Assignment 4

Jeff Gould

3/30/2021

Analysis

1

$$
\begin{aligned}
P(\beta|y_i) &= \mathcal{L}(y_i|\beta)\pi(\beta) \\
&\propto \exp\left[\sum[y_i\beta - \log(1+e^\beta)]\right] \\
\log P(\beta|y_i) &\propto \log\left(\exp\left[\sum[y_i\beta - \log(1+e^\beta)]\right]\right) \\
&= \sum[y_i\beta - \log(1+e^\beta)] \\
&= \sum y_i\beta - n\log(1_e^\beta)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial l}{\partial \beta} &= \sum y_i - \frac{ne^\beta}{1+e^\beta} \overset{set}{=} 0 \\
\sum y_i &= \frac{ne^\beta}{1+e^\beta} \\
\sum y_i &= e^\beta(n - \sum y_i) \\
\hat{\beta} &= \log\left(\frac{\sum y_i}{n - \sum y_i}\right)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 l}{\partial \beta^2} &= \frac{ne^\beta}{1+e^\beta} - \frac{ne^{2\beta}}{(1+e^\beta)^2} \\
&= -\frac{ne^\beta}{(1+e^\beta)^2}
\end{aligned}
$$

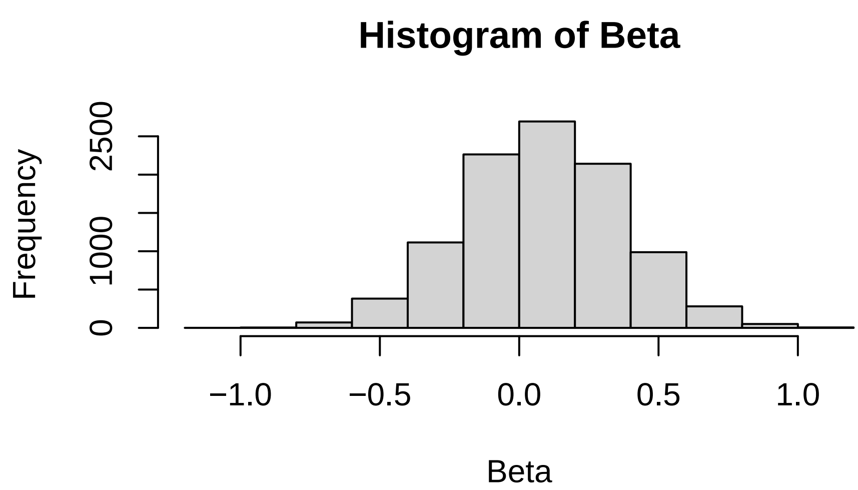
$I(\beta) = \frac{ne^\beta}{(1+e^\beta)^2}$

$$I(\hat{\beta}) = \frac{ne^{\log\left(\frac{\sum y_i}{n-\sum y_i}\right)}}{\left(1+e^{\log\left(\frac{\sum y_i}{n-\sum y_i}\right)}\right)^2}$$

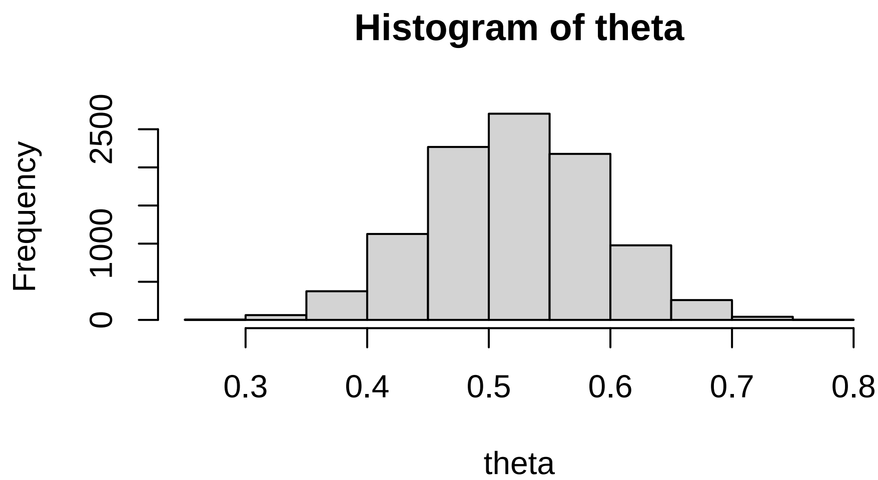
$$= \frac{n\left(\frac{\sum y_i}{n-\sum y_i}\right)}{\left(1+\left(\frac{\sum y_i}{n-\sum y_i}\right)\right)^2}$$

$\beta|y_i \dot{\sim} \mathcal{N}(\hat{\beta}, I(\hat{\beta}))$, where $\hat{\beta}$ and $I(\hat{\beta})$ are as defined above

```
forestfire <- read.delim("forestfire.txt")
n <- nrow(forestfire)

beta_hat <- log(sum(forestfire) / (n - sum(forestfire)))
var <- ((n * exp(beta_hat)) / (1 + exp(beta_hat)) - (2 * n * exp(beta_hat)) /
          ((1 + exp(beta_hat))^2))^(-1)


set.seed(812)
Beta <- rnorm(10000, mean = beta_hat, sd = sqrt(var))
hist(Beta)
```




```
theta <- exp(Beta) / ( 1+ exp(Beta))

hist(theta)
```



Jeffrey's Prior

$\pi(\theta) \propto [J(\theta)]^{1/2}$

$J[\theta] = -E\left[\frac{\partial^2 \log \mathcal{L}(Y|\theta)}{\partial^2 \theta}|\theta\right]$

$\mathcal{L} \propto \exp\left\{\sum_{i=1}^n y_i\beta - \log(1+e^\beta)\right\}$

$\log \mathcal{L} \propto \sum_{i=1}^n y_i\beta - \log(1+e^\beta) = \sum y_i\beta - n\log(1+e^\beta)$

$\frac{\partial \log \mathcal{L}}{\partial \beta} = \sum y_i - \frac{ne^\beta}{1+e^\beta}$

$\frac{\partial^2}{\partial \beta^2} = \frac{ne^\beta}{1+e^\beta} - \frac{ne^{2\beta}}{(1+e^\beta)^2} = -\frac{ne^\beta}{(1+e^\beta)^2}$

$J(\theta) = -E\left[-\frac{ne^\beta}{(1+e^\beta)^2}|\beta\right] = \frac{ne^\beta}{(1+e^\beta)^2}$

$\pi(\beta) \propto \left(\frac{ne^\beta}{(1+e^\beta)^2}\right)^{1/2} \propto \frac{e^{\beta/2}}{1+e^\beta}$

$P(\beta|Y) \propto \exp\{\sum y_i\beta - \log(1+e^\beta)\}\frac{e^{\beta/2}}{1+e^\beta}$

Take the log:

$\sum(y_i\beta) - n\log(1+e^\beta) + \log\left(\frac{e^{\beta/2}}{1+e^\beta}\right) = \sum(y_i\beta) - n\log(1+e^\beta) + \frac{\beta}{2} - \log(1+e^\beta) = \sum(y_i\beta) - (n+1)\log(1+e^\beta) + \frac{\beta}{2}$

$\frac{\partial l}{\partial \beta} = \sum y_i - \frac{(n+1)e^\beta}{1+e^\beta} + \frac{1}{2} \overset{set}{=} 0$

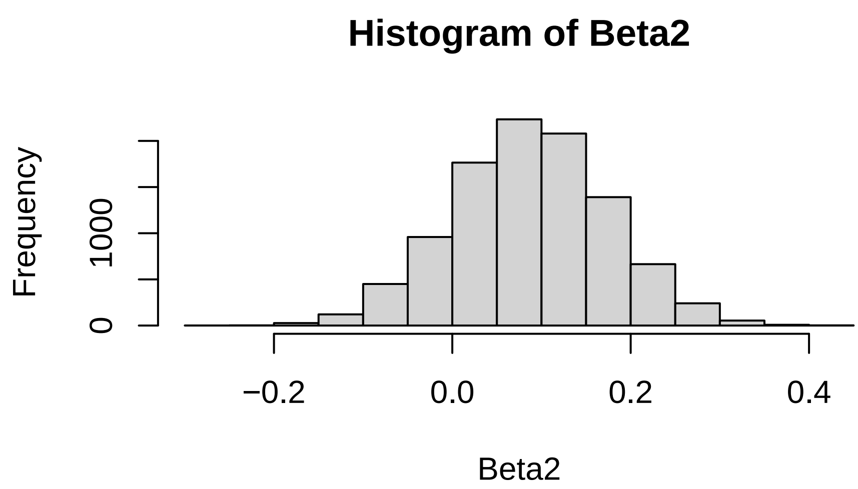
$$\begin{aligned}
\sum y_i + \frac{1}{2} &= \frac{(n+1)e^\beta}{1+e^\beta} \\
(n+1)e^\beta &= (1+e^\beta)(\sum y_i + 1/2) = \sum y_i + 1/2 + e^\beta \sum y_i + e^\beta/2 \\
e^\beta(n+1) - e^\beta \sum y_i - e^\beta(1/2) &= \sum y_i + 1/2 \\
e^\beta &= \frac{\sum y_i + 1/2}{(n+1) - \sum y_i - 1/2} \\
\hat{\beta} &= \log\left(\frac{\sum y_i + 1/2}{(n+1) - \sum y_i - 1/2}\right)
\end{aligned}$$

$$
\begin{aligned}
\frac{\partial^2}{\partial \beta^2} &= [(n+1)e^{\beta}][-1(1+e^{\beta})^{-2}e^{\beta}] + [(1+e^{\beta})^{-1}][(n+1)e^{\beta}] \\
&= \frac{(n+1)e^{\beta}}{1+e^{\beta}} - \frac{(n+1)e^{2\beta}}{(1+e^{\beta})^2} \\
&= -\frac{(n+1)e^{\beta}}{(1+e^{\beta})^2}
\end{aligned}
$$

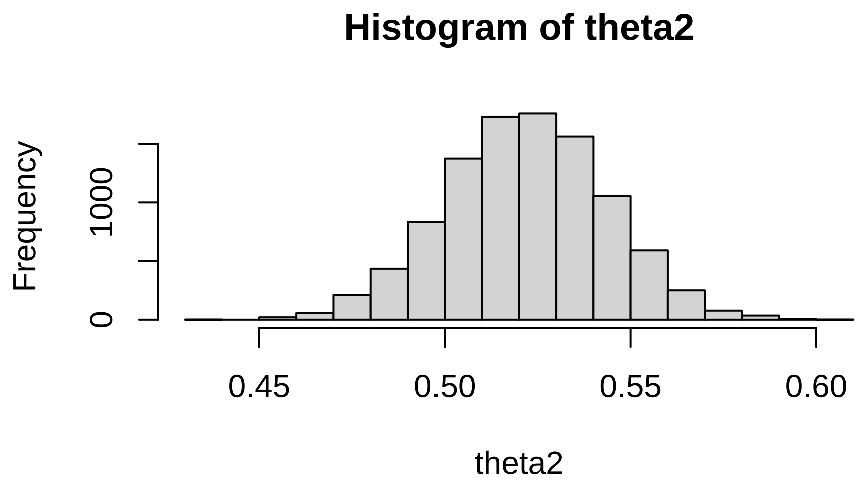
$I(\hat{\beta}) = -\frac{d^2}{d\beta^2}\log[p(\beta|y)] = \frac{(n+1)e^{\hat{\beta}}}{(1+e^{\hat{\beta}})^2}$

$$
P(\beta|y) \dot{\sim} \mathcal{N}\left(\log\left(\frac{\sum y_i + 1/2}{(n+1) - \sum y_i - 1/2}\right), \frac{(1+e^{\hat{\beta}})^2}{(n+1)e^{\hat{\beta}}}\right)
$$

```r
beta_hat <- log(
  (sum(forestfire) + 1/2) / (n+1 - sum(forestfire) - 1/2)
  )
var <- ((1 + exp(beta_hat))^2) / ((n+1) * exp(beta_hat))


set.seed(812)
Beta2 <- rnorm(10000, mean = beta_hat, sd = sqrt(var))
hist(Beta2)
```




```r
theta2 <- exp(Beta2) / ( 1+ exp(Beta2))

hist(theta2)
```




```r
# mean(theta)
# mean(theta2)
# sd(theta)^2
# sd(theta2)^2

mean(Beta)
```

```
## [1] 0.08424518
```

```r
mean(Beta2)
```

```
## [1] 0.08740173
```

```r
sd(Beta)^2
```

```
## [1] 0.08106654
```

```r
sd(Beta2)^2
```

```
## [1] 0.007534076
```

We do find that the model is sensitive to the choice of prior, as the second model, made using Jeffrey's Prior, has a much narrower variance than that of the first model. The means are pretty comparable, but the distribution is approximately on order of magnitude smaller for the second model

2

First let's build a model using all input variables

```r
library(mvtnorm)
bikeshare <- read.table("bikeshare.txt", header = T)

casual_model_1 <- glm(casual ~ yr  + workingday  + atemp + hum  + holiday + temp + windspeed,
                      data = bikeshare, family = poisson(link = "log"))
summary(casual_model_1)
```

```
##
## Call:
## glm(formula = casual ~ yr + workingday + atemp + hum + holiday +
##     temp + windspeed, family = poisson(link = "log"), data = bikeshare)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -53.563   -9.164   -1.384    5.818   44.896
##
## Coefficients:
##              Estimate Std. Error  z value Pr(>|z|)
## (Intercept)  6.293572   0.009885  636.705  < 2e-16 ***
## yr           0.353792   0.002611  135.496  < 2e-16 ***
## workingday  -0.876070   0.002590 -338.285  < 2e-16 ***
## atemp        2.347946   0.075263   31.196  < 2e-16 ***
## hum         -0.730709   0.009925  -73.624  < 2e-16 ***
## holiday     -0.244973   0.006946  -35.270  < 2e-16 ***
## temp         0.469600   0.066165    7.097 1.27e-12 ***
## windspeed   -1.170926   0.018862  -62.079  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 378438  on 730  degrees of freedom
## Residual deviance: 111688  on 723  degrees of freedom
## AIC: 117688
##
## Number of Fisher Scoring iterations: 5
```

In the summary, we see that `temp` and `holiday` are the two weakest variables, using the z-value. This makes intuitive sense, and there is a high correlation between `temp` and `atemp`, and most of the signal from `holiday` likely is found within `workingday`

```r
casual_model_2 <- glm(casual ~ yr  + workingday  + atemp + hum + windspeed,
                      data = bikeshare, family = poisson(link = "log"))
summary(casual_model_2)
```

```
##
## Call:
## glm(formula = casual ~ yr + workingday + atemp + hum + windspeed,
##     family = poisson(link = "log"), data = bikeshare)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
```

```
## -53.252   -9.499   -1.361    5.968   47.296
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  6.258586   0.009543  655.85   <2e-16 ***
## yr           0.352687   0.002611  135.09   <2e-16 ***
## workingday  -0.856483   0.002542 -336.89   <2e-16 ***
## atemp        2.882460   0.008474  340.17   <2e-16 ***
## hum         -0.745906   0.009910  -75.27   <2e-16 ***
## windspeed   -1.148244   0.018762  -61.20   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 378438  on 730  degrees of freedom
## Residual deviance: 113052  on 725  degrees of freedom
## AIC: 119048
##
## Number of Fisher Scoring iterations: 5
```

Now we see higher magnitude z-values, and notice for `atemp` it climbed from 31.196 to 340.17, comfirming our belief that `atempt` and `temp` were highly correlated.

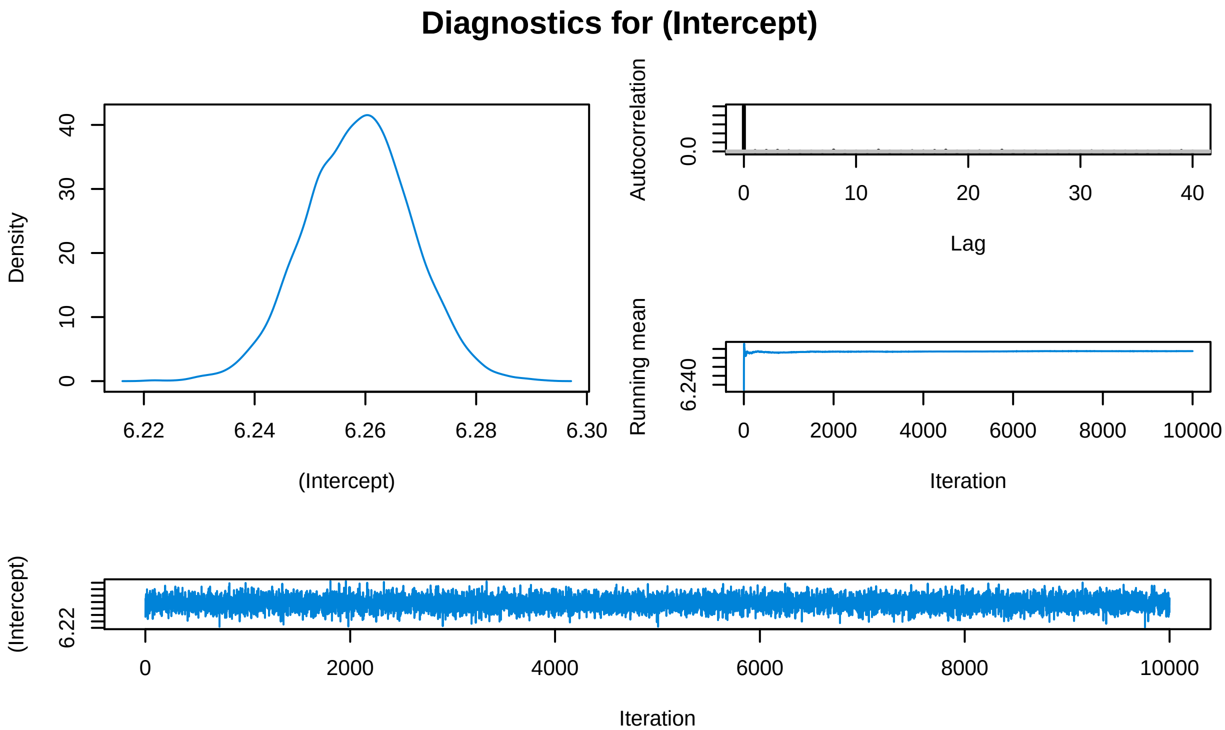
Now let's look at the confidence intervals for the coefficients:

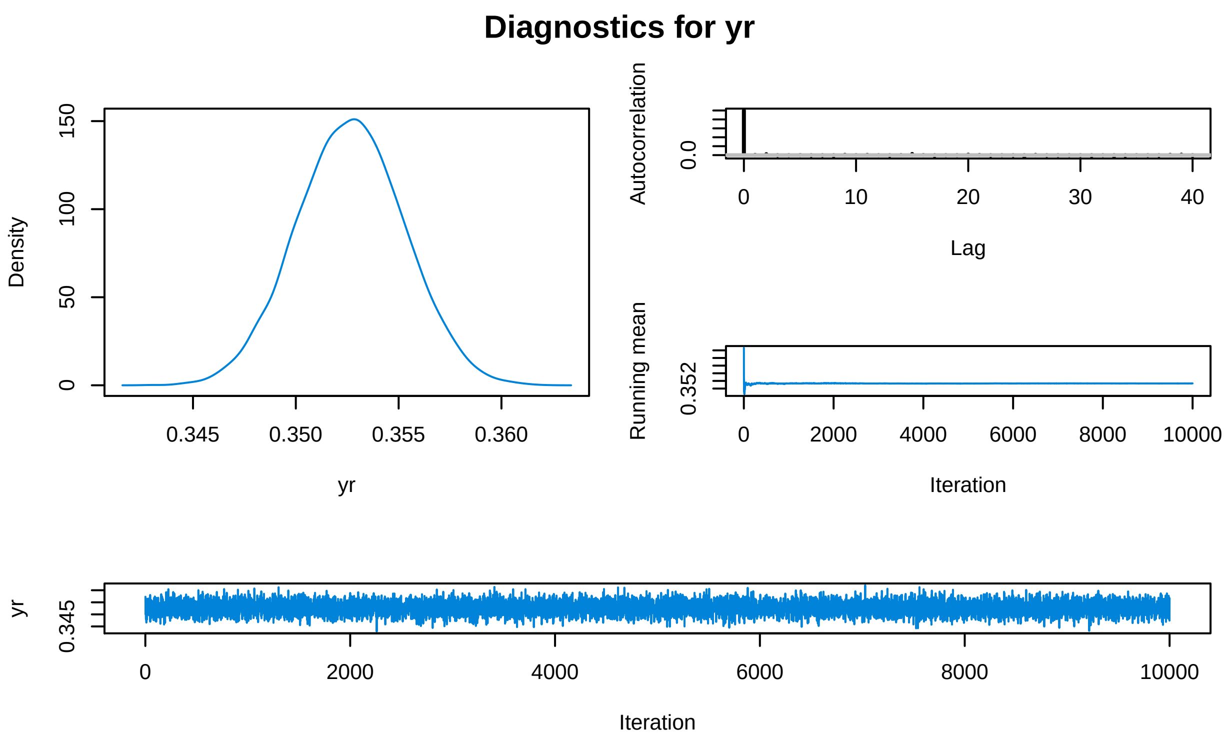
```r
bhat    <- coef(casual_model_2)
vbeta   <- vcov(casual_model_2)
B       <- 10000
set.seed(1959)
beta    <- rmvnorm(B, mean = bhat, sigma = vbeta)

round(t(apply(beta, 2, quantile, probs = c(0.5, 0.025, 0.975))), 4)
```

```
##                 50%    2.5%   97.5%
## (Intercept)  6.2590  6.2399  6.2773
## yr           0.3527  0.3476  0.3578
## workingday  -0.8566 -0.8615 -0.8515
## atemp        2.8824  2.8657  2.8993
## hum         -0.7461 -0.7654 -0.7265
## windspeed   -1.1484 -1.1853 -1.1111
```

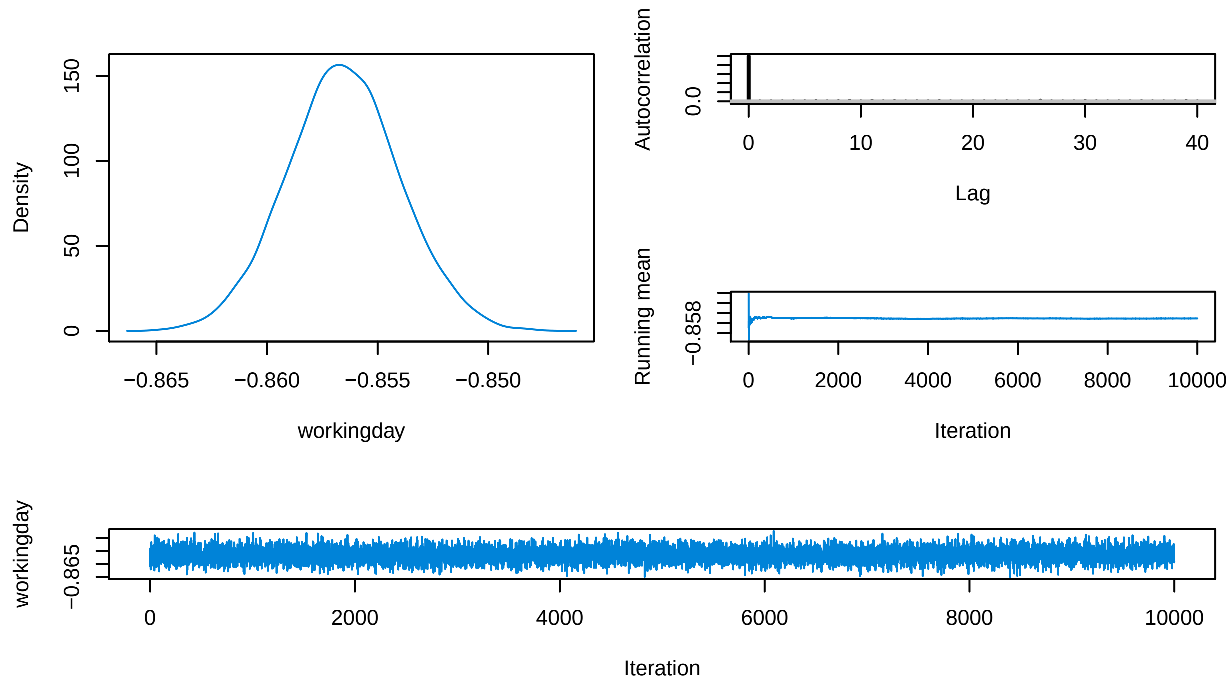
And now the diagnostic mcmc plots

```r
library(mcmcplots)

for (v in colnames(beta)) {
  x <- data.frame(var = beta[,v])
  colnames(x) <- v
  mcmcplot1(x, style = "plain")
}
```

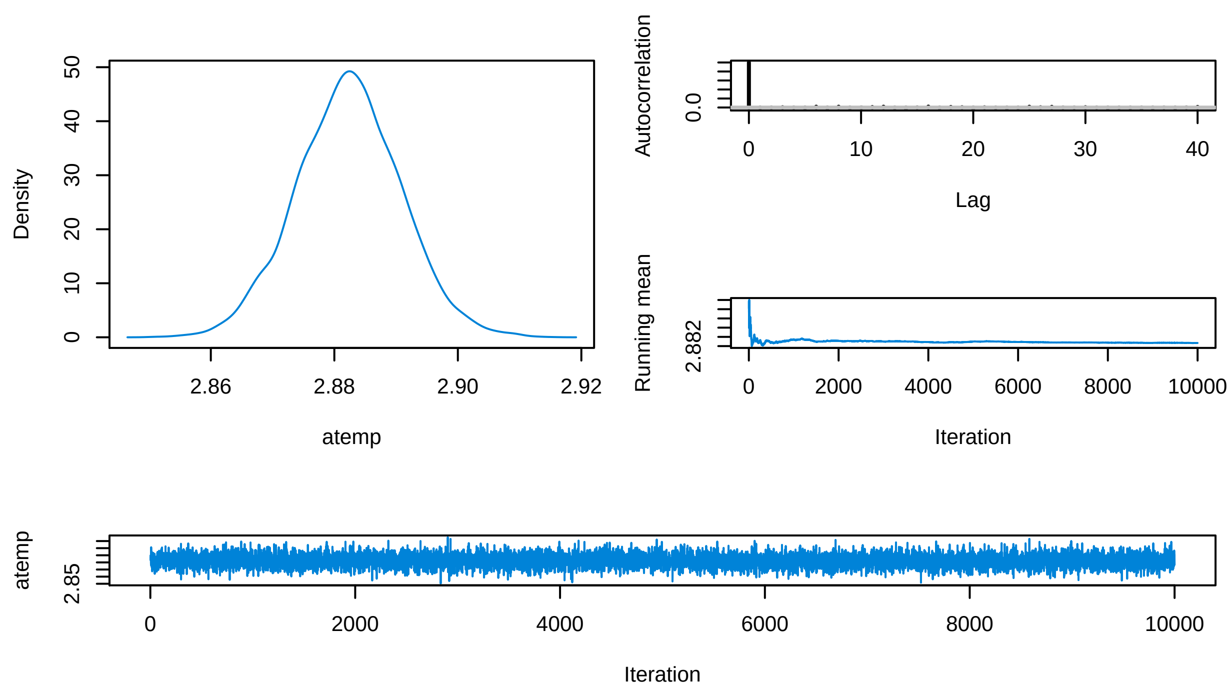

Diagnostics for (Intercept)
Density
(Intercept)
Autocorrelation
Lag
Running mean
Iteration
(Intercept)
Iteration

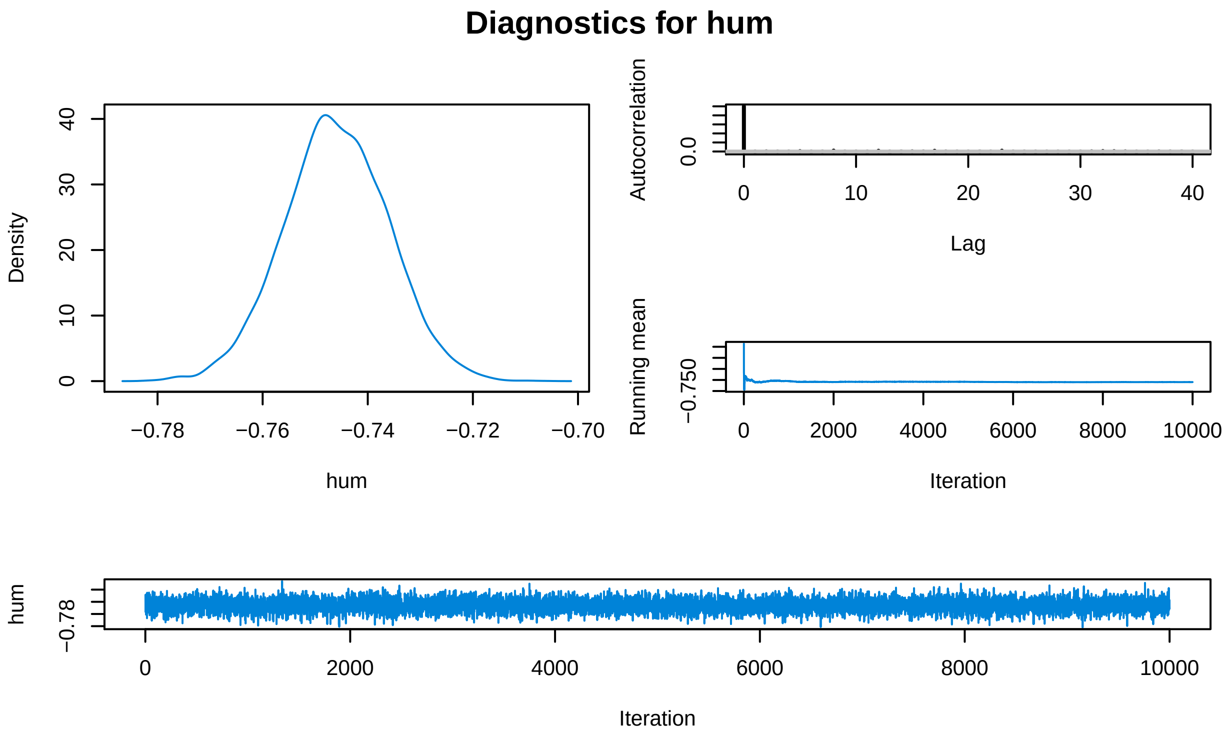

Diagnostics for yr
Density
yr
Autocorrelation
Lag
Running mean
Iteration
yr
Iteration

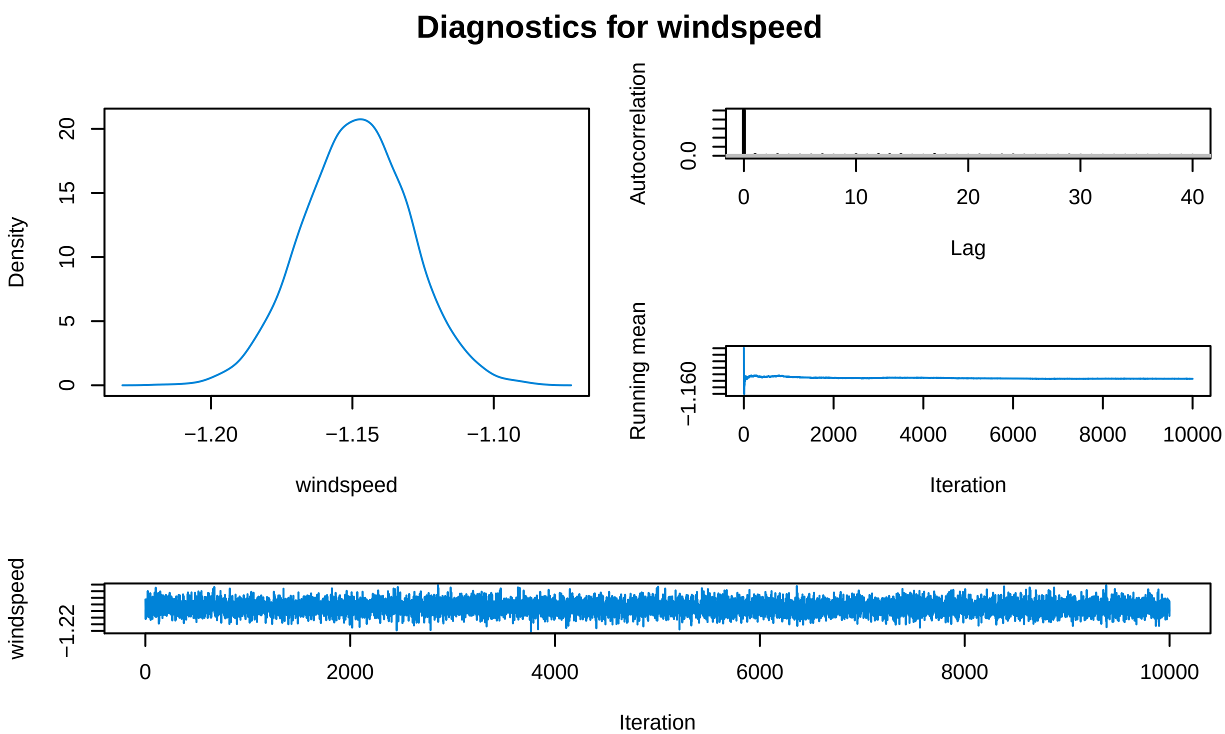
Diagnostics for workingday



Diagnostics for atemp







And lastly, Gewecke's Diagnostic

```
geweke.diag(mcmc(beta))
```

```
##
```

```
## Fraction in 1st window = 0.1
## Fraction in 2nd window = 0.5
##
## (Intercept)          yr  workingday       atemp         hum   windspeed
##    -2.84061    -0.28246    -0.01257     1.83148     1.61526     3.06513
```

Gewecke's Diagnostic mostly looks good using a threshold of $|z| \leq 3$, however `windspeed` is just over 3, suggesting we may want to exclude it from the final model

Now for the registered user model:

```
registered_model_1 <- glm(registered ~ yr  + workingday  + atemp + hum  + holiday + temp + windspeed,
                          data = bikeshare, family = poisson(link = "log"))
summary(registered_model_1)
```

```
##
## Call:
## glm(formula = registered ~ yr + workingday + atemp + hum + holiday +
##     temp + windspeed, family = poisson(link = "log"), data = bikeshare)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -83.054   -9.805   -0.049   10.771   39.311
##
## Coefficients:
##              Estimate Std. Error  z value Pr(>|z|)
## (Intercept)  7.480269   0.004662 1604.639  < 2e-16 ***
## yr           0.486662   0.001275  381.646  < 2e-16 ***
## workingday   0.265292   0.001449  183.057  < 2e-16 ***
## atemp        1.507289   0.031285   48.179  < 2e-16 ***
## hum         -0.402209   0.004695  -85.670  < 2e-16 ***
## holiday     -0.100041   0.004409  -22.688  < 2e-16 ***
## temp        -0.149680   0.027541   -5.435 5.48e-08 ***
## windspeed   -0.819787   0.008740  -93.794  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 526854  on 730  degrees of freedom
## Residual deviance: 165097  on 723  degrees of freedom
## AIC: 172364
##
## Number of Fisher Scoring iterations: 4
```

Same findings as with the casual model, with `temp` and `holiday` being the least important variables

Remove them from the model:

```
registered_model_2 <- glm(registered ~ yr  + workingday  + atemp + hum + windspeed,
                          data = bikeshare, family = poisson(link = "log"))
summary(registered_model_2)
```

```
##
## Call:
## glm(formula = registered ~ yr + workingday + atemp + hum + windspeed,
##     family = poisson(link = "log"), data = bikeshare)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -82.915   -9.634    0.049   10.900   34.705
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  7.477630   0.004507 1659.28   <2e-16 ***
## yr           0.486446   0.001275  381.49   <2e-16 ***
## workingday   0.273817   0.001403  195.11   <2e-16 ***
## atemp        1.339955   0.003961  338.28   <2e-16 ***
## hum         -0.401569   0.004688  -85.65   <2e-16 ***
## windspeed   -0.824850   0.008672  -95.11   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 526854  on 730  degrees of freedom
## Residual deviance: 165663  on 725  degrees of freedom
## AIC: 172926
##
## Number of Fisher Scoring iterations: 4
```

Now humidity and windspeed are much less significant predictors than the other variables, but they may be important enough to still include.

```
bhat    <- coef(registered_model_2)
vbeta   <- vcov(registered_model_2)
B       <- 10000
set.seed(1959)
beta    <- rmvnorm(B, mean = bhat, sigma = vbeta)

round(t(apply(beta, 2, quantile, probs = c(0.5, 0.025, 0.975))), 4)
```

```
##                 50%    2.5%   97.5%
## (Intercept)  7.4778  7.4689  7.4865
## yr           0.4864  0.4840  0.4889
## workingday   0.2738  0.2710  0.2766
## atemp        1.3399  1.3322  1.3478
## hum         -0.4017 -0.4108 -0.3924
## windspeed   -0.8249 -0.8420 -0.8076
```

```
geweke.diag(mcmc(beta))
```

```
##
## Fraction in 1st window = 0.1
## Fraction in 2nd window = 0.5
```

```
##
## (Intercept)          yr  workingday       atemp         hum   windspeed
##    -2.78400    -0.24612    -0.08971     1.82901     1.65302     3.07871
```

Gewecke's Diagnostic again suggests removing windspeed from our predictive variables

```r
registered_model_3 <- glm(registered ~ yr  + workingday  + atemp + hum,
                      data = bikeshare, family = poisson(link = "log"))
summary(registered_model_3)
```

```
##
## Call:
## glm(formula = registered ~ yr + workingday + atemp + hum, family = poisson(link = "log"),
##     data = bikeshare)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -90.376  -10.558   -0.552   11.249   33.590
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  7.220867   0.003620 1994.45   <2e-16 ***
## yr           0.492092   0.001274  386.28   <2e-16 ***
## workingday   0.275067   0.001403  196.01   <2e-16 ***
## atemp        1.404974   0.003914  358.96   <2e-16 ***
## hum         -0.295398   0.004528  -65.24   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 526854  on 730  degrees of freedom
## Residual deviance: 174827  on 726  degrees of freedom
## AIC: 182088
##
## Number of Fisher Scoring iterations: 4
```

```r
bhat    <- coef(registered_model_3)
vbeta   <- vcov(registered_model_3)
B       <- 10000
set.seed(1959)
beta    <- rmvnorm(B, mean = bhat, sigma = vbeta)

round(t(apply(beta, 2, quantile, probs = c(0.5, 0.025, 0.975))), 4)
```

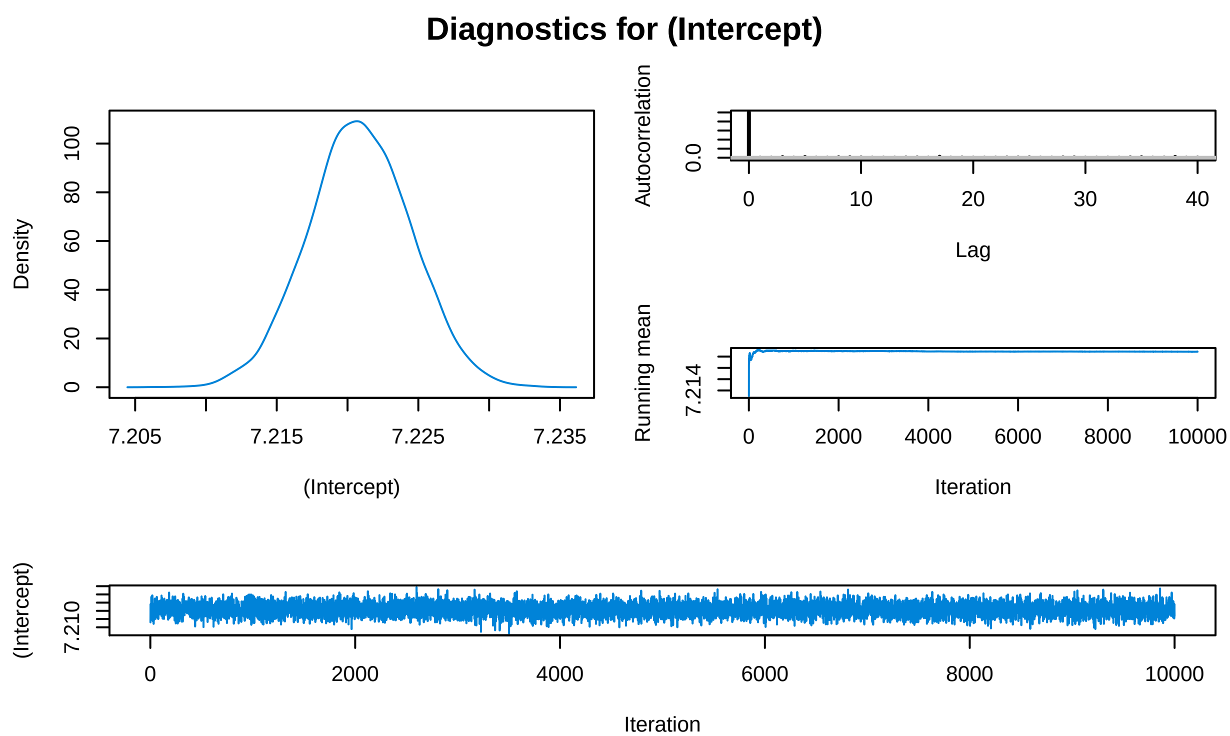
```
##                 50%    2.5%   97.5%
## (Intercept)  7.2208  7.2138  7.2280
## yr           0.4921  0.4896  0.4946
## workingday   0.2751  0.2724  0.2778
## atemp        1.4051  1.3973  1.4126
## hum         -0.2955 -0.3043 -0.2864
```

```r
geweke.diag(mcmc(beta))
```

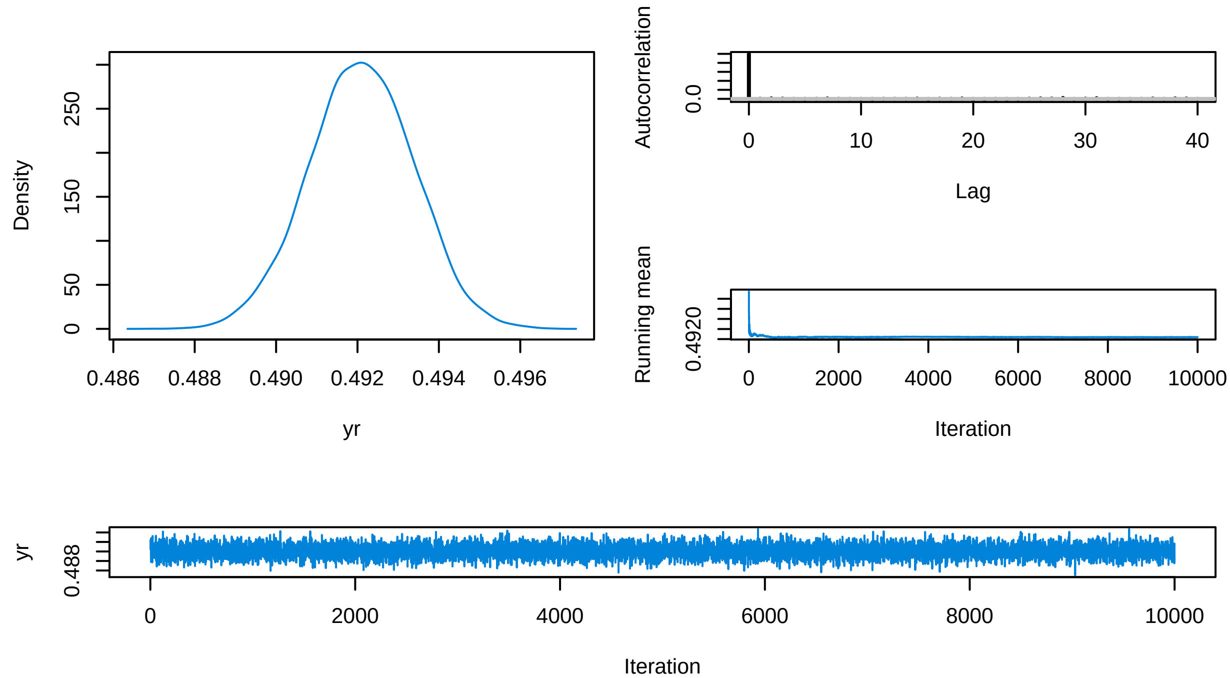
```
##
## Fraction in 1st window = 0.1
## Fraction in 2nd window = 0.5
##
## (Intercept)          yr  workingday       atemp         hum
##     1.44924    -0.21112     0.45479     0.01527    -1.41012
```

And lastly, the model diagnostic plots:

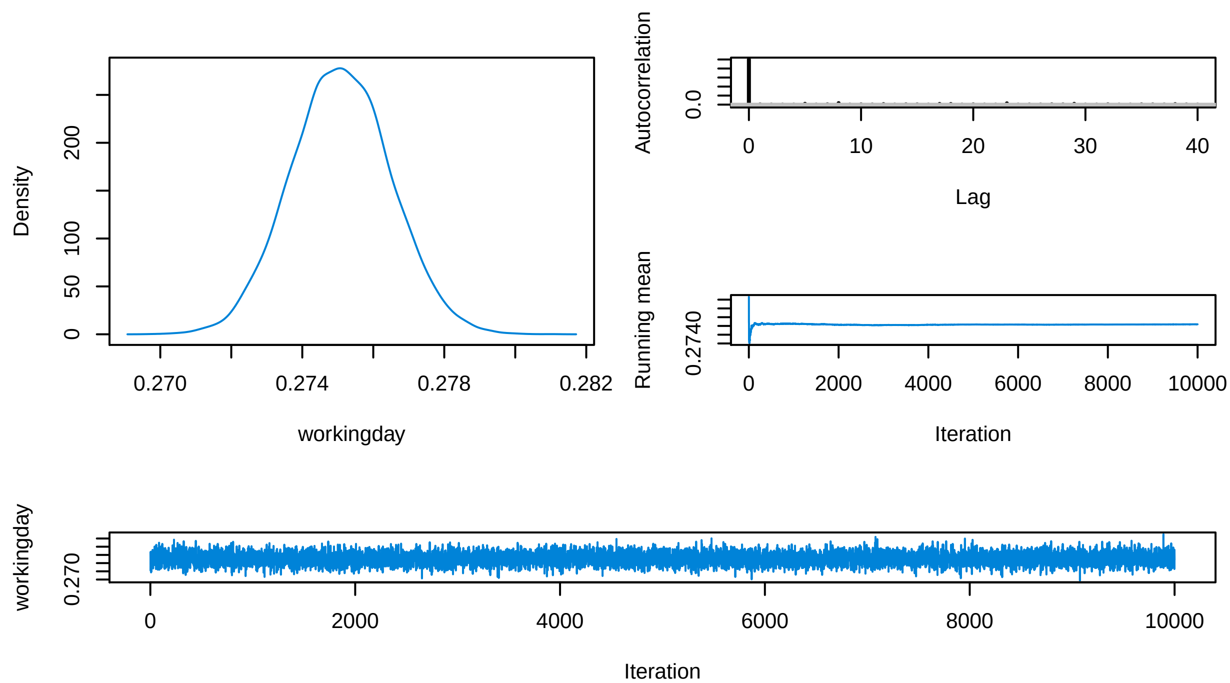
```r
for (v in colnames(beta)) {
  x <- data.frame(var = beta[,v])
  colnames(x) <- v
  mcmcplot1(x, style = "plain")
}
```



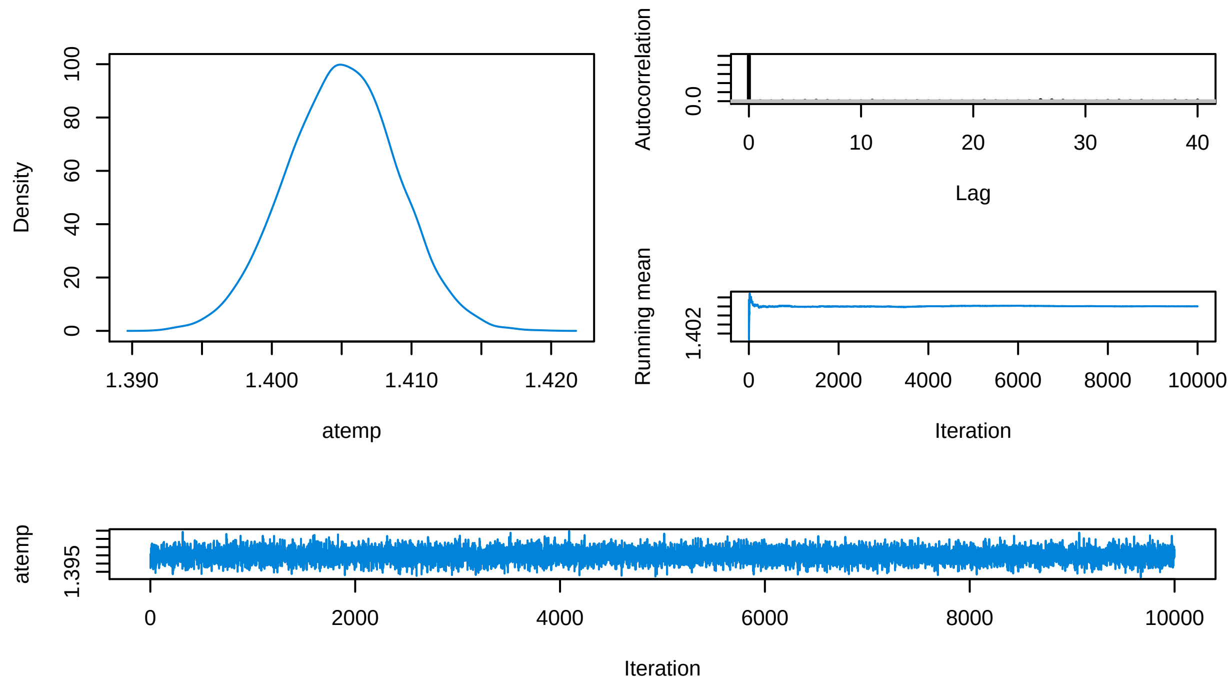
Diagnostics for yr



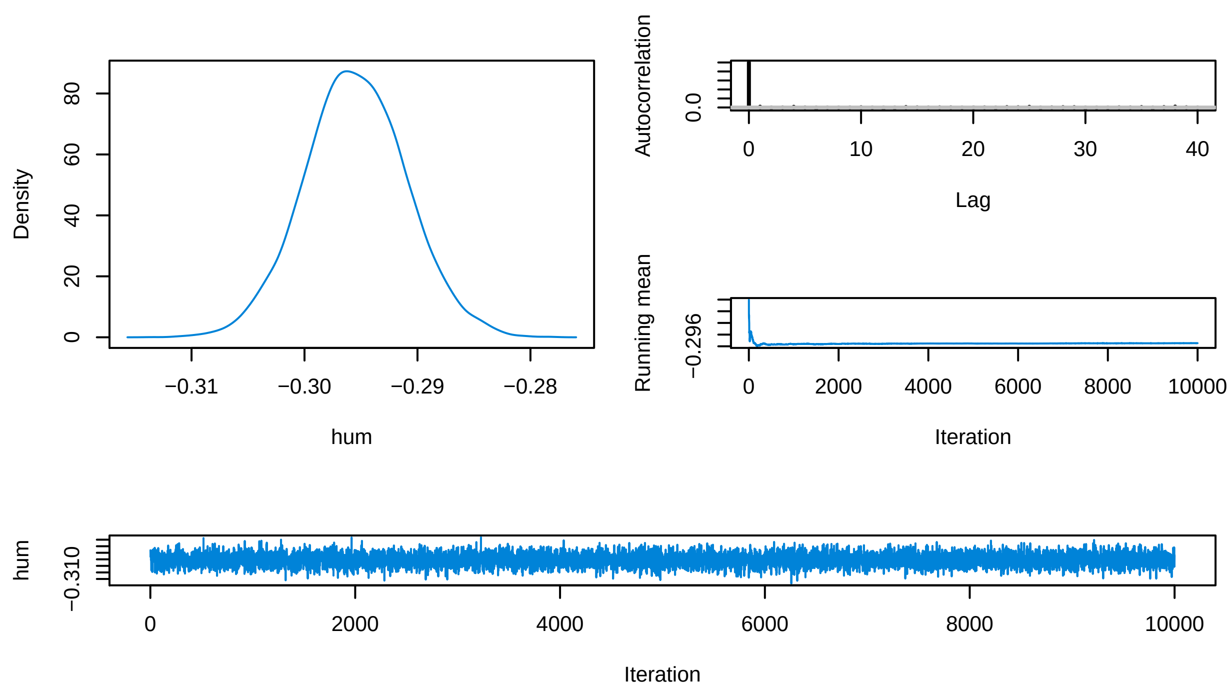
Diagnostics for workingday



Diagnostics for atemp



Diagnostics for hum



The models are very similar, with one glaring exception. For registered users, a `workingday` leads to higher user count, while casual users climb on non-working days. This matches our findings from the mid-term, which is good that we were able to replicate findings with different methods. And again, it make intuitive sense, as we would expect registered users to be higher volume and use the bikes on their commute, while people who ride casually are likely to have a separate commute routine, but ride the bikes to get around the city on the weekends.

Theoretical

1

$\mathcal{L}(y|\lambda) = \prod \frac{\lambda^{y_i} e^{-\lambda}}{y_i!} \propto e^{-n\lambda} \prod \lambda^{y_i} = e^{-n\lambda} \lambda^{\sum y_i}$

$\pi(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}$

$\pi(\alpha, \beta) \propto \frac{\beta^{\alpha s}}{\Gamma(\alpha)^r} p^{\alpha-1} e^{-\beta q}$

$$
\begin{aligned}
\mathcal{L}(Y|\lambda)\pi(\lambda|\alpha,\beta)\pi(\alpha,\beta) &\propto \left(e^{-n\lambda}\lambda^{\sum y_i}\right)\left(\frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda}\right)\left(\frac{\beta^{\alpha s}}{\Gamma(\alpha)^r}p^{\alpha-1}e^{-\beta q}\right) \\
&\propto \lambda^{\sum y_i + \alpha - 1} e^{-(n\lambda + \beta\lambda)} \frac{\beta^{\alpha(s+1)}}{\Gamma(\alpha)^{r+1}} e^{-\beta q} p^{\alpha-1} \\
&\propto \lambda^{\sum y_i + \alpha - 1} e^{-(n+\beta)\lambda}
\end{aligned}
$$

Which is the kernel for a Gamma distribution with parameters:

$\alpha' = \sum y_i + \alpha$

$\beta' = n + \beta$

So $P(\lambda|y) = \frac{(n+\beta)^{\sum y_i + \alpha}}{\Gamma(\sum y_i + \alpha)} \lambda^{\sum y_i + \alpha - 1} e^{-(n+\beta)\lambda}$

$$
\begin{aligned}
P(\alpha, \beta|y) &= \frac{P(\alpha, \beta, \lambda|y)}{P(\lambda|y, \alpha, \beta)} \\
&= \frac{\lambda^{\sum y_i + \alpha - 1} \frac{\beta^{\alpha(s+1)}}{\Gamma(\alpha)^{r+1}} e^{-(n\lambda + \beta\lambda + \beta q)} p^{\alpha-1}}{\lambda^{\sum y_i + \alpha - 1} \frac{(n+\beta)^{\sum y_i + \alpha}}{\Gamma(\sum y_i + \alpha)} e^{-(n+\beta)\lambda}} \\
&= \frac{\Gamma(\sum y_i + \alpha) \beta^{\alpha(s+1)} p^{\alpha-1} e^{-\beta q}}{(n+\beta)^{\sum y_i + \alpha} \Gamma(\alpha)^{r+1}}
\end{aligned}
$$

And this doesn't have a closed form distribution/solution, but is easy enough to calculate analytically

2

$$
\begin{aligned}
\mathcal{L}(y|\mu) &= \prod \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left[-\frac{(y_i - \mu)^2}{2\sigma_0^2}\right] \\
&= (2\pi\sigma_0^2)^{-n/2} \exp\left[-\frac{1}{2\sigma_0^2} \sum (y_i - \mu)^2\right] \\
&\propto \exp\left[-\frac{1}{2\sigma_0^2} \sum (y_i - \mu)^2\right]
\end{aligned}
$$

$$
P(\mu|\theta) = \frac{1}{\sqrt{2\pi\tau_0^2}} \exp\left[-\frac{(\mu - \theta)^2}{2\tau_0^2}\right]
$$

$$P(\theta) = \frac{1}{\sqrt{2\pi\gamma_0^2}} \exp\left[-\frac{\theta^2}{2\gamma_0^2}\right]$$

$$\begin{aligned}
P(\mu, \theta|y) &\propto P(\theta)P(\mu|\theta)\mathcal{L}(y|\mu) \\
&\propto \frac{1}{\sqrt{2\pi\gamma_0^2}} \exp\left[-\frac{\theta^2}{2\gamma_0^2}\right] \frac{1}{\sqrt{2\pi\tau_0^2}} \exp\left[-\frac{(\mu-\theta)^2}{2\tau_0^2}\right] \exp\left[-\frac{1}{2\sigma_0^2}\sum(y_i-\mu)^2\right]
\end{aligned}$$

$$\begin{aligned}
P(\mu|y) &\propto P(\mu|\theta)\mathcal{L}(y|\mu) \\
&\propto \frac{1}{\sqrt{2\pi\tau_0^2}} \exp\left[-\frac{(\mu-\theta)^2}{2\tau_0^2}\right] \exp\left[-\frac{1}{2\sigma_0^2}\sum(y_i-\mu)^2\right]
\end{aligned}$$

As shown in our notes:

$$\begin{aligned}
\exp\left[-\frac{1}{2\sigma_0^2}\sum(y_i-\mu)^2\right] &= \exp\left[-\frac{1}{2\sigma_0^2}\sum(y_i-\bar{y}+\bar{y}+\mu^2)\right] \\
&= \exp\left[-\frac{1}{2\sigma_0^2}\left\{\sum(y_i-\bar{y})^2 + \sum 2(y_i-\bar{y})(\bar{y}-\mu) + \sum(\bar{y}-\mu)^2\right\}\right] \\
&= \exp\left[-\frac{1}{2\sigma_0^2}\left\{\sum(y_i-\bar{y})^2 + n(\bar{y}-\mu)^2\right\}\right] \\
&\propto \exp\left[-\frac{n}{2\sigma_0^2}(\bar{y}-\mu)^2\right]
\end{aligned}$$

Plugging in:

$$\begin{aligned}
P(\mu|y) &\propto P(\mu|\theta)\mathcal{L}(y|\mu) \\
&\propto \frac{1}{\sqrt{2\pi\tau_0^2}} \exp\left[-\frac{(\mu-\theta)^2}{2\tau_0^2}\right] \exp\left[-\frac{1}{2\sigma_0^2}\sum(y_i-\mu)^2\right] \\
&\propto \exp\left[-\frac{(\mu-\theta)^2}{2\tau_0^2}\right] \exp\left[-\frac{n}{2\sigma_0^2}(\bar{y}-\mu)^2\right] \\
&= \exp\left[-\frac{1}{2}\left(\frac{n}{\sigma_0^2}(\bar{y}-\mu)^2 + \frac{1}{\tau_0^2}(\mu-\theta)^2\right)\right] \\
&= \exp\left[-\frac{1}{2}\left(\frac{n}{\sigma_0^2}(\bar{y}^2-2\mu\bar{y}+\mu^2) + \frac{1}{\tau_0^2}(\mu^2-2\mu\theta+\theta^2)\right)\right] \\
&= \exp\left[-\frac{1}{2}\left(\mu^2\left\{\frac{n}{\sigma_0^2}+\frac{1}{\tau_0^2}\right\}\right) - 2\mu\left\{\frac{\theta}{\tau_0^2}+\frac{n\bar{y}}{\sigma_0^2}\right\} + \frac{\theta^2}{\tau_0^2} + \frac{n\bar{y}^2}{\sigma_0^2}\right] \\
&= \exp\left[-\frac{1}{2}\left(\frac{1}{\tau_0^2}+\frac{n}{\sigma_0^2}\right)\left(\mu^2 - 2\mu\left\{\frac{\frac{\theta}{\tau_0^2}+\frac{n\bar{y}}{\sigma_0^2}}{\frac{1}{\tau_0^2}+\frac{n}{\sigma_0^2}}\right\} + \frac{\frac{\theta^2}{\tau_0^2}+\frac{n\bar{y}^2}{\sigma_0}}{\frac{1}{\tau_0^2}+\frac{n}{\sigma_0^2}}\right)\right] \\
&= \exp\left[-\frac{1}{2}\left(\frac{1}{\tau_0^2}+\frac{n}{\sigma_0^2}\right)\left(\left\{\mu - \frac{\frac{\theta}{\tau_0^2}+\frac{n\bar{y}}{\sigma_0^2}}{\frac{1}{\tau_0^2}+\frac{n}{\sigma_0^2}}\right\}^2 + \frac{\frac{\theta^2}{\tau_0^2}+\frac{n\bar{y}^2}{\sigma_0^2}}{\frac{1}{\tau_0^2}+\frac{n}{\sigma_0^2}} - \left\{\frac{\frac{\theta}{\tau_0^2}+\frac{n\bar{y}}{\sigma_0^2}}{\frac{1}{\tau_0^2}+\frac{n}{\sigma_0^2}}\right\}^2\right)\right] \\
&\propto \exp\left[-\frac{1}{2}\left(\frac{1}{\tau_0^2}+\frac{n}{\sigma_0^2}\right)\left(\mu - \frac{\frac{\theta}{\tau_0^2}+\frac{n\bar{y}}{\sigma_0^2}}{\frac{1}{\tau_0^2}+\frac{n}{\sigma_0^2}}\right)^2\right]
\end{aligned}$$

Which is the kernel for a normal distribution with parameters $\mu_1 = \frac{\frac{\theta}{\tau_0^2}+\frac{n\bar{y}}{\sigma_0^2}}{\frac{1}{\tau_0^2}+\frac{n}{\sigma_0^2}}$ and $\frac{1}{\tau_1^2} = \frac{1}{\tau_0^2} + \frac{n}{\sigma_0^2}$

Now we need to find $P(\theta|y_i's)$

Note that $P(y|\theta) \sim \mathcal{N}(\theta, \sigma_0^2 + \tau^2)$

$$\begin{aligned}
P(\theta|y_i) &\propto P(\theta)P(y|\theta) \\
&\propto \exp\left[-\frac{1}{2\gamma_0^2}(\theta-0)^2\right] \prod \frac{1}{\sqrt{2\pi(\sigma_0^2+\tau^2)}} \exp\left[-\frac{(y_i-\theta)^2}{2(\sigma_0^2+\tau^2)}\right] \\
&\propto \exp\left[-\frac{1}{2\gamma_0^2}(\theta-0)^2\right] \exp\left[-\frac{1}{2(\sigma_0^2+\tau^2)}\sum(y_i-\theta)^2\right]
\end{aligned}$$

Now follow the same process algebraically as we did for the conditional posterior, where we sub in 0 for θ, θ is subbed for μ, $\sigma_0^2 + \tau^2$ is subbed for σ_0^2, and γ_0^2 is subbed for τ_0^2

This leaves us with:

$$\begin{aligned}
P(\theta|y_i) &\propto P(\theta)P(y|\theta) \\
&\propto \exp\left[-\frac{1}{2\gamma_0^2}(\theta-0)^2\right] \prod \frac{1}{\sqrt{2\pi(\sigma_0^2+\tau^2)}} \exp\left[-\frac{(y_i-\theta)^2}{2(\sigma_0^2+\tau^2)}\right] \\
&\propto \exp\left[-\frac{1}{2\gamma_0^2}(\theta-0)^2\right] \exp\left[-\frac{1}{2(\sigma_0^2+\tau^2)}\sum(y_i-\theta)^2\right] \\
&\propto \exp\left[-\frac{1}{2}\left(\frac{1}{\gamma_0^2}+\frac{n}{\sigma_0^2+\tau_0^2}\right)\left(\theta - \frac{\frac{n\bar{y}}{\sigma_0^2+\tau_0^2}}{\frac{1}{\gamma_0^2}+\frac{n}{\sigma_0^2+\tau_0^2}}\right)^2\right]
\end{aligned}$$

So θ is normally distributed with mean $= \frac{n\bar{y}}{\sigma_0^2+\tau_0^2}$ and variance $\gamma_1^2 = \frac{1}{\gamma_0^2} + \frac{n}{\sigma_0^2+\tau_0^2}$

3

$$\begin{aligned}
P(\theta^{(b)} \le a) &= \int_{-\infty}^{a} p(\theta)d\theta \\
&= \int_{-\infty}^{\infty} 1\{\theta \le a\}\frac{p(\theta)}{g(\theta^*)}g(\theta^*)d\theta \\
&= M \cdot E_g\left[1\{\theta \le a\}\frac{p(\theta)}{Mg(\theta^*)}\right] \\
&= M \cdot E_g\left[1\{\theta \le a\}E\left[1\left\{U \le \frac{p(\theta)}{Mg(\theta^*)}\right\}|\theta\right]\right] \\
&= ME_gE\left[1\{\theta \le a\}1\left\{U \le \frac{p(\theta)}{Mg(\theta^*)}\right\}|\theta\right] \\
&= \frac{P(\theta \le a, \theta^* accepted)}{1/M} \\
&= \frac{P(\theta \le a, \theta^* accepted)}{P(\theta^* accepted)} \\
&= P(\theta^* \le a|accepted\theta^*)
\end{aligned}$$