UNKNOWNS in each characteristic group showed the same general trend. In other words, on the basis of these graphs, it looked as though there was a good possibility that the UNKNOWNS were no different from the KNOWNS, at least in the aggregate. It was decided to investigate this by the use of a statistical procedure called the "Chi Square Test".

The Chi Square Test is a statistical test of the likelihood that two distributions come from the same population, that is, it gives the probability that there is no difference in the make-up of the two distributions being measured.

The method is outlined as follows:

(1) Adjust the distributions by multiplying the KNOWNS in each characteristic group by the ratio of the total number of UNKNOWNS to the total number of KNOWNS. (The Chi Square Test is applicable only to distributions which have the same total number of elements.)

(2) Take the difference between the number of UNKNOWNS and the adjusted number of KNOWNS in each characteristic group.

(3) Square the remainder from Step 2.

(4) Divide the result of Step 3 by the corresponding number of adjusted KNOWNS.

This is the chi square for the particular group. Summing the individual chi squares over the groups of a characteristic gives the chi square for that characteristic. This number is then compared with a table of the distribution of chi square which can be found in many texts on elementary statistics.

It will be noted that chi square is tabulated in terms of degrees of freedom which in this case is one less than the number of groups of sightings for each characteristic.

The tabulations of KNOWNS and UNKNOWNS against the six characteristics and the Chi Square Test as it was applied are shown in Tables II through VII. In each case, the number of degrees of freedom is given, as is the value of chi squares corresponding to probabilities of 5 per cent and 1 per cent that two distributions with this number of degrees of freedom come from the same population. Since the greater the value of chi square the smaller the probability of homogeneity of two distributions, a calculated value of chi square greater than either the 5 per cent or 1 per cent values will indicate a probability less than 5 per cent or 1 per cent, respectively, that the two distributions are homogeneous. The term homogeneity is used here to indicate that two distributions could have come from the same population.