1 Probability, Conditional Probability and Bayes Formula

The intuition of chance and probability develops at very early ages.[1] However, a formal, precise definition of the probability is elusive.

If the experiment can be repeated potentially infinitely many times, then the probability of an event can be defined through relative frequencies. For instance, if we rolled a die repeatedly, we could construct a frequency distribution table showing how many times each face came up. These frequencies (n_i) can be expressed as proportions or relative frequencies by dividing them by the total number of tosses n : $f_i = n_i/n$. If we saw six dots showing on 107 out of 600 tosses, that face's proportion or relative frequency is $f_6 = 107/600 = 0.178$ As more tosses are made, we "expect" the proportion of sixes to stabilize around $\frac{1}{6}$.

Famous Coin Tosses: Buffon tossed a coin 4040 times. Heads appeared 2048 times. K. Pearson tossed a coin 12000 times and 24000 times. The heads appeared 6019 times and 12012, respectively. For these three tosses the relative frequencies of heads are 0.5049, 0.5016, and 0.5005.

What if the experiments can not be repeated? For example what is probability that Squiki the guinea pig survives its first treatment by a particular drug. Or "the experiment" of you taking ISyE8843 course in Fall 2004. It is legitimate to ask for the probability of getting a grade of an A. In such cases we can define probability **subjectively** as a measure of strength of belief.

Figure 1: A gem proof condition 1913 Liberty Head nickel, one of only five known and the finest of the five. Collector Jay Parrino of Kansas City bought the elusive nickel for a record $1,485,000, the first and only time an American coin has sold for over $1 million.

Tutubalin's Problem. In a desk drawer in the house of Mr Jay Parrino of Kansas City there is a coin, 1913 Liberty Head nickel. What is the probability that the coin is heads up?

The *symmetry* properties of the experiment lead to the classical definition of probability. An ideal die is symmetric. All sides are "equiprobable". The probability of 6, in our example is a ratio of the number of *favorable* outcomes (in our example only one favorable outcome, namely, 6 itself) and the number of all possible outcomes, 1/6.[2]

---

[1]Piaget, J. and Inhelder B. *The Origin of the Idea of Chance in Children*, W. W. Norton & Comp., N.Y.

[2]This definition is attacked by philosophers because of the fallacy called *circulus vitiosus*. One defines the notion of probability supposing that outcomes are equi**probable**.

> (**Frequentist**) An event's **probability** is the proportion of times that we expect the event to occur, if the experiment were repeated a large number of times.
>
> (**Subjectivist**) A subjective **probability** is an individual's degree of belief in the occurrence of an event.
>
> (**Classical**) An event's **probability** is the ratio of the number of favorable outcomes and possible outcomes in a (symmetric) experiment.

Term	Description	Example
Experiment	Phenomenon where outcomes are uncertain	Single throws of a six-sided die
Sample space	Set of all outcomes of the experiment	$S = \{1, 2, 3, 4, 5, 6\}$, (1, 2, 3, 4, 5, *or* 6 dots show)
Event	A collection of outcomes; a subset of S	$A = \{3\}$ (3 dots show), $B = \{3, 4, 5,$ *or* $6\}$ (3, 4, 5, or 6 dots show) or 'at least three dots show'
Probability	A number between 0 and 1 assigned to an event.	$P(A) = \frac{1}{6}$. $P(B) = \frac{4}{6}$.

Sure event occurs *every time* an experiment is repeated and has the probability 1. Sure event is in fact the sample space $\mathcal{S}$.

An event that *never* occurs when an experiment is performed is called **impossible event**. The probability of an impossible event, denoted usually by $\emptyset$ is 0.

> For any event A, the probability that A will occur is a number between 0 and 1, inclusive:
> $$0 \leq P(A) \leq 1,$$
> $$P(\emptyset) = 0, \quad P(\mathcal{S}) = 1.$$

The **intersection** (product) $A \cdot B$ of two events A and B is an event that occurs if both events A **and** B occur. The key word in the definition of the intersection is **and.**

In the case when the events A and B are independent the probability of the intersection is the product of probabilities: $P(A \cdot B) = P(A)P(B)$.

Example: The outcomes of two consecutive flips of a fair coin are independent events.

Events are said to be **mutually exclusive** if they have no outcomes in common. In other words, it is impossible that both could occur in a single trial of the experiment. For mutually exclusive events holds $P(A \cdot B) = P(\emptyset) = 0$.

In the die-toss example, events $A = \{3\}$ and $B = \{3, 4, 5, 6\}$ are not mutually exclusive, since the outcome $\{3\}$ belongs to both of them. On the other hand, the events $A = \{3\}$ and $C = \{1, 2\}$ are mutually exclusive.

The union $A \cup B$ of two events A and B is an event that occurs if at least one of the events A or B occur. The key word in the definition of the union is **or.**

For mutually exclusive events, the probability that at least one of them occurs is

$$P(A \cup C) = P(A) + P(C)$$

For example, if the probability of event $A = \{3\}$ is 1/6, and the probability of the event $C = \{1, 2\}$ is 1/3, then the probability of A or C is

$$P(A \cup C) = P(A) + P(C) = 1/6 + 1/3 = 1/2.$$

The *additivity* property is valid for any number of mutually exclusive events $A_1, A_2, A_3, \ldots$:

$$P(A_1 \cup A_2 \cup A_3 \cup \ldots) = P(A_1) + P(A_2) + P(A_3) + \ldots$$

What is $P(A \cup B)$ if the events A and B are not mutually exclusive.

> For any two events A and B, the probability that either A or B will occur is given by the **inclusion-exclusion** rule
>
> $$P(A \cup B) = P(A) + P(B) - P(A \cdot B)$$

If the events A abd B are exclusive, then $P(A \cdot B) = 0$, and we get the familiar formula $P(A \cup B) = P(A) + P(B)$.

The inclusion-exclusion rule can be generalized to unions of arbitrary number of events. For example, for three events A, Ba and C, the rule is:

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cdot B) - P(A \cdot C) - P(B \cdot C) + P(A \cdot B \cdot C).$$

For every event defined on $\mathcal{S}$, we can define a counterpart-event called its **complement**. The complement A^c of an event A consists of all outcomes that are in $\mathcal{S}$, but are *not* in A. The key word in the definition of an complement is **not.** In our example, A^c consists of the outcomes: $\{1, 2, 3, 4, 5\}$.

The events A and A^c are mutually exclusive by definition. Consequently,

$$P(A \cup A^c) = P(A) + P(A^c)$$

Since we also know from the definition of A^c that it includes all the events in the sample space, S, that are not in A, so

$$P(A) + P(A^c) = P(\mathcal{S}) = 1$$

> For any complementary events A and A^c,
> $$P(A) + P(A^c) = 1, \quad P(A) = 1 - P(A^c), \quad P(A^c) = 1 - P(A)$$

These equations simplify solutions of some probability problems. If $P(A^c)$ is easier to calculate than $P(A)$, then $P(A^c)$ and equations above let us obtain $P(A)$ indirectly.

This and some other properties of probability are summarized in table below.

Property	Notation
If event S will *always* occur, its probability is 1.	$P(\mathcal{S}) = 1$
If event $\emptyset$ will *never* occur, its probability is 0.	$P(\emptyset) = 0$
Probabilities are always between 0 and 1, inclusive	$0 \leq P(A) \leq 1$
If $A, B, C, \ldots$ are all mutually exclusive then $P(A \cup B \cup C \ldots)$ can be found by addition.	$P(A \cup B \cup C \ldots) = P(A) + P(B) + P(C) + \ldots$
If A and B are mutually exclusive then $P(A \cup B)$ can be found by addition.	$P(A \cup B) = P(A) + P(B)$
Addition rule: The general *addition rule* for probabilities	$P(A \cup B) = P(A) + P(B) - P(A \cdot B)$
Since A and A^c are mutually exclusive and between them include all possible outcomes, $P(A \cup A^c)$ is 1.	$P(A \cup A^c) = P(A) + P(A^c) = P(\mathcal{S}) = 1$, and $P(A^c) = 1 - P(A)$

2   Conditional Probability and Independence

A **conditional probability** is the probability of one event if another event occurred. In the "die-toss" example, the probability of event A, three dots showing, is $P(A) = \frac{1}{6}$ on a single toss. But what if we know that event B, at least three dots showing, occurred? Then there are only four possible outcomes, one of which is A. The probability of $A = \{3\}$ is $\frac{1}{4}$, *given* that $B = \{3, 4, 5, 6\}$ occurred. The *conditional probability of A given B* is written $P(A|B)$.

$$P(A|B) = \frac{P(A \cdot B)}{P(B)}$$

Event A is **independent** of B if the conditional probability of A given B is the same as the unconditional probability of A. That is, they are independent if

$$P(A|B) = P(A)$$

In the die-toss example, $P(A) = \frac{1}{6}$ and $P(A|B) = \frac{1}{4}$, so the events A and B are not independent.

4

> The probability that two events A and B will both occur is obtained by applying the **multiplication rule:**
>
> $$P(A \cdot B) = P(A)P(B|A) = P(B)P(A|B)$$
>
> where $P(A|B)$ $(P(B|A))$ means the probability of A given B (B given A).

For independent events only, the equation in the box simplifies to

$$P(A \cdot B) = P(A)P(B).$$

• Prove $P(A_1 A_2 \ldots A_n) = P(A_1|A_2 \ldots A_n)\, P(A_2|A_3 \ldots A_n)\, \ldots P(A_{n-1}|A_n)\, P(A_n)$.

Example: Let the experiment involves a random draw from a standard deck of 52 playing cards. Define events A and B to be "the card is ♠ and "the card is queen". Are the events A and B independent? By definition, $P(A\&B) = P(Q♠) = \frac{1}{52}$. This is the product of $P(♠) = \frac{13}{52}$ and $P(Q) = \frac{4}{52}$, and A and B in question are independent. The intuition barely helps here. Pretend that from the original deck of cards 2♡ is excluded prior to the experiment. Now the events A and B become dependent since

$$P(A) \cdot P(B) = \frac{13}{51} \cdot \frac{4}{51} \neq \frac{1}{51} = P(A\&B).$$

The multiplication rule tells us how to find probabilities for composite event $(A \cdot B)$. The probability of $(A \cdot B)$ is used in the general *addition rule* for finding the probability of $(A \cup B)$.

Rule	Notation
Definitions	
The *conditional probability* of A given B is the probability of event A, if event B occurred.	$P(A|B)$
A is *independent* of B if the conditional probability of A given B is the same as the unconditional probability of A.	$P(A|B) = P(A)$
Multiplication rule:	
The general *multiplication rule* for probabilities	$P(A \cdot B) = P(A)P(B|A) = P(B)P(A|B)$
For *independent events* only, the multiplication rule is simplified.	$P(A \cdot B) = P(A)P(B)$

3   Pairwise and Global Independence

If three events A, B, and C are such that any of the pairs are exclusive, i.e., $AB = \emptyset$, $AC = \emptyset$ or $BC = \emptyset$, then the events are mutually exclusive, i.e, $ABC = \emptyset$. In terms of independence the picture is different. Even if the events are *pairwise* independent for all three pairs A, B; A, C; and B, C, i.e., $P(AB) = P(A)P(B)$, $P(AC) = P(A)P(C)$, and $P(BC) = P(B)P(C)$, they may be dependent in the totality, $P(ABC) \neq P(A)P(B)P(C)$.

Here is one example.

The four sides of a tetrahedron (regular three sided pyramid with 4 sides consisting of isosceles triangles) are denoted by 2, 3, 5 and 30, respectively. If the tetrahedron is "rolled" the number on the basis is the outcome of interest. Let the three events are A-the number on the base is even, B-the number is divisible by 3, and C-the number is divisible by 5. The events are pairwise independent, but in totality dependent.

Algebra is clear here, but what is the intuition. The "trick" is that the events AB, AC, BC and ABC all coincide. That is $P(A|BC) = 1$ although $P(A|B) = P(A|C) = P(A)$.

The concept of independence (dependence) is not transitive. At first glance that may seem not correct, as one may argue as follows, "If event A depends on B, and event B depends on C, then event A should depend on the event C." A simple exercise proves the above statement wrong.

Take a standard deck of 52 playing cards and replace the $\clubsuit Q$ with $\diamondsuit Q$. Thus the deck now still has 52 cards, two $\diamondsuit Q$ and no $\clubsuit Q$. From such a deck draw a card at random and consider three events: A-the card is queen, B-the card is red, and C- the card is $\heartsuit$. It ie easy to see that A and B are dependent since $P(A\&B) = 3/52 \neq P(A) \cdot P(B) = 4/52 \cdot 27/52$. The events B and C are dependent as well since the event C is contained in B, and $P(BC) = P(C) \neq P(B) \cdot P(C)$. However, the events A and C are independent, since $P(AC) = P(\heartsuit Q) = 1/52 = P(A)P(C) = 13/53 \cdot 4/52$.

4 Total Probability

Events $H_1, H_2, \ldots, H_n$ form a partition of the sample space $\mathcal{S}$ if
(i) They are mutually exclusive ($H_i \cdot H_j = \emptyset$, $i \neq j$) and
(ii) Their union is the sample space $\mathcal{S}$, $\bigcup_{i=1}^n H_i = \mathcal{S}$.

The events $H_1, \ldots, H_n$ are usually called **hypotheses** and from their definition follows that $P(H_1) + \cdots + P(H_n) = 1 \ (= P(\mathcal{S}))$.

Let the event of interest A happens under any of the hypotheses H_i with a known (conditional) probability $P(A|H_i)$. Assume, in addition, that the probabilities of hypotheses $H_1, \ldots, H_n$ are known. Then $P(A)$ can be calculated using the *total probability formula*.

Total Probability Formula.

$$P(A) = P(A|H_1)P(H_1) + \cdots + P(A|H_n)P(H_n).$$

The probability of A is the weighted average of the conditional probabilities $P(A|H_i)$ with weights $P(H_i)$.

Stanley. Stanley takes an oral exam in statistics by answering 3 questions from an examination card drawn at random from the set of 20 cards. There are 8 favorable cards (Stanley knows answers on all 3 questions). Stanley will get a grade A if he answers all 3 questions. What is the probability for Stanley to get an A if he draws the card
(a) first

(b) second

(c) third?

Solution: Denote with A the event that Stanley draws a favorable card (and consequently gets an A).

(i) If he draws the card first, then clearly $P(A) = 8/20 = 2/5$.

(ii) If Stanley draws the second, then one card was taken by the student before him. That first card taken might have been favorable (hypothesis H_1) or unfavorable (hypothesis H_2). Obviously, the hypotheses H_1 and H_2 partition the sample space since no other type of cards is possible, in this context. Also, the probabilities of H_1 and H_2 are 8/20 and 12/20, respectively. Now, after one card has been taken Stanley draws the second. If H_1 had happened, probability of A is 7/19, and if H_2 had happened, the probability of A is 8/19. Thus, $P(A|H_1) = 7/19$ and $P(A|H_2) = 8/19$. By the total probability formula, $P(A) = 7/19 \cdot 8/20 + 8/19 \cdot 12/20 = 8/20 = 2/5$.

(iii) Stanley has the same probability of getting an A after two cards have been already taken. The hypotheses are H_1={ both cards taken favorable }, H_2={ exactly one card favorable }, and H_3={ none of the cards taken favorable }. $P(H_1) = 8/20 \cdot 7/19$, $P(H_3) = 12/20 \cdot 11/19$. and $P(H_2) = 1 - P(H_1) - P(H_3)$. Next, $P(A|H_1) = 6/18, P(A|H_2) = 7/18$, and $P(A|H_3) = 8/18$. Therefore, $P(A) = 6/18 \cdot 7/19 \cdot 8/20 + 7/18 \cdot ... + 8/18 \cdot 11/19 \cdot 12/20 = 12/20$.

Moral: Stanley's lack on the exam does not depend on the order in drawing the examination card.

Two-headed coin Out of 100 coins one has heads on both sides. One coin is chosen at random and flipped two times. What is the probability to get

(a) two heads?

(b) two tails?

Solution:

(a) Let A be the event that two heads are obtained. Denote by H_1 the event (hypothesis) that a fair coin was chosen. The hypothesis $H_2 = H_1^c$ is the event that the two-headed coin was chosen.

$$P(A) = P(A|H_1)P(H_1) + P(A|H_2)P(H_2) = 1/4 \cdot 99/100 + 1 \cdot 1/100 = 103/400 = 0.2575.$$

(b) Exercise. [Ans. 0.2475]

5   Bayes' Formula

Recall that multiplication rule claims:

$$P(AH) = P(A)P(H|A) = P(H)P(A|H).$$

This simple identity is the essence of **Bayes' Formula.**

Bayes Formula. Let the event of interest A happens under any of hypotheses H_i with a known (conditional) probability $P(A|H_i)$. Assume, in addition, that the probabilities of hypotheses $H_1, \ldots, H_n$ are known (*prior* probabilities). Then the conditional (*posterior*) probability of the hypothesis H_i, $i = 1, 2, \ldots, n$, given that event A happened, is

$$P(H_i|A) = \frac{P(A|H_i)P(H_i)}{P(A)},$$

where

$$P(A) = P(A|H_1)P(H_1) + \cdots + P(A|H_n)P(H_n).$$

Assume that out of N coins in a box, one has heads at both sides. Such "two-headed" coin can be purchased in Spencer stores. Assume that a coin is selected at random from the box, and without inspecting it, flipped k times. All k times the coin landed up heads. What is the probability that two headed coin was selected?

Denote with A_k the event that randomly selected coin lands heads up k times. The hypotheses are H_1-the coin is two headed, and H_2 the coin is fair. It is easy to see that $P(H_1) = 1/N$ and $P(H_2) = (N-1)/N$. The conditional probabilities are $P(A_k|H_1) = 1$ for any k, and $P(A_k|H_2) = 1/2^k$.

By total probability formula,

$$P(A_k) = \frac{2^k + N - 1}{2^k N},$$

and

$$P(H_1|A_k) = \frac{2^k}{2^k + N - 1}.$$

For $N = 1,000,000$ and $k = 1, 2, \ldots, 30$ the graph of posterior probabilities is given in Figure 2

It is interesting that our prior probability $P(H_1) = 0.000001$ jumps to posterior probability of 0.9991, after observing 30 heads in a row. The matlab code `bayes1_1.m` producing Figure 2 is given in the Programs/Codes on the GTBayes Page.

Prosecutor's Fallacy The prosecutor's fallacy is a fallacy commonly occurring in criminal trials but also in other various arguments involving rare events. It consists of subtle exchanging of $P(A|B)$ for $P(B|A)$. A zealous prosecutor has collected an evidence, say fingerprint match, and has an expert testify that the probability of finding this evidence if the accused were innocent is tiny. The fallacy is committed the prosecutor proceeds to claim that the probability of the accused being innocent is comparably tiny.

Why is this incorrect? Suppose there is a one-in-a-million chance of a match given that the accused is innocent. The prosecutor deduces that means there is only a one-in-a-million chance of innocence. But in a community of 10 million people, one expects 10 matches, and the accused is just one of those ten. That would indicate only a one-in-ten chance of guilt, if no other evidence is available.
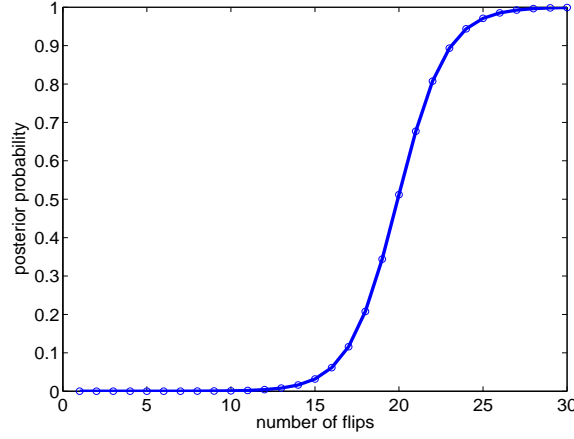
Figure 2: Posterior probability of the "two-headed" coin for $N = 1,000,000$ if k heads appeared.

The fallacy thus lies in the fact that the *a priori* probability of guilt is not taken into account. If this probability is small, then the only effect of the presented evidence is to increase that probability somewhat, but not necessarily dramatically.

As a real-life example, consider the case of Sally Clark [1], who was accused in 1998 of having killed her first child at 11 weeks of age, then conceived another child and killed it at 8 weeks of age. The prosecution had expert witness testify that the probability of two children dying from sudden infant death syndrome is about 1 in 73 million. To provide proper context for this number, the probability of a mother killing one child, conceiving another and killing that one too, should have been estimated and compared to the 1 in 73 million figure, but it wasn't. Ms. Clark was convicted in 1999, resulting in a press release by the UK Royal Statistical Society which pointed out the mistake. Sally Clark's conviction was eventually quashed on other grounds on appeal on 29th January 2003.

In another scenario, assume a burglary has been committed in a town, and 10,000 men in the town have their fingerprints compared to a sample from the crime. One of these men has matching fingerprint, and at his trial, it is testified that the probability that two fingerprint profiles match by chance is only 1 in 20,000. This does not mean the probability that the suspect is innocent is 1 in 20,000. Since 10,000 men were taken fingerprints, there were 10,000 opportunities to find a match by chance; the probability that there was at least one fingerprint match is $1 - \binom{10000}{0} \cdot \left(\frac{1}{20000}\right)^0 \cdot \left(1 - \frac{1}{20000}\right)^{10000}$, about 39% – considerably more than 1 in 20,000. (The probability that exactly one of the 10,000 men has a match is $\binom{10000}{1} \cdot \left(\frac{1}{20000}\right)^1 \cdot \left(1 - \frac{1}{20000}\right)^{9999}$ or about 30%, which certainly casts a reasonable doubt.)

Return to our example with N coins. Assume that out of $N = 1000000$ coins in a box, one has heads at both sides. Such "two-headed" coin is "guilty." Assume that a coin is selected at random from the box, and without inspecting it, flipped $k = 15$ times. All $k = 15$ times the coin landed up heads. Based on this evidence the "prosecutor" claims the selected coin is guilty since if it was "innocent," observed $k = 15$ heads in a row will appear with probability $\frac{1}{2^{15}} \approx 0.00003$. But the probability that the "guilty" coin was really selected is $\frac{1}{1+999999/2^{15}} \approx 0.03$ and prosecutor is accusing an "innocent" coin with probability of approximately 0.97.

In legal terms, the prosecutor is operating in terms of a presumption of guilt, something which is contrary to the normal presumption of innocence where a person is assumed to be innocent unless found guilty.

Two Masked Robbers. Two masked robbers try to rob a crowded bank during the lunch hour but the teller presses a button that sets off an alarm and locks the front door. The robbers realizing they are trapped, throw away their masks and disappear into the chaotic crowd. Confronted with 40 people claiming they are innocent, the police gives everyone a lie detector test. Suppose that guilty people are detected with probability 0.85 and innocent people appear to be guilty with probability 0.08. What is the probability that Mr. Smith was one of the robbers given that the lie detector says he is?

Guessing. Subject in an experiment are told that either a red or a green light will flash. Each subject is to guess which light will flash. The subject is told that the probability of a red light is 0.7, independent of guesses. Assume that the subject is a probability matcher- that is , guesses red with probability .70 and green with probability .30.
    (i) What is the probability that the subject guesses correctly?
    (ii) Given that a subject guesses correctly, what is the probability that the light flashed red?

False Positives. False positives are a problem in any kind of test: no test is perfect, and sometimes the test will incorrectly report a positive result. For example, if a test for a particular disease is performed on a patient, then there is a chance (usually small) that the test will return a positive result even if the patient does not have the disease. The problem lies, however, not just in the chance of a false positive prior to testing, but determining the chance that a positive result is in fact a false positive. As we will demonstrate, using Bayes' theorem, if a condition is rare, then the majority of positive results may be false positives, even if the test for that condition is (otherwise) reasonably accurate.

Suppose that a test for a particular disease has a very high success rate:

    • if a tested patient has the disease, the test accurately reports this, a 'positive', 99% of the time (or, with probability 0.99), and

    • if a tested patient does not have the disease, the test accurately reports that, a 'negative', 95% of the time (i.e. with probability 0.95).

Suppose also, however, that only 0.1% of the population have that disease (i.e. with probability 0.001). We now have all the information required to calculate the probability that, given the test was positive, that it is a false positive.

Let D be the event that the patient has the disease, and P be the event that the test returns a positive result. The probability of a true positive is

$$P(D|P) = \frac{0.99 \times 0.001}{0.99 \times 0.001 + 0.05 \times 0.999} \approx 0.019,$$

and hence the probability of a false positive is about (1 - 0.019) = 0.981.

Despite the apparent high accuracy of the test, the incidence of the disease is so low (one in a thousand) that the vast majority of patients who test positive (98 in a hundred) do not have the disease. Nonetheless, this is 20 times the proportion before we knew the outcome of the test! The test is not useless, and re-testing may improve the reliability of the result. In particular, a test must be very reliable in reporting a negative result when the patient does not have the disease, if it is to avoid the problem of false positives. In mathematical terms, this would ensure that the second term in the denominator of the above calculation is small, relative to the first term. For example, if the test reported a negative result in patients without the disease with probability 0.999, then using this value in the calculation yields a probability of a false positive of roughly 0.5.

Multiple Choice. A student answers a multiple choice examination question that has 4 possible answers. Suppose that the probability that the student knows the answer to the question is 0.80 and the probability that the student guesses is 0.20. If student guesses, probability of correct answer is 0.25.
   (i) What is the probability that the fixed question is answered correctly?
   (ii) If it is answered correctly what is the probability that the student really knew the correct answer.

Manufacturing Bayes. Factory has three types of machines producing an item. Probabilities that the item is I quality f it is produced on i-th machine are given in the following table:

machine	probability of I quality
1	0.8
2	0.7
3	0.9

The total production is done 30% on type I machine, 50% on type II, and 20% on type III.
One item is selected at random from the production.
   (i) What is the probability that it is of I quality?
   (ii) If it is of first quality, what is the probability that it was produced on the machine I?

Two-headed coin. 4 One out of 1000 coins has two tails. The coin is selected at random out of these 1000 coins and flipped 5 times. If tails appeared all 5 times, what is the probability that the selected coin was 'two-tailed'?

Kokomo, Indiana. In Kokomo, IN, 65% are conservatives, 20% are liberals and 15% are independents.
   Records show that in a particular election 82% of conservatives voted, 65% of liberals voted and 50% of independents voted.
   If the person from the city is selected at random and it is learned that he/she did not vote, what is the probability that the person is liberal?

Inflation and Unemployment. Businesses commonly project revenues under alternative economic scenarios. For a stylized example, inflation could be high or low and unemployment could be high or low. There are four possible scenarios, with the assumed probabilities:

Scenario	Inflation	Unemployment	Probability
1	high	high	0.16
2	high	low	0.24
3	low	high	0.36
4	low	low	0.24

(i) What is the probability of high inflation?
(ii) What is the probability of high inflation if unemployment is high?
(iii) Are inflation and unemployment independent?

Information Channel. One of three words AAAA, BBBB, and CCCC is transmitted via an information channel. The probabilities of these words are 0.3, 0.5, and 0.2, respectively. Each letter is transmitted independently of the other letters and it is received correctly with probability 0.6. Since the channel is not perfect, the letter can change to one of the other two letters with equal probabilities of 0.2. What is the probability that word AAAA had been submitted if word ABCA was received.

Jim Albert's Question. An automatic machine in a small factory produces metal parts. Most of the time (90% by long records), it produces 95% good parts and the remaining have to be scrapped. Other times, the machine slips into a less productive mode and only produces 70% good parts. The foreman observes the quality of parts that are produced by the machine and wants to stop and adjust the machine when she believes that the machine is not working well. Suppose that the first dozen parts produced are given by the sequence

s	u	s	s	s	s	s	s	s	u	s	u

where **s** – satisfactory and **u** – unsatisfactory. After observing this sequence, what is the probability that the machine is in its good state? If the foreman wishes to stop the machine when the probability of "good state" is under 0.7, when should she stop?

Medical Doctors and Probabilistic Reasoning. The following problem was posed by Casscells, Schoenberger, and Grayboys (1978) to 60 students and staff at Harvard Medical School: *If a test to detect a disease whose prevalence is 1/1000 has a false positive rate of 5%, what is the chance that a person found to have a positive result actually has the disease, assuming you know nothing about the person's symptoms or signs?*

Assuming that the probability of a positive result given the disease is 1, the answer to this problem is approximately 2%. Casscells et al. found that only 18% of participants gave this answer. The modal response was 95%, presumably on the supposition that, because an error rate of the test is 5%, it must get 95% of results correct.

Let's Make a Deal. In the popular television game show *Let's Make a Deal,* Monty Hall is the master of ceremonies. At certain times during the show, a contestant is allowed to choose one of three identical doors A, B, C, behind only one of which is a valuable prize (a new car). After the contestant picks a door (say, door A), Monty Hall opens another door and shows the contestant that there is no prize behind that door. (Monty Hall knows where the prize is and always chooses a door where there is no prize.) He then asks the contestant whether he or she wants to stick with their choice of door or switch to the remaining unopened door. Should the contestant switch doors? Does it matter?

A federalist paper resolved? Suppose that a work, the author of which is known to be either Madson or Hamilton, contains a certain key phrase. Suppose further that this phrase occurs in 60% of the papers known to have been written by Madison, but in only 20% of those by Hamilton. Finally, suppose a historian gives subjective probability .3 to the event that the author is Madison (and consequently .7 to the complementary event that the author is Hamilton). Compute the historian's posterior probability for the event that the author is Madison.

References

[1] Batt, J. Stolen Innocence: A Mother's Fight for Justice. The Authorised Story of Sally Clark, Ebury Press.

[2] Barbeau, E. (1993). The Problem of the Car and Goats, *CMJ*, 24:2, p. 149

[3] Casscells, W., Schoenberger, A., and Grayboys, T. (1978). Interpretation by physicians of clinical laboratory results. *New England Journal of Medicine*, **299**, 999-1000.

[4] Gillman, L. (1992). The Car and the Goats,*AMM* 99:1, p. 3

[5] Selvin, S. (1975). A Problem in Probability, *American Statistician*, 29:1, p. 67