The "100 year flood"

Douglas A. Kerr

Issue 2
March 21, 2025

PREFACE

We might read that some community, as a result of a severe series of rainstorms, was hit by a "100 year flood". That sounds like a pretty severe flood. But just what does that expression mean? This article discuses that and a number of related issues.

1 THE "100 YEAR" FLOOD

We might read that some community, as a result of a severe series of rainstorms, was hit by a "100 year flood". That sounds like a pretty severe flood. But just what does that expression mean?

If someone has no real background in statistics, he might explain that to a friend as meaning "the kind of flood that we could expect to only happen once 100 years". But, especially if we have some knowledge of statistics, we might wonder exactly what **that** means.

2 A PRECISE DEFINITION

2.1 The concept of the definition

To be more precise, the term "100 year flood" is usually taken to mean a flood of such severity that, for the locality being discussed, over a long time,[1] the average time between successive floods of such severity or greater is 100 years. I call that time the *mean recurrence interval* (MRI) for a flood of that severity)

2.2 A directly related metric

It can be shown rigorously (see Appendix A) that, for the situation just described, the average number of floods of that severity or greater per year is 0.01 (the reciprocal of the *mean recurrence interval*). This holds regardless of the nature of the statistical distribution of the occurrence of floods of various severity. This metric is sometimes called the *mean occurrence rate* of floods of that severity or greater.

[1] Assuming that the statistics of flood occurrence at that locality do not change over that "long time". This is of course not likely in reality, but is a necessary part of the "mathematical fiction" underlying the definition.

2.3 A temptation

In is then tempting to think that, for the situation described just above, the probability that there will be one or more floods of that severity or greater is also 0.01 per year.

That would in fact be rigorously so if the occurrence of floods followed the *binomial distribution*. But in fact an analysis of historical data suggests that the statistics of flood occurrence typically very closely follows a different statistical distribution, called the *Poisson distribution*. Much analytical work on flood occurrence is predicated on that distribution.[2] (Appendix B gives more information on the Poisson distribution.)

But under the Poisson distribution, the annual probability of occurrence is not inherently the same as the reciprocal of the mean recurrence interval.

However, as the mean recurrence interval increases, the Poisson distribution approaches the binomial distribution. Thus, for example, at a *mean recurrence interval* of 100 years, the *annual probability of occurrence* is very, very nearly the same as the reciprocal of the *mean recurrence interval*, so the distinction is of no real consequence..

But if we speak in terms of a flood with a substantially lower *mean recurrence interval* (perhaps 5 years), the probability of a flood of such severity or greater occurring in any given one year period (its AEP) is significantly different from the reciprocal of the *mean recurrence interval*.

2.4 Locality-based

It is important to realize that a "100 year" storm is not, in general, a storm of any defined severity. The severity of the "100 year storm" varies with the locality in which we are interested.

So an flood described as a "100 year" flood is a far less severe flood at a city astride a well contained small stream than would a "100 year" flood at a city astride a barely-contained normally-raging river. The "100 year flood" in the former location might cause a slight overflowing of the river onto the adjacent riverwalk, while the "100 year flood" in the latter location might fill Main Street with two feet of water.

---

[2] This can be seen in that many often-stated numerical implications are consistent with the Poisson distribution having been assumed)

3 ACCORDING TO THE USGS

3.1 Introduction

The United States Geological Survey (USGS) is the federal agency one of whose many responsibilities is the collection and analysis of data regarding to the occurrence of floods, and the publication, for numerous locations, of information on the likelihood that a flood of a certain magnitude or greater will occur.

3.2 The recurrence interval

In prior times, the USGS characterized floods at a specific locality in terms of what they call the *recurrence interval* of a flood having the severity of interest (or greater), **at that locality.** Thus has essentially has the same definition as the *mean recurrence interval* (MRI) of which I spoke earlier. I will use the USGS form of the term for a little while.

3.3 The annual exceedance probability (AEP)

Today, the USGS characterizes floods at a specific locality in terms of the probability that one or more floods of the severity of interest described will occur in any arbitrary one year period at that locality. The probability is called by USGS the *annual exceedance probability* (AEP) of such a flood.

3.4 But not quite

The USGS often states unequivocally that the AEP is the exact reciprocal of the *recurrence interval*. And in fact, for a recurrence interval of 100 years that is **almost** exactly true.

But because of the situation described in section 2.3, the AEP is not "inherently" the reciprocal of the recurrence interval.

Nevertheless, for a recurrence interval of 100 years, assuming the Poisson distribution, the calculated AEP is very close to 0.01 (0.00995), so the difference is not of any consequence in that case.

However, for lesser values of the recurrence interval, the calculated AEP can depart sufficiently from the reciprocal of the recurrence interval. For example, for a recurrence interval of 2 years (the reciprocal of which is 0.5), the calculated AEP is 0.393.

3.5 Locality-based

Keep in mind that, as with the "100 year flood" characterization, an AEP value does not in any way indicate in an absolute way (by any measure) the severity of the flood referred to. Rather it only tells us, *for a specific locality*, how "rare" is the flood being characterized.

4 A MISAPPLICATION OF AEP

The rather odd nature of the metrics (*mean*) *recurrence interval* (MRI) and *annual exceedance probability* (AEP) can lead writers to embrace faulty concepts of their meanings or to inappropriately apply the term.

An example is in defining the areas within some city that are considered Special Flood Hazard Areas (SFHA), which affects the matter of federal flood insurance and often local building code restrictions.

The official (and in my opinion “appropriate”) definition of these areas is (and I paraphrase), “any area for which there would be a probability of 0.01 or more that they would have any significant flooding in any arbitrary one year period”.

> As to what I imprecisely call, for conciseness, “significant flooding”, I note that for the National Flood Insurance Program the definition of “flood” is:
>
> > A flood is a general and temporary condition of partial or complete inundation of two or more acres of normally dry land or two or more properties, caused by overflow of inland or tidal waters, unusual surface water runoff, or mudflow

In any case, we frequently find that formal definition of an Special Flood Hazard Area paraphrased as “the areas for which AEP is 0.01 or more.”

Of course, as we saw earlier, AEP is not a measure of the probability that a certain area will experience some certain severity of flooding (in this case, any “flood” at all, as that is defined in this context).

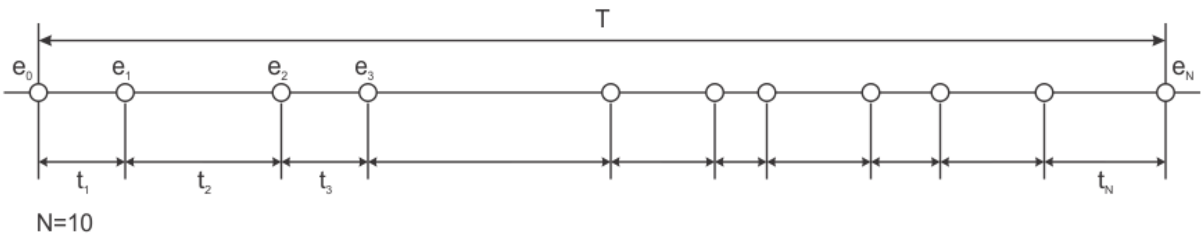
So we see that transforming the actual definition of a Special Flood Hazard Area into “the areas for which AEP is 0.01 or more” is not meaningful.

-#-

Appendix A
Mean recurrence interval and its implications

In the body of the article, I asserted that if t_0 is the average time between events (the *mean recurrence interval*) is t_0, the average number of events per unit time is $1/t_0$.

This is demonstrated using the following figure:



We see on the timeline a N+1 events (N=10 in this illustration), labeled e_0 through e_N. The time to event i (from the prior event) is t_i. The average time between two events in this scenario is t_{avg}.

The total time span we consider is T, and I have arranged the illustration so that this runs exactly between the times of occurrence of two events to avoid bothersome "end effects".

Since the average time between two events in this scenario is t_{avg}, then the total time from the first event to the last must be just Nt_{avg}.

It then follows that the mean rate of occurrence of the events must be just $1/t_{avg}$.

Accordingly, the average number of events occurring in some time T_x is T_x/t_{avg}.

It is tempting to think that the probability that an event will occur in an interval of length one unit is also $1/t_{avg}$..

That would in fact be so if the distribution involved were the *binomial distribution*.

But if the specific kind of events, as described in the body of the article, follow the Poisson distribution, it does not work out exactly that way. The formula that describes the probability that a certain number of events will occur in a certain time interval under the Poisson distribution is given in Appendix B

Appendix B
The Poisson distribution

B.1 INTRODUCTION

In the body of this article, it was reported that historical data shows that the occurrence of floods of at least some certain magnitude closely follows the *Poisson process* model, resulting in a *Poisson distribution* of these occurrences.

B.2 THE POISSON PROCESS

The important defining features of a Poisson process are:

- The average rate of an event occurring (per unit time) is constant over time. (At least over the period in which we are interested.)

- The probability an event will occur in an given interval of time is not affected by the time since the prior event.

B.3 THE POISSON DISTRIBUTION

The distribution of events in a Poisson process is called the Poisson distribution. The fundamental formula for, in a Poisson distribution, calculating the probability that there will be a certain number of events in a certain interval is:

$$P_k = \frac{\lambda^k e^{-\lambda}}{k!} \qquad (1)$$

where P_k is the probability that there will be k events in that certain interval and λ is the *mean event rate* in events per the duration of that certain interval. The factor $k!$ is the factorial of k, the product of all integers from 1 up through k. Note that k must be an integer; λ need not be.

B.4 EXAMPLES

Suppose we know that the average number of customers that, during the one-hour period 2:00-3:00 pm, enter a bakery is 2.5 (and that the behavior of this flow is Poissonian). (That is a rate, λ, of 2.5 customers per hour.)

Then, if we wish to know the probability that exactly 4 customers will enter the bakery during that one-hour period on any given day, we can use Formula 111, with λ=2.5 and $k=4$. The result is about 0.133.

What does that mean? Well, it means that over a long time (assuming the nature of the customer flow at that time of the day remains unchanged), we could expect (n the statistical sense) to have exactly

4 customers during that hour in about one day out of 8 (approximately 1/0.133).

What is the probability that, in some certain one-hour period (at that time of day) there will be no customers? Here we use λ=2.5 and *k*-0 (and note that 0! is defined as 1). The result is about 0.0821.

B.5 RELATIONSHIP BETWEEN THE MEAN RECURRENCE INTERVAL AND THE ANNUAL EXCEEDANCE PROBABILITY

In the body of the article I pointed out that, if the actual occurrence of floods of at least some certain severity or greater followed the Poisson distribution, **it was not true** that the probability that flood of at least that severity would occur, per year (the *annual exceedance probability* [3], was the reciprocal of the *mean recurrence interval* (MRI) of such a flood.

Nonetheless, for large values of the MRI, the probability that such a flood would occur, per year was very close to the reciprocal of the *mean recurrence interval* (MRI) of such a flood.

Thus, especially if we are speaking of the “100 year” flood (the severity most often spoken of), we can with only minuscule error reckon the annual probability of occurrence of such as flood as the reciprocal or the MRI of such a flood.

Oddly, the US Geological Survey often states unequivocally that the what they call the *annual exceedance probability* (AEP) of a certain severity of flood is exactly the reciprocal of the *recurrence interval* In some case, in their public information, they even give a table showing that equivalence, for values of the MRI as low as 2 years.

This table shows that supposed relationship, along with the actual annual probability of having one or more floods exceeding that implied severity, assuming the Poisson distribution.

---

[3] The USGS uses the term ***annual*** *exceedance probability* (AEP), but I mean my term to refer to exactly the same property.

(Mean) recurrence interval (years)	Annual Exceedance Probability (AEP) as given by USGS (%)	Calculated annual probability of exceedance under the Poisson distribution (%)
100	1	0.995
50	2	1.98
25	4	3.92
10	10	9.52
5	20	18.1
2	50	39.3

We see that only for the smaller values of the mean recurrence interval are the two expressions of annual probability substantially different.

Nonetheless, the USGS’ insistence that the AEP is the reciprocal of the recurrence interval (even for small values of the latter) is disppoiting.

B.6 AN OFTEN-CITED EXAMPLE

An example of the working of the Poisson distribution, in the matter of floods, that is often cited is this. A mortgage lender wants to know, with respect to a property in a certain locality, what is the probability that, over the 30 year term of the proposed mortgage, the property will be the subject of at least one “100 year flood”.

Formula 1 tells us that this probability is 25.9%.[4]

The narrative of the example is of course hard to believe. Why would the mortgage lender think that only a “100 year flood” would cause enough damage to be of concern?

Well, it might be that the property of interest had been declared to lie in a Special Flood Hazard Area (SFHA)[5], meaning that the mean recurrence time of “any significant flooding” (my term) is 100 years (see the discussion of this in Section 4). Thus the flood that causes “any significant flooding” is the “100 year” flood

-#-

[4] The fact that this is widely cited (with the identical result) is evidence that indeed the Poisson distribution is assumed in the matter of flood occurrence.

[5] See section 4.