Bayesian Analysis Module III: Deep Dive

Dr. Mark Williamson

DaCCoTA

University of North Dakota

DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UND
UNIVERSITY OF NORTH DAKOTA.

Introduction

- Last time, we covered some surface details about the computational process of Bayesian Analysis and looked at simple examples in R, SAS, and SPSS

- Today, we'll work on more details under the hood when it comes to computation
- We will also look at detailed examples in R and SAS

DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UND UNIVERSITY OF NORTH DAKOTA

Reviewing the Basics

Bayes' Theorem

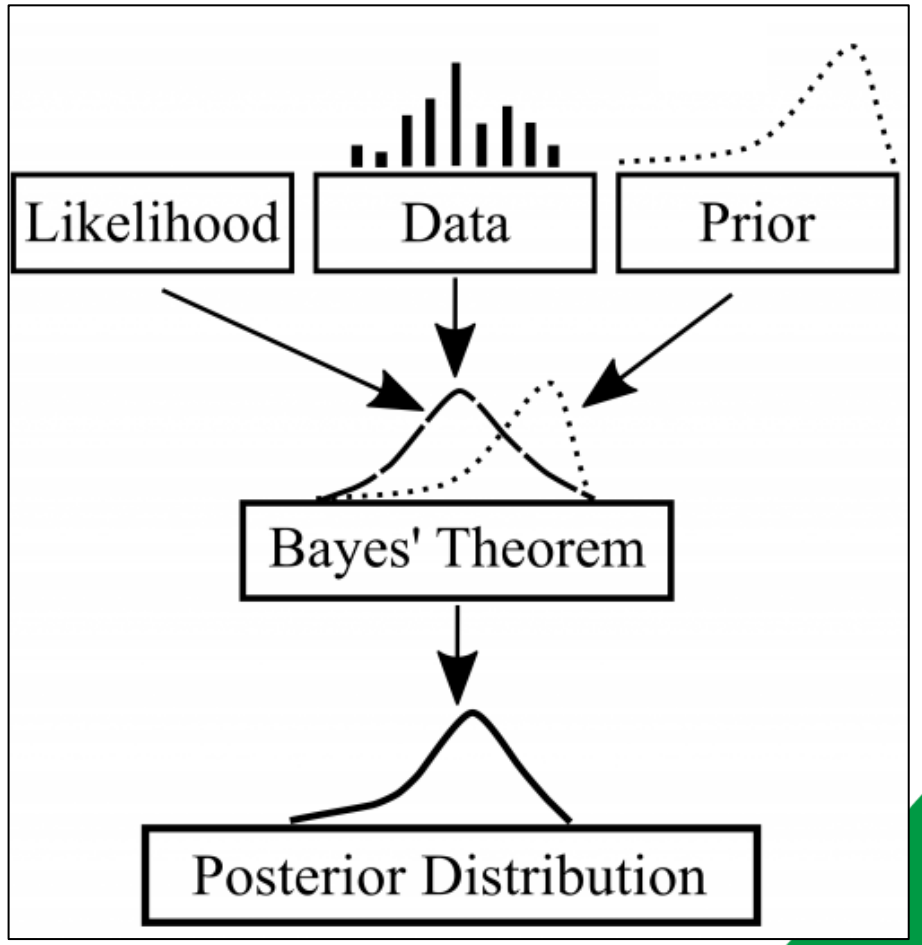
- Updates probabilities (degrees of belief) after obtaining new data
- Probability of A occurring given that B has occurred is equal to the probability that they have both occurred, relative to the probability that B has occurred

Bayes Theorem:

$$P(hyp|data) = \frac{P(data|hyp) \times P(hyp)}{P(data)}$$

Bayes Theorem:

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$



DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UND
UNIVERSITY OF NORTH DAKOTA.

Reviewing the Basics 2

Items needed for analysis

- Priors
- Model

Parameters

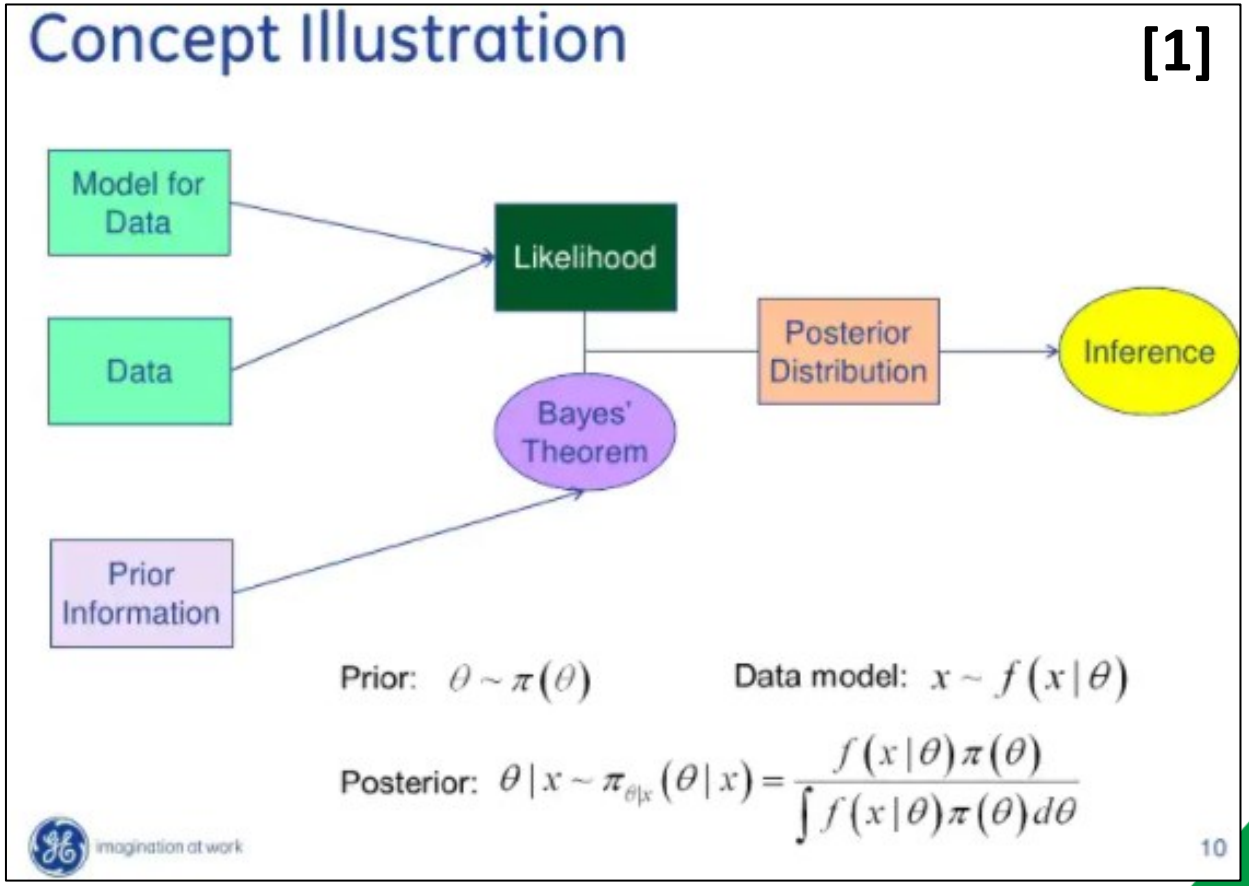
- Data

Steps for analysis (JAGS)

- Define (Model + Priors)
- Compile (Data + Model + Priors)
- Simulate (Posterior)

Posterior simulation details

- MCMC burn-in, iterations, thinning, chain number
- Trace plots, autocorrelation lag plots, and PDFs

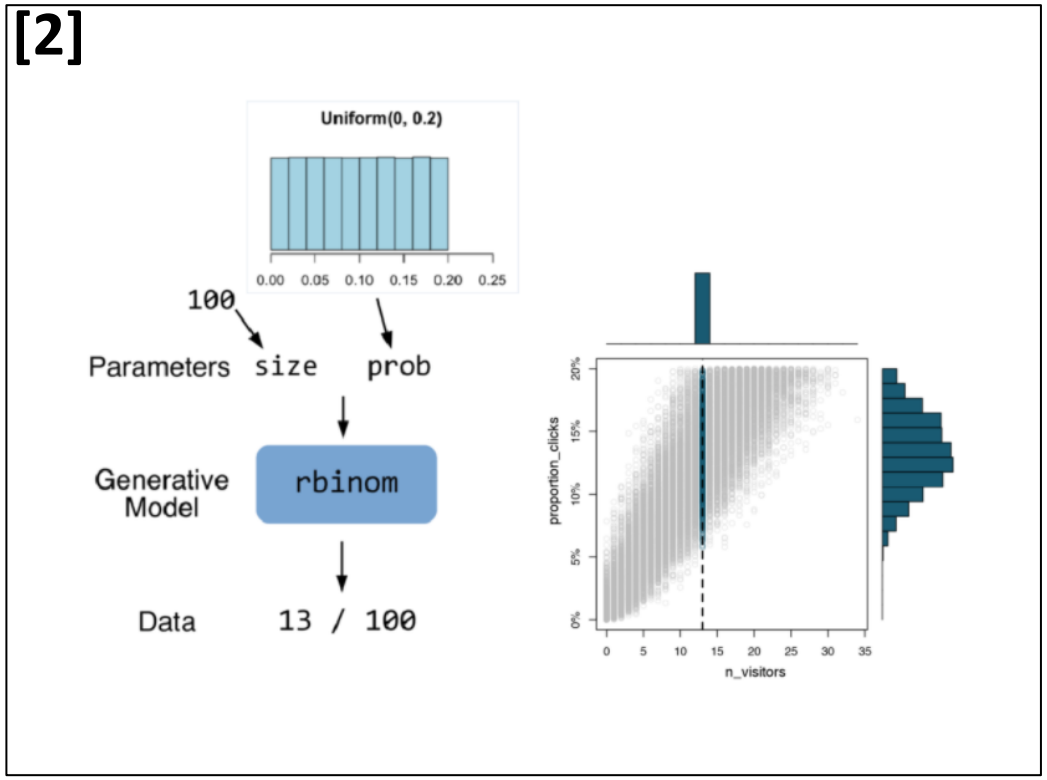


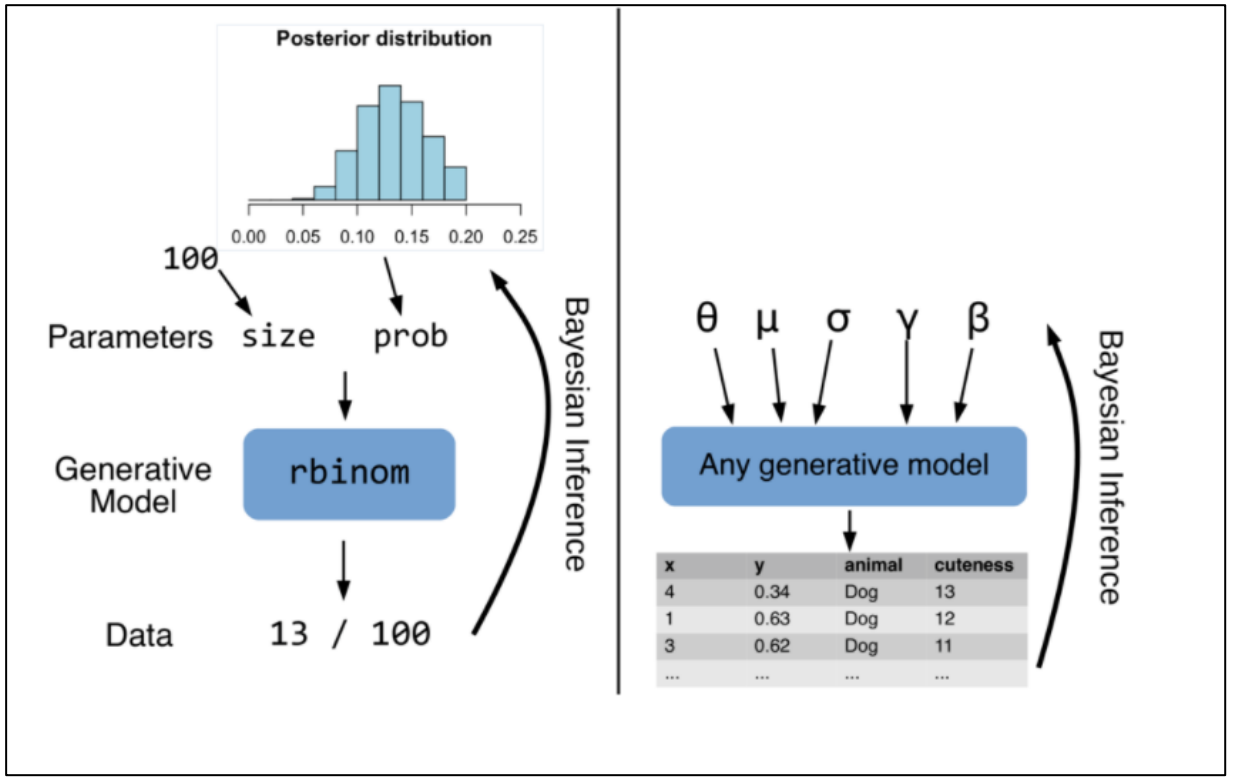
DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UNIVERSITY OF NORTH DAKOTA

Reviewing the Basics 3

[2]





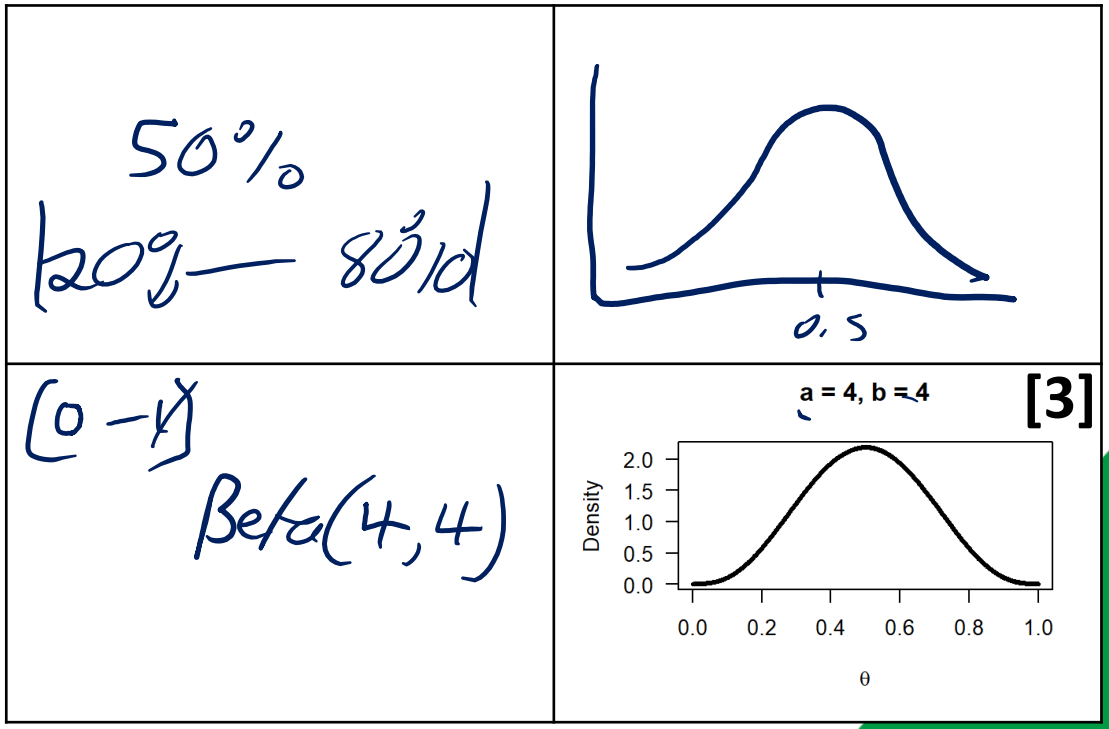
DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UND UNIVERSITY OF NORTH DAKOTA.

Details

Defining Prior

- Summary of what we know of the situation
- May come from expertise in the field or past experiments/observations
- May be very precise or very vague
- Fundamentally a probability distribution for a variable/parameter
- How to get it?
  1. Conceptualize it (mean, spread, limits)
  2. Draw it
  3. Determine distribution and parameters
  4. Model it in software package



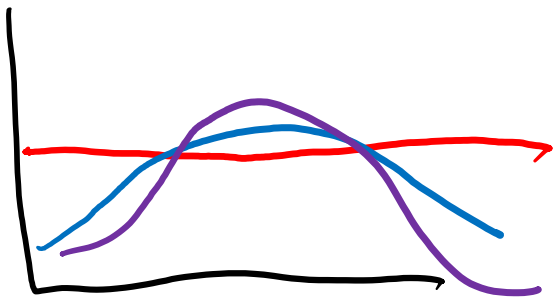
DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UND
UNIVERSITY OF NORTH DAKOTA.

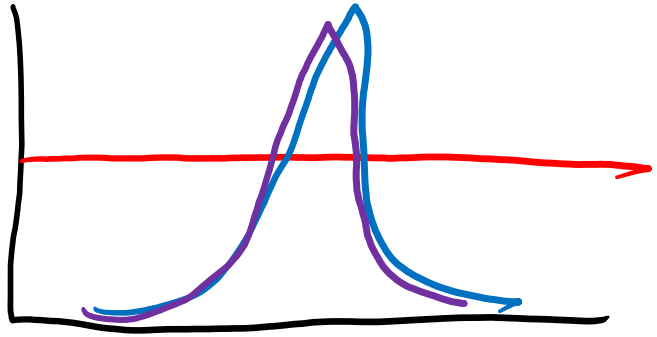
Details 2

Effects of priors and data on posterior

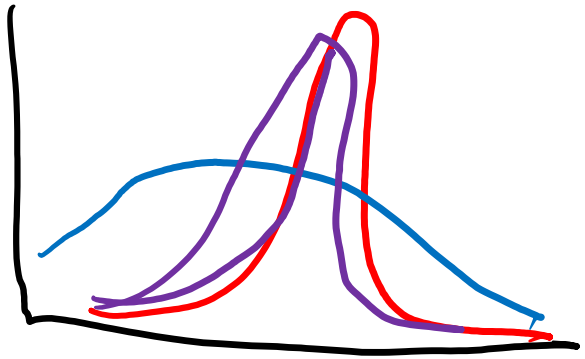
vague prior + vague data



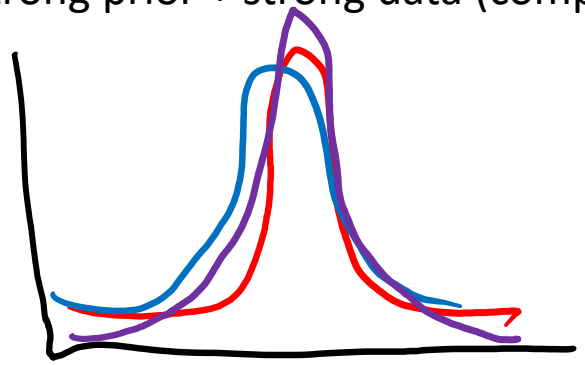
vague prior + strong data



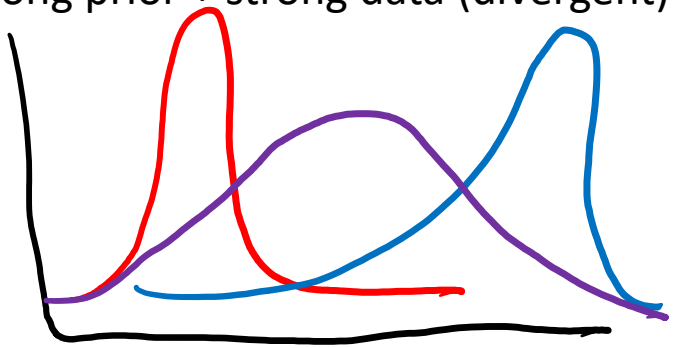
strong prior + vague data



strong prior + strong data (compliment)



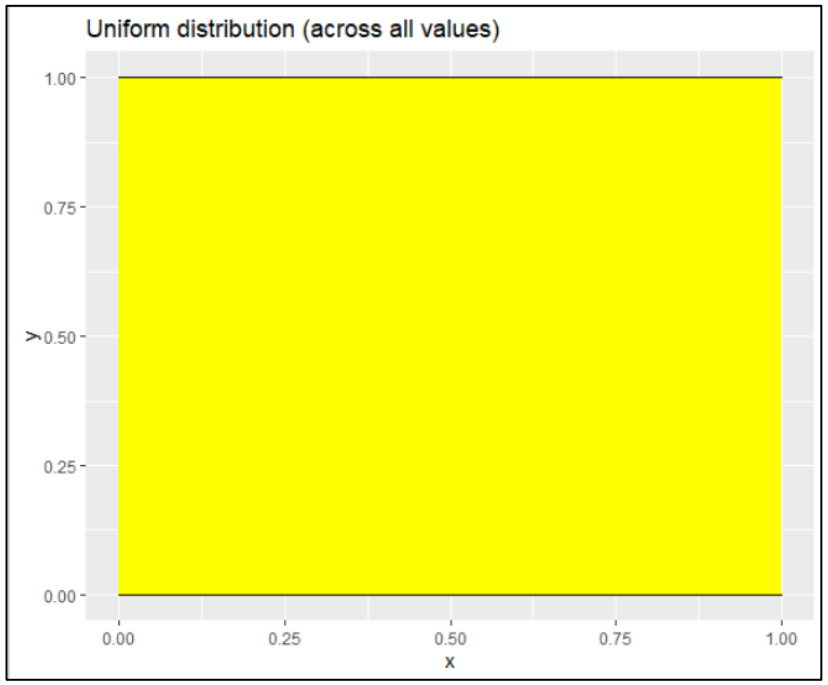
strong prior + strong data (divergent)



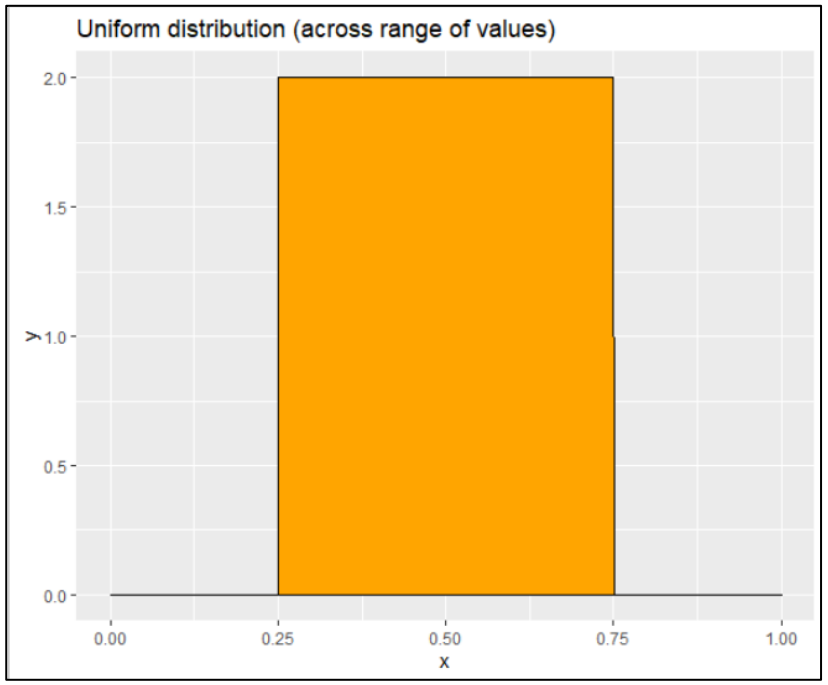
	Prior
	Data
	Posterior

Details 3.1

Prior Distribution Examples



Uniform(min=0, max=1)



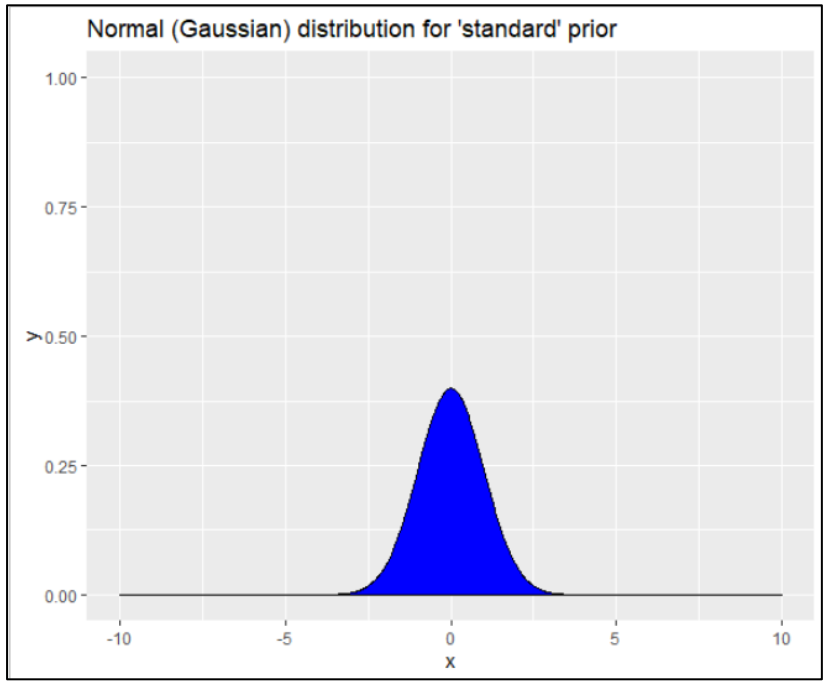
Uniform(min=0.25, max=0.75)

DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

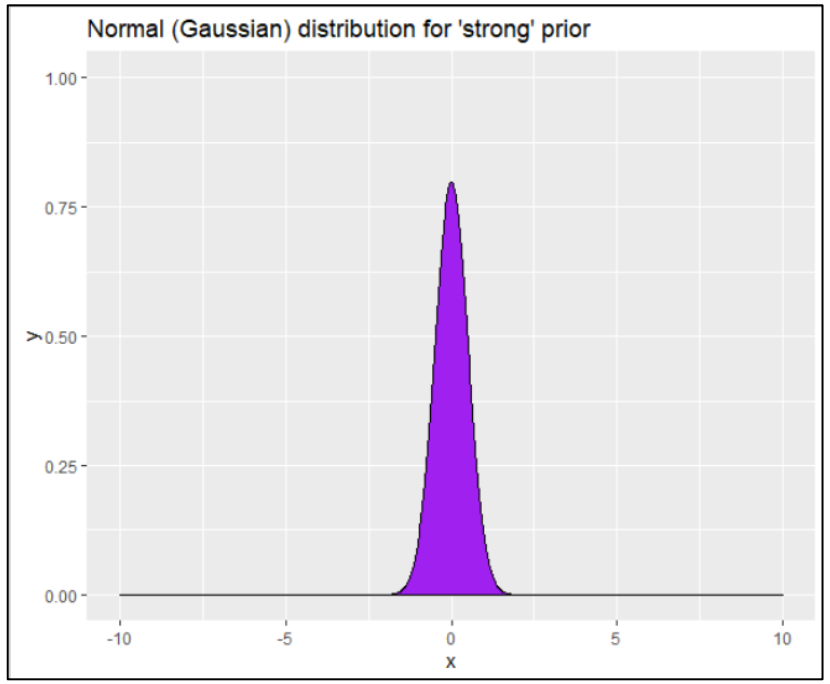
UND UNIVERSITY OF NORTH DAKOTA.

Details 3.2

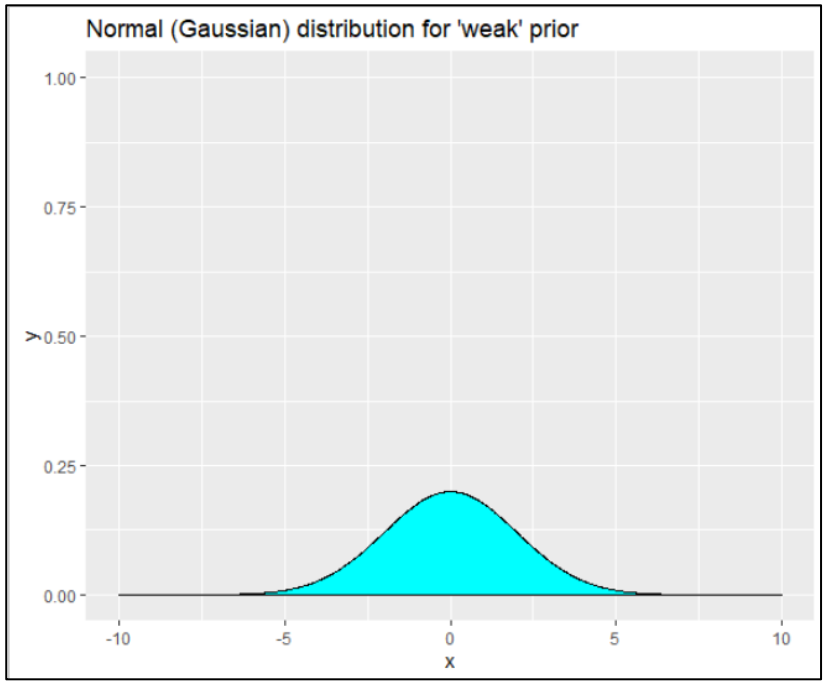
Prior Distribution Examples



Normal($\mu=0$, $\sigma^2=1$)



Normal($\mu=0$, $\sigma^2=2$)



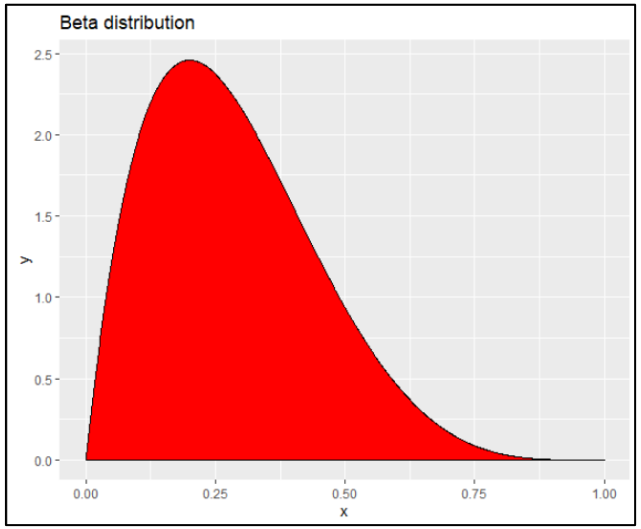
Normal($\mu=0$, $\sigma^2=0.5$)

DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

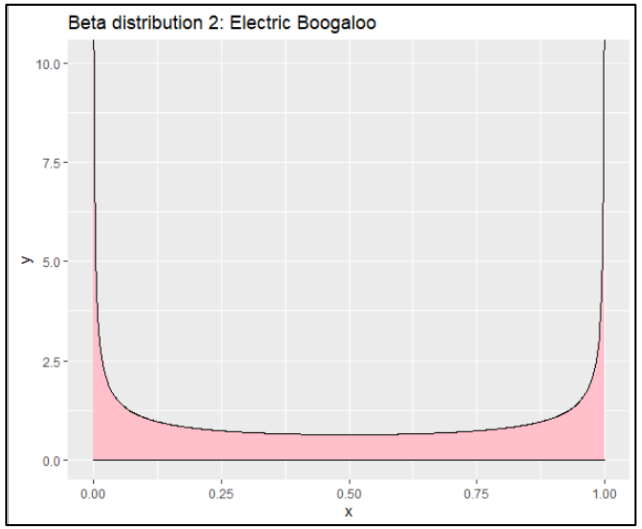
UND UNIVERSITY OF NORTH DAKOTA.

Details 3.3

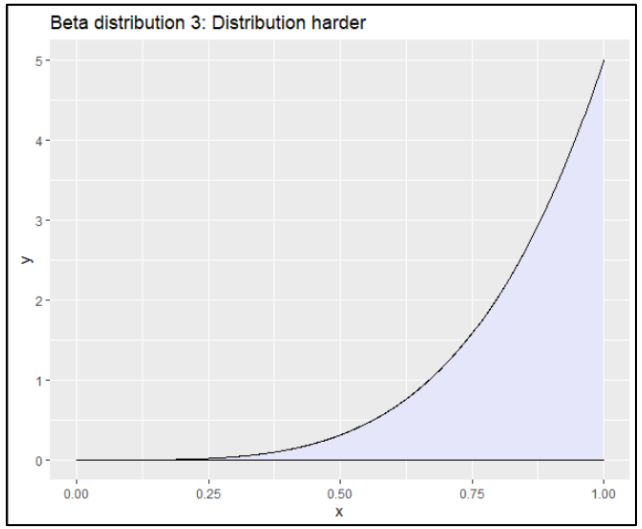
Prior Distribution Examples



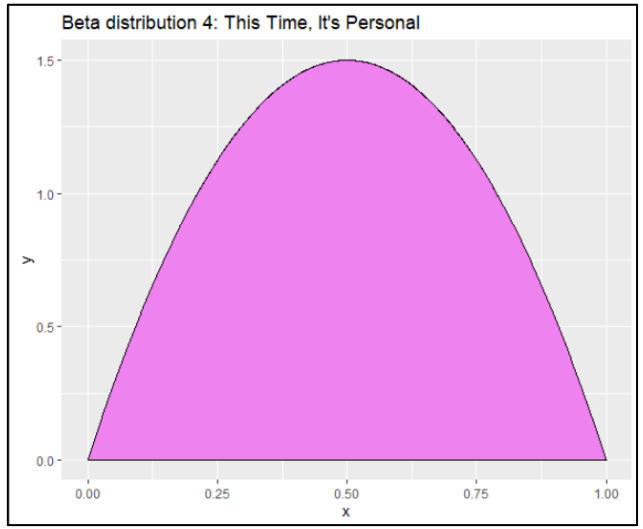
Beta(α=2, β=5)



Beta(α=0.5, β=0.5)



Beta(α=5, β=1)

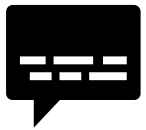


Beta(α=2, β=2)

DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UND UNIVERSITY OF NORTH DAKOTA.

Details 4



Likelihood Model

- AKA generative model
- Function defined by parameters and priors (probability)
- Model will run using the data and the defined parameters and priors
- Specific structure depends on what you are doing

$$y = mx + b$$

$$y = normal(mu, var = s2)$$

$$mu = b0 + b1 * x$$

…

Beta-binomial	$X \sim Bin(n, p)$	$p \sim Beta(a, b)$
Normal-normal	$Y_i \sim N(m, s^2)$	$m \sim N(m, s^2)$ $s \sim Unif(min, max)$
Regression	$Y_i \sim N(m_i, s^2)$	$m_i = a + bX_i$ $a \sim N(m, s^2)$ $b \sim N(m, s^2)$ $s \sim Unif(min, max)$
Multivariate	$Y_i \sim N(m_i, s^2)$	$m_i = a + bX_i + cZ_i$ $a \sim N(m, s^2)$ $b \sim N(m, s^2)$ $c \sim N(m, s^2)$ $s \sim Unif(min, max)$
Poisson	$Y_i \sim Pois(l_i)$	$\log(l_i) = a + bX_i + cZ_i$ $a \sim N(m, s^2)$ $b \sim N(m, s^2)$ $c \sim N(m, s^2)$

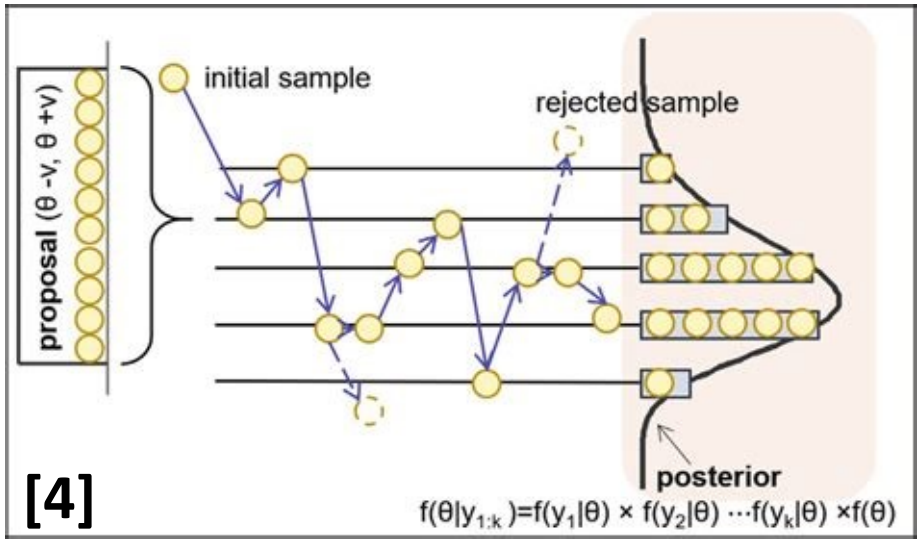
DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UNIVERSITY OF NORTH DAKOTA

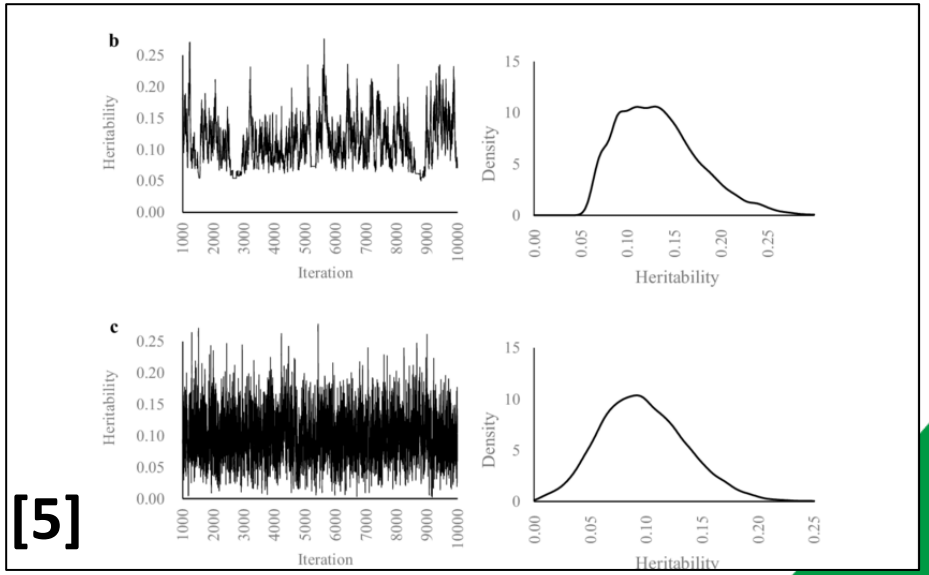
Details 5

Compiling and Simulating

- Model is passed to be compiled
- Includes likelihood and priors
- Simulation will produce posterior probability density functions, along with diagnostic information
  - MCMC -> Markov Chain Monte Carlo
  - Markov Chains -> random walks where each step depends on previous step
  - Explores spaces that can't be fully computed



[4]



[5]

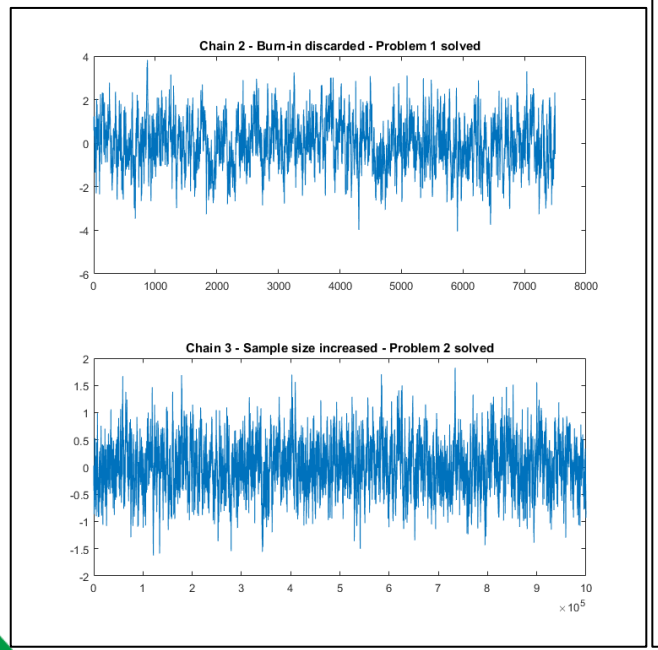
DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UND UNIVERSITY OF NORTH DAKOTA.

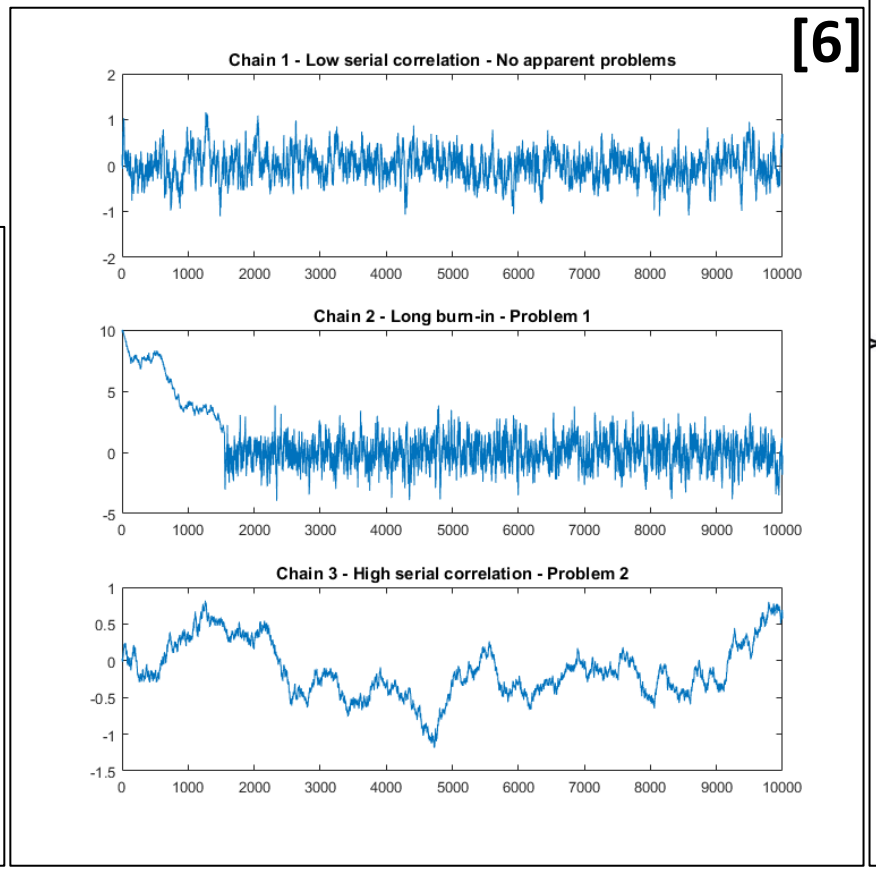
Details 6

Compiling and Simulating cont.

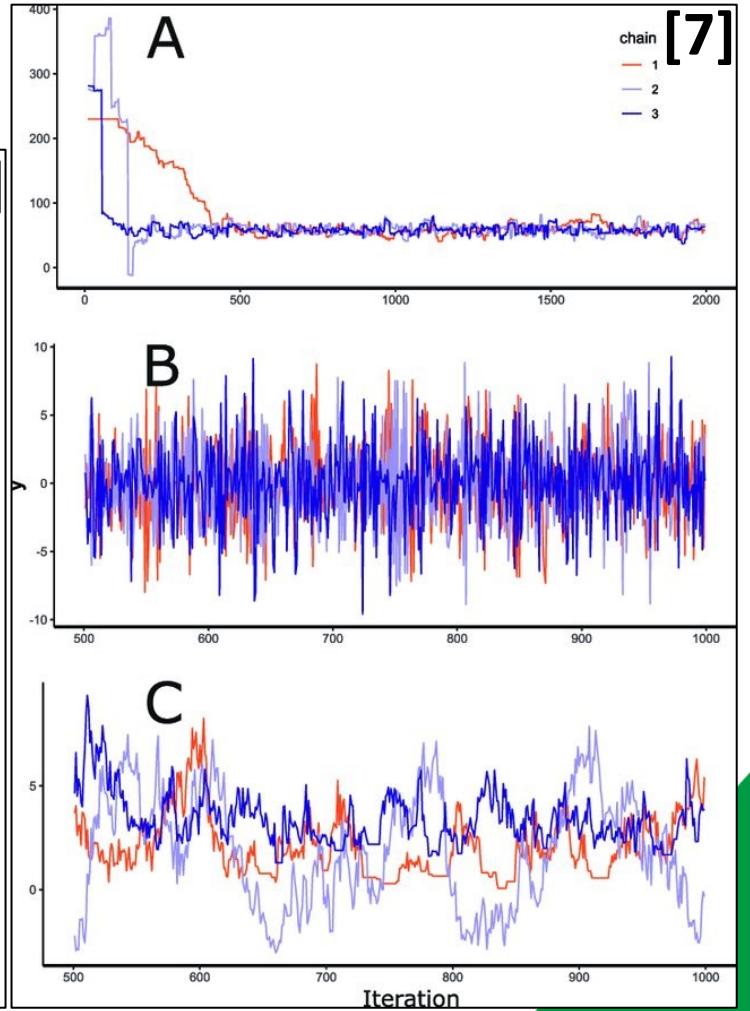
- Setting MCMC simulation parameters:
  - burn-in
  - iterations
  - chain number



[6]



[7]

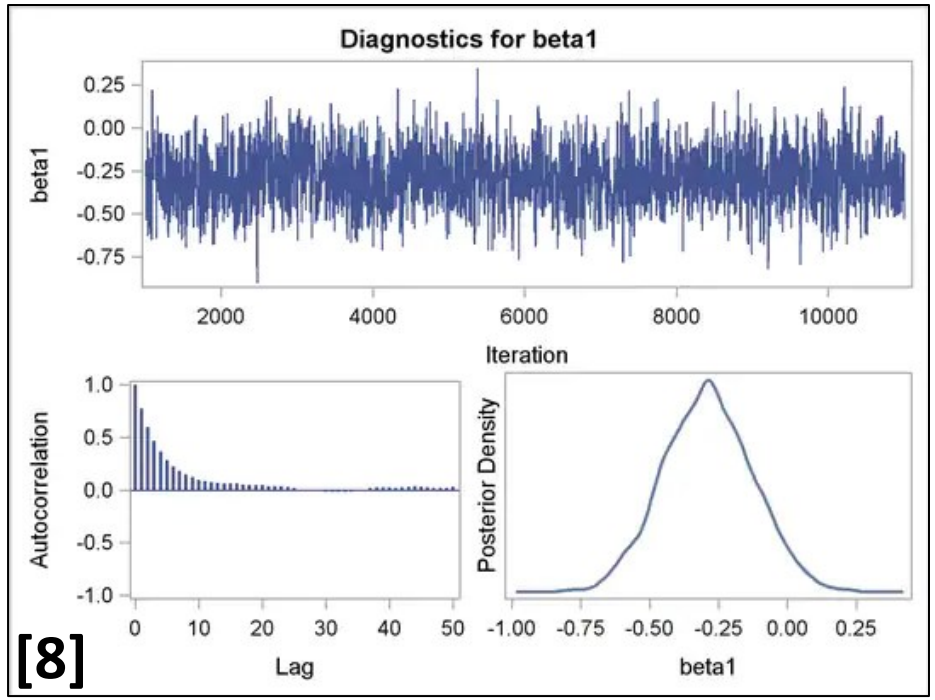


Details 7

Compiling and Simulating cont.

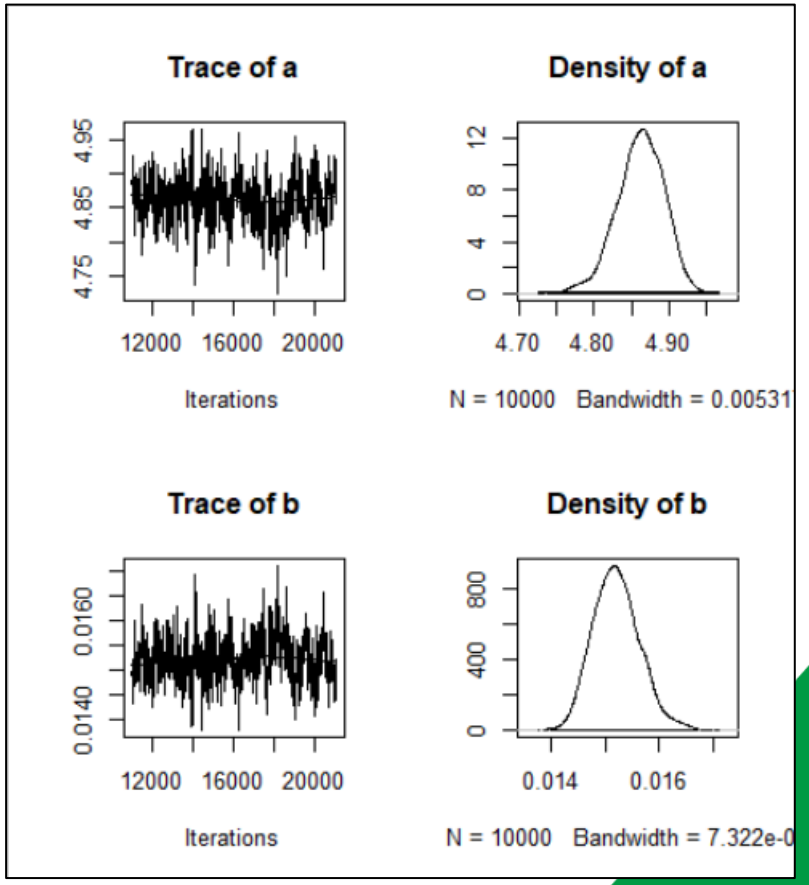
- Diagnostics:
  - trace plots
  - autocorrelation lag plots
  - PDFs and intervals

SAS



[8]

R



DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UND
UNIVERSITY OF NORTH DAKOTA.

Assessment 1

qualtrics XM

UNIVERSITY OF NORTH DAKOTA UND

https://und.qualtrics.com/jfe/form/SV_1QQ9MOt6bMsinnU

DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UND
UNIVERSITY OF NORTH DAKOTA.

Examples setup

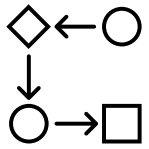
R[9]	**SAS**[10]
Package **rjags** • Probability of single value • Simple linear regression • Multiple linear regression • Poisson regression	PROC **MCMC** • Simple linear regression • Mixed-effects model • Logistic regression

R-code available at: https://med.und.edu/daccota/_files/docs/berdc_docs/bayesian_analysis_3_r_code.txt
SAS-code available at: https://med.und.edu/daccota/_files/docs/berdc_docs/bayesian_analysis_3_sas_code.txt

DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UNIVERSITY OF NORTH DAKOTA

Step-by-step Example 1



Bayesian Analysis in R

General rjags in R

Analysis using the rjags package proceeds in steps:

1. Define the model using the BUGS language in a separate file.
2. Read in the model file using the jags.model function. This creates an object of class “jags”.
3. Update the model using the update method for “jags” objects. This constitutes a ‘burn-in’ period.
4. Extract samples from the model object using the coda.samples function. This creates an object of class “mcmc.list” which can be used to summarize the posterior distribution. The coda package also provides convergence diagnostics to check that the output is valid for analysis (see Plummer et al 2006).

Our steps

1. View priors
2. View data
3. Define model
4. Compile model
5. Simulate posterior
6. Plot results

DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UNIVERSITY OF NORTH DAKOTA

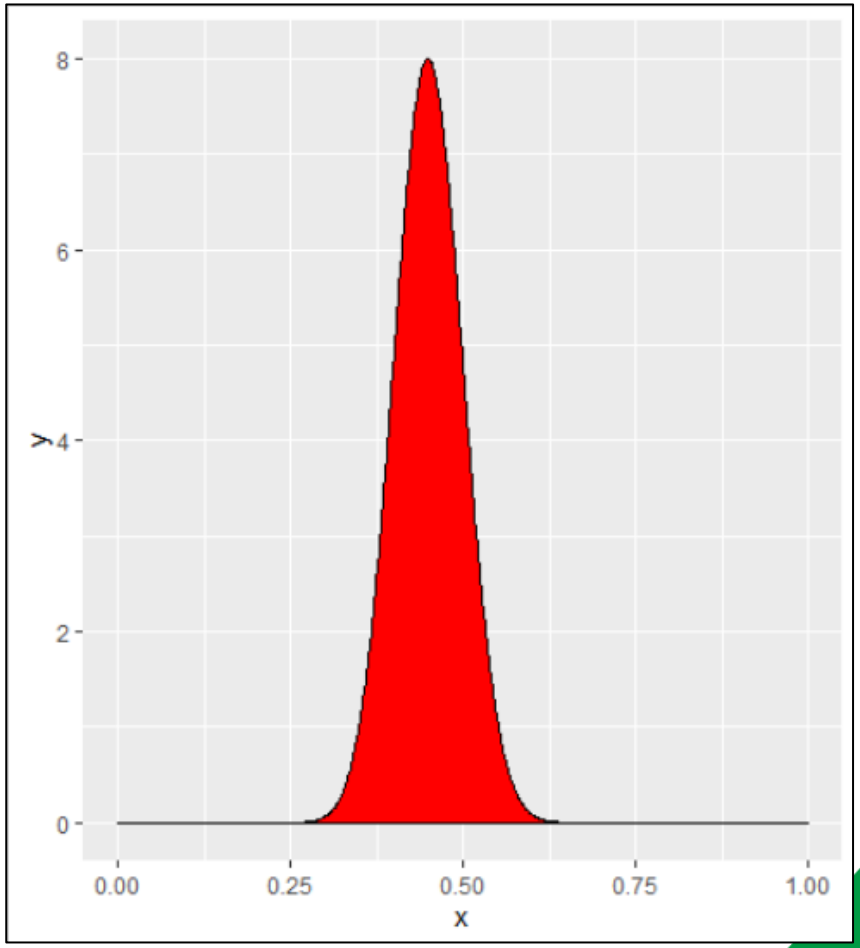
Step-by-step Example 1.1

Probability of single value: What are chances of winning a vote?

```
 #VIEW prior: think you fall under 50 percent of vote
ggplot(data = tibble(x = 0:1), aes(x)) +
  stat_function(fun = dbeta, n = 1000, args = list(shape1=45,
shape2=55),
              geom='area', colour='black', fill='red')

  #VIEW data: small poll (data) found 6/10 would vote for you
#N/A

  #DEFINE model
vote_model <-"model{
  #Likelihood model for X
  X ~dbin(p, n)
  #Prior model for p
  p ~dbeta(a, b)
}"
```



Step-by-step Example 1.1

Probability of single value cont.

```
 #COMPILE model
vote_jags <- jags.model(textConnection(vote_model),
   data  =list(a=45, b=55, X=6, n=10),
   inits =list(.RNG.name="base::Wichmann-Hill",
.RNG.seed=100))

  #SIMULATE posterior
vote_sim <- coda.samples(model=vote_jags,
   variable.names=c("p"), n.iter=10000)

  #PLOT results
summary(vote_sim)
plot(vote_sim)
```

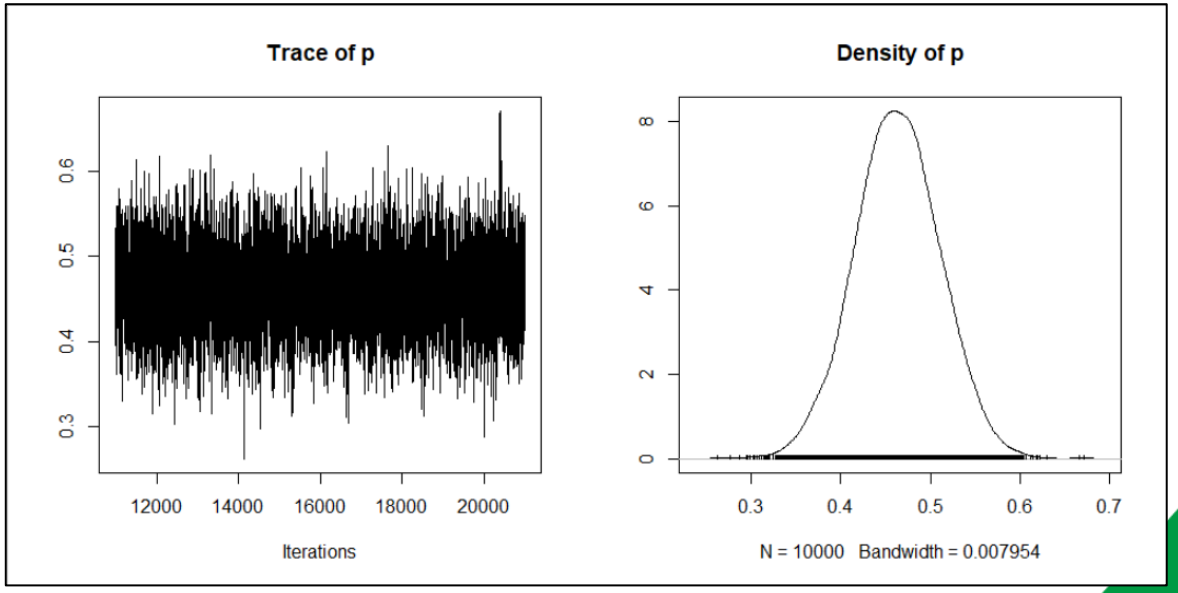
```
Iterations = 11001:21000
Thinning interval = 1
Number of chains = 1
Sample size per chain = 10000

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

          Mean             SD       Naive SE Time-series SE
     0.4631767      0.0473438      0.0004734      0.0006326

2. Quantiles for each variable:

  2.5%   25%   50%   75% 97.5%
0.3703 0.4313 0.4630 0.4949 0.5559
```



DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

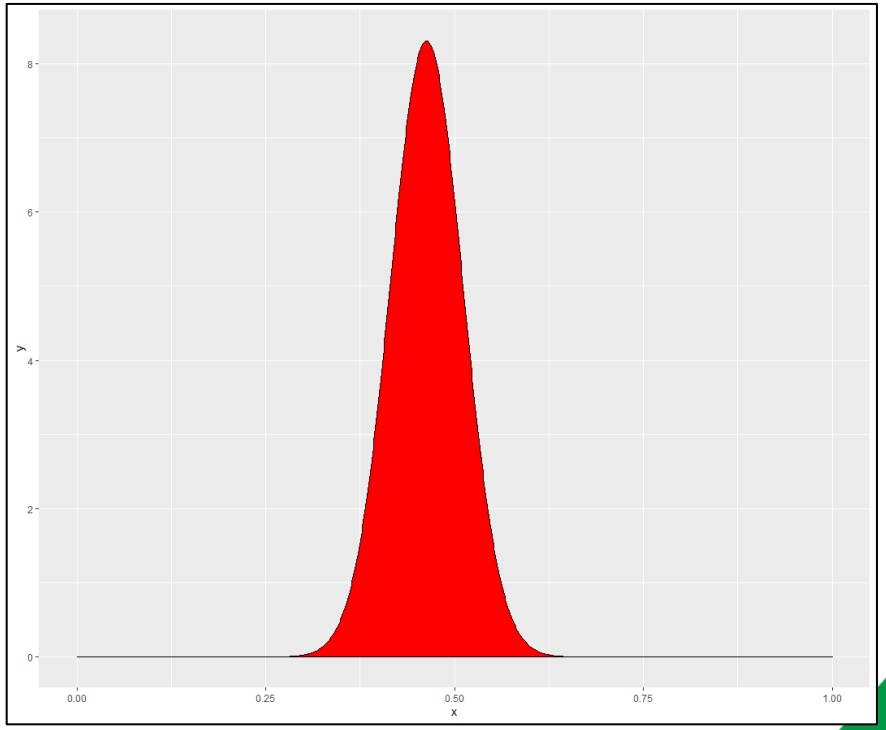
UNIVERSITY OF NORTH DAKOTA

Step-by-step Example 1.1

Probability of single value update

```
 #RETREAD USING UPDATED DATA
#new prior
ggplot(data = tibble(x = 0:1), aes(x)) +
  stat_function(fun = dnorm, n = 1000, args = list(mean=0.463, sd=0.048),
                geom='area', colour='black', fill='red')

vote_model_2 <-"model{
  #Likelihood model for X
  X ~ dbin(p, n)
  #Prior models for p
  p ~ dnorm(0.463, 0.048)
}"
```



Step-by-step Example 1.1

Probability of single value update cont.

```
#updated data -> now 688 out of 1000 people said they'd vote for you
vote_jags_2 <- jags.model(textConnection(vote_model_2),
   data  =list(X=688, n=1000),
   inits =list(.RNG.name="base::Wichmann-Hill", .RNG.seed=100))

vote_sim_2 <- coda.samples(model=vote_jags_2,
                  variable.names=c("p"), n.iter=10000)

 #PLOT results
summary(vote_sim_2)
plot(vote_sim_2)
```

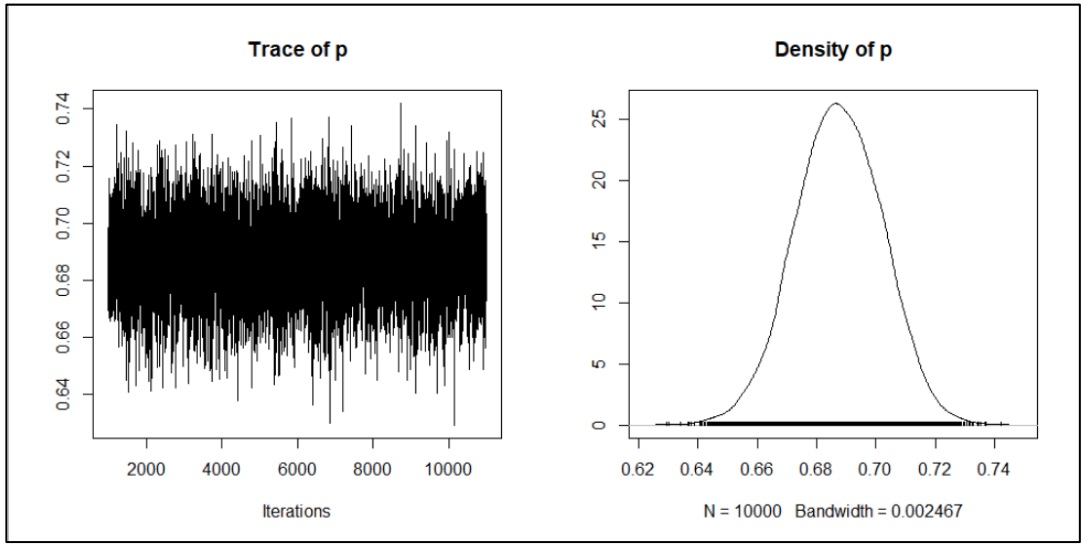
Iterations = 1001:11000
Thinning interval = 1
Number of chains = 1
Sample size per chain = 10000

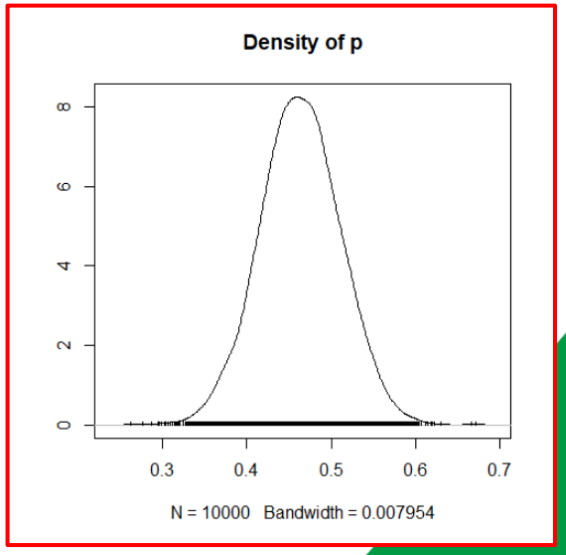
1. Empirical mean and standard deviation for each variable,
   plus standard error of the mean:

Mean	SD	Naive SE	Time-series SE
0.6874392	0.0146836	0.0001468	0.0001935

2. Quantiles for each variable:

2.5%	25%	50%	75%	97.5%
0.6581	0.6775	0.6874	0.6976	0.7155





DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UNIVERSITY OF NORTH DAKOTA

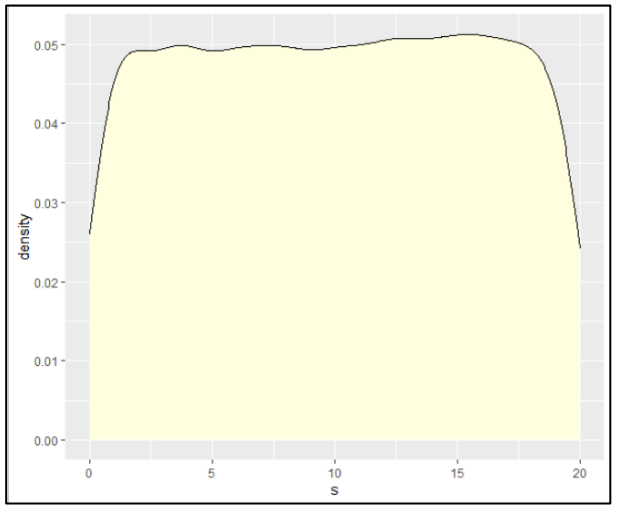
Step-by-step Example 1.2

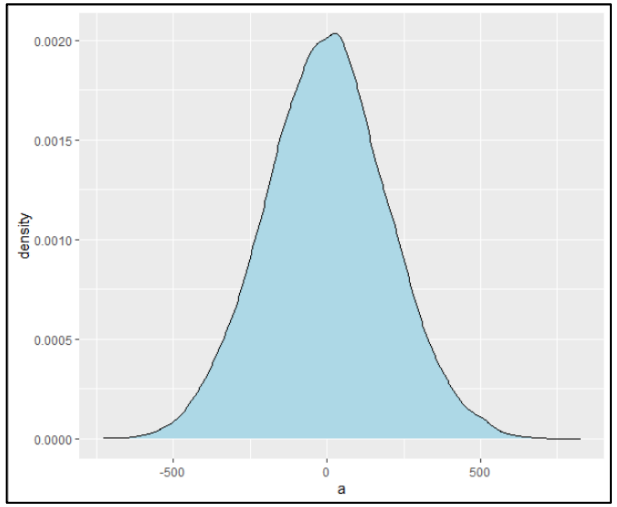
Simple Linear Regression: Can weight be modeled as a function of height?

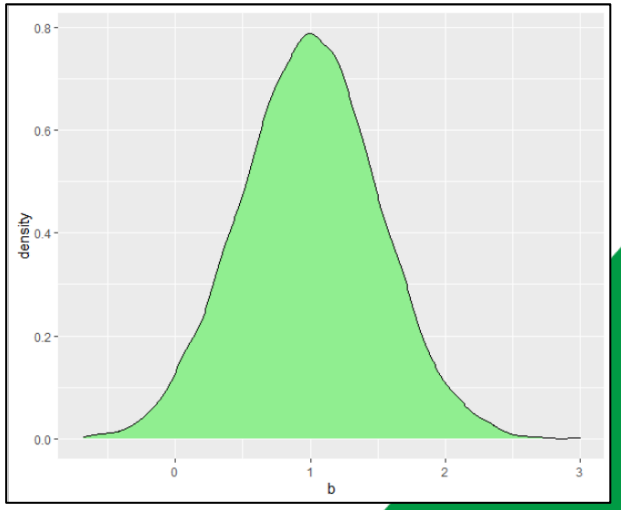
```
#VIEW priors: a (intercept), b (slope), & s (standard deviation)
a <- rnorm(n=10000, mean=0, sd=200)
b <- rnorm(n=10000, mean=1, sd=0.5)
s <- runif(n=10000, min=0, max=20)
samples <- data.frame('a'=a, 'b'=b, 's'=s)
head(samples)

ggplot(samples, aes(x=a)) +
  geom_density(fill="lightblue")
ggplot(samples, aes(x=b)) +
  geom_density(fill="lightgreen")
ggplot(samples, aes(x=s)) +
  geom_density(fill="lightyellow")
```

```
          a         b        s
1 -318.84317 1.2387111 13.981697
2   99.24168 1.1995112 19.119811
3 -284.48808 0.4500849  1.099107
4 -353.41394 1.2572241 17.689841
5  227.24244 1.1815163  5.972512
6   80.44609 0.5957011  6.890548
```







DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UNIVERSITY OF NORTH DAKOTA

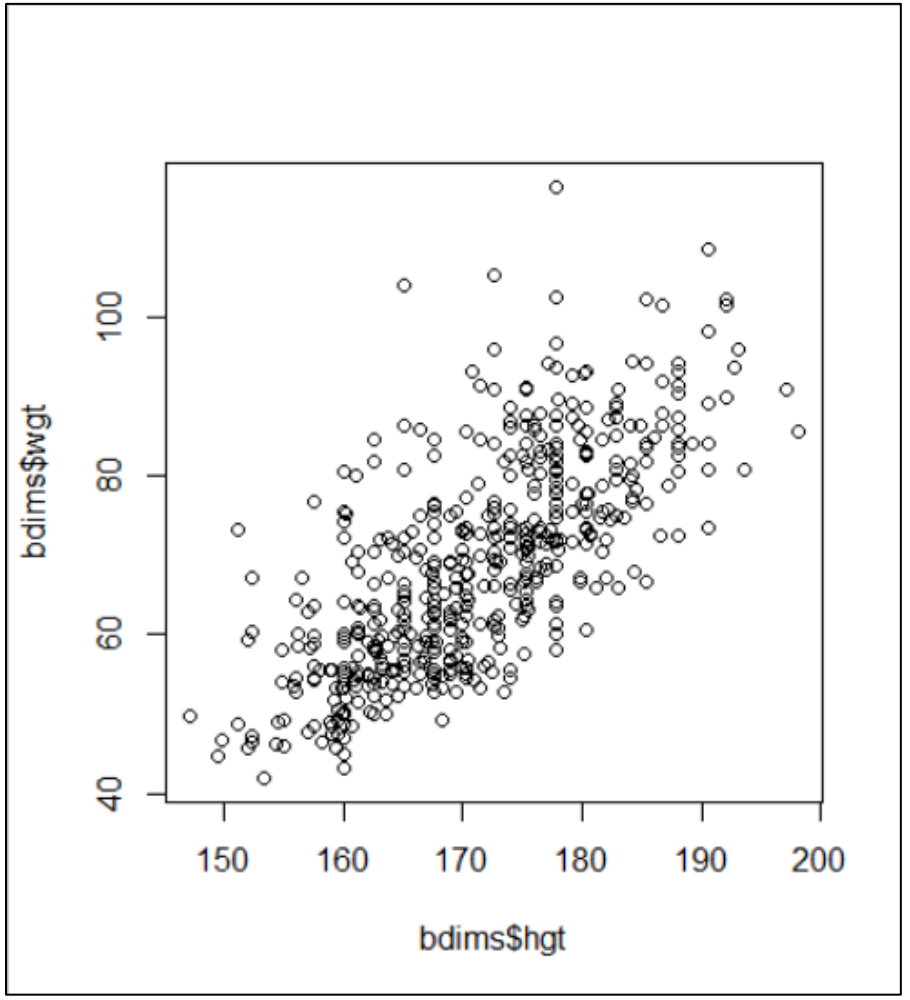
Step-by-step Example 1.2

Simple Linear Regression cont.

```
#VIEW data
data(bdims)
par(mfrow=c(1,1))
plot(x=bdims$hgt, y=bdims$wgt)

 #DEFINE model
weight_model <-"model{
   # Likelihood model for Y[i]
   for(i in 1:length(Y)) {
   Y[i] ~ dnorm(m[i], s^(-2))
   m[i] <- a + b * X[i]
   }

   # Prior models for a, b, s
   a ~ dnorm(0, 200^(-2))
   b ~ dnorm(1, 0.5^(-2))
   s ~ dunif(0, 20)
}"
```

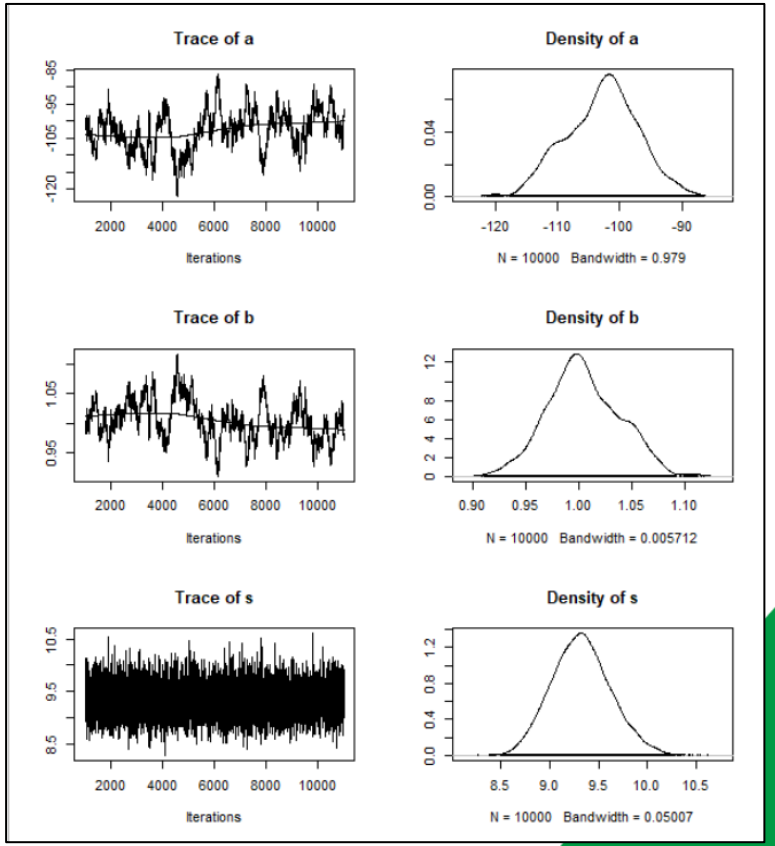


Step-by-step Example 1.2

Simple Linear Regression cont.

```
#COMPILE model
weight_jags <- jags.model(textConnection(weight_model),
                  data  =list(X=bdims$hgt, Y=bdims$wgt),
                  inits =list(.RNG.name="base::Wichmann-Hill",
.RNG.seed=100))
```

```
#SIMULATE posterior
weight_sim <- coda.samples(model=weight_jags,
                  variable.names=c("a", "b", "s"), n.iter=10000)
summary(weight_sim)
plot(weight_sim)
```



Iterations = 1001:11000
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
a	-102.836	5.8274	0.058274	1.147094
b	1.005	0.0340	0.000340	0.006693
s	9.325	0.3013	0.003013	0.003870

2. Quantiles for each variable:

	2.5%	25%	50%	75%	97.5%
a	-114.3834	-106.8077	-102.399	-98.904	-91.583
b	0.9396	0.9821	1.002	1.028	1.072
s	8.7608	9.1185	9.317	9.518	9.940

DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UNIVERSITY OF NORTH DAKOTA

Step-by-step Example 1.2

Simple Linear Regression update

```
#more iterations
weight_sim2 <- coda.samples(model=weight_jags,
                            variable.names=c("a", "b", "s"), n.iter=50000)
summary(weight_sim2)
plot(weight_sim2)
```

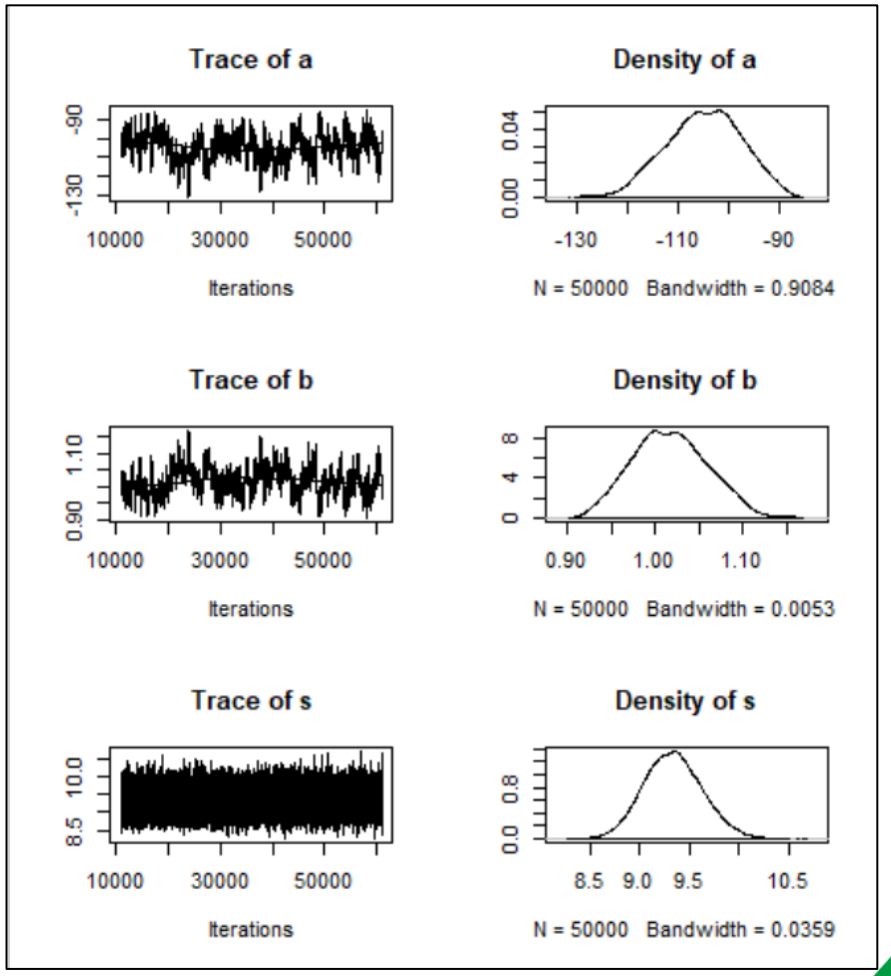
```
Iterations = 11001:61000
Thinning interval = 1
Number of chains = 1
Sample size per chain = 50000

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

     Mean      SD  Naive SE Time-series SE
a -104.969 7.46044 0.0333641      0.846349
b    1.017 0.04352 0.0001946      0.004915
s    9.334 0.29482 0.0013185      0.001718

2. Quantiles for each variable:

      2.5%      25%      50%     75%   97.5%
a -119.7142 -110.0296 -104.684 -99.623 -91.375
b    0.9378    0.9862    1.016   1.047   1.104
s    8.7816    9.1302    9.327   9.529   9.935
```



Step-by-step Example 1.2

Simple Linear Regression update 2

```
  #more chains
weight_jags2 <- jags.model(textConnection(weight_model),
                  data  =list(X=bdims$hgt, Y=bdims$wgt),
                  inits =list(.RNG.name="base::Wichmann-Hill", .RNG.seed=100),
                  n.chains=4)

weight_sim3 <- coda.samples(model=weight_jags2,
                   variable.names=c("a", "b", "s"), n.iter=10000)
summary(weight_sim3)
traceplot(weight_sim3, mfrow = c(2, 2), ask = F)
```

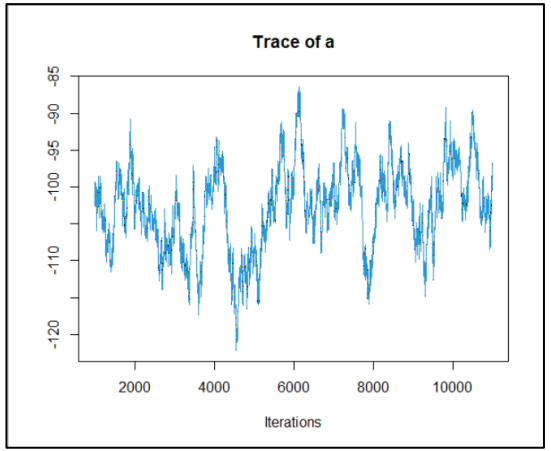
Iterations = 1001:11000
Thinning interval = 1
Number of chains = 4
Sample size per chain = 10000

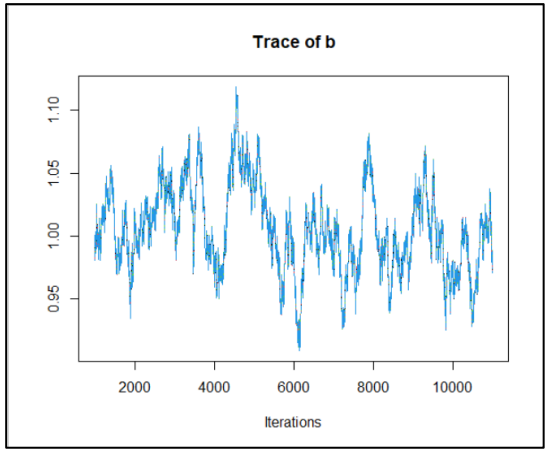
1. Empirical mean and standard deviation for each variable, plus standard error of the mean:

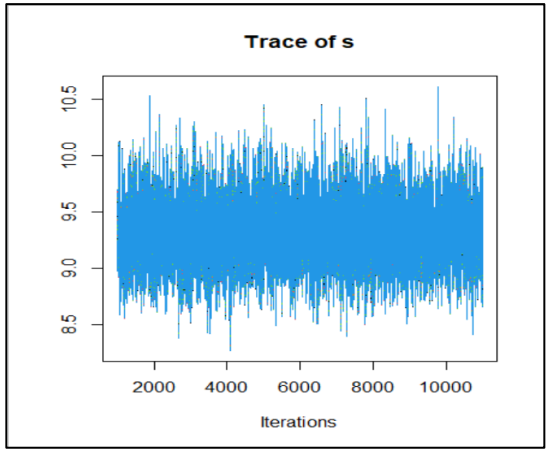
	Mean	SD	Naive SE	Time-series SE
a	-102.836	5.8272	0.029136	0.573547
b	1.005	0.0340	0.000170	0.003346
s	9.325	0.3012	0.001506	0.001935

2. Quantiles for each variable:

	2.5%	25%	50%	75%	97.5%
a	-114.3834	-106.8077	-102.399	-98.904	-91.583
b	0.9396	0.9821	1.002	1.028	1.072
s	8.7608	9.1185	9.317	9.518	9.940







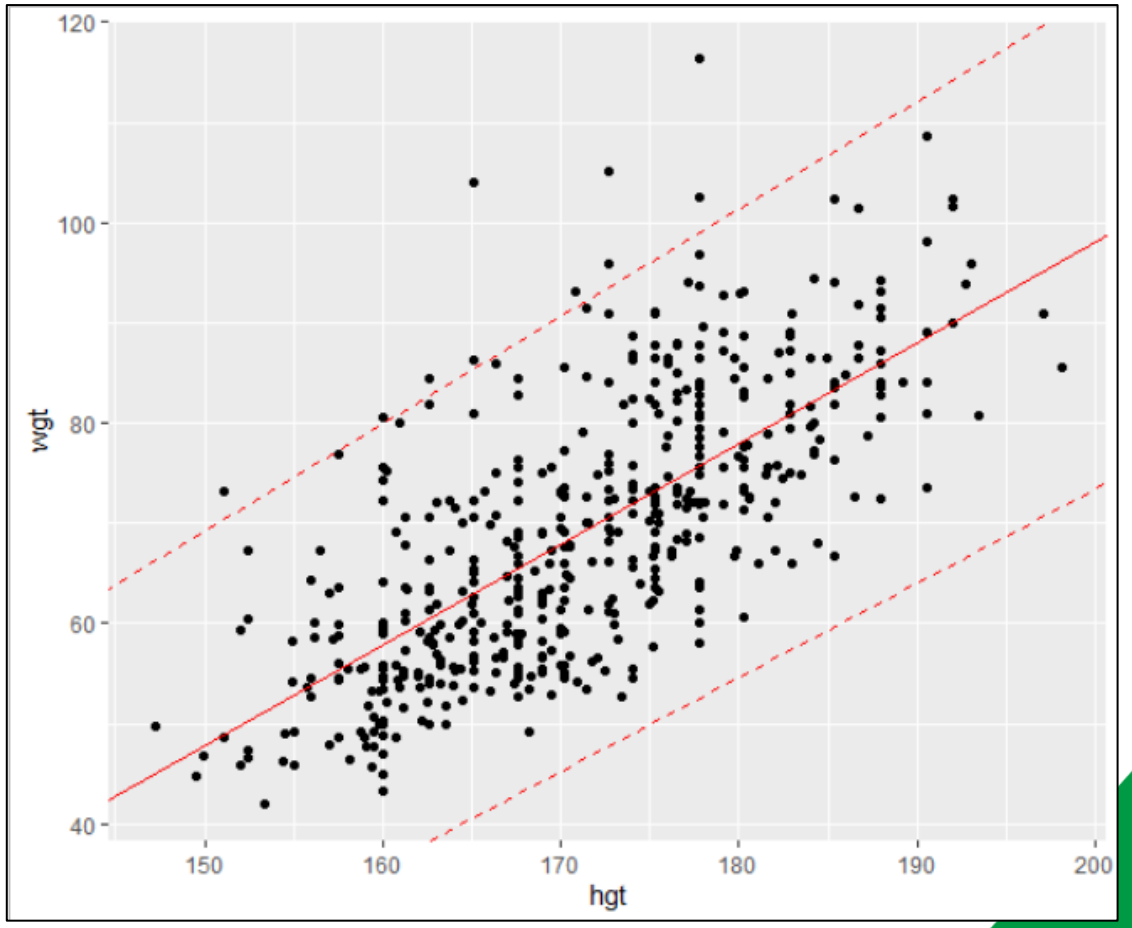
DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UND
UNIVERSITY OF NORTH DAKOTA.

Step-by-step Example 1.2

Simple Linear Regression cont.

```
#PLOT results
ggplot(data=bdims, aes(x=hgt, y=wgt))+
  geom_point() +
  geom_abline(intercept=-102.836, slope=1.005, color='red',) +
  geom_abline(intercept=-91.583,  slope=1.072, color='red', lty=2) +
  geom_abline(intercept=-114.3834,slope=0.9396, color='red', lty=2)
```



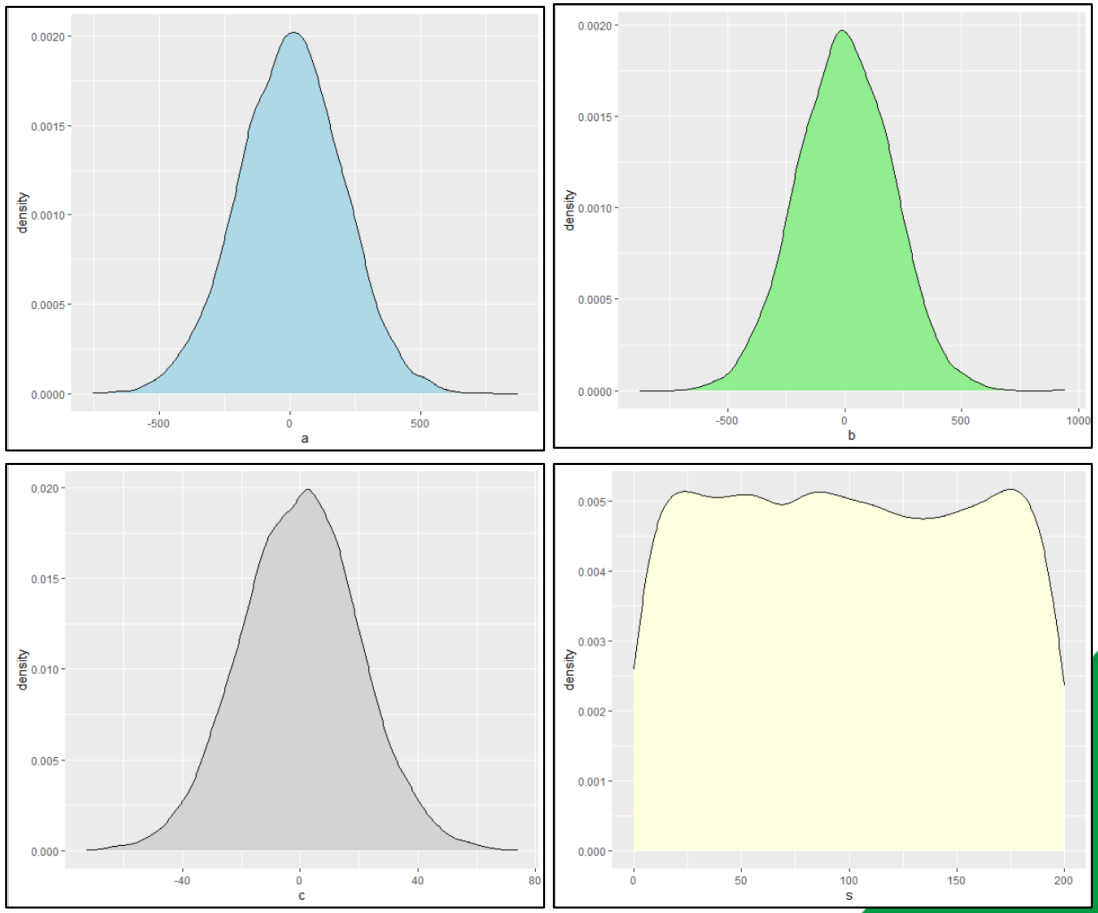
Step-by-step Example 1.3

Multiple Regression: Can we model volume by temperature and weekday?

```
#VIEW priors: a (weekend intercept), b (contrast b/t weekday/weekend
intercepts), c (common slope) & s (residual standard deviation)
a2 <- rnorm(n=10000, mean=0, sd=200)
b2 <- rnorm(n=10000, mean=0, sd=200)
c2 <- rnorm(n=10000, mean=0, sd=20)
s2 <- runif(n=10000, min=0, max=200)Ou
samples2 <- data.frame('a'=a2, 'b'=b2, 'c'=c2, 's'=s2)
head(samples)
```

```
ggplot(samples2, aes(x=a)) +
  geom_density(fill="lightblue")
ggplot(samples2, aes(x=b)) +
  geom_density(fill="lightgreen")
ggplot(samples2, aes(x=c)) +
  geom_density(fill="lightgrey")
ggplot(samples2, aes(x=s)) +
  geom_density(fill="lightyellow")
```

	a	b	s
1	-318.84317	1.2387111	13.981697
2	99.24168	1.1995112	19.119811
3	-284.48808	0.4500849	1.099107
4	-353.41394	1.2572241	17.689841
5	227.24244	1.1815163	5.972512
6	80.44609	0.5957011	6.890548



DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UNIVERSITY OF NORTH DAKOTA

Step-by-step Example 1.3

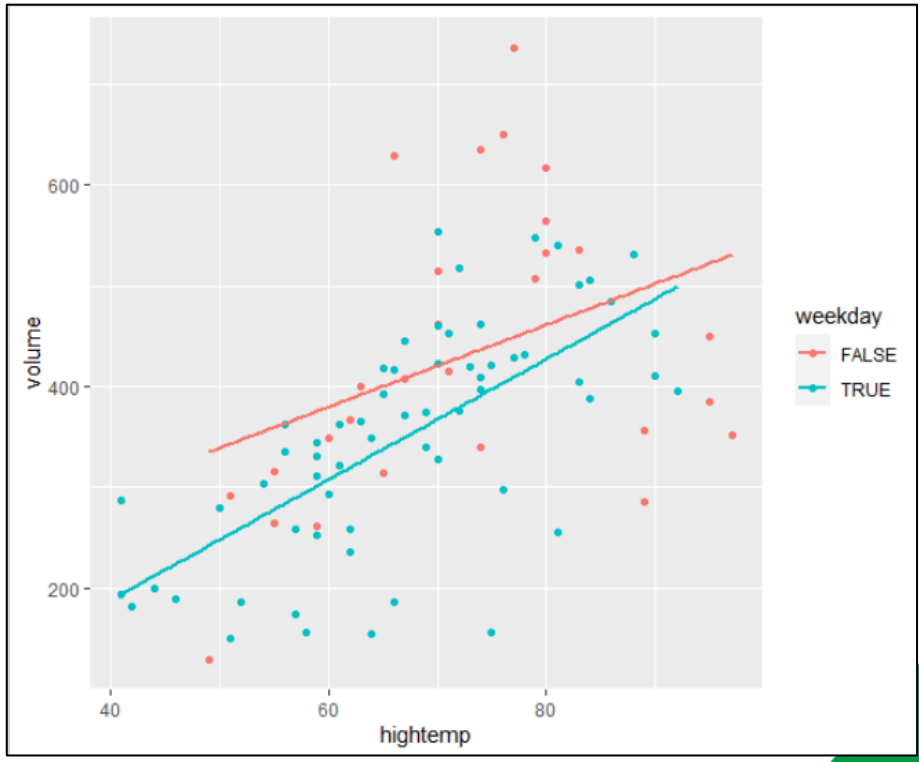
Multiple Regression cont.

```
#VIEW data
library("mosaic")
data(RailTrail)
head(RailTrail)

RailTrail <- RailTrail %>% mutate(weekday=as.factor(weekday))
class(RailTrail$weekday)

ggplot(data=RailTrail, aes(x=hightemp, y=volume, color=weekday))+
  geom_point() +
  geom_smooth(method=lm, se=FALSE, fullrange=FALSE)
```

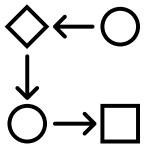
	hightemp	lowtemp	avgtemp	spring	summer	fall	cloudcover	precip	volume	weekday	dayType
1	83	50	66.5	0	1	0	7.6	0.00	501	TRUE	weekday
2	73	49	61.0	0	1	0	6.3	0.29	419	TRUE	weekday
3	74	52	63.0	1	0	0	7.5	0.32	397	TRUE	weekday
4	95	61	78.0	0	1	0	2.6	0.00	385	FALSE	weekend
5	44	52	48.0	1	0	0	10.0	0.14	200	TRUE	weekday
6	69	54	61.5	1	0	0	6.6	0.02	375	TRUE	weekday



DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UNIVERSITY OF NORTH DAKOTA

Step-by-step Example 1.3



Multiple Regression cont.

```
 #DEFINE model
rail_model <- "model{
   #Likelihood model for Y[i]
   for(i in 1:length(Y)) {
     Y[i] ~ dnorm(m[i], s^(-2))
     m[i] <- a + b[X[i]] + c * Z[i]
   }

   # Prior models for a, b, c, s
   a ~ dnorm(0, 200^(-2))
   b[1] <- 0
   b[2] ~ dnorm(0, 200^(-2))
   c ~ dnorm(0, 20^(-2))
   s ~ dunif(0, 200)
}"
```

```
#COMPILE model
rail_jags <- jags.model(textConnection(rail_model),
   data  =list(Y=RailTrail$volume, X=RailTrail$weekday,
Z=RailTrail$hightemp),
   inits =list(.RNG.name="base::Wichmann-Hill",
.RNG.seed=100))

 #SIMULATE posterior
rail_sim <- coda.samples(model=rail_jags,
                         variable.names=c("a", "b", "c", "s"), n.iter=10000)
summary(rail_sim)
plot(rail_sim)
rail_chains <-data.frame(rail_sim[[1]])
```

DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UNIVERSITY OF NORTH DAKOTA

Step-by-step Example 1.3

Multiple Regression cont.

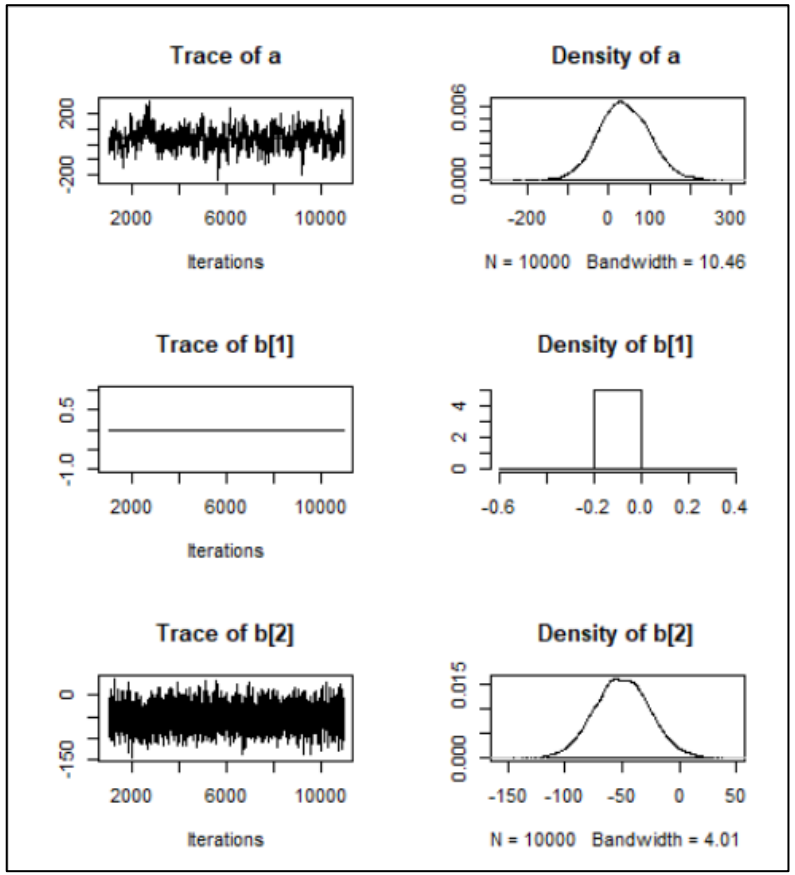
Iterations = 1001:11000
Thinning interval = 1
Number of chains = 1
Sample size per chain = 10000

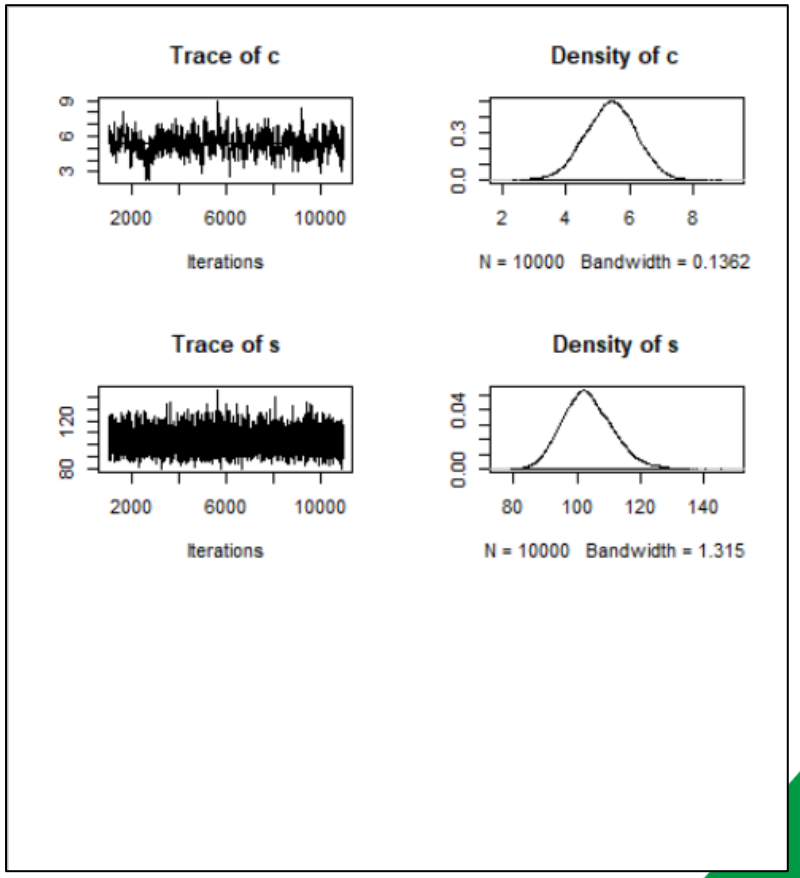
1. Empirical mean and standard deviation for each variable, plus standard error of the mean:

	Mean	SD	Naive SE	Time-series SE
a	37.36	62.2583	0.622583	4.79380
b[1]	0.00	0.0000	0.000000	0.00000
b[2]	-50.07	23.8714	0.238714	0.66501
c	5.41	0.8162	0.008162	0.06194
s	103.42	7.9216	0.079216	0.11666

2. Quantiles for each variable:

	2.5%	25%	50%	75%	97.5%
a	-82.553	-4.292	35.624	79.271	159.659
b[1]	0.000	0.000	0.000	0.000	0.000
b[2]	-96.407	-65.939	-50.294	-33.910	-2.939
c	3.796	4.869	5.422	5.955	6.966
s	89.510	97.935	102.900	108.427	120.719





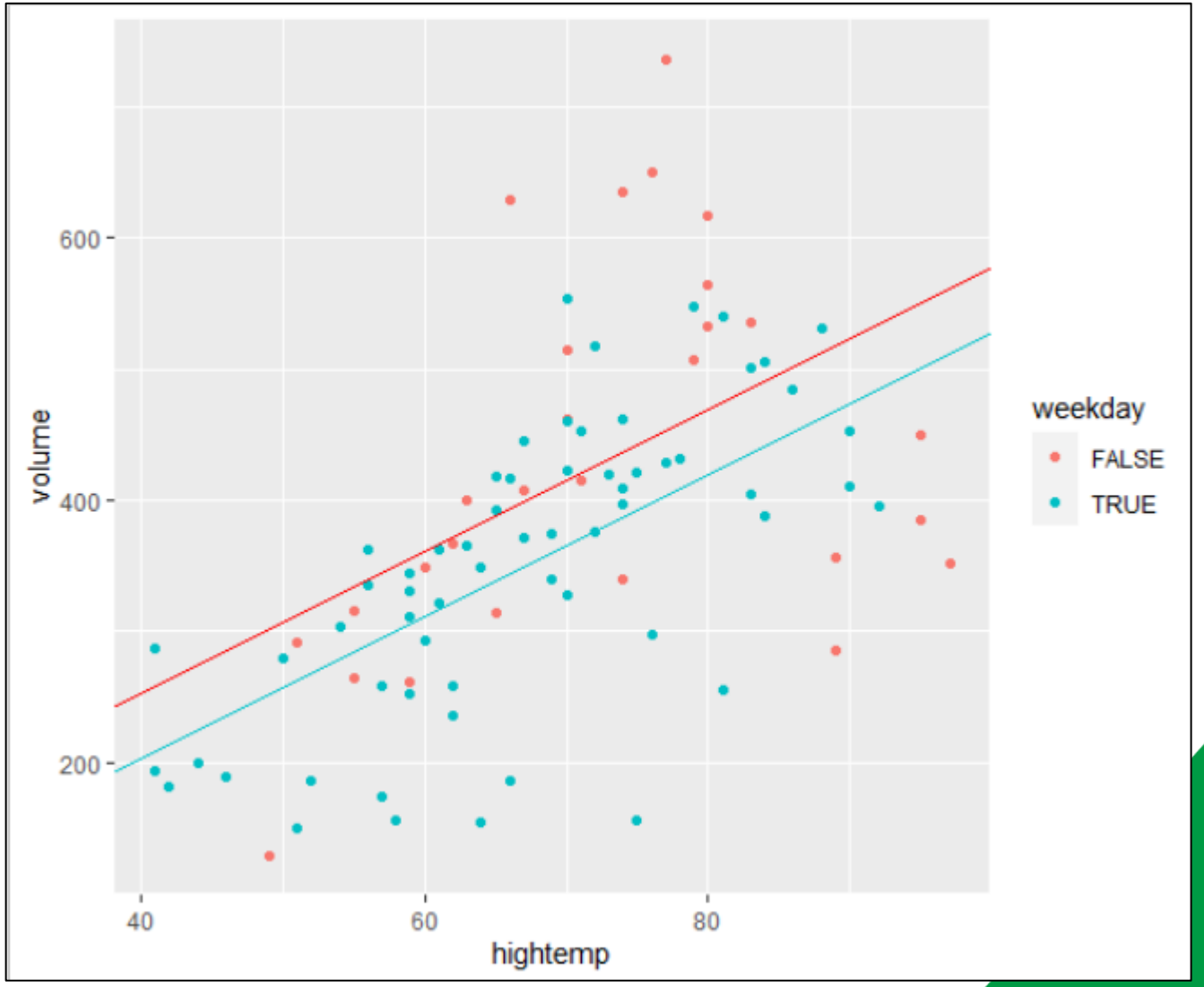
DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UNIVERSITY OF NORTH DAKOTA

Step-by-step Example 1.3

Multiple Regression cont.

```
  #PLOT results
ggplot(RailTrail, aes(y = volume, x = hightemp, color =
weekday)) +
  geom_point() +
  geom_abline(intercept = mean(rail_chains$a),
              slope = mean(rail_chains$c), color = "red") +
  geom_abline(intercept = mean(rail_chains$a) +
mean(rail_chains$b.2.),
              slope = mean(rail_chains$c), color = "turquoise3")
```



DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UNIVERSITY OF NORTH DAKOTA

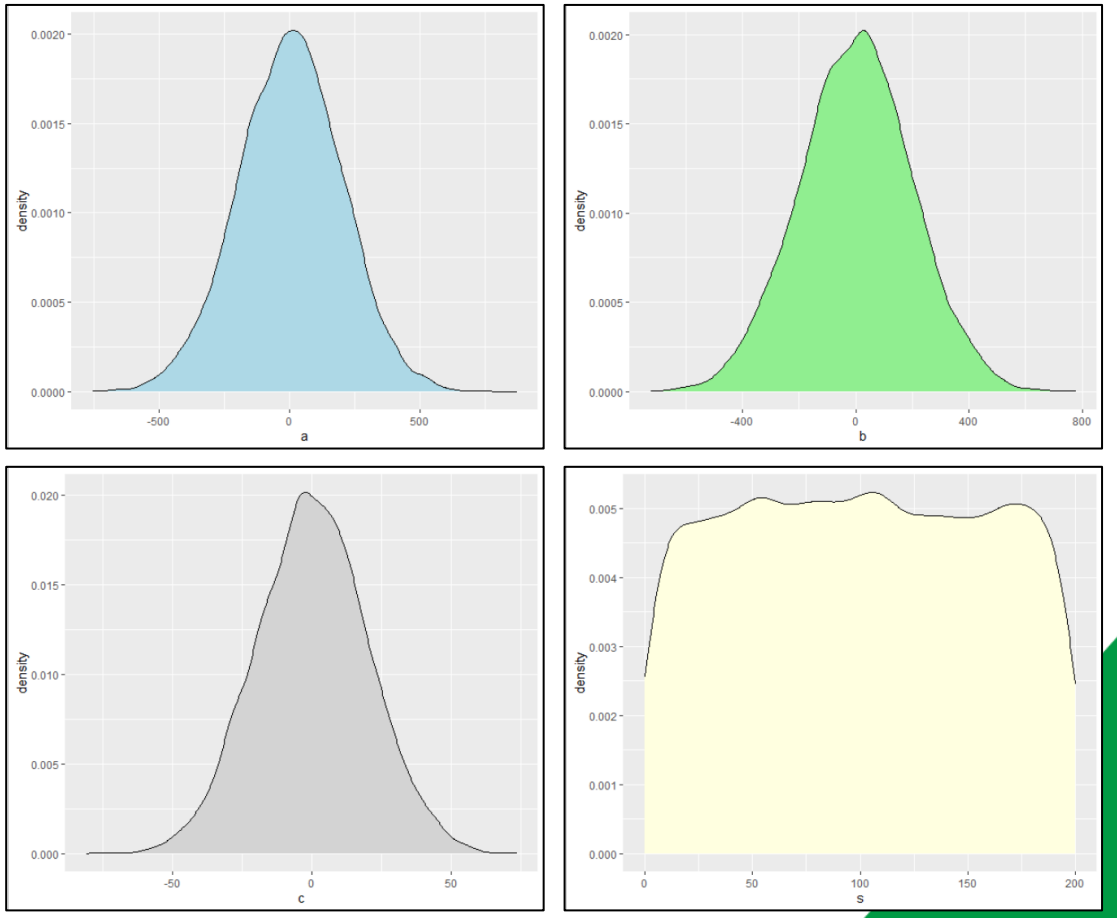
Step-by-step Example 1.4

Poisson Regression: Can we predict volume (as count) by temp. {and weekday}?

```
#VIEW priors: a (weekend intercept), b (contrast b/t weekday/weekend intercepts),
c (common slope) & s (residual standard deviation)
a3 <- rnorm(n=10000, mean=0, sd=200)
b2 <- rnorm(n=10000, mean=0, sd=200)
c2 <- rnorm(n=10000, mean=0, sd=20)
s2 <- runif(n=10000, min=0, max=200)
samples2 <- data.frame('a'=a2, 'b'=b2, 'c'=c2, 's'=s2)
head(samples)

ggplot(samples2, aes(x=a)) +
  geom_density(fill="lightblue")
ggplot(samples2, aes(x=b)) +
  geom_density(fill="lightgreen")
ggplot(samples2, aes(x=c)) +
  geom_density(fill="lightgrey")
ggplot(samples2, aes(x=s)) +
  geom_density(fill="lightyellow")
```

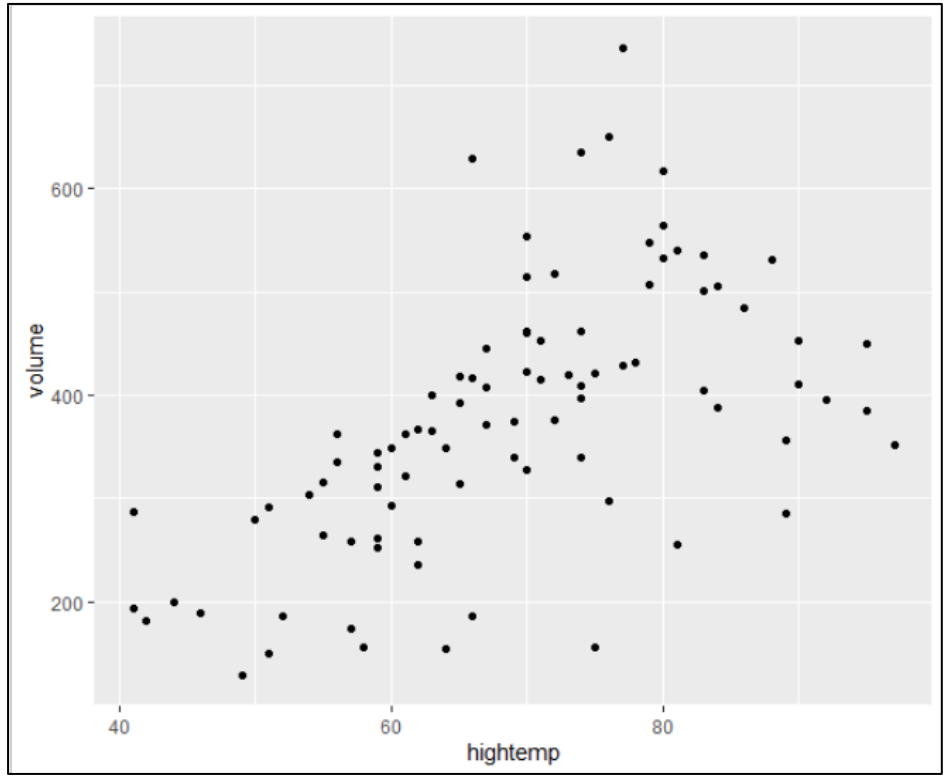
	a	b	s
1	-318.84317	1.2387111	13.981697
2	99.24168	1.1995112	19.119811
3	-284.48808	0.4500849	1.099107
4	-353.41394	1.2572241	17.689841
5	227.24244	1.1815163	5.972512
6	80.44609	0.5957011	6.890548

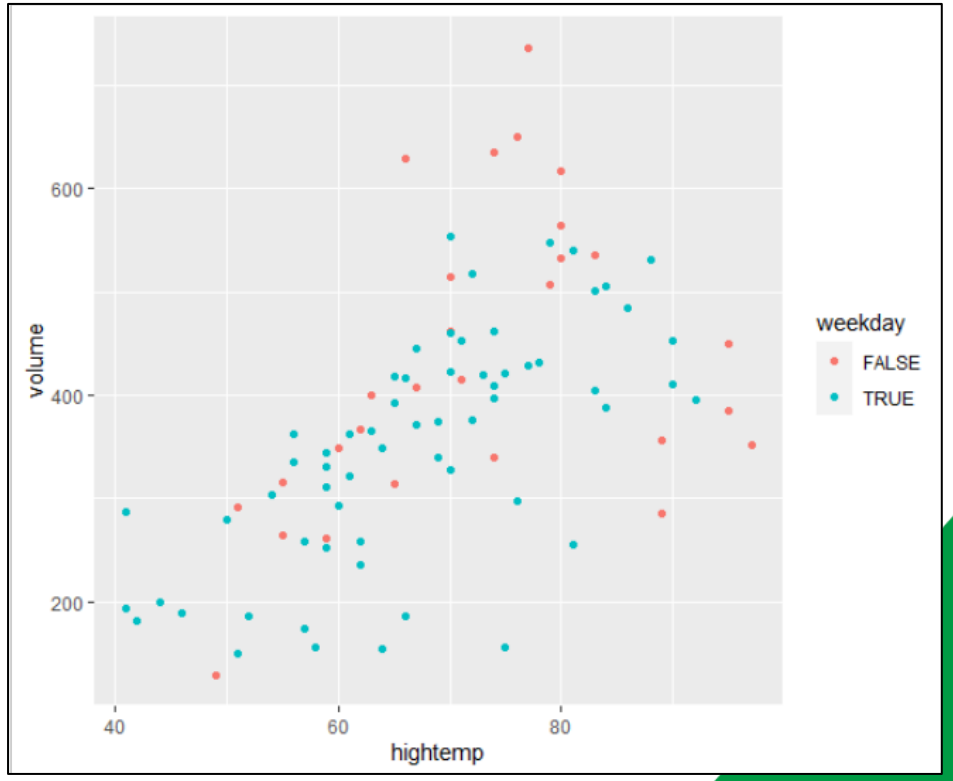


Step-by-step Example 1.4

Poisson Regression cont.

```
#VIEW data
ggplot(data=RailTrail, aes(x=hightemp, y=volume))+
  geom_point()
ggplot(data=RailTrail, aes(x=hightemp, y=volume,
color=weekday))+
  geom_point()
```

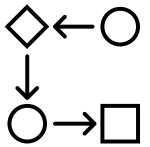




DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UNIVERSITY OF NORTH DAKOTA

Step-by-step Example 1.4



Poisson Regression cont.

```
 #DEFINE model
poisson_model <- "model{
 # Likelihood model for Y[i]
 for(i in 1:length(Y)) {
  Y[i] ~ dpois(l[i])
  log(l[i]) <- a + b * X[i]
 }
 # Prior models for a, b
 a ~ dnorm(0, 200^(-2))
 b ~ dnorm(0, 2^(-2))
}"

 #COMPILE model
poisson_jags <- jags.model(textConnection(poisson_model),
    data = list(Y = RailTrail$volume, X = RailTrail$hightemp),
    inits = list(.RNG.name =  "base::Wichmann-Hill", .RNG.seed = 100))
```

DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UNIVERSITY OF NORTH DAKOTA

Step-by-step Example 1.4

Poisson Regression cont.

```
#SIMULATE posterior
poisson_sim <- coda.samples(model = poisson_jags,
                            variable.names = c("a", "b"), n.iter = 10000)
summary(poisson_sim)
plot(poisson_sim)
poisson_chains <- data.frame(poisson_sim[[1]])
```

```
#PLOT results
ggplot(RailTrail, aes(y = volume, x = hightemp)) +
 geom_point() +
 stat_function(fun = function(x){exp(mean(poisson_chains$a) +
mean(poisson_chains$b) * x)},
              color = "red")
```

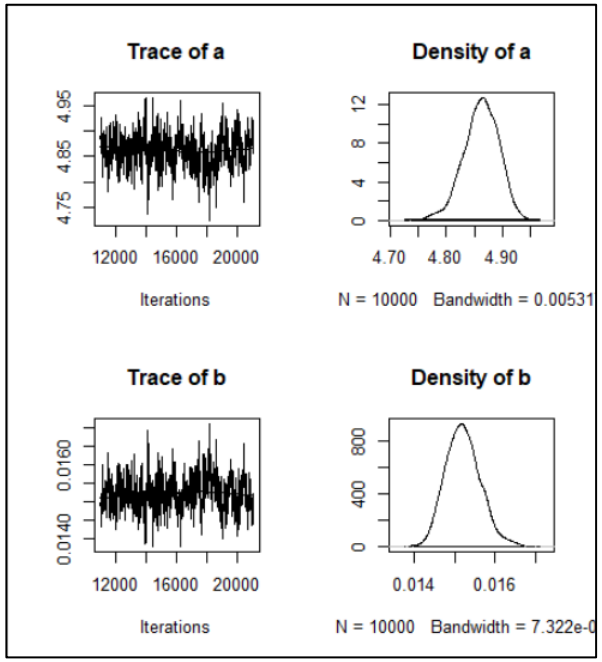
Iterations = 11001:21000
Thinning interval = 1
Number of chains = 1
Sample size per chain = 10000

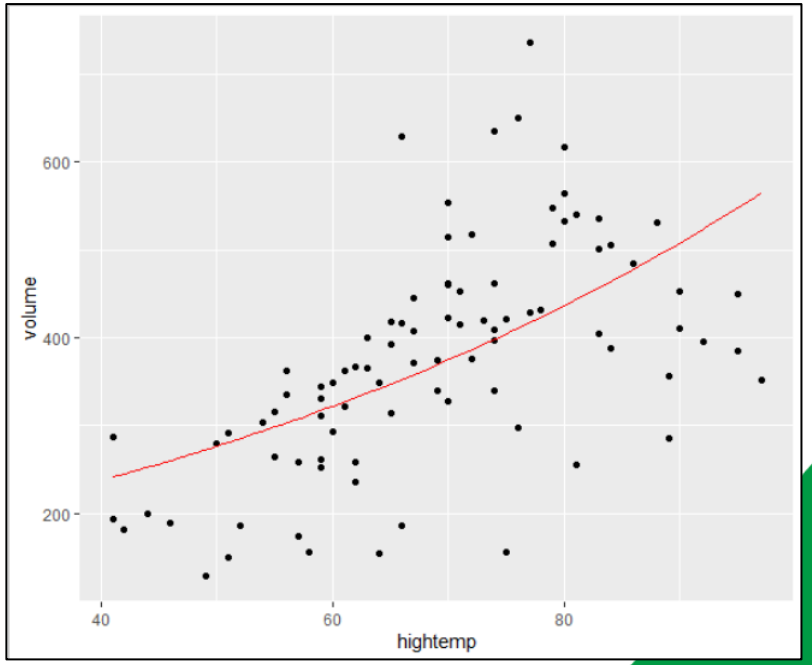
1. Empirical mean and standard deviation for each variable,
   plus standard error of the mean:

	Mean	SD	Naive SE	Time-series SE
a	4.86161	0.0322411	3.224e-04	3.182e-03
b	0.01521	0.0004439	4.439e-06	4.441e-05

2. Quantiles for each variable:

	2.5%	25%	50%	75%	97.5%
a	4.79176	4.8419	4.86327	4.88430	4.91900
b	0.01443	0.0149	0.01519	0.01548	0.01618

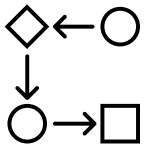




DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UNIVERSITY OF NORTH DAKOTA

Step-by-step Example 1.4



Poisson Regression update

```
#add Weekday

 #DEFINE model
poisson_model2 <- "model{
 # Likelihood model for Y[i]
 for(i in 1:length(Y)) {
  Y[i] ~ dpois(l[i])
  log(l[i]) <- a +b[X[i]] + c*Z[i]
 }
 # Prior models for a, b, c, s
 a ~ dnorm(0, 200^(-2))
 b[1] <- 0
 b[2] ~ dnorm(0, 2^(-2))
 c ~ dnorm(0, 2^(-2))
}"
```

```
#COMPILE model
poisson_jags2 <- jags.model(textConnection(poisson_model2),
   data = list(Y = RailTrail$volume, X = RailTrail$weekday,
               Z = RailTrail$hightemp),
   inits = list(.RNG.name =  "base::Wichmann-Hill", .RNG.seed =
100))

 #PLOT results
ggplot(RailTrail, aes(y = volume, x = hightemp, color =
weekday)) +
 geom_point() +
 stat_function(fun = function(x){exp(mean(poisson_chains$a) +
mean(poisson_chains$c) * x)},
               color = "red") +
 stat_function(fun = function(x){exp(mean(poisson_chains$a) +
mean(poisson_chains$b.2.) + mean(poisson_chains$c) * x)},
               color = "turquoise3")
```

Step-by-step Example 1.4

Poisson Regression update

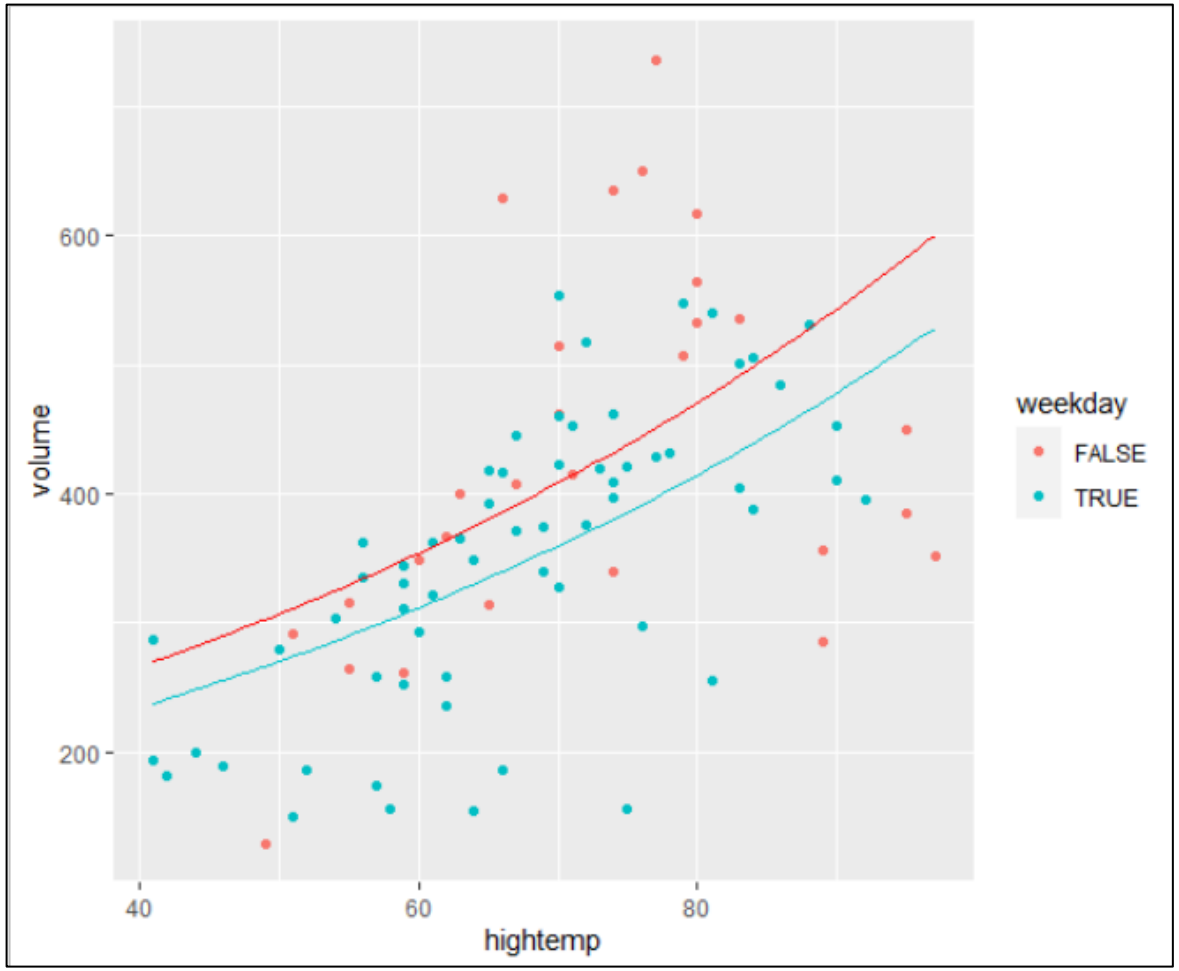
Iterations = 11001:21000
Thinning interval = 1
Number of chains = 1
Sample size per chain = 10000

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

	Mean	SD	Naive SE	Time-series SE
a	5.01207	0.0300023	3.000e-04	2.742e-03
b[1]	0.00000	0.0000000	0.000e+00	0.000e+00
b[2]	-0.12753	0.0114683	1.147e-04	3.368e-04
c	0.01428	0.0003838	3.838e-06	3.392e-05

2. Quantiles for each variable:

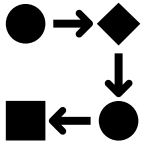
	2.5%	25%	50%	75%	97.5%
a	4.95455	4.99251	5.01149	5.03194	5.07355
b[1]	0.00000	0.00000	0.00000	0.00000	0.00000
b[2]	-0.14963	-0.13544	-0.12748	-0.11971	-0.10489
c	0.01351	0.01402	0.01428	0.01453	0.01501



DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UND
UNIVERSITY OF NORTH DAKOTA.

Step-by-step Example 2



Bayesian Analysis in SAS

Procedure for MCMC in SAS

The following statements can be used with PROC MCMC:

PROC MCMC options ;
ARRAY array specification ;
BEGINCNST/ENDCNST ;
BEGINNODATA/ENDNODATA ;
BY variables ;
MODEL statistical model specification ;
PARMS parameters and starting values ;
PRIOR/HYPERPRIOR prior or hyperprior specification ;
Program statements ;
UDS user defined sampler specification ;

Our steps

1. Get data
2. View priors
3. Run analysis
4. Plot results

DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UND
UNIVERSITY OF NORTH DAKOTA

Step-by-step Example 2.1

Simple Linear Regression: Can we model weight as a function of height? [11]

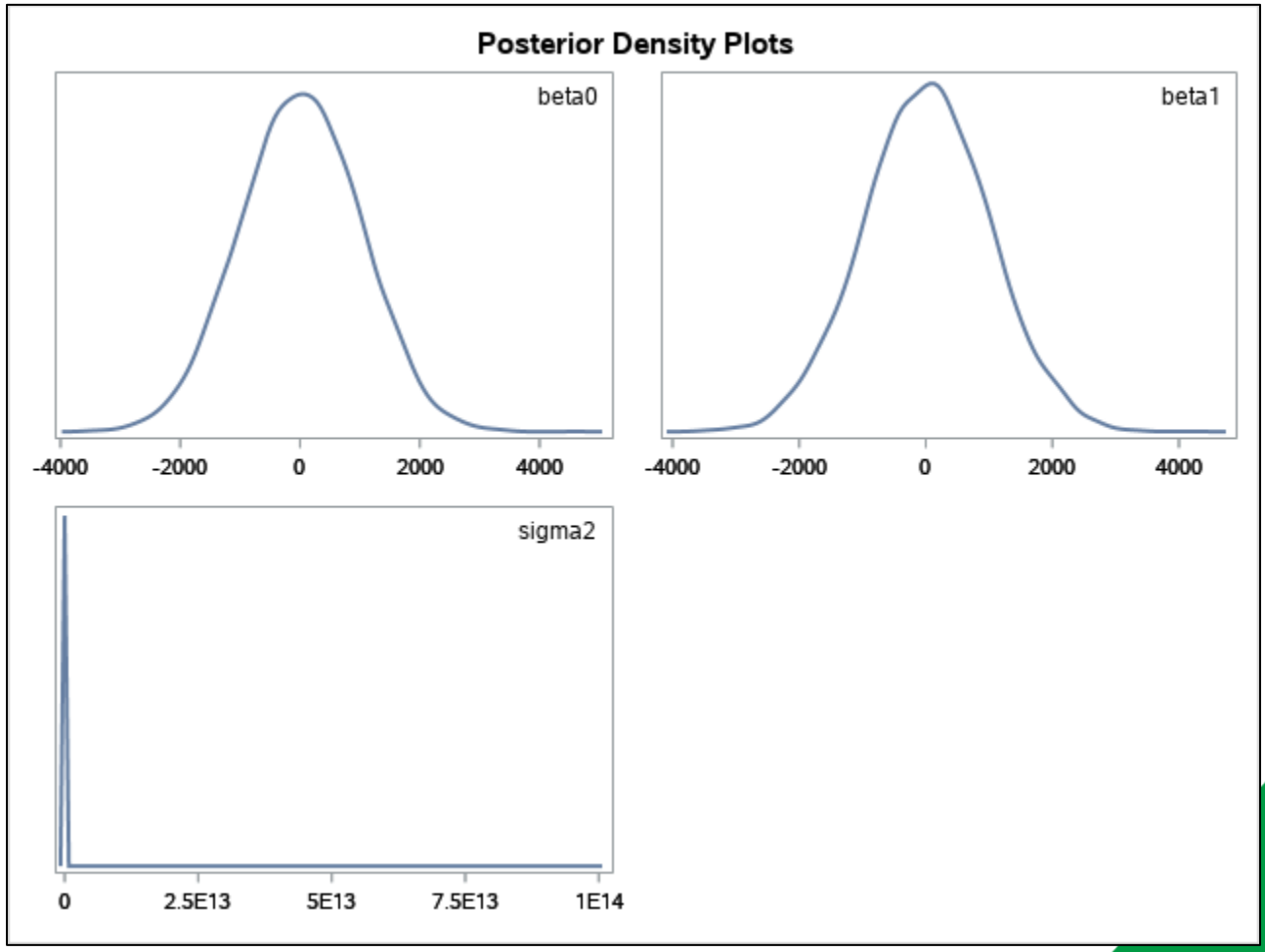
```
*Get data;
DATA Class;
        input Name $ Height Weight @@;
        datalines;
        Alfred  69.0 112.5   Alice  56.5  84.0   Barbara 65.3  98.0
        Carol   62.8 102.5   Henry  63.5 102.5   James   57.3  83.0
        Jane    59.8  84.5   Janet  62.5 112.5   Jeffrey 62.5  84.0
        John    59.0  99.5   Joyce  51.3  50.5   Judy    64.3  90.0
        Louise  56.3  77.0   Mary   66.5 112.0   Philip  72.0 150.0
        Robert  64.8 128.0   Ronald 67.0 133.0   Thomas  57.5  85.0
        William 66.5 112.0
        ;
PROC PRINT data=Class;
DATA a;
```

Obs	Name	Height	Weight
1	Alfred	69.0	112.5
2	Alice	56.5	84.0
3	Barbara	65.3	98.0
4	Carol	62.8	102.5
5	Henry	63.5	102.5
6	James	57.3	83.0
7	Jane	59.8	84.5
8	Janet	62.5	112.5
9	Jeffrey	62.5	84.0
10	John	59.0	99.5
11	Joyce	51.3	50.5
12	Judy	64.3	90.0
13	Louise	56.3	77.0
14	Mary	66.5	112.0
15	Philip	72.0	150.0
16	Robert	64.8	128.0
17	Ronald	67.0	133.0
18	Thomas	57.5	85.0
19	William	66.5	112.0

Step-by-step Example 2.1

Simple Linear Regression cont.

```
*View priors;
PROC MCMC data=a stats=none diag=none nmc=10000
outpost=gout
        plots=density seed=1;
        parms beta0 0 beta1 0;
        parms sigma2 1;
        prior beta0 beta1 ~ normal(mean = 0, var = 1e6);
        prior sigma2 ~ igamma(shape = 3/10, scale = 10/3);
        model general(0);
```

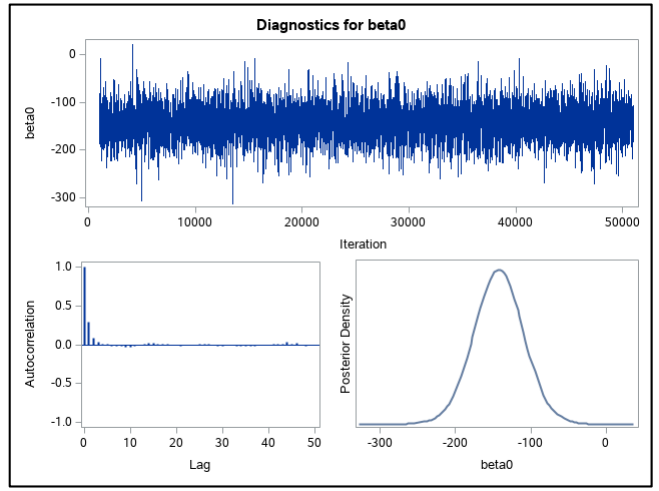


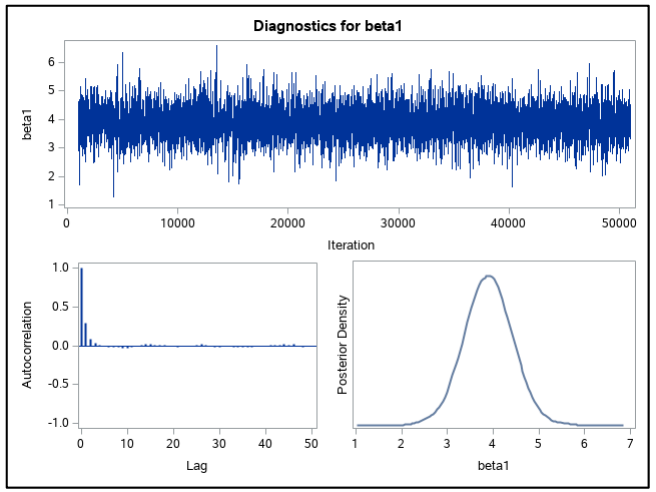
Step-by-step Example 2.1

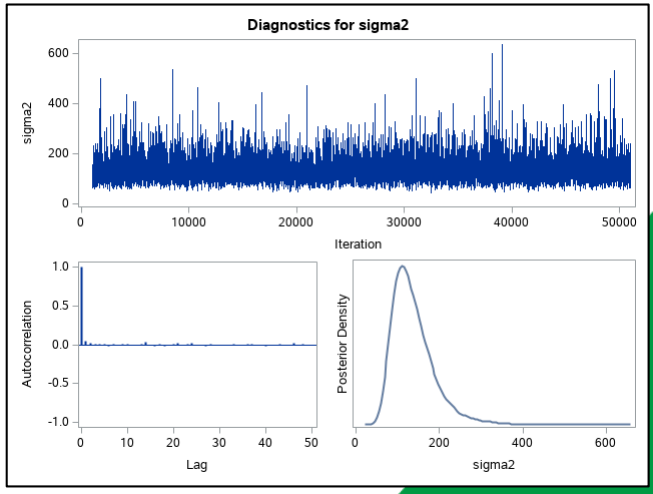
Simple Linear Regression cont.

```
*Run Analysis;
PROC MCMC data=class outpost=classout nmc=50000 thin=5
seed=246810;
        parms beta0 0 beta1 0;
        parms sigma2 1;
        prior beta0 beta1 ~ normal(mean = 0, var = 1e6);
        prior sigma2 ~ igamma(shape = 3/10, scale = 10/3);
        mu = beta0 + beta1*height;
        model weight ~ normal(mu, var = sigma2);
```

Posterior Summaries and Intervals					
Parameter	N	Mean	Standard Deviation	95% HPD Interval	
beta0	10000	-142.7	33.7973	-210.0	-77.5370
beta1	10000	3.8937	0.5400	2.7957	4.9106
sigma2	10000	137.6	51.2936	59.7617	236.5



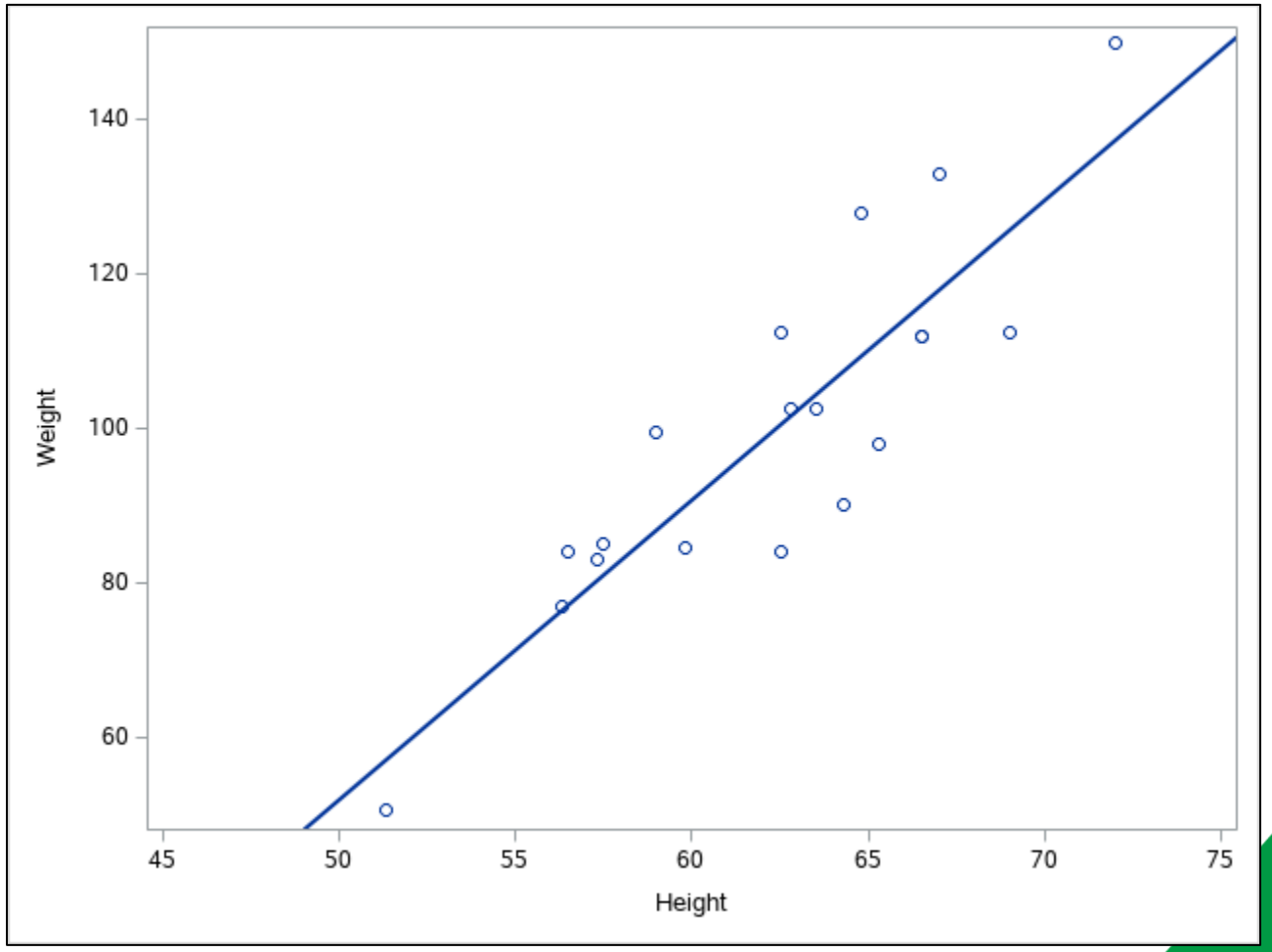




Step-by-step Example 2.1

Simple Linear Regression cont.

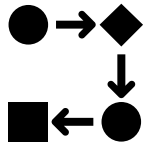
```
*Plot Results;
PROC SGPLOT data=Class noautolegend;
            scatter x=height y=weight;
            lineparm x=0 y=-142.7 slope=3.89/
lineattrs=(thickness=2);
            yaxis min=50 max=150;
            xaxis min=45 max=75;
```



DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UNIVERSITY OF NORTH DAKOTA

Step-by-step Example 2.2



Mixed-Effects Mode: Can we predict height by gender, while blocking for family? [12]

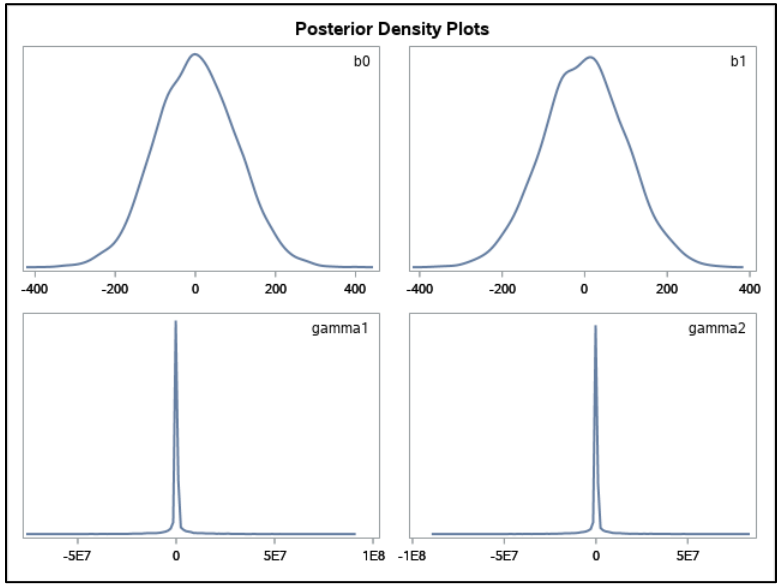
```
*Get data;
DATA heights;
        input family g$ height @@;
        datalines;
        1 F 67   1 F 66   1 F 64   1 M 71   1 M 72   2 F 63
        2 F 63   2 F 67   2 M 69   2 M 68   2 M 70   3 F 63
        3 M 64   4 F 67   4 F 66   4 M 67   4 M 67   4 M 69
        ;
DATA input;
        set heights;
        if g eq 'F' then gender = 1;
        else gender = 0;
        drop g;
PROC PRINT data=input;
```

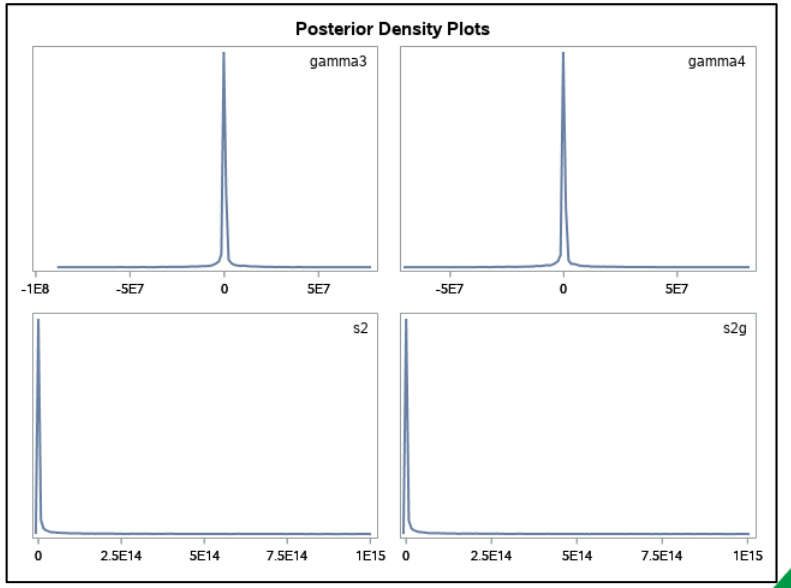
Obs	family	height	gender
1	1	67	1
2	1	66	1
3	1	64	1
4	1	71	0
5	1	72	0
6	2	63	1
7	2	63	1
8	2	67	1
9	2	69	0
10	2	68	0
11	2	70	0
12	3	63	1
13	3	64	0
14	4	67	1
15	4	66	1
16	4	67	0
17	4	67	0
18	4	69	0

Step-by-step Example 2.2

Mixed-Effects Mode cont.

```
*View priors;
PROC MCMC data=a stats=none diag=none nmc=10000
outpost=gout
            plots=density seed=1;
            array gamma[4];
   parms b0 0 b1 0 gamma: 0;
   parms s2 1 ;
   parms s2g 1;
   prior b: ~ normal(0, var = 10000);
   prior gamma: ~ normal(0, var = s2g);
   prior s2: ~ igamma(0.001, scale = 1000);
            model general(0);
```





DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UNIVERSITY OF NORTH DAKOTA

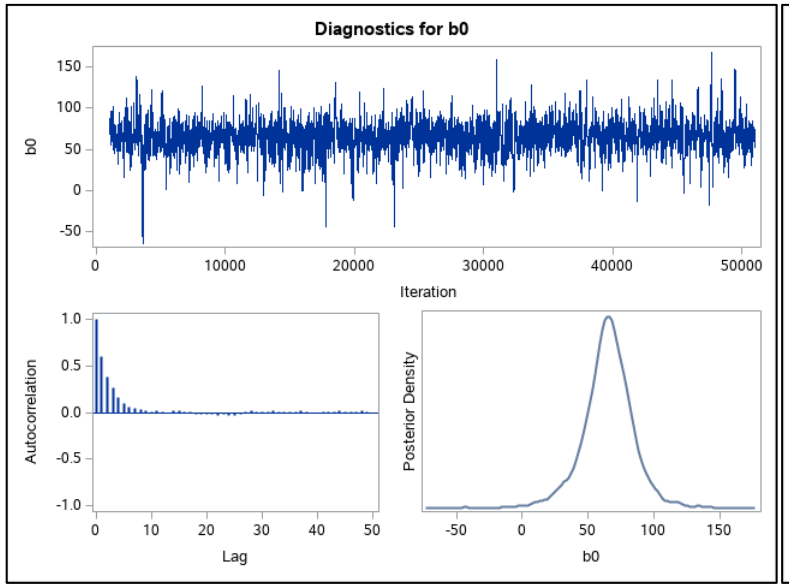
Step-by-step Example 2.2

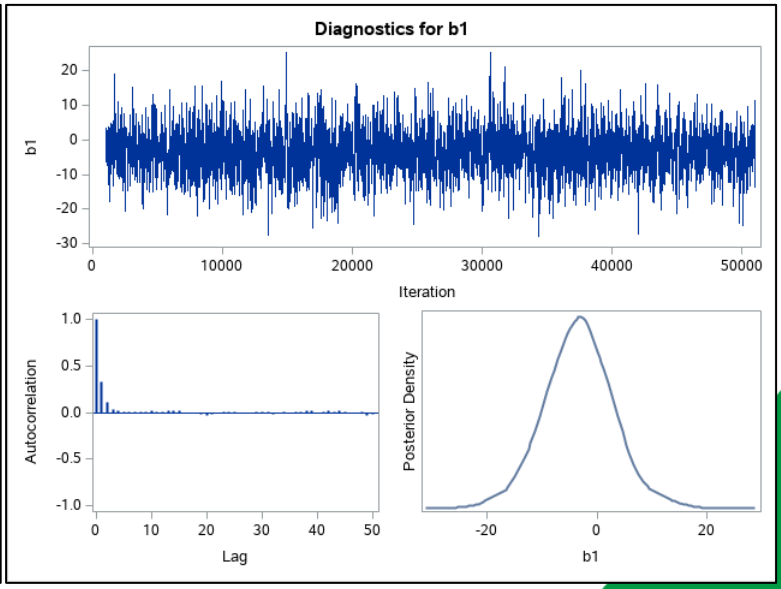
Mixed-Effects Mode cont.

```
*Run Analysis;
PROC MCMC data=input outpost=postout thin=10 nmc=50000
seed=7893 monitor=(b0 b1);
    array gamma[4];
    parms b0 0 b1 0 gamma: 0;
    parms s2 1 ;
    parms s2g 1;
    prior b: ~ normal(0, var = 10000);
    prior gamma: ~ normal(0, var = s2g);
    prior s2: ~ igamma(0.001, scale = 1000);
    mu = b0 + b1 * gender + gamma[family];
    model height ~ normal(mu, var = s2);

*Plot Results;
        *N/A;
```

Posterior Summaries and Intervals					
Parameter	N	Mean	Standard Deviation	95% HPD Interval	
b0	5000	64.9363	19.8441	24.2318	103.9
b1	5000	-3.3589	6.3557	-15.8245	9.8598

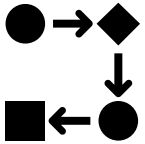




DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UND
UNIVERSITY OF NORTH DAKOTA

Step-by-step Example 2.3



Logistic Regression: can beetle death (y vs. n) be predicted by contaminant (x)? [13]

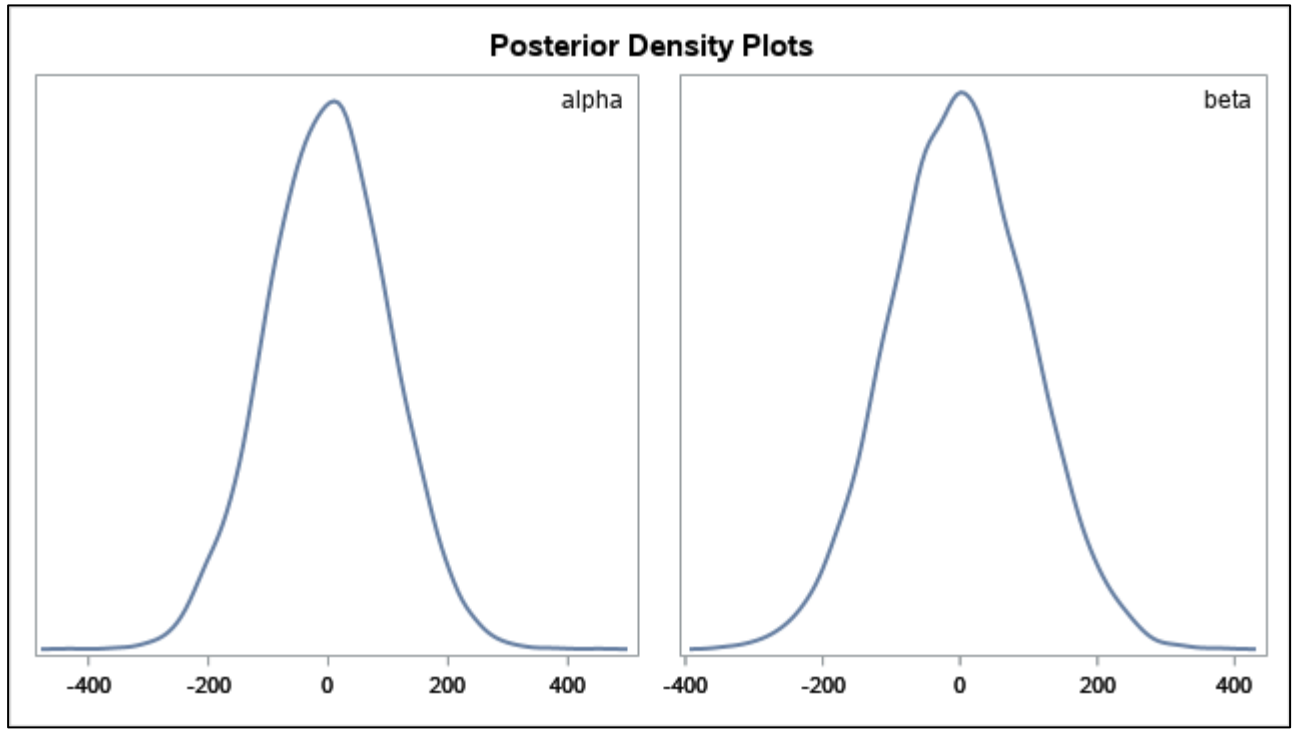
```
*Get data;
DATA beetles;
        input n y x @@;
        datalines;
        6  0  25.7   8  2  35.9   5  2  32.9   7  7  50.4   6  0  28.3
        7  2  32.3   5  1  33.2   8  3  40.9   6  0  36.5   6  1  36.5
        6  6  49.6   6  3  39.8   6  4  43.6   6  1  34.1   7  1  37.4
        8  2  35.2   6  6  51.3   5  3  42.5   7  0  31.3   3  2  40.6
        ;
PROC PRINT data=beetles;
```

Obs	n	y	x
1	6	0	25.7
2	8	2	35.9
3	5	2	32.9
4	7	7	50.4
5	6	0	28.3
6	7	2	32.3
7	5	1	33.2
8	8	3	40.9
9	6	0	36.5
10	6	1	36.5
11	6	6	49.6
12	6	3	39.8
13	6	4	43.6
14	6	1	34.1
15	7	1	37.4
16	8	2	35.2
17	6	6	51.3
18	5	3	42.5
19	7	0	31.3
20	3	2	40.6

Step-by-step Example 2.3

Logistic Regression cont.

```
*View priors;
PROC MCMC data=a stats=none diag=none nmc=10000
outpost=gout
        plots=density seed=1;
        parms (alpha  beta) 0;
        prior alpha beta ~ normal(0, var = 10000);
        model general(0);
```



Step-by-step Example 2.3

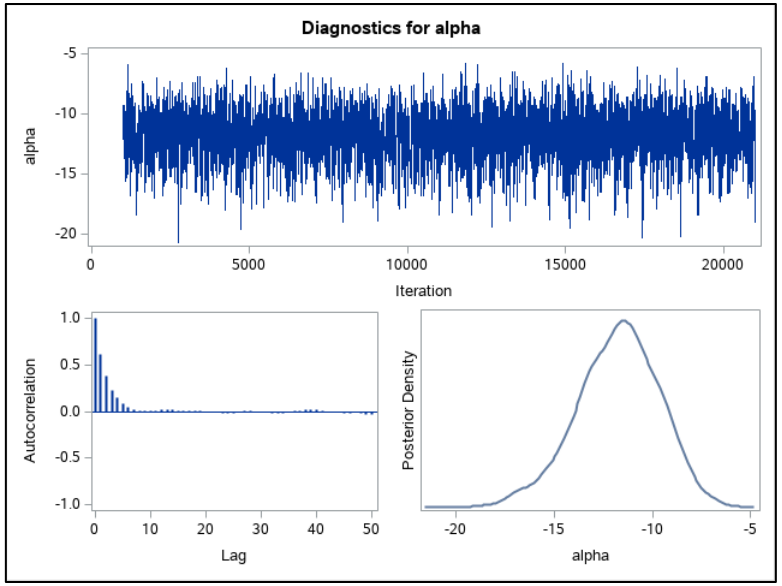
Logistic Regression cont.

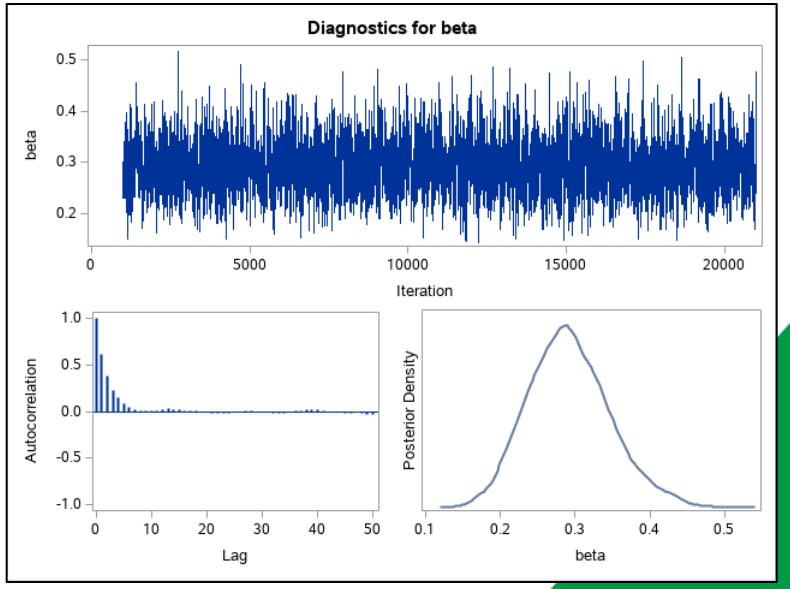
```
*Run Analysis;
PROC MCMC data=beetles ntu=1000 nmc=20000 nthin=2
propcov=quanew
        diag=(mcse ess) outpost=beetleout seed=246810;
            parms (alpha  beta) 0;
            prior alpha beta ~ normal(0, var = 10000);
            p = logistic(alpha + beta*x);
            model y ~ binomial(n,p);

*Plot Results;
            *N/A;
```

Posterior Summaries and Intervals

Parameter	N	Mean	Standard Deviation	95% HPD Interval	
alpha	10000	-11.7707	2.0997	-15.8821	-7.6674
beta	10000	0.2920	0.0542	0.1901	0.4027





DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UND
UNIVERSITY OF NORTH DAKOTA.

Assessment 2

qualtrics XM

UNIVERSITY OF NORTH DAKOTA UND

https://und.qualtrics.com/jfe/form/SV_eDsLcHid4yFnO50

DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UND
UNIVERSITY OF NORTH DAKOTA.

Caveats and Concerns 

- Modeling priors and parameters can get tricky
- Easy does it, start simple and build from there
- Don’t need to go through the Bayesian complexity for simple things
- More than one way to Bayesian a cat (especially in R)
- More parameters that can be set (ex. thinning, seed, tuning loops)
- Not always easy to visualize the results (more references to do so)
- Make sure to work hard to interpret outcomes
- Strong opinions abound the underlying philosophy [17]

DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UND UNIVERSITY OF NORTH DAKOTA.

Real World Examples

[14]

Yapi RB, Chammartin F, Hürlimann E, et al. Bayesian risk profiling of soil-transmitted helminth infections and estimates of preventive chemotherapy for school-aged children in Côte d'Ivoire. *Parasit Vectors*. 2016;9:162. Published 2016 Mar 21. doi:10.1186/s13071-016-1446-0

We modelled soil-transmitted helminth infection risks within a standard Bayesian geostatistical framework and used Markov chain Monte Carlo (MCMC) simulations algorithms to estimate model parameters [29]. In short, the prevalence of each infection at a given location is modelled on the logit scale as a linear function of rigorously chosen explanatory variables and a Gaussian spatial process that accounts for residual spatial correlation not otherwise captured by the covariates.

Table 4 Parameter estimates and validation measures of Bayesian geostatistical models for hookworm, *A. lumbricoides* and *T. trichiura* infection risks in Côte d'Ivoire, late 2011/early 2012

Hookworm	OR (95 % BCI)
Setting	
Rural	1.0
Urban	0.3 (0.2, 0.5)[a]
	Median (95 % BCI)
Variance	1.1 (0.5, 6.3)
Spatial range (km)	217.6 (94.4, 585.1)
	Model validation measures[b]
Mean error (ME) (%)	0.3
Sum standard deviation (SD) (%)	2.2

A. lumbricoides	OR (95 % BCI)
Soil acidity (pH)	
< 5.2	1.0
5.2–5.4	0.3 (0.1, 0.9)[a]
≥ 5.4	0.1 (0.0, 0.4)[a]
Soil moisture	2.4 (1.4, 4.0)[a]
	Median (95 % BCI)
Variance	1.5 (0.6, 3.8)
Spatial range (km)	75.8 (6.3, 429.2)
	Model validation measures[b]
Mean error (ME) (%)	−2.1
Sum standard deviation (SD) (%)	1.0

T. trichiura	OR (95 % BCI)
Rainfall coefficient of variation (CV)	2.0 (1.3, 3.2)[a]
	Median (95 % BCI)
Variance	2.3 (1.0, 5.0)
Spatial range (km)	20.5 (5.9, 80.5)
	Model validation measures[b]
Mean error (ME) (%)	−0.2
Sum standard deviation (SD) (%)	0.6

[a]Significant based on 95 % BCI
[b]Assessed by fitting the model on a subsample of the data (80 %)

DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UNIVERSITY OF NORTH DAKOTA

Real World Examples

[15]

Zhao L, Morgan MA, Parsels LA, Maybaum J, Lawrence TS, Normolle D. Bayesian hierarchical changepoint methods in modeling the tumor growth profiles in xenograft experiments. *Clin Cancer Res*. 2011;17(5):1057-1064. doi:10.1158/1078-0432.CCR-10-1935

A Bayesian hierarchical changepoint (BHC) method was used to model logarithmically transformed tumor volumes. Each tumor was assumed to have a growth profile, represented by a pre-nadir regression rate, a regression period, a nadir volume, and a post-nadir regrowth rate. Confidence intervals were calculated to compare these features between different treatments. We used data from a study assessing the effects of radiation, gemcitabine, and a Chk1/2 inhibitor on MiaPaCa-2 xenografts.

```
model
{
# likelihood
 for ( l in 1:N) {
        y[l] ~ dnorm(mu[l], tau)I(,t.cen[l])
        mu[l] <- a[ID[l],Loc[l]]
             -exp(b[ID[l],1])*(T[l]-(r[ID[l],Loc[l]]))*step((r[ID[l],Loc[l]])-T[l])
             +exp(b[ID[l],2])*(T[l]-(r[ID[l],Loc[l]]))*step(T[l]-(r[ID[l],Loc[l]]))
 }
#priors
  for (i in 1:NP){
       b[i,1:2] ~ dmnorm(mu.b[], OmegaB[,])
        for (k in 1:2){
          a[i,k] ~ dnorm(mu.a,tau.a)
          r[i,k] ~dnorm(mu.r,tau.r)
        }
  }
   mu.a ~ dnorm(0,0.01)
   tau.a ~ dgamma(0.1,0.1)
   mu.r  ~ dnorm(20,0.001)
   tau.r ~ dgamma(0.1,0.1)
   OmegaB[1:2,1:2] ~ dwish(OmegaTau[,],2)
   mu.b[1:2] ~ dmnorm(mub[],tau[,])
   tau ~ dgamma(0.01,0.01)
   mb1 <- exp(mu.b[1])
```

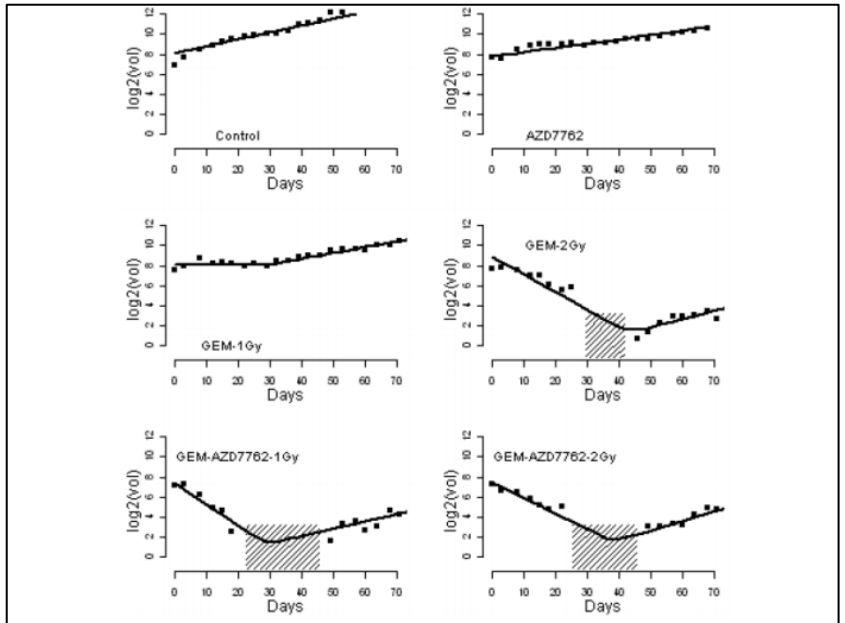


Figure 1. Estimated individual tumor growth profiles for 6 tumors
The observed growth profile (squares) of six selected tumors, with estimated tumor-specific profiles (solid line) from the BHC model; areas shaded with 45 degree lines represent the measurements that are below the limit of quantification. The y axis is the $\log_2$ tumor volume; the x axis is the days after the start of treatment.

DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

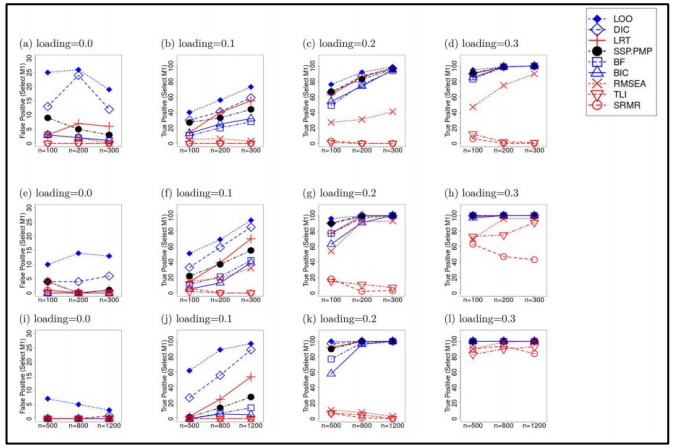
UND UNIVERSITY OF NORTH DAKOTA.

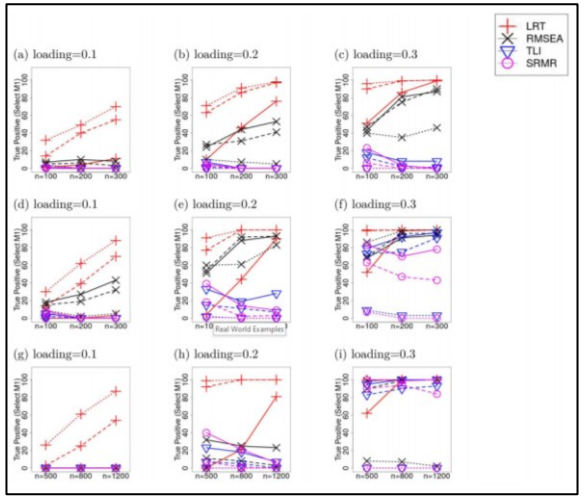
Real World Examples

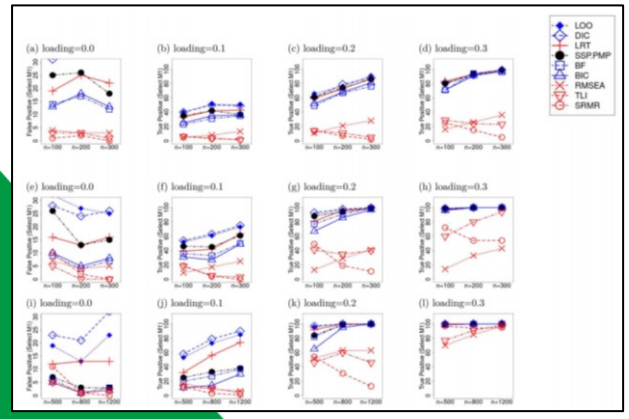
[16]

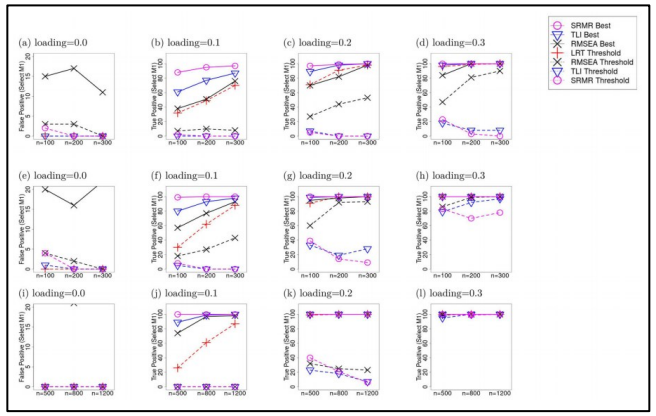
Lu ZH, Chow SM, Loken E. A comparison of Bayesian and frequentist model selection methods for factor analysis models. *Psychol Methods*. 2017;22(2):361-381. doi:10.1037/met0000145

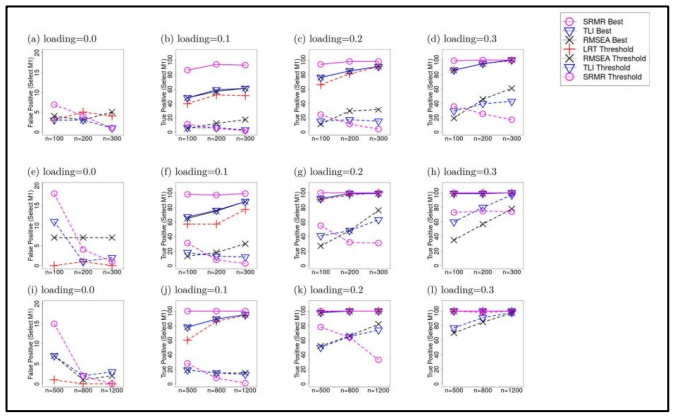
We compare the performances of well-known frequentist model fit indices (MFIs) and several Bayesian model selection criteria (MCC) as tools for cross-loading selection in factor analysis under low to moderate sample sizes, cross-loading sizes, and possible violations of distributional assumptions. The Bayesian criteria considered include the Bayes factor (BF), Bayesian Information Criterion (BIC), Deviance Information Criterion (DIC), a Bayesian leave-one-out approach with Pareto smoothed importance sampling (LOO-PSIS), and a Bayesian variable selection method using the spike-and-slab prior (SSP; Lu, Chow, & Loken, 2016).







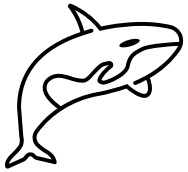




DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UND
UNIVERSITY OF NORTH DAKOTA.

Summary and Conclusion



- Bayesian analysis takes in prior beliefs (hypothesis) and evidence (data) to generate posterior beliefs (probability)
- The detailed method of running analysis is defining priors and parameters to construct a likelihood model
- Standard inference and diagnostic output include trace plots, autocorrelation lag plots, and probability density functions
- Things get complicated quickly, but there are great depths to explore [18-20]

DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

UND UNIVERSITY OF NORTH DAKOTA.

References

[1] https://www.slideshare.net/ASQwebinars/general-bayesian-methods-for-typical-reliability-data-analysis

[2] https://learn.datacamp.com/courses/fundamentals-of-bayesian-data-analysis-in-r

[3] https://lakens.github.io/statistical_inferences/bayes.html

[4] https://www.researchgate.net/publication/331494053_Prognostics_102_Efficient_Bayesian-Based_Prognostics_Algorithm_in_MATLAB/figures?lo=1

[5] https://www.researchgate.net/publication/337863468_Performance_of_Hamiltonian_Monte_Carlo_and_No-U-Turn_Sampler_for_estimating_genetic_parameters_and_breeding_values/figures?lo=1

[6] https://statlect.com/fundamentals-of-statistics/Markov-Chain-Monte-Carlo-diagnostics

[7] https://www.researchgate.net/publication/337806909_Bayesian_Model_Selection_in_Fisheries_Management_and_Ecology/figures?lo=1

[8] https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_mcmc_sect049.htm

[9] https://cran.r-project.org/web/packages/rjags/rjags.pdf

[10] https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#mcmc_toc.htm

[11] https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_mcmc_sect007.htm

[12] https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_mcmc_sect006.htm

[13] https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_mcmc_sect045.htm

[14] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4802658/pdf/13071_2016_Article_1446.pdf

[15] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3048927/pdf/nihms-256616.pdf

[16] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5527973/pdf/nihms867195.pdf

[17] https://www.allendowney.com/blog/2021/04/25/bayesian-and-frequentist-results-are-not-the-same-ever/

[18] https://bookdown.org/kevin_davisross/bayesian-reasoning-and-methods/introduction-to-jags.html

[19] https://www.slideshare.net/FrancescoCasalegno1/markov-chain-monte-carlo-methods-116880736

[20] https://astrostatistics.psu.edu/RLectures/diagnosticsMCMC.pdf

DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

Acknowledgements

UND
UNIVERSITY OF NORTH DAKOTA.

- The DaCCoTA is supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number U54GM128729.
- For the labs that use the Biostatistics, Epidemiology, and Research Design Core in any way, including this Module, please acknowledge us for publications. ***"Research reported in this publication was supported by DaCCoTA (the National Institute of General Medical Sciences of the National Institutes of Health under Award Number U54GM128729)".***

DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY