# Newcomb's speed of light

Week3-ex3, solution

R-template `ex_speed_of_light.Rmd`.

Data file `ex_speedOfLight.dat`.

(Here we redo the analysis from page 66 in BDA3.)

Simon Newcomb conducted experiments on speed of light in 1882. He measured the time required for light to travel a certain distance and here we will analyze a data recorded as deviations from 24,800 nanoseconds. The model used in BDA3 is

$$y_i \sim N(\mu, \sigma^2) \tag{1}$$
$$p(\mu, \sigma^2) \propto \sigma^{-2} \tag{2}$$

where $y_i$ is the $i$'th measurement, $\mu$ is the mean of the measurement and $\sigma^2$ the variance of the measurements. Notice that this prior is improper ("uninformative"). This corresponds to widely used uniform prior for $\mu$ in the range $(-\infty, \infty)$, and uniform prior for $\log(\sigma)$ (BDA3 pp. 66, 52, and 21). Both priors are improper and cannot be found from Stan. You can use instead

$$\begin{aligned} p(\mu) &\sim N(0, (10^3)^2) \\ p(\sigma^2) &\sim \text{Inv-}\chi^2(\nu = 4, s^2 = 1) \end{aligned} \tag{3}$$

In this exercise your tasks are the following:

1. Write a Stan model for the above model and sample from the posterior of the parameters. Report the posterior mean, variance and 95% central credible interval for $\mu$ and $\sigma^2$.
2. Additionally draw samples from the posterior predictive distribution of hypothetical new measurement $p(\tilde{y}|y)$. Calculate the mean, variance and 95% quantile of the posterior predictive distribution.
3. How does the posterior predictive distribution differ from the posterior of $\mu$ and Why?
4. Which parts of the model could be interpreted to correspond to aleatory and epistemic uncertainty? Discuss whether this distinction is useful here.
5. Instead of Inverse-$\chi^2$ distribution the variance parameter prior has traditionally been defined using Gamma distribution for the precision parameter $\tau = 1/\sigma^2$. By using the results in Appendix A of BDA3 derive the analytic form of a Gamma prior for the precision corresponding to the prior (3). This should be of the form Gamma$(\alpha, \beta)$, where $\alpha$ and $\beta$ are functions of $\nu$ and $s^2$.

**Note!** Many common distributions have multiple parameterizations, for which reason you need to be careful when interpreting others' works. The variance/precision parameter and their priors are notorious for this. The reason is mainly historical since different parameterizations correspond to different analytical solutions.

**Grading:** 2 points from correct answer for each of the above steps.

## Model answers

Load the needed libraries into R and set options for multicore computer.

```
library(ggplot2)
library(StanHeaders)
library(rstan)
```

```
## rstan (Version 2.21.2, GitRev: 2e1f913d3ca3)

## For execution on a local, multicore CPU with excess RAM we recommend calling
## options(mc.cores = parallel::detectCores()).
## To avoid recompilation of unchanged Stan programs, we recommend calling
## rstan_options(auto_write = TRUE)
```

```
set.seed(123)

options(mc.cores = parallel::detectCores())
rstan_options(auto_write = TRUE)
```

**Part 1 & 2.**

write the model description, set up initial values for 4 chains and sample from the posterior

```
speed_of_light_observations_model = "data {
    int n; //the number of observations
    vector[n] y; //data recorded as deviations from 24800 nanoseconds
  }
parameters {
  real mu; //mean
  real<lower=0> sigma2; //precision
}
model {
  mu~normal(0,sqrt(10^6));
  sigma2~scaled_inv_chi_square(4,1);
  y ~ normal(mu,sqrt(sigma2));  //or with for loop over the vector elements

}
generated quantities {
      // produce samples from the posterior predictive distirbution
      real ytilde;
      ytilde = normal_rng(mu,sqrt(sigma2));
}"

# Load the data and put it into a list format
dataset <- read.table ("ex_speedOfLight.dat", header=TRUE)
#dataset
y <- dataset$y
#hist(y,20)
# set data into named list
data = list(n =nrow(dataset), y =y)

# initialize parameters
init1 = list(mu = 10, sigma2 = 1)
init2 = list(mu = 20, sigma2 = 10)
init3 = list(mu = 30, sigma2 = 100)
init4 = list(mu = 40, sigma2 = 1000)
```

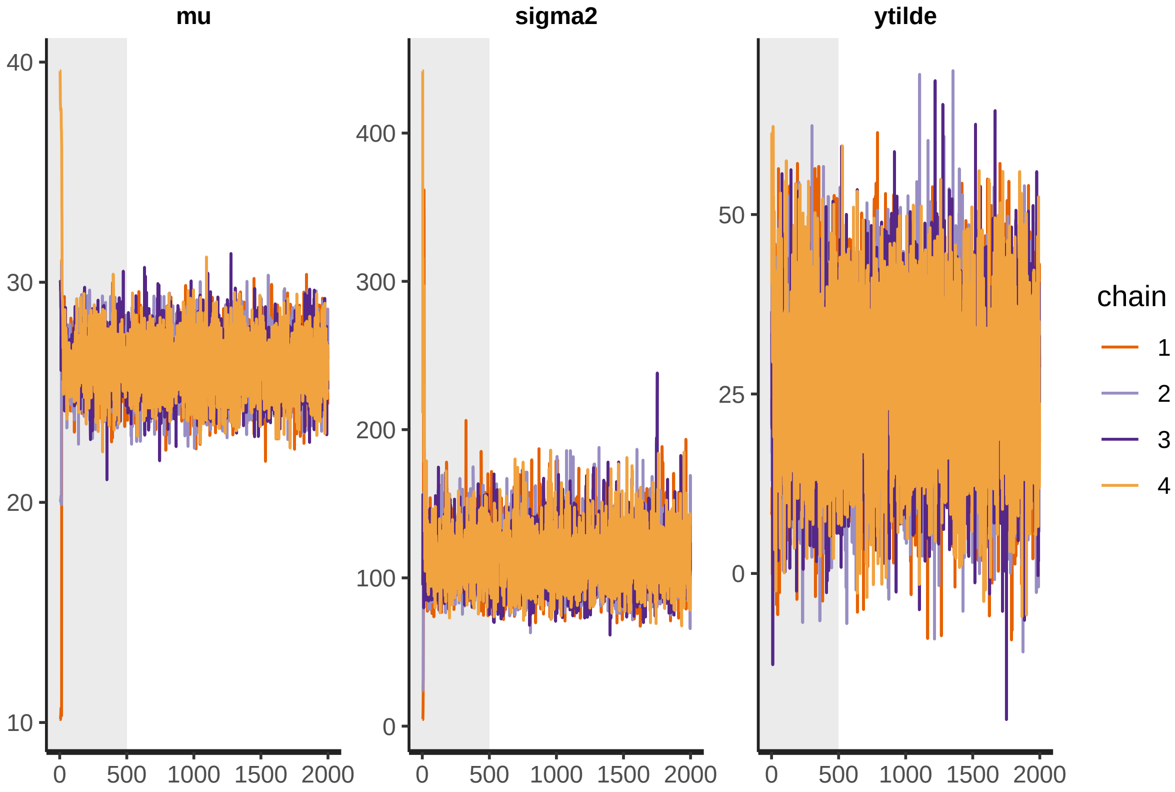
```
inits <- list(init1, init2, init3, init4)

# Fit a model defined in the Stan modeling language and return the fitted result as an instance of stan
set.seed(123)
post=stan(model_code=speed_of_light_observations_model,data=data,warmup=500,iter=2000,chains=4,thin=1,in
```

Let's then examine the convergence and autocorrelation of the chains.

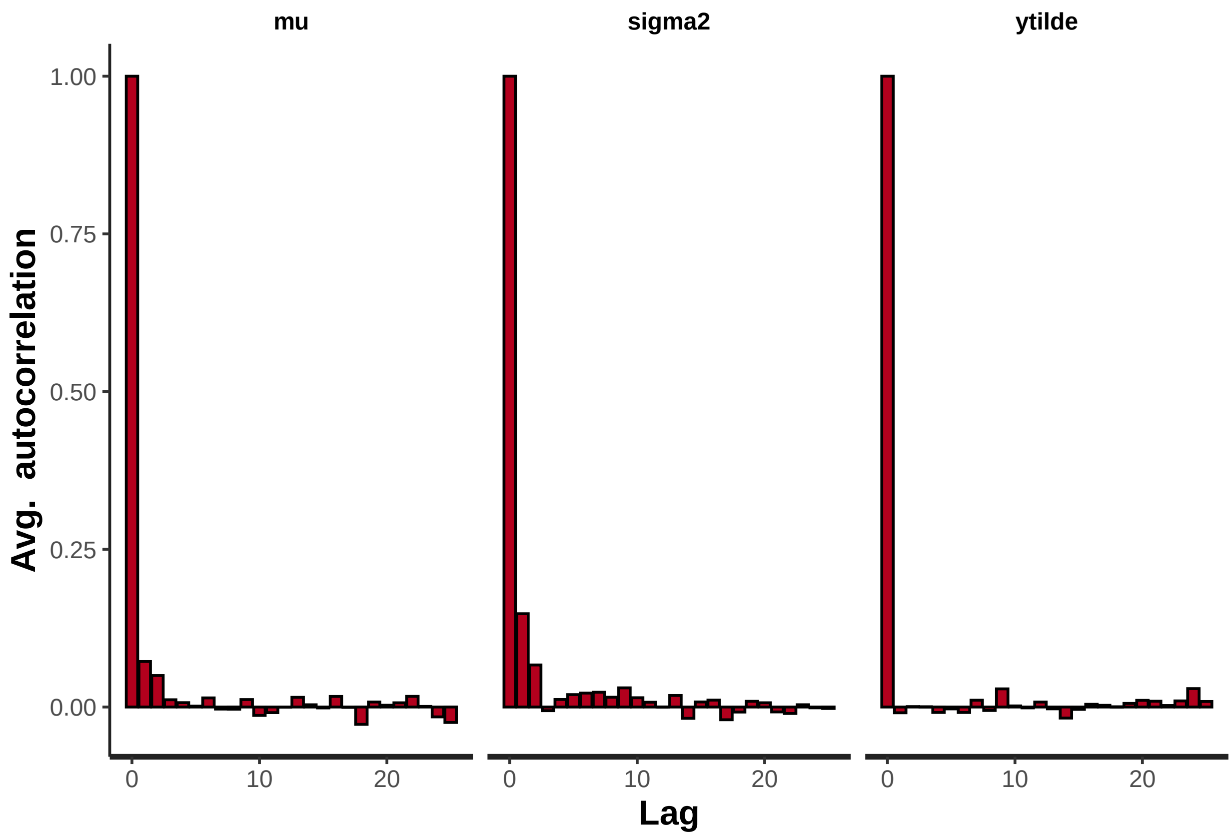
```
# Check the names of sampled variables
names(post)
```

```
## [1] "mu"     "sigma2" "ytilde" "lp__"
```

```
#check the convergence visually, plot the sample chains
plot(post, plotfun= "trace", inc_warmup = TRUE)
```




```
# plot autocorrelation function
stan_ac(post,inc_warmup = FALSE, lags = 25)
```




Note, print(post) returns also mean, sd (=square root of variance) and some quantiles (95% central credible interval is defined by 2.5% and 97.5% quantiles) for all parameters. The 95% quantile for $\tilde{y}$ has to be calculated separately though.

```
# print Rhat and other summary statistics
# summary(post)
print(post)
```

```
## Inference for Stan model: fc73293ba31c3f826629687a629d244d.
## 4 chains, each with iter=2000; warmup=500; thin=1;
## post-warmup draws per chain=1500, total post-warmup draws=6000.
##
##            mean se_mean    sd    2.5%     25%     50%     75%   97.5% n_eff Rhat
## mu        26.21    0.02  1.28   23.72   25.35   26.22   27.05   28.79  4584    1
## sigma2   112.15    0.33 19.49   80.08   98.46  110.11  123.31  156.44  3512    1
## ytilde    26.25    0.14 10.65    5.56   19.23   26.28   33.35   47.40  6099    1
## lp__    -199.63    0.02  0.99 -202.31 -200.00 -199.32 -198.93 -198.67  2845    1
##
## Samples were drawn using NUTS(diag_e) at Tue Nov 30 17:29:18 2021.
## For each parameter, n_eff is a crude measure of effective sample size,
## and Rhat is the potential scale reduction factor on split chains (at
## convergence, Rhat=1).
```

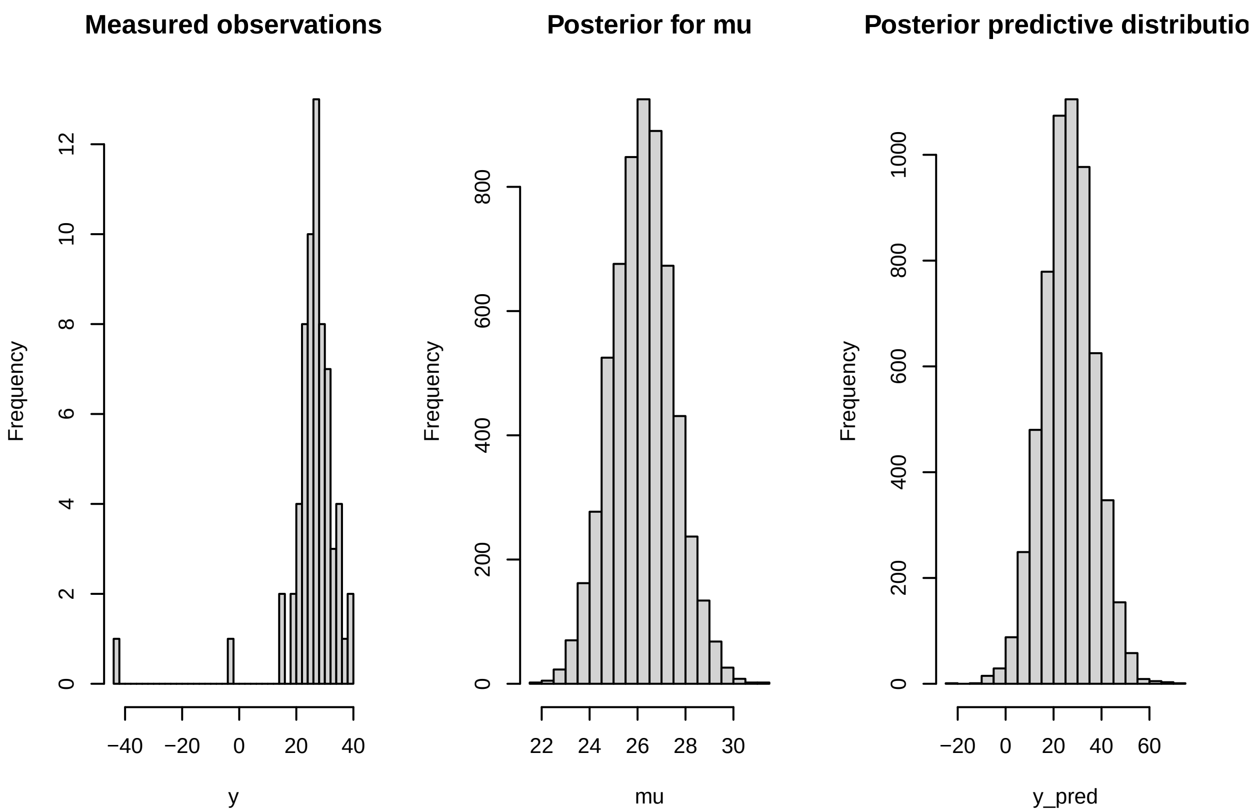
```
quantile(as.matrix(post, pars =c("ytilde")), 0.95 )
```

```
##      95%
## 43.67767
```

**Part 3.**

Note. We had defined the Stan model so that posterior predictive samples for $\tilde{y}$ where generated in the "generated quantities" block. Hence, we can examine them directly using the print(post) command. Let's additionally visualize the posteriors

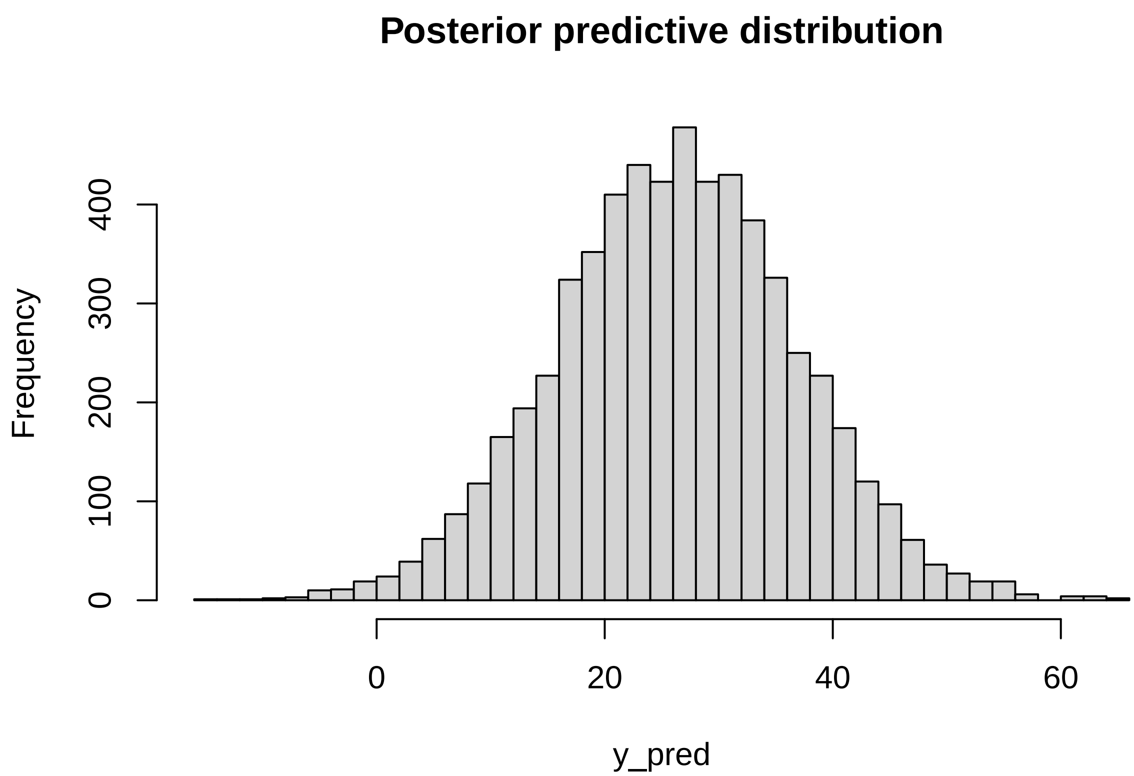
```
post_samples=as.matrix(post, pars =c("mu","sigma2","ytilde"))
# write.table(post_samples, file="param.txt", row.names=FALSE, col.names=TRUE)

par(mfrow=c(1,3))
hist(y,main='Measured observations',breaks=30)
hist(post_samples[,"mu"], main="Posterior for mu", xlab="mu",breaks=30)
hist(post_samples[,"ytilde"], main="Posterior predictive distribution", xlab="y_pred",breaks=30)
```




From the above histograms we see that the posterior mean of $\mu$ matches that of $\tilde{y}$ whereas the spread of the posterior distribution for $\tilde{y}$ is much wider than that for $\mu$. The reasons are that in the posterior of $\tilde{y}$ 1) we have marginalized over uncertainty in $\mu$ and $\sigma^2$ and 2) we have taken into account the observation error represented by the observation model $\tilde{y} \sim N(\mu, \sigma^2)$.

Below we demonstrate how to sample from the posterior predictive distribution outside Stan.

```
# Draw samples from the posterior predictive distribution
set.seed(123)
ytilde2 <- rnorm(length(post_samples[,"mu"]), post_samples[,"mu"], sqrt(post_samples[,"sigma2"]))
hist(ytilde2, main="Posterior predictive distribution", xlab="y_pred",breaks=30)
```




**Part 4**

The uncertainty about $\mu$ and $\sigma^2$ is epistemic in nature and it is represented by the prior and posterior distributions. The uncertainty on $y$ or $\tilde{y}$ contains also aleatory uncertainty which is described by the obseration model $\tilde{y} \sim N(\mu, \sigma^2)$.

**Part 5**

Here the question was to find parameter transformation between Scaled-Inverse-Chi distribution for variance parameter,

$\sigma^2 \sim \text{Inv-}\chi(\nu = 4, s^2 = 1)$

and Gamma distributions for the precision parameter

$1/\sigma^2 \sim \text{Gamma}(\alpha, \beta)$

From page 579 of BDA3 we get the following relationship. $Inv - Chi(\nu = 4, s^2 = 1)$ is the same as $\text{Inv-Gamma}(\alpha, \beta)$ with

$\alpha = \frac{\nu}{2} = 2$

$\beta = \frac{\nu}{2} s^2 = 2$

On the other hand, on page 583 of BDA3 it is told that if $1/\theta$ has a Gamma distribution with parameters $\alpha$, $\beta$ then $\theta$ has the inverse-Gamma distribution." (And the other way around). Hence, the prior for the precision is

$1/\sigma^2 \sim \text{Gamma}(2, 2)$

## Grading

**Total 10 points.** 2 points from correct answer for each of the above steps.