Dial-Jackson-homework4

Jackson Dial

2023-01-31

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.4     v dplyr   1.0.7
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   2.0.1     v forcats 0.5.1

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

Q1

A

1 a) $\text{Beta}(\alpha,\beta) = \dfrac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}$ $\quad B(\alpha,\beta) = \dfrac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$

$\psi = \log(\theta/(1-\theta))$

Find inverse of ψ, derivate, Abs. value, then multiply by pdf w/ $h(x)$ plugged in.

$\log(\theta/(1-\theta)) = \psi \rightarrow (\theta/(1-\theta)) = \exp(\psi)$

$\dfrac{\theta}{1-\theta} = \exp(\psi) \rightarrow \theta = (1-\theta)e^{\psi} \rightarrow \theta = e^{\psi} - \theta e^{\psi} \rightarrow$

$\theta + \theta e^{\psi} = e^{\psi} \rightarrow \theta(1+e^{\psi}) = e^{\psi} \rightarrow \theta = \dfrac{e^{\psi}}{1+e^{\psi}}$

$$\rightarrow \frac{\left(\dfrac{e^{\psi}}{1+e^{\psi}}\right)^{\alpha-1}\left(1-\dfrac{e^{\psi}}{1+e^{\psi}}\right)^{\beta-1}}{B(\alpha,\beta)}$$

$$\frac{d}{d\psi}\frac{e^{\psi}}{1+e^{\psi}} \rightarrow \frac{(1+e^{\psi})e^{\psi} - e^{\psi}(e^{\psi})}{(1+e^{\psi})^2} = \frac{e^{\psi} \cancel{+(e^{\psi})^2} \cancel{-(e^{\psi})^2}}{1+2e^{\psi}+e^{\psi 2}}$$

$$\rightarrow \frac{e^{\psi}}{(1+e^{\psi})^2}$$

$$\frac{\left(\frac{e^{\psi}}{1+e^{\psi}}\right)^{\alpha-1}\left(1-\frac{e^{\psi}}{1+e^{\psi}}\right)^{\beta-1}}{B(\alpha,\beta)} \cdot \frac{e^{\psi}}{(1+e^{\psi})^{2}}$$

$$\rightarrow \frac{\frac{e^{\psi(\alpha-1)}}{(1+e^{\psi})^{(\alpha-1)}} \cdot \frac{1^{(\beta-1)}}{(1+e^{\psi})^{(\beta-1)}}}{B(\alpha,\beta)} \quad \frac{e^{\psi}}{(1+e^{\psi})^{2}}$$

$$\rightarrow \frac{e^{\psi\alpha}}{e^{\psi}(1+e^{\psi})^{\alpha-1}\cdot(1+e^{\psi})^{\beta-1}\cdot B(\alpha,\beta)} \quad \frac{e^{\psi}}{(1+e^{\psi})^{2}}$$

$$\rightarrow \frac{e^{\psi\alpha}\,\cancel{(1+e^{\psi})}\,\cancel{(1+e^{\psi})}}{\cancel{e^{\psi}}(1+e^{\psi})^{\alpha}(1+e^{\psi})^{\beta}\cdot B(\alpha,\beta)} \cdot \frac{\cancel{e^{\psi}}}{\cancel{(1+e^{\psi})^{2}}}$$

$$\rightarrow \frac{e^{\psi\alpha}}{(1+e^{\psi})^{\alpha}(1+e^{\psi})^{\beta}B(\alpha,\beta)}$$

$$\rightarrow \frac{1}{B(\alpha,\beta)} \cdot \frac{e^{\psi\alpha}}{(1+e^{\psi})^{\alpha+\beta}}$$

$\alpha=\beta=1$

$$\rightarrow \frac{e^{\psi}}{(1+e^{\psi})^{2}}$$

B

$$\text{Gamma} \sim \frac{b^{a}}{\Gamma(a)} x^{a-1} e^{-bx}$$

$\psi=\log(\theta)$

$$\theta = e^{\psi} \qquad \frac{d}{d\psi} = e^{\psi}$$

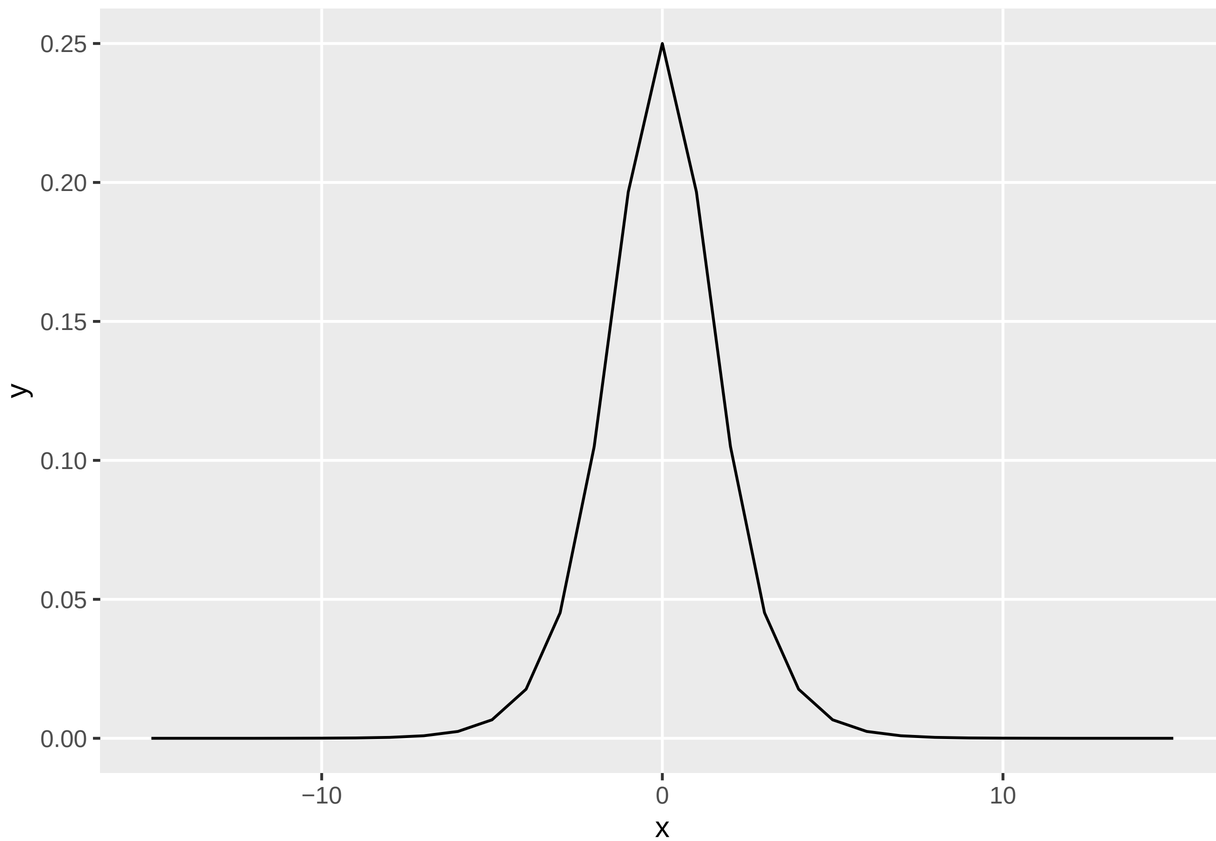
$$\frac{b^{a}}{\Gamma(a)} e^{\psi(a-1)} e^{-be^{\psi}} \cdot e^{\psi}$$

$$\rightarrow \frac{b^{a}}{\Gamma(a)} e^{\psi(a-1)-be^{\psi}+\psi} \rightarrow \frac{b^{a}}{\Gamma(a)} e^{\psi a-\psi-be^{\psi}+\psi}$$

$$\rightarrow \frac{b^{a}}{\Gamma(a)} e^{\psi a-be^{\psi}}$$

$a=b=1$

```r
x <- -15:15
y <- (exp(x) / (exp(x) + 1)^2)
dat <- cbind.data.frame(x,y)
ggplot(dat, aes(x = x, y = y))+
  geom_line()+
  theme(panel.grid.minor = element_blank())
```



B

$$\rightarrow \frac{1}{B(\alpha,\beta)} \cdot \frac{e^{\psi\alpha}}{(1+e^{\psi})^{\alpha+\beta}} \qquad \alpha=\beta=1$$

$$\rightarrow \frac{e^{\psi}}{(1+e^{\psi})^{2}}$$

B

$$\text{Gamma} \sim \frac{b^{a}}{\Gamma(a)} x^{a-1} e^{-bx} \qquad \psi=\log(\theta)$$

$$\theta = e^{\psi} \qquad \frac{d}{d\psi} = e^{\psi}$$

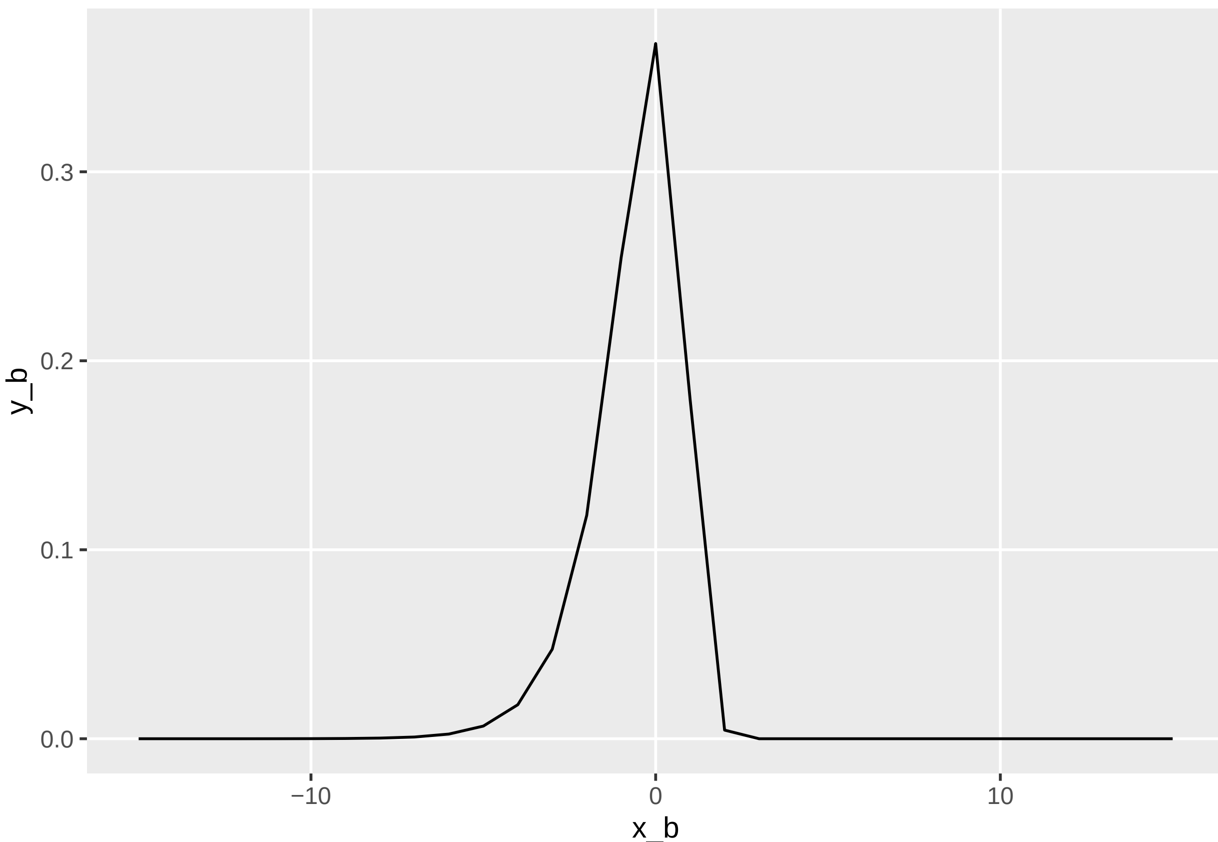
$$\frac{b^{a}}{\Gamma(a)} e^{\psi(a-1)} e^{-be^{\psi}} \cdot e^{\psi}$$

$$\rightarrow \frac{b^{a}}{\Gamma(a)} e^{\psi(a-1)-be^{\psi}+\psi} \rightarrow \frac{b^{a}}{\Gamma(a)} e^{\psi a-\psi-be^{\psi}+\psi}$$

$$\rightarrow \frac{b^{a}}{\Gamma(a)} e^{\psi a-be^{\psi}} \qquad a=b=1$$

$\rightarrow \quad e^{\psi} - e^{\psi}$

```r
x_b <- -15:15
y_b <- exp(x_b-exp(x_b))
dat_b <- cbind.data.frame(x_b,y_b)
ggplot(dat_b,aes(x_b,y_b))+
  geom_line()+
  theme(panel.grid.minor = element_blank())
```



Q2

Task 4

```r
#set seed for reproducibility
set.seed(123)

# spurters
x = c(18, 40, 15, 17, 20, 44, 38)

# control group
y = c(-4, 0, -19, 24, 19, 10, 5, 10,
      29, 13, -9, -8, 20, -1, 12, 21,
      -7, 14, 13, 20, 11, 16, 15, 27,
      23, 36, -33, 34, 13, 11, -19, 21,
```

```
      6, 25, 30,22, -28, 15, 26, -1, -2,
      43, 23, 22, 25, 16, 10, 29)

# store data in data frame
iqData = data.frame(Treatment = c(rep("Spurters", length(x)),
                                  rep("Controls", length(y))),
                                  Gain = c(x, y))
```

```
prior = data.frame(m = 0, c = 1, a = 0.5, b = 50)
findParam = function(prior, data){
  postParam = NULL
  c = prior$c
  m = prior$m
  a = prior$a
  b = prior$b
  n = length(data)
  postParam = data.frame(m = (c*m + n*mean(data))/(c + n),
                c = c + n,
                a = a + n/2,
                b =  b + 0.5*(sum((data - mean(data))^2)) +
                  (n*c *(mean(data)- m)^2)/(2*(c+n)))
  return(postParam)
}
# Find parameters for each of the distributions we are wanting to pull from
postS = findParam(prior, x)
postC = findParam(prior, y)

# library(lestat)
# postDistS <- normalgamma(postS[1], postS[2], postS[3], postS[4])
#
# sample(postDistS, size = 10)
#
# plot(normalgamma(3,4,5,6))
```

```
set.seed(456)
rnormgamma <- function(n, mu, lambda, alpha, beta) {
  if (length(n) > 1)
    n <- length(n)
  tau <- rgamma(n, alpha, beta)
  x <- rnorm(n, mu, sqrt(1/(lambda*tau)))
  data.frame(tau = tau, x = x)
}



postSsamps <- rnormgamma(1000000,postS[,1], postS[,2], postS[,3], postS[,4])
postCsamps <- rnormgamma(1000000,postC[,1], postC[,2], postC[,3], postC[,4])
names(postSsamps) <- c("S_tau", "S_x")
names(postCsamps) <- c("C_tau", "C_x")

all_x <- cbind(postSsamps, postCsamps)

final <- all_x %>%
```

```
  mutate(s_greater_c = case_when(
    S_x > C_x ~ 1,
    TRUE ~ 0
  ))

posterior_prob <- mean(final$s_greater_c)
#The posterior probability is 0.970645
```

This means the probability of the mean QI for spurters has a 97% probability of being higher than the mean QI score for the control group.

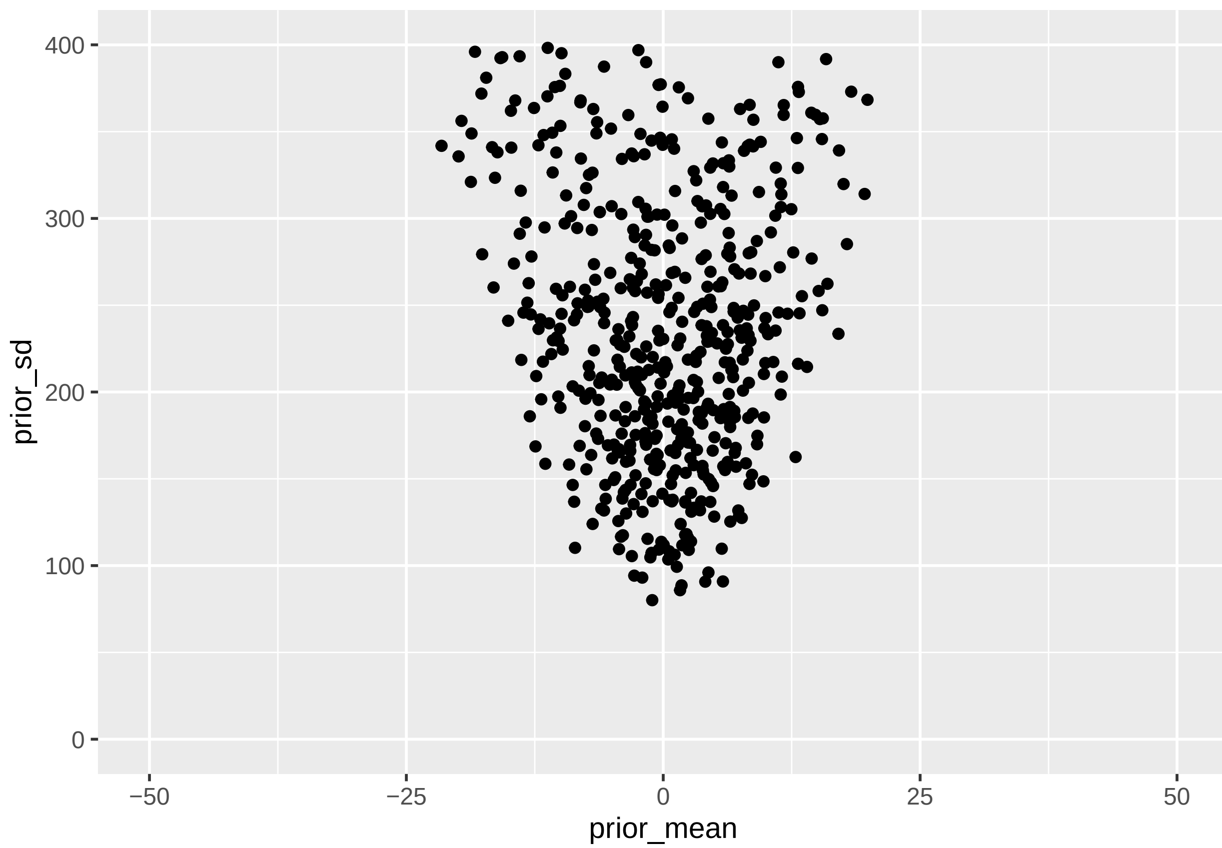
Task 5

```
#Priors are also Normal-Gamma
library(tictoc)
set.seed(789)
prior_df_loop <- data.frame()
tic()
for(i in 1:1000){
  prior_samps <- rnormgamma(1000, 0, 1, .5, 50)[,2]
  prior_mean <- mean(prior_samps)
  prior_sd <- sd(prior_samps)
  prior_row <- cbind(prior_mean, prior_sd)
  prior_df_loop <- rbind(prior_df_loop, prior_row)
}
toc()
```

```
## 1 sec elapsed
```

```
ggplot(prior_df_loop,aes(x = prior_mean, y = prior_sd))+
  geom_point()+
  xlim(-50,50)+
  ylim(0,400)
```

```
## Warning: Removed 477 rows containing missing values (geom_point).
```



It looks like the mean is about 0 and the standard deviation is likely between 20 and 100