Assignment 5.b for **STATGR6103**

submitted to Professor Andrew Gelman

Advait Rajagopal

10 October 2016

1 Question 1

Consider a simple one-parameter model of independent data, $y_i \sim \text{Cauchy}(\theta, 1)$, $i = 1, ..., n$, with uniform prior density on θ. Suppose $n = 2$.

1.1 Part A

Prove that the posterior distribution is proper.

A "proper" distribution is one that integrates to a finite value. We are given the following information;

$$p(\theta) \sim \text{Unif}[-\infty, +\infty]$$

$$p(y_i|\theta) \sim \text{Cauchy}(\theta, 1)$$

This means the likelihood function $p(y_i|\theta)$ is;

$$p(y_i|\theta) \propto [1 + (y_i - \theta)^2]^{-1}$$

Therefore;

$$p(y|\theta) \propto \prod_{i=1}^{n}[1 + (y_i - \theta)^2]^{-1}$$

We know that with a uniform prior on θ and a Cauchy likelihood, the posterior distribution of θ will be proportional to the likelihood.

$$p(\theta|y) \propto \prod_{i=1}^{n}[1 + (y_i - \theta)^2]^{-1}$$

To conclude that a posterior distribution $p(\theta|y)$ is proper we need to show that the integral $\int_\theta p(\theta|y)d\theta$ has a sensible form and a integrates to a finite value.[1] Consider the following intuition;

$$\lim_{\theta \to -\infty} p(\theta|y) = \lim_{\theta \to +\infty} p(\theta|y) = 0 \tag{1}$$

If the limiting values of a function as the argument approaches its extreme values in the domain are finite and in particular 0 we can conclude that the function integrates to a finite value over its domain. i.e; equation 1 implies;

$$\int_\theta p(\theta|y) = \int_{-\infty}^{+\infty} p(\theta|y)d\theta \tag{2}$$

$$= \int_{-\infty}^{+\infty} \prod_{i=1}^{n}[1 + (y_i - \theta)^2]^{-1}d\theta \propto 1 \tag{3}$$

Upon careful examination of equation 3 we find that since the θ term is quadratic and we know from equation 1 that the limiting value of $p(\theta|y)$ converges to 0, we can infer that the integral of the posterior distribution is finite and sensible and when multiplied by the appropriate normalizing constant the integral will equal 1. The posterior distribution is thus "proper".

1.2 Part B

Under what conditions will the posterior density be unimodal?
The aim here is to provide a geometric intuition. A unimodal distribution is one that has a single modal quantity.
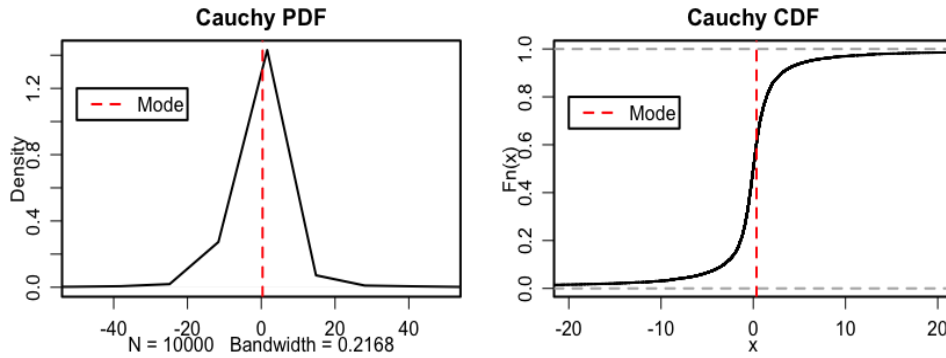

Figure 1: (a) Cauchy PDF, (b) Cauchy CDF

---

[1] Gelman's BDA 3, page 52.

It is clear from Figure 1(a) that the PDF is proper and the limiting values approach 0 on either side. Appropriately the smooth CDF follows in Figure 1(b). The conditions for a distribution to be unimodal are that the CDF function should be *convex* for values less than the modal value and *concave* for values greater than the value which is clearly evident to the left and the right of the red line in Figure 1(b).

2 Question 2

Suppose you have 100 data points that arose from the following model: $y = 3 + 0.1x_1 + 0.5x_2 + \epsilon$, with error ϵ having a t distribution with mean 0, scale 5, and 4 degrees of freedom. We shall explore the implications of fitting a standard linear regression to these data.

2.1 Part A

Simulate data from this model. For simplicity, suppose the values of x_1 are simply the integers from 1 to 100, and that the values of x_2 are random and equally likely to be 0 or 1. Use Stan to fit a linear regression (with normal errors) to these data and see if the 50% posterior intervals for the regression coefficients cover the true values.

The model is as follows;

$$y = 3 + 0.1x_1 + 0.5x_2 + \epsilon \tag{4}$$

$$\epsilon \sim t(4)$$

The posterior estimates for the Stan model are listed Table 1;

Table 1: Posterior distributions of parameters

	mean	2.5%	25%	50%	75%	97.5%	R-hat
β_0	2.11	-0.76	1.15	2.08	3.13	4.95	1
β_1	0.11	0.07	0.10	0.11	0.13	0.16	1
β_2	-0.09	-2.73	-1.02	-0.09	0.89	2.66	1
σ	6.69	5.88	6.34	6.67	6.97	7.71	1

The parameters of the posterior distribution were estimated with a noninformative uniform prior. The posterior estimates for the slope and intercept coefficients are relatively close to the true values described in equation 4. The posterior estimates belong to the 50% interval.

2.2 Part B

Put the above step in a loop and repeat 100 times. Calculate the confidence coverage for the 50% intervals for each of the three coefficients in the model. I ran the Stan model a 100 times[2] and obtained summaries of the mean, first quantile and third quantile of the three parameter estimates. These have been presented in the tables below.

Table 2: Summary of mean values of β

	Min.	1^{st}Qu.	Median	Mean	3^{rd}Qu.	Max.
β_0	1.991	2.088	2.130	2.128	2.168	2.285
β_1	0.1098	0.1118	0.1122	0.1124	0.1131	0.1149
β_2	-0.2174	-0.1279	-0.0939	-0.0932	-0.0547	0.0506

Table 3: Summary of Qu.1 values of β

	Min.	1^{st}Qu.	Median	Mean	3^{rd}Qu.	Max.
β_0	1.041	1.141	1.182	1.188	1.242	1.377
β_1	0.0941	0.09663	0.09728	0.09730	0.09792	0.09964
β_2	-1.1420	-0.994	-0.9655	-0.9627	-0.9280	-0.8020

Table 4: Summary of Qu.3 values of β

	Min.	1^{st}Qu.	Median	Mean	3^{rd}Qu.	Max.
β_0	2.910	3.018	3.059	3.055	3.104	3.281
β_1	0.1248	0.1268	0.1275	0.1275	0.1281	0.1300
β_2	0.666	0.7311	0.7707	0.7719	0.8150	1.000

[2]see code attached

4

3 Code

3.1 R Code

Listing 1: R Code

```r
rm(list = ls())
getwd()
setwd("/Users/Advait/Desktop/New_School/
          Fall16/BDA/Class9")
install.packages("beepr")
library(beepr)
x1 <- c(1:100)
x2 <- rbinom(100, 1, 0.5)
error <- 5*rt(100, 4)
y <- NULL
for (i in 1:100){
  y[i] = 3 + 0.1*x1[i] + 0.5*x2[i] + error[i]
}
plot(density(y))
library(rstan)
rstan_options(auto_write = TRUE)
options(mc.cores = parallel::detectCores())
stanc("5b.stan")
fit1 <- stan("5b.stan", data = list("x1", "x2", "y"),
            iter = 1000,
            chains = 3)
ext1 <- extract(fit1)
ext1$sigma
par(mfcol = c(1,1))
print(fit1)
plot(density(error),
     main = "Error_distributed_with_very_wide_tails")
abline(v=(mean(error)),lty  =2, col = "red")
```

Listing 2: R Code contd.

```r
meanextb0 <- NULL
meanextb1 <- NULL
meanextb2 <- NULL
q1extb0 <- NULL
q1extb1 <- NULL
q1extb2 <- NULL
q3extb0 <- NULL
q3extb1 <- NULL
q3extb2 <- NULL

for(i in 1:100){
  fit <- stan("5b.stan", data = list("x1", "x2", "y"),
                    iter = 1000,
                    chains = 3)
  meanextb0[i] <- quantile(extract(fit)$b0,
                          probs = 0.5)
  meanextb1[i] <- quantile(extract(fit)$b1,
                          probs = 0.5)
  meanextb2[i] <- quantile(extract(fit)$b2,
                          probs = 0.5)
  q1extb0[i] <- quantile(extract(fit)$b0,
                          probs = 0.25)
  q1extb1[i] <- quantile(extract(fit)$b1,
                          probs = 0.25)
  q1extb2[i] <- quantile(extract(fit)$b2,
                          probs = 0.25)
  q3extb0[i] <- quantile(extract(fit)$b0,
                          probs = 0.75)
  q3extb1[i] <- quantile(extract(fit)$b1,
                          probs = 0.75)
  q3extb2[i] <- quantile(extract(fit)$b2,
                          probs = 0.75)
}
```

Listing 3: Stan Code

```
data{
  int x1[100];
  int x2[100];
  real y[100];
  #real error[100];
  }


parameters{
  real b0;
  real b1;
  real b2;
  real <lower = 0> sigma;
}
model{
  for (i in 1: 100){
  y[i] ~ normal(b0 + b1*x1[i] + b2*x2[i], sigma);
  }
  }
```