Butler, R.W. and Wood, Andrew T.A. (2014) Laplace approximation of Lauricella functions F_A and F_D. Advances in Computational Mathematics . ISSN 1572-9044 (In Press)

Laplace approximation of Lauricella functions F_A and F_D

Ronald W. Butler
Statistical Science Department
Southern Methodist University
Dallas, TX 75275-0332 U.S.A.

Andrew T. A. Wood
School of Mathematical Sciences
University of Nottingham
Nottingham NG7 2RD U.K.

November 4, 2014

Abstract

The Lauricella functions, which are generalizations of the Gauss hypergeometric function ${}_2F_1$, arise naturally in many areas of mathematics and statistics. So far as we are aware, there is little or nothing in the literature on how to calculate numerical approximations for these functions outside those cases in which a simple one-dimensional integral representation or a one-dimensional series representation is available. In this paper we present first-order and second-order Laplace approximations to the Lauricella functions $F_A^{(n)}$ and $F_D^{(n)}$. Our extensive numerical results show that these approximations achieve surprisingly good accuracy in a wide variety of examples, including cases well outside the asymptotic framework within which the approximations were derived. Moreover, it turns out that the second-order Laplace approximations are usually more accurate than their first-order versions. The numerical results are complemented by theoretical investigations which suggest that the approximations have good relative error properties outside the asymptotic regimes within which they were derived, including in certain cases where the dimension n goes to infinity.

Key Words and Phrases: Gauss hypergeometric function; Lauricella functions; vector-argument hypergeometric functions.

1 Introduction

The Lauricella functions $F_A^{(n)}$, $F_B^{(n)}$, $F_C^{(n)}$ and $F_D^{(n)}$ were introduced in the case $n = 3$ by Lauricella (1893). Each of these functions is a generalization of the classical Gauss hypergeometric function ${}_2F_1$ (e.g. Abramowitz and Stegun, 1972, Chapter 15), and ${}_2F_1$ is recovered when $n = 1$. An extensive account of many of the mathematical properties of Lauricella functions for a general positive integer n, and discussion of problems in mathematics and statistics in which they arise, are given in the book by Exton (1976). These functions appear in a wide variety of settings; see, for instance, Dickey (1983), Lijoi and Regazzini (2004), Kerov and Tsilevich (2004) and Scarpello and Ritelli (2012).

Our starting point in this paper is the question of how to calculate good approximations for Lauricella functions, a problem which, so far as we are aware, has received little

or no attention in the literature. We derive first- and second-order Laplace approximations for the functions $F_A^{(n)}$ and $F_D^{(n)}$, focusing on those situations in which convenient one-dimensional integral representations or one-dimensional series representations are not available. Our numerical results indicate that it is nearly always preferable to use one of the second-order versions. These approximations are shown to be highly accurate for a broad range of argument values, and not just in the asymptotic regimes in which they were derived. In a companion paper, Butler and Wood (2014), we study statistical applications of $F_A^{(n)}$, $F_D^{(n)}$ and $\Phi_2^{(n)}$, a confluent form of $F_D^{(n)}$ which is important in various statistical and other contexts, and we derive Laplace approximations for $\Phi_2^{(n)}$.

Software in Matlab and R along with instructions are available for first- and second-order Laplace approximations for $F_A^{(n)}, F_D^{(n)}$, and $\Phi_2^{(n)}$. They may be downloaded from *http://faculty.smu.edu/rbutler* and are function programs supplied as M- and R-files.

A different type of generalization of the univariate hypergeometric functions are the hypergeometric functions of matrix argument; see Muirhead (1982) for a detailed account and Richards (2012) for more recent developments. The matrix-argument generalizations are quite different to the Lauricella generalizations of the Gauss hypergeometric function. Laplace approximations have been successfully used for various matrix-argument hypergeometric functions by Butler and Wood (2002, 2003).

The outline of this paper is as follows. In §2 we present mathematical definitions and relevant integral representations. In §3 and §4 Laplace approximations are presented for $F_A^{(n)}$ and $F_D^{(n)}$ respectively, and numerical results demonstrating the level of accuracy are presented in §6. Asymptotic properties of the approximations are discussed in §5. Many of the more detailed calculations are presented in the appendices.

2 Review of $F_A^{(n)}$ and $F_D^{(n)}$

2.1 Hypergeometric series representations

Let $(a, m) = \Gamma(a + m)/\Gamma(a)$ denote the Pochhammer symbol with Γ as the classical gamma function; note that when m is a positive integer we shall define $(a, m) = a \ldots (a + m - 1)$ when a is a non-positive integer. The Lauricella function $F_A^{(n)}$ has parameters given by scalar a and n-vectors $\mathbf{b} = (b_1, \ldots, b_n)^T$ and $\mathbf{c} = (c_1, \ldots, c_n)^T$ and an n-vector argument $\mathbf{x} = (x_1, \ldots, x_n)^T$. The defining expansion for $F_A^{(n)}$ is

$$F_A^{(n)}(a, \mathbf{b}; \mathbf{c}; \mathbf{x}) = \sum_{m_1=0}^{\infty} \cdots \sum_{m_n=0}^{\infty} (a, m_1 + \cdots m_n) \prod_{i=1}^{n} \left\{ \frac{(b_i, m_i) x_i^{m_i}}{(c_i, m_i) m_i!} \right\} \tag{1}$$

which is convergent for $\sum_{i=1}^{n} |x_i| < 1$; see Exton (1976, p. 41).

Function $F_D^{(n)}$ has parameters given by scalars a and c, an n-vector $\mathbf{b} = (b_1, \ldots, b_n)^T$, and an n-vector argument $\mathbf{x} = (x_1, \ldots, x_n)^T$. The defining expansion for $F_D^{(n)}$ is

$$F_D^{(n)}(a, \mathbf{b}; c; \mathbf{x}) = \sum_{m_1=0}^{\infty} \cdots \sum_{m_n=0}^{\infty} \frac{(a, m_1 + \cdots m_n)}{(c, m_1 + \cdots m_n)} \prod_{i=1}^{n} \left\{ \frac{(b_i, m_i) x_i^{m_i}}{m_i!} \right\} \tag{2}$$

and is convergent for $\max_{1 \le i \le n} |x_i| < 1$.

In the special case when $n = 1$, $b_1 = b$, $c_1 = c$, and $x_1 = x$,

$$F_A^{(1)}(a, b; c; x) = F_D^{(1)}(a, b; c; x) = {}_2F_1\,(a, b; c; x),$$

where ${}_2F_1$ denotes the Gauss hypergeometric function.

2.2 Integral representations

There are a number of integral representations of $F_A^{(n)}$ and $F_D^{(n)}$, each of which places different conditions on the function parameters; see Exton (1976) for a detailed account. We first consider one-dimensional representations. When $\mathrm{Re}(a) > 0$,

$$F_A^{(n)}(a, \mathbf{b}; \mathbf{c}; \mathbf{x}) = \Gamma(a)^{-1} \int_0^\infty e^{-t} t^{a-1} \prod_{i=1}^n {}_1F_1\left(b_i; c_i; x_i t\right) dt, \tag{3}$$

where ${}_1F_1$ is the confluent hypergeometric function. For arbitrary a, the function $F_A^{(n)}$ also has the following multi-dimensional Euler-type integral representation when $\mathrm{Re}(b_i) > 0$ and $\mathrm{Re}(c_i - b_i) > 0$ for $i = 1, \ldots, n$:

$$F_A^{(n)}(a, \mathbf{b}; \mathbf{c}; \mathbf{x}) = \prod_{i=1}^n \frac{\Gamma(c_i)}{\Gamma(b_i)\Gamma(c_i - b_i)} \int_0^1 \cdots \int_0^1 \left(1 - \mathbf{u}^T\mathbf{x}\right)^{-a} \prod_{i=1}^n \left\{u_i^{b_i - 1}(1 - u_i)^{c_i - b_i - 1}\right\} d\mathbf{u}, \tag{4}$$

where $\mathbf{u} = (u_1, \ldots, u_n)^T$, $\mathbf{u}^T\mathbf{x} = \sum_{i=1}^n u_i x_i$ and $d\mathbf{u} = \prod_{i=1}^n du_i$.

When $\mathrm{Re}(a) > 0$ and $\mathrm{Re}(c - a) > 0$,

$$F_D^{(n)}(a, \mathbf{b}; c; \mathbf{x}) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(c - a)} \int_0^1 u^{a-1}(1 - u)^{c-a-1} \prod_{i=1}^n (1 - u x_i)^{-b_i} du. \tag{5}$$

For arbitrary a, and when $\mathrm{Re}(b_i) > 0$ for $i = 1, \ldots, n$ and $\mathrm{Re}(c - b_+) > 0$, where $b_+ = \sum_{i=1}^n b_i$, the function $F_D^{(n)}$ has representation

$$F_D^{(n)}(a, \mathbf{b}; c; \mathbf{x}) = \frac{\Gamma(c)}{\Gamma(c - b_+) \prod_{i=1}^n \Gamma(b_i)} \int_{u_i \geq 0,\, u_+ \leq 1} \left(1 - u_+\right)^{c - b_+ - 1} \left(1 - \mathbf{u}^T\mathbf{x}\right)^{-a} \prod_{i=1}^n u_i^{b_i - 1} d\mathbf{u}, \tag{6}$$

where $u_+ = \sum_{i=1}^n u_i$.

3 Approximation of $F_A^{(n)}$

3.1 One-dimensional integral representation

If $\mathrm{Re}(a) > 0$ and an accurate method for calculating the confluent hypergeometric function ${}_1F_1$ is available, then approximation of $F_A^{(n)}$ through the one-dimensional integration of (3) will be difficult to beat. To see that such integration is generally well-behaved, note that the integrand in (3) has the dominant factor $e^{-t}t^{a-1}$ which decreases like a gamma density to offset the increasing ${}_1F_1$ factors. As $t \to \infty$, these terms have order ${}_1F_1\left(b_i; c_i; x_i t\right) \sim c_i e^{x_i t} t^{b_i - c_i}$ so that overall the integrand has a gamma-like tail of order $c_0 \exp[-\{1 - \mathrm{Re}(x_+)\}\, t] t^{a - (c_+ - b_+) - 1}$ when $\mathrm{Re}(x_+) < 1$. The subscript $+$ indicates summation over the relevant index, e.g. $x_+ = \sum_{i=1}^n x_i$. For general guidelines on computation of special functions, which is relevant in these cases, see Backeljauw et al. (2014).

3.2 Laplace approximation

If, however, $\mathrm{Re}(a) < 0$, the integral representation in (3) is not valid, so some other method must be used. Here, we will develop a Laplace approximation and, for simplicity,

assume that all the parameters are real. This is not a limitation, however, since the resulting approximation may be used with complex parameters and also may be justified by using analytic continuation arguments; cf. the discussion in Butler and Wood (2002, §6). First write $d = -a$ so that $d > 0$. The integrand in (4) is

$$h(\mathbf{u}) \exp\{-g(\mathbf{u})\} \tag{7}$$

where

$$h(\mathbf{u}) = \prod_{i=1}^{n} u_i^{-1}(1-u_i)^{-1} \tag{8}$$

$$g(\mathbf{u}) = -\sum_{i=1}^{n} \{b_i \log u_i + (c_i - b_i)\log(1-u_i)\} - d\log\left(1 - \mathbf{u}^T\mathbf{x}\right). \tag{9}$$

The first two sets of derivatives are

$$g_i := \frac{\partial g}{\partial u_i} = -\frac{b_i}{u_i} + \frac{c_i - b_i}{1-u_i} + \frac{dx_i}{1-\mathbf{u}^T\mathbf{x}} \tag{10}$$

$$g_{ij} := \frac{\partial^2 g}{\partial u_i \partial u_j} = d_{2i}\delta_{ij} + \frac{dx_i x_j}{\left(1-\mathbf{u}^T\mathbf{x}\right)^2}, \tag{11}$$

for $i, j = 1, \ldots, n$, where δ_{ij} is the indicator that $i = j$ and

$$d_{2i} = \frac{b_i}{u_i^2} + \frac{c_i - b_i}{(1-u_i)^2}.$$

If $\mathbf{D} = \mathrm{diag}(d_{2i})$ is a $n \times n$ diagonal matrix in $\{d_{2i}\}$ and $\mathbf{v} = \mathbf{x}/\left(1 - \mathbf{u}^T\mathbf{x}\right)$, then the $n \times n$ Hessian matrix is $\mathbf{G} = \mathbf{D} + d\mathbf{v}\mathbf{v}^T$. From this form, note that $\mathbf{G}$ is positive definite for all $\mathbf{u} \in [0,1]^n$ when $b_i > 0$ and $c_i - b_i > 0$ for all i, and $d > 0$. Consequently, g has a unique minimum on $[0,1]^n$, making this a promising setting for a Laplace approximation.

To implement Laplace's approximation we need to find this minimum using (10). We first of all solve

$$-\frac{b_i}{u_i} + \frac{c_i - b_i}{1-u_i} + \lambda x_i = 0 \qquad (i = 1, \ldots, n), \tag{12}$$

where λ will be chosen later. After some elementary calculations, and taking care to choose the root in $[0,1]$, we find that

$$\tilde{u}_i(\lambda) = \begin{cases} \left\{c_i + \lambda x_i - \sqrt{(c_i + \lambda x_i)^2 - 4\lambda x_i b_i}\right\}/(2\lambda x_i) & x_i \neq 0 \\ b_i/c_i & x_i = 0 \end{cases} \tag{13}$$

where the solution at $x_i = 0$ follows directly from (12). Note that the possible values of $\lambda = d/\left(1 - \mathbf{u}^T\mathbf{x}\right)$ satisfy

$$\frac{d}{1+x_-} \leq \lambda \leq \frac{d}{1-x_+}$$

where $x_- = \sum_{\{i:x_i<0\}} |x_i|$ and $x_+ = \sum_{\{i:x_i>0\}} x_i$. To solve for λ, we find the unique solution $d/(1+x_-) \leq \hat{\lambda} \leq d/(1-x_+)$ to the equation

$$\lambda = d/\left\{1 - \textstyle\sum_{i=1}^{n} \tilde{u}_i(\lambda) x_i\right\}. \tag{14}$$

To show uniqueness, implicitly differentiate (12) to show that $\operatorname{sgn}(\partial\tilde{u}_i/\partial\lambda) = -\operatorname{sgn}(x_i)$; thus the denominator of (14) is increasing in λ which leads to a unique root. Consequently, the exponent g is minimized at $\hat{\mathbf{u}} = (\hat{u}_1, ...\hat{u}_n)^T$, where $\hat{u}_i = \tilde{u}_i(\hat{\lambda})$.

The notation for the approximation requires evaluating various functions of $\mathbf{u}$ at $\hat{\mathbf{u}}$. For example, denote $\hat{\lambda} = d/(1 - \hat{\mathbf{u}}^T\mathbf{x})$ and $\hat{g}$, $\hat{h}$, and $\hat{\mathbf{D}}$, as g, h, and $\mathbf{D}$, evaluated at $\hat{\mathbf{u}}$. The resulting Laplace approximations are given in the following theorem.

Theorem 1 *Suppose* $\mathrm{Re}(a) < 0$ *and* $\mathrm{Re}(b_i) > 0 < \mathrm{Re}(c_i - b_i)$ *for* $i = 1, \ldots, n$. *The first-order calibrated Laplace approximation to the integral in (4) is*

$$\hat{F}_A^{(n)}(a, \mathbf{b}; \mathbf{c}; \mathbf{x}) = \left\{ \prod_{i=1}^{n} \frac{\hat{\Gamma}(c_i)}{\hat{\Gamma}(b_i)\hat{\Gamma}(c_i - b_i)} \right\} (2\pi)^{n/2} |\hat{\mathbf{G}}|^{-1/2} \hat{h} e^{-\hat{g}}, \tag{15}$$

where $\hat{h} = h(\hat{\mathbf{u}})$ *and* $\hat{g} = g(\hat{\mathbf{u}})$ *are given by the functions in (8) and (9). The determinant*

$$|\hat{\mathbf{G}}| = \{1 + d^{-1}\hat{\lambda}^2 \mathbf{x}^T \hat{\mathbf{D}}^{-1} \mathbf{x}\} |\hat{\mathbf{D}}|$$

and $\hat{\Gamma}(y) = \sqrt{2\pi} y^{y-1/2} e^{-y}$ *is Stirling's approximation for* $\Gamma(y)$. *Two second-order calibrated Laplace approximations are*

$$\hat{F}_{A2}^{(n)} = \hat{F}_A^{(n)}(a, \mathbf{b}; \mathbf{c}; \mathbf{x}) \frac{1 + \hat{O}_{A,\mathbf{x}}}{1 + \hat{O}_{A,\mathbf{0}}} \quad \textit{and} \quad \hat{F}_{A2e}^{(n)} = \hat{F}_A^{(n)}(a, \mathbf{b}; \mathbf{c}; \mathbf{x}) \exp(\hat{O}_{A,\mathbf{x}} - \hat{O}_{A,\mathbf{0}}), \tag{16}$$

where correction terms $\hat{O}_{A,\mathbf{x}}$ *and* $\hat{O}_{A,\mathbf{0}}$ *are given in (38) of the Appendix §A.1.2.*

Note: the expressions in Appendix §A.1.2 for the correction terms $\hat{O}_{A,\mathbf{x}}$ and $\hat{O}_{A,\mathbf{0}}$ are of computational order $O(n^2)$ and have been reduced from their original order of $O(n^6)$.

Proof. The raw Laplace approximation to $F_A^{(n)}$ is given by

$$\tilde{F}_A^{(n)}(a, \mathbf{b}; \mathbf{c}; \mathbf{x}) = \left\{ \prod_{i=1}^{n} \frac{\Gamma(c_i)}{\Gamma(b_i)\Gamma(c_i - b_i)} \right\} (2\pi)^{n/2} |\hat{\mathbf{H}}|^{-1/2} \hat{h} e^{-\hat{g}}.$$

The calibrated version of the approximation is

$$\hat{F}_A^{(n)}(a, \mathbf{b}; \mathbf{c}; \mathbf{x}) = \tilde{F}_A^{(n)}(a, \mathbf{b}; \mathbf{c}; \mathbf{x}) / \tilde{F}_A^{(n)}(a, \mathbf{b}; \mathbf{c}; \mathbf{0})$$

and has been arranged to be exact at $\mathbf{x} = \mathbf{0} = (0, \ldots, 0)^T$. After simplification, its value is given in (15), i.e. we replace $\Gamma(y)$ by the Stirling approximation $\hat{\Gamma}(y)$ so that

$$\hat{F}_A^{(n)}(a, \mathbf{b}; \mathbf{c}; \mathbf{0}) = F_A^{(n)}(a, \mathbf{b}; \mathbf{c}; \mathbf{0}) = 1.$$

Second-order correction terms $\hat{O}_{A,\mathbf{x}}$ and $\hat{O}_{A,\mathbf{0}}$ are derived in §A.1. ■

4 Approximation of $F_D^{(n)}$

If $\mathrm{Re}(a) > 0$ and $\mathrm{Re}(c-a) > 0$, so that integral representation (5) is valid, then a simple and direct approach to computation is to use one-dimensional numerical integration to evaluate (5). If $\max_i |x_i| < 1$, then numerical integration is well-behaved with the

integrand smooth and bounded over $u \in [0,1]$. A second possibility when $\mathrm{Re}(a) > 0$ is to use the one-dimensional series expansion given by van Laarhoven and Kalker (1988). See Backeljauw et al. (2014) for general computational guidance of relevance in the above case. Here, we focus on the more challenging case in which $\mathrm{Re}(a) < 0$.

If $\mathrm{Re}(a) < 0$, $\mathrm{Re}\,(c - b_+) > 0$ and also $\mathrm{Re}(b_i) > 0$ for $i = 1, \ldots, n$, then for (6) we may use Laplace's approximation with h and g in (7) replaced by

$$h(\mathbf{u}) = (1 - u_+)^{-1} \prod_{i=1}^{n} u_i^{-1} \tag{17}$$

$$g(\mathbf{u}) = -\textstyle\sum_{i=1}^{n} b_i \log u_i - (c - b_+) \log(1 - u_+) - d \log\left(1 - \mathbf{u}^T\mathbf{x}\right) \tag{18}$$

where $d = -a$ so that $d > 0$. Straightforward calculation yields

$$g_i := \frac{\partial g}{\partial u_i} = -\frac{b_i}{u_i} + \frac{c - b_+}{1 - u_+} + \frac{dx_i}{1 - \mathbf{u}^T\mathbf{x}}, \tag{19}$$

$$g_{ij} := \frac{\partial^2 g}{\partial u_i \partial u_j} = \frac{b_i}{u_i^2}\delta_{ij} + \frac{c - b_+}{(1 - u_+)^2} + \frac{dx_i x_j}{(1 - \mathbf{u}^T\mathbf{x})^2}, \tag{20}$$

for $i, j = 1, \ldots, n$. If $\mathbf{D} = \mathrm{diag}(b_i/u_i^2)$ is an $n \times n$ diagonal matrix, $\mathbf{v}_1 = \mathbf{1}/\left(1 - \mathbf{u}^T\mathbf{1}\right)$ and $\mathbf{v}_2 = \mathbf{x}/\left(1 - \mathbf{u}^T\mathbf{x}\right)$, where $\mathbf{1}$ is a vector of ones, then the $n \times n$ Hessian matrix $\mathbf{G}(\mathbf{u}) = \mathbf{D} + (c - b_+)\mathbf{v}_1\mathbf{v}_1^T + d\mathbf{v}_2\mathbf{v}_2^T$ so $\mathbf{G}$ is positive definite for all $\mathbf{u} \in [0,1]^n$.

To minimize g we solve

$$-\frac{b_i}{u_i} + \mu + \lambda x_i = 0$$

giving $\tilde{u}_i(\mu, \lambda) = b_i/(\mu + \lambda x_i)$ for given μ and λ. The minimum of g at $(\hat{\mu}, \hat{\lambda})$ must be unique because the Hessian $\mathbf{G}$ is positive definite and so g is convex. The values $\hat{\mu}$ and $\hat{\lambda}$ satisfy

$$\mu = \frac{c - b_+}{1 - \sum_{i=1}^{n} \tilde{u}_i(\mu, \lambda)} \quad \text{and} \quad \lambda = \frac{d}{1 - \sum_{i=1}^{n} \tilde{u}_i(\mu, \lambda) x_i}, \tag{21}$$

and are bounded as follows:

$$\frac{\mu d}{\mu(1 - x_{\min}) + x_{\min}(c - b_+)} \le \lambda \le \frac{\mu d}{\mu(1 - x_{\max}) + x_{\max}(c - b_+)} \tag{22}$$

$$c - b_+ < \mu < \infty.$$

It turns out, however, that $\hat{\mu}$ and $\hat{\lambda}$ can be more easily calculated by first determining $\hat{\rho} = \hat{\lambda}/\hat{\mu}$ as the root of a single transcendental equation in $\rho = \lambda/\mu$. To find this equation, rewrite the two equations in (21) as

$$\mu = \textstyle\sum_{i=1}^{n} b_i/(1 + \rho x_i) + c - b_+ \quad \text{and} \quad \lambda = \rho \sum_{i=1}^{n} b_i x_i/(1 + \rho x_i) + d. \tag{23}$$

Now set ρ equal to the ratio of the right summation in (23) divided by the left summation. After simplification this leads to the single equation

$$d = \rho\left(\sum_{i=1}^{n} \frac{b_i(1 - x_i)}{1 + \rho x_i} + c - b_+\right) \qquad \begin{array}{rl} 0 < \rho < \infty & \text{if} \quad \min_i x_i \ge 0 \\ 0 < \rho < -1/\min x_i & \text{if} \quad \min x_i < 0 \end{array} \tag{24}$$

with a unique root $\hat{\rho}$ as explained below. The values for $\hat{\mu}$ and $\hat{\lambda}$ are given by replacing ρ in (23) by $\hat{\rho}$. The constraints on ρ in the root-finding in (24) are those consistent

with $\{\tilde{u}_i(\mu,\lambda) : i = 1,\ldots,n\}$ lying in the interior of the simplex in $\Re^n$. It can be shown by differentiation with respect to ρ that the RHS of the equation in (24) is monotonic increasing in each of the regions for ρ given in (24). Moreover, the RHS of the equation in (24) has range $(0,\infty)$ in both cases of relevance. Consequently, the solution for ρ is unique when ρ lies in the relevant interval given in (24).

In what follows we use a 'hat' to indicate evaluation at $\hat{\mu}$ and $\hat{\lambda}$, e.g. $\hat{u}_i = \tilde{u}_i(\hat{\mu},\hat{\lambda})$ for $i = 1,\ldots,n$. The Hessian $\mathbf{G}$ evaluated at $\hat{\mathbf{u}}$ may be expressed in terms of $(\hat{\mu},\hat{\lambda})$ as

$$\hat{\mathbf{G}} = \mathbf{G}(\hat{\mathbf{u}}) = \hat{\mathbf{D}} + (c-b_+)^{-1}\hat{\mu}^2\mathbf{1}\mathbf{1}^T + d^{-1}\hat{\lambda}^2\mathbf{x}\mathbf{x}^T$$

where $\hat{\mathbf{D}} = \mathrm{diag}\{b_1/\hat{u}_1^2,\ldots,b_n/\hat{u}_n^2\}$. In §A.1.4, it is shown that

$$|\hat{\mathbf{G}}| = |\hat{\mathbf{D}}|\,\hat{\Xi} \tag{25}$$

where

$$\hat{\Xi} = \left\{1 + (c-b_+)^{-1}\hat{\mu}^2\sum_{i=1}^n \frac{\hat{u}_i^2}{b_i}\right\}\left\{1 + d^{-1}\hat{\lambda}^2\sum_{i=1}^n \frac{\hat{u}_i^2 x_i^2}{b_i}\right\} - (c-b_+)^{-1}d^{-1}\hat{\mu}^2\hat{\lambda}^2\left(\sum_{i=1}^n \frac{\hat{u}_i^2 x_i}{b_i}\right)^2.$$

The resulting Laplace approximations are given in the following theorem.

Theorem 2 *Suppose* $\mathrm{Re}(a) < 0 < \mathrm{Re}(c-b_+)$ *and* $\mathrm{Re}(b_i) > 0$ *for* $i = 1,\ldots,n$. *The first-order calibrated Laplace approximation to* $F_D^{(n)}$ *given in (6) is*

$$\hat{F}_D^{(n)}(a,\mathbf{b};c;\mathbf{x}) = \frac{\hat{\Gamma}(c)}{\hat{\Gamma}(c-b_+)\prod_{i=1}^n \hat{\Gamma}(b_i)}(2\pi)^{n/2}|\hat{\mathbf{G}}|^{-1/2}\hat{h}e^{-\hat{g}}, \tag{26}$$

where $\hat{h} = h(\hat{\mathbf{u}})$ *and* $\hat{g} = g(\hat{\mathbf{u}})$ *are given by the functions in (17) and (18). The determinant* $|\hat{\mathbf{G}}|$ *is given in (25) and* $\hat{\Gamma}(y) = \sqrt{2\pi}y^{y-1/2}e^{-y}$ *is Stirling's approximation for* $\Gamma(y)$. *Two second-order calibrated Laplace approximations are*

$$\hat{F}_{D2}^{(n)} = \hat{F}_D^{(n)}(a,\mathbf{b};\mathbf{c};\mathbf{x})\frac{1+\hat{O}_{D,\mathbf{x}}}{1+\hat{O}_{D,\mathbf{0}}} \qquad \text{and} \qquad \hat{F}_{D2e}^{(n)} = \hat{F}_A^{(n)}(a,\mathbf{b};\mathbf{c};\mathbf{x})\exp(\hat{O}_{D,\mathbf{x}} - \hat{O}_{D,\mathbf{0}}), \tag{27}$$

where correction terms $\hat{O}_{D,\mathbf{x}}$ *and* $\hat{O}_{D,\mathbf{0}}$ *are given in (40) of §A.1.3.*

Note: the expressions in Appendix §A.1.3 for the correction terms $\hat{O}_{D,\mathbf{x}}$ and $\hat{O}_{D,\mathbf{0}}$ are of computational order $O(n^2)$ and have been reduced from their original order of $O(n^6)$.

5 Relative errors of the approximations

In this section we discuss the theoretical accuracy of the Laplace approximations presented above. The discussion is incomplete because so many different limiting cases arise and it is difficult to summarise the full diversity of asymptotic regimes concisely. The relevant measure for assessing theoretical accuracy is the relative error. The relative error of an approximation $\hat{F}$ of F is defined by $(\hat{F}-F)/F$. We consider two types of result, those for fixed n and those for $n \to \infty$.

5.1 Fixed n results

Let us first clarify the asymptotic regimes for which the Laplace approximations were designed.

Case I: $F_A^{(n)}(a, \mathbf{b}; \mathbf{c}; \mathbf{x})$. Let the arguments of $F_A^{(n)}$ be fixed and satisfy

$$a < 0, \quad b_i > 0, \quad c_i - b_i > 0, \quad i = 1, \ldots, n, \quad \text{and} \quad \sum_{i=1}^{n} |x_i| < 1, \tag{28}$$

so that the integral representation (3) is valid, and let ν denote a large positive quantity. The approximation in Theorem 1 was designed to evaluate $F_A^{(n)}(\nu a, \nu \mathbf{b}; \nu \mathbf{c}; \mathbf{x})$ as $\nu \to \infty$.

Case II: $F_D^{(n)}(a, \mathbf{b}; c; \mathbf{x})$. Here, the arguments are fixed and satisfy

$$a < 0, \quad b_i > 0, \quad i = 1, \ldots, n, \quad c - b_+ > 0, \quad \text{and} \quad \max_{i=1,\ldots,n} |x_i| < 1, \tag{29}$$

so that integral representation (6) holds. The approximation in Theorem 2 was designed to evaluate $F_D^{(n)}(\nu a, \nu \mathbf{b}; \nu c; \mathbf{x})$ as $\nu \to \infty$.

For the asymptotic regimes considered in these cases, the main results, which can be proved using standard theorems on Laplace approximation (e.g. Hsu, 1951), are as follows.

Theorem 3 *As $\nu \to \infty$, the relative errors in the first-order approximations (15) and (26) under Cases I and II, respectively, are both $O(\nu^{-1})$; and the relative errors of the corresponding second-order approximation in (16) and (27) are both $O(\nu^{-2})$.*

It is important to recognise, however, that the usefulness of the Laplace approximations in §3 and §4 goes far beyond the asymptotic regimes considered in Cases I and II. With some further work, which we do not reproduce here, it can be shown that in many other asymptotic regimes, including the following, the relative error of the Laplace approximations remains bounded. Below, α, β and γ are binary variables with possible values 0 or 1 which satisfy $\beta \leq \gamma$ and $\alpha + \beta + \gamma \geq 1$.

(i) If a, $\mathbf{b}$, $\mathbf{c}$ and $\mathbf{x}$ satisfy (28), and $\gamma \geq \beta$, then

$$\lim_{\nu \to \infty} \frac{\hat{F}_A^{(n)}(\nu^\alpha a, \nu^\beta \mathbf{b}; \nu^\gamma \mathbf{c}; \mathbf{x})}{F_A^{(n)}(\nu^\alpha a, \nu^\beta \mathbf{b}; \nu^\gamma \mathbf{c}; \mathbf{x})} \in (0, \infty); \tag{30}$$

(ii) If a, $\mathbf{b}$, c and $\mathbf{x}$ satisfy (29), then

$$\lim_{\nu \to \infty} \frac{\hat{F}_D^{(n)}(\nu^\alpha a, \nu^\beta \mathbf{b}; \nu^\gamma \mathbf{c}; \mathbf{x})}{F_D^{(n)}(\nu^\alpha a, \nu^\beta \mathbf{b}; \nu^\gamma \mathbf{c}; \mathbf{x})} \in (0, \infty). \tag{31}$$

Our final remarks in this subsection are more speculative. First, although we have not explored this in the present paper, we believe there are grounds, based on analytic continuation arguments, for hoping that the Laplace approximations considered in this paper will do a good job of tracking the function it is approximating outside the domain of the integral representation and for complex values of the parameters and arguments; cf. Butler and Wood (2002, §6).

If this is indeed the case, then on the basis of (30) and (31), and the comments immediately above, there seem grounds for speculating that the following holds. We just formulate the conjecture for $F_D^{(n)}$; the corresponding conjecture for $F_A^{(n)}$ is formulated in similar fashion. Let Ω denote a subset of $\mathbf{C}^{2n+2}$ where $\mathbf{C}$ is the set of complex numbers and $2n+2$ is the (complex) dimension of $(a, \mathbf{b}, c, \mathbf{x})$ where all these quantities are allowed to be complex. Suppose that on Ω the Lauricella function $F_D^{(n)}(a, \mathbf{b}; c; \mathbf{x})$ is jointly analytic in all the variables. Then we conjecture that for any such Ω which is a closed set,

$$\sup_{(a,\mathbf{b},c,\mathbf{x}) \in \Omega} \left| \log \left\{ \frac{\hat{F}_D^{(n)}(a, \mathbf{b}; c; \mathbf{x})}{F_D^{(n)}(a, \mathbf{b}; c; \mathbf{x})} \right\} \right| < \infty,$$

where $\hat{F}_D^{(n)}$ is either the first-order or one of the second-order Laplace approximations presented in §4.

5.2 Results when $n \to \infty$

Here we show that there are interesting asymptotic regimes for $F_A^{(n)}$ and $F_D^{(n)}$ in which $n \to \infty$ and yet the relative error of the Laplace approximation goes to zero. This provides theoretical support for the excellent results seen in Tables 2 and 4 when n is relatively large. The key requirements for the results in this subsection to hold are negligibility conditions similar to those needed for central limit theorems for independent triangular arrays. Proofs of Theorems 4 and 5 are outlined in the appendix.

For $F_A^{(n)}$, a standard sufficient condition for the hypergeometric series expansion to be absolutely convergent is that $\sum_{i=1}^n |x_i^{(n)}| < \infty$. Theorem 4 is proved under weaker conditions, but if we were to assume this condition holds then assumptions (ii) and (iii) in Theorem 4 would not be required. A analogous comment applies to Theorem 5: a standard sufficient condition for the hypergeometric series expansion for $F_D^{(n)}$ to be absolutely convergent is that $\max_{i=1,\ldots,n} |x_i^{(n)}| < 1$. This condition is not assumed in Theorem 5 below, but if it were to hold then assumptions (ii), (iii) and (iv) of Theorem 5 would be redundant.

Theorem 4 *Consider* $\mathbf{b}^{(n)} = (b_1^{(n)}, \ldots, b_n^{(n)})^T$ *and* $\mathbf{c}^{(n)} = (c_1^{(n)}, \ldots, c_n^{(n)})^T$ *such that, for each* $n \geq 1$ *and* $i = 1, \ldots, n$, $0 < b_i^{(n)} < c_i^{(n)}$. *Define* $\mathbf{x}^{(n)} = (x_1^{(n)}, \ldots, x_n^{(n)})^T$ *and* $\boldsymbol{\pi}^{(n)} = (\pi_1^{(n)}, \ldots, \pi_n^{(n)})^T$, *where* $\pi_i^{(n)} = b_i^{(n)}/c_i^{(n)}$, *and suppose that the following conditions are satisfied as* $n \to \infty$*:*

(i) $\max_{1 \leq i \leq n} |x_i^{(n)}|/c_i^{(n)} \to 0$*;*

(ii) for all $n \geq 1$, $\sum_{i=1}^n \max(x_i^{(n)}, 0) \leq 1$*;*

(iii) for some $A < \infty$ *independent of* n, $\sum_{i=1}^n \max(-x_i^{(n)}, 0) \leq A$*;*

(iv) for some $\epsilon > 0$ *independent of* n*,*

$$\limsup_{n \to \infty} \max_{1 \leq i \leq n} \pi_i^{(n)} \leq 1 - \epsilon \quad \text{and} \quad \liminf_{n \to \infty} \pi_i^{(n)} \geq \epsilon;$$

(v) defining $\xi_n = \sum_{i=1}^n \pi_i^{(n)} x_i^{(n)}$, *it is assumed that* $\xi_n \to \xi_0 \in (-\infty, 1)$ *as* $n \to \infty$.

Then, for fixed $a < 0$,

$$\lim_{n\to\infty} \frac{\hat{F}_A^{(n)}(a, \mathbf{b}^{(n)}; \mathbf{c}^{(n)}; \mathbf{x}^{(n)})}{F_A^{(n)}(a, \mathbf{b}^{(n)}; \mathbf{c}^{(n)}; \mathbf{x}^{(n)})} = 1. \tag{32}$$

Moreover, the second-order terms satisfy $\hat{O}_{A,\mathbf{x}} - \hat{O}_{A,\mathbf{0}} \to 0$, *from which we conclude that both second-order approximations have limiting relative error* 0.

Assumption (iv) above is stronger than is needed but is included to avoid uninteresting complications in the proof. A similar type of result with appropriate modifications holds for the $F_D^{(n)}$ approximations.

Theorem 5 *Consider* $\mathbf{b}^{(n)} = (b_1^{(n)}, \ldots, b_n^{(n)})^T$ *and* $c^{(n)}$ *such that, for each* $n \geq 1$ *and* $i = 1, \ldots, n$, $b_i^{(n)} > 0$ *and* $b_+^{(n)} = \sum_{i=1}^n b_i^{(n)} < c^{(n)}$. *Define* $\mathbf{x}^{(n)} = (x_1^{(n)}, \ldots, x_n^{(n)})^T$ *and* $\boldsymbol{\pi}^{(n)} = (\pi_1^{(n)}, \ldots, \pi_n^{(n)})^T$, *where* $\pi_i^{(n)} = b_i^{(n)}/c^{(n)}$, *and suppose that the following conditions are satisfied as* $n \to \infty$:

(i) $b_+^{(n)} \to \infty$ *and, for some* $\epsilon > 0$ *independent of* n, $c^{(n)} - b_+^{(n)} \geq \epsilon$;

(ii) $\max_{1\leq i\leq n} |x_i^{(n)}| = o(c^{(n)})$;

(iii) $\limsup_{n\to\infty} \max_{1\leq i\leq n} x_i^{(n)} < 1$;

(iv) for some $A < \infty$ *independent of* n, $\min_{1\leq i\leq n} x_i^{(n)} \geq -A$;

(v) defining $\xi_n = \sum_{i=1}^n \pi_i^{(n)} x_i^{(n)}$, *it is assumed that* $\xi_n \to \xi_0 \in (-\infty, 1)$ *as* $n \to \infty$.

Then, for fixed $a < 0$,

$$\lim_{n\to\infty} \frac{\hat{F}_D^{(n)}(a, \mathbf{b}^{(n)}; c^{(n)}; \mathbf{x}^{(n)})}{F_D^{(n)}(a, \mathbf{b}^{(n)}; c^{(n)}; \mathbf{x}^{(n)})} = 1.$$

Moreover, the second-order terms satisfy $\hat{O}_{D,\mathbf{x}} - \hat{O}_{D,\mathbf{0}} \to 0$, *from which we conclude that both second-order approximations have limiting relative error* 0.

6 Numerical accuracy

6.1 Accuracy of approximations of $F_A^{(n)}$

For a limited number of examples, it is possible to compute $F_A^{(n)}(a, \mathbf{b}; \mathbf{c}; \mathbf{x})$ using exact arithmetic for the integration in Maple. Table 1 shows such computations. Also shown are the various first- and second-order Laplace approximations with their percentage relative errors computed using Maple carrying 100 digits in the computations. Apart from the third example, which has $n = 2$ and fractional entries for $\mathbf{b}$ and $\mathbf{c}$, the two second-order approximations achieve smaller relative errors. Large values for d and the entries of $\mathbf{b}$ and $\mathbf{c}$ replicate the asymptotics for the Laplace expansion where accuracy is expected so these values have been purposefully kept small in Table 1 to show more challenging examples for the approximations. Both second-order approximations achieve remarkable accuracy for these examples.

Table 2 examines accuracy when exact computation of $F_A^{(n)}$ is not possible. Monte Carlo simulation of these values was performed using programs written in R and working with double precision arithmetic. The integral representation in (4) characterizes $F_A^{(n)}$ as the expected value of $(1-\sum_{i=1}^n U_i x_i)^{-a}$ when $\{U_i\}$ are independent Beta $(b_i, c_i - b_i)$ random variables. The cell entries labelled $\breve{F}_A^{(n)} \pm 1.96SE$ provide 95% confidence intervals for $F_A^{(n)}$ based on averaging 10^7 values of the random quantity $(1-\sum_{i=1}^n U_i x_i)^{-a}$.

All approximations in Table 2 were computed using both the Matlab and R routines available at *faculty.smu.edu/rbutler/*. The Matlab and R computations agreed to $8-10$ digits in all instances. We suspect that this limitation in agreement is due more to inherent inaccuracy in the underlying programing for the R language.

Table 1. Exact value $F_A^{(n)}$ is compared to first-order $\hat{F}_A^{(n)}$ and second-order $\hat{F}_{A2}^{(n)}$ and $\hat{F}_{A2e}^{(n)}$ Laplace approximations for the listed values of $n, d, \mathbf{b}, \mathbf{c}$, and $\mathbf{x}$. Percentage relative errors $100(\hat{F}_{A2}^{(n)}/F_{A2}^{(n)} - 1)$ are given in the last three columns and the most accurate approximation in each row is emboldened. The notation used for $\mathbf{b}, \mathbf{c}$, and $\mathbf{x}$ is as follows: $5(1) = 1, 1, 1, 1, 1$; $1(1)5 = 1, 2, 3, 4, 5$; $5(1)$, $2(1)6 = 1, 1, 1, 1, 1, 2, 3, 4, 5, 6$, etc.

n	d	$\mathbf{b}$	$\mathbf{c}$	$\mathbf{x}$		% Relative Error		
		$F_A^{(n)}$	$\hat{F}_A^{(n)}$	$\hat{F}_{A2}^{(n)}$	$\hat{F}_{A2e}^{(n)}$	$\hat{F}_A^{(n)}$	$\hat{F}_{A2}^{(n)}$	$\hat{F}_{A2e}^{(n)}$
2	2	$2, 3$	$3, 4$	0.1, 0.2				
		0.6157	0.6146	**0.6158**	0.6160	−0.178	**0.0160**	0.0480
2	1	$1, 1$	$2, 2$	0.45, 0.55				
		0.5000	0.4952	0.4973	**0.4978**	−0.960	−0.540	**−0.440**
2	1	$1/2, 1/2$	$1, 1$	0.7, 0.25				
		0.5250	**0.5193**	0.4973	0.4978	**−1.08**	−5.28	−5.18
2	3	$4, 6$	$6, 8$	0.85, 0.15				
		0.05718	0.05802	0**.05718**	0.05711	1.463	**−0.0018**	−0.126
2	1	$1, 1$	$2, 2$	$-1, -3$				
		3.000	2.966	2.982	**2.987**	−1.13	−0.600	**−0.433**
5	1	5(1)	5(2)	2(0.1), 0.2, 2(0.3)				
		0.5000	0.4927	0.4977	**0.5010**	−1.46	−0.46	**0.200**
5	2	1(1)5	2(2)10	2(0.1), 0.2, 2(0.3)				
		0.2573	0.2561	**0.2572**	0.2576	−0.466	**−0.0389**	0.117
5	2	5(1/2)	5(1)	2(0.1), 0.2, 2(0.3)				
		0.2800	0.2606	0.2654	**0.2716**	−6.93	−5.21	**−3.00**
5	2	1(1)5	2(2)10	$2(-2), 2(-4), -6$				
		102.37	101.88	**102.35**	102.48	−0.479	**−0.020**	0.110

Table 2. Simulated 95% confidence interval $\breve{F}_A^{(n)} \pm 1.96SE$ for $F_A^{(n)}$ is compared to first-order $\hat{F}_A^{(n)}$ and second-order $\hat{F}_{A2}^{(n)}$ and $\hat{F}_{A2e}^{(n)}$ Laplace approximations for the listed values of $n, d, \mathbf{b}, \mathbf{c}$, and $\mathbf{x}$. Percentage relative errors $100(\hat{F}_{A2}^{(n)}/\breve{F}_{A2}^{(n)} - 1)$ are given in the last three columns and the most accurate approximation in each row is emboldened. See Table 1 for an explanation of the notation for $\mathbf{b}, \mathbf{c}$, and $\mathbf{x}$.

n	d	$\mathbf{b}$	$\mathbf{c}$	$\mathbf{x}$	
	$\breve{F}_A^{(n)} \pm 1.96SE$		$\hat{F}_A^{(n)}$	$\hat{F}_{A2}^{(n)}$	$\hat{F}_{A2e}^{(n)}$
			% Rel. Err.	% Rel. Err.	% Rel. Err.
10	5	1(1)10	2(3/2)31/2	2(1/20), 6(1/10), 2(3/20)	
	$0.009199 \pm 3.88 \times 10^{-6}$		0.009099	**0.009182**	0.009219
			−1.09	**−0.186**	0.220
10	10	1(1)10	6(3/2)41/2	2(1/20), 6(1/10), 2(3/20)	
	$0.004854 \pm 2.28 \times 10^{-6}$		0.004842	0.004852	**0.004855**
			−0.234	−0.03621	**0.0212**
20	5	1(1/2)21/2	2(2/3)44/3	4(1/40), 12(1/20), 4(3/40)	
	$0.003656 \pm 1.27 \times 10^{-6}$		0.003626	0.003645	**0.003663**
			−0.824	−0.295	**0.203**
40	10	10(1), 20(3), 10(5)	10(2), 20(6), 10(10)	8(1/80), 24(1/40), 8(3/80)	
	$0.001151 \pm 4.5 \times 10^{-7}$		0.001133	0.001140	**0.001156**
			−1.49	−0.897	**0.496**
60	10	20(1), 20(4), 20(6)	20(2), 20(7), 20(10)	15(1/90), 30(1/60), 15(1/45)	
	$0.0002780 \pm 1.0 \times 10^{-7}$		0.0002722	0.0002739	**0.0002803**
			−2.09	−1.48	**0.830**
120	10	40(1), 40(4), 40(6)	40(2), 40(7), 40(10)	30(.005), 60(.009), 30(.01)	
	$0.0002941 \pm 7.4 \times 10^{-8}$		0.0002909	0.0002915	**0.0002954**
			−1.06	−0.883	**0.468**

6.2 Accuracy of approximations of $F_D^{(n)}$

Table 3 shows computations for values of $F_D^{(n)}(a, \mathbf{b}; c; \mathbf{x})$, the various first- and second-order Laplace approximations, and percentage relative error when it is possible to do exact arithmetic for the integration in Maple. Our comments on Table 3 are similar to those given in §6.1 for Table 1.

Table 3. Exact value $F_D^{(n)}$ is compared to first-order $\hat{F}_D^{(n)}$ and second-order $\hat{F}_{D2}^{(n)}$ and $\hat{F}_{D2e}^{(n)}$ Laplace approximations for the listed values of $n, d, b, \mathbf{c}$, and $\mathbf{x}$. Percentage relative errors $100(\hat{F}_{D2}^{(n)}/F_{D2}^{(n)} - 1)$ are given in the last three columns and the most accurate approximation in each row is emboldened. See Table 1 for an explanation of the notation for $\mathbf{b}, \mathbf{c}$, and $\mathbf{x}$. Also, $2\{0.1(.1)0.5\} = 0.1(.1)0.5,\ 0.1(.1)0.5$.

n	d	$\mathbf{b}$	c	$\mathbf{x}$		% Relative Error		
		$F_D^{(n)}$	$\hat{F}_D^{(n)}$	$\hat{F}_{D2}^{(n)}$	$\hat{F}_{D2e}^{(n)}$	$\hat{F}_D^{(n)}$	$\hat{F}_{D2}^{(n)}$	$\hat{F}_{D2e}^{(n)}$
2	2	$1, 1$	3	0.45, 0.55				
		$2/3$	0.66178	**0.66671**	0.66783	−0.733	**0.00665**	0.174
2	1	$1/2, 1/2$	2	0.7, 0.25				
		0.76250	0.75913	0.76099	**0.76170**	−0.442	−0.198	**−0.105**
2	3	$4, 6$	12	0.85, 0.15				
		0.28221	**0.28221**	0.28220	0.28220	**0.00106**	−0.00474	−0.00515
2	3	$4, 6$	12	0.8, 0.9				
		0.030231	0.02997	0.030102	**0.030112**	−0.878	−0.425	**−0.392**
2	1	$1, 1$	3	$-3, -4$				
		$10/3$	3.3259	3.3297	**3.3305**	−0.222	−0.109	**−0.0840**
2	3	$1, 1$	4	$-1, -2$				
		$61/10$	5.99828	6.07679	**6.09224**	−1.67	−0.380	**−0.127**
5	2	5(1)	6	2(0.1), 0.2, 2(0.3)				
		0.69619	0.69581	0.69611	**0.69625**	−0.0546	−0.0116	**−0.00855**
5	2	5(1)	6	2(0.8), 3(0.9)				
		0.095238	0.093526	0.093593	**0.093626**	−1.80	−1.73	**−1.69**
5	4	1(1)5	16	2(0.1), 0.2, 2(0.3)				
		0.36276	0.36269	0.36275	**0.36276**	−0.0179	−0.00254	**0.00155**
5	4	1(1)5	16	0.5(0.1)0.9				
		.0078256	.0076330	**.0078127**	.0078643	−2.46	**−0.165**	0.494
10	2	10(1)	11	2(.05), 6(.1), 2(.15)				
		0.82659	0.82657	0.82658	**0.82659**	−0.00197	−0.000836	**0.000197**
10	2	10(1)	11	$2\{0.1(.1)0.5\}$				
		0.53106	0.53085	**0.53097**	0.53086	−0.0405	**−0.0174**	−0.0384

Table 4. Simulated 95% confidence interval $\breve{F}_D^{(n)} \pm 1.96SE$ for $F_D^{(n)}$ is compared to first-order $\hat{F}_D^{(n)}$ and second-order $\hat{F}_{D2}^{(n)}$ and $\hat{F}_{D2e}^{(n)}$ Laplace approximations for the listed values of $n, d, \mathbf{b}$,c, and $\mathbf{x}$. Percentage relative errors $100(\hat{F}_{D2}^{(n)}/\breve{F}_{D2}^{(n)} - 1)$ are given in the last three columns and the most accurate approximation in each row is emboldened. See Table 1 for an explanation of the notation for $\mathbf{b}, \mathbf{c}$, and $\mathbf{x}$.

n	d	**b**	c	**x**		
	$\breve{F}_D^{(n)} \pm 1.96SE$			$\hat{F}_D^{(n)}$	$\hat{F}_{D2}^{(n)}$	$\hat{F}_{D2e}^{(n)}$
				% Rel. Err.	% Rel. Err.	% Rel. Err.
10	5	1(1)10	60	$1/10, 1/10(1/10)9/10$		
$0.019333 \pm 5.0 \times 10^{-6}$				0.019328	0.019332	**0.019333**
				−0.0269	−0.00642	**−0.00110**
10	5	$3(1), 4(5), 3(9)$	65	$5(1/2), 5(3/4)$		
$0.0252249 \pm 6.4 \times 10^{-6}$				0.025214	0.025220	**0.025222**
				−0.0433	−0.0205	**−0.0125**
20	5	1(1/2)21/2	130	$5(1/4), 1(1/2), 5(3/4)$		
$0.0274947 \pm 3.7 \times 10^{-6}$				0.0274904	0.0274913	**0.0274917**
				−0.0155	−0.0122	**−0.0108**
40	10	$10(1), 20(3), 10(5)$	130	$10(1/4), 20(1/2), 10(3/4)$		
$0.00047302 \pm 1.3 \times 10^{-7}$				0.00047297	0.00047299	**0.00047302**
				−0.0108	−0.00677	**−0.000404**
60	10	$20(1), 20(4), 20(6)$	230	$15(1/4), 30(1/2), 15(3/4)$		
$0.000283352 \pm 5.1 \times 10^{-8}$				0.000283346	0.000283348	**0.000283353**
				−0.00229	−0.00151	**0.000332**
120	10	$40(1), 40(4), 40(6)$	450	$30(1/4), 60(1/2), 30(3/4)$		
$0.000208523 \pm 2.5 \times 10^{-8}$				0.000208515	0.000208515	**0.000208516**
				−0.00380	−0.00371	**−0.00329**

Table 4 examines accuracy when exact computation of $F_D^{(n)}$ is not possible. All computations were performed using both Matlab and R routines as described in Table 2. The integral representation in (6) characterizes $F_D^{(n)}$ as the expected value of $\left(1 - \sum_{i=1}^n U_i x_i\right)^d$ when $(U_1, \ldots U_n)$ is a Dirichlet $(\mathbf{b}, c - b_+)$ random vector. Cell entries labelled $\breve{F}_D^{(n)} \pm 1.96SE$ provide 95% confidence intervals for $F_D^{(n)}$ based on averaging 10^7 values of the random quantity $\left(1 - \sum_{i=1}^n U_i x_i\right)^d$. In these higher dimensions, second-order approximation $\hat{F}_{D2e}^{(n)}$ consistently demonstrates greater accuracy than $\hat{F}_{D2}^{(n)}$ if the true value of $F_D^{(n)}$ is taken to be the center of the confidence interval. However, the inherent randomness connected with the simulation limits the accuracy for the confidence interval centers so as to prevent any firm conclusions. In simulations with $n \leq 40$, confidence intervals capture $3 - 4$ significant digits of the true values which is also the degree to which all three approximations agree. For $n = 60$ and 120, the simulations capture 4 significant digits but the three approximations agree to 5 significant digits. The agreement between simulated estimates and first-order approximations is reflected in the p-values of one-sample two-tailed z-tests for the hypotheses that the first-order approximation $\hat{F}_D^{(n)}$ is the exact value of $F_D^{(n)}$, i.e. $H_0 : F_D^{(n)} = \hat{F}_D^{(n)}$. For the six examples in Table 4, these p-values in percentages are 32, 8.9, 25, 70, 91, and 75% respectively from $n = 10$ to 120. Such increasing accuracy in n demonstrates the asymptotic regime as $n \to \infty$ of Theorem 5 in which all three approximations become asymptotically correct.

A Derivations and proofs

A.1 Derivation of Laplace approximations

A.1.1 Second-order Laplace approximation

A second-order Laplace approximation is shown for the integral

$$\int_{\mathbf{u}\in\Re^p} h(\mathbf{u})\exp\{-\nu g(\mathbf{u})\}d\mathbf{u} = (2\pi)^{p/2}|\hat{\mathbf{G}}|^{-1/2}\hat{h}e^{-\hat{g}}\{1+\hat{O}_{\text{Cor}}+O(\nu^{-2})\}. \tag{33}$$

The correction term $\hat{O}_{\text{Cor}}$ has order $O(\nu^{-1})$ and is computed below so the overall approximation achieves the indicated $O(\nu^{-2})$ relative error. Our integral expressions, however, will assume $\nu = 1$. The correction term $\hat{O}_{\text{Cor}}$ can be derived by using Taylor expansions of h and g about $\hat{\mathbf{u}}$ to give

$$\hat{h}\hat{O}_{\text{Cor}} = \tfrac{1}{2}\hat{h}_{ij}\mu^{ij} - \tfrac{1}{6}\hat{h}_i\hat{g}_{jkl}\mu^{ijkl} - \tfrac{1}{24}\hat{h}\hat{g}_{ijkl}\mu^{ijkl} + \tfrac{1}{72}\hat{h}\hat{g}_{ijk}\hat{g}_{uvw}\mu^{ijkuvw}. \tag{34}$$

Einstein summation notation has been employed in this expression and the subscripted notation is, for example, $\hat{h}_{ij} = \partial^2 g/\partial u_i\partial u_j|_{\mathbf{u}=\hat{\mathbf{u}}}$, with $\mu^{\alpha\beta\cdots}$ given by

$$\mu^{\alpha\beta\cdots} = (2\pi)^{-p/2}|\hat{\mathbf{G}}|^{1/2}\int_{\mathbf{u}\in\Re^p}\exp\{-\tfrac{1}{2}u^iu^j\hat{g}_{ij}\}(u^\alpha u^\beta\cdots)d\mathbf{u}.$$

Properties of higher-order multinormal moments $\mu^{\alpha\beta\cdots}$ allow (34) to be rewritten as

$$\hat{h}\hat{O}_{\text{Cor}} = \tfrac{1}{2}\operatorname{tr}\hat{\mathbf{H}}\hat{\mathbf{G}}^{-1} - \tfrac{1}{2}\hat{h}_i\hat{g}_{jkl}\hat{g}^{ij}\hat{g}^{kl} - \tfrac{1}{8}\hat{h}\hat{g}_{ijkl}\hat{g}^{ij}\hat{g}^{kl} + \tfrac{1}{8}\hat{h}\hat{g}_{ijk}\hat{g}_{uvw}\hat{g}^{ij}\hat{g}^{uv}\hat{g}^{kw} + \tfrac{1}{12}\hat{h}\hat{g}_{ijk}\hat{g}_{uvw}\hat{g}^{iu}\hat{g}^{jv}\hat{g}^{kw} \tag{35}$$

where $\hat{\mathbf{H}} = (\hat{h}_{ij})$ and $\hat{\mathbf{G}}^{-1} = (\hat{g}^{ij})$ are $n\times n$ matrices. Defining

$$\hat{\varsigma}_i = \hat{g}_{ijk}\hat{g}^{jk} \qquad \text{and} \qquad \hat{t}_{ij} = \hat{g}_{ijkl}\hat{g}^{kl}$$

and n-vector $\hat{\boldsymbol{\varsigma}} = (\hat{\varsigma}_i)$, $n\times n$ matrix $\hat{\mathbf{T}} = (\hat{t}_{ij})$, and n-vector $\hat{\mathbf{h}} = (\hat{h}_i)$, then (35) is

$$\hat{h}\hat{O}_{\text{Cor}} = \tfrac{1}{2}\operatorname{tr}\hat{\mathbf{H}}\hat{\mathbf{G}}^{-1} - \tfrac{1}{2}\hat{\mathbf{h}}^T\hat{\mathbf{G}}^{-1}\hat{\boldsymbol{\varsigma}} - \tfrac{1}{8}\hat{h}\operatorname{tr}\hat{\mathbf{T}}\hat{\mathbf{G}}^{-1} + \tfrac{1}{8}\hat{h}\hat{\boldsymbol{\varsigma}}^T\hat{\mathbf{G}}^{-1}\hat{\boldsymbol{\varsigma}} + \tfrac{1}{12}\hat{h}\hat{g}_{ijk}\hat{g}_{uvw}\hat{g}^{iu}\hat{g}^{jv}\hat{g}^{kw}. \tag{36}$$

A.1.2 Expression for $\hat{O}_{A,\mathbf{x}}$ in Theorem 1

The components of the correction term $\hat{O}_{\text{Cor}} = \hat{O}_{A,\mathbf{x}}$ in (36) are easily evaluated by using simple calculus. However, it is of prime importance to find simple expressions to ease the amount of computation. The first two derivatives of h evaluated at $\hat{\mathbf{u}}$ are specified in terms of the n-vectors $\hat{\mathbf{a}}_1 = (\hat{a}_{1i})$ and $\hat{\mathbf{a}}_2 = (\hat{a}_{2i})$ with

$$\hat{a}_{1i} = -1/\hat{u}_i + 1/(1-\hat{u}_i) \quad \text{and} \quad \hat{a}_{2i} = 1/\hat{u}_i^2 + 1/(1-\hat{u}_i)^2.$$

The first derivative is n-vector $\hat{\mathbf{h}} = \hat{\mathbf{a}}_1\hat{h}$ and the Hessian is $\hat{\mathbf{H}} = (\hat{h}_{ij}) = \{\hat{\mathbf{a}}_1\hat{\mathbf{a}}_1^T + \operatorname{diag}(\hat{\mathbf{a}}_2)\}\hat{h}$. Third- and fourth-order derivatives for g at $\hat{\mathbf{u}}$ are specified in terms of the n-vectors $\hat{\mathbf{a}}_3 = (\hat{a}_{3i})$ and $\hat{\mathbf{a}}_4 = (\hat{a}_{4i})$ with

$$\hat{a}_{3i} = -2b_i/\hat{u}_i^3 + 2(c_i-b_i)/(1-\hat{u}_i)^3 \quad \text{and} \quad \hat{a}_{4i} = 6b_i/\hat{u}_i^4 + 6(c_i-b_i)/(1-\hat{u}_i)^4,$$

so that

$$\hat{g}_{ijk} = 2dx_ix_jx_k/(1-\hat{\mathbf{u}}^T\mathbf{x})^3 + \hat{a}_{3i}\delta_{ijk} \quad \text{and} \quad \hat{g}_{ijkl} = 6dx_ix_jx_kx_l/(1-\hat{\mathbf{u}}^T\mathbf{x})^4 + \hat{a}_{4i}\delta_{ijkl}, \tag{37}$$

where δ_{ijk} is the indicator that $i=j=k$, etc. Then

$$\hat{\varsigma}_i = \hat{g}_{ijk}\hat{g}^{jk} = 2d(\mathbf{x}^T\hat{\mathbf{G}}^{-1}\mathbf{x})x_i/(1-\hat{\mathbf{u}}^T\mathbf{x})^3 + \hat{a}_{5i}$$
$$\hat{t}_{ij} = \hat{g}_{ijkl}\hat{g}^{kl} = 6d(\mathbf{x}^T\hat{\mathbf{G}}^{-1}\mathbf{x})x_ix_j/(1-\hat{\mathbf{u}}^T\mathbf{x})^4 + \hat{a}_{6i}\delta_{ij}$$

where n-vectors $\hat{\mathbf{a}}_5 = \mathrm{diag}(\hat{\mathbf{G}}^{-1})\hat{\mathbf{a}}_3$ and $\hat{\mathbf{a}}_6 = \mathrm{diag}(\hat{\mathbf{G}}^{-1})\hat{\mathbf{a}}_4$. This yields

$$\hat{\boldsymbol{\varsigma}} = 2d^{-2}\hat{\lambda}^3(\mathbf{x}^T\hat{\mathbf{G}}^{-1}\mathbf{x})\mathbf{x} + \hat{\mathbf{a}}_5 \quad \text{and} \quad \hat{\mathbf{T}} = 6d^{-3}\hat{\lambda}^4(\mathbf{x}^T\hat{\mathbf{G}}^{-1}\mathbf{x})\mathbf{x}\mathbf{x}^T + \mathrm{diag}(\hat{\mathbf{a}}_6).$$

Using (37), the final term in (36) is

$$\tfrac{1}{12}\hat{h}\hat{g}_{ijk}\hat{g}_{uvw}\hat{g}^{iu}\hat{g}^{jv}\hat{g}^{kw} = \hat{h}\{\tfrac{1}{3}d^{-4}\hat{\lambda}^6(\mathbf{x}^T\hat{\mathbf{G}}^{-1}\mathbf{x})^3 + \tfrac{1}{12}\hat{\mathbf{a}}_3^T\hat{\boldsymbol{\Gamma}}\hat{\mathbf{a}}_3 + \tfrac{1}{3}\hat{\mathbf{a}}_3^T\hat{\mathbf{a}}_7\},$$

where $n\times n$ matrix $\hat{\boldsymbol{\Gamma}} = \{(\hat{g}^{ij})^3\}$ and n-vector $\hat{\mathbf{a}}_7 = (a_{7i})$ with $a_{7i} = d^{-2}\hat{\lambda}^3\{(\hat{\mathbf{G}}^{-1}\mathbf{x})_i\}^3$.

After including all these reduced summations, $O_{A,\mathbf{x}}$ in (36) is

$$\begin{aligned}\hat{O}_{A,\mathbf{x}} = {} & \tfrac{1}{2}\hat{\mathbf{a}}_1^T\hat{\mathbf{G}}^{-1}(\hat{\mathbf{a}}_1 - \hat{\boldsymbol{\varsigma}}) + (\hat{a}_2/2 - \hat{\mathbf{a}}_6/8)^T\,\mathrm{diag}(\hat{\mathbf{G}}^{-1})\mathbf{1} - \tfrac{3}{4}d^{-3}\hat{\lambda}^4(\mathbf{x}^T\hat{\mathbf{G}}^{-1}\mathbf{x})^2 \\ & + \tfrac{1}{8}\hat{\boldsymbol{\varsigma}}^T\hat{\mathbf{G}}^{-1}\hat{\boldsymbol{\varsigma}} + \tfrac{1}{3}d^{-4}\hat{\lambda}^6(\mathbf{x}^T\hat{\mathbf{G}}^{-1}\mathbf{x})^3 + \tfrac{1}{12}\hat{\mathbf{a}}_3^T\hat{\boldsymbol{\Gamma}}\hat{\mathbf{a}}_3 + \tfrac{1}{3}\hat{\mathbf{a}}_3^T\hat{\mathbf{a}}_7 \end{aligned} \tag{38}$$

where $\mathbf{1}$ is an n-vector of ones. The value of (38) when $\mathbf{x}=\mathbf{0}$ determines the second-order correction term $\hat{O}_{A,\mathbf{0}}$ computed with n-vector $\hat{\mathbf{u}}_0 = (b_i/c_i)$.

A.1.3 Expression for $\hat{O}_{D,\mathbf{x}}$ in Theorem 2

The computation of this correction term follows the same pattern of computation as for $F_A^{(n)}$ so only the main expressions are given. To express the derivatives of h and g at $\hat{\mathbf{u}}$, the following n-vectors $\hat{\mathbf{d}}_1 = (\hat{d}_{1i})$, etc. are needed with components

$$\hat{d}_{1i} = -1/\hat{u}_i + 1/(1-\hat{u}_+), \quad \hat{d}_{2i} = 1/\hat{u}_i^2, \quad \hat{d}_{3i} = -2b_i/\hat{u}_i^3, \quad \hat{d}_{4i} = 6b_i/\hat{u}_i^4. \tag{39}$$

For expansion (36), the Hessian of h is

$$\hat{\mathbf{H}} = (\hat{h}_{ij}) = \{\hat{\mathbf{d}}_1\hat{\mathbf{d}}_1^T + (c-b_+)^{-2}\hat{\mu}^2\mathbf{1}\mathbf{1}^T + \mathrm{diag}(\hat{\mathbf{d}}_2)\}\hat{h}.$$

Also

$$\begin{aligned}\hat{\boldsymbol{\varsigma}} &= 2(c-b_+)^{-2}\hat{\mu}^3(\mathbf{1}^T\hat{\mathbf{G}}^{-1}\mathbf{1})\mathbf{1} + 2d^{-2}\hat{\lambda}^3(\mathbf{x}^T\hat{\mathbf{G}}^{-1}\mathbf{x})\mathbf{x} + \hat{\mathbf{d}}_5 \\ \hat{\mathbf{T}} &= 6(c-b_+)^{-3}\hat{\mu}^4(\mathbf{1}^T\hat{\mathbf{G}}^{-1}\mathbf{1})\mathbf{1}\mathbf{1}^T + 6d^{-3}\hat{\lambda}^4(\mathbf{x}^T\hat{\mathbf{G}}^{-1}\mathbf{x})\mathbf{x}\mathbf{x}^T + \mathrm{diag}(\mathbf{d}_6)\end{aligned}$$

where n-vectors $\hat{\mathbf{d}}_5 = \mathrm{diag}(\hat{\mathbf{G}}^{-1})\hat{\mathbf{d}}_3$ and $\hat{\mathbf{d}}_6 = \mathrm{diag}(\hat{\mathbf{G}}^{-1})\hat{\mathbf{d}}_4$. The final term in (36) is

$$\begin{aligned}\tfrac{1}{12}\hat{h}\hat{g}_{ijk}\hat{g}_{uvw}\hat{g}^{iu}\hat{g}^{jv}\hat{g}^{kw} = \hat{h}\Big\{ & \tfrac{1}{3}(c-b_+)^{-4}\hat{\mu}^6(\mathbf{1}^T\hat{\mathbf{G}}^{-1}\mathbf{1})^3 + \tfrac{1}{3}d^{-4}\hat{\lambda}^6(\mathbf{x}^T\hat{\mathbf{G}}^{-1}\mathbf{x})^3 + \\ & \tfrac{2}{3}(c-b_+)^{-2}d^{-2}\hat{\mu}^3\hat{\lambda}^3(\mathbf{x}^T\hat{\mathbf{G}}^{-1}\mathbf{1})^3 + \tfrac{1}{12}\hat{\mathbf{d}}_3^T\hat{\boldsymbol{\Gamma}}\hat{\mathbf{d}}_3 + \tfrac{1}{3}\hat{\mathbf{d}}_3^T\hat{\mathbf{d}}_7\Big\},\end{aligned}$$

where $n\times n$ matrix $\hat{\boldsymbol{\Gamma}} = \{(\hat{g}^{ij})^3\}$ and n-vector $\hat{\mathbf{d}}_7 = (\hat{d}_{7i})$ has

$$\hat{d}_{7i} = (c-b_+)^{-2}\hat{\mu}^3\{(\hat{\mathbf{G}}^{-1}\mathbf{1})_i\}^3 + d^{-2}\hat{\lambda}^3\{(\hat{\mathbf{G}}^{-1}\mathbf{x})_i\}^3.$$

The final (36) expression reduces to

$$\begin{aligned}\hat{O}_{D,\mathbf{x}} = &\tfrac{1}{2}\hat{\mathbf{d}}_1^T\hat{\mathbf{G}}^{-1}(\hat{\mathbf{d}}_1 - \hat{\boldsymbol{\varsigma}}) + \tfrac{1}{2}(c-b_+)^{-2}\hat{\mu}^2(\mathbf{1}^T\hat{\mathbf{G}}^{-1}\mathbf{1}) + (\hat{\mathbf{d}}_2/2 - \hat{\mathbf{d}}_6/8)^T\,\mathrm{diag}(\hat{\mathbf{G}}^{-1})\mathbf{1} \\ &-\tfrac{3}{4}(c-b_+)^{-3}\hat{\mu}^4(\mathbf{1}^T\hat{\mathbf{G}}^{-1}\mathbf{1})^2 - \tfrac{3}{4}d^{-3}\hat{\lambda}^4(\mathbf{x}^T\hat{\mathbf{G}}^{-1}\mathbf{x})^2 + \tfrac{1}{8}\hat{\boldsymbol{\varsigma}}^T\hat{\mathbf{G}}^{-1}\hat{\boldsymbol{\varsigma}} + \tfrac{1}{3}(c-b_+)^{-4}\hat{\mu}^6(\mathbf{1}^T\hat{\mathbf{G}}^{-1}\mathbf{1})^3 \\ &+\tfrac{1}{3}d^{-4}\hat{\lambda}^6(\mathbf{x}^T\hat{\mathbf{G}}^{-1}\mathbf{x})^3 + \tfrac{2}{3}(c-b_+)^{-2}d^{-2}\hat{\mu}^3\hat{\lambda}^3(\mathbf{x}^T\hat{\mathbf{G}}^{-1}\mathbf{1})^3 + \tfrac{1}{12}\hat{\mathbf{d}}_3^T\hat{\boldsymbol{\Gamma}}\hat{\mathbf{d}}_3 + \tfrac{1}{3}\hat{\mathbf{d}}_3^T\hat{\mathbf{d}}_7. \end{aligned} \tag{40}$$

The value of (40) when $\mathbf{x} = \mathbf{0}$ determines the second-order correction term $\hat{O}_{D,\mathbf{0}}$. In this case, the solution to (21) is $\hat{\mu}_0 = c$ and $\hat{\lambda}_0 = d$ which leads to n-vector $\hat{\mathbf{u}}_0 = \mathbf{b}/c$.

A.1.4 Determinant of Hessian in (25)

If

$$\mathbf{A} = \mathbf{I}_n + \mathbf{y}\mathbf{y}^T + \mathbf{z}\mathbf{z}^T$$

where $\mathbf{I}_n$ is the $n \times n$ identity matrix and $\mathbf{y}$ and $\mathbf{z}$ are $n \times 1$ vectors, then

$$\mathbf{A} = (\mathbf{I}_n + \mathbf{y}\mathbf{y}^T)^{1/2}\{\mathbf{I}_n + (\mathbf{I}_n + \mathbf{y}\mathbf{y}^T)^{-1/2}\mathbf{z}\mathbf{z}^T(\mathbf{I}_n + \mathbf{y}\mathbf{y}^T)^{-1/2}\}(\mathbf{I}_n + \mathbf{y}\mathbf{y}^T)^{1/2}.$$

Using the standard results

$$|\mathbf{I}_n + \mathbf{y}\mathbf{y}^T| = 1 + \mathbf{y}^T\mathbf{y} \quad \text{and} \quad (\mathbf{I}_n + \mathbf{y}\mathbf{y}^T)^{-1} = \mathbf{I}_n - (\mathbf{1} + \mathbf{y}^T\mathbf{y})^{-1}\mathbf{y}\mathbf{y}^T,$$

it follows that

$$\begin{aligned}|\mathbf{A}| &= (1 + \mathbf{y}^T\mathbf{y})\{1 + \mathbf{z}^T(\mathbf{I}_n + \mathbf{y}\mathbf{y}^T)^{-1}\mathbf{z}\} = (1 + \mathbf{y}^T\mathbf{y})[1 + \mathbf{z}^T\{\mathbf{I}_n - (\mathbf{1} + \mathbf{y}^T\mathbf{y})^{-1}\mathbf{y}\mathbf{y}^T\}\mathbf{z}] \\ &= (1 + \mathbf{y}^T\mathbf{y})(1 + \mathbf{z}^T\mathbf{z}) - (\mathbf{y}^T\mathbf{z})^2. \end{aligned} \tag{41}$$

The first term in (25) is $|\hat{\mathbf{D}}|$ and the second term $\hat{\Xi}$ is the computation of (41) with

$$\mathbf{y} = (c - b_+)^{-1/2}\hat{\mu}\hat{\mathbf{D}}^{-1/2}\mathbf{1} \qquad \mathbf{z} = d^{-1/2}\hat{\lambda}\hat{\mathbf{D}}^{-1/2}\mathbf{x}.$$

A.2 Proof of Theorem 4

Conditions (ii) and (iii) of Theorem 4 imply that, for $a < 0$ and $n \geq 1$,

$$0 \leq \left(1 - \sum_{i=1}^n V_i^{(n)} x_i^{(n)}\right)^{-a} \leq (1 + A)^{-a},$$

where $V_i^{(n)} \sim \mathrm{Beta}(b_i^{(n)}, c_i^{(n)} - b_i^{(n)})$ and $V_1^{(n)}, \ldots, V_n^{(n)}$ are independent for $n \geq 1$. Therefore the family of random variables $\{W_n = (1 - \sum_{i=1}^n V_i^{(n)} x_i^{(n)})^{-a} : n \geq 1\}$ is uniformly integrable; see e.g. Rogers and Williams (1994, p. 115). Note that $F_A^{(n)}(a, \mathbf{b}^{(n)}; \mathbf{c}^{(n)}; \mathbf{x}^{(n)}) = E(W_n)$ for $n \geq 1$.

As $n \to \infty$, condition (v) of the theorem implies that

$$E\left(\sum_{i=1}^n V_i^{(n)} x_i^{(n)}\right) = \sum_{i=1}^n \pi_i^{(n)} x_i^{(n)} = \xi_n \to \xi_0;$$

and, using conditions (i)-(iii) and standard formulae for the mean and variance of the beta distribution,

$$\begin{aligned}\mathrm{Var}\left(\sum_{i=1}^{n} V_i^{(n)} x_i^{(n)}\right) &= \sum_{i=1}^{n} \pi_i^{(n)}(1-\pi_i^{(n)})(x_i^{(n)})^2/(c_i^{(n)}+1) \\ &\leq \max_{1\leq i\leq n} \frac{|x_i^{(n)}|}{c_i^{(n)}}(1+A) \to 0.\end{aligned}$$

Consequently, since L^2 convergence implies convergence in probability, $\sum_{i=1}^{n} V_i^{(n)} x_i^{(n)} \to \xi_0$ in probability. Therefore, since $\{W_n\}_{n\geq 1}$ is uniformly integrable, it follows from Rogers and Williams (1994, Theorem 21.2) that

$$\lim_{n\to\infty} F_A^{(n)}(a, \mathbf{b}^{(n)}; \mathbf{c}^{(n)}; \mathbf{x}^{(n)}) = (1-\xi_0)^{-a}.$$

Under the conditions of the theorem, $\hat{F}_A^{(n)}$ has the same limit. A key step in proving this is to establish that

$$\begin{aligned}\hat{u}_i^{(n)} &= \tilde{u}_i(\hat{\lambda}^{(n)}) \\ &= \pi_i^{(n)}\left[1 - (\hat{\lambda}^{(n)} x_i^{(n)}/c_i^{(n)})(1-\pi_i^{(n)}) + O\left\{\left(x_i^{(n)}/c_i^{(n)}\right)^2\right\}\right] \\ &= \pi_i^{(n)}\left[1 + a\phi_i^{(n)}\theta^{(n)} x_i^{(n)}(1-\pi_i^{(n)})/\{c_i^{(n)}(1-\xi_0)\} + O\{(x_i^{(n)}/c_i^{(n)})^2\}\right], \qquad (42)\end{aligned}$$

where $\hat{\lambda}^{(n)} = -a\theta^{(n)}/(1-\xi_0)$,

$$\max_{1\leq i\leq n} |\phi_i^{(n)} - 1| \to 0 \quad \text{and} \quad \theta^{(n)} \to 1. \qquad (43)$$

In (42), $\tilde{u}_i(\lambda)$ is defined in (13), and (43) follows from conditions (i), (iv) and (v); details of the proof are straightforward but laborious and are omitted. Further calculations of a similar nature, using (42) and (43), show that

$$|\hat{\mathbf{G}}^{(n)}| \sim |\hat{\mathbf{D}}^{(n)}| \sim \prod_{i=1}^{n} 1/\{\pi_i^{(n)}(1-\pi_i^{(n)})\},$$

and then, after substitution into the RHS of (15) and cancellation, (32) follows. The final part of the theorem is a consequence of the fact that all terms in $\hat{O}_{D,\mathbf{x}}$ involving $\mathbf{x}$ have factor $\mathbf{x}^\top \hat{\mathbf{G}}^{-1}\mathbf{x}$ which converges to 0 as a consequence of condition (i), (42) and (43), while the sum of terms not involving $\mathbf{x}$ converges to $\hat{O}_{D,\mathbf{0}}$.

A.3 Proof of Theorem 5

The structure of the proof of Theorem 5 is similar to that of Theorem 4 although some of the details are more complex. We shall only give selected details. Suppose $(U_1^{(n)}, \ldots, U_{n+1}^{(n)})^\top$ has a Dirichlet$(b_1^{(n)}, \ldots, b_n^{(n)}, c^{(n)} - b_+^{(n)})$ distribution. Then

$$F_D^{(n)}(a, \mathbf{b}^{(n)}; c^{(n)}; \mathbf{x}^{(n)}) = E(W_n)$$

where now $W_n = (1 - \sum_{i=1}^{n} U_i^{(n)} x_i^{(n)})^{-a}$. Then $E(\sum_{i=1}^{n} U_i^{(n)} x_i^{(n)}) = \xi_n \to \xi_0$ by condition (iv) and, using the well-known expressions for the second moments of the Dirichlet

distribution plus conditions (i)-(iii),

$$\begin{aligned}\mathrm{Var}\left(\sum_{i=1}^{n} U_i^{(n)} x_i^{(n)}\right) &= \left(\sum_{i=1}^{n} \frac{\pi_i^{(n)}(x_i^{(n)} - \xi_n)^2}{c^{(n)}+1}\right) + \frac{(1-\pi_+^{(n)})\xi_n^2}{c^{(n)}+1} \\ &\le \sum_{i=1}^{n} \frac{\pi_i^{(n)}\{x_i^{(n)}\}^2}{c^{(n)}+1} \\ &\le \frac{\max_{1\le i\le n}|x_i^{(n)}|}{c^{(n)}}(1+A) \to 0,\end{aligned}$$

where $\pi_+^{(n)} = \sum \pi_i^{(n)}$. Using reasoning the same as that in the proof of Theorem 4, it is seen that $\lim_{n\to\infty} F_D^{(n)}(a, \mathbf{b}^{(n)}; \mathbf{c}^{(n)}; \mathbf{x}^{(n)}) = (1-\xi_0)^{-a}$.

To show that $\hat{F}_D^{(n)}(a, \mathbf{b}^{(n)}; c^{(n)}; \mathbf{x}^{(N)})$ has the same limit we first derive an expansion for $\rho = \lambda/\mu$. From (24) it is seen that

$$\rho\left(1 - \pi_+^{(n)} + \sum_{i=1}^{n} \frac{\pi_i^{(n)}}{1+\rho x_i^{(n)}}\right) = \rho \sum_{i=1}^{n} \frac{\pi_i^{(n)} x_i^{(n)}}{1+\rho x_i^{(n)}} + \frac{d}{c^{(n)}}. \tag{44}$$

As d is fixed and $c^{(n)} \to \infty$ as $n \to \infty$ by condition (i) of the theorem, the second term on the RHS of (44) converges to 0, and it is appropriate to expand both sides of (44) about $\rho = 0$, from which we obtain $\rho = [d/\{c^{(n)}(1-\xi_n)\}]\{1+o(1)\}$, and consequently, using (23), we obtain

$$\hat{\mu} = \left(c^{(n)} - \frac{d\xi_n}{1-\xi_n}\right)\{1+o(1)\} \quad \text{and} \quad \hat{\lambda} = \frac{d\xi_n}{1-\xi_n}\{1+o(1)\}.$$

Therefore

$$\hat{u}_i^{(n)} = \tilde{u}_i(\hat{\mu}, \hat{\lambda}) = \frac{b_i^{(n)}}{\hat{\mu} + \hat{\lambda} x_i^{(n)}} = \pi_i^{(n)}\left(1 + \frac{d\xi_n(1-x_i^{(n)})}{c^{(n)}(1-\xi_n)} + O\{(c^{(n)})^{-2}\}\right),$$

and some further calculations show that

$$\begin{aligned}1 - \hat{u}_+^{(n)} &= 1 - \sum_{i=1}^{n} \hat{u}_i^{(n)} \\ &= 1 - \sum_{i=1}^{n} \pi_i^{(n)}\left(1 + \frac{d\xi_n}{c^{(n)}(1-\xi_n)} + O\{(c^{(n)})^{-2}\}\right) \\ &= (1-\pi_+^{(n)})\left(1 - \frac{d\xi_n(\pi_+^{(n)} - \xi_n)}{c^{(n)}(1-\xi_n)(1-\pi_+^{(n)})} + O\{(c^{(n)})^{-2}\}\right).\end{aligned}$$

Consequently,

$$\begin{aligned}&\left(\sum_{i=1}^{n} b_i^{(n)} \log \hat{u}_i^{(n)}\right) - (c^{(n)} - b_+^{(n)}) \log(1 - \hat{u}_+^{(n)}) \\ &\qquad = \left(\sum_{i=1}^{n} b_i^{(n)} \log \pi_i^{(n)}\right) - (c^{(n)} - b_i^{(n)}) \log(1 - \pi_+^{(n)}) + o(1).\end{aligned}$$

The remaining steps in the proof closely follow those of Theorem 4; the details are omitted.

A.4 Outline of proofs of (30) and (31)

Full details of the proof of (30) and (31) in all cases is very lengthy but at the same time fairly elementary. We outline the steps we have followed in proving these results.

1. In all cases of (30) and (31) considered, the point at which the integrand in (4) or (6) is maximized converges, as $\nu \to \infty$, to a point in $[0,1]^n$, $\mathbf{u}^\dagger = (u_1^\dagger, \ldots, u_n^\dagger)^\top$, say. To evaluate an asymptotic value of the denominator, proceed as follows: (i) for those i such that $u_i^\dagger = 0$, we use the transformation $u_i = s_i/\nu$; (ii) for those i such that $u_i^\dagger = 1$, we use the transformation $u_i = 1 - s_i/\nu$; and (iii) for the remaining u_i we use a suitable form of Laplace's approximation, with g and h in (33) chosen appropriately. It turns out the leading-order term of the denominator on the LHS of (30) and (31) is straightforward to determine in each case, using a suitable combination of (i), (ii) and (iii).

2. We now consider the numerator on the LHS of (30) and (31). It can be shown that, in all cases under consideration, $\hat{\mathbf{u}} = (\hat{u}_1, \ldots, \hat{u}_n)^\top$, defined to be $\tilde{\mathbf{u}}(\hat{\lambda})$ in §3 in the case of $\hat{F}_A^{(n)}$ and $\tilde{\mathbf{u}}(\hat{\lambda}, \hat{\mu})$ in §4 in the case of $\hat{F}_D^{(n)}$, converges to $\mathbf{u}^\dagger$ as $\nu \to \infty$. Moreover, further calculations show that, for each i, $u_i^\dagger = \hat{u}_i + \delta_i/\nu + O(\nu^{-2})$ for some real number δ_i.

3. With further calculations, and using point 2 above, it can be shown in each case that the LHS of (30) and (31) remains bounded away from 0 and ∞ as $\nu \to \infty$. However, in most cases the limit is not equal to 1, in which case the relative error does not go to 0.

Acknowledgment. We are grateful to Ayman Al-Rawashdeh for his help in writing the Matlab and R programs. We also thank two referees for reading the paper carefully. The first author has been supported by the U.S. National Science Foundation under grant DMS-1104474.

References

[1] Abramowitz, M. and Stegun, I.A. (1972). *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*. Dover: New York.

[2] Backeljauw, F., Becuwe, S., Cuyt, A., Van Deun, J. and Lozier, D. W. (2014). Validated evaluation of special mathematical functions. *Science of Computer Programming*, **90**, 2-20.

[3] Butler, R.W. and Wood, A.T.A. (2002). Laplace approximation for hypergeometric functions of matrix argument. *Ann. Statist.* **30**, 1155-1177.

[4] Butler, R.W. and Wood, A.T.A. (2003). Laplace approximation for Bessel functions of matrix argument. *Journal of Computational and Applied Mathematics*, **155**, 359–382.

[5] Butler, R.W. and Wood, A.T.A. (2014). On statistical applications of Lauricella functions with emphasis on the confluent $\Phi_2^{(n)}$ function. *In preparation*.

[6] Dickey, J.M. (1983). Multiple hypergeometric functions: probabilistic interpretations and statistical uses. *J. Amer. Statist. Assoc.*, **78**, 628–637.

[7] Exton, H. (1976). *Multiple Hypergeometric Functions and Applications. Foreword by L. J. Slater.* Ellis Horwood Ltd., Chichester.

[8] Hsu, L.C. (1951). On the asymptotic evaluation of a class of multiple integrals involving a parameter. *Amer. J. Math.*, **73**, 625—634.

[9] Kerov, S.V. and Tsilevich, N.V. (2004). The Markov-Krein correspondence in several dimensions. *Journal of Mathematical Sciences*, **121**, 2345–2359.

[10] Lauricella, G. (1893). Sulle funzioni ipergeometriche a più variabili (in Italian). *Rendiconti del Circolo Matematico di Palermo*, **7**, 111–158.

[11] Lijoi, A. and Regazzini, E. (2004). Means of a Dirichlet process and multiple hypergeometric functions. *Ann. Probab.*, **32**, 1469–1495.

[12] Muirhead, R.J. (1982). *Aspects of Multivariate Statistical Theory.* John Wiley: New York.

[13] Richards, D. St. P. (2012). Functions of matrix argument. *NIST Digital Library of Mathematical Functions, Chapter 35*, *http://dlmf.nist.gov*

[14] Rogers, L.C.G. and Williams, D. (1994). *Diffusions, Markov Processes, and Martingales. Volume 1: Foundations.* John Wiley: New York.

[15] Scarpello, G.M. and Ritelli, D. (2012). Legendre hyperelliptic intergrals, π new formulae and Lauricella functions through the elliptic singular moduli. *arXiv 1209:1940.*

[16] Van Laarhoven, J.M and Kalker, T.A.C.M. (1988). On the computation of Lauricella functions of the fourth kind. *Journal of Computational and Applied Mathematics* **21**, 369–375.