FIRST COURSE IN PROBABILITY AND STATISTICS (MS-A0501/2/3/4)
Department of Mathematics and Systems Analysis, Aalto University
Jukka Kohonen, Pekka Pere, Natalia Vesselinova

Exam 21.02.2024

The exam starts at 9.00 h and ends at 12.00 h.

You may use any kind of calculator in the exam without an Internet connection (mobile phones and tablet computers are not allowed). You may use your handwritten double-sided A4 note in the exam. **Present the intermediate steps of all your calculations, and justify all your answers in detail. A correct answer alone is worth zero points. All tasks are worth six points. Carefully review all your answers.** And once more, the best of luck!

QUESTION 1

A math exam at a university consists of written and oral questions. Only students who have passed the written exam are invited to a face-to-face interview (the oral part of the exam) by the teacher. At the interview, with equal probability the teacher asks the interviewed student no question at all, a question or at most two questions related to their written exam. The interviewed student passes the exam only if they correctly answer all (one or two) questions posed by the teacher or if the teacher does not ask any questions. If the student does not correctly answer the first of two questions (assuming the lecturer decides to ask two), they fail the exam (the lecturer does not continue with the second question).

a) Find the probability that a student will pass the oral part (which would also mean passing the exam), if the probability for correctly answering a question is π. (Hint: it might be helpful to draw the sequential process – a tree with branches – to visually depict the different possibilities.)

b) What is the probability to fail the exam when the teacher asks two questions and the student correctly answers the first of them?

QUESTION 2

In an animation a ball with eyes(!) pushed another ball, a green cube watched(!), entered, and stopped the aggression of the ball with eyes. In another animation a red cube looked at only or did not stop the aggression. Six-month-old infants ($n = 20$) watched the animations. Next the infants were let to choose a green or red cube. 17 out of 20 infants chose the green cube. Researches infered that already six-months-olds admire agents acting protectively.[1]

Let us assume that six-month-olds choose a green or a red cube with equal probability (0.5). An experiment is organised in which 20 six-month-olds are let to choose between a green and red cube independently.

a) What is the probability that 17 out of the 20 six-months-olds choose a green cube?

b) What is the probability that at least 17 out of the 20 six-months-olds choose a green cube?

[1] Y. Kanakogi, Y. Inoue, G. Matsuda, D. Butler, K. Hiraki ja M. Myowa–Yamakoshi (2017): Preverbal Infants Affirm Third-Party Interventions That Protect Victims from Aggressors. *Nature. Human Behaviour.* doi:10.1038/s41562-016-0037.

QUESTION 3

In a certain population, 1/100 of the population carries a particular disease. There is a test that has two possible results, "positive" and "negative". If a person with the disease is tested, the test will be positive with probability 0.95. If a healthy person is tested, the test will be positive with probability 0.02.

We select a random person from the population and let D be the event that the person has the disease, and T the event that the person tests positive. The complement events are D^c and T^c, respectively.

a) Calculate the following five quantities, each with at least four decimals:

$$P(T \mid D), \quad P(T \cap D), \quad P(D \mid T), \quad P(D^c \mid T), \quad \text{and} \quad P(D \mid T^c).$$

b) Which of the previous quantities expresses the probability that a randomly selected person obtains a positive test result?

c) If a randomly selected person obtains a positive test result, which of the quantities expresses the probability that the person has the disease?

d) Suppose that the prevalence of the disease increases from 1/100 to 1/10, but the test technology remains the same. For each of the previous five quantities, in order, express whether it will increase (+), decrease (−) or remain unchanged (0). For example, "00000" would mean that they all remain unchanged. Here it is enough to express the direction of the change (numerical answers not required).

QUESTION 4

Do people feel differently about people who Became Rich or who were Born Rich? Koo *et al.* (2023) asked 289 research participants to evaluate how much they like people in the respective groups.[2] The answers were given as integer numbers 1–7 (so-called Likert-scale). Integers 1, 4, and 7 corresponded to attitudes "do not like them at all", "moderately like them" (neutral evaluation), and "like them a lot", respectively. The evaluation was hence the more positive the larger the integer. The sample mean and sample standard deviation of the evaluations were 4.98 and 1.30 for the Became Rich people and 3.60 and 1.55 for the Born Rich people.

The question concerns the evaluations about the Became Rich people. You can assume that the usual assumptions for carrying out a statistical test apply (evaluations are independent, identically distributed *etc.*) and for simplicity that the standard deviation of the evaluation is know to be 1.30.

a) Formulate a test statistic for the null hypothesis that the mean evaluation is 4, and calculate the numerical value of the statistic.

b) Test the null hypothesis at significance level 0.001 (two-sided test). Explain carefully all stages of your test. (Hint: The 0.99-, 0.995-, 0.999-, 0.9995- and 0.9999-quantiles of the standard normal distribution are 2.326, 2.576, 3.090, 3.291, and 3.719, respectively.)

QUESTION 5 When a certain coin is tossed, the results "heads" (1) and

[2] H.J. Koo, P.K. Piff, and A.F. Shariff (2023): If I Could Do It, So Can They: Among the Rich, Those With Humbler Origins are Less Sensitive to the Difficulties of the Poor. *Social Psychological and Personality Science*, 14, 333–341.

“tails” (0) have probabilities Θ and $1 - \Theta$, respectively. The parameter Θ has prior distribution with continuous density

$$f_\Theta(\theta) = 30\theta^2(1-\theta)^2$$

for $0 \le \theta \le 1$, and zero otherwise. The coin is tossed eight times and the results are $1, 0, 0, 0, 1, 0, 0, 0$.

a) Calculate the maximum likelihood (ML) estimate for Θ. Express the result as a fraction, or a decimal number with at least three decimals.

b) Write an unnormalized density function for Θ, based on the prior distribution and the observations. (“Unnormalized” means that your function can have a constant coefficient whose value you need not determine.)

c) Using the unnormalized density function from (b), find the maximum a posteriori (MAP) estimate for Θ. Express the result as a fraction, or a decimal number with at least three decimals.

You must obtain the results by working with the functions and finding the points where they obtain their maxima (it is not enough just to recognize the distribution and appeal to its known properties).