The Annals of Statistics,
1983, Vol. 11, No. 4, 1100–1108

THE EFFECT OF DEPENDENCE ON CHI-SQUARED AND EMPIRIC DISTRIBUTION TESTS OF FIT[1]

By Leon J. Gleser and David S. Moore

Purdue University

Suppose that a test of fit to a parametric family of distributions is employed, with critical points determined from the limiting null distribution of the test statistic for IID observations. It is shown that if the observations are in fact a stationary process satisfying a positive dependence condition, the test will reject a true null hypothesis too often. This result is established for a broad class of chi squared and empiric df tests, including the Pearson, Kolmogorov-Smirnov and Cramér-von Mises tests with general estimators of unknown parameters. Furthermore, the method of proof is sufficiently general to apply also to other classes of tests. Confounding of positive dependence with lack of fit is therefore a general phenomenon in the use of omnibus tests of fit.

1. Introduction. In testing the fit of a sequence of observations to a parametric family of distributions, it is commonly assumed that the observations are independent and identically distributed (IID). In practice, however, the observations may have substantial dependence, as when the data are collected as a time series. Suppose, then, that $X_1, \cdots, X_n$ are observations on a (strictly) stationary stochastic process (SSP) and that G is the common univariate df of the X_i. A statistician who believes that the X_i are IID tests the hypothesis that G is a member of a parametric family $\{F(\cdot, \theta): \theta \text{ in } \Omega\}$, for Ω an open set in Euclidean m-space R^m. We will show that when $G = F(\cdot, \theta_0)$ for some θ_0, and the SSP satisfies a positive dependency condition, chi-squared and empiric distribution function (EDF) tests reject the true null hypothesis too often. That is, *positive dependence is confounded with lack of fit.* Since the class of tests for which this result holds is very broad, including the Pearson, Kolmogorov-Smirnov, and Crámer-von Mises tests with the parameter θ estimated in general ways, this confounding deserves recognition as a general phenomenon in applying omnibus tests of fit.

Chanda (1981) and, more generally, Moore (1982) have independently studied the limiting distribution of chi-squared statistics when the data are dependent, with emphasis on obtaining the form of the limiting covariance matrix of the standardized cell frequencies. Moore also proves the confounding of positive dependence with lack of fit in one case, that of testing the fit of a general Gaussian SSP to a specified normal distribution. The positivity condition that we impose on the bivariate distributions of (X_i, X_j) arises naturally from consideration of the covariance matrix of cell frequencies when the data form a SSP. We do not, however, make use of the detailed form of the covariance matrix. Some specific examples of such matrices can be found in Moore (1982) and Chanda (1981).

Our procedure in this paper is to abstract and generalize three essential steps from Moore (1982), which we now introduce in turn. The first two are asymptotic results. We will *assume* that these hold. Our goal is to avoid detailed convergence arguments, but to establish qualitative results about the limiting behavior of tests of fit that are true whenever appropriate asymptotics are available. The required asymptotic results are in fact widely true.

Received August 1982.

[1] Research supported by the National Science Foundation under Grant MCS 81-21948.

AMS 1980 *subject classifications.* Primary, 62G10; secondary, 60G10.

Key words and phrases. Tests of fit, positive dependence, stationary stochastic processes, empiric distribution function.

First, we require that the estimator $\theta_n = \theta_n(X_1, \cdots, X_n)$ used to estimate θ have an asymptotic expansion of central limit theorem type that is valid *both* for $\{X_i\}$ IID and for the SSP in question. When $\{X_i\}$ is IID, many common estimators θ_n have under $F(\cdot, \theta_0)$ the representation

$$(1.1) \qquad n^{1/2}(\theta_n - \theta_0) = n^{-1/2} \sum_{i=1}^{n} g(X_i, \theta) + o_p(1),$$

where g has zero mean and finite covariance matrix. For example, this is true in regular cases for maximum likelihood estimators (MLEs) and for Bayes estimators with respect to continuous priors. (See e.g. Ibragimov and Has'minskii (1981), Chapters I.8 and III.) In many cases, θ_n continues to satisfy (1.1) with the *same* g when $\{X_i\}$ is a SSP. This assertion must be checked in each case, but typically requires only repeating the IID-case proof and employing a law of large numbers for the SSP. Moore (1982) gives references for MLEs and a proof for minimum chi-squared estimators. Moreover, (1.1) can often be obtained directly for specific estimators without appealing to results for general classes of estimators such as MLEs. Note that while the asymptotic form of θ_n is assumed the same for IID and SSP observations, the limiting behavior will usually differ due to dependence among the $g(X_i, \theta_0)$'s.

Second, we require the availability of an appropriate central limit theorem for the SSP. In the chi-squared case, an ordinary (multivariate) central limit theorem for sums $\sum_1^n h(X_i)$ of functions of the SSP is needed. The EDF case requires in addition a weak convergence result for the EDF process. Gastwirth and Rubin (1975) give theorems that ensure that our results hold in many interesting examples, e.g. when $\{X_i\}$ is a Gaussian SSP with $\sum_1^\infty |\rho(X_1, X_{1+i})| < \infty$. Limit theorems for functions of SSPs are an active field of research; progress in this area will extend the class of processes for which our conclusions hold.

Third, we require a positivity condition on the bivariate distributions of (X_i, X_j). Section 2 discusses this condition, and places it in the context of the considerable literature on positive dependence of bivariate distributions. Useful equivalent conditions are obtained, and it is noted that a number of exchangeable bivariate distributions satisfy these conditions. In particular, our positivity condition is equivalent to

$$(1.2) \qquad E\{h(X_i)h(X_j)\} \geq 0 \quad \text{for all} \quad h \quad \text{with} \quad E|h(X_i)h(X_j)| < \infty.$$

Readers willing to accept (1.2) as a positive dependency condition without discussion may omit Section 2. Condition (1.2) is applied to study the effect of positive dependence on the large sample behavior of chi-squared statistics in Section 3, and of EDF statistics in Section 4. Section 5 comments on the generality of our methods, and applies them to several other tests of fit.

Rinott and Pollak (1980) employ condition (1.2) for a sequence of IID bivariate observations (X_i, Y_i) to study the effect of positive dependence between X and Y on the asymptotic level of tests that X and Y have equal marginal distributions. While their methods are similar to ours, their problem is quite different. And, as intuition suggests, we reach opposite qualitative conclusions: tests of equal marginal distributions for bivariate data are generally conservative under positive dependence *within* bivariate observations, while tests of fit have larger than nominal levels under positive dependence *across* observations.

2. A positivity condition. Throughout this section, (X, Y) will be an exchangeable bivariate r.v. with distribution function F. Any pair of variables from a SSP are exchangeable. Here is the required positivity condition.

DEFINITION 2.1. Exchangeable r.v.'s (X, Y), or their distribution F, are *positive dependent on intervals* (PDI) if for every integer $M \geq 2$ and every partition of $(-\infty, \infty)$

into intervals $A_1, \cdots, A_M$ the $M \times M$ matrix P with entries $p_{ij} = P[X \text{ in } A_i, Y \text{ in } A_j]$ is positive semidefinite (psd).

Since (1.2) is the essential tool employed in Sections 3 and 4, the following is the central fact about PDI distributions.

THEOREM 2.1. *(X, Y) are* PDI *if and only if* $E\{h(X)h(Y)\} \ge 0$ *for all measurable* h *such that* $E|h(X)h(Y)| < \infty$.

PROOF. If h is a step function, we can write $h = \sum_1^M b_i 1_i$ where 1_i are the indicator functions of intervals A_i that partition the line. Since then

$$E\{h(X)h(Y)\} = b'Pb$$

for $b' = (b_1, \cdots, b_m)$, clearly (1.2) implies PDI.

Suppose now that (X, Y) are PDI. If (1.2) can be established for h with $|h| \le C$, C constant, then truncating general h and employing the dominated convergence theorem will complete the proof. Moreover, if $|h| \le C$, there exists a sequence of bounded simple functions h_n converging to h such that $E\{h_n(X)h_n(Y)\} \to E\{h(X)h(Y)\}$. So (1.2) need only be proved for simple functions h satisfying $|h| \le C$.

For any such h and any $\varepsilon > 0$, there is a step function h^* with $P[h(X) \ne h^*(X)] < \varepsilon$. This follows from the result (Halmos (1950), page 56) that if μ is a σ-finite measure on a σ-field $\mathscr{F}$ generated by a field $\mathscr{F}_0$, then for any A in $\mathscr{F}$ and $\eta > 0$ there is an A^* and $\mathscr{F}_0$ with $\mu(A \,\Delta\, A^*) < \eta$, where $A \,\Delta\, A^*$ is the symmetric difference. Here, take $\mathscr{F}_0$ to be all finite unions of disjoint intervals, $\mathscr{F}$ the Borel sets on the line, and μ the distribution of X. PDI asserts that $E\{h^*(X)h^*(Y)\} \ge 0$, and (1.2) for h follows from this and

$$|E\{h(X)h(Y)\} - E\{h^*(X)h^*(Y)\}|$$
$$= |E\{h(X)[h(Y) - h^*(Y)]\} + E\{h^*(Y)[h(X) - h^*(X)]\}| \le 4C^2\varepsilon. \quad \square$$

Shaked (1979) discusses the relations among several concepts of positive dependence. He calls (X, Y) PDD if F is a positive semidefinite distribution function, i.e., if the $M \times M$ matrix with entries $F(a_i, a_j)$ is psd for all $a_1 < a_2 < \cdots < a_M$ and integers $M \ge 2$. Shaked states (Proposition 2.2) that PDD is equivalent to (1.2). However, his proof that PDD implies (1.2) involves integrating $E\{h(X)h(Y)\}$ by parts and therefore requires that h be of bounded variation. We therefore outline a direct proof of the fact that our condition PDI is equivalent to PDD. This result places PDI in the context of the relations discussed by Shaked, and with Theorem 2.1 establishes equivalence of PDD and (1.2).

THEOREM 2.2. *(X, Y) are* PDI *if and only if they are* PDD.

PROOF. For $a_1 < \cdots < a_M$, denote the matrix of $F(a_i, a_j)$ by F_M and the column vector of $F(a_i, \infty) = F(\infty, a_i)$ by f_M. First notice that F_M is psd for all a_i and M if and only if $F_M - f_M f_M'$, the matrix of $F(a_i, a_j) - F(a_i, \infty)F(\infty, a_j)$, is. The "if" assertion is obvious. To see "only if," take $a_M = \infty$ and expand $\det(F_M)$ by its last column to obtain $\det(F_M) = \det(F_{M-1} - f_{M-1}f_{M-1}')$. Since $\det(F_M) \ge 0$, all of the $(M-1) \times (M-1)$ upper left principal minors of $F_M - f_M f_M'$ (for arbitrary a_M) have nonnegative determinant, and this $M \times M$ matrix is therefore psd. PDD is thus equivalent to $F_M - f_M f_M'$ psd for all M and a_i.

On the other hand, PDI is immediately equivalent to $\text{Cov}\{h(X)h(Y)\} \ge 0$ for all step functions h. But if $h = \sum_1^M b_i 1_i$, where 1_i is the indicator of $A_i = (a_{i-1}, a_i]$ and $-\infty = a_0 < a_1 < \cdots < a_M = \infty$, then

$$\text{Cov}\{h(X)h(Y)\} = x'(F_{M-1} - f_{M-1}f_{M-1}')x,$$

where $x_i = b_i - b_{i+1}$, $i = 1, \cdots, M-1$, and F_{M-1} is formed from $a_1 \cdots a_{M-1}$. The theorem follows. $\square$

REMARKS. (1) It is easy to see that (1.2) is equivalent to $\mathrm{Cov}\{h(X)h(Y)\} \geq 0$ for all h for which the covariance exists. Shaked (1979) and Rinott and Pollak (1980) use the condition in this form.

(2) PDI implies that the correlation $\rho(X, Y) \geq 0$. For (X, Y) bivariate normal, PDI is equivalent to $\rho(X, Y) \geq 0$, but the equivalence does not hold for all bivariate exchangeable distributions.

(3) If (X, Y) are conditionally IID, they are PDI. Moore (1982) established PDI for symmetric bivariate normal (X, Y) with $\rho > 0$ in this way. Other distributions that are conditionally IID, and hence PDI, are listed with references by Shaked (1977), page 510; they include bivariate exponential, F, logistic, and χ^2 distributions. The bivariate t distributions are also in this class. Shaked (1979) shows that not all PDD distributions are conditionally IID, and that the class of PDD distributions is closed under convolution, mixture, and convergence in distribution. The class of PDD (or PDI) distributions is thus extensive.

(4) Total positivity of order infinity (TP_∞) for F implies PDI, but the bivariate t, for example, is PDI but not even TP_2.

3. Chi-squared statistics. Observations $X_1, \cdots, X_n$ are to be tested for fit to the family $\{F(\cdot, \theta): \theta \text{ in } \Omega\}$. Choose cells $A_k = (a_{k-1}, a_k]$, $k = 1, \cdots, M$ with boundaries $-\infty = a_0 < a_1 < \cdots < a_M = \infty$. Let 1_k be the indicator function of A_k, so that the kth cell frequency is $N_k = \sum_{i=1}^n 1_k(X_i)$. The corresponding cell probability is $p_k(\theta) = F(a_k, \theta) - F(a_{k-1}, \theta)$. Let $V_n(\theta)$ be the M-vector of standardized cell frequencies, having kth component $[N_k - np_k(\theta)]/[np_k(\theta)]^{1/2}$. Except in the simple null hypothesis case $\Omega = \{\theta_0\}$, the unknown parameter θ is estimated by $\theta_n = \theta_n(X_1, \cdots, X_n)$. Chi squared statistics are psd quadratic forms in $V_n(\theta_n)$. In particular, the Pearson statistic is the sum of squares $V_n(\theta_n)' V_n(\theta_n)$.

Suppose now that $X_1, X_2, \cdots$ are a SSP, and that the common univariate marginal distribution of the X_i is $F(\cdot, \theta_0)$ for some θ_0 in Ω. Suppose further that the estimator θ_n satisfies (1.1) both when $\{X_i\}$ is IID $F(\cdot, \theta_0)$ and for the SSP in question. Then Moore (1982) follows the IID-case development of Moore and Spruill (1975) to show that

$$V_n(\theta_n) = n^{-1/2} \sum_{i=1}^n h(X_i) + o_p(1). \tag{3.1}$$

Here $h(x) = \Delta(x) - Bg(x, \theta_0)$, where B is the $M \times m$ matrix with (i, j)th entry $p_i^{-1/2}\partial p_i/\partial\theta_j$ and $\Delta(x)$ is the M-vector with components $[1_k(x) - p_k]/p_k^{1/2}$. (When the argument θ is omitted, $\theta = \theta_0$ is assumed.) Since $E\{h(X_i)\} = 0$, a central limit theorem applied to (3.1) will imply

$$V_n(\theta_n) \to_{\mathscr{L}} N(0, \Sigma), \quad \Sigma = \lim_{n\to\infty}(1/n)\mathrm{Cov}\{\textstyle\sum_{i=1}^n h(X_i)\} < \infty. \tag{3.2}$$

Moore (1982) cites several applicable central limit theorems for SSPs, and notes that (3.2) often continues to hold even when data-dependent cells are employed. The limiting covariance matrix Σ will of course differ from Σ_{IID}, the limiting covariance matrix of $V_n(\theta_n)$ in the IID case. Chanda (1981) and Moore (1982) derive the form of Σ for several common estimators θ_n. Here is our main result on chi-squared tests.

THEOREM 3.1. *Suppose that $X_1, X_2, \cdots$ is a* SSP *such that (X_i, X_j) is* PDI *for all $i \neq j$, that X_i has distribution function $F(\cdot, \theta_0)$, and that* (3.2) *holds under $F(\cdot, \theta_0)$ both for $\{X_i\}$* IID *and for the* SSP *in question. Then if Σ_{IID} is the limiting covariance matrix of $V_n(\theta_n)$ in the* IID *case, $\Sigma - \Sigma_{\mathrm{IID}}$ is* psd.

PROOF. Write

$$\Sigma = E\{h(X_1)h(X_1)'\} + \lim_{n\to\infty}(1/n) \textstyle\sum_{i,j=1, i\neq j}^n C_{ij},$$

where $C_{ij} = E\{h(X_i)h(X_j)'\}$. The first term on the right is Σ_{IID} since all $C_{ij} = 0$ in the IID case. Theorem 2.1 implies that all C_{ij} are psd, since $a'C_{ij}a = E\{f(X_i)f(X_j)\}$ where $f(x) = \sum_1^M a_k h_k(x)$, a_k and h_k being the ith components of the M-vectors a and h, respectively. Thus $\Sigma - \Sigma_{\text{IID}} = \lim_{n\to\infty}(1/n) \sum C_{ij}$, which exists and is finite by (3.2), is psd. □

General statistics of chi-squared type have the form $T_n = V_n(\theta_n)' W_n V_n(\theta_n)$, where W_n is a (possibly data-dependent) psd $M \times M$ matrix converging in $F(\cdot, \theta_0)$-probability as $n \to \infty$ to a psd matrix $W = W(\theta_0)$. A number of useful examples of such statistics, in addition to the Pearson case $W = I$, are discussed in Moore and Spruill (1975) and Moore (1977). In all these cases, the centering matrix W is the same in the limit for $\{X_i\}$ IID and for SSP's such that θ_n remains a consistent estimator of θ_0. The limiting null distribution of T_n is that of $V'WV$ for $V \sim N(0, \Sigma)$. This is the distribution of $\Sigma_1^M \lambda_k Z_k^2$, where the Z_k are independent $N(0, 1)$ r.v.'s and λ_k are the characteristic roots of $W^{1/2}\Sigma W^{1/2}$. Theorem 3.1 implies that $W^{1/2}(\Sigma - \Sigma_{\text{IID}})W^{1/2}$ is psd, and hence (Bellman (1960), page 115) that $\lambda_k(W^{1/2}\Sigma W^{1/2}) \geq \lambda_k(W^{1/2}\Sigma_{\text{IID}} W^{1/2})$, where $\lambda_k(H)$ denotes the kth largest characteristic root of a matrix H. We have proved the following result.

COROLLARY 3.1. *Suppose that the conditions of Theorem* 3.1 *hold, and that* $T_n = V_n(\theta_n)' W_n V_n(\theta_n)$ *where* W_n *has the same* psd *limit in probability in the* IID *and* SSP *cases. Then the limiting null distribution of* T_n *in the* SSP *case is stochastically larger than in the* IID *case.*

In the limit, the test of fit with critical region $T_n > c$ rejects at least as often in the SSP case as in the IID case. This result applies in particular to the Pearson statistic with θ_n the minimum chi-squared estimator (the Pearson-Fisher statistic) or with θ_n the raw data maximum likelihood estimator (the Chernoff-Lehmann statistic). These common tests, when applied to SSP data by a naive user who believes the data to be IID, therefore reject too often whenever the SSP is positively dependent in the PDI sense and is sufficiently regular to be covered by a central limit theorem implying (3.2).

REMARKS. (1) The difference $\lambda_k(\text{SSP}) - \lambda_k(\text{IID})$ for some characteristic values λ_k, and therefore the difference in test level, is strict when Σ and Σ_{IID} have a common null space $\mathscr{N}$ and $\Sigma - \Sigma_{\text{IID}}$ is positive definite on $\mathscr{N}^{\perp}$. Examination of the form of Σ given by Moore (1982) shows that this is usually the case. Moore also shows that for several common Gaussian SSP's the characteristic values increase without bound, and the test level approaches 1, as the positive dependence of the SSP increases. In fact, the machinery of Section 2 can be used to show that the matrix C_{ij} in the proof of Theorem 3.1 is monotone in the incidence matrix P_{ij} of Definition 2.1 for (X_i, X_j), in the sense that $P_{ij}^{(1)} - P_{ij}^{(2)}$ psd implies $C_{ij}^{(1)} - C_{ij}^{(2)}$ psd. By a result of Rinott and Pollak (1980, page 194) it follows that the test levels in Moore's Gaussian 1-dependent and first order autoregressive examples are increasing functions of the correlation $\rho(X_i, X_{i+1})$. The confounding of positive dependence with lack of fit can therefore be arbitrarily serious in common cases.

(2) Corollary 3.1 is a general statement resulting from the assumption that all (X_i, X_j) are PDI. For some chi-squared tests based on M cells, the conclusion of Corollary 3.1 can be obtained assuming only that the incidence matrix P of Definition 2.1 is psd for partitions of the line into *exactly* M intervals. This is done for the Pearson statistic without estimated parameters in Theorem 3.1 of Moore (1982). A slight modification of the argument given there applies as well to the Pearson-Fisher statistic, $V_n(\theta_n)' V_n(\theta_n)$ with θ_n the minimum chi-squared estimator. The Pearson and Pearson-Fisher statistics are distinguished by the fact that Σ_{IID} is a projection matrix. We do not have a direct proof requiring only PDI for fixed M in other cases covered by Corollary 3.1, such as the Chernoff-Lehmann statistic.

(3) Common IID-case chi-squared statistics employ centering matrices $W_n(X_1, \cdots, X_n)$ having the same limit W for quite general SSPs $\{X_i\}$. By Corollary 3.1 and Remark 1,

such statistics have different limiting laws for different degrees of dependence among the X_i. It is sometimes possible to choose W_n to adjust for the dependence, and obtain a statistic having the same distribution for, e.g., any m-dependent SSP. Moore (1982) gives an example of such a statistic. In this example, W_n involves sample estimators of the incidence matrices P_{ij}, and does not have the same limit in the IID and dependent-data cases.

4. EDF statistics. The statistics considered in this section are functions of the EDF process with parameter θ estimated. Durbin (1973) laid down the outline for the large-sample theory of such statistics in the IID case. Neuhaus (1976) presents the theory in a manner very similar in outline and generality to the analogous chi-squared theory of Moore and Spruill (1975). We will show, without repeating details, that Neuhaus' development extends to suitable SSPs.

We remark first that a basic condition for the meaningfulness of EDF statistics for testing fit of the univariate marginal of a SSP $X_1, X_2, \cdots$ is that the Glivenko-Cantelli result $\sup | F_n(x) - F(x, \theta_0) | \to 0$ a.s. continues to hold, where F_n is the EDF of $X_1, \cdots, X_n$. This is clearly true for $\{X_i\}$ ergodic; see e.g. Tucker (1959). Ergodicity is stronger than the condition stated by Moore (1982) for $N_k/n \to p_k(\theta_0)$, which is required for meaningfulness of chi-squared tests. But ergodicity is weak relative to the conditions known to imply central limit theorems for functions of $\{X_i\}$, and carries with it the laws of large numbers that are usually sufficient to verify (1.1) for SSPs in regular cases. Of course, we do not invoke ergodicity explicitly because of our strategy of assuming that the required convegence results hold.

Suppose that $X_1, X_2, \cdots$ have common df $F(\cdot, \theta_0)$. Define, following Neuhaus,

$$\bar{F}(\cdot, \theta) = F(F^{-1}(\cdot, \theta_0), \theta), \quad V_i = F(X_i, \theta_0) \quad i = 1, 2, \cdots,$$

and let $\bar{F}_n$ be the EDF of $V_1, \cdots, V_n$. The EDF process is $\bar{Z}_n(t) = n^{1/2}[\bar{F}_n(t) - \bar{F}(t, \theta_n)]$ for $0 \le t \le 1$, and takes values in the Skorohod space $D[0, 1]$. If F is suitably regular and θ_n satisfies (1.1), Neuhaus' arguments apply in the SSP case, and show that under $F(\cdot, \theta_0)$

$$\bar{Z}_n(t) = n^{-1/2} \textstyle\sum_{i=1}^n h(t, V_i) + o_p(1), \tag{4.1}$$

where $E\{h(t, V_i)\} = 0$ and $o_p(1)$ now means uniform convergence to zero in probability over $0 \le t \le 1$. Here

$$h(t, v) = 1_t(v) - \bar{F}(t, \theta_0) - \bar{g}(v, \theta_0)' q(t, \theta_0),$$

where 1_t is the indicator function of $(-\infty, t]$, $\bar{g}(\cdot, \theta_0) = g(F^{-1}(\cdot, \theta_0), \theta_0)$ with g the function in (1.1), and $q(t, \theta)$ is the m-vector of derivatives $\partial F(s, \theta)/\partial\theta_k$ evaluated at $s = F^{-1}(t, \theta_0)$. The expression (4.1) is analogous to (3.1), and similarly holds with the same function h both for $\{X_i\}$ IID and for SSPs whenever θ_n satisfies (1.1) in both cases and F is sufficiently regular.

A suitable central limit theorem applied to (4.1) will imply the analog of (3.2):

$$\bar{Z}_n \to_{\mathscr{L}} \bar{Z}_0 \quad \text{in} \quad D[0, 1], \quad \text{where } \bar{Z}_0 \text{ is a Gaussian process}$$

(4.2) with a.s. continuous paths, zero mean, and covariance function

$$c(s, t) = \lim_{n\to\infty}(1/n)\text{Cov}\{\textstyle\sum_{i=1}^n h(s, V_i), \sum_{j=1}^n h(t, V_j)\} < \infty.$$

Examination of the form of $h(t, v)$ and of Neuhaus' proof of the weak convergence $\bar{Z}_n \to_{\mathscr{L}} \bar{Z}_0$ in the IID case (his Theorem 2.2) show that (4.2) follows from: (a) A finite-dimensional central limit theorem for $\sum_1^n h(t, V_i)$ that includes convergence of covariances in its conclusion; and (b) A weak convergence result for the EDF process $n^{1/2}[\bar{F}_n(t) - \bar{F}(t, \theta_0)]$ *without parameter estimation*. This separation occurs because the θ_n enter $h(t, V_i)$ only via the product of a function $\bar{g}$ of V_i and a function q of t that is the same for all n.

Since both (a) and (b) are known to hold for many SSP's, (4.2) will often be true in the SSP case as well as in the IID case. For example, Theorem 22.2 of Billingsley (1968), which has been considerably extended by later authors, implies (4.2) for certain φ-mixing processes. Since many common time series models are not φ-mixing, more useful results for our purposes are given by Gastwirth and Rubin (1975). They establish (a) for all h having finite variance, and also (b), for a class of mixing processes that includes all Gaussian SSP's with $\Sigma \mid \rho(X_1, X_{1+i}) \mid < \infty$.

THEOREM 4.1. *Suppose that $X_1, X_2, \cdots$ is a* SSP *such that (X_i, X_j) is* PDI *for all $i \neq j$, that X_i has* df *$F(\cdot, \theta_0)$, and that* (4.2) *holds under $F(\cdot, \theta_0)$ both for $\{X_i\}$* IID *and for the* SSP *in question. Then if $c_{\text{IID}}(s, t)$ is the covariance function of $\bar{Z}_0$ in the* IID *case, $c(s, t) - c_{\text{IID}}(s, t)$ is a* psd *function.*

PROOF. Following the proof of Theorem 3.1, we need only show that

$$c_{ij}(s, t) = E\{h(s, V_i)h(t, V_j)\}$$

is psd, all $i \neq j$. For any function f on [0, 1] for which the integral converges absolutely,

$$\int_0^1 \int_0^1 f(s)f(t)c_{ij}(s, t)\, ds\, dt = E\{h_f(V_i)h_f(V_j)\}$$

where $h_f(v) = \int_0^1 f(s)h(s, v)\, ds$. Since (V_i, V_j) is PDI, this integral is nonnegative. □

To obtain comparisons of asymptotic test levels for the SSP and IID case, we apply Theorem 4.1 together with a generalization by Rinott and Pollak (1980) of a lemma of T. W. Anderson.

LEMMA 4.1. (Rinott and Pollak). *Let Z_1, Z_2 be Gaussian processes in $C[0, 1]$ with zero means and covariance functions $c_1(s, t), c_2(s, t)$ respectively, such that $c_2(s, t) - c_1(s, t)$ is a* psd *function. Then $P[Z_1 \text{ in } A] \geq P[Z_2 \text{ in } A]$ for any closed, convex, symmetric set A in $C[0, 1]$.*

Taking $A_\Lambda = \{f \text{ in } C[0, 1]: \Lambda(f) \leq c\}$ for Λ a continuous functional on $C[0, 1]$, the large sample level of the test of fit with critical region $\Lambda(\bar{Z}_n) > c$ will be greater in the SSP case than in the IID case whenever A_Λ is closed, convex, and symmetric and Theorem 4.1 applies. Note that only continuity on $C[0, 1]$ (that is, under uniform convergence to a continuous limit) is required of Λ, since $\bar{Z}_0$ is in $C[0, 1]$ a.s. and $\Lambda(\bar{Z}_n) \to_{\mathscr{D}} \Lambda(\bar{Z}_0)$ follows from continuity a.s. with respect to the distribution of $\bar{Z}_0$.

The result above covers the Kolmogorov-Smirnov (KS) test, for which $\Lambda(f) = \sup |f|$. The Cramér-von Mises (CvM) statistic is not a fixed functional of $\bar{Z}_n$, but rather $\Lambda_n(\bar{Z}_n)$ where $\Lambda_n(f) = \int f^2(t)\, d\bar{F}(t, \theta_n)$. But Neuhaus (1976, page 76) shows that $\Lambda_n(\bar{Z}_n) \to_{\mathscr{L}} \Lambda(Z_0)$, where $\Lambda(f) = \int f^2(t)\, d\bar{F}(t, \theta_0)$, whenever $\bar{Z}_n \to_{\mathscr{L}} \bar{Z}_0$ in $D[0, 1]$ and $\int f(t)\, d\bar{F}(t, \theta_n) \to \int f(t)\, d\bar{F}(t, \theta_0)$ in probability for all f in $C[0, 1]$. The latter condition is satisfied under (1.1) and Neuhaus' regularity conditions on F. Lemma 4.1 applied to A_Λ now shows that the limiting level of the CvM critical regions $\Lambda_n(\bar{Z}_n) > c$ is larger for SSPs satisfying the conditions of Theorem 4.1 than for IID observations.

The Kolmogorov-Smirnov and Cramér-von Mises examples motivate, and show two different ways of applying, our concluding result.

COROLLARY 4.1. *Suppose that the conclusion of Theorem* 4.1 *holds for a* SSP *$\{X_i\}$ and that a test of fit of $X_1, \cdots, X_n$ has critical regions S_n such that $P[(X_1, \cdots, X_n) \text{ in } S_n] \to P[\Lambda(\bar{Z}_0) > c]$, where Λ is a functional on $C[0, 1]$ with $A_\Lambda = \{f: \Lambda(f) \leq c\}$ closed, convex, and symmetric. Then the limiting level of the test is at least as large in the* SSP *case as in the* IID *case.*

REMARKS. (1) Normal cdf's satisfy Neuhaus' regularity conditions and those needed to ensure that the MLE's $(\bar{X}, s)$ of the parameters (μ, σ) satisfy (1.1). Moreover, for Gaussian processes PDI is equivalent to $\rho_k \geq 0$ for $k \geq 1$, where $\rho_k = \rho(X_1, X_{1+k})$. The convergence results of Gastwirth and Rubin therefore ensure that Corollary 4.1 applies to any Gaussian SSP $\{X_i\}$ with $\rho_k \geq 0$ and $\sum \rho_k < \infty$ when $(\bar{X}, s)$ are used as estimators in testing normality. This class includes the first-order autoregressive and all m-dependent Gaussian processes.

(2) Corollary 4.1 is designed to apply to critical regions of the form $\{\Lambda(\bar{Z}_n) > c\}$ or $\{\Lambda_n(\bar{Z}_n) > c\}$. EDF tests are sometimes employed with critical regions of forms such as $\{\Lambda(\bar{Z}_n) > c_n\}$, where $c_n \to c$. When the distribution function of $\Lambda(\bar{Z}_0)$ is continuous, the conclusion of the corollary continues to apply.

(3) In addition to the usual KS and CvM statistics, Corollary 4.1 applies to weighted versions of these statistics, as well as to the extensions of the CvM statistic discussed by Neuhaus (1973). The treatment is similar to that of the CvM statistic above; the necessary analysis can typically be found in the literature on IID-case convergence.

(4) In the CvM case, the limiting null distribution is that of $\sum_{k=1}^{\infty} \lambda_k Z_k^2$, where the Z_k are independent $N(0, 1)$ r.v.'s and λ_k are the characteristic roots of the covariance function $c(s, t)$ considered as an operator on an appropriate L_2 space. (See Neuhaus (1979) for a survey.) In this case, one can obtain $\lambda_k(\text{SSP}) \geq \lambda_k(\text{IID})$ as in the chi-squared case.

5. Other statistics. The method of proof used in this paper is both simple and quite general in applicability. Consider any statistic of the form $T_n = \Lambda(U_n(\theta_n))$, where θ_n is an estimator of θ, $U_n(\theta)$ is an asymptotically Gaussian random variable in a space S, and $\Lambda: S \to [0, \infty)$ is a continuous, convex, symmetric functional. The limiting null distribution of T_n in the IID case is obtained by an analytic expansion, first of θ_n and then of $U_n(\theta_n)$, about the true θ_0, followed by application of a CLT on S to the dominant term of the expansion. Thus $U_n(\theta_n)$ is asymptotically Gaussian for IID data. Inspection typically reveals that for quite general SSP's, the estimator θ_n remains consistent and therefore the *same* analytic expansion of $U_n(\theta_n)$ remains valid. Whenever a suitable CLT for SSP's on S exists, $U_n(\theta_n)$ is therefore asymptotically Gaussian both for IID and for SSP data. The positive dependency condition (1.2) implies that the difference between the covariance functions for the SSP and IID cases is positive semidefinite. It then follows by Anderson's lemma (see Tong (1980), page 55) or its generalization to function space that the asymptotic null distribution of T_n is stochastically larger for SSP than for IID data. Thus any critical region $\{T_n > c\}$ has asymptotic size at least as large for SSP as for IID data.

In Section 4 we applied this method with $U_n(\theta_n) = \bar{Z}_n$ on $S = D[0, 1]$. In the setting of Section 3, with $U_n(\theta_n) = V_n(\theta_n)$ on $S = R^M$, Anderson's lemma for $N(0, \Sigma)$ provides an alternate proof of Corollary 3.1. There $\Lambda(x) = x'Wx$ for positive semidefinite W. Other possible choices for Λ when $x = (x_1, \cdots, x_M)$ are $\Lambda_1(x) = \max |x_k|$, $\Lambda_2(x) = \sum_1^M |x_k|$ and $\Lambda_3(x) = \max_{1 \leq m \leq M} |\sum_{k=1}^m x_k|$. Taking $x_k = n^{-1/2}(N_k - np_k)$, which does not change the applicability of our method, Λ_2 generates a statistic of Hoeffding equivalent to David's empty cell statistic, and Λ_3 generates a KS statistic for discrete or grouped data. Both of these statistics, particularly the latter, are discussed by Pettitt and Stephens (1977).

The confounding of positive dependence with lack of fit holds for tests based on convex, symmetric functionals of other asymptotically Gaussian quantities as well, provided only that the technical task of establishing the required CLT for the SSP case is successful. Candidates include tests of fit based on the quantile process and on spacings, for which Shorack (1972) establishes convergence to Gaussian processes in the IID case. In addition, some common test statistics have analytic expansions showing that under the null hypothesis they are asymptotically equivalent for both IID and SSP data to statistics of the classes treated here. Inspection of the analysis shows that our qualitative conclusion applies. For example, this is true of the log likelihood ratio statistic for grouped data because of its analytic relation to the Pearson statistic. Cressie and Read (1982) have proposed a family of statistics asymptotically equivalent to the Pearson statistic. This family includes the log likelihood ratio, Neyman modified chi-squared, and Freeman-Tukey statistics. Their analysis combined with our method shows that Corollary 3.1 holds for the entire class.

REFERENCES

BELLMAN, R. (1960). *Introduction to Matrix Analysis.* McGraw-Hill, New York.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

CHANDA, K. C. (1981). Chi-square goodness-of-fit tests based on dependent observations. In C. Taillie et al. (Eds.), *Statistical Distributions in Scientific Word* **5** 35–49. D. Reidel, Dordrecht, Netherlands.

CRESSIE, N. and READ, T. R. C. (1983). Multinomial goodness-of-fit tests. *J. Roy. Statist. Soc.*, to appear.

DURBIN, J. (1973). Weak convergence of the sample distribution function when parameters are estimated. *Ann. Statist.* **1** 279–290.

GASTWIRTH, J. L. and RUBIN, H. (1975). The asymptotic distribution theory of the empiric c.d.f. for mixing processes. *Ann. Statist.* **3** 809–824.

HALMOS, P. R. (1950). *Measure Theory.* Van Nostrand, New York.

IBRAGIMOV, I. A. and HAS'MINSKII, R. Z. (1981), *Statistical Estimation, Asymptotic Theory* (trans. S. Kotz). Springer, New York.

MOORE, D. S. (1977). Generalized inverses, Wald's method, and the construction of chi-squared tests of fit. *J. Amer. Statist. Assoc.* **72** 131–137.

MOORE, D. S. (1982). The effect of dependence on chi-squared tests of fit. *Ann. Statist.* **10** 1163–1171.

MOORE, D. S. and SPRUILL, M. C. (1975). Unified large-sample theory of general chi-squared statistics for tests of fit. *Ann. Statist.* **3** 599–616.

NEUHAUS, G. (1973). Zur Verteilungskonvergenz einiger Varianten der Cramér-von Mises-Statistik. *Math. Operationsforsch. u. Statist.* **4** 473–484.

NEUHAUS, G. (1976). Weak convergence under contiguous alternatives of the empirical process when parameters are estimated: the D_k approach. *Lecture Notes in Math.* **566** 68–82. Springer, New York.

NEUHAUS, G. (1979). Asymptotic theory of goodness of fit tests when parameters are present: a survey. *Math. Operationsforsch. Statist., Ser. Statist.* **10** 479–497.

PETTITT, A. N. and STEPHENS, M. A. (1977). The Kolmogorov-Smirnov goodness-of-fit statistic with discrete and grouped data. *Technometrics* **19** 205–210.

RINOTT, J. and POLLAK, M. (1980). A stochastic ordering induced by a concept of positive dependence and monotonicity of asymptotic test sizes. *Ann. Statist.* **8** 190–198.

SHAKED, M. (1977). A concept of positive dependence for exchangeable random variables. *Ann. Statist.* **5** 505–515.

SHAKED, M. (1979). Some concepts of positive dependence for bivariate interchangeable distributions. *Ann. Inst. Statist. Math.* **31** 67–84.

SHORACK, G. R. (1972). Convergence of quantile and spacings processes with applications. *Ann. Math. Statist.* **43** 1400–1411.

TONG, Y. L. (1980). *Probability Inequalities in Multivariate Distributions.* Academic, New York.

TUCKER, H. G. (1959). A generalization of the Glivenko-Cantelli Theorem. *Ann. Math. Statist.* **30** 828–830.

DEPARTMENT OF STATISTICS
PURDUE UNIVERSITY
MATHEMATICAL SCIENCES BUILDING
WEST LAFAYETTE, INDIANA 47907