Hints for Assignment 1

Question 1

c)

ii) The summation of all the values must be appended to the end of the bayes box. For instance, if your bayes box is a matrix “m”, you can find a sum of all the columns using “colSums(m)” and append it to the matrix “m” using “rbind”

iv) The posterior mean (the point estimate for p):

Classical methods often report the maximum likelihood estimator (MLE) or the method of moments estimator (MOME) of a parameter. In contrast, Bayesian approaches often use the posterior mean. The definition of the posterior mean is given by

$$E(\theta|\mathbf{y}) = \int \theta \, p(\theta|\mathbf{y}) \, d\theta$$

You can find it by summing the product of the ‘posterior’ and ‘p’ across the entire data frame.

post.mean = sum(df$p* df$posterior) in a d.f of the Bayesian box.

v) A (1-alpha)100% BCI (Bayesian credible interval) for “p”: The Bayesian Credible Interval is just a central interval made from the posterior. So, a 95% bci will contain .95 of the area. The Bayesian set estimates are called *credible sets*, which is also known as *credible intervals*. This is analogous to the concept of confidence intervals used in classical statistics. Given a posterior distribution $p(\theta|\mathbf{y})$, A is a credible set for θ if,

$$P(\theta \in A|\mathbf{y}) = \int_A p(\theta|\mathbf{y}) d\theta$$

For example, you can construct a 95%credible set for $\theta$by finding an interval, A, over which $\int_A p(\theta|\mathbf{y}) = 0.95$.

To compute it in R you can do the following:

- Find a cumulative summation of the posterior: cp=cumsum(df$posterior)
- Find the lower limit: L= max(which(cp<alpha/2))
- The upper limit: U = min (which (cp > 1-alpha/2))
- Get those values from ‘p’ to get the interval: BCI = df$p[c(L, U)].

Question 3

Definition. Let X be a discrete random variable with probability mass function $f(x)$ and support S. Then:

$$M(t) = E(e^{tX}) = \sum_{x \in S} e^{tx} f(x)$$

is the **moment generating function of** X as long as the summation is finite for some interval of t around 0. That is, $M(t)$ is the moment generating function ("**m.g.f.**") of X if there is a positive number h such that the above summation exists and is finite for $-h < t < h$.

a.) https://www.youtube.com/watch?time_continue=157&v=ez_vq23xWrQ This video will point you to the answer.
b.) https://canvas.ou.edu/files/31348325/download?download_frd=1. This is a document that will help you with the b) and c) question.

Question 4

A skeleton of the function expected is as follows :

mynorm <- function (mu,sigma,a,b,alpha)

#calculate the probability for given parameter (a and b) values using pnorm

##create normal curve with mu +/- 3 standard deviations using function: curve and dnorm

#Shade area of interest

xcurve = seq(a,b, length=1000) # length is arbitrary

ycurve = dnorm(xcurve, mu,sigma)

polygon(c(a,xcurve,b), c(0,ycurve,0), col="blue") # this adds the area

Now place legend on the above plot | Code: text(a+4,0.02,substitute(paste("Prob=",prob), list(prob = prob)))

#compute alpha/2 and 1-alpha/2 quantiles using qnorm and make a list }

obj = mynorm(mu = 10, sigma = 8, a = 8, b = 11,alpha=0.1)

The output should look as follows: