Laplace's Law of Succession

Peter Haggstrom
www.gotohaggstrom.com
mathsatbondibeach@gmail.com

October 31, 2012

1 What is the probability that the sun will rise tomorrow?

Laplace sought to apply probabilistic methods to the problem of working out the likelihood that the sun will rise tomorrow on the assumption that it has risen n times in succession. Hence this problem is referred to as Laplace's "Law of Succession". Note that this has got nothing to do with the physics of how the sun operates and as you will see in a later development, the probability can be calculated using urn model techniques. It is the intricate combinatorial arguments involved in that urn model derivation that may be of some interest to serious students of probability theory.

Laplace had to make some assumptions which were essentially these:
1. The probability of the sun rising on any day is a constant whose value is unknown.
2. It is assumed that this unknown probability is a random variable x which is uniformly distributed on [0,1]. This reflects the total ignorance of the probability of the sun rising.
3. Finally, it is assumed that successive sunrises are independent events.

Thus ξ has a density function f such that:

$$f(p) = 1 \text{ for } 0 \leq p \leq 1 \tag{1}$$

Condition (1) can be expressed unrigorously but meaningfully for what follows as:

$$P\{p \leq \xi \leq p + dp\} = 1 \text{ for } 0 \leq p \leq 1 \tag{2}$$

Now if S_n is the event that the sun rises n times in succession, then using the independence assumption we then have:

$$P\{S_n | \xi = p\} = p^n \tag{3}$$

Now recall that if $\Omega = \bigcup_{i=1}^{n} A_i$ and the A_i are disjoint, then $P\{B\} = \sum_{i=1}^{n} P\{A_i\}\, P\{B|A_i\}$, so that using it as a discrete analogy we have:

$$P\{S_n\} = \sum_{0 \le p \le 1} P\{\xi = p\} P\{S_n|\xi = p\} \tag{4}$$

But we now need to make the leap from (4) to the continuous case where we hope that:

$$P\{S_n\} = \int_0^1 P\{S_n|\xi = p\}dp = \int_0^1 p^n dp = \frac{1}{n+1} \tag{5}$$

Now applying (5) to n and n+1 we see that:

$$P\{S_{n+1}|S_n\} = \frac{P\{S_{n+1}S_n\}}{P\{S_n\}} = \frac{\frac{1}{n+2}}{\frac{1}{n+1}} = \frac{n+1}{n+2} \tag{6}$$

Now the above "proof" is not entirely kosher. For instance, there is a fundamental problem with (3) in that $P\{\xi = p\} = 0$ for all p so the conditional probability is not defined. You need the concept of the Radon-Nikodym derivative (for a discussion of this see *Steven E Shreve, "Stochastic Calculus for Finance II Continuous Time Finance", Springer, 2004 p.36).* A much more detailed and rigorous treatment of the Radon-Nikodym derivative can be found in *Patrick Billingsley, Probability and Measure, Third Edition, Wiley, pp 422-425.*

2 Kai Lai Chung's rederivation of Laplace's Law of succession using an urn model

In his book *"Chance and Choice Memorabilia", World Scientific Publishing Company, 2004* Kai Lai Chung begins with an article titled "Will the Sun Rise Again?". The problem posed in the article is that if the sun has risen on n successive days, what is the probability that it will rise on the next day? Laplace gave as the answer $\frac{n+1}{n+2}$ so that for $n = 1$ the probability is 2/3. Chung notes that since Laplace assumes that it is equally likely that the sun rises or not on each day, why is it then that the answer is not simply 1/2? This is in fact the question Chung posed to himself as a young high school student which he then solved, evidencing why he would become an internationally significant figure in probability theory. In what follows I have explained in detail all the compressed steps

in Chung's articles (one of which is in Chinese). He proceeds to develop an urn model as follows:
1. In the urn there are two balls either black or white.
2. Draw the first ball and see if it white. What is the probability that the other ball in the urn is also white?
3. The urn may contain i black balls where $i = 0, 1, 2$. Laplace called these probabilities "causes" which are denoted by C_i. E is the event that the first ball drawn is white and F is the event that the second ball is also white. It follows that the conditional probabilities are:

$$P\{E|C_i\} = \frac{2-i}{2} \qquad \text{i=0,1,2} \tag{7}$$

Laplace then argued using Bayes' Rule that the "inverse probabilities" ought to be proportional to these conditional probabilities. Chung then gives this formula:

$$P\{C_i|E\} = \frac{P\{E|C_i\}}{\sum_i P\{E|C_i\}} \tag{8}$$

If you can see this result straight away then you can move on, however, as you will see Chung has compressed a couple of steps. To start with we assume that the C_i are disjoint events which partition the sample space and $P\{C_i\} > 0$ for all i. Then for any event E we have the following:

$$E = (E \cap C_0) \cup (E \cap C_1) \cup \dots (E \cap C_n) \tag{9}$$

using the fact that the C_i partition the space ie they are disjoint.

Therefore,

$$P\{E\} = \sum_{i=0}^{n} P\{E \cap C_i\} \tag{10}$$

using the fact that the C_i are disjoint.

So:

$$P\{E\} = \sum_{i=0}^{n} P\{E \cap C_i\} = \sum_{i=0}^{n} P\{C_i\}\, P\{E|C_i\} \tag{11}$$

using the fact that $P\{E|C_i\} = \frac{P\{E \cap C_i\}}{P\{C_i\}}$

Now $P\{C_i|E\} = \frac{P\{C_i \cap E\}}{P\{E\}} = \frac{P\{E \cap C_i\}}{P\{E\}}$ and $P\{E|C_i\} = \frac{P\{E \cap C_i\}}{P\{C_i\}}$.

Therefore, $P\{C_i|E\} = \frac{P\{C_i\}\, P\{E|C_i\}}{P\{E\}}$

So finally:

$$P\{C_i|E\} = \frac{P\{C_i\}\, P\{E|C_i\}}{\sum_{i=0}^{n} P\{C_i\}\, P\{E|C_i\}} \quad \text{using (8) and (11)} \tag{12}$$

You will notice that (12) differs from (8). The reason for this is that Chung has assumed that $P\{C_i\} = \frac{1}{n+1}$ for all $i = 0, 1, 2, \ldots n$. As you will see in the continuous derivation, this step will be made explicitly clear. When that adjustment is taken into account the constant factor $P\{C_i\}$ cancels out and you get (8).

So at this stage Chung has the relationship (8) and he then states that:

$$P\{F|E\} = \sum_i P\{C_i|E\}\, P\{F|C_i \cap E\} \tag{13}$$

To see how (14) is derived it is easiest to start with just two disjoint C_i ie C and its complement C^*. The logic for the general case can be justified on inductive grounds. We first note that $F \cap E \cap C$ and $F \cap E \cap C^*$ are disjoint, ie start with:

$$\begin{aligned} P\{F|E\} &= \frac{P\{F \cap E\}}{P\{E\}} = \frac{P\{F \cap E \cap \Omega\}}{P\{E\}} = \frac{P\{(F \cap E) \cap (C \cup C^*)\}}{P\{E\}} \\ &= \frac{P\{(F \cap E \cap C) \cup (F \cap E \cap C^*)\}}{P\{E\}} \quad \text{using disjointness and De Morgan's Laws} \\ &= \frac{P\{F|E \cap C\}\, P\{E \cap C\}}{P\{E\}} + \frac{P\{F|E \cap C^*\}\, P\{E \cap C^*\}}{P\{E\}} \\ &= P\{C|E\}\, P\{F|E \cap C\} + P\{C^*|E\}\, P\{F|E \cap C^*\} \end{aligned} \tag{14}$$

Generalising (14) for an arbitrary number of C_i we get Chung's result $P\{F|E\} = \sum_i P\{C_i|E\}\, P\{F|C_i \cap E\}$.

His next assertion is that for $0 \leq i \leq m$ the following hold:

$$P\{E|C_i\} = \frac{\binom{n}{r}\binom{m}{i}}{\binom{n+m}{r+i}} \tag{15}$$

$$P\{F|E \cap C_i\} = \frac{\binom{m-1}{i}}{\binom{m}{i}} \tag{16}$$

Chung defines:

$$S = \sum_i P\{E|C_i\} \tag{17}$$

To see how (15) is derived, note that C_i is the event that the total number of black balls in the urn is $r+i$ while E is the event that the first n balls drawn contain exactly r balls. There are $n+m$ balls in total in the urn hence there are $\binom{n+m}{r+i}$ ways of drawing the $r+i$ black balls from $n+m$ balls. The "favourable" cases involve getting r black balls from n followed by i black balls from m balls ie $\binom{n}{r}\binom{m}{i}$

Similarly (16) is derived by noting that $P\{F|E \cap C_i\}$ is the probability that the $(n+1)^{st}$ ball is white given that the first n balls contain exactly r black balls and the total number of black balls in the urn is $r+i$. Now:

$$P\{F|E \cap C_i\} = \frac{P\{F \cap E \cap C_i\}}{P\{E \cap C_i\}} = \frac{\frac{\binom{n}{r}\binom{m-i}{i}}{\binom{n+m}{r+i}}}{\frac{\binom{n}{r}\binom{m}{i}}{\binom{m+n}{r+i}}} \tag{18}$$

The top of (18) is the probability that the $(n+1)^{st}$ ball is white and the first n balls contain exactly r black balls where there is a total of r+i black balls in the urn. Thus the r black balls out of n can be chosen in $\binom{n}{r}$ ways and then choosing the remaining i black balls from $m-1$ so that there is only one way of choosing the last ball to be white (there being $n+m$ balls in total). The factor $\binom{n+m}{r+i}$ reflects the fact that $r+i$ balls are chosen from $m+n$. Similarly, the bottom of (18) is the probability that in the first n balls, exactly r are black and the total number of black balls in the urn is $r+i$. Chung then states that after "obvious manipulations" you get:

$$S = \binom{n+m}{n}^{-1} \sum_{i=0}^{m} \binom{r+i}{r} \binom{n+m-r-i}{n-r} \tag{19}$$

The details are as follows:

$$\begin{aligned} S = \sum_{i=0}^{m} P\{E|C_i\} = \sum_{i=0}^{m} \frac{\binom{n}{r}\binom{m}{i}}{\binom{n+m}{r+i}} &= \binom{n}{r} \sum_{i=0}^{m} \frac{m!}{(m-i)!\, i!} \frac{(n+m-r-i)!\,(r+i)!}{(n+m)!} \\ = \binom{n}{r} \sum_{i=0}^{m} \frac{(n-r)!\, m!\, (r+i)!}{i!\,(n+m)!} \binom{n+m-r-i}{n-r} &= \frac{n!}{r!} \sum_{i=0}^{m} \frac{r!\, m!}{(n+m)!} \binom{r+i}{r} \binom{n+m-r-i}{n-r} \\ &= \binom{n+m}{n}^{-1} \sum_{i=0}^{m} \binom{r+i}{r} \binom{n+m-r-i}{n-r} \end{aligned} \tag{20}$$

Alternatively one can use the substitution $n = r+i$ in (19) so that $\binom{n+m}{r+i}$ becomes $\binom{n+m}{n}$, $\binom{n}{r}$ becomes $\binom{r+i}{r}$ and $\binom{m}{i}$ becomes $\binom{n+m-r-i}{n-r}$.

3 The Main Identity

Following Chung's development of how to simplify (20), we seek to prove the main identity which is this:

$$\sum_{i=0}^{z} \binom{x+i}{x} \binom{y+z-i}{y} = \binom{x+y+z+1}{x+y+1} \tag{21}$$

If we can prove (21) then (20) will become:

$$S = \binom{n+m}{n}^{-1} \binom{n+m+1}{n+1} = \frac{n+m+1}{n+1} \tag{22}$$

To see this, note from (20) that:

$$\binom{n+m}{n}^{-1} \sum_{i=0}^{m} \binom{r+i}{r} \binom{n+m-r-1}{n-r} = \binom{n+m}{n}^{-1} \sum_{i=0}^{m} \binom{r+i}{r} \binom{y+m-i}{y}$$

$$\text{where } y = n-r,\ z = m \text{ and } x = r$$

$$= \binom{n+m}{n}^{-1} \binom{x+y+m+1}{x+y+1} \quad \text{using (21)}$$

$$= \binom{n+m}{n}^{-1} \binom{n+m+1}{n+1} = \frac{n!\, m!\, (n+m+1)!}{(n+m)!\, m!\, (n+1)!} = \frac{n+m+1}{n+1} \tag{23}$$

Chung then makes the substitution:

$$f(n,m) = \binom{n}{r}^{-1} S = \frac{n+m+1}{n+1} \tag{24}$$

He then says that $P\{F|E\} = \frac{f(n+1,m-1)}{f(n,m)}$ (I will come to this shortly) which after "incredible" cancellations reduces to the desired result: $\frac{n-r+1}{n+2}$. When $r = 0$ you get $\frac{n+1}{n+2}$. He says: "Isn't it a miracle that the arbitrarily introduced number m disappears in the final result? Can we prove this a priori?"

To prove (21) Chung begins with the standard combinatorial identity (no magic here):

$$\binom{x+i}{x} = \binom{x-1+i}{x-1} + \binom{x-1+i}{x} \tag{25}$$

He then gets by induction:

$$\binom{x+i}{x} = \sum_{j=0}^{i} \binom{x-1+j}{x-1} \tag{26}$$

(26) is clearly true for $i = 0$. We assume (26) holds for any i. Then:

$$\binom{x+i+1}{x} = \binom{x+i}{x} + \binom{x+i}{x-1} = \sum_{j=0}^{i} \binom{x-1+j}{x-1} + \binom{x+i}{x-1} = \sum_{j=0}^{i+1} \binom{x-1+j}{x-1} \tag{27}$$

as required.

Using these results (21) becomes:

$$\sum_{i=0}^{z} \binom{x+i}{x} \binom{y+z-i}{y} = \sum_{j=0}^{z} \sum_{i=j}^{z} \binom{x-1+j}{x-1} \binom{y+z-i}{y} = \sum_{j=0}^{z} \binom{x-1+j}{x-1} \binom{y+1+z-j}{y+1} \tag{28}$$

where (x,y) in (21) has been replaced with (x-1,y+1) and by continuing the process the original sum is reduced to (0,y+x). The sum is valid for all relevant values of x and y and in particular those of the form $x-1$ and $y+1$ and so you get the following sum resulting:

$$\sum_{i=0}^{z} \binom{0+i}{i} \binom{y+x+z-i}{y+x} = \sum_{i=0}^{z} \binom{y+x+z-i}{y+x} = \binom{y+x+z+1}{x+y+1} \tag{29}$$

To see why $\sum_{i=0}^{z} \binom{y+x+z-i}{y+x} = \binom{y+x+z+1}{x+y+1}$ use (26) as follows with some substitutions:

$$\sum_{i=0}^{z} \binom{y+x+z-i}{y+x} = \sum_{i=0}^{z} \binom{u+z-i}{u} = \sum_{j=0}^{z} \binom{u+j}{u} = \binom{x+y+z+1}{x+y+1} \tag{30}$$

The aim in this whole project is to prove that $P\{F|E\} = \frac{n+1}{n+2}$ and to do so we can crank through (13) as follows and it will become clear where f(n,m) comes from.

$P\{F|E\} = \sum_{i=0}^{m} P\{C_i|E\}\, F\{F|C_i \cap E\}$ from(13)
$= \frac{1}{S} \sum_{i=0}^{m} P\{E|C_i\}\, F\{F|C_i \cap E\}$ from(13) using (8)

$$= \frac{\sum_{i=0}^{m} \frac{\binom{n}{r}\binom{m}{i}}{\binom{n+m}{r+i}} \frac{\binom{m-1}{i}}{\binom{m}{i}}}{\sum_{i=0}^{m} \frac{\binom{n}{r}\binom{m}{i}}{\binom{n+m}{r+i}}} \text{ using (15),(16) and (22)}$$

$$= \frac{\sum_{i=0}^{m} \frac{\binom{m-1}{i}}{\binom{n+m}{r+i}}}{\sum_{i=0}^{m} \frac{\binom{m}{i}}{\binom{n+m}{r+i}}} \tag{31}$$

Now let:

$$f(n,m) = \sum_{i=0}^{m} \frac{\binom{m}{i}}{\binom{n+m}{r+i}} = \sum_{i=0}^{m} \frac{m!\,(m+n-r-i)!\,(r+i)!}{(m-i)!\,i!\,(n+m)!}$$

$$= \frac{m!\,r!\,(n-r)!}{(n+m)!} \sum_{i=0}^{m} \binom{r+i}{r}\binom{n+m-r-i}{n-r} = \binom{n}{r}^{-1}\binom{n+m}{n}^{-1} \sum_{i=0}^{m} \binom{r+i}{r}\binom{n+m-r-i}{n-r}$$

$$= \binom{n}{r}^{-1} \frac{n+m+1}{n+1} \text{ using (23) .....} \quad (32)$$

Recall that $P\{F|E\} = \frac{f(n+1,m-1)}{f(n,m)}$. Note that the numerator in (31) is $\sum_{i=0}^{m} \frac{\binom{m-1}{i}}{\binom{n+m}{r+i}} = \sum_{i=0}^{m-1} \frac{\binom{m-1}{i}}{\binom{n+m}{r+i}}$ since $\binom{m-1}{m} = 0$ so $f(n+1,m-1) = \sum_{i=0}^{m-1} \frac{\binom{m-1}{i}}{\binom{n+1+m-1}{r+i}} = \sum_{i=0}^{m-1} \frac{\binom{m-1}{i}}{\binom{n+m}{r+i}}$ by plugging $n+1$ and $m-1$ into the expression for f(n,m).

When we substitute $n+1$ and $m-1$ into (32) we get:

$$f(n+1, m-1) = \binom{n+1}{r}^{-1} \frac{n+1+m-1+1}{n+2} = \binom{n+1}{r}^{-1} \frac{n+m+1}{n+2} \tag{33}$$

Hence:

$$P\{F|E\} = \frac{\binom{n+1}{r}^{-1} \frac{n+m+1}{n+2}}{\binom{n}{r}^{-1} \frac{n+m+1}{n+1}} \text{ using (32) and (33)}$$

$$= \frac{n+1}{n+2} \frac{n!\,(n+1-r)!\,r!}{(n-r)!\,r!(n+1)!} = \frac{n+1-r}{n+2} \tag{34}$$

so when $r = 0$ you get the required probability $\frac{n+1}{n+2}$.

Tucked away in the Chinese part of his article Chung derives the required probabiity $P\{F|E\} = \frac{n+1}{n+2}$ as follows:

$f(n, m) = \frac{n+m+1}{n+1}$ using (24) and $f(n+1, m-1) = \frac{n+m+1}{n+2}$

$$\text{Hence } P\{F|E\} = \frac{n+1}{n+2} \tag{35}$$

4 Laszlo Lovasz's counting argument

The main identity in Chung's proof can be proved by generating function techniques, for instance, by using something like:

$$(1-t)^{-x-1} = \sum_{i=0}^{\infty} \binom{x+i}{x} t^i \tag{36}$$

By checking the coefficient of t^z in $(1-t)^{-x-1}\,(1-t)^{-y-1} = (1-t)^{-(x+y+1)-1}$ the result

follows.
In his book "Combinatorial Problems and Exercises", Second Edition, AMS Chelsea Publishing,2000 Laszlo Lovasz poses problem 42(i) at page 21 which is essentially the main identity (21).
He solves this with a 3 line counting proof which can be reworded as follows.

In $\sum_{i=0}^{z} \binom{x+i}{x} \binom{y+z-i}{y} = \binom{x+y+z+1}{x+y+1}$ think of choosing in increasing order $x+y+1$ integers from the integers ranging from 1 to $x+y+z+1$. Let the $(x+1)^{st}$ choice be $x+i+1$ where $0 \le i \le z$. For each i, x integers are chosen from 1 to $x+i$ while the other y integers are chosen from $x+i+2$ to $x+y+z+1$. The total of these choices is the i^{th} term in $\sum_{i=0}^{z} \binom{x+i}{x} \binom{y+z-i}{y} = \binom{x+y+z+1}{x+y+1}$.

5 A less combinatorial proof

A much more straightforward proof of Laplace's Law of Succession can be found at pages 63-64 of Dimitri P Bertsekas and John N Tsitsikils, "Introduction to Probability", Athena Scientific, 2001.

There are m + 1 boxes with the k^{th} box containing r black balls and m-k white balls where $0 \le k \le m$. A box is chosen at random, all choices being equally likely - note the comment I made in the context of the derivation (8) of above - see the comment after (12)) and then a ball is chosen at random from that box.

There are m successive draws with replacement (a new ball is selected independently each time). Suppose that a black ball was drawn each of the n times. What is the probability that if one more ball is drawn it will be black?

Let E be the event of a black ball being drawn at time $n+1$ while B_n is the event a black ball is drawn in each of n previous draws. We have the following basic relationships:

$$P\{E|B_n\} = \frac{P\{E \cap B_n\}}{P\{B_n\}} \tag{37}$$

$$P\{B_n\} = \sum_{k=0}^{m} P\{k^{th} \text{ box is chosen}\} \, (\frac{k}{m})^n$$

this being an application of the total probability result in (13)

$$= \frac{1}{m+1} \sum_{k=0}^{m} (\frac{k}{m})^n \tag{38}$$

$$P\{E \cap B_n\} = P\{B_{n+1}\} = \frac{1}{m+1} \sum_{k=0}^{m} (\frac{k}{m})^{n+1} \tag{39}$$

To get an approximation for (38) we simply move to an integral as follows:

$$\frac{1}{m+1} \sum_{k=0}^{m} (\frac{k}{m})^{n} \approx \frac{1}{m+1} \int_0^m (\frac{x}{m})^n dx = \frac{m}{m+1} \frac{1}{n+1} \approx \frac{1}{n+1} \text{for large m} \tag{40}$$

Following the same process for (39) we get as its approximation $\frac{1}{n+2}$ and from (37) the required probability is approximated by $\frac{n+1}{n+2}$ as before. Note that the derivation given by Chung does not involve any approximations.

6 History

An earlier version of this document contained some typos which have been corrected while I was retypesetting this in Latex.