**International Journal for Social Studies**
Available at https://pen2print.org/index.php/ijss/

**ISSN: 2455-3220**
**Volume 04 Issue 12**
**December 2018**

# On The Average Number of Households with Malaria Control in Nigeria: A Bayesian Approach

**Oladoja, Oladapo Muyiwa**

oladojadapo@gmail.com

Department of Statistics,

University of Ibadan, Ibadan, Nigeria.

## Abstract

*Malaria, a common disease poses a great threat to human existence in Nigeria. Though various ways existed by which the endemic disease can be prevented amongst are the use of mosquito nets and indoor residual spraying (IRS). This study determine the average number of households with mosquito nets and IRS among the six geopolitical zones in Nigeria using the Bayesian approach. Assuming normality for the households with mosquito nets and IRS, and also using a conjugate normal prior, the posterior density follows a normal distribution. For households with mosquito nets, the posterior mean and standard deviation respectively for the six geopolitical zones are 5864.33 & 824.662 for North-Central, 5186.51 & 824.662 for North-East, 9293.89 & 824.662 for North-West, 4825.68 & 824.662 for South-East, 5287.67 & 824.662 for South-West and 7214.91 & 824.662 for South-South. For households with IRS, the posterior mean and standard deviation respectively for the six geopolitical zones are 4454.97 & 566.391 for North-Central, 4644.03 & 566.391 for North-East, 6820.33 & 566.391 for North-West, 4246.24 & 566.391 for South-East, 3664.52 & 566.391 for South-West and 4873.29 & 566.391 for South-South. The North-West geopolitical zone had the highest average number of households with mosquito nets and IRS, while the South-East geopolitical zone had the lowest average number of households with mosquito nets and the South-South geopolitical zone had the lowest number of households with IRS.*

**Keywords: Posterior Mean, Posterior Standard Deviation, Conjugate Normal Prior.**

## 1.0 INTRODUCTION

Nigeria is the most populous country in Africa and the seventh most populous country in the world. Nigeria has the third-largest youth population in the world, after India and China, with more than 90 million of its population under age 18. Malaria is endemic in Nigeria, with year-round transmission. Malaria remains an important cause of morbidity and mortality in Nigeria. Nigeria accounted for 32% of the global estimate of 655,000 malaria deaths in 2010 (World Health Organization, 2012). An estimated 97% of the country's approximate population of 160 million residents are at risk of malaria. Children under age 5 and pregnant women are the groups most vulnerable to illness and death from malaria infection in Nigeria.

The National Malaria Control Strategic Plan (NMCSP) addresses malaria control through three core interventions: prevention of malaria transmission through an integrated vector management strategy, prompt diagnosis and adequate treatment of clinical cases at all levels and in all sectors of health care, and prevention and treatment of malaria in pregnancy. The use of mosquito nets is the main method of malaria prevention employed in Nigeria. Free nets are distributed through mass campaigns, public health facilities, faith-based organizations, nongovernmental organizations (NGOs), retail commercial outlets, and maternal and child health

ISSN: 2455-3220
Volume 04 Issue 12
December 2018

weeks with the goal of achieving universal access. Indoor residual spraying (IRS) is another component of efforts to control malaria transmission in Nigeria. To obtain information on coverage of indoor residual spraying, all households interviewed in the 2013 NDHS were asked whether the interior walls of their dwelling had been sprayed against mosquitoes during the 12-month period before the survey and, if so, who had sprayed the dwelling. Part of the new strategy in the fight against malaria is to increase IRS coverage in Nigeria.

This study is aimed at determining the average number of households with mosquito nets and IRS in the six geopolitical zones of Nigeria using the Bayesian approach.

## 2.0 MATERIALS AND METHODS

In statistics, Bayesian inference is a method of inference in which Bayes' rule is used to update the probability estimate for a hypothesis as additional evidence is acquired. A typical Bayesian analysis can be outlined in the following steps:

**Formulation of a probability model for the data:** The first step in a Bayesian analysis is to choose a probability model for the data. This process, which is analogous to the classical approach of choosing a data model, involves deciding on a probability distribution for the data if the parameters were known. If the n data values to be observed are $y_1, y_2, ..., y_n$ and the vector of unknown parameters is denoted $\theta$ then, assuming that the observations are made independently, we are interested in choosing a probability function for the data.

**Deciding on the Prior Distribution:** Once the data model (probability model) is chosen, a Bayesian analysis requires the assertion of a prior distribution for the unknown model parameters. Parameters of prior distributions are called hyper parameters, to distinguish them from the model of the underlying data.

**Normal Density Prior:** We parameterize this family as $P(\theta) \propto e^{-\frac{1}{\sigma_0^2}(\theta-\theta_0)^2}$. That is, $\theta \sim N(\theta_0, \sigma_0^2)$ with hyper parameters $\theta_0$ and $\sigma_0^2$.

**Construction of the Likelihood Function:** The likelihood function is given by $L(\theta \mid y_i) = \prod_{i=1}^{n} P(y_i \mid \theta)$

**Normal Likelihood:** The normal density is fundamental to most statistical modeling. Assuming the normal model is appropriate we will derive the result for a sample of independent and identically distributed observations $y = (y_1, y_2, ..., y_n)$

$$f(y_1, y_2, ..., y_n \mid \theta) = \prod_{i=1}^{n} f(y_i \mid \theta) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\theta)^2}$$

**Posterior Distribution:** Mathematically, it is written as $P(\theta \mid y_i) \propto P(y_i \mid \theta) \times P(\theta)$

**Normal Density Prior and Normal Likelihood**

$$P(\theta \mid y) \propto \exp\{-\frac{1}{2\sigma_0^2}(\theta-\theta_0)^2\}.\exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\theta)^2\}$$

ISSN: 2455-3220
Volume 04 Issue 12
December 2018

$$\propto \exp\{-\frac{1}{2\sigma_0^2\sigma^2}[\sigma^2(\theta-\theta_0)^2+\sigma_0^2\sum_{i=1}^{n}(y_i-\theta)^2]\} \propto \exp\{-\frac{1}{2}[\theta^2(\frac{n}{\sigma^2}+\frac{1}{\sigma_0^2})-2\theta(\frac{n\bar{y}}{\sigma^2}+\frac{\theta_0}{\sigma_0^2})]\}$$

$$P(\theta\,|\,y)\propto \exp\{-\frac{1}{2}(\frac{1}{\sigma_n^2})[\theta^2-2\theta\theta_n]\}\propto\{-\frac{1}{2}\frac{(\theta-\theta_n)^2}{\sigma_n^2}\}$$

Therefore, $\theta\,|\,y \sim N(\theta_n,\sigma_n^2)$

## 3.0 RESULTS AND DISCUSSION

***Setting up the Prior Density***

Assuming normality for the households with mosquito nets and IRS, and using a conjugate normal prior density, if we are interested in estimating the true average number of households with mosquito nets and IRS in Nigeria. We have an intuition that the median of our prior distribution for households with mosquito nets is 5576 and 4549.5 for households with IRS. We see from the output that the normal density with mean $\theta_0$ = 5576 & 4549.5 and $\sigma_0$ = 2081.644 & 1489.798 for households with mosquito nets and IRS respectively matches our prior information. This implies that the parameter $\theta_0 \sim N(\theta_0,\sigma_0^2)$, so that for households with mosquito nets, the prior distribution is $N(5576, 2081.644^2)$ and for households with IRS, the prior distribution is $N(4549.5, 1489.798^2)$.

If we consider an alternative prior density to model our beliefs about the true average value of the number of households with mosquito nets and IRS in Nigeria. Any symmetric density instead of a normal could be used, so we use a t density with location $\theta_0$, scale $\tau$ and 2 degrees of freedom. Since our prior median is 5576 and 4549.5 for households with mosquito nets and IRS respectively, we let the median of our t density be equal to $\theta_0$ = 5576 & 4549.5. We find the scale parameter $\tau$, so the t density matches our prior belief that where $T$ is a standard t variate with two degrees of freedom. It follows that

$$\tau = 3500/t_2(0.95),$$

Where $t_v(p)$ is the pth quantile of a random variable with v degrees of freedom. We find $\tau$ by using the t quantile function.

We display the normal and t priors in a single graph each for both households with mosquito nets and IRS in figure 1 below. Although they have the same basic shape, note that the t density has irregular patterns.

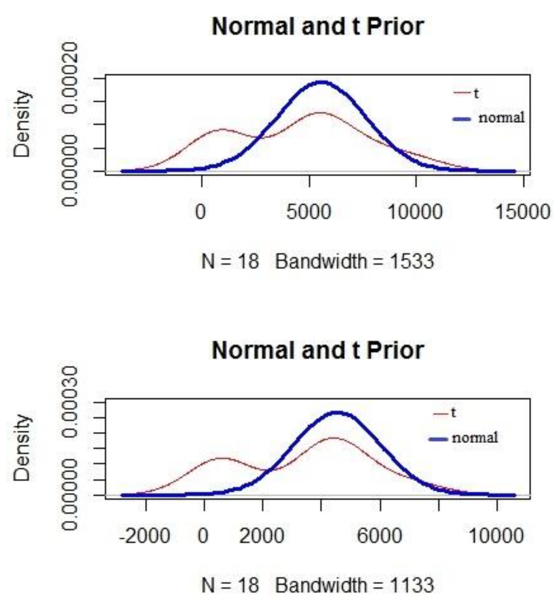



**Figure 1** Normal and t priors for representing prior opinion about households with mosquito nets and IRS in Nigeria respectively.

***Computation of the Posterior Density***

ISSN: 2455-3220
Volume 04 Issue 12
December 2018

| Zone | Household with Mosquito Nets | | |
|---|---|---|---|
| | $\bar{y}$ | $\theta_n$ | $\sigma_n$ |
| North-Central | 5918 | 5864.33 | 824.662 |
| North-East | 5114 | 5186.51 | 824.662 |
| North-West | 9986 | 9293.89 | 824.662 |
| South-East | 4686 | 4825.68 | 824.662 |
| South-South | 5234 | 5287.67 | 824.662 |
| South-West | 7520 | 7214.91 | 824.662 |

Since Nigeria comprises of six geopolitical zones, the breastfeeding values for the zones are $y_1, y_2, y_3, y_4, y_5, y_6$. Assume that an individual zone y is distributed as $N(\theta, \sigma^2)$ with known standard deviation, the observed mean of households with mosquito nets and IRS value $\bar{y}$ is $N(\theta, \sigma^2/6)$

With the use of a normal prior in this case, the posterior density of $\theta$ will also have a normal functional form. Recall that the precision is defined as the inverse of the variance. Then the posterior precision $P_n = \frac{1}{\sigma_n^2}$ is the sum of the data precision $P = \frac{n}{\sigma^2}$ and the prior precision $P_0 = \frac{1}{\sigma_0^2}$

$$P_n = P + P_0 = \frac{6}{\sigma^2} + \frac{1}{\sigma_0^2}$$

The posterior variance is given as,

$$\sigma_n^2 = 1/P_n = 1/(6/\sigma^2 + 1/\sigma_0^2)$$

The posterior mean of $\theta$ is expressed as a weighted average of the sample mean and the prior mean where the weights are proportional to the precisions.

$$\theta_n = \frac{\bar{y}P + \theta_0 P_0}{P + P_0} = \frac{\bar{y}(6/\sigma^2) + \theta_0(1/\sigma_0^2)}{6/\sigma^2 + 1/\sigma_0^2}$$

We illustrate the posterior calculations for the six geopolitical zones in Nigeria in year 2008 and 2013. In each case we compute the posterior mean and posterior standard deviation of households with mosquito nets and IRS in Nigeria geopolitical zones as displayed in Tables 1 and 2.

**Table 1**: Posterior mean and standard deviation of households with mosquito nets in Nigeria.

We perform the posterior calculations using the t prior for each of the possible sample results. Not e that the posterior density of $\theta_n$ is given, up to a proportionality constant, by

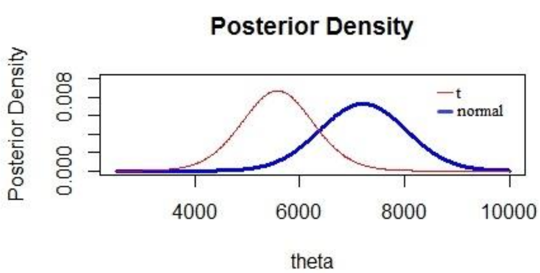
$$P(\theta_n \mid data) \propto P(\bar{y} \mid \theta, \sigma^2/n) P(\theta \mid v, \theta_0, \tau),$$

Where $P(y \mid \theta, \sigma^2)$ is a normal density with mean $\theta$ and standard deviation $\sigma$, and $P(\theta \mid v, \theta_0, \tau)$ is a t density with median $\theta_0$, scale parameter $\tau$ and degrees of freedom v. Since this density does not have a convenient functional form, we summarize it using a direct prior times likelihood approach. We construct a grid of $\theta$ values that covers the posterior density, compute the product of the normal likelihood and the t prior on the grid, and convert these products to probabilities by dividing by the sum. Essentially we are approximating the continuous posterior density by a discrete distribution on this grid.

ISSN: 2455-3220
Volume 04 Issue 12
December 2018

Figure 2 shows the posterior densities for the two priors for households with mosquito nets and IRS respectively. When a normal prior is used, the posterior will always be a compromise between the prior information and the observed data, even when the data result conflicts with one's prior beliefs about location of Nigeria's average value of households with mosquito nets and IRS. In contrast, when a t prior is used, the likelihood will be in flat tailed portion of the prior and the posterior will resemble the likelihood function.

**Table 2**: Posterior mean and standard deviation of households with IRS in Nigeria.



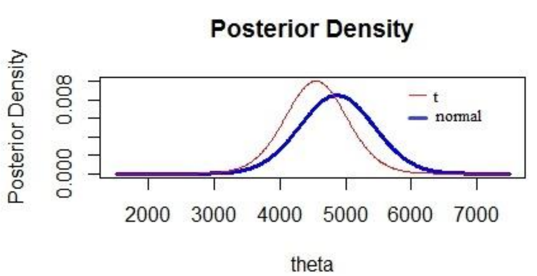




**Figure 2**. Posterior densities for average value of Nigeria's households with mosquito net and IRS using normal and t prior respectively.

## 4.0 CONCLUSION

This research outlines the basic goal of Bayesian inference inform of the posterior, which is understood as a weighted average between knowledge about the parameter before data is observed (which is represented by the prior distribution) and the information about the parameters contained in the observed data (which is represented by the likelihood function). Since malaria remains an important cause of morbidity and mortality in Nigeria. There is need to evaluate the measures of preventing it. In line with the aim of this work, we have been able to use Bayesian statistics to determine on average the number of households with mosquito nets and IRS in the six geopolitical zones of Nigeria. Having assumed that the number of households with mosquito nets and IRS follows a Normal density using a known standard deviation of 2200 and 1500 respectively.

| **Zone** | **Household with IRS** | | |
|---|---|---|---|
| | $\bar{y}$ | $\theta_n$ | $\sigma_n$ |
| North-Central | 4439 | 4454.97 | 566.391 |
| North-East | 4660 | 4644.03 | 566.391 |
| North-West | 7204 | 6820.33 | 566.391 |
| South-East | 4195 | 4246.24 | 566.391 |
| South-South | 3515 | 3664.52 | 566.391 |
| South-West | 4928 | 4873.29 | 566.391 |

Using a conjugate prior of normal density, we were able to show that the posterior density of households with mosquito nets and IRS follows a normal distribution. Thereafter, we obtained the posterior mean and standard deviation. The North-West geopolitical zone had the highest average number of households with mosquito nets and IRS, while the South-East geopolitical zone had the lowest average number of households with mosquito nets and the South-South geopolitical zone had the lowest number of households with IRS.

## 5.0 REFERENCES

Albert, J. (2009). Bayesian Computation with R. *Springer Science, Second Edition.*

Annest, A., Bumgarner, R. E., Raftery, A. E. and Yeung, K.Y. (2014). The Iterative Bayesian Model

Averaging Algorithm for Survival Analysis: an Improved Method for Gene Selection and Survival Analysis on Microarray Data.

Flanders, W. D. and Klein Baum, D. G. (1995). Basic models for disease occurrence in epidemiology. *International Journal of Epidemiology*. 1995; 24: 1-7.

Gelman, A., Carlin, J. B., Stern, H.S. and Rubin, D. B. (2004). Text in Statistical Science, 3rd ed. Bayesian Data Analysis. *Chapman & Hall, New York, USA*, pp. 1-163.

Glickman, M.E. and Van Dyk D.A. Methods in Molecular Biology, *Humana Press, Totowa, NJ, USA* pp. 319-338.