A Mathematician Explores the Gap Between Stories and Statistics, Logic and Language

John A. Paulos
"Mathematics and Narrative" Meeting
Mykonos, July 2005

Introduction

Borrowing from Isaih Berlin, I note that we're all foxes, and we need a hedgehog. We know many little things about the relation between mathematics and narrative, but lack one big comprehensive insight. Leaving aside the question of whether any such insight is possible or not, I offer here a few of the things I know about the complex interaction between mathematics and narrative. It is a topic I discussed at much greater length in my book, Once Upon a Number: The Hidden Mathematical Logic of Stories.

First, a little vignette: At his faint chuckle she turned and faced her once beloved uncle. Unceremoniously she ripped the papers from the pocket of his Hawaiian shirt as he warily stepped backwards toward the hotel room door and then, with unmitigated disgust at his blubber as well as his duplicity, she hissed, "22.8% of all bankruptcies filed between July, 1995, and June, 1997, were attributed to bad legal advice, up 9.2% over the last biennial period."

"I did the best I could," the 273 pound man answered faintly. He desperately wanted to avoid further rousing his enraged niece who, despite a lithe figure, a weight of 113 pounds, an angelic face, and long curly hair, was fully capable of physically hurting him. Once safely out in the hotel hallway, he took heart and offered, "A meta-analysis of several studies suggests that fewer than 40% of legal malpractice cases are due to malicious intent, the balance to simple incompetence." At this she lunged for him, scratching his neck and tearing the flimsy shirt from his fat back.

As this vignette is intended to illustrate, the stories we tell in our everyday lives often coexist uncomfortably with statistics of supposed relevance, even when the two don't outright contradict each other. Our stories are filled with people who do things and who have motives, desires, fears, and possibly an unnatural love of rigatoni. Their particular circumstances and situations loom large in any description of them. In statistics, on the other hand, there are rarely agents, but only demographics, general laws, processes. Particularities and details are usually dismissed as unimportant.

Stories and Statistics

Stories and statistics? Whatever might this juxtaposition be getting at? If pressed, most people would probably say something dismissive like stories and statistics go together like a horse and paper clip. But I argue that story-telling and informal discourse have given birth over time to the complementary modes of thinking employed in statistics, logic, and mathematics generally. Although the latter skills are generally more difficult to come by and occasionally run counter to our intuitions, in broad outline we can say that first we tell stories, and then, in the blink of an eon, we cite statistics.

This suggests that the chasm between them is more a matter of tradition, degree, and terminology than something untraversably deep. I believe this to be so, and because the gap between stories and statistics is a synecdoche for the gap between C.P. Snow's proverbial two cultures, some of my points may have a broader resonance than initial appearances would indicate.

Since synecdoche is a literary term for a figure of speech in which the part is substituted for the whole, or sometimes the other way around, it is somewhat analogous to a sample's substituting for the whole population. With this bit of pedantry, we've already landed our first bit of kite string across the chasm.

Notions of probability and statistics did not suddenly appear in the full dress regalia we encounter in mathematics classes. There were plebeian glimmerings of the concepts of average and variability in stories dating from antiquity. Bones and rocks were already in use as dice. References to likelihood appear in classical literature.The importance of chance in everyday life was clearly understood by at least a few. It's not hard to imagine thoughts of probability flitting through our ancestors' minds. "If I'm lucky, I'll get back before they finish eating the beast."

Millennia later, these ideas were formalized as Pascal and Fermat refined them to solve certain gambling problems in the 17th century, Laplace and Gauss further developed their applications to scientific concerns in the next century and a half, and Quetelet and Durkheim used them in the 19th century to help understand regularities in social phenomena. Soon people were saying things like: "The chances of getting at least one 6 in four rolls of a single die are greater than the chances of getting at least one 12 in twenty-four rolls of a pair of dice." "The probability of a particle decaying in the next minute is .927." "Exit polls show that 4 out of 5 voters in favor of gun control legislation cast their ballots for Gore."

After this bullet train through the history of statistics, let me slow down to note some of the many colloquial ancestors of the most salient ideas in probability and statistics. Consider first the notions of central tendency - average, median, mode, etcetera. They most certainly grew out of workaday words like "usual," "customary," "typical," "same," "middling," "most," "standard," "stereotypical," "expected," "nondescript," "normal," "ordinary," "medium," "conventional," "commonplace," "so-so," and so on. It is hard to imagine prehistoric humans, even those lacking the vocabulary above, not possessing some sort of rudimentary idea of the typical. Any situation or entities - storms, animals, rocks - that recurred again and again would, it seems, lead naturally to the notion of a typical or average recurrence.

Or examine the precursors of the notions of statistical variation - standard deviation, variance, and the like. These are words such as "unusual," "peculiar," "strange," "singular," "original," "extreme," "special," "unlike," "unique," "deviant," "dissimilar," "disparate," "different," "bizarre," "too much," and so on. The slang term "far-out" to indicate unconventionality is particularly interesting because an observation that is far out on the "tail" of a graph of a statistical distribution is rare and unusual and bespeaks a high degree of variability in the quantity in question. Again any recurrent situation or entity would, over time, suggest the notion of an unusual exception. If some events are common, others are rare.

Probability itself is present in such words as "chance," "likelihood," "fate," "odds," "gods," "fortune," "luck," "happenstance," "random," and many others. Note that a mere acceptance of the idea of alternative possibilities and of the open-endedness essential to story-telling almost entails some notion of probability; some alternatives will be judged more likely than others. And the need to single out aspects of recurrent situations and entities leads to the key statistical concept of sampling, reflected in words like "instance," "case," "example," "cross-section," "observation," "specimen," and "swatch." Likewise, the natural mental process of yoking together like things suggests the important idea of correlation, which has the following correlates (so to speak): "association," "connection," "relation," "linkage," "conjunction," "conformity," "dependence," "proportionate," and the ever too ready "cause."

Even less familiar statistical ideas such as control, standardization, hypothesis testing, so-called Bayesian analysis, and categorization correspond to common sense phrases and ideas that are an integral part of human cognition and story-telling. Like Moliere's character who was shocked to find that he'd been speaking prose his whole life, many people are surprised when told that much of what they characterize as common sense is statistics or, more generally, mathematics. It's telling too that the word "account" refers not only to numbers but to narratives as well.

Admit it or not, we're all statisticians, as when we're making grand inferences about a person from the tiny sample of behavior known as a first impression. The difference between mathematical statistics and the everyday variety is often simply the degree of formalization and objective rigor. Standard deviation is computed according to specific rules and definitions, and so are correlation coefficients, the rank-sum statistic, chi-square values, and averages. Their everyday cousins are not so formalized.

This does not mean that there cannot be stringent constraints on the mundane uses of these terms as well. The comedian Steven Wright tells a story about going into a clothing store and telling the clerk he's looking for a shirt that is "extra medium."

Stories as Context for Statistics

Unfortunately, people generally ignore the connections between the formal notions of statistics and the informal understandings and stories from which they grow. They consider numbers as coming from a different realm than narratives and not as distillations, complements, or summaries of them. They often cite statistics in bald form, without the supporting story and context needed to give them meaning.

Part of context is internal and attitudinal. People don't fully realize that how *we* characterize people and events, how we view their circumstances and context, and how we imbed them into stories often determines to a large extent what we think of them. If, for example, we describe a person Waldo as coming from country X, 45 percent of whose citizens have a certain characteristic, then if we know nothing else about him, it seems reasonable to assume that there is a 45 percent probability that Waldo shares this characteristic. But if we describe Waldo as belonging to a certain ethnic group 80 percent of whose members in the region comprising countries X, Y, and Z have the characteristic in question, then we'll probably conclude the chances are 80% that Waldo shares this characteristic. And if we describe Waldo as belonging to a nation-X-wide organization only 15 percent of whose members have this characteristic, then we're likely to state that his chances of having the characteristic are only 15%. Which (combination of) descriptions we employ is to an extent up to us, so the pleasingly precise statistics we confidently cite are often as revealing of us as they are of Waldo (who, just for the record, does not have the characteristic).

More commonly the problem is not with our attitudes, however, but with our knowledge. We're simply unaware of the external context of most statistics that we read or hear about. Since this is particularly true of news stories, let me begin with them. The contextual questions we should ask when reading a news story are the very ones statisticians ask when presented with a survey of some sort. We want the answers to questions such as how many, how likely, and what percentage, of course. But we also want to know if the numbers on homelessness or child abuse, for example, come from police blotter reports, in which case they're likely to be low; or if they come from scientifically controlled studies, in which case they're likely to be somewhat higher; or if they come from the press releases of groups with an ideological axe to grind, in which case they're liable to be extremely high (or extremely low depending on the ideology).

Without an ambient story, some background knowledge, and some indication of the provenance of the statistics, it's impossible to evaluate their validity. Common sense and informal logic are as essential to this task as an understanding of the formal statistical notions; both are preconditions for numeracy. Although many stories need no numbers, some accounts without supporting statistics run the risk of being dismissed as anecdotal. Conversely, although some figures are almost self-explanatory, statistics without any context always run the risk of being arid, irrelevant, and even meaningless.

The denial of the mutual dependence of stories and statistics, and the pedagogy that is a consequence of this denial, is one reason for the disesteem in which statistics, and mathematics and science generally, are widely held. Its practitioners are simultaneously saluted as awe-inspiring geniuses and summarily disregarded as ivory-tower-eccentrics. (Most of the time they are neither, sometimes one or the other, rarely both.) Describing the world may be thought of as a Olympic contest between simplifiers - scientists in general, statisticians in particular - and complicators - humanists in general, story-tellers in particular. It's a contest they should both win.

Stories not only provide context for statistical statements. They can also illustrate and vivify them as well. Consider that, as with philosophical abstractions, many of the ideas and problems in probability theory have associated with them standard vignettes. Examples are such stories as the gambler's fallacy and gambler's ruin, the Banach match box problem, the drunkard's random walks, the Monty Hall problem, the St. Petersburg paradox, the random chord problem, the hot hand, the Buffon needle problem, and many others.

More Contrasts

Zillions of more conventional stories, from the *Iliad* and *Odyssey* to art films and television soap operas, and zillions of surveys, polls, and studies demonstrate the many contrasts between stories and statistics. One major difference is that in story-telling the focus is almost always on individual people rather than on analyses, arguments, and averages. Such a focus is a necessary corrective to overweening abstraction and keeps the statistics in human perspective.

Even if they're true there's sometimes a hint of something inhuman and vaguely pornographic about statistics. A couple of extreme examples: since half the people in the U.S. are men and half are women, the average American adult has one ovary and one testicle, or the average resident of Dade County, Florida is born Hispanic and dies Jewish. Conversely, pornography, with its loosely bound sequences of many undifferentiated storyless sexual couplings (or triplings), often has the feel of a statistical survey.

But a focus on individuals can be deceptive and manipulative and can distort discussions of public policy issues, especially those involving health and safety. A poignant television story of a victim of a rare reaction to a vaccine can render invisible the vast good brought about by this same vaccine. There are countless examples of this media-induced bathos, instances of the so-called "tyranny of the anecdote."

Some writers try to enjoy the virtues of both individuals' stories and statistical surveys by improperly conflating them. The result doesn't so much bridge the gulf between them as fall into it. An example is the common convention of conjuring up some "representative" person, a fictional Jeremy, Linda, or Kevin (but never a Waldo or Gertrude), to endorse or exemplify whatever statistical conclusion a newspaper or magazine article has reached. This practice seems seems to be getting more common.

A number of other critical aspects of the gap between statistics-citing and story-telling derive from the fact that, as the proverbial writing teacher's maxim enjoins, a story shows, rather than tells. Stories usually employ dialogue and other devices and don't limit themselves to declarative pronouncements; they develop the context and relevant relationships instead of merely positing raw data; they are open-ended and metaphorical, whereas statistics and mathematics generally are determinate and literal; and stories unfold in time instead of being presented as timeless.

A computer analogy is helpful. If we think of conventional stories as being told from one point of view (just as a serial processor performs one calculation at a time), then statistics may be thought of as providing a view from nowhere in particular (very many parallel processors performing simultaneous calculations). Between them are these amalgams which may be thought of as a varying number of variously connected viewpoints (processors). Combining the virtues of these two very different ways of apprehending the world, through stories and statistics, can be thought of as a literary analogue to a common problem in computer design and architecture.

Too Many Characteristics, Not Enough People

Still, the right balance between depth of characterization and the number of characters isn't always clear. In stories as in everyday life there are relatively few people we interact with personally, but they are real 3-dimensional folk (actually in a mathematical sense, they're N-dimensional folk for large values of N). They possess or are associated with an indeterminately large number of possible traits, circumstances, relationships, informal rules, and agreements. We certainly don't know everything about the people closest to us or even about ourselves, but we are implicitly aware of so many details and so much richness of context that writing it all down would turn most of us into bad novelists. Even to those whom we don't know well, we can attach a dozen adjectives, a few adverbs, and a couple of anecdotes. Contrast this abundance of personal particulars with most scientific studies where, although there generally are a very large number of people (or other data), the people surveyed are flat, having only one or two dimensions - who they'll vote for, whether they smoke, or what brand of soft drink or laxative they prefer.

Stories and statistics offer us the complementary choices of knowing a lot about a few people or knowing a little about many people. The first option leads to the common observation that novels illuminate great truths of the human condition. Novels are multivalent and bursting with ironies, details, and metaphors, while social science and demographic statistics can seem simple-minded and repellingly earnest by comparison. We can, however, easily delude ourselves into thinking that more of a general nature is being revealed to us by some memoir or personal reminiscence or by a novel or short story than is really the case. Biased and small samples are always major problems, of course, but my caveat arises from something more specific: the technical, uneuphonic statistical notion of an adjusted multiple correlation coefficient.

If the number of traits considered is large compared to the number of people being surveyed, there will appear to be more of a relationship among the traits than actually obtains. To see why, imagine a study that examined only two people and two characteristics, say intelligence and shyness. Imagine further a graph with degrees of intelligence on one axis and degrees of shyness on the other and two points on it corresponding to the two people. If the shyer of the two were more intelligent, there would be a perfect correlation between the two traits and a straight line connecting the two points on the graph. More shy, more intelligent. But if the shyer of the two were less intelligent, there would still be a perfect correlation between the two traits and a straight line pointing in the opposite direction connecting the two points. More shy, less intelligent.

You can find perfect correlations that mean nothing for any three people and three characteristics and, in general, for any N people and N characteristics. But the number of characteristics needn't equal the number of people. Whenever the number of characteristics is a significant fraction of the number of people, the so-called multiple correlation among the characteristics will suggest spurious associations among them.

To tell us anything useful, multiple correlation analysis must be based on a very large number of people and a much, much smaller number of characteristics. Yet the insights commonly that come from stories and everyday life are precisely the opposite. We each know in a full-bodied way relatively few people, and for them the number of characteristics, relationships, characteristics of relationships, relationships of characteristics, and so on that we're aware of is indeterminately large. Thus we tend to overestimate our general knowledge of other people and are convinced of all sorts of associations (more complicated variants of "more shy, less intelligent") that are simply bogus. By failing to adjust downward our multiple correlation coefficients (so to speak), we convince ourselves that we know all manner of stuff that just isn't so.

Just as stories are sometimes a corrective to the excessive abstraction of statistics, statistics are sometimes a corrective to the misleading richness of stories.

Stereotypes, Whimsy, and Statistical Conservatism

The alternative in everyday life to probabilistic calculations and explanations is the amorphous "discipline" of common sense and rough appraisal. Rather than invoking precise probabilities, in our everyday approach to life we find it more natural to deal with rules of thumb and approximate categories - in other words, with stereotypes. Although many assume that stereotypes are always evil vestiges of benighted mind-sets, more often they are essential to effective communication and have themselves been unfairly stereotyped (assuming a concept can be treated unfairly). Many stereotypes permit the economy of expression necessary for rapid communication and effective functioning. "Chair" is a stereotype, but one never hears complaints from bar stools, recliners, bean bags, art deco pieces, high back dining room varieties, precious antiques, chaise longues, or kitchen instances of the notion. Stereotypes, of course, admit of all sorts of exceptions that upon further examination in individual cases are easily apparent, but that doesn't mean they should or even can be universally proscribed. Complexity, subtlety, and precision cost time and money, and they are often unnecessary and sometimes even obscuring.

Recognition of common stereotypes and knowledge of recurring stereotyped situations - restaurant behavior, retail purchasing, hygiene practices, audience deportment, etcetera - are essential for navigating through everyday life. Approaches to artificial intelligence, in particular that of computer scientist Roger Schank and others, have reinforced the observation that we chart our course and communicate with others by invoking common types, scenarios, and scripts as shorthand. Like statistical notions, stereotypes generally do violence to particular cases and individuals, but they pay their way by summarizing general information whose many exceptions would be too time-consuming to note.

In any case, statistical decision-making is a drab, conservative process unlike the spirited snap judgements that characterize personal appraisals. The so-called null hypothesis in statistics is the assumption that the phenomenon, relationship, or hypothesis under observation is not significant, but merely the result of chance. To reject the null hypothesis it is conventional to require that the probability of the phenomenon occurring merely by chance be less than five percent. (This is the source of the story about the statistician who witnessed the decapitation of 25 cows, noted that one survived the ordeal, and dismissed the survival as not significant.) In my peregrinations through this world, I've observed that few people regularly make decisions like this in their personal lives. It would be too boring even if such precision were possible. (Having offered a partial defense of stereotypes, I feel obliged to mention that a common related stereotype of statisticians is that they are people who chose their profession because they couldn't stand the excitement of accounting. Recal also the definition of the extroverted mathematician. He's the one who looks at your shoes while he's speaking.)

The idea of boredom suggests yet another difference between stories and statistics. In listening to stories we have an acknowledged inclination or desire to suspend an initial disbelief in order to be entertained, whereas in evaluating statistics we generally have an opposite inclination or desire to suspend an initial belief in order not to be beguiled. In statistics we're said to commit a Type I error when we reject a truth and a Type II error when we accept a falsehood. There is, of course, no way to always avoid both types of error, and we have different error thresholds in different endeavors. Nevertheless, the type of error people feel more comfortable making gives some indication of their intellectual personality type. People who like to be entertained and beguiled and who hate the prospect of making a Type I error might be more likely to prefer stories to statistics. Those who don't like being entertained or beguiled and who hate the prospect of making a Type II error might be more likely to prefer statistics to stories. In any case, this speculation is a little story with no statistics to back it up so make of it what you wish.

Although wrong much of the time, we have more confidence in our personal gut decisions than we do in more public ones. We all tend to distrust decisions that are made far from us. We insist on exacting statistical protocols in public decision-making yet often accept the sloppiest reasoning from those closest to us. In small groups there is trust and little perceived need for statistics. As Theodore Porter has shown in his *Trust in Numbers*, quantitative methods and controls often arise because of the political weakness of expert communities and a suspicion of their findings by the larger community. Those anticipating distrust are most likely to undergird their conclusions with substantial statistics or at least adorn them with fake statistical finery.

Between Subjective Viewpoint and Impersonal Probability

Part of the reason expressions of racism seem more common to minorities than they do to majorities is that they are merely by virtue of the arithmetic fact that they are a minority. Even assuming that, contrary to fact, blacks and whites hold positions of equal importance and influence in the US, for example, and that a similar percentage of each group is racist and that the country is both residentially and professionally integrated, blacks will be the recipients of more expressions of racism. Note that one of the conventional elements of narrative plays a role in the above analysis: the simple notion of an agent's viewpoint. The richness and complexity of most everyday situations make basic arithmetic insights less visible. Similar remarks hold for fictional situations. Seeing events from the point of view of a character within a story or through the eyes of the story's narrator is not conducive to probabilistic reasoning.

Another instance of the connection between the personalization of events and mild paranoia is Murphy's Law, which states that generally whatever can go wrong will go wrong. Contrary to the tongue-in-cheek aspect of this characterization, there is something profound about the phenomenon it describes. The failure of things to work right is in many situations not a result of personal bad luck but a consequence of the complexity and interdependence of many systems.

A homely, counterintuitive example of Murphy's Law comes from probability theory and was recently elaborated upon by science writer Robert Matthews. Imagine you have 10 pairs of socks and that despite your best efforts 6 of these socks disappear. The question is: what is more likely - that you're lucky and end up with 7 complete pair, i.e., the 6 missing socks come from only 3 pairs, or that you're unlucky and end up with only 4 complete pair, i.e., that the 6 missing socks come from 6 different pairs? The surprising answer is that you're more than 100 times as likely to end with the worst possible outcome, only 4 pair (plus 6 single socks), than you are to end up with the best possible outcome, 7 pair (and no singles). More precisely, the probability of 7 pair is .003, of 6 pair is .130, of 5 pair is .520, and of only 4 pair is .347.

The tendency of socks to shed their mates is certainly Murphy's Law with a vengeance! Nevertheless, it's what we should expect to happen and there is no need to invoke bad luck to account for our incomplete pairs of socks. I realize that most people who speak of a series of minor personal mishaps and their personal bad luck are merely embellishing a funny story or trying to establish a conversational rapport with others and don't necessarily subscribe to their literal pronouncements. Still many people claim that the world is conspiring against them, and a gentle mathematical debunking helps dispel this illusion.

Self-aggrandizement is the key to the appeal of the illusion and of paranoia in general. The feeling is a consequence of the perhaps subconscious inference that if the world is out to get me, I must be pretty important. It's hard for these people to come to grips with the likely fact that almost no one gives a pair of socks about them.

More generally, psychology supplies a number of the girders that help bridge the gap between statistics and stories, Murphy's Law being just one of a number of psychological roadblocks that often obscure our perceptions of chance phenomena and hence, eventually, our views of ourselves. The whole litany of foibles studied by cognitive psychologists is relevant: the availability error, anchoring effect, regression to the mean, confirmation bias, and so on. Of particular relevance in revising our estimates of the likelihood of various events is Bayes' Theorem, of which few people have a good grasp.

A belief in the significance of coincidence is another little-noted, but widespread illusion. This is especially so for world class coincidences, which are sometimes termed miracles. (A proviso is that the result be positive. It's generally not termed a miracle when a rare earthquake levels a building on the only day of the year it is full of schoolchildren.) As with their more mundane versions, the vast majority of "miracles" mean nothing, some point to valuable, yet overlooked connections, and a few suggest a void in our understanding. David Hume's insight about this latter variety of miracle is too little appreciated. Hume observed that every piece of evidence for a miraculous coincidence, i.e., for a violation of natural law, is also evidence for the proposition that the regularities that the alleged miracle violated are not really laws of nature after all.

The most amazing coincidence of all would be the complete absence of all coincidences.

One consequence of the mistaken belief that coincidences are quite special and almost always significant is their rarity in most modern fiction, where introducing one is considered a kind of cheating. We've moved *too* far away from Victorian novelists who regularly inserted outlandish coincidences into their works. If Charlotte Bronte stretched the long arm of coincidence to the breaking point, as was once remarked, most modern writers have reduced it to an unnatural stub incapable of reaching out to a larger world. Coincidences are the ubiquitous stuff of life and leaving them out of a novel or movie makes plot and character development necessarily more deterministic and less life-like.

Stories and Logic

Standard scientific and mathematical logic is termed *extensional* since objects and sets are determined by their extensions (that is, by their members). That is, entities are the same if they have the same members even if they are referred to differently. In everyday *intensional* (with an s) logic, this isn't so. Entities that are equal in extensional logic can't always be interchanged in intensional logic. "Creatures with hearts" and "creatures with kidneys" may refer to the same set of creatures extensionally (that is, all creatures that have hearts may happen to have kidneys and vice versa), but the terms certainly differ in intension or meaning. Likewise, one may promise to arrive in Philadelphia on the day of the wedding, but even though the wedding is President Millard Fillmore's birthday (that is, they're extensionally the same), it would be an odd person who would describe his date of arrival in this alternative way.

Betwixt the two logics lies a gap we cannot ignore. In mathematical contexts, the number 3 can always be substituted for, or interchanged with, the square root of 9 or the largest whole number smaller than the constant π without affecting the truth of the statement in which it appears. By contrast, although Lois

Lane knows that Superman can fly, and even though Superman equals Clark Kent, she doesn't know that Clark Kent can fly, and the substitution of one for the other can't be made. Oedipus is attracted to the woman Jocasta, not to the extensionally equivalent person who is his mother. The perspectives of Lois Lane and Oedipus may be limited, but in the impersonal realm of mathematics, one's ignorance or, in general, one's attitude toward some entity does not affect the validity of a proof involving it or the allowability of substituting equals for equals.

The logic of history is intensional. Take any historical account of a major event and substitute for incidents and entities in it any extensionally equivalent ones that come to mind. The result will likely be humorous or absurd, like substituting Millard Fillmore's birthday for any reference to one's wedding day. ("Ah, one of the happiest days of my life was Millard Fillmore's 172nd birthday.") Our view of the event and, in general, the verdict of history on it depends to an extent on which extensionally equivalent characterization we choose. And which characterization we choose depends on many things, including our particular psychologies, the general historical context, and the history subsequent to the event in question.

More generally, we all sometimes want, believe, expect, fear, or are embarrassed by something without wanting, believing, expecting, fearing, or being embarrassed by something else to which it's, for all (im)practical purposes, extensionally equivalent.

What is the relevance of all this to probability and statistics? As subdisciplines of pure mathematics, their appropriate logic is the standard extensional logic of proof and demonstration. But for *applications* of probability and statistics and the stories in which they play a crucial role, which are what most people mean when they refer to them, the appropriate logic is informal and intensional. The reason is that an event's probability, or rather our judgement of its probability, is almost always affected by its intensional description.

Recall, for example, the choice we have in assigning a likelihood to Waldo's possessing a given characteristic. If we describe him as employee 28-903 in a certain company in country X, 45 percent of whose citizens have a certain characteristic, then we'd be reasonable in assuming that there is a 45 percent probability that Waldo shares this characteristic. But if we describe him as the only person living at a given address and belonging to a certain ethnic group, 80 percent of whose members in the region comprising countries X, Y, and Z have the characteristic in question, then we'd probably conclude that the chances are 80% that Waldo shares this characteristic. And if we describe him as a specific low-level functionary in a nation-X-wide organization only 15 percent of whose members have this characteristic, then we'd likely to state that his chances of having said characteristic are only 15%.

These descriptions are extensionally equivalent; they each specify the same individual, Waldo. Which (combination of) inequivalent intensional descriptions we employ and which we take to be most basic is, to an extent, up to us. Yet it affects our assignments of probability and everything that flows from them.

Intensional logic is an ill-formed and incompletely understood collection of disciplines that includes certain outgrowths of mathematical logic and philosophy (so-called modal logic, situation semantics, inductive logic, and action theory), parts of linguistics, information theory, cognitive science, psychology, and, most important of all, the informal everyday logical intuitions and understandings we all have.

It is more tied to context, perspective, and experience than extensional logic and hence requires the use of indexicals - words such as "this," "that," "you," "now," "then," "here," "there," and last, but certainly not least, "I" and "me." When using it, we must situate the action and the people involved. We must take into account their traits, the people and things they know, and the circumstances they find themselves in. This situating and contextualizing is the analogue of establishing the initial conditions of a scientific law - the height and velocity of a projectile, the temperature and pressure of a gas, etc. However, unlike the case in science where the laws are numerous and the initial conditions are often minor details, in intensional logic, the contexts, connections, and conditions are much more important than the relatively few "laws" of behavior. The hackneyed refrain, "You really had to be there to understand," is often true.

Semantics and Common Knowledge

Moving beyond mathematical logic, a diverse band of scholars has worked to formalize different aspects of the context-bound, self-referential, metaphor-laden, agent-centered, opaquely referential nature of intensional logic. However ill-formed a subject it is, it nevertheless grounds our understanding of mathematical applications, specifically probabilistic and statistical ones (and of narratives and informal discourse as well). These applications are not always as clear-cut and unproblematic as people generally believe.

Many approaches also stress the self-referential aspect of everyday narrative and conversational situations. The notion of "Common ground" or "common knowledge" is particularly important. It is the information well from which each participant in a dialogue draws and to which each contributes. In the usual formulation of this notion, X is an element of information from the common ground occupied by Myrtle and Waldo if Myrtle and Waldo each know X, know that the other knows X, know that the other knows that the other knows X, and so on in a potentially infinite regress. Common ground or common knowledge is an inherently self-referential idea that entails more than two people merely knowing the same bit of information and knowing that the other knows it.

"The Parable of Furious Feminists" is relevant. It takes place in a benightedly sexist village of uncertain location. In this village there are 50 married couples and each woman immediately knows when another woman's husband has been unfaithful but never when her own has. The very strict feminist statutes of the village require that if a woman can *prove* that her husband has been unfaithful, she must kill him that very day. Assume also that the women are statute-abiding, intelligent, aware of the intelligence of the other women, and, mercifully, that they never inform other women of their philandering husbands. As it happens, all 50 of the men have been unfaithful, but since no woman can prove her husband has been so, the village proceeds merrily and warily along. Then one morning the tribal matriarch from the far side of the forest comes to visit. Her honesty is acknowledged by all and her word taken as law. She warns darkly that there is at least one philandering husband in the village. Once this fact, only a minor consequence of what they already know, becomes *common* knowledge, what happens? The answer is that the matriarch's warning will be followed by 49 peaceful days and then, on the 50th day, by a massive slaughter in which all the women kill their husbands. The proof by induction is left to you.

However formalized, the situational and self-referential aspects of ordinary human communication (including communication about probabilities and statistics) help to make story-telling and conversation essential parts of self-building and culture-construction. To communicate with someone it's necessary to empathize with him or her (which, of course, requires positing the existence of a him or her). One must make reference to the necessary cultural and background knowledge, to the common ground or common knowledge of the participants, and to the particular situation at hand. The understandings involved are delicate and evanescent and the requisite knowledge base Brobdingnagian. No computer, for example, has yet passed the Turing test.

The extensional logic of science is, as I've indicated, not adequate for describing this cognitive coupling that plays such a big role in story-telling and conversation. In general, we don't just impart information to one another and then draw static inferences from this information about the external world.

We "dance" with each other and establish a common ground in which the story or the conversation can proceed.

Pure mathematics and its extensional logic allow for, indeed even call for, personal detachment, for standing outside a relationship, a governmental policy, a biological phenomenon, an entire galaxy. Mathematics is *extricative*; it extricates and disentangles us. By contrast, informal intensional logic, whose squishy rules develop out of life itself, tends to involve us with others, to induce us to influence and be influenced by each other, to presuppose both a personal sovereignty and a shared social context. Intensional logic is i*mplicative*; it implicates and entangles us.

This entanglement can happen quickly. Although philosophers warn of the impossibility of a completely private language, semi-private languages are part of the common ground of any two significantly related people and appear in any extended story. How the members of a couple signal their intentions to purchase items and their attitudes toward money would fill a small book. The natural way these understandings arise is illustrated in the following story.

> A young man is on vacation and calls home to speak to his brother.
>
> "How's Oscar the cat?"
>
> "The cat's dead, died this morning."
>
> "That's terrible. You know how attached I was to him. Couldn't you have broken the news more gently?"
>
> "How?"
>
> "You could've said that he's on the roof. Then the next time I called you could have said that you haven't been able to get him down, and gradually like this you could've broken the news."
>
> "Okay, I see. Sorry."
>
> "Anyway, how's Mom?"

Meaning, Information, and Complexity Horizons

Information theory provides a number of oblique perspectives on stories, statistics, and selves. Boris Pasternak's well-known remark, "What is laid down, ordered, factual, is never enough to embrace the whole truth," underlines the necessarily limited information content of any story. It also brings to mind Kurt Godel's famous first theorem about the incompleteness of formal mathematical systems: in any sufficiently rich formal system, there will always be statements which will be neither provable nor disprovable. Greg Chaitin has shown that Godel's theorem follows from the fact that no program can generate a sequence with greater complexity than it itself possesses. (Complexity of a sequence and hence of anything that can be so encoded is defined roughly as the length of the shortest program capapble of generating the sequence.) As Chaitin has remarked, we can't prove a ten pound theorem (generate a very complex sequence) with five pounds of axioms (with a less complex program), and this limitation afflicts any information-bearing entity, human, electronic, or other.

The complexity of our brain and DNA connects information theory to the notion of Self. If we think of DNA as something like a computer program directing the building of an embryo, then rough estimates of the complexity of the embryo program reveal it to be grossly insufficient to delineate the trillions of connections in a human brain. These connections come largely from the experiences of a specific time and culture, and thus a large part of our identity is supplied to us by events outside our skulls.

However information is encoded in the brain, the brain's complexity - its factual knowledge, associations, reasoning ability - is necessarily limited. Once again a rough number - three billion bits has been proposed - can be attached to it, but the existence of the number is here more important than its value. The reason is that any phenomenon in nature that is more complex than the human brain is, by definition, too complex for us to comprehend. Alternatively, we can't make predictions (generate binary sequences) of greater complexity than (the information encoded within) our brains. Regularities may exist which provide a key to understanding the universe, but they may be beyond what I term the human brain's "complexity horizon" (a notion that with time will, I think, gain much greater currency and importance).

Alternatively put, there may be a relatively short "secret to the universe" program, a theory of everything having complexity, say, ten billion bits, that we're just too limited (i.e., too stupid) to understand. Although they differ ineradicably, both traditional religious and scientific approaches to a hoped-for theory of everything share the perhaps naive assumptions that such a theory can be found and that its complexity will be sufficiently limited to be understood by us. Why assume that?

Stories, Analogies, and Physical Entropy

Finally, an interesting correspondence between stories, information, and self is suggested by the notion of "physical entropy," which was defined by the physicist Wojciech Zurek in his "Thermodynamic Cost of Computation, Algorithmic Complexity, and the Information Metric," which appeared in *Nature*, September, 1989. He defined it to be the sum of Claude Shannon's information content, measuring the improbability or surprise inherent in a yet to be fully revealed entity, and Chaitin's complexity, measuring the algorithmic information content of what's already been revealed. The definition was designed to clarify certain classical problems in thermodynamics (in particular that of Maxwell's demon). But physical entropy can also be used to model the human-story system. Imagine two readers encountering a new short story or novel. One is a very sophisticated litterateur, while the other is quite naive. For the first reader, the story holds few surprises, its twists and tropes being old hat. The second naive reader, however, is amazed at the plot, the characters, the verbal artistry. The question is, How much information is in the story?

In attempting to answer this, it makes sense to consider the readers as well, readers who bring vastly different backgrounds to the story. The first reader's mind already has encoded within it large portions of the story's complexity; the second's is relatively unencumbered by such complexity. The Shannon information content of the story, its improbability or the surprise it engenders, is thus less for the first reader whose mind's complexity is, in this regard, greater, while the opposite holds for the second reader. As they read the story, both readers' judgments of improbability or surprise dwindle although at different rates, and their minds' complexity rises, again differentially. The sum of these two, the physical entropy, remains constant and is a measure of the information content of the mind-story system.

There are three aspects of this admittedly vague speculation that I like. One is that the notions involved are in the same conceptual ballpark as the second law of thermodynamics, which C.P. Snow used to illustrate the gulf between scientific elites and literary ones, the latter presumed not to understand the significance of the second law. Since this is, in part, also a more oblique look at the gap between the two cultures, this historical resonance is satisfying.

More importantly, the speculation above provides a sense (not incompatible, for example, with some neuroscientists' theories or even with Hamlet's metaphorical utterance "Within the book and volume of my brain") in which stories become a physical part of us. They become encoded somehow into fragments of mental routines capable of generating them at will, and, if integrated into our conceptual and emotional maps of the world, they change us forever. We are the stories we tell.

The third appealing aspect of this line of thought is that it seems to underline the extent to which context (in this case the human-story system rather than just the story) is needed when making assessments. A story makes no sense to people who don't have any of the relevant linguistic and psycho-social understandings that it presumes. Without a scientific and cultural matrix which supports these essential, but implicit understandings, theories and stories are meaningless.

Perhaps this is what postmodernist literary theorists mean when they refer to "the death of the author." Their reluctance to accept the author's view of his work as definitive (or in extreme formulations, as even very important) may indicate a realization (or an overestimation) of the fact that meaning is a socially-mediated phenomenon made possible by common cultural understandings. Wittgenstein once remarked aptly, "That Newtonian mechanics can be used to describe the world tells us nothing about the world. But it does tell us something - that it can be used to describe the world in the way in which we do in fact use it." The same might be said about Jamesian novels or Seinfeldian situation comedies.

Bridging the Gap

I'll close with the following little interchange between George and Martha.

> GEORGE: This talk of motives, promises, questions, fears, and wishes is sloppy. Why don't we simply cite facts, use extensional logic, do mathematics and forget all this messy stuff?
>
> MARTHA: I know what you mean. Let's just vow to each other to do that from now on. We both yearn for clarity and precision.

The joke, such as it is, is that George and Martha are planning to use only extensional logic, but intensional notions are built into the very fabric of their (and our) communication. The situation is rather like shouting about the importance of silence or claiming that one's brother is an only child. Although wishes, fears, promises, and motives are not (yet) the province of mathematics and science, they and the stories in which they're embedded are an essential ground for understanding these subjects and their applications.