Communications in Statistics—Simulation and Computation®, 34: 243–252, 2005

ISSN: 0361-0918 print/1532-4141 online
DOI: 10.1081/SAC-200055639

Taylor & Francis
Taylor & Francis Group

Distribution Theory

Comparison Between Two Quantiles: The Normal and Exponential Cases

HUIZHEN GUO[1] AND K. KRISHNAMOORTHY[2]

[1]Department of Mathematics and Computer Science, Xavier University, Cincinnati, Ohio, USA
[2]Department of Mathematics, University of Louisiana at Lafayette, Lafayette, Louisiana, USA

We consider the problem of interval estimation and testing for the difference between the quantiles of two populations. Inferential procedures based on the generalized variable approach are given for the normal and exponential cases. The generalized variable approach produced exact procedures for the normal case with some restrictions on the parameter space and exact procedures for the exponential case. Applications and two illustrative examples are given.

Keywords Generalized limit; Generalized p-value; Reliability parameter; Size; Stress-strength model.

Mathematics Subject Classification 62F03; 62F25.

1. Introduction

Comparison between two groups or populations is an important problem in statistics and is commonly used in practice. The populations are usually compared with respect to their means (or medians) to establish superiority of one population over the other or to check if the two populations are equivalent. For example, two drugs may be compared with respect to their mean effects to determine the better one. Even though comparing two populations with respect to their means is a common problem, there are situations where one needs to compare the populations instead of their means. We shall discuss some specific practical situations after formulating the problem of this article. Consider two independent random variables X_1 and X_2. Let x_{i,p_i} denote the p_ith quantile of X_i, $i = 1, 2$. That is, $x_{i,p_i} = \inf\{x : P(X_i \leq x) \geq p_i\}$, $i = 1, 2$. The problem of interest here is to make inference

Received March 19, 2004; Accepted January 10, 2005
Address correspondence to K. Krishnamoorthy, Department of Mathematics, University of Louisiana at Lafayette, Lafayette, LA 70504, USA; Fax: 337-482-5346; E-mail: krishna@louisiana.edu

about $x_{1,p_1} - x_{2,p_2}$ based on a sample of n_i observations on X_i, $i = 1, 2$. Let us consider first the hypothesis testing problem when

$$H_0 : x_{1,p_1} - x_{2,p_2} \leq 0 \quad \text{vs.} \quad H_a : x_{1,p_1} - x_{2,p_2} > 0. \tag{1.1}$$

To understand the practical significance of the testing problem, let $p_1 = p_2 = 0.90$, then the hypothesis testing in (1.1) leads to the comparison between the 90th percentiles (or top 10% of the populations) of the two populations. Albers and Löhnberg (1984) presented a biomedical problem where comparison between the pth quantiles of two populations arises. These authors provided an approximate distribution free confidence interval for $x_{1,p} - x_{2,p}$. Bristol (1990) suggested a modification to Albers and Löhnberg's method. If $p_1 < 0.5$ and $p_2 > 0.5$, and if the sample data provide sufficient evidence to support H_a, then we conclude that majority of the data on X_1 are larger than the majority of the data on X_2. This is the case in reliability testing in a stress-strength model. If X_1 denotes the strength and X_2 denotes the stress to which X_1 is subjected, then the reliability parameter $R = P(X_1 - X_2 > 0)$. Letting $p_1 = 1 - p$ and $p_2 = p$, we see that $x_{1,1-p} - x_{2,p} > 0$ implies that $R > p$; however, the converse may not hold. Hall (1984), Reiser and Guttman (1986), and Weerahandi and Johnson (1992) give approximate inferential procedures for R when X_1 and X_2 are independent normal random variables. Bhattacharyya and Johnson (1974), Tong (1974), and Kelley et al. (1976) considered this reliability problem when X_1 and X_2 are independent exponential random variables. Thus, we see that the problem of comparing two quantiles is related to the reliability testing in a stress-strength model.

In this article, we propose methods for comparing the quantiles of two normal populations and comparing the quantiles of two exponential distributions. Our methods are based on the recent concepts of generalized p-value and generalized limit. The generalized p-value has been introduced by Tsui and Weerahandi (1989) and the generalized confidence interval by Weerahandi (1993); see Weerahandi (1995b). This approach has turned out to be very useful for obtaining tests and confidence intervals in some complex problems; see Zhou and Mathew (1994), Weerahandi (1995a), and Weerahandi and Berger (1999). Using this generalized variable approach, we give inferential procedures for the difference between two normal quantiles in the following section. The generalized variable approach yields exact methods when the variances of the normal populations are unknown and equal; when the variances are unknown and arbitrary, confidence intervals and p-values can be obtained using Monte Carlo simulation. The performance of the test based on the generalized p-value is evaluated numerically for the normal case with unequal variances. The numerical studies show that the test is very satisfactory in controlling the sizes. In Sec. 2.4, we present two examples for the normal case. In Sec. 3, we present inferential procedures for $x_{1,p_1} - x_{2,p_2}$ when the distributions are exponential with different scale parameters. In this case, the procedures based on the generalized variable are exact. Some concluding remarks are given in Sec. 4.

2. Generalized Inference for the Normal Case

Let $X_{11}, \ldots, X_{1n_1}$ be a sample of observations from $N(\mu_1, \sigma_1^2)$, and $X_{21}, \ldots, X_{2n_2}$ be a sample of observations from $N(\mu_2, \sigma_2^2)$. Let $\overline{X}_i$ and S_i^2 denote, respectively, the mean and variance based on the ith sample, $i = 1, 2$. Let $(\bar{x}_1, \bar{x}_2, s_1^2, s_2^2)$ be an observed value of $(\overline{X}_1, \overline{X}_2, S_1^2, S_2^2)$.

Let Q_{i,p_i} denote the p_ith quantile of the $N(\mu_i, \sigma_i^2)$ distribution, $i = 1, 2$. Note that $Q_{i,p_i} = \mu_i + z_{p_i}\sigma_i$, $i = 1, 2$, where z_α, $0 < \alpha < 1$, denotes the αth quantile of the standard normal distribution. We want to test

$$H_0 : Q_{1,p_1} \leq Q_{2,p_2} \quad \text{vs.} \quad H_a : Q_{1,p_1} > Q_{2,p_2}. \tag{2.1}$$

The method of constructing *generalized pivot statistic* for obtaining confidence limits on a parameter of interest and the *generalized test variable* for hypothesis testing are explained for a general setup in the appendix.

In the following, we present inferential procedures for $Q_{1,p_1} - Q_{2,p_2}$ when the variances are equal. Even though these procedures can be easily deduced from the solutions to the one-sample problem given in Weerahandi (1995b, p. 64), we provide them (without details) for the sake of completeness and easy reference. A generalized pivot statistic for $Q_{1,p_1} - Q_{2,p_2}$ without assuming equality of variances is given in Sec. 2.2.

2.1. *Inference When the Variances are Equal*

Let $\mu_d = \mu_1 - \mu_2$. If it is assumed that $\sigma_1^2 = \sigma_2^2 = \sigma^2$, then $Q_{1,p_1} - Q_{2,p_2} = \mu_d + (z_{p_1} - z_{p_2})\sigma$. Let S_p^2 denote the usual pooled estimate of the common variance σ^2 so that $(n_1 + n_2 - 2)S_p^2/\sigma^2 \sim \chi^2_{n_1+n_2-2}$, and let s_p^2 be an observed value of S_p^2. Furthermore, let $\overline{X}_d = \overline{X}_1 - \overline{X}_2$ and $\bar{x}_d$ be an observed value of $\overline{X}_d$. Since $\overline{X}_d \sim N(\mu_d, \sigma^2(1/n_1 + 1/n_2))$ independently of S_p^2, the solution to this problem can be easily deduced from the solution to the one-sample case. In other words, the generalized pivot statistic for $\mu_d + (z_{p_1} - z_{p_2})\sigma$ can be readily obtained from the generalized pivot statistic for a quantile of a $N(\mu, \sigma^2)$ distribution. For the latter one-sample problem, Weerahandi (1995b, p. 64) provided a generalized pivot statistic; using that pivot statistic, we get the generalized pivot statistic for $Q_{1,p_1} - Q_{2,p_2}$ as

$$T_1 = \bar{x}_d + t_{n_1+n_2-2}(\delta_1)s_p(1/n_1 + 1/n_2)^{\frac{1}{2}}, \tag{2.2}$$

where $\delta_1 = (z_{p_1} - z_{p_2})/(1/n_1 + 1/n_2)^{\frac{1}{2}}$. The quantiles of T can be used to get confidence bounds on $Q_{1,p_1} - Q_{2,p_2}$. In particular, we see that

$$\bar{x}_d + t_{n_1+n_2-2,\alpha}(\delta_1)s_p(1/n_1 + 1/n_2)^{\frac{1}{2}} \tag{2.3}$$

is an exact $1 - \alpha$ lower limit for $Q_{1,p_1} - Q_{2,p_2}$, where $t_{m,p}(c)$ denotes the $100p$th percentile of the non central t distribution with df $= m$ and non centrality parameter c. At the level of significance α, the null hypothesis in (2.1) will be rejected whenever the limit in (2.3) is greater than zero or the p-value

$$P(T_1 < 0 | Q_{1,p_1} - Q_{2,p_2} = 0) = P\big(t_{n_1+n_2-2}(-\delta_1) > \bar{x}_d/\big(s_p(1/n_1 + 1/n_2)^{\frac{1}{2}}\big)\big) < \alpha. \tag{2.4}$$

Using the relation that $t_{m,\alpha}(-c) = -t_{m,1-\alpha}(c)$, it is easy to see that the $1 - \alpha$ confidence interval for $Q_{1,p_1} - Q_{2,p_2}$ is given by

$$\big(\bar{x}_d - t_{n_1+n_2-2,1-\alpha/2}(-\delta_1)s_p(1/n_1 + 1/n_2)^{\frac{1}{2}},\ \bar{x}_d + t_{n_1+n_2-2,1-\alpha/2}(\delta_1)s_p(1/n_1 + 1/n_2)^{\frac{1}{2}}\big). \tag{2.5}$$

We note that the generalized variable approach produced exact limits and tests for $Q_{1,p_1} - Q_{2,p_2}$ when the variances are equal.

Remark 2.1. When $p_1 = p_2$, $Q_{1,p_1} - Q_{2,p_2} = \mu_1 - \mu_2 + (z_{p_1} - z_{p_2})\sigma = \mu_1 - \mu_2$. In this case, the problem of comparing the quantiles is equivalent to the problem of comparing the means. Also, the confidence interval in (2.5) is the usual two-sample t interval for $\mu_1 - \mu_2$.

2.2. *Inference on $Q_{1,p_1} - Q_{2,p_2}$ When the Variances are Unknown and Arbitrary*

The generalized pivot statistic for obtaining confidence limits for $Q_{1,p_1} - Q_{2,p_2} = \mu_1 - \mu_2 + z_{p_1}\sigma_1 - z_{p_2}\sigma_2$ when the variances are unknown and arbitrary can be developed as follows. Let

$$\begin{aligned} T_2 &= \bar{x}_d - \frac{\overline{X}_d - \mu_d}{(\sigma_1^2/n_1 + \sigma_2^2/n_2)^{\frac{1}{2}}}\left(\frac{\sigma_1^2}{n_1}\frac{s_1^2}{S_1^2} + \frac{\sigma_2^2}{n_2}\frac{s_2^2}{S_2^2}\right)^{\frac{1}{2}} + z_{p_1}\frac{\sigma_1}{S_1}s_1 - z_{p_2}\frac{\sigma_2}{S_2}s_2 \\ &= \bar{x}_d - Z\left(\frac{v_1^2}{n_1 U_1^2} + \frac{v_2^2}{n_2 U_2^2}\right)^{\frac{1}{2}} + z_{p_1}\frac{v_1}{\sqrt{W_1^2}} - z_{p_2}\frac{v_2}{\sqrt{W_2^2}}, \end{aligned} \tag{2.6}$$

where $v_i^2 = (n_i - 1)s_i^2, i = 1, 2,$ $Z, U_1^2, U_2^2, W_1^2,$ and W_2^2 are independent random variables with $Z \sim N(0, 1)$ and $U_i^2 \sim \chi^2_{n_i-1}$, and $W_i^2 \sim \chi^2_{n_i-1}$, $i = 1, 2$.

Notice that T_2 is a function of $(\overline{X}_1, \overline{X}_2, S_1^2, S_2^2; \bar{x}_1, \bar{x}_2, s_1^2, s_2^2)$, and satisfies two conditions in (A.3) of the appendix. In particular, it follows from the first equality of (2.6) that the value of T_2 at $(\overline{X}_1, \overline{X}_2, S_1^2, S_2^2) = (\bar{x}_1, \bar{x}_2, s_1^2, s_2^2)$ is $Q_{1,p_1} - Q_{2,p_2}$ (the parameter of interest). From the second equality of (2.6), we see that, for a given $(\bar{x}_1, \bar{x}_2, s_1^2, s_2^2)$, the distribution of T_2 is independent of any unknown parameters. Therefore, its percentiles (computed using Monte Carlo simulation) can be used to interval estimate $Q_{1,p_1} - Q_{2,p_2}$. Specifically, $T_{2,\alpha}$, the lower αth quantile of T_2, is a $1 - \alpha$ lower limit for $Q_{1,p_1} - Q_{2,p_2}$, and $(T_{2,\alpha/2}, T_{2,1-\alpha/2})$ is a $1 - \alpha$ confidence interval for $Q_{1,p_1} - Q_{2,p_2}$.

As shown in the Appendix, the generalized test variable for $Q_{1,p_1} - Q_{2,p_2}$ is given by $G_2 = T_2 - (Q_{1,p_1} - Q_{2,p_2})$. It is clear that, for given $(\bar{x}_1, \bar{x}_2, s_1^2, s_2^2)$, G_2 is stochastically decreasing with respect to $Q_{1,p_1} - Q_{2,p_2}$. So, the test based on G_2 rejects the null hypothesis in (2.1), whenever the generalized p-value

$$P(G_2 < 0 | Q_{1,p_1} - Q_{2,p_2} = 0) = P(T_2 < 0 | Q_{1,p_1} - Q_{2,p_2} = 0) < \alpha. \tag{2.7}$$

The generalized p-value can be estimated using Monte Carlo simulation. For example, to compute the generalized p-value in (2.7), generate 100,000 $(Z, U_1^2, U_2^2, W_1^2, W_2^2)$ random numbers. For each simulated $(Z, U_1^2, U_2^2, W_1^2, W_2^2)$, compute T_2 using step 2 of (2.6). The proportion of the values of T_2 which are less than zero is a Monte Carlo estimate of the generalized p-value in (2.7).

2.3. *Size and Power Studies*

To appraise the accuracies of the test based on the generalized variable T_2 in (2.6), we estimated its sizes using Monte Carlo simulation. To estimate the sizes, we generated 2500 values of $(\bar{x}_1, s_1^2, \bar{x}_2, s_2^2)$, and for each such generated value,

Table 1a

Monte Carlo estimates of the sizes of the generalized test based on (2.7) when $n_1 = n_2 = n$ and $\alpha = 0.05$; $\mu_d = 0$, $\sigma_2^2 = 1.0$; normal case (a) $H_0 : Q_{1,0.25} \leq Q_{2,0.75}$ vs. $H_a : Q_{1,0.25} > Q_{2,0.75}$; (b) $H_0 : Q_{1,0.05} \leq Q_{2,0.95}$ vs. $H_a : Q_{1,0.05} > Q_{2,0.95}$

	(a)				(b)			
$\sigma_1^2 \backslash n$	10	15	20	25	10	15	20	25
0.1	0.04	0.05	0.04	0.05	0.04	0.04	0.04	0.05
0.2	0.05	0.04	0.05	0.05	0.04	0.05	0.04	0.05
0.3	0.04	0.04	0.05	0.06	0.04	0.05	0.05	0.05
0.4	0.04	0.04	0.05	0.04	0.03	0.04	0.05	0.04
0.5	0.04	0.05	0.05	0.05	0.04	0.04	0.05	0.04
0.6	0.04	0.04	0.04	0.04	0.03	0.03	0.04	0.05
0.7	0.04	0.04	0.05	0.05	0.04	0.04	0.04	0.04
0.8	0.05	0.04	0.05	0.05	0.04	0.05	0.04	0.04
0.9	0.03	0.04	0.04	0.05	0.03	0.04	0.04	0.04
1	0.04	0.05	0.04	0.04	0.03	0.04	0.04	0.04
2	0.04	0.04	0.05	0.05	0.04	0.04	0.05	0.04
3	0.04	0.04	0.04	0.06	0.04	0.04	0.05	0.04
4	0.04	0.05	0.05	0.05	0.04	0.04	0.05	0.05
5	0.05	0.05	0.05	0.06	0.03	0.05	0.04	0.04
6	0.05	0.05	0.04	0.05	0.04	0.04	0.05	0.04
7	0.05	0.05	0.06	0.05	0.04	0.04	0.04	0.05
8	0.06	0.06	0.05	0.05	0.04	0.05	0.05	0.05
9	0.05	0.06	0.05	0.05	0.05	0.05	0.05	0.05
10	0.06	0.06	0.06	0.05	0.05	0.04	0.05	0.06

we used 5000 simulated values of $(Z, U_1^2, U_2^2, W_1^2, W_2^2)$ to estimate the probability in (2.7). The percentage of the 2500 probabilities which are less than α is an estimate of the size of the generalized test based on T_2. For a good test, this proportion should be close to α. IMSL subroutines RNCHI and RNNOA are used to generate, respectively, the chi-square random numbers and the normal random numbers. Since the distribution of T_2 is location-scale invariant, without loss of generality, we can take $\mu_d = 0$ and $\sigma_2^2 = 1$.

The estimated sizes of the generalized test based on (2.7) are given in Table 1a when $n_1 = n_2$, and in Table 1b when $n_1 \neq n_2$. The sizes are given for testing (a) $H_0 : Q_{1,0.25} \leq Q_{2,0.75}$ vs. $H_a : Q_{1,0.25} > Q_{2,0.75}$ and for testing (b) $H_0 : Q_{1,0.05} \leq Q_{2,0.95}$ vs. $H_a : Q_{1,0.05} > Q_{2,0.95}$. We observe from these table values that the sizes vary from 0.03 to 0.06 when the nominal level is 0.05; however, they are in general very close to the nominal level 0.05. Thus, the generalized test is satisfactory for practical purposes.

To understand the power properties of the generalized test based on (2.7), we plotted powers as a function of $Q_{1,0.25} - Q_{2,0.75}$ for some values of n_1 and n_2 and $\alpha = 0.05$. For given n_1 and n_2, the powers are computed at the parameter values $\sigma_1^2 = 1$, $\sigma_2^2 = 2$, and various values of $\mu_d = \mu_1 - \mu_2$ so that $Q_{1,0.25} - Q_{2,0.75} \geq 0$. The power plots are given in Figure 1. It is clear from the plots that the test possesses some natural power properties. In particular, the power increases as $Q_{1,0.25} - Q_{2,0.75}$ increases; also, the power increases as the sample sizes increase.

Table 1b
Monte Carlo estimates of the sizes of the generalized test based on (2.7) when $n_1 \neq n_2$ and $\alpha = 0.05$; $\mu_d = 0$, $\sigma_2^2 = 1.0$; normal case (a) $H_0 : Q_{1,0.25} \leq Q_{2,0.75}$ vs. $H_a : Q_{1,0.25} > Q_{2,0.75}$; (b) $H_0 : Q_{1,0.05} \leq Q_{2,0.95}$ vs. $H_a : Q_{1,0.05} > Q_{2,0.95}$

	(a)			(b)		
$\sigma_1^2 \backslash (n_1, n_2)$	(10,15)	(8,20)	(7,32)	(10,15)	(8,20)	(7,32)
0.1	0.05	0.04	0.04	0.04	0.04	0.04
0.2	0.05	0.04	0.03	0.04	0.04	0.04
0.3	0.04	0.04	0.04	0.04	0.04	0.04
0.4	0.05	0.04	0.04	0.04	0.04	0.04
0.5	0.04	0.05	0.04	0.04	0.04	0.04
0.6	0.04	0.04	0.04	0.04	0.04	0.04
0.7	0.05	0.05	0.04	0.04	0.05	0.04
0.8	0.05	0.05	0.05	0.04	0.03	0.04
0.9	0.04	0.04	0.04	0.04	0.04	0.04
1	0.04	0.04	0.05	0.04	0.04	0.05
2	0.05	0.05	0.05	0.04	0.05	0.04
3	0.05	0.05	0.06	0.04	0.05	0.04
4	0.05	0.05	0.05	0.04	0.05	0.05
5	0.04	0.05	0.06	0.05	0.05	0.05
6	0.06	0.05	0.06	0.04	0.05	0.05
7	0.05	0.05	0.04	0.05	0.05	0.05
8	0.05	0.06	0.06	0.04	0.05	0.05
9	0.05	0.05	0.05	0.05	0.05	0.05
10	0.06	0.05	0.06	0.05	0.05	0.05

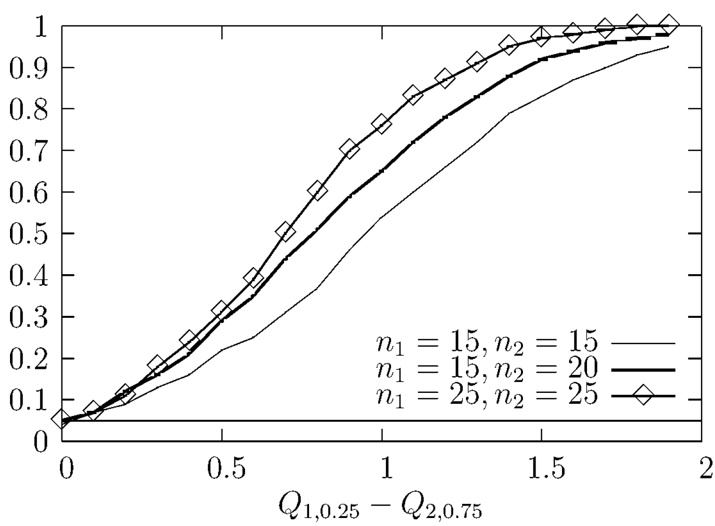

Figure 1. Powers of the generalized test based on (2.7).

2.4. *Examples*

We shall now illustrate the methods using the examples given in Hall (1984) and Weerahandi and Johnson (1992). These examples are considered in these articles for the purpose of estimating and testing the reliability parameter $R = P(X_1 > X_2)$. We here compute lower limits and test for $Q_{1,p_1} - Q_{2,p_2}$ assuming normality. The generalized p-values and limits are computed using 1,000,000 simulation runs.

Example 1 (Hall, 1984). A sample of $n_1 = 50$ capacitors yielded mean breakdown voltage $\bar{x}_1 = 6.75$ kV and $s_1^2 = 0.123$. The voltage output from $n_2 = 20$ transverters (power supplies) produced $\bar{x}_2 = 4.00$ kV and $s_2^2 = 0.53$. We have $\bar{x}_d = \bar{x}_1 - \bar{x}_2 = 2.75$. Using the methods in Sec. 2.2, we computed 95% lower limit for $Q_{0.05} - Q_{0.95}$ as 0.421. This means that 95% of the breakdown voltage exceeds the 95th percentile of the voltage output by 0.421 kV. We also computed 95% lower limit for $Q_{0.03} - Q_{0.97}$ as 0.112 and for $Q_{0.02} - Q_{0.98}$ as -0.119. In addition, we computed p-values for testing $H_a : Q_{0.05} > Q_{0.95}$ as 0.0037, for testing $H_a : Q_{0.03} > Q_{0.97}$ as 0.028, and for testing $H_a : Q_{0.02} > Q_{0.98}$ as 0.087. Thus, we conclude that the 0.03rd quantile of the breakdown voltage exceeds the 0.97th quantile of the voltage output when the nominal level is 0.05.

Example 2. In this example, we consider the rocket-motor experiment data set given in Weerahandi and Johnson (1992) for illustration purpose. We are interested in making inferences on the reliability of the rocket motor at the highest operating temperature of 59°C. At this temperature, the operating pressure X_2 distribution tends to be closest to the chamber burst strength X_1 distribution. Assumption of normality has been verified in the paper just cited. We want to test $H_0 : Q_{0.000001} - Q_{0.999999} \leq 0$ vs. $H_a : Q_{0.000001} - Q_{0.999999} > 0$. A sample of $n_1 = 17$ and a sample of $n_2 = 24$ observations yielded the statistics $\bar{x}_1 = 16.485$, $s_1^2 = 0.3409$, $\bar{x}_2 = 7.789$, and $s_2^2 = 0.05414$. Using (2.7), we computed the p-value as 0.0003. We see that the data provide strong evidence to indicate that the 0.000001th quantile of the burst strength distribution exceeds the 0.999999th quantile of the pressure distribution. We can also conclude that the reliability parameter $R = P(X_1 > X_2) > 0.999999$. The reported generalized p-value in Weerahandi and Johnson (1992) for testing $H_a : R \leq 0.999999$ vs. $H_a : R > 0.999999$ is 0.0000042. It should be noted that the result based on quantile testing not only indicates that $P(X_1 > X_2) > 0.999999$ but also that the 0.000001th quantile of X_1 exceeds the 0.999999th quantile of X_2.

3. Exponential Distributions

We shall now present the generalized inferential procedures for the difference between the quantiles of two exponential distributions. Note that the probability density function (pdf) of an exponential distribution with failure rate λ is given by $f(x|\lambda) = \lambda \exp(-\lambda x)$, $x > 0$ and $\lambda > 0$. The cumulative density function (cdf) is given by $F(x|\lambda) = 1 - \exp(-\lambda x)$. For any given $0 < p < 1$, the pth quantile is the root of $F(x|\lambda) = p$, and is given by $-\ln(1-p)/\lambda$. Thus, if $X_i \sim f(x|\lambda_i)$, $i = 1, 2$, then the p_ith quantile of X_i can be expressed as

$$\eta_i = \frac{-\ln(1-p_i)}{\lambda_i}, \quad i = 1, 2. \tag{3.1}$$

Because η_1 and η_2 are positive, testing $H_0 : \eta_1 \leq \eta_2$ vs. $H_a : \eta_1 > \eta_2$ is equivalent to

$$H_0 : \frac{\lambda_2}{\lambda_1} \leq c \quad \text{vs.} \quad H_a : \frac{\lambda_2}{\lambda_1} > c, \tag{3.2}$$

where $c = \ln(1-p_2)/\ln(1-p_1)$. Let $X_{i1}, \ldots, X_{in_i}$ be a sample from $F(x|\lambda_i)$, $i = 1, 2$. Define $Y_i = \sum_{j=1}^{n_i} X_{ij}$, $i = 1, 2$ and let y_i be an observed value of Y_i, $i = 1, 2$. Notice that Y_1 and Y_2 are independent and $2\lambda_i Y_i \sim \chi^2_{2n_i}$, $i = 1, 2$, and hence Y_2/Y_1 is distributed as a constant times F random variable. Using these results, it is not difficult to show that the p-value for testing (3.2) is given by

$$P\left(F_{2n_2,2n_1} < c\frac{n_1 y_2}{n_2 y_1}\right), \tag{3.3}$$

where $F_{a,b}$ denotes the F random variable with the numerator df $= a$ and the denominator df $= b$. The null hypothesis in (3.2) or equivalently $H_0 : \eta_1 - \eta_2 \leq 0$ will be rejected whenever this p-value is less than α.

Although the above testing approach for hypotheses in (3.2) is well known (see Lawless, 1982, Sec. 3.3), it is of interest to see that the generalized p-value is indeed equal to the p-value obtained above. Define

$$\begin{aligned} T_3 &= \frac{-\ln(1-p_1)y_1}{\lambda_1 Y_1} - \frac{-\ln(1-p_2)y_2}{\lambda_2 Y_2} \\ &= \frac{-\ln(1-p_1)y_1}{Y_1^*} - \frac{-\ln(1-p_2)y_2}{Y_2^*}, \end{aligned} \tag{3.4}$$

where $Y_1^* = \lambda_1 Y_1$ and $Y_2^* = \lambda_2 Y_2$. It can be easily verified that the generalized p-value $P(T_3 < 0|\eta_1 - \eta_2 = 0) = P(\frac{-\ln(1-p_1)y_1}{Y_1^*} - \frac{-\ln(1-p_2)y_2}{Y_2^*} < 0|\eta_1 - \eta_2 = 0)$ for testing (3.2) is given by (3.3).

The confidence limits for the ratio η_1/η_2 can be easily obtained using the distributional result that $\frac{n_1 Y_2 \lambda_2}{n_2 Y_1 \lambda_1} \sim F_{2n_2,2n_1}$. However, to obtain confidence limits for the difference $\eta_1 - \eta_2$, we need to compute the quantiles of T_3. Because the distribution of T_3, given y_1 and y_2, is free of any unknown parameters, Monte Carlo simulation can be used to find the quantiles of T_3. The quantiles of T_3 can be computed more accurately using a numerical integration, and a root searching method. Toward this, let $F(x; a)$ denote the gamma cdf with the shape parameter a. The αth quantile c_1 of T_3 is the solution of the equation

$$P(T_3 < c_1|n_1, n_2, a_1, a_2) = g(c_1) = \alpha, \tag{3.5}$$

where $a_i = -\ln(1-p_i)y_i$, $i = 1, 2$. Using some standard arguments, $g(c_1)$ can be expressed as

$$g(c_1) = \begin{cases} 1 - E_{Y_2^*} F\left(\left(\frac{a_1 Y_2^*}{c_1 Y_2^* + a_2}\right); n_1\right), & c_1 > 0, \\ E_{Y_1^*} F\left(\left(\frac{a_2 Y_1^*}{a_1 - c_1 Y_1^*}\right); n_2\right), & c_1 \leq 0, \end{cases} \tag{3.6}$$

where Y_i^* is a gamma random variable with the shape parameter n_i, $i = 1, 2$ and E_X denotes the expectation with respect to the distribution of X. We computed the values of c_1 for various values of (n_1, n_2, a_1, a_2) and $\alpha = 0.05$ using Monte Carlo simulation, and using the equation in (3.5). Both approaches yielded practically the same values (they are not reported here). Therefore, we recommend Monte Carlo simulation for computing the quantiles of T_3 because it is simpler than the numerical method based on (3.5) and (3.6).

4. Concluding Remarks

In this article, we developed inferential procedures for comparing two normal quantiles and for comparing two exponential quantiles using the generalized variable approach. The procedures given in Sec. 2 for normal distributions are applicable to compare the quantiles of two lognormal populations. Specifically, if the data are from lognormal distributions, then the methods can be applied to the logged data to construct confidence limits or hypothesis testing for the difference between two lognormal quantiles. The developed procedures are useful for establishing the superiority of one population over another provided the data satisfy the distributional assumptions of this article.

Appendix: The Generalized *p*-Value and Generalized Confidence Interval

A general setup in which the generalized p-value can be defined is as follows. Let X be a random variable (could be a vector) whose distribution depends on the parameters (θ, δ), where θ is a scalar parameter of interest, and δ is a nuisance parameter. Suppose we want to test

$$H_0 : \theta \le \theta_0 \quad \text{vs. } H_a : \theta > \theta_0, \tag{A.1}$$

where θ_0 is a specified quantity. Let x be an observed value of X. A *generalized test variable*, to be denoted by $G(X; x, \theta, \delta)$, is a function of X, x, θ, and δ, and it satisfies the following conditions:

(i) For a fixed x, the distribution of $G(X; x, \theta, \delta)$ is free of the nuisance parameter δ.
(ii) The observed value of $G(X; x, \theta, \delta)$, namely $G(x; x, \theta, \delta)$, is free of any unknown parameters.
(iii) For fixed x and δ, the distribution of $G(X; x, \theta, \delta)$ is stochastically increasing or stochastically decreasing in θ. That is, $P(G(X; x, \theta, \delta) \ge a)$ is an increasing function of θ, or is a decreasing function of θ, for every a. (A.2)

Now, let $g = G(x; x, \theta, \delta)$, the observed value of $G(X; x, \theta, \delta)$. If $G(X; x, \theta, \delta)$ is stochastically increasing in θ, the generalized p-value for testing the hypotheses in (A.1) is given by

$$\sup_{H_0} P(G(X; x, \theta, \delta) \ge g) = P(G(X; x, \theta_0, \delta) \ge g | \theta = \theta_0),$$

and if $G(X; x, \theta, \delta)$ is stochastically decreasing in θ, the generalized p-value for testing the hypotheses in (A.1) is given by

$$\sup_{H_0} P(G(X; x, \theta, \delta) \le g) = P(G(X; x, \theta_0, \delta) \le g | \theta = \theta_0).$$

Note that the computation of the generalized p-value is possible in view of the conditions (i) and (ii) in (A.2), i.e., the distribution of $G(X; x, \theta, \delta)$ is free of the nuisance parameter δ and $g = G(x; x, \theta, \delta)$ is free of any unknown parameters.

A generalized confidence interval for θ is computed using the percentiles of a *generalized pivot statistic*, say $T(X; x, \theta, \delta)$, satisfying the following conditions:

(i) The observed value of $T(X; x, \theta, \delta)$, namely $T(x; x, \theta, \delta)$ is θ, the parameter of interest.
(ii) For a fixed x, the distribution of $T(X; x, \theta, \delta)$ is free of unknown parameters. (A.3)

In general, $T(X; x, \theta, \delta) = G(X; x, \theta, \delta) + \theta$. The percentiles of $T(X; x, \theta, \delta)$ can be used to compute confidence intervals for θ. For example, the 95th percentile of $T(X; x, \theta, \delta)$ is a 95% generalized upper confidence limit for θ.

For further details on the concepts of generalized p-values and generalized confidence intervals, along with numerous examples, we refer to the original articles by Tsui and Weerahandi (1989) and Weerahandi (1995b).

References

Albers, W., Löhnberg, P. (1984). An approximate confidence interval for the difference between quantiles in a bio-medical problem. *Statistica Neerlandica* 38:20–22.

Bhattacharyya, G. K., Johnson, R. A. (1974). Estimation of reliability in a multicomponent stress-strength model. *J. Amer. Statist. Assoc.* 69:966–970.

Bristol, D. R. (1990). Distribution-free confidence intervals for the difference between quantiles. *Statistica Neerlandica* 44:87–90.

Hall, I. J. (1984). Approximate one-sided tolerance limits for the difference or sum of two independent normal variates. *J. Qual. Technol.* 16:15–19.

Kelley, G. D., Kelley, J. A., Schucany, W. R. (1976). Efficient estimation of $\Pr(Y < X)$ in the exponential ease. *Technometrics* 18:359–360.

Lawless, J. F. (1982). *Statistical Models and Methods for Lifetime Data*. New York: John Wiley.

Reiser, B. J., Guttman, I. (1986). Statistical inference for $\Pr(Y < X)$: the normal case. *Technometrics* 28:253–257.

Tong, H. (1974). A note on the estimation of $\Pr(Y < X)$ in the exponential case. *Technometrics* 16:625.

Tsui, K. W., Weerahandi, S. (1989). Generalized p-values in significance testing of hypotheses in the presence of nuisance parameters. *J. Amer. Statist. Assoc.* 84:602–607.

Weerahandi, S. (1993). Generalized confidence intervals. *J. Amer. Statist. Assoc.* 88:899–905.

Weerahandi, S. (1995a). ANOVA under unequal error variances. *Biometrics* 51:589–599.

Weerahandi, S. (1995b). *Exact Statistical Methods for Data Analysis*. New York: Springer-Verlag.

Weerahandi, S., Johnson, R. A. (1992). Testing reliability in a stress-strength model when X and Y are normally distributed. *Technometrics* 34:83–91.

Weerahandi, S., Berger, V. W. (1999). Exact inference for growth curves with intraclass correlation structure. *Biometrics* 55:921–924.

Zhou, L., Mathew, T. (1994). Some tests for variance components using generalized p-values. *Technometrics* 36:394–402.