EE 527: Machine Learning Laboratory

Assignment 2

Due date: 23 Jan 2023

---

1. Generate **N** points in an interval $[a, b]$**.** Evaluate the normalized frequency distribution of these points. Make m bins for constructing the distribution.
   - Let $a = -100$ and $b = 100$.
   - Experiment with **N** = 100, 1000 & 10000.
   - Choose the value of **m** appropriately.

   Plot the normalized frequency distribution.

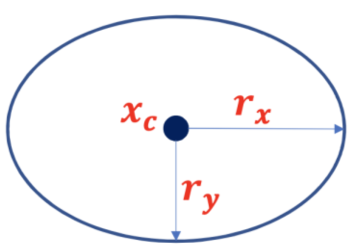
2. With **N** = 5000 and **m** = 400, construct a normalized frequency distribution (**Q1**). Treat this distribution as a weighted dataset $\{(x_i, p_i); i = 1,2, \dots m\}$, where $\sum_{i=1}^{m} p_i = 1$.

   Evaluate the weighted **AM**, **GM**, **HM**, **median** and **mode**.

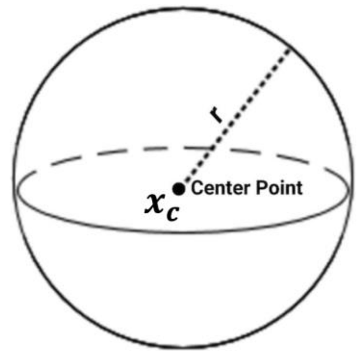
3. Write a general code that takes any weighted dataset $\{(x_i, \omega_i); i = 1,2, \dots n\}$ as input and provides the weighted AM, GM, HM, median and mode. In this particular case, $\omega_i \geq 0$ and $\sum_{i=1}^{n} \omega_i \neq 1$.

4. **Generation of Points**

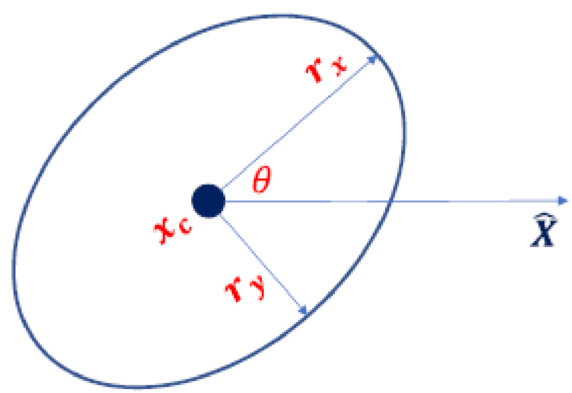
   a. Randomly generate **n**=1000 2D points $S_e = \{x_1, x_2, \dots x_n\}$ inside an a 2D ellipse of axes r_x= 150, r_y= 100 and centered at $x_c = (-10, 20)$. The axes of the ellipse are aligned with the co-ordinate system axes.



b. Randomly generate **n**=1000 points $S_{hs} = \{x_1, x_2, \ldots x_n\}$ inside a 10-Dimensional hypersphere of radius $r = 100$, centered at $x_c = (-1,2-1,0,0,0,3,4,9,0)$.



c. Randomly generate **n** = 1000 2D points $S_{eo} = \{x_1, x_2, \ldots x_n\}$ inside an oriented 2D ellipse of axes $r_x = 150, r_y = 100$ and centered at $x_c = (-10,20)$. The major axis makes an angle of $\theta = \frac{\pi}{3}$ with the horizontal axis $\hat{X}$.



5. **Covariance Matrix Computation**

Compute the 2x2 Covariance Matrix **C** using the points in S_{eo}. Plot the Eigen Vectors $(\hat{e}_1, \hat{e}_2)$ of **C** and the axes of the oriented ellipse, all originating from the center x_c. The lengths of $(\hat{e}_1, \hat{e}_2)$ should be respectively set to $\left(k\sqrt{\lambda_1},\ k\sqrt{\lambda_2}\right)$. Change the value of **n** and report observations. Try plotting with k = 3, 4, 5.