Exercise session 5 – Prob. & Stat. 2 — Nov 10, 2022

Bayesian statistics – basics

1. Bayesian coin: I have two coins in a bag. One is regular (one side Head, the other Tail). The other coin has a Head on both sides. I toss a coin and a Head shows. What is the probability that I have the regular coin? what if I toss the coin n times and I see Head every time?

2. We have two boxes, the first contains two white and one black ball, the second contains two black and one white ball. We pick one of the boxes at random (the probability of picking the first box is p) and then we pick a ball.

(a) Describe how to use the MAP method to decide from the color of the drawn ball whether we have chosen the first or the second box.

(b) For $p = 1/2$, calculate the probability of making the wrong decision. Compare with the error of the wrong decision without drawing the ball.

Bayesian statistics – two discrete RVs

3. Several students are taking the same test. Each of them belongs (uniformly at random) to one of three groups – everyone in the first group has learned a question with probability $\vartheta_1 = 0.3$, everyone in the second group has learned a question with probability $\vartheta_2 = 0.7$, and in the third with probability $\vartheta_3 = 0.95$. A randomly selected student answers k questions (out of ten) correctly.

(a) Use the MAP method to decide which group the student belongs to.

(b) Let L be the number of questions the student has learned. The student answers 5 questions correctly. Infer the posterior distribution, the MAP estimate of L, and the LMS estimate.

Bayesian statistics – two continuous RVs

4. Romeo is meeting Juliet. Juliet has a "maximum lateness feature" given by random variable $\Theta \sim U(0, 1)$ (time measured in hours). If $\Theta = \vartheta$ then Juliet is late for any meeting by a random variable $X \sim U(0, \vartheta)$. After one meeting when Juliet was late by time x, what can Romeo infer about distribution of Θ? Formally, what is $p_{\Theta|X}(\vartheta|x)$?

5. Suppose that a police radar measures the speed of a car plus an error $U \sim U(0, 5)$. Suppose that the actual speed of the car is $V \sim U(45, 65)$. Use the conditional probability method (LMS) to estimate the speed of the car based on the measured speed.

6. The number of minutes between bus arrivals at the bus stop has an exponential distribution with parameter Θ, we use a prior distribution with density $f_\Theta(\vartheta) = 10\vartheta$ for $\vartheta \in [0, 1/\sqrt{5}]$ (and zero elsewhere – verify that this is a density).

(a) We go to the bus stop and have to wait 30 minutes. What is the posterior density and estimate of θ (by MAP and LMS method).

(b) We take five measurements (we go on the bus five times and write down the waiting time). We have to wait 30, 25, 15, 40 and 20 minutes, assuming that the days are independent measurements with the same distribution. What is the posterior density and estimate of Θ (by MAP and LMS methods).

More problems for practice

7. A variant of the problem number 4: we have independent random variables $X_1, \dots, X_n$. All are uniform on the interval $[0, \vartheta]$, where ϑ is the value of n.v. Θ. The a priori distribution of Θ is uniform on $[0, 1]$. Before we had $n = 1$, now we will assume $n > 3$.

(a) Find the distribution of Θ using the conditional probability (given values $x_1, \dots, x_n$ of the variable $X_1, \dots, X_n$).

(b) Plot the conditional mean squared error (MSE) plot for the MAP estimates and conditional means, as a function of $\bar{x} = \max\{x_1, \ldots, x_n\}$, for $n = 5$.

(c) If $\bar{x} = 0.5$, how do the MAP estimate and conditional mean, and the corresponding mean squared error, evolve, as a function of $n \to \infty$?

8. Quido is writing a test with ten questions, each with three choices. For each of the questions, one of two equally likely choices occurs (independently of the other choices): Either Quido has studied the material, learned the question, and gets it right, or he hasn't studied the material and guesses a uniformly random choice.

(a) If Quido answered the first question correctly, what is the probability that he has learned the question?

(b) What is the a priori distribution (probability function) of the number of questions that has Quido learned (out of the ten in the test, not in total).

(c) Quido answered six out of ten questions correctly. What is the posterior distribution of the number of questions Quido has learned?

9. We toss a coin with probability of Head being Θ. Our a priori distribution is given by the function $f = f_\Theta$ such that $f(0) = f(1) = 0$, $f(0.5) = 2$, f is linear on the intervals $[0, 0.5]$ and $[0.5, 1]$, and f is zero outside $[0, 1]$.

Find the MAP estimate of Θ if out of n independent tosses we got k Heads.

10. The number of shopping carts in a store is Θ – a uniformly distributed random variable with values $\{1, 2, \ldots, 100\}$. Each cart has a number (1, ..., Θ). On our cart (assumed to be uniformly random) is a number X. Estimate Θ

(a) by the MAP method,

(b) by the LMs method.