

Least Square Approximation of Percentage Body Fat for Black Women

Eguasa O., Eraikhuemen I. B. and Ehigie Loveth

Department of Mathematics and Computer Science, Benson Idahosa University, Benin City
oeguasa@biu.edu.ng & ieraikhuemen@biu.edu.ng

ABSTRACT
This work analyzes the least square model describing the response variable (Percentage Body Fat) with two independent variables, the Body Mass Index (BMI) and the Skin-fold. We show that with little disturbance of the model, generated six parameters were obtained which approximate the percentage body fat. The result obtained from MATLAB inbuilt function was compared with that of MAPLE programming and the approximation of the model was significantly equal. We also considered the goodness of fit test, the test hypothesis at five percent level of significance for $z, \bar{z}\ and\ \hat{z}$. It was discovered that it was a good fit for the three values of Z. The significance of the test shows that $z, \bar{z}\ and\ \hat{z}$ where less than the critical points, hence falls within the acceptance region of the χ^2_{tab}.

Key words: Body Mass index, Skin-fold, percentage body fat, essential body fat, storage body fat, bioelectrical impedance analysis,

1.1 INTRODUCTION

The percentage body fat percentage of a person or animal is the total weight of fat divided by total weight Body fat includes essential body fat and storage body fat. Essential body fat is necessary to maintain life and reproductive functions. The percentage of essential body fat for women is greater than that of men, due to the demands of child bearing and other hormonal functions. The percentage of essential fat is 2 – 5% in men, and 10 – 13% in women. Storage body fat consists of fat accumulation in adipose tissue, part of which protects internal organs in the chest and abdomen. The minimum recommended total body fat percentage exceeds the essential fat percentage value reported above. A number of methods are available for determining body fat percentage, such as measurement with calipers or through the use of bioelectrical impedance analysis. The body fat percentage is also a measure of fitness level, since it is the only body measurement which directly calculates a person's relative body composition without regard to height or weight. The widely used body mass index (BMI) provides a measure that allows the comparison of adiposity of individuals of different heights and weights. While Body Mass Index (BMI) largely increases as adiposity increases, due to differences in body composition. It is not an accurate indicator of body fat. For example; individuals with greater muscle mass will have higher Body Mass Index (BMI). The thresholds between "normal" and "overweight" and between "overweight" and "obese" are sometimes disputed for this reason. In this paper, we will see the body composition as the amount of fat you have, relative to lean tissue (muscles, bones, body water, organs etc). This measurement is a clearer indicator of your fitness because regardless of what your weight's, the higher percentage body fat you have, the more likely you are to develop obesity – related diseases, including heart disease, high blood pressure and stroke and type 2-diabetese.

According to Alban De Schutter et al (2011) carried out an experiment to see how body mass index correlate with body fatness in the Body Composition in Coronary Heart Disease despite its many known shortcomings, body mass index (BMI) is the most widely used measure of obesity, because of its particularity other more physiologic measurements of obesity have been proposed, including percent body fat (BF). Few studies have compared BMI and BF, especially in patients with coronary heart disease (CHD). They studied 581 patients with CHD following major CHD events. They divided patients into low (≤ 25% in men and ≥ 35% in women) as determined by the sum of the skin-fold method and compared these findings with standard BMI determinations.

Debabrata et al (1984) obtained Regression equations for the estimation of percent body fat from less direct anthropometric measures. The selection of the best repressors from 17 independent variables was performed by the maximum R^2 improvement method which is superior to conventional stepwise selection techniques. Due to the non constant variance among random subgroups of the sample for percent body fat, a weighted least squares analysis with iterations was performed. There is considerable interest in the amount of body fat because of its direct relationship to obesity. The amount of body fat can be expressed as a percentage of total body weight (percent body fat; % BF) or as the absolute weight of body fat recorded in kg (total body fat;

TBF). Truly, direct measurements of body fat are possible only by Cadavers analysis. Therefore, alternative procedures have been employed based on densitometry, gamma ray spectrometry, or hydrometrics. The method used most commonly is densitometry which involves underwater weighing, weighing in air and the measurement of residual lung volume. The measurement of body density (BD) that is obtained is used to estimate % BF or TBF using either the equation of Siri (1961) or that of Brozek et al (1963).

According to Dale Wagner et al (2000) the measurement of body composition in blacks and whites obtained biological differences which exist in the body composition in of blacks and whites. Their reviewed literature on the differences and similarities between the two races relative to fat, Free body mass (water, mineral, and protein), fat patterning and body dimensions and proportion. It was discovered that the blacks have a greater bone mineral density and body protein content than whites, resulting in a greater fat free body. Total body water (TNW) makes up the largest portion of the racial differences in the hydration of the FFB could lead to a systematic error in the estimation of %BF. Fat patterning refers to the relative distribution of subcutaneous fat on the body as opposed to absolute amounts of fat. There are racial differences in body proportions. Blacks have a greater tendency toward metamorphic and, on average, have shorter trunks and longer extremities than whites. Their review shows that the FFB of blacks and whites differs significantly.

Sparling et al (2008); investigated the possibility of predicting body fat from height, weight and skin-fold measurement for black women. Percentage body fat can be estimated by two methods: hydrostatic weighing which is also known as underwater weighing and bio-electric impedance analysis as earlier discussed. As in standard practice, height and weight enter the prediction as the fixed combination of weight divided by height squared to form a factor called Body Mass Index.

This work is to use least square approximation to determine the response variable (BF) as against two independent variables (BMI (x) and skin-fold (y)). Emphasis is on the work of Sparling et al. We will extend the method of Least Squares to linear models of r generalized independent variables $x_{1,} x_{2,...,} x_r$ and one generalized dependent or response variable Y,

$$y = a_1 x_1 + a_2 x_2 + ... + a_r x_r \tag{1}$$

Note that we can recover the two variables case by taking $r = 2$ and $\chi_2 = 1$. Assume there are n data points $(x_i, ..., x_n, y_i), i = 1, ..., n$ as before let e_i denote the error between the experimental value y_i and the predicted value,

$$e_i = y_i - (\ a_1 x_1 + a_2 x_2 + ... + a_r x_r) \qquad i = 1, ..., n. \tag{2}$$

to minimize the squared error, $E(a_1, ..., a_r) = \sum_{i=1}^{n} e_i^{\,2}$ (3)

$$\sum_{i=1}^{n} f[y_i - (a_i x_i + ... + a_r x_r)]^2$$

Using matrix notation, M^T is the matrix of data values of the independent variables.

$$M^T = \begin{pmatrix} x_{11,} & x_{12} & x_{1n} \\ x_{21} & x_{22} & x_{2n} \\ . & & \\ . & & \\ . & & \\ x_{r1} & x_{r2} & x_{rn} \end{pmatrix} \tag{4}$$

The i^{th} row of this matrix is the vector of data values of x_i. Represent the data values of the dependent variable Y as a column vector and denote the whole column.

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_n \end{pmatrix} \quad (5)$$

the system of equations can be written in matrix from as

$$M^T Ma = M^T Y \quad (6)$$

Where a is the column vector of regression parameter.

$$a = M^{-1} Y \quad (7)$$

Where

$$M^T = \begin{pmatrix} 19.36 & 19.24 & 20.76 & 21.43 & 18.72 \ldots 24.19 \\ 86.0 & 94.5 & 105.3 & 91.5 & 75.2 \ldots 148.2 \\ 1 & 1 & 1 & 1 & 1 \ \ldots \ 1 \end{pmatrix}$$

And the response vector is

$$Y^T = [19.3 \quad 22.2 \quad 24.3 \quad 17.1 \quad 19.6 \ldots 31.0]$$

The result obtained by [Sparling et al] using MAPLE programming for the least square parameters was a = 0.0065, b = 0.1507 and c = 8.074 which now yield

Percentage body fat ($\hat{Z}$) = 0.0056x + 0.1507y +8.074. (8)

2.1 Model Design

Using the derivation of least squared approximation by Sparling et al, we also derive another method where the dependent (response variable) is a function of two or more independent variables.

$Z = f(x, y)$ is a bivariate function. (9)

We use the least squares multivariate approximation to solve this problem.

Given the set of n data points (x_i, y_i, z_i)

Where $i = 1(1)n$, consider the linear polynomial

$$Z = a + bx + cy \quad (10)$$

The sum of the squares of the deviation is given by;

$$L(a,b,c) = \sum_{i=1} (e_i)^2 \quad (11)$$

Differentiating L (a, b, c) partially with respect to (a, b, c) and dividing by 2 yields.

$$\sum_{i=1}^{N} z_i - Na - b\sum_{i=1}^{N} x_i - c\sum_{i=1}^{N} y_i = 0 \quad (12)$$

a, b and c can be obtained by Gaussian elimination method.

A linear fit to a set of bivariate data may not be good enough or inadequate.

Consider the parabolic bivariate polynomial

$$Z = a + bx + cy + dx^2 + ey^2 + fxy \quad (13)$$

The sum of the squares of the deviation is given in;

$$L(a,b,c,d,e,f) = \sum_{i=1}^{N} [z_i - (a + bx_i + cy_i + dx_i^2 + ey_i^2 + fx_i y_i) \qquad (14)$$

The function is at minimum when dividing by 2 and can be written as the matrix equation

$$A\tau = K \qquad (15)$$

$$\tau = A^{-1}K = \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix} \qquad (16)$$

Where A^{-1} is the inverse of A

$$A = \begin{pmatrix} N & \sum x_i & \sum y_i & \sum x_1^2 & \sum y_1^2 & \sum x_i y_i \\ \sum x_i & \sum x_1^2 & \sum x_i y_i & \sum x_1^3 & \sum x_i y_i^2 & \sum x_1^2 y_1 \\ \sum y_i & \sum x_i y_i & \sum y_i^2 & \sum x_1 y_1^2 & \sum y_1^3 & \sum x_1 y_1^2 \\ \sum x_1^2 & \sum x_1^3 & \sum x_1^2 y_i & \sum x_1^4 & \sum \chi_1^2 y_1^2 & \sum x_1^3 y_i \\ \sum y_1^2 & \sum x_i y_i & \sum y_1^3 & \sum x_1^2 y_1^2 & \sum y_1^4 & \sum x_1^3 y_1 \\ \sum x_1 y_i & \sum x_1^2 y_i & \sum xi\, y_i^2 & \sum x_1^3 y_1 & \sum x_i y_1^3 & \sum x_1^2 y_1^2 \end{pmatrix}$$

$$K = \begin{pmatrix} \sum Z_i \\ \sum x_i y_i \\ \sum y_i Z_i \\ \sum x_1^2 Z_i \\ \sum y_1^2 Z_i \end{pmatrix}$$

3.1 APPPLICATION AND ANALYSIS OF THE METHOD

The assumed relationship is taken as

$$Z = a + bx + cy + dx^2 + ey^2 + fxy \qquad (17)$$

For some constants a, b, c, d, e, and f, consider the data collected by Sparling et al.

Body Mass (Kg/m^2) = x, Skin-fold = y ,% body fat = Z

The method of least squares to linear models of r generalized independent variables X_1, ... X_r and one general dependent variable y.

Substitute the values into the equation

$$A = \begin{pmatrix} 12 & 255 & 1384 & 5476 & 173980 & 30226 \\ 255 & 5476 & 30226 & 118832 & 3898497 & 666785 \\ 1384 & 30226 & 173980 & 666785 & 23608286 & 3898497 \\ 5476 & 118832 & 666785 & 3690284638 & 88163923 & 14864681 \\ 173980 & 30226 & 23608286 & 88163923 & 3690284638 & 870982676 \\ 30226 & 666785 & 3898497 & 14864681 & 870982676 & 88163923 \end{pmatrix}$$

$$K = \begin{pmatrix} 307 \\ 6643 \\ 37587 \\ 145494 \\ 4751820 \\ 836450 \end{pmatrix}$$

$$Z = a + bx + cy + dx^2 + ey^2 + fxy \quad (18)$$

Applying MATLAB inbult function $\tau = A^{-1}K$

$>> A = [12\ 255 \ldots 30226;\ 255\ 5476 \ldots 666785;\ 1384\ 30226 \ldots 3898497;\ 5376\ 118832 \ldots 14864681;\ 173980$
$30226 \ldots 870982676;\ 30226\ 666785 \ldots 88163923];$

$>> K = [307; 6643; 37587; 145494; 4751820; 836450];$

$>> \tau = A / K$

$\tau =$

13.5863
1.4389
−0.1981
−0.0797
0.0000
0.0161

taking

a = 13.5863, b = 1.4389, c = − 0.1981, d = − 0.0797, e = 0.0000, f = 0.0161

substitute the values into the equation (18)

$$\bar{Z}_i = 13.5863 + 1.4389x - 0.1981y - 0.0797x^2 + 0.0000y^2 + 0.0161xy$$

$\bar{Z}_1 = 21.3, \bar{Z}_2 = 22.3, \bar{Z}_3 = 23.4, \bar{Z}_4 = 21.3, \bar{Z}_5 = 20.4, \bar{Z}_6 = 22.2, \bar{Z}_7 = 30.9,$
$\bar{Z}_8 = 19.8,\ \bar{Z}_9 = 25.4,\ \bar{Z}_{10} = 34.3,\ \bar{Z}_{11} = 34.4,\ \bar{Z}_{12} = 30.1.$

4.1 TEST STATISTICS

$$\chi^2_{cal} = \sum_{l=12}^{k} \frac{(z_o ? \bar{z}_e)^2}{z_e} \quad (19)$$

Z_0 = Observed frequency, $\bar{z}_e$ = Expected frequency
From the table,
Let % body fat

$\bar{z}_e = 25$

Substitute in the values into the equation

$$\chi^2_{cal} = \sum_{l=12}^{k} \frac{(z_o ? \bar{z}_e)^2}{z_e} = 13.164$$

$\therefore \chi^2_{cal} = 13.0000$

The results gotten from the least square method from the above solutions are;
∴ The test statistics for the above least square results are
(Ƶ) % body fat = 14.2703
($\bar{Ƶ}$) % of body fat = 13.000

$\bar{Z}_e = 26$

Taking $Ƶ_0$ from $Ƶ_1 – Ƶ_{12}$
$Ƶ_1 = 19.3$, $Ƶ_2 = 22.2$, $Ƶ_3 = 24$, $Ƶ_4 = 17.1$, $Ƶ_5 = 19.6$, $Ƶ_6 = 23.9$, $Ƶ_7 = 29.5$, $Ƶ_8 = 24.1$, $Ƶ_9 = 26.2$
$Ƶ_{10} = 33.7$, $Ƶ_{11} = 36.2$, $Ƶ_{12} = 31.0$
Substitute in the values into the evaluation

$$\chi^2_{cal} = \sum_{l=1}^{k} \frac{(z_o ? \bar{z}_e)^2}{z_e} = 14.2703$$

From Sparling et al, we investigating the possibility of predicting body from height, weight and Skin-fold measurement for black women. Using MAPLE program we get the following result.
a = 0.00656 b = 0.1507 c = 8.074
Thus, we find that
Percent-body-fat ≈0.0065 x body-mass-index + 0.1507 x skin-fold + 8.074
And we can use the table I we have above.

Let percent-body-fat = $\hat{z}$

∴ $\hat{Z}$ = 0.0065 x body-mass-index + 0.1507 x skin-fold + 8.074.

Substitute the values into the equation when body-mass-index = x
Skin-fold = y.

$\hat{Z}_1 = 21.2$, $\hat{Z}_2 = 22.4$, $\hat{Z}_3 = 24.1$, $\hat{Z}_4 = 22.0$, $\hat{Z}_5 = 19.5$, $\hat{Z}_6 = 22.2$, $\hat{Z}_7 = 31.7$, $\hat{Z}_8 = 19.5$

$\hat{Z}_9 = 26.3$, $\hat{Z}_{10} = 33.7$, $\hat{Z}_{11} = 33.9$, $\hat{Z}_{12} = 30.6$

Adding the results of Sparling et al, we now have another table as follows;

TEST STATISTICS FOR $\hat{Z}$

$$\chi^2_{cal} = \sum_{l=1}^{k} \frac{(z_o \; ? \; \hat{z}_e)^2}{\hat{z}_e}$$

$$\hat{z} = 26$$

Where $\hat{Z}_1 = 21.2, \hat{Z}_2 = 22.4, \hat{Z}_3 = 24.1, \hat{Z}_4 = 22.0, \hat{Z}_5 = 19.5, \hat{Z}_6 = 22.2, \hat{Z}_7 = 31.7,$

$\hat{Z}_8 = 19.5, \hat{Z}_9 = 26.3, \hat{Z}_{10} = 33.7, \hat{Z}_{11} = 33.9, \hat{Z}_{12} = 30.6$

Substitute in the values into the equation

$$\chi^2_{tab} = 13.000$$

4.1.1 TEST HYPOTHESIS

$H_0 : p_1 = p_2 = ... = p_{12}$

H_1: At least one of the equality does not hold.

H_0 = Null Hypothesis, H_1 = Alternative Hypothesis

∝ is the level of significance

K is the degree of freedom

$\chi^2_{1-\alpha,k-1} = \chi^2_{0.95,11} = 19.6751$

Conditions

1. If $\chi^2_{cal} > \chi^2_{tab}$ reject H_0. That is the % body fat is not Desirable

2. If $\chi^2_{cal} < \chi^2_{tab}$ we do not reject H_0. This is the % body is Desirable for all measurement for the estimated body fat. Hence, it is a good fit.

Body fat is the amount (usually expressed as a percentage) of fat in a person's body. Body fat plays an important role in human healthy and some is necessary for maintaining many bodily process and functions.

5 SUMMARY AND CONCLUSION

The estimation of percentage body fat was calculated only for black women. Using data from Sparling et al. We investigate body fat from body mass index and skin-fold. The parameter for the least square approximation was three to predict the response variable (z) with the aid of Maple programming language. We also analyse the same data using a modified least square approximation with six parameters and the result obtained from our analysis was same as that of the estimated values of Sparling et al. We hereby conclude that the results from both analysis (chi square and test statistics) was approximately equal signifying a good fit. The estimation of percentage body fat was calculated with the relationship between percentage body fat $z, \bar{z}\ and\ \hat{z}$ and a test statistics was carried out to know if the test for percentage body fat is desirable for all measurement for estimated body fat and the goodness of fit test. From the result of chi-square, we say that percentage body fat is desirable for all measurement for the estimated body fat, hence it is a good fit. In the estimation of percentage body fat, there is no single ideal of percentage of body fat for everyone. Levels of body fat dependent on sex, race, and age. We will recommend that the above levels of body should also be considered and other methods for calculating percentage body fat should also be considered.

ISSN 2224-5804 (Paper) ISSN 2225-0522 (Online)
Vol.3, No.14, 2013

APPENDIX

Table 1: Height, Weight, Skin-fold and Percentage Body Fat for Black Women

Height (m)	Weight (Kg)	Body Mass (Kg/m^2)	Skin-fold	% Body Fat
63. 0	109.3	19.36	86.0	19.3
65.0	115.6	19.24	94.5	22.2
61.7	112.4	20.76	105.3	24.3
65.2	119.6	21.43	91.5	17.1
66.2	116.7	18.72	75.2	19.6
65.2	114.0	18.85	93.2	23.9
70.0	152.2	21.84	156.0	29.5
63.9	115.6	19.90	75.1	24.1
63.2	121.3	21.35	119.8	26.2
68.7	167.7	24.98	169.3	33.7
68.0	160.9	24.46	170.0	36.2
66.0	149.9	24.19	148.2	31.0

Source: Sparling et al.

Table 2: Data values of the dependent variable Y as a column vector

S/N	~~Z~~	X ~~Y~~	Y ~~Z~~	X^2 ~~Z~~	Y^2 ~~Z~~
1	19.3	374	1660	7238	142743
2	22.2	421	2098	8214	198246
3	24.3	504	2559	10473	26438
4	17.1	366	1565	7849	143161
5	19.6	367	1474	6860	110838
6	23.9	451	2227	8485	207594
7	29.5	644	4602	14072	717912
8	24.1	480	1810	9544	135924
9	26.2	559	3139	11947	376022
10	33.7	842	5705	21029	965909
11	36.2	885	6154	21648	1046180
12	31.0	750	4594	18135	680853

Table 3: Data for Body Mass and Skin-Fold

S/N	X	Y	Xy	X²	Y²	X²y	Xy²	X³	Y³	x²y²	X³y	xy³	X⁴	Y⁴
1	19.36	86	1665	375	7396	32250	143187	7256	636056	2773500	624016	12314044	140482	54700816
2	19.24	94.5	1818	370	8930	34965	171813	7122	843909	3304100	673029	16236809	137031	79749365
3	20.76	105.3	2186	431	11088	45384	230187	8947	1167576	4778928	942119	24238878	185795	122945740
4	21.43	91.5	1961	459	8372	41999	179412	9842	766061	3842748	900543	16416687	210906	70094570
5	18.72	75.2	1408	350	5655	26320	105862	6560	425259	1979250	493312	7960848	122807	31979477
6	18.85	93.2	1757	355	8686	33086	163731	6698	809558	3083530	624254	15260168	126254	75450765
7	21.84	156	3407	477	24336	74412	531498	10417	3796416	11608272	1625052	82913725		592240896
8	19.9	75.1	1494	396	5640	29740	112236	7881	423565	2233440	591863	8428944	156824	31809713
9	21.35	119.8	2558	456	14352	54629	306415	9732	1719374	6544512	1165894	36708635	207774	205981052
10	24.98	169.3	4229	624	28662	105643	715977	15588	4852560	17885088	2639048	121216949	389376	821538333
11	24.46	170	4158	598	28900	101660	706894	14634	4913000227515	17282200	2487780	120171980	357953	835210000
12	24.19	148.2	3585	585	21963	86697	531285	14155	3254952	12848355	2097771	78737289	342407	482383911

Table 4: Estimated value for $\bar{Z}$

Body Mass (kg/m^2) (x)	Skin-fold (y)	% body fat (Z)	$\bar{Z}$
19.36	86.0	19.3	21.3
19.24	94.5	22.2	22.3
20.76	105.3	24.3	23.4
21.43	91.5	17.1	21.3
18.72	75.2	19.6	20.4
18.85	93.2	23.9	22.2
21.84	156.0	29.5	30.9
19.90	75.1	24.1	19.8
21.35	119.8	26.2	25.4
24.98	169.3	33.7	34.3
24.46	170.0	36.2	34.4
24.19	148.2	31.0	30.1

Table 5: Comparison of estimated percentage Body Fat for (z , $\bar{z}$, $\hat{z}$).

Body Mass (Kg/M2(x))	Skin-fold (y)	% body fat (z)	% body fat($\bar{z}$)	% body fat ($\hat{z}$)
19.36	86.0	19.3	21.3	21.2
19.24	94.5	22.2	22.3	22.4
20.76	105.3	24.3	23.4	24.1
21.43	91.5	17.1	21.3	22.0
18.72	75.2	19.6	20.4	19.5
18.85	93.2	24.9	22.2	22.2
21.84	156.0	29.5	30.9	31.7
19.90	75.1	24.1	19.8	19.5
21.35	119.8	26.2	25.4	26.3
24.98	169.3	33.7	34.3	33.7
24.46	170.0	36.2	34.4	33.9
24.19	148.2	31.0	30.1	30.6

REFERENCES

Alban De Schutter (2011): How Body Mass Index Correlate with Body Fatness in the Body Composition in Coronary Heart Disease.

Brozek, Josef, Grande, Francisco, Anderson, Joseph T, Keys Ancel (2006): “Densitometry Analysis of Body Composition. New York.

Debabrata Mukherjee and Alex F. Roche (1984): The Estimation of Percent Body Fat, Body Density and Total Body Fat by Maximum R^2 Regression Equations.

Dale R. Wagner and Vivian H. Heyward (2000): Measurement of Body Composition in Blacks and Whites. American Journal of clinical nutrition , June 2000, Vol. 71 (6) 1392 – 1402.

Pat Ferguson (2013): Researchers Weigh In On the Body Fat versus BMI Debate

Sparling, P. B., Melinda Millard-Stafford, Linda B.R., Linda D. AND Bryan T.H. (2008): **“**Body composition by bioelectric impedance and densitometry in black women” American Journal of Human Biology 5, pp 111-117, 1993.