Herring Investigations at Plymouth. I. Methods of Collection and Treatment of Data.

By

E. Ford, A.R.C.Sc.,

Naturalist at the Plymouth Laboratory.

With 5 Figures in the Text.

CONTENTS.

GENERAL INTRODUCTION ........................................ 238

THE USES AND LIMITATIONS OF THE STATISTICAL DATA COLLECTED .......................... 239

LENGTH OF FISH .................................................. 239

Selective Action of Fishing Nets ................................ 240

AGE AND GROWTH FROM SCALES .................................. 241

Estimations of Age from Scale-readings .......................... 241

The Part of the Body from which Scales are taken ............... 243

Methods of Recording Estimates of Age .......................... 243

Estimations of Growth from Scale Measurements ................. 246

Apparatus used in Study of Growth from Scale Measurements .... 246

Growth "Types" from Scale Measurements ........................ 246

Comparisons between Samples in different Areas ............... 250

Differences of Growth as Evidence of Migrations ............... 252

THE NUMBER OF VERTEBRÆ ........................................ 253

Preparation of Skeletons ........................................ 253

Use of Radiograms .............................................. 253

The Vertebral Count made .................................... 253

The Chances of drawing truly Representative Samples from a known Population ........ 254

The Arithmetic Mean and the Standard Deviation ............... 256

Errors due to Sampling ........................................ 258

Comparisons between Samples .................................. 259

Table of Notations and Formulae ................................ 261

BIOLOGICAL SIGNIFICANCE OF STATISTICAL RESULTS ............... 262

LITERATURE CITED .................................................. 265
GENERAL INTRODUCTION.

Towards the end of the year 1924, the present writer commenced a statistical study of random samples from commercial landings of herrings at Plymouth, in order to obtain information on the two following questions:

(a) Upon what year-class or year-classes of herring was the season's fishery of 1924–25 concentrated?
(b) Did the fishes landed throughout that season appear to conform to a single morphological type?

Throughout the work the size of the sample and the number of characters studied had to be such that all observations could be made by one investigator with the help of a single Laboratory Assistant. It was found by experience that if the sample did not greatly exceed one hundred fishes, the following routine procedure could be followed for each fish of two or three samples per week:

(a) Take a measurement of length.
(b) Clean and mount three scales on a numbered glass slide.
(c) Open up the body-cavity, note the sex and degree of maturity of the roe or milt.
(d) Attach a linen label to the head of the fish by means of a safety pin.
(e) When all the fishes had been treated as above, boil in lots of about fifty, and prepare clean skeletons to be dried for a subsequent determination of the number of vertebrae.

The results for the season showed quite definitely that herrings with five summer growth-zones and five winter-rings to the scale greatly predominated in the catches throughout the season November, 1924, to March, 1925. On the other hand, all observations seemed to suggest that the conception of a single "type" of herring for Plymouth was untenable. In the light of these results it appeared advisable to examine herrings from other areas, there being the possibility that, elsewhere, "types" would be recognised which would, in the event of their coming to Plymouth to spawn, produce a complex population similar to that observed in 1924–25. Accordingly, in addition to a series of samples from the Plymouth catches during each of the two following seasons, 1925–26 and 1926–27, samples from Milford Haven, Padstow, Newlyn, Mevagissey, Brixham, Brighton, and Lowestoft were studied.

This extensive sampling resulted in the accumulation of a formidable mass of statistical data from which the writer has been able to extract information outside the original scope of the research. The most convenient way of presenting this information is in the form of a series of
separate papers each of which deals with one particular section of the work done.

In the present part an endeavour has been made to provide the reader with a general account of the uses and limitations of the various statistical data collected for the study of the life-history of the herring. It is freely admitted that the larger portion of the subject-matter may make little or no appeal to the technical expert—it is not intended for him. It is hoped, however, that it may prove helpful to the general biological worker in assisting him to recognise the peculiar usefulness and importance of the statistical method of approach to the problems of herring biology. Probably in the case of no other economically valuable fish are we so dependent upon indirect evidence for our knowledge of its life and habits. If, then, the conclusions based on indirect evidence derived from statistical data are to carry conviction, it is essential that the general mathematical principles underlying the statistical treatment should be made clear. Since this report was prepared, the British Association Committee on Biological Measurements has issued a leaflet (British Association Report, Section D, Leeds, 1927), making recommendations for the taking and presentation of biological measurements. This leaflet should prove most helpful to all biologists whose work necessitates the compilation and interpretation of statistical data.

Acknowledgments for special services rendered will be made at times appropriate during the series of papers, but the writer is pleased here to record his thanks to Mr. F. J. Warren, Laboratory Assistant at Plymouth, who has acted as recorder throughout and rendered help generally in the collection of the data.

THE USES AND LIMITATIONS OF THE STATISTICAL DATA COLLECTED.

LENGTH OF FISH.

The standard measurement adopted was identical with character 9 of Orton (18, p. 73), viz. “from the tip of the snout to the distal end of the longest ray in the dorsal fluke of the caudal fin, when the fluke is placed so that its dorsal margin lies parallel to the line of measurement.” The actual measurement was made to the “nearest 1 mm. below”; thus, a fish whose length exceeded 25·5 cm. but was less than 25·6 cm. was recorded as of length 25·5 cm. When for statistical purposes it became necessary to group fishes into classes according to length, the grouping was made to the “nearest 1 cm. below”; thus, all fishes of length 25·0 cm. to 25·9 cm. both inclusive, were included in the 25 cm. group.

A convenient method of comparing samples with regard to the length
of the fishes is by means of the calculated values known as the median and the upper and lower quartiles. An illustration will serve to demonstrate the nature and significance of these statistical values. Imagine one hundred fishes arranged in serial order of length from the shortest in the first position to the longest at the end of the series. Note the lengths of the 25th and 26th fishes of the series; then the value lying midway between these two lengths is the lower quartile, \((Q_1)\). The value midway between the lengths of the 50th and the 51st fishes is the median, \((Q_2)\), while that between the lengths of the 75th and 76th is the upper quartile, \((Q_3)\). In other words, the values of \(Q_1, Q_2, Q_3\) merely indicate the lengths below which lie 25%, 50%, and 75% of the series.

An alternative method for comparison between samples is to calculate the values of the statistics known as the arithmetic mean and the standard deviation for each sample. Consideration of these two mathematical constants is given in the section of this paper which deals with the treatment of data on the number of vertebrae (page 256.)

If samples appear to differ with regard to the length of fish it is necessary to determine the extent to which the observed difference can be attributed to each of the following causes:

1. Errors, unavoidable or accidental, either in the actual determinations of length, or in the subsequent calculation of constants.
2. Errors of sampling due to the limitation of the size of the statistical sample.
3. Some form of selective action on the part of the apparatus used for the collection of the sample, which action has operated differently in the cases under comparison.
4. Real differences in population sampled.

The causes under heading 1 need not be enlarged upon, and those under headings 2 and 4 can be more conveniently dealt with at a later stage. Some observations on heading 3 are given in the sub-section which follows.

Selective Action of Fishing Nets.

It is important to remember that while random samples from commercial landings may give an excellent picture of the sizes of herring forming the mainstay of the local fisheries, they must be treated with the utmost caution as evidence concerning the actual herring stock on the grounds.

The outlook of the fisherman is entirely economic—he fishes for the maximum financial yield—so that in designing his nets he endeavours to obtain an apparatus which will catch the maximum quantity of fish of the best-paying sizes which can be caught in his local area. There
seems little reason to doubt that the fisherman himself is by far the best judge of the most profitable size of mesh required in his own district, even if we contend that past custom may have influenced him in his choice. Unlike the biologist, however, he has no reason to be disturbed by the fact that the mesh of his net allows smaller fishes to escape and may be too small to capture the largest, nor by the knowledge that the mesh of his net differs from that used in another area. When, therefore, the biologist wishes to compare the herring stock in one area with that in another, he must remember that the data for random samples from commercial catches are subject to correction for differences in mesh used in the two areas (vide Bjerkan, 1).

AGE AND GROWTH FROM SCALES.

As a rule, three scales from each fish, separately picked with forceps from the anterior region of the body near the tip of the pectoral fin, were cleaned between finger and thumb in water and mounted on a numbered glass slide, using a mixture of egg-albumen and glycerine as the adhesive.

Estimations of Age from Scale-readings.

The methods of scale-reading which have been followed so persistently and carefully by the Norwegians are based upon three important premises:

1. That for each winter passed through by a fish subsequent to the first formation of scales, a definite winter-ring is recorded on the scale.
2. That the scale-reader is able to distinguish between winter-rings and secondary-rings, the latter of which may not be in any way connected with age.
3. That the scale-reader has reliable information on the age of the fish when the first winter-ring was laid down.

The first of these three premises is obviously the most vital, and its proof has been derived from a study of the edge of the scale at different times of the year. The researches of Dahl (3), Lea (15), and later Hodgson (7) have shown that the herring scale grows by a definite annual cycle consisting of an initial phase of active growth (the physiological summer) followed by one of rest from growth (the physiological winter), and that the character of the edge of the scale changes accordingly. Thus all herrings caught during the summer have scales which exhibit a typical growing or "summer" edge, while those caught during the winter have scales which exhibit an equally typical resting or "winter" edge. The
commencement of the growing phase of every annual cycle is recorded on
the scale as a definite winter-ring.

The second premise, while it serves to impress the fact that some
practical experience in scale-reading and interpretation is necessary, is
intended rather to emphasize the truth that scale-readings are, to some
extent, personal estimates of age. In this connection the words of Johansen
(11, p. 23) may be usefully repeated, although, in the writer's opinion,
these are somewhat too severe. He says:—

"... Erroneous determinations are probably frequent, and mainly
on account of the fact that the winter-rings are often not more distinct
than the 'secondary-rings,' which have nothing to do with the winter-
rings. Each naturalist forms his own praxis in his distinction between
the two sorts of rings. He may be able to count the number of 'winter-
rings' approximately in the same way at different times, but this is
not a sufficient proof that the counting is correct. There is some
evidence that the first winter-ring is often feeble or wanting, and in
old specimens it is usually the case that there is an outer zone with
more or less indistinct rings which are lying very close to each other.
In this zone the distinction between winter-rings and secondary-rings
becomes highly arbitrary. . . ."

Although it is true that "each naturalist forms his own praxis" in his
distinction between true and false rings, tests have shown that two
skilled scale-readers independently examining the same sample of scales
arrive at very similar results. Lea (12, p. 17) provides two instances of
this. The age of each of one hundred herrings was estimated from their
scales by himself and Knut Dahl in turn, and the final results compared.
There was a discrepancy of one year in three instances where there was
only one scale available for examination, but absolute agreement in all
cases where three scales were available. The second test was conducted
by Lea and Hjort. Their separate age estimates for several hundreds
of herrings agreed in every case but one, in which Lea had undoubtedly
made a mistake. But, even this agreement between independent observers
does not dispose of the possibility of false rings being so similar in appear-
ance to true winter-rings that they defy detection in the routine inspection
of scales. The remarkable Norwegian results over an unbroken period
of many years (Lea, 16), however, surely constitute a striking argument
in favour of the view that the unavoidable personal error in age estima-
tion from scales cannot, on the whole, be regarded as unduly large.

That reliable information on the time of formation of the first winter-
ring is essential for accurate age estimations from scales will be realised
from the study of the problem presented to the Scottish investigators
(20). Two broods of herring larvae occur annually, the one in the spring
with its maximum in April, and the other in the autumn, at its height in September. The members of the spring brood and a proportion of the autumn brood are fully scaled by the end of the year. But a large—if not the larger—proportion of the autumn brood is carried over the winter months still in the unscaled condition. It is at once seen that when these latter fishes acquire their first winter-ring, their more advanced relatives of the same autumn spawning, as well as those spawned in the previous spring, will be recording their second winter-ring. A preliminary report on the growth of winter-spawned herrings at Plymouth is given in Part 4 of this series (see p. 305), in which this question is more fully considered. It will be sufficient here to state that in the Plymouth estimates of age for fish of the south-western area, the first winter-ring was in all cases regarded as formed during the first winter following that of birth.

The Part of the Body from which Scales are taken.

At a meeting of the Herring Committee of the International Council (2, p. 34) held in Oslo, June, 1923, a recommendation was made regarding the part of the body from which scales for age-estimation should be taken. When the rings on the scale are distinctly marked and few in number, it is of little importance from which part of the body the scales are taken; on the other hand, when the rings are many and closely set, accurate counting becomes more and more difficult as one proceeds from scales taken from the anterior region of the body near the tip of the pectoral fin, to those either from the posterior part of the body or on the dorsal surface. When, therefore, the object is to count the maximum number of rings, those scales (viz. those from the anterior part of the body near the tip of the pectoral fin) should be chosen which show the fullest development of the rings.

Methods of Recording Estimates of Age.

The manner in which scale-readings may be recorded needs some consideration. Hodgson (7, p. 13) differentiates between three annual phases in the appearance of the edge of the scale. During the physiological winter the edge is in the "A" phase with a clear margin continuous around the scale. At the onset of new growth the scale-striae are seen to extend right to the margin at the "top" of the scale, i.e. at the edge furthest from the base-line, leaving on the inside the first indications of a new winter-ring. On each side, however, the clear margin of the "A" phase persists. This is the "B" phase. In the height of the physiological summer, the striae extend to the margin throughout the whole of the edge, and the new winter-ring is left behind in its complete form. This is the "C" phase. A concise record of an age-estimate, therefore, would include
not only the number of completed winter-rings, but also some indication of the phase of the edge. Thus, the common practice in scale-recording of referring fishes to appropriate "age-groups" according to the number of completed winter-rings inside the edge of the scale does not differentiate between scales which are in different phases of edge. At Plymouth it has been found convenient to consider an "A" phase edge as a completed winter-ring, and to record scale-readings in the form of a double entry of the number of summer zones and winter-rings. The following tabular illustration will serve to demonstrate the method adopted:

<table>
<thead>
<tr>
<th>Phase of scale edge</th>
<th>No. of completed winter rings inside edge</th>
<th>Plymouth record of age</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>A</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>B</td>
<td>3</td>
<td>4 or 5*</td>
</tr>
<tr>
<td>C</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

An alternative method of procedure is to assign fishes to "year-groups" (Lea, 15, p. 104). When a fish is placed in, say, the 1920 year-group, this means that the fish in question formed its first summer zone on the scale in the year 1920. For all fishes spawned in the spring, the date of the year-group will be identical with that of the year of birth, but in the case of those spawned during the latter part of the year, such is not necessarily so. It has been shown on page 243 above that although some of the autumn-spawned Scottish herrings definitely acquire scales and so form a summer zone during the calendar year of birth, a greater proportion of them are carried over the winter still in the unscaled condition, so that the first summer zone is not formed until the calendar year following that of birth. Thus the Scottish herrings spawned in the autumn of 1920 would be assigned partly to the 1920 year-group and partly to the 1921 year-group. In a place like Plymouth, where the maximum spawning takes place during the months of December and January, fishes born in two calendar years will obviously form part of the same year-group. For example, the offspring of the spawning of December, 1919—January, 1920, formed their first summer zone in 1920, and therefore belong to the 1920 year-group.

During actual scale-reading it is often convenient to create a number of "interval age groups," which allow the inclusion of fishes of doubtful age in their approximately correct position in the series of age groups. For instance, an observer may feel certain that a scale is either 6-zoned 6-ringed or 7-zoned 7-ringed, but is unable to decide which; by assigning

* Personal judgment would be exercised in light of particular circumstances.
FIG. 1.—Vertical projector used at Plymouth.
A. Black cloth hood spread over wire framework.
B. Mirror set at 45°.
C. Measuring plate of ground glass. See Fig. 2 for details.
D. Lamp-box sheltering 100-watt electric lamp.
E. 200-c.c. flask acting as condenser. It is filled with distilled water to which has been added:
   5 drops of 0.04% Sol. of Brom-thymol blue.
   5 drops of 0.04% Sol. of Brom-cresol purple.
   2 drops of N/Sodium bicarbonate.
F. Wooden clamp for securing microscope stand.
G. Thumb-screws on bolts for raising or lowering projector.
H. Slot in vertical stand to allow passage of G.
this record to an interval group which includes fishes of age, 6-zoned 6-ringed or 7-zoned 7-ringed, he is able to give an approximate estimate of age.

Estimations of Growth from Scale Measurements.

A considerable volume of information has been accumulated in recent years on the growth of the herring, from scale measurements. The method is based on the presupposition that the growth of the scale takes place at a rate simply proportional to the rate of growth of the fish (Lea, 12). Probably this initial assumption is not entirely correct, but most valuable results may nevertheless be obtained by a cautious use of the method (vide Lea, 14). As Lea points out, however, it is necessary to remember that all sources of error which make themselves apparent in age determinations also apply to estimates of growth derived from scale measurements. If a winter-ring be overlooked, then two years’ growth will be taken together as that of one, and the growth of the succeeding years will be erroneously reckoned. If, on the other hand, a secondary-ring be wrongly interpreted as a true winter-ring, the year’s growth will be divided into two, and that of the following years again erroneously recorded.

Now, although all actual measurements of growth increments are made on the scales themselves, it is usual to express the results not in terms of the absolute scale measurements, but in terms of the length of the fish corresponding with the scale measurements. Thus, the values \( l_1, l_2, \ldots, l_x \) are the theoretical lengths of \( x \) the fish when the first, second \( \ldots \) \( x \) th winter-ring was formed on the scale.

Apparatus used in Study of Growth from Scale Measurements.

The determinations of the theoretical lengths \( l_1, l_2, \ldots, l_x \) of a fish at the time of the formation of the first, second \( \ldots, x \) th winter-ring on the scale were made at Plymouth on the principle of the Norwegian method (Lea, 15, p. 103); but by projecting the image of the scale on to a ground-glass plate fitted with a movable swing-arm, the use of slips of paper was avoided, and the values of \( l_1, l_2, \ldots, l_x \) read off direct. The general features of the simple home-made projector used for this purpose may be gathered from the accompanying Figs. 1 and 2 and their legends, without further comment.

Growth “Types” from Scale Measurements.

If in a given sample all the fishes of a particular age-group agreed closely in their values of \( l_1, l_2, \ldots, \) etc., it would be reasonable to suspect that they had lived and grown up under similar growth conditions.
FIG. 2.—Diagrams illustrating method of using measuring-plate (Fig. 1-C).

B. The actual plate with its movable swing-arm, and the projection of a herring-scale from a fish of length 22.5 cm.

A. The reflected view of B as seen by the observer in the mirror (Fig. 1-B). The swing-arm is first adjusted to register 22.5 cm. The measuring-plate is then moved until its base-line coincides with the base-line of the herring-scale, and one of its vertical lines, A B, passes through the scale-origin B, while the edge of the herring-scale is in register with the swing-arm at point C. To determine "I," for the fish, all that is necessary is to move the swing-arm until it is in register at point D along the line A B with the first ring of the herring-scale, and then read off the value (I₁ = 9.0 cm.).
Under such circumstances it would be permissible to use the average values of \( l_1, l_2, \) etc., as a "type" representing the stock of the particular age in the sampling area. In this way, Hodgson (8, p. 16) arrived at his "Southern North Sea Type," for which \( l_1=8 \) cm., \( l_2=16 \) cm., \( l_3=20-21 \) cm., and \( l_4=23 \) cm. Having suspected such a "type" steps must be taken to trace the series of events by which the type has been built up—the feeding grounds visited, the grounds where spawning, if any, has occurred, and ultimately the place where the fishes of that type were themselves spawned. Until a logical and consecutive account of these events can be given, a conceived "type" remains merely a more or less convenient summary of mathematical averages.

A common experience, however, is to find that the values of \( l_1, l_2, \) etc., for fishes of the same age in the same area vary over so wide a range that the calculated averages are not in the least representative.* For example, the following table shows the values of \( l_1, l_2, l_3, \) and \( l_4 \) for each of 374 herrings of the 4-zoned 4-ringed class taken off the Sussex Coast during the present investigations:

<table>
<thead>
<tr>
<th>LENGTH. CENTRAL VALUES OF 22 LENGTH CLASSES.</th>
</tr>
</thead>
<tbody>
<tr>
<td>------</td>
</tr>
<tr>
<td>( l_1 )</td>
</tr>
<tr>
<td>( l_2 )</td>
</tr>
<tr>
<td>( l_3 )</td>
</tr>
<tr>
<td>( l_4 )</td>
</tr>
<tr>
<td>Arith. Mean.</td>
</tr>
<tr>
<td>CM.</td>
</tr>
</tbody>
</table>

Mere inspection of these data will suffice to convince us that there is little justification for assuming that all the fishes belong to one "type" for which \( l_1=10:73 \) cm., \( l_2=18:04 \) cm., \( l_3=21:89 \) cm., and \( l_4=23:65 \) cm. It will be observed that the individual values of \( l_1 \) show a tendency to group themselves about two modes, 8-5 cm. and 11-5 cm., respectively. This suggests that in our data we may be dealing with a mixture of two "types," one of which has \( l_1=\text{ca.} \) 8-5 cm., and the other, \( l_1=\text{ca.} \) 11-5 cm.; but as there is no definite indication of "bimodality" in the corresponding values of \( l_2, l_3, \) or \( l_4, \) it is not possible from the data, as they stand, to make further progress in the definition of the two conceivable types.

If we analyse the data more closely and ascertain for each individual value of \( l_1 — \) the corresponding values of \( l_2, l_3, \) and \( l_4 — \) we find that fishes with the smaller values of \( l_1 \) tend on average to add relatively greater annual increments of length in subsequent years than those with the

* The reliability of the average may be tested by calculation of its "standard error." See page 259 for a consideration of this statistic. Also consult Lea (15, p. 133) for an account of the use of the standard error of the mean in this connection.
larger values of $l_1$. In illustration of this point, the following is a table correlating the values of $l_2$ with observed values of $l_1$:—

**Correlation Table.**

<table>
<thead>
<tr>
<th>Length $l_1$</th>
<th>Central Values of 13 Length Classes.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CM.</td>
<td>5-5 6-5 7-5 8-5 9-5 10-5 11-5 12-5 13-5 14-5 15-5 16-5 17-5</td>
</tr>
<tr>
<td>13-5</td>
<td>2 2 2 2 2 2 2 2 2 2 2 2 2</td>
</tr>
<tr>
<td>14-5</td>
<td>1 1 1 1 1 1 1 1 1 1 1 1 1</td>
</tr>
<tr>
<td>15-5</td>
<td>0 0 0 0 0 0 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>16-5</td>
<td>7 7 7 7 7 7 7 7 7 7 7 7 7</td>
</tr>
<tr>
<td>17-5</td>
<td>1 1 1 1 1 1 1 1 1 1 1 1 1</td>
</tr>
<tr>
<td>18-5</td>
<td>1 1 1 1 1 1 1 1 1 1 1 1 1</td>
</tr>
<tr>
<td>19-5</td>
<td>2 2 2 2 2 2 2 2 2 2 2 2 2</td>
</tr>
<tr>
<td>20-5</td>
<td>3 3 3 3 3 3 3 3 3 3 3 3 3</td>
</tr>
<tr>
<td>21-5</td>
<td>2 2 2 2 2 2 2 2 2 2 2 2 2</td>
</tr>
<tr>
<td>22-5</td>
<td>1 1 1 1 1 1 1 1 1 1 1 1 1</td>
</tr>
<tr>
<td>Totals</td>
<td>1 3 24 70 54 33 69 47 26 12 11 3 1</td>
</tr>
</tbody>
</table>

Mean $l_1$ = 14.5, Mean $l_1$ = 9.0

We observe from the values of the character (mean $l_2$-$l_3$) shown at the foot of the table that there is a steady decrease from left to right, which is to say that the smaller values of $l_1$ have added larger increments of length than the greater values of $l_1$ have done. In a similar manner it could be shown that fishes with the smaller $l_2$ added larger increments of length, in attaining the length $l_3$, than did the fishes with the greater $l_2$; and similarly those with the smaller $l_3$ in growing to the length $l_4$.

If the mean values of $l_2$ for successive values of $l_1$ be plotted graphically as in Fig. 3, it will be found that they conform reasonably well with a straight line AB, except at the extremes where the number of observations is small. It is possible to determine a mathematical equation representing the straight line which best fits the data, and it is:—

$$l_3 = 0.59 l_1 + 11.70$$

How nearly this equation fits the data may be seen when the observed values and the calculated values are tabulated in conjunction, thus:—

<table>
<thead>
<tr>
<th>$l_1$</th>
<th>5-5</th>
<th>6-5</th>
<th>7-5</th>
<th>8-5</th>
<th>9-5</th>
<th>10-5</th>
<th>11-5</th>
<th>12-5</th>
<th>13-5</th>
<th>14-5</th>
<th>15-5</th>
<th>16-5</th>
<th>17-5</th>
</tr>
</thead>
<tbody>
<tr>
<td>observed</td>
<td>14.50</td>
<td>14.83</td>
<td>14.06</td>
<td>15.17</td>
<td>16.93</td>
<td>17.69</td>
<td>17.26</td>
<td>17.93</td>
<td>18.59</td>
<td>19.10</td>
<td>19.73</td>
<td>20.00</td>
<td>20.50</td>
</tr>
<tr>
<td>$l_2$</td>
<td>0.59</td>
<td>1.70</td>
<td>2.80</td>
<td>3.90</td>
<td>4.90</td>
<td>5.90</td>
<td>6.90</td>
<td>7.90</td>
<td>8.90</td>
<td>9.90</td>
<td>10.90</td>
<td>11.90</td>
<td>12.90</td>
</tr>
</tbody>
</table>

In a similar way it can be shown that the values of $l_3$ and $l_4$ are "positively correlated" according to a straight line equation, and also those of $l_3$ and $l_4$, the two equations being:—

$$l_3 = 0.50 l_2 + 12.80$$

$$l_4 = 0.77 l_3 + 6.71$$
It is equally possible to calculate two further equations, the one linking $l_3$ and $l_5$, and the other $l_4$ and $l_6$:

$$l_3 = 277l_1 + 18.92 \quad \dots \quad (4)$$
$$l_4 = 221l_1 + 21.28 \quad \dots \quad (5)$$

From the above we thus learn that instead of endeavouring to segregate the fishes into two growth "types" based on average values of $l_1, l_2, l_3,$ and $l_4$, we can summarise the data in the form of five simple equations (1 to 5 above), from which we can calculate the length reached by a fish of any given value of $l_1$ at the end of any subsequent growth period.

**Fig. 3.—** Mean values of $L_2$ for observed values of $L_1$ (cm.). Fishes from Sussex Coast.

**Comparisons between Samples from different Areas.**

Equations 1 to 5 above enable us to compare the growth of the Sussex fishes with that of similar aged fishes from other places. Supposing, for example, that we wished to compare the growth of the Sussex fishes as indicated by equations 1 to 5 with that of fishes taken at Grimsby. Hjort (6) has given data on 141 herrings obtained from Grimsby on September 30th, 1906 (Collection, Nr. 26, p. 151), which may be used as an illustration. These 141 herrings were of a similar age to, and for
all practical purposes otherwise comparable with, the Sussex fishes.
The average values of \( l_1, l_2, l_3, \) and \( l_4 \) given by Hjort are:

<table>
<thead>
<tr>
<th>( l_1 )</th>
<th>( l_2 )</th>
<th>( l_3 )</th>
<th>( l_4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.2</td>
<td>17.1</td>
<td>21.6</td>
<td>24.1</td>
</tr>
</tbody>
</table>

The corresponding values for the Sussex fishes are:

<table>
<thead>
<tr>
<th>( l_1 )</th>
<th>( l_2 )</th>
<th>( l_3 )</th>
<th>( l_4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.73</td>
<td>18.04</td>
<td>21.89</td>
<td>23.65</td>
</tr>
</tbody>
</table>

Now we wish to learn how the growth of the Grimsby herrings from \( l_1 \) to \( l_2 \), from \( l_2 \) to \( l_3 \), and from \( l_3 \) to \( l_4 \) compares with that of the Sussex herrings for the corresponding periods. Let us turn to our equation (1) and calculate the length \( l_2 \), which would be reached by a Sussex fish commencing at a length \( l_1 = 9.2 \) cm., equal to the average value of \( l_1 \) for the Grimsby fishes:

\[
l_2 = 0.59 l_1 + 11.70 \quad (1)
\]

\[
= 0.59(9.2) + 11.70
\]

\[
= 17.13 \text{ cm.} \quad \text{(Compare with Hjort’s observed average } l_2 = 17.1 \text{ cm.)}
\]

Again, let us determine from our equation (2) the length \( l_3 \) which would be reached by a Sussex fish commencing at a length \( l_2 = 17.1 \) cm., which is the average value of \( l_2 \) for the Grimsby herrings:

\[
l_3 = 0.504 l_2 + 12.80 \quad (2)
\]

\[
= 0.504(17.1) + 12.80
\]

\[
= 21.42 \text{ cm.} \quad \text{(Compare with Hjort’s observed average } l_3 = 21.6 \text{ cm.)}
\]

Finally, from our equation (3) let us calculate the length \( l_4 \) which would be reached by a Sussex fish commencing at a length \( l_3 = 21.6 \) cm., the average value of \( l_3 \) for the Grimsby fishes:

\[
l_4 = 0.774 l_3 + 6.71 \quad (3)
\]

\[
= 0.774(21.6) + 6.71
\]

\[
= 23.42 \text{ cm.} \quad \text{(Compare with Hjort’s observed average } l_4 = 24.1 \text{ cm.)}
\]

From these results we are surely justified in suspecting that the Sussex fishes and the Grimsby fishes had, for the two years in which they grew from \( l_1 \) to \( l_3 \), been under growth conditions of a closely similar nature. In their fourth year, however, when changing from \( l_3 \) to \( l_4 \), it would seem
that the Grimsby fishes had experienced slightly the more favourable conditions for growth.

The fact that the average value of $l_1$ is greater for the Sussex fishes than for those from Grimsby shows that fishes with the greater values of $l_1$ were relatively more numerous on the Sussex ground than on the Grimsby ground at the time of sampling.

A more critical method for comparing these two sets of data would have been to determine equations for the Grimsby fishes similar to equations 1 to 5 and then make a comparison between corresponding equations. It does not seem necessary, however, here to do this as the above is sufficient to illustrate the usefulness of the equations in this instance.

**Differences of Growth as Evidence of Migrations.**

One further important application of these equations must be considered. Storrow (21, p. 13), following the earlier observations of Lea, has endeavoured to trace the movements of herring shoals by a study of the differences in the yearly increments of growth as exhibited on the scale. If we imagine a herring which, say, for the first three years of its life had remained within a region where unduly large growth does not occur, and assume that it migrates at the end of the third year to a region where it can grow at a relatively faster rate than hitherto, then its scale would show an unusually wide zone of growth for the latter period. Storrow has followed this conception and thereby endeavoured to trace the movements of herrings from "narrow seas" to "oceanic" conditions, and vice versa. Now, although there appears to be nothing inherently wrong with this principle, the study of our growth equations shows most clearly that care must be exercised in interpreting the differences between annual increments of growth as shown by the scales. Every difference must be considered in its relation to the actual lengths of the fishes at the commencement of the growth period in which the difference arose, for we have seen that fishes which are smaller at the beginning of any one growth period will add a greater increment of length by the end of that period than those which are larger.* Thus, if in an Area, A, three-year-old fishes were found to have added 2 cm. during the fourth year, while in a second area, B, similar-aged fishes were found to have added 3 cm. during the same period, it is incorrect to accept this observation as evidence of better growing conditions in area B unless it can be demonstrated that at the beginning of the fourth year the average size was the same in the two areas.

* Watkin (22, Part V, p. 50) refers to this phenomenon as "compensatory growth." See also D'Arcy Thompson, "Growth and Form" (p. 78).
THE NUMBER OF VERTEBRÆ.

Preparation of Skeletons.

At an early stage of the work a number of records were lost due to three causes:—

1. Identification labels became unreadable as the result of boiling.
2. Identification labels became detached during boiling or subsequent cleaning.
3. An undue amount of breakage of skeletons occurred during cleaning.

The first of these difficulties was overcome by writing the serial number of the fish in marking ink, sold under the commercial name of "Melanyl," on linen tape. The second was overcome by securing the safety-pin through the skull behind the eyes. Breakage of skeletons was reduced to a minimum by commencing with the fish immersed in cold water and raising almost but not quite to the boiling point. In this way the flesh was just set and came away easily and cleanly, leaving the skeleton intact with all the vertebrae secure. A final cleaning with a tooth-brush in cold water ensured a skeleton in excellent condition for counting vertebrae when dried.

Use of Radiograms.

An alternative method for determining the number of vertebrae has been tried, which promises to be useful and convenient when the fishes are too small to be boiled. Radiograms are taken and counts of the vertebrae made direct from them. Dr. J. B. Stauffer, of Devonport, was good enough to arrange for a trial of the method at the Royal Albert Hospital, and more recently Dr. H. A. Harris, of the Institute of Anatomy, University College, London, provided me with a series of most beautiful negatives of herrings ranging from about 35 mm. to 80 mm. in length.

The Vertebral Count made.

The actual count of the "number of vertebrae" was identical with that of Orton (18, p. 80); that is to say, exclusive of the terminal ossicle. Fig. 4, reproduced from Johansen (10, p. 6, Fig. 4), may be used to demonstrate the difference between the count made at Plymouth and that adopted by Johansen. At Plymouth the count concluded with the vertebra marked 2, whereas Johansen included No. 1 in his count. Thus, all Plymouth records of the total number of vertebrae need to be increased by one before they become comparable with those of Johansen.

A sample of one hundred fishes usually included one which exhibited an abnormality in the vertebral column (see Ford and Bull, 5). Such abnormal specimens were not used for the statistical count.
The Chances of drawing truly Representative Samples from a known Population.

If a random sample of herrings be taken and the number of vertebrae in each determined, a “frequency” table can be drawn up to show the number of fishes having \(a, b, \ldots x\) vertebrae respectively. Now it is no more reasonable to expect this table to depict an exact representation of the proportions in the population sampled than it would be to expect to draw the ace of spades from a pack of cards in a single trial. Either event might, of course, occur, but the “chances” are against it. In the case of the pack of cards which we know to consist of 52 dissimilar cards, we say that the “chance” of cutting the ace of spades at a single trial is \(\frac{1}{52}\), for any one of the pack is equally likely to turn up. Similarly, if we knew the exact proportions of fishes having \(a, b, \ldots x\) vertebrae in a population of herrings, we could calculate the chance of drawing a sample of any given size and of any given composition. For example, let us assume a very large population of herrings having 57, 56, or 55 vertebrae, the proportions of the three classes being:

<table>
<thead>
<tr>
<th>Fishes with 57 vertebrae</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot; 56 &quot;</td>
<td>60%</td>
</tr>
<tr>
<td>&quot; 55 &quot;</td>
<td>30%</td>
</tr>
</tbody>
</table>

![Diagram of caudal bones of a herring](image)

**Fig. 4.**—Caudal bones of a herring of 75 mm., Limfjord, Jutland, Jan. 1916 (after Johansen).
The formula given in the footnote below enables us to calculate the chance of drawing a sample of one hundred fishes in which the proportions of the vertebral classes are of a particular value. In the following table the values of the chance for nine conceivable samples are given:—

<table>
<thead>
<tr>
<th>No. of Vertebræ</th>
<th>A₁</th>
<th>A₂</th>
<th>A₃</th>
<th>A₄</th>
<th>A₅</th>
<th>A₆</th>
<th>A₇</th>
<th>A₈</th>
<th>A₉</th>
</tr>
</thead>
<tbody>
<tr>
<td>57</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td>15</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>56</td>
<td>78</td>
<td>58</td>
<td>60</td>
<td>70</td>
<td>60</td>
<td>50</td>
<td>60</td>
<td>60</td>
<td>40</td>
</tr>
<tr>
<td>55</td>
<td>21</td>
<td>41</td>
<td>35</td>
<td>25</td>
<td>36</td>
<td>85</td>
<td>25</td>
<td>20</td>
<td>40</td>
</tr>
<tr>
<td>n</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Value of Chance</th>
<th>10,000,000</th>
<th>1,000,000</th>
<th>1,000</th>
<th>1,000</th>
<th>1,000</th>
<th>1,000</th>
<th>10,000</th>
<th>10,000,000</th>
</tr>
</thead>
</table>

It is seen that sample A₅ is more likely to occur than any one of the others. This illustrates the general fact that in taking a random sample from a population, the one most likely to occur is that which is a faithful representation of the population sampled. The values of the chance also show that the nearer the sample approaches the true representation the better its chance of being drawn. The values of the chance for samples A₁ and A₉ which exhibit proportions of the vertebrae classes appreciably different from those of the population are, thus, extremely small.

The value of the “chance” is materially affected by an increase in the number of fishes comprising the sample. This important fact may be illustrated by comparing the values of the chance for samples constructed as follows:—

<table>
<thead>
<tr>
<th>Percentage No. of fishes with</th>
<th>Series 1</th>
<th>Series 2</th>
<th>Series 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>57 vertebrae</td>
<td>10%</td>
<td>10%</td>
<td>20%</td>
</tr>
<tr>
<td>56</td>
<td>40%</td>
<td>60%</td>
<td>60%</td>
</tr>
<tr>
<td>55</td>
<td>50%</td>
<td>30%</td>
<td>20%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Value of chance when n=100</th>
<th>0000001</th>
<th>0012</th>
<th>00003</th>
</tr>
</thead>
<tbody>
<tr>
<td>n= 50</td>
<td>000208</td>
<td>0023</td>
<td>00114</td>
</tr>
<tr>
<td>n= 20</td>
<td>008245</td>
<td>0056</td>
<td>01555</td>
</tr>
</tbody>
</table>

* The “chance” of a random sample of n articles giving a₁ of the first kind, a₂ of the second ... aₙ of the last kind is given by: \[ \frac{n!}{a₁! a₂! ... aₙ!} \cdot \frac{P₁ P₂ ... Pₙ}{Pₙ₁ Pₙ₂ ... Pₙₙₕ} \]
where P₁, P₂, ... Pₙ are the “probabilities” that a single item drawn from the population is of the 1st, 2nd ... nth kind (vide Fisher, 4, p. 66).
It will be seen that the effect of increasing (n) from 20 to 100 has been to lessen the absolute value of the chance in each of the series, but that the degree of lessening is very much more marked in series 1 and 3 than in series 2. In other words, although we have actually lessened our chance of drawing an exact representation of the population (series 2), we have at the same time reduced to a far greater degree the chance of drawing an "outside" sample (series 1 and 3). Thus, in practice, by making our statistical sample as large as possible, we very materially add to our chances of gaining a close approximation to the composition of the population sampled.

In the foregoing illustrations, however, we have been dealing with samples from a population of known composition, whereas normally we have to do with samples which come from a population whose composition we never can know, but concerning which we wish to obtain information. Having taken a sample, no matter how large, we are bound to admit that in our very first trial we may have drawn an "outside chance" of composition quite different from that of the population. But while this fact must not be overlooked, we are entitled to regard this event as highly improbable, and to suspect that the sample is a not unreasonable approximation to the true proportions in the population.

*The Arithmetic Mean and the Standard Deviation.*

The characterisation of a population in the form of a frequency table showing the proportions of its component vertebrae classes is somewhat unhandy for statistical purposes. As an alternative it has been found more convenient to specify the composition of a population in terms of certain purely mathematical characters or "parameters" (Fisher, 4, p. 7). There are two "parameters" which we need to consider at this stage, viz. the Arithmetic Mean Number of vertebrae (M) and the Standard Deviation (σ). Let us consider a sample of 363 Shetland herrings examined by Johansen (11, p. 8):

<table>
<thead>
<tr>
<th>No. of Vertebrae</th>
<th>51</th>
<th>54</th>
<th>55</th>
<th>56</th>
<th>57</th>
<th>58</th>
<th>59</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Fishes</td>
<td>1</td>
<td>1</td>
<td>10</td>
<td>170</td>
<td>163</td>
<td>17</td>
<td>1</td>
</tr>
</tbody>
</table>

Let us imagine 363 exactly similar weights suspended from a graduated rod in their appropriate positions, according to the above distribution, as in Fig. 5. Obviously there is but one point on the rod at which the latter could be supported so that this system of weights would remain in an exact state of "balance." This point is in the region of 56-504.
That this is so is shown by the fact that if we multiply the number of weights at each point of suspension by their distance from 56.504, the sum of the products to the left is approximately equal to that of the products to the right:

<table>
<thead>
<tr>
<th>Left.</th>
<th>Right.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1 \times 5.504 = 5.504$</td>
<td>$163 \times 0.496 = 80.848$</td>
</tr>
<tr>
<td>$1 \times 2.504 = 2.504$</td>
<td>$17 \times 1.496 = 25.432$</td>
</tr>
<tr>
<td>$10 \times 1.504 = 15.040$</td>
<td>$1 \times 2.496 = 2.496$</td>
</tr>
<tr>
<td>$170 \times 0.504 = 85.680$</td>
<td></td>
</tr>
</tbody>
</table>

| Sum $= 108.728$ | Sum $= 108.776$ |

The value, 56.504, is the Arithmetic Mean Number of Vertebrae for the sample. This statistic is, thus, simply and solely a numerical value representing the centre point of the distribution of the fishes among the several vertebrae classes; it has no concrete counterpart in nature.

![Diagram illustrating the nature and significance of the "Arithmetic Mean"](image_url)

But, just as there are many ways of redistributing the 363 weights along the rod while still retaining balance at point 56.504, so are there many different frequency distributions having the same value for the Arithmetic Mean. We therefore need some further character which will serve to specify the extent of the scattering or dispersion of individuals about the mean. One way of doing this is to add together all the individual deviations from the mean, irrespective of their + or - sign, and divide by the total number of individuals. The result is the average or mean deviation. Its value for the above sample is .599. A more convenient character, however, is the standard deviation, which, like the average deviation, is a form of average of individual deviations from the Arithmetic Mean, but derived in a different manner. Instead of ignoring the positive and negative signs of these deviations, the deviations are squared, thereby bringing all of them to the same positive sign. An average of these
squares is then taken. The square-root of this mean square is the standard deviation. Its value for the above sample is \( \pm 0.718 \).

Errors due to Sampling.

From the data for a sample (A) of 363 fishes, Johansen thus arrived at estimates \( m_a (=56.504) \) and \( \sigma_a (=\pm 0.718) \) of the parameters \( M \) and \( \sigma \) for the particular population he had sampled. Now, it is almost certain that if Johansen had been able to analyse a second sample (B), also of 363 fishes and taken from the same population, he would have arrived at a second pair of estimates, \( m_b \) and \( \sigma_b \), which differed from \( m_a \) and \( \sigma_a \), notwithstanding the fact that both samples had been taken from the same population. The estimates \( m_a \) and \( m_b \) (and similarly \( \sigma_a \) and \( \sigma_b \)) are, as we say, subject to "errors due to sampling," and it is necessary for us next to consider the question of the magnitude of such "errors." We have already seen that in drawing a single sample from a population, the chances are all in favour of its being one which fairly faithfully depicts the true composition of the population sampled. If, then, we proceed to draw a number of similar samples in succession, we may expect near estimates of \( M \) to occur more frequently than outside ones. As sampling is continued, the resulting estimates \( m_a, m_b, \ldots, m_x \) tend to group themselves in a frequency distribution about a central value, itself a new and probably closer estimate of \( M \) than any one of the series. The larger the number of samples taken, the more regular does the distribution become, and the nearer to \( M \) does the central value approach. In the extreme theoretical case of an infinitely large number of samples, the distribution is perfectly symmetrical and of a particular mathematical type, known as the "normal," while the central value is identical with \( M \). The standard deviation \( (\sigma_M) \) of this normal distribution of estimates of \( M \) is a convenient measure of the average amount by which individual sample estimates differ from \( M \) itself. Moreover, the amount \( (m_x - M) \) by which any one particular sample estimate differs from \( M \) can be expressed in terms of the average \( (\sigma_M) \). Now Fisher (4, p. 46) tells us that \( (m_x - M) \) exceeds twice \( \sigma_M \) about once in 22 trials, thrice \( \sigma_M \) only once in 370 trials, while nearly a thousand million trials are needed for \( (m_x - M) \) to exceed \( \sigma_M \) sixfold. With these results in mind, statisticians have formally agreed to assume that a random sample, drawn from a population whose arithmetic mean number of vertebrae is \( M \), is not likely to yield an estimate of \( M \), which falls outside the range \( (M - 2 \sigma_M) \) and \( (M + 2 \sigma_M) \). But, as we have seen, \( \sigma_M \) is a purely theoretical value representing the standard deviation of the distribution of an infinite series of estimates of \( M \), so that until we can obtain some knowledge of its probable value, the above formal agreement is of little practical use. Obviously
we cannot obtain the exact value of \( \sigma_M \), but fortunately we can derive an estimate of it from sample data. This estimate of \( \sigma_M \) is known as the standard error of the mean (\( \sigma_M \)).

Returning now to Johansen’s sample of 363 fishes, we have an estimate of \( M \) (\( m=56.504 \)) and an estimate of \( \sigma_M \) (\( \sigma_m=\pm0.0377 \)), so that we may say that the value of \( M \) lies most likely between 56.504 - \( 2(0.0377) \) and 56.504 + \( 2(0.0377) \).

**Comparisons between Samples.**

Supposing that, in practice, we have analysed two samples and arrived at the following estimates:

<table>
<thead>
<tr>
<th>Sample 1</th>
<th>Sample 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Fishes</td>
<td>( n_1 )</td>
</tr>
<tr>
<td>Arithmetic Mean No. of Vertebrae</td>
<td>( m_1 )</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>( \sigma_1 )</td>
</tr>
<tr>
<td>Standard Error of the Mean</td>
<td>( \sigma_m_1 = \frac{\sigma_1}{\sqrt{n_1}} )</td>
</tr>
</tbody>
</table>

We observe that there is a difference \( (m_1-m_2) \) between the means. How can we decide whether this difference is one which might reasonably arise between two samples from the same type of population; or is due to the fact that we have sampled different populations? If the two samples had been taken from the same population, the probability is that neither \( m_1 \) nor \( m_2 \) would have been greatly different from the parameter (\( M \)) of the population. From this it follows that their difference \( (m_1-m_2) \) would not have been large. Here again, by formal agreement, we say that probably \( (m_1-m_2) \) would not have exceeded twice its own standard error, viz. \( \sqrt{\sigma_m_1^2 + \sigma_m_2^2} \). If then we find that the difference \( (m_1-m_2) \) does actually exceed twice its own standard error, we suspect that the samples had been taken from different populations.

It will be instructive here to return to the question of the number of fishes \( (n) \) in our statistical sample. We have seen (p. 255) from the study of the “chance” for each of a series of similar composed samples in which \( (n) \) is 20, 50, and 100 respectively that the value of the chance is very materially affected by the value of \( (n) \). This same fact may be

---

\* The standard error of the mean (\( \sigma_M \)) is calculated by \( \sigma_M = \frac{\sigma}{\sqrt{n}} \) where \( \sigma \) = standard deviation of sample.

\( n \) = number of individuals in sample.

The mathematical argument for using the value of \( \sigma_M \) as an estimate of \( \sigma_M \) may be found in a standard textbook. See table of Notations and Formulae on page 261.

\† See table of Notations and Formulae on page 261.
demonstrated by comparing the values of \((m)\) and of their standard errors for samples of different sizes from two populations \(A\) and \(B\): 

<table>
<thead>
<tr>
<th>Number of Fishes with</th>
<th>Population A.</th>
<th>Population B.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sample (a_1)</td>
<td>Sample (a_2)</td>
</tr>
<tr>
<td>57 vertebræ</td>
<td>10</td>
<td>40</td>
</tr>
<tr>
<td>56 , ,</td>
<td>60</td>
<td>240</td>
</tr>
<tr>
<td>55 , ,</td>
<td>30</td>
<td>120</td>
</tr>
<tr>
<td>(n)</td>
<td>100</td>
<td>400</td>
</tr>
<tr>
<td>Arithmetic Mean (m)</td>
<td>55.80 (=m_1)</td>
<td>55.90 (=m_2)</td>
</tr>
<tr>
<td>No. of Vertebræ ((m))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard Deviation ((\sigma))</td>
<td>0.60 (=\sigma_1)</td>
<td>0.62 (=\sigma_2)</td>
</tr>
<tr>
<td>Standard Error of Mean</td>
<td>(\sigma_m = \frac{\sigma}{\sqrt{n}})</td>
<td>0.060</td>
</tr>
<tr>
<td>(\sigma_m^2)</td>
<td>0.0036</td>
<td>0.00090</td>
</tr>
</tbody>
</table>

Comparing samples \(b_1\) and \(a_1\): \((m_2-m_1) = 0.10\);

standard error of \((m_2-m_1) = \sqrt{\sigma_{m_2}^2 + \sigma_{m_1}^2} = 0.087\).

Comparing samples \(b_2\) and \(a_2\): \((m_2-m_1) = 0.10\);

standard error of \((m_2-m_1) = \sqrt{\sigma_{m_2}^2 + \sigma_{m_1}^2} = 0.043\).

For the samples of 100, viz. \(b_1\) and \(a_1\), the difference \((m_2-m_1)\) does not exceed twice its own standard error. Therefore we should not suspect any material difference in type between the two populations \(B\) and \(A\). But when the samples are increased to 400, as in samples \(b_2\) and \(a_2\), then the difference \((m_2-m_1)\) does exceed twice its own standard error, and we should be led to suspect a significant difference in composition between the two populations.

An alternative method of ascertaining whether or no two samples may be regarded as representative of distinct populations, is to apply a "test of independence" explained by Fisher (4, p. 77), in which the value of an "index of dispersion," \(\chi^2\) is calculated, and its significance noted from a table of "goodness of fit" (Fisher, p. 98). The reader is referred to Fisher for an explanation of these terms, and of the details.
of the test itself. One example, worked out in detail, will, however, illustrate the method:

<table>
<thead>
<tr>
<th></th>
<th>No. of Vertebras</th>
<th>No. of Fishes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample C</td>
<td>10, 40, 50</td>
<td>100</td>
</tr>
<tr>
<td>Sample D</td>
<td>15, 78, 27</td>
<td>120</td>
</tr>
<tr>
<td>Totals</td>
<td>25, 118, 77</td>
<td></td>
</tr>
</tbody>
</table>

No. of Vertebras:
- 57
- 56
- 55

No. of Fishes:
- 100
- 120

From Table III (Fisher, 4, p. 99) it is seen that the "chance" of such a discrepancy is less than one in a hundred; \( \chi^2 \) may thus be regarded as significant of a difference in type of population sampled.

**Table of Notations and Formulae.**

It has not been thought necessary to include in preceding paragraphs the proofs of the formulae given below. Such proofs can be found in standard textbooks. A recommended list of suitable works of reference are given on page 12 of the leaflet issued by the British Association on Biological Measurements (British Association, Section D, Leeds, 1927).

Standard Deviation \( \sigma = \pm \sqrt{\frac{\sum p \alpha^2}{(n-1)}} \)

where \( p = \) number of variates in the separate classes.
\( \alpha = \) deviation of a variate from the mean (m).
\( n = \) total number of variates.

Standard Error of the mean \( m \), calculated by \( \sigma_m = \pm \frac{\sigma}{\sqrt{n}} \)

Standard Error of the Standard Deviation, calculated by \( \sigma_\sigma = \frac{\sigma}{\sqrt{2n}} \)

Standard Error of a difference \((m_1 - m_2)\), calculated by \( \sqrt{\frac{\sigma_{m_1}^2}{(n_1-1)} + \frac{\sigma_{m_2}^2}{(n_2-1)}} \)
Biological Significance of Statistical Results.

The preceding paragraphs were devoted to a survey of the nature and collection of sample data, and, in them, an endeavour was made to explain the methods of treatment utilised. It is now necessary to consider how these raw statistical data serve to advance our knowledge of those vital events in the life-cycle of herrings upon which the great commercial fisheries are so completely dependent.

The annual life-cycle of an adult herring may be said to consist essentially of a movement inshore to a spawning ground, followed by an outward movement to a feeding ground. On the spawning ground the fish sheds its ova or milt, thereby playing its part in the determination of the morphological characters of a batch of larvae which in due course will themselves become adult. It is not yet known, however, to what extent the morphological characters of the larvae are due to heredity or to environmental influences. On a feeding ground the fish feeds and grows and records on its scale an index of the increase in length. There is no reason to suppose that the number of vertebrae of an adult fish, which was definitely fixed at a very early stage in its life, undergoes any subsequent change. Thus the scales of a fish will tell us what happened during its stay on a feeding ground, whereas the number of vertebrae in itself tells us nothing of its adult life, but gives us a clue to the combined effect of parental influence, and environmental influence soon after it was born. It will be appreciated, therefore, that by an organised scheme of intensive sampling carried out for a sufficient time, it should ultimately be possible from the study of scales to link up feeding grounds with spawning grounds, and from the study of the number of vertebrae to determine whether given feeding grounds and spawning places are visited by distinct herring stocks, as well as the extent to which this occurs. For instance, from the comparison of scale measurements of fishes spawning at Plymouth during the winter of 1925–26 with those of fishes taken during the feeding period in the deeper waters off the southwest coast of Cornwall in 1926, it would seem reasonable to suspect that some of the latter had come to Plymouth to spawn; on the other hand, the data on the average number of vertebrae for the same samples showed that in both these areas more than one type of stock occurred. The interpretation of these undoubtedly different types of stock as revealed by the study of such characters as the average of vertebrae, the number of keeled scales or bony fin-rays, presents many difficulties. Without the slightest doubt, the averages for any one of these characters often differ from sample to sample in the same sampling area, and from area to area, by amounts appreciably exceeding those which can be wholly attributed to errors due to sampling. In an endeavour to demonstrate
the difficulties of the problem presented, we will imagine an egg spawned by a herring having 56 vertebrae to be fertilised by milt from a male also having 56 vertebrae. What number of vertebrae will the resulting offspring possess? We will consider this question under two hypotheses:

**Hypothesis (1):**

That the average number of vertebrae is a character strictly inherited and thus unaffected by environmental conditions prevailing at the time of its fixation.

Under this hypothesis, the average number of vertebrae for our larval brood would invariably be 56-00.

**Hypothesis (2):**

That the average number of vertebrae is a character entirely decided by environmental conditions prevailing at the time of its fixation and thus not in the least dependent upon the parental number of vertebrae.

In this instance we should require to know the particular environmental conditions under which the vertebrae of the larvae were developed before we could state what the average number of vertebrae for the total brood would be.

Let us next enquire how, under these two hypotheses, we should interpret a marked difference between the average number of vertebrae for herrings in one area A and that for herrings in a second area B? Under hypothesis (1) it would seem quite reasonable to suspect that areas A and B were populated by different "genetic races" of herrings, each "race" carrying out an independent life-cycle within definite limits of distribution, and retaining its identity from generation to generation by the processes of heredity. This being so, herrings from area A, if transplanted to area B, would carry with them their distinct racial potentialities and react accordingly. Under hypothesis (2), however, it would be held that the observed differences in average number of vertebrae between the herrings in area A and those in area B were entirely due to the fact that the larvae developed their vertebrae under different environmental conditions. Fishes from area A if transplanted to area B would subsequently react to the new environmental conditions in precisely the same way as the normal inhabitants of the area did.

Now Johansen by comparing the values of the average number of vertebrae, keeled scales, and fin-rays for samples of herrings from different parts of the North Sea, and noting the magnitude of the differences, has created a considerable number of "races" which he regards as true races, the characters of which are due to heredity ([11], p. 90). He has expressed the view, however, that smaller differences, such as those between the "Scotch Summer Herring" and the "Dogger Bank Herring"
and between the "Scotch Spring Herring" and the "Icelandic Summer Herring" may not be of a hereditary nature. Johansen would thus seem to take a standpoint which is in the nature of a compromise between our hypotheses 1 and 2 above.

It seems to the present writer that until more definite knowledge of the respective parts played by heredity and environment in the establishment of the number of such characters as vertebrae, keeled scales, and fin-rays is gained, it is doubtful whether, in directing our sample analyses towards the immediate recognition of possible "races," we are following the method most likely to lead to a correct interpretation of observed sample differences. We are not without indications that environment can, and does, play its part in the establishment of the number of vertebrae, keeled scales and fin-rays.

Schnakenbeck (19, p. 192), by utilising the values of the average number of vertebrae for no less than twenty-eight of Johansen's races of herrings, has shown that the value increases from areas of low salinity to those of higher salinity. Schmidt (vide Johansen, 11, p. 89) has demonstrated by direct experiment that temperature and salinity materially affect the average value of the above morphological characters in Zoarces viviparus L., Lebistes reticulatus Regan., and Salmo trutta L. Johansen (11, p. 90) suggests that the herring is a less "plastic" species than those studied by Schmidt, but there is, as yet, no experimental proof of this.

In accepting appreciable differences between sample averages as evidence for differences in "racial" composition, we are, in practice, forced to conclude that racial admixture is of frequent occurrence (vide Part 2, p. 275). In a new area of observation the knowledge that such admixture may occur renders it most difficult to decide what races are actually represented. If, too, mixture of races, particularly if it occurs at spawning time, is a common feature, then it is not easy to understand how the distinctive characters of the different races are retained from one generation to the next.

As an alternative to the study of sample averages as "racial" characters, it is suggested that they should be examined as the defining morphological characters of temporary populations, regardless of "race," and used in conjunction with the data on age and growth from scale measurements as clues to the movements of shoals which brought about those temporary populations. Having established a working knowledge of these movements, the recognition of "races," if such are shown to exist, would follow as a natural sequence.

It is not proposed further to pursue this matter at this stage; the present object is to impress the case for the fullest possible study of the morphological differences which the statistical study of random samples has amply demonstrated. Sampling in a given area must be frequent
and continuous. Experience in the south-western area has shown most conclusively that it is very dangerous to attempt to draw conclusions as to the herring populations frequenting an area from the results of a few isolated samples. Furthermore, it becomes increasingly evident that one of the most desirable and urgent pieces of research is the study by experiment of the number of vertebrae in herrings reared from eggs of known parentage under controlled environmental conditions.

LITERATURE CITED.


