An Analysis of Some Physical Features of Babies at Birth (A Case Study of Federal Medical Centre, Umuahia)

Henrietta Ebele Oranye, Kingsley Chinedu Arum, Agnes Nneoma Kalu

Department of Statistics, Faculty of Physical Sciences, University of Nigeria, Nsukka, Nigeria

Email address:
henrietta.oranye@unn.edu.ng (H. E. Oranye), kingsley.arum@unn.edu.ng (K. C. Arum), perkysuccess@gmail.com (A. N. Kalu)

To cite this article:
Henrietta Ebele Oranye, Kingsley Chinedu Arum, Agnes Nneoma Kalu. 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

Received: July 1, 2019; Accepted: August 16, 2019; Published: July 6, 2020

Abstract: New born baby is a gift from God and what a newborn baby looks like is not a baby model, rather a newborn baby looks varies from baby to baby in terms of weight, height and head circumferences. In this research, a sample of 200 male and female babies was used for the analysis. The aim is to identify if there is a significant difference between the means of the variables considered. In this research, three variables were considered for both male and female babies at birth. The result showed that the mean birth weight is 3.55kg and 3.39kg for male and female babies respectively. The mean height and head circumference of female babies recorded higher than their male counterpart. The Hoteling’s $T^2$-test showed that there is a significant difference between the mean vectors of the variables considered; hence a discriminant analysis was conducted. The discriminant function obtained fairly classifies the group at 42% error rate. From the results gotten, there is significant difference between the height, weight and head circumference of male and female babies and conclude that male babies are heavier in terms of weight while female babies have bigger head circumference than the male babies.

Keywords: Babies Height, Babies Weight, Babies Head Circumference, Discriminant Function, Hoteling’s $T^2$-test

1. Introduction

A newborn baby looks varies from baby to baby in terms of weight, length and head circumferences. New born babies are checked during birth in case of signs of problems or complications. A thorough physical assessment is performed by doctors to access the health of the baby. Birth weight, head circumference and height are part of these physical assessment and they are important indicator to the health of the baby.

Babies’ size at birth is important determinant of birth outcome and perinatal morbidity and mortality. Being too big at birth puts the newborn at high risk of various health problems at or after birth. Measuring babies’ size at birth is important for monitoring of individuals growth and development and for public health in efforts to improve neonatal and maternal morbidity and mortality [7].

According to Wikipedia Birth weight is the body weight of a baby at its birth, measured with a baby scale where newborn is lying down in the weighting pan. Weights of the baby are measured to the nearest 0.01 kg. Newborn height is measured from top of the head to heels. The heights of the baby are measured to the nearest millimeter (mm). Head circumference is taken as the largest circumference of the head, above eyes and ears. Measurement tape is positioned just above ears and eyebrows and around the biggest part of the back of the head. Measurement is read to the nearest 0.1cm [4].

In this study, multivariate analysis was used to analyze data collected on measurement of babies at birth. The variables considered in this study include the height, weight and head circumference of babies.

In this study, discriminant function analysis was used to obtain a linear function of the variables as either belonging to male group or female group. We used the Hoteling’s $T^2$-test to examine if there is significant difference between the mean vectors of the variables (weight, length and head circumference) measures for male and female babies at birth. The misclassification error rate was also obtained using the Mahalanobis squared distance.
2. Literature Review

Several literatures have documented analysis on different features of babies with different approaches. Different features of babies contribute a lot to a newborn baby looks.

Head circumference as predictor of microcephaly among term infants born in a teaching hospital in Malaysia from 2011 to 2015. Using a cross-sectional study from the electronic birth census data, the independent variables were mothers’ age and height, parity, birth weight and birth length. All term newborns, both alive and stillbirth, with 37–41 completed gestational weeks, and a birth weight of at least 500 g was extracted from the census. A total of 26, 503 newborns fulfilled the inclusion criteria (13, 655 males, 12, 840 females). The mean head circumferences for male and female newborns were 32.93 cm and 32.56 cm respectively. The average head circumference for Malaysian newborns was found to be smaller than the World Health Organization Standard Growth Chart for Term Infant. There is an increasing trend of microcephaly across the years and the low birth weight was noted as the main predictor of microcephaly [8].

Figure 1: Scatter plot of features of the babies.

Chi-square test and t-test was used to analyze the prevalence of gender difference in body mass index with sample data of 426 high school students who completed sectional self-report questionnaires. The variables considered for their study include; body weight, eating habit, body height. The result of the test carried out shows that females are underweight compared to males [2].

“An Anthropometric studies of newborns at birth in Western Rajasthan”. In this study, they used some anthropometric measurement which include; length, weight, head circumference and chest circumference of normally healthy neonates at birth. Using independent sample T-test, it was found that there is significant difference between different parameter of male and female babies. And there is significant difference between mean length, head circumference of male and female babies [6].

A study was carried out to test for the equality of two covariance matrices from two independent multivariate normal populations with high-dimensional data. The test statistic was based on unbiased and consistent estimator of the ratio between the sums of squares of covariance matrix elements. Under the null hypothesis, the proposed test statistic is asymptotically standard normal distributed when the number of variables and the sample sizes go together to infinity. Simulation study was conducted to investigate the performance of the proposed test statistic. The results showed that the proposed test is superior to the other three tests appeared in the literature for various patterns of common covariance matrix. Finally, two real data sets are analyzed to illustrate the application of our theoretical results [5].

3. Methods

3.1. Data Presentation

The data for this study are measurements taken on babies
delivered at Federal Medical Centre (FMC), Umuahia for a period of 2yrs. The measurements obtained are: Babies weight (BW): weight of child at birth measured in kilogram. Babies Height (BH): height of child at birth measured in centimeter. Head circumference (HC): Head circumference of child at birth measured in centimeter.

\[ x_1 = \text{Head circumference of male babies}, \quad x_2 = \text{Birth height of male babies}, \quad x_3 = \text{Birth weight of male babies}, \quad y_1 = \text{Head circumference of female babies}, \quad y_2 = \text{Birth height of female babies}, \quad y_3 = \text{Birth weight of female babies}, \quad n_1 = \text{Number of observations for male babies}, \quad n_2 = \text{Number of observations for female babies}, \quad \overline{x} = \text{Mean vector for male babies}, \quad \overline{y} = \text{Mean vector for female babies}, \quad S_1 = \text{covariance matrix for male babies}, \quad S_2 = \text{covariance matrix for female babies}, \quad S = \text{Pooled covariance matrix}, \quad P = \text{Number of variables considered.}

Many univariate tests and confidence intervals are based on the univariate normal distribution. Similarly the majority of multivariate procedures have the multivariate normal distribution as their underpinning [9].

Some features of the multivariate normal distribution

i. The distribution can be completely described using only means, variance and covariance

ii. The linear functions of multivariate normal variables are also normal.

iii. Even when the data are not multivariate normal, the multivariate normal may serve as a useful approximation, especially in inferences involving sample mean vectors, which are approximately multivariate normal by the central limit theorem. The central limit theorem states that for large samples, the sampling distribution of means in the univariate case will approach normality [3].

### Table 1. Sample Mean Vector and Covariance of Male and Female.

| Gender | Variable           | Observation | Mean     | Covariance |
|--------|--------------------|-------------|----------|------------|
| Male   | Head Circumference | 200         | 34.355   | 7.7539     |
|        | Birth Height       | 200         | 46.08    | 15.53078   |
|        | Birth Weight       | 200         | 3.5594   | 0.253715   |
| Female | Head Circumference | 200         | 33.49    | 7.6199     |
|        | Birth Height       | 200         | 44.93    | 14.3851    |
|        | Birth Weight       | 200         | 3.39775  | 0.256361   |

3.2. Estimation of Parameters

The means, variances, and covariance’s are the natural parameters of multivariate normal distribution.

**Mean Vectors Estimation**

The sample mean vector for the male and female babies are given by \( \overline{x} = \frac{1}{n_1} \sum_{i=1}^{n_1} x_i \) and \( \overline{y} = \frac{1}{n_2} \sum_{i=1}^{n_2} y_i \) respectively.

\[
\begin{bmatrix}
    x_1 - \overline{x} \\
    x_2 - \overline{x} \\
    x_3 - \overline{x}
\end{bmatrix} =
\begin{bmatrix}
    37 - 34.355 \\
    50 - 46.08 \\
    4.60 - 3.5594
\end{bmatrix}
\]

\[
\begin{bmatrix}
    y_1 - \overline{y} \\
    y_2 - \overline{y} \\
    y_3 - \overline{y}
\end{bmatrix} =
\begin{bmatrix}
    36 - 33.49 \\
    47 - 44.93 \\
    3.30 - 3.39775
\end{bmatrix}
\]

Covariance matrix:

\[
S_1 = \frac{1}{n_1-1} \sum_{i=1}^{n_1} (x_i - \overline{x})(x_i - \overline{x})^T = \begin{bmatrix}
    7.7539 & 3.78485 & 0.555686 \\
    3.78485 & 15.53078 & 1.057014 \\
    0.555686 & 1.057014 & 0.253715
\end{bmatrix}
\]

\[
S_2 = \frac{1}{n_2-1} \sum_{i=1}^{n_2} (y_i - \overline{y})(y_i - \overline{y})^T = \begin{bmatrix}
    7.6199 & 1.9493 & 0.348003 \\
    1.9493 & 14.3851 & 0.708743 \\
    0.348003 & 0.708743 & 0.256361
\end{bmatrix}
\]

Pooled covariance matrix:

\[
S = \frac{1}{(n_1+n_2-2)} (S_1 + S_2) = \begin{bmatrix}
    7.706165 & 2.874261 & 0.453428 \\
    2.874261 & 14.99543 & 0.885091 \\
    0.453428 & 0.885091 & 0.255677
\end{bmatrix}
\]
The inverse of the pooled sample dispersion matrix

\[
S^{-1} = \begin{pmatrix}
0.14794235 & -0.01617625 & -0.2063688 \\
-0.01617625 & 0.08558076 & -0.2675720 \\
-0.20636879 & -0.26757197 & 5.2034361
\end{pmatrix}
\]

3.3. Multivariate Test for Normality

In testing for the multivariate normality, the Mardia’s test which is based on multivariate extension of Skewness and Kurtosis measures is used to test for both the male and female sample.

Hypothesis to be tested:

- \( H_0 \): The data are normally distributed
- \( H_1 \): The data are not normally distributed

Level of significance: \( \alpha = 0.05 \)

Test Statistic

For Skewness measure:

\[
\frac{1}{n} \sum_{i=1}^{n} (b_{1,i} - \bar{x})^3 \times S^{-1} (\bar{x} - \bar{y})^3
\]

And for Kurtosis measure:

\[
\frac{1}{n} \sum_{i=1}^{n} (b_{2,i} - \bar{x})^4 \times S^{-1} (\bar{x} - \bar{y})^4
\]

Decision: For Skewness measure: reject \( H_0 \) if \( \chi^2_{cal} > \chi^2_{\alpha/2} \), accept if otherwise, for Kurtosis measure: reject \( H_0 \) if \( |Z| > Z_{\alpha/2} \), accept if otherwise.

Table 2. Test for Normality.

| Test Statistic | P value | Decision |
|----------------|---------|----------|
| Skewness       | 0.7049  | NOT SIGNIFICANT |
| Kurtosis       | 0.7368  | NOT SIGNIFICANT |

For Female

| Test Statistic | P value | Decision |
|----------------|---------|----------|
| Skewness       | 0.5203  | NOT SIGNIFICANT |
| Kurtosis       | 0.9327  | NOT SIGNIFICANT |

For Male

Conclusion: since the null hypothesis for male and female babies is not rejected, the multivariate test for normality is satisfied, hence both male and female babies’ data are normally distributed.

3.4. Multivariate Test for Equality of Mean Vector

The test of equality of mean vector is used to check if the two samples differ significantly. To obtain this, the Hotelling’s \( T^2 \) test statistic is used, which is a multivariate extension of student’s t distribution. Hotelling’s \( T^2 \) statistic occurs when testing hypotheses regarding means in one- and two-sample multivariate normal populations.

Hypothesis to be tested:

\( H_0: \mu_1 = \mu_2 \)

Level of significance: \( \alpha = 0.05 \)

Test statistic:

\[
F = \frac{(n_1+n_2-P-1)T^2}{P(n_1+n_2-P-1)}
\]

Level of significance: \( \alpha = 0.05 \)

Test statistic:

\[
F = \frac{(n_1+n_2-P-1)T^2}{P(n_1+n_2-P-1)}
\]

Decision: Reject \( H_0 \) if \( F_{cal} > F_{tab} \), accept if otherwise.
Table 3. Test for Significance.

| Test  | Test statistic | Prob > F (3, 396) | Decision |
|-------|----------------|-------------------|----------|
| Fcal  | 5.4714885      | 0.117186          | SIGNIFICANT |

Conclusion: there is a significant difference between the mean vector of male and female babies at birth. The mean vectors of male and female babies are not the same.

3.5. Tests for Equality of Two Covariance Matrix

If the two mean vectors are heterogeneous, classification is carried out and also the covariance matrix will be tested for equality. If the two covariance matrices are equal, Fisher’s linear discriminant analysis is used for classification; otherwise the Fisher’s quadratic discriminant analysis is used.

Hypothesis to be tested:

\( H_0: \) The two covariance matrices are equal

\( H_1: \) The two covariance matrices are not equal

Level of significance \( \alpha = 0.05 \)

Test statistic:

\[
T = -\rho \ln \lambda \sim \chi^2_{a \times 2, \nu}
\]

\[
\rho = 1 - \left( \frac{1}{n_1} + \frac{1}{n_2} - \frac{1}{n_1 + n_2} \right) \left( \frac{2p^2 + 3p - 1}{6p(p+1)} \right)
\]

\[
\ln \lambda = \frac{n_1}{2} \ln(S_1) + \frac{n_2}{2} \ln(S_2) - \frac{(n_1 + n_2)}{2} \ln(S)
\]

Decision rule: Reject \( H_0 \) if \( T > \chi^2_{a \times 2, \nu} \), accept if otherwise.

Where \( \nu = \frac{1}{2}p(p+1) \)

Computation:

\[
\ln \lambda = \frac{200}{2} \ln(17.98774) + \frac{200}{2} \ln(22.51829) - \frac{200 + 200}{2} \ln(20.62022)/
\]

\[
= 288.47 + 311.433 - 612.54
\]

\[
\rho = 1 - \left( \frac{1}{200} + \frac{1}{200} - \frac{1}{400} \right) \left( \frac{2(3)^2 + 3(3) - 1}{6(3)(3+1)} \right)
\]

\[
\frac{1}{n_1} + \frac{1}{n_2} - \frac{1}{n_1 + n_2} \left( \frac{2p^2 + 3p - 1}{6p(p+1)} \right)
\]

Table 4. Test for Equality of Covariance Matrices across two Samples.

| Chi Square | Box F (6, 1147681.8) | Chi Square |
|------------|----------------------|------------|
| Value      | 7.586625             | 1.25       |
| Prob > F   | 0.2750               | 0.2750     |

Levene's Test of Equality of Error Variances

| F      | df1 | df2 | Sig. |
|--------|-----|-----|------|
| HeadCircumference | .022 | 1 | 398 | .883 |
| Height  | .582 | 1 | 398 | .446 |
| weight  | .028 | 1 | 398 | .867 |

In conclusion, there is no significant difference between the covariance matrix of male and female babies at birth. The covariance matrices of male and female babies are the same. Hence, the linear discriminant analysis would be used for classification.

3.6. Discriminant Function for Classification

The two mean vectors we tested are heterogeneous that is why classification will be carried out. And also, the covariance matrices are equal which will result in using firsher’s linear discriminant analysis.

We wish to predict group membership from the data. This analysis will create a set of discriminant functions that will enable us to classify into which group (male or female) as case falls, based on the predictor variables (head circumference, birth height and birth weight).

The linear discriminant function is given as:

\[
L(x) = (\overline{x} - \overline{y})^T S^{-1} x - \frac{1}{2} (\overline{x} - \overline{y})^T S^{-1} (\overline{x} - \overline{y})
\]

\[
(\overline{x} - \overline{y})^T = (0.82 \ 1.135 \ 0.15665) , S^{-1} = \begin{pmatrix}
0.14794 & -0.01618 & -0.20637 \\
-0.01618 & -0.26757 & 0.20637 \\
-0.20637 & -0.26757 & 5.20344
\end{pmatrix}
\]

\[
(\overline{x} - \overline{y})^T S^{-1} (\overline{x} - \overline{y}) = 10.9850
\]

\[
\frac{1}{2} (\overline{x} - \overline{y})^T S^{-1} (\overline{x} - \overline{y}) = 5.4925
\]

\[
LDA = 0.07063X_1 + 0.04195X_2 + 0.34220X_3 - 5.49255
\]
The critical value of the discriminate function is given as

\[ Y_{\text{cal}} = \frac{X + Y}{2} \]

Where \( X \) and \( Y \) are the mean discriminant value for group 1 and group 2 respectively.

\[
X = 0.07063(34.35) + 0.04195(46.08) + 0.34220(3.5594) - 5.49255 = 11.06460
\]

\[
Y = 0.07063(34.49) + 0.04195(44.93) + 0.34220(3.39775) - 5.49255 = 10.97612
\]

The critical value of the discriminant function is

\[
S_{TUV} = \frac{1}{2}(11.06460 + 10.97612) = 11.02036
\]

**Classification Rule**

The classification rule is given below:

Allocate male child to group 1 if values of \( X > 11.02036 \) or group 2 if values of \( X < 11.02036 \).

The classification of the groups is summarized in the table below:

| True Gender | Group 1 | Group 2 | Total |
|-------------|--------|--------|-------|
| Male        | 116    | 84     | 200   |
| Percentage  | 58.00  | 42.00  | 100   |
| Female      | 89     | 111    | 200   |
| Percentage  | 44.50  | 55.50  | 100   |
| Total       | 205    | 195    | 400   |
| Percentage  | 51.25  | 48.75  |       |
| Prior       | 0.5000 | 0.5000 |       |

**Table 5. Linear Discriminant Function Table.**

Discriminant Analysis | Stata Annotated Output [1].

The classifications with their frequencies of the true gender are summarized in the table above. The totals of male and female data are 200 dataset each that fall into group 1 and group 2 respectively. 116 out of 200 dataset are in group 1; meaning that 116 are classified appropriately by \( 0.07063X_1 + 0.04195X_2 + 0.34220X_3 - 5.49255 \) and the remaining 84 are classified wrongly in group 2.

The misclassification proportion for group 1 is \( \frac{84}{200} = 0.42 \) and group 2 gives \( \frac{89}{200} = 0.445 \) the total misclassification is given by \( 0.42 + 0.445 = 0.865 \).

Based on the analysis from the above table, there is need for misclassification error.

Error of misclassification

Let the total chance of misclassification (error rate = \( \epsilon \)) consisting of two components:

\( e_1 = \text{Probability that an individual belonging to group 2 is misclassified as belonging to groups 1} \)
\( e_2 = \text{Probability that an individual belonging to group 1 is misclassified as belonging to groups 2} \)

Hence the total chance of misclassification \( (e_1 + e_2) = \Phi(\frac{-1}{2}\Delta_p) \)

To detect the error rate involved, we use the mahalanobis’s distance.

Estimating \( \Delta_p \)

\[
\Delta_p = (\bar{x} - \bar{y})^T S^{-1} (\bar{x} - \bar{y})
\]

Where, \( \Delta_p^2 \) is the mahalanobis squared distance.

Computation:

\[
= [0.82 \quad 1.135 \quad 0.15665] \begin{pmatrix} 0.14794 & -0.01618 & -0.20637 \\ -0.01618 & -0.01618 & -0.26757 \\ -0.20637 & -0.26757 & 5.20344 \end{pmatrix} \begin{pmatrix} 0.82 \\ 1.135 \\ 0.15665 \end{pmatrix}
\]

\[
\Delta_p^2 = 0.1591367
\]

\[
\Delta_p = \sqrt{0.1591367} = 0.3989194
\]

\[
e_1 = \text{Pr}[Z < \frac{-1}{2}\Delta_p] = \Phi(-0.1994) = 0.4207
\]

Therefore, error rate is 42.072%.

The most discriminating variable is given by \( \frac{\Delta_p}{S} \), where \( S \) is the standard deviation of sample dispersion matrix

For computation for group 1

\[
X_1 = 0.3989194/2.7539 = 0.0514475, X_2 = 0.3989194/15.53078 = 0.0256857, X_3 = 0.3989194/0.253765 = 1.57200
\]

For group 2

\[
Y_1 = 0.3989194/7.6199 = 0.05235, Y_2 = 0.3989194/14.3851 = 0.027731, Y_3 = 0.3989194/0.256361 = 1.55608
\]
From the computations above, it was found that $X_3$ which is male birth weight and $Y_3$ which is female birth weight are the most discriminating variables in the groups.

4. Discussion

This study was conducted on data collected from FMC, Umuahia. The study was performed on 400 babies (200 males and 200 females). The study was performed with measurements on weight, height and head circumference of babies at birth.

The result showed that the mean weight for male babies is 3.559 kg and for female babies is 3.398 kg. The mean height for male babies is 46.08 cm and for female babies is 44.93 cm. The mean head circumference for male babies is 34.35 cm and for female babies is 33.49 cm. Using Hotelling’s t-test, it was observed that there is significant difference between the mean weight, height and head circumference of male and female babies at $\alpha = 0.05$ level of significance. These results agree with some literature mentioned above, which have a significant difference between the mean vector of male and female babies at birth.

5. Conclusion

From the statistical analysis carried out in these work the test for normality was satisfied. The data obtained for this study are normally distributed. This satisfies the normality assumption for multivariate analysis.

The Hotelling’s T-test was used to test for the significant difference between the mean vectors of the parameters considered for the two samples measured. The result showed that there is a significant difference between the mean vectors of the variables considered hence a discriminant analysis was conducted to distinguish the two groups. The test for equality of covariance matrix shows that the covariance matrix are the same, hence a linear discriminant analysis was used for the classification. The discriminant function obtained fairly separates the group at 42% error rate.

Finally, this work reveals that there is significant difference between the height, weight and head circumference of male and female newborns. The male babies are heavier, larger, higher and have bigger head circumference than the female babies.

References

[1] Discriminant Analysis Stata Annotated Output. UCLA Institute for Digital Research & Education. Link: https://stats.idre.ucla.edu/stata/output/discriminant-analysis/.

[2] Mohamed B. S. and Rihab A. S. (2014) “Prevalence of gender difference in body mass index”. Procedia-social and behavioral sciences. 159 pp 184-191.

[3] Rencher, A. C (2002) “methods of Multivariate Analysis, 2nd edition”. John Wiley and sons, Inc. New York.

[4] Salo J, Mäki P and Dunkel L. (2011). In Health inspections in child health clinics and school health care, method hand book Finnish National Institute for Health and Welfare, Tampere.

[5] Saowapha C. and Samruam C. (2013): A Test for Testing the Equality of Two Covariance Matrices for High-dimensional Data. Journal of Applied Sciences Volume 13 (2): 270-277, 2013.

[6] Sushama K K, Samta G (2014): An Anthropometric Study of Normal Full Term Newborns at Birth in Western Rajasthan International Journal of Advanced Research (2014), Volume 2, Issue 10, 671-675, ISSN 2320-5407.

[7] Jenni Vaarno (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.

[8] Rosnah S., May L. Y., Zaleha A. M., Ahmad S., Jaafar R., Shareena I., Khadijah S., Aniza I., Idayu B. I., Saperi S. (2018): Trend of head circumference as a predictor of microcephaly among term infants born at a regional center in Malaysia between 2011 - 2015. Research and Reports in Neonatology 2018: 8 9-17. https://www.dovepress.com/by197.210.52.36.

[9] Williams Chibuzo, (2012): A multivariate analysis of climatic measurements in two cities: a case study of Enugu and Onitsha (2006-2010).