Practical Sheet 1: Probability Refresher

APL 405 - 2023W (Machine Learning for Mechanics)

1. **[Practical 1: 10 marks]** It is frequently important to measure certain characteristics of a random variable, say X. For example, we might only be interested in the average or mean or expected value of X. Assume X is a standard Gaussian random variable $\mathcal{N}(0, 1)$.

   (a) Write a program to verify using *sample mean* estimate,

   $$\text{Sample mean, } \hat{\mu} = \frac{1}{M} \sum_{i=1}^{M} x_i$$

   that X would be zero on an average. Report results for M = 10, 100, 1000, 10000, and 100000

   (b) One of the most important problems in probability is to determine the PDF for a transformed random variable, i.e., one that is a function of X, say X^2 as an example. Plot the PDF of X^2 side by side to that of X.

   (c) Repeat the task for part (a) for computing the sample mean of the transformed random variable X^2.

2. **[Practical 2: 10 marks]** Expectation operator $\mathbb{E}[\cdot]$ is linear, meaning

   $$\mathbb{E}_X[a_1 g_1(X) + a_2 g_2(X)] = a_1 \mathbb{E}_X[g_1(X)] + a_2 \mathbb{E}_X[g_2(X)]$$

   for any two constants a_1 and a_2 and any two functions $g_1(X)$ and $g_2(X)$.

   Consider X is a univariate Gaussian random variable with mean 1 and variance 2. Now consider $g(X) = a_1 g_1(X) + a_2 g_2(X)$ such that $a_1 = 2$, $a_2 = 1$, $g_1(X) = X + 1$ and $g_2(X) = -4X - 5$.

   (a) Manually derive the expected value (or the mean) of $g(X)$.

   (b) Write a code to verify if the sample mean of $g(X)$ converges to the derived expected value (from previous part).

   (c) Plot the PDF of $g(X)$ using histogram with 100000 samples from X. Choose an appropriate number of bins to represent the PDF of $g(X)$.

3. **[Practical 3: 15 marks]** Consider a case where some *noisy* measurements of a constant θ are available. The random variable Y (termed here as the measurement) is related to θ in the following fashion:

$$Y = 3\theta + \epsilon$$

where ϵ denotes a noise term taken to be a Gaussian random variable with zero mean and variance $\sigma^2 = 0.1$. The true value of θ is 0.5.

(a) Derive the expected value (or mean) of Y. What relation does the expected value has with θ?

(b) Generate 100 measurements of Y using the given the 'noise' distribution of ϵ and true value of θ

(c) Now treat as if the true value of θ is unknown and use the generated 100 measurements of Y to calculate (or *estimate*) θ.

(d) The estimate of θ from the (c) is a random variable. What is the reason? Also, compute the variance of the estimate of θ

(e) Write a code to compute the mean of the estimate of θ, by considering different batches of measurements.