Freedman's Rabbit-Hat Theorem and Earthquake Probabilities

P.B. Stark

Department of Statistics

University of California, Berkeley

DRAFT June 25, 2013

The rabbit axioms:

1. For the number of rabbits in a closed system to increase, the system must contain at least two rabbits.

2. No negative rabbits.

From these, Freedman's Rabbit-Hat Theorem follows:

Theorem 1 (Freedman) *You cannot pull a rabbit from a hat unless at least one rabbit has previously been placed in the hat.*

Corollary 1 *You cannot "borrow" a rabbit from an empty hat, even with a binding promise to return the rabbit later.*

How does the Rabbit-Hat Theorem apply here? One cannot conclude that a process is random without making assumptions that amount to assuming that the process is random. (Something has to put the randomness rabbit into the hat.) Testing whether the process appears to be random using the *assumption* that it is random cannot prove that it is random. (You can't borrow a rabbit from an empty hat.) Similarly, one cannot conclude that a process is a stationary random process without assumptions strong enough to imply that the process is both random and stationary. Observing the process isn't enough.

If I understand what you wrote, there's (1) a tacit assumption that if failures at different locations are uncorrelated, we get a constant expected number of failures per year, and (2) an assertion that if a failure process has been going on for a long time, the failures will be approximately uncorrelated.

I claim that neither is true without additional assumptions: there's no rabbit in the hat yet. I'll give a couple of examples.

Let's just have a single type of bulb, installed at N locations. Suppose that the bulb installed at the ith location fails for the kth time at time T_{ik}. We define T_{i0} to be 0 for all i. For illustration, I'll assume that the failure times are random, but that is a huge assumption that I would not make lightly if this were a real scientific problem: It has no physical basis here. We will look at two different stochastic models for T_{ik} and what they imply.

Case 1. $T_{ik} = km+e_{ik}$, $k > 0$, where m is the mean time between failures and $\{e_{ik}\}$ are independent, identically distributed, zero-mean, random variables with finite variance σ^2. This is a "characteristic earthquake" model with a noisy clock, as opposed to a renewal model. In this model, failure times at the same location and across locations are uncorrelated. Indeed, the failure times are independent. But if σ is small compared to m, most of the failures will occur at times close to multiples of m, and if you stock spare bulbs based on the average rate of failure, N/m, you will have too many most years and far too few when you need them. This example shows that uncorrelated failures is not enough to give the uniformity you are relying on.

Case 2. $T_{ik} = T_{ik-1} + e_{ik}$, $k > 0$, where $\{e_{ik}\}$ are independent, identically distributed nonnegative random variables with mean m. This is a renewal model. Failure times at the ith location are now dependent, but failure times across locations are independent (and hence uncorrelated). Let $N_i(t)$ be the number of failures at location i in the time interval $(0, t]$, and let $U_i(t) = \mathbb{E}N_i(t)$. In this case, the Blackwell Renewal Theorem tells us that for each i and every fixed h,

$$\lim_{t\to\infty} U_i(t+h) - U_i(t) = h/m. \tag{1}$$

That is, at each location i, the rate of failures is asymptotically unconditionally uniform; since the bulb are exchangeable, the rates are the same at all N locations. Of course, the conditional rate of failures given the time of the last failure will not be uniform.

In this situation, your argument about "waiting long enough" is true because it holds for each i separately—it has nothing to do with correlation of failures across locations. But it is still not true conditional on the times of the most recent failures. That is, given that the most recent failure at location i was at time t_i, $i = 1, \ldots, N$, it is not true that the expected rate of failure per unit time does not depend on time nor that the expected rate of failures across all N sites does not depend on time. Again, the number of spares it makes sense to stock in a given year might have little relation to the average rate of failures.

By the way, there's an interesting theorem called the Waiting Time Paradox, which goes as follows. Suppose you are observing a renewal process with mean time between failures equal to m. Pick "a random point in time" t and look at the interval between the last failure at or before t and the next failure after t. The expected length of that interval is greater than m. (If we pick a point uniformly in time, we are more likely to hit a long interval than a short interval.) Hence, if it became our job to maintain the lightbulbs starting at some arbitrary time t, and lightbulb failures followed a renewal process with MTBF m, we should expect at the start that the rate of failures would be *less than* $1/m$ per location on average, and should stock spares accordingly.