Ministerie van Defensie

Department of Military Sciences

Student data	
Name:	
Peoplesoft number:	
Class:	
Signature:	

General			
Course:	Probability & Statistics (Resit)	**Course code:**	P&S
Date:	August 25, 2025	**Time:**	10:00-13:00
Examiner:	Dr. ir. D.A.M.P. Blom	**Number of Pages:**	5
Number of Questions	5	**Total Points:**	100

General instructions
- All answers must be supported by a clear explanation. Answers such as “yes” or “no” without justification will receive no credit. - Round final answers to four decimal places, where applicable. - If you are unable to solve a subquestion, you are encouraged to make a reasonable assumption and proceed. Partial credit may still be awarded for correct methodology, even if intermediate answers are incorrect. - The use of a graphical calculator without a CAS (Computer Algebra System) is permitted. - No exam-related material may be taken out of the examination room. - Please write your name and PeopleSoft number on each page and number all pages of your answers (e.g., 1/5, 2/5, etc. if you hand in five answer sheets). - The use of electronic devices capable of sending, receiving or storing information (e.g., mobile phones, smartwatches) is strictly prohibited. These must be left outside the exam room or handed in to the examiner, switched off or in airplane mode. - Ensure your handwriting is legible. Illegible or unclear answers will not be graded. - Toilet visits are only allowed with prior permission from the examiner. - Upon leaving the examination room, all materials (exam paper, scrap paper, formula sheets) must be handed in to the examiner.

Grading

- The final grade of the Probability and Statistics course is entirely based on this exam.
- The exam consists of five open-ended questions, each with subquestions.
- The number of points available for each (sub)question is indicated in brackets. A total of 100 points can be earned.
- Your final grade will be calculated by dividing the total points earned by 10.
- A minimum final grade of 5.5 is required to pass the course.

Procedure after the exam

- Exam results will be published within ten working dates after the exam date.
- If you have questions about the grading, you may contact the course coordinator within ten working days after the results have been released.

Good luck!

Problem 1 (20 points) A coastal rescue team is conducting search operations for missing swimmers lost at sea. Based on previous data and ocean current modeling, the joint probability density function (PDF) for the coordinates (X, Y) of a swimmer's location (in kilometers relative to the shore) is given by:

$$f_{X,Y}(x,y) = \begin{cases} C \cdot (3x + 2y), & \text{if } 0 \leq x \leq 2 \text{ and } 0 \leq y \leq 3 \\ 0, & \text{otherwise,} \end{cases}$$

where C is a constant.

1a [6pt] Determine the value of C so that $f_{X,Y}(x,y)$ satisfies the properties of a valid joint probability density function.

1b [6pt] Compute the marginal PDFs $f_X(x)$ and $f_Y(y)$, for X and Y respectively.

1c [8pt] Are X and Y independent random variables? Justify your answer with a computation.

Problem 2 (23 points) During CBRN readiness training, military standards require that personnel are able to don full protective gear in no more than 6.0 minutes. A commander collects data from a random sample of twelve soldiers. The sample mean donning time is 6.3 minutes. Assume donning times are normally distributed with a population standard deviation of 0.4 minutes.

2a [5pt] Construct a $95\,\%$-confidence interval for the true mean donning time μ. Interpret the result in context.

2b [4pt] What is the minimum sample size required to estimate the true mean donning time μ within ± 0.1 minutes, with $95\,\%$ confidence?

We would like to test if the true mean donning time is indeed up to military standards. For that purpose, a hypothesis test needs to be conducted. From the following questions onwards, we assume that the **sample** standard deviation (based on the twelve soldiers) is equal to 0.4 minutes, and the true standard deviation σ is unknown. Use a significance level $\alpha = 0,05$.

2c [4pt] State the null and alternative hypothesis of the hypothesis test. Explain your choice on the type of hypothesis test (left-sided, two-sided or right-sided).

2d [10pt] Perform the hypothesis test and state your conclusion based on the critical region.

Problem 3 (15 points) A set of n independent and identically distributed random variables $X_1, X_2, \dots, X_n$ are drawn from a distribution with the probability density function:

$$f(x;\theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, \quad x \geq 0, \quad \theta > 0$$

3a [3pt] Write down the likelihood function $L(x_1, x_2, \dots, x_n; \theta)$ given a sample of realizations of $X_1, X_2, \dots, X_n$.

3b [7pt] Derive the maximum likelihood estimator (MLE) of the parameter θ.

Hint: consider the natural logarithm (ln) of the likelihood function (the log-likelihood).

3c [5pt] Show that the sample mean $\overline{X} = \frac{(X_1+X_2+\dots+X_n)}{n}$ is an unbiased estimator for the parameter θ.

Problem 4 (20 points) An air force unit is testing a new type of radar used for detecting enemy drones. The manufacturer, Thales, claims that the radar has a $70\,\%$ success rate in detecting a drone. To test this claim, 1000 independent trials are conducted. In each trial, the radar attempts to detect 4 drones. The number of drones detected per trial is recorded as follows:

Number of detected drones	Observed frequencies
0	15
1	105
2	290
3	360
4	230
Total:	1000

To test the claim with the manufacturer, we would like to test whether the observed data indeed fits a binomial distribution using a chi-square goodness-of-fit test.

4a [2pt] Which values for the parameters n and p should we use for our hypothesis test. Explain your answer.

4b [3pt] State the null hypothesis H_0 and the alternative hypothesis H_1 of the hypothesis test. What does rejecting H_0 tell us about the distribution of the number of detected drones?

4c [3pt] Calculate the expected frequencies under the null hypothesis H_0.

4d [9pt] Perform the hypothesis test at a significance level $\alpha = 0,05$ by computing the p-value.

4e [3pt] Interpret the result of the hypothesis test using data from the tables of observed and expected frequencies.

Problem 5 (22 points) In a joint project between physical training instructors of the Royal Military Academy (KMA) and the medical service, the relationship is investigated between average sleep duration of a cadet (in hours) and the recovery time after intense physical exertion (in hours).

Recovery time is measured using a muscle soreness index; a cadet is considered fully recovered when this index falls below a certain threshold.

For nine cadets, the average sleep duration and recovery time after intense exertion are measured.

Sleep duration (in hours)	$4,0$	$4,5$	$5,0$	$5,5$	$6,0$	$6,5$	$7,0$	$7,5$	$8,0$
Recovery time (in hours)	66	61	63	62	65	64	57	59	60

5a [2pt] If we want to perform a regression analysis, which variable would be the dependent variable Y and which variable the independent variable X?

5b [5pt] Draw the corresponding scatter plot based on your answer to subquestion (a).

5c [8pt] Calculate Pearson's correlation coefficient $r(x, y)$. What can you conclude about the relationship between the two variables?

5d [7pt] Calculate the regression line $Y = a + b \cdot X$ by computing the coefficients a and b. Based on this regression line, give a statistically sound prediction for the recovery time of a cadet who has slept for 6 hours and 45 minutes.