Exponentials in Statistics

More Than You Wanted To Know

Suppose we want to model the following situation. We are checking if we are getting any hits on our brand new web page, and we can reasonably assume that such hits arrive totally at random. That is, while we might have a higher or lower *average* rate of hits, since people are connecting from all over the world, one connection does not "drag along" more connections: each arrives independently of every other.

Let us look for a function, which we'll call $a(t)$, describing the "probability" (in an intuitive sense), of not having received any hits between our start ($t = 0$), and time t. Now, suppose you have received no hits at time t, and you ask yourself what the probability will be not to receive any hot for another s units of time, *knowing that you had no hits in the first* t *units.*

Here's the tricky point. Think of all possibilities as a big set A. The probability that something in A happens is 1, (100%), since something, anything has to happen, by definition. Now, we know that no hits arrived between 0 and t, so we have shrunk the available possibilities to this subset, whose probability is $a(t)$. Since we are asking for the probability of not receiving hits in the next s units of time, given that we are in this smaller subset of possibilities, it is reasonable to accept that this probability is given by

$$\frac{\text{probability of no hits in } t+s \text{ units of time}}{\text{probability of the set of possibilities that we are in know}} = \frac{a(t+s)}{a(t)}$$

On the other hand, if hits arrive completely at random, without any dependency on whether any arrived before, the probability of not getting hits in the next s units starting now, at time t, should be the same as that of not receiving any hits in s units$<$1 of time, no matter when we start counting. Hence, we can reasonably assume that

$$\frac{a(t+s)}{a(t)} = a(s)$$

$$a(t+s) = a(t)a(s)$$

This is "functional equation", and we know one possible class of functions that has this property: exponential functions:

$$b^{t+s} = b^t b^s$$

Since probabilities are no larger than 1 (corresponding to 100%), and times are positive, we need the base b to be smaller than 1. It turns out (it's easy to show, with some calculus, but maybe not so easy with algebra alone) that exponential functions are the *only* functions with this property (at least if we require the solution to be reasonably nice, like with no jumps in its graph). Since (again, this is best understood within calculus) it's just inconvenient to use a base different from e, this result is stated as

> With the assumptions made at the beginning, the probability of not receiving any hit in t units of time is e^{-kt}, where k is a constant related to the expected number of hits per unit of time[1]. This model is called "the exponential probability distribution"

You hope that bus arrivals at a stop do *not* follow this model. If they did, the likelihood that a bus will show up in the next five minutes would be the same, whether you just arrived, or whether you had been standing there for an hour!

[1] Here $k > 0$, and the function is written as e^{-kt} to make sure that we have a number that is between 0 and 1 - if you prefer, we are using the base $e^{-1} < 1$.