Solution_HW5

Wei-Rou,Xu

2023-11-15

Given the sepal width of "iris" data, use the following steps to show if the mean of sepal width is 3. Assume the sepal widths are random samples from a normal distribution.

(a) Use the maximum likelihood estimation method to estimate the model parameters of the normal distribution.

```
###### MME ######
data.set=iris$Sepal.Width
obj.norm <- function(par){
  mu <- par[1]
  sigma2 <- par[2]
  y <- (mean(data.set)-mu)^2+(var(data.set)-sigma2)^2
  return(y)
}

est.normal <- optim(c(3.057 ,0.189),obj.norm)
est.normal$par
```

```
## [1] 3.0573334 0.1899794
```

```
###### MLE ######
likelihood.normal<-function(par,data){
  mu=par[1]
  sig2=par[2]

  if(sig2>0){ ## Ensure that the variance is greater than zero
    joint=dnorm(data, mean = mu, sd = sqrt(sig2))
    return(-sum(log(joint)))
  }else{
    return(1e+5)
  }
}

est.normal <- optim(c(3.057 ,0.189),likelihood.normal,data=data.set)
est.normal$par
```

```
## [1] 3.0573193 0.1887255
```

$\hat{\mu}$=3.0573, $\hat{\sigma}^2$=0.1887

(b) According to the question "if the mean of sepal width is 3", what is the estimate for the quantity of interest?

$\hat{\mu}$=3.0573

(c) Construct the 95% confidence interval for the the true quantity of interest by bootstrapping.

```
##### Non-parametric Bootstrapping sampling distribution #####
likelihood.mu=function(par,sig2,data){
  mu=par
  joint=dnorm(data, mean = mu, sd = sqrt(sig2))
  return(-sum(log(joint)))
}

set.seed(20231116)
data.collected <- data.set
est.collected <- optim(3, likelihood.mu, sig2 = est.normal$par[2], data = data.collected,
                       method = "Brent" ,lower = -10, upper = 10)
est <- est.collected$par

np.mu.hat.100<-rep(NA,1000)
for(j in 1:1000){
  dataY <- sample(data.collected, length(data.collected), replace = TRUE)
  estY <- optim(3,likelihood.mu, sig2 = est.normal$par[2], data = dataY,
                method = "Brent", lower = -10, upper = 10)
  np.mu.hat.100[j] <-estY$par[1]
}

quantile(np.mu.hat.100, c(0.025, 0.975))
```

```
##     2.5%    97.5%
## 2.989317 3.128017
```

The 95% confidence interval for μ by bootstrapping is [2.989,3.128].

```
Ln <- est + qnorm(0.025) * sqrt(var(np.mu.hat.100))
Un <- est + qnorm(0.975) * sqrt(var(np.mu.hat.100))
c(Ln, Un)
```

```
## [1] 2.987987 3.126679
```

(d) Based on the 95% confidence interval in (c), how would you conclude the question "if the mean of sepal width is 3"?

Because 3 is in the 95% confidence interval for μ, we would not reject the conclusion that the mean of sepal width is 3.

(e) If I use a one-sample t-test to test if $H_0 : \mu = 3$, is the conclusion as the same as the result in (d)?

```
t.test(data.collected,mu=3)
```

```
##
##  One Sample t-test
##
## data:  data.collected
## t = 1.611, df = 149, p-value = 0.1093
## alternative hypothesis: true mean is not equal to 3
## 95 percent confidence interval:
##  2.987010 3.127656
## sample estimates:
## mean of x
##  3.057333
```

According to one-sample t-test, p-value=0.1093>0.05, we would not reject the null hypothesis, which means that at the significance level of 0.05, we do not have enough evidence to say mean of sepal width is not equal to 3.