The Annals of Statistics
1975, Vol. 3, No. 3, 573-580

NECESSARY AND SUFFICIENT CONDITIONS FOR ASYMPTOTIC JOINT NORMALITY OF A STATISTIC AND ITS SUBSAMPLE VALUES

BY J. A. HARTIGAN

Yale University

1. Introduction. If $X_1, X_2, \cdots, X_n$ are independent identically distributed random variables with mean μ and variance σ^2, then the mean $\bar{X} = \sum_{j=1}^{n} X_j/n$ is asymptotically normal with mean μ and variance σ^2/n. Surprisingly, asymptotic normality also holds for such diverse statistics as order statistics, correlations, maximum likelihood estimates and Bayes estimates, and eigenvalues. This paper gives necessary and sufficient conditions that sequences of subsample values of a statistic be asymptotically joint normal. Also, the following generalization of the central limit theorem is proved:

Let $t_n(X_1, \cdots, X_n)$ be a sequence of symmetric measurable functions in $X_1, \cdots, X_n$, and suppose $n \operatorname{Cov}(t_n, t_m) \to \sigma^2$ whenever $n \geqq m \to \infty$. Then $n^{\frac{1}{2}}(t_n - Et_n)$, $m^{\frac{1}{2}}(t_m - Et_m)$ are asymptotically joint normal with variances σ^2 and correlation ρ whenever $m, n \to \infty$, $m/n \to \rho^2$, $0 \leqq \rho^2 \leqq 1$. The mean satisfies the conditions of the theorem since $n \operatorname{Cov}(\bar{X}_n, \bar{X}_m) = \sigma^2$ exactly.

The property of the mean which compels the normal limit is

$$n^{\frac{1}{2}}\bar{X}_{1,n} = (m/n)^{\frac{1}{2}}m^{\frac{1}{2}}\bar{X}_{1,m} + [(n-m)/n]^{\frac{1}{2}}(n-m)^{\frac{1}{2}}\bar{X}_{m+1,n},$$

where $\bar{X}_{r,s}$ denotes the mean of $X_r, X_{r+1}, \cdots, X_s$. Thus if the mean is to have a limiting distribution G after standardization, and if Y_1 and Y_2 are independently distributed as G, then $\alpha_1 Y_1 + \alpha_2 Y_2$ must have the distribution G after standardization. Of course this property defines the stable laws, of which only the normal has finite variance.

Generalizing this, a "mean-like" sequence of statistics t_n satisfies $n^{\frac{1}{2}}[t_{1,n} - (m/n)t_{1,m} - (n-m)/nt_{m+1,n}] \to 0$ as m, $n - m$, $n \to \infty$. This condition will ensure that t_n is asymptotically normal if $n^{\frac{1}{2}}(t_n - a_n)$ converges to a distribution with finite variance. The mean-like property is implied for $t_n - Et_n$ by the above condition $n \operatorname{Cov}(t_n, t_m) \to \sigma^2$ as $n \geqq m \to \infty$, provided that t_n is a symmetric function of the observations. To handle asymmetric functions, it is necessary to consider behavior of the statistic as a function of subsets $X_{i_1}, X_{i_2}, \cdots, X_{i_n}$. These subsets appear in the three conditions for centrality of a statistic, which are necessary and sufficient for joint asymptotic normality of sequences of statistics defined on the subsets, proved in Theorem 1, Section 2.

Theorem 3 presents a simpler set of sufficient conditions when nt_n^2 is uniformly integrable, and Theorem 4 is the generalization of the central limit theorem mentioned above.

Received November 1973; revised October 1974.