vrije Universiteit *amsterdam*

School of Business and Economics

Exam:	Multivariate Statistics
Code:	E_EOR2_MS
Examinator:	Andre Lucas
Co-reader:	Julia Schaumburg
Date:	March 26, 2019
Time:	12:00
Duration:	2 hours
Calculator allowed:	Yes
Graphical calculator allowed:	Yes
Number of questions:	4
Type of questions:	Open
Answer in:	English

Remarks:

Credit score:	See the credits for each separate question
Grades:	The grades will be made public on: April 9, 2019.
Inspection:	Wednesday, April 10, 2019 at 10am. Room tba via CANVAS.
Number of pages:	... (including front page)

Good luck!

Question 1. [35points]

You have a normally distributed random vector $X \sim N(\mu, \Sigma)$, with $X = (X_1, X_2, X_3, X_4)^\top$, and

$$\mu = (1, 3, 0, *)^\top,$$

$$\Sigma = \begin{pmatrix} 7 & 4 & -2 & * \\ 4 & 6 & -1 & * \\ -2 & -1 & 3 & * \\ * & * & * & * \end{pmatrix}.$$

A) What is your best guess of the value of $(X_1, X_2)^\top$ if you are given $X_3 = 2$? And what is the precision (=covariance matrix) corresponding to that? [10points]

$$\begin{pmatrix}1\\3\end{pmatrix} + \begin{pmatrix}-2\\-1\end{pmatrix}\frac{2}{3} = \begin{pmatrix}-\frac{1}{3}\\ 2\frac{1}{3}\end{pmatrix}.\ \text{5pt}$$

$$\begin{pmatrix}7 & 4\\4 & 6\end{pmatrix} - \frac{1}{3}\begin{pmatrix}-2\\-1\end{pmatrix}\begin{pmatrix}-2\\-1\end{pmatrix}^\top = \begin{pmatrix}5\frac{2}{3} & 3\frac{1}{3}\\ 3\frac{1}{3} & 5\frac{2}{3}\end{pmatrix}.\ \text{5pt}$$

B) If you know that

$$\left(\begin{pmatrix}X_3\\X_4\end{pmatrix}\middle|\begin{pmatrix}X_1\\X_2\end{pmatrix}\right) \sim N\left(\begin{pmatrix}\frac{-8X_1 + X_2 + 5}{26}\\ \frac{2X_1 + 3X_2 - 37}{26}\end{pmatrix}, \frac{1}{26}\begin{pmatrix}11 & 7\\7 & 8\end{pmatrix}\right)$$

Compute the remaining elements of μ and Σ. [15pt]

$$\mathcal{A} = \frac{1}{26}\begin{pmatrix}-8 & 1\\2 & 3\end{pmatrix}, \qquad a = \frac{1}{26}\begin{pmatrix}5\\-37\end{pmatrix}, \quad \left(\begin{pmatrix}X_3\\X_4\end{pmatrix}\middle|\begin{pmatrix}X_1\\X_2\end{pmatrix}\right) \sim N\left(a + \mathcal{A}\begin{pmatrix}X_1\\X_2\end{pmatrix}, \frac{1}{26}\begin{pmatrix}11 & 7\\7 & 8\end{pmatrix}\right).$$

You can now use Th5.4 and obtain for the mean elements:

$$a + \mathcal{A}\begin{pmatrix}1\\3\end{pmatrix} = \begin{pmatrix}0\\-1\end{pmatrix}.\ \text{5pt}$$

For the covariance elements (lower-left block):

$$\mathcal{A}\Sigma_{1:2,1:2} = \frac{1}{26}\begin{pmatrix}-8 & 1\\2 & 3\end{pmatrix}\begin{pmatrix}7 & 4\\4 & 6\end{pmatrix} = \begin{pmatrix}-2 & -1\\1 & 1\end{pmatrix}.\ \text{5pt}$$

The covariance lower-right block

$$\frac{1}{26}\begin{pmatrix}11 & 7\\7 & 8\end{pmatrix} + \frac{1}{26}\begin{pmatrix}-2 & -1\\1 & 1\end{pmatrix}\begin{pmatrix}-8 & 2\\1 & 3\end{pmatrix} = \frac{1}{26}\begin{pmatrix}26 & 0\\0 & 13\end{pmatrix} = \begin{pmatrix}1 & 0\\0 & 0.5\end{pmatrix}.\ \text{5pt}$$

You study the Trampoline distribution, which is given by a *mean-zero* normal random variable X with covariance matrix Σ, divided by an independent random variable U with pdf $f_U(u) = 1 - |u|$ for $|u| < 1$, and zero else, so $Y = X/U$.

C) Derive that the distribution of Y is given by

$$f_Y(y) = 2\int_0^1 \frac{\exp\left(-\frac{z^2}{2}y^\top\Sigma^{-1}y\right)}{|2\pi\Sigma|^{\frac{1}{2}}} \cdot (z^p - z^{p+1})\, dz,$$

[hint: use $(Y,Z) = (X/U, U)$]. [10pt]

No domain change. Jacobian of the transformation $(Y,Z) = (X/U,\ U) \leftrightarrow (X,U) = (YZ,Z)$ is $abs(Z^p) = abs(U^p)$. 5pt

$$f_Y(y) = \int_{-1}^{1} f_X(\mu + y \cdot z\, ; \mu, \Sigma) \cdot (1 - |z|)\, abs(z^p) dz$$

$= 2\int_0^1 \frac{\exp\left(-\frac{z^2}{2} y^\top \Sigma^{-1} y\right)}{|2\pi\Sigma|^{\frac{1}{2}}} (1-z) z^p dz$ 5pt.

Question 2. [30pt]

You model financial return data, and you want that the (co)variances of the returns to exist.
Your model consists of Student's t marginals with 5 and 8 degrees of freedom respectively, and a Student's t copula with only 1 degree of freedom.
Your friend warns you that this is a flawed model, because the covariances between the two returns does not exist in this case.

A) Argue whether your friend is right or wrong. [10pt]
[hint: you may use the inequality $E[|X_1 X_2|] \leq \sqrt{E[X_1^2] E[X_2^2]}$, such that a covariance exists if the variances exist for two random variables X_1 and X_2].

Your friend is wrong. The marginal behavior of the variables is t(5) and t(8), so their variances exist, so their covariance exists. [10pt]

Consider two random variables X_1 and X_2 with *pdf* $f_{X_1}(x_1) = 2x_1$ for $0 < x_1 < 1$, zero else, and $f_{X_2}(x_2) = \exp(-x_2)$ for $x_2 > 0$, zero else.

B) If the joint *cdf* is given by $F(x_1, x_2) = \exp\left[-\sqrt{(-\log(x_1^2))^2 + (-\log(1 - e^{-x_2}))^2}\right]$, then what is the corresponding copula? [10pt]
[hint: first compute the PITs]

The PITs are $u_1 = x_1^2$ and $u_2 = 1 - \exp(-x_2)$, so $F(x_1, x_2) = \exp[-\sqrt{(-\log u_1)^2 + (-\log u_2)^2}]$, which via Sklar's theorem must therefore be the copula.

Using the same pdfs as under B), and a copula *density* $c(u_1, u_2) = 2u_1^2 u_2^2 / (u_1 + u_2 - u_1 u_2)^3$.
C) Compute the density $f_X(0.5, 0.5)$. [10pt]

$$u_1 = 0.25, u_2 = 0.39,$$

$$(2 \cdot 0.5) \cdot (\exp(-0.5)) \cdot 2 \cdot 0.25^2 \cdot \frac{0.39^2}{(0.64 - 0.25 \cdot 0.39)^3} = 0.0725$$

Question 3. [20pt]

You estimated

$$\bar{x} = (1\ \ 3\ \ 5\ 0)^{\top},$$

$$\mathcal{S} = \begin{pmatrix} 7 & 4 & -2 & -2 \\ 4 & 6 & -1 & 0 \\ -2 & -1 & 3 & 1 \\ -2 & 0 & 1 & 5 \end{pmatrix}$$

You hypothesize that the means of X_1 and X_3 are equal, *and* that the average mean of the first two variables (X_1, X_2) is the *twice* the average mean of the last two variables (X_3, X_4).

A) Formulate this as a testing problem using matrix notation. [5pt]

$$\mathcal{A} = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 2 & 2 & -1 & -1 \end{pmatrix}$$
$$H0: \mathcal{A}\mu = 0$$

B) Formulate the corresponding test statistic [you do not need to compute the value(s) numerically]. [5pt]

$$(n-1)(\mathcal{A}\bar{x})^{\top}(\mathcal{A}\mathcal{S}\mathcal{A}^{\top})^{-1}(\mathcal{A}\bar{x})$$

Assume the numerical value of your test statistic is 5.48.
You are given the following table.

χ^2 critical values for ν degrees of freedom	5%
$\nu = 1$	3.84
$\nu = 2$	5.99
$\nu = 3$	7.81
$\nu = 4$	9.49

C) At a 5% significance level, what do you conclude on your hypothesis? Explain your answer. [5pt]

We have 2 restrictions, so we should look at 2 degrees of freedom. We do not exceed the critical value, so we do *not* reject.

You proceed with your analysis, and ask yourself whether variable X_4 is uncorrelated with (X_1, X_2, X_3). You compute the log-likelihoods under the null and the alternative and obtain the values -1238.1 and -1234.6, respectively.

D) Can reject your null hypothesis or not at the 5% significance level? Explain your answer. [5pt]

The likelihood ratio test in this case is 2*(-1234.6+1238.1)=7, with 3 restrictions (3 covariances are zero) and thus 3 degrees of freedom, so *NOT reject* the null hypothesis.

Question 4. [15pt]

You consider 6 stock returns from two different regions that you try to model using an orthogonal factor model structure with 2 factors.
The factors are

(i) a factor for the general market conditions; all stocks load on this factor and have the same loading λ_0;

(ii) a factor for the regional developments in the second region only; only the 3 stocks from region 2 load on the second (regional) factor, all with the same loading λ_1.

A) Formulate the orthogonal factor model for this data. Carefully specify the factor loading matrix satisfying conditions (i) and (ii) above, and the covariance matrix of all the random variables used. [10pt]

$X = \begin{pmatrix} \lambda_0 & 0 \\ \lambda_0 & 0 \\ \lambda_0 & 0 \\ \lambda_0 & \lambda_1 \\ \lambda_0 & \lambda_1 \\ \lambda_0 & \lambda_1 \end{pmatrix} F + E$, where F is bivariate standard normal, and $E = (E_1, \ldots, E_6)^\top$ has a diagonal covariance matrix and a multivariate normal distribution.

You estimate a standard orthogonal factor model and obtain an estimated loadings matrix $\hat{B}$. The matrix $\hat{B}$ does not look at all like what you expect theoretically. In particular, it does not have the zeroes for the stocks from region 1 with respect to the regional factor for region 2.

B) Abstracting from estimation error, explain why your estimates $\hat{B}$ can be so far off from what you expect theoretically. [5pt]

The loadings are unique up to rotation. You may therefore have found a different rotation, e.g. due to a different normalization during estimation.