ARITHMETIC MEAN AND STANDARD DEVIATION

By Dr. Charles H. Blake
Analysis of Bird Data

There are some simple statistical techniques for studying the consistency and accuracy of observed data for bird weights, wing lengths, etc., when these data are recorded for a group of similar birds. The following is an example of these calculations:

In studying data, for example: bird weights (or other data) for numbers of birds (for instance: all of the birds of a given species banded during an Operation Recovery), we are interested in determining the AVERAGE value (often called ARITHMETIC MEAN). We also need to know how much each individual bird weight differs from this AVERAGE value. If we know the amounts of these DIFFERENCES (technically called DEVIATIONS), we can make some simple computations to determine the AVERAGE DEVIATION and the STANDARD DEVIATION and these computed values help us measure the dispersion of the bird weight values. These deviation computations help in pointing out values which are extremely high or low and which, therefore, should be considered abnormal. When we know the ARITHMETIC MEAN and the STANDARD DEVIATION, we can say that 67% of the observed values should normally differ from the ARITHMETIC MEAN by less than one STANDARD DEVIATION. (98% will normally differ from the mean by less than two STANDARD DEVIATIONS.)

The following demonstrates a specific example of these statistical computations and also demonstrates a technique for substituting an adjusted (coded) value for each observation in order to simplify the computations.

The ARITHMETIC MEAN (or AVERAGE) is, in most cases, the single value best representing the data. The STANDARD DEVIATION is a measure of the dispersion, or spread, of the data. The methods of computation here given are applicable to such quantities as weights and wing lengths. We shall establish certain basic symbols to develop some general procedures.

As to symbolism, we can call any value (coded (adjusted) or uncoded) "x", and the number of observations of the value, "f". The total number of observations is "N". The sum of any group of figures is indicated by the capital Greek sigma, "Σ". The ARITHMETIC MEAN is often marked "X" and the STANDARD DEVIATION, "S".

Let us take the weights of Swainson's thrushes at Hillsborough, N. C. To simplify matters, each weight is recorded only as its whole number of grams, discarding all fractions. To simplify computation, an adjusted (or coded) value is substituted for the original value. The adjusted (coded) value for 25 gms. = 1, 26 gms. = 2, 24 gms. = 0, etc. This adjustment is appropriate since there are no values below 25. Because there is one bird of excessive weight (46 gms.) we have to test whether it should be rejected from the series. This involves finding the average deviation.

Our preliminary mean is 1001/139 = 7.2. Let us round this to 7 and use it to get column 5. Note that the sign of the deviation is ignored in this procedure. The average deviation is 330/139 = 2.4. The deviation of the 46 gr. bird is 15 while 4 times 2.4 is only 9.6. Always be conservative in rejecting an observation but in this case we reject the 46 gr/bird.

The simple way to obtain Col. 6 is to multiply each entry in Col. 4 by the corresponding entry in Col. 2.

Since we have rejected the 46 gr. bird we obtain a final ARITHMETIC MEAN (AVERAGE value), $\Sigma X/N$ = 979/138 = 7.1. The decoded (unadjusted) mean is 31.1.

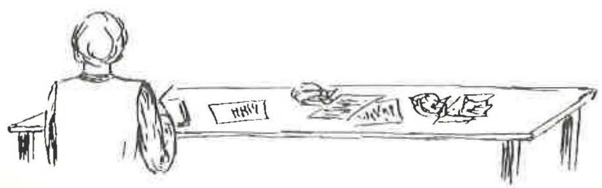
Now, $(\Sigma X/N)^2 = 50.4$ and $\Sigma X^2 f/N = 8271/138 = 60.00$. The difference is 9.6. The STANDARD DEVIATION is the square root of 9.6 or 3.1 g. We express the whole result as 31.1 = 3.1 g. Two-thirds of the observations should fall in the range 28 to 34 g. Actually 74% do so and we conclude that the data are a little more closely grouped about the mean than is called for by theory.

References: M. J. Moroney. Facts from Figures, Penguin Books, A236, 1951. Interesting and useful laymen's account of statistics.

F. E. Croxton, Elementary Statistics, Dover, S506, 1959. Particularly good programming of statistical methods with fully worked out examples.

Not Recommended: Arkin and Colton, Statistical Methods, Barnes & Noble - College Outline Series, 27, 1956 - (Intended to be used with a college textbook. Many methods given are incompetent or inapplicable.)

— Cont. on next page



Can you do a paper for the next Annual Meeting?

SWAINSON'S THRUSH DATA AND STATISTICAL COMPUTATIONS

Col 1	2*	3	4	5	6
wt gms	coded wt. (x)	no. birds (f)	(xf) (col 2 x col 3)	(deviation) x(f) (col 3 x deviation from 31 gms	(x^2f) (col 4 x col 2)
25	1	1	1	6	1
26	2	4	8	20	16
27	3	11	33	44	99
28	4	7	28	21	112
29	5	17	85	34	425
30	6	26	156	26	936
31	7	26	182	0	1274
32	8	11	88	11	704
33	9	10	90	20	810
34	10	5	50	15	500
35	11	5	55	20	605
36	12	6	72	30	864
37	13	3	39	18	507
38	14	2	28	14	392
39	15	1	15	8	225
40	16	2	32	18	512
41	17	1	17	10	289
Subtotal:		138 (N)	979 (Σx)		8271
46	22	1	22	15	
TOTAL:		139	1001	330	

Notes: *Coded weight is an adjusted value to simplify computation 24=0; 25=1; 26=2; ... 30=6, ... 46=22, etc. since no value is less than 25

Arithmetic Mean-Preliminary = $24 + \frac{1001}{139} = 24 + 7.2 = 31.2$

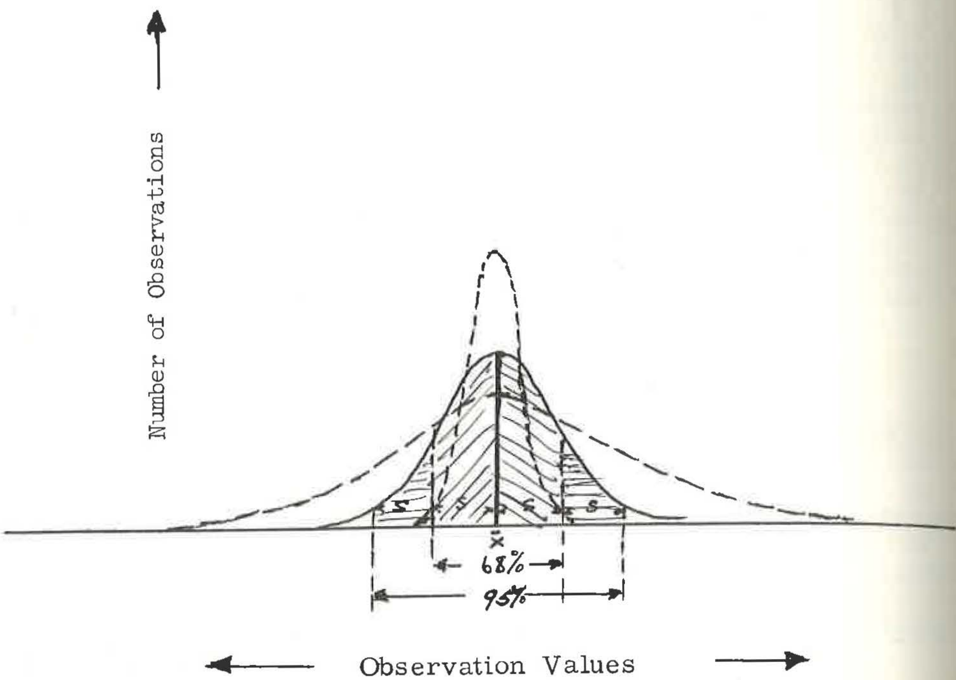
Average Deviation = $\frac{330}{139} = 2.4$

Arithmetic Mean-Final = $24 + \frac{979}{139} = 24 + 7.1 = 31.1$

Standard Deviation = square root of $\left(\frac{\Sigma x^2 f}{N} - \frac{\Sigma x}{N}\right) = \sqrt{\frac{8271}{138} = \frac{979}{138}}$

$= \sqrt{60.0 - 50.4} = \sqrt{9.6} = 3.1$

Diagram of typical curve enclosing an average distribution area. The diagram shows the arithmetic mean, and the percentage of the total area covered by one and two standard deviations.



——— Typical curve enclosing an area with average dispersion from the arithmetic mean.

— — — — Set of observations with the same arithmetic mean but more dispersion. (A larger standard deviation.)

- - - - - - Set of observations with same arithmetic mean but less dispersion. (A smaller standard deviation.)