Ann. Inst. Statist. Math.
Vol. 43, No. 3, 539–550 (1991)

CONSTRUCTING ELEMENTARY PROCEDURES FOR INFERENCE OF THE GAMMA DISTRIBUTION

TAKEMI YANAGIMOTO[1] AND EIJI YAMAMOTO[2]

[1] *The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106, Japan*
[2] *Department of Applied Mathematics, Okayama University of Science, Okayama 700, Japan*

(Received March 28, 1989; revised September 11, 1990)

Abstract. The aim of the present paper is to construct a series of estimators and tests in the one and the two sample problems in the gamma distribution through the Kullback-Leibler loss. Some of them are newly introduced here. When the approach is applied to the case of the normal distribution, the well known estimators and tests are derived. It is found that the conditional maximum likelihood estimator of the dispersion parameter plays a key role.

Key words and phrases: Chi-square test, conditional inference, exponential family, F-test, Kullback-Leibler loss.

1. Introduction

In practical applications data are often positive valued, and our main interest is placed on inference of the mean and the dispersion of a population. When a set of positive data can be assumed to be an approximate sample from a normal population, the widely employed standard methods based on the normality assumption can be applied. However, when the assumption is questionable, for example, when the coefficient of variation is not small, we need an alternative method. In this paper we discuss the use of the gamma distribution. The gamma distribution has favorable properties. The most desirable property among them is that the sample mean is the maximum likelihood estimator of the mean, and the maximum entropy property is also attractive (Kagan *et al.* (1973)).

The gamma distribution, $Ga(\mu, \theta)$, has the density function,

$$p(x;\ \mu, \theta) = \frac{1}{\Gamma(1/\theta)} \frac{x^{1/\theta - 1}}{(\mu\theta)^{1/\theta}} e^{-x/\mu\theta}, \tag{1.1}$$

where the parameters μ and θ represent the mean and the squared coefficient of variation, and they are orthogonal (Cox and Reid (1987)), that is, Fisher's information matrix is diagonal. This parametrization corresponds to the exponential dispersion model in Jorgensen (1987), and it is convenient for the numerical calculation of estimators (Yanagimoto (1988)). This parametrization is different from the conventional one, $\gamma = 1/\theta$ and $\lambda = \mu\theta$. Because of the popularity of this parametrization in terms of γ we occasionally use the parameter γ in addition to θ for the readers' convenience. Let a random variable X have the distribution $Ga(\mu, \theta)$, and assume that θ is small. Then the distribution of X/μ is well approximated by the normal distribution with mean 1 and variance θ. Such a normal approximation is improved by the Wilson and Hilferty approximation, $(X/\mu)^{1/3}$. Thus the normal distribution is an alternative candidate, when θ is fairly small.

In spite of much work on inference of γ and $\lambda = \mu\theta$ only little attention has been paid to elementary procedures for inference of μ in the gamma population. Grice and Bain (1980), Shiue and Bain (1983) and Jensen (1986) presented tests for $\mu = \mu_0$ in the one sample problems and those for $\mu_1 = \mu_2$ in the two sample problems. Inference of θ was developed in relation to that of the variance in the normal distribution. The likelihood ratio test based on the conditional likelihood was proposed as the uniformly most powerful similar test (Shorack (1972)). On the other hand, the familiar estimator of θ is the maximum likelihood estimator. However, Yanagimoto (1988) claimed superiority of the conditional maximum likelihood estimator over the (unconditional) maximum likelihood estimator. These estimators and tests were introduced individually based on different principles. Therefore it is useful to construct a series of elementary procedures for inference of μ and θ in a systematic way. They include the conditional maximum likelihood estimator of θ and the conditional likelihood ratio test of θ, since these estimator and test show good performance.

The aim of this paper is to construct the estimators and test statistics by using the Kullback-Leibler loss so as to satisfy the above requirements. Our main effort is devoted to presenting a test of the mean μ under an unknown θ in the one sample case, and that of the equivalence of means of two gamma populations under an unknown common θ. Extensions to the multisample problems are straightforward. Formal application of our approach to normal and inverse Gaussian populations yields known statistics for the ANOVA and the analysis of reciprocals (Tweedie (1957)).

Useful properties of the Kullback-Leibler loss are reviewed and developed in Section 2. The methods for the one sample problems are proposed in Section 3, and those for the two sample problems in Section 4. In Section 5 the accuracy of the approximation to the critical values employed in the tests is discussed. Some remarks are given in the final section.

2. Kullback-Leibler loss

We explore properties of the Kullback-Leibler loss, which are convenient for constructing elementary procedures for inference of μ and θ in a systematic way. Let $x_1, \ldots, x_n$ be a sample of size n from a population with a density function, $p(x;\ \mu, \theta)$. Consider an estimator $\hat{\mu}$ of the mean μ. The Kullback-Leibler loss is defined by

$$KL_n(\hat{\mu};\ \mu, \theta) = 2\int \log \frac{\prod p(z_i;\ \hat{\mu}, \theta)}{\prod p(z_i;\ \mu, \theta)} \prod p(z_i;\ \hat{\mu}, \theta) \prod dz_i \tag{2.1}$$
$$(= nKL(\hat{\mu};\ \mu, \theta)).$$

The multiplier 2 is attached for ease of comparison with the log likelihood ratio test statistic. This loss is equivalent with the minimum discrimination information statistic in Kullback (1959). When the density function is the gamma in (1.1), it is expressed explicitly as

$$KL_n(\hat{\mu};\ \mu,\ \theta) = \frac{2n}{\theta}\left(\frac{\hat{\mu}}{\mu} - 1 - \log\frac{\hat{\mu}}{\mu}\right). \tag{2.2}$$

This loss appears in Brown (1968) as a typical one for an estimator of the mean of a positive distribution. It is known that $KL_n(\hat{\mu};\ \mu,\ \theta)$ with the maximum likelihood estimator $\hat{\mu}$ for a distribution of the exponential family is equal to twice the log likelihood ratio test statistic, that is,

$$KL_n(\hat{\mu};\ \mu,\ \theta) = 2\log\frac{\prod p(x_i;\ \hat{\mu},\ \theta)}{\prod p(x_i;\ \mu,\ \theta)}. \tag{2.3}$$

This correspondence permits us the likelihood inference interpretation of some of the proposed procedures.

This loss has many favorable properties like the loss of standardized squared difference in the normal distribution. The most important one is

$$\begin{aligned} \sum KL(x_i;\ \mu,\ \theta) &= nKL(\bar{x};\ \mu,\ \theta) + \sum_i KL(x_i;\ \bar{x},\ \theta) \\ &= \frac{2n}{\theta}\left(\frac{\bar{x}}{\mu} - 1 - \log\frac{\bar{x}}{\mu}\right) + \sum\frac{2}{\theta}\left(\frac{x_i}{\bar{x}} - 1 - \log\frac{x_i}{\bar{x}}\right) \\ &= \frac{2n}{\theta}\left(\frac{\bar{x}}{\mu} - 1 - \log\frac{\bar{x}}{\mu}\right) + \frac{2n}{\theta}\log\frac{\bar{x}}{\tilde{x}}, \end{aligned} \tag{2.4}$$

with the geometric mean $\tilde{x}$. Recall that $\bar{x}$ and $\bar{x}/\tilde{x}$ are independent.

For ease of notations we write (2.4) as $TKL(\mu,\ \theta) = AKL(\mu,\ \theta) + RKL(\theta)$, interpreted as, that the total variation is the average variation plus the residual. This orthogonal decomposition becomes equal to that of deviance in the generalized linear model (McCullagh and Nelder (1989)) by disregarding the dispersion constant θ. Recent developments of the decomposition in the case of the reproductive exponential family can be seen in Jorgensen (1987). We will be careful with the role of each term in (2.4). Consequently, the proposed procedures are slightly different from those based on the deviance such as the analysis of deviance.

The above orthogonal decomposition of this loss holds for other distributions such as the normal and the inverse Gaussian distributions. For the normal distribution $N(\mu,\ \sigma^2)$ with mean μ and variance σ^2 and the inverse Gaussian distribution $IG(\mu,\ \theta)$ with the density function of $\sqrt{1/2\pi\theta x^3}\exp(-(x-\mu)^2/2\theta\mu^2 x)$, the explicit forms of the Kullback-Leibler loss are $n(\hat{\mu}-\mu)^2/\sigma^2$ and $n(\hat{\mu}/\mu+\mu/\hat{\mu}-2)/\theta\mu$, respectively. In these distributions there are no difficulties in constructing elementary procedures for inference of the mean and the dispersion parameters, because each term appearing in the decomposition corresponding to (2.4) has exactly the chi-square distribution.

In the gamma distribution, however, each term has an approximate chi-square distribution. Thus we construct elementary procedures by using appropriate chi-squared approximations if necessary.

3. Proposed procedures—one sample problems

All the estimators and the test statistics given here are derived through the Kullback-Leibler loss. It should be emphasized here that the proposed procedures correspond exactly with the standard elementary ones in the normal population. In fact the latter ones are derived by replacing the Kullback-Leibler loss of the gamma distribution by that of the normal distribution. Furthermore, the distributions of the estimators or test statistics in corresponding problems are similar to each other. All the procedures are summarized in Table 1.

In this section we concentrate on the derivation of the procedures, their optimality and other interpretations of them. We use approximations of which accuracies will be discussed in Section 5. The significance level, α, is assumed to be 0.1, 0.05 or 0.01. The first value is popular in the reliability theory.

Problem 1. *Estimation of* μ. The parameter μ is estimated so as to minimize $TKL(\mu,\theta) = \sum 2n(x_i/\mu - 1 - \log x_i/\mu)/\theta$, or equivalently to minimize $AKL(\mu,\theta)$. This yields the estimator $\hat{\mu} = \bar{x}$, which is free from θ. It is also the maximum likelihood estimator.

Problem 2. *Estimation of* θ. The parameter θ is estimated by solving the equation $RKL(\theta) = E(RKL(\theta)) = 2n\{\xi(\theta) - \xi(\theta/n)\}/\theta$, where $\xi(\theta) = -\psi(1/\theta) - \log\theta$ with $\psi(\cdot)$ being the digamma function. The similar idea was found in McCullagh ((1983), p. 63), though it was not developed widely enough. This estimator is the conditional maximum likelihood estimator given the sample mean $\bar{x}$. An optimality of the conditional maximum likelihood estimator is given in Godambe (1980), and superiority of it over the (unconditional) maximum likelihood estimator is discussed in Yanagimoto (1988).

Problem 3. *Test for* H_0: $\mu = \mu_0$ *against* $\mu \neq \mu_0$ *when* θ *is known.* The rejection region is given by $AKL(\mu_0,\theta) = 2n\{\bar{x}/\mu_0 - 1 - \log\bar{x}/\mu_0\}/\theta > c_\alpha$ for a suitable value c_α. This test is the likelihood ratio test. The critical value is approximated by $c_\alpha \fallingdotseq \{2n\xi(\theta/n)/\theta\}\cdot\chi_1^2(1-\alpha)$ with $\chi_1^2(1-\alpha)$ being the $(1-\alpha)$-th point of χ_1^2, if n/θ is greater than 1 for $\alpha = 0.1$ or 0.05, and 3 for $\alpha = 0.1$, 0.05 or 0.01. Recall that it is easy to calculate the exact critical value, if it is desired.

Problem 4. *Test for* H_0: $\mu = \mu_0$ *against* H_1: $\mu \neq \mu_0$ *when* θ *is unknown.* The rejection region is given by $AKL(\mu_0,\hat{\theta}) = 2n\{\bar{x}/\mu_0 - 1 - \log\bar{x}/\mu_0\}/\hat{\theta} > c_\alpha(\theta)$ with $\hat{\theta}$ being the estimator of θ in Problem 2, that is, the conditional maximum likelihood estimator. In the gamma case, different from the normal case, the critical value $c_\alpha(\theta)$ depends on the unknown parameter θ. The value is approximated by $c_\alpha(\theta) \fallingdotseq (2n/\theta)\xi(\theta/n)t^2_{f(\theta)}(1-\alpha/2)$, where the adjusted degrees of freedom $f(\theta)$ is

$$f(\theta) = 2n(\xi'(\theta) - \xi'(\theta/n)/n). \tag{3.1}$$

The value still depends on θ. A method for determining a critical value is to replace θ by $\hat{\theta}$. Another practical method is to select an appropriately fixed value of θ.

Table 1. Comparisons of the proposed procedures in the gamma model and the corresponding standard ones in the normal model.

Problem[1]	Gamma model		Normal model	
	Statistic	Distribution[2]	Statistic[3]	Distribution
1 Estimation of μ	$\bar{x}$	$Ga(\mu,\ \theta/n)$	$\bar{x}$	$N(\mu,\ \sigma^2/n)$
2 Estimation of θ	$\xi(\hat{\theta}) - \xi(\hat{\theta}/n) = \log \bar{x}/\tilde{x}$	$\theta c\chi_f^2$	$\hat{\sigma}^2 = s^2$	$\sigma^2\chi_{n-1}^2/(n-1)$
3 Test for $\mu = \mu_0$, θ known	$\frac{2n}{\theta}\left(\frac{\bar{x}}{\mu_0} - 1 - \log\frac{\bar{x}}{\mu_0}\right)$	$c\chi_1^2$	$n(\bar{x}-\mu_0)^2/\sigma^2$	χ_1^2
4 Test for $\mu = \mu_0$, θ unknown	$\frac{2n}{\hat{\theta}}\left(\frac{\bar{x}}{\mu_0} - 1 - \log\frac{\bar{x}}{\mu_0}\right)$	ct_f^2	$n(\bar{x}-\mu_0)^2/\hat{\sigma}^2$	t_{n-1}^2
5 Test for $\theta = \theta_0$	$\frac{2n}{\theta_0}\log \bar{x}/\tilde{x}$	$c\chi_f^2$	$(n-1)s^2/\sigma_0^2$	χ_{n-1}^2
6 Estimation of $\mu = \mu_1 = \mu_2$	$\bar{z} = (n\bar{x} + m\bar{y})/(n+m)$	$Ga(\mu,\ \theta/(n+m))$	$(n\bar{x} + m\bar{y})/(n+m)$	$N(\mu,\ \sigma^2/(n+m))$
7 Estimation of $\theta = \theta_1 = \theta_2$	$(n+m)\xi(\hat{\theta}) - n\xi(\hat{\theta}/n) - m\xi(\hat{\theta}/m)$ $= n\log\bar{x}/\tilde{x} + m\log\bar{y}/\tilde{y}$	$\theta c\chi_f^2$	$\hat{\sigma}^2 = \{(n-1)s_x^2 + (m-1)s_y^2\}$ $/(n+m-2)$	$\sigma^2\chi_{n+m-2}^2/(n+m-2)$
8 Test for $\mu_1 = \mu_2$, θ known	$2(n\log\bar{x}/\bar{z} + m\log\bar{y}/\bar{z})/\theta$	$c\chi_1^2$	$(\bar{x}-\bar{y})^2/(1/n+1/m)\sigma^2$	χ_1^2
9 Test for $\mu_1 = \mu_2$, θ unknown	$2(n\log\bar{x}/\bar{z} + m\log\bar{y}/\bar{z})/\hat{\theta}$	ct_f^2	$(\bar{x}-\bar{y})^2/(1/n+1/m)\hat{\sigma}^2$	t_{n+m-2}^2
10 Test for $\theta_1 = \theta_2$	$\frac{(n\log\bar{x}/\tilde{x})/(n-1)}{(m\log\bar{y}/\tilde{y})/(m-1)}$	cF_{f_1,f_2}	s_x^2/s_y^2	$F_{n-1,m-1}$

[1] In the case of the normal model θ is read as σ^2.

[2] The distribution is the approximated one except for the cases of 1 and 6.

[3] $s^2 = \Sigma(x_i - \bar{x})^2/(n-1)$.

When θ is expected to be small, we set $f(\theta) \doteqdot f(0) = n-1$. Such a conventional guess of θ looks useful in practice. As discussed later, $t^2_{f(\theta)}(1-\alpha)$ is not much sensitive to the variation of $f(\theta)$, and $f(\theta)$ changes slowly with respect to θ.

This test is a newly introduced one. The test statistic is equal to $2\log\{\prod p(x_i;\ \hat{\mu},\ \hat{\theta})/\prod p(x_i;\ \mu_0,\ \hat{\theta})\}$. Note that the rejection region based on the usual likelihood ratio is different from ours, though both the rejection regions are common in the normal population.

This test may be called the t^2-test of the mean in the gamma population to emphasize the relation with the t test of the mean in the normal population. A similar test was proposed in Jensen (1986) and a one-sided test was proposed in Grice and Bain (1980), which will be discussed briefly in the final section.

Problem 5. *Test for* H_0: $\theta = \theta_0$ *against* $\theta > \theta_0$. The rejection region is given by $RKL(\theta_0) = 2n(\log\bar{x}/\tilde{x})/\theta_0 > c_\alpha$ for a suitable value c_α. The test is the uniformly most powerful similar test (Shorack (1972)). Glaser (1976*a*) presented a method for calculating the exact value of c_α, but it requires a large amount of computation, and a special program is necessary. Using the first and second moments of $RKL(\theta_0)$, Bain and Engelhardt (1975) gave an approximate value of c_α as $(2n/\theta_0)\{\xi(\theta_0) - \xi(\theta_0/n)\}\chi^2_f(1-\alpha)/f$ where the adjusted degree of freedom f is

$$f = \frac{2n\{\xi(\theta_0) - \xi(\theta_0/n)\}^2}{\theta_0^2(\xi'(\theta_0) - \xi'(\theta_0/n)/n)}. \tag{3.2}$$

This approximation is sufficiently accurate if $n \geq 10$ and $\theta_0 \leq 2$.

4. Proposed procedures—two sample problems

Prior to presenting procedures for the two sample problems, we give the orthogonal decomposition of the loss function corresponding to (2.4). Let $x_1, \ldots, x_n$ and $y_1, \ldots, y_m$ be samples of sizes n and m from the gamma populations having $p(x;\ \mu_1,\ \theta)$ and $p(y;\ \mu_2,\ \theta)$, respectively. Then it holds that

$$\begin{aligned}
&\sum KL(x_i;\ \mu,\ \theta) + \sum KL(y_i;\ \mu,\ \theta) \\
&\quad = (n+m)KL(\bar{z};\ \mu,\ \theta) + \{nKL(\bar{x};\ \bar{z},\ \theta) + mKL(\bar{y};\ \bar{z},\ \theta)\} \\
&\qquad + \left\{\sum KL(x_i;\ \bar{x},\ \theta) + \sum KL(y_i;\ \bar{y},\ \theta)\right\} \\
&\quad = \frac{2(n+m)}{\theta}\left(\frac{\bar{z}}{\mu} - 1 - \log\frac{\bar{z}}{\mu}\right) + \frac{2}{\theta}\left(n\log\frac{\bar{x}}{\bar{z}} + m\log\frac{\bar{y}}{\bar{z}}\right) \\
&\qquad + \frac{2}{\theta}\left(n\log\frac{\bar{x}}{\tilde{x}} + m\log\frac{\bar{y}}{\tilde{y}}\right),
\end{aligned} \tag{4.1}$$

with $\bar{z} = (n\bar{x} + m\bar{y})/(n+m)$. We subsequently write (4.1) as $TKL(\mu,\ \theta) = AKL(\mu,\ \theta) + BKL(\theta) + RKL(\theta)$; the third term on the right-hand side is also written as $RKL_X(\theta) + RKL_Y(\theta)$. From the analogy with the usual ANOVA theory these terms are regarded as the total, average, between samples and residual variations. These four terms of the right-hand side are mutually independent. This decomposition can be extended straightforwardly to the general k sample problems, but we will not pursue the problem any further.

Problem 6. *Estimation of* $\mu = \mu_1 = \mu_2$. The parameter μ, under the assumption that $\mu = \mu_1 = \mu_2$, is estimated so as to minimize $TKL(\mu, \theta)$, or equivalently $AKL(\mu, \theta)$, which yields $\hat{\mu} = \bar{z}$. This is the maximum likelihood estimator.

Problem 7. *Estimation of* θ. The estimator of θ, which is common to both the populations, is given by $RKL(\theta) = E(RKL(\theta)) = [2n\{\xi(\theta) - \xi(\theta/n)\} + 2m\{\xi(\theta) - \xi(\theta/m)\}]/\theta$. This estimator is the conditional maximum likelihood estimator given the sample means, $\bar{x}$ and $\bar{y}$.

Problem 8. *Test for* H_0: $\mu_1 = \mu_2$ *against* H_1: $\mu_1 \neq \mu_2$ *when* θ *is known.* The rejection region is given by $BKL(\theta) > c_\alpha$. This is the likelihood ratio test. The critical value is approximated by $c_\alpha \fallingdotseq b(\theta;\ n, m)\chi_1^2(1-\alpha)$ with $b(\theta;\ n, m) = (2/\theta)\{n\xi(\theta/n) + m\xi(\theta/m) - (n+m)\xi(\theta/(n+m))\}$, if both n/θ and m/θ are greater than 1.

Problem 9. *Test for* H_0: $\mu_1 = \mu_2$ *against* H_1: $\mu_1 \neq \mu_2$ *when* θ *is unknown.* The rejection region is given by $BKL(\hat{\theta}) > c_\alpha(\theta)$. The critical value is approximated by $c_\alpha(\theta) \fallingdotseq b(\theta;\ n, m)t^2_{f(\theta)}(1-\alpha/2)$ where $f(\theta) = k_n(\theta) + k_m(\theta)$ with $k_n(\theta)$ being the right-hand side of (3.1). Again as in Problem 4, $c_\alpha(\theta)$ depends on θ. We can apply the same treatments to an unknown θ as those in Problem 4.

This test statistic is derived as a type of the likelihood ratio test as in Problem 4. The two sample test for the equivalence of the means is most important in practice. This test may be called the two sample t^2-test in the gamma population. A similar test was proposed in Jensen (1986), and another test was proposed in Shiue and Bain (1983).

Problem 10. *Test for* H_0: $\theta_1 = \theta_2$ *against* H_1: $\theta_1 > \theta_2$. In this problem we assume that the two samples come from the gamma populations having $Ga(\mu_1, \theta_1)$ and $Ga(\mu_2, \theta_2)$. The rejection region is given by $\{RKL_X(\theta)/(n-1)\}/\{RKL_Y(\theta)/(m-1)\} > c_\alpha(\theta)$ for a suitable value $c_\alpha(\theta)$. This test is equivalent to the conditional likelihood ratio statistic given the sample means, $\bar{x}$ and $\bar{y}$, and is the uniformly most powerful similar test. The critical value is approximated by $c_\alpha(\theta) \fallingdotseq F_{f_1(\theta), f_2(\theta)}(1-\alpha)$, where the adjusted degrees of freedom $f_1(\theta)$ is given by

$$f_1(\theta) = \frac{2n\{\xi(\theta) - \xi(\theta/n)\}^2}{\theta^2(\xi'(\theta) - \xi'(\theta/n)/n)}, \tag{4.2}$$

and $f_2(\theta)$ is given by replacing n by m in (4.2). We either replace θ by $\hat{\theta}$, or by a preassign value, say $\theta = 0$, as in Problem 4. The first procedure is recommended by Shiue *et al.* (1988).

5. The accuracy of approximations

In the previous two sections we used various approximations to obtain critical values of the test problems. We will discuss the accuracy of the approximations of Problems 3, 4, 5 and 10 case by case. It is known that the log likelihood ratio test has asymptotically the chi-square distribution. The accuracy of the approximation is sharply improved by adjusting the first one or two moments with $c\chi_d^2$ for suitable values c and d. These adjustment constants are usually referred to as Bartlett's adjustment constants. In all the cases the accuracy depends on both n and θ, and it becomes higher as either n or $1/\theta$ increases.

In Problem 3 the distribution of $AKL(\mu_0, \theta)$ is a function of $\theta/n = 1/n\gamma$. The proposed approximation is given by adjusting the first moment. Fortunately this approximation is accurate as seen in Table 2. The condition $\theta/n \leq 1$ presents sufficient accuracy for $\alpha = 0.05$, and that $\theta/n \leq .33$ does for $\alpha = 0.01$.

Table 2. The probability $\Pr\{AKL(\mu_0, \theta) > c\chi_1^2(1-\alpha)\}$ with $x_i \sim Ga(\mu, \theta)$, $i = 1, \ldots, n$, and the Bartlett adjustment factor $c = (2n/\theta)\xi(\theta/n)$.

θ/n	α		
	0.1	0.05	0.01
.2	.1000	.0500	.00997
.25	.1000	.0500	.00995
.33	.1000	.0499	.00992
.5	.1001	.0499	.00980
1.	.1001	.0493	.00921
2.	.0994	.0470	.00771

The approximate critical value of Problem 4, or equivalently that of Problem 9, looks less satisfactorily accurate than the others. However, since these problems are most important in practice, it is worthwhile to pursue the actual accuracy. We conducted simulation studies to confirm the accuracy. Prior to presenting the result of the simulation studies we give the backgrounds for yielding the proposed test statistic and approximation. A notable fact is that $\hat{\theta}$ given in Problem 2 only has a small bias, which was suggested in Yanagimoto (1988). We present simulation results on the bias of $\hat{\theta}$ in Table 3, which show that the relative bias is less than 2% for $n \geq 10$. Note that the relative bias of the unconditional maximum likelihood estimator of $1/\theta$ is much larger than that of our estimator of θ. The usual asymptotic theory of $\hat{\theta}$ leads to an approximate variance $2\theta^2/f(\theta)$ with $f(\theta)$ in (3.1). Thus the numerator is approximated as $(2n/\theta)\xi(\theta/n)\chi_1^2$, and the denominator as $\chi_f^2(\theta)$. Taking account of the independence of both the terms, we obtain the approximation suggested in the previous section.

Table 3. Estimated mean and variance of $\hat{\theta}/\theta$ with 10,000 replications. The last column presents the approximation of variance due to the usual asymptotic theory.

n	θ	$E(\hat{\theta}/\theta)$	$V(\hat{\theta}/\theta)$	Approx.
10	.2	.9803	.1934	.2072
	.5	.9827	.1769	.1891
	1.	.9812	.1624	.1686
	2.	.9882	.1450	.1470
20	.2	.9765	.0933	.0984
	.5	.9872	.0841	.0902
	1.	.9947	.0789	.0807
	2.	.9982	.0711	.0706

The approximate distribution still involves an unknown parameter. The degree of freedom, $f(\theta)$, is sensitive to θ, but fortunately the critical value $t^2_{f(\theta)}(1-\alpha/2)$ is not very sensitive to θ, which is illustrated in Fig. 1. The above reasoning of the approximation may not be sufficiently persuasive. To check the accuracy, we conducted simulation studies, using the technique in Grice and Bain (1980). Table 4 presents the result, which shows the satisfactorily accurate approximation when θ is moderate for $\alpha = 0.1$ or 0.05.

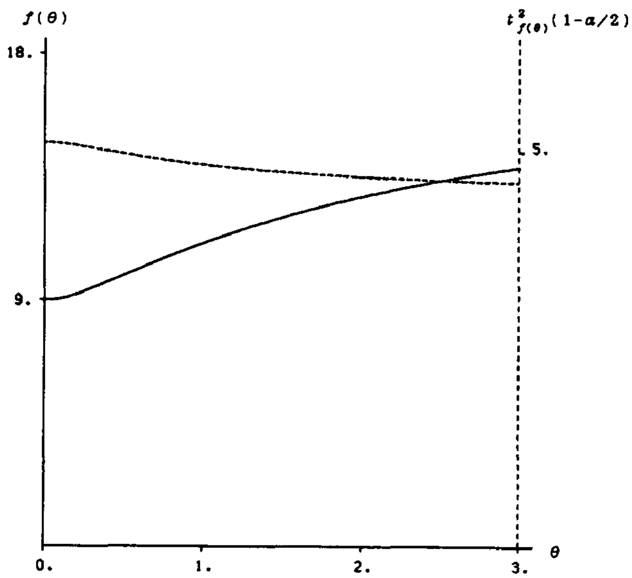

Fig. 1. Behaviors of degrees of freedom $f(\theta)$ (solid) and the critical value $t^2_{f(\theta)}(1-\alpha/2)$ (dotted) in the case of $n = 10$ and $\alpha = 0.05$.

Bain and Engelhardt (1975) discussed the accuracy of the approximation suggested in Problem 5. Although the exact critical value can be computed as in Glaser (1976*a*), the access to the computer program is restricted. Table 4 in Bain and Engelhardt (1975) and our study based on Glaser (1976*b*) show that the conditions $n \geq 3$ and $\theta \leq 2$ assure the satisfactory accuracy for $\alpha = 0.05$ as well as the acceptable one for $\alpha = 0.01$. Though their Table 1 contains adjusted coefficients for small values of $\gamma = 1/\theta$, we guess that the approximation is poor in such situations.

Table 4. The estimated rejection level of the approximated tests by simulations with 40,000 replications for $n = 10$ (the upper) and for $n = 20$ (the lower).

α	θ	Approximation of degrees of freedom: Estimated	Approximation of degrees of freedom: $\theta = 0$
.05	.2	.0506	.0494
		.0505	.0499
	.5	.0512	.0485
		.0508	.0494
	1.	.0517	.0473
		.0507	.0485
	2.	.0522	.0462
		.0506	.0478
.1	.2	.1013	.0999
		.1006	.0998
	.5	.1022	.0991
		.1009	.0994
	1.	.1028	.0979
		.1016	.0993
	2.	.1028	.0966
		.1013	.0987

A useful fact of Problem 10 is that the ratio of $(2n/\theta)\{\xi(\theta) - \xi(\theta/n)\}$ to $(2m/\theta)\{\xi(\theta) - \xi(\theta/m)\}$ is well approximated by $(n-1)/(m-1)$, if $\theta < 2$ and n/m is not largely different from 1. Thus the coefficient is approximately free from the parameter θ. Using this fact together with the approximation in Problem 5, we obtain (4.2).

6. Concluding remarks

The proposed procedures may at first appear to require elaborate computations for obtaining approximate values. However we can apply simple, accurate approximations of the values. Therefore no difficulty results in practice, if a personal computer is available. In terms of the asymptotic formula of the digamma function the function $\xi(\theta)$ is well approximated by $\xi(\theta) \fallingdotseq \theta/2+\theta^2/12-\theta^4/120$, if θ is small, say less than $3/4$. An approximation with wider range of applicability is given by the following,

$$\xi(\theta) \fallingdotseq \theta - \frac{1}{2}\log(1+\theta+0.33\theta^2).$$

The first derivative of $\xi(\theta)$ can be approximated by that of the above approximate functions. As Yanagimoto (1988) shows, an approximation of $\hat{\theta}$ in Problem 2 is given by

$$\hat{\theta} \fallingdotseq \tilde{\xi}((2n+1)z/2(n-1)) - \tilde{\xi}(z/2(n-1)),$$

with z being $\log \bar{x}/\tilde{x}$ and $\tilde{\xi}(z) = z + (1/3)\log(1+3z+5z^2/2)$, which is a good approximation of $\xi^{-1}(z)$.

Grice and Bain (1980) and Shiue and Bain (1983) gave other approximate one-sided tests similar to those of Problems 4 and 9. In the test for $\mu = \mu_0$ under an unknown dispersion parameter discussed in Problem 4, their test is based on the fact that $\bar{x}/\mu_0$ belongs to the gamma distribution with mean 1 and the dispersion parameter θ/n. Their first approximation is $\bar{x}/\mu_0 \sim Ga(1, \hat{\theta}_u/n)$ with the unconditional maximum likelihood estimator $\hat{\theta}_u$. Unfortunately, this approximation is poor as they showed. Therefore they obtained approximate critical values by evaluating the distribution at $\theta = 0$. The accuracy of the approximation looks incidental. Our methods have an advantage, that they are derived in a systematic way. Consequently the use of $\hat{\theta}_c$ is reasonable in relation to the estimation and the test of θ. In contrast no theoretical background is given for the use of $\hat{\theta}_u$ by the above authors.

Using the fact that the conditional distribution of $\bar{x}$ given $t = \bar{x}/\mu_0 - \log \bar{x}/\mu_0 - 1 + \log \bar{x}/\tilde{x}$ is free from θ, Jensen (1986) proposed a similar test for Problem 4. Since the distribution is complicated, he discussed an approximation. Although the numerical computation of the conditional distribution is possible (Glaser (1976*a*, 1976*b*)), the accuracy of the approximation is not studied widely enough. The theoretical background for this test is not strong. In fact, the statistic t contains μ_0, and the rejection region is not presented explicitly. Another defect is that the test is introduced separately from other problems such as the estimation of μ. In fact, it does not seem easy to show whether the sample mean attains the maximum of the conditional likelihood or not. Consequently, further research will be necessary to recommend this similar test for practice, though it looks promising.

Acknowledgements

The authors wish to express their thanks to the two referees, the associate editor and the chief editor for their constructive comments.

References

Bain, L. J. and Engelhardt, M. (1975). A two-moment chi-square approximation for the statistic $\log(\bar{x}/\tilde{x})$, *J. Amer. Statist. Assoc.*, **70**, 948–950.

Brown, L. (1968). Inadmissibility of the usual estimators of scale parameters in the problems with unknown location and scale parameter, *Ann. Math. Statist.*, **39**, 29–48.

Cox, D. R. and Reid, N. (1987). Parameter orthogonality and approximate conditional inference (with discussion), *J. Roy. Statist. Soc. Ser. B*, **49**, 1–39.

Glaser, R. E. (1976*a*). The ratio of the geometric mean to the arithmetic mean for a random sample from a gamma distribution, *J. Amer. Statist. Assoc.*, **71**, 480–487.

Glaser, R. E. (1976*b*). Exact critical values for Bartlett's test for homogeneity of variance, *J. Amer. Statist. Assoc.*, **71**, 488–490.

Godambe, V. P. (1980). A sufficiency and ancillarity in the presence of a nuisance parameter, *Biometrika*, **67**, 155–162.

Grice, J. V. and Bain, L. J. (1980). Inferences concerning the mean of the gamma distribution, *J. Amer. Statist. Assoc.*, **75**, 929–933.

Jensen, J. L. (1986). Inference for the mean of a gamma distribution with unknown shape parameter, *Scand. J. Statist.*, **13**, 135–151.

Jorgensen, B. (1987). Exponential dispersion models (with discussion), *J. Roy. Statist. Soc. Ser. B*, **49**, 127–162.

Kagan, A. M., Linnik, Y. V. and Rao, C. R. (1973). *Characterization Problems in Mathematical Statistics*, Wiley, New York.

Kullback, S. (1959). *Information Theory and Statistics*, Wiley, New York.

McCullagh, P. (1983). Quasi-likelihood functions, *Ann. Statist.*, **11**, 59–67.

McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models*, 2nd ed., Chapman and Hall, London.

Shiue, W.-K. and Bain, L. J. (1983). A two-sample test of equal gamma distribution scale parameters with unknown common shape parameter, *Technometrics*, **25**, 377–381.

Shiue, W.-K., Bain, L. J. and Engelhardt, M. (1988). Test of equal gamma-distribution means with unknown and unequal shape parameters, *Technometrics*, **30**, 169–174.

Shorack, G. R. (1972). The best test of exponentiality against gamma alternatives, *J. Amer. Statist. Assoc.*, **67**, 213–214.

Tweedie, M. C. K. (1957). Statistical properties of inverse Gaussian distributions I, *Ann. Math. Statist.*, **28**, 362–377.

Yanagimoto, T. (1988). The conditional maximum likelihood estimator of the shape parameter in the gamma distribution, *Metrika*, **35**, 161–175.