On Selecting the Best of k Lognormal Distributions

Shanti S. Gupta* and Klaus J. Miescke**
Purdue University
University of Illinois at Chicago

Technical Report #87-36

Department of Statistics
Purdue University

August 1987
Revised November 1988

* The research of this author was supported by the Office of Naval Research Contract N00014–88–K–0170 and NSF Grant Number DMS–8606964 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

** The research of this author was supported by the Air Force Office of Scientific Research Grant 85–0347 at the University of Illinois at Chicago.

⋆

On Selecting the Best of k Lognormal Distributions

Shanti S. Gupta and Klaus J. Miescke
Purdue University
University of Illinois at Chicago

Summary

For k lognormal populations, which differ only in one certain parameter θ, the problem of finding the population with the largest value of θ is considered. For two-parameter lognormal families, several natural choices of θ are treated, where the problem can be solved, through logarithmic transformation of the observations, within the framework of estimating parameters in k, possibly restricted, normal populations. For three-parameter lognormal families, this standard approach of selecting in terms of natural estimators fails to work if θ is the "guaranteed lifetime." For this case, a selection procedure is derived which is based on an L-statistic which has the smallest asymptotic variance. Of importance here is that it is location equivariant, whereas it does not matter what it actually estimates. Comparisons are made with other suitable selection rules, through the asymptotic relative efficiencies, as well as in an example of intermediate sample sizes. It is shown that only in the latter, the selection rule, which is based on the sample minima, compares favorably.

AMS 1980 Subject Classification: Primary 62F07 Secondary 62F12

Key Words: Selection procedures, the best lognormal distribution, L-estimators, finding the most reliable component.

1. Introduction

By definition, we say that a random variable X has a three-parameter lognormal distribution if $ln(X - \gamma)$ is normally distributed with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$, where $0 \leq \gamma < X < \infty$ is called the "guaranteed lifetime." This distribution reduces to the so-called two-parameter lognormal distribution if γ is known, in which case γ is set equal to zero for simplicity.

The lognormal family has been studied and discussed thoroughly in Aitchison and Brown (1969) and in Johnson and Kotz (1970). The latter authors state that, "It is quite likely that the lognormal distribution will be one of the most widely applied distributions in practical statistical work in the near future." Various applications have been considered in the literature, as is reported by Kane (1982). The main part of interest has been the estimation of parameters.

A special feature of this distribution in reliability applications, where X represents the timelength of life (or repair) of a piece of equipment, is that the failure rate (or repair rate) increases at first and then eventually decreases to zero, cf. Gupta, McDonald, and Galarneau (1974).

Estimation for the two-parameter family appears, at a first glance, to be quite easy, since the logarithmic transformation of observations in a given random sample reduces the model to the familiar normal family $N(\mu, \sigma^2)$. Especially, maximum likelihood estimation, which is invariant under transformations, apparently is straightforward. However, estimation of parameters other than μ and σ^2, under different criteria of estimation, may lead to substantial difficulties, as it is pointed out in Johnson and Kotz (1970). A profound analysis, which includes the Bayesian approach, is provided by Zellner (1971). Further Bayes estimators and illustrative examples can be found in Martz and Waller (1982), Chapter 9.3.

Estimation for the three-parameter family turns out to be much more complex. The concept of sufficiency does no longer reduce a sample to a two-dimensional sufficient statistic, but merely ends up with the order statistics. Even the maximum likelihood approach runs into serious difficulties since the maximized likelihood tends to infinity as γ tends to the sample minimum, cf. Johnson and Kotz (1970). Methods to circumvent this difficulty as well as other methods of estimation are considered by Cohen and Whitten (1980), Johnson and Kotz (1970), Kane (1982), and LaRiccia and Kindermann (1983). Furthermore, estimators for the reduced model, where $\sigma^2 = 1$, are discussed in Gibbons and McDonald (1975).

Let us assume that there are k lognormal populations π_i with parameters $(\gamma_i, \mu_i, \sigma_i^2)$, $i = 1, \ldots, k$, from which independent samples $\underline{X}_i$ of a common size $n \geq 2$ have been drawn, $i = 1, \ldots, k$. Assume that these populations differ only in one single parameter θ, say, which may be γ, μ, σ^2, or any other parameter of relevance to a lognormal distribution. The problem considered in this paper is to derive some suitable decision rules, based on the k samples, to find that population which is associated with the largest θ-value.

The standard approach to this problem is to compare populations $\pi_1, \ldots, \pi_k$ through estimators $\hat{\theta}_i = \hat{\theta}_i(\underline{X}_1, \ldots, \underline{X}_k)$, $i = 1, \ldots, k$, where $\hat{\theta}_i$ is an estimator of θ_i, for the k samples of size n, which is optimum in some reasonable way. This will be done in Section 2, where the two-parameter lognormal family case is treated. For various choices of θ, it will be shown that through logarithmic transformation of observations, the problem reduces to the selection of normal parameters which has been dealt with extensively in the literature, cf. Gupta and Panchapakesan (1979) for an overview. Furthermore, special emphasis will be given to the case where θ is the maximum failure (or repair) rate because of its importance to reliability applications.

Section 3 deals with the problem of $\theta = \gamma$ for three-parameter lognormal families, where we assume that $\mu_1 = \ldots = \mu_k$ and $\sigma_1^2 = \ldots = \sigma_k^2$. Because of the difficulties arising in an attempt to estimate γ, it seems natural to choose any other point of location in the three-parameter lognormal density as θ, which can be estimated in an easier and better way. Thinking perhaps first of the mean, median, mode, or some fixed quantile, one arrives eventually at the following nonstandard approach to the present problem, where all points of location are treated without prejudice, i.e. without regard to their actual meaning to the shape of the density. From the theory of L-estimators, which is thoroughly developed and discussed in Huber (1981) and in Lehmann (1983), a minimum variance L-estimator is derived for the three-parameter lognormal family, and then it is used for selecting the population with the largest "guaranteed lifetime" γ. What is actually done is to find a location-equivariant statistic which has a small variance. Here, it does not matter which point of location is actually estimated, and it will be seen that the theory of L-estimators can be successfully applied to the present situation where nonsymmetric densities are considered. The performance of this selection procedure is also examined in an intermediate sample size example of $n = 20$, and it is shown that for this case, the natural selection rule, which selects in terms of the largest of the k sample minima, compares still favorably. Asymptotically, however, it is seen that the latter procedure is inferior.

L-estimators have been used previously by Hustý (1981) to derive robust alternatives to Bechhofer's selection procedure in the indifference zone approach, and asymptotic relative efficiencies are given for normal, uniform, double exponential, and logistic distributions. In a similar approach, Hustý (1984) has found robust alternatives to Gupta's subset selection procedure. In the present paper, neither of the two approaches, which are discussed in full detail by Gupta and Panchapakesan (1979), is emphasized. The results derived in this paper are applicable and useful not only for both these approaches, which depend on very specific assumptions on the loss functions, but also to other decision-theoretic and especially to Bayes approaches.

2. Two-Parameter Lognormal Populations

At the beginning, let us briefly recall some basic properties of a two-parameter lognormal random variable X. First, $ln(X) \sim N(\mu, \sigma^2)$, i.e. the logarithm of X is normally distributed with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$. The density and the cumulative distribution function of X are given by

$$(1) \qquad f(x) = (\sigma x)^{-1}\varphi(\sigma^{-1}(ln x - \mu)), \; F(x) = \Phi(\sigma^{-1}(ln x - \mu)), \; x > 0,$$

where φ denotes the density and Φ the cumulative distribution function of the standard normal distribution $N(0, 1)$.

From the well-known identity for the moments of X, cf. Gupta (1962), $E(X^q) = exp(q\mu + q^2\sigma^2/2), q = 1, 2, \ldots$, we get the following relations between (m, v^2), say, the mean and variance of X, and (μ, σ^2).

$$m = exp(\mu + \sigma^2/2),\ v^2 = exp(2\mu + \sigma^2)[exp(\sigma^2) - 1], \tag{2}$$

and

$$\mu = ln(m^2[v^2 + m^2]^{-1/2}),\ \sigma^2 = ln(1 + v^2/m^2). \tag{3}$$

Furthermore, the median of X is $exp(\mu)$, the mode is equal to $exp(\mu - \sigma^2)$, and the α-th quantile of X is given by $F^{-1}(\alpha) = exp(\mu + \sigma\Phi^{-1}(\alpha))$, $0 < \alpha < 1$.

For applications in reliability problems, where X represents the length of life of a piece of equipment (or a component of a system), the failure rate of X is of importance. It is given by $\psi(x) = f(x)/(1 - F(x)), x > 0$. As mentioned earlier in the introduction, ψ increases at first and then eventually decreases to zero. The following result can be used to control the maximum failure rate in later applications.

Theorem 1. The maximum failure rate of X occurs at time $x_0 = exp(\mu + \sigma w_0)$, where w_0 is the unique solution of $\varphi(w)/(1 - \Phi(w)) = w + \sigma, w \in \mathbb{R}$. Moreover, the maximum value of ψ is equal to $\psi(x_0) = (1 + w_0/\sigma)exp(-\mu - \sigma w_0)$.

Proof: By means of the substitution $x = exp(\mu + \sigma w), w \in \mathbb{R}$, the failure rate can be written as

$$\psi(x) = \sigma^{-1} exp(-\mu + \sigma^2/2)\varphi(w + \sigma)/(1 - \Phi(w)),\ x > 0. \tag{4}$$

Its derivative with respect to w is zero if and only if w satisfies $\varphi(w)/(1 - \Phi(w)) = w + \sigma$. The ratio $(1-\Phi(w))/\varphi(w)$ is called Mill's ratio and has been studied repeatedly in the past literature. From one of the earlier resources, Sampford (1953), one can see that the function $H(w) = \varphi(w)/(1 - \Phi(w))$ is positive with $0 < H'(w) < 1$ and $H''(w) > 0, w \in \mathbb{R}$, is thus convex, and satisfies $\lim_{w \to -\infty} H(w) = 0, H(w) > w$ for $w > 0$, and $\lim_{w \to \infty} (H(w) - w) = 0$. The rest of the proof is straightforward.

Let there now be given k two-parameter lognormal populations π_i with parameters (μ_i, σ_i^2), or in view of (2) and (3), with (m_i, v_i^2), from which independent samples $\underline{X}_i = (X_{i1}, \ldots, X_{in})$ of a common size $n \geq 2$ have been drawn, $i = 1, \ldots, k$. By using the transformation $Z_{ij} = ln(X_{ij}), j = 1, \ldots, n, i = 1, \ldots, k$, one gets independent samples $\underline{Z}_i = (Z_{i1}, \ldots, Z_{in})$ from $N(\mu_i, \sigma_i^2), i = 1, \ldots, k$. Selection procedures for the latter model have been studied extensively in the literature, where an overview of the results can be found in Gupta and Panchapakesan (1979). In the following, it will be shown how various selection problems for the k lognormal populations transform to equivalent problems for k normal populations.

The most common selection problems are those where the k populations differ only in one single parameter θ, say, and the goal is to find that population which has the largest value of θ. The standard approach is then to compare $\pi_1, \ldots, \pi_k$ through estimators $\hat{\theta}_i = \hat{\theta}_i(\underline{X}_1, \ldots, \underline{X}_k), i = 1, \ldots, k$, where $\hat{\theta}_i$ is an estimator of θ_i, for the k samples of size n, which is optimum in some reasonable way. Several cases, which seem quite likely to occur in applications, are discussed below.

<u>Case I</u>: $\sigma_1^2 = \ldots = \sigma_k^2 = \sigma^2$, $\theta = \mu$.

Selecting the population which has the largest $\overline{Z}_i = (Z_{i1} + \ldots + Z_{in})/n$ is the natural procedure for $\theta = \mu$. It has various optimality properties as is reported in Gupta and Panchapakesan (1979). To mention at least one, it has the largest probability of correctly selecting the largest μ_i, among all permutation invariant selection rules. From (2) one can see that it is also optimum, in the same senses, for selecting the largest m_i, i.e. $\theta = m$, or the largest v_i^2, i.e. $\theta = v^2$. Furthermore, the same holds if θ is chosen to be the median, the mode, or any fixed quantile of the two-parameter lognormal family. Finally, from Theorem 1 one concludes that

<u>Corollary 1</u>. <u>Under the assumption of $\sigma_1^2 = \ldots = \sigma_k^2$, the population with the largest (smallest) value of μ has the smallest (largest) maximum failure rate.</u>

<u>Proof</u>: If the variances σ_i^2 are all equal, then the solution w_0, which is mentioned in

Theorem 1, is seen to be the same value for all populations. The proof is thus completed by examining the explicit form of the maximum failure rate, which is given in Theorem 1.

Various ways of implementing this selection procedure under given performance criteria are described in Gupta and Panchapakesan (1979), which include two-stage selection procedures for applications where the common value of σ^2 is unknown.

<u>Case II</u>: $\mu_1 = \ldots = \mu_k = \mu$, $\theta = \sigma^2$.

The natural procedure for $\theta = \sigma^2$ is to select the population with the largest maximum likelihood estimator $\hat{\sigma}_i^2$ of $\sigma_i^2, i = 1, \ldots, k$, under both situations where μ is known or unknown. It has analogous optimality properties as the procedure in Case I, and it can be implemented under standard performance criteria. This time however, no two-stage approach is needed if μ is unknown.

From (2) it can be seen that this selection rule is also appropriate for finding the largest m_i, i.e. $\theta = m$, or the largest v_i^2, i.e. $\theta = v^2$. Furthermore, the same holds if θ is chosen to be the negative mode or any fixed α-quantile, $\alpha \neq 1/2$, of the two-parameter lognormal family. However, a result analogous to Corollary 1 does not hold. A thorough examination of the results in Theorem 1 shows that the maximum failure rate is not a monotone function of σ^2. To be more specific, for $M(w) = (1 + w/\sigma(w))exp(-w\sigma(w))$, and for $\sigma(w) = H(w) - w, w \in \mathbb{R}$, one finds that

$$\sigma(w)^2 exp(w\sigma(w))M'(w) = H(w)[1 - (H(w) - w)^2(1 + wH(w))], \quad w \in \mathbb{R}, \tag{5}$$

which is seen to be positive at $w = 0$, whereas it is negative for sufficiently small w since in (5), [...] $\to -\infty$ as $w \to -\infty$. Noticing that $H(w) - w$ has a negative derivative on $\mathbb{R}$ completes the argumentation.

<u>Case III</u>: $v_1^2 = \ldots = v_k^2 = v^2$, $\theta = m$.

In this case, one can see from (2) and (3) that the problem of selecting the largest μ_i, i.e. $\theta = \mu$, is equivalent to the choice of $\theta = m$, and to the choice of $\theta = \sigma^{-2}$. However, the estimation of $\mu_1, \ldots, \mu_k$, and possibly v^2 in case of an unknown v^2, is a rather difficult task under the given condition of $exp(2\mu_1+\sigma_1^2)[exp(\sigma_1^2)-1] = \ldots = exp(2\mu_k+\sigma_k^2)\ [exp(\sigma_k^2)-1]$. For example, the maximum likelihood estimators cannot be given in closed form but rather have to be found through computer programs. This makes the given selection problem less attractive for applications, as long as the common sample size n is small.

For sufficiently large n, on the other hand, the natural procedure is to select in terms of the largest sample mean $\overline{X}_i = (X_{i1} + \ldots + X_{in})/n, i = 1, \ldots, k$. Since these sample means are independent and asymptotically normal with $N(m_i, v^2/n), i = 1, \ldots, k$, one arrives at the same situation which was considered in Case I, where now the $\overline{X}_i$'s play the role of the $\overline{Z}_i$'s considered there.

Case IV: $m_1 = \ldots = m_k = m,\ \theta = v^2$.

In this case, (2) and (3) show that the problem of selecting the largest σ_i^2, i.e. $\theta = \sigma^2$, is equivalent to the choice of $\theta = v^2$, and to the choice of $\theta = \mu^{-1}$. Similar to the case considered before, estimation of $\sigma_1^2, \ldots, \sigma_k^2$ under the constraint of $\mu_1 + \sigma_1^2/2 = \ldots = \mu_k + \sigma_k^2/2$ is difficult, and for example, the maximum likelihood estimators have no closed form. For sufficiently large n, however, the natural rule, based on asymptotically optimum estimators, selects in terms of the largest $\sum_{j=1}^{n}(X_{ij} - m)^2, i = 1, \ldots, k$, if m is known, and in case of an unknown m, it selects in terms of the largest $\sum_{j=1}^{n}(X_{ij} - \overline{X}), i = 1, \ldots, k$, where $\overline{X}$ is the overall average of the kn observations. Thus, one arrives at the situation considered before in Case II.

The General Case

Some comments have to be made at the end of this section about situations where both μ_i and $\sigma_i^2, i = 1, \ldots, k$ vary freely. Although selection with respect to one parameter $\theta = h(\mu, \sigma^2)$, say, without control of the complementing second parameter dimension, is usually not acceptable in applications, circumstances may arise where this type of problem is of relevance. To give two examples from reliability considerations, $\theta = \mu + \sigma\Phi^{-1}(\alpha)$ corresponds to selection of the largest α-th quantile for some fixed $0 < \alpha < 1$, and $\theta = (\mu - ln(t_0))/\sigma$ refers to finding the largest probability of surviving a fixed time $t_0 > 0$. Such problems can be treated, through the transformed variables Z_{ij}, within the framework of k independent samples from $N(\mu_i, \sigma_i^2), i = 1, \ldots, k$, and they deserve to be studied more carefully in the future.

3. Three-Parameter Lognormal Populations.

Let us assume that there are k lognormal populations π_i with parameters (γ_i, μ_i, σ_i^2), $i = 1, \ldots, k$, from which independent samples $\underline{X}_i$=(X_{i1}, ..., X_{in}) of a common size of $n \geq 2$ have been drawn, $i = 1, \ldots, k$. The problem to be considered is to select that population which has the largest "guaranteed lifelength" γ_i, i.e. $\theta = \gamma$. Throughout this section, we assume that $\mu_1 = \ldots = \mu_k = \mu$, say, and $\sigma_1^2 = \ldots = \sigma_k^2 = \sigma^2$, say, hold, where in the first part, μ and σ^2 are known. This situation is of statistical relevance, and it is complex enough to develop the main ideas of the approach completely. Later it will be seen clearly which adjustments have to be made to adapt the derived selection rule to the case where μ and σ^2 are unknown. The derived selection rule will actually be independent of μ, and for σ^2, a pooled-sample estimator will be found.

The standard approach to this problem has been described and applied in Section 2. For $i = 1, \ldots, k$, a suitable estimator $\hat{\theta}_i = \hat{\theta}_i(\underline{X}_1, \ldots, \underline{X}_k)$, based on the k samples of size n, with good performance properties has to be found, and then the selection is made in terms of the largest of the estimates $\hat{\theta}_1, \ldots, \hat{\theta}_k$. However, as mentioned already in the Introduction, estimation of parameters is in this case rather difficult, cf. Cohen and Whitten (1980), Johnson and Kotz (1970), Kane (1982), and LaRiccia and Kindermann (1983). This is so even if σ^2 is known, as it can be seen from Gibbons and McDonald (1975), where $\sigma^2 = 1$ is assumed.

In ranking and selection problems for location parameter families with nonsymmetric densities, one realizes quickly that comparing the k estimates of a "natural" point of location in these k densities may not be the best approach. Rather than estimating in the present model $\gamma_1, \ldots, \gamma_k$, or the k expectations, medians, modes, etc., and then selecting in terms of the largest of the k estimates, one may better search for any functional of the distribution which can be estimated sufficiently well by a location equivariant statistic.

In the following, we will consider the class of L-estimators, which are linear combinations of the order statistics of a sample. To simplify explanations and notation, let us first consider one sample from a lognormal population, $X_1, \ldots, X_n$, say, where $ln(X_j - \theta) \sim N(\mu, \sigma^2), j = 1, \ldots, n$, and $\mu \in \mathbb{R}$ and $\sigma^2 > 0$ are known. An L-estimator is now of the form

(6) $$L_n = n^{-1} \sum_{j=1}^{n} \lambda\left(j/(n+1)\right) X_{[j]},$$

where $X_{[1]} < \ldots < X_{[n]}$ denote the ordered values of the sample $X_1, \ldots, X_n$, and where λ is a bounded weight function, which is defined on the open unit interval (0,1), and satisfies $\int_0^1 \lambda(t)dt = 1$. It should be pointed out, however, that λ is allowed to assume also negative values. A thorough analysis and discussion of this class of estimators can be found in Huber (1981), and in Lehmann (1983), Chapter 5.5.

The main application of L-estimators is for location parameter families $\tilde{f} = \{f_\theta\}_{\theta \in \mathbb{R}}$, where $f_\theta(x) = f(x - \theta), x \in \mathbb{R}$, and where f is symmetric about the origin. Then for any λ which is symmetric about 1/2, L_n as given by (6) is a suitable candidate for an estimator of θ. As shown in Lehmann (1983), if one takes weights w_{jn} proportional to $\lambda(j/(n+1))$, which add up to one, then

(7) $$\tilde{L}_n = \sum_{j=1}^{n} w_{jn} X_{[j]}$$

is a location equivariant estimator of θ, which is asymptotically normal and unbiased, provided that some mild regularity conditions on f and λ are fulfilled, and the asymptotic distribution of L_n is the same as that one of $\tilde{L}_n$. The last step in this approach is to find a weight function λ which yields the minimum asymptotic variance.

To apply this technique to the lognormal model under concern, one notices the following facts. $\tilde{L}_n$ and L_n have still a common asymptotic normal distribution. Theorem 5.1 of Lehmann (1983) is formulated and proved without the assumption of symmetry of λ and f, and the asymptotic mean and variance are given explicitly in dependence of λ and f. The main idea for solving the k lognormal selection problem is now as follows. Since $\tilde{L}_n$ is location equivariant, it is reasonable to find its value for each of the k samples and then select that population which yields the largest $\tilde{L}_n$-value. And one realizes immediately that this selection procedure remains exactly the same if L_n instead of $\tilde{L}_n$ is employed. To find now a good selection procedure, one has to search for a suitable weight function λ, which guarantees that L_n has a small variance or asymptotic variance. In this paper, the latter approach will be followed.

At this point one can see that selection of populations is not always a comparison of estimates of the location of some natural point of shape in the densities. Varying λ to minimize the asymptotic variance of L_n makes no sense from an estimation point of view, since for a nonsymmetric f, the asymptotic mean varies substantially with λ. From Lehmann (1983) one finds that it is equal to

$$\nu(F,\lambda) = \int_0^1 \lambda(u)F^{-1}(u)du, \tag{8}$$

where F is the cumulative distribution function of f. But it does make sense from a selection point of view, since comparisons of θ_i's is equivalent to comparisons of the locations of any fixed point of the k densities. In other words, if location equivariant statistics are employed, it does not matter what they estimate as long as they are doing it well.

To minimize now the asymptotic variance $\tau^2(F,\lambda)$, say, of L_n, Theorem 5.2 of Lehmann (1983) provides the solution. A careful examination of its proof shows that, with some minor modifications, it can be used as the proof of a generalized version of this theorem, which does not require symmetry of λ and f, and can be applied to the two-parameter lognormal density

$$f(x) = (\sigma x)^{-1}\varphi\left(\sigma^{-1}(lnx-\mu)\right), \; x>0. \tag{9}$$

The construction of the weight function λ_0, say, which minimizes $\tau^2(F,\lambda)$, proceeds as follows. It involves an auxiliary function g, which is denoted by γ in Lehmann (1983).

$$g(x) = -f'(x)/f(x) = x^{-1}[1+\sigma^{-2}(lnx-\mu)], \; x>0, \tag{10}$$

which has the derivative

$$g'(x) = (\sigma x)^{-2}(1+\mu-\sigma^2-lnx), \; x>0. \tag{11}$$

Following along the lines of Lehmann (1983), with the appropriate modifications mentioned above, λ_0 is given by

(12) $$\lambda_0(t) = g'\left(F^{-1}(t)\right) / \int_0^\infty g^2(x)f(x)dx, \ t \in (0,1),$$

which can be further evaluated by means of

(13) $$F^{-1}(t) = exp\left(\mu + \sigma\Phi^{-1}(t)\right), \ t \in (0,1),$$

as well as

(14) $$\int_0^\infty g^2(x)f(x)dx = (1+\sigma^{-2})exp\left(2(\sigma^2 - \mu)\right).$$

The results can be summarized as follows, where the asymptotic normality of L_n for this unbounded $\lambda_0(t)$ is guaranteed by Shorack (1972).

Theorem 2. For a sample of size n from a distribution, which has the density given by (9), the statistic L_n (or $\tilde{L}_n$) which has the smallest asymptotic variance $\tau^2(F, \lambda)$ is based on the weight function

(15) $$\lambda_0(t) = (1+\sigma^2)^{-1}(1 - \sigma[\sigma + \Phi^{-1}(t)]) \cdot exp(-2\sigma[\sigma + \Phi^{-1}(t)]), \ t \in (0,1).$$

Moreover, the asymptotic variance and mean are given by

(16) $$\tau^2(F, \lambda_0) = \sigma^2(1+\sigma^2)^{-1}exp(2(\mu - \sigma^2)), \text{and}$$

(17) $$\nu(F, \lambda_0) = (1+\sigma^2)^{-1}exp(\mu - 3\sigma^2/2).$$

It should be noted that (17) has been stated only for the sake of completeness. It is irrelevant for applications in the selection problem under concern. It follows from (8), i.e.

(18) $$\nu(F, \lambda_0) = \int_0^1 \lambda_0(t)F^{-1}(t)dt.$$

More importantly, one can see from (15) that λ_0 does not depend on μ. Thus, it does not need to be known in applications.

Now we return to the original selection problem for k three-parameter lognormal populations, where the observations X_{ij} satisfy $ln(X_{ij} - \gamma_i) \sim N(\mu, \sigma^2), j = 1, \ldots, n, i = 1, \ldots, k$, and are independent. The goal is to select the population associated with the largest γ-value, i.e. $\theta = \gamma$. It is still assumed that μ and σ^2 are known. The case where they are unknown will be discussed later.

The selection procedure based on L-estimation is to select that population which yields the largest L_n-statistic, based on λ_0 given by (15), from the k samples. This is justified by recalling that the limiting distribution as $n \to \infty$ of $n^{1/2}[L_n(\underline{X}_i) - \gamma_i - \nu(F, \lambda_0)]/\tau(F, \lambda_0)$ is $N(0,1), i = 1, \ldots, k$. Thus, for sufficiently large n, the probability of correctly selecting the population with the largest γ-value is approximated by, say, $P(\lambda_0)$ which is simply the following.

$$(19) \qquad P(\lambda_0) = \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \Phi\left(z + n^{1/2}(\gamma_{[k]} - \gamma_{[i]})/\tau(F, \lambda_0)\right) \varphi(z) dz.$$

It is the well-known PCS for selecting the largest normal mean, which is tabulated for selected values of k and $\Delta = n^{1/2}(\gamma_{[k]} - \gamma_{[i]})/\tau(F, \lambda_0)$, cf. Gupta and Panchapakesan (1979).

Any other selection rule which employs L_n-statistics derived from a different weight function λ has analogous properties. Everything mentioned above holds analogously, λ_0 has simply to be replaced by λ. And since $\tau(F, \lambda_0) \leq \tau(F, \lambda)$, we see that the selection rule based on λ_0 is asymptotically superior to all others based on different λ's. To be more specific, we consider the concept of Pitman's asymptotic relative efficiency (ARE) for selection rules, which is described in details in Miescke (1979). For two L-statistics $L_n^{(1)}$ and $L_n^{(2)}$, say, which are based on λ_1 and λ_2, respectively, one finds that

$$(20) \qquad \text{ARE } (L^{(1)}, L^{(2)}) = \tau^2(F, \lambda_2)/\tau^2(F, \lambda_1).$$

To give two examples, let L_n^* be the sample median, and let L_n^{**} be the sample mean. Then for L_n, based on λ_0, one gets the following asymptotic relative efficiencies, which are greater than one.

$$\text{(21)} \qquad \text{ARE } (L, L^*) = 2^{-1}\pi(1+\sigma^2)exp(2\sigma^2), \text{ and}$$
$$\text{ARE } (L, L^{**}) = (1+\sigma^{-2})exp(3\sigma^2)[exp(\sigma^2)-1].$$

Thus, for the special case of $\sigma^2 = 1$, one finds that ARE $(L, L^*) = 23.2$. This means that asymptotically, the selection rule based on the k medians requires about 23.2 times as many observations as the rule based on L_n, to get the same probability of a correct selection. The result for the sample means procedure is even more impressive. Here one finds ARE $(L, L^{**}) = 69.0$.

To get an idea how well this selection procedure based on L_n through λ_0 performs at an intermediate sample size, the following example is considered.

Example: $n = 20,\ \mu = 0,\ \sigma^2 = 1$.

From the tables in Gupta, McDonald, and Galarneau (1974), the covariance matrix C, say, of the order statistics $X_{[1]} < \ldots < X_{[n]}$ of a sample of size $n = 20$ from a lognormal distribution with $\mu = 0$ and $\sigma^2 = 1$ can be derived. This allows to determine, with sufficient accuracy, the variance of any linear combination of these order statistics, and thus for any $\tilde{L}_n$ based on weights $w_{jn}, j = 1, \ldots, n$, as it is shown in (7), by means of

$$\text{(22)} \qquad \text{Var } (\tilde{L}_n) = \text{ Var } (\underline{w}^T \underline{Y}) = \underline{w}^T C \underline{w},$$

where $\underline{Y} = (X_{[1]}, \ldots, X_{[n]})^T$, and $\underline{w}^T$ denotes the transposed of the column vector $\underline{w}$ of weights.

Since n is not large, it is more appropriate to consider $\tilde{L}_n$, given by (7), since it is location equivariant and thus more suitable for comparison purposes.

The values of $\lambda_0(j/21), j = 1, \ldots, 20$, can be found by use of (15), where $\sigma^2 = 1$, and the values of $\Phi^{-1}(j/21), j = 1, \ldots, 20$, which are tabulated in Mueller, Neumann, and Storm (1977), p. 113. From here, it is easy to determine the weights $w_{j20}, j = 1, \ldots, 20$, which are proportional to $\lambda_0(j/21), j = 1, \ldots, 20$, and add up to one. The values of the weights w_{j20} of $X_{[j]}, j = 1, \ldots, 20$, have been found to be the following: .55690, .21305,

.10716, .05996, .03514, .02083, .01209, .00659, .00306, .00080, -.00063, -.00149, -.00196, -.00216, -.00217, -.00203, -.00180, -.00150, -.00113, -.00070.

The variance of $\tilde{L}_{20}$ is shown below, together with the variances of three other location equivariant statistics, which are natural candidates for selection purposes.

$\tilde{L}_{20}$	$X_{[1]}$	$(X_{[10]} + X_{[11]})/2$	$\overline{X}$
.0076	.0076	.0826	.2335

From this one can see that the asymptotically most efficient selection rule based on $\tilde{L}_{20}$ is already at $n = 20$, much better than the ones based on sample medians and on sample means. But, strikingly, the smallest order statistic provides a reasonable alternative. Thus, although it is not acceptable as an estimator of the "guaranteed lifetime" γ, it turns out to be a very suitable tool for ranking and selection of $\gamma_1, \ldots, \gamma_k$, at least for moderate values of n.

For larger values of n, it can be seen that the rule based on $X_{[1]}$ becomes more and more inferior. The asymptotic variance of $X_{[1]}$ can be derived along the lines of Leadbetter, Lindgren, and Rootzén (1983), and its ratio to the corresponding asymptotic variance of $\tilde{L}_n$ turns out to be 3.46 for $n = 10^3$, 7.47 for $n = 10^4$, 20.22 for $n = 10^5$, and infinity in the limit.

To complete this example, let us mention that the variance of $\tilde{L}_{20}$, approximated by $\tau^2(F, \lambda_0)/20$ through (16) with $\mu = 0$ and $\sigma^2 = 1$, turns out to be equal to .0034. The variance of $X_{[1]}$, approximated by the asymptotic variance of $X_{[1]}$, on the other hand, is 0.0090.

At the end of this section, let us deal with the more general situation where μ and σ^2 are unknown. Lack of knowing μ does not cause any trouble, since $\lambda_0(t)$ does not depend on μ. This can be seen immediately from (15). Moreover, it should be noted that the two specific asymptotic relative efficiencies, which are given by (21), do not depend on μ. This follows from the fact that μ contributes only a factor of $exp(2\mu)$ to the asymptotic variances of linear combinations of the order statistics of a sample from a lognormal distribution.

Now, estimating σ^2 is not such a difficult task as, for example, estimating $\gamma_1, \ldots, \gamma_k$. The moment estimator for one sample, as it is given by Johnson and Kotz (1970), p. 124, can be modified to a pooled-sample estimator for the k samples under concern as follows. Let

$$(23) \qquad \hat{m}_q = (kn)^{-1} \sum_{i=1}^{k} \sum_{j=1}^{n} (X_{ij} - \overline{X}_i)^q, \; q = 2, 3,$$

where $\overline{X}_1, \ldots, \overline{X}_k$ are the usual sample means. Since $\hat{m}_2$ estimates in (2), $v^2 = exp$ $(2\mu + \sigma^2)[exp(\sigma^2) - 1]$, and $\hat{m}_3$ estimates the third central moment, which is the same for all of the k populations, and is for all i, j equal to

$$(24) \qquad \begin{aligned} &E([X_{ij} - E(X_{ij})]^3) \\ &= exp(3\mu + 3\sigma^2/2)[exp(\sigma^2) - 1]^2[exp(\sigma^2) + 2], \end{aligned}$$

it follows that $(\hat{m}_3)^2(\hat{m}_2)^{-3}$ is a suitable estimator of $[exp(\sigma^2) - 1][exp(\sigma^2) + 2]^2$. The proposed estimator of σ^2 is thus the unique solution $\hat{\sigma}^2$ of

$$(25) \qquad [exp(\hat{\sigma}^2) - 1][exp(\hat{\sigma}^2) + 2]^2 = (\hat{m}_3)^2(\hat{m}_2)^{-3}.$$

Although one might have some reservations about the quality of this estimator for one sample of size n, in the present situation, where usually $k \geq 3$ samples contribute to $\hat{\sigma}^2$, this estimator is expected to perform sufficiently well to be used in the selection procedure under concern.

Acknowledgements

The authors are grateful to the referees for their helpful comments, for suggesting to account for the asymptotic variance of the smallest order statistic, and for bringing to their attention the work of Jaroslav Hustý.

References

[1] Aitchison, J., and Brown, J. A. C. (1969). The Lognormal Distribution. 4th ed., Cambridge University Press, London.

[2] Cohen, A. C., and Whitten, B. J. (1980). Estimation in the three-parameter lognormal distribution. *J. Amer. Statist. Assoc.* **75**, 399–404.

[3] Gibbons, D. I., and McDonald, G. C. (1975). Small-sample estimation for the lognormal distribution with unit shape parameter. *IEEE Transactions on Reliability*, **R–24**, 290–295.

[4] Gupta, S. S. (1962). Life test sampling plans for normal and lognormal distributions. *Technometrics* **4**, 151–175.

[5] Gupta, S. S., McDonald, G. C., and Galarneau, D. I. (1974). Moments, product moments and percentage points of the order statistics from the lognormal distribution for samples of size twenty and less. *Sankhyā Ser. B*, Vol. **36**, 230–260.

[6] Gupta, S. S., and Panchapakesan, S. (1979). Multiple Decision Procedures. J. Wiley, New York.

[7] Huber, P. J. (1981). Robust Statistics. J. Wiley, New York.

[8] Hustý, J. (1981). Ranking and selection procedures for location parameter case based on L-estimates. *Aplikace Matematiky* **26**, 377–388.

[9] Hustý, J. (1984). Subset selection of the largest location parameter based on L-estimates. *Aplikace Matematiky* **29**, 397–410.

[10] Johnson, N. L., and Kotz, S. (1970). Continuous Univariate Distributions–1. Houghton Mifflin, Boston.

[11] Kane, V. E. (1982). Standard and goodness–of–fit parameter estimation methods for the three-parameter lognormal distribution. *Commun. Statist.-Theor. Meth.*, **11**, 1935–1957.

[12] LaRiccia, V. N., and Kindermann, R. P. (1983). An asymptotically closed form estimator for the three-parameter lognormal distribution. *Commun. Statist.-Theor. Meth.* **12**, 243–261.

[13] Leadbetter, M. R., Lindgren, G., and Rootzén, H. (1983). Extremes and Related Properties of Random Sequences and Processes. Springer Verlag, New York.

[14] Lehmann, E. L. (1983). Theory of Point Estimation. J. Wiley, New York.

[15] Martz, H. F., and Waller, R. A. (1982). Bayesian Reliability Analysis. J. Wiley, New York.

[16] Mueller, P. H., Neumann, P., and Storm, R. (1977). Tables for Mathematical Statistics (German). Carl Hanser Verlag, Munich.

[17] Miescke, K. J. (1979). Identification and selection procedures based on tests. *Ann. Statist.* **7**, 207–219.

[18] Sampford, M. R. (1953). Some inequalities on Mill's ratio and related functions. *Ann. Math. Statist.* **24**, 130–132.

[19] Shorack, G. R. (1972). Functions of order statistics. *Ann. Math. Statist.* **43**, 412–427.

[20] Zellner, A. (1971). Bayesian and non-Bayesian analysis of the log-normal distribution and log-normal regression. *J. Amer. Statist. Assoc.* **66**, 327–330.

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release, distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) Technical Report #87-36	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Purdue University 6b. OFFICE SYMBOL *(If applicable)*	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS *(City, State, and ZIP Code)* Department of Statistics West Lafayette, IN 47907	7b. ADDRESS *(City, State, and ZIP Code)*
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Office of Naval Research 8b. OFFICE SYMBOL *(If applicable)*	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-88-K-0170 and NSF DMS-8606964
8c. ADDRESS *(City, State, and ZIP Code)* Arlington, VA 22217-5000	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

11. TITLE *(Include Security Classification)*
ON SELECTING THE BEST OF k LOGNORMAL DISTRIBUTIONS (Unclassified)

12. PERSONAL AUTHOR(S)
Shanti S. Gupta and Klaus J. Miescke

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Technical	FROM ____ TO ____	August 1987	18

16. SUPPLEMENTARY NOTATION

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*
Selection procedures, the best lognormal distribution, L-estimators, finding the most reliable component.

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

For k lognormal populations, which differ only in one certain parameter θ, the problem of finding the population with the largest value of θ is considered. For two-parameter lognormal families, several natural choices of θ are treated, where the problem can be solved, through logarithmic transformation of the observations, within the framework of estimating parameters in k, possibly restricted, normal populations. For three-parameter lognormal families, this standard approach of selecting in terms of natural estimators fails to work if θ is the "guaranteed lifetime." For this case, a selection procedure is derived which is based on an L-statistic which has the smallest asymptotic variance. Of importance here is that it is location equivariant, whereas it does not matter what it actually estimates. Comparisons are made with other suitable selection rules, through the asymptotic relative efficiencies, as well as in an example of intermediate sample sizes. In the latter, it is seen that the selection rule, which is based on the sample minima, compares favorably.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Shanti S. Gupta	22b. TELEPHONE *(Include Area Code)* 317-494-6031 22c. OFFICE SYMBOL