STATISTICS-Sem 2-Gen-Theorem of Total Probability-Saptarshi Mondal-30 Mar 2020

Law of Total Probability

Consider a very simple question: In a certain country there are three provinces, call them B1, B2, and B3 (i.e., the country is partitioned into three disjoint sets B1, B2, and B3). We are interested in the total forest area in the country. Suppose that we know that the forest area in B1, B2, and B3 are 100 sq km, 50 sq km, and 150 sq km, respectively. What is the total forest area in the country? If your answer is

$$100 \text{ sq km} + 50 \text{ sq km} + 150 \text{ sq km} = 300 \text{ sq km},$$

you are right. That is, you can simply add forest areas in each province (partition) to obtain the forest area in the whole country. This is the idea behind the law of total probability, in which the *area of forest* is replaced by *probability of an event* A. In particular, if you want to find P(A), you can look at a partition of S, and add the amount of probability of A that falls in each partition. We have already seen the special case where the partition is B and B': we saw that for any two events A and B,

$$P(A)=P(A\cap B)+P(A\cap B')$$

and using the definition of conditional probability, $P(A\cap B)=P(A|B)P(B)$, we can write

$$P(A)=P(A|B)P(B)+P(A|B')P(B').$$

We can state a more general version of this formula which applies to a general partition of the sample space S.

Law of Total Probability:

If $B1,B2,B3,\cdots$ is a partition of the sample space S, then for any event A we have

$$P(A)=\sum P(A\cap Bi)=\sum P(A|Bi)P(Bi).$$

Here is a proof of the law of total probability using probability axioms:

Proof

Since $B1,B2,B3,\cdots$ is a partition of the sample space S, we can write

$$S = \cup Bi$$

$$\begin{aligned} A &= A\cap S \\ &= A\cap(\cup Bi) \\ &= \cup(A\cap Bi) \text{ by the distributive law.} \end{aligned}$$

Now note that the sets $A\cap Bi$ are disjoint (since the Bi's are disjoint). Thus, by the third probability axiom,

$$P(A)=P(\cup(A\cap Bi))=\sum P(A\cap Bi)=\sum P(A|Bi)P(Bi).$$

Here is a typical scenario in which we use the law of total probability. We are interested in finding the probability of an event A, but we don't know how to find P(A) directly. Instead, we know the conditional probability of A given some events Bi, where the Bi's form a partition of the sample space. Thus, we will be able to find P(A) using the law of total probability, $P(A)=\sum P(A|Bi)P(Bi)$.

Example

I have three bags that each contain 100 marbles:

- Bag 1 has 75 red and 25 blue marbles;
- Bag 2 has 60 red and 40 blue marbles;
- Bag 3 has 45 red and 55 blue marbles.

I choose one of the bags at random and then pick a marble from the chosen bag, also at random. What is the probability that the chosen marble is red?

Solution: Let R be the event that the chosen marble is red. Let Bi be the event that I choose Bag i. We already know that

$$P(R|B1)=0.75,$$
$$P(R|B2)=0.60,$$
$$P(R|B3)=0.45$$

We choose our partition as B1,B2,B3. Note that this is a valid partition because, firstly, the Bi's are disjoint (only one of them can happen), and secondly, because their union is the entire sample space as one the bags will be chosen for sure, i.e., $P(B1\cup B2\cup B3)=1$. Using the law of total probability, we can write

$$\begin{aligned} P(R) &= P(R|B1)P(B1)+P(R|B2)P(B2)+P(R|B3)P(B3) \\ &= (0.75)1/3+(0.60)1/3+(0.45)1/3 \\ &= 0.60 \end{aligned}$$

Sometimes the probabilities needed for the calculation of total probability isn't specified in the exact way you need to solve the equation. **An alternate version of the total probability rule** (found with the multiplication rule) is:

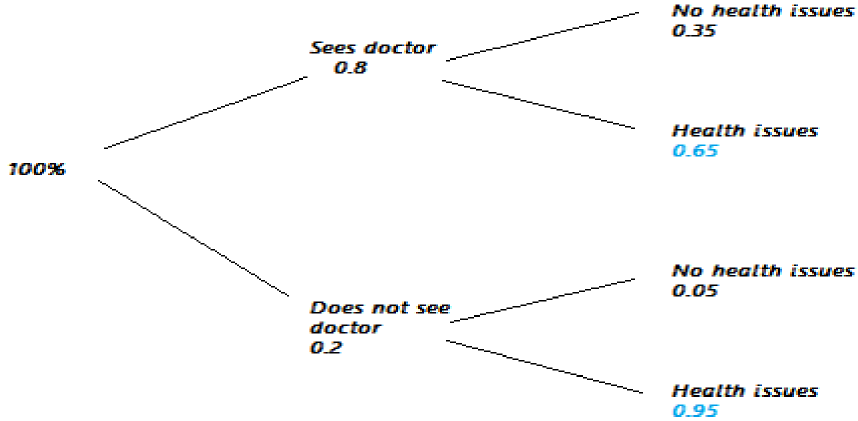
$$P(A \cap B) = P(A \mid B) * P(B) + P(A \cap B') = P(A \mid B')P(B').$$

In practical situations, it can be difficult to work with these equations. It's *much* easier to work with a tree or table.

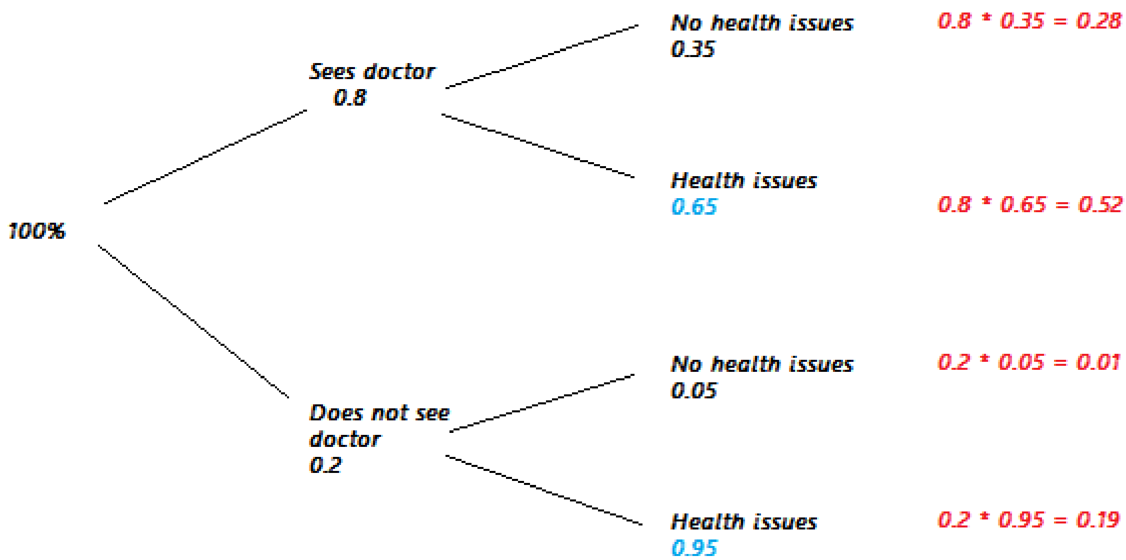
Using a Probability Tree to Find Total Probability

Sample question: 80% of people attend their primary care physician regularly; 35% of those people have no health problems crop up during the following year. Out of the 20% of people who don't see their doctor regularly, only 5% have no health issues during the following year. What is the probability a random person will have no health problems in the following year?

Step 1: Sketch out a tree. The following tree uses the information given in the question with the addition of two probabilities (in blue) obtained by the complement. For example, if 5% of people do not have health problems, that means 95% of people do have health problems.



Step 2: Multiply the probabilities for each branch. For example, the top branch has 0.8 on the first segment and 0.35 on the second. These calculations are shown in red on the graph below:



Step 3: Find the probabilities that answer the question. For this example, we want the probability a random person will have no health problems. If you look at the graph, the branches leading to "no health problems" are the top branch and the third branch down. The probabilities listed in red are 0.28 and 0.01, so the solution is: 0.28 + 0.01 = 0.29.

Extension:

Given n mutually exclusive events A_1, ..., A_n whose probabilities sum to unity, then

$P(B) = P(B \mid A_1) P(A_1) + \ldots + P(B \mid A_n) P(A_n),$

Where B is an arbitrary event, and $P(B \mid A_i)$ is the conditional probability of B assuming A_i.

Example:

A person has undertaken a mining job. The probabilities of completion of job on time with and without rain are 0.42 and 0.90 respectively. If the probability that it will rain is 0.45, then determine the probability that the mining job will be completed on time.

Solution:

Let A be the event that the mining job will be completed on time and B be the event that it rains. We have,

$P(B) = 0.45$,

$P(\text{no rain}) = P(B') = 1 - P(B) = 1 - 0.45 = 0.55$

By multiplication law of probability,

$P(A|B) = 0.42$

$P(A|B') = 0.90$

Since, events B and B′ form partitions of the sample space S, by total probability theorem, we have

$P(A) = P(B)\ P(A|B) + P(B')\ P(A|B')$

$=0.45 \times 0.42 + 0.55 \times 0.9$

$= 0.189 + 0.495 = 0.684$

So, the probability that the job will be completed on time is 0.684.

EXAMPLE.

A biased coin (with probability of obtaining a Head equal to $p > 0$) is tossed repeatedly and independently until the first head is observed. Compute the probability that the first head appears at an even numbered toss.

SOLUTION: Define:

- sample space Ω to consist of all possible infinite binary sequences of coin tosses;
- event H1 - head on first toss;
- event E - first head on even numbered toss.

We want P(E) : using the Theorem of Total Probability, and the partition of Ω given by {H1, H1'}

$P(E) = P(E|H1)P(H1) + P(E|H1')P(H1')$. Now clearly, $P(E|H1) = 0$ (given H1, that a head appears on the first toss, E cannot occur) and also $P(E|H1')$ can be seen to be given by $P(E|H1') = P(E') = 1 - P(E)$, (given that a head does not appear on the first toss, the required conditional probability is merely the probability that the sequence concludes after a further odd number of tosses, that is, the probability of E'). Hence P(E) satisfies $P(E) = 0 \times p + (1 - P(E)) \times (1 - p) = (1 - p)\ (1 - P(E))$, so that

$P(E) = (1 - p)/\ (2 - p)$.

EXAMPLE

Two players A and B are competing at a trivia quiz game involving a series of questions. On any individual question, the probabilities that A and B give the correct answer are α and β respectively, for all questions, with outcomes for different questions being independent. The game finishes when a player wins by answering a question correctly. Compute the probability that A wins if (a) A answers the first question, (b) B answers the first question.

SOLUTION: Define:

- sample space Ω to consist of all possible infinite sequences of answers;
- event A - A answers the first question;
- event F - game ends after the first question;
- event W - A wins.

We want $P(W|A)$ and $P(W|A')$.

Using the Theorem of Total Probability, and the partition given by $\{F, F'\}$

$P(W|A) = P(W|A \cap F)P(F|A) + P(W|A \cap F')P(F'|A)$. Now, clearly

$P(F|A) = P$ [A answers first question correctly] $= \alpha$,

$P(F'|A) = 1 - \alpha$, and

$P(W|A \cap F) = 1$, but $P(W|A \cap F') = P(W|A')$, so that

$P(W|A) = (1 \times \alpha) + (P(W|A') \times (1 - \alpha)) = \alpha + P(W|A')\,(1 - \alpha)$. (1)

Similarly, $P(W|A') = P(W|A' \cap F)P(F|A') + P(W|A' \cap F')P(F'|A')$. We have

$P(F|A') = P$ [B answers first question correctly] $= \beta$, $P(F'|A) = 1 - \beta$,

but $P(W|A' \cap F) = 0$. Finally $P(W|A' \cap F') = P(W|A)$,

so that $P(W|A') = (0 \times \beta) + (P(W|A) \times (1 - \beta))$.

$= P(W|A)\,(1 - \beta)$. (2)

Solving (1) and (2) simultaneously gives, for (a) and (b) $P(W|A) = \alpha / (1 - (1 - \alpha)(1 - \beta))$,

$P(W|A') = (1 - \beta)\,\alpha / (1 - (1 - \alpha)(1 - \beta))$

EXAMPLE

Patients are recruited onto the two arms (0 - Control, 1-Treatment) of a clinical trial. The probability that an adverse outcome occurs on the control arm is p0, and is p1 for the treatment arm. Patients are allocated alternately onto the two arms, and their outcomes are independent. What is the probability that the first adverse event occurs on the control arm?

SOLUTION: Define:

- sample space Ω to consist of all possible infinite sequences of patient outcomes;

• event E1 - first patient (allocated onto the control arm) suffers an adverse outcome;

• event E2 - first patient (allocated onto the control arm) does not suffer an adverse outcome, but the second patient (allocated onto the treatment arm) does suffer an adverse outcome;

• event E0 - neither of the first two patients suffer adverse outcomes;

• event F - first adverse event occurs on the control arm.

We want P(F). Now the events E1, E2 and E0 partition **Ω**, so, by the Theorem of Total Probability,

$P(F) = P(F|E1)P(E1) + P(F|E2)P(E2) + P(F|E0)P(E0)$.

Now, $P(E1) = p0$, $P(E2) = (1 - p0)\, p1$, $P(E0) = (1 - p0)\,(1 - p1)$ and $P(F|E1) = 1$, $P(F|E2) = 0$. Finally, as after two non-adverse outcomes the allocation process effectively re-starts, we have $P(F|E0) = P(F)$. Hence $P(F) = (1 \times p0) + (0 \times (1 - p0)\, p1) + (P(F) \times (1 - p0)\,(1 - p1))$

$$= p0 + (1 - p0)\,(1 - p1)\, P(F),$$

which can be re-arranged to give $P(F) = p0 / (p0 + p1 - p0p1)$.