First course in probability and statistics
Department of Mathematics and Systems Analysis
Aalto University

Final exam
03 Apr 2018
J. Virta/C. Webb

Guidelines: Write briefly and clearly, but justify your answers. A lone number as an answer does not yield points. The exam has 4 problems, each worth 0–6 points. Each answer sheet should contain:

- Course name and code
- LASTNAME and FIRSTNAMES (in block letters)
- Student number
- Study program and year
- Date and signature

Allowed equipment: A calculator and an a4 note card (hand written, text only on one side, own name in the upper right corner, no need to return)

P1 A new doping test detects 95% of users of a forbidden hormone product, but also produces (false) positive test results to 2% of athletes who do not use the product. From a national skiing team where 1% of the members use the forbidden product, an athlete is selected to be tested using random sampling. Determine the probabilities of:

(a) "The test result is positive", **(2p)**

(b) "An athlete who is tested positive does not use the hormone product", **(2p)**

(c) "An athlete who is tested negative does not use the hormone product". **(2p)**

P2 During an off-season in tennis Henri adopted a new training method. The aim was to improve the success probability of his first serve θ that was 0.50 during the previous season – we assume that each first serve succeeds independently of the other first serves. The effect of the new training method was tested by studying 6 first serves in the beginning of a new season. A null hypothesis $H_0 : \theta = 0.50$ was tested against an alternative hypothesis $H_1 : \theta \neq 0.50$, by choosing $X =$ "number of successful first serves" as the test statistic. A test result of $x = 5$ was observed.

(a) What is the expected number of successful first serves, under the null hypothesis? **(1p)**

(b) What is the probability of the event $X = 5$, under the null hypothesis? **(1p)**

(c) Determine the p-value of the observed test result, and based on that decide whether to reject or accept the null hypothesis. **(2p)**

(d) Determine the rejection region of the test on a 5% significance level. **(2p)**

P3 Consider the random variables X and Y. The random variable X takes values in $\{-1,0,1\}$ and the random variable Y takes values in $\{0,1\}$. Their joint distribution is given in the table below.

	Y	
X	0	1
-1	$\frac{1}{3}$	0
0	0	$\frac{1}{3}$
1	$\frac{1}{3}$	0

In other words, $\mathbb{P}(Y = 0 \text{ and } X = \pm 1) = \frac{1}{3} = \mathbb{P}(Y = 1 \text{ and } X = 0)$, and $\mathbb{P}(Y = 1 \text{ and } X = \pm 1) = 0 = \mathbb{P}(Y = 0 \text{ and } X = 0)$.

(a) Calculate the probability mass functions of the marginal distributions of X and Y. **(1p)**

(b) Calculate the expected values of the two random variables: $\mathbb{E}(X)$ and $\mathbb{E}(Y)$. **(1p)**

(c) Calculate the standard deviations of the two random variables: $\mathrm{SD}(X)$ and $\mathrm{SD}(Y)$. **(1p)**

(d) Calculate the covariance of the two random variables: $\mathrm{Cov}(X,Y)$. **(1p)**

(e) Calculate the correlation of the two random variables: $\mathrm{Cor}(X,Y)$. **(1p)**

(f) Are the two random variables independent? Justify your answer. **(1p)**

P4 An experimental physicist believes that she has found a new elementary particle. She has a device that produces these particles and allows measuring the lifetime of the produced particle. The physicist assumes that the lifetimes of the produced particles are random, independent from each other, and are exponentially distributed with probability density function $f(t|\lambda) = \lambda e^{-\lambda t}$, $t > 0$ with some unknown parameter $\lambda > 0$. She performs the measurement five times and observes the lifetimes $t_1 = 3.5$ s, $t_2 = 2.7$ s, $t_3 = 8.5$ s, $t_4 = 4.2$ s, and $t_5 = 4.7$ s.

(a) Help the physicist form the likelihood function of the parameter λ given the observed data set $t_1, ..., t_5$. **(1p)**

(b) Help the physicist form the maximum likelihood estimate for the parameter λ given the observed data set $t_1, ..., t_5$. **(2p)**

(c) The famous theoretical physicist Prof. Baloney arrives at the site of research and claims: "This is obviously the Ω-particle whose existence I have predicted and I believe its lifetime to be roughly 4 s". Prof. Baloney would like to form a Bayes-estimate for the parameter λ. He thinks that the belief of a lifetime of 4 seconds would be described well by a prior distribution whose density function is $p_0(\lambda) = 4e^{-4\lambda}$, $\lambda > 0$.

Help Prof. Baloney to form a Bayes-estimate for the value of the parameter λ by finding the point where the posterior density obtains its maximum value, when the posterior distribution is calculated by using the prior distribution p_0, the observed data $t_1, ..., t_5$, and the same stochastic model as in problem a. **(3p)**

(**Hint:** You do not need to calculate the normalizing constant of the posterior distribution. You are able to solve the problem simply using the un-normalized posterior distribution.)