A STATISTICAL NOTE ON OVERLAP IN MEASUREMENTS

By Jack P. Hailman

Actual published figures sometimes provide a more interesting immediate problem for introducing a statistical point in banding than would any consideration of abstract principles. Robertson's (EBBA News, 33(2): 79-80, 1970) comments on wing measurements of accipiters provides just such a problem. His tables of ranges show no overlap among the three species considered, and no overlap between the two sexes of a given species (with exceptions to the latter discussed in the text). The conclusion most banders would draw is that any accipiter in the hand can be identified unambiguously as to species and sex by wing measurement along. However, this conclusion is not necessarily valid from the data presented.

The problem is that the range in a series of measurements is not a good representative statistic for the actual variation that occurs. There are many "best estimates" of population variation that may be calcualated from a sample of measurements, but the most common one is the standard deviation(s). Three standard deviations to either side of the mean value ($\bar{x}$) embrace more than 99% of the values actually occurring in the population. Therefore, a convenient way of checking the best estimate of whether the hawk wing measurements really overlap would be to tabulate $\bar{x}$ -3s and $\bar{x}$ +3s. Only the person having the actual set of measurements can calculate s.

As a rapid check on the data presented, however, we can use the range (R) to estimate the sample standard deviation (s) according to the method presented by Natrella (Experimental Statistics, U.S. Dept. of Commerce, 1963, pp. 2-6 to 2-7). The range is defined as the difference between the highest and lowest value. The range is then divided by a tabulated factor d_n which is very nearly the square-root of the sample size (n), when n lies between 3 and 10. Thus, our estimate of s (s_e) is calculated by: $s_e = R/d_n$. As the sample size (n) becomes larger, the range (R) becomes a very inefficient estimator of the standard deviation.

(Another quick method that gives somewhat different results was suggested by Mosteller and Bush (Handbook of Social Psychology, Chapter 8, p. 323, 1954). For sample sizes up to n = 15, they suggest using $s_e = R/n$ for rough calculations.)

We can make a very crude estimate of the overlap by using the Natrella values for d_n up to n = 12, and the square root of n for n greater than 12, recalling that the latter is a poor estimate . These rough calculations, shown in the accompanying table, were made with a slide rule and may thus lack accuracy in the third digit; nevertheless they illustrate the statistical point as well as giving a rough check on the actual example of accipiter wing lengths.

Note first that the table shows that the span of expected values is in every case larger than the range of measurements given by Robertson. Second, these crude calculations support the conclusion that there is no overlap in measurement among the three species, Lastly, the calculations fail to support the distinction by sex: we expect both the Cooper's Hawk and the Goshawk to have some males that are larger than some females. Interestingly, these are the two species for which Robertson reports individuals having wing lengths between his ranges.

The reliability of wing lengths for judging the species and sex of an accipiter is still an open question that can be decided ultimately only on the more accurate calculations of the actual sample standard deviations.

(Department of Zoology, University of Wisconsin, Madison, Wis. 53706.)

TABLE -- WING MEASUREMENTS IN ACCIPITERS

Species	Sex	n	R	d_n*	$s_e = R/d_n$	$\bar{x} - 3s_e$	$\bar{x} + 3s_e$
Sharp-shinned	M	12	9	3.258	2.762	162.3	178.9
	F	37	16	6.082	2.628	194.6	210.4
Cooper's	M	7	17	2.704	6.280	218.8	256.5
	F	17	21	4.121	5.100	254.0	284.6
Goshawk	M	29	24	5.382	4.459	304.1	330.9
	F	20	24	4.472	5.365	329.3	361.5

* $d_n = \sqrt{n}$ for $n > 12$