Statistical Analysis of Global Mental Illness Prevalence

(Vinay Chandra Konda, Shivani Atul Bhinge, Tejas Naik)

December 6, 2025

Abstract

This report details a statistical analysis of the global prevalence of main mental health disorders like Schizophrenia, Depressive, Anxiety, Bipolar, and Eating disorders from 1990 to 2019. Utilizing a public dataset, the study conducts exploratory data analysis, time series analysis, correlation heatmaps, and inferential testing (t-tests, confidence intervals) to address key questions regarding global patterns and inter-disorder relationships. The findings consistently show that **Depression and Anxiety** have the highest mean prevalence globally, with a slight, stable upward trend over the decades. A significant positive correlation between Anxiety, Bipolar, and Eating disorders was observed, while Depressive prevalence exhibits a notably high negative correlation with Schizophrenia in decadal averages. Inferential testing confirmed a **statistically significant mean difference** between Depression and Anxiety prevalence, with Anxiety being consistently higher. The results provide an evidence-based foundation for prioritizing public health resources globally.

1 Introduction and Background

Mental health disorders have become one of the most serious and rapidly growing health challenges of the 21st century. As global attention shifts toward overall well-being and economic productivity, understanding how mental illnesses are changing over time has become increasingly important. This study aims to statistically examine global patterns, relationships, and long-term trends in major mental health conditions including schizophrenia, depression, anxiety, bipolar disorder, and eating disorders across different countries from 1990 to 2019.

1.1 Research Questions

Our statistical analysis is designed to answer the following primary questions:

1. How the global prevalence rates of major mental illnesses evolved between 1990 and 2019?

2. What is the relation between different types of mental disorders wrt nature and strength?

3. Are the observed differences in mean prevalence between the most common disorders (*Depressive* and *Anxiety*) statistically significant?

2 Methods and Results

2.1 Exploratory Data Analysis (EDA)

Initial summary statistics provided a foundational understanding of the data:

- **High Prevalence: Anxiety** (Mean = 4.102%) and **Depressive** (Mean = 3.767%) disorders show the highest mean prevalence globally.

- **Low and Less Variable Prevalence: Schizophrenia** (Mean = 0.2666%) and **Eating** disorders (Mean = 0.1957%) have the lowest and least variable prevalence.

2.1.1 Distribution of Depression Prevalence

The distribution of Depression Prevalence was visualized using a histogram and density curve.
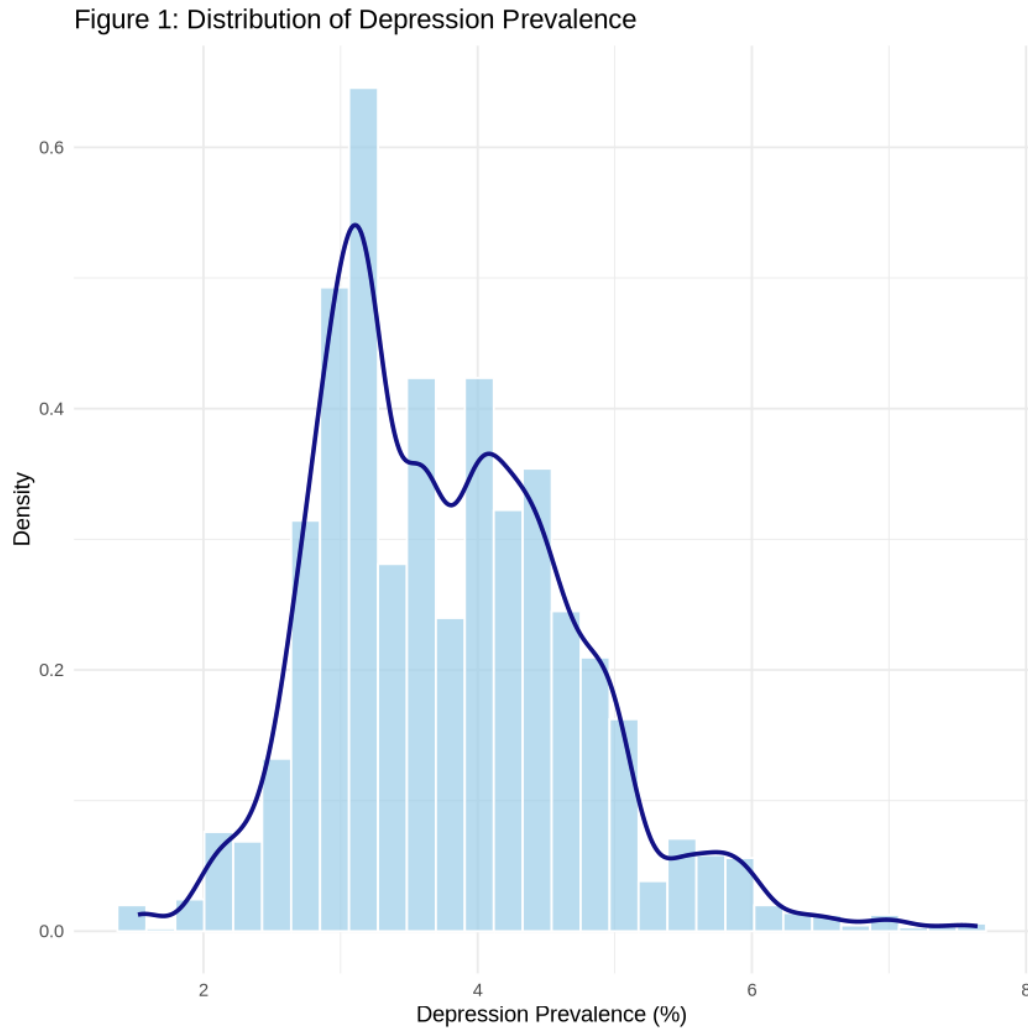

Figure 1: Distribution of Depression Prevalence

The distribution is **slightly right-skewed**, with most values clustering between 2% and 5%. The density curve exhibits a single main peak, indicating no extreme global outliers in the distribution of national depression rates.

2.1.2 Boxplot Comparison of Disorder Prevalence

A boxplot comparison across all five disorders visually confirms the descriptive statistics.
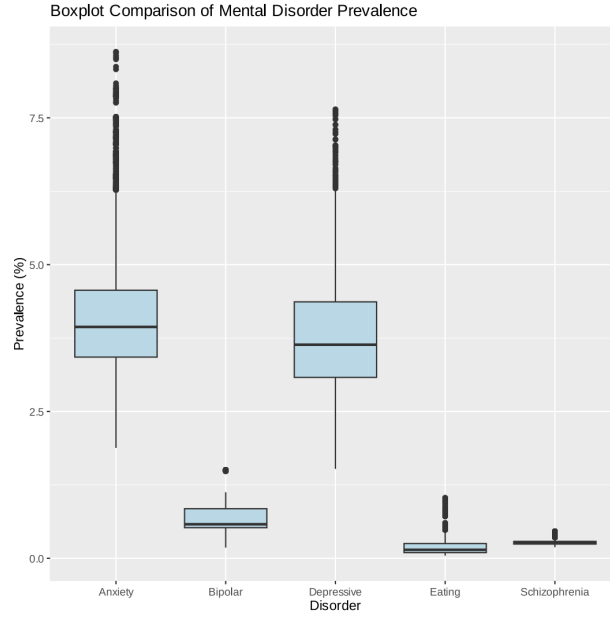
Figure 2: Boxplot Comparison of Mental Illness Prevalence

The boxplots clearly show that the median prevalence for Anxiety and Depressive disorders is significantly higher than that for Bipolar, Eating, and Schizophrenia disorders.

2.2   Temporal Analysis: Global Average Trend

The global average prevalence for each disorder was tracked over the years (1990-2019) to identify temporal trends.
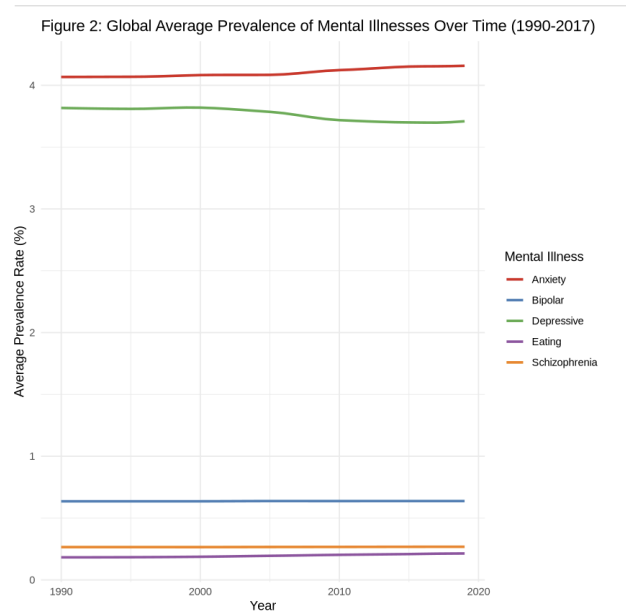

Figure 3: Global Average Prevalence of Mental Illnesses (Time Series Plot)

The results indicate that global trends for all disorders **remain relatively stable** over time. Anxiety and

Depressive disorders consistently maintain the highest prevalence, while Schizophrenia and Bipolar disorder remain low and stable. The plot suggests a very gradual, slight increase for the higher-prevalence disorders (Anxiety and Depressive) over almost three decades.

2.3   Geospatial Analysis: Top 10 Depression Prevalence

To investigate geographic variation, the top 10 countries with the highest average depression prevalence were identified.
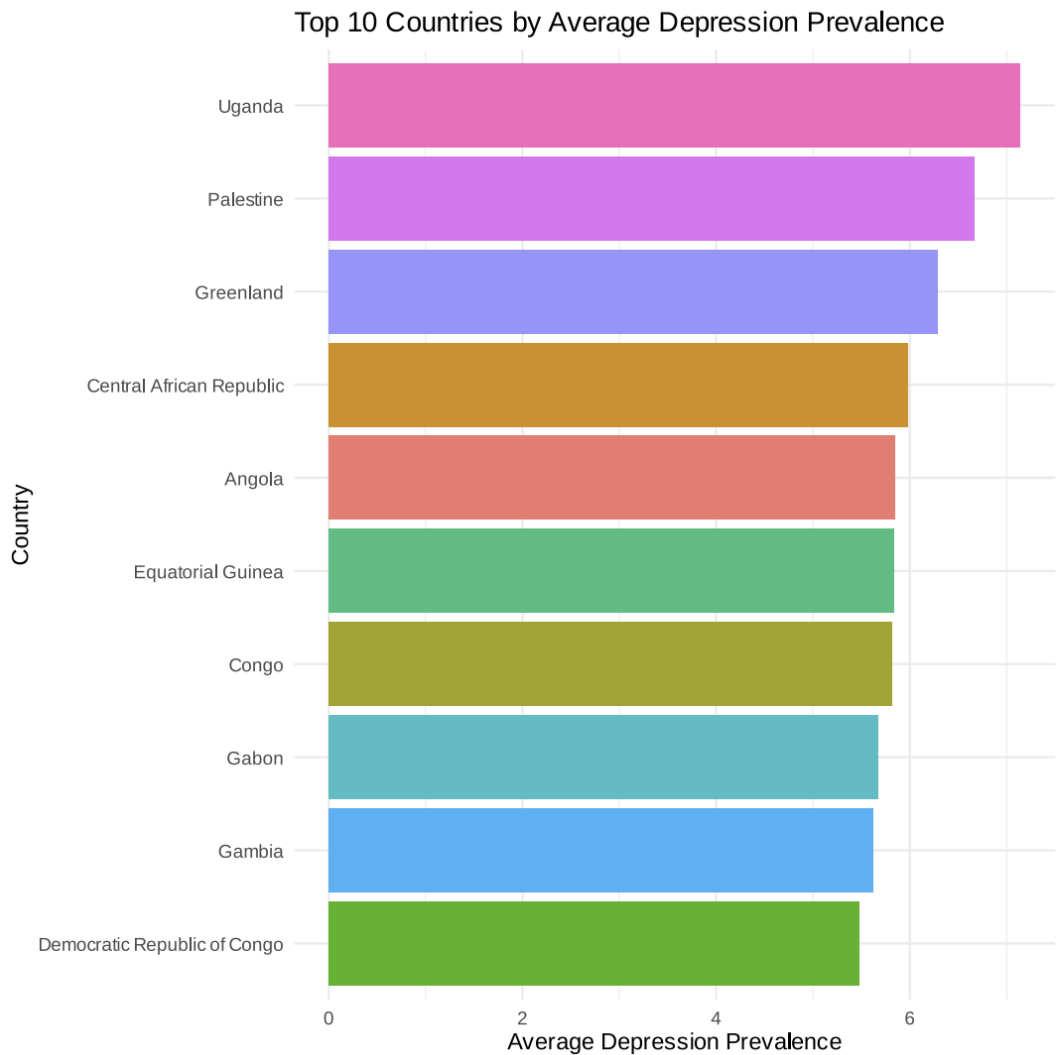

Figure 4: Top 10 countries with Highest Average Depression Prevalence

Nations like **Uganda** and **Palestine** are shown to have the highest average depression prevalence, potentially highlighting the influence of socio-political factors such as conflict, political instability, and underdeveloped healthcare on mental health outcomes.

2.4 Advanced Statistical Inference and Correlation

2.4.1 Correlation Heatmap of Mental Disorders

The pairwise correlations between all the five mental disorder prevalence rates were calculated and visualized to understand their co-occurrence and correlation.
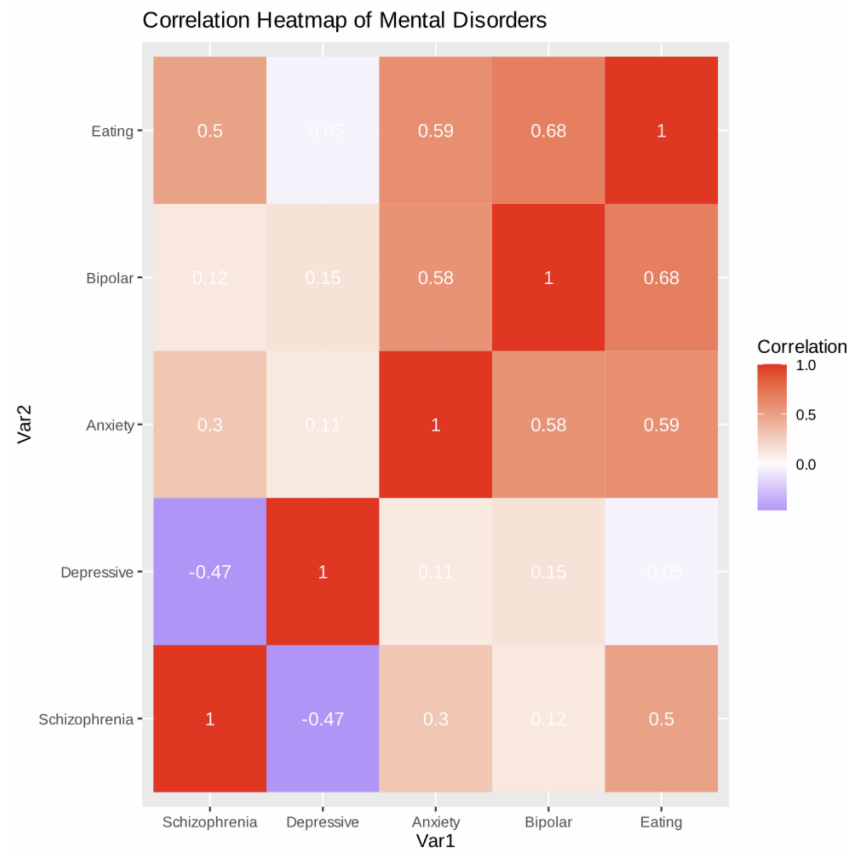

Figure 5: Correlation Heatmap of Mental Disorders Prevalence

- **Strongest Positive Correlation:** The prevalence of **Bipolar and Eating** disorders ($r = 0.68$) and **Anxiety and Eating** disorders ($r = 0.59$) show the strongest moderate-to-strong positive correlations. Depression and Anxiety also show a positive correlation ($r = 0.11$).

- **Negative Correlation:** A moderate negative correlation exists between **Schizophrenia and Depressive** disorders ($r = -0.47$).

- **Weakest Correlations: Schizophrenia** consistently exhibits the weakest correlations with other disorders.

2.4.2 Paired T-test (Inferential Technique 1)

To test the difference in mean prevalence between the two most common illnesses, a **paired t-test** was conducted.

- **Null Hypothesis (H_0):** $\mu_{\text{Depressive}} - \mu_{\text{Anxiety}} = 0$ (The true mean difference is zero).

- **Result:** The test yielded a t-statistic of $t = -20.352$ with a p-value $< 2.2e\text{-}16$.

- **Conclusion:** Since the p-value is extremely small (far less than $\alpha = 0.05$), we **reject the null hypothesis**. There is a statistically significant difference in the mean prevalence of Depression and Anxiety. The sample mean difference is -0.3348%, confirming that **Anxiety prevalence is consistently higher on average** than Depression prevalence.

2.4.3 Confidence Intervals and Bootstrapping (Inferential Technique 2)

A 95% Confidence Interval (CI) was calculated for the mean Depression Prevalence ($\mu_{\text{Depressive}}$) using both the standard T-method and the non-parametric Bootstrap Percentile Method.

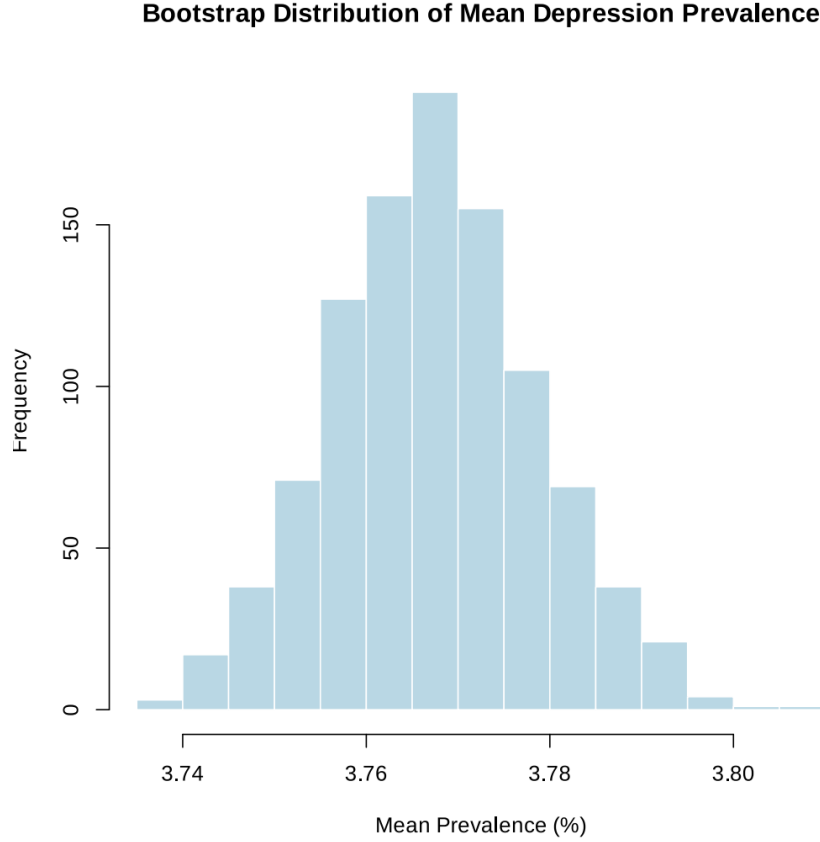**Bootstrap Distribution of Mean Depression Prevalence**

Figure 6: Bootstrap Distribution of Mean Depression Prevalence

- **T-Method CI:** $[3.7444\%, 3.7897\%]$.

- **Bootstrap CI (1000 replicates):** $[3.7462\%, 3.7905\%]$.

The close agreement between the two methods suggests that the estimated mean is **stable** and that the Central Limit Theorem (CLT) assumptions hold, as further supported by the nearly symmetric bootstrap distribution of sample means.

2.4.4 Variance Test (Inferential Technique 3)

An F-test was conducted to compare the variability or variance between Depression and Anxiety prevalence.

- **Result:** The test returned a variance ratio ($\frac{\sigma^2_{\text{Depressive}}}{\sigma^2_{\text{Anxiety}}}$) of $F = 0.77575$, with a p-value $< 2.2\text{e-}16$.

- **Conclusion:** The small p-value indicates that the ratio of variances is **significantly different from 1**. Since the ratio is less than 1, the variance of Depression prevalence ($\sigma^2_{\text{Depressive}}$) is significantly lower than that of Anxiety prevalence ($\sigma^2_{\text{Anxiety}}$).

2.4.5 Normal Distribution Fit for Depression Prevalence (Inferential Technique 4)

To further assess the normality assumption for the Depression Prevalence data, a histogram was generated with a theoretical normal distribution curve with the histogram. This method provides a visual check, complementing the findings from the bootstrap method.
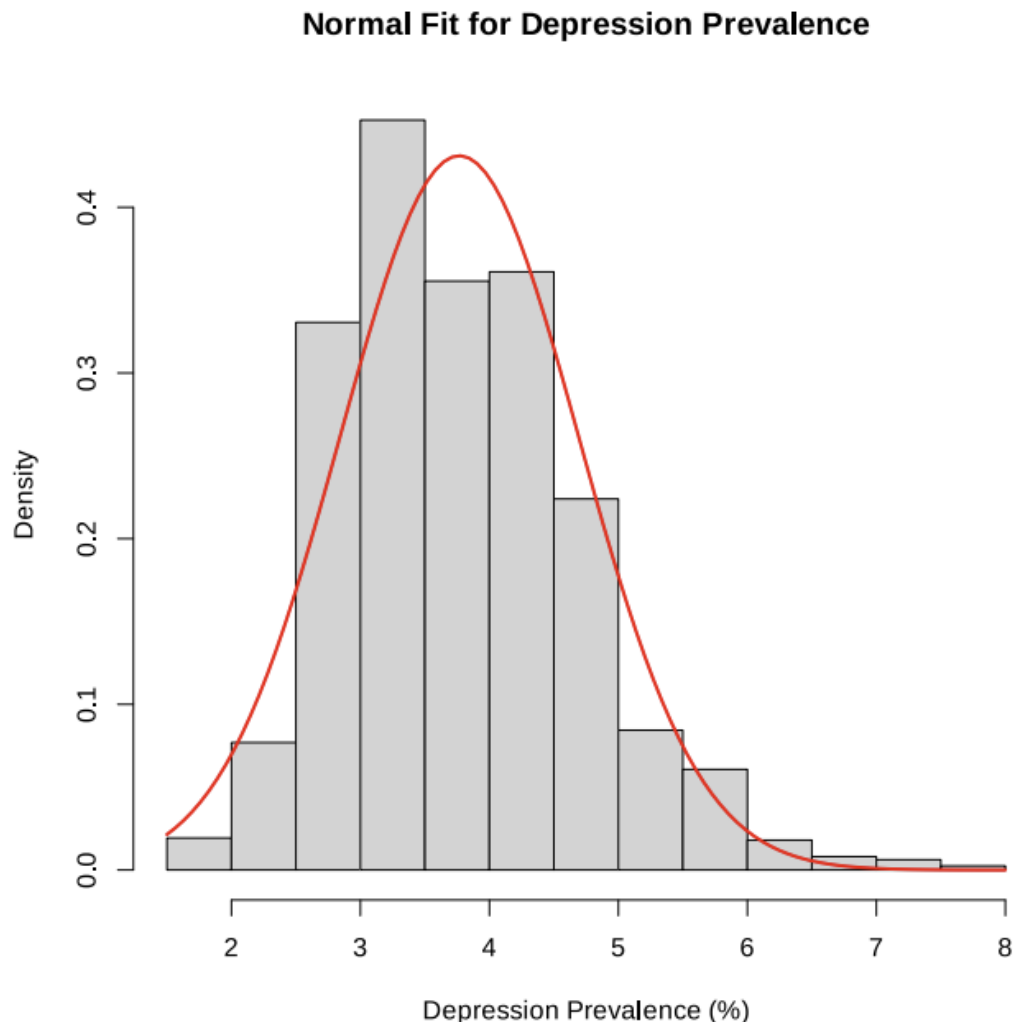

Figure 7: Normal Fit for Depression Prevalence

- From the plot we can see that red curve, representing the normal distribution defined by the sample mean and standard deviation, aligns well with the shape of the histogram.

- Hence this visual fit indicates that the distribution of global Depression Prevalence rates is **approximately normal**. This shows that the use of parametric inferential techniques, such as the T-method confidence interval and the paired T-test, on the Depression Prevalence data was appropriate.

2.4.6 QQ-Plot Analysis for Depression Prevalence (Inferential Technique 5)

To further analyze or validate the normality assumption of the Depression Prevalence data, a Quantile-Quantile (QQ) plot was generated. The QQ plot compares the quantiles of the sample data against the theoretical quantiles of a normal distribution, providing a visual diagnostic for normality.
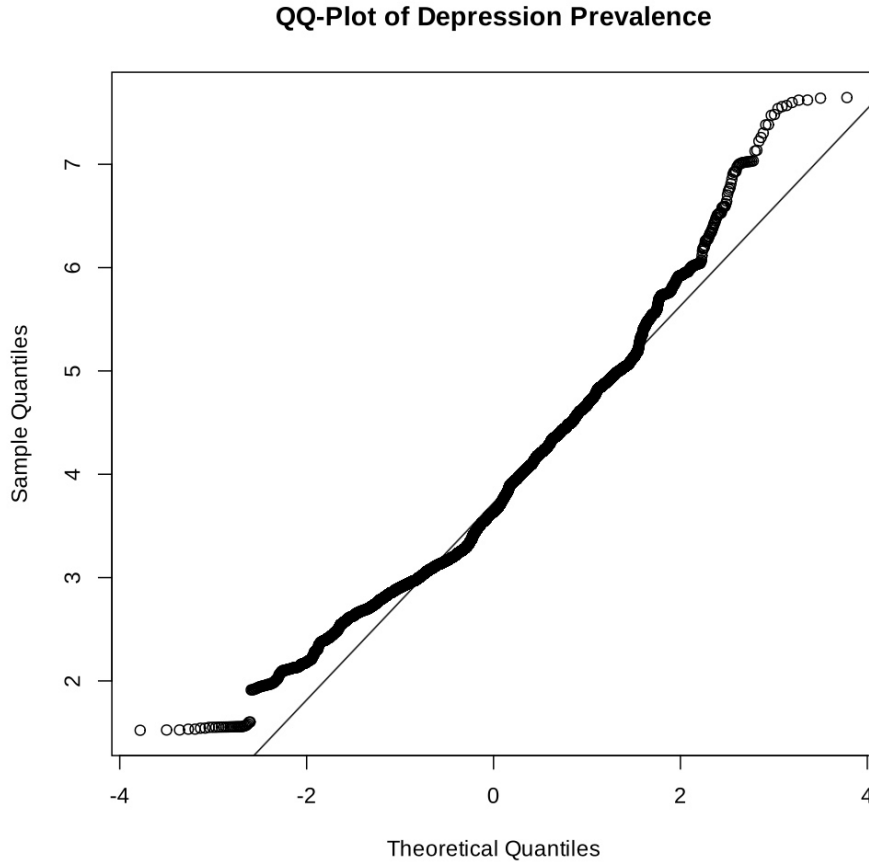
QQ-Plot of Depression Prevalence

Figure 8: QQ-Plot of Depression Prevalence

- From the QQ plot, it is evident that the points deviate noticeably from the straight reference line, particularly in the tails.

- This suggests that the data are not perfectly normally distributed, showing potential right skewness or the presence of outliers at higher prevalence values.

- Despite this deviation, the central portion of the data roughly follows the normal trend, indicating that the normality assumption holds reasonably well for the majority of observations.

- These findings complement the histogram-based analysis, together suggesting that while the Depression Prevalence distribution is approximately normal, it may exhibit slight non-normality in extreme values.

3 Conclusions

The statistical analysis provides a robust quantitative understanding of the global mental health landscape from 1990 to 2019.

3.1 Key Findings

- **Prevalence Dominance:** Depression (mean = 3.77%) and Anxiety (mean = 4.10%) are the most prevalent mental disorders globally.

- **Stability Over Time:** The global average prevalence rates for all five disorders have remained remarkably stable from 1990 to 2019, showing only a slight, gradual increase for Anxiety and Depression.

- **Correlation Structure:** A strong positive correlation exists among Bipolar, Eating, and Anxiety disorders, suggesting potential shared reporting factors. Schizophrenia prevalence appears largely independent of the other conditions.

- **Significant Difference:** The paired t-test confirmed that **Anxiety prevalence is statistically and significantly higher** than Depression prevalence, with a mean difference of approximately 0.33%.

4. References

1. Charlson FJ, Ferrari AJ, Santomauro DF, Diminic S, Stockings E, Scott JG, McGrath JJ, Whiteford HA. *Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019.* Lancet Psychiatry. 2022 Feb;9(2):137–150. doi: 10.1016/S2215-0366(21)00395-3. PMID: 35026139; PMCID: PMC8812183.

2. Liu J, Ning W, Zhang N, Zhu B, Mao Y. *Estimation of the Global Disease Burden of Depression and Anxiety between 1990 and 2044: An Analysis of the Global Burden of Disease Study 2019.* Healthcare (Basel). 2024 Aug 29;12(17):1721. doi: 10.3390/healthcare12171721. PMID: 39273745; PMCID: PMC11395616.

3. Polanczyk GV, Salum GA, Sugaya LS, Caye A, Rohde LA. *Worldwide prevalence and disability from mental and substance use disorders in children and youths: a systematic review and meta-analysis.* JAMA Psychiatry. 2024 Mar 1;81(3):256–268. doi: 10.1001/jamapsychiatry.2023.4611. PMID: 38310245.

4. Guo X, Chen Y, Li M, Zhao L, Wang H. *Global burden and trends of major mental disorders in the population aged under 24 years from 1990 to 2021: results from the Global Burden of Disease Study 2021.* Front Public Health. 2025;13:1635801. doi: 10.3389/fpubh.2025.1635801.

5. Zhang Y, Wu H, Chen X, Li Q, Wang Z. *Global burden of depression and anxiety disorders in older adults, 1990–2019: an age–period–cohort and temporal–spatial analysis.* J Affect Disord. 2025;360:45–55. doi: 10.1016/j.jad.2024.12.045. PMID: 39650218.

6. Steel Z, Marnane C, Iranpour C, Chey T, Jackson JW, Patel V, Silove D. *The global prevalence of common mental disorders: a systematic review and meta-analysis.* PLoS One. 2014 Apr 16;9(4):e92859. doi: 10.1371/journal.pone.0092859. PMID: 24740865; PMCID: PMC3997379.