

# Bayesian Analysis of Challenger Dataset

Class assignment for Statistical Computing I course

Subhrajyoty Roy (MB1911)

November 7, 2020

## / Introduction

The *Challenger* disaster occurred on the 28th January of 1986, when the NASA Space Shuttle orbiter *Challenger* broke apart and disintegrated at 73 seconds into its flight, leading to the deaths of its seven crew members. The accident had serious consequences for the NASA credibility resulting in an interruption of 32 months in any space shuttle programs. In order to investigate the cause of the failure, *Rogers Commission* was formed and they assigned the cause of the accident to the failure of an O-ring seal in unusually cold temperature during the launch. Although this problem with O-ring at low temperature was identified, an oversimplified analysis of the existing data on O-ring led NASA's engineers to conclude that the O-ring failure is independent of the change in temperature. However, later *Rogers Commission* found the error in the simplified analysis and concluded that,

"A careful analysis of the flight history of O-ring performance would have revealed the correlation of O-ring damage in low temperature".

Thus, the **Challenger** dataset was born, which stimulated the analysis of the effect of temperature on different O-rings of the space shuttle.

## // About Challenger Dataset

The Challenger dataset first appeared with a formal statistical analysis 3 years after the *Challenger* incident (see Dalal, Fowlkes, and Hoadley 1989). Several other authors also proposed different analysis of the data over the course of next 3-4 years (see Lavine 1991; Martz and Zimmer 1992).

Let us first take a look at the dataset.

```
library(tidyverse)
challenger <- read.table('./datasets/challenger.txt', sep = '\t', header = TRUE)
knitr::kable(challenger)
```

flight	date	nfails.field	nfails.nozzle	fail.field	fail.nozzle	temp	pres.field	pres.nozz
1	12/04/81	0	0	0	0	18.9	50	50
2	12/11/81	1	0	1	0	21.1	50	50
3	22/03/82	0	0	0	0	20.6	50	50

flight	date	nfails.field	nfails.nozzle	fail.field	fail.nozzle	temp	pres.field	pres.nozz
5	11/11/82	0	0	0	0	20.0	50	50
6	04/04/83	0	2	0	1	19.4	50	50
7	18/06/83	0	0	0	0	22.2	50	50
8	30/08/83	0	0	0	0	22.8	100	50
9	28/11/83	0	0	0	0	21.1	100	100
41-B	03/02/84	1	1	1	1	13.9	100	100
41-C	06/04/84	1	1	1	1	17.2	200	200
41-D	30/08/84	1	1	1	1	21.1	200	200
41-G	05/10/84	0	0	0	0	25.6	200	200
51-A	08/11/84	0	0	0	0	19.4	200	200
51-C	24/01/85	2	2	1	1	11.7	200	200
51-D	12/04/85	0	2	0	1	19.4	200	200
51-B	29/04/85	0	2	0	1	23.9	200	100
51-G	17/06/85	0	2	0	1	21.1	200	200
51-F	29/07/85	0	0	0	0	27.2	200	200
51-I	27/08/85	0	0	0	0	24.4	200	200
51-J	03/10/85	0	0	0	0	26.1	200	200
61-A	30/10/85	2	0	1	0	23.9	200	200
61-B	26/11/85	0	2	0	1	24.4	200	200
61-C	12/01/86	1	2	1	1	14.4	200	200

The **Challenger** dataset consists of 23 observations with 9 columns.

- **flight** column denotes an identifier for the space shuttle name.
- **date** denotes the date of launch.
- **nfails.field**, **nfails.nozzle** denotes the number of O-rings which experienced temperature specific stress in field joints and nozzle joints respectively.
- **fail.field** and **fail.nozzle** denotes whether there is a problem with the O-ring, specifically the binary versions of **nfails.field** and **nfails.nozzle**.
- **temp** is the temperature during the launch.

- `pres.field` and `pres.nozzle` are leak check pressure tests of the O-rings. These tests assured that the rings would seal the joint.

## // Goals

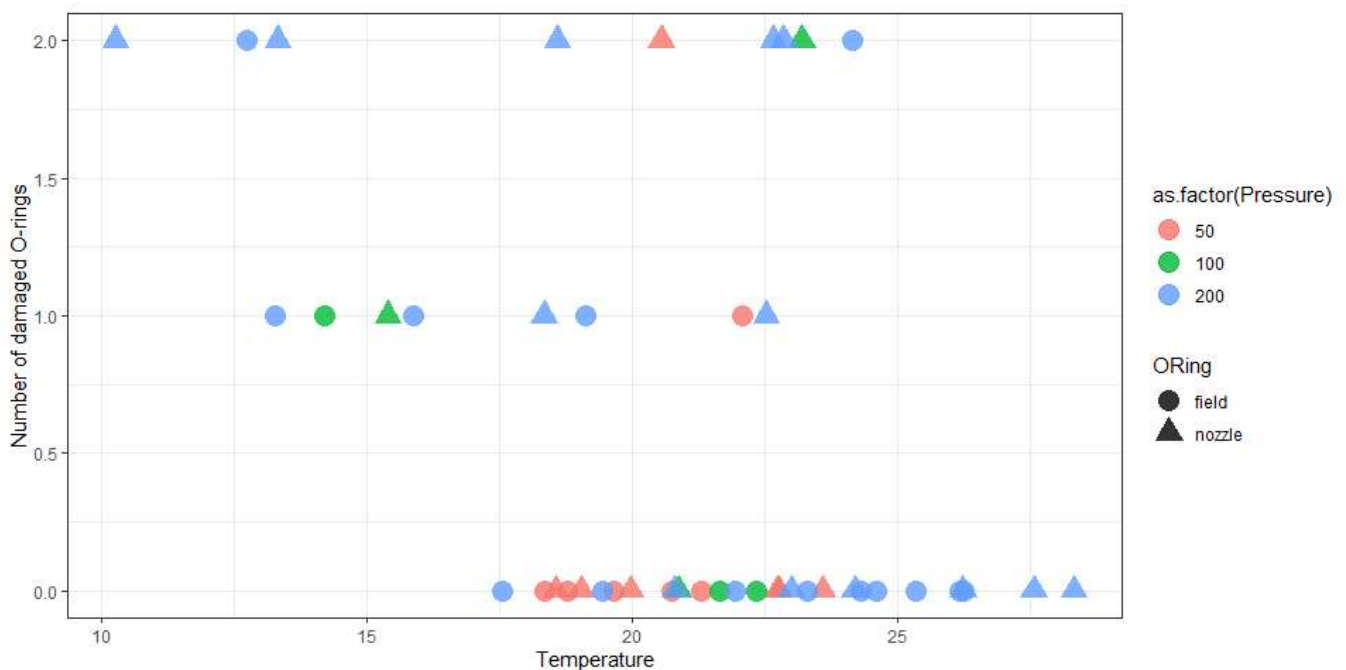
There are mainly three primary questions associated with the data.

1. Is the temperature associated with O-ring incidents? In other words, does temperature play a significant role in the determination of the failure probabilities of O-ring?
2. How a change in temperature induces variability in the probability of failure of an O-ring?
3. What was the predicted probability of failure of an O-ring for the temperature of the launch day? This can also be thought as an extrapolation problem since the launch on the Challenger incident happened at an unusually low temperature  $-0.6^{\circ}C$  or  $31^{\circ}F$ .

## / A Primary Exploration

We start by creating a plot of the two types of failure under different temperature.

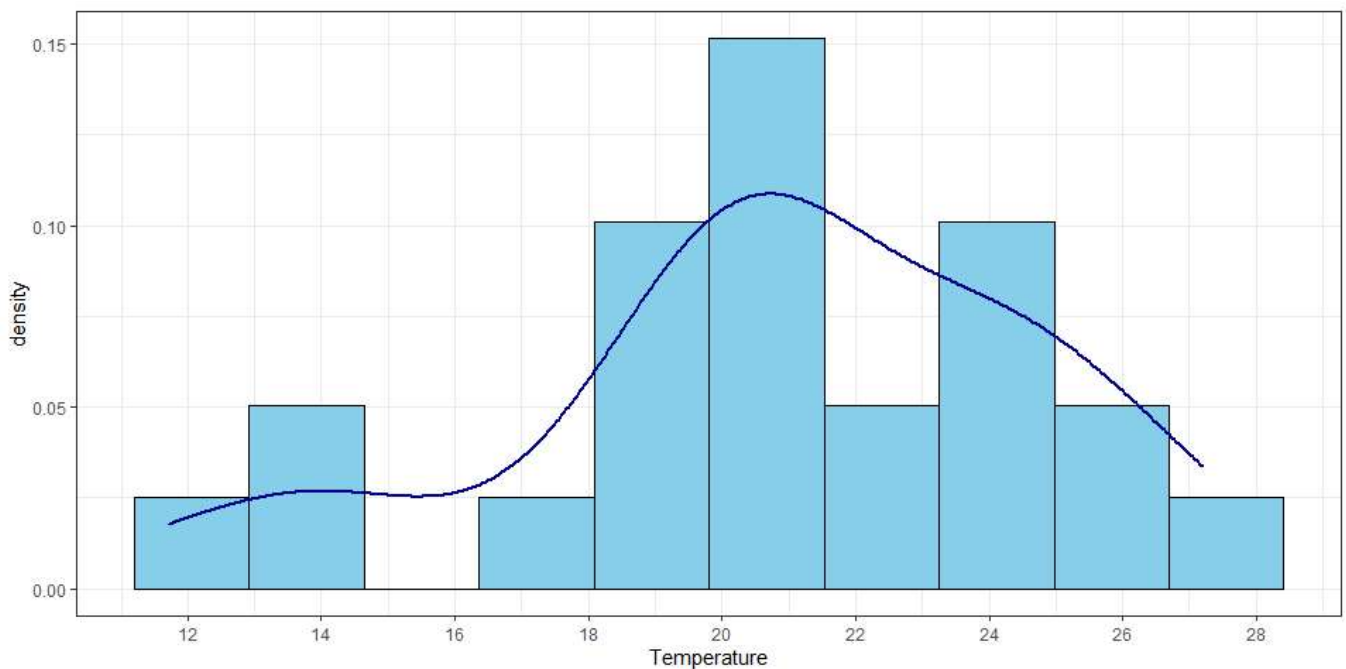
```
challenger %>%
  select(flight, Temperature = temp, nfails.field, nfails.nozzle) %>%
  pivot_longer(3:4, names_to = "ORing", values_to = "FailCount") %>%
  mutate(ORing = substring(ORing, 8)) %>%
  left_join(
    challenger %>%
      select(flight, pres.field, pres.nozzle) %>%
      pivot_longer(2:3, names_to = "ORing", values_to = "Pressure") %>%
      mutate(ORing = substring(ORing, 6))
  ) %>%
  ggplot(aes(x = Temperature, y = FailCount)) +
  geom_jitter(aes(color = as.factor(Pressure), shape = ORing), alpha = 0.8,
    size = 5, height = 0, width = 2) +
  theme_bw() + ylab('Number of damaged O-rings')
```



As the above figure shows, the number of damaged O-rings is lesser in higher temperature, and the number of damaged O-rings is higher for lower temperature. However, several O-ring failure incidents have also happened in higher temperature like  $22^{\circ}\text{C}$ . Also, it seems there is no inherent bias towards the failure of a particular O-ring either in the field joints or in the nozzle joints, as both of these seem to be equally affected by temperature. Also, different pressures do not seem to affect incident probability.

Now, the following histogram illustrates the distribution of temperature in launch dates.

```
challenger %>%
  ggplot(aes(x = temp)) +
  geom_histogram(aes(y = ..density..), fill = "skyblue", color = "black", bins = 10) +
  geom_density(color = "darkblue", size = 1) +
  scale_x_continuous(breaks = scales::pretty_breaks(n = 10)) +
  theme_bw() + xlab("Temperature")
```



Clearly, there are two very different distributions of temperature, namely a distribution of temperature about  $20^{\circ}\text{C}$  -  $26^{\circ}\text{C}$ , the natural temperature of the launch sites, while there are some launches at relatively colder days with  $12^{\circ}\text{C}$  -  $14^{\circ}$  temperature. Hence, it is worthwhile to take this two different distributions into account when a modelling of the data is proposed.

## / A Bayesian Model

Let,  $Y_i$  denotes the number of failed O-rings (combining both type of joints). The usual number of O-rings from pre-Challenger space shuttles were exactly two for each type of joints namely the nozzle and the field joints. However, following the Challenger incident, NASA started making shuttles with 3 O-ring joints instead only two (see Wikipedia contributors [2020](#)). With the usual logistic regression model in mind, we can thus model this as,

$$Y_i \sim \text{Bin}(4, p_i), \quad \text{logit}(p_i) = \alpha + \beta T_i$$

where  $p_i$  is the probability of damage of an O-ring,  $\text{Bin}(\cdot)$  denotes the binomial distribution,  $T_i$  is the temperature of the  $i$ -th launch, and  $\text{logit}(x) = \log(x/(1 - x))$  is the usual logistic link function.

However, such an assumption that  $Y_i \sim \text{Bin}(4, p_i)$  could be faulty, since the damage to the O-rings in the field joint or the nozzle joint are not independent of each other, but subjected to the same weather conditions. Hence, the Bernoulli trials describing the failure of each individual O-ring failures are confounded with each other. For this, we consider two response variables,  $Y_{i,\text{field}}$  and  $Y_{i,\text{nozzle}}$  which are the indicator variable for failure of at least one O-ring in the field joint and nozzle joint respectively.

$$\begin{aligned} Y_{i,\text{field}} &\sim \text{Ber}(p_{i,\text{field}}) & Y_{i,\text{nozzle}} &\sim \text{Ber}(p_{i,\text{nozzle}}) \\ \text{logit}(p_{i,\text{field}}) &= \alpha_1 + \beta_1 T_i & \text{logit}(p_{i,\text{nozzle}}) &= \alpha_2 + \beta_2 T_i \end{aligned}$$

As a generic model, we consider the following instead

$$Y_i \sim \text{Ber}(p_i) \quad \text{logit}(p_i) = \alpha + \beta T_i$$

For any kind of Bayesian modelling, we need to choose a prior on the parameters, here  $\alpha$  and  $\beta$ . However, because of the logistic modelling, such choice of prior on the parameters become difficult. In this regard, since the distribution of the temperature is found to be concentrated around  $14^\circ\text{C}$  and  $21^\circ\text{C}$ , we can choose a prior on the  $p_i$ 's at those specific temperatures.

For the specific purpose, we assume that  $p_i$  at  $T_i = 21^\circ\text{C}$  which is the natural temperature at the launch site, the O-ring is expected to work nicely. This assumption is based on the idea that the O-ring has been made and tested to work with the natural temperature, as guaranteed by the manufacturers themselves. In this case, we assume  $p_i \sim \text{Beta}(1, 2)$  distribution. On the other hand, at temperature  $T_i = 14^\circ\text{C}$  lower than usual temperature, the effect is unknown, hence a uniform prior  $p_i \sim \text{Beta}(1, 1)$  would be more appropriate.

Therefore, we have the following two equations

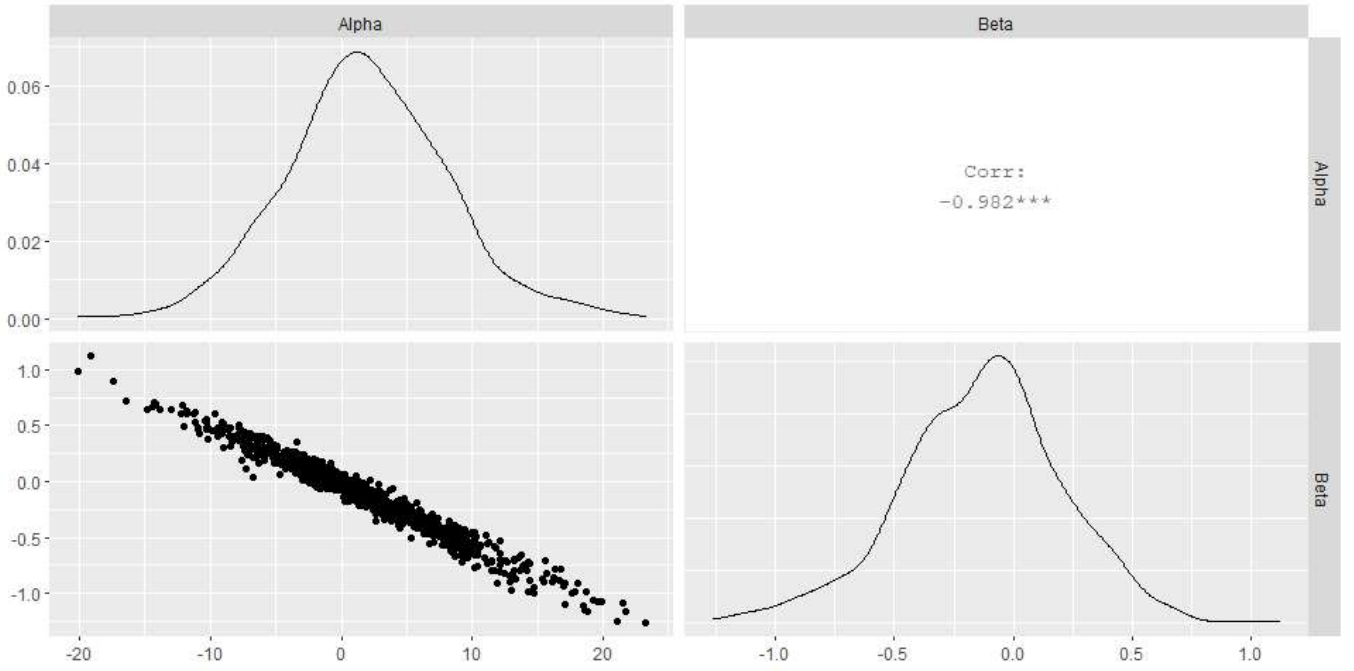
$$(\alpha + 14\beta) = \text{logit}(U) \quad (\alpha + 21\beta) = \text{logit}(V)$$

where  $U \sim \text{Beta}(1, 1)$  and  $V \sim \text{Beta}(1, 2)$ . In other words,

$$\beta = \frac{1}{7}(\log(U/(1 - U)) - \log(V/(1 - V))), \quad \alpha = \log(U/(1 - U)) - 14\beta$$

Using this relation, we can obtain the prior on  $(\alpha, \beta)$  numerically, by generating different samples  $U$  and  $V$ . The following code and the plot illustrates the approximated joint density and the marginal distributions.

```
set.seed(1911) # set a seed for reproducibility
U <- runif(1000, min = 0, max = 1) # this is Beta(1, 1) distribution
V <- rbeta(1000, shape1 = 1, shape2 = 2)
eU <- log(U/(1-U)) # take the logits
eV <- log(V/(1-V))
beta_vec <- (1 / 7) * (eV - eU)
alpha_vec <- eU - (14 * beta_vec)
GGally::ggpairs(tibble("Alpha" = alpha_vec, "Beta" = beta_vec))
```



Therefore, the proposed Bayesian model is as follows:

1.  $Y_i | p_i \sim \text{Ber}(p_i)$ .
2.  $\text{logit}(p_i) = \alpha + \beta T_i$ .
3.  $U = \frac{e^{\alpha+14\beta}}{1 + e^{\alpha+14\beta}} \sim \text{Unif}(0, 1)$ .
4.  $V = \frac{e^{\alpha+21\beta}}{1 + e^{\alpha+21\beta}} \sim \text{Beta}(1, 2)$ .
5.  $U$  and  $V$  are independently distributed.

In this case, the response  $Y_i$  denotes a binary indicator variable that takes value 1 if at least one of the two O-rings are damaged for  $i$ -th shuttle tests. However, with a slight abuse of notation, we use  $Y_i$  for damage in O-ring of either of field or nozzle joint.

## // Form of Posterior Distribution

First, we write the likelihood function,

$$L(Y_i, T_i | \alpha, \beta) = \prod_{i=1}^n \frac{e^{(\alpha+\beta T_i)Y_i}}{(1 + e^{(\alpha+\beta T_i)})} \propto e^{\sum_{i=1}^n Y_i(\alpha+\beta T_i)}$$

Next, we denote the prior distribution as  $p(\alpha, \beta)$ . To obtain this, we use a change of variable technique on the distribution of  $(U, V)$ .

$$p(U, V) = (1 - V) \quad \forall 0 \leq U, V \leq 1$$

$$\begin{aligned}
p(\alpha, \beta) &= \frac{1}{(1 + e^{\alpha+21\beta})} \times \det \begin{bmatrix} \frac{\partial U}{\partial \alpha} & \frac{\partial U}{\partial \beta} \\ \frac{\partial V}{\partial \alpha} & \frac{\partial V}{\partial \beta} \end{bmatrix} \\
&= \frac{1}{(1 + e^{\alpha+21\beta})} \det \begin{bmatrix} \frac{e^{\alpha+14\beta}}{(1+e^{\alpha+14\beta})^2} & \frac{14e^{\alpha+14\beta}}{(1+e^{\alpha+14\beta})^2} \\ \frac{e^{\alpha+21\beta}}{(1+e^{\alpha+21\beta})^2} & \frac{21e^{\alpha+21\beta}}{(1+e^{\alpha+21\beta})^2} \end{bmatrix}, \quad \alpha, \beta \in \mathbb{R} \\
&= \frac{7e^{2\alpha+35\beta}}{(1 + e^{\alpha+21\beta})^3 (1 + e^{\alpha+14\beta})^2}
\end{aligned}$$

Thus, the posterior distribution is,

$$p(\alpha, \beta \mid Y_i, T_i) \propto \frac{7e^{2\alpha+35\beta}}{(1 + e^{\alpha+21\beta})^3 (1 + e^{\alpha+14\beta})^2} \times \frac{e^{\sum_{i=1}^n Y_i(\alpha+\beta T_i)}}{\prod_{i=1}^n (1 + e^{\alpha+\beta T_i})}$$

Since integrating the posterior form in order to obtain the posterior density by normalizing is complicated, we shall use Metropolis-Hastings algorithm to obtain samples from this posterior distribution.

The following code calculates this scaled version of posterior density.

```
dens.posterior <- function(alpha, beta, Y_vec, T_vec) {
  tmp1 <- alpha + beta * T_vec
  likelihood <- exp(sum(Y_vec * tmp1)) / prod((1 + exp(tmp1)))
  prior <- 7 * exp(2 * alpha + 25 * beta) / ((1 + exp(alpha + 21 * beta))^3 * (1 +
    exp(alpha + 14 * beta))^2)
  return(prior * likelihood)
}
```

## // Obtaining Posterior Samples by MCMC

As seen before, the marginal distributions of  $\alpha$  and  $\beta$  seems close to a symmetric distribution, and hence we shall use a Gaussian distribution as the proposal distribution for our purpose. The MH algorithm then goes as follows:

1. Set  $t = 0$  and pick an initial state  $(\alpha_0, \beta_0)$ .
2. For  $t \geq 0$ , do the following:
  - Generate a random candidate  $(\alpha, \beta) \sim N((\alpha_t, \beta_t), I_2)$ .
  - Calculate the acceptance probability  $A_t((\alpha, \beta)) = \min \left( 1, \frac{p(\alpha, \beta)}{p(\alpha_t, \beta_t)} \right)$
  - Accept  $(\alpha_{(t+1)}, \beta_{(t+1)}) = (\alpha_t, \beta_t)$  with probability  $A_t((\alpha, \beta))$ .

Although the Markov chain eventually converges to the desired distribution, the initial samples may follow a very different distribution, especially if the starting point is in a region of low density. As a result, a burn-in period is typically necessary where an initial number of samples (e.g. the first 1000 or so) are thrown away. Also, the successive samples in the chain are correlated. So, one has to take only every  $n$ -th sample for some suitably chosen jumping width  $n$ .

The following code implements this MH algorithm.

```

doMH <- function(Y_vec, T_vec, burnin = 1000, thin = 25, nsamp = 200, seed = 2020) {
  set.seed(seed) # set a seed for reproducibility

  post.samples <- matrix(NA, nrow = nsamp, ncol = 2) # posterior samples
  nchain <- burnin + (thin * nsamp) - 1

  # Initialize
  alpha <- 0
  beta <- 0

  pb <- txtProgressBar(min = 0, max = nchain, style = 3)

  # Do metropolis hasting algorithm
  for (i in 1:nchain) {
    accept <- FALSE
    while (!accept) {
      new_samp <- MASS::mvrnorm(n = 1, mu = c(alpha, beta), Sigma = diag(2))
      post_ratio <- dens.posterior(new_samp[1], new_samp[2], Y_vec, T_vec) /
        dens.posterior(alpha, beta, Y_vec, T_vec)

      # if 0/0 form (due to numerical error), make it 1
      if (is.nan(post_ratio)) { post_ratio <- 0.5 }

      A <- min(1, post_ratio)
      check <- rbinom(1, size = 1, prob = A)

      if (check == 1) {
        # if accepted, update (alpha, beta)
        alpha <- new_samp[1]
        beta <- new_samp[2]
        accept <- TRUE
      }
    }

    # if we have thin-th sample after burnin, take that as iid posterior sample
    if ((i >= burnin) & ((i-burnin) %% thin == 0)) {
      index <- (i - burnin) %% thin + 1
      post.samples[index, 1] <- alpha
      post.samples[index, 2] <- beta
    }

    setTxtProgressBar(pb, value = i)
  }

  colnames(post.samples) <- c("Alpha", "Beta")
  return(post.samples)
}

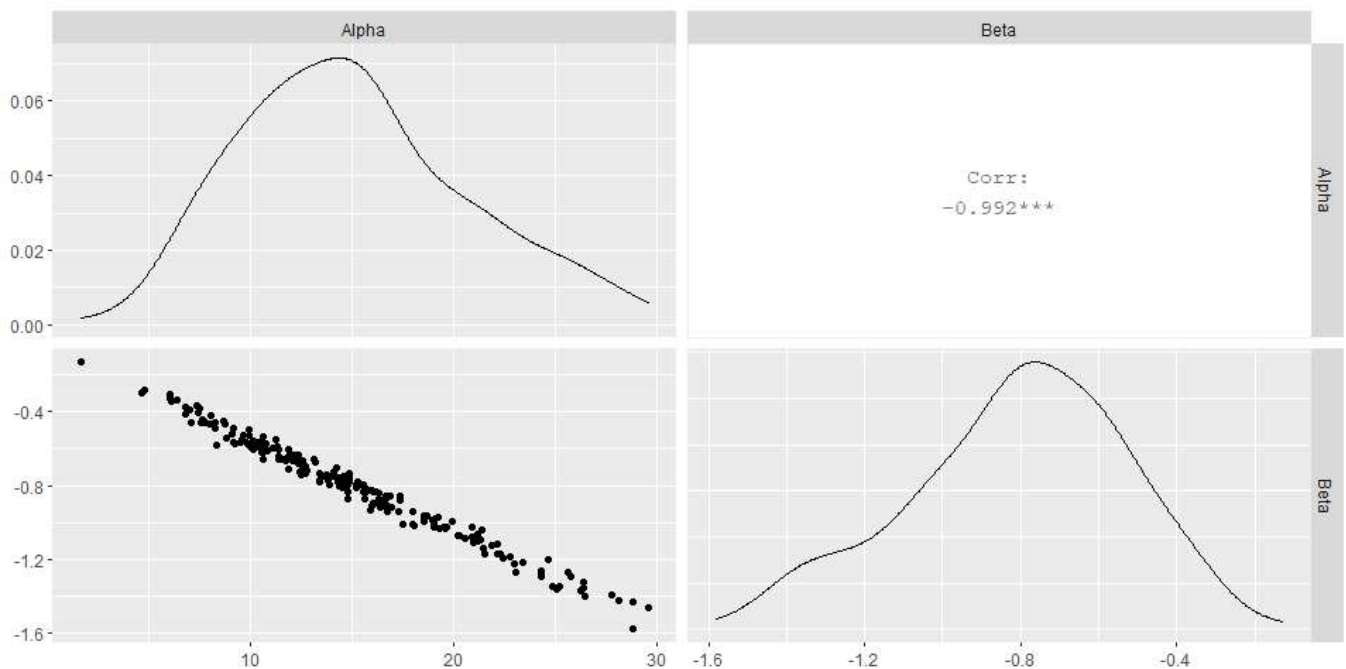
```

Now, we use this function `doMH` to obtain 200 posterior samples from the posterior distribution, with a burn-in period of 1000 iterations and take every 25-th sample in the long running Markov chain. We first obtain the posterior sample for  $(\alpha, \beta)$  pertaining to the O-ring of field joint.

```
post.field <- doMH(challenger$fail.field, challenger$temp, seed = 1911)
```



```
GGally::ggpairs(as.data.frame(post.field))
```



We see that the joint distribution in this case is similar to that of the prior, however, the marginal distributions of  $\alpha$  and  $\beta$  seems shifted. Based on this, we can now obtain the posterior mean, median and a 95% posterior credible interval for  $\alpha_{field}$ ,  $\beta_{field}$ , which are the parameters for O-ring failure at field joint.

```
apply(post.field, 2, FUN = function(x) {
  c("Mean" = mean(x), "Median" = median(x), "Lower" = quantile(x, probs = 0.025),
    "Upper" = quantile(x, probs = 0.975))
})
```

	Alpha	Beta
Mean	15.075104	-0.8083670
Median	14.636753	-0.7821916
Lower.2.5%	6.172486	-1.3991462
Upper.97.5%	26.471137	-0.3348110

So, assuming a squared error loss function, the estimate (i.e. posterior mean) is,

$$\hat{\alpha}_{field} = 15.075, \quad \hat{\beta}_{field} = -0.80837$$

and the 95% credible interval is

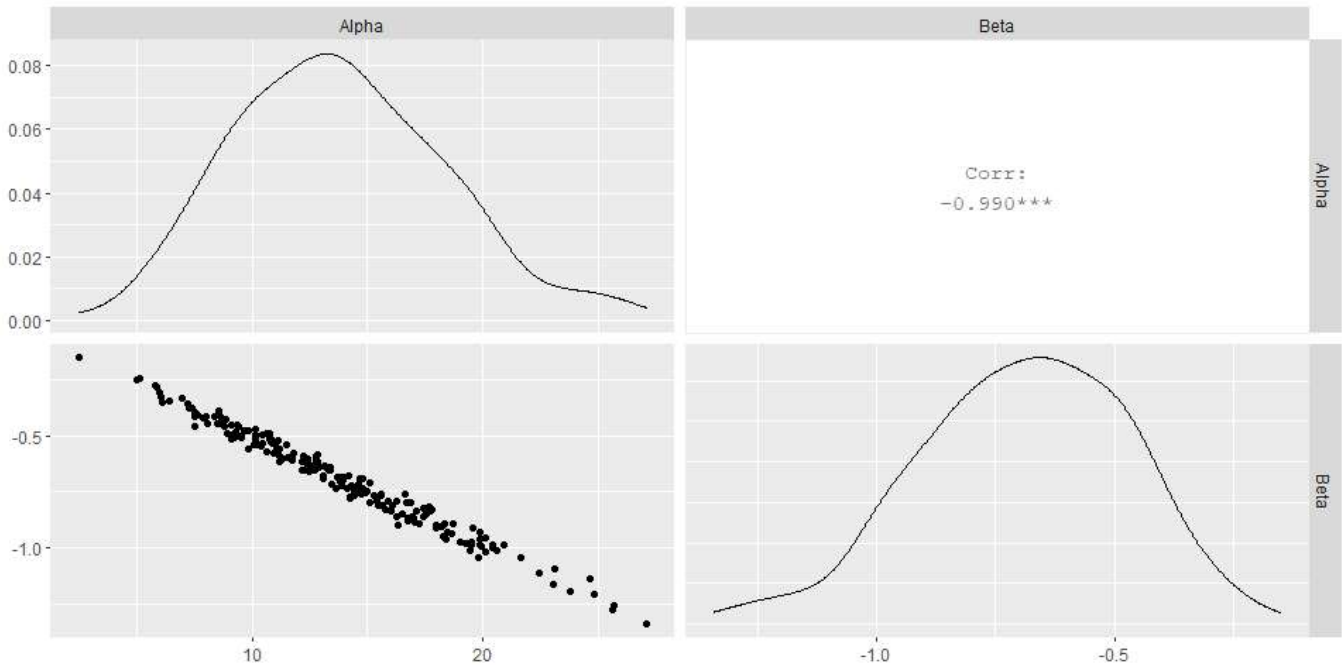
$$CI(\hat{\alpha}_{field}) = (6.172, 26.471), \quad CI(\hat{\beta}_{field}) = (-1.399, -0.3348)$$

In other words, since the credible interval of  $\beta$  does not contain 0, there is a significant negative effect of temperature on the failure of O-ring for field joint, as suggested by the above Bayesian analysis.

We also perform the same pertaining to the O-ring of the nozzle joint.

```
post.nozzle <- doMH(challenger$fail.nozzle, challenger$temp, seed = 2020)
```

```
GGally::ggpairs(as.data.frame(post.nozzle))
```



We see similar results as in the case of nozzle joint.

```
apply(post.nozzle, 2, FUN = function(x) {
  c("Mean" = mean(x), "Median" = median(x), "Lower" = quantile(x, probs = 0.025),
    "Upper" = quantile(x, probs = 0.975))
})
```

	Alpha	Beta
Mean	13.769010	-0.6919086
Median	13.366432	-0.6835910
Lower.2.5%	5.964624	-1.1679702
Upper.97.5%	23.929827	-0.3034682

In this case, denoting  $\alpha_{nozzle}$  and  $\beta_{nozzle}$  as the parameters, we have

$$\hat{\alpha}_{nozzle} = 13.769, \quad \hat{\beta}_{nozzle} = -0.6919$$

and the 95% credible interval is

$$CI(\hat{\alpha}_{nozzle}) = (5.964, 23.9298), \quad CI(\hat{\beta}_{nozzle}) = (-1.168, -0.3034)$$

In other words, since the credible interval of  $\beta$  does not contain 0, there is a significant negative effect of temperature on the failure of O-ring for nozzle joint as well.

## / Predicting Probability of O-ring Damage in Challenger Scenario

In order to predict the posterior probability of the incident at Challenger scenario, we can simply obtain  $P(Y = 1|T, \alpha, \beta)$  with  $T = (-0.6)$  and for each posterior sample of  $(\alpha, \beta)$ . These can be used to obtain a credible set for the probability of failure in the challenger incident. The following code does precisely this for the field joint.

```
ep <- post.field[, 1] * (-0.6) * post.field[, 2]
p <- exp(ep) / (1 + exp(ep))
c("Mean" = mean(p), "Median" = median(p), "Lower" = quantile(p, probs = 0.025), "Upper" =
  quantile(p, probs = 0.975))
```

Mean	Median	Lower.2.5%	Upper.97.5%
0.9728819	0.9988865	0.7801639	1.0000000

This shows there is a 78% to 100% chance of failure of O-ring in the field joint at the Challenger incident based on the pre-challenger space shuttle data. Performing similar analysis with nozzle joint, we have

```
ep <- post.nozzle[, 1] * (-0.6) * post.nozzle[, 2]
p <- exp(ep) / (1 + exp(ep))
c("Mean" = mean(p), "Median" = median(p), "Lower" = quantile(p, probs = 0.025), "Upper" =
  quantile(p, probs = 0.975))
```

Mean	Median	Lower.2.5%	Upper.97.5%
0.9648906	0.9955939	0.7476051	1.0000000

In this case also, a 95% Bayesian credible interval shows there is 74% to 100% chance of an O-ring failure in the temperature faced during Challenger shuttle launch.

## / Conclusion

Thus, the above Bayesian analysis clearly indicates that there is a significant negative association of O-ring failure with temperature during the launch. This negative association could range from  $(-1)$  to  $(-0.3)$ , assuming that the temperature is measured in celsius scale. Also, when the temperature is  $0^{\circ}\text{C}$ , since the intercept term  $\hat{\alpha}$  is positive, there is a high chance of failure of O-ring in both the joints, which is what caused the *Challenger* disaster. The effect on the O-rings of both nozzle joint and field joint are very similar, hence indicating a positive correlation of failure among them. Also, both the field joint O-ring and nozzle joint O-ring are found to have extremely high probability of failure in the circumstances and temperature faced by the *Challenger* space shuttle.

## / References

Dalal, Siddhartha R, Edward B Fowlkes, and Bruce Hoadley. 1989. "Risk Analysis of the Space Shuttle: Pre-Challenger Prediction of Failure." *Journal of the American Statistical Association* 84 (408): 945–57.

Lavine, Michael. 1991. "Problems in Extrapolation Illustrated with Space Shuttle O-Ring Data." *Journal of the American Statistical Association* 86 (416): 919–21.

Martz, HF, and WJ Zimmer. 1992. "The Risk of Catastrophic Failure of the Solid Rocket Boosters on the Space Shuttle." *The American Statistician* 46 (1): 42–47.

Wikipedia contributors. 2020. "Space Shuttle Challenger Disaster — Wikipedia, the Free Encyclopedia." [https://en.wikipedia.org/w/index.php?title=Space\\_Shuttle\\_Challenger\\_disaster&oldid=987153115](https://en.wikipedia.org/w/index.php?title=Space_Shuttle_Challenger_disaster&oldid=987153115).