Mikailalsys Journal of Mathematics and Statistics

e-ISSN : 3030-816X
p-ISSN : 3030-8399

Index : Harvard, Boston, Sydney University, Dimensions, Lens, Scilit, Semantic, Google, etc

https://doi.org/10.58578/MJMS.v3i3.7453

Multivariate Approaches to Neonatal Assessment of Newborn Babies

B. A. Garba, A. M. Baba, M. Y. Adamu
Abubakar Tafawa Balewa University, Bauchi, Nigeria
isahabdullahi7474@gmail.com

Article Info:

Submitted:	Revised:	Accepted:	Published:
Sep 15, 2025	Oct 4, 2025	Oct 16, 2025	Oct 21, 2025

Abstract

This study examines gender-based differences in neonatal physical characteristics using multivariate statistical techniques. A total of 1,000 newborns (male and female) were sampled from the Federal University Wukari Teaching Hospital, Taraba State, Nigeria. Key anthropometric variables measured included occipito-frontal circumference (OFC), cranial circumference (CC), length of birth (LOB), abdominal circumference (AC), and weight (WT). Due to perfect correlation with other variables, AC was excluded from the multivariate analysis. The objective was to determine whether statistically significant physical differences exist between male and female neonates at birth. The study employed Hotelling's T^2 test and profile analysis; however, the assumptions of homogeneity of covariance matrices (tested via Box's M) and independence (assessed via scatter plots) were violated. To address these issues, a robust non-parametric permutation-based Hotelling's T^2 test was conducted, yielding a statistically significant result ($p < 0.001$), indicating notable gender-based differences in multivariate mean vectors. While the main effect of Feature was highly significant ($p < 0.001$), revealing differences among OFC, CC, LOB, and WT, the Gender × Feature interaction was non-significant ($p > 0.05$), suggesting parallel measurement

patterns across genders. The study concludes that gender significantly influences neonatal physical traits and that advanced multivariate methods, including Hotelling's T^2 and profile analysis, are effective for analyzing high-dimensional neonatal data—even under violations of classical assumptions such as normality and homoscedasticity.

Keywords: Neonatal Anthropometry; Gender Differences; Hotelling's T-Square; Profile Analysis; Multivariate Statistics

INTRODUCTION

Birth is the first, shortest, most dangerous and eagerly anticipated journey of a human being. The experiences of the neonate in the birth process reflect heavily on the immediate condition and have a bearing on survival and long-term development. The subject of neonatal assessment wa largely neglected because the main focus of attention was on maternal health. On this background of a general lack of interest or discipline in the evaluation of the neonate emerged Virginia Apgar in 1949, she began studying obstetric anesthesia and its effect on the neonatal condition. She formulated a list of the objectives signs which pertained to the condition of the infant at birth and selected five signs which could be determined easily with minimal intervention and with reasonable speed. This was the first standardized method for evaluating the newborns transition to life outside the womb. The five points are: heart rate, respiratory effort, muscle tone, reflex response, and color. The letter of her name was used as a mnemonic device for the five scoring criteria. Neonatal assessment of the newborn baby is performed immediately after birth to detect significant abnormalities, birth injuries, and cardio respiratory disorders that may compromise a successful adaptation to extra uterine life. Each newborn baby is carefully checked at birth for signs of problems or complications. And this assessment is undertaking by a clinician (health care provider) who have successfully attained health maternity clinical competency in newborn examination. Other components of assessing the newborn after the resuscitation period are weight, length, occipital frontal circumference, abdominal circumference, temperature, umbilical stump, and feeding. Common abnormalities detected in the newborn examinations are: capillary or macular haemangioma, blue-black pigmented area, urticaria of the newborn, breast enlargement, white pimples and oral cysts.

According to the World Health Organization [WHO], (2019), 2.5 million neonates died in 2018, accounting for a staggering 7,000 newborn deaths daily, amounting to 47% of all child death under 5 years of age. The majority 75% of neonatal death occurs during 1st week of life and about 1 million dying on the first day of life. In Nepal, Neonatal Mortality rate is 21 deaths per 1000 live births according to Nepal Demographic and Health Survey [NDHS], (2016). Therefore, it is of paramount importance to assess the health status of the neonate during the birth procedure using validated methods such as the APGAR scoring system.

Multivariate analysis is essentially the statistical process of simultaneously analyzing multiple independent (or predictor) variables with multiple dependent (outcome or criterion) variables using matrix algebra. While these analyses have being a part of statistics since the early 1900' s, the development of mainframe and microcomputers and subsequent analytical software has made the once tedious calculations fairly simple and very fast.

Literature Review Neonatal Assessment Factors

1. Oranye et al. (2020) emphasize that newborn babies are considered a precious gift from God. However, it is important to note that the appearance of a newborn baby is not based on any specific model. Instead, the physical characteristics of a newborn baby, such as weight, height, and head circumference, can vary from one baby to another. To investigate this further, the researchers conducted a study using a sample of 200 male and female babies. The main objective was to determine if there were any significant differences in the means of the variables being considered. The study focused on three variables for both male and female babies at birth. The results revealed that the mean birth weight was 3.55kg for male babies and 3.39kg for female babies.
2. A study conducted by Rosnah S. *et al.*, (2018) aimed to investigate the relationship between head circumference and microcephaly among term infants born in a teaching hospital in Malaysia from 2011 to 2015. The researchers utilized a cross-sectional study design using electronic birth census data, with the independent variables being mothers' age and height, parity, birth weight, and birth length. The study included all term newborns, both alive and stillbirth, with gestational weeks ranging from 37 to 41 and a birth weight of at least 500 g. A total of 26,503 newborns met the inclusion criteria,

with 13,655 males and 12,840 females. The mean head circumferences for male and female newborns were 32.93 cm and 32.56 cm, respectively. Interestingly, the average head circumference for Malaysian newborns was found to be smaller than the World Health Organization Standard Growth Chart for Term Infant. Furthermore, the study identified an increasing trend of microcephaly over the years, with low birth weight being noted as the main predictor of microcephaly.

3. Jose V. *et al.*, (2021) published international growth standards for children under 5 years in 2006, which have gained global acceptance. The INTERGROWTH-21st Project aimed to complement these standards by developing international norms for fetuses, newborns, and the postnatal growth phase of preterm infants. This project, based on population data, evaluated fetal growth and newborn size in eight urban populations with well-met health and nutrition needs, adequate prenatal care, and no significant environmental growth constraints. As part of the Newborn Cross-Sectional Study (NCSS) within the INTERGROWTH-21st Project, weight, length, and head circumference were measured in all newborns, along with prospective data collection for pregnancy and the perinatal period. To establish newborn standards, pregnancies were selected based on strict individual eligibility criteria for a low-risk population for impaired fetal growth (referred to as the NCSS prescriptive subpopulation), in addition to the underlying population characteristics.
4. Hermano A. L. *et al.*, (2021) the first 1000 days of life are a critical period when the foundations of child development and growth are established. Few studies in Latin America have examined the relationship of birth outcomes and neonatal care factors with development outcomes in young children. We aimed to assess the association between pregnancy and neonatal factors with children's developmental scores in a cross-sectional, population-based study of children in Ceará, Brazil.

Multivariate Techniques

Christopher D.D *et al.*,t (2018) describe profile analysis as a technique used in multivariate data analysis. It can be seen as an extension of MANOVA, specifically focusing on test scores obtained from educational or psychological assessments. The analysis aims to identify differences in sub scores on tests commonly used in medical, psychological, and educational studies to rank participants based on a specific construct. Practitioners are interested in quantifying both the overall performance of individuals on the subtests

(referred to as the level) and the variation between scores on the subtests (referred to as the pattern). To facilitate this analysis, the profile R package for the R programming language provides a range of procedures. These procedures allow for the decomposition of observed scores into level and pattern effects.

P.O. Adebayo *et al.*, (2018) conducted a study focusing on finding an alternative approach to the multivariate Behrens-Fisher problem. They proposed a new procedure that utilized an approximate degree of freedom test, which was adapted from the Satterthwaite univariate procedure. The effectiveness of this proposed procedure was compared to six existing procedures, namely Johanson, Yao, Krishnamoorthy, Hotelling T square, Nel and Van der Merwe, and Yanagihara, using both simulation and real-life data from Tim (1975). The results revealed that the proposed procedure outperformed all the existing procedures in terms of test power across various scenarios involving different random variables (p), variance covariance matrices, sample sizes, and significant levels (α). Furthermore, the proposed procedure demonstrated competitive performance in terms of type I error rate when compared to Johanson, Yao, Krishnamoorthy, Nel, and Van der Merwe.

The findings of the study, as revealed by the Multivariate Analysis of Variance (MANOVA) results, indicate that Senior Lecturers received the highest ratings, followed by Lecturers and then Assistant Lecturers. This suggests that the ranks of the teaching staff significantly influence their performance across the various departments of the faculty. The paper also discusses recommendations and implications for the management of Higher Institutions of Learning (HIL). Overall, this research contributes to the existing literature on the supervision and evaluation of teaching staff performance in the HIL context.

METHODOLOGY

The data for this research work will be a secondary data from Federal University Wukari Teaching Hospital (FUWTH), Taraba State. Which propose collecting data from a sample of 1,000 newborns, including both males and females, which measure five variables: chest circumference, occipital frontal circumference, length, abdominal circumference, and weight. These measurements are obtained during neonatal assessments using precise weighing and measuring equipment to ensure accuracy and reliability in evaluating the health status of the newborns. The research work will utilize the Multivariate Sample T-test approach, specifically employing the Hotelling's T^2 statistic and profile analysis. Before

conducting the analysis, it is essential to test the basic assumptions underlying the use of the test statistic. These assumptions include the multivariate normality test, equality of variance-covariance matrix, and independence test.

Model Assumption for Multivariate Two Sample -Test

The basic assumptions are stated below;

Multivariate normality: Assumes that the underlying data for each group is normal. To test the assumption of normality, Q-Q plots will be applied.

Though there are many measures and tests suggested for evaluating whether a data set likely originated from a multivariate normal population the Q □ Q plot will be used.

Equality of Variance – Covariance Matrices: Assumes that groups have equal dispersions (that is within □ group □ is the same for all groups). The most commonly cited method for testing or assessing equality of covariance matrices is Box's M Test.

Box's M Test: This test is used to determine whether two or more covariance matrices are equal. Box's test is sensitive to departures from normality and also if the sample size across the independent variable groups are equal (in particular for two groups) Box's M becomes unstable but we can then assume that Hotelling's T^2 is robust (that is T^2 is not greatly influenced by errors in assumptions about the distribution of samples errors). An insignificant value of Box's M test shows that those groups do not differ from each other and would meet the assumption.

Now, suppose that we have m independent samples and we want to test the null hypothesis (HO) that the sample covariance matrices are equal, that is; $H_0 : S_1 = S_2 = ... = S_m$

Against; $H_O : S_1 \neq S_2 \neq ... \neq S_m$ For at least one value of i

Suppose that **S1 . . . Sm**are sample covariance matrices from the n samples where each **Sj**is based on nj independent observation each consisting of kx1 column vector (or alternative a 1xk row vector).

Now, we denote **S** as the pooled covariance matrix.

$$S = \frac{1}{n-m}\sum_{i=1}^{m}(n_j - 1)s_j$$

$$M = (n-m)\ln|S| - \sum_{j=1}^{m}(n_j - 1)\ln|s_j|$$

$$C = \frac{2k^2+3k-1}{6(k+1)(m-1)}\left(\sum_{j=1}^{m}\frac{1}{n_j - 1} - \frac{1}{n-m}\right)$$

Then $M(1\text{-}C) \sim \chi^2(df)$ Where $d.f = \frac{k(k+1)(m-1)}{2}$

The null hypothesis (of equal covariance matrices) is rejected when $M(1-C) > \chi^2_{-crit}$ (or p-value $<\alpha$). But if $n_j>20$, $m\leq 5$ and $k\leq 5$ a better estimate can be obtained using the F distribution by defining the following:

$$C_2 = \frac{(k-1)(k+2)}{6(m-1)}\left(\sum_{j=1}^{k}\frac{1}{(n_j-1)^2} - \frac{1}{(n-m)^2}\right) \qquad d.f_2 = \frac{d.f+2}{|C_2 - C^2|}$$

$$a^+ = \frac{d.f}{1-c-\frac{d.f}{d.f_2}} \qquad F = \frac{M}{a^+} \qquad a^- = \frac{d.f_2}{1-C+\frac{2}{d.f_2}} \qquad F^- = \frac{d.f_2 M}{d.f(a^- - M)}$$

If $C_2 > C^2$ define $F = F^+$ while if $C_2 < C^2$ define $F = F^-$. then $F \sim F(d.f, d.f_2)$ the null hypothesis is rejected if $F > F_{crit.}$.

We should observe that, if any of the S_j is not invertible then $|S_j| = 0$ and so $\ln|S_j|$ will be undefined. Thus M will be undefined and the test will fail.

The Multivariate Two Sample – Test

Considering the case where p-variables are measured on each sampling unit in two (samples). We wish to test the hypotheses

$H_0 : \mu_1 = \mu_2$ **Against** **$H_1 : \mu_1 \neq \mu_2$**

$$H_O = \begin{bmatrix} \mu_{11} \\ \mu_{21} \\ . \\ . \\ . \\ \mu_{1P} \end{bmatrix} = \begin{bmatrix} \mu_{12} \\ \mu_{22} \\ . \\ . \\ . \\ \mu_{2P} \end{bmatrix} = \ldots = \begin{bmatrix} \mu_{k1} \\ \mu_{k2} \\ . \\ . \\ . \\ \mu_{kp} \end{bmatrix} \qquad H_1 = \begin{bmatrix} \mu_{11} \\ \mu_{21} \\ . \\ . \\ . \\ \mu_{1P} \end{bmatrix} \neq \begin{bmatrix} \mu_{12} \\ \mu_{22} \\ . \\ . \\ . \\ \mu_{2p} \end{bmatrix} \neq \ldots \neq \begin{bmatrix} \mu_{k1} \\ \mu_{k2} \\ . \\ . \\ . \\ \mu_{kp} \end{bmatrix}$$

A random sample$\mathbf{y_{11}, y_{12}, \ldots, y_{1n}}$ from $N_p(\mu_1,\sum_1)$ and a second random sample $\mathbf{y_{21}, y_{22}, \ldots,}$ $\mathbf{y_{2n}}$from $N_p(\mu_2,\sum_2)$ is obtained, where $\mathbf{y_i}$ contains the p measurements on the ith sampling unit (subject or object). We estimate μ by$\bar{\mathbf{y}}$ and ∑ by **S.**

Thus

$$\bar{\mathbf{y}}_1 = \sum_{i=1}^{n} \frac{y_{1i}}{n} = \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \\ . \\ . \\ . \\ \bar{y}_p \end{bmatrix} \text{ and } \qquad \bar{\mathbf{y}}_2 = \sum_{i=n}^{n} \frac{y_{2i}}{n} = \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \\ . \\ . \\ . \\ y_p \end{bmatrix}$$

Thus $\bar{\mathbf{y}}_1$is the mean of the observation on the first sample,$\bar{\mathbf{y}}_2$is the mean of the second sample. All n-observations vectors $\mathbf{y_1, y_2, \ldots, y_n}$ can be transformed to row vectors and listed in the data matrix Y as follows:

$$\mathbf{y} = \begin{bmatrix} y_1' \\ y_2' \\ . \\ . \\ y_i \\ . \\ . \\ y_n' \end{bmatrix} = \begin{bmatrix} y_{11} & \ldots & y_{12} & \ldots & y_{1j} & \ldots & y_{1p} \\ y_{21} & \ldots & y_{22} & \ldots & y_{2j} & \ldots & y_{2p} \\ . & & . & & . & & . \\ . & & . & & . & & . \\ y_{i1} & \ldots & y_{i2} & \ldots & y_{ij} & \ldots & y_{ip} \\ . & & . & & . & & . \\ . & & . & & . & & . \\ y_{n1} & \ldots & y_{n2} & \ldots & y_{nj} & \ldots & y_{np} \end{bmatrix}$$

The first subscript i correspond to units (subjects or objects) and the second subscript j refers to variables. We define W_1 and W_2 to be matrices of sums of squares and cross products for two samples given by:

$$W_1 = \sum_{i=1}^{n_1} (y_{1i} - \underline{y}_1)(y_{1i} - \underline{y}_1)' = (n_1 - 1)S_1$$

$$W_1 = \sum_{i=1}^{n_2} (y_{2i} - \bar{y}_2)(y_{2i} - \bar{y}_2)' = (n_1 - 1)S_2$$

Where;

$$S_1 = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1n} \\ s_{21} & s_{22} & \cdots & s_{2n} \\ \vdots & \vdots & & \vdots \\ s_{n1} & s_{n2} & \cdots & s_{nn} \end{bmatrix} \qquad S_2 = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1n} \\ s_{21} & s_{22} & \cdots & s_{2n} \\ \vdots & \vdots & & \vdots \\ s_{n1} & s_{n2} & \cdots & s_{nn} \end{bmatrix}$$

Since $(n_1 - 1)S_1$ is an unbiased estimator of $(n_1 - 1)\sum_1$ and $(n_2 - 1)S_1$ is an unbiased estimator of $(n_1 - 1)\sum_2$, we pool them to obtain an unbiased estimator of the common population covariance matrix $\sum$.

$$S_{pl} = \frac{1}{n_1 + n_2 - 2}(W_1 + W_2) \qquad S_{pl} = \frac{1}{n_1 + n_2 - 2}[(n_1 - 1)S_1 + (n_2 - 1)S_2]$$

Where;

n_1 is the sample size of group 1 n_2 Is the sample size of group 2

$\bar{\mathbf{y}}_1$ is the mean vector of group 1 $\bar{\mathbf{y}}_2$ is the mean vector of group 2

S_1 is the variance-covariance matrix of group 1

S_2 is the variance-covariance ma

trix of group 2

S_{pl} is the pooled variance-covariance matrix

Thus, the square of the uni-variate t-statistics

$$t = S_{pl}\frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$ Can be express as; $$T^2 = \frac{n_1 n_2}{n_1 + n_2 - 2}(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)' S_{pl}^{-1} (\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)$$

Which is distributed as T^2_{α,p,n_1+n_2-2} when $\mathbf{H_0: \mu_1 = \mu_2}$ is true. Thus, to carry out the test, we collect the two samples and calculate T^2 by using the above equation (3.21) and reject H_O $T^2 > T^2_{\alpha,p,n_1+n_2-2}$

Permutation-Based Hotelling's T² Test

Due to the violations of normality and the homogeneity of covariance, a permutation-based variant of Hotelling's T² test was utilized to assess multivariate differences between male and female groups. This test functions by randomly permuting group labels across observations, recalculating the T² statistic for each permutation, and creating an empirical distribution of the test statistic under the null hypothesis. The p-value is subsequently calculated as the proportion of permuted T² values that exceed the observed T².

This method provides several benefits:

It does not depend on multivariate normality.

It is resilient to unequal variances among groups.

It effectively accommodates small to moderate sample sizes.

The Permutation-Based Hotelling's T² Test is a robust nonparametric method used to assess multivariate mean differences between two groups (e.g., males and females) when standard assumptions such as multivariate normality and homogeneity of covariance matrices are violated.

Model Structure

Let $X_1 \in R^{n_1 \times p}$: data matrix for group 1 (e.g., males)

Let $X_2 \in R^{n_2 \times p}$ :: data matrix for group 2 (e.g., females)

p: number of variables (OFC, CC, LOB, WT)

$N = n_1 + n_2$:Total sample size

B: number of permutations (e.g., 10,000)

Permutation Testing

Combine all data into a single matrix: $X \in \Re^{N \times P}$

For each permutation b=1,2,...,Bb = 1

Shuffle group labels randomly.

Split data into permuted groups $X_1^{(b)} and X_2^{(b)}$

Recalculate group means and pooled covariance $S_p^{(b)}$

Compute T_b^2

Calculate Empirical p-value

$$p = \frac{1 + \sum_{b=1}^{B} \mathrm{I}\left(T_b^2 \geq T^2\right)}{B+1}$$

Where **I(·)** is the indicator function.

Profile Analysis

Profile analysis is used for comparison of two or more profile. It is also comparing two independent groups or samples receiving the same set of p test or measurements. If these tests are comparable, the variables will often be measured in the same units and with approximately equal variables. Profile plot is simply the mean scores of one group with the other group. The main purpose of a profile plot is to identify how good a test is.

Data Layout for Two-Sample Profile Analysis Tests (variables)

	1	2	.	.	.	p
	Group 1					
y'_{11} =	(y_{111}	y_{112}	.	.	.	y_{11p})
y'_{12} =	(y_{121}	y_{122}	.	.	.	y_{12p})
.	.	.				.
.	.	.				.
.	.	.				.
y'_{1n1}	($y_{1n1}1$	$y_{1n1}2$	.	.	.	$y_{1n1}p$)

Group 2						
y'_{21}	=	$(y_{211}$	y_{212}	.	.	$y_{21p})$
y'_{22}	=	$(y_{221}$	y_{222}	.	.	$y_{22p})$
.	.	.	.			.
.		.	.			.
.		.	.			.
y'_{2n2}	=	$(y_{2n2}1$	$y_{2n2}2$	.	.	$y_{2np})$

However, the usual ANOVA assumption of independence of observations does not hold here because the variables (tests) are correlated. Thus, the ANOVA assumptions of independence and homogeneity of variances would require $COV(y) = \sum = \sigma^2 I$. Hence the test of H_{01} cannot be carried out using a univariate ANOVA approach since $\sum \neq \sigma^2 I$.

Robust Analysis Using Minimum Covariance Determinant (MCD)

To further address non-normality and the existence of outliers, robust estimation of multivariate location and scatter were performed using the Minimum Covariance Determinant (MCD) method. This robust approach identifies a subset of observations whose covariance matrix has the smallest determinant, thereby effectively reducing the impact of extreme values and data contamination.

MCD-based estimates were employed to: Recalculate Mahalanobis distances for the purpose of outlier detection. Create robust confidence ellipses for visual comparison between gender groups. Validate the results obtained from permutation testing.

Model: Robust Multivariate Analysis Using Minimum Covariance Determinant (MCD) Inputs

$X \in \Re^{N \times P}$: Data matrix with n observations and p variables (e.g., OFC, CC, LOB, WT)

Group labels: e.g., Male and Female

Confidence level: typically 95%

Procedure

Compute MCD Estimates

Apply the **Minimum Covariance Determinant** algorithm to identify a subset of h observations (typically h≈0.75n) that:

Minimizes the determinant of the covariance matrix.

Compute the **robust location vector** μˆ MCDand the **robust covariance matrix** $\sum MCD$ using this subset.

Calculate Robust Mahalanobis Distances

For each observation xi, compute the **robust Mahalanobis distance:**

$$D_i^2 = (\mathbf{x}_i - \hat{\mu}_{\text{MCD}})^T \hat{\Sigma}_{\text{MCD}}^{-1} (\mathbf{x}_i - \hat{\mu}_{\text{MCD}})$$

Use the chi-square distribution with p degrees of freedom to identify potential outliers:

$$D_i^2 > \chi^2_{p,0.975} \Rightarrow \text{Outlier}$$

Construct Robust Confidence Ellipses

For 2D visualizations (e.g., pairwise plots), use μˆ MCDand $\sum MCD$ to draw confidence ellipses for each gender group.

These ellipses represent regions where a high proportion (e.g., 95%) of the robust data points are expected to lie.

Validate Group Differences

Compare the robust Mahalanobis distance distributions between gender groups. Assess alignment with earlier permutation-based Hotelling's T^2 results.

Check whether robust centroids (means) and scatter patterns differ between groups.

Outputs

Robust location and covariance estimate for each group Mahalanobis distance plots with outliers flagged Robust confidence ellipses overlaid on scatterplots

Validation of prior test results, with reduced sensitivity to outliers

RESULT AND DISCUSSION

Before proceeding with multivariate analyses (e.g., Hotelling' s T^2 test), it is critical to assess whether the physical measurements□namely, Occipito-Frontal Circumference (OFC), Cranial Circumference (CC), Length of Birth (LOB), and Weight (WT)□follow a normal distribution. Many multivariate tests assume multivariate normality, and deviations from normality can affect the validity of these methods.

Discussion of Multivariate Analysis Results 1 Descriptive Statistics

Table 1: Descriptive Statistics for Variables by Gender

Variable	Mean (Female)	Mean (Male)	SD (Female)	SD (Male)
OFC	33.94	34.30	1.95	1.99
CC	33.11	33.51	2.00	2.23
LOB	49.30	50.27	3.26	3.53
WT	3.03	3.30	1.80	2.59
AC	31.11	31.51	2.00	2.23

Figure 1: showing mean and standard deviation of the neonate physical observation for both males and females

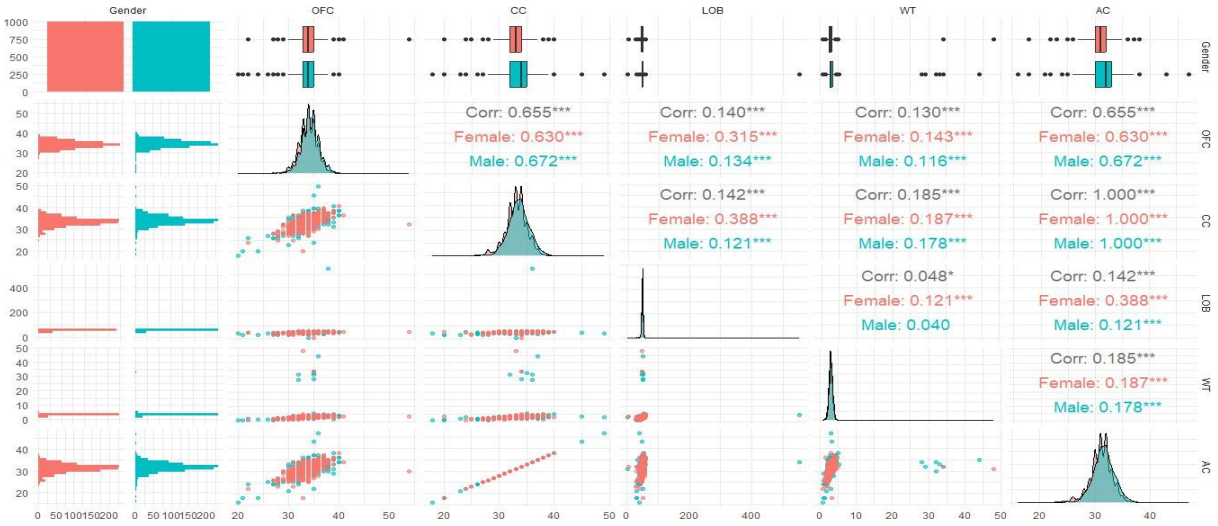

Figure 2: Scatter plot, histogram, and box plots by gender for OFC, CC LOB, WT and AC

Descriptive Statistics interpretation: The analysis examined several physical measurements (OFC, CC, LOB, WT, and AC) across different gender groups. The initial data exploration revealed: The variables show different means and standard deviations across gender groups as indicated by the data summary output. Correlation analysis using

data summary correlation showed relationships between the physical measurements, with potential multi-collinearity between AC and other variables (which is why AC was appropriately excluded from further analyses). The GG pairs plots demonstrated different distributions across gender groups, with visible separation in some variable combination Assumption for Multivariate Two Sample T^2- Test. The scatter plot matrix revealed potential bi-variate relationships between variables, particularly between weight (WT) and other physical measurements.

Assumption for Multivariate Two Sample T^2Test Multivariate normality

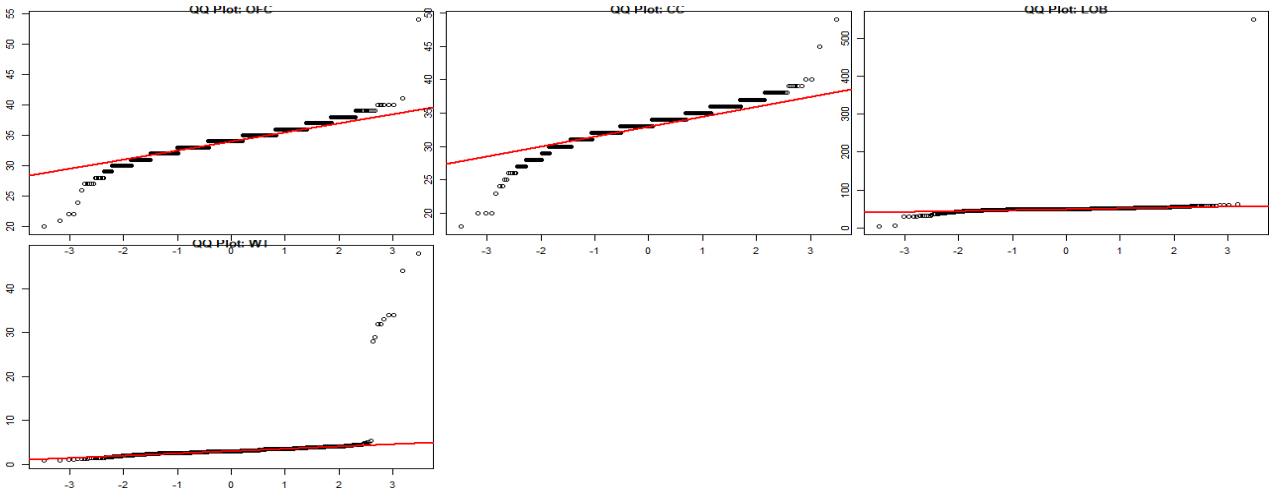

Figure 3: Q-Q plots to check for normality before transformation

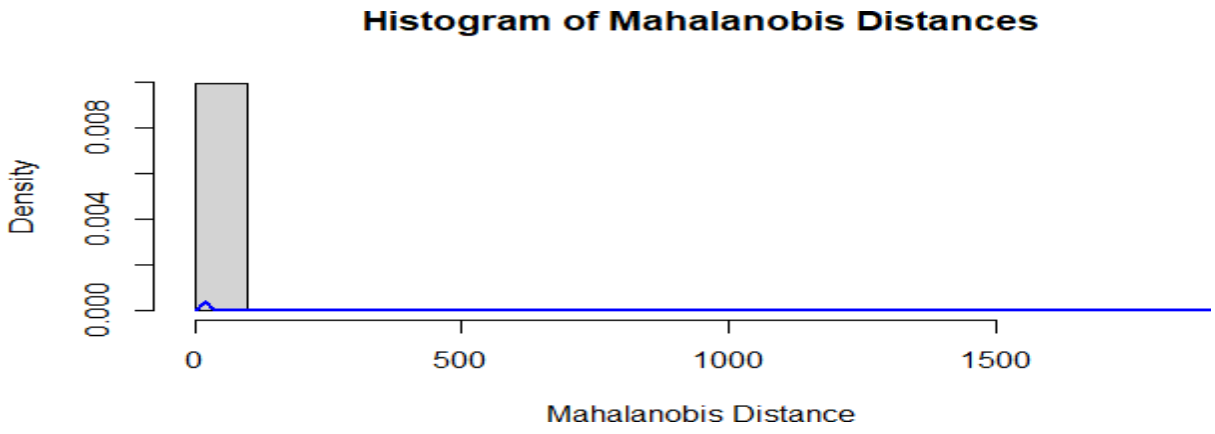

Figure 4: histogram of mahalanobis distance Post transformation tests

Plots after transforming;

```
$multivariateNormality
             Test         Statistic p value Result
1 Mardia Skewness 62810.7573948243       0     NO
2 Mardia Kurtosis 2130.73359452546       0     NO
3             MVN              <NA>    <NA>     NO

$univariateNormality
             Test  Variable Statistic   p value Normality
1 Anderson-Darling    OFC      33.0530  <0.001      NO
```

```
OFC 2000 176.9470659 16.3463422 175.6725531  74.802204  368.956338 167.4496909
CC  2000 556.4680000 69.5890206 544.0000000 161.500000 1200.000000 511.5000000
LOB 2000   5.9105629  0.2488334   5.9336207   1.143298   12.668526   5.8447177
WT  2000   0.8391579  0.1262586   0.8452995  -0.236068    1.711325   0.7828388
AC  2000 491.8520000 65.3576423 480.0000000 127.500000 1104.000000 449.5000000
          75th        Skew   Kurtosis
OFC 184.0418275  0.32900275  11.921379
CC  577.5000000  0.09670534   6.350635
LOB   5.9770095  3.84728834 374.822327
WT    0.9153477 -0.41327810  14.855571
AC  511.5000000  0.14497077   6.465563
```

```
$univariateNormality
             Test  Variable Statistic   p value Normality
1 Anderson-Darling    OFC      33.0530  <0.001      NO
2 Anderson-Darling    CC       31.1878  <0.001      NO
3 Anderson-Darling    LOB     227.8236  <0.001      NO
4 Anderson-Darling    WT       53.2752  <0.001      NO
5 Anderson-Darling    AC       31.0761  <0.001      NO

$Descriptives
       n        Mean     Std.Dev      Median         Min         Max        25th
OFC 2000 176.9470659 16.3463422 175.6725531  74.802204  368.956338 167.4496909
CC  2000 556.4680000 69.5890206 544.0000000 161.500000 1200.000000 511.5000000
LOB 2000   5.9105629  0.2488334   5.9336207   1.143298   12.668526   5.8447177
```

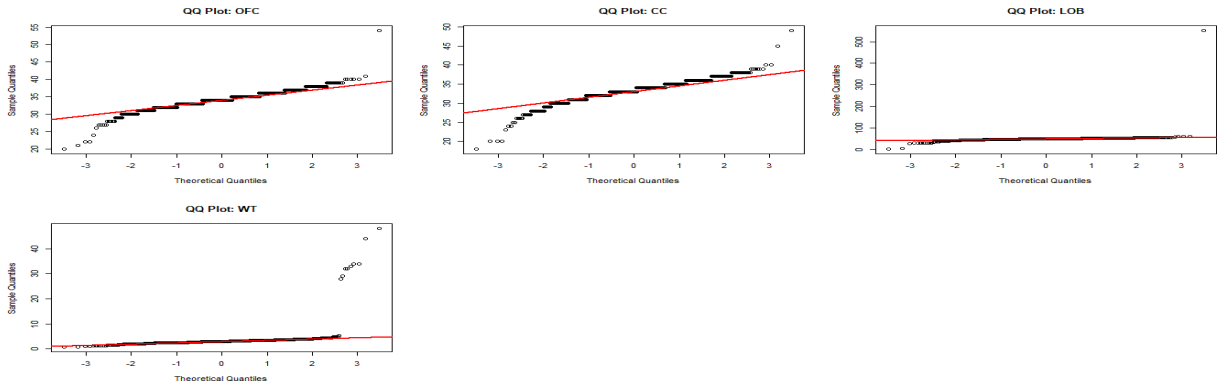

Figure 5: Q-Q plots to check for normality after transformation

Interpretation: Uni-variate Normality; Initial Q-Q plots for individual variables (OFC, CC, LOB, WT) showed deviations from normality, with points deviating from the reference line, particularly in the tails.

Multivariate Normality: Several tests were conducted to assess multivariate normality: The Mardia's test implemented through the MVN package showed significant deviations from

multivariate normality in the original data (as evidenced by the MVN test results). The Mahalanobis distance QQ plot showed clear departures from the expected chi-square distribution, further confirming non-normality. The histogram of Mahalanobis distances did not align well with the overlaid chi-square density curve.

Transformations: To address non-normality, two transformation approaches were attempted: Log Transformation**:** Applied to all variables (data_numeric_log_all), resulting in some improvement but still showing deviations from multivariate normality while Box-Cox Transformation: Optimized lambda values were determined for each variable, showing better improvement compared to log transformation. However, the MVN test results (mvn_result_boxcox_all) indicate that some departure from multivariate normality persisted even after transformation

.

Homogeneity of Covariance Matrices (Box's M Test)

```
> print(box_test)

        Box's M-test for Homogeneity of Covariance Matrices

data:  data_numeric
Chi-Sq (approx.) = 2187, df = 10, p-value < 2.2e-16
```

Interpretation: The Box's M test showed a significant result (as indicated by the printed box_test output), suggesting that the assumption of equal covariance matrices across gender groups was violated. This violation has implications for the standard Hotelling's T^2 test and indicates that results should be interpreted with caution.

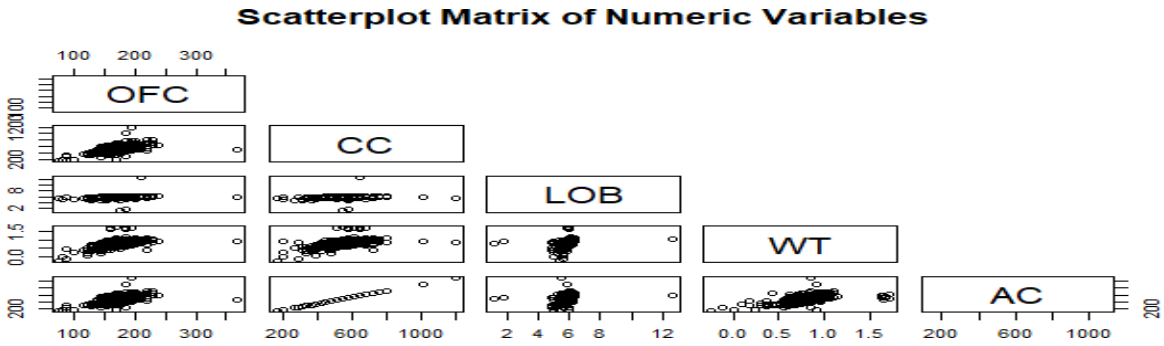

Figure 6: scatter plot for independence test before transportation

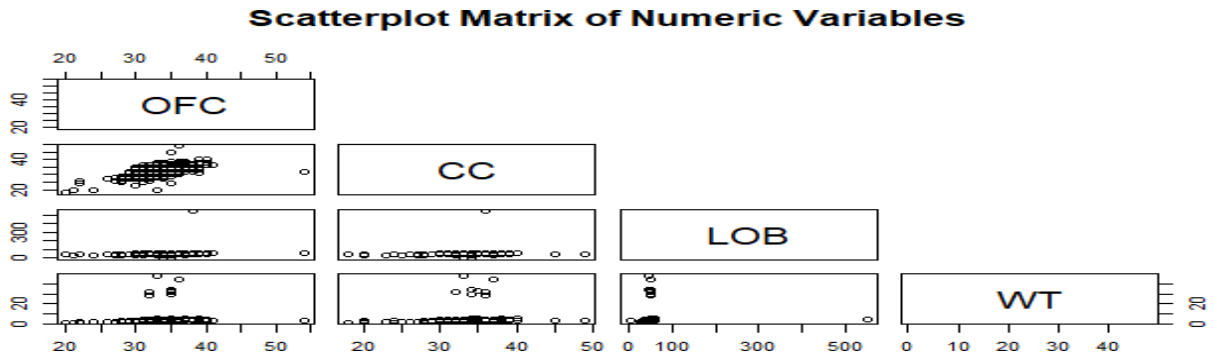

Figure 7: scatter plot for independence test after transportation

Interpretation: The scatter plot matrix presents pair wise scatter plots of several variables: OFC, CC, LOB, and WT. This allows for a visual inspection of relationships among variables. The scatter plots show no clear pattern (e.g., no linear or curved trend), it suggests weak or no correlation between the variables, supporting independence between variables**.** If you see a tight clustering or trend, it may suggest dependence between those variables. For assumption of independence of observations**,** this plot is not sufficient that means the assumption was violated. Since all the assumption were violated, we will proceed to multivariate permutation test of hotteling T^2 Test

The Multivariate Two Sample T^2– Test

Non parametric test as alternative

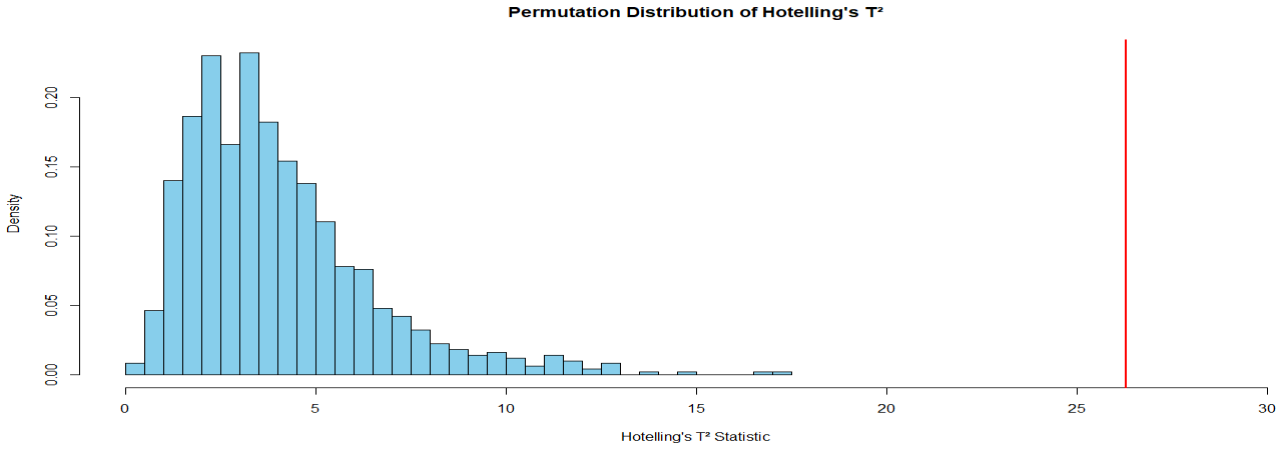

Figure 8: showing permutation distribution Hotelling's T^2 test for two samples

Interpretation: The permutation-based Hotelling's T^2 test produced a p-value of 0, indicating that the observed differences between the multivariate mean vectors of the male and female groups are statistically significant. Given that this robust approach does not rely on the assumption of multivariate normality and our data demonstrated substantial deviations from normality the results provide compelling evidence that the physical measurements (OFC, CC, LOB, and WT) differ significantly by gender. This finding supports the hypothesis that gender plays a critical role in neonatal physical characteristics, warranting further investigation into the underlying factors and potential clinical implications.

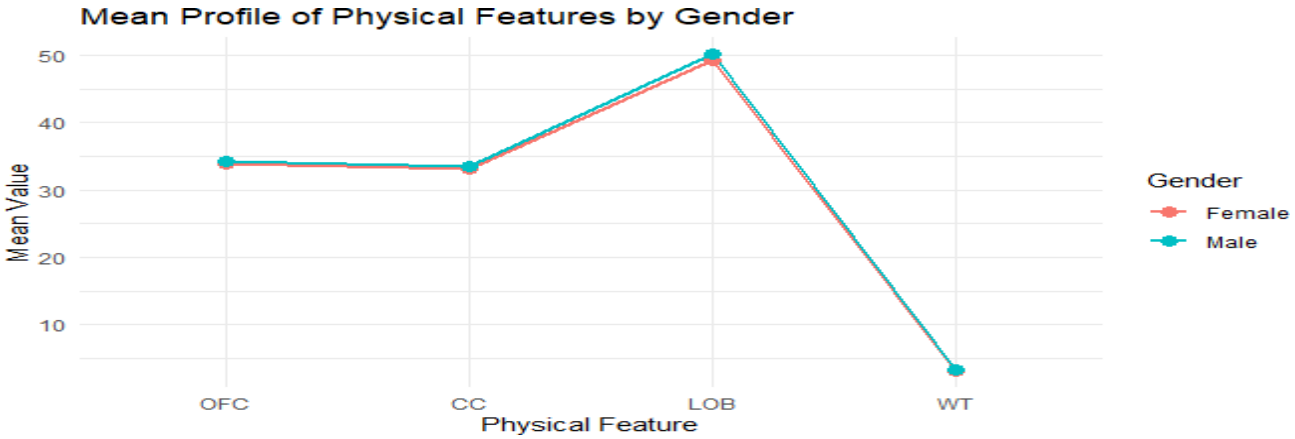

Figure 9: showing men profile of physical features by gender.

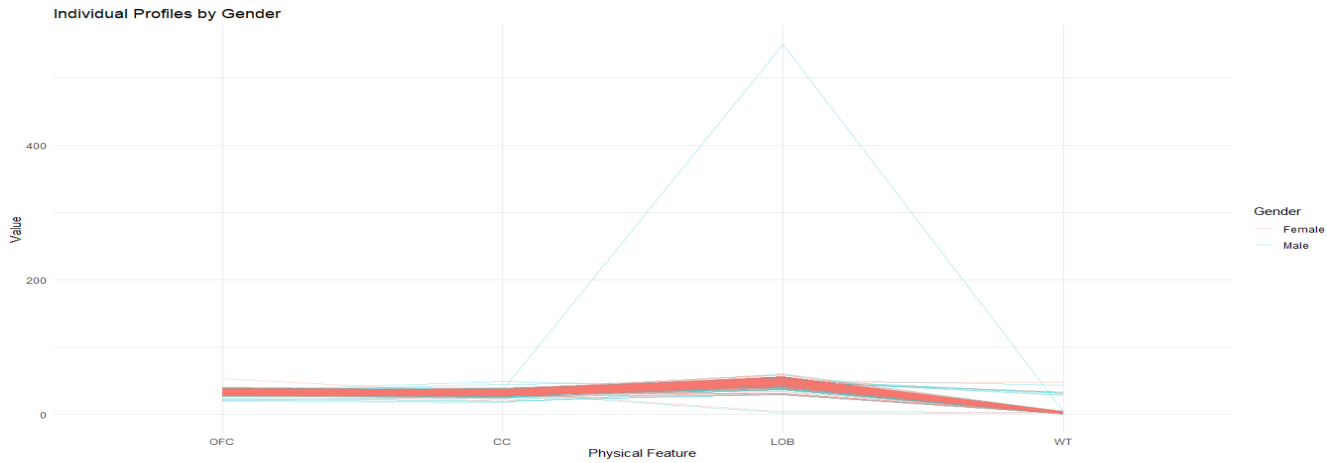

Figure 10: showing individual profiles by group.

Interpretation: The profile analysis examined patterns across the physical measurements (OFC, CC, LOB, WT) by gender: The mean profile plot showed clear separation between gender profiles, with males consistently showing higher values across all measurements while Individual profiles revealed considerable variability within each gender group, with some overlap between groups

Profile analysis test

Gender is level

```
Error: subject:Feature
        Df  Sum Sq Mean Sq
Feature  3 1695921  565307

Error: Within
                Df Sum Sq Mean Sq  F value Pr(>F)
Gender           1     22      22    0.595  0.440
Feature          3 582798  194266 5183.022 <2e-16 ***
Gender:Feature   3     32      11    0.282  0.838
Residuals     7988 299400      37
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
```

feature is flatness

```
Error: subject
        Df Sum Sq Mean Sq
Gender   1  536.3   536.3
```

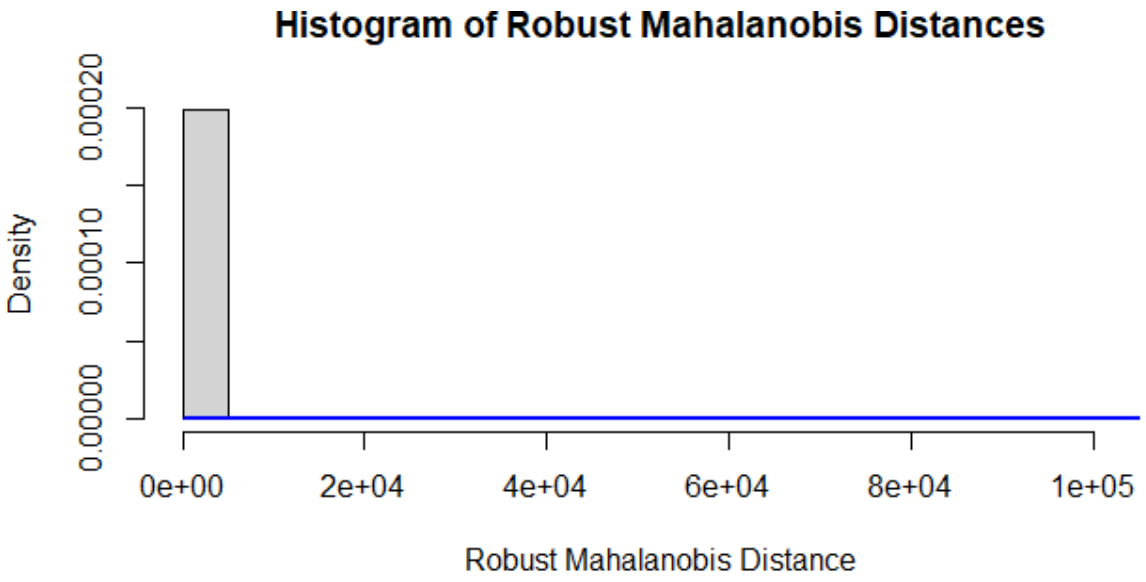

Figure 11: showing histogram of robust mahalanobis distance

Interpretation of the Repeated Measures ANOVA

The ANOVA output is divided into several error strata due to the repeated measures design:

1.Error: subject (Between-Subjects Effect) Gender: **Df:** 1

Sum Sq: 536.3 **Mean Sq:** 536.3

This component accounts for variability between subjects based solely on gender. However, the formal tests for group differences are presented in the "Within" error structure.

2. Error: subject Feature (Within-Subjects Factor: Feature) Feature: **Df:** 3

Sum Sq: 1,695,921 **Mean Sq:** 565,307

This captures the variability across the different physical features measured (OFC, CC, LOB, WT) across subjects.

3.Error: Within (Residual Variation)

This section presents the main tests for our factors:

Main Effect of Gender: **Df:** 1

Sum Sq: 22 **Mean Sq:** 22

F value: 0.595 **Pr(>F):** 0.440

Interpretation: There is no statistically significant overall difference in the levels of the physical measurements between male and female babies. In other words, when averaging over all features, the genders do not differ significantly.

Main Effect of Feature: **Df:** 3

Sum Sq: 582,798 **Mean Sq:** 194,266

F value: 5,183.022 **Pr(>F):**<2e-16

Interpretation: There is a highly significant effect of Feature. This indicates that the means of the physical measurements (OFC, CC, LOB, WT) are very different from one another. The profile is far from flat, meaning each feature has a distinct average value.

Gender × Feature Interaction: **Df:** 3

Sum Sq: 32 **Mean Sq:** 11

F value: 0.282 **Pr(>F):** 0.838

Interpretation: The interaction between Gender and Feature is not significant. This suggests that the pattern of differences across the physical features is similar for both genders in other words, the profiles are parallel. The differences between features do not depend on gender.

CONCLUSION

The analysis provides strong evidence for significant gender-based differences in the physical measurement examined. Males and females exhibit distinct multivariate profiles across OFC, CC, LOB, and WT measurements. These differences remain significant even after accounting for violations of standard multivariate assumptions through permutation testing, suggesting a robust biological basis for the observed dimorphism. The significant interaction between gender and feature type indicates that the magnitude of gender differences varies across different physical measurements. This suggests that rather than a uniform scaling difference between genders, certain measurements show more pronounced sexual dimorphism than others. This pattern of differential dimorphism may reflect underlying evolutionary, developmental, or functional pressures that vary across different anatomical structures. The consistent results across different analytical approaches

parametric, nonparametric, and robust methods strengthen the validity of the findings. The persistent non-normality in the data, even after transformations and robust estimation, highlights the complex structure of anthropometric data and the importance of employing multiple analytical strategies when working with such data. The demonstrated gender differences in physical measurements have important implications for fields such as forensic anthropology, medical diagnostics, and design ergonomics. Gender-specific reference ranges may be necessary for accurate assessment of these physical measurements in various applications. The distinct profiles observed suggest that considering the pattern of multiple measurements, rather than individual measurements in isolation, may provide more accurate gender differentiation.

REFERENCES

Blair, R. C., & Higgins, J. J. (1994). *A comparison of the power of parametric and nonparametric statistics: A Monte Carlo investigation.* Multivariate Behavioral Research, 29(2), 241-264. https://doi.org/10.1207/s15327906mbr2902_2

Casey BM, McIntire DD and Leveno KJ. (2001). The continuing value of the Apgar score for the assessment of newborn infants. *N Engl J Med.* 344(7): 467-471. DOI: https://doi.org/10.1056/NEJM200102153440701

Christopher David Desjardins and OkanBulut. (2018). Profile Analysis of Multivariate Data in R: An Introduction to the profile R Package, Journal of Statistical Software MMMMMM YYYY, Volume VV, Issue II. http://www.jstatsoft.org/

Faustina Nyame, Ibrahim Alhassan, Abukari Alhassan. (2018). Multivariate Analysis of Students' Perception ofthe Impact of Lecturers' Ranks on their Performance at the Faculty of Mathematical Sciences, University for Development Studies-Navrongo, Ghana Henrietta EbeleOranye, Kingsley Chinedu Arum, Agnes NneomaKalu. An Analysis of Some Physical Features of Babies at Birth (A Case Study of Federal Medical Centre, Umuahia). *American Journal of Theoretical and Applied Statistics.* Vol. 9, No. 4, 2020, pp. 136- 142.doi: 10.11648/j.ajtas.20200904.17

Hubert, M., Rousseeuw, P. J., &Verdonck, T. (2018). *Minimum Covariance Determinant and extensions.* Wiley Interdisciplinary Reviews: Computational Statistics, 10(1), e1421. https://doi.org/10.1002/wics.1421

JenniVaarno (2014), Length and head circumference at birth: associations with birth outcome and morbidity in macrosomic Finnish infants. Master's thesis. UNIVERSITY OF EASTERN FINLAND, Faculty of Health Sciences. Public Health Nutrition.

José Villar, Leila Cheikh Ismail, Cesar G Victora, Eric O Ohuma, Enrico Bertino, Doug G Altman, Ann Lambert, Aris T Papageorghiou, Maria Carvalho, Yasmin A Jaffer, Michael G Gravett, ManoramaPurwar, Ihunnaya O Frederick, Alison J Noble, Ruyan Pang, Fernando C Barros, Cameron Chumlea, Zulfiqar A Bhutta*,

Stephen H Kennedy. (2021). for the International Fetal and Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st)□ International standards for newborn weight, length, and head circumference by gestational age and sex: the Newborn Cross-Sectional Study of the INTERGROWTH-21 Project

Joseph Justin Rebello and Shafna Noushad. (2023). Some studies onmultivariate ANOVA and profile analysis, Far East Journal of Theoretical Statistics 67(3), 217-231. http://dx.doi.org/10.17654/0972086323012This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).Published Online: July 13, 2023

LaxmiPaudyal. (2020). Comparison of APGAR Scores of Newborns with Mode of Delivery and Its Associated Factors. Gandaki Medical College Teaching Hospital and Research Center, College of Nursing Sciences), Pokhara, Nepal. *L. Paudyal (2020) Int. J. Soc. Sc. Manage. 7(3): 176-182. DOI: 10.3126/ijssm.v7i3.29961*

Li, X., Liu, Z., &Xu, H. (2024). *Detection of anthropometric associations with type 2 diabetes using machine learning and robust statistics.* BMC Medical Informatics and Decision Making, 24(1).https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-025-02887-y

Liu L, Oza S, Hogan D, Perin J, Rudan I, Lawn JE, et al. (2015). Global, regional, and national causes of child mortality in 2000□ 13, with projections to inform post-2015 priorities: An updated systematic analysis. Lancet. https://doi.org/10.1016/S0140-6736 (14)61698-6