

Hazard Curves and Fragility Functions

Consider Figure 1, which shows the concept of hazard curves on the left, and a fragility function on the right. IM is a random variable that represents some site-specific ground motion intensity measure, with outcomes denoted by im . The failure event is denoted by F . The hazard curve is shown in two different versions. At the top of Figure 1a, it is given as a complementary cumulative distribution function, CCDF, denoted $G(im)$. Using that function, a classical application of the theorem of total probability is the calculation of the *probability* of failure as

$$P(F) = \int_0^{\infty} P(F|im) \cdot f(im) \cdot dim = \int_0^{\infty} P(F|im) \cdot \left| \frac{dG(im)}{dim} \right| \cdot dim \quad (1)$$

where $f(im)$ is the probability density function. The following abbreviated version is helpful for numerical evaluation of Eq. (1):

$$P(F) = \int_0^{\infty} P(F|im) \cdot |dG(im)| = \sum_{im} P(F|im) \cdot |dG(im)| \quad (2)$$

where the last equality implies that the im -axis has been divided into intervals. Midpoint-integration implies that $P(F|im)$ =fragility value at the midpoint of each interval and $dG(im)$ =change in G in that interval.

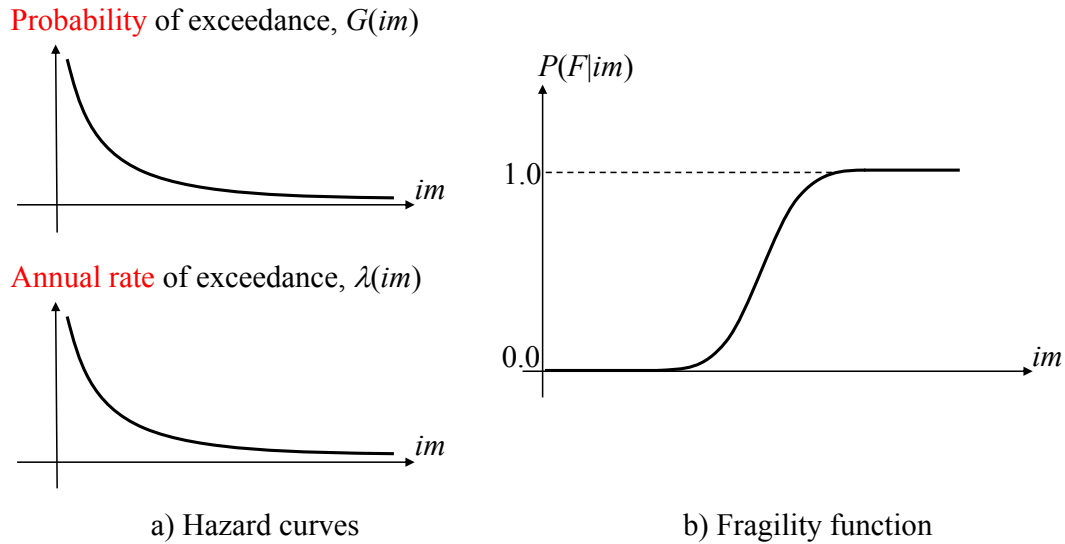


Figure 1: Hazard curves and fragility functions.

At the bottom of Figure 1a, the hazard curve is given as an annual *rate* of exceedance of the ground shaking intensity im . In that case, the theorem of total probability gives the annual *rate* of failures as

$$\lambda(F) = \int_0^{\infty} P(F | im) \cdot \left| \frac{d\lambda(im)}{dim} \right| \cdot dim = \sum_{im} P(F | im) \cdot |d\lambda(im)| \quad (3)$$

where $d\lambda(im)$ =change in λ in each interval along the im -axis.

Failure Rate with Multiple Sources

Above, only one source of earthquakes is considered. Now consider a situation with multiple sources. For example, in the Vancouver region it is common to consider three sources: Crustal, Subcrustal, and Subduction earthquakes. Let S be a discrete random variable with outcomes $s=1,2,3,\dots,K$, where K =number of considered earthquake sources. Furthermore, assume that each source is associated with a unique hazard curve and a unique fragility function. Importantly, in this section also assume that the hazard curve is given in terms of the annual rate of exceedance. That means the given fragility functions are $P(F|im,s)$ and the given hazard curves as $\lambda(im,s)$. In this case, Eq. (3) gives the *number* of failures, per time, from each source. As a result, summation yields the sought failure rate:

$$\lambda(F) = \sum_{s=1}^K \lambda(F | s) = \sum_{s=1}^K \left(\sum_{im} P(F | im, s) \cdot |d\lambda(im, s)| \right) \quad (4)$$

Failure Probability with Multiple Sources

A variation on the previous section is addressed here, now seeking the failure *probability*. That means we seek the probability of failure that is akin to the *rate* of failure in Eq. (4). While *rates* could be added over earthquake sources to obtain Eq. (4), probabilities cannot be added in the same manner.

Part I

As a first case, it is assumed that the intensity value, im , is known *a priori*. In this case the failure probability associated with each source is obtained without integration over im :

$$P(F | s) = P(F | IM = im, s) \quad (5)$$

where $P(F|IM=im,s)$ is the failure probability for source s . The result in Eq. (5) is conditioned upon source s . To remove this conditioning, i.e., to obtain the final $P(F)$, the theorem of total probability is applied:

$$P(F) = \sum_{s=1}^K (P(F | s) \cdot p(s)) \quad (6)$$

where $p(s)$ is the probability mass function for S . However, that PMF depends on im . To understand that, we recognize that the hazard curve formulated as an exceedance rate is, in a sense, the complementary integral of a histogram. One can imagine that histogram to count the number of im -occurrences in bins along the im -axis. That imaginary histogram, with a finite or infinite bin size, is essentially the slope of the hazard curve. Continuing the histogram analogy, that means that $p(s)$ for a given im is the number of occurrences in the bin around the value im , divided by the total number of occurrences from different sources:

$$p(s) = \frac{\left| \frac{d\lambda(s)}{dim} \right|}{\sum_{s=1}^K \left| \frac{d\lambda(s)}{dim} \right|} = \frac{|d\lambda(s)|}{\sum_{s=1}^K |d\lambda(s)|} \quad (7)$$

Substitution of Eq. (7) into Eq. (6) yields

$$P(F) = \frac{\left(\sum_{s=1}^K (P(F|s) \cdot |d\lambda(s)|) \right)}{\left(\sum_{s=1}^K |d\lambda(s)| \right)} \quad (8)$$

where $d\lambda(s)$ are the changes in the exceedance rate for various sources for a selected *im*-interval.

Part II

Next, assume that *IM* remains a random variable, and further suppose the occurrence probability for each source is known. In other words, suppose $p(s)$ is given, in addition to the fragility functions and the hazard curves. In contrast with the previous section, which did not involve integration along the *im*-axis, the hazard curves must now be considered. In order to obtain the *probability* of failure, we cannot employ the exceedance rates at the bottom of Figure 1a. Rather, it is necessary to employ hazard curves in the form of $G(im)$, shown at the top of that figure. That means we consider as given the fragility functions $P(F|im,s)$ and the hazard functions expressed as $G(im,s)$. First, the integral over *IM* is carried out for each source, using Eq. (2), to obtain $P(F|s)$. Next, the sum over sources is carried out, employing the expression in Eq. (6), yielding $P(F)$.

Part III

Now consider both *IM* and *S* to be random variables, but without any other information than the fragility functions and hazard functions. In other words, we do now know the PMF $p(s)$ beyond the information given by the hazard curves. Perhaps one way to address this problem is to employ the PMF from Eq. (6). Matching Eq. (7), but now employing the CCDF instead of exceedance rates, the PMF is

$$p(im,s) = \frac{|dG(im,s)|}{\sum_{s=1}^K |dG(im,s)|} \quad (9)$$

It is clear that this PMF varies with *im*. For that reason, we now consider the combined total probability integration over the continuous variable, *IM*, and sum over the discrete variable, *S*, producing the following combined version of Eqs. (2) and (6):

$$P(F) = \int_0^\infty \underbrace{\left(\sum_{s=1}^K P(F|im,s) \cdot p_s(s) \right)}_{P(F|im)} \cdot \underbrace{\left| \frac{dG(im)}{dim} \right| \cdot dim}_{|dG(im)|} \quad (10)$$

The total differential in Eq. (10) is not conditioned on the earthquake source. To obtain that unconditional total differential, the following application of the theorem of total probability is made:

$$|dG(im)| = \sum_{s=1}^K |dG(im, s)| \cdot p_s(s) \quad (11)$$

The PMF $p(s)$ appears in both Eq. (10) and Eq. (11). Matching Eq. (7), that PMF is

$$p_s(im, s) = \frac{|dG(im, s)|}{\sum_{s=1}^K |dG(im, s)|} \quad (12)$$

Substitution of Eq. (12) into Eqs. (10) and (11), followed by the substitution of Eq. (11) into Eq. (10) yields

$$P(F) = \sum_{im} \left(\sum_{s=1}^K \left(P(F | im, s) \cdot \frac{|dG(im, s)|}{\sum_{s=1}^K |dG(im, s)|} \right) \cdot \sum_{s=1}^K \left(|dG(im, s)| \cdot \frac{|dG(im, s)|}{\sum_{s=1}^K |dG(im, s)|} \right) \right) \quad (13)$$

Intermediate Response

The modelling behind the equations presented above assume the chain $IM \rightarrow F$. In other words, for a given im the probability of failure is known. A more general modelling approach is to assume the chain $IM \rightarrow D \rightarrow F$, where D is an intermediate continuous random variable, with realizations denoted by d , perhaps representing an inter-storey drift response. The first objective in this situation is to calculate the probability distribution for D , given the hazard curve. Suppose the hazard curve is given as the annual exceedance rate, $\lambda(im)$. Also assume that the CCDF of D , conditioned on im , is also available, denoted $G(d|im)$. The theorem of total probability says

$$\lambda(d) = \int_0^\infty G(d | im) \cdot \left| \frac{d\lambda(im)}{dim} \right| \cdot dim = \int_0^\infty G(d | im) \cdot |d\lambda(im)| \quad (14)$$

where $\lambda(d)$ is the annual rate of exceedance of the drift, D , over the threshold d . Provided a fragility function of the form $P(F|d)$, i.e., the failure probability is determined by the drift, d , the failure probability is, after another application of the theorem of total probability, as in Eq. (3):

$$\lambda(F) = \int_0^\infty \int_0^\infty P(F | d) \cdot |G(d | im)| \cdot |d\lambda(im)| \quad (15)$$

where Eq. (14) has been differentiated and substituted for $d\lambda(d)$.

Multiple Fragility Functions

In earthquake engineering the concept of a fragility function is often extended by replacing the failure event with the entrance into certain damage states. In that case, the horizontal axis is often taken to represent a structural demand, called engineering demand parameter, EDP, such as an inter-storey drift. Figure 2 illustrates fragility functions for damage states ds_1 , ds_2 , and ds_3 . Each curve in Figure 2 displays the probability that the damage is equal to or greater than a specific damage state, for a given demand value. For instance, the left-most curve displays the probability that the damage state is equal to or greater than ds_1 . The probability that the component is in a particular damage state is

$$P(DS = ds_i) = P(DS \geq ds_i) - P(DS \geq ds_{i+1}) \quad (16)$$

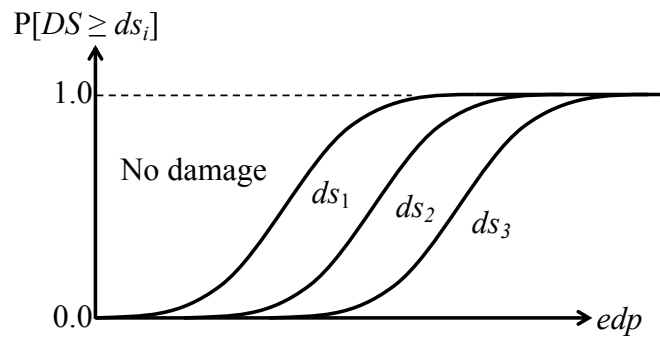


Figure 2: Fragility functions for damage.

Formulation of Fragility Functions

While hazard curves are essentially established by seismologists, a fragility function is ideally the result of a probabilistic analysis using reliability methods. However, fragility functions, such as those in Figure 2, are often described by the lognormal CDF. That choice is made because it is a convenient two-parameter function that increases monotonically from zero to unity. In this case, the lognormal CDF does not represent a probability distribution. As shown in the document on continuous random variables, the CDF of a lognormal variable X can be expressed in terms of the standard normal CDF as follows:

$$F(x) = \Phi\left(\frac{1}{\sigma_Y} \ln\left(\frac{x}{m_X}\right)\right) \quad (17)$$

where m_X =median of X and σ_Y =standard deviation of the underlying Normal random variable $Y=\ln(X)$. With a change in notation to let $d=X$ denote the demand, and applied as a fragility function, Eq. (17) is rewritten:

$$P(DS \geq ds_i) = \Phi\left(\frac{1}{\beta} \ln\left(\frac{d}{\theta}\right)\right) \quad (18)$$

where β =standard deviation of $\ln(d)$ and θ =median of d . It is reiterated that $\ln(d)$ has the normal distribution and that d thus have the lognormal distribution. The fragility function

in Eq. (18) has two parameters, β and θ . Suppose N observations of d are available, namely the demand at which the component enters the considered damage state. To determine θ , start by computing the median of the underlying normal random variable, $\ln(d)$, which is equal to its mean:

$$m_{\ln(d)} = \mu_{\ln(d)} = \frac{1}{N} \cdot \sum_{i=1}^N \ln(d_i) \quad (19)$$

Because $m_{\ln(d)} = \ln(m_d)$, the value of θ , i.e., the median of d , is

$$\theta = m_d = \exp(m_{\ln(d)}) = \exp\left(\frac{1}{N} \cdot \sum_{i=1}^N \ln(d_i)\right) \quad (20)$$

To determine β it is tractable to consider observations of d_i/θ , i.e., a lognormal variable with zero median. As a result, the realizations of $\ln(d_i/\theta)$ are normal with zero mean. In fact, β is indeed the standard deviation of $\ln(d_i/\theta)$. According to classical inference, the variance is:

$$\beta^2 = \text{Var}[\ln(d)] = \text{Var}\left[\ln\left(\frac{d}{\theta}\right)\right] = \frac{1}{N-1} \cdot \sum_{i=1}^N \left(\ln\left(\frac{d_i}{\theta}\right)\right)^2 \quad (21)$$

Thus,

$$\beta = \sqrt{\frac{1}{N-1} \cdot \sum_{i=1}^N \left(\ln\left(\frac{d_i}{\theta}\right)\right)^2} \quad (22)$$

Of course, a conceptually simpler approach is to compute the sample mean and sample standard deviation of d_i and thereafter compute the corresponding distribution parameters of whatever distribution is selected, such as the lognormal.

References

Cornell, C. A., Jalayer, F., Hamburger, R. O., and Foutch, D. A. (2002). "Probabilistic Basis for 2000 SAC Federal Emergency Management Agency Steel Moment Frame Guidelines." *ASCE Journal of Structural Engineering*, 128(4), 526–533.