EXERCISE 4

WEIYU LI

1. Read the first 4 chapters of *Empirical Likelihood* by Owen.

2. Apply the empirical likelihood method to draw the $50\%, 90\%, 95\%, 99\%$ confidence regions of *law* dataset in *bootstrap* package, and compare them with confidence regions of normal distribution. (Use scel.R on http://statweb.stanford.edu/ owen/empirical/)

```
# scel.R has two callable functions
# emplik    does one EL calculation,
# tracelr   calls emplik on a trajectory from mu0 to mu1 in N+1 steps

# However, one can only copy the part without tracelr in this problem
# We omit the copying part here

# Start from the drawing part
library(plotrix) # to draw ellipses
m <- colMeans(law)
s <- cov(law)
values <- eigen(s)$values
vectors <- eigen(s)$vectors
n <- nrow(law)
p <- 2

# EL confidence regions : all in red
ilabel <- seq(m[1] - 30, m[1] + 50, 1)
jlabel <- seq(m[2] - .3, m[2] + .3, 0.01)
test <- matrix(rep(1: 81 * 61), 81, 61)
for (i in 1 : 81) {
  for (j in 1 : 61) {
    test[i,j] <- emplik(law, c(ilabel[i], jlabel[j]))$logelr
  }
}
plot(c(560, 640), c(2.8, 3.4))
contour(ilabel, jlabel, exp(test), levels = c(0.5, 0.1, 0.05, 0.01), col = 'red', add = TRUE

# Normal confidence regions : all in blue
c <- (n - 1) * p / (n - p) * qf(0.5, p, n - p)
```

Date: 2019/10/14.
liweiyu@mail.ustc.edu.cn.

```
a <- sqrt(values[1] * c / n)
b <- sqrt(values[2] * c / n)
angle <- atan(vectors[2] / vectors[1])
draw.ellipse(m[1], m[2], a, b, angle, deg = FALSE, border = 'blue') # 50% CR
c <- (n - 1) * p / (n - p) * qf(0.9, p, n - p)
a <- sqrt(values[1] * c / n)
b <- sqrt(values[2] * c / n)
angle <- atan(vectors[2] / vectors[1])
draw.ellipse(m[1], m[2], a, b, angle, deg = FALSE, border = 'blue') # 90% CR
c <- (n - 1) * p / (n - p) * qf(0.95, p, n - p)
a <- sqrt(values[1] * c / n)
b <- sqrt(values[2] * c / n)
angle <- atan(vectors[2] / vectors[1])
draw.ellipse(m[1], m[2], a, b, angle, deg = FALSE, border = 'blue') # 95% CR
c <- (n - 1) * p / (n - p) * qf(0.99, p, n - p)
a <- sqrt(values[1] * c / n)
b <- sqrt(values[2] * c / n)
angle <- atan(vectors[2] / vectors[1])
draw.ellipse(m[1], m[2], a, b, angle, deg = FALSE, border = 'blue') # 99% CR
```

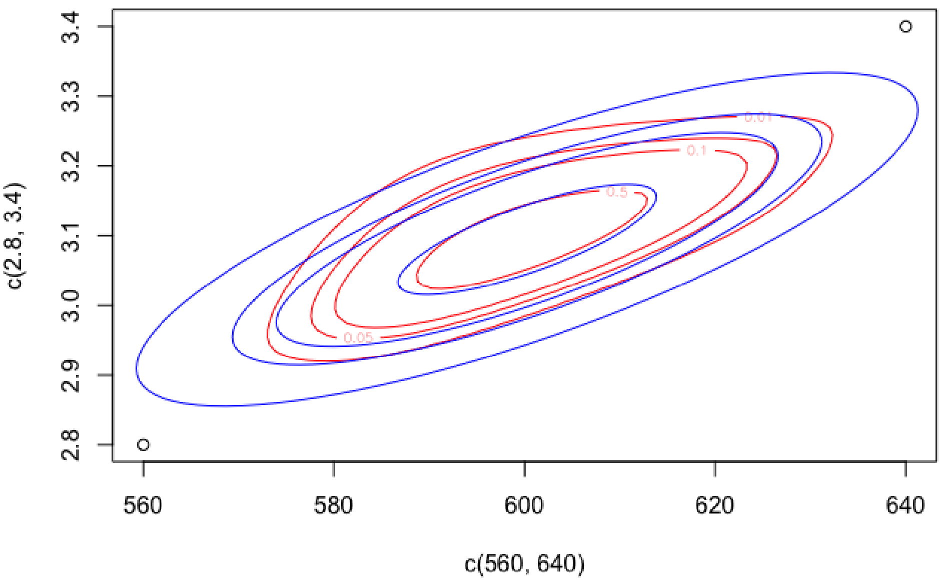

FIGURE 1. Red: EL confidence regions. Blue: normal confidence regions. Larger percentages lead to wider regions, respectively.

Remark 1. *According to Owen, one can also use F calibration to make the confidence regions narrower.*

3. Suppose i.i.d. samples $(X_i, Y_i), i = 1, \ldots, n$ from the population (X, Y). We are interested in the parameter $\tau = \sigma_X^2/\sigma_Y^2$, where σ_X^2 and σ_Y^2 are variances of X and Y, respectively. Denote $\theta = (\tau, \eta')'$, where η is a nuisance parameter. Solve estimate equations of an estimator for θ, and then give the empirical likelihood confidence interval of τ.

Solve. Let $\eta = (\sigma_Y^2, \mu_X, \mu_Y)$, then the parameters are determined by

$$\begin{cases} E(X_1 - \mu_X) = 0 \\ E(Y_1 - \mu_Y) = 0 \\ E[(Y_1 - \mu_Y)^2 - \sigma_Y^2] = 0 \\ E[(X_1 - \mu_X)^2 - \sigma_Y^2 \tau] = 0 \end{cases},$$

which gives the estimate equations

$$\begin{cases} \hat{\mu_X} = \frac{1}{n} \sum_i X_i \\ \hat{\mu_Y} = \frac{1}{n} \sum_i Y_i \\ \hat{\sigma_Y^2} = \frac{1}{n} \sum_i (Y_i - \hat{\mu_Y})^2 \\ \hat{\tau} = \frac{\frac{1}{n} \sum_i (X_i - \hat{\mu_X})^2}{\hat{\sigma_Y^2}} = \frac{\sum_i (X_i - \hat{\mu_X})^2}{\sum_i (Y_i - \hat{\mu_Y})^2} \end{cases}.$$

In the slides, we only know how to compute EL confidence set of θ, but for τ, we estimate the other elements of θ first, and then give the $(1-\alpha)$- EL confidence interval as

$$\{\tau : \log L(\tau) + n \log n > -0.5 \xi_{1,1-\alpha}^2\},$$

where $\log L(\tau) = \max \left\{ \sum_i \log p_i : p_i \geq 0, \sum_i p_i = 1, \tau = \frac{\sum_i p_i (X_i - \hat{\mu_X})^2}{\sum_i p_i (Y_i - \hat{\mu_Y})^2} \right\}$, $\hat{\mu_X} = \sum_i p_i X_i$, $\hat{\mu_Y} = \sum_i p_i Y_i$. □