Quiz 1

Daniel Detore

I Pledge my honor that I have abided by the Stevens Honor System.

```
srs <- c(170, 129, 172, 174, 166, 169, 188, 181, 175, 164, 175, 177,
182, 173, 179, 167, 202, 163, 160, 162)
```

Problem 1

```
boxplot(srs, horizontal=TRUE, main="x' samples boxplot")
# mean:
abline(v = mean(srs), col="red")
mean(srs)
```

```
## [1] 171.4
```

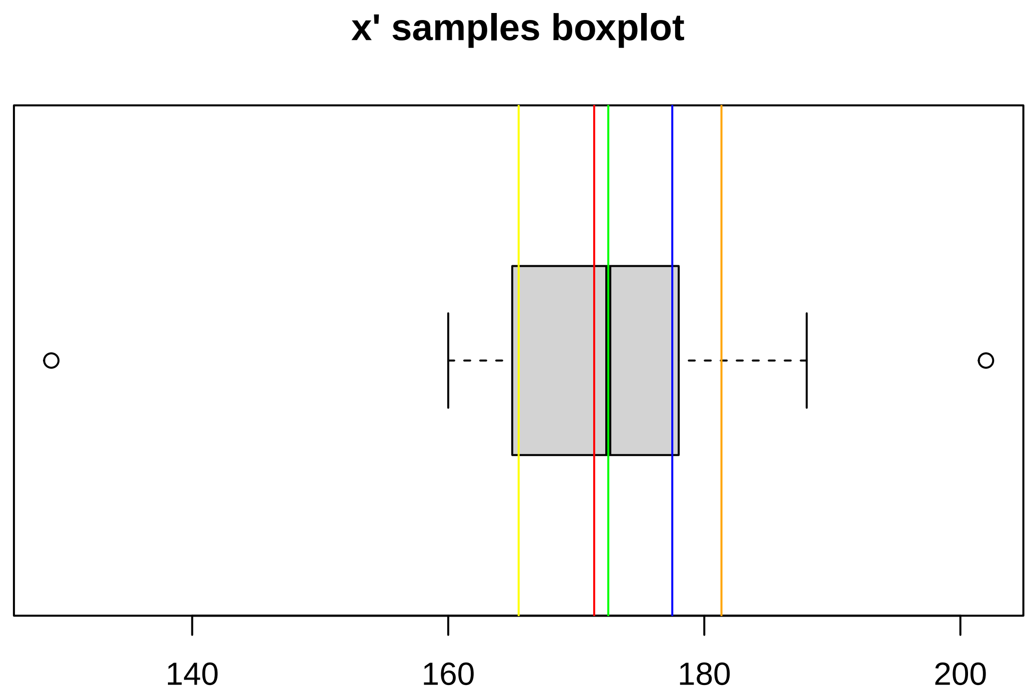
```
# 86th quantile:
quantile(srs, probs=.86)
```

```
##    86%
## 181.34
```

```
abline(v = quantile(srs, probs=.86), col="orange")
# all quartiles:
summary(srs)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   129.0   165.5   172.5   171.4   177.5   202.0
```

```
abline(v = quantile(srs, probs=.25), col="yellow")
abline(v = quantile(srs, probs=.5), col="green")
abline(v = quantile(srs, probs=.75), col="blue")
```



The outliers, based on the box plot, are the min and max values 129 and 202.

```r
max(srs)
```

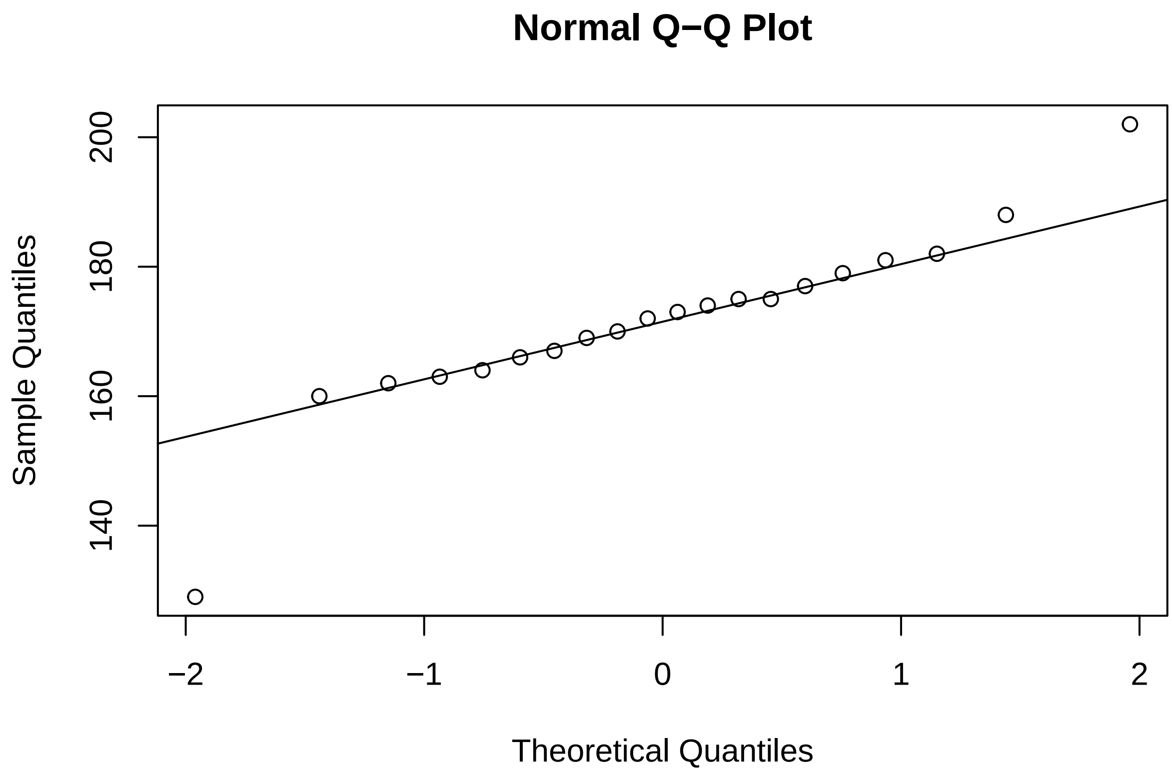
```
## [1] 202
```

```r
min(srs)
```

```
## [1] 129
```

Problem 2

```r
qqnorm(srs)
qqline(srs)
```




```r
IQR(srs)
```

```
## [1] 12
```

```r
var(srs)
```

```
## [1] 196.7789
```

```r
sd(srs)
```

```
## [1] 14.02779
```

```r
summary(srs)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   129.0   165.5   172.5   171.4   177.5   202.0
```

Based on this summary, the data is skewed left because the mean is smaller than the median.

Problem 3

i

$T_1 \sim \chi_n^2$ where $n = 1$ because T_1 follows the form $\Sigma_{i=0}^{n} X_i^2$ where $X \sim \mathcal{N}(0, 1)$ and all observations are independent.

This means that $P(T_1 \leq t_1) = pchisq(t_1, 1)$.

ii

According to the fundamental theorem (2): $\Sigma_{i=1}^{n} \frac{(X_i - \bar{X})^2}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$

This means that $P(T_2 \leq t_2) = pchisq(t_2, 19)$.

iii

The probability $P(T_2 \leq t_2)$ is larger because, as the degrees of freedom increase, more of the χ^2 distribution's area moves to its tails. Since this probability is cumulative, and therefore collects most of its value from the distribution's left tail, it will usually be greater than the probability $P(T_1 \leq t_1)$.

iv

The form of T_3 implies that it follows $\mathcal{N}(0, 1)$. This is the standardization of the original distribution.

$P(|T_3| > |t_3|) = P(T_3 < -|t_3|) + 1 - P(T_3 < |t_3|) = pnorm(-|t_3|, 0, 1) + 1 - pnorm(|t_3|, 0, 1)$.

$P(|T_3| > |t_3|) = P(|Z| > |t_3|)$, because $T_3 \equiv Z \sim \mathcal{N}(0, 1)$, therefore it cannot be that $P(|T_3| > |t_3|) > P(|Z| > |t_3|)$.

v

According to a corrollary of the fundamental theorem, T_4 is of the form $\frac{N(0,1)}{\sqrt{\chi^2_n/n}} \implies T_4 \sim \mathcal{T}_{n-1}$.

$P(|T_4| > |t_4|) = P(T_4 < -|t_4|) + 1 - P(T_4 < |t_4|) = pt(-|t_3|, 0, 1) + 1 - pt(|t_3|, 0, 1)$.

$P(|T_4| > |t_4|) > P(|Z| > |t_3|)$, because the $\mathcal{T}$ distribution has much more weight in its tails, even with $df = 1$, than the standard normal distribution.

Problem 4

In order of left-to-right then top-to-bottom:

$3X_2 + 4 \sim \mathcal{N}(4, 9)$

$-4X_1 + 3X_2 + 2X_3 - 5 \sim \mathcal{N}(-4(-1) + 2(1) - 5, (-4)^2(9) + 3^2(1) + 2^2(16)) \equiv \mathcal{N}(1, 217)$

$\equiv 3 * \chi^2_1 = \chi^2_3$

$\equiv \frac{\frac{1}{2}\chi^2_2}{\chi^2_1/1} \sim \mathcal{F}_{2,1}$

$\sim \mathcal{T}_2$

$\sim \mathcal{F}_{1,2}$

Problem 5

i

$\hat{\mu} = \bar{X} = \frac{\Sigma_{i=1}^{n} X_i}{n}$

$\hat{\sigma^2} = S^2 = \frac{\Sigma_{i=1}^{n}(X_i - \bar{X})^2}{n}$

ii

$\mathcal{T}_{n-1}$

χ^2_{n-1}

iii

It is expected that $\hat{\sigma^2} = \frac{\Sigma_{i=1}^{n}(X_i - \bar{X})^2}{n}$ will approximate the true σ^2 as the sample size increases.

iii

They are not independent because you need $\bar{X} = \hat{\sigma}$ to calculate $S^2 = \bar{\sigma^2}$. They are correlated.

Problem 6

It would most likely be near 0, since x is random. It is incredibly unlikely for all xs to fall into a perfect linearly increasing correlation.