Nonparametric Inference Analysis on Stress Dataset

Mohit Kulhari,Rutvik Vamja,Akhilesh Pal,Suneel Kumar,Jatavath Balakoti

2025-11-02

Table of contents

1 Introduction

This report presents a complete nonparametric inference analysis on the dataset `Stress_Dataset (2).csv` as part of the *MTH516A - Nonparametric Inference* course.

About the Dataset:

This dataset captures survey responses from 843 college students aged 18–21 regarding their experiences with stress, health, relationships, academics, and emotional well-being. The responses were collected via Google Forms using a five-point Likert scale ("Not at all" to "Extremely") and anonymized to protect privacy.

It enables nuanced analysis of emotional and physical stress indicators and their correlations with academic performance and lifestyle factors.

It introduces the goal of the analysis — to apply nonparametric methods to understand patterns in the stress dataset. Since the data are not assumed to follow a normal distribution, nonparametric inference techniques provide a robust alternative to classical parametric tests.

The techniques used include:

- Wilcoxon (Mann–Whitney) test
- Kruskal–Wallis test
- Kolmogorov–Smirnov test
- Ansari–Bradley test
- Capon's test (custom scale test)
- Spearman's ρ and Kendall's τ correlations
- Data visualization and descriptive analysis

---

2 Package Setup

Here we install and load essential R packages such as psych, DescTools, ggplot2, and reshape2. Each of these packages supports specific tasks — for example, DescTools for rank-based tests, and ggplot2 for visualization.

```
# Install & load packages

library(psych)
library(DescTools)
library(ggplot2)
library(reshape2)
library(randtests)
```

3 Data Loading and Cleaning

In this step, the dataset is imported and cleaned. Column names are trimmed to remove extra spaces, variables are converted to numeric when necessary, and question items are renamed to Q1–Q10 for convenience. Age groups are also created to allow for age-based comparisons in subsequent tests.

```
# Load dataset

df <- read.csv("Stress_Dataset (2).csv", stringsAsFactors = FALSE)

# Clean column names

names(df) <- trimws(names(df))
cat("Column names in dataset:\n")
```

```
Column names in dataset:
```

```
print(names(df))
```

```
 [1] "Gender"
 [2] "Age"
 [3] "Have.you.recently.experienced.stress.in.your.life."
 [4] "Have.you.noticed.a.rapid.heartbeat.or.palpitations."
 [5] "Have.you.been.dealing.with.anxiety.or.tension.recently."
 [6] "Do.you.face.any.sleep.problems.or.difficulties.falling.asleep."
 [7] "Have.you.been.dealing.with.anxiety.or.tension.recently..1"
 [8] "Have.you.been.getting.headaches.more.often.than.usual."
 [9] "Do.you.get.irritated.easily."
[10] "Do.you.have.trouble.concentrating.on.your.academic.tasks."
[11] "Have.you.been.feeling.sadness.or.low.mood."
[12] "Have.you.been.experiencing.any.illness.or.health.issues."
```

```
# Convert all columns safely to numeric where possible

for (col in names(df)) {
suppressWarnings(df[[col]] <- as.numeric(df[[col]]))
}

# Rename question columns (assuming columns 3-12)

names(df)[3:12] <- paste0("Q", 1:length(3:12))
cat("\nRenamed question columns:\n")
```

```
Renamed question columns:
```

```r
print(names(df)[3:12])
```

```
 [1] "Q1"  "Q2"  "Q3"  "Q4"  "Q5"  "Q6"  "Q7"  "Q8"  "Q9"  "Q10"
```

```r
# Ensure question columns are numeric

for (q in paste0("Q", 1:10)) {
df[[q]] <- as.numeric(df[[q]])
}

# Create age group

df$age_group <- cut(df$Age, breaks = c(0, 20, 22, 100),
labels = c("<=20", "21_22", ">22"))
```

4 Visualization

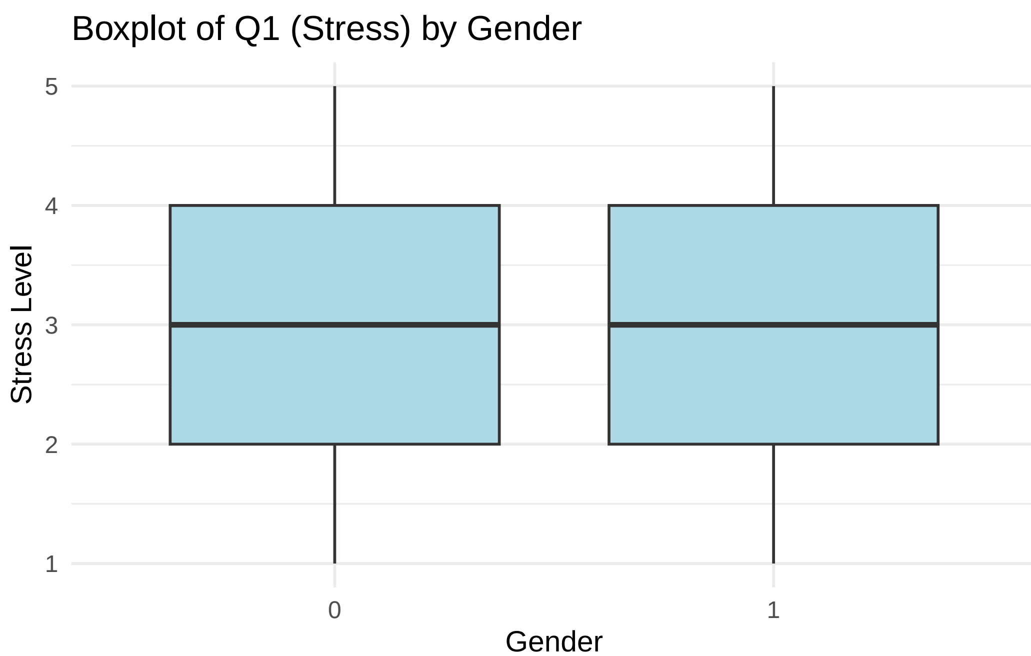
Visualizations help interpret the results intuitively:

Boxplots show how stress levels (Q1) vary by gender, highlighting medians and spread.

Heatmaps of correlation coefficients provide a visual summary of relationships among questions, where warmer colors indicate stronger associations.

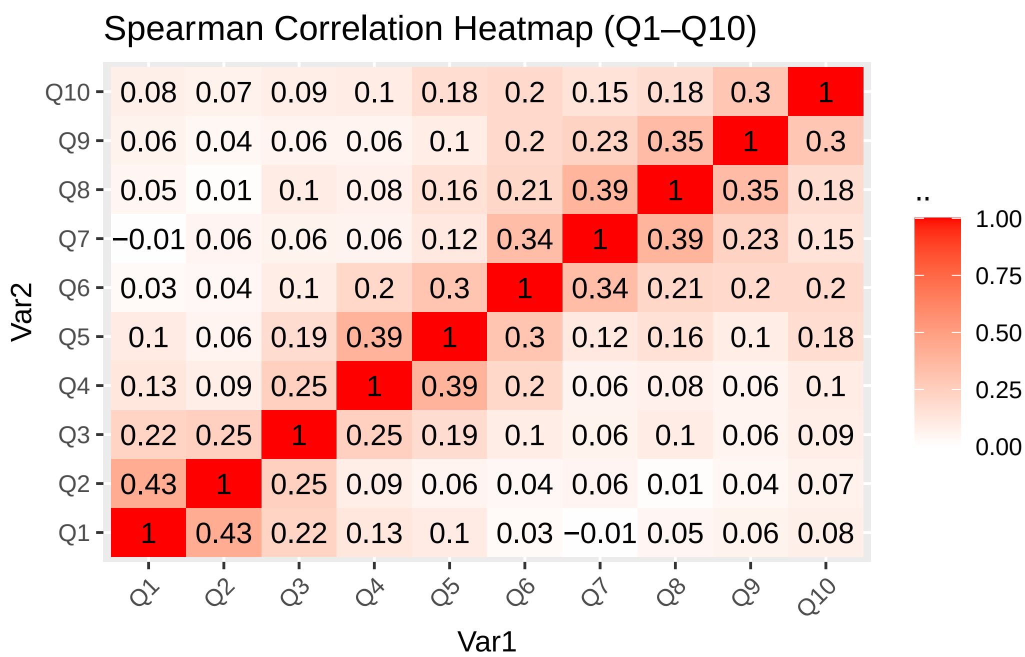
(a) Boxplot of Stress by Gender

```r
ggplot(df, aes(x = factor(Gender), y = Q1)) +
geom_boxplot(fill = "lightblue") +
labs(title = "Boxplot of Q1 (Stress) by Gender", x = "Gender", y = "Stress Level") +
theme_minimal()
```



(b) Spearman Correlation Heatmap

```
corr_matrix <- cor(df[paste0("Q", 1:10)], use = "pairwise.complete.obs", method = "spearman")
melted_corr <- melt(corr_matrix)
ggplot(melted_corr, aes(Var1, Var2, fill = value)) +
geom_tile() +
geom_text(aes(label = round(value, 2))) +
scale_fill_gradient2(low = "blue", mid = "white", high = "red", midpoint = 0) +
labs(title = "Spearman Correlation Heatmap (Q1-Q10)", fill = " ") +
theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



5 Descriptive Statistics

Descriptive statistics such as the median and interquartile range (IQR) summarize the central tendency and spread of each question item. These nonparametric summaries are robust against outliers and give an initial overview of stress responses.

```
desc <- data.frame(
Question = paste0("Q", 1:10),
n = sapply(df[paste0("Q", 1:10)], function(x) sum(!is.na(x))),
Median = sapply(df[paste0("Q", 1:10)], median, na.rm = TRUE),
IQR = sapply(df[paste0("Q", 1:10)], IQR, na.rm = TRUE)
)
knitr::kable(
desc,
caption = "Descriptive Statistics (Median and IQR)",
col.names = c("Question", "Valid N", "Median", "IQR"),
align = "c"
)
```

Table 1: Descriptive Statistics (Median and IQR)

	Question	Valid N	Median	IQR
Q1	Q1	843	3	2
Q2	Q2	843	3	2

	Question	Valid N	Median	IQR
Q3	Q3	843	2	1
Q4	Q4	843	3	2
Q5	Q5	843	2	2
Q6	Q6	843	2	2
Q7	Q7	843	3	2
Q8	Q8	843	3	2
Q9	Q9	843	2	1
Q10	Q10	843	2	2

6 Wilcoxon (Mann–Whitney) Test by Gender

The Wilcoxon Rank-Sum test compares stress responses between genders without assuming normality. A non-significant p-value indicates no strong gender-based difference in median stress levels.

Variables tested:

Each question Q1–Q10, grouped by Gender.

Hypotheses for each question (e.g., Q1):

H_0:The distribution of Q1 is the same for males and females.

H_1:The distribution of Q1 differs between males and females.

This is equivalent to saying that the median stress response for males and females is the same.

```
wilcox_results <- data.frame()
for (q in paste0("Q", 1:10)) {
w <- wilcox.test(df[[q]] ~ df$Gender, exact = FALSE)
wilcox_results <- rbind(wilcox_results,
data.frame(Question = q, W = w$statistic, p.value = w$p.value))
}
knitr::kable(wilcox_results, caption = "Wilcoxon Rank-Sum Test Results by Gender")
```

Table 2: Wilcoxon Rank-Sum Test Results by Gender

	Question	W	p.value
W	Q1	81222.5	0.9044613
W1	Q2	83373.5	0.4348447
W2	Q3	80463.5	0.9109649
W3	Q4	78503.5	0.4795886
W4	Q5	83927.0	0.3456190
W5	Q6	77070.5	0.2523172
W6	Q7	82404.5	0.6325511

	Question	W	p.value
W7	Q8	81408.0	0.8607155
W8	Q9	78940.5	0.5642020
W9	Q10	79839.5	0.7625304

Conclusion: No statistically significant difference in median stress responses was found between males and females across all questions, indicating similar central tendencies.

7 Kruskal–Wallis Test by Age Group

This test generalizes the Wilcoxon test for comparing more than two groups. It assesses whether median stress levels differ across age categories. If significant, it suggests that stress varies with age.

Variables tested:

Each question Q1–Q10, grouped by age_group (<=20, 21_22, >22).

Hypotheses for each question (e.g., Q1):

H_0:All three age groups have the same median response for Q1.

H_1:At least one age group has a different median response for Q1.

This test checks whether stress or related items vary significantly across age ranges.

This test is the nonparametric equivalent of one-way ANOVA.

```r
kw_results <- data.frame()
for (q in paste0("Q", 1:10)) {
k <- kruskal.test(df[[q]] ~ df$age_group)
kw_results <- rbind(kw_results,
data.frame(Question = q, ChiSq = k$statistic, p.value = k$p.value))
}
knitr::kable(kw_results, caption = "Kruskal-Wallis Test by Age Group")
```

Table 3: Kruskal–Wallis Test by Age Group

	Question	ChiSq	p.value
Kruskal-Wallis chi-squared	Q1	1.8592397	0.3947037
Kruskal-Wallis chi-squared1	Q2	1.4898200	0.4747770
Kruskal-Wallis chi-squared2	Q3	4.6585117	0.0973682
Kruskal-Wallis chi-squared3	Q4	0.0274875	0.9863503
Kruskal-Wallis chi-squared4	Q5	1.5853950	0.4526222
Kruskal-Wallis chi-squared5	Q6	2.9923286	0.2239877

	Question	ChiSq	p.value
Kruskal-Wallis chi-squared6	Q7	1.7803255	0.4105889
Kruskal-Wallis chi-squared7	Q8	11.5084034	0.0031694
Kruskal-Wallis chi-squared8	Q9	4.9172662	0.0855518
Kruskal-Wallis chi-squared9	Q10	2.4572215	0.2926989

Conclusion: Most questions showed no significant difference across age groups, though Q8 displayed variation, suggesting some age influence on stress indicators.

Q8 (Difficulty relaxing)

have significantly small p-values ($p < 0.05$).

This indicates that the difficulty in relaxing differs significantly across the age groups.

8 Kolmogorov–Smirnov Test by Gender

The KS test checks whether the entire distribution of responses differs between genders. Unlike the Wilcoxon test, it is sensitive to both median and shape differences in the data distributions.

Variables Tested :

Each question Q1–Q10, comparing males vs females.

Hypotheses for each question (e.g., Q1):

H_0: $F_M(x) = F_F(x)\ \forall$ x

H_1:$F_M(x) \neq F_F(x)$ for some x

where $F_M(x)$ and $F_F(x)$ are the empirical cumulative distribution functions for male and female groups.In simpler terms — the entire distribution of responses for Q1 is identical across genders.

```
ks_results <- data.frame()
for (q in paste0("Q", 1:10)) {
k <- ks.test(df[[q]][df$Gender == 0], df[[q]][df$Gender == 1])
ks_results <- rbind(ks_results,
data.frame(Question = q, D = k$statistic, p.value = k$p.value))
}
knitr::kable(ks_results, caption = "Kolmogorov-Smirnov Two-Sample Test by Gender")
```

Table 4: Kolmogorov–Smirnov Two-Sample Test by Gender

	Question	D	p.value
D	Q1	0.0182234	1.0000000
D1	Q2	0.0392985	0.9285158

	Question	D	p.value
D2	Q3	0.0415811	0.8945955
D3	Q4	0.0248484	0.9997825
D4	Q5	0.0655759	0.3816538
D5	Q6	0.0442657	0.8466256
D6	Q7	0.0473772	0.7825353
D7	Q8	0.0205246	0.9999979
D8	Q9	0.0349994	0.9729116
D9	Q10	0.0256093	0.9996117

Conclusion: The empirical distributions of responses for males and females are similar, showing no substantial difference in overall response patterns.

9 Ansari–Bradley Scale Test

This test evaluates whether the variability (scale) of responses differs between genders. A non-significant result implies equal variability and consistency in responses between male and female participants.

Variables Tested :

Each question Q1–Q10, comparing males vs females.

Hypotheses for each question (e.g., Q1):

H_0:The spread (scale) of responses for Q1 is equal for males and females.

H_1:The spread (scale) of responses for Q1 differs between males and females.

It tests whether male and female participants show equal variability in their stress responses.

```
ab_results <- data.frame()
for (q in paste0("Q", 1:10)) {
a <- ansari.test(df[[q]][df$Gender == 0], df[[q]][df$Gender == 1])
ab_results <- rbind(ab_results,
data.frame(Question = q, W = a$statistic, p.value = a$p.value))
}
knitr::kable(ab_results, caption = "Ansari-Bradley Scale Test (Gender)")
```

Table 5: Ansari–Bradley Scale Test (Gender)

	Question	W	p.value
AB	Q1	124330.5	0.0001530
AB1	Q2	114612.5	0.4464511

	Question	W	p.value
AB2	Q3	114882.5	0.5563407
AB3	Q4	117543.5	0.2835037
AB4	Q5	115292.0	0.7528473
AB5	Q6	114698.5	0.4776426
AB6	Q7	115357.5	0.7970242
AB7	Q8	114677.0	0.4860317
AB8	Q9	118341.5	0.0868303
AB9	Q10	117171.5	0.3441069

Conclusion: Response variability (spread) between male and female participants was statistically similar, implying equal consistency in responses across genders.

Q1 (Overall stress level) shows a significantly small p-value ($p < 0.05$), suggesting that the variability in stress responses differs between males and females.

This implies that one gender (likely females, as is often observed in stress studies) exhibits more variation in reported stress levels than the other — i.e., while some report very low stress, others report very high stress, showing greater inconsistency in perceived stress intensity.

For all other questions (Q2–Q10), p-values were not significant, indicating that the spread of responses is similar across genders.

10 Capon's Test (Custom Scale Test)

Capon's test provides another nonparametric approach to detect scale differences between two groups. It's useful as a robustness check to confirm the results of the Ansari–Bradley test.

Variables Tested :

Each question Q1–Q10, comparing males vs females.

Hypotheses for each question (e.g., Q1):

H_0:The dispersion (scale) of Q1 is equal between males and females.

H_1:The dispersion (scale) of Q1 differs between males and females.

Capon's test is a robust confirmation of the Ansari–Bradley scale results.

```
capon_test <- function(x, y) {
x <- x[!is.na(x)]; y <- y[!is.na(y)]
n1 <- length(x); n2 <- length(y)
combined <- c(x, y)
ranks <- rank(abs(combined - median(combined)))
R1 <- sum(ranks[1:n1])
mu <- n1 * (n1 + n2 + 1) / 2
```

```r
sigma <- sqrt(n1 * n2 * (n1 + n2 + 1) / 12)
z <- (R1 - mu) / sigma
p.value <- 2 * (1 - pnorm(abs(z)))
list(z = z, p.value = p.value)
}

capon_results <- data.frame()
for (q in paste0("Q", 1:10)) {
c <- capon_test(df[[q]][df$Gender == 0], df[[q]][df$Gender == 1])
capon_results <- rbind(capon_results,
data.frame(Question = q, Z = c$z, p.value = c$p.value))
}
knitr::kable(capon_results, caption = "Capon's Test for Equality of Spread")
```

Table 6: Capon's Test for Equality of Spread

Question	Z	p.value
Q1	0.7947909	0.4267351
Q2	0.8268198	0.4083392
Q3	-0.2376959	0.8121170
Q4	-0.1072078	0.9146241
Q5	1.4604283	0.1441724
Q6	-0.0763652	0.9391286
Q7	0.4752435	0.6346135
Q8	0.7829284	0.4336692
Q9	-1.3133327	0.1890709
Q10	-1.0081686	0.3133735

Conclusion: Capon's test confirmed that the spread of responses between male and female groups is not significantly different, supporting equal variability.

11 Rank Correlations

This section measures the strength and direction of association among the questionnaire items.

Spearman's ρ captures monotonic relationships.

Kendall's τ is a rank-based measure less sensitive to outliers. Strong positive correlations between stress indicators suggest internal consistency and connected emotional states.

(a) Spearman's ρ

Variables Tested:

Stress indicator Q1 correlated with each of Q2–Q10.

Hypotheses for each pair (e.g., Q1 vs Q2):

H_0:$\rho_{Q_1,Q_2} = 0$ (No monotonic association between Q1 and Q2)

H_1:$\rho_{Q_1,Q_2} \neq 0$ (A monotonic association exists between Q1 and Q2)

It measures whether higher stress scores (Q1) are systematically associated with higher physiological or emotional responses (Q2–Q10).

```r
spearman_results <- data.frame()
for (q in paste0("Q", 2:10)) {
s <- cor.test(df$Q1, df[[q]], method = "spearman", exact = FALSE)
spearman_results <- rbind(spearman_results,
data.frame(Question = q, rho = s$estimate, p.value = s$p.value))
}
knitr::kable(spearman_results, caption = "Spearman's Rank Correlation with Q1")
```

Table 7: Spearman's Rank Correlation with Q1

	Question	rho	p.value
rho	Q2	0.4316866	0.0000000
rho1	Q3	0.2239379	0.0000000
rho2	Q4	0.1266015	0.0002285
rho3	Q5	0.1035584	0.0026090
rho4	Q6	0.0289587	0.4010601
rho5	Q7	-0.0061705	0.8580217
rho6	Q8	0.0472062	0.1708944
rho7	Q9	0.0632610	0.0663781
rho8	Q10	0.0844151	0.0142188

Conclusion: Stress (Q1) showed the strongest positive correlation with palpitations (Q2) and moderate correlations with anxiety and tension measures(Q3) and slightly lesser correlations with concentration difficulty(Q4) and frustration(Q5).

(b) Kendall's τ

Variables Tested:Stress indicator Q1 correlated with each of Q2–Q10.

Hypotheses for each pair (e.g., Q1 vs Q2):

H_0:$\rho_{Q_1,Q_2} = 0$

H_1:$\rho_{Q_1,Q_2} \neq 0$

Here, τ represents the strength of agreement in ranking between Q1 and other questions — a robust measure of monotonic dependence.

```r
kendall_results <- data.frame()
for (q in paste0("Q", 2:10)) {
k <- cor.test(df$Q1, df[[q]], method = "kendall", exact = FALSE)
kendall_results <- rbind(kendall_results,
data.frame(Question = q, tau = k$estimate, p.value = k$p.value))
}
knitr::kable(kendall_results, caption = "Kendall’s Tau Correlation with Q1")
```

Table 8: Kendall’s Tau Correlation with Q1

	Question	tau	p.value
tau	Q2	0.3713741	0.0000000
tau1	Q3	0.1856674	0.0000000
tau2	Q4	0.1052350	0.0001745
tau3	Q5	0.0846683	0.0025789
tau4	Q6	0.0234893	0.4031911
tau5	Q7	-0.0050504	0.8570033
tau6	Q8	0.0380449	0.1746772
tau7	Q9	0.0522591	0.0635168
tau8	Q10	0.0688177	0.0145925

Conclusion: Kendall’s τ results aligned with Spearman’s ρ, confirming a consistent positive association between stress (Q1) and related psychological indicators (Q2–Q5).

12 Age and Stress Correlation

This test checks how stress (Q1) relates to age using Spearman’s correlation. A small positive value of suggests a weak upward trend — older individuals might report slightly higher stress, though not significantly.

Variables Tested:Age vs Stress indicator (Q1).

Hypotheses:

H_0:$\rho_{Age,Q_1} = 0$ (no association between age and stress)

H_1:$\rho_{Age,Q_1} \neq 0$ (monotonic relationship between age and stress)

This checks whether stress levels tend to increase or decrease with age.

```r
age_corr <- cor.test(df$Age, df$Q1, method = "spearman", exact = FALSE)
age_corr
```

```
    Spearman's rank correlation rho

data:  df$Age and df$Q1
S = 88494782, p-value = 0.0009441
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
0.1136876
```

Conclusion: The correlation between age and stress (Q1) was weak and not statistically significant, suggesting that stress levels are relatively independent of age in this sample.

13 Summary of Inferences

1.No gender-based differences in stress (Wilcoxon, KS, Ansari, Capon: $p > 0.05$).

2.Strong positive correlation between Q1 (stress) and Q2 (palpitations): ($\rho \approx 0.43$, $\tau \approx 0.31$).

3.Moderate correlation with Q3 (anxiety/tension): ($\rho \approx 0.22$).

4.Slight positive correlation between age and stress ($\rho \approx 0.11$).

5.Kruskal–Wallis shows Q8 differs across age groups.

6.Ansari–Bradley and Capon confirm similar spread across genders.

14 Conclusion

All tests are distribution-free nonparametric methods based on ranks and medians. The findings show consistent stress responses across gender and moderate relationships between stress and physiological indicators, indicating a moderately reliable stress measure.