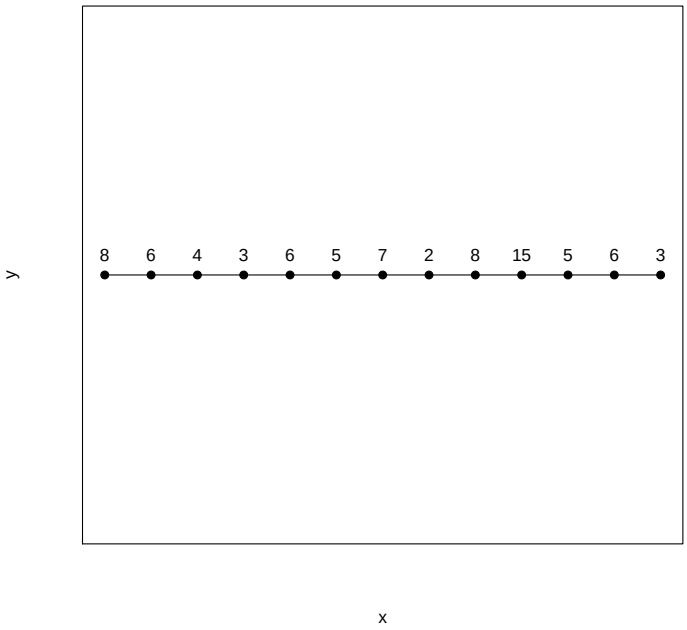
University of California, Los Angeles
Department of Statistics

Statistics C173/C273 **Instructor: Nicolas Christou**

Exam 1
23 January 2015

Name: _____

Problem 1 (30 points)
You are given the following spatial data in one dimension:



The distance between points is 5 meters, and the value of the variable of interest is shown at each data point. Answer the following questions:

a. Use the classical and the robust variogram estimator to compute by hand $2\hat{\gamma}(5)$. Please show all your work.

b. Compute the box plot of the cloud using the classical estimator when $h = 5$. Please submit your `R` code and the boxplot. Note: For the classical estimator we use $[z(s_i) - z(s_j)]^2$.

c. Compute the box plot of the cloud using the robust estimator when $h = 5$. Please submit your `R` code and the boxplot. Note: For the robust estimator we use $[z(s_i) - z(s_j)]^{0.5}$.

Problem 2 (35 points)
Answer the following questions:

a. Do not use R! Suppose that two points s_1 and s_2 are 100 m apart. Calculate the variance of a linear combination: $\hat{Z}(s_0) = Z(s_1) + Z(s_2)$. Note: $Z(s)$ is a second order stationary function with a spherical semivariogram with a range $\alpha = 250$ m and sill $c_1 = 3$. Assume the nugget is zero.

b. Use R! The table below gives 35 grades on a regular $100m \times 100m$ grid. Calculate the classical variogram in the two principal directions (NS and EW) using $h_1 = 100m, h_2 = 200m, h_3 = 300m$, without angular tolerance.

3.6	2.8	4.1	4.6	5.3	3.8	4.2
3.6	3.5	5.2	4.5	6.1	4.4	4.0
5.1	3.8	4.9	3.3	5.7	6.2	6.3
4.2	4.0	5.6	4.2	4.9	5.3	4.4
4.6	5.7	6.1	5.4	4.7	5.2	6.0

You can create the data set as follows:

```
x <- c(rep(0,5),rep(100,5),rep(200,5),rep(300,5),rep(400,5),rep(500,5),rep(600,5))

y <- c(rep(seq(0,400,100), 7))

z <- c(4.6,4.2,5.1,3.6,3.6,5.7,4.0,3.8,3.5,2.8,6.1,5.6,4.9,5.2,4.1,5.4,4.2,3.3,4.5,4.6,
4.7,4.9,5.7,6.1,5.3,5.2,5.3,6.2,4.4,3.8,6.0,4.4,6.3,4.0,4.2)
```

Problem 3 (35 points)

Part A

The following data give the location $(x, y$ coordinates) and the calcium content at depth 0-20 cm ($ca20$), for each data point. There are 178 data points. Please access the data at:

```
a <- read.table("http://www.stat.ucla.edu/~nchristo/statistics_c173_c273/
     soil_ca_mg_data.txt", header=TRUE)
```

a. Use `geoR` to construct a bubble plot of the data.

b. Plot the classical omnidirectional;semivariogram up to a maximum distance of 850 m.

c. Fit by eye the spherical semivariogram to the sample semivariogram.

Please submit:

1. The bubble plot.
1. The plot of the sample semivariogram with the fitted spherical semivariogram.
2. The entire `R` code that you used to solve this problem.

Part B

Access the California 2012 Presidential elections results and the results on Proposition 30:

```
a19 <- read.table("http://www.stat.ucla.edu/~nchristo/statistics_c173_c273/
       ca_elections_results_2012.txt", header=TRUE)
```

Create a data frame with the variables `longitude`, `latitude`, and `prop30yes`. But first convert the votes into percentages.

a. Plot the data points (bubble plot) on the map of California.

b. Compute the sample variogram (classical and robust) of the percent of yes votes for Proposition 30. Use direction $3\frac{\pi}{4}$.

c. Use `geoR` to fit the Gaussian model variogram by eye to the sample variogram (classical and robust).

Please submit:

1. The bubble plot.
1. The plot of the sample semivariogram with the fitted Gaussian semivariogram.
2. The entire `R` code that you used to solve this problem.