Laplace's "Memoire sur la Probabilite des Causes par le Evenemens." (1774)

Jeff Gill

University of California, Davis

jgill@ucdavis.edu

Inverse Probability

◇ **Claim**: …je me propose de déterminer la probabilité des causes par les événements matiére neuve à bien de égards et qui mérit d'autant plus d'être cultivee que c'est principalement sous ce point de vue que la science des hasards peut être utile la vie civile.

◇ **Setup**:

- ⧫ 2 events observed: E_1, E_2.
- ⧫ n causes: $A_1, A_2, \ldots, A_n$.

◇ **Laplace's Assumptions**:

- ⧫ E_i are *conditionally independent* given A_i.
- ⧫ A_i are equally probable.

Two Events:

◇ Definition of Conditional Probability:

$$P(E_2|E_1) = P(E_2, E_1)/P(E_1)$$

◇ Law of Iterated Probability:

$$\begin{aligned} P(E_2|E_1) &= \sum_{i=1}^{n} P(E_2, A_i|E_1) \\ &= \sum P(E_2|A_i, E_1)P(A_i|E_1) \\ &= \sum P(E_2|A_i)P(A_i|E_1) \\ &= \sum P(E|A_i)P(A_i|E) \end{aligned}$$

◇ Now Use Equiprobability:

$$P(E_1) = \frac{1}{n}\sum_{i=1}^{n} P(E_1|A_i) = \frac{1}{n}\sum_{i=1}^{n} P(E|A_i)$$

Two Events (cont.):

◇ Joint Probability:

$$
\begin{aligned}
P(E_1, E_2) &= \frac{1}{n} \sum_{i=1}^{n} P(E_1, E_2 | A_i) \\
&= \frac{1}{n} \sum P(E_1 | A_i) P(E_2 | A_i) \\
&= \frac{1}{n} \sum P(E | A_i) P(E | A_i) \\
&= \frac{1}{n} \sum P(E | A_i)^2
\end{aligned}
$$

◇ Match up Conditional with LIP:

$$
\begin{aligned}
\sum P(E|A_i) P(A_i|E) &= \frac{P(E_1, E_2)}{P(E_1)} \\
&= \frac{\sum P(E|A_i)^2}{\sum P(E|A_i)}
\end{aligned}
$$

$k+1$ Events:

$\diamondsuit$ Definition of Conditional Probability:

$$P(E_2|E_1, E_3, \ldots, E_{k+1}) = \frac{P(E_1, E_2, E_3, \ldots, E_{k+1})}{P(E_1, E_3, \ldots, E_{k+1})}$$

$\diamondsuit$ Law of Iterated Probability:

$$\begin{aligned} P(E_2|E_1, E_3, \ldots, E_{k+1}) &= \sum_{i=1}^{n} P(E_2, A_i|E_1, E_3, \ldots, E_{k+1}) \\ &= \sum P(E_2|A_i, E_1, E_3, \ldots, E_{k+1}) P(A_i|E_1, E_3, \ldots, E_{k+1}) \\ &= \sum P(E_2|A_i) P(A_i|E_1, E_3, \ldots, E_{k+1}) \\ &= \sum P(E_2|A_i) P(E_1|A_i) P(A_i|E_3, \ldots, E_{k+1}) \\ &= \sum P(E|A_i)^k P(A_i|E) \end{aligned}$$

$k+1$ Events (cont.):

◇ Joint Probability:

$$P(E_1, E_2, \ldots, E_{k+1}) = \frac{1}{n}\sum_{i=1}^{n} P(E_1, E_2, \ldots, E_{k+1}|A_i)$$
$$= \frac{1}{n}\sum P(E_1|A_i)P(E_2|A_i)\cdots P(E_{k+1}|A_i)$$
$$= \frac{1}{n}\sum P(E|A_i)^{k+1}$$

◇ Match up Conditional with LIP:

$$\sum P(E|A_i)^k P(A_i|E) = \frac{\sum P(E|A_i)^{k+1}}{\sum P(E|A_i)}$$

◇ This is now a set of linear equations for which there is only one solution:

$$P(A_i|E) = \frac{P(E|A_i)}{\sum_j P(E|A_j)} \tag{1}$$

Choice of Means

◇ Given three scalar observations: $a, b, c \in [A : B]$, find the best estimator of the true value V.

◇ Label the probability of an observation differing from V by the amount x as $y = \phi(x)$.

◇ Two questions:

- ◆ what shape to use for $y = \phi(x)$?
- ◆ what mean to use?

◇ Set three conditions:

- ◆ Elle doit être partagee en deux parties entièrement semblables par la droit VR, car il est tout aussi probable que l'observation s'écartera de la verité à droit comme à gauche;
- ◆ Elle doit avoir pour asymptote la ligne KP, parce que la probabité que l'observation s'éloigne de la vérité à une distance infinie est évidemment nulle;
- ◆ L'aire entié de cette courbe doit être égale à l'unité, puisqu'il est certain que l'observation tombera sur un de points de la droite KP.

Choice of Means (cont.)

♢ Let: $p = b - a$, $q = c - b$, $x = V - a$.

♢ The probability of the differences (x, p, q) to V:

$$\begin{aligned} y &= \phi(x)\phi(p-x)\phi(p+q-x) \\ &= \phi(V-a)\phi(b-a-x)\phi(b-a+c-b-x) \\ &= \phi(V-a)\phi(b-a-x)\phi(c-a-x) \end{aligned}$$

♢ Si nous supposons le véritable instat au point V':

$$\begin{aligned} y' &= \phi(x')\phi(p-x')\phi(p+q-x') \\ &= \phi(V'-a)\phi(b-a-x')\phi(c-a-x') \end{aligned}$$

♢ Ratio:

$$\frac{y}{y'} = \frac{\phi(x)\phi(p-x)\phi(p+q-x)}{\phi(x')\phi(p-x')\phi(p+q-x')}$$

♢ Modern Version:

$$\frac{y}{y'} = \frac{\phi(x_1-V)\phi(x_2-V)\phi(x_3-V)}{\phi(x_1-V')\phi(x_2-V')\phi(x_3-V')}$$

which would solve by MLE. Stigler notes here that x is a pivotal quantity and p, q ancillary statistics.

◇ Assume a uniform prior (more tacitly than explicitly), and evaluate 2 conditions:

- ⧫ La primiére est l'instant tel qu'en également probable que le véritable instant du phénomène tombe avan out après; on pourrait appeler cet instant milieu de probabilité.

  Meaning: use the median.

- ⧫ Le seconde est l'instant tel qu'en le prenant pour milieu, la somme des erreurs à craindre, multipliées par leur probabilité, soit un minimum; on pourrait l'appeler milieu d'erreur ou milieu astronomique, comm étant celui auquel les astronomes doivent s'arreter de préférence.

  Meaning: use the quantity at the "astronomical center of mass" that minizes: $f(x) = \int |x - V| f(x) dx$. That is, in modern terms, minimize the posterior expected loss.

◇ The really cool part:

- Laplace shows that these are identical.
- Found merely by bisecting the curve.
- Equivalence by: starting with $-\infty < \alpha < V < \beta < \infty$,

$$\frac{\partial}{\partial x} f(x) = \frac{\partial}{\partial x} \Bigg(\int_{-\infty}^{\alpha} |x - V| f(x) dx + \int_{\alpha}^{\beta} |x - V| f(x) dx + \int_{\beta}^{\infty} |x - V| f(x) dx \Bigg) \equiv 0$$

only by:

$$\int_{-\infty}^{V} |x - V| f(x) dx = \int_{V}^{\infty} |x - V| f(x) dx$$