

Take Home Quiz 2

Your Name

Due Tuesday Oct. 22 at 11:00am

There is no time limit, but this quiz should take you about 30-40 minutes. Place your answers into this markdown document, knit it, and hand in the result as a PDF or Word document. You may use R, any reference material, and information already available on the internet. Do not work together, do not use AI, and do not get help from anyone but Dr. Clair.

Problem 1 (10 points)

Here is a stan program:

```
data {  
  int<lower = 0, upper = 10> Y;  
}  
parameters {  
  real<lower = 0, upper = 1> pi;  
}  
model {  
  Y ~ binomial(10, pi);  
  pi ~ beta(2, 2);  
}
```

- What does this program do?
- If you pass $Y=3$ as data, what distribution will the simulation results have?

Solution This program computes the Beta-Binomial model for Y successes on 10 trials. If $Y=3$, the simulated π values would have the Beta(5,9) distribution.

Problem 2 (10 points)

Explain why grid approximation will not work for models with a large number of parameters.

Solution This is the “curse of dimensionality.” For a model with N parameters and a grid with k divisions, we’d need k^N calculations. To do a grid with 100 divisions and 20 parameters would require 10^{40} computations and take forever.

Problem 3 (10 points)

Amber Thomas collected data on the ages of actors in teen TV shows released between 2000 and 2021. We will investigate the question, how old are the actors playing 17-year-olds?

Let Y be the age of an actor playing a 17-year-old on TV. Model Y as a normal rv with unknown mean μ and sd σ , as:

- $Y \sim \text{Normal}(\mu, \sigma^2)$
- $\mu \sim \text{Normal}(17, 2^2)$
- $\sigma \sim \text{Gamma}(1,1)$

- a. What are the parameters in this model?
- b. Is this a conjugate prior model?
- c. Explain why the prior on μ is a reasonable choice.

Solution The parameters are μ and σ . It is not a conjugate prior model. Since the characters are 17, it's reasonable to think the actors might have a mean age near 17.

Problem 4 (10 points)

The teen actor data is available on our course web page at https://turtlegraphics.org/bayes/data/actor_character_age_difference.csv

As a first step, load the data and filter down to actors with `character_age` 17 which have a valid `actor_age` (some `actor_age` values are missing).

Make a histogram of the actor's ages for actors that have `character_age` equal to 17.

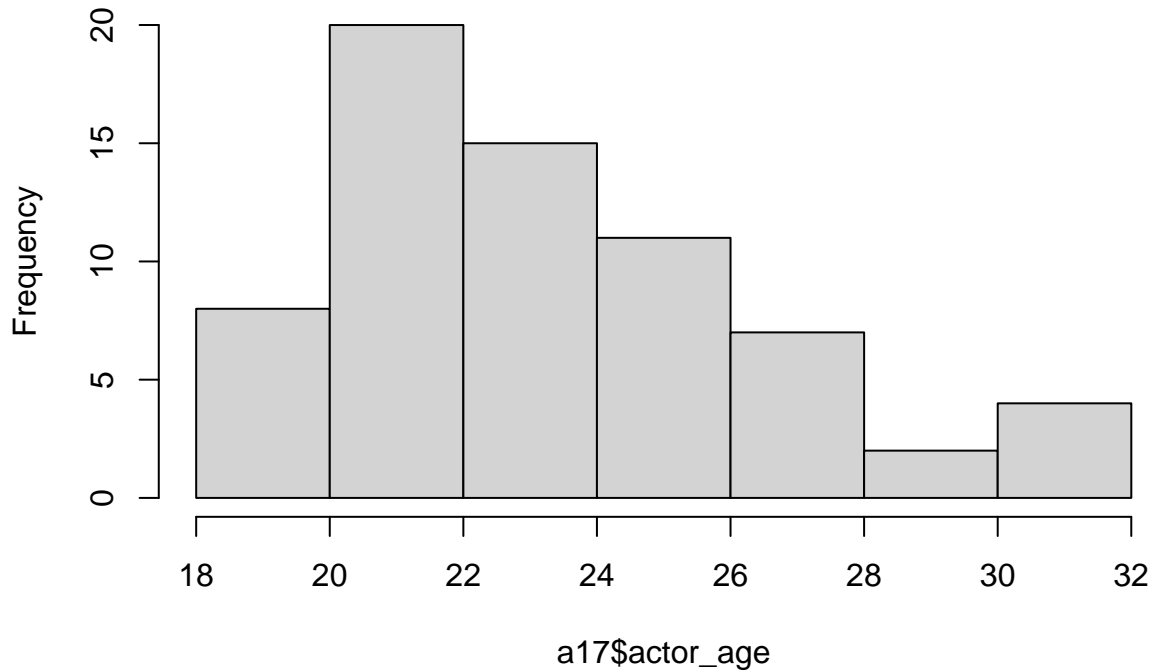
Solution

```
library(dplyr)

##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:stats':
##
##   filter, lag
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

actors <- read.csv("https://turtlegraphics.org/bayes/data/actor_character_age_difference.csv")
a17 <- actors |> filter(character_age == 17, !is.na(actor_age))
hist(a17$actor_age)
```

Histogram of a17\$factor_age



Problem 5 (10 points)

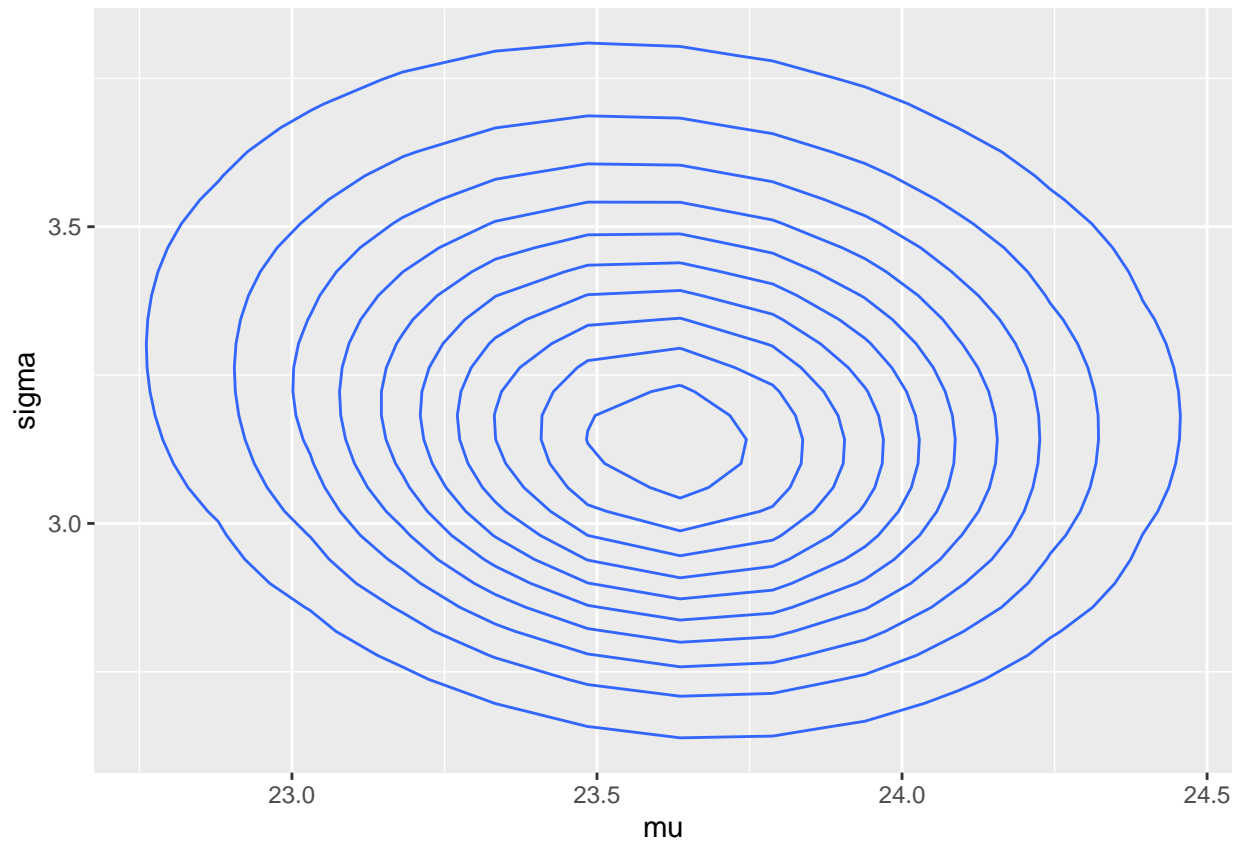
Use either a grid approximation or MCMC to estimate the posterior distribution of μ and σ for this model, with data given by the 67 actors who portrayed 17-year olds.

Make a plot of the posterior distribution.

Just by looking at your plot, give an estimate of both μ and σ at the mode of the posterior.

Solution At the mode, μ is between 23.5 and 23.7, while σ is between 3.0 and 3.2.

```
library(ggplot2)
grid <- expand.grid(list(mu=seq(15,30,length=100), sigma=seq(1,5,length=100)))
likelihood <- function(mu,sigma) {
  prod(dnorm(a17$factor_age,mu,sigma))
}
gapprox <- grid |> mutate(prior = dnorm(mu,17,2) * dgamma(sigma,1,1),
  likelihood = purrr::map2_dbl(mu,sigma,likelihood),
  post = prior*likelihood)
gapprox |> ggplot(aes(x=mu,y=sigma,z=post)) + geom_contour(show.legend = FALSE)
```



```
norm_prog <- "
data {
  int<lower=0> N;
  vector[N] y;
}
parameters {
  real mu;
  real<lower=0> sigma;
}
model {
  y ~ normal(mu, sigma);
  mu ~ normal(17,2);
  sigma ~ gamma(1,1);
}
"

#norm_sim <- rstan::stan(model_code = norm_prog, data = list(y = a17$factor_age, N = nrow(a17)))
#rstan::extract(norm_sim) |> as.data.frame() |>
# ggplot(aes(x=mu,y=sigma)) + geom_point(alpha=0.2)
```