to contain 12 dead martin nestlings and 21 that, though weak, had miraculously survived the period of near starvation. So, we proceeded to remove the dead birds from the compartments with a long wire hook. While I was busy with these tasks the temperature rose rapidly under the warm June sun. At long last insects started to fly, and the adult martins were in a frenzy as they swarmed into their compartments with precious food for their starving babies. So intent were they on getting food to their young quickly that they ignored me completely as I stood on the ladder examining nests within inches of where they were alighting to feed their nestlings. It appeared that the crisis was over. But during the next two hours four dead or dying adult male martins were found, two on the ground near the colony, one on the porch of the martin house, and one that evidently entered a compartment to feed its young and died there. Several other adult martins appeared to be on the verge of collapse and may have died. After at least two days of complete starvation and inactivity it would seem that the sudden burst of strenuous activity in getting food to the starving nestlings as quickly as possible was more than some of the birds could endure, so they simply collapsed and died of overexertion."

But there is a bright side. EBBA Members Connie Katholi and Anne Shreve have assured me that there are still Purple Martins. They know because they have banded each fall at the big Purple Martin roost at Charleston, West Virginia (approx. 140 air miles southwest of here). The martins from western Pa., probably stop at that roost each fall on their way to South America because Anne and Connie have recaptured 4 that were banded here at Clarksville, Pa. Three were direct recoveries and one of these was within 25 days.

This year the martins began congregating at the roost immediately after the rains stopped and they reported on July 1st that thousands of martins had just arrived and estimated 10,000 there two nights later - an unusually high number for that early date.

That 10,000 figure seems like a lot of Purple Martins but not when one considers the vast area they come from to converge at that roost. If an average of only 2 survived per colony of the estimated 150 colonies in my home county in southwestern Pennsylvania - that would be 300 Purple Martins when normally we would expect several thousand to leave this county each fall.

Some estimates run as high as 25 years before the martin population gets back to normal. My guess at present is from 5 to 10 years. EBBA member John Morgan (of Old Town, Maine) was here soon after the tragedy and he thinks the martins will make a quick comeback. I sincerely hope he is right!

--R.D. #1, Box 229, Clarksville, Pa. 15322

AN INTRODUCTION TO SOME STATISTICAL TERMS

By Mary Heimerdinger Clench

Several years ago I wrote a workshop paper on statistics ("Basic data interpretation for beginners," *EBBA NEWS* 33: 263-267, 1970) in which I used, but did not explain in detail, several statistical terms. In thinking over topics for this workshop issue, I decided it might be helpful now to discuss some of these terms. All are commonly encountered and are based on relatively simple concepts even if two are comparatively difficult to compute or to define mathematically. I won't concern myself with their rigorous aspects, but will try only to give EBBA NEWS readers some idea of what they mean. I hope that banders who are not mathematically inclined will be able to follow these explanations, thereby adding the terms to their working vocabulary. All of them deal with aspects of what statisticians call "frequency distributions"--how often and in what patterns certain phenomena such as measurements, numbers of birds, occur--the sort of things that banders most often analyze from their data. After reading this paper some banders may be stimulated to go on to learn exactly how to use these methods; but my primary aim is to get banders without a mathematical background over the language barrier, to know what is meant by a mean, a normal curve, and a standard deviation, so that when they see these terms used by others they will have a basic understanding of what the analyses are about.

Mean: A mean is the same thing you probably learned in grade school as an average. To calculate a mean you add up the numbers in a sample, say measurements of wing lengths, and divide by the number of individual birds measured. A simple example would be: 5 adult male American Goldfinches, measured in the spring, with wing lengths of 70, 71, 72, 72, and 73 mm; add these figures together (= 358) and divide by the number of birds in the sample (5) and you get a mean of 71.6 mm. Strictly speaking this kind of an average is called an arithmetic mean, but often is just called a "mean". Why isn't it called an "average"? Quite frankly, I'm not sure. Certainly if you look in statistics books you will find the word "average" discussed; you will also discover that there are many different kinds of averages, and that there are several different kinds of means as well. I suspect that mathematicians prefer to use "mean" ("arithmetic mean") because it has a precise definition in technical usage, whereas "average" is tainted with inexact usage in everyday language (as synonymous with typical, common, or ordinary.)

Normal Curve: This term has a large number of synonyms within statistics, including "bell-shaped curve," "Gaussian distribution," "Normal distribution," and even "Law of errors." Personally I prefer "bell-shaped curve" because that is what the curve looks like. "Normal," however, is the term in most general use so I shall use it here, spelling it with a capital N to remind you that it has a special meaning in statistics. The term Normal is unfortunate for it is not meant to imply that any of the numbers that form a Normal curve are, in fact, "normal" (with a small n) for whatever you may be measuring. A Normal curve is only the name given to a curve with a partucular shape (and derived from a particular equation.)

If you have measured a good many birds you already have the basic idea of what a Normal curve is like, whether you realize it or not. You know, for instance, that when you measure wings of a particular age or sex class of a species at any given season, most of the birds will give about the same measurement, but that there will also be a few individuals that are somewhat larger and a few that are somewhat smaller. You also know that one particular measurement (for example, 50 mm) may be more frequent than any other, that most of the other birds will have measurements close to the most frequent (*e.g.*, 47, 48, 49 and 51, 52, 53 mm) and that only a few will have particularly long or short wings (in this case, under 47 or over 53mm). If you then add up all your measurements and divide by the number of birds to calculate the mean wing length, and you find that the mean is the same as the most frequent wing length (the "mode," in this example 50mm), then you are on your way to saying that your measurements fit a Normal curve.

In simple terms, a Normal curve is a bell-shaped curve with its single peak in the center (at the mean), falling away symmetrically on the sides (see figure 1). The curve may be tall and narrow, low and wide, or anywhere in between. The important thing is that the peak is at the mean and that the decline away from the peak is similar on each side. If the decline is not similar (Normal) and the curve seems to "lean" somewhat to one side of the other, statisticians say that the curve is "skewed."

A skewed curve is one that is close to being Normal, but the mean falls to one side or the other of the single peak--the mean does not coincide with the mode (figure 2). Skewed curves are frequently found in biological measurements and can represent the normal (small n) condition. Take, for example, bird weights: any given species will have a minimum, fat free, weight; through most of the year most of the individuals will have a very small amount of fat and will weigh at least a little more than the minimum; during the migration period (if the species is migratory) the birds put on fat and consequently will weigh more, but how much more will depend on the individual, its physiological condition, the abundance and type of food available, etc. You know that when handling migrants some are fat, and a few seem packed almost to the bursting point. If you graphed the number of the various weights taken of a migratory species throughout the year, the most frequent (mode) might represent the migratory condition, but the right hand side of the curve would extend out to include the very fat individuals weighed during migration.

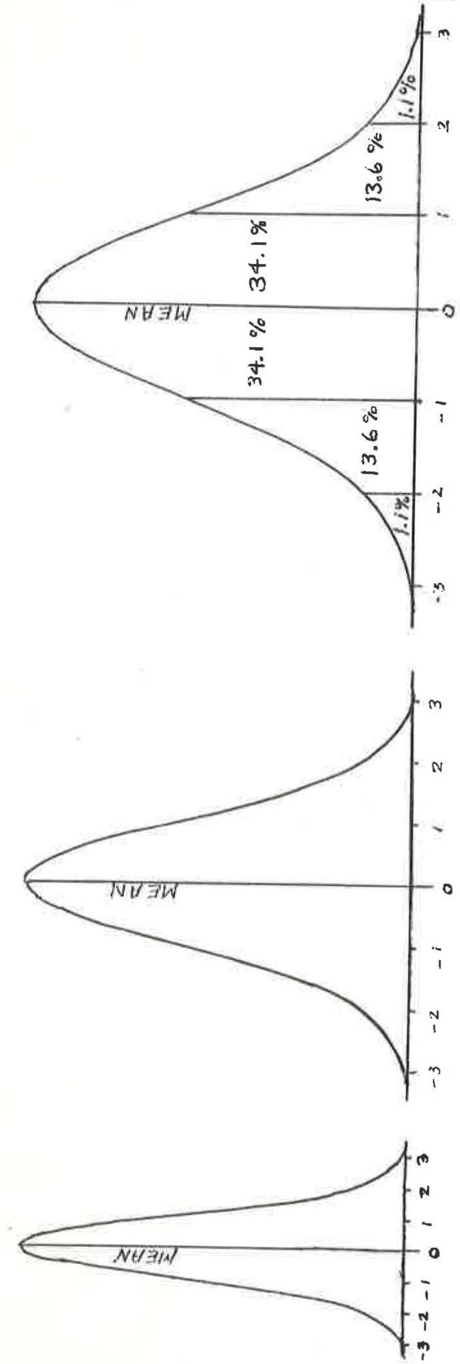

Figure 1. Normal Curves. The first of these curves has an arbitrary standard deviation value of 1, the second has a s.d. of 2, and the third has a s.d. of 3, to demonstrate the effect of the size of the standard deviation on the width (shape) of the curve. The third figure is divided into six units, from -3 standard deviations to +3 standard deviations, to show the percentage of the range encompassed by each standard deviation.

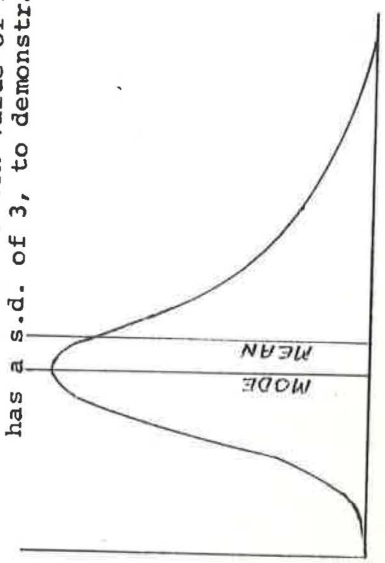

Figure 2. *(left)* A skewed Curve. Note that the mean does not fall on the same value as the mode.

These heavy migrants would serve to increase the mean weight for the year and thus might "skew" the curve so that the mean lies away from the mode, toward the heavy end.

In everyday use, the terms "Normal" and "skewed" are helpful because they are shorthand ways of describing data. There is a lot of information in the phrases "the wing lengths have a mean of 50mm and they form a Normal curve" or "they have a mean of 50mm but the curve is skewed to the right"; this allows the listener to form a mental image of what the curve looks like without having actually to see all the data and to draw the curve. By its definition, then, a Normal curve has a particular configuration--a single peak at the mean and symmetrically sloping sides. But simply saying that a curve is Normal is not enough. To allow the listener to form a more exact picture of the curve in his mind he must know how tall and wide the curve is around the mean. That is accomplished by giving the standard deviation.

Standard deviation: This term also has a few synonyms (such as "root-mean-square deviation") but more important to know are its various expressions in statistical shorthand, commonly S.D., s.d., σ (the small Greek letter sigma) or s (anglicized from the Greek), and sometimes as a plus and minus sign ($\pm$) before a number. Standard deviations can be rather tedious, if not difficult, to compute, and as I am trying to present the "field marks" of these statistical terms rather than a textbook lesson on how to compute them, I shall not try to derive a standard deviation here. You can find that in any statistics book. What is of interest to the non-statistician bander is to know that a standard deviation tells you how much a set of data varies around a mean.

As a rough rule, but one that works well in most curves that have only a single peak and are reasonably symmetrical on the sides (whether or not they are strictly Normal), one standard deviation will encompass approximately two-thirds of the observations (measurements) on either side of the mean, two standard deviations will take in about 95% of the observations, and three standard deviations will take in over 99% (figure 1). Thus if you know that a curve is Normal, with a mean at 50mm and a standard deviation of 2.3 mm, you can add and subtract the standard deviation to/from the mean and know that about 2/3 of the measurements will lie between 47.7 and 52.3 mm, that about 95% will lie between the mean plus or minus two standard deviations (45.4 and 54.6mm), and that essentially all of them will lie between the mean plus or minus three standard deviations (43.1 and 56.9mm). This, then, gives you a curve that you can visualize: it is rather tall and thin, Normal shape, with a spread of some 14mm from about 43-to-57mm. If, on the other hand, the Normal curve had a mean of 50mm but a standard deviation of 7.1 mm, you would picture it as rather flat and broad, extending from 3 S.D. on the left (3 x 7.1 = 21.3, subtracted

from 50 = 28.7mm) to 3 S.D. on the right (21.3 + 50 = 71.3mm), a spread of 42.6 mm.

You can see the usefulness of these terms: Normal tells you that the curve has a particular shape, the mean tells you where the curve's peak is located, and the standard deviation tells you how far on the base of the graph the curve will be spread and consequently (given the sample size) how high the peak will be.

To try a specific example: in a sample of 710 AHY-Male American Goldfinches banded at Carnegie Museum's Powdermill Nature Reserve during April, May, and June, the arithmetic mean of the wing length is 72.12mm, the curve is essentially Normal, and the standard deviation is 1.74 mm. First visualize what the graph should look like, then look at figure 3 to see if you were right. I have shown these data in the figure both as a bar graph (as you would be able to draw it) and as adjusted to the best fitting Normal curve. This is to indicate that although you may not have the statistical training to draw a Normal curve derived from your data, you can draw a bar graph and see from its shape (and calculation of the mean) whether or not it would form an approximately Normal curve. In actual fact very few sets of data taken from the natural world will form a precisely Normal curve, but a great many will form a curve that is close enough to the mathematical ideal that one can call them Normal for practical purposes.

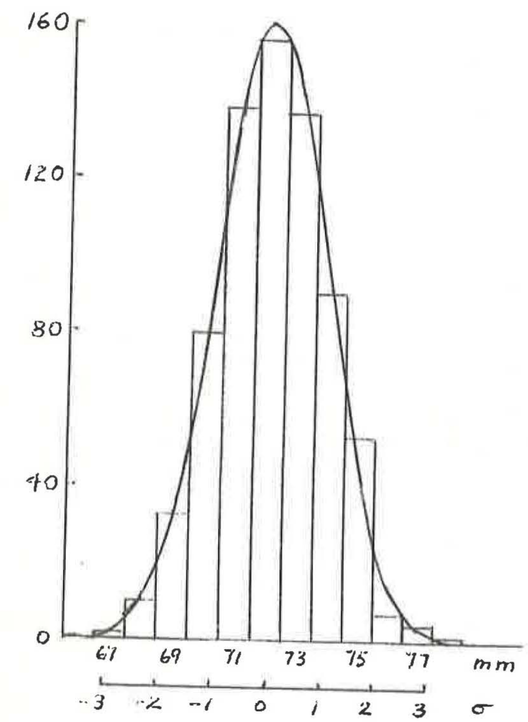

Figure 3. Wing lengths of AHY-M American Goldfinches banded at Powdermill Nature Reserve during Apr.-May-June. Sample size = 710;s.d. = 1.74mm;mean = 72.12mm. A Normal curve, fitted to these data, has been superimposed.

Of what use is all this? Primarily, knowing certain characteristics about your data, such as whether they form a Normal curve or not, and especially knowing the standard deviations (hence the usual variation) of the things you are measuring, will allow you to make predictions about related observations. For instance, having calculated the S.D. of Powdermill spring adult male goldfinches, we can predict that over 99% of the spring adult male goldfinches we band in the future will have wing lengths that range from 67 to 77mm. If, on the other hand, we caught a spring male goldfinch that measured only 65 mm we would know precisely how much smaller it was then normal; we would also inspect the bird with particular care, knowing how unusual it was, looking for some special explanation of its small size.

There are also kinds of data analyses that might enable you to predict "fill-in" data you do not have. For example, if you graphed the fall migration pattern of a species at your banding station and found you had a curve like figure 4, you would see that you had most of what looked like a Normal curve; you might then be able to extend the curve to the end by prediction even though you had closed your banding station before that migration had ended and hence did not actually have the data to complete the curve.

Conversely, when your results don't fit your predictions based on previous analyses, this too can be interesting. For instance, if you catch a group of birds that do not fall within the previously calculated curve for that age, sex, and season for that species, you would know that something had changed--perhaps that group of birds came from a different population than you normally sample (particularly suspect in migrants; unusual weather patterns can bring in birds from a population that does not usually pass through your area.)

Finally, a word of warning. I have given here very simple explanations of what Normal curves, means, and standard deviations are; what they look like; and an indication of their usefulness. This allows you to make some use of the terms yourself, but it particularly allows you to understand (at least partially) what a statistician means when he uses them. I have not, however, gone into how these measures are derived and how complicated they may be to understand within the rigors of statistical theory. Let me just give you the formula for the Normal distribution to indicate the iceberg of mathematics that this simple-looking curve sits on top of:

$$y=\frac{1}{\sigma\sqrt{2\pi}}\,e^{\frac{-(x-\bar{x})^2}{2\sigma^2}}$$

So don't try to prove that a curve is Normal, or calculate a standard deviation without learning exactly what these terms mean and what they are based on. I recommend Facts from Figures by M.J. Moroney (my copy is dated 1957), a paperback Pelican Book (Number A236), Pelican Books Inc., New York, or Quantitative Zoology by G.G. Simpson, A. Roe, and R.C. Lewontin (my copy is dated 1960), Harcourt, Brace and Co., New York, for thorough explanations of these terms.

I thank Miriam Stern for extracting the goldfinch data from our Powdermill banding files, and Robert C. Leberman and A.C. Lloyd for their diligent banding at the Reserve that produced the large sample of wing measurements. Harry K. Clench and Kenneth C. Parkes kindly read drafts of the manuscript and offered several helpful comments and suggestions.

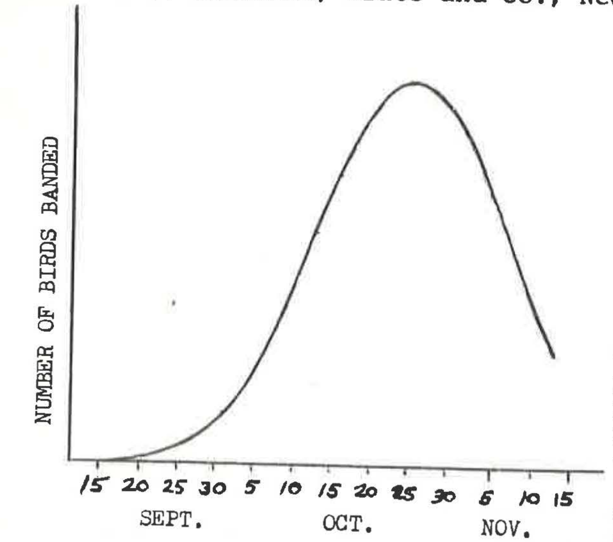

Figure 4. Hypothetical curve showing the number of individuals of a species banded during a fall season. The last banding day was November 13. If this migration pattern follows the Normal curve to its end, one would expect to band the last of the species in early December.

-- *Carnegie Museum, Pittsburgh, Pa. 15213.*