CIS5526: Homework 7

Assigned: November 23, 2020

due Monday, December 7th by noon through Canvas

Homework Policy

All assignments are INDIVIDUAL! You may discuss the problems with your colleagues, but you must solve the homework by yourself. Please acknowledge all sources you use in the homework (papers, code or ideas from someone else).Assignments should be submitted in class on the day when they are due. No credit is given for assignments submitted at a later time, unless you have a medical problem.

1. (20 points) You are given a sample $\{y_i\}$, i=1,… N, from an unknown probability distribution p(y). Show that value t that minimizes $\sum_{i=1}^{N} |y_i - t|^R$ is (a) median of the sample for R=1, (b) average of the sample for R=2, (c) average between the highest and smallest y for R=∞.

2. (20 points) Show that the maximum likelihood estimate of the mean and standard deviation of Gaussian distribution given the data set D = {x_i, i=1, 2, …N} is $\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$, $\sigma = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2$.

3. (20 points) You are given a data set D = {2, 3, 2} of numbers sampled randomly from Poisson distribution with parameter μ defined as $p(x|\mu) = e^{-\mu}\mu^x/x!$
   a) Calculate the log-likelihood that D is generated from Poisson distribution with parameter μ =1.
   b) Use the maximum likelihood approach to find the optimal value of μ.

4. (20 points: programming assignment) Generate N = 100 points from Gaussian distribution with mean 2 and standard deviation 2 (those values are the ground truth).
   Plot a histogram of D.
   a. Plot a histogram of D. What is the mean and standard deviation on this sample?
   b. Calculate the log-likelihood P[D| μ,σ] given (i) μ = 1, σ=2, (ii) μ = 1.5, σ=2, (iii) μ = 2, σ=2. Which of those results in the highest log-likelihood?
   c. Find the maximum likelihood estimation of μ, σ. How does this compare to the true parameters?
   d. Repeat the following procedure 200 times: (1) generate N = 100 points Gaussian distribution with mean 2 and standard deviation 2; (2) find the maximum likelihood estimates of μ,σ. Plot the histogram of those 50 outcomes. What is the average of the estimates over 50 experiments? Now, repeat the same procedure, but generate N = points in each iteration. How do the results change? What is the explanation for this (you can search the Web to answer)?

5. (20 points) After random sampling from an underlying distribution, you obtain a small square uniformly filled with 1000 CLASS 1 examples and a large square uniformly filled with 1000 CLASS 2 examples.

   a) What is the largest classification accuracy you can obtain by any machine learning method?

   b) What would be the accuracy of 1-nearest-neghbor classifier?

   c) What is the best achievable accuracy by a linear classifier? Why?