

Homework 7  
MATH 4830, Spring 2006  
due Thurs. March 9

1. p. 106, Problem 14.

To get you started, here are the 10 possible groupings:

1. 1,5  
4,8,9
2. 1,4  
5,8,9
3. 1,8  
5,4,9
4. 1,9  
5,4,8
5. 5,4  
1,8,9
6. 5,8  
1,4,9
7. 5,9  
1,4,8
8. 4,8  
1,5,9
9. 4,9  
1,5,8
10. 8,9  
1,5,4

2. Using the data from Problem 1, let's use the `myperm` function that we wrote in class to approximate the permutation test  $p$ -value. You can have the `myperm` function for your very own, by sourcing the file by:

```
> source("http://www-math.cudenver.edu/~bbailey/4830/myperm.r")
```

You can check that you have the function by typing `myperm` at the R command line and hitting return.

We need some data for the problem, so let's make 2 vectors in R, call the `myperm` function, and count how many differences are less than or equal to -4. The `set.seed` function will generate the same sequence of random numbers so that we all get the same answer!

```
> g1 <- c(1,5)
> g2 <- c(4,8,9)
> set.seed(1)
> temp <- myperm(100,g1,g2)
> sum(temp <= -4)/100
```

How does this answer compare to your exact answer in Problem 1?

3. Guinea Pig Lifetimes. We will analyze the treatment effect on the lifetimes of guinea pigs in Chapter 1, Exercise 11. The data can be read into R by:

```
> gpigs <- read.table("http://www-math.cudenver.edu/~bbailey/4830/gpigs.dat",
header=T)
```

Make a new R object `newgpig` that is a list with components `newgpigs$bacilli` and `newgpigs$control` and contains the survival times in days for the control and treatment groups. You can do this by the list command,

```
> newgpigs <- list(control=gpigs$LIFETIME[gpigs$GROUP=="control"],
bacilli=gpigs$LIFETIME[gpigs$GROUP=="bacilli"])
```

Type `newgpigs` at the R command line to see if your new dataset looks correct.

(a) Make side-by-side boxplots of the data. Does it look like the means and variances are the same for the two groups?

(b) Let's use the `attach` function so that we don't have to type so much! After you attach the dataset, you can use `control` and `bacilli`.

```
> attach(newgpigs)
> control
> bacilli
```

For fun, let's compare 4 different 2-sample tests and the one-sided  $p$ -values. You should be able to write down the null and alternative hypothesis for each tests.

```
> t.test(control, bacilla, alternative="g", var.equal=T)
```

(If you perform the `var.test` on the data, you will reject the null hypothesis of equal population variances! So, you should use the Welch's t-test.)

```
> t.test(control, bacilla, alternative="g", var.equal=F)
```

```
> wilcox.test(control, bacilla, alternative="g")
```

```
> test <- myperm(1000,control,bacilla)
> diffmean <- mean(control) - mean(bacilla)
> sum(test >= diffmean)/1000
```

What can you conclude from the above 4 tests?