FITTING GENERAL GRAM-CHARLIER SERIES

By Paul A. Samuelson

Massachusetts Institute of Technology

1. Introduction. Since the last part of the nineteenth century at least, it has been common to represent a probability distribution by means of a linear sum of terms consisting of a parent function and its successive derivatives. Usually the parent function is the Type A or normal curve, as discussed by Gram [1], Bruns [2], Charlier [3], and numerous others. In addition there have been generalizations in various directions: for example, the Type B expansion in terms of the Poisson parent function and its successive finite differences.

Unlike these two types, which have a definite probability interpretation, another generalization involves the use of other parent functions and their derivatives (or differences) to give an approximate representation of a given frequency curve. With this process is associated the names of Charlier, Carver [4], Roa [5], and many others. Two general methods by which the equating of moments of the fitted curve and the given distribution yield the appropriate coefficients have been given by Charlier and Carver respectively. An account of the latter's technique is more accessible to the average English speaking statistician.

It is the purpose of the present discussion to indicate how the Charlier method may be simplified, and can be used to replace the Carver method. In doing so, I am following up the oral suggestion made some years ago by Professor E. B. Wilson of Harvard, that repeated integration by parts will yield the requisite coefficients very simply. At the same time certain methods implicit in the work of Dr. A. C. Aitken [6] show how the use of a moment generating function can often lighten the algebraic analysis. There will also be a brief indication of analogous results for general finite difference parent families; and attention will be called to a troublesome historical blunder which has permeated the statistical literature.

2. Alternative methods. Avoiding the overburdened expression generating function, I shall consider parent functions, called $f(x)$, with the restrictive properties:

a) Moments of all order of $f(x)$ exist.

b) Derivatives of any required order exist with appropriate continuity.

c) There exist high order contact at the extremities of the distribution as defined below.

Mathematically,

a) $\int_{-\infty}^{\infty} x^k f(x)dx$ is finite for all positive integral values of k

and

c) $\lim\limits_{x\to\pm\infty} x^j f^k(x) = 0$ for all positive integral values of j and k.