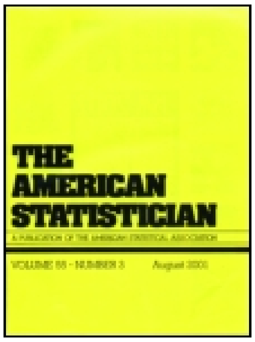

The American Statistician

ISSN: 0003-1305 (Print) 1537-2731 (Online) Journal homepage: https://www.tandfonline.com/loi/utas20

Confidence Bands for Quantile Functions: A Parametric and Graphic Alternative for Testing Goodness of Fit

Walter A. Rosenkrantz

To link to this article: https://doi.org/10.1080/00031305.2000.10474543

Confidence Bands for Quantile Functions: A Parametric and Graphic Alternative for Testing Goodness of Fit

Walter A. ROSENKRANTZ

This article derives simultaneous $100(1-\alpha)\%$ confidence intervals for the quantiles of a normal distribution using a method first proposed by Cheng and Iles and independently rediscovered by Satten. These methods yield a novel parametric/graphic alternative to the usual goodness-of-fit tests for normality based on normal Q-Q plots, or extensions of Kolmogorov's test based on the empirical distribution function: Draw the graph of the empirical quantile plot, then plot the corresponding upper and lower bounds for these quantiles. We thus obtain a $100(1-\alpha)\%$ confidence band for the empirical quantile plot. The more the parent distribution of the data departs from the normal, the more likely the empirical quantile plot will have points that lie outside these bounds. The empirical quantile plot and its confidence bands are pleasing to the eye and easy to interpret, even for a user unfamiliar with the underlying theory. Our bounds, although not optimal, are easily computable in terms of the normal, t, and χ^2 distributions. The test is applicable for all sample sizes n (no delicate asymptotic limit theorems are required) and it works when the mean and variance are estimated from the data. These methods are illustrated by applying them to two datasets analyzed by different methods.

KEY WORDS: Convex confidence sets; Normal probability plots; Simultaneous confidence intervals.

1. INTRODUCTION

Determining whether a random sample $\boldsymbol{X} = (X_1, \ldots, X_n)$ comes from a given distribution $F(x)$ or, more generally, from a family of distributions $F(x; \theta_1, \ldots, \theta_n)$ depending on one or more unknown parameters, is a fundamental problem in statistics, and a variety of graphical and formal methods have been proposed to study it (Bickel and Doksum 1977; Lehmann 1999). Among the more widely used methods are the venerable normal Q-Q plot, tests based on the empirical distribution function, such as the Kolmogorov test, and variations of it due to Lilliefors (1967), Durbin (1973), and Stephens (1974). We refer the reader to Lehmann (1999) for a more complete discussion as well as useful references to the current literature. More recently, Davison and Hinkley (1997, p. 148) obtained confidence bands for the quantiles using the parametric bootstrap. When the data come from a normal population, then the points of the Q-Q plot appear to lie along a straight line. But judging the straightness of the plot is quite subjective (see Figures 2 and 5) and, as is well known, determining the asymptotic behavior of the Kolmogorov test when the mean and variance must be estimated from the data is not an easy matter (Durbin 1973; Khmaladze 1981); which is why the Lilliefors confidence bands are obtained via Monte Carlo simulations. The confidence bands for the empirical quantile plots proposed here provide a simple, but useful, theoretical and graphical complement to these subjective judgments, and they are easily computed using only the standard statistical functions available in all the standard statistical software packages (we used SAS). To illustrate the method, it is applied to two datasets where the traditional methods mentioned above are unable to determine whether the data come from a normal population.

Walter A. Rosenkrantz is Professor, Department of Mathematics and Statistics, University of Massachusetts, Amherst, MA 01003-4515 (E-mail: rkrantz@math.umass.edu). In the course of obtaining the results presented here the author benefited from useful conversations with J. Buonacorsi, J. Horowitz, M. Kahn, and Marco Paolella. The impetus for studying this problem arose while supervising Todd Blessinger's master's project in statistics (Blessinger 1997).

2. EMPIRICAL QUANTILE PLOTS

Let $\hat{F}_n(x)$, $\hat{Q}_n(p)$, $(X_{(1)} \leq \cdots \leq X_{(k)} \leq \cdots \leq X_{(n)})$ denote the empirical distribution function, the p quantile, and the order statistics of the random sample $\boldsymbol{X} = (X_1, \ldots, X_n)$ taken from a continuous distribution. Then, as is well known, the kth order statistic is the $(k-.5)/n$ quantile of the empirical distribution function; that is,

$$\hat{Q}_n\left(\frac{k-.5}{n}\right) = X_{(k)}.$$

The empirical quantile plot is obtained by plotting the points $((k-.5)/n, X_{(k)})$, $(k = 1, \ldots, n)$. This is the discrete analog of the inverse of the empirical distribution function which is obtained by plotting the points $(X_{(k)}, k/n)$. The idea of replacing the empirical distribution function with the empirical quantile plot appears for the first time in Rosenkrantz (1969) and was subsequently used again to give a new approach to limit theorems for linear combinations of order statistics (Rosenkrantz and O'Reilly 1972). Let $Q(p)$ denote the p quantile of the normal distribution with unknown mean μ and variance σ^2. In Proposition 1 we present an explicit, easily computable formula for simultaneous $100(1-\alpha)\%$ confidence intervals $[Q_1(p), Q_2(p)]$ for $Q(p)$; that is,

$$P(\cap_{0<p<1}(Q_1(p) \leq Q(p) \leq Q_2(p)) \geq 1-\alpha.$$

The confidence bands are obtained by plotting the points $((k-.5)/n, Q_i(k-.5)/n))$, $(i = 1, 2)$, $(k = 1, \ldots, n)$.

Table 1. Dataset from Bickel and Doksum (1977, p. 384)

2.968	2.097	1.611	3.038	7.921	5.476	9.858	1.397	0.155	1.301
9.054	1.958	4.058	3.918	2.019	3.689	3.081	4.229	4.669	2.274
1.971	10.379	3.391	2.093	6.053	4.196	2.788	4.511	7.300	5.856
0.860	2.093	0.703	1.182	4.114	2.075	2.834	3.968	6.480	2.360
5.249	5.100	4.131	0.020	1.071	4.455	3.676	2.666	5.457	1.046
1.908	3.064	5.392	8.393	0.916	9.665	5.564	3.599	2.723	2.870
1.582	5.453	4.091	3.716	6.156	2.039				

A method for obtaining the simultaneous $100(1-\alpha)\%$ confidence intervals for the quantiles was discussed in great generality by Cheng and Iles (1983) and Satten (1995), so we shall content ourselves with a sketch, referring the reader to Section 3 of this article for the fine details. Let $\mathcal{R}(\boldsymbol{X})$ denote a $100(1-\alpha)\%$ compact, convex, joint confidence region for (μ, σ) of the sort discussed by Arnold and Shavelle (1998), Bickel and Doksum (1977), Cheng and Iles (1983), Rohatgi (1976), and Satten (1995). Let $F(x; \mu, \sigma)$ denote the cumulative distribution function of the normal distribution with mean μ and variance σ^2. The p quantile of F is $Q(p) = \mu + x(p)\sigma$, where $x(p) = \Phi^{-1}(p)$ is the p quantile of the standard normal distribution; that is $Q(p)$ satisfies the equation $F(Q(p); \mu, \sigma) = p$. Observe that the p quantile $Q(p)$ is a linear function of the parameters (μ, σ), so the line in the plane defined by $\mu + x(p)\sigma = Q(p)$ is called an *isoquantile* line. For p fixed, $Q_1(p)$ is defined to be the minimum and $Q_2(p)$ is the maximum value of $Q(p)$ on the confidence set $\mathcal{R}(\boldsymbol{X})$; that is,

$$\begin{aligned} Q_1(p) &= \min(\mu + x(p)\sigma : (\mu, \sigma) \in \mathcal{R}(\boldsymbol{X})) \\ Q_2(p) &= \max(\mu + x(p)\sigma : (\mu, \sigma) \in \mathcal{R}(\boldsymbol{X})). \end{aligned}$$

Because $Q(p)$ is linear and $\mathcal{R}(\boldsymbol{X})$is compact and convex, it is clear that its minimum and maximum values are attained on the boundary of the confidence region (see Figure 7). In particular, these equations are easily solved if one uses the trapezoidal confidence shaped confidence region used here or the rectangular shaped confidence region described in Bickel and Doksum (1977) and Rohatgi (1976).

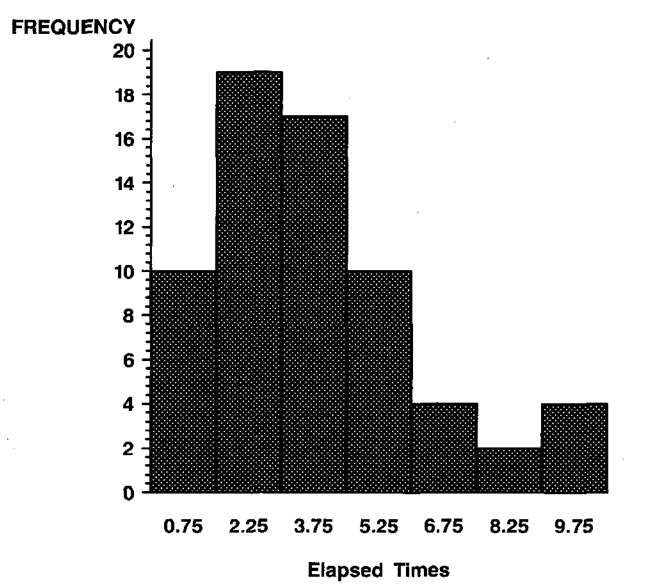

Figure 1. Histogram of Bickel and Doksum Data (Bickel and Doksum 1977, p. 384).

Postponing the formal derivation of the confidence intervals to Section 3, we begin with two examples of empirical quantile plots and their confidence bands.

Example 1. The dataset shown in Table 1 comes from Bickel and Doksum (1977, p. 384, tab. 9.6.3), and lists the elapsed times spent above a certain high level for a series of 66 wave records taken at San Francisco Bay.

After standardizing the variables we plotted the histogram (Figure 1) and the normal Q-Q plot (Figure 2). The histogram indicates that the data are skewed to the right, and the Q-Q plot suggests that the data do not come from a normal distribution; but the evidence against normality is inconclusive.

Using the Stephens confidence bands for the empirical distribution function, Bickel and Doksum tested the data for normality, at the level $\alpha = .10$, (Bickel and Doksum 1977, p. 384, Figure 9.6.3) and reproduced here as Figure 3. Looking at Figure 3 it is clear that this test fails to reject the hypothesis of normality. Furthermore, notice how wide the confidence bands become in the tails of the distribution. Now, consider the author's simultaneous 90% confidence band for the empirical quantile plot displayed in Figure 4. The hypothesis of normality is clearly rejected, and the confidence bands in the tails are much narrower than the Stephens bands. Finally, we note that these results are consistent with those obtained by the Shapiro–Wilk test (Bickel and Doksum 1977, p. 388).

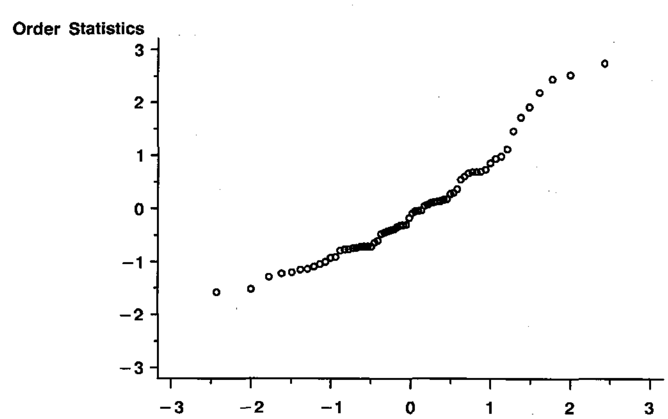

Figure 2. Normal Q-Q Plot. Bickel and Doksum (1977, tab. 9.6.3).

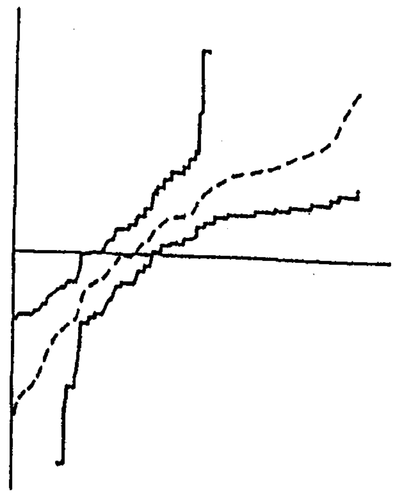

Figure 3. Kolmogorov Confidence Band for Data in Table 1.

Example 2. The dataset shown in Table 2 is the eighth in a series of measurements performed at the National Bureau of Standards to determine the acceleration of gravity g [the complete dataset is reproduced in Davison and Hinkley (1997), tab. 3.1, p. 72].

In their recent book, Davison and Hinkley applied parametric resampling methods to obtain a 90% confidence band for the Q-Q plot (Davison and Hinkley 1997, Figure 4.5, p. 154). Even though one of the sample quantiles falls outside the confidence band, the authors note that it is premature to conclude that the data come from a nonnormal population, The reader is referred to Davison and Hinkley (1997, Examples 3.2, p. 72; and Examples 4.6, 4.7, and 4.8, pp. 150–154) for a detailed discussion of their methods and their conclusions). In brief, their results are inconclusive because, as they themselves pointed out, their confidence bounds are pointwise and not simultaneous. In addition the normal probability plot itself, Figure 5, does not provide strong evidence for, or against, the null hypothesis. Although the data appear to lie on a straight line, there are departures from linearity that make one uneasy. On the other hand, the 90% confidence band for the empirical quantile plot, Figure 6, tells a different story—the data do not even come close to departing from normality.

2.1 A Numerical Example and Comparison With Other Methods

We now compare the bounds produced by our methods with nonparametric confidence intervals for p quantiles discussed by Wilks (1962, chap. 11), and with the bounds obtained by using the tables in Hahn and Meeker (1991).

To illustrate the basic ideas we compute the 90% confidence interval for the 90th percentile for the dataset in Table 2, using Equation (2) of Section 3. and the values $n = 13$, $\alpha = .10, \overline{X} = 80.3846$, $s = 3.3551$, obtaining the following results (rounded to four decimal places):

$$Q_1(.90) = 82.1707,\ Q_2(.90) = 90.4445.$$

We then computed the corresponding nonparametric confidence interval using the results given by Wilks (1962). It is well known that the interval $(X_{(j)}, X_{(k)})$ contains the quantile $Q(p)$ with confidence level $1-\gamma$ given by

$$1-\gamma = \sum_{j\le i\le k-1} \binom{n}{i} p^i(1-p)^{n-i}. \qquad (1)$$

Since the dataset in Table 2 contains $n = 13$ observations it follows that the interval $(X_{(1)}, X_{(13)})$ provides the maximum coverage probability for any quantile. In this case, as we now show, the maximum coverage probability for the 90th percentile is .75. To see this, insert the values $n = 13, p = .90, j = 1, k = 13$ into Equation (1) and use the fact that $X_{(1)} = 76$, $X_{(13)} = 86$. Therefore,

$$1-\gamma = \sum_{1\le i\le 12} \binom{n}{i} (.90)^i(.10)^{13-i} \approx 1-(.9)^{13} = .75.$$

Consequently, the nonparametric confidence interval [76, 86] has width 10 and coverage probability .75, while the 90% parametric confidence interval has width 8.2738, and coverage probability .90.

We obtain sharper bounds for a single quantile using the tables in Hahn and Meeker (1991; see sec. 4.4, p. 56, where the notation $g'(\alpha, p, n)$ is defined).

$$\begin{aligned} Q_1(.90) &= \overline{X} + g'(.05, .90, 13) \times s \\ &= 80.3846 + 0.772 \times 3.3551 = 82.9747, \end{aligned}$$

$$\begin{aligned} Q_2(.90) &= \overline{X} + g'(.95, .90, 13) \times s \\ &= 80.3846 + 2.155 \times 3.3551 = 87.6148. \end{aligned}$$

This calculation shows that, if only one or two confidence intervals are wanted, then the tables in Hahn and Meeker

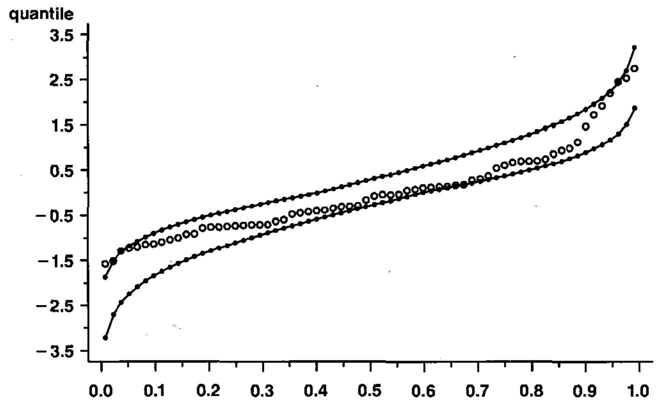

Figure 4. 90% Confidence Band for Quantile Plot. Bickel and Doksum (1977, tab. 9.6.3).

Table 2. Gravity Data

84	86	85	82	77	76	77	80	83	81	78	78	78

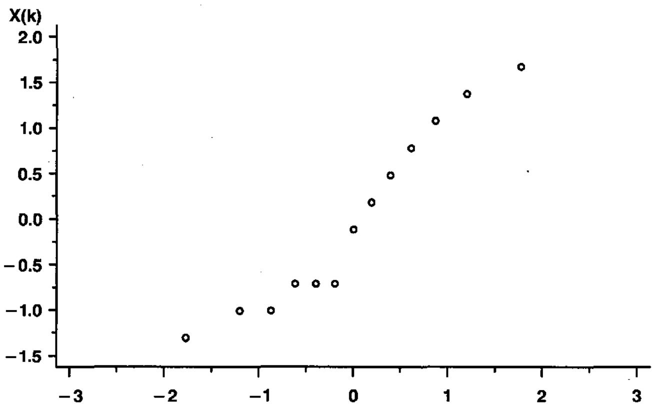

Figure 5. Normal Probability Plot. Davison and Hinckley (1997, tab. 3.1).

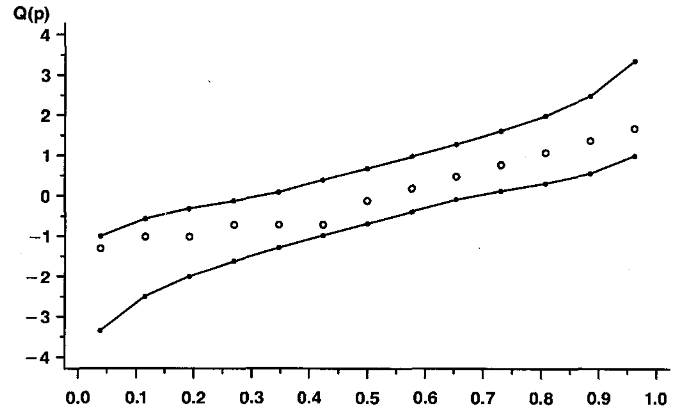

Figure 6. 90% Confidence Band for Quantile Plot. Davison and Hinckley (1997, tab. 3.1).

(1991) yield sharper bounds. But, if one wants simultaneous confidence intervals of the sort required to produce the confidence bands in Figures 4 and 6, then the bounds given in Proposition 1 will be much narrower, even for moderate values of the sample size.

Another graphical test for normality was developed by Lilliefors (1967). The idea is to compare the empirical distribution function of the sample to that of the standard normal distribution. In appearance it is similar to Figure 3. The advantages of our approach are clear: The empirical distribution function has jumps of fixed size at random times which give it a ragged, unattractive appearance. The empirical quantile plot is somewhat easier to study because it has random jumps at fixed times, making comparisons with the normal quantile plot easier to understand.

3. DERIVATION OF THE SIMULTANEOUS CONFIDENCE INTERVALS

Before stating our main result we recall the following notation: $x(p) = \Phi^{-1}(p)$ is the p quantile of the standard normal distribution, $z(\alpha) = x(1-\alpha)$, $t_\nu(\alpha)$, and $\chi^2_\nu(\alpha)$ are the upper tailed $100(1-\alpha)$ percentiles of the normal, t and chi-square distribution (with ν degrees of freedom), respectively. That is, $P(t_\nu > t_\nu(\alpha)) = \alpha$, $P(\chi^2_\nu > \chi^2_\nu(\alpha)) = \alpha$. The basic idea leading to Proposition 1 is a geometrical construction due to Cheng and Iles (1983); however, it undoubtedly goes back to Scheffé's derivation of simultaneous confidence intervals for all contrasts in the analysis of variance.

Proposition 1. Let $Q(p)$ denote the p quantile of the normal distribution with unknown mean μ and variance σ^2. Let $X_1, \ldots, X_n$ be a random sample taken from this distribution with sample mean $\overline{X}$ and sample variance s^2. The intervals $(Q_1(p), Q_2(p))$ defined by

$$Q_1(p) = \overline{X} - k_1(p)s,\ Q_2(p) = \overline{X} + k_2(p)s, \tag{2}$$

where the quantities $k_i(p)$, $(i = 1, 2)$ are defined in Equations (5)–(8), are simultaneous $100(1-\alpha)\%$ confidence intervals for $Q(p)$; that is

$$P(\cap_{0<p<1}(Q_1(p) \le Q(p) \le Q_2(p)) = (1-\alpha'/2)^2 = 1-\alpha. \tag{3}$$

We note for future reference that $(1-\alpha'/2)^2 = 1-\alpha$ implies that

$$\alpha' = 2(1-\sqrt{1-\alpha}) > \alpha. \tag{4}$$

1. If $0 < p \le \Phi\left(z(\alpha'/4)/\sqrt{n}\right)$, then

$$k_1(p) = z(\alpha'/4)\sqrt{\frac{(n-1)}{n\chi^2_{n-1}(1-\alpha'/4)}} - x(p)\sqrt{\frac{(n-1)}{\chi^2_{n-1}(1-\alpha'/4)}}. \tag{5}$$

If $\Phi\left(z(\alpha'/4)/\sqrt{n}\right) < p < 1$, then

$$k_1(p) = z(\alpha'/4)\sqrt{\frac{(n-1)}{n\chi^2_{n-1}(\alpha'/4)}} - x(p)\sqrt{\frac{(n-1)}{\chi^2_{n-1}(\alpha'/4)}}. \tag{6}$$

2. If $0 < p \le \Phi\left(-z(\alpha'/4)/\sqrt{n}\right)$, then

$$k_2(p) = z(\alpha'/4)\sqrt{\frac{(n-1)}{n\chi^2_{n-1}(\alpha'/4)}} + x(p)\sqrt{\frac{(n-1)}{\chi^2_{n-1}(\alpha'/4)}}. \tag{7}$$

If $\Phi\left(-z(\alpha'/4)/\sqrt{n}\right) < p < 1$, then

$$k_2(p) = z(\alpha'/4)\sqrt{\frac{(n-1)}{n\chi^2_{n-1}(1-\alpha'/4)}} + x(p)\sqrt{\frac{(n-1)}{\chi^2_{n-1}(1-\alpha'/4)}}. \tag{8}$$

Proof:

1. The first step is to construct the Mood exact joint confidence set $\mathcal{R}(\boldsymbol{X})$ for (μ, σ) defined as:

$$\mathcal{R}(\boldsymbol{X}) = (\mu, \sigma) : \overline{X} - z(\alpha'/4)\frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X} + z(\alpha'/4)\frac{\sigma}{\sqrt{n}},\quad s\sqrt{\frac{(n-1)}{\chi^2_{n-1}(\alpha'/4)}} \le \sigma \le s\sqrt{\frac{(n-1)}{\chi^2_{n-1}(1-\alpha'/4)}}. \tag{9}$$

It is easy to see that $\mathcal{R}(\boldsymbol{X})$ is a trapezoid with the top and bottom sides parallel to the μ axis (Figure 7). The parallel lines passing through the opposite corner points of

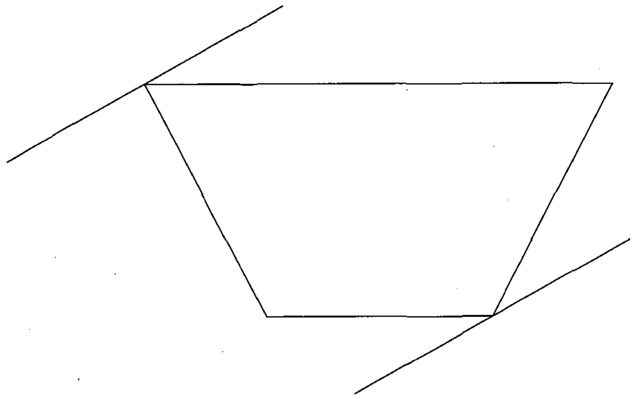

Figure 7. Mood Confidence Region and Isoquantile Lines.

the trapezoid are the isoquantile lines defined by Equations (13)–(14).

Our construction differs only slightly from the one given by Arnold and Shavelle; their joint confidence set is for (μ, σ^2) (see Arnold and Shavelle 1998, fig. 1, p. 134). Because $\overline{X}$ and s are independent it follows that

$$P((\mu, \sigma) \in \mathcal{R}(\boldsymbol{X})) = (1 - \alpha'/2)^2 = 1 - \alpha,$$

so $\mathcal{R}(\boldsymbol{X})$ is a $100(1-\alpha)\%$ joint confidence set for μ, σ.

We label the four corner points of the trapezoid as NE(northeast), SE(southeast), SW(southwest), and NW(northwest). Their coordinates are:

$$\mathrm{NW} = (\overline{X} - \delta_2,\, \sigma_2),\ \mathrm{NE} = (\overline{X} + \delta_2,\, \sigma_2),$$
$$\mathrm{SW} = (\overline{X} - \delta_1,\, \sigma_1),\ \mathrm{SE} = (\overline{X} + \delta_1,\, \sigma_1),$$

where

$$\delta_1 = \frac{sz(\alpha'/4)}{\sqrt{n}}\sqrt{\frac{n-1}{\chi^2_{n-1}(\alpha'/4)}},$$
$$\delta_2 = \frac{sz(\alpha'/4)}{\sqrt{n}}\sqrt{\frac{n-1}{\chi^2_{n-1}(1-\alpha'/4)}} \tag{10}$$
$$\sigma_1 = s\sqrt{\frac{n-1}{\chi^2_{n-1}(\alpha'/4)}},$$
$$\sigma_2 = s\sqrt{\frac{n-1}{\chi^2_{n-1}(1-\alpha'/4)}}. \tag{11}$$

We note for future reference that the slopes of the left- and right-hand sides of the trapezoid depend only on α' and n. More precisely, the slopes of the sides of the trapezoid are

$$\text{slope of trapezoid} = \pm(\sigma_2 - \sigma_1)/(\delta_2 - \delta_1) = \pm\frac{\sqrt{n}}{z(\alpha'/4)}. \tag{12}$$

We recall that the p quantile $Q(p)$ with p fixed defines the isoquantile line $Q(p) = \mu + x(p)\sigma$ in the (μ, σ) plane. For p fixed, $Q_1(p)$ and $Q_2(p)$ are, respectively, the minimum and maximum values attained by $Q(p)$ on the confidence set $\mathcal{R}(\boldsymbol{X})$; that is,

$$Q_1(p) = \min(\mu + x(p)\sigma : (\mu, \sigma) \in \mathcal{R}(\boldsymbol{X})), \tag{13}$$
$$Q_2(p) = \max(\mu + x(p)\sigma : (\mu, \sigma) \in \mathcal{R}(\boldsymbol{X})). \tag{14}$$

The strip between the two isoquantile lines $Q_1(p)$, $Q_2(p)$ determines a convex set $\mathcal{C}_p(\boldsymbol{X})$ in the (μ, σ) plane that contains the confidence set $\mathcal{R}(\boldsymbol{X})$) for every p, $0 < p < 1$. Therefore,

$$\cap_{0<p<1}\mathcal{C}_p(\boldsymbol{X}) \supseteq \mathcal{R}(\boldsymbol{X})$$

implies

$$P(\cap_{0<p<1}\mathcal{C}_p(\boldsymbol{X})) \geq P(\mathcal{R}(\boldsymbol{X})) = (1 - \alpha'/2)^2 = 1 - \alpha.$$

In particular, for every $(\mu, \sigma) \in \mathcal{R}(\boldsymbol{X})$ we must have $Q_1(p) \leq Q(p) \leq Q_2(p)$. Consequently,

$$P(\cap_{0<p<1}(Q_1(p) \leq Q(p) \leq Q_2(p)) = 1 - \alpha.$$

2. The final step is to give an explicit formula for $Q_i(p)$, $(i = 1, 2)$. Because the reasoning used to derive the lower and upper bounds are similar we shall give the proof for $Q_1(p)$ only.

We begin the proof for the case $0 < p \leq .5$, so the isoquantile line has positive slope $-1/x(p) > 0$. (When $p = .5$ the isoquantile line is a vertical line). Looking at Figure 7 we see that in this case the minimum is always attained when the isoquantile line passes through the NW corner. The equation of the isoquantile line passing through the NW corner is

$$\mu + x(p)\sigma = \overline{X} - \delta_2 + x(p)\sigma_2 = \overline{X} - k_1(p)s = Q_1(p).$$

Now suppose that $.5 < p < 1$, so $x(p) > 0$. In this case the isoquantile line has a negative slope. Consequently, the minimum value of $\mu + x(p)\sigma$ occurs either at the NW corner or the SW corner, depending on the magnitude of the slope. The slope of the left side of the trapezoid, as noted previously, is $-\sqrt{n}/z(\alpha'/4)$; therefore, when $-1/x(p) \leq -\sqrt{n}/z(\alpha'/4)$ the minimum value is attained at the NW corner, otherwise it is attained at the SW corner. The inequality $-1/x(p) \leq -\sqrt{n}/z(\alpha'/4)$ is equivalent to $x(p) \leq z(\alpha'/4)/\sqrt{n}$, and this implies that

$$p = \Phi(x(p)) \leq \Phi\left(z(\alpha'/4)/\sqrt{n}\right).$$

The equation of the isoquantile line with slope $-1/x(p)$ that passes through the southwest corner is

$$\mu + x(p)\sigma = \overline{X} - \delta_1 + x(p)\sigma_1 = Q_1(p).$$

Summing up, we have shown that

$$\min(\mu + x(p)\sigma : (\mu, \sigma) \in \mathcal{R}(\boldsymbol{X})) = \overline{X} - \delta_2 + x(p)\sigma_2,\ 0 < p \leq \Phi\left(z(\alpha'/4)/\sqrt{n}\right),$$

and

$$\min(\mu + x(p)\sigma : (\mu, \sigma) \in \mathcal{R}(\boldsymbol{X})) = \overline{X} - \delta_1 + x(p)\sigma_1,\ \Phi\left(z(\alpha'/4)/\sqrt{n}\right) < p < 1.$$

Equations (5)–(6) are obtained by substituting the values for δ_i, σ_i, $(i = 1, 2)$ displayed in Equations (10)–(11).

More generally, it is easy to see that for each $p \neq \Phi(z(\alpha'/4)/\sqrt{n})$, $\Phi(-z(\alpha'/4)/\sqrt{n})$ there is a unique isoquantile line that touches the trapezoid at exactly one of its corner points. For example, when $0 < p \leq \Phi(-z(\alpha'/4)/\sqrt{n})$ the maximum of $Q(p)$ is attained at the SE corner. The equation of the isoquantile line passing through the southeast corner with slope $-1/x(p)$ is

$$\mu + x(p)\sigma = \overline{X} + \delta_1 + x(p)\sigma_1 = Q_2(p);$$

its graph is plotted in Figure 7.

The other cases are treated in a similar fashion.

4. IMPROVEMENTS AND GENERALIZATIONS

1. The upper and lower confidence bounds obtained here used only the simplest possible joint confidence set. In a recent article that appeared in this journal, Arnold and Shavelle (1998) discussed a wide variety of joint confidence sets for the mean and variance of a normal distribution. Consequently, the isoquantile lines that are tangent to these confidence regions also determine the lower and upper confidence bounds. These confidence intervals are simultaneous $100(1-\alpha)\%$ confidence intervals for the quantiles because these sets are all convex, so they are the intersection of all the half spaces that contain them.

2. Cheng and Iles (1983) and Satten (1995) noted that this method can be applied in principle to any continuous distribution depending on a number of unknown parameters. However, with the exception of the lognormal and exponential distributions, constructing the $100(1-\alpha)\%$ confidence region leads to expressions that are difficult to evaluate numerically. And it is not easy to determine when the confidence set, obtained by using the likelihood ratio statistic, is convex. This is a question of independent interest. The advantages (and disadvantages) of our method are clear: By using the confidence trapezoid, or the confidence rectangle, we avoid all of these difficulties. In particular we derive simple closed-form expressions for the lower and upper bounds that are clearly superior to Q-Q plots and the confidence bands obtained using variations on the nonparametric Kolmogorov test—compare Figures 2, 3, and 4. The disadvantages are that the methods described here are asymptotically not optimal, since it is known that the maximum likelihood regions are asymptotically of smallest expected area (Arnold and Shavelle 1998).

3. The confidence bounds obtained here can also be used to benchmark the performance of the bootstrap and other bounds that have been proposed in the literature (see, e.g., Davison and Hinkley 1997; Efron and Tibshirani 1993; and Breiman, Stone, and Kooperberg 1990).

[Received July 1998. Revised May 1999.]

REFERENCES

Arnold, B.C., and Shavelle, R. M. (1998), "Joint Confidence Sets For the Mean and Variance of a Normal Distribution," *The American Statistician*, 52, 133–140.

Bickel, P. J., and Doksum, K. A. (1977), *Mathematical Statistics: Basic Ideas and Selected Topics*, San Francisco: Holden-Day.

Blessinger, T. (1997), "Confidence Intervals for Quantiles," unpublished master's project, Department of Statistics, University of Massachusetts.

Breiman, L., Stone, C. J., and Kooperberg, C. (1990), "Robust Confidence Bounds For Extreme Upper Quantiles," *Journal of Statistical Computing and Simulation*, 37, 127–149.

Cheng, R. C. H., and Iles, T. C. (1983), "Confidence Bands for Cumulative Distribution Functions of Continuous Random Variables," *Technometrics*, 25, 77–86.

Davison, A. C., and Hinkley, D. V. (1997), *Bootstrap Methods and Their Applications*, New York: Cambridge University Press.

Durbin, J. (1973), *Distribution Theory for Tests Based on the Sample Distribution Function, Conference Board of the Mathematical Sciences*, Philadelphia, PA: SIAM.

Efron, B., and Tibshirani, R. J. (1993), *An Introduction to the Bootstrap*, New York: Chapman and Hall.

Hahn, G. J., and Meeker, W. Q. (1991), *Statistical Intervals: A Guide for Practitioners*, New York: Wiley.

Khmaladze, E. V. (1981), "Martingale Approach in the Theory of Goodness-of-Fit Tests," *Theory of Probability and its Applications*, vol. xxvi, pp. 240–257.

Lehmann, E. H. (1999), *Elements of Large Sample Theory*, New York: Springer.

Lilliefors, H. W. (1967), "On the Kolmogorov–Smirnov Test For Normality with Mean and Variance Unknown," *Journal of the American Statistical Association*, 62, 399–402.

Rohatgi, V. K. (1976), *An Introduction to Probability theory and Mathematical Statistics*, New York: Wiley.

Rosenkrantz, W. A. (1969), "A Rate of Convergence for the Von Mises Statistic," *Transactions of American Mathematical Society*, 139, 329–337.

Rosenkrantz, W. A., and O'Reilly, N. E. (1972), "Applications of the Skorokhod Representation to Rates of Convergence of Linear Combinations of Order Statistics," *Annals of Mathematical Statistics*, 43, 1204–1212.

Satten, G. A. (1995) "Upper and Lower Bound Distributions That Give Simultaneous Confidence Intervals for Quantiles," *Journal of the American Statistical Association*, 90, 747–752.

Stephens, M. A. (1974), "EDF Statistics for Goodness-of-Fit and Some Comparisons," *Journal of the American Statistical Association*, 69, 730–737.

Wilks, S. S. (1962), *Mathematical Statistics*, New York: Wiley.