14th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP14
Dublin, Ireland, July 9-13, 2023

Poorly-informative priors in geotechnical risk analysis

Desmond N.D. Hartford
Principal Engineer, Dam Safety Management, BC Hydro, Burnaby BC, Canada

Gregory B. Baecher
Prof., Dept. of Civil and Environmental Engineering, University of Maryland, College Park MD, USA

ABSTRACT: Quantitative risk analysis has become a common part of geotechnical engineering. In the Bayesian context, we fuse prior beliefs about adverse outcomes with the Likelihood of observed data or modeling outputs to form updated beliefs about risk. The prior beliefs and the likelihoods are equal contributors to our qualifications of risk, especially in the context of sparse information, which is common in geotechnical projects. In domains such as dam safety, seismic hazard assessment, and flood damage reduction Bayesian priors have come to depend on subjective probabilities derived by quantifying engineering judgment. These may be poorly-informed and based on personal experience. Questions arise of validation and calibration of these priors and how to understand the role of reference priors.

1. INTRODUCTION

Every practitioner of Bayesian analysis recognizes the importance of prior probabilities. Yet, IJ Good (1983) famously estimated that there are 46,656 varieties of Bayesians, and thus one presumes as many approaches to priors. Bayes Rule describes how a prior probability (or distribution) can be updated to a consistent posterior probability after having observed data. Thus, we are confronted with the familiar equation,

$$f(A|B) \propto f(A)\, p(B|A) \qquad (1)$$

in which A is an uncertain entity and B is a set of observations, $f(A)$ is the marginal (prior) probability and $p(B|A)$ is the conditional probability (*Likelihood*) of the data. The entity A might be a hypothesis, model, parameter, data point, or something else. Bayes Rule provides a compact way of fusing the prior probability with data to arrive at a logically consistent updated probability. As Hacking (2001) noted, ‘Bayes Rule is trivial, but it is very tidy’.

Most applications of Bayesian methods focus on the machinery of probability updating rather than the meaning of the probabilities. As a result, the focus is on the Likelihood model and not on the prior probability. It is common to invest attention on the detailed stochastic representation of the Likelihood while relegating the prior to a minimally-informative placeholder. Yet, it is obvious that the prior probability is just as important as the Likelihood in drawing inferences, and the facility for fusing information of different sorts is a principal benefit of Bayesian thinking.

2. TYPES OF PRIOR PROBABILY

Many approaches to geotechnical priors have been suggested. The more common are:

Empirical priors are those based on pre-existing information about the inferences in question. These might include earlier data, engineering modeling, or reasoning from first principles. Such priors are in principle available but seem rarely used in geotechnical applications.

Diffuse priors include the uniform, locally-uniform, and Jeffreys (non-informative). The first two are flat over the region of the Likelihood. Such locally-uniform priors are used for regularization of ill-posed problems in data science. The Jeffreys prior intends to be invariant to parameter transformations. Each of these is suggested to be ‘objective’.

Minimally informative priors are those that maximize some measure of information. These include maximum entropy distributions (*e.g.,* Gilbert *et al* 2016), those which maximize other information measures, and Bernardo's 'reference priors' which maximize the Kullback-Leibler divergence (*i.e.,* maximum difference) of posterior *vs.* prior.

Conjugate priors are those closed under multiplication by the Likelihood. They are chosen, principally for convenience, but were popular before the advent of Markov chain Monte Carlo (MCMC) and other numerical methods. Conjugate-diffuse priors can be had by assuming a large variance.

Subjective priors are based on personal probabilities reflecting degrees-of-belief. Much has been written about this in the geotechnical literature. One should note that the choice of Likelihood is similarly subjective and equally influential.

3. REFERENCE PRIORS

While subjectivism is widely talked about, most Bayesian analyses in geotechnical applications use reference priors. Box and Tiao (1973) used the term *reference prior* for any prior dominated by the Likelihood and used as a standard. We adopt that meaning. These are chosen by a formal rule and intend to be objective. MCMC has motivated expanded use of reference priors. For large samples, asymptotic behavior makes the choice mostly unimportant. Obviously, this is not true for small samples, and even with large samples the choice of reference prior may at times affect the posterior (Kass and Wasserman 1996).

The obvious questions are: how and in what way are priors important, to what extent are reference priors useful, and when (and how) should informative priors be used? A significant advantage of Bayesian thinking is its facility for fusing information. This can include qualitative and quantitative data, categorical or scenario findings, and subjective information. In most geotechnical applications, data are sparse. Therefore, the prior carries important information for risk analysis.

3.1. Frequentist results in Bayesian clothing

Using a reference prior may yield Bayesian results which differ little from frequentist results, except in interpretation: (1) confidence intervals of the frequentist method may be numerically like the Bayesian credible intervals, (2) the maximum Likelihood estimate of the frequentist method will be the same as the mode of the Bayesian distribution, and (3) the Likelihood ratio of the frequentist method will be the same as the Bayes Factor of the Bayesian. Thus, to what purpose is Bayesian analysis? One could invoke Fisher's fiducial concept and use the frequentist results as if they were probabilities (Zabell 1992).

Uniform priors follow the *Principle of Insufficient Reason* arising from symmetry and classical probability. While intuitively appealing, the principle is not invariant to how sample space is partitioned. It is widely used in hypothesis testing for prior probabilities. Non-informative priors are attributed to Jeffreys (1931) who held that probability is 'not a matter for personal judgment'. In *Theory of Probability* (1939) he proposed two reference priors, one for estimation and one for hypothesis testing. For the former he proposed that the prior should be invariant under a transformation of variables. The practical implication was to use a uniform prior over θ for $\theta \in \mathcal{R}: (-\infty, \infty)$, and a uniform prior over $1/\theta$ for $\theta \in \mathcal{R}: [0, \infty)$. For hypothesis testing using Bayes Factors, he proposed treating the initial probabilities of the hypotheses as equal. Use of the reference priors meant that two researchers would reach the same conclusions from the same data.

The interpretation of reference priors is twofold. On the one hand, they may be taken as a formal representation of ignorance; on the other hand, they may be taken simply as a default. Often the reference prior is improper in the sense that its integral vanishes. While that seems a trivial problem, it may lead to inconsistencies in high-dimensional problems (Kass and Wasserman 1996). The latter are not uncommon in geotechnical applications. However, this is beyond the present scope. A workaround for this problem is to use a proper prior which is diffuse in the region of the Likelihood. However, to the extent that the reference prior is simply a default, what does that imply for using the posterior probability in risk analysis? The reference prior as default may be suitable for publishing in journals, but it is likely unsuitable for decisions since it is not a fused probability of all available information.

3.2. Telling us what we already know

A poorly-informative prior will not entirely determine a posterior probability or distribution. Figure 1 shows the relationship between prior and posterior probability in the search for a site anomaly. Does the anomaly exists if undetected (Tang 1987). The principle of insufficient reason suggests a prior of p=0.5 on H_0: the anomaly exists; and p=0.5 on the alternative H_a: it does not. The anomaly is not found (z). Presume the probability of finding an existing anomaly is high, p=0.8. The Bayes Factor in favor of H_0 given z is thus $[(1-0.8)/1.0] = 0.2$.

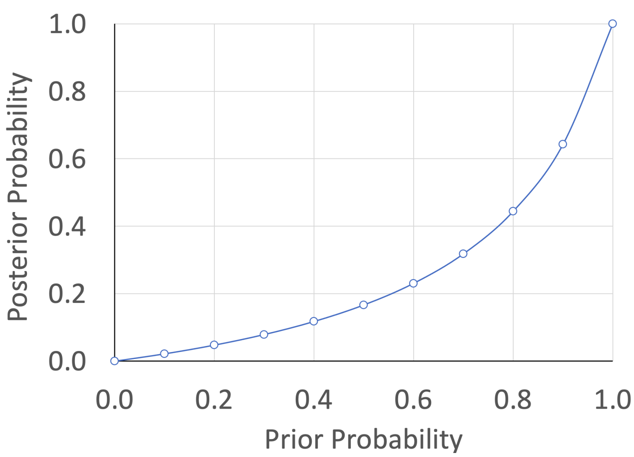

Figure 1. Posterior probability of an undetected anomaly given its prior probability.

The posterior probability of H_0 as a function of the prior probability is shown in Figure 1. What is the impact of our choice of priors? We could have chosen another prior and the posterior probability could have been anything from 0 to 1. To some extent the Bayes calculation simply reflects what we had believed *a priori*.

4. SUBJECTIVE PRIORS

A great deal has been written about subjective probability for geotechnical risk. Bayesian thinking demands only that subjective priors be coherent. It doesn't demand that they have predictive validity. In geotechnical use, however, the subjective prior may strongly affect the posterior. One wants the prior to be more than simply coherent: one wants it to be valid and calibrated. *Valid* means the prior reflects an intrinsic degree of belief; *calibrated* means the prior corresponds to frequencies in the world. The former is philosophically required; the latter is operational preferred.

4.1. Valid priors

The validity of subjective priors relies on the elicitation protocol and its execution. There is a large literature on this topic spread across many disciplines (Baecher 2019). It need not take space here.

4.2. Calibrated priors

When an individual assigns a subjective probability, the values are correct to the extent they reflect a valid degree of belief. Two individuals may each assign a valid subjective probability, and they may differ, but they will both be correct. Neither is 'wrong'. On the other hand, if a third-party wishes to use the expert's subjective probability as part of a prior for decision making, one would like the expert's subjective probability also to be calibrated. If the expert says, $p = 0.1$ of some event happening, then over repeated trials, one expects that 10% of the time that event obtains. Morris (1974) talks of this as treating the expert as an instrument with a calibration curve.

Creating an expert's calibration curve is not difficult and provides feedback for improvement. The personal calibration curve of Figure 2 for one of the authors (GBB) was constructed by making many predictions of "almanac" like events in the daily news and tracking how often they later happened: *did-happen = 1; did-not-happen = 0.* A simple logistic regression yields the curve. Here, the subject is overconfident. He overestimates the probabilities of likely events and underestimates those of rare events. The wise third-party would adjust this expert's probabilities accordingly.

Treating the expert as an instrument, provides a logical basis for developing consensus distributions. The respective calibrations for a committee of experts can be treated as marginal likelihoods in the Bayesian sense. An updated consensus probability or probability distribution can be formed from the joint Likelihood. Obviously, some consideration needs to be given to potential correlations among the experts' opinions, which are not uncommon.

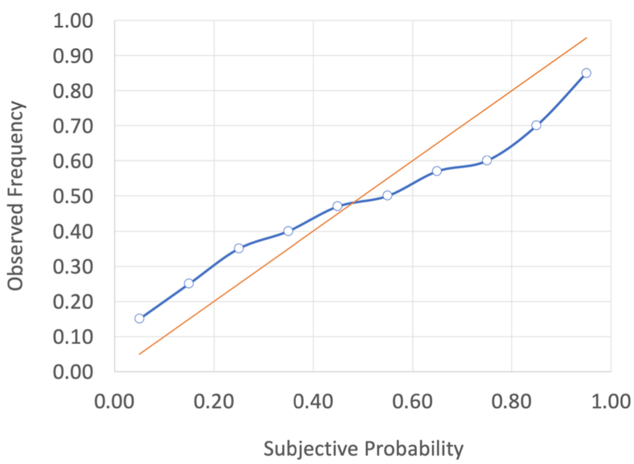

Figure 2. A personal calibration curve constructed by making repeated forecasts of events and recording the frequencies with which those events later occur.

4.3. Rare events

The necessity of expressing probabilities for rare events is a common challenge in risk analysis. These might be categorized in four sorts. The first involves statistical populations with a large denominator, *e.g.,* the background rate of failure of all modern dams. The second involves extreme values occurring at the tails of other distributions, *e.g.,* the occurrence of a large earthquake. The third involves events that can easily be subdivided into sub-events with more frequent probabilities, *e.g.,* the reliability of hydraulic gate systems. The fourth involves basic events for which there is no statistical history, that are not a tail-events, and that cannot be decomposed, *e.g.,* internal erosion events are arguably of this type although attempts have been made to transpose them into the other types.

An obvious difficulty with assigning probabilities to rare events is that they do not happen often, and most people including subject matter experts have little intuition for such low numbers. Human subjects seem to have reasonably calibrated intuition about probabilities in the range [0.1,0.9] (Fischhoff, Slovic, and Lichtenstein 1977), but outside that range our abilities wane. None of us—expert or novice—has an intuitive sense for probabilities in the range, say, $p \leq 0.01$ as frequently appear in dam safety studies, and no one should seriously accept estimates of native probabilities much smaller than that. Another difficulty with probabilities for rare events is that they cannot be validated because they seldom if ever happen. This has been a recurring problem in assigning probabilities to malicious attacks.

Consider that we wish to test the hypothesis that the prediction offered by an expert on the *probability of some rare event* is correct, against the alternate hypothesis that the probability is something different from what the expert assigned. Set the hypotheses, $H_0 : p = \hat{p}$ and $H_a : p \neq \hat{p}$, in which $\hat{p}$ is the expert's assigned value of the probability of the rare event. The test hypothesis is that there is no difference between the expert's probability and the population of real dams. The alternative hypothesis that the probability is something other than $\hat{p}$. What is our confidence in the expert if we observe the rare event?

Consider that we employ an expert to assign a probability to a rare event that is part of an event tree for a dam safety risk analysis. This might be, for example, the existence of an undetected discontinuity in an abutment (as in the example above). The expert considers the problem and returns with an assessment, "the chance of this rare event, in my opinion, is one-in-a-thousand." We take this to mean, $\hat{p} = 10^{-3}$.

Since this is an eminent consultant, we assign a high prior confidence to his opinion. Say, we assign a probability to the expert being correct of, $\Pr(H_0) = 0.99$. Now, the rare event is found to obtain. Our posterior probability in the expert's assignment by Bayes Rule reduces to 0.17. Our faith in the expert's assignment has dropped from 0.99 to 0.17, as should be expected because he or she was wrong. On the other hand, for the situation in which no occurrence is observed, the posterior probability in the expert's assignment increases slightly from 0.99 to 0.995. If the rare event occurs, the expert's credibility is profoundly impaired; but if it does not, the expert's credibility remains more or less unchanged.

4.4. Cromwell's rule and unanticipated data

Dennis Lindley (1985) invented this rule in remembrance of Cromwell's plea to the Commissioners of the Kirk of Scotland in 1650, 'I beseech you, in the bowels of Christ, consider it possible you may be mistaken'. The rule says, 'it is inadvisable to attach probabilities of zero to uncertain events, for if the prior probability is zero so is the posterior, whatever be the data'. The intention is to protect against unanticipated data. A prior probability of 1.0 is similarly inadvisable. The rule is obvious, and it may be that in the face of unanticipated data the practitioner simply throws out the prior, yet it is reasonable advice. That said, it is poor practice to hunt for priors that generate the most pleasing posterior distribution.

5. POORLY-INFORMATIVE PRIORS

In practical geotechnical settings with sparce data, both the prior and Likelihood are influential on posterior probabilities. The importance of the prior spans from weak to strong. A weak prior is one dominated by the Likelihood. A strong prior is one which dominates the Likelihood.

Ehrenberg (1986) noted the dilemma: we can have 'a weak […] prior, in which case why bother; or a strong prior, in which case why collect new data?' Presuming the priors to be valid, poorly-informed priors create their own challenges.

5.1. Unforeseen conditions

The validity of the prior is essential when unforeseen conditions arise. The decision to go one way or other rests on the prior which may be the only information upon which to decide before observing field performance. Differences in opinions at this stage can be problematic. An example illustrates issues that arise in practice (Hartford, 1994).

Figure 3. Pier K Foundation, Malibamatso River Bridge, Lesotho. (Hartford, 1994).

An inherently stiff incrementally launched concrete bridge deck was to be constructed across a valley where there is a large height differential between a short pier founded on the top of the cliff and a high pier founded in the valley floor. The bridge deck was to be launched from the cliff side, across the short pier towards the tall pier. The foundation of the shortest pier was unconfined on the cliff face and had to withstand the deck launching forces as well as long-term vehicle loads. The early investigation failed to identify a deep-seated, critically oriented failure surface. However, stereographic analysis pointed to an exfoliation discontinuity cluster with kinematically unfavourable orientation for shallow plane failure.

Two priors with different implications arose from the borehole data. One favoured doing nothing, while the other favoured intervention. Further study was not an option as the bridge was under construction. The complexity of the decision involved: a) how well calibrated the experts were, and b) how the decision-maker perceived this calibration. The experts' priors when transformed to the decision-maker's priors do not retain their uniqueness.

Whether the prior is in fact separated from the consequences of failure also becomes an issue, although from a risk perspective it should not be. During launching, deformations that exceeded design threshold were twice observed. Work stopped and construction was altered to mitigate the effects before proceeding on a "see what happens" basis. Data obtained during drilling the foundation improvements permitted updating of the priors, and a Monte Carlo simulation led to the conclusion that the probability of failure was 'remote'.

Predictions in site investigation and geotechnical construction involve feedback concerning the accuracy of the predictions when the "hypothesis" is tested in the field. This points to the need for predictions related to what will happen, rather than vague statements of possibly or probably (Lambe 1973). Decisions are definitive, probabilities in contrast are often impossible to validate.

5.2. Internal erosion

The above example illustrates different experts with different beliefs. The "rightness" of the beliefs remained unresolved. The subsequent case of *Coursier Lake Dam* involved internal erosion in an earth dam already experiencing distress (BCHydro, 1995). Two groups of experts were engaged. Group A comprised specialists in design and construction. Group B comprised specialists in monitoring and field data interpretation.

Four potential failure mechanisms were identified, and a common event tree of binary nodes was developed reflecting the common beliefs of both groups. Consensus probabilities were assigned during a workshop (Table 1). The overall results were in general agreement, as a factor of two in geotechnical reliability was deemed inconsequential. Reliance on a common event tree presumably contributed to the similarity. However, both results were alarmingly high.

Upon investigation, the two Groups had adopted different diagnostic strategies when interpreting the same information. Group A emphasized causal reasoning from borehole logs, field conditions, and laboratory data. Group B emphasized instrumentation records, response times, and trends. Of the 18 paths to failure in the event tree, just four were assigned similar probabilities. As an example, the groups disagreed completely on the probability that internal erosion in the foundation would occur. Group A considered it *very unlikely* ($p = 0.1$); Group B considered it *very likely* ($p = 0.9$).

Figure 4. Coursier Lake Dam and reservoir after decommissioning (courtesy BC Hydro).

The embankment was eventually repaired and returned to service. Shortly after reaching full reservoir, performance concerns again emerged to reveal a new sinkhole outside the repaired section. Forensic deconstruction revealed previously unknown features which controlled the performance of the structure. Poorly-informative priors can generate similar prior probabilities, but for reasons which are exceedingly different. Later updating by Bayes Rule must be influenced not only by the probabilities, but also by the causal reasoning leading to those probabilities.

Table 1. Consensus probabilities of event tree nodes by the two expert groups.

Failure mechanism	Group A	Group B
Erosion of embankment	0.49	0.23
Erosion of foundation	0.13	0.13
Piping around outlet	0.08	0.03
Piping into drains	0.002	0.01
Total probability of failure	0.70	0.40

6. CONCLUSION

Bayesian approaches to data analysis and risk modeling in geotechnical practice are reasonably accepted by specialists. Rapid advances in the use of Bayesian models are clearly underway. The present purpose is not to disparage that progress. The widespread use of reference priors has been useful in research but is less useful on practical projects. In the field, data are sparse and the poorly-informed priors based on sometimes unenumerated expert opinions are important. The capacity of Bayesian methods for fusing priors with other qualitative information and with data from multiple sensor types is among the strengths of the approach. It is important that these priors be valid and calibrated. Attention to the issues of poorly-informed priors in practical problems is an important frontier for continuing development of the Bayesian approach.

7. ACKNOWLEDGMENT

An earlier version of the manuscript for this paper was co-authored by the late John T. Christian. The authors are indebted to John for his insights and good spirit. We also acknowledge the support of BCHydro in allowing use of the Coursier Lake Dam case story.

8. REFERENCES

Baecher, G.B. 2019. “Putting Numbers on Geotechnical Judgment.” Buchanan Lecture, Texas A&M University. Available at pm.umd.edu/person/gregory-baecher/

BCHydro (1995). Coursier Lake Dam Deficiency Investigation. Supplement to the Deficiency Investigation. 1st Estimate of the Probability of Failure by Piping. BC Hydro Report MEP-11. April.

Box, George E. P., and George C. Tiao. 1973. *Bayesian Inference in Statistical Analysis*. 1 edition. Reading, Mass.: Addison-Wesley Publishing.

Ehrenberg, A. S. C. 1986. “Discussion of, ‘Bayesian Methods in Practice’.” *Applied Statistics* 35: 135–36.

Fischhoff, B., P. Slovic, and S. Lichtenstein. 1977. “Knowing with Certainty.” *J. Exp. Psych,* 3 (4): 552–64.

Gilbert, R.B., M. Habibi, and F. Nadim. 2016. “Accounting for Unknown Unknowns in Managing Multi-Hazard Risks.” In *Multi-Hazard Approaches to Civil Infrastructure Engineering*, edited by P. Gardoni and J.M. LaFave, 383–412. Cham: Springer International Publishing.

Good, I. J. 1983. *Good Thinking: The Foundations of Probability and Its Applications*. Minneapolis: University of Minnesota Press.

Hacking, Ian. 2001. *An Introduction to Probability and Inductive Logic*. New York: Cambridge University Press.

Hartford, DND. (1994). *Geotechnical Aspects of the Construction of the Northern Access Road (Pitseng to Katse Intake) LHDA Contract No. 103"* by D.N.D. Hartford. In Blight, G.E. et al (eds.) *Geotechnics in the African Environment Vol. 2* Balkema, 1994, pp 475 – 516.

Jeffreys, Harold. 1931. *Scientific Inference*. Cambridge: The University press.

______________. 1939. *Theory of Probability*. The International Series of Monographs on Physics. Oxford: The Clarendon Press.

Kass, Robert E., and Larry Wasserman. 1996. “The Selection of Prior Distributions by Formal Rules.” *Journal of the American Statistical Association* 91 (435): 1343.

Lambe, T. W. 1973. “Predictions in Soil Engineering.” *Géotechnique* 23 (2): 151–202.

Lindley, D. V. 1985. *Making Decisions*. 2nd ed. London; New York: Wiley.

Morris, Peter A. 1974. “Decision Analysis Expert Use.” *Management Science* 20 (9): 1233–41.

Tang, Wilson H. 1987. “Updating Anomaly Statistics-Single Anomaly Case.” *Structural Safety* 4: 151–63.

Zabell, S. L. 1992. “R. A. Fisher and Fiducial Argument.” *Statistical Science* 7 (3): 369–87.