Adequate Mathematical Models of the Cumulative Distribution Function of Order Statistics to Construct Accurate Tolerance Limits and Confidence Intervals of the Shortest Length or Equal Tails

NICHOLAS NECHVAL[1*], GUNDARS BERZINS[1], KONSTANTIN NECHVAL[2]
[1]BVEF Research Institute, University of Latvia, Riga LV-1586, LATVIA
[2]Transport and Telecommunication Institute, Riga LV-1019, LATVIA
Corresponding Author

Abstract: - The technique used here emphasizes pivotal quantities and ancillary statistics relevant for obtaining tolerance limits (or confidence intervals) for anticipated outcomes of applied stochastic models under parametric uncertainty and is applicable whenever the statistical problem is invariant under a group of transformations that acts transitively on the parameter space. It does not require the construction of any tables and is applicable whether the experimental data are complete or Type II censored. The exact tolerance limits on order statistics associated with sampling from underlying distributions can be found easily and quickly making tables, simulation, Monte-Carlo estimated percentiles, special computer programs, and approximation unnecessary. The proposed technique is based on a probability transformation and pivotal quantity averaging. It is conceptually simple and easy to use. The discussion is restricted to one-sided tolerance limits. Finally, we give practical numerical examples, where the proposed analytical methodology is illustrated in terms of the exponential distribution. Applications to other log-location-scale distributions could follow directly.

Key-Words: - anticipated outcomes, parametric uncertainty, unknown (nuisance) parameters, elimination, pivotal quantities, ancillary statistics, new-sample prediction, within-sample prediction.

Received: July 2, 2022. Revised: January 4, 2023. Accepted: February 2, 2023. Published: March 2, 2023.

1 Introduction

Statistical prediction and optimization (under parametric uncertainty) of future random quantities (future outcomes, order statistics, etc.) based on the past and current data is the most prevalent form of statistical inference. Predictive inferences for future random quantities are widely used in risk management, finance, insurance, economics, hydrology, material sciences, telecommunications, and many other industries. Predictive inferences (predictive distributions, prediction or tolerance limits (or intervals), confidence limits (or intervals) for future random quantities on the basis of the past and present knowledge represent a fundamental problem of statistics, arising in many contexts and producing varied solutions. Statistical prediction is the process by which values for unknown observables (potential observations yet to be made or past ones which are no longer available) are inferred based on current observations and other information at hand. The approach used here is a special case of more general considerations applicable whenever the statistical problem is invariant under a group of transformations, which acts transitively on the parameter space [1-12].

There are the following types of prediction problems:

1.1 New-Sample Prediction Problem

In this case, the data from a past sample of size n are used to make prediction on one or more future units in a second sample of size m from the same process or population. For example, based on previous (possibly censored) life test data, one could be interested in predicting the following: (1) time to failure of a new item (m = 1); (2) time until the kth failure in a future sample of m units, $m \geq k$; (3) number of failures by time $\tau^{\bullet}$ in a future sample of m units. Formally we call the problems in this category as two-sample problems.

1.2 Within-Sample Prediction Problem

In this case, the problem is to predict future events in a sample or process based on the early-failure data from that sample or process. For example, if n units are followed until censoring time τ_c and there are r observed ordered failure times, $X_1 \leq ... \leq X_r$, one could be interested in predicting the following: 1) time of next failure; 2) time until l additional failures, $l \leq n - r$; 3) number of additional failures in a future interval $\left(\tau_c, \tau^{\bullet}\right)$. Formally we call the problems in this category as one-sample problems.

2 Adequate Mathematical Models of the Cumulative Distribution Function of Order Statistics to Construct New-Sample Tolerance Limits (One-Sided)

Theorem 1.Let us assume that there is a random sample of t ordered observations $Z_1 \leq \ldots \leq Z_t$ from a known distribution with a probability density function (pdf) $f_\mu(z)$, cumulative distribution function (cdf) $F_\mu(z)$, where μ is the parameter (in general, vector), then the adequate mathematical models of the cumulative distribution function (cdf) of the rth order statistic Z_r, $r \in \{1, 2, \ldots, t\}$, to construct one-sided γ −content tolerance limits with confidence level β, are given (for a new sample) as follows:

2.1 Adequate Mathematical Model 2.1

$$\int_0^{F_\mu(z_r)} f_{r,t-r+1}(u)du = P_\mu(Z_r \leq z_r \mid t), \quad (1)$$

where

$$f_{r,t-r+1}(u) = \frac{1}{\mathrm{B}(r, t-r+1)} u^{r-1}(1-u)^{(t-r+1)-1},$$

$$0 < u < 1, \quad (2)$$

is the probability density function (pdf) of the beta distribution $(Beta(r, t-r+1))$ with shape parameters r and t-r+1,

$$P_\mu(Z_r \leq z_r \mid t) = \sum_{j=r}^{t} \binom{t}{j} [F_\mu(z_r)]^j \; [1 - F_\mu(z_l)]^{t-j}. \quad (3)$$

Proof. On the one hand, it follows from (1) that

$$\frac{d}{dz_r} \int_0^{F_\mu(z_r)} f_{r,t-r+1}(u)du$$

$$= \frac{d}{dy_r} \int_0^{F_\mu(z_r)} \frac{1}{\mathrm{B}(r, t-r+1)} u^{r-1}(1-u)^{(t-r+1)-1} du$$

$$= \frac{F_\mu(z_r)^{r-1}}{\mathrm{B}(r, t-r+1)} \left(1 - F_\mu(z_r)\right)^{(t-r+1)-1} f_\mu(z_r). \quad (4)$$

On the other hand, it follows from (1) that

$$\frac{d}{dz_r} P_\theta(Z_r \leq z_r \mid t) = \frac{F_\mu(z_r)^{r-1}}{\mathrm{B}(r, t-r+1)}$$

$$\times \left(1 - F_\mu(z_r)\right)^{(t-r+1)-1} f_\mu(z_r). \quad (5)$$

Thus, $F_\mu(z_r)$ is the generalized pivotal quantity:

$$F_\mu(z_r) = u \sim f_{r,t-r+1}(u)$$

$$= \frac{1}{\mathrm{B}(r, t-r+1)} u^{r-1}(1-u)^{(t-r+1)-1}, \quad 0 < u < 1. \quad (6)$$

This ends the proof.

2.2 Adequate Mathematical Model 2.2

$$\int_{1-F_\mu(z_r)}^{1} f_{t-r+1,r}(u)du = P_\mu(Z_r \leq z_r \mid t), \quad (7)$$

where

$$f_{t-r+1,l}(u) = \frac{1}{\mathrm{B}(t-r+1, r)} u^{(t-r+1)-1}(1-u)^{r-1},$$

$$0 < u < 1, \quad (8)$$

is the probability density function (pdf) of the beta distribution $(Beta(t-r+1, r))$ with shape parameters $t-r$+1 and r,

Proof. It follows from (7) that

$$\frac{d}{dz_r} \int_{1-F_\mu(z_r)}^{1} f_{t-r+1,r}(u)du$$

$$= \frac{d}{dz_r} \int_{1-F_\mu(z_r)}^{1} \frac{1}{\mathrm{B}(t-r+1, r)} u^{(t-r+1)-1}(1-u)^{r-1} du$$

$$= \frac{-1}{\mathrm{B}(t-r+1, r)} \left(1 - F_\mu(z_r)\right)^{(t-r+1)-1} F_\mu(z_r)^{r-1}(-f_\mu(z_r))$$

$$= \frac{1}{\mathrm{B}(r, t-r+1)} F_\mu(z_r)^{r-1} \left(1 - F_\mu(z_r)\right)^{t-r} f_\mu(z_r). \quad (9)$$

It follows from (5) and (9) that

$$\frac{d}{dz_r} P_\theta(Z_r \leq z_r \mid t) = \frac{d}{dz_r} \int_{1-F_\mu(z_r)}^{1} f_{t-r+1,r}(u)du. \quad (10)$$

Thus, $1 - F_\mu(z_r)$ is the generalized pivotal quantity:

$$1 - F_\mu(z_r) = u \sim f_{t-r+1,r}(u)$$

$$= \frac{1}{\mathrm{B}(t-l+1, l)} u^{(m-l+1)-1}(1-u)^{l-1}, \quad 0 < u < 1. \quad (11)$$

This ends the proof.

2.3 Adequate Mathematical Model 2.3

$$\int_0^{\frac{t-r+1}{r}\frac{F_\mu(z_r)}{1-F_\mu(z_r)}} \varphi_{r,t-r+1}(u)du = P_\mu(Z_r \le z_r \mid t), \quad (12)$$

where

$$\varphi_{r,t-r+1}(u) = \frac{1}{\mathrm{B}(r,t-r+1)}$$

$$\times \frac{\left[\frac{r}{t-r+1}u\right]^{r-1}}{\left[1+\frac{r}{t-r+1}u\right]^{t+1}} \frac{r}{t-r+1}, \quad u \in (0,\infty). \quad (13)$$

is the probability density function (pdf) of the F distribution $(F(r,t-r+1))$ with parameters r and $t-r+1$, which are positive integers known as the degrees of freedom for the numerator and the degrees of freedom for the denominator.

Proof. It follows from (12) that

$$\frac{d}{dz_r}\int_0^{\frac{t-r+1}{r}\frac{F_\mu(z_r)}{1-F_\mu(z_r)}} \varphi_{r,t-r+1}(u)du = \frac{d}{dz_r}P_\mu(Z_r \le z_r \mid t). \quad (14)$$

Thus, $\frac{t-r+1}{r}\frac{F_\mu(z_r)}{1-F_\mu(z_r)}$ is the generalized pivotal quantity:

$$\frac{t-r+1}{r}\frac{F_\mu(z_r)}{1-F_\mu(z_r)} = u \sim \varphi_{l,m-l+1}(u) = \frac{1}{\mathrm{B}(r,t-r+1)}$$

$$\times \frac{\left[\frac{r}{t-r+1}u\right]^{l-1}}{\left[1+\frac{r}{t-r+1}u\right]^{m+1}} \frac{r}{t-r+1}, \quad u \in (0,\infty). \quad (15)$$

This ends the proof.

2.4 Adequate Mathematical Model 2.4

$$\int_{\frac{r}{t-r+1}\frac{1-F_\mu(z_r)}{F_\mu(z_r)}}^{\infty} \varphi_{t-r+1,r}(u)du = P_\mu(Z_r \le z_r \mid t), \quad (16)$$

where

$$\varphi_{t-r+1,r}(u) = \frac{\frac{t-r+1}{r}}{\mathrm{B}(t-r+1,r)} \frac{\left[\frac{t-r+1}{r}u\right]^{t-r}}{\left[1+\frac{t-r+1}{r}u\right]^{t+r}},$$

$$u \in (0,\infty), \quad (17)$$

is the probability density function (pdf) of the F distribution $(F(t-r+1,r)$ with parameters $t-r+1$ and r, which are positive integers known as the degrees of freedom for the numerator and the degrees of freedom for the denominator.

Proof. It follows from (16) that

$$\frac{d}{dz_r}\int_{\frac{r}{t-r+1}\frac{1-F_\mu(z_r)}{F_\mu(z_r)}}^{\infty} \varphi_{t-r+1,r}(u)du = \frac{d}{dz_r}P_\mu(Z_r \le z_r \mid t). \quad (18)$$

Thus, $\frac{r}{t-r+1}\frac{1-F_\mu(z_r)}{F_\mu(z_r)}$ is the generalized pivotal quantity:

$$\frac{r}{t-r+1}\frac{1-F_\mu(z_r)}{F_\mu(z_r)} = u \sim \varphi_{t-r+1,r}(u) = \frac{1}{\mathrm{B}(t-r+1,r)}$$

$$\times \frac{\left[\frac{t-r+1}{r}u\right]^{t-r}}{\left[1+\frac{t-r+1}{r}u\right]^{t+1}} \frac{t-r+1}{r}, \quad u \in (0,\infty). \quad (19)$$

This ends the proof.

3 Adequate Mathematical Models of the Cumulative Distribution Function (Conditional) of Order Statistics to Construct Within-Sample Tolerance Limits (One-Sided)

Theorem 2. Let us assume that there is a random sample of t ordered observations $Z_1 \le \ldots \le Z_t$ from a known distribution with a probability density function (pdf) $f_\mu(z)$, cumulative distribution function (cdf) $F_\mu(z)$, where μ is the parameter (in general, vector), then the adequate mathematical models of the conditional cumulative distribution function (cdf) of the rth order statistic Z_r ($1 \le k < r \le t$) given $Z_k=z_k$ are determined (for the same sample) as follows:

3.1 Adequate Mathematical Model 3.1

$$\int_0^{1-\frac{\bar F_\mu(z_r)}{\bar F_\mu(z_k)}} f_{r-k,t-r+1}(u)du = P_\mu(Z_r \le z_r \mid Z_k = z_k; t)$$

$$= \sum_{j=r-k}^{t-k}\binom{t-k}{j}\left[1-\frac{\bar F_\mu(z_r)}{\bar F_\mu(z_k)}\right]^j \left[\frac{\bar F_\mu(z_r)}{\bar F_\mu(z_k)}\right]^{t-k-j} \quad (20)$$

where $\bar F_\mu(z) = 1-F_\mu(z)$,

$$f_{r-k,t-r+1}(u)=\frac{u^{r-k-1}(1-u)^{(t-r+1)-1}}{\mathrm{B}(r-k,t-r+1)}du,\quad 0<u<1,\quad (21)$$

is the probability density function (pdf) of the beta distribution ($Beta(r-k,t-r+1)$) with shape parameters $r-k$ and $t-r+1$.

Proof. On the one hand, it follows from (20) that

$$\frac{d}{dz_r}\int_0^{1-\frac{\bar F_\mu(z_r)}{\bar F_\mu(z_k)}} f_{r-k,t-r+1}(u)du$$

$$=\frac{d}{dz_r}\int_0^{1-\frac{\bar F_\mu(z_r)}{\bar F_\mu(z_k)}} \frac{u^{r-k-1}(1-u)^{(t-r+1)-1}}{\mathrm{B}(r-k,t-r+1)}du$$

$$=\frac{\left[1-\frac{\bar F_\mu(z_r)}{\bar F_\mu(z_k)}\right]^{r-k-1}\left[\frac{\bar F_\mu(z_r)}{\bar F_\mu(z_k)}\right]^{t-r}\frac{f_\mu(z_r)}{\bar F_\mu(z_k)}}{\mathrm{B}(r-k,t-r+1)}.\quad (22)$$

On the other hand, it follows from (20) that

$$\frac{d}{dz_r}P_\mu(Z_r\le z_r\mid Z_k=z_k;t)$$

$$=\frac{\left[1-\frac{\bar F_\mu(z_r)}{\bar F_\mu(z_k)}\right]^{r-k-1}\left[\frac{\bar F_\mu(z_r)}{\bar F_\mu(z_k)}\right]^{t-r}\frac{f_\mu(z_r)}{\bar F_\mu(z_k)}}{\mathrm{B}(r-k,t-r+1)}.\quad (23)$$

Thus, $1-\bar F_\mu(z_r)/\bar F_\mu(z_k)$ is the generalized pivotal quantity:

$$1-\frac{\bar F_\mu(z_r)}{\bar F_\mu(z_k)}=u\sim f_{r-k,t-r+1}(u)=\frac{u^{r-k-1}(1-u)^{(t-r+1)-1}}{\mathrm{B}(r-k,t-r+1)},$$

$$0<u<1.\quad (24)$$

This ends the proof.

3.2 Adequate Mathematical Model 3.2

$$\int_{\frac{\bar F_\mu(z_r)}{\bar F_\mu(z_k)}}^{1} f_{t-r+1,r-k}(u)du,=P_\mu(Z_r\le z_r\mid Z_k=z_k;t)$$

$$=\sum_{j=r-k}^{t-k}\binom{t-k}{j}\left[1-\frac{\bar F_\mu(z_r)}{\bar F_\mu(z_k)}\right]^j\left[\frac{\bar F_\mu(z_r)}{\bar F_\mu(z_k)}\right]^{t-k-j}\quad (25)$$

where

$$f_{m-l+1,l-k}(u)=\frac{u^{(m-l+1)-1}(1-u)^{l-k-1}}{\mathrm{B}(m-l+1,l-k)},\quad 0<u<1,\ (26)$$

is the probability density function (pdf) of the beta distribution ($Beta(t-r+1,r-k)$) with shape parameters $t-r+1$ and $r-k$.

Proof. On the one hand, it follows from (25) that

$$=\frac{d}{dz_r}\int_{\frac{\bar F_\mu(z_r)}{\bar F_\mu(z_k)}}^{1} f_{t-r+1,r-k}(u)du$$

$$=\frac{d}{dz_r}\int_{\frac{\bar F_\mu(z_r)}{\bar F_\mu(z_k)}}^{1}\frac{u^{(t-r+1)-1}(1-u)^{r-k-1}}{\mathrm{B}(t-r+1,r-k)}du$$

$$=\frac{\left[1-\frac{\bar F_\mu(z_r)}{\bar F_\mu(z_k)}\right]^{r-k-1}\left[\frac{\bar F_\mu(z_r)}{\bar F_\mu(z_k)}\right]^{(t-r+1)-1}\frac{f_\mu(z_r)}{\bar F_\mu(z_k)}}{\mathrm{B}(r-k,t-r+1)}.\quad (27)$$

On the other hand, it follows from (25) that

$$\frac{d}{dz_r}P_\mu(Z_r\le z_r\mid Z_k=z_k;t)$$

$$=\frac{\left[1-\frac{\bar F_\mu(z_r)}{\bar F_\mu(z_k)}\right]^{r-k-1}\left[\frac{\bar F_\mu(z_r)}{\bar F_\mu(z_k)}\right]^{t-r}\frac{f_\mu(z_r)}{\bar F_\mu(z_k)}}{\mathrm{B}(r-k,t-r+1)}.\quad (28)$$

Thus, $\bar F_\mu(z_r)/\bar F_\mu(z_k)$ is the generalized pivotal quantity:

$$\frac{\bar F_\mu(z_r)}{\bar F_\mu(z_k)}=u\sim f_{t-r+1,r-k}(u)=\frac{u^{(t-r+1)-1}(1-u)^{r-k-1}}{\mathrm{B}(t-r+1,r-k)},$$

$$0<u<1.\quad (29)$$

This ends the proof.

3.3 Adequate Mathematical Model 3.3

$$\int_0^{\frac{t-r+1}{r-k}\left(1-\frac{\bar F_\mu(z_r)}{\bar F_\mu(z_k)}\right)\Big/\frac{\bar F_\mu(z_r)}{\bar F_\mu(z_k)}}\varphi_{r-k,t-r+1}(u)du$$

$$=P_\mu(Z_r\le z_r\mid Z_k=z_k;t)$$

$$=\sum_{j=r-k}^{t-k}\binom{t-k}{j}\left[1-\frac{\bar F_\mu(z_r)}{\bar F_\mu(z_k)}\right]^j\left[\frac{\bar F_\mu(z_r)}{\bar F_\mu(z_k)}\right]^{t-k-j},\quad (30)$$

where

$$\varphi_{r-k,t-r+1}(u)=\frac{\frac{r-k}{t-r+1}}{\mathrm{B}(r-k,t-r+1)}\frac{\left[\frac{r-k}{t-r+1}u\right]^{r-k-1}}{\left[1+\frac{r-k}{t-r+1}u\right]^{t-k+1}},$$

$$u \in (0,\infty), \qquad (31)$$

is the probability density function (pdf) of the F distribution $(F(r-k,t-r+1))$ with parameters $r-k$ and $t-r+1$, which are positive integers known as the degrees of freedom for the numerator and the degrees of freedom for the denominator.

Proof. It follows from (30) that

$$= \frac{d}{dz_r} \int_0^{\frac{t-r+1}{r-k}\left(1-\frac{\bar{F}_\mu(z_r)}{\bar{F}_\mu(z_k)}\right)\Big/\frac{\bar{F}_\mu(z_r)}{\bar{F}_\mu(y_k)}} \varphi_{r-k,t-r+1}(u)du.$$

$$= \frac{d}{dz_r} P_\mu(Z_r \le z_r \mid Z_k = z_k; t) \qquad (32)$$

Thus,

$\frac{t-r+1}{r-k}\left(1-\bar{F}_\mu(z_r)/\bar{F}_\mu(z_k)\right)/\left(1-\bar{F}_\mu(z_r)/\bar{F}_\mu(z_k)\right)$ is the generalized pivotal quantity:

$$\frac{t-r+1}{r-k}\left(1-\frac{\bar{F}_\mu(z_r)}{\bar{F}_\mu(z_k)}\right)\Big/\frac{\bar{F}_\mu(z_r)}{\bar{F}_\mu(z_k)} = u \sim \varphi_{r-k,t-r+1}(u)$$

$$= \frac{1}{\mathrm{B}(r-k,t-r+1)} \frac{\left[\frac{r-k}{t-r+1}u\right]^{r-k-1}}{\left[1+\frac{r-k}{t-r+1}u\right]^{t-k+1}} \frac{r-k}{t-r+1},$$

$$u \in (0,\infty). \qquad (33)$$

This ends the proof.

3.4 Adequate Mathematical Model 3.4

$$\int_{\frac{r-k}{t-r+1}\frac{\bar{F}_\mu(z_r)}{\bar{F}_\mu(z_k)}\Big/\left(1-\frac{\bar{F}_\mu(z_r)}{\bar{F}_\mu(z_k)}\right)}^{\infty} \varphi_{t-r+1,r-k,}(u)du,$$

$$= P_\mu(Z_r \le z_r \mid Z_k = z_k; t)$$

$$= \sum_{j=r-k}^{t-k} \binom{t-k}{j} \left[1-\frac{\bar{F}_\mu(z_r)}{\bar{F}_\mu(z_k)}\right]^j \left[\frac{\bar{F}_\mu(z_r)}{\bar{F}_\mu(z_k)}\right]^{t-k-j} \qquad (34)$$

where

$$\varphi_{t-r+1,r-k}(u) = \frac{1}{\mathrm{B}(t-r+1,r-k)}$$

$$\times \frac{\left[\frac{t-r+1}{r-k}u\right]^{t-r+1}}{\left[1+\frac{t-r+1}{r-k}u\right]^{m-k+1}} \frac{t-r+1}{r-k}, \quad u \in (0,\infty). \qquad (35)$$

is the probability density function (pdf) of the F distribution $(F(t-r+1,r-k))$ with parameters $t-r+1$ and $r-k$, which are positive integers known as the degrees of freedom for the numerator and the degrees of freedom for the denominator.

Proof. It follows from (30) that

$$= \frac{d}{dz_r} \int_{\frac{r-k}{t-r+1}\frac{\bar{F}_\mu(z_r)}{\bar{F}_\mu(z_k)}\Big/\left(1-\frac{\bar{F}_\mu(z_r)}{\bar{F}_\mu(z_k)}\right)}^{\infty} \varphi_{t-r+1,r-k,}(u)du$$

$$= \frac{d}{dz_r} P_\mu(Z_r \le z_r \mid Z_k = z_k; t) \qquad (36)$$

Thus,

$\frac{r-k}{t-r+1}\left(\bar{F}_\mu(z_r)/\bar{F}_\mu(z_k)\right)/\left(1-\bar{F}_\mu(z_r)/\bar{F}_\mu(z_k)\right)$ is the generalized pivotal quantity:

$$\frac{r-k}{t-r+1}\frac{\bar{F}_\mu(z_r)}{\bar{F}_\mu(z_k)}\Big/\left(1-\frac{\bar{F}_\mu(z_r)}{\bar{F}_\mu(z_k)}\right)$$

$$= u \sim \varphi_{t-r+1,r-k}(u) = \left[\mathrm{B}(t-r+1,r-k)\right]^{-1}$$

$$\times \frac{\left[\frac{t-r+1}{r-k}u\right]^{t-r+1}}{\left[1+\frac{t-r+1}{r-k}u\right]^{t-r+1}} \frac{t-r+1}{r-k}, \quad u \in (0,\infty). \qquad (37)$$

This ends the proof.

4 Two-Parameter Exponential Distribution

Let $\mathbf{Z} = (Z_1 \le \ldots \le Z_r)$ be the first r ordered observations (order statistics) in a sample of size t from the two-parameter exponential distribution with the probability density function (pdf)

$$f_\mu(z) = \vartheta^{-1}\exp\left(-\frac{z-\delta}{\vartheta}\right), \quad \vartheta > 0,\ \mathrm{z} > 0, \qquad (38)$$

and the cumulative distribution function (cdf)

$$F_\mu(z) = 1-\exp\left(-\frac{z-\delta}{\vartheta}\right), \quad \bar{F}_\mu(z) = 1-F_\mu(z), \qquad (39)$$

where $\mu=(\delta,\vartheta)$, δ is the shift parameter and ϑ is the scale parameter. It is assumed that these parameters are unknown. In Type II censoring, which is of primary interest here, the number of survivors is fixed and Z_k is a random variable. In this case, the likelihood function is given by

$$L(\delta,\vartheta)=\prod_{i=1}^{r} f_\mu(z_i)\left(\bar{F}_\mu(z_r)\right)^{t-r}$$

$$=\frac{1}{\vartheta^r}\exp\left(-\left[\sum_{i=1}^{r}(z_i-\delta)+(t-r)(z_r-\delta)\right]\Big/\vartheta\right)$$

$$=\frac{1}{\vartheta^r}\exp\left(-\left[\sum_{i=1}^{r}\begin{matrix}(z_i-z_1+z_1-\delta)\\+(t-r)(z_r-z_1+z_1-\delta)\end{matrix}\right]\Big/\vartheta\right)$$

$$=\frac{1}{\vartheta^{r-1}}\exp\left(-\left[\sum_{i=1}^{r}(z_i-z_1)+(t-r)(z_r-z_1)\right]\Big/\vartheta\right)$$

$$\times\frac{1}{\vartheta}\exp\left(-\frac{t(z_1-\delta)}{\vartheta}\right)$$

$$=\frac{1}{\vartheta^{r-1}}\exp\left(-\frac{s_r}{\vartheta}\right)\times\frac{1}{\vartheta}\exp\left(-\frac{t(s_1-\delta)}{\vartheta}\right),\quad (40)$$

where

$$\mathbf{S}=\begin{pmatrix}S_1=Z_1,\\ S_r=\sum_{i=1}^{r}(Z_i-Z_1)+(t-r)(Z_r-Z_1)\end{pmatrix}\quad (41)$$

is the complete sufficient statistic for μ. The probability density function of $\mathbf{S}=(S_1, S_r)$ is given by

$$f_\mu(s_1,s_r)$$

$$=\frac{\frac{1}{\vartheta^{r-1}}\exp\left(-\frac{s_r}{\vartheta}\right)\times\frac{1}{\vartheta}\exp\left(-\frac{t(s_1-\delta)}{\vartheta}\right)}{\frac{1}{s_r^{r-2}}\int_0^\infty\frac{s_r^{r-2}}{\vartheta^{r-1}}\exp\left(-\frac{s_r}{\vartheta}\right)ds_r\times\frac{1}{t}\int_0^\infty\frac{t}{\vartheta}\exp\left(-\frac{t(s_1-\delta)}{\vartheta}\right)ds_1}$$

$$=\frac{\frac{1}{\vartheta^{r-1}}\exp\left(-\frac{s_r}{\vartheta}\right)\times\frac{1}{\vartheta}\exp\left(-\frac{t(s_1-\delta)}{\vartheta}\right)}{\frac{\Gamma(r-1)}{s_r^{r-2}}\times\frac{1}{t}}$$

$$=\frac{1}{\Gamma(r-1)\vartheta^{r-1}}s_r^{r-2}\exp\left(-\frac{s_r}{\vartheta}\right)$$

$$\times\frac{t}{\vartheta}\exp\left(-\frac{t(s_1-\delta)}{\vartheta}\right)=f_\vartheta(s_r)\,f_\mu(s_1),\quad (42)$$

where

$$f_\mu(s_1)=\frac{t}{\vartheta}\exp\left(-\frac{t(s_1-\delta)}{\vartheta}\right),\quad s_1\ge\delta,\quad (43)$$

$$f_\vartheta(s_r)=\frac{1}{\Gamma(r-1)\vartheta^{r-1}}s_r^{r-2}\exp\left(-\frac{s_r}{\vartheta}\right),\quad s_r\ge 0.\quad (44)$$

$$V_1=\frac{S_1-\delta}{\vartheta}\quad (45)$$

is the pivotal quantity, the probability density function of which is given by

$$f_1(v_1)=t\exp(-tv_1),\quad v_1\ge 0,\quad (46)$$

$$V_r=\frac{S_r}{\vartheta}\quad (47)$$

is the pivotal quantity, the probability density function of which is given by

$$f_r(v_r)=\frac{1}{\Gamma(r-1)}v_r^{r-2}\exp(-v_r),\quad v_r\ge 0.\quad (48)$$

4.1 Constructing One-Sided γ-Content ToleranceLimit with a Confidence Level β (where Model 2.1 is used)

Theorem 3. Let $Z_1 \le \ldots \le Z_r$ be the first r ordered observations from the preliminary sample of size t from a two-parameter exponential distribution defined by the probability density function (37). Then the lower one-sided γ-content tolerance limit with a confidence level β, $L_k = L_k(\mathbf{S})$ (on the kth order statistic Y_k from a set of n future ordered observations $Y_1 \le \ldots \le Y_n$ also from the distribution (37)), which satisfies

$$E\left\{\Pr\left(P_\mu(Y_k>L_k\,|\,n)\ge\gamma\right)\right\}=\beta,\quad (49)$$

is given by

$$L_k=L_k(\mathbf{S})$$

$$=\begin{cases}S_1+\frac{S_r}{t}\left[1-\left(\frac{\Delta_{1-\gamma}^t}{1-\beta}\right)^{\frac{1}{r-1}}\right], & \text{if}\quad t\ge\frac{\ln(1-\beta)}{\ln\Delta_{1-\gamma}},\\ S_1-\frac{S_r}{t}\left[\left(\frac{\Delta_{1-\gamma}^t}{1-\beta}\right)^{\frac{1}{r-1}}-1\right], & \text{if}\quad t<\frac{\ln(1-\beta)}{\ln\Delta_{1-\gamma}},\end{cases}\quad (50)$$

where

$$\Delta_{1-\gamma}=1$$

$$-\,q_{(k,n-k+1),1-\gamma}\left(Beta(k,n\text{-}k+1),\,1-\gamma\ \text{quantile}\right).\quad (51)$$

Proof. It follows from (1), (39) and (49) that

$$\Pr\left(P_\mu(Y_k>L_k\,|\,n)\ge\gamma\right)$$

$$=\Pr\left(\int_0^{F_\mu(L_k)} f_{k,n-k+1}(u)du \le 1-\gamma\right)$$

$$=\Pr\left(1-\exp\left(-\frac{L_k-\delta}{\vartheta}\right)\le q_{k,n-k+1;1-\gamma}\right)$$

$$=\Pr\left(\exp\left(-\frac{L_k-\delta}{\vartheta}\right)\ge 1-q_{k,n-k+1;1-\gamma}\right)$$

$$=\Pr\left(\frac{L_k-S_1}{S_r}\frac{S_r}{\vartheta}+\frac{S_1-\delta}{\vartheta}\le -\ln\left(1-q_{k,n-k+1;1-\gamma}\right)\right)$$

$$=\Pr\left(\frac{S_1-\delta}{\vartheta}\le -\frac{L_k-S_1}{S_r}\frac{S_r}{\vartheta}-\ln\left(1-q_{k,n-k+1;1-\gamma}\right)\right)$$

$$=\Pr\left(V_1\le -\eta_{L_k}V_r-\ln\Delta_{1-\gamma}\right)=\int_0^{-\eta_{L_k}V_r-\ln\Delta_{1-\gamma}} f_1(v_1)dv_1, \quad (52)$$

where

$$\eta_{L_k}=\frac{L_k-S_1}{S_r}. \quad (53)$$

It follows from (49) and (52) that

$$E\left\{\Pr\left(P_\omega(Y_k>L_k\mid n)\ge\gamma\right)\right\}=E\left\{\int_0^{-\eta_{L_k}V_r-\ln\Delta_{1-\gamma}} f_1(v_1)dv_1\right\}$$

$$=E\left\{\int_0^{-\eta_{L_k}V_r-\ln\Delta_{1-\gamma}} t\exp\left(-tv_1\right)dv_1\right\}$$

$$=E\left\{1-\exp\left(-t\left[-\eta_{L_k}V_r-\ln\Delta_{1-\gamma}\right]\right)\right\}$$

$$=E\left\{1-\exp\left(t\eta_{L_k}V_r\right)\exp\left(\ln\Delta_{1-\gamma}^t\right)\right\}$$

$$=E\left\{1-\Delta_{1-\gamma}^t\exp\left(t\eta_{L_k}V_r\right)\right\}$$

$$=\int_0^\infty\left(1-\Delta_{1-\gamma}^t\exp\left(t\eta_{L_k}v_r\right)\right)f_r(v_r)dv_r$$

$$=\int_0^\infty\left(1-\Delta_{1-\gamma}^t\exp\left(t\eta_{L_k}v_r\right)\right)\frac{1}{\Gamma(r-1)}v_r^{r-2}\exp\left(-v_r\right)dv_r$$

$$=1-\frac{\Delta_{1-\gamma}^t}{\left[1-t\eta_{L_k}\right]^{r-1}}=\beta. \quad (54)$$

It follows from (54) that

$$\eta_{L_k}=\frac{L_k-S_1}{S_r}=\frac{1}{t}\left(1-\left[\frac{\Delta_{1-\gamma}^t}{1-\beta}\right]^{\frac{1}{r-1}}\right). \quad (55)$$

It follows from (55) that

$$L_k=S_1+\frac{S_r}{t}\left(1-\left[\frac{\Delta_{1-\gamma}^t}{1-\beta}\right]^{\frac{1}{r-1}}\right). \quad (56)$$

It follows from (56) that

$$1-\left[\frac{\Delta_{1-\gamma}^t}{1-\beta}\right]^{\frac{1}{r-1}}\le 0 \quad \left(\text{if } t\ge\frac{\ln(1-\beta)}{\ln\Delta_{1-\gamma}}\right) \quad (57)$$

or

$$1-\left[\frac{\Delta_{1-\gamma}^t}{1-\beta}\right]^{\frac{1}{r-1}}>0 \quad \left(\text{if } t<\frac{\ln(1-\beta)}{\ln\Delta_{1-\gamma}}\right). \quad (58)$$

Then (50) follows from (56), (57) and (58). This ends the proof.

Corollary 3.1. Let $Z_1\le\ldots\le Z_r$ be the first r ordered observations from the preliminary sample of size t from a two-parameter exponential distribution defined by the probability density function (38). Then the upper one-sided γ-content tolerance limit with a confidence level β, $U_k\equiv U_k(\mathbf{S})$ (on the kth order statistic Y_k from a set of n future ordered observations $Y_1\le\ldots\le Y_n$ also from the distribution (38), which satisfies

$$E\left\{\Pr\left(P_\mu(Y_k\le U_k\mid n)\ge\gamma\right)\right\}=\beta, \quad (59)$$

is given by

$$U_k=\begin{cases} S_1+\dfrac{S_r}{t}\left[1-\left(\dfrac{\Delta_\gamma^t}{\beta}\right)^{\frac{1}{r-1}}\right], & \text{if} \quad t\ge\dfrac{\ln\beta}{\ln\Delta_\gamma}, \\ S_1-\dfrac{S_r}{t}\left[\left(\dfrac{\Delta_\gamma^t}{\beta}\right)^{\frac{1}{r-1}}-1\right], & \text{if} \quad t<\dfrac{\ln\beta}{\ln\Delta_\gamma}, \end{cases} \quad (60)$$

where

$$\Delta_\gamma=1-q_{(k,n-k+1),\gamma}\left(Beta(k,n-k+1),\ \gamma \text{ quantile}\right). \quad (61)$$

4.2 Numerical Practical Example

Let us assume that $k=1$, $r=m=n=15$, $\gamma=\beta=0.95$,

$$\mathbf{S}=\begin{pmatrix} S_1=Z_1=9, \\ S_r=\sum_{i=1}^{r=t}(Z_i-Z_1)+(t-r)(Z_r-Z_1)=192.2508 \end{pmatrix}. \quad (62)$$

Then the lower one-sided γ-content tolerance limit with a confidence level β, $L_{k=1}\equiv L_{k=1}(\mathbf{S})$ can be obtained from (50). Since

$$t=15<\frac{\ln(1-\beta)}{\ln\Delta_{1-\gamma}}=\frac{\ln(1-\beta)}{\ln\left(1-q_{(k,n-k+1),1-\gamma}\right)}=876, \quad (63)$$

where the quantile of $Beta(k, n-k+1), 1-\gamma$ is given by

$$q_{(k,n-k+1),1-\gamma} = 0.003414, \quad (64)$$

it follows from (50) and (64) that

$$L_1(\mathbf{S}) = S_1 - \frac{S_r}{t}\left[\left(\frac{\Delta_{1-\gamma}^t}{1-\beta}\right)^{\frac{1}{r-1}} - 1\right] = 9 - 3 = 6. \quad (65)$$

Statistical inference. From (65) it follows that there is a 95% certainty that failures will not occur in the proportion γ=0.95 or more of a set of n selected items before the end of the lower one-sided γ-content tolerance limit $L_1(\mathbf{S})$ = 6 monthly intervals.

5 Adequate Mathematical Models of the Cumulative Distribution Function of Order Statistics to Construct Equal Tails or Shortest Length Confidence Intervals

Let $\mathbf{Z} = (Z_1 \le \ldots \le Z_r)$ be the first r ordered observations (order statistics) in a sample of size t from the exponential distribution with the probability density function

$$f_\mu(z) = \mu^{-1}\exp(-z/\mu), \quad \mu > 0,\ z > 0, \quad (66)$$

and the cumulative probability distribution function

$$F_\mu(z) = 1 - \exp(-z/\mu), \quad (67)$$

where μ is the scale parameter. It is assumed that the parameter μ is unknown. In Type II censoring, which is of primary interest here, the number of survivors is fixed and Z_r is a random variable. It is known that

$$S_r = \sum_{j=1}^{r} Z_j + (t-r)Z_r \quad (68)$$

is the complete sufficient statistic for μ. The probability density function of S_r is given by

$$f_\mu(s_r) = \frac{1}{\Gamma(r)\mu^r} s_r^{r-1}\exp\left(-\frac{s_r}{\mu}\right), \quad s_r \ge 0. \quad (69)$$

$$V_r = S_r/\mu \quad (70)$$

is the pivotal quantity, the probability density function of which is given by

$$f(v_r) = \frac{1}{\Gamma(r)} v_r^{r-1}\exp(-v_r), \quad v_r \ge 0. \ (Gamma\ (r,1)). \quad (71)$$

Consider the above example, where t units, whose lifetimes are distributed according to the same exponential distribution (66), are put on test simultaneously, and where all units are observed until failure. In this case, $Z_1 \le \ldots \le Z_r$ are the first r ordered observations (the Type II censored sample and the parameter μ is unknown).

1) What is the 100(1−α)% shortest-length confidence interval for μ based on Z_r? *Answer 1:*

5.1 Application of Mathematical Model 2.1

It follows from (6) that $F_\mu(z_r)$ is the generalized pivotal quantity:

$$F_\mu(z_r) = u \sim f_{r,t-r+1}(u) = \frac{u^{r-1}(1-u)^{(t-r+1)-1}}{\mathrm{B}(r, t-r+1)},$$

$$0 < u < 1;\ (Beta(r, t-r+1)). \quad (72)$$

Using (72), it can be obtained a 100(1−α)% confidence interval for μ from

$$\Pr\left(u_1 \le F_\mu(z_r) \le u_2\right) = \Pr\left(u_1 \le 1 - \exp\left(-\frac{z_r}{\mu}\right) \le u_2\right)$$

$$= \Pr\left(1 - u_2 \le \exp\left(-\frac{z_r}{\mu}\right) \le 1 - u_1\right)$$

$$= \Pr\left(\ln\left(\frac{1}{1-u_1}\right) \le \frac{z_r}{\mu} \le \ln\left(\frac{1}{1-u_2}\right)\right)$$

$$= \Pr\left(\frac{z_r}{\ln(1/[1-u_2])} \le \mu \le \frac{z_r}{\ln(1/[1-u_1])}\right)$$

$$= 1 - \alpha \quad (73)$$

by suitably choosing the decision variables u_1 and u_2. Hence, the statistical confidence interval for μ is given by

$$\left[\frac{z_r}{\ln(1/[1-u_2])}, \frac{z_r}{\ln(1/[1-u_1])}\right]. \quad (74)$$

The length of the statistical confidence interval for μ is given by

$$L(u_1, u_2 \mid z_r) = \left(\frac{z_r}{\ln(1/[1-u_1])} - \frac{z_r}{\ln(1/[1-u_2])}\right). \quad (75)$$

In order to find the shortest length confidence interval $L(u_1, u_2 \mid z_r)$, we should find a pair of decision variables u_1 and u_2 such that $L(u_1, u_2 \mid z_r)$ is minimum.

It follows from (73) and (74) that

$$\int_{u_1}^{u_2} f_{r,t-r+1}(u)du = \int_0^{u_2} f_{r,t-r+1}(u)du - \int_0^{u_1} f_{r,t-r+1}(u)du$$

$$=(1-\alpha+p)-p=1-\alpha, \quad (76)$$

where p $(0 \le p \le \alpha)$ is a decision variable,

$$\int_0^{u_2} f_{r,t-r+1}(u)du = (1-\alpha+p) \quad (77)$$

and

$$\int_0^{u_1} f_{k,m-k+1}(u)du = p. \quad (78)$$

Then u_2 represents the $(1-\alpha+p)$ - quantile, which is given by

$$u_2 = q_{1-\alpha+p;(r,t-r+1)}, \quad (79)$$

u_1 represents the p - quantile, which is given by

$$u_1 = q_{p;(r,t-r+1)}. \quad (80)$$

The shortest length confidence interval for μ can be found as follows:

Minimize

$$L^2(u_1,u_2 \mid z_r) = \left[\frac{z_r}{\ln(1/[1-u_1])} - \frac{z_r}{\ln(1/[1-u_2])} \right]^2$$

$$= \left[\begin{array}{l} \dfrac{z_r}{\ln\left(1/\left[1-q_{p;(r,t-r+1)}\right]\right)} \\ -\dfrac{z_r}{\ln\left(1/\left[1-q_{1-\alpha+p;(r,t-r+1)}\right]\right)} \end{array} \right]^2 \quad (81)$$

subject to

$$0 \le p \le \alpha, \quad (82)$$

Numerical Solutions. The optimal numerical solution minimizing $L(u_1, u_2| z_r)$ can be obtained using the computer software "Solver". If, for example, t=10, r = 4, α = 0.05, then the optimal numerical solution is given by

$$p = 0.048394, \quad u_1 = q_{p;(r,t-r+1)} = 0.148512,$$

$$u_2 = q_{1-\alpha+p;(r,t-r+1)} = 0.779435 \quad (83)$$

with the 100(1−α)% shortest-length confidence interval

$$L(u_1,u_2 \mid z_r) = 1026.313 - 109.1584 = 917.1544. \quad (84)$$

The 100(1−α)% equal tails confidence interval is given by

$$L(u_1,u_2 \mid z_r; p = \alpha/2)$$

$$= 1273.159 - 156.1236 = 1117.036. \quad (85)$$

with

$$p = 0.025, \quad u_1 = 0.121552, \quad u_2 = 0.652453. \quad (86)$$

Relative efficiency. The relative efficiency of $L(u_1,u_2 \mid z_r ; p = \alpha/2)$ as compared with $L(u_1,u_2 \mid z_r)$ is given by

$$\text{rel.eff.}_L \left\{ L(u_1,u_2 \mid z_r; p = \alpha/2), L(u_1,u_2 \mid z_r) \right\}$$

$$= \frac{L(u_1,u_2 \mid z_r)}{L(u_1,u_2 \mid z_r; p = \alpha/2)}$$

$$= \frac{917.1544}{1117.036} = 0.821061. \quad (87)$$

2) What is the 100(1−α)% shortest-length confidence interval for μ based on S_r? *Answer 2:*

5.2 Application of Gamma (r,1)

It follows from (71) that $S_r/\mu = u$ represents the pivotal quantity:

$$\frac{S_r}{\mu} = u \sim f_{r,1}(u)$$

$$= \frac{1}{\Gamma(r)} u^{r-1} \exp(-u), \quad u \ge 0, \ (\mathit{Gamma}\ (r,1)). \quad (88)$$

Using (88), it can be obtained a 100(1−α)% confidence interval for μ from

$$\Pr\left(u_1 \le \frac{S_r}{\mu} \le u_2 \right) = \Pr\left(\frac{S_r}{u_2} \le \mu \le \frac{S_r}{u_1} \right) = 1-\alpha \quad (89)$$

by suitably choosing the decision variables u_1 and u_2. Hence, the statistical confidence interval for μ is given by

$$\left[s_r/u_2, s_r/u_1 \right]. \quad (90)$$

The length of the statistical confidence interval for μ is given by

$$L(u_1,u_2 \mid s_r) = \left(s_r/u_1 - s_r/u_2 \right). \quad (91)$$

In order to find the shortest length confidence interval $L(u_1,u_2 \mid s_r)$, we should find a pair of decision variables u_1 and u_2 such that $L(u_1,u_2 \mid s_r)$ is minimum.

It follows from (88) and (89) that

$$\int_{u_1}^{u_2} f_{r,1}(u)du = \int_0^{u_2} f_{r,1}(u)du - \int_0^{u_1} f_{r,1}(u)du$$

$$= (1-\alpha+p) - p = 1-\alpha, \quad (92)$$

where p $(0 \le p \le \alpha)$ is a decision variable,

$$\int_0^{u_2} f_{r,1}(u)du = (1-\alpha+p) \quad (93)$$

and

$$\int_0^{u_1} f_{r,1}(u)du = p. \tag{94}$$

Then u_2 represents the $(1-\alpha+p)$ - quantile, which is given by

$$u_2 = q_{1-\alpha+p;(r,1)}, \tag{95}$$

u_1 represents the p - quantile, which is given by

$$u_1 = q_{p;(r,1)}. \tag{96}$$

The shortest length confidence interval for μ can be found as follows:

Minimize

$$L^2(u_1,u_2 \mid s_r) = \left[\frac{s_r}{u_1} - \frac{s_r}{u_2}\right]^2$$

$$= \left[\frac{s_r}{q_{p;(r,1)}} - \frac{s_r}{q_{1-\alpha+p;(r,1)}}\right]^2 \tag{97}$$

subject to

$$0 \le p \le \alpha, \tag{98}$$

The optimal numerical solution minimizing $L(u_1, u_2 \mid s_r)$ can be obtained using the standard computer software "Solver" of Excel 2016. If, for example, t=10, r = 4, α = 0.05, then the optimal numerical solution is given by

$$p = 0.048393, \quad u_1 = q_{p;(r,1)} = 1.351362,$$

$$u_2 = q_{1-\alpha+p;(r,1)} = 12.45735 \tag{99}$$

with the 100(1−α)% shortest-length confidence interval

$$L(u_1,u_2 \mid s_r) = 1035.992 - 112.3884 = 923.6087. \tag{100}$$

The 100(1−α)% equal tails confidence interval is given by

$$L(u_1,u_2 \mid s_r; p = \alpha/2)$$

$$= 1284.562 - 159.6848 = 1124.878 \tag{101}$$

with

$$p = 0.025, \quad u_1 = 1.089865, \quad u_2 = 8.767273. \tag{102}$$

Relative efficiency. The relative efficiency of $L(u_1,u_2 \mid s_r\,;\, p = \alpha/2)$ as compared with $L(u_1,u_2 \mid s_r)$ is given by

$$\text{rel.eff.}_L \left\{ L(u_1,u_2 \mid s_r; p = \alpha/2), L(u_1,u_2 \mid s_r) \right\}$$

$$= \frac{L(u_1,u_2 \mid s_r)}{L(u_1,u_2 \mid s_r; p = \alpha/2)} = \frac{923.6087}{1124.878}$$

$$= 0.821075. \tag{103}$$

Inference. Two completely different versions of constructing confidence intervals of the shortest length and equal tails gave practically the same final results. This confirms the validity of the analytical conclusions and computational algorithms presented in this paper.

6 New Mathematical Approach to Constructing Statistical Estimates of the Probability Density and Cumulative Distribution Function

Let **Z** = ($Z_1 \le ... \le Z_r$) be the first r ordered observations (order statistics) in a sample of size t from the two-parameter exponential distribution with the probability density function (pdf) (38) and the cumulative distribution function (cdf) (39), where the parametric vector μ is equal to (δ, ϑ); the shift parameter δ and the scale parameter ϑ are unknown.

6.1 Example of Constructing Statistical Estimates for the Two-Parameter Exponential Distribution

Let us suppose that Z is a future observation from the same distribution (39), independent of **Z** = ($Z_1 \le ... \le Z_r$). Then a statistical estimate of (39) can be determined as follows.

Step 1. Invariant embedding of S_1 in (39) to isolate the unknown parameter δ from the problem through V_1 (45),

$$F_\mu(z) = 1 - \exp\left(-\frac{z - s_1 + s_1 - \delta}{\vartheta}\right)$$

$$= 1 - \exp\left(-\frac{z - s_1}{\vartheta}\right)\exp(-v_1), \quad z \ge s_1, \tag{104}$$

Step 2. Averaging (104) over the probability distribution of the pivotal quantity V_1 to eliminate unknown parameter δ from the problem. It follows from (104) and (46) that the pivot-based estimate of the cumulative distribution function (39) (obtained through the pivot-based method) is given by

$$F_{s_1,\vartheta}(z) = \int_0^\infty F_\mu(z) f_1(v_1) dv_1$$

$$= \int_0^\infty \left[1 - \exp\left(-\frac{z - s_1}{\vartheta}\right)\exp(-v_1)\right] t\exp(-tv_1) dv_1$$

$$= 1 - \exp\left(-\frac{z - s_1}{\vartheta}\right) \int_0^\infty t\exp(-v_1[t+1]) dv_1$$

$$= 1 - \exp\left(-\frac{z - s_1}{\vartheta}\right)\frac{t}{t+1}, \quad z \in (s_1, \infty). \tag{105}$$

Since

$$\frac{dF_{s_1,\vartheta}(z)}{dz}=\frac{t}{t+1}\frac{1}{\vartheta}\exp\left(-\frac{z-s_1}{\vartheta}\right) \qquad (106)$$

and

$$\frac{\frac{t}{t+1}\frac{1}{\vartheta}\exp\left(-\frac{z-s_1}{\vartheta}\right)}{\int_{s_1}^{\infty}\frac{t}{t+1}\frac{1}{\vartheta}\exp\left(-\frac{z-s_1}{\vartheta}\right)dz}=\frac{1}{\vartheta}\exp\left(-\frac{z-s_1}{\vartheta}\right), \qquad (107)$$

It follows from (107) that the probability density function (pdf) of Z is given by

$$f_{s_1,\vartheta}(z)=\frac{1}{\vartheta}\exp\left(-\frac{z-s_1}{\vartheta}\right), \quad z\geq s_1, \qquad (108)$$

with the cumulative distribution function

$$F_{s_1,\vartheta}(z)=1-\exp\left(-\frac{z-s_1}{\vartheta}\right). \qquad (109)$$

Step 3. Invariant embedding of S_r in (109) to isolate the unknown parameter δ from the problem through V_r (47),

$$F_{s_1,\vartheta}(z)=1-\exp\left(-\frac{z-s_1}{\vartheta}\right)=1-\exp\left(-\frac{z-s_1}{s_r}\frac{s_r}{\vartheta}\right)$$

$$=1-\exp\left(-\frac{z-s_1}{s_r}v_r\right), \quad z\geq s_1. \qquad (110)$$

Step 4. Averaging (111) over the probability distribution of the pivotal quantity V_r to eliminate unknown parameter ϑ from the problem. It follows from (110) and (48) that the pivot-based estimate of the cumulative distribution function (39) (obtained through the pivot-based method) is given by

$$\int_0^{\infty}F_{s_1,\vartheta}(z)f_r(v_r)dv_r=\int_0^{\infty}\left[1-\exp\left(-\frac{z-s_1}{s_r}v_r\right)\right]$$

$$\times\frac{1}{\Gamma(r-1)}v_r^{r-2}\exp(-v_r)dv_r$$

$$=1-\left(1+\frac{z-s_1}{s_r}\right)^{-(r-1)}=F_{\mathbf{s}}(z). \qquad (111)$$

where

$$\bar{F}_{\mathbf{s}}(z)=1-F_{\mathbf{s}}(z)=\left(1+\frac{z-s_1}{s_r}\right)^{-(r-1)}. \qquad (112)$$

The pivot-based estimate of the probability density function (38) is given by

$$f_{\mathbf{s}}(z)=\frac{dF_{\mathbf{s}}(z)}{dz}=\frac{r-1}{s_r}\left(1+\frac{z-s_1}{s_r}\right)^{-r}, \quad z\geq s_1. \qquad (113)$$

It follows from (111) that the cumulative distribution function of the ancillary statistic

$$W=\frac{Z-S_1}{S_r} \qquad (114)$$

is given by

$$F(w)=1-\frac{1}{(1+w)^{r-1}}. \qquad (115)$$

The probability density function of the ancillary statistic (114) is given by

$$f(w)=\frac{dF(w)}{dw}=\frac{r-1}{(1+w)^r}, \quad w\geq 0. \qquad (116)$$

Constructing Confidence Interval for Z. Using (114) and (115), it can be obtained a 100(1−α)% confidence interval for Z from

$$\Pr(w_1\leq W\leq w_2)=\Pr\left(w_1\leq\frac{Z-S_1}{S_r}\leq w_2\right)$$

$$=\Pr(w_1S_r+S_1\leq Z\leq w_2S_r+S_1)=1-\alpha. \qquad (117)$$

by suitably choosing the decision variables w_1 and w_2. Hence, the statistical confidence interval for Z is given by

$$[w_1s_r+s_1, w_2s_r+s_1]. \qquad (118)$$

The length of the statistical confidence interval for Z is given by

$$L(w_1,w_2\mid s_r)=(w_2s_r-w_1s_r)=(w_2-w_1)s_r. \qquad (119)$$

In order to find the shortest length confidence interval $L(w_1,w_2\mid s_r)$, we should find a pair of decision variables w_1 and w_2 such that $L(w_1,w_2\mid s_r)$ is minimum.

It follows from (116) and (117) that

$$\int_{w_1}^{w_2}f(w)dw=\int_0^{w_2}f(w)dw-\int_0^{w_1}f(w)dw$$

$$=F(w_2)-F(w_1)=(1-\alpha+p)-p=1-\alpha, \qquad (120)$$

where p $(0\leq p\leq\alpha)$ is a decision variable,

$$\int_0^{w_2}f(w)dw=F(w_2)=(1-\alpha+p) \qquad (121)$$

and

$$\int_0^{w_1}f(w)dw=F(w_1)=p. \qquad (122)$$

Then u_2 represents the $(1-\alpha+p)$ - quantile, which is given by

$$w_2=q_{1-\alpha+p}=\left(\frac{1}{\alpha-p}\right)^{1/(r-1)}-1, \qquad (123)$$

w_1 represents the p - quantile, which is given by

$$w_1 = q_p = \left(\frac{1}{1-p}\right)^{1/(r-1)} - 1. \qquad (124)$$

The shortest length confidence interval for *Z* can be found as follows:

Minimize

$$L^2(w_1, w_2 \mid s_r) = \left[\left(w_2 - w_1\right)s_r\right]^2 = \left[\left(q_{1-\alpha+p} -\right)\right]$$

$$= \left[\left(\frac{1}{\alpha - p}\right)^{1/(r-1)} - \left(\frac{1}{1-p}\right)^{1/(r-1)}\right]^2 s_r^2. \quad (125)$$

subject to

$$0 \le p \le \alpha, \qquad (126)$$

The optimal numerical solution minimizing $L(w_1, w_2 \mid s_r)$ can be obtained using the standard computer software "Solver" of Excel 2016. If, for example, $r = 4$, $\alpha = 0.05$, then the optimal numerical solution is given by

$$p = 0 \qquad (127)$$

with the 100(1−α)% shortest-length confidence interval

$$L(w_1, w_2 \mid s_r) = 1.114743 s_r. \qquad (128)$$

The 100(1−α)% equal tails confidence interval is given by

$$L(w_1, w_2 \mid s_r; p = \alpha/2) = 1.508517 s_r \qquad (129)$$

with

$$p = 0.025. \qquad (130)$$

Relative efficiency. The relative efficiency of $L(w_1, w_2 \mid s_r; p = \alpha/2)$ as compared with $L(w_1, w_2 \mid s_r)$ is given by

$$\text{rel.eff.}_L \left\{ L(w_1, w_2 \mid s_r; p = \alpha/2), L(w_1, w_2 \mid s_r) \right\}$$

$$= \frac{L(w_1, w_2 \mid s_r)}{L(w_1, w_2 \mid s_r; p = \alpha/2)} = \frac{1.114743 s_r}{1.508517 s_r}$$

$$= 0.738966. \qquad (131)$$

7 Conclusion

The new intelligent computational methods proposed in this paper are conceptually simple, efficient, and useful for constructing accurate statistical tolerance limits and shortest-length or equal-tailed confidence intervals under the parametric uncertainty of applied stochastic models. The methods listed above are based on adequate mathematical models of the cumulative distribution function of order statistics and constructive use of the principle of invariance in mathematical statistics. We have illustrated proposed intelligent computational methods for the exponential distribution. Applications to other log-location-scale distributions can follow directly.

References:

[1] Nechval, N.A. and Vasermanis, E.K., *Improved Decisions in Statistics*, Riga: Izglitibas soli, 2004.

[2] Nechval, N.A., Berzins, G., Purgailis, M., and Nechval, K.N., Improved estimation of state of stochastic systems via invariant embedding technique, *WSEAS Transactions on Mathematics*, Vol. 7, 2008, pp. 141–159.

[3] Nechval, N.A., Nechval, K.N., Danovich V., and Liepins, T., Optimization of new-sample and within-sample prediction intervals for order statistics, in *Proceedings of the 2011 World Congress in Computer Science, Computer Engineering, and Applied Computing*, WORLDCOMP'11, Las Vegas Nevada, USA, CSREA Press, July 18-21, 2011, pp. 91–97.

[4] Nechval, N.A., Nechval, K.N., and Berzins, G., A new technique for intelligent constructing exact γ-content tolerance limits with expected (1 − α)-confidence on future outcomes in the Weibull case using complete or Type II censored data, *Automatic Control and Computer Sciences (AC&CS)*, Vol. 52, 2018, pp. 476–488.

[5] Nechval, N.A., Berzins, G., Nechval, K.N., and Krasts, J., A new technique of intelligent constructing unbiased prediction limits on future order statistics coming from an inverse Gaussian distribution under parametric uncertainty, *Automatic Control and Computer Sciences (AC&CS)*, Vol. 53, 2019, pp. 223–235.

[6] Nechval, N.A. Berzins, G., and Nechval, K.N., A novel intelligent technique for product acceptance process optimization on the basis of misclassification probability in the case of log-location-scale distributions, in: F. Wotawa et al. (Eds.) *Advances and Trends in Artificial Intelligence. From Theory to Practice.* IEA/AIE 2019, Lecture Notes in Computer Science, vol. 11606, 2019, pp. 801–818, Springer Nature Switzerland AG.

[7] Nechval, N.A., Berzins, G., and Nechval, K.N. A novel intelligent technique of invariant statistical embedding and averaging via pivotal quantities for optimization or improvement of statistical decision rules under parametric uncertainty, *WSEAS Transactions on Mathematics*, Vol. 19, pp. 17-38, 2020.

[8] Nechval, N.A. Berzins, G., and Nechval, K.N., A new technique of invariant statistical embedding and averaging via pivotal quantities for intelligent constructing efficient statistical decisions under parametric uncertainty, *Automatic Control and Computer Sciences (AC&CS)*, Vol. 54, 2020, pp. 191-206.

[9] Nechval, N.A., Berzins, G., and Nechval, K.N., Cost-effective planning reliability-based inspections of fatigued structures in the case of log-location-scale distributions of lifetime under parametric uncertainty, in *Proceedings of the 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference*, Edited by Piero Baraldi, Francesco Di Maio and Enrico Zio, ESREL2020-PSAM15, 1-6 November, 2020, Venice, Italy, pp. 455-462.

[10] Nechval, N.A., Berzins, G., and Nechval, K.N., A new technique of invariant statistical embedding and averaging in terms of pivots for improvement of statistical decisions under parametric uncertainty, CSCE'20 - The 2020 World Congress in Computer Science, Computer Engineering, & Applied Computing, July 27-30, 2020, Las Vegas, USA, in: H. R. Arabnia et al. (eds.), *Advances in Parallel & Distributed Processing, and Applications*, Transactions on Computational Science and Computational Intelligence, pp. 257-274. Springer Nature Switzerland AG 2021.

[11] Nechval, N.A., Berzins, G., and Nechval, K.N., A new simplecomputational method of simultaneous constructing and comparing confidence intervals of shortest length and equal tails for making efficient decisions under parametric uncertainty. *Proceedings of Sixth International Congress on Information and Communication Technology – ICICT 2021*, Lecture Notes in Network and Systems (LNNS, volume 235), Yang X.-S., Sherratt S., Dey N., Joshi A. (eds), 25-26 February 2021, London, United Kingdom, pp. 473-482. Springer Nature Singapore 2022.

[12] Nechval, N.A., Berzins, G., Nechval, K.N., Tsaurkubule, Zh., and Moldovan, M., A new intelligent method for eliminating unknown (nuisance) parameters from underlying models as an alternative to the Bayesian approach. *Journal of Applied Mathematics and Computation*, Vol. 1, 2022, pp. 53-65.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The authors equally contributed in the present research, at all stages from the formulation of the problem to the final findings and solution.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The authors have no conflicts of interest to declare that are relevant to the content of this article.