

Some 19th century Arguments for the Rational Assignment of Probabilities for Possible events in Nature

Byron Wall
York University, bwall@yorku.ca

Abstract—English:

Countless decisions are made every day by each of us individually and collectively through our governments and other institutions, about what actions to take in the present in order to optimize a future in which many possible outcomes are more than moderately uncertain. At a personal level, we make these decisions intuitively, based on past experience. At the institutional and government level, we increasingly rely upon quantitative statistical projections and risk assessments. A great deal of interesting and well-worked out mathematics goes into these projections. Most of the mathematics is based upon models in which probabilities can be specified with precision. But the usefulness and reliability of these models depends crucially upon how well the tidy world of the model compares to an incompletely understood Nature. The history of probability and statistics is peppered with arguments, sometimes vociferous, over the assignment of a probability to events in Nature, both those that are agreed to be highly probable, such as whether the sun will rise tomorrow, and those that are deemed highly improbable, such as what the chances are of snow in

July, or living to the age of 200, or invasion from outer space. In the late 19th century, these arguments were carried on by respectable mathematicians and philosophers who were seeking to find solid ground for inference from incomplete information, the basis of statistics. This paper explores some of that debate.

Abstraite—Français:

De nombreuses décisions sont prises chaque jour par chacun de nous individuellement et collectivement par l'entremise de nos gouvernements et des autres institutions, au sujet des mesures à prendre dans le présent afin d'optimiser un avenir dans lequel de nombreux résultats possibles sont significativement incertains. Au niveau personnel, nous prenons ces décisions de manière intuitive, basée sur l'expérience. Au plan institutionnel et au niveau du gouvernement, nous nous appuyons de plus en plus sur des projections de données statistiques et des évaluations des risques. Beaucoup de travail intéressant et bien travaillé sur les mathématiques se trouve dans ces projections. La plupart des mathématiques est fondée sur des modèles dans lesquels les probabilités peuvent être spécifiées avec précision. Mais l'utilité et la fiabilité de ces modèles dépendent essentiellement de la façon dont le monde ordonné du modèle se compare aux forces naturels mal-compris. L'histoire de la probabilité et des statistiques est truffée d'arguments, parfois bruyantes, au sujet

d'accorder une probabilité aux évènements dans la nature, aussi bien ceux qui sont acceptés d'être hautement probables, comme si le soleil se lèvera demain, et ceux qui sont jugés très improbable, comme les chances d'une chute de neige en Juillet, ou vivre à l'âge de 200 ans, ou une invasion venant de l'espace extra-atmosphérique. À la fin du 19e siècle, ce débat a été mené par des mathématiciens et des philosophes respectés qui cherchaient un terrain solide pour l'inférence de la base de statistiques. Cet article explore une partie de ce débat.

The text:

The traditional histories of probability theory start with the correspondence between Blaise Pascal and Pierre de Fermat over the celebrated Problem of Points. According to this tradition, the gambler, the Chevalier de Méré posed the problem to Pascal in 1654, who then wrote to Fermat, and between the two of them, a satisfactory solution was reached, which then marked the beginning of the mathematical theory of probability. The problem concerns the fair division of prize money between two players who are interrupted before the game they are playing can reach the normal specified conclusion of the game. Though the Pascal-Fermat correspondence of 1654 is the agreed upon official starting point for probability theory, this very problem in slightly different format had been discussed by mathematicians before. Cardano and Tartaglia, for

example had both written about it, though without reaching the solution proposed by Pascal and Fermat.¹

The details of the game do not matter so much as the reasoning of Pascal and Fermat, but it is perhaps worth mentioning the basic format: the game is played in rounds, at which a point is gained by one player or the other, and the game is over when one player has accumulated a specified number of points. Each player is deemed to have the same likelihood of winning a round as the other player. Hence the game is either one of skill between two equally matched players, or is purely aleatory game, such as dice. In the traditional version, there are two players, A and B . When the game is interrupted, player A needs to gain a more points to win and player B needs b more points. Hence the game can go at most $a + b - 1$ further rounds. Pascal and Fermat together came to a resolution amounting to the following: A list of all possible future outcomes has size 2^{a+b-1} . The fair division of the stake will be the proportion of these outcomes that lead to a win by A versus the proportion that lead to a win by B .

That it is *this* solution to this essentially colourless game that set the stage for the development of probability theory is significant for a number of reasons. First, the stated premises of the game make it clear that this is an analysis of blind luck. The game might be one of exacting skill, such as a series of chess games, but the specifications are that the players are to be treated as of equal

skill, so we might just as well be tossing coins. Each outcome is treated as equally likely, and the essential mathematical analysis is that of counting up permutations and combinations. Probability theory is built upon a fundamental set of equally probable outcomes.

A typical historical question might be, why did this analysis spring forward in the middle of the 17th century? Games of chance have been around since the dawn of human civilization and the very problem that Pascal and Fermat discussed, the Problem of Points had first appeared in print in 1494, but a viable solution was not proposed until Pascal and Fermat did so in 1654. Was there some added insight that was not apparent until the middle of the seventeenth century?

Ian Hacking reports the suggestion that if the existence of equiprobable outcomes was not generally recognized, a theory built upon them would not be proposed.² There are all manner of reasons why earlier civilizations might not have considered the results of games of chance to have equiprobable outcomes. For example, an early precursor of dice was the *talus*, a knucklebone or heel bone that can land in any of four different ways. But each talus was different and the likelihood of any given talus landing in each possible way with the same frequency was remote. Certainly, it was not taken for granted that it would do so.

Moreover, if the outcome of any indeterminate result was relegated to Fate, then there would be no incentive to create a theory to understand outcomes that were viewed as controlled elsewhere. It is then perhaps understandable that probability theory is a child of the Scientific Revolution.

Be that as it may, probability theory began its development at just the same time that mathematical models came into their own to describe the everyday operation of things on Earth and the objects in the sky. The Scientific Revolution was followed by the Enlightenment, when humanity was encouraged to think that with the application of reason, all the puzzles of the world could be solved.

The standard probability model which gives us the normal distribution was developed first for dealing with astronomical observations. Error theory, as it was called, sought to determine the best possible estimate of the true value of, say, the position of a star or planet given a number of not identical observations by trained astronomers. The essential notion is that there is a "true" value for the position in question and applying mathematical analyses to the reported values will enable astronomers to come closest to that value.

Some of the same mathematicians who applied themselves to the questions of optimal values for astronomical measurements also used their mathematical approach to optimize other matters of measurement in civilized

life. Among the most important of these was Pierre Simon de Laplace. Laplace is memorable for having relegated God to the role of a stand-in for unknown causes – a hypothesis, he told Napoleon, for which he had no need, since he was confident that the laws of Nature, when fully understood, would account for everything.³ And, while Laplace did not require the hypothesis of God to complete his worldview, he was willing to entertain the metaphor of an omniscient being to drive home his thoroughly deterministic world view as the foundation of his conception of probability. For example, in his *Essay of Probabilities*, Laplace intones:

Given for one instant an intelligence which could comprehend all the forces by which nature is animated and the respective situation of the beings who compose it—an intelligence sufficiently vast to submit these data to analysis—it would embrace in the same formula the movements of the greatest bodies of the universe and those of the lightest atom; for it, nothing would be uncertain and the future, as the past, would be present to its eyes.⁴

Clearly, for Laplace, probability was a measure of one's relative knowledge or ignorance of the true state of affairs, since everything was fully determined from the get-go. But as is also clear from Laplace's willingness to dispense with God, he was quite confident that the intellectual tools at his disposal, particularly

mathematics, would enable humankind to ascertain the truth with a level of certainty that was simply not available to earlier, less enlightened generations.

An example of the value of mathematical analysis to determine the truth in matters of human affairs is the application of probability analyses to jurisprudence. A crime is committed and someone is arrested and charged with the crime. A trial ensues. At the trial witnesses are called, evidence is produced. Are the witnesses reliable? Is the evidence conclusive? Does mathematics have any bearing on determining the truth in such situations? A number of people thought so, beginning with Gottfried Leibniz in the 17th century. In the 19th century, Laplace eagerly carried on this tradition of mathematically assisted justice with his studies of the reliability of the jury system.

In France, trial by jury underwent frequent and extensive review in the years following the French Revolution. In 1790 the size of a jury had been fixed at twelve, but there was considerable disagreement over what level of agreement was required for a conviction. Various levels were set from time to time, beginning with a majority of 10 required. There were many changes, including a simple majority, i.e. 7:5, and then 2/3, i.e. 8:4. By 1836, the simple majority was re-established.⁵ Laplace was drawn into the question of the optimal jury configuration in particular because he was convinced that the jury deciding by a simple majority was far too prone to giving an incorrect verdict.

Common sense would suggest that a jury that decides by a majority of seven to five on the guilt or innocence of a suspect is likely to be in error more often than juries that were unanimous in their decision, or even just had a greater majority, but is it really possible to make this more precise? Laplace thought so, and the fact that he attempted to do so is strong evidence that Laplace believed that the sort of exactitude, precision, and determinism that he ascribed to the universe as a whole was, in principle, available to humankind, if not immediately, then just around the corner.

What seems extraordinary are the blithe assumptions that preceded the mathematical analysis. For example, Laplace began with an assumption that the probability of the guilt of an accused, any accused, is $\frac{1}{2}$. This appears to be on the basis of the principle of indifference—that the *a priori* probability of any unknown matter is $\frac{1}{2}$ until some *a posteriori* information is available. Recall that the solution to the Problem of Points proposed by Pascal and Fermat also stipulated that the players were of equal skill, so that the outcome of future rounds was equally likely to be a win for either player, or alternatively, that the game they were playing was one of no skill whatsoever, so that a probability of $\frac{1}{2}$ could be assumed. It does make it ever so much easier to calculate probabilities if such a simplifying assumption can be made.

We go on. Laplace declared that the reliability of a juror is somewhere

between $\frac{1}{2}$ and 1. What's the logic here? Well, here we start with an assumption that human judgment is worth something, and that, on average, a juror is at least as likely to get it right as a coin toss. Hence, the lower bound on a juror's reliability is that of the coin toss coming up heads, namely $\frac{1}{2}$. Remember that this is a lower bound for the *average* juror. Actually, that is my interpolation as to what Laplace must have been thinking. Surely, I am thinking, Laplace could imagine a juror whose instincts were so bad that his decisions could well be worse than that of a coin toss. But perhaps I am anachronistically reading in an interpretation from the age of psychology into the age of rationalism. Hence, I take that back; indeed, I think Laplace thought that *any* juror was, by being a rational human being, at least no worse than random. In any case, Laplace takes the lower bound to the unknown juror's reliability as $\frac{1}{2}$. The upper bound is absolute infallibility, represented by the upper bound probability of 1. So that sets the outer limits. Now what?

In order to go on from this without complicating the calculations unbearably, Laplace needs some further information as to how the reliability of the jurors is distributed between that of a coin toss and that of an omniscience generally associated with the sort of God Laplace had no need for. Once again, I am tempted to speculate anachronistically on how one might suppose this reliability trait to be distributed among these hypothetical jurors. Here, I am led

to the idea of the normal distribution of error theory, as used in astronomy. Surely, the astronomer Laplace would imagine the juror's attempts to discern the true verdict to be akin to the astronomical observer attempting to record the true position of a star. Therefore Laplace would imagine a mean value in the vicinity of a probability of $\frac{3}{4}$ with a nice, symmetrical bell curve descending from both sides of that figure, with 68% of the judgments falling within one standard deviation to either side of that mean value. Or so I might speculate.

But that again would be to apply a notion from a later age. The idea that human attributes themselves fell into predictable, normal distributions was articulated just a couple of decades later by the astronomer Adolphe Quetelet, about whom more later. Laplace takes an even simpler notion, namely that the reliability of the jurors is uniformly distributed between the boundary values of $\frac{1}{2}$ and 1. By what right does Laplace seize on this simplifying assumption? None, really, except the principle of indifference once again, a sort of presumptive Occam's Razor for probability theory: when no information is available, choose the probability that makes calculations easiest.

Pressing on, Laplace now works only with the average reliability, which certainly made the calculations easier. Among his conclusions are that a unanimous jury panel of n members has a chance of being wrong equal to $(\frac{1}{2})^{n+1}$. Ian Hacking has commented that "no tidier example of an *a priori* rabbit

out of a hat can be imagined.”⁶

Using these tools, Laplace then proceeded to announce the chances of a jury being in error, varying with the size of the jury panel and the split of votes.

His results look like this:

For a jury that divides:	The chance of error is:
12:0	$1/8192$
9:3	About $1/22$
8:4	About $1/8$
7:5	$2/7$
5:3	About $1/4$
9:0	$1/1024$
112:100	About $1/5$
501:500	About $1/2$

You can imagine the effect of such calculations and announcements of results with such precision. If you were an enthusiast for the Enlightenment and

an optimist about human capacity for plumbing all the mysteries of the world and wrestling them to the ground with exact reasoning, then maybe you found this line of reasoning convincing. Even if you did not agree with every assumption, you might view this approach as a promising first approximation: even matters of uncertainty and fickle human nature could be wrestled to the ground and contained within a confidence interval if aggregated data is available. This is what the jury system is all about, of course, and these numbers that Laplace has calculated merely express that with some precision.

On the other hand, if, like Ian Hacking, you viewed these calculations as “an *a priori* rabbit out of a hat,” then you would want to go no further down this road, but instead return to first principles. Indeed the response to Laplace’s pronouncements did divide rather sharply along these lines, and in a rough way, that division also characterized the different approaches of Continental Europe and the British Isles.

Laplacian analysis provides yet another fine illustration of the fundamental rift in epistemology in general and science in particular that is so well captured by my favourite commentary on the scope of ancient Greek philosophy, the fresco *The School of Athens*, by Raphael in the Vatican. The fresco includes figures from the whole range of ancient Greek culture: philosophers, mathematicians, scientists, playwrights, architects, and artists. In the centre of

the fresco are the figures of Plato and Aristotle, peripatetically walking and discussing philosophy. Plato, the teacher, gestures to the heavens, while Aristotle, his student, reaches forward, indicating the world around. The point of this characterization is that Plato argued that true knowledge was to be had only through an understanding of the abstract, unchanging and eternal Forms, which to a Renaissance Christian like Raphael meant what was in the heavens. The Forms included the objects of mathematics and mathematical reasoning was the stepping stone to apprehension of the Forms. Aristotle deeply distrusted this approach, holding that reasoning that is not grounded in sense experience of the world around us can end up reaching false conclusions. In consequence, Aristotle had little interest in mathematics.

In the case at hand, Laplace's reasoning would fit the Platonic model and serve as an excellent example of how abstract, quantitative reasoning and calculation can clarify and provide a more exact measure of the relative uncertainty that the jury system tries to minimize. And, correspondingly, those who object to Laplace's approach from the outset voiced objections along the lines that Aristotle had to his master's approach.

There was a chorus of Aristotle-like complaints from Britain to the trend in *a priori* assignments of probabilities in the works of Laplace and others who followed a similar line of reasoning that usually began with the arbitrary

assignment of the probability of $\frac{1}{2}$ to the occurrence or non-occurrence of some event or condition about which nothing at all is known.

One of the most pointed critics of Laplace's approach was Robert Leslie Ellis, a mathematician of great promise in the mid-19th century, whose potential was not realized due to a debilitating disease which halted his serious academic work at the age of 32 and took his life ten years later. Ellis was born in 1817, entered Trinity College, Cambridge, in 1836, graduated as Senior Wrangler in 1840, and immediately became a Fellow of Trinity College. During his short career, he was the Editor of the *Cambridge Mathematical Journal*, and Moderator, i.e., principal mathematical examiner, for Cambridge University. He wrote 6 papers on probability, mostly on technical details, but one, which has attracted the most notice, was a theoretical paper entitled "On the Foundations of the Theory of Probabilities," read to the Cambridge Philosophical Society in 1842.⁷

In that paper, Ellis specifically attacked the principle of insufficient reason invoked by Laplace and others, accusing them of circular reasoning from Bernoulli's Law of Large Numbers. Bernoulli's theorem states that if an event has probability p of a certain outcome and can be subjected to n trials, each of which is independent of each other, then the expected occurrence of the particular outcome after n trials, is pn , and for any arbitrary small number ϵ , the

probability that the actual frequency of outcome of the event is within the interval $p \pm e$ increases with n , and approaches 1 as n approaches infinity. So, for example, in a coin toss, the greater the number of trials, the smaller will be the percentage difference of the number of heads tossed out of all tosses from exactly 50%. Ellis's point is that Bernoulli's law starts with a known probability of the outcome, p , and deduces the convergence of the actual results to the expected value. But, he says, in the actual case of coin tosses, our reason for asserting that the probability of heads is $\frac{1}{2}$ is because we have experience of tossing coins many times and finding that they do, on average, tend to come out heads half of the time. In other words, it is the frequency of the long-term results obtained that has led us to the assertion that the chances of heads or tails are 50:50, not the pre-knowledge of the probability that leads to the assertion of the outcome. The argument is made that the symmetry of the coin, or the die, leads directly to the assertion of equiprobable outcomes, and that all the other factors involved in a toss—the spin given the coin or die, the height thrown, wind disturbance, etc.—will cancel each other and lead to one outcome as often as to another. Hence, the generalization, which Ellis objects to, of saying that whenever we don't have a reason to favour one outcome over another, they have equal probability.

But taking cases where it is generally agreed upon that there is equal

probability of one result over another in a single trial does not immediately reveal the problem, since the expected aggregate outcome will be the same regardless of the train of reasoning. It is more revealing to take a case where the *a priori* assignment of probability can lead to different results. Take the case of an induction which combines some *a priori* assumptions with some *a posteriori* information. An extension of the principle of indifference that was common in the 19th century was this: on m occasions, an event x has occurred. The presumptive probability that the next trial will produce x is given as $(m+1)/(m+2)$. This is the sort of reasoning applied to the question of whether the sun will rise tomorrow, given that it has risen every known day of human history. Aside from whether the application of this formula makes any sense whatsoever, it is at least fairly unambiguous what is meant by the sun rising.

What if it is an open question whether an event counts as the next occasion of the sort in question? Ellis gave the following example: 10 vessels sail up a river, all have flags. The presumptive probability that the next vessel will have a flag is then 11/12. But suppose all the vessels were of one type. Ellis gave the example of Indiamen, a class of merchant ship. Do we then assign this probability to the next vessel no matter what kind, Ellis asked, or only if it is of the same kind as the previous 10 vessels? If the probability is 11/12 for any kind of vessel, it cannot also be 11/12 for the next of the same kind. Ditto if all the

previous flags had been of one colour. Then the probability that the next will be any kind of flag would be the same as the probability that it will be of the same colour. If this is true then it must be impossible for the flag to be any other colour, that is, there would be zero probability for that event.

These kind of examples show the arbitrariness and inconsistency once the notion of a numerical probability gets extended beyond the artificial world of gaming, where the operative probabilities are built into the game through physical symmetries and mutually exclusive outcomes. But Nature is not so conveniently organized.

The criticism of Ellis was echoed, with slightly different emphases, by a host of other British mathematicians and philosophers, including George Boole, and John Stuart Mill. Mill particularly was incensed by the hocus pocus of assigning a probability value in situations where we knew nothing at all. In such situations, Mill said, "to attempt to calculate chances is to convert mere ignorance into dangerous error by clothing it in the garb of knowledge." ⁸

All these examples, both the assertions by Laplace and similar ones by others of the same persuasion, and the rebuttals by British empiricists, concerned the assertion or denial of *a priori* probabilities to situations where there was not an existing body of data. The debate here is whether one gains anything by attributing a probability value to situations of ignorance. But just as

these debates are taking place, statistical data was beginning to be collected, and people were beginning to make assertions about future outcomes on the basis of trends seen in existing data.

The most illustrious example of this is that of the astronomer/sociologist Adolphe Quetelet, whose *Treatise on Man* established the practice of drawing conclusions about all manner of characteristics of human beings on the basis of collected statistics. Quetelet's work was published in French in 1835 and then translated into English in 1842.⁹ It caused a huge stir because of its conclusions. Quetelet's thesis was that both in physical characteristics and in behaviour, human beings closely resembled each other. To Quetelet, the average value for any statistic on humanity that he collected was, as it was in astronomy, the best estimate of the "true value" for that statistic. Quetelet had the idea that the variations that exist in the human frame (and likewise in human behaviour) were "errors" from Nature's model. Just as the astronomer's reported observations of star positions were treated as attempts to report the true position of said star and any deviations were errors, Quetelet viewed the variations in human statistics from the mean as the copying process in Nature missing the mark somewhat and not quite getting it right. It was no accident that Quetelet became the Royal Astronomer of Belgium. The mathematical training he received on the way to becoming an astronomer is what he put to use on the data he collected

on human statistics.

For Quetelet, all attention was focused on the apex of the bell curve, because that was our best estimate of the blueprint for humanity. Here is Quetelet's explanation of his focus:

The *social* man, whom I here consider, resembles the centre of gravity in bodies: he is the centre around which oscillate the social elements—in fact, so to speak, he is a fictitious being, for whom every thing proceeds conformably to the medium results obtained for society in general. It is this being whom we must consider in establishing the basis of social physics, throwing out of view peculiar or anomalous cases, and disregarding any inquiry tending to show that such or such an individual may attain a greater or less development in one of his faculties.¹⁰

It is important here to note the import of the last sentence, namely that outliers are to be disregarded. For Quetelet, the exceptions are Nature's blunders, and they tell us nothing about Nature's plan. Hence the grossly overweight, the anorexic, the giant, the midget, the genius, and the mentally defective are all to be ignored. They are in effect accidents and not important. It's eerily reminiscent of the passage in Aristotle's *Physics*¹¹ where Aristotle avers that to understand Nature we should ignore all things that happen randomly or coincidentally since they do not repeat. Instead we should focus on

what happens regularly and repeatedly, since this is how things happen purposefully in Nature. By such reasoning, Aristotle effectively brushed aside the notion of evolution by natural selection, which had been suggested by Empedocles in the fifth century BCE.

Quetelet's focus on the mean led to a great deal of attention being paid to human characteristics and much debate over whether we are all cast from a mould that predetermines not just our physical characteristics, but also our behaviour. The debate centered on the implications for determinism and freedom of the will. On the one hand, it did establish that assertions about the most likely values for human data should be based upon extrapolations from existing data about real people, not from *a priori* calculations on philosophical grounds. And what was true of human statistics began to spread to all statistics. Forecasts of the future would be based upon data collected about the past. But then clearly the better forecasts would result from situations where there was the most comparable data. Quetelet's emphasis on the mean and his outright assertion that outliers weren't worth worrying about led to fewer data being assembled for the rare events.

How then, would probabilities be calculated for events that are unlikely, but do happen? This is what I hope to discover during my next sabbatical leave, starting this coming January. What I expect to find is that the probabilities that

get assigned to rare events are simply those that are predicted by a normal distribution for outliers that would be out in the tails of the distribution. This is another form of *a priori* pronouncement. Events that we know the least about, because they happen only rarely, will be given a precise numerical probability and that probability will be used, alongside the better established probabilities for which there exists supporting data.

Somehow, in the late 19th century or in the early 20th century, the pattern was established for calculating probabilities for rare events, and, I suspect, that pattern has not materially changed. The calculated probabilities would be used, for example, to establish cost benefit analyses for public projects—to build or not to build levees in New Orleans capable of withstanding a Category 5 hurricane. Or to establish insurance premiums against certain rare perils. Or to evaluate the likelihood of cascading collapses of ill-secured mortgages bringing down the financial system. We rely upon the evaluation of the likelihood of these hard to predict events to protect us from undue risk. Recent experience suggests that these probabilities have not been evaluated correctly and as a result, we live in a world that faces more uncovered risk than it might otherwise face.

¹ Almost any history of probability tells the story of the Pascal-Fermat correspondence on this game. A convenient summary is in Ian Hacking's *The Emergence of Probability: A Philosophical Study of Early Ideas About Probability, Induction and Statistical Inference*, 2nd ed. (Cambridge: Cambridge University Press, 2006), Chapter 7, pp. 57-62.

² Hacking reports this idea, though he then disagrees with it. *Ibid.*, pp. 3-4.

³ Cited in many places, e.g. C. B. Boyer, *A History of Mathematics*, 2nd ed. (New York: Wiley, 1968), p. 538.

⁴ Laplace, *A Philosophical Essay on Probabilities* [orig. 1814], trans F. W. Truscott and F. L. Emory [New York, 1951], p. 3. Quoted in Ian Hacking, *The Taming of Chance* (Cambridge: Cambridge University Press, 1990), pp. 11-12.

⁵ Hacking, *Taming of Chance*, p. 91.

⁶ *Ibid.*, p. 92.

⁷ R. L. Ellis, "On the Foundations of the Theory of Probabilities," *Transactions of the Cambridge Philosophical Society*, 8, pt. 1:(1844) 1-6. Read 14 February 1842.

⁸ J. S. Mill, *System of Logic*, cited in Theodore M. Porter, *The Rise of Statistical Thinking, 1820-1900*, (Princeton: Princeton University Press, 1986), pp. 82-83.

⁹ Adolphe Quetelet, *Sur L'Homme et le Développement de ses Facultés* (1835). Translated as *A Treatise on Man and the Development of his Faculties* (Edinburgh: William and Robert Chambers, 1842).

¹⁰ Quetelet, *Treatise on Man*, p. 8.

¹¹ Aristotle, *The Physics*, Book II Ch 8, 198b 17-33.