ON INDICES OF DISPERSION

By Paul G. Hoel

University of California, Los Angeles

1. Introduction. In biological sciences the index of dispersion for the binomial and Poisson distributions is very useful for testing homogeneity of certain types of data. For example, the dilution technique in making blood counts finds it useful. Recently there have been attempts to use it to determine allergies by observing the change in the blood count after allergic foods have been taken. Here the sample may consist of only a few readings; consequently it is important to know how accurate this index is when applied to small samples. After inspecting the application of the Poisson index to such counts, I was surprised to see the lack of agreement with theory. At first it appeared that the fault lay with the chi-square approximation which is used on this index, but later it was clear that the assumption of a basic Poisson distribution was at fault. It now appears that statisticians will need to be careful about citing blood counts as examples of data following a Poisson distribution.

This paper is the result of investigating the accuracy of the chi-square approximation for the distribution of these indices. Previous work on this problem seems to have consisted in some sampling experiments [1] for small values of the parameters involved, and in some theoretical work [2] in which the sampling distribution is considered only for a fixed sample mean. Although sampling distributions ordinarily differ very little from the distributions obtained by assuming the mean of the sample fixed, for small degrees of freedom the difference may be appreciable and therefore requires investigation. In this paper the accuracy of the chi-square approximation is investigated by finding expressions for the descriptive moments of the distribution which are correct to terms of order N^{-3}. These expressions are obtained by means of Fisher's semi-invariant technique.

2. Moments of the distribution. Employing Fisher's notation [3], let the binomial index of dispersion be denoted by z, then z may be written as:

$$z = \frac{\Sigma(x-\bar{x})^2}{\bar{x}\left(1-\frac{\bar{x}}{n}\right)} = \frac{(N-1)k_2}{k_1\left(1-\frac{k_1}{n}\right)} = \frac{N-1}{\kappa_1\left(1-\frac{\kappa_1}{n}\right)}\,\frac{k_2}{\left(1+\frac{k_1-\kappa_1}{\kappa_1}\right)\left(1-\frac{k_1-\kappa_1}{n-\kappa_1}\right)}.$$

Letting $w = k_1 - \kappa_1$, $y = k_2$, $a = n - \kappa_1$, $b = \dfrac{N-1}{\kappa_1\left(1-\frac{\kappa_1}{n}\right)}$, z may be ex-