Lecture/text homework assignment # 8

Note: Please circle your answers when appropriate!

For problems 2 - 4 you need to clearly state H_0, H_1, give α (if not given), and clearly state your conclusion (in words!).

1) ***Use R for all parts of this problem.*** Use the data on heights you collected in recitation a few weeks ago. Some of this is a bit repetitive from homework 7, but these are important concepts.

(a) Do a *t*-test and test the following:

H_0: $\mu = \bar{y}$ *for your recitation section* vs. H_1: $\mu \neq \bar{y}$

Use $\alpha = .01$

(if, for example, your $\bar{y} = 64.72$, then you would do H_0: $\mu = 64.72$)

(c) now repeat but add 4.75 to your $\bar{y}$ (so you'd do H_0: $\mu = 69.47$ (= 64.72 + 4.75))

(d) repeat (c) but use $\alpha = .00000000001$ *(but see part (e) before you do this)*.

(e) ***Why do you not need to tell R about α?*** ***Make sure you understand this!!***

(f) what kind of error did you probably make in part (d)? (Just write out the answer for this question).

R instructions:

To test, for example, that $\mu = 64.72$, you need to give R the value of μ:

```
t.test(height,mu = 64.72)
```

(assuming you named your variable "height"; of course, you may have to do `data$height` or something similar depending on how you have your data in R).

Since R gives you *p*-values (= probabilities) you don't need to calculate or use t^* (although R automatically prints this if you want to use it).

Do not use R for the rest of the problems:

2) You want to figure out if there is a difference in lengths between male and female kestrels (kestrels are small falcons found all over Northern Virginia). You catch 12 females and 10 males with the following results (in cm):

Females:	28.3	28.2	25.2	27.2	26.2	26.8	27.3	28.9	24.0	26.9	28.3	27.2
Males:	23.4	24.3	20.5	24.2	22.1	22.3	22.6	23.1	22.0	23.9		

(a) Is there a difference in lengths? Use $\alpha = .01$. Note that you have to calculate *everything* yourself. Make sure you give H_0 and H_1 (symbols are okay), and clearly write out your conclusion.

3) A study of estrogen levels in two different groups of women finds the following results (in pg/mL):

Group A: 18.7 20.6 20.7 19.7 19.9 19.4 20.2 21.6 18.8 14.1 21.6 16.2 21.7 20.8 19.3 21.3 19.9 20.8 23.2

Group B: 15.2 36.2 27.5 4.7 24.5 29.4 25.9 62.8

Some summary statistics to help you:

	$\bar{y}$	s	n
Group A:	19.921	2.0457	19
Group B:	28.275	16.9169	8

Is there a difference in estrogen levels? (Note: $d.f. = 7.0864$ for Welch's t-test). Use $\alpha = 0.05$. Make sure you give H_0 and H_1 (symbols are okay), and clearly write out your conclusion.

4) Repeat (3), but this time assume equal variances (i.e., use the classic t-test). Use the same level of α you used before. Make sure you give H_0 and H_1 (symbols are okay), and clearly write out your conclusion.

5) Now let's compare the tests from problems (3) and (4)

(a) Which test (problem (3) or problem (4)) lets you reject the null hypothesis?

(b) Which test do you **think* has more power?

Usually, but not always(!!), the test with the most power has a lower p-value.

(c) If you don't know that the population variances are equal, which test should you use?

(d) Which test ***should*** you use here? (Refer to (c)).

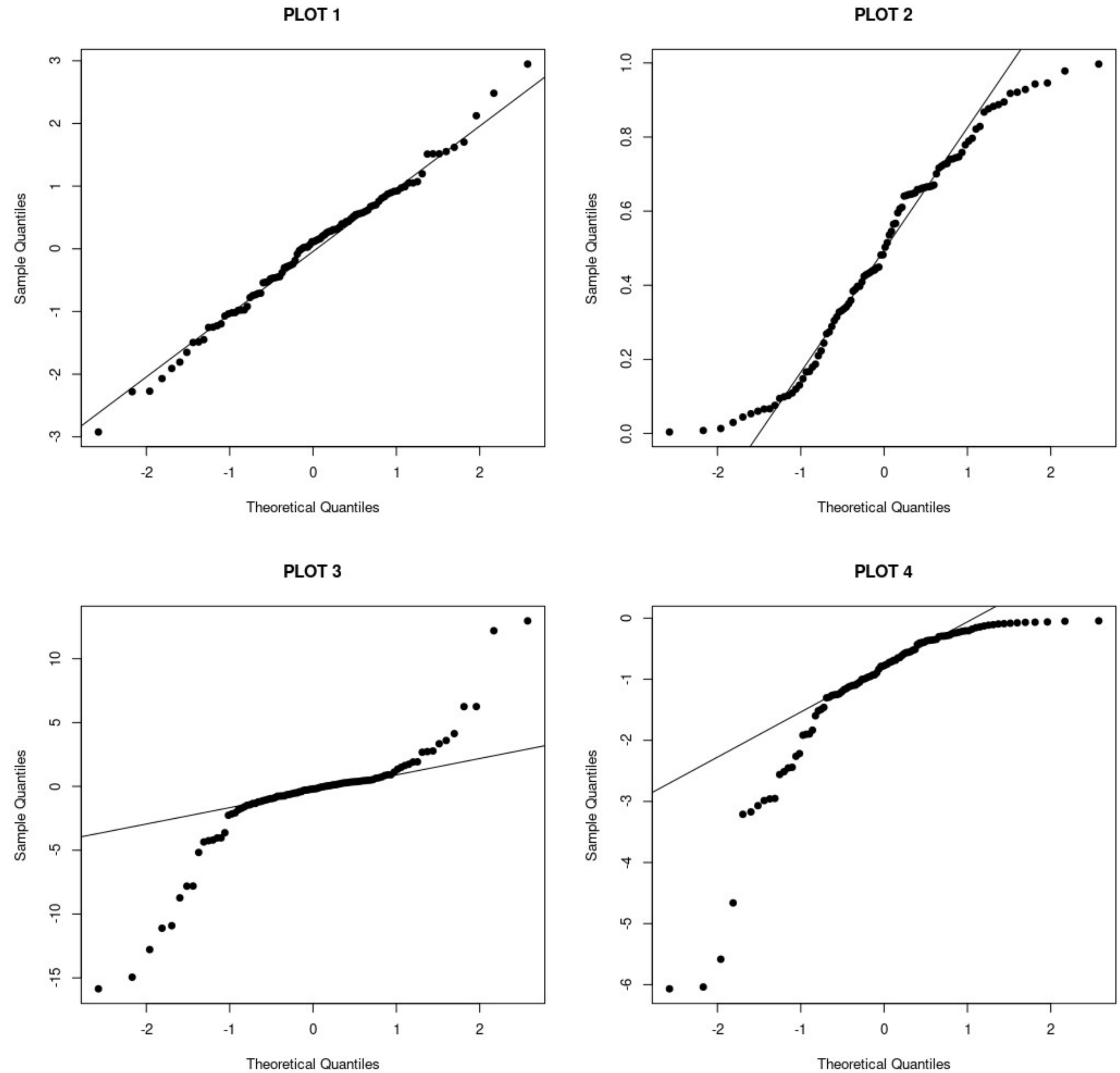
Big hint and comment:

This is an example of when the classic (= equal variance) t-test can make a serious mistake. Rejecting a null hypothesis when it is not appropriate is a pretty serious mistake. Note also that the sample sizes are very different.

(If you're theoretically inclined, here's what happens: even though we set $\alpha = 0.05$, the classic t-test makes a type I error at a much higher rate than 5%. In other words, despite setting $\alpha = 0.05$, the actual value of $\alpha > 0.05$, which is obviously not good!)

6) Match the *q-q* plots below with the appropriate letter:

(A) long tailed but symmetrical
(B) approximately normal
(C) skewed right
(D) skewed left
(E) short tailed
(F) inverse normal



7) Let's use the some more data on male kestrel length (a different sample than in problem (2). this time we have 12 birds ($n = 12$). The data are already, sorted to help you. Most of the normal scores are also given:

Length(cm):	21.0	21.1	22.2	22.7	22.9	23.2	23.7	23.9	24.0	24.8	28.5	29.1
Z-scores:			-0.812	-0.549	-0.319	-0.105	0.105	0.319	0.549	0.812		

(a) Calculate the missing normal scores (the first two and the last two).

(b) Now construct a *q-q* plot (normal probability plot). Do this by hand.

Be prepared to discuss these problems in recitation the week of March 31st.