

**THE VARIANCE OF THE PROPORTIONS OF SAMPLES FALLING  
WITHIN A FIXED INTERVAL FOR A  
NORMAL POPULATION**

BY G. A. BAKER

*University of California, Davis*

Suppose that we have a normal population

$$(1) \quad y = \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ -\frac{(x - m)^2}{2\sigma^2} \right\}$$

and we draw samples of  $N$  from this population. We wish to estimate the proportion,  $p$ , of the population between two fixed limits,  $m + \lambda\sigma$  and  $m + \mu\sigma$ . One way to make this estimate is simply to count the number of observed  $x$ 's which fall in this interval. We shall denote this number by  $n$ . Then the ratio

$$(2) \quad n/N$$

is an estimate of  $p$ . If this is done the variance of  $p$  is well known to be

$$(3) \quad \frac{p(1 - p)}{N}.$$

The method of estimating  $p$  by counting the number in a definite interval is nonparametric and requires no assumption of normal or other specified type of sampled population for validity. However, if we know that the sampled population is normal then we may make use of this knowledge in estimating  $p$  and possibly obtain an improved estimate.

Another way to estimate  $p$  which makes use of the form of the sampled population is to compute

$$(4) \quad \begin{aligned} \bar{x} &= \frac{1}{N} \sum_{i=1}^N x_i \\ s^2 &= \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2, \end{aligned}$$

and hence the integral

$$(5) \quad \int_{m+\lambda\sigma}^{m+\mu\sigma} \frac{e^{-(x-\bar{x})^2/(2s^2)}}{s \sqrt{2\pi}} dx.$$

It is implied in elementary texts that (5) is a better estimate of  $p$  than is (2) although this point is not discussed.

It is the purpose of the present note to discuss the variance of the estimate (5) and compare this variance with (3).

Now (5) is a function of the first two moments of the sample and it follows from an application of a theorem stated by H. Cramér [1] that (5) is asymptotically normal with mean  $p$  and variance given by