LING82100: homework 3 solution

```
> library(e1071)
> data <- read.table("casillas.tsv", header = TRUE)
```

1 /d/

```
> d <- subset(data, consonant == "d")
> shapiro.test(d$vot)

Shapiro-Wilk normality test

data:  d$vot
W = 0.81233, p-value < 2.2e-16

> skewness(d$vot)
[1] -0.05618133
> kurtosis(d$vot)
[1] -1.82546
> median(d$vot)
[1] -16.035
> wilcox.test(d$vot)

Wilcoxon signed rank test with continuity correction

data:  d$vot
V = 16274, p-value = 2.275e-16
alternative hypothesis: true location is not equal to 0
```

This data set does not satisfy the normality assumptions of the t-test: the Shapiro-Wilk test is significant at $\alpha = .05$ ($W = 0.812$, $p < .001$) and the distribution is substantially platykurtic (-1.83). The sample median is -16.04. We apply the one-sample Wilcoxon test, for which the result is significant at $\alpha = .05$ ($W = 16274$, $p < .001$), leading us to reject the null hypothesis.

2 Spanish /t/

```
> es.t <- subset(data, language == "spanish" & consonant == "t")
```

```
> shapiro.test(es.t$vot)

Shapiro-Wilk normality test

data:  es.t$vot
W = 0.9948, p-value = 0.7853
> skewness(es.t$vot)
[1] -0.008857618
> kurtosis(es.t$vot)
[1] 0.1939441
> mean(es.t$vot)
> mean(es.t$vot)
[1] 10.71572
> t.test(es.t$vot)

One Sample t-test

data:  es.t$vot
t = 15.653, df = 179, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
  9.364799 12.066646
sample estimates:
mean of x
 10.71572
```

This data set appears to be consistent with the normality assumptions of the *t*-test; the Shapiro-Wilk test does not reject the null hypothesis ($W = 0.995$, $p = .785$) at $\alpha = .05$. and skewness and excess kurtosis are both quite mild. The sample mean is 10.72. The results of a one-sample *t*-test are significant at $\alpha = .05$ ($t = 15.65$, d.f. $= 179$, $p < .001$; 95% CIs: 9.36, 12.066), leading us to reject the null hypothesis.

3 English coronal stops

```
> en <- subset(data, language == "english")
> qqnorm(en$vot)
> qqline(en$vot)
> skewness(en$vot)
[1] 0.4022722
> kurtosis(en$vot)
[1] -1.050855
> with(en, tapply(vot, consonant, median))
     d      t
15.715 50.200
```

```
> with(en, wilcox.test(vot ~ consonant))

Wilcoxon rank sum test with continuity correction

data:  vot by consonant
W = 549, p-value < 2.2e-16
alternative hypothesis: true location shift is not equal to 0
```

This data set does not satisfy the normality assumptions of the *t*-test: the QQ-plot does not fit it well, it has a slight positive skew (0.40), and is platykurtic (-1.05). The sample medians are 15.72 for /d/ and 50.20 for /t/. We apply the two independent samples Wilcoxon test, for which the result is significant at $\alpha = .05$ ($W = 549$, $p < .001$), leading us to reject the null hypothesis.