Homework 3

ST557

Nick Sun

October 21, 2019

Question 1

Part a.

We want to test if the social science + history, verbal, and science scores are different from the null hypothesis $H_0 : \mu = (500, 50, 30)$.

We can compute this in R using the formula $T(\mu_0) = n(\bar{X} - \mu_0)^T S^{-1}(\bar{X} - mu_0) \sim \frac{n-p}{(n-1)p} F_{p,n-p}$

```
mu0 <- c(500, 50, 30)
xbar <- apply(clep, 2, mean)
n <- 87
p <- 3

scaling <- function(n, p) {
  return((n-p)/((n-1)*p))
}

(T2 <- scaling(n,p)*87*t(xbar - mu0)%*%solve(cov(clep))%*%(xbar - mu0))
```

```
##          [,1]
## [1,] 72.70564
```

```
qf(.95, 3, 87-3)
```

```
## [1] 2.713227
```

Our Hotelling's T^2 is *much* larger than our critical F-statistic of 2.71.

Therefore we reject the null hypothesis that the mean vector of this distribution is $(500, 50, 30)$. In the context of the problem, we have evidence to say that the distribution of students scores in 2011 is different than in prior years.

Part b.

To get the length and direction of the axes of the 95% confidence ellipsoid, we will use the following formula:

$$\sqrt{\lambda_i \left(\frac{p(n-1)}{n(n-p)}\right) F_{p,n-p}(\alpha) e_i}$$

where λ_i and e_i are the eigenvalues and eigenvector respectively for the sample covariance matrix.

```
##               A           B           C
## [1,] 0.99390539  0.103731534 -0.037307396
## [2,] 0.10344339 -0.994589227 -0.009577815
## [3,] 0.03809906 -0.005660238  0.999257936
```

The eigenvectors of the sample covariance matrix show the directions of the confidence ellipse. We see that the first eigenvector is primarily compsed of the SocSciHistory score, the second eigenvector is primarily composed of the Verbal score, and the third eigenvector is made of the Science score.

```r
axis_length <- function(lambda, n, p, alpha = .95) {
  return(sqrt(lambda*(p*(n-1))/(n*(n-p))*qf(alpha, p, n-p)))
}

axis_length(lambdas[1], n, p)
```

```
## [1] 23.73
```

```r
axis_length(lambdas[2], n, p)
```

```
## [1] 2.472768
```

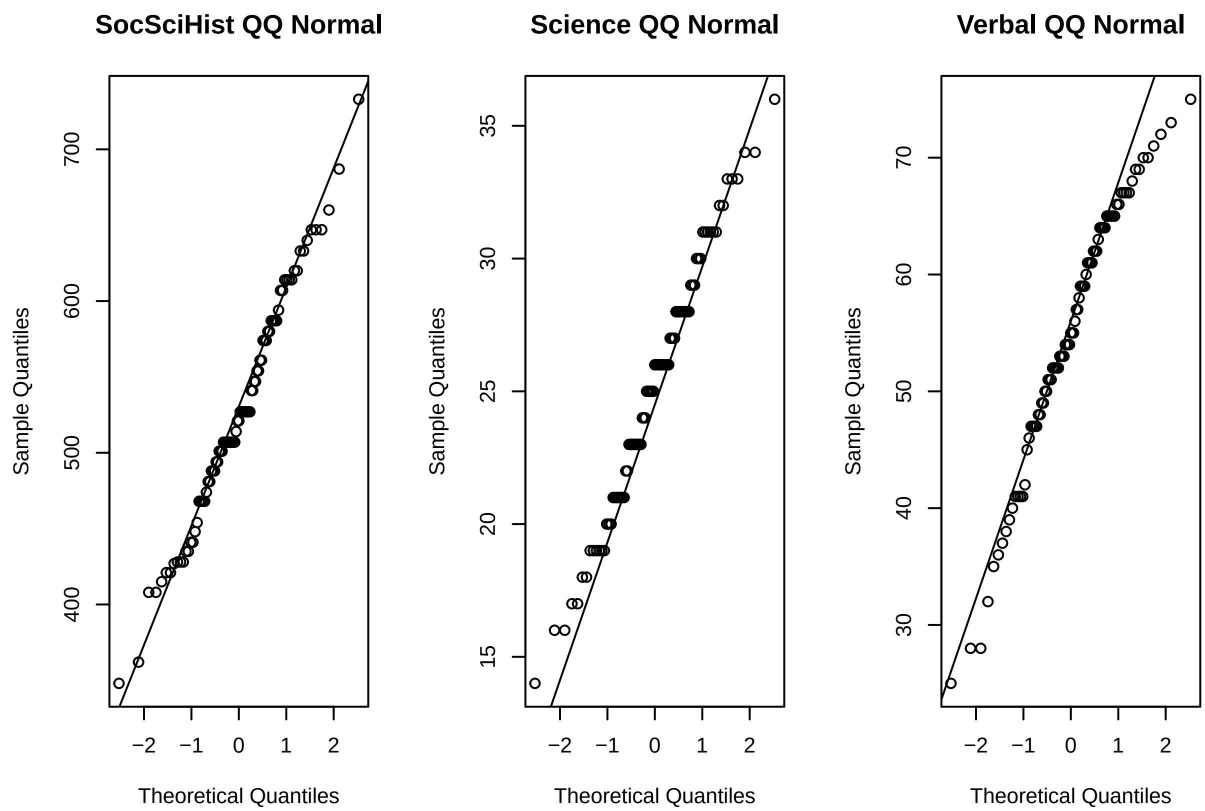
```r
axis_length(lambdas[3], n, p)
```

```
## [1] 1.1825
```

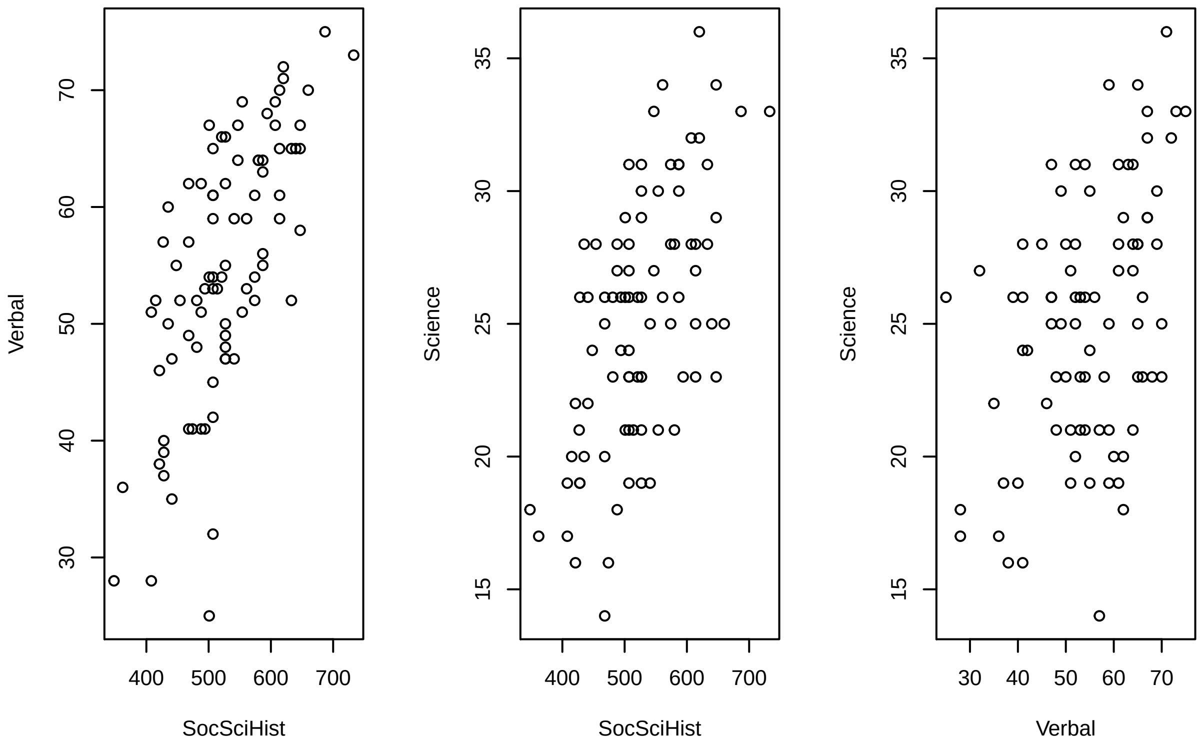
As we can see, the length of our first axis is quite long relative to the other two axes.

Part c.

Constructing QQ-plots and the marginal distributions for the scores of the three tests is easy enough in R.



The Verbal scores have a noticeably short right tail indicating that the marginal distribution may not be normal. Creating three pairwise scatterplots for our data produces the following:



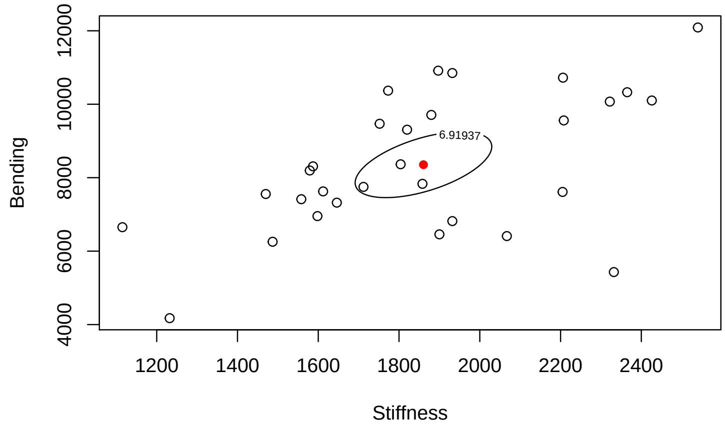
There appears to be a positive correlation between all pairs of the test scores. This is further confirmed by looking at the correlation matrix of the data - all of the correlation coefficients are positive and most are above .600.

```
##            SocSciHist    Verbal   Science
## SocSciHist  1.0000000 0.6986958 0.6060092
## Verbal      0.6986958 1.0000000 0.4333201
## Science     0.6060092 0.4333201 1.0000000
```

Question 2

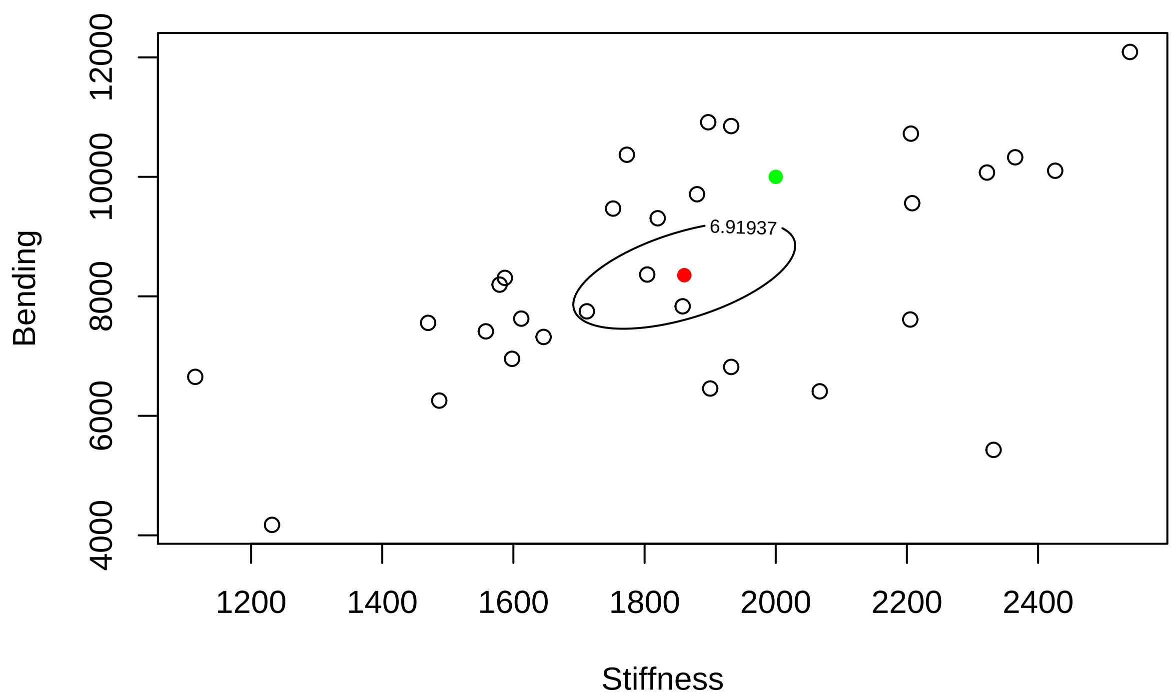
Part a.

We are asked to construct and plot a 95% confidence ellipse for the mean vector for this 2 variable lumber data.



Part b.

If $\mu_0 = [2000, 10000]$ is the typical stiffness and bend strength and we want to check if this value would be rejected at $\alpha = .05$, then we just need to check if that point is contained within our confidence interval.



The green point is outside, so we would reject at $\alpha = .05$ that the true population mean vector is $[2000, 10000]$.

Question 3

Part a.

We want to construct 95% Bonferroni joint confidence intervals for the means of the individual variables in the bone mineral dataset. There are 6 variables in the dataset, so we essentially want to make 6 $1 - \frac{.05}{6}$ 100% confidence intervals.

```
##        dRadius   nRadius dHumerus nHumerus     dUlna     nUlna
## [1,] 0.9093662 0.8797634 1.955683 1.886414 0.7662471 0.7530394
## [2,] 0.7782338 0.7568766 1.629677 1.583266 0.6425529 0.6346406
```

Part b.

Let's compute Hotelling's T^2 simulataneous confidence intervals.

The formula for these simultaneous confidence intervals is

$$a_k^T \bar{X} \pm \sqrt{\frac{p(n-1)}{n(n-p)} F_{p,n-p}(\alpha) a_k^T S a_k}$$

where a_k is a vector of coefficients (in our case, it is a vector of all 1s).

```
##        dRadius  nRadius dHumerus nHumerus     dUlna     nUlna
## [1,] 0.9455821 0.913702 2.045718 1.970137 0.8004086 0.7857386
## [2,] 0.7420179 0.722938 1.539642 1.499543 0.6083914 0.6019414
```

We can see that the Hotelling's T^2 simulanteous confidence intervals are slightly wider than the Bonferroni simultaneous confidence intervals.

Question 4

We are being asked here about bags of flour and their weight on different scales.

Part a.

To test whether the population mean vector is $\mu = (10, 10, 10)$, let's use Hotelling's T^2, $\alpha = .05$.

```r
mu0 <- c(10, 10, 10)
xbar <- apply(flour, 2, mean)
n <- nrow(flour)
p <- ncol(flour)

(T2 <- scaling(n,p)*n*t(xbar - mu0)%*%solve(cov(flour))%*%(xbar - mu0))
```

```
##          [,1]
## [1,] 4.600899
```

```r
qf(.95, p, n-p)
```

```
## [1] 3.196777
```

Our T^2 statistic is larger than our critical F statistc of 3.196 so we reject the null hypothesis. We have enough evidence to say that the population mean vector is different thn $(10, 10, 10)$.

Part b.

Now we can try the Bonferroni multiple testing method. This approach is the same as conducting multiple t-tests on the individual columns with a modified level of significance $\alpha^* = \frac{\alpha}{p}$

```
alphastar <- .05/p
t.test(flour$Scale1, mu = 10, alternative = "two.sided", conf.level = 1-alphastar)$p.value > alphastar
```

```
## [1] TRUE
```

```
t.test(flour$Scale2, mu = 10, alternative = "two.sided", conf.level = 1-alphastar)$p.value > alphastar
```

```
## [1] TRUE
```

```
t.test(flour$Scale3, mu = 10, alternative = "two.sided", conf.level = 1-alphastar)$p.value > alphastar
```

```
## [1] TRUE
```

We fail to reject the null hypothesis for any of the indiviudal t-tests. It should be noted that we did get a low p-value for the measurements on Scale 1, but due to the p-value correction we were not able to reach our level of significance.

Part c.

The T^2 simultaneous confidence intervals are going to be wider than the Bonferroni confidence intervals, so if a point is contained in a the Bonferroni simulatenous intervals it is guaranteed to be in the T^2 confidence intervals. We showed in the previous part that $(10, 10, 10)$ is contained in Bonferroni since we failed to reject for any of our t-tests, so the point is contained in the T^2 intervals as well.

Part d.

We want to perform a univariate two-sided test of $H_0 : \frac{1}{3}\mu_1 + \frac{1}{3}\mu_2 + \frac{1}{3}\mu_3 = 10$. We will need to calculate a repeated measures tests for this.

```
S <- cov(flour)
C <- rep(1/3, 3)
xbar <- xbar - 10

n*xbar%*%C%*%solve(t(C)%*%S%*%C)%*%C%*%xbar
```

```
##          [,1]
## [1,] 3.280655
```

```
((p-1)*(n-1)/(n-p+1))*qf(.95, p-1, n-p+1)
```

```
## [1] 7.504065
```

Our repeated measures modified F-test fails to reject at $\alpha = .05$, telling us that the average of the three scales is significantly different from 10.

Question 5

Part a.

Since the sample is known to be drawn from a multivariate normal distribution, we know that a property of this distribution is that a covariance of 0 is actually equivalent to independence.

Therefore, if $\Sigma = I_p$ then the elements of $\bar{X}$ are independent since the covariances are all zero.

Part b.

If we have p level α^* hypothesis tests, then given the null hypotheses are all true we can except that the probability of rejecting a single hypothesis test is α^*. This is assuming the test is well calibrated so that the p-values have a uniform distribution when the null hypothesis is true.

Since these tests are independent, the probability that we do not reject on any of the p tests is $(1-\alpha^*)^p$. However, we want the probability we reject *at least one test*. That will be $1-(1-\alpha^*)^p$.

Part c.

If we set the familywise error rate $1-(1-\alpha^*)^p$ equal to α, we can derive what we should set α^* to in order to control the overall signficance level at α

$$\begin{aligned}
1-(1-\alpha^*)^p &= \alpha \\
1-\alpha^* &= (1-\alpha)^{1/p} \\
\alpha^* &= 1-(1-\alpha)^{1/p}
\end{aligned}$$

Part d.

i.

Since both the regular Bonferroni confidence intervals and these modified confidence intervals will use the same standard deviation in practice, all we have to do is compare the ratio of the critical t-statistics being used.

```r
n <- 10
p <- 4
alpha_bonferroni <- .05/p
alpha_star <- 1-(1-.05)^(1/p)

qt(1-alpha_star/2, df = n-1)/qt(1-alpha_bonferroni/2, df = n-1)
```

```
## [1] 0.9961878
```

The modified confidence interval is almost the same size as the original Bonferroni interval.

ii.

```r
n <- 10
p <- 8
alpha_bonferroni <- .05/p
alpha_star <- 1-(1-.05)^(1/p)

qt(1-alpha_star/2, df = n-1)/qt(1-alpha_bonferroni/2, df = n-1)
```

```
## [1] 0.9959848
```

The modified confidence interval is still around the same size as the original Bonferroni interval.

iii.

Since both the regular Bonferroni confidence intervals and these modified confidence intervals will use the same standard deviation in practice, all we have to do is compare the ratio of the critical t-statistics being used.

```r
n <- 20
p <- 4
alpha_bonferroni <- .05/p
alpha_star <- 1-(1-.05)^(1/p)

qt(1-alpha_star/2, df = n-1)/qt(1-alpha_bonferroni/2, df = n-1)
```

```
## [1] 0.9968024
```

The modified confidence interval is *still* almost the same size as the original Bonferroni interval.

iv.

Since both the regular Bonferroni confidence intervals and these modified confidence intervals will use the same standard deviation in practice, all we have to do is compare the ratio of the critical t-statistics being used.

```r
n <- 20
p <- 8
alpha_bonferroni <- .05/p
alpha_star <- 1-(1-.05)^(1/p)

qt(1-alpha_star/2, df = n-1)/qt(1-alpha_bonferroni/2, df = n-1)
```

```
## [1] 0.9967338
```

From these four investigations, we see that our modified α^* doesn't change the size of the simultaneous confidence intervals too much, but they are slightly smaller in all cases. With further investigation, we see that as p increases, the size of the modified confidence interval approaches the size of the Bonferroni interval.