

Empiricisms of the Baroque

computing the folds

The likelihood of what happens in our probable lives: Baroque empiricism in probabilistic modelling What probably happens The curve of the world:

Adrian Mackenzie Cesagen/Department of Sociology, Lancaster University
a.mackenzie@lancaster.ac.uk

Abstract

Along with an investment in intricate machinery, artificiality, automation and calculation, contemporary power formations inherit probability from the Baroque. As Ian Hacking writes, ‘there was no concept of probability until about 1660’ (Hacking, 1975). Probability as a numbered relation to what happens is a Baroque invention. Whether in insurance risk calculation, population or infection modelling, high-frequency trades, management of fish stocks, forensic evidence, estimates of genetic susceptibility or email spam filtering (and this list could be much more extended), many contemporary forms of power work with probability-based numbers. Probabilities, likelihoods and expected values show no sign of abating amidst the growth of data associated with digital devices. It would be plausible to argue that we are actually experiencing an intensified relation to probability in settings where different flows of data are brought together. Computations of probability are being augmented and distributed into many different data-driven processes ranging from mundane to supra-mundane. The social sciences are no exception to this since most forms of calculation and modelling used in social sciences depend on probability. But if it is possible and desirable to work empirically on probabilistic processes and their shaping of events, where would we turn? The case study in this paper is a statistical modelling technique called Markov Chain Monte Carlo (MCMC). MCMC, a way of simulating samples of data from complicated probability distributions, is used widely across social and natural sciences, and in many domains of government, commerce and industry. The analysis of MCMC supports a heightened awareness of the heterogeneous artefactuality of probability. In MCMC, it is possible to follow a faultline that has run through the calculations of probability since its 17th century inception. Along this faultline, probability wavers between generating numbers that index something happening in the world and numbers that index our belief about the world. While philosophers and statisticians have often been passionately divided over subjective and objective renderings of probability, the mixtures of event and belief found in MCMC is more interestingly convoluted. In tracking how MCMC simulates samples in order to make sense of observed data, can we intensify our sense of what happens?

Introduction

But the essence of the Baroque entails neither falling into nor emerging from illusion but rather realizing something in illusion itself. (Deleuze, 1993, 125)

In an essay on how Gilles Deleuze's notion of the fold might inform the social studies of science and technology, Geoff Bowker argues that what Deleuze's work urges us to do is 'is to take the impossible to be true and see what can be done with it' (Bowker, 2010, 125). Like Deleuze's claim that the essence of the Baroque lies in 'realizing something in illusion itself' (Deleuze, 1993, 125), Bowker's injunction to take the impossible 'to be true' suggests that we might treat the Baroque as a sensibility alive to impossibility, aware of illusion, and yet empirically, practically inclined to make something of that. In a prosaic sense, the impossible, as that which cannot occur or happen, was a matter of interest in the Baroque invention of concepts and practices of probability.

If we seek to understand probability in terms of a Baroque sensibility, two possible paths are open for investigation. Probability has a historical lineage leading back into the Baroque. In his history of chance, Ian Hacking writes, 'there was no concept of probability until about 1660' (Hacking, 1990). Probability as a numbered relation to what happens is a largely Baroque invention (associated with Blaise Pascal, G.W.F Leibniz, etc). Probability is also an ongoing problem, played out in many forms, concerned with how we make sense of what happens. In the three or so centuries since then, probability has taken on an extraordinary power in many social settings. Whether in insurance risk calculation, particle physics, population or infection modelling, high-frequency trades, management of fish stocks, forensic evidence, image processing, estimates of genetic susceptibility or email spam filtering (and this list could be vastly more extended), many contemporary styles of power count on notions of probability flowing from Baroque thought. Along with an investment in ornate facades, indirectly light interiors, intricate machinery, artificiality, automation, number and calculation, contemporary power formations inherit probability from the Baroque.

Regardless of this historical lineage, there is a second way we might approach numbers via the Baroque. The Baroque, writes Gilles Deleuze, 'refers not to an essence but to an operative function' (Deleuze, 1993,3). Probability is an 'operative function' and arguably increasingly important operative function concerned with events. As an operative function, probability attributes numberings to what happens. While Deleuze has almost nothing to say about probability in *The Fold: Leibniz and the Baroque*, he has much to say about number. I would suggest that we are actually experiencing an intensified relation to probability in many settings, and this affects, as it folds back in potentially quite profound ways, our sense of what happens. Recent transformations in probability are intimately associated with computational practice, and in particular with the pressing problem of what we do with more data. What happens as we

get more data (as in ‘big data,’ in the ‘data deluge,’ etc.)? As various authors have recently argued (Silver, 2012), more data brings with it much more uncertainty. More data widens margins of uncertainty because it fosters ambitions to know what happens much more widely and comprehensively. This entails, in turn, many forms of assimilation, augmentation and imputation of diverse data flows. Massively available data actually makes it harder to find out what happens because we encounter constantly changing numbers on various scales. These numbers have to be constantly updated and brought into relation with other numbers. Furthermore, out of the myriad of available numbers, representative, significant or important tendencies have to be isolated and highlighted. All of this compels recourse to many numbers of the probability kind. Rather than more data diminishing chance, in a certain sense, as we will see, it tremendously intensifies recourse to chance. In recent years, practice of inference based on the generation of massive supplies of random numbers are being extended and ramified into many different data-driven processes ranging from mundane to supra-mundane, from banal to exceptional. As we will see, the social sciences are no exception to this since many of the forms of statistical modelling and testing used in social sciences to address problems of one and many, part and whole, of belonging, relationality and inclusion, depend on shaping flows of computer-generated random numbers.

Number has long had an ambivalent reputation in many parts of the social sciences, and especially in strands of European social and critical theory. For complicated reasons, prominent voices in social and philosophical thought have decried numbering, calculation or mathematisation. Critical theory objected to mathematical thought in its alliance with instrumental reason (Habermas, 1970), classic phenomenological re-appraisals voiced by Edmund Husserl (Husserl, 1970) or Martin Heidegger (Heidegger & Lovitt, 1977; Elden, 2006) saw number and calculation as covering-over more immediate and primary modalities of experience; or, to take a quite different example, deconstructionist thinkers such as Jacques Derrida or Bernard Stiegler (Derrida & Stiegler, 2002; Stiegler, 1998) regarded the entwining of number with technology as provoking disruptive transformations (in memory, subjectivity, politics and economy). In social studies of science and technology from the 1970s onwards, number and calculation associated with sciences and technology have also received much attention, in order to identify how numbers, models, and calculations are shaped by social interests of various kinds (Mackenzie, 1990; Porter, 1996; Desrosieres, 1998; Holmberg, Bischof & Bauer, 2012), but usually in the interests of centralised forms of power such as states, commerce or industry. More generally, social sciences have both relied on quantification. Finally, researchers in sociology, anthropology history, media studies and other social sciences-humanities disciplines have appraised numbers from a wide range of different angles, in order to show how numbers and quantification develop, mutate, hybridize, converge with or destabilise various domains of knowledge and practice science, industry, government, and everyday life. The work on numbers in relation to risk, crime, disease, land, climate, economics,

biodiversity, market exchange, audit cultures, finance or state power is ongoing (see (Espeland & Stevens, 2008) for a broad description).

To transit a century of social and philosophical thought on number so rapidly inevitably foreshortens any sense of the differences that run across this work. Yet it highlights too that the problem of how to live with numbers, and how to include numbers remains troublesome. More recent attempts to re-kindle an affirmative relation to number, for instance, in the work of Alain Badiou (Badiou, 2008), in the re-activation of Tardean sociology (Barry & Thrift, 2007; Latour, Jensen, Venturini, et al., 2012), or Helen Verran's work (Verran, 2011) offer a very different orientation. For Badiou, number and associated formalisms such as mathematical set theory provide vital resources for philosophical invention. Mathematical thought indelibly remains, for Badiou, ontologically primary. A different rapprochement with number can be found in Tarde's monadology or Verran's recent work on enumerated entities (Verran, 2011). The term 'enumerated entity,' she writes, '(rather than merely number) is a way of insisting on the importance of locating number' in relation to the work of naturalizing socially constructed categories (66). Seeking a way between the naturalization of numbers as counting what is, and the rejection of numbers as overtly designed or fabricated, Verran suggests that we with particular forms of number such as indices that suture different realities to each other in many everyday settings.

A Baroque treatment of number, and in this case, a certain form of number, probability, offers another alternative. It radicalises number in several ways. As Deleuze writes, in Baroque mathematics, 'the object of the discipline is a 'new affection' of variable sizes, which is variation itself' (Deleuze, 1993, ??). While Deleuze, following the polymathic and inordinately inventive swerves of Leibniz's work, largely presents variation in terms of curves and functions, almost everything he will say about curves directly concerns variations in numbers. In contrast to geometrical treatments where numbers run in lines to infinity, the curved lines, surfaces or folds that Deleuze discusses in such detail, are less easy to analyse, predict or control. Variation is intrinsic to them, and these variations can be seen in a proliferation of series, inflections, and infinitesimal differences. Working with variations, I suggest, not only occasions the production of many numbers, it also opens up some different ways of living amidst many numbers.

Number in variation is radicalised in Leibniz's famously impossible maxim that each monad includes the whole world. It is through numbers, understood in a somewhat unorthodox way, that monads — the parts of the world — can include the whole world. This is an especially slippery point in Deleuze's analysis of Leibniz's work, but one that has strong contemporary resonances in many settings. Deleuze says: 'for Leibniz, the monad is clearly the most "simple" number, that is, the inverse, reciprocal, harmonic number. ... The inverse number has special traits: it is infinite or infinitely small, but also, by opposition to the natural number, which is collective, it is individual and distributive' (Deleuze, 1993, 129). These kinds of philosophical formulations are rather indigestibly remote from much social theory. For instance, what is a 'distributive

number’, and how does it differ from a collective number? I will argue below that probabilities are superlative distributive numbers. The very possibility and impossibility of making the numbers into the root, or the radix of the monad, seems to me quite important if we are interested in developing ways of engaging affirmatively with number, without granting it the pure ontological primacy of mathesis or exiling it to the badlands of alienated reason.

To say that having a world — for the monad is a mode of having a world by including it — is a number is to suggest either a very different notion of having or a different notion of number (‘inverse, reciprocal, harmonic’), or some combination of both. Support for all three possibilities can be found in Deleuze’s Leibniz. For instance, Deleuze writes that ‘the painting-window [of Renaissance perspective] is replaced by tabulation, the grid on which lines, numbers and changing characters are inscribed. ... Leibniz’s monad would be just a such grid’ (27). This suggests a different notion of the subject, no longer the subject of the world-view, the subject as locus of reason, experience or intentionality, but as ‘the truth of a variation’ (20) played out in numbers and characters tabulated on gridded screens. We might think of border control officers viewing numerical, predictions of whether a passenger arriving on a flight is likely to present a security risk (Amoore, 2009), financial traders viewing changing prices on their screens (Knorr-Cetina & Bruegger, 2002), a genomic research deciding whether the alignment scores between two different DNA sequences suggests a phylogenetic relationship, or a player in a large online multiplayer games such as World of Warcraft quickly checking the fatigue levels of their character before deciding what to do, as all typical cases where numbers in variation populate the monadic grid. But there is also a shift in the understanding of number here. In relation to the monad as inverse number, Deleuze writes ‘the inverse number has special traits: it is infinite or infinitely small, but also, by opposition to the natural number, which is collective, it is individual and distributive’ (129). Here too, the suggestion that numbers possess traits such as being ‘individual and distributive’ hints at a different conception of numbers, and a supple line along which we might track numbers through experience, and across coarse distinctions between subject-object, nature-culture, self-other. Like the tabular numerical grid of experience, the distributive character of number resonates with our situation in which there are many numbers operating, but in which the potential of numbers to do something seems increasingly inchoate and uncertain.

Markov Chain Monte Carlo

While it departs from a typical Baroque event — the invention of probability — the focus here is on how contemporary probability has become entwined with a particular mode of computation. This paper will not in any way trace the complicated emergence of probability and its development in various statistical approaches to knowing or deciding. Historians of statistics have documented this in great detail, and tracked how statistics is implicated in power-knowledge in

various settings (MacKenzie, 1978; Stigler, 1986; Hacking, 1990; Daston, 1994; Porter, 1996). My concern here is the problem of invention of forms of thought able to critically affirm probability. This problem arises, I suggest, in many, perhaps all, contemporary settings where population is implicated. In seeking to unfold ways of thinking probability for social theory, risks of scientism or scientocentrism abound. On this score, a Baroque sense of what happens offers at least tentative pointers to a different way of thinking about what is likely to happen.

There are very many ways we might explore contemporary forms of the individual, distributive, and somehow infinite number. In this paper, I focus on a single statistical simulation technique called MCMC - Markov Chain Monte Carlo simulation. MCMC is sometimes called algorithm: a series of operations that transform or reshape something. Actually, MCMC has been called one of the most influential algorithms in twentieth century science and engineering (Andrieu, De Freitas, Doucet, et al., 2003, 5). But MCMC is not really an algorithm, or at least, if it is, it is an algorithm subject to various algorithmic implementations (for instance, Metropolis-Hastings MCMC; Gibbs Sampler MCMC are just implementations). Invented during the 1950s, the MCMC technique is important in contemporary statistics, and especially in Bayesian statistics. It plays significant roles in areas such as images, speech and audio processing, computer vision, computer graphics, molecular biology and genomics, robotics, decision theory and information retrieval (Andrieu, De Freitas, Doucet, et al., 2003, 37-38). In all cases, MCMC is a way of simulating a sample of points on a complicated curve or surface in order to estimate the area or volume that it encompasses. What MCMC has added to the world is subtle yet indicative. The statisticians Christian Robert and George Casella write that MCMC ‘Markov chain Monte Carlo changed our emphasis from “closed form” solutions to algorithms, expanded our impact to solving “real” applied problems and to improving numerical algorithms using statistical ideas, and led us into a world where “exact” now means “simulated”’ (Robert & Casella, 2008, 18). This shift from ‘closed form’ solution to algorithms and a world where ‘exact means simulated’ might be all too easily framed by a post-modern sensibility as another example of the primacy of the simulacra over the original. But here, a Baroque sensibility might allow us to approach MCMC less adversely. MCMC is an example of the kinds of operations that we might associate with individual, distributive number. In MCMC, shifting paths, curved movement, convergence and divergence, and especially, convoluted forms, are encountered everywhere. Rather than the infinite, straight lines of geometry, or the crowds of points or particles that can be statistically treated en-masse, MCMC techniques traverse a middle ground where the closed form analysis nor statistical aggregates work easily. In this world, vertiginous architectures attempt to create convergent beliefs (as do the iconic examples of counter-Reformation religious architecture found especially in Italy), yet that convergence is fabricated dynamically precisely because there is no straight line along which belief can run (for instance, between earth and heaven).

The MCMC technique addresses the problem of what to do with very uneven or folded distributions of numbers. It is a way of calculating areas or volumes whose curves, convolutions and hidden recesses elude perception. Accounts of MCMC emphasise the high-dimensional spaces in which the algorithm works: ‘there are several high-dimensional problems, such as computing the volume of a convex body in d dimensions, for which MCMC simulation is the only known general approach for providing a solution within a reasonable time’ (Andrieu, De Freitas, Doucet, et al., 2003,5). Indeed, we could say that MCMC increasingly facilitates the fabrication of high-dimensional, convoluted data spaces. As an algorithmic technique, MCMC distils into computer calculations a set of practices relating to number that are very intricately interwoven with the history of sciences such as physics and statistics. (This poses problems for reading, because like Leibniz’s writings, the texts and documents describing MCMC are replete with mathematical and scientific references that are not immediately familiar.) The numbers that MCMC works on are found in particularly kinds of scientific settings where there are many variables. The ‘high-dimensional’ spaces referred to above by the MCMC algorithms were devised by physicists working at Los Alamos in the 1940s (Metropolis, 1949)}, generalised by statisticians (Hastings, 1970)}, taken up by ecologists working on spatial interactions in plant communities during the 1970s (Besag, 1974), by computer scientists working on image reconstruction (Geman & Geman, 1984), and then subsequently seized on again by statisticians in the early 1990s (Gelfand & Smith, 1990). In the 1990s, it became clear that the algorithm could make Bayesian inference — a style of statistical reasoning that differs substantially from mainstream statistics in its treatment of probability (Mcgrayne, 2011) — practically useable in many situations. A vast, still continuing, expansion of Bayesian statistics followed, nearly all of which relied on MCMC in some form or other. (Thompson Reuters Web of Knowledge shows 6 publications on MCMC in 1990, but over 1000 a year for the last five years in areas ranging from agricultural economics to zoology, from wind-power capacity prediction to modelling the decline of lesser sand eels in the North Sea; similarly NCBI Pubmed lists close to 4000 MCMC-related publications since 1990 in biomedical and life sciences, ranging from classification of new-born babies EEGs to within-farm transmission of foot and mouth disease; searches on Bayesian methods lead to even more results). In the social sciences too, political scientists regularly use MCMC in their work (Gill, 2011) because their research terrain — elections, opinions, voting patterns — little resembles the image of events projected by mainstream statistics: independent, identically distributed (‘iid’) events. When brought together with Bayesian inference, MCMC allows, as the political scientist Jeff Gill observes, all unknown quantities to be ‘treated probabilistically’ (Gill, 2011,1).

This treatment of all things — the model, the parameters, the results, the missing data — in terms of probability seems especially significant in relation to data practices in contemporary sciences and network media, but it also bears in vaguely defined ways on the many different aspects of contemporary lives to the extent that they are shaped, tinged or shaded by probabilistic numbers. In all

of these settings, what it might mean to include the world as a distributive number? My question here, based partly on the exceptionally convergent popularity of MCMC in science and other domains, is what we might do with MCMC? In that it allows all numbers to be treated as probabilities, that is, as distributions of values, or as ‘random variables,’ MCMC bears within it a particularly rich and generative instance of probability practice. In its Bayesian applications, it retains the shifting and generative overlaps between ‘internal’ and ‘external’ evidence associated with Baroque probability. It also practically and very concretely engages with large flows of numbers in the form of data, with problems of convergence, and with problems of apprehending variations and fluctuations associated with convoluted or folded forms (the high-dimensional ‘convex volumes’ mentioned above). I will not trace in any great detail the diversity of applications of MCMC, especially since the so-called ‘Bayesian revolution’ of the 1990s (Mcgrayne, 2011). My interest here lies in the forms of movement, the treatments of variation, and the operations with number that take place in the algorithm. For the purposes of highlighting these movements, I leave aside the various social interests, and cultural and economic contexts associated with MCMC (for instance, its advent in an atomic weapons research laboratory at Los Alamos). While I draw on histories of statistics and other sciences, as well as on contemporary literature on MCMC and Bayesian statistics, this paper tentatively proposes that we should treat MCMC as a way of thinking about how to include the world, as an operation that effects quasi-monadic expression of a world, a world that we all inhabit to some degree or rather, at least in so far as we live in worlds increasingly cross-hatched by data distributions. Inasmuch as we inhabit worlds furnished with data-derived actions, signs, inscriptions and operations, we find ourselves inflected by probabilities and probability distributions. The shifts in agency associated with such numbers are not readily obvious. Tracking patterns of movement through algorithms such as MCMC offers at least one way to imagine participating differently in the making of numbers. Reading techniques like MCMC, I suggest, might help us understand what it means to inhabit numbers distributively.

Reading distributively

As we saw, Deleuze contrasts collective and distributive numbers. For Leibniz, monads are ‘individual and distributive numbers’ (Deleuze, 1993, 129), while natural numbers are ‘collective’. This term ‘distributive’ can be read in different ways (for instance, distributivity in mathematical operations describes the way they apply to the variables they operate on), but here I connect, or actually, equate it with probability. Typical textbook introductions to statistics often begin with explorations of the distribution of data. Only after introducing distributions — what values a variable takes and ‘how often’ it takes those values (Moore, 2009, 25) — will statistics textbooks begin to address questions of probability. Here they refer to random events such as rolls of the dice or hands of cards, etc, in order to begin to formalise descriptions of how often things occur

(typically, the rules of probability introduced include how to add and multiply independent and disjoint events (Moore, 2009, 330)). In mainstream statistics, variables take on different values because they are measured or observed more or less often (in surveys, in polls, in instrument readings, by sensors, in observations). The numbers that results from counting and measuring have nothing intrinsically random about them. Rather they way they are obtained through sampling brings randomness into play. Conversely, although the data may result from random sampling, the numbers themselves count what happens in the world. Statistics, at least in the textbook examples, treats probability as a way of describing variations in repeated measurements. Measurement is a practice affected by many sources of variation, error or noise, and probability comes into play as a way of registering the extent of that affect.

A Baroque genealogy of probability harmonises much more easily with the MCMC-based shift in statistical practice. Rather than chance events, the Baroque notion of probability includes credibility, or degrees of belief. According to Hacking's account, the emergence of the concept of probability can be seen in terms of an unstable convergence between 'external' and 'internal' evidence. For mid-17th century sciences, external evidence comes from outside the thing. Testimony, for instance, is external evidence, and is well understood already in the Renaissance (Hacking, 1975,34). Internal evidence consists in a 'thing pointing beyond itself' (34), and is specifically new to the 17th century. The model for internal evidence comes from the 'low sciences' of alchemy, geology, medicine and astrology, knowledges that cannot argue through demonstration as geometry does, but must rely on diagnoses, exploration or 'adventure' to form opinions (36). This new empiricism works with signs, especially 'natural signs.' Signs-as-evidence, as they inform a physician's diagnosis for instance, pass over into a more widespread conception of the sign as the basis of the empirical knowledge that will 'come to dominate European thought' (46). Natural signs are not conventional or conjectural; instead they are frequently observed in experience. In the ways they point beyond themselves, frequently observed signs lie somewhere between demonstration (the ideal of highly mathematical forms of knowing such as geometry and algebra that predate the Baroque) and testimony provided by a witness or evinced in a document. Only because the naturalness of such signs stems from their repetition does counting and number come central. During the second half of the 17th century, in the wake of this mutation in what counts as evidence, calculations of probability based on counting what happens can gain ground. Games of chance (lotteries, dices games, etc.) had long been subject to attempts at calculation, but now calculations of annuities based on expected lifespans become amenable to the new empiricism of countable sign-as-evidence. Interestingly, the emergence of a concept of probability was not dependent on games of chance. While Pascal and Huyghens, for instance, departed from games of chance, Leibniz, as Hacking reports, working in Germany, not knowing much about games of chance, developed an account of probability based on numerical degrees of belief derived from jurisprudence (89; Lorraine Daston develops a similar argument at greater length (Daston, 1988)).

This forked emergence of probability, split between something in the subject and something in the world continues to flow through subsequent developments, and arguably, displays fresh permutations today.

At its core, in as much as it refers to something that can be evidenced only by multiple signs, the notion of probability implies and perhaps generates accumulations of numbers. Consistent with its provenance in low-profile epistemic practices, probability expresses the limits or constraints in our power to know what is fully there. Conversely, it also imposes on us the need to repeat or replicate measurements many times in order to manage those limits, or to relax their constraints. Things vary in many ways, and events don't always happen in the same way. For instance, the height of adult individuals in human populations range widely. In comparing different national populations, for instance, how does one know the height of people in a given population? Only by measuring the heights of many individuals in the different populations, counting the different heights, and then finding ways of expressing how they are distributed between minimum and maximum values. The latter are the so-called population parameters, such as the mean, the variance, standard deviation, the mode, or the median. Population became perhaps the central operational conception of the totality or whole in statistical reasoning during the 17-20th centuries. Populations were countenanced through techniques of observation, sampling, measuring, and increasingly, inference, that sought to describe the distribution of different outcomes. At the same time, the notion of population was generalised well beyond its original referent, the population, to include any collection of measurable outcomes (for example, in the nineteenth century, physicists started treating matters — gases in particular — as populations of particles).

The problem of estimating the parameters of populations, which by definition cannot be measured directly but only estimated on the basis of sampling, triggers the construction of many different forms of statistical practice. At core, these practices all seek to estimate population parameters, and to describe the totality of outcomes in terms of a distribution of values. Statisticians have devised various tests to see whether data derived from observation or experiment fits a particular distribution, and then various data transformations so that skewed data more closely resembles a known distribution, as well as many different methods of fitting data to curves, and then evaluating the goodness-of-fit. The details of these techniques are less relevant here than the underpinning notion of the probability distribution that generates, it is assumed, the range of observed or measured outcomes. This commitment to an underlying distribution is captured in the well in the concept of the random variable. Whatever the differences between schools of statistical practice (R.A. Fisher's frequentists vs. Bayesian inference (Mcgrayne, 2011)), there is little disagreement on the underlying nature of variation. All variation in what happens can be understood in terms of random variables. The textbook definition of a random variable runs: 'a random variable is a mapping that assigns a real number to each outcome' (Wasserman, 2003,19). This simple definition links outcomes — what happens

— to numbers through a ‘mapping.’ Mapping is a form of one-to-one correspondence, usually expressed as a mathematical function. A random variable links events to numbers through functions.

Perhaps the most famous function or mapping is the normal or Gaussian distribution. Statistics uses dozens of different probability distributions express continuous and discrete variations to map outcomes to real numbers. There are so many probability distributions — normal (Gaussian), uniform, Cauchy exponential, gamma, beta, hypergeometric, binomial, Poisson, chi-squared, Boltzmann-Gibbs distributions, etc (see [nist_2012] for a gallery of distributions) — because outcomes occur in widely differing patterns. The queuing times at airport check-ins do not, for instance, easily fit a normal distribution. Queues are usually modelled using a Poisson distribution, which unfortunately, distributes waiting times very differently. While knowledge of how to mathematically express the probabilities associated with particular outcomes has accreted over several centuries, they all share a common purpose: to express the distribution of outcomes found in some population (e.g. height, age, gender, etc). The diverse range of probability distributions — and we will see below some reasons why we can expect them to proliferate in certain settings — attests to the way in which populations resist measure.

These practices of defining the mapping real numbers (numbers that can take any continuous value) all focus on the shape of the probability distribution(s). One of the most important mathematical descriptions in any data analysis setting will be the probability density function (for continuously varying quantities) or the probability mass function (for variables that have discrete values; for instance: 1,2,3,4,5). The *probability density function* (pdf) is a function, usually graphed as a curve, that describes how likely a variable is to take on a particular value. In many cases, statistical practice seeks to estimate distribution functions such as pdfs (or their close relatives, cdfs — *cumulative distribution functions*) for the given data. The underlying probability distribution is ‘unobservable,’ but assumed to give rise to all the data gathered through experiments and observations. The task is to estimate the shape of that curve, and its parameters (means, variance, etc.). Given that curve, areas under the pdf equate to the likely range of value of a variable. While the total area under the probability density function curve always must be equal to one (since the probability of each individual outcome ranges between 0 and 1), finding the area under particular parts of the curve is a key issue. Scientists, economists, and engineers constantly use and re-use standard distributions such as Gaussian, exponential, Poisson, or Beta partly because these functions distribute numbers in ways that can be calculated more easily. The contours of a curve — its slopes, peaks and troughs — not only describe the distribution of values of relevant variable, they afford a way of treating a population as a function that generates outcomes. Hence, finding the area under probability density or mass function curves becomes the way in which many epistemic processes link mathematical functions to lived states of affairs (populations).

From the perspective of a Baroque sensibility, the most striking feature of the way in which the concept of probability has unfolded into a variety of probability distribution functions would be less the sheer proliferation of mathematical functions for working with probability distributions than the way in which the different shapes, areas, densities and masses of probability distributions could be used to support estimations, inferences and predictions of change, growth as well as many other processes. Through the generating role played by probability distributions in almost field of science, government, industry, technology and increasingly media and commerce we could name, probability mixes through almost all forms of relationality. In calculations of insurance risk, in algorithms for error correction, in psychological testing, in climate models or biodiversity surveys, just to name a few, probability distributions function ground all inference. Although certain distributions, such as the normal, Poisson or binomial, etc., have dominated in these developments, this was largely because it was easier to calculate estimates of their main parameters (mean, variance, etc). In terms of shape, area and hence probability density, the normal distribution is one of the most tractable curves to work with. Even with the various data transformations and normalizations developed over several centuries, other probability distributions are harder to work with. This occasions many disputes in the history of statistics over ‘curve-fitting’ to normal or other mathematically comfortable distributions as arbitrary and unjustified (Hacking, 1990, 164). However these disputes have been resolved (see (MacKenzie, 1978) for an early 20th century example), in one shape or another, the practical problem of finding the area under all or some part of the curve continues to deeply affect what we believe about many different things (sub-atomic particles, climate change, likelihood of glaucoma, the chances of rain today, etc.).

The mathematical difficulties posed by non-standard distributions relates to that other great Baroque mathematical invention, calculus. If calculus made possible so many different calculations of rates of change, calculations that profoundly affected senses of space, time, and increasingly growth and change more generally (hence, Deleuze’s work both on Leibniz and in his philosophical conceptualisation of difference more generally is deeply imbricated with differential calculus), it also ran into many obstacles in relations to calculations of probability. Calculating the area under a curve in order to estimate variables is an problem of integration. That is, the area under a curve is given the by integral of the probability density function. If the probability distribution cannot be normalized, then the area under the curve is much harder to estimate. The ornate and at times bewildering apparatus of statistical tests and procedures, as well as the debates between different schools of statistics (Bayesian vs frequentist), largely obscures the continuous trajectory in which what happens is transformed into a problem of measuring areas under curves, or volumes under surfaces. Sometimes the estimates are understood as a measure of our belief about what happens (as in Bayesian analysis) and sometimes it is understood as a measure of the frequency with which events occur in the world (as in frequentist statistics). While there is now a very extensive technical and philosophical

literature on the differences between Bayesian and frequentist statistics, more or less the same computations can be in the service of either standpoint. So this is not the main point I want to pursue here.

This situation has been changing recent decades, and this has begun to open a different gamut of realizations of probability. Nevertheless, more data actually increase the number of random variables. As statistical practices, coupled with information and communication systems, have been more widely mobilised, the desire to bring more and more events within the purview of probability functions generates many more random variables. At the same time, as hinted already in my introduction of MCMC, random variables have dilated to include all numbers. What used to be treated as a fixed parameter in mainstream statistics, for instance, the mean age of a national population (as contrasted with the mean of the sample) can now be treated itself as a random variable in Bayesian statistics (as Gill observes). So, not only do ages vary according to some distribution function in a population, the parameters that once described the shape of that distribution themselves can be understood in terms of a probability distribution. Furthermore, as soon as statisticians, scientists, engineers and others begin to try to model the relations between different random variables, they encounter non-standard distributions. That is, when there are multiple random variables, as is very commonly the case, their joint probability density function becomes much more complicated.

The problem of what to do with confluences of different data is a pressing issue. What happens when sets or series of random variables are given to us? It is one thing to countenance events or outcomes in terms of a normal curve. But if events — life expectancy, degree of atmospheric warming, an insurance risk — entail outcomes generated jointly by different, intersecting probability distributions, then our position, our capacity to grasp variation, and our ability to inhabit a situation, becomes much more slippery. In statistical terminology, we deal then with joint (probability) distributions. Rather than curves, joint probability distributions define more complex concave volumes, sometimes in many dimensional probability spaces. These multi-dimensional probability distributions distribute numbers in much more convoluted patterns that are difficult, if not impossible to graphically visualize. Finding the areas or volume of important regions of the joint probability density function is difficult yet essential. The so-called ‘marginal densities’ (Wasserman, 2003) 34} for joint probability distributions express relations between subsets of the random variables that might be measured or observed in a given experiment or observation. Usually the subset chosen for analysis refers to some particularly important dependency in a given situation (lung cancer mortality rates depend on smoking, levels of atmospheric carbon dioxide depend on hydrocarbon fuel consumption and deforestation, rates of swine flu infection during epidemics depend on school holiday dates and levels of existing immunity) against the background noise of many other events. Calculating a marginal density distribution means selects a space of heightened interest or importance that cuts across the possibly mountainous terrain of the overall joint probability distribution. This qualifies distributive

numbers in an important way: we all live in marginal densities, conditioned by many variables, and only occasionally does the joint distribution become in any way obvious.

Living in the margins

Marginal densities and joint probability distributions bring us back to MCMC. MCMC is, as mentioned above, a technique for simulating samples from high-dimensional or complicated concave volumes. In other words, it is a way of exploring the contoured and folded spaces when flows of data come together in one joint probability space. These spaces are barely inhabitable in the sense they are generated by the combinations of mathematical functions or probability distributions. They are barely inhabitable because it is not easy to see, find or feel how they are shaped, except in the exceptional cases where calculus can deliver a deductive analytical ‘closed form’ solution to the problems of integration (finding the area) and differentiation (finding the distribution function for one variable). By contrast, MCMC effectively simulates some part of the joint distribution, not by using complicated mathematical functions, but by shaping prodigious quantities of random numbers to approximate the shape of joint distribution. In doing this, data no longer comes from an outside world. In a genuinely monadic move, MCMC delineates the shape of the curve of the world by sampling numbers from the pseudo-random number generator supplied by the computer itself. How does it do this?

It is interesting to think why scientists working at the epicentre of the ‘closed world’ (Edwards, 1996) of nuclear weapons research should develop a technique that allows the world to open. In 1953, Metropolis, Rosenbluth and Teller were calculating ‘the properties of any substance which may be considered as composing of interacting individual molecules’ (Metropolis, Rosenbluth, Rosenbluth, et al., 1953, 1087) (for instance, the flux of neutrons in a hydrogen bomb detonation). In their short, but still widely cited paper (over 20,000 citations according to Google Scholar; over 14,000 according to Thomson Reuters Web of Knowledge), they describe how they used statistical simulation to deal with the number of possible interactions in a substance, and to thereby come up with a statistical description of the properties of the substance. Their model system consists of a square containing only a few hundred particles. These particles are at various distances from each other and exert forces (electric, magnetic, etc.) on each other dependent on the distance. In order to estimate the probability that the substance will be in any particular state (fissioning, vibrating, crystallising, cooling down, etc.), they needed to integrate over the many dimensional space comprising all the distance and forces between the particles. (This space is a typical multivariate joint distribution.) As they write, ‘it is evidently impossible to carry out a several hundred dimensional integral by the usual numerical methods, so we resort to the Monte Carlo method’ (1088), a method that Metropolis and Stanislaw Ulam had already described in an earlier

paper (Metropolis, 1949). They immediately go on to say, however, that they cannot just sample a random set of points to simulate because the probability distribution of the system is quite uneven, and randomly sampled points are very likely to lie in the low probability regions. Instead they propose a move which becomes the *modus operandi* of the subsequent MCMC work (and hence justifies the high citation count): ‘we place the N particles in any configuration ... then we move each of particles in succession’ (1949). Here is the beginning of the ‘random walk’ technique that underpins MCMC. By moving each particle by a small random amount, they can calculate a slight change in the system state, and then decide whether that move puts the system in a more or less probable state. If that state is more likely, the move is allowed; otherwise the particle stays where it is. Having carried out this process of small moves for all the particles, they can calculate the overall system state or property. The process of randomly displacing the particles by a small amount opens the bumpy topography of the joint probability density to exploration. The pattern of slightly different system configuration represented by the different particle positions taken together tends to approximate the probability distribution of the probability space more generally.

For several decades after this paper, the Metropolis Monte Carlo technique was applied to fluids, gases, plasmas, radiation, magnets, planets, stars, crystals and ions. But in 1970, the Canadian statistician W.K Hastings published a paper entitled ‘Monte Carlo Sampling Methods Using Markov Chains and Their Applications’ (Hastings, 1970). He generalised the Metropolis technique beyond the properties of physical substances, and supplements the first MC — Monte Carlo — by another MC — Markov Chain. In coining this second MC, Hastings’ paper formalises the random walk part of the Metropolis technique in terms of the mathematics of Markov chains. Markov chains are mathematical models that specify changes of states in some system according to probabilities. The statistical properties of Markov chains had been heavily investigated by statisticians and mathematicians, particularly in relation to how quickly the transitions in state converge to ‘stationary distributions.’ Given ways of constructing what the overall pattern of a particular random walk will look like, it also becomes possible to design Monte Carlo simulations that explore the complex topography multidimensional probability distributions without getting bogged or lost in sparsely populated, low probability regions. Hastings’ paper is less heavily cited than Metropolis’ (~ 3500 according to Web of Knowledge; 5800 according to Google Scholar), but represents a fork in the path of the technique. The application of MCMC begins to diverge from the physical properties of substances and turns towards a much more general set of situations. In Hastings’ paper, physical terms such as energy or distance retreat into the background, and instead the paper talks only in terms of distributions, and sampling from distributions. Although clearly indebted to physics, the technique has become broadly statistical. Even as use of Metropolis MC continues to broaden in the physical sciences in analyses of many physical processes, statisticians and others began to treat the technique in much more general terms.

This is not to say that all reference to physical properties disappears. In many ways, the references to physics become more important, even as the technique begins to be used outside physics. For instance, Stuart and Donald Geman’s ‘Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images’ (Geman & Geman, 1984) has been cited many thousands of times (~ 6200 according to Web of Knowledge; 13,500 according to Google Scholar) in and across diverse fields such as artificial intelligence, genetics, medical imaging, oncology, and forestry. On the one hand this paper concerns a topic seemingly distant from physics — ‘the restoration of images.’ Again, the forest of citations associated with paper on MCMC suggest that the technique has somehow taken on a much relevance than the fields of statistical physics or imaging processing in which it first took shape. On the other hand, the title of the paper still points to physics: the ‘Gibbs distribution’ is one of the most well-known distributions in the physics of particle movements. Indeed, Geman and Geman make a direct analogy between atoms and images. They treat blurry or distorted images as overheated physical systems. The task is to simulate the cooling of the system in order to isolate and estimate the most probable image, where an image is understood as a spatial lattice of light intensities and edges. They use this approach (although the name MCMC has not yet been applied to it) to effectively generate a sequence of images, each of which is more likely than the last to be the originally observed image. One reason this paper continues to be so heavily cited is that they name their method of overcoming the the computational problem of generating this sequence of images ‘the Gibbs Sampler’ (Geman & Geman, 1984, 722).

Finally in 1990, an article entitled ‘Sampling-Based Approaches to Calculating Marginal Densities’ appeared in *Journal of the American Statistical Association* (Gelfand & Smith, 1990). The statisticians Alan Gelfand and Adrian Smith state that the problem they are addressing is how ‘to obtain numerical estimates of nonanalytically available marginal densities of some or all [the collection of random variables] simply by means of simulated samples from available conditional distributions, and without recourse to sophisticated numerical analytic methods’ (Gelfand & Smith, 1990, 398). They take up the Gibbs sampler algorithm as developed by (Geman & Geman, 1984) for image-processing, investigate some of its formal properties (convergence), and then set out a number of mainstream statistical problems that could be done differently using MCMC and the Gibbs sampler in particular. This paper is sometimes said to have announced the ‘Bayesian revolution’ (Robert & Casella, 2010, 9) because it made clear the links between MCMC and statistical inference more generally through six illustrative mainstream examples: multinomial models, hierarchical models, multivariate normal sampling, variance components, and the k-group normal means model. The details of these examples need not detain, but each of the illustrations in the paper shows how previously difficult problems of Bayesian inference can be carried out by sampling simulations. As they state in another paper from the same year, ‘the potential of the methodology is enormous, rendering straightforward the analysis of a number of problems hitherto regarded

as intractable' (Gelfand, Hills, Racine-Poon, et al., 1990, 984). A rapid convergence on MCMC follows from the 1990s onwards. Gibbs samplers appear in desktop computer software such as the widely used WinBUGS ('Windows Bayes Using Gibbs Sampler') written by statisticians at Cambridge University in the early 1990s (Lunn, Thomas, Best, et al., 2000), and MCMC quickly moves into the different disciplines and applications found today.

Where does the development of MCMC techniques, and in particular the Gibbs sampler, leave numbers? A recent textbook of MCMC methods writes:

A Gibbs sampler is a simulated Markov chain X_1, X_2, \dots that is constructed to have a desired long-run distribution. By observing enough simulated values of X_n after the simulation has stabilized, we hope that the distribution of sampled X_n will approximate its long-run distribution (Suess & Trumbo, 2010, 150)

Like the many other contemporary descriptions of the workings of the Gibbs sampler, and the Metropolis-Hastings algorithm (the two principal implementations of MCMC), this formulation brings us back to the connection between distribution and simulation. MCMC simulates distributions by generating samples from them. The samples themselves, rather than coming from the world, come from computations. The chain of values produced by the Markov Chain, if everything is working, move towards a stable distribution of values. This stable or repeated distribution of values approximates the desired but inaccessible joint probability density distribution, the distribution that bears within it the relations between different random variables. In all these settings, of which I have listed only a few, versions of the same problem could be found: the extreme difficulty of numerically or analytically calculating the marginal densities given joint densities of multiple random variables. In all these settings, massive computing power ('brute force') by itself would not help because it would have no way of closely following the shape of the joint density.

The central idea is to use available information about a prior distribution and data to construct an ergodic Markov chain whose limiting distribution is the desired posterior distribution. Then we simulate enough steps of the chain to obtain a good approximation to the limiting distribution (Suess & Trumbo, 2010, 21).

In practice, MCMC can be used to combine existing data, a prior estimate of the distribution of values of the variable under consideration (the prior distribution) to generate an updated estimate (the posterior distribution) of the value of the variable. The countless Markov Chains that have been constructed and run in the last two decades largely relate to this application of MCMC.

Conclusion

In the wake of the refining and popularisation of MCMC, particularly in the form of the Gibbs Sampler during the 1990s, distributive numbers varied more widely. MCMC has been principally applied to a particular kind of statistical inference, Bayesian inference, inference that combines observed data with prior estimates of the value of variables. There is a very abundant literature on Bayesian inference. While many accounts of Bayesian inference emphasise the contrast between frequentist and subjective interpretations of probability, they also curtail the variation in distributive numbers by subsuming them into an object-subject difference. For our purposes, however, neither the objectivist (frequentists) or subjective (Bayesian) interpretations of probability work well. The Bayesian interpretation has the virtue of treating all variables or parameters as probability distributions. It allows everything to vary, and makes ‘the truth of a variation appears to a subject’ (Deleuze, 1993, 20). All parameters and parameters in a Bayesian framework can become random variables subject to degrees of belief or credibility. Nevertheless, while Bayesian statistics seem to me to lie closer to a Baroque treatment of data, understanding it as ‘subjective’ tends to gloss over the ways in which MCMC supports inference by the generation of vast quantities of random numbers. It still needs numbers from the world. It draws on constant supplies of numbers that vary without an obvious pattern or predictability. This is not a problem with the technique itself, a technique whose capacity to generate numbers in the world has been fabulously productive. But it does suggest that we need to think more about what happens with more data.

Many of Deleuze’s formulations of the Baroque converge on the curves. He suggests, for instance, that the world ‘is the infinite curve that touches at an infinity of points an infinity of curves, the curve with a unique variable, the convergent series of all series’ (Deleuze, 1993, 24). This description of the world as curve, or the curve of the world, resonates strongly in the scene I have been describing, both in the general account of probability as distributive number and in MCMC as a technique that crafts distributive numbers by playing on the convergence of series of numbers. In Deleuze’s account of the fold, curves act as causes: ‘the presence of a curved element acts as a cause’ (Deleuze, 1993, 17). Like so many others in this strangely impossible book, this claim begins to make more sense as we see curves proliferate, and as we see how finding, generating, exploring and shaping curves become more common practices in so many settings (asthma studies, multiplayer game coordination, epidemiological modelling, spam filtering, etc.). The particles, maps, images and populations comprising our world figure in a Baroque sensibility as curves. When Deleuze writes, following Leibniz, ‘the world is the infinite curve that touches at an infinity of points,’ he could be describing how curves generated by distributive numbers populate our worlds with shapes, textures, images, sounds and movements derived from convergent series of numbers.

In all almost any setting - government, civil society, health, entertainment,

the military, the or private life — we face ‘more data.’ Many different things might happen in the wake of more data. I have suggested that amidst the many different things we might do with more data, one common feature is conspicuous: data does not end uncertainty. Rather, as the pervasiveness of MCMC indicates, more data actually triggers much greater recourse to randomness. It draws on much great quantities of random numbers to support exploration of data. This is not because observations or measurements pertaining to what happens in the world are random, as if everything uniformly happens by chance. A ‘uniform distribution’ of event would actually make things a lot simpler since then everything would happen according to dice rolls, and dice rolls are more or less ‘iid’ — independent and identically distributed. Just the opposite, so much happens relationally that numbers ineluctably point to joint or conditional probabilities. More data, then, suggests the need to track more interactions or relations between the different events that gave rise to the data. When different flows of data coming from the world come together, they generate the complex volumes or surfaces that Deleuze describes so well in *The Fold*. These surfaces enfold events, and reshape them.

Where are we in the folded volumes of data? By tracing some of the curved surfaces that have been generated in the high-dimensional spaces, we have some chance of finding what in our sensation of change, movement, texture or image is attributable to distributive numbers. What I have been seeking to do in relation MCMC is to find ways of internalizing or including numbers. It would be possible to trace other data-related techniques along the same lines. For instance, the rise of bootstrap sampling as a way of estimating statistical errors (Efron, 1979), or the emergence of the Expectation Maximisation (EM) algorithm (Dempster, Laird & Rubin, 1977) would be other possible, perhaps less rich, examples, that shows how simulation supports distributive numbers. The example of MCMC offers something that lies closer to the power of curves. Deleuze describes what how monads include the world in many different ways in the fold. In nearly all of them, the monad is differentiating and integrating: ‘each monad includes the world as infinite series of infinitely small units, but establishes differential relations and integrations only upon a limited portion of the series, such that the monads themselves enter in an infinite series of inverse numbers’ (Deleuze, 1993, 130). The curves of the marginal densities that MCMC estimates are integrations that include a limited portion of the world, albeit in a grainier, less smoothly flowing form than that envisioned by Deleuze. Those curves express the world (and indeed the state of the world is often expressed as a series of curves: commodity prices, population growth/decline, economic growth, etc.), and those curves are in the world. ‘The world is what the soul expresses,’ as Deleuze points out, just ‘because those curves are in the world’ (Deleuze, 1993). They tessellate what we see, and what happens. In much of what I have been describing about distributive numbers, random variables, probability distributions, and the MCMC techniques of sampling from distributions, there has been some excitement, some tentative desire, to say that it would be good, from the perspective of the Baroque sensibility, not to diminish or flatten these curves,

but to find how we are included in them. By this I don't mean everyone should be using MCMC in order to update their beliefs on various matters. We are all affected by MCMC in various ways (as for instance, when an Xbox-Live player finds themselves pitted against a new opponent online, the match has been arranged by an MCMC-based player-matching system that is constantly updating its profile of player abilities using game data), and could potentially react to that in interesting ways. If we are the most simple numbers, inverse numbers, individual and distributive, if we inhabit simulated surfaces, then the problem becomes how to integrate and differentiate well. Rather we need to find what in those operative functions that grid our world can be reused politically and philosophically.

References

- Amoore L (2009) Lines of sight: on the visualization of unknown futures. *Citizenship Studies*, 13, 17–30, Available from: <http://search.ebscohost.com/login.aspx?direct=true&db=a9h&live> (accessed 19 May 2010).
- Andrieu C, De Freitas N, Doucet A, et al. (2003) An introduction to MCMC for machine learning. *Machine learning*, 50, 5–43.
- Badiou A (2008) *Number and Numbers*. Cambridge, Polity Press.
- Barry A and Thrift N (2007) Gabriel Tarde: imitation, invention and economy. *Economy and Society*, 36, 509–525.
- Besag J (1974) Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, 192–236.
- Bowker GC (2010) A Plea for Pleats. In: Jensen CB (ed.), *Deleuzian Intersections: Science, Technology, Anthropology*, Oxford, Berghahn Books, p. 123.
- Daston L (1988) *Classical Probability in the Enlightenment*. New Brunswick, N.J., Princeton University Press.
- Daston L (1994) How probabilities came to be objective and subjective. *Historia mathematica*, 21, 330–344.
- Deleuze G (1993) *The fold : Leibniz and the Baroque*. London, Athlone Press.
- Dempster AP, Laird NM and Rubin DB (1977) Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39, 1–38, Available from: <http://www.jstor.org/stable/2984875> (accessed 19 October 2012).
- Derrida J and Stiegler B (2002) *Echographies of television : filmed interviews*. Cambridge, Polity Press.
- Desrosieres A (1998) *The Politics of Large Numbers: A History of Statistical Reasoning*. Cambridge, Mass, Harvard University Press.

- Edwards PN (1996) *The closed world : computers and the politics of discourse in Cold War. Inside technology*, Cambridge, Mass. ; London, MIT Press.
- Efron B (1979) Bootstrap methods: another look at the jackknife. *The annals of Statistics*, 7, 1–26, Available from: <http://projecteuclid.org/euclid.aos/1176344552> (accessed 19 October 2012).
- Elden S (2006) *Speaking against number: Heidegger, language and the politics of calculation*. Edinburgh, Edinburgh University Press.
- Espeland WN and Stevens ML (2008) A Sociology of Quantification. *European Journal of Sociology / Archives Européennes de Sociologie*, 49, 401–436.
- Gelfand AE and Smith AFM (1990) Sampling-Based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association*, 85, 398–409, Available from: <http://www.jstor.org.ezproxy.lancs.ac.uk/stable/2289776> (accessed 14 January 2011).
- Gelfand AE, Hills SE, Racine-Poon A, et al. (1990) Illustration of Bayesian inference in normal data models using Gibbs sampling. *Journal of the American Statistical Association*, 85, 972–985, Available from: <http://amstat.tandfonline.com/doi/abs/10.1080/01621459.1990.10474968> (accessed 19 October 2012).
- Geman S and Geman D (1984) Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 721–741.
- Gill J (2011) Introduction to the Virtual Issue: Bayesian Methods in Political Science.
- Habermas J (1970) *Toward a rational society; student protest, science, and politics*. Boston, Beacon Press.
- Hacking I (1975) *The emergence of probability*. Cambridge ; New York, Cambridge University Press.
- Hacking I (1990) *The taming of chance*. Cambridge University Press.
- Hastings WK (1970) Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57, 97–109.
- Heidegger M and Lovitt W (1977) *The question concerning technology, and other essays*. Harper colophon books, New York ; London, Harper and Row.
- Holmberg C, Bischof C and Bauer S (2012) Making Predictions: Computing Populations. *Science, Technology & Human Values*, Available from: <http://sth.sagepub.com/content/early/2012/04/27/0162243912439610> (accessed 08 October 2012).
- Husserl E (1970) *Crisis of European sciences and transcendental phenomenology*. Chicago, Ill., Northwestern University Press.

- Knorr-Cetina K and Bruegger U (2002) Traders' Engagement with Markets. A Postsocial Relationship. *Theory, Culture and Society*, 19, 161–185.
- Latour B, Jensen P, Venturini T, et al. (2012) The Whole is Always Smaller than its Parts. How Digital Navigation May Modify Social Theory. *British Journal of Sociology*.
- Lunn DJ, Thomas A, Best N, et al. (2000) WinBUGS-a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing*, 10, 325–337.
- MacKenzie D (1978) Statistical Theory and Social Interests A Case-Study. *Social Studies of Science*, 8, 35–83, Available from: <http://sss.sagepub.com/content/8/1/35.short> (accessed 08 October 2012).
- Mackenzie DA (1990) *Inventing accuracy: an historical sociology of nuclear missile guidance. Inside technology*, Cambridge, Mass., MIT Press.
- Mcgrayne SB (2011) *The Theory That Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines, and Emerged Triumphant from Two Centuries of Controversy*. New Haven, CT, Yale University Press.
- Metropolis N (1949) The Monte Carlo Method. *Journal of the American Statistical Association*, 44, 335–341.
- Metropolis N, Rosenbluth AW, Rosenbluth MN, et al. (1953) Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21, 1087, Available from: <http://link.aip.org/link/%3FJCPSA6/21/1087/1> (accessed 10 October 2012).
- Moore DS (2009) *The Basic Practice of Statistics*. New York; London, W. H. Freeman.
- Porter TM (1996) *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton Univ Pr.
- Robert C and Casella G (2008) A History of Markov Chain Monte Carlo-Subjective Recollections from Incomplete Data. *0808.2902*.
- Robert CP and Casella G (2010) *Introducing Monte Carlo Methods with R*. New York, Springer.
- Silver N (2012) *The Signal and the Noise: The Art and Science of Prediction*. New York; London, Allen Lane.
- Stiegler B (1998) *Technics and time, 1: the fault of Epimetheus. Meridian: crossing aesthetics*, Stanford, Calif., Stanford University Press.
- Stigler SM (1986) *The history of statistics: the measurement of uncertainty before 1900*. Cambridge, Mass., Harvard University Press.
- Suess EA and Trumbo BE (2010) *Introduction to Probability Simulation and Gibbs Sampling with R*. New York, Springer.

Verran H (2011) The changing lives of measures and values: from centre stage in the fading “disciplinary” society to pervasive background instrument in the emergent “control” society. *The Sociological Review*, 59, 60–72.

Wasserman L (2003) *All of statistics: a concise course in statistical inference*. New York, Springer.