**EJERCICIOS JUEVES**

**Clase 8: Funciones**

8.1 Write a function to calculate the mean of a series of numbers. Compare the answer to the inbuilt R function

8.2 Write a function to calculate standard error of a series of numbers

8.3 Write a function that prints out the mean and standard deviation of a series of numbers

8.4 Modify the above function to include standard error

8.5 Write a function to convert Fahrenheit to Celsius

8.6 Write a function to calculate the Simpson index

$$D = \sum_{i=1}^{S} p_i^2.$$

where $p_i$ is the fraction of all organisms which belong to the i-th species. Calculate the Simpson index for the species richness data in the vegetation2 dataset, for all observations and by transect.

8.7 Write a function to calculate the Shannon index

$$H' = -\sum_{i=1}^{S} p_i \ln p_i$$

where $p_i$ is the fraction of individuals belonging to the i-th species. Calculate the Simpson index for the species richness data in the vegetation2 dataset, for all observations and by transect.

8.8. Write a single function to calculate the mean, sd, se, Simpson index and Shannon index. Output the results in a list.

EXTRA: 8.9 Make a nice looking figure (dot or strip plot) that shows the Simpson index Shannon index, mean and error bars.

**Clase 9: Loops**

9.1 Write a for loop to print the values 1 to 10.

9.2 Write a for loop to print Celsius degrees in Fahrenheit for the degrees 25 to 32.

9.2 Instead of tapply, or other functions, write a loop to calculate the Simpson index for each transect in the vegetation data

**Clase 10: RANDOM DATA**

10.1 Generate 10 random numbers from a uniform distribution on [0,10]. Use R to find the maximum and minimum values.x

10.2 Generate 10 random normal numbers with mean 5 and standard deviation 5 (normal(5,5)). How many are less than 0? (Use R)

10.3 Generate 100 random normal numbers with mean 100 and standard deviation 10. How many are 2 standard deviations from the mean (smaller than 80 or bigger than 120)?

10.4 Toss a fair coin 50 times (using R). How many heads do you have?

10.5 Roll a \die" 100 times. How many 6's did you see?

10.6 Select 6 numbers from a lottery containing 49 balls. What is the largest number? What is the smallest? Answer these using R.

10.7 For normal(0,1), find a number z* solving P(Z <= z*) = .05 (use qnorm).

10.8 For normal(0,1), find a number z* solving P(-z* <= Z <= z*) = 0.05 (use qnorm and symmetry).

10.9 How much area (probability) is to the right of 1.5 for a normal(0,2)?

10.10 Make a histogram of 100 exponential numbers with mean 10. Estimate the median. Is it more or less than the mean?

10.11 Can you figure out what this R command does?

```
rnorm(5,mean=0,sd=1:5)
```

10.12 Use R to pick 5 cards from a deck of 52. Did you get a pair or better? Repeat until you do. How long did it take?

**Clase 11: Confidence intervals**

11.1 Create 15 random numbers that are normally distributed with mean 10 and s.d. 5. Find a 1-sample z-test at the 95% level. Did it get it right?

11.2 Do the above 100 times. Compute what percentage is in a 95% confidence interval. Hint: The following might prove useful

```
f=function () mean(rnorm(15,mean=10,sd=5))
SE = 5/sqrt(15)
xbar = simple.sim(100,f)
alpha = 0.1;zstar = qnorm(1-alpha/2);sum(abs(xbar-10) < zstar*SE)
alpha = 0.05;zstar = qnorm(1-alpha/2);sum(abs(xbar-10) < zstar*SE)
alpha = 0.01;zstar = qnorm(1-alpha/2);sum(abs(xbar-10) < zstar*SE)
```

11.3 The t-test is just as easy to do. Do a t-test on the same data. Is it correct now? Comment on the relationship between the confidence intervals.

11.4 Find an 80% and 95% confidence interval for the median for the `exec.pay` dataset.

11.5 For the simple data set `rat` do a t-test for mean if the data suggests it is appropriate. If not, say why not. (This records survival times for rats.)

11.6 Repeat the previous for the simple data set `puerto` (weekly incomes of Puerto Ricans in Miami.).

11.7 The median may be the appropriate measure of center. If so, you might want to have a confidence interval for it too. Find a 90% confidence interval for the median for the Simple data set malpract (on the size of malpractice awards). Comment why this distribution doesn't lend itself to the z-test or t-test.

DATASETS:

```
puerto <-
c(150, 280, 175, 190, 305, 380, 290, 300, 170, 315, 280, 255,
335, 180, 200, 210, 350, 360, 550, 225, 245, 175, 180, 260, 320,
345, 180, 225, 275, 280, 325, 350, 190, 220, 265, 270, 335, 310,
270, 240, 235, 355, 370, 260, 280, 350, 390, 250, 375, 250)


rat <-
c(152, 152, 115, 109, 137, 88, 94, 77, 160, 165, 125, 40, 128,
123, 136, 101, 62, 153, 83, 69)
```

```
exec.pay <-
  c(136, 74, 8, 38, 46, 43, 9, 9, 12, 11, 20, 9, 95, 34, 7, 14,
    39, 12, 29, 21, 60, 35, 17, 36, 29, 162, 88, 31, 6, 135, 13,
    20, 9, 14, 28, 42, 10, 35, 2, 16, 28, 42, 142, 33, 134, 23, 34,
    16, 13, 167, 9, 22, 39, 28, 30, 22, 14, 9, 25, 106, 32, 30, 89,
    89, 47, 17, 26, 1231, 6, 103, 48, 24, 11, 19, 13, 29, 20, 45,
    3, 33, 41, 7, 11, 10, 22, 36, 7, 19, 41, 40, 10, 15, 93, 67,
    29, 25, 91, 38, 2510, 5, 32, 65, 0, 13, 27, 16, 21, 6, 0, 28,
    8, 13, 71, 36, 11, 106, 37, 41, 13, 900, 38, 24, 15, 27, 12,
    12, 22, 40, 49, 22, 118, 48, 10, 1, 36, 155, 9, 34, 29, 12, 0,
    28, 21, 32, 18, 52, 29, 13, 199, 40, 11, 51, 45, 43, 31, 5, 18,
    15, 25, 9, 18, 13, 58, 22, 40, 34, 16, 31, 27, 15, 23, 49, 60,
    28, 74, 42, 24, 17, 9, 61, 20, 23, 26, 31, 167, 19, 14, 13, 146,
    283, 12, 53, 26, 16, 29, 51, 15, 22, 27)
```