Simultaneous Confidence Bands and Regions for Log-Location-Scale Distributions with Censored Data

Luis A. Escobar
Dept. of Experimental Statistics
Louisiana State University
Baton Rouge, LA 70803

Yili Hong and William Q. Meeker
Dept. of Statistics
Iowa State University
Ames, IA 50011

April 4, 2008

Abstract

In many areas of application, especially life testing and reliability, it is often of interest to estimate an unknown cumulative distribution (cdf). A simultaneous confidence band (*SCB*) of the cdf can be used to assess the statistical uncertainty of the estimated cdf over the entire range of the distribution. Cheng and Iles (1983) presented a general approach of constructing an *SCB* for the cdf of a continuous random variable. For the log-location-scale family of distributions, there are explicit forms for the upper and lower boundaries of the *SCB*. In this article, we extend the work of Cheng and Iles (1983). We study the *SCB*s based on local information, expected information, and estimated expected information for both the "cdf method" and the "quantile method." We also study the effects of exceptional cases where a simple *SCB* does not exist. We describe calibration of the bands to provide exact coverage for complete data and type II censoring and better approximate coverage for other kinds of censoring. We also extend these procedures to regression analysis.

Keywords: Confidence region; information matrix; maximum likelihood; Quantile; Wald confidence interval.

1 Introduction

1.1 Background

It is often of primary interest to estimate an unknown cumulative distribution (cdf). Particular areas of application include life testing and reliability. Usually, it is important to assess the precision of the cdf estimate. Jeng and Meeker (2001) present two example applications for single distribution local-scale models; one is on life data and the other one is on probability of detection in which the usual simple regression model is replaced by a physics-based computer model so that there is only an unknown location parameter and an unknown scale parameter. One approach of describing the uncertainty of the estimated cdf is to construct a simultaneous confidence band (SCB) that contains the entire unknown cdf with a certain confidence level.

Cheng and Iles (1983) described a general method of constructing an SCB for the cdf of a continuous random variable. Their method is well suited for the location-scale and log-location-scale models, which include the most popular distribution families used in lifetime modeling. Their approach consists of two steps. First, one identifies a $100(1-\alpha)\%$ simultaneous confidence region (SCR), denoted by $CR(\boldsymbol{\theta})$, for the unknown parameters $\boldsymbol{\theta}$. The second step consists of obtaining the graph of the cdf $F(y;\boldsymbol{\theta})$ for all $\boldsymbol{\theta}$ in $CR(\boldsymbol{\theta})$, the S-shaped region in the plane swept by the graph is a SCB. Because the $CR(\boldsymbol{\theta})$ captures the true value of $\boldsymbol{\theta}$ with probability $1-\alpha$, the probability that the S-shaped region will capture the true cdf is at least $1-\alpha$. For the log-location-scale family there are closed-form expressions for the upper and lower boundaries of the SCB and with some mild conditions on the SCR, the coverage probability for the SCB is exactly $1-\alpha$.

1.2 Model, data, quantiles, and probabilities

The results of this paper apply to location-scale and log-location-scale distributions. A random variable Y belongs to the location-scale family of distributions, with location μ

and scale σ, if

$$F_Y(y;\mu,\sigma) = \Phi\left(\frac{y-\mu}{\sigma}\right), \quad -\infty < y < \infty \tag{1}$$

where $\Phi(z)$ is a cdf that does not depend on any unknown parameters, $-\infty < \mu < \infty$, and $\sigma > 0$. It can be shown that $\Phi(z)$ is the cdf of $(Y-\mu)/\sigma$. The normal (NOR), the smallest extreme value (SEV), and the logistic distributions are location-scale distributions.

A positive random variable T belongs to the log-location-scale family distribution if $Y = \log(T)$ is a member of the location-scale family. The lognormal, the Weibull, and the loglogistic are among the important distributions of this family. For example, the cdf and pdf of the Weibull random variable T are

$$F_T(t;\mu,\sigma) = \Phi_{\text{sev}}\left[\frac{\log(t)-\mu}{\sigma}\right] \quad \text{and} \quad f_T(t;\mu,\sigma) = \frac{1}{\sigma t}\phi_{\text{sev}}\left[\frac{\log(t)-\mu}{\sigma}\right]$$

where $\Phi_{\text{sev}}(z) = 1 - \exp[-\exp(z)]$ and $\phi_{\text{sev}}(z) = \exp[z - \exp(z)]$ are the standard smallest extreme value cdf and pdf, respectively. For the lognormal distribution, replace Φ_{sev} and ϕ_{sev} above with Φ_{nor} and ϕ_{nor}, the standard normal cdf and pdf, respectively.

Life tests often result in censored data. Type I (time) censored data result when unfailed units are removed from test at a prespecified time, perhaps due to limited time for study completion. Type II (failure) censored data result when a test is terminated after a specified number r of failures, say $2 \le r \le n$. If all units fail, the data are called "complete" or "uncensored" data.

Suppose that T is a lifetime from a log-location-scale distribution. Frequently, interest is on quantities like the failure probability $F_T(t_e;\mu,\sigma)$ at t_e or the p quantile t_p of the distribution. Define $y_e = \log(t_e)$ and $y_p = \log(t_p)$. It follows that $F_T(t_e;\mu,\sigma) = F_Y(y_e;\mu,\sigma) = \Phi\left[(y_e - \mu)/\sigma\right]$ and the p quantile of $F_Y(y)$ is $y_p = \mu + z_p\,\sigma$, where $z_p = \Phi^{-1}(p)$ is the p quantile of $\Phi(z)$.

1.3 Maximum likelihood estimation and information matrices

For a censored sample with n independent exact and right censored observations from a log-location-scale distribution, the likelihood of the data at $\boldsymbol{\theta} = (\mu, \sigma)'$ is

$$L(\boldsymbol{\theta}) = \mathcal{C} \prod_{i=1}^{n} \left\{ \frac{1}{\sigma t_i} \phi \left[\frac{\log(t_i) - \mu}{\sigma} \right] \right\}^{\delta_i} \left\{ 1 - \Phi \left[\frac{\log(t_i) - \mu}{\sigma} \right] \right\}^{1-\delta_i},$$

where $\delta_i = 1$ if t_i is an exact observation and $\delta_i = 0$ if t_i is a right censored observation, and $\mathcal{C}$ is a constant that does not depend on the unknown parameters. Standard computer software (e.g., JMP, MINITAB, SAS, S-PLUS/SPLIDA) provide maximum likelihood (ML) estimates of $\boldsymbol{\theta}$ and functions of $\boldsymbol{\theta}$ such as quantiles and probabilities. We denote the ML estimator of $\boldsymbol{\theta}$ by $\widehat{\boldsymbol{\theta}} = (\widehat{\mu}, \widehat{\sigma})'$. By the invariance property of ML estimators, the ML estimator of y_p is $\widehat{y}_p = \widehat{\mu} + z_p \widehat{\sigma}$. Similarly, the ML estimator of the cumulative probability of Y at y_e is $\widehat{p} = \Phi\left[(y_e - \widehat{\mu})/\widehat{\sigma}\right]$. See, for example, Chapter 8 in Meeker and Escobar (1998) for more details.

There are three kinds of information matrices that are used in statistical inference.

- The expected information matrix (also known as the Fisher information matrix) which usually depends on unknown parameters.

- The estimated information matrix is the ML estimator of the Fisher information matrix, obtained by evaluating the expected information matrix at the ML estimate $\widehat{\boldsymbol{\theta}}$.

- The observed information matrix, another estimator of the Fisher information matrix, is the negative Hessian matrix of the log-likelihood function, evaluated at the ML estimate $\widehat{\boldsymbol{\theta}}$.

The expected information matrix for $\boldsymbol{\theta}$ is

$$I_{\boldsymbol{\theta}} = \mathrm{E}\left[-\frac{\partial^2 \mathcal{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right] = \left(\frac{n}{\sigma^2} \right) \begin{bmatrix} f_{11} & f_{12} \\ f_{12} & f_{22} \end{bmatrix} = \left(\frac{n}{\sigma^2} \right) M. \tag{2}$$

Here $\mathcal{L}(\boldsymbol{\theta}) = \log[L(\boldsymbol{\theta})]$ is the log-likelihood of the data, M is the scaled information matrix with elements f_{ij}, $i, j = 1, 2$. Also, define the scaled covariance matrix Λ by

$$\Lambda = \begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{12} & \lambda_{22} \end{bmatrix} = \begin{bmatrix} f_{11} & f_{12} \\ f_{12} & f_{22} \end{bmatrix}^{-1} = M^{-1}. \tag{3}$$

Let $\widehat{I}_{\boldsymbol{\theta}}$ denote the estimated expected information matrix which is $\widehat{I}_{\boldsymbol{\theta}} = (n/\widehat{\sigma}^2)\, M$. Note that in the presence of censoring, f_{11}, f_{12}, f_{22} are functions of the proportion censored (Type II or failure censoring), or on the expected proportion censored (Type I or time censoring) but they do not depend on $\boldsymbol{\theta}$. When M depends on unknown parameters (e.g. Type I censoring), M is evaluated at $\widehat{\boldsymbol{\theta}}$ in which case, we would use notation $\widehat{M}, \widehat{f}_{ij}, \widehat{\Lambda}_{ij}$, and $\widehat{\lambda}_{ij}$ to denote the ML estimators of theses respective quantities. In the development of this paper, however, we avoid this extra notational burden.

The observed information matrix is given by

$$\breve{I}_{\boldsymbol{\theta}} = -\frac{\partial^2 \mathcal{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\bigg|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}} = \left(\frac{n}{\widehat{\sigma}^2}\right) \begin{bmatrix} \breve{\iota}_{11} & \breve{\iota}_{12} \\ \breve{\iota}_{12} & \breve{\iota}_{22} \end{bmatrix} = \left(\frac{n}{\widehat{\sigma}^2}\right) \breve{I} \tag{4}$$

where $\breve{I}$ is the scaled local information matrix with elements $\breve{\iota}_{ij}, i, j = 1, 2$. We use the notation $\breve{\Lambda}$ for the local estimate of the scaled covariance matrix, that is

$$\breve{\Lambda} = \begin{bmatrix} \breve{\lambda}_{11} & \breve{\lambda}_{12} \\ \breve{\lambda}_{21} & \breve{\lambda}_{22} \end{bmatrix} = \begin{bmatrix} \breve{\iota}_{11} & \breve{\iota}_{12} \\ \breve{\iota}_{12} & \breve{\iota}_{22} \end{bmatrix}^{-1} = \breve{I}^{-1}.$$

1.4 General approaches and contributions of this work

There are two alternative approaches for obtaining *SCB*s for the log-location-scale family using the general method proposed by Cheng and Iles (1983). We call them the "quantile method" and the "cdf method," respectively.

The "quantile method" obtains *SCB*s directly for quantiles. That is, for each $0 < p < 1$, the *SCB* for the p quantile is the solution to the optimization problems

$$\max_{\mu,\sigma} (\mu + z_p\sigma) \quad \text{and} \quad \min_{\mu,\sigma} (\mu + z_p\sigma) \tag{5}$$
$$\text{Subject to:} \quad (\mu, \sigma)' \in CR(\mu, \sigma).$$

The "cdf method" obtains SCB directly for cumulative probabilities. That is, for each $-\infty < y_e < \infty$, find the solution for the optimization problems

$$\max_{\mu,\sigma} \Phi\left(\frac{y_e - \mu}{\sigma}\right) \quad \text{and} \quad \min_{\mu,\sigma} \Phi\left(\frac{y_e - \mu}{\sigma}\right) \tag{6}$$
$$\text{Subject to:} \quad (\mu, \sigma)' \in CR(\mu, \sigma).$$

As shown in the appendix of Jeng and Meeker (2001), for a given CR, there is equivalence between optimization procedures (5) and (6), which means (5) and (6) give the same SCB for the cdf.

Cheng and Iles (1983) provided closed form solutions of (5) and (6) using expected information for complete data. In this work, we extend the work of Cheng and Iles (1983) in the following ways. We show how to compute SCBs based on local information, expected information, and estimated expected information for both the "cdf method" and the "quantile method." We show the effects of exceptional cases where the SCBs have non-finite boundaries. Cheng and Iles (1983) considered only complete data. We describe calibration of the intervals to provide exact coverage for type II censoring and approximate coverage for other kinds of censoring. We also show how to extend these procedures to regression analysis.

1.5 Related literature

Statistical methods for log-location-scale distributions, especially with application to lifetime studies are given, for example, in Meeker and Escobar (1998), and Lawless (2003). Cheng and Iles (1983) presented a general approach to construct SCBs for cdf of a continuous random variable. Cheng and Iles (1988) gave one-sided SCBs for a location-scale cdf with complete data. Jeng and Meeker (2001) compared coverage probabilities of SCBs based on a normal approximation, using observed information and expected information, likelihood, and bootstrap procedures. Their paper described the geometry of one versus two sided SCBs, and presented two examples. Sa and Lee (1998) developed SCB for the p quantile and the expected lifetime of the Weibull regression model by using the SCR

for the parameters and the Lagrange multiplier procedure. There is a substantial amount literature on simultaneous confidence intervals for a regression function (e.g. Miller, 1981), a problem different from that considered here.

1.6 Overview

The rest of paper is organized as follows. Sections 2 through 4 give single distribution *SCR*s and *SCB*s based on expected information, estimated expected information, and local Fisher information, respectively. Section 5 provides calibration of these *SCB*s. Section 6 shows how to extend the results to the regression model. Section 7 gives some concluding remarks and some areas for future research.

2 Single Distribution Simultaneous Confidence Regions Based on Expected Information

Here, we consider simultaneous inference for parameters, and functions of the parameters, of the single location-scale distribution $F_Y(y;\mu,\sigma) = \Phi\left[(y-\mu)/\sigma\right]$ based on the expected information matrix $I_{\boldsymbol{\theta}}$. An approximate $100(1-\alpha)\%$ *SCR* for $\boldsymbol{\theta} = (\mu,\sigma)'$ is given by $(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})' I_{\boldsymbol{\theta}} (\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}) \leq \gamma_F$ and can be re-expressed as

$$(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})' M (\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}) \leq \gamma_F^s \sigma^2 \tag{7}$$

where $\sigma > 0$, $\gamma_F^s = \gamma_F/n$, and $\gamma_F > 0$ is a constant chosen to target a required confidence level.

Theorem 1 *The SCR for $\boldsymbol{\theta} = (\mu,\sigma)'$ in (7) based on expected information has the following properties:*

1. *The simultaneous region is convex.*

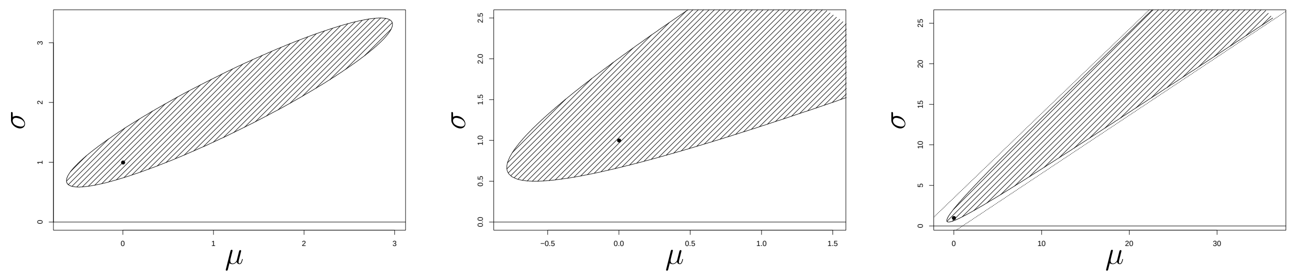

Figure 1: Possible shapes for the expected information *SCR*s.

2. *The minimum and maximum values of σ in the simultaneous region are*

$$\sigma_F^{\min} = \frac{\widehat{\sigma}}{1+\sqrt{\gamma_F^s \lambda_{22}}}, \qquad \sigma_F^{\max} = \begin{cases} \dfrac{\widehat{\sigma}}{1-\sqrt{\gamma_F^s \lambda_{22}}} & \text{if} \quad \gamma_F^s \lambda_{22} < 1 \\ \infty & \text{if} \quad \gamma_F^s \lambda_{22} \geq 1 \end{cases} \tag{8}$$

where λ_{22} is given in (3).

3. *The shape of the region is determined by the sign of $1 - \gamma_F^s \lambda_{22}$*

- *If $\gamma_F^s \lambda_{22} < 1$ the SCR is an ellipse and $\widehat{\sigma}/2 < \sigma_F^{\min} < \widehat{\sigma}$.*
- *If $\gamma_F^s \lambda_{22} = 1$ the SCR is the content of a parabola and $\sigma_F^{\min} = \widehat{\sigma}/2$.*
- *If $\gamma_F^s \lambda_{22} > 1$ the SCR is the content of a single branch from a hyperbola and $0 < \sigma_F^{\min} < \widehat{\sigma}/2$.*

Figure 1 illustrates the three possible shapes for the *SCR*s.

Proof: First we prove part *(1)*. Let $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ be two points in the region. Observe that $\boldsymbol{\theta}_i = (\mu_i, \sigma_i)'$ and $\sigma_i > 0$. Consider a convex combination of those two points say $a\boldsymbol{\theta}_1 + b\boldsymbol{\theta}_2$ where $0 \leq a \leq 1$ and $b = 1 - a$. Then

$$(\widehat{\boldsymbol{\theta}} - a\boldsymbol{\theta}_1 - b\boldsymbol{\theta}_2)' M (\widehat{\boldsymbol{\theta}} - a\boldsymbol{\theta}_1 - b\boldsymbol{\theta}_2) = \left[a(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_1) + b(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_2)\right]' M \left[a(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_1) + b(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_2)\right]$$

$$= a^2(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_1)' M (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_1)' + 2ab(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_1)' M (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_2) + b^2(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_2)' M (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_2)'$$

$$\leq a^2 \gamma_F^s \sigma_1^2 + a^2 \gamma_F^s \sigma_2^2 + 2ab\sqrt{\gamma_F^s \sigma_1^2 \gamma_F^s \sigma_2^2} = \gamma_F^s (a\sigma_1 + b\sigma_2)^2.$$

The above inequality holds because $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ are in the region and M is positive definite. Thus the region is convex.

To prove part *(2)*, note that the maximum and minimum of σ must be attained at the boundary of the *CR*. By implicit differentiation, $(\widehat{\mu}-\mu)f_{11}+(\widehat{\sigma}-\sigma)f_{12}+[(\widehat{\mu}-\mu)f_{12}+(\widehat{\sigma}-\sigma)f_{22}+\gamma_F^s\sigma](d\sigma/d\mu)=0$. Setting $d\sigma/d\mu=0$, one gets $(\widehat{\mu}-\mu)=-(\widehat{\sigma}-\sigma)f_{12}/f_{11}$. Substituting this into (7), and after some simplification, one gets $(1-\gamma_F^s\lambda_{22})\sigma^2-2\widehat{\sigma}\sigma+\widehat{\sigma}^2=0$. If $\gamma_F^s\lambda_{22}\neq 1$, the two roots are $\widehat{\sigma}/\left(1\pm\sqrt{\gamma_F^s\lambda_{22}}\right)$. If $\gamma_F^s\lambda_{22}=1$, the root is $\widehat{\sigma}/2$. After noting that only positive values of σ are allowed, one can see that the results in part *(2)* hold.

Now, we proceed to prove part *(3)*. Direct computations show that $(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})'M(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})=\gamma_F^s\sigma^2$ can be written as

$$(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})'R(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})+2\widehat{\sigma}\gamma_F^s\boldsymbol{\delta}'(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})=\gamma_F^s\widehat{\sigma}^2 \tag{9}$$

where $\boldsymbol{\delta}=(0,1)'$ and

$$R=\begin{bmatrix} f_{11} & f_{12} \\ f_{12} & (f_{22}-\gamma_F^s) \end{bmatrix}. \tag{10}$$

Now, let $\zeta_1\geq\zeta_2$ be the eigenvalues of R and let O be the corresponding matrix of eigenvectors. Thus $O'RO=\mathrm{diag}(\zeta_1,\zeta_2)$ and OO' is equal to a 2×2 identity matrix. Thus from (9)

$$\begin{aligned}(w_1,w_2)\mathrm{diag}(\zeta_1,\zeta_2)(w_1,w_2)'-2(v_1,v_2)(w_1,w_2)'&=\gamma_F^s\widehat{\sigma}^2\\ \zeta_1w_1^2+\zeta_2w_2^2-2v_1w_1-2v_2w_2&=\gamma_F^s\widehat{\sigma}^2\end{aligned} \tag{11}$$

where $(w_1,w_2)=(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})'O$ and $(v_1,v_2)=-\gamma_F^s\widehat{\sigma}\boldsymbol{\delta}'O$.

Let $D=\det(R)$. If $D=0$ then $f_{22}-\gamma_F^s=f_{12}^2/f_{11}$. In this case, $\zeta_1=f_{11}+f_{12}^2/f_{11}$, and $\zeta_2=0$. A matrix of eigenvectors for R is

$$O=\frac{1}{\sqrt{f_{11}\zeta_1}}\begin{bmatrix} f_{11} & -f_{12} \\ f_{12} & f_{11} \end{bmatrix}.$$

Therefore, $v_1 = -\gamma_F^s \widehat{\sigma} f_{12}/\sqrt{f_{11}\zeta_1}$ and $v_2 = -\gamma_F^s \widehat{\sigma} f_{11}/\sqrt{f_{11}\zeta_1} \neq 0$, and the quadratic in (11) has the parabolic shape

$$\zeta_1 \left(w_1 - \frac{v_1}{\zeta_1} \right)^2 = 2v_2 \left(w_2 + \frac{\gamma_F^s \widehat{\sigma}^2 + v_1^2/\zeta_1}{2v_2} \right). \tag{12}$$

If $D \neq 0$ the two eigenvalues are different from 0 and the quadratic in (11) can be expressed as

$$\zeta_1 \left(w_1 - \frac{v_1}{\zeta_1} \right)^2 + \zeta_2 \left(w_2 - \frac{v_2}{\zeta_2} \right)^2 = \gamma_F^s \widehat{\sigma}^2 + (v_1, v_2)\mathrm{diag}(\zeta_1^{-1}, \zeta_2^{-1})(v_1, v_2)' = \frac{\gamma_F^s \widehat{\sigma}^2}{1 - \gamma_F^s \lambda_{22}}. \tag{13}$$

When $D < 0$, ζ_1 and ζ_2 have different signs ($\zeta_2 < 0 < \zeta_1$) and the curve in (13) is a hyperbola. When $D > 0$, the two eigenvalues are positive and the curve in (13) is an ellipse.□

Lemma 1 *The ellipsoidal expected information based SCR for $\boldsymbol{\theta}$ in (7) can be expressed as the set of all the values $\boldsymbol{\theta} = (\mu, \sigma)'$, such that $(R^{-1}M\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})'R(R^{-1}M\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \leq k$ where R is defined in (10) and $k = \widehat{\sigma}^2 \gamma_F^s/(1 - \gamma_F^s \lambda_{22})$.*

When $D = f_{11}(f_{22} - \gamma_F^s) - f_{12}^2 > 0$, R is positive definite and $k > 0$. In this case, as shown in **Theorem 1**, the *SCR* includes only positive values of σ. Cheng and Iles (1983) used the same condition to ensure that R is positive definite.

Proof: Using $O'RO = \mathrm{diag}(\zeta_1, \zeta_2)$ the ellipse in (13) can be written as follows

$$\left(w_1 - \frac{v_1}{\zeta_1}, \quad w_2 - \frac{v_2}{\zeta_2} \right) O'RO \left(w_1 - \frac{v_1}{\zeta_1}, \quad w_2 - \frac{v_2}{\zeta_2} \right)' = \frac{\gamma_F^s \widehat{\sigma}^2}{1 - \gamma_F^s \lambda_{22}} \; . \tag{14}$$

Now,

$$O \left(w_1 - \frac{v_1}{\zeta_1}, \quad w_2 - \frac{v_2}{\zeta_2} \right)' = O(w, w_2)' - O \left(\frac{v_1}{\zeta_1}, \quad \frac{v_2}{\zeta_2} \right)' = \widehat{\boldsymbol{\theta}} - \boldsymbol{\theta} - O\mathrm{diag}(\zeta_1^{-1}, \zeta_2^{-1})(v_1, v_2)'$$
$$= (I + \gamma_F^s R^{-1} \boldsymbol{\delta}\boldsymbol{\delta}')\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta} = R^{-1}M\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}.$$

Substituting this expression into (14), one gets that the expected information *SCR* is $\left(R^{-1}M\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right)' R \left(R^{-1}M\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right) = \gamma_F^s \widehat{\sigma}^2/(1 - \gamma_F^s \lambda_{22})$.□

Define

$$\widehat{\boldsymbol{\tau}} = (\widehat{\tau}_1, \widehat{\tau}_2)' = R^{-1}M\widehat{\boldsymbol{\theta}} = \left[I + \gamma_F^s R^{-1}\boldsymbol{\delta}\boldsymbol{\delta}'\right]\widehat{\boldsymbol{\theta}} = \widehat{\boldsymbol{\theta}} + \frac{\widehat{\sigma}\gamma_F^s}{1-\gamma_F^s\lambda_{22}}\,(\lambda_{12},\, \lambda_{22})'. \qquad (15)$$

Then $\widehat{\tau}_1 = \widehat{\mu} + \lambda_{12}\widehat{\sigma}\gamma_F^s/[1-\gamma_F^s\lambda_{22}]$ and $\widehat{\tau}_2 = \widehat{\sigma}/[1-\gamma_F^s\lambda_{22}]$. The center of the ellipse in (14) is away from $\widehat{\boldsymbol{\theta}}$ by $(\widehat{\boldsymbol{\tau}} - \widehat{\boldsymbol{\theta}}) = (\widehat{\sigma}\gamma_F^s)R^{-1}\boldsymbol{\delta}$ which is proportional to $\widehat{\sigma}$ and does not depend on $\widehat{\mu}$.

2.1 Confidence bands for quantiles

Theorem 2 *An approximate $100(1-\alpha)\%$ SCB for the quantiles $y_p, 0 < p < 1$, based on expected information is*

1. *If $\gamma_F^s\lambda_{22} < 1$ (i.e., the SCR for $(\mu, \sigma)'$ is ellipsoidal)*

$$[\underset{\sim}{y_p}, \quad \widetilde{y}_p] = \widehat{y}_p + \widehat{\sigma}\left[h_1(\Lambda, p) \mp h_2(\Lambda, p)\right] \qquad (16)$$

where

$$h_1(\Lambda, p) = \frac{\gamma_F^s(\lambda_{12} + z_p\lambda_{22})}{1-\gamma_F^s\lambda_{22}} \qquad (17)$$

$$h_2(\Lambda, p) = \frac{\sqrt{\gamma_F^s\left(\lambda_{11} + 2z_p\lambda_{12} + z_p^2\lambda_{22}\right) - (\gamma_F^s)^2\left(\lambda_{11}\lambda_{22} - \lambda_{12}^2\right)}}{1-\gamma_F^s\lambda_{22}} \qquad (18)$$

and λ_{ij}, $i, j = 1, 2$ are the elements of the scaled covariance matrix defined in (3).

2. *If $\gamma_F^s\lambda_{22} = 1$ (i.e., the SCR for $(\mu, \sigma)'$ is parabolic) the SCB for y_p is*

$$[\underset{\sim}{y_p}, \quad \widetilde{y}_p] = \begin{cases} [-\infty, \quad \widehat{y}_p + \widehat{\sigma}\, g_1(\Lambda, p)] & \text{if} \quad p < p_b \\ [\widehat{y}_p - \widehat{\sigma}\, g_1(\Lambda, p), \quad \infty] & \text{if} \quad p > p_b \\ [-\infty, \quad \infty] & \text{if} \quad p = p_b \end{cases} \qquad (19)$$

where $p_b = \Phi\left(-\lambda_{12}/\lambda_{22}\right)$ is the critical value of the SCB , and $g_1(\Lambda, p) = \gamma_F^s[\det(\Lambda) + (\lambda_{12} + \lambda_{22}z_p)^2]/(2|\lambda_{12} + \lambda_{22}z_p|)$. The SCB boundaries in (19) are the limit as $\gamma_F^s \to (1/\lambda_{22})^-$ of the boundaries in (16). The boundaries in (19) have the following property $\lim_{p\to p_b^+} \underset{\sim}{y_p} = -\infty$ and $\lim_{p\to p_b^-} \widetilde{y}_p = \infty$.

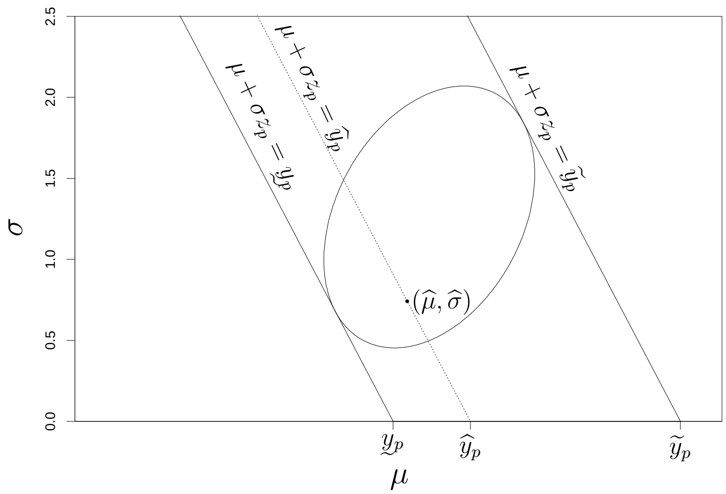

Figure 2: Graphical illustration of the quantile method for finding the boundaries of an *SCB* for one particular value of p for an ellipsoidal expected information *SCR*.

3. If $\gamma_F^s \lambda_{22} > 1$ (i.e., the SCR for $(\mu, \sigma)'$ is hyperbolic) the SCB for y_p is

$$[\underset{\sim}{y_p}, \quad \underset{\sim}{y_p}] =$$
$$\begin{cases} (-\infty, \quad \widehat{y}_p + \widehat{\sigma}\left\{h_1(\Lambda, p) + h_2(\Lambda, p)\right\}] & \textit{if} \quad p < \Phi\left[\dfrac{-\lambda_{12} - \sqrt{(\gamma_F^s \lambda_{22} - 1)\det(\Lambda)}}{\lambda_{22}}\right] \\ [\widehat{y}_p + \widehat{\sigma}\left\{h_1(\Lambda, p) - h_2(\Lambda, p)\right\}, \infty) & \textit{if} \quad p > \Phi\left[\dfrac{-\lambda_{12} + \sqrt{(\gamma_F^s \lambda_{22} - 1)\det(\Lambda)}}{\lambda_{22}}\right] \\ (-\infty, \quad \infty) & \textit{otherwise} \end{cases} . \tag{20}$$

Figure 2 illustrates the quantile method for finding boundaries of an *SCB* for a specific value of p based on an ellipsoidal *SCR*. Figure 3 shows the corresponding *SCB* using the quantile method and an ellipsoidal expected information *SCR*. Figure 4 illustrates the quantile method for finding boundaries of an *SCB* for two particular values of p based on a parabolic *SCR*. Figure 5 shows the corresponding *SCB* from (19) based on a parabolic *SCR*. Figure 6 illustrates the quantile method for finding boundaries of an *SCB* for two particular values of p based on a hyperbolic *SCR*. Figure 7 shows the corresponding *SCB*

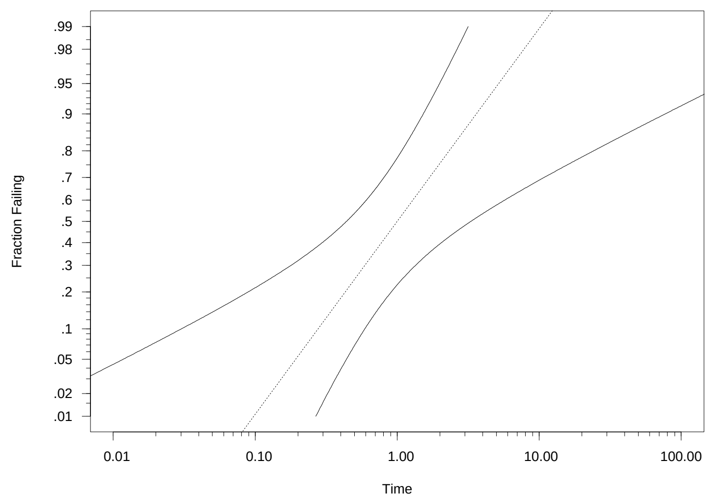

Figure 3: Lognormal probability plot illustrating the *SCB* for a lognormal distribution using the quantile method or the cdf method based on an ellipsoidal expected information *SCR*.

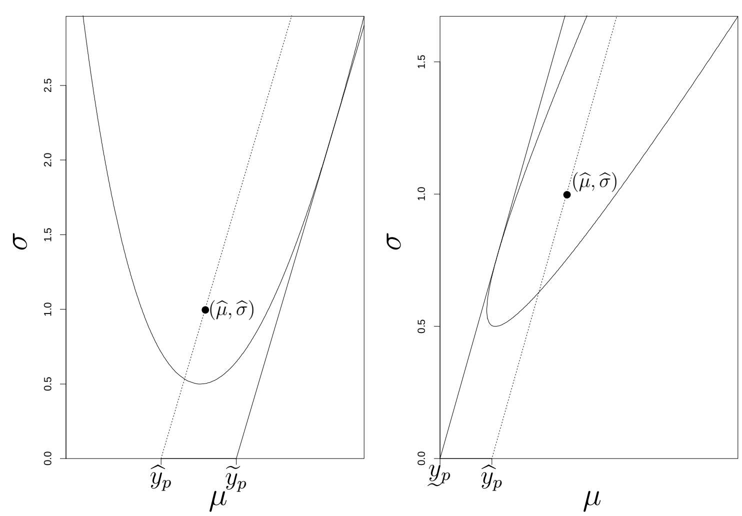

Figure 4: Graphical illustration of the quantile method for finding boundaries of an *SCB* for two particular values of p based on a parabolic *SCR* ($p = 0.2, p_b = 0.429$ on the left for the upper boundary and $p = 0.2, p_b = 0.054$ on the right for the lower boundary).

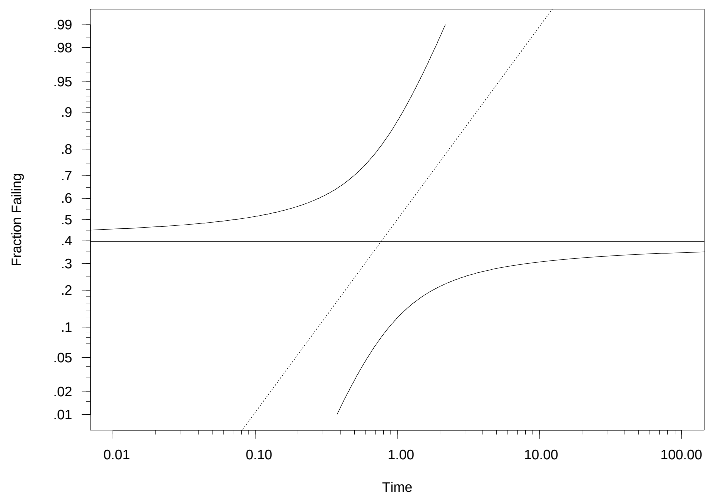

Figure 5: Lognormal probability plot illustrating an *SCB* for a lognormal distribution using the quantile method with critical value $p_b = 0.397$ based on a parabolic *SCR*.

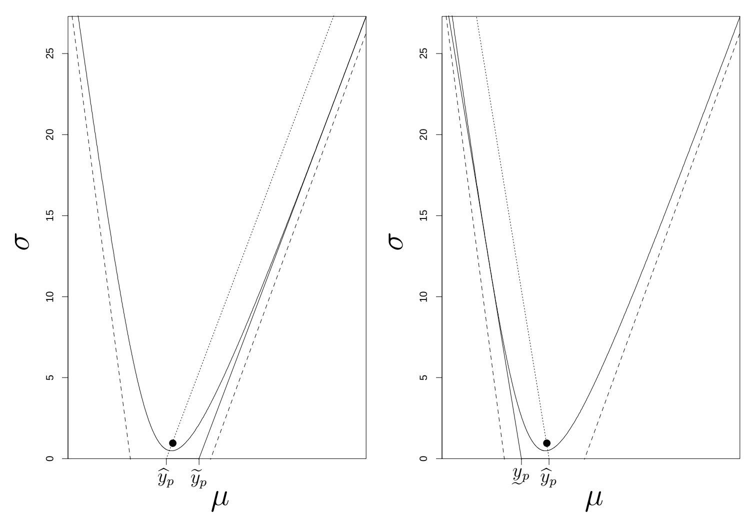

Figure 6: Graphical illustration of the quantile method for finding boundaries of an *SCB* for two particular values of p based on a hyperbolic *SCR* (the dots in the plot are the ML estimates $(\widehat{\mu}, \widehat{\sigma})'$, the dashed lines are the asymptotes of the hyperbolic curves, $p = 0.2$ on the left for the upper boundary and $p = 0.6$ on the right for the lower boundary. The $x-$intercept of the dotted lines give the ML estimates of the quantiles for these particular p's).

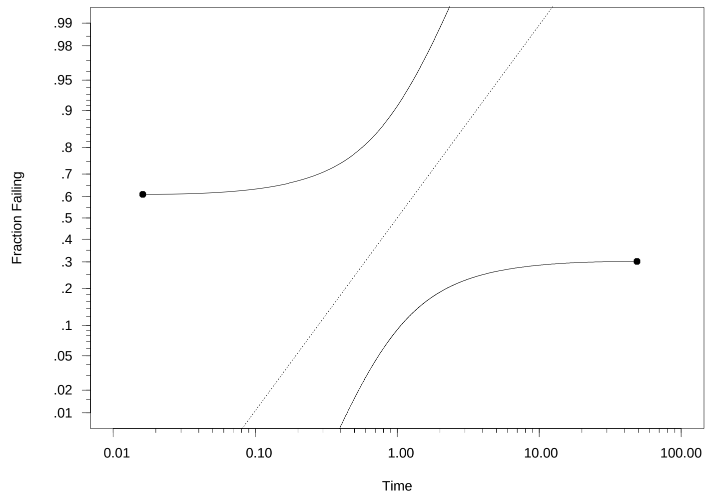

Figure 7: Lognormal probability plot illustrating an *SCB* for a lognormal distribution using the quantile method based on a hyperbolic *SCR*. The *SCB* boundaries for those values of p between the limits in equation (20) are $(-\infty, \infty)$. These limits on p are indicated by dots in this figure.

based on a hyperbolic *SCR*. Appendix A.3 gives a proof of **Theorem 2**.

It can be shown that the boundaries of the *SCB*s in **Theorem 2** are all monotone increasing in p. The *SCB* in (16) is equivalent to, but expressed differently than the *SCB* in Cheng and Iles (1983). To see the equivalence of the formulas for the upper boundary, from equation (16), we have

$$\widetilde{y}_p = \widehat{y}_p + \widehat{\sigma}[h_1(\Lambda, p) + h_2(\Lambda, p)] = \widehat{y}_p + \widehat{\sigma}\left(\frac{[h_1(\Lambda, p)]^2 - [h_2(\Lambda, p)]^2}{h_1(\Lambda, p) - h_2(\Lambda, p)}\right).$$

Using (17)-(18) and after simplifications

$$\widetilde{y}_p = \widehat{y}_p + \widehat{\sigma}\frac{(f_{22} - 2f_{12}z_p + f_{11}z_p^2)}{\sqrt{\left(\frac{n}{\gamma}\left(f_{22} - 2f_{12}z_p + f_{11}z_p^2\right) - 1\right)(f_{11}f_{22} - f_{12}^2)} - (-f_{12} + f_{11}z_p)}$$

which is the formula given by Cheng and Iles (1983). The proof for $\underset{\sim}{y}_p$ is similar.

2.2 Confidence bands for cumulative probabilities

Theorem 3 *An approximate* $100(1-\alpha)\%$ *SCB for the cumulative probabilities* $p = F(y_e;\mu,\sigma), -\infty < y_e < \infty$, *based on the expected information, is given by*

1. *If* $\gamma_F^s\lambda_{22} < 1$ *(i.e., the SCR for* $(\mu,\sigma)'$ *is ellipsoidal)*

$$[\underset{\sim}{p}, \quad \widetilde{p}] = \left[\Phi(\underset{\sim}{a}_F), \quad \Phi(\widetilde{a}_F)\right] \tag{21}$$

where $\underset{\sim}{a}_F = z_{\widehat{p}} - h_3(\Lambda,\widehat{p})$, $\widetilde{a}_F = z_{\widehat{p}} + h_3(\Lambda,\widehat{p})$,

$$h_3(\Lambda,\widehat{p}) = \sqrt{\gamma_F^s\left(\lambda_{11} + 2z_{\widehat{p}}\lambda_{12} + z_{\widehat{p}}^2\lambda_{22}\right)}, \tag{22}$$

$\widehat{p} = \Phi[(y_e - \widehat{\mu})/\widehat{\sigma}]$, *and* $z_{\widehat{p}} = \Phi^{-1}(\widehat{p}) = (y_e - \widehat{\mu})/\widehat{\sigma}$.

2. *If* $\gamma_F^s\lambda_{22} = 1$ *(i.e., the SCR for* $(\mu,\sigma)'$ *is parabolic) the upper and lower bounds for* p *are still given by (21) with* $\gamma_F^s = 1/\lambda_{22}$. *In this case,* $\lim_{y_e\to\infty}\underset{\sim}{p} = \lim_{y_e\to-\infty}\widetilde{p} = \Phi\left(\lambda_{12}/\lambda_{22}\right)$. *This SCB is the limit as* $\gamma_F^s \to (1/\lambda_{22})^+$ *of the SCB in (23).*

3. *If* $\gamma_F^s\lambda_{22} > 1$ *(i.e., the SCR for* $(\mu,\sigma)'$ *is hyperbolic) then*

$$(\underset{\sim}{p}, \quad \widetilde{p}) = \begin{cases} (\Phi(\underset{\sim}{a}_F), \quad p_U) & \text{if} \quad \widehat{p} \le p_1 \\ (\Phi(\underset{\sim}{a}_F), \quad \Phi(\widetilde{a}_F)) & \text{if} \quad p_1 < \widehat{p} < p_2 \\ (p_L, \quad \Phi(\widetilde{a}_F)) & \text{if} \quad p_2 < \widehat{p} \end{cases} \tag{23}$$

where

$$p_L = \Phi\left[\frac{-\lambda_{12} - \sqrt{(\lambda_{22}\gamma_F^s - 1)\det(\Lambda)}}{\lambda_{22}}\right], \quad p_U = \Phi\left[\frac{-\lambda_{12} + \sqrt{(\lambda_{22}\gamma_F^s - 1)\det(\Lambda)}}{\lambda_{22}}\right]$$

$$p_1 = \Phi\left[-\frac{\lambda_{12}}{\lambda_{22}} - \frac{1}{\lambda_{22}}\sqrt{\frac{\det(\Lambda)}{\lambda_{22}\gamma_F^s - 1}}\right], \qquad p_2 = \Phi\left[-\frac{\lambda_{12}}{\lambda_{22}} + \frac{1}{\lambda_{22}}\sqrt{\frac{\det(\Lambda)}{\lambda_{22}\gamma_F^s - 1}}\right].$$

Appendix A.3 gives a proof of **Theorem 3**. Figures 8, 9, and 10 illustrate the cdf method for finding boundaries of an *SCB* for one particular value of y_e for an elliptical *SCR*, a parabolic *SCR*, and a hyperbolic *SCR*, respectively.

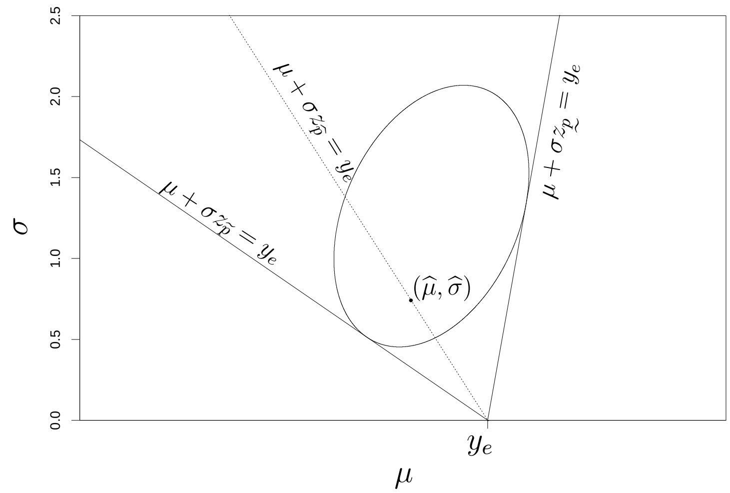

Figure 8: Graphical illustration of the cdf method for finding the boundaries of an *SCB* for one particular value of y_e, based on an elliptical *SCR*.

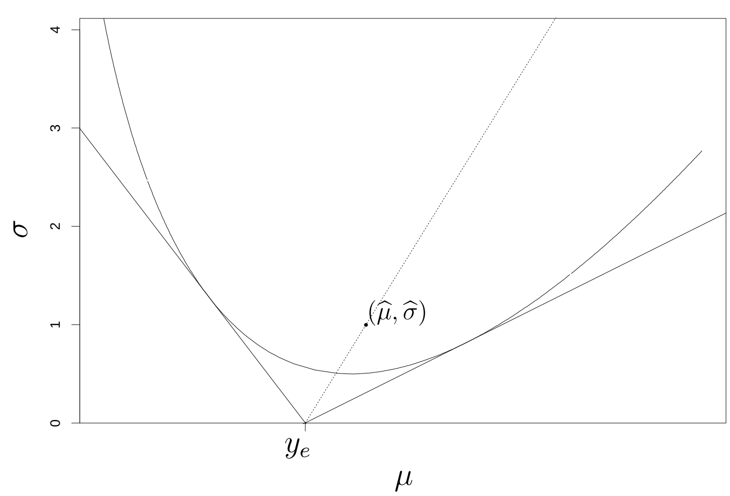

Figure 9: Graphical illustration of the cdf method for finding the boundaries of an *SCB* for one particular value of y_e based on a parabolic *SCR*.

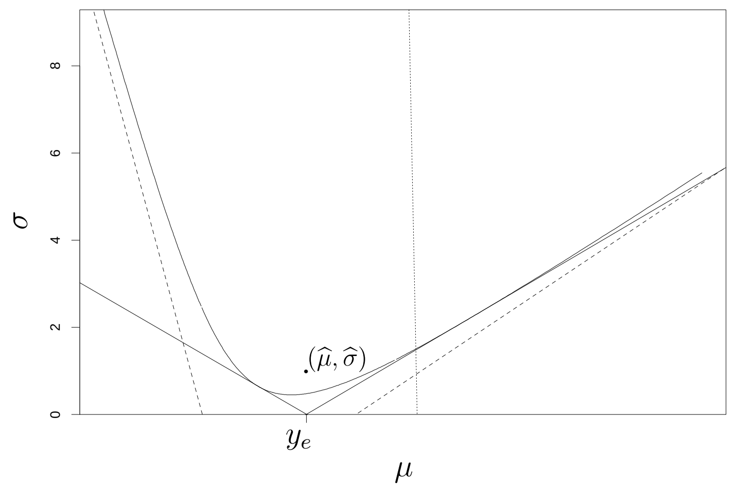

Figure 10: Graphical illustration of the cdf method for finding the boundaries of an *SCB* for one particular value of y_e based on a hyperbolic *SCR* (the dashed lines are the asymptotes of the hyperbolic curve).

As shown in **Lemma 4**, the boundaries of the *SCB* in (21) are obtained by inverting the boundaries of the *SCB* in (16). In particular, the *SCB* boundaries $\underset{\sim}{p}$ and $\tilde{p}$, as shown in Figure 8 are the solutions to (see equation (16)) $y_e = \widehat{\mu} + \widehat{\sigma}[z_{\underset{\sim}{p}} + h_1(\Lambda, \underset{\sim}{p}) - h_2(\Lambda, \underset{\sim}{p})]$ and $y_e = \widehat{\mu} + \widehat{\sigma}\left[z_{\widetilde{p}} + h_1(\Lambda, \widetilde{p}) + h_2(\Lambda, \widetilde{p})\right]$. The *SCB* in (21) is equivalent to, but expressed differently than, that given in Cheng and Iles (1983). To see this, using the definition of h_3 given in (22), we get

$$\widetilde{a} = z_{\widehat{p}} + \sqrt{\gamma_F^s\left(\lambda_{11} + 2z_{\widehat{p}}\lambda_{12} + z_{\widehat{p}}^2\lambda_{22}\right)} = z_{\widehat{p}} + \sqrt{\frac{\gamma}{nf_{11}}\left(\frac{f_{11}f_{22} - 2z_{\widehat{p}}f_{11}f_{12} + z_{\widehat{p}}^2 f_{11}^2}{f_{11}f_{22} - f_{12}^2}\right)}$$

$$= z_{\widehat{p}} + \sqrt{\frac{\gamma}{nf_{11}}\left[1 + \frac{(f_{11}z_{\widehat{p}} - f_{12})^2}{f_{11}f_{22} - f_{12}^2}\right]}$$

which agrees with the formula in Cheng and Iles (1983). The proof for $\underset{\sim}{a}$ is similar. It can be shown that the boundaries of the *SCB* in **Theorem 3** are all monotone increasing in y_e.

3 Single Distribution Simultaneous Confidence Regions Based on Estimated Expected Information

An approximate $100(1-\alpha)\%$ *SCR* for $\boldsymbol{\theta} = (\mu, \sigma)'$ based on the estimated expected information $\widehat{I}_{\boldsymbol{\theta}}$ is given by $(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})' \widehat{I}_{\boldsymbol{\theta}} (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \leq \gamma_E$ and can be re-expressed as

$$(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})' M (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \leq \gamma_E^s \widehat{\sigma}^2 \tag{24}$$

where $\widehat{I}_{\boldsymbol{\theta}}$ is the ML estimate of the information matrix, γ_E is a properly chosen constant to provide the required confidence, and $\gamma_E^s = \gamma_E / n$. To ensure that the confidence region does not include negative values of σ one must choose γ_E to be small enough that

$$f_{11}(f_{22} - \gamma_E^s) - f_{12}^2 > 0. \tag{25}$$

When γ_E satisfies this condition, the minimum and maximum values of σ in the confidence region are $\sigma_E^{\min} = \widehat{\sigma}(1 - \sqrt{\gamma_E^s \lambda_{22}})$ and $\sigma_E^{\max} = \widehat{\sigma}(1 + \sqrt{\gamma_E^s \lambda_{22}})$, respectively.

3.1 Confidence band for quantiles

Theorem 4 *An approximate* $100(1-\alpha)\%$ *SCB for the quantiles* $y_p = \mu + z_p \sigma$ *of* $F(y; \mu, \sigma), 0 < p < 1$, *based on estimated expected information can be expressed as*

$$[\underset{\sim}{y_p}, \quad \widetilde{y}_p] = \widehat{\mu} + \widehat{\sigma}\left[z_p \mp h_3(\Lambda, p)\right] \tag{26}$$

where $h_3(\Lambda, p) = \sqrt{\gamma_E^s \left(\lambda_{11} + 2 z_p \lambda_{12} + z_p^2 \lambda_{22}\right)}$.

See Appendix A.4 for the proof.

3.2 Confidence band for cumulative probabilities

Theorem 5 *An approximate* $100(1-\alpha)\%$ *SCB for the probabilities* $p = F(y_e; \mu, \sigma), -\infty < y_e < \infty$, *based on estimated expected information is obtained by inverting the SCB in (26). Appendix A.4 shows that the SCB is*

$$[\underset{\sim}{p}, \quad \widetilde{p}] = [\Phi(\underset{\sim}{a_E}), \quad \Phi(\widetilde{a}_E)] \tag{27}$$

where $\underset{\sim}{a}_E = z_{\widehat{p}} + h_1(\Lambda, \widehat{p}) - h_2(\Lambda, \widehat{p})$, $\widetilde{a}_E = z_{\widehat{p}} + h_1(\Lambda, \widehat{p}) + h_2(\Lambda, \widehat{p})$, $\widehat{p} = \Phi[(y_e - \widehat{\mu})/\widehat{\sigma}]$, *and* $z_{\widehat{p}} = \Phi^{-1}(\widehat{p}) = (y_e - \widehat{\mu})/\widehat{\sigma}$. *The functions* $h_1(\cdot,\cdot)$ *and* $h_2(\cdot,\cdot)$ *are as defined in (17)-(18), but here they are evaluated at* $(\Lambda, \widehat{p})$.

It can be shown that the *SCB* in (27) is obtained by inverting the *SCB* in (26).

4 Single Distribution Simultaneous Confidence Regions Based on Observed Information

A "Wald" simultaneous approximate $100(1-\alpha)\%$ *SCR* for $\boldsymbol{\theta} = (\mu, \sigma)'$ based on the observed information $\breve{I}_{\boldsymbol{\theta}}$ has the form $(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})'\breve{I}_{\boldsymbol{\theta}}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \leq \gamma_O$ and can be re-expressed as

$$(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})'\breve{I}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \leq \gamma_O^s \widehat{\sigma}^2 \tag{28}$$

where $\breve{I}_{\boldsymbol{\theta}}$ is the observed information matrix and $\breve{I}$ is defined in (4). To ensure that the Wald region does not include negative values of σ, it is required to choose γ_O such that elements of $\breve{I}$ satisfies the restriction

$$\breve{\iota}_{11}(\breve{\iota}_{22} - \gamma_O^s) - \breve{\iota}_{12}^2 > 0 \tag{29}$$

where $\gamma_O^s = \gamma_O/n$. When γ_O satisfies this condition, the minimum and maximum values of σ in the *SCR* are $\sigma_E^{\min} = \widehat{\sigma}\left(1 - \sqrt{\gamma_O^s \breve{\lambda}_{22}}\right)$ and $\sigma_E^{\max} = \widehat{\sigma}\left(1 + \sqrt{\gamma_O^s \breve{\lambda}_{22}}\right)$.

4.1 Confidence band for quantiles

When the *SCR* is based on observed information, the *SCB* for quantiles is given as follows.

Theorem 6 *An approximate* $100(1-\alpha)\%$ *SCB for the quantiles* $y_p = \mu + z_p\sigma$ *of* $F(y;\mu,\sigma), 0 < p < 1$, *based on observed information is*

$$[\underset{\sim}{y}_p, \quad \widetilde{y}_p] = \widehat{\mu} + \widehat{\sigma}\left[z_p \mp h_3(\breve{\Lambda}, p)\right] \tag{30}$$

where $h_3(\breve{\Lambda}, p) = \sqrt{\gamma_O^s\left(\breve{\lambda}_{11} + 2z_p\breve{\lambda}_{12} + z_p^2\breve{\lambda}_{22}\right)}$.

4.2 Confidence band for cumulative probabilities

When the *SCR* is based on observed information, the *SCB* for cumulative probabilities is given as follows.

Theorem 7 *An approximate* $100(1-\alpha)\%$ *SCB for the probabilities* $p = F(y_e; \mu, \sigma)$, $-\infty < y_e < \infty$, *based on observed information is*

$$[\underset{\sim}{p}, \quad \widetilde{p}] = [\Phi(\underset{\sim}{a}_O), \quad \Phi(\tilde{a}_O)] \tag{31}$$

where $\underset{\sim}{a}_O = z_{\widehat{p}} + h_1(\breve{\Lambda}, \widehat{p}) - h_2(\breve{\Lambda}, \widehat{p})$, $\widetilde{a}_O = z_{\widehat{p}} + h_1(\breve{\Lambda}, \widehat{p}) + h_2(\breve{\Lambda}, \widehat{p})$, $\widehat{p} = \Phi[(y_e - \widehat{\mu})/\widehat{\sigma}]$, *and* $z_{\widehat{p}} = \Phi^{-1}(\widehat{p}) = (y_e - \widehat{\mu})/\widehat{\sigma}$. *The functions* $h_1(\cdot, \cdot)$ *and* $h_2(\cdot, \cdot)$ *are defined in (17) and (18), but here they are evaluated at* $(\breve{\Lambda}, \widehat{p})$.

It can be shown that the *SCB* in (31) is obtained by inverting the *SCB* in (30). The proofs of **Theorems 6** and **7** are very similar to the proofs of **Theorems 4** and **5**, and thus are omitted.

5 Calibration of the Simultaneous Regions

For the log-location-scale family, it can be shown that the *SCB* has the same coverage probability as the corresponding *SCR* if only if the *SCR* is convex and there exists a $p_0 \in (0, 1)$ such that at least one of the boundaries $\min_{(\mu,\sigma)' \in SCR}(\mu + z_{p_0}\sigma)$ or $\max_{(\mu,\sigma)' \in SCR}(\mu + z_{p_0}\sigma)$ is finite. All the *SCR*s considered in this paper satisfy these two conditions. Thus it suffices to calibrate the *SCR*s. Sections 5.1 to 5.3 show how to use simulation to obtain the values of $\gamma_F, \gamma_E, \gamma_O$, respectively, needed in equations (7), (24) and (28) to obtain $100(1-\alpha)\%$ *SCR*s. The coverage probability of the *SCR*s is exact for complete or Type II censored data and approximate for the Type I censored data. Section 5.4 provides an analytical procedure to obtain γ when the data are a complete (uncensored) sample from a normal distribution. Our Section 5.4 is related to Section 3.2 of Cheng and Iles (1983) who considered only the *SCR* based on expected information.

5.1 The expected information *SCR*

For Type II (failure) censored data, the ellipsoidal *SCR* based on expected information in (7) can be written as

$$CR_F = \left\{(\mu, \sigma)' : f_{11}\, L_F^2 + 2\, f_{12}\, L_F S_F + f_{22}\, S_F^2 \leq \gamma_F\right\} \tag{32}$$

where $L_F = \sqrt{n}\,[(\widehat{\mu} - \mu)/\sigma]$ and $S_F = \sqrt{n}\,[(\widehat{\sigma} - \sigma)/\sigma]$ have distributions that depend on the number of failures, r, the sample size, n, and the distribution $\Phi(z)$, but they do not depend on the unknown parameters $(\mu, \sigma)'$. Therefore, for given (n, r) and $\Phi(z)$ one can approximate the distribution of CR_F by using simulation. Thus with γ_F equal to the $100(1-\alpha)\%$ quantile of the distribution of CR_F, (7) provides an exact (except for Monte Carlo error) $100(1-\alpha)\%$ *SCR* for $\boldsymbol{\theta} = (\mu, \sigma)'$.

5.2 The estimated expected information *SCR*

For Type II (failure) censored data, the *SCR* in (24) based on estimated expected information can be written as

$$CR_E = \left\{(\mu, \sigma)' : f_{11}\, L_E^2 + 2\, f_{12}\, L_E S_E + f_{22}\, S_E^2 \leq \gamma_E\right\} \tag{33}$$

where $L_E = \sqrt{n}\,[(\widehat{\mu} - \mu)/\widehat{\sigma}]$ and $S_E = \sqrt{n}\,[(\widehat{\sigma} - \sigma)/\widehat{\sigma}]$ have distributions that depend on the number of failures, r, the sample size, n, and the distribution $\Phi(z)$, but they do not depend on the unknown parameters $(\mu, \sigma)'$. Therefore, for given (n, r) and $\Phi(z)$ one can approximate the distribution of CR_E by using simulation. With γ_E equal to the $100(1-\alpha)\%$ quantile of the distribution of CR_E, (33) provides an exact $100(1-\alpha)\%$ *SCR* for $\boldsymbol{\theta} = (\mu, \sigma)'$.

5.3 The observed information *SCR*

For Type II (failure) censored data, the *SCR* in (28) based on observed information can be written as

$$CR_O = \left\{(\mu, \sigma)' : \breve{\imath}_{11}\, L_E^2 + 2\, \breve{\imath}_{12}\, L_E S_E + \breve{\imath}_{22}\, S_E^2 \leq \gamma_O^s\right\} \tag{34}$$

where $L_E = \sqrt{n}\,[(\widehat{\mu}-\mu)/\widehat{\sigma}]$ and $S_E = \sqrt{n}\,[(\widehat{\sigma}-\sigma)/\widehat{\sigma}]$. The estimates $\breve{\iota}_{ij}$ are defined in (4) and their distributions depends on n, r, and $\Phi(z)$ only and do not depend on the unknown parameters $(\mu, \sigma)'$. For given (n, r) and $\Phi(z)$ one can approximate the distribution of CR_O by using simulation. With γ_O^s equal to the $100(1-\alpha)\%$ quantile of the distribution of CR_O, (34) provides an exact $100(1-\alpha)\%$ SCR for $\boldsymbol{\theta} = (\mu, \sigma)'$.

5.4 Coverage probabilities in the case of uncensored normal distribution data

Here we consider the calibration for the region (7) with complete normal distribution data (the same setting considered in Section 3.2 of Cheng and Iles 1983). For this setting, the ML estimates of the parameters $(\mu, \sigma)'$ are $\widehat{\mu} = \bar{Y} = \sum_{i=1}^{n} Y_i/n$, and $\widehat{\sigma} = \sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2/n}$.

5.4.1 Calibration of the ellipsoidal expected information *SCR*

In this case, the expected information matrix is $I_{\boldsymbol{\theta}} = nM/\sigma^2 = n\text{diag}(1,2)/\sigma^2$. The SCR in (7) can be expressed as $(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})' I_{\boldsymbol{\theta}} (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) = Z^2 + (R - \sqrt{2n})^2 \leq \gamma$ where $Z = \sqrt{n}(\widehat{\mu} - \mu)/\sigma$, and $R \sim \sqrt{2n\widehat{\sigma}^2/\sigma^2}$. Here Z and R are independent, $Z \sim \text{NOR}(0,1)$, and $R \sim \sqrt{2\,\chi^2_{(n-1)}}$. Then $f_R(r) = r f_{\chi^2_{(n-1)}}(r^2/2)$, $r > 0$. It can be shown that the coverage probability of this region is given by

$$\text{CPE}(\gamma, n) = \int_{-\sqrt{\gamma}}^{\sqrt{\gamma}} \left(w + \sqrt{2n}\right) f_{\chi^2_{(n-1)}} \left[\frac{(w + \sqrt{2n})^2}{2}\right] F_{\chi^2_1}\left(\gamma - w^2\right) dw\,. \tag{35}$$

For a large sample, the coverage probability in (35) is approximately equal to $F_{\chi^2_2}(\gamma)$. A value of γ that provides an SCR with exactly $100(1-\alpha)\%$ confidence is obtained by solving $\text{CPE}(\gamma, n) = 1 - \alpha$ for γ.

5.4.2 Estimated expected and observed information *SCR* calibrations

For uncensored data from a normal distribution, the SCR based on estimated expected information and observed information are the same and given by $(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})' \widehat{I}_{\boldsymbol{\theta}} (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) = W^2 +$

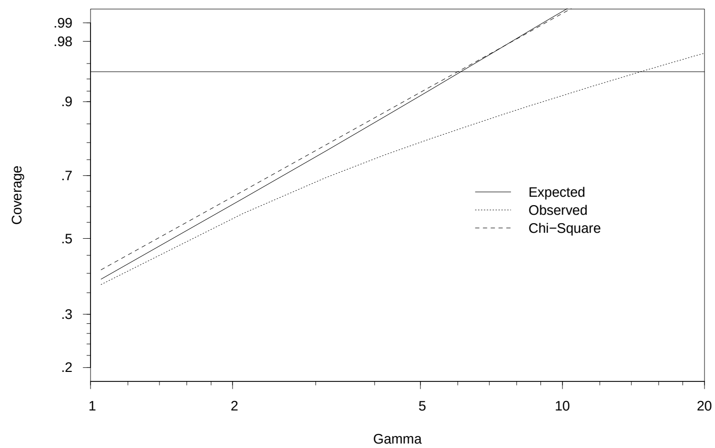

Figure 11: Weibull probability plot of the coverage probability for the ellipsoidal expected information and the observed information based *SCR*s as a function of γ for samples of size $n = 10$ from a normal distribution.

$(V - \sqrt{2n})^2 \le \gamma$ where $W = ZV/\sqrt{2n}$, $Z = \sqrt{n}(\widehat{\mu} - \mu)/\sigma$, and $V = \sqrt{2n\sigma^2/\widehat{\sigma}^2}$. Here Z and V are independent, $Z \sim \text{NOR}(0,1)$, $V \sim n\sqrt{2/\chi^2_{(n-1)}}$, and $W|(V = v) \sim \text{NOR}(0, v^2/2n)$. Then $f_V(v) = (4n^2/v^3)\, f_{\chi^2_{(n-1)}}(2n^2/v^2)$, $v > 0$. The coverage probability of the *SCR* that includes only positive values of σ is

$$\text{CPL}(\gamma, n) = \int_{-\sqrt{\gamma}}^{\sqrt{\gamma}} \left[\frac{4n^2}{(w+\sqrt{2n})^3}\right] f_{\chi^2_{(n-1)}}\left[\frac{2n^2}{(w+\sqrt{2n})^2}\right] F_{\chi^2_1}\left[\frac{2n(\gamma - w^2)}{(w+\sqrt{2n})^2}\right] dw\,. \qquad (36)$$

For a large sample, the coverage probability in (36) is approximately equal to $F_{\chi^2_2}(\gamma)$. A value of γ that provides a $100(1-\alpha)\%$ *SCR* is obtained by solving $\text{CPL}(\gamma, n) = 1 - \alpha$ for γ.

Figure 11 gives a plot of the simultaneous coverage probability as a function of γ for samples from a normal distribution of size $n = 10$ for ellipsoidal expected information and observed information based *SCR*s.

6 Generalization to Regression Problems

The procedures given in Sections 2 through 4 can be extended directly to regression problems, under the usual model that assumes fixed explanatory variables and independent observations. The only difference is the dimensionality of the matrices and vectors involved in the computations. We only obtain the *SCB*s for the local information matrix, but similar results could be derived for the expected information and estimated expected information matrix.

6.1 Simultaneous inference for quantiles of regression model using observed information

The generalization to regression problems with an *SCR* for $\boldsymbol{\theta}$ based on observed information is as follows. For a given vector of the explanatory variables $\boldsymbol{x}$, the p quantile of $F(y;\mu(\boldsymbol{x}),\sigma)$ is $y_p = \mu(\boldsymbol{x}) + z_p\sigma = \boldsymbol{x}'\boldsymbol{\beta} + z_p\sigma$ where $\boldsymbol{x} = (x_1,\ldots,x_k)'$, $\boldsymbol{\beta} = (\beta_1,\ldots,\beta_k)'$, and $\boldsymbol{\theta} = (\boldsymbol{\beta}',\sigma)'$. Suppose that we have a "Wald" observed information *SCR* for $\boldsymbol{\theta}$ of the form

$$(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})'\breve{I}_{\boldsymbol{\theta}}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \leq \gamma \tag{37}$$

where $\breve{I}_{\boldsymbol{\theta}}$ is the observed information matrix. We write

$$\breve{I}_{\boldsymbol{\theta}} = \left(\frac{n}{\widehat{\sigma}^2}\right)\begin{bmatrix} \breve{\Delta}_{11} & \breve{\Delta}_{12} \\ \breve{\Delta}_{12} & \breve{\Delta}_{22} \end{bmatrix} \quad \text{and} \quad \breve{\Lambda} = \left(\frac{n}{\widehat{\sigma}^2}\right)\breve{I}_{\boldsymbol{\theta}}^{-1} = \begin{bmatrix} \breve{\Lambda}_{11} & \breve{\Lambda}_{12} \\ \breve{\Lambda}_{21} & \breve{\Lambda}_{22} \end{bmatrix}.$$

Here $\breve{\Delta}_{11}$ and $\breve{\Lambda}_{11}$ are $k \times k$ matrices and $\breve{\Delta}_{22}$ and $\breve{\Lambda}_{22}$ are scalars. To ensure that the Wald region does not include negative values of σ, one must choose γ such that $\breve{I}_{\boldsymbol{\theta}}$ in (37) satisfies the restriction $\breve{\Delta}_{22} - \breve{\Delta}_{21}\breve{\Delta}_{11}^{-1}\breve{\Delta}_{12} > \gamma_F^s$ where $\gamma_F^s = \gamma/n$, or equivalently that $\left[1 - \gamma_F^s\breve{\Lambda}_{22}\right] > 0$.

Theorem 8 *An approximate* $100(1-\alpha)\%$ *SCB for the quantiles* $y_p = \mu(\boldsymbol{x}) + z_p\sigma$ *at* $\boldsymbol{x}$ *(*$0 < p < 1$*) is given by*

$$[\underset{\sim}{y_p}, \quad \widetilde{y}_p] = \widehat{\mu} + \widehat{\sigma}\left[z_p \mp h_4(\breve{\Lambda}, \boldsymbol{x}, p)\right] \tag{38}$$

where $h_4(\breve{\Lambda}, \boldsymbol{x}, p) = \sqrt{\gamma_F^s \left(\boldsymbol{x}'\breve{\Lambda}_{11}\boldsymbol{x} + 2z_p\boldsymbol{x}'\breve{\Lambda}_{12} + z_p^2\breve{\Lambda}_{22}\right)}$ *and* $\widehat{\mu} = \boldsymbol{x}'\widehat{\boldsymbol{\beta}}$.

6.2 Simultaneous inference for cumulative probabilities of regression model using observed information

For given values of the explanatory variables $\boldsymbol{x}$, the cumulative probability at log time y_e is $p = \Pr(y \le y_e) = F[y_e; \mu(\boldsymbol{x}), \sigma] = \Phi\left[(y_e - \mu(\boldsymbol{x}))/\sigma\right]$.

Theorem 9 *An approximate* $100(1-\alpha)\%$ *SCB for the probabilities* $p = F(y_e; \mu(\boldsymbol{x}), \sigma)$ $(-\infty < y_e < \infty)$ *based on observed information is*

$$[\underset{\sim}{p}, \quad \widetilde{p}] = [\Phi(\underset{\sim}{a}), \quad \Phi(\widetilde{a})] = \Phi\left[z_{\widehat{p}} + h_5(\breve{\Lambda}, \boldsymbol{x}, \widehat{p}) \mp h_6(\breve{\Lambda}, \boldsymbol{x}, \widehat{p})\right] \tag{39}$$

where $h_5(\breve{\Lambda}, \boldsymbol{x}, \widehat{p})$ *and* $h_6(\breve{\Lambda}, \boldsymbol{x}, \widehat{p})$ *are given by* $h_5(\breve{\Lambda}, \boldsymbol{x}, \widehat{p}) = \gamma_F^s(\boldsymbol{x}'\breve{\Lambda}_{12} + z_{\widehat{p}}\breve{\Lambda}_{22})/(1 - \gamma_F^s\breve{\Lambda}_{22})$ *and* $h_6(\breve{\Lambda}, \boldsymbol{x}, \widehat{p}) = \left(\sqrt{\gamma_F^s(\boldsymbol{x}'\breve{\Lambda}_{11}\boldsymbol{x} + 2z_{\widehat{p}}\boldsymbol{x}'\breve{\Lambda}_{12} + z_{\widehat{p}}^2\breve{\Lambda}_{22}) - (\gamma_F^s)^2\boldsymbol{x}'(\breve{\Lambda}_{11}\breve{\Lambda}_{22} - \breve{\Lambda}_{12}\breve{\Lambda}_{21})\boldsymbol{x}}\right)/(1 - \gamma_F^s\breve{\Lambda}_{22})$ *and* $z_{\widehat{p}} = [y_e - \widehat{\mu}(\boldsymbol{x})]/\widehat{\sigma} = \Phi^{-1}(\widehat{p})$.

This *SCB* is obtained by inverting the *SCB* for the quantiles in (38). **Theorems 8** and **9** can be proved in a straightforward manner by using **Lemma 4** in Appendix A.2.

7 Concluding Remarks and Areas for Further Research

Here we provide some general comments about the results developed in this article. The boundaries of the *SCB* based on expected information are monotone increasing, as indicated in the paper. Under the required conditions (25) and (29), the boundaries of the *SCB*s based on estimated information and local information are also monotone increasing. For more information about the coverage probability of these *SCB*s, see Jeng and Meeker (2001), who used extensive simulations to compare the coverage probability for different *SCB*s.

The results in this article are based on Type I and Type II censoring. These results can be directly applied to other kinds of censoring, such as, staggered entry censoring, as long as the information matrix can be expressed as in (2). For more complicated censoring, similar methods could be developed without too much additional work when the distribution assumption is log-location-scale. For distributions not from this family, it may be necessary to use numerical methods to find the *SCB* from a given *SCR*.

When the inference of interest is a general function of $\boldsymbol{\theta}$, (e.g. the hazard function), Cheng and Iles' idea still can be applied. The procedure to get the *SCB*s on the function may, however, have to be done numerically.

A Technical Details

A.1 Maximization and minimization over quadratic regions

In this section, we present two results for maximization and minimization over quadratic regions. **Lemma 2** gives a general result in maximization and minimization over a spherical region. The dimensionality of the spherical region is arbitrary. **Lemma 3** gives results in maximization and minimization over two dimensional parabolic and hyperbolic regions.

Lemma 2 *(A general result in maximization and minimization over a spherical region) Consider the maximization problem:* $\max_{\boldsymbol{\alpha}} \boldsymbol{d}' \boldsymbol{\alpha}$ *subject to:* $\boldsymbol{\alpha}'\boldsymbol{\alpha} \leq k,\ k > 0$. *The maximum is obtained at* $\widetilde{\boldsymbol{\alpha}} = \boldsymbol{d}\sqrt{k/(\boldsymbol{d}'\boldsymbol{d})}$ *and* $\max_{\boldsymbol{\alpha}} \boldsymbol{d}' \boldsymbol{\alpha} = \sqrt{k\,(\boldsymbol{d}'\boldsymbol{d})}$.

A simple consequence of this result is that the constrained $\max(\boldsymbol{d}'\boldsymbol{\alpha})$ occurs at $\widetilde{\boldsymbol{\alpha}}$ and the constrained $\min(\boldsymbol{d}'\boldsymbol{\alpha})$ occurs at $-\widetilde{\boldsymbol{\alpha}}$. This result is illustrated in Figure 12.

Proof: Using the Schwartz inequality, $\boldsymbol{d}'\boldsymbol{\alpha} \leq \sqrt{(\boldsymbol{\alpha}'\boldsymbol{\alpha})(\boldsymbol{d}'\boldsymbol{d})} \leq \sqrt{k(\boldsymbol{d}'\boldsymbol{d})}$. The upper bound is obtained (i.e., the left hand side is equal to the right hand side) when $\boldsymbol{\alpha} = \tilde{\boldsymbol{\alpha}}$ with $\tilde{\boldsymbol{\alpha}} = a\,\boldsymbol{d}$ and $\tilde{\boldsymbol{\alpha}}'\tilde{\boldsymbol{\alpha}} = k$. Then $a = \sqrt{k/(\boldsymbol{d}'\boldsymbol{d})}$. Thus the optimal solution and the maximum are $\tilde{\boldsymbol{\alpha}} = \boldsymbol{d}\,\sqrt{k/(\boldsymbol{d}'\boldsymbol{d})}$ and $\max_{\boldsymbol{\alpha}} \boldsymbol{d}'\,\boldsymbol{\alpha} = \sqrt{k\,(\boldsymbol{d}'\boldsymbol{d})}$.□

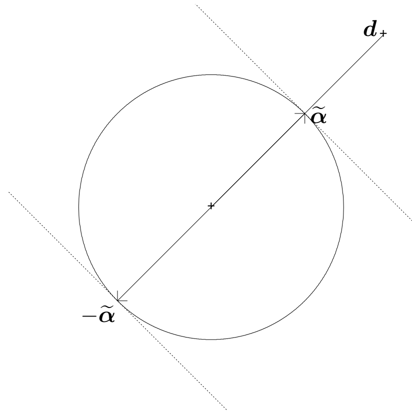

Figure 12: Minimization on a circle.

Lemma 3 *1. (parabolic region) Consider* $\min_{\boldsymbol{\alpha}} t = \boldsymbol{d}'\boldsymbol{\alpha}$ *and* $\max_{\boldsymbol{\alpha}} t = \boldsymbol{d}'\boldsymbol{\alpha}$ *subject to:* $\alpha_2 \geq k\alpha_1^2,\ k > 0,$ *where* $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)'$ *and* $\boldsymbol{d} = (d_1, d_2)'$*. Then*

$$\left[\min_{\boldsymbol{\alpha}} \boldsymbol{d}'\boldsymbol{\alpha}\,,\ \max_{\boldsymbol{\alpha}} \boldsymbol{d}'\boldsymbol{\alpha}\right] = \begin{cases} \left[-\infty,\ -\dfrac{d_1^2}{4kd_2}\right], & d_2 < 0, \\ [-\infty,\ \infty\,], & d_2 = 0, \\ \left[-\dfrac{d_1^2}{4kd_2},\ \infty\right], & d_2 > 0. \end{cases}$$

2. (hyperbolic region) Consider $\min_{\boldsymbol{\alpha}} t = \boldsymbol{d}'\boldsymbol{\alpha}$ *and* $\max_{\boldsymbol{\alpha}} t = \boldsymbol{d}'\boldsymbol{\alpha}$ *subject to:* $\alpha_2^2 - \alpha_1^2 \geq k,\ k > 0,$ *where* $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)', \alpha_2 > 0$ *and* $\boldsymbol{d} = (d_1, d_2)'$*. Then*

$$\left[\min_{\boldsymbol{\alpha}} \boldsymbol{d}'\boldsymbol{\alpha}\,,\ \max_{\boldsymbol{\alpha}} \boldsymbol{d}'\boldsymbol{\alpha}\right] = \begin{cases} \left[-\infty,\ -\sqrt{k(d_2^2 - d_1^2)}\ \right], & d_2 < 0, \textit{and } |d_1| < |d_2|, \\ \left[\ \sqrt{k(d_2^2 - d_1^2)},\ \infty\ \right], & d_2 > 0, \textit{and } |d_1| < |d_2| \\ [-\infty,\ \infty\,], & \textit{otherwise}. \end{cases}$$

$d_2 < 0,$ *and* $|d_1| < |d_2|$ *is equivalent to* $d_2 < 0,$ *and* $d_2^2 - d_1^2 > 0$*.* $d_2 > 0,$ *and* $|d_1| < |d_2|$ *is equivalent to* $d_2 > 0,$ *and* $d_2^2 - d_1^2 > 0$*.*

Proof: First we prove part *(1)*. When $d_2 = 0$, $t = d_1\alpha_1$ then $\max t = \infty$ and $\min t = -\infty$ because $-\infty < \alpha_1 < \infty$. Now consider $d_2 > 0$ (the proof for $d_2 < 0$ is similar).

Note that $\max_{\boldsymbol{\alpha}} t \geq \max_{\alpha_1} d_1\alpha_1 + d_2\alpha_1^2 \to \infty$, as $\alpha_1 \to \infty$. Hence $\max_{\boldsymbol{\alpha}} t = \infty$. Also note that $t = d_1\alpha_1 + d_2\alpha_2 \geq d_1\alpha_1 + kd_2\alpha_1^2 = kd_2[\alpha_1 - d_1/(2d_2)]^2 - d_1^2/(4kd_2)$. Hence $\min_{\boldsymbol{\alpha}} t = -d_1^2/(4kd_2)$, proving part *(1)*.

Now we prove part *(2)*. The proof for $d_2 = 0$ is like the proof of part *(1)* and $d_2 = 0$. Next consider the case of $d_2 > 0, |d_1| < |d_2|$ (the proof for $d_2 < 0, |d_1| < |d_2|$ is similar). Direct computations show that $d_2 > 0, |d_1| < d_2$ is equivalent to $d_2 > 0, d_2^2 - d_1^2 > 0$. Because $t = d_1\alpha_1 + d_2\alpha_2 = (-d_1)(-\alpha_1) + d_2\alpha_2$, it suffices to consider the case of $d_1 > 0, d_2 > 0, d_1 < d_2$ case. Note that $\max_{\boldsymbol{\alpha}} t \geq d_1\alpha_1 + d_2\sqrt{\alpha_1^2 + k}) \to \infty$, as $\alpha_1 \to \infty$. Hence $\max_{\boldsymbol{\alpha}} t = \infty$. Also note that $t = d_1\alpha_1 + d_2\alpha_2 \geq d_1\alpha_1 + d_2\sqrt{\alpha_1^2 + k}$. Taking first and second derivatives with respect to α_1, it can be shown that the minimum point of $d_1\alpha_1 + d_2\sqrt{\alpha_1^2 + k}$ occurs at $(\alpha_1, \alpha_2) = \left[-\sqrt{kd_1^2/(d_2^2 - d_1^2)}, \sqrt{kd_2^2/(d_2^2 - d_1^2)}\right]$. Direct computations show that $\min_{\boldsymbol{\alpha}} t = \sqrt{k(d_2^2 - d_1^2)}$. □

A.2 A simultaneous confidence band based on a Wald type confidence region

Here, we present a general result for an *SCB* obtained from a Wald type *SCR*. This is **Lemma 4**, which is needed to prove **Theorems 2-7**.

Lemma 4 *Suppose that* $\widehat{\boldsymbol{\theta}} = (\widehat{\boldsymbol{\beta}}', \widehat{\sigma})'$ *are the ML estimators of the unknown parameters* $\boldsymbol{\theta} = (\boldsymbol{\beta}', \sigma)' = (\theta_1, \cdots, \theta_k)'$ *in a location-scale model with cdf* $F(y; \mu_i, \sigma) = \Phi[(y - \mu_i)/\sigma]$, $i = 1, \ldots, n$, *where* $\mu_i = \boldsymbol{x}_i'\boldsymbol{\beta}$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_{k-1})'$, *the* $\boldsymbol{x}_i$*'s are vectors of known explanatory variables, and* $\Phi(z)$ *is a continuous cdf that does not depend on unknown parameters. Consider an approximate* $100(1-\alpha)\%$ *SCR for* $\boldsymbol{\theta}$ *of the form*

$$(\widehat{\boldsymbol{\tau}} - \boldsymbol{\theta})'\,\Omega\,(\widehat{\boldsymbol{\tau}} - \boldsymbol{\theta}) \leq \vartheta \tag{40}$$

where the k *dimensional vector* $\widehat{\boldsymbol{\tau}} = (\widehat{\tau}_1, \ldots, \widehat{\tau}_k)'$*, the scalar* ϑ*, and the completely specified* $k \times k$ *positive definite matrix* Ω *may depend on* $\widehat{\boldsymbol{\theta}}$ *and* α *but they do not depend on* $\boldsymbol{\theta}$*. Then*

1. *The SCR defined by (40) contains only positive values of* θ_k *if and only if* $\widehat{\tau}_k > 0$ *and* $\widehat{\tau}_k^2 \det(\Omega) > \vartheta \det(\Omega_{11})$*, or equivalently if* $\widehat{\tau}_k > 0$ *and* $(1 - c\Upsilon_{22}) > 0$*, where*

$$\Upsilon = \begin{bmatrix} \Upsilon_{11} & \Upsilon_{12} \\ \Upsilon_{12} & \Upsilon_{22} \end{bmatrix} = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{12} & \Omega_{22} \end{bmatrix}^{-1},$$

$c = \vartheta/\widehat{\tau}_k^2$*, and* Ω_{11} *and* Υ_{11} *are* $(k-1) \times (k-1)$ *matrices.*

2. *For a fixed vector of explanatory variables, say* $\boldsymbol{x}$*, an approximate* $100(1-\alpha)\%$ *SCB for the quantile* p *at* $\boldsymbol{x}$*,* $y_p = \boldsymbol{x}'\boldsymbol{\beta} + \sigma z_p$*, is given by* $\underset{\sim}{y}_p = \min(\boldsymbol{x}'\boldsymbol{\beta} + \sigma z_p)$ *and* $\widetilde{y}_p = \max(\boldsymbol{x}'\boldsymbol{\beta} + \sigma z_p)$ *where the* min *and* max *are taken over the constrained region (40). Then*

$$[\underset{\sim}{y}_p, \quad \widetilde{y}_p] = \boldsymbol{x}'\widehat{\boldsymbol{\tau}}_1 + \widehat{\tau}_k \left[z_p \mp \sqrt{c \left(\boldsymbol{x}'\Upsilon_{11}\boldsymbol{x} + 2z_p\boldsymbol{x}'\Upsilon_{12} + z_p^2\Upsilon_{22} \right)} \right] \tag{41}$$

where $\widehat{\boldsymbol{\tau}}_1$ *are the first* $(k-1)$ *components of* $\widehat{\boldsymbol{\tau}}$*, i.e.,* $\widehat{\boldsymbol{\tau}}_1 = (\widehat{\tau}_1, \ldots, \widehat{\tau}_{k-1})'$*.*

3. *For a fixed vector of explanatory variables, say* $\boldsymbol{x}$*, and* log *time* y_e*, an approximate* $100(1-\alpha)\%$ *SCB for the cumulative probability* $p = \Phi[(y_e - \mu)/\sigma]$ *has boundaries* $\underset{\sim}{p} = \Phi(\underset{\sim}{a}) = \min \Phi[(y_e - \mu)/\sigma]$ *and* $\widetilde{p} = \Phi(\widetilde{a}) = \max \Phi[(y_e - \mu)/\sigma]$*, where the* min *and* max *are taken over the constrained region (40). Then*

$$\begin{aligned}[\underset{\sim}{a}, \quad \widetilde{a}] = \widehat{\xi} + \frac{c(\boldsymbol{x}'\Upsilon_{12} + \widehat{\xi}\Upsilon_{22})}{1 - c\Upsilon_{22}} \\ \mp \frac{\sqrt{c\left(\boldsymbol{x}'\Upsilon_{11}\boldsymbol{x} + 2\widehat{\xi}\boldsymbol{x}'\Upsilon_{12} + \widehat{\xi}^2\Upsilon_{22}\right) - c^2\boldsymbol{x}'\left(\Upsilon_{11}\Upsilon_{22} - \Upsilon_{12}\Upsilon_{21}\right)\boldsymbol{x}}}{1 - c\Upsilon_{22}}\end{aligned} \tag{42}$$

where $\widehat{\xi} = (y_e - \boldsymbol{x}'\widehat{\boldsymbol{\tau}}_1)/\widehat{\tau}_k$*.*

Proof: To prove part *(1)*, one finds $\min \theta_k = \min \boldsymbol{\delta}'\boldsymbol{\theta}$ (where $\boldsymbol{\delta} = (\mathbf{0}, 1)'$) for the constrained region $(\widehat{\boldsymbol{\tau}} - \boldsymbol{\theta})'\Omega(\widehat{\boldsymbol{\tau}} - \boldsymbol{\theta}) \le \vartheta$. We use **Lemma 2**. In this case, $\arg\min_{\boldsymbol{\theta}} \boldsymbol{\delta}'\boldsymbol{\theta} = \arg\min_{\boldsymbol{\theta}} \boldsymbol{\delta}'(\boldsymbol{\theta} - \widehat{\boldsymbol{\tau}}) = \arg\min_{\boldsymbol{\theta}} \left\{ \left(\sqrt{\Omega^{-1}}\boldsymbol{\delta}\right)' \left[\sqrt{\Omega}\,(\boldsymbol{\theta} - \widehat{\boldsymbol{\tau}})\right] \right\}$ subject to the constraint

$$(\boldsymbol{\theta} - \widehat{\boldsymbol{\tau}})'\Omega(\boldsymbol{\theta} - \widehat{\boldsymbol{\tau}}) = \left[\sqrt{\Omega}(\boldsymbol{\theta} - \widehat{\boldsymbol{\tau}})\right]' \left[\sqrt{\Omega}(\boldsymbol{\theta} - \widehat{\boldsymbol{\tau}})\right] = \vartheta. \tag{43}$$

Using **Lemma 2** with $\boldsymbol{\alpha} = \sqrt{\Omega}(\boldsymbol{\theta} - \widehat{\boldsymbol{\tau}})$, $\boldsymbol{d} = \sqrt{\Omega^{-1}}\,\boldsymbol{\delta}'$, $k = \vartheta$, it follows that the minimum occurs at $-\boldsymbol{d}\sqrt{\vartheta/(\boldsymbol{d}'\boldsymbol{d})}$. Then $\sqrt{\Omega^{-1}}(\widetilde{\boldsymbol{\theta}} - \widehat{\boldsymbol{\tau}}) = -\left(\sqrt{\vartheta}/\sqrt{\boldsymbol{\delta}'\Omega^{-1}\boldsymbol{\delta}}\right)\sqrt{\Omega^{-1}}\,\boldsymbol{\delta}$ or equivalently $\widetilde{\boldsymbol{\theta}} = \widehat{\boldsymbol{\tau}} - \left(\sqrt{\vartheta}/\sqrt{\boldsymbol{\delta}'\Omega^{-1}\boldsymbol{\delta}}\right)\Omega^{-1}\boldsymbol{\delta}$. Therefore $\min\theta_k = \boldsymbol{\delta}'\widetilde{\boldsymbol{\theta}} = \widehat{\tau}_k - \sqrt{\vartheta\,\boldsymbol{\delta}'\Omega^{-1}\boldsymbol{\delta}} = \widehat{\tau}_k - \sqrt{\vartheta\,\Upsilon_{22}}$. Thus $\min\theta_k > 0$ if and only if $\widehat{\tau}_k > 0, \widehat{\tau}_k - \sqrt{\vartheta\,\Upsilon_{22}} > 0$, which implies that $(1 - c\Upsilon_{22}) > 0$. Because $\det(\Omega) = \det(\Omega_{11})\,\det(\Omega_{22} - \Omega_{21}\Omega_{11}^{-1}\Omega_{21}) = \det(\Omega_{11})/\det(\Upsilon_{22})$, the *SCR* includes only positive values of θ_k if and only if $\widehat{\boldsymbol{\tau}}_k > 0$ and $\widehat{\tau}_k^2\det(\Omega) > \vartheta\det(\Omega_{11})$.

Without lost of generality, to prove parts *(2)* and *(3)*, we assume that $\boldsymbol{x}$ is the scalar 1 and we write $\boldsymbol{x}'\boldsymbol{\beta} = \mu, \widehat{\boldsymbol{\tau}} = (\widehat{\boldsymbol{\tau}}_1, \widehat{\boldsymbol{\tau}}_k)'$ (i.e., $k = 2$). Similar to the proof of part *(1)*, finding the boundaries of *SCB* for quantiles consists of finding $\arg\max_{\boldsymbol{\theta}} \left\{\left(\sqrt{\Omega^{-1}}\boldsymbol{c}\right)'\left[\sqrt{\Omega}\,(\boldsymbol{\theta} - \widehat{\boldsymbol{\tau}})\right]\right\}$, subject to the constraint (43), with $\boldsymbol{c} = (\mu, z_p)'$. Using **Lemma 2** with $\boldsymbol{\alpha} = \sqrt{\Omega}(\boldsymbol{\theta} - \widehat{\boldsymbol{\tau}})$, $\boldsymbol{d} = \sqrt{\Omega^{-1}}\boldsymbol{c}$, $k = \vartheta$, we see that the maximum occurs at $\widetilde{\boldsymbol{\alpha}} = \sqrt{\Omega}(\boldsymbol{\theta} - \widehat{\boldsymbol{\tau}}) = \left(\sqrt{\vartheta}/\sqrt{\boldsymbol{c}'\Omega^{-1}\boldsymbol{c}}\right)\sqrt{\Omega^{-1}}\,\boldsymbol{c}$ or $\boldsymbol{\theta} = \widehat{\boldsymbol{\tau}} + \left(\sqrt{\vartheta}/\sqrt{\boldsymbol{c}'\Omega^{-1}\boldsymbol{c}}\right)\Omega^{-1}\boldsymbol{c}$. Therefore $\max_{\boldsymbol{\theta}}\boldsymbol{c}'\boldsymbol{\theta} = \boldsymbol{c}'\widehat{\boldsymbol{\tau}} + \sqrt{\vartheta\;\boldsymbol{c}'\Omega^{-1}\boldsymbol{c}}$ and $\min_{\boldsymbol{\theta}}\boldsymbol{c}'\boldsymbol{\theta} = \boldsymbol{c}'\widehat{\boldsymbol{\tau}} - \sqrt{\vartheta\;\boldsymbol{c}'\Omega^{-1}\boldsymbol{c}}$. Straightforward substitutions yield, $\boldsymbol{c}'\widehat{\boldsymbol{\tau}} \mp \sqrt{\vartheta\;\boldsymbol{c}'\Omega^{-1}\boldsymbol{c}} = \widehat{\tau}_1 + \widehat{\tau}_k\left[z_p \mp \sqrt{c\left(\Upsilon_{11} + 2z_p\Upsilon_{12} + z_p^2\Upsilon_{22}\right)}\right]$.

To prove part *(3)*, let $\widetilde{p} = \max\Phi[(y_e - \mu)/\sigma]$ where the maximization is over the constrained region (43). This implies $y_e \le \mu + \sigma z_{\widetilde{p}}$, which shows that y_e is the lower boundary of the *SCB* for the $\widetilde{p}$ quantile (i.e., $y_e = y_{\underset{\sim}{\widetilde{p}}}$). Then solving this equation for $\widetilde{p}$ provides the upper bound for p. Letting $\widetilde{a} = \Phi^{-1}(\widetilde{p})$, one needs to solve for $\widetilde{a}$ in the equation $y_e = \widehat{\tau}_1 + \widehat{\tau}_k\left[\widetilde{a} - \sqrt{c\,(\Upsilon_{11} + 2\widetilde{a}\Upsilon_{12} + \widetilde{a}^2\Upsilon_{22})}\right]$. Similarly, the lower bound $\underset{\sim}{a}$ is the solution to the equation, $y_e = \widehat{\tau}_1 + \widehat{\tau}_k\left[\underset{\sim}{a} + \sqrt{c(\Upsilon_{11} + 2\underset{\sim}{a}\Upsilon_{12} + \underset{\sim}{a}^2\Upsilon_{22})}\right]$. Then $\underset{\sim}{a}$ and $\widetilde{a}$ are the roots of the quadratic equation $(\widehat{\xi} - a)^2 - c\,(\Upsilon_{11} + 2a\Upsilon_{12} + a^2\Upsilon_{22}) = 0$ where $\widehat{\xi} = (y_e - \widehat{\tau}_1)/\widehat{\tau}_k$. Straightforward but lengthly computations give $[\underset{\sim}{a},\quad \widetilde{a}] = \widehat{\xi} + c(\Upsilon_{12} + \widehat{\xi}\,\Upsilon_{22})/(1 - c\Upsilon_{22}) \mp \sqrt{c(\Upsilon_{11} + 2\widehat{\xi}\Upsilon_{12} + \widehat{\xi}^2\Upsilon_{22}) - c^2(\Upsilon_{11}\Upsilon_{22} - \Upsilon_{12}^2)}/(1 - c\Upsilon_{22})$.□

A.3 Proofs of Theorems 2 and 3

Proof: [**Theorem 2**] To prove part *(1)*, use **Lemma 1** to get the representation (40) for the *SCR*. In this case $\boldsymbol{x} = 1$ and $\boldsymbol{\beta} = \beta_0$ are both scalars and we write $\boldsymbol{x}'\boldsymbol{\beta} = \mu$. Use the results in **Lemma 4** with $\Omega = R, \vartheta = \gamma_F^s\widehat{\sigma}^2/(1 - \gamma_F^s\lambda_{22}), \boldsymbol{\tau} = (\widehat{\boldsymbol{\tau}}_1, \widehat{\boldsymbol{\tau}}_2)', c = \gamma_F^s/(1 -$

$\gamma_F^s\lambda_{22})$ to obtain $c\Upsilon_{11} = \gamma_F^s\Lambda_{11} - (\gamma_F^s)^2(\Lambda_{11}\Lambda_{22} - \Lambda_{12}^2)\,, c\Upsilon_{22} = \gamma_F^s\Lambda_{22}\,,$ and $c\Upsilon_{12} = \gamma_F^s\Lambda_{12}$. Substituting these values of $\widehat{\boldsymbol{\tau}}, \Upsilon$, and c into (41), using $\Lambda = \lambda_{ij}$ (because they are scalars), and after some simplification, one gets $[\underset{\sim}{y}_p, \quad \widetilde{y}_p] = \widehat{\mu} + \widehat{\sigma}\,[z_p + h_1(\Lambda, p) \mp h_2(\Lambda, p)]\,.$

To prove part *(2)*, we use the parabolic expression of the *SCR* in (12). Thus,

$$y_p = (1, z_p)\widehat{\boldsymbol{\theta}} - (1, z_p)O\left(\frac{v_1}{\zeta_1}, -\frac{\gamma_F^s\widehat{\sigma}^2 + v_1^2/\zeta_1}{2v_2}\right) + (1, z_p)O\text{diag}(-1, 1)\boldsymbol{\alpha}. \tag{44}$$

Here, O, ζ_1, v_1, v_2 are defined in (11). To use **Lemma 3**, let $\boldsymbol{d'} = (1, z_p)O\text{diag}(-1, 1), k = -\zeta_1/(2v_2)$. Hence $d_1 = (-f_{11} - z_p f_{12})/\sqrt{f_{11}\zeta_1}, d_2 = (-f_{12} + z_p f_{11})/\sqrt{f_{11}\zeta_1}$. $d_2 > 0$ implies that $z_p > f_{12}/f_{11} = -\lambda_{12}/\lambda_{22}$, which means $p > \Phi(-\lambda_{12}/\lambda_{22})$. When $d_2 > 0$, by **Lemma 3**, $[\,\min_{\boldsymbol{\alpha}} \boldsymbol{d'\alpha}\,,\ \max_{\boldsymbol{\alpha}} \boldsymbol{d'\alpha}\,] = [\,-d_1^2/(4kd_2),\ \infty\,]$. Substituting this into (44), after some simplification, one obtains the second case of (19). The proofs for other cases are similar. By L'Hospital's Rule, after some simplification, we can show that the *SCB* in (19) is the band in (16) when $\gamma_F^s \to (1/\lambda_{22})^-$.

To prove part *(3)*, write $R = S'\text{diag}(1, -1)S$ where $S = \text{diag}\left(\sqrt{|\zeta_1|}, \sqrt{|\zeta_2|}\,\right)O'$. The *SCR* in **Lemma 1** can be re-expressed as $(R^{-1}M\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})'S'\text{diag}(-1, 1)S(R^{-1}M\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \geq \widehat{\sigma}^2\,\gamma_F^s/(\gamma_F^s\lambda_{22} - 1)$. Thus,

$$y_p = (1, z_p)R^{-1}M\widehat{\boldsymbol{\theta}} - (1, z_p)S^{-1}\boldsymbol{\alpha}. \tag{45}$$

By (15), $(1, z_p)R^{-1}M\widehat{\boldsymbol{\theta}}$ can be simplified as $\widehat{y}_p + \widehat{\sigma}h_1(\Lambda, p)$. To use **Lemma 3**, let $\boldsymbol{d'} = -(1, z_p)S^{-1}, k = \widehat{\sigma}^2\,\gamma_F^s/(\gamma_F^s\lambda_{22} - 1)$. When $d_2 > 0, |d_1| < |d_2|$, this condition can be simplified (details omitted here) to $p > \Phi\left[\left(-\lambda_{12} + \sqrt{(\gamma_F^s\lambda_{22} - 1)\det(\Lambda)}\right)/\lambda_{22}\right]$. By **Lemma 3**, $[\,\min_{\boldsymbol{\alpha}} \boldsymbol{d'\alpha}\,,\ \max_{\boldsymbol{\alpha}} \boldsymbol{d'\alpha}\,] = \left[\,\sqrt{k(d_2^2 - d_1^2)}\,,\ \infty\,\right]$. Substituting this into (45), after some simplification, one obtains the second case of part *(3)* . The proofs for other cases are similar. □

Proof: [**Theorem 3**] We only prove part *(1)*. The proofs for parts *(2)* and *(3)* are similar but lengthy, and they are omitted here. To obtain the confidence bounds for the cumulative probabilities, notice that $\widehat{\xi} + c(\boldsymbol{x}'\Upsilon_{12} + \widehat{\xi}\Upsilon_{22})/(1 - c\Upsilon_{22}) = (\widehat{\xi} + c\Upsilon_{12})/(1 - c\Upsilon_{22}) = (y_e - \widehat{\mu})/\widehat{\sigma} = z_{\widehat{p}}$. Substituting this into (42) and after some simplification, one get $[\underset{\sim}{p}, \quad \widetilde{p}] = \Phi\left[z_{\widehat{p}} \mp h_3(\Lambda, \widehat{p})\right]$ where $h_1(\cdot,\cdot), h_2(\cdot,\cdot), h_3(\cdot,\cdot)$ were defined in (17), (18), and (22), respectively.□

A.4 Proofs of Theorems 4 and 5

Proof: To prove **Theorem 4**, we use the results in **Lemma 4**. In this case $\boldsymbol{x} = 1$ and $\boldsymbol{\beta} = \beta_0$ are both scalars and we write $\boldsymbol{x}'\boldsymbol{\beta} = \mu$. Also $\Omega = \widehat{\boldsymbol{I}}_{\boldsymbol{\theta}}, \vartheta = \gamma, \boldsymbol{\tau} = (\widehat{\mu}, \widehat{\sigma})', c = \gamma/\widehat{\sigma}^2$. Using these equivalences and $\widehat{\boldsymbol{I}}_{\boldsymbol{\theta}} = (n/\widehat{\sigma}^2)M$, direct computations give $c\Upsilon_{ij} = \gamma_F^s \Lambda_{ij}$.

To prove **Theorem 5**, substituting these values of $\widehat{\boldsymbol{\tau}}, \Upsilon$, and c into (41), using $\Lambda = \lambda_{ij}$ (because they are scalars), and after some simplification, one gets $[\underset{\sim}{y}_p, \quad \widetilde{y}_p] = \widehat{\mu} + \widehat{\sigma}\left[z_p \mp h_3(\Lambda, p)\right]$ and $[\underset{\sim}{p}, \quad \widetilde{p}] = \Phi\left[z_{\widehat{p}} + h_1(\Lambda, \widehat{p}) \mp h_2(\Lambda, \widehat{p})\right]$.□

References

Cheng, R., Iles, T., 1983. Confidence bands for cumulative distribution functions of continuous random variables. Technometrics 25 (1), 77–86.

Cheng, R., Iles, T., 1988. One-sided confidence bands for cumulative distribution functions. Technometrics 20 (2), 155–159.

Jeng, S.-L., Meeker, W. Q., 2001. Parametric simultaneous confidence bands for cumulative distributions from censored data. Technometrics 43 (4), 450–461.

Lawless, J. F., 2003. Statistical Models and Methods for Lifetime Data, 2nd Edition. John Wiley and Sons Inc.

Meeker, W. Q., Escobar, L. A., 1998. Statistical Methods for Reliability Data. John Wiley and Sons Inc.

Miller, R. G., 1981. Simultaneous Statistical Inference, 2nd Edition. Springer.

Sa, P., Lee, S. J., 1998. Simultaneous confidence bands for the pth percentile and the mean lifetime in exponential and weibull regression models. Tech. rep., http://interstat.statjournals.net/YEAR/1998/articles/9804001.pdf.