MATH 5310: Probability
Fall 2024
Homework 6
Due before 11:59pm on Thursday 11/14

Part1:

An important extension to how we have used distributions creating new distributions that are *mixtures* of other distributions. For example, a *mixture distribution* created from two normal distributions means that the random variable of interest for a proportion of the population can be described with one normal distribution's probability density function and the remaining part of the population can be described with a different normal distribution's probability density function.

Here is a nice general discussion of mixture distributions

https://www.statisticshowto.com/mixture-distribution/

Here is a more technical discussion of mixture distributions

https://fkorona.github.io/ATML/2017_2/Lecture_notes/03E_Mixtures.pdf

Here is an example specific to mixing two normal distributions

https://ajeyvenkataraman.wordpress.com/2020/05/30/some-notes-on-mixture-of-normal-distributions/

Here is a more technical discussion more specific to mixing normal distributions:

https://www.value-at-risk.net/mixtures-of-distributions/

(1) Write the pdf for a mixture of two normal distributions. Define all notation.

(2) Imagine you want to simulate the outcome for one person from population where the outcome is from a mixture distribution of two normal pdfs. Explain which of the following is the correct strategy:

   i. Use a Bernoulli distribution (based on the mixing proportion) to randomly assign your simulated person to one of the two subpopulations, and then use that subpopulations normal pdf to simulate their outcome

   ii. Randomly simulate an outcome from each of the two normal pdfs, and then create the final simulated outcome equal to the weighted average of these two with the weights being the mixing proportions (that is, the final outcome equals the sum of each simulated outcome times the proportion of the population in that subgroup)

(3) OPTIONAL – The classic example for mixing two normal distributions is eruption times of the Old Faithful geyser at Yellowstone National Park. Further, Bayes is an often used tool in estimating the parameters of mixture distributions.

If you want to see an example of how estimation of the individual normal distributions comprising the mixture is done with Bayes, download the following R code and work through the first two parts (through the top of page 7)

https://www.jstatsoft.org/article/view/v032i06

Part 2: Casella and Berger Problems:

3.17

3.23

3.24 (d) only

3.25 and 3.26 (a) only