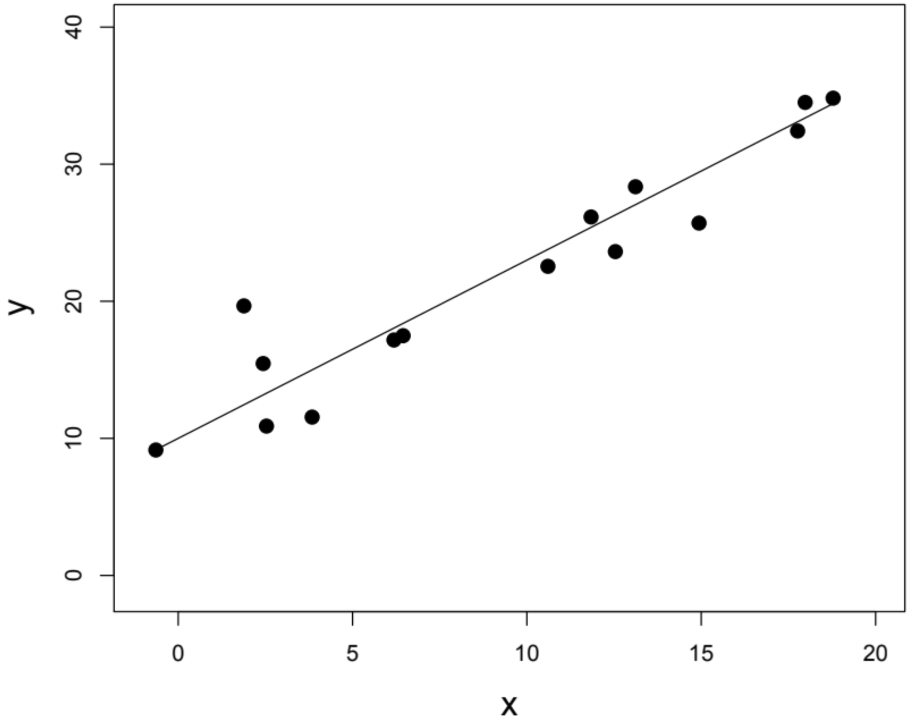
Inference from a Single Model

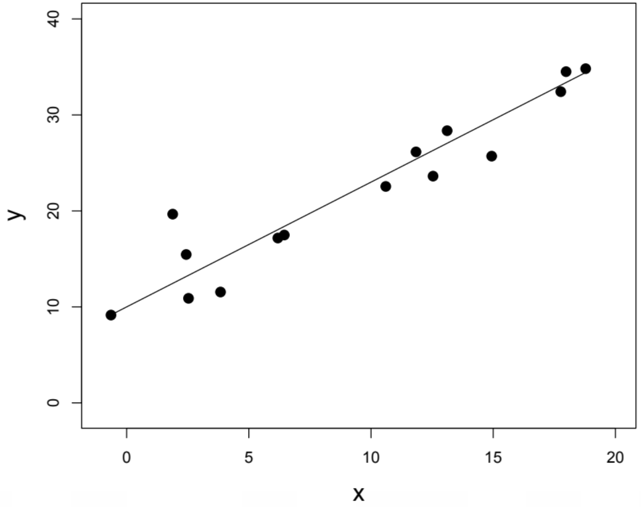
Alison Ketz

Slides adapted from Tom Hobbs, Chris Che-Castaldo, Mary B. Collins

A simple example: Linear Regression



A simple example: Linear Regression



$$
\begin{aligned}
g(\beta, x_i) &= \beta_0 + \beta_1 x_i \\
[\beta, \sigma^2 | \mathbf{y}] &\propto \prod_{i=1}^{15} \text{normal}\left(y_i | g\left(\beta, x_i\right), \sigma^2\right) \text{normal}\left(\beta_0 | 0, 1000\right) \times \\
&\quad \text{normal}\left(\beta_1 | 0, 1000\right) \times \\
&\quad \text{inverse gamma}\left(\sigma^2 | .001, .001\right)
\end{aligned}
$$

A simple example: Linear Regression

$$
\begin{aligned}
g(\beta, x_i) &= \beta_0 + \beta_1 x_i \\
[\beta, \sigma^2|\mathbf{y}] &\propto \prod_{i=1}^{15} \text{normal}\left(y_i|g\left(\beta, x_i\right), \sigma^2\right) \text{normal}\left(\beta_0|0, 1000\right) \times \\
& \quad \text{normal}\left(\beta_1|0, 1000\right) \times \\
& \quad \text{inverse gamma}\left(\sigma^2|.001, .001\right)
\end{aligned}
$$

```
model {

  # priors
  beta0 ~ dnorm(0,.001)
  beta1 ~ dnorm(0,.001)
  tau ~ dgamma(.001, .001)
  sigma_sq  <- 1/tau

  # likelihood
  for(i in 1:n) {
    mu[i] <- beta0 + beta1*x[i]
    y[i] ~ dnorm(mu[i], tau)
  }
}
```

A simple example: Linear Regression

Coda summary output includes posterior means and credible intervals

```
> summary(jc)

Iterations = 35001:45000
Thinning interval = 1
Number of chains = 1
Sample size per chain = 10000

1. Empirical mean and standard deviation for each variable,
   plus standard error of the mean:

           Mean     SD Naive SE Time-series SE
b0       10.486 1.3992 0.013992       0.032676
b1        1.227 0.1239 0.001239       0.002812
sigma.sq  9.065 4.4158 0.044158       0.055806

2. Quantiles for each variable:

           2.5%   25%    50%    75%  97.5%
b0       7.6559 9.607 10.489 11.371 13.264
b1       0.9832 1.148  1.226  1.305  1.476
sigma.sq 3.9819 6.206  8.054 10.666 20.128
```

A simple example: Linear Regression

Naive SE: The standard error assuming i.i.d. samples, calculated as SD/sqrt(K) where K = # of samples

```
> summary(jc)

Iterations = 35001:45000
Thinning interval = 1
Number of chains = 1
Sample size per chain = 10000

1. Empirical mean and standard deviation for each variable,
   plus standard error of the mean:

           Mean     SD Naive SE Time-series SE
b0       10.486 1.3992 0.013992       0.032676
b1        1.227 0.1239 0.001239       0.002812
sigma.sq  9.065 4.4158 0.044158       0.055806

2. Quantiles for each variable:

           2.5%   25%    50%    75%  97.5%
b0       7.6559 9.607 10.489 11.371 13.264
b1       0.9832 1.148  1.226  1.305  1.476
sigma.sq 3.9819 6.206  8.054 10.666 20.128
```

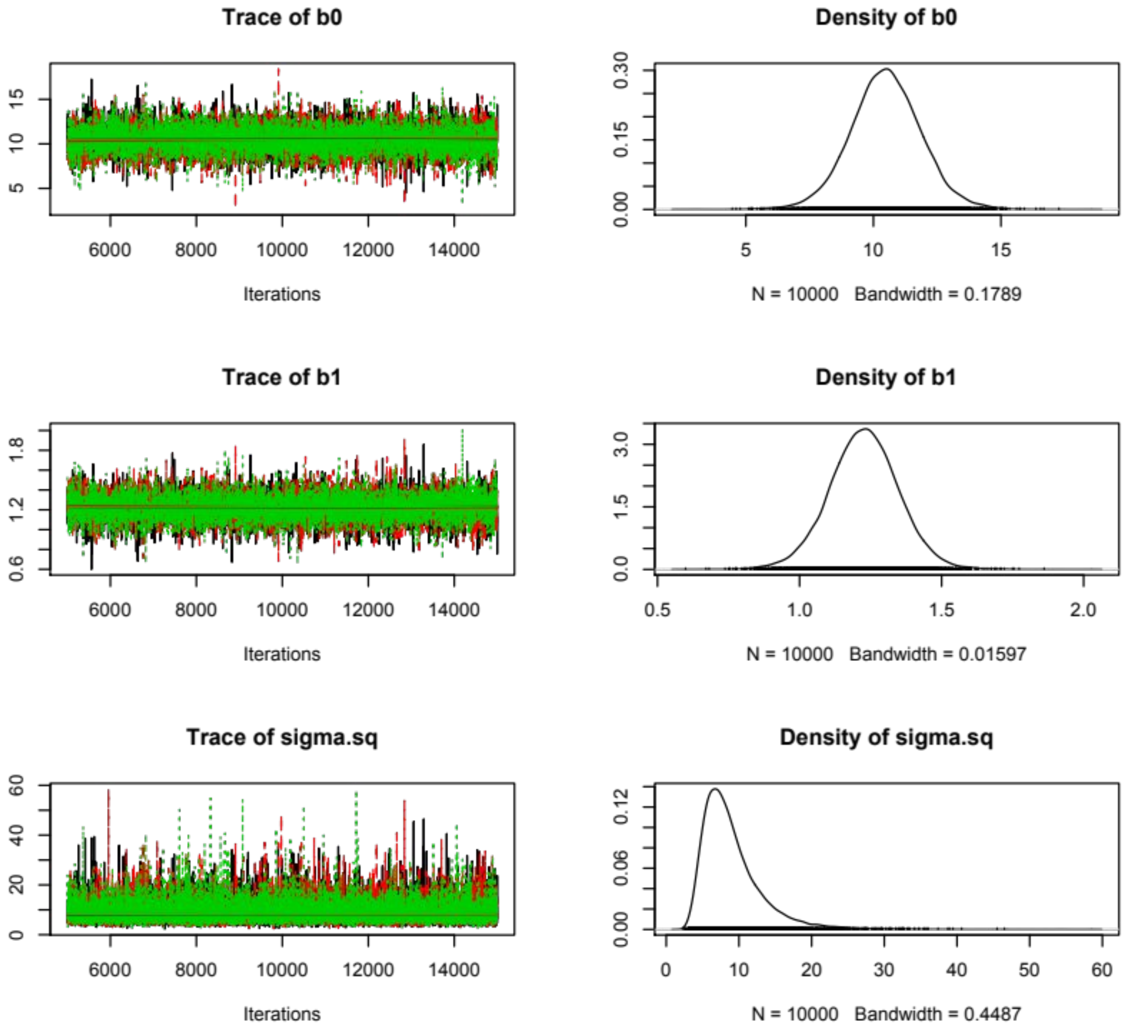
A simple example: Linear Regression

Time-series SE: The standard error adjusted for autocorrelation, often calculated using the effective sample size.

```
> summary(jc)

Iterations = 35001:45000
Thinning interval = 1
Number of chains = 1
Sample size per chain = 10000

1. Empirical mean and standard deviation for each variable,
   plus standard error of the mean:

           Mean     SD Naive SE Time-series SE
b0       10.486 1.3992 0.013992       0.032676
b1        1.227 0.1239 0.001239       0.002812
sigma.sq  9.065 4.4158 0.044158       0.055806

2. Quantiles for each variable:

           2.5%   25%    50%    75%  97.5%
b0       7.6559 9.607 10.489 11.371 13.264
b1       0.9832 1.148  1.226  1.305  1.476
sigma.sq 3.9819 6.206  8.054 10.666 20.128
```

A simple example: Linear Regression

Output from JAGS

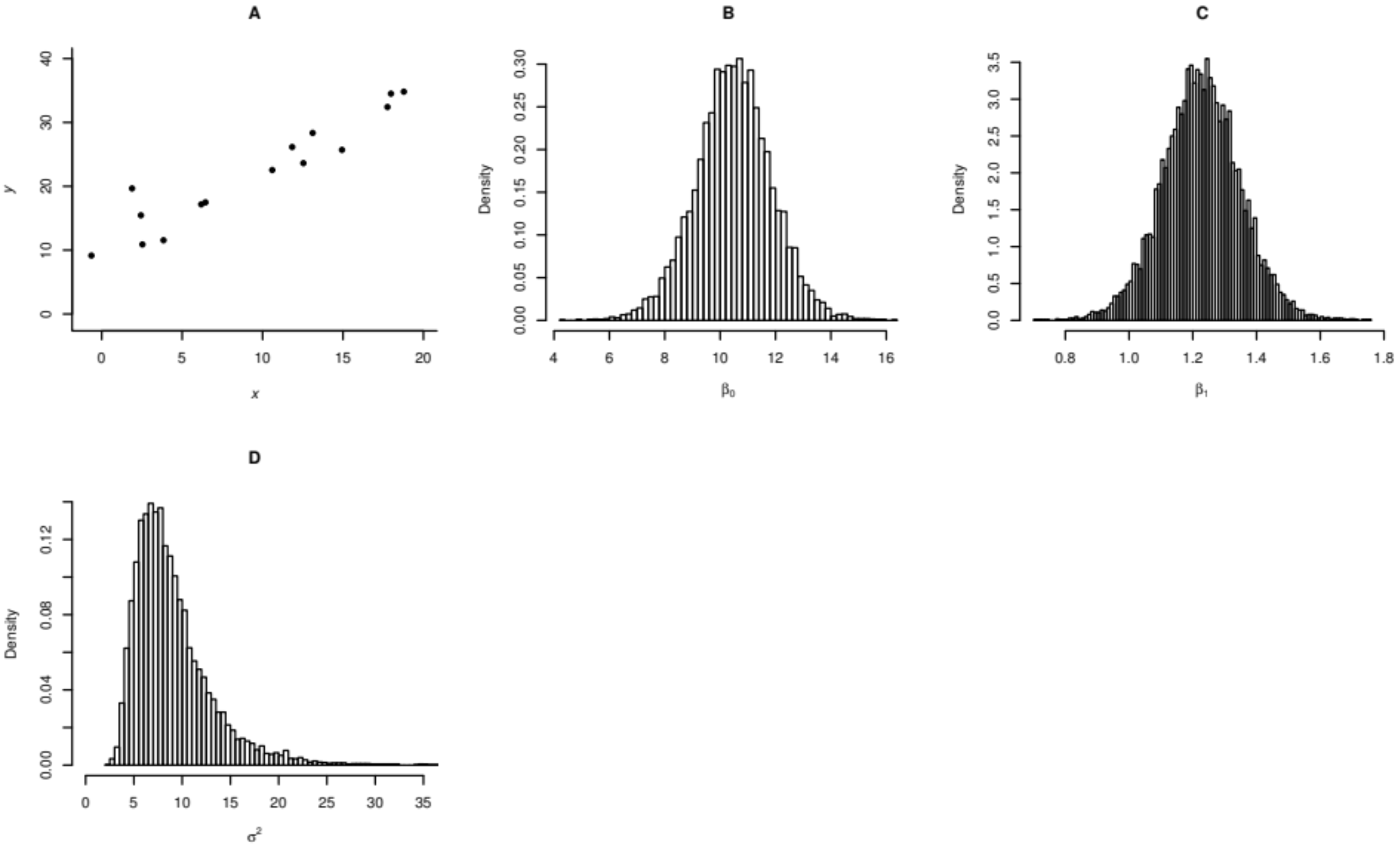


A simple example: Linear Regression

Output from JAGS

k	5001	5002	5003	5004	5005	…	9996	9997	9998	9999	10000
β_0	9.84	10.9	10.6	10.7	11.9	…	12.5	9.84	11.1	10.9	11.7
β_1	1.38	1.22	1.25	1.12	1.14	…	1.04	1.23	1.27	1.06	1.14
σ^2	10.8	6.32	4.96	4.57	5.76	…	12.4	9.77	8.19	6.88	13.1

A simple example: Linear Regression
Marginal Distributions of Parameters



Monte Carlo Integration
Marginal Distributions of Parameters

Use the row for each parameter, for example, β_0 to approximate moments of its marginal posterior distribution. For example, the mean is given analytically by the integral

$$
\begin{aligned}
[\beta_0|\mathbf{y}] &= \int_{\beta_1}\int_{\sigma^2}[\beta_1,\beta_0,\sigma^2 \mid \mathbf{y}]d\beta_1 d\sigma^2 \\
\mathrm{E}(\beta_o|\mathbf{y}) &= \int_{\beta_o}\beta_0\,[\beta_0|\mathbf{y}]\,d\beta_0,
\end{aligned}
$$

which is approximated by $$\mathrm{E}(\beta_0|\mathbf{y}) \approx \frac{1}{K}\sum_{k=1}^{K}\beta_0^{(k)}.$$

A simple example: Linear Regression
Marginal Distributions of Parameters

Another marginal posterior distribution of the moment of a distribution, for example, the variance, σ^2, is given analytically by the integral

$$\mathrm{var}(\beta_0|\mathbf{y}) \approx \frac{\sum_{k=1}^{K}\left(\beta_0^{(k)} - \frac{1}{K}\sum_{k=1}^{K}\beta_0^{(k)}\right)^2}{K}$$

We obtain other statistics (e.g., medians, quantiles, credible intervals, highest posterior density intervals) by applying the appropriate function to the row.

Bayesian Credible Intervals

- 95% equal-tailed Bayesian credible intervals
  - Convention from frequentist statistics
  - (Lower quantile, Upper quantile) = (.025,.975)
  - works for symmetric marginal posteriors
  - expected in journals

- Highest posterior density interval (board)

Median height of willows in dammed and fenced plots was 184 cm (highest posterior density interval, HPDI = 173, 192)

Median height of willows in dammed and fenced plots was 184 cm (equal-tailed Bayesian Credible Interval, BCI = 174, 194)

What else can we do with a model?... Predictions

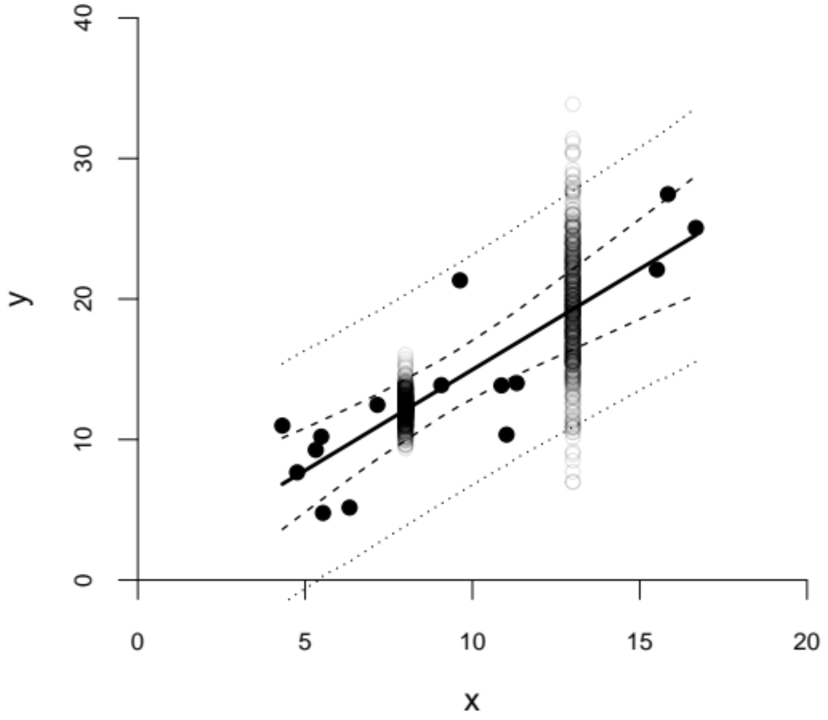
We want to know the distribution of the mean of the response when the predictor variable equals x_4

$$\mu_4 = \beta_0 + \beta_1 x_4.$$

We also want to know the distribution of a new observation at x_4

$$y_4^{new} \sim \text{normal}(\mu_4, \sigma^2)$$

What can we do with a model?... Predictions



Predicting a mean response:

The **posterior predictive distribution** of the mean

$$[\mu_4|\mathbf{y}] = \int_{\beta_o}\int_{\beta_1}\int_{\sigma^2} [\mathrm{E}(\ \mu_4\ )|\beta_0,\beta_1,\sigma^2]\,[\beta_0,\beta_1,\sigma^2|\mathbf{y}]\,d\beta_0 d\beta_1 d\sigma^2,$$

which we approximate by calculating

$$\mu_4^{(k)} = \beta_0^{(k)} + \beta_1^{(k)} x_4$$

at each iteration of the MCMC algorithm. Statistics can be calculated from each iteration of

Predicting a new observation: Composition Sampling

The **posterior predictive distribution** of y_4

$$[y_4^{new}|\mathbf{y}] = \int_{\beta_o}\int_{\beta_1}\int_{\sigma^2} \left[y_4^{new}|\beta_0,\beta_1,\sigma^2\right]\left[\beta_0,\beta_1,\sigma^2|\mathbf{y}\right] d\beta_0 d\beta_1 d\sigma^2,$$

We approximate this integral by drawing a new observation, y_4^{new}, at each MCMC iteration

$$y_4^{new(k)} \sim \text{normal}(\beta_0^{(k)} + \beta_1^{(k)} x_4, \sigma^{2(k)})$$

Note: this is used for posterior predictive checks (coming soon).

Predicting a new observation with MCMC

```
> y.new[,4]
     2.5%       50%     97.5%
 7.146763 13.618321 19.975031
> mu[,4]
    2.5%      50%    97.5%
11.30646 13.60504 15.92383
>
```

What else could this be used for?

```
model {

  # priors
  beta0 ~ dnorm(0,.001)
  beta1 ~ dnorm(0,.001)
  tau ~ dgamma(.001, .001)
  sigma.sq  <- 1/tau

  # likelihood
  for(i in 1:n) {
    mu[i] <- beta0 + beta1*x[i]
    y[i] ~ dnorm(mu[i], tau)
  }

  #derived quantities
  #mu4.new <- beta0 + beta1*x[4]
  y.new ~ dnorm(mu[4],tau)
}
```

A simple example: Linear Regression

Output from JAGS

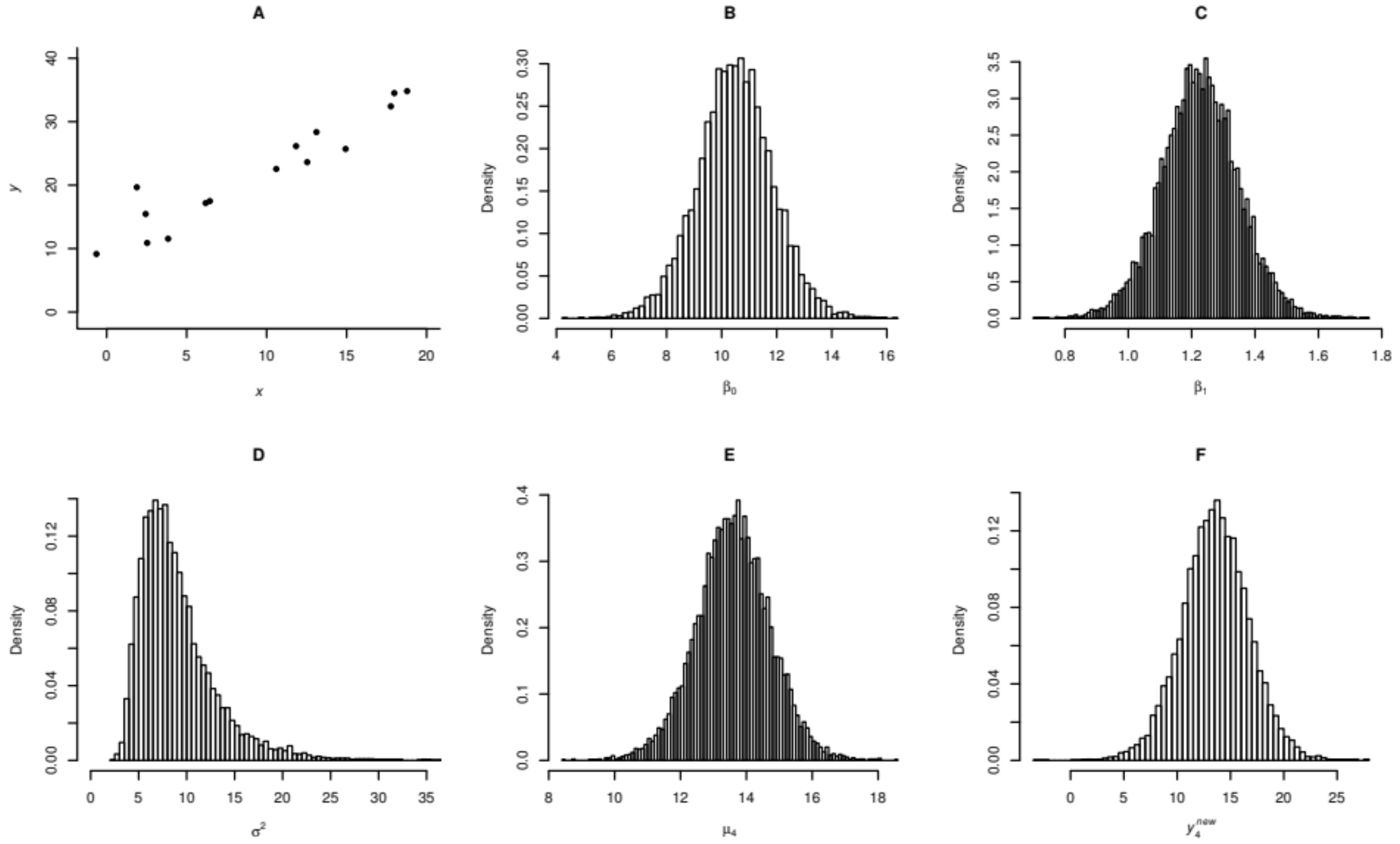
k	5001	5002	5003	5004	5005	...	9996	9997	9998	9999	10000
β_0	9.84	10.9	10.6	10.7	11.9	...	12.5	9.84	11.1	10.9	11.7
β_1	1.38	1.22	1.25	1.12	1.14	...	1.04	1.23	1.27	1.06	1.14
σ^2	10.8	6.32	4.96	4.57	5.76	...	12.4	9.77	8.19	6.88	13.1
μ_4	13.3	14	13.8	13.6	14.7	...	15.1	12.9	14.3	13.6	14.6
y_4^{new}	9.29	14.3	9.44	15.3	18.5	...	11.4	16.8	12.8	14.5	15.8

A simple example: Linear Regression

Output from JAGS

k	5001	5002	5003	5004	5005	...	9996	9997	9998	9999	10000
β_0	9.84	10.9	10.6	10.7	11.9	...	12.5	9.84	11.1	10.9	11.7
β_1	1.38	1.22	1.25	1.12	1.14	...	1.04	1.23	1.27	1.06	1.14
σ^2	10.8	6.32	4.96	4.57	5.76	...	12.4	9.77	8.19	6.88	13.1
μ_4	13.3	14	13.8	13.6	14.7	...	15.1	12.9	14.3	13.6	14.6
y_4^{new}	9.29	14.3	9.44	15.3	18.5	...	11.4	16.8	12.8	14.5	15.8

A simple example: Linear Regression
Marginal Distributions of Parameters



Derived Quantities

- Any quantity derived from a random variable becomes a random variable

- All random variables can have their own posterior distribution

- Obtain by calculating the value at each iteration based on current values of other parameters at that iteration

- Can be done within model code, or after the MCMC has run

- Inference on any function of parameters:
    1. Difference between means
    2. Ratios of means
    3. Eigen analysis (population growth rate)
    4. Forecasts in time series
    5. Indices (Shannon diversity index)

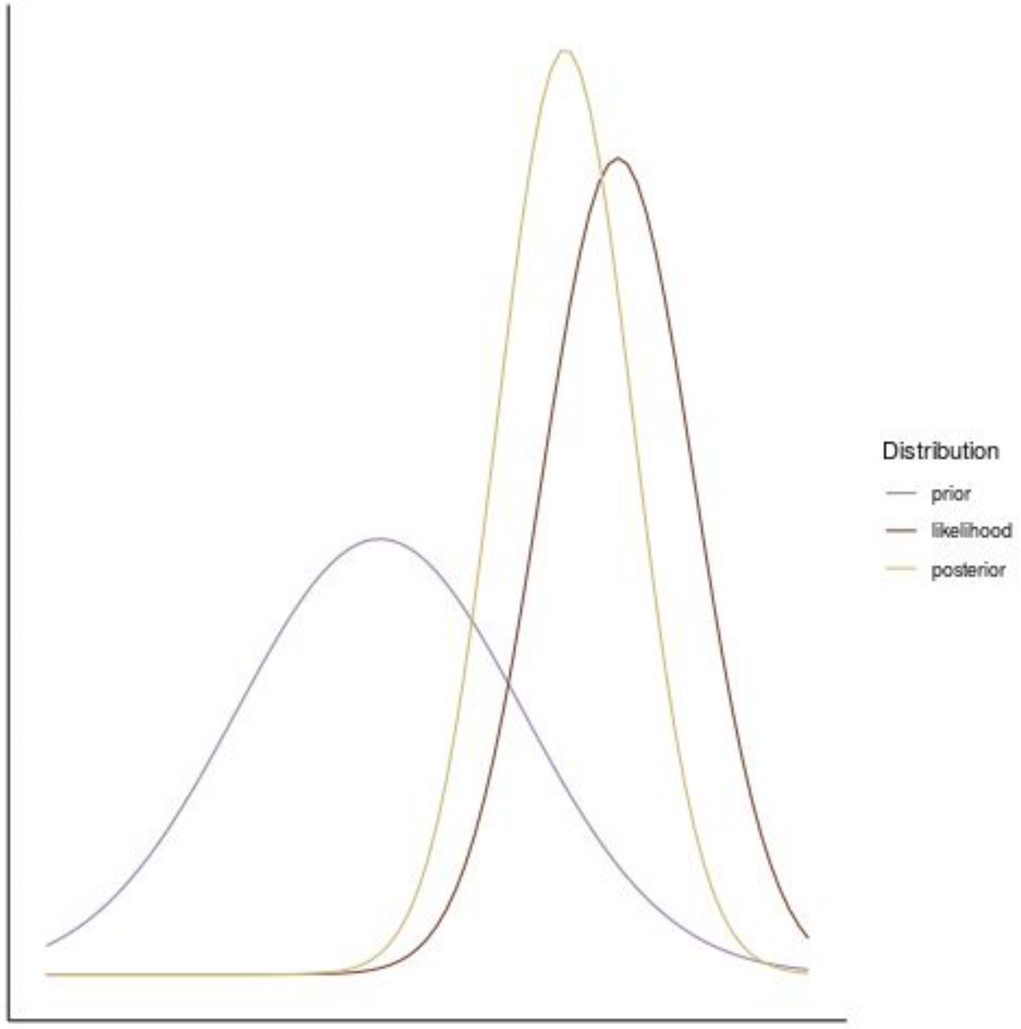
Alternative library for fitting JAGS models:
jagsUI

https://cran.r-project.org/web/packages/jagsUI/index.html

Another library for fitting Bayesian models for more advanced users:

nimble

https://r-nimble.org/


Distribution
prior
likelihood
posterior