DEPARTMENT OF STATISTICAL SCIENCES, UNIVERSITY OF TORONTO

Survival Analysis on Accidental Deaths According to Handedness Based on Bayesian Inference

Jiaqi Bi

November 19, 2021

ABSTRACT This report analyzes the left-handed people's accidental deaths comparing to right-handed people. These two groups have different rate of accidental deaths that left handed people tend to have higher rate of being accidental dead with the time increasing than right handed people. Using the method of survival analysis comparing censoring data and accidental deaths data illustrates the difference of the hazard.

1 INTRODUCTION

With the society developing rapidly, people tend to regard their lifetimes as important as older days. More researches on how daily habits and life style will affect people's life are substantially increasing. This report will illustrate a discussion on how left-handed people's lifespans are influenced by accidental deaths comparing to the right-handed with the data sourced from "Lifespans of UK 1st class cricketers born 1840-1960". We will analyze the data using survival analysis to see the hazards of the left-handed people with Bayesian inference including censoring. The data visualization may not show a significant difference between two groups as Figure 1.1 shows that, their lifespans in a boxplot are relatively identical. Therefore, we are going to use survival analysis to expand our data to see the hazard between two groups.

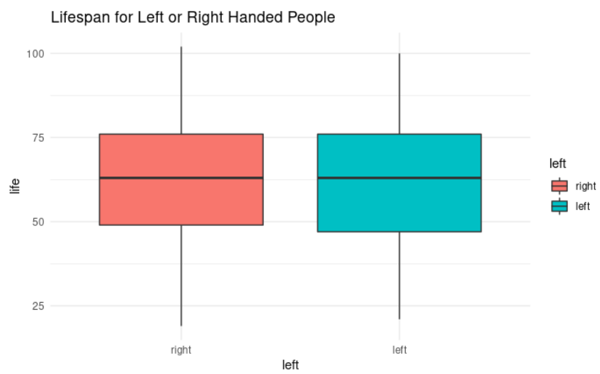

Figure 1.1: Lifespan of Left and Right Handed People

2 METHODS

2.1 SURVIVAL ANALYSIS ON POPULATION OF ACCIDENTAL DEATHS

Generally, survival analysis involves Weibull distribution. The data for lifetimes are usually left-skewed with the mean lies on around 70, and there will be very few deaths below 40 or above 110. Therefore, we choose the prior for α to be between 5 and 10 as shown in Figure 2.1, so we choose 7.5 with the standard deviation of 2/3.

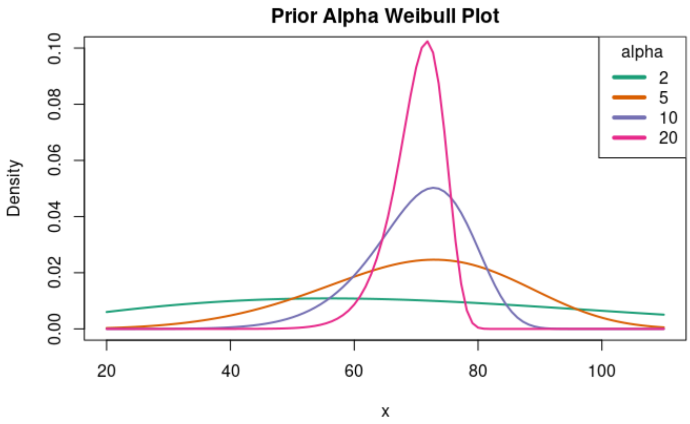

Figure 2.1: Prior Alpha

So, we will use log-Normal(log(7.5), 2/3) to be our prior. Hence, we can construct our model of the Weibull regression using Bayesian inference with the population of accidental deaths:

$$Y_i \sim \text{Weibull}(\lambda_i, \alpha) \tag{2.1}$$

where Y_i is the response of lifetime of the individual i following Weibull distribution, such that

$$\lambda_i = \exp(-\eta_i) \tag{2.2}$$

with the scale parameter λ_i influenced by predictors, where the predictors are linear combination representing as η_i,

$$\eta_i = X_i\beta \tag{2.3}$$

An increase in X_i by one unit is associated with an increase in hazard by $\exp(\beta)$ times. X_i represents the time period for the individual i.

As Table 2.1 shows, the positive value of mean on left-handed people represent the time scale for dying is 7.6% smaller than right-handed, that is, they age 7.6% more quickly. Using the proportional hazards model gives us

$$h_{right}(t) = 0.076 h_{left}(t) \tag{2.4}$$

Table 2.1: Parameter Rate Ratio for Accidental Deaths without Censoring

	Mean	Standard Deviation	2.5% Quantile	97.5% Quantile
Intercept	0.613	0.180	0.239	0.947
Decade	0.131	0.060	0.018	0.254
Left-Handed	0.076	0.135	-0.198	0.336
Alpha Parameter	1.727	0.110	1.521	1.953

People who are alive when the data is collected are considered as censoring as their lifetimes are unknown, which we cannot see their hazards using only data containing who are labeled as accidental deaths.

2.2 Survival Analysis on Accidental Deaths with Censoring

Now, we are adding the data for those who are not dead yet as censoring including those dead in bed. We consider their lifetimes unknown, so the method of censoring is applied to the survival analysis. Therefore, we are going to analyze the data using hierarchical model.

$$Z_i|Y_i, A_i = \min(Y_i, A_i) \tag{2.5}$$

$$E_i|Y_i, A_i = I(Y_i < A_i) \tag{2.6}$$

We introduced A_i that is individual i's age in 1992. Z_i is the individual's age we are going to use whether he is dead or alive. E_i includes an event indicator where $(Y_i < A_i)$ that Y_i may be unknown for this individual i. Without loss of generality, we have Y_i still follows Weibull distribution as lifetimes are typically left-skewed:

$$Y_i \sim \text{Weibull}(\exp(\eta_i), \kappa) \tag{2.7}$$

$$\eta_i = X_i\beta \tag{2.8}$$

Note that the modelling for lifetimes response variable does not change from the model without censoring such that larger β means η_i is larger and $\exp(-\eta_i)$ is smaller. This indicates smaller λ_i which means the death happens more quickly. When $E_i = 1$ and it represents a dead individual. When $E_i = 0$ then $Z_i < Y_i < \infty$ and it represents an alive individual with unknown lifetimes. We now have a likelihood function:

$$L(\mathbf{Z}, \mathbf{E}; \beta, \kappa) = \prod_{i,E_i=1} f(Z_i) \prod_{i;E_i=0} \int_{Z_i}^{\infty} f(u)du \tag{2.9}$$

with the unknown Y_i is integrated out to be the probability density of lifetime.

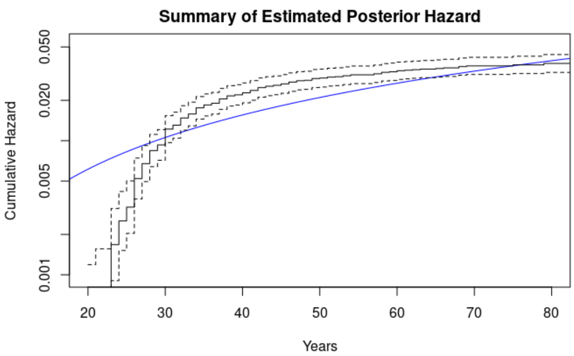

Figure 2.2: Posterior Hazard Summary

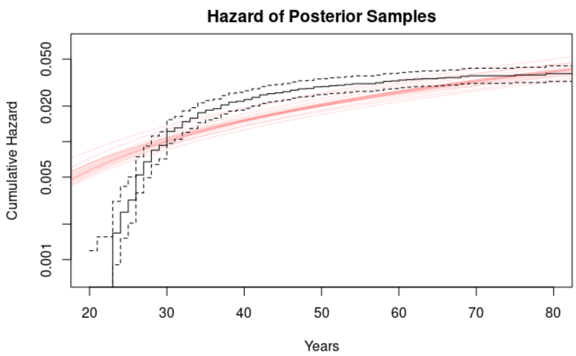

Figure 2.3: Posterior Samples Hazard

The cumulative hazard increases with the age increases as shown in Figure 2.2, which means the more likely to die at an accident with the age increases. The estimate seems reasonable as the sample of posterior does not expand the uncertainty too much and within a relatively legit interval as shown in Figure 2.3. Thus, with the censoring data included in the analysis, we can draw conclusions based on the survival analysis and the rate ratio of each parameter.

3 Results

As Table 3.1 shows the information about parameter rate ratio for accidental deaths with censoring data, we conclude left handed people live 32.2% shorter with the 95% confidence interval of 7.3% to 56.6% than right handed people. That is, left handed people have a Weibull scale parameter which is 32.3% smaller. The negative value at decade parameter represents that being born a decade later increases the lifetime by 56.2% with 95% confidence interval of 20.8% to 92.8%. We can also see the relative lifetime rate ratio comparisons between accidental deaths data and all data shown in Figure 3.1, that left handed people have higher rate of being accidental dead comparing to right handed people.

Table 3.1: Parameter Rate Ratio for Accidental Deaths with Censoring

	Mean	2.5% Quantile	97.5% Quantile
Intercept	-2.195	-2.490	-1.926
Decade	-0.562	-0.929	-0.208
Left-Handed	0.322	0.073	0.566
Alpha Parameter	1.348	1.221	1.483

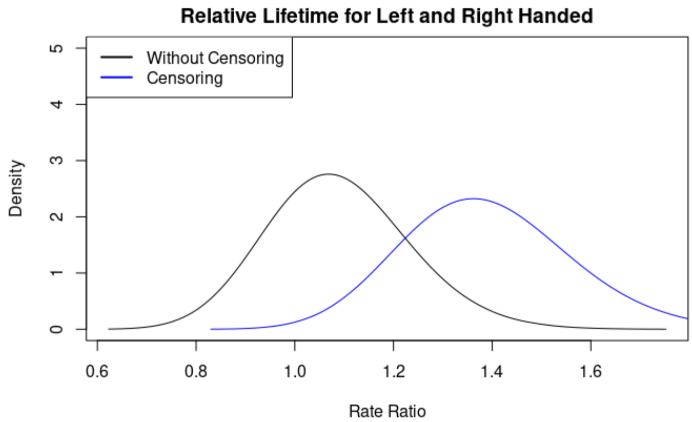

Figure 3.1: Relative Lifetime for Left vs. Right Handers

Thus, we conclude that left handed people tend to have higher accidental death rate than right handed people despite adding censoring to the data or not. Left handed people have the mean of 32.2% shorter lifetimes than right handed people.