Journal of the Nigerian Society for Mathematical Biology
Volume 2 (2019), pp. 32 – 40

The effect of mothers' age, sex and weight of the child in predicting caesarean delivery among women

V. Adah[*†] and D. Achimugu[*]

Abstract

Labor in child bearing is the onset of regular contractions and cervical change. The study aimed at investigating the probability of normal delivery given the sex, weight of Child and Age of mother. This research work covers the record of babies delivered in the Federal Medical Center Makurdi between the months of June to September 2016. Data was collected from the delivery register of the maternity ward/Labor room. The study employed the logistic regression to determine the probability of caesarean section among women given the Sex and Weight of the baby, and Age of the mother. The study revealed that factors such as the sex of child, weight of child, and age of the mother fits the model with p-values: 0.39089, 0.00422, and 0.00739 respectively, with the weight of the child and age of the mother are significant in predicting caesarean delivery. The likelihood ratio test shows that the fitted model adequately fits the data. In conclusion pregnant mothers should avoid foods that increases baby's weight and aged pregnant mothers should attend antenatal regularly so they could be given the necessary medical attention.

Keywords: Labor; logistic regression; odds ratio; response probability

1 Introduction

Labor is defined as the onset of regular contractions and cervical change. It is traditionally divided into three stages. The first stage encompasses the onset of labor to the complete dilatation of the cervix, and is subdivided into latent and active phases. The active phase begins when the rate of cervical dilatation accelerates, which occurs at 4 cm on average. The second stage consists of the time from complete dilatation of the cervix to delivery of the infant. The third stage is complete at the delivery of the placenta. The original labor curves were plotted by Friedman in the 1950s and are the basis for describing prolonged labor [14].

One of the challenges of modern obstetrics in developing countries is the large number of women who expect a normal vaginal delivery without obstetric intervention, only to end up with a highly

[*] Department of Mathematics/Statistics/Computer Science, University of Agriculture Makurdi, Nigeria.
[†] Corresponding Author: adahvictoria14@yahoo.com

complicated birth, resulting in a view of personal failure and/or suboptimal care. In addition, some of these women are left with long-term disease due to strain on the pelvic floor [4].
There is an increasing demand for optional cesarean section, driven by a desire to avoid traumatic childbirth and future pelvic floor injuries ([9],[10]). However, it is currently impossible to identify those women who are mostly at risk of these outcomes. Previous attempts using clinically or sonographically predicted birth weight have been as favorable as it should be [8]. There are widespread fears surrounding vaginal delivery of the breech babies and lack of information generally available on safe vaginal delivery of such babies. These factors, together with the prevailing myths and beliefs that caesareans guarantee healthy babies, more often than not leave the woman with no option but to accept the decisions made for her by her obstetrician. Mothers may not be aware much earlier than 36-37 weeks that their baby remaining in a breech position is a problem. The prevalence of breech presentation decreases from about 15% at 29-32 weeks gestation to between 3-4%.
Many hospitals have a policy of elective caesarean section at 38 weeks gestation for all breech presentations [12]. For many mothers, particularly those who have made great efforts to maximize the chances of 'as natural a birth as possible', such a position is extremely upsetting - loss of control of or involvement in the delivery of her baby is often total. In fact, a mother in such a position does have three main choices although these are unlikely to be made known to her: Elective caesarean section; Vaginal breech delivery or vaginal breech extraction using forceps and Natural active breech birth.
Despite the widespread acceptance that breech babies should be delivered by caesarean section, it has not been proven to be safer for the baby than natural active breech birth. An international multi-centre Term Breech Trial is currently being undertaken to look at the question of which is the better approach for management of the breech baby at term: planned caesarean section or planned vaginal birth [11].

2 Research methodology

This research study makes use of secondary data collected from the Federal Medical Centre Benue State for the year 2016.The predictor variables used are the age, sex and location of the patients.

2.1 Logistic Regression Analysis

Regression is the study of the relationship that exists between a dependent variable, Y, and a set of independent variables, X, using a linear function.

Logistic regression is a statistical method for analyzing a dataset in which there is one or more independent variables that determine an outcome. The outcome is measured with a dichotomous variable (in which there are only two(2) possible outcomes). That is the dependent variable which is dichotomous, contains data coded as 1(true, success, yes) or 0(false, failure, no) [13].

The logistic regression analysis is said to be Multiple, when we have a single dependent variable, Y, and multiple independent variables, denoted $x_{1,}x_2 \dots x_n$. The multiple linear regression equation is mathematically represented below:

$$\text{logit}\left(\mu_{ij}\right) = B_0 + B_1X_1 + B_2X_2 + B_3X_3 \qquad (3)$$
$$i = 1,2,\dots,n$$

Where;
β_0 = the Y intercept
β_i = are the Slope Parameters. i= 1, 2, 3.
X_1 = Sex of the Child (Male/Female).
X_2 = Weight of the Child
X_3 = Age of Mother.
μ_{ij} = the response probability of Caesarean Section for the individual (i= 1,2 ,…,n) the variable Y is an indicator for Caesarean Section, that is:

$$Y_i = \begin{cases} 1, & \text{if } \textit{a woman has caesarean Section} \\ 0, & \text{otherwise} \end{cases}$$

The general multiple logistic regression models in terms of μ_{ij} is as shown:

$$\mu_{ij} = \frac{e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n+\varepsilon_i}}{1+e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n+\varepsilon_i}} \quad (4).$$

The response variable μ_{ij} is a member of the exponential family.
The probability density function of the exponential family takes the following general form:

$$f(y;\theta) = b(y)\exp[\eta(\theta)T(y) - a(\theta)] = \exp[\eta(\theta)T(y) - a(\theta) + b(y)] \quad (5)$$

The probability mass function of the binomial random variable y is given as:

$$f\left(y; n, \mu_{ij}\right) = \binom{n_i}{y_i} \mu_{ij}^{\,y_i}\left(1-\mu_{ij}\right)^{n_i - y_i} \qquad y_i = 0,1,2,\dots,n \quad (6)$$

$$= \binom{n_i}{y_i}\left(\frac{\mu_{ij}}{1-\mu_{ij}}\right)^{y_i}\left(1-\mu_{ij}\right)^{n_i}$$

$$= \exp\left[y_i \log\left(\frac{\mu_{ij}}{1-\mu_{ij}}\right) + n_i \log\left(1-\mu_{ij}\right) + \log\binom{n_i}{y_i}\right] \quad (7)$$

$$\Rightarrow g(u) = \log\left(\frac{\mu_{ij}}{1-\mu_{ij}}\right) \quad - \text{the link function} \quad [1]$$

$$\Rightarrow \log\left(\frac{\mu_{ij}}{1-\mu_{ij}}\right) = X\beta = \eta_j(x)$$

Thus, it can be seen that the binomial distribution is an exponential family distribution with:

$$\eta(\theta) = \log\left(\frac{\mu_{ij}}{1-\mu_{ij}}\right),$$

$$T(y) = y_i\,,$$

$$a(\theta) = -n_i \log\left(1-\mu_{ij}\right) \text{ and } b(y) = \binom{n_i}{y_i}$$

It can be shown that if Y_i has a distribution in the exponential family then it has mean, μ_i and variance, V_i:

$$E(Y_i) = a'(\theta)$$

$$Var(Y_i) = a''(\theta)$$

where $a'(\theta)$ and $a''(\theta)$ are the first and second derivatives of $a(\theta)$.

The Logistic regression model has the form:

$$logit(\mu_{ij}) = B_0 + B_1X_1 + \cdots + B_pX_p \quad i = 1,2 \dots n \tag{8}$$

$$\ln\left[\frac{\mu_{ij}}{1-\mu_{ij}}\right] = \log(\text{odds}) = \eta_j(\text{x})$$

B_0 is the Y intercept, B_s are the slope parameters, and the X_s are a set of predictors. The regression coefficients B_s can be interpreted along the same lines as in linear models, bearing in mind that the left-hand-side is a logit rather than a mean.[1] Thus, B_s represents the change in the logit of the probability associated with a unit change in the ith predictor holding all other predictors constant.
Hence the probability of an event occurring is expressed by the equation:

$$\mu_{ij} = \frac{e^{B_0 + B_1X_1 + \cdots + B_pX_p}}{1 + e^{B_0 + B_1X_1 + \cdots + B_pX_p}}$$

$$= \frac{e^{\sum_{k=0}^{p} X_{ip}B_p}}{1 + e^{\sum_{k=0}^{K} X_ipB_p}} \tag{9}$$

and $1 - \mu_{ij} = \frac{1}{1+e^{B_0 + B_1X_1 + \cdots + B_pX_p}}$

$$\text{a}(\theta) = n_i * -\log\left(1 - \mu_{ij}\right) = n_i \log\left(1 + e^{\eta_j(\text{x})}\right)$$

$$a'(\theta) = n_i \frac{e^{\eta_j(\text{x})}}{1 + e^{\eta_j(\text{x})}} = n_i\,\mu_{ij}$$

$$a''(\theta) = n_i \frac{e^{\eta_j(\text{x})}}{(1+e^{\eta_j(\text{x})})^2} = n_i\,\mu_{ij}(1 - \mu_{ij})$$

$$\eta_j(\text{x}) = logit\left(\mu_{ij}\right)$$

The link function can be written in terms of the mean as:

$$\eta_j(\text{x}) = \log\left[\frac{\mu_{ij}}{1-\mu_{ij}}\right]$$

$$\eta_j(\text{x}) = \log\left[\frac{\mu_i}{n_i - \mu_i}\right]$$

which can also be written as $\eta_j(\text{x}) = \log(\mu_i) - \log(n_i - \mu_i)$

Differentiating with respect to μ_i

$$\frac{d\,\eta_j(\text{x})}{d\mu_i} = \frac{1}{\mu_i} + \frac{1}{n_i - \mu_i} = \frac{n_i}{\mu_i(n_i - \mu_i)} = \frac{1}{n_i\,\mu_{ij}(1-\mu_{ij})}$$

The working dependent variable is

$$z_i = \eta_i + (y_i - \mu_i)\frac{d\,\eta_j(\text{x})}{d\mu_i} \tag{10}$$

$$= \eta_i + \left((\frac{y_i}{n_i} - \mu_{ij})\frac{1}{\mu_{ij}(1 - \mu_{ij})}\right)$$

The iterative weight is

$$w_i = \frac{1}{\left[a''(\theta)\left(\frac{d\,\eta_j(\mathrm{x})}{d\mu_i}\right)^2\right]} \tag{11}$$
$$= n_i \mu_{ij}\left(1-\mu_{ij}\right)$$

$$(y/B) = \prod_{i=1}^{N} \frac{n_i\,!}{y_i!\,(n_i\ - y_i)!} {\mu_{ij}}^{y_i} \left(1-\mu_{ij}\right)^{n_i - y_i} \tag{12}$$

Thus, the equation to be maximized can be written as:

$$\prod_{i=1}^{N} \left(\frac{\mu_{ij}}{1-\mu_{ij}}\right)^{y_i} \left(1-\mu_{ij}\right)^{n_i} \tag{13}$$

Substituting μ_{ij} into Equation 13 yields

$$\prod_{i=1}^{N} \left(\mathrm{e}^{\sum_{k=0}^{K} X_{ik}B_k}\right)^{y_i} \left(1 - \frac{\mathrm{e}^{\sum_{k=0}^{K} X_{ik}B_k}}{1+\mathrm{e}^{\sum_{k=0}^{K} X_{ik}B_k}}\right)^{n_i}$$

$$\prod_{i=1}^{N} \left(\mathrm{e}^{y_i \sum_{k=0}^{K} X_{ik}B_k}\right)\left(1+\mathrm{e}^{\sum_{k=0}^{K} X_{ik}B_k}\right)^{-n_i} \tag{14}$$

the log likelihood function:

$$l(\beta) = \sum_{i=1}^{N} y_i \left(\sum_{k=0}^{K} X_{ik}B_k\right) - n_i\ .\log\left(1+\mathrm{e}^{\sum_{k=0}^{K} X_{ik}B_k}\right) \tag{15}$$

Thus, differentiating Equation 15 with respect to B_k

$$\frac{dl(\beta)}{dB_k} = \sum_{i=1}^{N} y_i x_{ik} - n_i\ \ .\ \frac{1}{1+\mathrm{e}^{\sum_{k=0}^{K} x_{ik}B_k}} . \frac{d}{dB_k}\left(1+\mathrm{e}^{\sum_{k=0}^{K} x_{ik}B_k}\right)$$
$$\sum_{i=1}^{N} y_i x_{ik} - n_i\ \ .\ \frac{1}{1+\mathrm{e}^{\sum_{k=0}^{K} x_{ik}B_k}} .\,\mathrm{e}^{\sum_{k=0}^{K} x_{ik}B_k}\ \frac{d}{dB_k}\sum_{k=0}^{K} x_{ik}B_k$$
$$= \sum_{i=1}^{N} y_i x_{ik} - n_i\ \ .\ \frac{1}{1+\mathrm{e}^{\sum_{k=0}^{K} x_{ik}B_k}} .\,\mathrm{e}^{\sum_{k=0}^{K} x_{ik}B_k} . x_{ik}$$
$$= \sum_{i=1}^{N} y_i x_{ik} - n_i\,\mu_{ij} x_{ik} \tag{16}$$

The maximum likelihood estimates for B can be found by setting each of the K + 1 equations equal to zero ([7],[6]). However, solving a system of nonlinear equations is not easy, the solution cannot be derived algebraically as it can in the case of linear equations. The solution must be numerically estimated using an iterative process.

The estimate of β can be obtained from the Iterative Weighted Least Squares method (IWLS).An iterative method is used to solve the equations and the resulting values of β's are called the maximum likelihood estimates of those parameters.

Iterative weighted least square estimates is obtained as:

$$\beta = (X'WX)^{-1}X'WZ \tag{17}$$

2.2 Odds Ratio

Odd ratio(O R) is a measure of association between an exposure(age, sex and weight) and an outcome(caesarean delivery). The OR represents the odds that an outcome will occur given a particular exposure, compared to the odds of the outcome occurring in the absence of that exposure. Thus, from logistic regression equation:

$$\log\left(\frac{\mu_{ij}}{1-\mu_{ij}}\right) = \beta_0 + \beta_1 x_1 + B_2 X_2 + B_3 X_3$$

$$\log\left(\frac{\mu_{ij}}{1-\mu_{ij}}\right) = \text{odds}$$

Therefore,

$$\log(\text{odds}) = \log\left(\frac{\mu_{ij}}{1-\mu_{ij}}\right) = \ln(\text{odds of event})$$

$\ln(\mu_{ij})$ =presence of event

$\ln(1-\mu_{ij})$ = absence of event

2.3 Test of Significance

2.3.1 Wald's Test

To test the null hypothesis that the single parameter estimate equals 0, the Wald statistic is given by:

$$\text{Wald} = \left(\frac{\hat{\beta}}{\hat{\sigma}_{\hat{\beta}_i}}\right)^2$$

Where,

$\hat{\sigma}_{\hat{\beta}_i}$ = Estimated Standard error of the i^{th} estimated coefficient.

The Wald statistic is asymptotically distributed as chi-square with 1 degree of freedom. The estimated standard error of the i^{th} estimated coefficient, $\hat{\sigma}_{\hat{\beta}_i}$, is the square root of the i^{th} diagonal element of the estimated covariance matrix $\sum \hat{\beta}$, that is, $\widehat{\sigma_{ii}}$, ([5], [15]).

2.3.2 Likelihood Ratio Test

In general, the likelihood ratio test is used to compare the fit of two models, one of which is nested within the other. This is typically performed to determine if a simpler model can be used to adequately model the data. The test is based on a comparison of full and reduced models where both models are fitted to the data and their log-likelihoods are calculated.

$$\begin{aligned} G = \chi^2 &= D_{null} - D_k \\ &= -2LL_k - (-2LL_k) \end{aligned}$$

where D_{null} is the deviance for the constant only model and D_k is the deviance for the model containing k number of predictors. If the model deviance is significantly smaller than the null deviance then one can conclude that the predictor or set of predictors significantly improved model fit([1] [3])

3 Result

The data used for this research analysis was collected from the Federal Medical Centre Makurdi. It comprises of the type of delivery, the Sex of the child, the weight of the child, and Age of Mother. The variable Y is binary coded 1or 0 indicating whether or not a woman delivers through and Caesarean section. Sex of the child is also dichotomous coded 1 for male and 0 for female, while the weight of the child and the age of the mother are continuous variables. The analysis was computed using a Statistical Software, R-Software version 3.2.

	Estimate	Std. Error	z-value	Pr(/z/)	2.5%	97.5%	Odds ratio
(Intercept)	-2.4531	0.5989	-4.096	4.2e-05	0.0266	0.2782	0.0860
SEX	0.1366	0.1592	0.858	0.39089	0.8391	1.5659	1.1463
WEIGHT	0.3790	0.1325	2.861	0.00422	1.1268	1.8940	1.4609
AGEOFMOTHER	0.0075	0.0133	1.962	0.00739	0.9815	1.0342	1.0075

Table1: Estimated Parameters.

The fitted model is:
logit(p) = $-2.4531 + 0.1366X_1 + 0.3790X_2 + 0.0075X_3$
where;
X_1 = the sex of the Baby
X_2 = the weight of the Baby
X_3 = the age of the mother.

1	2	4	5	6	7	8	9	10
0.248	0.308	0.237	0.304	0.379	0.179	0.357	0.196	0.201
11	12	14	15	16	17	18	19	20
0.215	0.2	0.321	0.323	0.289	0.283	0.262	0.323	0.265
21	22	24	25	26	27	28	29	30
0.243	0.212	0.274	0.236	0.272	0.334	0.252	0.229	0.281
31	32	34	35	36	37	38	39	40
0.306	0.229	0.246	0.292	0.207	0.372	0.367	0.394	0.259
41	42	44	45	46	47	48	49	50
0.278	0.287	0.303	0.257	0.25	0.257	0.243	0.252	0.264
51	52	54	55	56	57	58	59	60
0.339	0.357	0.334	0.281	0.257	0.243	0.252	0.264	0.25
61	62	64	65	66	67	68	68	70
0.303	0.29	0.378	0.273	0.26	0.262	0.306	0.297	0.321
71	72	74	75	76	77	78	79	80

0.367	0.326	0.308	0.339	0.266	0.283	0.227	0.269	0.243
81	82	84	85	86	87	88	89	90
0.216	0.336	0.278	0.278	0.2	0.268	0.358	0.284	0.287
91	92	94	95	96	97	98	99	100
0.251	0.27	0.263	0.269	0.289	0.228	0.262	0.284	0.289

Table 2: Fitted Probability Values

Residual	Df	Resid. Dev	Df Dev	Pr(chi)
1	805	946.87	3	
2	802	936.94	9.921	0.01925

Table 3: Analysis of Deviance

4 Discussion

Table 1 shows the estimated parameter and Standard errors. The sex of the child has no significant ($p>0.05$) impact on the probability of caesarean section among the study population, while controlling for weight and age of mother. The weight of the child appears to have a significant ($p<0.05$) on the probability of having caesarean section while controlling for sex of the child and age of the mother. The age of the mother also have significant ($p>0.05$) on the probability of having caesarean section while controlling for sex and weight of the child.
The odd for a male child leading to caesarean section is 1.1463 times the odds for a female child leading to caesarean section. The weight of the child is continuous; hence for a unit increase in the weight of the child, the probability of having caesarean section will also increase by 1.4609. The age of the mother is also continuous thus a unit increase in the age of the mother will lead to an increase of the probability of having caesarean Section by 1.0075.
Table 2 shows the estimated probabilities. From the result on this table it can be seen that a male child with much weight (say > 3.5kg) from a mother age (say 35 and above) leads to high probabilities of having caesarean section.
Table 3 shows the analysis of deviance. To test the hypothesisH_0: $\beta_1 = \beta_2 = \beta_3 = 0$, the likelihood ratio test is used. The test statistic for the model is $\chi^2 = 9.921$ with p- value = 0.0193. Hence the fitted model adequately fit the data.

Conclusion

The findings of the data indicate there exist relationship between the probability of caesarean delivery and the covariates (mothers' age, sex and weight of the child). It shows categorically from the result that the analysis and the test of significance are useful in analyzing, comparing and predicting the probability of caesarean delivery. The study shows that the weight of the baby is significant to caesarean delivery, with sex of the baby and age of mother having positive effect to having caesarean section. We therefore recommend that pregnant mothers should avoid any thing that increases baby's weight and aged pregnant mothers should attend antenatal regularly so that they could be given the necessary medical attention.

References

[1] Agresti, A., (2002). Categorical Data Analysis. A John Wiley & Sons, Inc. 2nd edition.
[2] Agresti , A., (2007). An introduction to Categorical Data Analysis. A John Wiley & Sons.
[3] Ait-Sahalia, Y., and Kimmel, R. (2007). Maximum likelihood estimation of stochastic volatility models. *Journal of Financial Economics*. **83**: 413-452.
[4] Bedford, J., Gandhi, M., Admassu, M., and Girma, A (2012). A normal delivery takes place at home: A Qualitative study of the location of childbirth in rural Ethiopia.
[5] Berkson, J. (2003). Tests of significance considered as evidence. *International journal of epidemiology*.**32**(5): 687-691.
[6] Czepiel, S. A. (2002). Maximum likelihood estimation of logistic regression models: theory and implementation.
[7] Dacheng, X. (2010). Quasi-maximum likelihood estimation of volatility with high frequency data. *Journal of Econometrics*.**159**:235-250.
[8] Dietz, H. P., Lanzarone, V., and Simpson, J. M. (2006). Predicting operative delivery. *Ultrasound Obstet Gynecol*, **27**: 409–415.
[9] Gjonej, R., Poloska, A., Keta, M., Delija, Z., Zyberaj, F., Bezhani, V., and Smakaj, E. (2015). The reasons of rising trend of cesarean section rate year after year. A retrospective study. *An international journal of Nursing and Midwifery*. **7**(1).
[10] Goel, P., Radotra, A., Sigh, I., Aggarwal, A., and Dua, D (2004). Effects of passive smoking on outcome in pregnancy. *J Postgrad Med*. **50**(1).
[11] Lowdon, G. (1998). Breech presentation: Caesarean operation versus normal birth. AIMS journal Autumn. **10**(3).
[12] Odetola, T. D. (2014). Health care utilization among rural women of child-bearing age: a Nigerian experience. *Pan African Medical Journal*.
[13] Tomoyoshi, T. (2014). Risk of performing multiple logistic regression analysis without considering multiplicity: an overview for clinicians and practitioners. *Journal of Medical Statistics and Informatics*.**2**(7).
[14] Winslow,M., Matus, C. D., and Patterson, D. A. (2008).Labor and delivery.www.aafp.org

[15] Yanqing, Y., and Wang, X. (2011). Comparison of wald, score, and likelihood ratio tests for response adaptive designs. *Journal of Statistical Theory and Applications*. **10** (4): 553-569.