Ministerie van Defensie

Faculty of Military Sciences

Student data	
Name:	
Peoplesoft number:	
Class:	
Signature:	

General			
Course:	Probability & Statistics (resit)	**Course code:**	P&S
Date:	4 juli 2025	**Time:**	10:00-13:00
Examinator:	Dr. ir. D.A.M.P. Blom	**Number of pages:**	6
Number of questions	5	**Number of points:**	80

General instructions

- All answers need to be motivated. If an answer “yes” or “no” is given without any textual explanation, no points will be awarded.
- Round your final answers to four decimals (if necessary)
- If you cannot solve a certain subquestion, please continue working with a reasonable estimate of the answer. In most cases, you can still get partial points if your computations are correct although working with incorrect intermediate answers.
- You are allowed to use a graphical calculator without a CAS (Computer Algebra System).
- Answers, in any form whatsoever, cannot leave the room.
- Please indicate your name, Peoplesoft number on each page and make a page numbering of your answer sheets (e.g., 1/5, 2/5, 3/5, 4/5 and 5/5)
- Each form of electronic devices capable of sending / receiving / carrying information (telephone, smartwatch, etc.) is not allowed during the entire duration of the exam. These devices need to be left outside of the exam room or be handed in to the examinator while turned off / in airplane mode.
- Please write legibly to avoid confusion in the grading of your work. If your answer is unclear or illegible, you will not be awarded any points for your answer.
- Visiting the toilet during the exam is permitted only after explicit permission of the examinator.
- Hand in all exam-related items (exam questions, scratch papers, formula sheets) at the examinator when leaving the exam room.

Grading

- The final grade of the course Probability and Statistics is fully determined by this exam.
- The exam consists of five open questions, with subquestions.
- For each (sub)question, the number of points that can be obtained are given between brackets. In total, you can earn 80 points.
- The final grade will be determined by dividing the number of awarded points by 8.
- The final grade must be at least 5.5 to successfully complete the course Probability and Statistics.

Procedure after the exam

- The grades of this exam will be published within ten working dates after the examination.
- When you have questions regarding the grading, you can contact the course coordinator within 10 working days after publication of the grades.

Good luck!

Problem 1 (15 points) During routine patrols in the North Sea, a helicopter unit detects multiple Russian vessels suspected of attempting to damage undersea communication cables. The helicopter record two discrete variables hourly:

- X: the number of radar contacts detected per hour (values: $1, 2, 3$)
- Y: the number of confirmed hostile aircraft identified per hour (values: $1, 2, 3$)

The joint probability distribution of X and Y reflects the likelihood of these events occurring simultaneously during patrol hours.

X Y	1	2	3
1	0.20	0.12	0.08
2	0.05	0.18	0.10
3	0.07	0.03	0.17

1a [5pt] Calculate the marginal distributions of X and Y, as well as $E[X]$, Var(X), $E[Y]$ and Var(Y).

1b [3pt] Calculate the covariance Cov(X, Y).

1c [5pt] Calculate the following probabilities:

- $P(X = 3, Y \leq 2)$
- $P(X = 3 \mid Y \leq 1)$
- $P(Y \leq 1 \mid X \leq 2)$

1d [2pt] Are X and Y independent? Explain your answer

Problem 2 (15 points) In a military exercise, measurements are conducted regarding the time a certain communication signal needs to reach a control post. These times $X_1, X_2, \ldots, X_n$ are assumed to be independent and identically distributed according to a uniform distribution over the interval $[\theta; \theta + 1]$, where θ is an unknown real parameter.

2a [5pt] Write down the likelihood function $L(x_1, x_2, \ldots, x_n; \theta)$ for the parameter θ in terms of a sample of realizations $x_1, x_2, \ldots, x_n$.

2b [5pt] Derive the maximum likelihood estimator (MLE) of the parameter θ.

2c [5pt] Show that the estimator $\overline{X} - \frac{1}{2} = \frac{(X_1 + X_2 + \ldots + X_n)}{n} - \frac{1}{2}$ is an unbiased estimator for the parameter θ.

Problem 3 (15 points) A naval research unit investigates the signal strength (in decibels, dB) of sonar pulses reflected from a newly designed stealth submarine hull.

Under standard conditions, the mean reflected signal strength from a conventional hull is $\mu_0 = 65$ dB, with unknown standard deviation σ dB.

To assess whether the new stealth design reduces detectability, a test series is conducted. Signal strength is measured in 10 independent trials with the following results (in dB):

$$\{63.2,\ 64.1,\ 62.5,\ 63.8,\ 65.0,\ 64.7,\ 63.5,\ 62.9,\ 63.6,\ 64.2\}$$

Assume the measurements follow a normal distribution. Use a significance level of $\alpha = 0.05$.

3a [4pt] Determine the mean and standard deviation of this sample.

3b [3pt] State the null and alternative hypotheses, and explain the direction of the test.

3c [9pt] Perform the hypothesis test for the mean reflected signal strength μ based on the critical region.

3d [4pt] Calculate the probability of a Type-II error β if the true reflected signal strength is actually normally distributed with $\mu = 64.5$ en $\sigma = 0.8$ dB (for a single observation).

Problem 4 (20 points) Over a six-month period, NATO cyber experts monitor cyberattacks targeting Estonia. The following five types of attacks are recorded: phishing, malware injection, DDoS (Distributed Denial-of-Service), brute-force login attempts and supply chain exploits. The Estonian team observed 500 attacks in total. Furthermore, the expected distribution of cyberattack types is assessed based on global intelligence reports:

Attack type	**Observed frequencies**	**Expected proportion**
Phishing	90	20%
Malware injection	160	30%
DDoS	120	25%
Brute-force	70	15%
Supply chain exploits	60	10%

The cyber team would like to test whether the observed distribution of the types of incoming cyber attacks differs significantly from the expected distribution

4a [3pt] Which kind of hypothesis test do we need to perform. State the null hypothesis H_0 and the alternative hypothesis H_1 of this test.

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

4c [9pt] Perform the hypothesis test at a significance level $\alpha = 0.05$, and compute the p-value.

4d [5pt] State a conclusion for the hypothesis test in the original context of the problem and interpret it using data from the table.

Problem 5 (15 points) In a test for the endurance of soldiers, research is conducted on the relationship between the load weight carried by the soldier (X) and the time needed for completing a 3 km speed march (Y). The following data were collected on 12 soldiers:

X	12	14	16	18	20	22	24	26	28	30	15	25
Y	16.75	16.29	17.97	19.78	17.65	18.15	21.37	20.65	19.3	21.31	16.05	18.55

The research team believes that there is a linear relationship between the weight of the vehicle and the deployment time.

5a [6pt] Calculate the least-squares estimates $\hat{\beta}_0$ and $\hat{\beta}_1$ for the slope and intercept of the linear regression line $Y = \beta_0 + \beta_1 X$, where Y is the deployment time, and X is the vehicle weight.

5b [4pt] Interpret the slope β_1 of the regression model in the context of this exercise. What does the slope suggest about the relationship between load weight and speed march time?

5c [5pt] Calculate the correlation coefficient $R(X, Y)$ between the load weight and speed march time. Based on the value of the correlation coefficient, what can you conclude about the strength and direction of the relationship between the two variables?