Homework 6

Steve Bronder
Statistical Inference

December 2, 2014

Exercise 1. Let Y have a t-distribution with 9 degrees of freedom. Using (only) repeated samples from a standard normal distribution, find the following:

- $P(Y < -1)$
- $P(-1 < Y < 1)$
- $P(Y > 2)$

Answer 1. For this exercise a for loop is used that takes ten samples from the normal distribuion and calculates the t statistic over ten thousand iterations. We then use the `ecdf()` function to compute the emirical cumulative distribution for each probability.

```
Y.t <- NULL

for (i in 1:10000) {
    Ynorm <- rnorm(10)

    Ynorm.mean <- mean(Ynorm)

    Ynorm.var <- var(Ynorm)

    Y.t[i] <- (Ynorm.mean)/(Ynorm.var/sqrt(10))
}

# Y < -1
ecdf(Y.t)(-1)

## [1] 0.1808

# -1 < Y < 1
1 - ecdf(Y.t)(-1) - (1 - ecdf(Y.t)(1))

## [1] 0.6386
```

```
# Y > 2
1 - ecdf(Y.t)(2)

## [1] 0.0618
```

- $P(Y < -1) = 17.72\%$
- $P(-1 < Y < 1) = 64.27\%$
- $P(Y > 2) = 6.24\%$

Exercise 2. Use the `pt()` command in **R** to confirm your answers for exercise 1. Then, state how each of the three probabilities would change (greater/less than) if the random variable Y were replaced with:

- a standard normal variable Z
- a t-distributed random variable w, with 4 degrees of freedom

Answer 2. Checking answers

```
#
# Y < -1
pt(-1,9)

## [1] 0.1717182

# -1 < Y < 1
pt(-1,9,lower.tail=FALSE)-pt(1,9,lower.tail=FALSE)

## [1] 0.6565636

# Y > 2
1-pt(2,9)

## [1] 0.03827641
```

These answers are close the answers received from the resampling method. Next, Z is subsituted into Y in order to see the change.

```
# Standard normal random variable Z
# Z < -1
pnorm(-1)

## [1] 0.1586553
```

```r
# -1 < Z < 1
pnorm(-1,lower.tail=FALSE)-pnorm(1,lower.tail=FALSE)

## [1] 0.6826895

# Z > 2
1-pnorm(2)

## [1] 0.02275013
```

The tail probabilites grow smaller and the within probabilities grow larger. This is because a standard normal random variable is not as fat at the tails as the t distribution. Now, w is substituted into Y in order to see the change.

```r
# t distribution random variable w, with 4 degrees of freedom

# W < -1
pt(-1,4)

## [1] 0.1869505

# -1 < w < 1
pt(-1,4,lower.tail=FALSE)-pt(1,4,lower.tail=FALSE)

## [1] 0.626099

# Y > 2
1-pt(2,4)

## [1] 0.05805826
```

The tail probabilites grew fatter. This is because as the number of degrees of freedom increases, the tails grow thinner.

Exercise 3. State clearly the null and alternative hypothesis, calculate the test statistic, and report the p-value using **R**

Answer 3. To test the claim that the population variance is greater than 6.2, we will perform a χ^2 with a null of the population variance being less than orequal to 6.2 and an alternative that it is greater than 6.2

$$H_0 : \sigma^2 \leq 6.2 \tag{1}$$
$$H_A : \sigma^2 > 6.2 \tag{2}$$

```r
((18-1)*6.5)/6.2

## [1] 17.82258

pchisq(17.82258,17,lower.tail=FALSE)

## [1] 0.4001157
```

With a p-value of .4 and $\alpha = .01$ we do not reject the null and conclude that the popluation variance is less than or equal to 6.2.

Exercise 4. State clearly the null and alternative hypothesis, calculate the test statistic, and report the p-value using **R**.

Answer 4. To test the claim that the there is a difference in the standard deviations, we will perform a f test with a null of the standard deviation of fast food one (sd_1) and standard deviation of fast food two (sd_2) being the same an alternate that they are different.

$$H_0 : sd_1 = sd_2 \tag{3}$$
$$H_A : sd_1 \neq sd_2 \tag{4}$$

```r
4.8/3.5

## [1] 1.371429

pf(1.371429,12,10)/2

## [1] 0.3435938
```

With a p-value of .34 and $\alpha = .01$ we do not reject the null and conclude the standard deviations are equal.

Exercise 5. State clearly the null and alternative hypothesis, calculate the test statistic, and report the p-value using **R**.

Answer 5. To test the claim that the average single-person household received at least thirty seven phone calls per month, we will perform a t test with a null of the sample mean $(\bar{X})$ being equal to the population mean (μ) and an alternative that they are not equal.

$$H_0 : \bar{X} = \mu \tag{5}$$
$$H_A : \bar{X} \neq \mu \tag{6}$$

```r
# Exercise 5

(34.9-37)/(6/sqrt(29))

## [1] -1.884808

pt(-1.88,28)/2

## [1] 0.0176375
```

With a p-value of .017 and $\alpha = .05$, we reject the null and conclude that there is a difference in the sample mean and population mean.