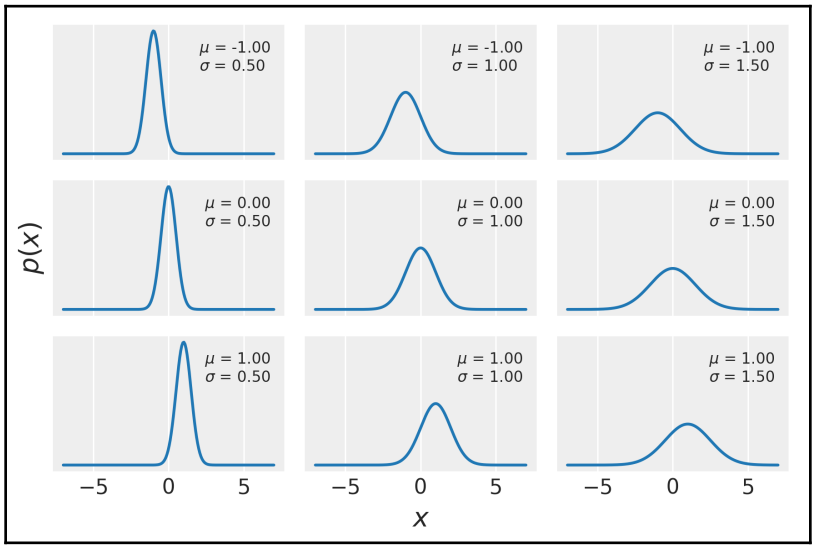
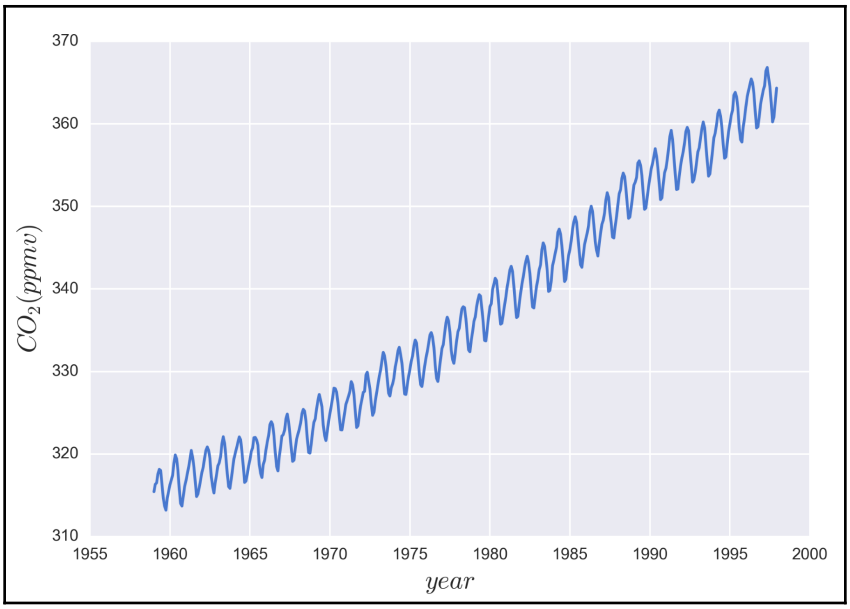
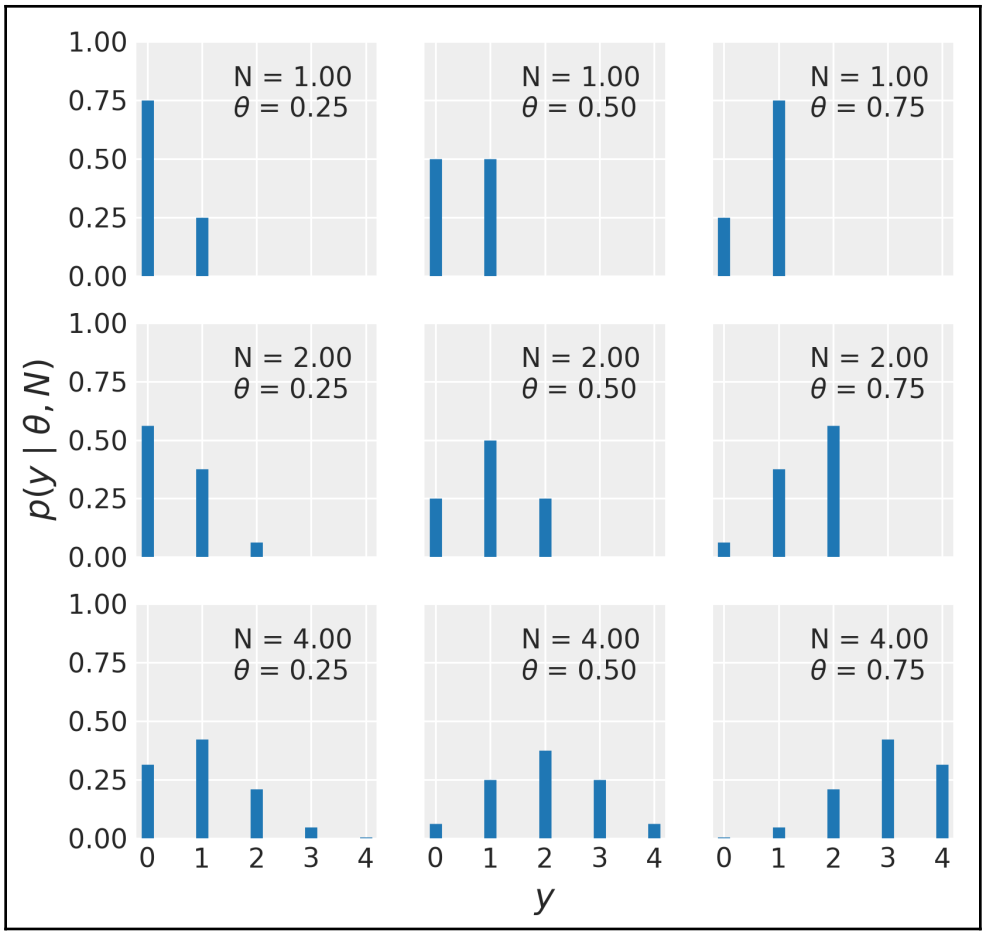
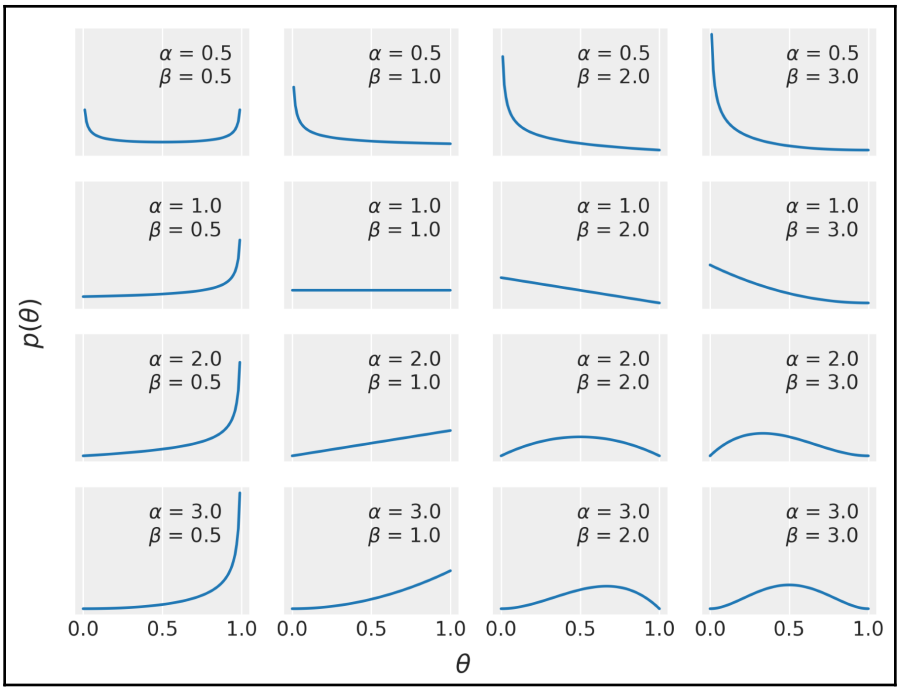
# 1
# Graphics

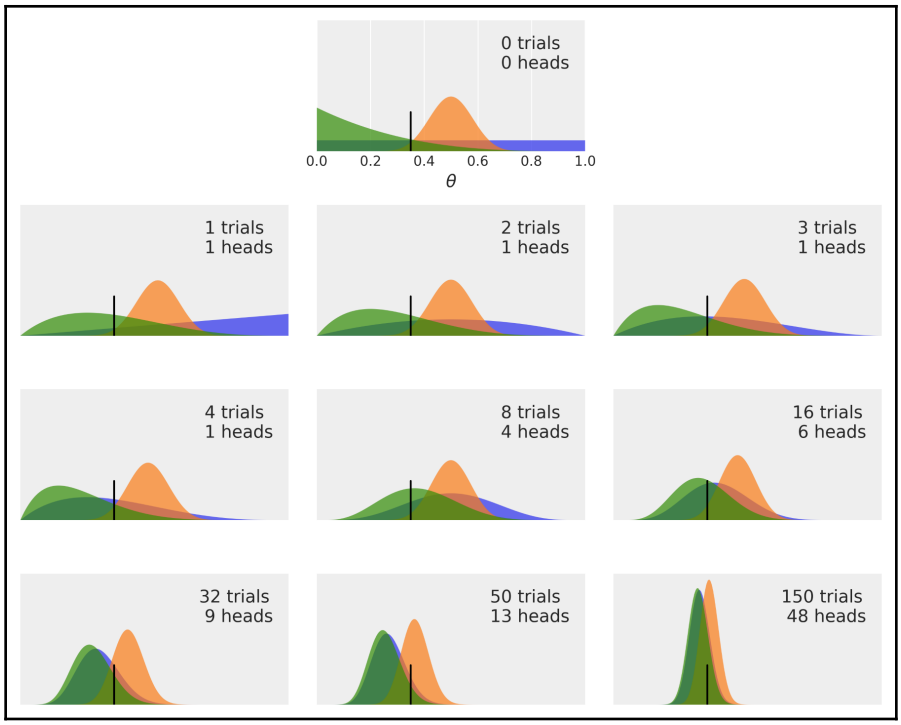
## Chapter 1: Thinking Probabilistically

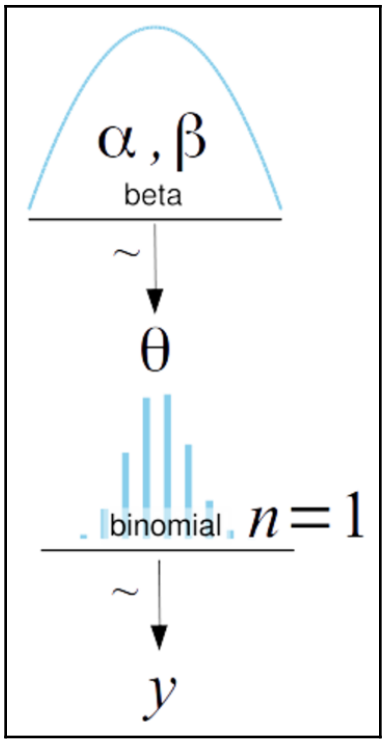


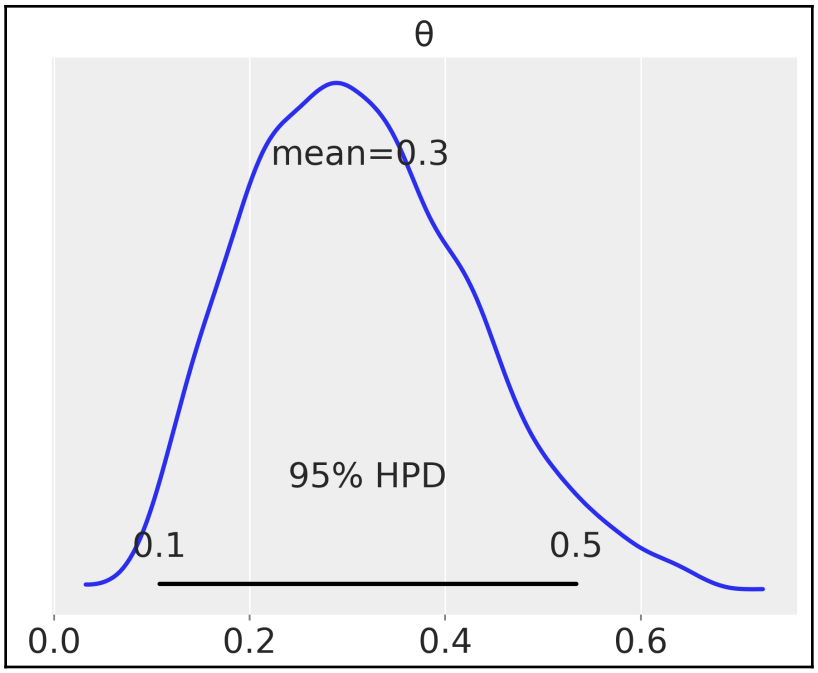

370
360
350
340
330
320
310
$CO_2(ppmv)$
1955
1960
1965
1970
1975
1980
1985
1990
1995
2000
$year$

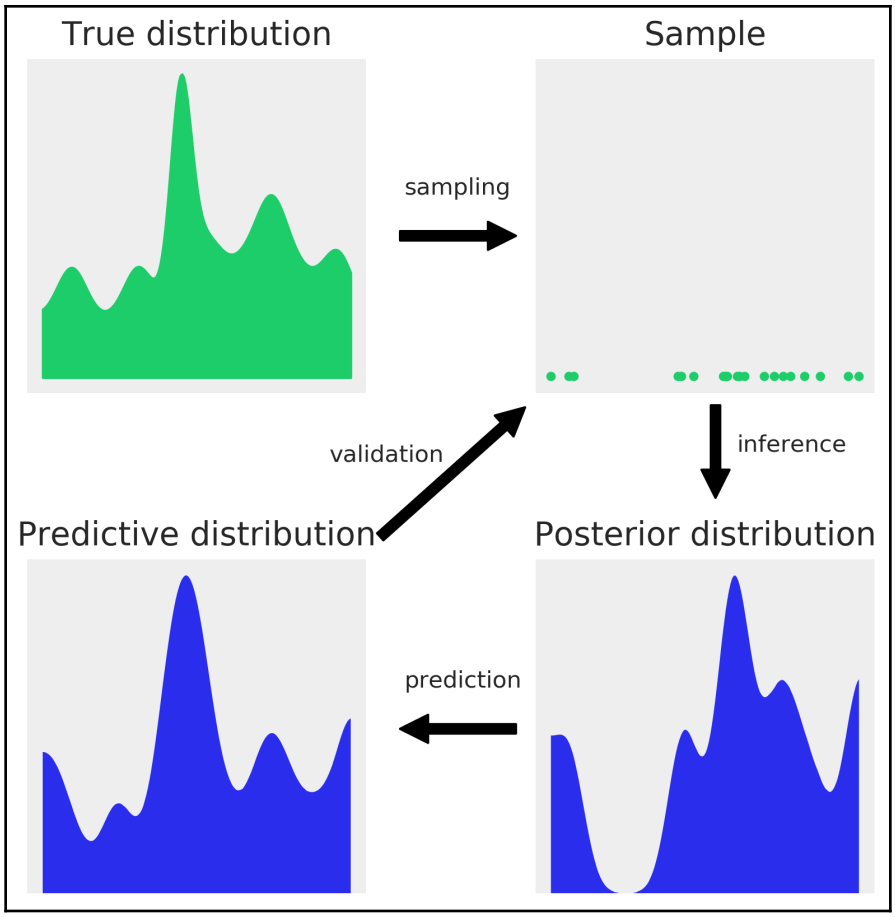

N = 1.00
θ = 0.25
N = 1.00
θ = 0.50
N = 1.00
θ = 0.75
N = 2.00
θ = 0.25
N = 2.00
θ = 0.50
N = 2.00
θ = 0.75
N = 4.00
θ = 0.25
N = 4.00
θ = 0.50
N = 4.00
θ = 0.75
p(y | θ, N)
y
0.00
0.25
0.50
0.75
1.00
0
1
2
3
4


α = 0.5
β = 0.5
α = 0.5
β = 1.0
α = 0.5
β = 2.0
α = 0.5
β = 3.0
α = 1.0
β = 0.5
α = 1.0
β = 1.0
α = 1.0
β = 2.0
α = 1.0
β = 3.0
α = 2.0
β = 0.5
α = 2.0
β = 1.0
α = 2.0
β = 2.0
α = 2.0
β = 3.0
α = 3.0
β = 0.5
α = 3.0
β = 1.0
α = 3.0
β = 2.0
α = 3.0
β = 3.0
p(θ)
0.0
0.5
1.0
θ

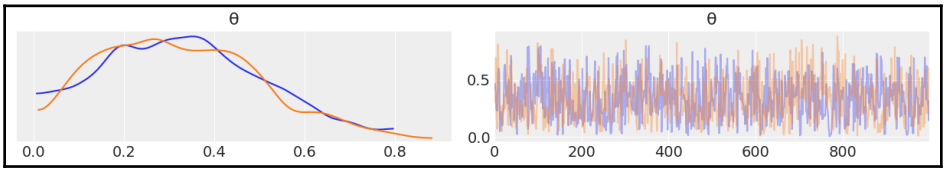

0 trials
0 heads
0.0
0.2
0.4
0.6
0.8
1.0
θ
1 trials
1 heads
2 trials
1 heads
3 trials
1 heads
4 trials
1 heads
8 trials
4 heads
16 trials
6 heads
32 trials
9 heads
50 trials
13 heads
150 trials
48 heads

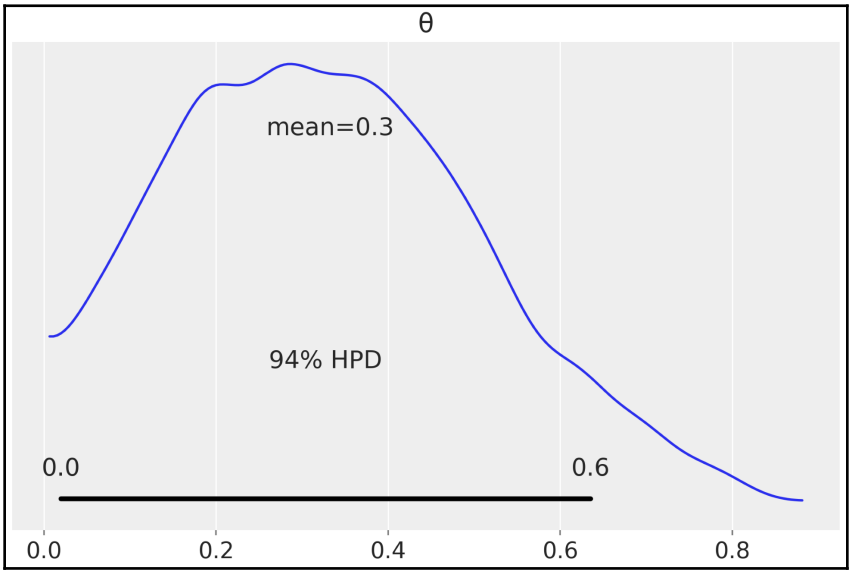

α , β
beta
~
θ
binomial
n = 1
~
y

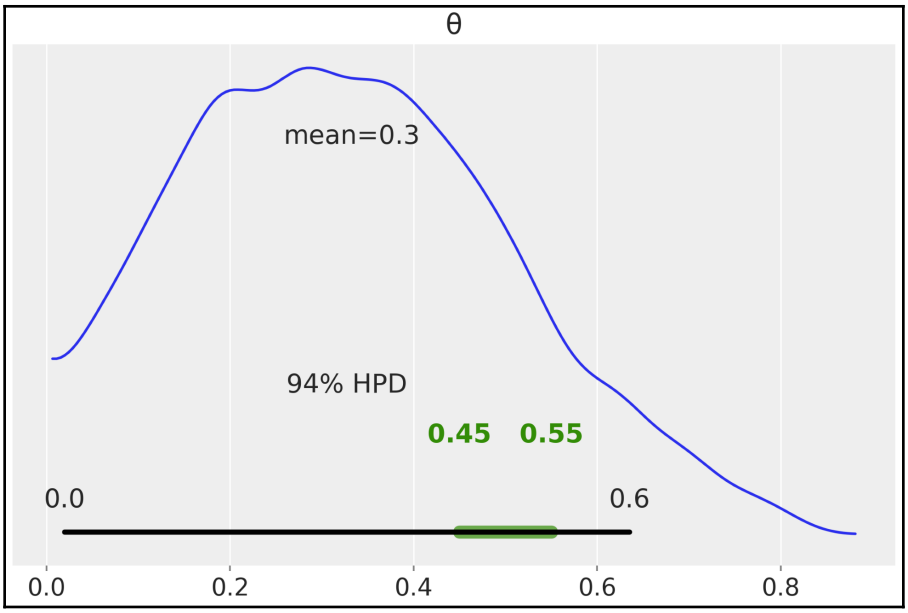

θ
mean=0.3
95% HPD
0.1
0.5
0.0
0.2
0.4
0.6

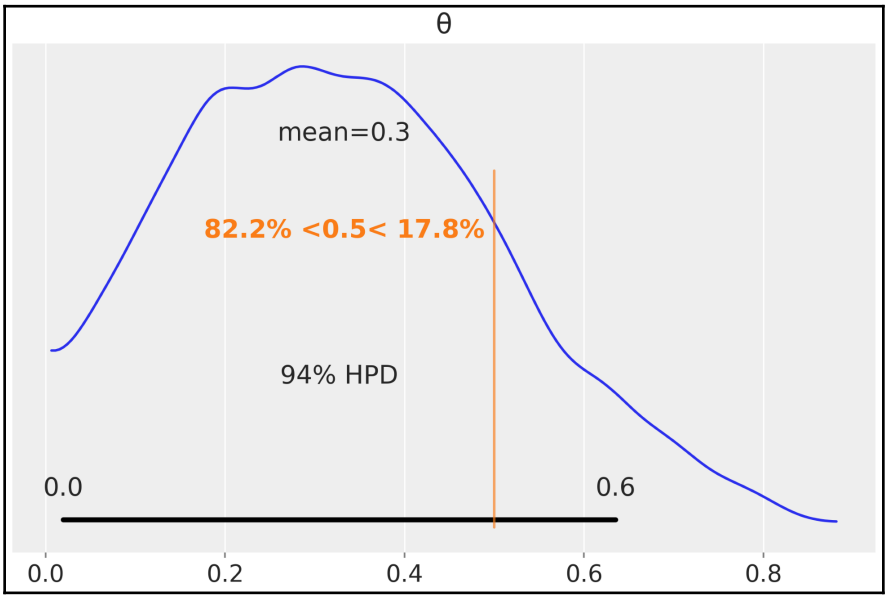

True distribution
Sample
sampling
validation
inference
Predictive distribution
Posterior distribution
prediction

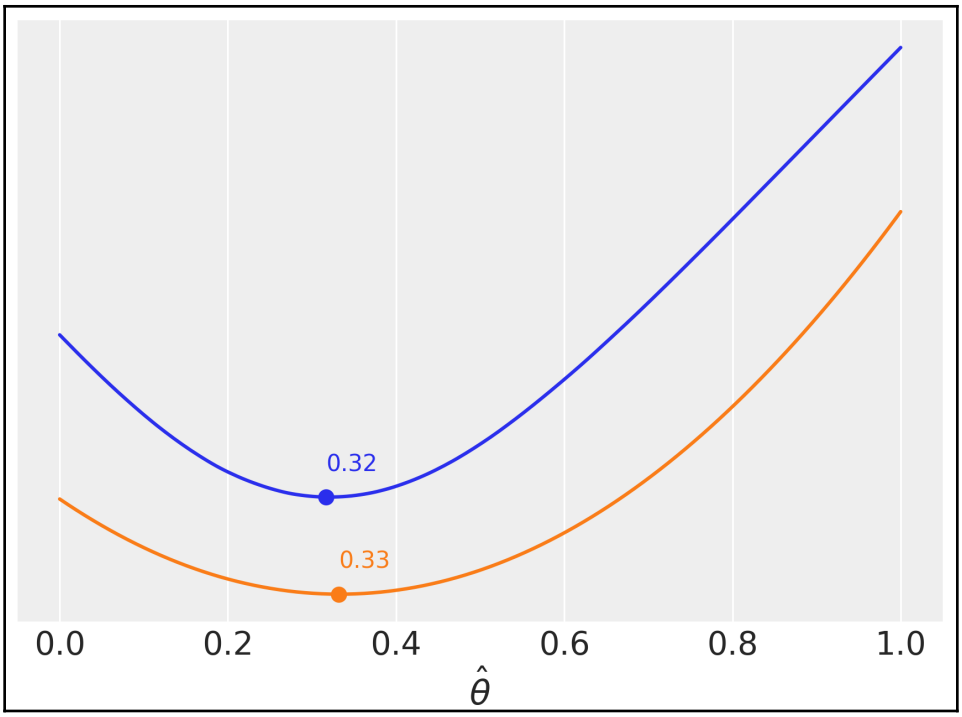
# Chapter 2: Programming Probabilistically

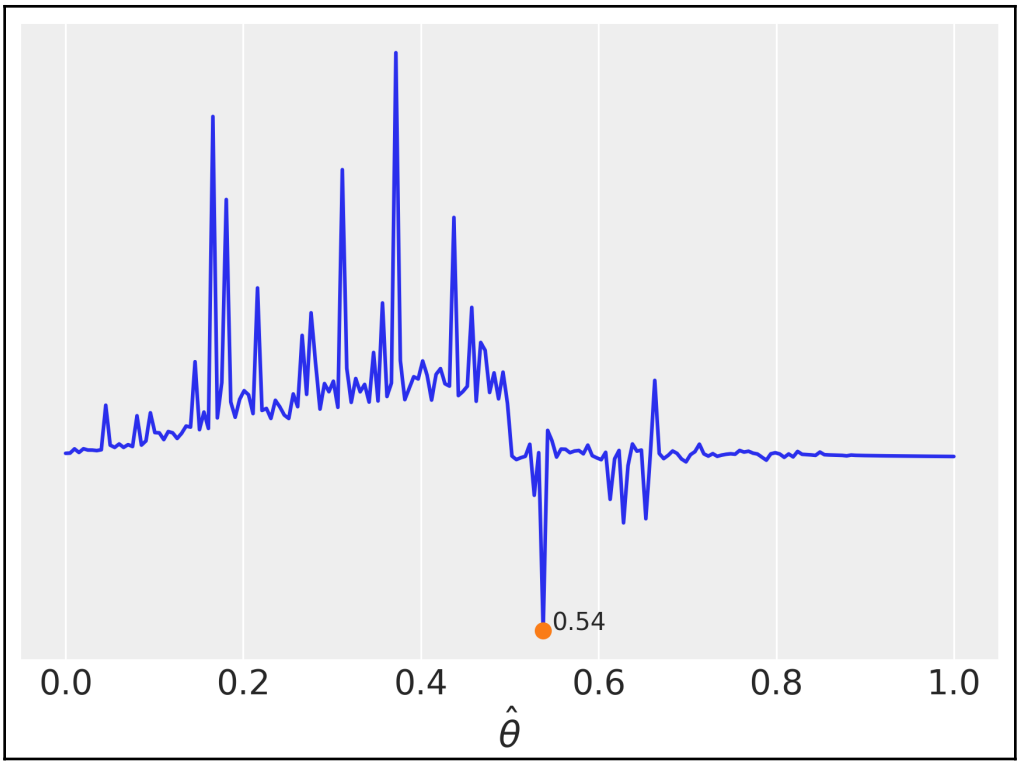


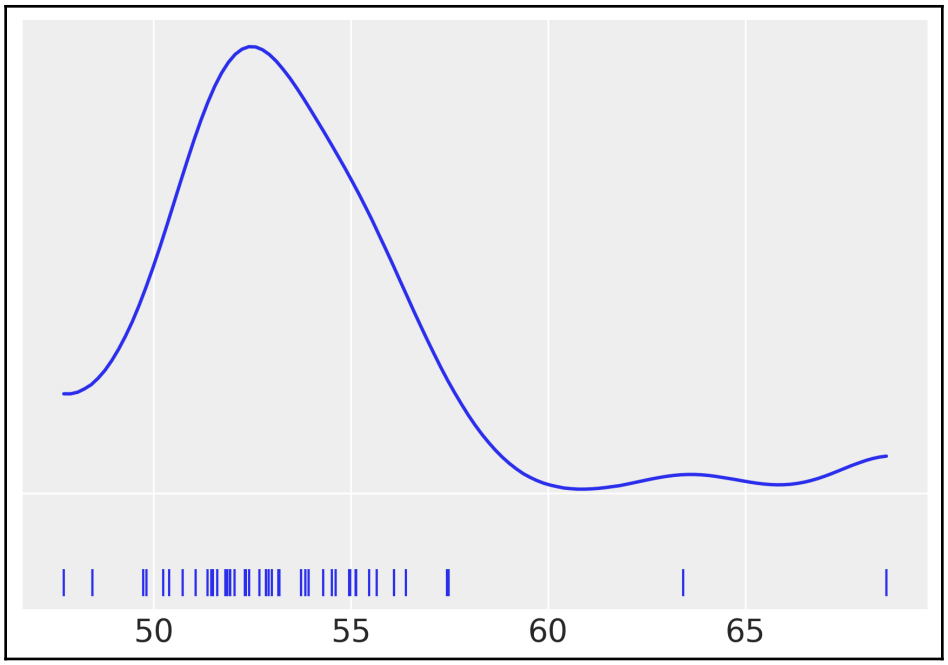



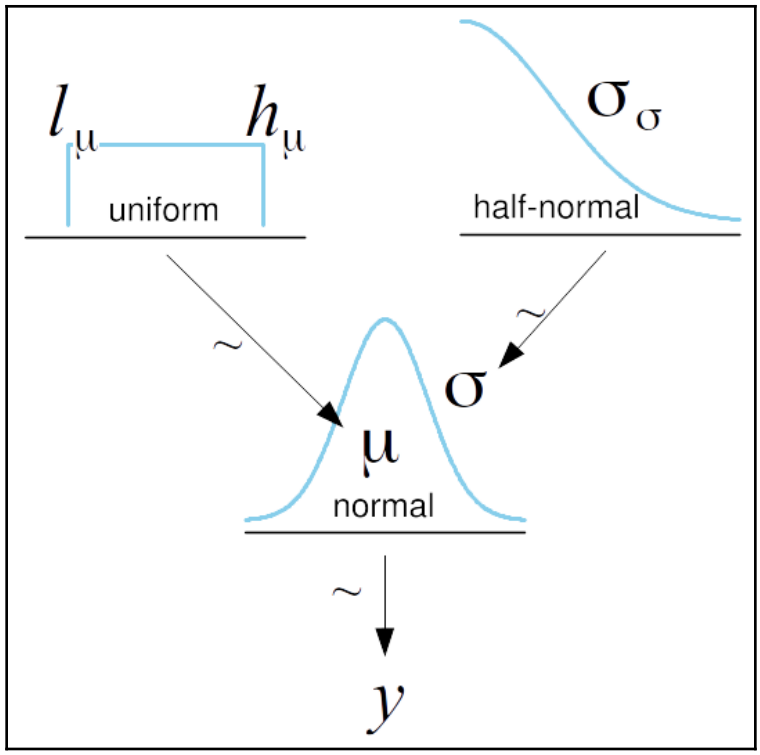

θ
mean=0.3
94% HPD
0.45 0.55
0.0
0.6
0.0
0.2
0.4
0.6
0.8

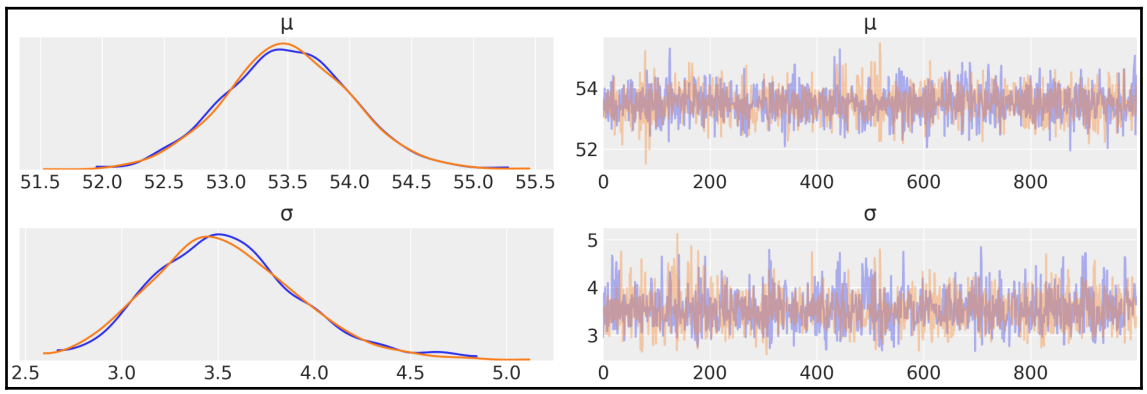

θ
mean=0.3
82.2% <0.5< 17.8%
94% HPD
0.0
0.6
0.0
0.2
0.4
0.6
0.8

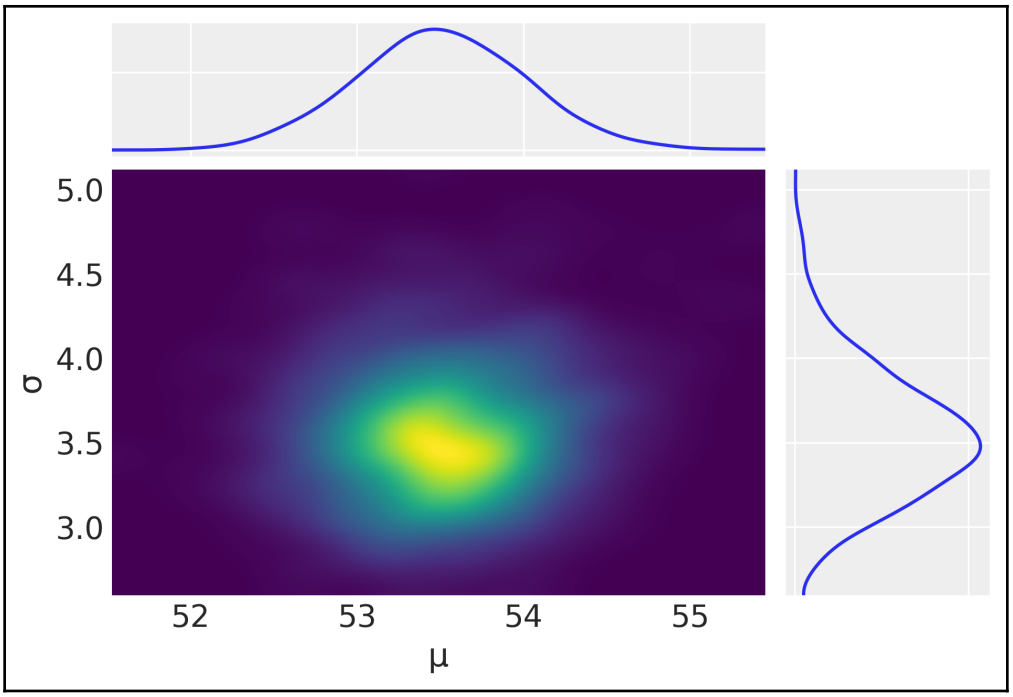

0.32
0.33
0.0
0.2
0.4
0.6
0.8
1.0
$\hat{\theta}$

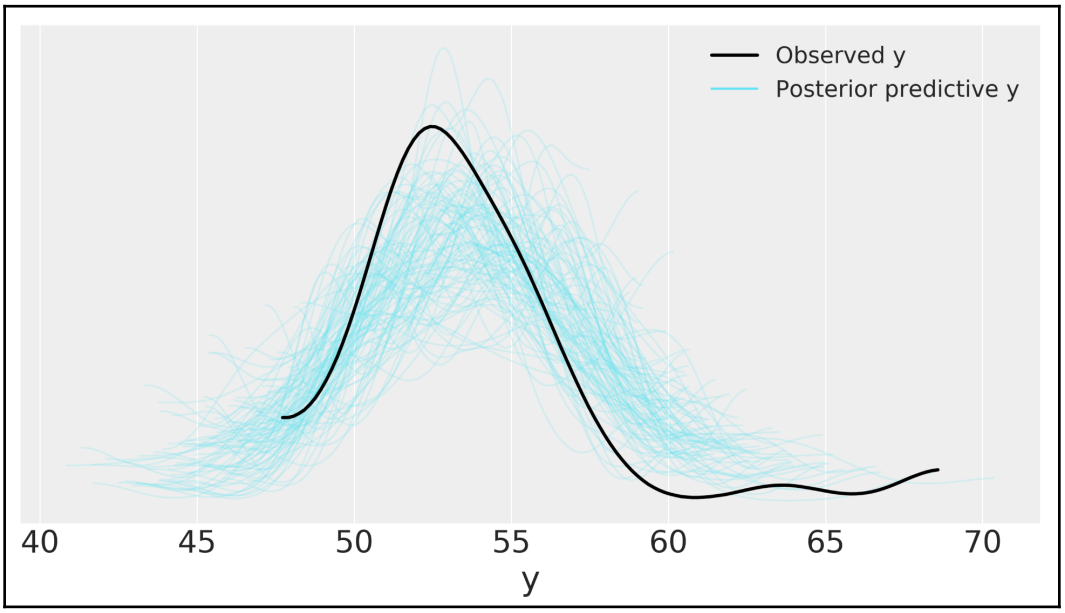

0.54
0.0
0.2
0.4
0.6
0.8
1.0
$\hat{\theta}$

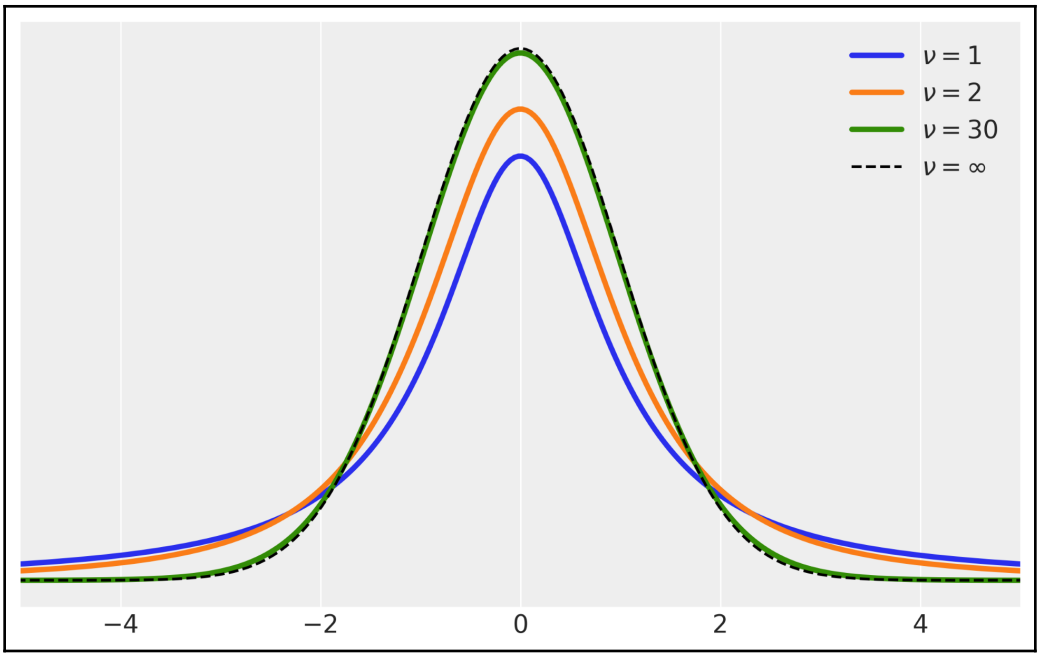

50
55
60
65

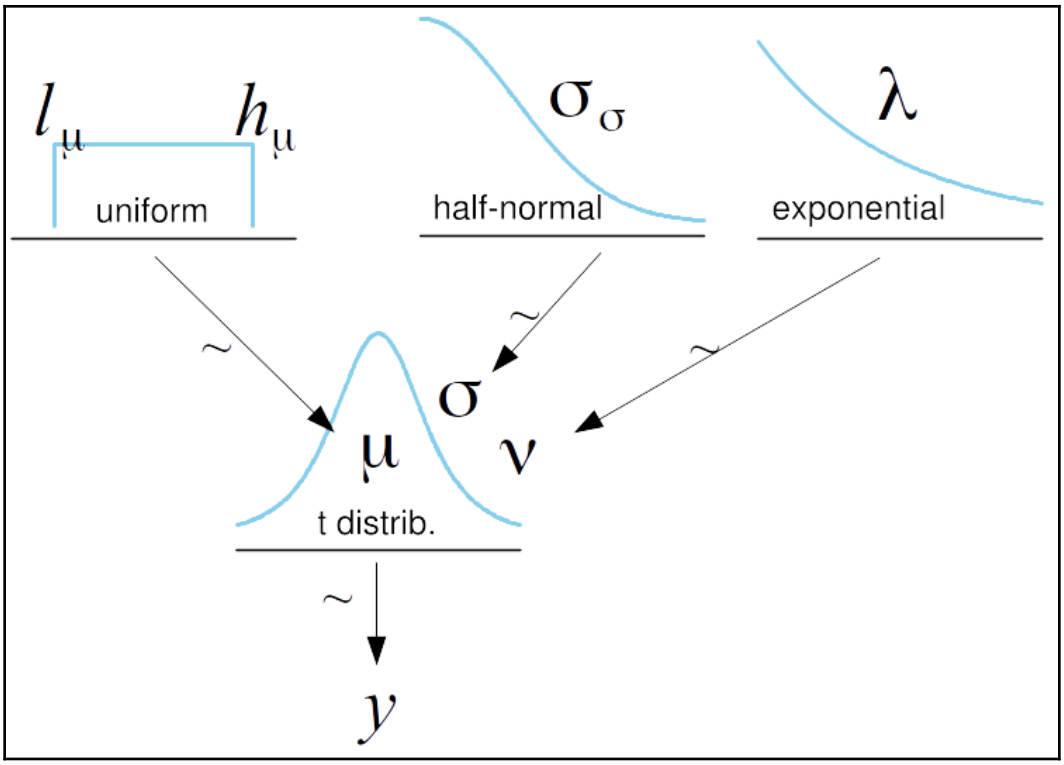

$l_\mu$
$h_\mu$
uniform
$\sigma_\sigma$
half-normal
~
~
$\sigma$
$\mu$
normal
~
$y$

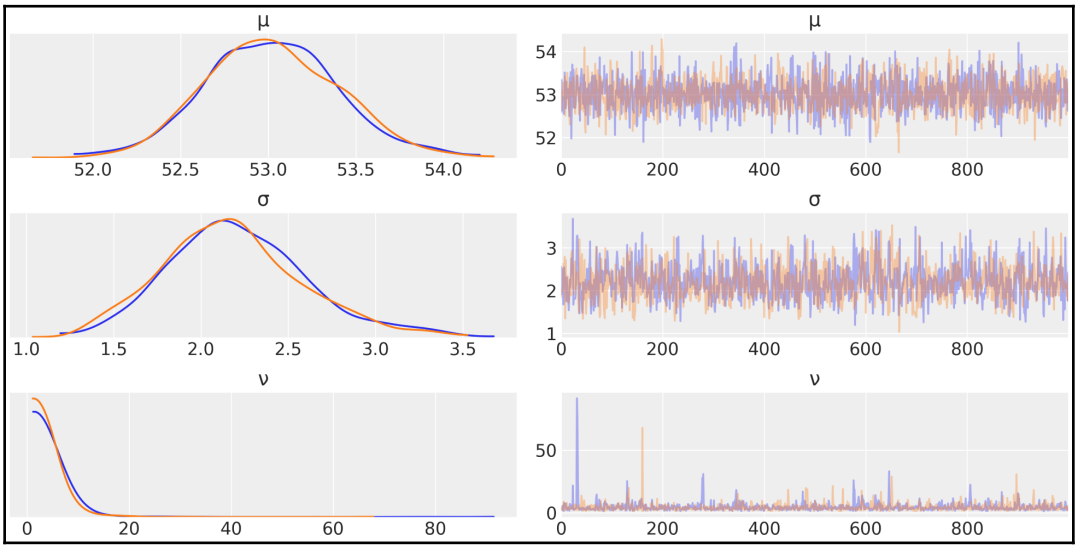

μ
51.5 52.0 52.5 53.0 53.5 54.0 54.5 55.0 55.5
σ
2.5 3.0 3.5 4.0 4.5 5.0
μ
54
52
0 200 400 600 800
σ
5
4
3
0 200 400 600 800

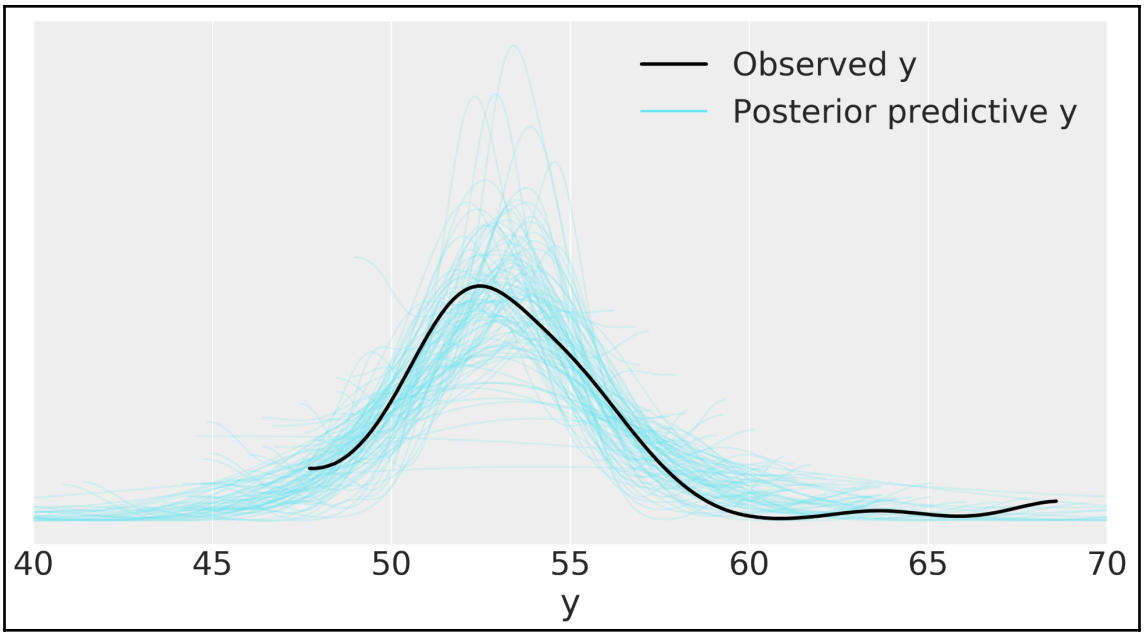

5.0
4.5
4.0
3.5
3.0
σ
52
53
54
55
μ

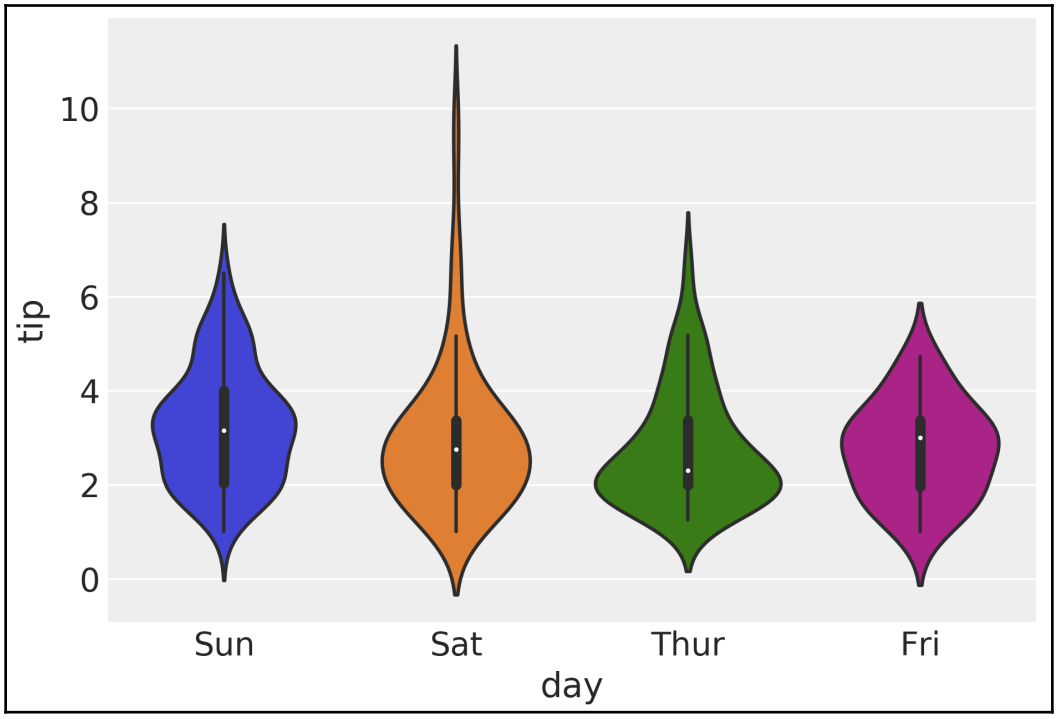

Observed y
Posterior predictive y
40
45
50
55
60
65
70
y

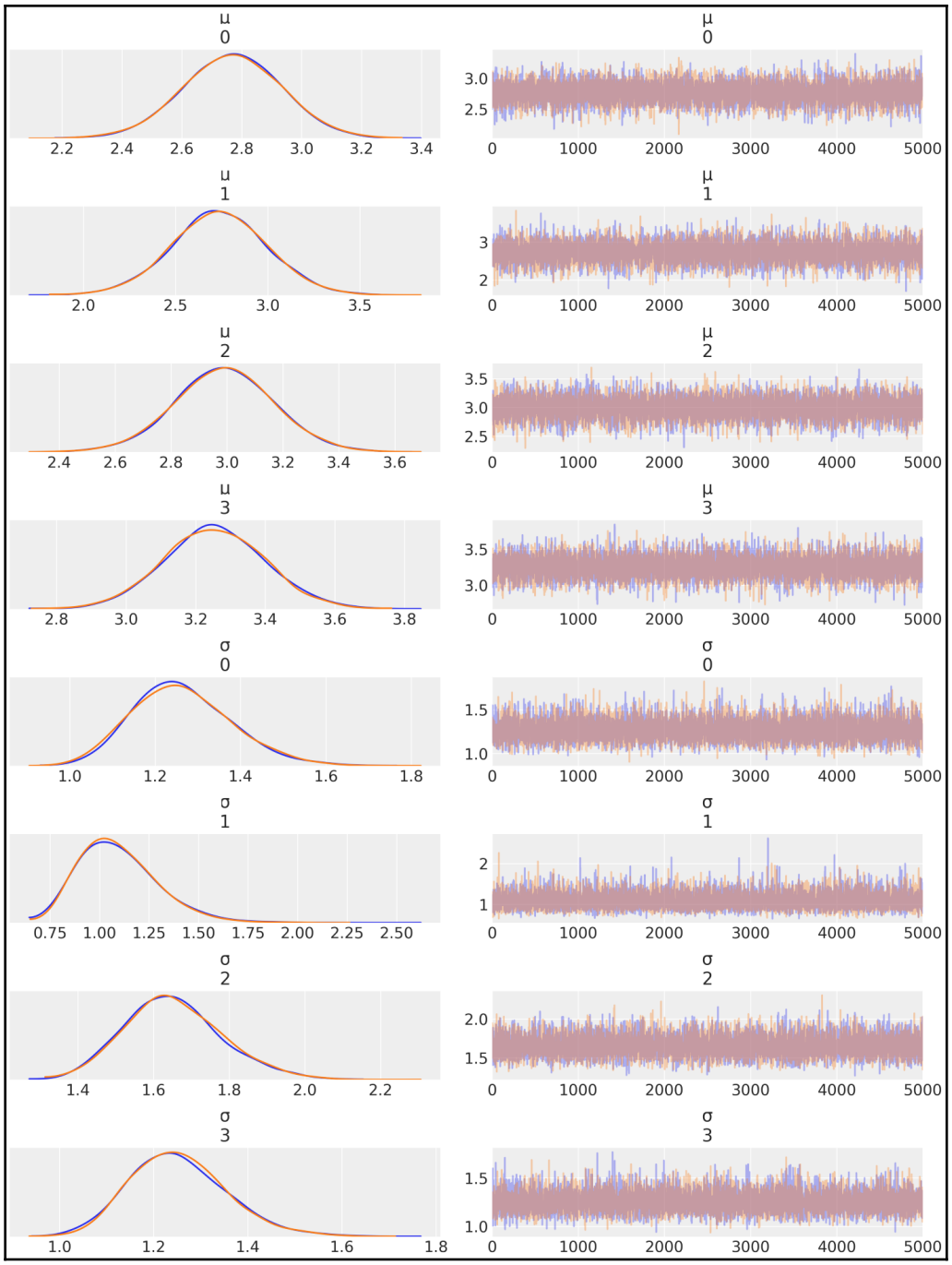

$\nu = 1$
$\nu = 2$
$\nu = 30$
$\nu = \infty$
−4
−2
0
2
4

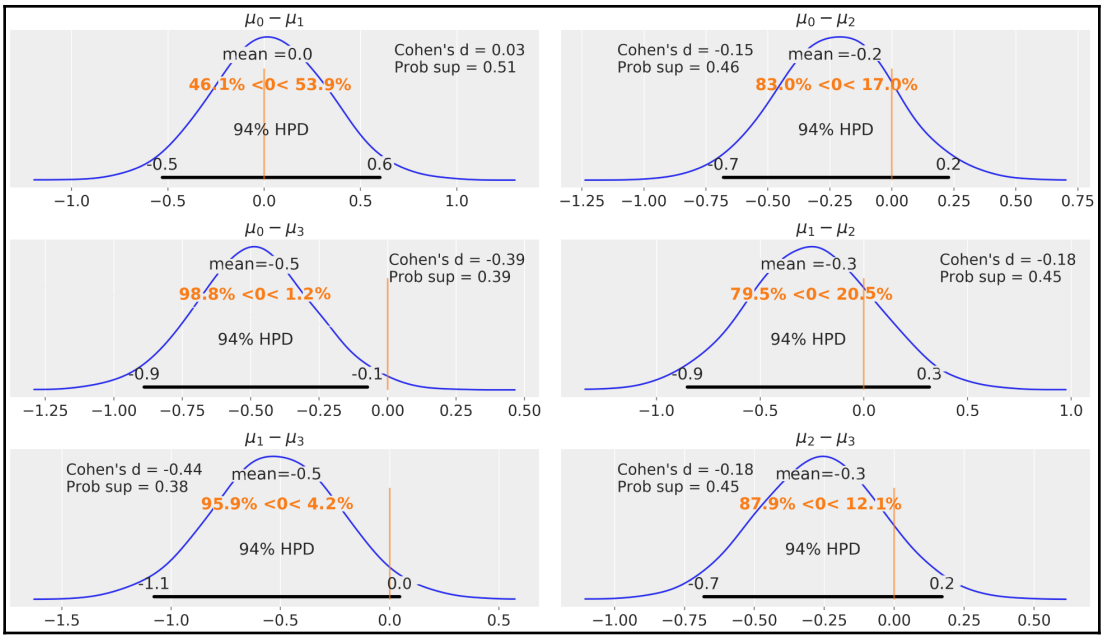

$l_\mu$
$h_\mu$
uniform
$\sigma_\sigma$
half-normal
$\lambda$
exponential
~
~
~
$\sigma$
$\mu$
$\nu$
t distrib.
~
$y$

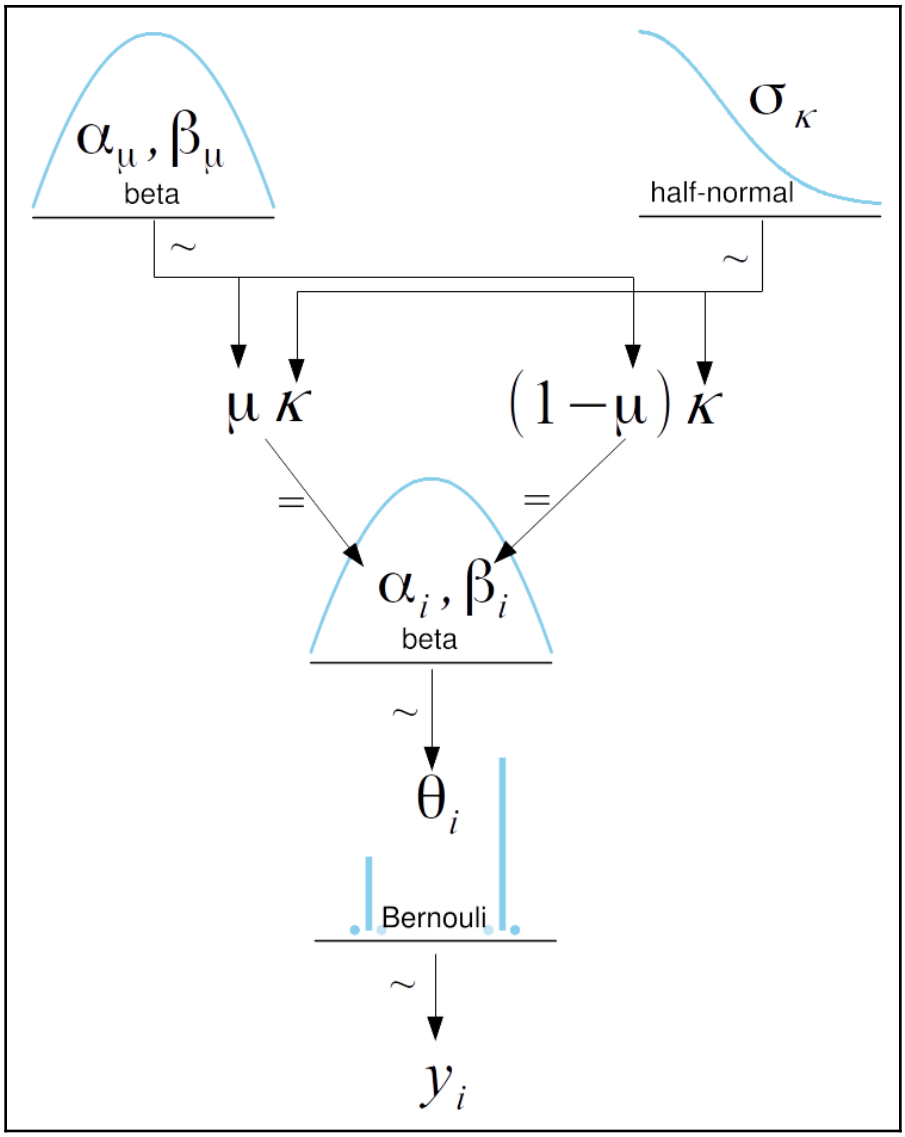

μ
52.0
52.5
53.0
53.5
54.0
μ
54
53
52
0
200
400
600
800
σ
1.0
1.5
2.0
2.5
3.0
3.5
σ
3
2
1
0
200
400
600
800
ν
0
20
40
60
80
ν
50
0
0
200
400
600
800

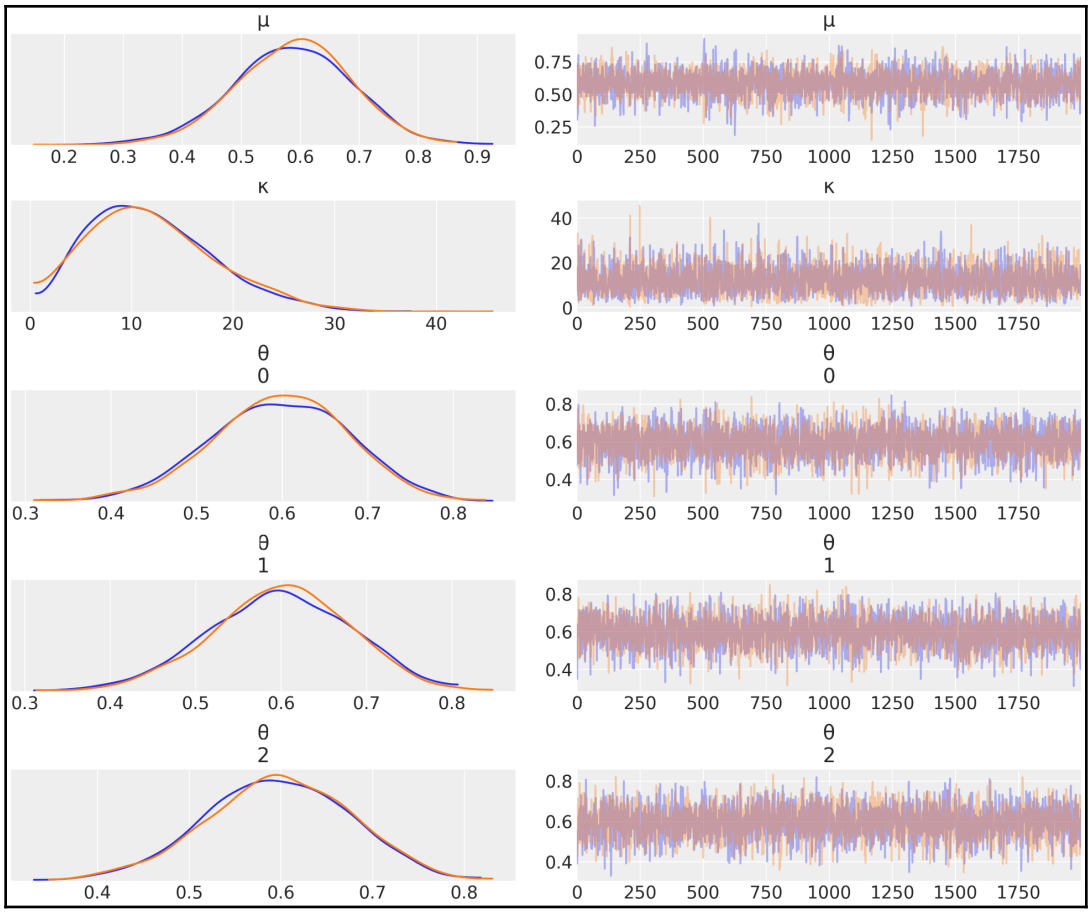

Observed y
Posterior predictive y
40
45
50
55
60
65
70
y

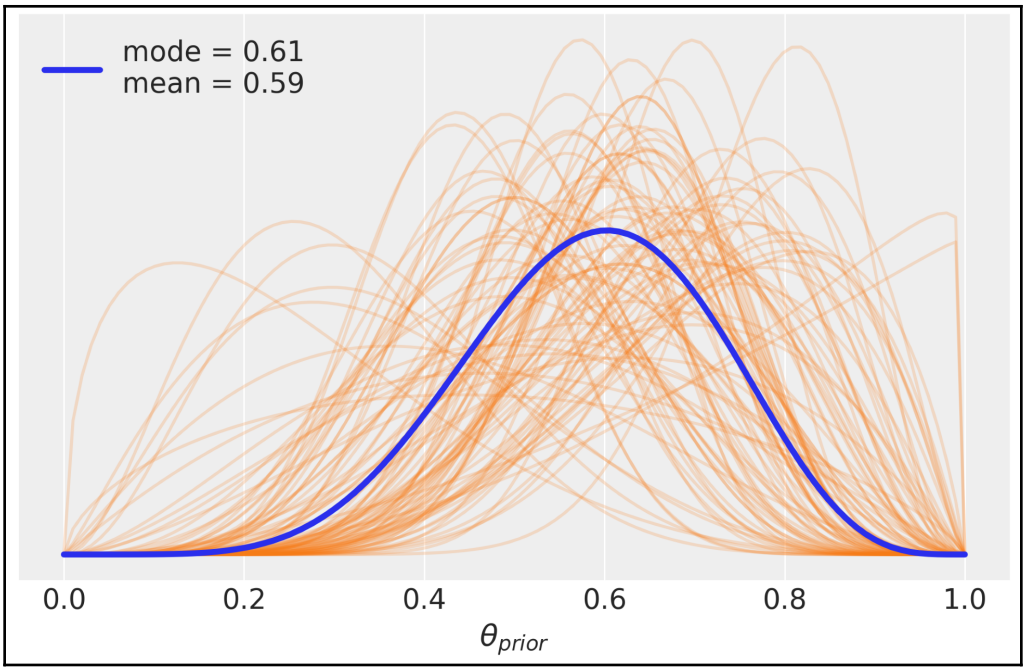

tip
10
8
6
4
2
0
Sun
Sat
Thur
Fri
day

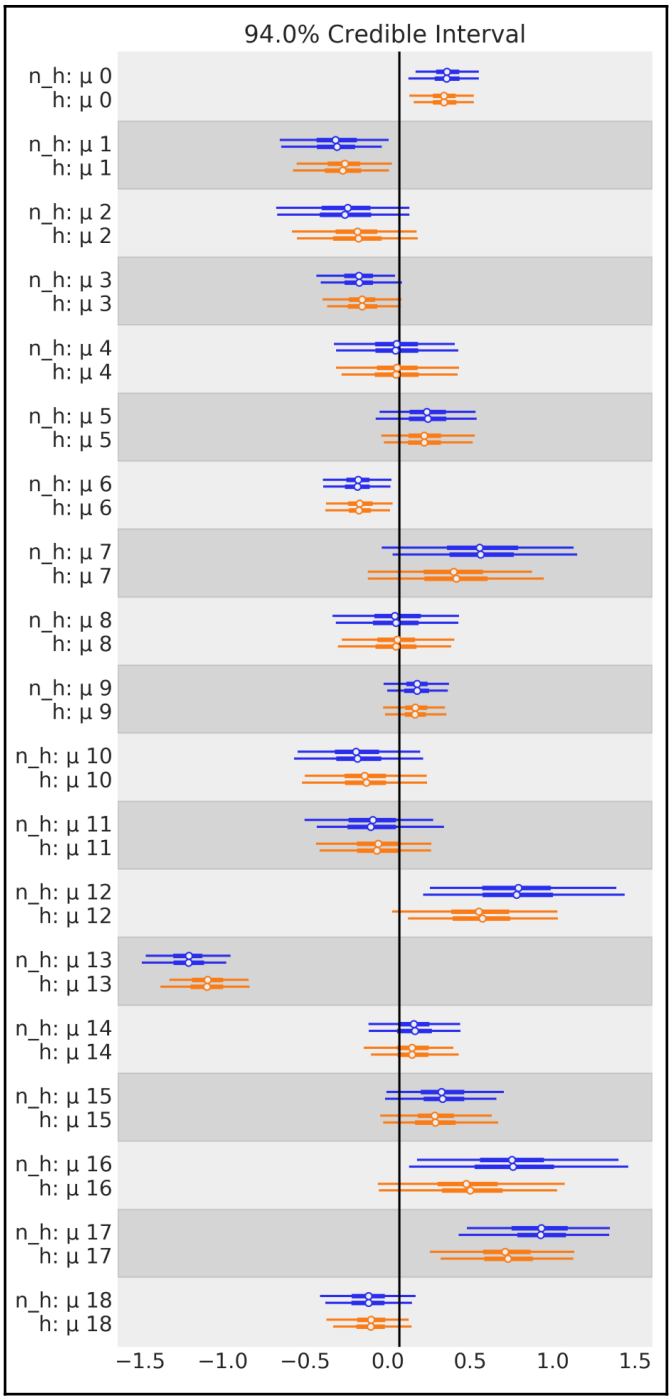

μ
0
2.2 2.4 2.6 2.8 3.0 3.2 3.4
μ
0
3.0
2.5
0 1000 2000 3000 4000 5000
μ
1
2.0 2.5 3.0 3.5
μ
1
3
2
0 1000 2000 3000 4000 5000
μ
2
2.4 2.6 2.8 3.0 3.2 3.4 3.6
μ
2
3.5
3.0
2.5
0 1000 2000 3000 4000 5000
μ
3
2.8 3.0 3.2 3.4 3.6 3.8
μ
3
3.5
3.0
0 1000 2000 3000 4000 5000
σ
0
1.0 1.2 1.4 1.6 1.8
σ
0
1.5
1.0
0 1000 2000 3000 4000 5000
σ
1
0.75 1.00 1.25 1.50 1.75 2.00 2.25 2.50
σ
1
2
1
0 1000 2000 3000 4000 5000
σ
2
1.4 1.6 1.8 2.0 2.2
σ
2
2.0
1.5
0 1000 2000 3000 4000 5000
σ
3
1.0 1.2 1.4 1.6 1.8
σ
3
1.5
1.0
0 1000 2000 3000 4000 5000


$\mu_0 - \mu_1$
mean =0.0
Cohen's d = 0.03
Prob sup = 0.51
46.1% <0< 53.9%
94% HPD
-0.5
0.6
$\mu_0 - \mu_2$
Cohen's d = -0.15
Prob sup = 0.46
mean =-0.2
83.0% <0< 17.0%
94% HPD
-0.7
0.2
$\mu_0 - \mu_3$
mean=-0.5
Cohen's d = -0.39
Prob sup = 0.39
98.8% <0< 1.2%
94% HPD
-0.9
-0.1
$\mu_1 - \mu_2$
mean =-0.3
Cohen's d = -0.18
Prob sup = 0.45
79.5% <0< 20.5%
94% HPD
-0.9
0.3
$\mu_1 - \mu_3$
Cohen's d = -0.44
Prob sup = 0.38
mean=-0.5
95.9% <0< 4.2%
94% HPD
-1.1
0.0
$\mu_2 - \mu_3$
Cohen's d = -0.18
Prob sup = 0.45
mean=-0.3
87.9% <0< 12.1%
94% HPD
-0.7
0.2

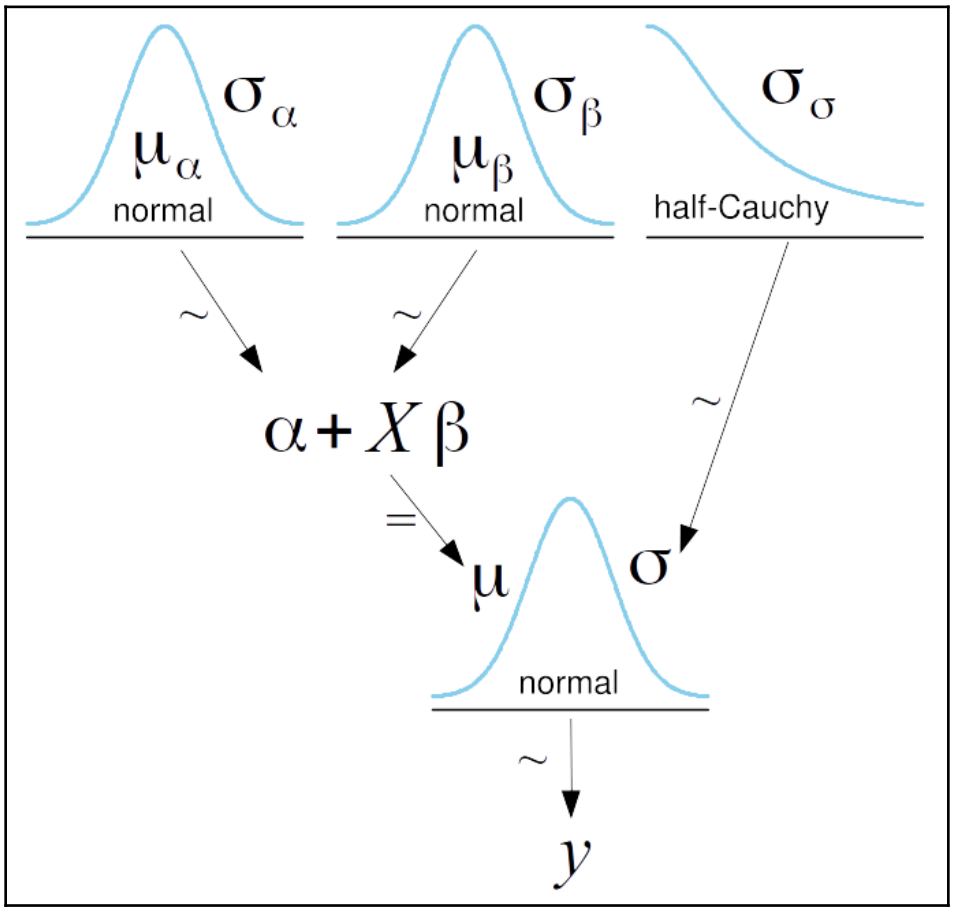

$\alpha_\mu, \beta_\mu$
beta
$\sigma_\kappa$
half-normal
~
~
$\mu\,\kappa$
$(1-\mu)\,\kappa$
=
=
$\alpha_i, \beta_i$
beta
~
$\theta_i$
Bernouli
~
$y_i$

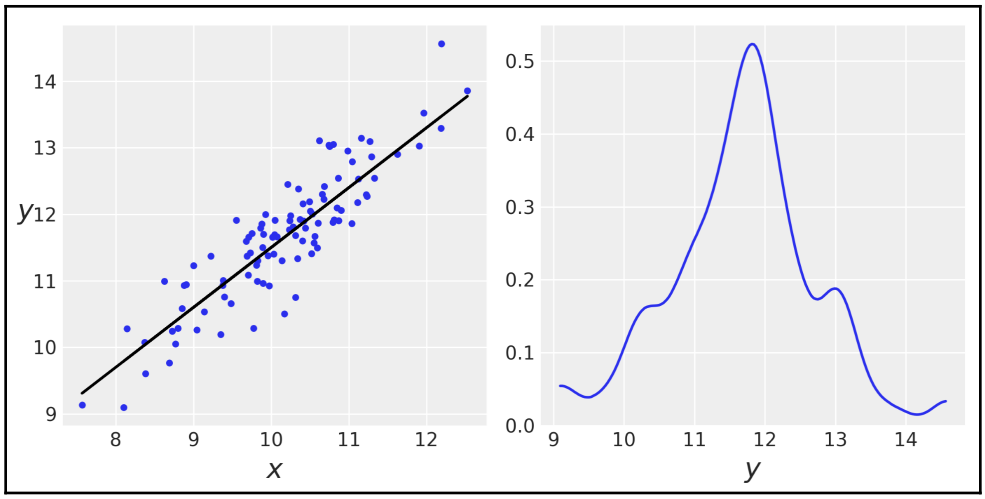

μ
0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9
μ
0.75 0.50 0.25
0 250 500 750 1000 1250 1500 1750
κ
0 10 20 30 40
κ
40 20 0
0 250 500 750 1000 1250 1500 1750
θ 0
0.3 0.4 0.5 0.6 0.7 0.8
θ 0
0.8 0.6 0.4
0 250 500 750 1000 1250 1500 1750
θ 1
0.3 0.4 0.5 0.6 0.7 0.8
θ 1
0.8 0.6 0.4
0 250 500 750 1000 1250 1500 1750
θ 2
0.4 0.5 0.6 0.7 0.8
θ 2
0.8 0.6 0.4
0 250 500 750 1000 1250 1500 1750

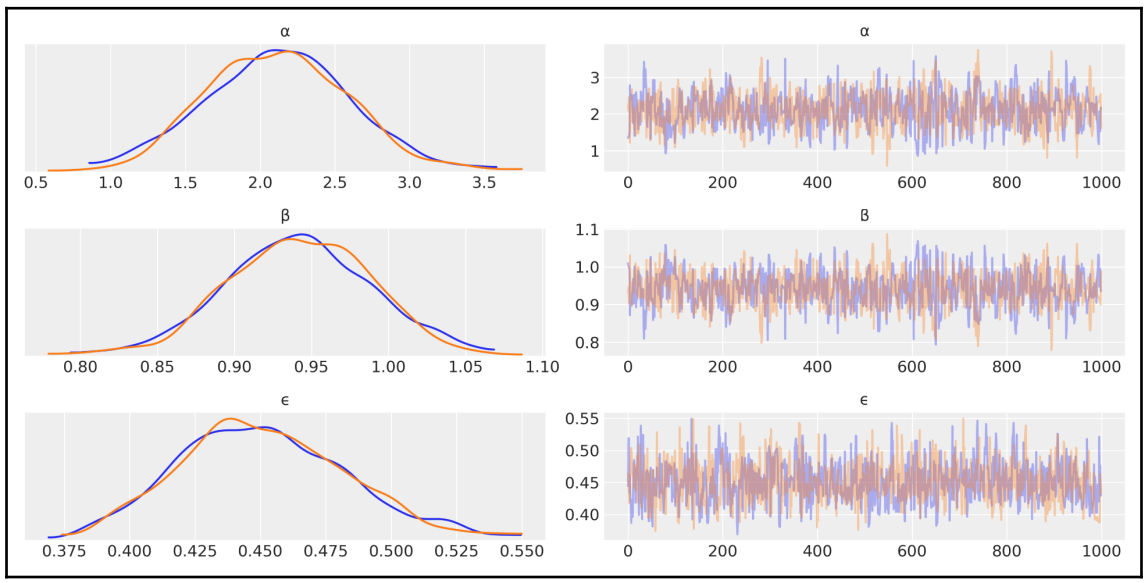

mode = 0.61
mean = 0.59
0.0
0.2
0.4
0.6
0.8
1.0
$\theta_{prior}$

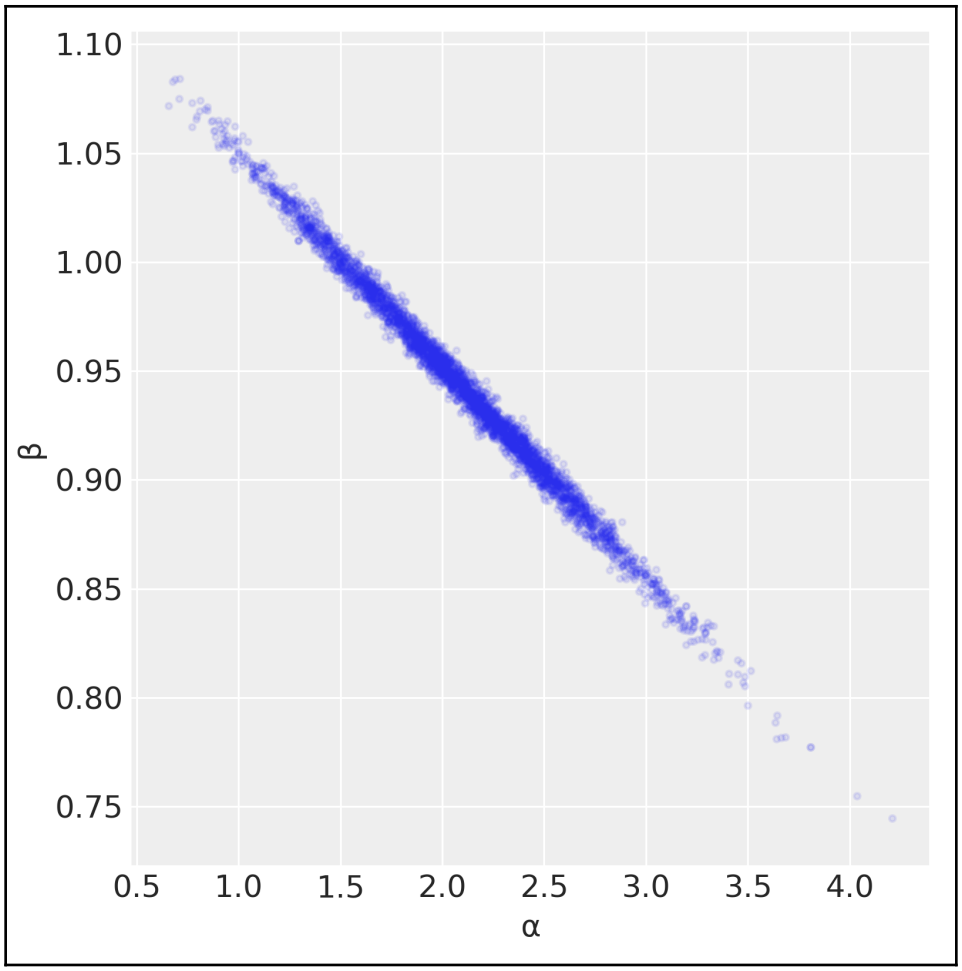

94.0% Credible Interval
n_h: μ 0
h: μ 0
n_h: μ 1
h: μ 1
n_h: μ 2
h: μ 2
n_h: μ 3
h: μ 3
n_h: μ 4
h: μ 4
n_h: μ 5
h: μ 5
n_h: μ 6
h: μ 6
n_h: μ 7
h: μ 7
n_h: μ 8
h: μ 8
n_h: μ 9
h: μ 9
n_h: μ 10
h: μ 10
n_h: μ 11
h: μ 11
n_h: μ 12
h: μ 12
n_h: μ 13
h: μ 13
n_h: μ 14
h: μ 14
n_h: μ 15
h: μ 15
n_h: μ 16
h: μ 16
n_h: μ 17
h: μ 17
n_h: μ 18
h: μ 18
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5

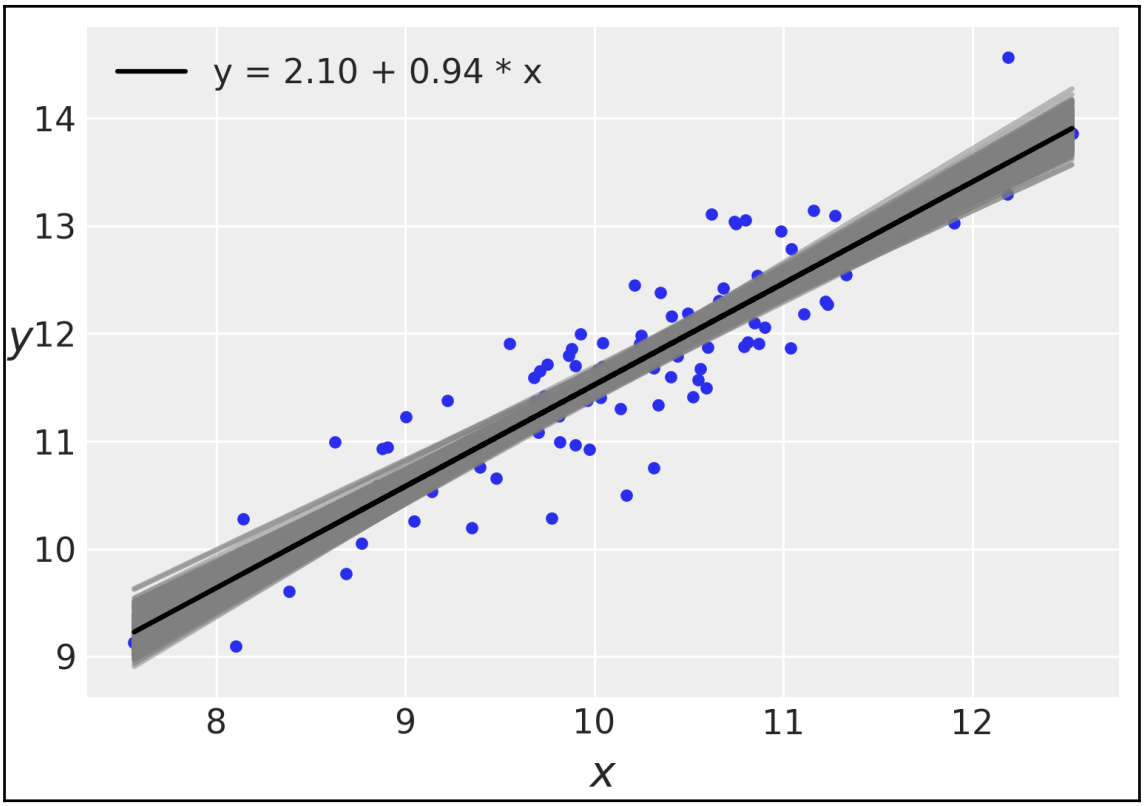
# Chapter 3: Modeling with Linear Regression

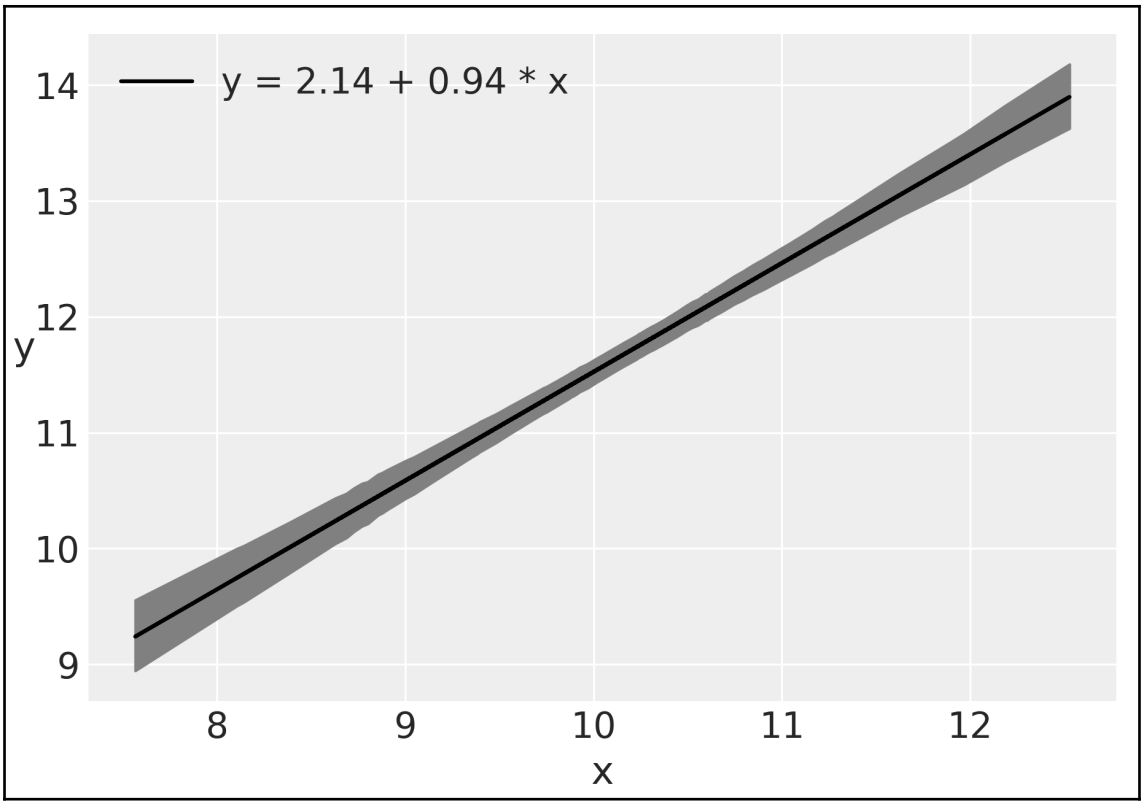


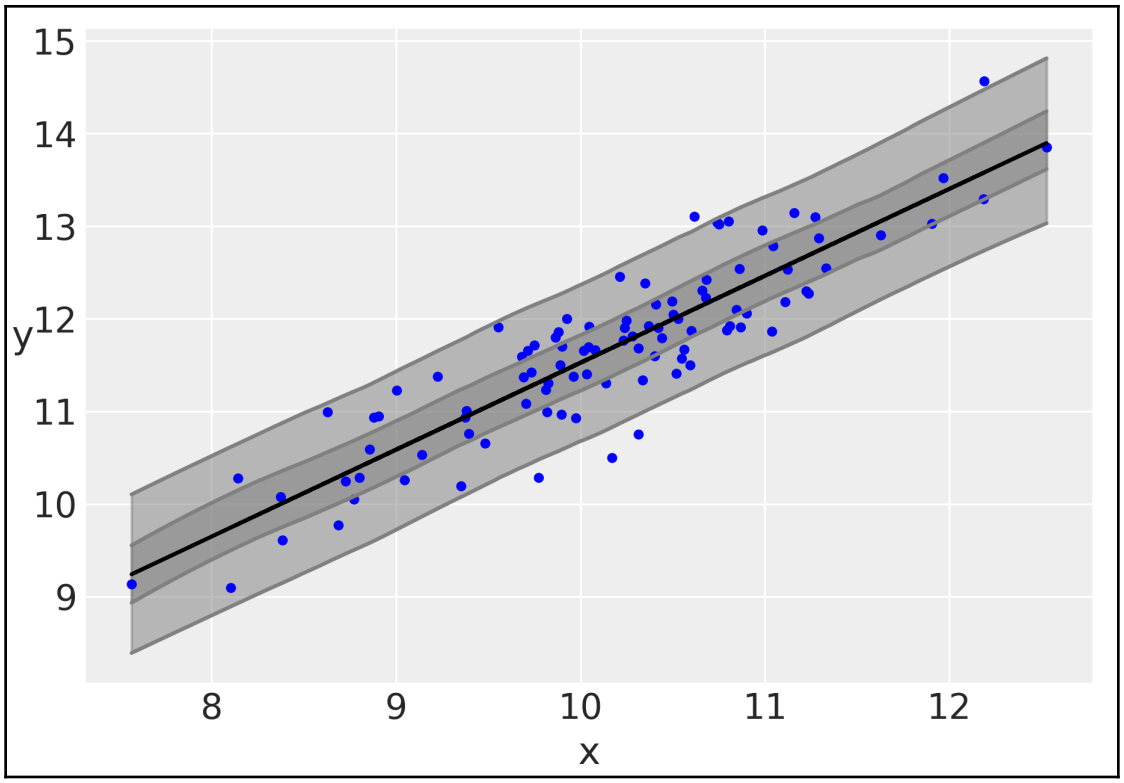

y
x
14
13
12
11
10
9
8
9
10
11
12
0.5
0.4
0.3
0.2
0.1
0.0
9
10
11
12
13
14
y

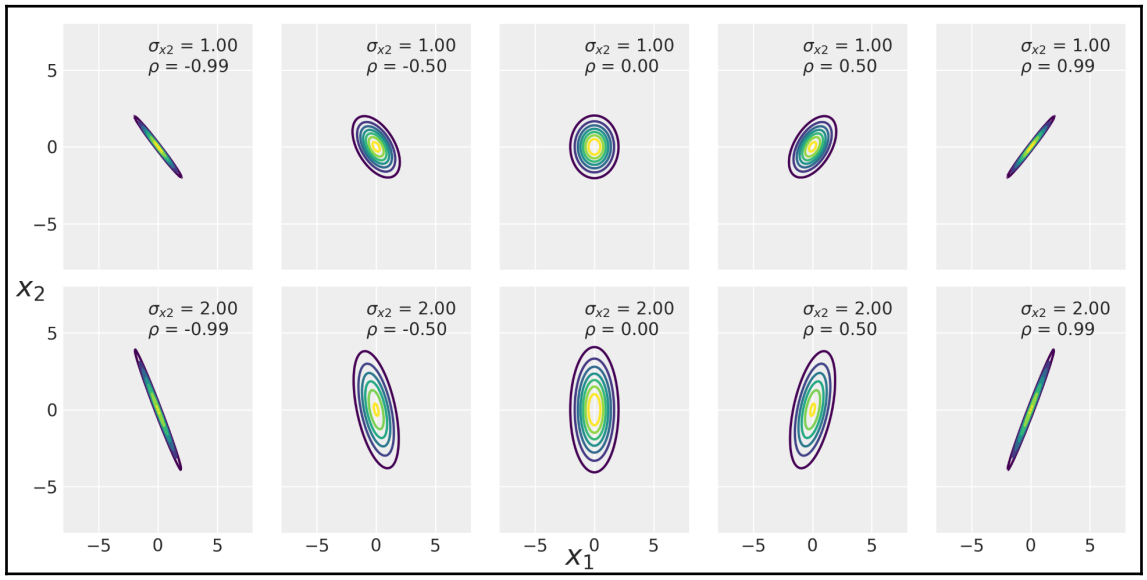

α
β
ϵ
0.5
1.0
1.5
2.0
2.5
3.0
3.5
0.80
0.85
0.90
0.95
1.00
1.05
1.10
0.375
0.400
0.425
0.450
0.475
0.500
0.525
0.550
3
2
1
1.1
1.0
0.9
0.8
0.55
0.50
0.45
0.40
0
200
400
600
800
1000

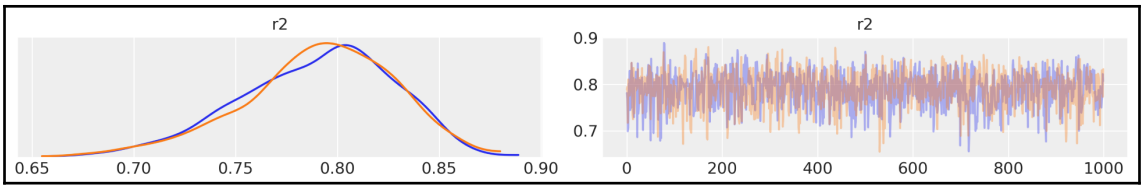

1.10
1.05
1.00
0.95
0.90
0.85
0.80
0.75
β
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
α

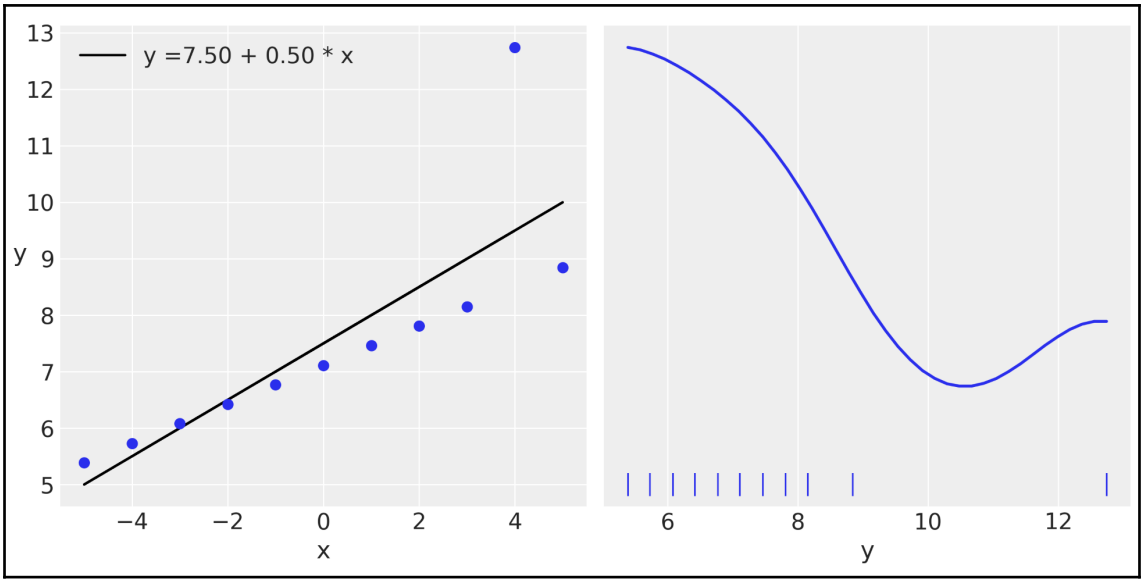

y = 2.10 + 0.94 * x
14
13
12
11
10
9
y
8
9
10
11
12
x

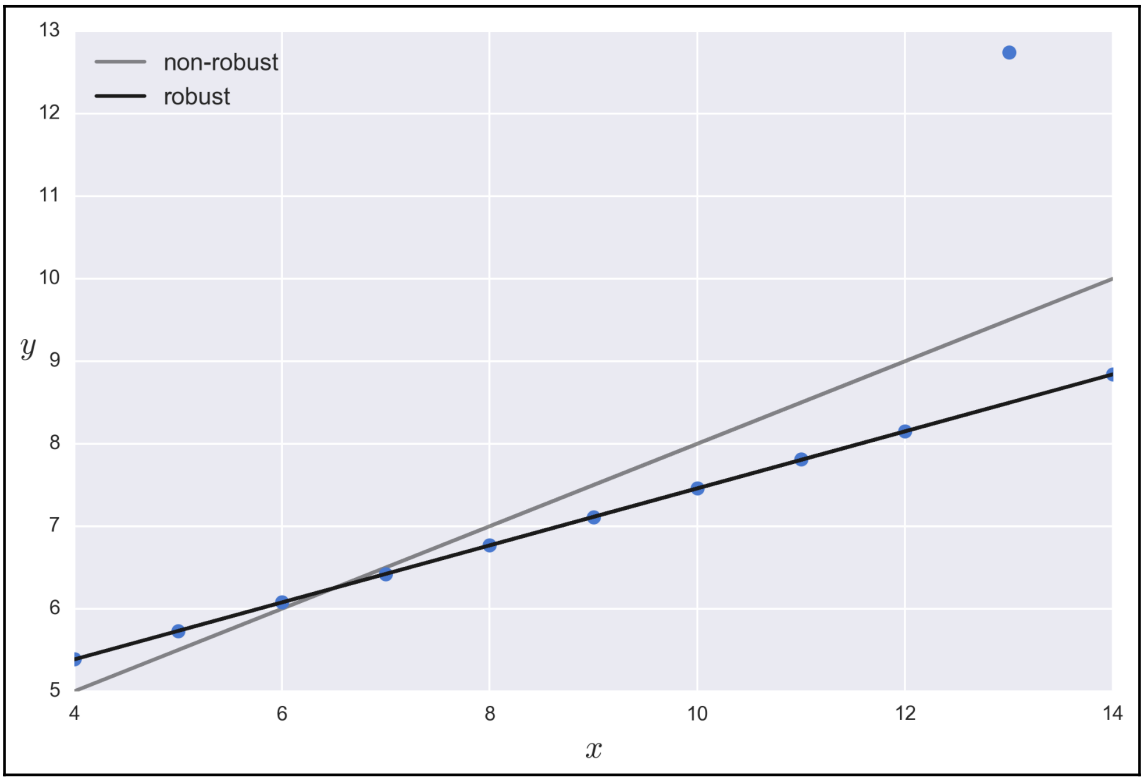

y = 2.14 + 0.94 * x
14
13
12
11
10
9
y
8
9
10
11
12
x

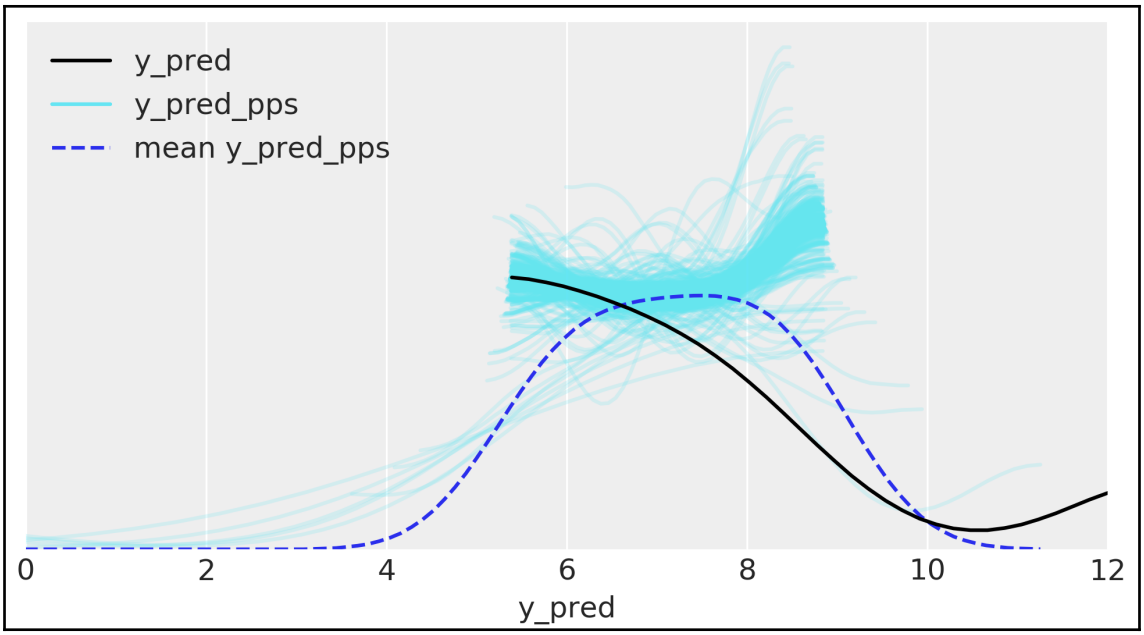

15
14
13
12
11
10
9
y
8
9
10
11
12
x

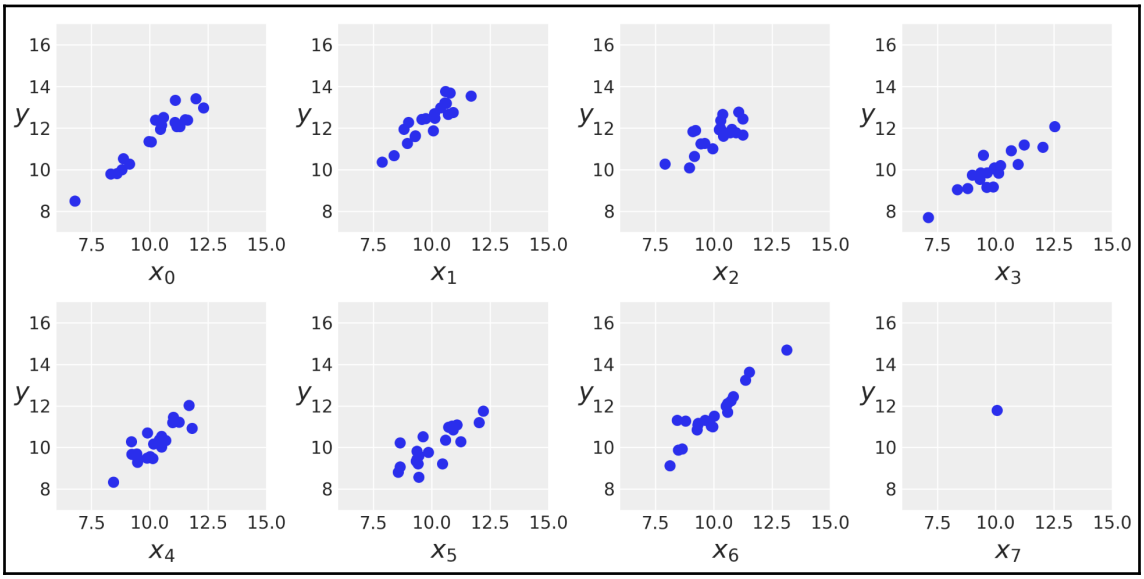

$\sigma_{x2}$ = 1.00
$\rho$ = -0.99
$\sigma_{x2}$ = 1.00
$\rho$ = -0.50
$\sigma_{x2}$ = 1.00
$\rho$ = 0.00
$\sigma_{x2}$ = 1.00
$\rho$ = 0.50
$\sigma_{x2}$ = 1.00
$\rho$ = 0.99
$\sigma_{x2}$ = 2.00
$\rho$ = -0.99
$\sigma_{x2}$ = 2.00
$\rho$ = -0.50
$\sigma_{x2}$ = 2.00
$\rho$ = 0.00
$\sigma_{x2}$ = 2.00
$\rho$ = 0.50
$\sigma_{x2}$ = 2.00
$\rho$ = 0.99
$x_2$
$x_1$

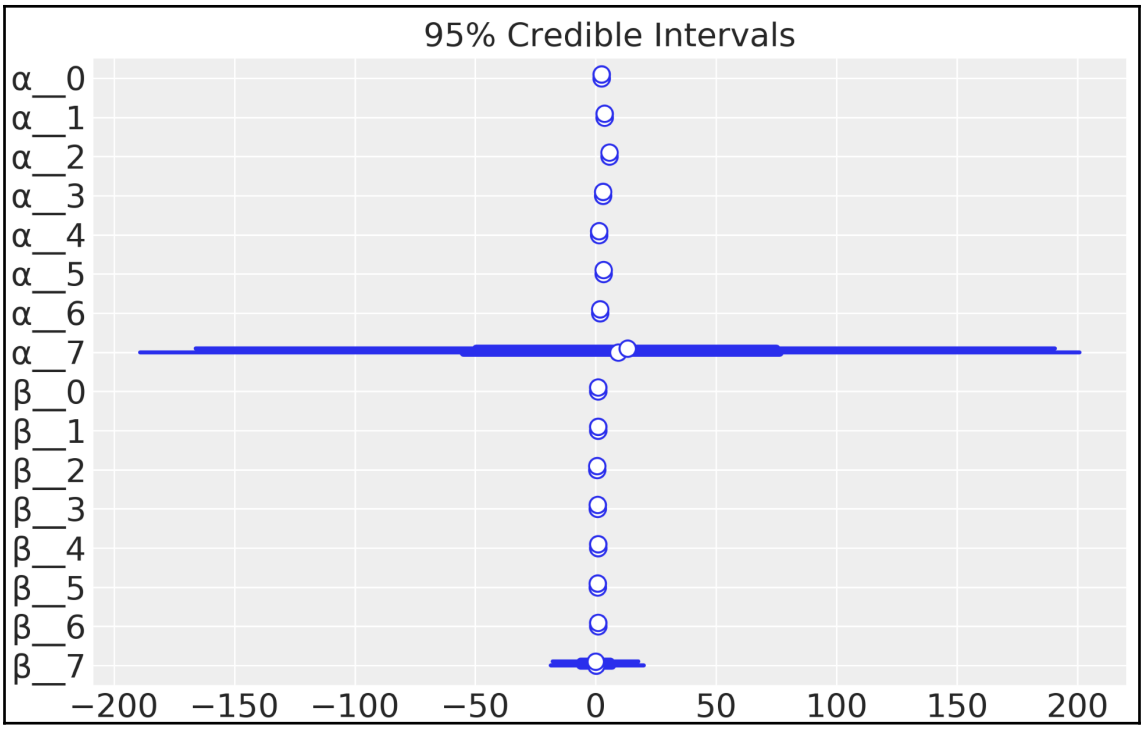

r2
0.65
0.70
0.75
0.80
0.85
0.90
r2
0.9
0.8
0.7
0
200
400
600
800
1000

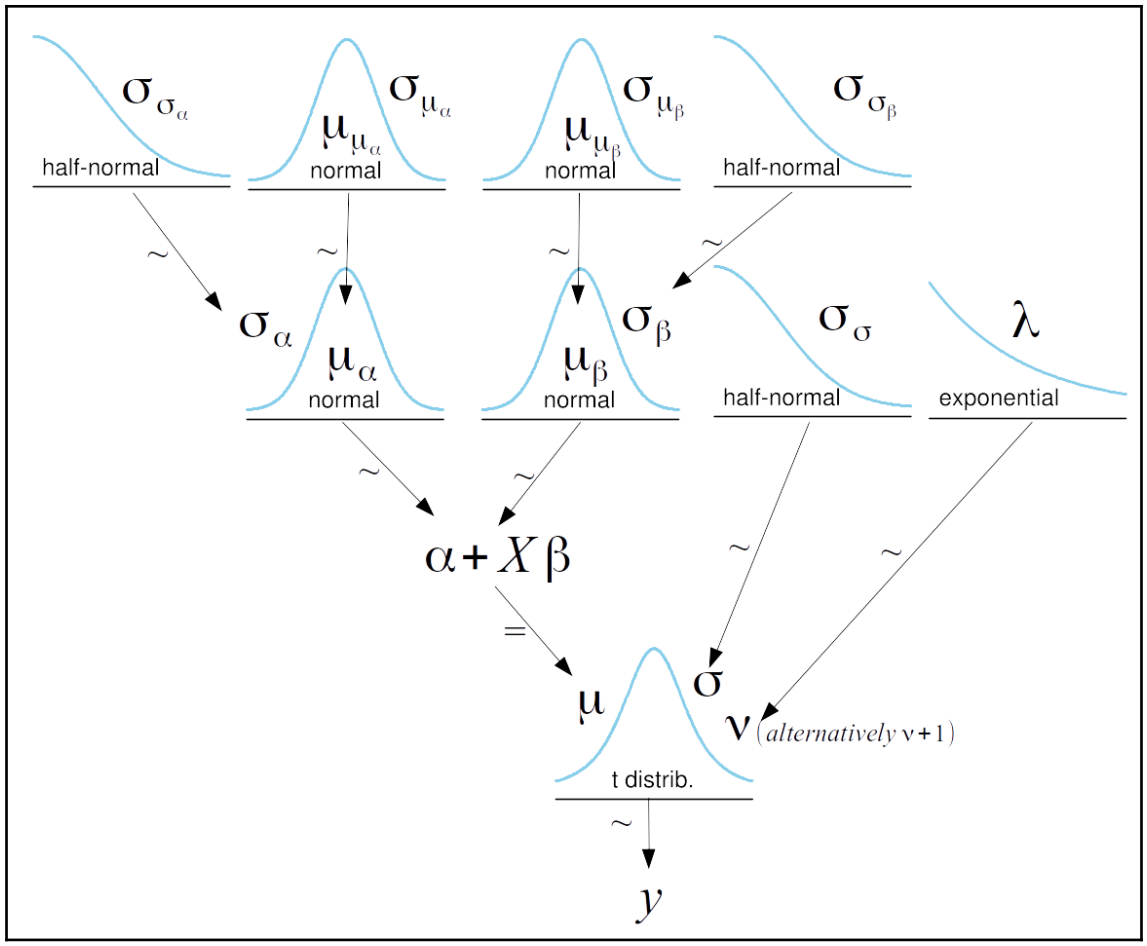

y =7.50 + 0.50 * x
13
12
11
10
y 9
8
7
6
5
−4
−2
0
2
4
x
6
8
10
12
y

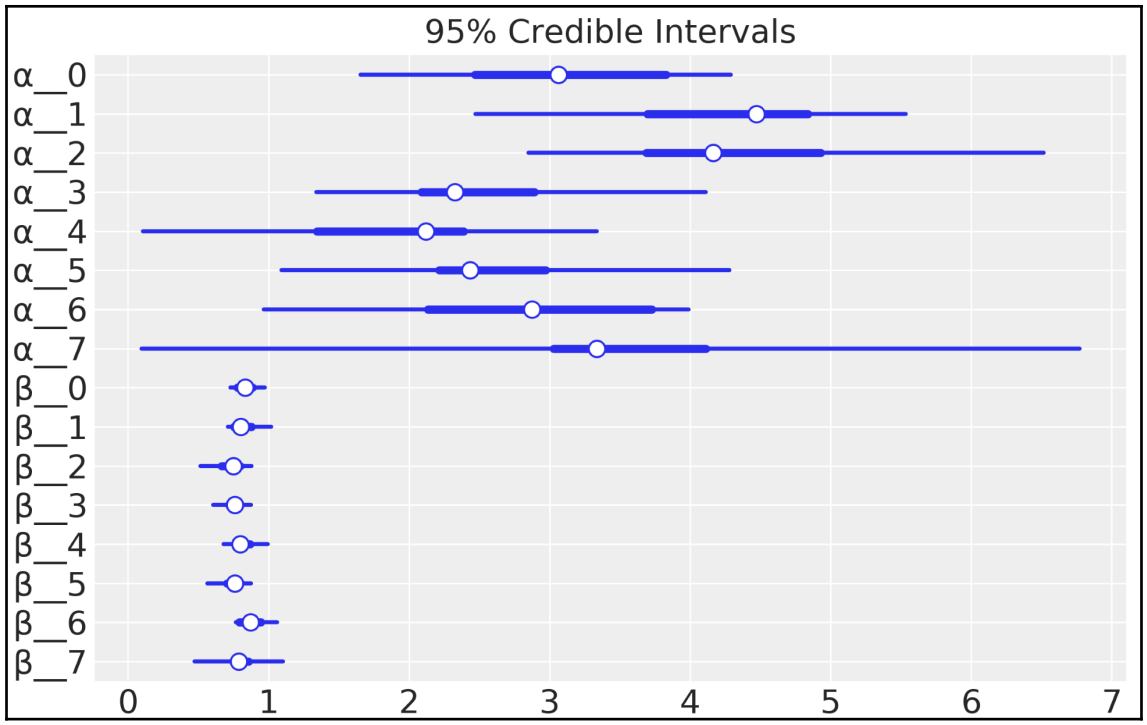

non-robust
robust
13
12
11
10
9
8
7
6
5
y
4
6
8
10
12
14
x

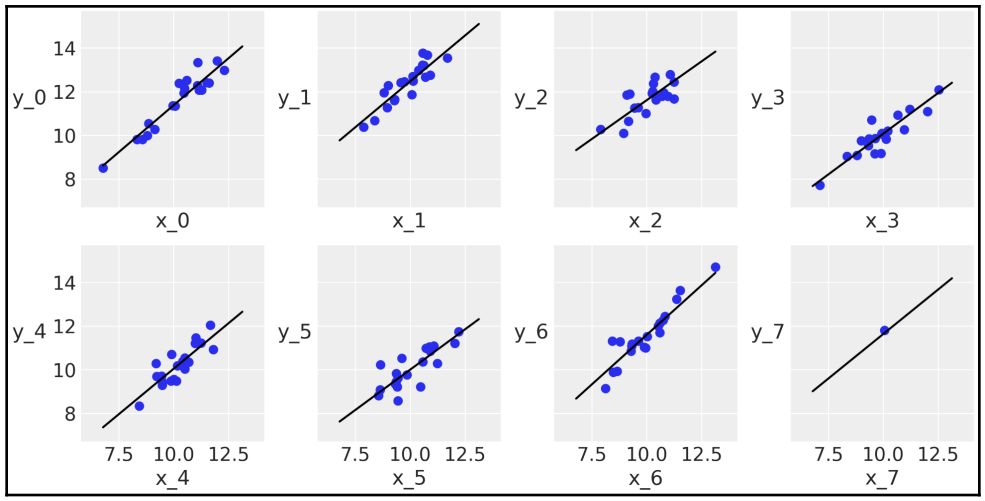

y_pred
y_pred_pps
mean y_pred_pps
0
2
4
6
8
10
12
y_pred

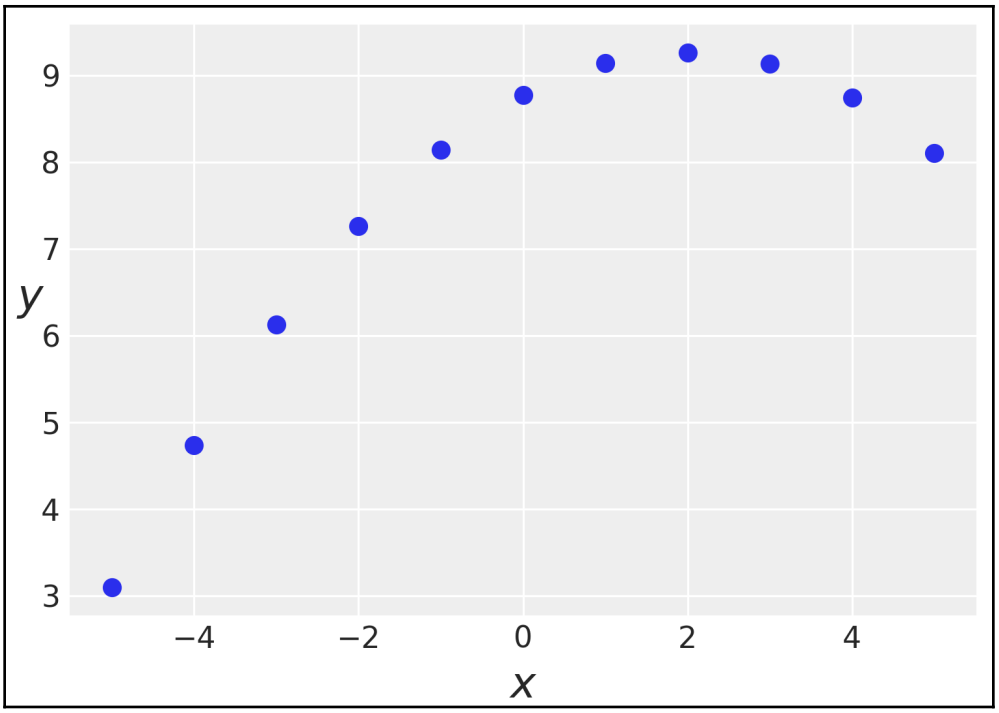

16
14
12
10
8
y
7.5
10.0
12.5
15.0
$x_0$
$x_1$
$x_2$
$x_3$
$x_4$
$x_5$
$x_6$
$x_7$

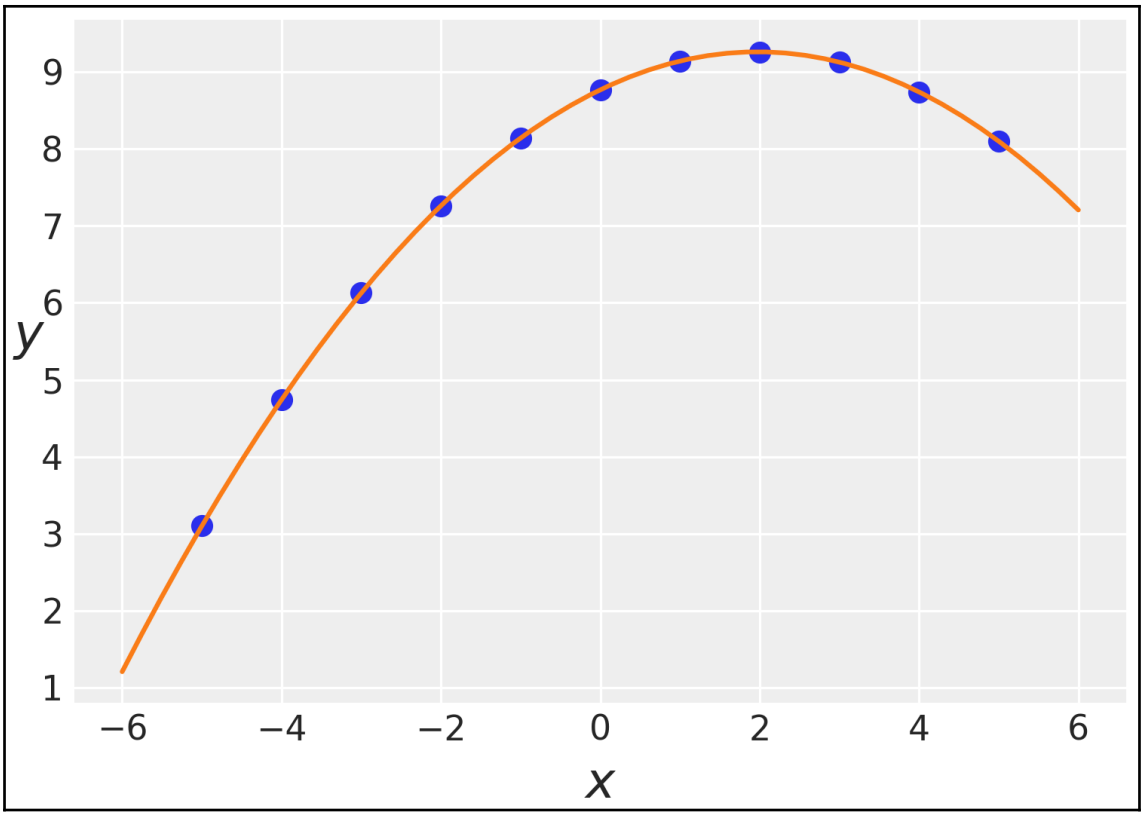

95% Credible Intervals
α__0
α__1
α__2
α__3
α__4
α__5
α__6
α__7
β__0
β__1
β__2
β__3
β__4
β__5
β__6
β__7
−200
−150
−100
−50
0
50
100
150
200

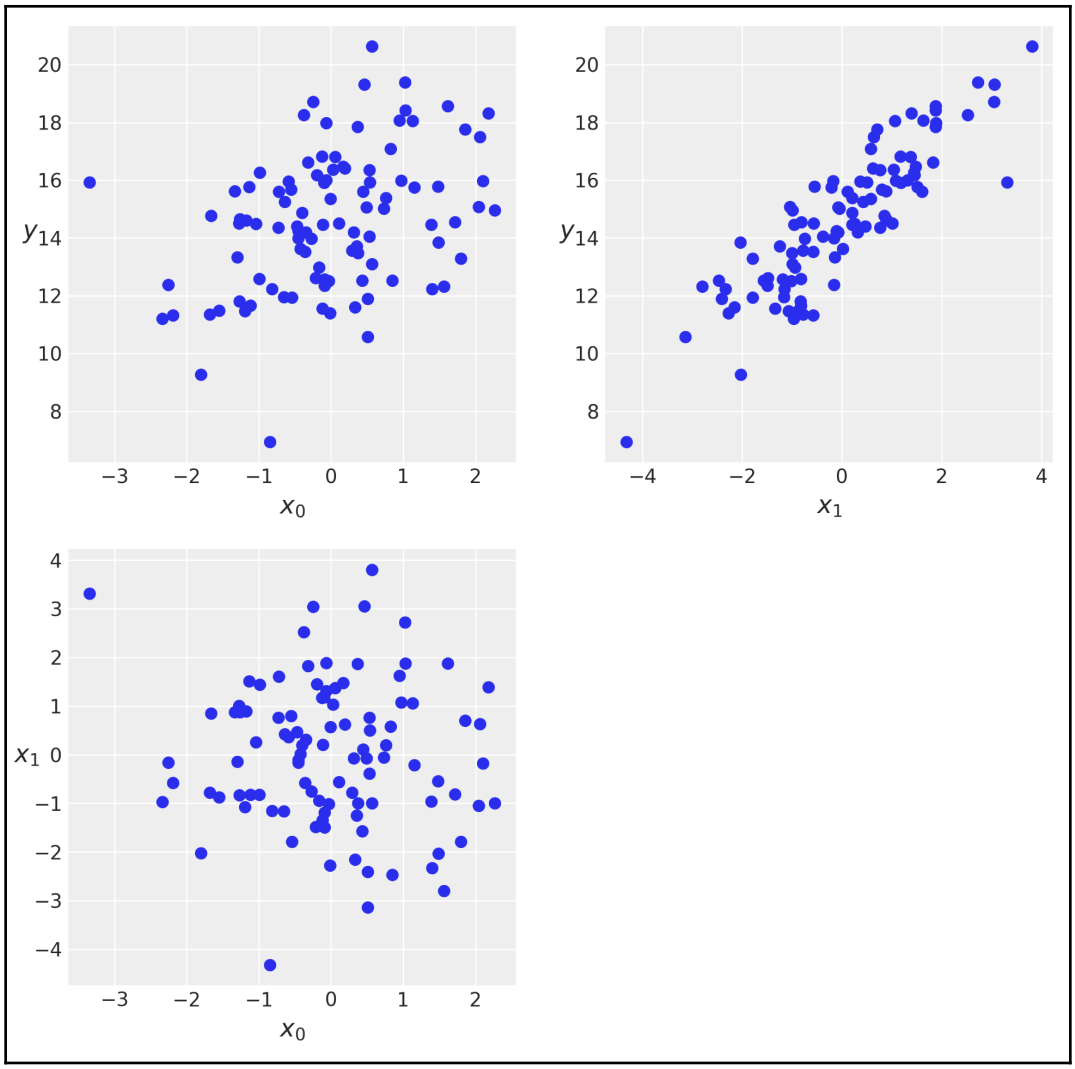

σ σα
half-normal
μ μα
σ σμα
normal
μ μμβ
σ σμβ
normal
σ σβ
half-normal
σ α
μ α
normal
μ β
σ β
normal
σ σ
half-normal
λ
exponential
α + X β
=
μ
σ
ν (alternatively ν + 1)
t distrib.
y

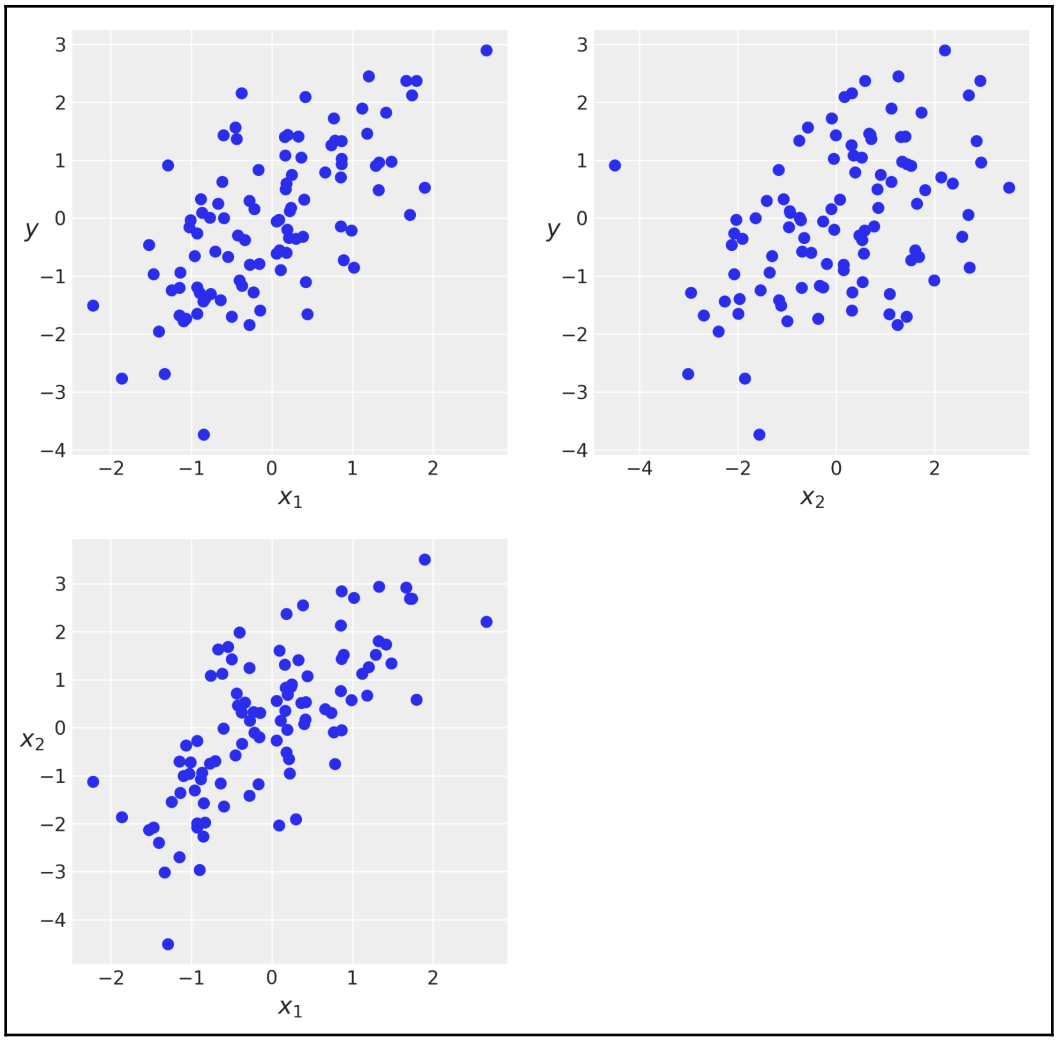

95% Credible Intervals
α__0
α__1
α__2
α__3
α__4
α__5
α__6
α__7
β__0
β__1
β__2
β__3
β__4
β__5
β__6
β__7
0
1
2
3
4
5
6
7

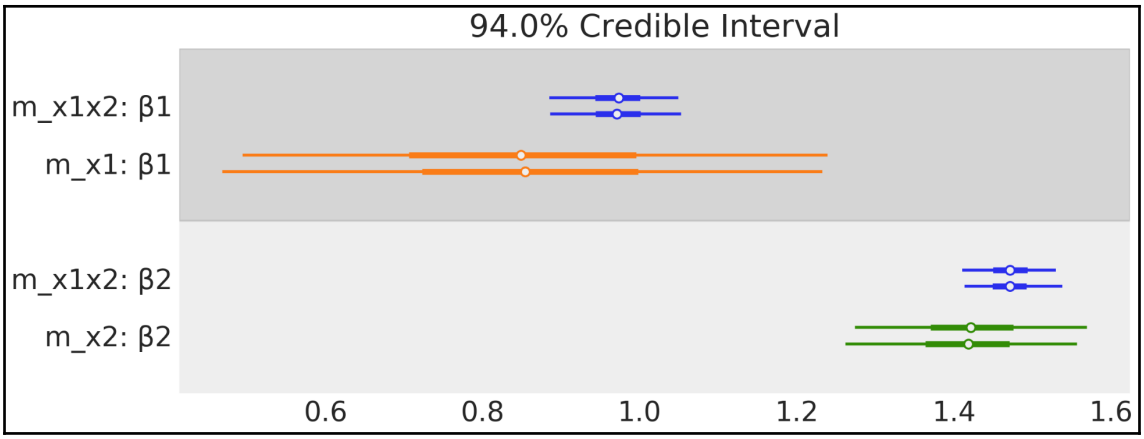

y_0
y_1
y_2
y_3
y_4
y_5
y_6
y_7
x_0
x_1
x_2
x_3
x_4
x_5
x_6
x_7
8
10
12
14
7.5
10.0
12.5

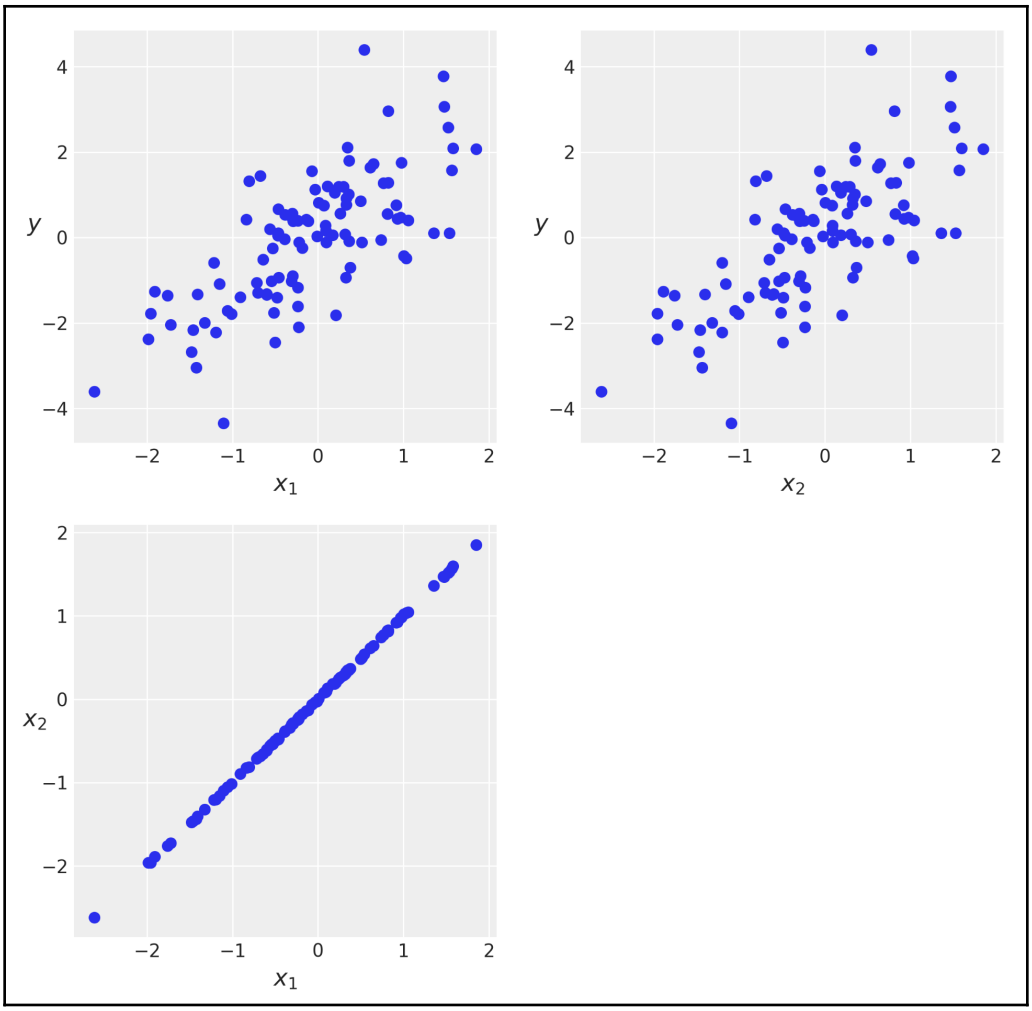

9
8
7
y
6
5
4
3
−4
−2
0
2
4
x

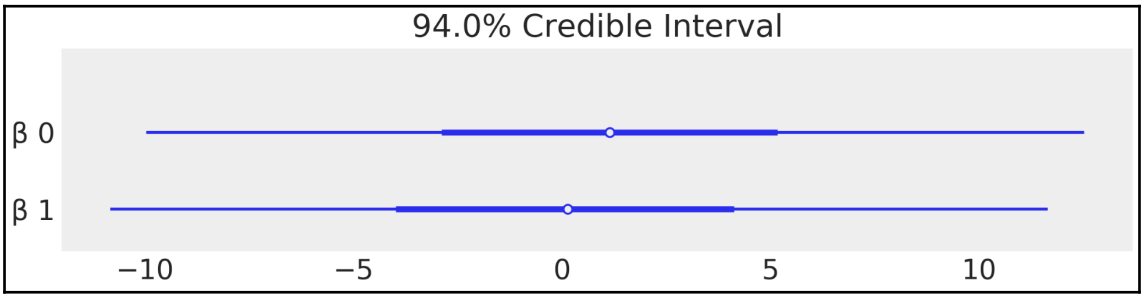

9
8
7
6
5
4
3
2
1
y
−6
−4
−2
0
2
4
6
x

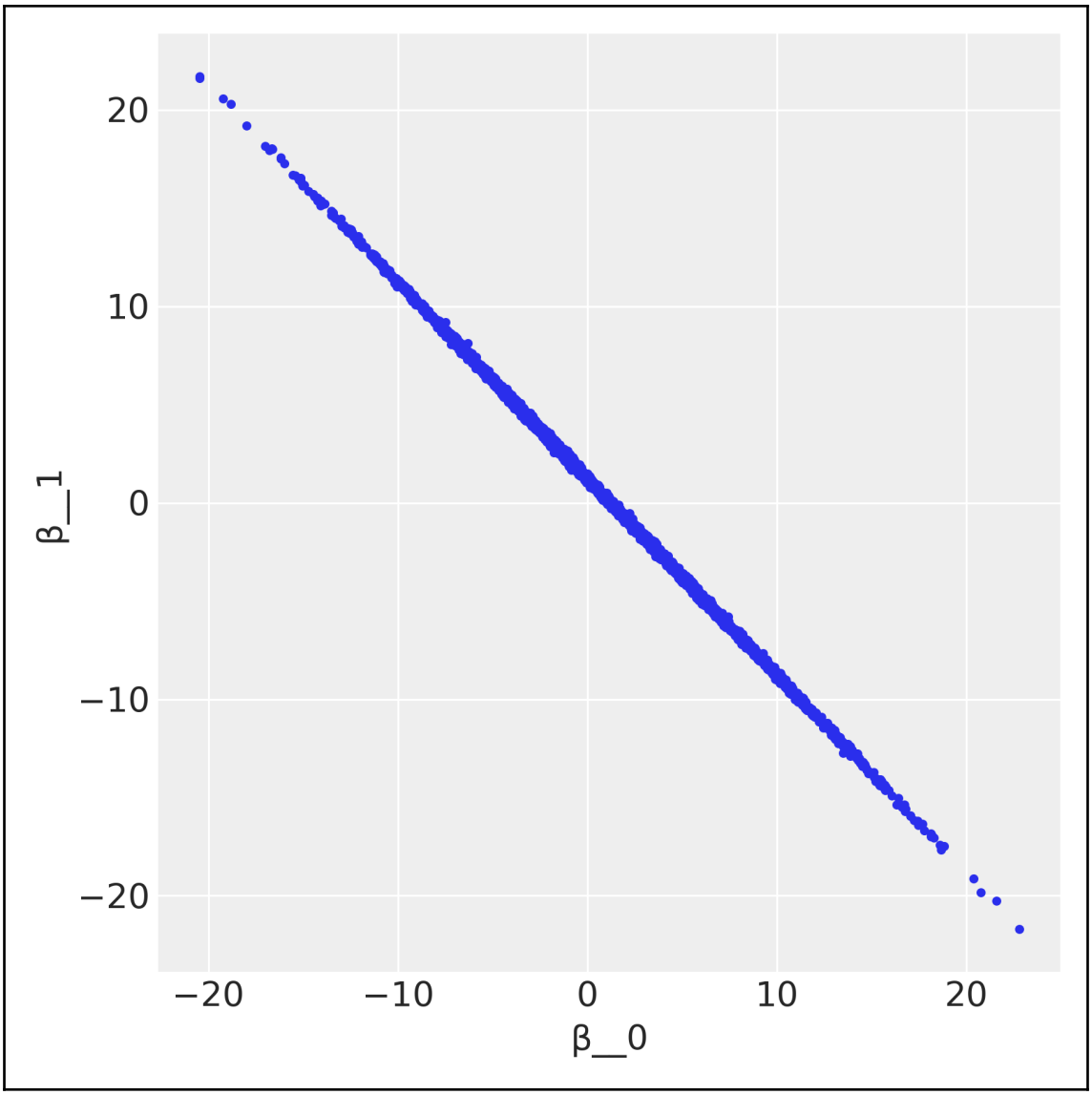

20
18
16
14
12
10
8
$y$
−3 −2 −1 0 1 2
$x_0$
20
18
16
14
12
10
8
$y$
−4 −2 0 2 4
$x_1$
4
3
2
1
0
−1
−2
−3
−4
$x_1$
−3 −2 −1 0 1 2
$x_0$

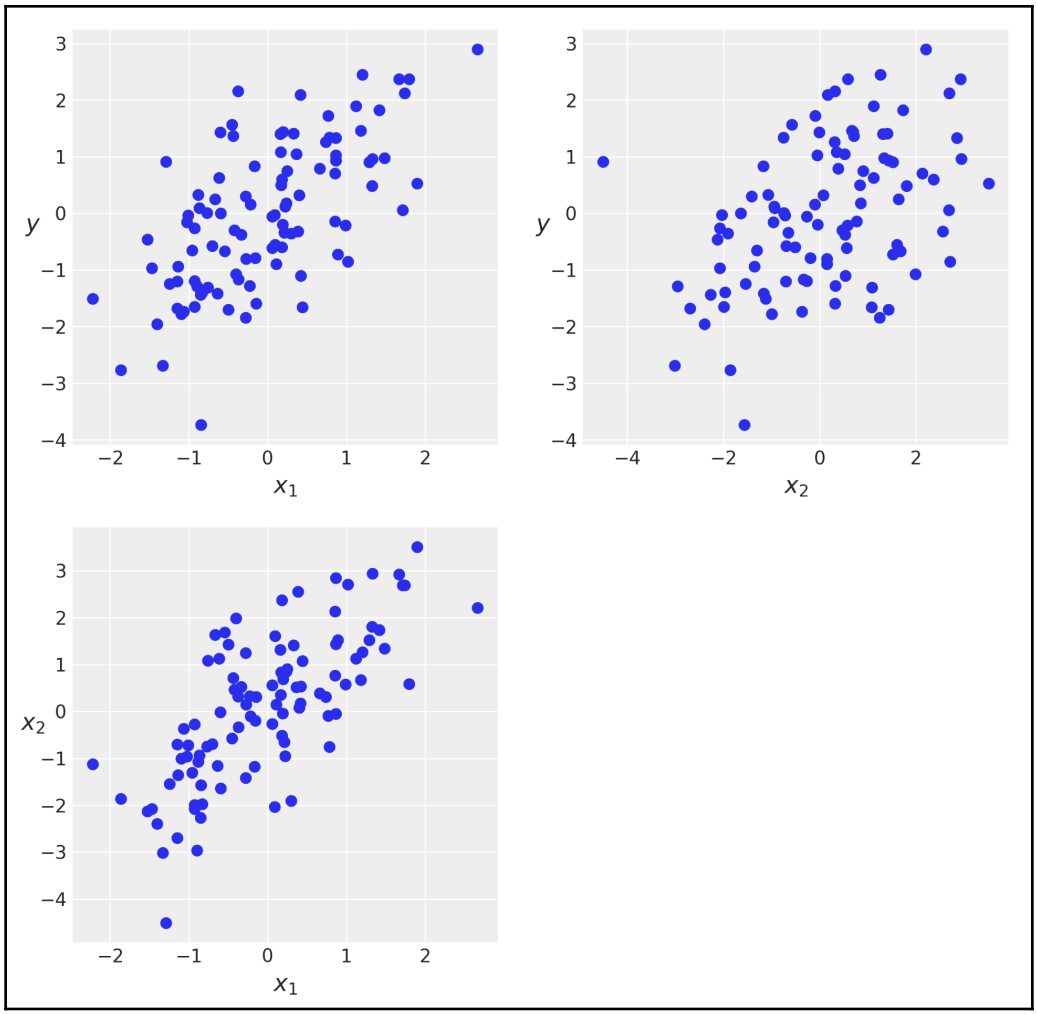

y
$x_1$
y
$x_2$
$x_2$
$x_1$

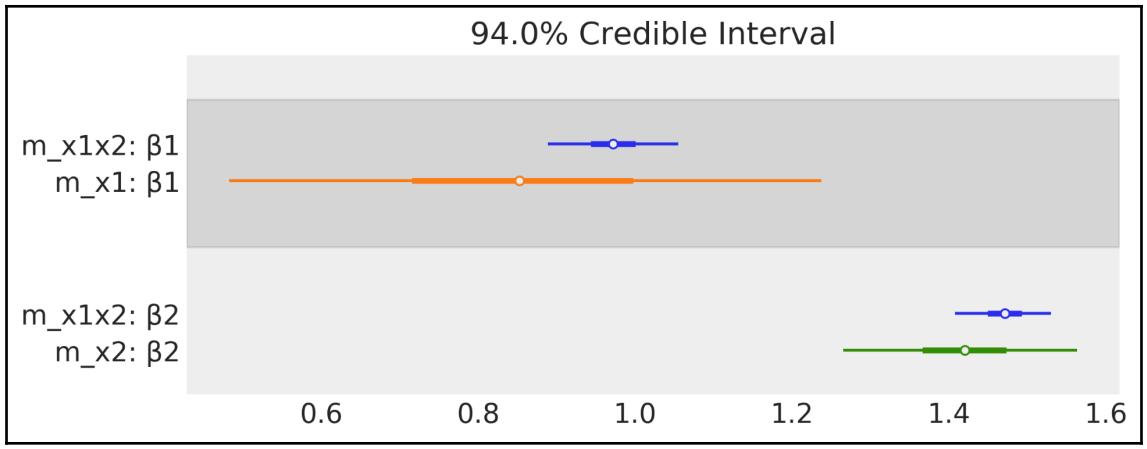

94.0% Credible Interval
m_x1x2: β1
m_x1: β1
m_x1x2: β2
m_x2: β2
0.6
0.8
1.0
1.2
1.4
1.6

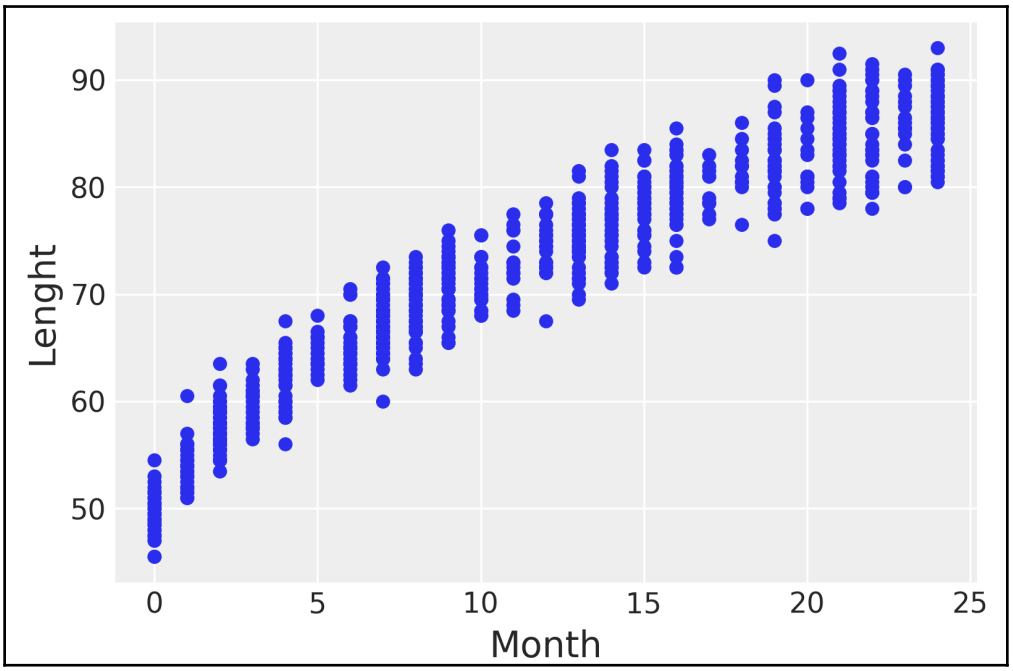

y
$x_1$
y
$x_2$
$x_2$
$x_1$

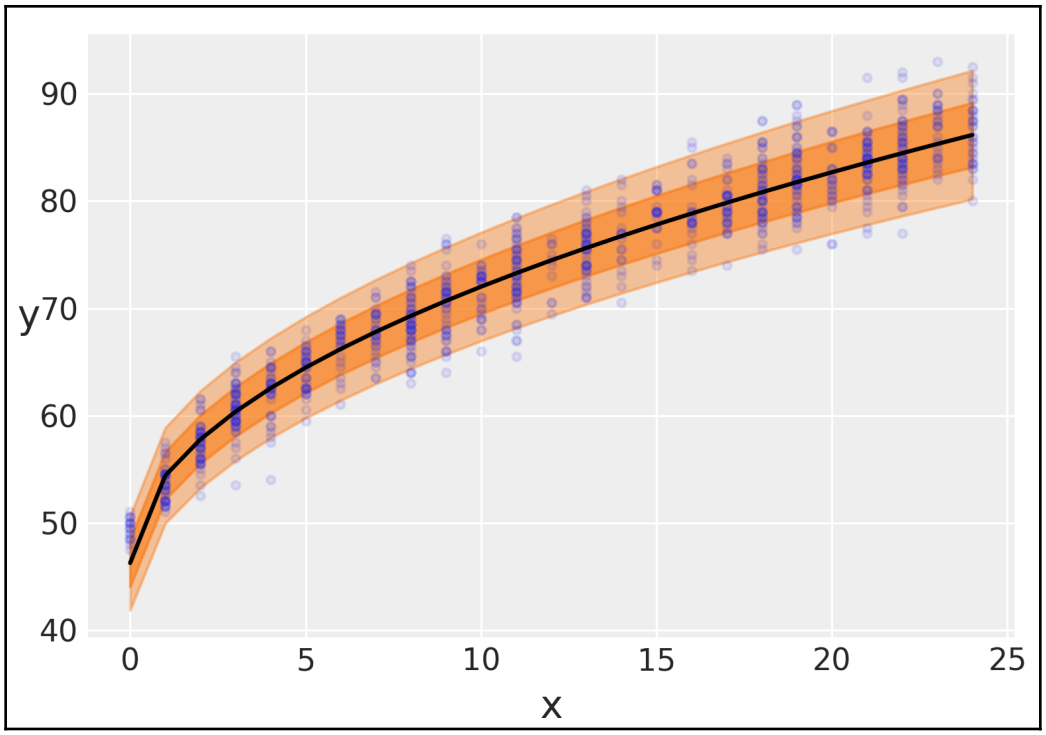

94.0% Credible Interval
β 0
β 1
−10
−5
0
5
10

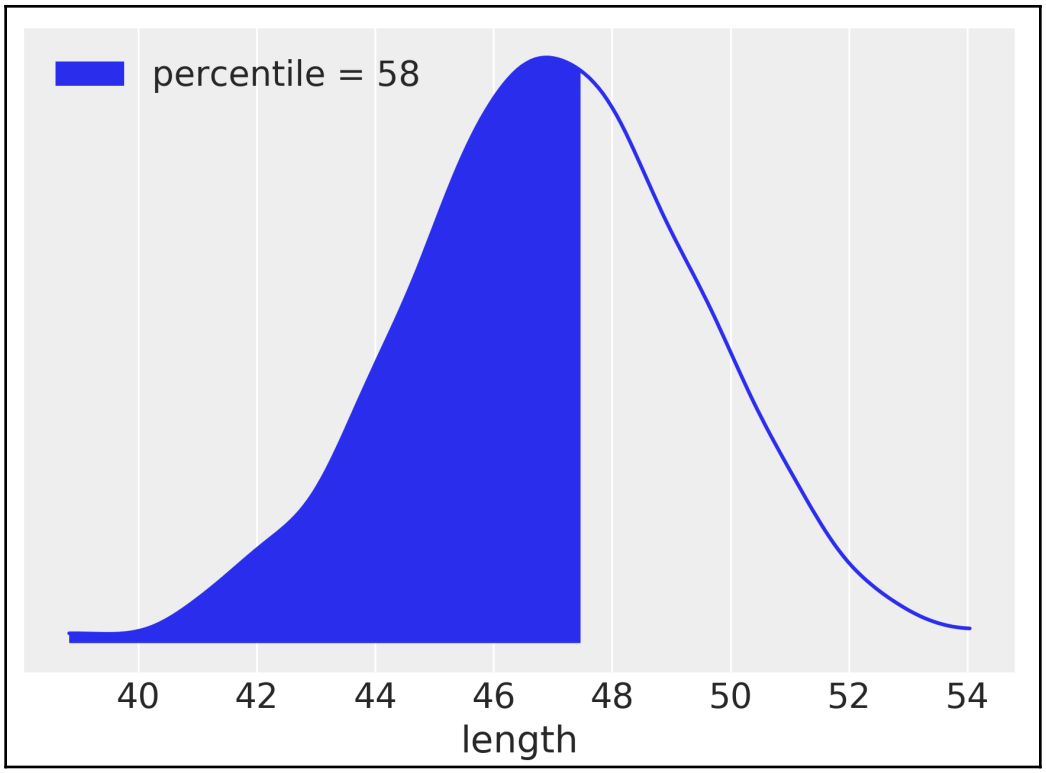

20
10
0
−10
−20
β__1
−20
−10
0
10
20
β__0


y
x1
y
x2
x2
x1

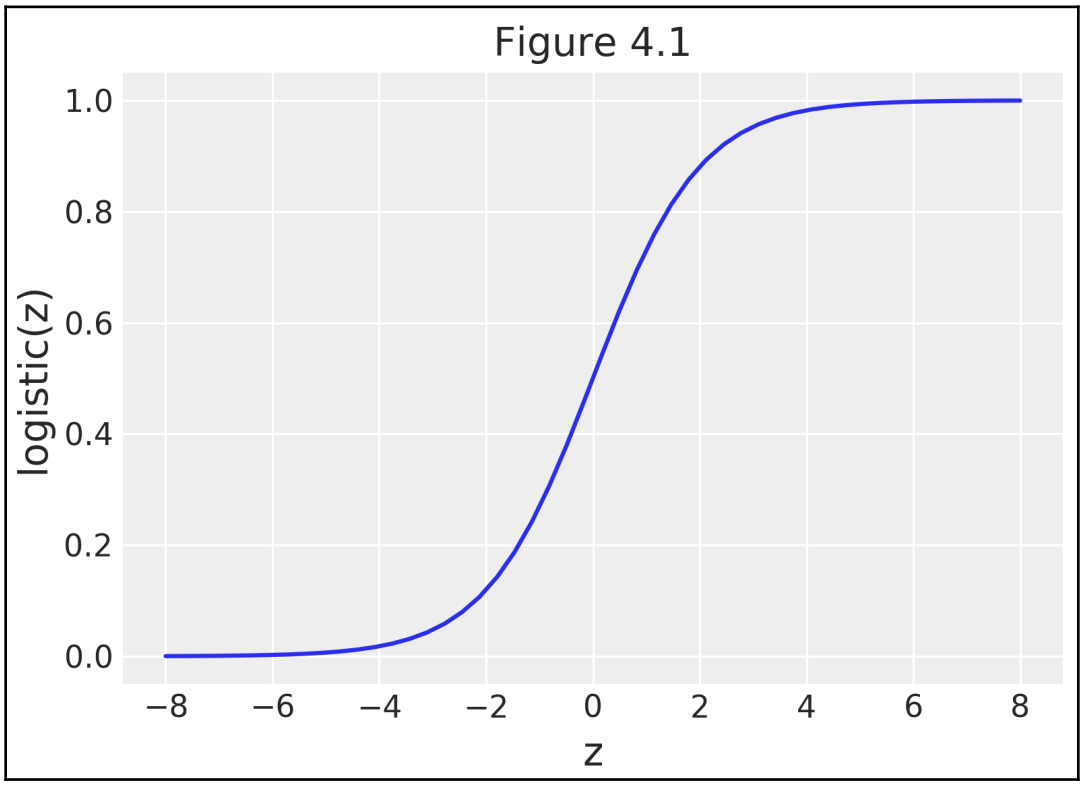

94.0% Credible Interval
m_x1x2: β1
m_x1: β1
m_x1x2: β2
m_x2: β2
0.6
0.8
1.0
1.2
1.4
1.6


Lenght
90
80
70
60
50
0
5
10
15
20
25
Month


90
80
y70
60
50
40
0
5
10
15
20
25
x


percentile = 58
40
42
44
46
48
50
52
54
length

# Chapter 4: Generalizing Linear Models



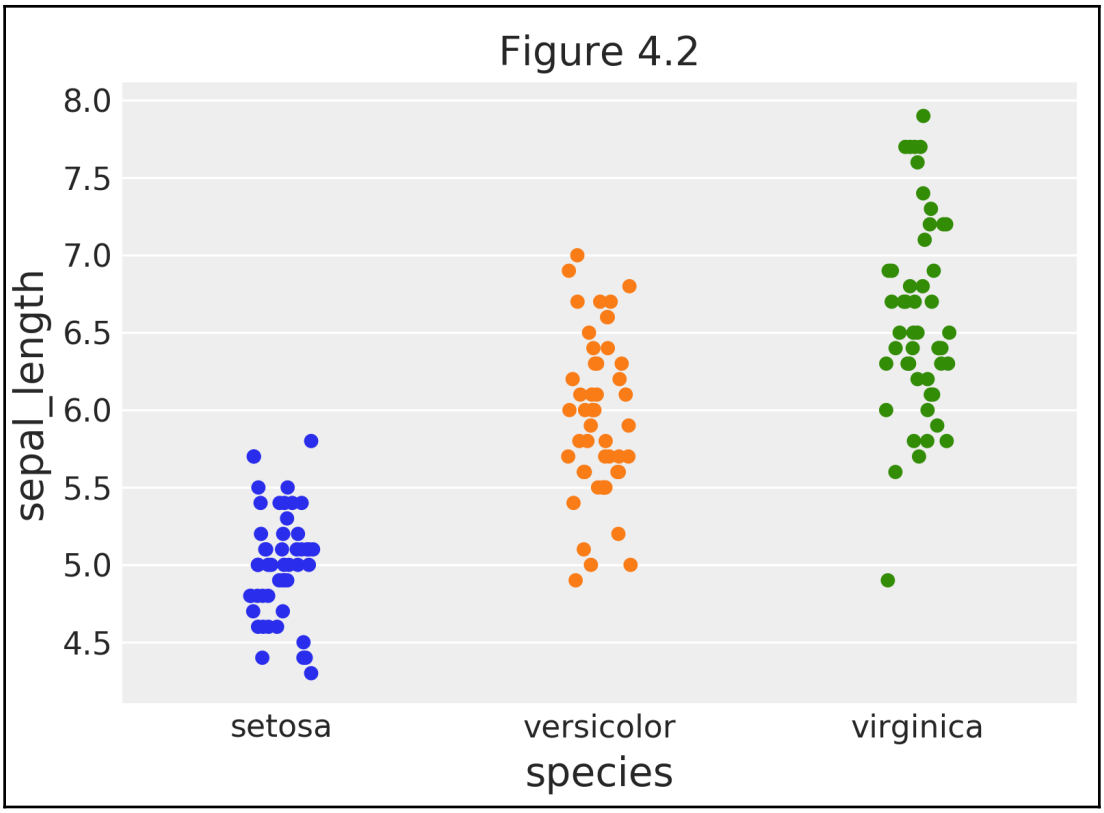


Figure 4.2
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
sepal_length
setosa
versicolor
virginica
species

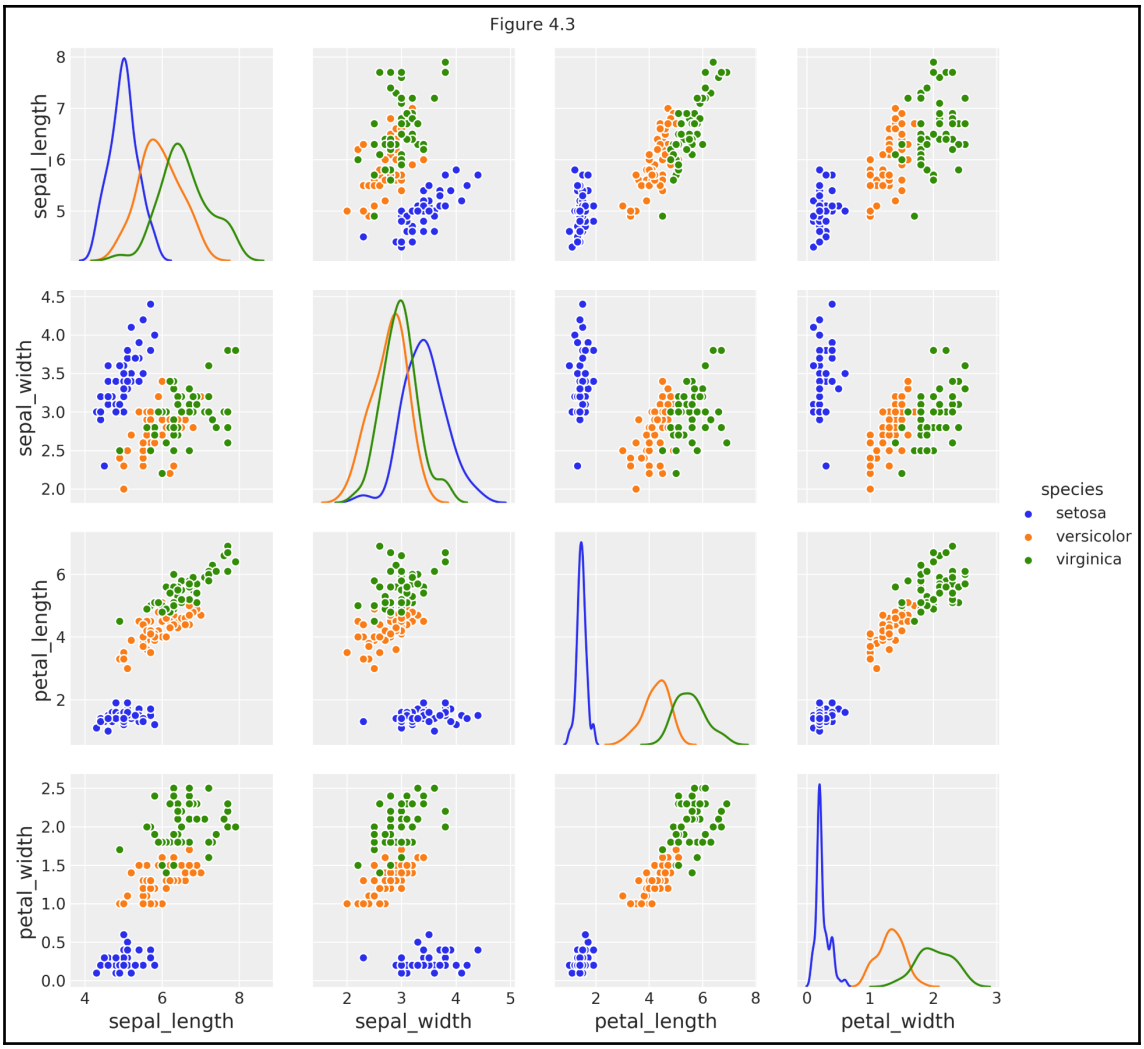


Figure 4.3
sepal_length
sepal_width
petal_length
petal_width
species
setosa
versicolor
virginica

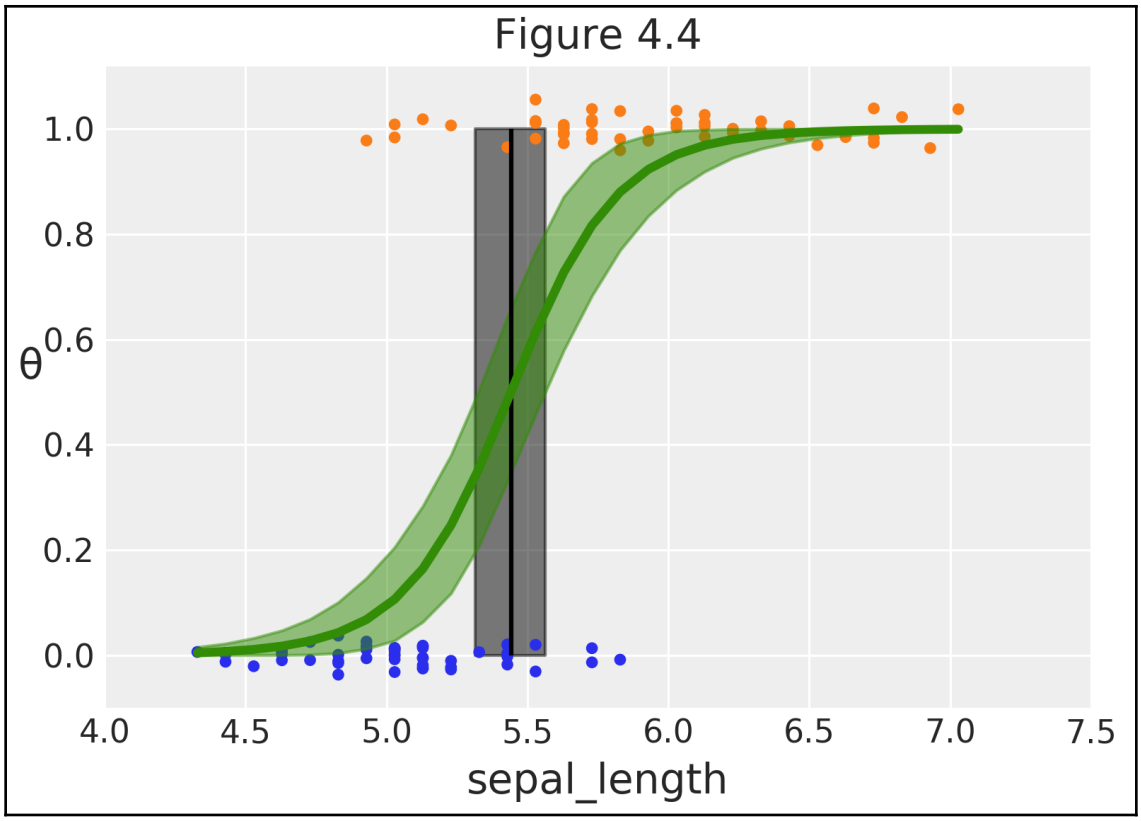


Figure 4.4
1.0
0.8
0.6
0.4
0.2
0.0
θ
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
sepal_length

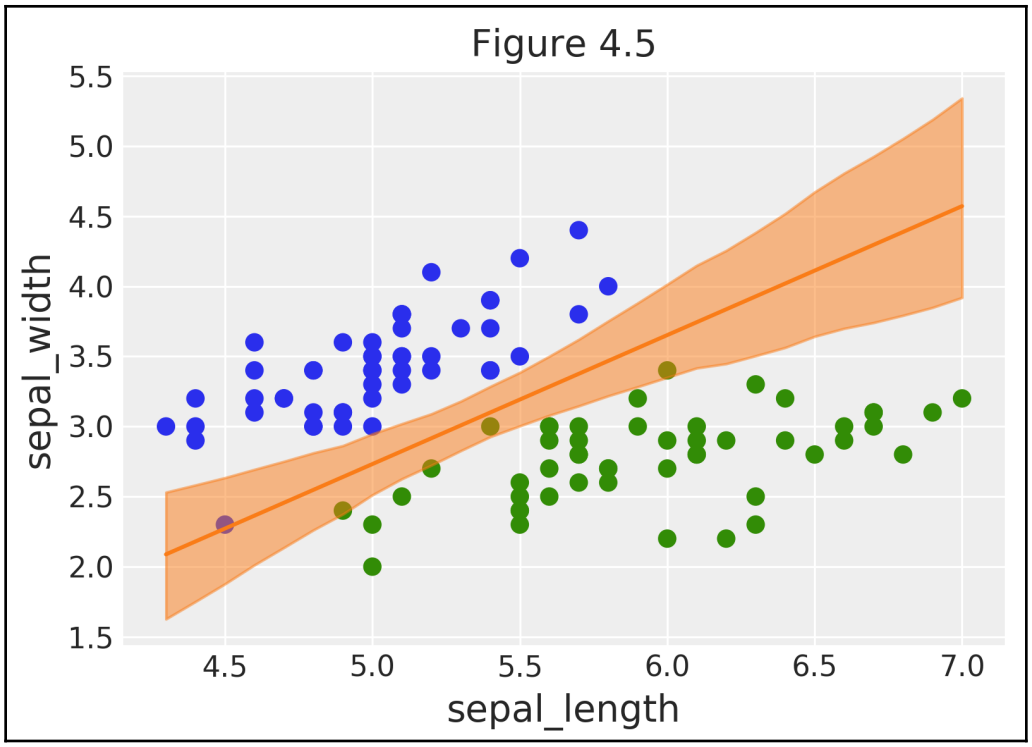


Figure 4.5
sepal_width
sepal_length
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
4.5
5.0
5.5
6.0
6.5
7.0

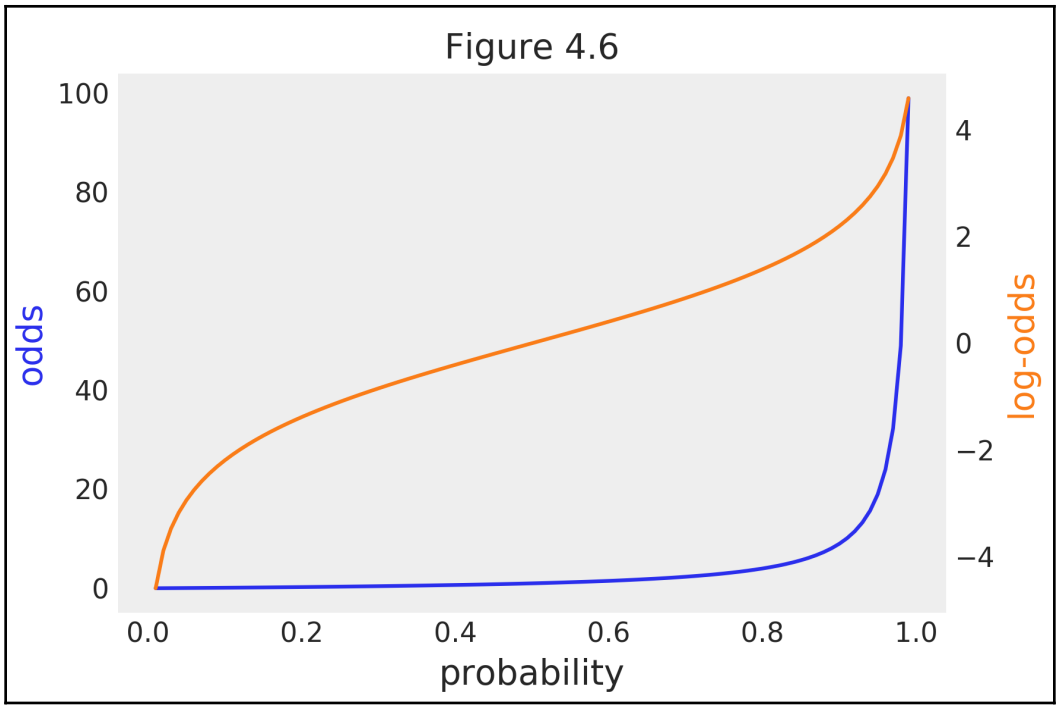


Figure 4.6
100
80
60
40
20
0
odds
4
2
0
−2
−4
log-odds
0.0
0.2
0.4
0.6
0.8
1.0
probability

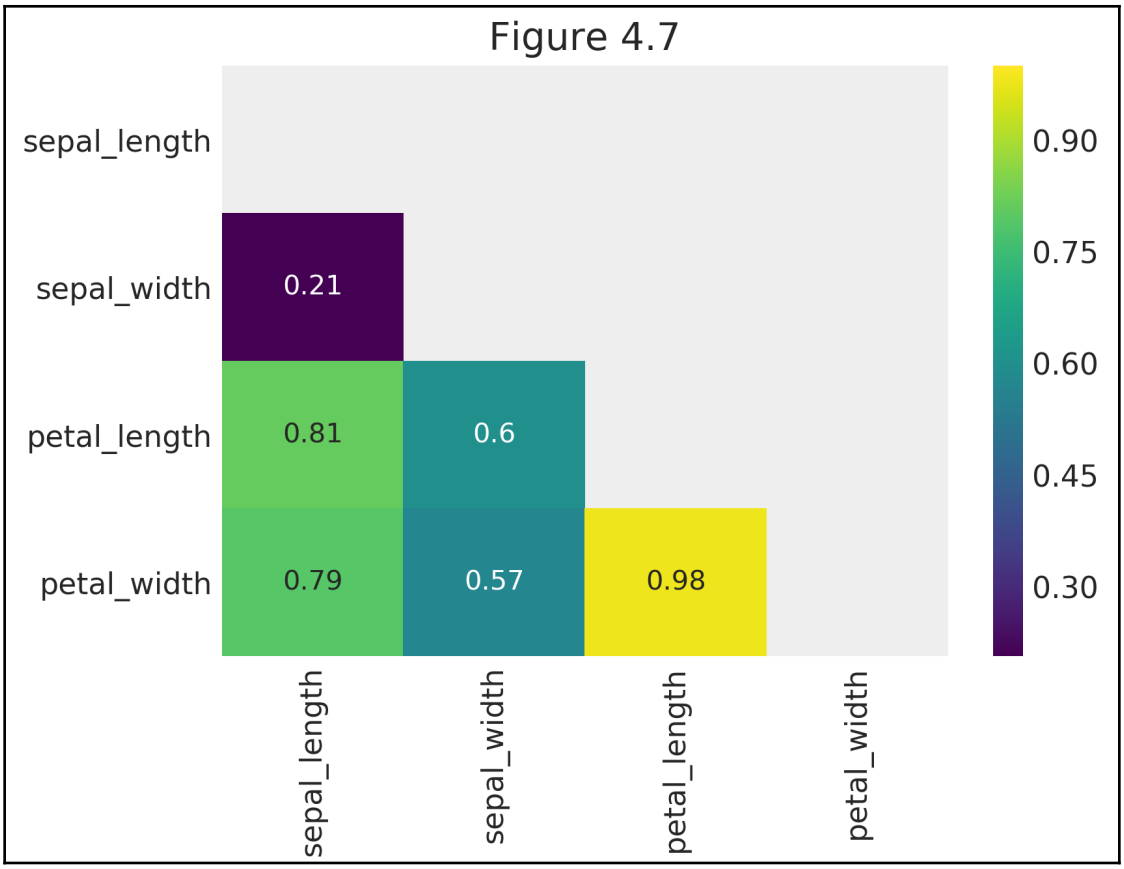


Figure 4.7
sepal_length
sepal_width
petal_length
petal_width
0.21
0.81
0.6
0.79
0.57
0.98
sepal_length
sepal_width
petal_length
petal_width
0.90
0.75
0.60
0.45
0.30

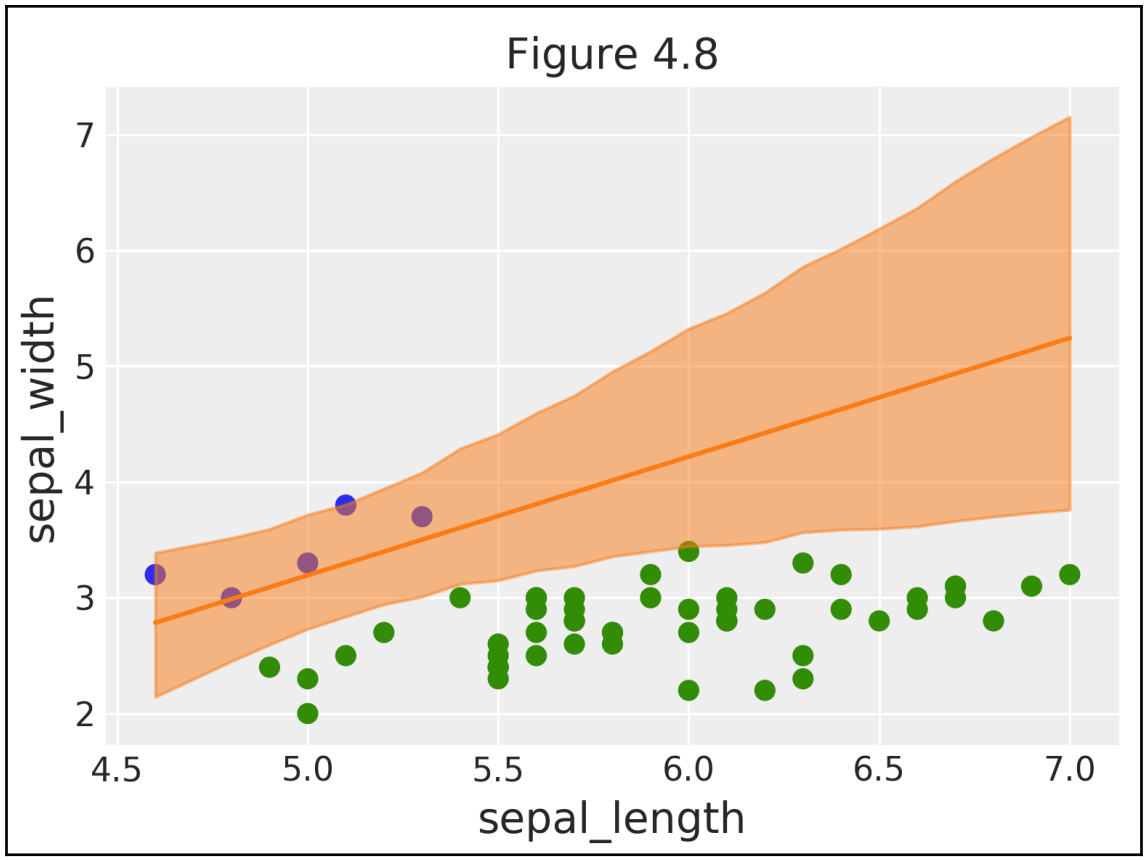


Figure 4.8
sepal_width
sepal_length
7
6
5
4
3
2
4.5
5.0
5.5
6.0
6.5
7.0

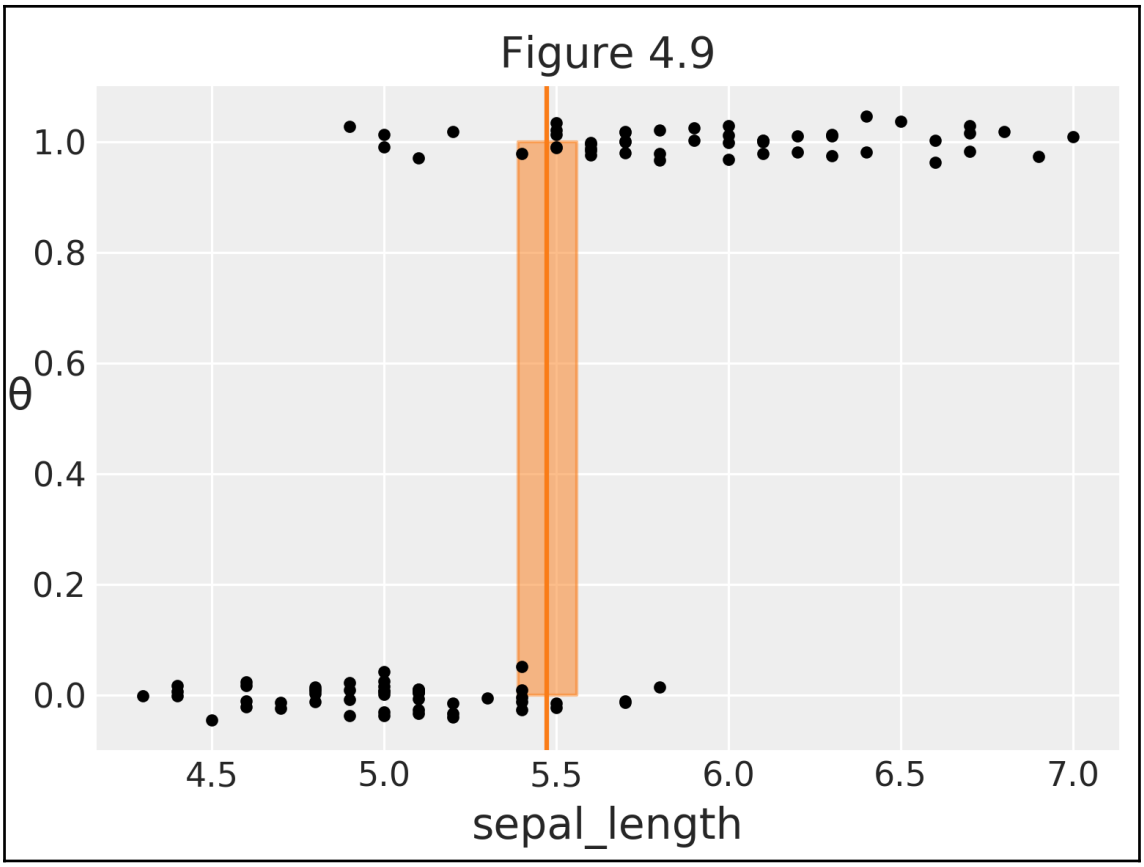


Figure 4.9
1.0
0.8
0.6
θ
0.4
0.2
0.0
4.5
5.0
5.5
6.0
6.5
7.0
sepal_length

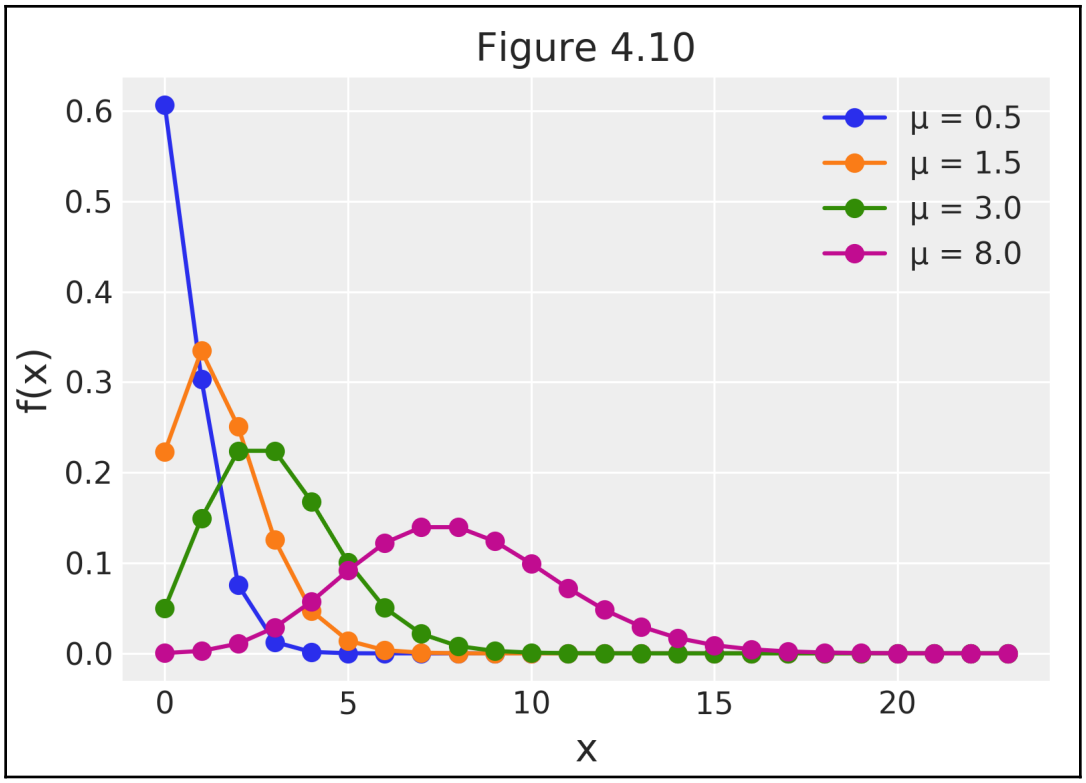


Figure 4.10
μ = 0.5
μ = 1.5
μ = 3.0
μ = 8.0
f(x)
x
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0
5
10
15
20

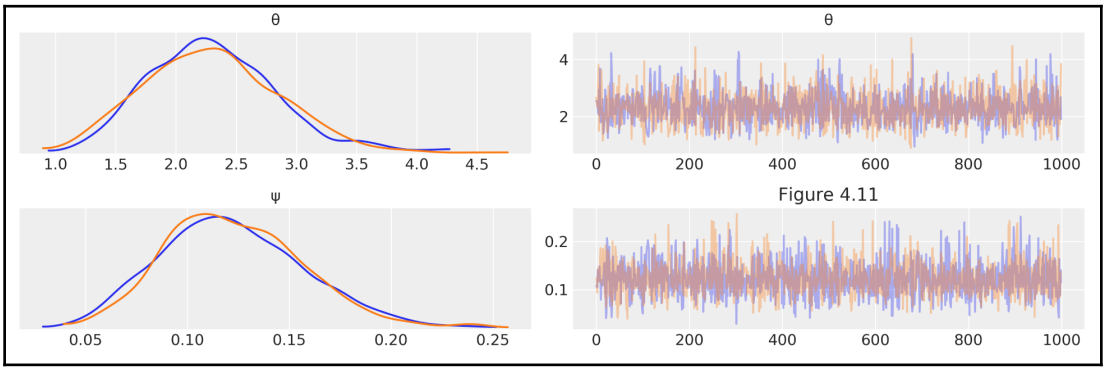

θ
ψ
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
0.05
0.10
0.15
0.20
0.25
0
200
400
600
800
1000
2
4
0.1
0.2

Figure 4.11

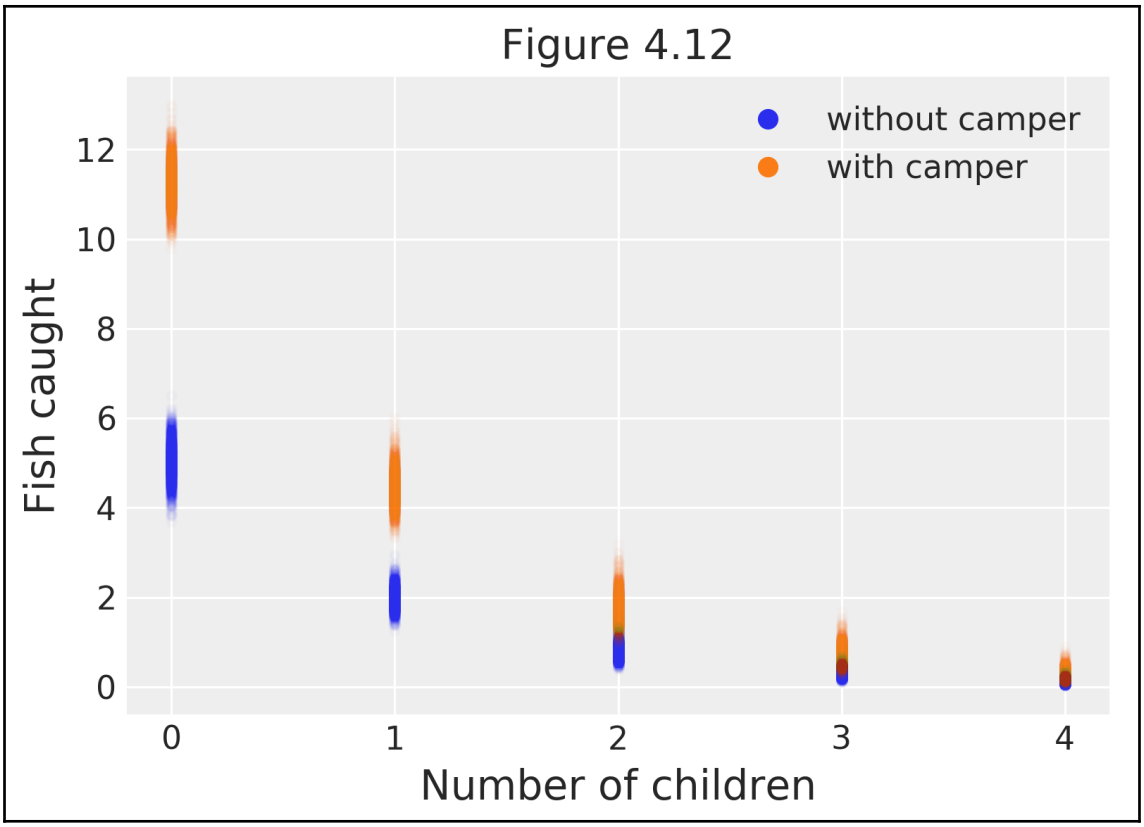


Figure 4.12
without camper
with camper
Fish caught
0
2
4
6
8
10
12
0
1
2
3
4
Number of children

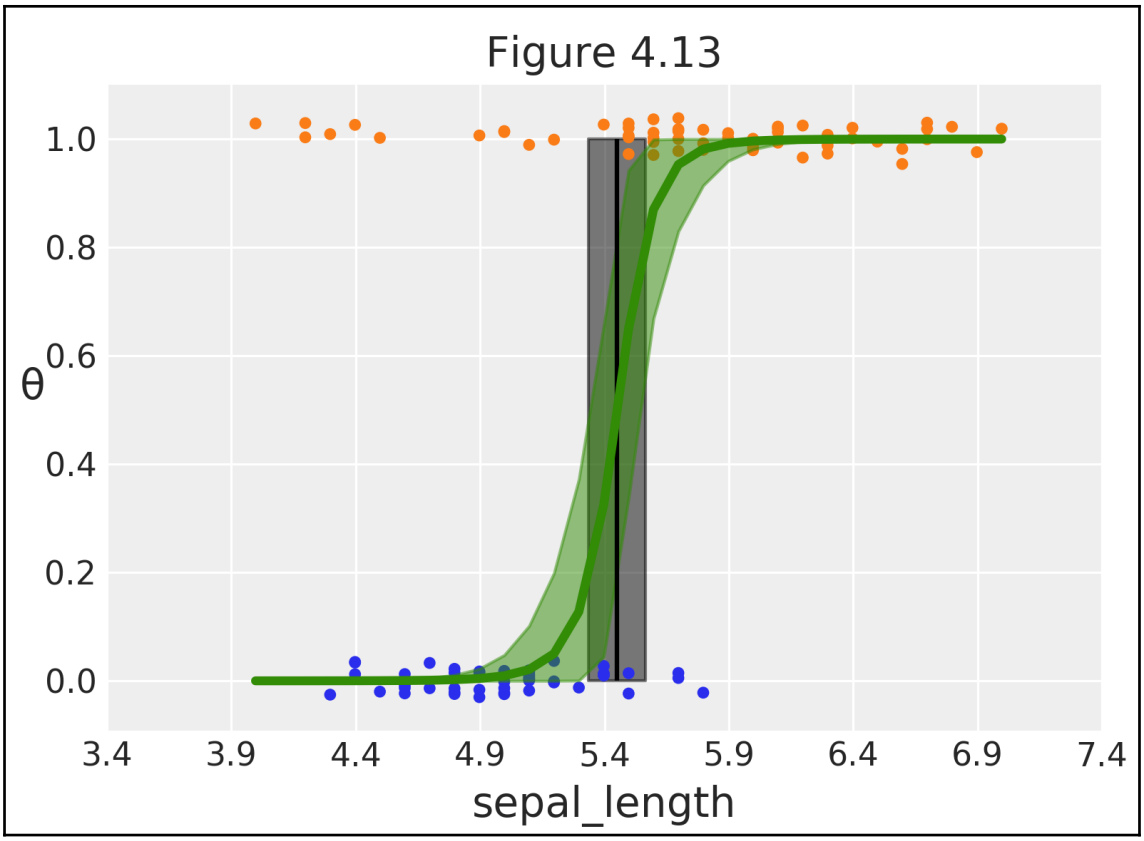
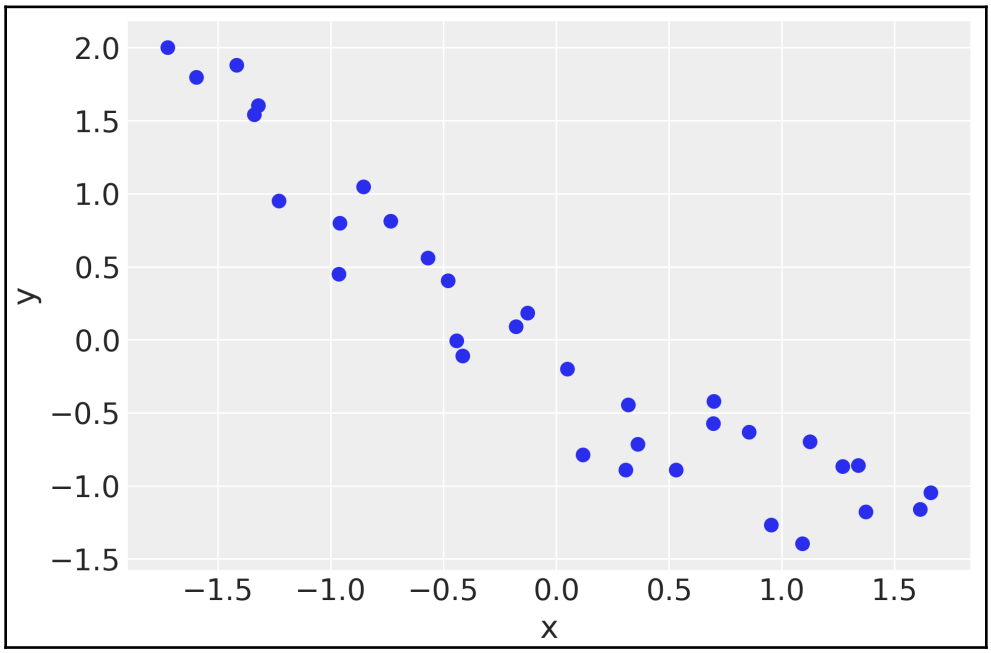
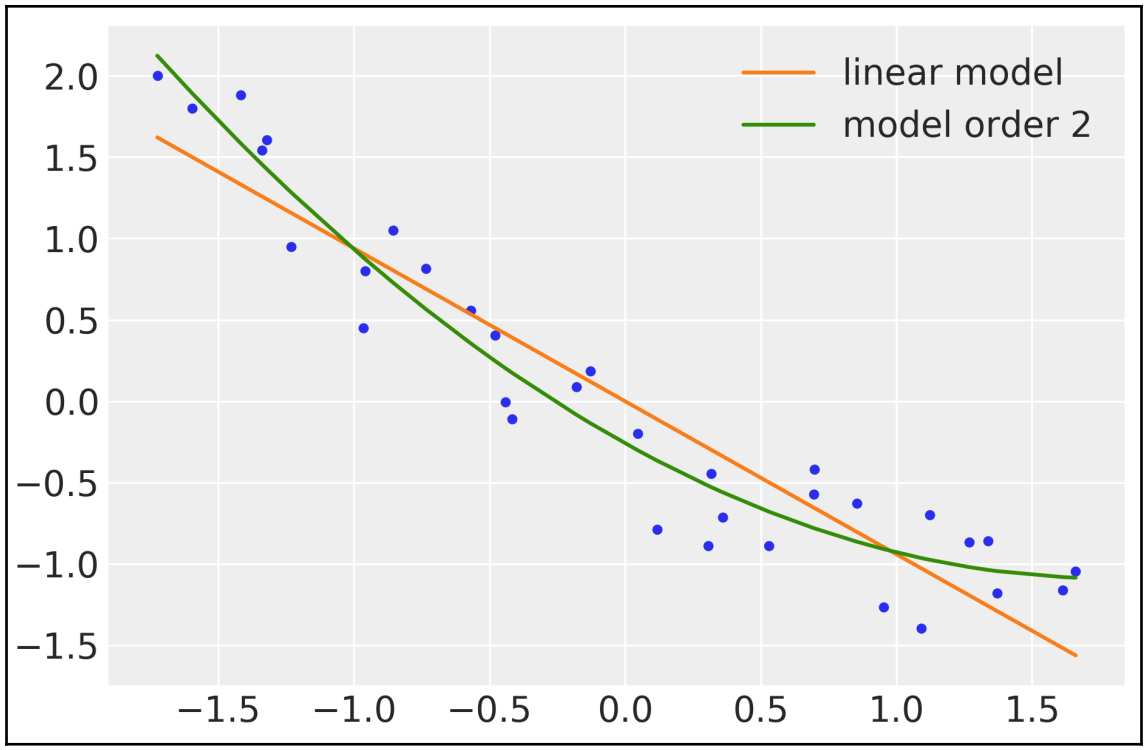
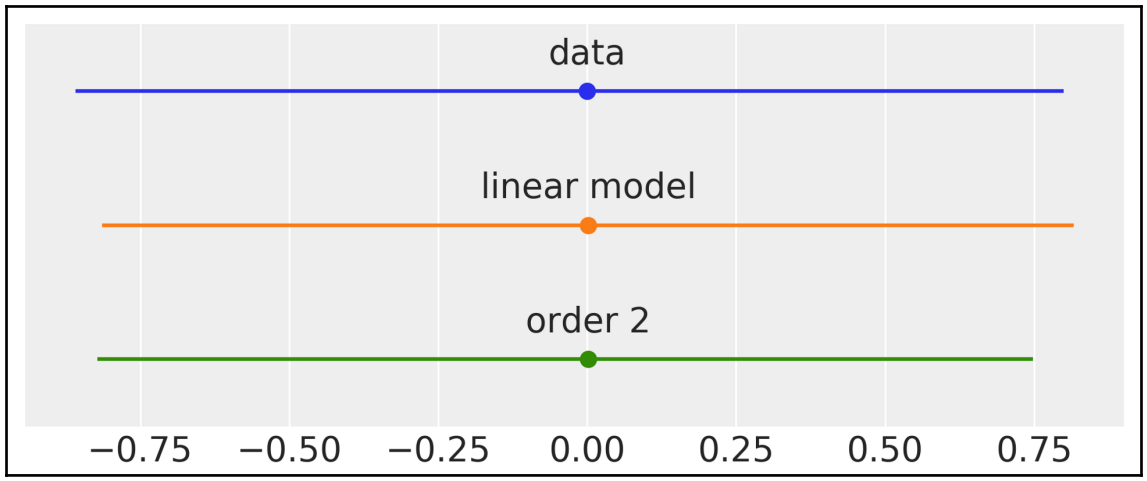


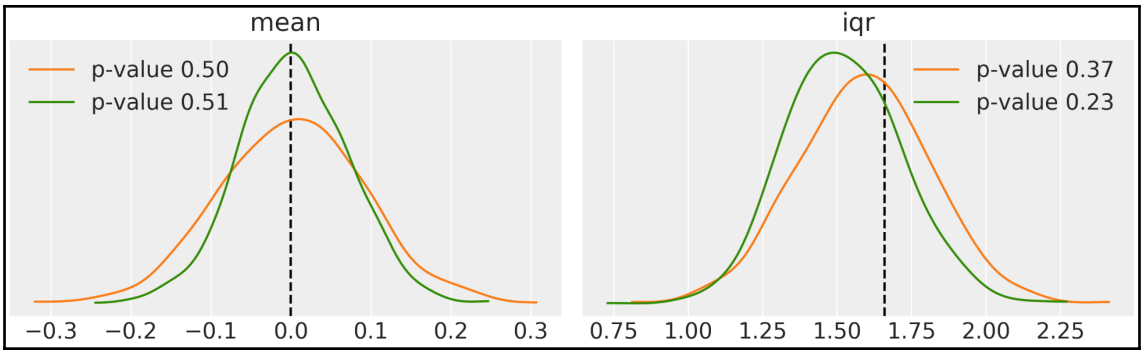
Figure 4.13
1.0
0.8
0.6
0.4
0.2
0.0
θ
3.4
3.9
4.4
4.9
5.4
5.9
6.4
6.9
7.4
sepal_length

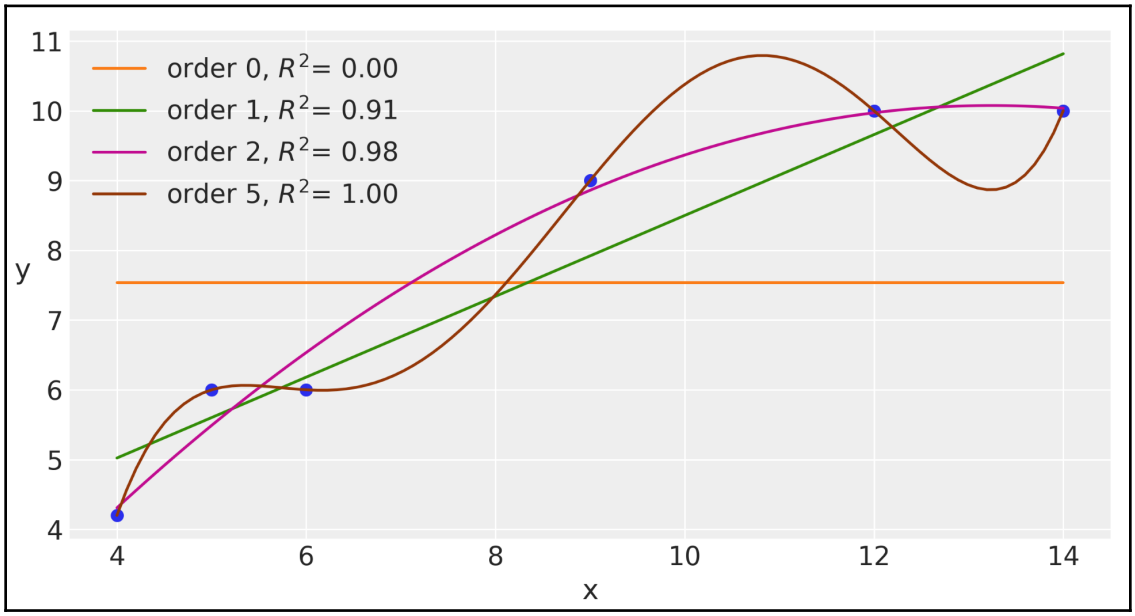
# Chapter 5: Model Comparison

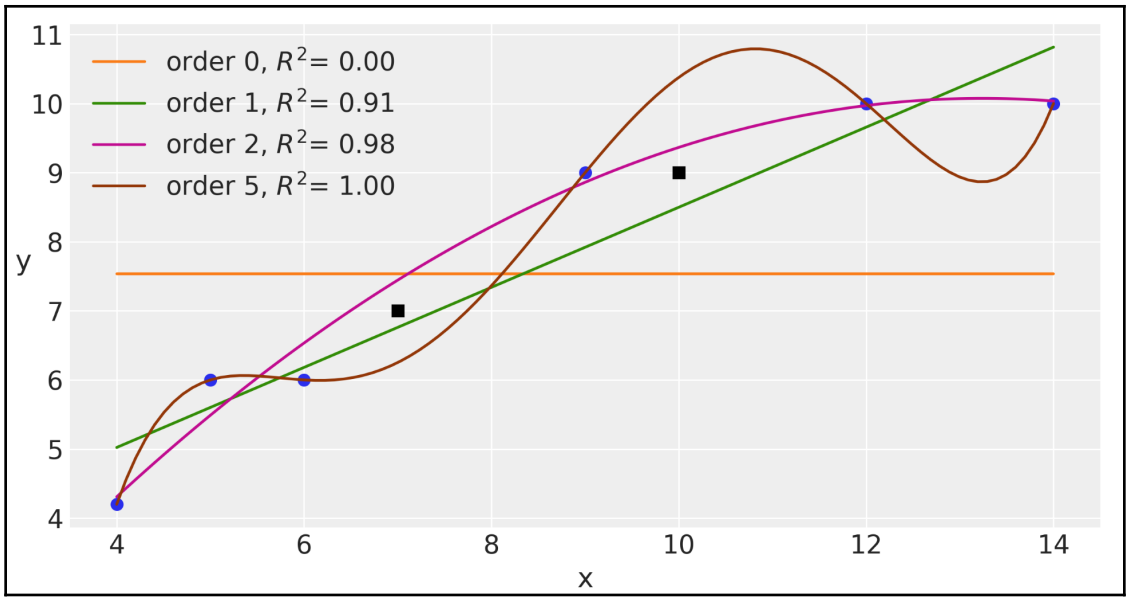



linear model
model order 2
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5

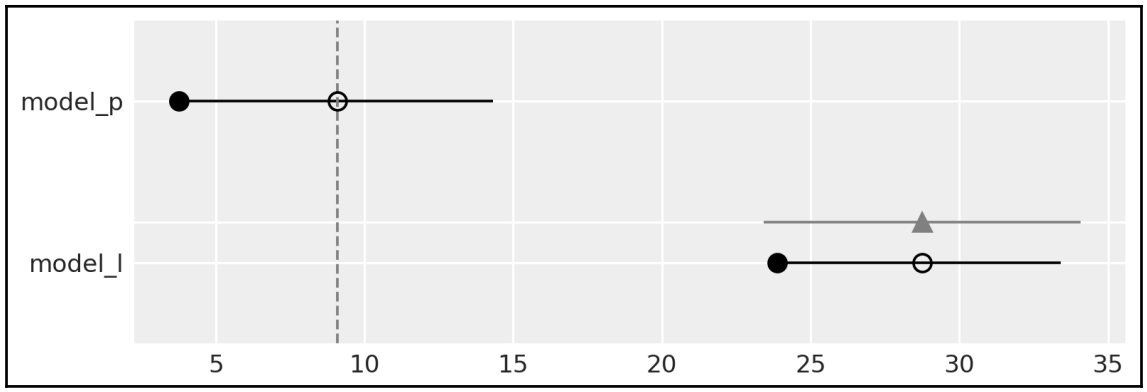

data
linear model
order 2
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75

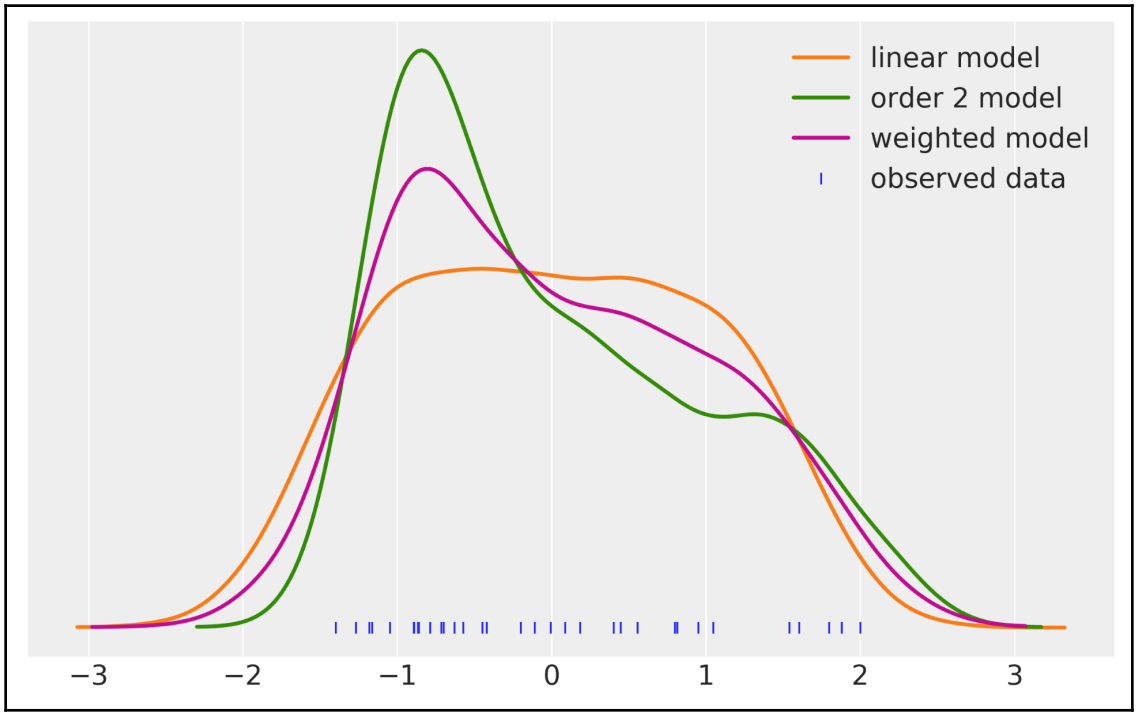

mean
p-value 0.50
p-value 0.51
−0.3 −0.2 −0.1 0.0 0.1 0.2 0.3
iqr
p-value 0.37
p-value 0.23
0.75 1.00 1.25 1.50 1.75 2.00 2.25

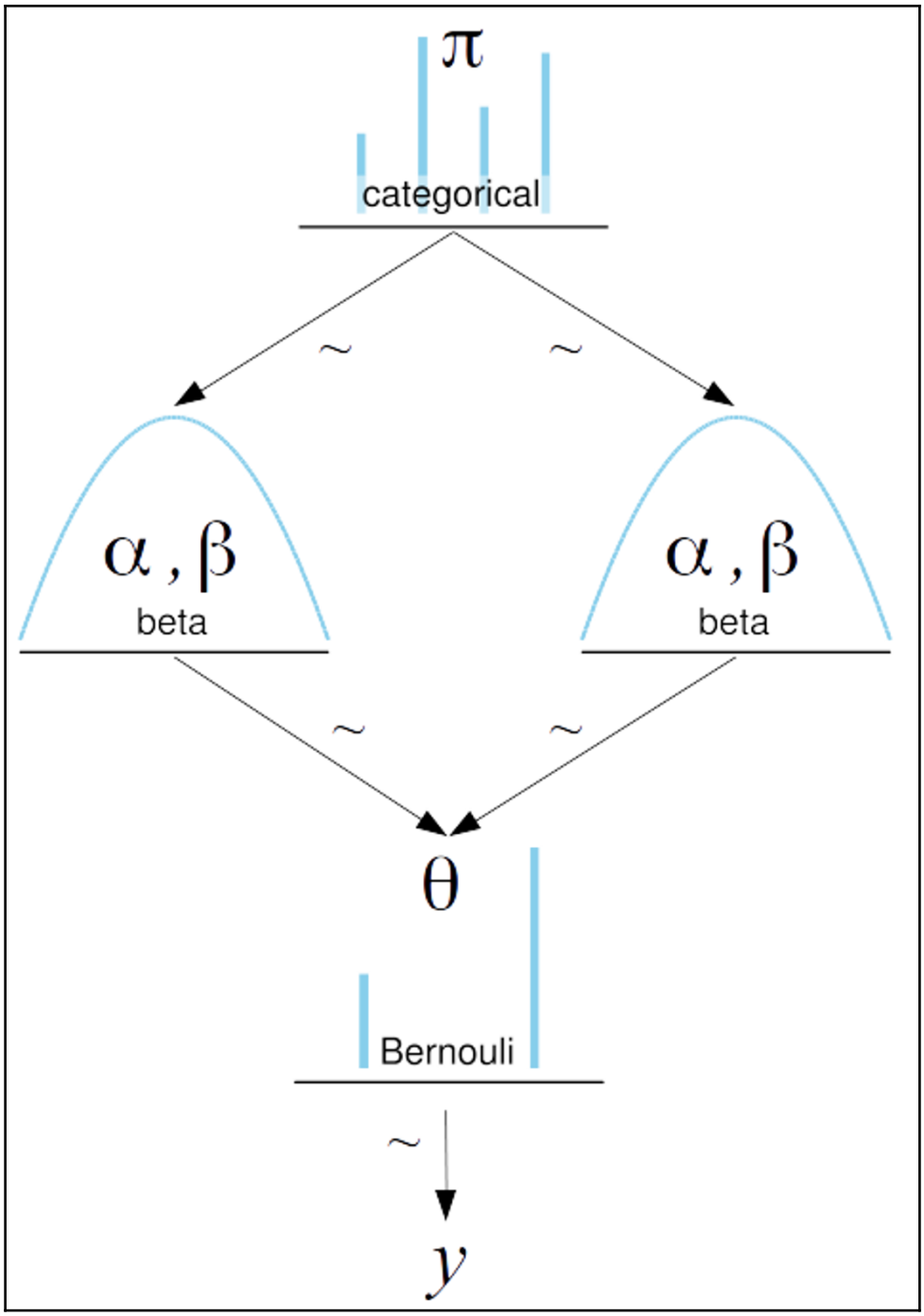

order 0, $R^2$= 0.00
order 1, $R^2$= 0.91
order 2, $R^2$= 0.98
order 5, $R^2$= 1.00
y
x
4 6 8 10 12 14
4 5 6 7 8 9 10 11

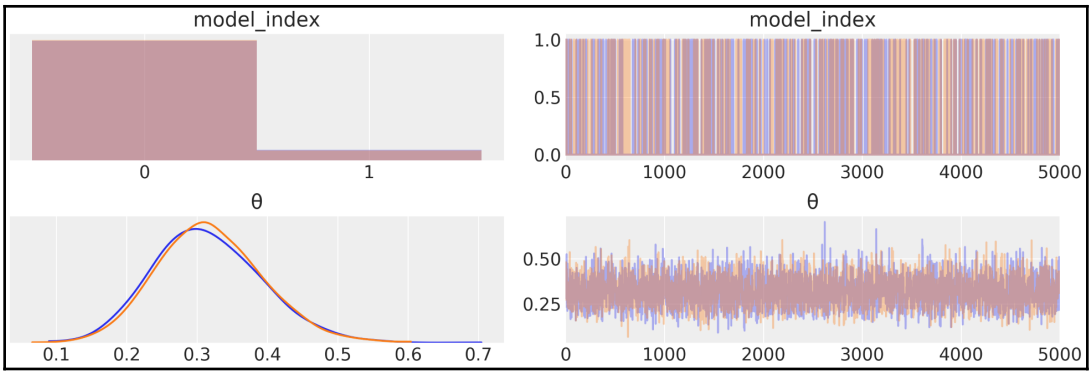

order 0, $R^2$= 0.00
order 1, $R^2$= 0.91
order 2, $R^2$= 0.98
order 5, $R^2$= 1.00
y
x
11
10
9
8
7
6
5
4
4
6
8
10
12
14

1st round
Test
Training
A1
2nd round
A2
3rd round
A3
Kth round
Ak

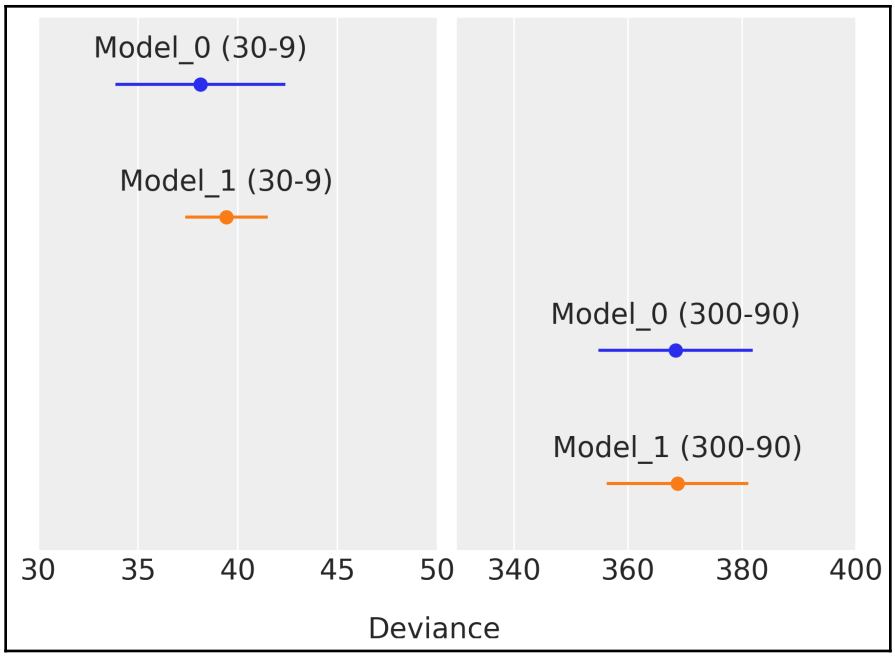

model_p
model_l
5
10
15
20
25
30
35

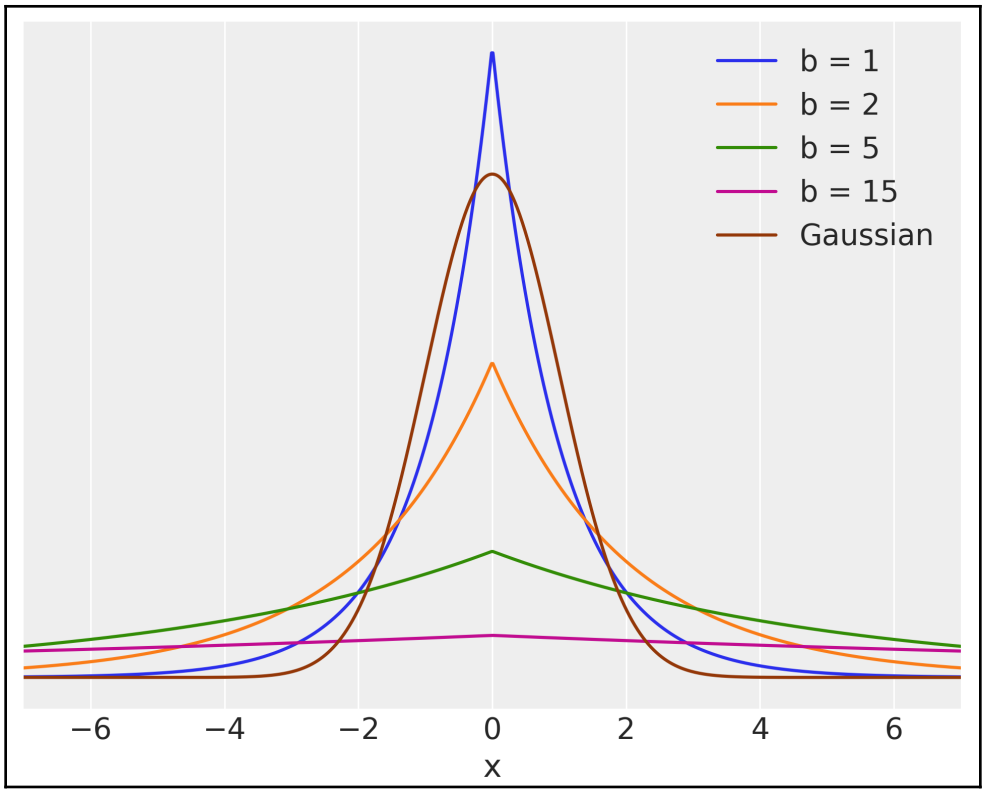

linear model
order 2 model
weighted model
observed data
−3
−2
−1
0
1
2
3

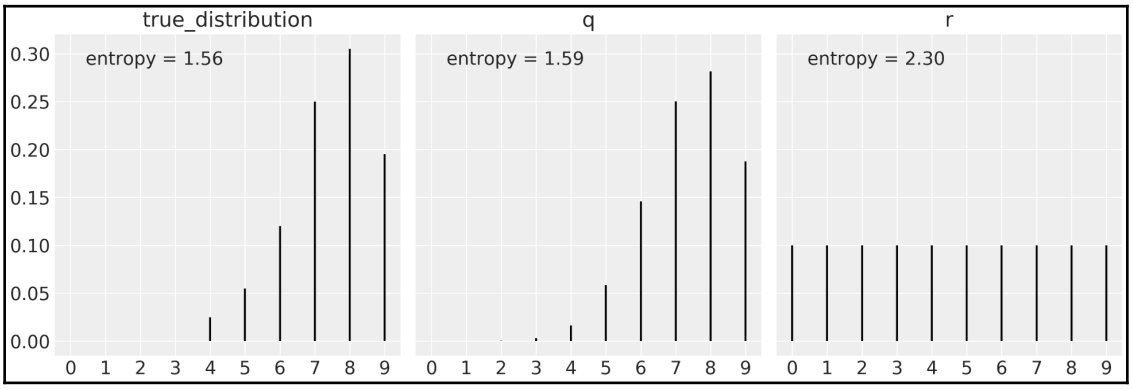

π
categorical
~
~
α , β
beta
α , β
beta
~
~
θ
Bernouli
~
y


model_index
0
1
model_index
1.0
0.5
0.0
0
1000
2000
3000
4000
5000
θ
0.1
0.2
0.3
0.4
0.5
0.6
0.7
θ
0.50
0.25
0
1000
2000
3000
4000
5000

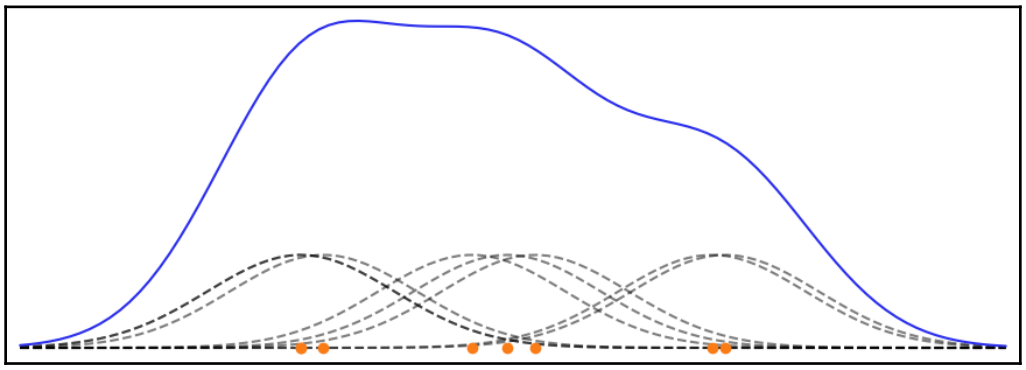
94.0% Credible Interval
Model_0 (30-9): θ
Model_1 (30-9): θ
Model_0 (300-90): θ
Model_1 (300-90): θ
0.2
0.3
0.4
0.5

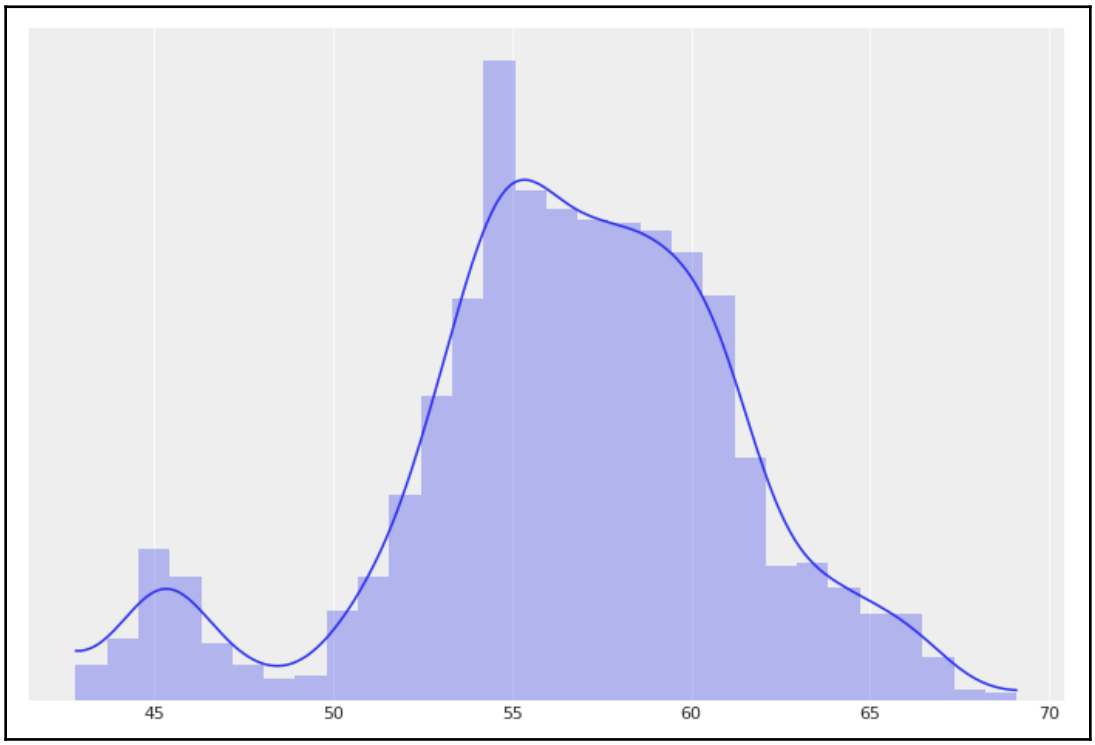

Model_0 (30-9)
Model_1 (30-9)
Model_0 (300-90)
Model_1 (300-90)
30
35
40
45
50
340
360
380
400
Deviance

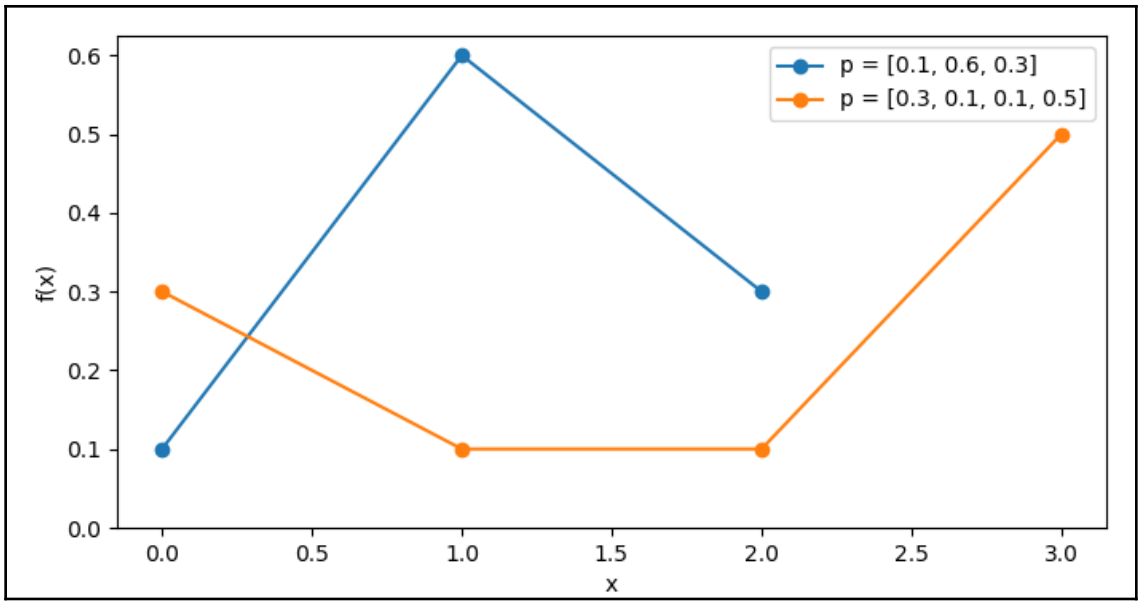

b = 1
b = 2
b = 5
b = 15
Gaussian
−6
−4
−2
0
2
4
6
x

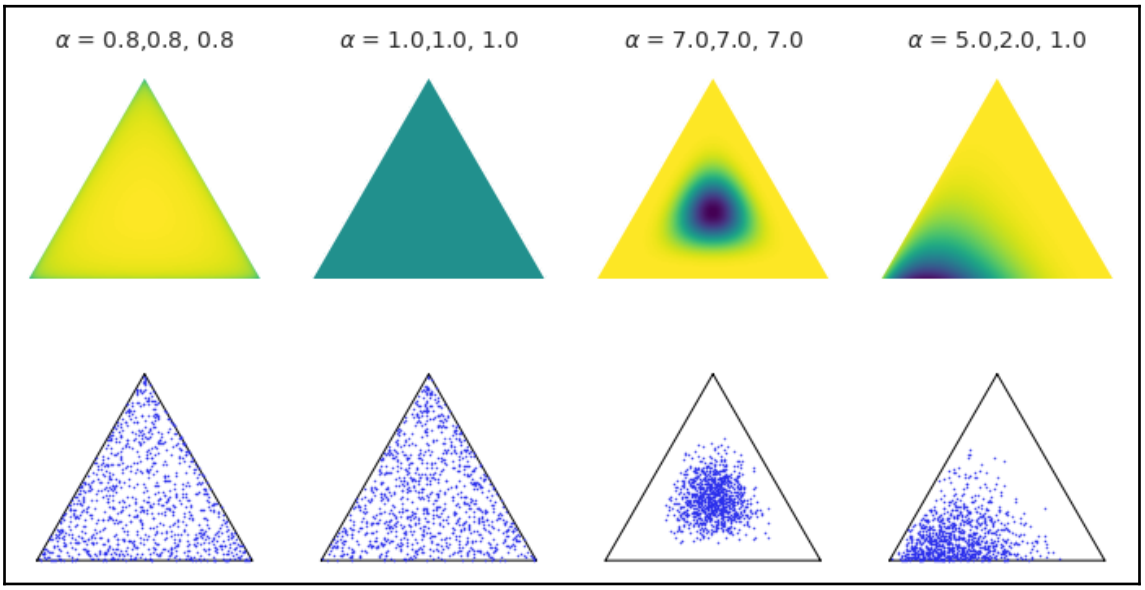

true_distribution
entropy = 1.56
q
entropy = 1.59
r
entropy = 2.30
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0 1 2 3 4 5 6 7 8 9

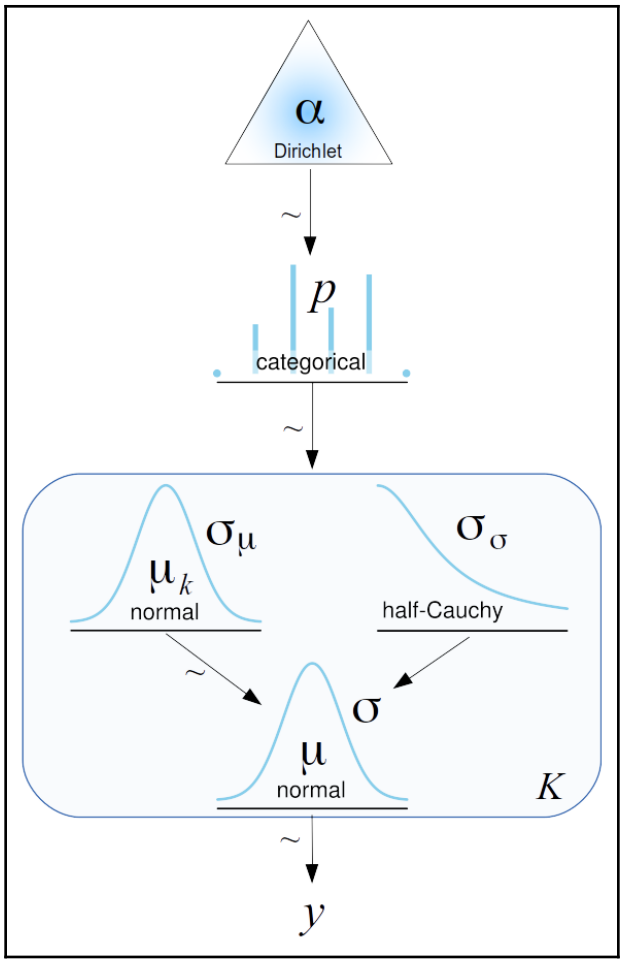
# Chapter 6: Mixture Models






p = [0.1, 0.6, 0.3]
p = [0.3, 0.1, 0.1, 0.5]
f(x)
x
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.0
0.5
1.0
1.5
2.0
2.5
3.0


α = 0.8,0.8, 0.8
α = 1.0,1.0, 1.0
α = 7.0,7.0, 7.0
α = 5.0,2.0, 1.0


α
Dirichlet
~
p
categorical
~
σμ
μk
normal
σσ
half-Cauchy
~
σ
μ
normal
K
~
y

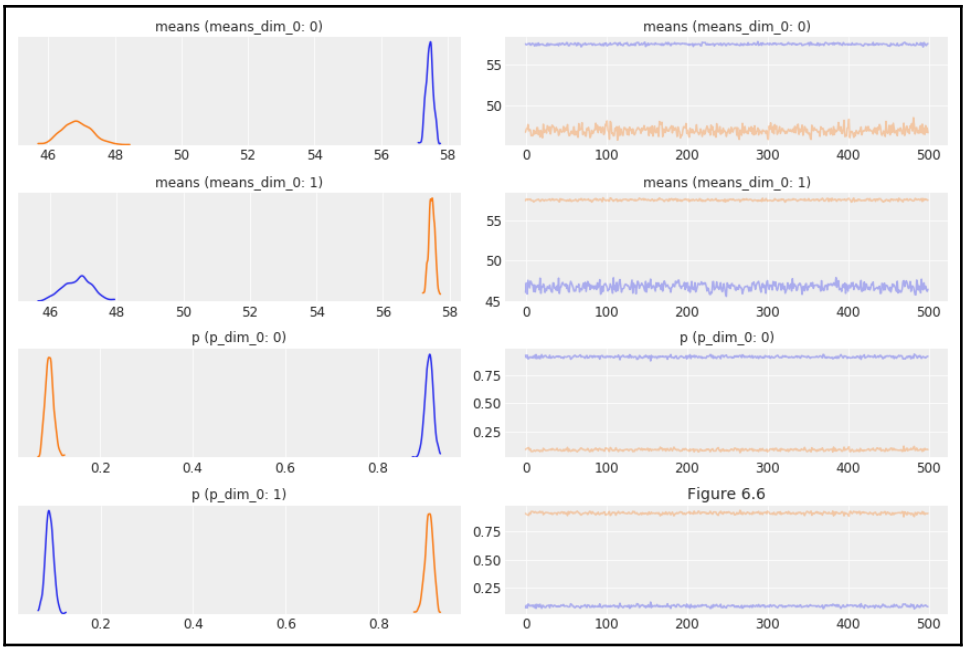
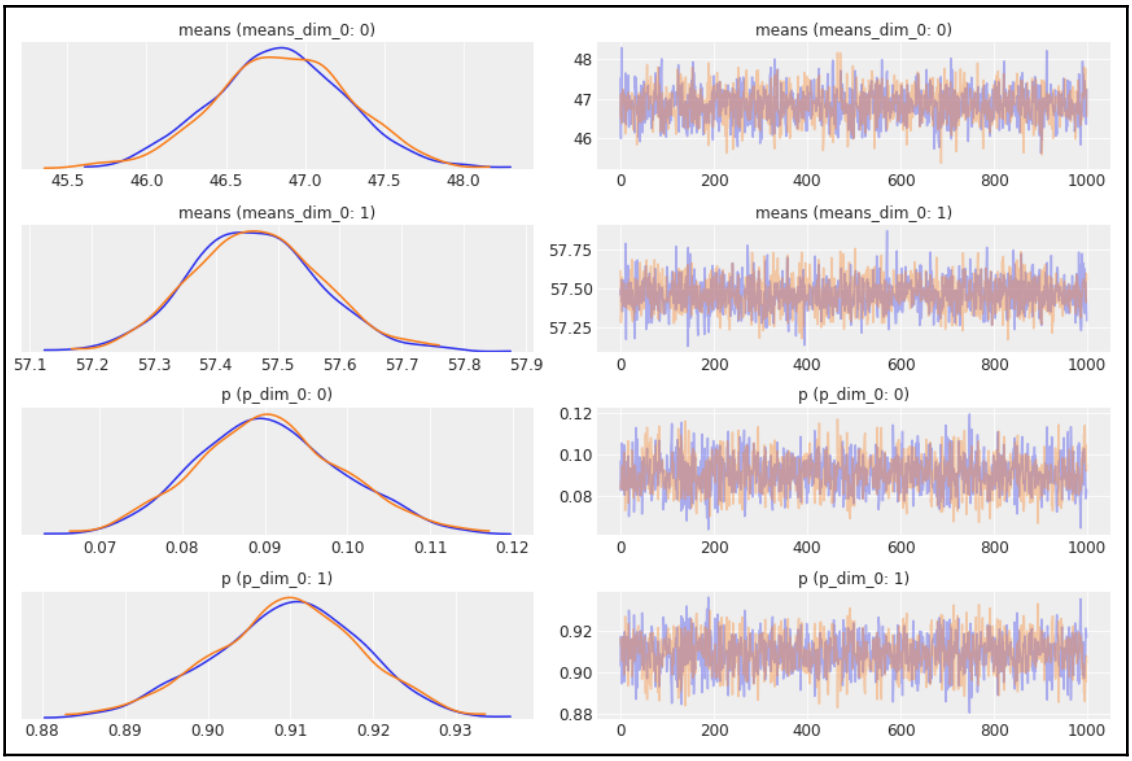
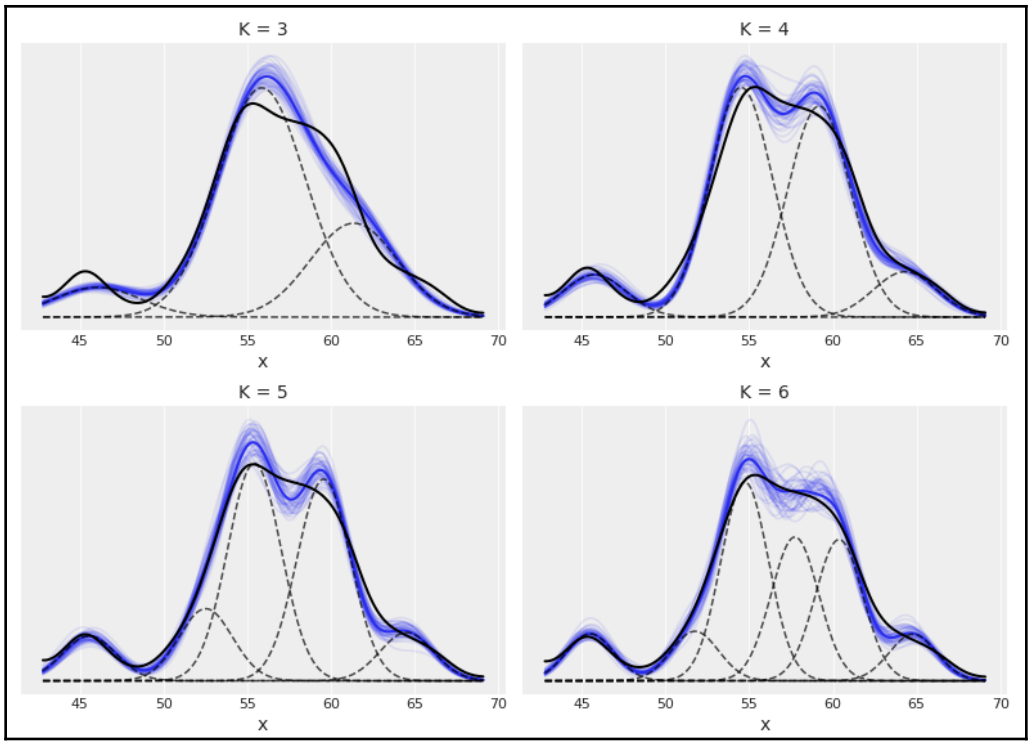
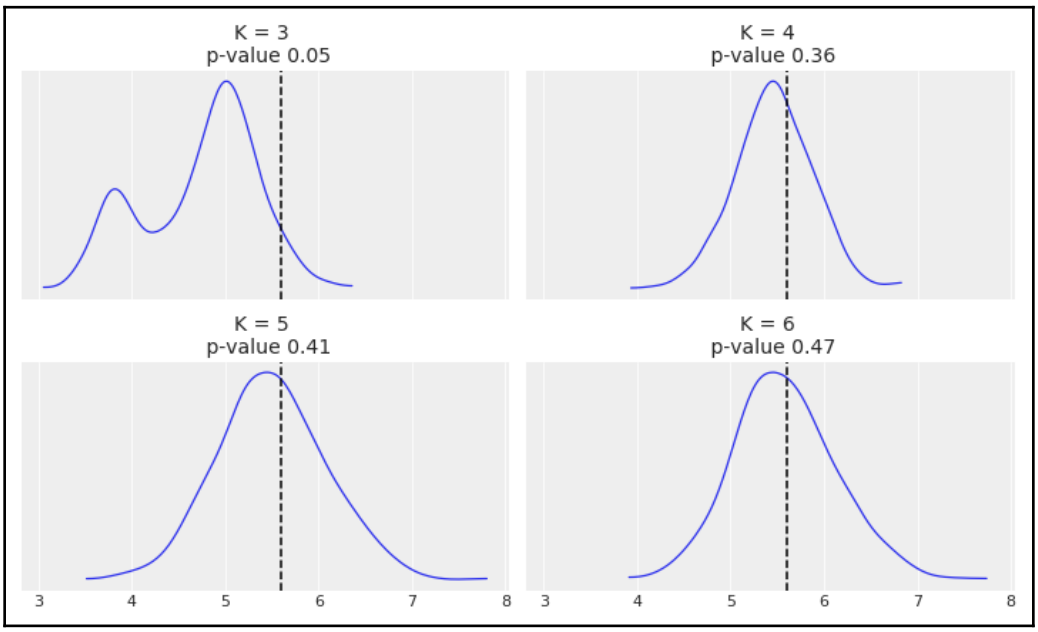
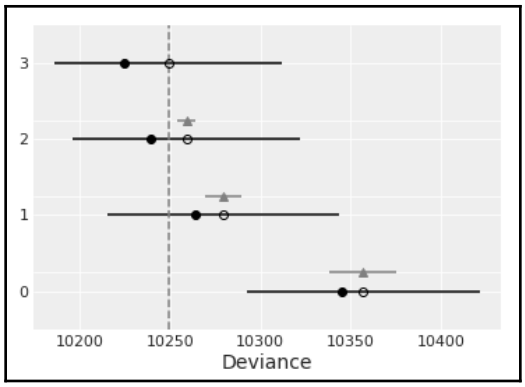

means (means_dim_0: 0)
means (means_dim_0: 0)
means (means_dim_0: 1)
means (means_dim_0: 1)
p (p_dim_0: 0)
p (p_dim_0: 0)
p (p_dim_0: 1)

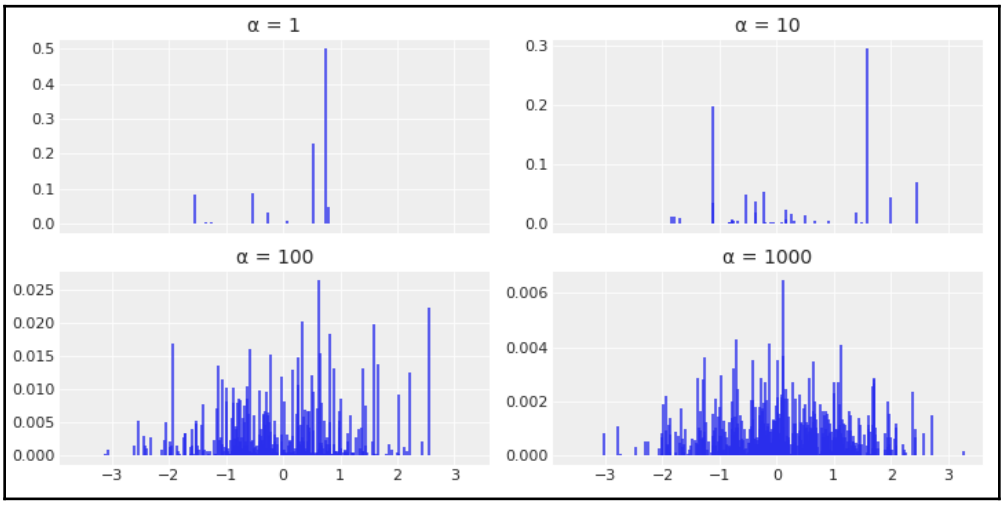
Figure 6.6

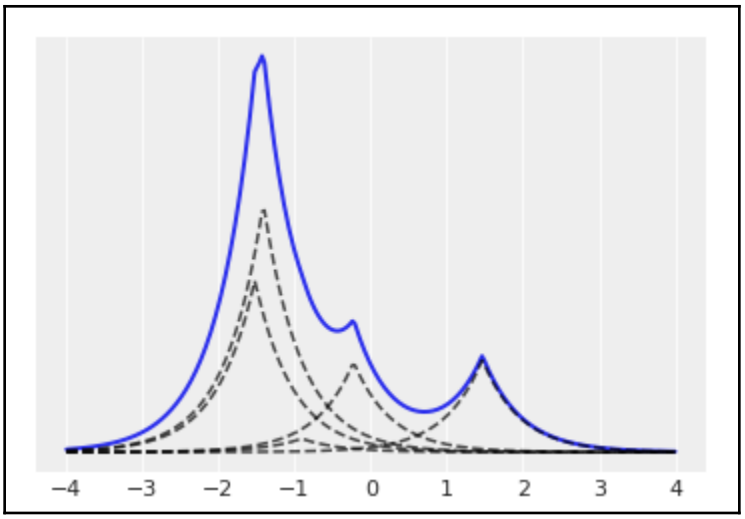

means (means_dim_0: 0)
45.5 46.0 46.5 47.0 47.5 48.0
means (means_dim_0: 0)
48 47 46
0 200 400 600 800 1000
means (means_dim_0: 1)
57.1 57.2 57.3 57.4 57.5 57.6 57.7 57.8 57.9
means (means_dim_0: 1)
57.75 57.50 57.25
0 200 400 600 800 1000
p (p_dim_0: 0)
0.07 0.08 0.09 0.10 0.11 0.12
p (p_dim_0: 0)
0.12 0.10 0.08
0 200 400 600 800 1000
p (p_dim_0: 1)
0.88 0.89 0.90 0.91 0.92 0.93
p (p_dim_0: 1)
0.92 0.90 0.88
0 200 400 600 800 1000

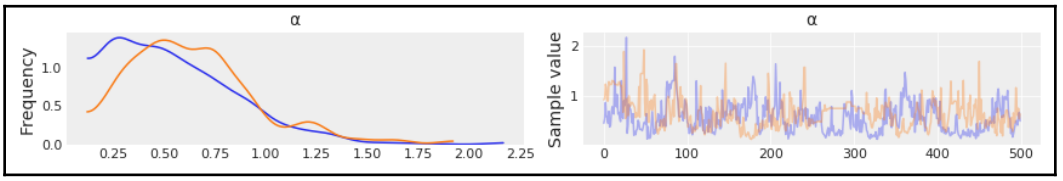

K = 3
K = 4
K = 5
K = 6
45
50
55
60
65
70
x

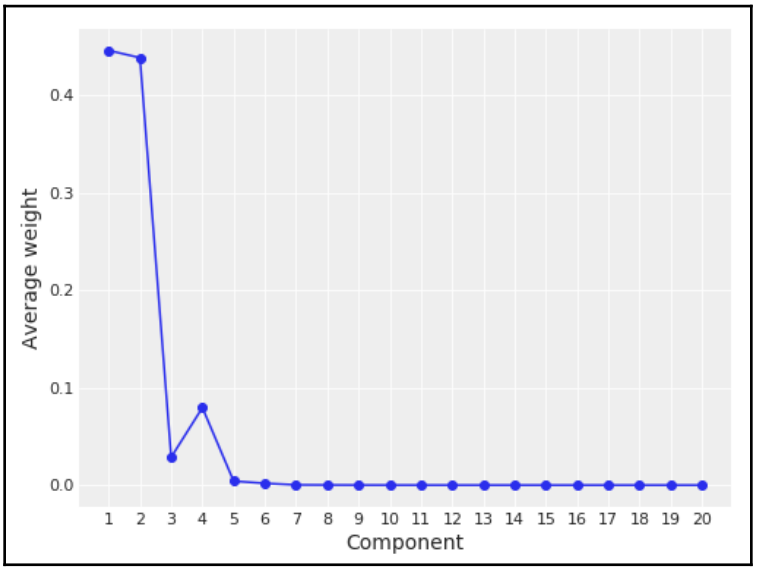

K = 3
p-value 0.05
K = 4
p-value 0.36
K = 5
p-value 0.41
K = 6
p-value 0.47
3
4
5
6
7
8

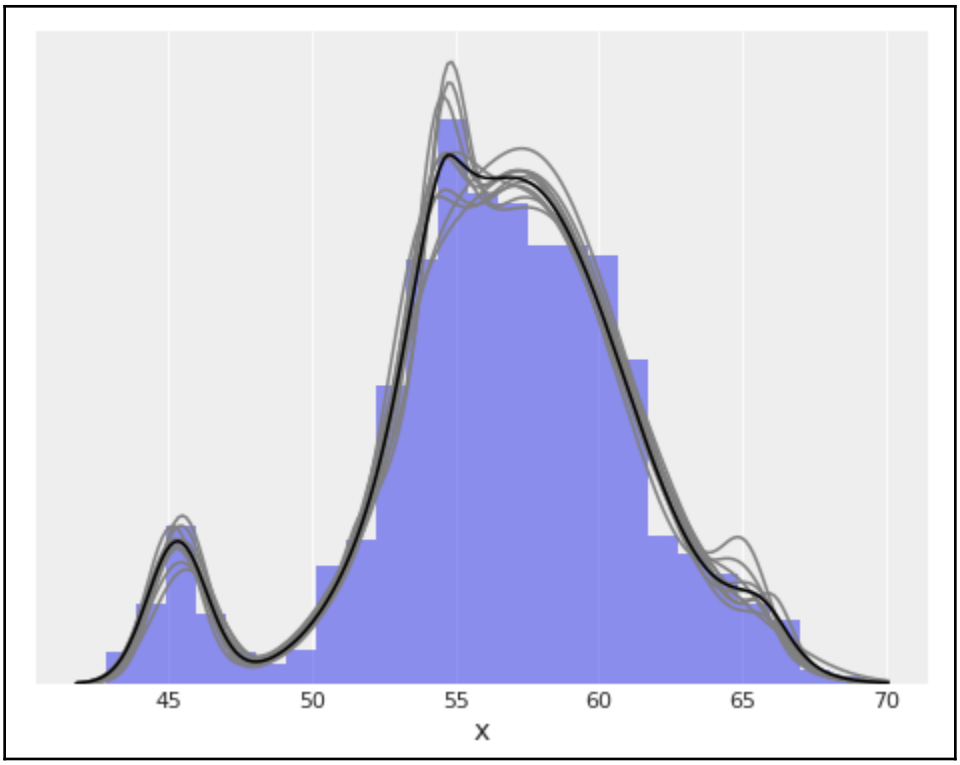

3
2
1
0
10200
10250
10300
10350
10400
Deviance


α = 1
α = 10
α = 100
α = 1000

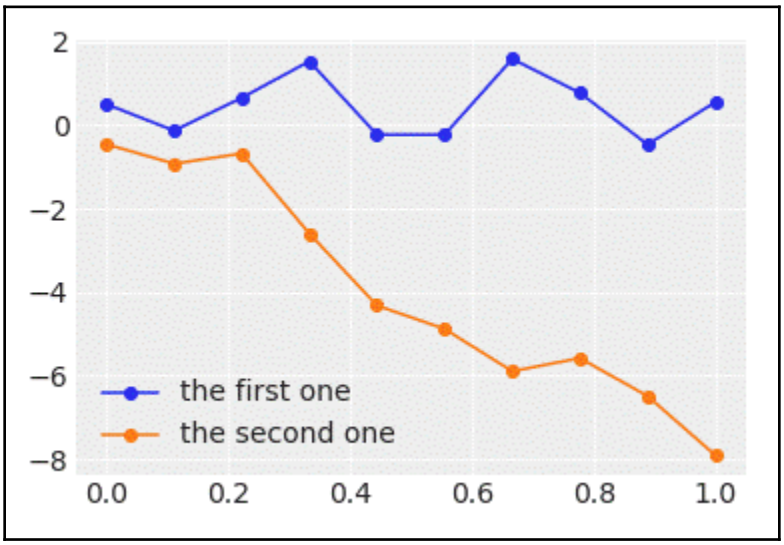

−4
−3
−2
−1
0
1
2
3
4

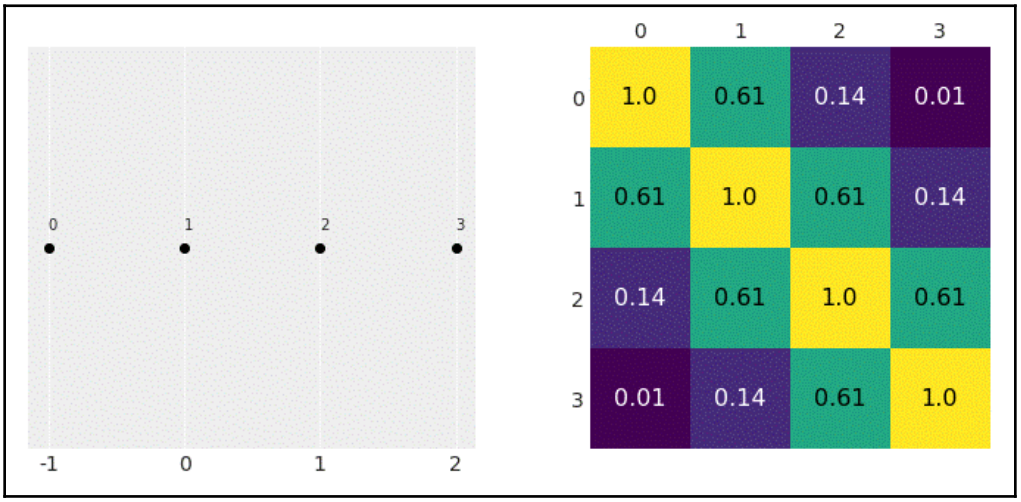

α
Frequency
1.0
0.5
0.0
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
α
Sample value
2
1
0
100
200
300
400
500

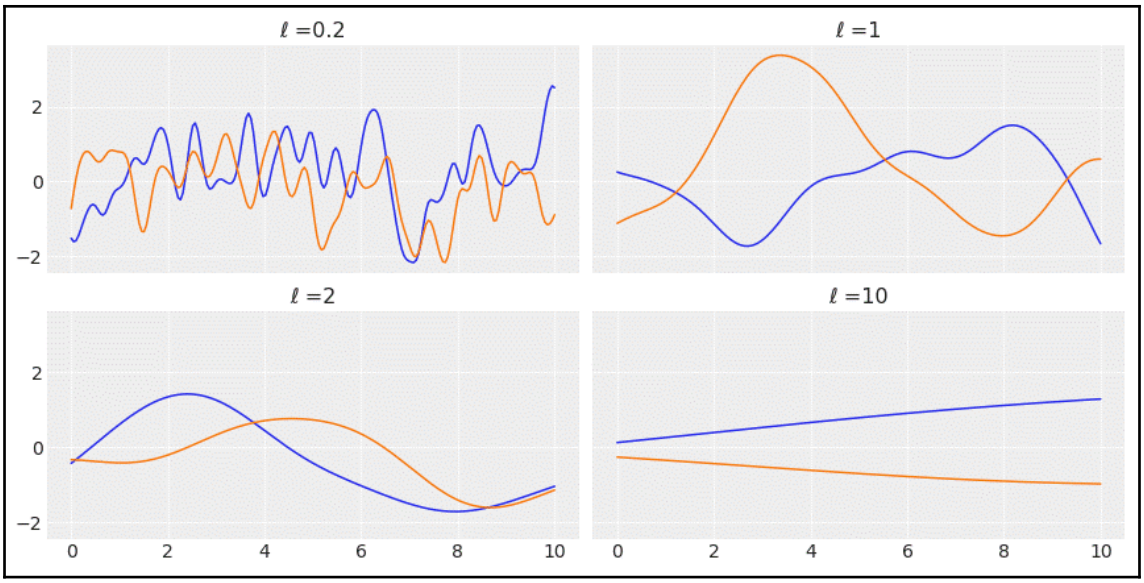

Average weight
0.4
0.3
0.2
0.1
0.0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
Component

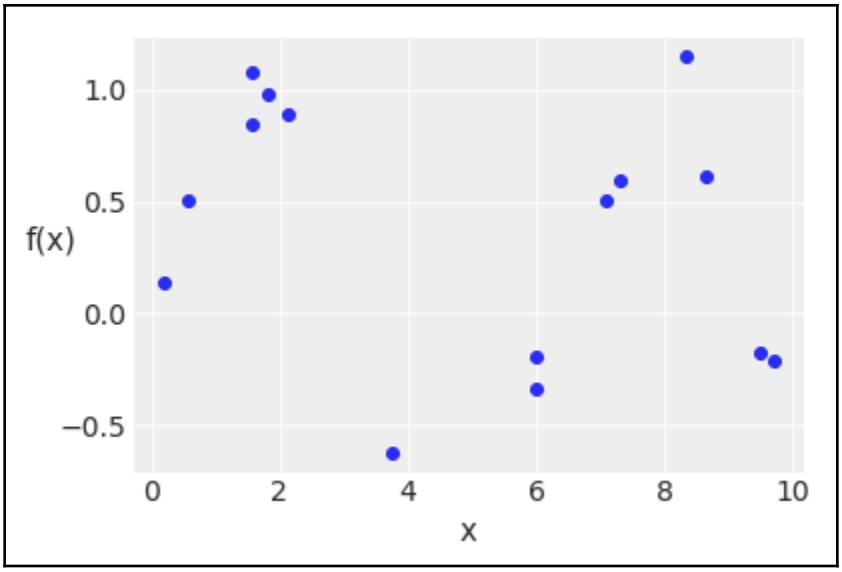

45
50
55
60
65
70
X

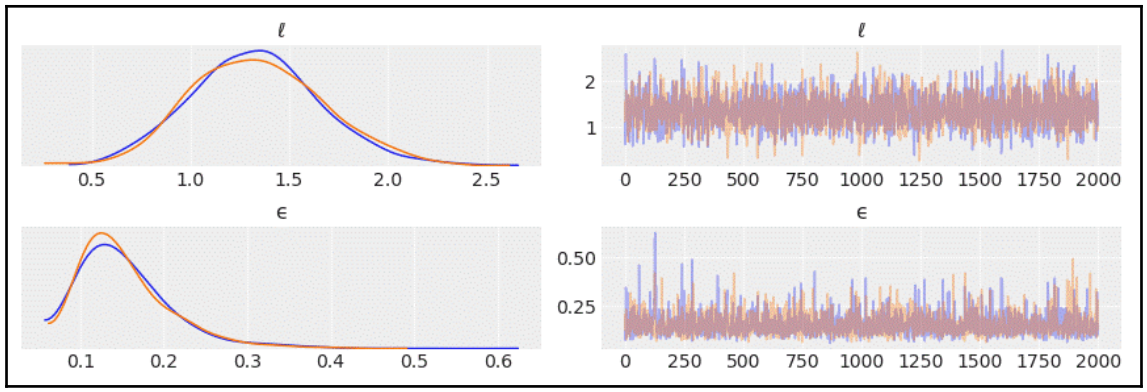
# Chapter 7: Gaussian Processes

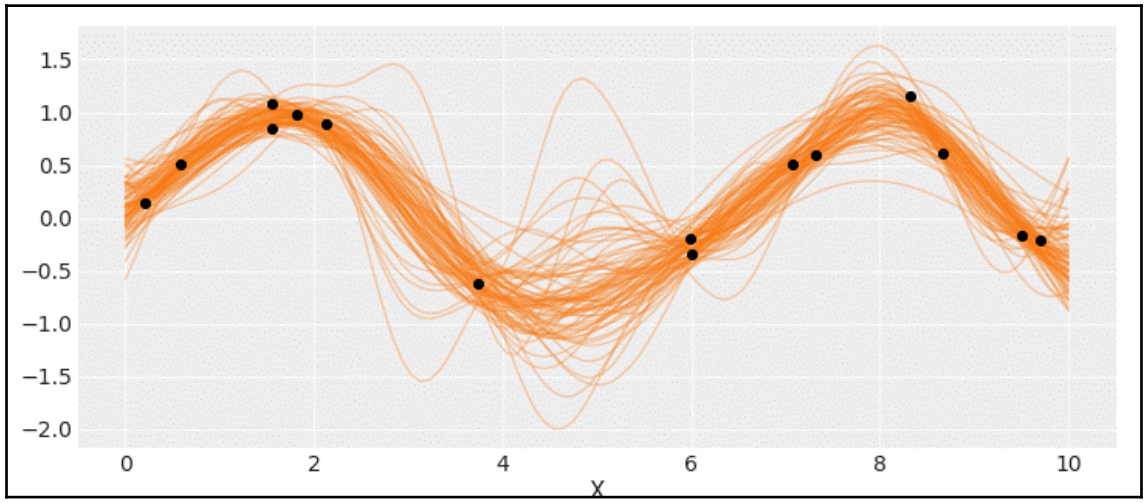


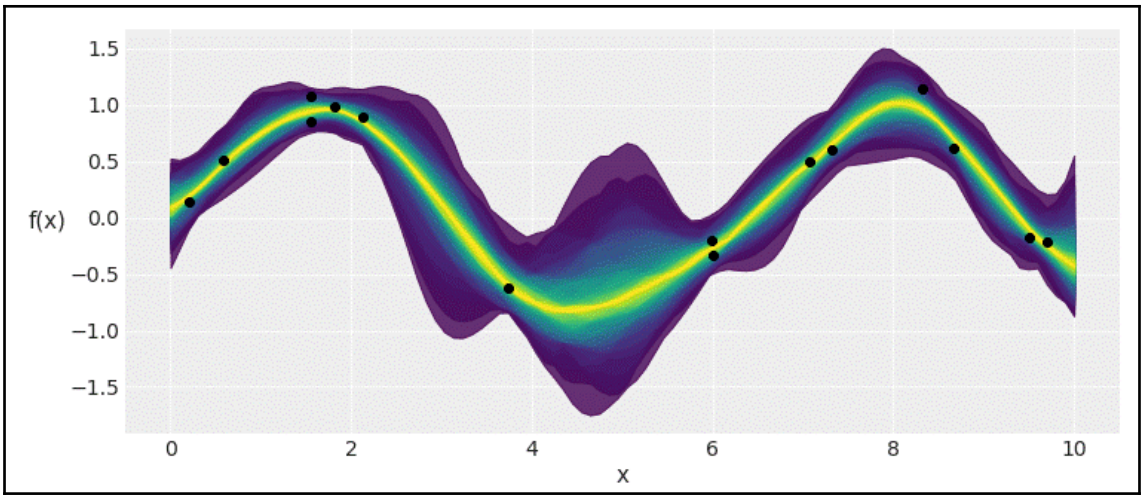



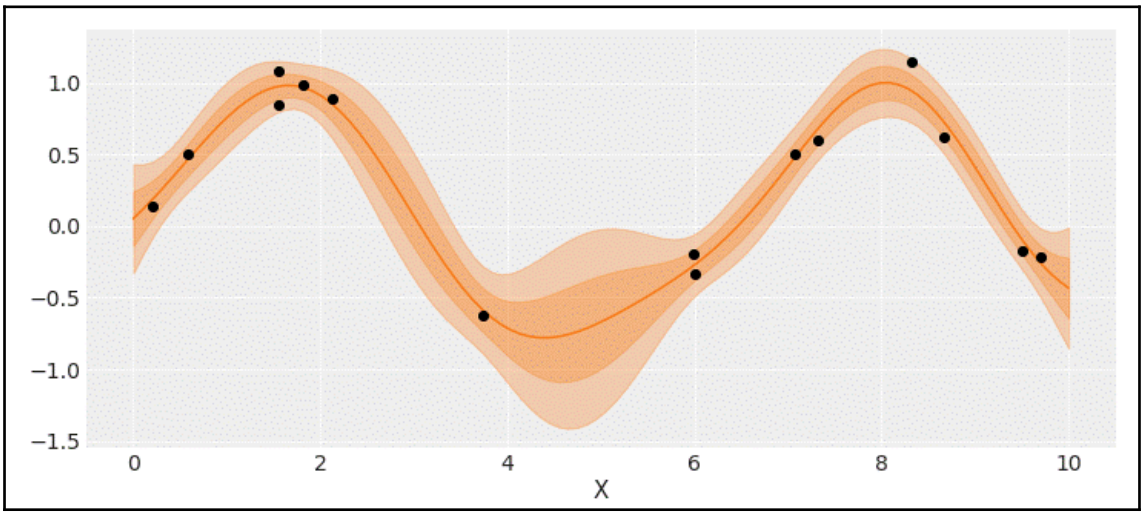

ℓ =0.2
ℓ =1
ℓ =2
ℓ =10
2
0
−2
0
2
4
6
8
10

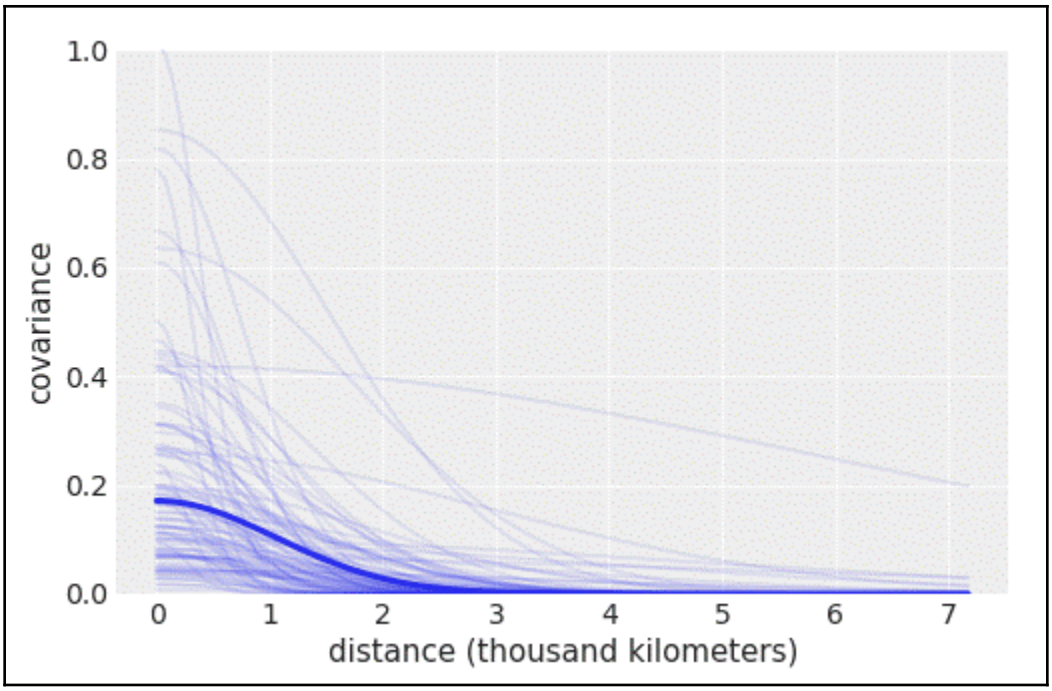

1.0
0.5
0.0
−0.5
f(x)
0
2
4
6
8
10
x

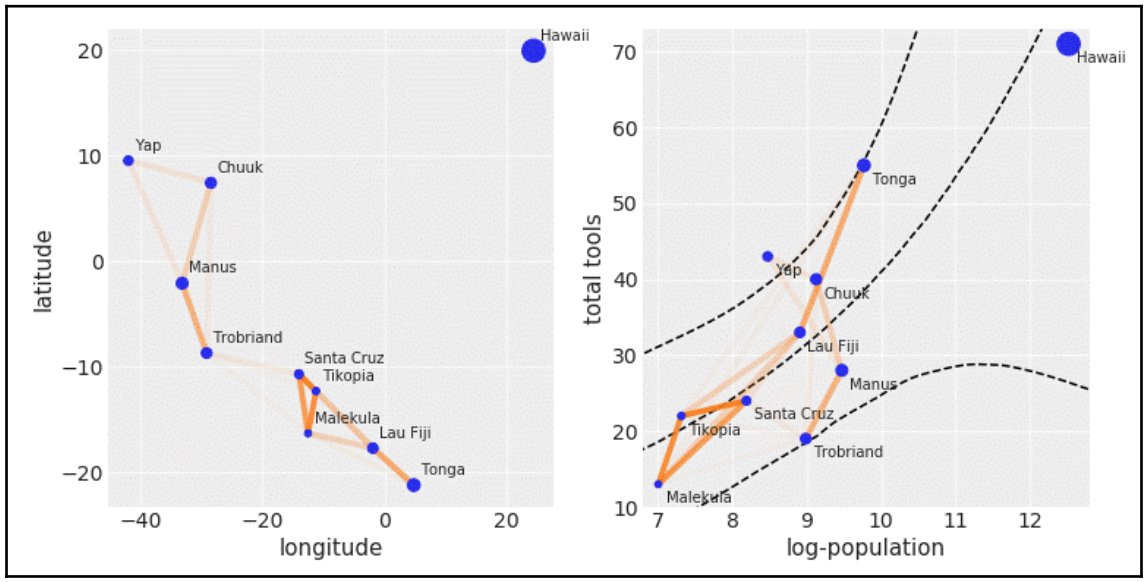

ℓ
0.5
1.0
1.5
2.0
2.5
ℓ
2
1
0
250
500
750
1000
1250
1500
1750
2000
ϵ
0.1
0.2
0.3
0.4
0.5
0.6
ϵ
0.50
0.25
0
250
500
750
1000
1250
1500
1750
2000

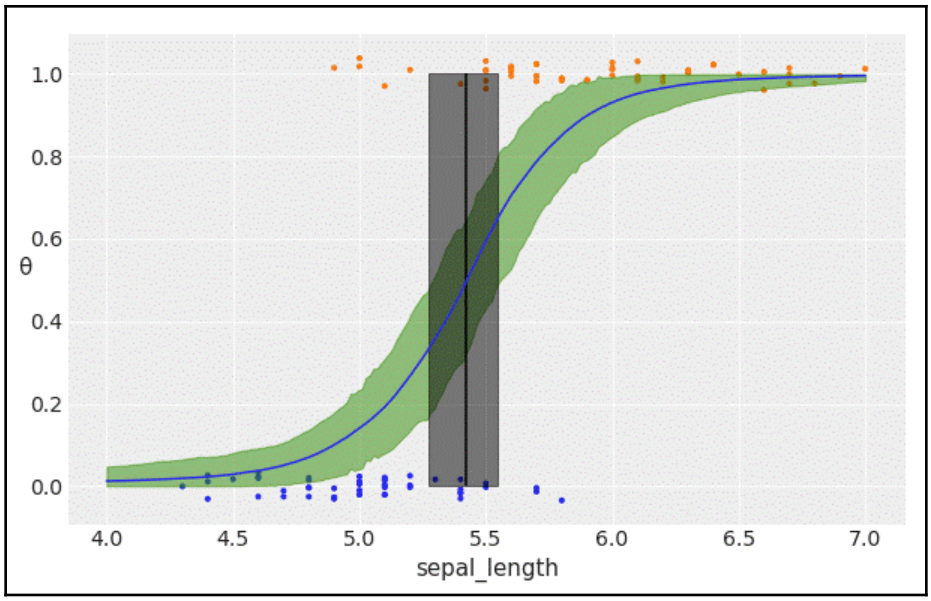

1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
0
2
4
6
8
10
X

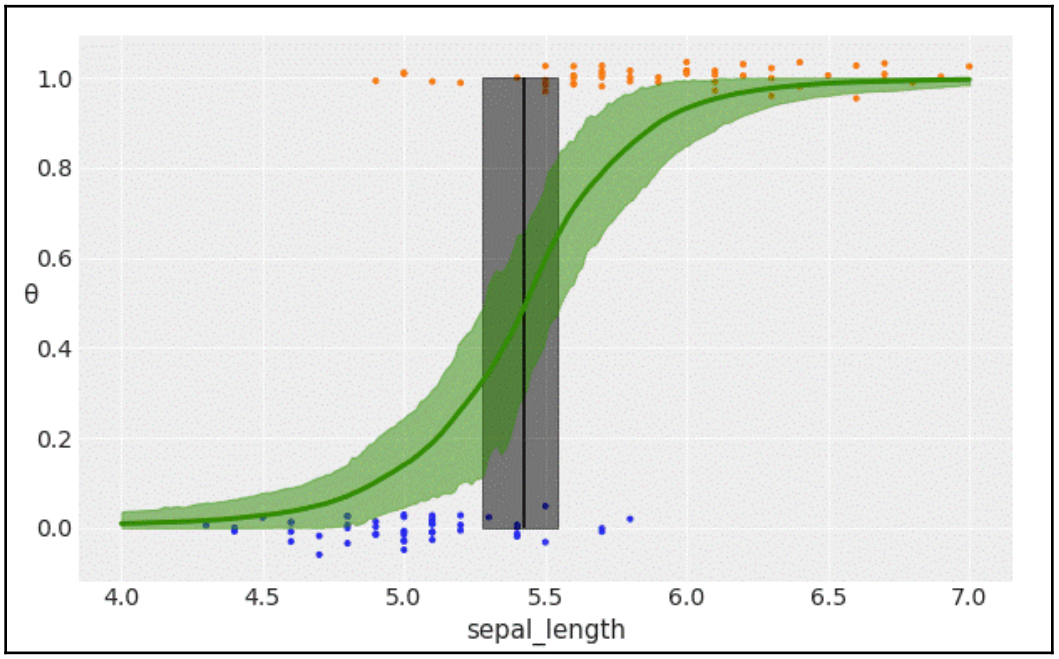

1.5
1.0
0.5
f(x) 0.0
−0.5
−1.0
−1.5
0
2
4
6
8
10
x

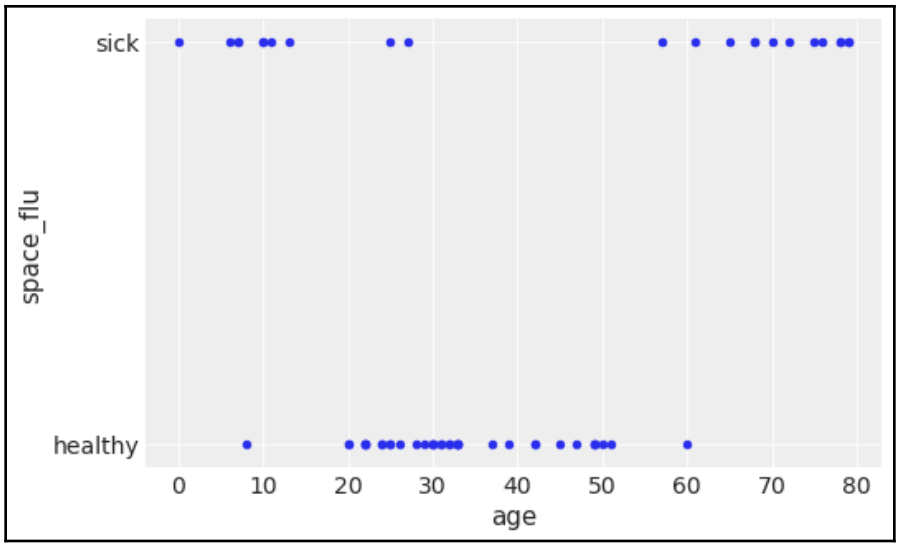

1.0
0.5
0.0
−0.5
−1.0
−1.5
0
2
4
6
8
10
X

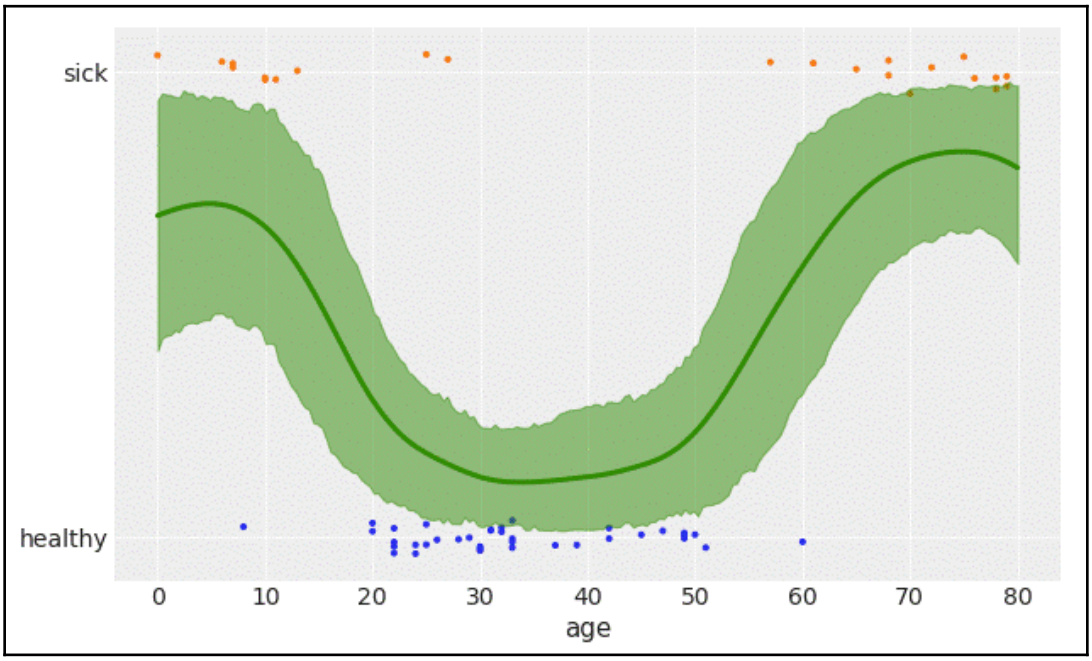

covariance
1.0
0.8
0.6
0.4
0.2
0.0
0
1
2
3
4
5
6
7
distance (thousand kilometers)

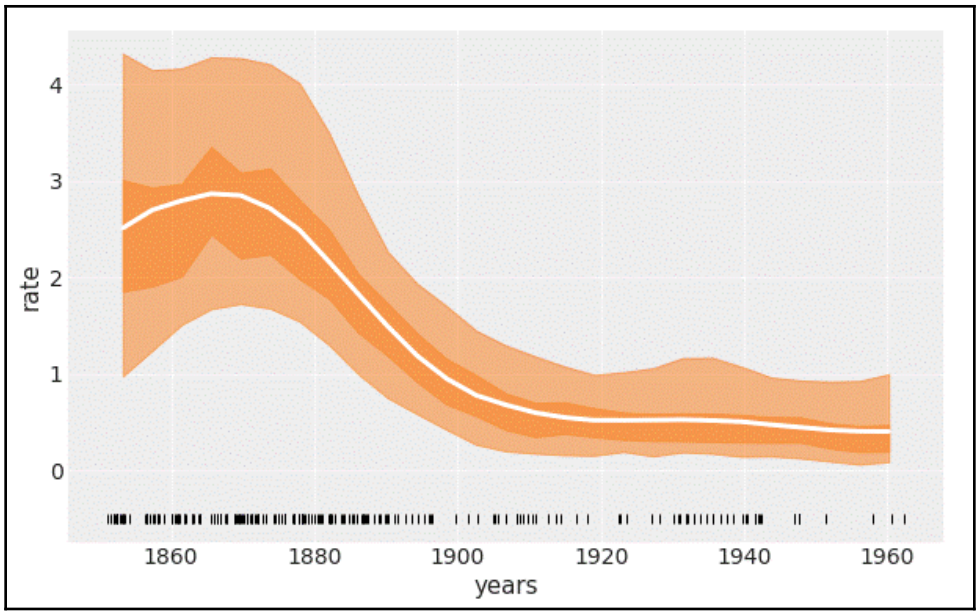

latitude
longitude
Hawaii
Yap
Chuuk
Manus
Trobriand
Santa Cruz
Tikopia
Malekula
Lau Fiji
Tonga
total tools
log-population

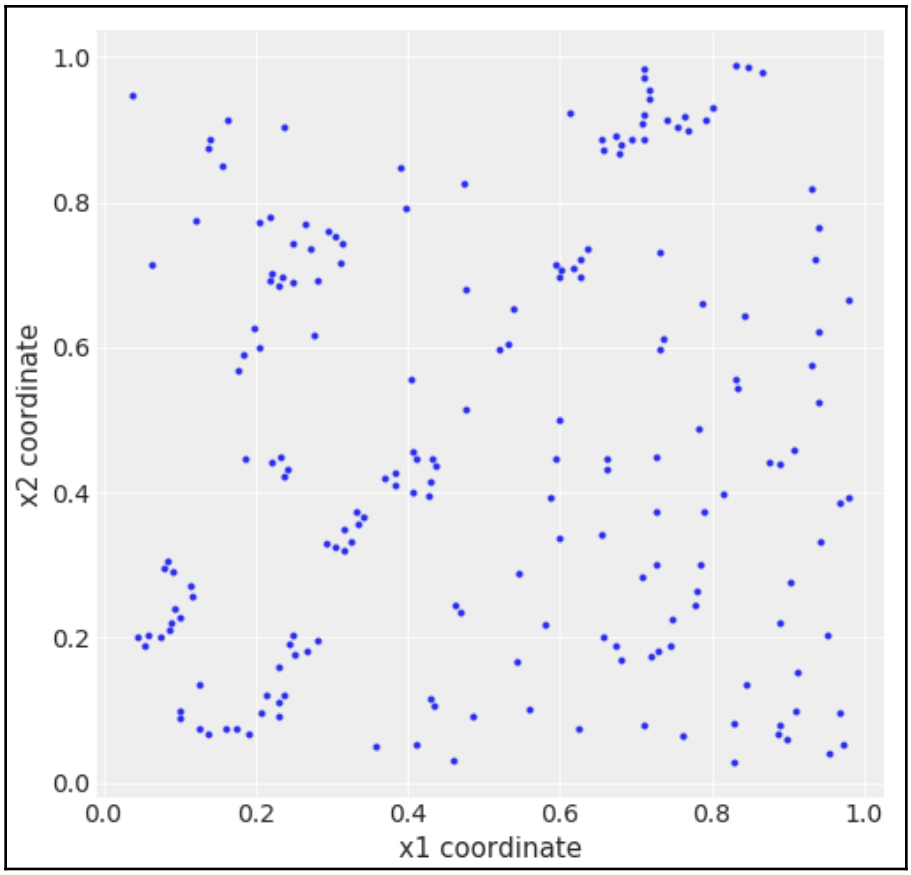

1.0
0.8
0.6
θ
0.4
0.2
0.0
4.0
4.5
5.0
5.5
6.0
6.5
7.0
sepal_length

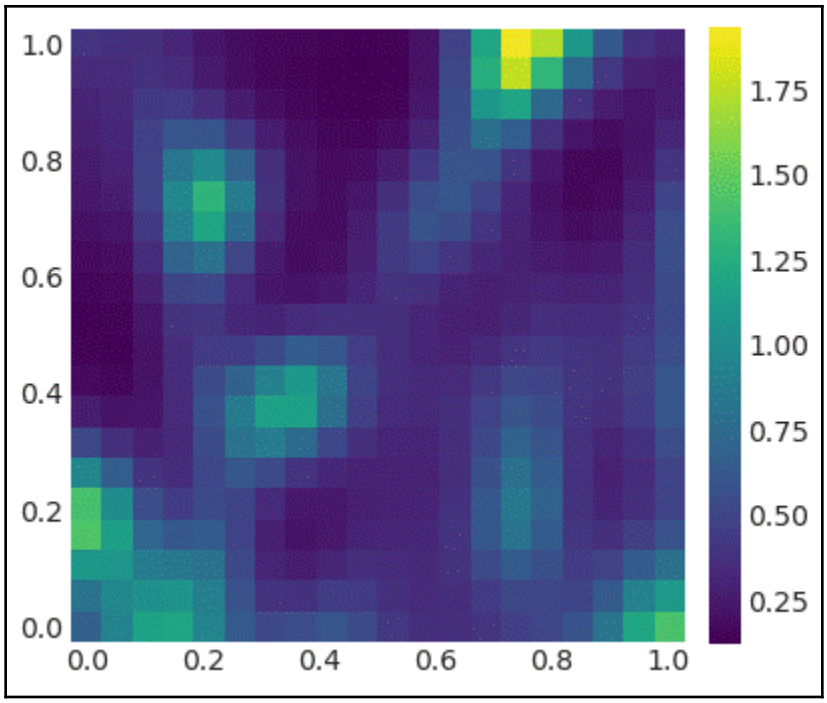

1.0
0.8
0.6
θ
0.4
0.2
0.0
4.0
4.5
5.0
5.5
6.0
6.5
7.0
sepal_length


sick
healthy
space_flu
0
10
20
30
40
50
60
70
80
age

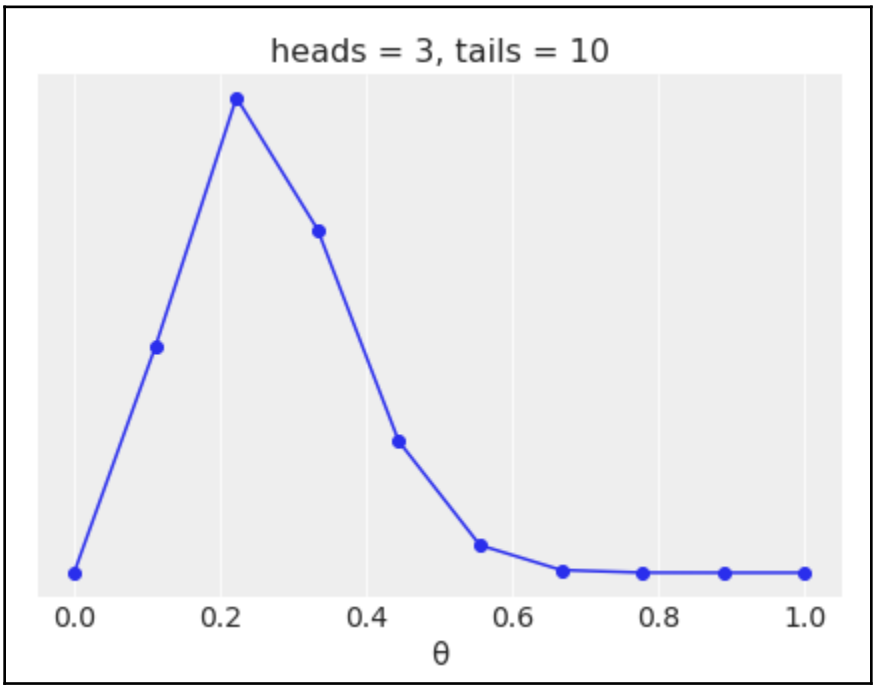

sick
healthy
0
10
20
30
40
50
60
70
80
age

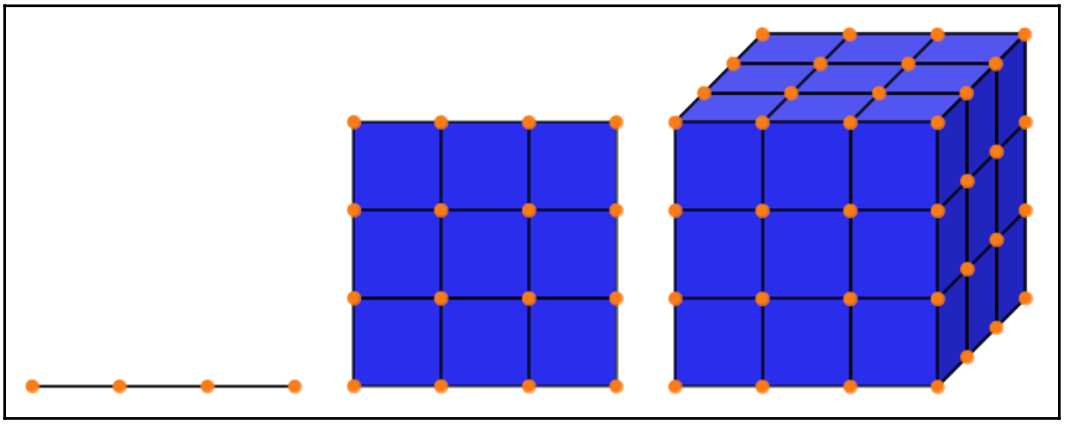

4
3
2
1
0
rate
1860
1880
1900
1920
1940
1960
years

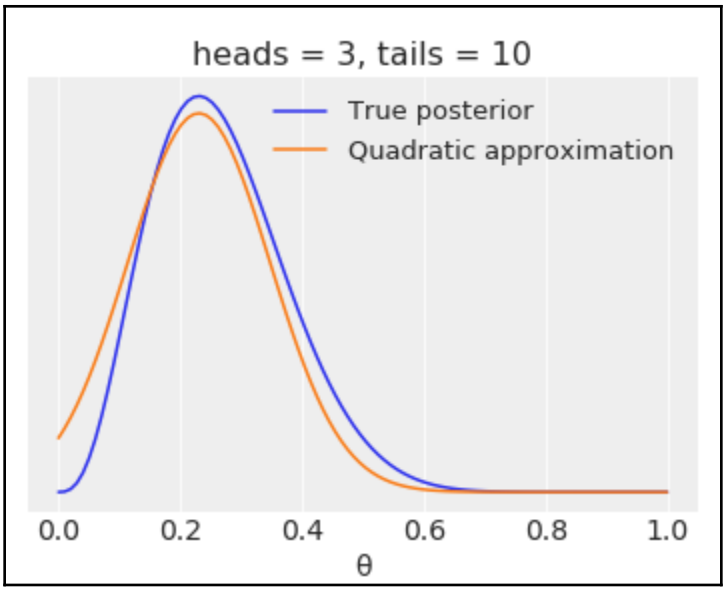

1.0
0.8
0.6
0.4
0.2
0.0
x2 coordinate
0.0
0.2
0.4
0.6
0.8
1.0
x1 coordinate

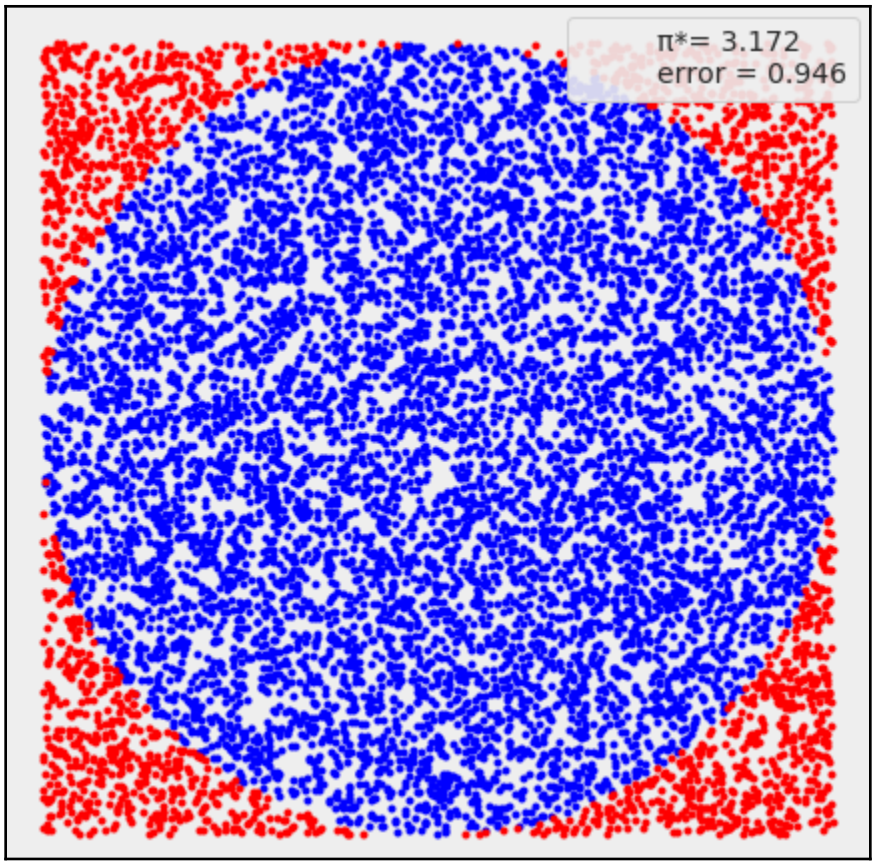

1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
1.75
1.50
1.25
1.00
0.75
0.50
0.25

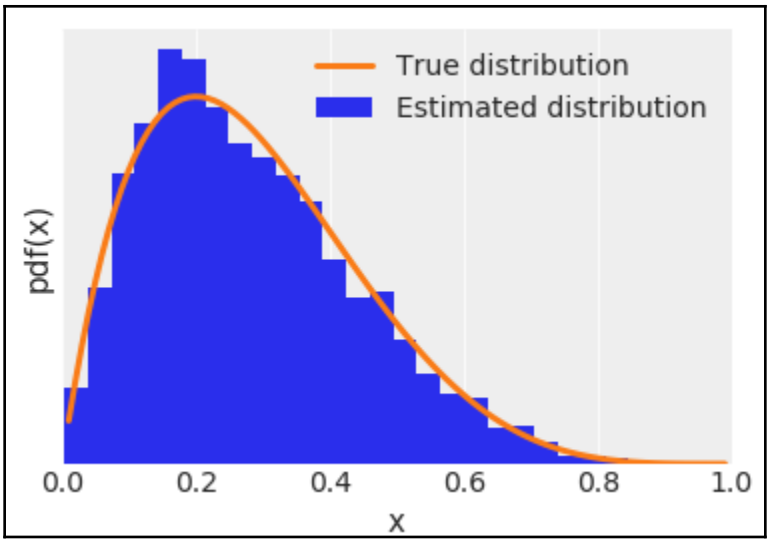
# Chapter 8: Inference Engines

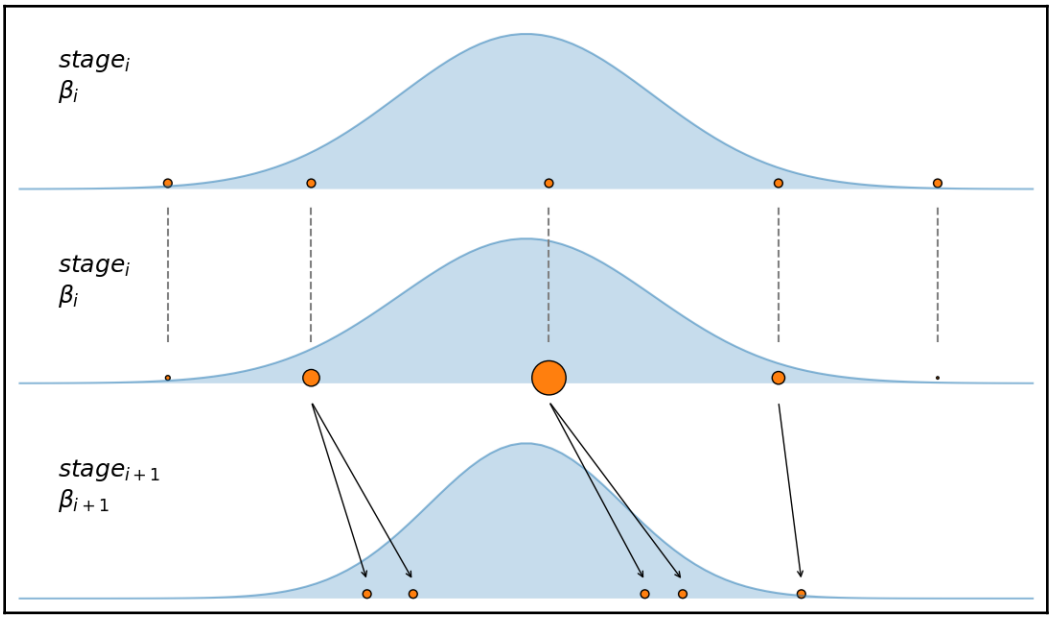


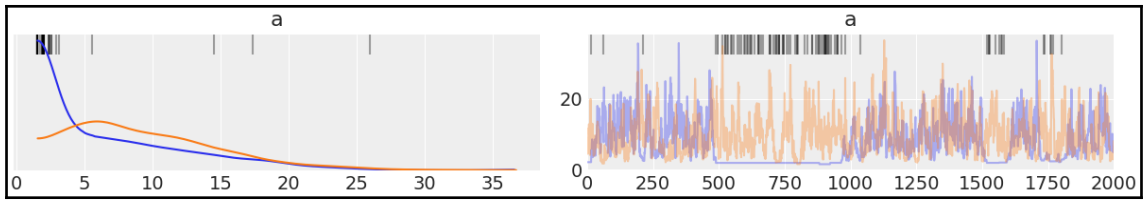



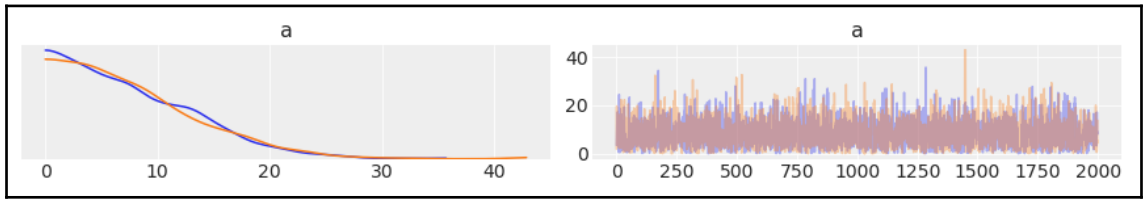

heads = 3, tails = 10
True posterior
Quadratic approximation
0.0
0.2
0.4
0.6
0.8
1.0
θ

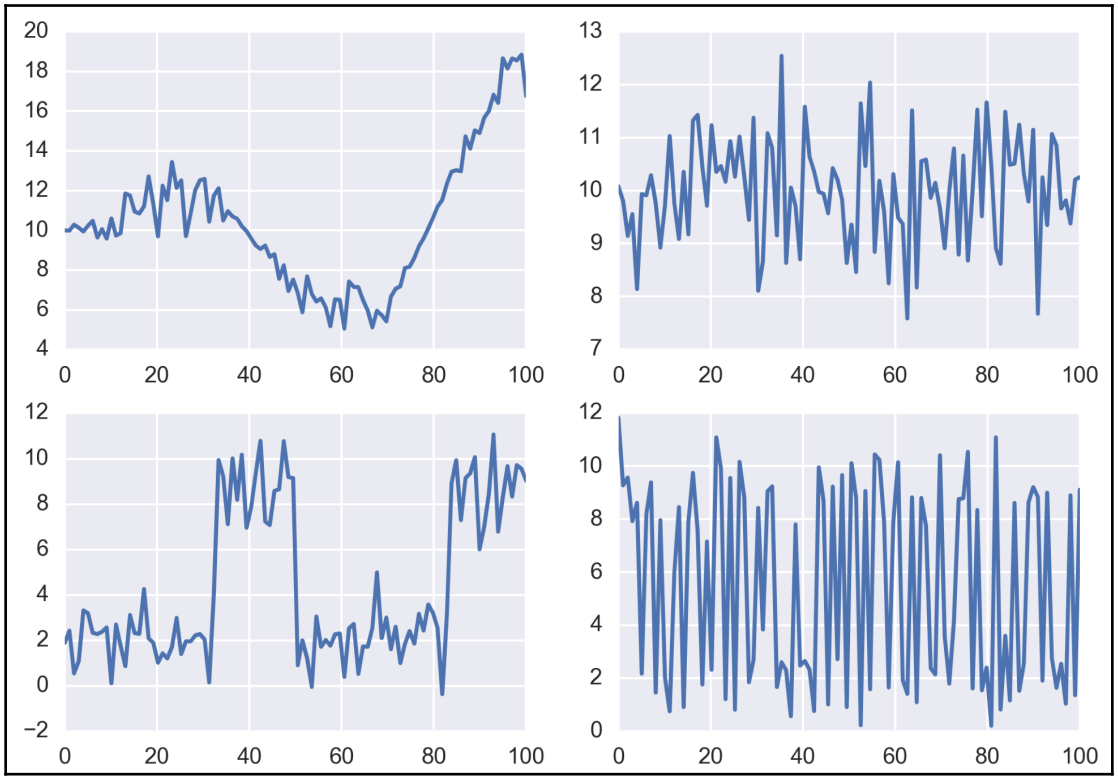

π*= 3.172
error = 0.946

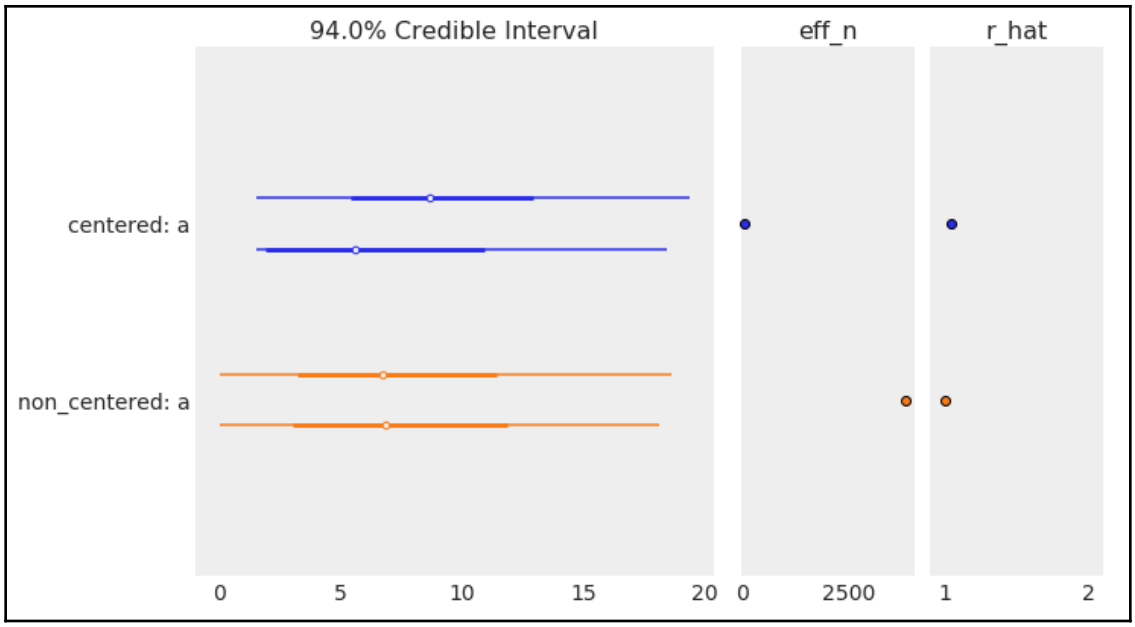

True distribution
Estimated distribution
pdf(x)
0.0
0.2
0.4
0.6
0.8
1.0
x

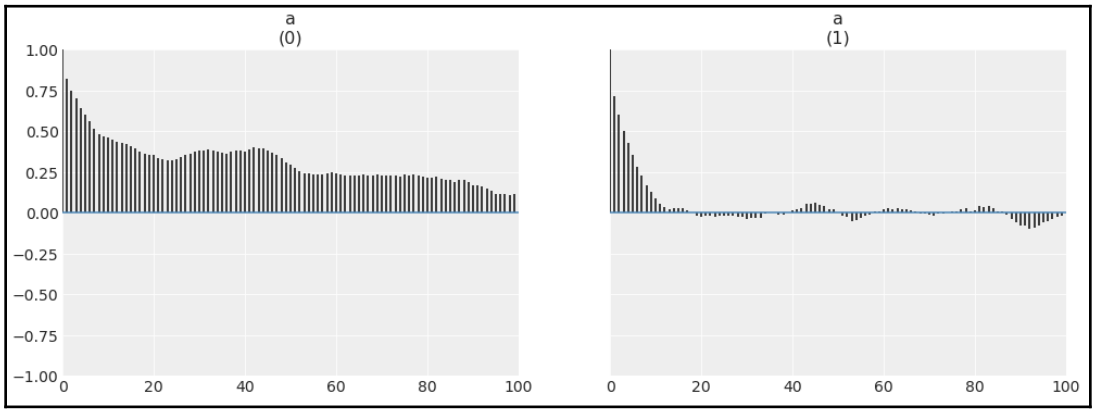

$stage_i$
$\beta_i$
$stage_i$
$\beta_i$
$stage_{i+1}$
$\beta_{i+1}$

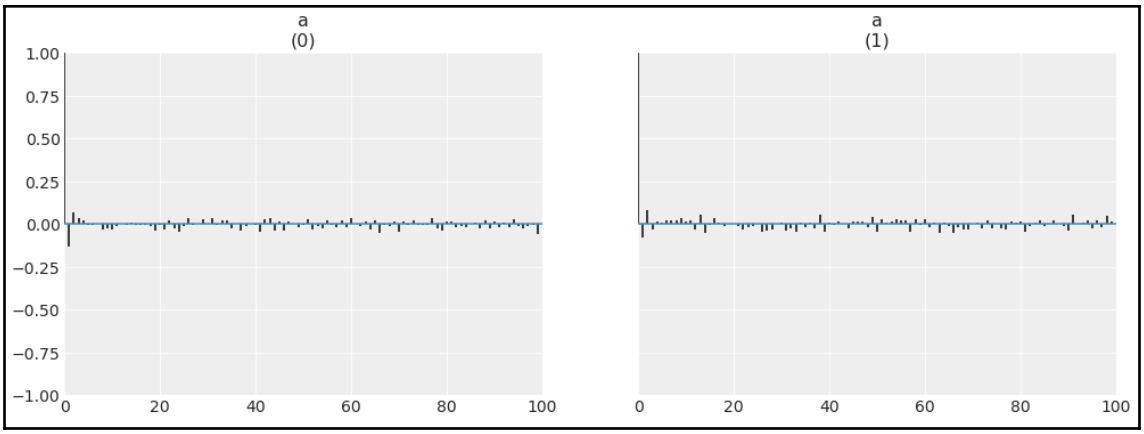

a
0 5 10 15 20 25 30 35
a
0 20
0 250 500 750 1000 1250 1500 1750 2000

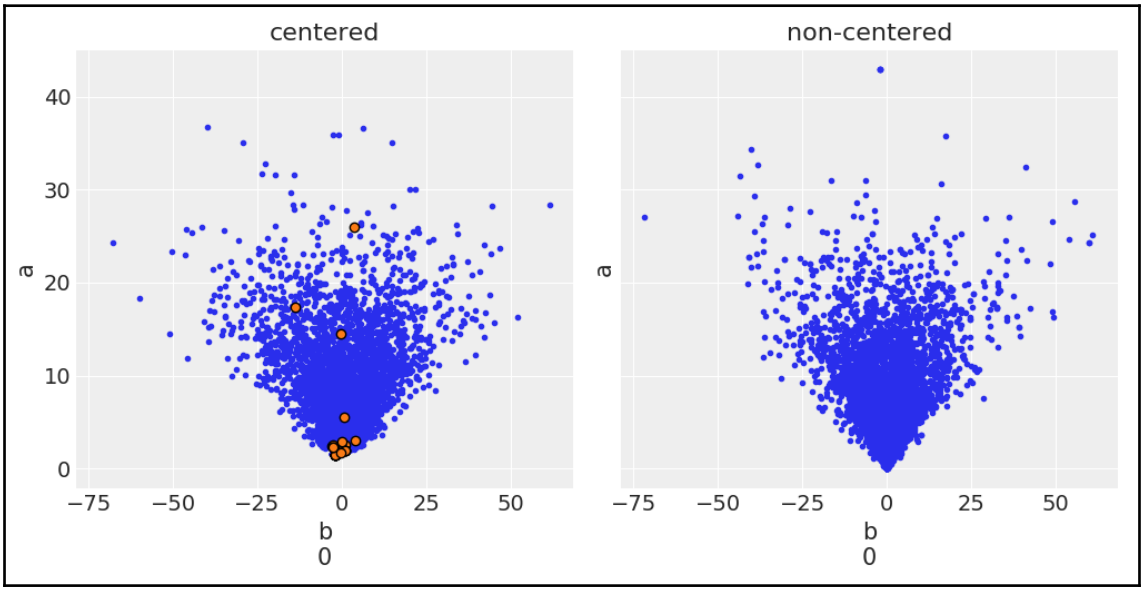

a
0 10 20 30 40
a
0 20 40
0 250 500 750 1000 1250 1500 1750 2000

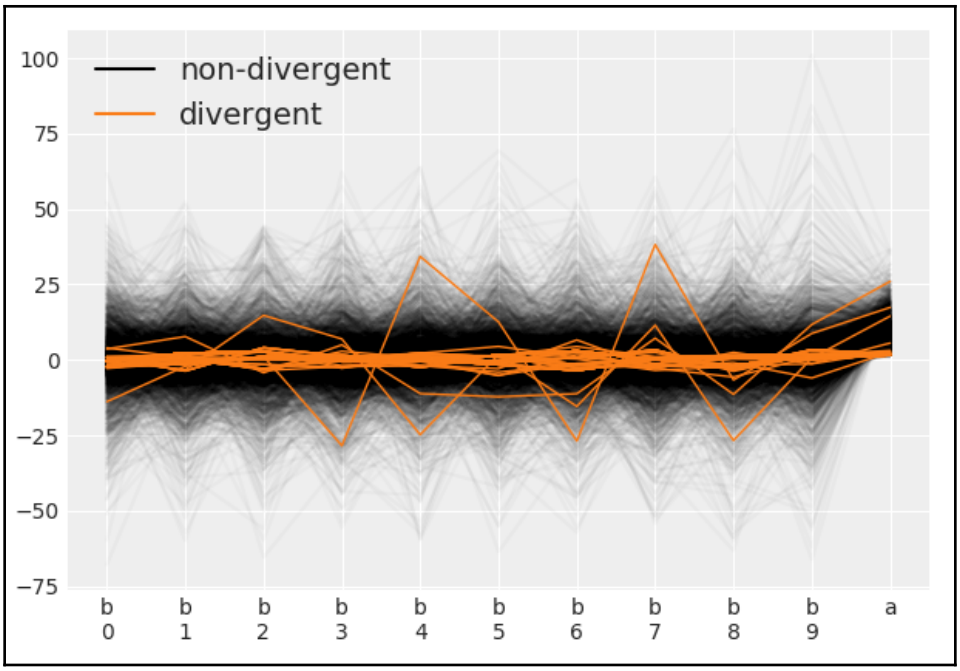

20 18 16 14 12 10 8 6 4
0 20 40 60 80 100
13 12 11 10 9 8 7
0 20 40 60 80 100
12 10 8 6 4 2 0 −2
0 20 40 60 80 100
12 10 8 6 4 2 0
0 20 40 60 80 100


94.0% Credible Interval
eff_n
r_hat
centered: a
non_centered: a
0
5
10
15
20
0
2500
1
2


a
(0)
a
(1)
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
0
20
40
60
80
100


a
(0)
a
(1)
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
0
20
40
60
80
100


centered
non-centered
a
b
0
40
30
20
10
0
−75
−50
−25
25
50


non-divergent
divergent
100
75
50
25
0
−25
−50
−75
b 0
b 1
b 2
b 3
b 4
b 5
b 6
b 7
b 8
b 9
a