SINGLE EARTHQUAKE LOSS PROBABILITIES

by

Betsy Schumacker[I] and Robert V. Whitman[II]

SYNOPSIS

A computer program gives the probability that n buildings in a region will fail during a single earthquake. As input it is necessary to specify the usual seismic risk parameters (source zones, recurrence rates, attenuation laws) plus failure probability as a function of intensity of failure for each type of building and soil condition. Z-transforms are used to combine the several binomial probability distribution functions. An example illustrates the application of the program and the effect of varying the parameters.

INTRODUCTION

When losses from earthquakes are averaged over the years, such losses appear modest - at least in the United States - compared to those caused by other natural hazards. However, earthquakes exceed other hazards in their capacity to cause losses of catastrophic proportions in a few moments of intense shaking that occurs without warning. Hence, in considering measures to mitigate the earthquake haard, it is desirable to estimate the probability that different thresholds of loss might be exceeded, in a city or a region, during any one single earthquake. Since any one earthquake may cause different intensities of shaking within the city or region and different types of buildings will respond differently to a given intensity of shaking, it is necessary to combine together several different sets of probabilities. This paper describes a method of analysis developed for this purpose, with focus on failures of buildings as an indirect measure of the potential for fatalities.

THE MODEL

The model is logically divided into the following components:

- Definition of building types and a failure probability vector for each type. The building types are based upon the quality of construction of the structure and the type of soil on which it is built. Each type thus indicates a general soil/structure response pattern. The failure probability vector is based on the type of effect one wishes to measure and is a 2-state measuring device (success or failure). Failure could be total collapse, moderate damage or greater, a certain level of repair cost, or a measure of fatalities such as 50% or more of occupants killed. The failure probability vectors give the probability of a single structure failure for a set of MMI or ranges of ground accelerations.

- Definition of the geometry of the region and the step size for spatial discretization, The entire region - seismic source zones and the target areas - is discretized into an equal-sized grid pattern

I Lecturer, Massachusetts Institute of Technology, Cambridge, Mass.
II Professor of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Mass.

- Definition of the target areas. The target areas can be spatially distributed and need to have their geometry defined and for each target cell the number of buildings of each type (construction and soil) which are in that cell must be given.

- Definition of the earthquake models. This consists of defining the geometry of all source zones to be considered, the frequency of occurrence and attenuation models, and the boundary conditions on epicentral magnitude for each zone, and the step size for discretization of epicentral magnitude.

- Computation of the probability mass function and cumulative distribution function of building failure over the target areas. The mass function gives the annual probability of n out of N buildings failing, where N is the total number of buildings in the target areas and n = 0, 1,...N. The cumulative gives the probability of n or more failing.

PROBABILITY COMPUTATION

Each building is assumed to be statistically independent of all other buildings; thus, since we are dealing with a 2-state occurrence of each building (failure or non-failure), our basic probability distribution is a binomial. (For multi-state occurrences, such as 5 levels of building damage, a multinomial distribution would be used.) For each event generated, a distribution function for n of N buildings failing is computed. It is multiplied by the probability of the event occurring (at an epicentral location, of a magnitude interval) and these are summed over all magnitude intervals in the magnitude range of interest and over all epicentral locations (all cells in the source zones). The computation of the distribution function for n of N buildings failing, given an event, will now be described.

Let $p_n(n_o)$ be the probability of n_o buildings failing given an event (at a given location, of a given magnitude, with a given attenuation rate). If we have only one target cell and only one kind of building in that cell (i.e., the probability of failure of a single building is the same for all buildings), then we have

$$p_n(n_o) = \binom{N}{n_o} p^{n_o} (1-p)^{N-n_o}$$

which is the standard binomial distribution function for n_o of N buildings failing, given the probability of failure for one building = p. However, when working with area (or distibuted) targets, one does not have this kind of situation. Rather, one has more than one kind of building in a cell as well as having different levels of ground shaking at different target cells. Hence, one has a total population with subgroups each of which has a different failure probability caused by being of different kinds, or receiving different excitation levels, or both. Thus, for any event, the population of buildings will be a non-homogeneous set in that the probability of failure varies from building to building. The distribution remains binomial, however, in the following fashion.

Since the failure curve (probability of failure vs. level of ground shaking) for each kind of building is discrete (by breaking the ground shaking into intensity of acceleration ranges), the total population (given an event) can be grouped into k homogeneous sets, each having its own probability of failure and each binomially distributed. Thus, since the buildings are assumed independent of one-another, the probability of n out of N failing can be written

$$p_n(n_o) = p_n(n_{o_1}, n_{o_2}, \text{--}, n_{o_k})$$

and is itself a random variable. The i^{th} random variable of this set, n_{oi}, would have a binomial mass function $p_{ni}(n_{oi})$ where $n_{oi} = 0, 1, \text{--}, N_i$. The sum of all $N_i = N (i = 1, 2, \text{--}, k)$ the total number of buildings in all target cells.

The combining of these k distributions can be performed by convolution or by discrete transform using the z-transform. The z-transform of binomial mass function $p_{ni}(n_{oi})$ is $p_{ni}^T(z_i)$ and is given by

$$p_{n_i}^T(z_i) \equiv E(z_i^{n_i}) = \sum_{n_i=0}^{N_i} z_i^{n_i}\, p_{n_i}(n_{o_i})$$

where $E(z_i^{n_i})$ is the expected value of the transform variable z_i for distribution on n_i. The coefficients of this polynomial, namely $p_{n_i}(n_{oi})$, are the probabilities of exactly n_{oi} out of N_i failing in the i^{th} group above, i.e. they are values of the function for group i. Since our desired distribution is based on the sum of random variables, we apply transform theory again to compute our random variable sum as the product of the transforms of the individual variables. In other words, if

$$n_o = n_{o1} + n_{o2} + \text{---}\ n_{ok}$$

is a sum of random variables, then its discrete transform is

$$p_{n_o}^T(z) = p_{n_{o1}}^T(z)\, p_{n_{o2}}^T(z) \cdots p_{n_{ok}}^T(z)$$

$$= \sum_{n_{o1}=0}^{N_1} p_{n_1}(n_{o1})\, z^{n_{o1}} \sum_{n_{o2}=0}^{N_2} p_{n_2}(n_{o2})\, z^{n_{o2}} \cdots \sum_{n_{ok}=0}^{N_k} p_{n_k}(n_{ok})\, z^{n_{ok}}$$

The coefficient of z^{n_o} above is equal to our wanted probability, i.e. is equal to $p_n(n_o)$ or to exactly n_o failing out of the total N.

This coefficient of z^{n_o} can be computed as one of the terms in the products of the polynomials from the transforms of the probability mass functions of the k groups.

Thus the general procedure is to (given an event and an M_o)

a) Determine the number of homogeneous groups
b) Compute the probability mass function for each group.
c) Compute the product of the transforms of the k mass functions as the polynomial products of the k mass functions.
d) The coefficients just computed are the desired probability $p_n(n_o)$, $n_o = o, 1, \text{---}, N$.

AN EXAMPLE

Eastern Massachusetts was chosen as an example, and 6 cities in that area were defined as the torget areas of interest. The whole region (including the 4 source zones) was discretized into 5 square cells. At this level, 5 of the cities were defined as 1 target cell each, and 1 was defined as 4 target cells. Fig. 1 shows the area of interest with source zones and target areas. One construction type (brick) and 2 soil types (good and bad) were chosen and the number of buildings of each of the 2 types (brick on good soil, brick on bad soil) in each target cell was determined. The numbers of buildings used in the computations were the actual number divided by 20. Table 1 shows the numbers used for each type for each target cell. Type 1 corresponds to good soil, type 2 to bad soil.

Failure probability vectors were specified for each type. Two classes of computations were made. The first class had a failure defined as having on the average 1% or more of the buildings occupants killed by an event, the second class had a failure defined as collapse, with a large fraction of the occupants killed. The failure vectors used are shown in Table 2.

Figures 2 and 3 show the computed probability functions for exactly n buildings failing - Fig. 2 for class 1, Fig. 3 for class 2. Fig. 2 is truncated at n = 700. Fig. 2 shows the general type of curve which results from this kind of computation. Since the basic distributions are binomial, the final distribution will have combinations of binomials, evidenced by the cusps in the curve. If the total number of buildings were large enough, each binomial-like cusp would be a separate and distinct binomially-shaped curve. In the one shown in Fig. 2, each cusp corresponds to the expected value of the binomials from one event.

If the number of buildings is small with respect to the difference in value of the failure probabilities for the MMI's, then all the cusps are not distinct since one will overlap near the high point of the other. This has happened in the curve in Fig. 3.

EFFECT OF CHANGES TO PARAMETERS

Changes to various parameters in the model can have different effects on the final distribution. These are summarized here.

- Small variations in the rate of occurrence cause virtually no change in the final distribution.

- Small variations ($\pm$ 10%) in the total population has little effect on the final distribution.

- An increase in the total population causes more cusps to come into view, a lowering of the probability of the expected value and an increase in the spread of each binomial curve, an increase in the mode point (expected value). The shifting to the right of each binomial is a cumulative shift to the right.

- Variations in the values in the failure probability vectors cause major changes in the final distribution. This is caused by the change in shape of the binomial when the value for p changes. As p goes from .1 to .5 to .9 the binomial goes from a right-tailed Poisson to a normal to a left-tailed Poisson. The mode point also changes greatly, since mean = pN.

BIBLIOGRAPHY

Schumacker, B. "Models for Threshold Exceedence and Loss Computation of Non-Homogeneous, Spatially Distributed Facilities," Department of Civil Engineering Research Report R76-31, M.I.T. Cambridge, Massachusetts.

CITY	TYPE 1	TYPE 2
1	470	70
2	500	100
3	500	100
4	470	70
5	500	120
6a	500	200
6b	150	100
6c	600	370
6d	500	180

TABLE 1: Number of Buildings of Each Type in Each Target Cell

CLASS	TYPE	V	VI	VII	VIII
1	1	.0045	.0545	.2	.425
	2	.0092	.1565	.45	.8
2	1	0	0	.004	.0115
	2	0	0	.0045	.0345

TABLE 2: Failure Probability Vectors for the Two Classes of Computations

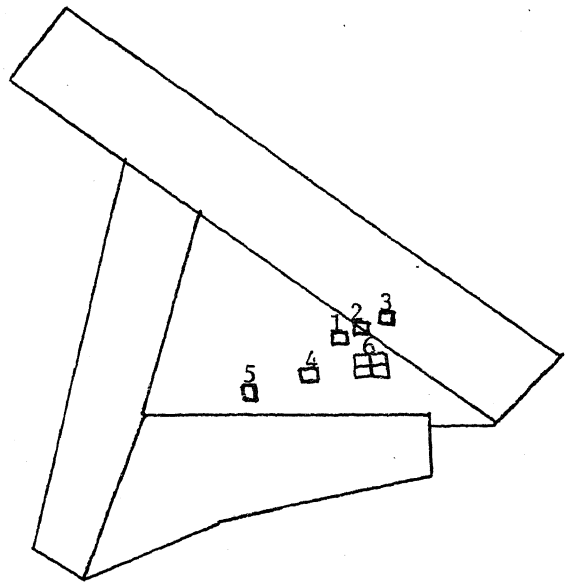

FIGURE 1 : AREA OF INTEREST

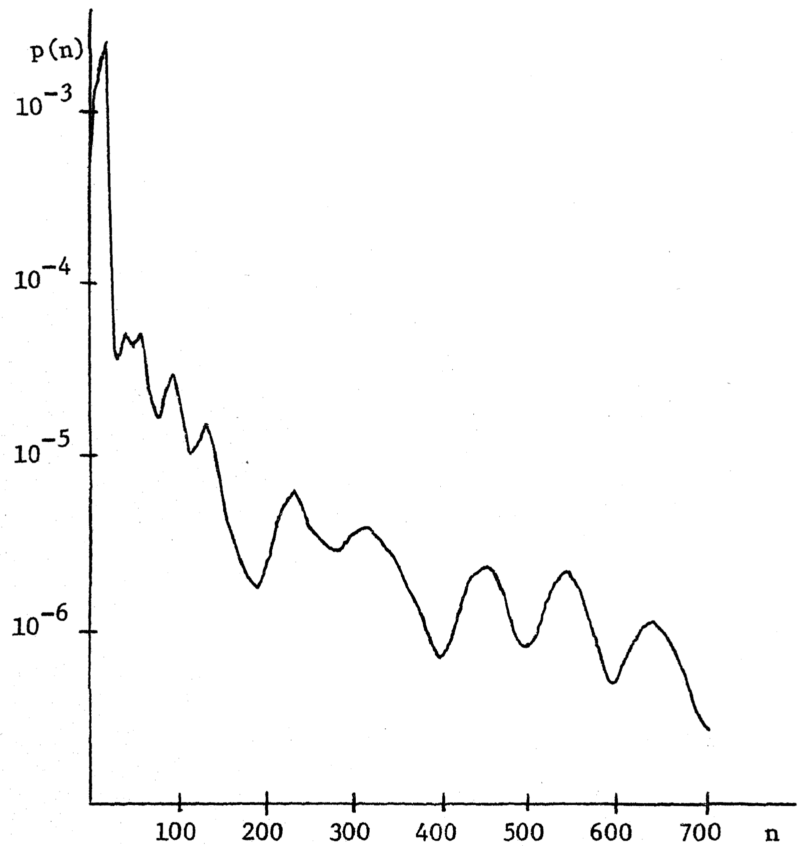

FIGURE 2: CLASS 1 DISTRIBUTION

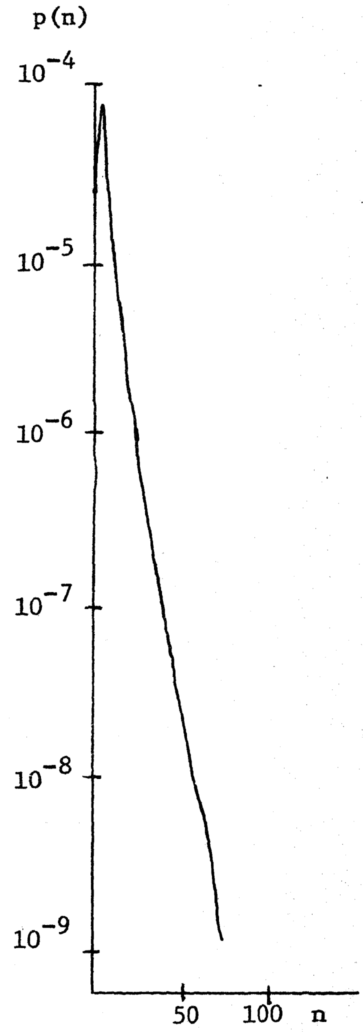

FIGURE 3:
CLASS 2
DISTRIBUTION