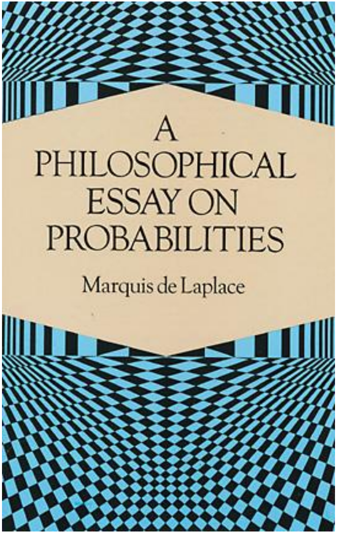
Lecture 5
Laplace and the Classical Theory

Patrick Maher

Philosophy 517
Spring 2007

The book



First edition 1814

English translation 1902

This philosophical essay on probabilities is the development of a lecture on probabilities which I delivered in 1795 . . . I present here without the aid of analysis the principles and general results of this theory, applying them to the most important questions of life, which are indeed for the most part only problems of probability. (p. 1)

Rejection of physical probability

All events are determined (p. 3)

All events, even those which account of their insignificance do not seem to follow the great laws of nature, are a result of it just as necessarily as the revolutions of the sun. In ignorance of the ties which unite such events to the entire system of the universe, they have been made to depend upon final causes or upon hazard, according as they occur and are repeated with regularity, or appear without regard to order; but these imaginary causes have gradually receded with the widening bounds of knowledge and disappear entirely before sound philosophy, which sees in them only the expression of our ignorance of the true causes.

Final causes = purposes; hazard = chance.

Probability is inductive (p. 6)

The curve described by a simple molecule of air or vapor is regulated in a manner just as certain as the planetary orbits; the only difference between them is that which comes from our ignorance.

Probability is relative, in part to this ignorance, in part to our knowledge. We know that of three or a greater number of events a single one ought to occur; but nothing induces us to believe that one of them will occur rather than the others. In this state of indecision it is impossible for us to announce their occurrence with certainty. It is, however, probable that one of these events, chosen at will, will not occur because we see several cases equally possible which exclude its occurrence, while only a single one favors it.

The biased coin

Let us take the game of heads and tails, and let us suppose that it is equally easy to throw the one or the other side of the coin. Then the probability of throwing heads at the first throw is 1/2 and that of throwing it twice in succession is 1/4. But if there exist in the coin an inequality which causes one of the faces to appear rather than the other without knowing which side is favored by the inequality, the probability of throwing heads at the first throw will always be 1/2; because of our ignorance of which face is favored by the inequality . . . But in this same ignorance the probability of throwing heads twice in succession is increased. Indeed this probability is that of throwing heads at the first throw multiplied by the probability that having thrown it at the first throw it will be thrown at the second; but its happening at the first throw is a reason for belief that the inequality in the coin favors it. (pp. 56–57)

Quantitative treatment of the biased coin (p. 57)

In order to submit this matter to calculus let us suppose that this inequality increases by a twentieth the probability of the simple event which it favors. If this event is heads, the probability will be $\frac{1}{2}$ *plus* $\frac{1}{20}$*, or* $\frac{11}{20}$*, and the probability of throwing it twice in succession will be the square of* $\frac{11}{20}$*, or* $\frac{121}{400}$*. If the favored event is tails, the probability of heads will be* $\frac{1}{2}$ *minus* $\frac{1}{20}$*, or* $\frac{9}{20}$*, and the probability of throwing it twice in succession will be* $\frac{81}{400}$*. Since we have at first no reason for believing that the inequality favors one of these events rather than the other, it is clear that in order to have the probability of the compound event heads heads it is necessary to add the two preceding probabilities and take the half of their sum, which gives* $\frac{101}{400}$ *for this probability, which exceeds* $\frac{1}{4}$ *by* $\frac{1}{400}$*.*

- $\frac{11}{20}$, $\frac{121}{400}$, $\frac{9}{20}$, $\frac{81}{400}$: These depend on facts about the world and aren't relative to evidence, hence are pp's.
- $\frac{101}{400}$: This is relative to evidence and doesn't depend on facts about the world, hence is ip.

The rule of succession

Laplace's statement (p. 19)

An event having occurred successively any number of times, the probability that it will happen again the next time is equal to this number increased by unity divided by the same number, increased by two units.

Clarifications (by me)

- The rule says: If an event has occurred n times in a row, the probability that it will happen the next time is

$$\frac{n+1}{n+2}.$$

- Laplace is assuming that there is no relevant evidence other than what is stated.
- The name "rule of succession" comes from John Venn (1888).

Example (p. 19)

Placing the most ancient epoch of history at five thousand years ago, or at 1826213 days, and the sun having risen constantly in the interval at each revolution of twenty-four hours, it is a bet of 1826214 to one that it will rise again tomorrow. But this number is incomparably greater for him who, recognizing in the totality of phenomena the principal regulator of days and seasons, sees that nothing at the present moment can arrest the course of it.

The probability of the sun rising tomorrow is

$$\frac{1826213 + 1}{1826213 + 2} = 0.999999$$

What Laplace assumes to derive the rule

- *When the probability of a single event is unknown we may suppose it equal to any value from zero to unity.* (p. 18)
- Notation (by me):
  - H_x: Hypothesis that the unknown probability of the event is x.
  - E: Evidence that the event has occurred n times in a row.
  - F: The event will occur next time (the future event).
- Laplace assumes:
  (i) Ip's of the H_x are initially uniform, for $x \in [0, 1]$.
  (ii) $ip(E|H_x) = x^n$.
  (iii) $ip(F|E.H_x) = x$.

The "unknown probability" is a fact about the world (we learn about it empirically) and not relative to evidence; hence it is a pp. Also (ii) and (iii) make sense if it is a pp, not otherwise.

Questions

1. What was Laplace's (official) view about the nature of probability? What was his argument for this?
2. Is Laplace's treatment of the biased coin is consistent with his official view about the nature of probability? Explain.
3. Is Laplace's derivation of the rule of succession consistent with his official view about the nature of probability? Explain.

Measurement of probability

The rule (pp. 6–7)

The theory of chance consists in reducing all the events of the same kind to a certain number of cases equally possible, that is to say, to such as we may be equally undecided about in regard to their existence, and in determining the number of cases favorable to the event whose probability is sought. The ratio of this number to that of all the cases possible is the measure of this probability, which is thus simply a fraction whose numerator is the number of favorable cases and whose denominator is the number of all the cases possible.

For equally possible cases:

$$\text{Probability} = \frac{\#\text{ of favorable cases}}{\#\text{ of possible cases}}.$$

Example (p. 8)

Let us suppose, for example, that we have three urns, A, B, C, one of which contains only black balls while the two others contain only white balls; a ball is to be drawn from the urn C and the probability is demanded that this ball will be black. If we do not know which of the three urns contains black balls only, so that there is no reason to believe that it is C rather than B or A, these three hypotheses will appear equally possible, and since a black ball can be drawn only in the first hypothesis, the probability of drawing it is equal to one third. If it is known that the urn A contains white balls only, the indecision then extends only to the urns B and C, and the probability that the ball drawn from the urn C will be black is one half. Finally this probability changes to certainty if we are assured that the urns A and B contain white balls only.

General principles of the calculus of probabilities:

First principle (p. 11)

The first of these principles is the definition itself of probability, which, as has been seen, is the ratio of the number of favorable cases to that of all the cases possible.

Second principle (p. 11)

But that supposes the various cases equally possible. If they are not so, we will determine first their respective possibilities, whose exact appreciation is one of the most delicate points of the theory of chance. Then the probability will be the sum of the possibilities of each favorable case.

Example for the second principle (pp. 11–12)

Let us suppose that we throw into the air a large and very thin coin whose two large opposite faces, which we call heads and tails, are perfectly similar. Let us find the probability of throwing heads at least one time in two throws ...

We can count at this game only three different cases, namely, heads at the first throw, which dispenses with throwing a second time; tails at the first throw and heads at the second; finally, tails at the first and second throw. This would reduce the probability to 2/3 if we should consider with d'Alembert these three cases as equally possible. But it is apparent that the probability of throwing heads at the first throw is 1/2, while that of the other two cases is 1/4 ... If we then, conforming to the second principle, add the possibility of 1/2 of heads at the first throw to the possibility of 1/4 of tails at the first throw and heads at the second, we shall have 3/4 for the probability sought.

"Possibility" is just another word for probability

1. The second principle implies that the probability of a case equals its possibility.
2. In the example to illustrate the second principle, Laplace interchanges "probability" and "possibility."
    - *the probability of throwing heads at the first throw is 1/2*
    - *the possibility of 1/2 of heads at the first throw* (both p. 12)
3. He interchanges "probability" and "possibility" in different statements of the second principle.
    - *the sum of the possibilities of each favorable case* (*Essay* p. 11)
    - *the sum of the probabilities of each favorable case* (1886 p. 181)

The classical theory

P1 as a definition

- Commentators generally accept that P1 really is a definition. The concept that it defines is called *classical probability*.
- If this is right, Laplace would be best construed as advocating classical probability as an explicatum for inductive probability.
- This is *the classical theory of probability*; and Laplace is taken to be its leading exponent.
- Standard criticisms of the classical theory:
  - The definition is circular, since "equally possible" just means "equally probable."
  - Inductive probabilities exist in many situations that can't be reduced to equally probable cases. Hence the concept of classical probability is too narrow.
- Laplace is also criticized for not adhering to his own definition, since he talked about probability in situations lacking equally probable cases.

Defects of this interpretation of Laplace

1. We have to suppose that Laplace was unaware he was giving a circular definition, which it implausible.
2. We have to suppose that Laplace was unaware he wasn't adhering to his own definition. This is implausible, especially since he allows that the cases might not be equally possible immediately after stating P1.
3. In other places where Laplace stated P1 he didn't say it was a definition.
    - *Essay* pp. 6–7.
    - *The Analytical Theory of Probabilities* (1886) p. 181.

These is strange if he regarded it as a definition.

P1 as a rule

- If P1 is not a definition, it would be a rule that inductive probability is supposed to obey. It concerns the explicandum, rather than defining an explicatum. And it is correct.
- Advantages of this interpretation:
  - Laplace didn't give a circular definition.
  - Laplace wasn't being inconsistent in applying probability to situations lacking equally probable cases.
  - It is understandable why Laplace sometimes states P1 without calling it a definition.
- Problems with this interpretation:
  - Laplace did (once) say that P1 is "the definition itself of probability."
  - He did say that "the theory of chance consists in reducing" events to equally possible cases.

My conclusion

- If we take P1 to be a definition, then:
  - We have to suppose that Laplace was massively confused.
  - It is odd why he doesn't consistently call it a definition.
- If we take it as a rule governing inductive probability, then:
  - A few things he says are odd and hard to explain.
  - These are isolated remarks that might have an explanation; there is no massive confusion.
- The second alternative seems much more satisfactory to me.
- If that is right, *Laplace didn't endorse the classical theory of probability*.

Questions

4. State Laplace's first and second principles of probability. Include any necessary provisos.
5. What is the classical theory of probability? What are two standard objections to it?
6. Laplace said his first principle was the definition of probability. State three reasons for doubting that he really meant that.
7. If Laplace's first principle isn't a definition, what would it be? What are the advantages and disadvantages of interpreting the principle in this other way?

Appendix: Derivation of rule of succession

Applying a continuous form of Bayes's theorem, we have:

$$\begin{aligned} p(H_x|E) &= \frac{p(E|H_x)p(H_x)}{\int_0^1 p(E|H_y)p(H_y)\,dy} \\ &= \frac{p(E|H_x)}{\int_0^1 p(E|H_y)\,dy}, \text{ by (i)} \\ &= \frac{x^n}{\int_0^1 y^n\,dy}, \text{ by (ii)} \\ &= (n+1)x^n. \end{aligned}$$

Applying a continuous form of the law of total probability, we have:

$$\begin{aligned} p(F|E) &= \int_0^1 p(F|E.H_x)p(H_x|E)\,dx \\ &= \int_0^1 x(n+1)x^n\,dx, \text{ by (iii) and the preceding result} \\ &= \frac{n+1}{n+2}. \end{aligned}$$

References

- Gillies, Donald. 2000. *Philosophical Theories of Probability,* Routledge. Chapter 2 gives conventional account of Laplace and the classical theory.
- Laplace, Pierre Simon. 1886. *Théorie analytique des probabilités*, 3rd ed, in *Œuvres complètes de Laplace*, vol. 7, Gauthier-Villars. The Introduction is a reprint of Laplace's *Essay* (in French).
- Laplace, Pierre Simon. 1995. *Philosophical Essay on Probabilities*, translated by Andrew I. Dale, Springer. This is a newer translation than the one I've used, with extensive notes.
- Gillispie, Charles Coulston, et al. 1997. *Pierre-Simon Laplace 1749–1827: A Life in Exact Science*, Princeton University Press. Says more about Laplace's work than his life.
- Venn, John. 1888. *The Logic of Chance*, 3rd ed., Macmillan. Chapter 8 names and criticizes the rule of succession.