hw 2 3303

Tyler Poelking

1/22/2018

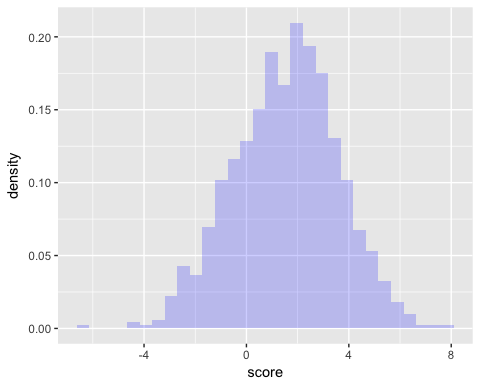
Problem 1: See Handwritten Page

Problem 2: a) See Handwritten Page for Marginal Density Formula

# marginal probabilities  
prob <- c( .5, .5 )  
  
# conditional means and standard deviations. Set sigma = 2 for both cases  
mu <- c(1, 2)  
sigma <- c( 2, 2)

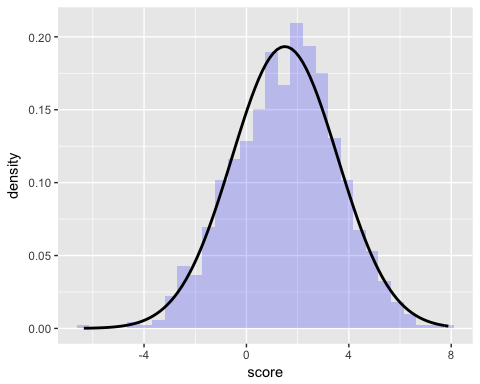
# number of samples  
n.samples <-1000  
  
# simulate genotypes  
param <- sample( c( "Equals 1", "Equals 2"), n.samples, replace = T, prob = prob )  
  
# simulate scores conditional on param  
score.sim <- data.frame(score = rnorm(n.samples, mu[as.numeric(as.factor(param))],   
 sigma[as.numeric(as.factor(param))]))  
  
# plot histogram of the samples  
p <- ggplot(score.sim, aes( x = score)) +  
 geom\_histogram( aes(y = ..density.. ), fill = "blue", alpha = 0.2)  
   
  
p

## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



# function that calculates the marginal density of scores  
marginal.density <- function(x)  
 {  
 prob[1] \* dnorm( x, mu[1], sigma[1] ) +  
 prob[2] \* dnorm( x, mu[2], sigma[2] )  
 }  
  
# add marginal density to the histogram  
p + stat\_function( fun = marginal.density, lwd = 1)

## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



1. What is P(θ=1|y=1)P(θ=1|y=1) when σ=2 ?

#calculate P(y=1| param =1)  
pnorm(q= 1, mean=1, sd = 2)

## [1] 0.5

#calculate P(y=1| param =2)  
pnorm(q= 1, mean=2, sd = 2)

## [1] 0.3085375

print("When sd is 2: ")

## [1] "When sd is 2: "

result = (.5\*.5)/(.5\*.5 + .3085\*.5)  
result

## [1] 0.6184292

1. Describe how P(θ=1|y=1) changes as you vary σσ.

#INCREASE sd  
  
#calculate P(y=1| param =1)  
a = pnorm(q= 1, mean=1, sd = 5 )  
#calculate P(y=1| param =2)  
b = pnorm(q= 1, mean=2, sd = 5 )  
  
result = (a\*.5)/(a\*.5 + b\*.5)  
print("When sd is 5 (larger): ")

## [1] "When sd is 5 (larger): "

result

## [1] 0.5430413

#Decrease sd  
  
#calculate P(y=1| param =1)  
a = pnorm(q= 1, mean=1, sd = .1 )  
#calculate P(y=1| param =2)  
b = pnorm(q= 1, mean=2, sd = .1 )  
  
result = (a\*.5)/(a\*.5 + b\*.5)  
print("When sd is .1 (smaller): ")

## [1] "When sd is .1 (smaller): "

result

## [1] 1

When sd gets larger, the width of the distribution increases, making it more likely that y = 1 when θ is 2. Thus, the P(θ=1|y=1) is less. When sd gets smaller, the width of the distribution decreases, making it less likely that y = 1 when θ is 2. Thus, the P(θ=1|y=1) is more.