Homework # 3

Foundations of Computer and Data Science CS-596

Problem 1: Let χ_1, χ_2 be random variables which 99% of the time are independent and Normally (Gaussian) distributed, both with mean 0 and variance 1 and 1% of the time they are independent and Normally distributed both with mean 0 and variance $\sigma^2 \neq 1$. a) Compute the joint pdf of the two random variables. b) Examine if the two random variables are *independent*. c) Give an example of two random variables that are *uncorrelated* but not independent.

Problem 2: Let χ, ζ be random variables that are related through the equality

$$\zeta = |\chi + s|.$$

a) If the pdf of χ is $f_\chi(x)$ compute the pdf of ζ when s is a deterministic quantity. b) Repeat the previous question when s is a random variable independent from χ and takes only the two values 0 and 1 with probability 0.2 and 0.8 respectively. c) Under the assumptions of question b) compute the posterior probability $\mathsf{P}(s = 0|\zeta = z)$. *Hint: For the computation of the pdf of a random variable the simplest way is to start with the computation of the cdf and then take the derivative. For b) use total probability.*

Problem 3: You are given an *unfair* coin where every time you throw it the probability to observe "head" is $p \in (0, 1)$ (and "tail" $1 - p$) where p is unknown. Assume that you throw the coin N times and you report the sequence of heads (H) and tails (T). If N_h is the number of heads you observed and $N - N_h$ the number of tails then: a) Find the probability of the specific sequence you have observed as a function of p, N, N_h. b) Compute the maximum likelihood estimator (MLE) of p if you are given such a sequence. Is the result familiar/expected? c) Compute the average of your estimate and the mean square error from the true unknown value p. What do you conclude about your estimate as $N \to \infty$?

Problem 4: Consider the random data $\{x_1, \ldots, x_N\}$ and assume that they are Markov related as follows

$$x_n = \alpha x_{n-1} + w_n, \quad n = 2, \ldots, N,$$

where $|\alpha| < 1$, $\{w_n\}$ are independent and identically distributed Gaussian random variables with mean 0 and variance 1 and independent from x_1. Random variable x_1 is also Gaussian with mean 0 and variance $\frac{1}{1-\alpha^2}$. a) Using the fact that linear combinations of Gaussians is also Gaussian show that all x_n are Gaussian. b) Find the joint probability density function of $\{x_1, \ldots, x_N\}$ assuming α is given. c) Find the maximum likelihood estimator (MLE) of parameter α. *Hint: For b) use the fact that* $\{x_1, x_2 - \alpha x_1, \ldots, x_N - \alpha x_{N-1}\}$ *are independent Gaussians.*

Problem 5: Assume that you have two hypotheses $\mathsf{H}_0, \mathsf{H}_1$. Under Hypothesis H_0 your observation vector $X = [x_1, \ldots, x_N]^T$ is Gaussian with mean vector μ_0 and covariance matrix Σ_0, while under H_1 it is again Gaussian with mean μ_1 and covariance matrix Σ_1. Assume that the two hypotheses are equiprobable ($\mathsf{P}(\mathsf{H}_0) = \mathsf{P}(\mathsf{H}_1) = 0.5$). a) Find the Likelihood ratio test (LRT) and its equivalent form by taking the logarithm on both sides of the inequality. b) What is the latter form reduced to when the two covariance matrices are equal ($\Sigma_0 = \Sigma_1$)? c) What if additionally the two means are equal? Does this make sense?

There is a meeting on Friday, November 15, at 5PM, in CoRE-101 to discuss YOUR questions.

Your reports, in *hard copy*, are due on Monday, November 18, in CBIM, between 10:00 and 11:00AM. Mr Stathopoulos will collect them.

Please STRICTLY comply with the indicated period because Mr. Stathopoulos has many other tasks, besides TAing. Late/Early submission will NOT be accepted neither by the TA nor by me.