INFORMATION MATRICES FOR SOME BIVARIATE PARETO DISTRIBUTIONS

by

Samuel Kotz
Department of Engineering Management and Systems Engineering
The George Washington University
Washington, D.C. 20052

Saralees Nadarajah
Department of Statistics
University of Nebraska
Lincoln, NE 68583

ABSTRACT: The Fisher information matrices for four bivariate Pareto distributions are derived.

Keywords and Phrases: Bivariate Pareto distributions, Euler psi function.


1 INTRODUCTION

Pareto distributions are the most popular and the most applied distributions in the field of income and wealth modelling. They are very versatile and a variety of uncertainties can be usefully modeled by them. Some of the other application areas include extreme values, failure times, modeling of birth rates and infant mortality rates, and reliability. In this paper, we consider four of the most popular bivariate Pareto distributions:

- the bivariate Lomax distribution with the joint survivor function and joint pdf

\[ \bar{F}(x, y) = \frac{1}{(1 + \theta x + \phi y)^a} \]  

and

\[ f(x, y) = \frac{a(a + 1)\theta \phi}{(1 + \theta x + \phi y)^{a+2}}, \]  

respectively, for \( x > 0, y > 0, \theta > 0, \phi > 0 \) and \( a > 0 \).

- a variant of the above with the joint pdf

\[ f(x, y) = \frac{b^{2a}\Gamma(2a + c)}{\Gamma(c)\Gamma^2(a)} \frac{x^{a-1}y^{a-1}}{(1 + bx + by)^{2a+c}} \]  

for \( x > 0, y > 0, a > 0, b > 0 \) and \( c > 0 \).
• Muliere and Scarsini (1987)'s bivariate Pareto distribution with the joint pdf

\[
f(x, y) = \begin{cases} 
\frac{\lambda_2 (\lambda_0 + \lambda_1)}{\beta^2} \left( \frac{x}{\beta} \right)^{-(1+\lambda_0+\lambda_1)} \left( \frac{y}{\beta} \right)^{-(1+\lambda_2)}, & \text{if } x > y, \\
\frac{\lambda_0}{\beta} \left( \frac{x}{\beta} \right)^{-(1+\lambda_0+\lambda_1)}, & \text{if } x = y, \\
\frac{\lambda_1 (\lambda_0 + \lambda_2)}{\beta^2} \left( \frac{y}{\beta} \right)^{-(1+\lambda_0+\lambda_2)} \left( \frac{x}{\beta} \right)^{-(1+\lambda_1)}, & \text{if } y > x 
\end{cases}
\]

(4)

for \( x \geq \beta, y \geq \beta, \lambda_0 > 0, \lambda_1 > 0, \lambda_2 > 0 \) and \( \beta > 0 \).

• De Groot (1970)'s bivariate Pareto distribution with the joint pdf

\[
f(x, y) = \gamma (\gamma + 1) (\xi - \eta)^\gamma (y - x)^{-(\gamma + 2)}
\]

(5)

for \( 0 < x < \eta < \xi < y \) and \( \gamma > 1 \).

The aim is to calculate the Fisher information matrix corresponding to each of the distributions above. For a given observation \((x, y)\), the Fisher information matrix is defined by

\[
(I_{jk}) = \left\{ E \left( \frac{\partial \log L(\theta)}{\partial \theta_j} \frac{\partial \log L(\theta)}{\partial \theta_k} \right) \right\}
\]

for \( j = 1, 2, \ldots, p \) and \( k = 1, 2, \ldots, p \), where \( L(\theta) = f(x, y) \) and \( \theta = (\theta_1, \theta_2, \ldots, \theta_p) \) are the parameters of the pdf \( f \). It has the meaning “information about the parameters \( \theta \) contained in the observation \((x, y)\).” The information matrix plays a significant role in statistical inference in connection with estimation, sufficiency and properties of variances of estimators. It is related to the covariance matrix of the estimate of \( \theta \) (being its inverse under certain conditions). See Cox and Hinkley (1974) for details.

The exact forms of the information matrix are derived in Sections 2, 3, 4 and 5. Some technical results required for the derivations are noted in Appendix. The calculations use the Euler psi function defined by

\[
\Psi(a) = \frac{\log \Gamma(a)}{da}.
\]

The properties of this special function can be found in Gradshteyn and Ryzhik (2000).

2 INFORMATION MATRIX FOR BIVARIATE LOMAX

If \((x, y)\) is a single observation from (2) then the log-likelihood function can be written as

\[
\log L(a, \theta, \phi) = \log \{a(a + 1)\theta \phi\} - (a + 2) \log(1 + \theta x + \phi y).
\]

The first-order derivatives are:

\[
\frac{\partial \log L}{\partial \theta} = \frac{1}{\theta} - (a + 2) \frac{x}{1 + \theta x + \phi y},
\]

\[
\frac{\partial \log L}{\partial \phi} = \frac{1}{\phi} - (a + 2) \frac{y}{1 + \theta x + \phi y},
\]
and
\[
\frac{\partial \log L}{\partial a} = \frac{1}{a} + \frac{1}{a+1} - \log (1 + \theta x + \phi y).
\]
The second-order derivatives are:
\[
\frac{\partial^2 \log L}{\partial \theta^2} = -\frac{1}{\theta^2} + (a + 2) \frac{x^2}{(1 + \theta x + \phi y)^2},
\]
\[
\frac{\partial^2 \log L}{\partial \theta \partial \phi} = (a + 2) \frac{xy}{(1 + \theta x + \phi y)^2},
\]
\[
\frac{\partial^2 \log L}{\partial \theta \partial a} = -\frac{x}{1 + \theta x + \phi y},
\]
\[
\frac{\partial^2 \log L}{\partial \phi^2} = -\frac{1}{\phi^2} + (a + 2) \frac{y^2}{(1 + \theta x + \phi y)^2},
\]
\[
\frac{\partial^2 \log L}{\partial \phi \partial a} = -\frac{y}{1 + \theta x + \phi y},
\]
and
\[
\frac{\partial^2 \log L}{\partial a^2} = -\frac{1}{a^2} - \frac{1}{(a+1)^2}.
\]
Using the well-known formula
\[
E (X^m Y^n \mid a) = mn \int_0^\infty \int_0^\infty x^{m-1} y^{n-1} (1 + \theta x + \phi y)^{-a} dydx
\]
(6) (where \(a\) denotes the shape parameter in (1)), we can express the elements of the Fisher information matrix as
\[
E \left( -\frac{\partial^2 \log L}{\partial \theta^2} \right) = \frac{1}{\theta^2} - \frac{a + 2}{3} E (X^3 Y \mid a + 2),
\]
\[
E \left( -\frac{\partial^2 \log L}{\partial \theta \partial \phi} \right) = -\frac{a + 2}{4} E (X^2 Y^2 \mid a + 2),
\]
\[
E \left( -\frac{\partial^2 \log L}{\partial \theta \partial a} \right) = \frac{1}{2} E (X^2 Y \mid a + 1),
\]
\[
E \left( -\frac{\partial^2 \log L}{\partial \phi^2} \right) = \frac{1}{\phi^2} - \frac{a + 2}{3} E (XY^3 \mid a + 2),
\]
\[
E \left( -\frac{\partial^2 \log L}{\partial \phi \partial a} \right) = \frac{1}{2} E (XY^2 \mid a + 1),
\]
and
\[ E \left( -\frac{\partial^2 \log L}{\partial a^2} \right) = \frac{1}{a^2} + \frac{1}{(a+1)^2}. \]

By application of Lemma 1, the expectations above can be calculated as
\[ E \left( X^3Y \mid a+2 \right) = \frac{3B(1,a+1)B(3,a-2)}{\theta^3 \phi}, \]
\[ E \left( XY^3 \mid a+2 \right) = \frac{3B(3,a-1)B(1,a-2)}{\theta \phi^3}, \]
\[ E \left( X^2Y^2 \mid a+2 \right) = \frac{4B(2,a)B(2,a-2)}{\theta^2 \phi^2}, \]
\[ E \left( X^2Y \mid a+1 \right) = \frac{2B(1,a+1)B(2,a-1)}{\theta^2 \phi}, \]
and
\[ E \left( XY^2 \right) = \frac{2B(2,a)B(1,a-1)}{\theta^2 \phi}. \]

### 3 INFORMATION MATRIX FOR A VARIANT OF BIVARIATE LOMAX

If \((x, y)\) is a single observation from (3) then the log-likelihood function can be written as
\[
\log L(a, \theta, \phi) = \log \left[ b^{2a} \Gamma(2a+c)/\left\{ \Gamma^2(a)\Gamma(c) \right\} \right] - (a-1) \log x + (a-1) \log y - (2a+c) \log(1 + bx + by).
\]

The first-order derivatives are:
\[
\frac{\partial \log L}{\partial a} = 2 \log b + 2\Psi(2a+c) - 2\Psi(a) + \log(xy) - 2 \log(1 + bx + by),
\]
\[
\frac{\partial \log L}{\partial b} = \frac{2a}{b} - \frac{(2a+c)(x+y)}{1 + bx + by},
\]
and
\[
\frac{\partial \log L}{\partial c} = \Psi(2a+c) - \Psi(c) - \log(1 + bx + by).
\]

The second-order derivatives are:
\[
\frac{\partial^2 \log L}{\partial a^2} = 4\Psi'(2a+c) - 2\Psi'(a),
\]
\[
\frac{\partial^2 \log L}{\partial a \partial b} = \frac{2}{b} - \frac{2(x + y)}{1 + bx + by},
\]
\[
\frac{\partial^2 \log L}{\partial a \partial c} = 2\Psi'(2a + c),
\]
\[
\frac{\partial^2 \log L}{\partial b^2} = -\frac{2a}{b^2} + \frac{(2a + c)(x + y)^2}{(1 + bx + by)^2},
\]
\[
\frac{\partial^2 \log L}{\partial b \partial c} = -\frac{x + y}{1 + bx + by},
\]
and
\[
\frac{\partial^2 \log L}{\partial c^2} = \Psi'(2a + c) - \Psi'(c).
\]

The elements of the Fisher information matrix corresponding to the above second-order derivatives can be computed as follows: Since certain second-order derivatives are constants, it is clear that
\[
E \left( -\frac{\partial^2 \log L}{\partial a^2} \right) = 2\Psi'(a) - 4\Psi'(2a + c),
\]
\[
E \left( -\frac{\partial^2 \log L}{\partial a \partial c} \right) = -2\Psi'(2a + c),
\]
and
\[
E \left( -\frac{\partial^2 \log L}{\partial c^2} \right) = \Psi'(c) - \Psi'(2a + c).
\]

Using the relation
\[
E (X^m Y^n | c) = \frac{b^{2a} \Gamma(2a + c)}{\Gamma(c) \Gamma^2(a)} \int_0^\infty \int_0^\infty \frac{x^{m+a-1} y^{n+a-1}}{(1 + bx + by)^{2a+c}} dy dx
\]  
(7)

(where \(c\) denotes the shape parameter in (3)), the remaining elements of the Fisher information matrix can be expressed as
\[
E \left( -\frac{\partial^2 \log L}{\partial a \partial b} \right) = -\frac{2}{b} + \frac{2c}{2a + c} \left\{ E (X | c + 1) + E (Y | c + 1) \right\},
\]
\[
E \left( -\frac{\partial^2 \log L}{\partial b^2} \right) = \frac{2a}{b^2} - \frac{c(c + 1)}{2a + c + 1} \left\{ E (X^2 | c + 2) + E (Y^2 | c + 2) + 2E (XY | c + 2) \right\},
\]
and
\[
\frac{\partial^2 \log L}{\partial b \partial c} = \frac{c}{2a + c} \left\{ E (X | c + 1) + E (Y | c + 1) \right\}.
\]
By application of Lemma 2, the expectations above can be calculated as

\[ E (X^2 | c + 2) = \frac{a(a + 1)}{b^2 c(c + 1)}, \]

\[ E (Y^2 | c + 2) = \frac{a(a + 1)}{b^2 c(c + 1)}, \]

\[ E (XY | c + 2) = \frac{a^2}{b^2 c(c + 1)}, \]

\[ E (X | c + 1) = \frac{a}{b c}, \]

\[ E (Y | c + 1) = \frac{a}{b c}. \]

4 INFORMATION MATRIX FOR MULIERE AND SCARSINI’s BIVARIATE PARETO

If \((x, y)\) is a single observation from (4) then the log-likelihood function can be written as

\[
\log L(\lambda_0, \lambda_1, \lambda_2, \beta) = \begin{cases} 
\log \left\{ \lambda_2 (\lambda_0 + \lambda_1) \beta^{\lambda_0 + \lambda_1 + \lambda_2} \right\} - (\lambda_0 + \lambda_1 + 1) \log x - (\lambda_2 + 1) \log y, & \text{if } x > y \geq \beta, \\
\log \left\{ \lambda_0 \beta^{\lambda_0 + \lambda_1 + \lambda_2} \right\} - (\lambda_0 + \lambda_1 + \lambda_2 + 1) \log x, & \text{if } x = y \geq \beta, \\
\log \left\{ \lambda_1 (\lambda_0 + \lambda_2) \beta^{\lambda_0 + \lambda_1 + \lambda_2} \right\} - (\lambda_0 + \lambda_2 + 1) \log x - (\lambda_1 + 1) \log y, & \text{if } y > x \geq \beta.
\end{cases}
\]

The first-order derivatives are:

\[
\frac{\partial \log L}{\partial \lambda_0} = \begin{cases} 
\frac{1}{\lambda_0 + \lambda_1} + \log \beta - \log x, & \text{if } x > y \geq \beta, \\
\frac{1}{\lambda_0} + \log \beta - \log x, & \text{if } x = y \geq \beta, \\
\frac{1}{\lambda_0 + \lambda_2} + \log \beta - \log y, & \text{if } y > x \geq \beta,
\end{cases}
\]

\[
\frac{\partial \log L}{\partial \lambda_1} = \begin{cases} 
\frac{1}{\lambda_0 + \lambda_1} + \log \beta - \log x, & \text{if } x > y \geq \beta, \\
\log \beta - \log x, & \text{if } x = y \geq \beta, \\
\frac{1}{\lambda_1} + \log \beta - \log x, & \text{if } y > x \geq \beta,
\end{cases}
\]

\[
\frac{\partial \log L}{\partial \lambda_2} = \begin{cases} 
\frac{1}{\lambda_2} + \log \beta - \log y, & \text{if } x > y \geq \beta, \\
\log \beta - \log x, & \text{if } x = y \geq \beta, \\
\frac{1}{\lambda_0 + \lambda_2} + \log \beta - \log y, & \text{if } y > x \geq \beta,
\end{cases}
\]
\[ \frac{\partial \log L}{\partial \beta} = \frac{\lambda_0 + \lambda_1 + \lambda_2}{\beta}, \]

\[ \frac{\partial^2 \log L}{\partial \lambda_0^2} = \begin{cases} 
- \frac{1}{(\lambda_0 + \lambda_1)^2}, & \text{if } x > y \geq \beta, \\
- \frac{1}{\lambda_0^2}, & \text{if } x = y \geq \beta, \\
- \frac{1}{(\lambda_0 + \lambda_2)^2}, & \text{if } y > x \geq \beta,
\end{cases} \]

\[ \frac{\partial^2 \log L}{\partial \lambda_0 \partial \lambda_1} = \begin{cases} 
- \frac{1}{(\lambda_0 + \lambda_1)^2}, & \text{if } x > y \geq \beta, \\
0, & \text{if } x = y \geq \beta, \\
0, & \text{if } y > x \geq \beta,
\end{cases} \]

\[ \frac{\partial^2 \log L}{\partial \lambda_0 \partial \lambda_2} = \begin{cases} 
0, & \text{if } x > y \geq \beta, \\
0, & \text{if } x = y \geq \beta, \\
- \frac{1}{(\lambda_0 + \lambda_2)^2}, & \text{if } y > x \geq \beta,
\end{cases} \]

\[ \frac{\partial^2 \log L}{\partial \lambda_1 \partial \lambda_2} = 0, \]

\[ \frac{\partial^2 \log L}{\partial \lambda_1 \partial \beta} = \frac{1}{\beta}, \]

\[ \frac{\partial^2 \log L}{\partial \lambda_2 \partial \beta} = \frac{1}{\beta}, \]

\[ \frac{\partial^2 \log L}{\partial \lambda_0 \partial \lambda_1} = \begin{cases} 
- \frac{1}{(\lambda_0 + \lambda_1)^2}, & \text{if } x > y \geq \beta, \\
0, & \text{if } x = y \geq \beta, \\
- \frac{1}{\lambda_1^2}, & \text{if } y > x \geq \beta,
\end{cases} \]

\[ \frac{\partial^2 \log L}{\partial \lambda_0 \partial \lambda_2} = \begin{cases} 
- \frac{1}{(\lambda_0 + \lambda_2)^2}, & \text{if } x > y \geq \beta, \\
0, & \text{if } x = y \geq \beta, \\
- \frac{1}{(\lambda_0 + \lambda_2)^2}, & \text{if } y > x \geq \beta,
\end{cases} \]

\[ \frac{\partial^2 \log L}{\partial \lambda_2 \partial \beta} = \frac{1}{\beta}, \]
and
\[ \frac{\partial^2 \log L}{\partial \beta^2} = -\frac{\lambda_0 + \lambda_1 + \lambda_2}{\beta^2}. \]

Since the second derivative with respect to \( \lambda_1 \) and \( \lambda_2 \) is zero it follows that the maximum likelihood estimates of the two parameters are independent. The elements of the Fisher information matrix corresponding to the non-zero second-order derivatives can be computed as follows: Since certain second-order derivatives are constants, it is clear that
\[
E \left( -\frac{\partial^2 \log L}{\partial \lambda_0 \partial \beta} \right) = -\frac{1}{\beta},
\]
\[
E \left( -\frac{\partial^2 \log L}{\partial \lambda_1 \partial \beta} \right) = -\frac{1}{\beta},
\]
\[
E \left( -\frac{\partial^2 \log L}{\partial \lambda_2 \partial \beta} \right) = -\frac{1}{\beta},
\]
and
\[
E \left( -\frac{\partial^2 \log L}{\partial \beta^2} \right) = \frac{\lambda_0 + \lambda_1 + \lambda_2}{\beta^2}.
\]

Since
\[
\Pr(X < Y) = \lambda_1 (\lambda_0 + \lambda_2) \beta^{\lambda_0 + \lambda_1 + \lambda_2} \int_{\beta}^{\infty} \int_{x}^{\infty} x^{-(\lambda_1 + 1)} y^{-(\lambda_0 + \lambda_2 + 1)} dy dx
\]
\[
= \lambda_1 \beta^{\lambda_0 + \lambda_1 + \lambda_2} \int_{\beta}^{\infty} x^{-(\lambda_0 + \lambda_1 + \lambda_2 + 1)} dx
\]
\[
= \frac{\lambda_1}{\lambda_0 + \lambda_1 + \lambda_2},
\]
the elements of the information matrix can be computed as
\[
E \left( -\frac{\partial^2 \log L}{\partial \lambda_0^2} \right) = \frac{\lambda_2}{(\lambda_0 + \lambda_1)^2 (\lambda_0 + \lambda_1 + \lambda_2)} + \frac{1}{\lambda_0 (\lambda_0 + \lambda_1 + \lambda_2)} + \frac{\lambda_1}{(\lambda_0 + \lambda_2)^2 (\lambda_0 + \lambda_1 + \lambda_2)},
\]
\[
E \left( -\frac{\partial^2 \log L}{\partial \lambda_0 \partial \lambda_1} \right) = \frac{\lambda_2}{(\lambda_0 + \lambda_1)^2 (\lambda_0 + \lambda_1 + \lambda_2)},
\]
\[
E \left( -\frac{\partial^2 \log L}{\partial \lambda_0 \partial \lambda_2} \right) = \frac{\lambda_1}{(\lambda_0 + \lambda_2)^2 (\lambda_0 + \lambda_1 + \lambda_2)},
\]
\[
E \left( -\frac{\partial^2 \log L}{\partial \lambda_1^2} \right) = \frac{\lambda_2}{(\lambda_0 + \lambda_1)^2 (\lambda_0 + \lambda_1 + \lambda_2)} + \frac{1}{\lambda_1 (\lambda_0 + \lambda_1 + \lambda_2)},
\]
and
\[
E \left( -\frac{\partial^2 \log L}{\partial \lambda_2^2} \right) = \frac{1}{\lambda_2 (\lambda_0 + \lambda_1 + \lambda_2)} + \frac{\lambda_1}{(\lambda_0 + \lambda_2)^2 (\lambda_0 + \lambda_1 + \lambda_2)}.
\]
5 INFORMATION MATRIX FOR DE GROOTS’s BIVARIATE PARETO

If \((x, y)\) is a single observation from (5) then the log-likelihood function can be written as

\[
\log L(\gamma, \xi, \eta) = \log \{\gamma (\gamma + 1)(\xi - \eta)\} - (\gamma + 2) \log(y - x).
\]

The first-order derivatives are:

\[
\frac{\partial \log L}{\partial \gamma} = \frac{1}{\gamma} + \frac{1}{\gamma + 1} + \log(\xi - \eta) - \log(y - x),
\]

\[
\frac{\partial \log L}{\partial \xi} = \frac{\gamma}{\xi - \eta},
\]

\[
\frac{\partial \log L}{\partial \eta} = -\frac{\gamma}{\xi - \eta},
\]

\[
\frac{\partial^2 \log L}{\partial \gamma^2} = -\frac{1}{\gamma^2} - \frac{1}{(\gamma + 1)^2},
\]

\[
\frac{\partial^2 \log L}{\partial \gamma \partial \xi} = \frac{1}{\xi - \eta},
\]

\[
\frac{\partial^2 \log L}{\partial \gamma \partial \eta} = -\frac{1}{\xi - \eta},
\]

\[
\frac{\partial^2 \log L}{\partial \xi^2} = -\frac{\gamma}{(\xi - \eta)^2},
\]

\[
\frac{\partial^2 \log L}{\partial \xi \partial \eta} = \frac{\gamma}{(\xi - \eta)^2},
\]

\[
\frac{\partial^2 \log L}{\partial \eta^2} = -\frac{\gamma}{(\xi - \eta)^2}.
\]

Since all of the second-order derivatives above are constants expressions for the elements of the Fisher information matrix are straight-forward.
6 APPENDIX

We need the following technical lemmas to calculate the elements of the Fisher information matrix.

**LEMMA 1** If $X$ and $Y$ are jointly distributed according to (2) then

\[
E(X^m Y^n) = \frac{mnB(n, a-n)B(m, a-m-n)}{\theta^m \phi^n}
\]

for $m \geq 1$ and $n \geq 1$.

**PROOF:** Using the formula (6), one can express

\[
E(X^m Y^n) = mn \int_0^\infty \int_0^\infty x^{m-1} y^{n-1} (1 + \theta x + \phi y)^{-a} \, dy \, dx
\]

\[
= mn \int_0^\infty x^{m-1} \phi^{-a} \int_0^\infty y^{n-1} \left( y + \frac{1 + \theta x}{\phi} \right)^{-a} \, dy \, dx
\]

\[
= mn \int_0^\infty x^{m-1} \phi^{-a} \left( \frac{1 + \theta x}{\phi} \right)^{n-a} B(n, a-n) \, dx
\]

\[
= mnB(n, a-n)\theta^{n-a} \phi^{-n} \int_0^\infty x^{m-1} \left( x + \frac{1}{\theta} \right)^{n-a} \, dx
\]

\[
= mnB(n, a-n)\theta^{n-a} \phi^{-n} \left( \frac{1}{\theta} \right)^{m-a+n} B(m, a-m-n).
\]

The result of the lemma follows. ▲

**LEMMA 2** If $X$ and $Y$ are jointly distributed according to (3) then

\[
E(X^m Y^n) = \frac{\Gamma(m + a)\Gamma(n + a)\Gamma(c - m - n)}{b^{m+n}\Gamma^2(a)\Gamma(c)}
\]

for $m \geq 1$ and $n \geq 1$.

**PROOF:** Using the formula (7), one can express

\[
E(X^m Y^n) = \frac{\Gamma(2a + c)}{b^{m+n}\Gamma(c)\Gamma^2(a)} \int_0^\infty \int_0^\infty (bx)^{m+a-1}(by)^{n+a-1} \, d(by)d(bx)
\]

\[
= \frac{\Gamma(2a + c)}{b^{m+n}\Gamma(c)\Gamma^2(a)} \int_0^\infty \int_0^\infty u^{m+a-1}v^{n+a-1} \, dv \, du
\]

\[
= \frac{\Gamma(2a + c)}{b^{m+n}\Gamma(c)\Gamma^2(a)} \int_0^\infty u^{m+a-1} \int_0^\infty \frac{v^{n+a-1}}{(1 + u + v)^{2a+c}} \, dv \, du
\]

\[
= \frac{\Gamma(2a + c)}{b^{m+n}\Gamma(c)\Gamma^2(a)} \int_0^\infty u^{m+a-1} \int_0^\infty \frac{v^{n+a-1}}{(1 + u + v)^{2a+c}} \, dv \, du
\]

\[
= \frac{\Gamma(2a + c)}{b^{m+n}\Gamma(c)\Gamma^2(a)} B(a + c - n, n + a) \, du
\]

\[
= \frac{\Gamma(2a + c)}{b^{m+n}\Gamma(c)\Gamma^2(a)} B(a + m, c - m - n)B(a + c - n, n + a).
\]

The result of the lemma follows. ▲

**REFERENCES**

