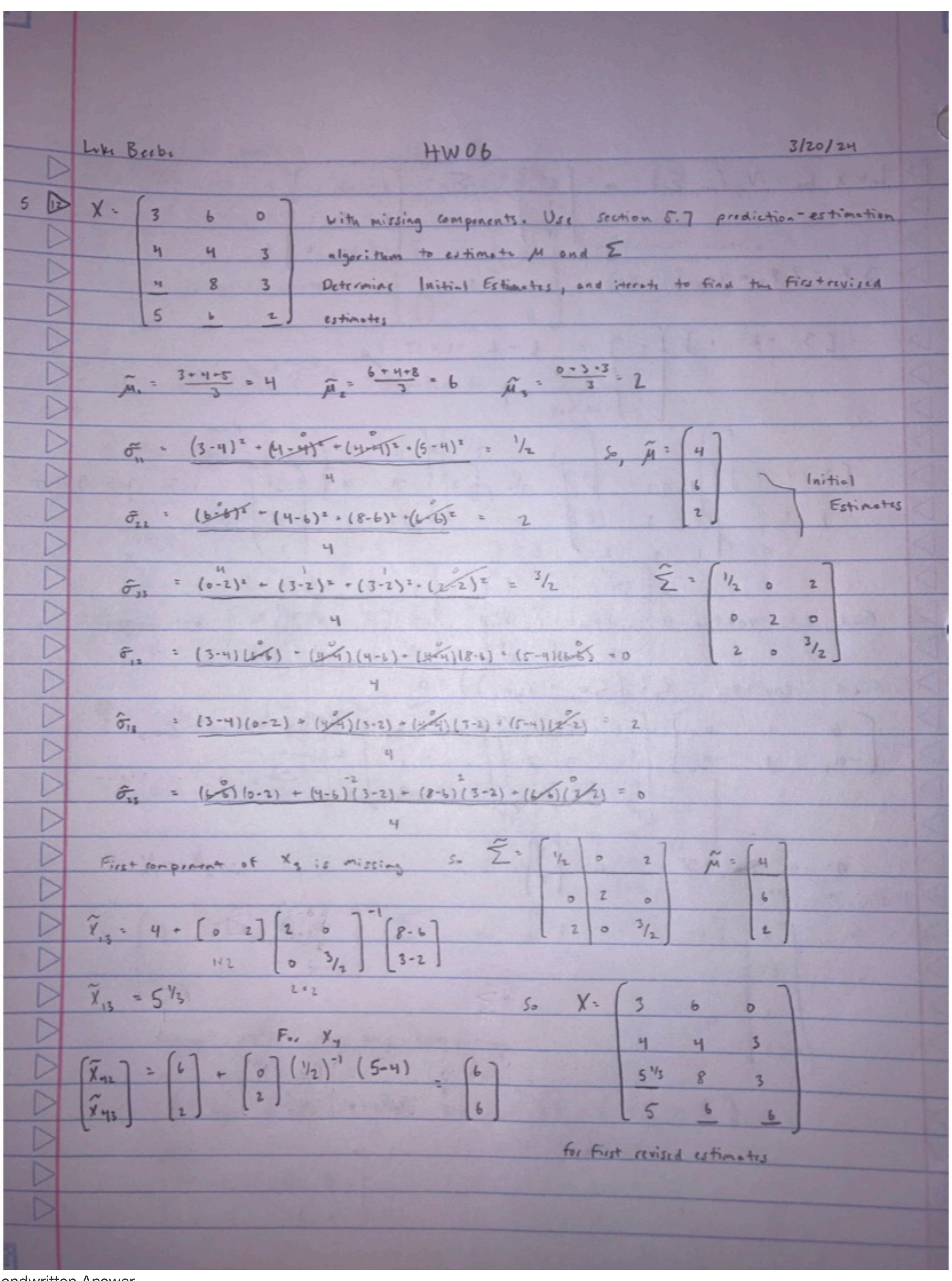
HW06

Luke Beebe

2024-03-20

5.12



Handwritten Answer

5.20

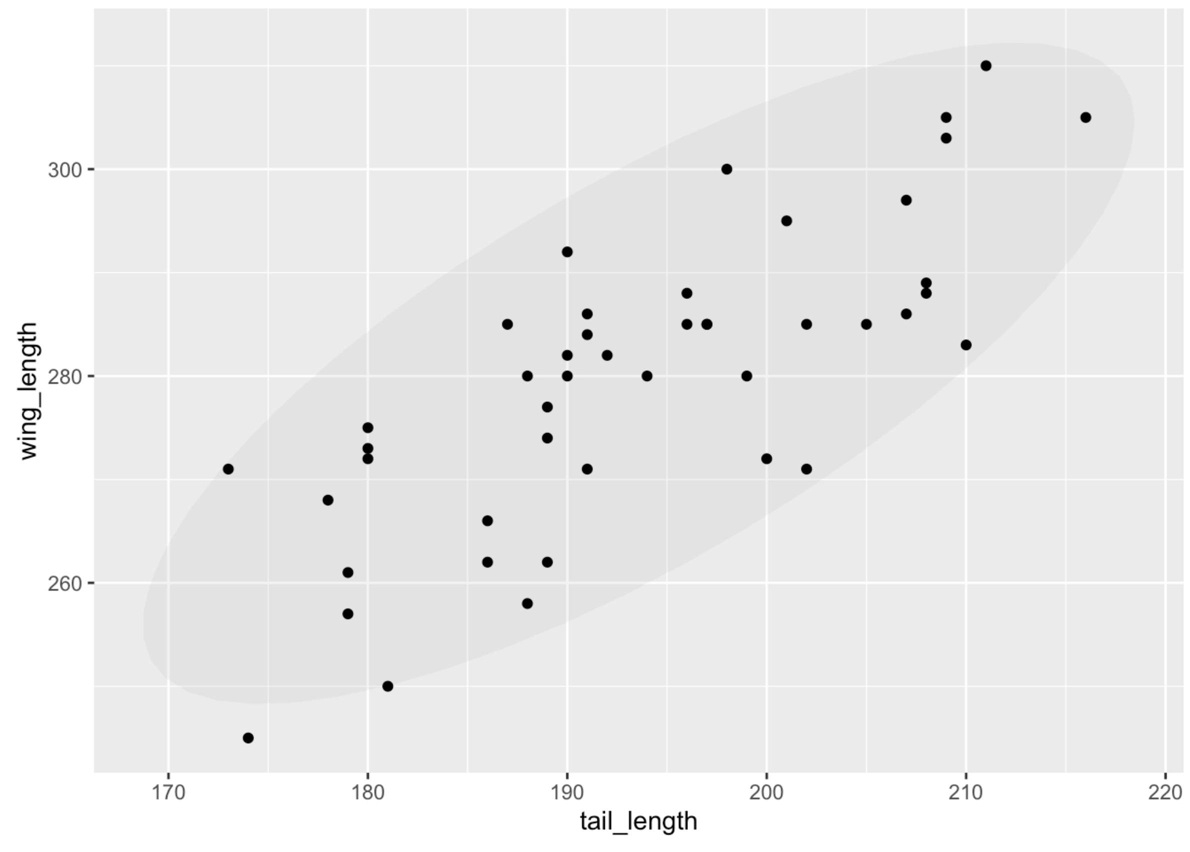
A wildlife ecologist measured tail length (mm) and wing length (mm) for a sample of n=45 female hook-billed kites.

```
tail_length <- c(191, 197, 208, 180, 180, 188, 210, 196, 191, 179, 208, 202, 200, 192, 199,
                 186, 197, 201, 190, 209, 187, 207, 178, 202, 205, 190, 189, 211, 216, 189,
                 173, 194, 198, 180, 190, 191, 196, 207, 209, 179, 186, 174, 181, 189, 188)

wing_length <- c(284, 285, 288, 273, 275, 280, 283, 288, 271, 257, 289, 285, 272, 282, 280,
                 266, 285, 295, 282, 305, 285, 297, 268, 271, 285, 280, 277, 310, 305, 274,
                 271, 280, 300, 272, 292, 286, 285, 286, 303, 261, 262, 245, 250, 262, 258)
female_kites <- data.frame(tail_length, wing_length)
```

a. Find and sketch the 95% confidence elipse for the population means u1 and u2. Suppose it is known that u1 = 190mm and u2 = 275mm for male hook-billed kites. Are these plausible values for the mean tail length and mean wing length for the female birds? Explain.

```
ggplot(female_kites, aes(tail_length, wing_length)) + geom_point() +
  geom_polygon(stat="ellipse", alpha=0.05)
```



They are plausible values for mean tail length and mean wing length for female birds as it falls within the 95% confident ellipse. This suggests that we cannot reject the null that the true mean of tail_length=190 and wing_length=275.

b. Construct the simultaneous 95% T^2 intervals and the 95% Bonferroni intervals for u1 and u2. Compare the two sets of intervals. What advantage, if any, do they T^2 intervals have over the Bonferroni intervals?

```
mu <- apply(female_kites,2,mean)
Sigma <- cov(female_kites)
n=45
p=2
T2 <- ((p*(n-1))/(n-p))*qf(p=0.95, p, n-p)
T2span <- sqrt(T2)*sqrt(Sigma[1,1]/n)

print("Hotelling's T2 CIs")
```

```
## [1] "Hotelling's T2 CIs"
```

```
paste("tail length:", paste(mu[1] - T2span, mu[1] + T2span))
```

```
## [1] "tail length: 189.421724184031 197.822720260413"
```

```
paste("wing length:", paste(mu[2] - T2span, mu[2] + T2span))
```

```
## [1] "wing length: 275.577279739587 283.978275815969"
```

```
print("Bonferroni CIs")
```

```
## [1] "Bonferroni CIs"
```

```
t <- qt(p=1-(0.05/(2*p)), n-1)
Tspan <- t*sqrt(Sigma[1,1]/n)
paste("tail length:" , paste(mu[1] - Tspan, mu[1] + Tspan))
```

```
## [1] "tail length: 189.821559660897 197.422884783548"
```

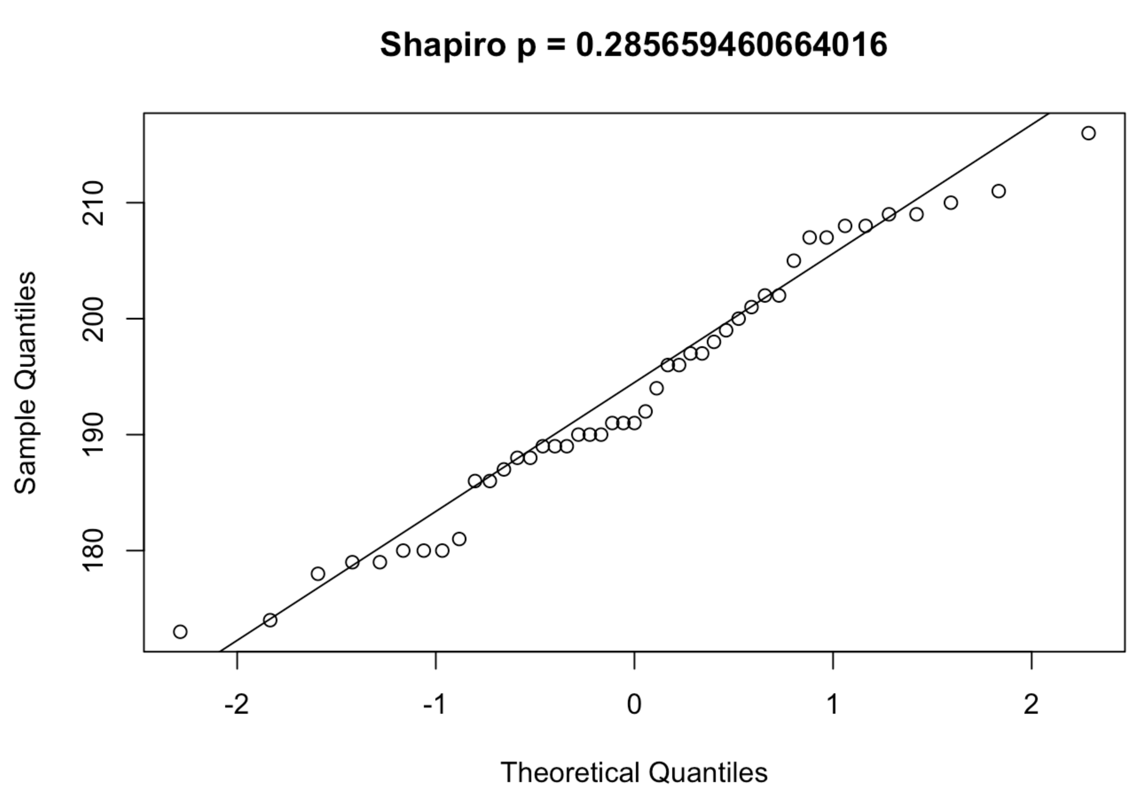
```
paste("wing length:",paste(mu[2] - Tspan, mu[2] + Tspan))
```

```
## [1] "wing length: 275.977115216452 283.578440339103"
```

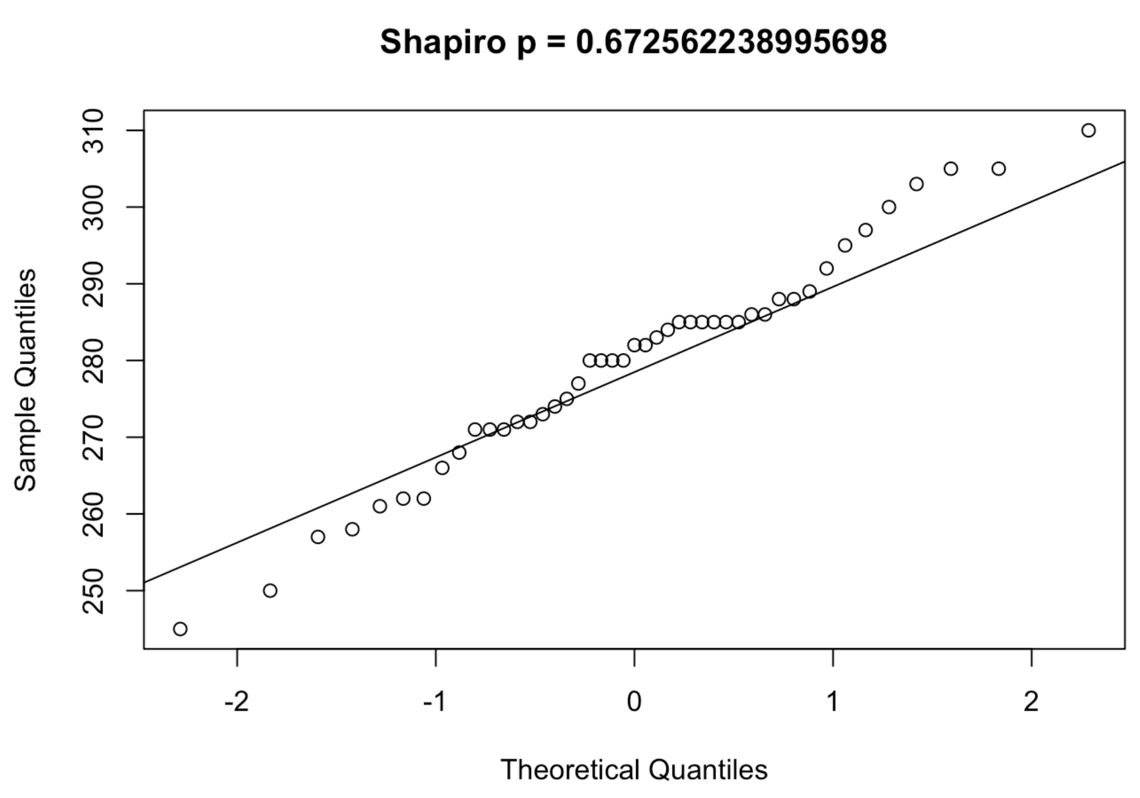
Bonferroni CIs are shorter than Hotelling's T2 CIs, providing more precise estimates. If we are interested only in the component means, the Bonferroni intervals will do. Otherwise, Hotelling's T2 takes the correlation between the measured variables into account.

c. Is the bivariate normal distribution a viable population model? Explain with reference to Q-Q plots and a scatter diagram.

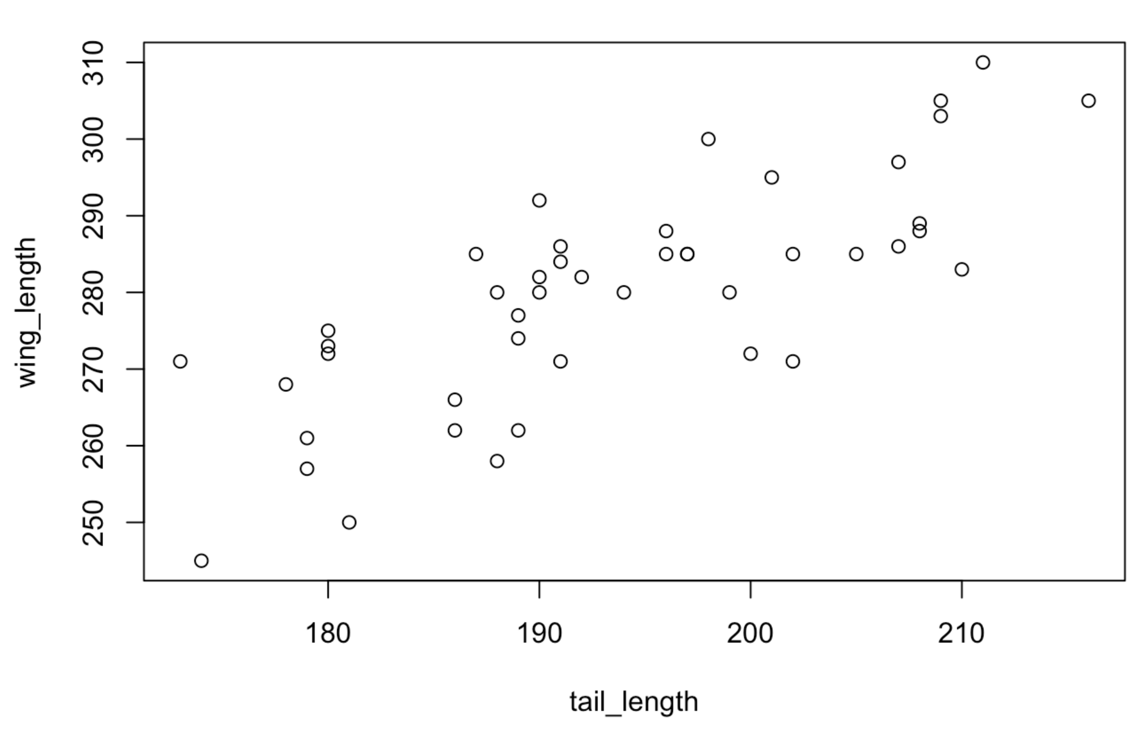
```
qqnorm(tail_length, main = paste("Shapiro p =",shapiro.test(tail_length)$p))
qqline(tail_length)
```




```
qqnorm(wing_length, main = paste("Shapiro p =",shapiro.test(wing_length)$p))
qqline(wing_length)
```




```
plot(tail_length, wing_length)
```



The bivariate normal distribution is a viable population model. The Q-Q plots look approximately normal, and a quick Shapiro-Wilkes test confirms its normality in the univariate case. I would have to use MVN to do further testing to see if it follows for the multivariate case; but given the plots, it is viable.