Fraser D. Neiman **University of Virginia**
Spring 2007

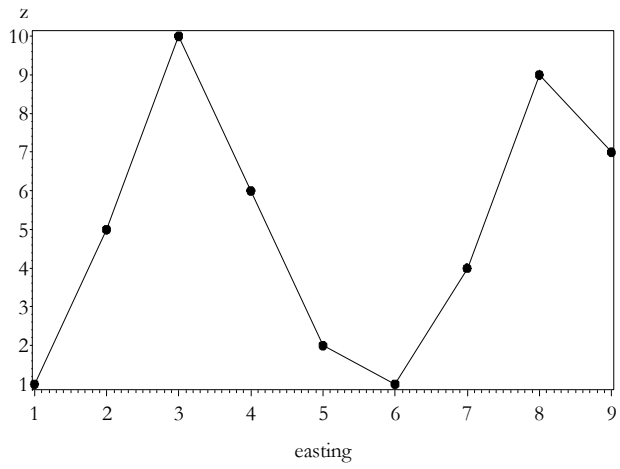
Correlograms Made Easy

1. First, some inspiration:

2. To keep it simple, we'll stick with a single dimension – so the northing coordinates are all the same:

Obs	northing	easting	z
1	1	1	1
2	1	2	5
3	1	3	10
4	1	4	6
5	1	5	2
6	1	6	1
7	1	7	4
8	1	8	9
9	1	9	7

Here is a plot:



3. Start with summary stats:
The mean: 5
The observation re-expressed as deviations from the mean: -4 0 5 1 -3 -4 -1 4 2
The variance ("mean squared deviations from the mean") = 11

4. Get the matrix of distances between pairs of points:

DISTMAT

0	1	2	3	4	5	6	7	8
1	0	1	2	3	4	5	6	7
2	1	0	1	2	3	4	5	6
3	2	1	0	1	2	3	4	5
4	3	2	1	0	1	2	3	4
5	4	3	2	1	0	1	2	3
6	5	4	3	2	1	0	1	2
7	6	5	4	3	2	1	0	1
8	7	6	5	4	3	2	1	0

5. And the matrix of squared differences between the z-values at the pairs of points
DIFFMAT

0	16	81	25	1	0	9	64	36
16	0	25	1	9	16	1	16	4
81	25	0	16	64	81	36	1	9
25	1	16	0	16	25	4	9	1
1	9	64	16	0	1	4	49	25
0	16	81	25	1	0	9	64	36
9	1	36	4	4	9	0	25	9
64	16	1	9	49	64	25	0	4
36	4	9	1	25	36	9	4	0

6. And the matrix of cross-products of deviations from the mean of the z-values at the pairs of points.

CPMAT

16	0	-20	-4	12	16	4	-16	-8
0	0	0	0	0	0	0	0	0
-20	0	25	5	-15	-20	-5	20	10
-4	0	5	1	-3	-4	-1	4	2
12	0	-15	-3	9	12	3	-12	-6
16	0	-20	-4	12	16	4	-16	-8
4	0	-5	-1	3	4	1	-4	-2
-16	0	20	4	-12	-16	-4	16	8
-8	0	10	2	-6	-8	-2	8	4

7. We have decided that the lag distance for our analysis will be 1. So the 0/1 weights matrix for the *first* spatial lag is :

W

0	1	0	0	0	0	0	0	0
1	0	1	0	0	0	0	0	0
0	1	0	1	0	0	0	0	0
0	0	1	0	1	0	0	0	0
0	0	0	1	0	1	0	0	0
0	0	0	0	1	0	1	0	0
0	0	0	0	0	1	0	1	0
0	0	0	0	0	0	1	0	1
0	0	0	0	0	0	0	1	0

This is derived from the distance matrix. For the first spatial lag, **W**[i,j] = 1 if **DISTMAT**[i,j] > 0 and **DISTMAT**[i,j]<= 1.

8. Now to compute the *semivariance* for the first spatial lag:

-multiply the corresponding elements of **W** and **DISTMAT**.
-add up the products.
-divide by the sum of the elements of **W.**
-divide by 2.

Or for the brave:

$$\gamma_h = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\left(z_i - z_j\right)^2}{2\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}}$$

Where *w[i,j]= 1* if the distance between points i and *j* falls in the *h*'th spatial lag, 0 otherwise.

In this case, for the first lag the *semivariance* = *7*.

9. Geary's c is just the *semivariance* divided by the *variance: 7/15 = .636.*

10. Now let's do the *autocovariance* for the first spatial lag:

-multiply the corresponding elements of **W** and **CPMAT.**
-add up the products.
-divide by the sum of **W**

Phormula phobia not allowed:

$$C_h = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\left(z_i - \bar{z}\right)\left(z_j - \bar{z}\right)}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}}$$

11. Moran's I is just the *autocovariance* divided by the *variance* = .281

12. Now we get the weights matrix for the second spatial lag:

W

```
0     0     1     0     0     0     0     0     0
0     0     0     1     0     0     0     0     0
1     0     0     0     1     0     0     0     0
0     1     0     0     0     1     0     0     0
0     0     1     0     0     0     1     0     0
0     0     0     1     0     0     0     1     0
0     0     0     0     1     0     0     0     1
0     0     0     0     0     1     0     0     0
0     0     0     0     0     0     1     0     0
```

This is also derived from the distance matrix. For the second spatial lag, **W**[i,j] = 1 if **DISTMAT**[i,j] > 1 and **DISTMAT**[i,j]<= 2.

13. Now to compute the *semivariance* for the second spatial lag,
-multiply the corresponding elements of **W** and **DISTMAT**.
-add up the products.
-divide by the sum of the elements of **W.**
-divide by 2.
In this case, for the first lag the *semivariance* = 1*7.7*

14. Geary's c is just the *semivariance* divided by the variance: *17.7/15 = 1.61.*

15. Now let's do the *autocovariance* for the second spatial lag:
-multiply the corresponding elements of **W** and **CPMAT.**
-add up the products.
-divide by the sum of **W**

16. Moran's I is just the *autocovariance* divided by the *variance* = -.78

17. We could go on like this for lags 3, 4 , and 5. But you get the idea! Here are the results:

Obs	SEMIVAR	GEARYS_C	MORANS_I	LAGDIST	MEANDIST	NEDGES
1	7.0000	0.63636	0.28125	1	1	8
2	17.7143	1.61039	-0.78896	2	2	7
3	17.0000	1.54545	-0.76705	3	3	6
4	8.7000	0.79091	0.10227	4	4	5
5	0.3750	0.03409	0.97159	5	5	4

Note that NEDGES is the number of point pairs at each lag. That's half the sum of elements in the weights matrix for that lag.