TD 1

September 2, 2024

Exercise 1

Consider X, Z two independent random variables $X \sim \mathcal{N}(0,1), Z \sim \mathcal{N}(0,1)$. We denote $\epsilon = 1$ if sg(X) = sg(Z) and $\epsilon = -1$ otherwise. Then we denote $Y = \epsilon Z$.
In other words, $Y = Z$ if X and Z have the same sign and $Y = -Z$ otherwise.

1 Write a code to generate N samples of X, Z, ϵ, Y and plot $\{(x_1, y_1), \ldots, (x_n, y_n)\}$ samples for $N = 1000$.

2 Plot histogram of samples $\{y_1, \ldots, y_n\}$. Prove that Y follows a normal distribution.

3 Print empirical correlation of x_i and y_i. Compute formal expression of $\mathrm{Cov(X, Y)}$.

4 Plot histogram of samples $\{x_1 + y_1, \ldots, x_n + y_n\}$. What do you observe ?

5 Is (X, Y) a Gaussian vector ? (Justify) Compute the joint density $f_{X,Y}$.

6 Compute the density f_{X+Y}. Is $X + Y$ Gaussian ?

Exercise 2

Consider $X \sim \mathcal{N}(0,1), N \sim \mathcal{N}(0, \sigma_N^2)$ independent and $W = -X + N$.

1 Write a code to generate N samples of X, N, W and plot $\{(x_1, w_1), \ldots, (x_n, w_n)\}$ samples for $\sigma_N = 0.25$ and $N = 1000$.

2 Compute the density f_W. Is (X, W) a Gaussian vector ? (Justify)

3 What are equi-probability curves for (X, W) joint distribution ?

4 What is the covariance matrix associated with (X, W) ? Deduce the joint density $f_{X,W}$.

5 Assume now that σ_N is unknown. What would be an unbiased estimator of σ_N ? (using samples from (X, W))

Exercise 3

We consider same random variables as previous exercises. Let B be a Bernouilli variable of parameter 0.5. Finally, $V = Y$ if $B = 0$ and $V = W$ otherwise.

This is a particular case of **mixture** of random variables. Let $(X_1, \dots, X_d)$ be random variables of densities $(f_1, \dots, f_d)$ and consider an independent discrete random variable A with values in $\{1, \dots, d\}$ associated with probabilities $(p_1, \dots, p_d)$. Then U is defined by $U = X_i$ iif $A = i$.

1 Prove that $\sum_{k=1}^{d} p_k f_k$ is still a probability density and that it corresponds to the distribution of U.

2 Deduce that $\mathbb{E}[U] = \sum_{k=1}^{d} p_k \mathbb{E}[X_k]$ and $\mathrm{Var(U)} = \sum_{\mathrm{k=1}}^{\mathrm{d}} \mathrm{p_k Var(X_k)}$.

3 Write a code to generate N samples of B, V and plot $\{(x_1, v_1), \dots, (x_n, v_n)\}$ samples for $N = 1000$. What is $\mathbb{E}[V]$ and $\mathrm{Var(V)}$?

4 Plot histogram of $\{v_1, \dots, v_n\}$ samples and compute the density f_V.

5 Show that $\mathrm{Cov(X, V)} = \frac{1}{2}(\mathrm{Cov(X, Y)} + \mathrm{Cov(X, W)})$.

6 How can we modify the definition of W to obtain $\mathrm{Cov(X, V)} = 0$?

7 In this case, are X and W independent ? Is (X, V) a Gaussian vector ?

8 Compute densities $f_{X,V}$ and f_{X+V}.

9 Plot histogram of $\{x_1 + v_1, \dots, x_n + v_n\}$ samples as well as the function f_{X+V} and compare them for 10^2,10^3 and 10^4 samples.

10 Assume now that samples coming from W are not observable. How can find an estimator of σ_N defined on samples from (X, V) ?

11 We denote $\chi^2_\beta(n)$ the β-quantile of khi-squared distribution with n liberty degrees. What would be a confidence interval of probability $1-\alpha$? (Justify)