Generalizations of the Wishart Distributions Arising From Monotone Incomplete Multivariate Normal Data

Wan-Ying Chang (Washington Dept. of Fish and Wildlife)

D.R. (Penn State University and SAMSI)

Background

We have a *population* of “patients”

We draw a *random sample* of N patients, and measure m variables on each patient:

1	Visual acuity
2	LDL
3	Systolic blood pressure
4	Glucose intolerance
5	Insulin response to oral glucose
6	Actual weight $\div$ Expected weight
$\vdots$	$\vdots$
m	Red blood cell count

We obtain data:

$$\begin{array}{cccccc} \text{Patient} & 1 & 2 & 3 & \cdots & N \\ & \begin{pmatrix} v_{1,1} \\ v_{1,2} \\ \vdots \\ v_{1,m} \end{pmatrix} & \begin{pmatrix} v_{2,1} \\ v_{2,2} \\ \vdots \\ v_{2,m} \end{pmatrix} & \begin{pmatrix} v_{3,1} \\ v_{3,2} \\ \vdots \\ v_{3,m} \end{pmatrix} & \cdots & \begin{pmatrix} v_{N,1} \\ v_{N,2} \\ \vdots \\ v_{N,m} \end{pmatrix} \end{array}$$

Vector notation for the data: $V_1, V_2, \ldots, V_N$

V_1: The m measurements on patient 1, stacked into a column

Classical multivariate analysis

Statistical analysis of data consisting of N vectors, each containing m entries

Common assumption: The population has a multivariate normal distribution

V: The vector of measurements on a randomly chosen patient

Multivariate normal populations are characterized by:

μ: The population mean vector

Σ: The population covariance matrix

For a given data set, μ and Σ are unknown

We wish to perform inference about μ and Σ

Construct confidence regions for, and test hypotheses about, μ and Σ

Anderson (2003). *An Introduction to Multivariate Statistical Analysis*

Eaton (1984). *Multivariate Statistics: A Vector-Space Approach*

Johnson and Wichern (2002). *Applied Multivariate Statistical Analysis*

Muirhead (1982). *Aspects of Multivariate Statistical Theory*

Standard notation: $V \sim N_p(\mu, \Sigma)$

The usual formula for the density function of V

$$f(v) = (2\pi)^{-m/2} |\Sigma|^{-1/2} \exp\big(- \tfrac{1}{2}(v-\mu)' \Sigma^{-1} (v-\mu) \big), \quad v \in \mathbb{R}^m$$

$V_1, V_2, \ldots, V_N$: Measurements on N randomly chosen patients

Estimate μ and Σ using Fisher's maximum likelihood principle

Likelihood function: $L(\mu, \Sigma) = \prod_{j=1}^{N} f(v_j)$

Maximize the likelihood function L w.r.t. μ and Σ

Maximum likelihood estimator: The value of μ, Σ which maximizes L

$\widehat{\mu} = \frac{1}{N} \sum_{j=1}^{N} V_j$: The sample mean

$\widehat{\Sigma} = \frac{1}{N-1} \sum_{j=1}^{n} (V_j - \bar{V})(V_j - \bar{V})'$: The sample covariance matrix

What are the probability distributions of $\widehat{\mu}$ and $\widehat{\Sigma}$?

$\widehat{\mu} \sim N_p(\mu, \frac{1}{N}\Sigma)$

As $N \to \infty$, $\frac{1}{N}\Sigma \to 0$, so $\widehat{\mu} \to \mu$

$\widehat{\Sigma}$ has a “Wishart” distribution, a generalization of the χ^2

Monotone incomplete data

Some patients were not measured completely (or the dog ate some of the data, or a measuring machine broke down)

The resulting data set, with $*$ denoting missing data

$$\begin{pmatrix} v_{1,1} \\ v_{1,2} \\ v_{1,3} \\ \vdots \\ v_{1,m} \end{pmatrix} \quad \begin{pmatrix} v_{2,1} \\ v_{2,2} \\ v_{2,3} \\ \vdots \\ v_{2,m} \end{pmatrix} \quad \begin{pmatrix} * \\ * \\ v_{3,2} \\ \vdots \\ v_{3,m} \end{pmatrix} \quad \cdots \quad \begin{pmatrix} * \\ * \\ v_{N,2} \\ \vdots \\ v_{N,m} \end{pmatrix}$$

Monotone data: Each $*$ is followed by $*$'s all the way to the end

We may need to renumber patients to see if data are monotone

Physical Fitness Data

Patients: Men taking a physical fitness course at NCSU

Three variables were measured

Oxygen intake rate (ml per kg body weight per minute)

RunTime (time taken, in minutes, to run 1.5 miles)

RunPulse (heart rate while running)

```
Oxygen RunTime RunPulse

44.609   11.37   178     |     39.407   12.63   174
45.313   10.07   185     |     46.080   11.17   156
54.297    8.65   156     |     45.441    9.63   164
51.855   10.33   166     |     54.625    8.92   146
49.156    8.95   180     |     39.442   13.08   174
40.836   10.95   168     |     60.055    8.63   170
44.811   11.63   176     |     37.388   14.03   186
45.681   11.95   176     |     44.754   11.12   176
39.203   12.88   168     |     46.672   10.00     *
45.790   10.47   186     |     46.774   10.25     *
50.545    9.93   148     |     45.118   11.08     *
48.673    9.40   186     |     49.874    9.22     *
47.920   11.50   170     |     49.091   10.85     *
47.467   10.50   170     |     59.571     *       *
50.388   10.08   168     |     50.541     *       *
                         |     47.273     *       *
```

Monotone data have a staircase pattern; we will consider the two-step pattern

Partition V into an incomplete part of dimension p and a complete part of dimension q

$$\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix}, \begin{pmatrix} X_2 \\ Y_2 \end{pmatrix}, \dots \begin{pmatrix} X_n \\ Y_n \end{pmatrix}, \begin{pmatrix} * \\ Y_{n+1} \end{pmatrix}, \begin{pmatrix} * \\ Y_{n+2} \end{pmatrix}, \dots \begin{pmatrix} * \\ Y_N \end{pmatrix}$$

Assume that the individual vectors are independent and are drawn from $N_m(\mu, \Sigma)$

Goal: Carry out maximum likelihood inference for μ and Σ

Obtain formulas as explicit as in the classical context

Where do monotone incomplete data arise?

Panel survey data (Census Bureau, Bureau of Labor Statistics)

Astronomy

Early detection of diseases

Wildlife survey research

Covert communications

Mental health research

Climate and atmospheric studies

$\ldots$

Sometimes we are given n independent observations on $\begin{pmatrix} X \\ Y \end{pmatrix}$ in addition to $N - n$ independent observations on Y

Difficulty: The likelihood function is more complicated

$$\begin{aligned} L &= \prod_{i=1}^{n} f_{X,Y}(x_i, y_i) \cdot \prod_{i=n+1}^{N} f_Y(y_i) \\ &= \prod_{i=1}^{n} f_Y(y_i) f_{X|Y}(x_i) \cdot \prod_{i=n+1}^{N} f_Y(y_i) \\ &= \prod_{i=1}^{N} f_Y(y_i) \cdot \prod_{i=1}^{n} f_{X|Y}(x_i) \end{aligned}$$

Partition μ and Σ similarly:

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

Define

$$\mu_{1\cdot 2} = \mu_1 - \Sigma_{12}\Sigma_{22}^{-1}\mu_2$$

$$\Sigma_{11\cdot 2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$

$$Y \sim N_q(\mu_2, \Sigma_{22}), \quad X|Y \sim N_p(\mu_{1\cdot 2}, \Sigma_{11\cdot 2})$$

Wilks, Anderson, Morrison, Olkin, Jinadasa, D. Tracy, ... : Their hard work produced $\widehat{\mu}$ and $\widehat{\Sigma}$

Sample means:

$$\bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_j, \quad \bar{Y}_1 = \frac{1}{n}\sum_{j=1}^{n} Y_j$$

$$\bar{Y}_2 = \frac{1}{N-n}\sum_{j=n+1}^{N} Y_j, \quad \bar{Y} = \frac{1}{N}\sum_{j=1}^{N} Y_j$$

Sample covariance matrices:

$$A_{11} = \sum_{j=1}^{n}(X_j - \bar{X})(X_j - \bar{X})', \quad A_{12} = \sum_{j=1}^{n}(X_j - \bar{X})(Y_j - \bar{Y}_1)'$$

$$A_{22,n} = \sum_{j=1}^{n}(Y_j - \bar{Y}_1)(Y_j - \bar{Y}_1)', \quad A_{22,N} = \sum_{j=1}^{N}(Y_j - \bar{Y})(Y_j - \bar{Y})'$$

The MLEs of μ and Σ

Notation: $\tau = n/N$, $\bar{\tau} = 1 - \tau$

$$\widehat{\mu}_1 = \bar{X} - \bar{\tau} A_{12} A_{22,n}^{-1} (\bar{Y}_1 - \bar{Y}_2), \qquad \widehat{\mu}_2 = \bar{Y}$$

$\widehat{\mu}_1$ is called the *regression estimator* of μ_1

In sample surveys, extra observations on a subset of variables are used to improve estimation of a parameter

$\widehat{\Sigma}$ is more complicated:

$$\begin{aligned} \widehat{\Sigma}_{11} &= \frac{1}{n}(A_{11} - A_{12} A_{22,n}^{-1} A_{21}) + \frac{1}{N} A_{12} A_{22,n}^{-1} A_{22,N} A_{22,n}^{-1} A_{21} \\ \widehat{\Sigma}_{12} &= \frac{1}{N} A_{12} A_{22,n}^{-1} A_{22,N} \\ \widehat{\Sigma}_{22} &= \frac{1}{N} A_{22,N} \end{aligned}$$

Problems which have been unsolved for up to seventy years:

Find the *exact* distributions of $\widehat{\mu}$ and $\widehat{\Sigma}$

Explicit confidence levels for elliptical confidence regions for μ

In testing hypotheses on μ or Σ, determine whether the likelihood ratio test statistics are unbiased

Describe the moments (means, variances, correlations) of the components of $\widehat{\mu}$

How does $\widehat{\mu}$ behave as n and/or $N \to \infty$?

The crucial obstacle: The distribution of $\widehat{\mu}$ for fixed n, N

The exact distribution of $\widehat{\mu}$

For $n > p + q$,

$$\widehat{\mu} \overset{\mathcal{L}}{=} \mu + V_1 + (n^{-1} - N^{-1}) \sqrt{Q_1/Q_2} \begin{pmatrix} V_2 \\ \mathbf{0} \end{pmatrix}$$

with $V_1 \sim N_{p+q}(\mathbf{0}, \Omega)$, $V_2 \sim N_p(\mathbf{0}, \Sigma_{11\cdot 2})$, $Q_1 \sim \chi^2_q$, $Q_2 \sim \chi^2_{n-q}$,

$$\Omega = N^{-1}\Sigma + (n^{-1} - N^{-1}) \begin{pmatrix} \Sigma_{11\cdot 2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

and V_1, V_2, Q_1, and Q_2 are independent

$n = N$: $\widehat{\mu} \equiv$ the sample mean; it is $N_m(\mu, N^{-1}\Sigma)$-dist'd.

If $\Sigma_{12} = 0$ then $\widehat{\mu}_1$ and $\widehat{\mu}_2$ are independent

$\widehat{\mu}$ is unbiased: $E(\widehat{\mu}) = \mu$

The covariance matrix of $\widehat{\mu}$

$$\mathrm{Cov}(\widehat{\mu}) = \frac{1}{N}\Sigma + \frac{(n-2)\bar{\tau}}{n(n-q-2)} \begin{pmatrix} \Sigma_{11\cdot 2} & 0 \\ 0 & 0 \end{pmatrix}$$

Higher moments of $\widehat{\mu}$ now are straightforward, however ...

Because of the term $1/Q_2$ in the distribution, no even moment of order $n-q$ or higher is finite

The asymptotic distribution of $\widehat{\mu}$

Let $n, N \to \infty$ with $n/N \to \delta$, where $0 < \delta \leq 1$. Then

$$\sqrt{N}(\widehat{\mu} - \mu) \xrightarrow{\mathcal{L}} N_{p+q}\left(0, \Sigma + (\delta^{-1} - 1)\begin{pmatrix} \Sigma_{11\cdot 2} & 0 \\ 0 & 0 \end{pmatrix}\right)$$

Many other asymptotic results can be obtained from the exact distribution of $\widehat{\mu}$

If n and $N \to \infty$ with $n/N \to 0$ then $\widehat{\mu}_2 \to \mu_2$, almost surely, and the distribution of $\widehat{\mu}_1$ is $\ldots$

The analog of Hotelling's statistic:

$$T^2 = (\widehat{\mu} - \mu)' \widehat{\mathrm{Cov}}(\widehat{\mu})^{-1} (\widehat{\mu} - \mu)$$

where

$$\widehat{\mathrm{Cov}}(\widehat{\mu}) = \frac{1}{N}\widehat{\Sigma} + \frac{(n-2)\bar{\tau}}{n(n-q-2)} \begin{pmatrix} \widehat{\Sigma}_{11\cdot 2} & 0 \\ 0 & 0 \end{pmatrix}$$

An obvious ellipsoidal confidence region for μ is

$$\left\{\nu \in \mathbb{R}^{p+q} : (\widehat{\mu} - \nu)' \widehat{\mathrm{Cov}}(\widehat{\mu})^{-1} (\widehat{\mu} - \nu) \leq c\right\}$$

What is the level of confidence associated with this region?

Reminder

$$\bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_j, \quad \bar{Y}_1 = \frac{1}{n}\sum_{j=1}^{n} Y_j$$

$$\bar{Y}_2 = \frac{1}{N-n}\sum_{j=n+1}^{N} Y_j, \quad \bar{Y} = \frac{1}{N}\sum_{j=1}^{N} Y_j$$

$$A_{11} = \sum_{j=1}^{n}(X_j - \bar{X})(X_j - \bar{X})', \quad A_{12} = \sum_{j=1}^{n}(X_j - \bar{X})(Y_j - \bar{Y}_1)'$$

$$A_{22,n} = \sum_{j=1}^{n}(Y_j - \bar{Y}_1)(Y_j - \bar{Y}_1)', \quad A_{22,N} = \sum_{j=1}^{N}(Y_j - \bar{Y})(Y_j - \bar{Y})'$$

$\tau = n/N, \;\; \bar{\tau} = 1 - \tau$

$$\begin{aligned}
\widehat{\Sigma}_{11} &= \frac{1}{n}(A_{11} - A_{12}A_{22,n}^{-1}A_{21}) + \frac{1}{N}A_{12}A_{22,n}^{-1}A_{22,N}A_{22,n}^{-1}A_{21} \\
\widehat{\Sigma}_{12} &= \frac{1}{N}A_{12}A_{22,n}^{-1}A_{22,N} \\
\widehat{\Sigma}_{22} &= \frac{1}{N}A_{22,N}
\end{aligned}$$

Anderson and Olkin (1985) gave an elegant derivation of $\widehat{\Sigma}$

A decomposition of $\widehat{\Sigma}$

Notation: $A_{11\cdot 2,n} := A_{11} - A_{12}A_{22,n}^{-1}A_{21}$

$$\begin{aligned} n\widehat{\Sigma} &= \tau \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22,n} \end{pmatrix} + \bar{\tau} \begin{pmatrix} A_{11\cdot 2,n} & 0 \\ 0 & 0 \end{pmatrix} \\ & \quad + \tau \begin{pmatrix} A_{12}A_{22,n}^{-1} & 0 \\ 0 & I_q \end{pmatrix} \begin{pmatrix} B & B \\ B & B \end{pmatrix} \begin{pmatrix} A_{22,n}^{-1}A_{21} & 0 \\ 0 & I_q \end{pmatrix} \end{aligned}$$

where

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22,n} \end{pmatrix} \sim \mathrm{W}_{p+q}(n-1, \Sigma) \quad \text{and} \quad B \sim \mathrm{W}_q(N-n, \Sigma_{22})$$

are independent. Also, $N\widehat{\Sigma}_{22} \sim \mathrm{W}_q(N-1, \Sigma_{22})$

$$
\begin{aligned}
A_{22,N} \quad &= \quad \sum_{j=1}^{n}(Y_j - \bar{Y}_1 + \bar{Y}_1 - \bar{Y})(Y_j - \bar{Y}_1 + \bar{Y}_1 - \bar{Y})' \\
&\quad + \sum_{j=n+1}^{N}(Y_j - \bar{Y}_2 + \bar{Y}_2 - \bar{Y})(Y_j - \bar{Y}_2 + \bar{Y}_2 - \bar{Y})'
\end{aligned}
$$

$$A_{22,N} = A_{22,n} + B$$

$$B = \sum_{j=n+1}^{N}(Y_j - \bar{Y}_2)(Y_j - \bar{Y}_2)' + \frac{n(N-n)}{N}(\bar{Y}_1 - \bar{Y}_2)(\bar{Y}_1 - \bar{Y}_2)'$$

Verify that the terms in the decomposition of $\widehat{\Sigma}$ are independent

Even the distribution of $\widehat{\Sigma}_{11}$ is non-trivial

If $\Sigma_{12} = 0$ then $A_{22,n}$, B, $A_{11\cdot 2,n}$, $A_{12}A_{22,n}^{-1}A_{21}$, $\bar{X}$, $\bar{Y}_1$, and $\bar{Y}_2$ are independent

Matrix F-distribution: $F_{a,b}^{(q)} = W_2^{-1/2} W_1 W_2^{-1/2}$

where $W_1 \sim \mathrm{W}_q(a, \Sigma_{22})$ and $W_2 \sim \mathrm{W}_q(b, \Sigma_{22})$

Theorem: Suppose that $\Sigma_{12} = 0$. Then

$$\Sigma_{11}^{-1/2}\widehat{\Sigma}_{11}\Sigma_{11}^{-1/2} \overset{\mathcal{L}}{=} \frac{1}{n}W_1 + \frac{1}{N}W_2^{1/2}\big(I_p + F\big)W_2^{1/2}$$

where $W_1 \sim \mathrm{W}_p(n-q-1, I_p)$, $W_2 \sim \mathrm{W}_p(q, I_p)$, $F \sim F^{(p)}_{N-n,n-q+p-1}$, and W_1, W_2, and F are independent

$$N\Sigma_{11}^{-1/2}\widehat{\Sigma}_{11}\Sigma_{11}^{-1/2} \overset{\mathcal{L}}{=} \frac{N}{n}\Sigma_{11}^{-1/2}A_{11\cdot 2,n}\Sigma_{11}^{-1/2}$$
$$+ \Sigma_{11}^{-1/2}A_{12}A_{22,n}^{-1}(A_{22,n} + B)A_{22,n}^{-1}A_{21}\Sigma_{11}^{-1/2}$$

With no assumptions on Σ_{12}:

$$\widehat{\Sigma}_{12}\widehat{\Sigma}_{22}^{-1} \overset{\mathcal{L}}{=} \Sigma_{12}\Sigma_{22}^{-1} + \Sigma_{11\cdot 2}^{1/2} W^{-1/2} K \Sigma_{22}^{-1/2}$$

where $W \sim \mathrm{W}_p(n-q+p-1, I_p)$, $K \sim N_{pq}(0, I_p \otimes I_q)$, and W and K are independent

In particular, $\widehat{\Sigma}_{12}\widehat{\Sigma}_{22}^{-1}$ is an unbiased estimator of $\Sigma_{12}\Sigma_{22}^{-1}$

Define

$$\Delta_{11} = \Sigma_{11\cdot 2}, \quad \Delta_{12} = \Sigma_{12}\Sigma_{22}^{-1}, \quad \Delta_{22} = \Sigma_{22}$$

The matrix LDU decomposition of any positive definite matrix

$$\Sigma = \begin{pmatrix} I_p & \Delta_{12} \\ 0 & I_q \end{pmatrix} \begin{pmatrix} \Delta_{11} & 0 \\ 0 & \Delta_{22} \end{pmatrix} \begin{pmatrix} I_p & 0 \\ \Delta_{21} & I_q \end{pmatrix}$$

There is a 1:1 correspondence between Σ and

$$\Delta = \begin{pmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{21} & \Delta_{22} \end{pmatrix}$$

The inverse transformation from Δ to Σ:

$$\begin{aligned} \Sigma_{11} &= \Delta_{11} + \Delta_{12}\Delta_{22}\Delta_{21} \\ \Sigma_{12} &= \Delta_{12}\Delta_{22} \\ \Sigma_{22} &= \Delta_{22} \end{aligned}$$

Starting with $\widehat{\Sigma}$, we construct $\widehat{\Delta}$ exactly as above

A hypergeometric function of matrix argument

Herz (1955), Muirhead (1982)

Multivariate gamma function: For $a > (q-1)/2$,

$$\Gamma_q(a) = \pi^{q(q-1)/4} \prod_{j=1}^{q} \Gamma(a - \tfrac{1}{2}(j-1))$$

Multivariate beta function: For $a, b > (q-1)/2$,

$$B_q(a,b) = \frac{\Gamma_q(a)\Gamma_q(b)}{\Gamma_q(a+b)},$$

M is $q \times q$ and symmetric; $a, b-a > (q-1)/2$

$$
\begin{aligned}
&{}_1F_1^{(q)}(a;b;M) \\
&= \frac{1}{B_q(a,b-a)} \int_{0<u<I_q} |u|^{a-\frac{1}{2}(q+1)} \, |I_q - u|^{b-a-\frac{1}{2}(q+1)} \, e^{\text{tr}\, Mu} \, du,
\end{aligned}
$$

${}_1F_1(KMK') = {}_1F_1(M)$ for all $q \times q$ orthogonal matrices K

Conclude: ${}_1F_1(M)$ depends only on the eigenvalues of M

Herz (1955): If rank$(M) = r$ then ${}_1F_1^{(q)}(a;b;M) = {}_1F_1^{(r)}(a;b;M_0)$ where M_0 is any $r \times r$ matrix whose non-zero eigenvalues coincide with those of M

The distribution of $\widehat{\Delta}$

Start with

$$\begin{aligned}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22,n} \end{pmatrix} &= \sum_{j=1}^{n} \begin{pmatrix} X_j - \bar{X} \\ Y_j - \bar{Y}_1 \end{pmatrix} \begin{pmatrix} X_j - \bar{X} \\ Y_j - \bar{Y}_1 \end{pmatrix}' \\ &\sim \mathrm{W}_{p+q}(n-1, \Sigma)\end{aligned}$$

A well-known property of the Wishart distribution:

$A_{11\cdot 2,n}$ and $\{A_{12}, A_{22,n}\}$ are mutually independent

Conclude: $A_{11\cdot 2,n}$ and $\{A_{12}, A_{22,N}\}$ are mutually independent

Express the $\widehat{\Delta}$'s in terms of the A's

$$n\widehat{\Delta}_{11} = A_{11\cdot 2,n} \sim \mathrm{W}_p(n-q-1, \Delta_{11})$$

$$N\widehat{\Delta}_{22} = A_{22,N} \sim \mathrm{W}_q(N-1, \Delta_{22})$$

$$\widehat{\Delta}_{12} = A_{12}A_{22,n}^{-1}$$

$$\widehat{\Delta}_{12}|A_{22,n} = A_{12}A_{22,n}^{-1}|A_{22,n} \sim N_{pq}(\Delta_{12}, \Delta_{11} \otimes A_{22,n}^{-1})$$

We shall evaluate the p.d.f. of $\widehat{\Delta}$ at

$$T = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix}$$

$\widehat{\Delta}_{11}$ and $\{\widehat{\Delta}_{12}, \widehat{\Delta}_{22}\}$ are independent

All that remains to be done is to evaluate the conditional p.d.f. of $\widehat{\Delta}_{12}|\widehat{\Delta}_{22}$ at T_{12}

Ξ_1, Ξ_2, and Ξ_3: Random matrices of the same dimension such that (Ξ_1, Ξ_2) and Ξ_3 are independent. Then the conditional p.d.f. of Ξ_1 given $\Xi_2 + \Xi_3 = \xi$, is

$$\begin{aligned} & f_{\Xi_1|\Xi_2+\Xi_3=\xi}(\xi_1) \\ & \qquad = \frac{1}{f_{\Xi_2+\Xi_3}(\xi)} \int f_{\Xi_1|\Xi_2=\xi_2}(\xi_1)\, f_{\Xi_2}(\xi_2)\, f_{\Xi_3}(\xi - \xi_2)\, d\xi_2 \end{aligned}$$

Set $\Xi_1 = A_{12}A_{22,n}^{-1} \equiv \widehat{\Delta}_{12}$, $\Xi_2 = N^{-1}A_{22,n}$, $\Xi_3 = N^{-1}B$

$\Xi_2 + \Xi_3 = N^{-1}(A_{22,n} + B) \equiv \widehat{\Delta}_{22}$

The lemma gives us the distribution of $\widehat{\Delta}_{12}|\widehat{\Delta}_{22}$

Substitute the densities of $\Xi_1|\Xi_2$ and Ξ_2 into the integral and evaluate:

$$
\begin{aligned}
&f_{\widehat{\Delta}_{12}|\widehat{\Delta}_{22}=T_{22}}(T_{12}) \\
&= c\,|\Delta_{11}|^{-q/2}|\Delta_{22}|^{-(N-1)/2} \\
&\times \exp(-\tfrac{1}{2}\text{tr}\,N\Delta_{22}^{-1}T_{22}) \\
&\times |T_{22}|^{\frac{1}{2}(N+p-1)-\frac{1}{2}(q+1)} \\
&\times {}_1F_1^{(q)}\left(\begin{matrix}\frac{1}{2}(n+p-1) \\ \frac{1}{2}(N+p-1)\end{matrix}; -\tfrac{1}{2}N(T_{12}-\Delta_{12})'\Delta_{11}^{-1}(T_{12}-\Delta_{12})T_{22}\right)
\end{aligned}
$$

Conclude:

$$f_{\widehat{\Delta}}(T) = f_{\widehat{\Delta}_{11}}(T_{11})\, f_{\widehat{\Delta}_{22}}(T_{22})\, f_{\widehat{\Delta}_{12}|\widehat{\Delta}_{22}=T_{22}}(T_{12})$$

To transform back to $\widehat{\Sigma}$, the Jacobians are

$$J(\widehat{\Delta}_{11} \to \widehat{\Sigma}_{11}) J(\widehat{\Delta}_{12} \to \widehat{\Sigma}_{12}) J(\widehat{\Delta}_{22} \to \widehat{\Sigma}_{22}) = 1 \cdot |\widehat{\Sigma}_{22}^{-1}|^{p} \cdot 1$$

The density function of $\widehat{\Sigma}$ is

$$f_{\widehat{\Sigma}}(T) = f_{\widehat{\Delta}_{11},\widehat{\Delta}_{12},\widehat{\Delta}_{22}}(T_{11} - T_{12}T_{22}^{-1}T_{21}, T_{12}T_{22}^{-1}, T_{22})\; |T_{22}|^{-p}$$

What can we do with, or obtain from, this expression for the p.d.f.?

Saddlepoint approximations: Apply Butler-Wood formulas

Can we let $q \to \infty$? Free probability ...

$\widehat{\Delta}_{11}$ and $\widehat{\Delta}_{22}$ are Wishart matrices, so we know their eigenvalue distributions

We can apply interlacing theorems to bound the distribution of the eigenvalues of $\widehat{\Delta}$

Open problem: Find the distribution of the eigenvalues of $\widehat{\Sigma}$

The distribution of $\widehat{\Sigma}$ is non-trivial but the distribution of $|\widehat{\Sigma}|$ is simple

$$|\widehat{\Sigma}| = |\widehat{\Sigma}_{11\cdot 2}| \cdot |\widehat{\Sigma}_{22}| = |\widehat{\Delta}_{11}| \cdot |\widehat{\Delta}_{22}|$$

These two are independent and each is a product of χ^2 random variables

Hao and Krishnamoorthy (2001): The sample generalized variance $|\widehat{\Sigma}|$ is distributed as

$$n^{-p} N^{-q} |\Sigma| \cdot \chi^2_{n-q-1} \chi^2_{n-q-2} \cdots \chi^2_{n-q-1-p} \chi^2_{N-1} \chi^2_{N-2} \cdots \chi^2_{N-q}.$$

This result raises the possibility of good results in hypothesis testing on Σ

Testing that $\Sigma = \Sigma_0$

Sample: Same as before, 2-step monotone incomplete

Σ_0: A given, positive definite matrix

Test $H_0 : \Sigma = \Sigma_0$ against $H_a : \Sigma \neq \Sigma_0$

Hao and Krishnamoorthy (2001): The LRT statistic for testing H_0 against H_a is

$$\begin{aligned}\lambda_1 \quad &\propto \quad |A_{22,N}|^{N/2} \exp\Big(-\tfrac{1}{2}\mathrm{tr}\, A_{22,N}\Big) \\ &\qquad \times |A_{11\cdot 2,n}|^{n/2} \exp\Big(-\tfrac{1}{2}\mathrm{tr}\, A_{11\cdot 2,n}\Big) \\ &\qquad \times \exp\Big(-\tfrac{1}{2}\mathrm{tr}\, A_{12}A_{22,n}^{-1}A_{21}\Big)\Big).\end{aligned}$$

Is the LRT unbiased? If C is a critical region of size α, is

$$P(\lambda_1 \in C|H_a) \geq P(\lambda_1 \in C|H_0)?$$

In the case of complete data, it is well-known that the LRT is not unbiased

E. J. G. Pitman: λ_1 becomes unbiased if the sample sizes are replaced by the corresponding degrees of freedom

With monotone incomplete data, perhaps a similarly modified statistic, λ_2, is unbiased?

Answer: Not always; λ_2 is unbiased if $|\Sigma_{11}| < 1$

With monotone incomplete data, a further modification is necessary: The modified LRT

$$\begin{aligned}\lambda_3 \quad \propto \quad & |A_{22,N}|^{(N-1)/2} \exp\big(-\tfrac{1}{2}\text{tr}\, A_{22,N}\big) \\ & \times |A_{11\cdot 2,n}|^{(n-q-1)/2} \exp\big(-\tfrac{1}{2}\text{tr}\, A_{11\cdot 2,n}\big) \\ & \times |A_{12}A_{22,n}^{-1}A_{21}|^{q/2} \exp\big(-\tfrac{1}{2}\text{tr}\, A_{12}A_{22,n}^{-1}A_{21}\big),\end{aligned}$$

is unbiased. Also, λ_1 is never unbiased for 2-step monotone data.

For diagonal $\Sigma = \mathrm{diag}(\sigma_{jj})$, the power function of λ_3 increases monotonically as any $|\sigma_{jj} - 1|$ increases, $j = 1, \ldots, p+q$.

μ_0 and Σ_0 are completely specified

With monotone 2-step data, test

$H_0 : (\mu, \Sigma) = (\mu_0, \Sigma_0)$ vs. $H_a : (\mu, \Sigma) \neq (\mu_0, \Sigma_0)$

The LRT is

$$\lambda_4 = \lambda_1 \exp\big(-\tfrac{1}{2}(n\bar{X}'\bar{X} + N\bar{Y}'\bar{Y})\big)$$

Remarkably, λ_4 is unbiased

The sphericity test, $H_0 : \Sigma \propto I_{p+q}$

The LRT statistic is known but whether or not it is unbiased remains an open question