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

## Renewal examination 7.6.2024

You may use a scientifiç calculator namely, a calculator that can have operations for trigonometry, power, exponential function, logarithm, and binomial coefficients. However, no other type of calculator is allowed. In particular, you may not use: 1) programmable calculators, which can run program code; 2) symbolic calculators, which can manipulate symbolic expressions; 3) graphical calculators, which are able to plot functions; 4) calculators with Internet connection. You may use your handwritten double-sided A4 note at 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!**

1. A patient is depressed. A medical doctor can prescribe medication or psychoanalysis or both for treatment.

a) Let us assume that psychoanalysis cures depression with probability 0.6, medication cures depression with probability 0.3, and that psychoanalysis and medication act independently.[1] What is the probability that the patient recovers from depression if the doctor prescribes both psychoanalysis and medication as treatments?

b) Let us relax the assumptions that the two treatments act independently. Let us assume that depression is cured by

- psychoanalysis alone with probability 0.6.
- medication alone with probability 0.3.
- medication and psychoanalysis in tandem with probability 0.8.

A medical doctor prescribes treatment after careful individual examination of what she sees fit for the patient. Depending on the patient, she decides that the appropriate treatment is

- psychoanalysis only with probability 0.05.
- medication only with probability 0.90.
- both psychoanalysis and medication with probability 0.05.

What is the probability that the patient recovers from depression by the prescribed treatment(s)?

c) Let us add the possibility that the patient recovers from depression with probability 0.2 independently from whether he gets treatment or not. What

[1] Probability 0.6 is taken from P. Knekt, M.A. Laaksonen, T. Härkänen, T. Maljanen, E. Heinonen, E. Virtala, and O. Lindfors (2012): The Helsinki Psychotherapy Study: Effectiveness, Sufficiency, and Suitability of Short- and Long-Term Psychotherapy. In R. Levy, J.S. Ablon, and H. Kächele (eds.): *Psychodynamic Psychotherapy Research.* Humana Press. Pp. 71–94. I thank Olavi Lindfors (Finnish Institute for Health and Welfare) for the information. The other probabilities or the independence assumption are not based on research. There is no reliable evidence of long-term effectiveness of medication.

is the probability that the patient recovers from depression if the assumptions of item a) apply otherwise and the doctor prescribes both psychoanalysis and medication as treatments?

2. Lotteries have been organised in Europe since the 16th or even the 15th century. Lotteries gained popularity in Europe towards the end of the 18th century. Apparently the first Finnish lottery was organised in 1745 to finance the restoration of the tower of the Turku cathedral. The tower had burned 1738.

A 1772 decree described a lottery to finance a hospital of the diocese of Turku. Lottery tickets had not sold well in previous lotteries so the expected pay off from participating in the lottery was increased. 3 200 tickets were to be sold at the price of 1.5 thalers. The main prize was 900 thalers. There were 400 other winning tickets each with prize 8.31 thalers. The hospital was to receive 12% (=allowance) from the sales of the tickets (before deducting expenses). It is assumed in the following that all tickets are sold.[2]

a) What is the expected payoff of a ticket? (Payoff = prize or money yield ignoring the price of the ticket.)

b) What is the variance of the payoff of a ticket? What is the corresponding standard deviation?

c) What is the expected return of a ticket? (Return = payoff − price of the ticket.)

d) What is the variance of the return of a ticket? What is the corresponding standard deviation?

e) What is the expected allowance for the hospital?

f) What is the variance of the allowance? What is the corresponding standard deviation?

3. The European Social Survey (ESS) was carried out in 2012. Face-to-face interviews were arranged for 1 700 Finns. 77.11765% or 1 311 of them told the interviewer that they had voted in the parliamentary election 2011. The actual voter turnout at the election was 70.5 %.[3]

The interviews are independent. Let us test if the voting proportion 0.7711765 is in accordance, within random variation, with the actual voting proportion 0.705 in the election.

a) Explain carefully why the surveyed voting proportion, appropriately standardised, follows approximately the Standard Normal distribution. (Hint: The Central limit theorem.)

b) Formulate the test statistic for the null hypothesis "voting proportion is 0.705" using the standardised voting proportion statistic of item a). Calculate the value of the test statistic.

c) What do you infer? Use significance level 0.01 (two-sided test). (Hint: The 0.95th, 0.975th, 0.99th, 0.995th, 0.999th, and 0.9995th quantiles of the

[2] S.M. Stigler (2022): *Casanova's Lottery*. University of Chicago Press. Chapters 5, 14. O.E.A. Hjelt (1893): Svenska och finska medicinalverkets historia 1663–1812, Part 3. Helsingfors central-tryckeri. P. 89. Kotimaisten kielten keskus: `https://kaino.kotus.fi/korpus/vks/meta/lait/as1700_rdf.xml` (read 24.5.2024). I thank Katariina Lehto for pointing out the first Finnish lottery and the book by Hjelt. The interpretation of the decree is by the composer of the question.

[3] `http://tilastokeskus.fi/til/evaa/2011/evaa_2011_2011-04-29_tie_001_en.html` (read 26.11.2020). I thank professor emeritus Seppo Laaksonen for pointing out the phenomena and the data.

Standard Normal distribution are 1.645, 1.960, 2.326, 2.576, 3.090, and 3.291, respectively.)

4. Patients arrive at the emergency unit of a hospital with unknown mean rate $\lambda$. The arrivals are assumed independent of each other and their number is modelled by a Poisson distribution with parameter $\lambda > 0$.

The probability mass function of a Poisson random variable $X$ with parameter $\lambda$, is given by:

$$p_X(x) = \mathsf{P}(X = x) = e^{-\lambda}\frac{\lambda^x}{x!},$$

where $x = 0, 1, \ldots$ denotes the number of events during a time interval.

Before recording any data, it is believed that the unknown parameter $\lambda$ has an exponential distribution of the form:

$$p(\lambda) = \begin{cases} e^{-\lambda} & \lambda > 0, \\ 0 & \text{else.} \end{cases}$$

a) The number of patients during a given hour of one work week is $x_1 = 100$, $x_2 = 110$, $x_3 = 90$, $x_4 = 50$, and $x_5 = 150$. Having observed these data and using the prior beliefs about $\lambda$, determine the density function of $\lambda$ (namely, the unnormalised posterior, which is proportional to the posterior by a constant).

b) New data is recorded during the weekend: $x_6 = 250$, and $x_7 = 200$. These additional observations are independent of the earlier ones $(x_1, \ldots, x_5)$ and follow the same Poisson distribution. Update the posterior (again, consider the unnormalised posterior).