STAT3023 Statiscal Inference

Lab Week 12: quiz

Tutor: Wen Dai

SID: 470408326

School of Mathematics and Statistics

The University of Sydney

Semester 2, 2021

1. We perform out setup code here.

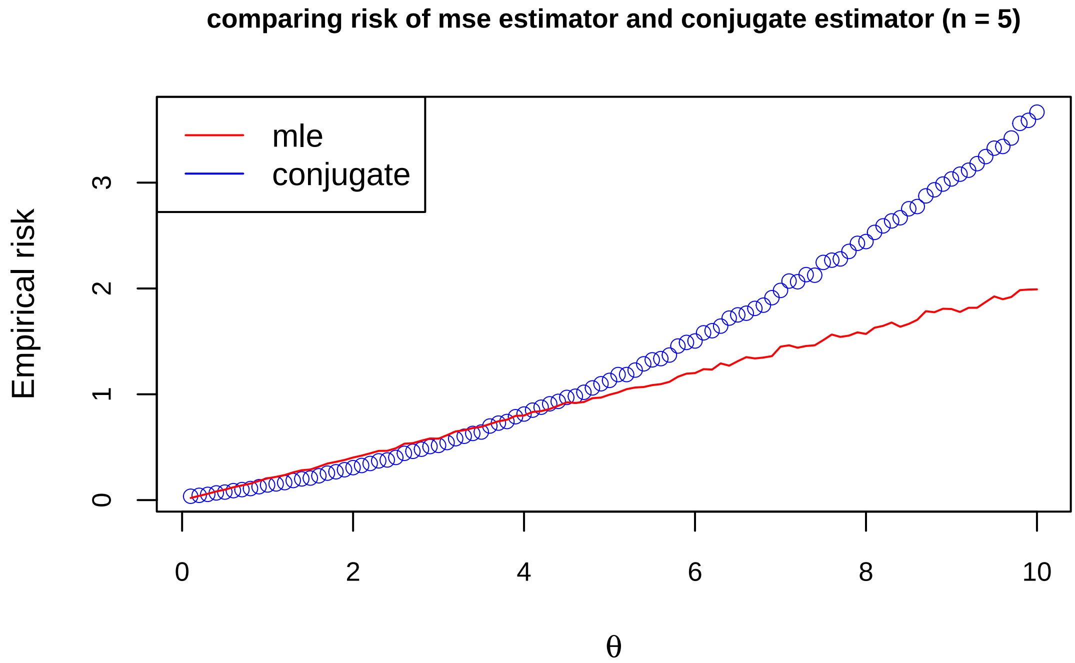
```
# Define a vector of 100 theta values from 0.1 to 10
th = 1:100/10
# the length of th which is 100 in this case
L.th = length(th)
# sample size of 5 so far
n = 5
# n times the mean squared error for the MLE and the conjugate. these are vectors
nMSE.mle = nMSE.conj = 0
# the number of simulation iterations for each theta value
B = 10000
```

2. Compute the mean squared errors for for each θ value

```
# We iterate through all the theta values (what the outer loop is for)
for (j in 1:L.th) {
    # intialise the vectors
    mle = conj = 0
    # For each sample...
    for (i in 1:B) {
        # We draw a size of 'n' from the poisson distribution with mean th[j]
        observations = rpois(n, th[j])
        # We compute the estimators and store them in the ith position of the vectors mle
        mle[i] = mean(observations)
        conj[i] = (n * mean(observations) + 1)/(n + 1)
    }
    # compute n times the average mean squared error and store it in the jth position of o
    # vectors
    nMSE.mle[j] = mean((mle - th[j])^2)
    nMSE.conj[j] = mean((conj - th[j])^2)
}
```

3. We now plot n times the mean squared error for the conjugate estimator in blue and we n times the mean squared error for the mle estimator in red. We plot them both against our θ values.

```
plot(th, nMSE.conj, col = "blue", ylab = "Empirical risk", xlab = expression(theta), main
    cex.main = 0.85, cex.axis = 0.8, lwd = 0.5)
lines(th, nMSE.mle, col = "red")
legend("topleft", legend = c("mle", "conjugate"), col = c("red", "blue"), lty = c(1, 1))
```

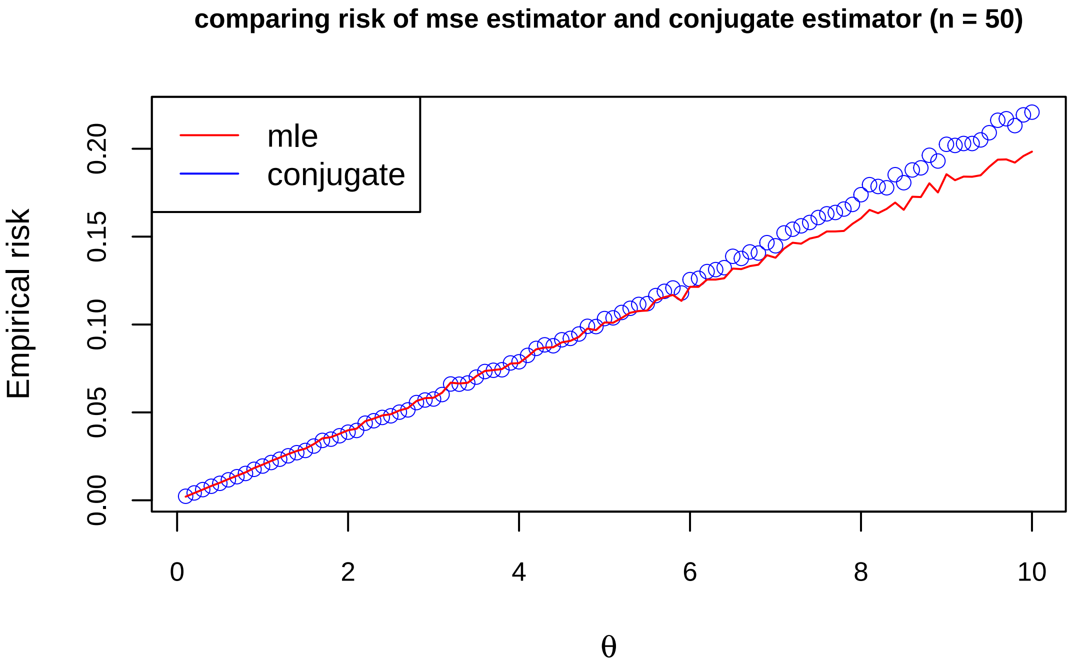


4. d_{conj} seems to do better than d_{MLE} for some θ values (that is, in θ values in the lower region) because for the expression

$$\frac{n\overline{X}+1}{n+1} = \frac{n}{n+1}\overline{X} + \frac{1}{n+1}$$

it is closer to the true θ values for smaller values of θ

5.
```
# change the size of n
n = 50
# We iterate through all the theta values (what the outer loop is for)
for (j in 1:L.th) {
    # intialise the vectors
    mle = conj = 0
    # For each sample...
    for (i in 1:B) {
        # We draw a size of 'n' from the poisson distribution with mean th[j]
        observations = rpois(n, th[j])
        # We compute the estimators and store them in the ith position of the vectors mle
        mle[i] = mean(observations)
        conj[i] = (n * mean(observations) + 1)/(n + 1)
    }
    # compute n times the average mean squared error and store it in the jth position of o
    # vectors
    nMSE.mle[j] = mean((mle - th[j])^2)
    nMSE.conj[j] = mean((conj - th[j])^2)
}
plot(th, nMSE.conj, col = "blue", ylab = "Empirical risk", xlab = expression(theta), main =
    cex.main = 0.85, cex.axis = 0.8, lwd = 0.5)
lines(th, nMSE.mle, col = "red")
legend("topleft", legend = c("mle", "conjugate"), col = c("red", "blue"), lty = c(1, 1))
```



6. What we see is that:

$$d_{\text{conj}} = \frac{n\overline{X}+1}{n+1} = \frac{\overline{X}+\frac{1}{n}}{1+\frac{1}{n}} \to \overline{X}$$

As $n \to \infty$