Lab. 4 Bayes' Box and R.

In this lab you will need to be able to make the Bayes' Box for Binomial experiments.

Objectives:

- Review classical methods
- Learn how the components of a Bayes' box are constructed using R.
- Learn how to construct a Bayes' box using an R function.
- Learn how to make graphical output from the Bayes box
- Learn how to make both point and interval estimates from a Bayes' box
- Learn how to automate the process with an all inclusive function.

Output:

- Make an Rmd document for this lab (ALL formulae need to be put in latex)
- At the conclusion render into the 3 file types
- Upload the 4 files to the server.

Bayes' Rule

$$p(\theta|x) = \frac{p(\theta)p(x|\theta)}{\sum_{\Theta} p(\theta)p(x|\theta)}$$

Experiment

A biased coin is tossed

- $n = 10$ times
- $x = 3$ successes (Heads)
- Each trial is independent
- There are only two possible outcomes per trial
- $P(H) = \theta$ is constant

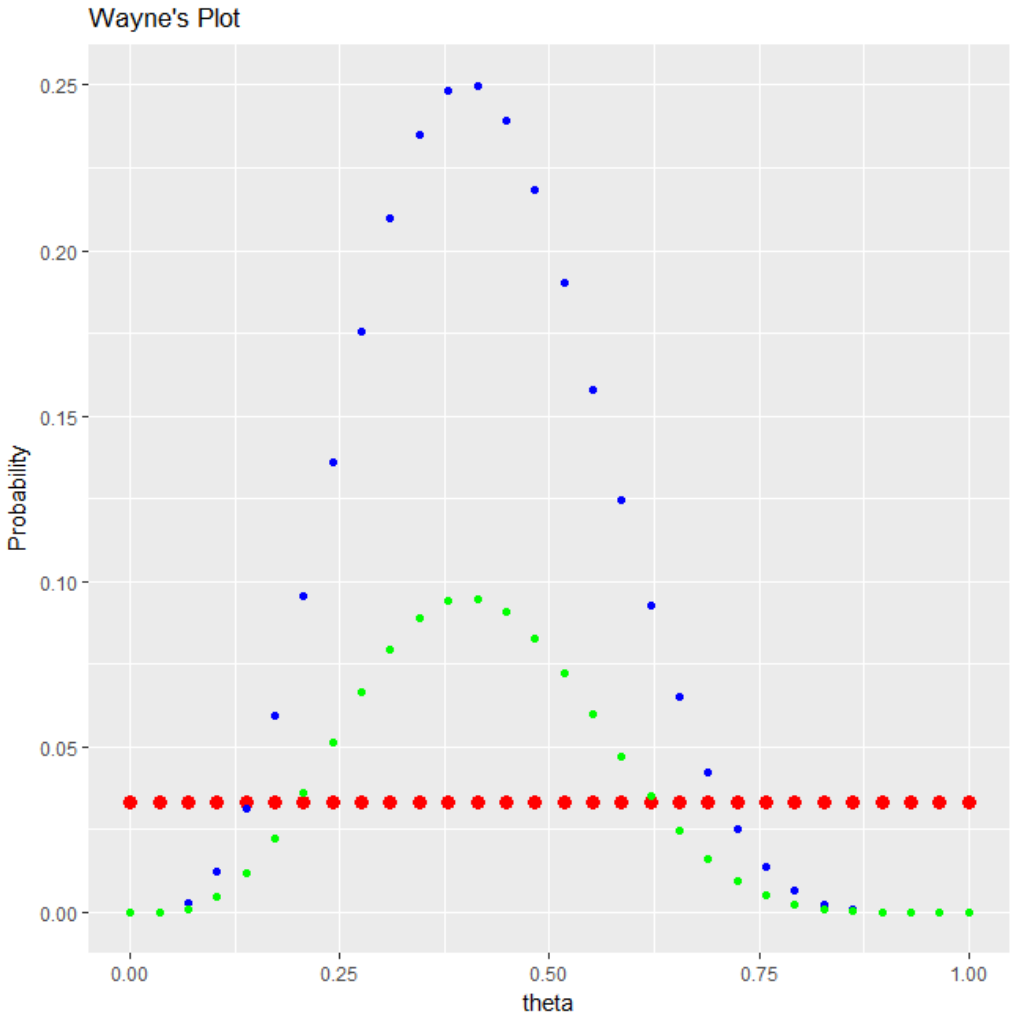
Tasks:

1. Classical estimates for θ
    a. What is the classical point estimate for θ?
    b. Using $\hat{\theta} \pm 1.95\sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}}$
        i. find the classical 95% ci for θ
        ii. interpret the interval (see http://onlinestatbook.com/2/estimation/confidence.html)
        iii. How many samples were used to make the interval?
        iv. How many samples are assumed to give the interpretation?

2. Bayesian methods – Bayes' rule. We will assume discrete values of θ and a uniform prior on the values of θ
    a. Suppose we take theta = seq(0, 1, length = 20)
    b. Plot the prior using base R.
    c. Write down the formula for the likelihood in this case – you can substitute the n and x so that it is a function of θ
    d. Plot the likelihood over the discrete values of theta. Place in your document.
    e. Now add the prior plot onto the same graph. Place in your document.
    f. We need the posterior in order to obtain the third and last plot – what formula will we use to obtain this? Write it down.
3. Bayes' Box – below is Bayes' box – it will need more rows!

theta	P(theta)	P(x|theta)	h(theta)	P(theta|x)
Total				

    a. Make a matrix in R that will hold the 5 column vectors. Show here.
    b. Rewrite Bayes' theorem in terms of h = prior*lik. Show formula here.
    c. Now make a plot similar to the one below (except it will be for the data given above) from the matrix using ggplot
        i. Put your name on it
        ii. Make sure it has the right colors
        iii. Make an improvement to the plot by changing the x ticks to represent the number of theta values.



4. Now make a function called mybinpost() that will do the following:
    a. Take arguments n, x, theta, alpha
        i. n = number of trials
        ii. x = number of successes
        iii. theta = vector of possible theta values (sequence from 0 to 1)
        iv. alpha/2 = tail probability (used to obtain 1-alpha Bayesian credible interval)
        v. Assume uniform prior
    b. The function will produce
        i. the above plot.
        ii. A list containing
            1. the Bayes box as a matrix
            2. the Bayes box in latex (package “xtable”)
            3. the number of trials n
            4. the number of successes x
            5. the Bayesian point estimate
            6. The 95% Bayesian credible interval (BCI)
5. Now use your function and include all output into your document for the following invocations

a. mybinpost(n=20,x=12,theta = seq(0,1,length=50), alpha=0.05)
b. mybinpost(n=50, x=25, theta=seq(0,1,length=30), alpha=0.1)