Poisson Processes

Point Process

The popular Poisson *point* process, often called the "Poisson process" for short, is a Bernoulli sequence in which trials are carried out at every time instant. I.e., the time between trials is zero and the number of trials is infinite. From here on, a trial that yields success will be called "occurrence." Under the assumptions that 1) an occurrence is equally likely to occur at any time instant, 2) any occurrence is independent of what happened before, and 3) only one occurrence can happen at a particular time, the number of successes, *x*, in a time interval, *T*, is given by the Poisson distribution (Ang and Tang 2007)

$$p(x) = \frac{(\lambda \cdot T)^x}{x!} e^{-\lambda \cdot T} \tag{1}$$

where λ is the rate of occurrences, i.e., the mean number of occurrences per unit time. Provided this basis, the Poisson process is referred to as a counting process; it counts occurrences in specific time intervals. The time between occurrences, *t*, has the exponential distribution:

$$f(t) = \lambda \cdot e^{-\lambda \cdot t} \tag{2}$$

The mean time between occurrences is $1/\lambda$ and is habitually called the "return period." Notice that realizations of a Poisson process is easily generated by generating outcomes of *t*, i.e., the random time between occurrences. It is also noted that the Poisson process has only one parameter, i.e., λ. However, there are different ways of expressing this rate. For example, expressions like "2% in 50" appear in earthquake engineering as proxies for λ. To understand their meaning, consider the probability of any non-zero number of occurrences during a time interval *T*, which is provided by Eq. (1):

$$p(1) + p(2) + \cdots = 1 - p(0) = 1 - e^{-\lambda \cdot T} \tag{3}$$

From Eq. (3) one can solve for the rate that yields a 2% probability of occurrence in a 50-year time interval. Expressions like "1-in-50" are also encountered in building codes. This is a direct expression of the rate, i.e., λ=1/50 and consequentially the return period is 50 years. Table 1 exemplifies that the rate is not equal to the annual probability.

Table 1: Return periods, rates, and annual probabilities.

Return period, in years	Rate, i.e., mean annual frequency	Annual probability of occurrence
1	1	1/1.582
5	1/5	1/5.517
10	1/10	1/10.508
50	1/50	1/50.502
100	1/100	1/100.501
500	1/500	1/500.500

1,000	1/1,000	1/1000.500
10,000	1/10,000	1/10,000.500

From Annual to Multi-year Probability

Suppose the annual probability of occurrence of some event is known. Denoting that probability by p_1, the objective here is to determine p_T, i.e., the probability of occurrence in T number of years. Observe that T is an integer in this subsection. Four equivalent alternatives are offered to achieve the objective.

First, p_T is determined from p_1 via the determination of the annual rate of occurrence. Solving $p_1=1-e^{-\lambda}$ for the rate yields $\lambda=-\ln(1-p_1)$. That rate is then substituted into the expression $p_T=1-e^{-\lambda T}$, which yields $p_T=1-e^{\ln(1-p1)T}$, which in turn simplifies to $p_T=1-(1-p_1)^T$.

Second, p_T is determined from p_1 via the expectation of the "system function," ϕ, for a series system of independent components. The expectation of the system function for a series system, i.e., the failure probability for a series system of independent components, is directly $p_T=1-(1-p_1)^T$.

Third, p_T is determined from p_1 via the Poisson distribution, but without determining λ. For the two probabilities, the Poisson distribution says: $p_1=1-e^{-\lambda}$ and $p_T=1-e^{-\lambda T}$. The first of those expressions is rearranged to read $e^{-\lambda}=1-p_1$ and the second expression is written $p_T=1-(e^{\lambda})^T$. Substituting the first into the second yields $p_T=1-(e^{\lambda})^T=1-(1-p_1)^T$.

Fourth, p_T is determined from p_1 by considering each year as a trial of a Bernoulli sequence. The probability of occurrence of "success" in a Bernoulli sequence is $1-p(x)=\binom{n}{x}p^x(1-p)^{n-x}$. The following probability emerges for $x=0$ and $n=T$: $1-p(0)=1-1\,p^0(1-p)^{T-0}=1-(1-p)^T$.

Bayesian Inference

There are several ways to estimate the occurrence rate λ of a Poisson process. The simplest and least precise approach is to divide the number of observations in a time interval by the length of the interval. Another approach is to explore a fit between the probability distribution in Eq. (2) and the observed values of time between occurrences. A good fit would indicate that the exponential distribution is an appropriate probability distribution, and thus that the underlying model is indeed a Poisson process.

Another option is Bayesian updating. In accordance with the general Bayesian principle, the distribution parameter λ is then considered a random variable. A conjugate prior for λ in Eq. (1) is the gamma distribution:

$$f(\lambda)=\frac{\nu(\nu\lambda)^{k-1}}{\Gamma(k)}\exp(-\nu\lambda) \tag{4}$$

Conjugate priors retain their distribution type as a posterior, and the parameters of the distributions are in this case updated by the formulas

$$\begin{aligned} k'' &= k' + x \\ \nu'' &= \nu' + t \end{aligned} \tag{5}$$

where double-prime and prime identifies posterior and prior parameters, respectively, while x is the number of observed occurrences in t.

Pulse Process

In contrast with a point process, a pulse process includes information about the duration and intensity of an occurrence. For this reason, each occurrence is referred to as a pulse. The variation of the intensity within an occurrence can take a variety of shapes, such as the ones shown in Figure 1 (Wen 1990). Each pulse process has only one of the pulse types shown in Figure 1, but the pulses within a realization of a process are usually different in duration and maximum intensity.

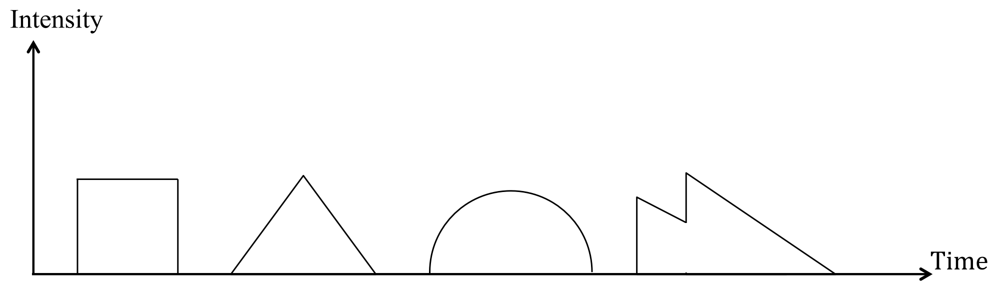

Figure 1: Possible shapes of the "pulse" within an occurrence in a pulse process.

The left-most pulse type in Figure 1 is assumed in the following, because it is most common in practical applications. This constant-intensity pulse type is also employed because it readily facilitates the modelling of loads that are "always on," but with sudden changes in the load intensity. An example of such a process is shown in Figure 2; other details in this figure will be described shortly.

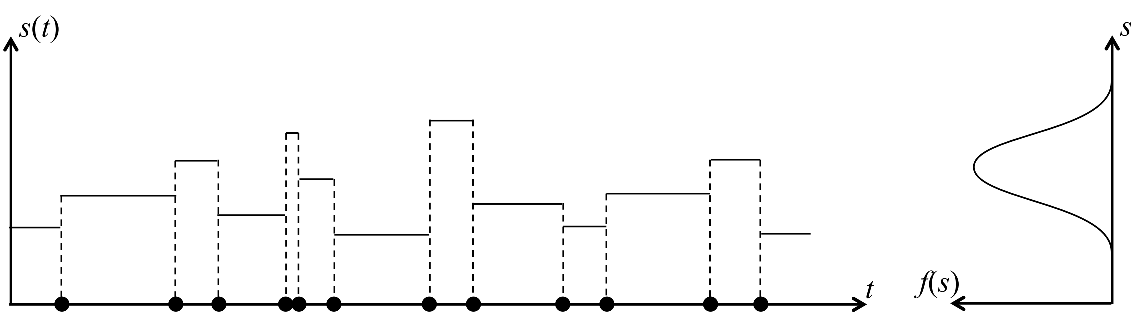

Figure 2: Realization of a Poisson pulse process that is "always on."

While a Poisson point process is characterized by one parameter, the rate, three quantities are needed to define a Poisson pulse process. To understand this model it is useful to think of an underlying Poisson point process with rate denoted by ν. Occurrences of this process are shown by solid circles in Figure 2. For each of these occurrences, the intensity value is "drawn" from the probability distribution at the right-hand side of Figure 2. This is because the intensity of an occurrence, denoted by S, is represented by a random variable, which is usually continuous. A generic probability distribution for S is shown in the right-hand side of Figure 2. For this process, which is always on, every occurrence of the underlying Poisson point process is associated with a change in the intensity. A different case is shown in Figure 3, where the process is "intermittent." In this case, the probability distribution for S has two components: 1) An ordinary PDF with area p_o, and 2) a lumped probability mass equal to $1-p_o$ at s=0. Therefore, when the intensity is drawn from this distribution then there is a probability equal to $1-p_o$ that the intensity is zero. As a result, several of the occurrences of the underlying processes in Figure 3 are associated with zero intensity. In contrast, for the process in Figure 2, which is always on, each occurrence of the underlying process is associated with an actual change in the intensity.

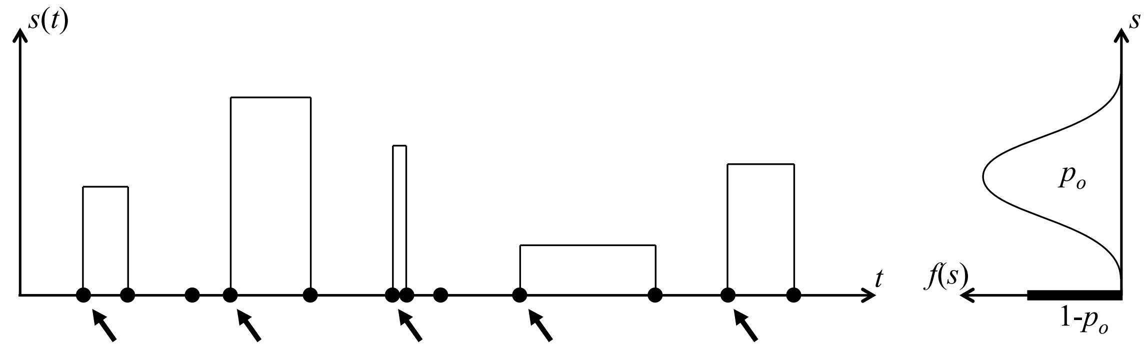

Figure 3: Realization of an intermittent Poisson pulse process.

In passing, it is noted that the PDF at the right-hand side of Figure 3, which is referred to as the "arbitrary point in time" (APIT) distribution, is written

$$f(s) = (1-p_o)\cdot\delta(s) + p_o \cdot f_Y(s) \tag{6}$$

where $\delta(s)$ is Dirac's delta function and $f_Y(s)$ is the probability distribution for the auxiliary random variable Y, which is the intensity of non-zero occurrences. As a result, the corresponding CDF is:

$$F(s) = (1-p_o)\cdot H(s) + p_o \cdot F_Y(s) \tag{7}$$

where $H(s)$ is the Heaviside function. Although the three quantities ν, p_o, and $f(s)$ represent a complete description of a Poisson pulse process, there are alternative representations that are often preferred. In particular, for intermittent processes it is common to employ the rate of occurrence of non-zero intensities. This is another Poisson point process, with rate denoted by λ. Occurrences of this process are identified by slanted arrows in Figure 3. Because the probability is p_o of non-zero intensity at every occurrence of the underlying process with rate ν, the rate of the new point process is:

$$\lambda = p_o \cdot \nu \tag{8}$$

Furthermore, it is common to work with the mean duration of the pulses, μ_d. From knowledge of Poisson point processes, this quantity, which is equal to the mean time between occurrences of the underlying process with rate ν, is:

$$\mu_d = \frac{1}{\nu} \tag{9}$$

In summary, a Poisson pulse process is uniquely described by $f(s)$ together with any pairing of the parameters ν, λ, p_o, and μ_d. A particularly telling parameter is p_o, which by combination of Eqs. (8) and (9) can be written:

$$p_o = \lambda \cdot \mu_d \tag{10}$$

This quantity is tells how frequent the process is "on." If $\lambda\mu_d$=1 then the process is always on. Conversely, if $\lambda\mu_d$ is substantially smaller than unity then the pulses are either infrequent or brief, or both.

Lifetime Maximum

While the APIT distribution describes the intensity at any point in time, an equally important parameter in practical design is the "lifetime" maximum intensity of a pulse process. In the following, the random variable R denotes this quantity for a time period equal to T. By denoting the number of occurrences in T by x, the theorem of total probability yields the following CDF for R:

$$F_R(r) = \sum_{x=0}^{\infty} P(R \le r \mid x \text{ occurrences}) \cdot P(x \text{ occurrences}) \tag{11}$$

Given independence in the occurrence of intensity changes, the first factor is obtained directly from the APIT distribution:

$$P(R \le r \mid x \text{ occurrences}) = \left(F_S(r)\right)^{x+1} \tag{12}$$

because with x occurrences in the underlying Poisson process there are x+1 intensities to take into account, including both ends of the process. The second factor in Eq. (11) is obtained directly from the underlying Poisson process, with rate ν:

$$P(x \text{ occurrences}) = \frac{(\nu \cdot T)^x}{x!} \exp\left(-\nu \cdot T\right) \tag{13}$$

By pulling factors that are independent of x outside the summation, Eq. (11) becomes:

$$F_R(r) = F_S(r) \cdot \exp\left(-\nu \cdot T\right) \cdot \sum_{x=0}^{\infty} \left(F_S(r)\right)^x \cdot \frac{(\nu \cdot T)^x}{x!} \tag{14}$$

It is next recognized that the summation is a series expansion of an exponential function:

$$\sum_{x=0}^{\infty} \frac{(F_S(r) \cdot \nu \cdot T)^x}{x!} = \exp\left(F_S(r) \cdot \nu \cdot T\right) \tag{15}$$

As a result, Eq. (11) condenses to:

$$\begin{aligned} F_R(r) &= F_S(r) \cdot \exp\left(-\nu \cdot T\right) \cdot \exp\left(F_S(r) \cdot \nu \cdot T\right) \\ &= F_S(r) \cdot \exp\left(-\nu \cdot T(1 - F_S(r))\right) \end{aligned} \tag{16}$$

This probability distribution for the life-time maximum value is further refined by introducing the expression for the APIT distribution in Eq. (7), which according to the previous discussion can be written:

$$F_S(s) = \left(1 - \lambda \cdot \mu_d\right) \cdot H(s) + \lambda \cdot \mu_d \cdot F_Y(s) \tag{17}$$

By also replacing ν by $1/\mu_d$ in Eq. (16) according to Eq. (9) for uniformity of notation, as well as substituting Eq. (17), Eq. (16) reads:

$$\begin{aligned} F_R(r) &= F_S(r) \cdot \exp\left(-T / \mu_d (1 - F_S(r))\right) \\ &= \left[(1-\lambda\mu_d)H(r) + \lambda\mu_d F_Y(r)\right] \cdot \exp\left(-\frac{T}{\mu_d} \cdot \left[1 - H(r) + \lambda\mu_d H(r) - \lambda\mu_d F_Y(r)\right]\right) \end{aligned} \tag{18}$$

For r>0, i.e. for threshold above zero intensity the Heaviside function is unity, thus:

$$F_R(r) = \left[1 - \lambda\mu_d + \lambda\mu_d F_Y(r)\right] \cdot \exp\left(-T\lambda\left[1 - F_Y(r)\right]\right) \tag{19}$$

Furthermore, for high thresholds, i.e., large values of r, the CDF $F_Y(r)$ is close to unity, which leads to the following approximation (Wen 1990):

$$F_R(r) \approx \exp\left(-T\lambda\left[1 - F_Y(r)\right]\right) \tag{20}$$

This expression for the CDF for the lifetime maximum intensity is often employed in practice because many reliability applications deal with rare failure, i.e., high values of r.

References

Ang, A. H.-S., and Tang, W. H. (2007). Probability concepts in engineering: emphasis on applications in civil & environmental engineering. Wiley.

Wen, Y. K. (1990). Structural load modeling and combination for performance and safety evaluation. Elsevier.