STT 380

In-Class Activity 19

A data science model trained on data predicts that 30% of customers will buy a product, but the uncertainty in the model is such that the standard deviation of this proportion is 0.2.

1. Use the EstBetaParams function at the top of page 39 to convert these values into parameters for a Beta distribution.
    a. estBetaParams <- function(mu, var) {
    b. alpha <- ((1 - mu) / var - 1 / mu) * mu ^ 2
    c. beta <- alpha * (1 / mu - 1)
    d. return(params = list(alpha = alpha, beta = beta))
    e. }
    f. estBetaParams(0.3, 0.2^2)
    g. $alpha
    h. [1] 1.275
    i. $beta
    j. [1] 2.975
2. Use qbeta to construct a 95% confidence interval for the true proportion.
    a. alpha <- 1.275
    b. beta <- 2.975
    c. print(qbeta(0.025,alpha,beta)) = 0.02026175
    d. print(qbeta(0.975,alpha,beta)) = 0.7469699
    e. x <- seq(0,1, by = 0.01)
    f. y <- dbeta(x,alpha,beta)
3. A soft release of the product is made to 10 customers, resulting in 5 sales. Construct the beta posterior from this information.
    a. release <- 10
    b. sales <- 5
    c. alpha_post <- alpha + sales
    d. beta_post <- beta + release - sales
    e. print(alpha_post) = 6.275

f. print(beta_post) = 7.975

g. y_post <-dbeta(x,alpha_post,beta_post)

4. Calculate the new mean and standard deviation for the posterior.

   a. mean(y_post) = 0.990099

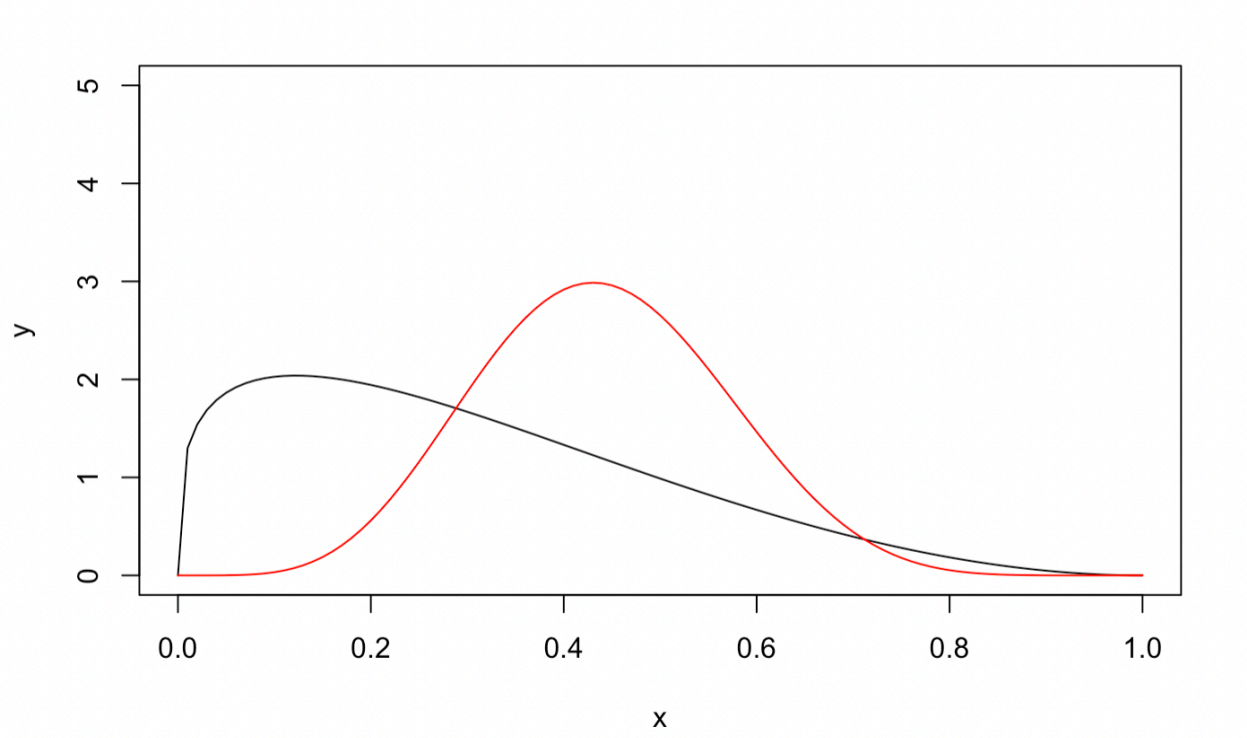
   b. sd(y_post) = 1.087197

5. Plot the posterior and prior beta distributions together. How do they compare?

   a. plot(x,y, ylim = c(0,5), type="l")

   b. lines(x,y_post,col="red")

   c. The black line is the prior distribution; it hits a maximum y value at a lower x value than the other graph. the posterior graph has a bigger max value.



6. The product will be profitable if at least 30% of customers purchase the product. Based on the posterior, what is the probability that at least 30% will purchase the product?
    a. greater_or_equal <- length(y_post[y_post>=0.3])
    b. total <- length(y_post)
    c.
    d. prob <- greater_or_equal/total
    e. prob = 0.5544554
    f. So 55%