STA 601/360 Homework 2

Yifei Wang

12 September, 2019

Exec 1. Hoff 3.1 Sample survey: Suppose we are going to sample 100 individuals from a county (of size much larger than 100) and ask each sampled person whether they support policy Z or not. Let $Y_i = 1$ if person i in the sample supports the policy, and $Y_i = 0$ otherwise.

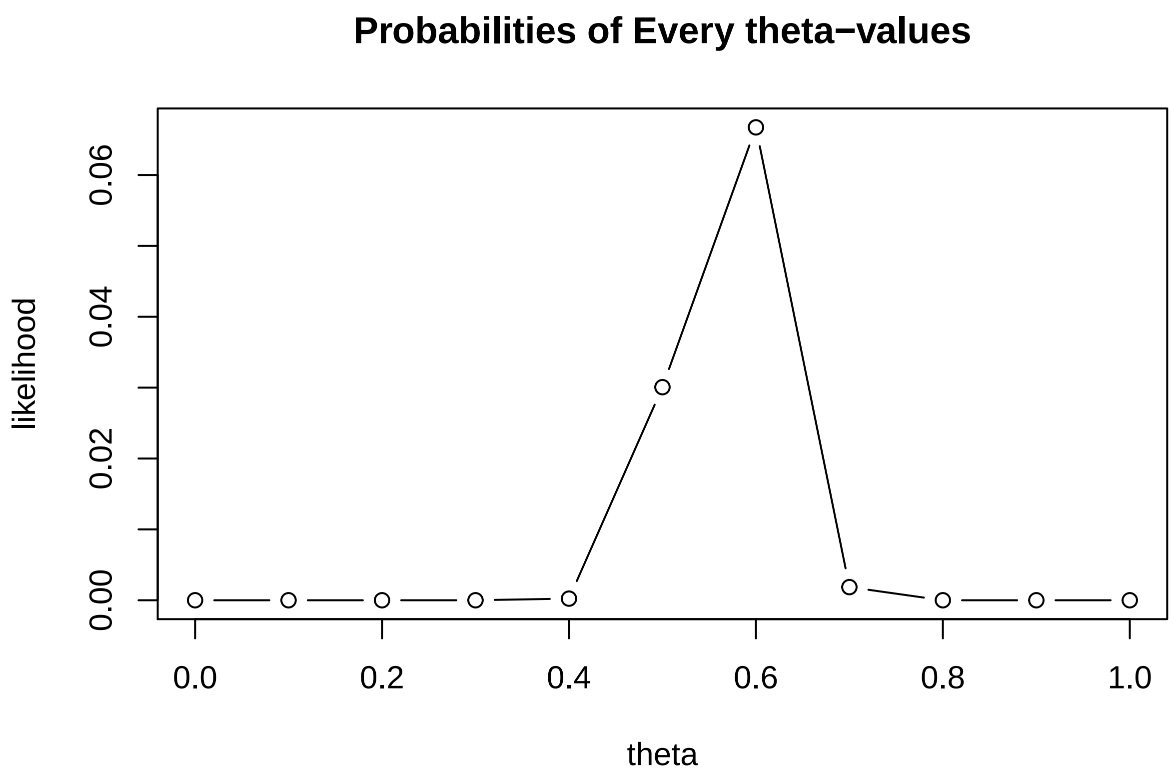
a) Assume $Y_1, \ldots, Y_{100}$ are, conditional on θ, i.i.d. binary random variables with expectation θ. Write down the joint distribution of $Pr(Y_1 = y_1, ..., Y_{100} = y_{100} \mid \theta)$ in a compact form. Also write down the form of $Pr(\sum Y_i = y \mid \theta)$.

$$P(Y_1 = y_1, ..., Y_{100} = y_{100} \mid \theta) = \theta^{\sum y_i}(1-\theta)^{(100-\sum y_i)}$$

$$P(\sum Y_i = y \mid \theta) = \binom{100}{y}\theta^{\sum y_i}(1-\theta)^{(100-\sum y_i)}$$

(b) For the moment, suppose you believed that $\theta \in \{0.0, 0.1, ..., 0.9, 1.0\}$. Given that the results of the survey were $\sum_{i=1}^{100} Y_i = 57$, compute $\sum Y_i = 57 \mid \theta$ for each of these 11 values of θ and plot these probabilities as a function of θ.

```
N = 100
y_sum = 57
theta = seq(0, 1, 0.1)
likelihood = choose(N, y_sum) * theta^y_sum * (1 - theta)^(N-y_sum)

plot(theta, likelihood, 'b', main = "Probabilities of Every theta-values")
```



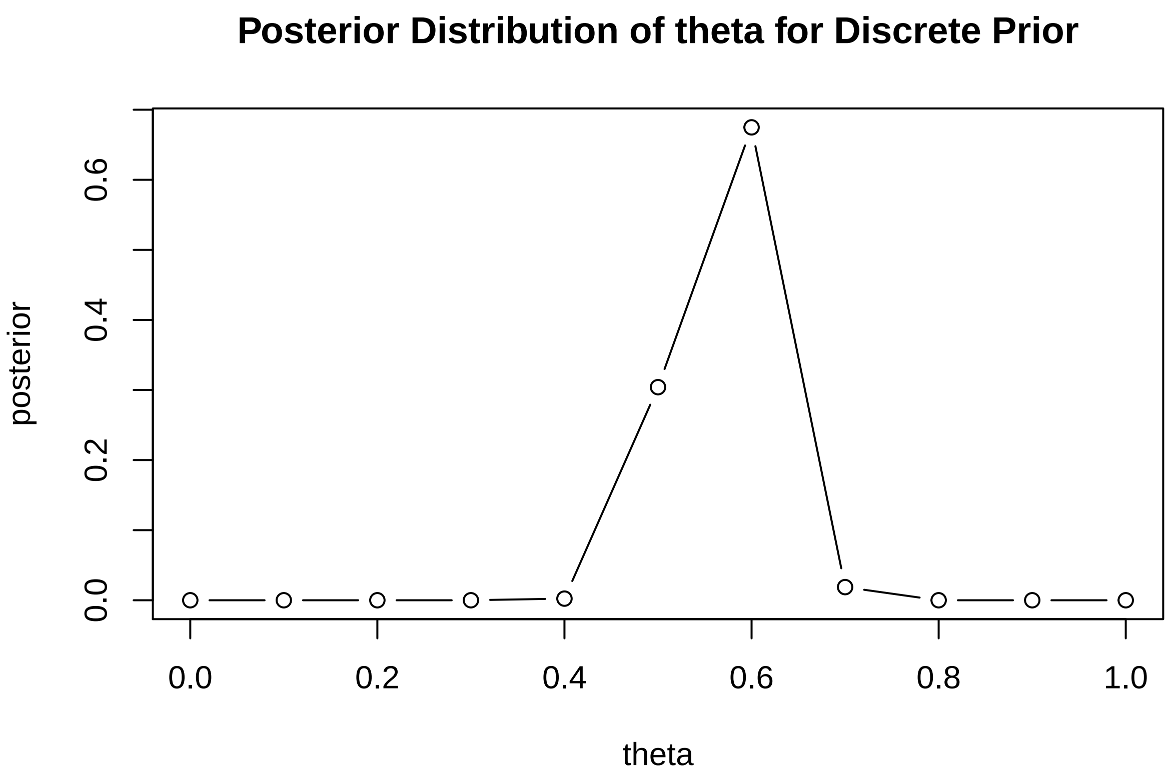
(c) Now suppose you originally had no prior information to believe one of these θ-values over another, and so $Pr(\theta = 0.0) = Pr(\theta = 0.1) =$ ůůů $= Pr(\theta = 0.9) = Pr(\theta = 1.0)$. Use Bayes' rule to compute $p(\theta \mid \sum_{i=1}^{n} Y_i = 57)$ for each θ-value. Make a plot of this posterior distribution as a function of θ.

$$Pr(\theta = 0.0) = Pr(\theta = 0.1) = \ldots = Pr(\theta = 0.9) = Pr(\theta = 1.0) = \frac{1}{11}$$

$$p(\theta \mid \sum_{i=1}^{n} Y_i = 57) = \frac{p(\sum_{i=1}^{n} Y_i = 57 \mid \theta) p(\theta)}{p(\sum_{i=1}^{n} Y_i = 57)} \propto p(\sum_{i=1}^{n} Y_i = 57 \mid \theta)$$

We need the posterior distribution to be integrated to 1. So we have,

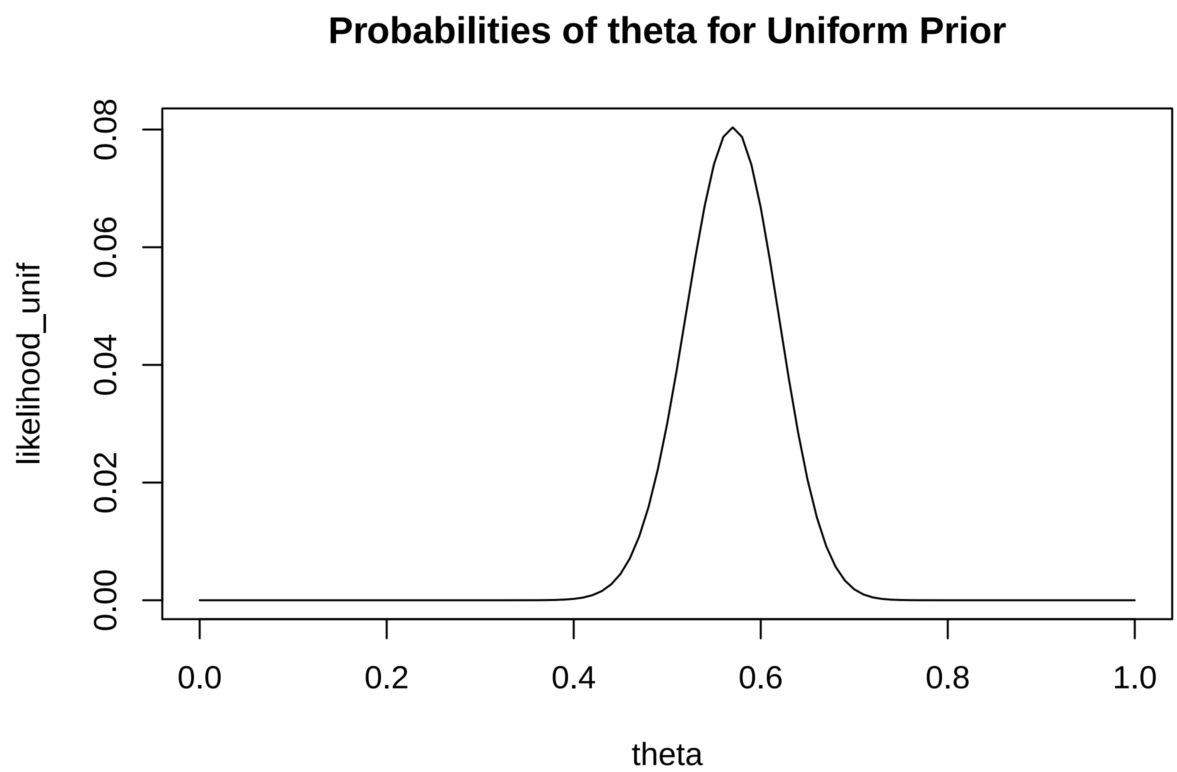
```r
posterior = likelihood / sum(likelihood)

plot(theta, posterior, 'b', main = "Posterior Distribution of theta for Discrete Prior")
```



(d) Now suppose you allow θ to be any value in the interval $[0, 1]$. Using the uniform prior density for θ, so that $p(\theta) = 1$, plot $p(\theta) \times Pr(\sum_{i=1}^{n} Y_i = 57 \mid \theta)$ as a function of θ

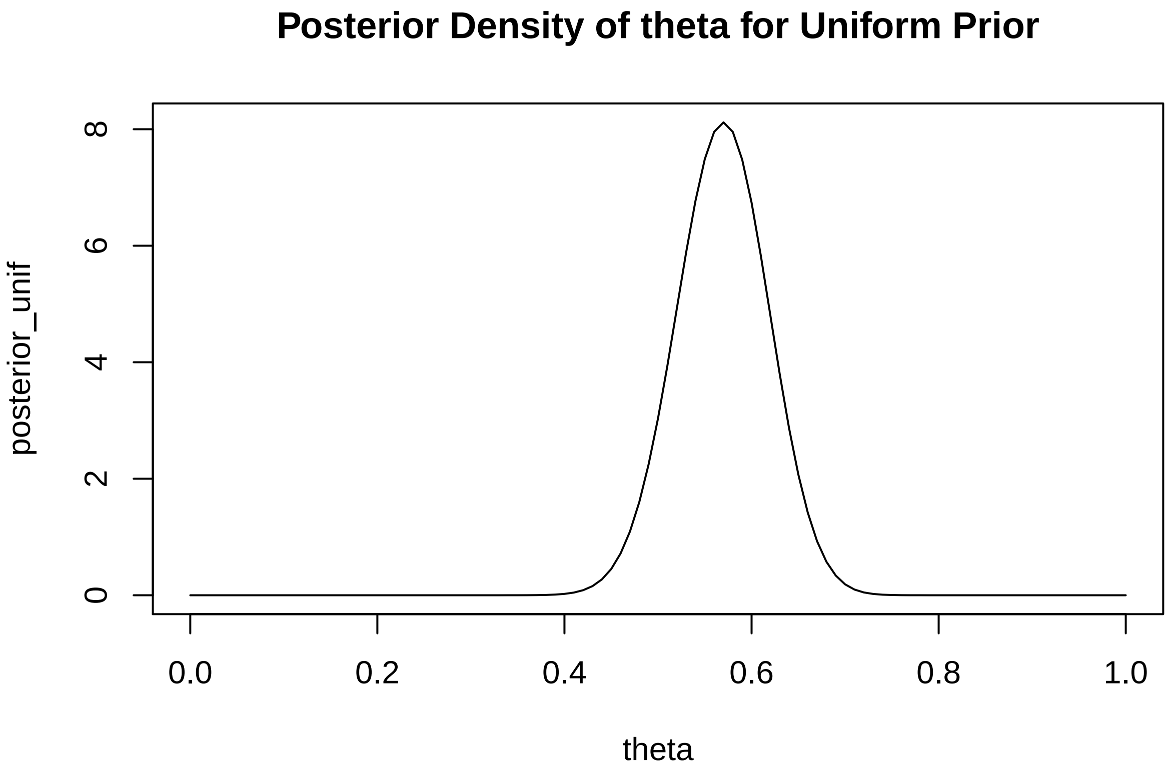
$$p(\theta) \times p(\sum_{i=1}^{n} Y_i = 57 \mid \theta) = \binom{100}{57} \theta^{57} (1-\theta)^{43}$$

```
theta = seq(0, 1, 0.01)
y_sum = 57
N = 100
likelihood_unif = choose(N, y_sum) * theta^y_sum * (1 - theta)^(N-y_sum)

plot(theta, likelihood_unif, 'l', main = "Probabilities of theta for Uniform Prior")
```



(e) As discussed in this chapter, the posterior distribution of θ is $Beta(1+57, 1+100-57)$. Plot the posterior density as a function of θ. Discuss the relationships among all of the plots you have made for this exercise.

```
theta = seq(0, 1, 0.01)
y_sum = 57
N = 100
a = 1
b = 1
posterior_unif = beta(a+y_sum, b+N-y_sum)^(-1) * theta^y_sum * (1 - theta)^(N-y_sum)

plot(theta, posterior_unif, 'l', main = "Posterior Density of theta for Uniform Prior")
```



We have 2 different priors, discrete uniform and continuous uniform. For the discrete one, the posterior probabiliy mass function has the identical shape with the sample method $P(\sum Y_i \mid \theta)$ (likelihood). For the continuous one, the posterior probability density function has the identical shape with the product of sample method and priors (because $p(\theta) \equiv 1$). Also, the shape of posterior probability function of the continuous one is similar to the discrete one, because both priors are uniformly and equivantly spaned from 0 to 1.

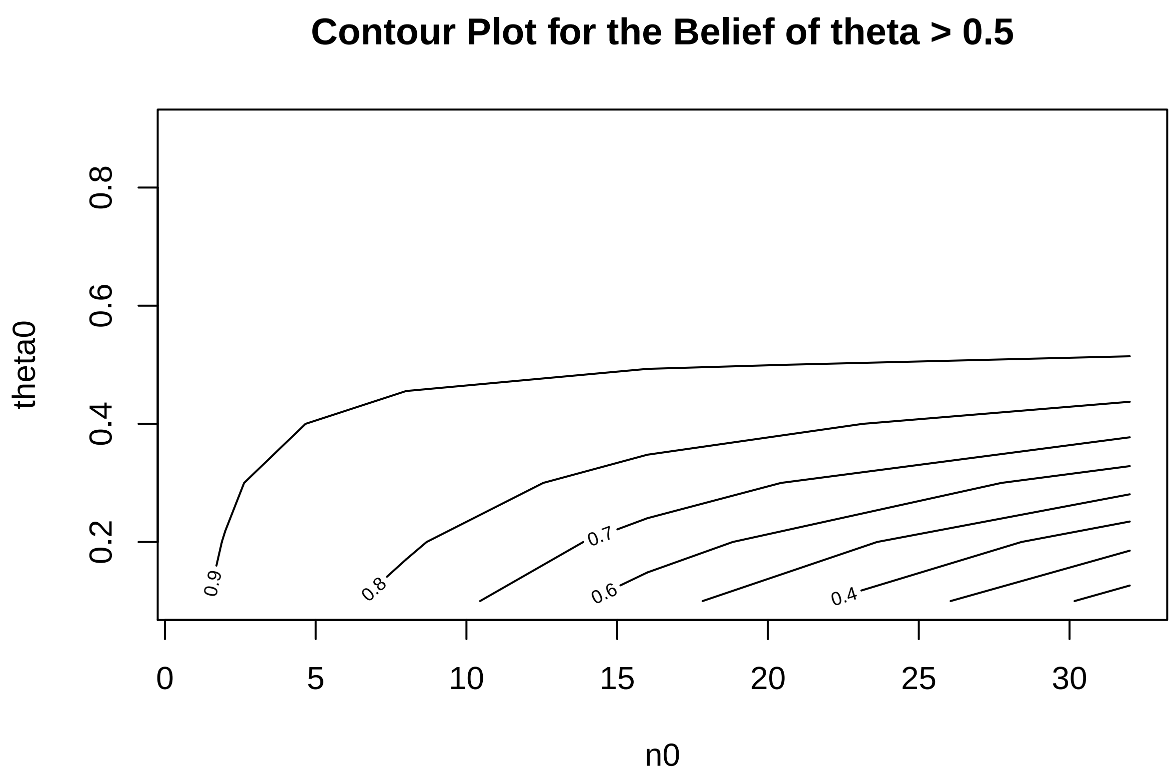
Exec 2. Hoff 3.2 Sensitivity analysis: It is sometimes useful to express the parameters a and b in a beta distribution in terms of $\theta_0 = \frac{a}{a+b} and n_0 = a + b$, so that $a = \theta_0 n_0$ and $b = (1-\theta_0)n_0$. Reconsidering the sample survey data in Exercise 3.1, for each combination of $\theta_0 \in \{0.1, 0.2, ..., 0.9\}$ and $n_0 \in \{1, 2, 8, 16, 32\}$ find the corresponding a, b values and compute $Pr(\theta > 0.5 \mid \sum Y_i = 57)$ using a $Beta(a, b)$ prior distribution for θ. Display the results with a contour plot, and discuss how the plot could be used to explain to someone whether or not they should believe that $\theta > 0.5$, based on the data that $\sum_{i=1}^{n} Y_i = 57$.

(a, b)	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	(0.1, 0.9)	(0.2, 0.8)	(0.3, 0.7)	(0.4, 0.6)	(0.5, 0.5)	(0.6, 0.4)	(0.7, 0.3)	(0.8, 0.2)	(0.9, 0.1)
2	(0.2, 1.8)	(0.4, 1.6)	(0.6, 1.4)	(0.8, 1.2)	(1.0, 1.0)	(1.2, 0.8)	(1.4, 0.6)	(1.6, 0.4)	(1.8, 0.2)
8	(0.8, 7.2)	(1.6, 6.4)	(2.4, 5.6)	(3.2, 4.8)	(4.0, 4.0)	(4.8, 3.2)	(5.6, 2.4)	(6.4, 1.6)	(7.2, 0.8)
16	(1.6, 14.4)	(3.2, 12.8)	(4.8, 11.2)	(6.4, 9.6)	(8.0, 8.0)	(9.6, 6.4)	(11.2, 4.8)	(12.8, 3.2)	(14.4, 1.6)
32	(3.2, 28.8)	(6.4, 25.6)	(9.6, 22.4)	(12.8, 19.2)	(16.0, 16.0)	(19.2, 12.8)	(22.4, 9.6)	(25.6, 6.4)	(28.8, 3.2)

$$
\begin{aligned}
P(\theta \mid \sum Y_i = 57) &= \frac{P(\sum Y_i = 57 \mid \theta)\ P(\theta)}{P(\sum Y_i = 57)} \\
&= \frac{\binom{100}{57}\theta^{57}(1-\theta)^{43} Beta(a,b)}{P(\sum Y_i = 57)} \\
&= Beta(a+y, b+n-y) \\
&= Beta(a+57, b+43)
\end{aligned}
$$

$$
Pr(\theta > 0.5 \mid \sum Y_i = 57) = Pr(Beta(a+57, b+43) > 0.5)
$$

```
theta0 = seq(0.1, 0.9, 0.1)
n0 = c(1, 2, 8, 16, 32)
a = outer(n0, theta0)
b = outer(n0, 1-theta0)

Pr = 1 - pbeta(0.5, a+y_sum, b+N-y_sum)
contour(x=n0, y=theta0, z=Pr, xlab="n0", ylab="theta0",
        main="Contour Plot for the Belief of theta > 0.5")
```



Given the fact that $\sum_{i=1}^{100} Y_i = 57$, this plot could be interpretated as how the prior knowledge influence on the posterior belief of whether θ is greater than 0.5. n_0 could be interpretated as the number of prior sample size and θ_0 could be treated as the prior θ. The value of θ_0 will affect the value of θ, by dragging the likelihood function towards the prior. Given $n = 100$, this "dragging" effect is determined by the values of n_0 and the differences between θ and θ_0. The greater the differences, or the greater the n_0, the greater the effect People should probably believe that $\theta > 0$ when the combination of $theta_0$ and n_0 falls into the upper part of the 0.5 contour line.