MCMC Diagnostics - IFLS data

Sarah Teichman

04/30/2020

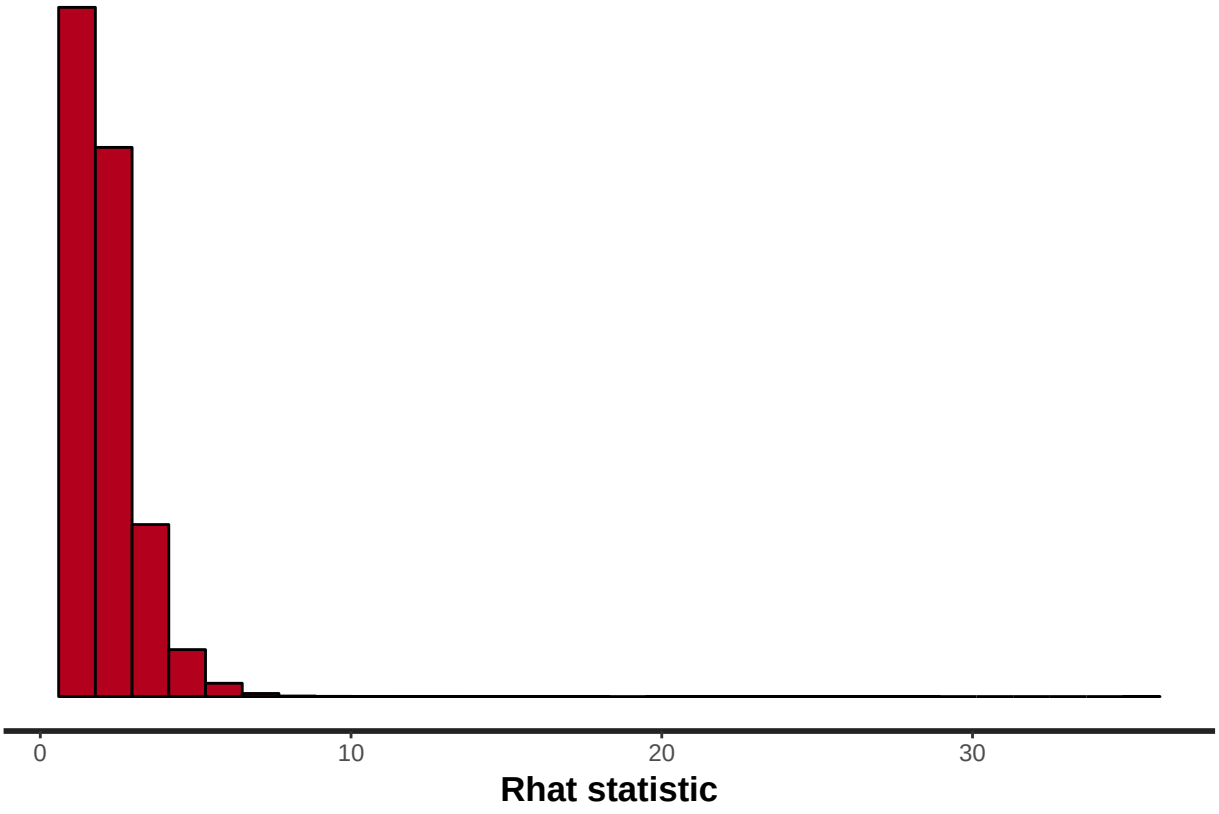
```
K <- 7
Ti <- 3
N <- 1973
```

General MCMC diagnostic plots

Overall model diagnostics from rstan package.

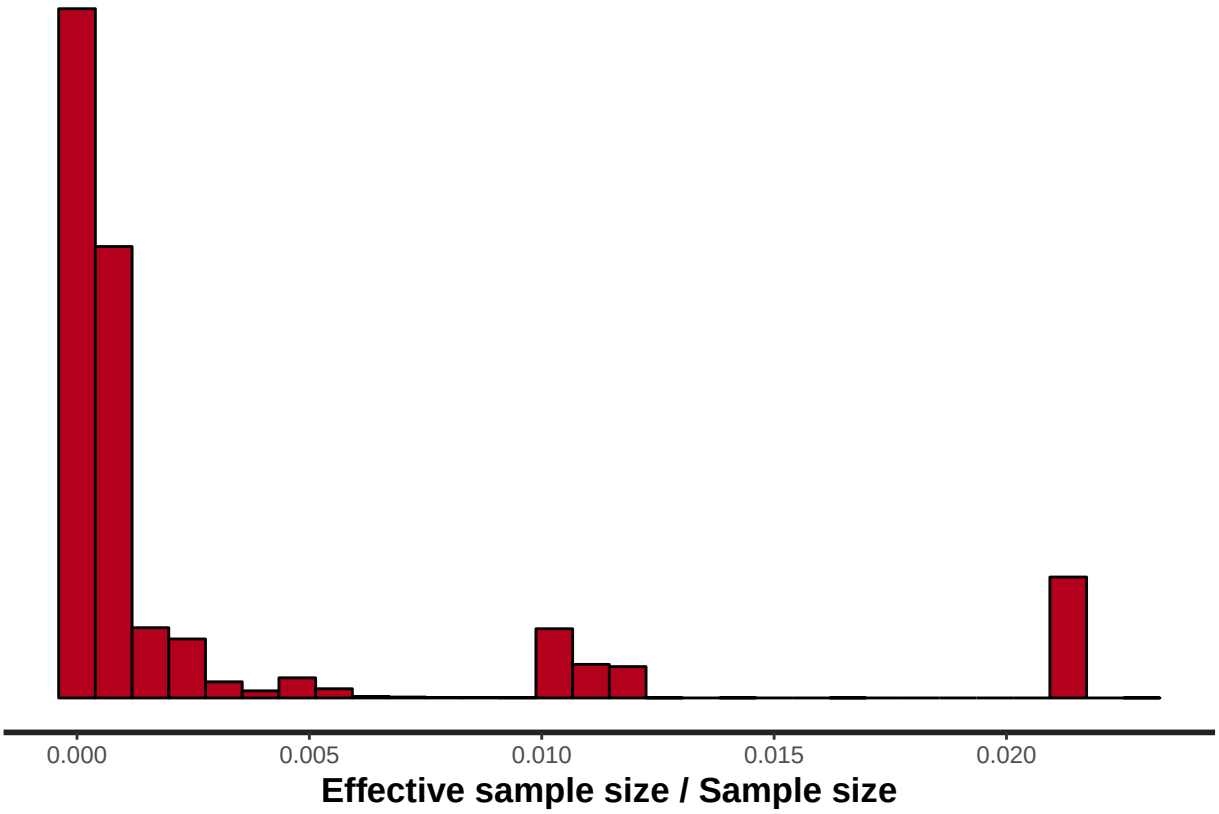
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
## Warning: Removed 2 rows containing non-finite values (stat_bin).
```



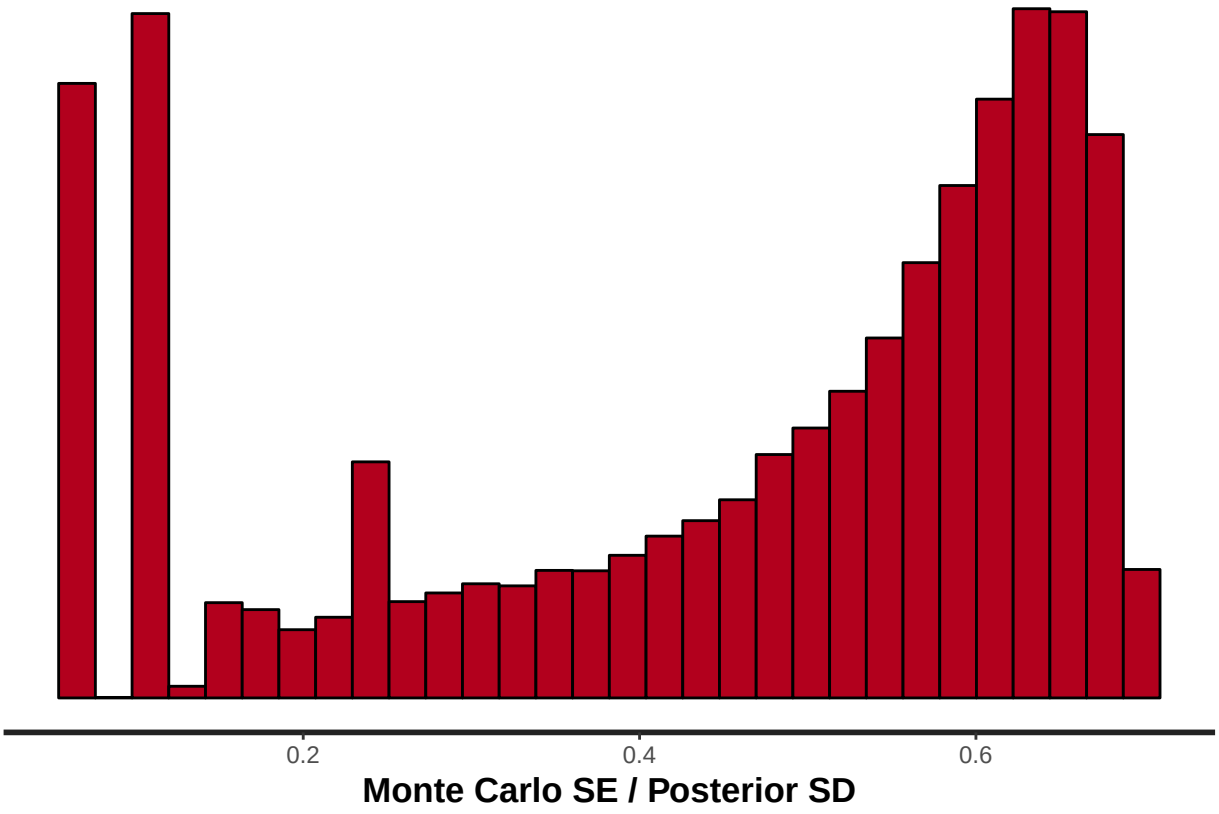

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
## Warning: Removed 2 rows containing non-finite values (stat_bin).
```




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

```
## Warning: Removed 2 rows containing non-finite values (stat_bin).
```



Individual Parameter Diagnostics

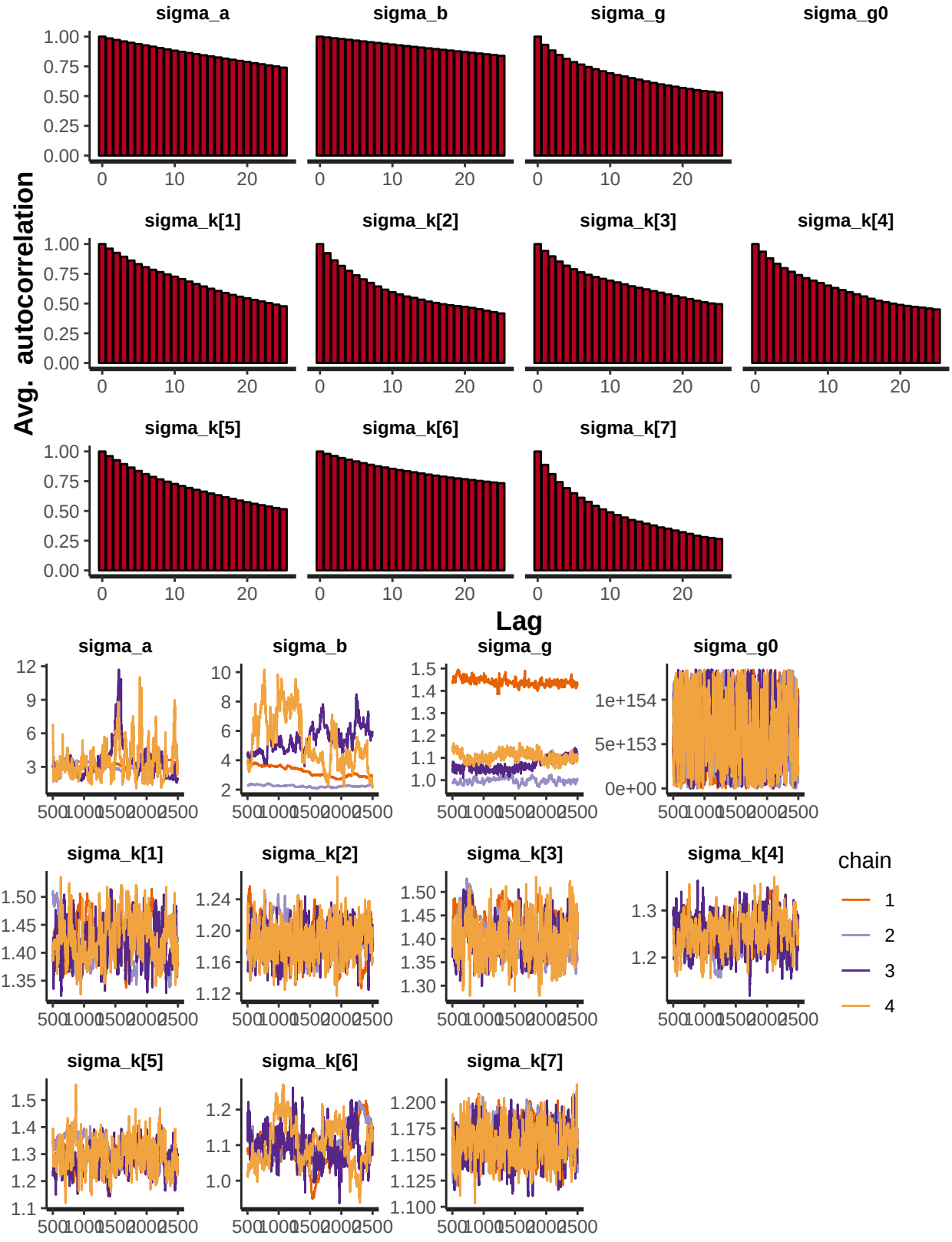
Individual parameter plots. Autocorrelation and trace plots for individual parameters, and histograms of posterior medians for group parameters.

```
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)$`50%`[ind])
  print(paste0("Summary statistics for posterior medians of ",param))
  print(summary(medians))
  title <- paste0("Posterior Medians of ",param)
  print(ggplot(medians, aes(x = avg)) + geom_histogram(bins = 60) + ggtitle(title))
  if (trim == T) {
    lim <- quantile(abs(medians$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),"%")))
  }
}
plot_fit <- function(fit) {
  get_single_plots(fit, sigma_params)
  get_single_plots(fit, beta)
```

```
  get_single_plots(fit, gamma)
  get_single_plots(fit, alpha)
  get_aggreg_plots(fit, "w")
  get_aggreg_plots(fit, "z")
  get_aggreg_plots(fit, "gamma_it")
  get_aggreg_plots(fit, "p")
}
plot_fit(fit)
```

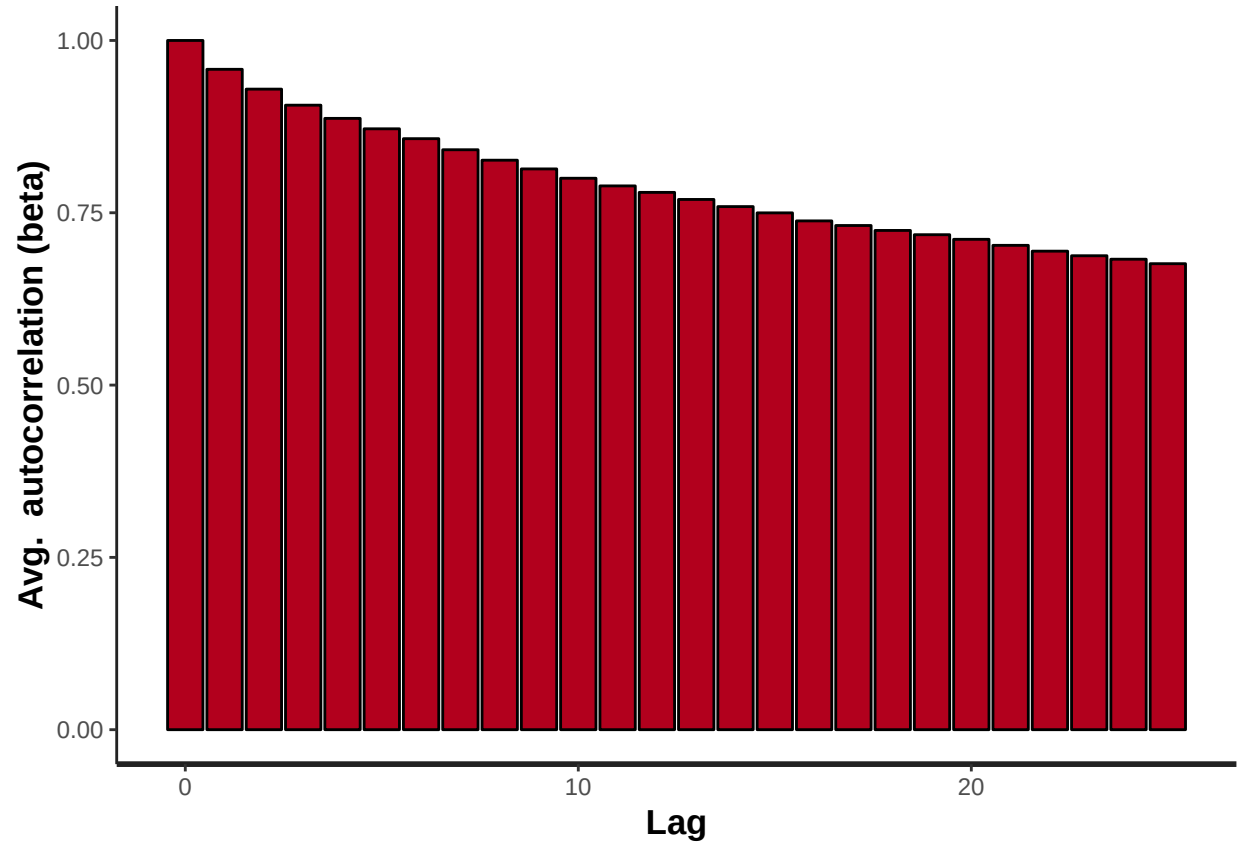
```
##                      mean     se_mean            sd           25%
## sigma_a      3.265677e+00 0.144158133  1.098957e+00  2.765422e+00
## sigma_b      4.055763e+00 0.940963942  1.702495e+00  2.475082e+00
## sigma_g      1.153199e+00 0.120467150  1.714019e-01  1.022594e+00
## sigma_g0   6.587623e+153         NaN 3.921638e+153 3.156491e+153
## sigma_k[1]   1.421888e+00 0.003969372  3.642606e-02  1.395548e+00
## sigma_k[2]   1.187516e+00 0.002419540  2.244841e-02  1.171265e+00
## sigma_k[3]   1.410643e+00 0.010180010  4.243788e-02  1.376933e+00
## sigma_k[4]   1.248610e+00 0.003556082  3.420344e-02  1.225341e+00
## sigma_k[5]   1.296761e+00 0.007580274  4.887123e-02  1.264975e+00
## sigma_k[6]   1.101167e+00 0.008823257  5.283775e-02  1.065404e+00
## sigma_k[7]   1.163920e+00 0.001264195  1.658818e-02  1.153217e+00
##                      50%           75%      n_eff      Rhat
## sigma_a      3.084065e+00  3.443387e+00  58.114360  1.059341
## sigma_b      3.620854e+00  5.264316e+00   3.273601  2.433936
## sigma_g      1.091902e+00  1.221684e+00   2.024388 10.819284
## sigma_g0   6.476256e+153 9.956583e+153        NaN       NaN
## sigma_k[1]   1.419992e+00  1.446708e+00  84.213349  1.055356
## sigma_k[2]   1.187630e+00  1.203000e+00  86.080588  1.029805
## sigma_k[3]   1.411809e+00  1.441534e+00  17.378446  1.231346
## sigma_k[4]   1.248866e+00  1.271147e+00  92.511590  1.046571
## sigma_k[5]   1.292108e+00  1.331151e+00  41.565863  1.154241
## sigma_k[6]   1.094827e+00  1.132726e+00  35.861688  1.058950
## sigma_k[7]   1.164591e+00  1.175332e+00 172.174739  1.027836
```

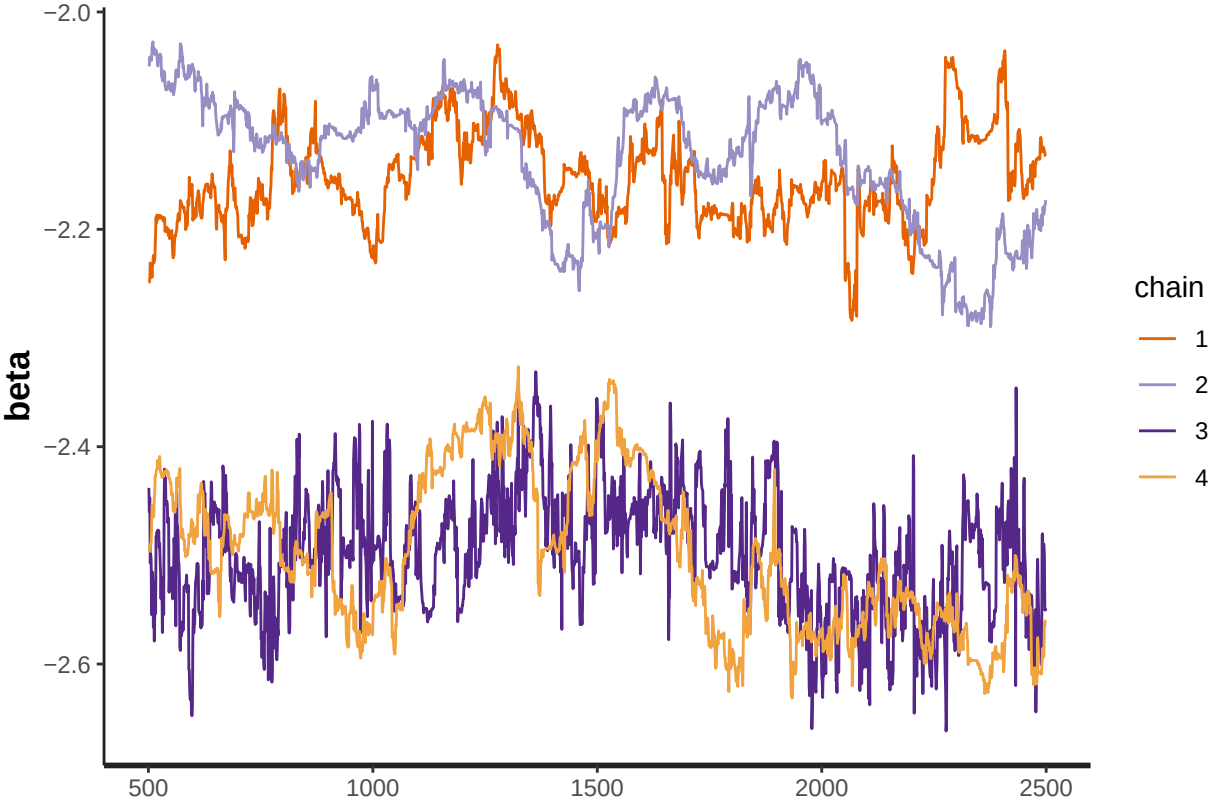
```
## Warning: Removed 104 rows containing non-finite values (stat_summary).
```



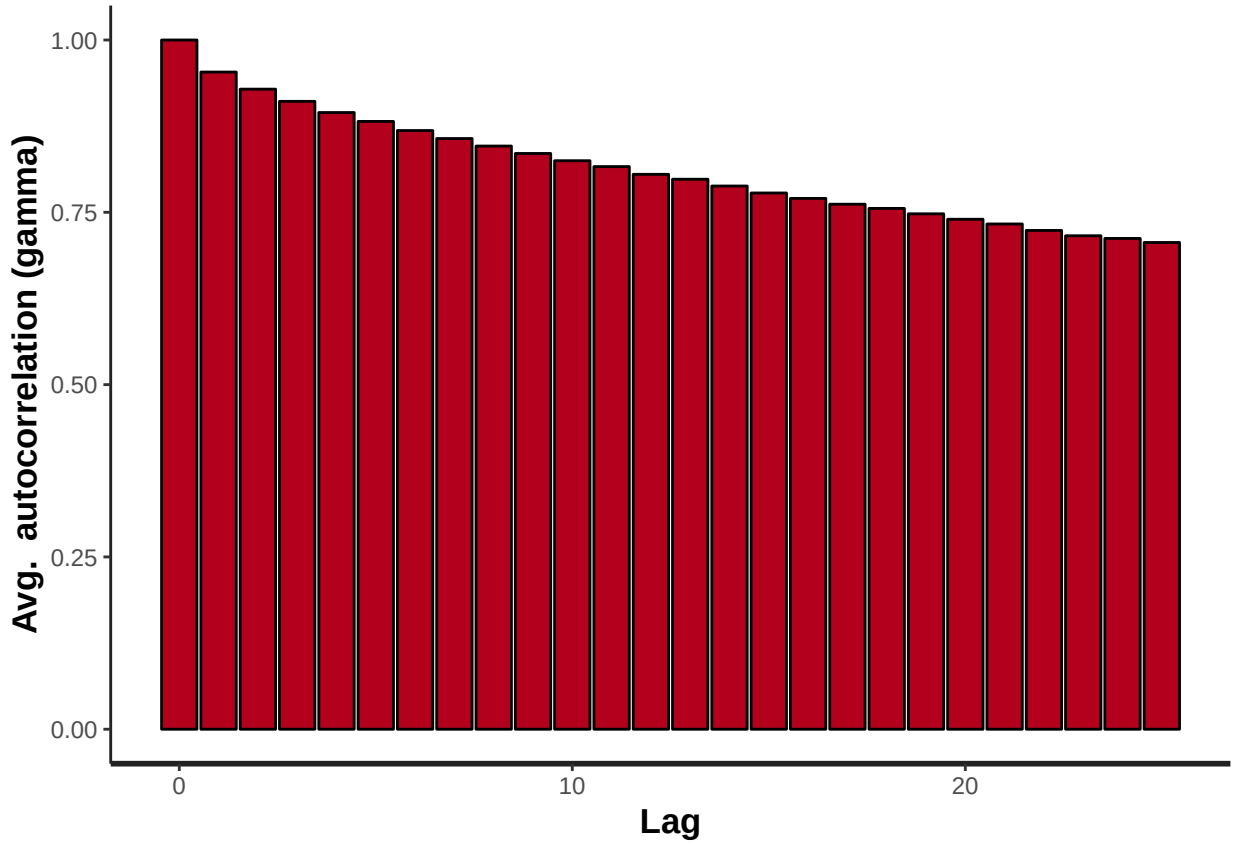

```
##           mean    se_mean        sd       25%       50%       75%    n_eff
## beta -2.320691 0.1258875 0.1867678 -2.500263 -2.308036 -2.144254 2.201094
##          Rhat
```

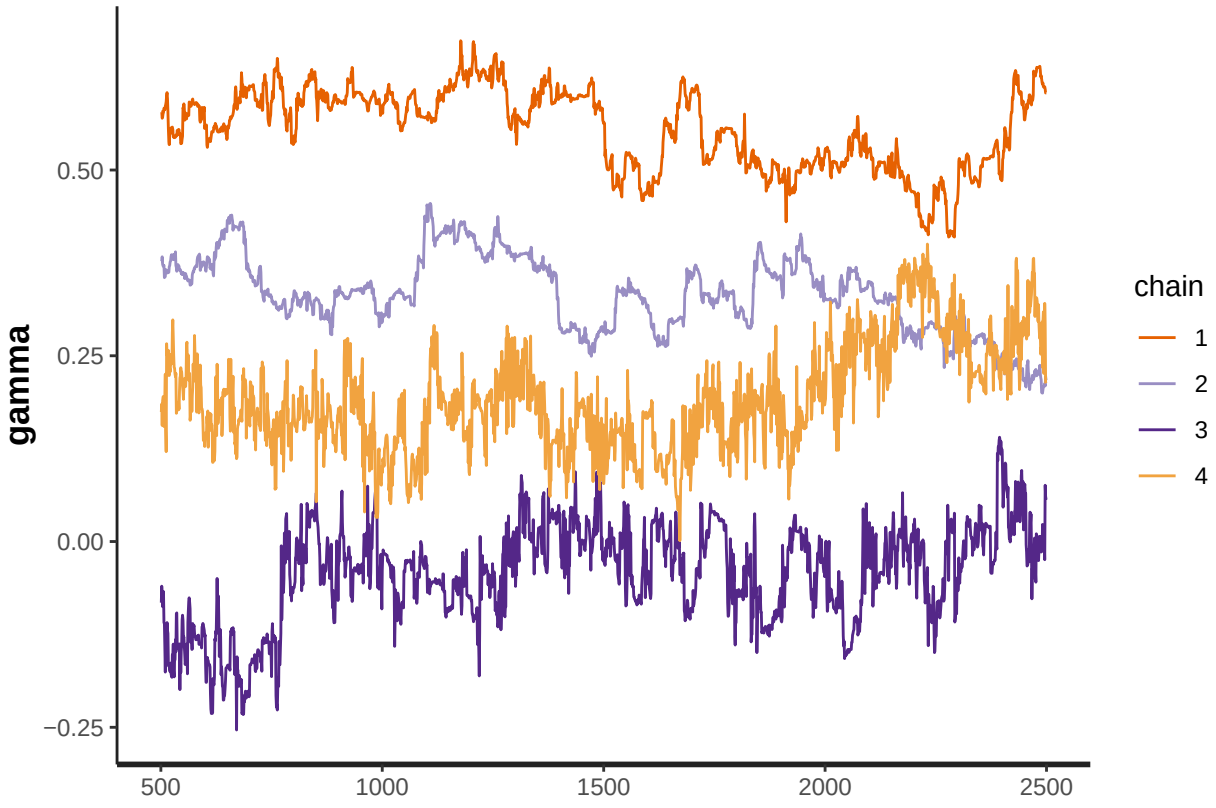
```
## beta 3.642755
```



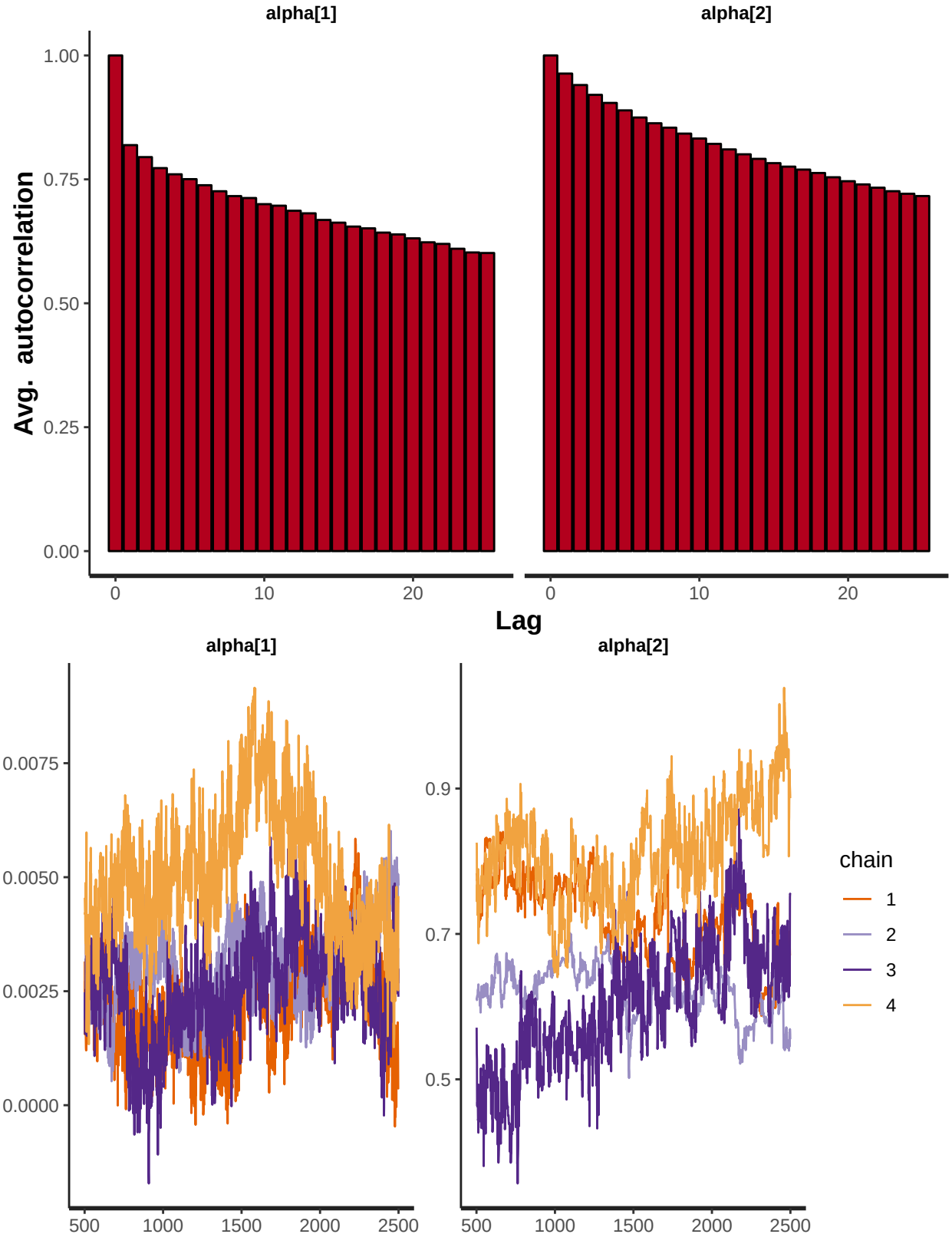


```
##           mean   se_mean        sd        25%       50%       75%    n_eff
## gamma 0.259342 0.1540008 0.2255122 0.07143885 0.2752672 0.4301783 2.144343
##           Rhat
## gamma 4.406152
```



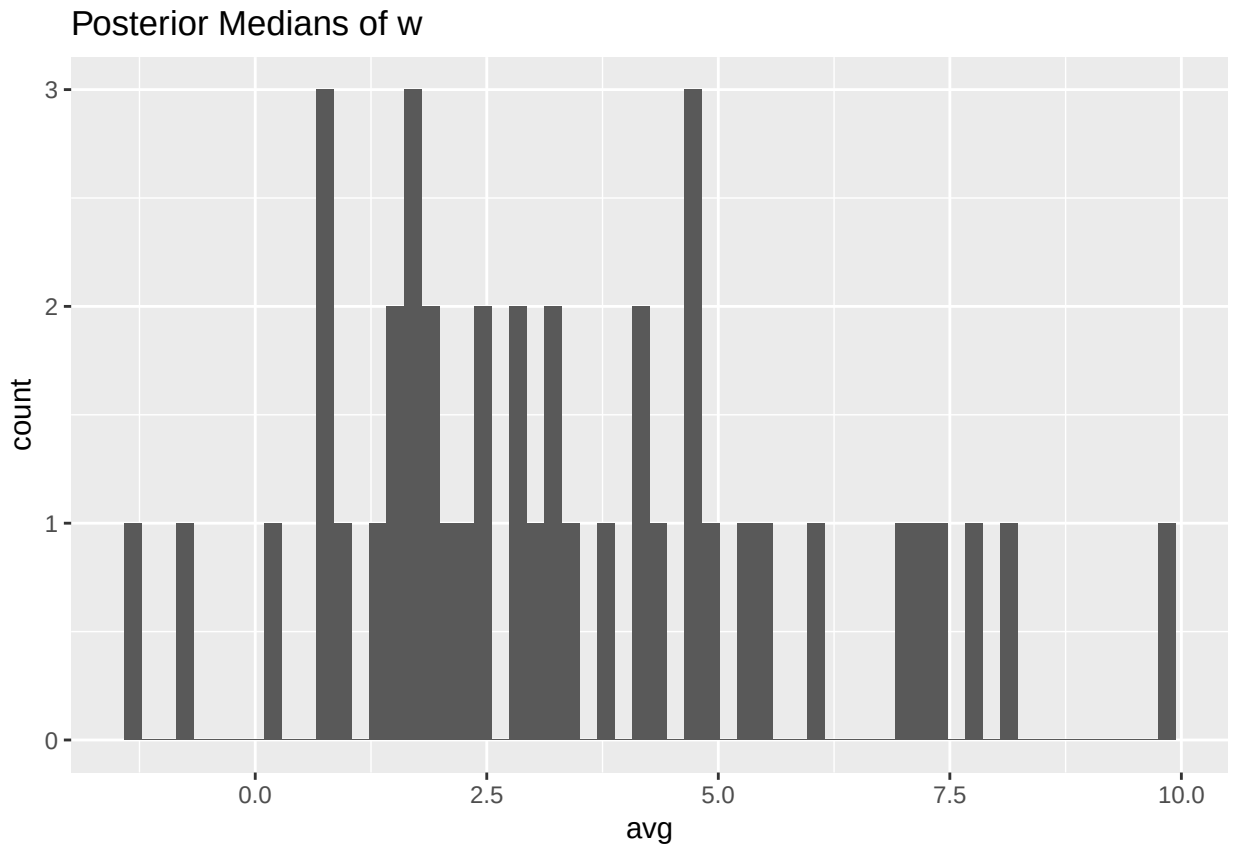



```
##    mean se_mean sd 25% 50% 75% n_eff Rhat
## NA   NA      NA NA  NA  NA  NA    NA   NA
```

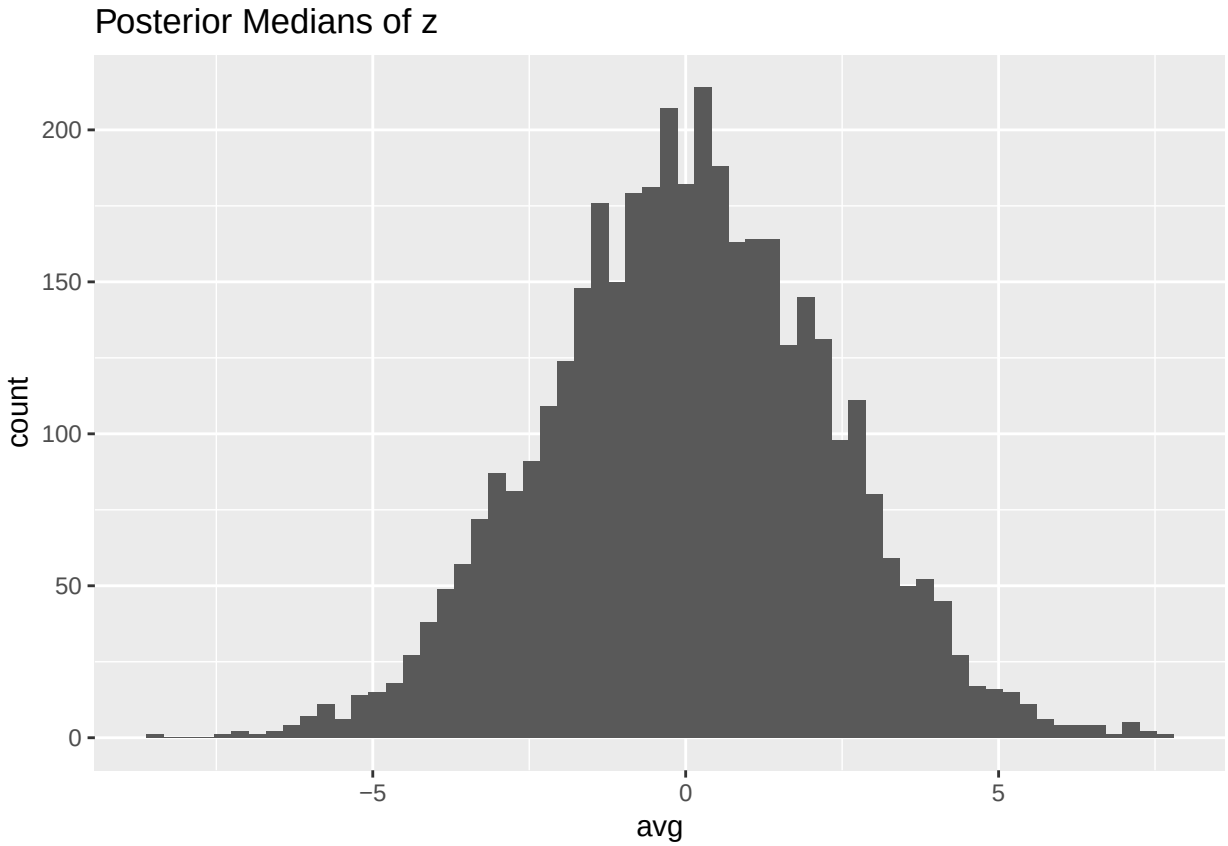



```
## [1] "Summary statistics for posterior medians of w"
##       avg
##  Min.   :-1.254
```

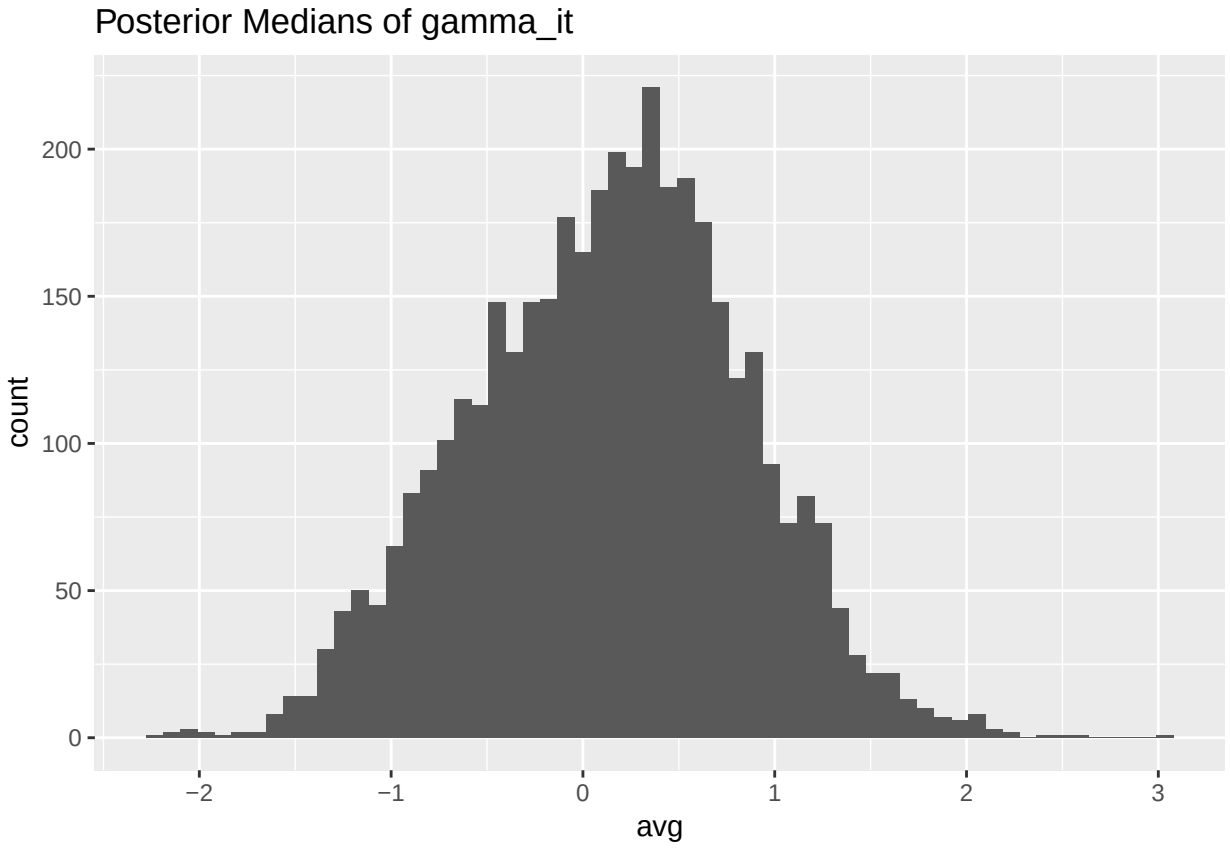
```
##  1st Qu.: 1.673
##  Median : 2.925
##  Mean   : 3.381
##  3rd Qu.: 4.668
##  Max.   : 9.913
```



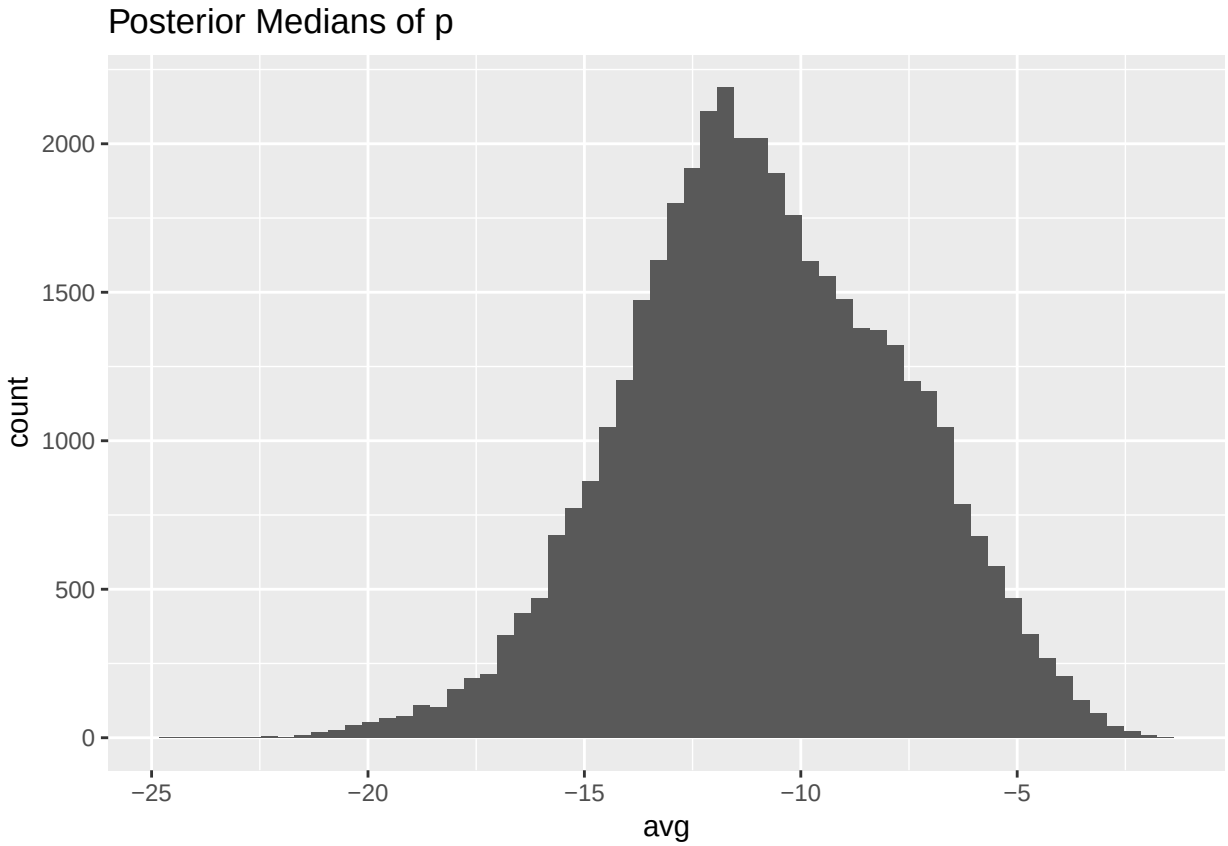

```
## [1] "Summary statistics for posterior medians of z"
##       avg
##  Min.   :-8.53602
##  1st Qu.:-1.46561
##  Median : 0.01939
##  Mean   : 0.03222
##  3rd Qu.: 1.56139
##  Max.   : 7.60880
```




```
## [1] "Summary statistics for posterior medians of gamma_it"
##       avg
##  Min.   :-2.1956
##  1st Qu.:-0.3570
##  Median : 0.1661
##  Mean   : 0.1356
##  3rd Qu.: 0.6185
##  Max.   : 3.0723
```




```
## [1] "Summary statistics for posterior medians of p"
##       avg
##  Min.   :-24.717
##  1st Qu.:-13.010
##  Median :-11.009
##  Mean   :-10.880
##  3rd Qu.: -8.592
##  Max.   : -1.647
```


Posterior Medians of p
count
2000
1500
1000
500
0
−25
−20
−15
−10
−5
avg