Last updated on February 24, 2026

BM5163 Bayesian Inference in Bioengineering

Problem Set 1

Instructions

1. You are expected to work on these problems on your own and not submit the solutions.

Questions

1. Suppose an experimentalist is trying to identify relationships between risk indicators (lifestyle, genetic disposition, etc.) and a particular disease. For this significance tests are employed at significance level α and statistical power $1-\beta$. Suppose there are total p equally probable relationships and the researcher randomly tests one of them
   (a) Create the error matrix and calculate rates.
   (b) If the significance test identifies a relationship to be true, what are the chances of it being true positive?
   (c) Can you estimate bounds on α and β to ensure that the probability of a relationship identified in this experiment is indeed a true relationship is greater than ζ?
2. Once you have solved the first question, read this paper.
3. Suppose a person is tested positive for a condition (with prevalence c) in a diagnostic test with sensitivity s_1 and specificity p_1. On the doctor's recommendation, he/she gets tested again by a different test (s_2, p_2) and is tested positive.
   (a) At the end of the two tests, what is the probability that the person has the particular condition?
   (b) Will the probability change if the order of two tests is reversed?
   (c) Analyze the problem for all possible combinations based on the order of tests and outcomes of two tests.
4. Write a `Python` function for sequential Bayesian updating in such diagnotic tests.
5. Redo the third question with n tests (s_i, p_i) out of which m give positive result. Test this with the `Python` program from the last question.
6. A portable biosensor is being used to screen individuals for a chronic metabolic disorder. This biosensor has sensitivity and specificity of 88% and 93%, respectively. For the chronic metabolic disorder, the population consists of two risk groups. The first group consists of 20% of the total population belonging to a high-risk group, where the prevalence of the disorder is 15%, and the second one, with 80% of the total population belong to a low-risk group, where the prevalence is 1%. A randomly selected individual from the population undergoes two independent tests using this biosensor, and both tests return positive. Initially, the clinician does not know whether the individual belongs to the high-risk or low-risk group.
   (a) Assuming the person is from the high-risk group, what is the probability they have the disorder after two positive tests?
   (b) Assuming the person is from the low-risk group, what is the probability they have the disorder after two positive tests?
   (c) If group membership is unknown, what is the overall probability that the person has the disorder after the two positive tests?
   (d) Does it matter whether we first update based on risk group and then incorporate test results, or first update based on test results and then incorporate risk information? Explain mathematically.
7. A clinician is evaluating a patient whose blood test shows an elevated inflammatory biomarker. This biomarker can be elevated due to either an autoimmune disorder (A) or a bacterial infection (I). These two conditions are assumed to be mutually exclusive for this problem (a patient cannot have both simultaneously). In the general population, the prevalence of A and I are 5% and 3%, respectively. The inflammatory biomarker test has the following characteristics. For a patient with A, it gives a positive result with a probability of 0.90. For a patient with I, the probability of positive test is 0.80. For a healthy person, it is 0.10. A patient tests positive for the biomarker.

(a) What is the probability that the patient has A?

(b) What is the probability that the patient has I?

(c) What is the probability that the patient is healthy despite the positive test?

(d) Which diagnosis is most likely after observing the positive result?

8. A hospital is screening patients for a rare neurodegenerative disorder using a novel imaging biomarker. 10% of the population carries a specific genetic variant (G^+) that increases the risk of the disease. Among individuals with the genetic variant (G^+), the probability of developing the disease is 20%. Among individuals without the variant (G^-), the probability of developing the disease is 2%. If a patient has the disease, it can manifest in two subtypes S_1(aggressive subtype) and S_2(mild subtype). The prevalence of these two subtypes, $S_1$and S_2, in the diseased population is 40% and 60%, respectively. The imaging biomarker behaves differently across subtypes. For cases with S_1, the test is positive with probability of 0.95, and for S_2it is 0.80. For a healthy individuals the false positive rate is 0.05. A randomly selected person from the population undergoes the imaging test and receives a positive result.

(a) What is the probability that the patient truly has the disease given a positive result?

(b) Given that the test is positive, what is the probability that the patient has the aggressive subtype?

(c) What is the probability that the patient carries the genetic variant given a positive test?