Paradoxes in Probability

The paradox of Mrs Smith's children

Mrs Smith has two children and (at least) one of them is a girl. What is the probability that the other child is a boy or a girl? The surprising answer is that the other child is *twice as likely to be a boy* as it is to be a girl.

How can that be? I hear you ask. Surely the other child is equally likely to be a boy or a girl. How can it be twice as likely to be a boy? Just knowing that one of the children is female tells us absolutely nothing about the other one.

The argument is persuasive, I know, but it is not correct all the same. The correct argument goes like this: there are four equally likely possibilities concerning Mrs Smith's children listed in the table below:

Elder child	Younger child
Boy	Boy
Boy	Girl
Girl	Boy
Girl	Girl

[1]

Since we are told that (at least) one of them is a girl we can rule out option 1 (Boy, Boy) leaving *three* equally likely options. Of these three options, 2 consist of one boy and one girl; only 1 option has both girls. The probability of the other child being a boy is therefore twice as likely as it being a girl.

Many people find this argument difficult to accept but I assure you, it is correct. The probability of the other child being a boy is 2/3 while the probability that it is a girl is 1/3.

Just to make things even more puzzling, suppose you are told that Mrs Smith has two children and the *elder* one is a girl. What then? The table now looks like this:

Elder child	Younger child
Boy	Boy
Boy	Girl
Girl	Boy
Girl	Girl

[2]

We have had to rule out *two* of the four options and now we can see that the probability of the other child being a boy is equal to the probability that it is a girl: 1/2 each.

But this sounds crazy. How is it that being told that the *elder* child is a girl makes a difference to the probability of the other one is a boy? It doesn't seem to make sense.

Try these two statements:

A) Mrs Smith has two children and (at least) one of them is a red-headed girl.
B) Mrs Smith has two children and the red-headed one is a girl.

Statement A is basically the same as the original statement. The fact that the girl mentioned has red hair is irrelevant to the argument and the probability that the other one is a boy is 2/3 not 1/2.

Statement B allows us to distinguish the two children because we infer that *only one of them has red hair*. We can therefore draw up the table as follows:

Red child	Fair child
Boy	Boy
Boy	Girl
Girl	Boy
Girl	Girl

[3]

and we can see that the probability that the other one is a boy is 1/2.

The crucial difference is that in the second case, the property of having red hair allows us to *distinguish the two children unambiguously* whereas this is not the case with the first statement. If you saw in the distance Mrs Smith with a red headed child, statement B would allow you to infer with certainty that the child was a girl. If all you knew was that one of her children was red headed, however, you would not be able to deduce this because it is perfectly possible that both children were red headed.

This apparent paradox is an example of what is called Bayesian or conditional probability and concerns the task of determining the probability of a certain outcome **X** given a certain condition **C**. In the original scenario we are asked to determine the probability that the other child is a boy (or girl) *given that* one of the children is a girl.

Bayes' Theorem

Bayes' theorem states that the probability of **X** *given* **C** is equal to the *a priori* probability of **X and C** divided by the *a priori* probability of **C** on its own. In symbols

$$p(\mathbf{X}\,|\,\mathbf{C}) = \frac{p(\mathbf{X}\,\&\,\mathbf{C})}{p(\mathbf{C})} \quad (1)$$

Lets calculate the probability that Mrs Smith's other child is a boy. First we must define **X** and **C** very carefully.

X : the other child is a boy
C : one of the children is a girl.

Now *p*(**X & C**) is the *a priori* probability that one child is a girl **and** the other child is a boy. By *a priori* we mean 'from first principles' i.e. without any other conditions. Since there are four equally likely possibilities and of these, two have one boy and one girl, this probability is1/2.

p(**C**) is the *a priori* probability that (at least) one of the children is a girl. We know that the probability of this outcome is 3/4.

Applying Bayes' theorem we find that *p*(**X | C**), the probability that the other child is a boy *given that* one child is a girl is1/2 over 3/4 = 2/3.

To make absolutely sure that the theorem works, lets calculate the probability that the other child is a boy given that the red-headed child is a girl.

X : the other child is a boy
C : the red-headed child is a girl.

p(**X & C**) = 1/4 (of the four possibilities, there is only one which has a red headed girl and a fair boy – see table [3])
p(**C**) = 1/2
p(**X | C**) = 1/4 over 1/2 = 1/2.

You can see how vital it is to specify exactly what **X** and **C** are and how careful you must be to calculate the two a priori probabilities.

Wednesday's child

We are not finished with this puzzle yet. Suppose you are now told that Mrs Smith has two children and that (at least) one of them is a girl born on Wednesday. What is the probability now that the other child is a boy?

Now if you have followed the arguments so far and are prepared to accept them you will probably want to say the fact that the child was born on Wednesday is no more relevant than the fact that she is red headed and makes no difference to the probabilities. All we really know is that Mrs Smith has two children and that at least one of them is a girl and under these circumstances the probability that the other child is a boy is 2/3.

Amazingly this is not correct. Even more surprising, the correct answer is not even 1/2 – it is somewhere in between. To see why we need to set out *all* the equally likely possibilities. It turns out that there are rather a lot of them – 196 in fact. (See table [4]).

I have highlighted those potential families which contain (at least.) one girl born on a Wednesday. There are 27 of them, of which 14 consist of a girl and a boy. These are indicated in yellow. The correct answer to the puzzle is therefore 14/27.

Let's see if Bayes theorem gives the same answer and if it throws any more light on the puzzle.

X : the other child is a boy
C : one of the children is a girl born on a Wednesday

Now the probability that a family of two contains a girl born on a Wednesday is *not* equal to the probability that the family contains a girl (3/4) multiplied by the probability that a given child is born on a Wednesday (1/7). This would be 3/28 or 21/196. Looking again at the table we have seen that the odds are slightly greater than this – 27/196. The reason for this is that, contrary to what you might expect, the two factors are not actually independent. When you expand the table [1] by 7, the number of possibilities in the BG and GB blocks increase by 7 but the number of possibilities in the GG block increase to 13. A litte more thought will convince you that if there had been 8 days in a week or n days, the total number of possibilities would increase to $n + n + n + (n - 1) = 4n - 1$, of which $2n$ would be families in which the other child is a boy. This gives us:

$p(\mathbf{X} \mathbin{\&} \mathbf{C}) = 2n / 4n^2$
$p(\mathbf{C}) = (4n - 1) / 4n^2$
$p(\mathbf{X} \mid \mathbf{C}) = 2n / (4n - 1)$

This formula gives us a nice way of interpreting the situation. n is the reciprocal of the *a priori* probability of the extra condition. In the case of Wednesday's child, $n = 7$ and $p(\mathbf{X} \mid \mathbf{C}) = 14/27$.

If there is no extra condition (or, if you like, the extra condition is a certainty) then $n = 1$ and $p(\mathbf{X} \mid \mathbf{C}) = 2/3$.

The greater the unlikelihood of the extra condition, the nearer the probability approaches 1/2. For example, if we knowe that Mrs Smith's girl was born on Christmas day, the probability of the other child being a boy would be $2\times 365 / (4 \times 365 - 1) = 730 / 1459 = 0.50034$. Indeed, we were wrong about the red headed girl. Since only about 1 in 20 people are red headed, the correct answer to this problem should have been $2\times 20 / (4 \times 20 - 1) = 40 / 79 = 0.506$.

To summarise what we have learnt we can say that *the more the extra condition serves to distinguish the two children*, the closer the probability that the other child is a boy approaches 1/2.

BM BM	BT BM	BW BM	BT BM	BF BM	BS BM	BS BM
BM BT	BT BT	BW BT	BT BT	BF BT	BS BT	BS BT
BM BW	BT BW	BW BW	BT BW	BF BW	BS BW	BS BW
BM BT	BT BT	BW BT	BT BT	BF BT	BS BT	BS BT
BM BF	BT BF	BW BF	BT BF	BF BF	BS BF	BS BF
BM BS	BT BS	BW BS	BT BS	BF BS	BS BS	BS BS
BM BS	BT BS	BW BS	BT BS	BF BS	BS BS	BS BS
BM GM	BT GM	BW GM	BT GM	BF GM	BS GM	BS GM
BM GT	BT GT	BW GT	BT GT	BF GT	BS GT	BS GT
BM GW	BT GW	BW GW	BT GW	BF GW	BS GW	BS GW
BM GT	BT GT	BW GT	BT GT	BF GT	BS GT	BS GT
BM GF	BT GF	BW GF	BT GF	BF GF	BS GF	BS GF
BM GS	BT GS	BW GS	BT GS	BF GS	BS GS	BS GS
BM GS	BT GS	BW GS	BT GS	BF GS	BS GS	BS GS
GM BM	GT BM	GW BM	GT BM	GF BM	GS BM	GS BM
GM BT	GT BT	GW BT	GT BT	GF BT	GS BT	GS BT
GM BW	GT BW	GW BW	GT BW	GF BW	GS BW	GS BW
GM BT	GT BT	GW BT	GT BT	GF BT	GS BT	GS BT
GM BF	GT BF	GW BF	GT BF	GF BF	GS BF	GS BF
GM BS	GT BS	GW BS	GT BS	GF BS	GS BS	GS BS
GM BS	GT BS	GW BS	GT BS	GF BS	GS BS	GS BS
GM GM	GT GM	GW GM	GT GM	GF GM	GS GM	GS GM
GM GT	GT GT	GW GT	GT GT	GF GT	GS GT	GS GT
GM GW	GT GW	GW GW	GT GW	GF GW	GS GW	GS GW
GM GT	GT GT	GW GT	GT GT	GF GT	GS GT	GS GT
GM GF	GT GF	GW GF	GT GF	GF GF	GS GF	GS GF
GM GS	GT GS	GW GS	GT GS	GF GS	GS GS	GS GS
GM GS	GT GS	GW GS	GT GS	GF GS	GS GS	GS GS

[4]

The paradox at last

You may be surprised to learn that we still haven't got to the real paradox yet! Here is the absolute cruncher.

Mrs Smith tells you that she has two children and shows you a photograph of her on holiday beside the seaside with one of her children. The child is a girl. What is the probability that the other child is a boy?

Is this the same as saying that (at least) one of the children is a girl – in which case that answer is that the probability of the other child being a boy is 2/3.

Or is saying that 'the child in the photograph is a girl' is like saying that 'the elder child is a girl' or 'the red-headed child is a girl' – in which case the probability is 1/2. Does the fact that there is only one child in the photograph distinguish the two children uniquely in the same way that being the elder or being red-headed differentiates them?

Let us see if Bayes' theorem can help resolve the issue. First the definitions:

X : the child not shown in the photograph is a boy
C : the child in the photograph is a girl

This is a surprisingly difficult problem, Take statement **C**. We want to know what is the *a priori* probability that the child in the photograph is a girl. Obviously it is the case that *before you see the photograph* there is an equal chance that it is either a boy or a girl and that there is an equal chance that the child not in the photo is a boy or a girl. Mathematically $p(\mathbf{X} \,\&\, \mathbf{C}) = 1/4$; $p(\mathbf{C}) = 1/2$ so the probability that the child not in the photo is a boy is 1/4 over 1/2 = 1/2.

But *as soon as you look at the photo and see that the child in the photo is a girl*, the situation changes. You now know something that you didn't know before; namely that at least one of Mrs Smith's children is a girl and we have seen that with this knowledge, the probability that the other child is a boy is 2/3, not 1/2.

But this sounds very odd. The factor of 3 which appears in the figure of 2/3 comes about mainly because $p(\mathbf{C}) = 3/4$. (You will recall that in 3 out of the 4 equally likely situations one of Mrs Smith's children is a girl). So what we are, in effect assuming is that, as soon as you see the photo, the *a priori* probability that the child in the photo is a girl becomes 3/4. But this is ridiculous. As soon as you see the photo you *know* that the child in the photo is a girl! In other words $p(\mathbf{C})$ is not 3/4 it is 1 and the probability that the child not in the photo is a boy reverts to being 1/2 again.

My interpretation of the paradox is therefore as follows: as soon as you see the photo you not only know that at least one of Mrs Smith's children is a girl, you also know that the child she took to the seaside and was photographed is a girl. This extra information serves to distinguish the two children uniquely (in effect $n = \infty$). The calculation is different but the result is the same. The probability that the child not in the photograph is a boy is still 1/2. I suspect, however, that the debate will continue!

Usually, obtaining extra knowledge about a situation alters the probabilities but this is not always true as my next example shows.

The Monty Hall paradox

In the American TV game show '*Let's make a deal*' hosted by Monty Hall the winning contestant was given the opportunity to win a car by choosing one of three closed doors. Behind one door lurked the car and a goat behind each of the other two. After choosing one door the host would open one of the other doors to reveal a goat. (Of course the host knows which door the car is behind so there is always at least one door he can open.) The contestant is given the choice of either sticking with his original guess or switching to the other remaining closed door. What should he do?

Most people would argue that, once the host has opened the door revealing the goat, there is an equal chance of the car being behind either of the two remaining doors so there is no advantage in switching your choice but this is false. It turns out that you can double your chances of winning by switching to the other door.

One way of arguing correctly is to say that, initially the probability of the car being behind the door you chose first is 1/3 and that the probability of it being behind one of the other doors is therefore 2/3. When the host opens one of these doors revealing the goat, he is, in fact giving you extra information about what is behind these doors. The chances of the car being behind the chosen door remain at 1/3 so the chances of the car being behind the door that the host did *not* open must be 2/3. It is therefore better to switch doors. It has to be admitted that the majority of people including many university professors do not think this argument has greater force than the first so lets see how Bayes' theorem gets the right answer.

Let us call the door which the contestant chooses first door A and the other two doors B and C. We shall calculate the probability that the car is behind one of the other doors given that the host has revealed a goat behind one of them. We therefore have

X : the car is behind one of the other doors B or C
C : either the car is not behind door B or it is not behind door C

Note carefully what the condition is. The host has opened one of the doors to reveal a goat. But there is always one door which the host can open so condition C is always satisfied. This means that

$p(\mathbf{C}) = 1$

But if $p(\mathbf{C}) = 1$ then $p(\mathbf{X} \ \&\ \mathbf{C}) = p(\mathbf{X})$. (the probability of **X** and a certainty is simply the probability of **X**) so

$p(\mathbf{X} \ \&\ \mathbf{C}) = 2/3$

$p(\mathbf{X} \mid \mathbf{C}) = 2/3$ over $1 = 2/3$.

So when the host opens one of the doors, *the probabilities do not change* because he only tells us what we knew already (that one of the doors B and C conceals a goat). It remains true that the probability that the car is behind the chosen door is 1/3 and the probability that it is behind one of the other doors is still 2/3. He does, however, give the contestant further information on the basis of which he can make a better choice.

The Three Card Trick

You have three cards: card A is white on both sides; card B is black on both sides; card C is black on one side and white on the other. While your back is turned your friend chooses one of the cards at random and places it on the table. When you turn round you see that the upper surface of the card is black. What is the chance that the other side of the card is black too?

One's immediate reaction is that the whole system is symmetrical with respect to black and white. There are exactly the same number of black sides as white sides and the fact that you can see a black side can't make any difference to the colour of the other side so it must be equally likely that the other side of the card is black or white. But is this correct? It sounds suspiciously like the paradox of Mrs Smith's children.

Let's use Bayes Theorem to calculate the answer.

X: The other side of the card on the table is black
C: The upper side of the card on the table is black

p(**X & C**) is the probability that both sides of the card on the table are black. Since only one card out of the three satisfies this condition, this probability = 1/3.

p(**C**) is the *a priori* condition that the upper face of the card on the table is black. Now there are 6 equally likely possibilities concerning the way the card is dealt of which 3 show a black face. This probability is therefore 1/2.

p(**X | C**) = 1/3 over 1/2 = 2/3.

The situation is indeed very like that paradox of Mrs Smith's children.

The Faulty Cancer Screening Test

Women are regularly screened for breast cancer but in any one test the probability of a false positive (i.e. a positive indication of cancer when there is, in fact none) is around 5%. The probability of missing a genuine cancer (a false negative) is around 10%. In addition, approximately 0.5% of women tested actually have cancer. Given these figures, two important questions present themselves – specifically

a) Given that the test result was positive, what is the chance that the subject has cancer?
b) Given that the test result was negative, what is the chance that the subject has cancer?

It would seem reasonable to conclude that, because 5% of positive results are faulty, if a woman is tested positive there is a 95% chance of her having cancer. Likewise, if she tests negative there is still a 10% chance of her having the disease because 10% of the negative results are faulty. But are these conclusions correct?

Suppose we test 100 women. On average 0.5 of these women will actually have cancer and 99.5 will not

Of the 0.5 women who have cancer, 90% (on average 0.45) will test positive and 10% (0.05) will test negative.

Of the 99.5 women who do not have cancer, 5% (i.e.4.975) will test positive and 95% (94.525) will test negative

We can summarise these results in a table.

	Test positive	Test negative	**Totals**
Have cancer	0.45	0.05	**0.5**
Do not have cancer	4.975	94.525	**99.5**
Totals	**5.425**	**94.575**	**100**

Now, taking the first question:

X: The subject has cancer
C: The test result was positive

p(**X & C**) is the *a priori* probability that the subject has cancer **and** that the test was positive. A glance at the above figures shows that out of the 100 women tested, 0.45 both have cancer and test positive so *p*(**X & C**) = 0.0045

p(**C**) is the *a priori* probability that the test result is positive. We can see from the table that the total number of positive test results we can expect is 5.425. $p(\mathbf{C})$ is therefore equal to 0.05425

$p(\mathbf{X} \mid \mathbf{C}) = 0.0045 / 0.05425 = 0.083$ or 8.3%

This is an awful lot smaller than our original guess of 95%! Why is this? Looking at the figures it is clear that most of the positive test results are due due to the number of false positives which in turn is due to the fact that the great majority of women tested do not actually have cancer. A positive test result should not therefore be taken as any sort of indication that the subject actually has cancer – only a recommendation that further tests are required.

What about the second question? If anything the situation is even worse here.

X: The subject has cancer
C: The test result was negative

$p(\mathbf{X} \,\&\, \mathbf{C})$ is the *a priori* probability that the subject has cancer and that the test was negative. Out of 0.5 women who actually have cancer only 0.05 will test negative so $p(\mathbf{X} \,\&\, \mathbf{C}) = 0.0005$

$p(\mathbf{C})$ is the *a priori* probability that the test result is negative. There are 94.575 negative results per 100 women so $p(\mathbf{C}) = 0.94575$

This means that $p(\mathbf{X} \mid \mathbf{C}) = 0.0005 / 0.94575 = 0.00053$ or approximately 0.05%

The argument that since 10% of the negative results are faulty, a women who has a negative result still has a 10% chance of the disease is catastrophically wrong (and yet a woman with a negative result who has innocently asked the question 'what is the chance that the test result is wrong?' may still go home with completely the wrong impression).

Before she had the test the chance of her having cancer was 0.5%. After the test, her chances of having cancer are 10 times smaller. She should go home well pleased.

The prosecutor's fallacy

(This fallacy is discussed in greater mathematical detail in another of my articles: 'Bayes Theorem and the Prosecutor's Fallacy')

DNA evidence is frequently used in rape cases to prove that the defendant is guilty on the basis that his DNA matches that of DNA recovered from the semen found on the victim.

Now consider the following two statements:[1]

> A1: “The probability of a match if the semen came from another person is one in a billion.”
>
> therefore
>
> B1: “The probability that the semen came from another person is one in a billion.”

Statement **A1** is simply a statement about the reliability of DNA matching and reflects the fact that it is extremely unlikely (but not impossible) that the DNA from two unrelated people will match.

Statement **B1** is something entirely different. It asserts that it is extremely unlikely that the semen collected from the victim came from someone other than the defendant and that the defendant is very unlikely to be innocent.

And yet surely statement **B1** follows from statement **A1**, doesn't it? After all, the two statements only differ in a few unimportant words, don't they?

Again, the error here is monumental and could have tragic consequences.

Let's consider two statement describing a rather less emotive situation.

---

1 Taken from 'A Guide to DNA' published by the Forensic Science Service

A2: “The probability of being colour blind if you are male is about 8%”

therefore

B2: “The probability that a certain colour blind person is male is about 8%”

Few people will be taken in by this. Just because approximately 8% of males are colour blind, it makes no sense to infer that the probability of a colour blind person being male is 8% because the argument takes no account of the incidence of colour blindness in the female population.

But consider the following argument:

A3: “The probability of being colour blind if you are female is about 0.5%”

therefore

B3: “The probability that a certain colour blind person is female is about 0.5%”

Do you think you could be persuaded by this? Why does it sound more likely than the argument about colour blind males? The best that can be said of it is that the result sounds at least plausible – any given colour blind person is, indeed, more likely to be male than female – but lets work out the exact value using Bayes Theorem. We have the following statements

X: The person is female
C: The person ins colour blind

Now we know from statement **A3** that $p(\mathbf{C} \mid \mathbf{X})$ i.e. the probability that you are colour blind if you are female, is 0.005 (or 0.5%).

What we really want to know is $p(\mathbf{X} \mid \mathbf{C})$ – the probability that the person is female given that they are colour blind. In other words, we want to switch round **X** and **C**. A little algebra is needed here:

$$\begin{aligned} p(\mathbf{X} \mid \mathbf{C}) &= p\frac{(\mathbf{X} \,\&\, \mathbf{C})}{p(\mathbf{C})} \\ p(\mathbf{C} \mid \mathbf{X}) &= p\frac{(\mathbf{X} \,\&\, \mathbf{C})}{p(\mathbf{X})} \\ \frac{p(\mathbf{X} \mid \mathbf{C})}{p(\mathbf{C} \mid \mathbf{X})} &= \frac{p(\mathbf{X})}{p(\mathbf{C})} \\ p(\mathbf{X} \mid \mathbf{C}) &= p(\mathbf{C} \mid \mathbf{X}) \times \frac{p(\mathbf{X})}{p(\mathbf{C})} \end{aligned} \quad (2)$$

In order to do the calculation, therefore, we need to know $p(\mathbf{X})$ and $p(\mathbf{C})$.

$p(\mathbf{X})$ is easy. This is the *a priori* probability that the person is female. Obviously this is 0.5.

$p(\mathbf{C})$, the *a priori* probability that the person is colour blind is a bit more difficult but we can do it in two stages. First, we know from statement **A3** that $p(\mathbf{C} \mid \mathbf{X}) = 0.005$. We also know from statement **A2** that $p(\mathbf{C} \mid \mathbf{notX}) = 0.08$. ($p(\mathbf{C} \mid \mathbf{notX})$ is the probability the you are colour blind if you are not female – i.e. male)

Now there is an important identity which is not difficult to prove which enables us to calculate $p(\mathbf{C})$ from these figures, namely:

$$p(\mathbf{C}) = p(\mathbf{C} \mid \mathbf{X}) \times p(\mathbf{X}) + p(\mathbf{C} \mid \mathbf{notX}) \times p(\mathbf{notX})$$

(All we are saying here is that the probability of something happening (**C**) is equal to the probability of (**C**) given (**X**) plus the probability of (**C**) given (**notX**))

Putting in the numbers that we know, $p(\mathbf{C}) = 0.005 \times 0.5 + 0.08 \times 0.5 = 0.0425$. We are now in a position to calculate the answer:

$$p(\mathbf{X} \mid \mathbf{C}) = 0.005 \times \frac{0.5}{0.0425} = 0.0588 \quad \text{or} \quad 5.88\%$$

How does this calculation apply to the rape case? Statement **A1** tells us that *given that the subject is innocent* there is a one in a million chance of a match between the samples. This allows us to identify **X** and **C**

X: The subject is innocent
C: The DNA samples match

and that $p(\mathbf{C} \mid \mathbf{X}) = 1/1{,}000{,}000{,}000$. This is a rather small number so lets call this figure p_{match} for short.

We want to know $p(\mathbf{X} \mid \mathbf{C})$ using equation (2).

$p(\mathbf{X})$ is the *a priori* probability that the subject is innocent. Now if other evidence resatricts the number of possible suspects to a certain group with N mambers, then the probability that he is innocent is $(1 - 1/N)$

To calculate $p(\mathbf{C})$ we need to know $p(\mathbf{C} \mid \mathbf{notX})$ i.e. the probability that the DNA samples match given that the subject is guilty. We can safely assume that this is a certainty.

$$p(\mathbf{C}) = p(\mathbf{C} \mid \mathbf{X}) \times p(\mathbf{X}) + p(\mathbf{C} \mid \mathbf{notX}) \times p(\mathbf{notX})$$
$$p(\mathbf{C}) = p_{match} \times (1 - 1/N) + 1 \times (1/N)$$
$$p(\mathbf{X} \mid \mathbf{C}) = p(\mathbf{C} \mid \mathbf{X}) \times \frac{p(\mathbf{X})}{p(\mathbf{C})}$$
$$p(\mathbf{X} \mid \mathbf{C}) = p_{match} \times \frac{p_{match} \times (1 - 1/N)}{p_{match} \times (1 - 1/N) + 1 \times (1/N)}$$
$$p(\mathbf{X} \mid \mathbf{C}) = \frac{p_{match} \times (N - 1)}{p_{match}(N - 1) + 1}$$

Now since N is a large number, we can put $(N - 1)$ equal to N so this simplifies to:

$$p(\mathbf{X} \mid \mathbf{C}) = \frac{p_{match} \times N}{p_{match} \times N + 1}$$

The crucial feature of this result is to note the relative sizes of $p_{match} \times N$ and 1. If N is so large than $p_{match} \times N$ is greater than 1, then the fraction tends to 1 – i.e. the subject is very probably innocent; but if we can restrict the number of possible subjects to less than a few million, say the population of New York, then the denominator is virtually equal to 1 and the formula becomes simpler still:

$$p(\mathbf{X} \mid \mathbf{C}) = p_{match} \times N$$

If, in order to secure a conviction, we need to establish a probability of less than 1 in a million, then with p_{match} equal to 1 in a billion, N must be less than 1000. DNA evidence is useless on its own, but if other evidence can reduce the number of possible suspects to less than 1000 individuals then the DNA evidence can be crucial.