1 Easy

1.1

In the model definition below, which line is the likelihood?

$$y_i \sim \mathcal{N}(\mu, \sigma) \tag{1}$$
$$\mu \sim \mathcal{N}(0, 10) \tag{2}$$
$$\sigma \sim Exp(1) \tag{3}$$

y_i is the likelihood.

1.2

In the model definition just above, how many parameters are in the posterior distribution?

Two, μ and σ.

1.3

Using the model definition above, write down the appropriate form of Bayes' theorem that includes the proper likelihood and priors.

$$Pr(\mu, \sigma|y) = \frac{\Pi_i \mathcal{N}(y|\mu, \sigma) * \mathcal{N}(\mu|0, 10) * Exp(\sigma|1)}{\int \int \Pi_i \mathcal{N}(y|\mu, \sigma) * \mathcal{N}(\mu|0, 10) * Exp(\sigma|1) d\mu d\sigma} \tag{4}$$

1.4

In the model definition below, which line is the linear model?

$$y_i \sim \mathcal{N}(\mu, \sigma) \tag{5}$$
$$\mu_i = \alpha + \beta x_i \tag{6}$$
$$\alpha \sim \mathcal{N}(0, 10) \tag{7}$$
$$\beta \sim \mathcal{N}(0, 1) \tag{8}$$
$$\sigma \sim Exp(2) \tag{9}$$

The line $\mu_i = \alpha + \beta x_i$ specifies the linear model.

1.5

In the model definition just above, how many parameters are in the posterior distribution?

Three - α, β and σ.

1.6

Translate the quap model formula below into a mathematical model definition.

```
flist <- alist(
  y ~ dnorm( mu , sigma ), mu <- a + b*x,
  a ~ dnorm( 0 , 10 ),
  b ~ dunif( 0 , 1 ), sigma ~ dexp( 1 )
)
```

$$y_i \sim \mathcal{N}(\mu, \sigma) \tag{10}$$
$$\mu_i = \alpha + \beta x_i \tag{11}$$
$$\alpha \sim \mathcal{N}(0, 10) \tag{12}$$
$$\beta \sim \mathcal{U}(0, 1) \tag{13}$$
$$\sigma \sim Exp(2) \tag{14}$$

1.7

A sample of students is measures for height each year for 3 years. After the third year, you want to fit a linear regression predicting height using year as a predictor. Write down the mathematical model definition for this regression, using any variable names and priors you choose. Be prepared to defend your choice of priors.

$$
\begin{align}
h_i &\sim \mathcal{N}(\mu, \sigma) \tag{15} \\
\mu_i &= \alpha + e^{\beta} x_i \tag{16} \\
\alpha &\sim \mathcal{N}(165, 30) \tag{17} \\
\beta &\sim \mathcal{N}(0, 10) \tag{18} \\
\sigma &\sim \mathcal{U}(0, 20) \tag{19} \\
& \tag{20}
\end{align}
$$

I'm assuming the question implies we are only interested in predicting year 3 growth for these adolescents during a stage in life in which they are growing; otherwise one might prefer an asymptotically flat curve or maybe a multi stage to account for shrinking in late age. But, to keep it simple and in line with the material covered in the chapter, I choose a linear regression; α should correspond to the mean marginal height of the students; I'm guessing this is around 155 but I leave a wide variance to account for the little information we have about these students. Since they are "students," I assume they are not in the shrinking phase of life and model with log-normal β which prevents negative growth over years, but use 0 as the mean for β because they may be older students, in which case they wouldn't grow that much.

1.8

Now suppose I remind you that every student got taller each year. Does this information lead you to change your choice of priors? How?

I would be inclined to set a higher mean for β than 0.

1.9

Now suppose I tell you that the variance among heights for students of the same age is never more than 64cm. How does this lead you to revise your priors?

I would go for $\sigma \sim \mathcal{U}(0, 8)$, bounding σ^2 at 64.