

STA135 homework6

Zhen Zhang

February 6, 2016

5.10

```
growth <- read.table("/Users/Elliot/Github/UCDavis/STA135/resource/STA135_data/T1-4.dat")
growth_length <- growth[, 5:8]
names(growth_length) <- paste0("Length", 2:5)
```

(a)

Simultaneous confidence intervals:

```
calculate_simultaneous_interval <- function(sample_data, test_mean,
  alpha, p, flag = T, print_flag = T) {
  names_sample_data <- names(sample_data)
  sample_data <- as.matrix(sample_data)
  n <- dim(as.matrix(sample_data))[1]
  sample_mean <- apply(sample_data, 2, mean)
  deviation <- sample_mean - test_mean
  covariance_matrix <- cov(sample_data)
  F_star <- sqrt((n - 1) * p / (n - p) * qf(1 - alpha, p, n - p))
  sd_deviation <- sqrt(diag(covariance_matrix) / n)
  part2 <- F_star * sd_deviation
  result <- matrix(c(deviation - part2, upper = deviation + part2),
    n <- dim(as.matrix(sample_data))[2])
  rownames(result) <- names_sample_data
  colnames(result) <- c("lower", "upper")
  if (print_flag)
    print(result)
  if (flag)
    return(F_star)
  return(result)
}

invisible(calculate_simultaneous_interval(growth_length, 0, 0.05,
  4))
```

```
##           lower    upper
## Length2 130.6851 155.8863
## Length3 127.0216 191.5498
## Length4 160.3082 185.9776
## Length5 155.3749 198.9108
```

(b)

simultaneous intervals for changes in length

```

growth_length_change <- as.data.frame(sapply(1:3, function(i) growth_length[,
  i + 1] - growth_length[, i]))
names(growth_length_change) <- paste0("Length", 3:5, "-Length", 2:4)
invisible(calculate_simultaneous_interval(growth_length_change, 0,
  0.05, 4))

```

```

##               lower    upper
## Length3-Length2 -21.22649 53.22649
## Length4-Length3 -22.73077 50.44505
## Length5-Length4 -20.65385 28.65385

```

(c)

The simultaneous interval is:

```

growth_length_compare <- as.data.frame(cbind(growth_length[, 2] -
  growth_length[, 1], growth_length[, 4] - growth_length[, 3]))
names(growth_length_compare) <- paste0("Length", c(3, 5), "-Length",
  c(2, 4))
invisible(calculate_simultaneous_interval(growth_length_compare, 0,
  0.05, 4))

```

```

##               lower    upper
## Length3-Length2 -21.22649 53.22649
## Length5-Length4 -20.65385 28.65385

```

The ellipse:

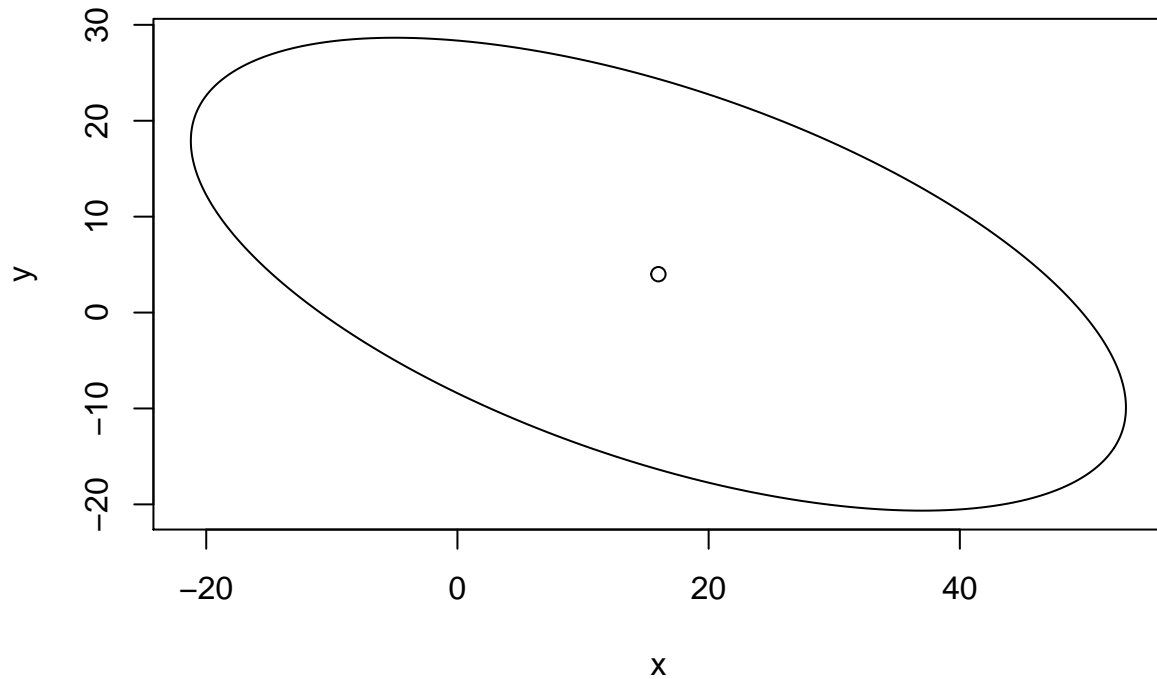
```

plot_ellipse <- function(sample_data, F_star, test_point) {
  n <- dim(sample_data)[1]
  sample_data_var <- cov(sample_data)
  sample_data_eigenvalue <- eigen(sample_data_var)$values
  main_vector <- eigen(sample_data_var)$vectors[, 1]
  lambda1 <- sqrt(sample_data_eigenvalue)[1]
  lambda2 <- sqrt(sample_data_eigenvalue)[2]
  alpha <- atan(main_vector[2]/main_vector[1])
  theta <- seq(0, 2 * pi, length = (1000))
  F_star <- F_star/sqrt(n)
  a <- lambda1 * F_star
  b <- lambda2 * F_star
  x_bar <- unname(apply(sample_data, 2, mean))
  x <- x_bar[1] + a * cos(theta) * cos(alpha) - b * sin(theta) *
    sin(alpha)
  y <- x_bar[2] + a * cos(theta) * sin(alpha) + b * sin(theta) *
    cos(alpha)
  plot(x, y, type = "l")
  points(x_bar[1], x_bar[2], type = "p")
  points(test_point[1], test_point[2], type = "p", col = "red")
}

plot_ellipse(growth_length_compare, calculate_simultaneous_interval(growth_length_compare,
  0, 0.05, 4), c(-110, -110))

```

```
##               lower    upper
## Length3-Length2 -21.22649 53.22649
## Length5-Length4 -20.65385 28.65385
```



I will revisit the plot in part (e).

(d)

```
calculate_Bonferroni_interval <- function(sample_data, test_mean,
  alpha, pair, print_flag = T) {
  names_sample_data <- names(sample_data)
  sample_data <- as.matrix(sample_data)
  n <- dim(as.matrix(sample_data))[1]
  sample_mean <- apply(sample_data, 2, mean)
  deviation <- sample_mean - test_mean
  covariance_matrix <- cov(sample_data)
  B <- qt(1 - alpha/(2 * pair), n - 1)
  sd_deviation <- sqrt(diag(covariance_matrix)/n)
  part2 <- B * sd_deviation
  result <- matrix(c(deviation - part2, upper = deviation + part2),
    n <- dim(as.matrix(sample_data))[2])
  rownames(result) <- names_sample_data
  colnames(result) <- c("lower", "upper")
  if (print_flag)
```

```

    print(result)
    return(result)
}

invisible(calculate_Bonferroni_interval(growth_length, 0, 0.05, 7))

```

```

##           lower    upper
## Length2 137.3884 149.1831
## Length3 144.1854 174.3860
## Length4 167.1359 179.1498
## Length5 166.9550 187.3307

```

```

invisible(calculate_Bonferroni_interval(growth_length_change, 0, 0.05,
7))

```

```

##           lower    upper
## Length3-Length2 -1.422784 33.42278
## Length4-Length3 -3.266772 30.98106
## Length5-Length4 -7.538521 15.53852

```

(e)

The bonferroni interval:

```

invisible(calculate_Bonferroni_interval(growth_length_compare, 0,
0.05, 7))

```

```

##           lower    upper
## Length3-Length2 -1.422784 33.42278
## Length5-Length4 -7.538521 15.53852

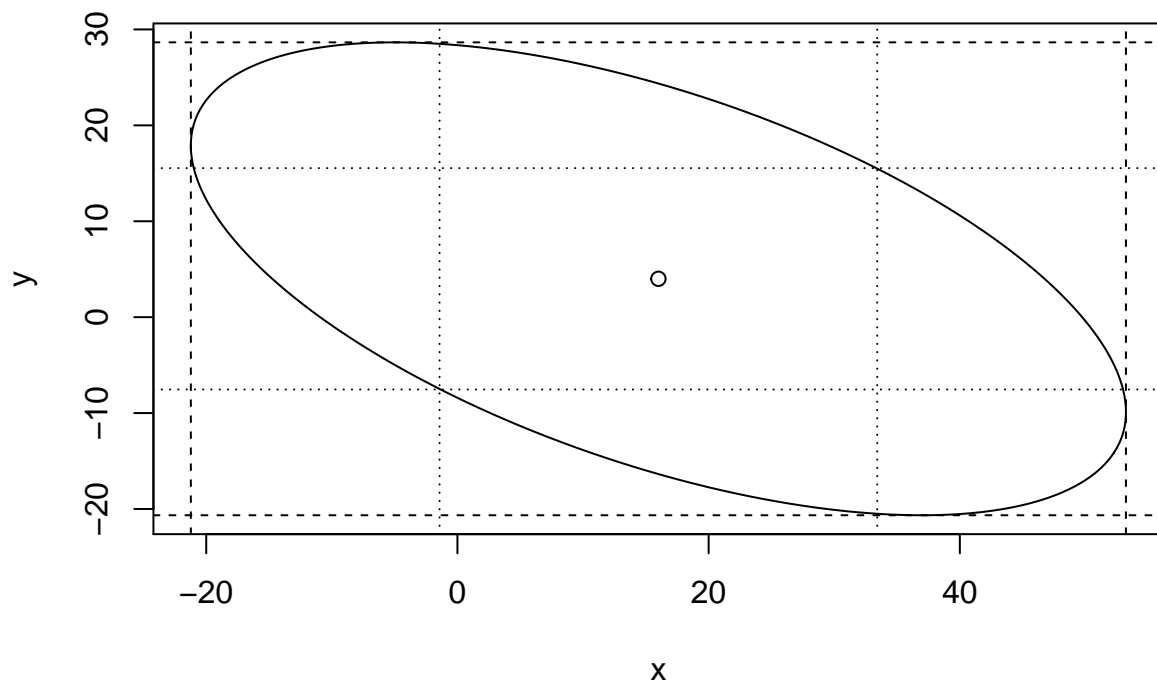
```

```

plot_ellipse(growth_length_compare, calculate_simultaneous_interval(growth_length_compare,
0, 0.05, 4, print_flag = F), c(-110, -110))
s_interval <- as.vector(t(calculate_simultaneous_interval(growth_length_compare,
0, 0.05, 4, FALSE, print_flag = F)))
b_interval <- as.vector(t(invisible(calculate_Bonferroni_interval(growth_length_compare,
0, 0.05, 7, print_flag = F))))
for (i in 1:2) {
  abline(v = s_interval[i], lty = 2)
  abline(v = b_interval[i], lty = 3)
}

for (i in 3:4) {
  abline(h = s_interval[i], lty = 2)
  abline(h = b_interval[i], lty = 3)
}

```



The bigger square with line type `lty = 2` is the simultaneous interval, while the inner square with line type `lty = 3` is the Bonferroni interval. Clearly the Bonferroni interval is much smaller.

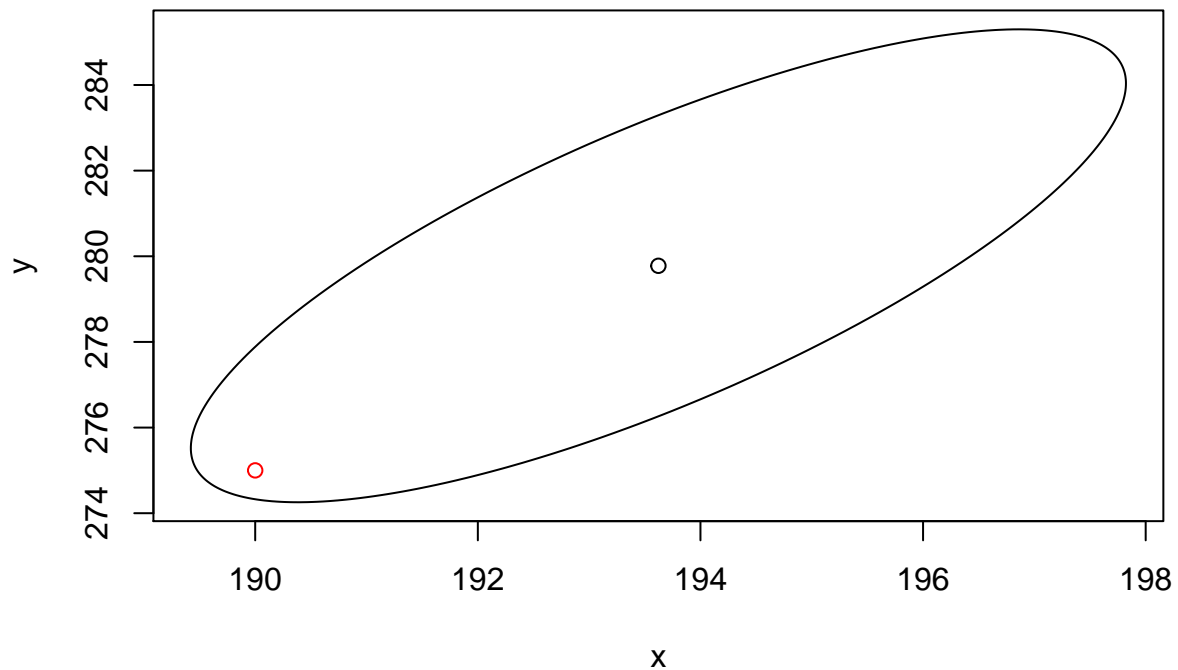
5.20

```
kites <- read.table("/Users/Elliott/Github/UCDavis/STA135/resource/STA135_data/T5-12.dat")
names(kites) <- c("tail length", "wing length")
```

(a)

```
plot_ellipse(kites, calculate_simultaneous_interval(kites, 0, 0.05,
2), c(190, 275))
```

```
##           lower  upper
## tail length 189.4217 197.8227
## wing length 274.2564 285.2992
```



Here is the ellipse. The point is in the ellipse, so it is plausible.

(b)

Simultaneous interval:

```
print(calculate_simultaneous_interval(kites, 0, 0.05, 2))
```

```
##           lower    upper
## tail length 189.4217 197.8227
## wing length 274.2564 285.2992
## [1] 2.564853
```

Bonferroni interval:

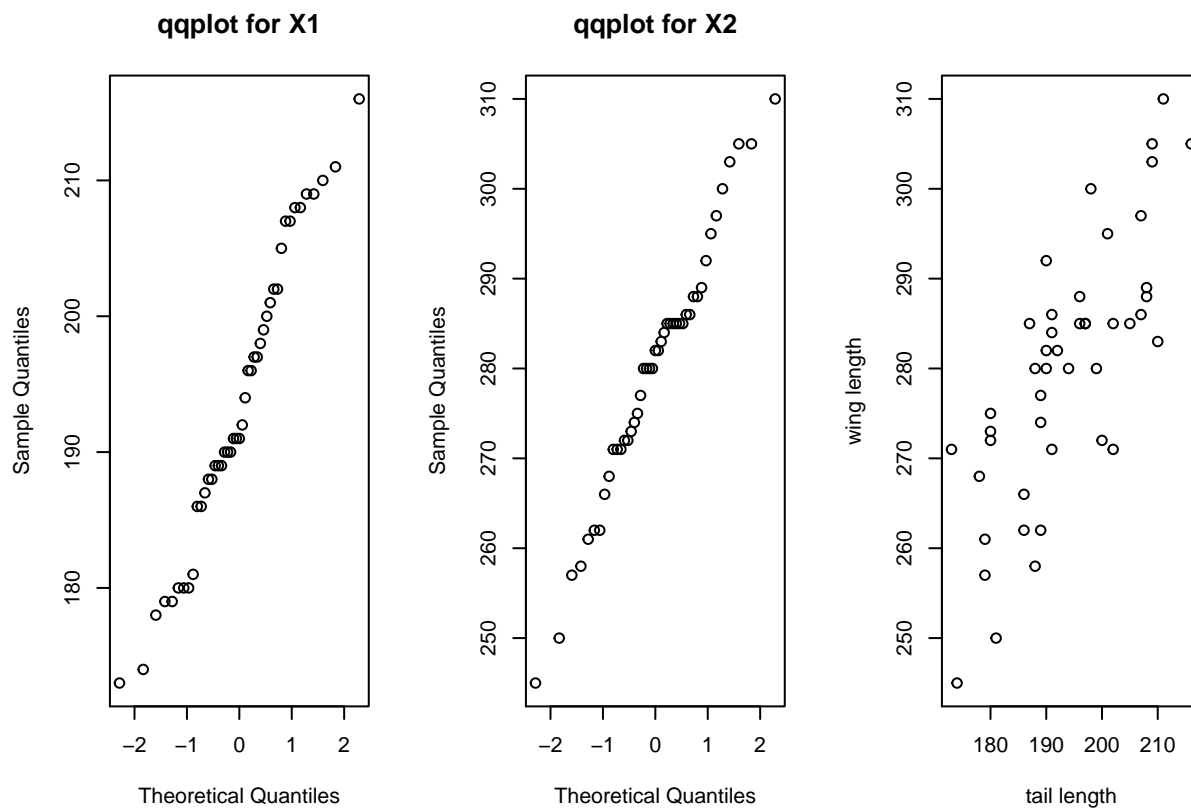
```
print(calculate_Bonferroni_interval(kites, 0, 0.05, 2))
```

```
##           lower    upper
## tail length 189.8216 197.4229
## wing length 274.7819 284.7736
##           lower    upper
## tail length 189.8216 197.4229
## wing length 274.7819 284.7736
```

The simultaneous interval is a little large than the bonferroni interval. The simultaneous confidence interval will touch the simultaneous confidence region from outside.

(c)

```
par(mfrow = c(1, 3))
qqnorm(kites$`tail length`, main = "qqplot for X1")
qqnorm(kites$`wing length`, main = "qqplot for X2")
with(kites, plot(`tail length`, `wing length`))
```



From the plot we know X_1 and X_2 are not normal. We could transform X_1 .