IT585 Advanced Machine Learning

Lab2

Understanding Hoeffding's Bound for Learning Theory

Instructions:

1. You have to code the solution in Google colab
2. You can use inbuilt libraries from python
3. Your plots, code, any insights, observations written as text should be submitted as one ipynb file to google classroom
4. Deadline : January 29,2024 11:59 PM IST
5. Name of your file should be : yourrollno_lab2.ipynb

Run a computer simulation for flipping $1,000$ fair coins. Flip each coin independently 10 times. Let's focus on 3 coins as follows: c_1 is the first coin flipped; c_{rand} is a coin you choose at random; $c_{\min}$ is the coin that had the minimum frequency of heads (pick the earlier one in case of a tie). Let ν_1, ν_{rand} and $\nu_{\min}$ be the fraction of heads you obtain for the respective three coins.

(a) What is μ for the three coins selected?

(b) Repeat this entire experiment a large number of times (e.g., $100,000$ runs of the entire experiment) to get several instances of ν_1, ν_{rand} and $\nu_{\min}$ and plot the histograms of the distributions of ν_1, ν_{rand} and $\nu_{\min}$. Notice that which coins end up being c_{rand} and $c_{\min}$ may differ from one run to another.

(c) Using (b), plot estimates for $\mathbb{P}[|\nu-\mu| > \epsilon]$ as a function of ϵ, together with the Hoeffding bound $2e^{-2\epsilon^2 N}$ (on the same graph).

(d) Which coins obey the Hoeffding bound, and which ones do not? Explain why.