MCMC Diagnostics - IFLS data

Sarah Teichman

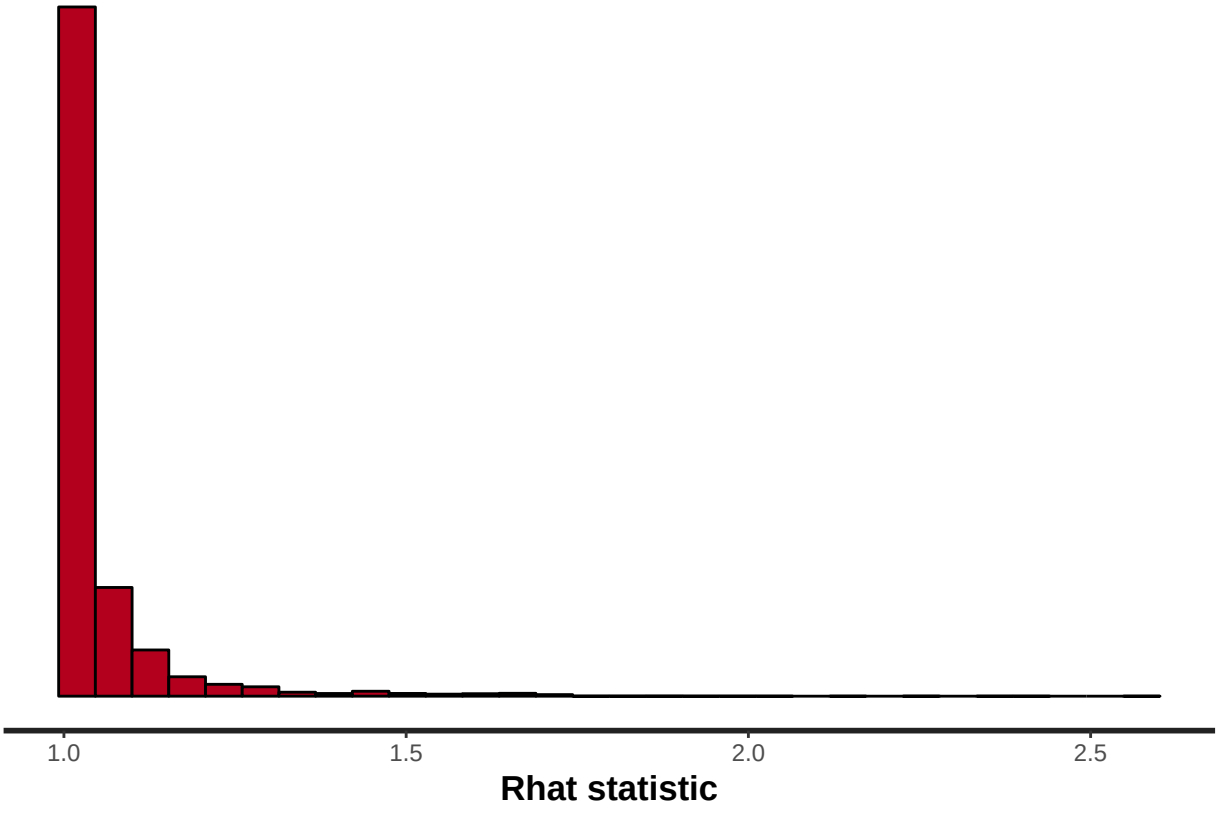
05/29/2020

```
K <- 7
Ti <- 3
N <- 1973
```

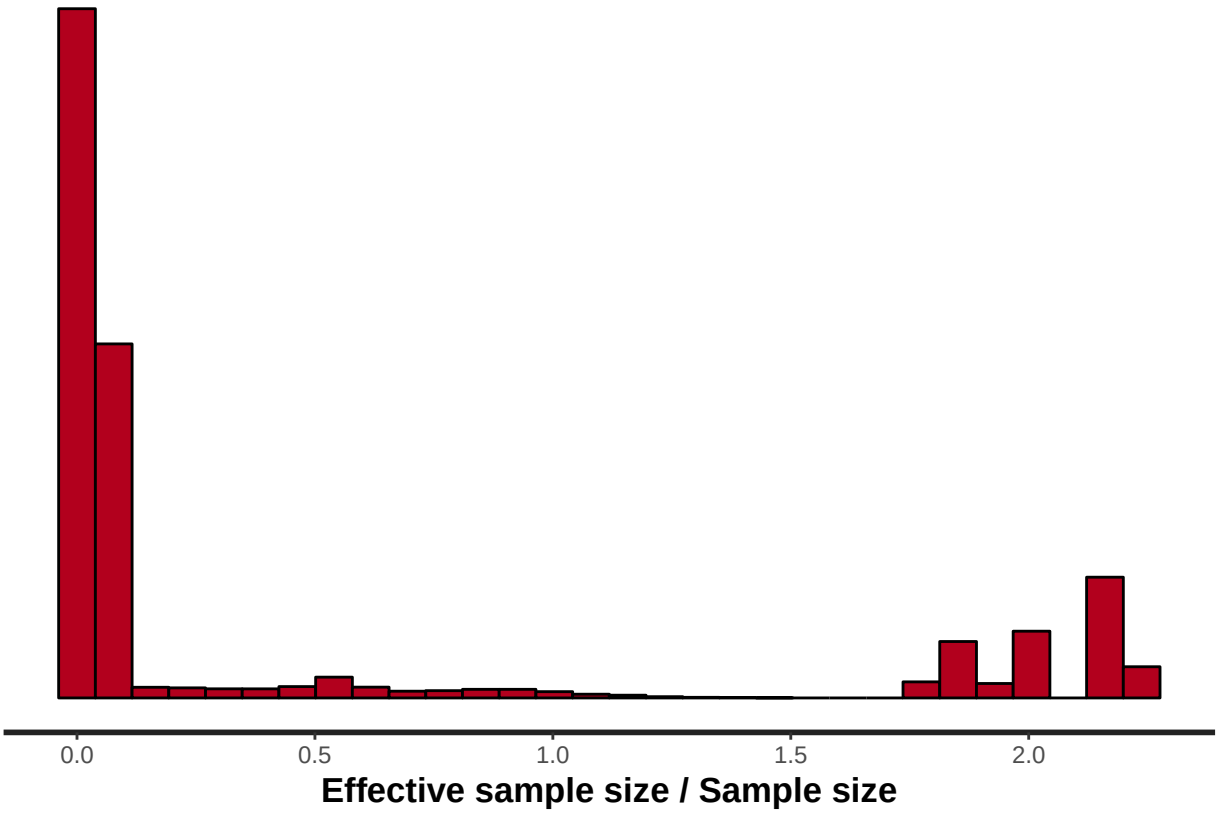
General MCMC diagnostic plots

Overall model diagnostics from rstan package.

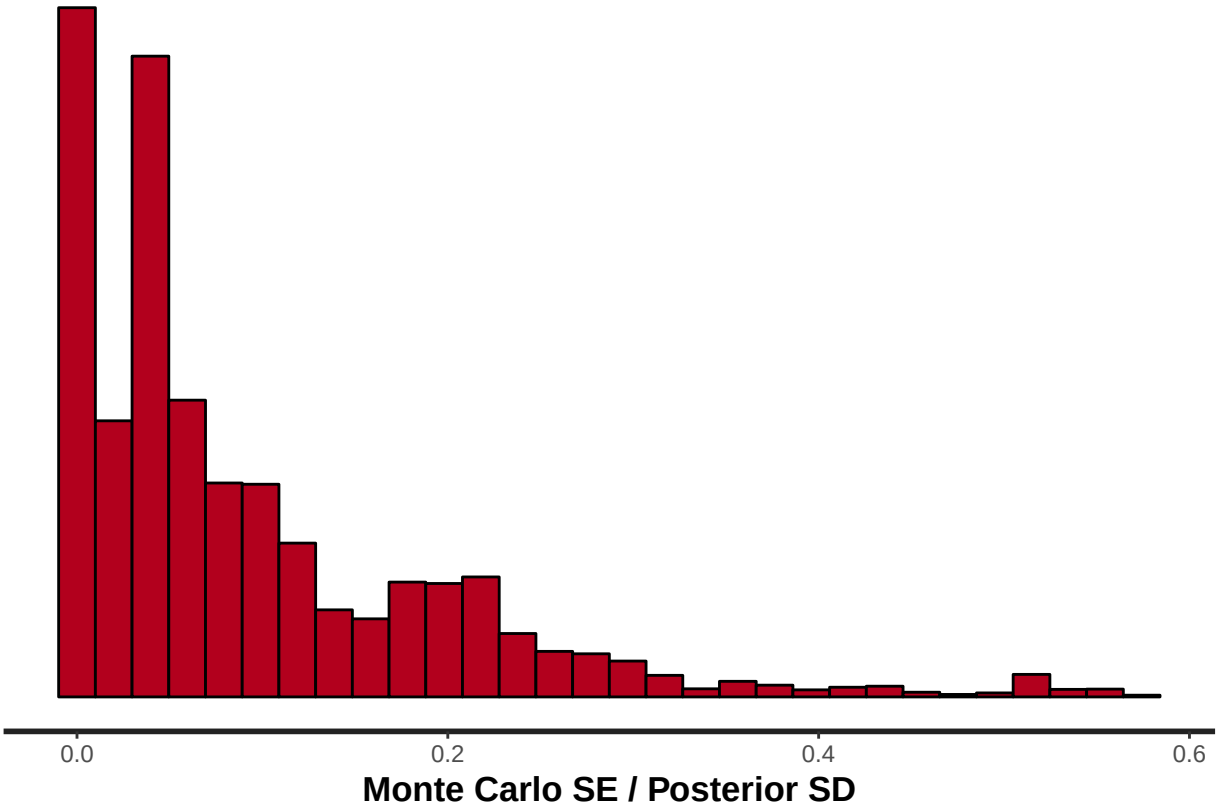
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

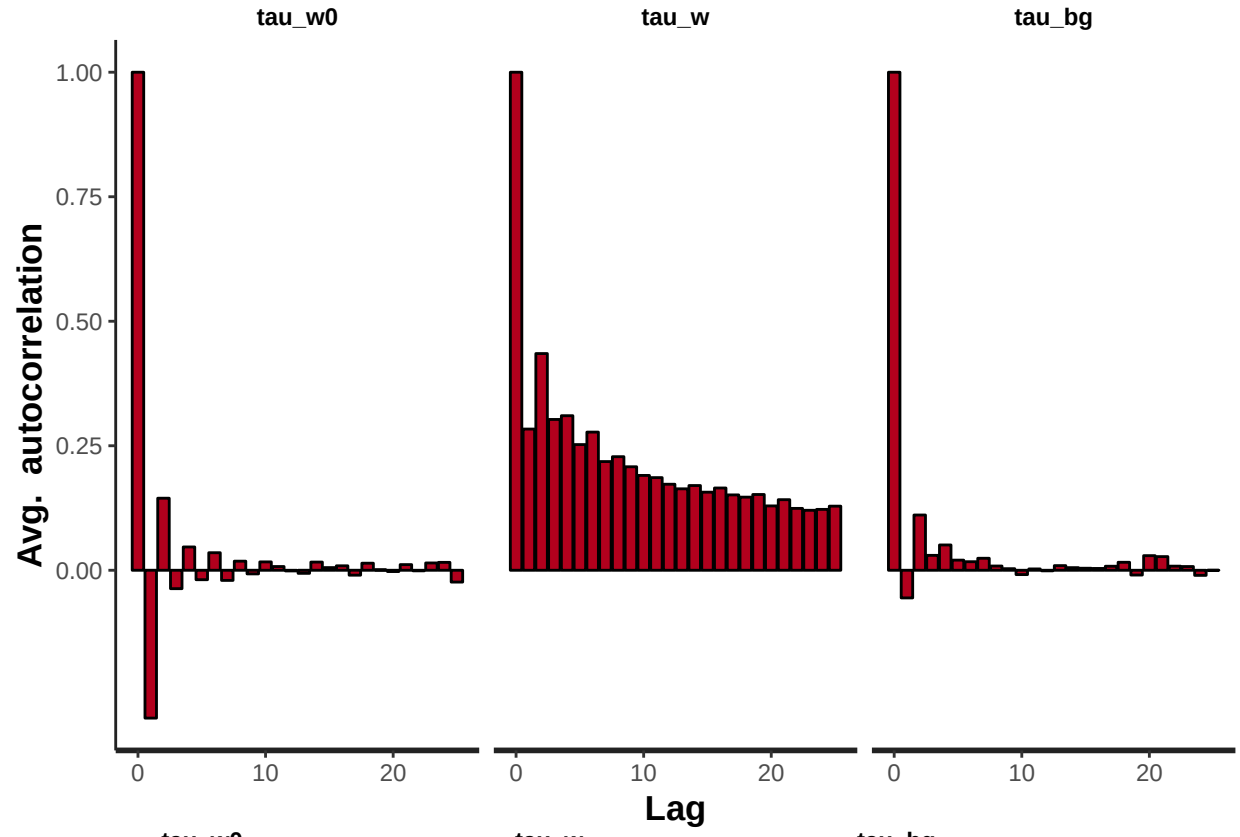


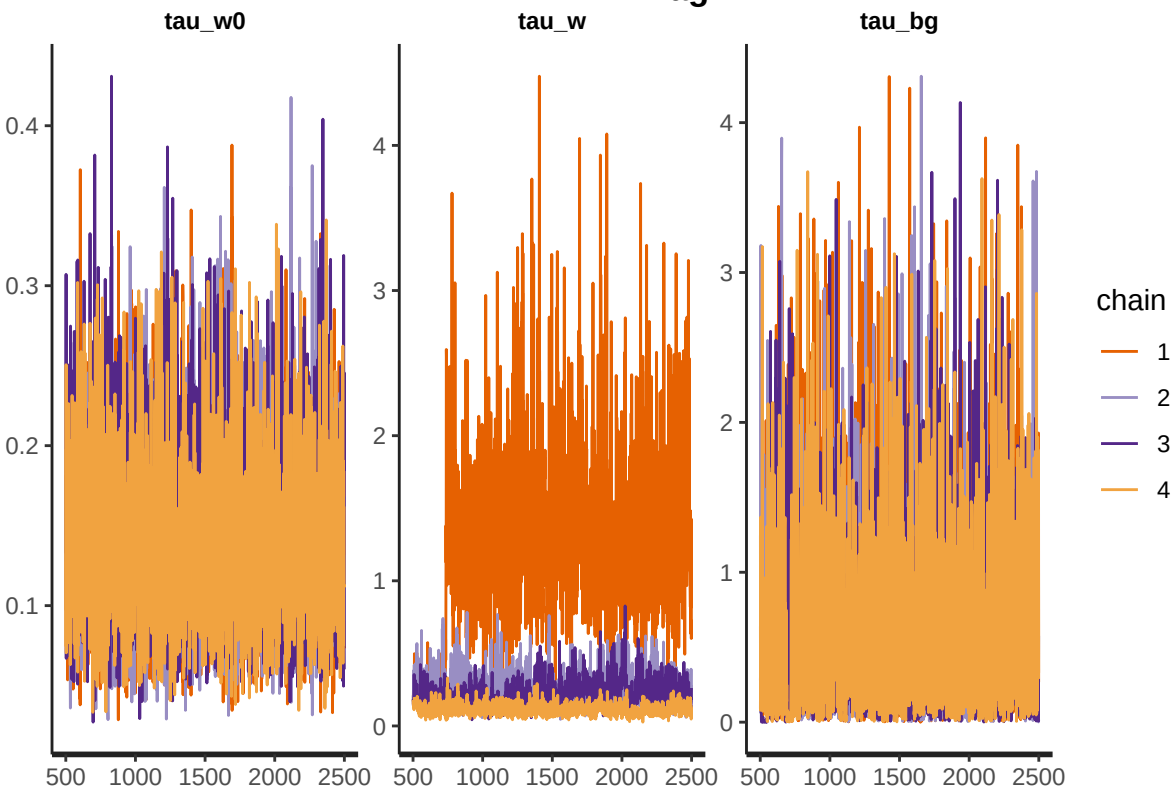
Individual Parameter Diagnostics

Individual parameter plots. Autocorrelation and trace plots for individual parameters, and histograms of posterior medians for group parameters.

```
get_single_plots <- function(fit, param) {
  print(fit_summ[param,c(1,2,3,5,6,7,9,10)])
  print(stan_ac(fit, pars = param))
  print(rstan::traceplot(fit, pars = param))
}
get_aggreg_plots <- function(fit, param, trim = F, trim_amount) {
  ind <- grep(paste0("^",param), rownames(as.data.frame(summary(fit)$summary)))
  medians <- data.frame(avg = as.data.frame(summary(fit)$summary)$`50%`[ind])
  title <- paste0("Posterior Medians of ",param)
  print(ggplot(medians, aes(x = avg)) + geom_histogram(bins = 30) + ggtitle(title) +
    xlab("Medians") + ylab("Count"))
  print("       ")
  if (trim == T) {
    lim <- quantile(abs(medians$avg), probs = trim_amount)
    meds_trim <- medians %>% filter(abs(medians$avg) < lim)
    print(ggplot(meds_trim, aes(x = avg)) + geom_histogram(bins = 60) +
            ggtitle(paste0(title, " Without Extreme ",100*(1-trim_amount),"%")))
  }
  means <- data.frame(avg = as.data.frame(summary(fit)$summary)$`mean`[ind])
  title <- paste0("Posterior Means of ",param)
  print(ggplot(means, aes(x = avg)) + geom_histogram(bins = 30) + ggtitle(title) +
    xlab("Means") + ylab("Count"))
  print("       ")
  sds <- data.frame(avg = as.data.frame(summary(fit)$summary)$`sd`[ind])
  title <- paste0("Posterior Standard Deviations of ",param)
  print(ggplot(sds, aes(x = avg)) + geom_histogram(bins = 30) + ggtitle(title) +
    xlab("Standard Deviations") + ylab("Count"))
}
plot_fit <- function(fit) {
  get_single_plots(fit, tau_params)
  get_single_plots(fit, sigma_params)
  get_single_plots(fit, beta)
  get_single_plots(fit, gamma)
  get_aggreg_plots(fit, "w")
  get_aggreg_plots(fit, "z")
  get_aggreg_plots(fit, "p")
}
plot_fit(fit)
```

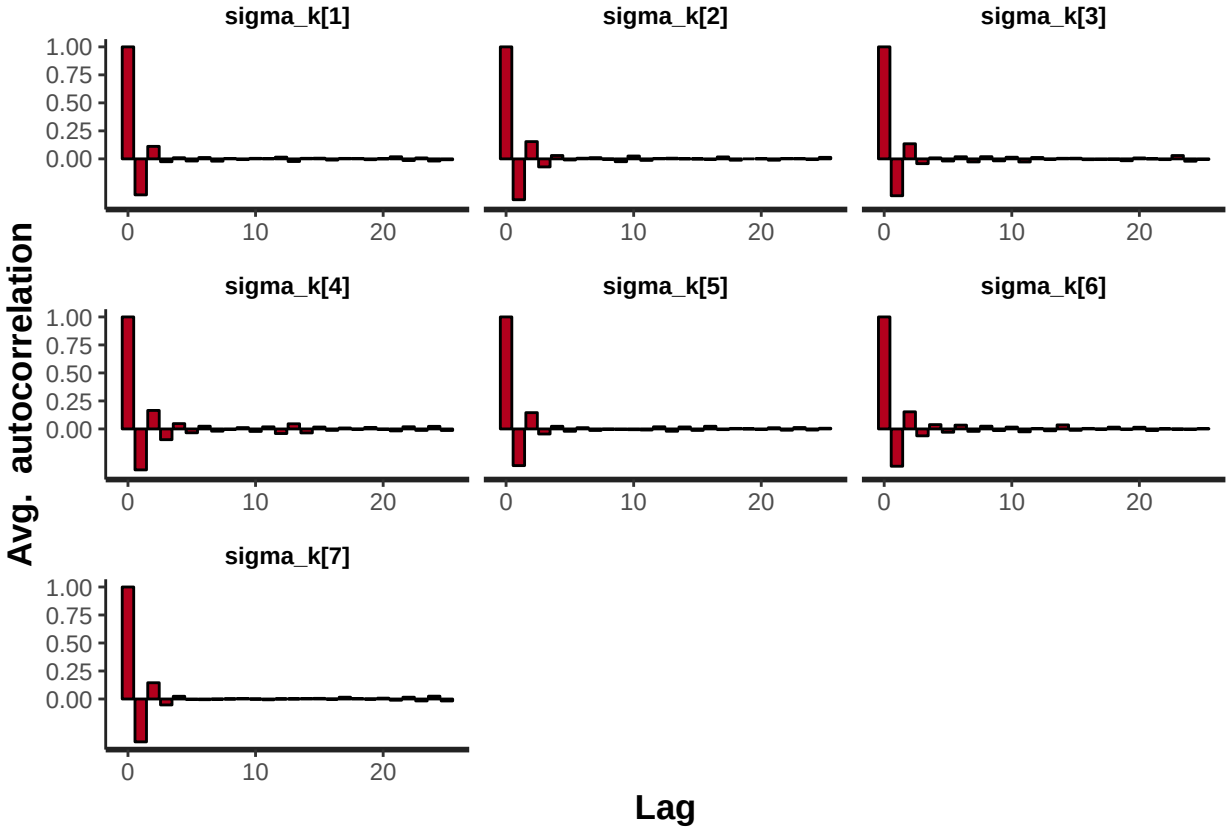
```
##             mean     se_mean         sd       25%       50%       75%
## tau_w0 0.1404302 0.000629097 0.05267168 0.1023252 0.1344418 0.1712962
## tau_w  0.4505817 0.321285436 0.55232299 0.1332295 0.2119770 0.4076351
## tau_bg 0.5997456 0.008798181 0.58737368 0.1759276 0.4216803 0.8366089
##              n_eff     Rhat
## tau_w0 7010.016802 1.004218
## tau_w     2.955318 1.829302
## tau_bg 4457.005542 1.004322
```

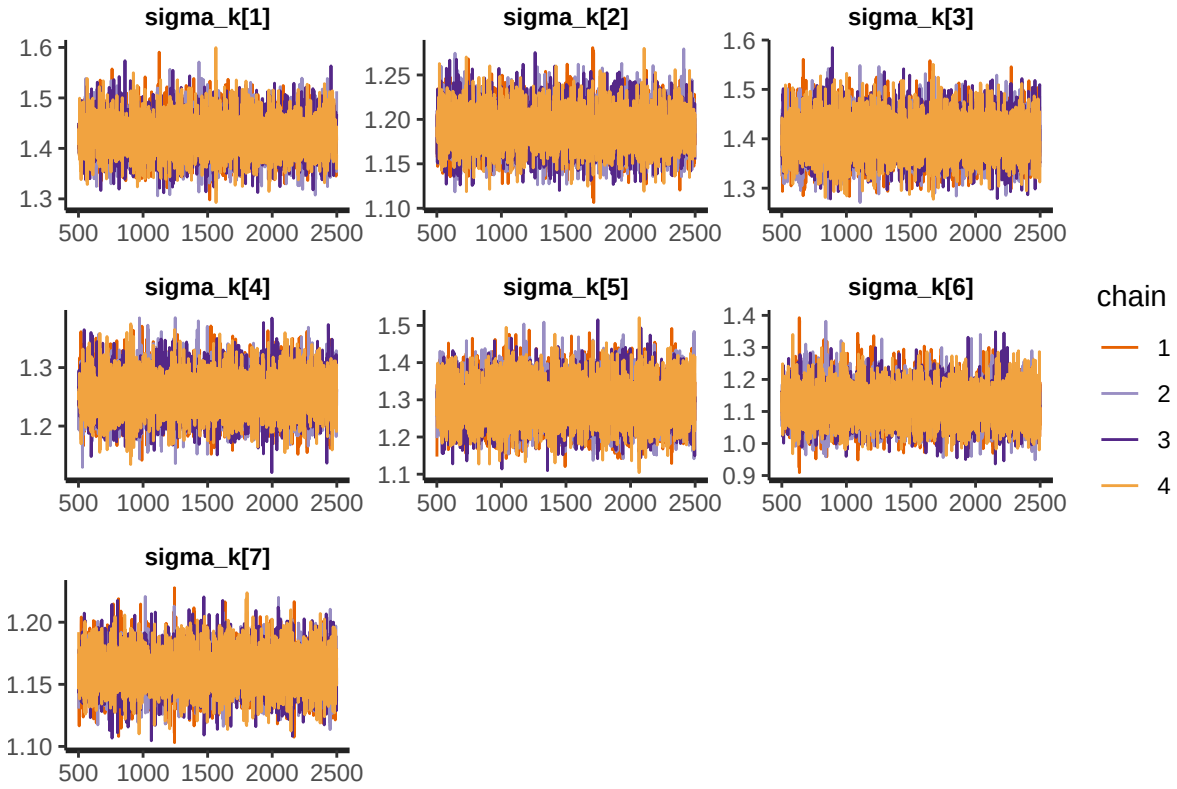




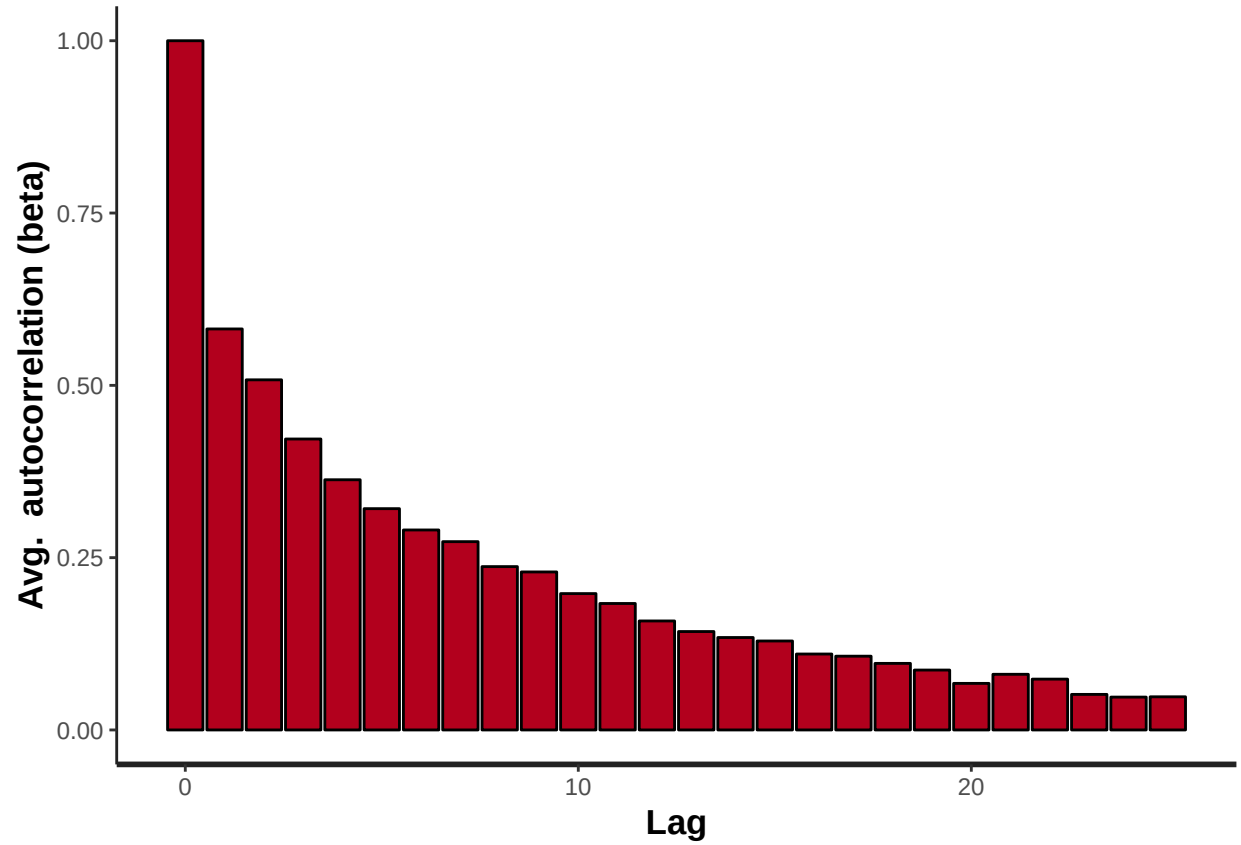

```
##                  mean      se_mean          sd      25%      50%      75%
## sigma_k[1] 1.424869 0.0003126555 0.03838279 1.399525 1.423677 1.449577
## sigma_k[2] 1.190219 0.0001862930 0.02358073 1.174096 1.189831 1.205875
```

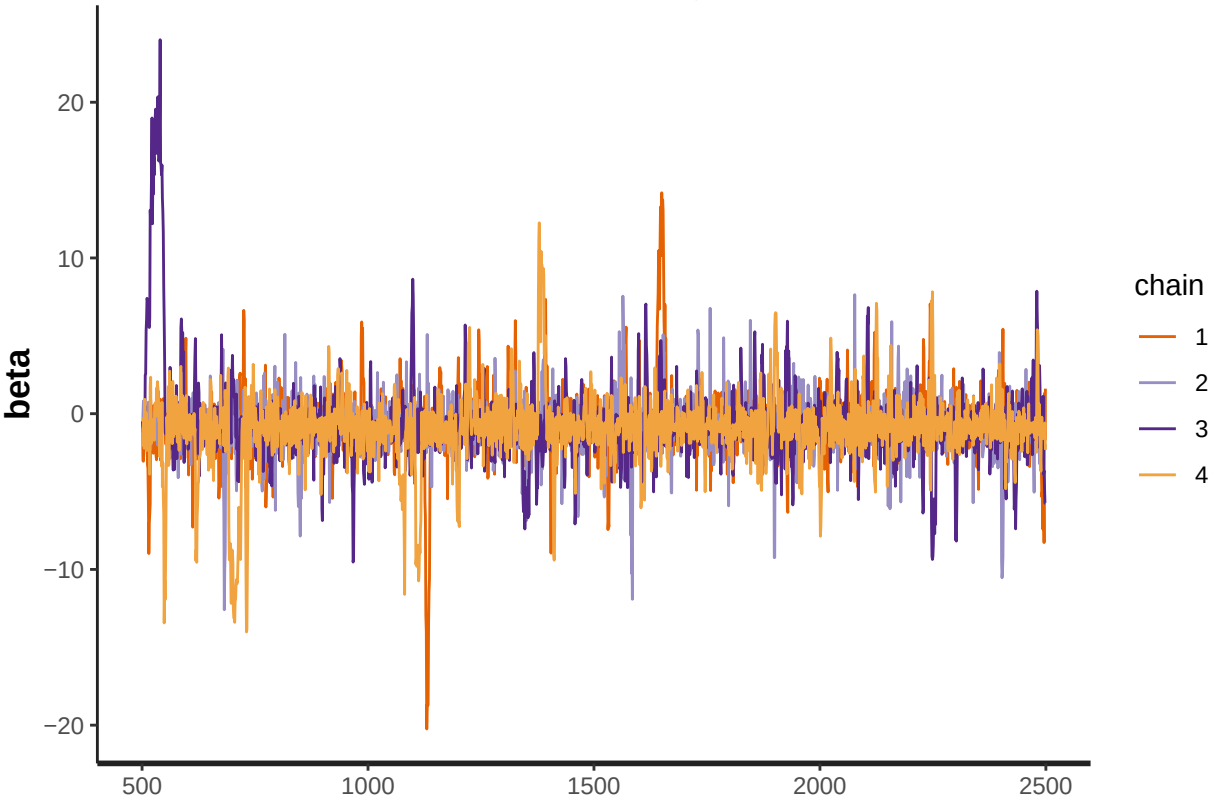
```
## sigma_k[3] 1.402604 0.0003434197 0.04246313 1.373491 1.401714 1.430532
## sigma_k[4] 1.253296 0.0002734876 0.03659210 1.228115 1.252198 1.277790
## sigma_k[5] 1.291914 0.0004870145 0.05911008 1.251370 1.289333 1.331149
## sigma_k[6] 1.119904 0.0004944651 0.05934707 1.078986 1.117127 1.158999
## sigma_k[7] 1.161318 0.0001263901 0.01670015 1.149788 1.161067 1.172417
##               n_eff      Rhat
## sigma_k[1] 15070.97 0.9997338
## sigma_k[2] 16022.18 0.9997046
## sigma_k[3] 15288.81 0.9996279
## sigma_k[4] 17901.91 0.9996729
## sigma_k[5] 14731.24 0.9997935
## sigma_k[6] 14405.46 0.9996860
## sigma_k[7] 17458.82 0.9997206
```





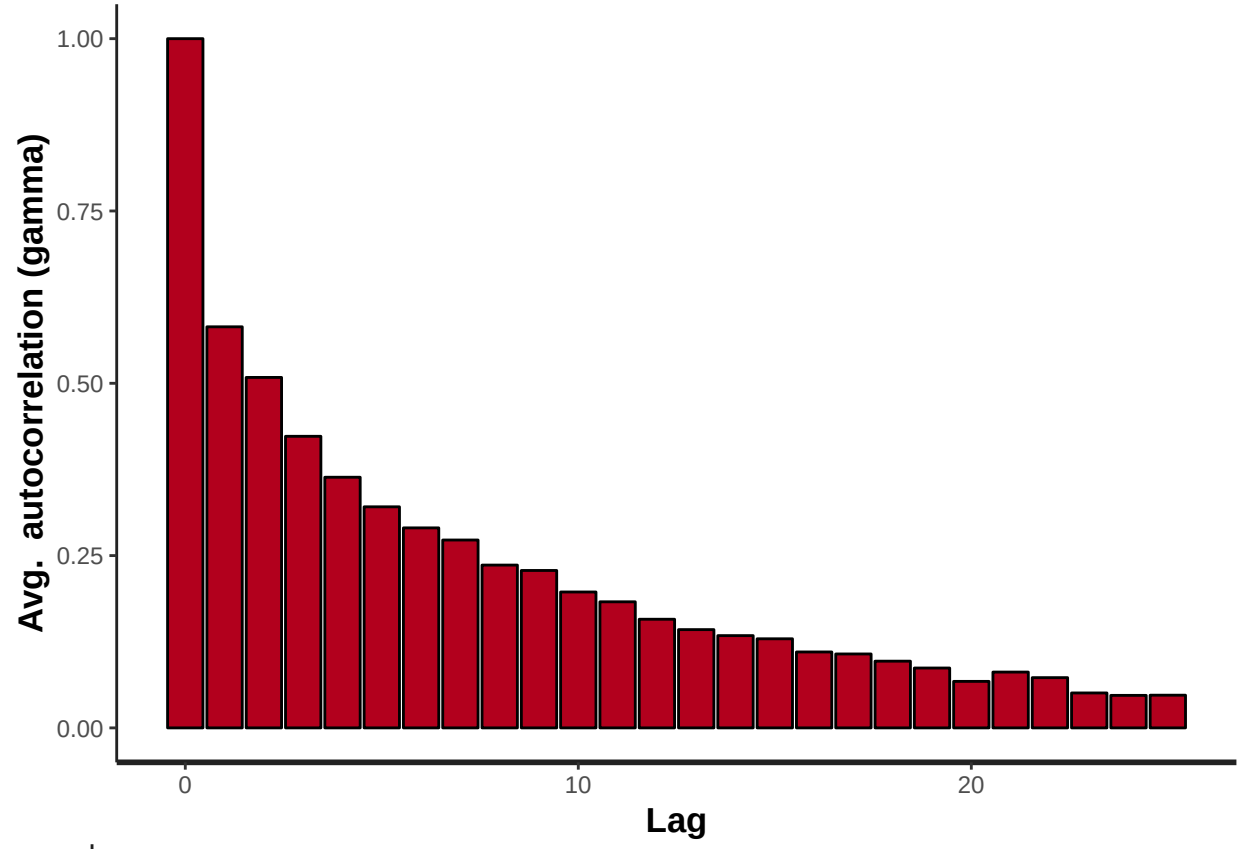

```
##              mean  se_mean       sd      25%       50%         75%    n_eff
## beta -0.7627081 0.107477 2.317934 -1.54405 -0.785995 -0.03670036 465.1264
##           Rhat
## beta 1.007923
```

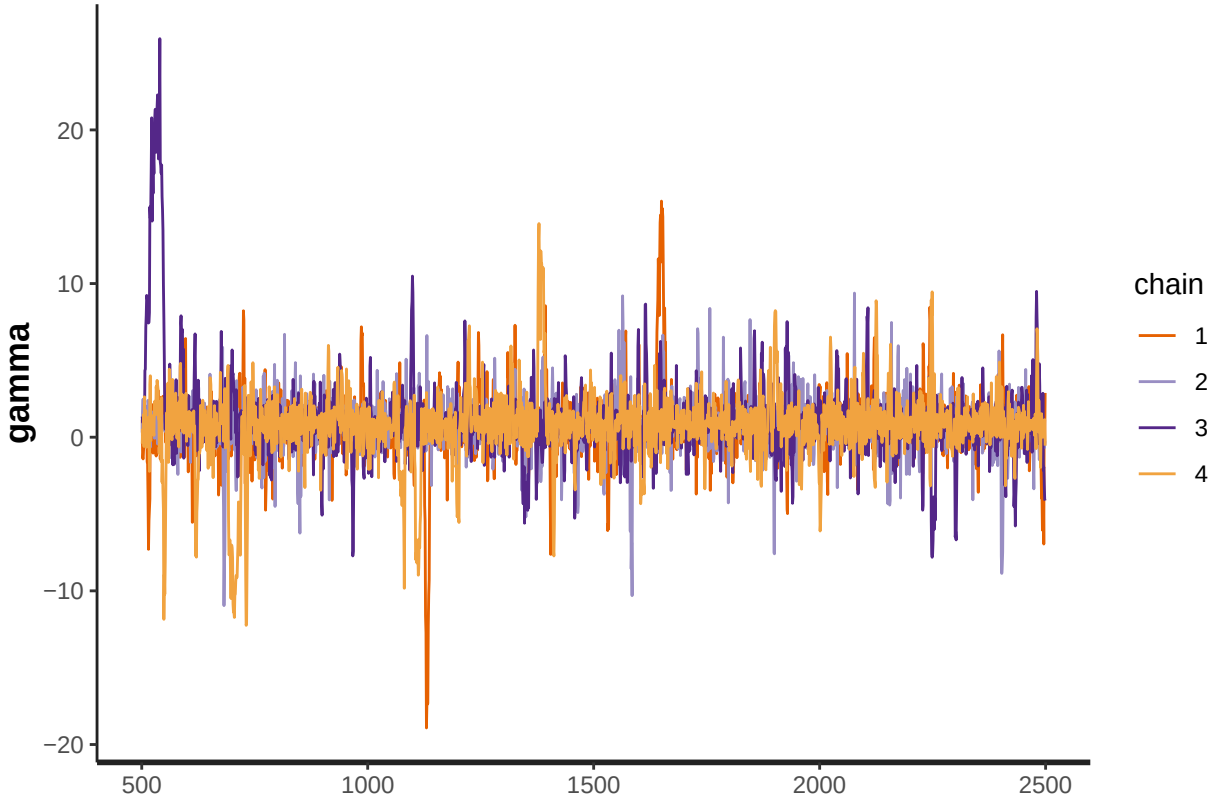


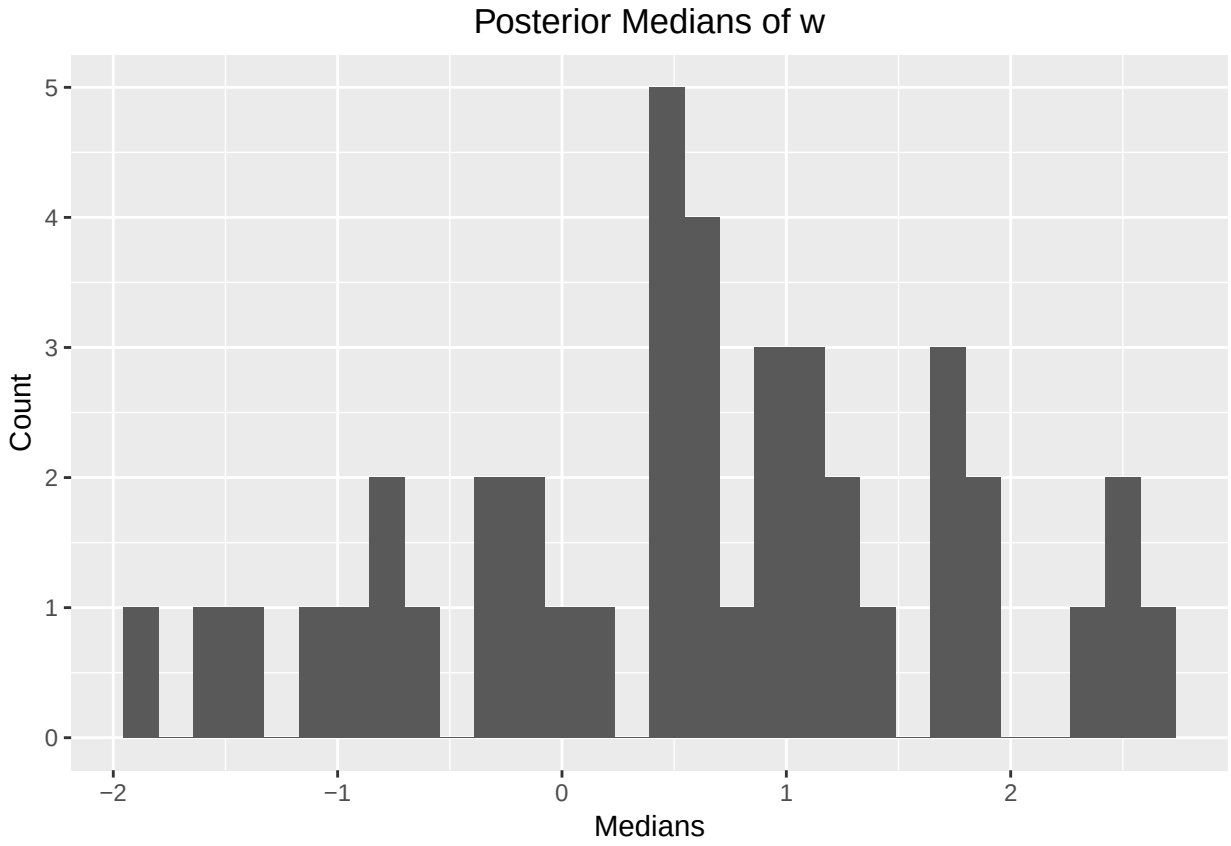



```
##              mean    se_mean       sd        25%       50%      75%    n_eff
## gamma 0.8456179 0.1082575 2.318785 0.06215549 0.8140468 1.575259 458.7802
```

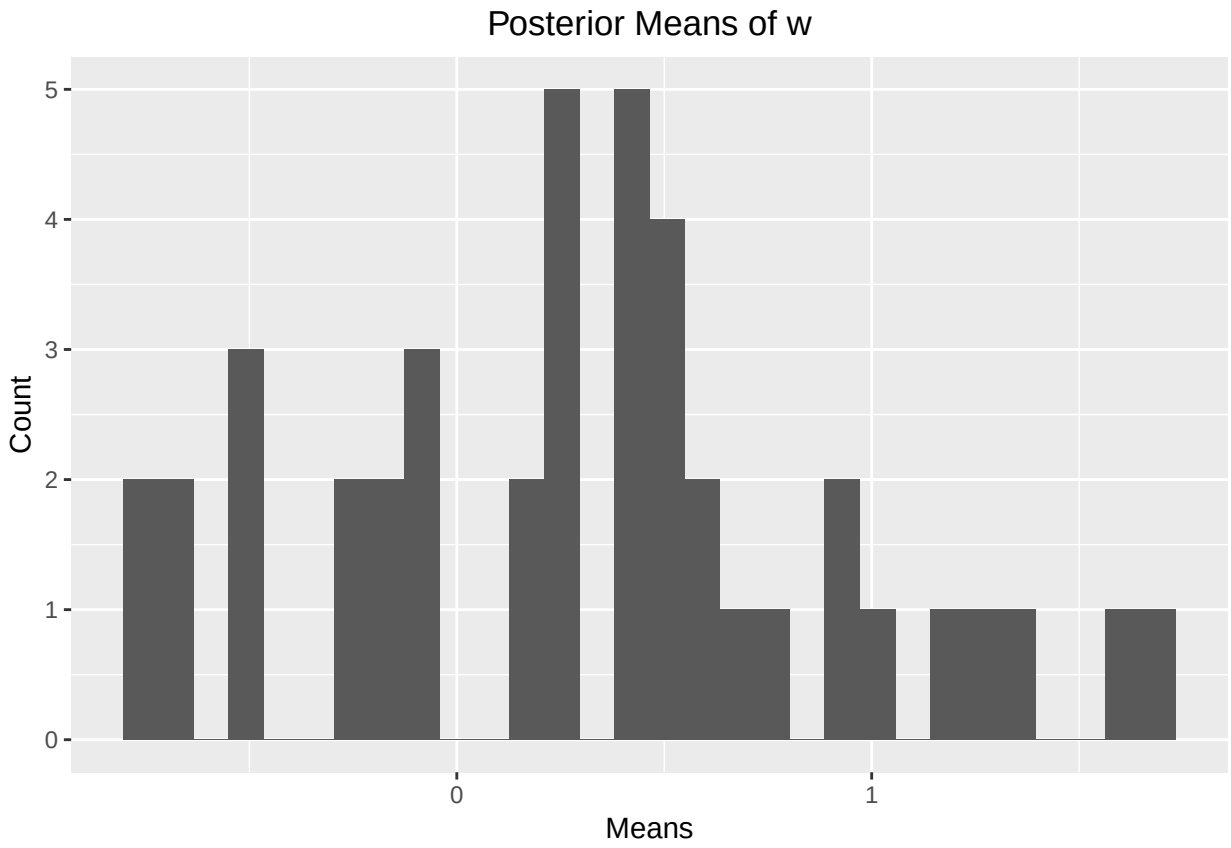
```
##           Rhat
## gamma 1.008998
```



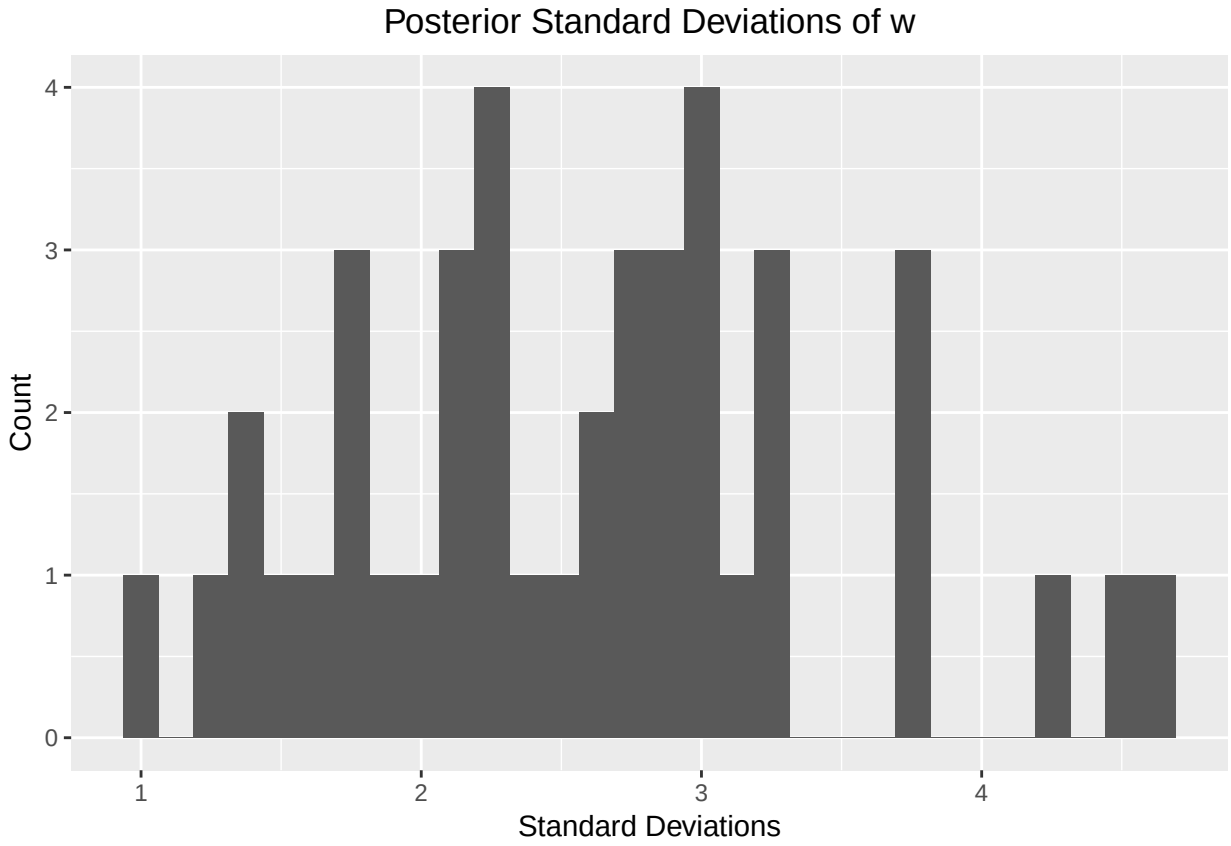


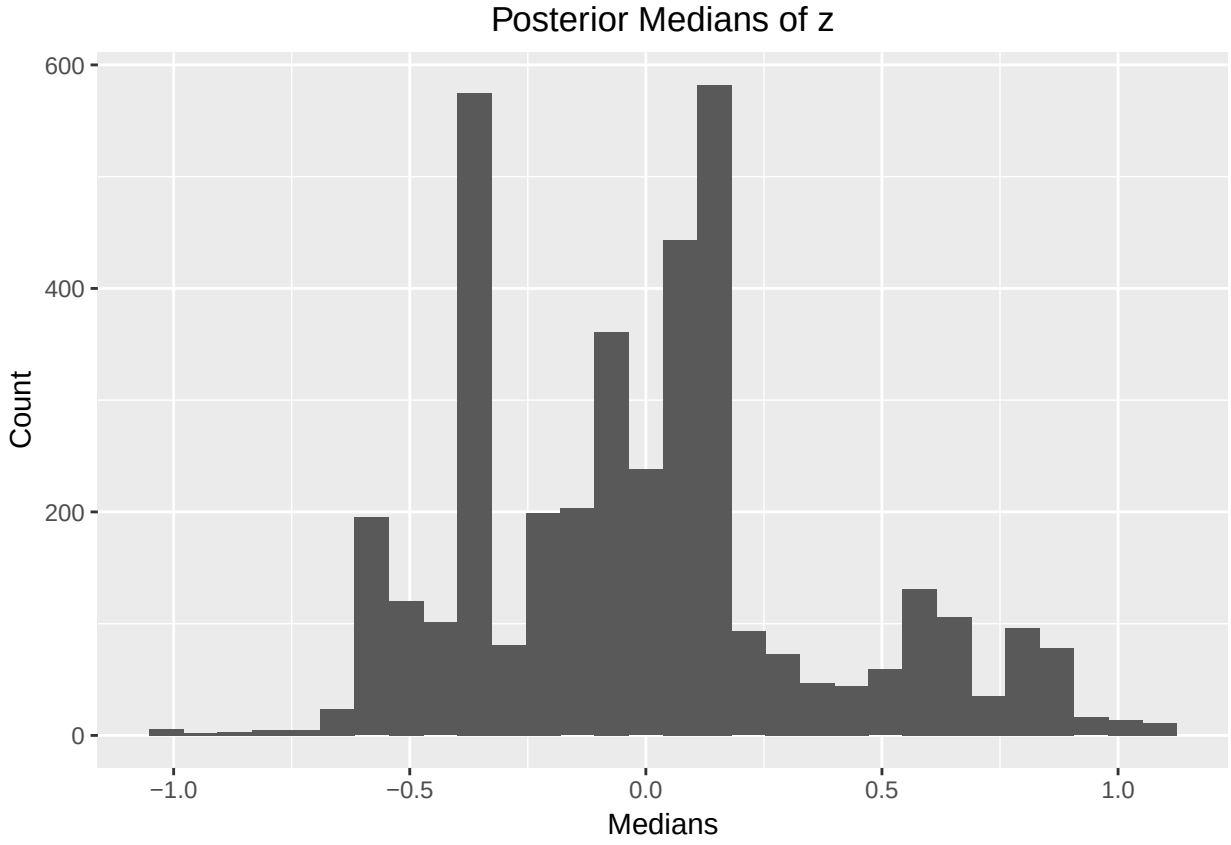



```
## [1] "       "
```

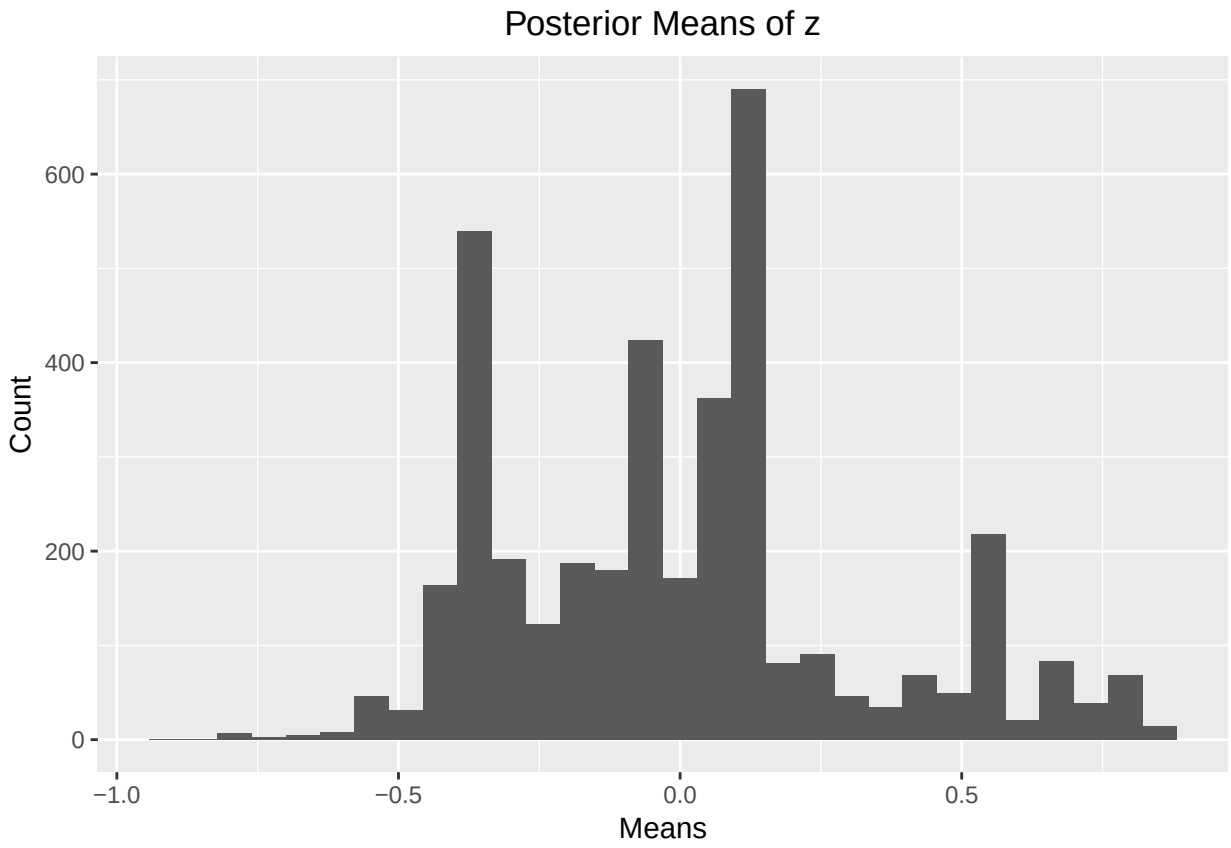



```
## [1] "      "
```

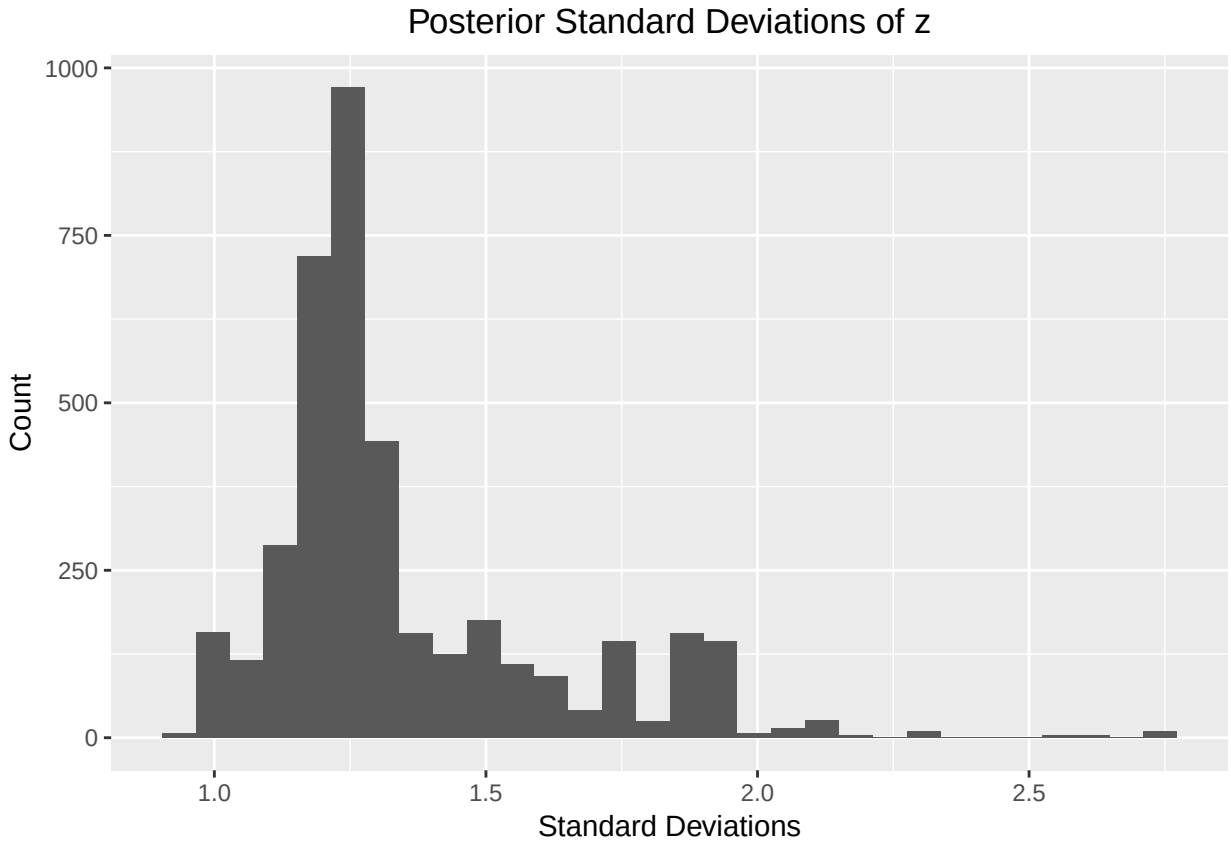

Posterior Standard Deviations of w
Count
0
1
2
3
4
1
2
3
4
Standard Deviations

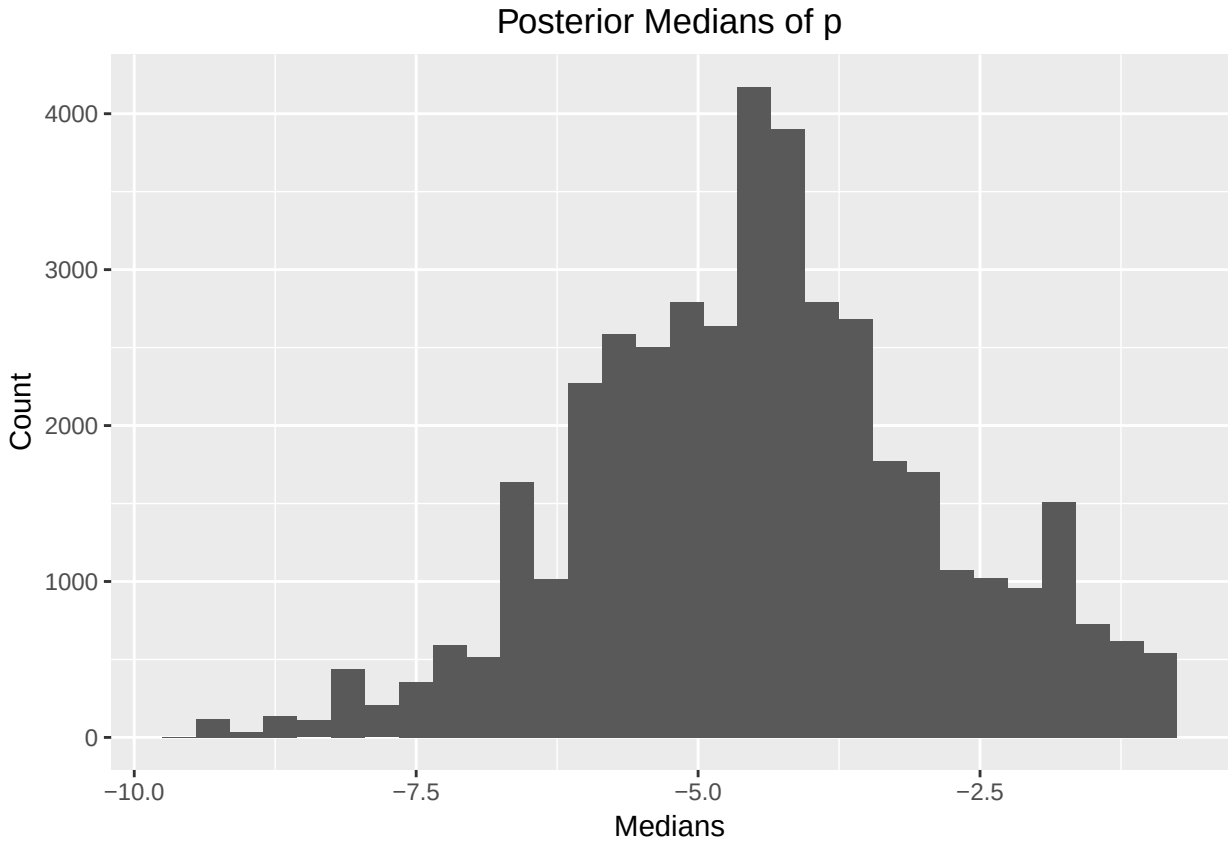

Posterior Medians of z
Count
0
200
400
600
−1.0
−0.5
0.0
0.5
1.0
Medians

```
## [1] "       "
```

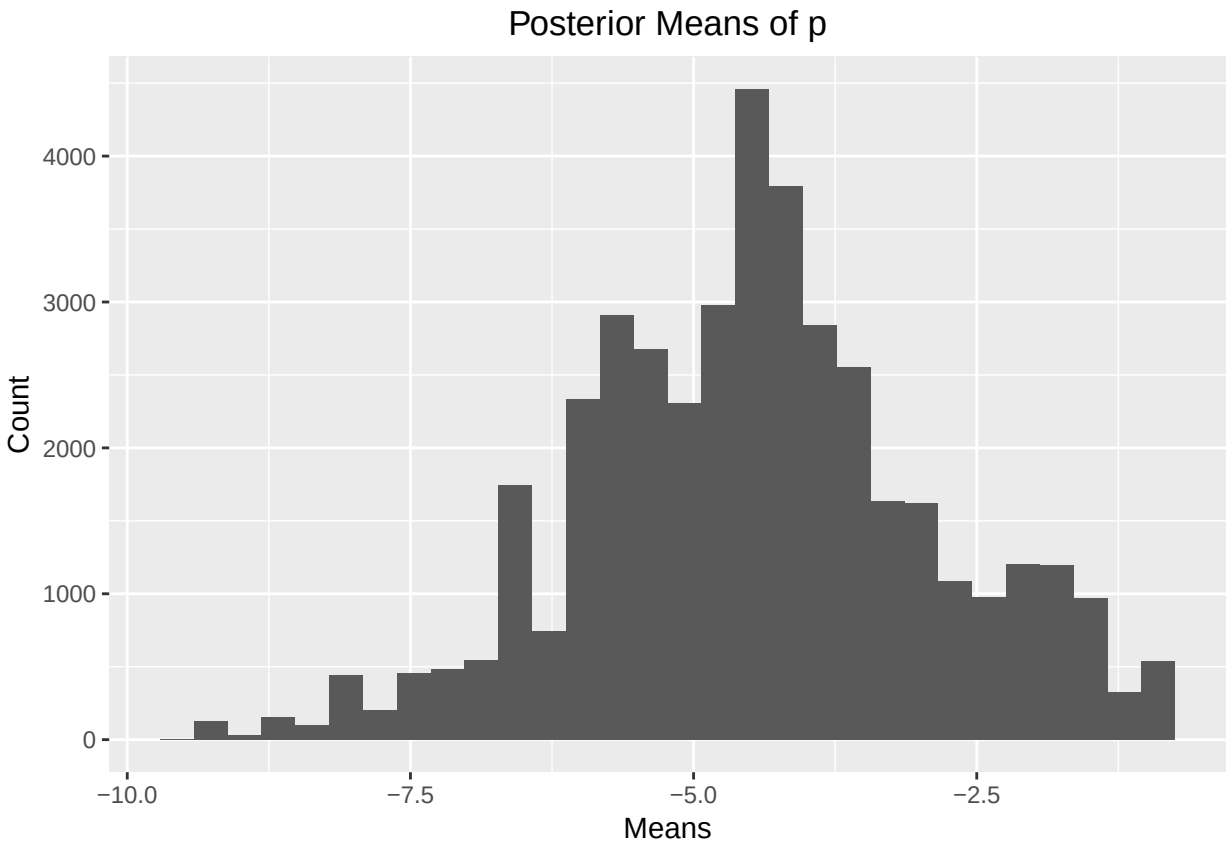



```
## [1] "       "
```

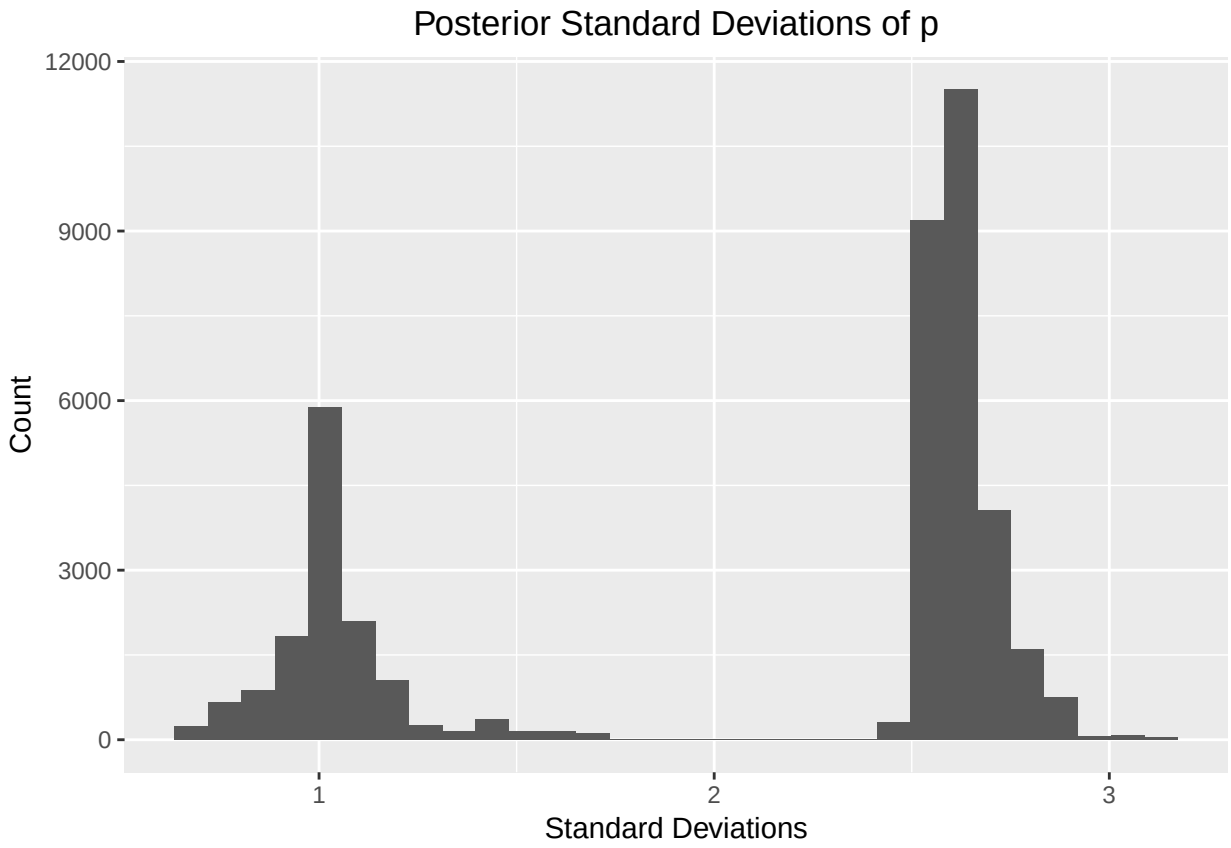

Posterior Standard Deviations of z
Count
1000
750
500
250
0
1.0
1.5
2.0
2.5
Standard Deviations


Posterior Medians of p
Count
4000
3000
2000
1000
0
−10.0
−7.5
−5.0
−2.5
Medians

```
## [1] "       "
```


Posterior Means of p
Count
4000
3000
2000
1000
0
−10.0
−7.5
−5.0
−2.5
Means

```
## [1] "       "
```


Posterior Standard Deviations of p
12000
9000
6000
3000
0
Count
1
2
3
Standard Deviations