International Journal of Statistical Sciences *ISSN 1683–5603*
Vol. 5, 2006, pp 9-26

Statistical Inference for a Bivariate Normal Population with a Common Mean

Ngamphol Sunthornworasiri

Montip Tiensuwan
Department of Mathematics, Faculty of Science, Mahidol University,
Bangkok, Thailand 10400

Bimal K. Sinha
Department of Mathematics and Statistics, University of Maryland,
Baltimore County, MD 21250, USA

[Received August 8, 2006; Revised December 28, 2007; Accepted Janaury 1, 2007]

Abstract

In this paper we derive the maximum likelihood estimates (MLE) and the method of moments (MM) estimates of the parameters of a bivariate normal population with a common mean. We also discuss large sample tests for these parameters. An environmental application is indicated.

Keywords and Phrases: Bivariate normal population, common mean, maximum likelihood, method of moments, large sample tests.

AMS Classification: Primary 62H12; Secondary 62H15.

1 Introduction

In this paper we address the statistical inference for the common mean of a bivariate normal population with unequal variances. We discuss estimation and tests for the common mean, the variances and the correlation coefficient. Both maximum likelihood and method of moments estimates are derived and some important features of both the methods are pointed out. Our application includes an EPA small data set on Reid Vapor Pressure (RVP) (Nussbaum and Sinha, 1997; Yu *et al.*, 2002).

Section 2 describes the underlying model and the inference problems as well as the data set. In Section 3, details of data analysis under this model are given. In

Section 4, we provide the numerical results of the analysis of the data set. Finally, the simulated results of power of the proposed tests are shown in Section 5.

2 Model specification and description of the data set

We assume that a random sample $\{(x_i, y_i),\ i = 1, 2, ..., n\}$ is drawn from a bivariate normal population of (X, Y) with a common mean, with the parameters $(\mu, \mu, \sigma_1^2, \sigma_2^2, \rho)$. There are practical situations where the assumption of a common mean is valid. Our goal is to provide estimates and tests for the parameters under the above model. We consider both the maximum likelihood estimates as well as the moment estimates, and discuss their properties. We also provide relevant test statistics for testing hypothesis about the common mean, the variances, and the correlation coefficient.

Our data set illustrating a common mean scenario deals with the paired observations (x_i, y_i) representing field and lab data on RVP for 15 locations. This problem is motivated by the following practical issue in the context of the attempt by the Environmental Protection Agency (EPA) of the United States to evaluate the gasoline quality based on what is known as Reid Vapor Pressure (RVP) (Nussbaum and Sinha, 1997). Occasionally, an EPA inspector would visit gas pumps in a city, take samples of gasoline of a particular brand, and measure RVP right at the spot which produces cheap and quick measurements. Once in a while, the inspector after measuring RVP at the spot will also ship a gasoline sample to a laboratory for a measurement of presumably higher precision at a higher cost, thus getting the pair (field, lab). Since usually laboratory measurements (Y) are much more expensive than field measurements (X) because of special packaging to be used to ship a gasoline sample from a field to a laboratory, not all the gasoline samples will be shipped to the laboratory and hence the resulting data would consist of many field measurements with occasional paired measurements obtained from both the field and laboratory. Our statistical analysis here is based on only the paired data reported below in Table 1. The scenario is such that the means are equal, but the variances are different. We will return to the analysis of this data set in Section 4.

Table 1. The field and lab data on RVP for new reformulated gasoline

X	Y	X	Y	X	Y
8.03	8.28	9.28	9.14	7.88	7.89
8.64	8.63	7.86	7.86	8.56	8.48
9.14	9.28	7.83	7.90	7.83	7.95
7.86	7.85	8.60	8.52	7.99	8.32
8.70	8.62	7.83	7.92	7.56	7.60

3 Statistical analysis of the model $N(\mu, \mu, \sigma_1, \sigma_2, \rho)$

3.1 Estimation of parameters

We discuss both the method of maximum likelihood (MLE) and the method of moments (MM) for estimating the four parameters: $\mu, \sigma_1^2, \sigma_2^2$ and ρ.

3.1.1 Method of maximum likelihood (MLE)

Since the joint p.d.f. of x and y is

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu)^2}{\sigma_1^2}+\frac{(y-\mu)^2}{\sigma_2^2}-\frac{2\rho}{\sigma_1\sigma_2}(x-\mu)(y-\mu)\right]},$$

for the given data set $\{x_i, y_i\}$, $i = 1, 2, ..., n$, the joint p.d.f. or the likelihood can be simplified to

$$L\left(\mu, \sigma_1, \sigma_2, \rho | \underline{x}, \underline{y}\right) = \frac{1}{\left(2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}\right)^n} e^{-\frac{1}{2(1-\rho^2)}\left[\frac{S_x^2+n(\bar{x}-\mu)^2}{\sigma_1^2}+\frac{S_y^2+n(\bar{y}-\mu)^2}{\sigma_2^2}-\frac{2\rho}{\sigma_1\sigma_2}(S_{xy}+n(\bar{x}-\mu)(\bar{y}-\mu))\right]}$$

where $\bar{x} = \sum x_i/n$, $\bar{y} = \sum y_i/n$, $S_x^2 = \sum (x_i - \bar{x})^2$, $S_y^2 = \sum (y_i - \bar{y})^2$ and $S_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y})$.

Obviously, the sufficient statistics are $\bar{x}, \bar{y}, S_x^2, S_y^2$ and S_{xy}, and these are not complete. Equating the first derivatives of $\ln L$ with respect to the parameters μ, σ_1, σ_2 and ρ to zero, and solving these equations, we get the MLEs as

$$\hat{\mu}_{MLE} = \frac{\bar{x}\left(S_y^2 - S_{xy}\right) + \bar{y}\left(S_x^2 - S_{xy}\right)}{S_x^2 + S_y^2 - 2S_{xy}}, \tag{1}$$

$$\hat{\sigma}_{1,MLE}^2 = \frac{1}{n}\left[S_x^2 + \frac{n\left(\bar{x}-\bar{y}\right)^2\left(S_x^2 - S_{xy}\right)^2}{\left(S_x^2 + S_y^2 - 2S_{xy}\right)^2}\right], \tag{2}$$

$$\hat{\sigma}_{2,MLE}^2 = \frac{1}{n}\left[S_y^2 + \frac{n\left(\bar{x}-\bar{y}\right)^2\left(S_y^2 - S_{xy}\right)^2}{\left(S_x^2 + S_y^2 - 2S_{xy}\right)^2}\right], \tag{3}$$

$$\hat{\rho}_{MLE} = \frac{S_{xy} - \frac{n(\bar{x}-\bar{y})^2\left(S_x^2-S_{xy}\right)\left(S_y^2-S_{xy}\right)}{\left(S_x^2+S_y^2-2S_{xy}\right)^2}}{\sqrt{S_x^2 + \frac{n(\bar{x}-\bar{y})^2\left(S_x^2-S_{xy}\right)^2}{\left(S_x^2+S_y^2-2S_{xy}\right)^2}} \cdot \sqrt{S_y^2 + \frac{n(\bar{x}-\bar{y})^2\left(S_y^2-S_{xy}\right)^2}{\left(S_x^2+S_y^2-2S_{xy}\right)^2}}}. \tag{4}$$

3.1.2 Method of Moments (MM)

From $E(X) = \mu$, $E(Y) = \mu$, $E(X^2) = \mu^2 + \sigma_1^2$, $E(Y^2) = \mu^2 + \sigma_2^2$ and $E(XY) = Cov(X,Y) + \mu^2 = \rho\sigma_1\sigma_2 + \mu^2$, we can write:

$\mu^2 + \sigma_1^2 = \sum_{i=1}^{n} x_i^2 \Big/ n$, $\mu^2 + \sigma_2^2 = \sum_{i=1}^{n} y_i^2 \Big/ n$ and $\mu^2 + \rho\sigma_1\sigma_2 = \sum_{i=1}^{n} x_i y_i \Big/ n$.

Since an estimate of μ seems arbitrary and not unique, we choose to use the MLE of μ under the assumption that σ_1^2, σ_2^2 and ρ are known, and then replace σ_1^2, σ_2^2 and ρ by MM estimates.
Since

$$\hat{\mu}_{MLE}\left(\sigma_1^2, \sigma_2^2, \rho\right) = \frac{\bar{x}\sigma_2^2 + \bar{y}\sigma_1^2 - \rho\sigma_1\sigma_2\left(\bar{x} + \bar{y}\right)}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} = \frac{\bar{x}\left(\sigma_2^2 - \rho\sigma_1\sigma_2\right) + \bar{y}\left(\sigma_1^2 - \rho\sigma_1\sigma_2\right)}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2},$$

using $\sigma_1^2 - \rho\sigma_1\sigma_2 = \frac{\sum_{i=1}^{n} x_i^2}{n} - \frac{\sum_{i=1}^{n} x_i y_i}{n}$ and $\sigma_2^2 - \rho\sigma_1\sigma_2 = \frac{\sum_{i=1}^{n} y_i^2}{n} - \frac{\sum_{i=1}^{n} x_i y_i}{n}$,
we get

$$\hat{\mu}_{MM} = \frac{\bar{x}\left(\sum_{i=1}^{n} y_i^2 - \sum_{i=1}^{n} x_i y_i\right) + \bar{y}\left(\sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i y_i\right)}{\sum_{i=1}^{n} (x_i - y_i)^2}. \quad (5)$$

Using (5), the MM estimates of σ_1^2, σ_2^2 and ρ are obtained as

$$\hat{\sigma}_{1,MM}^2 = \frac{\sum_{i=1}^{n} x_i^2}{n} - \left[\frac{\bar{x}\left(\sum_{i=1}^{n} y_i^2 - \sum_{i=1}^{n} x_i y_i\right) + \bar{y}\left(\sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i y_i\right)}{\sum_{i=1}^{n} (x_i - y_i)^2}\right]^2, \quad (6)$$

$$\hat{\sigma}_{2,MM}^2 = \frac{\sum_{i=1}^{n} y_i^2}{n} - \left[\frac{\bar{x}\left(\sum_{i=1}^{n} y_i^2 - \sum_{i=1}^{n} x_i y_i\right) + \bar{y}\left(\sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i y_i\right)}{\sum_{i=1}^{n} (x_i - y_i)^2}\right]^2, \quad (7)$$

$$\hat{\rho}_{MM} = \frac{\frac{\sum_{i=1}^{n} x_i y_i}{n} - \hat{\mu}_{MM}^2}{\sqrt{\hat{\sigma}_{1,MM}^2}\sqrt{\hat{\sigma}_{2,MM}^2}}. \quad (8)$$

Expressed in terms of $\bar{x}, \bar{y}, S_x^2, S_y^2$ and S_{xy}, the above estimates reduce to

$$\hat{\mu}_{MM} = \frac{\bar{x}\left(S_y^2 - S_{xy}\right) + \bar{y}\left(S_x^2 - S_{xy}\right)}{S_x^2 + S_y^2 - 2S_{xy} + n(\bar{x}-\bar{y})^2}, \tag{9}$$

$$\hat{\sigma}_{1,MM}^2 = \frac{S_x^2 + n\bar{x}^2}{n} - \left[\frac{\bar{x}\left(S_y^2 - S_{xy}\right) + \bar{y}\left(S_x^2 - S_{xy}\right)}{S_x^2 + S_y^2 - 2S_{xy} + n(\bar{x}-\bar{y})^2}\right]^2, \tag{10}$$

$$\hat{\sigma}_{2,MM}^2 = \frac{S_y^2 + n\bar{y}^2}{n} - \left[\frac{\bar{x}\left(S_y^2 - S_{xy}\right) + \bar{y}\left(S_x^2 - S_{xy}\right)}{S_x^2 + S_y^2 - 2S_{xy} + n(\bar{x}-\bar{y})^2}\right]^2, \tag{11}$$

$$\hat{\rho}_{MM} = \frac{\frac{S_{xy}+n\bar{x}\bar{y}}{n} - \left[\frac{\bar{x}\left(S_y^2-S_{xy}\right)+\bar{y}\left(S_x^2-S_{xy}\right)}{S_x^2+S_y^2-2S_{xy}+n(\bar{x}-\bar{y})^2}\right]^2}{\sqrt{\frac{S_x^2+n\bar{x}^2}{n} - \left[\frac{\bar{x}\left(S_y^2-S_{xy}\right)+\bar{y}(S_x^2-S_{xy})}{S_x^2+S_y^2-2S_{xy}+n(\bar{x}-\bar{y})^2}\right]^2} \cdot \sqrt{\frac{S_y^2+n\bar{y}^2}{n} - \left[\frac{\bar{x}\left(S_y^2-S_{xy}\right)+\bar{y}(S_x^2-S_{xy})}{S_x^2+S_y^2-2S_{xy}+n(\bar{x}-\bar{y})^2}\right]^2}}. \tag{12}$$

3.1.3 Properties of estimates

In this section we study the large sample properties of the MLE and MM estimates. A brief description of the necessary theory is given in the Appendix.

The results given in the Appendix can be used to derive the expressions of large sample means and variances of the estimates of the parameters. This is done below by first expressing the ML and MM estimates in terms of

$$\underline{T} = (T_1, T_2, T_3, T_4, T_5) = \left(\bar{x}, \bar{y}, \frac{S_x^2}{n-1}, \frac{S_y^2}{n-1}, \frac{S_{xy}}{n-1}\right).$$

MLEs:

$$\hat{\mu}_{MLE} = \frac{T_1(T_4 - T_5) + T_2(T_3 - T_5)}{T_3 + T_4 - 2T_5},$$

$$\hat{\sigma}_{1,MLE}^2 = T_3 + \frac{(T_1 - T_2)^2 (T_3 - T_5)^2}{(T_3 + T_4 - 2T_5)^2}, \quad \hat{\sigma}_{2,MLE}^2 = T_4 + \frac{(T_1 - T_2)^2 (T_4 - T_5)^2}{(T_3 + T_4 - 2T_5)^2},$$

$$\hat{\rho}_{MLE} = \frac{T_5 - \frac{(T_1-T_2)^2(T_3-T_5)\,(T_4-T_5)}{(T_3+T_4-2T_5)^2}}{\sqrt{T_3 - \frac{(T_1-T_2)^2(T_3-T_5)^2}{(T_3+T_4-2T_5)^2}} \cdot \sqrt{T_4 - \frac{(T_1-T_2)^2(T_4-T_5)^2}{(T_3+T_4-2T_5)^2}}}.$$

Moment Estimates:

$$\hat{\mu}_{MM} = \frac{T_1(T_4-T_5)+T_2(T_3-T_5)}{T_3+T_4-2T_5+n(T_1-T_2)^2}, \quad \hat{\sigma}^2_{1,MM} = T_3 + T_1^2 - \left[\frac{T_1(T_4-T_5)+T_2(T_3-T_5)}{(T_3+T_4-2T_5)+(T_1-T_2)^2}\right]^2,$$

$$\hat{\sigma}^2_{2,MM} = T_4 + T_2^2 - \left[\frac{T_1(T_4-T_5)+T_2(T_3-T_5)}{(T_3+T_4-2T_5)+(T_1-T_2)^2}\right]^2,$$

$$\hat{\rho}_{MM} = \frac{T_5 + T_1T_2 - \left[\frac{T_1(T_4-T_5)+T_2(T_3-T_5)}{(T_3+T_4-2T_5)+(T_1-T_2)^2}\right]^2}{\sqrt{T_3 + T_1^2 - \left[\frac{T_1(T_4-T_5)+T_2(T_3-T_5)}{(T_3+T_4-2T_5)+(T_1-T_2)^2}\right]^2} \cdot \sqrt{T_4 + T_2^2 - \left[\frac{T_1(T_4-T_5)+T_2(T_3-T_5)}{(T_3+T_4-2T_5)+(T_1-T_2)^2}\right]^2}}.$$

We now apply the general result for the mean and variance of $\phi(\underline{T})$ given in the Appendix and readily get the following results.

Method of MLE:

$$E(\hat{\mu}_{MLE}) = \mu, \quad Var(\hat{\mu}_{MLE}) = \frac{(1-\rho^2)\sigma_1^2\sigma_2^2}{(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)n} + \frac{(1-\rho^2)\sigma_1^2\sigma_2^2}{(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)n^2} + o\left(1/n^2\right),$$

$$E(\hat{\sigma}^2_{1,MLE}) = \sigma_1^2 + \frac{\sigma_1^2(\sigma_1-\rho\sigma_2)^2}{(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)n} + o\left(1/n\right), \quad Var(\hat{\sigma}^2_{1,MLE}) = \frac{2\sigma_1^4}{n} + o\left(1/n\right),$$

$$E(\hat{\sigma}^2_{2,MLE}) = \sigma_2^2 + \frac{\sigma_2^2(\sigma_1-\rho\sigma_2)^2}{(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)n} + o\left(1/n\right), \quad Var(\hat{\sigma}^2_{2,MLE}) = \frac{2\sigma_2^4}{n} + o\left(1/n\right),$$

$$E(\hat{\rho}_{MLE}) = \rho + \frac{\sigma_1\sigma_2(1-\rho^2)^2}{(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)n} + o\left(1/n\right), \quad Var(\hat{\rho}_{MLE}) = \frac{(1-\rho^2)^2}{n} + o\left(1/n\right).$$

Method of Moments:

$$E(\hat{\mu}_{MM}) = \mu - \frac{\mu}{n}, \quad Var(\hat{\mu}_{MM}) = \frac{(1-\rho^2)\sigma_1^2\sigma_2^2}{(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)n} + o\left(1/n\right),$$

$$E(\hat{\sigma}^2_{1,MM}) = \sigma_1^2 + O\left(1/n\right),$$

$$Var(\hat{\sigma}^2_{1,MM}) = \frac{2\sigma_1^2\left(\sigma_1^2(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2) + 2\mu^2(\sigma_1-\rho\sigma_2)^2\right)}{(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)n} + o\left(1/n\right),$$

$$E(\hat{\sigma}^2_{2,MM}) = \sigma_2^2 + O\left(1/n\right),$$

$$Var(\hat{\sigma}^2_{2,MM}) = \frac{2\sigma_2^2\left(\sigma_2^2(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)+2\mu^2(\sigma_2-\rho\sigma_1)^2\right)}{(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)n} + o\left(1/n\right),$$

$$E(\hat{\rho}_{MM}) = \rho + O\left(1/n\right),$$

$$Var(\hat{\rho}_{MM}) = \frac{(1-\rho^2)(1+\rho^2)\left[\sigma_1^2\sigma_2^2(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)+\mu^2(\sigma_1^2-\sigma_2^2)^2(\sigma_1^2+\sigma_2^2)^2\right]}{n\sigma_1^2\sigma_2^2(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)} + o\left(1/n\right)$$

$$\approx \frac{1-\rho^4}{n} \text{ if } \sigma_1^2=\sigma_2^2.$$

3.1.4 Comparison of the estimates

In this section we compare the MLEs and MMs of $\mu, \sigma_1^2, \sigma_2^2$ and ρ based on their large sample properties. We first discuss on the basis of terms up to $O\left(1/n\right)$. Details are omitted.

(i) Estimation of μ :

Since $Var(\hat{\mu}_{MLE}) = \dfrac{(1-\rho^2)\sigma_1^2\sigma_2^2}{(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)n} = Var(\hat{\mu}_{MM})$, $\hat{\mu}_{MLE} \approx \hat{\mu}_{MM}$, up to $O\left(1/n\right)$.

(ii) Estimation of σ_1^2 :

The variance of $\hat{\sigma}_1^2$ for method of MLE and MM, up to $O\left(1/n\right)$, are $Var(\hat{\sigma}^2_{1,MLE}) = 2\sigma_1^4/n$ and

$$Var(\hat{\sigma}^2_{1,MM}) = \frac{2\sigma_1^2\left(\sigma_1^2(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)+2\mu^2(\sigma_1-\rho\sigma_2)^2\right)}{(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)n}.$$

Obviously, when $\mu=0$, $Var(\hat{\sigma}^2_{1,MLE}) \approx Var(\hat{\sigma}^2_{1,MM})$. So $\hat{\sigma}^2_{1,MLE} \approx \hat{\sigma}^2_{1,MM}$, up to $O\left(1/n\right)$.

When $\mu \neq 0$, $Var(\hat{\sigma}^2_{1,MLE}) < Var(\hat{\sigma}^2_{1,MM})$. So $\hat{\sigma}^2_{1,MLE}$ is better than $\hat{\sigma}^2_{1,MM}$, up to $O\left(1/n\right)$.

(iii) Estimation of σ_2^2 :

Similarly, $Var(\hat{\sigma}^2_{2,MLE}) = 2\sigma_2^4/n$ and

$$Var(\hat{\sigma}^2_{2,MM}) = \frac{2\sigma_2^2\left(\sigma_2^2(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)+2\mu^2(\sigma_2-\rho\sigma_1)^2\right)}{(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)n}.$$

Again, when $\mu=0$, $Var(\hat{\sigma}^2_{2,MLE}) \approx Var(\hat{\sigma}^2_{2,MM})$. So $\hat{\sigma}^2_{2,MLE} \approx \hat{\sigma}^2_{2,MM}$, up to $O\left(1/n\right)$.

When $\mu \neq 0$, $Var(\hat{\sigma}^2_{2,MLE}) < Var(\hat{\sigma}^2_{2,MM})$. So $\hat{\sigma}^2_{2,MLE}$ is better than $\hat{\sigma}^2_{2,MM}$, up to $O\left(1/n\right)$.

(iv) Estimation of ρ :

The variance of $\hat{\rho}$ for method of MLE and MM, up to $O\left(1/n\right)$, are the following $Var(\hat{\rho}_{MLE}) = (1-\rho^2)^2/n$ and

$$Var(\hat{\rho}_{MM}) = \frac{(1-\rho^2)(1+\rho^2)\left[\sigma_1^2\sigma_2^2(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)+\mu^2(\sigma_1^2-\sigma_2^2)^2(\sigma_1^2+\sigma_2^2)^2\right]}{n\sigma_1^2\sigma_2^2(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)}.$$

Obviously, when $\mu = 0$, $Var(\hat{\rho}_{MLE}) \approx Var(\hat{\rho}_{MM})$. So $\hat{\rho}_{MLE} \approx \hat{\rho}_{MM}$, up to $O(1/n)$. When $\mu \neq 0$, $Var(\hat{\rho}_{MLE}) < Var(\hat{\rho}_{MM})$. So $\hat{\rho}_{MLE}$ is better than $\hat{\rho}_{MM}$, up to $O(1/n)$.

Next we discuss the comparison of the estimates based on terms up to $O\left(1/n^2\right)$. Our observations are the following. Details are omitted.

(i)' For $\hat{\mu}_{MLE}$ and $\hat{\mu}_{MM}$,

when $\mu = 0$, $Var(\hat{\mu}_{MLE}) \approx Var(\hat{\mu}_{MM})$. So $\hat{\mu}_{MLE} \approx \hat{\mu}_{MM}$,
when $\mu \neq 0$, $Var(\hat{\mu}_{MLE}) < Var(\hat{\mu}_{MM})$. So $\hat{\mu}_{MLE}$ is better than $\hat{\mu}_{MM}$.

(ii)' For $\hat{\sigma}^2_{1,MLE}$ and $\hat{\sigma}^2_{1,MM}$,

when $\mu = 0$, $Var(\hat{\sigma}^2_{1,MLE}) < Var(\hat{\sigma}^2_{1,MM})$. So $\hat{\sigma}^2_{1,MLE}$ is better than $\hat{\sigma}^2_{1,MM}$,
when $\mu \neq 0$, the comparison is not straightforward.

(iii)' For $\hat{\sigma}^2_{2,MLE}$ and $\hat{\sigma}^2_{2,MM}$,

when $\mu = 0$, $Var(\hat{\sigma}^2_{2,MLE}) < Var(\hat{\sigma}^2_{2,MM})$. So $\hat{\sigma}^2_{2,MLE}$ is better than $\hat{\sigma}^2_{2,MM}$,
when $\mu \neq 0$, the comparison is not straightforward.

(iv)' For $\hat{\rho}_{MLE}$ and $\hat{\rho}_{MM}$,
when $\mu = 0$, $Var(\hat{\rho}_{MLE}) < Var(\hat{\rho}_{MM})$. So $\hat{\rho}_{MLE}$ is better than $\hat{\rho}_{MM}$,
when $\mu \neq 0$, the comparison is not straightforward.

3.2 Tests of hypotheses

In this section we consider the problem of constructing large sample tests for suitable hypotheses of the basic parameters $\mu, \sigma_1^2, \sigma_2^2$ and ρ.

A. Test for $H_0 : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$ where μ_0 is a given constant.
The test statistic is given by

Method of MLE: $$T_{1\mu} = \frac{\hat{\mu}_{MLE} - \mu_0}{\sqrt{\hat{V}ar(\hat{\mu}_{MLE})}}$$

Method of Moments: $$T_{2\mu} = \frac{\hat{\mu}_{MM} - \mu_0}{\sqrt{\hat{V}ar(\hat{\mu}_{MM})}}$$

where $\hat{V}ar(\hat{\mu}_{MLE})$ and $\hat{V}ar(\hat{\mu}_{MM})$ are the estimated variances up to order $O(1/n)$.
We reject H_0 if $|T_{1\mu}| > Z_{\alpha/2}$ in case of $T_{1\mu}$, and if $|T_{2\mu}| > Z_{\alpha/2}$ in case of $T_{2\mu}$.

B. Test for $H_0 : \sigma_1^2 = \sigma_2^2$ against $H_1 : \sigma_1^2 \neq \sigma_2^2$.
The test statistic is given by

Method of MLE: $$T_{1\sigma} = \frac{\hat{\sigma}^2_{1,MLE} - \hat{\sigma}^2_{2,MLE}}{\sqrt{\hat{V}ar(\hat{\sigma}^2_{1,MLE} - \hat{\sigma}^2_{2,MLE})}}$$

Method of Moments: $$T_{2\sigma} = \frac{\hat{\sigma}^2_{1,MM} - \hat{\sigma}^2_{2,MM}}{\sqrt{\hat{V}ar(\hat{\sigma}^2_{1,MM} - \hat{\sigma}^2_{2,MM})}}$$

where $\hat{V}ar(\hat{\sigma}^2_{1,MLE} - \hat{\sigma}^2_{2,MLE})$ and $\hat{V}ar(\hat{\sigma}^2_{1,MM} - \hat{\sigma}^2_{2,MM})$ are the estimated variances up

to order $O(1/n)$. We reject H_0 if $|T_{1\sigma}| > Z_{\alpha/2}$ in case of $T_{1\sigma}$, and if $|T_{2\sigma}| > Z_{\alpha/2}$ in case of $T_{2\sigma}$.

C. Test for $H_0 : \rho = \rho_0$ against $H_1 : \rho \neq \rho_0$ where ρ_0 is a given constant.
In case of $\hat{\rho}_{MLE}$, we test this hypothesis based on the variance stabilizing transformation
$h(\rho) = \text{arctanh } (\rho)$ since $E(\hat{\rho}_{MLE}) \approx \rho$ and $Var(\hat{\rho}_{MLE}) \approx (1-\rho^2)^2/n$.
We use the test statistic:

$$T_{1\rho} = \sqrt{n}\left(h(r) - h(\rho_0)\right).$$

Also, based on $\hat{\rho}_{MM}$, we propose to use the test statistic

$$T_{2\rho} = \frac{\hat{\rho}_{MM} - \rho_0}{\sqrt{\hat{V}ar(\hat{\rho}_{MM})}}$$

where $\hat{V}ar(\hat{\rho}_{MM})$ is the estimated variance up to order $O(1/n)$, Under H_0, $T_{1\rho}$ and $T_{2\rho}$ follow asymptotic standard normal distributions.

4 An Application

In this section we provide the numerical results of the analysis of the data set given in Table 1. The summary statistics are shown in Table 2.

Table 2. Summary statistics for Data in Table 1

N	$\bar{x}$	$\bar{y}$	s_x^2	s_y^2	r	S_x^2	S_y^2	S_{xy}
15	8.2393	8.2827	0.2848	0.2454	0.9716	3.9868	3.4355	3.5957

4.1 Analysis of data set

We apply the basic bivariate model given earlier to analyze this data set. It may be noted that, as expected, the sample lab variability is smaller than the sample field variability. From the summary statistics in Table 2, we have T_1= 8.2393, T_2= 8.2827, T_3= 0.2848, T_4= 0.2454 and T_5= 0.2568. Based on these statistics, we can compute the estimated values of the parameters as shown in Table 3.

Table 3. The estimated values of the MLE and MM estimates

MLE	estimated values	MM	estimated values
$\hat{\mu}_{MLE}$	8.3127	$\hat{\mu}_{MM}$	7.4631
$\hat{\sigma}^2_{1,MLE}$	0.2902	$\hat{\sigma}^2_{1,MM}$	12.4728
$\hat{\sigma}^2_{2,MLE}$	0.2463	$\hat{\sigma}^2_{2,MM}$	13.1494
$\hat{\rho}_{MLE}$	0.9690	$\hat{\rho}_{MM}$	0.9996

Since the sample size of the above data set is small (only 15), rather than using the large sample theory as developed in Section 3, we use the famous resampling technique to derived the cut-off points of the various test statistics mentioned in Section 3. Table 4 shows the 5% and 95% cut-off points of the three test statistics from 800 resample data sets.

Table 4. The 5% and 95% cut-off points for each test statistic

Hypothesis Testing	Method of MLE		Method of moments	
	5%	95%	5%	95%
$H_0 : \mu = 8$	0.7117	4.2795	-138.3927	3.6289
$H_0 : \sigma_1^2 = \sigma_2^2$	-0.4879	2.1540	-39.4765	0.5802
$H_0 : \rho = 0.9$	0.3964	4.5923	1.7341	143.1814

We now use the above cut-off points to carry out three hypotheses of interest.

A. Test for $H_0 : \mu = 8$ against $H_1 : \mu \neq 8$.

The value of the test statistic is given by

Method of MLE: $$T_{1\mu} = \frac{\hat{\mu}_{MLE} - 8}{\sqrt{\hat{V}ar(\hat{\mu}_{MLE})}} = 2.4856$$

Method of Moments: $$T_{2\mu} = \frac{\hat{\mu}_{MM} - 8}{\sqrt{\hat{V}ar(\hat{\mu}_{MM})}} = \text{-}12.7933$$

Since $0.7117 < T_{1\mu} < 4.2795$ and $-138.3927 < T_{2\mu} < 3.6289$, so we accept H_0 by both the tests.

B. Test for $H_0 : \sigma_1^2 = \sigma_2^2$ against $H_1 : \sigma_1^2 \neq \sigma_2^2$.

The value of the test statistic is given by

Method of MLE: $$T_{1\sigma} = \frac{\hat{\sigma}^2_{1,MLE} - \hat{\sigma}^2_{2,MLE}}{\sqrt{\hat{V}ar(\hat{\sigma}^2_{1,MLE} - \hat{\sigma}^2_{2,MLE})}} = 1.1642$$

Method of Moments: $$T_{2\sigma} = \frac{\hat{\sigma}^2_{1,MM} - \hat{\sigma}^2_{2,MM}}{\sqrt{\hat{V}ar(\hat{\sigma}^2_{1,MM} - \hat{\sigma}^2_{2,MM})}} = \text{-}3.9138$$

Since -0.4879 $< T_{1\sigma} <$ 2.1540 and -39.4765 $< T_{2\sigma} <$ 0.5802, so we accept H_0 by both the tests.

C. Test for $H_0 : \rho = 0.9$ against $H_1 : \rho \neq 0.9$.
The value of the test statistic is given by

Method of MLE: $$T_{1\rho} = \sqrt{n}\,(h(r) - h(\rho_0)) = 2.3371$$

Method of Moments: $$T_{2\rho} = \frac{\hat{\rho}_{MM} - \rho_0}{\sqrt{\hat{V}ar(\hat{\rho}_{MM})}} = 19.0788$$

Since 0.3964 $< T_{1\rho} <$ 4.5923, and 1.7341 $< T_{2\rho} <$ 143.1814, so we accept H_0 by both the tests.

5 Power of the proposed tests

In this section we provide some simulated results of power of the proposed tests in order to compare the tests derived by two methods: MLE and MM. We generate paired data form a bivariate normal distribution with $n = 5$ and 10. For testing $H_0 : \mu = 0$, we generate data under $\mu = 0$ (size) and different values of μ (power). For testing $H_0 : \sigma_1^2 = \sigma_2^2$ and $H_0 : \rho = 0.5$, we take $\mu = 0$ without loss of generality. We have computed $T_{1\mu}$, $T_{2\mu}$, $T_{1\sigma}$, $T_{2\sigma}$, $T_{1\rho}$ and $T_{2\rho}$ with 1,000 runs. The results are shown in Tables 5 – 7 below.

Table 5. Simulated power of the test: $H_0 : \mu = 0$ when $\sigma_1^2 = 1$, $n = 5$ and 10

σ_2^2	ρ	Test	n = 5						n = 10					
			μ						μ					
			-1.5	-1	-0.5	0.5	1	1.5	-1.5	-1	-0.5	0.5	1	1.5
0.6	0.2	$T_{1\mu}$	0.051	0.051	0.051	0.051	0.051	0.051	0.049	0.049	0.049	0.049	0.049	0.049
		$T_{2\mu}$	0.168	0.116	0.070	0.066	0.099	0.150	0.141	0.103	0.065	0.053	0.069	0.100
	0.4	$T_{1\mu}$	0.051	0.051	0.051	0.051	0.051	0.051	0.050	0.050	0.050	0.050	0.050	0.050
		$T_{2\mu}$	0.192	0.117	0.071	0.063	0.106	0.164	0.148	0.104	0.067	0.055	0.075	0.110
	0.6	$T_{1\mu}$	0.051	0.051	0.051	0.051	0.051	0.051	0.051	0.051	0.051	0.051	0.051	0.051
		$T_{2\mu}$	0.221	0.142	0.080	0.070	0.122	0.194	0.143	0.100	0.065	0.063	0.086	0.127
1.5	0.2	$T_{1\mu}$	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050
		$T_{2\mu}$	0.109	0.080	0.056	0.060	0.072	0.109	0.103	0.078	0.056	0.054	0.068	0.090
	0.4	$T_{1\mu}$	0.049	0.049	0.049	0.049	0.049	0.049	0.050	0.050	0.050	0.050	0.050	0.050
		$T_{2\mu}$	0.115	0.079	0.056	0.060	0.077	0.119	0.112	0.085	0.060	0.054	0.070	0.095
	0.6	$T_{1\mu}$	0.051	0.051	0.051	0.051	0.051	0.051	0.050	0.050	0.050	0.050	0.050	0.050
		$T_{2\mu}$	0.127	0.090	0.056	0.061	0.090	0.130	0.118	0.091	0.062	0.051	0.073	0.102

Table 6. Simulated power of the test: $H_0 : \sigma_1^2 = \sigma_2^2$ when $\sigma_1^2 = 1$, $n = 5$ and 10

ρ	Test	n = 5				n = 10			
		σ_2^2				σ_2^2			
		0.6	0.8	1.5	2	0.6	0.8	1.5	2
0.2	$T_{1\sigma}$	0.064	0.053	0.065	0.090	0.101	0.060	0.071	0.152
	$T_{2\sigma}$	0.061	0.050	0.059	0.095	0.144	0.076	0.077	0.142
0.4	$T_{1\sigma}$	0.071	0..059	0.059	0.078	0.128	0.063	0.083	0.171
	$T_{2\sigma}$	0.057	0.044	0.071	0.095	0.141	0.085	0.076	0.134
0.6	$T_{1\sigma}$	0.079	0.063	0.062	0.085	0.159	0.071	0.101	0.200
	$T_{2\sigma}$	0.045	0.041	0.058	0.071	0.151	0.071	0.077	0.148

Table 7. Simulated power of the test: $H_0 : \rho = 0.5$ when $\sigma_1^2 = 1$, $n = 5$ and 10

σ_2^2	Test	n = 5						n = 10					
		ρ						ρ					
		0.1	0.3	0.4	0.6	0.7	0.8	0.1	0.3	0.4	0.6	0.7	0.8
0.6	$T_{1\rho}$	0.045	0.044	0.045	0.050	0.046	0.041	0.046	0.044	0.047	0.046	0.049	0.051
	$T_{2\rho}$	0.083	0.062	0.055	0.048	0.046	0.039	0.085	0.065	0.058	0.047	0.051	0.048
0.8	$T_{1\rho}$	0.039	0.044	0.045	0.049	0.044	0.0.38	0.045	0.046	0.045	0.051	0.053	0.060
	$T_{2\rho}$	0.081	0.061	0.055	0.047	0.038	0.036	0.085	0.067	0.058	0.048	0.047	0.046
1.5	$T_{1\rho}$	0.053	0.048	0.048	0.046	0.045	0.043	0.047	0.045	0.047	0.048	0.055	0.063
	$T_{2\rho}$	0.085	0.068	0.056	0.048	0.036	0.041	0.085	0.066	0.055	0.052	0.048	0.042
2	$T_{1\rho}$	0.043	0.048	0.048	0.049	0.046	0.041	0.050	0.048	0.050	0.050	0.056	0.068
	$T_{2\rho}$	0.087	0.062	0.053	0.044	0.034	0.042	0.083	0.065	0.054	0.051	0.048	0.051

We can see from the tables above that for testing $H_0 : \mu = 0$, the test based on $T_{2\mu}$ is better than the test based on $T_{1\mu}$. However, for testing $H_0 : \sigma_1^2 = \sigma_2^2$, it turns out that the test based on $T_{1\sigma}$ is better than the test based on$T_{2\sigma}$. For testing $H_0 : \rho = 0.5$, we see that the test based on $T_{2\rho}$ is better than the test based on $T_{1\rho}$ for $\rho < 0.5$. The conclusion reverses for $\rho > 0.5$.

Acknowledgement

We thank the Thailand Research Fund (TRF) for research grant.

References

1. Anderson, T. W. (2003). An Introduction to Multivariate Statistical Analysis (3rd ed.), John Wiley and Sons, New York, USA.

2. Nussbaum B. D. and Sinha B. K. (1997). Cost effective gasoline sampling using ranked set sampling, Proceeding of the Section on Statistics and Environment. American Statistical Association, August issue, 83-87.

3. Rao, C. R. (1973). Linear Statistical Inference and Its Applications (2nd ed.), John Wiley and Sons, New York, USA.

4. Yu, P. L. H., Sun, Y. and Sinha, B. K. (2002). Estimation of the common mean of a bivariate normal population, Ann. Inst. Statist. Math, 54(4), 861-878.

Appendix

A brief description of large sample theory

Let $\underline{T} = (T_1, T_2, T_3, T_4, T_5) = \left(\bar{x}, \bar{y}, \frac{S_x^2}{n-1}, \frac{S_y^2}{n-1}, \frac{S_{xy}}{n-1}\right)$

where $\bar{x} = \sum x_i/n$, $\bar{y} = \sum y_i/n$, $S_x^2 = \sum (x_i - \bar{x})^2$, $S_y^2 = \sum (y_i - \bar{y})^2$ and $S_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y})$.

Note that the five basic statistics are all computed from the original bivariate data: $\{x_i, y_i\}$, $i = 1, 2, ..., n$, and our MLE and MM estimates are all function of $\underline{T}$.

Large Sample Technique:

Let $\phi(T_1, T_2, T_3, T_4, T_5) = \phi(\underline{T})$ be a smooth function of $\underline{T}$ (derivatives up to some orders exist). Our goal is to find the large sample mean and variance of $\phi(\underline{T})$.

Facts to be used:

§ 1. $\begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} = \begin{pmatrix} T_1 \\ T_2 \end{pmatrix} \sim N\left[\begin{pmatrix} \mu \\ \mu \end{pmatrix}, \frac{1}{n}\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}\right]$.

§ 2. $(n-1)\begin{pmatrix} T_3 & T_5 \\ T_5 & T_4 \end{pmatrix} = \begin{pmatrix} S_x^2 & S_{xy} \\ S_{xy} & S_y^2 \end{pmatrix} \sim W\left(\sum, \ n-1\right)$, where $\sum = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$.

§ 3. $\begin{pmatrix} T_1 \\ T_2 \end{pmatrix}$ and $\begin{pmatrix} T_3 & T_5 \\ T_5 & T_4 \end{pmatrix}$ are independent.

Taylor series expansion

Let $E(T_1) = \theta_1$, $E(T_2) = \theta_2$, $E(T_3) = \theta_3$, $E(T_4) = \theta_4$ and $E(T_5) = \theta_5$. To derive the large sample mean and variances of $\phi(\underline{T})$, we use the standard Taylor expansion as follows.

$$\phi(\underline{T}) = \phi(T_1, T_2, T_3, T_4, T_5)$$
$$= \phi(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5) + \left[(T_1 - \theta_1) \left.\frac{\partial \phi}{\partial T_1}\right|_{\underline{T}=\underline{\theta}} + (T_2 - \theta_2) \left.\frac{\partial \phi}{\partial T_2}\right|_{\underline{T}=\underline{\theta}} + (T_3 - \theta_3) \left.\frac{\partial \phi}{\partial T_3}\right|_{\underline{T}=\underline{\theta}} \right.$$
$$\left. + (T_4 - \theta_4) \left.\frac{\partial \phi}{\partial T_4}\right|_{\underline{T}=\underline{\theta}} + (T_5 - \theta_5) \left.\frac{\partial \phi}{\partial T_5}\right|_{\underline{T}=\underline{\theta}} \right] + \frac{1}{2} \left[(T_1 - \theta_1)^2 \left.\frac{\partial^2 \phi}{\partial T_1^2}\right|_{\underline{T}=\underline{\theta}} + (T_2 - \theta_2)^2 \left.\frac{\partial^2 \phi}{\partial T_2^2}\right|_{\underline{T}=\underline{\theta}} \right.$$
$$+ (T_3 - \theta_3)^2 \left.\frac{\partial^2 \phi}{\partial T_3^2}\right|_{\underline{T}=\underline{\theta}} + (T_4 - \theta_4)^2 \left.\frac{\partial^2 \phi}{\partial T_4^2}\right|_{\underline{T}=\underline{\theta}} + (T_5 - \theta_5)^2 \left.\frac{\partial^2 \phi}{\partial T_5^2}\right|_{\underline{T}=\underline{\theta}}$$
$$+ 2(T_1 - \theta_1)(T_2 - \theta_2) \left.\frac{\partial^2 \phi}{\partial T_1 \partial T_2}\right|_{\underline{T}=\underline{\theta}} + 2(T_1 - \theta_1)(T_3 - \theta_3) \left.\frac{\partial^2 \phi}{\partial T_1 \partial T_3}\right|_{\underline{T}=\underline{\theta}}$$
$$+ 2(T_1 - \theta_1)(T_4 - \theta_4) \left.\frac{\partial^2 \phi}{\partial T_1 \partial T_4}\right|_{\underline{T}=\underline{\theta}} + 2(T_1 - \theta_1)(T_5 - \theta_5) \left.\frac{\partial^2 \phi}{\partial T_1 \partial T_5}\right|_{\underline{T}=\underline{\theta}}$$
$$+ 2(T_2 - \theta_2)(T_3 - \theta_3) \left.\frac{\partial^2 \phi}{\partial T_2 \partial T_3}\right|_{\underline{T}=\underline{\theta}} + 2(T_2 - \theta_2)(T_4 - \theta_4) \left.\frac{\partial^2 \phi}{\partial T_2 \partial T_4}\right|_{\underline{T}=\underline{\theta}}$$
$$+ 2(T_2 - \theta_2)(T_5 - \theta_5) \left.\frac{\partial^2 \phi}{\partial T_2 \partial T_5}\right|_{\underline{T}=\underline{\theta}} + 2(T_3 - \theta_3)(T_4 - \theta_4) \left.\frac{\partial^2 \phi}{\partial T_3 \partial T_4}\right|_{\underline{T}=\underline{\theta}}$$
$$\left. + 2(T_3 - \theta_3)(T_5 - \theta_5) \left.\frac{\partial^2 \phi}{\partial T_3 \partial T_5}\right|_{\underline{T}=\underline{\theta}} + 2(T_4 - \theta_4)(T_5 - \theta_5) \left.\frac{\partial^2 \phi}{\partial T_4 \partial T_5}\right|_{\underline{T}=\underline{\theta}} \right] + \cdots \tag{A.1}$$

We now take the expectation of both sides of (A.1) to find the large sample mean of $\phi(\underline{T})$, Therefore, using fact 3, the general result for mean of$\phi(\underline{T})$, up to $O(1/n)$, is

$$E(\phi(\underline{T})) = \phi(\underline{\theta}) + \frac{1}{2} \left[Var(T_1) \left.\frac{\partial^2 \phi}{\partial T_1^2}\right|_{\underline{T}=\underline{\theta}} + Var(T_2) \left.\frac{\partial^2 \phi}{\partial T_2^2}\right|_{\underline{T}=\underline{\theta}} + Var(T_3) \left.\frac{\partial^2 \phi}{\partial T_3^2}\right|_{\underline{T}=\underline{\theta}} \right.$$
$$\left. + Var(T_4) \left.\frac{\partial^2 \phi}{\partial T_4^2}\right|_{\underline{T}=\underline{\theta}} + Var\ (T_5) \left.\frac{\partial^2 \phi}{\partial T_5^2}\right|_{\underline{T}=\underline{\theta}} \right] + Cov(T_1, T_2) \left.\frac{\partial^2 \phi}{\partial T_1 \partial T_2}\right|_{\underline{T}=\underline{\theta}} + Cov(T_3, T_4) \left.\frac{\partial^2 \phi}{\partial T_3 \partial T_4}\right|_{\underline{T}=\underline{\theta}}$$
$$+ Cov\ (T_3, T_5) \left.\frac{\partial^2 \phi}{\partial T_3 \partial T_5}\right|_{\underline{T}=\underline{\theta}} + Cov(T_4, T_5) \left.\frac{\partial^2 \phi}{\partial T_4 \partial T_5}\right|_{\underline{T}=\underline{\theta}}. \tag{A.2}$$

For the variance of $\phi(\underline{T})$, up to $O\left(1/n^2\right)$, we consider $\phi(\underline{T})$ as consisting of three parts: the first part is the linear terms, the second part is the quadratic terms and the last part is the cross-product terms and then compute the variance of each part separately. So, the general result for variance of $\phi(\underline{T})$ is as follow

$$Var(\phi(\underline{T})) = Var(1^{st} part + 2^{nd} part + 3^{rd} part)$$
$$= Var(1^{st} part) + Var(2^{nd} part) + Var(3^{rd} part) + 2Cov(1^{st} part, 2^{nd} part)$$
$$+ 2Cov(1^{st} part, 3^{rd} part) + 2Cov(2^{nd} part, 3^{rd} part).$$

To compute the above variances and covariances, we use some standard properties of a bivariate normal distribution, Helmert's orthogonal transformation and the following theorem.

Theorem (Anderson, 2003) Let $S = \begin{pmatrix} S_x^2 & S_{xy} \\ S_{xy} & S_y^2 \end{pmatrix} \sim W\left(\sum,\ n-1\right)$,where $\sum = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$. Then

(i) $E(S_{ij}) = (n-1)\,\sigma_{ij}$.

(ii) $Cov(S_{ij}, S_{kl}) = (n-1)\,(\sigma_{ik}\sigma_{jl} + \sigma_{il}\sigma_{jk})$.

(iii) $E(SAS) = n(n-1)\,(\sum A \sum) + (n-1)\,(tr \sum A) \sum$, where $A > 0$ a matrix of constants.

The following results then follow immediately.

$$Var(T_1) = Var(\bar{x}) = \frac{\sigma_1^2}{n}, \quad Var(T_2) = Var(\bar{y}) = \frac{\sigma_2^2}{n},$$

$$Var(T_3) = Var\left(\frac{S_x^2}{n-1}\right) = \frac{2\sigma_1^4}{n-1} \approx \frac{2\sigma_1^4}{n},\ Var(T_4) = Var\left(\frac{S_y^2}{n-1}\right) = \frac{2\sigma_2^4}{n-1} \approx \frac{2\sigma_2^4}{n},$$

$$Var(T_5) = Var\left(\frac{S_{xy}}{n-1}\right) = \frac{\sigma_1^2\sigma_2^2(1+\rho^2)}{n-1} \approx \frac{\sigma_1^2\sigma_2^2(1+\rho^2)}{n}$$

$$Cov(T_1, T_2) = Cov(\bar{x}, \bar{y}) = \frac{\rho\sigma_1\sigma_2}{n},$$

$$Cov(T_3, T_4) = Cov\left(\frac{S_x^2}{n-1}, \frac{S_y^2}{n-1}\right) = \frac{2\sigma_1^2\sigma_2^2\rho^2}{n-1} \approx \frac{2\sigma_1^2\sigma_2^2\rho^2}{n},$$

$$Cov(T_3, T_5) = Cov\left(\frac{S_x^2}{n-1}, \frac{S_{xy}}{n-1}\right) = \frac{2\rho\sigma_1^3\sigma_2}{n-1} \approx \frac{2\rho\sigma_1^3\sigma_2}{n},$$

$$Cov(T_4, T_5) = Cov\left(\frac{S_y^2}{n-1}, \frac{S_{xy}}{n-1}\right) = \frac{2\rho\sigma_1\sigma_2^3}{n-1} \approx \frac{2\rho\sigma_1\sigma_2^3}{n},$$

$$Var(T_1-\theta_1)^2 = Var(\bar{x}-\mu_1)^2 = \frac{2\sigma_1^4}{n^2}, \quad Var(T_2-\theta_2)^2 = Var(\bar{y}-\mu_2)^2 = \frac{2\sigma_2^4}{n^2},$$

$$Var(T_3-\theta_3)^2 = Var\left(\frac{S_x^2}{n-1} - \sigma_1^2\right)^2 \approx \frac{8\sigma_1^8}{n^2},$$

$$Var(T_4-\theta_4)^2 = Var\left(\frac{S_y^2}{n-1} - \sigma_2^2\right)^2 \approx \frac{8\sigma_2^8}{n^2},$$

$$Var(T_5-\theta_5)^2 = Var\left(\frac{S_{xy}}{n-1} - \sigma_{12}\right)^2 \approx \frac{2\sigma_1^4\sigma_2^4(1+\rho^2)^2}{n^2},$$

$$Cov\left((T_1-\theta_1)^2, (T_2-\theta_2)^2\right) = Cov\left((\bar{x}-\mu_1)^2, (\bar{y}-\mu_2)^2\right) = \frac{2\rho^2\sigma_1^2\sigma_2^2}{n^2},$$

$$Cov\left((T_3-\theta_3)^2,(T_4-\theta_4)^2\right)=Cov\left(\left(\frac{S_x^2}{n-1}-\sigma_1^2\right)^2,\left(\frac{S_y^2}{n-1}-\sigma_2^2\right)^2\right)\approx\frac{4\rho^4\sigma_1^4\sigma_2^4}{n^2},$$

$$Cov\left((T_3-\theta_3)^2,(T_5-\theta_5)^2\right)=Cov\left(\left(\frac{S_x^2}{n-1}-\sigma_1^2\right)^2,\left(\frac{S_{xy}}{n-1}-\sigma_{12}\right)^2\right)\approx\frac{4\rho^2\sigma_1^6\sigma_2^2}{n^2},$$

$$Cov\left((T_4-\theta_4)^2,(T_5-\theta_5)^2\right)=Cov\left(\left(\frac{S_y^2}{n-1}-\sigma_2^2\right)^2,\left(\frac{S_{xy}}{n-1}-\sigma_{12}\right)^2\right)\approx\frac{4\rho^2\sigma_1^2\sigma_2^6}{n^2},$$

$$Var(T_1-\theta_1)(T_2-\theta_2)=Var(\bar{x}-\mu_1)(\bar{y}-\mu_2)=\frac{\sigma_1^2\sigma_2^2(1+\rho^2)}{n^2},$$

$$Var(T_1-\theta_1)(T_3-\theta_3)=Var(\bar{x}-\mu_1)\left(\frac{S_x^2}{n-1}-\sigma_1^2\right)\approx\frac{2\sigma_1^6}{n^2},$$

$$Var(T_1-\theta_1)(T_4-\theta_4)=Var(\bar{x}-\mu_1)\left(\frac{S_y^2}{n-1}-\sigma_2^2\right)\approx\frac{2\sigma_1^2\sigma_2^4}{n^2},$$

$$Var(T_1-\theta_1)(T_5-\theta_5)=Var(\bar{x}-\mu_1)\left(\frac{S_{xy}}{n-1}-\sigma_{12}\right)\approx\frac{\sigma_1^4\sigma_2^2(1+\rho^2)}{n^2},$$

$$Var(T_2-\theta_2)(T_3-\theta_3)=Var(\bar{y}-\mu_2)\left(\frac{S_x^2}{n-1}-\sigma_1^2\right)\approx\frac{2\sigma_1^4\sigma_2^2}{n^2},$$

$$Var(T_2-\theta_2)(T_4-\theta_4)=Var(\bar{y}-\mu_2)\left(\frac{S_4^2}{n-1}-\sigma_2^2\right)\approx\frac{2\sigma_2^6}{n^2},$$

$$Var(T_2-\theta_2)(T_5-\theta_5)=Var(\bar{y}-\mu_2)\left(\frac{S_{xy}}{n-1}-\sigma_{12}\right)\approx\frac{\sigma_1^2\sigma_2^4(1+\rho^2)}{n^2},$$

$$E\left((T_3-\theta_3)^2(T_4-\theta_4)^2\right)=E\left(\left(\frac{S_x^2}{n-1}-\sigma_1^2\right)^2\left(\frac{S_y^2}{n-1}-\sigma_2^2\right)^2\right)\approx\frac{4\sigma_1^4\sigma_2^4(1+\rho^4)}{n^2},$$

$$Var(T_3-\theta_3)(T_4-\theta_4)=Var\left(\frac{S_x^2}{n-1}-\sigma_1^2\right)\left(\frac{S_y^2}{n-1}-\sigma_2^2\right)\approx\frac{4\sigma_1^4\sigma_2^4}{n^2},$$

$$E\left((T_3-\theta_3)^2(T_5-\theta_5)^2\right)=E\left(\left(\frac{S_x^2}{n-1}-\sigma_1^2\right)^2\left(\frac{S_{xy}}{n-1}-\sigma_{12}\right)^2\right)\approx\frac{2\sigma_1^6\sigma_2^2(1+3\rho^2)}{n^2},$$

$$Var(T_3-\theta_3)(T_5-\theta_5)=Var\left(\frac{S_x^2}{n-1}-\sigma_1^2\right)\left(\frac{S_{xy}}{n-1}-\sigma_{12}\right)\approx\frac{2\sigma_1^6\sigma_2^2(1+\rho^2)}{n^2},$$

$$E\left((T_4-\theta_4)^2(T_5-\theta_5)^2\right)=E\left(\left(\frac{S_y^2}{n-1}-\sigma_1^2\right)^2\left(\frac{S_{xy}}{n-1}-\sigma_{12}\right)^2\right)\approx\frac{2\sigma_1^2\sigma_2^6(1+3\rho^2)}{n^2},$$

$$Var(T_4-\theta_4)(T_5-\theta_5)=Var\left(\frac{S_y^2}{n-1}-\sigma_1^2\right)\left(\frac{S_{xy}}{n-1}-\sigma_{12}\right)\approx\frac{2\sigma_1^2\sigma_2^6(1+\rho^2)}{n^2},$$

$$E\left((T_3-\theta_3)^2(T_4-\theta_4)(T_5-\theta_5)\right)$$

$$= E\left(\left(\frac{S_x^2}{n-1}-\sigma_1^2\right)^2\left(\frac{S_y^2}{n-1}-\sigma_2^2\right)\left(\frac{S_{xy}}{n-1}-\sigma_{12}\right)\right) \approx \frac{4\sigma_1^5\sigma_2^3\rho(1+\rho^2)}{n^2},$$

$$E\Big((T_3-\theta_3)(T_4-\theta_4)^2(T_5-\theta_5)\Big)$$
$$= E\left(\left(\frac{S_x^2}{n-1}-\sigma_1^2\right)\left(\frac{S_y^2}{n-1}-\sigma_2^2\right)^2\left(\frac{S_{xy}}{n-1}-\sigma_{12}\right)\right) \approx \frac{4\sigma_1^3\sigma_2^5\rho(1+\rho^2)}{n^2},$$

$$E\Big((T_3-\theta_3)(T_4-\theta_4)(T_5-\theta_5)^2\Big)$$
$$= E\left(\left(\frac{S_x^2}{n-1}-\sigma_1^2\right)\left(\frac{S_y^2}{n-1}-\sigma_2^2\right)\left(\frac{S_{xy}}{n-1}-\sigma_{12}\right)^2\right) \approx \frac{2\sigma_1^4\sigma_2^2\rho^2(3+\rho^2)}{n^2},$$

$$E(T_3-\theta_3)^3 = E\left(\frac{S_x^2}{n-1}-\sigma_1^2\right)^3 \approx \frac{8\sigma_1^6}{n^2}, \quad E(T_4-\theta_4)^3 = E\left(\frac{S_4^2}{n-1}-\sigma_2^2\right)^3 \approx \frac{8\sigma_2^6}{n^2},$$

$$E(T_5-\theta_5)^3 = E\left(\frac{S_{xy}}{n-1}-\sigma_{12}\right)^3 \approx \frac{2\sigma_1^3\sigma_2^3\rho(3+\rho^2)}{n^2},$$

$$E\Big((T_3-\theta_3)(T_4-\theta_4)^2\Big) = E\left(\left(\frac{S_x^2}{n-1}-\sigma_1^2\right)\left(\frac{S_y^2}{n-1}-\sigma_2^2\right)^2\right) \approx \frac{8\sigma_1^2\sigma_2^4\rho^2}{n^2},$$

$$E\Big((T_3-\theta_3)(T_5-\theta_5)^2\Big) = E\left(\left(\frac{S_x^2}{n-1}-\sigma_1^2\right)\left(\frac{S_{xy}}{n-1}-\sigma_{12}\right)^2\right) \approx \frac{2\sigma_1^4\sigma_2^2(1+3\rho^2)}{n^2},$$

$$E\Big((T_3-\theta_3)^2(T_4-\theta_4)\Big) = E\left(\left(\frac{S_x^2}{n-1}-\sigma_1^2\right)^2\left(\frac{S_y^2}{n-1}-\sigma_2^2\right)\right) \approx \frac{8\sigma_1^4\sigma_2^2\rho^2}{n^2},$$

$$E\Big((T_4-\theta_4)(T_5-\theta_5)^2\Big) = E\left(\left(\frac{S_y^2}{n-1}-\sigma_1^2\right)\left(\frac{S_{xy}}{n-1}-\sigma_{12}\right)^2\right) \approx \frac{2\sigma_1^2\sigma_2^4(1+3\rho^2)}{n^2},$$

$$E\Big((T_3-\theta_3)^2(T_5-\theta_5)\Big) = E\left(\left(\frac{S_x^2}{n-1}-\sigma_1^2\right)^2\left(\frac{S_{xy}}{n-1}-\sigma_{12}\right)\right) \approx \frac{8\sigma_1^5\sigma_2\rho}{n^2},$$

$$E\Big((T_4-\theta_4)^2(T_5-\theta_5)\Big) = E\left(\left(\frac{S_y^2}{n-1}-\sigma_1^2\right)^2\left(\frac{S_{xy}}{n-1}-\sigma_{12}\right)\right) \approx \frac{8\sigma_1\sigma_2^5\rho}{n^2},$$

$$E\left((T_3-\theta_3)(T_4-\theta_4)(T_5-\theta_5)\right)$$
$$= E\left(\left(\frac{S_x^2}{n-1}-\sigma_1^2\right)\left(\frac{S_y^2}{n-1}-\sigma_2^2\right)\left(\frac{S_{xy}}{n-1}-\sigma_{12}\right)\right) \approx \frac{4\sigma_1^3\sigma_2^3\rho(1+\rho^2)}{n^2},$$

$$E\big((T_1-\theta_1)^3(T_2-\theta_2)\big) = E\left((\bar{x}-\mu_1)^3(\bar{y}-\mu_2)\right) = \frac{3\sigma_1^3\sigma_2\rho}{n^2},$$

$$E\big((T_1-\theta_1)(T_2-\theta_2)^3\big) = E\left((\bar{x}-\mu_1)(\bar{y}-\mu_2)^3\right) = \frac{3\sigma_1\sigma_2^3\rho}{n^2},$$

$$E\Big((T_3-\theta_3)^3(T_4-\theta_4)\Big) = E\left(\left(\frac{S_x^2}{n-1}-\sigma_1^2\right)^3\left(\frac{S_y^2}{n-1}-\sigma_2^2\right)\right) \approx \frac{12\sigma_1^6\sigma_2^2\rho^2}{n^2},$$

$$E\left((T_3-\theta_3)^3(T_5-\theta_5)\right)=E\left(\left(\frac{S_x^2}{n-1}-\sigma_1^2\right)^3\left(\frac{S_{xy}}{n-1}-\sigma_{12}\right)\right)\approx\frac{12\sigma_1^7\sigma_2\rho}{n^2},$$

$$E\left((T_4-\theta_4)^3(T_5-\theta_5)\right)=E\left(\left(\frac{S_y^2}{n-1}-\sigma_1^2\right)^3\left(\frac{S_{xy}}{n-1}-\sigma_{12}\right)\right)\approx\frac{12\sigma_1\sigma_2^7\rho}{n^2},$$

$$E\left((T_3-\theta_3)(T_4-\theta_4)^3\right)=E\left(\left(\frac{S_x^2}{n-1}-\sigma_1^2\right)\left(\frac{S_y^2}{n-1}-\sigma_2^2\right)^3\right)\approx\frac{12\sigma_1^2\sigma_2^6\rho^2}{n^2},$$

$$E\left((T_3-\theta_3)(T_5-\theta_5)^3\right)=E\left(\left(\frac{S_x^2}{n-1}-\sigma_1^2\right)\left(\frac{S_{xy}}{n-1}-\sigma_{12}\right)^3\right)\approx\frac{6\sigma_1^5\sigma_2^3\rho(1+\rho^2)}{n^2},$$

$$E\left((T_4-\theta_4)(T_5-\theta_5)^3\right)=E\left(\left(\frac{S_y^2}{n-1}-\sigma_1^2\right)\left(\frac{S_{xy}}{n-1}-\sigma_{12}\right)^3\right)\approx\frac{6\sigma_1^3\sigma_2^5\rho(1+\rho^2)}{n^2}.$$

It therefore follows from the above discussion that for a given $\phi(\underline{T})$, we can eventually compute $E\left(\phi(\underline{T})\right)$ up to $O\left(1/n\right)$ and $Var\left(\phi(\underline{T})\right)$ up to $O\left(1/n^2\right)$. This is precisely done in Section 3 for various choices of $\phi(\underline{T})$ which represent MLEs and MMs of the common mean μ, variances σ_1^2 and σ_2^2, and correlation ρ.