NOTE.

THE LAWS OF PROBABILITY AND THEIR MEANING.—The point has more than once been raised, in conversation and in correspondence, that the constancy of form shown by the frequency distributions of sizes of genera—the charts showing numbers of genera of 1, 2, 3, . . . species—has no bearing on vital phenomena, but is in some way or other a necessary outcome of 'the laws of probability'.

But the laws or rules of probability are only rules for the combination of chances and can give us nothing by themselves, any more than the rules of multiplication and division. If the chance of a given event occurring is p_1 and the chance of an independent event occurring is p_2, the chance of both events occurring is $p_1 p_2$: if the two events are not independent but alternative, the chance of *either* the first *or* the second occurring is $p_1 + p_2$—such rules alone can only lead us to a form of frequency distribution by applying them to given assumptions concerning the facts. A few illustrations may make the matter clearer.

If we have n dice, and we assume (a) that the chance of throwing sixes is the same for every die, (b) that the same dice are used throughout the experiment, so that the chance of throwing sixes is the same at every throw, (c) that the result of the throw of every die is independent of that of every other, then the laws or rules of probability show that the frequencies of 0, 1, 2, 3, . . . sixes in N throws of the n dice are given by the successive terms of $N(q+p)^n$. If we find the results of a long series of throws differ significantly from this theoretical expression, we may conclude that our assumptions are in error: the dice are not all the same, or two different sets of dice with different values of p have been used, or it may be that we have merely taken an incorrect value of p.

Again, if we assume (a) that an error of observation is compounded of a number of elementary errors, (b) that the number of these elementary errors is large, (c) that positive and negative elementary errors are equally frequent, (d) that elementary errors are independent, the form of frequency distribution deduced is the normal curve of errors. If we make different assumptions as to the way in which an error of observation is built up, we will—or may—arrive at a different form of distribution, and conversely, if the curve observed is not the normal curve of errors, we are justified in concluding that our assumptions as to the genesis of an error of observation are wrong.

The following case, taken from some recent work,[1] is particularly illuminating as regards the light which the form of the frequency distribution may throw on the facts. Data are given as to the number of girls in a munition works who have met with 0, 1, 2, 3, . . . accidents (trivial accidents, sufficient to send the girl to the Welfare Room) during a certain period. The actual form of frequency distribution is not unlike that

[1] Greenwood and Yule, Journn. Roy. Stat. Soc., March, 1920.

observed for sizes of genera, girls with no accidents being most frequent, but there is not so long a 'tail' towards high numbers of accidents.

A. It is assumed that every girl is equally likely to meet with an accident during any short interval of time. The form of frequency distribution deduced does not fit the facts.

B. It is assumed that the girls may be regarded as forming two groups, careful and careless, a girl of the second class having a higher chance of meeting with an accident during any short interval of time than a girl of the first class. The form of frequency distribution deduced still cannot be made to give a good fit to the facts.

C. It is assumed that the girls are of all degrees of 'carelessness', i. e. of chance of accident during a short interval of time, and that the form of frequency distribution of carelessness is given by one of Pearson's curves with origin at zero, so that there are or may be girls of zero carelessness. The form of frequency distribution now deduced for numbers of girls with 0, 1, 2, 3, . . . accidents is found to give an excellent fit to the facts, and from the constants of the fitted distribution we can deduce the distribution of carelessness amongst the girls. The interpretation thus arrived at is confirmed by finding that the number of accidents met with by a girl during one interval is correlated with the number she shows during a subsequent interval.

We may conclude that assumptions A and B are in error, whereas C *may* be right, and since it is reasonable and probable and otherwise confirmed it probably *is* the right interpretation.

Similarly, from the view of evolution now discussed, the form of the frequency distribution for sizes of genera can be deduced, and it appears to give a very good fit to the facts. This is, in so far, confirmatory evidence of the truth of the view.

G. UDNY YULE.