Einige Bemerkungen zur mechanischen Ausgleichung statistischer Reihen.

Von Dr. **O. Schenker**, Beamter des Eidgenössischen statistischen Bureaus, Interlaken.

Wenn man einen Kollektivgegenstand (d. h. eine Vielheit gleichartiger Dinge) nach einem veränderlichen Merkmal statistisch ordnet (z. B. die Bevölkerung eines Landes nach dem Alter), so erhält man eine Zahlenreihe, die man passend auch als statistische Reihe zu bezeichnen pflegt. Der Anschaulichkeit halber denken wir uns das veränderliche Merkmal in der zugrunde liegenden Einheit (als die wir die kleinste noch zu berücksichtigende Grösse, z. B. ein Jahr, wählen wollen) auf der Abszissenachse eines rechtwinkligen Koordinatensystems als Abszisse x (oder Argument x) und die zugehörige Anzahl von Gliedern des Kollektivgegenstandes zwischen den Abszissen $x - \frac{1}{2}$ und $x + \frac{1}{2}$ liegend oder anders ausgedrückt, welche das veränderliche Merkmal von der Grösse $x - \frac{1}{2}$ weg bis zur Grösse $x + \frac{1}{2}$ enthalten, als Ordinate y_x aufgetragen. Diese Ordinaten y_0, y_1, $y_2 \ldots y_x$, $y_{x+1} \ldots$ werden im allgemeinen keine regelmässig verlaufende Reihe bilden. Der Grund hierfür ist zu suchen in dem Vorhandensein von zufälligen und systematischen Fehlern, sowie in dem Mangel einer genügend grossen Zahl von Beobachtungen. Wenn wir z. B. die Bevölkerung eines Landes nach Altersjahren ordnen wollen, so werden wir als Abszissen die Alter $0, 1, 2, 3, \ldots x, x+1, \ldots$ und als Ordinaten die Zahl der Personen, welche bzw. $0 - \frac{1}{2}$, $\frac{1}{2} - \frac{3}{2}$, $\frac{3}{2} - \frac{5}{2}$, $\frac{5}{2} - \frac{7}{2}$, $\ldots (x - \frac{1}{2}) - (x + \frac{1}{2})$, $(x + \frac{1}{2}) - (x + \frac{3}{2}) \ldots$ Jahre alt sind, auftragen. Stützt sich diese Darstellung, der Zeitersparnis wegen, bloss auf die Geburtsjahre, so ist sie bloss genau, wenn die Volkszählungen am Jahresende stattfinden; finden sie aber am 1. Dezember statt, so müsste man als Abszissen eigentlich die Alter $0, \frac{11}{12}, \frac{23}{12}, \frac{35}{12}, \ldots x - \frac{1}{12}, x + \frac{11}{12}, \ldots$ und als Ordinaten die Zahl der Personen, welche bzw. $0 - \frac{5}{12}, \frac{5}{12} - \frac{17}{12}, \frac{17}{12} - \frac{29}{12}, \frac{29}{12} - \frac{41}{12}, \ldots \left(x - \frac{7}{12}\right) - \left(x + \frac{5}{12}\right), \left(x + \frac{5}{12}\right) - \left(x + \frac{17}{12}\right), \ldots$ Jahre alt sind, auftragen. Es ist übrigens bei der geometrischen Darstellung statistischer Reihen logischer und im Interesse der Kontinuität erwünscht, die Zahl der Personen einer Altersklasse nicht durch äquidistante Ordinaten, sondern durch die Rechtecke über den zugehörigen Strecken der Abszissenachse darzustellen, z. B. über den Strecken $0 - \frac{1}{2}$, $\frac{1}{2} - \frac{3}{2}$, $\frac{3}{2} - \frac{5}{2}$, $\frac{5}{2} - \frac{7}{2}$, $\ldots (x - \frac{1}{2}) - (x + \frac{1}{2})$, $(x + \frac{1}{2}) - (x + \frac{3}{2}) \ldots$. Die Punkte der Abszissenachse, in welchen diese Rechtecke zusammenstossen, pflegt man als Wechselpunkte zu bezeichnen, weil sie den Übergang von einem Rechteck zum andern bezeichnen. Der Mensch als ein Teil der Natur liebt bekanntlich das Gesetzmässige und Symmetrische, und darum muss für den Statistiker auch der Wunsch entstehen, die Unregelmässigkeiten im Verlaufe einer statistischen Reihe zu entfernen. Auf einem Gebiete, welches Anspruch auf eine Wissenschaft, oder doch auf eine wissenschaftliche Methode machen kann, darf man sich selbstverständlich nicht von blossen Wünschen leiten lassen, oder anders ausgedrückt, das Gewünschte muss auch den Tatsachen entsprechen; dies trifft nun hier zu, denn es ist ja allgemein bekannt, dass statistische Reihen um so regelmässiger verlaufen, je grösser das zugrunde gelegte Beobachtungsmaterial ist; das umfangreichere Beobachtungsmaterial liefert bekanntlich auch zuverlässigere Resultate, woraus folgt: mit steigender Zahl der Beobachtungen nimmt eine statistische Reihe einen regelmässigeren und damit zuverlässigeren Verlauf. Dass eine von allen Fehlern befreite statistische Reihe regelmässig verlaufen muss, kann a priori wie folgt eingesehen werden: Sei x die Abszisse eines Wechselpunktes, $x + \Delta x$ diejenige des nächstfolgenden, $x - \Delta x$ diejenige des vorangehenden, dann werden die im Punkte x zusammenstossenden Rechtecke nur wenig voneinander verschieden sein, falls Δx klein genug ist, sie werden gleich sein, wenn $\Delta x = 0$ ist, denn der Ursachenkomplex, welcher die in x zusammenstossenden Rechtecke bestimmt, erscheint offenbar als Funktion der Abszisse x und Δx; ist er für das eine Rechteck dargestellt durch $f(x, + \Delta)$, so erscheint er für das andere Rechteck in der Form $f(x, - \Delta)$. Die beiden Komplexe sind also nur wenig verschieden und damit ihre Wirkungen, d. h. die im Punkte x zusammenstossenden Rechtecke selbst. Diese Schlussfolgerung hat zur notwendigen Voraussetzung, dass die Ursachen auf völlig übereinstimmende Elemente einwirken. In der Praxis wird dies sozusagen nie der Fall sein. Es wird z. B. nicht *zwei* Menschen geben, die in jeder Beziehung völlig gleichwertig sind. Die menschliche Unvollkommenheit macht sich eben auf dem Gebiete der Statistik geltend wie auf allen andern Gebieten menschlicher Tätigkeit, und so muss man sich mit Approximationen begnügen. Die Elemente, welche Gegenstand der Statistik sind, werden im allgemeinen unendlich viele Eigenschaften besitzen, sie werden aber bloss in mehr oder weniger grossen Gruppen in gewissen Eigenschaften übereinstimmen, z. B. im Alter, im Einkommen, in der Heimat. Denkt man sich diese Elemente (z. B. die Bevölkerung eines Landes) nach einer solchen Eigenschaft, wie eingangs angegeben, statistisch geordnet, so können zwei Fälle eintreten: 1. Die übrigen Merkmale dieser Elemente treten nach dem Gesetz des Zufalls auf, d. h. sie sind homogen gemischt. 2. Ein weiteres Merkmal oder mehrere weitere Merkmale weisen eine von der Abszisse abhängige Häufigkeit auf. Im ersten Fall wird bei genügender Beobachtungszahl die statistische Reihe sich so verhalten, als ob die zur Abszisse x gehörigen Elemente in allen Merkmalen völlig miteinander übereinstimmen würden, indem die übrigen homogen gemischten Merkmale sich neutralisieren, d. h. es bilden sich aus ihnen Mittelwerte, welche für alle Abszissen konstant sind. Im zweiten Fall gilt dasselbe, nur zeigen hier zwei oder mehr Merkmale eine von der Abszisse x abhängige Häufigkeit. A priori kommt man also unter Zuhilfenahme des Begriffs des Zufalls ebenfalls zu dem Satz: eine statistische Reihe nimmt mit steigender Zahl der Beobachtungen einen regelmässigeren und damit zuverlässigeren Verlauf. Denkt man sich die benachbarten Wechselpunkte einer statistischen Reihe sehr nahe zusammengerückt, so tun es auch die gemeinsamen Seiten zweier benachbarter Rechtecke, und zwar beschreiben ihre Gipfelpunkte eine Kurve, die man als Verteilungskurve zu bezeichnen pflegt; sie zeigt, wie sich die einzelnen Abszissen oder Argumente nach ihrer Häufigkeit verteilen; man könnte dafür auch die Bezeichnung Häufigkeitskurve gebrauchen.

Was wir hier vorausgeschickt haben, dient als notwendige Voraussetzung der mechanischen Ausgleichung, indem sie gebunden ist an das Vorhandensein regelmässig verlaufender und damit genügend genauer Verteilungskurven.

Die Fehler, welche den Gliedern einer statistischen Reihe anhaften, können momentan wirkenden Ursachen entspringen (zufällige Fehler), sie können auf konstant wirkende Ursachen zurückführen (systematische Fehler), und sie können ihren Grund in der ungenügenden Zahl von Beobachtungen haben. Die Glieder einer statistischen Reihe werden daher bald über die fehlerfreie Verteilungskurve hinausragen, bald werden sie dieselbe nicht erreichen, immer unter der Voraussetzung, dass die benachbarten Wechselpunkte nahe genug liegen. Bekanntlich tritt der Einfluss der zufälligen Fehler mit wachsender Beobachtungszahl zurück, während dasselbe von den systematischen Fehlern wenigstens nicht in gleichem Masse zu erwarten ist. Die zufälligen Fehler sind übrigens auch dadurch charakterisiert, dass sie gleich leicht positiv wie negativ eintreten können. Zu den systematischen Fehlern gehören z. B. die Angaben in runden Zahlen. Natürlich muss der Statistiker auf Fehler dieser Art ein besonderes Augenmerk richten; am besten ist es, wenn sie vermieden werden können. Es gibt statistische Reihen, deren Verlauf eine typische Regelmässigkeit aufweist. Wir brauchen nur an die Fehlerkurve von Gauss zu erinnern; im weitern können wir verweisen auf die Absterbeordnung; die Verteilungskurve hat hier allerdings nicht den gewöhnlichen Verlauf, indem nicht beide äussersten Ordinaten null sind; wenn man aber die Ordinaten mit den zugehörigen Altern multipliziert, so erhält man eine Verteilungskurve dieser Form.

Die mechanische Ausgleichung hat nun den Zweck, die Unregelmässigkeiten einer statistischen Reihe zu entfernen und gleichzeitig genauere Resultate zu erhalten. Bevor wir untersuchen können, inwieweit dieser Zweck auch erreicht wird, ist eine Definition notwendig:

Die Glieder $y_0, y_1, y_2, \ldots y_n, y_{n+1}, \ldots$ einer statistischen Reihe mechanisch ausgleichen, heisst, dieselben durch Mittelwerte aus dem auszugleichenden Glied und den benachbarten Gliedern ersetzen, wobei nach einem festen Gesetz verfahren wird; es ist üblich, ein Gesetz zu verwenden, das die Symmetrie in bezug auf das auszugleichende Glied wahrt, z. B. kann man y_n durch

$$\frac{y_{n-1} + y_n + y_{n+1}}{3}$$

ersetzen oder durch

$$\frac{y_{n-2} + y_{n-1} + y_n + y_{n+1} + y_{n+2}}{5}$$

oder durch die entsprechenden arithmetischen Mittel mit ungleichen Gewichten usw. Auf den ersten Blick mag die Wahl solcher Ausgleichungsformeln mehr oder weniger willkürlich erscheinen. Dem ist aber nur zum Teil so. Auf alle Fälle wird durch solche Mittelwerte der erste Ausgleichungszweck, ein regelmässigerer Verlauf der statistischen Reihe erreicht; hierüber braucht wohl keine nähere Erörterung stattzufinden. Wie gestaltet sich aber die Wirkung der Ausgleichung in bezug auf die Genauigkeit der Reihe? Die Ausgleichung muss diese Genauigkeit erhöhen, wenn sie ihren Zweck erreichen soll. Zur Prüfung dieser zweiten Frage brauchen wir ein Mass, um die Genauigkeit der Reihe messen zu können; ein solches erhalten wir offenbar, wenn die sämtlichen Gliedern der Reihe anhaftenden Fehler ohne Rücksicht auf das Vorzeichen addiert werden, oder anders ausgedrückt, wir wählen die Summe der absoluten Fehler sämtlicher Glieder als Mass für die Genauigkeit der statistischen Reihe. Man wird einwenden, die Fehler der einzelnen Glieder seien unbekannt, dies trifft zu, man kann aber die fehlerfreie Verteilungskurve, sowie die Verteilung und Grösse der Fehler als bekannt voraussetzen, um die Wirkung durch die mechanische Ausgleichung genau beurteilen zu können. Die mechanische Ausgleichung denken wir uns in zwei Etappen ausgeführt, indem wir ausgleichen: 1. die fehlerfreien Reihenglieder, für die wir die entsprechenden Ordinaten der Verteilungskurve setzen können; 2. die Fehler der Reihenglieder selbst. Ist die Verteilungskurve eine Gerade, so ergibt die Anwendung einer Ausgleichungsformel von der erwähnten üblichen Gestalt (wir werden in Zukunft diese symmetrische Gestalt benützen) auf die Ordinaten der Verteilungskurve keine Veränderung der letztern. Das folgende Täfelchen gibt hierfür ein Beispiel.

Ausgleichungsformel:

$$\frac{y_{n-1} + y_n + y_{n+1}}{3}$$

Abszisse	Ordinate	Ausgeglichene Ordinate
0	103	103
1	105	105
2	107	107
3	109	109
4	111	111
5	113	113
6	115	115

Die erste und letzte Ordinate muss man unausgeglichen lassen, weil hier der Anschluss an das vorausgehende, bzw. nachfolgende Glied fehlt. Für Verteilungskurven, die von der Geraden abweichen, aber nur schwach gekrümmt sind, wird näherungsweise dasselbe gelten, z. B.:

Ausgleichungsformel:

$$\frac{y_{n-1} + y_n + y_{n+1}}{3}$$

Abszisse der Verteilungskurve	Ordinate der Verteilungskurve	Ausgeglichene Ordinate der Verteilungskurve	Abszisse der Verteilungskurve	Ordinate der Verteilungskurve	Ausgeglichene Ordinate der Verteilungskurve
0	103,5	103,5	4	111	111
1	104,5	105	5	112,5	112,8
2	107	107	6	115	115
3	109,5	109,2			

Wenn die Verteilungskurve stark von der Geraden abweicht, so ist die Wahl des ausgleichenden Mittelwertes nicht mehr freigestellt, wie das folgende Beispiel zeigt:

Ausgleichungsformel:

$$\frac{y_{n-1} + y_n + y_{n+1}}{3}.$$

Abszisse der Verteilungskurve	Ordinate der Verteilungskurve	Ausgeglichene Ordinate der Verteilungskurve	Abszisse der Verteilungskurve	Ordinate der Verteilungskurve	Ausgeglichene Ordinate der Verteilungskurve
0	100	100	4	122	126
1	101	101,7	5	146	150
2	104	105	6	182	186
3	110	112	7	230	230

Wie man sieht, entstehen hier bei der Ausgleichung systematische Fehler, indem sämtliche ausgeglichenen Ordinaten zu gross sind. Dasselbe ist der Fall für jede Verteilungskurve, die über ein grösseres Stück konvex oder dann über ein grösseres Stück konkav gegen die Abszissenachse gekrümmt ist. Bei Wiederholung des Ausgleichungsprozesses an den ausgeglichenen Ordinaten treten die systematischen Fehler noch deutlicher hervor, im letzten Beispiel wie folgt:

Abszisse der Verteilungskurve	Ordinate der Verteilungskurve	Ausgeglichene Ordinate der Verteilungskurve	Abszisse der Verteilungskurve	Ordinate der Verteilungskurve	Ausgeglichene Ordinate der Verteilungskurve
0	100	100	4	126	129,3
1	101,7	102,2	5	150	154
2	105	106,2	6	186	188,6
3	112	114,3	7	230	230

Im folgenden Beispiel macht sich der Einfluss der Ausgleichung durch einen konstanten Fehler bemerkbar:

Ausgleichender Mittelwert:

$$\frac{y_{n-1} + y_n + y_{n+1}}{3}.$$

Abszisse der Verteilungskurve	Ordinate der Verteilungskurve	Ausgeglichene Ordinate der Verteilungskurve	Abszisse der Verteilungskurve	Ordinate der Verteilungskurve	Ausgeglichene Ordinate der Verteilungskurve
0	1	1	0	1	1
1	4	4,67	1	4,67	5,1
2	9	9,67	2	9,67	10,33
3	16	16,67	3	16,67	17,33
4	25	25,67	4	25,67	26,33
5	36	36,67	5	36,67	37,33
6	49	49,67	6	49,67	50,33
7	64	64,67	7	64,67	65,1
8	81	81	8	81	81

Wenn die Verteilungskurve stets konvex oder stets konkav gegen die Abszissenachse gekrümmt ist, so werden durch die Ausgleichung die ausgeglichenen Ordinaten sämtlich zu gross bzw. zu klein, und die neue Verteilungskurve teilt die Eigenschaft der ursprünglichen. Dies ist leicht einzusehen. Die Ausgleichung bewirkt also hier eine Abflachung der Verteilungskurve und damit ein falsches Bild derselben. Um diese Wirkung zu vermeiden, muss bei der Wahl des ausgleichenden Mittelwertes auf die Krümmung der Verteilungskurve tunlichst Rücksicht genommen werden. Die Literatur hierüber ist ziemlich reichhaltig [1]). Darum wollen wir uns begnügen, die Wirkung der mechanischen Ausgleichung auf die Fehler der Glieder einer statistischen Reihe etwas näher ins Auge zu fassen, und zwar vorzüglich mit Rücksicht auf das Vorzeichen dieser Fehler. Eine Aufeinanderfolge von lauter positiven bzw. negativen Fehlern wollen wir als eine Fehlergruppe bezeichnen. Man hat also Fehlergruppen mit 1 Fehler, mit 2, 3, 4, 5 ... Fehlern. Zunächst soll angenommen werden, die äussersten Glieder der statistischen Reihe können nicht ausgeglichen werden, die Fehler seien der absoluten Grösse nach für jedes Glied gegeben und positiv gleich leicht möglich wie negativ. Eine Statistik über die verschiedenen Fehlerkombinationen mit Rücksicht auf das Vorzeichen allein ergibt das folgende Resultat:

Gliederzahl der statistischen Reihe	Zahl der		Zahl der Fehlergruppen mit									
	pos.	neg.	1	2	3	4	5	6	7	8	9	10
	Fehler		Fehler	nur positiven oder nur negativen Fehlern								
1	1	0	1	—	—	—	—	—	—	—	—	—
1	0	1	1	—	—	—	—	—	—	—	—	—
2	2	0	0	1	—	—	—	—	—	—	—	—
2	1	1	4	0	—	—	—	—	—	—	—	—
2	0	2	0	1	—	—	—	—	—	—	—	—
3	3	0	0	0	1	—	—	—	—	—	—	—
3	2	1	5	2	0	—	—	—	—	—	—	—
3	1	2	5	2	0	—	—	—	—	—	—	—
3	0	3	0	0	1	—	—	—	—	—	—	—
4	4	0	0	0	0	1	—	—	—	—	—	—
4	3	1	6	2	2	0	—	—	—	—	—	—
4	2	2	12	6	0	0	—	—	—	—	—	—
4	1	3	6	2	2	0	—	—	—	—	—	—
4	0	4	0	0	0	1	—	—	—	—	—	—
5	5	0	0	0	0	0	1	—	—	—	—	—
5	4	1	7	2	2	2	0	—	—	—	—	—
5	3	2	21	10	3	0	0	—	—	—	—	—
5	2	3	21	10	3	0	0	—	—	—	—	—
5	1	4	7	2	2	2	0	—	—	—	—	—
5	0	5	0	0	0	0	1	—	—	—	—	—
6	6	0	0	0	0	0	0	1	—	—	—	—
6	5	1	8	2	2	2	2	0	—	—	—	—
6	4	2	32	14	6	3	0	0	—	—	—	—
6	3	3	48	24	8	0	0	0	—	—	—	—
6	2	4	32	14	6	3	0	0	—	—	—	—
6	1	5	8	2	2	2	2	0	—	—	—	—
6	0	6	0	0	0	0	0	1	—	—	—	—
7	7	0	0	0	0	0	0	0	1	—	—	—
7	6	1	9	2	2	2	2	2	0	—	—	—
7	5	2	45	18	9	6	3	0	0	—	—	—
7	4	3	90	44	17	4	0	0	0	—	—	—
7	3	4	90	44	17	4	0	0	0	—	—	—
7	2	5	45	18	9	6	3	0	0	—	—	—
7	1	6	9	2	2	2	2	2	0	—	—	—
7	0	7	0	0	0	0	0	0	1	—	—	—
8	8	0	0	0	0	0	0	0	0	1	—	—
8	7	1	10	2	2	2	2	2	2	0	—	—
8	6	2	60	22	12	9	6	3	0	0	—	—
8	5	3	150	70	30	12	4	0	0	0	—	—
8	4	4	200	100	40	10	0	0	0	0	—	—
8	3	5	150	70	30	12	4	0	0	0	—	—
8	2	6	60	22	12	9	6	3	0	0	—	—
8	1	7	10	2	2	2	2	2	2	0	—	—
8	0	8	0	0	0	0	0	0	0	1	—	—
9	9	0	0	0	0	0	0	0	0	0	1	—
9	8	1	11	2	2	2	2	2	2	2	0	—
9	7	2	77	26	15	12	9	6	3	0	0	—
9	6	3	231	102	47	24	12	4	0	0	0	—
9	5	4	385	190	80	26	5	0	0	0	0	—
9	4	5	385	190	80	26	5	0	0	0	0	—
9	3	6	231	102	47	24	12	4	0	0	0	—
9	2	7	77	26	15	12	9	6	3	0	0	—
9	1	8	11	2	2	2	2	2	2	2	0	—
9	0	9	0	0	0	0	0	0	0	0	1	—
10	10	0	0	0	0	0	0	0	0	0	0	1
10	9	1	12	2	2	2	2	2	2	2	2	0
10	8	2	96	30	18	15	12	9	6	3	0	0
10	7	3	336	137	70	40	24	12	4	0	0	0
10	6	4	672	325	140	57	20	5	0	0	0	0
10	5	5	840	420	180	60	12	0	0	0	0	0
10	4	6	672	325	140	57	20	5	0	0	0	0
10	3	7	336	137	70	40	24	12	4	0	0	0
10	2	8	96	30	18	15	12	9	6	3	0	0
10	1	9	12	2	2	2	2	2	2	2	2	0
10	0	10	0	0	0	0	0	0	0	0	0	1

[1]) Über verschiedene Typen von Verteilungskurven s. *A. Knibbs*, The Mathematical Theory of Population, Melbourne 1917 (Census of the Commonwealth of Australia, Appendix A, Vol. 1), S. 47 ff. — Über verschiedene Methoden der Ausgleichung s. *A. Knibbs*, op. c., S. 85 ff.; ferner *Westergaard*, Theorie der Statistik, 2. Ausgabe, dänisch, 1915, S. 265 ff.; ferner *E. Czuber*, Wahrscheinlichkeitsrechnung, Leipzig 1903, S. 403 ff. — Ein praktisches Beispiel für die mechanische Ausgleichung bietet die schweizerische Sterbetafel für die Jahre 1901 bis 1910 (Die Ergebnisse der eidgenössischen Volkszählung vom 1. Dez. 1910, II. Bd.), S. 43* ff.

Die Fortführung dieser Statistik auf rein induktivem Weg ist beschwerlich. Sie wird aber überflüssig, wenn die Gesetze, nach welchen die gegebenen Zahlenreihen fortschreiten, bekannt sind. Ein interessantes Gesetz ergibt sich bei der Bildung der Kolonnensummen, wie folgt:

Gliederzahl der statistischen Reihe	Zahl der		Zahl der Fehlergruppen mit									
	pos.	neg.	1	2	3	4	5	6	7	8	9	10
	Fehler		Fehler	nur positiven oder nur negativen Fehlern								
1	1–0	0–1	2	—	—	—	—	—	—	—	—	—
2	2–0	0–2	4	2	—	—	—	—	—	—	—	—
3	3–0	0–3	10	4	2	—	—	—	—	—	—	—
4	4–0	0–4	24	10	4	2	—	—	—	—	—	—
5	5–0	0–5	56	24	10	4	2	—	—	—	—	—
6	6–0	0–6	128	56	24	10	4	2	—	—	—	—
7	7–0	0–7	288	128	56	24	10	4	2	—	—	—
8	8–0	0–8	640	288	128	56	24	10	4	2	—	—
9	9–0	0–9	1408	640	288	128	56	24	10	4	2	—
10	10–0	0–10	3072	1408	640	288	128	56	24	10	4	2

Die mechanische Ausgleichung einer statistischen Reihe wird durchschnittlich um so erfolgreicher, d. h. die Summe der absoluten Fehler nach der Ausgleichung wird um so kleiner sein, je grösser unter sonst gleichen Verhältnissen (will sagen ohne den absoluten Fehler eines einzelnen Gliedes zu ändern, wie schon anfangs vorausgesetzt) die Zahl der Fehlergruppen ist. Über die durchschnittliche Zahl der Fehlergruppen auf eine Fehlerkombination gibt die folgende Tabelle Auskunft:

Gliederzahl der statistischen Reihe	Zahl aller Fehlergruppen	Zahl der Fehlerkombinationen (infolge Änderung d. Vorzeichens	Zahl der Fehlergruppen auf eine Fehlerkombination	Gliederzahl der statistischen Reihe	Zahl aller Fehlergruppen	Zahl der Fehlerkombinationen (infolge Änderung d. Vorzeichens	Zahl der Fehlergruppen auf eine Fehlerkombination
1	2	2	1	6	224	64	3,5
2	6	4	1,5	7	512	128	4
3	16	8	2	8	1152	256	4,5
4	40	16	2,5	9	2560	512	5
5	96	32	3	10	5632	1024	5,5

Etwas detaillierter ist die folgende Darstellung gehalten.

Aus derselben ergibt sich, dass bei gegebener Gliederzahl der statistischen Reihe diejenigen Fehlerkombinationen für die Ausgleichung am vorteilhaftesten sind, für welche die Zahl der positiven Fehler mit der Zahl der negativen möglichst übereinstimmt. In der Praxis wird jeder positive Fehler einen gleich grossen negativen hervorrufen und vice versa. Wenn z.B. jemand sein Alter um ein Jahr zu niedrig angibt, 50 statt 51, so wird in der Altersklasse 50 eine Person zu viel und in der Altersklasse 51 eine Person zu wenig erscheinen. Man darf daher sagen, dass für eine statistische Reihe die Summe der positiven Fehler gleich ist der Summe der negativen, vorausgesetzt, dass Auslassungen und mehrfache Zählungen mit ge-

Gliederzahl d. statistischen Reihe	Zahl der		Zahl aller Fehlergruppen	Zahl aller Fehlerkombinationen	Zahl der Fehlergruppen auf eine Fehlerkombination
	pos.	neg.			
	Fehler				
10	10	0	1	1	1
10	9	1	28	10	2,8
10	8	2	189	45	4,2
10	7	3	623	120	5,2
10	6	4	1219	210	5,8
10	5	5	1512	252	6
10	4	6	1219	210	5,8
10	3	7	623	120	5,2
10	2	8	189	45	4,2
10	1	9	28	10	2,8
10	0	10	1	1	1
8	4	4	350	70	5
6	3	3	80	20	4
4	2	2	18	6	3
2	1	1	4	2	2

nügender Sorgfalt vermieden worden sind; denn eine Kompensation braucht hier nicht einzutreten; wenn ausserdem die Fehler der Reihe vom Zufall regiert werden, so darf man zudem noch gleichviel positive wie negative Fehler voraussetzen, weil zu einer andern Annahme kein Grund vorhanden ist. Wir kommen also zu dem Schlusse: Sofern die Fehler einer statistischen Reihe vom Zufall regiert werden, sind die Voraussetzungen zu einer möglichst erfolgreichen Ausgleichung von selbst erfüllt. Es ist also von Bedeutung, systematische Fehler bei jeder Statistik tunlichst zu vermeiden, oder, wenn dies nicht möglich ist, diese selbst einer besondern Untersuchung zu unterwerfen, die ihre Elimination möglich macht [1]). Auf jeden Fall ist durch die mechanische Ausgleichung einer Reihe immer eine Reduktion des durchschnittlichen Fehlers eines Gliedes zu erwarten, auch wenn die günstigsten Chancen nicht vorhanden sind. Die Beispiele am Schlusse werden dies widerspruchslos zeigen.

Bis jetzt wurde angenommen, die äussersten Glieder der statistischen Reihe können nicht ausgeglichen werden. Sofern die Verteilungskurve zum Typus der Fehlerkurve von Gauss gehört, d. h. wenn die Glieder der Reihe nach beiden Seiten gegen Null abnehmen, kann man diese Annahme fallen lassen, überhaupt jedesmal, wenn die Glieder einer statistischen Reihe sich mit fallendem und wachsendem Index demselben Grenzwerte nähern. Eine solche Ausgleichung kann man als durchgehende bezeichnen. Wir lassen auch für diesen Fall

[1]) Über systematische Fehler (Altersangaben in runden Zahlen, zu niedrige Altersangaben beim weiblichen Geschlecht etc.) s. *A. Knibbs*, op. c., S. 108 ff. — Über das normal-zufällige Geschehen und seine Gesetze s. auch *G. Pólya*, Anschauliche und elementare Darstellung der Lexisschen Dispersionstheorie, Zeitschrift für schweiz. Statistik und Volkswirtschaft, Heft 2, 1919.

eine Statistik über die Fehlerkombinationen folgen; auch hier soll jedem Gliede ein Fehler von unveränderlichem absolutem Werte zugeordnet sein. Die Statistik gestaltet sich so, als ob die Abszissenachse den Umfang eines Kreises mit sehr grossem Radius darstellen würde; diese Vorstellung über die kreisförmige Anordnung der Reihenglieder hat den Vorteil, dass damit sämtliche Glieder als gleichwertig, d. h. für die Ausgleichung in gleicher Weise zugänglich, erscheinen:

Fehlerkombinationen bei restloser (durchgehender) Ausgleichung.

Gliederzahl der statistischen Reihe	Zahl der		Zahl der Fehlergruppen mit									
	pos.	neg.	1	2	3	4	5	6	7	8	9	10
	Fehler		Fehler	nur positiven oder nur negativen Fehlern								
1	1	0	0	—	—	—	—	—	—	—	—	—
1	0	1	0	—	—	—	—	—	—	—	—	—
2	2	0	0	0	—	—	—	—	—	—	—	—
2	1	1	4	0	—	—	—	—	—	—	—	—
2	0	2	0	0	—	—	—	—	—	—	—	—
3	3	0	0	0	0	—	—	—	—	—	—	—
3	2	1	3	3	0	—	—	—	—	—	—	—
3	1	2	3	3	0	—	—	—	—	—	—	—
3	0	3	0	0	0	—	—	—	—	—	—	—
4	4	0	0	0	0	0	—	—	—	—	—	—
4	3	1	4	0	4	0	—	—	—	—	—	—
4	2	2	8	8	0	0	—	—	—	—	—	—
4	1	3	4	0	4	0	—	—	—	—	—	—
4	0	4	0	0	0	0	—	—	—	—	—	—
5	5	0	0	0	0	0	0	—	—	—	—	—
5	4	1	5	0	0	5	0	—	—	—	—	—
5	3	2	15	10	5	0	0	—	—	—	—	—
5	2	3	15	10	5	0	0	—	—	—	—	—
5	1	4	5	0	0	5	0	—	—	—	—	—
5	0	5	0	0	0	0	0	—	—	—	—	—
6	6	0	0	0	0	0	0	0	—	—	—	—
6	5	1	6	0	0	0	6	0	—	—	—	—
6	4	2	24	12	6	6	0	0	—	—	—	—
6	3	3	36	24	12	0	0	0	—	—	—	—
6	2	4	24	12	6	6	0	0	—	—	—	—
6	1	5	6	0	0	0	6	0	—	—	—	—
6	0	6	0	0	0	0	0	0	—	—	—	—
7	7	0	0	0	0	0	0	0	0	—	—	—
7	6	1	7	0	0	0	0	7	0	—	—	—
7	5	2	35	14	7	7	7	0	0	—	—	—
7	4	3	70	42	21	7	0	0	0	—	—	—
7	3	4	70	42	21	7	0	0	0	—	—	—
7	2	5	35	14	7	7	7	0	0	—	—	—
7	1	6	7	0	0	0	0	7	0	—	—	—
7	0	7	0	0	0	0	0	0	0	—	—	—
8	8	0	0	0	0	0	0	0	0	0	—	—
8	7	1	8	0	0	0	0	0	8	0	—	—
8	6	2	48	16	8	8	8	8	0	0	—	—
8	5	3	120	64	32	16	8	0	0	0	—	—
8	4	4	160	96	48	16	0	0	0	0	—	—
8	3	5	120	64	32	16	8	0	0	0	—	—
8	2	6	48	16	8	8	8	8	0	0	—	—
8	1	7	8	0	0	0	0	0	8	0	—	—
8	0	8	0	0	0	0	0	0	0	0	—	—
9	9	0	0	0	0	0	0	0	0	0	0	—
9	8	1	9	0	0	0	0	0	0	9	0	—
9	7	2	63	18	9	9	9	9	9	0	0	—
9	6	3	189	90	45	27	18	9	0	0	0	—
9	5	4	315	180	90	36	9	0	0	0	0	—
9	4	5	315	180	90	36	9	0	0	0	0	—
9	3	6	189	90	45	27	18	9	0	0	0	—
9	2	7	63	18	9	9	9	9	9	0	0	—
9	1	8	9	0	0	0	0	0	0	9	0	—
9	0	9	0	0	0	0	0	0	0	0	0	—
10	10	0	0	0	0	0	0	0	0	0	0	0
10	9	1	10	0	0	0	0	0	0	0	10	0
10	8	2	80	20	10	10	10	10	10	10	0	0
10	7	3	280	118	68	40	24	20	10	0	0	0
10	6	4	560	302	142	70	36	10	0	0	0	0
10	5	5	700	400	200	80	20	0	0	0	0	0
10	4	6	560	302	142	70	36	10	0	0	0	0
10	3	7	280	118	68	40	24	20	10	0	0	0
10	2	8	80	20	10	10	10	10	10	10	0	0
10	1	9	10	0	0	0	0	0	0	0	10	0
10	0	10	0	0	0	0	0	0	0	0	0	0

Gewisse Gesetzmässigkeiten treten deutlich zutage; auch bewährt sich hier das bekannte statistische Gesetz, dass Gesetzmässigkeiten bei grosser Beobachtungszahl besser zu erkennen sind als bei kleinem Umfang der Beobachtungen; so ergibt die Summation der Vertikalkolonnen eine nach einfachem Gesetz fortschreitende Zahlenreihe:

Gliederzahl der statistischen Reihe	Zahl der		Zahl der Fehlergruppen mit									
	pos.	neg.	1	2	3	4	5	6	7	8	9	10
	Fehler		Fehler	nur positiven oder nur negativen Fehlern								
1	1-0	0-1	0	—	—	—	—	—	—	—	—	—
2	2-0	0-2	4	0	—	—	—	—	—	—	—	—
3	3-0	0-3	6	6	0	—	—	—	—	—	—	—
4	4-0	0-4	16	8	8	0	—	—	—	—	—	—
5	5-0	0-5	40	20	10	10	0	—	—	—	—	—
6	6-0	0-6	96	48	24	12	12	0	—	—	—	—
7	7-0	0-7	224	112	56	28	14	14	0	—	—	—
8	8-0	0-8	512	256	128	64	32	16	16	0	—	—
9	9-0	0-9	1152	576	288	144	72	36	18	18	0	—
10	10-0	0-10	2560	1280	640	320	160	80	40	20	20	0

Die Zahl der Fehlergruppen gibt hier an, wie oftmals die Fehler das Vorzeichen wechseln. Für die durchschnittliche Zahl der Zeichenwechsel erhält man daher das folgende Bild:

Gliederzahl der statistischen Reihe	Zahl der pos. Fehler	Zahl der neg. Fehler	Zahl der Zeichenwechsel	Zahl der Fehlerkombinationen	Zahl der Zeichenwechsel auf eine Fehlerkombination
1	1—0	0—1	0	2	0
2	2—0	0—2	4	4	1
3	3—0	0—3	12	8	1,5
4	4—0	0—4	32	16	2
5	5—0	0—5	80	32	2,5
6	6—0	0—6	192	64	3
7	7—0	0—7	448	128	3,5
8	8—0	0—8	1024	256	4
9	9—0	0—9	2304	512	4,5
10	10—0	0—10	5120	1024	5

Bei der durchgehenden Ausgleichung hat man also auf je 2 Glieder der Reihe einen Zeichenwechsel zu erwarten.

Von besonderem Interesse ist die Frage: Bei welchem Verhältnis der positiven und negativen Fehler wird die Zahl der Zeichenwechsel oder was dasselbe ist, die Zahl der Fehlergruppen, am grössten? Die Antwort gibt die folgende Zusammenstellung:

Gliederzahl der statistischen Reihe	Zahl der pos. Fehler	Zahl der neg. Fehler	Zahl der Zeichenwechsel	Zahl der Fehlerkombinationen	Zahl der Zeichenwechsel auf eine Fehlerkombination
10	10	0	0	1	0
10	9	1	20	10	2
10	8	2	160	45	3,55...
10	7	3	560	120	4,66...
10	6	4	1120	210	5,33...
10	5	5	1400	252	**5,55...**
10	4	6	1120	210	5,33...
10	3	7	560	120	4,66...
10	2	8	160	45	3,55...
10	1	9	20	10	2
10	0	10	0	1	0

Zu demselben Resultat gelangt man in etwas anderer Form, wenn man die Fehlerkombinationen nach der Zahl der Zeichenwechsel ordnet. Wir nehmen zunächst wieder an, die äussersten Reihenglieder bleiben von der Ausgleichung unberührt und erhalten:

Gliederzahl der statistischen Reihe	Zahl der pos. Fehler	Zahl der neg. Fehler	Zahl der Fehlerkombinationen mit 1 Z.-W.	2 Zeichenwechseln	3	4	5	6	7	8	9	10
1	1	0	—	—	—	—	—	—	—	—	—	—
1	0	1	—	—	—	—	—	—	—	—	—	—
2	2	0	0	—	—	—	—	—	—	—	—	—
2	1	1	2	—	—	—	—	—	—	—	—	—
2	0	2	0	—	—	—	—	—	—	—	—	—
3	3	0	0	0	—	—	—	—	—	—	—	—
3	2	1	2	1	—	—	—	—	—	—	—	—
3	1	2	2	1	—	—	—	—	—	—	—	—
3	0	3	0	0	—	—	—	—	—	—	—	—
4	4	0	0	0	0	—	—	—	—	—	—	—
4	3	1	2	2	0	—	—	—	—	—	—	—
4	2	2	2	2	2	—	—	—	—	—	—	—
4	1	3	2	2	0	—	—	—	—	—	—	—
4	0	4	0	0	0	—	—	—	—	—	—	—
5	5	0	0	0	0	0	—	—	—	—	—	—
5	4	1	2	3	0	0	—	—	—	—	—	—
5	3	2	2	3	4	1	—	—	—	—	—	—
5	2	3	2	3	4	1	—	—	—	—	—	—
5	1	4	2	3	0	0	—	—	—	—	—	—
5	0	5	0	0	0	0	—	—	—	—	—	—
6	6	0	0	0	0	0	0	—	—	—	—	—
6	5	1	2	4	0	0	0	—	—	—	—	—
6	4	2	2	4	6	3	0	—	—	—	—	—
6	3	3	2	4	8	4	2	—	—	—	—	—
6	2	4	2	4	6	3	0	—	—	—	—	—
6	1	5	2	4	0	0	0	—	—	—	—	—
6	0	6	0	0	0	0	0	—	—	—	—	—
7	7	0	0	0	0	0	0	0	—	—	—	—
7	6	1	2	5	0	0	0	0	—	—	—	—
7	5	2	2	5	8	6	0	0	—	—	—	—
7	4	3	2	5	12	9	6	1	—	—	—	—
7	3	4	2	5	12	9	6	1	—	—	—	—
7	2	5	2	5	8	6	0	0	—	—	—	—
7	1	6	2	5	0	0	0	0	—	—	—	—
7	0	7	0	0	0	0	0	0	—	—	—	—
8	8	0	0	0	0	0	0	0	0	—	—	—
8	7	1	2	6	0	0	0	0	0	—	—	—
8	6	2	2	6	10	10	0	0	0	—	—	—
8	5	3	2	6	16	16	12	4	0	—	—	—
8	4	4	2	6	18	18	18	6	2	—	—	—
8	3	5	2	6	16	16	12	4	0	—	—	—
8	2	6	2	6	10	10	0	0	0	—	—	—
8	1	7	2	6	0	0	0	0	0	—	—	—
8	0	8	0	0	0	0	0	0	0	—	—	—
9	9	0	0	0	0	0	0	0	0	0	—	—
9	8	1	2	7	0	0	0	0	0	0	—	—
9	7	2	2	7	12	15	0	0	0	0	—	—
9	6	3	2	7	20	25	20	10	0	0	—	—
9	5	4	2	7	24	30	36	18	8	1	—	—
9	4	5	2	7	24	30	36	18	8	1	—	—

Gliederzahl der statistischen Reihe	Zahl der		Zahl der Fehlerkombinationen mit									
	pos.	neg.	1	2	3	4	5	6	7	8	9	10
	Fehler		Z.-W.	Zeichenwechseln								
9	3	6	2	7	20	25	20	10	0	0	—	—
9	2	7	2	7	12	15	0	0	0	0	—	—
9	1	8	2	7	0	0	0	0	0	0	—	—
9	0	9	0	0	0	0	0	0	0	0	—	—
10	10	0	0	0	0	0	0	0	0	0	0	—
10	9	1	2	8	0	0	0	0	0	0	0	—
10	8	2	2	8	14	21	0	0	0	0	0	—
10	7	3	2	8	24	36	30	20	0	0	0	—
10	6	4	2	8	30	45	60	40	20	5	0	—
10	5	5	2	8	32	48	72	48	32	8	2	—
10	4	6	2	8	30	45	60	40	20	5	0	—
10	3	7	2	8	24	36	30	20	0	0	0	—
10	2	8	2	8	14	21	0	0	0	0	0	—
10	1	9	2	8	0	0	0	0	0	0	0	—
10	0	10	0	0	0	0	0	0	0	0	0	—

Sieht man sich diese Zahlen näher an, so wird man bemerken, dass die Zeichenwechsel um so dichter auftreten, je kleiner der Unterschied zwischen der Zahl der positiven und negativen Fehler ist. Wir wollen uns beschränken auf die zehngliedrige Reihe, indem wir den relativen Betrag der Fehlerkombinationen notieren unter Berücksichtigung der 2 Kombinationen ohne Zeichenwechsel:

Gliederzahl der statistischen Reihe	Zahl der		Zahl der Fehlerkombinationen (in Prozenten der Horizontalsumme) mit									
	pos.	neg.	0	1	2	3	4	5	6	7	8	9
	Fehler		Z.-W.	Z.-W.	Zeichenwechseln							
10	10	0	100	0	0	0	0	0	0	0	0	0
10	9	1	0	20	80	0	0	0	0	0	0	0
10	8	2	0	4,4	17,8	31,1	46,7	0	0	0	0	0
10	7	3	0	1,7	6,6	20	30	25	16,7	0	0	0
10	6	4	0	1,0	3,8	14,3	21,4	28,6	19,0	9,5	2,4	0
10	5	5	0	0,8	3,2	12,7	19,0	28,6	19,0	12,7	3,2	0,8
10	4	6	0	1,0	3,8	14,3	21,4	28,6	19,0	9,5	2,4	0
10	3	7	0	1,7	6,6	20	30	25	16,7	0	0	0
10	2	8	0	4,4	17,8	31,1	46,7	0	0	0	0	0
10	1	9	0	20	80	0	0	0	0	0	0	0
10	0	10	100	0	0	0	0	0	0	0	0	0

In den vorstehenden Vertikalreihen (0—9) stellen die mittleren Glieder, indem man von links nach rechts fortschreitet, zuerst Minimalwerte, dann aber Maximalwerte dar, womit aber offenbar gezeigt ist, dass bei 5 positiven und 5 negativen Fehlern die Zeichenwechsel am dichtesten auftreten.

Im Falle durchgehender Ausgleichung erhält man analoge Resultate; die Statistik zeigt, dass die Zeichenwechsel nur in gerader Zahl auftreten, (was übrigens unmittelbar einleuchtet) und dass man die Zahl der Fehlerkombinationen mit 2, 4, 6, 8, 10... Zeichenwechseln erhält, wenn die früher erhaltenen Zahlen der Reihe nach paarweise addiert werden. Wir beschränken uns darum auf eine statistische Reihe von 10 Gliedern:

Restlose Ausgleichung.

Gliederzahl der statistischen Reihe	Zahl der		Zahl der Kombinationen mit									
	pos.	neg.	1	2	3	4	5	6	7	8	9	10
	Glieder		Z.-W.	Zeichenwechseln								
10	10	0	0	0	0	0	0	0	0	0	0	0
10	9	1	0	10	0	0	0	0	0	0	0	0
10	8	2	0	10	0	35	0	0	0	0	0	0
10	7	3	0	10	0	60	0	50	0	0	0	0
10	6	4	0	10	0	75	0	100	0	25	0	0
10	5	5	0	10	0	80	0	120	0	40	0	2
10	4	6	0	10	0	75	0	100	0	25	0	0
10	3	7	0	10	0	60	0	50	0	0	0	0
10	2	8	0	10	0	35	0	0	0	0	0	0
10	1	9	0	10	0	0	0	0	0	0	0	0
10	0	10	0	0	0	0	0	0	0	0	0	0

oder bei Darstellung in relativen Zahlen, mit Einschluss der 2 Kombinationen ohne Zeichenwechsel:

Gliederzahl de. statistischen Reihe	Zahl der		Zahl der Fehlerkombinationen (in Prozenten der Horizontalsumme) mit										
	pos.	neg.	0	1	2	3	4	5	6	7	8	9	10
	Glieder		Z.-W.	Z.-W.	Zeichenwechseln								
10	10	0	100	0	0	0	0	0	0	0	0	0	0
10	9	1	0	0	100	0	0	0	0	0	0	0	0
10	8	2	0	0	22,2	0	77,8	0	0	0	0	0	0,
10	7	3	0	0	8,3	0	50	0	41,7	0	0	0	0
10	6	4	0	0	4,8	0	35,7	0	47,6	0	11,9	0	0
10	5	5	0	0	4	0	31,7	0	47,6	0	15,9	0	0,8
10	4	6	0	0	4,8	0	35,7	0	47,6	0	11,9	0	0
10	3	7	0	0	8,3	0	50	0	41,7	0	0	0	0
10	2	8	0	0	22,2	0	77,8	0	0	0	0	0	0
10	1	0	0	0	100	0	0	0	0	0	0	0	0
10	0	10	100	0	0	0	0	0	0	0	0	0	0

Auch hier beweist der Verlauf der Vertikalreihen, dass die Zeichenwechsel umso dichter auftreten, je kleiner die Differenz zwischen der Zahl der positiven und negativen Fehler ist.

Es erübrigt noch, an Beispielen zu zeigen, dass die Fehlersumme einer statistischen Reihe durch die Ausgleichung um so kleiner wird, je mehr Zeichenwechsel auftreten. Wir wählen hierzu den allerdings sehr einfachen aber recht typischen Fall, dass die Fehler der statistischen Reihe sämtlich den absoluten Wert 1 besitzen, und beschränken uns auf 8 Glieder, sowie auf den ausgleichenden Mittelwert: $\dfrac{y_{n-1} + y_n + y_{n+1}}{3}$; die äussersten Glieder bleiben zunächst unausgeglichen:

Gliederzahl der statistischen Reihe	Zahl der		Zahl der Zeichenwechsel									
	pos.	neg.	0	1	2	3	4	5	6	7	8	9
	Fehler		Durchschnittliche Fehlersumme einer Fehlerkombination nach der Ausgleichung									
8	8	0	8	0	0	0	0	0	0	0	—	—
8	7	1	0	7,3	6,2	0	0	0	0	0	—	—
8	6	2	0	6,7	5,8	5,6	4,8	0	0	0	—	—
8	5	3	0	6,7	5,8	5	4,5	4,3	4	0	—	—
8	4	4	0	6,7	5,8	4,9	4,4	4	4	4	—	—
8	3	5	0	6,7	5,8	5	4,5	4,3	4	0	—	—
8	2	6	0	6,7	5,8	5,6	4,8	0	0	0	—	—
8	1	7	0	7,3	6,2	0	0	0	0	0	—	—
8	0	8	8	0	0	0	0	0	0	0	—	—

Die Horizontalreihen nehmen ab, wenn man von links nach rechts fortschreitet (0—7), wie es erwartet wurde. Im Falle der restlosen Ausgleichung konstatiert man dasselbe, gemäss der folgenden Darstellung.

Wir haben uns anhand des Vorstehenden wiederholt davon überzeugen können, dass die statistische Methode nicht bloss zur Erforschung gesellschaftlicher Zustände und Vorgänge geeignet ist, sondern dass sie überall da von Bedeutung sein kann, wo es sich um das Studium von Gesetzmässigkeiten handelt, d. h. im Bereiche jeder wissenschaftlichen Tätigkeit. Kann man sich z. B. die

Restlose Ausgleichung; ausgleichender Mittelwert:

$$\frac{y_{n-1} + y_n + y_{n+1}}{3}.$$

Gliederzahl der statistischen Reihe	Zahl der		Zahl der Zeichenwechsel									
	pos.	neg.	0	1	2	3	4	5	6	7	8	9
	Fehler		Durchschnittliche Fehlersumme einer Fehlerkombination nach der Ausgleichung									
8	8	0	8	0	0	0	0	0	0	0	0	—
8	7	1	0	0	6	0	0	0	0	0	0	—
8	6	2	0	0	5,3	0	4,3	0	0	0	0	—
8	5	3	0	0	5,3	0	3,6	0	3	0	0	—
8	4	4	0	0	5,3	0	3,6	0	2,7	0	2,7	—
8	3	5	0	0	5,3	0	3,6	0	3	0	0	—
8	2	6	0	0	5,3	0	4,3	0	0	0	0	—
8	1	7	0	0	6	0	0	0	0	0	0	—
8	0	8	8	0	0	0	0	0	0	0	0	—

Methode der induktiven Forschung ohne ein gewisses Mass statistischer Tätigkeit denken? Kaum.

Es wäre daher verfehlt, wollte man die statistische Methode als solche angreifen. Was die Statistik vielfach zum Gegenstand der Kritik macht, ist der Umstand, dass ihre Resultate oft zu sehr verallgemeinert werden; eine Statistik, deren Zahlen sich nach bestimmten Tendenzen richten müssen, kann natürlich keinerlei Anspruch auf wissenschaftlichen Wert machen.