Review, R. Styron, “The impact of earthquake cycle variability on neotectonic and paleoseismic slip rate estimates.”

Uncertainty in estimates of earthquake recurrence and fault slip rate are important parameters, pursued by conscientious investigators of seismic hazard. The author perceives a lack of statistical support, and offers in this paper “insights” in variously tones of “friend of the practitioner” and “trust me, I’m the numerate one here”. Neither is convincing.

A couple of omissions in this paper are particularly striking. First, how do we have a paper addressed to “variability” in fault slip rate, addressing particularly the problem of small samples, without mentioning the methods of estimation for censored samples? There is an extensive statistical literature to estimate parameters and uncertainties and numerous recent papers applying it in paleoseismic contexts. This literature provides real quantitative methods to deal with the open intervals, long or short, that affect the geologist’s estimate of fault slip rate and recurrence estimation. These are real equations, with real uncertainties. One would look in vain in this paper for anything of similar substance. Second, pages of this paper could be replaced (and improved) by a presentation and discussion of the properties of the standard error. E.g., given an estimate a sample-based estimate of the mean, how far might the population (or true) mean be from the estimate? S.E. is estimated by the sample standard deviation divided by the square root of the number of samples. So, of course, estimates from small samples from a fuzzy log normal converge more slowly than from a well-defined (quasi-periodic) lognormal. Instead of a small equation (SE=s/sqrt(n)), our paper back-calculates the result using 2 million years of samples, and presents the results like a new discovery. And again, with little by way of meaningful uncertainties (e.g., p1, lines 11-13, 14-16).

A few particulars

1 L9,10: We read that the most important parameter is the coefficient of variation. First, this equation is the arithmetic coefficient of variation, and not the CV for a lognormal distribution. The CV of a lognormal does not depend on the mean. We could stop here, but a central flaw in the paper is exposed – nothing in this paper addresses how to obtain this most important parameter. If attempted, the essential emptiness of a 2,000,000 year sample would emerge. No real data set in paleoseismology resolves the mean and standard deviation to better than maybe 50%. Typical sites do well to resolve it to a factor of 2. P.3, line 22-23 reflect this reality.

It is not obvious that the author has material experience words “aleatoric” and “epistemic”. Line 1, “aleatoric uncertainty” is a contradiction in terms. Bird, Zechar and Frankel all know better than to use the method the author alleges in lines 21 and 22 to arrive at epistemic uncertainty in slip rate. They would more likely consider the allegation a misreading of their work. I could multiply examples. More broadly, the lack of care in writing makes one wonder how to understand this paper. p2, L5: A perturbation in slip rate would mean it was slipping at rate X, then changes to Y. p.2, L14-17 have careful paleoseismologists doing reasonable things in one sentence, then imply they would make plainly rookie mistakes in the next. From here these read like inexperienced generalizations.

p.4, L15-23: The descriptions of the lognormal variables here give one pause. First, log-normal parameters do not have units. Second, the mean recurrence interval is not the location parameter of a log normal. This is just wrong. Third, if one uses the CV equation for the lognormal distribution (e.g., https://en.wikipedia.org/wiki/Log-normal_distribution#Arithmetic_coefficient_of_variation), the CV will not match the COV alleged here.   Given that the study depends on these distributions, we can't really use subsequent conclusions.

p.6, L20. If the number of samples is really n = N – t + 1, the samples are correlated by virtue of the overlap in the windows. No accounting has been made of the correlation structure.

p.6, L26: Starts a narrative of the consequences of the standard error, as though the standard error was never invented. The fuzzy, back-of-the-envelope estimates start to get thick here. Real uncertainty estimates would serve better.

p.7, L1-3: Two observations: First, as written, the practicing geologist is being asked to believe that 60 earthquake cycles have passed with zero displacement. I can guess what was intended, but should not have to. Second, what probability is associated with this 60-cycle thing? I ask because practicing hazard geologists have to make estimates, and give weights to extreme events. What is the probability of 60 cycles, a CV of 2.0, …? Hard to imagine that the author has thought much about what these results would mean or how to use them if they were true.