1912. Myrmecophilous notes for 1911 (Entomologists Record, XXIV, Nr. 1—2) (Phora formicarum p. 36).
1914. Myrmecophilous notes for 1913 (Entomologists Record, XXVI, Nr. 2, p. 37—45) (Phora formicarum p 42).
Lubbock, John (Lord Avebury), 1883. Ameisen, Bienen und Wespen. Beobachtungen über die Lebensweise der geselligen Hymenopteren. Autoris. deutsche Ausgabe, Brockhaus, Leipzig (p. 22 u. 61; Beschreibung der Phora formicarum Verr. p. 370).
Schmitz, Hermann, 1913. Zusammenstellung der bis Ende 1913 beschriebenen myrmecophilen und termitophilen Phoriden (Jaarboek d. Natuur-hist. Genootsch. Limburg) (p. 6—7 Separ. Plastophora formicarum Verr.).
1914. Die myrmecophilen Phoriden der Wasmann'schen Sammlung (Zool. Jahrb. System. XXXVII, 6. Heft, p. 509—566 und Taf. 29 u. 30) (p. 532 u. 557 Plastophora formicarum[5])).
Strobl, Gabriel, 1910. Die Dipteren von Steiermark (Mitteil. Naturw. Ver. Steiermark XLVI, p. 45—293) (p. 125 erste Angabe über Phora formicarum auf dem Kontinent),
Verrall, 1877. (Erste Beschreibung von Phora formicarum) Journ. Linn. Soc. London XIII, p. 258.
Westwood, J. O., 1840. Introduction to the modern classification of Insects, Vol. II, p. 234 ff.
Wood, J. H., 1908. On the British Species of Phora. Part 2 (Entomol. Monthl. Mag. (2) XIX) p. 174.

Einige zahlenkritische Bemerkungen zu den Mendelschen Regeln.

Von P. Riebesell, Hamburg.

1. Aufgabe der vorliegenden Arbeit.

Bei der Entdeckung und der Wiederentdeckung der Mendelschen Regeln handelte es sich um die Deutung gewisser Zahlenverhältnisse, in denen bestimmte Eigenschaften von Tieren und Pflanzen bei der Kreuzung auftraten. Zur Erklärung für das Auftreten der meisten beobachteten Zahlenverhältnisse genügten die folgenden 4 Hypothesen: 1. Eindeutige Zuordnung von Erbfaktoren zu den äußeren Merkmalen, 2. Vorhandensein von Faktorenpaaren in den Zygoten, 3. Vollkommene Spaltung der Faktoren bei der Gametenbildung, 4. Anwendbarkeit der einfachsten Regeln der Wahrscheinlichkeitslehre. Später wurden jedoch Zahlenverhältnisse beobachtet, die sich mit den genannten 4 Grundhypothesen nicht mehr erklären ließen. Es soll im folgenden zunächst untersucht werden, ob die Mendelschen Regeln eine notwendige Folge aus den Beobachtungs-

5) Die Angabe p. 532, daß diese Art in der Wasmann'schen Sammlung nicht vertreten sei, ist zu berichtigen. Das oben (S. 318) erwähnte Exemplar von R. Handmann aus Bosnien hatte ich bei der Durchsicht der Sammlung für obige Arbeit von P. Schmitz übersehen und ihm erst später zugesandt.

ergebnissen und den genannten Hypothesen sind, vor allem ob bei der Deutung der beobachteten Zahlenverhältnisse die Regeln beachtet sind, die die Mathematik für die Ableitung von allgemeinen Gesetzen aus Beobachtungen aufgestellt hat. Sodann soll gezeigt werden, daß die zahlreichen Ergänzungshypothesen mit den Grundannahmen nicht mehr in Einklang stehen und eine eindeutige Erklärung der Tatsachen nicht ermöglicht wird.

2. Beobachtete und erwartete Zahlen.

Wenn aus einer vorliegenden Beobachtungsreihe unter Benutzung von Hypothesen ein quantitatives Gesetz abgeleitet wird, ist es unerläßlich, die Genauigkeit anzugeben, mit der das Gesetz die Beobachtungen wiedergibt, um dadurch einen Schluß zu ermöglichen, ob das gewonnene Gesetz den Anforderungen der Fehlertheorie genügt. Die Beobachtungsfehler selbst sollen hierbei unberücksichtigt bleiben und die beobachteten Zahlen als Tatsachen hingenommen werden. Sowohl in den meisten Originalarbeiten[1]) als auch in den Hauptwerken über Vererbungslehre[2]) werden nun Untersuchungen über die Größe der Fehler zwischen den beobachteten und den nach den abgeleiteten Gesetzen erwarteten Zahlen nicht angestellt.

Erst Johannsen und Lang[3]) haben darauf hingewiesen, wie wichtig zahlenkritische Untersuchungen der Ergebnisse sind. Sie haben aber als einziges Kriterium für die Brauchbarkeit der abgeleiteten Formeln den mittleren Fehler benutzt. Inzwischen hat I. A. Harris[4]) darauf aufmerksam gemacht, daß in vielen Fällen dieses Kriterium versagt, und er hat im Anschluß an eine Arbeit von K. Pearson[5]) eine neue Größe eingeführt, mit deren Hilfe die Mendelschen Regeln auf ihre Eignung zur Darstellung der Beobachtungen geprüft werden sollen. Es soll im folgenden untersucht werden, wie weit diese Formeln exakten Anforderungen entsprechen.

---

1) Vgl. z. B. R. C. Punnett, Reduplication Series in Sweet Peas, Journal of Genetics, Vol. III, Nr. 2, 1913. — G. H. Shull, Duplicate genes for capsuleform in Bursa bursa-pastoris. Ztschr. f. ind. Abst. u. Vererb.-Lehre, Bd. XII, Heft 2, 1914.

2) Vgl. z. B. W. Bateson, Mendels Vererbungstheorien. Leipzig 1914.

3) W. Johannsen, Elemente der exakten Erblichkeitslehre. 1. Aufl., Jena 1909, 2. Aufl., Jena 1913. — A. Lang, Experimentelle Vererbungslehre in der Zoologie seit 1900. 1. Hälfte, Jena 1914.

4) I. A. Harris, A simple test of the goodness of fit of Mendelian ratios. The American Naturalist, Vol. 46, 1912.

5) K. Pearson, On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. Phil. Magazine, Vol. 50, 1900.

3. Die Binonialformel.

Sind m genotypische Differenzpunkte (Faktorenpaare) vorhanden, so ergeben sich die erwarteten Häufigkeiten in der F_2-Generation aus der Formel:

$$\left(\frac{1}{4}+\frac{2}{4}+\frac{1}{4}\right)^m. \tag{1}$$

Herrscht bei allen Faktorenpaaren vollkommene Dominanz, so lautet die Formel:

$$\left(\frac{3}{4}+\frac{1}{4}\right)^m. \tag{2}$$

Für m = 1 ergeben sich aus (2) die Werte: 3 : 1,
für m = 2: 9 : 3 : 3 : 1,
für m = 3: 27 : 9 : 9 : 9 : 3 : 3 : 3 : 1.

Es hat sich nun aber gezeigt, daß in sehr vielen Fällen, soweit ein äußeres Merkmal ins Auge gefaßt wurde, sich nicht die Zahlenverhältnisse 1 : 2 : 1 oder 3 : 1 ergaben, und man stellte die Theorie auf, daß eine Eigenschaft durch mehrere Faktoren bestimmt wird, die unabhängig voneinander spalten. Die Zahlenverhältnisse für m = 1, m = 2 u. s. w. sollen daher nicht nur für mehrere Merkmalspaare gelten, sondern je nach der Art, wie die Faktoren sich zu einer äußerlich erkennbaren Eigenschaft zusammensetzen, ergeben sich für m = 2 und m = 3 u. s. w. Zahlen, die sich auch auf ein Merkmal beziehen können,

z. B. für m = 2: 9 : 3 : 4, 9 : 6 : 1, 9 : 7, 15 : 1 u. s. w.
und für m = 3: 27 : 37, 55 : 9 u. s. w.

4. Die Bestimmung der Faktorenzahl.

Gehen wir zunächst von einem äußeren Merkmal aus und liegt ein beobachtetes Zahlenverhältnis $n_1 : n_2$ vor, so fragt es sich, welcher Mendelsche Bruch ihm entspricht. Läßt man nur die einfachste der vielen Möglichkeiten zu, daß nämlich zwei verschiedene Phänotypen vorhanden sind, von denen der eine durch das Vorhandensein sämtlicher dominanten Faktoren bedingt ist und der andere in allen übrigen Fällen auftritt, so müßte die Gleichung gelten:

$$\frac{3^m}{4^m - 3^m} = \frac{n_1}{n_2}. \tag{3}$$

Aus dieser Exponentialgleichung ist m auf einfache Weise zu berechnen. Es ergibt sich:

$$m = \frac{\log n - \log n_1}{\log 4 - \log 3}, \tag{4}$$

wo $n_1 + n_2 = n$ gesetzt ist.

Es sind natürlich nur ganzzahlige m zu gebrauchen, und es müßte jedesmal untersucht werden, ob die dem errechneten m benachbarte ganze Zahl für die erwarteten Häufigkeiten Werte ergibt, deren Fehlergrenzen die beobachteten Zahlen umfassen. Daß das im allgemeinen der Fall sein wird, geht aus den folgenden Betrachtungen hervor.

Zunächst soll an einem Beispiel gezeigt werden, daß die Formel (4) bessere Ergebnisse liefert, als die bisher angewandte Methode des Probierens. Es sollen die von Johannsen[6]) gegebenen Zahlen von Miss Saunders genommen werden. Bei der Kreuzung gewisser weißer und cremefarbiger Levkoyenrassen mit ungefärbtem Zellsaft war F_1 saftgefärbt. In F_2 waren auf 223 Individuen 128 saftgefärbt und 95 saftfarblos. Für m ergibt sich nach der obigen Formel:

$$m = \frac{\log 223 - \log 95}{\log 4 - \log 3} = 2{,}996.$$

Die nächste ganze Zahl ist 3, und die erwarteten Häufigkeiten wären 27 : 37. Die Abweichung ist in der Tat denkbar gering, da

$$\frac{95}{223} - \frac{27}{64} = 0{,}004 \text{ ist.}$$

Nun hat Miss Saunders aus biologischen Gründen, wegen des Auftretens des Farbfaktors neben dem Saftfaktor, $m = 2$ angenommen und demnach als erwartete Zahlen 7 : 9 zugrunde gelegt. In diesem Falle beträgt aber die Abweichung das Dreifache, da

$$\frac{7}{16} - \frac{95}{223} = 0{,}012 \text{ ist.}$$

Die Größe der Abweichung des beobachteten Verhältnisses vom erwarteten könnte in erster Annäherung als Genauigkeitsmaß angesehen werden.

5. Der mittlere Fehler.

Sind e_1 und e_2 die erwarteten Zahlen und ist n die Gesamtzahl der Beobachtungen, so ist der mittlere Fehler f (berechnet pro e) gegeben durch die Formel:

$$f = \sqrt{\frac{e_1 \cdot e_2}{n}}. \tag{5}$$

Ganz abgesehen davon, daß diese Formel, wie in der biologischen Literatur Harris betont hat, nur gilt, wenn n eine große Zahl ist und $\frac{e_1}{e}$ sowie $\frac{e_2}{e}$, $(e = e_1 + e_2)$, nicht weit von $\frac{1}{2}$ abweichen, muß bei der Anwendung der Formel auf das vorliegende Problem noch etwas Grundsätzliches berücksichtigt werden. Die Formel sagt zunächst nur aus, daß bei einer großen Zahl von Be-

6) A. a. O. (2. Aufl.), S. 506, 513.

obachtungen über Ereignisse, denen a priori die Wahrscheinlichkeiten $\frac{e_1}{e}$ bezw. $\frac{e_2}{e}$ zukommen, die mittlere Abweichung von dem wahrscheinlichsten Wert f beträgt. Daraus darf man aber nicht ohne weiteres schließen, daß, wenn bei einem beobachteten Zahlenverhältnis der Fehler unter dem nach Formel (5) berechneten liegt, daraus ein Schluß auf die Richtigkeit der vermuteten Zahlen e_1 und e_2 gezogen werden kann. Es ist ohne Schwierigkeit zu sehen, daß nach den Formeln (3) und (5) für ein gegebenes Verhältnis $n_1 : n_2$ zahlreiche Werte von m, d. h. von e_1 und e_2, zulässig sind. Das ist ohne weiteres aus der im 8. Kapitel abgeleiteten Gleichung (13) ersichtlich.

In dem angezogenen Beispiel ergibt sich für die Annahme 27 : 37 der Fehler $f = 2,12$. Rechnen wir die von Miss Saunders gefundenen Werte auf die Kombinationszahl $e = 4^3 = 64$ um, so ergibt sich das Verhältnis $27,26 : 36,74 = (27 + 0,26) : (37 - 0,26)$. Die Abweichungen liegen also weit unter dem mittleren Fehler. Für die Annahme 7 : 9 ergibt sich für f der Wert 0,53; und die beobachteten Zahlen lauten, wenn sie auf $4^2 = 16$ umgerechnet werden, $6,82 : 9,18 = (7 - 0,18) : (9 + 0,18)$. Auch hier bleiben daher, wenn auch nicht so weit wie vorher, die Fehler unterhalb des mittleren Wertes.

In zweiter Annäherung könnte das Verhältnis der beobachteten Abweichung zur mittleren Abweichung als Genauigkeitsmaß dienen.

Wie das Beispiel zeigt, sind also nach dem Kriterium des mittleren Fehlers allein beide Annahmen zulässig, und wenn auch diejenige mit 2 Faktorenpaaren die einfachere ist, so gibt diejenige mit 3 Faktoren die Möglichkeit, auf mehrfache Weise das Zustandekommen der äußeren Merkmale aus dem Genotypus zu erklären, da sie ja außer dem Verhältnis 27 : 37 z. B. noch die Möglichkeit 28 : 36 zuläßt, abgesehen davon, daß sie für den ersten Fall die geringsten Fehler liefert.

Die Ursache für die an dem Beispiel erläuterte Eigentümlichkeit ist darin zu sehen, daß die verschiedenen auf Grund der Binomialformel (2) sich ergebenden Mendelschen Brüche so nahe beieinander liegen, daß der Fehlerbereich des einen den andern mit umfaßt. Nehmen wir z. B. wieder die Verhältnisse 7 : 9 und 27 : 37 und fragen, ob der zweite Wert innerhalb des mittleren Fehlers des ersten liegen kann. Es muß dann sein:

$$\left|\frac{37}{64}\cdot 16 - 9\right| < \sqrt{\frac{9\cdot 7}{n}}.$$

Daraus ergibt sich:

$$n < 31,$$

d. h. so lange die Versuchszahl unter 31 bleibt, ist zwischen den beiden Möglichkeiten nicht zu unterscheiden. Daß dann Zahlenwerte, die zwischen den beiden Verhältnissen liegen, noch für eine größere Anzahl von Beobachtungen innerhalb der Fehlergrenze liegen können, versteht sich von selbst. Noch auffälliger wird dies, wenn wir zu größeren Werten von e_1 und e_2 übergehen. Kehren wir das obige Beispiel um und fragen, wie lange der Wert 7 : 9 innerhalb der mittleren Fehlergrenze von 27 : 37 bleibt, so ergibt sich:

$$\left|\frac{9}{16}\cdot 64 - 37\right| < \sqrt{\frac{37\cdot 27}{n}},$$

d. h. $n < 999$.

6. Die exakte Bestimmung des geeignetsten Mendelschen Bruches.

Soll ich zu einem gegebenen Zahlenverhältnis den geeignetsten Mendelschen Bruch suchen, und lasse ich zunächst alle biologischen Gesichtspunkte außer Acht, so ist vom rein mathematischen Standpunkt aus folgendermaßen zu verfahren: Es ist die Wahrscheinlichkeit zu bestimmen, mit der sich das beobachtete Verhältnis unter Zugrundelegung des erwarteten Gesetzes als Versuchsergebnis nach der Fehlertheorie ergeben würde. Es sind hierbei drei Wege gangbar.

a) Zunächst kann das sogenannte Hauptproblem der aposteriorischen Wahrscheinlichkeitsrechnung benutzt werden [7]).

Hat sich aus n Beobachtungen das Zahlenverhältnis $n_1 : n_2$ ergeben, so ist die Wahrscheinlichkeit, daß sich in $e = e_1 + e_2$ weiteren Versuchen das Verhältnis $e_1 : e_2$ ergibt, durch folgenden Wert gegeben:

$$W = \frac{e!\,(n_1 + e_1)!\,(n_2 + e_2)!\,(n + 1)!}{e_1!\,e_2!\,(n + e + 1)!\,n_1!\,n_2!}. \tag{6}$$

Die aposteriorische Wahrscheinlichkeit des Ereignisses selbst wird:

$$W_1 = \frac{n_1 + 1}{n + 2}, \tag{7}$$

ihr Unterschied von der wahrscheinlichsten Hypothese $\frac{n_1}{n}$ ist also

$$\frac{n_1}{n} - \frac{n_1 + 1}{n + 2} = \frac{2\,n_1 - n}{n\,(n + 2)}. \tag{8}$$

Wenden wir diese Formeln auf unser Beispiel an, so ergibt sich mit Hilfe der Stirlingschen Formel aus (6) für die Annahme 7 : 9 (für $e = 64$ berechnet)

$$W = 0{,}088$$

und für die Annahme 27 : 37

$$W = 0{,}089.$$

7) Vgl. z. B. E. Czuber, Wahrscheinlichkeitsrechnung, 3. Aufl., Leipzig 1914, Bd. I, S. 219.

Wir sehen also auch hier wieder, daß die letztere Annahme die größere Wahrscheinlichkeit hat, während nicht ohne weiteres in allen Fällen der kleineren Abweichung auch die größere Wahrscheinlichkeit zukommt. Gleichzeitig aber bemerken wir, wie außerordentlich klein die Wahrscheinlichkeiten, daß ich mit der Binomialformel das wahre Zahlenverhältnis dargestellt habe, überhaupt sind. Das liegt natürlich daran, daß das geprüfte Material der Zahl nach viel zu gering ist, um sichere Schlüsse zuzulassen.

b) Eine zweite Methode, die nach der Binonialformel erhaltenen Werte auf ihren Genauigkeitsgrad zu prüfen, besteht in der Anwendung des Bernoullischen Theorems.

Liegen dem Vorgang die Wahrscheinlichkeiten $w_1 = \frac{e_1}{e}$ bezw. $w_2 = \frac{e_2}{e}$ zugrunde, so verteilen sich die bei verschiedenen Versuchen zu erwartenden Häufigkeiten n_1 nach dem Gauß'schen Verteilungsgesetz

$$H = \frac{1}{\sqrt{2\pi n w_1 w_2}} \cdot \varepsilon^{\frac{-y^2}{2 n w_1 w_2}}, \tag{9}$$

wo ε die Basis des natürlichen Logarithmensystems ist und y die Abweichung vom wahrscheinlichsten Wert $w_1 \cdot n - n_1$ bedeutet.

Da der wahrscheinlichste Wert $(y = 0)$ mit der Wahrscheinlichkeit

$$H_1 = \frac{1}{\sqrt{2\pi n w_1 w_2}} \tag{10}$$

auftritt, kann

$$G_1 = \frac{H}{H_1} = \varepsilon^{\frac{-y^2}{2 n w_1 w_2}} \tag{11}$$

als Maß der Genauigkeit für das errechnete Verhältnis $e_1 : e_2$ dienen.

Für die in (11) auftretende Exponentialfunktion finden sich in allen größeren Lehrbüchern über Wahrscheinlichkeitsrechnung oder über Variationsstatistik Tabellen, mit denen in einfacher Weise die Werte von G_1 zu berechnen sind. Für die Annahme 7 : 9 ergibt sich für das beobachtete Verhältnis 95 : 128 die Wahrscheinlichkeit 0,051, während bei der Annahme 27 : 37 das beobachtete Ergebnis mit der Wahrscheinlichkeit 0,054 zu erwarten ist. Die kleine Differenz ist von großer Bedeutung, da die Wahrscheinlichkeit für die wahrscheinlichste Annahme selbst nur unwesentlich größer als 0,054 ist. Für den zweiten Fall ergibt sich also $G_1 = 1$, während im ersteren Fall $G_1 = 0{,}94$ ist.

c) Das hier aufgestellte Genauigkeitsmaß unterscheidet sich von dem von Harris vorgeschlagenen dadurch, daß durch das letztere die Wahrscheinlichkeit dafür angegeben wird, daß eine Beobachtung als ein ganz beliebiges Versuchsergebnis aus einer Reihe von Ereignissen, die einem Gesetz gehorchen, angesehen werden kann. Das Kriterium dafür ist zwar von Pearson auf verschiedene Beispiele angewandt und auf ein System von Beobachtungen ausgedehnt, aber es ist nicht, wie Harris angibt, von Pearson zuerst aufgestellt, sondern in der Wahrscheinlichkeitsrechnung bereits durch das Bernoullische Theorem gelöst worden. Letzteres läßt sich nämlich folgendermaßen formulieren:

Sind n Versuche gemacht, so ist die Wahrscheinlichkeit dafür, daß die Häufigkeit des einen Ereignisses (n_1) ebenso weit oder weiter von der wahrscheinlichsten Zahl $\frac{e_1}{e} \cdot n$ abweicht als die beobachtete Zahl n_1, gegeben durch den Ausdruck:

$$G_2 = 1 - \left(\frac{2}{\sqrt{\pi}} \int_0^{\gamma} \varepsilon^{-t^2} \, dt + \frac{\varepsilon^{-\gamma^2} \cdot e}{\sqrt{2 n e_1 e_2}} \right), \tag{12}$$

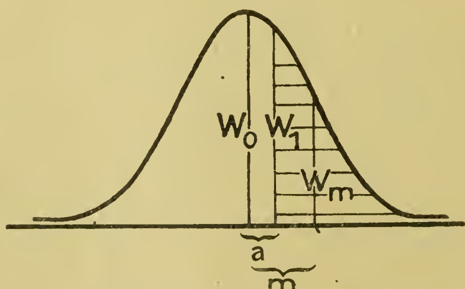
$$\text{wo } \gamma = \frac{y \cdot e}{\sqrt{2 n e_1 e_2}} \text{ ist.}$$

Sowohl für das Integral wie auch für die Exponentialfunktion sind ausführliche Tafelwerke vorhanden.

Für unser Beispiel ergibt sich bei der Annahme 7 : 9 der Wert $G_2 = 0{,}68$ und bei 27 : 37 der Wert $G_2 = 0{,}85$.

7. Kritik der Genauigkeitskriterien.

Alle Kriterien sind unter der Voraussetzung abgeleitet, daß den Ergebnissen feste Wahrscheinlichkeiten zugrunde liegen, d. h. daß die Abweichungen bei großer Versuchszahl sich nach der normalen Gauß'schen Verteilungskurve ordnen. Diese sei in der nebenstehenden Abbildung veranschaulicht.



Dann ist das im 4. Kapitel abgeleitete Kriterium durch die Abweichung a gegeben, die angibt, um wieviel die beobachtete Häufigkeit von der wahrscheinlichsten abweicht. Das im 5. Kapitel aufgestellte Genauigkeitsmaß wird dargestellt durch das Verhältnis a : m, wo m den mittleren Fehler bedeutet. Im 6. Kapitel ist durch (9) die Größe W_1 gegeben, die die Wahrscheinlichkeit, mit der das beobachtete Verhältnis aus der zu W_0 gehörigen Variationskurve sich ergibt, darstellt. Durch (11) ist das Verhältnis $W_1 : W_0$ gegeben. (12) liefert schließlich das Verhältnis der schraffierten Fläche zur halben Gesamtfläche. Aus den Lehren der Variationsstatistik folgt ohne weiteres, daß nur die im 6. Kapitel abgeleiteten Kriterien Ansprüche auf Exaktheit machen können. Von ihnen steht das Harrissche an erster Stelle. Es ist aber zu betonen, daß die Herleitung der Formel (12) nur unter zahlreichen für das Bernoullische Theorem geltenden vereinfachenden Annahmen möglich ist. Außerdem sagt sie nur aus, wie groß die Wahrscheinlichkeit ist, daß bei einem beliebigen Versuch die Abweichung ebenso groß oder größer ist als die beobachtete. Daß diese Wahrscheinlichkeit ein Maß dafür ist, daß ein beobachtetes Verhältnis aus einer bestimmten Reihe von erwarteten stammt, ist nicht ohne Weiteres erwiesen, sondern kann nur als eine Definition von Pearson-Harris bezeichnet werden. Über die Wahrscheinlichkeit mit der gerade die beobachtete Abweichung selbst zu erwarten ist, sagt sie nichts aus. Es muß daher das Kriterium (9) bezw. (11) hinzugenommen werden, das eine Vereinfachung des exakt gültigen Genauigkeitsmaßes (6) darstellt. (9) teilt mit (12) den Vorteil, daß es nicht nur auf das Verhältnis von zwei Zahlen anwendbar ist, sondern das gesamte Beobachtungsergebnis auf einmal zu prüfen gestattet.

Alle Genauigkeitsmaße setzen uns in den Stand, verschiedene Annahmen auf ihren Wahrscheinlichkeitsgrad zu prüfen, sie gestatten aber nicht zu entscheiden, ob eine Annahme die allein richtige ist. Das Versuchsergebnis, und sei es noch so groß, ist immer nur als ein Versuch anzusehen, und vom Standpunkte der Wahrscheinlichkeitstheorie aus ist es ganz unzulässig, daraus sichere Schlüsse zu ziehen. Erst zahlreichere Versuchsergebnisse von gleicher Individuenzahl könnten über die wahre Gestalt der Verteilungskurve Aufschluß geben. Ergibt sich eine schiefe Kurve — und die Zusatzhypothesen der neueren Vererbungslehre über die Koppelung der Faktoren setzen diese Annahme geradezu voraus —, so sind alle Kriterien ungültig.

Vom mathematischen Standpunkt aus ganz unzulässige Schlüsse finden sich in der Literatur in großer Zahl. Hier seien nur einige aufgeführt. So wird bei Bateson (a. a. O. S. 134) aus dem beobachteten Verhältnis 70 : 21 : 36 auf 9 : 3 : 4 geschlossen ($m = 2$), während das Verhältnis 141 : 42 : 73, welches sich für $m = 3$ ergeben kann, eine viel bessere Übereinstimmung von erwarteten und beobachteten Zahlen ergibt. An einer andern Stelle (S. 151) wird aus 627 : 27 : 17 : 214 auf 637 : 27 : 27 : 194 geschlossen. Ferner einige Beispiele bei Lang (a. a. O.):

S. 507: Aus 33 : 10 : 8 : 2 : 12 wird gefolgert 27 : 9 : 9 : 3 : 16.

S. 529: Aus 21 : 6 : 0 : 4 : 6 wird gefolgert 9 : 3 : 1 : 2 : 1.

S. 562 ergibt sich 66 : 46 statt 56 : 56.

Allgemein gibt Lang an, daß die Fehlergrenze durch den dreifachen mittleren Fehler gegeben ist. Das ist zwar für eine große Anzahl von Versuchsgruppen gültig. Wie aber die Verteilungskurve zeigt, kommen auch größere Abweichungen, wenn auch nur selten, vor. Wenn also Lang bei einem erwarteten Verhältnis 9 : 3 : 3 : 1 eine Zahl von 8 Individuen untersucht und das Ergebnis als Stütze für die aus theoretischen Erwägungen geschlossenen Erwartungszahlen benutzen will (a. a. O. S. 364), so ist das unzulässig. Aus diesem einen Versuch sind überhaupt keine Schlüsse zu ziehen. Er kann sowohl an der Grenze wie in der Mitte der Variationskurve liegen. Ist doch selbst bei der doppelten Anzahl von Individuen, d. h. bei 16, die Wahrscheinlichkeit, daß wirklich das erwartete Verhältnis 9 : 3 : 3 : 1 auftritt, nicht etwa 1, wie die landläufige Auffassung annimmt — die beispielsweise behauptet, daß unter sechs Würfen mit einem Würfel eine bestimmte Zahl einmal auftritt — sondern nur 0,2. Leider gibt auch das von Lang zur Veranschaulichung seiner Ausführungen herangezogene wahrscheinlichkeitsrechnerische Beispiel (a. a. O. S. 367) insofern ein falsches Bild, als bei der Gegenwahrscheinlichkeit für das betrachtete Ereignis nicht sämtliche übrigen Ereignisse (statt „mehr als" muß es immer heißen „mehr oder weniger als") berücksichtigt sind.

8. Schlußfolgerungen.

Im 4. Kapitel wurde gezeigt, daß für zwei ganz beliebige Zahlen n_1 und n_2 sich immer eine zugehörige Faktorenzahl m berechnen läßt, allein unter der Voraussetzung, daß der eine Phaenotypus dadurch charakterisiert ist, daß alle dominanten Faktoren mindestens in der Einzahl vertreten sind. Läßt man diese einschränkende Voraussetzung fallen, so kann man der Formel (3) zahlreiche andere an die Seite stellen. So würde beispielsweise die Formel $(3+1)^{m_1} \cdot (2+2)^{m_2}$ zahlreiche Wertepaare m_1 und m_2 ergeben. Aber auch allein mit der Formel (3) ergibt sich für jedes beliebige Verhältnis nicht nur ein Wert von m sondern mehrere, die nach den Gesetzen der Wahrscheinlichkeitslehre nicht als unmögliche Werte zu gelten haben. Es lassen sich zwar unter den verschiedenen errechneten Werten solche von größerer und geringerer Wahrscheinlichkeit unterscheiden, aber da es sich nur um Wahrscheinlichkeitswerte handelt, ist nicht zu sagen, ob nicht gerade der Wert mit der größeren Abweichung der richtigere ist. So könnte etwa der Wert mit der kleinsten mittleren Abweichung aus der Formel

$$3^x - \frac{n_1}{n} \cdot 4^x = \delta \cdot \sqrt{\frac{3^x (4^x - 3^x)}{n}} \tag{13}$$

(wo δ zwischen 0 und 1) durch graphische Darstellung ermittelt werden. Selbst eine größere Zahl von Beobachtungen würde eine sichere Entscheidung zwischen nahe aneinander liegenden Zahlenverhältnissen nicht ermöglichen.

Daß beliebige Zahlenverhältnisse auch ohne Zuhilfenahme weiterer Hypothesen durch einen oder mehrere Mendelsche Brüche dargestellt werden können, geht auch aus folgender Betrachtung hervor. Die Binomialformel, die den Mendelschen Brüchen zugrunde liegt

$$\left(\frac{1}{2} + \frac{1}{2}\right)^n$$

geht für große n über in

$$\left(\frac{1}{2} + \frac{1}{2}\right)^\infty.$$

Das ist aber nichts anderes als die Gaußsche Verteilungskurve, die ein ganz beliebiges durch Zufall erhaltenes Beobachtungsmaterial darstellt. Daß nun für den Gesamtphänotypus zahlreiche Faktoren m maßgebend sind, ist ohne Zweifel. Man könnte daraus den Schluß ziehen, daß die Verallgemeinerung der Mendelschen Regel auf die Variationskurve führt und daß damit eine neue Bestätigung für diese Regel erbracht ist. Wenn man aber bedenkt, daß die Gaußsche Kurve nur die Darstellung für eine ganz zufällige Verteilung ist und bei der Prüfung der Mendelschen Regeln nicht die sämtlichen Variationen berücksichtigt sondern gewisse äußere Merkmale herausgegriffen werden, d. h. Ordinaten der Variationskurve in beliebiger Weise addiert werden, so ist daraus nur der Schluß zu ziehen, daß mit diesen Formeln beliebige Verhältnisse dargestellt werden können. Der Übergang von der quantitativen zur qualitativen Variation und zur alternativen Vererbung erscheint so in einem ganz anderen Lichte.

Sind schließlich die Faktoren nicht unabhängig voneinander, so verliert die Binomialformel ihre Gültigkeit. Die Grundlage für sie ist der Multiplikationssatz der Wahrscheinlichkeitsrechnung. Dieser ist aber nur anwendbar, wenn die Wahrscheinlichkeiten unabhängig voneinander sind. Die zahlreichen Ergänzungshypothesen über die Koppelung der Faktoren, ihre Anziehung beziehungsweise Abstoßung, die Epi- und Hypostasie, die Annahme geringerer oder größerer Lebensfähigkeit bestimmter Kombinationen widersprechen daher geradezu den Grundlagen der Mendelchen Regeln. Erschien bereits durch die Aufstellung der An- und Abwesenheitstheorie die Annahme, daß die Faktoren immer in doppelter Zahl auftreten, einigermaßen gekünstelt, so scheint sie sich bei den neueren Hypothesen nur noch durch die Absicht, mit Gewalt das Spaltungsprinzip aufrecht zu erhalten, rechtfertigen zu lassen.

Eine Entscheidung, welche Wahrscheinlichkeiten dem Vorgang zugrunde liegen und wie sie voneinander abhängen, ist nur mit der Methode auszuführen, die Kapteyn [8]) für beliebige Variationskurven ausgearbeitet hat. Leider ist das Beobachtungsmaterial in jedem bisher beobachteten Falle viel zu klein, um diese Methode in Anwendung zu bringen.

Nun ist es natürlich häufig möglich, aus biologischen Gründen zwischen den verschiedenen mathematisch möglichen Mendelschen Brüchen eine Auswahl zu treffen. Häufig wird die F_3-Generation entscheidend sein. So zeigt die Rechnung, daß in unserm Beispiel unter den gemachten Voraussetzungen bei der Annahme 7 : 9 in F_3 das Verhältnis 156 : 100 auftreten müßte und bei der Annahme 27 : 37 würde sich 3060 : 1036 ergeben. Bei einem großen Zahlenmaterial würde sich also in diesem Falle die Unhaltbarkeit der einen Annahme durch die Untersuchung von F_3 erweisen lassen. Abgesehen davon, daß nur in den seltensten Fällen Beobachtungen über F_3 gemacht sind, lassen sich auch leicht Fälle finden, wo die Unterscheidung schwierig oder unmöglich ist, so etwa bei 9 : 3 : 4 und 8 : 4 : 4 oder 61 : 3 und 62 : 2 und vielen anderen. Jedenfalls muß verlangt werden, daß in allen Fällen die Unhaltbarkeit der benachbarten Mendelschen Brüche, die häufig größere mathematische Wahrscheinlichkeit für sich haben als die behaupteten, dargetan wird.

Das Verhalten der Landinsekten dem Wasser gegenüber.

Von J. S. Szymanski, Wien.

(Mit 1 Textfigur.)

Das Wohngebiet vieler Landinsekten ist häufig Überschwemmungen ausgesetzt. Denn nach jedem Gußregen bildet sich auf den Wiesen und Äckern eine Unzahl von Lachen, die kleine Inseln umschließen.

Es war nun von Interesse zu untersuchen, wie sich die auf solchen Inseln befindliche Insekten ans „Land“ herüberretten können.

Fast fünfzig Insektenarten wurden mit Rücksicht auf diese Frage untersucht.

---

8) Vgl. J. C. Kapteyn, Skew frequency curves in biology and statistics. Teil I, Groningen 1904, Teil II, Groningen 1916.

Zeitschrift/Journal: Biologisches Zentralblatt

Jahr/Year: 1918

Band/Volume: 38

Autor(en)/Author(s): Riebesell P.

Artikel/Article: Einige zahlenkritische Bemerkungen zu den Mendelschen Regeln. 329-340