Why is there Math in my Archaeology?

The modern foundations of quantitative archaeology, written decades too soon

Introduction

Fifty years ago, what was arguably one of the most important papers written for modern work in quantitative archaeology was published in American Antiquity. Unfortunately for its author, and generations of archaeologists, it received relatively little attention. With a small number of citations, more than half of which have occurred in just the last few years, its elegance and mathematical precision went largely unappreciated – even by the growing cohorts of computational and quantitative archaeologists whose work would have greatly benefited from John's brilliant work.

John Justeson's 1973 article "Limitations of Archaeological Inference" was not only correct, even if the field at large hadn't realized it, but also still very much at the forefront of digital archaeology.

LIMITATIONS OF ARCHAEOLOGICAL INFERENCE: AN INFORMATION-THEORETIC APPROACH WITH APPLICATIONS IN METHODOLOGY

JOHN S. JUSTESON

ABSTRACT

A framework is established for the application of information-theoretic concepts to the study of archaeological inference, ultimately to provide an estimate of the degree to which archaeologists, or anthropologists in general, can provide legitimate answers to the questions they investigate. Particular information-theoretic measures are applied to the design elements on the ceramics of a southwestern pueblo to show the methodological utility of information theory in helping to reach closer to that limit.

Department of Anthropology
University of California, Berkeley
September, 1971

Figure 1: Abstract of Justeson (1973) "The Limitations of Archaeological Inference"

The premise was actually quite straightforward – behavioral information is "encoded" in the material artifacts deposited within an archaeological site, and the archaeologist's motive is to "decode" that information on the other end.

The novelty was that John saw this "encoding-decoding" process as an information flow, which could be described by what was (at the time) a relatively esoteric set of mathematical tools known as *information theory*.

The foundations of information theory were developed by Shannon (1948) as a way to analyze the transmission of information *independently* of the content of the message.

> "The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have meaning; that is they refer to or are correlated according to some system with certain physical or conceptual entities." (Shannon 1948, 1)

John saw that this theory established a quantifiable "upper limit" on how interpretable archaeological data could be. We could actually *calculate* the limits of archeological inference.

> "If the empirically measured parameters are not consistent with the relationship between them that is required by the theory for a given material or behavioral system, then the data by which that system is to be interpreted cannot have a consistent susceptibility to decoding; that is, there will be no basis for deriving a coherent archaeological interpretation of the data that will accurately reflect the prehistoric situation." (Justeson 1973, 136)

In particular, he was addressing two *inherent* limitations of the archaeological record:

1. limitations imposed by the degree of preservation of culturally significant remains and by the skewing of their relationships through time until their recovery; and
2. limitations on the interpretability of archaeological data for the cultural descriptions.

The first limitation is analogous degradation of a signal due to noise or interference affecting a transmission, and the second to the encoding and decoding of that signal between sender and receiver.

Information Theory

At the heart of information theory is a single equation that aimed to quantify a measure of information based on the concept of *entropy* from the thermodynamics of physical systems:

$$H(\mathcal{X}) = -\sum_{\rangle=\infty}^{\backslash} \surd^{(\S_\rangle)} \log_{\in} \surd^{(\S_\rangle)}$$

The total information entropy H of a system $\mathcal{X}$, which contains some number of discrete features or attributes $(x_1, x_2, \ldots x_n)$, is defined as the negative sum over all features of each attribute's probability of occurrence $p(x_i)$ times the log of its probability.

What this does, although not immediately obvious to most of us, is to tell us the minimum number of "events" of that system that it would take before we could start detecting a pattern. The more events it would take, the less information each observation is actually giving us.

One easy way to think of it is that, for a high-entropy event we can't reasonably predict any individual occurrence of a completely random event – i.e., *any* outcome is equally likely – so we are likely to be *surprised* each time.

After *a lot* of observations we could make a fair prediction of the outcomes over a *large* number of events, but still couldn't accurately predict any single event. Each event gives us only a *small* amount of information, so we would need a large number of observations before we could distinguish it from some other system.

Conversely, it wouldn't take us too long to notice something that regularly (or never) occurs. Each event would tell us a *lot* of information, so we need fewer observations to see a pattern.

Archaeological Information

What John did was to apply this concept of how "surprising" information concerning the "event" of an artifact or its attributes would be, given a certain frequency of occurrence. By considering the archaeological record as a transmission channel, one could reasonably compute the integrity of the archaeological "signal" based on the coherency of the codes received.

Consider the artifacts, features, or attributes as encoding the meaningful categories as would be recognized by the archaeological society. On the other end, we (as archaeologists) establish a system of "decoding" that signal through our categorization and classification schema. The limits of inference, then, would be to determine the translatability between encoding and decoding.

Channels, Classification, and Signal

Most are familiar with the now canonical work in Schiffer (1987) "Formation Processes of the Archaeological Record" that introduced the idea of various transformations of the archaeological record. Few, however, are likely aware that these transformations were in fact already quantified more than a decade earlier in John's paper.

The analogy between the archaeological record and a communication channel was a somewhat recent fascination of the "New" archaeology, but John took that analogy to its logical limits by formally applying the mathematics Shannon's information theory to the problem – something that had only recently been attempted and, at the time of its writing, had only one published example (i.e., Gorman 1970). As John noted at the time, information theory *concepts* had been introduced but not the actual mathematical formalisms.

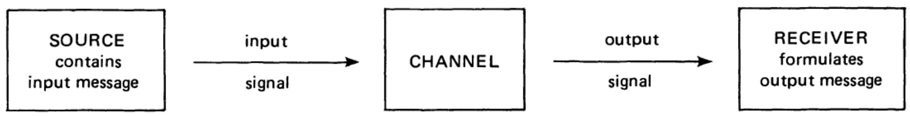

Figure 2: Schematic representation of information transmission (Fig. 1 Justeson 1973, 133).

$$C = \max_{\pi} \left\{ \sum_{j} \left[\sum_{i} \pi_i w(j|i) \log_2 \sum_{i} \pi_i w(j|i) - \sum_{i} \pi_i w(j|i) \log_2 \sum_{i} \pi_i w(j|i) \right] \right\}$$

Applications

$$M_t = \sum_{i=t}^{T} N_i$$

$$L_t = \sum_{i=t}^{T} M_i$$

References Cited

Gorman, Frederick. 1970. "Symbol as Strategy: Toward a Synthetic View of Material Culture." Field Museum of Natural History.

Justeson, John S. 1973. "Limitations of Archaeological Inference: An Information-Theoretic Approach with Applications in Methodology." *American Antiquity* 38 (2): 131–49. https://doi.org/10.2307/279360.

Schiffer, Michael B. 1987. *Formation Processes of the Archaeological Record*. University of New Mexico Press.

Shannon, Claude E. 1948. "A Mathematical Theory of Communication." *Bell System Technical Journal* 27 (4): 623–56. https://doi.org/10.1002/j.1538-7305.1948.tb00917.x.