About the statistical validation of probability generators

G. GRANDORI[(1)], E. GUAGENTI[(1)] and L. PETRINI[(2)]

[(1)] *Dipartimento di Ingegneria Strutturale, Politecnico di Milano, Italy*
[(2)] *Dipartimento di Meccanica Strutturale, Università di Pavia, Italy*

(Received April 22, 2003; accepted July 22, 2003)

Abstract - The procedures that generate probability statements are called probability generators (or simply generators). In this paper, we consider the generators that yield the magnitude distribution function F_M in the frame of the probabilistic seismic hazard analysis at a site. From an engineering point of view, the behaviour of F_M in the range of strong earthquakes is of decisive importance. In general, however, the statistical validation of a generator in the range of interest is not feasible because of the limited number of occurrences of strong earthquakes. In spite of this, in this paper we show that a simple empirical generator of F_M can be statistically validated from an engineering point of view thanks to: 1) a new approach to the comparison between competing generators, and 2) the comparison, following this new approach, of the empirical generator with a particular class of generators based on mathematical models.

1. Introduction

Following Lind (1996), we call "probability generators" or simply "generators" those procedures (algorithms, models, methods, etc.) that generate probability statements. In particular, a generator may yield a probability distribution function (which is equivalent to an infinity of probability statements) or simply an elementary ("point") probability statement.

When dealing with seismic hazard at a site, the generator of the magnitude distribution F_M $(m;\ \vartheta)$ of the earthquakes that can significantly affect the site, plays an important role. From now on, a distribution F_M with known vector of parameters ϑ will be briefly indicated with an upper index (e.g. F^i); on the other hand, a distribution that contains free parameters will be called a "model" and indicated with a lower index (e.g. F_i). A generator must yield a completely defined F^i.

Corresponding author: G. Grandori, Dip. di Ingegneria Strutturale, Politecnico di Milano, p.za L. da Vinci 32, 20133 Milano, Italy; phone: +39 0223994336, +39 0223994207, fax: +39 0226688998; e-mail: petrini@unipv.it

Grandori (1991) and Lind (1996) discuss the difficulties that interfere with obtaining the statistical validation of a probability generator in detail. These difficulties may become insurmountable in the case of some practical applications. Consider, for instance, the evaluation of the peak ground acceleration (PGA) with a 500-year return period at a given site. Call a this quantity, which is a synthetic measure of local hazard, and, is generally suggested by seismic codes as an appropriate value for the design of ordinary buildings. The quantity a is dominated by the magnitude distribution of strong earthquakes, while the statistical data on these events are scarce. It follows that the available data are generally not sufficient for the statistical validation of the generator of a magnitude distribution F^i, as well as for a significant comparison between different generators, even if they lead (*coeteris paribus*) to important differences in the evaluation of a (see for instance Kagan, 1993; Wu et al., 1995).

The condition *coeteris paribus* means that, in the comparison between generators of magnitude distribution, all other elements that contribute to the evaluation of a at the considered site (like the distribution of the earthquakes in space and time, and the attenuation law) are supposed to be known and independent of the type of generator. As a consequence, if a magnitude distribution function F^i is defined (both form and parameters), then a known procedure Z, applied to F^i, gives the value of a at the site:

$$a^i = Z\,(F^i). \tag{1}$$

When the statistical validation of plausible generators of magnitude distribution is not feasible because of a limited number of occurrences of strong earthquakes, Kagan (1993) suggests that "in practical evaluations of seismic risk a range of results should be presented to display possible alternatives. These calculations should be performed using various models of size distribution with indication of possible variations of parameter estimates". In other words, a scenario of both epistemic and aleatory uncertainties should be displayed.

As far as the comparison between competing generators is concerned, it has been observed (Grandori et al. 1997, 1998) that to shift the attention from the fitting of catalogue data to the expected error in the estimate of a opens new statistical prospects.

Precisely, call F^o the true magnitude distribution, a^o the true value of a, and S_o a random sample drawn from F^o (the data of the catalogue constitute an S_o). The generator G_i of a distribution F^i leads, through Eq. (1), to an estimate $\hat{a}_i$ of the quantity a. We define the "credibility" of the generator G_i the probability Δ_i^o that, starting from the information contained in one random sample S_o, the generator G_i leads to estimate a with an error smaller than a given limit; say, for instance, less than 20%:

$$\Delta_i^o = \mathrm{P}\,\{[a^o - \hat{a}_i \,|\leq 0.2\, a^o\}. \tag{2}$$

Note that, in the case of two competing generators G_γ and G_s, the fact $\Delta_\gamma^o > \Delta_s^o$ does not mean that, compared with G_s, the generator G_γ leads to a magnitude distribution in some way "closer" to F^o. It means that, compared with G_s, the generator G_γ leads to an estimate of a closer to a^o in probabilistic terms.

It is appropriate, here, to lay the stress on the limits of the above-mentioned credibility. In fact, Δ_i^o refers to a specific quantity at a given site; it is far from being a characteristic of the generator G_i as such. Moreover, and above all, can the credibility Δ_i^o be helpful given that the real distribution F^o is not known? The answer is: yes, at least in some problems.

In Grandori et al. (2003b) the comparison between two generators G_1 and G_2 is considered, with reference to an Italian site, given the catalogue of the earthquakes in the zone during the last 300 years. Two hypotheses are on the table:

$$a)\ \Delta_1^o > \Delta_2^o \quad \text{and} \quad b)\ \Delta_2^o > \Delta_1^o \tag{3}$$

In the above paper, it is shown that the application of Eq. (2) with many different conjectural F^o can lead to the evaluation of the relative likelihood L_a/L_b of the two hypotheses a) and b).

In the present paper, the comparison between generators, with Eq. (2), is carried out following a particular technique; the result is, in practice, the statistical validation of a simple generator applied to the hazard analysis at the same Italian site previously considered. We do not claim that the simple generator and the technique leading to its statistical validation would be effective for any other site. We will merely describe the numerical experiments that show the effectiveness of the procedure “at least” in the considered case.

A straightforward presentation of the first pilot - experiment is the description of a simulated competition between experts who support different generators.

2. The competition between experts

The competition ground is shown in Fig. 1. The goal is the evaluation of the quantity *a* at the Alfa site.

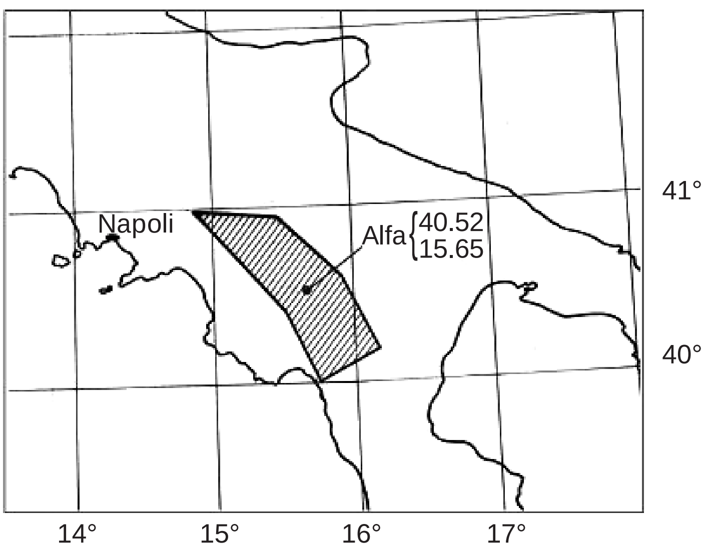

Fig. 1 - The competition ground.

The distribution of the earthquakes in space and time, the attenuation law, the mean annual number of events n and the number ν of events contained in the catalogue are known. The competitors are supplied with all these elements, so that only the magnitude distribution is needed in order to evaluate the quantity a with Eq. (1).

Actually, the Jury knows, too, the true magnitude distribution F^o, both form F_o and parameters b and m_1 (see Fig. 2) for $M > 4$; smaller magnitudes are neglected. So the Jury can: 1) calculate the true value a^o, and 2) draw from the true magnitude distribution as many size-ν random samples S_o as wanted. Each S_o is one of the possible real catalogues given that the real distribution is F^o.

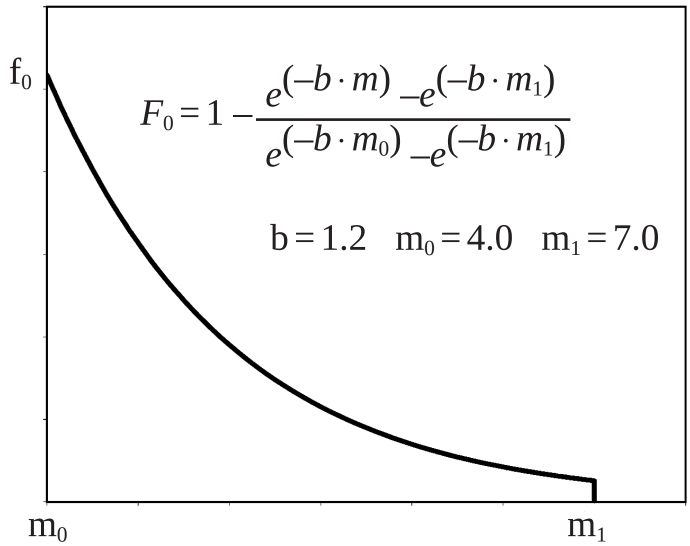

Fig. 2 - The true magnitude distribution.

However, the competitors are not supplied with the form F_o or with the parameters b and m_1; they are requested to answer the following question: "How would you estimate the quantity a on the basis of the information contained in just one of the random samples S_o?"

The first competitor is a famous champion, who carried off the prize in many similar competitions. This is due to the fact that, thanks to a special supernatural talent, he "divines" the real form F_o, which can be classified as the "right" model.

Then, the answer of the first competitor is: "From the random sample S_o I estimate, by maximum likelihood, the parameters of the mathematical model F_o. So, I obtain a completely defined magnitude distribution $\hat{F}$ and, from Eq. (1), the requested estimate $\hat{a}_o$ of the quantity a". Note that the generator adopted by the first competitor is composed of the right model and of the method used for the estimate of the parameters; we will call it the "right model generator".

The Jury repeats the application of this procedure with one thousand random samples S_o, thus obtaining the relative frequency of the event $\{ |a^o - \hat{a}_o| \leq 0.2\, a^o\}$; i.e. (with good approximation) the credibility Δ_o^o of the right model generator. The numerical result is:

$$\Delta_o^o = P\,\{ |a^o - \hat{a}_o| \leq 0.2\, a^o\} = 0.63. \tag{4}$$

A second competitor does not have special talents; moreover, he does not like mathematical models. His naïve answer is: "From the sample S_o I derive the cumulative frequency polygon, which provides a completely defined empirical distribution F^*; from this distribution I derive an estimate a_*, which is my estimate of the quantity a".

By repeating this procedure with one thousand random samples, the Jury obtains the credibility of this "polygon generator":

$$\Delta^o_* = P\{\,|a^o - \hat{a}_*| \leq 0.2\, a^o\} = 0.67. \qquad (5)$$

The results of the competition show that, unexpectedly, the credibility Δ^o_* of the empirical generator is of the same order as the credibility Δ^o_o of the generator based on the "right" mathematical model; i.e. the two procedures are affected by the same aleatory uncertainty.

However, the procedure based on a mathematical model (in the absence of special supernatural talents) is also affected by an epistemic uncertainty which is difficult to control.

Anyhow, for a few cents, the second competitor takes off with the prize.

Now a question arises: is the result $\Delta^o_* \cong \Delta^o_o$ just an accident? In order to find an answer, we repeated the comparison between the two generators with different conjectural "true" values of the parameters b, m_1. The comparisons of Table 1 (columns 1 and 2) show that, for the considered site, and if the true F^o is an exponential truncated distribution, the result $\Delta^o_* \cong \Delta^o_o$ is systematic, with small fluctuations in the range of plausible couples of parameters.

Table 1 - Truncated exponential distribution.

			1	**2**	**3**	**4**
m_1	b	a^o	Δ^o_o	Δ^o_*	φ_*	Δ^o_φ
7.4	0.8	.41	.74	.64	1.26	.77
	1.0	.36	.63	.54	1.25	.65
	1.2	.31	.53	.49	1.23	.55
7.0	0.8	.33	.73	.72	1.24	.80
	1.0	.30	.68	.68	1.23	.77
	1.2	.27	.63	.67	1.23	.74
6.6	0.8	.27	.85	.88	1.23	.89
	1.0	.25	.78	.86	1.22	.88
	1.2	.23	.70	.80	1.22	.86
		average	.70	.70	1.23	.77

3. The modified polygon generator

A second question is: can the polygon generator be improved? The answer is: yes, in more than one way.

A first way is described in Grandori et al. (2003a).

A simpler way of increasing the credibility of the polygon generator is based on the evaluation, for each sample, of the 0.9 percentile $\xi^*_{0.9}$ of its empirical distribution F^*. A numerical analysis shows that, for all the couples of parameters of Table 1, the mean value (over one thousand samples) of the adimensional percentile

$$\varphi_* = \xi^*_{0.9} \,/\, \mu_* \tag{6}$$

(μ_* is the mean of the sample) is rather stable and close to 1.23.

Generally speaking, a sample with a high value of φ_* leads to overestimating the quantity a and viceversa. Then, we assumed a modified estimate $\hat{a}^!_*$ instead of $\hat{a}_*$ in the following cases:

$$\hat{a}^!_* = 0.8\ \hat{a}_* \quad \text{if } \varphi_* > 1.28$$
$$\hat{a}^!_* = 1.2\ \hat{a}_* \quad \text{if } \varphi_* < 1.18.$$

The credibility Δ^o_φ of this modified generator is always higher than both Δ^o_o and Δ^o_* (Table 1, column 4).

A first conclusion. Suppose that for the considered site a traditional best-fit analysis led to the choice of the truncated exponential model for magnitude distribution. To adopt the modified polygon generator, as an alternative, would lead to a higher credibility even if the truth were exactly an exponential truncated distribution.

4. About the statistical validation of the modified polygon generator

Suppose now that, following Kagan's (1993) suggestion quoted in the introduction, the decision is to display the results that would be obtained by adopting, besides the exponential distribution, two further conjectural "true" distributions as possible alternatives; precisely:

1) a double exponential distribution

$$F^o = 1 - \exp\left[\exp \beta\,(m_o - u) - \exp \beta\,(m - u)\right]; \tag{7}$$

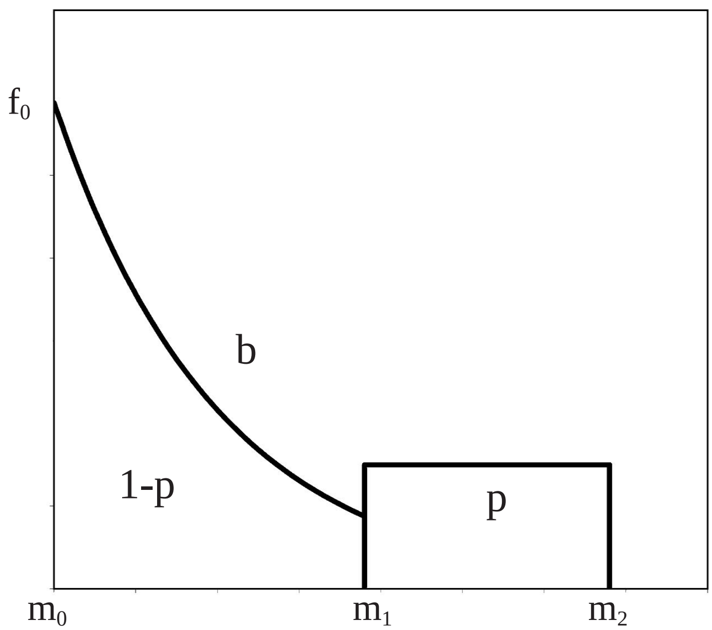

Fig. 3 - The characteristic type magnitude distribution.

2) a characteristic type distribution as shown in Fig. 3, with four parameters (the relative frequency *p* of characteristic magnitudes, the *b* value of the exponential part of the distribution, m_1 and m_2).

Tables 2 and 3 show, for each case, the comparison between the modified polygon generator and the right model generator. Note that in the last case the right model, besides the right mathematical form, has the right parameters *b* and m_1; i.e. the values of Δ_o^o are actually optimistic.

Again, on the average Δ_*^o is of the same order as Δ_o^o, while Δ_φ^o is larger than Δ_o^o.

Table 2 - Double exponential distribution.

β	u	a^o	Δ_o^o	Δ_*^o	φ_*	Δ_φ^o
0.32	0	.23	.64	.69	1.21	.74
	0.8	.32	.56	.48	1.22	.51
	1.6	.48	.48	.39	1.27	.46
0.30	– 0.8	.21	.65	.71	1.17	.81
	0	.29	.57	.50	1.22	.53
	0.8	.42	.51	.40	1.25	.46
0.28	– 1.6	.20	.66	.70	1.19	.81
	– 0.8	.27	.57	.55	1.22	.58
	0	.38	.49	.39	1.24	.42
		average	.57	.53	1.22	.59

Table 3 - Characteristic type distribution b = 1, p = 0.1 “right model” with true *b* and m_1.

m_1	m_2	a^o	Δ_o^o	Δ_*^o	φ_*	Δ_φ^o
5.6	6.6	.24	.68	.83	1.22	.88
	7.0	.30	.65	.63	1.23	.69
	7.4	.37	.51	.50	1.24	.57
5.9	6.6	.27	.78	.86	1.23	.88
	7.0	.31	.72	.68	1.23	.75
	7.4	.39	.56	.53	1.24	.61
6.2	6.6	.30	.80	.90	1.24	.90
	7.0	.35	.81	.79	1.25	.80
	7.4	.42	.61	.64	1.26	.68
		average	.68	.71	1.24	.75

At this point, we would venture a second more general conclusion. If, for the magnitude distributions that are considered plausible for the site, and for a wide fan of plausible parameters, the credibility of the modified polygon generator is always of the same order, or even larger, than the one of the right model generators (as in the case of our Italian site), then the modified polygon generator can be considered, in practice, as statistically validated.

References

Grandori G.; 1991: *Paradigms and falsification in Earthquake Engineering.* Meccanica, **26**, 17-21.

Grandori G., Guagenti E. and Tagliani A.; 1997: *Una discussione delle incertezze nella valutazione della pericolosità sismica.* In: Atti dell'8° Convegno Nazionale ANIDIS, Ed. Priulla, Palermo, pp. 19-26.

Grandori G., Guagenti E. and Tagliani A.; 1998: *A proposal for comparing the reliabilities of alternative seismic hazard models.* Journal of Seismology, **2**, 27-35.

Grandori G., Guagenti E. and Petrini L.; 2003a: *Modeling magnitude distribution for local hazard evaluation: a new approach.* In: Proc. of Pacific Conference on Earthquake Engineering, New Zealand Society for Earthquake Engineering, Christchurch New Zealand, CD-Rom Paper 025.

Grandori G., Guagenti E. and Tagliani A.; 2003b: *Magnitude distribution versus local seismic hazard.* Bull. Seism. Soc. Am., **3**, 1091-1098.

Kagan Y.Y.; 1993: *Statistics of characteristic earthquakes*. Bull Seism. Soc. Am., **83**, 7-24.

Lind N.C.; 1996: *Validation of probabilistic models*. Civil Eng. Syst., **13**, 175-183.

Wu S.-C., Cornell C.A. and Winterstein S.T.; 1995: *A hybrid recurrence model and its implication on seismic hazard results*. Bull. Seism. Soc. Am., **85**, 1-16.