Question 11.24

a. $\alpha = 0.01$.
b. $H_0 : \mu_{did} = \mu_{didnot}$ (or $\mu_{did} - \mu_{didnot} = 0$) where *did* corresponds to the "did evacuate" and *didnot* corresponds to the "did not evacuate" groups. The H_A is $\mu_{did} > \mu_{didnot}$ (or $\mu_{did} - \mu_{didnot} > 0$).
c. A two-sample t-test is required because a quantitative variable (commitment level score) was measured on two populations ("did evacuate" and "did not evacuate") that were **IN**dependent (no connection between individuals that did and those that did not evacuate their pets) and two population means are being compared in the null hypothesis.
d. The data appear to be part of a voluntary response observational study without clear randomization.
e. The assumptions are met because the two groups appear to be independent, the sample size ($n_{did} + n_{didnot}$=241) is > 40 and we were told that the p-value for Levene's test of the homogeneity of variance test is "large" (i.e., > 0.01) so that it appears that the variances between groups are equal. Therefore, the test statistic computed below should reasonably follow a t-distribution with $n_{did}+n_{didnot}-2 = 239$ df.
f. The statistic is $\bar{x}_{did} - \bar{x}_{didnot} = 7.694 - 6.640 = 1.054$. The pooled sample variance is,

$$s_p^2 = \frac{(116-1)3.410^2 + (125-1)3.102^2}{116+125-2} = 10.58749$$

The standard error of the statistic is,

$$SE_{\bar{x}_{did}-\bar{x}_{didnot}} = \sqrt{10.58749\left(\frac{1}{116}+\frac{1}{125}\right)} = 0.4194894$$

g. The t test statistic is $\frac{1.054-0}{0.4194894} = 2.513$ with $116 + 125 - 2 = 239$ df.
h. The p-value is $p = 0.0063$.
i. The H_0 is rejected because the $p - value < \alpha = 0.01$.
j. The average commitment to animals was significantly greater for those people that did evacuate their pets as compared to those people that did not evacuate their pets.
k. The 99% lower bound is $1.054 - 2.342051 * 0.419 = 0.073$. One is 99% confident that the mean level of commitment for those that evacuated their pets is at least 0.073 greater than those that did not evacuate their pets.

Question 11.25

a. $\alpha = 0.05$.
b. $H_0 : \mu_i = \mu_r$ (or $\mu_i - \mu_r = 0$) where i corresponds to the "impacted" and r corresponds to the "reference" groups. The H_A is $\mu_i \neq \mu_r$ (or $\mu_i - \mu_r \neq 0$).
c. A two-sample t-test is required because quantitative variable (methyl mercury level) was measured on two populations ("impacted" and "reference" sites) that were **IN**dependent (no connection between sites at the two locations) and two population means are being compared in the null hypothsis.
d. The data appear to be part of an observational study without clear randomization.
e. The assumptions are not quite met but we will continue anyway. The two groups appear to be independent. The variances appear to be equal because the p-value from the Levene's test is "large" (i.e., $p = 0.3595 > 0.05$; Table 1). The sample size ($n_i + n_r$=13) is not > 15 and there is not enough data to make useful histograms. Therefore, the test statistic computed below may not follow a t-distribution with $n_c + n_d - 2$=11) df.

Table 1. Results from the Levene's test for testing equal variances among impacted and reference groups.

```
      Df F value Pr(>F)
group  1  0.9144 0.3595
      11
```

f. The statistic is $\bar{x}_c - \bar{x}_d = 0.0573$ - $0.0386 = 0.0188$ (Table 2).

Table 2. Results from the two-sample t-test for testing that the mean mercury levels in mussels differed between the impacted and reference sites.

```
t = 1.5259, df = 11, p-value = 0.1553
95 percent confidence interval:
 -0.008301233  0.045825042
sample estimates:
 mean in group impacted mean in group reference
             0.05733333              0.03857143
```

g. The t test statistic is 1.5259 with 11 df (Table 2).
h. The p-value is $p = 0.1553$ (Table 2).
i. The H_0 is not rejected because the $p - value > \alpha$.
j. The methyl mercury levels do not appear to be different in mussels from the impacted and reference sites.

Appendix – R Commands

```
( distrib(2.512578,distrib="t",df=239,lower.tail=FALSE) )
( distrib(0.99,distrib="t",df=239,type="q",lower.tail=FALSE) )
```

```
impacted <- c(0.011,0.054,0.056,0.095,0.051,0.077)
reference <- c(0.031,0.040,0.029,0.066,0.018,0.042,0.044)
df <- data.frame(mhg=c(impacted,reference),
grp=factor(rep(c("impacted","reference"),c(length(impacted),length(reference)))))
( m.lt <- levenesTest(mhg~grp,data=df) )
( m.t <- t.test(mhg~grp,data=df,var.equal=TRUE) )
```

Notes From Professor

- If you do not use informative subscripts, then make sure that you explicitly define what the “1” and “2” subscripts mean.
- Make sure that you provide evidence for the fact that the two groups are independent.
- Note in 11.25 that the sample size is so small that the assumption is that the sample distributions must be normal; “not strongly skewed” is not adequate with this small sample size.
- Note that in 11.25, it is easier to create a CSV file to load the data.
- Note that in 11.25, if you decide to use `Summarize()` to get the sample means then do NOT set `digits=1` because this rounds the means to too few significant digits. In other words, make sure to take note of the scale that the measurements were made on when rounding. Note that `digits=3` would be more appropriate.