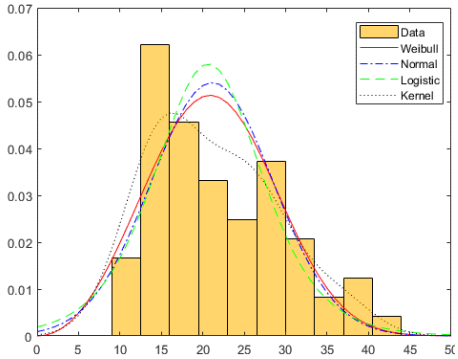
Group Assignment 2

Fitting a probability distribution



In this exercise each group will explore further the data collected in Part A of Group Assignment 1.

Recall that in Part B of Group Assignment 1, every group compiled nine datasets, representing three different times of the day across three locations. For this exercise, you will combine the datasets from the various times into one, resulting in three consolidated datasets, each corresponding to one of the three locations (A, B, & C).

The objective of this assignment is to fit an appropriate continuous probability distribution to each of the three datasets.

Fit the following distributions to each of the datasets:

- Log-normal distribution
- Weibull distribution
- Exponential distribution

(A) *Visualisation:* For each dataset, plot histograms and overlay the fitted distribution curves on the corresponding histogram to illustrate the fit.

[30 Marks]

(B) *Parameter Estimation:* For each fit on each dataset, estimate the parameters of the distribution. Write the probability density function (PDF) using the estimated parameter values.

[40 Marks]

(C) *Goodness-of-Fit Comparison:* Compare the goodness-of-fit for each distribution to each dataset. Utilize appropriate statistical tests and criteria to evaluate and compare how well each distribution fits the data.

[30 Marks]

Prepare a report containing the following:

Visual Plots: Histograms with overlayed distribution fits for each dataset.

Estimated Parameters: A list of the estimated parameters for each distribution fit, alongside the corresponding PDFs written out with these values.

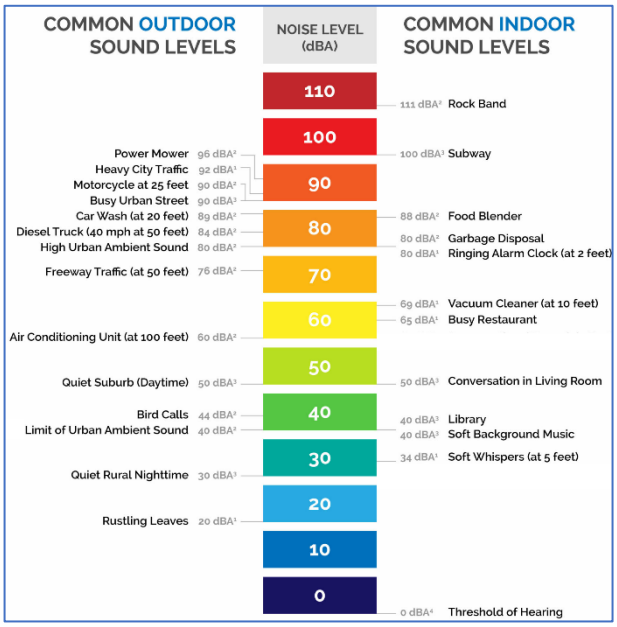
Comparison Report: A detailed comparison of the goodness-of-fit for the distributions across all three datasets, including statistical test results and your interpretation.

In your report, please ensure all plots are clearly labelled with titles, axis labels, and legends (where applicable). Present the estimated parameters and the PDFs in a clear and organized manner, preferably in tabular form. In your comparison report, succinctly discuss the implications of your findings and suggest which distribution provides the best fit for each dataset, supported by your statistical analysis.

Group Assignment 3
Comparison of datasets

In this exercise each group will explore further question *'which location is noisier on campus?'*.

This involves not only examining your own data, data collected in Part A of Group Assignment 1, but also comparing it with data collected by other groups under different scenarios.



Recall in Part A, there were two data collection scenarios:

- Groups with odd numbers (Scenario 1) measured noise levels in student cafeterias.
(Scenario 1-Student cafeteria, Locations: The Pav - location A; The Buttery - location B; The Dining Hall -location C;)

- Groups with even numbers (Scenario 2) focused on noise levels near campus gates.
(Scenario 2-Campus Gate, Locations: Front gate - location A; Nassau St. gate - location B; Pearse St. gate -location C;)

The objective of this exercise is to compare the noise levels at the locations A,B & C to compare the noise levels within a scenario and between scenarios.

Step 1: Data Preparation

Each group should start with the three datasets prepared in Group Assignment 2. Further, download a dataset prepared by any other group recording noise levels in the alternate scenario: Odd-numbered groups download data from an even-numbered group, and vice versa.

Please compile the newly downloaded dataset to create three location-specific datasets, following the approach taken with the data collected by your own group.

Step 2: Location comparison

- Compare noise levels recorded at locations B & C. You may use hypothesis testing methods for comparison. Please comment if the means are comparable between locations B & C.
- Compare noise levels recorded at locations A & C. You may use hypothesis testing methods for comparison. Please comment if the variances are comparable between locations A & C.
- Which location is noisier?

[50 Marks]

Step 3: Scenario comparison

Please compare the datasets recorded in scenarios 1 & 2. You may use hypothesis testing methods for comparison. Answer the following:

Are the campus gates noisier than student cafeterias?

[50 Marks]

Please prepare a report containing the results of steps 2 & 3. Your report should include:

- **Null and Alternative Hypotheses:** Clearly state these for each comparison.
- **Test Type:** Specify the statistical tests used for each comparison.
- **Results and Conclusions:** Present the outcomes of your analyses, including any statistical significance or lack thereof, and interpret the results in the context of the hypotheses.

Submission

Each group is required to submit a report in .pdf format compiling the results of group assignment 2 & 3 by **19th April, 2024** on BB.

Note: Teamwork is an essential part of this assignment, thus, make sure to distribute the tasks evenly among team members. Clearly mention the contributions and role of each team member in your submission. Upon request, your submission may require an interview session.