Predicting the Lifespan of World Leaders

Cathy Lee, Alice Liao, Ashley Murray, and Matty Pahren

8-31-20

I. Introduction

World leaders are influential people that have the power to impact millions of people's lives. Some leaders are elected for terms while others serve for life, so their expected survival is of great interest. In this paper, we seek to examine Popes, U.S. Presidents, Dalai Lamas, Chinese Emperors, and Japanese Emperors to see how their lifespans compare. Additionally, we want to analyze the impact of a leader's birth year on survival, and whether or not this impact might vary depending on the type of leadership. After we build a model to analyze the effects, we also wish to obtain predictions for the lifespans of leaders that are currently alive, as well as make comparative statements about the survival times of these leaders. We will accomplish this by using survival analysis. We use Bayesian Inference to estimate the model parameters with the help of the JAGS program.

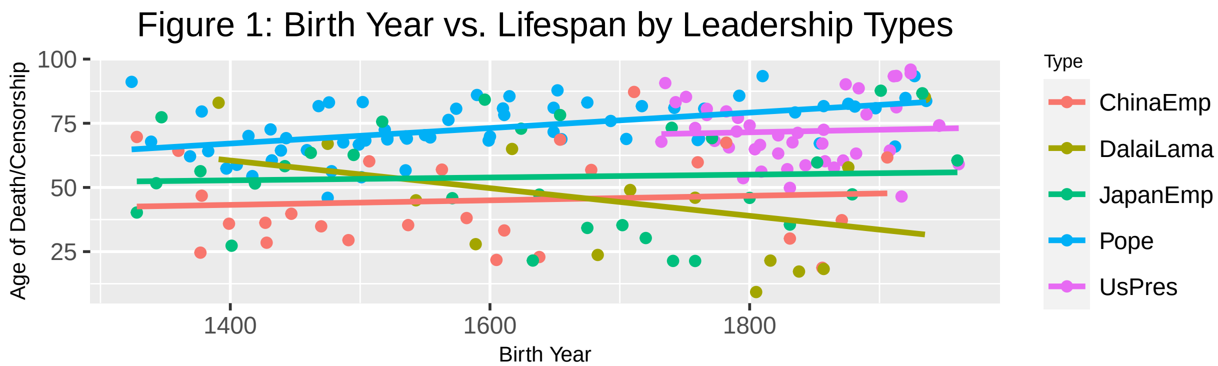
The rest of the paper is structured as follows. First, we will describe the data used to carry out our analysis. Then, we will discuss the Weibull model and Bayesian framework used to ground our analysis plan. Then, we will show how we carried out our analysis plan and recount our results. After that, we discuss our conclusions and areas for further research. Finally, our code and some additional information can be found in the Appendix.

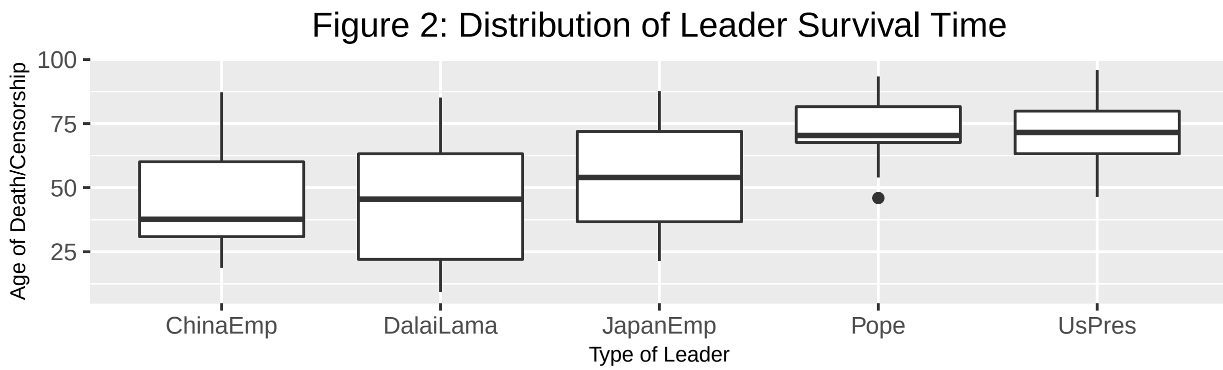
II. Data

II a. Description of the Data

The data used in this analysis contains entries for 177 different world leaders. The types of world leaders present in the data include Popes, US Presidents, Dalai Lamas, Chinese Emperors, and Japanese Emperors. Each individual's birth date and their leadership position is recorded. We calculated each leader's birth year from the birth date column. For some of these groups, we have data dating all the way back to the 14th century. Additionally, leaders who have passed away also have their death date and age of death recorded. For leaders that are still living, these columns instead contain the date the dataset was created (July 31, 2020) and their current age on that date. We updated that date to the due date of the report, August 31, 2020. In order to further clarify who is dead or alive, there is a column titled "Censored," which takes a value of 0 if the person is dead and a value of 1 if that person is still alive. Related to this, there is another column called "Fail," which takes on a value of 1 if the person is dead and a value of 0 if the person is alive. The problem with censored data is that we don't know exactly when a person will die. Thus, we have to come up with a way to model a censored person's death date if we use them in building our model. In total, there are 10 living leaders in our dataset, whose ages we are trying to predict.

II b. Exploratory Data Analysis





Based on Figure 1, it appears that birth year's association with lifespan does depend on type of leadership. In particular, Dalai Lamas' lifespans are negatively correlated with birth year, whereas the rest of the leader types have positive correlations with birth year. From the boxplot (Figure 2), we see that the distribution of the lifespans of Popes and US Presidents is centered higher, which makes intuitive sense given that these leaders are elected later in life, while Dalai Lamas and Emperors can be elected at much younger ages. For instance, if a person is a toddler and dies at age five, that person can never be president (and therefore would not have their age be part of the dataset). However, a person can become a Dalai Lama as a toddler and have their age at death be a lot smaller.

These same patterns can also be observed in the histograms in Figure 3. Also, in the histograms, there is no extreme skewing, so it does not appear that transformations are needed.

III. Methods

III a. Motivating the Model

We use survival analysis as a way to model T_i, the lifespan of a given leader i depending on the leader's year of birth and type of leadership. Survival analysis is useful in that it allows the consideration of "censored" data. This means that we do not always observe the outcome for each data point. For example, we do not know the death date of leaders that are still alive. Thus, we do not know their lifespans, we just know that their survival time T_i will be greater than their current age.

We model the lifespan T_i of an individual i after the Weibull distribution specified below. We choose to use the Weibull distribution because it is often utilized in survival analysis and allows the user to specify a flexible shape parameter for the distribution. The first parameter r, also known as the shape parameter, is a positive scalar, and the second parameter μ, the scale parameter, is a linear function of the covariates (in this case, the century of the birth year and types of leadership as well as their interactions).

$$T_i \sim Weibull(r, \mu_i)$$

$$\log(\mu_i) = \beta_0 + \sum_{j=1,\cdots,6} \beta_j I(Birth\ Century_i = j) + \sum_{k=7,\cdots,10} \beta_k I(Leadership_i = k)$$

$$+ \sum_{j=4,\cdots,6,\ k=7} \beta_{j,k} I(Birth\ Century_i = j) * I(Leadership_i = k)$$

$$+ \sum_{j=1,\cdots,6,\ k=8,\cdots,10} \beta_{j,k} I(Birth\ Century_i = j) * I(Leadership_i = k)$$

where the indicator functions $Birth\ Century_i = 1, \cdots, 6$ corresponds to a leader i born in the 15th, 16th, 17th, 18th, 19th and 20th century respectively. The 14th century is the baseline for comparison. $Leadership_i = 7, \cdots, 10$ corresponds to a leader i being a U.S. President, a Chinese Emperor, a Dalai Lamai and a Japanese Emperor respectively. $Leadership_i$ uses Pope as the baseline for comparison. $\beta_{j,k}$ is the parameter for the interaction effect between brith year century j and leadership type k.

For example, for Pope Francis who was born in 1963, his birth year century is the 20th century and his leadership type is Pope. $I(Birth\ Century_{Francis} = j)$ will all be 0 except for $j = 6$ (which indicates he was born in the 20th century) and $I(Leadership_i = k)$ will all be 0 since Pope is the baseline for comparison for leadership types. We can write $\log(\mu_{Francis}) = \beta_0 + \beta_6$ for Pope Francis.

Note that the first U.S. President, George Washington, was born in 1732 (18th century, $j = 4$) and hence the interactions between $Birth\ Century_i = 1, 2, 3$ (born in the 15th century, 16th century or 17th century) and $Leadership_7$ (being a U.S. President) are meaningless. As such, we didn't include those interaction terms in our model.

Prior to using this model, we considered a model that used leaders' birth years as a continuous variable instead of binning them into centuries. However, that model resulted in a less desirable model predictive ability as we ran model diagnostics. The residual plot showed a clear trend between residuals and predicted lifespans. Using a quadratic polynomial improved the residuals, but our current model is still superior. Model diagnostics of the current model can be found in section V.c and more details on previous models can be found in the Appendix A.2. *make sure we include them in the Appendix*

III b. Addressing Censored Data

To handle censored observations, we specify their contribution to the likelihood function using the Poisson "zeros trick" and set the upper limit of age of death to 100. Details on this method and how to implement it can be found in Appendix A.1.

III c. Prior Choice

We assume that our priors are independent because intuitively a given observation could not be two types of leaders. Additionally, we do not have sufficiently strong prior knowledge about the relationship between birth year and type of leadership to specify an informative prior. Therefore, we use uninformative priors for the betas in our model, in which we chose normal betas with mean 0 and precision 1.0E-3, this includes the coefficients and the intercept in our model.

For our prior for r, we chose a prior of $r \sim exp(0.1)$, as we believe the hazard of death increases with time. This is supported by the fact that older people have higher risk of mortality. For example, in an article by Eurostat, in the European Union in 2016, 82.9% of all deaths happened among people over 65. Within the Appendix A.5, different Weibull distributions are plotted with different values of r in order to showcase the fact that $r \sim exp(0.1)$ gives an increasing rate of hazard, as the shape of the Weibull Distribution starts to have an increasing level of density toward the upper x values for $r \sim exp(0.1)$, compared to smaller values of r.

IV. Analysis

As Stander et al (2018) pointed out, following this Weibull distribution, $\log(T) = \frac{1}{r}(\mu) + \frac{1}{r}(\epsilon)$, or

$$\log(T) \sim -\frac{1}{r}(\beta_0 + \beta_1 x_1 + \cdots + \beta_6 x_6 + \beta_7 x_7 + \cdots \beta_{10} x_{10} + \beta_{1,8} x_1 x_8 + \cdots + \beta_{6,10} x_6 x_{10}) + \frac{1}{r}\log(\epsilon)$$

$$= \alpha_0 + \alpha_1 x_1 + \cdots + \alpha_6 x_6 + \alpha_7 x_7 + \cdots \alpha_{10} x_{10} + \alpha_{1,8} x_1 x_8 + \cdots + \alpha_{6,10} x_6 x_{10}) + \frac{1}{r}\log(\epsilon)$$

where $\epsilon \sim exp(1)$ and $\epsilon > 0$. $\alpha = -\beta/r$ is thus a monotone transformation of β. $x_1, \cdots, x_6$ are indicator values (i.e., 0s and 1s) corresponding to birth year's century, and $x_7, \cdots, x_{10}$ are indicator values corresponding to types of leadership.

The interpretation of coefficients depends on the interaction terms (i.e. both birth year century and the types of leadership). For example, while keeping all else constant, if a Pope was born in the 20th century, he is expected to live longer by a multiplicative factor of $exp(\alpha_6)$ compared to a Pope born in the 14th century, or his lifespan is expected to increase by a percentage of $100 * exp(\alpha_6 - 1)$. If a U.S. President was born in the 18th century, he is expected to live $exp(\alpha_0 + \alpha_4 + \alpha_7 + \alpha_{4,7})$ years; if the president were born in the 20th century instead, he is expected to live $exp(\alpha_0 + \alpha_6 + \alpha_7 + \alpha_{6,7})$ years. In other words, his lifespan is expected to increase by a multiplicative factor of $exp(\alpha_6 - \alpha_4 + \alpha_{6,7} - \alpha_{4,7})$. If a Chinese Emperor and a Dalai Lama were both born in the 19th century, the Chinese emperor is expected to live longer by a multiplicative factor of $exp(\alpha_8 - \alpha_9 + \alpha_{5,8} - \alpha_{5,9})$ and so on.

V. Results

V a. Output and Interpretation of the Model

The table below shows the model output of the multiplicative factors, αs. Note that since $\alpha = -\beta/r$ is a monotone transformation, if β is statistically significant (i.e. 95% credible interval doesn't contain 0), it should be the same for the corresponding α.

Since the baseline for comparison is a Pope born in the 14th century, it's meaningful to interpret the intercept in this case. The intercept shows that we expect a Pope born in the 14th century to live $\exp(\alpha_0) = \exp(4.3)$ = 75.1 years. $\alpha_{6,9}$ which corresponds to the interaction between born in the 20th century and being a Dalai Lama has an abnormally large point estimate, largely because there is only one Dalai Lamai born the last century and he is a censored data point.

While $\alpha_1, \cdots, \alpha_5$ have 95% credible intervals containing 0 (an indication that these coefficients are statistically insignificant), α_6 has a positive 95% credible interval. Hence birth year overall is an significant predictor in this model. If we only look at the main effect of leadership types ($\alpha_7, \cdots, \alpha_{10}$), these parameters' 95% credible intervals all contain 0. However, some of their interactions with birth year are statistically significant, for example, $\alpha_{1,8}$'s CI is always negative. Due to model hierarchy, we need to keep main effects of leadership types in the model as well.

The specific interpretation of these parameters depends on inputs from birth year and type of leadership. Below are some examples.

- A Japanese emperor born in different centuries
  - 14th century: $\log(T_{14th,JapanEmp}) = \alpha_0 + \alpha_{10} + \frac{1}{r}\log(\epsilon)$
  - 15th century: $\log(T_{15th,JapanEmp}) = \alpha_0 + \alpha_1 + \alpha_{10} + \alpha_{1,10} + \frac{1}{r}\log(\epsilon)$
  - 16th century: $\log(T_{16th,JapanEmp}) = \alpha_0 + \alpha_2 + \alpha_{10} + \alpha_{2,10} + \frac{1}{r}\log(\epsilon)$
  - 17th century: $\log(T_{17th,JapanEmp}) = \alpha_0 + \alpha_3 + \alpha_{10} + \alpha_{3,10} + \frac{1}{r}\log(\epsilon)$
  - 18th century: $\log(T_{18th,JapanEmp}) = \alpha_0 + \alpha_4 + \alpha_{10} + \alpha_{4,10} + \frac{1}{r}\log(\epsilon)$
  - 19th century: $\log(T_{19th,JapanEmp}) = \alpha_0 + \alpha_5 + \alpha_{10} + \alpha_{5,10} + \frac{1}{r}\log(\epsilon)$
  - 20th century: $\log(T_{20th,JapanEmp}) = \alpha_0 + \alpha_6 + \alpha_{10} + \alpha_{6,10} + \frac{1}{r}\log(\epsilon)$

The summary table is shown below if we compare the expected lifespans of later Japanese emperors with that of a Japanese emperor born in the 14th century.

Japanese Emperor	Multiplicative Factor	% Increase / Decrease in Lifespan
Born in 14th Century	1.00	0
Born in 15th Century	0.91	-9
Born in 16th Century	1.20	20
Born in 17th Century	1.02	2
Born in 18th Century	0.89	-11
Born in 19th Century	0.83	-17
Born in 20th Century	1.75	75

Surprisingly, the model suggests that longevity doesn't always increase with calendar time. The biggest increase in lifespan is in the 20th century, where the expected lifespan of a Japanese emperor increases by 75% compared to a Japanese emperor born 6 centuries ago. This might make sense, given the last century saw the biggest number of technological breakthroughs and improvement in healthcare.

- Dalai Lamai born in different centuries

We can also compare how different Dalai Lamas born in different centuries are expected to live. The percentage increase in lifespan becomes generally more negative, which means later born Dalai Lamas are expected to live shorter, confirming the negative correlation between lifespan and birth year exihibited in Figure 1 of section II.b EDA. The Dalai Lama born in the 20th century, however, is expected to live abnormally longer than the 1st Dalai Lama born in the 14th century, due to the extremely large point estimate of $\alpha_{6,9}$ that corresponds to the interaction between born in the 20th century and being a Dalai Lama. As explained earlier, this might be because there is only one Dalai Lama born in the 20th century and he is censored.

Dalai Lama	Multiplicative Factor	% Increase / Decrease in Lifespan
Born in 14th Century	1.00	0
Born in 15th Century	0.80	-20
Born in 16th Century	0.45	-55
Born in 17th Century	0.64	-36
Born in 18th Century	0.53	-47
Born in 19th Century	0.46	-54
Born in 20th Century	132.96	13196

- Different leaders born in the 18th and 19th centuries

Leaders born in 18th Century	Multiplicative Factor	% Increase / Decrease in Lifespan
Pope	1.00	0
U.S. President	0.98	-2
Chinese Emperor	0.99	-1
Dalai Lama	0.63	-37
Japanese Emperor	0.73	-27

Leaders born in 19th Century	Multiplicative Factor	% Increase / Decrease in Lifespan
Pope	1.00	0
U.S. President	0.84	-16
Chinese Emperor	0.40	-60
Dalai Lama	0.52	-48

Leaders born in 19th Century	Multiplicative Factor	% Increase / Decrease in Lifespan
Japanese Emperor	0.64	-36

Popes born in the 18th and 19th centuries are expected to live longer than other leaders born in the same century. If born in the 18th century, a U.S. president and a Chinese emperor are expected to enjoy comparable lifespans as a Pope. However, if born in the 19th century instead, the expected lifespans of a U.S. president and a Chinese emperor are significantly shorter than that of a Pope (16% and 60% shorter respectively).

The table below shows the estimate and 95% credible interval (the bounds are the 2.5% quantile and 97.5% quantile) for all 10 living leaders in the dataset.

Variable	Mean	Standard Deviation	2.5% Quantile	Median	97.5% Quantile
age_Francis_predictive	94.1	3.8	87.1	94.3	99.8
age_Obama_predictive	96.6	1.9	93.6	96.6	99.8
age_Dalai_predictive	87.3	7.3	74.9	87.3	99.3
age_Naruhito_predictive	97.9	1.2	96.0	97.9	99.9
age_Benedict_predictive	87.4	7.3	74.8	87.7	99.2
age_Carter_predictive	92.7	4.3	85.5	92.8	99.6
age_Clinton_predictive	91.9	4.6	84.1	91.9	99.5
age_Bush_predictive	86.4	10.4	63.3	88.8	99.6
age_Trump_predictive	81.9	11.0	61.3	82.8	99.1
age_Akihito_predictive	87.3	7.3	74.9	87.3	99.4

V b. Posterior Inference for Selected Leaders

Figure 3: Posterior Distribution Density Plots

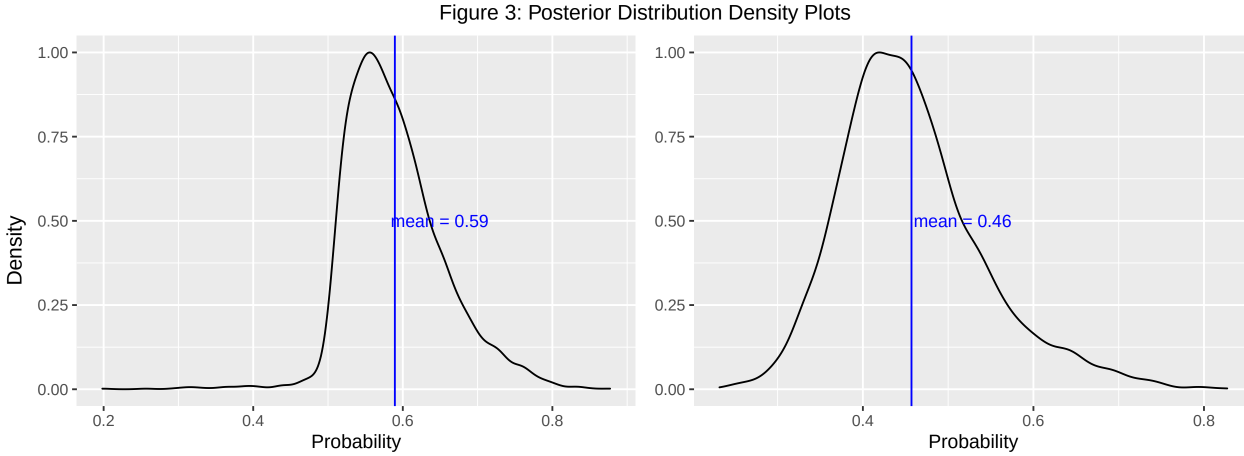


Fig. 3a This density plot shows the posterior distribution of the probability that the 14th Dalai Lama outlives Pope Francis.

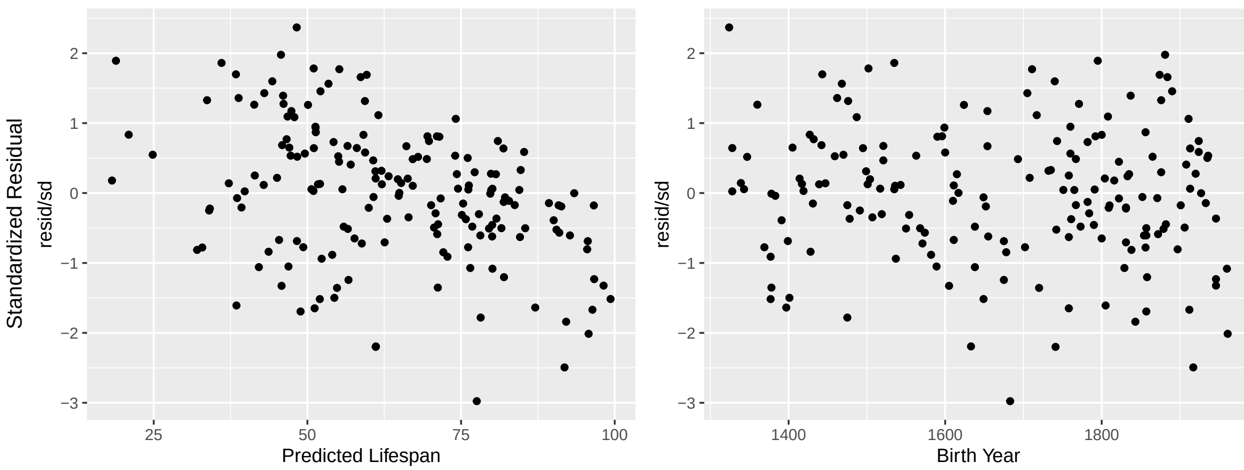
Fig. 3b This density plot shows the posterior distribution of the probability that Obama outlives Emperor Naruhito.

Figure 4 depicts the posterior distribution of the probability that Dalai Lama outlives Pope Francis, and that Obama outlives Emperor Naruhito. The probability that the 14th Dalai Lama will have a longer lifespan than Pope Francis is 0.59 with a 95% confidence interval of (0.5, 0.74). The probability that President Obama will have a longer lifespan than Emperor Naruhito is 0.46 with a 95% confidence interval of (0.32, 0.66). For histograms on the posterior predictive distribution of lifespans for the 4 leaders, see Appendix A.6.

V c. MCMC and Model Diagnostics

The combination of traceplots, lag-1 scatterplots, and acf plots suggest the chain for each parameters converges. Code for those plots are included in Appendix A.2. The Rhat's are all close to 1, which is another indicator of convergence. Most of the effective sample sizes are greater than 1000 (the effective sample size of 400 for r is lower, but the number of iterations was kept at 50,000 for computational efficiency).

Figure 4: Residual Plots



The standardized residuals plotted against birth year are randomly scattered, showing that our model doesn't systematically over- or under-predict lifespan based on birth year. The standardized residuals vs. predicted lifespan are all still fairly randomly scattered, however there might be some trending for predicted lifespans greater than 75. We tried several different model fits, including using age as a continuous variable and using polynomial transformations, and our choice of using binned age as opposed to one of these transformations gave us the best residual plot. Some of this trending that is left over might be because we do not have enough predictors in our model. It is also worth pointing out that we have a few standardized residuals above 2 and below -2, meaning we might have some outliers. However, none of them cross the 3 or -3 threshold, so none of them are extreme outliers.

In order to perform posterior predictive checks, ten simulated datasets were created by first extracting the leaders that were non-censored. There were ten different iterations that performed the same task: every single row of the non-censored leaders and a random sample of the model parameters (from the MCMC chain) was used to define a set of μ's, and those μ's were used to simulate a set of lifespans from the Weibull distribution.

Figure 5: Posterior Predictive Checks Density Plot

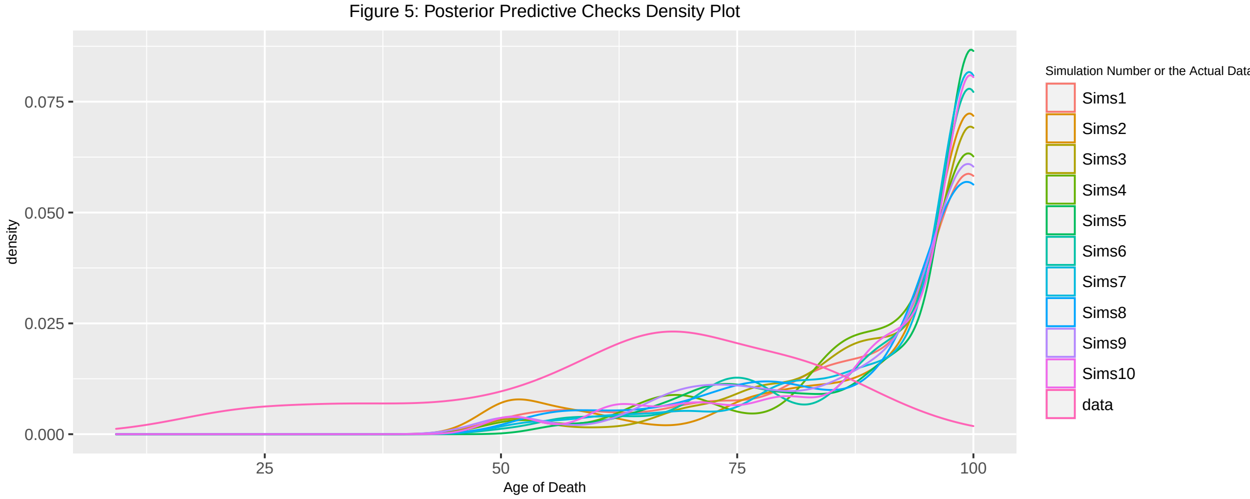
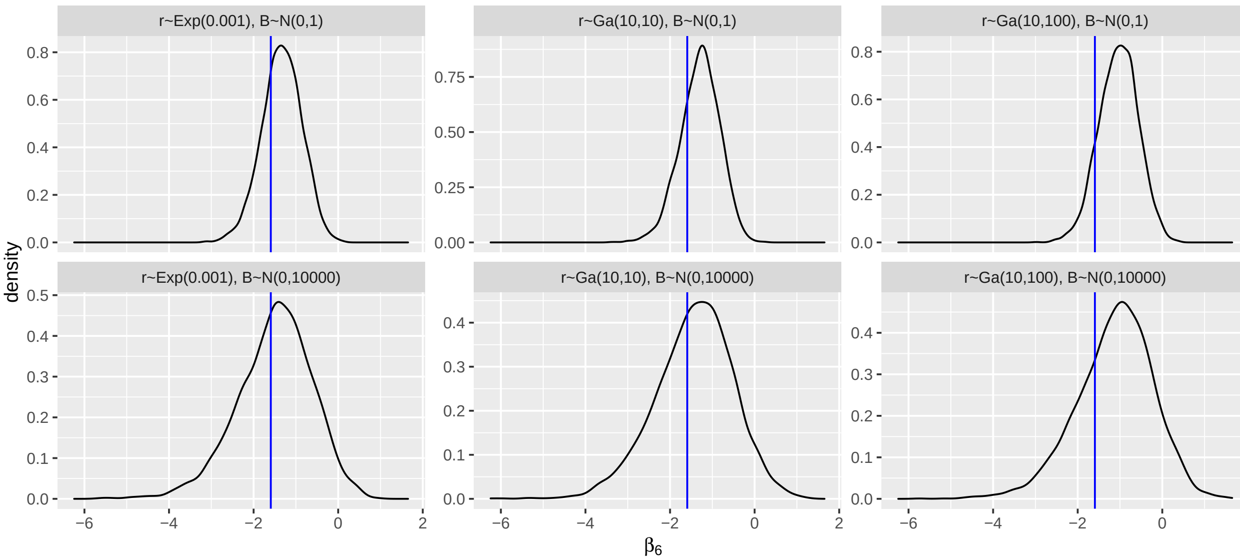


Figure 6 showcases the curvatures of the ten different simulations done in the posterior predictive checks and how they relate to the actual age of the non-censored individuals. There appears to be a peak within the non-censored individuals at around 70, but most of the simulations are seeing a peak at near 100. For the posterior predictive checks, it appears that our model returns values that are more concentrated around the upper truncation limit, which is why the dataframe's Inf values are encoded as 100 (the upper truncation limit defined in our function). This follows what the standardized residuals found, where some of the values from the model are left-skewed, and the posterior predictive checks is showing that a lot of the leaders are being predicted to having a larger lifespan.

VI. Sensitivity Analysis

Figure 6: Sensitivity Analysis for Different Priors



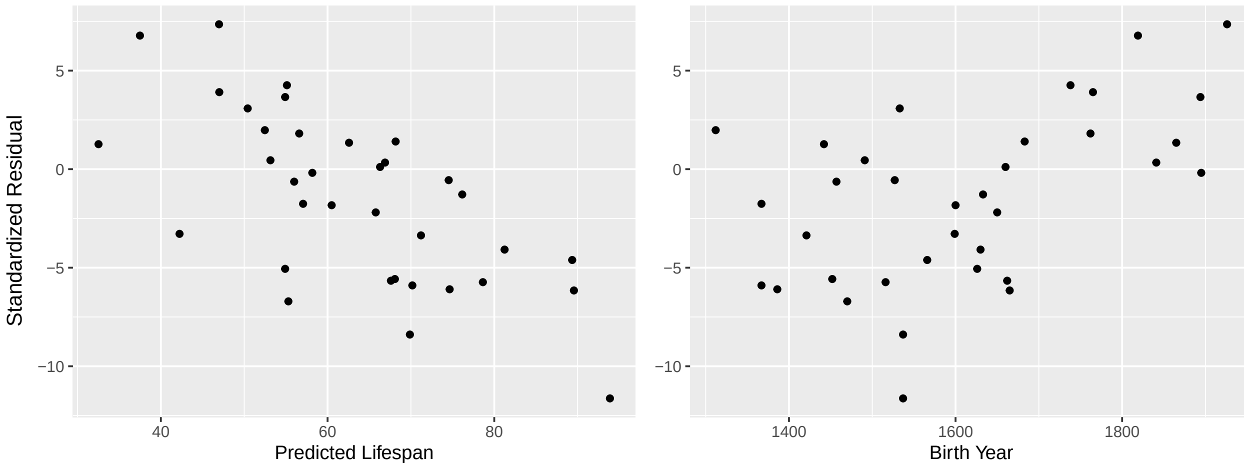
The blue line indicates the estimate of the coefficient from our model. It appears that the analysis is generally not sensitive to different priors.

We tested $N(0, 1)$ and $N(0, 10000)$ priors for the βs and $Exp(0.001)$ (Stander et al prior), $Gamma(10, 10)$ (mean = 1), and $Gamma(10, 100)$ (mean = 0.1) priors for r. Distribution of β_6 was approximated using those different priors and plotted above. The blue line, which is the estimate for β_6 from our model, crosses (or is close to) the mode of all but one of the distributions, so our analysis is mostly robust against changes in prior. However, it is somewhat sensitive to the prior of $r \sim Gamma(10, 100)$, which makes sense since that prior implies a decreasing hazard while we assumed increasing hazard. Code to conduct sensitivity analysis for the remaining betas can be found in Appendix A.3.

VII. Cross Validation

We sought to perform external cross-validation on our model to assess how useful our model would be to predict lifespans of other leaders. The data that we used for the cross-validation was from all of England's monarchs born in or after the 1300s, accessed from Wikipedia. This dataset only included one person that was still alive (Queen Elizabeth II). We assigned leader type `JapanEmp` to all British monarchs since we felt this might be the closest match, given China doesn't have any emperors after 1911.

Figure 7: Residual Plots for British Monarchs



As shown in the standardized residuals plots above, the our model didn't do a good job predicting how long British monarchs could live based on their birth year - standardized residuals are very large. In the plot showing Predicted Lifespan and the standardized residuals, it appears to be random. As the birth year of the monarchs increases, there is an upward trend. It seems that our model is overpredicting for leaders in the later centuries.

The result, however, is perhaps not surprising to see because British monarchs and Japanese emperors are different types of leaders in different cultures and geographical regions, not to mention our model itself lacks other important variables such as leaders' health conditions, access to healthcare, etc.

VI. Conclusion and Further Discussion

We sought to compare Popes, US Presidents, Dalai Lamas, Chinese Emperors, and Japanese Emperors to see how their lifespans might be explained by their birth year and types of leadership. Our model has found that both variables are significant predictors and the impact of birth year on lifespan does change based on leadership type.

One limitation worth noting is that our data does not include election year, so we have no way to control for the fact that Presidents and Popes must have a minimum lifespan. On the other hand, Dalai Lamas can be chosen close to birth and thus have a younger death date.

Additionally, there is a lack of previous existing work in the field to inform stronger prior choices, especially for the β_i coefficients.

This model could be improved by including more predictors; for example, health varies widely among people, and could lead to a better model for prediction. For further implementations of this research question, it would be valuable to interpret economic or quality of life data as it concerns the different countries that the leaders grew up in, as the conditions that a person live in lead to changes in a person's lifespan.

VII. References

Stander, J., Dalla Valle, L., and Cortina-Borja, M. (2018). A Bayesian Survival Analysis of a Historical Dataset: How Long Do Popes Live? The American Statistician 72(4):368-375.

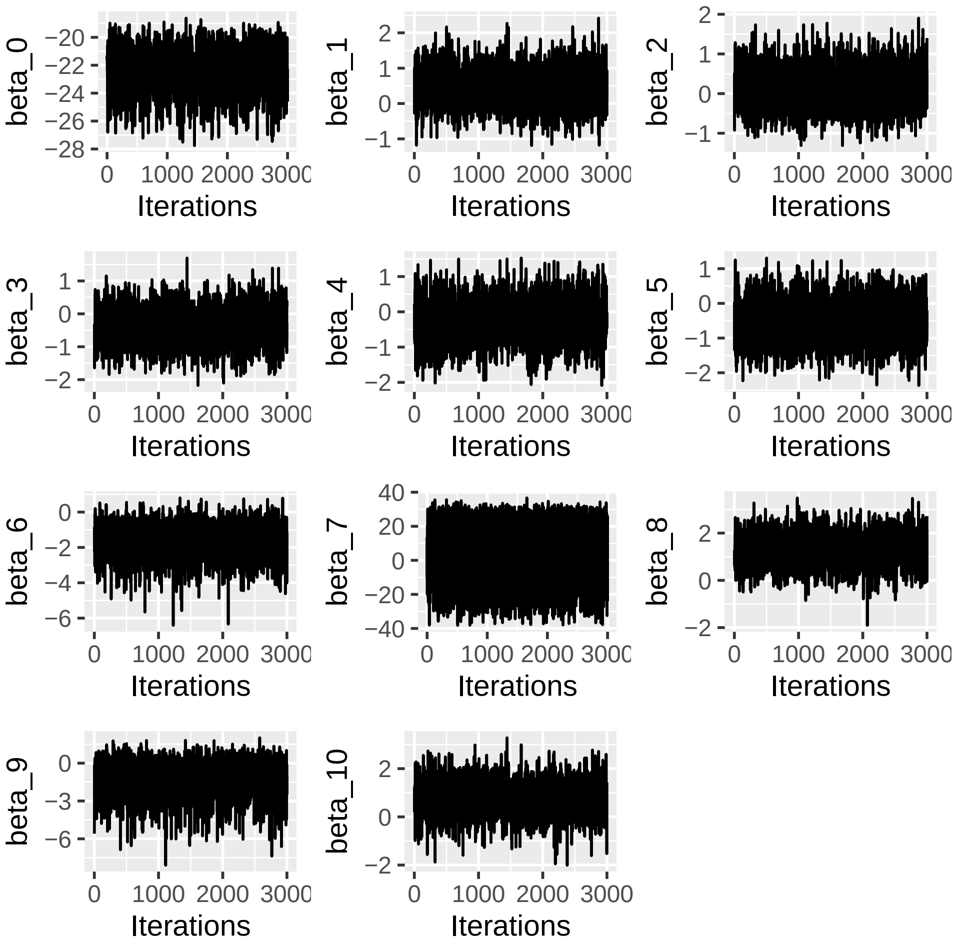
Causes of death statistics - people over 65. (n.d.). Retrieved September 08, 2020, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Causes_of_death_statistics_-_people_over_65

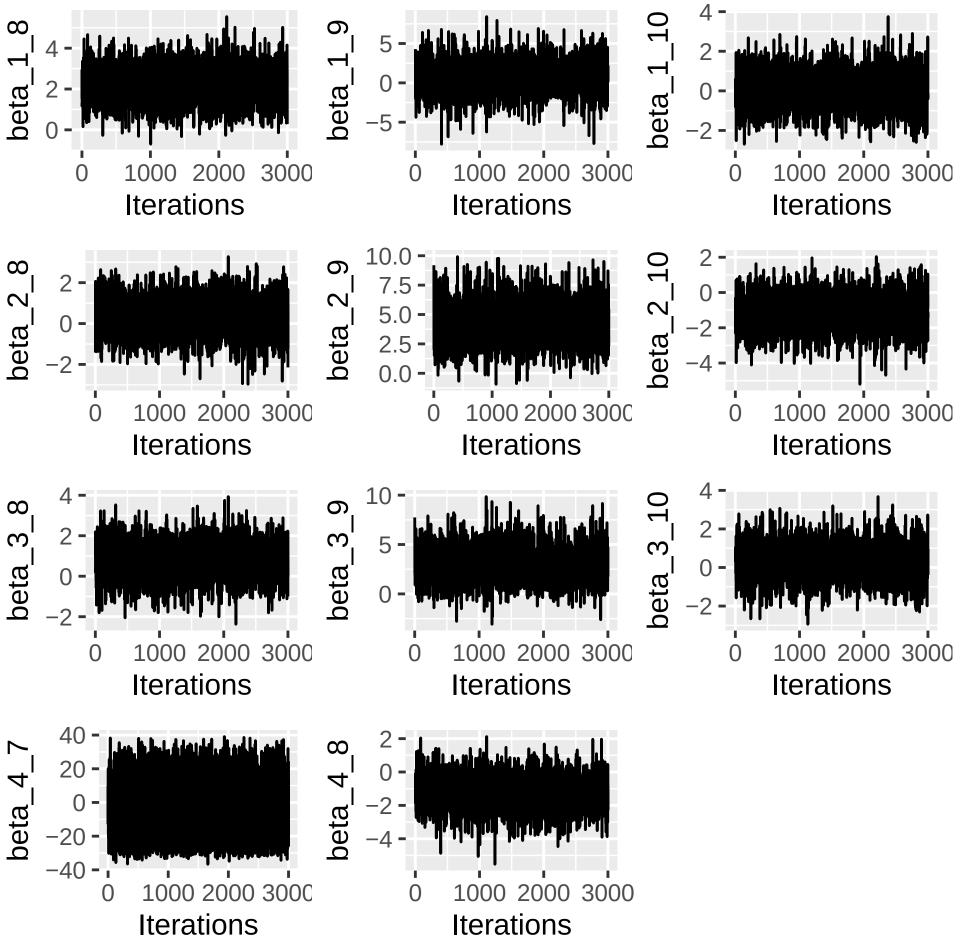
VIII. Appendix

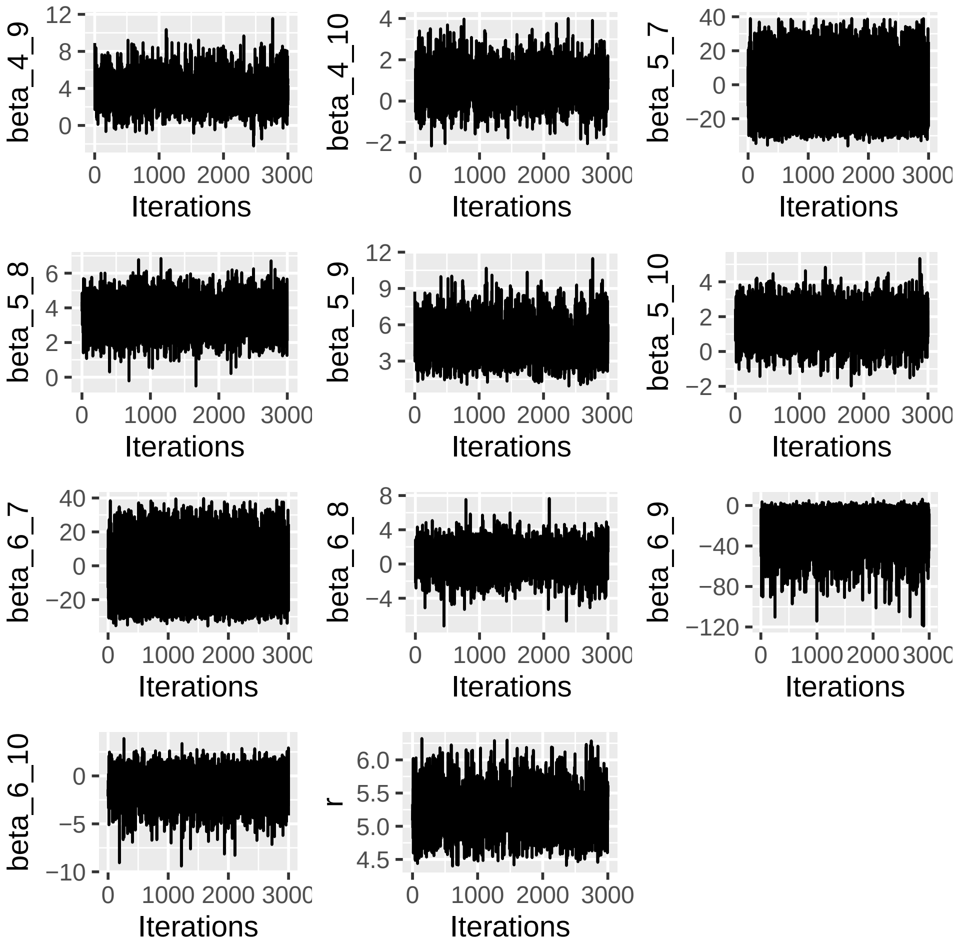
A.1 Details on the Poisson "Zero-Trick" Method

We use the Poisson "zeros trick," where a set of 0's are created and the likelihood of observing an 0 follows a $Poisson(\phi_i)$ distribution. The likelihood L_i is $Pr(z_i = 0) = exp(-\phi)$. In other words, ϕ_i equals $-log(L_i)$. This mechanism is necessary because we don't know the lifespan of currently living people, so we need a way to model their expected survival times. The censored vector has 0 for dead leaders and 1 for alive leaders. The vector censoring_limits contains 100 (chosen as an upper bound on age) for dead leaders and current age for alive leaders. The non-100 observations in censoring_limits are saved in t_censored, and t_censored is taken to the power of r then multiplied by mu_censored to calculate phi_censored, which is the mean parameter of the Poisson likelihood described above.

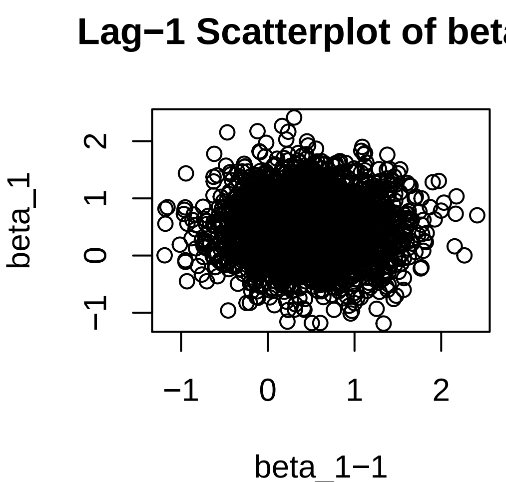
A.2 Code for Traceplots, Lag-1 Scatterplots, and Acf Plots

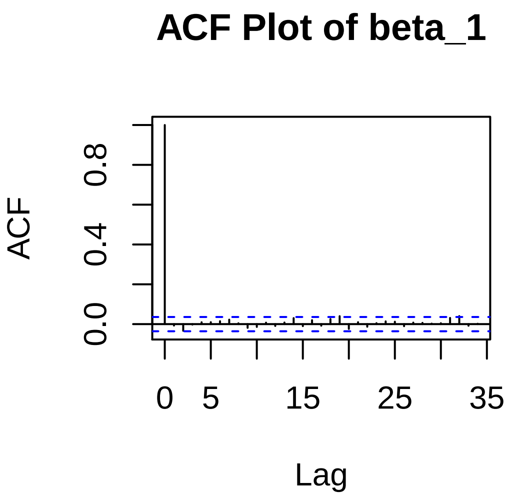



beta_1_8
beta_1_9
beta_1_10
beta_2_8
beta_2_9
beta_2_10
beta_3_8
beta_3_9
beta_3_10
beta_4_7
beta_4_8
Iterations
0
1000
2000
3000



All of the traceplots and code for plotting is above.





Display of lag-1 scatterplot and acf plot for β_1 is shown above, code provided allows generations of plots for all parameters.

Prior to arriving at our current model, we explored to other models for $\log(\mu_i)$:

1. Birth year as a continuous variable.

$$\log(\mu_i) = \beta_0 + \beta_1(Year\ of\ Birth_i) + \beta_2(Leadership_i = UsPres) + \beta_3(Leadership_i = ChinaEmp)$$
$$+\beta_4(Leadership_i = DalaiLama) + \beta_5(Leadership_i = JapanEmp)$$
$$+\beta_6(Year\ of\ Birth_i * (Leadership_i = UsPres))$$
$$+\beta_7(Year\ of\ Birth_i * (Leadership_i = ChinaEmp))$$
$$+\beta_8(Year\ of\ Birth_i * (Leadership_i = DalaiLama))$$
$$+\beta_9(Year\ of\ Birth_i * (Leadership_i = JapanEmp))$$

where $Leadership_i = Pope$ is the baseline for comparison.

2. Birth year as a continuous variable and quadratic polynomial model.

$$\log(\mu_i) = \beta_0 + \beta_1(Year\ of\ Birth_i) + \beta_2(Year\ of\ Birth_i)^2 + \beta_3(Leadership_i = UsPres)$$
$$+\beta_4(Leadership_i = ChinaEmp) + \beta_5(Leadership_i = DalaiLama) + \beta_6(Leadership_i = JapanEmp)$$
$$+\beta_7(Year\ of\ Birth_i * (Leadership_i = UsPres)) + \beta_8(Year\ of\ Birth_i * (Leadership_i = ChinaEmp))$$
$$+\beta_9(Year\ of\ Birth_i * (Leadership_i = DalaiLama)) + \beta_9(Year\ of\ Birth_i * (Leadership_i = JapanEmp))$$
$$+\beta_{10}((Year\ of\ Birth_i)^2 * (Leadership_i = UsPres)) + +\beta_{11}((Year\ of\ Birth_i)^2 * (Leadership_i = ChinaEmp))$$
$$+\beta_{12}((Year\ of\ Birth_i)^2 * (Leadership_i = DalaiLama)) + +\beta_{13}((Year\ of\ Birth_i)^2 * (Leadership_i = JapanEmp))$$

where $Leadership_i = Pope$ is the baseline for comparison.

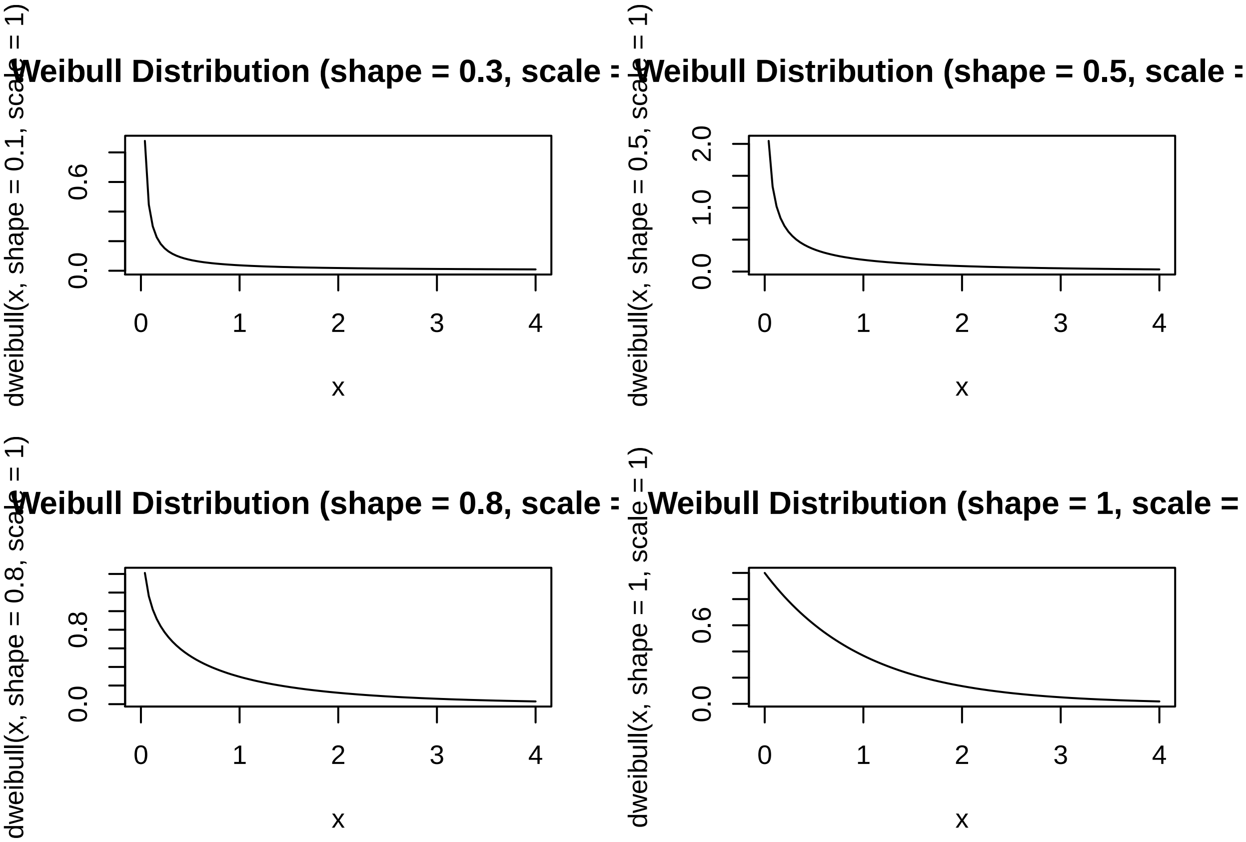
A.3 Model Output (the alphas)

Leaders born in 19th Century	Multiplicative Factor	% Increase / Decrease in Lifespan
Pope	1.00	0
U.S. President	0.84	-16
Chinese Emperor	0.40	-60
Dalai Lama	0.52	-48
Japanese Emperor	0.64	-36

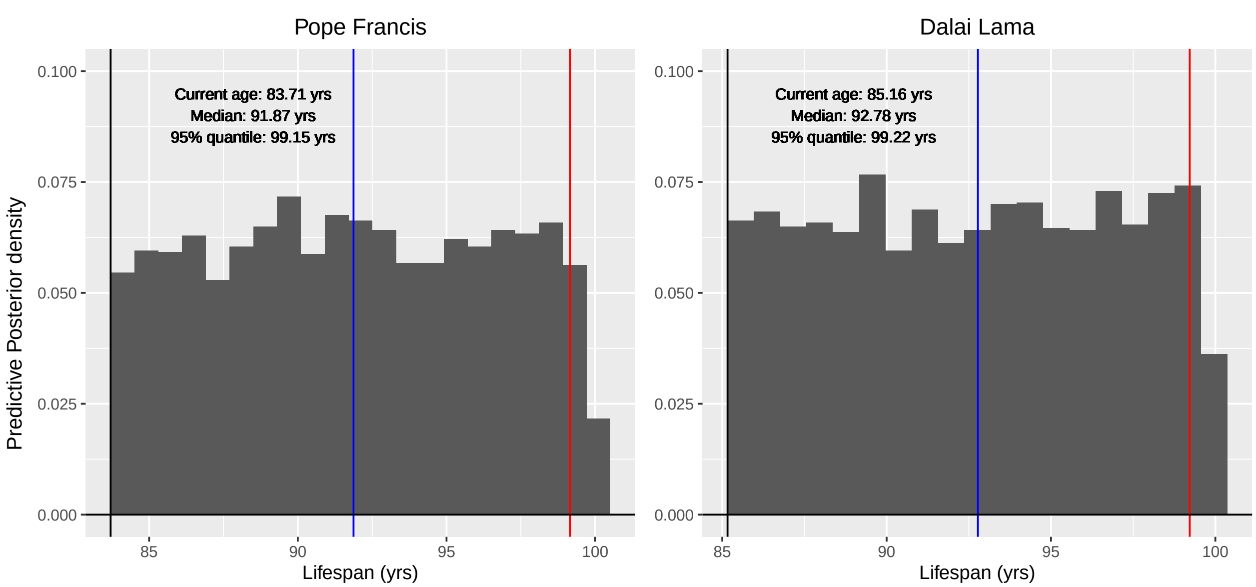
A.3 Code for Sensitivity Analysis

```
# parameter = "beta_6" # change parameter here to run for different betas
#
# sa1 = data.frame(sensitivity_1$BUGSoutput$sims.matrix) %>%
#   select(parameter)
# sa2 = data.frame(sensitivity_2$BUGSoutput$sims.matrix) %>%
#   select(parameter)
# sa3 = data.frame(sensitivity_3$BUGSoutput$sims.matrix) %>%
#   select(parameter)
# sa4 = data.frame(sensitivity_4$BUGSoutput$sims.matrix) %>%
#   select(parameter)
# sa5 = data.frame(sensitivity_5$BUGSoutput$sims.matrix) %>%
#   select(parameter)
# sa6 = data.frame(sensitivity_6$BUGSoutput$sims.matrix) %>%
#   select(parameter)
#
# sa_df = data.frame(vals = rbind(sa1, sa2, sa3, sa4, sa5, sa6),
#                    type = c(rep("r~Exp(0.001), B~N(0,1)", nrow(sa1)),
#                             rep("r~Ga(10,10), B~N(0,1)", nrow(sa2)),
#                             rep("r~Ga(10,1000), B~N(0,1)", nrow(sa3)),
#                             rep("r~Exp(0.001), B~N(0,100)", nrow(sa4)),
#                             rep("r~Ga(10,10), B~N(0,100)", nrow(sa5)),
#                             rep("r~Ga(10,1000), B~N(0,100)", nrow(sa6))))
#
# neworder <- c("r~Exp(0.001), B~N(0,1)", "r~Ga(10,10), B~N(0,1)", "r~Ga(10,1000), B~N(0,1)",
#               "r~Exp(0.001), B~N(0,100)", "r~Ga(10,10), B~N(0,100)", "r~Ga(10,1000), B~N(0,100)")
#
# sa_df_plot <- arrange(mutate(sa_df, type=factor(type,levels=neworder)), type)
#
# sampled_vals = model_output$BUGSoutput$summary[parameter, ]
# sampled_beta = sampled_vals["mean"]
#
# ggplot(sa_df_plot) + geom_density(aes(x = sa_df_plot[,1])) +
#   facet_wrap(~type,  scales = "free_y", nrow = 2) +
#   geom_vline(xintercept = sampled_beta, color = "blue") +
#   labs(x = expression(beta[6]), caption = paste("Fig. 8 The above plots show sensitivity analysis for
#   theme(plot.caption = element_text(hjust = 0.5, vjust = -0.5, size = 4))
```

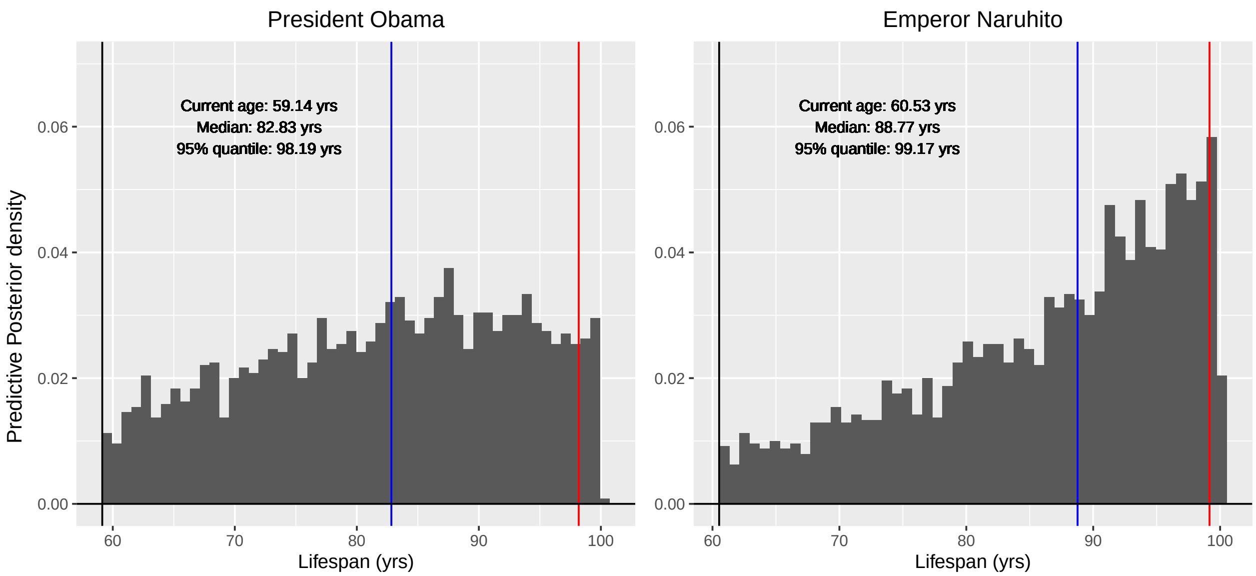
A.5 Weibull Distributions with Differing r Values



A.6 Histograms of Posterior Predictive Distribution for Lifespan of the 4 Selected Leaders



The posterior predictive probability density function of the lifespan for hypothetical leaders with the same attributes as Pope Francis (right) and the 14th Dalai Lama (left). The black vertical line marks current age the blue marks posterior median, and the red marks the posterior 95% quantile.



The posterior predictive probability density function of the lifespan for hypothetical leaders with the same attributes as Obama (right) and Emperor Naruhito (left). The black vertical line marks current age, the blue marks posterior median, and the red marks the posterior 95% quantile.

The histogram of the posterior predictive distributions of the lifespans for a leader with the same birth year and leadership type as Pope Francis and that of a leader with the same above attributes as the 14th Dalai Lama are more uniform. The histogram for a leader with the same birth year and leadership type as President Obama and that of a leader with the same above attributes as Emperor Naruhito are both left skewed, with modes around the 90s.