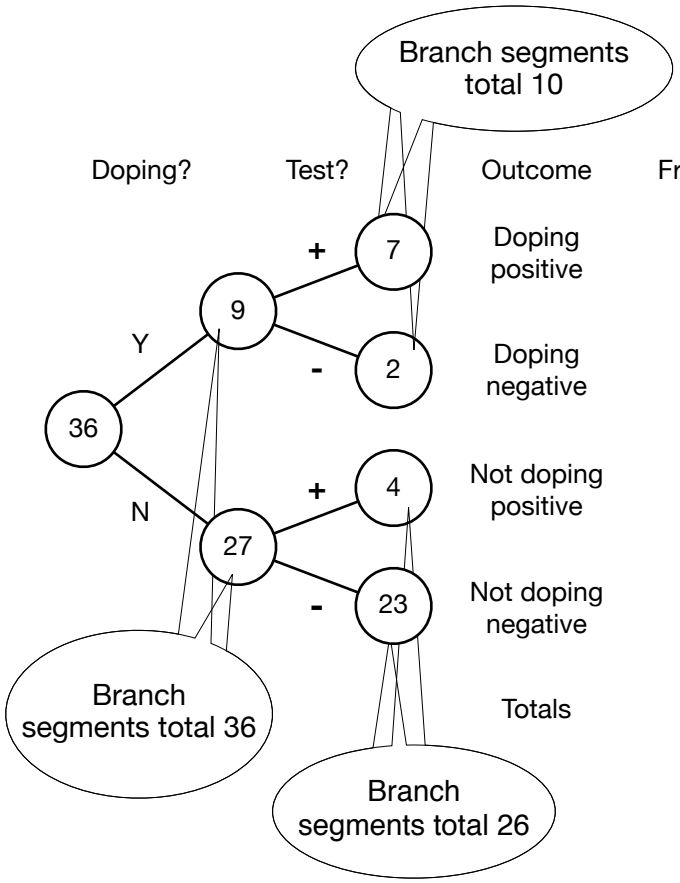
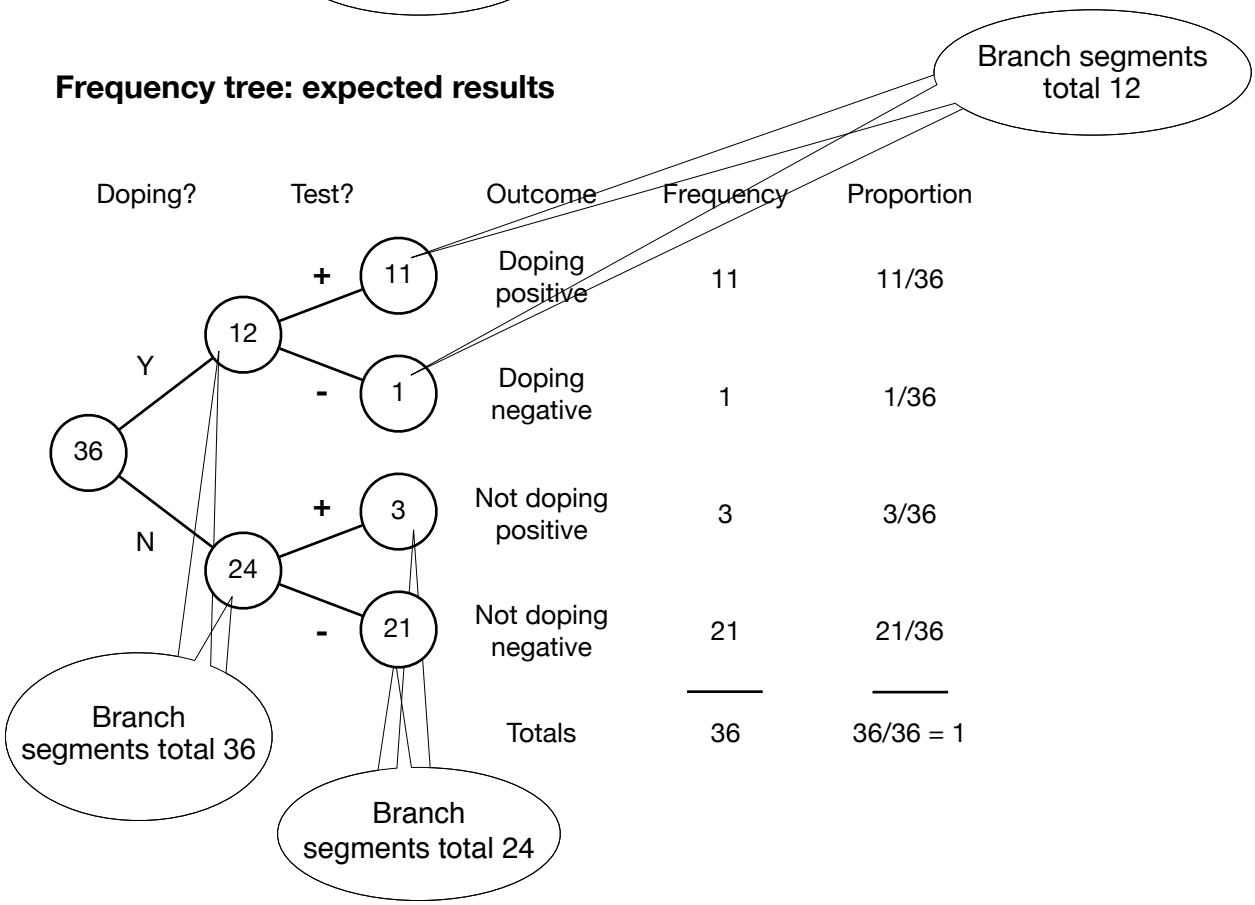
Frequency tree: sample experimental results

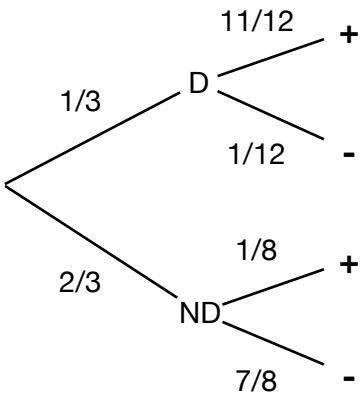


The structure is better preserved by keeping proportions out of 36 - a denominator which has meaning in this experiment (unlike18 or 9).

Frequency tree: expected results



Probability Tree



Outcome	Expected proportion	Probability
Doping positive	11/36	11/36 = 1/3 x 11/12
Doping negative	1/36	1/36 = 1/3 x 1/12
Not doping positive	3/36 = 1/12	1/12 = 2/3 x 1/8
Not doping negative	21/36 = 7/12	7/12 = 2/3 x 7/8
Totals	1	1

On a frequency tree, the total for each pair of branch segments is equal to the previous value, from which they fork. On a probability tree the sum of the probabilities on each pair of branch segments is 1. In this problem, the probabilities change, because the events on the branch segments are different - whether the athlete is taking the banned substance or not on the first set, and whether the athlete tests positive or negative on the second.

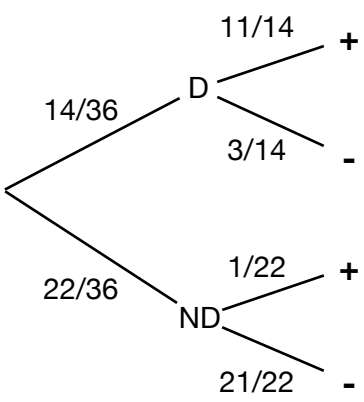
From expected proportions to probability calculations

The expected proportion of the 36 trials is equivalent to the probability that a particular outcome occurs. Starting from this as the answer enables students to see how the probability calculation (multiplying along the branches) mirrors the process by which they obtained the expected frequency of each outcome:

On average, 1/3 of the athletes will be doping because there are 2 chances in 6 of the die giving red. Hence, on average, 2/3 or 24 of them will not be doping.

Of the 12 athletes we expect to be doping, on average 3/36, or 1/12, will test negative, because there are 3 chances in 36 of two dice giving a total of 4. Hence, on average, 11 of the 12 we expect to be doping will test positive and 1 will test negative. 11/36 is equivalent to multiplying 1/3 by 11/12, and 1/36 is equivalent to multiplying 1/3 by 1/12.

Of the 24 athletes we expect not to be doping, on average 1/8, will test positive, because there is a 1 chance in 8 of getting three heads when a coin is flipped three times. Hence, on average, 3 of the 24 who are not doping will test positive, and the other 21 we expect to test negative. 3/36 is equivalent to multiplying 2/3 by 1/8, and 21/36 is equivalent to multiplying 2/3 by 7/8.



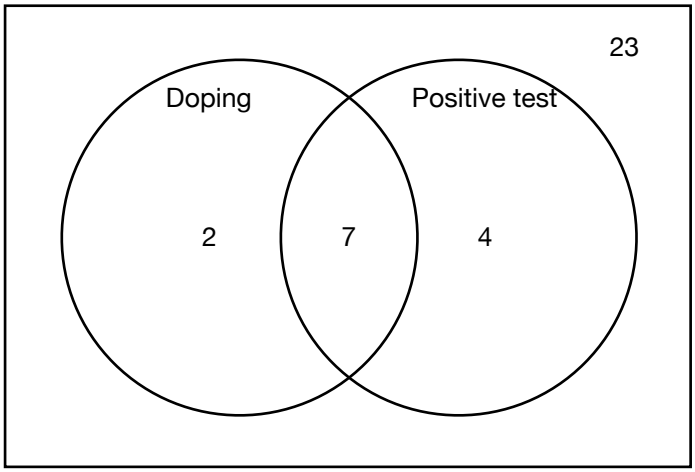
Outcome	Expected proportion	Probability
Positive doping	11/36	11/36 = 14/36 x 11/14
Negative doping	3/36 = 1/12	3/36 = 14/36 x 3/14
Positive not doping	1/36	1/36 = 22/36 x 1/22
Negative not doping	21/36 = 7/12	21/36 = 22/36 x 21/22
Totals	1	1

2-way Tables and Venn Diagrams

Sample experimental results

		Test?		
		Positive	Negative	Test totals
Doping?	Y	7	2	9
	N	4	23	27
	Test totals	11	25	36

Notice where figures from the tree are displayed on the 2-way table. Which are not on the tree? How do we find them there?



It can be tricky to label the circles correctly on a Venn diagram. They correspond to the two stages of the tree diagram - so the circles represent athletes who are doping and athletes who test positive.

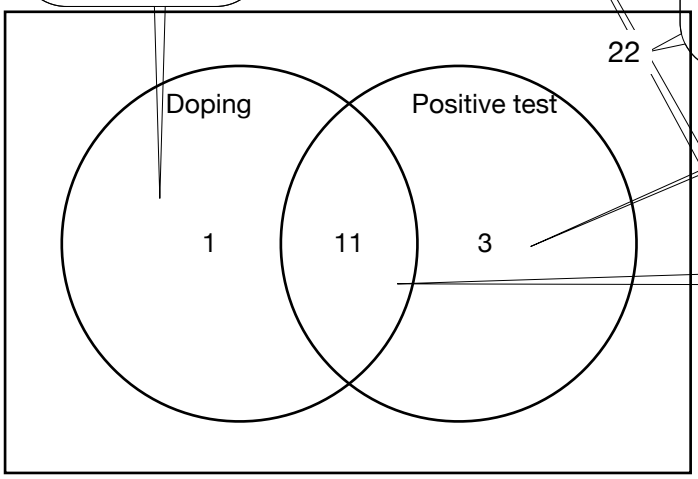
Expected results

		Test?		
		Positive	Negative	Test totals
Doping?	Y	11	1	12
	N	3	21	24
	Test totals	14	22	36

Doping, negative test

Not doping negative test

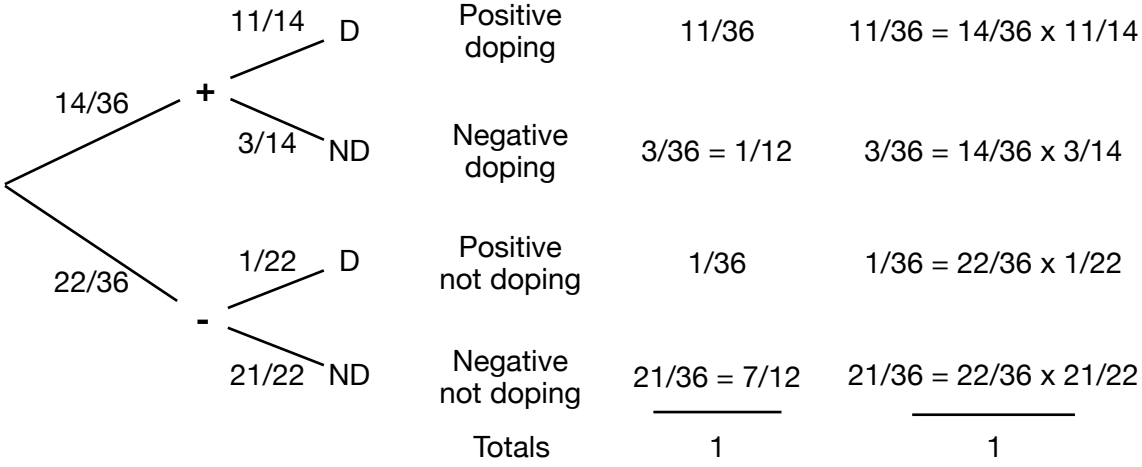
Not doping positive test



Doping positive test

Notice where figures are displayed on the 2-way table which are not on the Venn diagram. How do we find them here?

Reverse Probability Tree



The reverse probability tree enables us to answer questions like:

> What is the probability that an athlete who tests negative is in fact taking the banned substance? (Answer: 1/12)
>
> What is the probability that an athlete who tests positive is in fact not taking the banned substance? (Answer: 3/36 = 1/12)

The 2-way table mediates between the two tree diagrams, since all the information is on the table. The crucial step is to identify correctly the subset of the data which provides the denominator. The three questions below all relate to the same subgroup of athletes, but their number is expressed as a proportion of a different subset of the data:

> What proportion of the athletes are taking the banned substance but test negative?
> Answer: 1/36 = P(doping and negative
>
> What proportion of athletes who are doping do we expect to test negative?
> Answer: 1/12 = P(negative given that they are doping)
>
> What proportion of athletes who test negative are in fact taking the banned substance?
> Answer: 1/22 = P(doping given that they test negative)