ESTIMATION FROM CAPTURE ALONE: HOW TO ESTIMATE THE TOTAL NUMBER OF ILLEGAL IMMIGRANTS

Dhirendra N. Ghosh, SAIC

Each year, hundreds of thousands of foreign citizens enter the United States. A large number of these people enter legally to become permanent, legal residents and often, eventually naturalized citizens. Along with this legal immigration, each year an equally large and steadily growing volume of people are illegally entering this country. The U.S. Immigration and Naturalization Service (INS) is the federal agency responsible for overseeing the legal immigration of foreign naturalists to this country. INS has approximately 200 offices divided into 4 regions across the U.S. and Puerto Rico.

There are thirty-five investigative services district offices and twenty Border Patrol sector offices. The service operates seven processing centers where apprehended immigrants may be detained, in addition to using three contract facilities and more than 900 state and local jails for detention. Firm figures exist on the number of illegal aliens apprehended by the INS in a given year, About ten percent of all illegal immigrants apprehended had resided in U.S. for more than six months.

The objective is to estimate the total number of illegal immigrants living in U.S. Only subjective estimates are available for the total size of the resident illegal alien population.

The Capture recapture methodology is well known for estimating the total size of a population. However, in this case we have a devise a methodology for estimating the total size from capture alone.

It is possible to interview a stratified random sample of the apprehended illegal immigrants before deportation and determine their year of arrival to U.S. Assuming this information is available, it is possible to estimate the total size of illegal immigrants living in U.S. using a statistical model.

Let us consider the simplest model first. It assumes that a fixed number (say N) of illegal immigrants enter U.S. every year. It further assumes that a fixed proportion (say P) of these illegal immigrants are apprehended every year. (We do not consider illegal immigrants apprehended within a short time of their arrivals). For example: of the N illegal immigrants who entered U.S. last year, NP will be apprehended this year. Now take the year before last, NP of them were apprehended last year and deported. Hence N NP of them were left this year, and so P(N - NP) of them are apprehended this year.

Now if n_1 is the number of immigrants apprehended this year who admit to have entered the U.S. last year and n_2 is the corresponding number for the year before last

then,

$$\frac{NP(1-P)}{NP} = \frac{n_2}{n_1}$$

or, $$1 - P = \frac{n_2}{n_1}$$

Thus, P can be statistically estimated from data obtained for immigrants entering U.S. in two consecutive years. N can be estimated then.

Obviously, if such a survey is conducted, it will include data for different year of arrival to U.S.

The estimate can then be improved by combining estimates obtained from two successive years. The total number of immigrants apprehended for the i_{th} year (counting backwards with the last year as year 1) is

$$NP(1 - P)^{i-1}.$$

The above model can be generalized in various ways:

1) The assumption that a fixed proportion is apprehended every year can be slackened. We may assume that the proportion apprehended is C_iP for the i^{th} where C_i is a known constant. This is realistic since it may be assumed that the longer an immigrant lives in U.S., the less is his/her chance of being apprehended.

2) The assumption that the number of illegal immigrants entering the U.S. every year is a constant, can also be slackened. We can assume that the number in the i^{th} year is K_iN where K_i is a known constant.

3) We may also introduce a mortality factor (due to death and legalization of stay) in the total population of illegal immigrants.

Obviously, this methodology can be applied to estimate the total size of any population subject to periodic capturing, when the age of a unit can be determined at the time of capture. For example, if we want to estimate the total number of fish in a tank, we could catch some number of fish at two points in time and determine the age of each fish. We could then estimate the total number of fish in the tank.