Some Tests Of Normality Based On Transforms.

Sucharita Ghosh.

Department of Statistics, University of North Carolina, Chapel Hill.

Summary: In this paper, some graphical tests of normality are proposed involving methods based on empirical transforms. The suggested techniques use derivatives of the empirical moment generating function and the empirical characteristic function and can also be used as formal testing procedures. Good power properties are indicated in power calculations. The proposed methods are invariant under location and scale transformations of the data, the graphical techniques in particular are found to be attractive and intuitively appealing in nature. Weak convergence results are obtained under general conditions. Some optimality problems are handled via maximisation of Bahadur slopes of test statistics.

Key Words: *Testing normality; Empirical Transforms; Probability plots; Bahadur slopes.*

Introduction

Let $X_1,...,X_n$ denote observations on a univariate random variable X with distribution function $F(x) = Pr(X \leq x)$, density function $f(x)$, characteristic function $\phi(t) = E(e^{itX})$ and moment generating function $m(t) = E(e^{tX})$ where $t \in R$. The problem here is to test $H_0 : X \sim N(\mu,\sigma^2)$ against the alternative H_1 : *negation of* H_0. Here $N(\mu,\sigma^2)$ denotes the normal distribution with mean μ and variance σ^2. If either μ or σ is not specified, H_0 is a composite null hypothesis, otherwise as in many practical situations, these parameters may be known with a reasonably high degree of accuracy, H_0 may be called a simple null hypothesis. A fairly large amount of work has been done in this area. Mardia in his 1980 paper discusses the existing tests beautifully. Other references on this topic are Gnanadesikan (1977) and D'Agostino and Stephens (1986). However the uniqueness of characteristic function and the moment generating function (when it exits) suggest goodness-of-fit methods based on these transforms. Define the empirical characteristic function by $\phi_n(t) = \int e^{itx} dF_n(x)$ and the empirical moment generating function by $m_n(t) = \int e^{tx} dF_n(x)$ where $F_n(x)$ denotes the empirical distribution function for the data $X_1, \cdots X_n$. Because of excellent properties available for the empirical characteristic function (ecf) and the empirical moment generating function (emgf) (see Csorgo 1980, 1981, Feuerverger and McDunnough 1981a, 1981b, 1984, Feuerverger and Mureika 1977), a number of authors have considered using these for goodness-of-fit problems (see Heathcote 1972, Feigin and Heathcote

1977, Koutrovellis 1980,1981, Koutrovellis and Kellermeier 1981, Kellermeier 1980, Murota and Takeuchi 1981, Hall and Welsh 1983, Epps and Pulley 1983, Epps, Singleton and Pulley 1982, Csorgo 1986, Epps 1987). A survey of testing by empirical characteristic functions is available in Csorgo (1984). In the present paper we develop some graphical methods for testing normality based on certain stochastic processes which can also be used as formal tests. Several graphs are exhibited in the appendix along with some power calculations for some of the suggested tests which indicate good power properties.

[A] Use of the empirical moment generating function:

Section [A1]: A general goodness-of-fit problem:

In this section we consider a more general goodness- of-fit problem; for simplicity we restrict ourselves to simple hypotheses. Thus assuming that the mgf exists both for the distribution specified by H_0 as well as the one specified by H_1, we consider the following testing problem:

$$H_0\colon m(t) = m_0(t) \tag{I}$$

versus

$$H_1\colon m(t) = m_1(t)$$

where $m_0(t) \neq m_1(t)$ for at least one $t \in R$. Then it is easily seen that under H_i $(i=0,1)$, the finite dimensional distributions of

$$T_n(t) = \sqrt{n}\,(m_n(t) - m_i(t))$$

converge to the finite dimensional distributions of a zero mean Gaussian process having the covariance structure

$$K_i(s,t) = m_i(s+t) - m_i(s)m_i(t)$$

[A1.1]: Although, two mgf's may intersect at a finite collection of values of the parameter t, examination of the empirical mgf at a finite number of t-points provides substantial information regarding the population. For simplicity here, a test based on the statistic $T_n(t)$ at a single value of t, may be based upon an approximate rejection region given by:

$$|\ T_n(t)\ | > \sqrt{K_0(t,t)}\cdot z_{\alpha/2}$$

where $z_{\alpha/2}$ is the upper $\alpha/2$ point of the $N(0,1)$ distribution.

For different situations note that several optimality criteria may be used to at least partially resolve this problem; in particular, Bahadur slopes of test statistics may be maximized with respect to t to decide about the values of the parameter t to be chosen. Noting that $\left\{ T_n(t) \right\}$ is a 'standard sequence' (Bahadur 1960), the approximate slope of the sequence $\left\{ T_n(t) \right\}$ may be calculated as:

$$C^{(a)}(t) = \frac{[m_1(t) - m_1(t)]^2}{K_0(t,t)}.$$

On the other hand, since $T_n(t)$, for fixed t is just $\sqrt{n}$ times the sample mean of n iid random variables Chernoff's large deviation theorem (Chernoff 1952) can be used to calculate the exact slope of the same sequence $\left\{ T_n(t) \right\}$ and is given by,

$$C^{(e)}(t) = 2 \, \underset{s}{Sup} \left\{ (m_1(t) - m_0(t))s - \log M_0(s \mid t) \right\}$$

where $M_0(s \mid t)$ is the mgf of e^{tX} under H_0, evaluated at s.

Maximization of $C^{(e)}(t)$ with respect to t may suggest a 'good' choice of t for testing (I), while simple algebraic manipulations show that for maximizing $C^{(e)}(t)$, an initial value of t may be taken as the value that maximizes $C^{(a)}(t)$. This is not surprising since asymptotically in the neighbourhood of H_0, the approximate slope and the exact slope are the same.

[A1.2]: Use of weighted integrals: The above discussion generalizes to the case where one considers a weight function $w(t)$, $[-T,T]$ $0 < T < \infty$ and define the test statistic

$$S_n = \int_t w(t) \, T_n(t) \, dt$$

where the parameter t may be restricted to a finite interval $[-T,T]$, $T < \infty$.

Asymptotic null distribution of S_n follows from central limit theorem, so that under H_0, as $n \rightarrow \infty$,

$$S_n \sim N(0, \sigma^2) \ \ where \ \ \sigma^2 = \int\int w(s)w(t)K_0(s,t)dsdt.$$

Clearly the properties of such a test will depend on the choice of $w(t)$ and hence a reasonable choice of $w(\cdot)$ may be of interest although such a choice will in general depend on H_1.

In particular, Bahadur's approximate slope for the sequence $\left\{ S_n \right\}$ is given by:

$$C^{(a)} = \frac{(\int w(t)[m_1(t) - m_0(t)] \, dt)^2}{\int\int w(s)w(t)K_0(s,t) \, ds \, dt}$$

which when maximized with respect to $w(\cdot)$ under the constraint (w.l.o.g.) $\int\int w(s)w(t)K_0(s,t) \, ds \, dt = 1.$ gives rise to the solution:

$$w(t) = L^{-1}\left[\frac{f_0(\cdot)}{f_1(\cdot)}\right]$$

where L^{-1} denotes the inverse Laplace transform operator.

Proof: Define $U(x) = \int w(t)e^{tx}\,dt$. Noting that

$$Var_{H_0}U(X) = \int\int w(s)w(t)K_0(s,t)\,ds\,dt,$$

the problem reduces to the maximization of $C^{(a)} = \frac{[E_{H_1}U(X) - E_{H_0}U(X)]^2}{Var_{H_0}U(X)}$ with respect to $w(\cdot)$ such that $Var_{H_0}U(X) = 1$. Equivalently w.l.o.g.,

$$Maximise \ E_{H_1}U(X) \tag{P1}$$

subject to:

$$(i) \quad E_{H_0}U(X) = 0 \tag{A}$$

and

$$(ii) E_{H_0}U^2(X) = 1 \tag{B}$$

Introducing variants $\delta U(x)$ to the solution of (P1) we have:

$$\int \delta U(x) f_1(x)\,dx = 0 \tag{a}$$

$$\int \delta U(x) f_0(x)\,dx = 0 \tag{b}$$

and

$$\int \delta U(x) U(x) f_0(x)\,dx = 0 \tag{c}$$

Equations (a),(b) and (c) together imply that there exists constants α and β such that

$$f_1(x) = \alpha\, U(x) f_0(x) + \beta\, f_0(x). \tag{d}$$

Using (A) and (B),

$$\int U(x) f_1(x)\,dx = \alpha$$

so that

$$f_1(x) = U(x) f_0(x) \int U(x) f_1(x)\,dx + \beta f_0(x).$$

Integrating both sides of the last equation and applying (A) we have $\beta = 1$, so that (d) implies $1 + \alpha U(x) = \frac{f_1(x)}{f_0(x)}$. Since $C^{(a)}$ is invariant under transformations of the type $aU(x) + b$, these constants can be removed and we may write

$$U(x) = \frac{f_1(x)}{f_0(x)}$$

or

$$w(t) = L^{-1}\frac{f_1(x)}{f_0(x)}$$

On the other hand, Bahadur's exact slope for the same sequence $\left\{ S_n \right\}$ may be calculated as:

$$C^{(e)} = -2 \log \underset{s}{Inf}\, e^{-su} M_0(s)$$

where, $u = \int w(t)[m_1(t) - m_0(t)]\, dt$ and $M_0(s)$ is the moment generating function of $\int w(t)[e^{tx} - m_0(t)]\, dt$, under H_0, evaluated at s assuming that it exists. In this situation it can be shown that the function $w(t)$ that maximizes $C^{(e)}$ is given by the inverse Laplace transform of the logarithm of the ratio $\frac{f_1(\cdot)}{f_0(\cdot)}$ where as before f_j is the probability density function of the underlying random variable under H_j.

One may notice at this point that maximization of the exact slope gives rise to the best (most powerful) test suggested by the Neyman-Pearson lemma whereas the approximate slope does not give rise to the most efficient test in the Neyman-Pearson sense.

We give a simple example for the approximate case.

Example: Let $X \sim N(0,\sigma^2)$. We want to test

$$H_0 : \sigma = \sigma_0$$

vs.

$$H_1 : \sigma = \sigma_1 \quad (\sigma_1 > \sigma_0).$$

To find the best test according to the approximate Bahadur slope, we need to find $w(\cdot)$ such that

$$\int_{-\infty}^{\infty} w(t)\, e^{tx}\, dt = e^{-\frac{1}{2}(x\,\delta)^2}$$

where

$$\delta^2 = \frac{1}{\sigma_0^{\,2}} - \frac{1}{\sigma_1^{\,2}}.$$

Solution: Under the constraint the $\int_{-\infty}^{\infty} w(t)\, dt = 1$, the function $w(t)$ maximizing $C^{(a)}$ is given by:

$$w(t) = \frac{1}{\sqrt{2\pi}\delta} \exp[-\frac{t^2}{2\delta^2}], \quad -\infty < t < \infty.$$

Section [A2]: Tests of normality and Some Derivative Processes: Simple null hypothesis of normality.

Let $X_1, X_2, \ldots, X_n$ be a set of iid observations with distribution function $F(\cdot)$ and mgf $m(t)$.

Although $m(t)$ itself characterizes the population distribution, as a *graphical* procedure, examination of $m_n(t)$ alone is not informative enough since $m_n(t)$ is a strictly convex function in t for any data set. However it is possible to characterize the normal distribution through some of the derivative processes associated with the mgf. Some such processes are described below.

[A2.1]: First Derivative of the Cumulant Generating Function:

The following is a characterization of normality.

A random variable X has $N(0,1)$ distribution iff $\frac{d}{dt}m(t) = t$ where $m(t)$ is the mgf of X.

Therefore to test the simple null hypothesis of normality (i.e. $H_0: m(t) = e^{t^2/2}$. departure of $m_n'(t) - t$ from zero may indicate non-normality where $m_n(t) = \int_{-\infty}^{\infty} e^{tx}\, dF_n(x)$.

We may now state the following

Theorem: A(2.1) Define

$$\sqrt{n}\, q_n(t) = \sqrt{n}\left\{ \frac{d}{dt}\log m_n(t) - \frac{d}{dt}\log m(t)\right\}$$

where $m(t)$ is the population (from which the sample is taken) mgf of the underlying random variable X and $m_n(t)$ is its empirical version in a random sample of size n. Then for $t \epsilon [-T, T]$, the stochastic process $\sqrt{n}\, q_n(t)$ converges weakly to a zero mean Gaussian process with covariance function:

$$L(s,t) = \sum_{j=0}^{1}\sum_{k=0}^{1} a_j(s) a_k(t)[m^{(j+k)}(s+t) - m^{(j)}(s) m^{(k)}(t)]$$

where

$$\frac{d^j}{dt^j} m(t) = m^{(j)}(t)$$

$$a_j(t) = \frac{\partial}{\partial U_j(t)} g(U_n(t)) \mid_{\varsigma(t)}$$

$$U_n(t) = (m_n(t), m_n'(t))'$$

$$\varsigma(t) = (m(t), m'(t))'$$

and $g(\cdot)$ is defined by:

$$g(U_n(t)) = \frac{d}{dt}\log m_n(t) = \frac{m_n'(t)}{m_n(t)}$$

and equivalently

$$g(\varsigma(t)) = \frac{m'(t)}{m(t)}.$$

Proof: Define $Z_n^{(j)}(t) = \sqrt{n}\,(m_n^{(j)}(t) - m^{(j)}(t)),\ j = 0,1.$ Multivariate central limit theorem can be used to prove the convergence of the finite dimensional distributions of $Z_n(t)$ and $Z_n'(t)$ to the finite dimensional distributions of the zero mean Gaussian processes $Z_0(t)$ and $Z_1(t)$ respectively with respective covariance structures given by:

$$K_0(s,t) = m(s+t) - m(s)m(t)$$

and

$$K_1(s,t) = m^{(2)}(s+t) - m'(s)m'(t).$$

Also by the multivariate central limit theorem, the finite dimensional distributions of the vector Gaussian process $(Z_0(t), Z_1(t))'$ whose covariance structure is given by:

$$cov(Z_0(s), Z_1(t)) = m'(s+t) - m(s)m(t)$$

and

$$cov(Z_i(s), Z_i(t)) = m^{(2i)}(s+t) - m^{(i)}(s)m^{(i)}(t)$$

where $i = 0,1$. Now for fixed t, a Taylor's series expansion can be used to write

$$\sqrt{n}\, q_n(t) = \sum_{i=0}^{1} Y_i(t) + R_n(t)$$

where

$$Y_i(t) = Z_i(t) \frac{\partial}{\partial U_i(t)} g(U_n(t)) \mid_{\varsigma(t)}$$

and

$$R_n(t) = \frac{1}{\sqrt{n}} \sum_{i=0}^{1} \sum_{j=0}^{1} Z_i(t) Z_j(t) \frac{\partial^2}{\partial U_i(t) \partial U_j(t)} g(U_n(t)) \mid_{\varsigma_n^*(t)}$$

where $| \varsigma_n^*(t) - \varsigma(t) | < | U_n(t) - \varsigma(t) |$.

First of all $\underset{t \epsilon [-T,T]}{Sup} R_n(t)$ converges to zero in probability. To see this note that

$$\underset{t \epsilon [-T,T]}{Sup} | R_n(T) | \leq \sum_{i=0}^{1} \sum_{j=0}^{1} v_i^n(t)\, v_j^n(t)\, u_{ij}^n(t)$$

where

$$v_i^n(t) = \underset{t \epsilon [-T,T]}{Sup} Z_i(t)$$

and

$$u_{ij}^n(t) = \underset{t \epsilon [-T,T]}{Sup} \frac{\partial^2}{\partial U_i(t) \partial U_j(t)} g(U_n(t)) \mid_{\varsigma_n^*(t)}.$$

Applying the law of iterated logarithm to each of $v_i^n(t)$ and $v_j^n(t)$ and noting that uniform convergence of each of $U_i(t)$ over the finite interval $[-T,T]$ implies that $u_{ij}^n(t)$ converges to a constant almost surely over $[-T,T]$, $\underset{t \epsilon [-T,T]}{Sup} R_n(t)$ is bounded by $O(n^{-1/2} \log \log n)$ almost surely. Tightness follows from a one-term Taylor expansion of $Y_i(t)$ which easily gives

$$E(\,Y(s)\,-\,Y(t)\,)\ \leq K\ |\,s-t\,|^2$$

where $K \geq 0$ is a constant. The theorem now follows by the application of Slutsky's theorem along with the application of the multivariate central limit theorem.

Therefore, one may reject the null hypothesis $H_0\ :m(t)\ = m_0(t)$ with an approximate level α if

$$|\sqrt{n}\ q_n(t)\ \ |\ \ > \sqrt{L_0(t,t)}z_{\alpha/2}$$

where $L_0(t,t)$ is now calculated from $L(t,s)$ by substituting $m(t)\ = m_0(t)$ in the expression for $L(t,s)$.

Now consider also as before a statistic of the type

$$S_n^{\,1}\ = \sqrt{n}\int_t w(t)\,q_n(t)\,dt$$

where the parameter t may be restricted to a finite interval [-T,T], $T\ <\ \infty$.

The asymptotic null distribution of $S_n^{\,1}$ is then Gaussian with mean zero and variance $\int\int w(s)w(t)L_0(s,t)dsdt$, by the central limit theorem. Denoting this distribution by $\overset{*}{P}_{H_0}$, the approximate Bahadur's slope for the sequence $\left\{S_n^{\,1}\right\}$ is given by:

$$C_1^{(a)}\ = \frac{[\int\left\{\frac{d}{dt}\log m_1(t)\ -\ \frac{d}{dt}\log m(t)\right\}w(t)dt\,]^2}{\int\int L_0(s,t)w(s)w(t)dsdt}$$

which is obtained by taking the limit in probability of

$$-\frac{2}{n}\log\overset{*}{P}_{H_0}\left\{\ |\,S_n^{\,1}\,|\ \ >\ \ |\,S_n^{\,1}\ (obs)\,|\ \right\}$$

where $S_n^{\,1}\ (obs)$ denotes an observed value of $S_n^{\,1}$ when the observations are sampled from the distribution specified by H_1.

A special case: Under the simple null hypothesis of normality (both mean and variance known), $m(t)\ = e^{t^2/2}$ which gives

$$C_1^{(a)}\ = \frac{[\int\left\{\frac{d}{dt}\log m_1(t)\ -\ t\right\}w(t)dt\,]^2}{\int\int e^{ts}(1\ +\ 3ts)w(t)w(s)dtds}.$$

For notational simplicity define

$$g(t)\ = \frac{d}{dt}\log m_1(t)\ -\ t$$

Note that given simple alternate hypothesis H_1, $g(t)$ is a known function. An optimum choice of

$w(\cdot)$

for this special case can be obtained by the maximisation of

$$\int g(t)w(t)dt \tag{P2}$$

with respect to $w(\cdot)$ such that

$$\int\int e^{ts}(1 + 3ts)w(t)w(s)dtds = 1 \tag{**}$$

$$\int w(t)dt < \infty$$

and

$$\int w(t)e^{ts}\, ds < \infty$$

for all t lying in an open interval around the origin.

Solution: Introducing variants $\delta w(t)$ to the solution of problem (P2) we have

$$\int \delta w(t)g(t)dt = 0 \tag{1}$$

Similarly the constraint (**) suggests, after ignoring the term involving $\delta w(t)\delta w(s)$,

$$\int\int e^{ts}(1 + 3ts)[w(t)\delta w(s) + w(s)\delta w(t)]dtds = 0$$

or

$$\int\int e^{ts}(1 + 3ts)w(t)\delta w(s)dtds + \int\int e^{ts}(1 + 3ts)w(s)\delta w(t)dtds = 0 \tag{2}$$

Now (2) implies

$$\int \psi(t)\delta w(t)dt = 0$$

where

$$\psi(t) = \int w(s)e^{ts}(1 + 3ts)ds. \tag{3}$$

(1) and (3) together imply that there exists α such that

$$g(t) = \alpha\psi(t)$$

or

$$\frac{g(t)}{\alpha} = \int w(t)e^{ts}\, ds + 3t\int sw(s)e^{ts}\, ds \tag{4}$$

$$= L_t[w(s)] + 3t\frac{d}{dt}L_t[w(s)]$$

where

$$L_t[(w(t)] = \int w(s)e^{ts}\, ds$$

To solve for α without loss of generality assume that

$$\int w(s)ds = 1$$

Then taking $t = 0$ in (4)

$$g(0) = \alpha \int w(s) ds = \alpha$$

or

$$\alpha = \frac{d}{dt} \log m_1(t) \mid_{t=0}$$

Thus given simple alternate hypothesis H_1, α is known. Thus the solution for problem (P2) is given by the equation

$$L_t[w(\cdot)] + 3t\frac{d}{dt}L_t[w(\cdot)] - \frac{g(t)}{g(0)} = 0$$

where

$$L_t[w(\cdot)] = \int w(s) e^{ts} ds$$

Now define

$$W(t) = L_t[w(\cdot)]$$

and

$$a(t) = \frac{g(t)}{g(0)}$$

where recall that

$$g(t) = \frac{d}{dt} \log m_1(t) - t$$

is a known function in t provided that H_1 is completely specified. Then we may rewrite the last differential equation as:

$$W(t) + 3tW'(t) - a(t) = 0$$

the solution for which also defines the solution for the problem (P2).

[A2.2]: Third derivative of the cumulant generating function:

One of the more striking defining properties of the normal distribution is that all its cumulants of order ≥ 3 are equal to zero. Therefore it is of natural interest to base a test on the third derivative of the cumulant generating function since it allows one to obtain a visual representation of all the cumulants beyond the second one.

Thus let $K(t) = log\ m(t)$. Then

$$\frac{d^3}{dt^3}K(t) = \frac{d^3}{dt^3} \log m(t)$$
$$= K_3 + K_4 \cdot t + K_5 \cdot \frac{t^2}{2!} + K_6 \frac{t^3}{3!} + \ldots$$

Clearly $m(t)$ corresponds to a normal distribution if and only if $\frac{d^3}{dt^3}K(t) \equiv 0$.

Let us define

$$r_n(t) = f(U_n(t)) - f(\varsigma(t))$$

where

$$U_n(t) = (m_n(t), m_n'(t), m_n^{(2)}(t), m_n^{(3)}(t))'$$
$$\varsigma(t) = E_F\, U_n(t)$$
$$f(U_n(t)) = \frac{d^3}{dt^3} K_n(t)$$
$$= \frac{m_n^{(3)}(t) m_n^2(t) - 3m_n^{(2)}(t) m_n'(t) + 2(m_n')^3}{(m_n(t))^3}$$

and equivalently

$$f(\varsigma(t)) = \frac{d^3}{dt^3} K(t)$$

and for every fixed t define $a_j(t) = \frac{\partial}{\partial U_{jn}}(t) f(U_n(t)) \mid_{\varsigma(t)};$ where $U_{jn}(t) = \frac{d^j}{dt^j} m_n(t)$ $\quad j = 0,1,2,3$. We may now state the following

Theorem: A(2.2) For $t \in [-T, T]$ the stochastic process $\sqrt{n}\, r_n(t)$ converges weakly to a zero mean Gaussian process with covariance structure

$$V(s,t) = \sum_{j=0}^{3} \sum_{k=0}^{3} \left[m^{(j+k)}(s+t) - m^{(j)}(s) m^{(k)}(t) \right] \cdot a_j(s) a_k(t)$$

The method of proof is the same as in the previous theorem with appropriate changes.

Thus in particular, under the simple null hypothesis of normality, $f(\varsigma(t)) \equiv 0$ so that in this case, significant departure of $\sqrt{n}\frac{d^3}{dt^3}K_n(t)$ from zero will indicate non-normality of the data.

As with any good graphical procedure, the study of $\sqrt{n}\frac{d^3}{dt^3}K_n(t)$ (in a neighbourhood of 0), as a function of t offers much other valuable information. For example, the value of $r_n(t)$ at $t = 0$ gives the third cumulant which apart from a multiplicative constant is equivalent to the skewness coefficient of the data, the slope at $t = 0$ is the fourth cumulant (equivalent to the kurtosis coefficient), the curvature at $t = 0$ is the fifth cumulant an so on. As a point of interest it should be noted that a graphical procedure or a test based on $r_n(t)$ may be dominated by the skewness present in the data set. In particular, if there are extreme observations in the data set in one particular direction, then the procedure will be sensitive to such outliers. Therefore if detection of outliers is of interest, and the statistician is not concerned with just the majority of the data set, then this procedure can be particularly useful.

To compare this procedure with the existing normal probability plot method, we consider the chi square family of distribution with varying degrees of freedom. We take samples of size 200 and plot $r_n(t) = \frac{d^3}{dt^3}K_n(t)$ along with its $\pm 2\ \sigma_n$ and $\pm 1\ \sigma_n$ bounds as calculated from the asymptotic null distribution of $r_n(t)$, where σ_n denotes s.d., vs. t, on the same graph paper. To make comparison, we also study the normal probability plots for the same data sets. Although we have one data set for each of the degrees of freedom, it is clear that $r_n(t)$ clearly offers a very attractive method for detecting departure from normality. The graphs are presented in the appendix (pages A1-A8).

Section [A3]: Composite null hypothesis of normality. The above results easily generalize to the composite null hypothesis situation. Thus consider the hypothesis (for unknown μ and σ)

$$H_0 : f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{1}{2\sigma^2}(x-\mu)^2} \qquad -\infty < x,\mu < \infty,\ 0 < \sigma < \infty.$$

vs.

$$H_1 : \textit{negation of } H_0$$

In this case define the studentized mgf and its derivatives as:

$$\tilde{m}_n(t) = \frac{1}{n}\sum_{j=1}^{n} e^{t(\frac{X_j - \bar{X}}{s})}$$

and

$$\tilde{m}^{(i)}{}_n(t) = \frac{d^i}{dt^i}\tilde{m}_n(t) = \frac{1}{n}\sum_{j=1}^{n}\left\{\frac{X_j - \bar{X}}{s}\right\}^j e^{t(\frac{X_j - \bar{X}}{s})}$$

where $\bar{X}$ and s are respectively the sample mean and the sample standard deviation in a sample of size n. Correspondingly we define the respective population quantities as:

$$\tilde{m}(t) = E\left[e^{(\frac{X-\mu}{\sigma})}\right]$$

and

$$\tilde{m}^{(i)} = \frac{d^i}{dt^i}\tilde{m}(t).$$

Restricting the parameter t in a finite interval $[-T,T]$ for an arbitrary positive number T, we have the following

Lemma A(3.1): Under H_0, and for $t \epsilon\ [-T,T]$, $0 < T < \infty$, as $n \to \infty$, the vector stochastic process

$$\tilde{Z}_n(t) = (\tilde{Z}_{0n}(t), \tilde{Z}_{1n}(t), \tilde{Z}_{2n}(t), \tilde{Z}_{3n}(t))'$$

where

$$\tilde{Z}_{jn}(t) = \sqrt{n}\left\{\tilde{m}_n^{(j)}(t) - \tilde{m}^{(j)}(t)\right\}$$

converges weakly to a zero mean vector Gaussian process

$$Z(t) = (Z_0(t), Z_1(t), Z_2(t), Z_3(t))'$$

with covariance function defined by:

$$E(Z_i(s)Z_j(t)) = \tilde{C}_{ij}(s,t), \quad i,j=0,1,2,3;$$

where

$$\begin{aligned}\tilde{C}_{jj}(t,s) &= \tilde{m}^{(2j)}(t+s) - \tilde{m}^{(j)}(t)\tilde{m}^{(j)}(s)\\ &- (\tilde{m}^{(j)}(s)+j\tilde{m}^{(j-1)}(s))\tilde{m}^{(j+1)}(t)-(\tilde{m}^{(j)}(t)+j\tilde{m}^{(j-1)}(t))\tilde{m}^{(j+1)}(s)\\ &-(j\tilde{m}^{(j)}(t)+t\tilde{m}^{(j+1)}(t))(\tilde{m}^{(j+2)}(s)-\tilde{m}^{(j)}(s))/2-(j\tilde{m}^{(j)}(s)+s\tilde{m}^{(j+1)}(s))(\tilde{m}^{(j+2)}(t)-\tilde{m}^{(j)}(t))/2\\ &+ (t\tilde{m}^{(j)}(t)+j\tilde{m}^{(j-1)}(t))(s\tilde{m}^{(j)}(s)+j\tilde{m}^{(j-1)}(s))\\ &+ \frac{3}{4}(t\tilde{m}^{(j-1)}(t)+j\tilde{m}^{(j)}(t))(s\tilde{m}^{(j+1)}(s)+j\tilde{m}^{(j)});\end{aligned}$$

and

$$\begin{aligned}\tilde{C}_{ij}(t,s) &= \tilde{m}^{(i+j)}(s+t) - \tilde{m}^{(i)}(t)\tilde{m}^{(j)}(s)\\ &- \frac{1}{2}\left[j\tilde{m}^{(j)}(s)+s\tilde{m}^{(j+1)}(s)\right]\cdot\left[\tilde{m}^{(i+2)}(t)-\tilde{m}^{(i)}(t)\right]\\ &-\frac{1}{2}\left[i\tilde{m}^{(i)}(t)+t\tilde{m}^{(i+1)}(t)\right]\cdot\left[\tilde{m}^{(j+2)}(s)-\tilde{m}^{(j)}(s)\right]\\ &- \left[s\tilde{m}^{(j)}(s)+j\tilde{m}^{(j-1)}(s)\right]\tilde{m}^{(i+1)}(t)-\left[\tilde{m}^{(i)}(t)+i\tilde{m}^{(i-1)}(t)\right]\tilde{m}^{(j+1)}(s)\\ &+ \frac{3}{4}\left[j\tilde{m}^{(j)}(s)+s\tilde{m}^{(j+1)}(s)\right]\\ &+ \left[s\tilde{m}^{(j)}(s)+j\tilde{m}^{(j-1)}(s)\right]\cdot\left[t\tilde{m}^{(i)}(t)+i\tilde{m}^{(i-1)}(t)\right]\end{aligned}$$

where under the null hypothesis of normality, $\tilde{m}(t) = e^{t^2/2}$ etc.

The proof of the lemma is outlined below.

Proof: The proof follows by noting that the stochastic process $\tilde{Z}_{kn}(t)$ can be written as:

$$\tilde{Z}_{kn}(t) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n} W_j(t) + R_n(t)$$

where

$$\begin{aligned}W_k(t) &= \left(\frac{X_j - \mu}{\sigma}\right)^k e^{it\left(\frac{X_j-\mu}{\sigma}\right)} - \tilde{m}^{(k)}(t)\\ &- \frac{1}{2}\left[k\tilde{m}^{(k)}(t) + t\tilde{m}^{(k+1)}(t)\right]\cdot\left[\left(\frac{X_j - \mu}{\sigma}\right)^2 - 1\right]\\ &- \left(\frac{X_j - \mu}{\sigma}\right)\left[t\tilde{m}^{(k)}(t) + k\tilde{m}^{(k-1)}(t)\right]\end{aligned}$$

and the remainder term $R_n(t)$ converges to zero in probability, uniformly in t over the finite interval $[-T,T]$. Tightness follows by a Taylor series expansion of $W(s)$ so that

$$E(W(s) - W(t))^2 \leq K \mid s-t \mid^2$$

where $K>0$ is a finite constant. The lemma follows by the application of the multivariate central limit along with Slutsky's theorem.

Development of the test statistic:

Define

$$\tilde{U}_n(t) = \left(\tilde{U}_{0n}(t),\ldots,\tilde{U}_{3n}(t)\right).$$

where $\tilde{U}_{jn}(t) = \tilde{m}^{(j)_n}(t)$, $j = 1,2,..$ and $\tilde{U}_{0n}(t) = \tilde{m}_n(t)$. Similarly define

$$\tilde{\varsigma}(t) = \left(\tilde{\varsigma}_0(t),\ldots,\tilde{\varsigma}_3(t)\right)'$$

where $\varsigma_j(t) = \tilde{m}^{(j)}(t)$, $j = 1,2,..$ and $\varsigma_0(t) = \tilde{m}(t)$. Finally define

$$f_j(\tilde{U}_n(t)) = \frac{d^j}{dt^j} log\ \tilde{m}_n(t),$$

$$f_j(\tilde{\varsigma}(t)) = \frac{d^j}{dt^j} log\ \tilde{m}(t),$$

and

$$\tilde{r}^j{}_n(t) = \sqrt{n}\left[f_j(\tilde{U}_n(t)) - f_j(\tilde{\varsigma}(t))\right]$$

where j is an integer. Using the above lemma we have the following

Theorem A(3.1): Under H_0 for $t \epsilon\ [-T,T]$ as $n \rightarrow \infty$, the stochastic process $\tilde{r}^1{}_n(t)$ weakly converges to the zero mean Gaussian process $Z_1(t)$ whose covariance function is given by:

$$\begin{aligned} E(Z_1(s)Z_1(t)) &= L_1(s,t) \\ &= \frac{1}{\tilde{m}(t)\tilde{m}(s)}\tilde{K}_1(s,t) - \frac{\tilde{m}'(s)}{\tilde{m}(t)\tilde{m}^2(s)}\tilde{C}_{01}(t,s) \\ &- \frac{\tilde{m}'(t)}{\tilde{m}(s)\tilde{m}^2(t)}\tilde{C}_{01}(s,t) + \tilde{K}_0(s,t)\frac{\tilde{m}'(s)\tilde{m}'(t)}{\tilde{m}^2(t)\tilde{m}^2(s)}\tilde{K}_0(s,t) \end{aligned}$$

Theorem A(3.2): Under H_0 for $t \epsilon\ [-T,T]$, as $n \rightarrow \infty$ the stochastic process $r^3{}_n(t)$ weakly converges to the zero mean Gaussian process $Z_3(t)$ whose covariance function is given by:

$$E(Z_3(s)Z_3(t)) = \sum_{j=0}^{3}\sum_{k=0}^{3} a_j(s)a_k(t)\tilde{C}_{jk}(s,t)$$

where

$$a_j(t) = \frac{\partial}{\partial U_{jn}(t)} f_3(\tilde{U}_n(t)) \mid_{\tilde{\varsigma}(t)}$$

where under H_0, $\tilde{m}(t) = e^{t^2/2}$ etc.

Several test procedures such as chi-square type statistics (Koutrovellis 1981), maximal deviation statistics (Csorgo 1986), simple projection statistics (Csorgo 1986), Cramer-von Mises type statistics, may be developed using our basic stochastic processes of Theorems 3.1 and 3.2. However it is of particular interest to obtain graphical procedures for assessing the characteristics of an underlying population distribution. To study these procedures, a large number of data sets from a variety of distributions were generated using NAG library routines although only a few of these results can be shown here. In particular we show graphs for 6 samples of size 50 each generated from N(0,1), Cauchy, and χ^2_2 distributions. For each sample, each of the test statistics (namely, $\tilde{r}^1_n(t)$, and $\tilde{r}^3_n(t)$ as in theorems A[3.1] and A[3.2] respectively) and the corresponding $\pm\sigma_n(t)$ and $\pm 2\sigma_n(t)$ limits were plotted against t. All these calculations were done for the standardized data (standardized with respect to the sample mean and the sample standard deviation). Departure from normality may be suggested whenever the plot of the statistic, as a function of t crosses the corresponding $\pm 2\sigma_n(t)$ limits. The graphs are presented in the appendix (pages A9-A14).

Not surprisingly the skewness in the data tends to affect these graphs quite markedly. This is particularly evident from the graph for sample # 6 from N(0,1). The histogram for this data (page A15) shows that there are two outliers present in this data set. The presence of these two large negative observations is creating negative skewness which is clearly depicted in the graph of $\frac{d^3}{dt^3} \log \tilde{m}_n(t)$. Also it is to be noted here that while plotting the graph of $\frac{d}{dt} \log \tilde{m}_n(t) - t$, the asymptotic variance vanishes at $t = 0$ which suggests examination of this process for values of t away from zero, although, for large values of t, the behavior of this process will be dominated by the largest (smallest for negative values of t) order statistic in the data set.

[B] Use of the empirical characteristic function.

A test of normality may be conducted in two stages. First one tests for symmetry. If the hypothesis of symmetry is not rejected, one proceeds to test for the shape of the underlying distribution.

Since a distribution function is symmetric (around zero) iff the characteristic function is real, a test of symmetry may be based on the imaginary part of the characteristic function (also see Feuerverger and Mureika 1977).

Thus let

$$\phi(t) = E_F(e^{itX}) = re(\phi(t)) + i\, im(\phi(t))$$

and

$$\phi_n(t) = \int e^{itx} dF_n(x)\, dx = re(\phi_n(t)) + i\, im(\phi_n(t))$$

define the characteristic functions in the population and in the sample where $F(\cdot)$ and $F_n(\cdot)$ are respectively the underlying population distribution function and the empirical distribution function based on a set of iid observations $X_1,\ldots,X_n$.

To develop a test for symmetry for the composite null hypothesis of symmetry, (unknown location) one could consider (a functional based on)

$$\tilde{Y}_n(t) = \frac{1}{n}\sum_{i=1}^{n}\sin\left[t\left(\frac{X_j - \bar{X}}{s}\right)\right], \quad t\epsilon R.$$

However if the first $2r+1$ moments of the underlying variable X exist, then

$$E(\sin X) = t\mu - \frac{t^3}{3!}\mu_3' \cdots + (-1)^{2r+1}\frac{t^{2r+1}}{(2r+1)!}\mu_{2r+1}' + 0(t^{2r+3})$$

where $\mu_j' = E(X^j)$, $\mu_1' = \mu$, $j = 2,3,\ldots$ which gives

$$\frac{d^3}{dt^3}E(\sin X) = -\mu_3' + \frac{t^2}{2!}\mu_5' + \ldots$$

examination of which in a small neighbourhood of zero (as a function of t) is equivalent to the examination of the skewness term (or μ_3') only, although for values of t away from zero, the above function provides more information about the population. Therefore one may consider the stochastic process

$$\frac{d^3}{dt^3}\tilde{Y}_n(t) = -\frac{1}{n}\sum_{i=1}^{n}\left[\frac{X_j - \bar{X}}{s}\right]^3 \cos\left[t\left(\frac{X_j - \bar{X}}{s}\right)\right]$$

Under the null hypothesis of symmetry (about an unknown constant μ), where we assume $E_F(X) = \mu$ and $Var(X) = \sigma^2$, we have $E_F \cos t\left(\frac{X - \mu}{\sigma}\right) = \tilde{\phi}(t) = E_F\left(e^{it\frac{X-\mu}{\sigma}}\right)$.We have the following

Theorem (B1) For $t\epsilon[-T,T]$, the stochastic process $\tilde{S}_n(t) = -\sqrt{n}\frac{d^3}{dt^3}\tilde{Y}_n(t)$ converges weakly to a zero mean Gaussian process with covariance function

$$K(s,t) = \frac{1}{2}\left[\tilde{\phi}^{(6)}(s+t) + \tilde{\phi}^{(6)}(s-t)\right] - \left[3\tilde{\phi}^{(2)}(t) - t\tilde{\phi}^{(3)}(t)\right]\tilde{\phi}^{(4)}(t)$$
$$- \left[3\tilde{\phi}^{(2)}(s) - s\tilde{\phi}^{(3)}(s)\right]\tilde{\phi}^{(4)}(s) + \left[3\tilde{\phi}^{(2)}(s) - s\tilde{\phi}^{(3)}(s)\right]\cdot\left[3\tilde{\phi}^{(2)}(t) - t\tilde{\phi}^{(3)}(t)\right].$$

Proof: The proof follows by a Taylor's series expansion of $\tilde{S}_n(t)$ and by noting that

$$e^{it\left(\frac{X_j - \bar{X}}{s}\right)} = e^{it\left(\frac{X_j - \bar{X}}{\sigma}\right)}\left\{1 + it\left(\frac{X_j - \bar{X}}{\sigma}\right)\left(\frac{\sigma}{s} - 1\right) + \frac{(it)^2}{2!}\left(\frac{X_j - \bar{X}}{\sigma}\right)^2\left(\frac{\sigma}{s} - 1\right)^2\theta_j\right\}, \qquad \text{(i)}$$

where $|\theta_j| \le 1$,

$$\sqrt{n}\,(\frac{\sigma}{s}-1) = -\frac{1}{\sigma(s+\sigma)}\cdot\frac{n}{n-1}\Big[n^{\frac{-1}{2}}\sum_{j=1}^{n}[(X_j-\mu)^2-\sigma^2] - \sqrt{n}\,(\bar{X}-\mu)^2 + n^{-\frac{1}{2}}\sigma^2\Big], \qquad \text{(ii)}$$

and thus writing

$$\tilde{S}_n(t) = \frac{1}{n}\sum_{j=1}^{n} Q_j(t) + R_n(t)$$

where the remainder term $R_n(t)$ converges to zero in probability an uniformly in t over an arbitrary finite interval $[-T,T]$ and $cov(Q_j(s),Q_j(t)) = K(s,t)$. The tightness follows by another one term Taylor series expansion of $Q(t)$ which gives

$$E(Q(s) - Q(t))^2 \le K\,|s-t|^2$$

where K is a finite positive constant.

Let us now turn our attention to testing normality in a symmetric class of distributions with the first two moments existing. Note that for a general distribution $F(\cdot)$, if the first $2r$ moments exist, then

$$re\,\phi(t) = 1 - \frac{t^2}{2!}\mu_2' + \frac{t^4}{4!}\mu_4' - \dots (-1)^r\frac{t^{2r}}{(2r)!}\mu_{2r}' + O(t^{2r+2})$$

so that

$$\frac{d^4}{dt^4} re\,\phi(t) = \mu_4' - \frac{t^2}{2!}\mu_6' + \dots + \frac{d^4}{dt^4}O(t^{2r+2})$$
$$= \mu_4' + O(t^2)$$

where $\mu_j' = E_F(X^j)$, $j = 1,2,\dots$ This implies that examination of the fourth derivative of the real part of the characteristic function for values of t very near zero is 'equivalent' to a kurtosis test although examination of this function for other values of t potentially offers more information.

Thus using the previous notation, define

$$\tilde{C}_n(t) = \sqrt{n}\left\{\frac{d^4}{dt^4} re\,\tilde{\phi}_n(t) - \frac{d^4}{dt^4} re\,\tilde{\phi}(t)\right\}$$

where $\tilde{\phi}_n(t) = \frac{1}{n}\sum_{j=1}^{n} e^{it(\frac{X_j-\bar{X}}{s})}$.

Using arguments exactly as before, we have the following

Theorem (B2): For $t\epsilon[-T,T]$, where T is an arbitrary finite positive constant, the stochastic process $\tilde{C}_n(t)$ converges weakly to a zero mean stochastic process with covariance structure:

$$L(s,t) = \frac{1}{2}\left[\tilde{\phi}^{(8)}(s+t) + \tilde{\phi}^{(8)}(s-t)\right] + (4\frac{\mu_4}{\sigma^4} - 9)\tilde{\phi}^{(4)}(t)\tilde{\phi}^{(4)}(s)$$
$$- 2\tilde{\phi}^{(4)}(t)\tilde{\phi}^{(6)}(s) - 2\tilde{\phi}^{(4)}(s)\cdot\tilde{\phi}^{(6)}(s)$$

where under the hypothesis of normality, $\frac{\mu_4}{\sigma^4} = 3$, $\tilde{\phi}(t) = e^{-t^2/2}$ etc.

As before, although several types of test statistics may be constructed based on these stochastic processes whose merrits will depend on the particular functional used, we shall emphasize graphical procedures particularly.

In particular we consider use of the process described in theorem (B2) for testing for the shape of the distribution. The same data sets as in the previous examples are used for this purpose and the graphs are presented in the appendix (pages A16-A18). Note that the effect of the negative outliers in sample # 6 from the $N(0,1)$ distribution is not visible, which is not so surprising since the dominating term in the expansion of $\frac{d^4}{dt^4} re\,\tilde{\phi}(t)$ is the fourth central moment of the distribution and not the third central moment which accounts for the skewness effect in the distribution.

All the stochastic processes considered above may be used formally for testing purposes. Clearly the power of such tests will depend on the particular functional used. In the appendix (page A19) we present some very modest power calculations based on simulations, for the test based on the third derivative of the cumulant generating function for the standerdized data. We consider tests carried out at single t points. Here the null hypothesis of interest is normal with unknown mean and variance. (Details of these calculations may be obtained from the author). One may refer to the power calculations done by Shapiro et al (1968) to compare these performances over other available tests. Based on these limited simulations one can see that the test based on the third derivative of the cumulant generating function is highly informative and rather promising. Tests based on the first derivative of the cumulant generating function and the empirical moment generating function were omitted from our simulations study, although we note that the graphs for these processes also showed promising results.

Acknowledgements: A major part of this work is based on the my Ph.D. thesis at the University of Toronto, Canada. Some new results were added during the my visit to the Department of

Statistics at the University of North Carolina at Chapel Hill. I also wish to thank Dr. Andrey Feuerverger and Dr. Jan Beran for many valuable discussions and suggestions.

References

Bahadur, R. R. (1960) Stochastic comparison of tests. *Ann. Math. Stat. 31, 276-295.*

Bahadur, R. R. (1967) Rates of convergence of estimates and test statistics. *Ann. Math. Stat. 38, 303-324.*

Bahadur, R. R. (1971) Some limit theorems in Statistics. *Conference Board of the Mathematical Sciences.*

Billingsley, P. (1971) Weak convergence of measures: Applications in probability. *Conference Board of Mathematical Sciences.*

Csorgo, S. (1980) The empirical moment generating function. *Colloquia Mathematica Soietatis Janos Bolyai 32, Nonparametric Statistical Inference, Budapest (Hungary).*

Csorgo, S. (1981) Limit behaviour of the empirical characteristic function. *Ann. Probab. 9, 130-144.*

Csorgo, S. (1984) Testing by empirical characteristic function: A survey. *Asymptotic Statistics 2, 45-56.*

Csorgo, S. (1986) Testing for normality in arbitrary dimension. *Ann. Stat. 14, 708-723.*

D'Agostino, R. B. and Stephens, M. A. (1986) Goodness-of-fit techniques. *Marcel Dekker, Inc. N.Y.*

Epps, T. W. (1987) Testing that a stationary time series is Gaussian. *Ann. Stat. 15, 1683-1698.*

Epps, T. W. and Pulley, L. B. (1983) A test for normality based on the empirical characteristic function. *Biomatrica 70, 723-726.*

Epps, T. W. and Pulley, L. B. (1986) A test of exponentiality vs monotone-hazard alternatives derived from the empirical characteristic function. *J. Roy. Stat. Soc. B 48, 206-213.*

Epps, T. W., Singleton, K. J. and Pulley, L. B. (1982) A test for separate families of distributions based on the empirical moment generating function. *Biometrika 69, 391-399.*

Feigin, P. D. and Heathcote, C. R. (1976) The empirical characteris function and the Cramer-von Mises statistic. *Indian J. Stat. A 38, 309-325.*

Feuerverger, A. (1987) New derivations of the maximum likelihood estimator and the likelihood ratio test. *Comm. Stat. Theor. Meth. 16, 1241-1251.*

Feuerverger, A. and McDunnough, P. (1981) On the efficiency of the empirical characteristic function procedures. *J. Roy. Stat. Soc. B 43, 20-27.*

Feuerverger, A. and McDunnough, P. (1981) On some Fourier methods of inference. *J. Amer. Stat. Assoc. 76, 379-387.*

Feuerverger, A. and Mureika, R. A. (1977) The empirical characteristic function and its applications. *Ann. Stat. 5, 88-97.*

Gnanadesikan, R. (1977) Methods for statistical data analysis for multivariate observations. *John Wiley & Sons.*

Hall, P. and Welsh, A. H. (1983) A test for normality based on the empirical characteristic function. *Biometrika 70, 485-489.*

Heathcote, C. R. (1972) A test of goodness of fit for symmetric random variables. *Austr. J. Stat. 14, 172-171.*

Kellermeier, J. (1980) The empirical characteristic function and large sample hypothesis testing. *J. Mult. Anal. 10, 78-87.*

Koutrouvelis, I. A. (1980) A goodness of fit test of simple hypotheses based on the empirical characteristic function. *Biometrika 67, 238-240.*

Koutrouvelis, I. A. and Kellermeier, J. (1981) A goodness of fit test based on the empirical characteristic function when the parameters must be estimated. *J. Roy. Stat. Soc. B 43, 173-176.*

Mardia, K. V. (1980) Tests of univariate and multivariate normality. *Handbook of Stat. 1, (ed. P. R. Krishnaiah), 279-320.*

Murota, K. and Takeuchi, K. (1981) The studentized empirical characteristic function and its application to test for the shape of the distribution. *Biometrika 68, 55-65.*

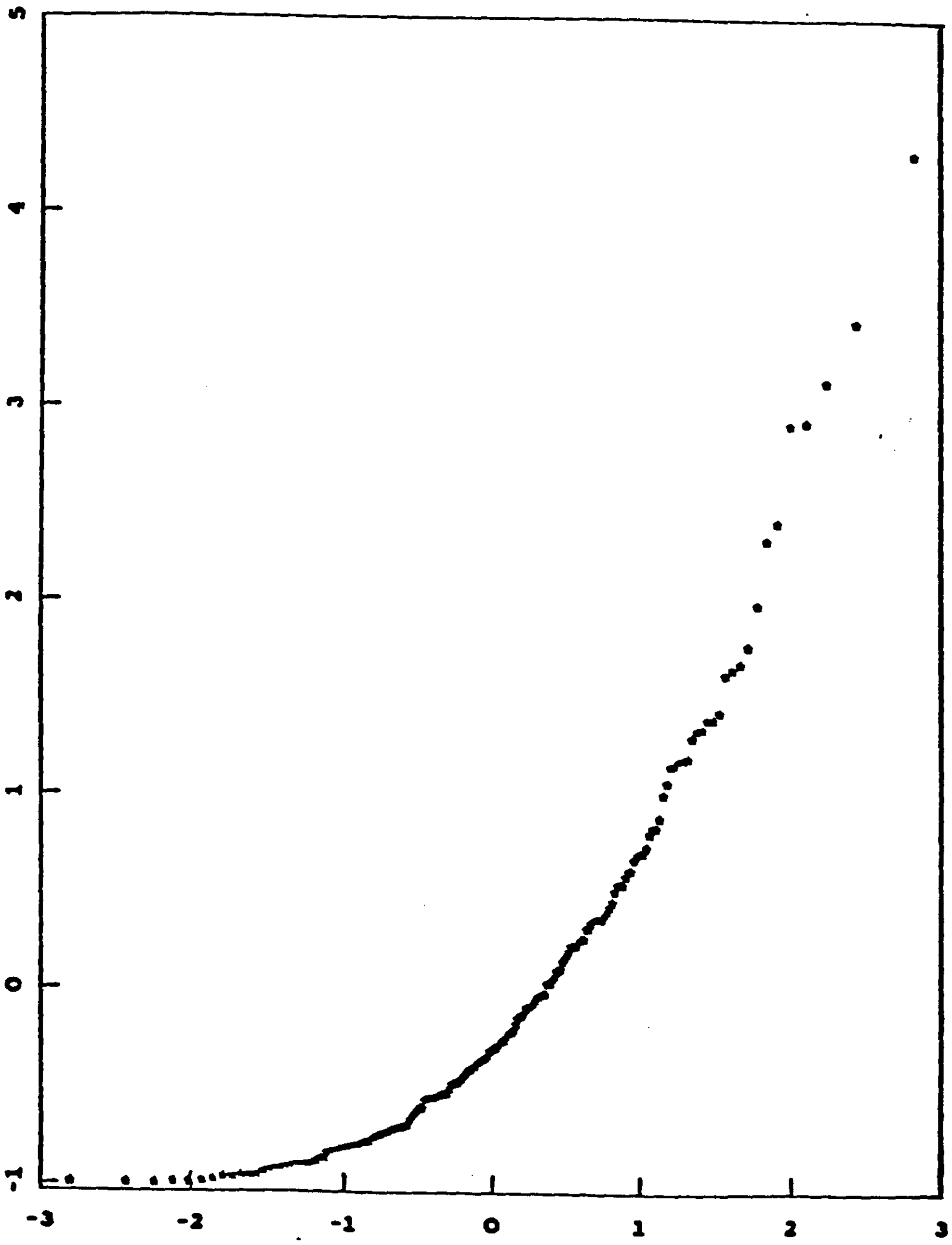
Serfling, R. J. (1980) Approximation theorems of mathematical statistics. *John Wiley & Sons. N.Y.*

Shapiro, S. S. and Wilk, M. B. (1965) An analysis of variance test for normality (Complete samples). *Biometrika 52, 591-611.*

Shapiro, S. S., Wilk, m. B. and Chen, H. J. (1968) A comparative study of various tests of normality. *J. Amer. Stat. Assoc. 63, 1343-1372.*

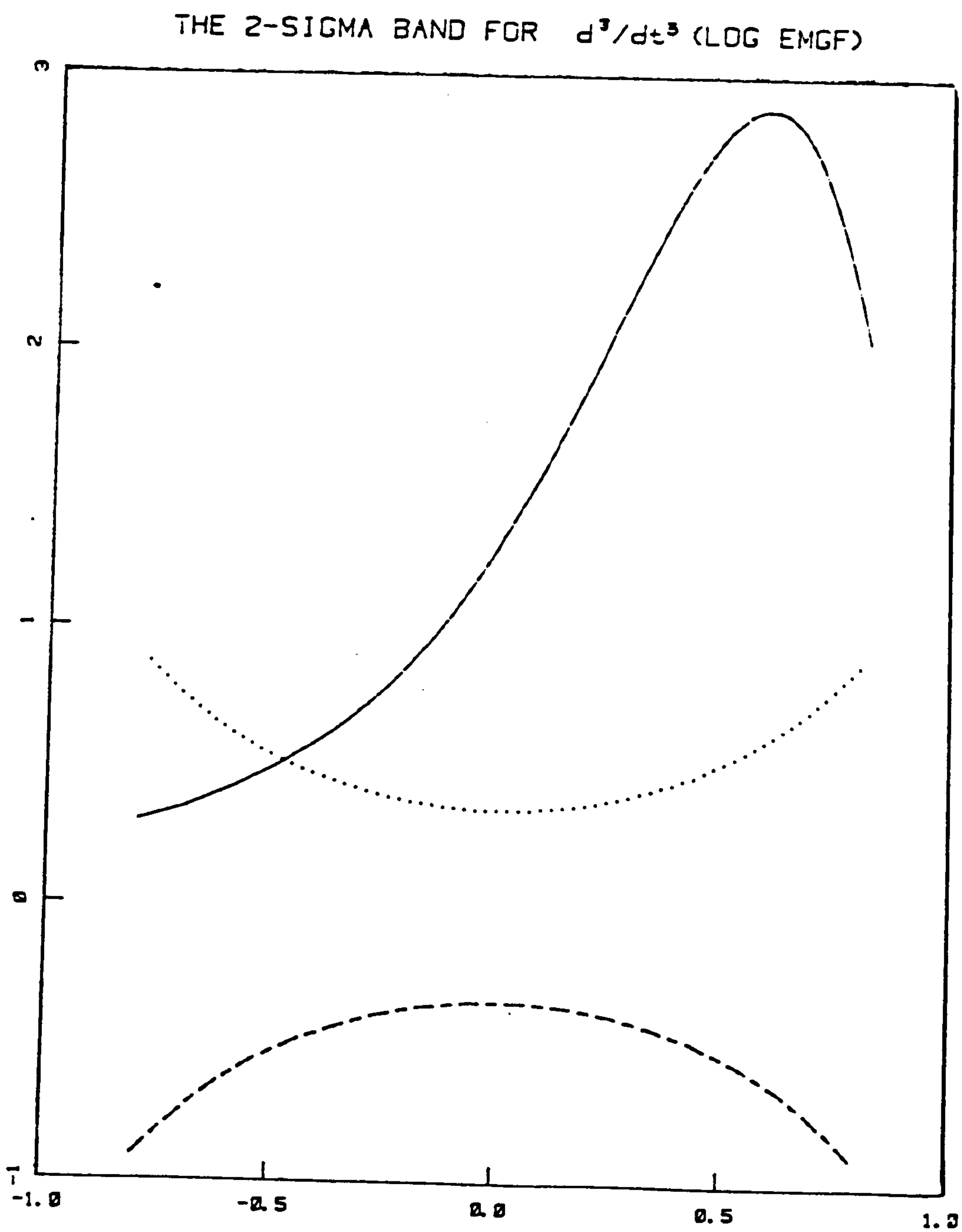
Welsh, A. H. (1984) A note on the scale estimates based on the empirical characteristic function and their application to test for normality. *Stat. & Prob. Letters 2, 345-348.*

Normal Probability Plot



Sample of size 200 from a chi-square population

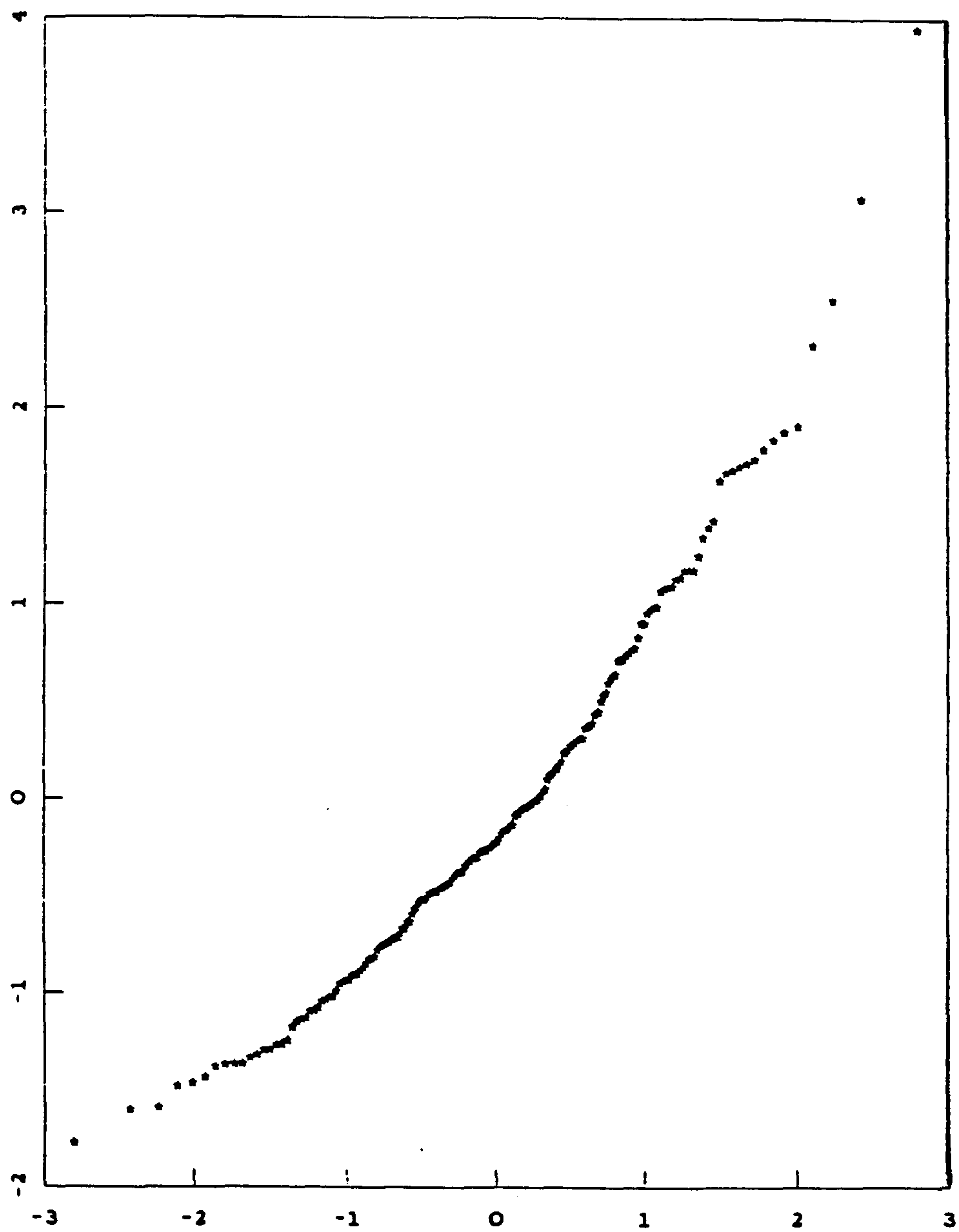
Sample # 1: Chi-square (2 d.f.)



Sample of size 200 from a chi-square population

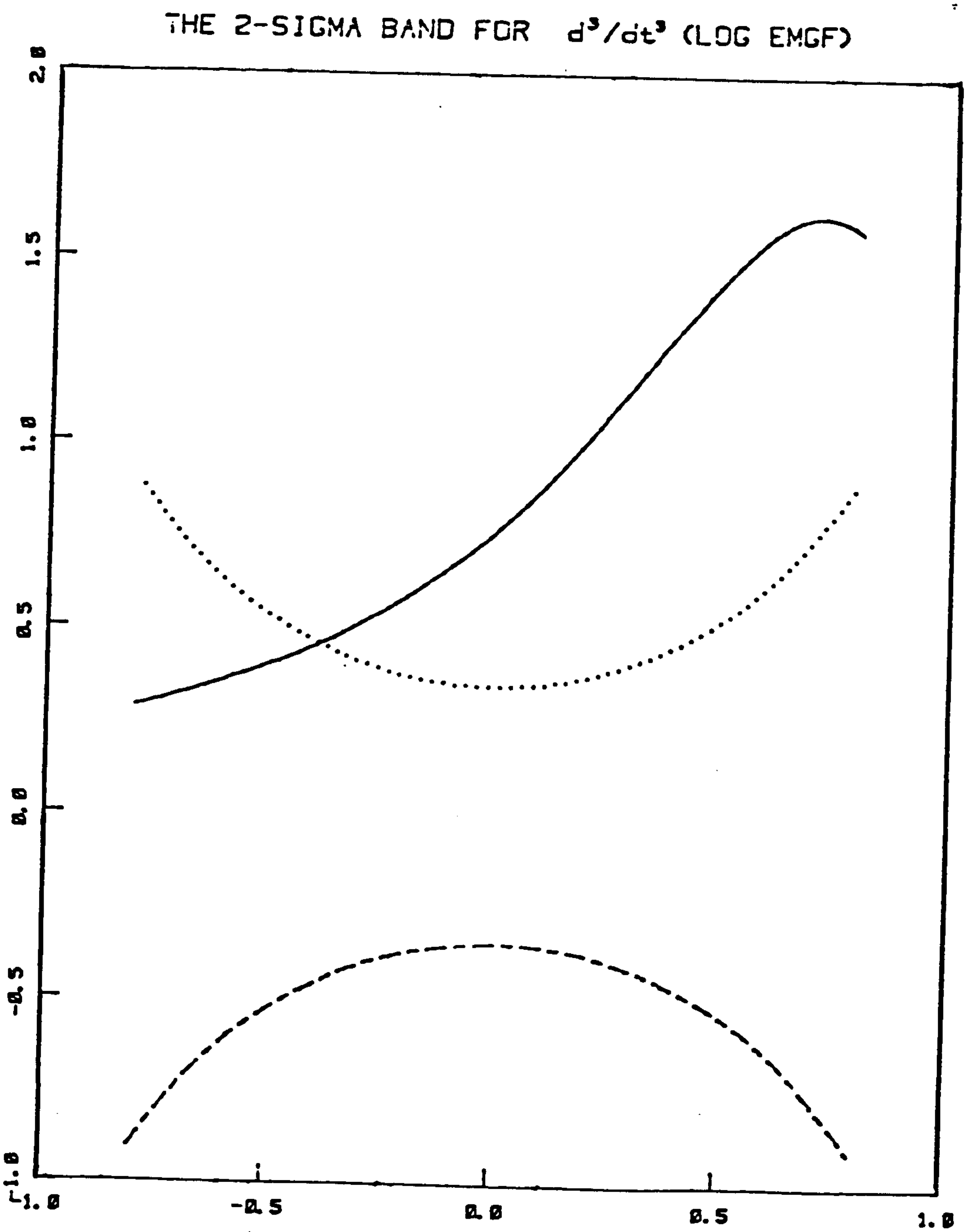
Sample # 1: Chi-square (2 d.f.)

Normal Probability Plot



Sample of size 200 from a chi-square population

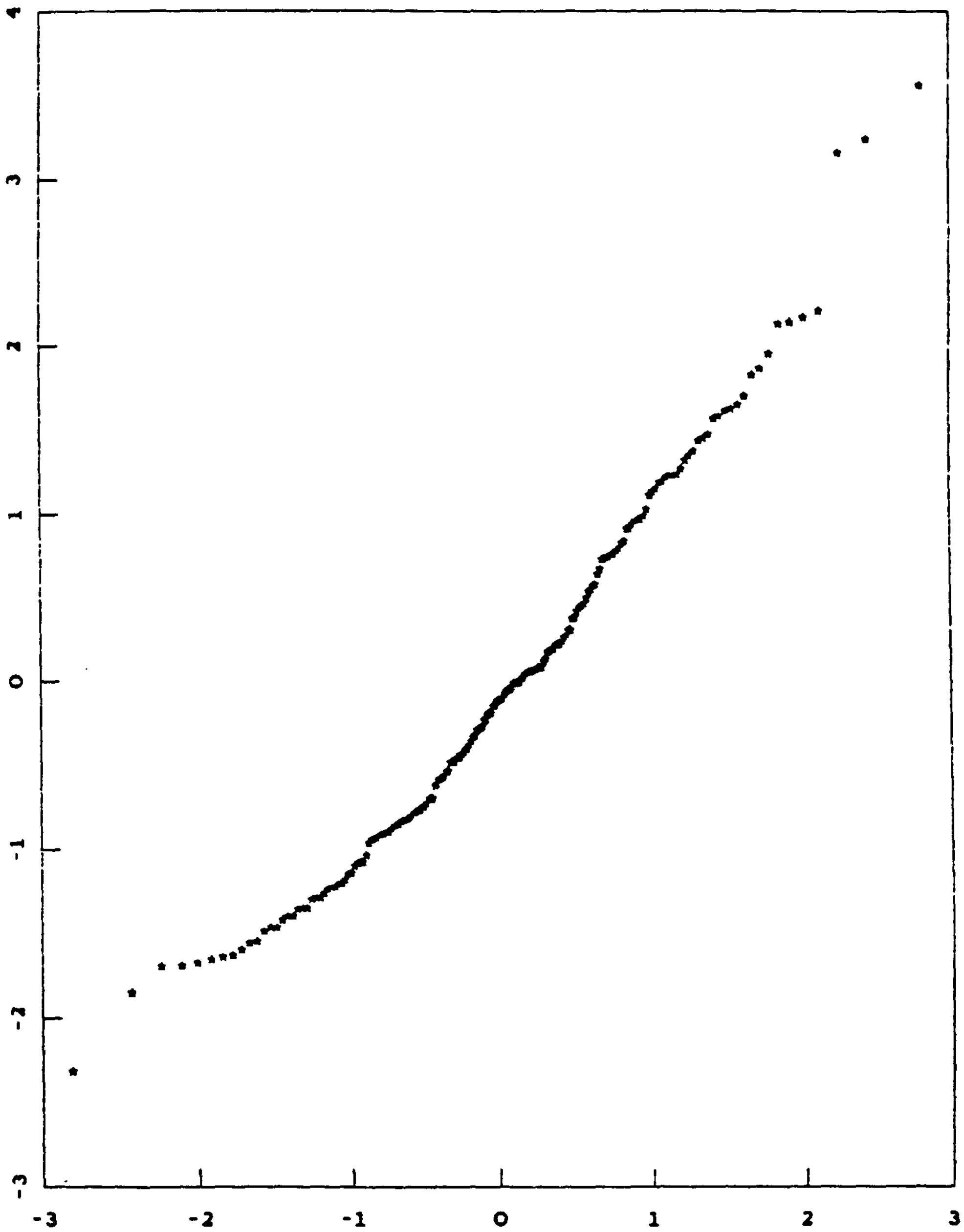
Sample # 2: Chi-square (10 d.f.)



Sample of size 200 from a chi-square population

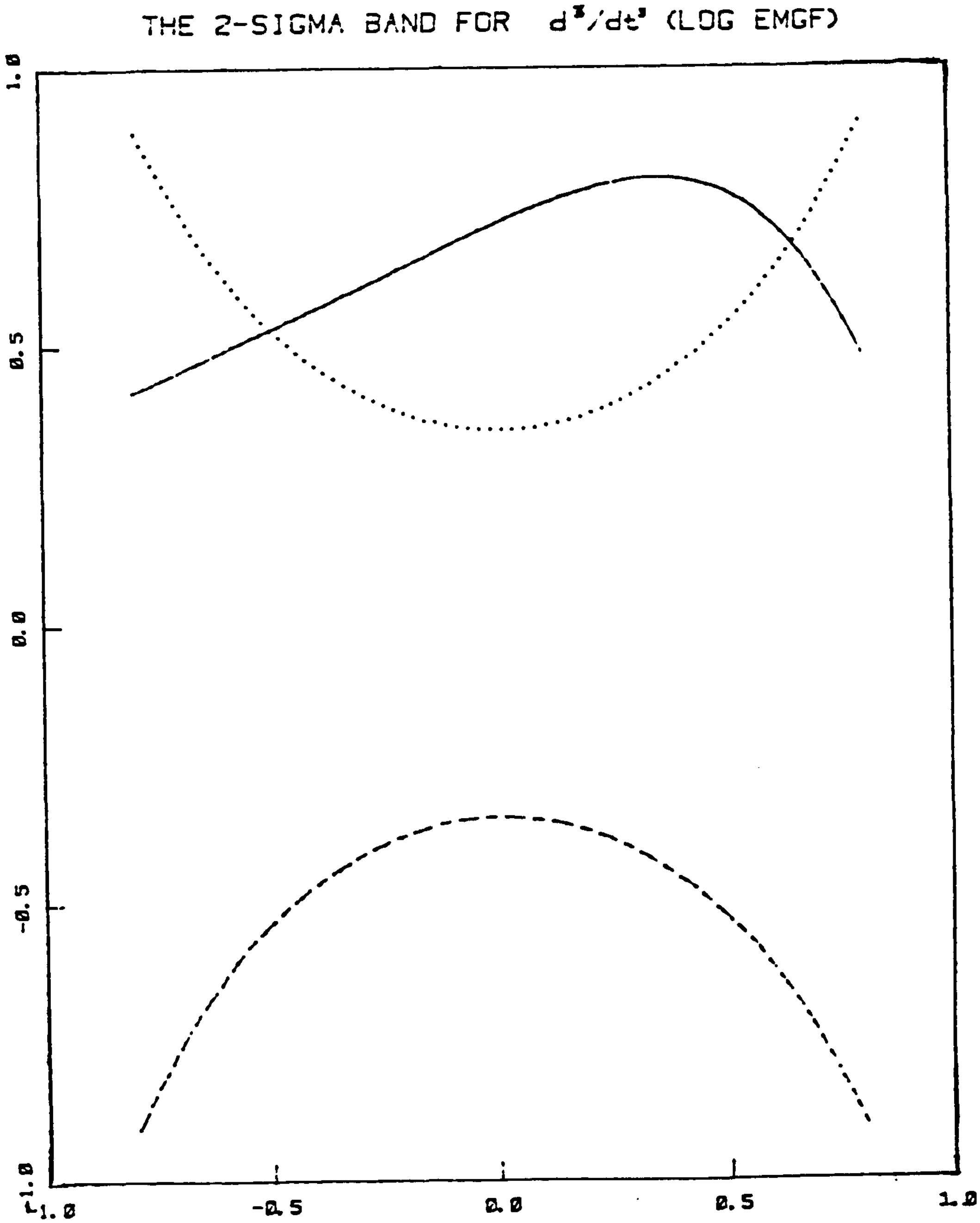
Sample # 2: Chi-square (10 d.f.)

Normal Probability Plot



Sample of size 200 from a chi-square population

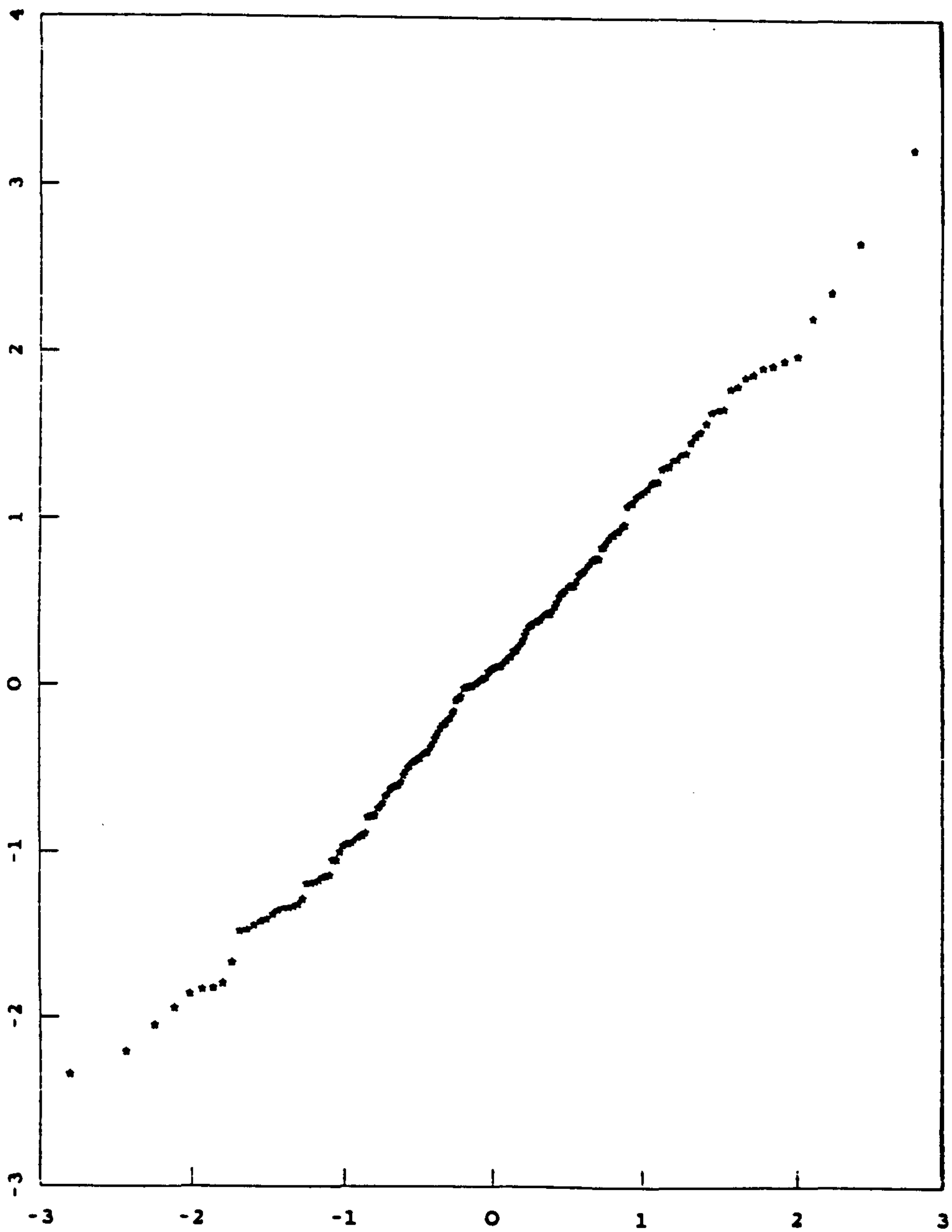
Sample # 3: Chi-square (50 d.f.)



Sample of size 200 from a chi-square population

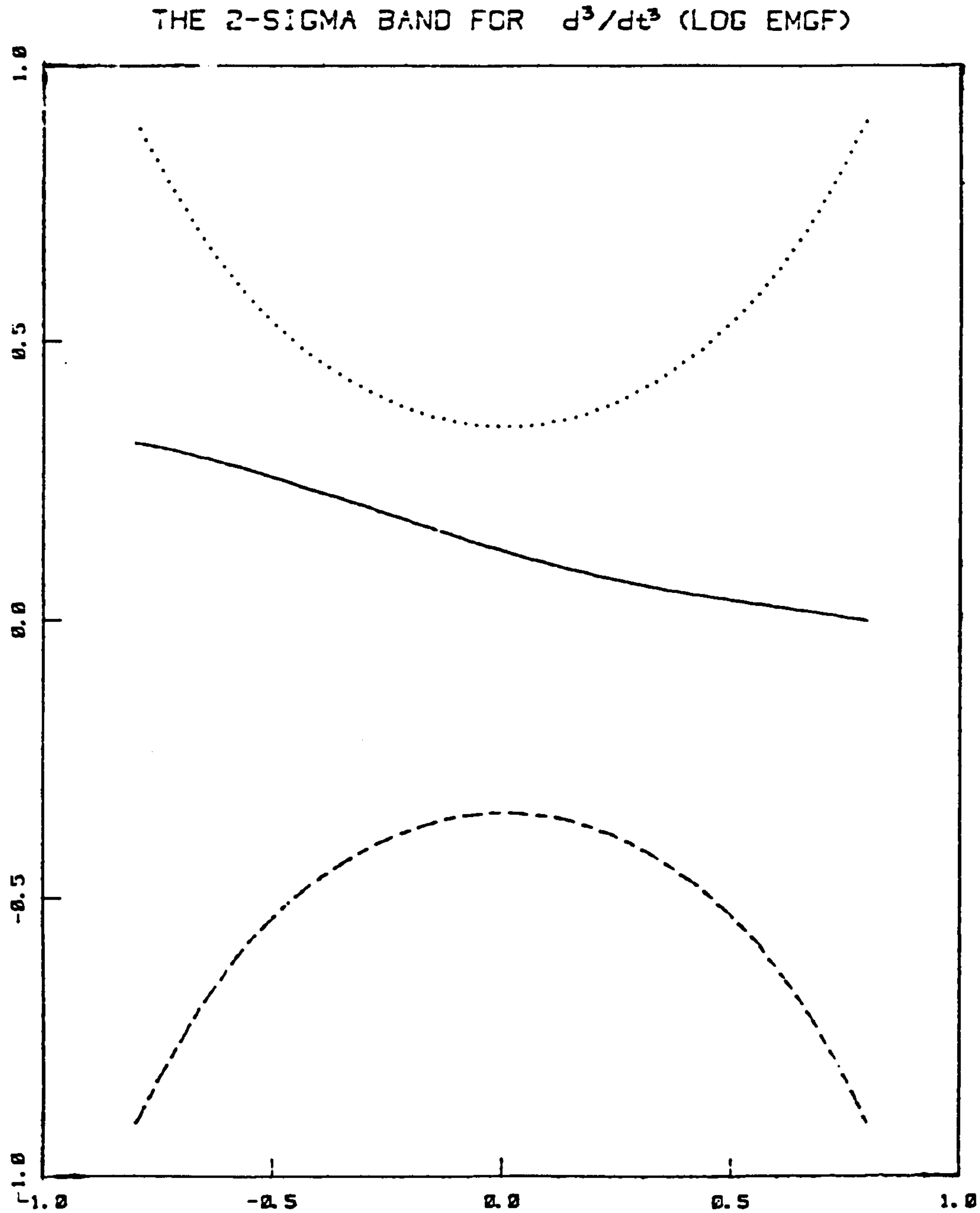
Sample # 3: Chi-square (50 d.f.)

Normal Probability Plot



Sample of size 200 from a chi-square population

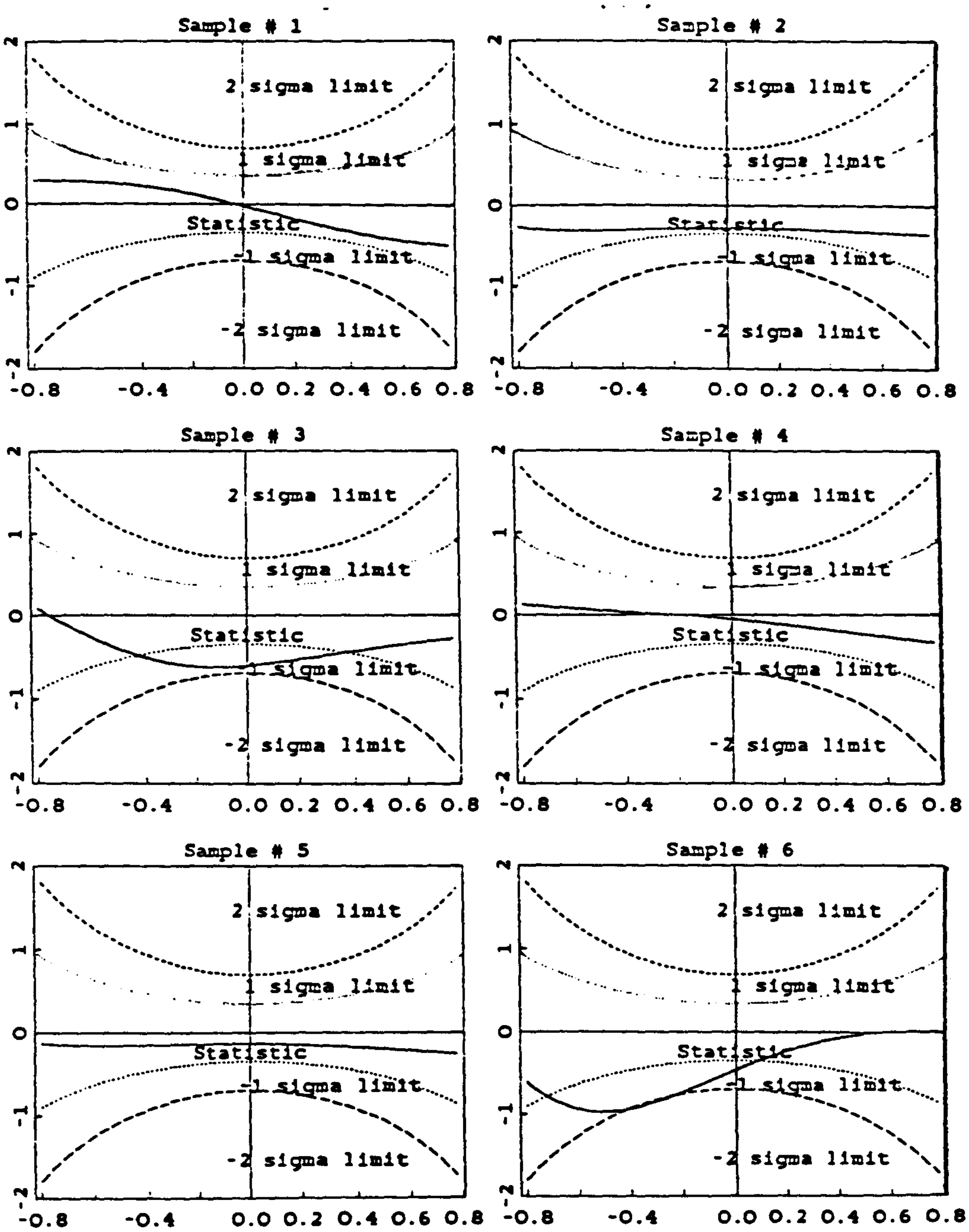
Sample # 4: Chi-square (150 d.f.)



Sample of size 200 from a chi-square population

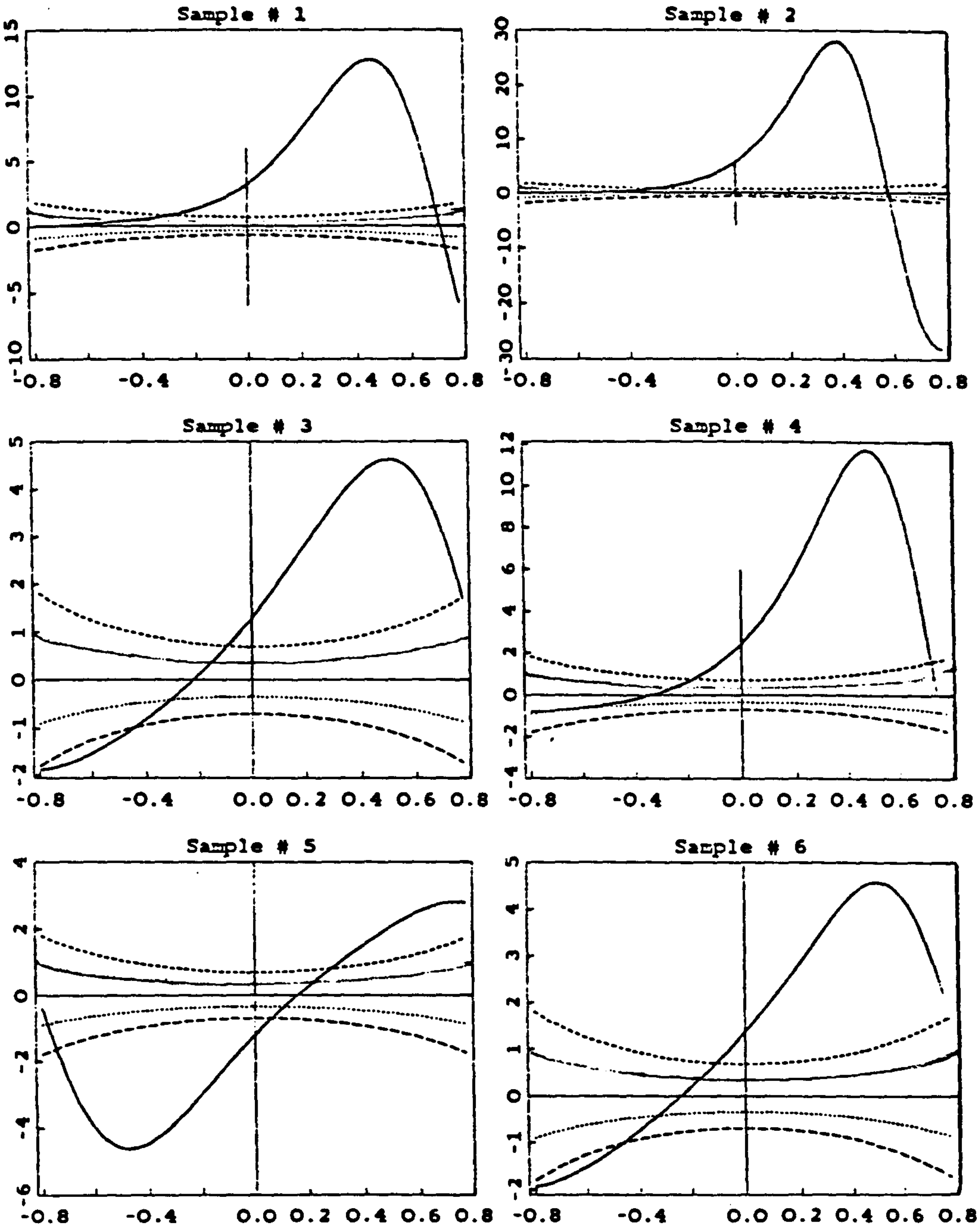
Sample # 4: Chi-square (150 d.f.)

Samples of size 50 from a N(0,1) population



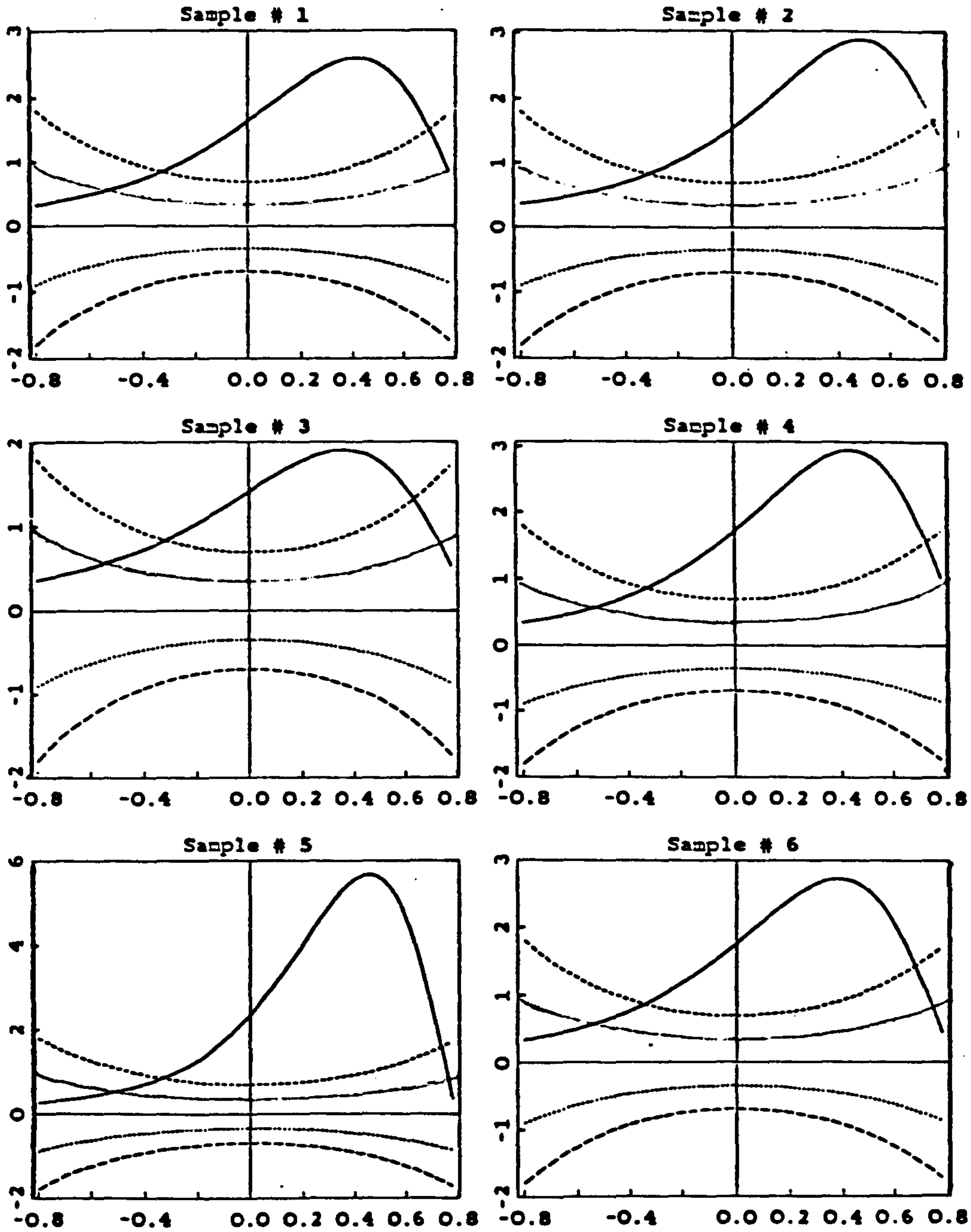
$$\frac{d^3}{dt^3} \log \tilde{m}_n(t) \quad vs. \quad t$$

Samples of size 50 from a Cauchy population



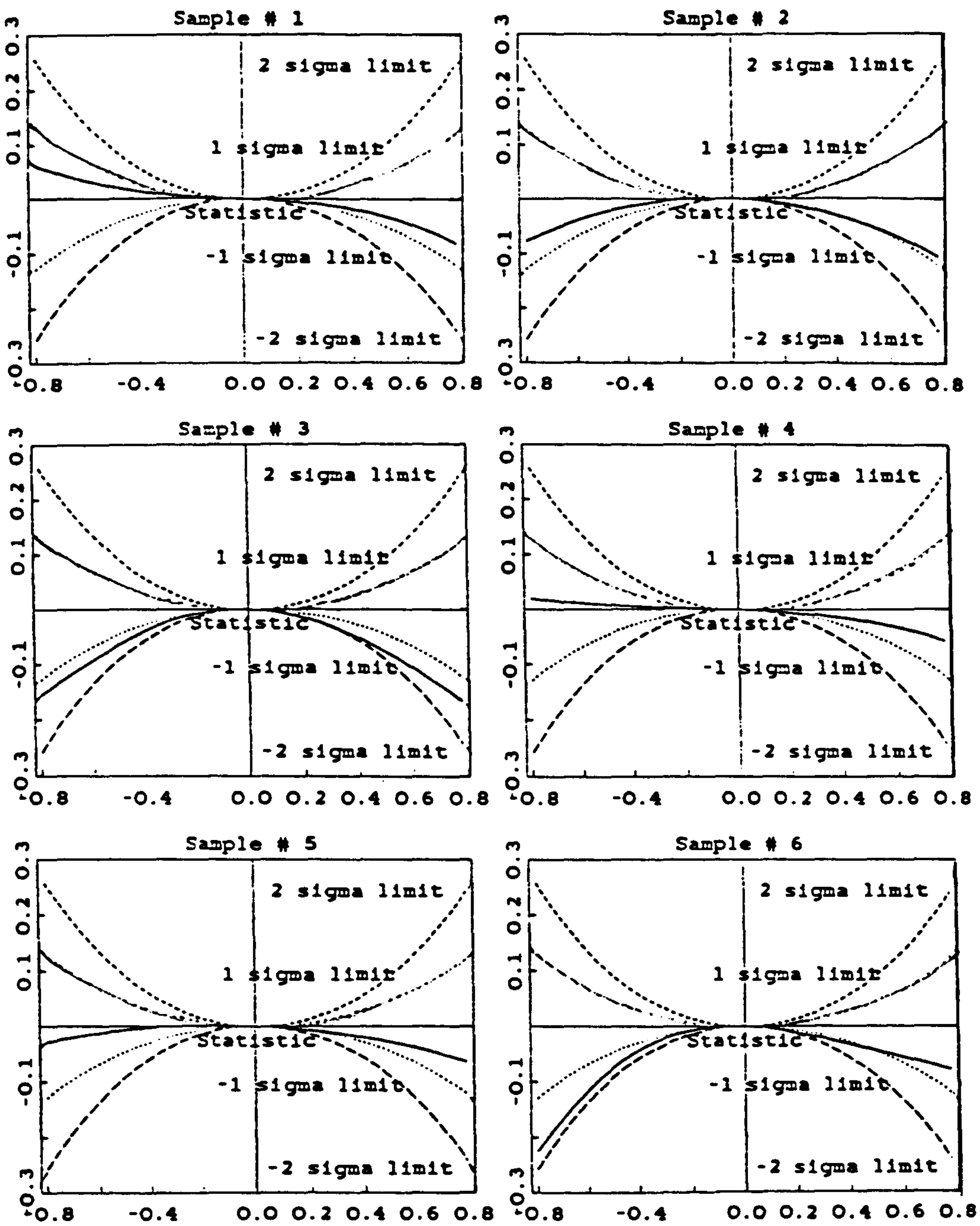
$$\frac{d^3}{dt^3} \log \tilde{m}_n(t) \quad vs. \quad t$$

Samples of size 50 from a Chi-square (2 d.f.) population



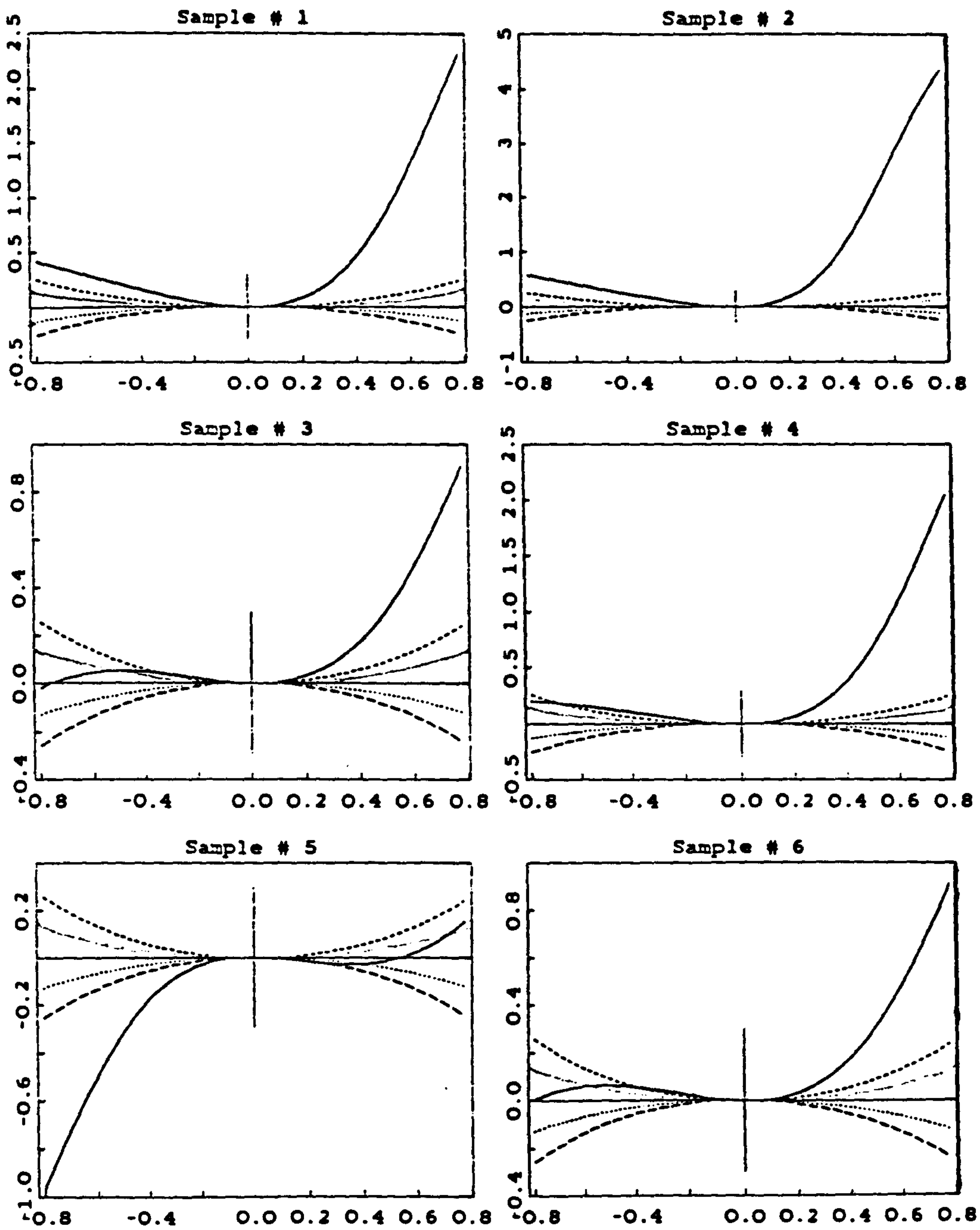
$$\frac{d^3}{dt^3} \log \tilde{m}_n(t) \quad vs. \quad t$$

Samples of size 50 from a N(0,1) population



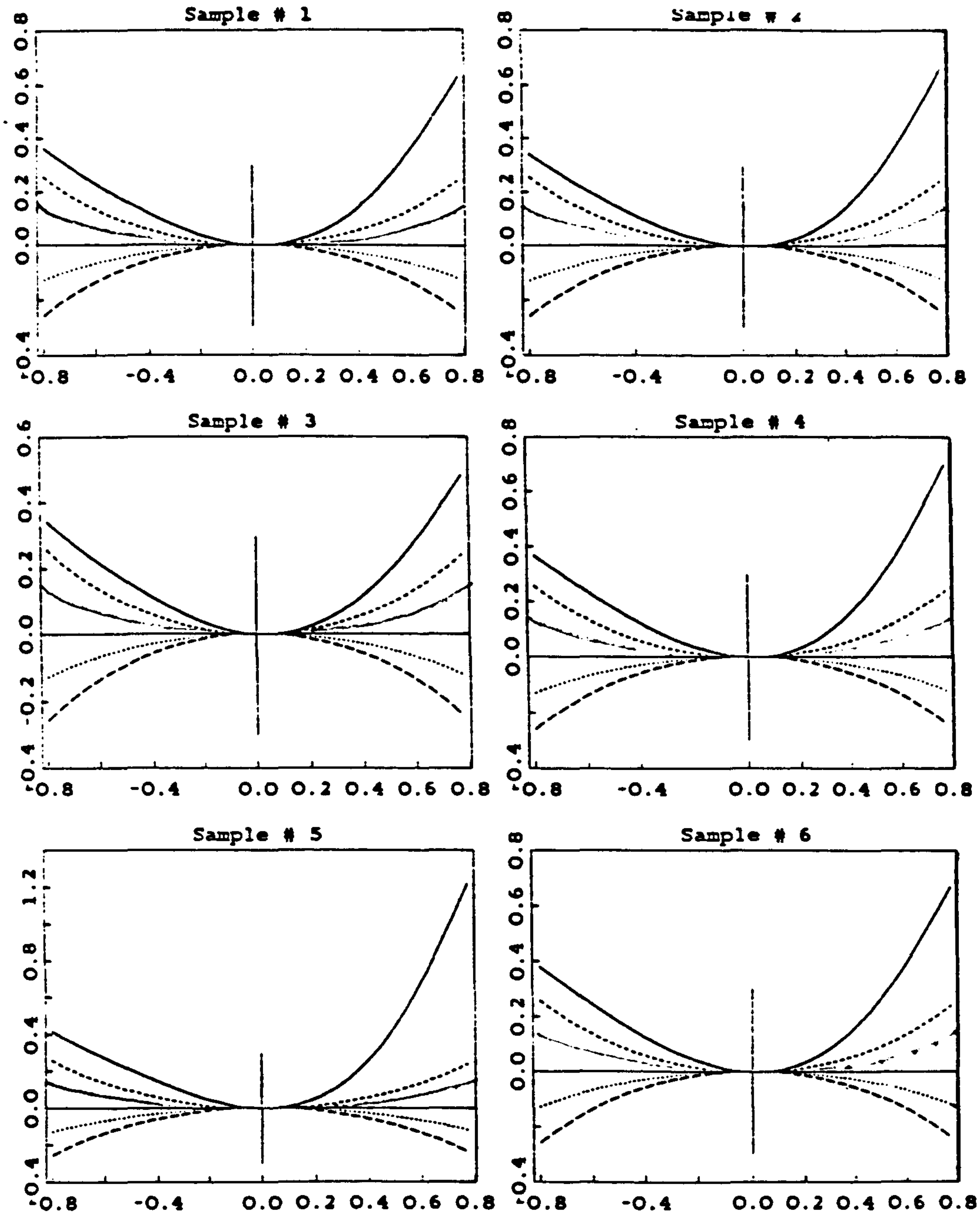
$$\frac{d}{dt} \log \tilde{m}_n(t) - t \quad vs. \quad t$$

Samples of size 50 from a Cauchy population



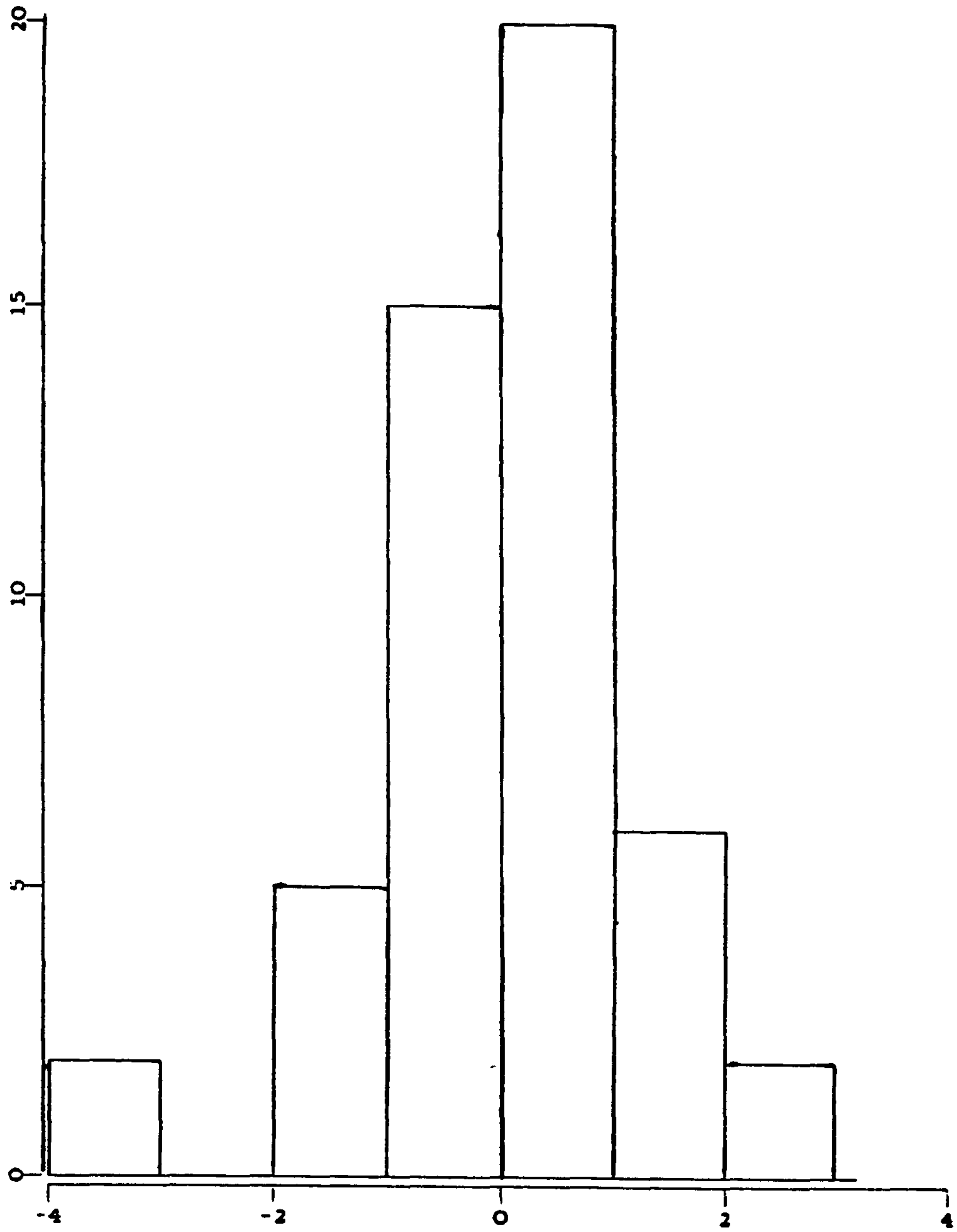
$$\frac{d}{dt} \log \tilde{m}_n(t) - t \quad vs. \quad t$$

Samples of size 50 from a Chi-square (2 d.f.) population

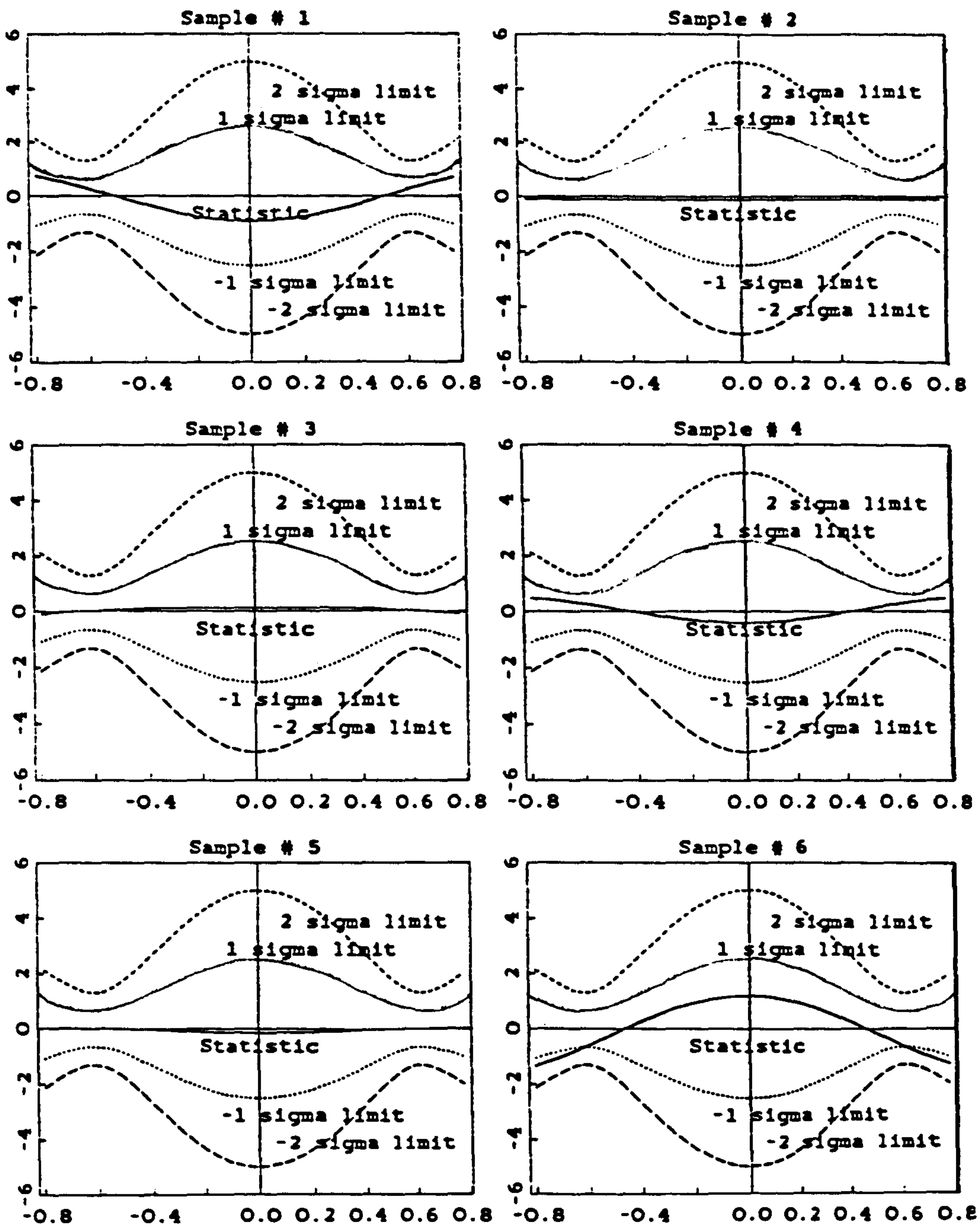


$$\frac{d}{dt} \log \tilde{m}_n(t) - t \quad vs. \quad t$$

Histogram for sample # 6 from N(0,1)

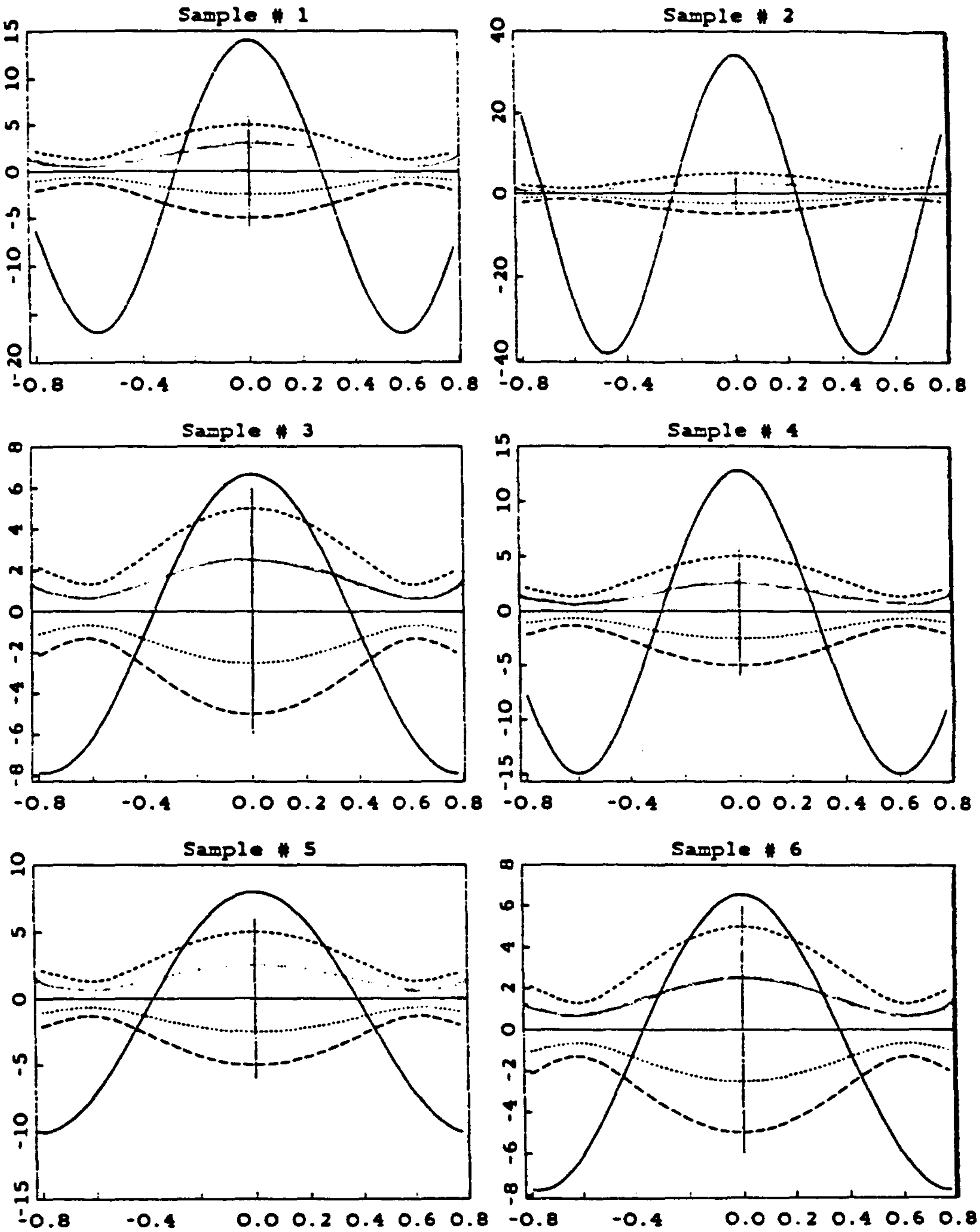


Samples of size 50 from a N(0,1) population



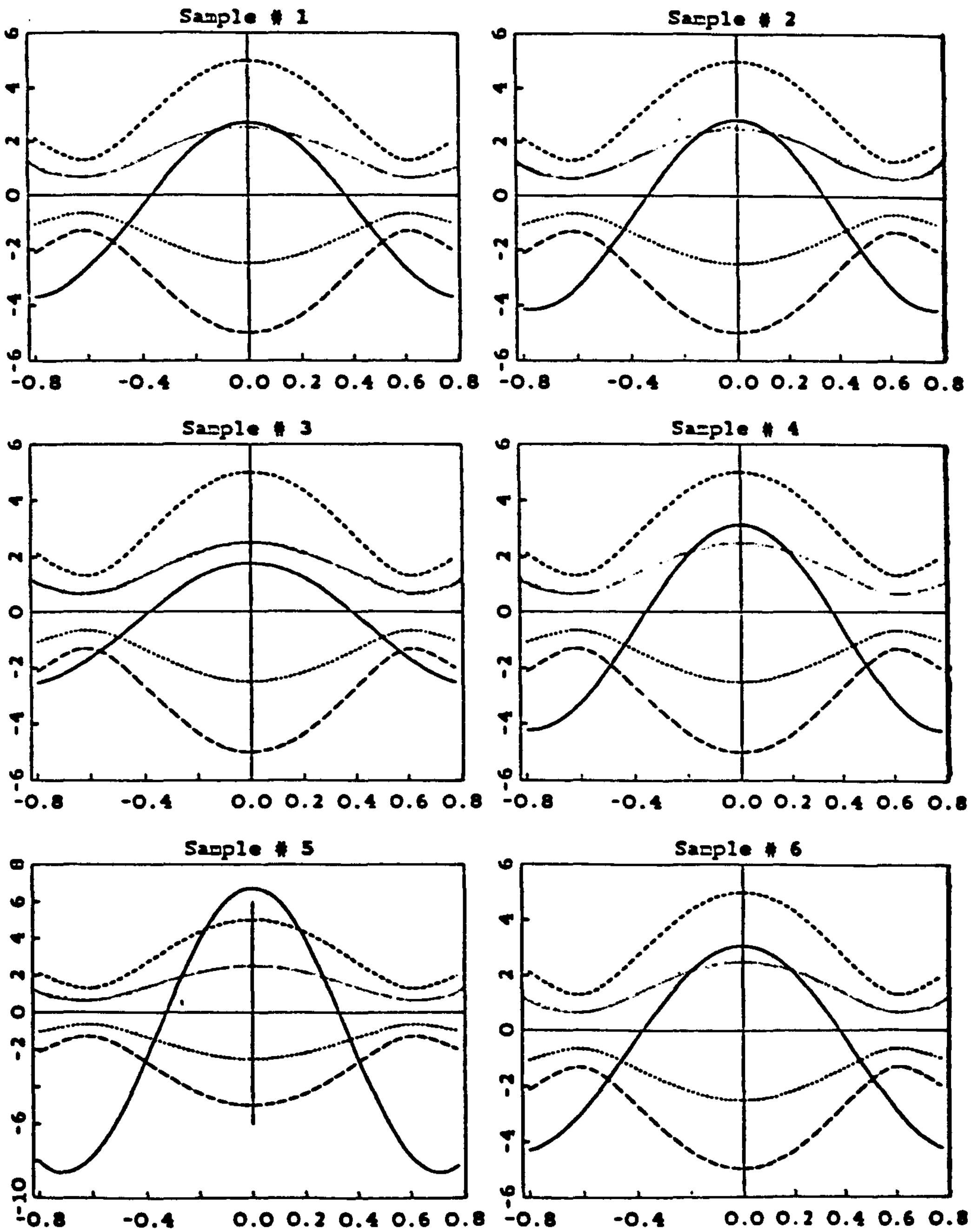
$$\frac{d^4}{dt^4}\, re\ \tilde{\phi}_n(t) - (t^4 - 6t^2 + 3)\, e^{-\frac{t^2}{2}} \quad vs. \quad t$$

Samples of size 50 from a Cauchy population



$$\frac{d^4}{dt^4}\, re\ \tilde{\phi}_n(t) - (t^4 - 6t^2 + 3)\, e^{-\frac{t^2}{2}} \quad vs. \quad t$$

Samples of size 50 from a Chi-square (2 d.f.) population



$$\frac{d^4}{dt^4}\, re\ \bar{\phi}_n(t) - (t^4 - 6t^2 + 3)\, e^{-\frac{t^2}{2}} \quad vs. \quad t$$

Power Calculations For $n \cdot \frac{[\frac{d^3}{dt^3} \log \tilde{m}_n(t)]^2}{K_0(t,t)}$.

Level Of Significance = 0.10.

Sample Size = 50.

# of Samples Used = 500.

Random Number Generators: G05DHF and G05DJF (from the NAG subroutine library.)

Distribution	t	Power
χ^2_1	0.5	0.972
χ^2_2	"	0.872
χ^2_4	"	0.714
χ^2_{10}	"	0.510
χ^2_1	0.2	1.000
χ^2_2	"	0.984
χ^2_4	"	0.894
χ^2_{10}	"	0.712
χ^2_1	0.0	1.000
χ^2_2	"	1.000
χ^2_4	"	0.970
χ^2_{10}	"	0.800
t_1	0.0	0.916
t_4	"	0.460
t_{10}	"	0.290
t_2	0.4	0.584
t_4	"	0.420
t_{10}	"	0.224