

CrimeStat

Frances F. Burden

Basics

CrimeStat is a spatial statistics package for the analysis of point data such as crime incidents. It is a good tool for analyzing one, or at most, two variables affecting incidents, the incidents themselves, and a secondary variable that might be used for comparison. CrimeStat is not a package for modeling correlates or determinants of crime incidents (i.e., CrimeStat does not offer regression techniques; appropriate packages for this would include SpaceStat or S+SpatialStats). The program calculates various spatial statistics and writes graphical objects as map features to ArcView, ArcGIS, MapInfo, AtlasGIS, Surfer for Windows, and ArcView Spatial Analyst.

CrimeStat Functionality

CrimeStat can perform a wide variety of different statistical routines, and many of the parameters can be user-defined. In addition, CrimeStat has Dynamic Data Exchange (DDE) capabilities, which means that CrimeStat can be accessed from within other programs.

CrimeStat includes capabilities for calculating statistics describing

- (a) the [spatial distribution](#) of incidents (such as the mean center, center of minimum distance, standard deviational ellipse, Moran's I spatial autocorrelation index, or directional mean);
- (b) [properties of distances](#) between incidents (including nearest neighbor analysis, linear nearest neighbor analysis, and Ripley's K statistic);

CrimeStat also includes routines for conducting

- (c) [basic 'hot spot' analysis](#) (including the mode, the fuzzy mode, hierarchical nearest neighbor clustering, and risk-adjusted nearest neighbor hierarchical clustering), and
- (d) [more advanced hot spot analysis](#) (including the Spatial and Temporal Analysis of Crime (STAC), K-means clustering, and Anselin's local Moran statistics).

In addition, CrimeStat includes [interpolation](#) and [journey to crime analysis](#) routines. The latter is a criminal justice method for estimating the location of a serial offender given the distribution of incidents and a model for travel distance. This routine allows the user to estimate a travel model with a calibration file and apply it to the serial events. It can also be used to identify an optimal central location given a distribution of 'points' and assumptions about travel behavior.

Lastly, CrimeStat includes a set of tools for analyzing [clustering in time and in space](#) (including the Knox and Mantel indices and the Correlated Walk Analysis module, which analyzes and predicts the behavior of a serial offender).

Data Setup in CrimeStat

The [primary](#) data file is required for CrimeStat. It is a point file with X and Y coordinates (example: a file containing the location of street robberies, each of which have an associated X and Y coordinate). You may wish to have weights and intensities for this file, for instance, perhaps you would like to weight the robbery coordinates according to how much was stolen. However, it is not necessary to use weights.

A [secondary](#) data file (optional) is also a point file with X and Y coordinates. It is usually used in comparison with the primary file. For example, if you were examining robberies with your first file, then perhaps your second file might contain the locations of banks and ATMs. It is also possible, although not necessary, to weight your secondary file as well.

A [reference file](#) is a grid that overlays the study area. Normally it is a regular grid, however it is possible to import irregular grids. CrimeStat is able to create a grid for you if given the X and Y coordinates for the lower left and upper right corners. The reference file is used in the risk-adjusted nearest hierarchical clustering routine, journey to crime estimation, and single and dual variable kernel density estimation routines. The [measurement parameters](#) define the measurement units of the coverage (area or length of the street network) and the type of distance measurement (indirect or direct) to be used. Direct Measurement means that CrimeStat will calculate the shortest distance between two points, while Indirect Measurement calculates the shortest distance between two points along the grid horizontally and vertically (no diagonal distance measurement). If no area is defined, then CrimeStat will define a rectangle by the minimum and maximum X and Y coordinates.

CrimeStat Availability

CrimeStat I and II were created by Ned Levine and Associates (<http://www.nedlevine.com/>) under an NIJ grant, freeware is available at (<http://www.icpsr.umich.edu/nacjd/crimestat.html>). CrimeStat comes with three datasets and an excellent manual that gives the background for the examples and statistics. The program and sample data sets are Windows-based zipped files for any Windows NT, Windows 2000 or Windows XP system (also on Windows 98 and 95). Depending on the size of the data files to be processed, there may be hundreds of millions of calculations on any one run, therefore it is critical that the computer be fast and have a sufficient amount of RAM. The program was designed on a Windows 2000 system with 256 MB of RAM running a dual-processor Pentium III 800 MHz computer.

CrimeStat at Penn State

CrimeStat II is available on GIA Lab machines and will be available on all of the computers in 806 Oswald (Spring 2003). The GIA resources page (www.pop.psu.edu/gia-core/resources.htm) includes two PowerPoint presentations on CrimeStat 1.0 (a) an introduction and (b) a tutorial. For other crime mapping and analysis tools see <http://www.ojp.usdoj.gov/nij/maps/tools.html>.

For further information about point pattern analysis, see GIS Resource Document 02-41 (GIS_RD-02-41).