

Fitting the Ehrenberg's Law-Like Equation Using the Age, Weight and Height of Secondary School Students

Nwagwu Chibuikem Chrysogonus^{*}, Nduka Ethelbert Chinaka, Ogoke Uchenna Petronilla

Department of Mathematics and Statistics, University of Port Harcourt, Port Harcourt, Nigeria

Email address:

pristan14@gmail.com (N. C. Chrysogonus)

^{*}Corresponding author

To cite this article:

Nwagwu Chibuikem Chrysogonus, Nduka Ethelbert Chinaka, Ogoke Uchenna Petronilla. Fitting the Ehrenberg's Law-Like Equation Using the Age, Weight and Height of Secondary School Students. *International Journal of Statistical Distributions and Applications*. Vol. 5, No. 1, 2019, pp. 10-14. doi: 10.11648/j.ijstd.20190501.13

Received: April 4, 2019; **Accepted:** May 21, 2019; **Published:** June 5, 2019

Abstract: The Ehrenberg's law like equation is deemed to fit (within minor error) universally among children aged 11 – 17. This paper adopts this equation to confirm or deny the universal applicability based on selected schools in South-South region in Nigeria which is believed to be environmentally degraded. Data on Age, Sex and anthropometric measures (weight and height) of the cohort in public and private schools were used to fit the model. After the data were pooled from the public and private school, Microsoft Excel Package was used to test how different conditions may influence the fit of the model to the data. The conditions considered were: Gender and Social class (determined by school type), hence there were five different data classifications. Results obtained showed that although the average deviation recorded for female students fell out of the expected limit of ± 0.01 although marginally by another $+0.01$; all other subsets of the data set were found to fit the given equation appropriately. This goes to show that irrespective of gender or social class, provided the students are within the specified age bracket, the Ehrenberg's law like equation will be a good model for use. This is equally verified using the test of significance as well as the test for equality of the regression models.

Keywords: Anthropometry, Model Adequacy, Law-Like Relationships, Statistical Models

1. Introduction

Scientists of all civilization, who existed until today, examined the human body using anthropometric methods. For these reasons, anthropometric data are used in many contexts to screen for or monitor disease. Anthropometry, a branch of morphometry, is the study of the size and shape of the components of biological forms and their variations in populations [1]. Among the works of Andrew Ehrenberg was the propounding of the Ehrenberg's Law-like Equation which sought to establish a relationship between the weight and height of secondary school students within given age ranges. This equation has been shown to be effective in different countries provided the conditions are strictly obeyed. Due to its law-like nature, the equation is not merely a summary of a specific set of data or a simple correspondence. Unlike other equations of the physical sciences (and others), the Ehrenberg's equation must hold under the given conditions. To this effect, extrapolation may prove ineffective outside the range of ages.

1.1. Statement of the Problem

The model being quite new in modern research has seen only a few countries run comparative study with it. Hence, the presumed universal applicability of this law-like equation has not been absolute with respect to all situations and countries. Using the Cole's Lambda-Mu-Sigma (LMS) method for smoothing the growth curves, [2] showed that human growth differs among the various ethnic groups since Chinese boys' weight is strikingly heavier than that of the WHO standard at age 6-10 years. The height is taller than that of the WHO for boys below 15 years and for girls below 13, but is significantly lower when boys over 15 years and girls over 13. Also, those within the same age bracket (11 – 17) living in the hilly region of India did not fit the Ehrenberg's equation even though previous researches for children in other regions do. If within the same country topographic difference could affect the results, other environmental factors can equally do same. Some studies

done in some African countries obey this law-like equation – Botswana [3] and Ghana [4]. It cannot however be assumed to hold true for a vast country like Nigeria with six geopolitical zones and different topographic, demographic and environmental patterns. This led to the need to localise the research to a South-South community especially due to the environmental degradation (pollution) caused by oil exploration and exploitation activities. Also, although a Nigerian community (South East) has been tested by Uche, he only considered the overall data set in the work; this research went further by classifying the data with respect to gender and social class in order to reduce the ambiguity when it comes to using the results of the research in decision making [5]. The new information could be used to assess the growth and nutrition of children and adolescents, especially those with disease that affects their body weight.

1.2. Scope of the Research

This research was conducted for public and private secondary school students the South-South region of Nigeria. It was unbiased as it included both male and female students randomly selected in each case. The information obtained as required were: Age, Weight, Height and Gender. The results of this research are applicable primarily to the community under study. The research went as far as classifying the data set into five (5) subsets and comparing each with the given equation. With a larger sample size (given the number of schools in the community), a more adequate and reliable model may have been established. However, this research was limited to the aforementioned scope due to time constraint given the period within which it must be concluded and finance, considering the location of the schools.

1.3. Review of Literature

Forchheh contributed greatly to this topic especially through his work on Ehrenberg Law-like relationship and anthropometry where he said: “In anthropometry, a weight-height relationship is used to measure the nutritional status of children. Differences between the weight-for-height relationship in anthropometry and the law-like relationship between weight and height for children are small and can be attributable to functional differences” [3]. Government and child welfare officers routinely use anthropometric indices to monitor and assess the nutritional status of children. In Botswana, for example, each child is expected to attend monthly medical checkups from birth until they are 5 years old. The main purpose is to weigh the child and record the weight on a growth chart which shows the expected weight of the child at each time point together with upper and lower limits of this expected value. If the child falls outside the lower limit, then the child qualifies for the Government-sponsored intervention programme on malnutrition. Weight-for-height is particularly useful because it can still be calculated when the age of the child is either unknown or is unreliable, a situation often faced by aid workers during intervention programmes in rural and war-torn communities”.

The formal use of a law-like relationship to characterise the relationship between the weight and height of children was first proposed by A.S.C. Ehrenberg, stating that the log-weight of a child lies between $0.8H + 0.36$ and $0.8H + 0.44$ [6]. Ehrenberg's law-like relationship was derived by using data collected on Birmingham boys aged 5 to 13 years and from one social group. It was then validated by using data from samples of:

- Birmingham boys of the same age group from two social classes.
- Canadian (Ottawa) boys ages 6 – 11 from three time periods (1933 – 1935, 1938 – 1940, 1953 – 1955) and classified as either above or below average socio-economic status.
- Canadian (Ottawa) girls in 1933 – 1935 also classified as above or below average socio-economic status.
- French boys and girls aged $5\frac{1}{2}$, $8\frac{1}{2}$, and $11\frac{1}{2}$ year old classified into seven groups according to the occupation of the head of the household [3].

The definitive statement of Andrew Ehrenberg's position on generalisation came much later in “Predictability and Prediction”, read to a meeting of the Royal Statistical Society. The summary stated boldly: “A result can be regarded as routinely predictable when it has occurred consistently under a known range of different conditions. This depends on the previous analysis of many sets of data, drawn from different populations. There is no such basis of extensive experience when a prediction is derived from the analysis of only a single set of data” [7].

The objective of a Pierre's paper was to define simple formulas to calculate the ideal weight for height of children and adolescents aged 2 to 18 years and came up with the rule being that a gain in H of 5cm would lead to a weight gain of 10% and a gain of 35cm a doubling of the body weight [8].

When Lovell carried out the test on Hutu and Tutsi boys living in Rwanda in 1957 – 1958, the slope was the same but the intercept was slightly different [9]. Drawing from Kpedekpo's results, Lovell suggested a modification of the law-like relationship to $\log W = 0.8H + c$ for age specific groups of children where c was dependent on the community. The value of c was 0.39 for Birmingham children [6] but only 0.36 for non-privileged and rural Ghanaian children [4].

Ehrenberg's law-like equation has been fitted to data from many countries. Taking a case study of Nigerian children, it was found that they fall within the range observed for other countries [5].

Height-weight relationship in young healthy Indian males (19 – 22 yrs.) has been established. The results of the present study show that Ehrenberg's equation of the form: $\log W = 0.8H + 0.4$ cannot be extended beyond the age of 18 years in young healthy Indian males. [10]

2. Method

This research work is centered on the Ehrenberg's model given by:

$$\log \bar{W} = 0.8\bar{H} + 0.4 \pm 0.01 \quad (1)$$

The above model was arrived at using the simple linear regression. In the above model,

1. $\log \bar{W}$ represents the response or dependent variable (weight).
2. 0.8 is the regression coefficient
3. \bar{H} represents the predictor or explanatory or independent variable (height).
4. 0.4 is the constant term of the regression.

When a mathematical equation is used to describe a relationship between two variables x and y , the observed values do not generally lie exactly on the line. For this reason, an equation like $\log W = 0.8H + 0.4$ should actually be $\log W \cong 0.8H + 0.4$; where the symbol \cong means *approximately* equal. The equation could also be written.

$$\log \bar{W} = 0.8\bar{H} + 0.4 \pm c \quad (2)$$

Here, the “plus or minus” symbol \pm indicates that some observed readings have positive and others negative deviations from the theoretical line $\log \bar{W} = 0.8\bar{H} + 0.4$, and that these deviations have an average size of c units. In the initial stages of data analysis, a systematic pattern in the

deviations means that the wrong descriptive relationship has been fitted, even if the deviations are small, only in more advanced work are systematic deviations sometimes acceptable as deliberate oversimplifications, as long as their nature is understood.

The Ehrenberg's model is an example of the semi-log model with the log taken to base-10. It is a linear relationship hypothesized between a log transformed outcome variable and predictor variable (s). R^2 has the same meaning here as usual – the proportion of variance in y explained by the model.

Tests of Hypothesis: A test of hypothesis was carried out to examine if there is any difference in the parameters of two different models. This is usually necessary in cases where the given data set has varying predicted influential factors like fertilizer type, social class, gender, age, etc. The researcher may be interested in comparing the different models in other to determine how much these extraneous factors actually affect the outcome of the experiment. In more practical terms, the procedure is simply to compute the F-statistic:

3. Result

Table 1. Overall data summary.

Age	11	12	13	14	15	16	17	Average
\bar{W}	36.563	40.286	42.844	47.337	52.015	55.432	57.355	47.405
\bar{H}	1.433	1.516	1.547	1.592	1.620	1.641	1.672	1.574
$\log \bar{W}$	1.563	1.605	1.632	1.675	1.716	1.744	1.759	1.671
$0.80\bar{H} + 0.41$	1.546	1.612	1.638	1.673	1.696	1.712	1.738	1.659
Deviation	0.019	-0.007	-0.006	0.002	0.020	0.031	0.019	0.011

Table 2. Private school data summary.

Age	11	12	13	14	15	16	17	Average
\bar{W}	36.143	40.971	45.387	51.341	55.821	59.333	64.500	50.499
\bar{H}	1.423	1.521	1.575	1.604	1.645	1.680	1.726	1.596
$\log \bar{W}$	1.558	1.612	1.657	1.710	1.747	1.773	1.810	1.695
$0.82\bar{H} + 0.39$	1.557	1.637	1.681	1.705	1.739	1.767	1.806	1.699
Deviation	0.002	-0.024	-0.024	0.005	0.008	0.006	0.002	-0.0037

Table 3. Public school data summary.

Age	11	12	13	14	15	16	17	Average
\bar{W}	37.364	38.733	40.455	43.689	46.519	51.737	55.640	44.877
\bar{H}	1.451	1.504	1.522	1.545	1.584	1.604	1.659	1.553
$\log \bar{W}$	1.572	1.588	1.607	1.640	1.668	1.714	1.745	1.648
$0.80\bar{H} + 0.40$	1.561	1.603	1.617	1.636	1.667	1.683	1.727	1.642
Deviation	0.016	-0.015	-0.010	0.004	0.000	0.031	0.016	0.006

Table 4. Female students data summary.

Age	11	12	13	14	15	16	17	Average
\bar{W}	36.571	41.259	44.364	51.756	53.867	55.600	53.556	48.139
\bar{H}	1.427	1.538	1.562	1.598	1.613	1.599	1.643	1.568
$\log \bar{W}$	1.563	1.616	1.647	1.714	1.731	1.745	1.729	1.678
$0.77\bar{H} + 0.48$	1.541	1.630	1.649	1.679	1.690	1.680	1.714	1.655
Deviation	0.022	-0.015	-0.002	0.035	0.041	0.066	0.014	0.023

Table 5. Male students data summary.

Age	11	12	13	14	15	16	17	Average
\bar{W}	36.722	39.091	41.226	43.311	50.472	55.235	58.955	46.430
\bar{H}	1.437	1.488	1.532	1.550	1.627	1.689	1.684	1.573
$\log \bar{W}$	1.565	1.592	1.615	1.637	1.703	1.742	1.771	1.661
$0.83\bar{H} + 0.35$	1.550	1.591	1.626	1.640	1.701	1.751	1.747	1.658
Deviation	0.015	0.002	-0.011	-0.004	0.002	-0.009	0.023	0.003

Table 6. Test of significance of social class.

Pr	Pu	X	EPr	EPu	ePr	e ² Pr	ePu	e ² Pu	EPo	ePoPr	ePoPu	e ² PoPr	e ² PoPu
1.76	1.71	1.67	1.76	1.74	0.004	0.000	0.053	0.003	1.748	0.016	-0.040	0.000	0.002
1.74	1.81	1.70	1.78	1.76	-0.044	0.002	0.065	0.004	1.772	-0.032	0.041	0.001	0.002
1.73	1.79	1.65	1.74	1.72	-0.011	0.000	-0.022	0.000	1.732	0.001	0.054	0.000	0.003
1.79	1.78	1.75	1.83	1.80	-0.033	0.001	0.012	0.000	1.813	-0.020	-0.034	0.000	0.001
1.86	1.72	1.63	1.73	1.70	0.137	0.019	0.051	0.003	1.715	0.148	0.001	0.022	0.000
1.78	1.76	1.64	1.73	1.71	0.043	0.002	-0.037	0.001	1.723	0.055	0.040	0.003	0.002
1.62	1.64	1.60	1.70	1.68	-0.079	0.006	-0.033	0.001	1.691	-0.068	-0.048	0.005	0.002
1.60	1.62	1.57	1.68	1.66	-0.075	0.006	0.011	0.000	1.667	-0.065	-0.043	0.004	0.002
1.66	1.64	1.54	1.65	1.63	0.010	0.000	0.007	0.000	1.642	0.020	0.001	0.000	0.000
1.64	1.62	1.52	1.64	1.62	0.007	0.000	0.000	0.000	1.626	0.017	-0.003	0.000	0.000

Table 7. Test of significance of gender.

Female	Male	X	EFe	EMa	eFe	e ² Fe	eMa	e ² Ma	EPo	ePoFe	ePoMa	e ² PoFe	e ² PoMa
1.57	1.56	1.45	1.5965	1.5535	-0.027	0.001	-0.088	0.008	1.58	-0.010	-0.020	0.000	0.000
1.63	1.64	1.66	1.7582	1.7278	-0.128	0.016	0.012	0.000	1.748	-0.118	-0.108	0.014	0.012
1.79	1.79	1.72	1.8044	1.7776	-0.014	0.000	0.025	0.001	1.796	-0.006	-0.006	0.000	0.000
1.69	1.67	1.56	1.6812	1.6448	0.009	0.000	-0.015	0.000	1.668	0.022	0.002	0.000	0.000
1.62	1.58	1.5	1.635	1.595	-0.015	0.000	-0.051	0.003	1.62	0.000	-0.040	0.000	0.002
1.64	1.61	1.58	1.6966	1.6614	-0.057	0.003	-0.035	0.001	1.684	-0.044	-0.074	0.002	0.005
1.76	1.66	1.62	1.7274	1.6946	0.033	0.001	-0.026	0.001	1.716	0.044	-0.056	0.002	0.003
1.69	1.61	1.55	1.6735	1.6365	0.017	0.000	0.000	0.000	1.66	0.030	-0.050	0.001	0.003

where: Pr = Private Pu = Public EPr = Estimate of Private EPu = Estimate of Public ePr = Error of Private ePu = Error of Public EPo = Estimate of Pooled data. Ma = Male Fe = Female EMa = Estimate of Male EFe = Estimate of Female eMa = Error of Male eFe = Error of Female EPo = Estimate of Pooled data.

4. Discussion of Results

4.1. Models for Each Data Set

After obtaining the coefficient and intercept for the data sets using Excel, the following models were got:

Overall model: $\log \bar{W} = 0.80\bar{H} + 0.41$ (Table 1).

Public School model: $\log \bar{W} = 0.80\bar{H} + 0.40$ (Table 3).

Private School model: $\log \bar{W} = 0.82\bar{H} + 0.39$ (Table 2).

Male Students model: $\log \bar{W} = 0.83\bar{H} + 0.35$ (Table 5).

Female Students model: $\log \bar{W} = 0.77\bar{H} + 0.48$ (Table 4).

Comparing with the Ehrenberg model: $\log \bar{W} = 0.8\bar{H} + 0.4 \pm 0.01$, we see that all the models resemble it although some deviate significantly (like the gender-based models). However, the overall data set fits the Ehrenberg model perfectly.

4.2. Average Deviation for Each Data Set

In line with a study of Ehrenberg's model, we check to see if the data sets are within the limit ± 0.01 .

Overall model: 0.011 (Table 1).

Public School model: 0.006 (Table 3).

Private School model: -0.0037 (Table 2).

Male Students model: 0.0026 (Table 5).

Female Students model: 0.023 (Table 4).

It is clear from the above that all the models are within the limit except that of the female students. This can be accounted for when we make recourse to a marginal increase in the limit or when we consider that taking the confidence interval of the parameters may reveal a lower value.

4.3. Test of Hypothesis

To do this, the data set was grouped based on social class and gender in order to test if these factors have any influence on the parameters.

From Table 6, we have the Sum of Square of Error for Private as 0.036 and that of Public is 0.013. The F-calculated and F-tabulated are 0.179 and 4.49 respectively while from Table 7, the Sum of Squares of Error for Female students is 0.022 and that of the Male is 0.013. The F-calculated and F-tabulated are 2.939 and 4.75 respectively.

i. Hypothesis:

H_0 : Social Class (or Gender) does not influence the model parameters.

H_1 : Social Class (or Gender) influences the model parameters.

- ii. Decision: In both Table 6 and 7, the $F\text{-calc} < F\text{-tab}$. Hence, we accept H_0 which implies that both factors do not influence the model parameters. This buttresses the fact that irrespective of the factor, the data still fit the Ehrenberg model.

4.4. Summary

From the results, it was evident that the data set for the students fit the Ehrenberg model and the average deviation was equally within the specified limit. The tests of equality also showed that whether the student is in the public or private school; is male or female, the relationship between the weight and height will fit the model. R^2 of 0.653 was obtained which implies that 65.3% of the total variation in the data set is explained by the model.

5. Conclusion

Based on the results of this research work, it is clear that the Ehrenberg equation can be extended to the population of secondary school students in the South-South region of Nigeria. The case of the female students is similar to an example given by a researcher who concluded that since the Body Mass Index (BMI) of certain people deviate only a little from the normal, they may be considered as overweight instead of obese [11]. We can conclude that irrespective of gender or social class, provided the students are within the specified age bracket, the Ehrenberg's law like equation will be a good model for use. This was equally verified using the test of hypothesis which showed that those demographic indices (gender and social class) have no significance on the model. This research can be extended to other communities to further test its breaking point. Also, further research may look to perturb the model so as to obtain a model with least deviation.

References

- [1] N. Utkualp and I. Ercan, "Anthropometric Measurements Usage in Medical Sciences," *BioMed Research International*, vol. 2015.
- [2] X. N. Zong and H. Li, "Construction of a New Growth References for China Based on Urban Chinese Children: Comparison with the WHO Growth Standards," *PLoS One*, vol. 8, no. 3, March 2013.
- [3] N. Forchheh, "Ehrenberg Law-like Relationship and Anthropometry," *Journal of the Royal Statistical Society Series A*, vol. 165, no. 1, pp. 155-172, 2002.
- [4] G. M. K. Kpedekpo, "Heights and Weights of Children in Ghana," *Journal of the Royal Statistical Society, Series A*, vol. 133, pp. 86-93, 1970.
- [5] P. I. Uche, "Heights and Weights of Nigerian Children: The Fit of Ehrenberg's Equation," *Biometrics*, vol. 37, pp. 149-151, 1981.
- [6] S. C. Ehrenberg, "The Elements of Law-like Relationships," *Journal of the Royal Statistical Society, Series A*, vol. 131, pp. 280-329, 1968.
- [7] S. C. and Bound, J. A. Ehrenberg, "Predictability and Prediction," *Journal of the Royal Statistical Society, Series A*, vol. 156, no. 2, pp. 167-206, 1993b.
- [8] M. B. Pierre, "Weight for Height during Growth, Useful Formulas to Insert into the Pediatrician's Smart Phone," *Academic Journal of Pediatrics and Neonatology*, vol. 4, no. 2, 2017.
- [9] H. G. Lovell, "Heights and Weights of West Indian Children," *Journal of the Royal Statistical Society Series A*, vol. 135, pp. 569-578, 1972.
- [10] S. S. Verma, N. Kishore, S. S. Purkayasha, and R. K. Ghagiok, "Height-Weight Relationship in Young Healthy Indian Males," *Indian Journal of Medical Sciences*, vol. 44, no. 4, pp. 90-92, April 1990.
- [11] Lewis Tanya. (2013) Live Science. [Online]. www.livescience.com/39097-bmi-not-accurate-health-measure.html
- [12] J. A Bound, "The Contribution of Andrew Ehrenberg to Social and Marketing Research," *Journal of Empirical Generalisations in Marketing Science*, vol. 12, no. 1, pp. 2-5, 2009.
- [13] S. C. Ehrenberg, "Even the Social Sciences Have Laws," *Nature*, vol. 365, p. 385, 1993a.
- [14] S. C. Ehrenberg, "Children's Heights and Weights in 1905," *Journal of the Royal Statistical Society, Series A*, vol. 138, pp. 239-241, 1975.
- [15] C., Wang, H., Lu, N., Chen, T., He, H., Lu, Y., and Tu, X. M. Feng, "Log-transformation and Its Implications for Data Analysis," *Shanghai Archives of Psychiatry*, vol. 26, no. 2, pp. 105-109, 2014.
- [16] Charkravarthy P. K., Suresh G., Chenna D., and Chenna V., "Relationship Between Anthropometric Measures and Dental Caries Among Adolescent National Cadets Corps of Udipi District, South India," *Journal of Natural Science, Biology and Medicine*, January 2013.
- [17] E. C. Nduka and U. P. Ogoke, *Principles of Applied Statistics: Regression and Correlation Analysis*, 2nd ed. Owerri, Imo State, Nigeria: Munagenesis Concepts, 2016.