Saisonbereinigte Arbeitslosenzahlen, Teil 2

Unterrichtssequenz für den 10.Schuljahrgang

Von GÜNTER NORDMEIER, Bad Essen

Zusammenfassung: *Es wird ein Unterrichtsgang vorgestellt, der über das einfache exponentielle Glätten zum exponentiellen Glätten mit Saisonbereinigung führt. Die entstehenden Glättungslinien werden als Näherungen für Trendkurven aufgefaßt und entsprechend interpretiert. Saisonfiguren werden herausgearbeitet und diskutiert. Das Verfahren wird mit seinen Stärken und Schwächen als ein mögliches mathematisches Modell für die Analyse von Zeitreihen mit deutlichen saisonalen Schwankungen dargestellt.*

1. Einleitung

1.1 Fächerübergreifende Aspekte

Im vorangehenden Aufsatz "Saisonbereinigte Arbeitslosenzahlen I" in "Stochastik in der Schule" 14 (H. 3/1994) wurde die Bedeutung anwendungsorientierten Lernens herausgearbeitet und begründet, warum wir uns mit der Saisonbereinigung von Arbeitslosenzahlen befassen. Diese im thematischen Zusammenhang wichtigen Ausführungen möchten wir hier nicht wiederholen, jedoch die Leserinnen und Leser freundlich darauf hinweisen.

Zwei Bemerkungen sollen jedoch diesem zweiten Aufsatz zur Saisonbereinigung von Arbeitslosenzahlen noch vorangeschickt werden:

1. Die Analyse von Zeitreihen ist eine relativ junge Spezialdisziplin der Beschreibenden Statistik. Sie hat beispielsweise für die Wirtschaftsprognostik eine außerordentlich große Bedeutung, aber nicht nur dort. In den meisten Lehrbüchern zur Wirtschaftsstatistik findet man entsprechende Kapitel zur Zeitreihenanalyse. Darüber hinaus findet man im Bereich der Wirtschaftswissenschaften eine Fülle von spezieller Literatur zu diesem Thema. Doch dies sind Quellen, die eine Lehrerin oder ein Lehrer der allgemeinbildenden Schule in der Regel nicht nutzt. Aus den elementaren Methoden der Zeitreihenanalyse wird in diesem Aufsatz das einfache und das erweiterte exponentielle Glätten "didaktisch aufbereitet" dargestellt. Dabei erleben die Schülerinnen und Schüler experimentelle Mathematik wie selten in ihrer Schulzeit. Sie erfahren, was es heißt, ein mathematisches Modell auszuwählen und ggf. Parameter zu variieren. Die saisonbereinigt geglätteten Werte schließlich können nicht mit dem Raster "falsch oder richtig" beurteilt werden. Man muß vielmehr darüber nachdenken, ob sie hinreichend genau der Wirklichkeit angepaßt sind und welche Interpretationen sie zulassen oder nicht.

2. Im aktuellen politischen Tagesgeschehen ist die Entwicklung der Arbeitslosenzahlen eines der zentralen Themen. Im Bundestagswahlkampf 1994 konnte man z.B. erleben, daß der leichte Rückgang der absoluten Arbeitslosenzahlen in den Monaten Juli bis September 1994 von den Vertretern der Parteien durchaus kontrovers interpretiert wurde. Die einen sahen darin einen Indikator für eine anspringende Konjunktur, die anderen erklärten, der leichte Rückgang sei nur auf saisonale Einflüsse zurückzuführen, und die hohe Arbeitslosigkeit halte an bzw. stiege noch an. In dieser Situation hätte eine Diskussion der saisonbereinigten Zahlen für eine Versachlichung gesorgt. Die dazu notwendigen Daten stellt das Statistische Bundesamt in bestimmten periodisch erscheinenden Veröffentlichungen laufend zur Verfügung. Hier - wie auch sonst im Leben - gilt sicher: Man versteht das Prinzip einer Saisonbereinigung besser, wenn man selbst (z.B. in seiner Schulzeit!) einmal ganz konkret versucht hat, diesen Schritt einer Zeitreihenanalyse zu vollziehen.

3. Für eine Zeitreihenanalyse benutzt man am besten den Computer und setzt dabei integrierte Software ein, die mindestens ein Tabellenkalkulationsprogramm und ein Grafikprogramm enthält. In diesem Aufsatz greifen wir auf MS-WORKS zurück.

Stochastik in der Schule 15(1995), Nr. 2, S. 4-26

1.2 Einfaches exponentielles Glätten

Die Glättungsformel

Die einzelnen Daten der zu untersuchenden Zeitreihe werden hier mit y_1, y_2, y_3, ... , y_n bezeichnet. Dabei ist y_t der Wert der Zeitreihe zum Zeitpunkt t. Die durch exponentielles Glätten bestimmten glatten Werte bezeichnen wir mit g_t.

Beim einfachen exponentiellen Glätten setzt man:

$g_1 = y_1$

$g_2 = \alpha \cdot y_2 + (1-\alpha) \cdot g_1$

$g_3 = \alpha \cdot y_3 + (1-\alpha) \cdot g_2$

.

.

.

$g_t = \alpha \cdot y_t + (1-\alpha) \cdot g_{t-1}$ mit $0 < \alpha < 1$		[1]
"Glättungsformel"		

Die Bedeutung des Glättungsparameters α

Der Einfluß des Glättungsparameters α läßt sich im Unterricht etwa wie folgt beschreiben:

1. Je stärker die jüngeren Zeitreihenwerte berücksichtigt werden sollen, desto größer muß α gewählt werden.
2. Je stärker die weiter zurückliegenden Werte berücksichtigt werden sollen, desto kleiner muß α gewählt werden.
3. Je größer α gewählt wird, desto stärker wird die Anpassung an die aktuelle Entwicklung vollzogen.
4. Je kleiner α gewählt wird, desto stärker wird die Reihe geglättet.

Diese Einsichten lassen sich experimentell gewinnen, indem man eine ausgewählte kurze Zeitreihe mit verschiedenen Werten für α (z.B. $\alpha_1=0{,}2$; $\alpha_2=0{,}4$; $\alpha_3=0{,}6$; $\alpha_4=0{,}8$) exponentiell glättet und die Ergebnisse vergleicht. Besonders augenfällig ist dies bei der Auswertung der zugehörigen Diagramme. Man kann den Einfluß des Glättungsparameters α jedoch auch direkt in der "Glättungsformel" ablesen, wenn man sich klarmacht, daß der erste Summand $\alpha \cdot y_t$ für die Angleichung an die aktuelle Entwicklung sorgt, während der zweite Summand $(1-\alpha) \cdot g_{t-1}$ bewirkt, daß aufeinanderfolgende Werte nicht zu sehr voneinander abweichen.

Zur Namensgebung

Die folgende - von Schülerinnen und Schülern des 10.Schuljahrgangs nachzuvollziehende - Umformung führt zur Gleichung [2]. Durch fortlaufendes Einsetzen erhält man:

$$\begin{aligned}
g_2 &= \alpha \cdot y_2 + (1-\alpha) \cdot y_1 \\
g_3 &= \alpha \cdot y_3 + (1-\alpha) \cdot [\alpha \cdot y_2 + (1-\alpha) \cdot y_1] \\
&= \alpha \cdot y_3 + \alpha \cdot (1-\alpha) \cdot y_2 + (1-\alpha)^2 \cdot y_1 \\
g_4 &= \alpha \cdot y_4 + (1-\alpha) \cdot [\alpha \cdot y_3 + \alpha \cdot (1-\alpha) \cdot y_2 + (1-\alpha)^2 \cdot y_1] \\
&= \alpha \cdot y_4 + \alpha \cdot (1-\alpha) \cdot y_3 + \alpha \cdot (1-\alpha)^2 \cdot y_2 + (1-\alpha)^3 \cdot y_1
\end{aligned}$$

usw.

Allgemein gilt:

$$g_t = \alpha \cdot y_t + \alpha \cdot (1-\alpha) \cdot y_{t-1} + \alpha \cdot (1-\alpha)^2 \cdot y_{t-2} + \ldots + (1-\alpha)^{t-1} \cdot y_1 \qquad [2]$$

In Gleichung [2] wird deutlich, warum das hier behandelte Verfahren "exponentielles Glätten" heißt. Man kann jeden glatten Wert g_t aus den Werten der Zeitreihe von y_1 bis y_t berechnen. Dabei nimmt die Gewichtung der verwendeten Daten rückwärtsschreitend exponentiell ab.

Literaturhinweis: Ein für den Unterricht aufbereitetes Beispiel für exponentielles Glätten findet man in meinem Aufsatz "Beiträge zur elementaren Zeitreihenanalyse - Teil 2" in "Stochastik in der Schule" 14(1994), Heft 1, Seite 49ff.

2. Exponentielles Glätten nach Bereinigung von Saisoneinflüssen (1.Erweiterung des exponentiellen Glättens)

In der Regel wird man sich bei der Unterrichtsplanung entscheiden müssen, ob man zur Zeitreihenanalyse ein Zerlegungsverfahren (siehe Aufsatz "Saisonbereinigte Arbeitslosenzahlen I") oder exponentielles Glätten oder gar bestimmte Wachstumsfunktionen einsetzen will. Schon aus Zeitgründen ist es nicht möglich, alle zugänglichen Methoden im Unterricht des Pflichtbereichs anzubieten. Deshalb wird im folgenden Unterrichtsentwurf, der für die Analyse der Zeitreihe das exponentielle Glätten vorsieht, nicht auf Begriffsbildungen und

Fachmethoden der im letzten Aufsatz beschriebenen Zerlegungsverfahren zurückgegriffen.

2.1 Das einleitende Unterrichtsgespräch

Unsere Unterrichtssequenz sollte mit einem Unterrichtsgespräch zum Problemkreis "Arbeitslosigkeit" beginnen. Als Anlaß könnten aktuelle Medienberichte und Kommentare über die Entwicklung der Arbeitslosenzahlen in der Bundesrepublik Deutschland, über regionale Entwicklungen am Arbeitsmarkt oder über den Abbau von Belegschaften einzelner Unternehmen dienen. Die Schülerinnen und Schüler werden über sozialpolitische Aspekte von Arbeitslosigkeit sprechen, Informationen über die Tätigkeit der Arbeitsämter erhalten und auch wissen wollen, auf welche Weise die Arbeitslosenzahlen festgestellt werden. An geeigneter Stelle dieses Gesprächs wird Tab.1 vorgelegt.

Tab.1: A r b e i t s l o s e (Früheres Bundesgebiet)
Ursprungswerte in 1000

	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994
Jan.	2539	2619	2590	2497	2519	2335	2191	1874	1875	2257	2736
Febr.	2357	2611	2593	2488	2517	2305	2153	1869	1863	2288	2742
März	2393	2474	2448	2412	2440	2178	2013	1731	1768	2223	2640
April	2253	2305	2230	2216	2262	2035	1915	1652	1747	2197	2590
Mai	2133	2193	2122	2099	2149	1947	1823	1604	1704	2148	2506
Juni	2113	2160	2078	2097	2131	1915	1808	1593	1715	2166	2478
Juli	2202	2221	2132	2176	2199	1973	1864	1694	1828	2326	2570
Aug.	2202	2217	2120	2165	2167	1940	1813	1672	1822	2315	2531
Sept.	2143	2152	2046	2107	2100	1881	1728	1610	1784	2288	2452
Okt.	2145	2149	2026	2093	2074	1874	1687	1599	1830	2359	2446
Nov.	2189	2211	2068	2133	2091	1950	1685	1618	1885	2408	2450
Dez.	2325	2347	2218	2308	2190	2052	1784	1731	2025	2514	2545
Jahr	2266	2304	2228	2229	2242	2038	1883	1689	1808	2270	

Quelle: Statistisches Bundesamt (Herausgeber):
- Fachserie 1, Reihe 4.3 "Erwerbstätigkeit und Arbeitsmarkt" (erscheint monatlich)
- "Statistischer Wochendienst" (Aktualisierung der Monatszahlen)
- "Statistisches Jahrbuch für die Bundesrepublik Deutschland" (nur Jahresdurchschnitte)

Die monatlichen Arbeitslosenzahlen werden auch in der Tagespresse veröffentlicht.

Wir gehen davon aus, daß sich in diesem Unterrichtsgespräch u.a. die folgenden Fragen ergeben:

1. Wie haben sich die Arbeitslosenzahlen langfristig entwickelt ?
2. Wie kann man den Trend möglichst genau beschreiben ?
3. Wie ist die aktuelle Entwicklung zu beurteilen ?
4. Welche Prognosen sind aufgrund der vorliegenden Daten möglich ?

Das Bemühen, Antworten auf diese Fragen zu erhalten, soll unsere Untersuchungen leiten.

2.2 Exponentielles Glätten der Jahresdurchschnitte

Die Schülerinnen und Schüler werden voraussichtlich vorschlagen, zunächst die Entwicklung der Jahreszahlen zu untersuchen, um einen Überblick über die langfristige Entwicklung zu geben. Dies ist eine gute Gelegenheit, um das exponentielle Glätten "einzuführen" (siehe auch Kap.1.2).
Man legt mit Hilfe des Computers und eines integrierten Tabellenkalkulations und Grafikprogramms eine Arbeitstabelle an (Tab.2).
Als Glättungsparameter haben wir 0,2; 0,3; 0,4; 0,5 und 0,6 ausgewählt. In dieser Phase des Unterrichts kann in den einzelnen Gruppen der Klasse durchaus experimentiert und die Wirkung verschiedener Glättungsparameter erprobt werden. Ist es das erste Umgehen mit dem exponentiellen Glätten, so ist diese Phase außerordentlich wichtig. Die vorn zusammengestellten Gesetzmäßigkeiten exponentiellen Glättens müssen bewußt gemacht werden. Mehrfach wird dabei außerdem auch die Eingabe einer "Formel" und das Ausfüllen von Spalten aufgrund einer Bearbeitungsvorschrift geübt, also wesentliche Tätigkeiten beim Einsatz eines Tabellenkalkulationsprogramms.

Die Durchmusterung der Tab.2 führt sicher zu der Einsicht, daß mit wachsendem Glättungsparameter die Anpassung der glatten Werte an die gegebenen Jahresdurchschnitte immer besser gelingt. Ob die glatten Werte schließlich auch "schöne" Ausgleichskurven liefern, ist noch die Frage. Dies muß man schon in einer Grafik sehen. Und diese ist beim eingesetzten Programmpaket mit wenigen Befehlen innerhalb kurzer Zeit auf dem Bildschirm erstellt und ggf. ausgedruckt. Damit alles übersichtlich bleibt, haben wir in Abb.1 nur die Glättungslinien für

α=0,2; α=0,4 und α=0,6 neben den als "Balken" dargestellten ursprünglichen Jahreswerten ausdrucken lassen.

Tab.2: Exponentielles Glätten der Jahresdurchschnitte

Jahr (1)	t (2)	Jahres-durchchnitt (3)	Geglättete Zeitreihe $\alpha = 0{,}2$ (4)	$\alpha = 0{,}3$ (5)	$\alpha = 0{,}4$ (6)	$\alpha = 0{,}5$ (7)	$\alpha = 0{,}6$ (8)
1984	1	2266000	2266000	2266000	2266000	2266000	2266000
1985	2	2304000	*2273600	**2277400	2281200	2285000	2288800
1986	3	2228000	2264480	2262580	2259920	2256500	2252320
1987	4	2229000	2257384	2252506	2247552	2242750	2238328
1988	5	2242000	2254307	2249354	2245331	2242357	2240531
1989	6	2038000	2211046	2185948	2162399	2140188	2119012
1990	7	1883000	2145437	2095064	2050639	2011594	1977405
1991	8	1689000	2054149	1973244	1905984	1850297	1804362
1992	9	1808000	2004919	1923671	1866790	1829148	1806545
1993	10	2270000	2057936	2027570	2028074	2049574	2084618
1994	11	2556000	2157548	2186099	2239244	2302787	2367447

* = 0,2·ZS(-2)+0,8·Z(-1)S (Befehle für MS-WORKS)
** = 0,3·ZS(-3)+0,7·Z(-1)S usw.

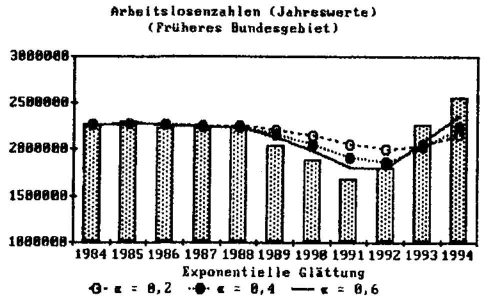

Abb.1: Exponentielles Glätten der Jahresdurchschnitte

Wir sehen:
Die langfristige, über mehr als ein Jahrzehnt gehende Entwicklung der Arbeitslosenzahlen ist in Abb.1 gut zu erkennen. Nach einem auffallenden Rückgang in den Jahren 1989 bis 1991 erfolgte ein steiler Anstieg mit einem(absoluten) Maximum im Jahre 1994. Man kann die Glättungslinien in erster Näherung als Trendkurven auffassen. Die drei Kurven in Abb.1 unterscheiden sich jedoch in der Größe der Ausschläge und in der Lage des Minimums.

Die aktuelle Entwicklung der Arbeitslosenzahlen ist aus den Jahreswerten mit Hilfe der exponentiellen Glättung nicht zu beschreiben. Die Gründe dafür lassen sich im Unterrichtsgespräch herausarbeiten. Die Schülerinnen und Schüler werden vorschlagen, nunmehr doch die monatlichen Arbeitslosenzahlen auszuwerten. Unter der Voraussetzung, daß die Glättung mit Hilfe von n-gliedrigen gleitenden symmetrischen Mitteln bekannt ist, kann der nachfolgende Exkurs sinnvoll sein.

Exkurs:
In Abb.2 sind die Jahresdurchschnitte der Arbeitslosenzahlen mit Hilfe n-gliedriger gleitender symmetrischer Mittel geglättet.

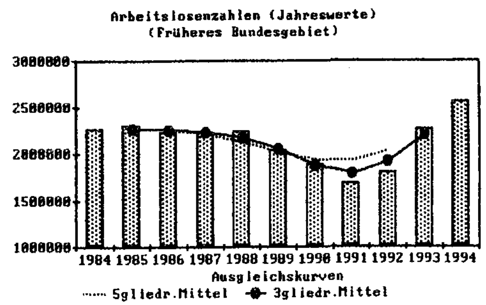

Abb.2: Glättung der Jahresdurchschnitte mit Hilfe gleitender symmetrischer Mittel

Die dabei entstandenen Trendlinien sind zwar "schöne" Ausgleichskurven, brechen jedoch vor dem aktuellen Rand der Zeitreihe ab. Mit ihnen könnte man also auch nicht die aktuelle Entwicklung beschreiben. Beim Vergleich der Abb.1 und 2 werden bestimmte Eigenschaften der benutzten Glättungsverfahren deutlich, z.B.: Beim Glätten mit gleitenden Mitteln fehlen grundsätzlich Werte an den Rändern. Dies ist beim exponentiellen Glätten nicht so, was man als Vorteil dieses Verfahrens auffassen kann.

Die durch exponentielles Glätten erzeugten Kurven haben jedoch den Nachteil, daß sie die aktuellen Trendänderungen stets etwas verspätet aufzeigen, wie man in Abb.1 gut sehen kann. Die Anpassung an die aktuelle Entwicklung gelingt umso besser, je größer der Glättungsparameter α gewählt wird.
Überhaupt läßt sich der Einfluß des Glättungsparameters auf den Verlauf der Trendkurve mit Hilfe der Abb.1 recht gut diskutieren (Siehe auch Kapitel 1.2).

2.3 Exponentielles Glätten der monatlichen Arbeitslosenzahlen

Findige Köpfe unter den Schülerinnen und Schülern benutzen die bisherige Arbeitstabelle weiter und ändern nur die Eingaben in den drei ersten Spalten. Auf diese Weise sparen sie sich das Eingeben der Glättungsformeln in den nachfolgenden Spalten.

Überhaupt sollte man an dieser Stelle des Unterrichtsganges den Schülerinnen und Schülern einen möglichst großen Spielraum zum eigenen Forschen einräumen. Dann werden in den Kleingruppen mit großer Wahrscheinlichkeit unterschiedliche Ansätze verfolgt. Mögliche Alternativen sind u.a.:

1. Eine Gruppe untersucht die Wirkung des exponentiellen Glättens zunächst nur für den Zeitraum der ersten 3 bis 4 Jahre, beispielweise mit der Zielsetzung, einen sinnvollen Glättungsparameter α für die neuen Querschnittsdaten zu bestimmen. Man vergleiche dazu Tab.3 und Abb.3.
2. Eine andere Gruppe wendet sich schwerpunktmäßig der aktuellen Entwicklung zu und gibt deshalb die monatlichen Arbeitslosenzahlen erst ab Januar 1991 ein. Die entsprechende Arbeitstabelle wird hier nicht abgedruckt. Die Abb.4 soll jedoch mögliche Ergebnisse aufzeigen.

3. In anderen Gruppen wird die Zeitreihenanalyse aller 132 Daten versucht. Die Abb.5 zeigt mögliche Arbeitsergebnisse.

Tab.3: Exponentielles Glätten der monatlichen Arbeitslosenzahlen (Neue Arbeitstabelle)

		Monatliche	Geglättete Zeitreihe				
Monat	t	Arbeitslosen-	$\alpha = 0{,}2$	$\alpha = 0{,}3$	$\alpha = 0{,}4$	$\alpha = 0{,}5$	$\alpha = 0{,}6$
(1)	(2)	zahl (3)	(4)	(5)	(6)	(7)	(8)
Jan. 1984	1	2539000	2539000	2539000	2539000	2539000	2539000
Feb. 1984	2	2537000	2538600	2538400	2538200	2538000	2537800
Mär. 1984	3	2393000	2509480	2494780	2480120	2465500	2450920
Apr. 1984	4	2253000	2458184	2422246	2389272	2359250	2332168
Mai 1984	5	2133000	2393147	2335472	2286763	2246125	2212667
Jun. 1984	6	2113000	2337118	2268731	2217258	2179563	2152867
Jul. 1984	7	2202000	2310094	2248711	2211155	2190781	2182347
Aug. 1984	8	2202000	2288475	2234698	2207493	2196391	2194139
Sep. 1984	9	2143000	2259380	2207789	2181696	2169695	2163455
Okt. 1984	10	2145000	2236504	2188532	2167017	2157348	2152382
Nov. 1984	11	2189000	2227003	2188672	2175810	2173174	2174353
Dez. 1984	12	2325000	2246603	2229571	2235486	2249087	2264741
Jan. 1985	13	2619000	2321082	2346399	2388892	2434043	2477296
Feb. 1985	14	2611000	2379066	2425780	2477735	2522522	2557519
Mär. 1985	15	2474000	2396053	2440246	2476241	2498261	2507407
...	..	...	...	...	...	...	...

Nach einer gewissen Bearbeitungszeit werden sich die ersten Gruppen mit Zwischenergebnissen melden, die überraschend sind und hinterfragt werden wollen.

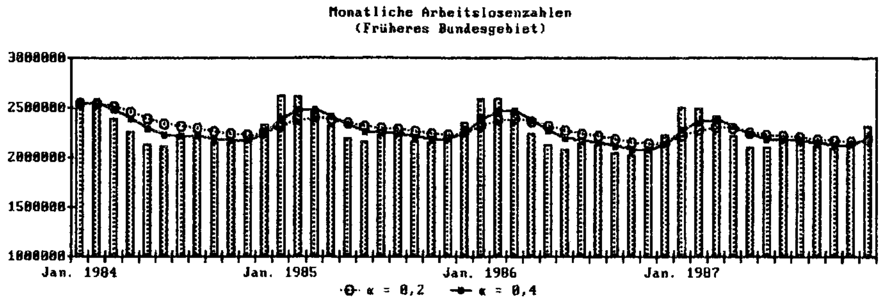

Abb.3: Exponentielles Glätten der monatlichen Arbeitslosenzahlen von 1984 bis 1987

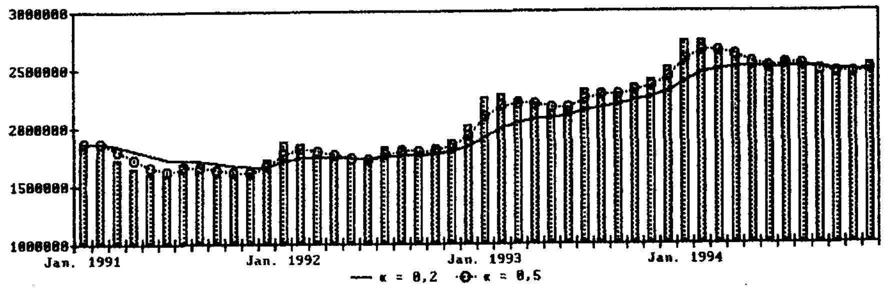

Abb.4: Exponentielles Glätten der monatlichen Arbeitslosenzahlen von 1991 bis 1994

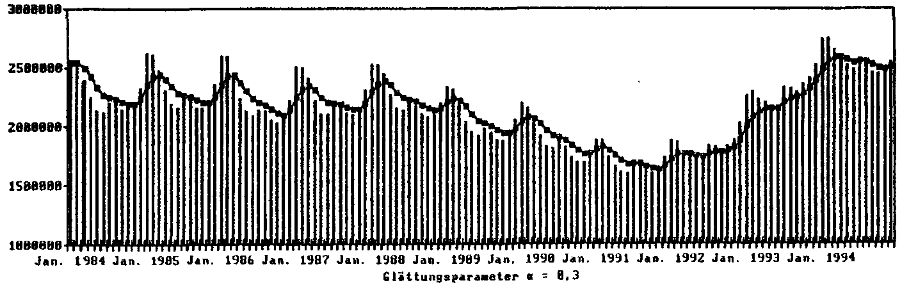

Abb.5: Exponentielles Glätten der monatlichen Arbeitslosenzahlen von 1984 bis 1994

Die einfache exponentielle Glättung der monatlichen Arbeitslosenzahlen führt, welchen Zeitraum man auch auswählt, zu Glättungslinien, die stärker die jahreszeitlichen Schwankungen als die langfristige Entwicklung aufzeigen. Man wird die entstandenen Linien nicht als Trendkurven bezeichnen können. Dieses Phänomen sollte ausführlich diskutiert werden. Die Auswertung der Abbildungen 3 bis 5 ist dabei hilfreich.
Wir haben jedoch ganz wichtige Tatsachen gefunden: Die jahreszeitlichen Schwankungen überlagern in einem sehr starken Maße die über mehrere Jahre laufenden langfristigen Schwingungen, die entweder konjunkturell oder struk-

turell bedingt sind. Es ist üblich, die jahreszeitlich bedingten Zu und Abnahmen der Arbeitslosenzahlen als Saisoneinflüsse zu erklären. Abb.3 zeigt z.B., daß die Auswirkungen dieser Saisoneinflüsse in den Jahren 1984 bis 1987 von Jahr zu Jahr fast gleich waren. Doch sowohl in Zeiten einer starken Abnahme der Arbeitslosigkeit (1989 bis 1991) als auch in Perioden eines enormen Anstiegs der Arbeitslosenzahlen (1992 bis 1994) zeigen sich diese Saisoneinflüsse, und zwar stets zu bestimmten Jahreszeiten mit der gleichen Tendenz. Dies sollte im einzelnen diskutiert und mit Beispielen belegt werden.

Für unsere weiteren Überlegungen ist von zentraler Bedeutung: Will man aus den monatlichen Arbeitslosenzahlen die wirkliche langfristige Entwicklung der Arbeitslosigkeit herausarbeiten, also den Trend aufzeigen, dann müssen die Saisoneinflüsse in geeigneter Weise herausgefiltert werden. Ein solches Verfahren heißt Saisonbereinigung, und wir stellen es im nächsten Kapitel vor.

2.4 Die Erweiterung des exponentiellen Glättens

Vorüberlegungen

Wir wollen nunmehr die saisonalen Einflüsse beim Glätten der monatlichen Arbeitslosenzahlen herausfiltern, und zwar mit dem Ziel, den langfristigen Trend herauszuarbeiten.
Im Unterrichtsgespräch sollte vorab geklärt werden:

1. Eine Durchmusterung der Tab.3 und die Analyse der Abbildungen 3 bis 5 verdeutlicht, daß der Glättungsparameter α möglichst klein gewählt werden sollte (Glättungseffekt!).
2. Wir brauchen Meßzahlen für die saisonalen Einflüsse.
3. Vor allem aber ist eine Idee nötig, wie das Filterungsverfahren vor sich gehen soll.

In dieser Situation hilft folgende Betrachtung weiter:
Wir stellen uns vor, wir hätten die langfristige Entwicklung der Arbeitslosenzahlen durch die ideale Trendlinie dargestellt - sehr wohl wissend, daß wir ein Verfahren, die ideale Trendlinie zu bestimmen, noch suchen. Danach versuchen wir, Wege zu dieser Ideallinie zu finden.

Anmerkung: Diese Art zu forschen, wird in der Schulgeometrie oft als die "Methode der Probefigur" bezeichnet.

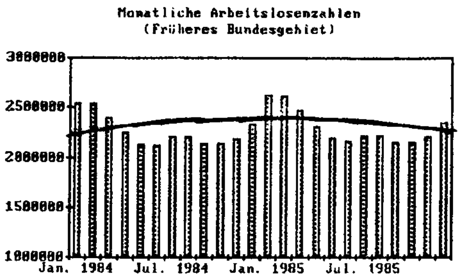

Abb.6: Monatliche Arbeitslosenzahlen und die angenommene ideale Trendlinie

In Abb.6 sind die monatlichen Arbeitslosenzahlen für die Zeit von Jan. 1984 bis Dez. 1985 wie üblich durch "Stäbe" dargestellt. Außerdem ist die angenommene Trendlinie eingezeichnet. In einer solchen Probefigur zeichnet man natürlich wichtige Linien so ein, wie sie in etwa zu erwarten sind. Als Anhaltspunkt für unsere Kurve dienten uns die Jahresdurchschnitte der Arbeitslosenzahlen für 1984 und 1985.

Es liegt nahe, und die Schülerinnen und Schüler werden dies beim Auswerten der Abb.6 auch vorschlagen, die Differenz $y_t - g_t$ als Meßzahl für den saisonalen Einfluß zu definieren.

Es soll gelten:
Meßzahl für den saisonalen Einfluß: $s_t = y_t - g_t$ [3]

Von August bis November sind die s_t - Werte beispielsweise in allen ausgewählten Jahren negativ, während sie im Januar und Februar positiv sind. Dies entspricht der Tatsache, daß die Arbeitslosenzahlen jedes Jahr ab Spätsommer zurückgehen und im Winter sprunghaft ansteigen.

Näherungen für die ideale Trendlinie

Da wir beim derzeitigen Stand des Unterrichtsganges für Zeitreihen mit auffallenden jahreszeitlichen Schwankungen noch kein Verfahren zur Bestimmung der idealen Trendlinie bzw. zur Berechnung der Trendwerte kennen, das auf dem exponentiellen Glätten aufbaut, ersetzen wir die gesuchte ideale Trendlinie zunächst durch eine möglichst gute Näherungskurve.
Dazu könnten Schülerinnen und Schüler folgende Wege anbieten:

1.Weg: Man wählt als Näherung für die ideale Trendlinie eine Treppenkurve, die durch die Werte der Jahresdurchschnitte festgelegt ist (Abb.7).

2.Weg: Man wählt aus den bereits durch einfaches exponentielles Glätten der monatlichen Arbeitslosenzahlen bestimmten Kurven diejenige mit den kleinsten jahreszeitlichen Schwankungen aus. Das ist eindeutig die Kurve mit dem Glättungsparameter $\alpha = 0{,}2$ (Abb.8).

An ihr kann man die langfristige Entwicklung noch am ehesten erkennen. Man könnte noch eine exponentielle Glättung mit $\alpha = 0{,}1$ erwägen, wird jedoch bei der Realisierung finden, daß die zugehörige Glättungslinie etwa ab 1990 der tatsächlichen Entwicklung zu sehr hinterherhinkt.

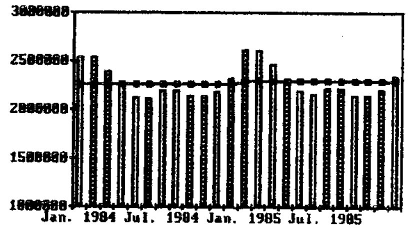

Abb.7: Treppenkurve der Jahresdurchschnitte als Näherung für die Trendkurve

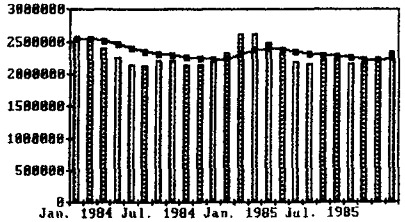

Abb.8: Glättungslinie mit $\alpha = 0{,}2$ als Näherung für Trendkurve

In Tab.4 sind die auf diesen Wegen bestimmten Näherungswerte für die Meßzahlen s_t zusammemgestellt. Dabei ist berücksichtigt, daß beim 2.Weg solche Näherungswerte für st frühestens für 1985 zu bestimmen sind. Anfangs weicht die Glättungslinie zu sehr von der Trendlinie ab. Dies sollte diskutiert werden.

Vergleicht man die nach Weg 1 und Weg 2 bestimmten Näherungswerte für s_t aus Tab.4, so stellt man fest, daß beide Reihen eine ziemlich gleiche Saisonfigur ergeben (Abb.9). Es besteht also die Chance, auch im Zusammenhang mit dem exponentiellen Glätten eine Saisonbereinigung durchzuführen.

Tab.4: Näherungswerte für die Meßzahlen der saisonalen Einflüsse

			1.Weg		2.Weg	
Monat	t	Monatliche Arbeitslosenzahl y_t	Jahres-durchschn. d_t	Näherungswert für s_t $y_t - d_t$	Glatter Wert für $\alpha = 0{,}2$	Näherungswert für s_t $y_t - g_t$
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Jan. 1984	1	2539000	2266000	273000		
Feb. 1984	2	2537000	2266000	271000		
Mär. 1984	3	2393000	2266000	127000		
Apr. 1984	4	2253000	2266000	- 13000		
Mai 1984	5	2133000	2266000	- 133000		
Jun. 1984	6	2113000	2266000	- 153000		
Jul. 1984	7	2202000	2266000	- 64000		
Aug. 1984	8	2202000	2266000	- 64000		
Sep. 1984	9	2143000	2266000	- 123000		
Okt. 1984	10	2145000	2266000	- 121000		
Nov. 1984	11	2189000	2266000	- 77000	(aus Tab.3)	
Dez. 1984	12	2325000	2266000	59000		
Jan. 1985	13	2619000	2304000	315000	2321082	297918
Feb. 1985	14	2611000	2304000	307000	2379066	231934
Mär. 1985	15	2474000	2304000	170000	2398053	75947
Apr. 1985	16	2305000	2304000	1000	2379442	- 74442
Mai 1985	17	2193000	2304000	- 111000	2342154	- 149154
Juni 1985	18	2160000	2304000	- 144000	2305723	- 145723
Juli 1985	19	2221000	2304000	- 83000	2288778	- 67778
Aug. 1985	20	2217000	2304000	- 87000	2274423	- 57423
Sep. 1985	21	2152000	2304000	- 152000	2249938	- 97938
Okt. 1985	22	2149000	2304000	- 155000	2229751	- 80751
Nov. 1985	23	2211000	2304000	- 93000	2226000	- 15000
Dez. 1985	24	2347000	2304000	43000	2250200	96800

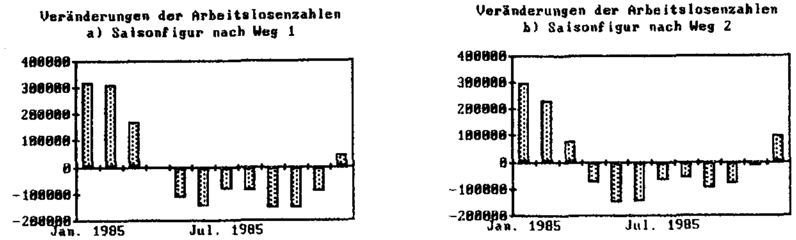

Abb.9: Saisonfiguren für 1985

2.4 Die Saisonbereinigung - Erweiterung des exponentiellen Glättens

Wir entscheiden uns für eine Saisonbereinigung nach Weg 2, weil dort konsequent mit den Monatswerten gearbeitet wird. Der entscheidende Schritt ist das Herausfiltern der Meßzahlen für die saisonalen Einflüsse während des Glättungsprozesses. Ziel dieses Vorhabens ist es, glatte Werte zu erhalten, die den Trend hinreichend gut beschreiben.
Wir zeigen dies zunächst am Beispiel der entsprechenden Werte für den Monat Januar 1986. Gegeben sind:

Monatliche Arbeitslosenzahl
für Jan.1986 (Ursprungswert): y(Jan.1986) = 2590000 (nach Tab.1)

Glättungswert für Dez.1985
(Glättungsparameter $\alpha = 0{,}2$): g(Dez.1985) = 2250200 (nach Tab.4)

Meßzahl für den saisonalen
Einfluß aus dem Vorjahreswert: s(Jan.1985) = s_{t-12} = 297918 (nach Tab.4)

In Analogie zum einfachen exponentiellen Glätten bestimmen wir jetzt den saisonbereinigten glatten Wert für Jan.1986 wie folgt:

$$g(\text{Jan.}1986) = 0{,}2\cdot(\underbrace{y(\text{Jan.}1986) - s(\text{Jan.}1985)}_{\uparrow}) + 0{,}8\cdot g(\text{Dez.}1985)$$

Aktuelle Arbeitslosenzahl vermindert um die
Meßzahl des Saisoneinflusses im entspr. Vorjahresmonat

$$g(\text{Jan.}1986) = 0{,}2\cdot(2590000 - 297918) + 0{,}8\cdot 2250200 = 2258576$$

Allgemein gilt:

$$g_t = \alpha\cdot(y_t - s_{t-12}) + (1-\alpha)\cdot g_{t-1} \qquad [4]$$

Formel des erweiterten exponentiellen Glättens von Zeitreihen mit 12monatiger Saisonperiode

Auf diese Weise bestimmen wir saisonbereinigte glatte Werte und damit hinreichend gute Schätzwerte für den langfristigen Trend.

Bei der praktischen Durchführung der Saisonbereinigung mit Hilfe einer Tabellenkalkulation kann man sich Arbeit sparen, wenn man die Tab. 3 noch einmal lädt, die Spalten (1) bis (4) beibehält, die Spalten (5) ff. löscht und eine neue Spalte (5) für die s_t-Werte einfügt. Damit hat die Tabelle zur Bestimmung der saisonbereinigten glatten Werte folgen-des Aussehen:

Monat		Monatliche Arbeitslosenzahl	Glatter Wert bei $\alpha = 0{,}2$	Veränderung durch saison. Einfluß	
	t	y_t	g_t	s_t	
(1)	(2)	(3)	(4)	(5)	
Jan. 1984	1	2539000	2539000		
⋮				bleibt leer!	
Dez. 1984	12	2325000	2246603		
Jan. 1985	13	2619000	2321082	297918	
⋮					$y_t - g_t$
Dez. 1985	24	2347000	2250200	96800	
Jan. 1986	25	2590000	....	....	
			Berechnung nach Gleichung [4]	Siehe unten!	

An dieser Stelle des Unterrichtsganges könnte der Einwand kommen, daß es im Sinne der Glättungsmethode sei, auch die Meßzahlen s_t ab 1986 zu glätten. Andernfalls sollte man einen Denkanstoß geben, diesen Gedanken zu verfolgen.
Außer der soeben erwähnten Differenz $y_t - g_t$ müßte eine Gleichung zur Glättung der s_t-Werte noch einen "Glättungssummanden" enthalten. Ferner brauchen wir einen zweiten Glättungsparameter. Nennen wir ihn γ.
Als Ergebnis der Überlegungen soll die Gleichung

$$s_t = \gamma\cdot(y_t - g_t) + (1-\gamma)\cdot s_{t-12} \qquad [5]$$

erarbeitet werden.
Hinweis: In der Fachliteratur wird der Glättungsparameter für den saisonalen Einfluß in der Regel mit gamma bezeichnet.

Bei einer Zeitreihe mit deutlich ausgeprägten saisonalen Schwankungen und einer Periodelänge von 12 Monaten erfolgt die erweiterte exponentielle Glättung demnach mit Hilfe des folgenden Gleichungssystems:

$$g_t = \alpha \cdot (y_t - s_{t-12}) + (1-\alpha) \cdot g_{t-1} \qquad t > 24$$

$$s_t = \gamma \cdot (y_t - g_t) + (1-\gamma) \cdot s_{t-12}$$

Seine Anwendung erfordert, wie es oben dargestellt wurde, einen Vorlauf von 2 Perioden mit einfachem exponentiellen Glätten. Bei einer langen Zeitreihe kann man den Vorlauf auch auf 3 bis 4 Jahre ausdehnen. Dies verbessert die Glättung. Wir schlagen vor, auch die saisonbereinigt glatten Werte mit g_t zu bezeichnen. Man könnte dafür auch eine andere Notation einführen.
In der angegebenen weiterführenden Literatur findet man Verfahren, wie man optimale Werte für die Glättungskonstanten rechnerisch bestimmt. Dort gibt es auch Hinweise, wie man die Güte einer Glättung bewertet. Dies alles sind jedoch Spezialverfahren, die dem Studium vorbehalten bleiben sollten.

Aus dem Gedanken, das einfache exponentielle Glätten in mehreren Richtungen zu erweitern und auf spezielle Fallunterscheidungen auszudehnen, entstanden vor etwa 30 Jahren im Bereich der Wirtschaftswissenschaften einsatzfähige Methoden, die heute in Form von Spezialprogrammen (fertiger Software) in der Praxis vielfältig angewendet werden. Die hier vorgestellte Glättung mit dem Herausfiltern saisonaler Einflüsse ist übrigens ein Sonderfall des in den Wirtschaftswissenschaften sehr beachteten "HOLT-WINTERS-Verfahren".

Mit Hilfe einer Arbeitstabelle (siehe Tab.5) wird die Saisonbereinigung und die Berechnung der saisonbereinigt glatten Werte durchgeführt.

Tab.5: Saisonbereinigt glatte Werte
Beide Glättungsparameter = 0,2; Zeitraum 1986 bis 1994

Monat	t	Monatliche Arbeitslosenzahl y_t	saisonbereinigt glatter Wert g_t	Veränderung durch saison. Einfluß s_t
(1)	(2)	(3)	(4)	(5)
⋮				
Jan. 1986	25	2590000	2258577	304619
Feb. 1986	26	2593000	2279075	248333
Mär. 1986	27	2448000	2297670	90824
Apr. 1986	28	2230000	2299025	- 73359
Mai 1986	29	2122000	2293450	- 153613
Jun. 1986	30	2078000	2279505	- 156879
Jul. 1986	31	2132000	2263560	- 80535
Aug. 1986	32	2120000	2246332	- 71205
Sep. 1986	33	2046000	2225853	- 114321
Okt. 1986	34	2026000	2202033	- 99807
Nov. 1986	35	2068000	2178226	- 34046
Dez. 1986	36	2218000	2166821	87676
⋮				
Jan. 1994	121	2736000	2397715	253000
Feb. 1994	122	2742000	2422708	239314
Mär. 1994	123	2640000	2445330	122280
Apr. 1994	124	2590000	2475395	18396
Mai 1994	125	2506000	2499335	- 69940
Jun. 1994	126	2478000	2515755	- 90300
Jul. 1994	217	2570000	2530544	- 7868
Aug. 1994	128	22531000	2539406	- 36765
Sep. 1994	129	2452000	2542097	- 98707
Okt. 1994	130	2446000	2541448	- 93371
Nov. 1994	131	2450000	2533662	- 58748
Dez. 1994	132	2545000	2523156	55465

In Abb.10 ist die zugehörige Glättungslinie dargestellt. Damit sind wir am Ziel unserer Arbeit angelangt. Aus diesem Grunde sind die Stäbe für die Originalwerte der Arbeitslosenzahlen in Abb.10 abweichend von den vorangehenden Abbildungen einmal ohne den vorher gewählten "Sockelbetrag" in "voller Länge" dargestellt. Es bleibt die Frage, ob die gewählten Glättungsparameter optimal sind. Wenn an dem Thema noch weiter gearbeitet werden soll, kann man zunächst den 2.Glättungsparameter γ variieren. Dazu müßte die Arbeitstabelle erweitert werden. Abb.11 zeigt, daß man saisonalen Einflüsse noch stärker herausfiltert, wenn man etwas größere γ-Werte wählt.

In Abb. 12 sind die saisonal bedingten Veränderungen der Arbeitslosenzahlen von Monat zu Monat dargestellt. Man sieht, daß die Saisonfiguren stets einige Jahre lang ziemlich gleich bleiben, sich mit der Zeit jedoch auch ändern.

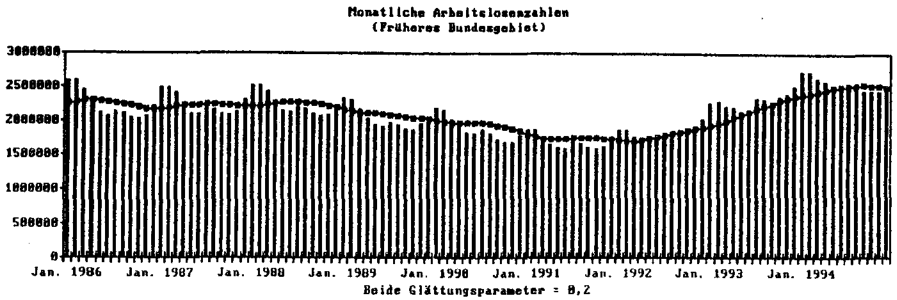

Abb.10: Die Kurve der saisonbereinigt glatten Arbeitslosenzahlen

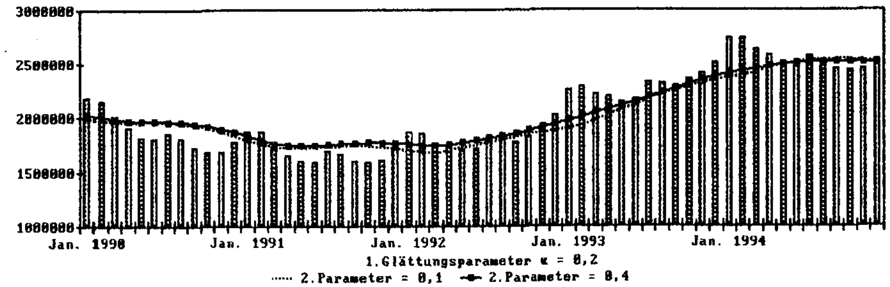

Abb.11: Glättungslinien für verschiedene Glättungsparameter γ (1990-1994)

Offensichtlich ist das exponentielle Glätten mit Saisonbereinigung etwas anfällig gegen starke Trendänderungen. Dies sieht man daran, daß die negativen Werte in Abb.12 in den Jahren 1990 bis 1992 dominieren, während ab 1993 die starken Zunahmen ins Auge fallen. Die Saisonmuster sind in den Jahren 1990 bis 1992 gewissermaßen nach unten verschoben und in den nachfolgenden Jahren nach oben.

Außerdem bleiben die Veränderungen in den Saisonmustern etwas hinter der wirklichen Entwicklung zurück und wirken auch jeweils noch etwas nach. Hier wird eine Eigenschaft des exponentiellen Glättens deutlich, die man kennen muß, damit es bei der Anwendung dieses mathematischen Modells nicht zu Fehlinterpretationen kommt.

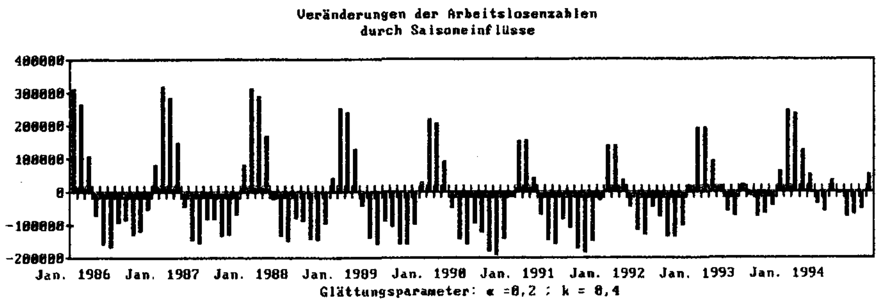

Abb.12: Die beim erweiterten exponentiellen Glätten entstandenen Saisonmuster

Die Bestimmungsgründe für die auffallenden Trendänderungen und die Veränderungen in den Saisonmustern suchen die Schülerinnen und Schüler zu recht in den Ereignissen der jüngsten deutschen Geschichte und in sich ändernden Strukturen in der deutschen Wirtschaft und damit auch am Arbeitsmarkt.

2.5 Abschlußdiskussion

Die Glättungslinien in den Abbildungen 10 und 11 werden als Näherungen für die Trendkurve interpretiert. Die Schülerinnen und Schüler können den gesamten Verlauf beschreiben, die Frage nach dem Minimum der Arbeitslosigkeit der

Wir haben uns hier schwerpunktmäßig mit dem ersten Hauptanliegen einer Zeitreihenanalyse befaßt, die Entwicklung der untersuchten Daten im fraglichen Zeitraum mit mathematischen Mitteln möglichst genau zu beschreiben. Das dargestellte erweiterte exponentielle Glätten liefert saisonbereigt glatte Werte, die von Autoren fachwissenschaftlicher Werke häufig auch als Meßzahlen bzw. als Schätzwerte für das "Niveau der Zeitreihe" bezeichnet werden.

Während des Unterrichtsganges war ein tiefer Einblick in die Problematik der Saisonbereinigung möglich. Das gewählte Verfahren lieferte brauchbare Näherungen für eine Trendbeschreibung. Daß es auch kleine Schwächen aufweist, hat für den Lernprozeß Vorteile: Die Schülerinnen und Schüler erleben hautnah, daß die Wahl eines mathematischen Modells in jedem Fall bedeutet, über die Stärken und Schwächen und den Anwendungsbereich eines Modells nachzudenken.

Es bleibt noch das zweite Hauptanliegen einer Zeitreihenanalyse, eine über das Ende der Zeitreihe hinausgehende Prognose zu ermöglichen. Auch dieses Problem ist auf der Basis exponentiellen Glättens zu lösen. In den Wirtschaftswissenschaften wurden dazu in den letzten drei Jahrzehnten unterschiedliche Methoden entwickelt und zur Anwendungsreife gebracht. Elementare Prognosemethoden werden wir in einem späteren Aufsatz - methodisch-didaktisch für die Schule aufbereitet - vorstellen.

Literatur

Siehe Teil 1 in "Stochastik in der Schule" 14 (1994) Heft 3

Günter Nordmeier
Platanenallee 9
D-49152 Bad Essen

aradoxien der beschreibenden Statistik

Von JÖRG MEYER, Hameln

wird eine größere Reihe von Sachverhalten vorgestellt, die der ik zugerechnet werden können und die dem gesunden lox oder doch sehr verblüffend erscheinen, jedenfalls auf den ersten spiele sind oft sehr einfache Situationen, die aber auf praktisch wichtige Fälle übertragen werden können.

Manchmal sehen wir etwas, geben aber nicht acht. Seneca.

Die Genese der gesamten elementaren Stochastik wird geradezu beherrscht von Auseinandersetzungen über Paradoxien. Winter.

0. Einleitung

Es gibt verschiedene Klassen von Paradoxien (Ansätze einer Klassifikation liefert Poundstone 1991, S. 16-19). Die hier behandelten sind solche, die dem gesunden Menschenverstand widersprechen. Bei hinreichend naher sachgemäßer Betrachtung entpuppen sie sich als Banalitäten.

Dies soll in diesem Aufsatz anhand einiger einfacher Beispiele aus der beschreibenden Statistik illustriert werden. Dabei wird oft der umgekehrte Weg durchlaufen: Banalitäten werden schief beleuchtet, so daß sie paradox erscheinen.

Die Anwendungen sind mitunter überraschend. Sie sind ausgesprochen oberflächlich und zeigen so etwas von der Buntheit der Welt.

Zu den hier behandelten Paradoxien gehören auch die von Blyth und Simpson, die trotz ihrer Einfachheit erst 1972 bzw. 1951 beschrieben wurden. Sie werden in gestufter Form innerhalb eines gemeinsamen Kontextes dargestellt.

Selbstverständlich ist es nicht sinnvoll, alle Kapitel dieses Aufsatzes geballt und zusammenhängend im Unterricht zu behandeln. Vielmehr möge man sie als *Materialsammlung* ansehen, aus der dosiert und kontextabhängig Beispiele zur Illu-

Stochastik in der Schule **15**(1995), Nr. 2, S. 27-50