

An Exact Test for Determining Elevated Risk of Disease within Contiguous Neighborhoods of High Toxin Concentration

D. D. Smith

October 3, 1995

Abstract

We develop an exact test that determines whether communities exposed to a plume, a drift of highly concentrated toxins, have an elevated risk of disease compared to communities unexposed to a plume. A central source producing a plume may affect health indicators of nearby neighborhoods. However, a comparison of the health indicators of the potentially affected neighborhoods with the unaffected neighborhoods must be made. Under minimal assumptions, it can be determined whether there is evidence that the plume elevates the risk of developing disease symptoms. We also discuss applications of this test to geographical information systems.

The Plume Problem

Consider a city with a central site that disperses toxins into the environment. Generally, such a site leaves a plume, a trail of heavily concentrated toxins, from the site. The toxins may be transported by air or water. Depending on environmental factors, such as weather, or chemical factors, such as density, the plume may be quite dense or quite disperse. Some examples of sites that that may produce plumes are smelters and sites that dispose of toxic waste. The "plume problem" is this: Do communities that lie within the plume's

path develop adverse health effects more frequently than communities not directly in the plume's path?

Before such technologies as geographical information systems (GIS), the solution to the plume problem required the collection and management of an enormous amount of data. GIS systems are able to handle such geographic and descriptive data conveniently. Plots of city districts may be overlaid with plumes or plume concentrations, and these may be compared with descriptive health statistics and indicators.

The Assumptions

This paper introduces a hypothesis test that will decide between H_0 : There is no indication of abnormally high rates of an adverse health effect in those areas within the path of the plume versus H_a : There is a negative association between the plume and adverse health effects that is not explained by chance alone. In order to test this hypothesis, we need the following assumptions:

- The community-at-large (e.g. the city under study) is divided into the equivalent of census tracts, in which all tracts have approximately the same number of individuals.
- Other than the influence of the plume, the rates of disease within each census tract is independent of all other census tracts.
- For the adverse health effect in question, there is a pre-established threshold that indicates if a census tract is at risk.

The Test

Let N equal the number of census tracts for the community-at-large. Then each census tract has a rate for the health indicator. Let X equal the number of census tracts that are at risk with respect to the established threshold of the health indicator. $P = \frac{X}{N}$ = the population proportion of census tracts at risk within the community-at-large. Let n equal the number of census tracts within the plume's influence. Finally, let x equal the number of tracts at risk and also within the plume's influence.

If the above assumptions hold, then x is a binomial random variable defined on $0, \dots, n$ and with success probability P . The p-value for this test

is $P(\mathcal{X} \geq x)$ where \mathcal{X} is a random variable distributed $\text{Bin}(n, P)$. Hence, the rejection region of this test is X^* where $P(\mathcal{X} \geq X^*) \leq \alpha$ and α equals the significance level.

An Example

Consider a city with $N = 18$ census tracts. We want to see if the plume, emanating from near the center of the city, has an effect on birth weight. Each census tract has an average birthweight, and 5 out of the 18 census tracts have average birth weights that are “low” according to a pre-established threshold of birth weights. The proportion of census tracts with low birth weights is $P = \frac{5}{18} = .2778$. The dark cloud in Figure 1 represents a plume. By overlaying the plume onto the census tracts, it appears that $n = 5$ tracts (the shaded ones) may be potentially affected by the plume. Out of the 5 tracts, 4 of them have low birth weights. The p-value that corresponds to the test of hypothesis outlined above is

$$\begin{aligned} P(\mathcal{X} \geq 4) &= \binom{5}{4} P^4(1 - P)^1 + \binom{5}{5} P^5(1 - P)^0 \\ &= 0.02315 \end{aligned}$$

At the $\alpha = 0.05$ level, we reject H_0 : and conclude that there is an association between low birth weights within the census tracts and the plume’s path through the city.

Discussion

This test has several advantages. It is exact, in that if the assumptions are met, it doesn’t rely on asymptotic or limiting results. It is also readily portable into the GIS context because every bit of information needed for the test may be extracted from current GIS packages. Information on health indicators within census tracts is one of the specialties of GIS and the overlay features of most GIS packages allow the researcher to count the number of tracts under a plume.

There are several disadvantages with this hypothesis test, however. For large areas, it is unreasonable to assume independence of tracts with respect to developing disease. Diseases usually develop in clusters and risk factors

such as income and age are not generally homogeneous from census tract to census tract. This method cannot control for such factors.

A decision that an investigator must make while using the technique is how to define which census tracts are potentially affected by the plume. If the plume only covers a small portion of a census tract, the researcher must decide whether the plume can affect residents' health substantially within the tract. There is some subjectivity inherent in deciding this, but such factors as plume concentration and population concentration may help to make a reasonable decision.

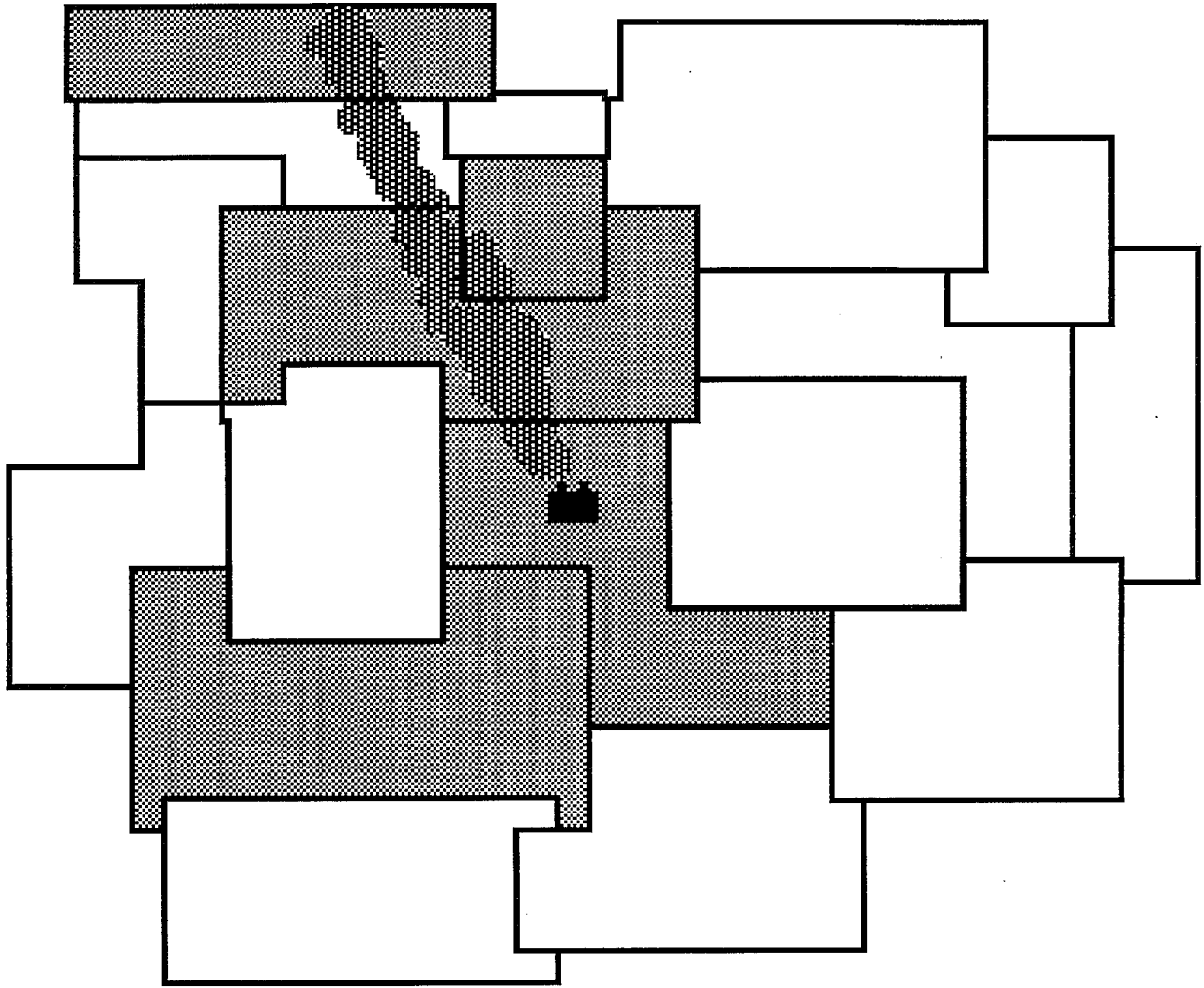


Figure 1: An example of a city with 18 census tracts and a plume originating near the center. The shaded regions indicate areas with low average birth weights.