

Biometrika Trust

The Law of Error by F. Y. Edgeworth; The Generalised Law of Error, or Law of Great Numbers; Ueber das Fehlergesetz by C. V. L. Charlier; Die Zweite Form des Fehlergesetzes; Ueber die Darstellung Willkürlicher Functionen; Researches into the Theory of Probability
Review by: W. P. E.

Biometrika, Vol. 5, No. 1/2 (Oct., 1906), pp. 206-210

Published by: [Biometrika Trust](#)

Stable URL: <http://www.jstor.org/stable/2331663>

Accessed: 18/06/2014 23:35

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
<http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Biometrika Trust is collaborating with JSTOR to digitize, preserve and extend access to *Biometrika*.

<http://www.jstor.org>

NOTICES AND BIBLIOGRAPHY.

NOTICES.

1. EDGEWORTH, F. Y. *The Law of Error*. Transactions of the Cambridge Philosophical Society, 1905, Vol. xx. pp. 36—65 and 113—141.
2. ——. *The Generalised Law of Error, or Law of Great Numbers*. Journal of the Royal Statistical Society, 1906, Vol. LXIX. pp. 497—539.
3. CHARLIER, C. V. L. *Ueber das Fehlergesetz*. Arkiv för Matematik, Astronomi och Fysik, Vol. II. Stockholm, 1905.
4. ——. *Die zweite Form des Fehlergesetzes*. Ibid.
5. ——. *Ueber die Darstellung Willkürlicher Functionen*. Ibid.
6. ——. *Researches into the Theory of Probability*. Meddelanden från Lunds Astronomiska Observatorium, Serie II. No. 4. Lund, 1906.

In (1) Professor Edgeworth, starting from various conditions, some of which he afterwards shows can be relaxed, gives four methods by which one can reach an “approximate expression of the frequency with which in the long run different values are assumed by a quantity which is dependent on a number of variable items or elements.” These conditions are that the elements assume different values in random fashion and in the long run recur with a proportionate frequency capable of being represented by a single definite frequency curve; that the variations are independent of each other*; that the method of aggregation by which the elements are compounded is summation, etc. etc.

Professor Edgeworth first gives a method which consists of equating the i^{th} moment of the frequency with the same moment of the given locus. He then shows that the same curve can be reached by working on the lines followed by Professor Morgan Crofton and by the method originated by Laplace and developed by Poisson. He then gives confirmatory evidence by using Laplace’s analysis with some of the conditions used by Crofton and inserts the fresh condition that if there be two or more magnitudes each fluctuating according to the law of error, then the sum of each must also fluctuate according to that law.

* [The assumptions that the elementary cause-groups are independent and that the aggregate is obtained by summation have yet to be justified. In particular the first assumption is opposed to the basis of every determinantal theory of heredity, and accordingly the frequency distributions of characters, which result from the fusion and throwing out of chromasomes, i.e. characters in living organisms, are extremely unlikely to comply closely with Professor Edgeworth’s form of frequency. I have repeatedly urged the necessity for considering contributions to the aggregate as *correlated*, i.e. the hypergeometrical as distinguished from the binomial form of series, as the basis of frequency distributions. The skew curves I have introduced proceed from the basis that the “contributory cause-groups” give contributions to the aggregate which are *correlated*. See *Biometrika*, Vol. iv. pp. 196, 203 *et seq.* K. P.]

The general form reached is written

$$e^{-\kappa_1 \frac{1}{3!} \left(\frac{d}{dx}\right)^3 + \kappa_2 \frac{1}{4!} \left(\frac{d}{dx}\right)^4 - \dots + (-1)^t \frac{1}{(t+2)!} \left(\frac{d}{dx}\right)^{t+2}} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2\sigma}}$$

where

$$\kappa_1^2 = \frac{\mu_3^2}{8\mu_2^3}; \quad \kappa_2 = \frac{\mu_4 - 3\mu_2^2}{4\mu_2^2}; \text{ etc.}$$

If this form be rewritten as

$$F(x) = A_0\phi(x) + A_3\phi^{\text{iii}}(x) + A_4\phi^{\text{iv}}(x) + \dots$$

where

$$\phi(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-(x-b)^2/2\sigma}$$

it becomes the same as that called Type A by Dr Charlier in (3), (5) and (6) and it is also the same as that given by Dr Thiele in "Theory of Observations" (London, C. and E. Layton, 1903) p. 35. Charlier's method of reaching his form is by following Hagen's development of Laplace. The same writer also gives in (4), and considers more minutely in (5) and (6), the form (Type B) which he writes

$$F(x) = B_0\psi(x) + B_1\Delta\psi(x) + B_2\Delta^2\psi(x) + \dots$$

where

$$\psi(x) = \frac{e^{-\lambda} \sin \pi x}{\pi} \left[\frac{1}{x} - \frac{\lambda}{1!(x-1)} + \frac{\lambda^2}{2!(x-2)} \dots \right].$$

This curve with a range limited in one direction is, we believe, new though Thiele has given a form very closely allied to it (*loc. cit.* p. 21).

Charlier uses the method of moments for fitting his curves, but though both Edgeworth and he do this, and their series finally take the same form, different graduation results will be reached owing to the index form being used in the one case and not in the other; the difference may, in some cases, be negligible but in others it becomes of more importance and we shall therefore refer to it later.

It will be noticed that in all cases it is proposed to use a series to describe the frequency distribution and there seem to us so many objections to this course in practice that it is well to take this opportunity of examining it. The objections to it are as follows :

(i) If one of the later coefficients has a large value the neglect of later terms of the series may involve a considerable error, while their inclusion demands the use of the higher moments which are untrustworthy owing to their large probable errors.

(ii) In some cases the series lead to negative frequencies, which is objectionable. This can often occur with Type A and is noticeable with Thiele's example (*loc. cit.* p. 50).

(iii) It is necessary to make successive graduations using an increasing number of terms in order to find how many terms of the series are required to give a satisfactory graduation.

(iv) As we cannot tell at the first how many terms to use, it is necessary to base the solution of the equations for finding the constants on integrations over the whole series from $-\infty$ to $+\infty$ and then neglect terms which may or may not be significant, or else to make successive trials with an increasing number of terms from equations formed from the actual number of terms used. The latter method would be better if the position of negative terms could be decided at the outset and if integration could be effected between any limits that might be indicated. This would however seem to be impossible and Charlier uses the former method; the objection does not apply to Edgeworth's series.

The effect of these objections in the case of Charlier's work is interesting as it is quite impossible to reproduce one of his frequency curves (the bi-modal curve, fig. 5 of (6)) statistically because the negative frequencies play so important a part in the series that if positive frequency only be taken (which is what would happen in practice) an entirely different curve is obtained. We are by no means satisfied that in such cases the integration for moments from $-\infty$ to $+\infty$ is

an unsatisfactory result. In all the graduations we could doubtless improve the agreement by using a greater number of terms in the series, but we think a considerable increase in the number would be required to give what we should consider a satisfactory graduation.

Size of family	Observations	Edgeworth's third Approximation	Charlier Type A† $\sigma = 2.928$; $\beta_3 = -.1214$ $\beta_4 = .0104$	Charlier Type B	Pearson Type I‡
-3	—	1*	-2	—	—
-2	—	9*	4	—	—
-1	—	30*	15	12	2
—	64	64	38	64	67
1	116	102	71	104	116
2	140	130	108	129	138
3	145	135	137	134	139
4	134	130	148	128	128
5	106	111	135	116	110
6	82	92	108	93	89
7	72	73	78	73	69
8	49	53	54	53	51
9	37	36	37	36	35
10	25	20	27	25	24
11	13	10	18	14	15
12	10	4	12	10	9
13	5	—	7	5	5
14	2	—	4	2	2
15	.4	—	2	1	1
Totals	1000	1000	1001	1000	1000

* Approximation by help of diagram in Edgeworth (2).

† Notation of Charlier (6), mid-ordinates, found by Charlier's tables, being used.

‡ "Chances of Death," Vol. I. p. 74.

To the actuary, influenced perhaps by professional bias, the justification of a formula for graduating frequency distributions is its width of application; to some extent we feel that such is also the justification of any theoretical conditions from which a curve is evolved. Edgeworth's series and Charlier's Type A will be found to give good graduations provided the distributions are not markedly skew but they become less satisfactory as the range of the observations takes a definite limit. Charlier's Type B on the other hand is certainly capable of graduating some distributions having a range limited in one direction but, though it can hardly be criticised fully at present, as the author states in (6) that his work on it is not yet complete, it may be well to point out that the solutions he gives are approximate and the choice of solution in any particular case seems somewhat arbitrary. The comparatively poor agreement reached above may be due to this approximate fitting and not to the failure of the curve itself. A statistical criterion to show whether Type A or Type B should be used in any particular case is certainly needed before these types can be used extensively in practice, but even then it would seem impossible to graduate the U-shaped distributions or those that rise abruptly from the axis of x at both ends.

One or two examples, besides that already mentioned, are given in (2), while there is a plentiful supply of statistical examples in (6) and most of them show a close agreement between the theoretical and actual frequencies; some are less satisfactory and fig. 9 of (6) gives so poor a fit that the odds against the graduation are more than 50 to one. There are many other points of interest in (6) beside the main subject, such as a proof, on the basis of Type A, of the relative positions of the mode, mean and median, a method of checking the numerical calculation of

moments, tables of the areas, ordinates and third and fourth differential coefficients of the normal curve, a table of $\psi(x)$ for Type B and a discussion of the dissection of a frequency distribution into components in which some approximate results are given and the suggestion of shortening the solution of the fundamental nomic by means of graphical work is made.

We have put forward the above criticisms to show the practical difficulties we have met in using the suggested methods; though these difficulties seem very important to us they do not blind us to the energy and ingenuity expended on the papers.

WILHELM FLIESS. *Der Ablauf des Lebens. Grundlegung zur exakten Biologie.* Leipzig, 1906, pp. 584 + viii.

As this is hardly the type of statistical work that will appeal to our readers it is unnecessary to criticise it.

W. P. E.

BIBLIOGRAPHY.

ANDERS, J. M. & MORGAN, A. C. Tetanus. A Preliminary Report of a Statistical Study. Jour. Amer. Med. Assoc. Vol. XLV, pp. 314—322. 1905.

Statistical data on 1201 cases of tetanus.

BATESON, WILLIAM. Presidential Address. Rep. 74th Meet. Brit. Ass. Adv. Sc. pp. 574—589. 1905.

BATESON, W. & GREGORY, R. P. On the Inheritance of Heterostylism in *Primula*. Proc. R. Soc. London, Vol. 76, B, pp. 581—588. 1905.

Mendelian.

BATESON, W., SAUNDERS, E. R., PUNNETT, R. C. & HURST, C. C. Experimental Studies in the Physiology of Heredity. 2nd Rep. Evol. Comm. R. Soc. 154 pp. 1905.

Mendelian results on plants and poultry.

CASTLE, W. E. Heredity of Coat Characters in Guinea-Pigs and Rabbits. Washington. Published by the Carnegie Institution, 8°, 78 pp., 6 pls. 1905.

Mendelian.

——. Inbreeding, Cross-Breeding and Sterility in *Drosophila*. Science, N. S., Vol. XXIII, p. 153. 1906.

Preliminary report of a study dealing mainly with the variation and inheritance of fertility in *Drosophila*.

CORRENS, C. Gregor Mendels Briefe an Carl Nägeli 1866–73. Ein Nachtrag zu den veröffentlichten Bastardierungsversuchen Mendels. Abh. math. phys. Kl. sächs. Ges. Wiss. Bd. XXIX, No. 3, pp. 187—265, 1 Facsimile. 1905.

CRAMPTON, HENRY EDWARD. On a General Theory of Adaptation and Selection. Journ. exper. Zool. Vol. II, pp. 425—430. 1905.

Theoretical.

DALY, R. A. Machine-made Line Drawings for the Illustration of Scientific Papers. Science, N. S., Vol. XXII, pp. 91—93. 1905.

Use of Hammond typewriter in making and lettering line diagrams in statistical and other work.

DANDENO, J. B. The Parachute Effect of Thistle-Down. Science, N. S., Vol. XXII, pp. 568—572. 1905.

An attempt to determine quantitatively the weight and surface area of the different parts of the down of the Canada thistle (*Carduus arvensis*), with special reference to the mechanics of seed dispersal.

DAVENPORT, C. B. Evolution without Mutation. Journ. exper. Zool. Vol. II, pp. 137—143. 1905.

——. The Origin of Black Sheep in the Flock. Science, N. S., Vol. XXII, pp. 674 & 675. 1905.

Using the data provided in Dr Alex. Graham Bell's "Sheep Catalogue" the author comes to the conclusion that "black wool colour in sheep behaves like a Mendelian recessive characteristic."