THE INTERNATIONAL JOURNAL OF HUMANITIES & SOCIAL STUDIES

Modelling the Determinants of Babies Weight Using Quantile Regression in Bolgatanga Municipality

Donatus Nyaaba
Lecturer, Department of Statistics, Bolgatanga Polytechnic Ghana
John Ayuekanbey Awaab
Assistant Lecturer, Department of Secretaryship and Management Studies,
Bolgatanga Polytechnic, Ghana
Tule Clement Kani
Teacher, Department of Mathematics, Zuarungu Senior High School, Ghana
Martin Amalitinga Akansuke
Teacher, Department of Mathematics, Zuarungu Senior High School, Ghana

Abstract
In recent years, birth outcomes have resulted in large economic cost, both direct medical cost and long-term developmental consequences. It is therefore not surprising that governments and cooperate entities focus their attention to prenatal-care improvements. Both low and high birth weights are examined to see the long-term effects on cognitive development, educational outcomes, and labor-market outcomes. However, several data including birth weight data involves the analysis of skewed data. When a distribution of variables is highly skewed, it implies the mean is sensitive to outliers and certainly not a good measure of central tendency. In this case, quantile regression is one approach that is appropriate to analyse such data. The performance of the best quantile regression model is compared with other skewed models such as the Weibull and lognormal regressions. The best regression model that fits the birth weight data is determined using the Q-Q plot and the AIC.

Keywords: Quantile Regression, Q-Q plot, baby's weight, Akaike Information Criteria (AIC), Bayesian information criterion (BIC) and Hannan-Quinn (HQ)

1. Background of the Study

Traditionally, linear regression analyses have detected increasing trends in the incidence of overweight/obesity among new born babies. However, these previous regression methods were limited in their ability to capture cross-distribution variations among effects.

Birth weight according to a USA National Library for Medicine is the first weight of a baby, taken just after he or she is born. How much a baby weights at birth can have a profound impact on the child's life? Babies who weigh less than 5.5 pounds at birth – the cut-off for the low birth weight –may face health challenges after birth including serious digestive issues, difficulty in breathing, and bleeding in the brain. Their risk of developing diabetes, heart diseases, high blood pressure and obesity later in life is also increased.

High birth weight babies are those that weigh more than 8.8 pounds at birth and can also face an increased risk of obesity, diabetes, and high blood pressure in adulthood. Genetics certainly play a role in birth weight. Parental health therefore has a lot to do with babies' weight at birth, especially when factors such as high blood pressure, diabetes, heart diseases and the likes are not well checked.

The prevalence of childhood obesity increased dramatically during the last decade in industrialized countries (Toschke et al., 2005).

TV watching, formula feeding, smoking in pregnancy, maternal obesity, parental social class are well known environmental constitutional or socio-demographic risk factors (Toschke et al., 2005).

However, it is uncertain as to whether these factors have effects on the entire weight distribution or just a part of it. Several data such as the birth weight data involves the analysis of highly skewed data. When a distribution of variables is highly skewed, it implies that the mean is sensitive to outliers and definitely not a good measure of central tendency.

Quantile regression therefore is one method to analyze such data. Quantile regression as proposed by Koenker and Bassett (1978) has immerged as an important statistical approach for addressing the limitations of simple linear regression. Quantile regression model is a natural extension of the linear regression model estimating various conditional quantile functions. This offers a strategy for determining how the covariates influence the entire response distribution.

The log-normal and the Weibull distributions by far are the most popular distributions for modeling skewed data. Suppose a research has observed n data points, say $x_1, x_2, ... x_n$ and he wants to use either two parameter log-normal model or two parameter Weibull model, which one is preferable?

It is well known that both models can be used effectively to analyze such skewed data set. Although these two models may provide similar data fit, it's desirable to select the correct or more nearly correct model.

The Log-normal distribution usually is preferred widely in situations where data points are skewed positively. That is, in instances of financial analysis, economics as well as health.

Similarly, the Weibull distribution describes data resulting from life and fatigue tests. It is a family of distribution that can assume the properties of several other distributions.

One most important approach in applied statistics in measuring a relationship between a response variable and covariates is through regression. This model can be written as:

$Y = \beta_0 + \beta_1 + ... + \beta_p X_p + E$, where $E \sim N(0, \sigma^2)$.

Ordinary linear regression estimates the average relationship between a set of regressors and the dependent variable based on the conditional mean, $E(Y/X)$ (Koenker and Basset, 1978). In quantile regression we use the conditional median $Qq(Y/X)$ to estimate the relationship between dependent and independent variables, the quantile $'q'E(1,0)$, where the median is the 50th quantile (Chen, et al, 2009). There is therefore an increasingly extension of quantile regression literature in economics. Buchinsky (1998), Koenker and Hallock (2001) and Koenker (2005) surveyed this literature in recent times.

1.1. Problem Statement

One area in recent times that have resulted in huge economic costs is adverse birth outcomes. These costs come in the form of direct medical cost and long-term developmental consequences, which necessitated the public-health community to focus efforts on prenatal-care improvements that are thought to improve birth outcomes.

A leading indicator of infant health is believed to be birth weight; classifying those weighing less than 2500grams (less than 5.5 pounds) at birth as having low birth weight and those weighing more than 4000grams and 4500grams (more than 8.8 pounds) as having high birth weight and very high birth weight respectively. Most studies estimated the effects of birth inputs on the fraction of births below thresholds.

Researchers over the years have tried to compare quantile regression to ordinary linear and multiple regressions. The mean in such regressions does not appear as a good measure of central tendency especially when such distributions are skewed positively, as in birth weight data. Thus, ordinary linear regression only estimates the average relationship between a set of regressors and the dependent variable based on the conditional mean (Koenker and Basette, 1978) which appears not good enough for the distributions.

Quantile regression comes in therefore as a good approach for such distributions as it seeks to determine a relationship between a response variable and covariates at every quantile. Since any quantile can be used, it makes it possible to model any predetermined location of the distribution. These adverse birth outcomes, especially in the Bolgatanga municipality is therefore what this study seeks to achieve for the birth weight data using quantile regression.

1.2. Research Questions

Following the researcher's familiarization with the research topic and study area, the following questions were asked, the answers to which is hoped would help explain the determinants of birth weights of babies using quantile regression.

- What factors are associated with birth outcomes in the area?
- What quantile appropriately describes the skewed data?
- Why do babies with low/high birth weight need special care?
- What is the best quantile regression model when compared to other skewed models?

1.4. Objectives of the Study

1.4.1. General Objective

The general objective of the study is to investigate the determinants of birth weight of babies in the Bolgatanga municipality.

1.4.2. Specific Objectives

- To identify the appropriate quantile that best describes the birth weight of the babies.
- To explore how prognostic factors, influence the determinants of birth weight of babies.
- To compare the performance of the best quantile regression model with other skewed models such as the Weibull and Lognormal regressions.
- To predict the probability of a baby falling into different categories of birth weight using multinomial logistic regression.

1.5. Significance of the Study

The purpose of this study is to obtain appropriate quantiles that best describes the distribution of positively skewed data. In light of this, researchers can make pragmatic decisions in relation to positively skewed data. The research would therefore be crucial to the health sector especially the increasing trends in the incidence of overweight/obesity among new born babies as well as academia. Researchers in academia would use the study as literature in other related areas.

2. Literature Review

2.1. Introduction

This section reviews the work of general authors relating to the research as well as concept definitions. Research works, empirical work and authors' opinion on Quantiles, Weibull and Lognormal Regressions are looked at.

2.2. Quantile Regression

2.2.1. Historical Background and Related Works

Koenker (1978) introduced quantile regression which models, conditional quantiles as functions of predictors. The quantile regression model is a natural extension of the linear regression model. While the linear regression model specifies the change in the conditional mean of the dependent variable associated with a change in the covariates (Wooldridge, 2013) the quantile regression model specifies changes in the conditional. Since any quantile can be used, it is possible to model any predetermined location of the distribution. Ordinary linear regression estimates the average relationship between a set of regressors and the dependent variable based on the conditional mean $E(Y/X)$ (Koenker and Bassett 1978).

Observable measures of poor prenatal care, such as smoking, have strong negative associations with birth weight. For instance, according to a report by the Surgeon General, mothers who smoke during pregnancy have babies that, on average, weigh 250 grams less (Centers for Disease Control and Prevention (2001)).

The direct medical costs of low birth weight are quite high. Based upon hospital-discharge data from New York and New Jersey, Almond et al., (2005) report that the hospital costs for newborns peaks at around $150,000 (in 2000 dollars) for infants that weigh 800 grams; the costs remain quite high for all "low birth weight" outcomes, with an average cost of around $15,000 for infants that weigh 2000 grams. The infant-mortality rate also increases at lower birth weights.

Other research has examined the long-term effects of low birth weight on cognitive development, educational outcomes, and labor-market outcomes. LBW babies have developmental problems in cognition, attention, and neuromata functioning that persist until adolescence (Hack et al., (1995)). LBW babies are more likely to delay entry into kindergarten, repeat a grade in school, and attend special-education classes Corman (1995); Corman and Chaikind (1998)). LBW babies are also more likely to have inferior labor-market outcomes, being more likely to be unemployed and earn lower wages (Behrman and Rosenzweig (2004); Case et al., (2005); Currie and Hyson (1999)).

Although it has received less attention in the economics literature, high-birth weight outcomes can also represent adverse outcomes. For instance, babies weighing more than $4,000$ grams (classified as high birth weight (HBW)) and especially those weighing more than $4,500$ grams (classified as very high birth weight (VHBW)) are more likely to require cesarean-section births, have higher infant mortality rates, and develop health problems later in life.

A difficulty in evaluating initiatives aimed at improving birth outcomes is to accurately estimate the causal effects of prenatal activities on these birth outcomes. Unobserved heterogeneity among child bearing women makes it difficult to isolate causal effects of various determinants of birth outcomes. Whether or not a mother smokes, for instance, is likely to be correlated with unobserved characteristics of the mother.
Another approach has been to utilize panel data (i.e., several births for each mother) to identify these effects from changes in prenatal behavior or maternal characteristics between pregnancies (Abrevaya (2006); Currie and Moretti (2002); Rosenzweig and Wolpin (1991); Royer (2004)). One concern with the panel-data identification strategy is the presence of "feedback effects," specifically that prenatal care and smoking in later pregnancies may be correlated with birth outcomes in earlier pregnancies. Royer (2004) provides an explicit estimation strategy to deal with such feedback effects (using data on at least three births per mother). Abrevaya (2006) shows that feedback effects are likely to cause the estimated (negative) smoking effect to be too large in magnitude.

Since the costs associated with birth weight have been found to exist primarily at the low end of the birth weight distribution (with costs increasing significantly at the very low end), most studies have estimated the effects of birth inputs on the fraction of births below various thresholds (e.g., $2,500$ grams for LBW and $1,500$ grams for "very low birth weight"). As an alternative, their paper considers a quantile-regression approach to estimating the effects of birth inputs on birth weight, so it is useful to compare the two approaches. The threshold-crossing approach fixes a common unconditional threshold for the entire sample, whereas the quantile-regression approach focuses upon particular

conditional quantiles of the birth weight distribution. Denoting birth weight by bw and a birth input vector by x, a probit-based threshold-crossing model for LBW outcomes would be $\Pr(bw < 2500/x) = \phi(x'\gamma)$. For each x, there is a conditional probability of the LBW outcome (bw below the common threshold) and estimates of γ can be used to infer the marginal effects of the birth inputs upon these conditional probabilities. For the quantile approach, a simple (linear) model for, say, the 5% conditional quantile would be $Q_{5\%}(bw/x) = x'\beta$. The value of the conditional quantile $Q_{5\%}(bw/x)$ may be below the LBW threshold of $2,500$ grams for some x values and above it for other x values. The estimated marginal effects (inferred from the estimates of β) would indicate how the 5% conditional quantile would be affected at all x values. These effects are not directly comparable to the probit-based effects.

A recent literature on estimation of quantile treatment effects, including Abadie, Angrist, and Imbens (2002) and Bitler, Gelbach, and Hoynes (2006), has argued that traditional estimation of average (mean) treatment effects may miss important causal impacts. Specifically, an average treatment effect inherently combines the magnitudes of causal effects upon different parts of the conditional distribution. It is quite possible, as in our birth weight application (and also in wage-distribution applications), that societal costs and benefits are more pronounced at the lower quantiles of the conditional distribution. As an example, if one estimated the average causal effect of smoking to be a reduction in birth weight of 150 grams, it could be the case that the effect of smoking on lower quantiles is substantially higher or lower than 150 grams. If a $200-\text{gram}$ effect were estimated at lower quantiles and a $100-\text{gram}$ effect at higher quantiles, this would argue for a stronger policy response than if the effects were instead stronger at the higher quantiles. Ultimately, consideration of how effects vary over the quantiles is an empirical question and one which they attempt to answer in the context of birth weight regressions in their paper.

Previous quantile-estimation approaches to estimating birth-outcome regressions have used cross-sectional data and, therefore, have suffered from an inability to control for unobserved heterogeneity. For instance, Abrevaya (2001) (see also Koenker and Hallock (2001) and Cher- nozhukov (2005)) uses cross-sectional federal natality data and finds that various observables have significantly stronger associations with birth weight at lower quantiles of the birth weight distribution; unfortunately, one cannot interpret these "effects" as causal since the estimation has a purely reduced-form structure that does not account for unobserved heterogeneity.

2.2.2. Historical Background and Review of Related Works on Lognormal Regression

The lognormal distribution finds its beginning in 1879. It was at this time that F. Galton noticed that if $X_1, X_2, ..., X_n$ are independently positive random variables such that:

$$T_n = \prod_{i-1}^{n} xi .$$

Then the log of their product is equivalent to the sum of their logs,

$$Ln(T_n) = \sum_{i=1}^{n} xi .$$

Due to this fact, Galton concluded that the standardized distribution of ln (Tn) would tend to a unit normal distribution as n goes to infinity.

After Galton, the roots of the lognormal distribution remained virtually untouched until 1903, when Kapteyn derived the lognormal distribution as a special case of the transformed normal distribution.

The lognormal is sometimes called the anti-lognormal distribution, because it is not the distribution of the logarithm of a normal variable, but instead the anti-log of a normal variable (Brezina 1963; John and Kotz, 1970).

An important characteristic of the lognormal distribution is its multiplicative property. It states that if two independent random variables, X_1 and X_2, are distributed respectively as lognormal (μ_1, σ^2_1) and lognormal (μ_2, σ^2_2) then the product of X_1 and X_2 is distributed as;

Lognormal $\left(\mu_1\mu_2, \sqrt{\sigma^2_1 + \sigma^2_2}\right)$. This multiplicative property for independent lognormal random variable stems from the additive properties of normal random variables (Ante 1985).

Another important characteristic of the lognormal distribution is the fact that for very small values of σ (e.g., less than 0.3), the lognormal is nearly indistinguishable from the normal distribution (Ante 1985). However, unlike the normal distribution, the lognormal does not possess a moment generating function. Instead, its moments are given by the following equation defined by Casella Berger (2002).

The lognormal distribution is useful in modeling continuous random variables which are greater than or equal to zero. This distribution is also useful in modeling data which would be considered normally distributed except

for the fact that it may be more or less skewed. Such skewness occurs frequently when means are low, variances are large, and values cannot be negative (Limpert, Stahet, and Abbt 2001).

2.2.3. Historical Background and Review of Related Works on Weibull Regression
There are many applications for the Weibull distribution in statistics. Although it was first identified by Fréchet in 1927, it is named after Waloddi Weibull and is a cousin to both the Fréchet and Gumbel distributions. Waloddi Weibull was the first to promote the usefulness of this distribution by modeling data sets from various disciplines (Murthy, Xie, and Jiang 2004).Extreme value theory is a unique statistical discipline that develops "models for describing the unusual rather than the usual (Coles 2001)." Perhaps the simplest example of an extreme value distribution is the exponential distribution. The Weibull distribution is specifically used to model extreme value data. One example of this is the frequent use of the Weibull distribution to model failure time data (Murthy et al. 2004). Its use is also applicable in various situations including data containing product failure times to data containing survival times of cancer patients. The extreme values in these analyses would be the unusual longevity of the shelf life of a product or survival of affected patients. One reason it is widely used in reliability and life data analysis is due to its flexibility. It can mimic various distributions like the normal. Special cases of the Weibull include the exponential ($\gamma = 1$) and Rayleigh ($\gamma = 2$) distributions. Because of its uses in lifetime analysis, a more useful function is the probability that the lifetime exceeds any given time, (i.e. $P(T > t)$). This is called the survival function or in the case of a product, the reliability.

2.2.4. Discriminating Between Weibull and Lognormal Regressions
Suppose an experimenter has observed n data points, say $x_1, ..., x_n$ and he wants to use either two-parameter log-normal model or two parameter Weibull model, which one is preferable?

It is well known that both the log-normal and Weibull models can be used quite effectively to analyze skewed data set. Although, these two models may provide similar data fit for moderate sample sizes, but still it is desirable to select the correct or more nearly correct model, since the inferences based on the model will often involve tail probabilities, where the effect of the model assumptions are very critical. Therefore, even if we have small or moderate samples, it is still very important to make the best possible decision based on whatever data are available.

The problem of testing whether some given observations follow one of the two probability distributions is quite old in the statistical literature. See for example the work of Atkinson (1969, 1970), Bain and Englehardt (1980), Chambers and Cox (1967), Chen (1980), Cox (1961, 1962), Dyer (1973), Fearn and Nebenzahl (1991), Gupta and Kundu (2003, 2004), Kundu, Gupta and Manglick (2004), Wiens (1999) and the references therein. Considering the problem of discriminating between Weibull and log-normal distributions, the ratio of the maximized likelihood (RML) is used in discriminating between the two distribution functions and using the approach of White (1982a, b), the asymptotic distribution of the logarithm of RML is obtained. It is observed that the asymptotic distribution is asymptotically normal and it is independent the asymptotic distribution can be used to compute the probability of correct selection (PCS) and it is observed in the simulation study that the asymptotic distribution works quite well even for small sample sizes. Dumonceaux and Antle (1973) proposed a likelihood ratio of the unknown parameters. test in discriminating between the log-normal and Weibull distributions. The asymptotic results can be used to obtain the critical regions of the corresponding testing of hypotheses problem also.

The minimum sample size required to discriminate between the two distribution functions for a given PCS was also determined. Using the asymptotic distribution of the logarithm of RML, it was obtained, the minimum sample size required to discriminate between the two distribution functions for a given user specified protection level, i.e. the PCS.

3. Research Methods

3.1. Introduction
This section examines the concepts of quantile regression, Weibull regression and the lognormal regression as well as elaboration of statistical methods and the tools used in the data analysis.

3.2. Data
Secondary data is obtained from the Bolgatanga Regional Hospital on birth weight of babies for 2017. The data contains the weight and gender of the babies as well as other maternal characteristics. Some of these maternal characteristics include marital status of the mother, educational level of the mother, occupation of the mother, age of mother at birth of child and some other determinants. In obtaining the statistical model for the data, the weight of the baby is used as the main response variable, whilst all other determinants and characteristics are the covariates. Bolgatanga Regional Hospital is located in the Bolgatanga municipality, which lies within the Guinea Savannah woodlands and provides health services to the populace and its environs. The municipality is one of the fifteen (15) districts/municipalities in the Upper East Region. The municipality has a total land area of about 1,674 sq.km and shares boundaries to the North with Burkina Faso, to the East with Bolgatanga and Bongo, West, the Builsa District and Sissala District and to the South with West Manprusi district. The settlement population of Bolgatanga is estimated to be 27,306 (population and housing census, 2010).

3.3. Basic Concepts and Model Description

3.3.1. Quantile Regression

We say that a student scores at the τth quantile of a standardized exam, if he performs better than the proportion τ, and worse than the proportion $(1-\tau)$, of the reference group of students. Thus, half of the students perform better than the median student, and half perform worse. Similarly, the quantiles divide the population into four segments with equal proportions of the population in each segment; the deciles into 10 equals. The quantile or percentile refers to the general case. More formally, any real valued random variable, Y, maybe characterized by the distribution function,

$$F(y) = prob(Y \leq y), \quad \text{while for any } 0 \angle \tau \angle 1,$$

$$Q(\tau) = \text{Inf} \left\{ y : F(y) \geq \tau \right\}.$$

Is called the τth quantile of Y. The median, Q (1/2), plays the central role. Like the distribution function, the quantile function provides a complete characterization of the random variable, Y. for any $0 < \tau < 1$, define the piecewise linear "check function", illustrated in figure 1.
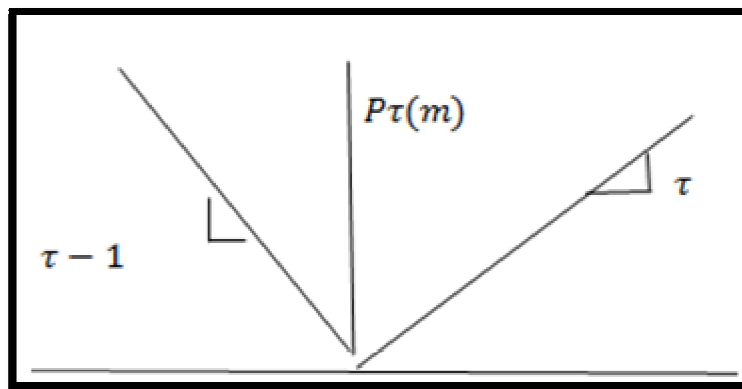

Figure 1: Quantile Regression P Function

Then, the function f (b) =E [p τ(X-b)] is minimized when b is the τ th quantile of F. To see that this is indeed true, observe that:

$$E\left[p(\tau)(X - b) \right] = (\tau - 1)\int (\tau - b)df(x) + \tau \int (x - b)df(x).$$

Then, differentiating equation above with respect to b, we find that;

$$(\tau - 1)\int df(x) - \tau \int (df)(x) = F(b) - \tau = 0$$

Since F is strictly increasing, the minimizer b is unique. Koenker and Basset used the above result as motivation to consider alternatives to the sample mean, which is the least squares estimator for the linear regression model. We can therefore generalize the aforementioned minimization problem to the regression context whereby the conditional quantile function is modeled as a linear function of covariates. To do so, we use some basic results from least squares theory, and naturally and intuitively extend this to the quantile case.

We define $y = \left[y1, ... yn \right]$ as a vector of response, and $\left[x1 .... xn \right]$ as an $n \times p$ covariate matrix. A maximum likelihood Type Estimate, also called an M-Estimate of location, for some parameter μ is defined by a minimization problem of the form:

$$\min_{\mu \in R} \sum_{i=1}^{n} \Psi(y_i - \mu) \qquad (3.1)$$

A least squares estimate is a typical example of an M-estimate.

By setting $\psi = (y_1 - \mu)^2$, expression (3.1) becomes

$$\min_{\mu \in R} \sum_{i=1}^{n} (y_i - \mu)^2 \qquad (3.2)$$

The solution of which is the sample mean;

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^{n} y$$

We can readily extend the theory of M-estimation to the regression context. By expressing the conditional mean of a random variable Y as;

$E[Y/X = x_i] = x^\tau_i \beta,$ Where x^T_i represents the ith row of the matrix X, the estimate of β is given as: $\hat{\beta}$

$$= \arg\min_{\beta \in R^p} \sum_{i=1}^{n} \Psi(y_i - x^\tau_i \beta) \qquad (3.3)$$

$$= \arg\min_{\beta \in R^p} \sum (y_i x^\tau_i \beta)^2 \qquad (3.4)$$

Furthermore, we have already shown that the τth sample quantile, say $\hat{\alpha}(\tau)$, is defined by a minimization problem of the form:

$$\min_{\alpha \in R} \sum_{i=1}^{n} p\tau(y_i - \alpha). \qquad (3.5)$$

As in the case of least squares theory expression (3.5) is just another example of M-estimator with $\Psi = P\tau(\cdot)$ If we specify the conditional quantile function for a random variable Y as
$Qy(T|X = x_i) = x^T_i \beta(T),$
We can estimate the parameters $\beta(T)$ by solving:

$$\min_{\beta \in R} \sum_{i=1}^{n} p\tau(y_i - x^T_i \beta(T)) \qquad (3.6)$$

3.3.2. Essential Properties of Quantile Regression

The practical usefulness of any estimation is determined, besides other factors, by its invariance and robustness properties, because they are essential for coherent interpretation of regression results. Although some of these properties are often perceived as granted, thus probably because of their validity in the case of the least square's regression, it does not have to be the case for more evolved regression procedures. Fortunately, quantile regression preserves many of these invariance properties, and even adds to them several other distinctive quartiles.

3.3.3. Equivariance

In many situations it is preferable to adjust the scale of original variables or reparametrize a model so that its result has a more natural interpretation. Such changes should not affect our quantitative and qualitative conclusions based on the regression output. Invariance to a set of some elementary transformations of the model is called equivariance in this context. Koenker and Bassett (1978) formulated four equivariance properties of quantile regression. Once we denote the quantile regression estimate for a given Tϵ (0, 1) and observations (y, X) byβ (T; y, X), then for ay p×p nonsingular matrix A, yR and a>0 holds

• $[\beta b(T; ay, X) = a\beta b(T; y, X)]$

• $[\beta b(T; -ay, X) = a\beta \beta b(1 - T; y, X)]$

• $\beta b(T; y + X_y X) = \beta \beta b(T; y, X) + y$

This means, for example, that if we use the measurement unit of y millimeters by 1000, then our estimate scales appropriately:

$\beta b(T; y[mm], X) = 1000.\beta b(T; y[m], X).$

3.3.4. Invariance to Monotonic Transformations

Quantile exhibit besides usual equivariance properties also equivariance to monotone transformations. Let f (.) be a non-decreasing function on R, then it immediately follows from the definition of the quantile function that for any random variable Y:

$Qf(Y)(T) = fQY(T).$

In other words, the quantile of the transformed random variable f(Y) are the transformed quantiles of the original Y. But, this is not the case of the conditional expectation
E {f(Y)} ≠ f (EY) unless f (.) is a linear function. This is why a careful choice of the transformation of the dependent variable is so important in various econometrics models when the ordinary least squares method is applied.
We can illustrate the strength of equivariance with respect to monotone transformation on the so-called censoring models. If we assume therefore that there exists, a simple linear regression model with i.i.d errors, and that the response variable yi is unobservable for some reason. Because of censoring, the standard least squares method is not consistent anymore but, a proper formulated maximum likelihood estimator can be applied. On the contrary, the quantile regression estimator, thanks to the equivariance to monotone transformations, does not run into problems of such.

3.4. Robustness

Sensitivity of an estimator to departures from its distributional assumptions is another important issue. The discussion concerning relative qualities of the mean and median is an example of how significant this kind of robustness or sensitivity can be. The sample mean, being a superior estimate of the expectation under the normality of the error distribution, can be adversely affected even by a single observation if it sufficiently far from the rest of data points. Likewise, the effect of such a distant observation on the sample median is bounded no matter how far the outlying observation is. This robustness of the median is, of course, outweighed by lower efficiency in some cases. Other quantile enjoys similar properties thus, the effect of outlying observations on the T-th sample quantile is bounded, in that the number of outliers is lower than

nmin{T, 1-T}. Quantile regression inherits these robustness properties since the minimized objective functions in the case of sample quantiles and in the case of quantile regression are the same. The only difference is that regressions residuals are used instead of deviations from mean

yi –u. Hence, quantile regression estimates are reliable in the presence of outlying observations that have large residuals.

3.5. Linear Programming of Quantile Regression

This is the computational aspect of quantile regression. Quantile regression estimates are found through numerical methods. More specifically estimates can be found through the techniques of linear programming.

Linear programming is a mathematical technique where the aim is to find a vector $x \in R_+^n$ that minimizes (or maximizes) the value of a given linear function subject to contain constraints. In the matrix notation, a linear program with n unknows and m constraints can be written in the following form:

$Minimize C^T x$

$Subject\ to\ Ax \le b$

$And\ x \ge 0,$

Where x is an $n \times 1$ vector of variables to be determined, c and d are known vectors of dimension $n \times 1$ and $m \times 1$ respectively, and A is a known $m \times n$ matrix. While an in-depth review of linear programming method is outside the scope of the dissertation, there are various algorithms which can be used to solve the linear program, such as simplex or interior point methods.

Following Davimo et al. [8], page 27, we consider a simple two parameter medium regression formulation:

$m(Y|x) = \beta_0 + \beta_1 x,$

Where $m(\cdot)$ denotes the median. So, given any sample $\{y_1, \ldots\ldots, y_n\}$ observations, the parameters β_0 and β_1 can be estimated as the solution to the following minimization problem:

$$\min_{\beta_0 \beta_1} \sum_{i=1}^{n} |\beta_0 + \beta_1 x_1 - y_i|$$

The resulting linear program is:

$Minimize \sum_{i=0}^{n} e_i$

$Subject\ to\ e_i \ge \beta_0 + \beta_1 x_1 - y_i \qquad i=1,\ldots\ldots,n$

$and\ e_i \ge -(\beta_0 + \beta_1 x_i - y_i) \qquad i=1,\ldots\ldots,n$

Where each e_i is an auxiliary variable that represents the error term at the ith point.

If we now consider the P parameter quantile regression modal given by:

$y = X\beta(\tau) + \epsilon,$

where y is an $n \times 1$ vector of response, x an $n \times P$ matrix of covariates, $\beta(\tau)$ is a $P \times 1$ vector of unknown parameter at a given quantile τ, and ϵ is an $n \times 1$ vector of unknown errors, the estimate $\hat{\beta}(\tau)$ can be found as the solution to the minimization problem:

$$\min_{\beta(\tau) \in R^P} \sum_{i=1}^{n} p\tau \left(y_i x_i^T \beta(T) \right).$$

Weibull Regression model

If X~Weibull $(\propto, \beta)$ then its density function is given by:

$$f\left(t_i \mid \alpha, \beta\right) = \frac{\alpha}{\beta^{\alpha}} t_i^{(\alpha-1)} \exp\left\{-\left(\frac{t_i}{\beta}\right)^{\alpha}\right\}, \; for \; \alpha \succ 0, \; and \; \beta \succ 0 \tag{3.7}$$

Location (L), shape (α), and scale (β) are the distributional parameters.

Input requirement;

$\beta \succ 0$ and can be any positive value

$\alpha \geq 0.05$

Location can take on any value

3.6. Parameters Estimation of Weibull Regression

3.6.1. Maximum Likelihood Estimation

This is computed by first assuming that T_i iid Weibull$(\propto, \beta)$, with probability given by its density function;

$f(t_i \mid \propto, \beta) = \frac{\propto}{\beta^{\propto}} t_i^{(\propto-1)} exp\left\{-\left(\frac{t_i}{\beta}\right)^{\propto}\right\}$, for $\propto > 0$, and $\beta > 0$

The joint density of the likelihood is the product of the density of each data point.

$$L(\propto, \beta \mid t) = \prod_{i=1}^{n} f(t_i \mid \propto, \beta) \tag{3.8}$$

$$L(\propto, \beta \mid t) = \prod_{i=1}^{n} \left[\frac{\propto}{\beta^{\propto}} t_i^{(\propto-1)} exp\left\{-\left(\frac{t_i}{\beta}\right)^{\propto}\right\}\right] \tag{3.9}$$

$$L(\propto, \beta \mid t) = \left(\frac{\propto}{\beta^{\propto}}\right)^n \prod_{i=1}^{n} t_i^{(\propto-1)} \times exp\left\{-\left(\frac{\Sigma t_i}{\beta}\right)^{\propto}\right\} \tag{3.10}$$

Taking natural log

$$L(\propto, \beta \mid t) = log\left[\left(\frac{\propto}{\beta^{\propto}}\right)^n \prod_{i=1}^{n} t_i^{(\propto-1)} \times exp\left\{-\left(\frac{\Sigma t_i}{\beta}\right)^{\propto}\right\}\right] \tag{3.11}$$

$$L(\propto, \beta \mid t) = log\left(\frac{\propto}{a^{\propto}}\right)^n + log\left[\prod_{i=1}^{n} t_i^{(\propto-1)}\right] - \left(\frac{\Sigma t_i}{\beta}\right)^2 \tag{3.12}$$

$$L(\propto, \beta \mid t) = nlog\left(\frac{\propto}{\beta^{\propto}}\right)^n + (\propto-1) log\left[\prod_{i=1}^{n} t_i\right] - \sum_{i-1}^{n}\left(\frac{t_i}{\beta}\right)^{\propto} \tag{3.13}$$

$$L(\propto, \beta \mid t) = n \; log \propto -n \; log\beta^{\propto} + (\propto-1)\sum_{i=1}^{n} log(t_i) - \sum_{i-1}^{n}\left(\frac{t_i}{\beta}\right)^{\propto} \tag{3.14}$$

$$L(\propto, \beta \mid t) = n \; log(\propto) - n \; log(\beta) + (\propto-1)\sum_{i=1}^{n} log(t_i) - \sum_{i-1}^{n}\left(\frac{t_i}{\beta}\right)^{\propto} \tag{3.15}$$

3.6.2. Method of Moments Estimator

The MME is defined by computing the sample moments, $U^k = \frac{1}{n}\sum_{i=1}^{n} t_i^{\,k}$,

And setting them equal to the theorical moments from the moment generating function, $M_T(t)$. The moment generating function for the Weibull is as follows; $M_T(t) = \beta^k \, r\left(1 + \frac{K}{\propto}\right)$, where K represents the kth theoretical moment, and $r \, (\cdot)$ represents the gamma function.

Only the first two moments are needed to derive the parameter estimates for the Weibull. Using the equation:

$$\frac{\sigma^2}{\mu_1^2} = \frac{M_2(t) - \left[M_1(t)\right]^2}{\left[M_1(t)\right]^2},$$

From Murthy, Xie and Jrang (2004), we see that it can be simplified to the following equations;

$$\frac{\sigma^2}{\mu_1^2} = \frac{M_2(t) - \left[M_1(t)\right]^2}{\left[M_1(t)\right]^2}, \tag{3.16}$$

$$\frac{\mu_2 - \mu_1^2}{\mu_1^2} = \frac{M_2(t)}{\left[M_1(t)\right]^2} - 1 \tag{3.17}$$

$$\frac{\mu_2}{\mu_1^2} - 1 = \frac{M_2(t)}{\left[M_1(t)\right]^2} - 1 \tag{3.18}$$

$$\frac{\mu_2}{\mu_1^2} = \frac{M_2(t)}{\left[M_1(t)\right]^2} \tag{3.19}$$

Next taking the second sample moment divided by the square of the first sample moment, a function of the theorical using our data is approximated.

$$\frac{\mu_2}{\mu_1^2} = \frac{M_2(t)}{\left[M_1(t)\right]^2}$$

$$\frac{\frac{1}{n}\sum_{i=1}^{n} t_i^2}{\left[\sum_{i=1}^{n} t_i\right]^2} = \frac{\beta^2 r\left(1 + \frac{2}{\alpha}\right)}{\beta^2 r^2\left(1 + \frac{1}{\alpha}\right)} \qquad (3.20)$$

$$\frac{n\sum_{i=1}^{n} t_i^2}{\left[\sum_{i=1}^{n} t_i\right]^2} = \frac{r\left(1 + \frac{2}{\alpha}\right)}{r^2\left(1 + \frac{1}{\alpha}\right)} \qquad (3.21)$$

3.6.3. The Lognormal Regression

Suppose a random variable X is lognormally distributed, then Y=In(X) has a normal distribution. Likewise, if Y has a normal distribution, then the exponential function of Y, X= exp(Y), has a lognormal distribution. A random variable which is lognormally distributed takes only positive real values.

The mathematical constructs for the lognormal distribution are as follows:

$$f\left(X \mid \mu, \sigma^2\right) = \frac{1}{\sqrt{2\pi\sigma^2}X}\exp\left[-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right] \quad \text{for } x \succ 0;\ \mu \succ 0 \text{ and } \sigma \succ 0$$

$$\text{Mean} = \exp\left(\mu + \frac{\sigma^2}{2}\right)$$

$$S.D = \sqrt{\exp\left(\sigma^2 + 2\mu\right)\left[\exp\left(\sigma^2\right)-1\right]}$$

$$Skewness = \left[\mid\sqrt{\exp\left(\sigma^2\right)-1}\right]\left(2 + \exp\left(\sigma^2\right)\right).$$

3.6.4. Properties of the Lognormal Distribution
- The mean (u) and standard deviation (σ) are the distributional parameters.
- Input requirements: the mean and standard deviation are both > 0 and can be any positive value
- Geometric moments:

The geometric mean of the lognormal distribution is:

GM[X] = e^μ and the geometric standard deviation is:

GSD[x] = e^σ

By analogy with the arithmetic statistics, one can define a geometric variance;

$G\,Var\,[x] = e^{\sigma^2}$ and a geometric coefficient of variation; $\qquad\qquad\qquad G\,C\,V\,[X] = e^6 - 1$

Because the log-transformed variable $Y = \ln x$ is symmetric and quantiles are preserved under monotonic transformations, the geometric mean of a lognormal distribution is equal to its median, Med(X).

3.6.5. Mode and Median

The mode is the point of global maximum of the probability density function. In particular it solves the equation;

$(\ln f)^1 = 0$

Mode [x] = $e^{\mu - \sigma^2}$

The Median is such a point where

$Fx = 0.5$:

Med[x] = e^μ

3.7. Arithmetic coefficient of variation

The arithmetic coefficient of variation $CV[x]$ is the ratio:

$\frac{SD[X]}{E[X]}$, on the natural scale.

For a lognormal distribution, it is equal to:

$CV[X] = \sqrt{e^{\sigma^2} - 1}$

Contrary to the arithmetic $s.d$, the arithmetic coefficient of variation is independent of the arithmetic mean.

3.8. Arithmetic Moments

For any real or complex number n, the nth moment of a lognormally distribution variable X is given by:

$$E(X^n) = e^{n\mu + \frac{1}{2}n^2 \sigma^2}$$

Specifically, the arithmetic mean, expected square, arithmetic variance and arithmetic standard deviation of a lognormal distribution variable x are given by:

$$E(X) = e^{\mu + \frac{1}{2}\sigma^2}$$
$$E(X^2) = e^{2\mu + 2\sigma^2}$$
$$Var[x] = E[x^2] - [Ex]^2$$
$$SD[x] = \sqrt{Var[x]} \qquad \text{Respectively.}$$

The parameters μ and σ can be obtained if the arithmetic mean and the arithmetic variance are known:

$$\mu = ln\left(\frac{E[X^2]}{\sqrt{E[X]^2}}\right)$$
$$\sigma^2 = ln\left(\frac{E[X]^2}{E[X^2]}\right)$$

3.9. Parameters Estimation of Lognormal Regression

The most frequent methods of parameter estimation for the lognormal are; Maximum likelihood and method of moments.

3.10. Maximum Likelihood Estimation

This is a popular estimation technique for many distributions because it picks the values of the distribution parameters that make the data "more likely" than any other values the parameters would make them. This is accomplished by maximizing the likelihood function of the parameters given the data. Some appealing features of maximum likelihood estimators include that they are asymptotically unbiased bias tends to zero as the sample size n increases; they are asymptotically efficient, in that they achieve the Cramer-Rao lower bound as n approaches ∞; and they cut asymptotically normal.   The ML function of the lognormal distribution for a series of X is $(i = 1,2,3,\ldots\ldots.n)$ Is devised by taking the product of the prob. Densities of the individual Xis:

$$L\left(\mu, \sigma^2 \mid x\right) = \prod_{i=1}^{n}\left[f\left(x_i \mid \mu, \sigma^2\right)\right]$$

$$L\left(\mu, \sigma^2 \mid x\right) = \prod_{i=}^{n}\left(2\pi\sigma^2\right)^{-\frac{1}{2}} x_i^{-1} \exp\left[\frac{-\left(\ln\left(x_i\right) - \mu\right)^2}{2\sigma^2}\right] \qquad (3.22)$$

$$L\left(\mu, \sigma^2 \mid x\right) = \left(2\pi\sigma^2\right)^{\frac{-1}{2}} x_i^{-1} \exp\left[\sum_{i=1}^{n} \frac{-\left(\ln\left(x_i\right) - \mu\right)^2}{2\sigma^2}\right] \qquad (3.23)$$

Taking natural log of the likelihood function;

We now find $\hat{\mu}$ and $\hat{\sigma}^2$, to maximize the $L(\mu, \hat{\sigma}^2 \mid x)$. Take partial directive L with respect to μ and $\hat{\sigma}^2$ and set it equal to 0. With resent to μ;

$$\frac{\partial L}{\partial \mu} = \frac{\sum_{i=1}^{n} ln(x_i)}{\hat{\sigma}^2} - \frac{2n\hat{\mu}}{2\hat{\sigma}^2} = 0$$
$$= \frac{n\hat{\mu}}{\hat{\sigma}^2} = \frac{\sum_{i=1}^{n} ln(x_i)}{\hat{\sigma}^2}$$
$$= n\hat{\mu} = \sum_{i=1}^{n} ln(x_i)$$
$$= \hat{\mu} = \frac{\sum_{i=1}^{n} ln(x_i)}{n}$$

With nearest to σ^2;

$$\frac{\partial^2}{\partial\sigma^2} = -\frac{n}{2}\frac{1}{\hat{\sigma}^2} - \frac{\sum_{i=1}^{n}(ln(x_i))^2}{2}(-\hat{\sigma}^2)^2$$
$$= -\frac{n}{2\hat{\sigma}^2} + \frac{\sum_{i=1}^{n}(ln(x_i) - \hat{\mu})^2}{2(\hat{\sigma}^2)^2} = 0$$
$$= \frac{n}{2\hat{\sigma}^2} = \frac{\sum_{i=1}^{n}(ln(x_i) - \hat{\mu})^2}{2\hat{\sigma}^4}$$
$$= n = \frac{\sum_{i=1}^{n}(ln(x_i) - \hat{\mu})^2}{\hat{\sigma}^2}$$

$$= \hat{\sigma}^2 = \frac{\sum_{i=1}^n (ln\ (x_i) - \hat{\mu})^2}{n}$$

$$= \hat{\sigma}^2 = \frac{\sum_{i=1}^n (ln\ (x_i) - \sum_{i=1}^n ln\ (x_i))^2}{n}$$

4. Method of Moment's Estimators

Another popular estimation technique, method of moments estimation equates sample moments with unbelievable population moments. To compute the MOM estimators $\hat{\mu}$ and $\hat{\sigma}^2$, we first need to find $E(x)$ and $E(x^2)$ for the x_2 lognormal ($\hat{\mu}$, $\hat{\sigma}^2$). We derive these using Casella and Berger's (2002) equation for the moments of the lognormal distribution.

$E(x^2) = exp[\ n\mu + n^2\ {\sigma^2}/{2}\ ]$;

$E(x) = exp[\ \mu + n\ {\sigma^2}/{2}\ ]$;

$E(x^2) = exp[\ 2\mu + 2\sigma^2]$.

So; $E(x) = e^{\mu + ({\sigma^2}/{2})}$ and $E(x^2) = e^{2(\mu + \sigma^2)}$

Now, we set $E(x)$ equal to the 1st sample moment m_1 and $E(x^2)$ equal to the 2nd sample moment m_2 where;

$$m_1 = \frac{\sum_{i=1}^n x_i}{n}$$

$$m_2 = \frac{\sum_{i=1}^n x_i^2}{n}$$

Setting $E(x) = m_1$

$$= e^{(\hat{\mu} + \hat{\sigma}^2/2} = \frac{\sum_{i=1}^n x_i}{n}$$

$$= \hat{\mu} + \frac{\hat{\sigma}^2}{2} = ln\ [\frac{\sum_{i=1}^n x_i}{n}]$$

$$= \hat{\mu} + \frac{\hat{\sigma}^2}{2} = ln\ [\frac{\sum_{i=1}^n x_i}{n}]$$

$$= \hat{\mu} + \frac{\hat{\sigma}^2}{2} = ln(\sum_{i=1}^n x_i) - ln\ (n)$$

$$= \hat{\mu} = ln\left(\sum x_i\right) - ln(n) - \frac{\hat{\sigma}^2}{2}$$

Setting $E(x^2) = m_2$

$$= e^{2(\hat{\mu} + \hat{\sigma}^2)} = \frac{\sum_{i=1}^n x_i^2}{n}$$

$$= 2\hat{\mu} + 2\hat{\sigma}^2 = ln\ [\frac{\sum_{i=1}^n x_i^2}{n}]$$

$$= 2\hat{\mu} + 2\hat{\sigma}^2 = ln(\sum_{i=1}^n x_i^2) - ln(n)$$

$$= \hat{\mu} = [ln(\sum_{i=1}^n x_i^2) - ln(n) - 2\hat{\sigma}^2] \times \frac{1}{2}$$

$$= \hat{\mu} = \frac{ln\ (\sum_{i=1}^n x_i^2)}{2} - \frac{ln(n)}{2} - \hat{\sigma}^2$$

Now we set the two $\hat{\mu}_s$ equal to each other and solve $\hat{\sigma}^2$.

$$= ln(\sum_{i=1}^n x_i) - ln(n) - \frac{\hat{\sigma}^2}{2} = \frac{ln\ \sum_{i=1}^n x_i^2}{2} - \frac{ln(n)}{2} - \hat{\sigma}^2$$

$$= 2\ ln(\sum_{i=1}^n x_i) - 2ln(n) - \hat{\sigma}^2 = ln\ \sum_{i=1}^n x_i^2 - ln(n) - 2\hat{\sigma}^2$$

$$= \hat{\sigma}^2 = ln(\sum x_i^2) - 2\ ln(\sum x_i) + ln\ (n)$$

Inserting the above value of $\hat{\sigma}^2$ into either the equation for $\hat{\mu}$ yields;

$$\hat{\mu} = ln\left(\sum_{i=1}^n x_i\right) - ln(n) - \frac{\hat{\sigma}^2}{2}$$

$$= ln(\sum_{i=1}^n x_i) - ln(n) - \frac{1}{2}[\ ln(\sum_{i=1}^n x_i^2) - 2\ ln(\sum_{i=1}^n x_i + ln(n)]$$

$$= ln\left(\sum x_i\right) - ln(n) - \frac{\sum_{i=1}^n x_i^2}{2} + ln(\sum_{i=1}^n x_i) - \frac{ln(n)}{2}$$

$$= 2\ ln\left(\sum_{i=1}^n x_i\right) - \frac{3}{2}ln(n) - \frac{ln(\sum_{i=1}^n x_i^2)}{2}$$

4.1. Choosing Best Fit Model

According to Anderson and Burnham, 2004, no model in the set of models is true; hence selection of the best approximating model is the ultimate goal. This therefore is the problem of selecting a few representative models from a large set of computational models for the purpose of decision making or optimization under uncertainty. The task can also involve the design of experiments such that the data collected is well suited to the problem of model selection. Given candidate models of similar predictive or explanatory power, the simplest model is most likely to be the choice.

However, in comparing the performance of the best quantile regression model with other skewed regression models such as the Weibull and lognormal regressions, a two parameter Weibull and lognormal distributions are obtained using the method maximum likelihood. Hence, the criteria for choosing one distribution out of the two is also based on the values of the estimated maximum likelihood estimates, the larger the likelihood, the better the model (Fiete, 2005).

Again, further sufficient test on selecting the best regression model from among these two is tested using the quantile quantile plot (Q-Q plots) and the Akaike Information criterion (AIC), using a five percent significance level (α = 0.05.).

4.2. Akaike Information Criterion (AIC)

This compares the quality of a set of statistical models to each other. The AIC will take each model and rank them from best to worse. The "best" model will be the one that neither under-fits nor over-fits. Therefore, in A.I.C, the model with the smallest value of A.I.C is selected because this model is estimated to be closest to the unknown truth among the candidate models considered.

Although the AIC will choose the best model from a set, it won't say anything about absolute quality. In other words, if all of your models are poor, then it will choose the best of a bad bunch.

Therefore, once you have selected the best model, consider running a hypothesis test to figure out the relationship between the variables in your model and the outcome of interest.

Akaike Information Criterion is usually calculated with software. The basic formula is defined as:

AIC = -2(log-likelihood) + 2K

Where:

- K is the number of model parameters (the number of variables in the model plus the intercept).
- Log-likelihood is a measure of model fit. The higher the number, the better the fit. This usually obtained from statistical output.

4.3. Quantile-Quantile Plot (Q-Q PLOTS)

A Q-Q (quantile -quantile) plot is a probability plot, which is a graphical method for comparing two probability distributions by plotting their quantiles against each other. A set of intervals for the quantiles is chosen and a point (x, y) on the plot corresponds to one of the quantiles of the second distribution (i.e., the y- coordinate) plotted against the same quantile of the first distribution (x- coordinate).

If the two distributions being plotted and compared are similar, the points in the Q-Q plot will approximately lie on the line y=x.

A Q-Q plot is used to compare the shapes of distributions, providing a graphical view of how properties such as location, scale, and skewness are similar or different in the two distributions.

With Q-Q plots, the points plotted are always non-decreasing when viewed from left to right. If the two distributions being compared are identical, the Q-Q plot follows the 45-line **y=x.**

If the two distributions agree after linearly transforming the values in one of the distributions, then the Q-Q plot follows some line, but not necessarily the line y=x. if the general trend of the Q-Q plot is flatter than the line y=x, the distribution plotted on the horizontal axis is more dispersed than the distribution plotted on the vertical axis. Conversely, if the general trend of the Q-Q plot is steeper than the line y=x, then the distribution plotted on the vertical axis is more dispersed than the distribution plotted on the horizontal axis.

4.4. Multinomial Logistic Regression Model

Multinomial logistic regression is the regression analysis to conduct when the dependent variable is nominal with more than two levels. It is a classification method that generalizes logistic regression to multiclass problems, that is, with more than two possible discrete outcomes. Multinomial regression is a model used to predict the probabilities of different possible outcomes of a categorically distributed dependent variable, given a set of independent variables. It is used when the dependent variable in question is nominal (equivalently categorical, meaning that it falls into any one of a set of categories that cannot be ordered in any meaningful way) and for which there are more than two categories.

4.5. Assumptions

The multinomial regression model assumes that data are case specific; that is, each independent variable has a single value for each case.

Again, the multinomial regression model also assumes that the dependent variable cannot be perfectly predicted from the independent variables for any case.

is the independent variables. These maternal characteristics includes: age of mother at birth of child, age of mother at first birth, educational status, gender of child etc.

5. Results and Analysis

5.1. Introduction
This section introduces the analysis of the various models and discussion of findings. It comprises the preliminary analysis, further analysis as well as the modeling sections. The Quantile regression model, the Lognormal regression model and the Gamma regression model as well as the Q-Q plot and AIC shall be discussed.

5.2. Preliminary Analysis
The birth weight data has a sample size of 906 with six variables. Weight represents the birth weight of babies, Age (Age of mother at delivery), Gender (Gender of baby), Marital (Marital status of mother), Education (Educational level of mother), and Employment (Employment status of mother). The weight of the baby is the response variable. Table 4.1 illustrates the variables in the data.

Variable	Description
Weight	Weight of baby
Age	Age of mother at birth
Gender	Gender of baby
Marital	Marital status of mother
Education	Educational level of mother
Employment	Employment status of mother

Table 1: Data Variables

Table 4.2 indicates the summary statistics of the quantitative variables in the data. The data has a sample size of 906 with minimum weight of 1.1kg and maximum weight of 4.6kg. The mean, median and mode of the birth weight data are respectively 2.908, 2.900 and 2.8. Since the mean is higher than the median and mode, it indicates that the birth weight data is positively skewed.

Variable	N	Minimum Weight/Age	Maximum Weight/Age	Mean	Median	Mode	Std. Dev
Weight	906	1.1	4.6	2.908	2.9	2.8	0.499
Age	906	14	47	26.386	26	23	6.286

Table 2: Summary Statistics of Baby's Birth Weight and Age of Mother

From table 4.3 below, female babies are 436, representing 48.1% less than babies who are males (470), and also representing 51.9%. On educational level of mothers, 97 mothers completed primary representing 10.7%, 237 mothers completed JHS representing 26.2%, 173mothers completed SHS representing 19.1%, 150 mothers completed tertiary representing 16.6% but 249 mothers did not go to school at all also representing 27.5%. In effect, mothers who did not go to school at all are more than those who attended school in various categories. This is likely to have a negative impact on the baby's weight. Employment status of mothers is 539 representing 59.5% as against 367 representing 40.5% of mothers who are not employed. Marital status of mothers has a hundred percent representation.

Variable			Frequency	Percent
Gender		Female	436	48.1
		Male	470	51.9
Marital		Yes	906	100
Education		None	249	27.5
		Primary	97	10.7
		JHS	237	26.2
		SHS	173	19.1
		Tertiary	150	16.6
Employment		No	367	40.5
		Yes	539	59.5

Table 3: Summary Statistics of Gender, Marital, Education and Employment

Table 3 categorizes the birth weight of babies into three. Babies with weight below 2.5kg are considered to have low weight, between 2.5kg and 4.5kg and considered to have normal weight and babies weighing above 4.5kg are considered to have high weight. From the table, 130 babies representing 14.3% of the data are categorized as low birth

weight, 775 representing 85.4% of the data are categorized as having normal weight and only 1 baby is categorized as having high birth weight representing 0.1%.

Category	Frequency	Percent	Cumulative%
Low Bweight	130	14.3	14.3
Normal Bweight	775	85.4	99.9
High Bweight	1	0.1	100
Total	906	100%	

Table 4: Categories of Birth Weight

Tables 5a and 5b illustrate the significant relationship between the weight categories and the gander of the baby as well as the Chi Square Test. In Table5, 66 female babies had low birth weight as against 64male babies, both representing 50.8% and 49.2% respectively. Again, 369 female babies representing 47.6% had normal weights as compared to 406 male babies also representing 52.4%. Only one female baby representing 100% was identified as having high birth weight. Table 4.5b is the Chi Square test used to evaluate the test of independence when using cross tabulation. The crosstabulation presents the distributions of two categorical variables simultaneously. The test of independence assesses whether an association exists between the two variables.

At 95% confidence with the null hypothesis that; no relationship exists on the categorical variables, the p-value for the Pearson Chi Square is 0.467 the Likelihood Ratio is 0.385 which are both greater than 0.05. This therefore suggests that there is no significant relationship between the categorical variables; hence the variables are independent of each other.

	Gender		
	Female	**Male**	**Total**
Category Low count	66	64	130
% within category	50.8%	49.2%	100.00%
Normal count	369	406	775
% within category	47.6%	52.4%	100.00%
High count	1	0	1
% within category	100.00!	0.00%	100.00%
Total count	436	470	906
% within category	48.1%	51.9%	100.00%

*Table 5: Category*Gender Crosstabulation*

	Value	Df	Asymptotic Significance
Pearson Chi Square	1.523	2	0.467
Likelihood Ratio	1.908	2	0.385

Table 6: Chi Square Test

Table 4.6a below is the crosstabulation of weight category and education level of mothers that test the significant relationship of these categorical variables. From the table, mothers without formal education recorded the highest figure of low birth weight of 52 representing 40.0% as against 29 for mothers who attain JHS level representing 22.3%, and 25for mothers who are SHS levers also representing 19.2%.
Mothers who attain tertiary level are 15representing only 11.5% and that of primary level are 9 also representing just 6.9%.

For babies with normal birth weight, it's seen from the table that, mothers who attain primary level recorded the least figure of 88 representing 11.4% whilst mothers at JHS level recorded the highest figure, thus 208 representing 26.8%. Also, mothers who attain SHS and tertiary levels are respectively 148 and 134 both representing 19.1% and 17.3%. Again, Table 8 is the Chi Square test. It is seen from the table that both p-values of the Pearson Chi Square and Likelihood ratio are respectively 0.016 and 0.030 which are both less than the standard 0.05(5%), that is the alpha level associated with a 95% confidence level. So, we conclude that, the variables are not independent of each other and that there is statistical relationship between the categorical variables, thus weight and educational level of mothers.

		JHS	NONE	PRIM.	SHS	TERT	Total
Category Low	count	29	52	9	25	15	130
	% within category	22.3%	40%	6.9%	19.2%	11.5%	100%
Normal	count	208	197	88	148	134	775
	% within category	26.8%	25.4%	11.4%	19.1%	17.3%	100%
High	count	0	0	0	0	0	1 100%
	% within category	0%	0 %	0%	0%	0%	
Total count		237	249	97	173	150	906
% within category		26.2%	27.5%	10.7%	19.1%	16.6%	100%

*Table 7: Category*Education Crosstabulation*

	Value	Df	Asymptotic Significance
Pearson Chi Square	18.799	8	0.016
Likelihood Ratio	16.980	8	0.030

Table 8: Chi Square Test

5.3. Further Analysis

This section depicts detailed analysis of the quantile regression model. The best quantile obtain is compared to other positively skewed distributions such as the Weibull and lognormal distributions. The AIC's of the distributions are also determined as well as the quantile-quantile plot to further elaborate on the appropriate distribution for the birth weight data.

5.4. Quantile Regression

Quantile regression offers a strategy for determining how the covariates influence the entire response distribution. It models the relationship between a response variable and the independent variables, specifying changes in the quantiles of the response variable produced by one-unit change in the predictor variables. Thus, making comparison of some percentiles of the birth weight more likely affected than others by some maternal characteristics.

The table below depicts clearly the 5th, 15th, 25th, 50th, 75th and 95th quantile regression coefficients estimates.
From table 4.7 below, the 5th, 15th, 25th, 50th, 75th, and 95th quantiles are considered. The 5th quantile is believed to be the best quantile, since it has the highest pseudo r square coefficient of 0.0439 with the 95th quantile being the worse performing with coefficient of 0.0120. The 5th quantile of birth weight of babies without employment is 0.1181818 lower than mothers with employment. However, the estimate rose in the 15th and 25th quantiles, but dropped again from the median quantile. This means that, babies born to mothers with employment contribute 0.03(15th), 0.0714286(25th), 0.0352941(50th) and 0.03333(75th) to the weight of babies. The age of mothers from the quantile coefficient estimates above all contribute positively to the weight of babies in the 5th to 95th quantiles. Again, the marital status of mothers all contributed positively to the weight of the babies and much contribution is seen in the 95th quantile (3.426667kg). For the educational level of the mother, thus, Primary, Jhs, Shs and Tertiary, mothers who completed primary on contributed positively to the weight of babies in the 75th quantile (0.0166667kg), but contributed negatively throughout the quantiles. Mothers who attained Jhs level only contributed positively to the baby's weight in the 15th, 25th, 50th and 95th quantiles. Those that completed Shs contribute negatively to the baby's weight in the 5th and 25th quantiles, but positively in the 25th, 50th, and 75th quantiles but experienced a negative contribution towards the 95th quantile. For mothers who attain tertiary level, apart from the 5th and 95th quantiles where mothers contribute negatively to babies' weight, there is positive contribution to the babies' weight in the 15th, 25th, 50th, 75th quantiles.
In effect, only the variables gender and marital contributes positively throughout the quantiles.

5.5. Quantile Regression Coefficient Estimates

Variable	5th	15th	25th	50th	75th	95th
Pseudo R2	0.0439	0.0222	0.0189	0.0253	0.031	0.012
Gender	0.1090909	0.03	0.09286	0.05882	0.1167	0.06
Educational Primary	-0.290909	-0.18	-0.1143	-0.0117	0.0167	-0.03333
Educational JHS	-0.1	0.01	0.01429	5.63E-1	-0.033	-7.70E-0
Educational SHS	-0.381818	-0.13	0.03571	0.04705	0.05	-0.09333
Educational Tertiary	0.127272	-5.50E-08	0.07857	0.07647	0.05	-0.06
Employment	-0.118181	0.03	0.07143	0.03529	0.033	-0.01
Age	0.0090909	0.01	0.00714	0.01176	0.066	0.01333
Marital	1.918182	2.24	2.37143	2.55882	2.666	3.42666

Table 9

5.6. Modeled Equations of Quantiles

$QBWEIGHT(0.05|X) = 0.0439 + 0.1090909 Gender - 0.2909091 Education(\Pr imary) - 0.1 Education(JHS)$

$-0.3818182 Education(SHS) + 0.1272727 Education(Tertiary) - 0.1181818 Employment + 0.0090909 Age +$

$1.918182(Marital)$ (4.1)

Model (4.1) measures the relationship between the dependent variable (BW eight) and the independent variables at the 5th quantile of the birth weight data. The model indicates that, the mother's educational level (primary, JHS and SHS) and employment status contribute negatively to the baby's weight, whilst all other variables (gender, educational level tertiary, age and marital status) all contribute positively to the baby's weight.

This implies that, when there is one-unit increment in the predictor variable, then the response variable (B Weight) will decrease by 0.2909091, 0.1, 0.3818182, and 0.1181818 for Education primary, JHS, SHS and Employment respectively.

$QBWEIGHT(0.15|X) = 0.0222 + 0.3 Gender - 0.18 Education(\Pr imary) + 0.01 Education(JHS) -$

$0.13 Education(SHS) - 5.50e - 08 Education(Tertiary) + 0.03 Employment + 0.1 Age + 2.24 Marital$

         (4.2)

Model (4.2) measures the relationship between the response variable and the predictor variables at the 15th quantile. Two covariates (Education SHS, Tertiary) decrease the response variable with a unit increase the predictor variables. Hence, the response variable (B weight) is decreased by 0.13 and 5.50e-08 respectively.

$QBWEIGHT(0.25|X) = 0.0189 + 0.0928571 Gender - 0.1142857 Education(\Pr imary) + 0.0142857 Education$

$(JHS) + 0.035714 Education(SHS) + 0.0785714 Education(Tertiary) + 0.0714286 Employment + 0.0071429 Age$

$+2.371429 Marital$ (4.3)

Model (4.3) also represents the relationship between the response variable and the predictor variables at the 25th percentile. It represents the first 25% of the birth weight data. It indicates that, a unit increase in the predictor variable will cause the response variable to decrease by only 0.1142857 for Education primary.

$QBWEIGHT(0.50|X) = 0.0253 + 0.588235 Gender - 0.0117647 Education(\Pr imary) + 5.63e - 10 Education(JHS)$

$+0.0470588 Education(SHS) + 0.0764706 Edacation(Tertiary) + 0.0352941 Employment + 0.0117647 Age +$

$2.558824 Marital$ (4.4)

Model (4.4) also represents the relationship existing between response variable and predictor variables at the median quantile (50th percentile). The variable that decreases the response variable with a unit increase in the predictor variables at the 25th quantile is the same variable that decreases the response variable at the 50th percentile, thus, the variable education (primary).

$QBWEIGHT(0.75|X) = 0.0310 + 0.1166667 Gender + 0.0166667 Education(\Pr imary) - 0.033333 Education$

$(JHS) + 0.05 Education(SHS) + 0.05 Education(Tertiary) + 0.03333 Employment + 0.0166667 Age + 2.66667$

$Marital$ (4.5)

This model (4.5) represents the relationship between the response variable and the predictor variables at the 75th percentile. It indicates that, only the covariate education$_{jhs}$ decrease the response variable by 0.03333 with a unit increase in the predictor variables.

$QBWEIGHT(0.90|X) = 0.0120 + 0.06 Gender - 0.033333 Education(\Pr imary) - 7.70e - 10 Education(JHS)$

$-0.093333 Education(SHS) - 0.06 Education(Tertiary) - 0.1 Employment + 0.013333 Age + 3.426667 Marital$

         (4.6)

The last model (4.6) measures the relationship that exists between the dependent variable (birth weight) and the independent variables at the 95th percentile. That is, the data points within last 95 percent of the data. At this stage, the variables Education$_{jhs}$, Education$_{shs}$, Education$_{tertiary}$, and Employment contribute negatively to the baby's weight whilst all other covariates contribute positively to the baby's weight. This means that with one-unit increment in the predictor variables, then the response variable birthweight will decrease by 0.03333, 7.70e-10, 0.093333, 0.06 and 0.1 for Education$_{primary}$, Education$_{jhs}$, Education$_{shs}$, Education$_{tertiary}$, and Employment.

5.7. Selecting the Best Performing Quantile

The 5th, 15th, 25th, 50th, 75th, are 95th quantiles are used to the measure the relationship that exist between the dependent variable and the independent variables. Their respective pseudo R square estimates are 0.0439, 0.0222, 0.0189, 0.0253, 0.0310, and 0.0120. The best quantile is determined using the particular quantile with the highest pseudo R square estimate. From table 4.3, the quantile with highest pseudo R square estimate is the 5th quantile with a coefficient of 0.0439. Its model (4.1), measures the relationship between the dependent variable (BWeight) and the independent variables at the 5th quantile of the birth weight data. The model indicates that, the mother's educational level (primary, jhs and shs) and

employment status contribute negatively to the baby's weight, whilst all other variables (gender, educational level$_{tertiary}$, age and marital status) all contribute positively to the baby's weight. This implies that, when there is one-unit increment in the predictor variable, then the response variable (BWeight) will decrease by 0.2909091, 0.1, 0.3818182, and 0.1181818 for Education$_{primary}$, jhs, shs and Employment respectively.

5.8. Parameter Estimation

       The birth weight data is used to estimate the parameters of the Weibull and the lognormal regressions. This is aimed at fitting the appropriate distribution to the data. For the Weibull, α and β are the scale and shape parameters respectively whilst for the Lognormal, u and σ are respectively the scale and shape parameters as well. However, in comparing the performance of the best quantile regression model with other skewed regression models such as the Weibull and lognormal regressions, a two parameter Weibull and lognormal distributions are obtained using the method of maximum likelihood. Table 4 below indicates the parameter values as well as the AIC's of the two distributions.

Table 4.8 compares the quality of a set of statistical models to each other. The AIC will take each model and rank them from best to worse. The "best" model will be the one that neither under-fits nor over-fits. Therefore, in A.I.C, the model with the smallest value of A.I.C is selected because this model is estimated to be closest to the unknown truth among the candidate models considered. From the table therefore, the AIC of the Weibull is 1270.1 and that of the Lognormal is 1292.3. Since the Weibull distribution is having the least AIC therefore, it is considered the preferred distribution to the Lognormal distribution. The two parameter Weibull and Lognormal distributions are as well computed. The Weibull parameters estimates are respectively 3.115 and 6.741 for α and β. Again, the parameters estimate for the lognormal distribution are ų (2.860) and σ (0.188).

Distribution	AIC	α/ų	β/σ	Std Error	t-value	Pr(>|t|)
Weibull	1270.1	3.115	6.741	0.0304	29.261	<2e- 16
Lognormal	1292.3	2.86	0.188	0.0261	36.054	<2e-16

Table 10: Parameter Estimates of Weibull and Lognormal Distributions

5.9. Model Fitness

5.9.1. Quantile Quantile (Q-Q) Plots of Weibull and Lognormal Distributions

       A Q-Q plot is used to compare the shapes of distributions, providing a graphical view of how properties such as location, scale, and skewedness are similar or different in the two distributions.

       That is, a set of intervals for the quantiles is chosen and a point (x, y) on the plot corresponds to one of the quantiles of the second distribution (i.e., the y- coordinate) plotted against the same quantile of the first distribution (x-coordinate).
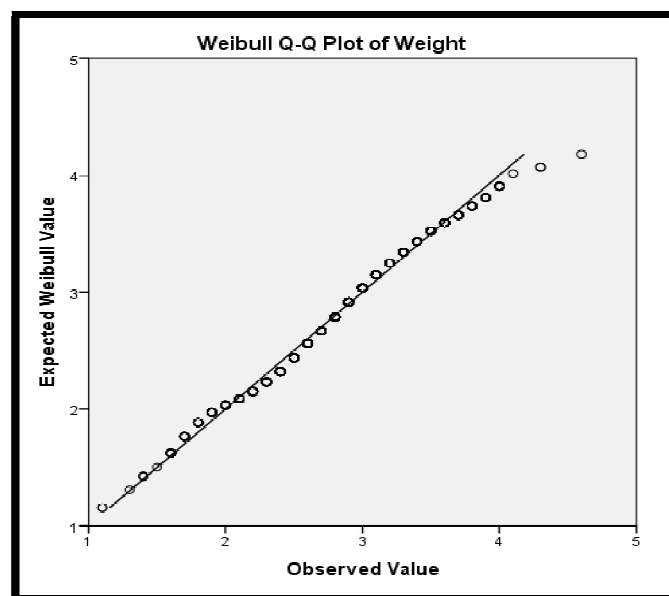

Figure 2: Q-Q Plot of Weibull Distribution

       Figure 4.2 above shows the Q-Q plot of the Weibull distribution indicating the line Y=X. The plot (black dots) from the original seem normal but only deviated from the upper tail end of the line Y=X.
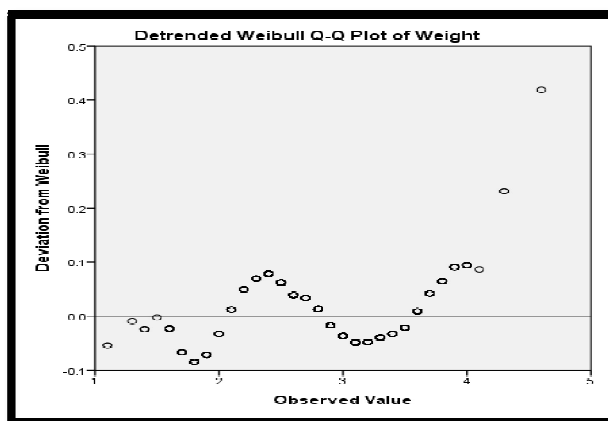
Figure 3: Detrended Weibull Q-Q plots

Figure 4 shows the detrended Weibull Q-Q plot. This plot shows the same information as the Q-Q plot for the Weibull distribution, in a different manner. In the detrended plot, the horizontal line at the origin represents the quantiles. The dots represent the magnitude and direction of deviation in the observed quantiles. Each dot is calculated by subtracting the expected quantile from the observed quantile. This implies that if a dot is below the trend line of the Weibull Q-Q plot, then it will appear above the line on the detrended Weibull Q-Q plot, because observed – expected > 0.
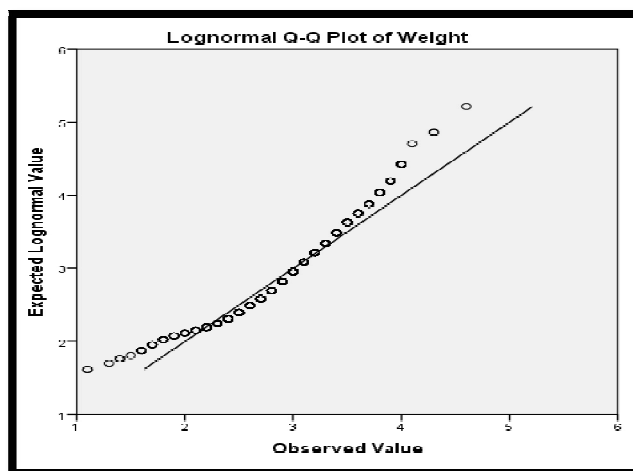

Figure 4: Q-Q Plot of Lognormal Distribution

Figure 4 above also depicts the Q-Q plot of the Lognormal distribution on the line Y=X. the plot (dotted line) started moving away from the line Y=X but moved towards the line a bit in the middle and again deviated from the line towards the tail end.

From the two plots, the Weibull looks well-constructed in the line Y=X than the Lognormal distribution. Hence, we conclude that, again, the Weibull distribution is a better fit than the Lognormal distributions confirming the AIC's.
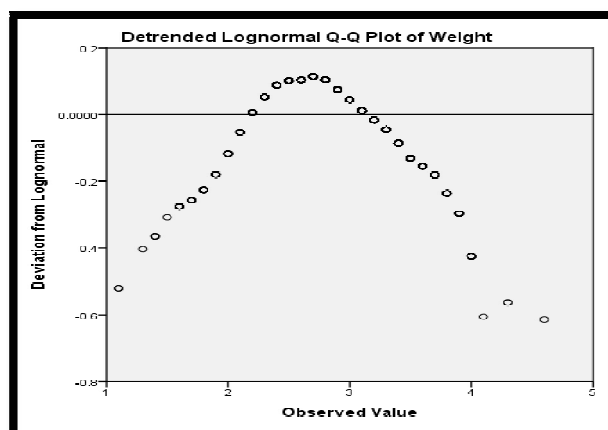

Figure 5: Detrended Q-Q Plot of Lognormal Distribution

5.9.2. Comparing Performance of Best Quantile to the Weibull and Lognormal Models

From table 4.3, the 5th quantile was selected as the best quantile with pseudo R square of 0.0439. The best quantile is selected using the quantile with the highest pseudo R square value. Again, comparing the performance of the best quantile to other positively skewed distributions such as the Weibull and Lognormal, the AIC of the 5th quantile is obtained. The AIC of the 5th quantile is 2198.717, compared to the AIC's of the Weibull and Lognormal distributions which are respectively 1270.1 and 1292.3. Although the 5th quantile is believed to be the best performing quantile, when compared to the Weibull and Lognormal models which are among the most popular models used in analyzing positively skewed data, the Weibull model is better than the Lognormal and the 5th quantile in terms of their AIC's with respect to the birth weight data.

5.9.3. The Multinomial Logistic Regression

Moreover, to model the probability of a baby falling into the different categories of birth weight (high, normal or low), factors associated with the mothers that affect the birth weight were considered and a multinomial logistic regression was fitted to the data using stepwise algorithm. The fitting of the model was done without the intercept. This is because the intercept turned out not significant at the 5% significance level. Table 9 indicated that only Age of the mother is useful in determining the probability of a baby falling into any of the weight categories. The parameter estimate of the Age of the mother is positive, indicating that as the age of the mother increases, the probability of the baby's weight increasing is higher. However, Age is only significant for low birth weight. The odds ratio estimates revealed that a mother's age is 82.21% less likely to result in low birth weight as compared to high birth weight.

Low vs High					
Variable	Estimates (B)	Standard error	Wald	Significance	Odd ratio (Exp. B)
Age	.643	.141	20.911	.000	1.901
Gender (F)	−1.372	3572.250	.000	1.000	.254
Gender (M)	0^c	–	–	–	–
Marital (Yes)	0^c	–	–	–	–
Educ (None)	16.676	3234.462	.000	.996	17471595.969
Educ (Prim)	6.120	5669.170	.000	.999	454.946
Educ (JHS)	3.530	8282.331	.000	1.000	34.125
Educ (SHS)	−16.665	6319.076	.000	.998	$5.787E-8$
Educ (Tert)	0^c	–	–	–	–
Emp (No)	17.653	5241.693	.000	.997	46427925.602
Emp (Yes)	0	–	–	–	–
	Normal vs High				
Age	.576	.000	–	–	1.779
Gender (F)	−1.679	3572.250	.000	1.000	.186
Gender (M)	0^c	–	–	–	–
Marital (Yes)	0^c	–	–	–	–
Educ (None)	17.683	3234.462	.000	.996	47812416.044
Educ (Prim)	6.701	5669.170	.000	.999	812.984
Educ (JHS)	16.097	8230.113	.000	.998	9795122.757
Educ (SHS)	8.893	6179.448	.000	.999	7277.995
Educ (Tert)	0^c	–	–	–	–
Emp (No)	−6.471	5072.495	.000	.999	0.002
Emp (Yes)	0^c	–	–	–	–

Table 11: Parameter Estimates of the Multinomial Logistic Regression Model

6. Summary of Findings

This study was aimed at assessing the determinants of birth weight of babies using quantile regression methods. The quantiles; 5th, 15th, 25th, 50th, 75th and 95th were obtained with their respective pseudo R square estimates as follows from table 4.3: (0.0439), (0.0222), (0.0189), (0.0253), (0.0310), and (0.0120). The birth weight data was tested to proof positive skewness. In table 4.2 of the summary statistics, the mean of the data was higher than the mode and median indicating that the data is positively skewed. This positive skewedness was affirmed using the boxplot in figure 4.1. The equation of the best performing quantile was modeled using the data variables. The two parameter Weibull and Lognormal distributions as well as their AIC are estimated as seen table 4.4.

The shape of the Q-Q plots and Detrended Q-Q plots for the Weibull and Lognormal distributions were plotted as seen in figures 4.2, 4.3, 4.4, and 4.5 respectively.

7. Conclusions

The objectives of this research are among other things to assess the determinants of birth weight of babies in the Bolgatanga municipality. To identify the best quantile that best describes the weight of babies and as well compare the performance of this best quantile to other positive skewed models such as the Weibull and Lognormal models. From table 4.1 the mean, median and mode are respectively 2.908, 2.900 and 2.8, indicating that the birth weight data is positively skewed since the mean has the highest value. Again, in identifying the appropriate quantile that best describes the data, the pseudo R square estimates of the selected quantiles are obtained as depicted in table 4.4. The 5th quantile was identified as having the highest pseudo R square value of 0.0439 compared to the 15th, 25th, 50th, 75th, and the 95th quantiles. Hence, the 5th quantile is selected as the appropriate quantile than best describes the birth weight data. However, with the use of AIC's and Q-Q plots, the Weibull regression is best fit for modeling, compared to the Lognormal and the 5th quantile models in terms of performance.

Again, Table 4.9 indicates that only Age of the mother is useful in determining the probability of a baby falling into any of the weight categories. The parameter estimate of the Age of the mother is positive, indicating that as the age of the mother increases, the probability of the baby's weight increasing is higher.

8. Recommendations

On the basis of the research, the following recommendations were made:

- Other statistical information should be employed from the results obtained from the various quantile regression and the other skewed distributions.
- Other model selection procedures should be employed.
- Several variables should be included in future research works such as nutritional intake of mothers.
- Future research works should expand the scope and size of the data points.

9. References

i. Abrevaya, J. (2006). Estimating the effect of smoking on birth outcomes using a matched panel-data approach. Journal of Applied Econometrics, (21): 489–519
ii. Almond, D., Kenneth, Y. C., and David, S.L. (2005). The costs of low birth weight. Quarterly Journal of Economics (120): 1031–1083.
iii. Angrist, J., Victor, C., and Ivan, F.V. (2006). Quantile regression under misspecification, with an application to the U. S. wage structure Econometrica (74): 539–563.
iv. Casella, G., and Berger, R. L. (2002). Statistical Inference (2nd ed.), Duxbury.
v. Celik, A. N. (2003). A Statistical Analysis of Wind Power Density Based on the Weibull and Rayleigh Models at the Southern Region of Turkey. Renewable Energy (29): 593–604, (http://www.elsvier.com/locate/renene).
vi. Coles, S. (2001). An Introduction to Statistical Modeling of Extreme Values, Springer - Verlag
vii. Jones, O., Maillardet, R., and Robinson, A. (2009). Introduction to Scientific Programming and Simulation Using R, Chapman and Hall/CRC.
viii. Koenker, R. and Gilbert B. Jr., (1982). Robust tests for heteroscedasticity based on re- gression quantiles. Econometrica 50, 43–62.
ix. Koenker, R. and Gilbert, B. (1978). Regression quantiles. Econometrica 46, 33–50.
x. Koenker, R. (2004). Quantile regression for longitudinal data. Journal of Multivariate Analy- sis (91): 74–89.
xi. Koenker, R. and Kevin, F. H. (2000). Quantile regression: an introduction, unpublished working paper.
xii. Koenker, R. and Kevin, F. H. (2001). Quantile regression, Journal of Economic Perspec- tives. 15, 143–156.
xiii. Lien, D. S. and William, N. E. (2005). Estimating the impact of large cigarette tax hikes: the case of maternal smoking and infant birth weight. Journal of Human Resources. 40, 373–392. London Limited.
xiv. Murthy, D. P., Xie, M., and Jiang, R. (2004). Weibull Models, John Wiley and Sons, Inc.
xv. Permutt, T. and Richard, H. (1989). Simultaneous-equation estimation in a clinical trial of the effect of smoking on birth weight. Biometrics 45, 619–622.
xvi. Shah, N.R. and Bracken, M.D. (2000) A Systematic Review and Meta-Analysis of Prospective. Studies on the Association between Maternal Cigarette Smoking and Preterm Delivery.
xvii. Sr´am, R. J., Binkov´a, B., Dejmek, J., and Bobak, M. (2005). Ambient air pollution and pregnancy outcomes: a review of the literature. Environmental health perspectives 113 Wooldridge, J. M., (2002). Econometric analysis of cross section and panel data, Cambridge, MA: MIT Press.
xviii. Wu, S.J. (2002). Estimations of the Parameters of the Weibull Distribution with Progres- sively Censored Data. Journal of The Japan Statistical Society, 32, 155–163.
xix. Zhang, L., Xie, M., and Tang, L. (2007). A study of two estimation approaches for param- eters of Weibull distribution based on WPP. Reliability Engineering and System Safety,