohnquote nicht nur mit einer Standardstruktur hinsichtlich der Beschäftigten (unselbständig oder nicht), sondern auch hinsichtlich der Arbeitszeit zu standardisieren, d.h. die Lohnquote zu berechnen, die sich ergeben hätte, wenn die Produktivitätssteigerung voll in Lohnsteigerungen und nicht z.T. auch in Arbeitszeitverkürzungen weitergegeben wäre.

Abb. 9.6 (oben) und 9.7 (unten): Erläuterungen in der Lösung 9.18



e) Aggregation von Messzahlen und Wachstumsraten

1. Aggregation von Messzahlen

Die Variable Y_t, deren Messzahlen $m_{0t} = Y_t/Y_0$ zu betrachten sind, soll ein Aggregat darstellen, d.h. es gelte zu jeder Periode t jeweils $Y = Sy_j$. Wie man leicht sieht, gilt

$$\text{(9.22a)} \qquad m_{0t} = \frac{Y_t}{Y_0} = \frac{y_{1t}}{y_{10}}\frac{y_{10}}{Y_0} + \frac{y_{2t}}{y_{20}}\frac{y_{20}}{Y_0} + \ldots = \Sigma\, m_{0t}^{j}\, g_{j0}$$

wobei die Gewichte g_{j0} jeweils die Anteile y_{j0}/Y_0, also der Teilbeträge y_j am Gesamtbetrag Y zur Basiszeit sind. Gl. 9.22a ist in Analogie zu Gl. 9.22 zu sehen und man kann analog zu Gl. 9.23 auch die Messzahl eines Aggregats als harmonisches Mittel darstellen mit Gewichten g_{jt}, die sich auf die Berichtszeit beziehen, also

$$\text{(9.23a)} \quad \frac{1}{m_{0t}} = m_{t0} = \sum_j \frac{1}{m_{0t}^j} g_{jt} = \sum_j m_{0t}^j g_{jt} \quad .$$

Gl. 9.26 bzw. 9.25a ist entsprechend bei Differenzen anzuwenden.

Beispiel 9.19:

Es gilt die Gleichung: Inlandsprodukt (P) plus Außenbeitrag (A) ist das Sozialprodukt Y. Für ein Land mögen die folgenden Zahlen gelten (in Mrd DM):

Jahr	P	A	Y
1970	660	15	675
1990	2000	120	2120

Man verifiziere Gl. 9.22a!

Lösung 9.19:

Die Messzahlen des Inlandsprodukts und Außenbeitrags betragen $m_{0t}^P = 3{,}0303$ und $m_{0t}^A = 8$. Die Gewichte betrugen zur Basiszeit 660/675 und 15/675. Die Messzahl $m_{0t}^Y = 2120/675 = 3{,}14074$ ist ein hiermit gewogenes arithmetisches Mittel der einzelnen Messzahlen.

2. Aggregation von Wachstumsfaktoren und Wachstumsraten

a) Aggregation

Es ist unmittelbar einsichtig, dass für Wachstumsfaktoren Gl. 9.22a analog gilt, wobei der Wachstumsfaktor w_t des Aggregats $Y=\Sigma y_j$ als gewogenes arithmetisches Mittel der Wachstumsfaktoren der Summanden j =1,2,...J darzustellen ist

$$\text{(9.22b)} \quad w_t = \frac{Y_t}{Y_{t-1}} = \Sigma(y_{jt}/y_{j,t-1})g_{j,t-1} = \Sigma w_{jt} g_{j,t-1}$$

wobei die Gewichte die Anteile $g_{j,t-1} = y_{j,t-1}/Y_{t-1}$ der Summanden am Gesamtmerkmalsbetrag in der Vorperiode sind. Natürlich kann man auch Gl. 9.23a analog anwenden, wenn die Gewichte $g_{j,t}$ die entsprechenden Anteile
zur Zeit t darstellen.

Ausgehend von den Wachstumsfaktoren w können mit r = w - 1 entsprechende Aussagen über die Wachstumsraten (r) gemacht werden.

$$\text{(9.27)} \quad r_t = \Sigma r_{jt}\, g_{j,t-1}.$$

b) Struktureffekt

Ist eine Größe Z eine Summe von X und Y, so gilt für die **Wachstumsrate** der Linearkombination Z nach Gl. 9.27:

$$r_z = r_x g_x + r_y g_y \quad \text{mit} \quad g_x = \frac{X_{t-1}}{Z_{t-1}} \quad \text{und} \quad g_y = \frac{Y_{t-1}}{Z_{t-1}}.$$

Entsprechend kann man die **Differenz von Wachstumsraten** einer Linearkombination zerlegen:

Die Differenz aufeinanderfolgender Wachstumsraten $r_{zt} - r_{z,t-1}$ ist darstellbar als die folgende Summe:

der erste Summand $[(r_{xt} - r_{x,t-1})g_{x,t-1} + (r_{yt} - r_{y,t-1})g_{y,t-1}]$ ist Ausdruck des Wachstumsef-
fekts und der zweite Summand $[r_{x,t-1}(g_{x,t-1} - g_{x,t-2}) + r_{y,t-1}(g_{y,t-1} - g_{y,\, t-2})]$ ist Ausdruck des Struktureffekts.

Dies ist eine zu Gl. 9.25 analoge Betrachtung mit Wachstumsraten.

Man kann ferner die Wachstumsrate einer Linearkombination, etwa von Z = X + Y für den Zeitraum von 0 bis t zerlegen in einen Teil, der auf den isolierten Einfluß von X und einen Teil der auf den isolierten Einfluß von Y zurückgeführt werden kann. Mit den Gewichten $g_x = X_0/Z_0$ und $g_y = Y_0/Z_0$ erhält man

$$(9.28) \quad r_{z,t} = \frac{Z_t}{Z_0} - 1 = (g_x w_x + g_y) - 1 + (g_y w_y + g_x) - 1.$$

Interpretation:
Die erste Klammer stellt den isolierten Einfluß von X auf die Summe Z dar, denn $g_x w_x + g_y = (X_t + Y_0)/Z_0$. Entsprechend ist der zweite Klammerausdruck der isolierte Einfluss
von Y auf die Summe Z: $g_y w_y + g_x = (Y_t + X_0)/Z_0$. Berücksichtigt man, dass gilt $g_x + g_y =$
1 und r = w - 1, so ist Gl. 9.28 nur eine andere Darstellung von Gl. 9.27.
Betrachtungen dieser Art haben seinerzeit für die Statistik-Ausbildung von Ökonomen in der DDR eine große Bedeutung gehabt. In dem bis zur "Wende" 1989 als "Bibel" geltenden Statistik-Lehrbuch von Donda et al. wird dieser Zusammenhang über viele Seiten als "additive Faktorenanalyse mit Indizes" (d.h. Messzahlen) ausgeführt. Das folgende Beispiel 9.20 zeigt eine Anwendung dieser Art.

Beispiel 9.20:

Die Kosten Z eines Betriebes setzen sich aus Material- (X) und Arbeitskosten (Y) zusammen. Es mögen die folgenden Zahlen gelten:

Jahr	X	Y	Z
1970	75	25	100
1980	80	40	120
1990	100	80	180

a) Wie lässt sich die Zunahme der Kosten von 1980 bis 1990 um 50% von 120 auf 180 auf zunehmende Arbeits- und zunehmende Materialkosten verteilen?

b) Man zerlege die **Veränderung** der Wachstumsrate von 20% (1970-80) auf 50% (1980-90) in Wachstums- und Struktureffekt.

Lösung 9.20:

a) Nach Gl. 9.28 gilt mit $g_x = 2/3$ und $g_y = 1/3$. Die Zunahme von 50% ist dann: 0,5 = (2/3·10/8 + 1/3) - 1 + (1/3·2 + 2/3) - 1. Die erste Klammer führt zu 7/6 = (100 + 40)/120 und ist Ausdruck der Erhöhung der Materialkosten von 80 auf 100. Der zweite Klammerausdruck beträgt 4/3 und dies ist zugleich (80 + 80)/120, also die Wirkung der Erhöhung der Lohnkosten von 40 auf 80. Man kann also sagen, dass die Zunahme der Gesamtkosten um 50% zurückzuführen ist auf 1/6 Material- und 1/3 Arbeitskosten (denn 1/6 + 1/3 = 1/2).

b) Die Differenz der Wachstumsraten von 0,5 und 0,2 (also 0,3) läßt sich zerlegen in einen Wachstumseffekt (20/80 - 5/75)(2/3) + (40/40 - 15/25)(1/3) = 23/90
und in einen Struktureffekt (5/75)(2/3 - 3/4) + (15/25)(1/3 - 1/4) = 4/90 [23/90 + 4/90 = 0,3].
Das Gewicht von X war zur Zeit t-1 (1980): 2/3, dagegen zur Zeit t-2 (1970): 3/4. Es ist von 1980 bis 1990 beständig zurückgegangen und entsprechend ist das Gewicht der Arbeitskosten gestiegen (von 1/4 1970 über 1/3 im Jahr 1980 bis 4/9 im Jahre 1990). Dies ist für den Struktureffekt verantwortlich, dessen Bedeutung mit 4/90 ganz wesentlich geringer ist als der Wachstumseffekt (23/90).