Ein Beitrag zur Theorie der Attribute.

Von Dr. V. Furlan, Basel.

1.

Im folgenden sollen einige Ideen zu der Theorie der Attribute ausgeführt und durchweg an einem praktischen Beispiel näher erläutert werden, wobei nur die Elemente der Theorie, wie sie von Pierson, Yule und anderen entwickelt wurden, vorausgesetzt werden. Die Masse oder das Grundmaterial sei mit N , die Attribute mit $a, b, c, \dots$ bezeichnet; ferner bedeute (a) das Mass für das Vorkommen des Attributs a in der Masse N , (b) das Mass für das Vorkommen des Attributs b in der Masse N , (c) das Mass für das Vorkommen des Attributs c in der Masse N , usw.; endlich sei das Mass des gleichzeitigen Vorkommens der beiden Attribute a und b mit $(a b)$, das Mass des gleichzeitigen Vorkommens der drei Attribute a, b und c mit $(a b c)$ bezeichnet usw.

Wir wollen nun in einem gegebenen Grundmaterial zwei Attribute a und b unterscheiden und zunächst ein Mass für die Konzentration dieser Attribute aufstellen. Zu diesem Zwecke stellen wir fest, dass sinngemäss die Konzentration der beiden Attribute um so grösser sein wird, je häufiger die Koexistenz derselben stattfindet, das heisst mit andern Worten, je grösser $(a b)$ ist im Verhältnis zu dem Vorkommen der beiden Attribute überhaupt. Das Mass für das Vorkommen der beiden Attribute a und b überhaupt ist aber

$$(a) + (b) - (a b).$$

Der negative Summand in dem letzteren Ausdruck ist deswegen notwendig, weil sonst die Fälle, in denen

$$\gamma_1 \cdot \frac{(a b)}{(a) - (a b) + (b)} + \gamma_2 \cdot \frac{(a c)}{(a) - (a c) + (c)} + \gamma_3 \cdot \frac{(b c)}{(b) - (b c) + (c)}$$

$$\gamma_1 + \gamma_2 + \gamma_3$$

Es bleibt nur noch übrig, die drei Gewichte in folgerichtiger Weise festzulegen. Und da drängt sich von selbst die Notwendigkeit auf, den Konzentrationsindex für die beiden Attribute a und b nur in dem Verhältnis in den Gesamtindex eingehen zu lassen, in dem die beiden Attribute überhaupt vorkommen. Nun ist das Mass für das Vorkommen der beiden Attribute, wie bereits bemerkt,

beide Attribute gleichzeitig vorkommen oder koexistieren, und die sowohl in (a) als auch in (b) enthalten sind, doppelt gezählt würden.

Nach dem eben Gesagten ergibt sich als Mass der Konzentration der beiden Attribute a und b :

$$\frac{(a b)}{(a) - (a b) + (b)}$$

Falls wir drei Attribute haben, tauchen neue Fragen auf. Wir können zunächst die Konzentration zwischen a und b , zwischen b und c und zwischen c und a für sich messen. Wir erhalten dann drei Konzentrationsmasse, nämlich

$$\frac{(a b)}{(a) - (a b) + (b)}$$

$$\frac{(a c)}{(a) - (a c) + (c)} \quad \text{und}$$

$$\frac{(b c)}{(b) - (b c) + (c)}$$

und wir können uns fragen, ob sich aus diesen drei Grössen nicht in geeigneter Weise ein Durchschnitt ziehen lässt, welcher aufzufassen wäre als Indexziffer für die Konzentration der drei Attribute zu je zweien. Wir werden also in geeigneter Weise drei Gewichte $\gamma_1, \gamma_2, \gamma_3$ bestimmen und danach den folgenden gewogenen Durchschnitt bilden:

$$(a) - (a b) + (b)$$

somit wird

$$\gamma_1 = (a) - (a b) + (b)$$

sein, und analog werden wir haben

$$\gamma_2 = (a) - (a c) + (c)$$

$$\gamma_3 = (b) - (b c) + (c)$$

Der Zähler des obigen Ausdruckes 1) geht daher über in

$$(ab) + (ac) + (bc)$$

während der Nenner die folgende Form annimmt

$$2(a) + 2(b) + 2(c) - (ab) - (ac) - (bc)$$

Die Indexziffer für die Konzentration der drei Attribute zu je zweien wird daher sich in der folgenden Form darstellen lassen

$$\frac{r}{2s - r}$$

worin

$$r = (ab) + (ac) + (bc)$$

und

$$s = (a) + (b) + (c)$$

gesetzt wurde.

Bei drei Attributen taucht aber noch eine andere Frage auf. Neben der Konzentration zu zweien spielt eine Konzentration zu dreien eine Rolle. Ein Mass für diese Konzentration zu dreien wird gefunden werden, wenn wir berücksichtigen, dass sinngemäss die Konzentration der drei Attribute um so grösser sein wird, je häufiger die Koexistenz derselben stattfindet, das heisst mit andern Worten, je grösser (abc) ist im Verhältnis zu dem Vorkommen der drei Attribute überhaupt. Das Mass für das Vorkommen der drei Attribute a , b und c überhaupt ist aber

$$(a) + (b) + (c) - (ab) - (bc) - (ac) + (abc)$$

Der Bau dieses Ausdrucks liegt auf der Hand. Da die Elemente der Masse, bei denen die beiden Merkmale a und b koexistieren, in der Summe $(a) + (b)$ zweimal gezählt werden, muss ein (ab) abgezogen werden; ebenso ist es notwendig (bc) und (ac) abzuziehen. Diejenigen Elemente, bei denen gleichzeitig die drei Merkmale a , b und c sich vorfinden, werden zunächst dreimal mitgezählt, nämlich bei (a) , bei (b) und bei (c) , und sie werden hierauf dreimal abgezogen, bei (ab) , bei (bc) und bei (ac) ; damit sie überhaupt zur Berücksichtigung gelangen, muss dann noch der Summand (abc) hinzugefügt werden.

Der obige Ausdruck für das Mass des Vorkommens der drei Attribute a , b und c überhaupt lässt sich auch schreiben, indem wir von den bereits eingeführten Zeichen r und s Gebrauch machen:

$$s - r + (abc).$$

Die Indexziffer für die Konzentration der drei Attribute zu dreien wird daher lauten:

$$\frac{(abc)}{s - r + (abc)}$$

2.

Von den obigen Ausführungen soll zunächst eine einfache Anwendung gemacht werden.

A. Die Grundmasse bestehe aus den Ende 1896 an den Börsen Basel, Genf und Zürich kotierten schweizerischen Obligationen, gemessen in Millionen Franken. Insgesamt waren damals an diesen drei Börsen 30,396.¹ Millionen Franken Obligationen kotiert, und hiervon waren 1605 Millionen Franken schweizerische Obligationen. Diese Summe verteilte sich folgendermassen:

1. kotiert gleichzeitig in Basel, Genf und Zürich: 801.³¹ Millionen Franken;
2. kotiert gleichzeitig nur in Basel und Zürich: 169.¹² Millionen Franken;
3. kotiert gleichzeitig nur in Basel und Genf: 73.⁸⁵ Millionen Franken;
4. kotiert gleichzeitig nur in Genf und Zürich: 135.⁹⁷ Millionen Franken;
5. kotiert nur in Basel: 247.⁹² Millionen Franken;
6. kotiert nur in Zürich: 93.²⁴ Millionen Franken;
7. kotiert nur in Genf: 83.⁵⁹ Millionen Franken.

Es fragt sich nun:

- a) wie gross ist die Konzentration zwischen Basel und Zürich allein;
- b) wie gross ist die Konzentration zwischen Basel und Genf allein;
- c) wie gross ist die Konzentration zwischen Genf und Zürich allein;
- d) wie gross ist die durchschnittliche Konzentration der drei Plätze Basel, Genf und Zürich zu je zweien;
- e) wie gross ist die Konzentration der drei Plätze Basel, Genf und Zürich zu dreien?

Die Antwort liegt in den obigen Formeln skizziert. Um die Konzentration zwischen Basel und Zürich allein zu berechnen, werden wir die Höhe des an beiden Börsen gemeinsam kotierten Kapitals dividieren durch das in Basel und Zürich kotierte Gesamtkapital, also:

$$\frac{169,12 + 801,31}{801,31 + 169,12 + 73,85 + 135,97 + 247,92 + 93,24}$$

Dieser Ausdruck berechnet sich, in Prozenten ausgedrückt, zu 63.⁸ %.

Es beträgt mithin, in der gekennzeichneten Weise gemessen, die Konzentration zwischen

Basel und Zürich . . .	63. ⁸ %
Basel und Genf . . .	57. ⁹ %
Genf und Zürich . . .	69. ¹ %

Dass diese drei Indizes nicht sehr voneinander abweichen, hängt damit zusammen, dass im Zähler

derselben jeweilen der Betrag von 801.³¹ Millionen Franken, der an allen drei Börsen gemeinsam kotiert wird, als Summand auftritt und den grössten Teil des Zählers ausmacht, während die Nenner nicht sehr verschieden sind.

Die durchschnittliche Konzentration der drei Plätze Basel, Genf und Zürich zu je zweien berechnet sich nach der oben angegebenen Formel zu 63.4 %; dieser Wert ist, wie man sieht, der Konzentration zwischen Basel und Zürich sehr nahe.

Endlich die Konzentration der drei Plätze Basel, Genf und Zürich zu dreien ergibt sich nach dem obigen zu

$$\frac{801,31}{1605,0} \text{ gleich } 49,9 \text{ \%}.$$

B. Die Grundmasse bestehe aus den Ende 1913 an den Börsen Basel, Genf und Zürich kotierten schweizerischen Obligationen, gemessen wiederum in Millionen Franken. Insgesamt waren damals an diesen drei Börsen 35,088.⁹ Millionen Franken Obligationen kotiert und hiervon stellten 3725.⁴ Millionen Franken schweizerische Titel dar. Diese Summe verteilte sich in der folgenden Weise:

1. kotiert gleichzeitig in Basel, Genf und Zürich: 2509.³⁸ Millionen Franken;
2. kotiert gleichzeitig nur in Basel und Zürich: 430.⁴² Millionen Franken;
3. kotiert gleichzeitig nur in Basel und Genf: 177.²⁰ Millionen Franken;
4. kotiert gleichzeitig nur in Genf und Zürich: 64.⁷³ Millionen Franken;
5. kotiert nur in Basel: 183.³⁵ Millionen Franken;
6. kotiert nur in Zürich: 139.²³ Millionen Franken;
7. kotiert nur in Genf: 221.⁰⁷ Millionen Franken.

Wie bei A. können wir auch hier die Fragen a) bis e) stellen, und wir erhalten dann das folgende Resultat:

Es beträgt die Konzentration zwischen

Basel und Zürich . . .	83.9 %
Basel und Genf . . .	74.9 %
Genf und Zürich . . .	72.7 %

Die durchschnittliche Konzentration der drei Plätze Basel, Genf und Zürich zu je zweien berechnet sich zu 77.2 % und endlich die Konzentration der drei Plätze zu dreien zu 67.4 %.

Interessant ist der Vergleich dieses Ergebnisses mit dem analogen für das Jahr 1896. In der Zwischenzeit ist also die durchschnittliche Konzentration der drei Plätze zu zweien von 63.4 % auf 77.2 % und die Konzentration zu dreien von 49.9 % auf 67.4 % gestiegen. Diese Steigerung ist zum grössten Teile auf

die Verstaatlichung der Eisenbahnen zurückzuführen, wodurch ein grosser Posten von Bundesbahnobligationen geschaffen wurde, der gleichzeitig an den drei Plätzen zur Kotierung gelangt. Im übrigen: Während im Jahre 1896 in bezug auf die Konzentration zu je zweien Genf-Zürich den ersten Platz einnahm, steht es im Jahre 1913 an dritter Stelle, während Basel-Zürich, das damals in zweiter Linie kam, nunmehr den ersten Rang einnimmt.

C. Die Grundmasse bestehe aus den Ende 1896 an den Börsen Basel, Genf und Zürich kotierten schweizerischen Aktien, ausgedrückt wie in den beiden vorhergehenden Beispielen nach dem Nennwert in Millionen Franken. Insgesamt waren damals an diesen drei Börsen 3684.² Millionen Franken Aktien kotiert, und hiervon stellten 754.¹ Millionen Franken schweizerische Titel dar. Diese Summe gliederte sich wie folgt:

1. kotiert gleichzeitig in Basel, Genf und Zürich: 369.¹² Millionen Franken;
2. kotiert gleichzeitig nur in Basel und Zürich: 89.⁷⁶ Millionen Franken;
3. kotiert gleichzeitig nur in Basel und Genf: 45.⁹ Millionen Franken;
4. kotiert gleichzeitig nur in Genf und Zürich: 30 Millionen Franken;
5. kotiert nur in Basel: 71.⁴⁵ Millionen Franken;
6. kotiert nur in Zürich: 82.²² Millionen Franken;
7. kotiert nur in Genf: 65.⁶⁵ Millionen Franken.

Wie oben bei A. können wir auch hier die Fragen a) bis e) wiederholen, und es ergibt sich dann das folgende Resultat:

Es beträgt die Konzentration zwischen

Basel und Zürich . . .	66.6 %
Basel und Genf . . .	61.8 %
Genf und Zürich . . .	58.5 %

Die durchschnittliche Konzentration der drei Plätze Basel, Genf und Zürich zu je zweien berechnet sich zu 62.3 % und endlich die Konzentration zu dreien auf 48.9 %.

Vergleicht man dieses Ergebnis mit dem analogen für die Obligationen (Beispiel A.), so ist zu bemerken, dass im Jahre 1896 die Konzentration der drei Plätze sowohl durchschnittlich zu je zweien als auch zu dreien in Aktien annähernd ebenso gross war wie in Obligationen. Allerdings ist im Verhältnis der Plätze zu je zweien ein grosser Unterschied zu konstatieren: in Aktien steht das Paar Genf-Zürich an letzter Stelle, während es in Obligationen an erster Stelle stand, und Basel-Genf, das bei den Obligationen den letzten Rang einnimmt, figuriert bei den Aktien an erster Stelle.

D. Die Grundmasse bestehe aus den Ende 1913 an den Börsen Basel, Genf und Zürich kotierten Aktien, ausgedrückt in Millionen Franken des Nennwerts. Insgesamt waren an dem genannten Zeitpunkt an den drei Börsen 10,928.2 Millionen Franken Aktien kotiert, und hiervon stellten 1281 Millionen Franken schweizerische Titel dar. Diese Summe lässt sich in der folgenden Weise zerlegen:

1. kotiert gleichzeitig in Basel, Genf und Zürich: 559.75 Millionen Franken;
2. kotiert gleichzeitig nur in Basel und Zürich: 159.56 Millionen Franken;
3. kotiert gleichzeitig nur in Basel und Genf: 20 Millionen Franken;
4. kotiert gleichzeitig nur in Genf und Zürich: 12.1 Millionen Franken;
5. kotiert nur in Basel: 96.15 Millionen Franken;
6. kotiert nur in Zürich: 278.54 Millionen Franken;
7. kotiert nur in Genf: 154.91 Millionen Franken.

Wie in den drei vorhergehenden Beispielen können wir auch hier die Frage nach der Konzentration aufstellen, und wir erhalten dann das folgende Ergebnis:

Es beträgt die Konzentration zwischen

Basel und Zürich . . .	63.9 %
Basel und Genf . . .	57.8 %
Genf und Zürich . . .	48.2 %

Die durchschnittliche Konzentration der drei Plätze Basel, Genf und Zürich zu je zweien berechnet sich auf 56.5 % und schliesslich die Konzentration zu dreien auf 43.7 %.

Vergleicht man dieses Resultat mit dem analogen für das Jahr 1896, so sieht man, dass die durchschnittliche Konzentration zu je zweien und ebenso die Konzentration zu dreien in der Periode 1896—1913 für die Aktien ziemlich abgenommen hat. Der Grund hierfür ist derselbe wie der oben angeführte Grund für die in demselben Zeitraum konstatierte Zunahme der Konzentration für die Obligationen: Durch die Verstaatlichung der Eisenbahnen wurde den schweizerischen Börsen ein sehr erhebliches, an den einzelnen Börsen gemeinsam gehandeltes Effektenmaterial entzogen, das später durch Neueinführungen nicht in demselben Masse ersetzt werden konnte.

Beim Vergleich der Konzentration der Aktien einerseits und der Obligationen andererseits im Jahre 1913 findet man, dass in beiden Fällen die Verbindung Basel-Zürich am engsten ist, während umgekehrt die Verbindung Genf und Zürich die lockerste ist.

3.

Es ist nicht schwer, das sub 1. Ausgeführte auf den Fall mehrerer Attribute zu verallgemeinern. Wir

mögen zunächst den Fall ins Auge fassen, dass in der gegebenen Grundmasse vier Attribute a, b, c und d unterschieden werden. Dann können wir in erster Linie die Konzentration der vier Attribute zu je zweien messen, also zwischen

a und b

a „ c

a „ d

b „ c

b „ d

c „ d

Wir werden insgesamt sechs Konzentrationskoeffizienten dieser Art haben, nämlich:

$$\varrho_1 = \frac{(ab)}{(a) - (ab) + (b)}$$

$$\varrho_2 = \frac{(ac)}{(a) - (ac) + (c)}$$

$$\varrho_3 = \frac{(ad)}{(a) - (ad) + (d)}$$

$$\varrho_4 = \frac{(bc)}{(b) - (bc) + (c)}$$

$$\varrho_5 = \frac{(bd)}{(b) - (bd) + (d)}$$

$$\varrho_6 = \frac{(cd)}{(c) - (cd) + (d)}$$

Auch hier taucht die Idee auf, aus diesen sechs Koeffizienten in geeigneter Weise einen Durchschnitt zu bilden, welcher aufzufassen wäre als Indexziffer für die Konzentration der vier Attribute zu je zweien. Wir nehmen also sechs Gewichte

$$\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5, \gamma_6$$

und bilden danach den gewogenen Durchschnitt:

$$\mu_1 = \frac{\varrho_1 \gamma_1 + \varrho_2 \gamma_2 + \varrho_3 \gamma_3 + \varrho_4 \gamma_4 + \varrho_5 \gamma_5 + \varrho_6 \gamma_6}{\gamma_1 + \gamma_2 + \gamma_3 + \gamma_4 + \gamma_5 + \gamma_6}$$

Die Gewichte werden nun folgerichtigerweise in derselben Weise bestimmt wie oben: Wir werden ϱ_1 nur in demselben Verhältnis in den Gesamtindex eingehen lassen, in dem die beiden Attribute a und b überhaupt vorkommen; und da das Mass für das Vorkommen der beiden Attribute gleich

$$(a) - (ab) + (b)$$

ist, so wird

$$\gamma_1 = (a) - (ab) + (b)$$

sein, und analog werden wir haben:

$$\gamma_2 = (a) - (ac) + (c)$$

$$\gamma_3 = (a) - (ad) + (d)$$

$$\gamma_4 = (b) - (bc) + (c)$$

$$\gamma_5 = (b) - (bd) + (d)$$

$$\gamma_6 = (c) - (cd) + (d)$$

Der Zähler von μ_1 geht also über in die Summe

$$r = (ab) + (ac) + (ad) + (bc) + (bd) + (cd)$$

während der Nenner die folgende Gestalt annimmt:

$$3(a) + 3(b) + 3(c) + 3(d) - \\ - (ab) - (ac) - (ad) - (bc) - (bd) - (cd)$$

Führt man nun weiter die Bezeichnung ein

$$s = (a) + (b) + (c) + (d)$$

so schreibt sich μ_2 , das ist der durchschnittliche Konzentrationskoeffizient der vier Attribute zu je zweien, in der folgenden Weise:

$$\mu_2 = \frac{r}{3s - r}$$

In zweiter Linie wird die Konzentration der Attribute in Gruppen von je dreien zu bestimmen sein, also zwischen

$$\begin{array}{l} a \text{ und } b \text{ und } c \\ a \text{ " } b \text{ " } d \\ a \text{ " } c \text{ " } d \\ b \text{ " } c \text{ " } d \end{array}$$

Wir haben oben die Methode kennen gelernt, wonach die Konzentration zwischen drei Elementen gemessen wird. Wir erhalten danach vier Konzentrationsindizes, die wir bzw. mit τ_1, τ_2, τ_3 und τ_4 bezeichnen mögen, und für die sich die nachstehenden Ausdrücke ergeben:

$$\tau_1 = \frac{(abc)}{(a) + (b) + (c) - (ab) - (bc) - (ac) + (abc)}$$

$$\tau_2 = \frac{(abd)}{(a) + (b) + (d) - (ab) - (ad) - (bd) + (abd)}$$

$$\tau_3 = \frac{(acd)}{(a) + (c) + (d) - (ac) - (ad) - (cd) + (acd)}$$

$$\tau_4 = \frac{(bcd)}{(b) + (c) + (d) - (bc) - (bd) - (cd) + (bcd)}$$

Auch hier können wir daran gehen, aus τ_1, τ_2, τ_3 und τ_4 in geeigneter Weise einen Durchschnitt zu ziehen, welcher dann als durchschnittlicher Konzentrationsindex der vier Attribute zu je dreien anzu-

sprechen wäre. Bezeichnen wir diesen in Analogie zu oben mit μ_3 , so ergibt sich hierfür die Formel

$$\mu_3 = \frac{\gamma_1 \tau_1 + \gamma_2 \tau_2 + \gamma_3 \tau_3 + \gamma_4 \tau_4}{\gamma_1 + \gamma_2 + \gamma_3 + \gamma_4}$$

worin $\gamma_1, \gamma_2, \gamma_3$ und γ_4 in geeigneter Weise festzulegende Gewichte sind. Und nun wird am folgerichtigsten τ_1 nur in dem Verhältnis in den Gesamtindex eingehen, in dem die drei Attribute a, b, c überhaupt vorkommen. Nun ist, wie oben ausgeführt, das Mass für das Vorkommen der drei Attribute überhaupt gleich

$$(a) + (b) + (c) - (ab) - (ac) - (bc) + (abc)$$

Wählen wir diesen Ausdruck als Gewicht γ_1 , so ist in Analogie damit

$$\gamma_2 = (a) + (b) + (d) - (ab) - (ad) - (bd) + (abd)$$

$$\gamma_3 = (a) + (c) + (d) - (ac) - (ad) - (cd) + (acd)$$

$$\gamma_4 = (b) + (c) + (d) - (bc) - (bd) - (cd) + (bcd)$$

und der Zähler von μ_3 geht über in die Summe

$$(abc) + (abd) + (acd) + (bcd)$$

Wir mögen diese Summe in Analogie zu oben mit q bezeichnen. Der Nenner von μ_3 nimmt die folgende Form an:

$$\gamma_1 + \gamma_2 + \gamma_3 + \gamma_4 = 3(a) + 3(b) + 3(c) + \\ + 3(d) - 2(ab) - 2(ac) - 2(ad) - 2(bc) - 2(bd) - \\ - 2(cd) + (abc) + (abd) + (acd) + (bcd)$$

Diese Summe ist aber nichts anderes als

$$3s - 2r + q$$

so dass der durchschnittliche Konzentrationsindex der vier Attribute zu je dreien sich in der folgenden einfachen Gestalt ergibt:

$$\mu_3 = \frac{q}{3s - 2r + q}$$

Es bleibt endlich noch die Konzentration der vier Attribute zu vierten zu bestimmen. Um ein Mass für diese Konzentration zu finden, braucht man sich nur zu vergegenwärtigen, dass sinngemäss die Konzentration der vier Attribute um so grösser sein wird, je häufiger die Koexistenz derselben stattfindet, das heisst mit andern Worten, je grösser $(abcd)$ ist im Verhältnis zu dem Vorkommen der vier Attribute überhaupt. Das Mass für das Vorkommen der vier Attribute überhaupt ist aber

$$(a) + (b) + (c) + (d) - (ab) - (ac) - (ad) - (bc) - \\ - (bd) - (cd) + (abc) + (abd) + (acd) + (bcd) - \\ - (abcd)$$

Der Bau dieses Ausdruckes ist unschwer zu erklären. Jedes Element der Masse, dem eines oder gleichzeitig mehrere von den vier Attributen entsprechen, wird darin einmal und nur einmal mitgezählt. In der Tat: Diejenigen Elemente, denen nur eines von den vier Attributen, beispielsweise a , zukommt, werden unter (a) gezählt und sonst nirgends; diejenigen Elemente, denen gleichzeitig zwei und nur zwei von den Attributen entsprechen, also beispielsweise a und b , werden einmal bei (a) und dann wiederum bei (b) positiv und einmal bei (ab) negativ, also im ganzen nur einmal mitgezählt; ferner diejenigen Elemente, denen gleichzeitig drei und nur drei von den vier Attributen entsprechen, also beispielsweise a , b und c werden viermal positiv gezählt, nämlich unter (a) , unter (b) , unter (c) und unter (abc) , und dreimal negativ, nämlich unter (ab) , unter (ac) und unter (bc) , so dass sie ebenfalls im ganzen nur einmal erscheinen; endlich diejenigen Elemente, denen gleichzeitig alle vier Elemente zukommen, werden achtmal positiv gezählt, nämlich unter (a) , (b) , (c) , (d) , (abc) , (abd) , (acd) und (bcd) , und ferner siebenmal negativ, nämlich unter (ab) , (ac) , (ad) , (bc) , (bd) , (cd) und $(abcd)$, so dass auch diese im ganzen nur einmal in das Gesamtergebnis aufgehen. *q. e. d.*

Das Mass für das Vorkommen der vier Attribute überhaupt lässt sich, wenn wir von den bereits eingeführten Zeichen s , r , q Gebrauch machen und in Analogie damit weiter das Zeichen

$$p = (abcd)$$

introduzieren, auch so schreiben

$$s - r + q - p$$

so dass das Mass für die Konzentration der vier Attribute zu vierten sich in der folgenden Form ergibt:

$$\mu_4 = \frac{p}{s - r + q - p}$$

Zusammenfassend kann man also sagen: Untersucht man eine Masse in bezug auf die Konzentration von vier Attributen a , b , c und d , so ist zu unterscheiden zwischen dem

durchschnittlichen Konzentrationsindex der vier Elemente zu je zweien: μ_2 ;

ferner dem durchschnittlichen Konzentrationsindex der vier Elemente zu je dreien: μ_3 ;

und endlich dem Konzentrationsindex der vier Elemente zu vierten: μ_4 .

Führt man die Bezeichnungen

$$p, q, r \text{ und } s$$

durch die folgenden Definitionen ein:

$$p = (abcd)$$

$$q = (abc) + (abd) + (acd) + (bcd)$$

$$r = (ab) + (ac) + (ad) + (bc) + (bd) + (cd)$$

$$s = (a) + (b) + (c) + (d)$$

so schreiben sich die Formeln für μ_2 , μ_3 und μ_4 in eleganter Weise wie folgt:

$$\mu_2 = \frac{r}{3s - r}$$

$$\mu_3 = \frac{q}{3s - 2r + q}$$

$$\mu_4 = \frac{p}{s - r + q - p}$$

Nur kurz sei schliesslich angedeutet, wie man zu verfahren hat, falls man in der Masse fünf Merkmale unterscheidet und die Konzentration nach diesen untersucht. Die fünf Merkmale seien bzw. a , b , c , d und e . Wir werden diese zunächst zu Paaren zusammenfassen und die Konzentration der fünf Attribute zu je zweien bestimmen. Wir werden darnach insgesamt 10 Konzentrationsindizes erhalten, und zwar je einen für

$$(ab), (ac), (ad), (ae)$$

$$(bc), (bd), (be)$$

$$(cd), (ce)$$

$$(de)$$

Aus diesen 10 Indizes können wir dann ebenso wie oben einen geeigneten gewogenen Durchschnitt ziehen, und wir erhalten danach den durchschnittlichen Konzentrationsindex der fünf Attribute zu je zweien.

Ferner können wir die fünf Attribute in Gruppen zu je dreien zusammenfassen und die Konzentration für jede einzelne Gruppe bestimmen. Wir werden so insgesamt 10 Konzentrationsindizes erhalten, und zwar je einen für

$$(abc), (abd), (abe), (acd), (ace), (ade)$$

$$(bcd), (bce), (bde)$$

$$(cde)$$

Auch aus diesen 10 Indizes können wir einen geeigneten Durchschnitt ziehen, welcher dann als der durchschnittliche Konzentrationsindex der fünf Attribute zu je dreien anzusprechen sein wird.

Weiter können wir die fünf Attribute in Gruppen zu je vierten zusammenfassen und die Konzentration zu vierten für jede einzelne Gruppe bestimmen. Wir erhalten so insgesamt fünf Konzentrationsindizes, und zwar je einen für

$$(abcd), (abce), (abde), (acde), (bcde)$$

Ziehen wir aus diesen fünf Indizes in geeigneter Weise den gewogenen Durchschnitt, so erhalten wir einen Ausdruck, der als Konzentrationsindex der fünf Attribute zu je vieren anzusprechen sein wird.

Endlich können wir noch einen Konzentrationsindex der fünf Attribute zu fünfen bestimmen, indem wir in genau derselben Weise verfahren, wie wir früher bei der Aufstellung des Konzentrationsindex der vier Attribute zu vieren bzw. der drei Attribute zu dreien vorgegangen sind.

Es bietet keine Schwierigkeiten, diese Betrachtungen auf den Fall von sechs, sieben und mehr Attributen zu erweitern. Bei einiger Übung gelingt es un schwer, die betreffenden Formeln direkt hinzuschreiben.

4.

Gehen wir noch einmal auf den relativ einfachen, in Abschnitt 1 entwickelten Fall dreier Attribute a , b und c zurück. Wir haben da zu unterscheiden zwischen der durchschnittlichen Konzentration der drei Attribute zu je zweien und zwischen der Konzentration derselben zu dreien; für die erstere ergab sich das Mass in der Formel

$$\mu_2 = \frac{r}{2s - r}$$

und für die zweite in der Formel

$$\mu_3 = \frac{q}{s - r + q}$$

wobei die Zeichen q , r und s der Reihe nach bedeuten:

$$s = (a) + (b) + (c)$$

$$r = (ab) + (ac) + (bc)$$

$$q = (abc)$$

Es fragt sich nun, welcher Zusammenhang zwischen μ_2 und μ_3 besteht. Zu diesem Zwecke bilden wir die Differenz

$$\mu_2 - \mu_3 = \frac{r}{2s - r} - \frac{q}{s - r + q}$$

und beobachten, dass diese Differenz sich wie folgt umformen lässt:

$$\begin{aligned} \mu_2 - \mu_3 &= \frac{r(s - r + q) - q(2s - r)}{(2s - r)(s - r + q)} \\ &= \frac{rs - r^2 + rq - 2qs + rq}{(2s - r)(s - r + q)} \\ &= \frac{2q(r - s) - r(r - s)}{(2s - r)(s - r + q)} \\ &= \frac{(s - r)(r - 2q)}{(2s - r)(s - r + q)} \end{aligned}$$

Der zuletzt hingeschriebene Ausdruck ist relativ leicht zu behandeln und gestattet unmittelbar die Beantwortung der Frage, ob und in welchen Fällen μ_2 und μ_3 zusammenfallen können. Dies wird dann und nur dann der Fall sein, wenn entweder

$$s - r = 0$$

oder aber

$$r - 2q = 0$$

ist, also mit andern Worten, wenn

$$r = s$$

oder

$$r = 2q$$

ist. Falls nun die Merkmale nicht zusammenfallen, ist aber jederzeit

$$(ab) < (a)$$

da die Zahl der Elemente der Grundmasse, in denen die beiden Attribute a und b sich vereinigt finden, kleiner sein muss als die Zahl der Elemente, die nur das Attribut a aufweisen (ansonsten wäre b nur ein Teilattribut von a). Ebenso ist

$$(ac) < (c)$$

und

$$(bc) < (b)$$

Aus diesen drei Ungleichungen folgt aber

$$(ab) + (ac) + (bc) < (a) + (b) + (c)$$

oder

$$s < r$$

Analog ist

$$(abc) < (ab)$$

und

$$(abc) < (ac)$$

folglich a fertiori

$$2(abc) < (ab) + (ac) + (bc)$$

oder

$$2q < r$$

Aus diesen Ungleichungen folgt also, dass der durchschnittliche Konzentrationskoeffizient der drei Attribute zu zweien (μ_2) und der Konzentrationskoeffizient der drei Attribute zu dreien nicht zusammenfallen können, falls die drei Attribute a , b und c nicht mit einander zusammenfallen, so dass etwa alle Elemente mit den Merkmalen b und c in den Elementen, die das Merkmal a enthalten, aufgehen usf.

Aus dem Ausdruck

$$\mu_2 - \mu_3 = \frac{(s - r)(r - 2q)}{(2s - r)(s - r + q)}$$

lässt sich aber noch mehr herauslesen. Zunächst wird man beobachten, dass sowohl die beiden Faktoren des Zählers als auch die beiden Faktoren des Nenners positiv sind; daher folgt, dass auch die Differenz $\mu_2 - \mu_3$ positiv ist, oder mit andern Worten, dass

$$\mu_2 > \mu_3$$

ist.

Bei gegebenem r und s ist μ_2 eine gegebene Grösse, während μ_3 noch von q abhängig ist; und zwar ist μ_3 mit q steigend und fallend. Man kann diese Tatsache einfach wie folgt in Worte fassen: Je grösser die Konzentration der drei Attribute zu dreien ist, um so kleiner ist die stets positive Differenz zwischen dem durchschnittlichen Konzentrationskoeffizienten der drei Attribute zu zweien und dem Konzentrationskoeffizienten der drei Attribute zu dreien.

Auch die zuletzt entwickelten Sätze lassen interessante und leicht zu erweisende Verallgemeinerungen auf den Fall von mehr als drei Attributen zu.

5.

Ein Gegenstück zu dem eingangs dieses Aufsatzes aufgestellten Begriffe der Konzentration liegt in dem Begriffe der Isolierung. Beschränken wir uns auf zwei Attribute a und b . Das Mass für das Vorkommen des Merkmals a überhaupt ist (a) , und das Mass für das isolierte Vorkommen von a ist offensichtlich

$$(a) - (ab)$$

Wir werden also als Indexziffer für das isolierte Vorkommen der a ohne weiteres den Quotienten

$$\frac{(a) - (ab)}{(a)}$$

hinschreiben können. Analog wird die Masszahl für das isolierte Vorkommen von b lauten:

$$\frac{(b) - (ab)}{(b)}$$

Auch hier entsteht weiterhin die Aufgabe, aus diesen beiden Masszahlen eine geeignete Durchschnittsziffer zu bilden, welche als gleichzeitiges Mass für das isolierte Vorkommen von a und b zu betrachten wäre. Wir nehmen also die Existenz von zwei Gewichten g_1 und g_2 an und erhalten dann den gesuchten Durchschnitt in der Form:

$$\frac{1}{g_1 + g_2} \left\{ g_1 \frac{(a) - (ab)}{(a)} + g_2 \frac{(b) - (ab)}{(b)} \right\}$$

Was nun die Bestimmung von g_1 anlangt, so liegt auf der Hand, dass es sich zweckmässigerweise empfiehlt, die Masszahl für das isolierte Vorkommen von a nur im Verhältnis zu dem Gesamtvorkommen von a in den

gemeinsamen Durchschnitt eingehen zu lassen, das ist mit andern Worten, $g_1 = (a)$ zu wählen. Aus demselben Grunde bestimmt sich $g_2 = (b)$ und wir erhalten den Isolierungskoeffizienten der zwei Attribute, den wir mit μ_1 bezeichnen mögen, in der folgenden Form:

$$\begin{aligned} \mu_1 &= \frac{(a) + (b) - 2(ab)}{(a) + (b)} \\ &= 1 - 2 \frac{(ab)}{(a) + (b)} \end{aligned}$$

Andererseits schreibt sich der Konzentrationskoeffizient der beiden Attribute zu zweien, wie oben gezeigt, wie folgt:

$$\mu_2 = \frac{(ab)}{(a) - (ab) + (b)}$$

Zwischen μ_1 und μ_2 besteht nun eine interessante Beziehung. Indem man die Gleichung für μ_1 umformt, findet man nämlich der Reihe nach

$$\begin{aligned} 1 - \mu_1 &= 2 \cdot \frac{(ab)}{(a) + (b)} \\ \frac{(a) + (b)}{(ab)} &= \frac{2}{1 - \mu_1} \end{aligned}$$

Und ebenso ergibt sich durch Umformung der Gleichung für μ_2 :

$$\begin{aligned} \frac{1}{\mu_2} &= \frac{(a) - (ab) + (b)}{(ab)} \\ &= \frac{(a) + (b)}{(ab)} - 1 \end{aligned}$$

Indem man nun die beiden Ausdrücke für

$$\frac{(a) + (b)}{(ab)}$$

einander gleich setzt, findet man:

$$\begin{aligned} \frac{2}{1 - \mu_1} &= 1 + \frac{1}{\mu_2} \\ &= \frac{1 + \mu_2}{\mu_2} \end{aligned}$$

oder weiterhin

$$\frac{(1 - \mu_1)(1 + \mu_2)}{2 \mu_2} = 1$$

Durch Umformung findet man auch

$$1 - \mu_1 = \frac{2 \mu_2}{1 + \mu_2}$$

und

$$\mu_1 = \frac{1 - \mu_1}{1 + \mu_2}$$

Diese letztere Gleichung stellt die gesuchte Beziehung zwischen μ_1 und μ_2 dar.

Etwas komplizierter gestalten sich die Verhältnisse im Falle von drei Attributen, a , b und c . Das Mass für das isolierte Vorkommen der a wird lauten:

$$\frac{(a) - (ab) - (ac) + (abc)}{(a)}$$

Der Zähler dieses Quotienten erklärt sich so: von den (a) kommt ein Teil gleichzeitig mit dem Attribut b und ein anderer Teil gleichzeitig mit dem Attribut c vor; will man das Mass für das isolierte Vorkommen der a haben, so muss von (a) also zunächst (ab) und (ac) abgezogen werden; hierbei werden jedoch diejenigen unter den ab , welche gleichzeitig auch c sind, und diejenigen unter den ac , welche gleichzeitig auch b sind, zweimal in Abzug gebracht; daher muss zu dieser Differenz noch ein (abc) hinzuaddiert werden, und wir erhalten so den Zähler des obigen Ausdrucks. Da das Mass für das Vorkommen des Merkmals a überhaupt (a) ist, so muss (a) noch als Nenner darunter geschrieben werden.

Analog erhalten wir das Mass für das isolierte Vorkommen der b in der Form:

$$\frac{(b) - (ab) - (bc) + (abc)}{(b)}$$

und endlich das Mass für das isolierte Vorkommen der c in der Gestalt:

$$\frac{(c) - (ab) - (ac) + (abc)}{(c)}$$

Wir können nun in Analogie zu oben aus diesen drei Masszahlen für das isolierte Vorkommen der Merkmale a , b und c eine geeignete Durchschnittsziffer bilden, welche als gleichzeitiges Mass für das isolierte Vorkommen von a , b und c zu betrachten sein wird. Wir nehmen wiederum die Existenz von drei Gewichten g_1 , g_2 und g_3 an, und erhalten dann den gewünschten Durchschnitt in der folgenden Form:

$$\frac{1}{g_1 + g_2 + g_3} \left\{ g_1 \cdot \frac{(a) - (ab) - (ac) + (abc)}{(a)} + g_2 \cdot \frac{(b) - (ab) - (bc) + (abc)}{(b)} + g_3 \cdot \frac{(c) - (ac) - (bc) + (abc)}{(c)} \right\}$$

Es ist klar, dass es sich empfehlen wird, die Masszahl für das isolierte Vorkommen von a nur im Verhältnis zum Gesamtvorkommen von a in den gemeinsamen Durchschnitt eingehen zu lassen, also $g_1 = (a)$ zu setzen. Ebenso wählen wir $g_2 = (b)$ und $g_3 = (c)$ und erhalten schliesslich den gesuchten Durchschnitt in der folgenden Form:

$$\mu_1 = \frac{(a) + (b) + (c) - 2(ab) - 2(ac) - 2(bc) + 3(abc)}{(a) + (b) + (c)}$$

Wenn wir von unseren oben eingeführten Bezeichnungen für den Fall dreier Attribute

$$\begin{aligned} q &= (abc) \\ r &= (ab) + (ac) + (bc) \\ s &= (a) + (b) + (c) \end{aligned}$$

Gebrauch machen, so lässt sich μ_1 , das ist also der durchschnittliche Isolierungskoeffizient dreier Merkmale, auch so schreiben:

$$\mu_1 = \frac{s - 2r + 3q}{s}$$

Andererseits schreibt sich der durchschnittliche Konzentrationskoeffizient der drei Attribute zu je zweien

$$\mu_2 = \frac{r}{2s - r}$$

und der Konzentrationskoeffizient der drei Attribute zu dreien

$$\mu_3 = \frac{q}{s - r + q}$$

Zwischen μ_1 , μ_2 und μ_3 besteht nun eine Beziehung, die sich auf folgende einfache Weise ergibt:

Es ist

$$\frac{1}{\mu_3} = \frac{s - r + q}{q}$$

und

$$\frac{3q}{\mu_3} = 3(s - r + q) = 3s - 3r + 3q$$

und es ist ferner

$$\frac{1}{\mu_2} = \frac{2s - r}{r}$$

und

$$\frac{r}{\mu_2} = 2s - r$$

Die Differenz

$$\frac{3q}{\mu_3} - \frac{r}{\mu_2}$$

stellt sich also in der Form

$$s - 2r + 3q$$

dar, und wir erkennen ohne weiteres, dass die Beziehung besteht:

$$\mu_1 = \frac{1}{s} \left(\frac{3q}{\mu_3} - \frac{r}{\mu_2} \right)$$

Die Formeln gestatten uns, den Isolierungskoeffizienten zu berechnen, falls wir die beiden Konzentrationskoeffizienten kennen. Wir kommen darauf im nächsten Abschnitt zurück.

6.

Knüpfen wir nochmals an den Fall zweier Attribute a und b in der gegebenen Grundmasse an. Das Mass für das Vorkommen von a war mit (a) und das Mass für das Vorkommen von b mit (b) bezeichnet worden. Dementsprechend wollen wir Zeichen für das isolierte Vorkommen eines oder mehrerer Attribute einführen, und zwar soll (A) das Mass für das isolierte Vorkommen von a und (B) das Mass für das isolierte Vorkommen von b , sowie endlich (AB) das Mass für das gleichzeitige isolierte (d. h. von weiteren Merkmalen freie) Vorkommen der beiden Attribute a und b bedeuten. Dann ist noch den Ausführungen in Abschnitt 5:

$$\begin{aligned}(a) &= (A) + (AB) \\ (b) &= (B) + (AB) \\ (ab) &= (AB)\end{aligned}$$

und umgekehrt:

$$\begin{aligned}(A) &= (a) - (ab) \\ (B) &= (b) - (ab) \\ (AB) &= (ab)\end{aligned}$$

In den neuen Zeichen drückt sich der Konzentrationskoeffizient der beiden Attribute, der oben mit

$$\frac{(ab)}{(a) - (ab) + (b)}$$

ermittelt worden war, nunmehr so aus:

$$\mu_2 = \frac{(AB)}{(A) + (AB) + (B)}$$

Der Isolierungskoeffizient μ_1 andererseits geht mit Hülfe der neuen Bezeichnungen über in:

$$\mu_1 = \frac{(A) + (B)}{(A) + 2(AB) + (B)}$$

Gehen wir nun auf den Fall dreier Attribute a , b und c über und halten an den neuen Bezeichnungen fest, so wird beispielsweise (A) das Mass für das isolierte Vorkommen von a , (AB) , das Mass für das isolierte gleichzeitige Vorkommen der beiden Attribute a und b usw. bedeuten. Danach werden sich die alten Bezeichnungen in den neuen wie folgt ausdrücken:

$$\begin{aligned}(a) &= (A) + (AB) + (AC) + (ABC) \\ (b) &= (B) + (AB) + (BC) + (ABC) \\ (c) &= (C) + (AC) + (BC) + (ABC) \\ (ab) &= (AB) + (ABC) \\ (ac) &= (AC) + (ABC) \\ (bc) &= (BC) + (ABC) \\ (abc) &= (ABC)\end{aligned}$$

Umgekehrt drücken sich die neuen Bezeichnungen mit Hülfe der alten in der nachstehenden Weise aus:

$$\begin{aligned}(A) &= (a) - (ab) - (ac) + (abc) \\ (B) &= (b) - (bc) - (ab) + (abc) \\ (C) &= (c) - (ac) - (bc) + (abc) \\ (AB) &= (ab) - (abc) \\ (AC) &= (ac) - (abc) \\ (BC) &= (bc) - (abc) \\ (ABC) &= (abc)\end{aligned}$$

Die neuen Bezeichnungen eignen sich — was nach ihrer Entstehung ohne weiteres verständlich sein dürfte — weniger gut zur Darstellung der Konzentrationskoeffizienten, allein sie leisten zur Darstellung der Isolierungskoeffizienten bessere Dienste als die alten. So schrieb sich, im Falle dreier Attribute, der Isolierungskoeffizient mit den alten Bezeichnungen in der Form

$$\mu_1 = \frac{(a) + (b) + (c) - 2(ab) - 2(ac) - 2(bc) + 3(abc)}{(a) + (b) + (c)}$$

während sich diese Formel an der Hand der neuen Bezeichnungen wie folgt darstellt:

$$\mu_1 = \frac{(A) + (B) + (C)}{(A) + (B) + (C) + 2(AB) + 2(BC) + 2(AC) + 3(ABC)}$$

Aus dieser Art der Darstellung geht ohne weiteres hervor, dass der Isolierungskoeffizient um so grösser ist, je grösser ceteris paribus (A) , (B) und (C) und je kleiner (AB) , (AC) , (BC) und (ABC) ausfallen.

Berechnen wir den Isolierungskoeffizient in den vier oben (Abschnitt 2) angeführten Beispielen, so finden wir die folgenden Zahlenwerte:

- A) $\mu_1 = 11.8 \%$
- B) $\mu_2 = 5.8 \%$
- C) $\mu_3 = 13.2 \%$
- D) $\mu_4 = 20.4 \%$

Das Resultat ist dem für die Konzentrationskoeffizienten errechneten konform: im Jahre 1896 war die Isolierung der drei Börsenplätze Basel, Genf und Zürich sowohl in bezug auf das kotierte schweizerische Obligationenmaterial als auch in bezug auf das kotierte schweizerische Aktienmaterial ungefähr gleich gross. Von da bis zum Jahre 1913 hat sich die Isolierung der Obligationen verkleinert, die der Aktien dagegen vergrössert. Die Gründe sind natürlich dieselben, wie die oben für die umgekehrte Bewegung der Konzentrationskoeffizienten dargelegten.

7.

Wir haben uns im obigen hauptsächlich mit den Merkmalen der Konzentration und Isolierung von Attributen beschäftigt, während die einschlägige Literatur in erster Linie das Merkmal der Abhängigkeit bzw. Unabhängigkeit der Attribute von einander ins Auge

fasst. Knüpfen wir wiederum an den Fall zweier Attribute a und b an, deren Vorkommen in der Masse N durch die Zahlen (a) und (b) gemessen werden möge. In allen jenen Elementen der Masse, in denen das Merkmal a nicht vorhanden ist, ist das Merkmal „non- a “ vorhanden, und die Häufigkeit des Vorkommens von non- a soll durch α bezeichnet werden; analog möge mit β die Häufigkeit des Vorkommens von „non- b “ ausgedrückt werden.

Das Merkmal der Unabhängigkeit der beiden Attribute a und b wird nun gemeinhin daran erkannt, dass das Attribut b unter den a ebenso häufig vertreten ist wie unter den non- a . Nun ist die Masszahl für das Vorkommen des Attributes b unter den a absolut gleich (ab) und relativ gleich $\frac{(ab)}{(a)}$, und ebenso die Masszahl für das Vorkommen des Attributes b unter den non- a absolut gleich (αb) und relativ gleich $\frac{(\alpha b)}{(\alpha)}$. Sollen also die beiden Attribute unabhängig sein, so muss die Beziehung statt haben

$$\frac{(ab)}{(a)} = \frac{(\alpha b)}{(\alpha)}$$

Aus dieser Beziehung lässt sich ohne weiteres folgern, dass auch das Attribut a unter den b ebenso häufig vertreten sein muss wie unter den non- b , also

$$\frac{(ab)}{(b)} = \frac{(\alpha \beta)}{(\beta)}$$

Für den Fall, dass in der Masse überhaupt nur die beiden Attribute a und b vorkommen, so dass jedes Element derselben entweder das Merkmal a oder das Merkmal b , oder beide zusammen besitzt, nimmt die obige Beziehung eine besonders einfache Form an: da alle Elemente, welche das Merkmal non- a aufweisen, eo ipso das Merkmal b besitzen müssen, ist

$$(ab) = (\alpha)$$

und ebenso gilt

$$(a\beta) = (\beta)$$

da ja alle Elemente der Masse, welche das Merkmal non- b zeigen, auch das Merkmal a besitzen müssen. In diesem Falle ist also

$$(a) = (b) = (ab)$$

Der oben berechnete Konzentrationskoeffizient der beiden Attribute, für den wir die Formel

$$\frac{(ab)}{(a) - (ab) + (b)}$$

aufgestellt hatten, wird für den Fall der Unabhängigkeit der beiden Attribute gleich 1, und ebenso wird der Isolierungskoeffizient, für den die Formel

$$\frac{(a) + (b) - 2(ab)}{(a) + (b)}$$

gewonnen worden war, in diesem speziellen Falle gleich Null.

Allein die ganze Problemstellung für den Fall, dass die beiden Attribute a und b die ganze Masse ausfüllen, hat bloss formales Interesse, da ja schon die Beziehung $(a) = (ab) = (b)$ besagt, dass in diesem Falle beide Attribute zusammenfallen müssen. Nehmen wir deshalb den allgemeineren Fall vor, dass in der Masse drei und nur drei Attribute a , b und c vorkommen, so zwar, dass jedes Element derselben entweder das Merkmal a , oder das Merkmal b , oder das Merkmal c , oder ein Paar von den drei Merkmalen, oder aber alle drei enthält. Sollen die drei Merkmale paarweise untereinander unabhängig sein, so werden, falls man das Merkmal „non- a “ mit α , das Merkmal „non- b “ mit β und das Merkmal „non- c “ mit γ bezeichnet, die drei Gleichungen gleichzeitig bestehen müssen:

$$\frac{(ab)}{(a)} = \frac{(\alpha b)}{(\alpha)}$$

$$\frac{(ac)}{(a)} = \frac{(\alpha c)}{(\alpha)}$$

$$\frac{(bc)}{(b)} = \frac{(\beta c)}{(\beta)}$$

Die Aufgabe besteht nun darin, die Grössen (a) , (β) , (ab) , (αc) und (βc) hinzuschreiben.

Zunächst (a) , die Masszahl für non- a . Non- a bedeutet aber mit anderen Worten das Vorkommen des Merkmals b allein, oder des Merkmals c allein, oder endlich der beiden Merkmale b und c allein zusammengenommen. Nun ist das Mass für das Vorkommen von b allein gleich

$$(b) - (ab) - (bc) + (abc)$$

ferner das Mass für das Vorkommen von c allein

$$(c) - (ac) - (bc) + (abc)$$

und endlich das Mass für das gemeinsame Vorkommen von b und c allein gleich

$$(bc) - (abc)$$

Addiert man diese drei Bestandteile zusammen, so findet man, da sich mehrere Glieder aufheben

$$(a) = (b) + (c) - (ab) - (bc) - (ac) + (abc)$$

Ebenso lässt sich (β) , die Masszahl für non- b darstellen. Denn non- b ist nichts anderes als das Vorkommen des Merkmals a allein, oder des Merkmals c allein, oder endlich der beiden Merkmale a und c allein zusammengenommen. Wie oben haben wir nun das Mass für das Vorkommen von a allein gleich

$$(a) - (ab) - (ac) + (abc)$$

ferner das Mass für das Vorkommen von c allein gleich

$$(c) - (ac) - (bc) + (abc)$$

und endlich das Mass für das gemeinsame Vorkommen von a und c allein gleich

$$(ac) - (abc)$$

Durch Addition dieser drei Bestandteile findet man

$$(\beta) = (a) + (c) - (ab) - (bc) - (ac) + (abc)$$

Die Differenz $(\alpha) - (\beta)$ nimmt danach eine sehr einfache Form an:

$$(\alpha) - (\beta) = (b) - (a)$$

Es bleibt noch übrig, die Ausdrücke für (ab) und (ac) hinzuschreiben. Aus (a) wird (ab) , indem wir in der Formel für (a) bei allen Bestandteilen, welche b noch nicht enthalten, dieses hinzufügen; daraus wird:

$$\begin{aligned} (ab) &= (b) + (bc) - (ab) - (bc) - (abc) + (abc) \\ &= (b) - (ab) \end{aligned}$$

und analog

$$\begin{aligned} (ac) &= (c) - (ac) \\ (\beta c) &= (c) - (bc) \end{aligned}$$

Damit erhalten wir die drei Bedingungen für die gegenseitige Unabhängigkeit der drei Merkmale einfach in der Form:

$$\begin{aligned} \frac{(ab)}{(a)} &= \frac{(b) - (ab)}{(b) + (c) - (ab) - (bc) - (ac) + (abc)} \\ \frac{(ac)}{(a)} &= \frac{(c) - (ac)}{(b) + (c) - (ab) - (bc) - (ac) + (abc)} \\ \frac{(bc)}{(b)} &= \frac{(c) - (bc)}{(a) + (c) - (ab) - (ac) - (bc) + (abc)} \end{aligned}$$

Aus diesen drei Gleichungen kann man nun, falls

$$(a), (b) \text{ und } (c)$$

gegeben sind, irgendein Tripel der vier übrigen Grössen

$$(ab), (ac), (bc), (abc)$$

berechnen. Eine von den letzteren wird aber unbekannt bleiben, und dieses ist der Grund, weshalb man aus der Voraussetzung allein, dass die Merkmale a , b und c paarweise unabhängig sind, nicht einen bestimmten Wert der Konzentrationsindizes oder des Isolierungskoeffizienten ableiten kann. Konzentration und Isolierung einerseits, Unabhängigkeit (bzw. Assoziation) andererseits sind vielmehr Dinge, die miteinander nur in losem Zusammenhang stehen.

8.

Für die Unabhängigkeit zweier Attribute a und b war oben etwa die Bedingung

$$\frac{(ab)}{(a)} = \frac{(ab)}{(a)}$$

benützt worden. Diese Formel lässt sich aber ohne weiteres in eine Reihe weiterer Formeln umformen, welche eindeutig aus ihr folgen, und aus denen sie umgekehrt selbst abgeleitet werden kann. Unter diesen Formeln ist die am häufigsten gebrauchte die folgende:

$$\frac{(ab)}{(N)} = \frac{(a)}{(N)} \cdot \frac{(b)}{(N)}$$

wo (N) das Mass der Grundmasse bedeutet. In allen Fällen, in denen die Unabhängigkeit der beiden Attribute nicht statt hat, ist entweder der rechte Teil der obigen Gleichung grösser als der linke oder aber umgekehrt; im ersteren Falle sprechen wir von einer positiven, im letzteren von einer negativen Assoziation der beiden Attribute. Zur Messung der Assoziation hat Udny Yule den folgenden Koeffizienten vorgeschlagen:

$$Q = \frac{(ab)(a\beta) - (a\beta)(ab)}{(ab)(a\beta) + (a\beta)(ab)}$$

Dieser Koeffizient ist gleich Null, wenn die beiden Attribute unabhängig sind, positiv (bis zum Höchstwerte $+1$), wenn die beiden Attribute positiv assoziiert sind, und negativ (bis zum Tiefstwerte -1), wenn sie negativ assoziiert sind. Wir wollen diese Formel umformen für den Fall, dass die ganze Grundmasse aus drei und nur drei Merkmalen a , b und c zusammengesetzt ist, und mithin in die folgenden Bestandteile zerfällt:

$$(A), (B), (C), (AB), (AC), (BC), (ABC)$$

Dann wird:

$$(ab) = (AB) + (ABC)$$

$(a\beta)$ ist die Koexistenz der beiden Attribute non- a und non- b und mithin nichts anderes als gleich (C) .

$$(a\beta) = (A) + (AC)$$

$$(ab) = (B) + (BC)$$

Der Zähler des Assoziationskoeffizienten Q wird danach gleich

$$\begin{aligned} (AB)(C) + (ABC)(C) - (A)(B) - (A)(BC) \\ - (AC)(B) - (AC)(BC) \end{aligned}$$

während der Nenner die folgende Form annimmt:

$$\begin{aligned} (AB)(C) + (ABC)(C) + (A)(B) + (A)(BC) \\ + (AC)(B) + (AC)(BC) \end{aligned}$$

Es ist eine rein rechnerische Aufgabe, diese Formel auf das in Abschnitt 2 angeführte numerische Beispiel (kotierte schweizerische Aktien und Obligationen an den schweizerischen Börsen) anzuwenden, auf welches wir nicht eingehen wollen. Der Leser, der sich der leichten Mühe unterzieht, wird aber zu einem recht belehrenden Ergebnis kommen.