Homework 04 - 4.4, 4.14, 4.24, 4.26, 4.34, 4.40, 4.48

Max Wagner

September 9, 2015

4.4

   a. The point estimate is the mean in this case, which is 171.1, and the median is 170.3.

   b. The point estimate of the SD is 9.4, the IQR is 177.8 - 163.8 = 14.

   c. Anything over mean + 2SD is considered abnormal. So anything over 189.1 in this case. 180cm is below this level, so the person is considered a normal height.

   d. The sample mean and sd would be somewhat similar, but it's unlikely they would be identical. When taking a sample, depending on the CI, the mean can vary, but will usually be in a relative range.

   e. SE will be used in this case. It is 0.4174687.

4.14

   a. False, we are 95% confident that the population mean is within the CI.

   b. False, one of the conditions state that the population should not be strongly skewed.

   c. False, we are looking at the population mean, not the sample means.

   d. True, this is what the CI means.

   e. True, the "net" is narrower, and is less likely to contain the population mean in this case.

   f. False, it would need to be 9 times larger, as the n is under a square root.

   g. True, the margin of error is half the CI.

4.24

   a. The individuals are independent, it is assumed to be a SRS, and it likely less than 10% of the population.

   b. null: mean = 32, alternative: mean != 32. I will reject the null hypothesis if the p-value is less than 0.10.

```
z <- (30.69 - 32) / (4.31 / sqrt(36)) # z score
p <- pnorm(z, 0 ,1); p # p value
```

```
## [1] 0.0341013
```

The p value is less than 0.10, which means I reject the null hypothesis.

   c. The p value is essentially the probabilty of incorrectly rejecting the null hypothesis. In this case, there is evidence that the p-value is less than the significance level, which means the null can be rejected.

   d.

```r
mean <- 30.69
SE <- 4.31 / sqrt(36)
z <- 1.645
lower <- mean - z * SE
upper <- mean + z * SE
```

The interval is 29.5083417, 31.8716583.

    e. The results disagree, as the p value test and CI test both conclude that 32 is unlikely to happen with a 90% CI.

4.26

    a. The two means are very different, so the z score / p value might be weird.

null: mean = 100, alternative: mean > 100. I will reject the null hypothesis if the p-value is less than 0.10.

```r
z <- (118.2 - 100) / (6.5 / sqrt(36)) # z score
p <- pnorm(z, 0 ,1); p # p value
```

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

The p value is higher than .10, which means we fail to reject the null hypothesis.

    b.

```r
mean <- 118.2
SE <- 6.5 / sqrt(36)
z <- 1.645
lower <- mean - z * SE
upper <- mean + z * SE
```

The interval is 116.4179167, 119.9820833.

    c. The value is not in the CI, but the p value test does not reject the null.

4.34 The sampling distribution is the distribution of point estimates taken from all possible samples of a fixed size from a population. With a larger sample size, the sampling distribution becomes more normal, the center will closer to the population mean, and the spread will be smaller.

4.40

    a.

```r
z <- (10500 - 9000) / (1000 / sqrt(1))
pnorm(z)
```

```
## [1] 0.9331928
```

    b.

```
mean <- 9000
sd <- (1000 / sqrt(15))
```
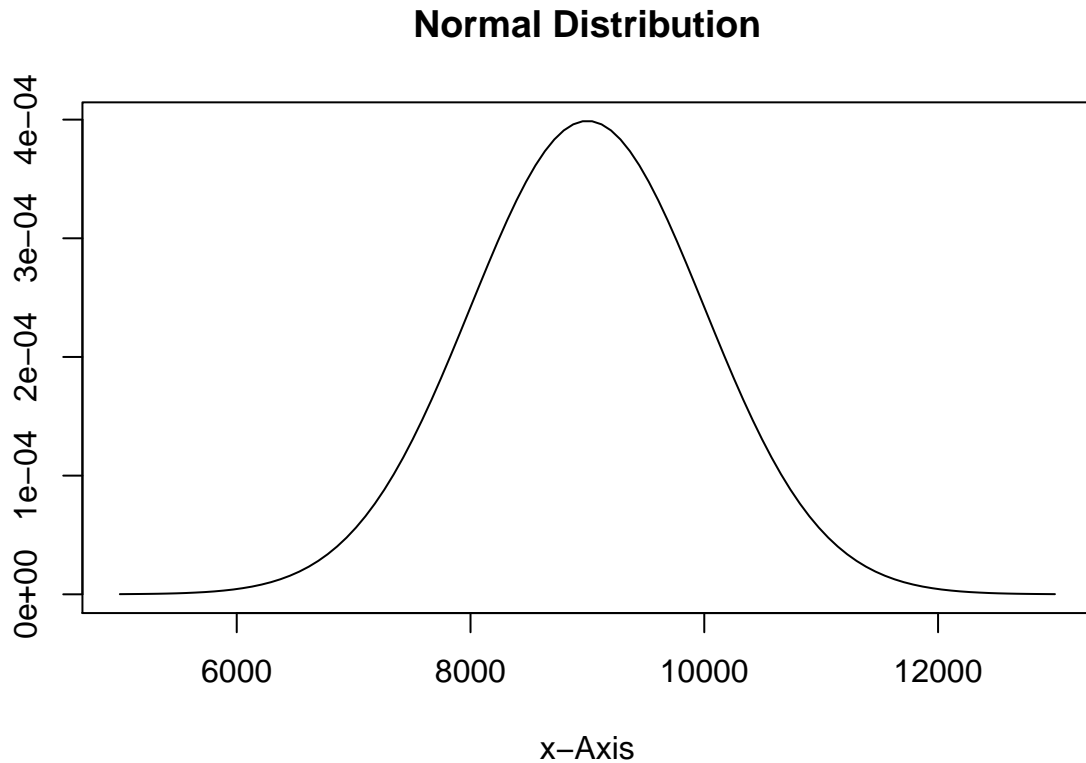
The distribution is N(9000, 258.1988897).

   c.

```
z <- (10500 - 9000) / (1000 / sqrt(15))
1 - pnorm(z)
```

```
## [1] 3.133452e-09
```

The probabilty of the mean being higher than 10500, is nearly 0.

   d. Using a snip of the code from the normalPlot code, and changing it a bit.

```
mean = 9000
sd = 1000
x <- seq(-4, 4, length = 100) * sd + mean
hx <- dnorm(x, mean, sd)
plot(x, hx, type = "l", xlab = "x-Axis", ylab = "", main = "Normal Distribution", xlim = c(5000, 13000))
```

```
mean = 9000
sd = (1000 / sqrt(15))
x <- seq(-4, 4, length = 100) * sd + mean
hx <- dnorm(x, mean, sd)
plot(x, hx, type = "l", xlab = "x-Axis", ylab = "", main = "Normal Distribution", xlim = c(5000, 13000))
```

Normal Distribution

x–Axis
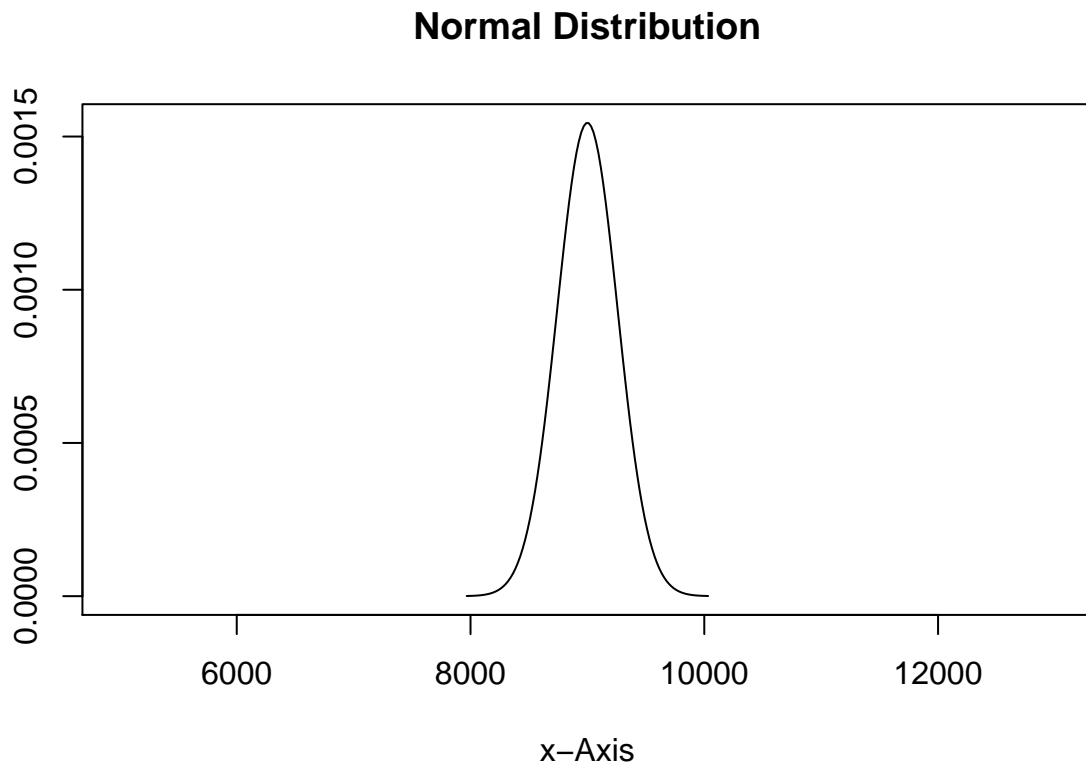
    e. You could estimate, but the the numbers would likely be off.

4.48

    a. As sample size increases, the p value decreases. The coin example is perfect here. If you flip a coin 10 times, and get 7 heads it's somewhat normal. If you flip it 1000 times, and get 700 heads it's much stranger. So in this case, the p value would be much lower to compensate.