McGinnis has been for nine years professor of physics and electrical engineering in the University of New Brunswick, Fredericton, N. B.

DR. HARRY B. YOCOM, of the department of biology of the College of the City of New York, has been appointed assistant professor of zoology in the University of Oregon.

DR. F. FRANCIS, professor of chemistry, has been appointed pro-vice-chancellor of Bristol University, in succession to Professor C. Lloyd Morgan, who is about to resign office. Dr. Lloyd Morgan has been appointed emeritus professor of psychology and ethics.

DISCUSSION AND CORRESPONDENCE

A PRIORI USE OF THE GAUSSIAN LAW

TO THE EDITOR OF SCIENCE: Mr. Michael[1] in interpreting Dr. Johnstone's results[2] for twenty counts of bacteria in polluted shellfish deplores certain naive errors to which the lay statistician is prone, but is not, so it seems to me, free from statistical illusion himself. I had hoped, at least, that the identification of the Gaussian law with the ideal "chance" distribution was a custom of the past, and that the prevalence of this practise in the literature was simply due to the inertia of thinking. May I submit the following relevant observations?

1. The sole condition of "change" is ignorance.[3] In science the thing to do with ignorance is to admit it, not to posit the form of distribution that a variable assumes under it.

2. Biological and mental phenomena, of whose conditions of variability we are thus ignorant, do not necessarily give symmetrical distributions when observed. Pearl showed that the amount and direction of skewness and the dependence of skewness on known conditions might be the significant biological fact.[4] The Gaussian law does hold for coin-tossing, but the relationship has been scientifically observed,[5] not posited a priori.

3. Moreover, there can be no reason to expect a Gaussian distribution a priori when we are ignorant. A form of distribution is always function of the unit of measurement; and, since the choice of a biological unit is ordinarily arbitrary, the chances of getting the normal distribution are very small.[6] Galton pointed out, furthermore, that chance distributions of two related variables, when the relationship is not linear, can not both be Gaussian.[7]

4. When we observe a skew distribution and are in ignorance of the conditions that cause the variation, it is useless labor to factor the skew distribution into a Gaussian "chance" distribution and a skewing factor, as Mr. Michael does. The two factors that we so obtain are meaningless. The Gaussian function is biologically meaningless because there is neither a priori nor observational ground for taking it as the curve of chance (ignorance). Mr. Michael's logarithmic function is biologically meaningless because it is merely a measure of the manner in which the observed data depart from the meaningless Gaussian law. Pearson saw this point in 1900 and noted the fallacy.[8] He also made fun of the Gaussian "fetish," although the position of the Biometric School has since become less definite.

5. Probability in science means frequency and nothing more. Fundamentally in science

1 E. L. Michael, "Concerning Application of the Probable Error in Cases of Extremely Asymmetrical Frequency Curves," SCIENCE, N. S., 51, 89–91.

2 J. Johnstone, "The Probable Error of a Bacteriological Analysis," cited as Rept. Lanc. Sea-Fish. Lab., No. 27, 1919, 64–85.

3 *Cf.* J. Venn, "Logic of Chance," 1888, espec. 119 ff.; B. Bosanquet, "Logic," 1911, I., 322 ff. If the scientist prefers not to go to the logician, let him see if he can formulate for himself, with scientific rigor, the conditions of "chance."

4 R. Pearl, "Variation and Differentiation in *Ceratophyllum*," 1907, espec. 90 f.

5 *E. g.*, see H. Westergaard, "Grundzüge der Theorie der Statistik," 1890, 21–38.

6 J. Bertrand, "Calcul des probabilités," 1889, 180 f.

7 F. Galton, *Proc. Roy. Soc.*, 29, 1879, 365–367.

8 K. Pearson, *Philos. Mag.*, 5th ser., 50, 1900, 173.

it means observed frequency. The value of the statistical constants is simply that they provide a conventional method of summarizing frequencies of observed data. To shift the meaning of probability from observed frequency to predicted frequency is precarious, although we are always attempting it in scientific generalization. However, it takes more than a process of division by the square root of the number of cases—the obtaining of the probable error of the mean—to bridge the gulf between observation and prediction. The lay conviction that the probable error of the mean is actually a prophecy is hard to overcome. That it is not prophetic will become clear to any one who will take the trouble to fractionate a large body of data, compute the probable errors of the means of each fraction and note how they vary, and then compare all these discordant predictions with the actual probable error of the means computed from the array of means. The probable error of the mean is a useful constant since it summarizes the variability of data in relation to their amount; but it is not a key to the future.

All this is negative. Actually what was Dr. Johnstone to do? First, observe and report, I should say; and let him predict who will. Certainly there is no need for much statistics to summarize his twenty cases. He wishes to know the most probable number of bacteria per cc. in this emulsion. Scientifically by the most probable number is meant the most frequent number; and his data show that 6–10 counts were more frequent than any other. Why obscure the simple fact by a statistical superstructure? If now he wishes to risk prediction on the basis of 20 cases, he may say that 6–10 counts will occur more often in his 250 c.c. than any other group, 16–20 counts next most often, 11–15 and 21–25 counts less often, and so on. This course has the simple merit of telling the observed truth and doing very little more.

In predicting the total number of bacteria within the 250 c.c. one must multiply the arithmetic mean of the counts by 250. We have given the distribution of 20 counts and we have no alternative to assuming that it is the most probable distribution of 20 counts. Hence we must take the observed distribution as many times over ($12\frac{1}{2}$ times) as 20 will go into 250 and sum all the frequencies. Dr. Johnstone found 366 bacteria in 20 c.c. The most probable number in 250 c.c. must be $250/20 \times 366 = 4{,}575$. Mr. Michael gets 4,005 by the erroneous assumption that the most probable (most frequent) logarithm is the logarithm of the most probable (most frequent) count, which is plainly impossible since the logarithmic relation is not linear. The illusion arises because we take it for granted that any most probable natural number must be inseparably connected with the most probable logarithm. When we substitute the word "frequent" for "probable" we may see our mistake, for the logarithms of the small numbers are more frequent than the logarithms of the large numbers.[9]

Concerning the general problem of obtaining "the probable error of extremely asymmetrical frequency curves," I would urge that in simple cases it is unnecessary to depart far from the observed facts. Usually one is most interested in the value of the most frequent (most probable) case and in the amount of deviation on either side. The values of the mode and of the upper and lower quartiles give this information, as well as the range within which half the cases have fallen and an indication of the skewness. Except the gift of prophecy, what more could one want?[10]

Edwin G. Boring

Clark University,
Worcester, Mass.

ALBINO VERTEBRATES

In July 1919, on the Beaver River near the mouth of the Dore River in Saskatchewan, I shot a pure albino grackle (*Quiscalus quiscula æneus*). It was a young male, 10.5 inches long, and was associated with a flock of grackles. It seemed much less shy than the

[9] S. Newcomb, *Amer. J. Math.*, 4, 1881, 39 f.

[10] See in general, "The Logic of the Normal Law of Error in Mental Measurement," *Amer. J. Psychol.*, 31, 1920, 1.