Bayesian Statistics Week 04

Nick Stracke & Pleun Scholten

1 Exercise 1

1.1

There are a few dependencies in the model:

- $\theta_k \rightarrow x_i$: The probability of survival for each member is dependant on the probability of survival of their family.
- $Z_i \rightarrow x_i$: The probability survival is dependent on the probability of being in family Z_k.
- $a_k, b_k \rightarrow \theta_k$: The probability of survival of the family is dependant on the hyperparameters a and b, the pseudo counts for survival and deceasing.
- $\alpha \rightarrow Z_k$: The probability of being in family Z_k is dependent on the distribution α.

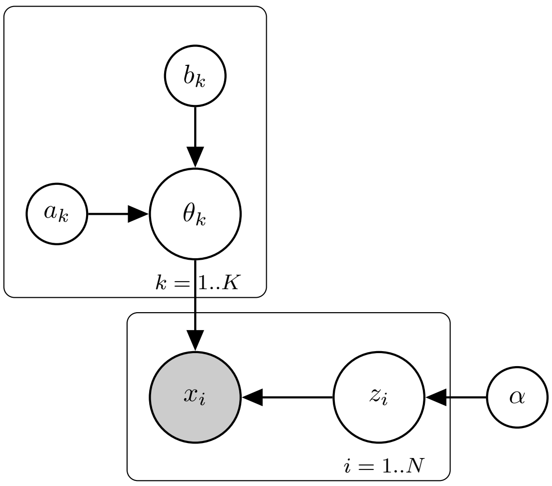


Figure 1: The model for survival.

1.2

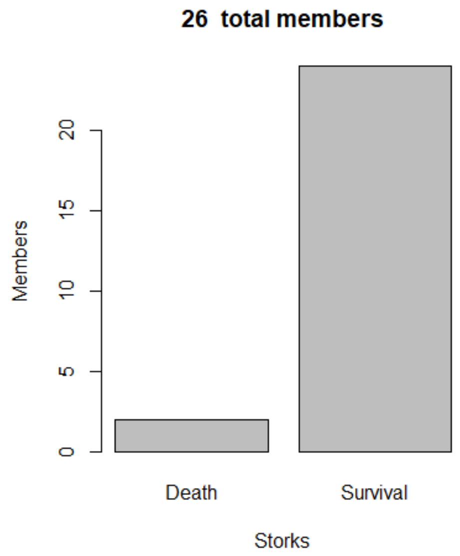
$$
\begin{aligned}
z_i|\alpha &\sim \mathrm{Cat}\left(z_i \mid \alpha\right) && \boldsymbol{\alpha} = 1/K \times \mathbf{1}_K, i = 1..N \\
\theta_k|a_k, b_k &\sim \mathrm{Beta}\left(\theta_k \mid a_k, b_k\right) && k = 1..K \\
x_i|\theta_k, z_i &\sim \mathrm{Bernoulli}\left(x_i \mid \theta_{z_i}\right) && i = 1..N
\end{aligned}
$$

Z_i is simply a categorical distribution, with parameter $\alpha = 1/K \times \mathbf{1}_K$, as described in the exercise. θ_k is dependant on the two hyperparameters a_k and b_k, and θ_k is a prior for outcome x_i, so the distribution used is a beta distribution. Because x_i is dependent on both θ_k and Z_i and because we want to get a binary outcome, we use a Bernoulli distribution.

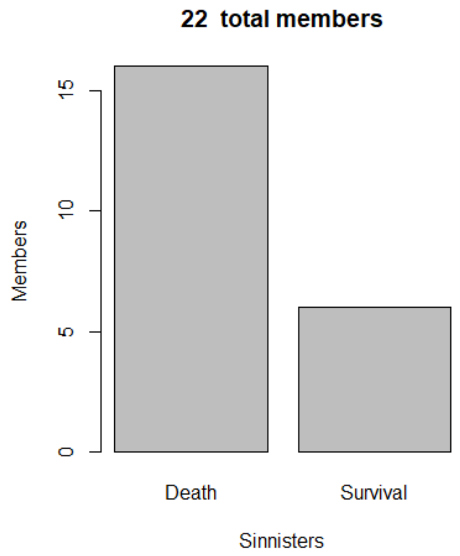
1.3

Please find attached the corresponding R script in case of questions. The way we understand the exercise is that α is 'set by the gods', i.e. uniform. Thus we can only change the hyper-parameters a and b.

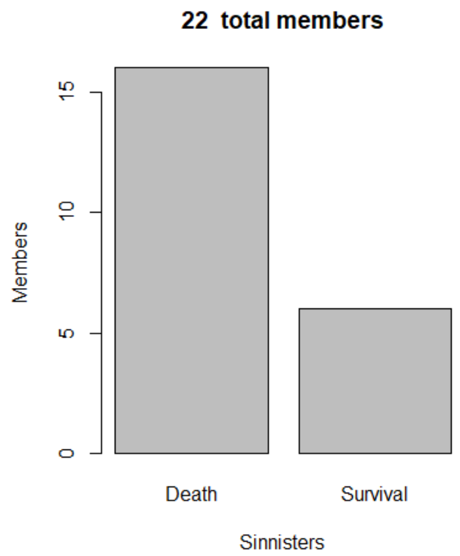
a Having a strong believe that the Storks will survive is encoded by giving the Storks a high pseudo-survivor count a and a low pseudo-deceased count b.



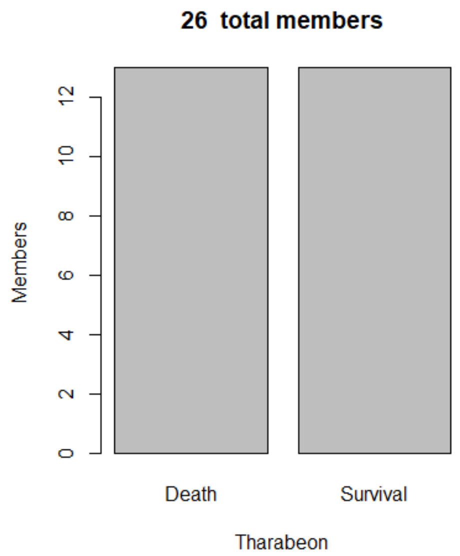
(a) a = 100, b = 4



(b) a = 20, b = 40



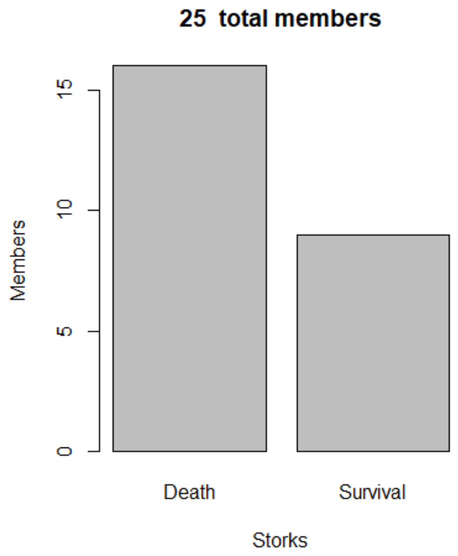
(c) a = 34, b = 15



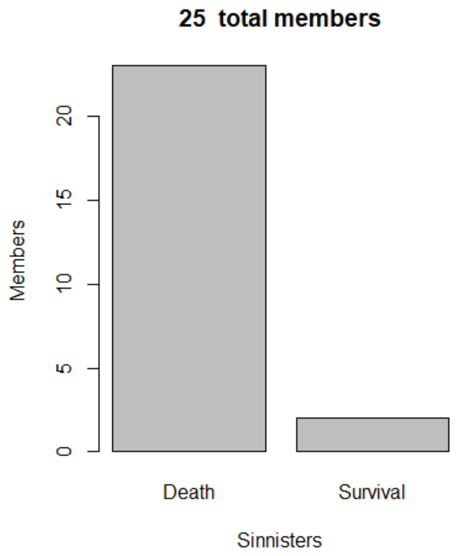
(d) a = 22, b = 22

Figure 2: Strong believe that Storks will survive.

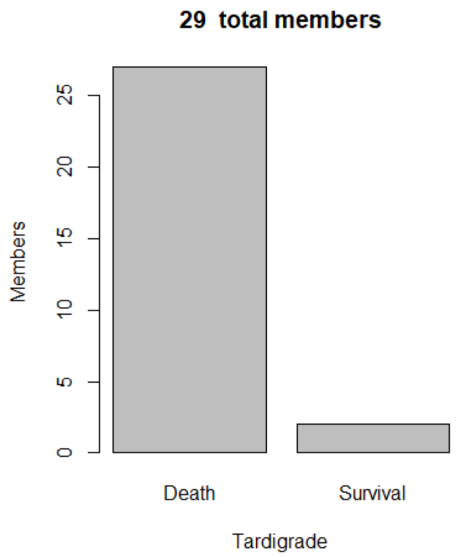
b We are not really sure what this means. We simply gave every the same low pseudo-survivor count and pseudo-survivor count.



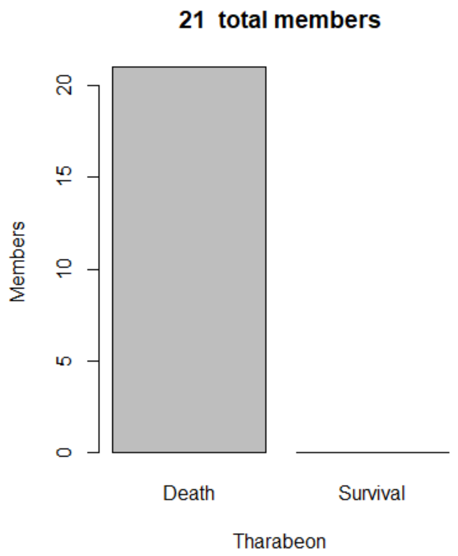
(a) a = 1, b = 10



(b) a = 1, b = 10



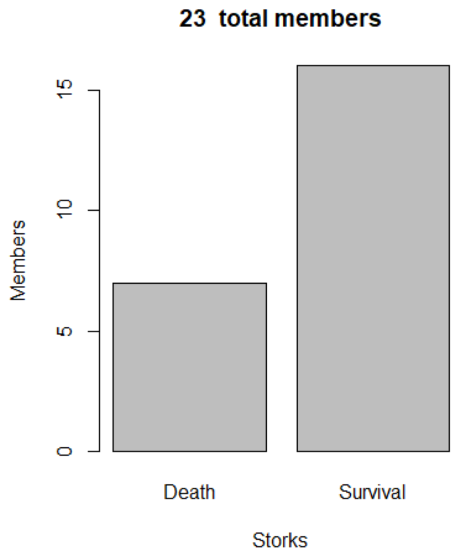
(c) a = 1, b = 10



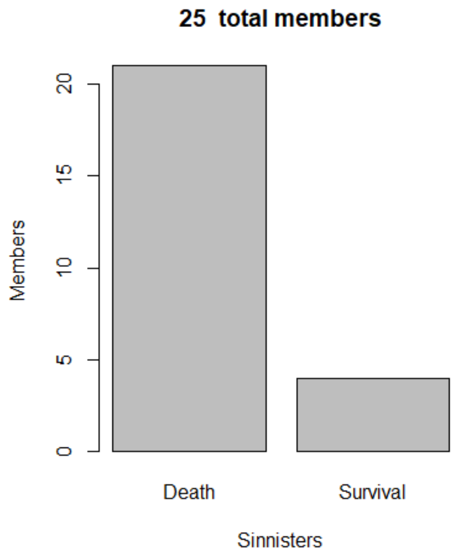
(d) a = 1, b = 10

Figure 3: Does not matter which family one belongs to; most likely going to die anyway

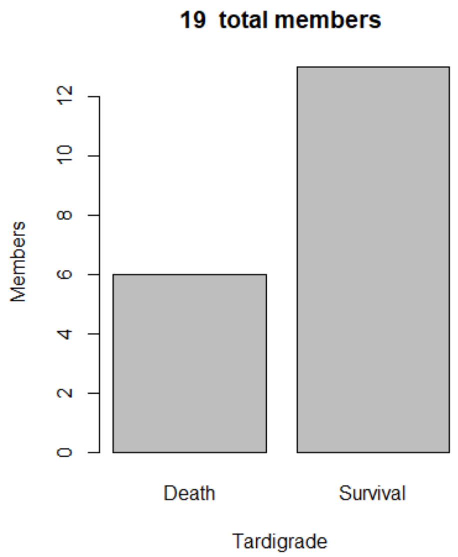
c Having uniform prior beliefs about each of the families survival probabilities refers to the pseudo-survivor count as well as the pseudo-deceased count being one for all the families.



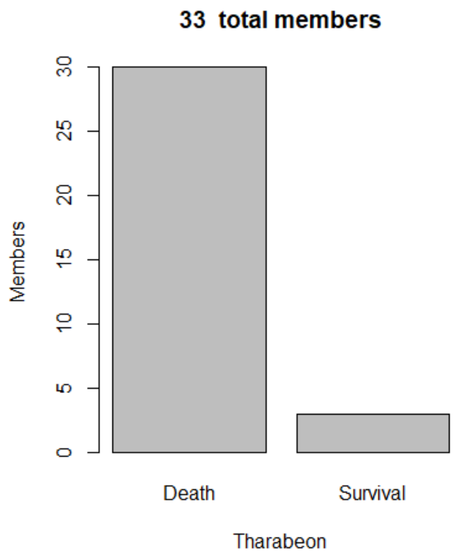
(a) a = 1, b = 1



(b) a = 1, b = 1



(c) a = 1, b = 1



(d) a = 1, b = 1

Figure 4: Uniforms believes for the survival probabilities

2 Exercise 2

2.1

$\boldsymbol{\pi}$ is the Dirichlet distribution of $\boldsymbol{\alpha}$, where $\boldsymbol{\alpha} = \alpha \times 1_K$. 1_K is a vector of length K with just one's. The vector $\boldsymbol{\pi}$ determines in what cluster a data point i is put. With a very low value for α, the probability distribution of $\boldsymbol{\pi}$ is very skewed towards one of the K values, i.e. there is a single cluster which is very close to 1, and the rest of the $K-1$ clusters are very close to 0. Hence, one cluster gets way more data points assigned than other clusters. When the value of α becomes much larger, all K values are more evenly distributed, towards the value $\frac{1}{K}$.

2.2

a is a hyperparameter of σ_k^2. When a is very small, the vector σ_k^2 also only contains very small numbers. The values of vector σ_k^2 tend to be around $\frac{a}{b}$. The larger the value of a, the more the shape of the data tends towards a Gaussian distribution.

2.3

With a very low value for σ_k^2, there will only be one peak in the histogram and the variance in x will be very low. With a very high σ_k^2-value however, the variance (distance) between the different clusters, so the range of x is very large. With a very small value for σ_k^2, the vector $\boldsymbol{\pi}$ has way less influence on the shape of the histogram.