Research Article **Environmetrics**

Received: 27 March 2015, Revised: 3 August 2015, Accepted: 4 August 2015, Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/env.2357

Likelihood ratio tests for comparing several gamma distributions

Kalimuthu Krishnamoorthy[a*], Meesook Lee[b] and Wang Xiao[a]

Likelihood ratio tests (LRTs) for comparing several independent gamma distributions with respect to shape parameters, scale parameters, and means are derived. The LRT for testing homogeneity of several gamma distributions is also derived. Extensive simulation studies indicate that the null distributions of the LRT statistics for all four problems depend on the parameters only weakly, and for practical purposes, they are independent of the parameters. Furthermore, our simulation studies suggest that the null distribution of the LRT statistic for testing the equality of shape parameters and that of the LRT statistic for testing the equality of scale parameters are essentially the same. Numerical algorithms to compute the maximum likelihood estimates and *p*-values are given. Percentiles of the null distributions are tabulated for some selected sample sizes to compare three and five gamma distributions. The methods are illustrated using two practical examples.

Keywords: constrained MLEs; large sample test; power; rainfall distribution; type I error

1. INTRODUCTION

The gamma distribution is one of the most commonly used distributions for analyzing meteorological data. This distribution is also used to model pollution/workplace exposure data and lifetime data. In meteorology, gamma distributions are used to model the amounts of daily rainfall in a region (Das, 1955 and Stephenson *et al.*, 1999), and to model the distribution of raindrop sizes (Brawn and Upton, 2007). Applications of gamma models to estimate the percentiles of annual maximum flood series data are reported in Ashkar and Ouarda (1998) and to compare the scale parameters of rainfall distributions for different seasons are provided in Schickedanz and Krause (1970). For other applications in environmental monitoring, ground water monitoring, industrial hygiene, lifetime data analysis, see Gibbons (1994), Lawless (2003), Bhaumik and Gibbons (2006), Krishnamoorthy *et al.* (2008), and Bhaumik *et al.* (2009).

There are difficulties associated with the problems of estimating or testing the gamma parameters. The standard methods based on pivotal quantities do not work for gamma distributions as the family of gamma distributions is not a location-scale family. The maximum likelihood estimates (MLEs) are not available in closed-form, and they can be evaluated only numerically. Most of the results available in the literature are approximate or based on large sample methods. For example, Shiue and Bain (1983) proposed an approximate test for the equality of two scale parameters assuming that the shape parameters are equal. Shiue, Bain and Engelhardt (1988) have proposed an approximate test for the equality of two gamma means. However, the problem of comparing the parameters of more than two gamma distributions has not received much attention in the literature. Tripathi *et al.* (1993) have proposed asymptotic tests based on the generalized minimum chi-squared method. In particular, they have presented asymptotic tests for equality of means, scale parameters, and shape parameters. Recently, Chang *et al*. (2011) have proposed a parametric bootstrap test for the equality of means. This test is based on a statistic whose percentiles are estimated by a Monte Carlo method based on simulated samples from gamma distributions using the MLEs under the null hypothesis of equal means as parameters. This parametric bootstrap test seems to be satisfactory for moderate to large samples. Thiagarajah (2013) has proposed an asymptotic test for homogeneity of several gamma distributions on the basis of Fisher's method of combining independent tests. Such asymptotic tests are valid for large samples.

In some situations, sample sizes are typically small. For example, Bhaumik and Gibbons (2006) have pointed out that assessing environmental impact on the basis of a small number of samples obtained from an area of concern is quite common in environmental monitoring. In exposure data analysis, the sample sizes are often small because of the cost of sampling and burden on workers. Application of large sample methods for small samples often leads to incorrect decisions, and so methods that are accurate for small samples are really warranted. Towards this, we propose tests based on the empirical distributions of the LRT statistics. As shown in Section 4, extensive simulation studies indicate that the percentiles of the LRT statistics depend only on sample sizes and the number of distributions to be compared. These empirical results suggest that the null distributions of the LRT statistics only weakly depend on parameters, and so the null distributions can be evaluated empirically by assuming some parameter values; for example, by assuming all the shape and scale parameters are equal to one.

* *Correspondence to: K. Krishnamoorthy, Department of Mathematics, University of Louisiana at Lafayette, Lafayette, LA, U.S.A. E-mail: krishna@louisiana.edu*

a *Department of Mathematics, University of Louisiana at Lafayette, Lafayette, LA, U.S.A.*

b *Department of Mathematics, South Louisiana Community College, Lafayette, LA, U.S.A.*

The LRTs based on such empirical distributions are practically exact in the sense that the empirical distribution can be used to test any parameter combinations. This is a generalization of the tests for the two-sample problems by Krishnamoorthy and Luis (2014) whose simulation studies indicated that such tests are almost exact for testing equality of two shape parameters or equality of two scale parameters.

In this article, we derive the likelihood ratio test (LRT) statistics for testing equality of shape parameters of several gamma distributions and for testing equality of several scale parameters. As the coefficient of variation of a gamma distribution is the reciprocal of the square root of the shape parameter, comparison of several shape parameters is the same as comparing coefficients of variation. We also develop tests for the equality of several gamma means and for testing homogeneity of several independent gamma distributions. The latter problem arises where data were collected from different sites or workplaces, and one wants to test if the pollution/exposure distributions are the same gamma distribution. Our LRT for homogeneity of gamma distributions is practically exact and is valid for any sample sizes.

The rest of the article is organized as follows. In the following section, we provide some preliminary results. In Section 3, we describe the LRTs for the equality of shape parameters, equality of scale parameters, and equality of means and for the equality of several independent gamma distributions. The empirical distributions of the LRT statistics are studied and evaluated in Section 4. The proposed LRT for the means is compared with the parametric bootstrap test proposed by Chang *et al.* (2011). A power comparison clearly indicates that the LRT for the equality of means is more powerful than the parametric bootstrap test for small samples. The LRTs are illustrated using two examples in Section 6, and some concluding remarks are given in Section 7.

2. PRELIMINARIES

The two-parameter gamma distribution, denoted by gamma(a, b), has the probability density function

$$f(x|a,b) = \frac{1}{\Gamma(a)b^a} e^{-(x)/b} x^{a-1}, \quad a > 0,\ b > 0$$

where a is the shape parameter and b is the scale parameter. Let $X_1, \ldots, X_n$ be a sample from a gamma(a, b) distribution. Let $\bar{X}$ and $\tilde{G}$ denote respectively the arithmetic mean and geometric mean of the samples. That is,

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{and} \quad \tilde{G} = \left(\prod_{i=1}^{n} X_i\right)^{1/n} \tag{1}$$

The log-likelihood function is expressed as

$$l(a,b) = -n\ln\Gamma(a) - na\ln b - n\bar{X}/b + (a-1)n\ln\tilde{G} \tag{2}$$

The MLE $\widehat{a}$ of the shape parameter a is the solution of the equation

$$\ln(a) - \psi(a) = \ln(\bar{X}/\tilde{G}) \tag{3}$$

where ψ denotes the digamma function. Letting $s = \ln(\bar{X}/\tilde{G})$, an approximation[†] to $\widehat{a}$ is given by

$$\widehat{a} \approx \frac{3 - s + \sqrt{(s-3)^2 + 24s}}{12s} \tag{4}$$

The absolute error of the aforementioned approximate MLE is no more than 1.5% of the true MLE. Using this approximate MLE as the initial value, say a_0, the MLE can be evaluated by the Newton–Raphson iterative scheme

$$a_{j+1} = a_j - \frac{\ln a_j - \psi(a_j) - s}{1/a_j - \psi'(a_j)}, \quad j = 0, 1, 2, \ldots$$

where $\psi'(x) = \partial\psi(x)/\partial x$ is the trigamma function. The aforementioned iterative scheme, with $a_0 = \widehat{a}$ defined in (4), converges in a few iterations (in most cases, three or less). The MLE of b is $\widehat{b} = \bar{X}/\widehat{a}$.

3. LIKELIHOOD RATIO TESTS

Let $(\bar{X}_i, \tilde{G}_i)$ denote the (arithmetic mean, geometric mean) based on a sample of size n_i from gamma(a_i, b_i) distribution, $i = 1, \ldots, k$.

3.1. The likelihood ratio test for equality of shape parameters

Consider testing

$$H_0 : a_1 = \ldots = a_k \text{ versus } H_a : a_i \neq a_j \text{ for some } i \neq j. \tag{5}$$

[†] Gamma distribution, Wikipedia.

Following (2), the log-likelihood function under H_0 is expressed as

$$\sum_{i=1}^{k} l(a, b_i) = -\sum_{i=1}^{k} n_i \left(\ln \Gamma(a) + a \ln(b_i) + \frac{\bar{X}_i}{b_i} - (a-1) \ln \tilde{G}_i \right) \tag{6}$$

where a is the unknown common shape parameter. It can be readily checked that the partial differential equation with respect to b_i yields $b_i = \bar{X}_i / a$, $i = 1, \ldots, k$. After replacing b_i in (6) by $\bar{X}_i / a$, the equation $\partial \sum_{i=1}^{k} l(a, b_i) / \partial a = 0$ yields

$$\ln a - \psi(a) = \sum_{i=1}^{n} w_i \ln \frac{\bar{X}_i}{\tilde{G}_i} \tag{7}$$

where $w_i = n_i / \sum_{j=1}^{k} n_j$. Noting that the aforementioned equation is similar to (3), the root can be found using the Newton–Raphson method with the starting value as defined in (4) with $s = \sum_{i=1}^{n} w_i \ln \frac{\bar{X}_i}{\tilde{G}_i}$. The root of the aforementioned equation, denoted by $\widehat{a}_c$, is the MLE of the unknown common shape parameter under $H_0 : a_1 = \ldots = a_k$. The MLEs of b_i's under H_0 are given by $\widehat{b}_{ic} = \bar{X}_i / \widehat{a}_c$, $i = 1, \ldots, k$.

Using the MLEs and the constrained MLEs, the LRT statistic for testing the equality of shape parameters is expressed as

$$\Lambda_a = 2 \left\{ \sum_{i=1}^{k} l(\widehat{a}_i, \widehat{b}_i) - \sum_{i=1}^{k} l(\widehat{a}_c, \widehat{b}_{ic}) \right\} = 2 \sum_{i=1}^{k} n_i \left(\ln \frac{\Gamma(\widehat{a}_c)}{\Gamma(\widehat{a}_i)} - (\widehat{a}_c \ln \widehat{a}_c - \widehat{a}_i \ln \widehat{a}_i) + (\widehat{a}_c - \widehat{a}_i)[\ln(\bar{X}_i / \tilde{G}_i) + 1] \right) \tag{8}$$

where $l(a, b)$ is given in (2), $(\widehat{a}_i, \widehat{b}_i)$ is the MLE of (a_i, b_i) based on $(\bar{X}_i, \tilde{G}_i)$, $i = 1, \ldots, k$, and $\widehat{a}_c$ is the constrained MLE of a satisfying (7).

For a specified level of significance α, the large sample approximate test rejects H_0 in (5) when $\Lambda_a > \chi^2_{k-1;1-\alpha}$, where $\chi^2_{m;p}$ denotes the $100p$ percentile of a chi-squared distribution with degrees of freedom m. For small samples, the null distribution of the LRT statistic can be estimated as shown in Section 4.1.

3.2. The likelihood ratio test for equality of scale parameters

Consider testing

$$H_0 : b_1 = \ldots = b_k \quad \text{versus} \quad H_a : b_i \neq b_j \text{ for some } i \neq j. \tag{9}$$

The log-likelihood function under H_0 can be expressed as

$$\sum_{i=1}^{k} l(a_i, b) = -\sum_{i=1}^{k} n_i \left(\ln \Gamma(a_i) + a_i \ln(b) + \frac{\bar{X}_i}{b} - (a_i - 1) \ln(\tilde{G}_i) \right) \tag{10}$$

where b is the unknown common scale parameter under H_0. The partial differential equation $\partial \sum_{i=1}^{k} l(a_i, b) / \partial b = 0$ yields $b = \sum_{i=1}^{k} n_i \bar{X}_i / \sum_{i=1}^{k} n_i a_i$. Substituting this expression for b in (10), and then differentiating with respect to a_i's, we obtain the following set of equations:

$$n_i \ln \left(\sum_{j=1}^{k} n_j a_j \right) - n_i \psi(a_i) - n_i \ln \left(\sum_{j=1}^{k} n_j \bar{X}_j \right) + n_i \ln(\tilde{G}_i) = 0, \quad i = 1, \ldots, k \tag{11}$$

By solving the aforementioned set of equations (Appendix A) for a_i's, we obtain the constrained MLEs, denoted by $\widehat{a}_{ic}$'s, for shape parameters a_i's.

The LRT statistic for testing the equality of scale parameters is given by

$$\Lambda_b = 2 \left\{ \sum_{i=1}^{k} l(\widehat{a}_i, \widehat{b}_i) - \sum_{i=1}^{k} l(\widehat{a}_{ic}, \widehat{b}_c) \right\} = 2 \sum_{i=1}^{k} n_i \left(\ln \frac{\Gamma(\widehat{a}_{ic})}{\Gamma(\widehat{a}_i)} + \widehat{a}_{ic} \ln \left(\widehat{b}_c\right) - \widehat{a}_i \ln \left(\widehat{b}_i\right) + (\widehat{a}_{ic} - \widehat{a}_i) \left(1 - \ln(\tilde{G}_i)\right) \right) \tag{12}$$

where $l(a, b)$ is given in (2), $(\widehat{a}_i, \widehat{b}_i)$ is the MLE of (a_i, b_i) based on $(\bar{X}_i, \tilde{G}_i)$, $i = 1, \ldots, k$, and $\widehat{a}_{ic}$'s are the constrained MLEs of a_i's satisfying (11), and $\widehat{b}_c = \sum_{i=1}^{k} n_i \bar{X}_i / \sum_{i=1}^{k} n_i \widehat{a}_{ic}$. For large samples, the LRT rejects the null hypothesis of equal scale parameters when $\Lambda_b > \chi^2_{k-1;1-\alpha}$. For small samples, the null distribution of Λ_b can be evaluated empirically as shown in Section 4.2.

3.3. The likelihood ratio test for equality of means

Let $\mu_i = a_i b_i$, $i = 1, \ldots, k$, and consider testing

$$H_0 : \mu_1 = \ldots = \mu_k \text{ versus } H_a : \mu_i \neq \mu_j \text{ for some } i \neq j. \tag{13}$$

Denoting the unknown common mean under H_0 by μ, the log-likelihood function under H_0 can be expressed as

$$\sum_{i=1}^{k} l(a_i, \mu) = -\sum_{i=1}^{k} n_i \left(\ln \Gamma(a_i) + a_i \ln \frac{\mu}{a_i} + \bar{X}_i \frac{a_i}{\mu} - (a_i - 1) \ln \tilde{G}_i \right) \tag{14}$$

The equation $\partial \sum_{i=1}^{k} l(a_i, \mu)/\partial a_i = 0$ yields

$$\ln a_i - \psi(a_i) = \ln \frac{\mu}{\tilde{G}_i} + \frac{\bar{X}_i}{\mu} - 1 = \ln \frac{\bar{X}_i}{\tilde{G}_i} - \ln \frac{\bar{X}_i}{\mu_i} + \frac{\bar{X}_i}{\mu} - 1 \tag{15}$$

The equation $\partial \sum_{i=1}^{k} l(a_i, \mu)/\partial \mu = 0$ yields

$$\mu = \frac{\sum_{i=1}^{k} n_i a_i \bar{X}_i}{\sum_{i=1}^{k} n_i a_i} \tag{16}$$

The constrained MLEs of the shape parameters are the roots of the (15), and these roots can be obtained numerically; see Appendix B. Let us denote the constrained MLEs of the shape parameters by $\widehat{a}_{1c}, \ldots, \widehat{a}_{kc}$. The constrained MLE $\widehat{\mu}_c$ is given by (16) with a_i replaced by $\widehat{a}_{ic}$, $i = 1, \ldots, k$.

The LRT statistic is defined as

$$\Lambda_\mu = 2 \left\{ \sum_{i=1}^{k} l(\widehat{a}_i, \widehat{b}_i) - \sum_{i=1}^{k} l(\widehat{a}_{ic}, \widehat{\mu}_c) \right\} = 2 \sum_{i=1}^{k} n_i \left(\ln \frac{\Gamma(\widehat{a}_{ic})}{\Gamma(\widehat{a}_i)} + \widehat{a}_{ic} \ln \widehat{b}_{ic} - \widehat{a}_i \ln \widehat{b}_i + \widehat{a}_i \left(\frac{\widehat{b}_i}{\widehat{b}_{ic}} - 1 \right) - (\widehat{a}_{ic} - \widehat{a}_i) \ln \tilde{G}_i \right) \tag{17}$$

where $\widehat{b}_{ic} = \widehat{\mu}_c / \widehat{a}_{ic}$, $l(a_i, b_i)$ is defined in (2) and $l(a_i, \mu)$ is defined in (14). For large samples, the LRT rejects the null hypothesis of equal means whenever $\Lambda_\mu > \chi^2_{k-1;1-\alpha}$.

3.4. The likelihood ratio test for homogeneity of several gamma distributions

The hypotheses for testing homogeneity of several gamma distributions are

$$H_0 : (a_1, b_1) = \ldots = (a_k, b_k) \quad \text{versus} \quad H_a : (a_i, b_i) \neq (a_j, b_j) \text{ for some } i \neq j. \tag{18}$$

The log-likelihood function under H_0 is simply the log-likelihood function based on a single sample of size $N = \sum_{i=1}^{k} n_i$ from a gamma(a, b) distribution and is expressed as

$$l(a, b) = -N \ln \Gamma(a) - N a \ln b - N \frac{\bar{\bar{X}}}{b} + N(a - 1) \ln \tilde{\tilde{G}} \tag{19}$$

where $(\bar{\bar{X}}, \tilde{G})$ is the (mean, geometric mean) based on all N observations. The LRT statistic for testing H_0 in (18) is given by

$$\Lambda_E = 2 \left(\sum_{i=1}^{k} l(\widehat{a}_i, \widehat{b}_i) - l(\widehat{a}, \widehat{b}) \right) = \sum_{i=1}^{k} n_i \left[\ln \left(\frac{\widehat{a}}{\widehat{a}_i} \right) + \ln \left(\frac{\widehat{b}^{\widehat{a}}}{\widehat{b}_i^{\widehat{a}_i}} \right) + \left(\frac{\bar{\bar{X}}}{\widehat{b}} - \frac{\bar{X}_i}{\widehat{b}_i} \right) - \ln \left(\frac{\tilde{\tilde{G}}^{\widehat{a}-1}}{\tilde{G}_i^{\widehat{a}_i - 1}} \right) \right]$$

where $l(a, b)$ is given in (19), $\widehat{a}_i$ and $\widehat{b}_i$ are the MLEs based on the ith sample.

For large samples, the LRT rejects H_0 in (18) when $\Lambda_E > \chi^2_{2k-2;1-\alpha}$. For small samples, the distribution of Λ_E can be evaluated empirically as shown in Section 4.3.

4. NULL DISTRIBUTIONS OF LRT STATISTICS

In order to study the null distribution of the LRT statistic Λ_a, we can assume without loss of generality that $b_1 = \ldots = b_k = 1$, because the LRT statistics in the preceding sections are invariant under the transformation $X_{ij} \to c_i X_{ij}$, $j = 1, \ldots, n_i$, $i = 1, \ldots, k$, and $c_i > 0$ for all i. To see this for the case of shape parameters, note that the MLE $\widehat{a}_i$ is a function of $\bar{X}_i/\tilde{G}_i$ (3), and so $\widehat{a}_i$ is invariant under the aforementioned scale transformation. It follows from (7) that the constrained MLE $\widehat{a}_c$ is also invariant under the scale transformation. As the LRT statistic Λ_a in (8) is a function of only $(\widehat{a}_c, \widehat{a}_i, \bar{X}_i/\tilde{G}_i)$, it is invariant under the scale transformation. This invariance property for other testing problems can be verified similarly. Null distributions of the LRT statistics may depend on the unknown shape parameters. However, our extensive simulation studies for each of the testing problems indicate that the percentiles of the null distributions of the LRT statistics Λ_a, Λ_b, Λ_μ, and Λ_E are affected mainly by the number of distributions to be compared and the sample sizes. These simulation results simply imply that the null distributions depend on the parameters only weakly. In the following, we shall provide some simulation results for each of the testing problems.

4.1. Empirical distribution of Λ_a

To show some evidence for our claim that the null distribution is not much affected by the parameter values, we estimated the percentiles of the null distribution of Λ_a using a Monte Carlo method consisting of 100,000 runs for some values of the common unknown shape parameter under H_0 in (5) and some sample sizes. The percentiles are plotted in Figure 1 for the cases of $k = 3$ and 5. The first plot shows the percentiles of Λ_a when all three shape parameters are small. For other two plots, we chose values of shape parameters that are very much different. These three plots clearly indicate that the percentiles of the LRT statistic Λ_a for various values of common a under H_0 are almost identical. In other words, we have strong simulation evidence to indicate that the null distribution of Λ_a does not depend much on any parameters, it depends mainly on the number of shape parameters being tested and sample sizes, and so the LRT is practically exact.

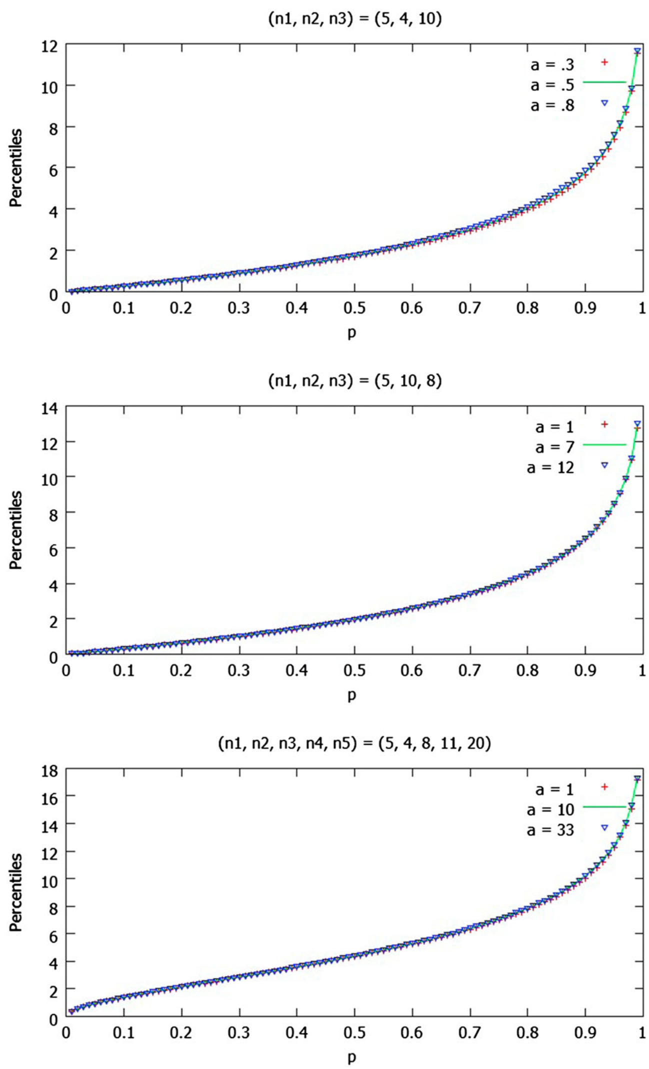

Figure 1. $100p$ percentiles of the likelihood ratio test statistic Λ_a for different common shape parameters under H_0

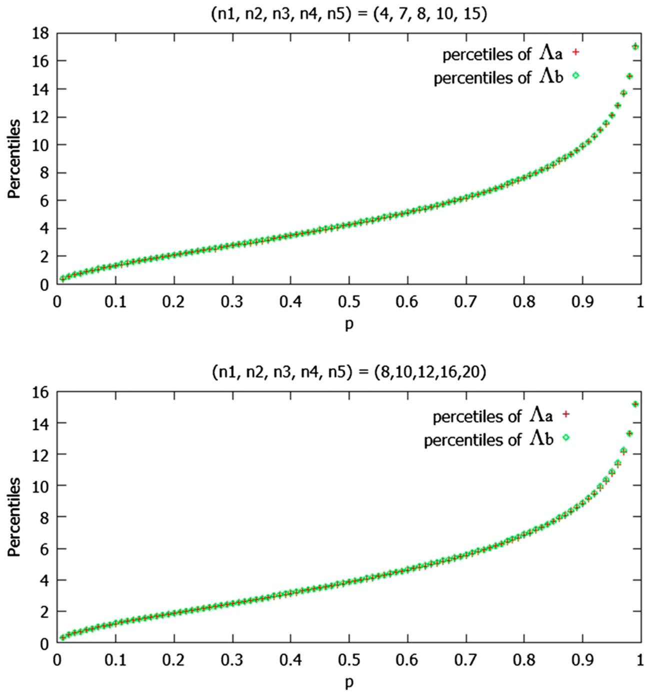

Figure 2. $100p$ percentiles of the likelihood ratio test statistics Λ_a and Λ_b; percentiles of Λ_a when $a_1 = \ldots = a_5 = 3, b_1 = 2, b_2 = 4, b_3 = 7, b_4 = 8, b_5 = 11$; percentiles of Λ_b when $b_1 = \ldots = b_5 = 5$ and $a_1 = 4, a_2 = 1, a_3 = 5, a_4 = 6, a_5 = 10$

The necessary percentiles or p-values to carry out the test can be estimated by Monte Carlo simulation based on independent samples generated from the gamma$(1, 1)$ distribution, that is, from the exponential distribution with location parameter zero and the scale parameter one, exponential$(0, 1)$. For instance, to estimate the percentile of the null distribution of Λ_a, the LRT statistic (8) is calculated for each set of samples generated from exponential$(0, 1)$, and the $100(1-\alpha)$ percentile of the simulated LRT statistics is an estimate of the $100(1-\alpha)$ percentile of the null distribution of Λ_a.

Remark Instead of generating exponential random numbers and then computing $\bar{X}$ and $\tilde{G}$, we can directly generate $\bar{X} \sim$ gamma$(n, 1)/n$ and generate $\tilde{G}/\bar{X}$ using the following distributional result. It is known that the ratio $\tilde{G}/\bar{X}$ and $\bar{X}$ are independent and

$$\frac{\tilde{G}}{\bar{X}} \sim \left(\prod_{j=1}^{n-1} Y_j\right)^{\frac{1}{n}} \tag{20}$$

where Y_j's are independent with $Y_j \sim$ beta$(a, j/n)$, $j = 1, \ldots, n-1$. This simulation approach is very similar to the one described in the preceding paragraph in terms of speed, and so we do not pursue this approach.

4.2. Null distribution of Λ_b

As noted earlier, the null distribution of Λ_b also depends on parameters only weakly. Furthermore, our simulation studies indicate that the null distribution of Λ_b and that of Λ_a are practically identical, and both distributions depend mainly on sample sizes. To provide some evidence to our claim, Monte Carlo estimates (based on 100,000 runs) of the percentiles of Λ_a and Λ_b when $(n_1, \ldots, n_5) = (4, 7, 8, 10, 15)$ and $(n_1, \ldots, n_5) = (8, 10, 12, 15, 20)$ are plotted in Figure 2. For both sets of sample sizes, the percentiles of Λ_a are estimated when $a_1 = \ldots = a_5 = 3$ and $(b_1, \ldots, b_5) = (2, 4, 7, 8, 11)$, and the percentiles of Λ_b are estimated when $b_1 = \ldots = b_5 = 5$ and $(a_1, \ldots, a_5) = (4, 1, 5, 6, 10)$. Both plots in Figure 2 clearly indicate that the null distributions of Λ_a and Λ_b are practically identical.

We estimated the 90th, 95th, and 99th percentiles of the null distribution of Λ_a (or of Λ_b) for some selected values of $(n_1, \ldots, n_k)$, and $k = 3$ and 5. The estimated values are given in Table 1. The percentiles for large samples are given in the last row and are based on the χ^2_{k-1} distribution. Notice that the percentiles for small samples are larger than the χ^2 percentiles, as a result, the large sample test based on the asymptotic chi-squared distribution could be liberal for small to moderate samples. Specifically, the type I error rates of the large sample tests for the equality of shape parameters or those for the equality of scale parameters could be larger than the nominal level if sample sizes are around 50 or smaller.

4.3. Null distributions of Λ_μ and Λ_E

To show that the distribution of the LRT statistic Λ_μ does not depend on any parameters, we estimated the percentiles of Λ_μ under $H_0 : a_1 b_1 = \ldots = a_k b_k$ using the Monte Carlo simulation with 100,000 runs. For some values of $(n_1, \ldots, n_k)$, $k = 3$, and $k = 5$. The

Table 1. Percentiles of the null distributions of the likelihood ratio test statistic Λ_a and Λ_b

$k=3$				$k=5$			
$(n_1,\ldots,n_k)$	0.90	0.95	0.99	$(n_1,\ldots,n_k)$	0.90	0.95	0.99
(4, 4, 4)	6.95	8.99	13.79	(4 , 4, 4, 4, 4)	11.73	14.23	19.76
(4, 6, 7)	6.32	8.22	12.63	(4 , 7, 8, 10, 15)	9.95	12.09	16.81
(4, 6, 10)	6.30	8.19	12.57	(8 , 10, 12, 16, 20)	8.90	10.87	15.15
(4, 8, 15)	6.16	8.00	12.30	(8 , 12, 16, 22, 30)	8.70	10.67	14.93
(8, 12, 16)	5.31	6.91	10.61	(10, 13, 16, 21, 25)	8.61	10.50	14.64
(8, 12, 20)	5.23	6.77	10.41	(10, 15, 20, 25, 30)	8.52	10.39	14.48
(10, 15, 20)	5.16	6.71	10.35	(10, 17, 21, 27, 30)	8.49	10.38	14.47
(10, 15, 30)	5.15	6.70	10.29	(10, 19, 21, 21, 28)	8.54	10.44	14.47
(15, 20, 30)	4.97	6.47	9.98	(12, 15, 18, 18, 27)	8.54	10.44	14.51
(20, 20, 30)	4.91	6.38	9.81	(15, 17, 20, 25, 30)	8.41	10.29	14.42
(20, 25, 30)	4.89	6.36	9.77	(17, 21, 24, 28, 30)	8.31	10.15	14.17
(25, 25, 30)	4.86	6.34	9.71	(20, 23, 25, 27, 30)	8.23	10.05	14.09
(30, 30, 30)	4.83	6.28	9.66	(17, 22, 27, 29, 30)	8.29	10.12	14.15
(50, 50, 50)	4.76	6.16	9.41	(50, 50, 50, 50, 50)	8.03	9.79	13.76
(70, 70, 70)	4.69	6.10	9.46	(70, 70, 70, 70, 70)	7.96	9.68	13.42
(∞,∞,∞)	4.61	5.99	9.21	$(\infty,\infty,\infty,\infty,\infty)$	7.78	9.49	13.28

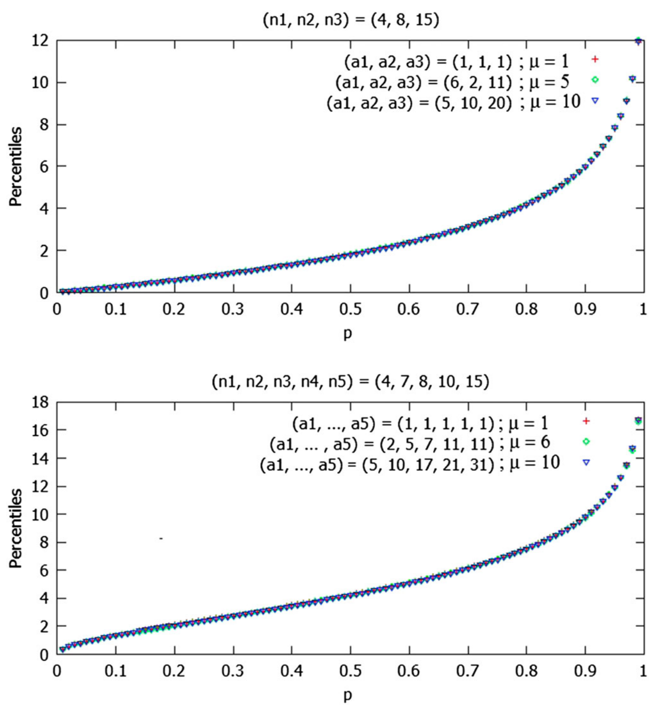

Figure 3. $100p$ percentiles of the likelihood ratio test statistics Λ_μ

estimated percentiles for the cases of $k=3$, $(n_1,n_2,n_3)=(4,8,15)$ and $k=5$, $(n_1,\ldots,n_5)=(4,7,8,10,15)$ are plotted in Figure 3. These two plots indicate that the null distributions are identical for various parameters satisfying $a_1b_1=\ldots=a_kb_k$, and they depend only on the sample sizes. So the p-values of the LRT for equality of means can be estimated by Monte Carlo method as described in Algorithm 1 for the test of equal shape parameters.

For the sake of illustration, we estimated the 90th, 95th, and 99th percentiles of the null distribution of Λ_μ for some selected sample sizes, and $k=3$ and 5. The estimated values are reported in Table 2. We shall use these percentiles to study the size and power properties of the LRT for the equality of means in the following section.

The plots in Figure 4 indicate that the null distribution of the LRT statistic Λ_E is almost free of parameters. As in the preceding problems, the percentiles of Λ_E can be estimated by Monte Carlo simulation. We estimated the percentiles of the LRT statistic Λ_E for testing $H_0:(a_1,b_1)=\ldots=(a_k,b_k)$ and present them in Table 3. Percentiles are given for some selected sample sizes when $k=3$ and 5. The percentiles for large samples are given in the row $(\infty,\ldots,\infty)$, which are the percentiles of the χ^2_{2k-2} distribution. The percentiles for small

Table 2. Percentiles of the null distribution of the likelihood ratio test statistic Λ_{μ}

$k=3$				$k=5$			
$(n_1,\ldots,n_k)$	0.90	0.95	0.99	$(n_1,\ldots,n_k)$	0.90	0.95	0.99
(4, 4, 4)	6.64	8.54	12.87	(4, 4, 4, 4, 4)	11.45	13.87	19.27
(4, 6, 7)	6.13	7.94	12.13	(4, 7, 8, 10, 15)	9.79	11.94	16.66
(4, 8,15)	6.00	7.80	11.99	(8, 12, 16, 22, 30)	8.71	10.59	14.84
(8, 12, 16)	5.28	6.86	10.50	(10, 13, 16, 21, 25)	8.61	10.47	14.69
(8, 12, 20)	5.26	6.87	10.52	(10, 15, 20, 25, 30)	8.48	10.30	14.46
(10, 15, 20)	5.12	6.64	10.24	(10, 17, 21, 27, 30)	8.49	10.32	14.46
(10, 15, 30)	5.14	6.69	10.24	(10, 19, 22, 29, 30)	8.44	10.27	14.35
(15, 20, 30)	4.96	6.47	9.96	(12, 15, 18, 26, 27)	8.46	10.34	14.45
(20, 20, 30)	4.90	6.41	9.89	(15, 17, 20, 25, 30)	8.33	10.12	14.11
(20, 25, 30)	4.84	6.30	9.63	(17, 21, 24, 28, 30)	8.30	10.08	14.08
(25, 25, 30)	4.88	6.32	9.68	(19, 23, 26, 27, 30)	8.26	10.07	13.99
(30, 30, 30)	4.79	6.23	9.71	(22, 24, 25, 28, 30)	8.22	10.02	14.03
(50, 50, 50)	4.74	6.16	9.44	(50, 50, 50, 50, 50)	8.00	9.76	13.65
(70, 70, 70)	4.70	6.13	9.42	(70, 70, 70, 70, 70)	7.94	9.69	13.58
$(\infty,\ldots,\infty)$	4.61	5.99	9.21	$(\infty,\ldots,\infty)$	7.78	9.49	13.28

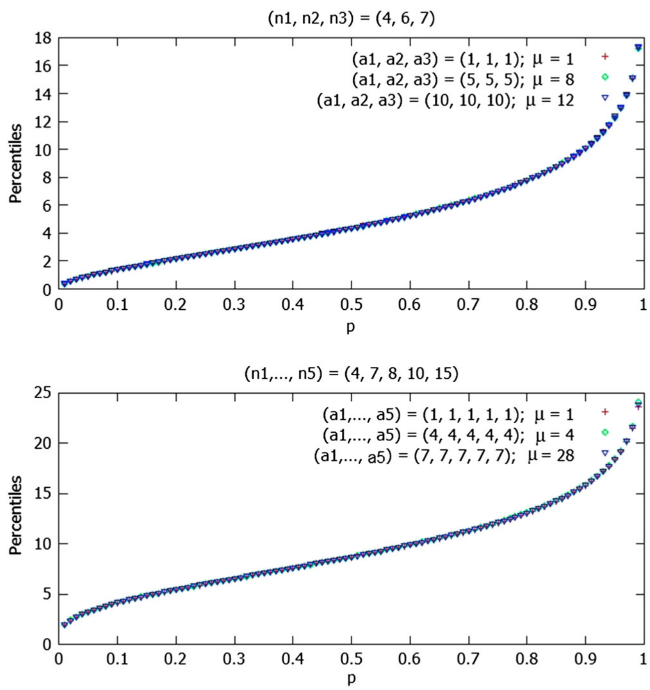

Figure 4. $100p$ percentiles of the null distribution of Λ_E; $b_i = \mu/a_i$, $i = 1,\ldots,k$

samples on the basis of the empirical distribution of Λ_E are smaller than the corresponding large sample approximation. This comparison indicates that if the large sample approximate test is used for small samples, then the type I error rates will be larger than the nominal level.

5. POWER STUDIES AND COMPARISON

Simulation studies in the preceding section clearly indicate that the LRTs based on the empirical null distributions are exact, and so type I error studies are not necessary. Comparison of percentiles of the LRT statistics with those of the large sample chi-squared approximation (Tables 1, 2, and 3) shows that the percentiles based on the chi-squared approximation are smaller than those based on the empirical distributions. This comparison implies that the type I error rates of the large sample tests when applied to small samples could be much larger than the nominal level. We also observe from these tables just cited that the empirical percentiles are close to the chi-squared percentiles for large samples, which indicate that large sample approximate tests and the tests based on the empirical distribution should be similar in terms of size and power for large samples.

Table 3. Percentiles of the null distribution of the likelihood ratio test statistic Λ_E

$k=3$				$k=5$			
$(n_1,\dots,n_k)$	0.90	0.95	0.99	$(n_1,\dots,n_k)$	0.90	0 .95	0.99
(4, 4, 4)	11.09	13.47	18.82	(4, 4, 4, 4, 4)	18.46	21.39	27.70
(4, 6, 7)	10.11	12.34	17.26	(4, 7, 8,10,15)	15.88	18.43	23.89
(4, 8,15)	9.63	11.76	16.50	(8, 12, 16, 22, 30)	14.52	16.84	21.85
(8, 12, 16)	8.77	10.70	14.94	(10, 13, 16, 21, 25)	14.43	16.75	21.73
(8, 12, 20)	8.76	10.66	14.92	(10, 15, 20, 25, 30)	14.29	16.59	21.53
(10, 15, 20)	8.57	10.42	14.63	(10, 17, 21, 27, 30)	14.26	16.58	21.33
(10, 15, 30)	8.48	10.35	14.54	(10, 19, 22, 29, 30)	14.23	16.57	21.63
(15, 20, 30)	8.28	10.07	14.09	(12, 15, 18, 26, 27)	14.24	16.58	21.59
(20, 20, 30)	8.21	10.06	14.19	(15, 17, 20, 25, 30)	14.12	16.37	21.21
(20, 25, 30)	8.22	10.02	14.05	(17, 21, 24, 28, 30)	14.02	16.26	21.00
(25, 25, 30)	8.14	9.95	13.94	(19, 23, 26, 27, 30)	14.03	16.25	21.11
(30, 30, 30)	8.13	9.92	13.95	(22, 24, 25, 28, 30)	13.93	16.16	20.77
(50, 50, 50)	7.95	9.74	13.61	(50, 50, 50, 50, 50)	13.68	15.87	20.56
(70, 70, 70)	7.91	9.69	13.46	(70, 70, 70, 70, 70)	13.61	15.78	20.49
(100, 100, 100)	7.87	9.58	13.43	(100, 100, 100, 100, 100)	13.46	15.63	20.14
(∞,∞,∞)	7.78	9.49	13.28	$(\infty,\dots,\infty)$	13.36	15.51	20.09

Table 4. Powers of the likelihood ratio test and computational approach test (in parentheses) for testing equality of means at the level 0.05

		$k=3;\ a_1=1$			
$(n_1,\dots,n_k)$	$b_1=b_2=b_3$	$a_2=1,a_3=1$	$a_2=1,a_3=2$	$a_2=2,a_3=4$	$a_2=1,a_3=3$
(5, 5, 5)	2	0.050 (0.024)	0.195 (0.098)	0.564 (0.405)	0.758 (0.314)
(5, 5, 5)	3	0.050 (0.032)	0.198 (0.099)	0.565 (0.390)	0.754 (0.345)
(5, 5, 5)	5	0.050 (0.028)	0.197 (0.111)	0.567 (0.404)	0.757 (0.331)
(5, 10, 5)	2	0.050 (0.027)	0.196 (0.098)	0.644 (0.483)	0.755 (0.329)
(5, 10, 5)	3	0.050 (0.036)	0.197 (0.108)	0.644 (0.489)	0.759 (0.313)
(5, 10, 5)	5	0.049 (0.030)	0.195 (0.090)	0.646 (0.492)	0.763 (0.316)
		$b_1=1$			
$(n_1,\dots,n_k)$	$a_1=a_2=a_3$	$b_2=1,b_3=1$	$b_2=1,b_3=2$	$b_2=2,b_3=4$	$b_2=1,b_3=3$
(5, 5, 5)	1	0.050 (0.033)	0.151 (0.073)	0.341 (0.224)	0.333 (0.168)
(5, 5, 5)	2	0.047 (0.037)	0.260 (0.121)	0.604 (0.492)	0.590 (0.348)
(5, 5, 5)	4	0.048 (0.037)	0.468 (0.111)	0.894 (0.815)	0.868 (0.598)
(5, 10, 5)	1	0.050 (0.035)	0.180 (0.103)	0.354 (0.278)	0.390 (0.181)
(5, 10, 5)	2	0.048 (0.034)	0.302 (0.167)	0.634 (0.597)	0.652 (0.382)
(5, 10, 5)	4	0.048 (0.030)	0.524 (0.320)	0.904 (0.888)	0.900 (0.581)

The tests based on a generalized minimum chi-squared procedure (Tripathi *et al.*, 1993) are also valid only for large samples, and they are also not simple to use. Recently, Chang *et al.* (2011) have proposed a test for equality of several gamma means, referred to as the computational approach test (CAT), which is based on the test statistic $\sum_{i=1}^{k}(\ln\widehat{\mu}_i - \overline{\ln\widehat{\mu}})^2$, where $\widehat{\mu}_i$ is the MLE of μ_i, $i=1,\dots,k$. The percentiles (under $H_0:\mu_1=\dots=\mu_k$) of the test statistic are estimated based on simulated samples from gamma$(\widehat{a}_{1c},\widehat{\mu}_c/\widehat{a}_{1c}),\dots,$ gamma$(\widehat{a}_{kc},\widehat{\mu}_c/\widehat{a}_{kc})$, where a_{ic} and μ_c are the constrained MLEs. Thus, the CAT is essentially based on the parametric bootstrap approach in which the samples are generated from gamma distributions using the constrained MLEs as the parameters. The test rejects the null hypothesis in (13) at the level α, if an observed value of the test statistic is larger than the estimated $100(1-\alpha)$ percentile of the $\sum_{i=1}^{k}(\ln\widehat{\mu}_i - \overline{\ln\widehat{\mu}})^2$.

We estimated the powers of the LRT for equality of means based on the empirical distribution and those of the CAT and presented them in Table 4. The powers of the LRT are estimated using simulations consisting of 100,000 runs. The power estimation of the CAT involves two nested "do loops," the inner "do loop" for estimating the percentiles and the outer one for estimating the proportion of times the test rejects the null hypothesis. We used 5000 runs for the outer "do loop" and 2500 for the inner "do loop." The type I error rates and powers are given only for small samples because for large samples these two tests are similar in terms of power. In the first part of Table 4, the powers are given assuming that the scale parameters are equal, and the mean differences are only due to the differences among the shape parameters.

We see in Table 4 that the type I error rates of the CAT are smaller than the nominal level, and as a result, the CAT is less powerful than the LRT. The powers of the LRT are much larger than those of the CAT for some cases. In the second part of the table, the powers are given for the case of common shape parameters, and the differences in the means are due to the differences among the scale parameters. We once again see that the CAT is conservative and is less powerful than the LRT.

6. EXAMPLES

Example 1 Schickedanz and Krause (1970) fitted normal, lognormal, and gamma distributions to weekly rainfall data from Springfield, Illinois, during the seasons of summer, fall, and winter for 1960–1964. They found that the gamma distribution is the best fit among these three distributions. The sample sizes along with MLEs are given for each season in Table 5.

To test the equality of the shape parameters of the rainfall distributions during these three seasons, we calculated the constrained MLEs $\widehat{a}_c = 0.8430$, $\widehat{b}_{1c} = 1.0723$, $\widehat{b}_{2c} = 0.9057$, and $\widehat{b}_{3c} = 0.4370$. The LRT statistic in (8) is 1.2673 with a p-value of 0.540. This p-value clearly indicates that the shape parameters are not significantly different.

To test the equality of the scale parameters of the rainfall distributions, we calculated the constrained MLEs as $\widehat{b}_c = 0.9480$, $\widehat{a}_{1c} = 0.8474$, $\widehat{a}_{2c} = 0.6396$, and $\widehat{a}_{3c} = 0.8345$. The LRT statistic for testing the equality of the scale parameters is 13.46 with a p-value of 0.002, which indicates that the scale parameters of the rainfall distributions are significantly different.

To test the equality of the means, we computed the constrained MLEs as $\widehat{\mu}_c = 0.6694$, $\widehat{a}_{1c} = 0.7421$, $\widehat{a}_{2c} = 0.7556$, and $\widehat{a}_{3c} = 0.8002$. The LRT statistic for testing the equality of the means is 20.19 with a p-value less than 0.001. This p-value provides strong evidence to conclude that the means are quite different.

Finally, to test if these rainfall distributions are the same, we found the MLEs based on all three samples are $\widehat{a} = 0.7741$ and $\widehat{b} = 0.8745$. The LRT statistic in (20) is 20.26 with a p-value less than 0.001.

The R functions that were used to obtain the results for the aforementioned example are posted at www.ucs.louisiana.edu/~kxk4695 and also available at the Environmetrics Web site.

Example 2 The data for this example are taken from Chang *et al.* (2011). In a ground water monitoring study on availability of water to neighboring domestic wells near Bel Air, Harford County, Maryland, the data were collected from various types of wells. The following data sets represent the well yields (in gallon per minute per foot) based on four topographic settings: flood plain, hilltop, hillside, and upland. As noted by Chang *et al.*, the data sets are heavily unbalanced with data sets from hillside and upland draw being small. The data sets from flood plain and hilltop seem to satisfy the assumption of gamma distributions, and so it is reasonable to assume that the other two small data sets also fit gamma models. For the four data sets given in Table 9 of Chang *et al.* (2011), we computed the statistics as in Table 6.

To test the equality of shape parameters of distributions of well yields, we calculated the LRT statistic in (8) as 1.450 with the p-value of 0.174. Thus, the equality of shape parameters is tenable. The LRT statistic in (12) for testing the equality of scale parameters is 11.196 with a p-value less than 0.001. Thus, the data provide enough evidence to indicate that the scale parameters are unequal.

To test the equality of means, the constrained MLEs are calculated as $\widehat{\mu}_c = 0.7834$, $\widehat{a}_{1c} = 0.4172$, $\widehat{a}_{2c} = 0.4321$, $\widehat{a}_{3c} = 0.3332$, and $\widehat{a}_{4c} = 0.4297$. The LRT statistic in (17) is 13.92 with a p-value 0.026. The p-value of the CAT reported in Chang *et al.* (2011) is 0.1034. Notice that the LRT rejects the null hypothesis of equal means, whereas the CAT does not. The test results are consistent with our power studies that indicated that the LRT is more powerful than the CAT.

Table 5. Weekly rainfall data with sample statistics and maximum likelihood estimates (MLEs)

Sample statistics	Summer $n_1 = 58$	Fall $n_2 = 51$	Winter $n_3 = 57$
$\bar{X}$	0.9040	0.7635	0.3684
$\ln \tilde{G}$	−0.8471	−1.0417	−1.5850
MLEs $(\widehat{a}, \widehat{b})$	(0.7959, 1.1358)	(0.7725, 0.9884)	(0.9860, 0.3736)

Table 6. Statistics and the maximum likelihood estimates (MLEs) for the ground water data sets in Table 9 of Chang *et al.* (2011)

Topographic settings	Sample size	$\bar{X}$	$\log(\tilde{G})$	MLEs $(\widehat{a}, \widehat{b})$
Flood plain	41	0.9501	−0.5979	(0.4218, 2.252)
Hilltop	17	0.2394	−2.4424	(0.6087, 0.3933)
Hillside	4	2.2175	−0.4482	(0.5090, 4.357)
Upland draw	3	0.1900	−2.5155	(0.7062, 0.2691)

7. CONCLUDING REMARKS

We have proposed LRTs based on empirical null distributions for comparing several gamma distributions with respect to scale parameters, shape parameters, or means. We also considered the problem of testing homogeneity of several gamma distributions. On the basis of strong simulation evidence, we have observed that the null distributions of the LRTs depend on parameters only weakly, and such dependence is not noticed in simulation results. These LRTs based on such empirical distributions are exact for practical purposes. Analytical proof for weak dependence appears to be difficult. Using scale invariance of the gamma family, we argued that the null distributions of the LRT statistics do not depend on the scale parameters. However, proving that they barely depend on the shape parameters seems to be difficult.

Even though we have provided a computational algorithm for implementing the LRTs, calculation of the empirical distributions is numerically involved. In order to help readers/practitioners, we have provided the R codes at www.ucs.louisiana.edu/~kxk4695 and have also made them available online at the journal's Web site.

Acknowledgements

The authors are grateful to two reviewers and an associate editor for providing useful comments and suggestions.

REFERENCES

Ashkar F, Ouarda TBMJ. 1998. Approximate confidence intervals for quantiles of gamma and generalized gamma distributions. *Journal of Hydrologic Engineering* **3**:43–51.
Brawn D, Upton G. 2007. Closed-form parameter estimates for a truncated gamma distribution. *Environmetrics* **18**:633–645.
Bhaumik DK, Gibbons RD. 2006. One-sided approximate prediction intervals for at least p of m observations from a gamma population at each of r locations. *Technometrics* **48**:112–119.
Bhaumik DK, Kapur K, Gibbons RD. 2009. Testing parameters of a gamma distribution for small samples. *Technometrics* **51**:326–334.
Chang C-H, Lin J-J, Pal N. 2011. Testing the equality of several gamma means: a parametric bootstrap method with applications. *Computational Statistics* **26**:55–76.
Das SC. 1955. Fitting truncated type III curves to rainfall data. *Australian Journal of Physics* **8**:298–304.
Gibbons RD. 1994. *Statistical Methods for Groundwater Monitoring*. Wiley: New York.
Krishnamoorthy K, Mathew T, Mukherjee S. 2008. Normal based methods for a gamma distribution: prediction and tolerance interval and stress-strength reliability. *Technometrics* **50**:69–78.
Krishnamoorthy K, Luis N. 2014. Small sample inference for gamma distributions: one- and two-sample problems. *Environmetrics* **25**:107–126.
Lawless JF. 2003. *Statistical Models and Methods for Lifetime Data*. Wiley: Hoboken, NJ.
Schickedanz PT, Krause GF. 1970. A test for the scale parameters of two gamma distributions using the generalized likelihood ratio. *Journal of Applied Meteorology* **9**:13–16.
Shiue W-K, Bain LJ. 1983. A two-sample test of equal gamma distribution scale parameters with unknown shape parameters. *Technometrics* **25**:377–381.
Shiue WK, Bain LJ, Engelhardt M. 1988. Test of equal gamma-distribution means with unknown and unequal shape parameters. *Technometrics* **30**:169–174.
Stephenson DB, Kumar KR, Doblas-Reyes FJ, Royer JF, Chauvin E, Pezzulli S. 1999. Extreme daily rainfall events and their impact on ensemble forecasts of the Indian monsoon. *Monthly Weather Review* **127**:1954–1966.
Thiagarajah K. 2013. Testing homogeneity of gamma populations. *Journal of Statistics & Management Systems* **16**:433–444.
Tripathi RC, Gupta RC, Pair RK. 1993. Statistical tests involving several independent gamma distributions. *Annals of the Institute of Mathematical Statistics* **45**:777–786.

APPENDIX A. CALCULATION OF THE CONSTRAINED MAXIMUM LIKELIHOOD ESTIMATES FOR TESTING THE EQUALITY OF SCALE PARAMETERS

The constrained MLEs of the shape parameters when $H_0 : b_1 = \ldots = b_k$ are the roots of the following equations

$$f_i(a_1, \ldots, a_k) = n_i \left[\ln \left(\sum_{i=1}^{k} n_i a_i \right) - \psi(a_i) - \ln \frac{\sum_{i=1}^{k} n_i \bar{X}_i}{\tilde{G}_i} \right] = 0, \quad i = 1, \ldots, k \tag{A.1}$$

The partial derivative

$$f_{ii}(a_1, \ldots, a_k) = \frac{\partial f_i}{\partial a_i} = \frac{n_i^2}{\sum_{j=1}^{k} n_j a_j} - n_i \psi'(a_i), \quad i = 1, \ldots, k$$

and

$$f_{ij} = \frac{\partial f_i}{\partial a_j} = f_{ji} = \frac{\partial f_j}{\partial a_i} = \frac{n_i n_j}{\sum_{i=1}^{k} a_i n_i}, \quad i \neq j$$

Letting $\boldsymbol{F} = (f_{ij})$, we obtain the following Newton–Raphson iterative scheme:

$$\begin{pmatrix} a_{1,j+1} \\ \vdots \\ a_{k,j+1} \end{pmatrix} = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{kj} \end{pmatrix} - (\boldsymbol{F})^{-1} \begin{pmatrix} f_1(a_{1j}, \ldots, a_{kj}) \\ \vdots \\ f_k(a_{1j}, \ldots, a_{kj}) \end{pmatrix}, \quad j = 0, 1, 2, \ldots$$

The roots can be obtained using the aforementioned iterative scheme with the starting value

$$a_{i0} \simeq \frac{3 - \tilde{s}_i + \sqrt{(\tilde{s}_i - 3)^2 + 24\tilde{s}_i}}{12\tilde{s}_i}$$

where $\tilde{s}_i = \ln[\sum_{j=1}^{k} n_j \bar{X}_j / \tilde{G}_i]$, $i = 1, \ldots, k$

APPENDIX B. CALCULATION OF THE CONSTRAINED MAXIMUM LIKELIHOOD ESTIMATES FOR TESTING THE EQUALITY OF MEANS

The constrained MLEs of the a_i's when $\mu_1 = \ldots = \mu_k$ are the solutions of the equations

$$g_i(a_1, \ldots, a_k) = n_i \left(\ln a_i - \psi(a_i) - \ln \frac{\mu}{\bar{X}_i} - \frac{\bar{X}_i}{\mu} + 1 \right) = 0, \quad i = 1, \ldots, k$$

where $\mu = \frac{\sum_{i=1}^{k} n_i a_i \bar{X}_i}{\sum_{i=1}^{k} n_i a_i}$. Let $g_{ii} = \frac{\partial g_i}{\partial a_i}$, $g_{ij} = \frac{\partial g_i}{\partial a_j} = g_{ji} = \frac{\partial g_j}{\partial a_i}$. Then, it is not difficult to check that

$$g_{ii} = \frac{n_i}{a_i} - n_i \psi'(a_i) + \frac{n_i^2 (\bar{X}_i - \mu)^2}{\mu^2 \sum_{l=1}^{k} n_l a_l}, \quad i = 1, \ldots, k$$

and

$$g_{ij} = g_{ji} = \frac{n_i n_j (\bar{X}_j - \mu)(\bar{X}_i - \mu)}{\mu^2 \sum_{l=1}^{k} n_l a_l}, \ i \neq j$$

where ψ' is the trigamma function. Let $\mathbf{G} = (g_{ij})$. In terms of the aforementioned derivatives, the Newton–Raphson scheme is

$$\begin{pmatrix} a_{1,l+1} \\ \vdots \\ a_{k,l+1} \end{pmatrix} \leftarrow \begin{pmatrix} a_{1l} \\ \vdots \\ a_{kl} \end{pmatrix} - \begin{pmatrix} g_{11}^l & g_{12}^l & \cdots & g_{1k}^l \\ g_{21}^l & g_{22}^l & \cdots & g_{2k}^l \\ \vdots & \vdots & \vdots & \vdots \\ g_{k1}^l & g_{k2}^l & \cdots & g_{kk}^l \end{pmatrix}^{-1} \begin{pmatrix} g_1(a_{1l}, \ldots, a_{kl}) \\ g_2(a_{1l}, \ldots, a_{kl}) \\ \vdots \\ g_k(a_{1l}, \ldots, a_{kl}) \end{pmatrix}, \quad l = 0, 1, 2, \ldots \tag{B.1}$$

where $g_{ij}^l = g_{ij}(a_{1l}, \ldots, a_{kl})$. We chose initial value $\mu_0 = \sum_{i=1}^{k} n_i \bar{X}_i / \sum_{i=1}^{k} n_i$ for μ in the aforementioned iterative scheme, and chose the initial values

$$a_{i0} = \frac{3 - s_i^* + \sqrt{(s_i^* - 3)^2 + 24 s_i^*}}{12 s_i^*}, \ i = 1, \ldots, k$$

where

$$s_i^* = \ln \mu_0 - \ln \tilde{G}_i + \bar{X}_i / \mu_0 - 1, \ i = 1, \ldots, k$$

An alternative iterative scheme that avoids calculation of the inverse of a matrix is to write the iterative scheme (B.1) as follows: Let $\mathbf{a} = (a_1, \ldots, a_k)'$, $\mathbf{a}_0 = (a_{10}, \ldots, a_{k0})'$, $\mathbf{G}^0 = (g_{ij}^0)$ and

$$\mathbf{g}_0 = (g_1(a_{10}, \ldots, a_{k0}), \ldots, g_k(a_{10}, \ldots, g(a_{k0}))'$$

By pre-multiplying both sides of (B.1) by $\mathbf{G}^0$, we obtain

$$\begin{aligned} \mathbf{G}^0 \mathbf{a}_1 &= \mathbf{G}^0 \mathbf{a}_0 - \mathbf{g}^0 \\ &= \mathbf{b}^0, \text{ say.} \end{aligned} \tag{B.2}$$

For a given $\mathbf{a}_0$, $\mathbf{b}^0$ can be computed, and the linear system $\mathbf{G}^0\mathbf{a}_1 = \mathbf{b}^0$ can be solved to find $\mathbf{a}_1$. For example, R function, 'solve(A,b)' can be used to find $\mathbf{a}_1$. This process should be continued until $||\mathbf{a}_1 - \mathbf{a}_0||$ becomes small. The iterative process based on (B.2) converges slightly faster than the one based on (B.1).

In our simulation studies, we noticed that both iterative processes produce negative values for $\mathbf{a}_1$, for very small sample sizes. These negative values cause computational issues because the function g_i involves natural logarithm of a_i's. In such situations, we can use the following simple iterative scheme:

1. For a given set of k samples, calculate the MLEs $\widehat{a}_i$'s and $\widehat{b}_i$'s.
2. Calculate $\mu_0 = \sum_{i=1}^{k} n_i \widehat{a}_i \bar{X}_i / \sum_{i=1}^{k} n_i \widehat{a}_i$, and set $s_i^* = \ln \mu_0 - \ln \tilde{G}_i + \bar{X} - i/\mu_0 - 1, i = 1, \ldots, k$.
3. Calculate $a_{i0} = \frac{3 - s_i^* + \sqrt{(s_i^* - 3)^2 + 24 s_i^*}}{12 s_i^*}$, $i = 1, \ldots, k$.
4. Recalculate $\mu_0 = \sum_{i=1}^{k} n_i \widehat{a}_{i0} \bar{X}_i / \sum_{i=1}^{k} n_i \widehat{a}_{i0}$, and $s_i^* = \ln \mu_0 - \ln \tilde{G}_i + \bar{X} - i/\mu_0 - 1, i = 1, \ldots, k$.
5. Recalculate a_{i0}'s using s_i^* in step 4.

The μ_0 in step 4 and the a_{i0}'s in step 5 are approximate constrained MLEs for the common unknown mean μ under H_0 and the shape parameters a_i's, respectively. Comparison of these approximate MLEs with the exact ones indicates that the approximation is quite satisfactory.