Asymptotic Expansions for Scale Mixtures of F-Distribution and Their Error Bounds

Yasunori Fujikoshi* and Ryoichi Shimizu**

**Department of Mathematics, Hiroshima University*
1-3-1 Kagamiyama, Higashi Hiroshima, Hiroshima 739-8626, Japan

***The Institute of Statistical Mathematics*
10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan

Abstract

A scale mixture variable is defined by $X = SZ$ or $X = S^{-1}Z$, where the scale factor S is a positive random variable, and X and S are independent. When Z is normal or chi-square distribution, asymptotic expansions of X and their error bounds have been extensively studied. Some of the results are found in Fujikoshi et al. (2010). In this paper we give some basic results on asymptotic expansions and their error bounds when Z is F-distribution. This paper is based on Fujikoshi and Shimizu (2009).

AMS 2000 subject classification: primary 62E20; secondary 62H10

Key Words and Phrases: Asymptotic expansions, Error bounds, F-distribution, Scale mixtures.

1. Introduction

In this paper we consider a scale mixture of F-distribution which is defined by

$$X_{q,n} = S_n^\delta Z_{q,n}, \tag{1.1}$$

where $Z_{q,n}/q$ is a random variable with an F-distribution with (q, n) degrees of freedom, S_n is a positive random variable, and S_n and $Z_{q,n}$ are independent. Here $\delta = \pm 1$ is a constant used to distinguish two types of scale mixtures: $X_{q,n} = S_n Z_{q,n}$ and $X_{q,n} = S_n^{-1} Z_{q,n}$. When $Z_{q,n}$ is distributed to the standard normal $N(0, 1)$ or a chi-square distribution χ_q^2 with q degrees of freedom, asymptotic expansions of $X_{q,n}$ and their error bounds have been extensively studied. For the results, see, e.g., Fujikoshi et al. (2010).

A scale mixture of F-distribution appears, for example, in profile analysis (see, Fujikoshi et al. (2010), Srivastava (2012)). In fact, consider profile analysis of k p-variate normal populations based on N samples. Then, it is known that a simultaneous confidence interval for differences in the levels of k profiles is based on a statistic T whose distribution is expressed as

$$T = \left(1 + \frac{\chi_{p-1}^2}{\chi_{n+1}^2}\right) \frac{\chi_q^2}{\chi_n^2/n}, \tag{1.2}$$

where $q = k - 1$ and $n = N - k - p + 1$. Here χ_q^2, χ_n^2, χ_{p-1}^2 and χ_{n+1}^2 are independent. Then, we can express T as a sale mixture of F-distribution in two ways with

$$\delta = 1, \quad S_n = 1 + \frac{\chi_{p-1}^2}{\chi_{n+1}^2}, \quad Z_{q,n} = \frac{\chi_q^2}{\chi_n^2/n}, \tag{1.3}$$

and

$$\delta = -1, \quad S_n = \frac{\chi_{n+1}^2}{\chi_{n+1}^2 + \chi_{p-1}^2}, \quad Z_{q,n} = \frac{\chi_q^2}{\chi_n^2/n}, \tag{1.4}$$

If we use (1.3) with $\delta = 1$, $S_n > 1$, and if we us (1.4) with $\delta = -1$, $S_n < 1$.

In this paper we gives asymptotic expansions of $X_{q,n}$ in (1.1) and their error bounds. The results are applied to the distribution of T in (1.2).

2. Preliminaries

With $\delta = \pm 1$, consider a scale mixture defined by $X = S^{\delta} Z$. Let G be the distribution function of Z. Depending on whether Z is symmetric or not, there are two types of asymptotic approximations and their error bounds. Here, since we are interesting in scale mixtures of F distribution, we summarize the main results on the case when Z is not symmetric.

Suppose that Z has a cumulative distribution function (cdf) $G(x)$. Then, the cdf of the scale mixture $X = S^{\delta} Z$ is given by

$$F(x) \equiv P(X \leq x) = P(Z \leq xS^{-\delta}) = \mathrm{E}_S\{G(xS^{-\delta})\}. \tag{2.1}$$

Assuming that the scale factor S is close to 1 in some sense, we consider approximating the cdf $F(x)$. Our interest also lies in evaluating possible errors of approximations. We assume without further notice that $G(x)$ is $k+1$ times continuously differentiable on its support $D = \{x \in R : g(x) > 0\}$ and that the scale factor S and its reciprocal S^{-1} have moments of required order.

Let $g(x)$ be the probability density function (pdf) of Z. For any $y > 0$, the conditional distribution of $X = S^{\delta} Z$ given $S = y$ has a cdf that is expressible as $G(xy^{-\delta})$, and mathematical induction shows that its derivatives with respect to y can be put in the form

$$\frac{\partial^j G(xy^{-\delta})}{\partial y^j} = y^{-j} c_{\delta,j}(xy^{-\delta}) g(xy^{-\delta}), \quad j = 1, 2, \ldots, k, \tag{2.2}$$

where the $c_{\delta,j}$ are real functions obtained by

$$c_{\delta,j}(x) g(x) = \left. \frac{\partial^j G(xy^{-\delta})}{\partial y^j} \right|_{y=1}. \tag{2.3}$$

We can use

$$G_{\delta,k}(x,y) = G(x) + \sum_{j=1}^{k-1} \frac{1}{j!} c_{\delta,j}(x) g(x) (y-1)^j \tag{2.4}$$

as an approximation to $G(xy^{-\delta})$, which would in turn induce an approximation to $F(x)$ by means of

$$G_{\delta,k}(x) \equiv \mathrm{E}\{G_{\delta,k}(x,S)\} = G(x) + \sum_{j=1}^{k-1} \frac{1}{j!} c_{\delta,j}(x) g(x) \mathrm{E}\{(S-1)^j\}. \tag{2.5}$$

For $x \in D$ and $y > 0$, write

$$\Delta_{\delta,k}(x,y) \equiv G(xy^{-\delta}) - G_{\delta,k}(x,y) \tag{2.6}$$

and

$$J_{\delta,k}(x,y) \equiv \begin{cases} \dfrac{|\Delta_{\delta,k}(x,y)|}{|y-1|^k}, & \text{for } y \neq 1, \\ \\ \dfrac{1}{k!}|c_{\delta,k}(x)|\; g(x), & \text{for } y = 1. \end{cases} \tag{2.7}$$

By applying Taylor's theorem to $G(xy^{-\delta})$ as a function of y and using (2.4), the remainder term $\Delta_{\delta,k}(x,y)$ can be put in the form

$$\Delta_{\delta,k}(x,y) = \frac{1}{k!} c_{\delta,k}(u) g(u) y_0^{-k} (y-1)^k, \tag{2.8}$$

where $u = xy_0^{-\delta/\rho}$, and y_0 is a positive number lying between 1 and y.

We define the positive constants $\alpha_{\delta,k}$ and $\beta_{\delta,k}$ by

$$\begin{aligned} \alpha_{-1,0} &= \max\left\{1 - G(0),\ G(0)\right\}, \\ \alpha_{\delta,k} &= \frac{1}{k!} \sup_x \left|c_{\delta,k}(x) g(x)\right|, \\ \beta_{\delta,k} &= \sup_{x \in D, y \leq 1} J_{\delta,k}(x,y). \end{aligned} \tag{2.9}$$

It follows from (2.8) that we have the inequalities $\alpha_{\delta,k} \leq \beta_{\delta,k}$ and

$$\sup_x \left|\Delta_{\delta,k}(x,y)\right| \leq \begin{cases} \alpha_{\delta,k}\ \left(y \vee y^{-1} - 1\right)^k, \\ \beta_{\delta,k}\ \left|y - 1\right|^k, \end{cases} \tag{2.10}$$

where $y \vee y^{-1} = \max\{y, y^{-1}\}$. From (2.8) we obtain the following upper bound for the approximation error:

$$\begin{aligned} \left|F(x) - G_{\delta,k}(x)\right| &\leq \alpha_{\delta,k} \mathrm{E}\{(S \vee S^{-1} - 1)^k\}, \\ &\leq \alpha_{\delta,k} \left[\mathrm{E}\left\{|S-1|^k\right\} + \mathrm{E}\left\{|S^{-1}-1|^k\right\}\right]. \end{aligned} \tag{2.11}$$

If $S > 1$, we obtain

$$\left|F(x) - G_{\delta,k}(x)\right| \leq \alpha_{\delta,k} \mathrm{E}\left\{|S-1|^k\right\}. \tag{2.12}$$

Further, it holds that

$$\left|F(x) - G_{\delta,k}(x)\right| \leq \beta_{\delta,k} \mathrm{E}\left\{|S-1|^k\right\}. \tag{2.13}$$

Note that for use of (2.11), it is required that both of $\mathrm{E}\{(S-1)^k\}$ and $\mathrm{E}\{(S^{-1}-1)^k\}$ exist. For use of (2.12) and (2.13), it is required that $\mathrm{E}\{(S-1)^k\}$ only exists .

3. Main Results

In this section we consider a scale mixture of F-distribution given by (1.1). Let $F_q(x;n)$ and $f_q(x;n)$ be the cdf and the pdf of $Z_{q,n}$. Then, since $Z_{q,n}/q$ is an F-distribution with (q,n) degrees of freedom, with $\delta = \pm 1$, the density is given by

$$f_q(x;\, n) = B_0(q,n)\frac{1}{n}\Big(\frac{x}{n}\Big)^{q/2-1}\Big(1+\frac{x}{n}\Big)^{-(q+n)/2},$$

where

$$B_0(q,n) \equiv \frac{\Gamma((q+n)/2)}{\Gamma(q/2)\,\Gamma(n/2)}.$$

This means that $Z_{q,n}/q$ follows an F-distribution $F(q,n)$ with (q,n) degrees of freedom. For an example of a sale mixture of F-distribution, see Section 2. Using (2.5), we have an approximation to the cdf $G_q(x;\, n)$ of the distribution of $X_{q,n}$ for large n. To this end, write $Y = S_n$ and consider

$$F_q(xy^{-\delta};n) \equiv \int_0^{xy^{-\delta}} f_q(x;\, n)dx,$$

which is the conditional cdf given the condition $Y = y$. Note that the cdf of $X_{q,n}$ is expressed as

$$G_q(x;n) \equiv P(X_{q,n} \leq x) = \mathrm{E}\{F_q(xS_n^{-\delta};n)\}. \tag{3.1}$$

The function $F_q(xy^{-\delta}; n)$ is formally approximated by

$$G_{\delta,k;q}(x,y;n) = F_q(x;n) - \delta \sum_{j=1}^{k-1} \frac{1}{j!} c_{\delta,j,q}(x;n)(y-1)^j f_q(x;n),$$

where $c_{\delta,j,q}(x;n)$'s are defined by

$$c_{\delta,j,q}(x;n) f_q(x;n) = -\frac{\partial^j F_q(xy^{-\delta};n)}{\partial y^j}\bigg|_{y=1}. \tag{3.2}$$

Here, note that the definition of $c_{\delta,j,q}(x;n)$ is different from the one of $c_{\delta,j}(x)$ (see (2.3)) in their signs. The approximation $G_{\delta,k;q}(x;y,n)$ induces an approximation to $F_q(x;n)$ by means of

$$G_{\delta,k;q}(x;n) = F_q(x;n) - \delta \sum_{j=1}^{k-1} \frac{1}{j!} c_{\delta,j}(x) f_q(x;x) \mathrm{E}\{(S-1)^j\}. \tag{3.3}$$

Naturally, the functions $c_{\delta,j,q}(x;n)$'s and the approximating function $G_{\delta,k;q}(x;n)$ depend on the choice of $\delta = 1$ and $\delta = -1$. We use the notation $a^{(\ell)}$ for a nonzero a and an integer ℓ to mean

$$a^{(\ell)} = a \cdot (a-1) \cdot \cdots \cdot (a-\ell-1), \quad \text{if} \quad \ell > 0, \quad \text{and} \quad a^{(0)} = 1.$$

Assume first that $\delta = 1$. Then, the defining equation (3.2) of $c_{1,j,q}(x;n)$ is equivalent to

$$\begin{aligned} c_{1,j,q}(x;n) &= x\Big(1+\frac{x}{n}\Big)^{(q+n)/2} \frac{\partial^{j-1}}{\partial y^{j-1}} y^{n/2-1}\Big(y+\frac{x}{n}\Big)^{-(q+n)/2}\bigg|_{y=1} \\ &= x\sum_{i=0}^{j-1} \binom{j-1}{i} \Big(\frac{n-2}{2}\Big)^{(j-i-1)} \Big(\frac{-(q+n)}{2}\Big)^{(i)} \Big(1+\frac{x}{n}\Big)^{i}, \end{aligned} \tag{3.4}$$

which gives, for example,

$$
\begin{aligned}
c_{1,1,q}(x;n) &= x,\\
c_{1,2,q}(x;n) &= \frac{x}{2(x+n)}[(n-2)x - n(q+2)],\\
c_{1,3,q}(x;n) &= \frac{x}{4(x+n)^2}\big[(n-2)(n-4)x^2 - 2(n-2)n(q+4)x\\
&\quad +n^2(q+2)(q+4))\big]. \qquad (3.5)\\
c_{1,4,q}(x;n) &= \frac{x}{8(x+n)^3}\big[(n-2)(n-4)(n-6)x^3\\
&\quad -3n(n-2)(n-4)(6+q)x^2\\
&\quad +3(-2+n)n^2(q+4)(q+6)x - n^3(q+2)(q+4)(q+6))\big].
\end{aligned}
$$

If $\delta = -1$, then, the defining equation (3.2) of $c_{-1,j,q}(x;n)$ is equivalent to equation (3.4), which leads to

$$
\begin{aligned}
c_{-1,j,q}(x;n) &= x\Big(1+\frac{x}{n}\Big)^{(q+n)/2} \frac{\partial^{j-1}}{\partial y^{j-1}} y^{n/2-1}\Big(y+\frac{x}{n}\Big)^{-(q+n)/2}\Bigg|_{y=1}\\
&= x\sum_{i=0}^{j-1}\binom{j-1}{i}\Big(\frac{q-2}{2}\Big)^{(j-i-1)}\Big(\frac{-(q+n)}{2}\Big)^{(i)}\Big(1+\frac{n}{x}\Big)^{i}, \qquad (3.6)
\end{aligned}
$$

which gives, for example,

$$
\begin{aligned}
c_{-1,1,q}(x;n) &= x,\\
c_{-1,2,q}(x;n) &= \frac{x}{2(x+n)}[(q-2)(x+n) - (q+n)x],\\
c_{-1,3,q}(x;n) &= \frac{x}{4(x+n)^2}\big[(q-2)(q-4)(x+n)^2 - 2\,(q-2)(q+n)x(x+n)\\
&\quad +(q+n)(q+n+2)x^2\big], \qquad (3.7)\\
c_{-1,4,q}(x;n) &= \frac{x}{8(x+n)^3}\big[(q-2)(q-4)(q-6)(x+n)^3\\
&\quad -3(q-2)(q-4)(q+n)x(x+n)^2\\
&\quad +3(q-2)(q+n)(q+n+2)x^2(x+n)\\
&\quad -(q+n)(q+n+2)(q+n+4)x^3\big].
\end{aligned}
$$

Setting

$$
\alpha_{\delta,k;q}(n) = \sup_x \Big|\frac{1}{k!}c_{\delta,k,\,q}(x;n)f_q(x;n)\Big|, \qquad (3.8)
$$

from (2.10) we obtain the following upper bound for the approximation error:

$$\begin{aligned}\left|G_q(x;n) - G_{\delta,k;q}(x;n)\right| &\le \alpha_{\delta,k;q}(n)\mathrm{E}\{(S_n \vee S_n^{-1} - 1)^k\}, \\ &\le \alpha_{\delta,k;q}(n)\left[\mathrm{E}\left\{|S_n - 1|^k\right\} + \mathrm{E}\left\{|S_n^{-1} - 1|^k\right\}\right]. \end{aligned} \tag{3.9}$$

If $S_n > 1$, we obtain

$$\left|G_q(x;n) - G_{\delta,k;q}(x;n)\right| \le \beta_{\delta,k;q}(n)\mathrm{E}\left\{|S_n - 1|^k\right\}. \tag{3.10}$$

Further, it holds that

$$\left|G_q(x;n) - G_{\delta,k;q}(x;n)\right| \le \alpha_{\delta,k;q}(n)\mathrm{E}\left\{|S_n - 1|^k\right\}, \tag{3.11}$$

under the assumption that S_n has k-momnet. The positive constant $\beta_{\delta,k;q}(n)$ is defined by

$$\beta_{\delta,k;q}(n) = \sup_{0<x,0<y\le 1} J_{\delta,k;n}(x,y;n), \tag{3.12}$$

where

$$J_{\delta,k;q}(x,y;n) \equiv \begin{cases} \dfrac{|\Delta_{\delta,k;q}(x,y;n)|}{|y-1|^k}, & \text{for } y \ne 1, \\ \\ \dfrac{1}{k!}|c_{\delta,k;q}(x;n)|\ f_q(x;n), & \text{for } y = 1, \end{cases} \tag{3.13}$$

and

$$\Delta_{\delta,k;q}(x,y;n) \equiv F_q(xy^{-\delta};n) - G_{\delta,k;q}(x,y;n) \tag{3.14}$$

For an approximation to the cdf of T in (1.2), it is suggested to use the sale mixture expressed as in (1.3), and to use (3.10). Then, we need moments of

$$U_{r,m} = S_n - 1 = \frac{\chi_r^2}{\chi_m^2},$$

where $r = p-1$ and $m = n+1$. It is well krnown that for $m-r+1-2j > 0$

$$\mathrm{E}(U_{r,m}^j) = \frac{r(r+2)\cdots(r+2(j-1))}{(m-r-1)(m-r+1)\cdots(m-r+1-2j)}. \tag{3.15}$$

For numerical values of $\alpha_{\delta,k;q}(n)$'s, see Tables 1 and 2.

Table 1: Numerical values of $\alpha_{\delta,k;q}(n)$ $(\delta = 1)$

n	10	20	50	100	200	300
$\alpha_{1,2;2}(n)$	0.2067	0.2180	0.2254	0.2280	0.2293	0.2301
$\alpha_{1,2;4}(n)$	0.3154	0.3487	0.3731	0.3822	0.3870	0.3899
$\alpha_{1,2;6}(n)$	0.3942	0.4540	0.5019	0.5207	0.5308	0.5371
$\alpha_{1,4;2}(n)$	0.1175	0.1245	0.1291	0.1308	0.1316	0.1321
$\alpha_{1,4;4}(n)$	0.1980	0.2224	0.2407	0.2476	0.2512	0.2534
$\alpha_{1,4;6}(n)$	0.2674	0.3170	0.3630	0.4121	0.4401	0.4582

Table 2: Numerical values of $\alpha_{\delta,k;q}(n)$ $(\delta = -1)$

n	10	20	50	100	200	300
$\alpha_{-1,2;2}(n)$	0.2278	0.2468	0.2604	0.2654	0.2680	0.2696
$\alpha_{-1,2;4}(n)$	0.3371	0.3881	0.4291	0.4452	0.4539	0.4593
$\alpha_{-1,2;6}(n)$	0.4102	0.4952	0.5703	0.6018	0.6190	0.6300
$\alpha_{-1,4;2}(n)$	0.1446	0.1648	0.1814	0.1880	0.1916	0.1939
$\alpha_{-1,4;4}(n)$	0.2317	0.2907	0.3471	0.3721	0.3862	0.3953
$\alpha_{-1,4;6}(n)$	0.2950	0.3987	0.5114	0.5657	0.5976	0.6185

Acknowledgements

The author's research is partially supported by the Ministry of Education, Science, Sports, and Culture, a Grant-in-Aid for Scientific Research (C), 16K00047, 2016-2018.

References

[1] Fujikoshi, Y., and Shimizu, R. (2009). Error bounds for asymptotic expansions for scale mixtures of F-distribution and their applications. Proceeding, 2009-Japanese Joint Statistical Meeting.

[2] Fujikoshi, Y., Ulyanov, V. V. and Shimizu, R. (2010). *Multivariate Statistics: High-Dimensional and Large-Sample Approximations*. Wiley, Hobeken, N. J.

[3] Srivastava, M.S. (2002). *Methods of Multivariate Statistics*. Wiley, Hobohen, N. J.