MA 471: Lab Assignment 04

Due on Monday, August 28, 2017

A. K. Dey

Sai Teja Talasila
140123040

Contents

Problem 1

R Code :

```
A = read.table("d-csp0108.txt",header = T)
A[,c(2,3)] = log(1+A[,c(2,3)])
A = as.data.frame(A)
colMeans(A[,c(2,3)])
m = colMeans(A[,c(2,3)])
s = c(sd(A[,2]),sd(A[,3]))

count = 0
for(i in 1:1000)
{
     Z = rnorm(1000,m[1],s[1])
     mu = mean(Z)
     sigma = sd(Z)
     L = mu-1.96*sigma/sqrt(length(Z))
     U = mu+1.96*sigma/sqrt(length(Z))
     if(L < m[1] && m[1] < U)
          count = count + 1
}
cat("Coverage Probability : ",count/1000)
```

```
Coverage Probability :  0.948
```

Problem 2

R Code :

```
data = read.table('d-csp0108.txt',header=T)
attach(data)
C_log = log(C+1)
SP_log = log(SP+1)

h = function(p)
{
    log(p/(1-p))
}

h_prime = function(p)
{
    1/(p*(1-p))
}

h_inv = function(p)
{
    exp(p)/(1 + exp(p))
}

conf_interval = function(X,dist="normal")
{
    mu = mean(X)
    sigma = sd(X)
    n = length(X)
    if(dist=="normal")
    {
        return(c(mu-1.96*sigma/sqrt(n),mu+1.96*sigma/sqrt(n)))
    }
    else if(dist=="bernoulli")
    {
        p_hat = mu
        return(c(p_hat-1.96*sqrt(p_hat*(1-p_hat)/n),p_hat+1.96*sqrt(p_hat*(1-p_hat)/n)
            ))
    }
    else if(dist=="normalized_bernoulli")
    {
        p_hat = mu
        if(p_hat==0)
          return(c(0,0))
        L = h(p_hat) - (1.96*(h_prime(p_hat))*sqrt(p_hat*(1-p_hat)/n))
        U = h(p_hat) + (1.96*(h_prime(p_hat))*sqrt(p_hat*(1-p_hat)/n))
        return(c(h_inv(L),h_inv(U)))
    }
}

interval = conf_interval(C)
cat("Confidence Interval of C: [",interval[1],",",interval[2],"]\n")
mu_c = mean(C)
```

```
sigma_c = sd(C)
n = length(C)
N = 1000
count = 0
for(i in 1:N)
{
    sample = rnorm(n,mu_c,sigma_c)
    interval = conf_interval(sample)
    if(interval[1]<=mu_c && mu_c<=interval[2])
        count = count + 1
}

coverage_prob_C = count/N
cat("Coverage Probability of C = ",coverage_prob_C)

p = 0.1
sample_size = c(20,50,100,1000)
for(size in sample_size)
{
    count = 0
    false_count = 0
    for(i in 1:1000)
    {
        sample = rbinom(size,1,p)
        interval = conf_interval(sample,dist="bernoulli")
        if(interval[1]<0 || interval[2]>1)
            false_count = false_count + 1
        if(interval[1]<=p && p<=interval[2])
            count = count + 1
    }
    cat("\nFor sample size = ",size)
    cat("\nCoverage Probability = ",count/1000)
    cat("\nNo. of intervals outside parameter space = ",false_count)
}

for(size in sample_size)
{
    count = 0
    false_count = 0
    for(i in 1:1000)
    {
        sample = rbinom(size,1,p)
        interval = conf_interval(sample,dist="normalized_bernoulli")
        if(interval[1]<0 || interval[2]>1)
            false_count = false_count + 1
        if(interval[1]<=p && p<=interval[2])
            count = count + 1
    }
    cat("\nFor sample size = ",size)
    cat("\nCoverage Probability = ",count/1000)
    cat("\nNo. of intervals outside parameter space = ",false_count)
}
```

Problem 2 continued on next page...

```
Confidence Interval of C: [ -0.001735229 , 0.0009793866 ]
Coverage Probability of C =  0.947
For sample size =  20
Coverage Probability =  0.884
No. of intervals outside parameter space =  744
For sample size =  50
Coverage Probability =  0.894
No. of intervals outside parameter space =  234
For sample size =  100
Coverage Probability =  0.941
No. of intervals outside parameter space =  8
For sample size =  1000
Coverage Probability =  0.947
No. of intervals outside parameter space =  0
For sample size =  20
Coverage Probability =  0.831
No. of intervals outside parameter space =  0
For sample size =  50
Coverage Probability =  0.971
No. of intervals outside parameter space =  0
For sample size =  100
Coverage Probability =  0.961
No. of intervals outside parameter space =  0
For sample size =  1000
Coverage Probability =  0.944
No. of intervals outside parameter space =  0
```