A COMPARISON OF THREE METHODS OF FITTING THE NORMAL OGIVE*

ELLiot M. CRAMER†

THE JOHNS HOPKINS UNIVERSITY

The Mueller-Urban method of fitting the normal ogive is derived, and the inadequacies of its inherent assumptions are discussed. This and the unweighted least squares method are compared to the maximum likelihood solution which is shown to be very close to the "ideal" least squares solution. As an empirical demonstration of the superiority of the maximum likelihood solution, random ogives are fitted by all three methods and they are compared on the basis of the expected values and the standard errors of the estimates. It is concluded that the maximum likelihood solution is uniformly superior to the others in all respects.

The normal ogive (normal integral) has been commonly used as a model for various types of psychological data for many years. Several methods have been proposed for fitting this function [7], but the conditions under which these methods are appropriate have not been precisely stated. The purpose of this paper is to indicate the nature of the approximations involved in the two most commonly used methods and to compare them with an exact procedure which has been little used in psychological work.

In a typical experiment observations are taken at each of several points on a continuum x, such that each observation represents a choice between two classes. The proportion of responses in one of the classes is plotted against x, and it is this function that is assumed to be, within sampling error, the normal ogive. Precisely, the relation is

\[ p = \int_{-\infty}^{x} Z(m, \sigma) \, dv = \int_{-\infty}^{x} Z(0, 1) \, dv, \]

where

\[ Z(m, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left[ -\frac{(v - m)^2}{2\sigma^2} \right], \]

*This research was done under Contract Nonr-248(55) between the Office of Naval Research and The Johns Hopkins University. This is Report No. 18 under that contract. Reproduction in whole or in part is permitted for any purpose of the United States Government. This paper is part of a dissertation submitted to The Johns Hopkins University. Part of this work was done while the author was a National Institutes of Health Research Fellow.

†Now at The Biometric Laboratory, The George Washington University. The author is indebted to Dr. Wendell R. Garner for his valuable advice and encouragement, and to Jerome Cornfield for several helpful discussions.
$y$ is the normal deviate $(x - m)/\sigma$ corresponding to $p$, and where $m$ and $\sigma$ are the mean and standard deviation of the normal curve. The parameters to be estimated are $m$ and $\sigma$, given the observed proportions $p$.

**Methods of Estimating $m$ and $\sigma$**

**Unweighted Least Squares**

The simplest method for obtaining estimates is to use the unweighted least squares procedure, where the basic data are taken as the normal deviates corresponding to the observed proportions. We can write

\[ y = (x - m)/\sigma = a + bx. \]

This is a linear relation between $y$ and $x$, where $a = -m/\sigma$ and $b = 1/\sigma$. A least squares line may be then computed in the ordinary manner to obtain estimates $m$ and $\sigma$. The function minimized is

\[ S = \sum (y - \bar{y})^2. \]

This is an unweighted least squares fit to the observed normal deviates. ($p$, $\hat{p}$, and $\bar{p}$ indicate, respectively, the observed, estimated, and true values of the proportion $p$. A similar convention is used for functions of $p$, such as $y$ and $Z$.)

Two objections have been noted to the above procedure. First, the reliability of $\bar{p}$ varies as a function of $\bar{p}$ and more weight should obviously be given to the more reliable points; second, this method minimizes the squared deviations of the observed from the predicted normal deviates, rather than the squared deviations of the proportions as would be desired.

**Mueller-Urban**

The Mueller-Urban solution (or weighted least squares solution) has been frequently used because of the above noted deficiencies of the unweighted least squares procedure. The Urban weight $n/p(1 - p)$ is the reciprocal of an estimate of the variance of the true proportion $\bar{p}$ and serves to equalize the deviations of the proportions with respect to their reliabilities. The Mueller weight, $Z^2$, the squared normal ordinate corresponding to the observed proportion $p$, serves to account partially for the transformation from proportions to normal deviates. The combined weight $nZ^2/p(1 - p)$ is used to obtain a weighted least squares solution. The function minimized is

\[ S = \sum \frac{nZ^2}{p(1 - p)} (y - \bar{y})^2. \]

One may think of a true Mueller-Urban weight as being determined from the true proportion $\bar{p}$. Such a weight is between 0.34 and 0.64 when $\bar{p}$ is between 0.1 and 0.9. For proportions of 0.01 and 0.001 it is 0.08 and 0.0004.