Estimation of parametric functions in Downton's bivariate exponential distribution

George Iliopoulos

Department of Mathematics

University of the Aegean

83200 Karlovasi, Samos, Greece

e-mail: *geh@aegean.gr*

Abstract

This paper considers estimation of the ratio of means and the regression function in Downton's (1970) bivariate exponential distribution. Unbiased estimators are given and, by presenting improved estimators, they are shown to be inadmissible in terms of mean squared error. The results are derived by conditioning on an unobserved random sample from a geometric distribution which provides conditional independence for the statistics involved.

AMS 2000 subject classifications: 62F10, 62C99.

Key words and phrases: Downton's bivariate exponential distribution, unbiased estimation, ratio of means, regression function, mean squared error, inadmissibility.

1   Introduction

One of the most important bivariate distributions in reliability theory is the bivariate exponential. There are various bivariate exponential distributions in the literature. A recent review can be found in the book of Kotz, Balakrishnan and Johnson (2000). In this paper we are interested in Downton's bivariate exponential distribution with probability

density function (pdf)

$$f(x, y; \lambda_1, \lambda_2, \rho) = \frac{\lambda_1 \lambda_2}{1 - \rho} \exp\left\{ -\frac{\lambda_1 x + \lambda_2 y}{1 - \rho} \right\} I_0 \left\{ \frac{2(\rho \lambda_1 \lambda_2 xy)^{1/2}}{1 - \rho} \right\} , \qquad (1.1)$$

where x, y, λ_1, $\lambda_2 > 0$, $0 \leqslant \rho < 1$, and $I_0(z) = \sum_{k=0}^{\infty} (z/2)^{2k}/k!^2$ is the modified Bessel function of the first kind of order zero. The above density was initially derived in a different form by Moran (1967). The form (1.1) is derived by Downton (1970) in a reliability context and is a special case of Kibble's (1941) bivariate gamma distribution.

Let (X, Y) be an observation from (1.1). The marginal distributions of X, Y are exponential with means (scale parameters) $1/\lambda_1$, $1/\lambda_2$ respectively. Since $I_0(0) = 1$, it is clear that X and Y are independent if and only if $\rho = 0$. Downton (1970) showed that ρ is the correlation coefficient of the two variates. By expanding in a series, the pdf can be written in the form

$$f(x, y; \lambda_1, \lambda_2, \rho) = \sum_{k=0}^{\infty} \pi(k; \rho) \, g_{k+1}\left(x; \frac{1-\rho}{\lambda_1} \right) g_{k+1}\left(y; \frac{1-\rho}{\lambda_2} \right) ,$$

where $g_\alpha(\cdot; \beta)$ denotes the pdf of a Gamma(α, β) random variable and $\pi(k; \rho) = (1 - \rho)\rho^k$, $k = 0, 1, 2, \ldots$, is the geometric probability mass function. Let K be a random variable having the above geometric distribution. Then, conditionally on $K = k$, X, Y are independent gamma variates with shape parameter $k+1$ and scale parameters $(1 - \rho)/\lambda_1$, $(1-\rho)/\lambda_2$ respectively. The most common algorithm for generating observations from (1.1) (see Downton, 1970 and Al-Saadi, Scrimshaw and Young, 1979) as well as the extension of the above distribution in more than two dimensions (see Al-Saadi and Young, 1982) are based on this well–known property.

Statistical inference for the parameters of (1.1) is restricted mainly on the correlation coefficient ρ. Nagao and Kadoya (1971), Al Saadi and Young (1980), and Balakrishnan and Ng (2001) considered the estimation problem of ρ, and Al Saadi, Scrimshaw and Young (1979) the problem of testing the hypothesis $\rho = 0$. However, another interesting problem is the estimation of $\lambda = \lambda_2/\lambda_1$, which represents the ratio of the means of the two components. For example, an estimated value greater than one indicates that on the average the first component is more reliable than the second one. Note that λ is also the ratio of the scale parameters of X and Y. Estimation of λ in general scale families including among others normal, exponential and inverse Gaussian has been considered by many authors in the past. For a decision theoretic approach, see Gelfand and Dey (1988), Madi and Tsui (1990), Kubokawa (1994), Madi (1995), Ghosh and Kundu (1996),

Kubokawa and Srivastava (1996) (who assume independence of the two components) and Iliopoulos (2001) (who considers the problem of estimation of the ratio of variances in the bivariate normal distribution).

Next, we outline the rest of the paper. In Section 2, an unbiased estimator, $\hat{\lambda}_U$, of λ is derived based on a random sample from (1.1). Then, a class of inadmissible estimators with respect to the mean squared error is constructed and it is shown that this class contains $\hat{\lambda}_U$. Furthermore, some alternative biased estimators dominating $\hat{\lambda}_U$ are presented. In Section 3, unbiased estimators of the regression of X on Y, as well as of the conditional variance of X given $Y = y$, are given. They are also shown to be inadmissible; improved estimators are presented as well. Finally, an Appendix contains useful expressions for expectations of geometric and negative binomial distributions as well as of the statistics involved in the derivation of the results.

2 Estimation of the ratio of means

Let $(X_1, Y_1), \ldots, (X_n, Y_n)$, $n \geqslant 2$, be a random sample from (1.1) and $\mathbf{K} = (K_1, \ldots, K_n)$ be the associated (unobserved) random sample from the geometric distribution $\pi(\cdot\,; \rho)$ such that, given $K_i = k_i$, X_i is independent of Y_i, $i = 1, \ldots, n$. Since each K_i is related only to (X_i, Y_i), it is clear that, conditionally on $\mathbf{K} = \mathbf{k} = (k_1, \ldots, k_n)$, all X's and Y's are independent. Set $K = \sum K_i$, $k = \sum k_i$ and note that K follows a negative binomial distribution.

By considering the joint distribution of the data, it is easily seen that the sufficient statistic is $(X_1 Y_1, \ldots, X_n Y_n, \sum X_i, \sum Y_i)$. Setting $S_1 = \sum X_i$, $S_2 = \sum Y_i$, $\mathbf{U} = (U_1, \ldots, U_n) = (X_1 S_1^{-1}, \ldots, X_n S_1^{-1})$, and $\mathbf{V} = (V_1, \ldots, V_n) = (Y_1 S_2^{-1}, \ldots, Y_n S_2^{-1})$ we obtain the one–to–one transformation $(U_1 V_1, \ldots, U_n V_n, S_1, S_2)$, which is also sufficient. Conditionally on $\mathbf{K} = \mathbf{k}$, S_1, S_2 are independent and $S_i \sim \mathrm{Gamma}(n + k, (1 - \rho)\lambda_i^{-1})$, $i = 1, 2$. Moreover, from a well–known characterization of the gamma distribution, (S_1, S_2) is independent of $(\mathbf{U}, \mathbf{V})$, and $\mathbf{U}$, $\mathbf{V}$ are iid from a $(n - 1)$-variate Dirichlet distribution with parameters $k_1 + 1, \ldots, k_{n-1} + 1, k_n + 1$.

Consider the estimation problem of $\lambda = \lambda_2/\lambda_1$. Nagao and Kadoya (1971) showed that the maximum likelihood estimators (mles) of λ_1 and λ_2 are $1/\bar{X}$ and $1/\bar{Y}$ respectively, thus the mle of λ is $\hat{\lambda}_{\mathrm{mle}} = S_1/S_2$. Using Lemma 4.1(vii) in the Appendix, we obtain the expectation of this estimator,

$$\mathbb{E}[S_1/S_2] = \mathbb{E}[\mathbb{E}(S_1/S_2 | \mathbf{K})] = \lambda\, \mathbb{E}\left[\frac{n + K}{n + K - 1}\right] = \lambda\left(\frac{n - \rho}{n - 1}\right). \qquad (2.1)$$

Hence, $\hat{\lambda}_{\mathrm{mle}}$ is biased. For deriving an unbiased estimator of λ it is necessary to employ an estimator of the correlation coefficient ρ. There are two classes of estimators of ρ in the literature: (i) estimators based on the statistic $T = \sum X_i Y_i / S_1 S_2 = \sum U_i V_i$ (such as the moment estimator) and (ii) estimators based on the sample correlation coefficient R, see Al-Saadi and Young (1980) and Balakrishnan and Ng (2001). However, R is not a function of the sufficient statistic, whereas T is. Therefore, T has been chosen for our purposes. Note also that the problem of estimation of λ remains invariant under the group of transformations $(X_i, Y_i) \longrightarrow (c_1 X_i, c_2 Y_i)$, $i = 1, \ldots, n$, and equivariant estimators of λ are of the form $\psi(U_1 V_1, \ldots, U_n V_n) S_1 / S_2$. A particular choice for ψ can be of the form $\psi(T)$, giving more justification to T.

The conditional expectation of T given $\mathbf{K} = \mathbf{k}$ is

$$\mathbb{E}[T | \mathbf{K} = \mathbf{k}] = \sum_{i=1}^{n} \mathbb{E}[U_i V_i | \mathbf{K} = \mathbf{k}] = \sum_{i=1}^{n} \mathbb{E}(U_i | \mathbf{K} = \mathbf{k})^2 = \sum_{i=1}^{n} \frac{(k_i + 1)^2}{(k + n)^2} \ .$$

Since T is a function of $\mathbf{U}$ and $\mathbf{V}$ solely, it follows that, conditionally on $\mathbf{K}$, it is also independent of S_1, S_2. Therefore,

$$
\begin{aligned}
\mathbb{E}[T S_1 / S_2] &= \mathbb{E}[\mathbb{E}(T S_1 / S_2 | \mathbf{K})] = \mathbb{E}[\mathbb{E}(T | \mathbf{K}) \mathbb{E}(S_1 / S_2 | \mathbf{K})] \\
&= \lambda \mathbb{E}\left[\sum_{i=1}^{n} \frac{(K_i + 1)^2}{(n + K)(n + K - 1)} \right] \\
&= \lambda \mathbb{E}\left\{ n(n + K)^{-1}(n + K - 1)^{-1} \mathbb{E}\left[(K_1 + 1)^2 \middle| K \right] \right\} \\
&= \lambda \mathbb{E}\left[\frac{n + 1 + 2K}{(n + 1)(n + K - 1)} \right] = \lambda \left(\frac{1}{n - 1} + \frac{n - 3}{n^2 - 1} \rho \right) \quad (2.2)
\end{aligned}
$$

(see Lemma 4.1). From (2.1), (2.2) it can be seen that each of $\mathbb{E}[S_1 / S_2]$, $\mathbb{E}[T S_1 / S_2]$ equals λ times a first degree polynomial in ρ. The derivation of an unbiased estimator of λ which is a function of S_1, S_2 and T is equivalent to finding c_0, c_1 such that $\mathbb{E}[c_0 S_1 / S_2 + c_1 T S_1 / S_2] = \lambda$. Solving the linear equations

$$
\begin{aligned}
\frac{n}{n - 1} c_0 + \frac{1}{n - 1} c_1 &= 1 \\
-\frac{1}{n - 1} c_0 + \frac{n - 3}{n^2 - 1} c_1 &= 0 \ ,
\end{aligned}
$$

we obtain $c_0 = (n - 3)/(n - 1)$, $c_1 = (n + 1)/(n - 1)$. Thus, we have proved the following proposition.

Proposition 2.1. *The estimator*

$$\hat{\lambda}_U = \left(\frac{n - 3 + (n + 1) T}{n - 1} \right) \frac{S_1}{S_2} \quad (2.3)$$

is unbiased for $\lambda = \lambda_2 / \lambda_1$.

4

For $n \geqslant 3$, the variance of $\hat{\lambda}_U$ is

$$
\begin{aligned}
\mathrm{Var}(\hat{\lambda}_U) &= \mathbb{E}\left(\frac{n-3+(n+1)T}{n-1}\frac{S_1}{S_2}\right)^2 - \lambda^2 \\
&= \left(\tfrac{n-3}{n-1}\right)^2 \mathbb{E}\left[S_1^2/S_2^2\right] + \tfrac{2(n-3)(n+1)}{(n-1)^2}\mathbb{E}\left[TS_1^2/S_2^2\right] + \left(\tfrac{n+1}{n-1}\right)^2 \mathbb{E}\left[T^2 S_1^2/S_2^2\right] - \lambda^2 \ ,
\end{aligned}
$$

and substituting the expectations from Lemma 4.2 we get

$$
\mathrm{Var}(\hat{\lambda}_U) = \left[\frac{2n^2 - 5n + 5}{(n-2)(n-1)^2} - \frac{2(n^3 - 3n + 10)}{(n^2 - 4)(n-1)^2}\rho + \frac{n^3 + 6n^2 - 5n + 38}{(n+3)(n^2 - 4)(n-1)^2}\rho^2\right]\lambda^2 \ .
$$

Consider the class of estimators of λ,

$$
\mathcal{C} = \{\hat{\lambda}_{a_1,a_2} = (a_1 + a_2 T)S_1/S_2, \ a_1, a_2 \in \mathbb{R}\} \ .
$$

The unbiased estimator $\hat{\lambda}_U$ as well as the mle $\hat{\lambda}_{\mathrm{mle}}$ are members of $\mathcal{C}$ for $a_1 = a_{1U} = (n-3)/(n-1)$, $a_2 = a_{2U} = (n+1)/(n-1)$ and $a_1 = 1$, $a_2 = 0$, respectively. We would like to characterize inadmissible estimators within $\mathcal{C}$ in terms of mean squared error (mse). By invariance, the (scaled) mse $\lambda^{-2}\mathbb{E}_{\lambda_1,\lambda_2,\rho}(\hat{\lambda}_{a_1,a_2} - \lambda)^2$ does not depend on λ_1, λ_2. Thus, without loss of generality, we assume for the rest of the section that $\lambda_1 = \lambda_2 = 1$ and denote the mse of $\hat{\lambda}_{a_1,a_2}$ as $\mathrm{mse}(\mathbf{a}, \rho)$, where $\mathbf{a} = (a_1, a_2)'$.

Fix $\rho \in [0,1]$. Then, for $n \geqslant 3$, $\mathrm{mse}(\mathbf{a}, \rho)$ is strictly convex in $\mathbf{a}$ and there exists a minimizing point $\mathbf{a}_0(\rho) = (a_{10}(\rho), a_{20}(\rho))'$ with

$$
\begin{aligned}
a_{10}(\rho) &= (n-2)q_1(\rho)/q_2(\rho) \\
a_{20}(\rho) &= 3(n-2)(n+1)(n+2)(n+3)\rho(1-\rho)^2/q_2(\rho) \ ,
\end{aligned}
$$

where

$$
\begin{aligned}
q_1(\rho) &= (n+1)(n+2)(n+3) + 4(n-6)(n+1)(n+3)\rho \\
&\quad - (n-5)(3n^2 + 29n + 30)\rho^2 + 2(n^3 - 11n - 46)\rho^3 \ , \\
q_2(\rho) &= (n+1)^2(n+2)(n+3) + 4(n-6)(n+1)^2(n+3)\rho \\
&\quad - (3n^4 + 32n^3 - 77n^2 - 382n - 312)\rho^2 \\
&\quad + 2(n^4 + 4n^3 + 25n^2 - 126n - 256)\rho^3 - 3(n^3 + 13n - 94)\rho^4 \ .
\end{aligned}
$$

As expected, it holds $\mathbf{a}_0(0) = ((n-2)/(n+1), 0)'$, i.e., in the case of two independent exponential samples, the best estimator within $\mathcal{C}$ coincides with the best equivariant estimator of λ. On the other hand, $\mathbf{a}_0(1) = (1, 0)'$, that is, the best estimator in this case is the mle. Notice here that the mse of the mle tends to zero as $\rho \to 1$. To see that

without evaluating it, observe that since the support of X_i, Y_i is $(0, \infty)$, $\rho = 1$ implies that $X_i/\lambda_1 = Y_i/\lambda_2$ with probability one.

Setting

$$\mathbf{B}(\rho) = \begin{pmatrix} \mathbb{E}_{\lambda_1=\lambda_2=1,\rho}(S_1^2/S_2^2) & \mathbb{E}_{\lambda_1=\lambda_2=1,\rho}(TS_1^2/S_2^2) \\ \mathbb{E}_{\lambda_1=\lambda_2=1,\rho}(TS_1^2/S_2^2) & \mathbb{E}_{\lambda_1=\lambda_2=1,\rho}(T^2S_1^2/S_2^2) \end{pmatrix}$$

the mse of $\hat{\lambda}_{a_1,a_2}$ can be expressed as

$$\mathrm{mse}(\mathbf{a}, \rho) = [\mathbf{a} - \mathbf{a}_0(\rho)]'\mathbf{B}(\rho)[\mathbf{a} - \mathbf{a}_0(\rho)] + \mathrm{mse}(\mathbf{a}_0(\rho), \rho) .$$

Let

$$E(\mathbf{a}, \rho) = \left\{ \mathbf{c} \in \mathbb{R}^2 : [\mathbf{c} - \mathbf{a}_0(\rho)]'\mathbf{B}(\rho)[\mathbf{c} - \mathbf{a}_0(\rho)] < \mathrm{mse}(\mathbf{a}, \rho) - \mathrm{mse}(\mathbf{a}_0(\rho), \rho) \right\}$$

be the interior of the ellipse that consists of the points $\mathbf{c} = (c_1, c_2)'$, such that $\hat{\lambda}_{c_1,c_2}$ has equal mse with $\hat{\lambda}_{a_1,a_2}$ for the particular ρ. Then, $\hat{\lambda}_{a_1,a_2}$ is admissible within $\mathcal{C}$ if and only if

$$\bigcap_{\rho \in [0,1)} E(\mathbf{a}, \rho) = \varnothing .$$

This condition is clearly satisfied by $\hat{\lambda}_{a_{10}(\rho),a_{20}(\rho)}$, $\forall \, \rho \in [0, 1)$, implying that these estimators are admissible within $\mathcal{C}$. By the continuity of the mse, this holds also for the mle. However, the determination of the above intersection is in general a problem which does not seem to allow for an analytical solution. Instead of that, we can find a subclass of $\mathcal{C}$ containing inadmissible estimators $\hat{\lambda}_{a_1,a_2}$, by fixing a_1 or a_2 one at a time.

Fix first a_1. Then, the mse of $\hat{\lambda}_{a_1,a_2}$ is quadratic in a_2 and uniquely minimized at $a_2 = a_2^*(a_1, \rho)$ given by

$$a_2^*(a_1, \rho) = \frac{(n-2)[n+1-(n-3)\rho] + [(n+1)^2 + (n^2-5n-12)\rho - 3(n-5)\rho^2]a_1}{(n+2)(n+3)^2 + 2(n+3)(n^2+n-26)\rho + (n^3-8n^2-27n+178)\rho^2} . \tag{2.4}$$

Since the denominator in (2.4) is positive for every $\rho \in [0, 1]$ and $n \geqslant 3$, $a_2^*(a_1, \rho)$ is bounded. Let $\underline{a_2^*}(a_1) = \inf_{\rho \in [0,1)} a_2^*(a_1, \rho)$ and $\overline{a_2^*}(a_1) = \sup_{\rho \in [0,1)} a_2^*(a_1, \rho)$. Then we have the following.

Proposition 2.2. *(i) If $a_2 \notin A_2^*(a_1) = [\underline{a_2^*}(a_1), \overline{a_2^*}(a_1)]$ then $\hat{\lambda}_{a_1,a_2}$ is inadmissible being dominated by $\hat{\lambda}_{a_1,\underline{a_2^*}(a_1)}$ if $a_2 < \underline{a_2^*}(a_1)$ or $\hat{\lambda}_{a_1,\overline{a_2^*}(a_1)}$ if $a_2 > \overline{a_2^*}(a_1)$.*
(ii) In particular, if $a_1 \leqslant a_{11} = (n-7)/(n-1)$ then

$$\underline{a_2^*}(a_1) = a_2^*(a_1, 1) = \frac{(n+2)(n+3)}{2(n+5)}(1 - a_1) , \tag{2.5}$$

$$\overline{a_2^*}(a_1) = a_2^*(a_1, 0) = \frac{(n+1)^2}{n+3}\left(\frac{n-2}{n+1} - a_1\right) , \qquad (2.6)$$

whereas, if $a_1 \geqslant a_{12} = (n^3 + 2n^2 - 41n - 34)/[(n-1)(n^2 + 9n + 10)]$,

$$\underline{a_2^*}(a_1) = a_2^*(a_1, 0) , \quad \overline{a_2^*}(a_1) = a_2^*(a_1, 1) ,$$

where $a_2^*(a_1, 0)$, $a_2^*(a_1, 1)$ *are as in (2.5), (2.6), respectively.*

Proof. Part (i) is a consequence of the convexity of the mse in a_2. Part (ii) arises from the monotonicity of $a_2^*(a_1, \rho)$ with respect to ρ. Specifically, for $a_1 \leqslant a_{11}$, $a_2^*(a_1, \rho)$ is strictly decreasing in ρ whereas for $a_1 \geqslant a_{12}$ it is strictly increasing. This can be seen by examining the sign of the derivative of $a_2^*(a_1, \rho)$ with respect to ρ which is proportional to the quadratic

$$(n+3)[(n-1)(n^2 + 9n + 10)a_1 - (n^3 + 2n^2 - 41n - 34)]+$$
$$2[(n-1)(n^3 - 35n - 46)a_1 - (n+1)(n^3 - 8n^2 - 27n + 178)]\rho+$$
$$[(n-1)(n^3 - 4n^2 - 19n + 102)a_1 - (n-3)(n^3 - 8n^2 - 27n + 178)]\rho^2 .$$

The rest of the proof is elementary (although messy) and therefore omitted. $\qquad \square$

Remark 2.1. When $a_2 < \underline{a_2^*}(a_1)$, by the convexity of the mean squared error, $\hat{\lambda}_{a_1, a_2}$ is dominated not only by $\hat{\lambda}_{a_1, \underline{a_2^*}(a_1)}$, but by any estimator $\hat{\lambda}_{a_1, a_2'}$ with $a_2' \in (a_2, \underline{a_2^*}(a_1)]$ (a similar argument occurs when $a_2 > \overline{a_2^*}(a_1)$). Nevertheless, $\hat{\lambda}_{a_1, \underline{a_2^*}(a_1)}$ is the best among these estimators, therefore is the only one mentioned in Proposition 2.2.

In a similar way, by fixing a_2 and letting a_1 to vary, one can obtain an analogous result. In this case the mse is quadratic in a_1 and uniquely minimized in $a_1^*(a_2, \rho)$ given by

$$a_1^*(a_2, \rho) = \frac{(n+1)(n-2)(n-\rho) - [(n+1)^2 + (n^2 - 5n - 12)\rho - 3(n-5)\rho^2]a_2}{(n+1)[n(n+1) - 4(n+1)\rho + 6\rho^2]} .$$

The denominator is always positive, thus $a_1^*(a_2, \rho)$ is bounded for $\rho \in [0, 1]$. Setting $\underline{a_1^*}(a_2) = \inf_{\rho \in [0,1)} a_1^*(a_2, \rho)$ and $\overline{a_1^*}(a_2) = \sup_{\rho \in [0,1)} a_1^*(a_2, \rho)$, we derive the following.

Proposition 2.3. *(i) If* $a_1 \notin A_1^*(a_2) = [\underline{a_1^*}(a_2), \overline{a_1^*}(a_2)]$ *then* $\hat{\lambda}_{a_1, a_2}$ *is inadmissible being dominated by* $\hat{\lambda}_{\underline{a_1^*}(a_2), a_2}$ *if* $a_1 < \underline{a_1^*}(a_2)$ *or* $\hat{\lambda}_{\overline{a_1^*}(a_2), a_2}$ *if* $a_1 > \overline{a_1^*}(a_2)$.
(ii) In particular, if $a_2 \leqslant a_{21} = 3n(n+1)/[(n-1)(n+2)]$ *then*

$$\underline{a_1^*}(a_2) = a_1^*(a_2, 0) = \frac{n-2}{n+1} - \frac{a_2}{n} , \qquad (2.7)$$

$$\overline{a_1^*}(a_2) = a_1^*(a_2, 1) = 1 - \frac{2a_2}{n+1} \; , \tag{2.8}$$

whereas, if $a_2 \geqslant a_{22} = 3(n+1)/(n-1)$,

$$\underline{a_1^*}(a_2) = a_1^*(a_2, 1) \; , \quad \overline{a_1^*}(a_2) = a_1^*(a_2, 0) \; ,$$

where $a_1^*(a_2, 0)$, $a_1^*(a_2, 1)$ are as in (2.7), (2.8), respectively.

Propositions 2.2 and 2.3 provide necessary conditions for the admissibility of $\hat{\lambda}_{a_1, a_2}$ within $\mathcal{C}$ as stated in Corollary 2.1 below.

Corollary 2.1. *Two necessary conditions for the admissibility of $\hat{\lambda}_{a_1, a_2}$ within $\mathcal{C}$ are $a_1 \in A_1^*(a_2)$ and $a_2 \in A_2^*(a_1)$.*

Typically, unbiased estimators of scale parameters (as is λ for the distribution of S_1/S_2) are inadmissible in terms of mean squared error. In our case, the inadmissibility of the unbiased estimator $\hat{\lambda}_U$ follows from Proposition 2.2, since $a_{1U} > a_{12}$ and $a_{2U} > \overline{a_2^*}(a_{1U}) = (n+2)(n+3)/(n-1)(n+5)$.

Corollary 2.2. *The unbiased estimator $\hat{\lambda}_U$ is inadmissible in terms of mean squared error being dominated by*

$$\hat{\lambda}_U^* = \hat{\lambda}_{a_{1U}, \overline{a_2^*}(a_{1U})} = \left(\frac{n-3}{n-1} + \frac{(n+2)(n+3)}{(n-1)(n+5)} T \right) \frac{S_1}{S_2} \; . \tag{2.9}$$

Consider now the broader class of estimators

$$\mathcal{D} = \{ \hat{\lambda}_\phi = \phi(T) S_1/S_2 \} \; ,$$

where $\phi(\cdot)$ is any function such that $\hat{\lambda}_\phi$ has finite mse. Using Stein's (1964) technique, originally presented for improving the best equivariant estimator of a normal variance when the mean is unknown, one concludes that $\hat{\lambda}_U^*$ in (2.9) as well as a large subset of $\mathcal{C}$ are inadmissible estimators. To be specific, consider the conditional mean squared error of $\hat{\lambda}_\phi$ given $T = t$, $\mathbf{K} = \mathbf{k}$, $\mathbb{E}\left\{ [\phi(T) S_1/S_2 - \lambda]^2 \middle| T = t, \mathbf{K} = \mathbf{k} \right\}$, which is quadratic in $\phi(t)$ and uniquely minimized at

$$\phi_\mathbf{k}(t) = \frac{\lambda \, \mathbb{E}[S_1/S_2 | T = t, \mathbf{K} = \mathbf{k}]}{\mathbb{E}[S_1^2/S_2^2 | T = t, \mathbf{K} = \mathbf{k}]} = \frac{n+k-2}{n+k+1} = \phi_k^* \; ,$$

say. Note that it does not depend on t, since conditionally on $\mathbf{K} = \mathbf{k}$, S_1, S_2 and T are mutually independent. Moreover ϕ_k^* is strictly increasing in k with $\phi_0^* = (n-2)/(n+1)$

n	3	5	10	20	50	100
ρ	0.8463	0.8317	0.8433	0.8663	0.9004	0.9239

Table 1. Values of ρ for which $\hat{\lambda}_{\mathrm{mle}}$ and $\hat{\lambda}_U^*$ have equal mean squared errors.

and $\lim_{k\to\infty} \phi_k^* = 1$. As a consequence, each estimator of the form $\hat{\lambda}_\phi$ with $\mathbb{P}[\phi(T) \notin [(n-2)/(n+1), 1]] > 0$ is inadmissible being dominated by the estimator $\phi^*(T)S_1/S_2$, where $\phi^*(T) = \max\{(n-2)/(n+1), \min[\phi(T), 1]\}$. Application of the above argument to the class $\mathcal{C}$ leads to the following proposition.

Proposition 2.4. *(i) If $a_1 \notin [(n-2)/(n+1), 1]$ or $a_2 \notin [(n-2)/(n+1) - a_1, 1 - a_1]$, then the estimator $\hat{\lambda}_{a_1,a_2}$ is inadmissible being dominated by*

$$\max\{(n-2)/(n+1), \min[a_1 + a_2 T, 1]\}S_1/S_2 .$$

(ii) In particular, $\hat{\lambda}_U^*$ in (2.9) is dominated by $\hat{\lambda}_U^{**} = \min\{\hat{\lambda}_U^*, \hat{\lambda}_{\mathrm{mle}}\}$.

The mse of $\hat{\lambda}_U^{**}$ cannot be derived in a closed form, therefore an analytical comparison with $\hat{\lambda}_{\mathrm{mle}}$ is impossible. However, it is easy to compare the latter with $\hat{\lambda}_U^*$. Table 1 shows, for selected sample sizes, the corresponding values of the correlation coefficient for which both estimators have equal mean squared errors. When ρ is less than the reported value, $\hat{\lambda}_U^*$ is superior to $\hat{\lambda}_{\mathrm{mle}}$ and vice-versa. Since $\hat{\lambda}_U^{**}$ dominates $\hat{\lambda}_U^*$, it follows that for ρ less than the reported value, $\hat{\lambda}_U^{**}$ dominates $\hat{\lambda}_{\mathrm{mle}}$ as well. (In fact, a Monte Carlo study showed that $\hat{\lambda}_U^{**}$ and $\hat{\lambda}_{\mathrm{mle}}$ have equal mean squared errors when ρ is approximately 0.05 higher than the values given in Table 1.) It can be concluded that $\hat{\lambda}_U^{**}$ should be preferred, unless almost perfect linear correlation is suspected.

3 Estimation of the regression and the conditional variance

Consider now estimation of the regression of X on Y based on a random sample from (1.1). Downton (1970) showed that the conditional expectation of X given $Y = y$ is linear in y, specifically,

$$\eta(y) = \mathbb{E}[X|Y = y] = \frac{1 - \rho}{\lambda_1} + \rho \frac{\lambda_2}{\lambda_1} y .$$

Obviously, for deriving an unbiased estimator of $\eta(y)$ it suffices to derive unbiased estimators of $\eta_1 = (1 - \rho)/\lambda_1$ and $\eta_2 = \rho\lambda_2/\lambda_1$.

Proposition 3.1. *(i) The estimator*

$$\hat{\eta}_{1U} = \frac{2 - (n+1)T}{n-1} S_1$$

is unbiased for $\eta_1 = (1-\rho)/\lambda_1$.

(ii) The estimator

$$\hat{\eta}_{2U} = \frac{n+1}{n-1}(nT-1)\frac{S_1}{S_2}$$

is unbiased for $\eta_2 = \rho\lambda_2/\lambda_1$.

Proof. (i) The problem is similar to that of the derivation of $\hat{\lambda}_U$ in (2.3). We have to find c_0, c_1 such that $\mathbb{E}[(c_0 + c_1 T)S_1] = (1-\rho)\lambda_1^{-1}$. Using Lemma 4.2 (i), (ii), it can be seen that it suffices to solve the equations $nc_0 + c_1 = 1$, $\frac{n-1}{n+1}c_1 = -1$, for c_0 and c_1. The solution is $c_0 = 2/(n-1)$ and $c_1 = -(n+1)/(n-1)$, hence $\hat{\eta}_{1U}$ is an unbiased estimator of $\eta_1 = (1-\rho)/\lambda_1$.

(ii) Similarly, we need to find c_0, c_1 such that $\mathbb{E}[(c_0 + c_1 T)S_1/S_2] = \rho\lambda_2/\lambda_1$. Using (2.1) and (2.2), we get the equations

$$\begin{aligned}
\frac{n}{n-1}c_0 + \frac{1}{n-1}c_1 &= 0 \\
-\frac{1}{n-1}c_0 + \frac{n-3}{n^2-1}c_1 &= 1 \,,
\end{aligned}$$

whose solution is $c_0 = -(n+1)/(n-1)$ and $c_1 = n(n+1)/(n-1)$, yielding $\hat{\eta}_{2U}$ as an unbiased estimator of $\eta_2 = \rho\lambda_2/\lambda_1$. $\qquad\square$

Corollary 3.1. *The estimator*

$$\hat{\eta}_U(y) = \frac{2 - (n+1)T}{n-1}S_1 + \frac{n+1}{n-1}(nT-1)\frac{S_1}{S_2}y$$

is unbiased for $\eta(y)$.

The estimator $\hat{\eta}_U(y)$ is inadmissible for every y, since it assumes negative values with positive probability. A rather crude improved estimator is its positive part, $\hat{\eta}_U^+(y) = \max\{0, \hat{\eta}_U(y)\}$, which has smaller risk for any convex loss function. However, the same occurs for $\hat{\eta}_{1U}$ and $\hat{\eta}_{2U}$, and it seems rational to improve first on them and use their improvements to estimate the regression.

An estimator dominating $\hat{\eta}_{1U}$ in terms of mean squared error can be derived using Stein's (1964) technique. Consider the conditional mean squared error of estimators of

the form $\phi(T)S_1$ given $T = t$, $\mathbf{K} = \mathbf{k}$, $\mathbb{E}\left\{ \left[\phi(T)S_1 - (1-\rho)\lambda_1^{-1}\right]^2 \Big| T = t, \mathbf{K} = \mathbf{k} \right\}$, which is quadratic in $\phi(t)$ and uniquely minimized at

$$\phi_{\mathbf{k}}(t) = \frac{\lambda_1^{-1}(1-\rho)\mathbb{E}[S_1|T = t, \mathbf{K} = \mathbf{k}]}{\mathbb{E}[S_1^2|T = t, \mathbf{K} = \mathbf{k}]} = \frac{1}{n+k+1} = \phi_k^* ,$$

say. Now, ϕ_k^* is positive, attaining its maximum when $k = 0$, i.e. $0 < \phi_k^* \leqslant \phi_0^* = (n+1)^{-1}$. As a consequence, each estimator of the form $\phi(T)S_1$ with $\mathbb{P}[\phi(T) \notin [0, (n+1)^{-1}]] > 0$ is inadmissible being dominated by the estimator $\phi^*(T)S_1$, where $\phi^*(T) = \max\{0, \min[\phi(T), (n+1)^{-1}]\}$. Since $\mathbb{P}[(2 - (n+1)T)/(n-1) \notin [0, (n+1)^{-1}]] > 0$, $\hat{\eta}_{1U}$ is dominated by the estimator

$$\hat{\eta}_1^* = \begin{cases} S_1/(n+1) & , \quad T < (n+3)/(n+1)^2 , \\ \hat{\eta}_{1U} & , \quad (n+3)/(n+1)^2 \leqslant T \leqslant 2/(n+1) , \\ 0 & , \quad T > 2/(n+1) . \end{cases} \tag{3.1}$$

In a similar fashion we can improve on $\hat{\eta}_{2U}$. Note that it contains the quantity $nT - 1$, which is the estimator of ρ obtained by Nagao and Kadoya (1971) using the method of moments. Using the condition $0 \leqslant \rho < 1$, Al-Saadi and Young (1980) modified this estimator to

$$\tilde{\rho} = \begin{cases} 0 & , \quad T < 1/n , \\ nT - 1 & , \quad 1/n \leqslant T \leqslant 2/n , \\ 1 & , \quad T > 2/n . \end{cases}$$

The replacement of $nT - 1$ in $\hat{\eta}_{2U}$ by $\max\{nT - 1, 0\}$ leads to its positive part, $\hat{\eta}_{2U}^+ = \max\{0, \hat{\eta}_{2U}\}$, which is an improved estimator of $\rho\lambda_2/\lambda_1$. Replacement of $nT - 1$ by $\tilde{\rho}$ seems also reasonable, leading to the estimator

$$\tilde{\eta}_2 = \begin{cases} 0 & , \quad T < 1/n , \\ \hat{\eta}_{2U} & , \quad 1/n \leqslant T \leqslant 2/n , \\ \dfrac{n+1}{n-1}\dfrac{S_1}{S_2} & , \quad T > 2/n . \end{cases} \tag{3.2}$$

However, using Stein's (1964) technique we can find an estimator dominating all these estimators. Consider the class of estimators of $\rho\lambda_2/\lambda_1$ having the form $\psi(T)S_1/S_2$. The conditional mean squared error given $T = t$, $\mathbf{K} = \mathbf{k}$ of such an estimator is uniquely minimized with respect to $\psi(t)$ at

$$\psi_{\mathbf{k}}(t) = \frac{\rho\lambda_2\lambda_1^{-1}\mathbb{E}[S_1/S_2|T = t, \mathbf{K} = \mathbf{k}]}{\mathbb{E}[S_1^2/S_2^2|T = t, \mathbf{K} = \mathbf{k}]} = \rho\frac{n+k-2}{n+k+1} = \psi_k^*(\rho) ,$$

11

ρ	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$n = 10$ η_1^*	39.4	41.8	43.1	44.3	46.2	49.4	53.9	58.9	64.1	67.7
η_2^*	46.9	45.4	43.3	42.1	42.3	44.3	47.9	53.0	59.4	65.2
$n = 20$ η_1^*	32.1	31.5	27.9	25.8	27.5	32.7	41.0	50.5	58.7	64.9
η_2^*	42.8	36.4	28.9	24.8	25.4	29.4	36.7	45.6	54.5	62.5
$n = 50$ η_1^*	28.1	22.1	11.6	6.5	7.5	13.2	22.5	34.2	47.4	58.7
η_2^*	43.6	27.9	12.5	6.3	6.9	11.8	20.0	30.7	43.8	56.4

Table 2. Simulated percentage risk improvement of the mean squared error of $\hat\eta_1^*$ in (3.1), $\hat\eta_2^*$ in (3.3) over $\hat\eta_{1U}$, $\hat\eta_{2U}$ respectively.

say. Since $0 \leqslant \psi_k^*(\rho) \leqslant \rho < 1$, any estimator of the form $\psi(T)S_1/S_2$ satisfying $\mathbb{P}[\psi(T) \notin [0,1]] > 0$ is inadmissible. Indeed, it is dominated by $\psi^*(T)S_1/S_2$, where $\psi^*(T) = \max\{0, \min[\psi(T), 1]\}$. Thus, $\hat\eta_{2U}$, $\hat\eta_{2U}^+$ are dominated by

$$
\hat\eta_2^* = \begin{cases} 0 & , \ T < 1/n \ , \\ \hat\eta_{2U} & , \ 1/n \leqslant T \leqslant 2/(n+1) \ , \\ S_1/S_2 & , \ T > 2/(n+1) \ . \end{cases} \tag{3.3}
$$

From (3.2) and (3.3), it is obvious that $\hat\eta_2^*$ dominates also $\tilde\eta_2$.

Remark. The estimators $\hat\eta_1^*$, $\hat\eta_2^*$ in (3.1), (3.3) respectively, have the property of "pre-testing" for ρ. For example, when T is "small" (smaller than $(n+3)/(n+1)^2$), indicating $\rho = 0$, $\hat\eta_1^*$ equals to the best equivariant estimator of $1/\lambda_1$ with respect to squared error loss, $S_1/(n+1)$. On the other hand, when T is "large" (greater than $2/(n+1)$), indicating ρ to be very close to one, $\hat\eta_1^*$ equals zero. Analogous comments hold for $\hat\eta_2^*$.

The percentage improvements in terms of mean squared error of the estimators $\hat\eta_1^*$, $\hat\eta_2^*$ over $\hat\eta_{1U}$, $\hat\eta_{2U}$ respectively, have been evaluated by Monte Carlo sampling from (1.1), for sample sizes $n = 10, 20, 50$ and $\rho = 0(.1).9$. We have taken 10000 replications for each pair (n, ρ). The results are shown in Table 2. It can be seen that the improvements are remarkable even for $n = 50$. Generally, they are larger for extreme values of ρ. This can be explained by the nature of the improved estimators as indicated in the above remark.

The conditional variance of X given $Y = y$ is also linear in y. Specifically,

$$
\theta(y) = \text{Var}(X|Y = y) = \left(\frac{1-\rho}{\lambda_1}\right)^2 + 2\rho(1-\rho)\frac{\lambda_2}{\lambda_1^2} y \ .
$$

Let $\theta_1 = (1-\rho)^2\lambda_1^{-2}$, $\theta_2 = 2\rho(1-\rho)\lambda_2\lambda_1^{-2}$. Then we have the following proposition.

Proposition 3.2. *(i) The estimator $\hat{\theta}_{1U} = h_1(T)S_1^2$, where*

$$h_1(T) = \frac{4(n+5) - 4(n+1)(n+5)T + (n+1)(n+2)(n+3)T^2}{(n-1)(n^2+5n+2)} \, ,$$

is unbiased for $\theta_1 = (1-\rho)^2\lambda_1^{-2}$.

(ii) The estimator $\hat{\theta}_{2U} = h_2(T)S_1^2/S_2$, where

$$h_2(T) = \frac{-4(n^2+7n+8) + 2(n+1)(3n^2+19n+18)T - 2(n+1)^2(n+2)(n+3)T^2}{(n-1)(n^2+5n+2)} \, ,$$

is unbiased for $\theta_2 = 2\rho(1-\rho)\lambda_2\lambda_1^{-2}$.

Proof. Similarly to the proof of Proposition 3.1, the problem reduces in finding c_0, c_1, c_2, d_0, d_1, d_2 such that $\mathbb{E}[(c_0 + c_1T + c_2T^2)S_1^2] = \theta_1$, $\mathbb{E}[(d_0 + d_1T + d_2T^2)S_1^2/S_2] = \theta_2$ for part (i), (ii) respectively. Using Lemma 4.2 and equating the coefficients of the appropriate second degree polynomials in ρ, we obtain $\hat{\theta}_{1U}$, $\hat{\theta}_{2U}$ as unbiased estimators of θ_1, θ_2. $\quad\square$

Corollary 3.2. *The estimator*

$$\hat{\theta}_U(y) = h_1(T)S_1^2 + y\,h_2(T)S_1^2/S_2$$

is unbiased for $\theta(y)$.

The estimators $\hat{\theta}_{1U}$, $\hat{\theta}_{2U}$ and hence $\hat{\theta}(y)$ assume negative values with positive probability. As in the estimation problem of $\eta(y)$, we can improve on them by truncating h_1, h_2 in suitable intervals. Omitting the details, an estimator of $\theta_1 = (1-\rho)^2\lambda_1^{-2}$ of the form $\phi(T)S_1^2$ satisfying $\mathbb{P}[\phi(T) \notin [0, 1/(n+2)(n+3)]] > 0$ is dominated by $\phi^*(T)S_1^2$ where $\phi^*(T) = \max\{0, \min[\phi(T), 1/(n+2)(n+3)]\}$, whereas an estimator of $\theta_2 = 2\rho(1-\rho)\lambda_2\lambda_1^{-2}$ of the form $\psi(T)S_1^2/S_2$ with $\mathbb{P}[\psi(T) \notin [0, 2(n-2)/(n+2)(n+3)]] > 0$ is dominated by $\psi^*(T)S_1^2/S_2$ where $\psi^*(T) = \max\{0, \min[\psi(T), 2(n-2)/(n+2)(n+3)]\}$, provided $n \geqslant 6$. The functions h_1, h_2 satisfy the above conditions for $n \geqslant 3$, thus $\hat{\theta}_{1U}$, $\hat{\theta}_{2U}$ are dominated by suitable estimators.

4 Appendix

Lemma 4.1. *Let $K_1, K_2, \ldots, K_n$ be a random sample from a geometric distribution with probability mass function*

$$\pi_1(k_1; \rho) = \mathbb{P}(K_1 = k_1; \rho) = (1-\rho)\rho^{k_1} \, , \quad k_1 = 0, 1, 2, \ldots \, . \tag{4.1}$$

and $K = \sum K_i$. Then,

(i) $\mathbb{P}(K_1 = k_1 | K = k) = \binom{n-2+k-k_1}{k-k_1} \binom{n+k-1}{k}^{-1}, \; 0 \leqslant k_1 \leqslant k,$

(ii) $\mathbb{P}(K_1 = k_1, K_2 = k_2 | K = k) = \binom{n-3+k-k_1-k_2}{k-k_1-k_2} \binom{n+k-1}{k}^{-1}, 0 \leqslant k_1, k_2, k_1 + k_2 \leqslant k,$

(iii) $\mathbb{E}[(K_1 + 1)^2 | K = k] = 1 + \frac{3n+1}{n(n+1)} k + \frac{2}{n(n+1)} k^2,$

(iv) $\mathbb{E}[(K_1 + 1)^2 (K_1 + 2)^2 | K = k] =$

$$4 \left(1 + \frac{8n^2+13n+3}{n(n+1)(n+3)} k + \frac{19n^2+41n+18}{n(n+1)(n+2)(n+3)} k^2 + \frac{18}{n(n+2)(n+3)} k^3 + \frac{6}{n(n+1)(n+2)(n+3)} k^4 \right),$$

(v) $\mathbb{E}[(K_1 + 1)^2 (K_2 + 1)^2 | K = k] =$

$$1 + \frac{(2n+3)(3n+1)}{n(n+1)(n+3)} k + \frac{13n^2+29n+14}{n(n+1)(n+2)(n+3)} k^2 + \frac{12}{n(n+2)(n+3)} k^3 + \frac{4}{n(n+1)(n+2)(n+3)} k^4,$$

(vi) $\mathbb{E}K = \dfrac{n\rho}{1-\rho}, \; \mathbb{E}K^2 = \dfrac{n\rho(1+n\rho)}{(1-\rho)^2},$

(vii) $\mathbb{E}\left[\dfrac{n+K}{n+K-1}\right] = \dfrac{n-\rho}{n-1}.$

Proof. Parts (i), (ii) are applications of the Bayes Theorem, whereas parts (iii)–(vi) are straightforward. We will prove only part (vii).

Since $K = \sum K_i$ follows a negative binomial distribution with probability mass function

$$\pi_n(k\,;\rho) = \binom{n+k-1}{k} \rho^k (1-\rho)^n \; , \;\; k = 0, 1, 2, \dots \;,$$

one has

$$
\begin{aligned}
\mathbb{E}\left[\frac{n+K}{n+K-1}\right] &= \sum_{k=0}^{\infty} \frac{n+k}{n+k-1} \binom{n+k-1}{k} \rho^k (1-\rho)^n \\
&= \sum_{k=0}^{\infty} (n+k) \frac{(n+k-2)!}{k!(n-1)!} \rho^k (1-\rho)^n \\
&= \frac{n(1-\rho)}{n-1} \sum_{k=0}^{\infty} \binom{n-1+k-1}{k} \rho^k (1-\rho)^{n-1} + \rho \sum_{k=0}^{\infty} \binom{n+k-1}{k} \rho^k (1-\rho)^n \\
&= \frac{n(1-\rho)}{n-1} + \rho = \frac{n-\rho}{n-1} \;.
\end{aligned}
$$

$\square$

Lemma 4.2. *Let* $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ *be a random sample from (1.1), and* $S_1 = \sum X_i, \; S_2 = \sum Y_i, \; T = \sum X_i Y_i / (S_1 S_2)$. *Then,*

14

(i) $\mathbb{E}[S_1] = n\,\lambda_1^{-1}$,

(ii) $\mathbb{E}[TS_1] = \left(1 + \frac{n-1}{n+1}\,\rho\right)\lambda_1^{-1}$,

(iii) $\mathbb{E}[S_1^2] = n(n+1)\,\lambda_1^{-2}$,

(iv) $\mathbb{E}[TS_1^2] = \left(n+1 + \frac{(n-1)(n+2)}{n+1}\,\rho - \frac{n-1}{n+1}\,\rho^2\right)\lambda_1^{-2}$,

(v) $\mathbb{E}[T^2S_1^2] = \left(\frac{n+3}{n+1} + \frac{2(n-1)(n+6)}{(n+1)(n+2)}\,\rho + \frac{(n-1)(n^2+n-18)}{(n+1)(n+2)(n+3)}\,\rho^2\right)\lambda_1^{-2}$,

(vi) $\mathbb{E}[S_1^2/S_2] = \left(\frac{n(n+1)}{n-1} - \frac{2(n+1)}{n-1}\,\rho + \frac{2}{n-1}\,\rho^2\right)\lambda_2\,\lambda_1^{-2}$,

(vii) $\mathbb{E}[TS_1^2/S_2] = \left(\frac{n+1}{n-1} + \frac{n^2-2n-7}{n^2-1}\,\rho - \frac{2(n-3)}{n^2-1}\,\rho^2\right)\lambda_2\,\lambda_1^{-2}$,

(viii) $\mathbb{E}[T^2S_1^2/S_2] = \left(\frac{n+3}{n^2-1} + \frac{2(n^2+3n-16)}{(n^2-1)(n+2)}\,\rho + \frac{n^3-4n^2-27n+78}{(n^2-1)(n+2)(n+3)}\,\rho^2\right)\lambda_2\,\lambda_1^{-2}$,

(ix) $\mathbb{E}[S_1^2/S_2^2] = \left(\frac{n(n+1)}{(n-1)(n-2)} - \frac{4(n+1)}{(n-1)(n-2)}\,\rho + \frac{6}{(n-1)(n-2)}\,\rho^2\right)\lambda_2^2\,\lambda_1^{-2}$,

(x) $\mathbb{E}[TS_1^2/S_2^2] = \left(\frac{n+1}{(n-1)(n-2)} + \frac{n^2-5n-12}{(n^2-1)(n-2)}\,\rho - \frac{3(n-5)}{(n^2-1)(n-2)}\,\rho^2\right)\lambda_2^2\,\lambda_1^{-2}$,

(xi) $\mathbb{E}[T^2S_1^2/S_2^2] = \left(\frac{n+3}{(n^2-1)(n-2)} + \frac{2(n^2+n-26)}{(n^2-1)(n^2-4)}\,\rho + \frac{n^3-8n^2-27n+178}{(n^2-1)(n^2-4)(n+3)}\,\rho^2\right)\lambda_2^2\,\lambda_1^{-2}$,

Proof. The marginal distribution of S_1 is Gamma$(n, 1/\lambda_1)$, thus (i), (iii) are immediate. From the rest, we will prove only (v) since the proofs of the other parts are similar.

Let $\mathbf{K} = (K_1, \ldots, K_n)$ be a random sample from the geometric distribution (4.1) and set $U_i = X_i S_1^{-1}$, $V_i = Y_i S_2^{-1}$, $i = 1, \ldots, n$. Then

$$
\begin{aligned}
\mathbb{E}[T^2|\mathbf{K} = \mathbf{k}] &= \mathbb{E}\left[\left(\sum_{i=1}^{n} U_iV_i\right)^2 \middle| \mathbf{K} = \mathbf{k}\right] \\
&= \mathbb{E}\left[\left(\sum_{i=1}^{n} U_i^2V_i^2 + \sum_{i=1}^{n}\sum_{\substack{j=1\\j\neq i}}^{n} U_iV_iU_jV_j\right)\middle| \mathbf{K} = \mathbf{k}\right] \\
&= (n+k)^{-2}(n+k+1)^{-2} \\
&\quad \times \left[\sum_{i=1}^{n}(k_i+1)^2(k_i+2)^2 + \sum_{i=1}^{n}\sum_{\substack{j=1\\j\neq i}}^{n}(k_i+1)^2(k_j+1)^2\right] , \\
\mathbb{E}[S_1^2|\mathbf{K} = \mathbf{k}] &= (n+k)(n+k+1)(1-\rho)^2\lambda_1^{-2} ,
\end{aligned}
$$

15

yielding

$$
\begin{aligned}
\mathbb{E}[T^2 S_1^2] &= \mathbb{E}[\mathbb{E}(T^2 S_1^2 | \mathbf{K})] = \mathbb{E}[\mathbb{E}(T^2 | \mathbf{K}) \mathbb{E}(S_1^2 | \mathbf{K})] \\
&= \mathbb{E}\left\{ \mathbb{E}\left[\left. \frac{n(K_1 + 1)^2(K_1 + 2)^2 + n(n-1)(K_1 + 1)^2(K_2 + 1)^2}{(n + K)(n + K + 1)} \right| K \right] \right\} \left(\frac{1 - \rho}{\lambda_1} \right)^2 \\
&= \mathbb{E}\left[\frac{n + 3}{n + 1} + \frac{4(n + 5)}{(n + 1)(n + 2)(n + 3)} K + \frac{4(n + 5)}{(n + 1)(n + 3)} K^2 \right] \left(\frac{1 - \rho}{\lambda_1} \right)^2 .
\end{aligned}
$$

Here the last equality follows from Lemma 4.1(iv), (v). Substituting the moments of K in the last expression from Lemma 4.1(vi), we obtain the desired result. $\qquad\square$

Acknowledgment

References

Al-Saadi, S. D., Scrimshaw, D. G. and Young, D. H. (1979). Tests for independence of exponential variables. *J. Statist. Comput. Simul.*, **9**, 217–233.

Al-Saadi, S. D. and Young, D. H. (1980). Estimators for the correlation coefficient in a bivariate exponential distribution. *J. Statist. Comput. Simul.*, **11**, 13–20.

Al-Saadi, S. D. and Young, D. H. (1982). A test for independence in a multivariate exponential distribution with equal correlation coefficient. *J. Statist. Comput. Simul.*, **14**, 219–227.

Balakrishnan, N. and Ng, H. K. T. (2001). Improved estimation of the correlation coefficient in a bivariate exponential distribution. *J. Statist. Comput. Simul.*, **68**, 173–184.

Downton, F. (1970). Bivariate exponential distributions in reliability theory. *J. Roy. Statist. Soc.* B, **32**, 408–417.

Gelfand, A. E. and Dey, D. K. (1988). On the estimation of a variance ratio. *J. Statist. Plann. Inference*, **19**, 121–131.

Ghosh, M. and Kundu, S. (1996). Decision theoretic estimation of the variance ratio. *Statist. Decisions*, **14**, 161–175.

Iliopoulos, G. (2001). Decision theoretic estimation of the ratio of variances in a bivariate normal distribution. *Ann. Inst. Statist. Math.*, **53**, 436–446.

Kibble, W. F. (1941). A two-variate gamma type distribution. *Sankhyā*, **5**, 137–150.

Kotz, S., Balakrishnan, N. and Johnson, N. L. (2000). *Continuous Multivariate Distributions*, **1**, Second edition. New York, Wiley.

Kubokawa, T. (1994). Double shrinkage estimation of ratio of scale parameters. *Ann. Inst. Statist. Math.*, **46**, 95–116.

Kubokawa, T. and Srivastava, M. S. (1996). Double shrinkage estimators of ratio of variances. *Multidimensional Statistical Analysis and Theory of Random Matrices* (eds. A.K. Gupta and V.L. Girko), 139–154, VSP, Netherlands.

Madi, T. M. (1995). On the invariant estimation of a normal variance ratio. *J. Statist. Plann. Inference*, **44**, 349–357.

Madi, T. M. and Tsui, K. W. (1990). Estimation of the ratio of the scale parameters of two exponential distributions with unknown location parameters. *Ann. Inst. Statist. Math.*, **42**, 77–87.

Moran, P. A. P. (1967). Testing for correlation between non-negative variates. *Biometrika*, **54**, 385–394.

Nagao, M. and Kadoya, M. (1971). Two-variate exponential distribution and its numerical table for engineering application. *Bulletin of the Disaster Prevention Institute*, Kyoto University, **20**, 183–215.

Stein, C. (1964). Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. *Ann. Inst. Statist. Math.*, **16**, 155–160.