Bayesian Analysis of Average Age Prevalence of Hypertensive Patients in South Western Part of Nigeria

Olawale Basheer Akanbi
Department of Statistics, University of Ibadan, Ibadan

*Corresponding Author: Olawale Basheer Akanbi, Department of Statistics, University of Ibadan, Ibadan

ABSTRACT

Hypertension is a non-communicable disease which is mostly related to age distribution and is rampant in Nigeria. One is likely to develop high blood pressure as long he is getting older. This paper was aimed to determine the average age that one is likely to develop hypertension in Ibadan North Local Government (IBLG). Bayesian Inference was used to analysis this problem. Both uniform and Normal priors were used for the probability update. The month that had the highest number of cases was January and the minimum was recorded in November. It was also observed that 862 (73.4%) of hypertensive patients in the IBLG were female while 312 (26.6%) of them were male. Age group with the highest case was between (52 and 61) years (29.1%) while only one person (0.1%) fell in the age group of 92 years and above. For the Classical, uniform and conjugate priors, the posterior means age of hypertensive patients were (53.5, 53.4 & 53.7) years respectively. The 95% credible interval for the age using conjugate prior was [47.23, 60.17]. It was recommended that an informative normal prior and normal likelihood should be considered for analysing average age prevalence of hypertensive patients in IBLG.

Keywords: Conjugate prior, posterior distribution, Normal distribution, Uniform Distribution, Blood pressure.

INTRODUCTION

Infectious and Non Communicable Diseases (NCDs) are serious issues affecting the health sector. The NCDs Cardiovascular, diabetes, cancer, and chronic respiratory diseases are common challenges to human health and development. These four diseases are the world’s biggest killers, causing in excess of 36 million deaths each year of which 80% are in low- and middle-income countries (World Health Organization (WHO) 2011). Unhealthy diet, physical inactivity, tobacco use and the harmful use of alcohol are the behavioral risk factors responsible for about 80% of coronary heart disease and cerebrovascular disease deaths (WHO 2013). The WHO also predicted globally, between 2006 and 2015, NCDs deaths to increase by 17% with highest in the African region by 27%. The most rampant NCDs is hypertension. Hypertension is a common and a major public health problem (Murray et al 1997). It is a serious threat to the health of adults usually in Africa. Hypertension was defined as systolic Blood Pressure (BP) ≥ 140mmHg or as diastolic BP ≥ 90mmHg or being on drug therapy. Recently, the prevalence of hypertension was observed to be 28% in North America, 44% in Western Europe (Wolf-Maier et al 2003), 28.3% for crude and 27.3% for age-standardized in urban Accra of Ghana West Africa (Amoah AGB 2003).

Blood pressure is therefore the force of blood against the artery walls as it circulates through the body. Blood pressure is written as two numbers. The first (systolic) number represents the pressure in blood vessels when the heart beats. The second (diastolic) number represents the pressure in the vessels when the heart rests between beats. High blood pressure can cause serious damage to health. It can harden the arteries, decreasing the flow of blood and oxygen to the heart. Perhaps, this reduction of the blood flow can cause chest pain, angina or heart failure which occurs when the heart cannot pump enough blood and oxygen to other organs, heart attack which occurs when the blood supply to the heart is blocked and heart muscle cells die from lack of oxygen. The longer the blood flow is blocked, the greater the damage to the heart. High blood pressure can burst or block arteries that supply blood and oxygen to the brain causing stroke. High blood...
Bayesian Analysis of Average Age Prevalence of Hypertensive Patients in South Western Part of Nigeria

pressure cannot be cured. It can however, be managed very effectively through lifestyle changes and, when needed medication. In most cases, the cause of high blood pressure is not known. In fact, high blood pressure usually doesn’t have symptoms. This is why it is sometimes called “the silent killer.” However, there are known risk factors for high blood pressure in which age is included. This deadly disease is not only common among the non-medical staff but also the Healthcare Workers (HCW) especially the nurses. For instance, the report of the Medical Research Council (MRC) of the South Africa showed that NCDs (hypertension) accounted for 37% of deaths (Steyn et al 2006) and recently 20% of deaths were attributed to hypertension among the HCW in South Africa (Skaal L and Pengpid S 2011). This compared with that of the Mexico HCW of 22% of deaths caused by the NCDs (Gonzalez-Valazquez et al 2006).

**METHODOLOGY**

The use of Bayesian reasoning in statistics and economics has roughly tripled since 1970 (Poier, 2006). Bayesian inference uses the Bayes’ theorem to update the degree of belief in the occurrence of certain events. Therefore, the Bayes theorem states that,

**Posterior \propto Likelihood \times Prior**

Mathematically given as:

\[
P(\theta|y_i) \propto P(y_i|\theta) \times P(\theta)
\]

\[
P(\theta|y_i) = \frac{P(y_i|\theta) \times P(\theta)}{P(y_i)} \tag{1}
\]

Where,

\[
P(y_i)
\]

Is the evidence (normalization constant useful for selection of Bayesian Model)

\[
P(y_i) = \int P(y_i|\theta) \times P(\theta) \, d\theta \tag{2}
\]

**Constructing the Likelihood Function**

The likelihood function is constructed once the data has been observed. The likelihood function measures the support provided for each possible value of the parameter by the data. then, the likelihood function is given by;

\[
P(y_i|\theta) = L(\theta|y_1, ..., y_n) = P(y_1, ..., y_n|\theta) = \prod_{i} P(y_i|\theta)
\]

\[
p(y|\mu) \propto e^{-\frac{1}{2\sigma^2}(y-\mu)^2} \tag{3}
\]

Equation (3) is the kernel distribution for the normal density function.

**Posterior Mean and Variance for the Conjugate Prior**

The observation \( y \) is a random variable taken from a normal distribution with mean \( \mu \) and variance \( \sigma^2 \) which is assumed known. We have a prior distribution that is normal with mean \( m \) and variances \( \sigma^2 \). The shape of the prior density is given by

\[
g(\mu) \propto e^{-\frac{1}{2\sigma^2}(\mu-m)^2} \tag{4}
\]

Where, we are ignoring the part that doesn’t involve \mu because multiplying the prior by any constant of proportionality will cancel out in the posterior. The shape of the likelihood is

\[
f(y|\mu) \propto e^{-\frac{1}{2\sigma^2}(y-\mu)^2} \tag{5}
\]

Where, we ignore the part that doesn’t depend on \mu because multiplying the likelihood by any constant will cancel out in the posterior. The prior times likelihood is

\[
P(\mu|y) = g(\mu) \times f(y|\mu) \propto e^{-\frac{1}{2\sigma^2}(\mu-m)^2} \tag{6}
\]

Putting the terms in exponent over the common denominator and expanding them out gives

\[
e^{-\frac{1}{2\sigma^2}(\mu^2-2\mu m+m^2)+s^2(y^2-2\mu y+\mu^2)}
\]

We combine the like terms

\[
e^{-\frac{1}{2\sigma^2}[(\sigma^2+m^2)\mu^2-2\sigma^2m\mu+y^2\mu+\mu^2]} \tag{7}
\]

and factor out \( \sigma^2+s^2/\sigma^2s^2 \). Completing the square and absorbing the part that doesn’t depend on \mu into the proportionality constant, we have

\[
e^{-\frac{1}{2\sigma^2+s^2/\sigma^2s^2}[(\sigma^2+m^2)\mu^2-2\sigma^2m\mu+y^2\mu+\mu^2]} \tag{7}
\]

Thus, the posterior mean,

\[
m' = \frac{\sigma^2 m + s^2 y}{\sigma^2 + s^2} \tag{8}
\]

and

The posterior variance,

\[
(s')^2 = \frac{(\sigma^2 m + s^2 y)}{\sigma^2 + s^2} \tag{9}
\]
Bayesian Analysis of Average Age Prevalence of Hypertensive Patients in South Western Part of Nigeria

Hence, the prior has a normal(m, s^2) and also, the posterior has a normal(m', (s')^2). This shows that the normal(m, s^2) distribution is the conjugate family for the normal observation distribution with known variance.

**Credible Interval**

A credible interval is an interval that contains the probability of the parameter. It is an interval in the domain of a posterior probability distribution used for interval estimation.

\[
E(\theta/y) \pm Z_{\frac{\alpha}{2}}SD(\theta/y) \tag{10}
\]

Where,

\[
E(\theta/y) = \text{Posterior Mean}, \quad SD(\theta/y) = \text{Posterior Standard Deviation}
\]

\[
d = Z_{\frac{\alpha}{2}} = z - \text{value}
\]

**DATA ANALYSIS**

Table I shows cross tabulation of sex and age. It is clearly seen that for each age group, women have more cases compare to men. The female average age was 53.10 while the male average age was 53.50. Also, the age group 52-61 years had the highest occurrence of hypertension.

<table>
<thead>
<tr>
<th>Age Distribution</th>
<th>Female</th>
<th>Male</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 - 11yr</td>
<td>5</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>12 - 21yr</td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>22 - 31yr</td>
<td>24</td>
<td>18</td>
<td>42</td>
</tr>
<tr>
<td>32 - 41yr</td>
<td>134</td>
<td>29</td>
<td>163</td>
</tr>
<tr>
<td>42 - 51yr</td>
<td>227</td>
<td>72</td>
<td>299</td>
</tr>
<tr>
<td>52 - 61yr</td>
<td>252</td>
<td>90</td>
<td>342</td>
</tr>
<tr>
<td>62 - 71yr</td>
<td>151</td>
<td>86</td>
<td>237</td>
</tr>
<tr>
<td>72 - 81yr</td>
<td>52</td>
<td>10</td>
<td>62</td>
</tr>
<tr>
<td>82 - 91yr</td>
<td>12</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>92yr+</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>862</td>
<td>312</td>
<td>1174</td>
</tr>
</tbody>
</table>

**Determination of Distribution of the Data Using Normal Curve**

From the figure below, it is observed that the data is normally distributed. Therefore, the normal distribution density function will be employed to obtain the data likelihood.

![Histogram](Image)

Figure 1. Normal Curve Depicting Data Distribution

Therefore,

\[
\mu = 53.23 \quad \text{and} \quad \sigma = 13.168
\]

The mean and standard deviation of the data are 53.23 and 13.168 respectively.
Bayesian Analysis of Average Age Prevalence of Hypertensive Patients in South Western Part of Nigeria

H₀ = The distribution of the age follows normal distribution
H₁ = The distribution of the age does not follow normal distribution.

Table 2. Chi-square Easyfit Output

<table>
<thead>
<tr>
<th>Chi-Squared</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Deg. of freedom</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistic</td>
<td>19.312</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P-Value</td>
<td>0.03648</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rank</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reject?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

From Table II shows that normal distribution was rank the best among other distribution that was in line with the data and was accepted at 98% and 99% Confidence Interval.

Bayesian Analysis

A non-informative and informative prior will be used to ascertain the average age of hypertensive patients in the area under coverage.

Posterior Mean and Standard Deviation Using Uniform Prior

The Bayes Estimate (Posterior mean) of the age distribution of the patients is;

\[ E(\theta | y) = m = \frac{(\sigma^2 m + s^2 y)}{\sigma^2 + s^2} = 53.7014 \]

Where Posterior Variance is

\[ V(\theta | y) = (s')^2 = \frac{(\sigma^2 m + s^2 y)}{\sigma^2 + s^2} = 10.8921 \]

Therefore, the Posterior Standard Deviation is = 3.3003

The credible interval

\[ E(\theta / y) = 53.7014, \quad SD(\theta / y) = 3.3003 \text{ and } Z_{\frac{\alpha}{2}} = 1.96 \]

Therefore,

The 95% credible interval for y (the average age of hypertensive patients): [47.23, 60.17]

RESULTS AND SUMMARY OF FINDINGS

A total number of 1174 patients’ information was used in the analysis. The data was obtained between January and November 2016. The age of the patients range from 2 to 99 years which was grouped into 10 classes of 10 class size. Majority 342(29.1%) of the patients fall within the age group 52 and 61 years while only 1 patient fall within the age group 92 and 99 years. More than three-quarters of the patients’ ages 1041(88.7) fall within four age groups (i.e between 32 and 71 years) while 133(11.3%) which less than one-quarter fall within the remaining six age groups (i.e. between 2 and 31 years, and 62 and 99 years). The average age of the patients is 53 years ± 13.4568. 862(73.4%) of
the patients were female while 312(26.6%) of them were male. The female average age was 53.10±13.203 while the male average age was 53.50±13.301. Chi-square and normal curve was used to show that the age distribution of the respondents follow normal distribution. Using a positive uniform prior (which is non-informative), the posterior follows a normal distribution. The posterior mean (Bayes estimate) of the age distribution of hypertensive patients is 53.402 while the posterior standard deviation is 9.098. Using a Conjugate prior (which is informative), the prior, the likelihood and the posterior follow a normal distribution. The posterior mean (Bayes estimate) of the age distribution of hypertensive patients is 53.7014 while the posterior standard deviation is 3.3003. The posterior confidence interval for average age of hypertensive patients is [47.23, 60.17].

**CONCLUSION**

This research outlines the basic goal of Bayesian inference, which is understood as a weighted average between knowledge about the parameter before data is observed (which is represented by prior distribution) and the information about the parameters (average age of hypertensive patients) contained in the observed data (which is represented by the likelihood function). In all, the following findings were made: Having grouped data into 10 age groups, the frequency distribution of age groups of hypertensive patients in the area under study was obtained and the age group that had the highest occurrence is 52-61 years. Age distribution of hypertensive patients in the area under study follows a normal distribution. The average age of hypertensive patients in the area was obtained to be 53.23. For the classical approach, non-informative and informative prior give the same average age (53 years) but 13.168, 9.098 and 3.3003 standard deviation respectively. Credible Interval for posterior mean was obtained to be [47.23, 60.17].

**REFERENCES**


Bayesian Analysis of Average Age Prevalence of Hypertensive Patients in South Western Part of Nigeria


Bayesian Analysis of Average Age Prevalence of Hypertensive Patients in South Western Part of Nigeria

Communicable Diseases in the WHO African Region. Meeting held in Johannesburg, South Africa


Copyright: © 2018 Olawale Basheer Akanbi. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.