Homework # 4:

Note: Please circle your answers as appropriate.

Remember - for all problems, show your work!

(Do *not* use R for problems 1 - 4)

1) Find the following probabilities:

a) $\Pr\{Z > -0.39\}$ **b)** $\Pr\{Z \leq -0.19\}$ **c)** $\Pr\{-0.45 < Z < 1.00\}$

2) The heart of a blue whale (*Baleonoptera musculus*) averages about 181.4 kg in weight. So let's assume μ = 181.4 kg. We'll also assume that σ = 23 kg. Let Y = weight of a blue whale heart, and figure out the following:

a) $\Pr\{Y < 170\}$ **b)** $\Pr\{Y > 170\}$ **c)** $\Pr\{Y \geq 170\}$ (Hint: what is $\Pr\{Y = 170\}$)

d) $\Pr\{Y < 80\}$ **e)** $\Pr\{Y < 165\}$ **f)** $\Pr\{165 < Y < 181.4\}$

g) add (e) and (f); are you surprised? ***Why or why not?***

h) for (d), is the probability really 0? ***Why or why not?***

i) add (a) and (b). Are you surprised? ***Why or why not?***

3) Refer to problem (2). Give the following percentiles (=quantiles):

a) 3 **b)** 39 **c)** 50 **d)** 61 **e)** 97.5 **f)** 99.5 **g)** 99.9

4) The average diastolic blood pressure for someone in their 30's is about 74.3 mm/Hg (let's assume μ = 74.3 mm/Hg). For these people the standard deviation for systolic blood pressure is about 12.3 mm/Hg (again, let's assume σ = 12.3 mm/Hg).

a) Give the values for diastolic blood pressure for the middle 80% of people.
(Hint/comment: if, for example, you want values for the middle 65%, you need to figure out many percent go in each tail. In other words, if you want the middle 65% that implies that you have 35% left that you need to divide equally into each tail. You will need two numbers in your answer. Draw a picture to help you.)

b) Give the values for diastolic blood pressure for the middle 95% of people.

c) Give the 97.5th percentile.

d) Are you surprised by the answers to (b) and (c)? ***Explain***. If you're not sure what's going on, draw some pictures of the normal curves.

e) Calculate the 00.1th percentile. **(f)** *(See next page.)*

f) Suppose you come across a person with a diastolic blood pressure of 35 mm/Hg. Would you think anything was unusual about this person? ***Why or why not (make sure you explain this!!).***

Now let's do some R:

5) We want to explore the effect of parameters on a binomial distribution (this is a bit similar to last week's assignment).

Start up R/RStudio

a) First we'll do our coin example from lecture but use 23 tosses instead of 10. Let's calculate those probabilities and plot a graph so we know what our "distribution" looks like:

In the script window or on the command line do the following:

`y <- c(0:23)`	we're giving y the numbers 0 through 18.
`y`	if you want to see "y", you can do this step (not necessary).
`pr <- dbinom(y,23,.5)`	this gives us all the probabilities for the binomial. ".5" is the probability of success, "23" is the number of trials, and y is the number we're interested in (in this case, all the values from 0 to 23). Make sure you put in the numbers in the right order.
`pr`	this should give you all the probabilities.
`barplot(pr,names.arg = y)`	this generates a plot of "pr" (our probabilities) and tells it to use the contents of "y" to label our x axis.

What does the graph look like?

Make sure you present your graph as well. Instructions for cutting and pasting graphs were given in the last homework assignment.

b) Let's try some different "parameters". In the instructions for (a), let's do 9 tosses (change the first step to `y <- c(0:9)` and change .5 to .15 (the third step above) and see what happens. This is kind of like having a coin that comes up heads 15% of the time.

What does the graph look like now? What changed? Why? Which probabilities are now higher? Why?

Again, make sure you present your graph.

6) Now let's try a normal distribution. The first plot will be easy, but the second will be more difficult to do, so follow the instructions carefully. We'll us our blue whale hearts from problem 2 and see what happens when we change the parameters.

```
x <- seq (66.4, 296.4, length.out=350)
```

This gives us 350 equally spaced values of x between 66.4 and 296.4. "seq" let's us generate a sequence of numbers, and "length.out" tells us how many numbers we want in the sequence.

```
plot(x, dnorm(x, mean = 181.4, sd = 23), col = "red", type = "l", ylim = c(0,0.018),
  ylab = "density", xlab = "Weight of blue whale hearts (kg)",
  main = "Several normal curves using different parameters for blue whale hearts")
```

"plot" is a command to generate graphs (it's quite versatile, and we can't explain everything here). "dnorm" is essentially the equation for the normal distribution. This command will plot x and the value of x at each of the 350 points using the normal equation. The "mean", "sd" and "col" functions should be self-explanatory. type = "l" (that's lower case "L", not "one") generates lines instead of points (try it without type = "l" to see the difference). "ylim" sets the limits on the y-axis; we do this so that the curve doesn't touch the top of the box that's around the graph. "ylab" and "xlab" were covered in a previous homework but should be fairly obvious. "main" gives the main title over the graph.

You can hold off on copying and pasting your graph until you've done all of the steps below (you only need to hand in one graph for problem 6). See also the optional step at the end of problem 6 if you want to make your plot look nicer.

At this point you should have a standard normal curve. Let's see what happens when we change the parameters.

(a) First, let's change the mean from 181.4 to 204.4:

```
lines(x, dnorm(x, mean = 204.4, sd = 23), col = "blue")
```

"lines" will plot lines, but without erasing the previous plot. In other words, it'll plot the lines on the same graph as before. We changed the color to blue so you can see which plot is which.

What happened to the normal distribution? How is it different from the first distribution you plotted?

(b) Now let's adjust the standard deviation:

```
lines(x, dnorm(x, mean = 181.4, sd = 46), col = "dark green")
```

You should know how this works at this point.

Again, what happened to the normal distribution? How is it different from the previous normal distributions you plotted? (What happens to the height of our distribution?)

(c) Now let's adjust both the mean and standard deviation.

```
lines(x, dnorm(x, mean = 192.9, sd = 34.5), col = "purple")
```

How is this different from the above normal distributions?

(d) ***Can you understand how important the parameters of a distribution are? What do the parameters of a distribution tell us about the distribution?***

(f) At this point your graph should have four curves plotted on top of each other in your final graph. Copy this graph and paste it into a word processor or whatever you want to use so that you can print a copy.

Optional: if you want to make your graph look even nicer, you can add the following:

```
legend (70,.017, c("mean = 181.4, sd = 23","mean = 204, sd = 23",
        "mean = 181.4, sd = 46", "mean = 192.9, sd = 34.5"),
        col = c("red","blue","darkgreen","purple"), lty = c(1,1,1,1))
```

The legend command can get pretty complicated. Here's a breakdown of what's going on:

The first two numbers (70,.017) give the x and y coordinates of the top left corner of the legend box. These coordinates are the same ones used in the plot.

The first "c(...)" contains the actual labels for the legend. In this case we want to combine four "labels", one for each line in our plot ("c" means combine or concatenate).

The "col = c(...)" part gives the four colors we want to use (we need to make sure they match the ones used in the plot and are in the same order as the first c(...) which has the labels).

Finally, we need to tell R what we used for plotting - we used lines, and "lty(...)" tells R to draw lines of the appropriate colors in our legend. The ones (1's) inside the lty command tell R to draw simple lines (R can draw many different kinds of lines - try using 2's or 3's here and you'll see what this means).

****

Copying and pasting graphs from R:

See instructions from the previous set of homework.

****

Problems are due in recitation Monday, June 15th.