Joarder, A.H. (2007). Some useful integrals and their applications in correlation analysis. *Statistical Papers, 49, 211-214* (52SPw)

Department of Mathematical Sciences
King Fahd University of Petroleum and Minerals
Dhahran 31261 Saudi Arabia
Email: anwarj@kfupm.edu.sa

The distribution of the product moment correlation coefficient based on the bivariate normal distribution is well known. Recently in many business and economic data, fat tailed distributions especially some elliptical distributions have been considered as parent populations. The normal and *t*-distributions are well known special cases of elliptical distribution. In this paper we derive some theorems involving double integrals and apply them to derive the probability distribution of the correlation coefficient for some elliptical populations. The general nature of the theorems indicates their potential use in probability distribution theory.

AMS Mathematics Subject Classification: 60-01, 60E05, 62E15, 62F03

Key Words and Phrases: Correlation coefficient, double integrals, elliptical distribution, bivariate *t*-distribution

1. Introduction

The distribution of the product moment correlation coefficient based on the bivariate normal distribution was derived by Fisher (1915). A recent interest among the applied scientists is the use of fat tailed distribution for modeling business data especially stock returns. Since the bivariate *t*-distribution has fatter tails, it has been increasingly applied for modeling business data especially stock returns.

The distribution is said to be robust if it remains the same under violation of normality. The robustness of the distribution of the correlation for elliptical population was proved by Fang and Anderson (1990, 10) by stochastic representation. Fang (1990) derived the null distribution whereas Ali and Joarder (1991) derived the nonnull distribution of the correlation coefficient for bivariate elliptical distribution. It should be pointed out that in the case of bivariate elliptical distribution, the observations in the sample are not necessarily independent.

In this paper we derive some theorems containing bivariate integrals that help derive the distribution of the correlation for different populations. The general nature of the theorems indicate their potential use for many other applications in the distribution theory of bivariate elliptical distribution. The objective is to provide insight to those experts in business, science and engineering who use elliptical models as models for samples. See e.g. Sutradhar and Ali (1986), Lange, Little and Taylor (1989), Sutradhar and Ali (1989), Fang (1990), Fang and Anderson (1990), Fang, Kotz and Ng (1990), Joarder and Ahmed (1998), Kibria and Haq (1999), Billah and Saleh (2000), Kibria (2003), Kibria(2004), Kibria and Saleh (2004), and Kotz and Nadarajah (2004).

2. The Bivariate Normal, T and Elliptical Distributions

The bivariate elliptical distribution which includes bivariate normal distribution as *t*-distribution is outlined in this section.

(i) The Bivariate Normal Distribution

Let $X = (X_1, X_2)'$ be bivariate normal random vector with probability density function (pdf)

$$f_1(x) = (2\pi)^{-1} |\Sigma|^{-1/2} \exp\left[\frac{-1}{2}\left((x-\theta)'\Sigma^{-1}(x-\theta)\right)\right] \qquad (2.1)$$

where $\theta = (\theta_1, \theta_2)'$ is unknown vector of location parameters and Σ is the 2×2 unknown positive definite matrix of population variances and covariance. The probability density function of the bivariate normal distribution will be denoted by $N_2(\theta, \Sigma)$. Now consider a sample $X_1, X_2, \cdots X_N \ (N > 2)$ having the joint probability density function

$$f_2(x_1, x_2, \cdots, x_N) = \frac{|\Sigma|^{-N/2}}{(2\pi)^{-N}} \exp\left(\frac{-1}{2}\sum_{j=1}^{N}(x_j-\theta)'\Sigma^{-1}(x_j-\theta)\right). \qquad (2.2)$$

The mean vector is $\bar{X} = (\bar{X}_1, \bar{X}_2)'$ so that the sums of squares and cross products matrix is given by $\sum_{j=1}^{N}(X_j - \bar{X})(X_j - \bar{X})' = A$. The symmetric bivariate matrix A can be written as $A = \left(a_{ik}\right), i = 1,2; k = 1,2$ where $a_{ii} = mS_i^2 = \sum_{j=1}^{N}(X_{ij} - \bar{X}_i)^2, \quad m = N - 1, (i = 1,2)$ and $a_{12} = \sum_{j=1}^{N}(X_{1j} - \bar{X}_1)(X_{2j} - \bar{X}_2) = mRS_1S_2$. Fisher (1915) derived the distribution of A for $p = 2$ in order to study the distribution of correlation coefficient from a normal sample. Wishart (1928) obtained the joint distribution of sample variances and covariances from the multivariate normal population.

The distribution of the bivariate Wishart matrix based on the bivariate normal distribution is given by

$$f_3(A) = \frac{2^{-m} \,|\Sigma|^{-m/2}}{\sqrt{\pi}\Gamma(m/2)\Gamma((m-1)/2)} |A|^{(m-3)/2} \exp\left(-\frac{1}{2} tr\Sigma^{-1}A\right), \qquad (2.3)$$

$A > 0, m > p$ (See e.g. Anderson, 2003, 252).

(ii) The Bivariate T-Distribution

Let $X = (X_1, X_2)'$ be a bivariate t-random vector with probability density function

$$f_4(x) = (2\pi)^{-1} |\Sigma|^{-1/2} \left(1 + (x - \theta)'(\nu\Sigma)^{-1}(x - \theta)\right)^{-\nu/2-1} \qquad (2.4)$$

where $\theta = (\theta_1, \theta_2)'$ is unknown vector of location parameters and Σ is the 2×2 unknown positive definite matrix of scale parameters while the scalar ν is assumed to be a known positive constant (Muirhead, 1982, 48). Notice that though the components X_1 and X_2 are uncorrelated, they are not independent unless $\nu \to \infty$.

Now consider a sample $X_1, X_2, \cdots X_N \ (N > 2)$ having the joint probability density function

$$f_5(x_1, x_2, \cdots, x_N) = \frac{\Gamma(\nu/2 + N)\,|\Sigma|^{-N/2}}{(\nu\pi)^N\,\Gamma(\nu/2)} \left(1 + \sum_{j=1}^{N}(x_j - \theta)'(\nu\Sigma)^{-1}(x_j - \theta)\right)^{-\nu/2-1}. \qquad (2.5)$$

which is the bivariate *t*-model for the sample. Note that the observations in the sample are uncorrelated and not independent unless $\nu \to \infty$. The random symmetric positive definite matrix A is said to have a Wishart distribution based on the bivariate t-population with $m = N - 1 > 2$ and $\Sigma(2 \times 2) > 0$, written as $A \sim W(m, \Sigma; \nu)$ if its probability density function is given by

$$f_6(A) = C_\nu(m,2)\,|\Sigma|^{-m/2} \quad |A|^{(m-3)/2}\,(1 + tr(\nu\Sigma)^{-1}A)^{-\nu/2-m}, \tag{2.6}$$

$A > 0, m > 2$ where

$$C_\nu(m,2) = \frac{\nu^{-m}\ \Gamma(\nu/2 + m)}{\sqrt{\pi}\ \Gamma(\nu/2)\Gamma(m/2)\Gamma\left((m-1)/2\right)}$$

(See Sutradhar and Ali, 1989, 160).

By the use of the duplication formula for gamma function given by

$$\Gamma(z) = \pi^{-1/2}\ 2^{z-1}\ \Gamma((z+1)/2)\ \Gamma(z/2) \tag{2.7}$$

(Anderson, 2003, 125) with $z = m - 1$ we have

$$C_\nu(m,2) = \frac{2^{m-2}\ \nu^{-m}\ \Gamma(m + \nu/2)}{\pi\ \Gamma(m-1)\Gamma(\nu/2)}. \tag{2.8}$$

(iii) The Bivariate Elliptical Distribution

The probability density function for the bivariate elliptical distribution is given by

$$f_7(x) = K(N,2)\,|\Sigma|^{-1/2}\ g_{N,2}((x-\theta)'\Sigma^{-1}(x-\theta)) \tag{2.9}$$

where $\theta = (\theta_1, \theta_2)'$ is unknown vector of location parameters and Σ is the 2×2 unknown positive definite matrix of scale parameters while the normalizing constant $K(N,2)$ is determined by the form of g (cf. Sutradhar and Ali, 1989).

Now consider a sample $X_1, X_2, \cdots X_N$ $(N > 2)$ having the joint probability density function

$$f_8(x_1, x_2, \cdots, x_N) = \frac{K(N,2)}{|\Sigma|^{N/2}}\, g_{N,2}\left(\sum_{j=1}^{N}(x_j - \theta)'\Sigma^{-1}(x_j - \theta)\right) \tag{2.10}$$

which is the bivariate elliptical model.

Theorem 2.1 (Sutradhar and Ali, 1989, 158) Consider the pdf of the bivariate Wishart matrix based on the bivariate elliptical model given by (2.10). Then the pdf of the Wishart matrix is given by

$$f_9(A) = C(m,2)\,|\Sigma|^{-m/2}|A|^{(m-3)/2}\; g_{m,2}(tr\Sigma^{-1}A),$$
$$C(m,2) = 2^{m-2}\pi^{m-1}\; K(m,2)/\Gamma(m-1), \quad m>2 \quad . \qquad (2.11)$$

3. Main Results

Lemma 3.1 Let $V = \sqrt{U_1 U_2}$ be the geometric mean of two independent chisquare random variables $U_i \sim \chi^2_m$ $(i=1,2)$. Then the moment generating function of V at ρ is given by

$$M_V(\rho) = \sum_{k=0}^{\infty} \frac{(2\rho)^k}{k!}\frac{\Gamma^2((m+k)/2)}{\Gamma^2(m/2)}, \quad -1<\rho<1.$$

Proof. By definition, the moment generating function of $V = \sqrt{U_1 U_2}$ at ρ is given by

$$E\left(e^{\rho\sqrt{U_1 U_2}}\right) = \int_0^\infty\int_0^\infty e^{\rho\sqrt{u_1 u_2}}\,\frac{1}{2^m\,\Gamma^2(m/2)}\left(u_1 u_2\right)^{m/2-1} e^{-(u_1+u_2)/2}du_1 du_2.$$

Then the lemma is obvious by virtue of

$$e^{\rho\sqrt{u_1 u_2}} = \sum_{k=0}^{\infty}\frac{\rho^k}{k!}(u_1 u_2)^{k/2} \text{ and } E\left(U_i^{k/2}\right) = \frac{2^k\;\Gamma(m+k/2)}{\Gamma(m/2)},\ (i=1,2).$$

Lemma 3.2 Let $I(\rho,m) = \int_0^\pi (\sin\theta)^{m-1}(1-\rho\sin\theta)^{-m}d\theta, \quad -1<\rho<1.$ Then

$$(i)\; I(\rho,m) = \frac{2^{m-1}}{\Gamma(m)}\sum_{k=0}^{\infty}\frac{(2\rho)^k}{k!}\;\Gamma^2\left((m+k)/2\right),$$

$$(ii)\; I(\rho,m) = \frac{2^{m-1}\;\Gamma^2(m/2)}{\Gamma(m)}\;M_V(\rho).$$

Proof. Since $|\rho\sin\theta|<1$, we have

$$(1-\rho\sin\theta)^{-m} = \sum_{k=0}^{\infty}\frac{\rho^k}{k!}\frac{\Gamma(m+k)}{\Gamma(m)}\;(\sin\theta)^k \text{ so that}$$

$$I(\rho,m)=\sum_{k=0}^{\infty}\frac{\rho^k}{k!}\frac{\Gamma(m+k)}{\Gamma(m)}\int_0^{\pi}(\sin\theta)^{m+k-1}d\theta$$
$$=\sum_{k=0}^{\infty}\frac{\rho^k}{k!}\frac{\Gamma(m+k)}{\Gamma(m)}\frac{\sqrt{\pi}\ \Gamma((m+k)/2)}{\Gamma((m+k+1)/2)}$$

by virtue of

$$\int_0^{\pi}(\sin\theta)^m d\theta=\frac{\sqrt{\pi}\Gamma((m+1)/2)}{\Gamma(m/2+1)}.$$

Next, replacing $\Gamma(m+k)$ by the duplication formula of gamma function, given by (2.7), with $z=m+k$, we have (i), which can also be written as (ii) by virtue of Lemma 3.1.

Theorem 3.1 For $-1<\rho<1$ and $\nu>0$, let

$J(\rho,m,\nu)=\int_0^{\infty}\int_0^{\infty}(u_1u_2)^{m/2-1}\left(1+u_1+u_2-2\rho\sqrt{u_1u_2}\right)^{-\nu-m}du_1du_2$. Then

$$(i)\ J(\rho,m,\nu)=\frac{\Gamma(\nu)}{\Gamma(m+\nu)}\sum_{k=0}^{\infty}\frac{(2\rho)^k}{k!}\Gamma^2\left(\frac{m+k}{2}\right),$$

$$(ii)\ J(\rho,m,\nu)=\frac{\Gamma(\nu)\Gamma^2(m/2)}{\Gamma(m+\nu)}\ M_V(\rho).$$

Proof. The integral in the theorem can be written as

$$J(\rho,m,\nu)=4\int_0^{\infty}\int_0^{\infty}(y_1y_2)^{m-1}\left(1+y_1^2+y_2^2-2\rho y_1y_2\right)^{-\nu-m}dy_1dy_2. \qquad (3.1)$$

The transformation $y_1=w\cos\theta,\ y_2=w\sin\theta$ with Jacobian $J(y_1,y_2\to w,\theta)=w$ yields

$$J(\rho,m,\nu)=2^{-m+3}\int_{\theta=0}^{\pi/2}\int_{w=0}^{\infty}\left(\sin 2\theta\right)^{m-1}w^{2m-1}\left(1+w^2-\rho w^2\sin 2\theta\right)^{-\nu-m}\ dwd\theta.$$

Next, the transformations $w^2=u,\ 2\theta=\alpha$ yield

$$J(\rho, m, \nu) = \int_{\alpha=0}^{\pi} \left(\frac{1}{2}\sin\alpha\right)^{m-1} \int_{u=0}^{\infty} u^{m-1}\left[1+(1-\rho\sin\alpha)u\right]^{-\nu-m} du d\alpha. \quad (3.2)$$

Then by virtue of

$$\int_0^{\infty} \frac{x^{m-1}}{(a+bx)^{m+n}} dx = \frac{\Gamma(m)\Gamma(n)}{a^n b^m \Gamma(m+n)} ,$$

the last integral of (3.2) becomes

$$\frac{\Gamma(m)\Gamma(\nu/2)}{\Gamma(m+\nu/2)} (1-\rho\sin\theta)^{-m}$$

so that $J(\rho, m, \nu) = \frac{1}{2^{m-1}} \frac{\Gamma(m)\Gamma(\nu)}{\Gamma(m+\nu)} \int_{\alpha=0}^{\pi} (\sin\alpha)^{m-1} (1-\rho\sin\alpha)^{-m} d\alpha.$

Then the theorem follows by Lemma 3.2.

Lemma 3.3 For a bivariate elliptical probability model given by (2.10), we have $\xi(N) = \int_0^{\infty} w^{N-1} g_{N,2}(w) dw = \pi^{-N} \Gamma(N) K^{-1}(N,2)$ (cf. Fang, Kotz and Ng, 1990, 66).

Proof. Make the transformation $\Sigma^{-1/2}(x_j - \theta) = z_j \quad (j = 1, 2, \cdots, N)$. Then the probability density function of $Z_1, Z_2, \cdots, Z_N$ is given by

$$f_{10}(z_1, z_2, \cdots, z_N) = K(N,2)\, g_{N,2}\left(\sum_{j=1}^{N} z_j' z_j\right). \quad (3.3)$$

The pdf of $z_{11} = u_1, z_{12} = u_2, \cdots, z_{N1} = u_{2N-1},\ z_{N1} = u_{2N}$ is given by

$$f_{10}(u_1, u_2, \cdots, u_{2N}) = K(N,2)\, g_{N,2}\left(\sum_{j=1}^{N} u_j^2\right)$$

and hence $\int_0^{\infty} \cdots \int_0^{\infty} K(N,2)\, g_{N,2}\left(\sum_{j=1}^{2N} u_j^2\right) du_1 \cdots du_{2N} = 1.$

Make the polar transformation

$$u_j = w\left(\prod_{k=1}^{j-1}\sin\theta_k\right)\cos\theta_j \ , \ \ (j = 1,2,\cdots,2N-1)$$

$$u_{2N} = w\prod_{k=1}^{2N-1}\sin\theta_k$$

where $w \in [0,\infty)$, $\theta_k \in [0,\pi)$ for $k = 1,2,\cdots,2n-2$; $\theta_{2N-1} \in [0,2\pi)$ with Jacobian

$$J(u_1,u_2,\cdots,u_{2N} \to \theta_1,\cdots\theta_{2N-1},w) = w^{2N-1}\prod_{k=1}^{2N-2}(\sin\theta_k)^{2N-k-1} \ .$$

Then

$$K(N,2)\int_0^{\pi}\cdots\int_{\theta_{2N-2}=0}^{\pi}\int_{\theta_{2N-1}=0}^{2\pi}\int_{w=0}^{\infty} w^{2N-1}\prod_{k=1}^{2N-2}(\sin\theta_k)^{2N-k-1}$$

$$\times g_{N,2}\left(w^2\right)d\theta_1\cdots d\theta_{2N-2}d\theta_{2N-1}\ dw = 1$$

or, $K(N,2)\int_0^{\infty} w^{N-1}\ g_{N,2}(w)dw = \pi^{-N}\ \Gamma(N)$.

Theorem 3.2 Let

$J_g(\rho,m) = \int_0^{\infty}\int_0^{\infty}(u_1u_2)^{m/2-1}g_{m,2}\left(u_1+u_2-2\rho\sqrt{u_1u_2}\right)du_1du_2$. Then

$$(i)\ J_g(\rho,m) = \frac{\xi(m)}{\Gamma(m)}\sum_{k=0}^{\infty}\frac{(2\rho)^k}{k\,!}\Gamma^2\left((m+k)/2\right),$$

$$(ii)\ J_g(\rho,m) = \frac{\xi(m)\Gamma^2(m/2)}{\Gamma(m)}\ M_V(\rho)$$

where $\xi(m) = \pi^{-m}\Gamma(m)K^{-1}(m,2) = 2^{m-2}(m-1)C^{-1}(m,2)/\pi$.

Proof. The integral in the theorem can be written as

$$J_g(\rho,m) = 4\int_0^{\infty}\int_0^{\infty}(y_1y_2)^{m/2-1}g_{m,2}\left(y_1^2+y_2^2-2\rho\sqrt{y_1y_2}\right)dy_1dy_2. \qquad (3.4)$$

The transformation $y_1 = w\cos\theta,\ y_2 = w\sin\theta$ with Jacobian $J(y_1,y_2 \to w,\theta) = w$ yields

$$J_g(\rho,m)$$
$$= 2^{-m+3}\int_{\theta=0}^{\pi/2}\int_{w=0}^{\infty}\left(\sin 2\theta\right)^{m-1} w^{2m-1} g_{m,2}\left(w^2-\rho w^2\sin 2\theta\right)^{-\nu-m} dw\,d\theta.$$

Next, letting $\int_{w=0}^{\infty} w^{m-1} g_{m,2}(w)dv = \xi(m)$, the independent transformations $w^2 = u,\ 2\theta=\alpha$ yield

$$J_g(\rho,m) = 2^{-m+1}\int_{\alpha=0}^{\pi}(\sin\alpha)^{m-1}\int_{u=0}^{\infty} u^{m-1} g_{m,2}\left[(1-\rho\sin\alpha)u\right]du\,d\alpha$$
$$= 2^{-m+1}\left[\int_{\alpha=0}^{\pi}\frac{(\sin\alpha)^{m-1}}{(1-\rho\sin\alpha)^m}d\alpha\right]\left[\int_{w=0}^{\infty} w^{m-1} g_{m,2}(w)dw\right]$$
$$= 2^{-m+1}\left[\frac{2^{m-1}\,\Gamma^2(m/2)}{\Gamma(m)}M_V(\rho)\right]\xi(m).$$

That is $J_g(\rho,m) = \dfrac{\xi(m)\Gamma^2(m/2)}{\Gamma(m)}M_V(\rho).$ Then the theorem follows by (2.11) and Lemma 3.3 in the following way:

$$\xi(m) = \int_0^{\infty} u^{m-1} g_{m,2}(u)du$$
$$= \pi^{-m}\Gamma(m)K^{-1}(m,2)$$
$$= 2^{m-2}(m-1)C^{-1}(m,2)/\pi.$$

4. Applications in Correlation Analysis

The long proof of the distribution of correlation coefficient by Fisher (1915) has been made shorter and elegant in this section. Further the distribution of correlation coefficient has been derived, along Fisher (1915), for bivariate *t*-distribution as well as bivariate elliptical distribution.

Theorem 4.1 The probability density function of the correlation coefficient R based on a bivariate normal population, *t*-population or elliptical population is given by

$$h(r) = \frac{2^{m-2}\ \Gamma^2(m/2)(1-\rho^2)^{m/2}}{\pi\ \Gamma(m-1)}\ \ (1-r^2)^{(m-3)/2} M_V\ (\rho r)$$

$$= \frac{2^{m-2}(1-\rho^2)^{m/2}}{\pi\ \Gamma(m-1)}\ (1-r^2)^{(m-3)/2} \sum_{k=0}^{\infty} \frac{(2\rho r)^k}{k\ !}\ \Gamma^2\left(\frac{m+k}{2}\right),\ \ -1 < r < 1$$

where $m > 2, -1 < \rho < 1$ and $M_V\ (\rho)$ is defined in Lemma 3.1 (cf. Johnson, Kotz and Balakrishanan, 1995, 548).

Proof. (i) Bivariate Normal Distribution Case

The probability density function of the elements of A given by (2.3) can be written as

$$f_3(a_{11}, a_{22}, a_{12}) = \frac{(1-\rho^2)^{-m/2}(\sigma_1\sigma_2)^{-m}}{2^m\sqrt{\pi}\Gamma(m/2)\Gamma((m-1)/2)}\ (a_{11}a_{22} - a_{12}^2)^{(m-3)/2}$$
$$\times \exp\left(-\frac{1}{2(1-\rho^2)}\left(\frac{a_{11}}{\sigma_1^2} + \frac{a_{22}}{\sigma_2^2} - \frac{2a_{12}}{\sigma_1\sigma_2}\right)\right) \tag{4.1}$$

where $a_{11} > 0, a_{22} > 0, -\infty < a_{12} < \infty, -1 < \rho < 1$, $m > 2, \sigma_1 > 0, \sigma_2 > 0$. Under the transformation $a_{11} = ms_1^2, a_{22} = ms_2^2, a_{12} = mrs_1s_2$ with Jacobian $m^3 s_1 s_2$, followed by the transformation $m s_1{}^2 = \sigma_1{}^2 u_1,\ m s_2{}^2 = \sigma_2{}^2 u_2$, with Jacobian $J(s_1{}^2, s_2{}^2 \to u_1 u_2) = (\sigma_1\sigma_2/m)^2$ and the integration over u_1 and u_2, the probability density function of R will be

$$h(r) = \frac{(1-r^2)^{(m-3)/2}}{4\pi\ \Gamma(m-1)(1-\rho^2)^{m/2}}$$
$$\times \int_0^\infty\int_0^\infty (u_1 u_2)^{m/2-1} \exp\left[-\frac{u_1 + u_2 - 2\rho r\sqrt{u_1 u_2}}{2\left(1-\rho^2\right)}\right] du_1 du_2. \tag{4.2}$$

Then the transformation $u_1 = (1-\rho^2)y_1,\ u_2 = (1-\rho^2)y_2$ with Jacobian $J(u_1, u_2 \to y_1, y_2) = (1-\rho^2)^2$ yields

$$h(r) = \frac{(1-\rho^2)^{m/2}}{4\pi\ \Gamma(m-1)} \ (1-r^2)^{(m-3)/2}$$

$$\times \int_0^\infty \int_0^\infty e^{\rho r \sqrt{y_1 y_2}} (y_1 y_2)^{m/2-1} e^{-(y_1+y_2)/2} dy_1 dy_2 \qquad (4.3)$$

The theorem is thus complete by Lemma 3.1.

Corollary 4.1 For $-1 < \rho, r < 1$, we have

$$\int_0^\infty \int_0^\infty e^{\rho r \sqrt{y_1 y_2}} (y_1 y_2)^{m/2-1} e^{-(y_1+y_2)/2} dy_1 dy_2 = \frac{4\pi\ \Gamma(m-1)}{(1-\rho^2)^{m/2}} (1-r^2)^{-(m-3)/2}.$$

(ii) Bivariate T-Distribution Case

Proof. The pdf of the elements of A given by (2.6) can be written as

$$f_6(a_{11}, a_{22}, a_{12}) = \frac{C_\nu(m,2)}{(1-\rho^2)^{m/2}(\sigma_1\sigma_2)^m} \ (a_{11}a_{22} - a_{12}^2)^{(m-3)/2}$$

$$\times \left(1 + \frac{1}{\nu(1-\rho^2)}\left(\frac{a_{11}}{\sigma_1^2} + \frac{a_{22}}{\sigma_2^2} - \frac{2\rho a_{12}}{\sigma_1\sigma_2}\right)\right)^{-\nu/2-m} \qquad (4.4)$$

where $a_{11} > 0, a_{22} > 0, -\infty < a_{12} < \infty, -1 < \rho < 1$, $m > 2, \sigma_1 > 0, \sigma_2 > 0$.
Under the transformation $a_{11} = ms_1^2, a_{22} = ms_2^2, a_{12} = mrs_1s_2$ with Jacobian $J(a_{11}, a_{22}, a_{12} \to r, s_1^2, s_2^2) = m^3 s_1 s_2$, followed by the transformation $ms_1^2 = \sigma_1^2 u_1$, $ms_2^2 = \sigma_2^2 u_2$ with Jacobian $J(s_1^2, s_2^2 \to u_1, u_2) = \left(\sigma_1\sigma_2 / m\right)^2$ and then integrating out u_1 and u_2 we have the probability density function of R as follows:

$$h(r) = \frac{C_\nu(m,2)(1-r^2)^{(m-3)/2}}{(1-\rho^2)^{m/2}}$$

$$\times \int_0^\infty \int_0^\infty (u_1 u_2)^{m/2-1} \left[1 + \frac{u_1 + u_2 - 2\rho r\sqrt{u_1 u_2}}{\nu(1-\rho^2)}\right]^{-\nu/2-m} du_1 du_2. \qquad (4.5)$$

Then the transformation $u_1 = (1-\rho^2) y_1$, $u_2 = (1-\rho^2) y_2$ with Jacobian $J(u_1, u_2 \to y_1, y_2) = (1-\rho^2)^2$ yields

$$h(r) = \ \nu^m \ C(m,\nu) \ (1-\rho^2)^{m/2} (1-r^2)^{(m-3)/2} J(\rho r, m, \nu/2) \qquad (4.6)$$

where $J(\rho, m, \nu)$ is defined in Theorem 3.1 and the Theorem 4.1 follows.

Corollary 4.1 For $-1 < \rho, r < 1$, we have

$$\int_0^\infty\int_0^\infty (u_1u_2)^{m/2-1}\left[1+\frac{u_1+u_2-2\rho r\sqrt{u_1u_2}}{\nu(1-\rho^2)}\right]^{-\nu/2-m} du_1du_2$$

$$=\frac{C_\nu^{-1}(m,2)}{(1-\rho^2)^{m/2}}\ (1-r^2)^{-(m-3)/2}.$$

(iii) Bivariate Elliptical Distribution Case

We now demonstrate how Theorem 3.2 eases the derivation of the distribution of the correlation coefficient (Ali and Joarder, 1991). The general nature of Theorem 3.2 indicates its potential application in the sampling distribution theory of elliptical population.

Proof. The pdf of the elements of A given by Theorem 2.1 can be written as

$$f_9(a_{11},a_{22},a_{12}) = \frac{C(m,2)(a_{11}a_{22}-a_{12}^2)^{(m-3)/2}}{(1-\rho^2)^{m/2}(\sigma_1\sigma_2)^m} \times g_{m,2}\left[\frac{1}{1-\rho^2}\left(\frac{a_{11}}{\sigma_1^2}+\frac{a_{22}}{\sigma_2^2}-\frac{2\rho a_{12}}{\sigma_1\sigma_2}\right)\right] \quad (4.7)$$

where $a_{11} > 0, a_{22} > 0, -\infty < a_{12} < \infty, -1 < \rho < 1$, $m > 2, \sigma_1 > 0, \sigma_2 > 0$.
Under the transformation $a_{11} = ms_1^2, a_{22} = ms_2^2, a_{12} = mrs_1s_2$ with Jacobian $J(a_{11},a_{22},a_{12} \to r, s_1^2, s_2^2) = m^3s_1s_2$, followed by the transformation $ms_1^2 = \sigma_1^2u_1$, $ms_2^2 = u_2\sigma_2^2$ with Jacobian $J(s_1^2,s_2^2 \to u_1,u_2) = (\sigma_1\sigma_2/m)^2$ and the integrating out u_1 and u_2, we have the probability density function of R given by

$$h(r) = C(m,2)\ (1-\rho^2)^{-m/2}(1-r^2)^{(m-3)/2} \times\int_0^\infty\int_0^\infty (u_1u_2)^{m/2-1}\ g_{m,2}\left(\frac{u_1+u_2-2\rho r\sqrt{u_1u_2}}{1-\rho^2}\right)du_1du_2. \quad (4.8)$$

Then the transformation $u_1 = (1-\rho^2)y_1$, $u_2 = (1-\rho^2)y_2$ with Jacobian $J(u_1,u_2 \to y_1,y_2) = (1-\rho^2)^2$ yields

$$h(r) = C(m,2)\ (1-\rho^2)^{m/2}(1-r^2)^{(m-3)/2}J_g(\rho r, m) \quad (4.9)$$

where $J_g(\rho, m)$ is defined in Theorem 3.2 and then the Theorem 4.1 follows.

Corollary 4.1 For $-1 < \rho, r < 1$, we have

$$\int_0^\infty \int_0^\infty (u_1 u_2)^{m/2-1} g_{m,2}(u_1 + u_2 - 2\rho r \sqrt{u_1 u_2}) du_1 du_2 = \frac{(1-r^2)^{-(m-3)/2}}{C(m,2)(1-\rho^2)^{m/2}}$$

where $g_{m,2}(.)$ is defined in Theorem 2.1.

5. Robustness of Some Tests on Correlation Coefficient

The results in Section 4 indicate robustness of the correlation coefficient in the bivariate elliptical population only. Thus the assumption of bivariate normality under which tests on correlation coefficient are developed can be relaxed to a broader class of bivariate elliptical distribution. The likelihood ratio test of the hypothesis $H_0 : \rho = 0$ against all alternatives $H_1 : \rho \neq 0$ is performed by $T = \sqrt{m-1}\, R(1-R^2)^{-1/2}$ which has a Student *t*-distribution with $(m-1) > 0$ degrees of freedom (d.f.). One significant lesson of the paper is that the acceptance of the null hypothesis does not mean independence unless the sample is from the bivariate normal distribution. The most popular test is based on $Z = \tanh^{-1} R = \ln\sqrt{(1+R)/(1-R)}$ has an approximate normal distribution with mean $\ln\sqrt{(1+\rho)/(1-\rho)}$ and variance $1/(m-2)$.

Muddapur (1988) proved that the statistic T has an exact *t*-distribution with $m-1$ degrees of freedom where

$$T = \frac{(\upsilon R - \rho S^*)\sqrt{m-1}}{\zeta\sqrt{(1-\rho^2)\left(1-R^2\right)}},$$

$$S^*\sqrt{a_{11}a_{22}} = \sigma_1^2 a_{22} + \sigma_2^2 a_{11} + (a_{22} + a_{11})\sigma_1^2\sigma_2^2\sqrt{1-\rho^2}, \tag{5.1}$$

$$\upsilon = 2\sigma_1\sigma_2 + (\sigma_1^2 + \sigma_2^2)\sqrt{(1-\rho^2)}, \zeta = 2\sigma_1\sigma_2\sqrt{1-\rho^2} + (\sigma_1^2 + \sigma_2^2) .$$

In particular, if the population variances are same $\sigma_1{}^2 = \sigma_2{}^2$, then the *t* statistic defined as

$$T = \frac{(R - \rho S^*)\sqrt{m-1}}{\sqrt{(1-\rho^2)(1-R^2)}}, \quad S^* = \frac{a_{11} + a_{22}}{2\sqrt{a_{11}a_{22}}} \tag{5.2}$$

has an exact t distribution with $m-1$ d.f. If the population variances are the same $\sigma_1^2 = \sigma_2^2$ and sample variances are the same i.e. $s_1^2 = s_2^2$, the above statistic simplifies to

$$T = \frac{(R - \rho)\sqrt{m-1}}{\sqrt{(1-\rho^2)(1-R^2)}} \tag{5.3}$$

which has an exact t –distribution with $m-1$ d.f. The above statistic was shown to have an approximate t distribution without the assumptions of $\sigma_1^2 = \sigma_2^2$ or of $s_1^2 = s_2^2$ by Samiuddin (1970). Muddapur (1988) also noted that the quantity

$$f = \frac{(1+r)(1-\rho)}{(1-r)(1+\rho)} \tag{5.4}$$

has an approximate F distribution with $(m-1, m-1))$ degrees of freedom for any ρ, and an exact F -distribution for $\rho = 0$.

6. Concluding Remarks

We warn that the distribution of R is not necessarily robust for independent observations from elliptical population. The models for samples considered in Section 2 imply that the observations X_j $(j = 1, 2, \cdots, N)$ are uncorrelated but not necessarily independent. The asymptotic distribution of R for independent observations from bivariate elliptical population was obtained by Muirhead (1982, 157). For the distributions of R in nonelliptical populations, the reader is referred to Johnson, Kotz and Balakrishnan (1995) and the references therein.

It is conjectured that the distribution of correlation coefficient may have a nicer form if the following representation is used in the derivation:

$$(N-1)R = \sum_{j=1}^{N} (T_{1j} - \bar{T}_1)(T_{2j} - \bar{T}) \tag{6.1}$$

where $T_{ij} = (X_{ij} - \mu_i)/S_i$, $(i = 1, 2; j = 1, 2, \cdots, N)$. The conditional expectation can also be employed on T_{1j} and T_{2j} to have possibly better forms for the moments of R.

Acknowledgements

The author acknowledges the excellent research support provided by King Fahd University of Petroleum and Minerals, Saudi Arabia through the Project FT 2004/22. The author also thanks the referee for constructive suggestions that have helped improve the quality of the paper.

References

1. Ali, M.M. and Joarder, A.H. (1991). Distribution of the correlation coefficient for the class of bivariate elliptical models, *Canadian Journal of Statistics,* 19, 447-452.

2. Anderson, T.W. (2003). An *Introduction to Multivariate Statistical Analysis*. John Wiley and Sons. New York.

3. Billah, M.B. and Saleh, A.K.M.E. (2000). Performance of the large sample tests in the actual forecasting of the pretest estimators for regression model with *t*-error, *Journal of Applied Statistical Science,* 9(3), 237-251.

4. Fang, K.T. (1990). *Generalized Multivariate Analysis.* Springer Verlag.

5. Fang, K.T. and Anderson, T.W (1990). *Statistical Inference in Elliptically Contoured and Related Distributions.* Allerton Press.

6. Fang, K.T; Kotz, S. and Ng, K.W. (1990). *Symmetric Multivariate and Related Distributions.* Chapman and Hall.

7. Fisher, R.A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population, *Biometrika,* 10, 507-521.

8. Joarder, A.H. and Ahmed, S.E. (1998). Estimation of the trace of the scale matrix of a class of elliptical distributions, *Metrika,* 48, 149-160.

9. Johnson, N.L.; Kotz, S. and Balakrishnan, N. (1995). C*ontinuous Univariate Distributions,* Vol. 2, John Wiley.

10. Kibria, B.M.G. (2003). Robust predictive inference for the multivariate

linear models with elliptically contoured distribution, *Far East Journal of Theoretical Statistics*, 10(1), 11-24.

11. Kibria, B. M. G. (2004). Conflict among the shrinkage estimators induced by W, LR and LM tests under a Student's *t* regression model, *Journal of the Korean Statistical Society*, 33 (4), 411-433.

12. Kibria, B.M.G. and Haq, M.S. (1999). Predictive inference for the elliptical linear model, *Journal of Multivariate Analysis*, 68, 235-249.

13. Kibria, B.M.G. and Saleh, A.K.M.E. (2004). Preliminary test ridge regression estimators with Student's *t* errors and conflicting test statistics, *Metrika*, 59(2), 105-124.

14. Kotz, S. and Nadarajah, S. (2004). *Multivariate t Distributions and Their Applications.* Cambridge University Press, Cambridge, UK.

15. Lange, K.L.; Little, R.J.A. and Taylor, J.M.G. (1989) Robust statistical modeling using the *t* distribution. *Journal of American Statistical Association*, 84, 881-896.

16. Muddapur, M.V. (1988). A simple test for correlation coefficient in a bivariate normal distribution, *Sankhya*, Series B, 50, 50-68.

17. Muirhead, R.J. (1982). *Aspects of Multivariate Statistical Theory*. John Wiley and Sons, New York.

18. Samiuddin, M. (1970). On a test for an assigned value of correlation in a bivariate normal distribution, *Biometrika*, 57, 461-464.

19. Sutradhar, B.C. and Ali, M.M (1986). Estimation of the parameters of regression model with a multivariate *t* error variable. *Communications in Statistics: Theory and Methods*, 15, 429-450.

20. Sutradhar, B.C. and Ali, M.M. (1989). A generalization of the Wishart distribution for the elliptical model and its moments for the multivatiate *t* model, *Journal of Multivariate Analysis*, 29(1), 155-162.

21. Wishart, J. (1928). The generalized product moment distribution in samples from a normal multivariate population, *Biometrika*, A20, 32-52.