ABSTRACTS OF PAPERS

I. Presented on December 27, 1941, at the New York Meeting of the Institute

A Generalized Analysis of Variance. FRANKLIN E. SATTERTHWAITE, University of Iowa and Aetna Life Insurance Company.

This paper examines the fundamental principals underlying designs for the analysis of variance. Given several statistics of the type, $\chi_j^2 = \Sigma_i \theta_{ij}^2$ where the θ's are arbitrary orthogonalized linear functions of certain underlying normal data, x_k; a rule is set up for determining a set of m_k as linear functions of the x_k such that $\chi_0^2 = \Sigma(x_k - m_k)^2$ will be independent of the remaining χ_j^2's. Further it is shown that simultaneously with the above, the x's and the θ's may be subjected to certain types of linear restrictions (for the purpose of estimating parameters or otherwise) without disturbing the distributions or the independence relations except for the appropriate reduction in degrees of freedom. The rule used to determine the m's gives results consistent with the standard designs for the analysis of variance. However, it goes further in that one may use weighted rather than simple averages in setting up his design. A practical application of this is the two way analysis of data which are averages and lack homogeniety of variance through constants of proportionality between the variances are known. The two way analysis of incomplete data is another practical problem which is solved by the simple expedient of a zero weight. The use of weighted averages frequently introduces difficulties in estimating parameters, particularly the mean. The combination of the linear restriction concept with standard analysis of variance methods solves this difficulty.

On the Power Function of the Analysis of Variance Test. ABRAHAM WALD, Columbia University.

It is known that the power function of the analysis of variance test depends only on a single parameter, say λ, where λ is a certain function of the parameters involved in the distribution of the sample observations. Let Z be any critical region (subset of the sample space) whose size does not depend on unknown parameters, i.e., it has the same size for all values of the parameters which are compatible with the hypothesis to be tested. It is shown that for any positive c the average power (a certain weighted integral of the power function) of the region Z over the surface $\lambda = c$ cannot exceed the power of the analysis of variance test on the surface $\lambda = c$ (the power of the latter test is constant on the surface $\lambda = c$). P. S. Hsu's result, *Biometrika*, January, 1941, pp. 62–68, follows from this as a corollary.

Definition of the Probable Error. E. J. GUMBEL, The New School for Social Research.

The probable error is usually defined either as the semi-interquartile range or as $\frac{2}{3}$ of the standard error. We define it as half of the smallest interval that has the probability $\frac{1}{2}$. For distributions which never increase (decrease), the beginning (end) of this interval is the origin (the median), and the end is the median (the end of the distribution). In general the probable error ρ is the solution of the equations $W(\xi + \rho) - W(\xi - \rho) = \frac{1}{2}$ and $w(\xi + \rho) = w(\xi - \rho)$ where ξ denotes the midpoint of the interval. For symmetrical distributions the first definition remains valid. For the Gaussian distribution the second definition holds besides. The numerical values for the midpoint ξ and the probable error ρ are given for some distributions usual in statistics. The calculation of the standard error of the probable error, which depends upon the distribution $w(x)$, determines whether the probable error is more or less precise than the standard error. For the asymmetrical exponential