INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact Factor: 6.078
(Volume 11, Issue 3 - V11I3-1394)
Available online at: https://www.ijariit.com

Musical Analysis Conforming to the Normal Distribution

Satpreet Singh Chadha
satpreetchadha9@gmail.com
Prabhavati Padamshi Soni International Junior College, Mumbai, Maharashtra

Dhruv Tahiliani
tahilianidhruv@gmail.com
Prabhavati Padamshi Soni International Junior College, Mumbai, Maharashtra

ABSTRACT

In music, note durations and inter-onset intervals (IOIs) are crucial parameters that influence the rhythm and flow of a composition. Understanding their distribution helps in uncovering underlying patterns in numerous musical performances. This paper explores the hypothesis that these parameters conform to a normal distribution, which utilises a bell-shaped curve and is commonly used in statistics to model data that clusters around a mean value with symmetrical dispersion.

Keywords: *Math, Mathematics, Note durations, inter-onset intervals, IOIs, Music, Normal Distribution, Statistics*

INTRODUCTION

In music, note durations and inter-onset intervals (IOIs) are essential parameters that affect the rhythm and the flow of a composition. Understanding their distribution helps in uncovering underlying patterns in numerous musical performances. This paper explores the hypothesis that these parameters conform to a normal distribution, which utilises a bell-shaped curve and is commonly used in statistics to model data that clusters around a mean value with symmetrical dispersion.

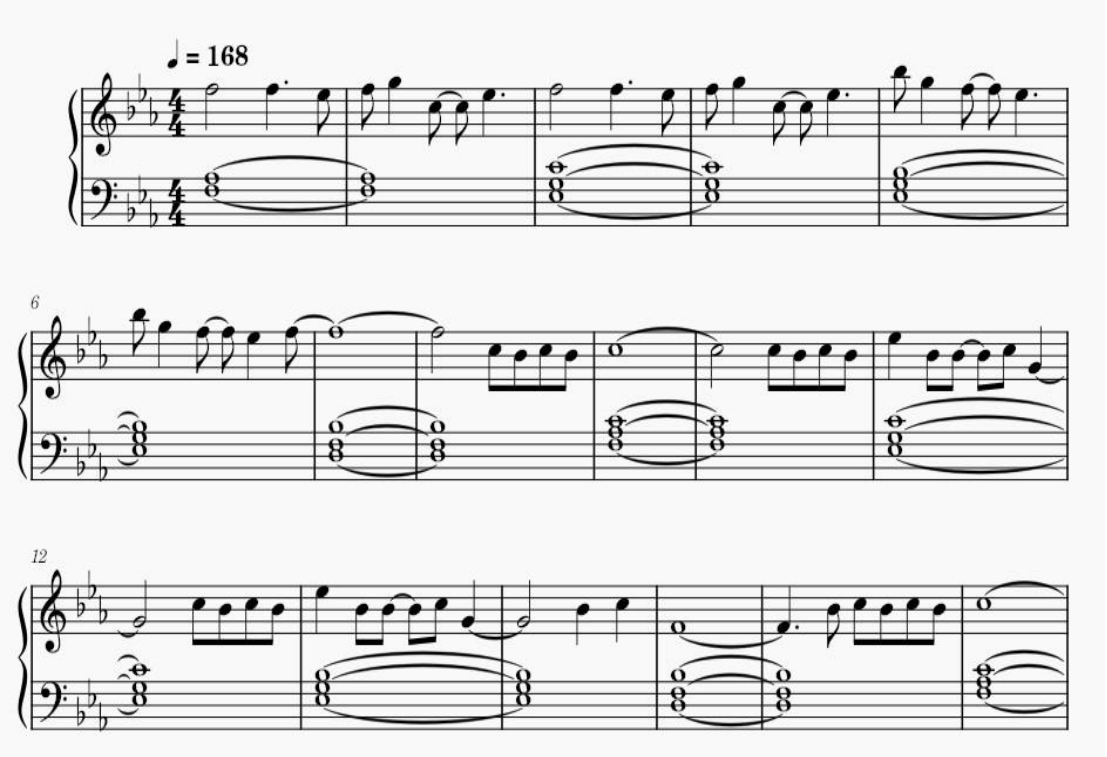
Note durations and IOIs are often expected to follow a normal distribution due to the natural variance in human performance and the general structure of musical compositions. This understanding can help us analyse variations in these parameters and how they reflect performance styles and compositional trends. Our research aims to offer insights into the probabilistic nature of musical expression and structure promoting musical articulation and creativity among artists

Key terms

Note Durations: The length of time for which a note is played is called its note duration ("Note Duration")
Inter onset intervals(IOIs): Inter-onset intervals (IOIs) are the intervals between onsets of adjacent notes and provide a measure of tempo ("Time point")

RESEARCH METHODOLOGY

1. **Data Sampling**: To ensure an unbiased sample, a diverse set of 100 musical pieces were chosen from each genre. For each genre, we assigned a unique number from 1-100 to each musical composition. Then using a random number generator, one song was chosen. The process was repeated for each of the five genres we selected including classical, jazz, pop, rock and Hindustani Classical to ensure a broad analysis of rhythmic patterns. This variety allows us to capture a wide range of rhythmic practices and styles.
2. **Data Collection and Formatting:** We first obtained a MIDI file of each song, then using MuseScore Studio, an open-source music notation software, we converted the MIDI file to an XML file type. An example of a file converted to the XML file format is displayed below:



3. **Data Cleaning and Formatting:** Before using the extracted data for note durations collected from different musical compositions, we ensured that no anomalies were present and the note duration data was accurate. To accomplish this, we generated the note durations from MuseScore several times for each individual composition, scanning for any missing values or gaps present and comparing each outcome obtained from a composition to verify that they are identical.
Then using an algorithm, we obtained accurate note duration data for the song. The algorithm for the same is shown below:

```python
import xml.etree.ElementTree as ET

def analyze_musicxml(file_path):
    tree = ET.parse(file_path)
    root = tree.getroot()

    durations = []

    # Iterate through all note elements
    for note in root.iter('note'):
        duration = note.find('duration')
        note_type = note.find('type')

        if duration is not None and note_type is not None:
            durations.append((duration.text, note_type.text))

    return durations

def print_durations(durations):
    print("Note Durations and Types:")
    for duration, note_type in durations:
        print(f"Duration: {duration}, Type: {note_type}")

if __name__ == "__main__":
    file_path = r"C:\Users\Satpreet\Downloads\Blinding Lights Normal Distribution.py"
    durations = analyze_musicxml(file_path)
    print_durations(durations)

```

The data was then formatted to fit the syntax of the second algorithm that plotted the histogram, normal distribution curve and q-q plot for the data. This process was repeated for each song. A part of the algorithm for the same can be found below:

```python
# Convert note data to DataFrame
df_notes = pd.DataFrame(note_data, columns=['Value', 'Type'])
df_notes['Duration'] = df_notes['Type'].map(note_durations) * df_notes['Value']

# Extract note durations
durations = df_notes['Duration'].values

# Fit a normal distribution
mean, std_dev = np.mean(durations), np.std(durations)

# Generate histogram and Q-Q plot
def plot_histogram_and_qq_with_normal_curve(durations):
    plt.figure(figsize=(14, 6))

    # Histogram
    plt.subplot(1, 2, 1)
    count, bins, ignored = plt.hist(durations, bins='auto', edgecolor='black', alpha=0.6, density=True)

    # Plot normal distribution curve
    xmin, xmax = plt.xlim()
    x = np.linspace(xmin, xmax, 100)
    p = stats.norm.pdf(x, mean, std_dev)
    plt.plot(x, p, 'k', linewidth=2)

    plt.title('Histogram of Note Durations with Normal Distribution Curve')
    plt.xlabel('Duration (beats)')
    plt.ylabel('Density')

    # Q-Q plot
    plt.subplot(1, 2, 2)
    stats.probplot(durations, dist="norm", plot=plt)
    plt.title('Q-Q Plot')

    plt.tight_layout()
    plt.show()

plot_histogram_and_qq_with_normal_curve(durations)
```

4. **Pattern Identification and analysis**: Using the graphs generated by the algorithm, we analysed how closely musical rhythms conform to statistical norms, specifically the normal distribution and what certain deviations might indicate about the musical style or composer's intent. Furthermore, we used graphs of commonly found IOIs of each of the genres to study the impact of the same on musical composition and performance.

Mathematical Working

1. Normal Distribution Theory: the normal distribution, also known as the Gaussian distribution or bell curve, is one of the most important continuous probability distributions in statistics. It is used to describe data that clusters around a central mean value with a symmetrical dispersion. ("Normal Distribution (Definition, Formula, Table, Curve, Properties & Examples)"). Using the functions in the algorithm given above we automatically calculated mean, standard deviation and the probability density function (PDF) of a normal distribution but the following can be calculated through:
   Mean can be calculated using the formula:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$$

Where,

- x_i is the variable
- μ the population mean
- N is the total number of values in the population
- $\sum_{i=1}^{N} x_i$ is the standard deviation

Standard deviation can be calculated using the formula:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$$

Where,

- x_i is the variable
- μ the mean
- N is the total number of values in the population
- $\sum_{i=1}^{N}(x_i - \mu)^2$is the sum of the squared deviations of each value from the population mean.

The probability density function (PDF) of a normal distribution is given by:

$$f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

Where,

- x is the variable
- μ the mean
- σ is the standard deviation

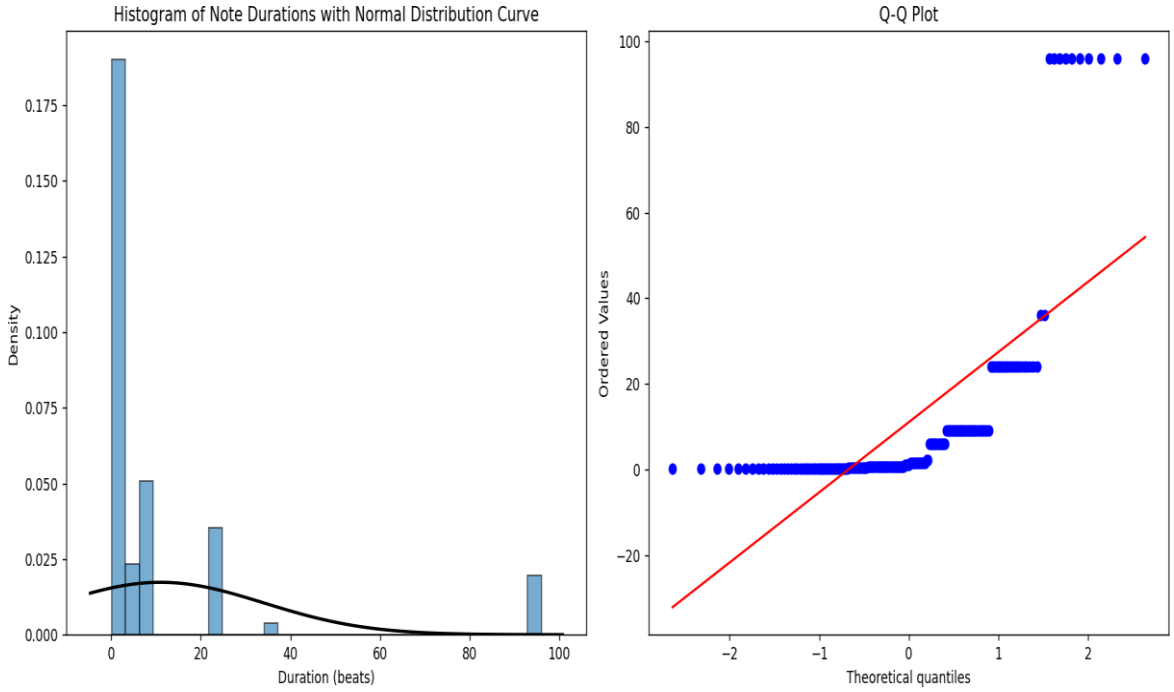
2. Histograms: A histogram is a visual representation of the distribution of quantitative data. The term was first introduced by Karl Pearson. ("Histogram"). Using the procedure clearly labelled and indicated under the histogram part of the algorithm, the histograms for the data were generated.

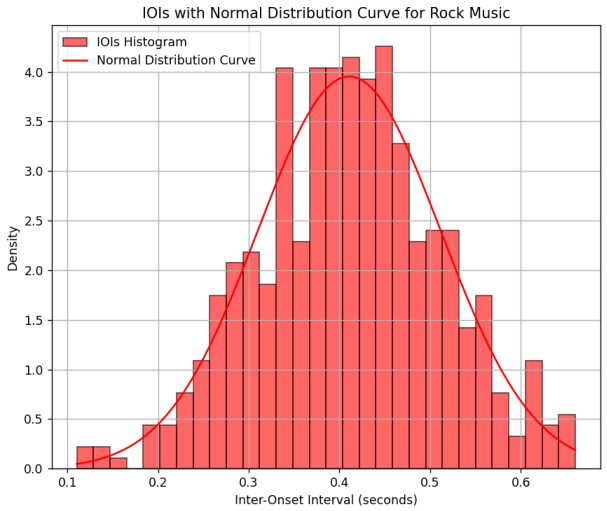
3. Quantile-Quantile (Q-Q) plots: Quantile-Quantile (Q-Q) plots are used to assess whether or not a set of data potentially came from some theoretical distribution. In most cases, this type of plot is used to determine whether or not a set of data follows a normal distribution. (Bobbitt)

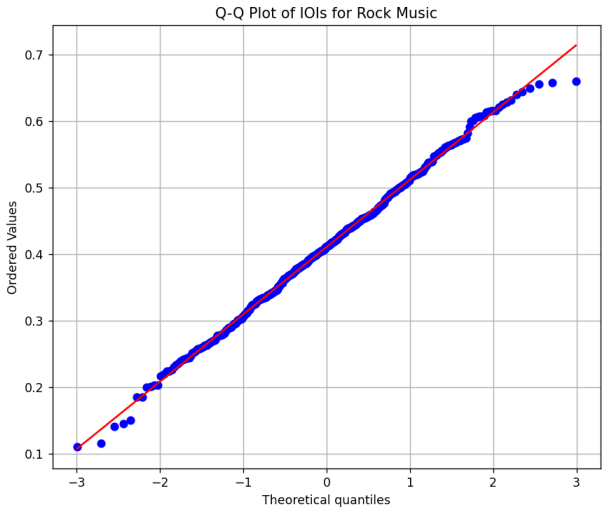
   Using the procedure clearly labelled and indicated under the histogram part of the algorithm, the Q-Q plots were generated.

Analysis

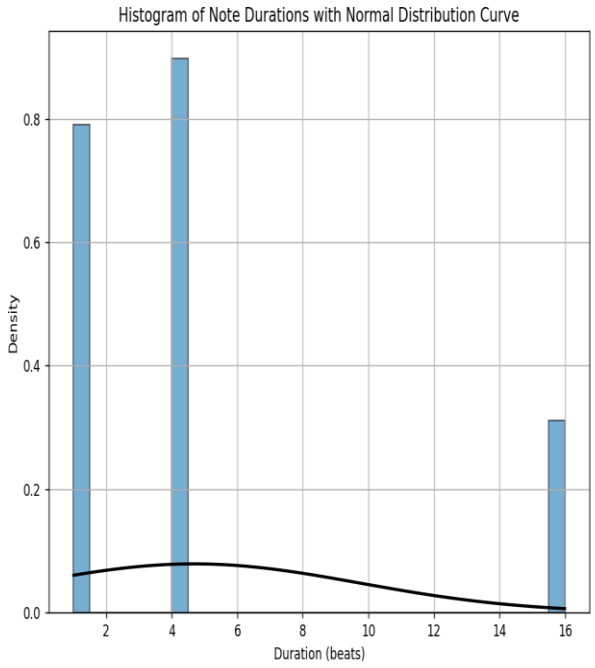
We Will Rock You("We Will Rock You") —> Rock Genre

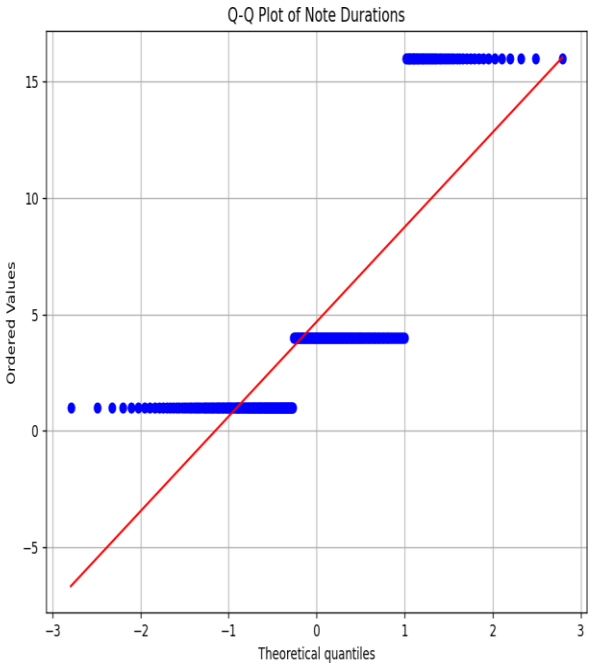


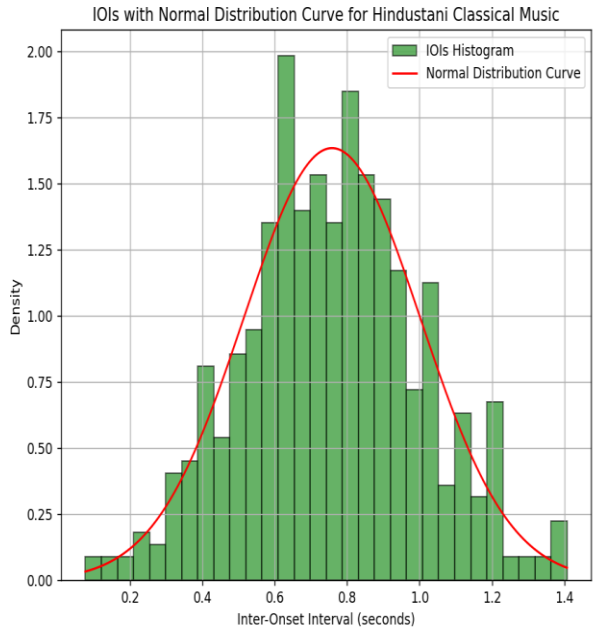


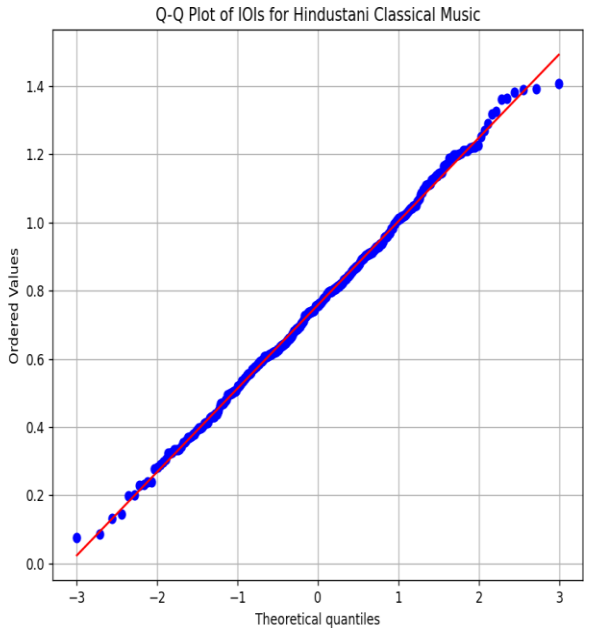


Raag Bhairav("Raag Bhairav") —> Hindustani Classical Genre

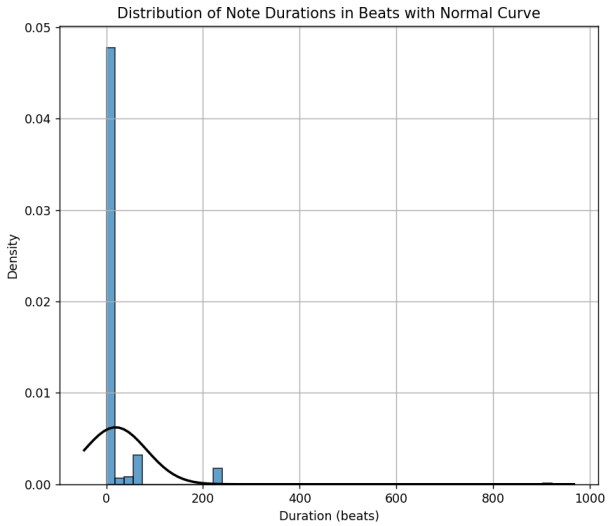


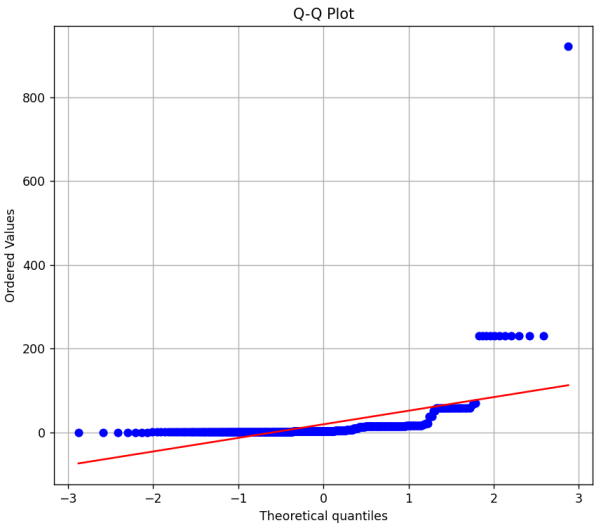


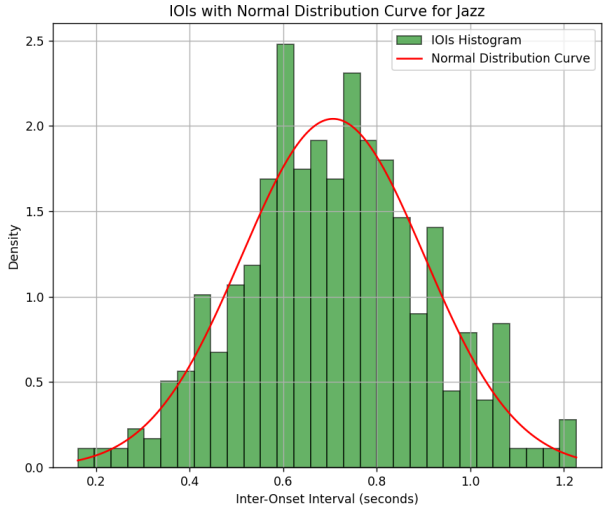


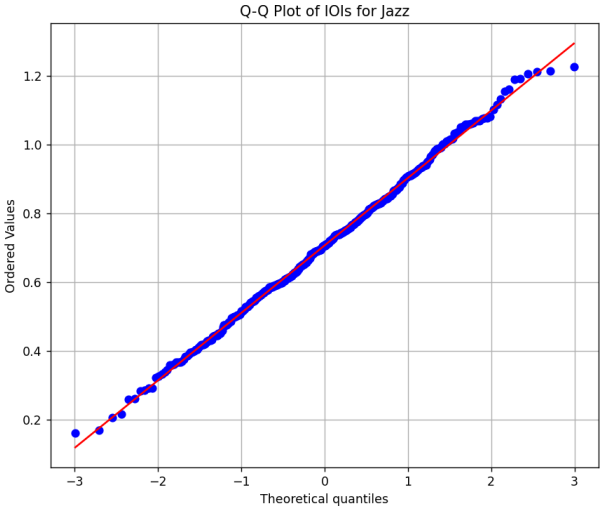


Songbird("Songbird") —> Jazz Genre

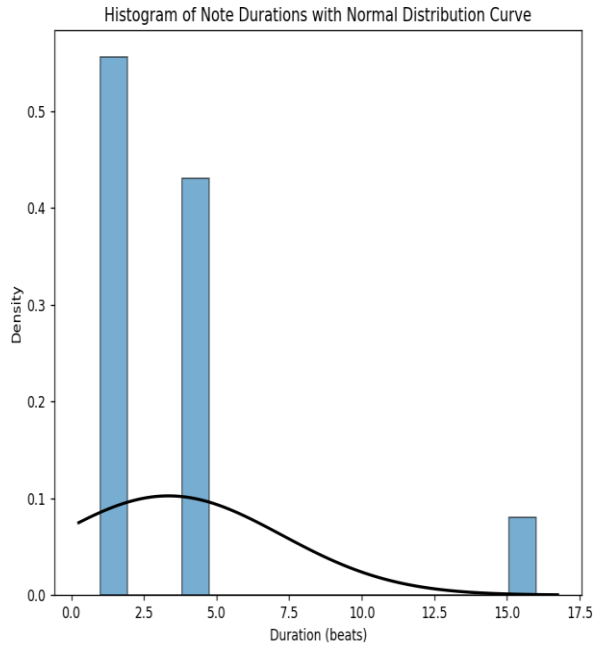


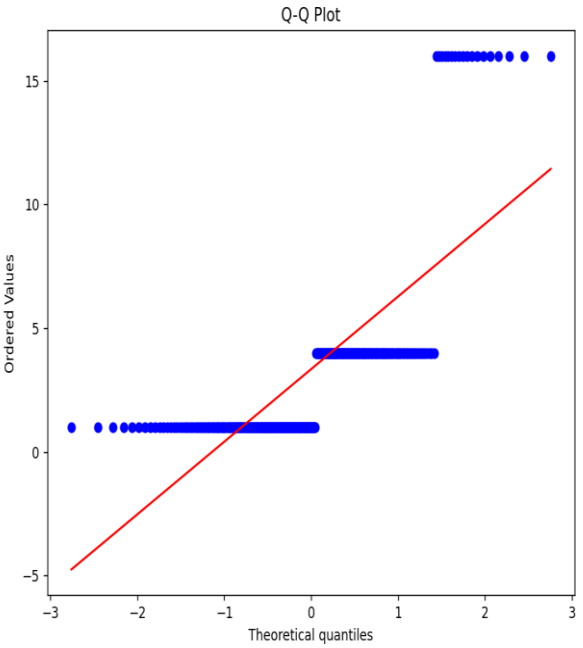


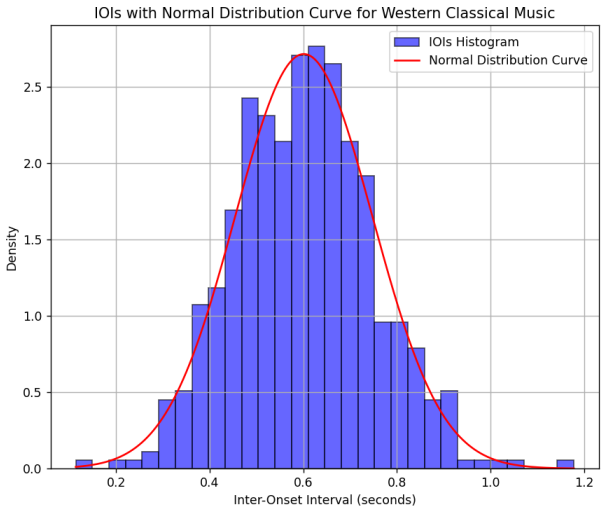


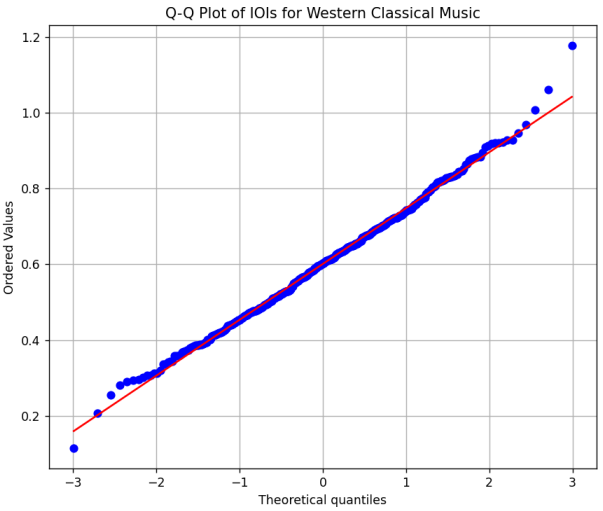


Beethoven's Symphony No. 5 in C Minor, Op. 67, Part I —> Classical Genre

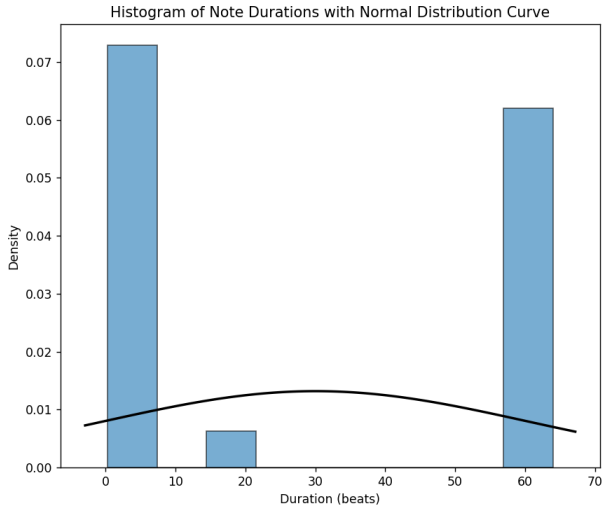


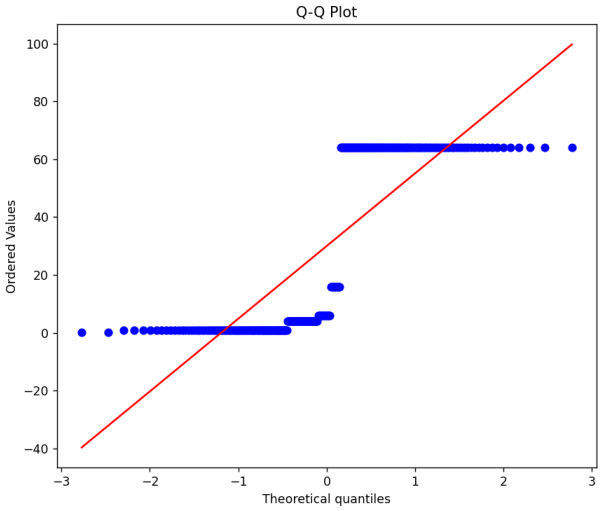


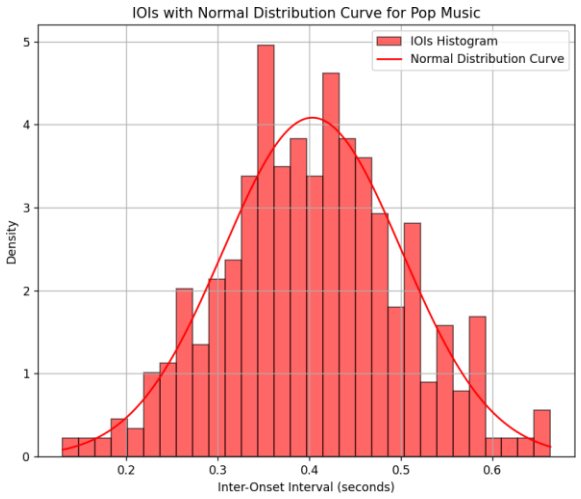


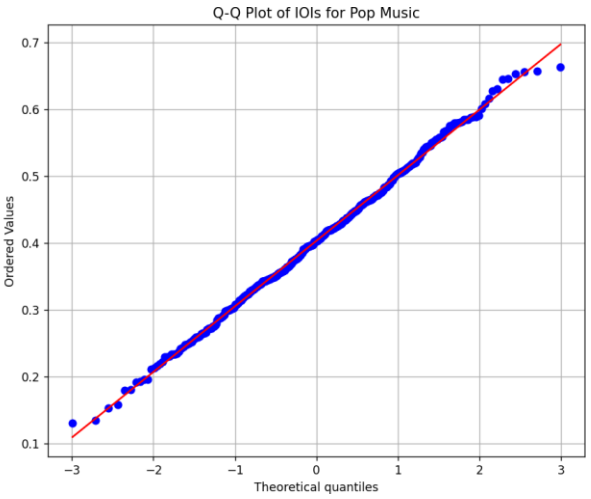


Blinding Lights("Blinding Lights") —> Pop Genre









CONCLUSION

By analysing various musical performances and compositions, we aimed to understand the underlying patterns and variations in these parameters, shedding light on the probabilistic nature of musical expression. Our findings suggest that while note durations and IOIs generally exhibit a tendency towards a normal distribution, there are noteworthy deviations which could have been influenced by factors such as the performance style of the human, genre, and the musician's intent. For example, using the graphs, we can determine the median of both note durations and inter-onset intervals for the five selected genres. The highest median for note durations has been displayed by the song "Songbird", which belongs to the jazz genre, giving insights into the probable note duration of the genre.

The normal distribution serves as a useful model for understanding general trends and variances in musical timing, however real-world data reveals additional complexities that may be due to human variability and artistic expression. This research highlights the importance of considering both statistical models and individual performance characteristics when analysing musical rhythm and flow, contributing to our understanding of musical articulation and creativity.

IMPLICATIONS

i. Performance Analysis: Understanding the normal distribution of note durations and IOIs can help musicians and composers refine their techniques and achieve desired expressive effects.
ii. Composition: Composers can use insights from statistical analysis to craft rhythmic patterns and structures that align with or deviate from traditional norms.
iii. This research may also be used to improve theoretical models on music composition or by AI music generating models to enhance music composition by improving existing machine learning algorithms on music analysis and aid the upcoming AI models in generating music.

EVALUATION

After a thorough assessment of our research paper and the approach to it, we arrived at the following judgement, identifying the merits and shortcomings of our research as well as future prospects for research.

STRENGTHS

This study uses a number of reliable statistical techniques, such as histograms, Q-Q plots, and normal distribution theory. All of these methods work together to ensure dependable and understandable results.
Data cleaning which was a part of the research methodology ensures maximum accuracy of note duration as well as inter-onset intervals (IOIs) values.
The use of a fair sampling method using number assignment followed by random number generation aims to reduce bias in the study, providing reliable results.

LIMITATION

The study's reliance on specific musical genres and data sources may limit the generalizability of the findings.
Further research could explore additional probability distributions or analyse a broader range of musical styles.
A larger sample size could have been used to arrive at more accurate results.

FUTURE RESEARCH

The use of other probability distributions to musical data can be researched.
We could explore the impact of different musical genres and cultures on the distribution of rhythmic parameters, how they are affected by genre, culture, region.

REFERENCE

[1] “Note Duration.” *musictheory.net*, https://www.musictheory.net/lessons/11 “Time point.” Wikipedia, https://en.wikipedia.org/wiki/Time_point.
[2] Queen. "We Will Rock You." News of the World, EMI Records, 1977.

[3] The Weeknd. "Blinding Lights." After Hours, XO and Republic Records, 2020. Fleetwood Mac. "Songbird." Rumours, Warner Bros. Records, 1977.
[4] Ustad Rashid Khan. "Raag Bhairav." Morning Ragas, Sony Music India, 1994.

[5] Bobbitt, Zach. “The Complete Guide: How to Interpret Q-Q Plots.” Statology, 19 January 2024, https://www.statology.org/qq-plot-interpretation/. Accessed 4 August 2024.
[6] “Histogram.” Wikipedia, https://en.wikipedia.org/wiki/Histogram. Accessed 4 August 2024. “Normal Distribution (Definition, Formula, Table, Curve, Properties & Examples).” BYJU'S,
[7] https://byjus.com/maths/normal-distribution/. Accessed 4 August 2024.