Name and Surname:
Pseudonym:

1	2	3	4	5	6

3A. Exam Paper NMAI059 Probability and Statistics 1 – 12.6.2024

Write the problem number and your surname on each sheet of paper.

You can also write the selected pseudonym on this paper, under which your results will be published. (Otherwise, they will be published with your initials.) **Submit the assignment as well (it will be available on the web).**

Do not write more problems on the same page!

You have **150 minutes** to complete the work.

No calculators, phones, etc. are allowed during the work. (Please turn off the sound on your phones in advance.)

If the result contains expressions that are difficult to compute without a calculator, do not calculate them: $137 \cdot 173$ is as good, if not better, an answer than 23701.

Provide detailed justification for all calculations.

You can use one (handwritten) cheat sheet in A4 format.

After grading the test, a grade of 1, …, 5 will be suggested to everyone via email. You can improve this grade by one level during the oral part – i.e., a 4 can be improved to 3, but a 5 means failure in this exam term. The oral part of the exam can take place preferably on Friday morning.

You may find the following table of quantile functions useful

p	0.9	0.95	0.975	0.99	0.995
$\Phi^{-1}(p)$	1.28	1.64	1.96	2.33	2.58
$\Psi_1^{-1}(p)$	3.08	6.31	12.71	31.82	63.66
$\Psi_2^{-1}(p)$	1.89	2.92	4.3	6.96	9.92
$\Psi_9^{-1}(p)$	1.38	1.83	2.26	2.82	3.25
$\Psi_{10}^{-1}(p)$	1.37	1.81	2.23	2.76	3.17
$\Psi_{11}^{-1}(p)$	1.36	1.8	2.2	2.72	3.11
$\Psi_{12}^{-1}(p)$	1.36	1.78	2.18	2.68	3.05
$\Psi_{13}^{-1}(p)$	1.35	1.77	2.16	2.65	3.01

Provide detailed justification for all calculations!

1. (10 points)

$x \backslash y$	1	2	3
0	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{12}$
1	$\frac{1}{8}$	$\frac{1}{12}$	$\frac{1}{24}$
2	$\frac{1}{8}$	$\frac{1}{12}$	$\frac{1}{24}$

The two-dimensional random vector (X, Y) has the probabilities of individual values given by the table on the right.

(a) Calculate $P(X < Y)$.

(b) Determine the marginal prob. functions p_X and p_Y.

(c) Are the random variables X and Y dependent or independent?

(d) Determine $\mathbb{E}(X + Y)$.

(e) Determine $\mathbb{E}(XY)$.

(f) Determine $\mathbb{E}(\frac{X}{Y})$.

2. (10 points)

In the Wild West, there is a gunfight between Rapid Richard and Precise Paul. Richard starts and has a probability of 0.5 to hit Paul (and the duel ends). Then Paul shoots – he has a probability of 0.8 to hit Richard (and the duel ends). The duel continues until someone is hit.

(a) What is the probability that Richard wins?

(b) What is the expected number of shots Richard fires?

3. (10 points) The measured values 1.5, 2, 2.8 are assumed to come from a random sample from the distribution $Exp(\lambda)$. (This is the time a web server waits for the next query.) That is, $P(T > t) = e^{-\lambda t}$.

(a) Propose a point estimate of the parameter λ using the method of moments.

(b) Propose a point estimate of the parameter λ using the maximum likelihood method.

4. (10 points) (a) Define the concept of the density of a random variable.

Decide whether a random variable X has the density given by the function $f(x) = c/x^3$ for $x > 1$ (and $f(x) = 0$ otherwise) for some $c \in \mathbb{R}$. If so, determine $\mathbb{E}(X)$ and $\mathrm{var}(X)$.

(b) Define the concept of conditional mean of a discrete random variable.

Let X_1, X_2 be the results of two rolls of a standard die (i.e., numbers 1, …, 6), denote $X = X_1 X_2$. Compute $\mathbb{E}(X \mid X_2 > 3)$.

5. (10 points) Explain the principle of hypothesis testing. Clarify, among other things, the following terms: Type 1 and Type 2 errors, significance level, power of the test, critical region, one-sample, two-sample, and paired tests.

Illustrate the concepts with the following example: A software engineer wants to test whether a new algorithm increases sorting efficiency compared to the existing algorithm. He collects a sample of 10 datasets and records the average sorting time of each dataset using both algorithms. He finds that the average reduction in run time is 2.5 ms with a standard deviation of 1 ms. Use the table on the previous page if necessary.

6. (10 points) State and prove the law of total expectation. Define the terms used.