Started on	Tuesday, 25 October 2022, 2:04 PM
State	Finished
Completed on	Monday, 31 October 2022, 4:11 PM
Time taken	6 days 2 hours
Marks	38.00/41.00
Grade	**9.27** out of 10.00 (**92.68**%)

Information

Declaration

The University regards plagiarism as a form of academic misconduct, and has very strict rules regarding plagiarism. See UNSW policies, penalties, and information to help you avoid plagiarism, as well as the guidelines in the online ELISE tutorials for all new UNSW students.

By submitting this assessment item, you declare that this assessment item is your own work, except where acknowledged, and acknowledge that the assessor of this item may, for the purpose of assessing this item:

1. Reproduce this assessment item and provide a copy to another member of the University; and/or
2. Communicate a copy of this assessment item to a plagiarism checking service (which may then retain a copy of the assessment item on its database for the purpose of future plagiarism checking)

You certify that you have read and understood the University Rules with respect to Student Conduct.

Information

Software

You may use any software of your choice to help you in this assessment. Use of R and computer algebra systems is permitted and encouraged. You are not required to derive exact answers where numerical answers are requested.

However, any code you use to get your results must be submitted as a part of your working.

Warning

For technical reasons, there may be problems marking your submission if not all answers are filled in; if you do not know how to answer a question, simply put in 0.

Level of significance

It is recommended that you do not round intermediate calculations at all. For final answers, usually four decimal places suffice, but you won't lose marks for being more precise.

Question 1

Correct

Mark 7.00 out of 7.00

Consider a joint density $f(x_1, x_2) = \frac{1}{2x_2}\exp[-(\frac{x_2}{4} + \frac{2x_1}{x_2})]$, for $x_1, x_2 > 0$.

For the following, enter the answers as algebraic expressions in terms of x1, x2, and/or, basic operations such as +, -, *, /, and ** (for power), common functions such as `exp`, `log`, and `cos`. Do **not** use parentheses other than `()` or implicit multiplication. Do **not** specify the domain/range of the the functions. For example, 3^2 would be expressed as `3**2`, and the joint density above would be expressed as

```
1/(2*x2) * exp(-(x2/4 + 2*x1/x2))
```

(a) Derive $f_{X_2}(x_2)$: (1/4) * exp(-x2/4)

✓

One possible correct answer is: exp(-x2/4)/4

(b) Derive $f_{X_1|X_2}(x_1|x_2)$: 2 * exp(-2*x1/x2) / x2

✓

One possible correct answer is: 2*exp(-2*x1/x2)/x2

(c) Give the best approximation $g^*(X_2)$ in mean square sense for X_1 (i.e. find explicitly $g^*(X_2)$ that minimises $\mathrm{E}[(X_1 - g(X_2))^2|X_2 = x_2]$ over all possible choices of $g(X_2)$ such that $\mathrm{E}[g(X_2)^2]$ is finite). x2 / 2

✓

One possible correct answer is: x2/2

(d) For a given realisation x_2, calculate the mean square error of the best approximation. (That is, calculate $\mathrm{E}\{[X_1 - g^*(X_2)]^2|X_2 = x_2\}$.) (x2)**2 / 4

✓

One possible correct answer is: x2**2/2**2

Your answer is correct.

Back to Top

Consider the matrix $\Sigma = \begin{bmatrix} 8 & 3 \\ 3 & 9 \end{bmatrix}$.

(a) Its characteristic polynomial (in terms of `lambda`) is: lambda**2 - 17*lambda + 6

One possible correct answer is: (8-lambda)*(9-lambda)-3**2

(b) The eigenvalues of this matrix (in decreasing order) are: 11.5414 and 5.4586 .

The corresponding normalised eigenvectors of this matrix are: [0.6464 , 0.763] and [-0.763 , 0.6464].

One possible correct answer is: 11.541381265149, 5.4586187348509, 0.6463748961302, 0.76301998247273, -0.76301998247273, 0.6463748961302

(c) Calculate $\Sigma^{1/2}$:

2.7796	0.5232
0.5232	2.9540

One possible correct answer is: 2.7796099804065, 0.52322878057753, 0.52322878057753, 2.9540195739323

(d) Calculate $\Sigma^{-1/2}$:

0.3722	-0.0659
-0.0659	0.3502

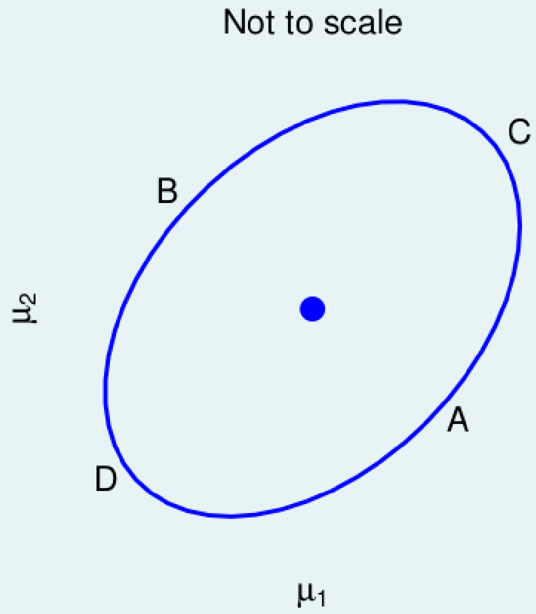
One possible correct answer is: 0.37217148384009, -0.065920630104749, -0.065920630104749, 0.35019794047184

Your answer is correct.

A random sample of 12 observations yielded a sample mean [-1.2, -3.9] and a sample variance-covariance matrix $S = \begin{bmatrix} 6 & 1 \\ 1 & 6 \end{bmatrix}$.

Assume that the population is multivariate normal.

(a) Consider a confidence ellipse for the mean vector at level 90%. What are the coordinates of points A–D, representing the "southeast", the "northwest", the "northeast", and the "southwest" of the ellipse respectively, depicted below?



	μ_1	μ_2
A (southeast)	-0.0423	-5.0577
B (northwest)	-2.3577	-2.7423
C (northeast)	0.1699	-2.5301
D (southwest)	-2.5699	-5.2699

✔

One possible correct answer is: -0.042252085847966, -5.057747914152, -2.357747914152, -2.742252085848, 0.16986580576812, -2.5301341942319, -2.5698658057681, -5.2698658057681

(b) Now, suppose that S above is actually the value of *known* population covariance matrix Σ. What are the coordinates of points A–D now?

	μ_1	μ_2
A (southeast)	2.1931	-7.2931
B (northwest)	-4.5931	-0.5069
C (northeast)	2.8147	0.1147
D (southwest)	-5.2147	-7.9147

✘

One possible correct answer is: -0.22050499980133, -4.8794950001987, -2.1794950001987, -2.9205049998013, -0.041045886335503, -2.7410458863355, -2.3589541136645, -5.0589541136645

(c) Under the original premise (i.e., S a sample covariance matrix, and Σ is unknown), use the projection method to obtain simultaneous 90% confidence intervals for the population means.

	Lower	Upper
μ_1	-2.9936	0.5936

(d) Under the original premise (i.e., S a sample covariance matrix, and Σ is unknown), use the Bonferroni method to obtain simultaneous 90% confidence intervals for the population means.

	Lower	Upper
μ_1	-2.7563	0.3563
μ_2	-5.4563	-2.3437

One possible correct answer is: -2.7563315319918, 0.35633153199176, -5.4563315319918, -2.3436684680082

Your answer is partially correct.

You have correctly answered 3 part(s) of this question.

Information

Pittsburgh, a city in western Pennsylvania, USA, was built at the intersection of three rivers: Monongahela, Allegheny, and Ohio. This dataset contains information about historical bridges crossing these three rivers.

river
Which river does it cross? Allegheny, Monogahela, or Ohio?

erected
year it was built

purpose
walking, aqueduct, railroad, or highway

length
length in feet

material
iron, wood, or steel

They have been provided in the dataset `bridges.csv`. (Click on the file to download.)

Question 4

Correct

Mark 4.00 out of 4.00

(a) Produce pairwise scatterplots, and perform the normality tests for variables year built (erected) and length. Which of the following quantitative measurements are consistent with normality?

- Year built: Yes
- Length: No
- Multivariate skewness: No
- Multivariate kurtosis: No

(b) Regardless of your answer to the previous question, assume that normality is satisfied. Use a T^2 test to test (at the conventional $\alpha = 0.05$) whether the population means of year built and length differ (jointly) between bridges built over Monogahela and bridges built over Allegheny:

- The T^2 test statistic:

  1.1960

- Degrees of freedom for the F distribution:

  2

  and

  66

- *P*-value:

  0.5577

- Conclude:

  A. There is sufficient evidence to conclude that population means of year built and length differ between the bridges over Monongahela and Allegheny.

  B. There is not sufficient evidence to conclude that population means of year built and length differ between the bridges over Monongahela and Allegheny.

  C. There is sufficient evidence to conclude that population means of year built and length do not differ between the bridges over Monongahela and Allegheny.

  D. There is not sufficient evidence to conclude that populations means of year built and length do not differ between the bridges over Monongahela and Allegheny.

  B

Information

Unless otherwise stated, multiple choice and True/False questions in the following section have a *guessing penalty*. This means that a correct answer gains you 1 mark, but an incorrect answer *loses* you around $1/(\#\text{ choices} - 1)$ marks. This means that if guessing randomly, your *expected* mark will be 0.

Do **not** let it prevent you from answering questions, because unless you are completely randomly guessing, you are far more likely to gain marks by answering than to lose them.

True or False? The joint distribution of two normal random variables is always multivariate normal.

Select one:

- True
- False ✓

Your answer is correct.

The correct answer is: False

Information

Universal or Normal only?

Let $\boldsymbol{X} \in \mathbb{R}^p$ be a random vector with $\mathrm{E}(\boldsymbol{X}) = \boldsymbol{\mu}$ and $\mathrm{Var}(\boldsymbol{X}) = \Sigma$, and let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ be an independent sample from the same distribution as $\boldsymbol{X}$. Which of the following are only guaranteed to be true if $\boldsymbol{X}$ has a multivariate normal distribution? (That is, they *may* still be true if $\boldsymbol{X}$ is not normal for some specific cases, but are not *guaranteed* to be.)

Question 7

Correct

Mark 0.50 out of 0.50

If $\boldsymbol{X}$ is written $\begin{pmatrix} X_{(1)} \\ \boldsymbol{X}_{(2)} \end{pmatrix}$ with $X_{(1)} \in \mathbb{R}^1$ and $\boldsymbol{X}_{(2)} \in \mathbb{R}^{p-1}$ and we wish to predict $X_{(1)}$ from $\boldsymbol{X}_{(2)}$ (i.e., we want a function $g(\boldsymbol{x}_{(2)})$ that predicts $X_{(1)}$ as well as possible), then the function that minimises the squared loss (i.e., $\mathrm{E}_{X_{(1)}}[(X_{(1)} - g(\boldsymbol{X}_{(2)}))^2 | \boldsymbol{X}_{(2)} = \boldsymbol{x}_{(2)}]$) is $g^*(\boldsymbol{x}_{(2)}) = \mathrm{E}_{X_{(1)}}[X_{(1)} | \boldsymbol{X}_{(2)} = \boldsymbol{x}_{(2)}]$.

Select one:

- Always ✓
- Normal

Your answer is correct.

The correct answer is: Always

Given the sample $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$, the maximum likelihood estimator (MLE) of ρ_{ij} (the correlation between X_i and X_j) is $\hat{\sigma}_{ij}/\sqrt{\hat{\sigma}_{ii}\hat{\sigma}_{jj}}$, where $\hat{\sigma}_{kl}$, $k, l = 1, \ldots, p$ are elements of matrix $\hat{\boldsymbol{\Sigma}}$, the MLE of Σ.

Select one:

- Always
- Normal ✖

Your answer is incorrect.

The correct answer is: Always

Question 9

Correct

Mark 0.50 out of 0.50

For a fixed matrix $A \in \mathcal{M}_{q,p}$, $\mathrm{Var}(A\boldsymbol{X}) = A\Sigma A^\top$.

Select one:

- Always ✔
- Normal

Your answer is correct.

The correct answer is: Always

Question 10

Correct

Mark 0.50 out of 0.50

If X_i and X_j are independent, then $\Sigma_{ij} = 0$.

Select one:

- Always ✔
- Normal

Your answer is correct.

The correct answer is: Always

Previous Activity

Next Activity