Bayesian Learning Lab 1

Mohamed Ali

2023-05-06

Question 1 Daniel Bernoulli

First we calculate the mean and the standar deviation from the below equation to compare the with sample means and standar deviations of θ as function of the accumulating number of drawn values.

$$E(\theta|y) \;=\; a/a+b$$

$$E(\theta|y) \;=\; ab/(a+b)^2(a+b+1)$$

A ### Figure 1.1

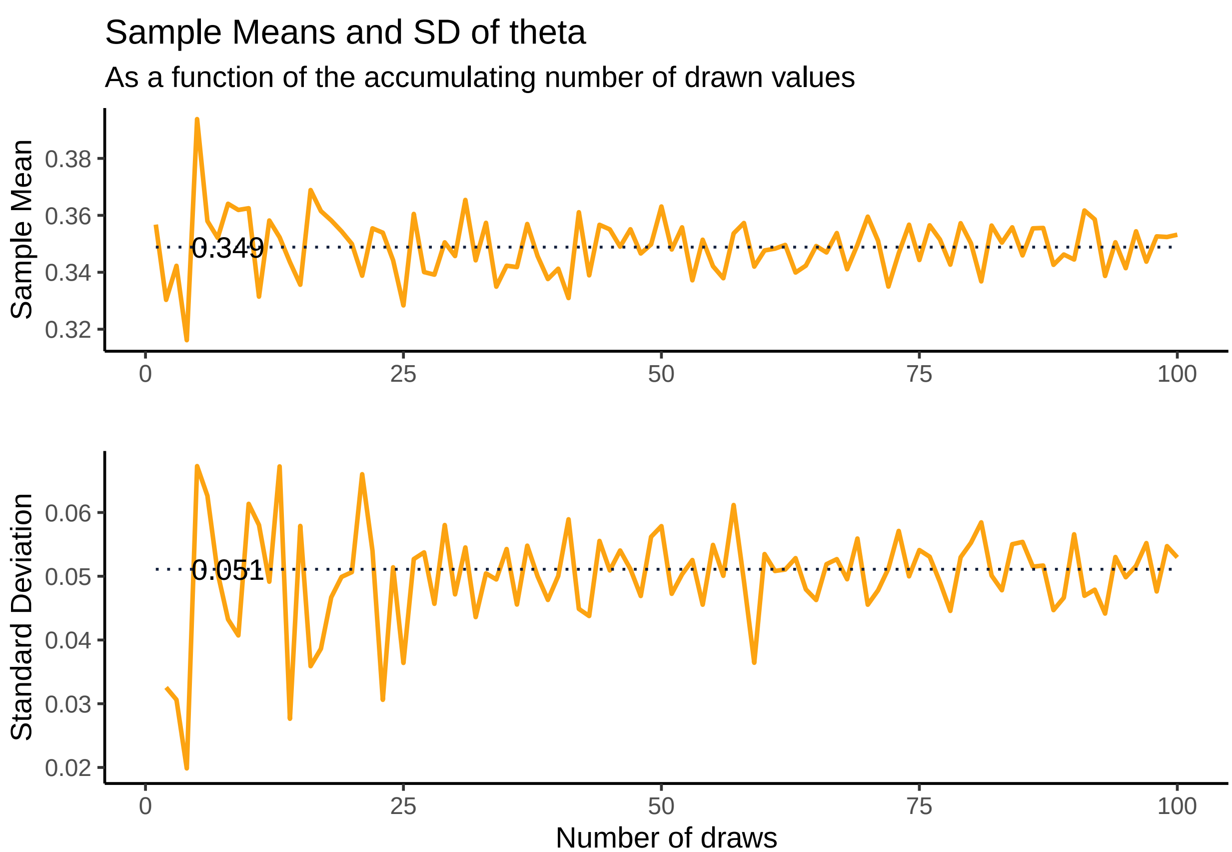
The figure below shows how the values of the Mean and Sd converges to the true values as the number of draws grows large

```
ber_fun<-function(n_d,n,s,a,b) {
  #Initial Value of the function givan from the question
  t_n = n
  s   = s
  f   = n-s
  a   = a
  b   = b
  #Beta(alpha+s,beta+s)
  a_new = a+s
  b_new = b+f
  Mean_true= a_new/(a_new+b_new)
  #we take the sqrt to get the Sd insted of the Var
  Sd_true=  sqrt((a_new*b_new)/(((a_new+b_new)^2) * (a_new+b_new+1)))
  # Beta(alpha+s,beta+s)
  mean_theta = c()
  sd_theta = c()
  n_draws = 1:n_d
  #For loop to fill the values of mean and sd based on the draws
  for (i in 1:n_d){
    mean_theta[i]=mean(rbeta(i,a_new,b_new))
    sd_theta[i]=sd(rbeta(i,a_new,b_new))}
  #Binding everything togther
  df<-cbind.data.frame(n_draws,mean_theta,sd_theta)
  #Plot of the mean
  mean<-ggplot(df,aes(x=n_draws))+geom_line(aes(y=mean_theta)
                                           , color='#FCA311', size=.8)+
              geom_line(aes(y=Mean_true), color='#14213D',linetype=3)+
              annotate(geom = "text", x = 8, y = Mean_true,
              label = paste0(format(round(Mean_true, 3), nsmall = 3)))+
```

```
            labs(title = 'Sample Means and SD of theta ',
        subtitle = 'As a function of the accumulating number of drawn values',
                      x= ' ', y='Sample Mean')+ theme_classic()
  #Plot of the Sd
  sd<-ggplot(df,aes(x=n_draws))+geom_line(aes(y=sd_theta),
                                        color='#FCA311', size=.8)+
            geom_line(aes(y=Sd_true), color='#14213D',linetype=3)+
            annotate(geom = "text", x = 8, y = Sd_true,
            label = paste0(format(round(Sd_true, 3), nsmall = 3)))+
            labs(x= 'Number of draws', y='Standard Deviation')+
            theme_classic()
  #grid.arrange(mean,sd)
  return(grid.arrange(mean,sd))
}


ber_fun(100,70,22,8,8)
```

```
## Warning: Removed 1 row(s) containing missing values (geom_path).
```



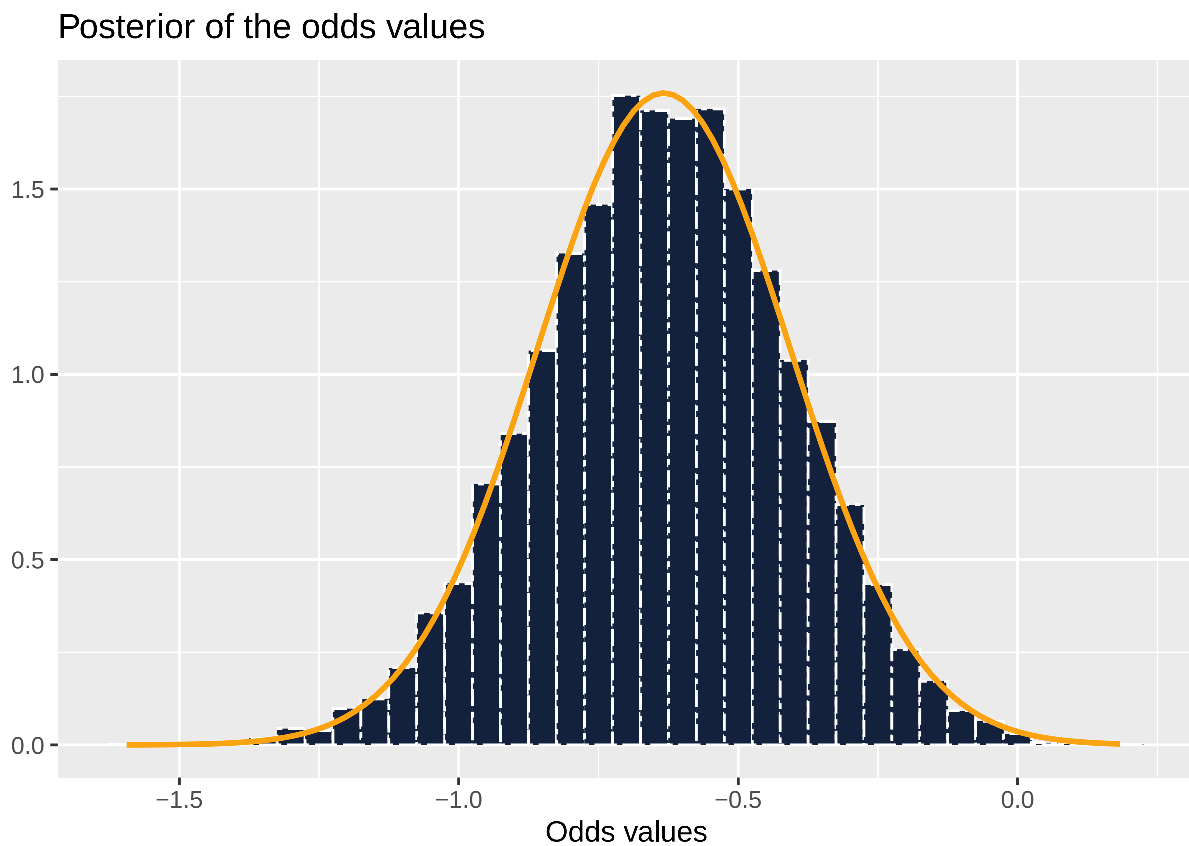
B

First we find the value from the beta posterior using the function *pbeta* we get the value of theta which can be used to compute the probability that a random variable from a beta distribution is less than or equal to a given value, or greater than a given value, depending on the value of the lower.tail argument. in our case we use lower.tail as False because we need the values grater than.

```
## [[1]]
## [1] "random values from the posterior with the given condition 0.86"
##
## [[2]]
## [1] "The exact value from the Beat posterior 0.828593587338586"
```

C

Now we want to draw 10000 random values from the posterior of the odds



Question 1 Daniel Bernoulli

First we find the value of tau, the question assume that we have 8 randomly selected persons thus we have n = 8 and we have the value of mu = 3.6 Given, to find the value of tau we use this equation:

$$\tau^2 = \sum_1^n (log\ y_i - \mu)^2/n$$

'. *Note* A non-informative prior is a prior distribution that is chosen to express little to no prior information about the parameters. Non-informative priors are often chosen to avoid introducing bias or strong assumptions into the model, and to allow the data to have a greater influence on the posterior distribution.

Examples of non-informative priors include the uniform distribution, the Jeffreys prior, and the reference prior. It's important to note that a non-informative prior is not necessarily a prior with no information at all, but rather one that expresses a minimal amount of information that is consistent with our knowledge and beliefs before observing the data. In practice, the choice of prior distribution often depends on the specific problem and the available prior knowledge.

in the question we have inverse chi distribution is our posterior, the inverse-chi-squared distribution (or inverted-chi-square distribution[1]) is a continuous probability distribution of a positive-valued random variable. It is closely related to the chi-squared distribution. It arises in Bayesian inference, where it can be used as the prior and posterior distribution for an unknown variance of the normal distribution.

the inverse chi distribution is The inverse chi-square distribution with degrees of freedom n and scale parameter s^2 is closely related to the inverse gamma distribution with shape parameter $\alpha = \mathrm{n}/2$ and scale parameter $\beta = 1/(2s^2)$. In fact, if X is Inv-χ^2(n,s^2) distributed, then Y = (ns^2)/X is Inv-γ(α,β) distributed, and vice versa. then we use the function rinvgamma from r and we change on the pramters shape and scale.

```
## Warning: Removed 143 rows containing non-finite values (stat_bin).
```

```
## Warning: Removed 2 rows containing missing values (geom_bar).
```