https://doi.org/10.17895/ices.pub.9465

2

NOTES ON STATISTICAL METHODS OF GROWTH STUDIES.

BY

PER OTTESTAD.

A. The problem of the growth of populations was most generally considered and discussed by Malthus and it was he who devised the growth function

$$y = m \cdot a^t \tag{1}$$

in which y represents the size of the population at the time denoted by t. This formula can be accepted as a logically correct expression of the growth of a population under free conditions, that is to say, conditions in which there are no restrictive factors influencing the course of the growth.

B. The formula, however, does not correctly express the growth which we observe in nature or as a result of experiments. The rule in such cases is that the growth follows a sigmoid course. The growth increments are small at first but rise to a maximum, and then fall to zero. A number of attempts have been made to find a mathematical function capable of expressing this sigmoid growth in a concise manner. The most interesting formula which has been evolved is

$$y = \frac{b \cdot e^{at}}{1 + c \cdot e^{at}} \tag{2}$$

which is based on the analogy of chemical processes, autocatalysis being regarded as a specific example of the mass law. In this connexion, however, it has been argued that this function is an expression of the results of the underlying biological processes. To this it may be said that the number of factors which determine the course of growth is so great and their mutual correlation and influence so complicated that, until more is known about these underlying biological processes, it is not practicable to seek for logical sigmoid growth functions. The immediate problem is that of finding a function which can serve to define the sigmoid growth with sufficient accuracy. If formula (2) is regarded from this angle the only objection which can be raised to it is that it is not a suitable instrument.

In a previous paper[1]) I have suggested that the construction of a theoretical curve should be founded on the growth increments rather than on the total growth. I demonstrated that in cases of symmetrical growth the increments can be described by the Gaussian frequency function

$$f(t) = \int_{t-\Delta t}^{t+\Delta t} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(t-m)^2}{2\sigma^2}} \tag{3}$$

It will be seen that there are only two constants (m and σ) to be calculated in the actual case, and these constants measure the two most interesting parameters of the growth. m is a parameter which indicates the point of time at which the growth-rate attains the optimum value and 6σ is a measure of the span of time from the inception of growth to its fulfilment.

When the necessary observations of a growth line have been made and it is desired to describe this growth by the Gaussian function (3) the method of procedure is as follows:—

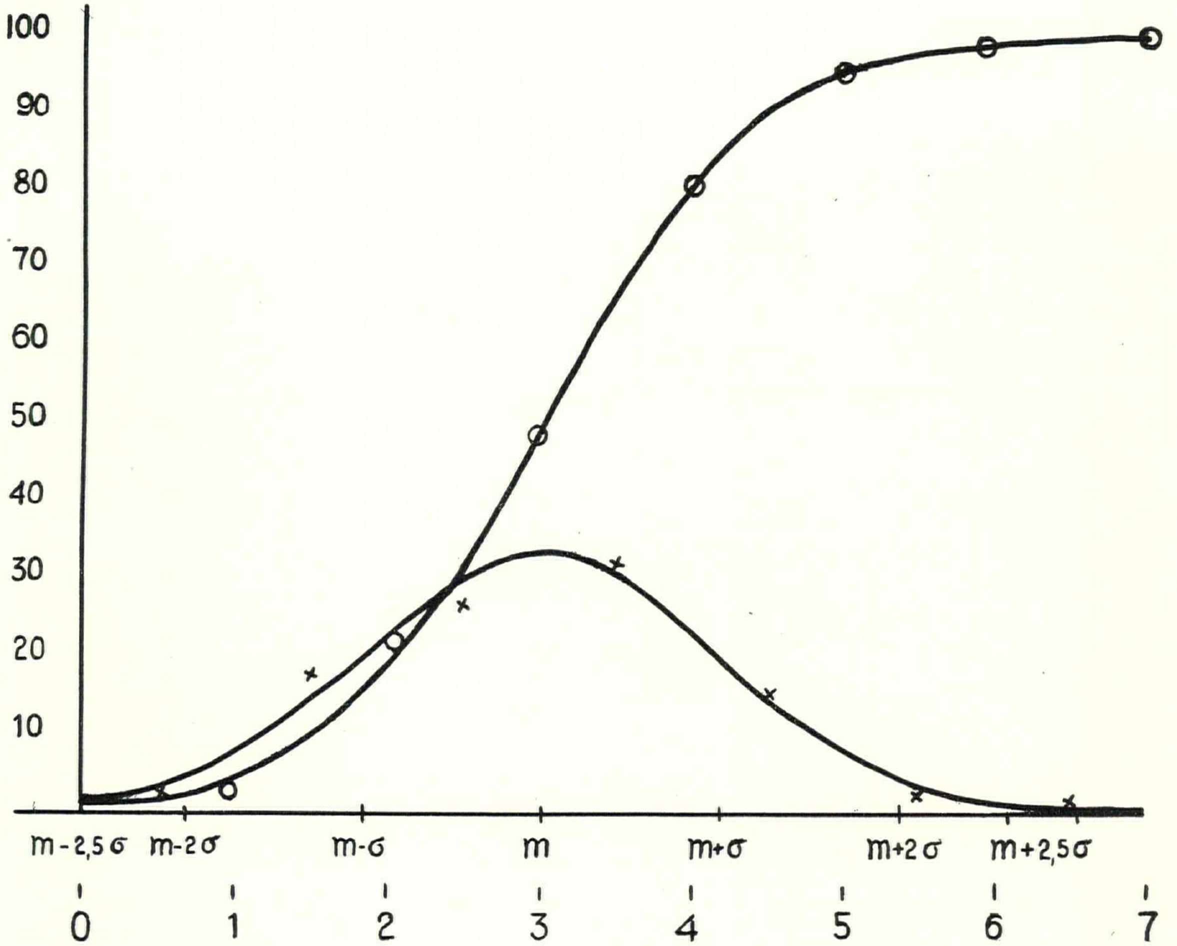
1. From these growth observations, which can be illustrated by an empiric sigmoid curve, the growth increments in the given time intervals are calculated. The growth increments are now regarded as the frequencies in a distribution series in which the time is the variable quantity. In Table 1[2]) particulars of the observations of the growth of a yeast cell population are given, and in the same table the growth increments are indicated.

Table 1.
The number of yeast cells in 1/250 mm³., reduced to an assymptote of 100.

t in 10 hr.	Number of cells	Time-interval	Mid-point in interval	Increment $h(t)$
0	1·1	0—1	0·5	2·9
1	4	1—2	1·5	18
2	22	2—3	2·5	27
3	49	3—4	3·5	32
4	81	4—5	4·5	15
5	96	5—6	5·5	3
6	99	6—7	6·5	1
7	100			

[1]) Per Ottestad: A Mathematical Method for the Study of Growth. Hvalrådets Skr., No. 7. Oslo, 1933.

[2]) Oscar W. Richards: The Growth of the Yeast Saccaromycus cerevisiae. Annals of Botany, Vol. XLII, No. CLXV, 1928.



2. From such a frequency distribution the arithmetic mean (m) and the standard deviation (σ) are now calculated in the usual manner. As regards the actual method of calculating these two parameters in practice almost any text-book of statistics[3]) may be consulted. The calculation of m and σ for the distribution series in our example is shown in Table 2:—

Table 2.
Calculation of m and σ.

t	$h(t)$	$t' = t - 2{\cdot}5$	$h(t) \cdot t'$	$h(t) \cdot t'^2$
0·5	2·9	—2	— 5·8	11·6
1·5	18	—1	—18	18
2·5	27	0	0	0
3·5	32	+1	+32	32
4·5	15	+2	+30	60
5·5	3	+3	+ 9	27
6·5	1	+4	+ 4	16
	98·9		+51·2	164·6

[3]) See for example: G. Udny Yule: An Introduction to the Theory of Statistics. p. 133, London, 1929.

$$m = 2{\cdot}5 + \frac{51{\cdot}2}{98{\cdot}9} = 3{\cdot}0177$$

$$\sigma^2 = \frac{164{\cdot}6}{98{\cdot}9} - \left(\frac{51{\cdot}2}{98{\cdot}9}\right)^2 = 1{\cdot}3963$$

$$\sigma = 1{\cdot}18$$

3. m and σ now being known the values of the Gaussian function (3) corresponding to $h(t)$ can be calculated. I do not propose to deal with the different methods by which these calculations can be made. The method employed in the present case has been described in a previous paper[4]).

In the accompanying figure the values of the Gaussian function are compared graphically with the empiric growth increments. By simple summation the total growth can now be calculated from these theoretical growth increments. The theoretical growth-curve is entered in the same figure in which

[4]) Per Ottestad: The Exponential Frequency Function and Frequency Distribution. Metron, Vol. XIII, No. 1, Roma, 1937.

the original empiric growth-data are also included for the purpose of comparison.

The difference between formula (2) and formula (3) is not that one of them is more logical than the other, but that formula (3) is undoubtedly a more suitable technical instrument than formula (2). The technical advantage of using formula (3) becomes apparent when the problem is that of studying the influence which a certain factor exerts upon the growth, for instance, when two or more growth-curves are to be compared.

Both formula (2) and formula (3) can only be used in cases of symmetrical growth. If the growth is not symmetrical, however, we can use a skew frequency function as a description of the growth increments.

C. In a great number of cases the question is one of comparing the growth-rate in various parts of the same population growth-curve. If observations of the growth increments can be produced for every short span of time over the whole range of growth, it may be practicable to evolve a mathematical function which can serve as a description of the growth. In many cases, however, it is difficult to provide such observations, and where this is so mathematical functions may lead us astray. It is, however, more useful, and also more simple, to treat the observations statistically and the task thus becomes one of finding statistics by which the growth can best be described. These statistics are always conventional but among those which are applicable there is one method which, in the present case, is preferable to the others. It is necessary, however, to distinguish between

1. absolute growth increment, and
2. relative growth increment.

Both measures must be referred to a constant interval of time (Δt). The absolute growth increment is merely the difference between the size of the population (in weight or number) at the time $(t+\Delta t)$ and the size at the time t. If the size at the time $(t+\Delta t)$ is $y\ (t+\Delta t)$ and the size at the time t is $y(t)$, the absolute growth increment is

$$z = y\,(t+\Delta t) - y\,(t)$$

The relative growth increment may be measured in various ways, for example,

$$g = \frac{y\,(t+\Delta t) - y\,(t)}{y\,(t)} \qquad \text{(a)}$$

$$g = \frac{y\,(t+\Delta t) - y\,(t)}{\frac{1}{2}\,[y\,(t+\Delta t) + y\,(t)]} \qquad \text{(b)}$$

c) Another measure which has frequently been used is

$$g = ln\,y\,(t+\Delta t) - ln\,y\,(t) \qquad \text{(c)}$$

This formula has been derived from the formula

$$C_2 = C_1 \cdot e^g$$

where C_1 represents the capital invested at the beginning of a period and g the rate of interest. The formula is based on the assumption that g is a constant during the period of investment. The capital considered as a function of time is an exponential function. In the majority of biological cases, however, the "capital" function of time follows a more or less regular sigmoid course. My opinion, therefore, is that in most cases formulae (a) or (b) are more suitable than formula (c).

There is no doubt that formula (c) gives excellent and useful expression to the relative growth-rate in those cases in which the course of growth conforms to a certain type. Formulae (a) and (b) are nothing more than statistical recording instruments and they contain no assumptions in regard to the type of growth. It is quite immaterial which of them is used, the choice being a matter of personal predilection. General agreement, however, as to the formula to be used for describing kindred types of growth would be convenient and obviate misunderstanding of these technical terms or conceptions. But the choice of a formula must, naturally, also be determined for the particular problem in each case and the material at disposal. It is, however, most expedient that, in each case, the formula used should be definitely stated.

In spite of the fact that formula (c) is restricted by certain assumptions as to growth it will nevertheless prove, in many cases, to be a useful one. It must be borne in mind, however, that a definite relation can be set up between formula (a) and formula (c). If the relative growth-rate measured by formula (a) is represented by g_a and measured by formula (c) it is represented by g_c. The following equation can be stated:—

$$g_c = g_a - \frac{g_a^{\,2}}{2} + \frac{g_a^{\,3}}{3} - \frac{g_a^{\,4}}{4} + \cdots$$

It will be seen from this that when g_a is small, that is to say, the growth is small, g_c and g_a will be very nearly equally large, and will express one and the same thing. In such cases it is perfectly justifiable to employ formula (c) as a measure for the relative growth-rate.