MITx: 6.041x Introduction to Probability - The Science of Uncertainty

Bookmarks

Bookmark

Problem 5: Hypothesis test between two normals

(4/4 points)

Conditioned on the result of an unbiased coin flip, the random variables $T_1, T_2, \ldots, T_n$ are independent and identically distributed, each drawn from a common normal distribution with mean zero. If the result of the coin flip is Heads this normal distribution has variance 1, otherwise it has variance 4. Based on the observed values $t_1, t_2, \ldots, t_n$, we use the MAP rule to decide whether the normal distribution from which they were drawn has variance 1 or variance 4. The MAP rule decides that the underlying normal distribution has variance 1 if and only if

$$\left| c_1 \sum_{i=1}^{n} t_i^2 + c_2 \sum_{i=1}^{n} t_i \right| < 1.$$

Find the values of $c_1 \geq 0$ and $c_2 \geq 0$ such that this is true. Express your answer in terms of n, and use 'ln' to denote the natural logarithm function, as in 'ln(3)'.

$c_1 =$ 3/(8*n*ln(2)) **Answer:** 3/(8*n*ln(2)) $c_2 =$ 0

Answer: 0

Answer:
Let $\Theta = 0$ denote that the observations $t_1, t_2, \ldots, t_n$ were generated from a normal distribution with variance 1, and let $\Theta = 1$ denote that they were generated from a normal distribution with variance 4. For simplicity, let us use the notation $N(t_1, \ldots, t_n; 0, \sigma^2)$ to denote the joint PDF of n i.i.d. normal random variables with mean 0 and variance σ^2, evaluated at $t_1, \ldots, t_n$.

Therefore, given the observations $t_1, \ldots, t_n$, the posterior probability that the samples are generated from a normal distribution with variance 1 is

$$\mathbf{P}(\Theta = 0 \mid T_1 = t_1, \ldots, T_n = t_n) = \frac{(1/2) \cdot N(t_1, \ldots, t_n; 0, 1)}{(1/2) \cdot N(t_1, \ldots, t_n; 0, 1) + (1/2) \cdot N(t_1, \ldots, t_n; 0, 4)}.$$

Similarly, the probability that the samples are generated from a normal distribution with variance 4 is given by

$$\mathbf{P}(\Theta = 1 \mid T_1 = t_1, \ldots, T_n = t_n) = \frac{(1/2) \cdot N(t_1, \ldots, t_n; 0, 4)}{(1/2) \cdot N(t_1, \ldots, t_n; 0, 1) + (1/2) \cdot N(t_1, \ldots, t_n; 0, 4)}.$$

The MAP rule favors $\Theta = 0$ if the following inequality holds:

$$\mathbf{P}(\Theta = 0 \mid T_1 = t_1, \ldots, T_n = t_n) > \mathbf{P}(\Theta = 1 \mid T_1 = t_1, \ldots, T_n = t_n)$$

We notice that the denominators in the expressions for $\mathbf{P}(\Theta = 0 \mid \ldots)$ and $\mathbf{P}(\Theta = 1 \mid \ldots)$ are the same, so it suffices to compare the numerators. Therefore, the MAP rule favors $\Theta = 0$ if the following inequality holds:

$$\begin{aligned} N(t_1, \ldots, t_n; 0, 1) &> N(t_1, \ldots, t_n; 0, 4) \\ \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi \cdot 1}} e^{-\frac{t_i^2}{2\cdot 1}} &> \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi \cdot 4}} e^{-\frac{t_i^2}{2\cdot 4}}. \end{aligned}$$

With a little bit of algebra, we obtain

$$\left| \frac{3}{8} \sum_{i=1}^{n} t_i^2 \right| < n \cdot \ln(2).$$

You have used 1 of 2 submissions

DISCUSSION

Click “Show Discussion" below to see discussions on this problem.