ON THE INTERPRETATION OF FREQUENCY CURVES IN BIOLOGY

by

E. F. DRION.

(from the Botanical Institute Government University Leyden)

TABLE OF CONTENTS.

I. THEORETICAL INTRODUCTION.

§ 1. If one measures the same characteristic of a great number of individuals of one species (or of a great number of equal organs of one individual), one generally will find different values. If all values occurred equally often life would be impossible: rats as big as elephants and men as small as fleas would exist. And if the size of all the characteristics was distributed at random all individuals would have a different form. Happily there seems always to exist a certain system in the results of our measurements. Individuals showing a certain average of the measured characteristic are most frequent, while individuals with a deviating value occur the rarer the greater the deviation from this average is. In order to compare two different characteristics of one species, or the same characteristic of two different species, it is necessary to measure a great number of individuals, for if one has measured only one individual for each characteristic (or for each species) it would be impossible to deduce a general law from the results. Perhaps one would find just the reverse in repeating the experiment. To get a good survey of the results it is desirable to plot them in a frequency curve, this means that one plots the results of the measurements against the number of individuals in which that result is found.

To make the argumentation less abstract we will limit ourselves to the discussion of the frequency curves of the length of plants, and we will suppose that this length is only the result of a growth-process. This means that we suppose that the differences in the initial-length are small compared with the differences in the final length. In macroscopic methods of measuring this will generally hold, in microscopic often not. The length of a cell for instance will not only depend on its growth but also on its length directly after the division of its mother cell. Almost all frequency curves of gradually changing objects, however, originate in a manner comparable to a growth-process.

As we have said already, a measurement performed on a single individual has but little value. When repeating the measurement on another individual, or even if possible on the same individual, probably another result would have been found, on account of the variability of biological material. Therefore one tries to eliminate this variability by measuring a great number of individuals. But the result of the measurements of a great number of individuals may always be plotted as a frequency curve. Therefore the measuring of a great number of individuals means that one supposes that, under certain circumstances, a frequency curve is less variable than a single individual, and that one expects to find a nearly identical frequency curve when repeating the experiment. For only in that case a frequency curve will have any value. Therefore the questions that should be answered before any quantitative biological work is possible are:

1. Which circumstances will enable us to draw any conclusions from frequency curves?
2. What may be concluded from given frequency curves?

To answer these questions we will follow this way: As it is supposed that the length of a plant is only the result of a growth-process, the factors which determine the speed of growth at a certain moment must be determined. Then it will be possible to answer the question how it is that, when measuring a great number of individuals, one finds on the one hand many different lengths, on the other hand many individuals of the same length, in other words how a frequency curve is *generated*, and also which causes determine its form. It will turn out that under certain conditions the form of the frequency curve does not depend upon the exact magnitude of the factors which determine the growth velocity.

An accurate study of these conditions will give the answer to the first question and even a qualitative answer to the second question, since it will turn out that under these conditions a repetition of the experiment shall give a nearly identical frequency curve.

Question 2 generally may be put in this form:

Two frequency curves are given. It may be that the conditions mentioned are fulfilled. Are the differences between the two curves small enough to make it probable, or large enough to make it improbable that this hypothesis is true? A complete answer to this problem may only be given after a further discussion of the form of the frequency curve. It will turn out that it is impossible to determine this form 'a priori', and that all previous theories which attempt to do this were incomplete or premature. Methods for the comparison of frequency curves that are based on those theories are not acceptable. The drawbacks of the method of curve-fitting will then be discussed.

Finally the method of Kapteyn and van Uven will be discussed as a method to compare frequency curves. It will be shown that with its help it is often possible to give a positive answer to question 2.

§ 2. The growth rate of a plant depends on two kinds of factors: the internal and the external.

The most important internal factors are the hereditary qualities of the plant. It is clear that generally two plants with different hereditary qualities shall have different growth velocities, for differences in hereditary qualities appear only as differences in reaction to internal or external stimuli.

Age is also a very important internal factor. This is very clearly shown by some perannial plants, which have a long period of growth. The growth velocity of these plants, when grown under constant external conditions, is not constant. At first they grow slowly, then their growth rate increases, but finally it decreases again. As the external conditions did not change this proves that the internal conditions of the plants were changed during their development.

But if two plants have the same hereditary qualities and the same age *even then* they generally will not have the same internal conditions. For even if their internal conditions were equal originally, differences in the external conditions will make them unequal after a certain time.

We shall call the internal conditions as far as they are only the result of age and of the external conditions the *constitution*. Differences in constitution (in this sense) are therefore *never* a result of differences in hereditary qualities. It is *impossible* to compare the constitutions of plants with different hereditary qualities.

The differences in constitution of two plants as a result of differences in the external conditions may be permanent. This will happen if the plants were influenced by different external conditions during a susceptible period. The susceptible period is a time during which e.g. an organ is formed. If the external conditions influence this formation, two plants which were influenced by different external conditions during this time will remain different during their entire life.

The difference in constitution between two plants, as the result of different external conditions during a non-susceptible period generally may be annulled, and even reversed by opposite differences in the external conditions. But if the external conditions were very bad for one of the plants, it may be that its constitution has deteriorated permanently, an extreme case of which is *death.*

§ 3. The external factors: water, light etc. also have a great influence on the speed of growth. Generally every change in the magnitude of these factors will cause a change in the growth velocity. The magnitude of this change in the growth rate resulting from a certain definite change in the magnitude of one of the factors is a function of the magnitude of all factors. The value of this function becomes very small if the magnitude of any of the factors, except the varying factor, drops below a certain limit, or if the varying factor surpasses a certain limit. In other words, if all the factors have a certain magnitude, the speed of growth will only change appreciably if the magnitude of one well-defined factor changes. This law is called the law of the limiting factors, or Blackman's law. In our case the factor whose change results in a appreciable change of the speed of growth is called the limiting factor. If two plants of the same age, the same heredity and the same constitution, are placed under the same external conditions, then their speed of growth will be equal. If for one of the plants one of the external conditions is altered this only will noticeably influence its speed of growth if that condition happened to be a limiting factor.

§ 4. We have supposed that the length of the plant is only determined by a process of growth. Therefore it will depend primarily on the time that process lasted, and secondarily on the growth rate at every moment of its existence. In general plants of different age will have different length. There are cases where the difference in length is primarily caused by a difference in age as in the case of the length of *Spirogyra*- and *Chaetomorpha*-cells. (Investigations of Baas Becking and Baker). After the cells of those algae have reached a certain length they divide, and start growth again. So the length of their cells is chiefly determined by the time lapsed since the last division.

On the other hand it happens that organisms (or organs) do not grow any more after a certain age: they are "full-grown". The length of these organisms will not depend on their total age, but only on their age when mature. The moment of maturity may depend on internal as well as on external conditions, as e.g. the flowering of the terminal-bud, which process may be determined by the external conditions if the bud has reached a certain development.

The length of plants of the same age depends on their speed of growth on all preceding moments, which, on its turn, depends on the internal and external conditions during those moments. Therefore, the length at a certain moment will also depend on those conditions. Most important of the internal conditions are the hereditary factors. This has been proved by Johannsen in his experiments with beans. Johannsen planted a great number of beans belonging to one sample. The mean length of the beans was different from plant to plant. If he planted a number of beans originating from the same plant, then the mean length of the beans from those plants was the same for all plants. And if he now planted beans originating from different plants (but having the same P_2) the mean length of the beans which came from those plants was also the same for all plants. As beans are always self-fertilized, all the descendents of one bean will have the same hereditary qualities. This experiment proves that the mean length of plants of the same heredity is equal, when grown under practically equal external conditions. Therefore we may conclude that the hereditary qualities of the beans of the original sample were unequal, for the mean length of the beans of the F_1 generation was unequal. Consequently the hereditary qualities of plants of the same species are not necessarily the same. In order to be certain that the hereditary qualities are equal, one must work with plants that are the vegetative offspring of one mother plant. And even then it is always possible to obtain a bud-variation! A reasonable certainty as to the genetic homogeneity exists in plants which are self-fertile, and if one uses the offspring of one plant.

In the second place the length of plants of the same age will depend on their constitution on every previous moment. To recapitulate: Constitution is the internal condition as far as it is determined by the age and the external conditions on all previous moments. Because the sprouts or seeds have passed through a long period of development on the mother plant it is possible that the plants have a different constitution on the moment of planting. In our experiments with *Kalanchoë verticillata* S. Elliot we have measured, three weeks after planting, the width of the leaves which were formed on the mother plant. After eight months the length of the largest plant where this leaf was narrower than 5.45 m.m. was not larger than the length of the smallest plant where this leaf was broader than 7.60 m.m.

In each group there were more than 60 plants. This proves that plants with a different width of the first leaf have a different constitution, and this constitution may differ so much that after eight months it is still apparent from the length of the plants.

If during a susceptible period, the plants grow under different external conditions this will of course highly influence the ultimate length of the plants, for in that case the constitution will always be different.

The external conditions.

§ 5. Except in experiments in the laboratory the external conditions are rarely the same during the whole development of the plant. But even on the same moment they generally will not be the same for different plants. Therefore, even if the constitution and the heredity of two plants are equal, their speed of growth will as a rule differ, because of the differences in the external conditions. On the other hand it is possible that two plants with different constitutions, or even different hereditary qualities, have the same speed of growth as a consequence of differences in the external conditions. But as the external conditions are usually not quite constant this situation will not be permanent as a rule.

It is perhaps impossible to grow plants under absolutely equal external conditions. But it is certainly not impossible to make the conditions so similar that none of the plants is always exposed to more favourable conditions than the other, and that for no period one is able to indicate a condition that is better for one plant than for any other plant. Such conditions we will call *„practically equal"*.

But even if the plants are exposed to practically equal conditions, yet there remain differences that one time will favour one of the plants, another time an other. Besides it is very well possible that two plants may not equally utilize the same favourable circumstance. E.g. the sun is shining some moments. By chance, one of the plants has its leaves perpendicular to the direction of the sun rays, the leaves of the other plant are parallel to this direction.

An other example is given by the budding of new roots. Sometimes the rootlet will find a grain of sand just opposite itself. Than it will take some time before it has passed along that grain. In an other plant a young root will appear between some grains and at once it may start to grow. This are all fortuitous differences. For fortuitous I want to give the definition of Poincaré:

> „Une cause très petite, qui nous échappe, détermine un effet considérable que nous ne pouvons pas ne pas voir, et alors nous disons que cet effet est dû au hasard. Si nous connaissions exactement les lois de la nature et la situation de l'univers à l'instant initial, nous pourrions prédire exactement la situation de ce même univers à un instant ultérieur. Mais lors même que les lois naturelles n'auraient plus de secret pour nous, nous ne pourrons connaître la situation initiale que *approximativement*. Si cela nous permet de prévoir la situation ultérieure *avec la même approximation*, c'est tout ce qu'il nous faut, nous disons que le phénomène a été prévu, qu'il est régi pas des lois; mais il n'en est pas toujours ainsi, il peut arriver que des petites différences dans les conditions initiales en engendrent de très grandes dans les phénoménes finaux; une petite erreur sur les premières produirait une erreur énorme sur les dernières. La prédiction devient impossible et nous avons le phénomène fortuit. (Science et Méthode, pag. 68).

It is clear that in biological events we have often such a disproportion between the differences in the causes and the differences in the effects: a very small difference in the direction of the first named rootlet — and it would have grown at once.

According to Poincaré there is still another case where we speak of fortuitous events, i.e. if the causes are very complex. In that case the final situation is largely independent of the initial situation. But when are the causes complex enough to cause this effect, and when are they too simple to cause this effect?

> „On voit ainsi ce que l'on doit entendre par conditions *trop simples;* ce sont celles qui conservent quelque chose, qui laissent subsister un invariant. Les équations différentielles du problème sont-elles trop simples pour que nous puissions appliquer les lois du hasard? Cette question paraît, au premier abord, dénuée de sens précis; nous savons maintenant ce qu'elle veut dire. Elles sont trop simples, si elle conservent quelque chose, si elle admettent une intégrale uniforme; si quelque chose des conditions initiales demeure inaltéré, il est clair que la situation finale ne pourra plus être indépendante de la situation initiale. (Science et Méthode, pag. 82).

In biology we have this great complexity of the causes. Not only there are a great number of causes that all influence at the same time one biological process, but, moreover, their influence is different at different times.

§ 6. Suppose that one plots, for two different plants, length against age. Probably the two ensuing lines will not coincide, for the speed of growth of the plants is hardly ever equal. But if they have equal hereditary qualities, practically equal constitutions, and if they are grown under practically equal external conditions, their speed of growth will not differ much. One time the growth curve of the first plant will show a steeper rise, another time the second plant will grow faster. It is impossible to determine 'a priori' which of the two plants will be larger at a certain moment, nor is it possible to predict which will be growing faster. But if at a certain moment, one plant is much smaller than the other, one may be sure that, accidently, the external conditions were very unfavourable for that plant (or that from the beginning its constitution was inferior) and that, moreover, the external conditions have caused an inferior constitution. Therefore one may expect that this plant will remain smaller, or at least that the time to overcome its handicap will probably be longer than the time to get it.

On the points of intersection of the growthcurves the growthrate of one of the plants is greater than that of the other. This may result from a difference in the constitution, a difference in the external conditions or from a difference in both. It is even possible that the plant which grows quicker has an inferior constitution but that its external conditions were more favourable (or the reverse).

Will the constitution of the more active plant improve on that moment? Not even that is certain. If that plant grows quicker by reason of its superior constitution, while the external conditions were not so favourable, it is probable that its constitution will be influenced unfavourably. But even if it grows quicker as a result of better external conditions it is not certain that its constitution will improve. This is a consequence of the phenomenon of the limiting factors. Suppose that at a certain moment the temperature is limiting factor. Consequently water and nutritive salts are sufficiently present. The plant which has a warmer place will grow quicker. If this plant has difficulties in pushing out its young root (e.g. caused by a grain of sand), its constitution will be inferior if water becomes limiting factor. It is therefore impossible to say 'a priori' that the constitution of the faster growing plant improves. Often, however, the same factor is limiting at a certain moment for all, or nearly all, processes. In that case the constitution of the fast-growing plant will improve.

§ 7. Suppose three plants *a*, *b*, *c*, with the same hereditary qualities and the same constitution when planted are grown under practically equal external conditions. If the length is determined at a certain moment and the plants are placed in the order of their length, all six possible orders, *abc*, *acb*, *bac*, *bca*, *cab*, *cba* are equally probably 'a priori'.

In two combinations out of six a plant has the same rank. Therefore the chance that a plant has a certain rank is 1/3. What is the chance that a plant *a* has a certain rank if the length of the two other plants is known? Then it is known which of the plants *b* and *c* is smaller and therefore has a lower rank, suppose *b*. The possibilities *abc*, *bac* and *bca* are now excluded. Only the possibilities *acb*, *cab* and *cba* remain. They are of course still equally likely. In each of them *a* has another rank. Therefore the probability of each rank of *a* is in this case also 1/3.

The probability a priori of a certain rank of plants *a* at any moment is therefore, under the above-mentioned conditions always 1/3, whether the length of the other plants at that moment is known, or not. The same holds good for plant *b* or *c*.

If a similar reasoning is applied to (n+1) plants of equal hereditary qualities, equal constitution when planted, and grown under equal external conditions, one finds that the probability 'a priori' for a certain rank for a certain plant is always $\frac{1}{n+1}$ whatever is known about the lengths of the other plants. Even in the case that the length of the n other plants is known, the rank of the (n+1)th plant may have all values from 1 to n+1 (inclusive). An the probability of any rank is still $\frac{1}{n+1}$. If the n measured plants are arranged in order of their magnitude the chance that the (n+1)th plant is smaller than the smallest is $\frac{1}{n+1}$ that it is larger than the greatest is $\frac{1}{n+1}$, and that its length lies between any two given plants which are nearest in rank is also $\frac{1}{n+1}$. The chance that its length lies between the length of two plants differing (g+1) in rank is $\frac{g+1}{n+1}$ (if two plants differ (g+1) in rank there are (g+1) intervals to each of which the (n+1)th plant may belong). If the length of the plant limiting the group at the left side be l_0, and the length of the plant limiting it at the right side be l_1, the above mentioned result may be expressed thus: the change that the length of the (n+1)th plant lies between l_0 and l_1 is $\frac{g+1}{n+1}$.

How large is the probability that the length of the (n+1)th plant lies between L_0 and L_1, if not one of the n measured plants has one of those lengths? Suppose that the length L_1 lies between the lengths of the plants with rank a and a+1, and that the length L_0 lies between the lengths of the plants with rank a+k and a+k+1. Then the probability W that the length of the (n+1)th plant lies between L_0 and L_1 will be smaller than the probability that its length lies between the lengths of the plants of rank a and a+k+1. The probability W is larger than the probability that its length lies between the lengths of the plants of rank a+1 and a+k. The first named probability is $\frac{k+1}{n+1}$ the second $\frac{k-1}{n+1}$. Therefore: $\frac{k-1}{n+1} < W < \frac{k+1}{n+1}$

Likewise: $\frac{k-1}{n+1} < \frac{k}{n} < \frac{k+1}{n+1}$.

If one takes $\frac{k}{n}$ for W the absolute error is less than $\frac{2}{n+1}$.

Bij taking n large enough this error may be made as small as needed.

The relative error is:

$$\frac{\left|W-\frac{k}{n}\right|}{W} < \frac{\left|W-\frac{k}{n}\right|}{\frac{k-1}{n+1}} < \frac{\frac{2}{n+1}}{\frac{k-1}{n+1}} = \frac{2}{k-1}$$

The relative error is therefore smaller than $\frac{2}{k-1}$.

It may now be endeavoured to determine which conditions are sufficient so that significance may be attached to a frequency curve. The argumentation is made simpler by considering the relative frequency curves, wherein the part of the total number of plants which have a certain length is plotted against that length. Such a frequency curve we will denote by (f.c.).

It is clearly impossible to say that a plant has exactly a certain length, but only that its length lies between certain limits. In reality one plots the number of plants, or the part of the total number, against the size-classes.

§ 8. A frequency curve is clearly without significance if the material gives two very different (f.c.)'s when divided at random in two parts (each containing of course a great number of individuals). It is now possible to establish the conditions which are sufficient for the practical identity of the (f.c.)'s of two random parts of the material:

1. The individuals should have the same hereditary qualities.
2. When planted they should have had the same initial constitution.
3. They should be grown under practically equal external conditions.

For if the first lot contains n individuals (n very great) and if there are g individuals between the lengths l_0 and l_1, the probability that the length of any individual of the second lot lies between l_0 and l_1, will be $\backsim \frac{g}{n}$. If the second lot contains also very many individuals, the $\left(\frac{g}{n}\right)^{th}$ part thereof will lie between the lengths l_0 and l_1, according to the law of Bernouilli. Consequently the same part of the whole material as in the first lot will be situated between those lengths. Therefore the two frequency curves are (nearly) identical.

Consequently those conditions are sufficient in order that the frequency curve be of any significance. But are they necessary and are they realizable? The second question certainly should be answered in the negative; the condition that the plants should have the same constitution when planted cannot be realized. For, as already stated, the constitution of a plant at a certain moment depends upon the external conditions during its development up to that time; and sprouts as well as seeds have developped on the motherplant during a certain time. And even if the external conditions (sensu stricto) were absolutely equal for two sprouts or seeds (which never happens), even then their constitutions probably would be unequal, for their connection with the mother plant would not be absolutely equal.

It may be proved that two sufficiently large groups will give the same (f.c.)'s if the differences in constitution are very small, so small in fact, that for no sprout (seed) one has a reason to suppose that its constitution is better, or worse, than for any other plant. (It is supposed of course that all plants have the same hereditary qualities and that they grow under practically equal conditions).

Suppose that among the material, from which the sprouts are chosen, occur a certain number of different constitutions. If the differences in constitution are so small as to be indetectable, it is possible to choose at random. (If they are not so small this is not always possible). Therefore if one selects twice a very great number of plants (from the same material), the proportion of the number of plants of one constitution to the whole number of plants chosen will be nearly the same in both groups, according to the law of Bernouilli. The (f.c.)'s of plants of the same constitution are equal. In both groups the (f.c.)'s of the entire material are the sums of the (f.c.)'s of plants of the same (initial) constitution, multiplied by the proportion of the plants of that constitution to the whole material. As the partial (f.c.)'s for the two groups are equal, and as they are each multiplied by the same number, their sums, the total (f.c.)'s, will be equal as well.

It is thus proved that a frequency curve has only practical significance if:

1. All plants have the same hereditary qualities.
2. All plants were of a nearly equal constitution when planted.
3. All plants have been grown under practically equal external conditions.

We shall call those conditions the *conditions A.*

The foregoing reasoning could also be applied if the original material, i.e. the material from which the sprouts or seeds are chosen, consisted of sprouts or seeds of unequal hereditary qualities. It would be found, by a similar reasoning, that, if two groups were chosen at random from the same material, and were grown under practically equal external conditions, those two groups would give (nearly) equal frequency curves. But it may not be thought judiceous to give so much latitude to the first condition, for if the original sample of sprouts is exhausted, the experiment cannot be repeated, unless the composition of the material was exactly known, in which case one could have worked equally well with homogeneous material.

If one considers a material with unequal hereditary qualities it will therefore be impossible to deduce general laws from the experiments. The same drawback does not hold for a material with small differences in the constitution, which involves gradual changes. Whenever the mother plants were not exposed to extremely variable external conditions the sprouts or seeds will have, in part, the same constitution. If one does not select extreme variants all the sprouts or seeds, therefore, should have practically equal constitution.

§ 9. Our first problem is now completely outlined. Only when the conditions A are satisfied a frequency curve will have any quantitative value. Sometimes, however, the inverse problem is given. One has a certain frequency curve, but does not know if the first of the conditions A (genotypic homogeneity) is satisfied. If the frequency curve is bimodal this may be a result of genotypic inhomogeneity. But only by experiments genotypic inhomogeneity may be proved. For genotypic homogeneous material may also give a bimodal frequency curve, as e.g. in the experiments of miss Tammes on the stalklength of Linum crepitans (Kapteyn and v. Uven p. 68). We have also found, when working with homogeneous material, a case of a bimodal frequency curve. The frequency curve of the difference in blade-length of the two leaves of the third pair of *Kalanchoë verticillata* S. Elliot was bimodal.

To be able to answer the second question:

„What may be concluded from given frequency curves?" one must at first ascertain with which aim they are composed. Sometimes one wants to ascertain if there are reasons to suspect genotypic inhomogeneity.

The frequency curves of a characteristic of plants with a different heredity will in general be different. Therefore the frequency curve of a mixture of those plants will be bimodal if the positions of the maxima of the frequency curves of plants of equal genotype *differed enough.* If the positions of those maxima differ but little, the frequency curve will be unimodal. Therefore bimodality *may* indicate genotypic inhomogeneity. Unimodality does not indicate homogeneity.

In most cases however frequency curves are set up to ascertain if, for two groups of plants, the same conditions A are satisfied. This problem is usually given in this form: A certain characteristic is measured for two groups of plants. It is known that the plants of both groups satisfy two of the three conditions A. Is it possible to ascertain from the frequency curves if they also satisfy the third condition? If the two curves are nearly identical the probability that all plants satisfy the third condition is very great. If the difference between the (f.c.)'s is very great this probability is very small.

In other words, if it is known that

1. There are no reasons to suspect that the sprouts or seeds had a different constitution when planted,
2. The plants were grown under practically equal external conditions

a great similarity between the two (f.c.)'s signifies that the plants of both groups were probably of the same heredity. A great difference between the curves signifies that the plants were probably of different heredity.

If it is known that

1. The plants are all of the same heredity

2. The plants were grown under practically equal external conditions

a great similarity between the two (f.c.)'s signifies that the plants of both groups had probably the same constitution when planted. A great difference between the curves signifies that the plants had probably a different constitution when planted.

If it is known that

1. The plants are all of the same heredity,
2. The plants had all the same constitution when planted

a great similarity between the two (f.c.)'s signifies that the plants of both groups were grown under practically equal external conditions. A great difference between the two curves signifies that probably the plants of the one group were grown under other external conditions than the plants of the second group.

Usually the last-mentionned problem is given in this form:

„It is known that all plants of two groups have equal heredity, had equal constitution when planted and that the external conditions were all equal except one". A great similarity in the (f.c.)'s of the two groups signifies in this case either that the factor which was different in both groups had no influence on the process or at least that it was not the limiting factor for that process. A great difference in both curves signifies that the unequal factor has an influence on the process.

The most important question which must be answered is therefore: It is possible that all plants of two groups satisfy the conditions A. The (f.c.)'s of a certain characteristic of the two groups have a certain difference. How great is the probability 'a posteriori' that the plants of both groups satisfy the conditions A?

The solutions which are given for this problem generally are founded on certain hypotheses on the origin of frequency curves. We will discuss the most important of them.

§ 10. The oldest hypothesis about the form of the frequency curve supposes that all frequency curves are identical with the error curve of Gauss. As far as we know this hypothesis has been set up by Quetelet. There are two demonstrations that the frequency curve must be identical with the error curve, but both are based on incorrect suppositions.

In the first demonstration (mentioned e.g. in Bauer, Vererbungslehre p. 18) it is supposed that:

1. A plant is influenced during its development by a great number of factors.

2. Each plant is influenced by the same number of factors.
3. Each factor is either favourable or infavourable for a certain plant.
4. The probability that a certain factor is favourable for a certain plant is always ½.
5. The length of a certain plant is a linear function of the number of favourable factors that influenced its growth.

Two plants which were influenced by the same number of favourable factors will have therefore an equal length whether those factors were the same or not.

Suppose that n factors influence the plants during their growth. The probability 'a priori' that for a certain plant (a) factors were favourable and therefore (n—a) unfavourable is

$$\frac{n!}{a!(n-a)!} \cdot \frac{1}{2^n}$$

All plants for which (a) factors were favourable will have the same length $(pa+q)$, (in which p and q are constants) according to condition 5.

Therefore the probability that a plant will have the length $(pa+q)$ is $\frac{n!}{a!(n-a)!} \cdot \frac{1}{2^n}$

If there are many plants, the number of plants with the length $(pa+q)$ will be $\backsim$ the $\left(\frac{n!}{a!(n-a)!} \cdot \frac{1}{2^n}\right)^{th}$ part of the total number of plants, according to the law of Bernouilli. If the part of the total number of plants which have a certain length $(pa+q)$ is plotted against that length and if those points are joined by a smooth line the theoretical (f.c.) is obtained. In this case the curve is called „binomial curve".

If n is very large the binomial curve approximates the error curve of Gauss $y = \frac{k}{\sqrt{\pi}} e^{-k^2(M-x)^2}$, at least if the units for the abscissa and the ordinate are suitably chosen. In this formula M is the mean length of all plants, x is the length of an individual plant, and k depends on the unit of measuring.

It is clear that the fundamental suppositions are hardly ever realized. The supposition 4 (the probability that a certain factor is favourable for a certain plant is always ½) implies that all factors are independent. This is very improbable.

But supposition 5 (the length of a plant is a linear function of the number of favourable factors) is in direct contradiction with the phenomenon of the limiting factors. E.g. suppose there are only two factors. Each factor may be favourable or unfavourable; *A* (favourable), *a* (unfavourable), *B* (favourable), *b* (unfavourable). There are four possible combinations: *AB*, *Ab*, *aB*, *ab*. According to hypothesis 4 these are equally probable. According to hypothesis 5 the length of the plants would be respectively (2p+q), (p+q), (p+q), (q).

But if in the cases *AB* and *Ab* the second factor is limiting, and in the cases *aB* and *ab* the first, the length of the plants will be respectively (2p+q), (p+q), (q), (q). The resulting frequency curve, in this case, will be non-binomial. This reasoning is therefore incompatible with reality.

§ 11. There is still another reasoning to prove that in biology the frequency curve should be the error curve of Gauss. First, the whole growth period is divided into a large number of elementary periods. It is supposed that the increase in length of a plant during an elementary period depends only upon the external conditions during that period; furthermore that the probability of a certain constellation of the external conditions at a certain moment is equal for all plants. This equality implies that the probability of a certain constellation for a certain plant does not depend upon the constellation during a former period. Finally it is supposed that the mean increase of all plants during an elementary period is equal to the median increase.

The total length of a plant is the sum of the increases in length during the elementary periods. Such an increase in length may be conceived as the algebraic sum of the mean increase in length and the deviation from that mean increase for that particular plant. The total length of a plant, therefore, will be equal to the mean length of all plants plus the (algebraic) sum of its deviations from the mean increase during the elementary periods. The hypothesis that the mean increase during an elementary period is equal to the median increase implicates that the probability of negative and positive deviations from that mean are equal. Therefore the length of a plant is a constant (the mean length of all plants) plus the sum of a great number of quantities which are as likely positive as negative. The probability that a certain plant has a definite deviation from the mean length of all plants is therefore equal to the probability that the sum of the quantities aforementionned has a certain value. It has been proved by Bessel that this probability follows the law of Gauss. Therefore the frequency curve of the length of the plants will also be equal to the error curve of Gauss.

There are two serious objections against the hypotheses upon which this deduction of the form of the frequency curve is based:

1. The supposition that the increase in length of a plant depends only upon the external conditions will be rarely true. If two plants have a different constitution their increase in length during an elementary period will be different if the external conditions are equal.

And generally the constitutions of two plants are different. Therefore the action on a plant of a certain constellation of the external conditions will depend upon the constellation of those conditions at previous moments as this constellation has influenced the constitution of the plant.

2. But, moreover, the constellation of the external conditions during successive periods will not be independent as those conditions change gradually.

Therefore the suppositions on which this deduction of the form of the frequency curve is based are not realized in nature.

§ 12. Why has it so often been tried to prove this identity with the error curve of Gauss? The reason seems to be primarily historical. About 1835 Quetelet had observed that the frequency curve of the breast circumference of Scottish conscripts was nearly identical with the binomial curve, which passes into the error curve when the exponent n becomes very large. Quetelet says:

> L'exemple que je viens de citer (viz. the frequency curve of the breast circumference of Scottish conscripts) mérite je crois notre attention: il nous montre que les choses se passent absolument comme si les poitrines qui ont été mesurées avaient été modeleées sur un même type, sur un même individu, ideal si l'on veut, mais dont nous pouvons saisir les proportions par une expérience suffisamment prolongée. Si telle n'était pas la loi de la nature, les mesures ne se grouperaient pas, malgré leurs défectuosités, avec l'étonnante symétrie que leur assigne la loi de possibilité. (Quetelet, 1846, p. 137).

In point of fact Quetelet gives only an experimental justification of the identity of the frequency curve and the binomial curve. Of the modern writers E. Borel shares this point of view (E. Borel, 1924. p. 127, p. 178). As soon as it is shown that often this identity does not exist the reasonings of Quetelet and Borel lose their value. And many frequency curves are found which are no binomial curves.

How is it then that in computing the results of biological experiments the methods used are generally based on the normal law of error, that is to say: it is supposed 'a priori' that the frequency curve is identical with the error curve of Gauss? Probably this is a consequence of the fact that most biologists forget that in biological experiments deviations from the mean usually are caused by the variability of the material. Those deviations are therefore not true errors, as the errors in physical experiments for instance. But it is usual to give them the name of error and so it is forgotten what they really are.

In answer to the objections of Bertrand (Bertrand 1907, p. 170 et sqq.) Poincaré gives a very general proof of the identity between the error- and the Gaussian curve. He puts the problem in the following way: Suppose z is the real value of the quantity which is measured. How large is the probability that the result of a measurement will fall between (x) and $(x+dx)$? Without any hypothesis this problem is insoluble. Poincaré supposes that:

1. This probability is proportional to dx.
2. The mean value of x is equal to the probable value of z.

He puts the probability that, the true value being z the result of a measurement will fall between (x) and $(x+dx)$

$$dW = \phi\ (x,z)\ dx$$

Clearly $y = \phi\ (x,z)$ is the equation of the error curve.

Which meaning may be attached to (x), (z) and $y = \phi\ (x,z)$ in variation statistics?

x is clearly the length of a single individual. In § 7 it has been demonstrated that a frequency curve has only any practical value if the material satisfies the conditions *A*. In that case a repetition of the experiment will give a practically identical frequency curve. If the (f.c.)'s of two groups of plants are equal the mean length of those plants will also be equal.

If the external conditions had been different probably the maxima of those (f.c.)'s would have been at another place. Then their mean length would have been different. Therefore the mean length may be taken as a measure for the external conditions. In that case $y = \phi\ (x,z)$ is the form of the frequency curve if the external conditions were such that the mean length of the plants is (z).

We have now identified our problem with the problem of the form of the error curve as put by Poincaré. Therefore the solutions of both problems should also be equal. Poincaré finds for the formula of the error curve

$$\phi\ (x,z) = \Theta\ (x)\ e - \int \psi\ (z)(z-x)\ dz.$$

$\Theta(x)$ is a function of x only, which indicates to what extent some values of x have a larger probability than other values, irrespective the value of z. E.g. in discontinuous variability $\theta (x)$ will be zero for any fractional value of x, and it will be equal to $\frac{1}{dx}$ for any integer value of x. In continuous variability $\theta (x)$ generally will be constant.

$\psi (z)$ is the probability 'a priori' of a certain value of z. In the calculus of errors $\phi (x,z)$ signifies the probability of committing an error $(x$-$z)$. Poincaré reasons that the probability 'a priori' of an error will not depend upon the probability 'a priori' of the magnitude of the measured quantity. Generally this will be true, but not always. E.g. the error made when the place of a star low at the horizon is determined probably will be greater than the error made when the star shines high above the horizon. But if one accepts the hypothesis of Poincaré that the probability of an error does not depend upon the probability 'a priori' of the magnitude of the measured quantity, then the error curve will be the curve of Gauss multiplied by $\theta (x)$. For in that case $\psi (z)$ is constant and $\int \psi (z) (z—x)\, dx = k^2 (z—x)^2$.

$(z—x)$ is the error.

In variation-statistics $\phi (x,z)$ is a measure for the *absolute* variability. But it is certain that the *absolute* variability depends upon the mean length of the plants. The absolute variability of oaks hundred years old will be larger and much larger than the absolute variability of oaks a hundred days old. Therefore, in the case of variation-statistics, one may not put ψ (z) = constant, but then the frequency curve will not be the error curve of Gauss. Only in cases where the absolute variability does not depend upon the mean length it may be expected that the form of the frequency curve becomes identical with the form of the error curve of Gauss.

§ 13. Rather early in literature it was observed that there are cases where the normal curve does not fit a given frequency distribution. Pearson has given a number of equations to fit any given frequency distribution: Though the original idea was to generalize the binomial curve by supposing that not all factors were independent, the constants of the found frequency curves have often such a value that a biological interpretation is rarely possible. Virtually all those frequency curves are only graduation curves as Pearson has now admitted in a letter to the Editor of 'Nature' (24 Aug. 1935 p. 296). A serious drawback to all such graduation curves is the absence of any causal relation between the constants of the curve and the factors that have caused the frequency distribution. It is possible that a small difference in those factors brings about a great difference in the constants of the curve, or even a difference in the formula, thougth the form of the frequency curve has changed but little.

§ 14. Edgeworth and Kapteyn have used another method to determine the form of the frequency curve. Their startingpoint is the thesis that a non-linear function of a quantity which is distributed according to the normal or Gaussian law cannot be distributed according to that law. Bertrand has also made this remark. He based upon it a severe criticism of the error curve of Gauss. Pearson criticizes the method of Kapteyn and Edgeworth. According to Pearson the relation between two quantities is, in biology, never causal but only correlational. His illustration of this remark gives the impression that he is trying to split straws. He says (Biometrika 4, 1905 p. 199):

> The sagittal arc in English women is almost exactly normal in its distribution and nasal breadth is very asymmetrical. Shall we take x sagittal arc and X nasal breadth and make $x = F(X) - M$?

Anyone who has read the paper of Kapteyn wil know that Kapteyn means causal relation as e.g. in the case of the diameter and volume of spherical berries. If the frequency curve of the diameter of those berries is a Gaussian curve, the frequency curve of the volumes should be asymmetrical. If the range of the diameter is large the asymmetry will be very pronounced.

That the frequency curve of the diameter and the volume of a group of nearly spherical berries are never at the same time Gaussian curves is easily demonstrated by an example. Suppose there are n berries with a diameter between 1 and 2, and also n with a diameter between 5 and 6 (at both sides of the maximum of the frequency curve). The volumes of the first-mentioned berries will lie between $\frac{4}{3}\pi$ and $\frac{4}{3}\pi.\ 8$. The volumes of the other group will lie between $\frac{4}{3}\pi.\ 125$ and $\frac{4}{3}\pi.\ 216$. Suppose the class width of the volumes is $\frac{4}{3}\pi.\ 7$. The first named n plants will all fall into one class, the other n (whose diameter lies between 5 and 6) will fall into 13 classes. Therefore the frequency curve of the volumes will be skew. It is clear that in this case the quantities which follow the normal law are no shadow quantities (Pearson p. 201).

In the same paper Pearson gives $\frac{1}{y} \cdot \frac{dy}{dx} = - \frac{x}{\sigma_0^2 F(x)}$ as the only form of the differential equation of the frequency curve. As Kapteyn has shown (Kapteyn 1916) this is simply the Edgeworth—Kapteyn formula in the differential form. Only the interpretation of F(x) is different. Pearson calls $\sqrt{F(x)}$ the standard deviation of the 'instantaneous Gaussian curve'. Edgeworth and Kapteyn suppose that there is a quantity $y = F(x)$ which is distributed according to the normal law. The difference is thus only a question of the interpretation of F (x).

There is however one point in the theory of Kapteyn which hardly has been discussed by Pearson and this is its most important part. As we have seen the supposition that the increase in length of a plant during an elementary period depends only upon the external conditions, and that the external conditions on a certain moment are independent of the conditions during any previous moment, leads to a normal or Gaussian frequency curve. The fault in this reasoning is that it does not take into account the constitution of the plant. Kapteyn supposes that the length is a good measure for the constitution of a plant. The growth rate of two plants of the same length should only depend upon the external conditions. If the external conditions were equal for two plants the growth rate would only depend upon the length of those plants. From those hypotheses and the additional one that the external conditions on two successive moments are independent, Kapteyn is able to deduce the form of the frequency curve.

$$y = \frac{h\,F^1(x)}{\sqrt{\pi}}\,e^{-h^2\,[F(x) - F(x_0) - M]^2}$$

$\frac{k}{F'(x)}$ is the growth rate of a plant of the length x (k denotes the magnitude of the external conditions). From a given frequency curve F (x) may be deduced, as Van Uven explains on p. 35 (Kapteyn and Van Uven 1916). By graphical differentiation $F'(x)$ and $\frac{1}{F'(x)}$ may be deduced from F (x). If the suppositions of Kapteyn were true it would thus be possible to determine, from a given frequency curve, how the growth rate of a plant depends upon its length under constant external conditions. But the fundamental supposition of Kapteyn, that the constitution of a plant is unambiguously determined by its length is certainly not correct. The constitution of plants of the same length but of different ages is different. But if the constitution is different the growth rate will be different as well. If the growth velocity as function of the length is calculated at different moments according to the method of Kapteyn and v. Uven, the results will be different, as happened in our experiments (p. 119).

There is still another objection against the deduction of Kapteyn. The external conditions at two consecutive elementary periods are supposed to be independent, and this never happens. Almost all external factors change gradually. If the soil around a plant is humid, it will not be dry some moments later. The reverse may sometimes happen but than the soil will remain more or less humid for some time. Neither does the temperature change abruptly. And even the light will usually alter gradually. Only at certain moments when the sun appears suddenly from behind the clouds, the changes in the intensity of the light will be very large. Though theoretically these objections are very grave they have probably no influence on the *results* of Kapteyn. For it is possible to replace the supposition that the external conditions on consecutive moments are independent by another hypothesis which is a better image of nature, and still obtain the same result as Kapteyn did.

Suppose that during a certain time factor *a* is limiting. The magnitude of that factor may have 2 k different values. We shall suppose that the growth rate is a linear function of the magnitude of factor *a*. If the magnitude during a certain elementary period is g we will suppose that the probability that during the following period it is (g—1) is $\frac{g}{2k}$, and that during the following period it is (g+1) is $\frac{2n-g}{2k}$. If k is very large the magnitude of the factor *a* changes gradually. A further advantage of this hypothesis is that deviations from the mean of the magnitude of the factor tend to become smaller. If the magnitude of a limiting factor changes always in this manner from elementary period to elementary period the calculations of Kapteyn and van Uven still hold good. (For the calculation vide p. 123).

Our general conclusion on the theories which explain the form of the frequency curve may be stated as follows.

The growth rate of a plant depends upon its heredity, its constitution at all moments of its growth and also depends upon the external conditions. The constitution of a plant depends upon its age and on its previous history. Therefore it is impossible:

1. To determine 'a priori' the form of the frequency curve.

2. To deduct 'a posteriori' the relation of length and growth-rate from the form of the frequency curve.

§ 15. As we have said the problem of the interpretation of frequency curves is usually given in this form: „The frequency curves of the same characteristic of two groups of plants are given. It is 'a priori' possible that for both groups the conditions A were the same. How large is the probability 'a posteriori' that this equality of the conditions A has existed?"

The school of 'Student' and Fisher formulates this problem in another way. They consider a given frequency distribution as a random sample from a very numerous assemblage of plants. It is supposed that all those plants satisfy the conditions A. Such an assemblage they call 'Universe'. If a random sample is taken from that 'Universe' how large is then the probability that the mean of the sample differs a certain number of times the standard deviation (of the sample) from the mean of the 'Universe'. To answer this problem the law underlying the frequency distribution of the 'Universe' must be known. If the means of two given frequency curves differ much, the probability that they are random samples of the same 'Universe' will be very small, independent of the way in which the composition of this 'Universe' is chosen. In that case it is very improbable that the conditions A were satisfied for both groups. But if the difference between the frequency curves is not large it will be possible to find such a composition of the 'Universe' that it is not improbable that both are random samples from that 'Universe'. If it was 'a priori' possible that the conditions A were satisfied, one may now conclude that it is very probable that they were satisfied.

Is it possible to attach any biological significance to the notion 'Universe', and in how far is it possible to regard a given frequency curve as a random sample from that 'Universe'? As we have demonstrated, if the conditions A are satisfied, the (f.c.) of any characteristic will tend to a certain limit if the number of individuals is very great. This limit-(f.c.) may be regarded as a description of the law underlying the frequency distribution of the 'Universe'. And if one has only measured a limited number of individuals those may be regarded as a random sample from that 'Universe'. It is thus possible to attach significance to the notions 'Universe' and 'random sample'. But why are those notions necessary?

The whole theory of errors is based upon the supposition that there are a great number of observations. (As we remarked already this theory is in an erroneous way also applied by biologists, though here the cause of discordance between the observations lies in the variability of the material). As in general the real value of the quantities measured and of the errors is not known, they are simply equalized to the mean value of the observations etc. According to the law of Bernouilli the difference between the probable and the mean value of a quantity tends to zero if the number of observations tends to ∞. Therefore in this case $(n \to \infty)$ the probable value of the observed quantity and of the error are known with great accuracy. If the real value of mean and probable error are known it is possible to determine the probability that, in a new set of observations, the mean will differ a certain quantity from the original mean. But if the number of observations is not very great all foregoing reasonings lose their value and it is impossible to determine with any accuracy the probable value of the observed quantity and of the probable error. But then it is also impossible to determine the probability that in a second set of observations the mean will differ a certain quantity from the original mean. This inconvenience does not exist in the method of 'Student' and Fisher, for they work only with the mean and standard deviation of the sample, which are entirely known, and with the mean of the 'Universe', which is given 'a priori'. This seems a great improvement but....there is one great drawback. To apply this method the law of the frequency distribution of the 'Universe' must be known 'a priori', and in biology this rarely, if ever, happens. Therefore in biology the method of 'Student' and Fisher seems in general to be unapplicable.

§ 16. It seems thus that there is no unobjectionable method to compare frequency curves by calculation. A direct comparison has also great drawbacks, a difference between two (f.c.)'s in the neighbourhood of their maxima has another significance as a difference at their ends. Particularly when there is 'high contact'. Furthermore the form of the (f.c.) is generally rather complicated, differences in form are therefore difficult to observe. The most practical method is therefore to compare the deviations of the (f.c.)'s from any well-defined standard curve. This curve should have the following properties:

1. Its range must be unlimited. It is true that the range of a biological characteristic is never unlimited, but at the other hand it is hardly ever possible to state of any individual that the characteristic has a definite maximum value.

2. Its ordinate should be known or should be easily calculated for each value of the abcissa.

3. Its form must not deviate too much from the general form of (f.c.)'s, otherwise the curves of the deviations from the standard curve will have a complicated form, and then our goal is not reached. The error curve of Gauss satisfies all those conditions. Therefore the best method of comparing (f.c.)'s is to compare their deviations from the error curve of Gauss. But the method of doing this is equal to the method of Kapteyn and van Uven to determine the function $z = F(x)$ which is distributed according to the normal law described in their second paper (Kapteyn and v. Uven 1916). In the same paper the formula of the probable error of the ordinate z of the curve $F(x)$ is given. It is thus possible, of two (f.c.)'s are given, to determine if their abscissal distance between points of equal ordinates is at all points less than the probable error, in which case conditions A were probably satisfied, or if their 'distance' is at any point more than three times the probable error, in which case the conditions A were, probably, not satisfied. If the form of the two curves differs much it is also very improbable that the conditions A were satisfied. In this case, moreover, generally some points of the second curve will deviate more than three times the probable error from the corresponding points of the first curve.

II. STATEMENT OF THE PROBLEM, MATERIAL AND METHODS.

§ 1. In the introduction it has been demonstrated that it is impossible to determine 'a priori' the form of the frequency curve of the length of plants. The demonstrations that the frequency curve must be identical with the error curve of Gauss were unacceptable because they were based upon suppositions which are not realized in biology, e.g. the constitution of the plants is not taken into account. And even the theory of Kapteyn, which is a great improvement (for it takes this constitution into account) is unacceptable; its measure for the constitution is incorrect. We have tried to demonstrate experimentally that those objections against the theory of Quetelet c.s. and of Kapteyn are of practical importance.

The easiest method to demonstrate the validity of our objections against the theory of Quetelet c.s. is to prove that the conclusions are faulty. Therefore it is necessary to determine the (f.c.) of any characteristic (e.g. the length) of a number of plants satisfying the conditions A. According to Quetelet one should find the error curve of Gauss. In none of our experiments this happened. Neither have we found in the litterature any example of the Gaussian curve if the conditions A were satisfied (e.g. Johannsen 1903. Pearl 1907. Dennert 1926).

A second method to justify our objections is to demonstrate directly that the growth-rate depends upon age and previous history. For this method it is necessary to measure regularly the length of a great number of plants which satisfy the conditions A. In this manner it is possible to determine the growth-rate of plants of equal age but of different length and of plants of different age and of equal length. According to the adherents of the theory of Quetelet the growthrate must be independent of age and of the length of the plants. According to Kapteyn it must be independent of age but may depend upon the length of the plants. It will appear that it depends upon both.

A third method is to compare the (f.c.)'s of a certain characteristic if the plants are classified according to another characteristic. If the magnitude of a certain characteristic is different for two groups of plants their previous history will have been different. But then their constitution is unequal at that moment. If after that moment another characteristic develops it will originate from plants with different constitution. Therefore the (f.c.)'s of the two groups of plants will be different for the second characteristic.

In our experiments we have measured the breadth of the first leaf a short time after the sprouts of *Kalanchoë* were planted. There existed practically no stem at that moment. But even at the concluding measurement there was an obvious difference in the frequency curves of the length of those plants, whose breadth of the first leaf was different some 8 months ago.

§ 2. To execute those experiments it is necessary to have a number of plants satisfying the conditions A. The first condition is *hereditary homogeneity*. There are two ways to obtain this homogeneity. Either one should measure the same organs of one plant, or one should measure the length of plants which are the vegetative ofspring of one mother plant. (Virtually those sprouts may also be regarded as the same organs of one plant). We have followed both ways.

§ 3. In the first place we have measured the length and the breadth of the foliola of full-grown leaves of *Robinia Pseudacacia* L. The experimental plant was a shrub growing at a distance of about 2 M. from some other shrubs which screened it against the W. and N. winds. At the E. and S. side it was screened by wood. The shrub originated probably from the root of a Robinia-tree some 4 m.· distant. Its age was about 4 years. The leaves were gathered and measured in the last week of August. It was rather difficult to determine whether a leaf was full-grown. But as the total length of the full-grown leaves increases from the base to the apex of each branch, those leaves were considered full-grown which were not smaller than their predecessors.

The foliola were measured by placing them on a piece of ruled paper. The distance of the lines was 1 mm., ½ mm.'s were estimated. The number of observations ending in entire mm.'s was twice the number ending in ½ mm.'s. This is of cours caused by a faulty estimation. Dennert (working with the needles of *Taxus baccata* L) has probably made the same mistake, but he ascribes the differences to a real fluctuation or 'Rythmus' of the frequency distribution. (Dennert 1926, e.g. fig. 11). For the greater number of computations we have therefore joined two succeeding classes to one class.

§ 4. Our second set of observations were made on *Kalanchoë verticillata* S. Elliot, a Crassulacea. *Kalanchoë verticillata* is a herbaceous perennial plant. Its lower leaves are decussate, but higher on the stem the leaves alternate. At the top of the leaves there are 4-12 dents. On these dents the sprouts are formed. These sprouts consist of two leaves placed on a 'disc'. The smallest of those leaves is directed toward the mother plant. (Fig. 1). If those sprouts are not too young they are easily separated from

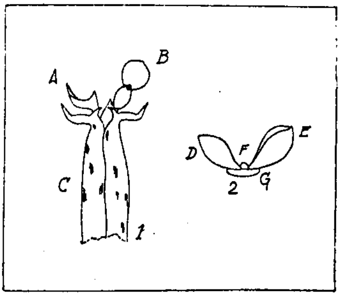

Fig. 1. 1. Apex of leaf of *Kalanchoë verticillata* S. Elliot. A. 'Dent'. B. Sprout. C. Blade of motherplant. Nat. Size.
2. Sprout. D. Smallest 'cotyledon'. E. Largest 'cotyledon'. F. Apical bud. G. 'Disc'. x 2.

the mother plant. If such a detached sprout is laid on the soil the disc forms roots and the bud between the leaves develops into a stem. The first year the plant generally bears no flowers. As long as the apical bud exists the lateral buds do not develop.

In December 1933 450 sprouts originating from one plant were placed on moist sand. In march 1934 each young plant was placed in a flowerpot. Those flowerpots were placed in a hot bed. By the beginning of July the length of the stems was measured, and the number of leaves counted. As the plants did not all grow on the same distance from the glass their external conditions were not practically equal. Therefore those frequency curves have no very great value, but they were not Gaussian.

96 of those plants which showed about the same length were used as mother plants for our definitive experiment. From the apex of three full-grown leaves of each of those plants two sprouts were taken. These sprouts were planted in 48 wooden boxes. In each box 12 plants were placed deriving from two mother plants.

The wooden boxes had an internal diameter of 5 × 5 × 55 cm. Five holes were drilled in the bottom for draining purposes. On the bottom was a layer of well-washed gravel about 1 cm. high. The boxes were filled with sand which was washed seven times in ordinary water.

They were placed in eight rows of six before the window of a room on the S. side of the laboratory. All boxes stood parallel to the window, the rows were perpendicular to the window. The temperature of this room was not constant, but it sunk never below 4°C. Only during the end of December and the beginning of January it was necessary to heat the room.

As the posterior boxes were further from the window every other day all boxes were displaced. If the original order was 1, 2, 3, 4, 5, 6 then after the displacement the order was 6, 1, 2, 3, 4, 5. Moreover each box was turned 180° and each row was displaced as a whole one place to the right. The row at the extreme right side being placed at the left side.

If the soil was dry the boxes were placed in water which was sucked up through the holes in the bottom. About once in six weeks a solution of v. d. Crone (to which was added traces of $H_3 BO_4$ and of $Mn SO_4$) was used instead of ordinary water. If the weather was very sunny the plants were, moreover, sprinkled with water.

After two weeks two of the boxes were upset and their plants spoiled. There remained therefore 46 boxes containing 552 plants. Of those plants one has died, all other developped more or less. The experiment was terminated on the 4th of July 1935. On October 24 the breadth of the largest of the first leaves (which have developped on the mother plant) was measured, with the aid of measuring calipers. The accuracy of this instrument was 0.05 mm. At November 16, January 11, February 6, March 11, March 27, April 30, May 15, June 11, July 5 the length of the plants was measured from the lowest leaves to the terminal-bud. A piece of ruled paper was placed besides the plants on which the length was read. Tenths of mm.'s were estimated as long as the plants were lower than 3 cm. From plants which were higher than 3 cm. only the mm.'s were noted.

On December 10 the length of the leaves of the third pair (when counted from the bottom) were measured.

On November 16, January 16 (F), February 13 (H), March 21 (L), April 8 (N), May 2, May 21 the number of full-grown leaves of each plant was counted. Sometimes a plant dropped one or more leaves but the leafscars, which were always readily recognizable, were then counted as leaves. It was not easy to determine what should be reckoned as the highest full-grown leaf pair, for the terminal-bud continually forms new leaves. We reckoned those leaves as full-grown which did not touch another.

How far did our plants satisfy the conditions A? The first condition was certainly satisfied for all plants had the same heredity, because they all originated from the same P_2. Our second condition was that the plants must have had a practically equal constitution when planted. Although all the mother plants grew under practically equal conditions the constitution of the sprouts was not equal. For the age of the sprouts, when detached from the mother plants was probably unequal, as it was very difficult to determine if the leaves of the mother plants on which they grew were just full grown. The older a sprout the more reserve food it has stored, and the better its constitution (and inversely). In all those cases in which the quantity of food stored in sprouts or seeds is small, this quantity will chiefly determine the constitution of those sprouts or seeds. For this food has to supply the energy to form roots. And the more and the stronger roots are formed, the better the plant will grow. But if the food in the sprouts or seeds is superabundant all plants may form their roots equally well, and the differences in the quantity of the food do not cause a difference in constitution.

As sprouts formed on leaves are generally small their quantity of reserve food (and therefore their magnitude) will generally be a measure for their constitution. The best manner to obtain plants of equal constitution would be to weigh them, but as they are very small this would take much time. It would be impossible to plant many sprouts on the same day, in other words the plants would not have all the same age. The second best method to select a great number of plants of the same constitution is to measure, some weeks after the sprouts are planted, and before the stem develops, the breadth of the largest leaf. If one chooses now for the experiment only those plants whose leafbreadth was nearly equal, they will probably have a nearly equal constitution. We have followed more or less this way in our experiments.

The third of the conditions A was that the plants should have been grown under practically equal external circumstances. By practical equal external circumstances was meant that for no period it was possible to indicate a factor which was more favourable for one as for another plant. By our regular displacements each box occupied all possible positions in relation to the windows, in the course of time. As far as light and temperature is concerned the conditions may therefore be designated as practically equal. As the plants were always watered on the same day the humidity and the supply of mineral salts was also practically equal.

The only question that has not been answered is therefore: „Has *Kalanchoë verticillata* a susceptible period in its growth?" We have not found any indication of it, and as the growth is a continuous process it does not seem probable. It is clear that the susceptible period will have only any influence on the form of the frequency curve if the external conditions are not constant. For two plants generally are not at the same time in a susceptible period, and if the external conditions are not constant, those plants will be influenced during their susceptible period by different external conditions. And then their constitution will be different during the remainder of their life. To eliminate this possibility we have tried twice to grow *Kalanchoë verticillata* under constant external conditions. The first time we raised the plants in watercultures in a constant temperature room in the Botanical Laboratory at Leyden. But all plants died in consequence of an fungous infection. The second time we raised the plants on sand cultures, also in the constant temperature room. The plants were constantly illuminated by six „window case" lamps, for if the susceptible period should happen to be very short or intermittent with a period of 24 hours, the difference between light and dark would have influenced the constitution of the plants. Moreover if intermittent light was used it would have been necessary to use very strong lamps, to give the plants a quantity of light equivalent to the daylight. But such strong lamps give so much heat that it is not possible to keep the temperature of the room constant.

This constant illumination however, has caused a luxuriant growth of Cyanophyceae, which so hardened the soilsurface that the plants died. So this second experiment under constant external conditions did also fail.

Summarizing our discussion of the material and methods we may conclude that:

Kalanchoë verticillata has some great advantages for this kind of investigations. It is very easy to get any number of sprouts of equal heredity. The plants grow well on sand if occasionally a v. d. Crone solution is given instead of water. It is easy to measure the plants as they do not branch as long as the terminal-bud is uninjured. And finally this species does not make high demands as to temperature (which in our case varied between 4° and 25° C.) and humidity.

Kalanchoë also has some drawbacks. It is very difficult to obtain sprouts of equal constitution. Moreover they are not easy to grow on water cultures, and only if plants are grown on water culture the external conditions (at least as far as the roots are concerned) are practically equal with certainly.

III. A BRIEF SURVEY OF THE RESULTS.

1. The results of the measurements of the foliola of *Robinia Pseudacacia.*

How far do the leaves of a plant satisfy the conditions A? If the plant is not a periclinal chimera all leaves should have, of course, equal heredity (bud variations are very rare). So the first condition is satisfied.

The second condition was that the sprouts or the seeds must have had an equal constitution when planted, that is to say when they began to grow independently. But does an organ of a plant grow independently? Most organs have a certain autonomy, but still their connection with the rest of the plant is of primordial importance. And when does a leaf for instance begin to grow 'independently'? Perhaps one may choose the moment on which the cells are formed which will only develop into a leaf as the moment on which the leaf begins its 'independent' existence. But in that case there are no reasons to suppose that those newly-formed cells have not always practically the same constitution. So the second condition is also satisfied for the leaves of one plant.

It is not so certain that the third condition, equality of the external circumstances has been satisfied. For, in general, the leaves formed in spring will have been under other external conditions during their development as the leaves formed in the summer. But at the other hand the climate in Holland is so variable that there are periods in the summer which are less favourable for the development of leaves than some periods in spring. It is therefore not possible to say 'a priori' that a well defined leaf (e.g. the 4th leaf of the third branch) has grown under more favourable circumstances than any other leaf. But as this was our condition of practical equal conditions we may conclude, if with a slight hesitation, that the third condition was also satisfied in this case.

268 leaves, with 3929 foliola, were measured. We have first put the results of our measurements in a correlation table[1]). (length of the foliola against breadth). As the terminal leaflet occupies a position apart we have not classed them with the other leaflets in the same table. The two tables (of the paired and of the unpaired leaflets) differed indeed. If the length of a paired leaflet was equal to the length of an terminal leaflet its breadth was generally smaller.

The correlation between length and breadth was in both cases high, but as the correlation was certainly not linear we have not computed the correlation coëfficient. In general the correlation coëfficient does not seem to have a great importance for biological research, except if its absolute value is very high (above 0,9) or very low (under 0,1). For the whole theory of correlations is based on the same suppositions as the theory of the normal frequency curve, and has the same theoretical value as this theory. Moreover it is impossible to gather from the correlation coëfficient if there are any systematical deviations. One has a whole set of valuable observations and one exchanges it for a number of doubtful value.

The frequency curve of the length of the paired leaflets was

---

[1]) The records of our measurements are kept in the Botanical Laboratory. They are available for all investigators who are interested in them. (Please apply to Botanical Laboratory Government University, Leyden, Holland.)

asymmetrical. The frequency curve of the breadth of the leaflets was also slightly asymmetric and moreover leptokurtic (fig. 2).

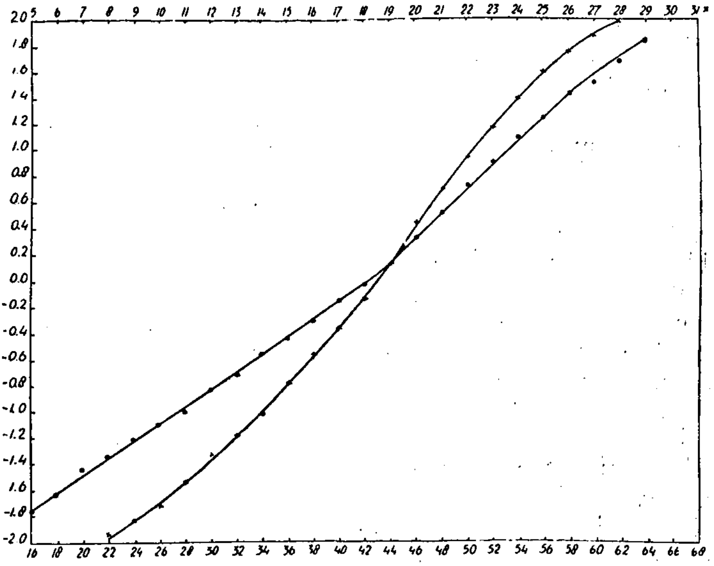

Fig. 2. z-function for length of 3013 (paired) foliola of *Robinia pseudacacia* L. Abscissa length in mm. (lower side). (....)
z-function for the breadth of 3029 (paired) foliola. Abscissa breadth in mm. (upper side). (x x x) Ordinate z.

We have also traced the frequency curve of the difference in length and of the difference in breadth of two opposite leaflets. If anywhere then here it may be expected that a Gaussian curve should appear. For the difference in length and in breadth of two opposite leaflets is very small. In more than half the number of cases the differences in length were less than 1.5 mm. and in more than three fourth of the cases the differences in breadth were less then 1.5 mm. As the differences in the total length were very small it may be expected that the differences in growth rate were also very small, the more so as the external conditions of two „partners” cannot have differed much. The difference in growth rate of the two leaflets would have been almost as often positive as negative if the constitution was always equal, for the external conditions will hardly ever favour always the same leaflet. In that case (equality of the constitution) the final difference in length would be the sum of a great number of independent quantities which were as probable positive as negative. The frequency curve of the differences in length of two opposite leaves would thus have been a Gaussian curve. But as is shown in fig. 3 the frequency curve

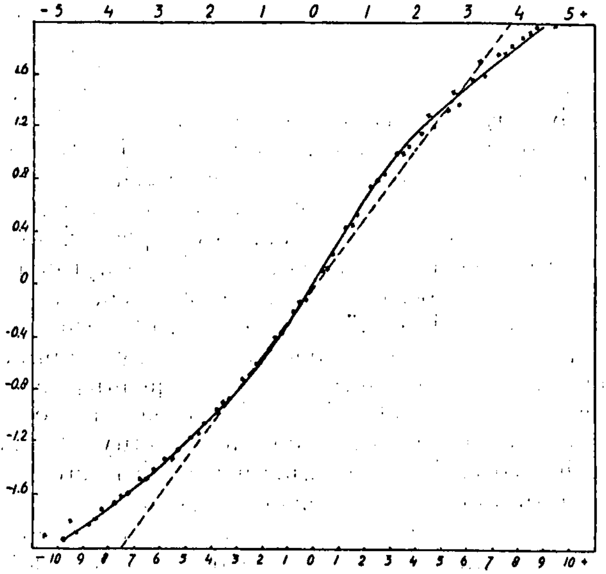

Fig. 3. z-function for the differences in length .... and for the differences in breadth x x x of two opposite foliola of *Robinia*. The line is the z-function of the Gaussian curve with the same mean and σ as the curve of the differences in breadth. Abscissa difference in length in mm (lower side) and of the differences in breadth (upper side) in mm. Ordinate z.

is clearly leptokurtic. Therefore the constitution has evidently influenced the growth rate, and the suppositions on which are based the derivation that the frequency curve must be a Gaussian curve were not satisfied.

The frequency curves of the differences in length and of the differences in breadth are equal if the unities of length are suitably chosen. This means probably that there is an internal factor which determines the form of the leaflet and that if the constitution is unfavourable for the growth in length it is generally as unfavourable for the growth in breadth.

The general conclusions on the measurements of the leaves of a *Robinia Pseudacacia* shrub are therefore that the frequency curve is in no case a Gaussian curve and that there are strong indications that this is a consequence of the influence of the constitution on the growth rate of the foliola.

§ 2. The results of the experiments with Kalanchoë verticillata.

On October 24, 3 weeks after the sprouts were planted the breadth of the largest „cotyledon" was measured. We call „cotyledons" the two leaves which were formed while the sprout was still attached to the mother plant. The breadth of this largest „cotyledon" is a measure for the quantity of reserve food the sprout has stored. But as the sprouts were planted three weeks before the measurements, the breadth of the „cotyledons" is also a measure for the velocity with which the roots were formed. For it is clear that all parts of a sprout will shrink as long as the quantity of water and foot it loses by respiration and transpiration is not compensated by the quantity of water it absorbs from the soil. And only if the sprout has formed roots it will be able to absorb any important quantity of water from the soil. If the „cotyledons" are wide it may therefore be expected that the constitution of the plant is good at that moment. For probably it has formd its roots early, and the quantity of reserve food it contained when planted was not small. In our experiments it was visible during the whole life of the plants that plants whose „cotyledons" were large were generally longer than plants whose „cotyledons" were small.

The frequency curve of the breadth of the largest „cotyledon" was asymmetric to the right and, moreover bradykurtic (fig. 4). As the „cotyledons" have grown during two different periods (on the mother plant and on the soil) which periods were separated by a period of shrinking (on the soil before the roots were formed) we have here a very complicated process. Therefore those who mean that in general the frequency curve must be Gaussian will perhaps not accept this frequency curve as a refutation of their theories.

On the tenth of December the length of both leaves of the third pair was measured. Those are the first leaves which are situated just above the „cotyledons". (*Kalanchoë verticillata* is decussate). The mean leaf length of the leaves above the largest „cotyledon" (which will be called D_1) was 0.31 mm. less than the mean leaf length of the other leaves (which will be called D_2). As the frequency curve of the length of D_2 is situated (practically) entirely to the right of the frequency curve

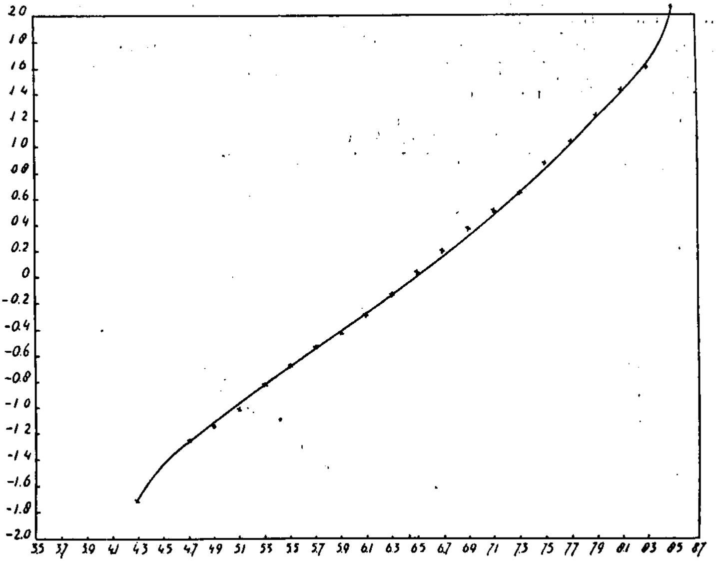

Fig. 4. z-function of the width of the largest 'cotyledon'. Abscissa width in mm. Ordinate z.

of D_1 this difference is real (fig. 6). Moreover in only 101 cases out of 531 the difference D_2—D_1 was negative. The frequency curve of D_2—D_1 was very curious (fig. 5).

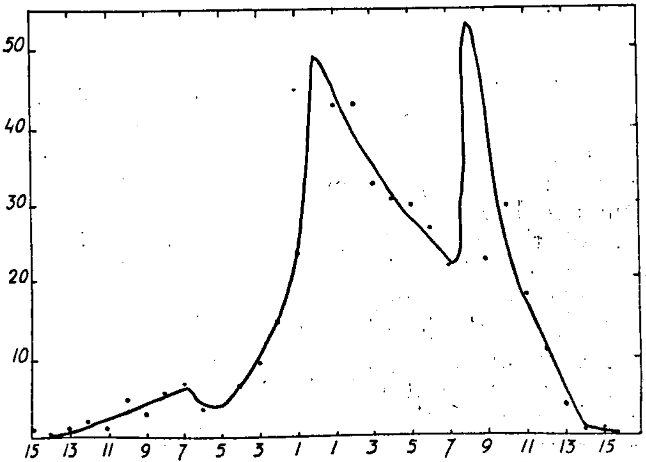

Fig. 5. Frequency curve of D_2—D_1. Abscissa D_2—D_1 in mm. Ordinate number of plants.

It had two maxima, one at 0 mm. and one at 0.8 mm. We have not found any explanation of this phenomenon. But it illustrates very clearly that a frequency curve of homogeneous material may be bimodal.

Neither the frequency curve of D_1 nor of D_2 was normal. They were both skew and bradykurtic (fig. 6).

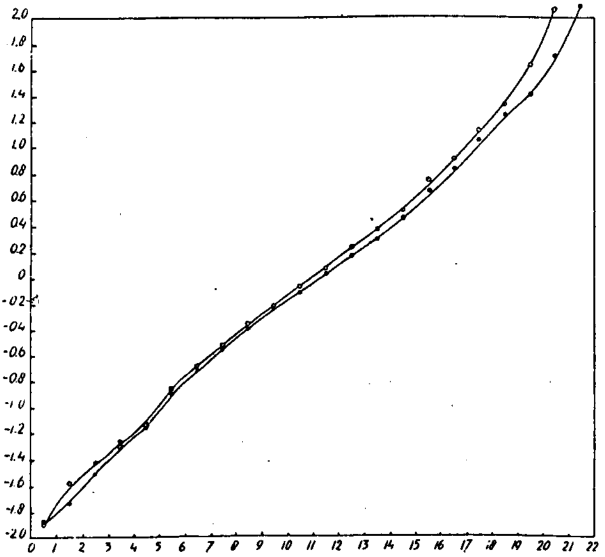

Fig. 6. z-function of D_1 o—o—o. z-function of D_2 —·—· Abscissa length of leaves in mm. Ordinate z.

It is clear that those three frequency curves (of D_1, D_2 and D_2—D_1) disprove absolutely the theory that the normal frequency curve must be Gaussian.

Nine times we have measured, in the course of our experiments, the length of the plants. This happened to yield a sufficient material to prove the validity of our objections against the theories of Quetelet and of Kapteyn. The form of the frequency curve of the length of the plants was never Gaussian. The first measurement B gave an irregular frequency curve, all others however were distinctly bradykurtic. This should be a sufficient refutation of the theory of Quetelet (fig. 7).

In our discussion on the frequency curve of the breadth of the largest „cotyledon" we have made plausible that plants with a different breadth of the „cotyledon" have a different constitu-

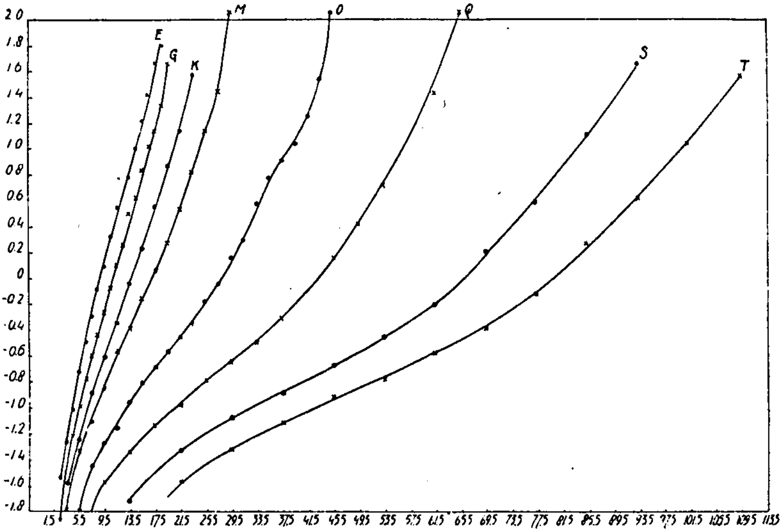

Fig. 7. z-functions for the length of *Kalanchoë* (all plants) after different periods.
E. measurements of Jan. 11. O. measurements of April 30.
G. measurements of Febr. 6. Q. measurements of May 15.
K. measurements of March 11. S. measurements of June 11.
M. measurements of March 27. T. measurements of July 4.
Abscissa length in mm. Ordinate z.

tion. If one chooses plants with a nearly equal breadth of the „cotyledon" (e.g. between 6.65. — 7.10 mm.) and therefore with an equal constitution on October 24, which will be in that case the form of the frequency curve of the length of those plants? From fig. 8 it is clear, that even in this case the frequency curve is not Gaussian.

We may conclude: The frequency curve of the length of a number of plants of *Kalanchoë verticillata,* which were (1) all of equal heredity, (2) of an equal constitution when planted, and (3) were grown under practically equal external conditions, is not Gaussian. Therefore the theory of Quetelet c.s. is unacceptable in this case. And as we have found no examples, where a frequency curve is Gaussian, if all individuals satisfy the conditions A, the idea has to be rejected on other grounds as well.

A second method to disprove the theories of Quetelet c.s. and of Kapteyn is to demonstrate that the suppositions on

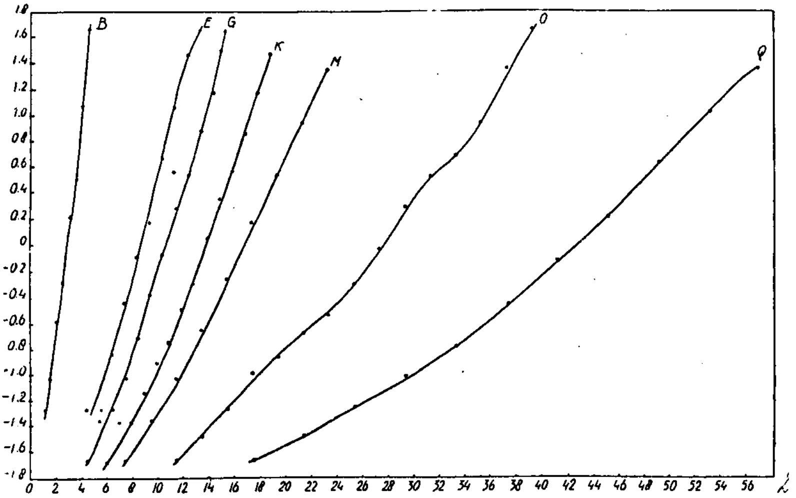

Fig. 8. z-functions of length of plants which width of the largest 'cotyledon' on October 24 was from 6.20 mm. to 6.65 mm. Abscissa length in mm. Ordinate z. B measurements of 16 Nov., for E—Q vide fig. 7. (107 Individuals).

which they are based are not realized in nature. The theory of the adherents of Quetelet is based upon the supposition that the increase in length during a certain period does not depend upon the increase in length during any previous period. Therefore it must be independent of the length of the plants. That is to say the mean growth rate during a certain period of plants, which had a different length at the beginning of the period, must be equal. Moreover, this mean growth rate must be the same for different periods.

In the theory of Kapteyn it is supposed that the mean growth rate during a certain period, is a function of the length of the plants. Further it is supposed, that this mean growth rate of plants is equal in different periods. A corollary of the first supposition of Kapteyn is, that the variability of the growth rate of plants of the same length is the same function of the length as the mean growth rate. In fig. 9 and fig. 10 we have plotted, for the different periods between our measurements, the mean growth rate in μ per diem against the length of the plants. Two things are evident from this figure:

1. The mean growth rate is not constant. Therefore the theory of the adherents of Quetelet does not seem to apply.

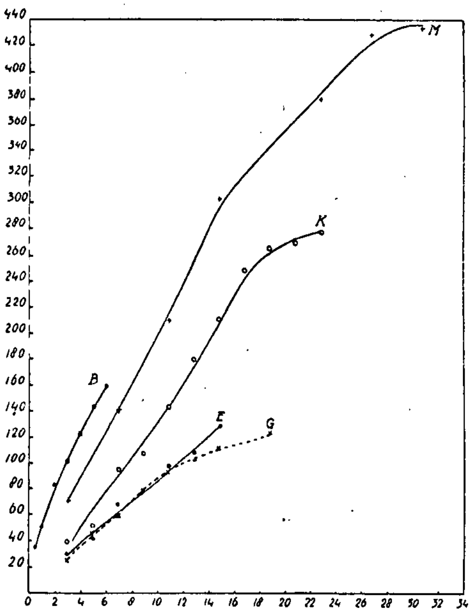

Fig. 9. Mean growth rate as a function of the length for different periods. Abscissa length of the plants. Ordinate mean growth rate in μ per diem.
B period between Nov. 16 and Jan. 11.
E „ „ Jan. 11 and Feb. 6.
G „ „ Feb. 6 and March. 11.
K „ „ March 11 and March 27.
M „ „ March 27 and April 30.

2. The mean growth rate is not only a function of the length of the plants, but it depends also upon their age. Therefore also the theory of Kapteyn does not hold here.

Moreover the theory of Kapteyn does not suffice three other sets of data:

1. According to Kapteyn the variability in the growth rate of plants of equal length is proportional to the mean growth rate. From table 1 (p. 119) (which gives the mean growth rate during the period G of plants of different length and the variability of the growth rate) it is evident, that variability and growth rate are not proportional. The variability is practically constant, the growth rate increases with the length of the plants. Our other measurements confirm this result. This should be a sufficient proof that the theory of Kapteyn is inapplicable.

2. Kapteyn has given a method to determine, from the form of the frequency curve, the mean growth rate of plants of a certain length. This mean growth rate is proportional to the cotangent of the line-tangent to the z-curve. If one has plotted the z-curve of the same material for two different moments, the calculated mean growth rate of plants of a certain length must be the same for both curves. (The mean growth rate does not depend on the time, according to Kapteyn.). Therefore the slope of both z-curves should be the same for an equal abscissa. From fig. 7 and fig. 8 it is clear that this is not true.

3. According to Kapteyn the constitution of a plant is unambiguously determined by its length. The probability that

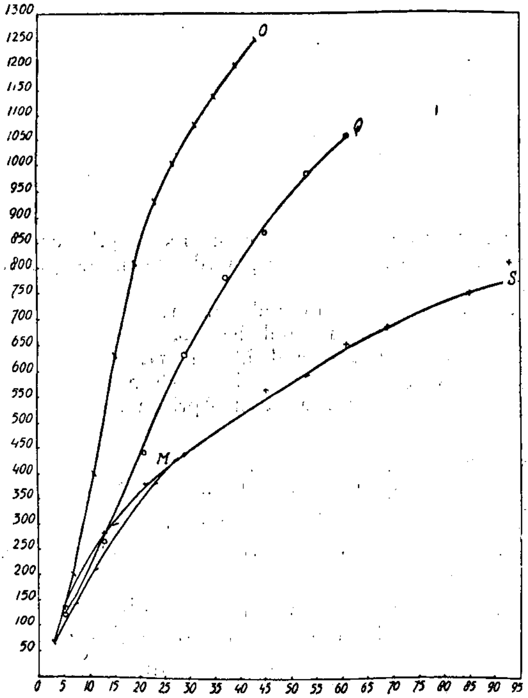

Fig. 10. Mean growthrate as a function of the length. Abscissa length of the plants. Ordinate mean growth rate in μ per diem.
M period between March 27 and April 30.
O „ „ April 30 and May 15.
Q „ „ May 15 and June 11.
S „ „ June 11 and July 4.

Table I.

Length of plants on Febr. 6.	Mean growth rate Febr. 6.-March 11. in μ per diem.	Standard Deviation of mean growth rate	Number of plants.
2— 4	25.4		17
4— 6	44.9		39
6— 8	59.8	13	69
8—10	78.2	21	88
10—12	93.2	19	103
12—14	100.4	21	104
14—16	111.4	20	61
16—18	111.8	19	41
18—20	122.7		20
20—22	121.2		2

the growth rate of a plant is less than the growth rate of other plants of the same length is therefore always ½. If the plants are classified according to their length after each measurement, the probability that the difference in rank of the same plant at two successive measurements is positive (or negative) is therefore ½. If there are 8 measurements, the probability that there will be a positive (and (7-a.) negative) changes in rank, will be:

$$\frac{7!}{a!\ (7-a)!} \cdot \frac{1}{2^7}$$

In table 2 is given the number of plants with *(a)* negative changes in rank as we have found in our experiments, and (in the third column) the theoretical number of negative changes

Table II.

Number of negative differences in rank. *a*	Number of plants, as found.	Number of plants, as calculated.
0	9	3.5
1	43	24.5
2	84	73.5
3	90	122.5
4	89	122.5
5	88	73.5
6	36	24.5
7	10	3.5

in rank. It is clear that there is a systematic difference between the theoretical and the experimental numbers. 3 and 4 negative deviations are too rare (¾ the theoretical number) all other occur too frequently. The explanation of this difference is probably that if, during a certain period a plant diminished in rank, the probability that during the following period it will again diminish in rank is more than ½. This means that the constitution of a plant is not unambiguously determined by its length. Therefore this result proves that in our case at least the theory of Kapteyn is inapplicable.

3. A third method to show the inapplicability of Quetelet's theory is the following:

We have classified the plants according to the breadth of their largest „cotyledon" on October 24. The frequency curves of the length of the plants on July 4 plotted for those different classes, showed to be distinctly unequal (fig. 11). This proves that the breadth of the „cotyledon" may be taken as a measure of the

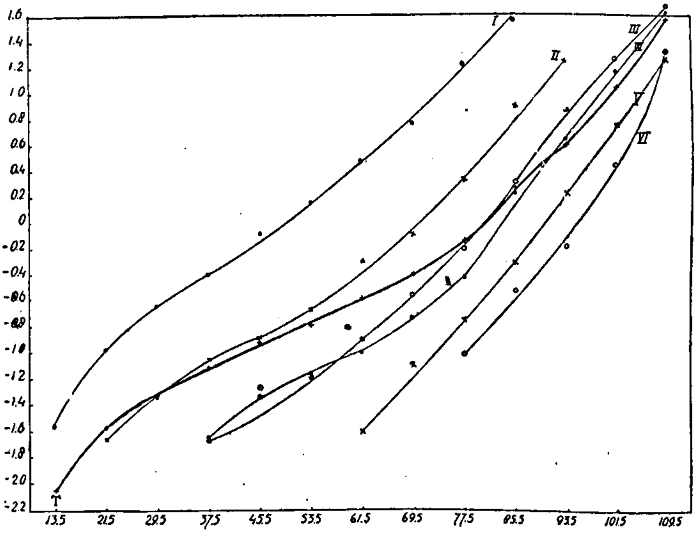

Fig. 11. z-function for the length of the plants on July 4, classified according to the width of the largest 'cotyledon on October 24.
T. All plants (536 plants).
I. Width of 'cotyledon' < 5.40 mm. (73 plants).
II. „ „ „ 5.40 mm. — 6.20 mm. (105 plants).
III. „ „ „ 6.20 m.m. — 6.65 mm. (108 plants).
IV. „ „ „ 6.65 mm. — 7.10 mm. (102 plants).
V. „ „ „ 7.10 mm. — 7.60 mm. (83 plants).
VI. „ „ „ > 7.60 m.m. (65 plants).

constitution of the plants on October 24, and that, if the difference in constitution was very large on that date, this difference was not annihilated after nine months.

From this result it is clear that the theory of Quetelet which does not take into account the constitution of the organism, is unacceptable.

The results of our countings of the number of leaves corroborate the results of our measurements. As is clear from fig. 12 a Gaussian curve was not obtained.

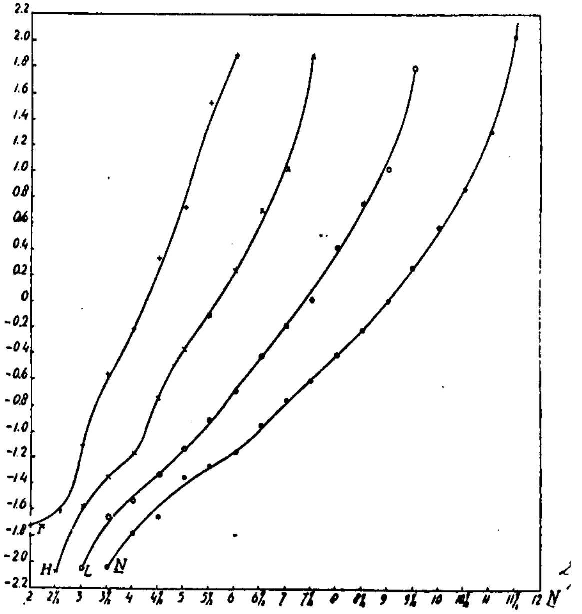

Fig. 12. z-function for number of leaves. Abscissa number of leaves. Ordinate z. For F, H, L, N vide p. 106.

SUMMARY AND CONCLUSIONS.

I. *Theoretical Introduction.*

1. In § 1—8 of the introduction we have determinated which were the necessary conditions (p. 89 conditions A.) which must be satisfied by a group of plants that the frequency curve of any characteristic of those plants may have any practical value. If those conditions are satisfied it is possible to obtain the same frequency curve if the experiment is repeated. To simplify the dis-

cussion we have only considered the frequency curve of the length of the plants.

2. In § 10—13 we have discussed the theory that the frequency curve should be identical with the error curve of Gauss. This theory was partly based upon experimental facts (e.g. by Quetelet and Borel). Later investigations have shown, however, that if the conditions A are satisfied the frequency curve is rarely, if ever, Gaussian.

Other investigators suppose that the frequency distribution originates in the same manner as the error curve does. They divide the whole period of growth in a great number of elementary periods. If it is supposed that the increase in length of a certain plant during an elementary period is as likely more as less than the mean increase in length of all plants during that period, the frequency curve of the total length of the plants will be Gaussian. This supposition is equivalent with the hypothesis that the growth rate of a plant is only determined by the external factors, and not by the previous history of the plants. But this is certainly not true. Therefore this proof of the theory that the form of the frequency curve is identical with the form of the error curve is unacceptable.

3. In § 14 the theory of Kapteyn is discussed. Though a great improvement on the theory of Quetelet, the theory of Kapteyn was nevertheless found not to apply. The fundamental supposition that *the mean growth of plants of the same length but of different age is equal* had to be rejected.

4. In § 13 we discussed the drawbacks of Pearson's method of curve fitting.

5. In § 15 it was demonstrated that the method of Fisher and 'Student' is generally not applicable to biological problems. For to apply this method it is necessary to know 'a priori' the frequency distribution of the 'Universe'.

6. In § 16 we have discussed a practical method to compare frequency curves (the method of Kapteyn).

II. *The experiments.*

1. The length and breadth of the foliola of all fullgrown leaves of a shrub of *Robinia Pseudacacia* L. were measured. Neither the frequency curves of the length and of the breadth of the foliola nor the frequency curves of the differences in length and in breadth of two opposite foliola was Gaussian. As the foliola satisfy probably the conditions A this proves that the theory of Quetelet does not hold.

2. 552 plants of *Kalanchoë verticillata* S. Elliott orginating from the same P_2 were grown under practically equal external conditions. The length of the plants was measured after nine consecutive periods. The mean growth rate during the eight subsequent periods depended both on the length as well as on the age of the plants, so both the theory of Quetelet and of Kapteyn did not apply to this case.

The frequency curves of the length were never Gaussian.

3. Three weeks after the sprouts were planted the breadth of their largest 'cotyledon' was measured. This breadth happened to be a good measure of the constitution (the internal condition as far as determined. by age and previous history) of the plants at that moment. For if the plants were classified according to the breadth the frequency curves of the length was always different for the different classes.

We may conclude from our theoretical and experimental investigation that it is impossible to determine 'a priori' the form of the frequency curve of any characteristic of a group of plants.

The foregoing investigations were carried out in the Botanical Laboratory of the Government University of Leyden. I am glad to be able to express here my sincere thanks to Professor Dr. L. G. M. Baas Becking for his constant help, suggestions and for his criticisms. I also want to thank Dr. C. D. Verrijp for his valuable advice and his help with the designing of the experiments, and Dr. J. B. D. Derksen for his advice in the mathematical part.

APPEND.

Derivation of formulae

Van Uven supposes in his derivation of the formula of Kapteyn that the increase in length of a plant of length x during an elementary period is $m\beta\psi(x)$. m indicates here the influence of the external conditions during this elementary period. $\psi(x)$ is the mean growth of plants of the length x. β is a constant of such a magnitude that both m and $\psi(x)$ may assume finite values.

The magnitude of m during an elementary period is assumed to be independent of its magnitude during any former period. As we have said this will hardly ever happen in nature. We have assumed therefore (p. 99) that m may assume the values 0 to 2k. If m has, during a certain period, the magnitude a the probability that during the succeeding period its magnitude will be a—1 is $\frac{a}{2k}$ and the probability that it will be a+1 is $\frac{2k-a}{2k}$.

Firstly it will be demonstrated that if, during a certain period, the frequency distribution is (symmetrically) binomial it will always remain (symmetrically) binomial.

If the frequency distribution is binomial the probability that during a certain period *m* has the magnitude a+1 is $\frac{1}{2^{2k}}\binom{2k}{a+1}$ that *m* has the magnitude a is $\frac{1}{2^{2k}}\binom{2k}{a}$, and that m has the magnitude a—1 is $\frac{1}{2^{2k}}\binom{2k}{a-1}$. The probability that during the following period *m* has the magnitude *a* is:

$$\frac{1}{2^{2k}}\left\{\frac{a+1}{2k}\binom{2k}{a+1}+\frac{2k-(a-1)}{2k}\binom{2k}{a-1}\right\}$$

$$=\frac{1}{2^{2k}}\left\{\frac{a+1}{2k}\binom{2k}{a+1}+\frac{2k-(a-1)}{2k}\binom{2k}{2k-(a-1)}\right\}$$

$$=\frac{1}{2^{2k}}\left\{\binom{2k-1}{a}+\binom{2k-1}{2k-a}\right\}$$

$$=\frac{1}{2^{2k}}\left\{\binom{2k-1}{a}+\binom{2k-1}{a-1}\right\}$$

$$=\frac{1}{2^{2k}}\binom{2k}{a}$$

Therefore the probability that *m* has (during this period) the magnitude a is equal to the probability that *m* had the magnitude a during the previous period.

In his derivation of the formula of Kapteyn Van Uven has to calculate the sum of the quantities *m* over the whole period of growth. As he supposes that the quantities *m* are (during each elementary period) distributed according to the normal law, and that moreover the magnitude of *m* during an elementary period is independent of its magnitude during a previous period the sum of *m* (summed over the whole period of growth) will be distributed according the normal law as well. It will now be demonstrated that if the quantities *m* are dependent of one another as stated above their sum is still distributed according to the normal law. To do this we shall make use of a theorem of S. Bernstein (Bernstein 1927 p. 24):

'Soit $S_n = x_1 + x_2 + \dots + x_n$ une somme de quantités dépen-
'dantes entre elles jouissant des propriétés suivantes:

'1. $M\ (S_n^2) = B_n > Mn^{\lambda}$, où $\lambda > \frac{2}{3}$

'2. Quelles que soient les valeurs déjà connues de certaines des
'x_k on peut fixer un nombre L, tel que, pour $i > k$, l'espérance
'mathématique de $|x_i^3|$ reste inférieure à L.

'3. Dans les mêmes conditions, on peut fixer un nombre N
'tel que l'on ait: $M\ (x_{i+1} + x_{i+2} + \dots + x_{i+g})^2 < N\ B_n\ \left(\frac{g}{n}\right)^{\lambda}$
'quel que soit g.

'4. Dans les mêmes conditions, pour $i—k > n^{\rho}$, où $\rho > \frac{\lambda}{2}$ est un
'nombre fixe, la variation de l'espérance mathématique de x_i
'ne dépasse pas $\frac{1}{n^{\mu}}$ où $\mu > 1 - \frac{\lambda}{2}$, et lorsqu'on a de plus $j—k > n^{\rho}$
'la variation de l'espérance mathématique du produit $x_i\ x_j$
'ne dépasse pas $\frac{1}{n^{2-\lambda}}$.

'Dans ces conditions la probabilité de l'inégalité

$$z_0 \sqrt{2B_n} < S_n < z_1 \sqrt{2B_n}$$

'a pour limite

$$\frac{1}{\sqrt{\pi}} \int_{z}^{z_1} e^{-z^2} dz$$

M (x) means the mathematical expectation of x. As the quantities x are supposed to have 'a priori' the probable value o we will consider the summation of m—k which has also the probable value o 'a priori'. We will prove that the quantities m—k satisfy the four conditions of Bernstein.

1. $M\ (S_n^2) = M\ (x_1^2 + x_2^2 + \dots + 2\ x_1\ x_2 + \dots)$

$$= M\ (x_1^2) + M\ (x_2^2) + \dots + 2\ M\ (x_1\ x_2) + \dots$$

The expectation of x_i^2 is always a positive quantity. The expectation of $x_i\ x_j$ is certainly not negative for the probability that x_i and x_j have the same sign is greater than the probability that they have a different sign (vide 4). Therefore $M\ (S_n^2)$ is the sum of n positive quantities and ½ n (n—1) quantities which are certainly not negative. Therefore

$M\ (S_n^2) > Mn^{\lambda}$ where $\lambda = 1$ and
where M is the smallest of the quantities $M\ (x_i^2)$.

The quantities $x = m - k$ satisfy, therefore, the first condition of Bernstein.

2. The expectation of $|x^3|$ is always finite as x_i can only assume finite values.

3. $$\frac{M\,(x_{i+1} + x_{i+2} + \ldots + x_{i+g})^2}{M\,(x_1 + x_2 + \ldots\ldots + x_n)^2} < N\left(\frac{g}{n}\right)^\lambda$$

If g is finite it is of course always possible to chose N so that this inegality is satisfied. If g tends to ∞ one has, in the numerator and in the denominator, functions of the same form. It will be shown in 4. that the expectation of x_{i+c} tends to $(a-k)\left(\frac{k-1}{k}\right)^c$ whatever the value of x_i is, when c tends to ∞. Therefore the expectation of $x_{i+k}\ x_j$ tends to zero if $j-i$ tends to ∞ and k remains finite. If k tends also to ∞ this product tends to the same value whatever the value of x is. The expectation of all except a finite number of products $x_i\ x_j$ tends therefore to the same value in numerator and denominator. The expectation of x_j^2 of the numerator is less than k^2. Therefore it is possible to choose for N such a value that condition 3 is satisfied.

4. To calculate the mathematical expectation of x_i if the value of x_k is known, it is easiest to replace our model by another, equivalent, model.

Suppose there are *2k* numbered balls, part of which are in a vase. The number of balls in the vase is equal to the magnitude of *m* during a certain period. There is moreover a second vase with *2k* tickets numbered from *1* to *2k*. To determine the magnitude of *m* at a certain moment if at the previous moment it was *m* a ticket is drawn. After the number is read the ticket is replaced in its vase. If the ball with the same number was in the vase it is now taken out. The number of balls in the vase, and therefore the magnitude of m_{i+1}, is in this case $m_i - 1$. The probability that this happens is $\frac{m_i}{2k}$. If the ball with the same number as the ticket was outside the vase it is now placed in it. The number of balls and therefore the magnitude of m_{i+1} is in this case $m_i + 1$. The probability that the above happens is $\frac{2k - m_i}{2k}$. The probability that at a certain moment there are *a* balls in the vase is therefore equal to the probability that *m* has the magnitude *a* on that moment.

The probability that a certain ball which was in the vase will be again in the vase after 2m tickets are drawn is the probability that its ticket is drawn an even number of times:

$$\frac{1}{(2k)^{2m}} + \frac{2m(m-1)}{1.2}\left(\frac{1}{2k}\right)^{2m-2}\left(\frac{2k-1}{2k}\right)^{2} + \ldots\ldots$$

The probability that a ball which was in the vase will not be in the vase is the sum of the probabilities that its ticket is drawn an odd number of times:

$$\frac{2m}{1}\left(\frac{1}{2\,k}\right)^{2m-1}\left(\frac{2k-1}{2k}\right) + \frac{2m\ (2m-1)\ (2m-2)}{1.2.3.}\left(\frac{1}{2\,k}\right)^{2m-3}\left(\frac{2k-1}{2k}\right)^{3} + \ldots\ldots$$

The differense of the probabilities that the ball is in the vase and that it is not in the vase is:

$$\left(\frac{1}{2k}\right)^{2m} - \frac{2m}{1}\left(\frac{1}{2k}\right)^{2m-1}\left(\frac{2k-1}{2k}\right) + \ldots$$

$$= \left(\frac{1}{2k} - \frac{2k-1}{2k}\right)^{2m} = \left(\frac{2k-2}{2k}\right)^{2m} = \left(1 - \frac{1}{k}\right)^{2m}$$

If m tends to ∞ the difference of those probabilities tends to zero. This means that, if a certain ball was originally in the vase, it is after a sufficient long time nearly as probable that it should be inside as outside.

If $2m + 1$ tickets are drawn the probability of this difference is

$$\left(1 - \frac{1}{k}\right)^{2m+1}$$

If a certain ball is originally not in the vase one may prove in the same manner that the probability that it is in the vase after a sufficiently long time is (practically) as great as the probability that it is not in the vase.

The mathematical expectation of the number of balls in the vase after *m* periods, if originally there were *a* balls in the vase is

$$a\left\{\frac{1}{2} + \frac{1}{2}\left(1 - \frac{1}{k}\right)^{m}\right\} + (2k-a)\left\{\frac{1}{2} - \frac{1}{2}\left(1 - \frac{1}{k}\right)^{m}\right\}$$

$$= \frac{1}{2}\ a + \frac{1}{2}\,a\left(1 - \frac{1}{k}\right)^{m} + k - k\left(1 - \frac{1}{k}\right)^{m} - \frac{1}{2}a + \frac{1}{2}\ a\left(1 - \frac{1}{k}\right)^{m}$$

$$= k + (a-k)\left(1 - \frac{1}{k}\right)^{m}$$

As the mathematical expectation of *m* is *k* the variation of the mathematical expectation is therefore $(a-k)\ \left(1 - \frac{1}{k}\right)^{m}$. Therefore if

$m > n^\rho$ this variation is certainly less than $\frac{1}{n^\mu}$ if n tends to ∞.

The variation of M $(x_i\ x_j)$, if the value of x_k is known, will be of the same order of magnitude, for the value of x depends (directly) only upon the value of x_i (if $j > i$). Therefore the condition 4 of Bernstein is also satisfied.

This means that the sum of the quantities m which depend on another in the manner stated, satisfies the Gaussian law also in this case, and that, therefore, the result of Van Uven may be retained.

LITERATURE.

(Papers marked by an asterisk are mentioned in the text.)

1. Afanasjewa, T. Ehrenfest. 1926. Over een misverstand betreffende de toepassing van de waarschijnlijkheidstheorie op de irreversibiliteit der natuurverschijnselen. Verslag van de gewone vergadering der afdeling natuurkunde Kon. Ak. v. Wetenschappen. Deel 34, pag. 598.

2. Arbon, J. G. 1844. Verhandeling over de Binomiaal Coëfficienten. Rotterdam. 113 pag.

3. Baart de la Faille, C. J. 1914. Statistische onderzoekingen bij Senecio vulgaris. Dissertatie Groningen.

*4. Baas Becking, L. G. M. and Leland S. Baker. 1926. Studies on Growth. Parts I—II. Stanford University Publication. University Series Biological Sciences. Volume IV. Number 2. pag. 61—132.

5. Baas Becking, L. G. M. and E. F. Drion. 1936. Acta Biotheoretica I. (in press).

6. Bauer, Erwin. 1922. Einführung in die experimentelle Vererbungslehre. Berlin. 5te u. 6te Auflage.

*7. Bernstein, Serge. 1927. Sur l'extension du théorème limite du calcul des probabilités aux sommes de quantités dépendantes. Mathematische Annale. Bnd. 97. pag. 1.

*8. Bertrand, J. 1907. Calcul des Probabilités. 2ième Edition (conforme à la 1ière) Gauthier-Villars. Paris.

9. Blaauw, A. H. en Mej. M. C. Versluys. 1926. De gevolgen van de temperatuurbehandeling in de zomer voor de Darwintulp. Verslag van de gewone vergadering der afdeeling natuurkunde Kon. Ak. v. Wetenschappen. Deel 34. pag. 755. pag. 919 (samen met Mej. G. Joustra). pag. 1067 (A. H. Blaauw en R. Mulder).

10. Blakeslee, A. F. Variability curve following Law of Chance. Journal of Heredity. Vol. VII. pag. 280.

11. Bodmer, Helene. 1927. Beiträge zum Heterostylie-Problem bei Lythrum Salicaria L. Flora 122. pag. 306—341.

*12. Borel, Emile. 1924. Eléments de la théorie des Probabilités. Paris. Librairie Scientifique J. Hermann. 4ième éd. 221 pag.

13. Bruyker, Dr. C. de. 1906. De gevoelige periode van den invloed der voeding op het aantal randbloemen van het eindhoofdje van Chrysanthemum carinatum. Handelingen van het Tiende Vlaamsche Natuur- en Geneeskundig Congres. Brugge. pag. 1—6.

14. Bruyker, Dr. C. de. 1906. Bemerkingen aangaande de Galtonsche curve. Handelingen van het Tiende Vlaamsche Natuur- en Geneeskundig Congres. Brugge. 30 September.

15. Bruyker, Dr. C. de. 1906. De polymorphe variatiecurve van het aantal bloemen bij Primula elatior Jacq.; hare beteekenis en hare beïnvloeding door uitwendige factoren. Tiende Vlaamsch Natuur- en Geneeskundig Congres. pag. 1—29.

16. Bruyker, Dr. C. de. 1907. Een nieuw geval van omkeering eener „halve Galtoncurve". Handelingen van het elfde Vlaamsch Natuur- en Geneeskundig Congres. Mechelen. pag. 74—82.

17. Bruyker, Dr. C. de. 1908. Over dubbele halve curven. Proefondervindelijke studie bij Calliopsis bicolor. Handelingen van het Twaalfde Vlaamsch Natuur- en Geneeskundig Congres. St. Niklaas. pag. 215—224.

18. Bruyker, Dr. C. de. 1908. De heterostylie bij Primula elatior Jacq. Statistische gegevens. Handelingen van het Twaalfde Natuur- en Geneeskundig Congres. St. Niklaas. pag. 242—248.

19. Bruyker, Dr. C. de. 1910. De statistische methode in de plantkunde en hare toepassing op de studie van den invloed der levensvoorwaarden. 226 pag. Gent. Koninklijke Vlaamsche Academie voor Taal- en Letterkunde. Uitgave van het Van de Ven-Heremans' Fonds. No. 6.

20. Buchanan-Wollaston, H. J. 1935. Statistical Tests. Nature 136. 3431. pag. 182.

21. Buchanan-Wollaston, H. J. 1935. Statistical Tests. Nature. vol. 136. pag. 722.

22. Cool, C. and A. N. Koopmans. Variation and correlation of the number of umbel rays of some Umbelliferae. Biometrika. Vol. XI. pag. 38.

23. Czuber, Emanuel. 1910. Wahrscheinlichkeitsrechnung und ihre Anwendung auf Fehlerausgleichung, Statistik und Lebensversicherung. 1ster Bnd. Wahrscheinlichkeitstheorie — Fehlerausgleichung — Kollectivmasslehre. 436 pag. 2ter Bnd. Mathematische Statistik — Mathematische Grundlagen der Lebensversicherung. 443 pag. Teubner. Leipzig und Berlin.

24. Derksen, J. B. D. 1935. Inleiding tot de correlatierekening. Diss. Leiden.

25. Elderton, W. Palin. 1906. Frequency and Correlation. London. C. and E. Layton.

26. Fisher, R. A. 1932. Statistical Methods for Research Workers. Oliver and Boyd. London, Edinburgh.

27. Fisher, R. A. 1935. Statistical Tests. Nature. 3438. Vol. 136. pag. 474.

28. Goebel, K. 1902. Ueger Regeneration im Pflanzenreich. Biologisch Centralblatt. XXII. pag. 385—505.

29. Günther, H. Die statistische Berechnung von Mittelwert und Streuung nach meinem Summenverfahren. Biol. Zentralbl. 54 (11/12), 588.

30. Hack, Prof. Dr. Franz. 1911. Wahrscheinlichkeitsrechnung. Leipzich. Sammlung Göschen 508.

31. Hamet, Raymond. 1907, 1908. Monographie du genre Kalanchoë. Bulletin de l'Herbier Boissier, 2me Série. Tome VII. No. 11. pag. 869—900. Tome VIII. No. 1. pag. 17—48.

32. Heyer, A. 1909. Ueber die Längenvariation der Coniferennadeln. Biometrika Vol. VI. pag. 354—365.

33. Irwin, J. O. 1927. On the frequency distribution of the means of samples from a population having any law of frequency with finite moments, with special reference to Pearson's Type II. Biometrika vol. XIX. pag. 225.

34. Hotelling, Harold. 1925. The distribution of correlation ratios calculated from random data. Proc. of the Nat. Ac. of Sciences. Vol. 11. No. 10. pag. 657—662. October 1925.

*35. Johannsen, W. 1903. Ueber Erblichkeit in Populationen und Reine Linien. G. Fischer — Jena.

36. Johannsen, W. 1926. Exakte Erblichkeitslehre mit biologischer Variationsstatistik. 3ter Aufl.

37. Kapteyn, J. C. 1903. Skew frequency curves in Biology and Statistics. Published by the Astronomical Laboratory at Groningue. P. Noordhoff. Groningen.

38. Kapteyn, J. C. 1906. Reply to Prof. Pearsons criticisms. Recueil des Travaux Botaniques Néerlandais. Vol. XI. pag. 216—222.

*39. Kapteyn, J. C. 1916. Skew frequency curves in Biology and Statistics. Rec. d. Trav. Bot. Néer. Vol. XIII. pag. 105—154.

*40. Kapteyn, Dr. J. C. and Dr. M. J. van Uven. 1916. Skew frequency curves in Biology and Statistics. 2nd. Paper. Gebr. Hoitsema. Groningen.

41. Kelley, Truman L. 1923. Statistical Method. New York. The Macmillan Company.

42. Koriba, K. 1909. Ueber die individuelle Verschiedenheid in der Entwicklung einiger fortwachsende Pflanze mit besonderer Rücksicht auf die Aussenbedingungen. Journal of the College of Science. Imperial University of Tokyo. Vol. XXVII. Art. 3.

43. McEwen, G. F. 1921. Rapid methods of approximating to terms in a binomial expansion

44. Mac Leod, J. 1899. Over de correlatie tusschen het aantal meeldraden en het aantal stampers bij het speenkruid (Ficaria Ranunculoides) Botanisch Jaarboek, uitgegeven door Dodonaea. Elfde jaargang, pag. 91—107.

45. Mac Leod, J. 1900. Over de veranderlijkheid van het aantal randbloemen en het aantal schijfbloemen bij de Korenbloem (Centaurea Genus) en over correlatieverschijnselen. Botanisch Jaarboek, uitgegeven door Dodonaea. Twaalfde jaargang, pag. 40—70.

46. Michael, E. L. 1920. Concerning application of the probable error in cases of extremely asymetrical frequency curves. Science N.S. no. 1308, pag. 89—90 Jan. 1920.

47. Nemec, B. 1907 1 Maart, 1911 10 Maart, 1911 24 Nov. Weitere Untersuchungen über Regeneration. Bulletin international de l'Academie des Sciences de Bohême.

48. Ossenbeck, Carola. 1927. Kritische und experimentelle Untersuchungen an Bryophyllum. Flora. 122. pag. 342—387.

*49. Pearl, Raymond. 1907. Variation and Differentiation in Ceratophyllum. Carnegie Institution of Washington.

*50. Pearson, Karl. 1905. „Das Fehlergesetz und seine Verallgemeinerungen durch Fechner und Pearson." A rejoinder. Biometrika vol. 4, pag. 169.

51. Pearson, K. 1935. Statistical Tests. Nature, 136. pag. 296.

52. Pearson, K. 1935. Statistical Tests. Nature. vol. 136, pag. 550.

53. Poincaré, H. 1904. La Science et l'Hypothèse. Paris. Ernest Flammarion.

*54. Poincaré, H. 1909. Science et Méthode. Paris. E. Flammarion.

*55. Poincaré, H. 1912. Calcul des Probabilités. Rédaction de A. Quiquet. 2ième édition. Paris. Gauthier-Villars 333 blz.

56. Polya, G. 1930. Sur quelques points de la théorie des probabilités. Annales de l'Institut Henri Poincaré, vol. 1, pag. 117.

*57. Quetelet, A. 1846. Lettres sur la théorie des probabilités. Bruxelles.

58. Quetelet, A. 1871. Anthropométrie. Paris.

59. Rahusen, A. E. 1903. Over een uitbreiding van het theorema van Tchebycheff. Nieuw Archief voor Wiskunde. Tweede Reeks, Deel VI, eerste stuk pag. 56—62.

60. Reed, H. S. 1923. A note on the Statics of Cyclic Growth. Proceedings of the National Academy of Sciences. (Graduate School of Tropical Agriculture and Citrus Experiment Station University of California). vol. 9, no. 3, pag. 65—67.

61. Riebesell. 1932. Mathematische Statistik und Biometrik.

62. Schols. 1893. De wet van de fouten van waarneming. Verslagen der zittingen van de Kon. Academie van Wetenschappen. pag. 194.

63. Seitz, W. und K. Hamacher-Odenhausen. 1934. Untersuchungen über das Galtonbrett. Die Naturwissenschaften. 22. Jahrgang, pag. 494.

64. Spruit, Dr. C. 1922. Is de biometrische frequentiekromme een normale frequentiekromme? Voordracht gehouden in de biologische sectie van het 2de Ned. Indische Natuurwetenschappelijk Congres, 11—14 Mei 1922. Bandoeng. Handelingen van het 2de Ned. Ind. Natuurwetensch. Congres 1922, pag. 143.

65. Tammes, Tine. 1907. Der Flachsstengel. Ein Statistisch-Anatomische Monographie. Natuurkundige Verhandelingen van de Hollandsche Maatschappij der Wetenschappen. Derde verzameling, deel VI, vierde stuk, 285 pag.

66. Teissier, Georges. 1933. Les lois élementaires de la Croissance. Annales de la Société Royal des Sciences Médicales et Naturelles de Bruxelles. Année 1933, liv. 3—4.

67. Uven, M. J. van. 1926. Over het bewerken van scheeve correlatie. Verslag van de gewone vergadering der afdeeling natuurkunde. Kon. Academie van Wetenschappen. Deel 34, pag. 787, pag. 965, Deel 35, pag. 129.

68. Uven, Prof. M. J. van. 1927. Linear Adjustment of a Set of Pairs of Numbers. Kon. Academie van Wetenschappen. Proceedings vol. XXX, pag. 1021—1038.

69. Verrijp, Dr. C. D. 1932. Variatie-Statistiek en haar toepassingen in de oogheelkunde. Ned. Tijdschrift voor geneeskunde. Jaargang 76, pag. 2018—2031.

70. Vries, H. de. 1898. Over het omkeeren der halve Galtoncurven. Botanisch Jaarboek X, pag. 27—61.

71. Vries, H. de. 1899. Ueber Curvenselection bei Chrysanthemum segetum. Ber. d. D. Bot. Ges. XVII, pag. 86—98.

72. Vries, H. de. 1901. Die Mutationstheorie. Leipzig, Verlag von Veit & Co.

73. Went, F. A. F. C. 1930. Ueber wurzelbildende Substanzen bei Bryophyllum calycinum Salisb. Zeitschrift für Botanik 23, pag. 19—26.