

MOCK Final Exam
Exam Duration: 2 Hours
Total Marks Available: 100

Instructions

1. Except for the **Base R** cheat sheet, this is a closed-book exam.
2. All answers and explanations must be provided in the answer book.
3. Keep your answers succinct and to the point. Long rambling answers with irrelevant details may work against you.

Advice: Read the entire exam before starting on your answers.

Question 1. (18 marks)

Answer each of the following questions.

- (a) (5 marks) Let \mathbf{A} be an $m \times n$ matrix with column vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ and row vectors $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m$. Let $\mathbf{x} = (x_1 \dots x_m)^\top$. Give an expression for $\mathbf{A}^\top \mathbf{x}$ in terms of the column or row vectors.
- (b) (5 marks) Suppose $\mathbf{A} \in \mathbb{R}^{n \times n}$ and $\text{rank}(\mathbf{A}) = n$. Is there always a solution to the equation $\mathbf{A}^\top \mathbf{x} = \mathbf{b}$ for any $\mathbf{b} \in \mathbb{R}^n$? Justify your answer.
- (c) (8 marks) A projection matrix is a matrix $\mathbf{P} \in \mathbb{R}^{n \times n}$ such that $\mathbf{P}^2 = \mathbf{P}$. What are the possible eigenvalues of \mathbf{P} ? Justify your answer.
-

Question 2. (12 marks)

Consider the matrix

$$\mathbf{A} = \begin{pmatrix} 13 & 0 & -15 \\ -3 & 4 & 9 \\ 5 & 0 & -7 \end{pmatrix}$$

and answer the following questions.

- (a) (4 marks) Use R to compute and then report the eigenvalues of \mathbf{A} ?
- (b) (3 marks) True or false: \mathbf{B} is a symmetric matrix where $\mathbf{B} := \mathbf{A}^\top \mathbf{A}$.
Note: Please use the properties from the module to get full marks. If you arrive at the correct answer using R, you receive 2 marks.
- (c) (5 marks) Use R to compute and then report (to two decimal places) the eigenvalues of \mathbf{B} ?
-

Question 3. (10 marks)

A group of 20 people go out to dinner. 10 go to an Italian restaurant, 6 to a Chinese restaurant and 4 to a French restaurant. The fractions of people satisfied with their meals were $4/5$, $2/3$ and $1/2$ respectively. The next day the person you are talking to (one of the group of 20) was satisfied with their meal. What is the probability that they went to the Italian restaurant?

Question 4. (20 marks)

Suppose that the lifespan of a piece of equipment is normally distributed with mean 10 years and standard deviation 3 years. A particular piece of equipment is now aged 11 years. The firm is considering two options:

- Replace it now;
- Wait until age 13 to replace it (unless it dies before then, in which case an unscheduled replacement is necessary).

Suppose that an immediate replacement will cost the company $\mathcal{L}a$. Replacement at age 13 – if the equipment lasts that long – will cost $\mathcal{L}b$, where $b < a$. An unscheduled replacement between now and age 13 – if the equipment dies suddenly during the next two years – will cost $\mathcal{L}c$, where $c > a$.

- (8 marks) What is the probability that, barring replacement now, the equipment will last until age 13?
- (8 marks) What is the expected cost of replacement if the equipment is not replaced now?
- (4 marks) Under what condition on a , b and c is the expected cost lower if the equipment is replaced now than under the age 13 policy?

Question 5. (25 marks)

As a production manager, you are investigating the processing times of machines in the assembly line. In the production of the most essential product of your company, there is a specific machine of high importance. The processing time of the machine follows a normal distribution with an assumed standard deviation of 10 minutes according to the chief production engineer. A test run of the machine produces 100 items with an averaging processing time of 90 minutes per item. Answer the following questions.

- (5 marks) Give a 95% confidence interval for the mean processing time.
- (5 marks) Is it true that with 95% probability the true (population) mean processing time lies in the interval from part (a)?
- (5 marks) How many items should the test run produce if we want the half width of the confidence interval (i.e., the margin of error) to be no more than 1 minute for a 95% confidence interval?
- (10 marks) Would you agree with the engineer's further claim that the average processing time is 88 minutes at a $\alpha = 0.05$ significance level?

Question 6. (15 marks)

Data for 51 US states (50 states plus the District of Columbia) was used to examine the relationship between violent crime rate (violent crimes per 100,000 persons per year) and the predictor variables of urbanization (percentage of the population living in urban areas) and poverty rate. A predictor variable indicating whether or not a state is classified as a Southern state (1 = Southern, 0 = not) was also included. Some R output for the linear regression analysis is shown below.

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-321.90	148.20	-2.17	0.035 .
urban	4.689	1.654	2.83	0.007 *
poverty	39.34	13.52	2.91	0.006 *
south(s=1)	-649.30	266.96	-2.43	0.019 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

- (a) **(5 marks)** What is the estimated linear model?
- (b) **(5 marks)** How to interpret the estimated coefficient for south, -649.30 ?
- (c) **(5 marks)** From the model summary, which predictor variables are significant at 0.01 level? Explain your reasoning.