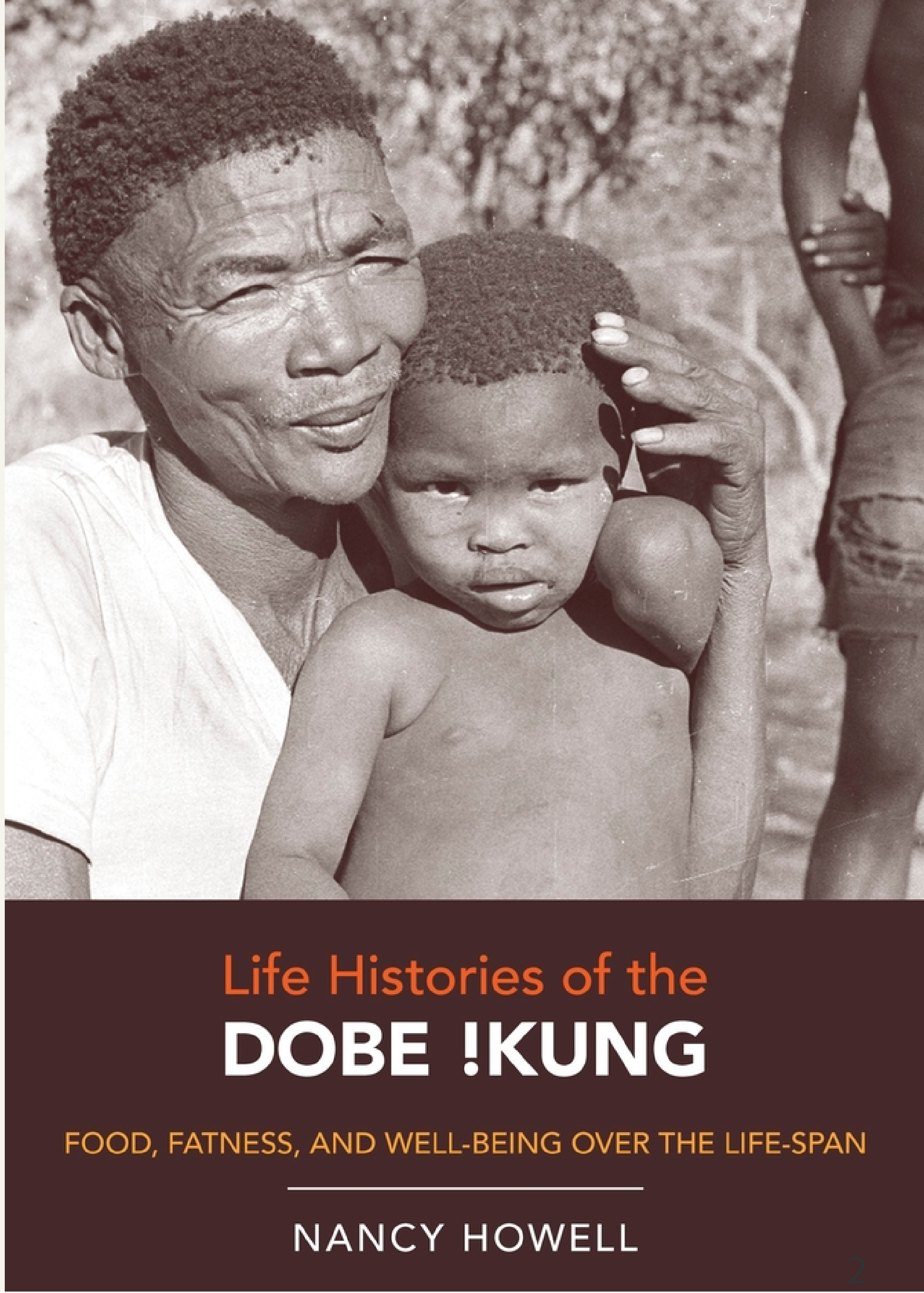
Bayesian regression

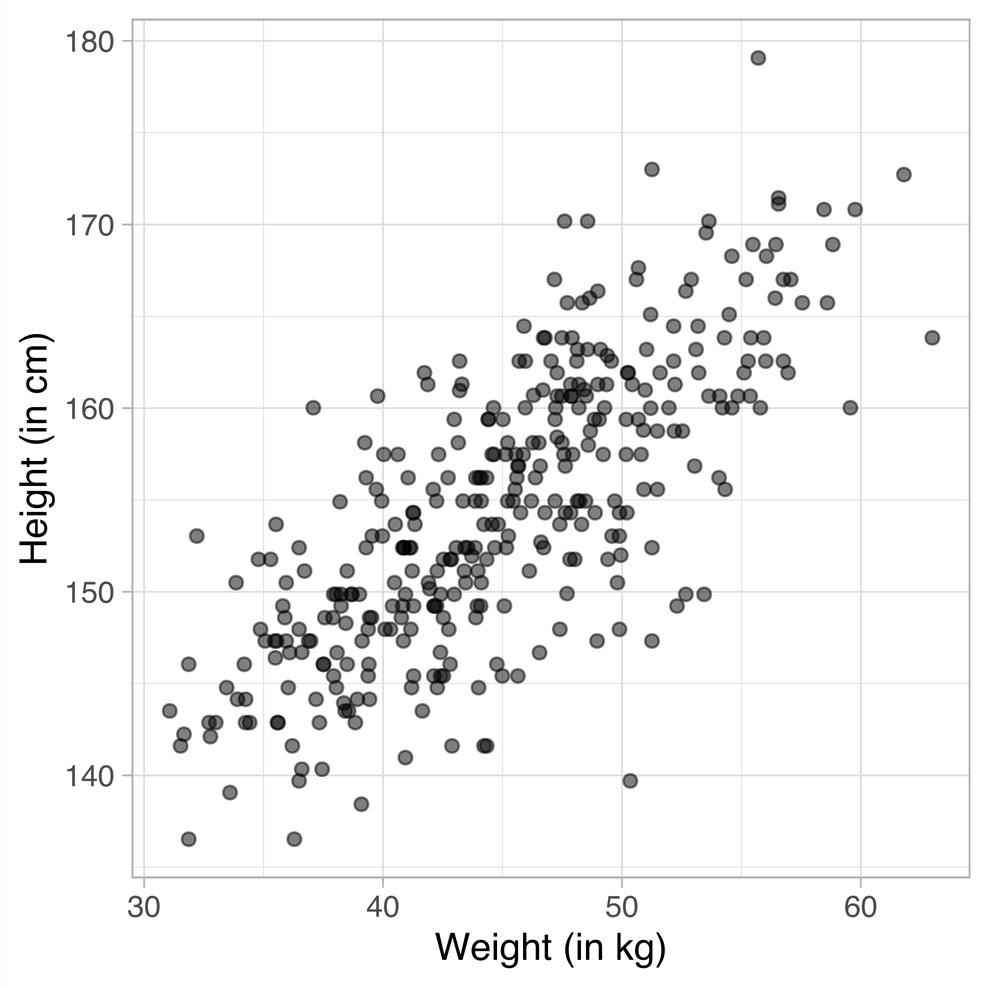
Stat 340: Bayesian Statistics

Example

- Partial census data for the Dobe area !Kung San, a foraging population
- Compiled from Nancy Howell's interviews

```
##              mean         sd      5.5%     94.5%     histogra
## height 154.59709  7.7423321 142.8750 167.00500
## weight  44.99049  6.4567081  35.1375  55.76588
## age     41.13849 15.9678551  20.0000  70.00000
## male     0.46875  0.4997328   0.0000   1.00000
```





How can we write a **Bayesian** model to relate weight and height of the Kalahari foragers?

Observation-specific mean

We can adapt our normal model for the mean to use an observation-specific mean, μ_i:

Sampling model: $$Y_i|\mu_i, \sigma \overset{\text{ind}}{\sim} \mathcal{N}(\mu_i, \sigma)$$

Now we need to link μ_i and x_i

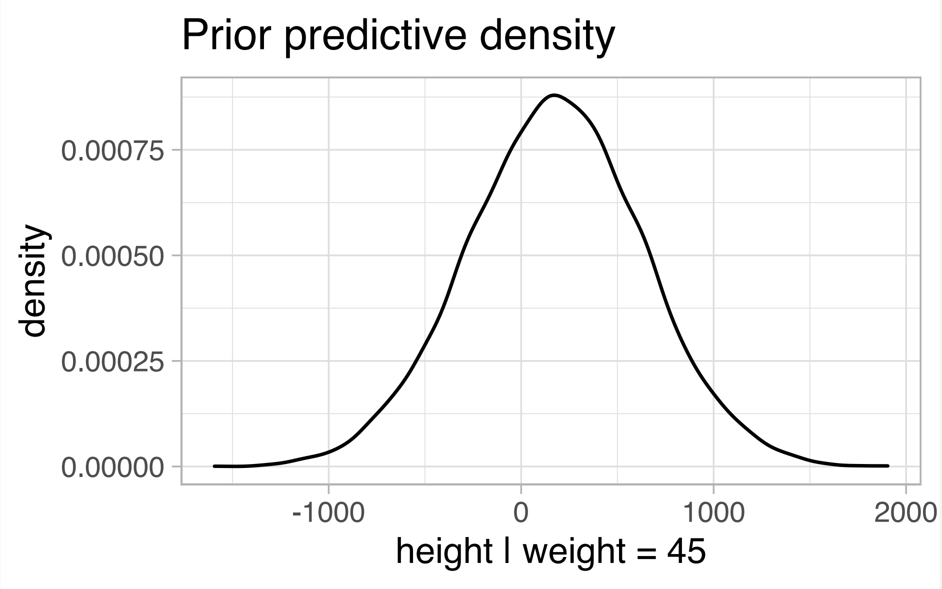
A weakly informative prior

We may have limited prior information about the regression coefficients, β_0 and β_1, and/or the standard deviation σ

Assume independence: $\pi(\beta_0, \beta_1, \sigma) = \pi(\beta_0, \beta_1)\pi(\sigma)$

1. Assume β_0 and β_1 are independent
   - $\pi(\beta_0, \beta_1) = \pi(\beta_0)\pi(\beta_1)$
   - Assign a weakly informative prior each coefficient: $\beta_i \sim \mathcal{N}(m_i, s_i)$
   - example: $\mathcal{N}(0, 100)$
2. Assign a weakly informative prior to σ
   - example: $1\sigma^2 \sim \mathrm{Gamma}(1, 1)$

Prior predictive as "sanity check"



Simulate the prior predictive:

1. Draw parameters from their prior distributions

2. Draw data from the sampling model plugging in these simulated parameters

Sampling from the prior predictive distribution

```
nsim <- 1e4                                         # no. of simulations

prior.sims <- data_frame(                           # simulate from ind. priors
  beta0 = rnorm(nsim, 178, 100),
  beta1 = rnorm(nsim, 0, 10),
  sigma = runif(nsim, 0, 50)
)
```

```
weight.value <- 45                                  # condition on value of x
```

```
prior.pred <- prior.sims %>%
  mutate(
    mu = beta0 + beta1 * weight.value,              # calculate E(y|x)
    height = rnorm(n(), mean = mu, sd = sigma)      # draw sample from normal
  )
```

Deriving the posterior

Sampling model: $Y_i | x_i, \beta_0, \beta_1, \sigma \sim \mathcal{N}(\beta_0 + \beta_1 x_i, \sigma)$

Joint likelihood:

$$\begin{aligned} L(\beta_0, \beta_1, \sigma) &= \prod_{i=1}^{n} \left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{1}{2\sigma^2} (y_i - \beta_0 - \beta_1 x_i)^2 \right\} \right] \\ &\propto \left(\frac{1}{\sigma^2} \right)^{n/2} \exp\left\{ -\frac{1/\sigma^2}{2} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 \right\} \end{aligned}$$

Joint posterior: $\pi(\beta_0, \beta_1, \sigma | \text{data}) \propto \pi(\beta_0, \beta_1, \sigma) L(\beta_0, \beta_1, \sigma)$

JAGS for Bayesian SLR

Write down the model string

```
slr_model <-"model {
## sampling model
for (i in 1:N){
y[i] ~ dnorm(beta0 + beta1 * x[i], invsigma2)
}

## priors
beta0 ~ dnorm(mu0, g0)
beta1 ~ dnorm(mu1, g1)
invsigma2 ~ dgamma(a, b)
sigma <- sqrt(pow(invsigma2, -1))
}"
```

JAGS for Bayesian SLR

Define the data and set prior parameters

```
the_data <- list(
  y = adults$height,    # response variable
  x = adults$weight,    # explanatory variable
  N = nrow(adults),     # sample size
  mu0 = 0,              # prior mean for beta0
  g0 = 0.0001,          # prior precision for beta0
  mu1 = 0,              # prior mean for beta1
  g1 = 0.0001,          # prior precision for beta1
  a = 1,                # prior shape for 1/sigma2
  b = 1                 # prior scale for 1/sigma2
)
```

JAGS for Bayesian SLR

Generate samples from the posterior

```
posterior <- run.jags(
  slr_model,
  data = the_data,
  n.chains = 1,
  monitor = c("beta0", "beta1", "sigma"),
  adapt = 1000,
  burnin = 5000,
  sample = 5000,
  silent.jags = TRUE
)
```

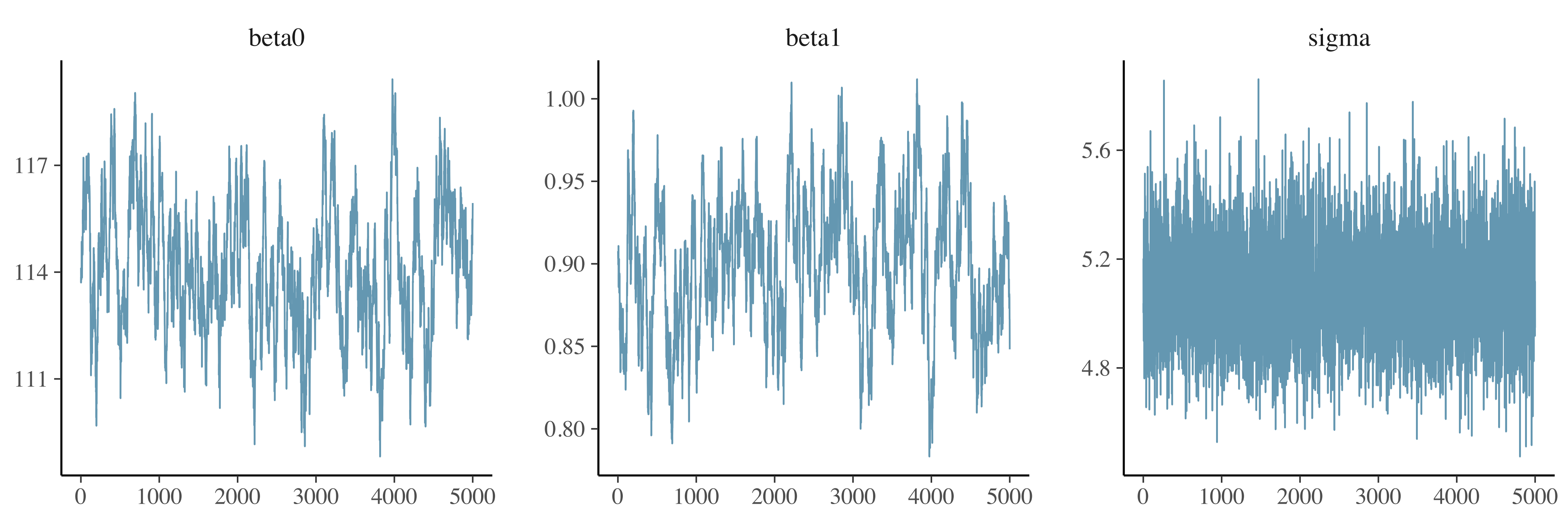
Summary of the fitted model

```
summary(posterior)
```

```
##                 Lower95      Median     Upper95        Mean        SD        Mode
## beta0 110.6005490 114.0250921 117.6466865 114.1046727 1.8346303 113.8677869
## beta1   0.8222891   0.9017146   0.9775988   0.9000988 0.0403727   0.9064987
## sigma   4.7315885   5.0787794   5.4687267   5.0857117 0.1917283   5.0777442
##             MCerr MC%ofSD SSeff      AC.10 psrf
## beta0 0.247707195    13.5    55 0.79203877   NA
## beta1 0.005459015    13.5    55 0.79115579   NA
## sigma 0.002711448     1.4  5000 0.01892964   NA
```

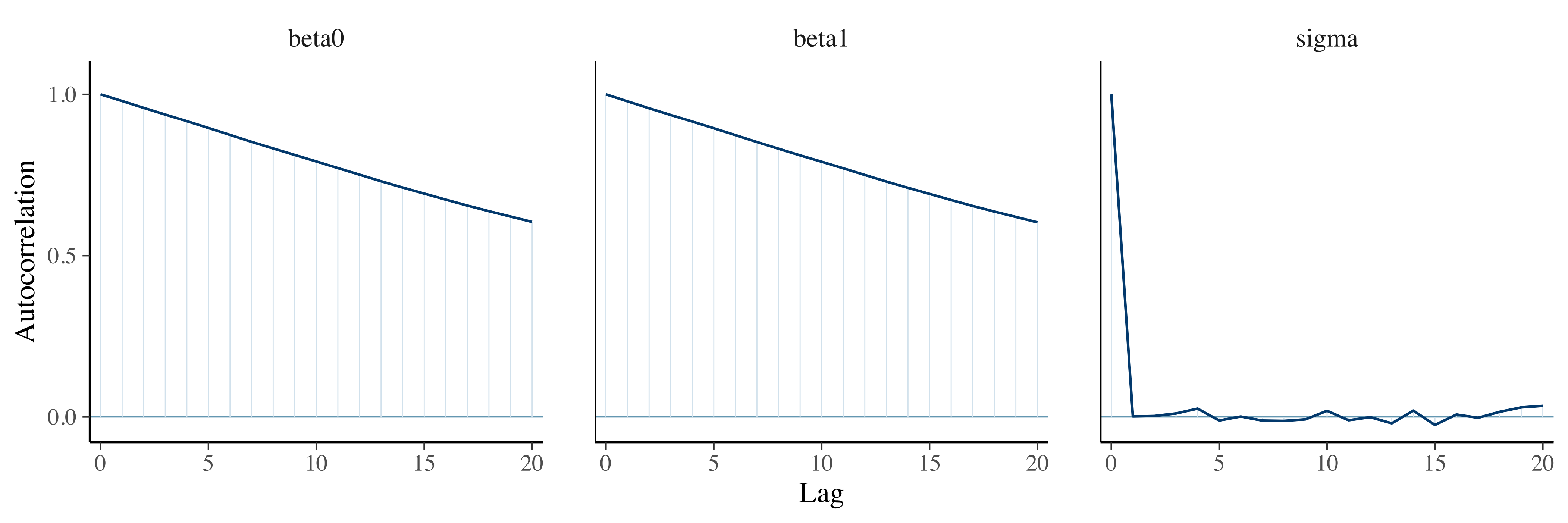
MCMC diagnostics

```
mcmc_trace(posterior$mcmc)
```



MCMC diagnostics

```
mcmc_acf(posterior$mcmc)
```

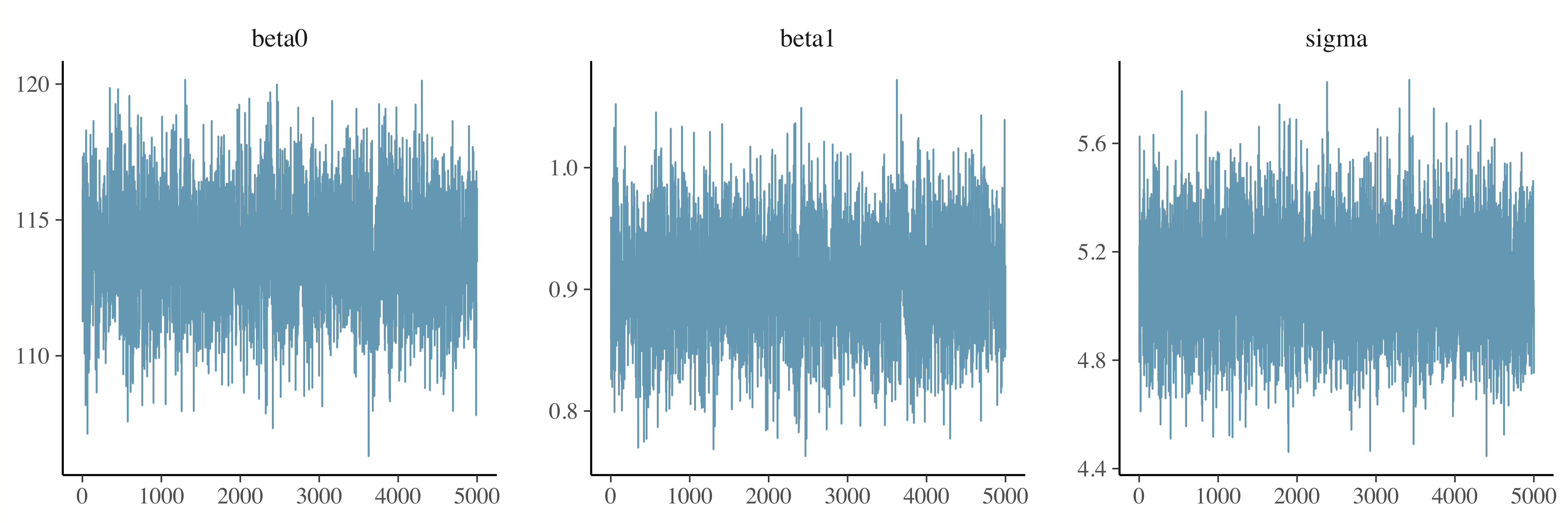


Setting `thin = 50`

```
posterior <- run.jags(
  slr_model,
  data = the_data,
  n.chains = 1,
  monitor = c("beta0", "beta1", "sigma"),
  adapt = 1000,
  burnin = 5000,
  sample = 5000,
  thin = 50,
  silent.jags = TRUE
)
```

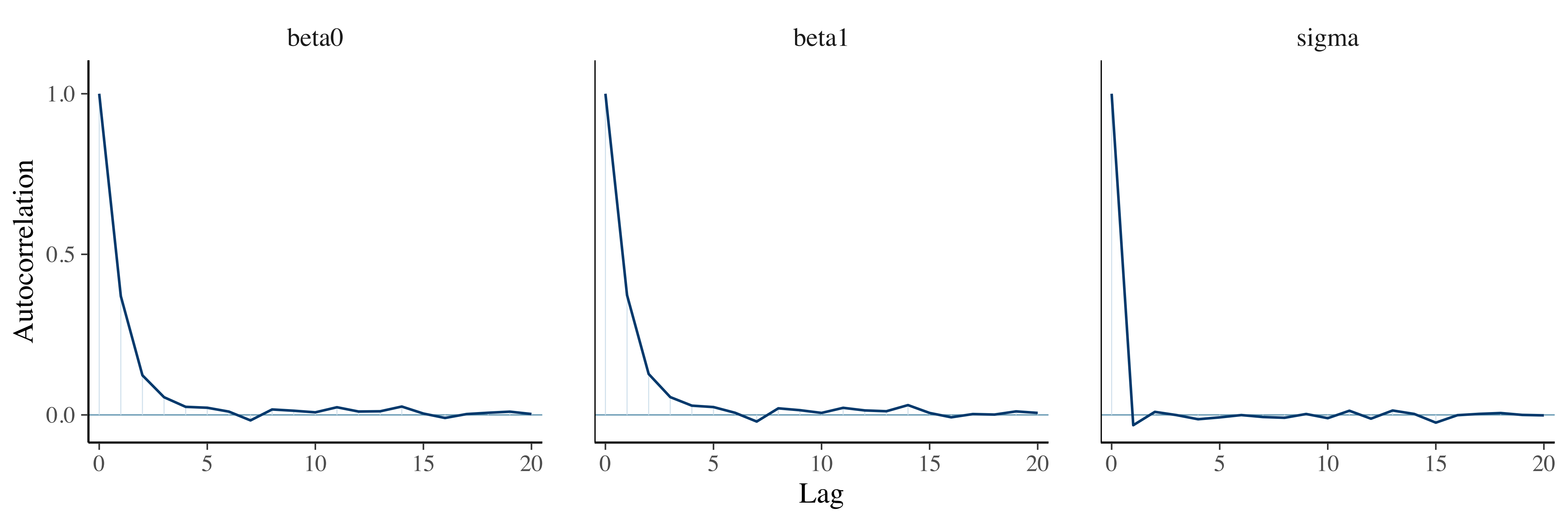
MCMC diagnostics

```
mcmc_trace(posterior$mcmc)
```



MCMC diagnostics

```
mcmc_acf(posterior$mcmc)
```



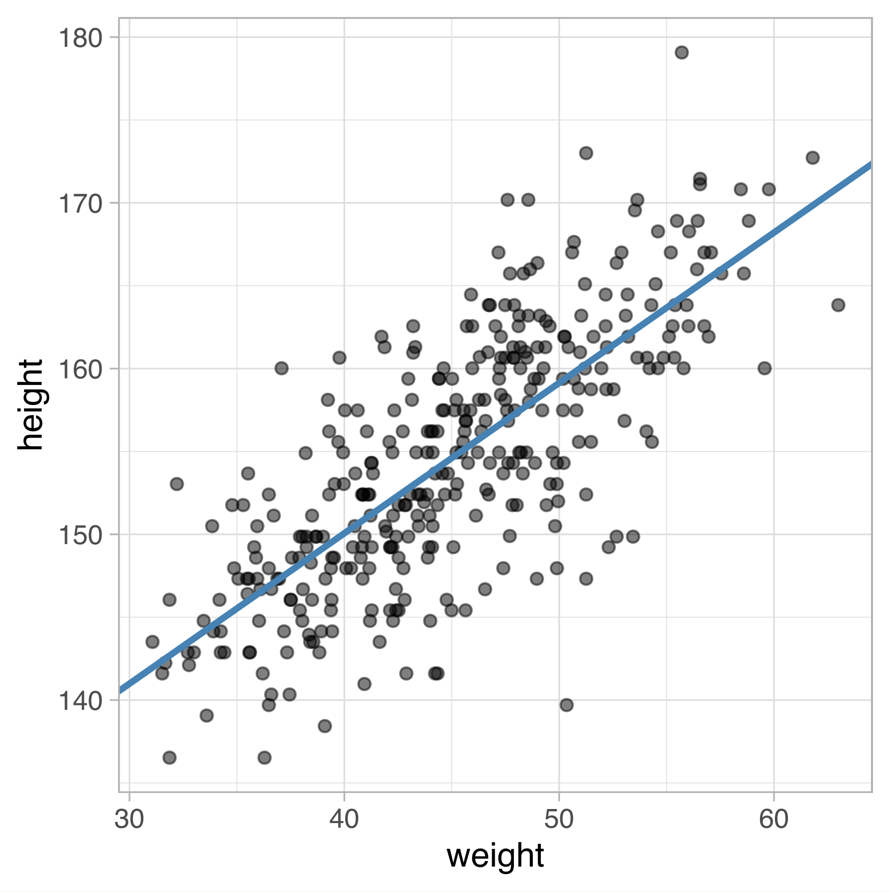
Summary of the fitted model

```
summary(posterior)
```

```
##          Lower95      Median     Upper95        Mean          SD         Mod
## beta0 110.1966997 113.7337743 117.6820902 113.7703854 1.94096529 113.854709
## beta1   0.8216034   0.9082515   0.9880905   0.9073003 0.04272439   0.912202
## sigma   4.6997300   5.0820890   5.4454647   5.0856568 0.19255547   5.067288
##              MCerr MC%ofSD SSeff       AC.500 psrf
## beta0 0.0404676912     2.1  2300  0.008030463   NA
## beta1 0.0009197745     2.2  2158  0.006204312   NA
## sigma 0.0026375062     1.4  5330 -0.010200225   NA
```

Plotting the fitted model

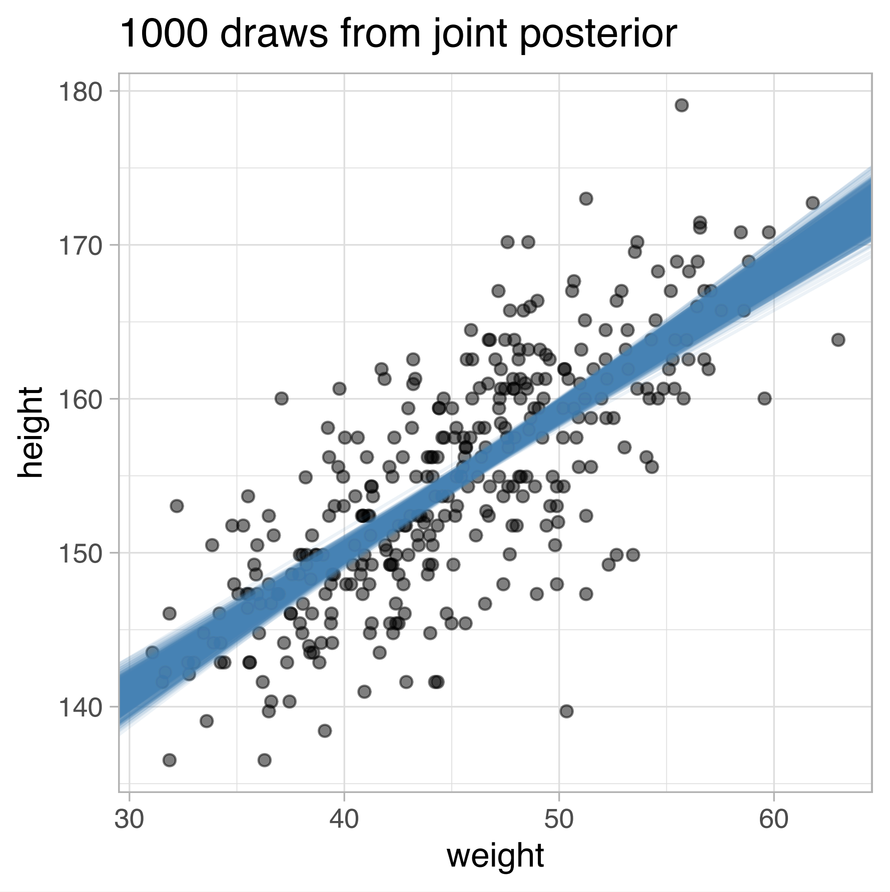
```
post_means <- apply(
  posterior$mcmc[[1]], 2, mean
)
```



Sampling from the joint posterior

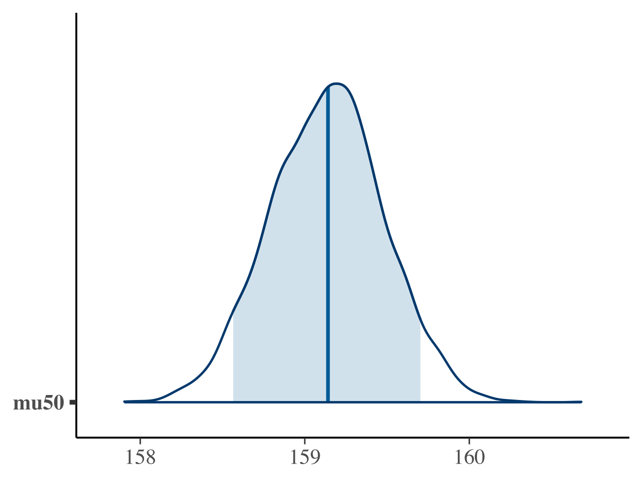
```r
post_draws <- as.data.frame(
  posterior$mcmc[[1]]
)
head(post_draws)
```

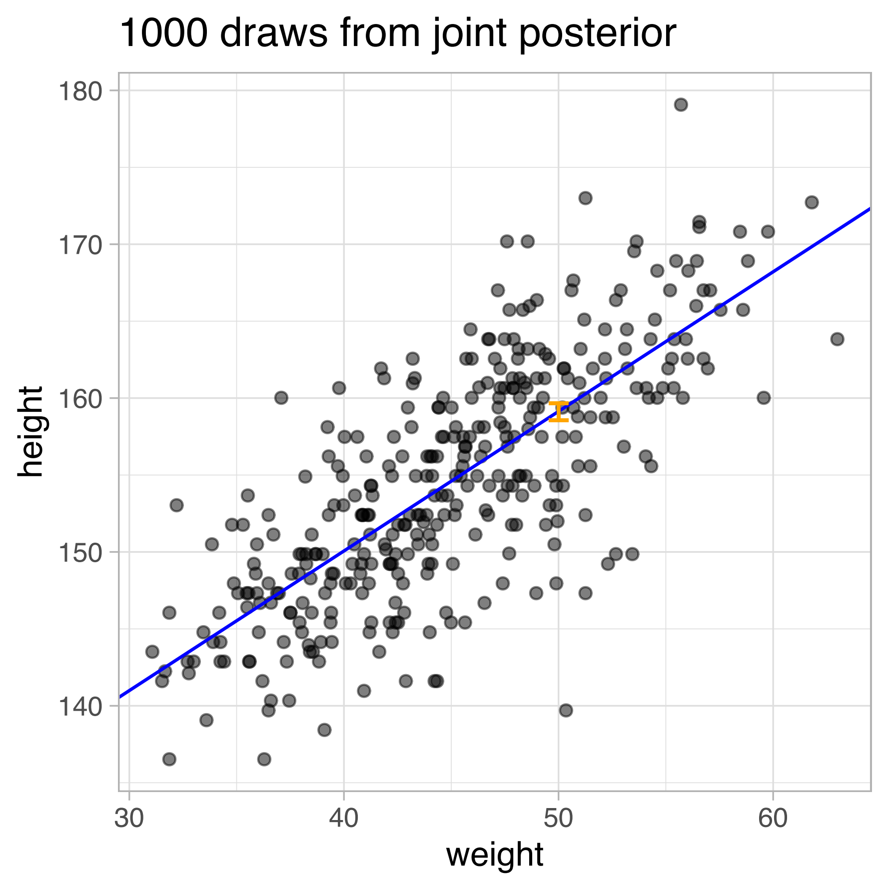
```
##         beta0     beta1    sigma
## 6001 116.1242 0.8590673 5.221635
## 6051 113.7029 0.9053650 4.954524
## 6101 111.2623 0.9591552 4.792757
## 6151 113.9549 0.9026097 4.987096
## 6201 117.3223 0.8261667 5.039425
## 6251 113.3996 0.9095483 5.626173
```



Generating mean responses

```r
mu_at_50 <- with(post_draws,
                 beta0 + beta1 * 50)
```






```r
quantile(mu_at_50, probs = c(0.05, 0.95))
##       5%      95%
## 158.5648 159.7029
```