A Worksheet about Probability Tree Diagrams

Math-270: Fall of 2016

September 30, 2016

Directions

Please do the problems in order. For each problem:

1. Read the problem.
2. Draw a probability tree diagram.
3. Double check that the leaf probabilities add to one.
4. Check your tree against mine. (The trees are provided between the questions and the answers.)
5. Try to answer all the questions marked (a), (b), etc...
6. Check all those answers in the back of the worksheet.
7. Congratulate self.
8. Continue by reading the next problem.

Questions

Note: Problems 1–7 have the most classic shape of a probability tree, with three levels and four leaves. Problems 8–13 introduce other tree shapes.

1. This problem is meant to capture the situation about HIV, as it was when I was in college during the 1990s. Suppose that an HIV-test is 98% accurate—that means it gives a true answer 98% of the time, and a false answer 2% of the time. Further suppose that among a certain college-aged population, 5% were HIV-positive and 95% were HIV-negative.

   (a) What is the probability that a random person (from this particular population) tests HIV positive?

   (b) What is the probability that a random person (from this particular population) tests HIV negative?

   (c) What is the probability that a random person (from this particular population) who is holding a positive result in their hands is actually HIV negative?

   (d) What is the probability that a random person (from this particular population) who is holding a positive result in their hands is actually HIV positive?

   (e) What is the probability that a random person (from this particular population) who is holding a negative result in their hands is actually HIV negative?

   (f) What is the probability that a random person (from this particular population) who is holding a negative result in their hands is actually HIV positive?

2. Now we're going to re-examine the previous question, but with a different batch of college students. Among this second population, 1% are HIV positive and 99% are HIV negative.

3. **Citation:** The following problem is a paraphrase of one found in *Finite Mathematics & Its Applications*, 10th edition, by Larry Goldstein, David Schneider, and Martha Siegel, where it was Example 6-6-2.

   Let's talk about Tuberculosis now. This disease has thankfully become rare. For the purposes of this problem, we will say that 0.25% of the population has Tuberculosis. On a patient with Tuberculosis, a positive result is returned with probability 99%, and a negative result with probability 1%. On a patient without Tuberculosis, a negative result is returned with probability 98%, and a positive result with probability 2%.

Warning: Do not round off your numbers while doing this problem. Use every digit that you have.

   (a) What is the probability that a random person (from this particular population) tests positive for tuberculosis?
   (b) What is the probability that a random person (from this particular population) tests negative for tuberculosis?
   (c) What is the probability that a random person (from this particular population) who is holding a positive result in their hands is actually not infected with tuberculosis?
   (d) What is the probability that a random person (from this particular population) who is holding a negative result in their hands is indeed not infected with tuberculosis?
   (e) What is the probability that a random person (from this particular population) who is holding a positive result in their hands is indeed infected with tuberculosis?
   (f) What is the probability that a random person (from this particular population) who is holding a negative result in their hands is actually infected with tuberculosis?

Note: I have to admit that I've made a terrible error in the previous problem. The original problem had the basic rate of tuberculosis at 2 in 10,000, which was true long ago when that book was first written in the very late 1970s. Today, the basic rate is 3 in 100,000. It is nice to know that something is improving. For some reason, I thought this went the other way, and I wrote 1 in 400 into the problem. Oops. I realized this error too late, and there is no time now to fix it. Darn.

4. For this problem, use fractions, and not decimals. A factory has two machines for making a particular part. Because the newer machine is twice as fast as the older machine, the older machine manufactures 1/3 of the stockpile of this particular part, and the newer manufactures makes 2/3 of the stockpile. The older machine has a defect rate of 1/20, and the newer machine has a defect rate of 1/30. The defective parts are detected in a testing phase prior to being installed further down the assembly line. Overall, what is the probability that one of these particular parts happens to be defective?

   The boss orders that the defective parts be analyzed, to determine which machine they came from. What fraction of the defective parts were from the older machine? What fraction were from the newer machine?

5. This problem was recommended by Prof. Seth Dutter. Outside of certain specialized agencies of the US Department of Defense, polygraphs (lie-detector tests) are about 90% effective. When I worked for the NSA, I was polygraphed several times, as are all employees, by extremely focused professional polygraphers. The effective rate is surely better, but almost nobody knows what their true effectiveness rate actually is. Let's suppose that at a particular firm, 1% of the employees are actively embezzling or stealing from the firm. The CEO decides that every single employee be given a polygraph examination, to determine if they are stealing or not.

   There are actually laws in most states of the USA that make such a decision completely illegal. You are about to find out why.

(a) What percentage of the firm's employees will fail the polygraph examination?
(b) What is the probability that an employee who failed the polygraph exam is innocent?
(c) What is the probability that an employee who failed the polygraph exam is guilty?

6. This problem is about a fast-food business. As it turns out, 60% of their business is via the drive-through window, and 40% is counter-service. Data from the registers reveals that 70% of the transactions at the drive-through are via credit/debit cards, while 80% of the transactions at the counter are via credit/debit cards. (The other transactions are via cash.)

   - What is the probability that a random transaction uses cash?
   - What is the probability that a random transaction uses credit cards?
   - What is the probability that a random credit-card transaction was at the drive-through window?
   - What is the probability that a random cash transaction was at the drive-through window?
   - What is the probability that a random credit-card transaction was at the counter?
   - What is the probability that a random cash transaction was at the counter?

Background: The computation in the following problem came up in an email from November 9th, 2004, with Richard Huzzey, then a doctoral student of Oxford University, now a Senior Lecturer of the University of Liverpool. He was president of the Oxford Diplomacy Society when I was a visiting student at Oxford, during roughly April to July of 2004. We were discussing the US presidential election of 2004.

The exit polls had indicated that Kerry had won the presidency with some margin, and the actual votes indicated that George W. Bush had won the presidency. (This is not the disputed election, with the confusion in Florida—that was in the year 2000.)

Many people around the world were shocked that the exit polls had indicated one outcome, and the official tabulation indicated the opposite outcome. How could this be? One hypothesis that was circulated had to do with electronic voting machines, which were (at that time) relatively new. Some were worried that Kerry had legitimately won, but that the machines had been hacked to indicate that Bush had won.

However, the discrepancy has a much more mundane explanation, which we will explore in the next problem.

7. Suppose that in a particular state (in the 2004 US presidential election), Bush received 52% of the vote, and Kerry received 48%. Further suppose that while 80% of the Bush voters were willing to speak to reporters conducting an exit poll, 95% of the Kerry voters were willing to speak to a reporter conducting an exit poll. (In that election, the Kerry crowd was enthusiastic about possibly evicting Bush from The White House.)

   Keeping in mind that the exit poll can only include people willing to speak to those reporters, what would the exit poll indicate?

   - What is the probability that a random voter ignores the reporters conducting the exit polls?
   - What is the probability that a random voter speaks to the reporters conducting the exit poll?
   - What is the probability that a voter who speaks to the reporters is a Kerry voter?
   - What is the probability that a voter who speaks to the reporters is a Bush voter?
   - What is the probability that a voter who ignores the reporters is a Kerry voter?
   - What is the probability that a voter who ignores the reporters is a Bush voter?
   - What does the exit poll indicate? Note, reporters often use three significant digits.

Note: In the following problems, we're now shifting to probability tree diagrams that are not in the classic shape.

8. Here is some *fictional* data about a university with three constituent colleges:

Question	Science, Technology, Engineering & Math	Social/Behavioral Sciences Sciences	Arts & Humanities
What percentage of the student body is this college?	20%	45%	35%
What percentage of the students in this college use marijuana?	15%	40%	65%

   (a) What percentage of the students at this college are marijuana users?
   (b) What percentage of the students at this college are marijuana abstainers?
   (c) If you pick a random marijuana-using student, what is the probability that they are majoring in the Arts & Humanities?
   (d) If you pick a random marijuana-using student, what is the probability that they are majoring in the Social/Behavioral Sciences?
   (e) If you pick a random marijuana-using student, what is the probability that they are majoring in S.T.E.M.?
   (f) If you pick a random marijuana-abstaining student, what is the probability that they are majoring in the Arts & Humanities?
   (g) If you pick a random marijuana-abstaining student, what is the probability that they are majoring in the Social/Behavioral Sciences?
   (h) If you pick a random marijuana-abstaining student, what is the probability that they are majoring in S.T.E.M.?

9. Let's suppose that a smartphone factory gets a particular chip from three different manufacturers. Company A provides 45%, Company B provides 35%, and Company C provides 20%. The defect rate turns out to be 1%, 2%, and 1% for those three companies, respectively.

   - What percentage of the chips, overall, are defective?
   - If I select a random defective chip, what is the probability that it is from Company A? from Company B? from Company C?
   - If I select a random non-defective chip, what is the probability that it is from Company A? from Company B? from Company C?

10. Let's imagine that a repair facility (for example, for a major car or computer manufacturer) frequently gets a specific circuitboard sent to it. For this circuitboard, they have three diagnostic procedures. Procedure 1 will locate the fault with probability 70%, and is done first. Procedure 2 will locate the fault with probability 60%, and is done if Procedure 1 does not locate the fault. Procedure 3 will locate the fault with probability 50%, and is done if both Procedures 1 and 2 fail to locate the fault. If all three procedures fail, they throw the circuitboard out and send the customer a brand new board.

   (a) What is the probability that the fault is not located at all?
   (b) If the fault is located, what is the probability that it was found by Procedure 1?
   (c) If the fault is located, what is the probability that it was found by Procedure 2?
   (d) If the fault is located, what is the probability that it was found by Procedure 3?
   (e) For a fault that is not located by Procedure 1, what is the probability that it is found by Procedure 2?
   (f) For a fault that is not located by Procedure 1, what is the probability that it is found by Procedure 3?

(g) For a fault that is not located by Procedure 1, what is the probability that it ends up never being found?

11. Let's consider an engineering program, where certain courses are causing students to fail out. To keep the problem relatively small and tractable, we will make the simplifying assumption that failed courses cannot be repeated. (In other words, any failure is terminal for a student in the program.) The students take Calculus I, Calculus II, Physics I, Calculus III, and Physics II, one at a time, and in that order. The pass rates are 90%, 70%, 80%, 80%, and 70%, respectively. Make a tree to model a students path through these critical courses.

    (a) What is the probability that a student who starts the program will successfully finish it?

    Note: Calculations of the following type are actually used at universities at diagnosing curriculum issues. It is called "cause of death" analysis.

    (b) Given that a student flunks out, what is the probability that they failed Calculus I?

    (c) Given that a student flunks out, what is the probability that they failed Calculus II?

    (d) Given that a student flunks out, what is the probability that they failed Physics I?

    (e) Given that a student flunks out, what is the probability that they failed Calculus III?

    (f) Given that a student flunks out, what is the probability that they failed Physics II?

    (g) If we call the course that a student failed their "cause of death," what is the leading "cause of death" among these five courses?

12. **Citation:** The following problem appeared in *Finite Mathematics & Its Applications*,10th edition, by Larry Goldstein, David Schneider, and Martha Siegel, where it was Exercises 6-6-4, 6-6-6, 6-6-7, and 6-6-8.

    A training program is used by a corporation to direct hirees to appropriate jobs. The program consists of two steps. Step I identifies 30% as management trainees, 60% as non-managerial workers, and 10% to be fired. In Step II, 75% of the management trainees are assigned to managerial positions, 20% are assigned to non-managerial positions, and 5% are fired. Also, In Step II, 60% of the non-managerial workers are kept in the same category, 10% are assigned to management positions, and 30% are fired.

    (a) What is the probability that a randomly chosen hiree will be assigned to management position at the end of the trainee period?

    (b) What is the probability that a randomly chosen hiree will be fired at the end of the trainee period?

    (c) What is the probability that a randomly chosen hiree will be designated a management trainee (in Step I), but not be appointed to a management position (after Step II)?

    Note: The remaining questions are my own.

    (d) After all is done, if you select someone appointed to a managerial position, what is the probability that they were assigned to the management track during Step II?

    (e) After all is done, if you select someone appointed to a non-managerial position, what is the probability that they were assigned to the non-management track during Step II?

    (f) After all is done, if you select someone who has been fired, what is the probability that they were fired during Step II?

13. The grand-slam closing problem of this worksheet is to "confirm" the Bernoulli Distribution Formula, by looking at a case that's kind of large, but that still can be done with a probability tree diagram. We are going to consider a website with four servers globally. Since they are far apart, the probability of one being up or down is independent of the probability of another being up or down. To keep the numbers simple, we will say that there is a 9/10 probability of the server being up, and a 1/10

probability of the server being down. Note, that even with horribly old and poorly maintained servers, you'd still expect much more than a 90% uptime—this scandalous approximation is done to make the arithmetic clean and simple.

We have to ask each server, one at a time, if it is up or down. To give you a hint, we will have a 5-layer tree, with 1 root branch, then 2 branches, 4 branches, 8 branches, and 16 branches, ending with 16 leaves. (That's a total of 31 branches!) Maybe it is too much work for you to draw this tree yourself. Instead, look at my tree, provided with the others. Then compute the following probabilities.

(a) 4 up and 0 down

(b) 3 up and 1 down

(c) 2 up and 2 down

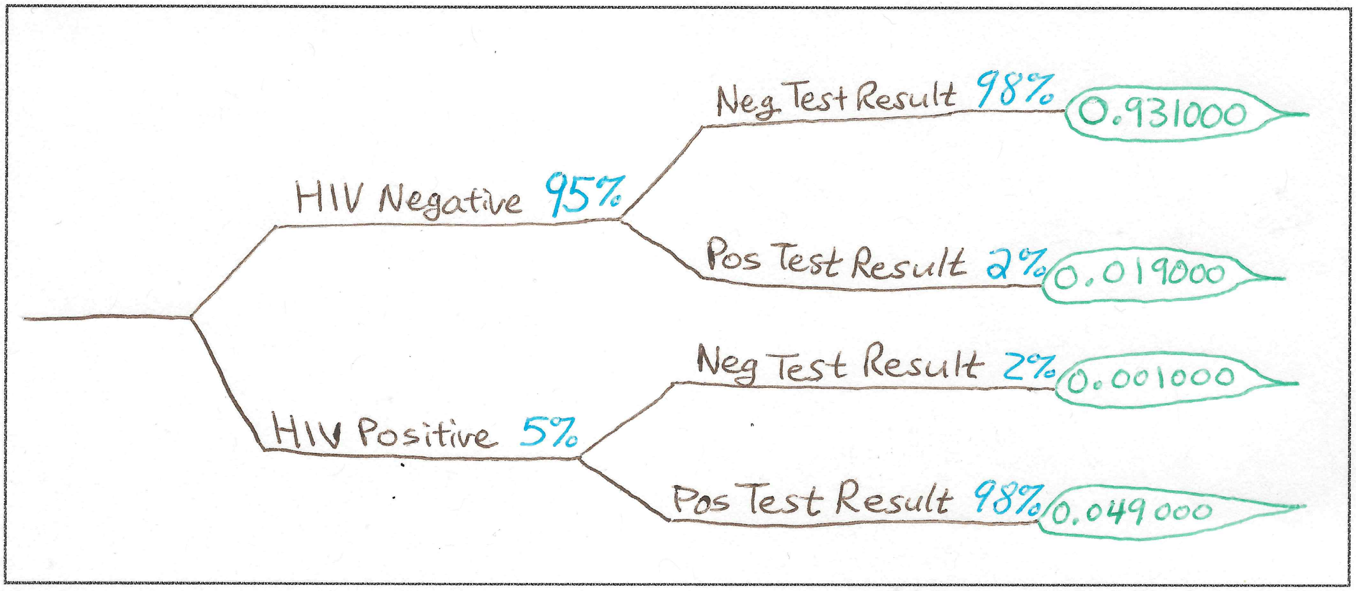
(d) 1 up and 3 down

(e) 0 up and 4 down

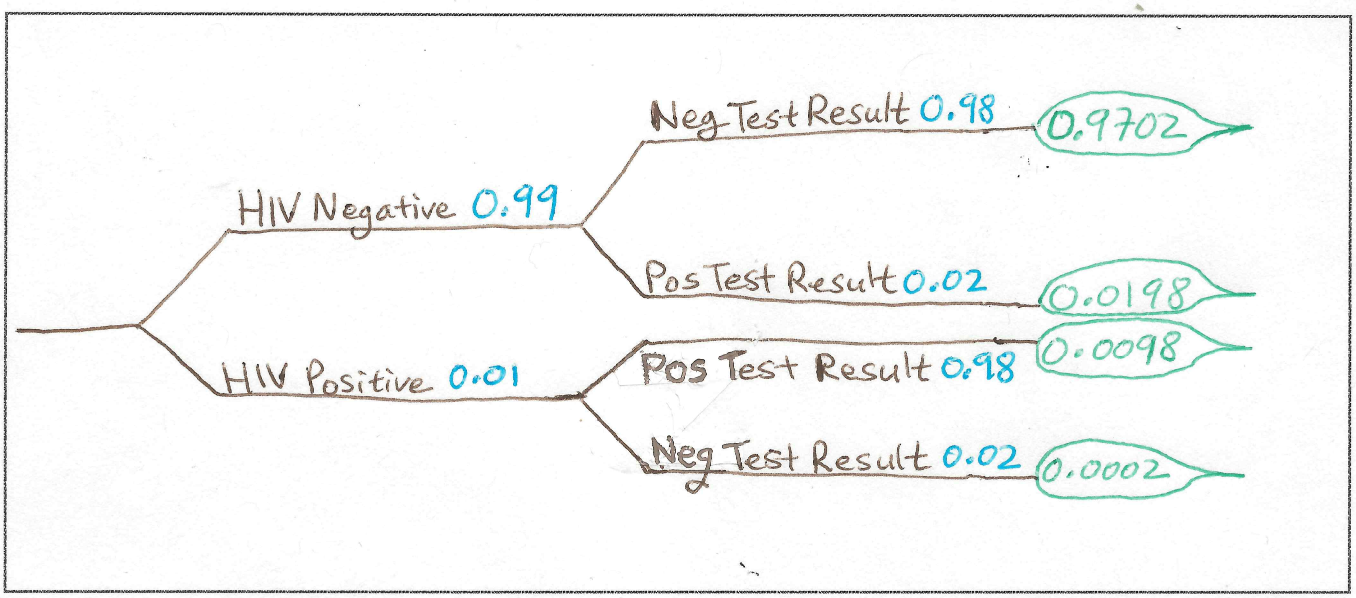
Note: It is extremely worthwhile to also compute these five probabilities with Bernoulli's Distribution Formula, so that you can see with your own eyes that indeed, all the numbers match.

The Trees!

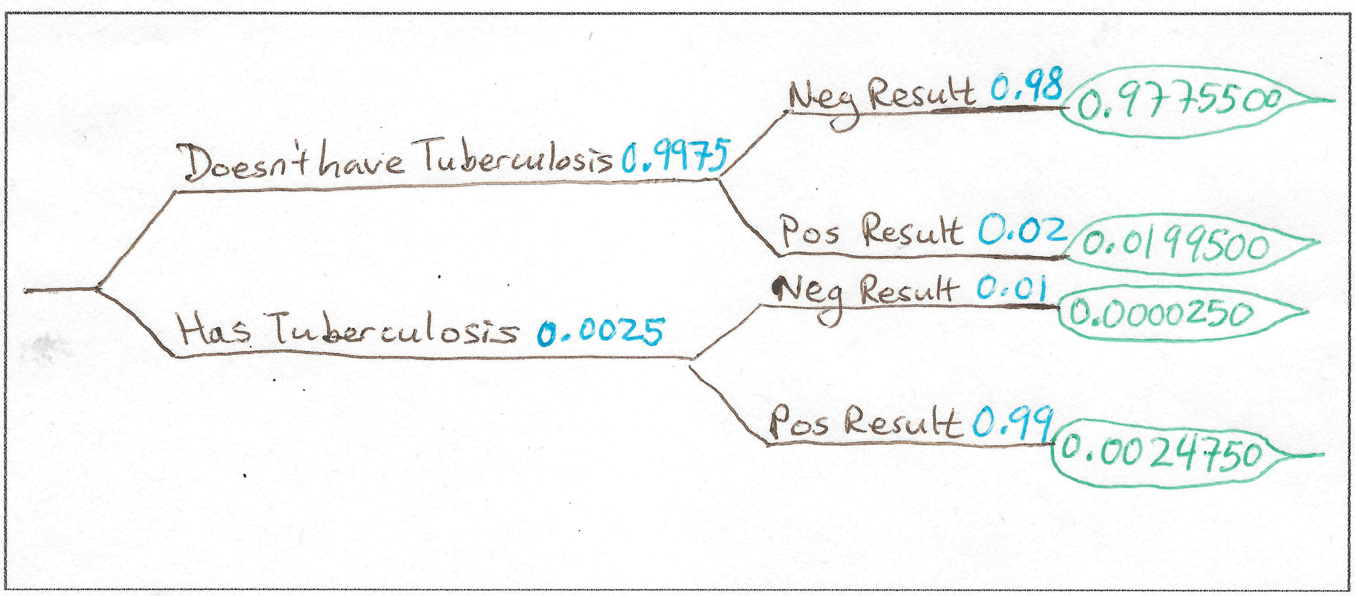
The probability tree diagrams will be given on the next few pages. After that, the answers will follow.



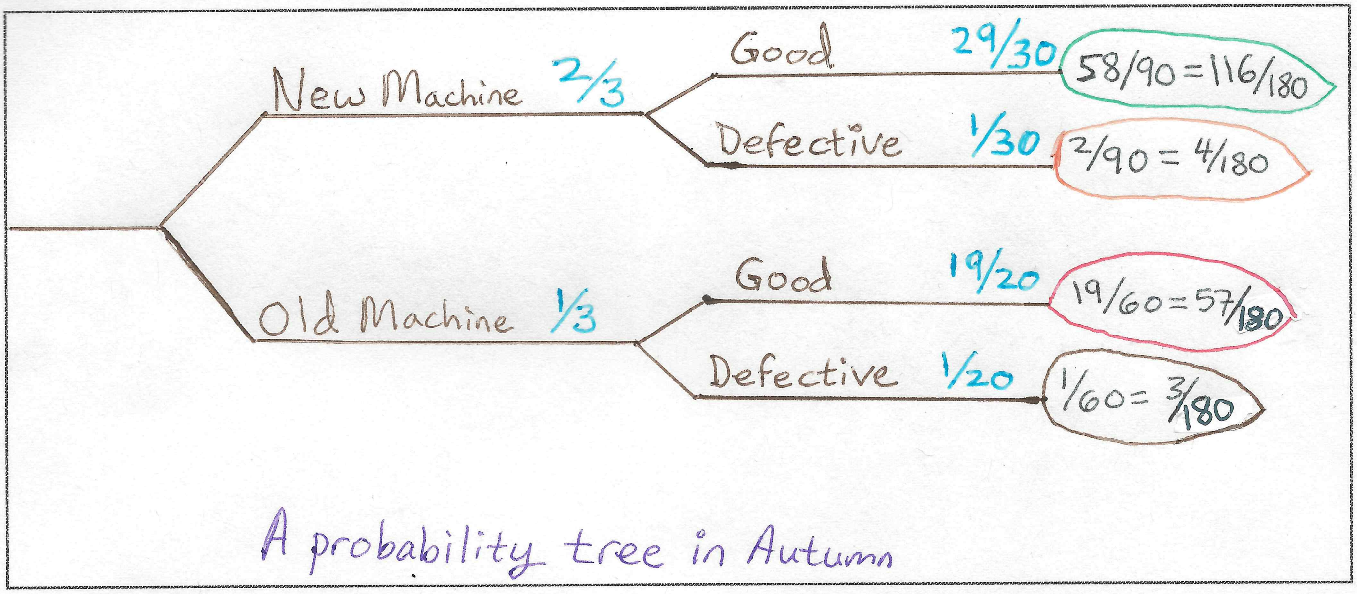
HIV Negative 95%
Neg Test Result 98%
0.931000
Pos Test Result 2%
0.019000
HIV Positive 5%
Neg Test Result 2%
0.001000
Pos Test Result 98%
0.049000



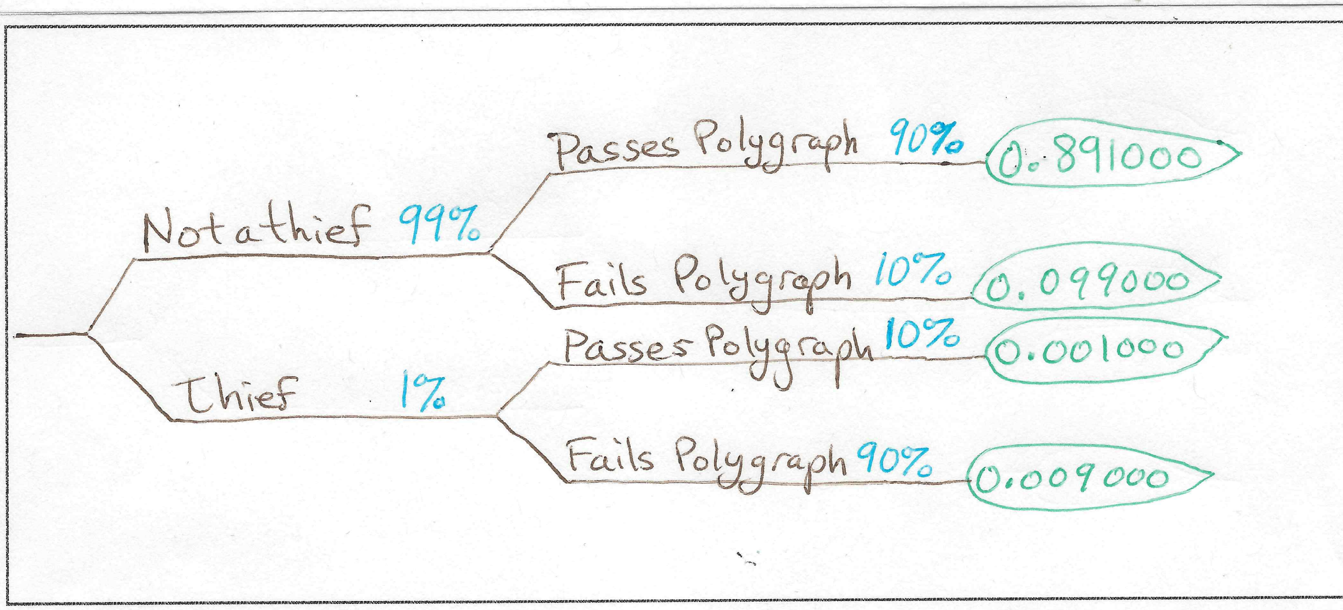
HIV Negative 0.99
Neg Test Result 0.98
0.9702
Pos Test Result 0.02
0.0198
HIV Positive 0.01
Pos Test Result 0.98
0.0098
Neg Test Result 0.02
0.0002



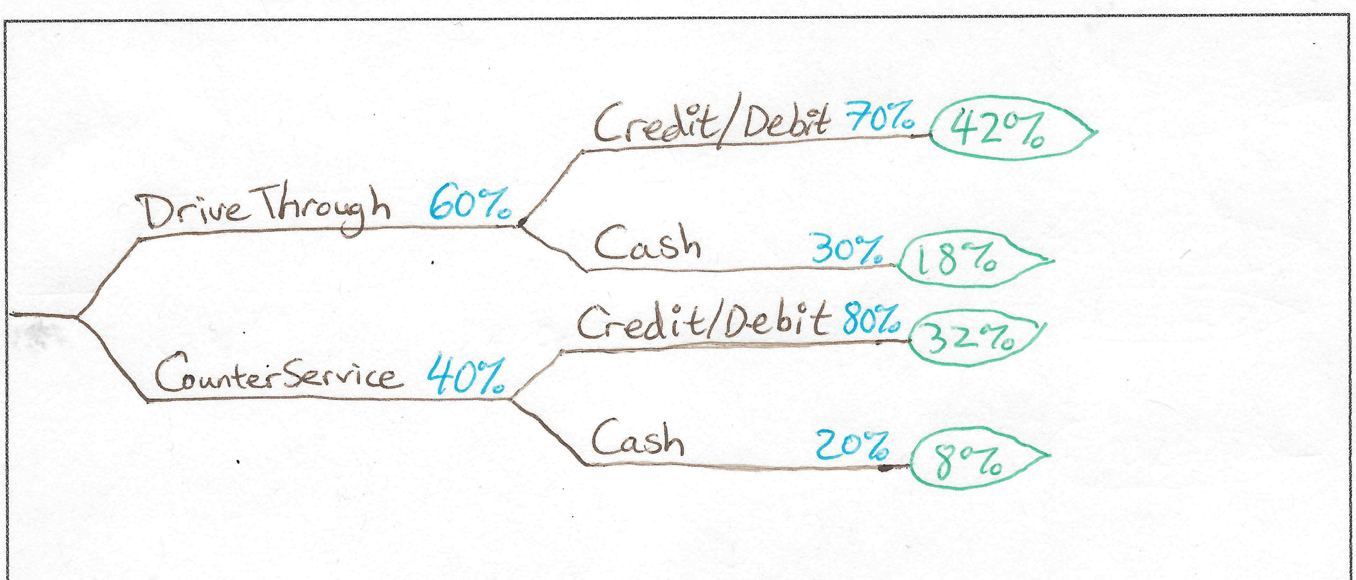
Doesn't have Tuberculosis 0.9975
Neg Result 0.98
0.977550
Pos Result 0.02
0.0199500
Has Tuberculosis 0.0025
Neg Result 0.01
0.0000250
Pos Result 0.99
0.0024750



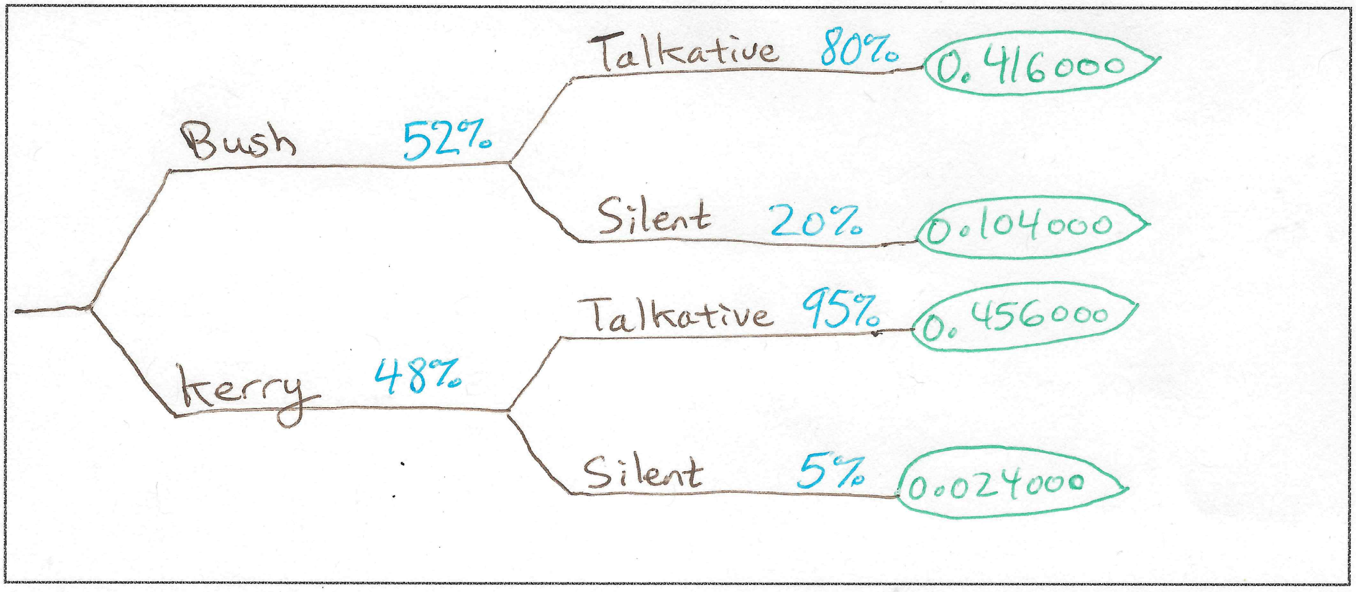
New Machine 2/3
Good 29/30
58/90 = 116/180
Defective 1/30
2/90 = 4/180
Old Machine 1/3
Good 19/20
19/60 = 57/180
Defective 1/20
1/60 = 3/180
A probability tree in Autumn

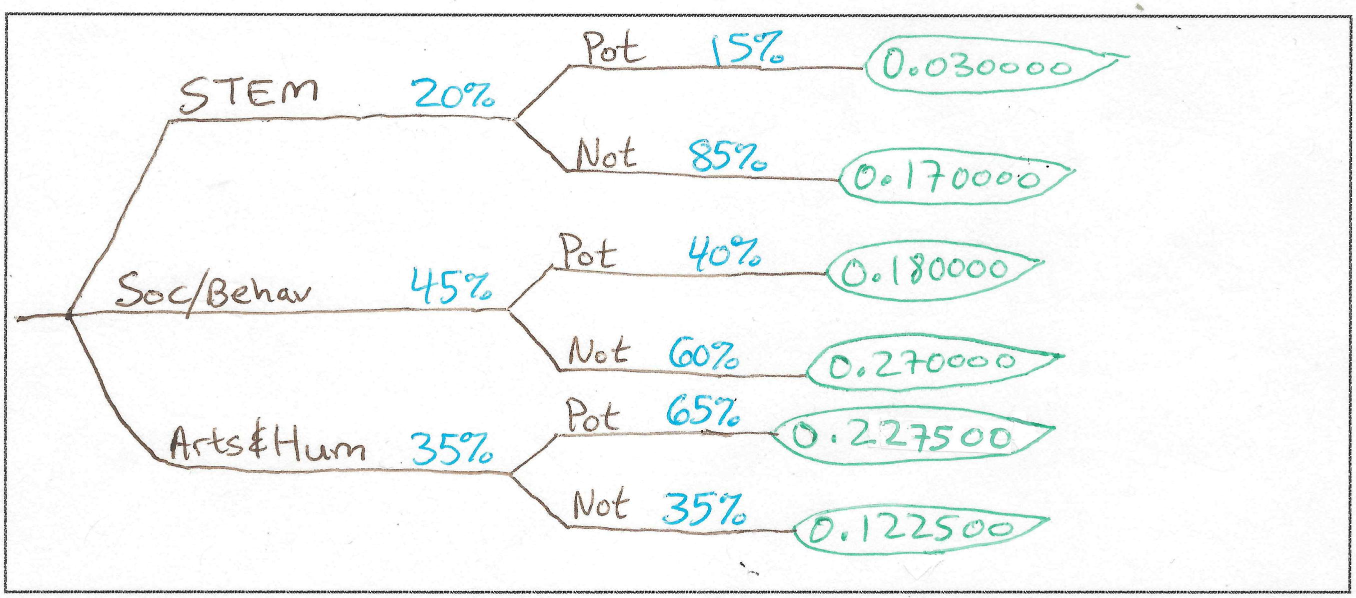


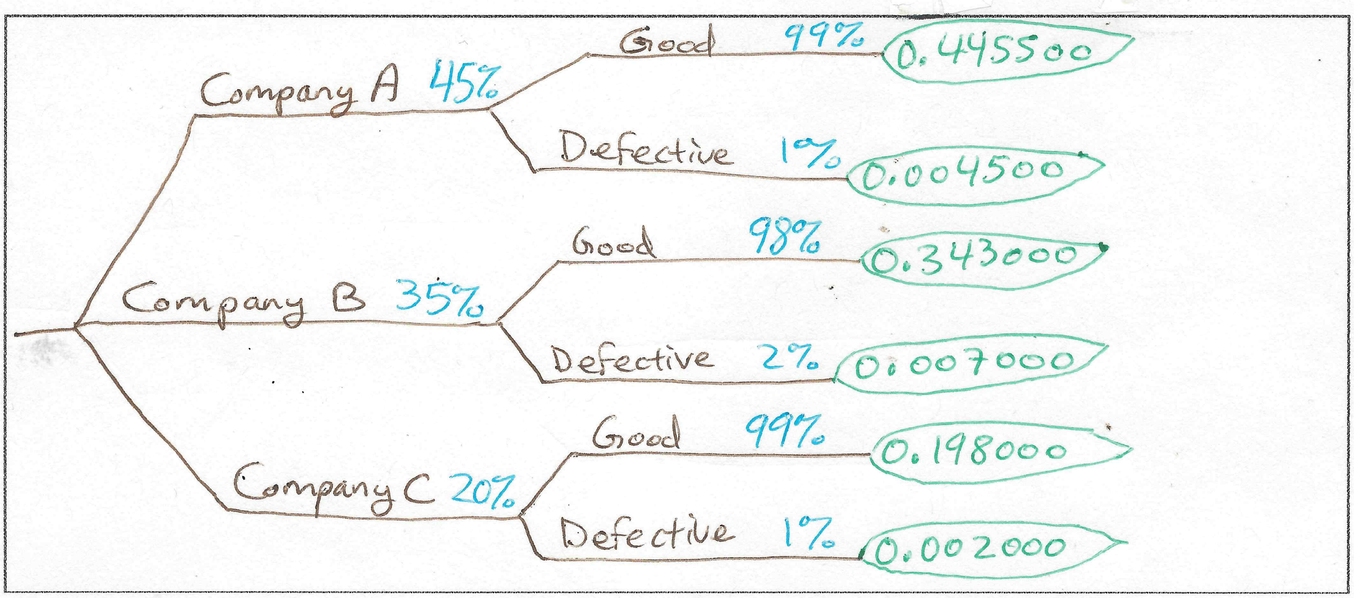
Not a thief 99%
Passes Polygraph 90%
0.891000
Fails Polygraph 10%
0.099000
Thief 1%
Passes Polygraph 10%
0.001000
Fails Polygraph 90%
0.009000

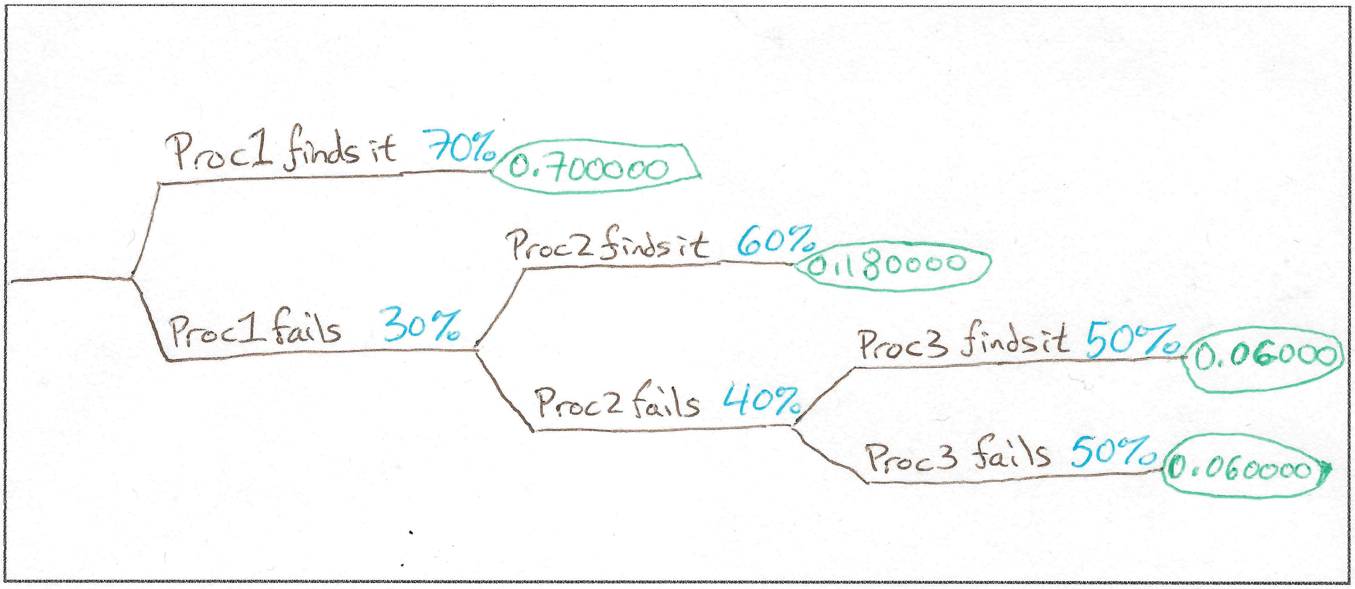


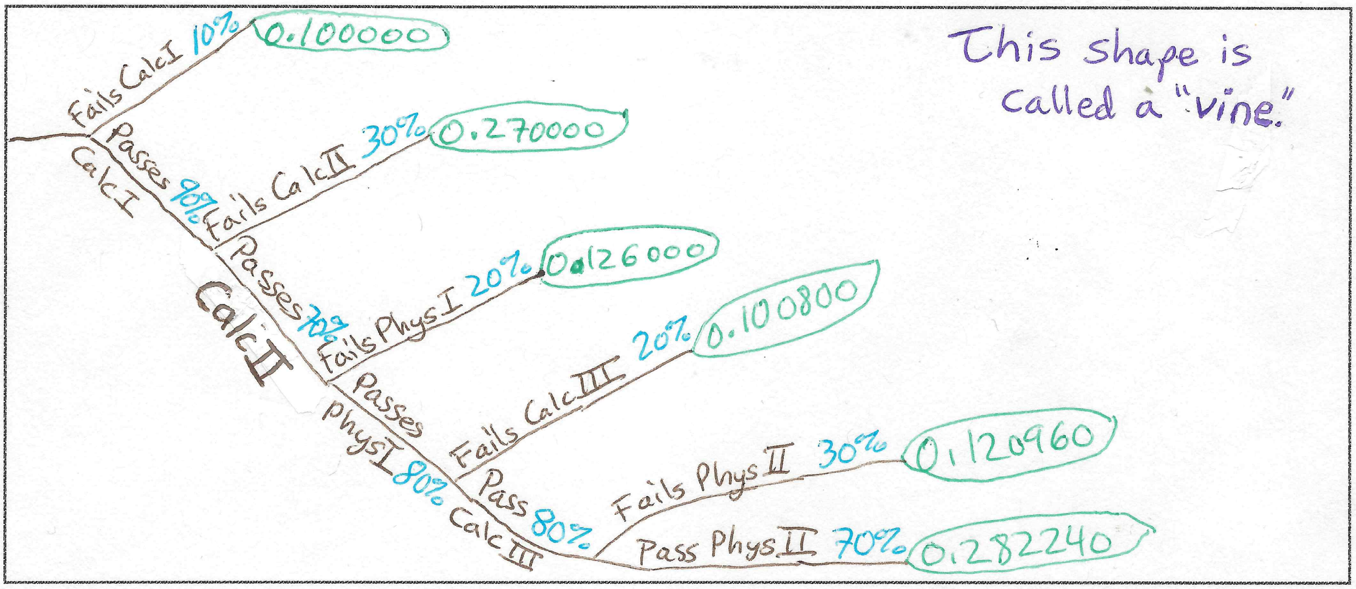
Drive Through 60%
Credit/Debit 70%
42%
Cash 30%
18%
Counter Service 40%
Credit/Debit 80%
32%
Cash 20%
8%

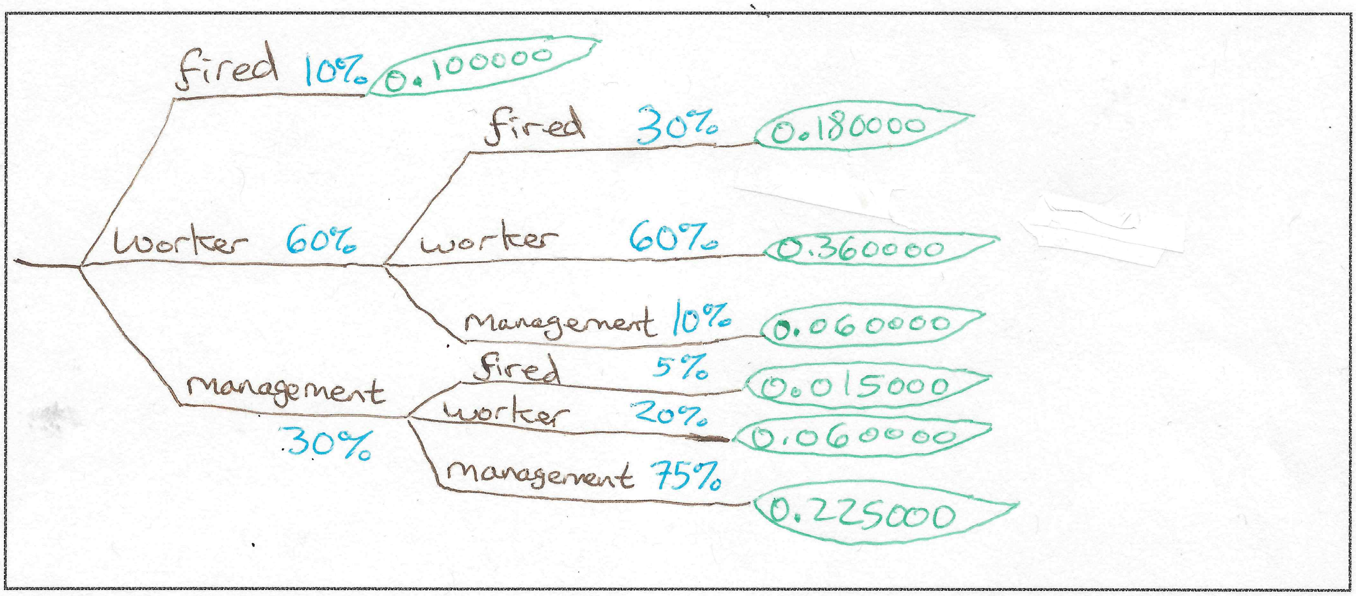

Bush 52%
Talkative 80%
0.416000
Silent 20%
0.104000
Kerry 48%
Talkative 95%
0.456000
Silent 5%
0.024000

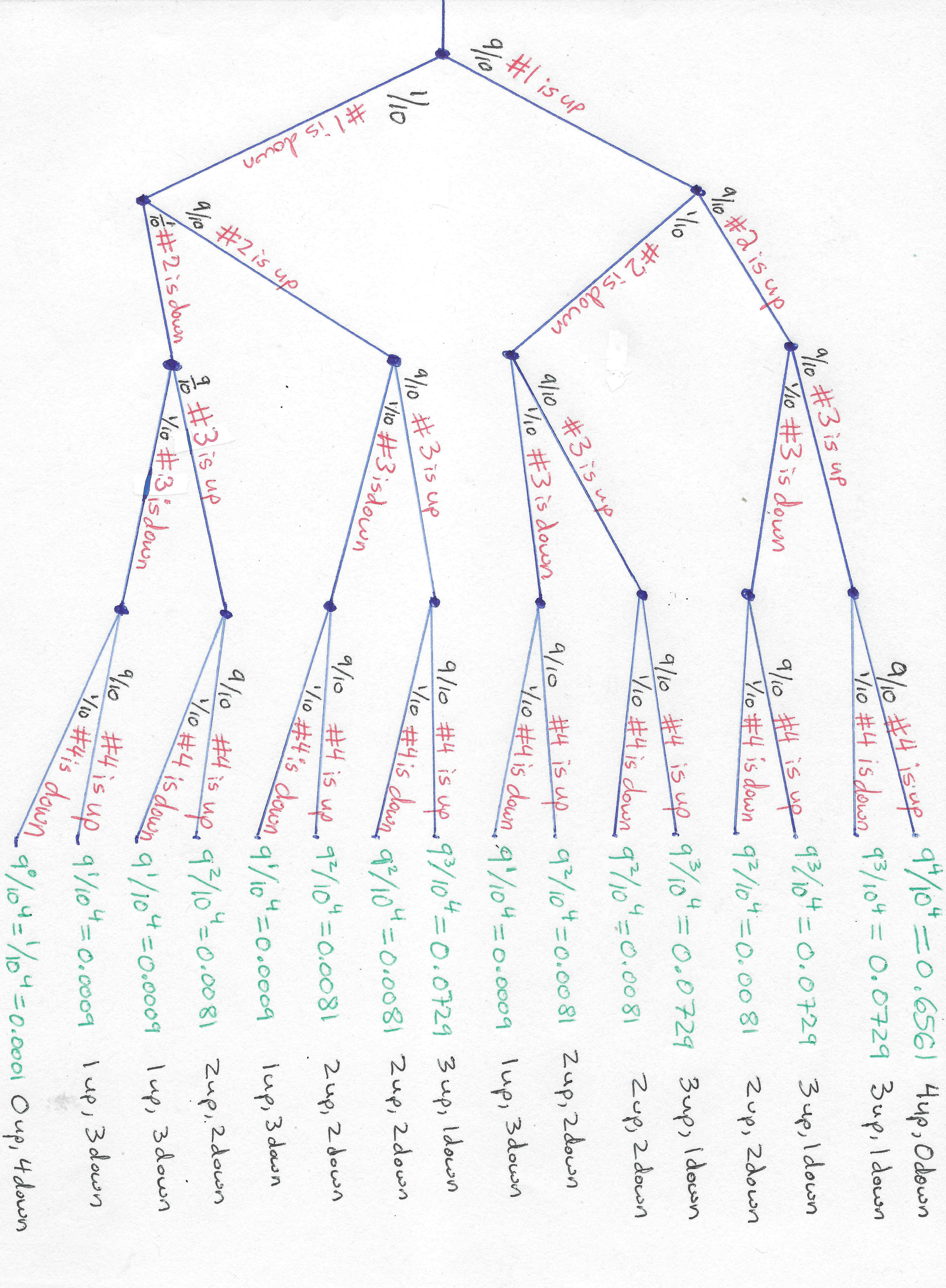

STEM 20%
Pot 15%
0.030000
Not 85%
0.170000
Soc/Behav 45%
Pot 40%
0.180000
Not 60%
0.270000
Arts&Hum 35%
Pot 65%
0.227500
Not 35%
0.122500


Company A 45%
Good 99%
0.445500
Defective 1%
0.004500
Company B 35%
Good 98%
0.343000
Defective 2%
0.007000
Company C 20%
Good 99%
0.198000
Defective 1%
0.002000


Proc1 finds it 70%
0.700000
Proc1 fails 30%
Proc2 finds it 60%
0.180000
Proc2 fails 40%
Proc3 finds it 50%
0.06000
Proc3 fails 50%
0.060000


Fails Calc I 10%
0.100000
Passes Calc I 90%
Fails Calc II 30%
0.270000
Passes Calc II 70%
Fails Phys I 20%
0.126000
Passes Phys I 80%
Fails Calc III 20%
0.100800
Pass Calc III 80%
Fails Phys II 30%
0.120960
Pass Phys II 70%
0.282240
This shape is called a "vine."


fired 10%
0.100000
worker 60%
fired 30%
0.180000
worker 60%
0.360000
management 10%
0.060000
management 30%
fired 5%
0.015000
worker 20%
0.060000
management 75%
0.225000



9/10
#1 is up
1/10
#1 is down
9/10
#2 is up
1/10
#2 is down
9/10
#3 is up
1/10
#3 is down
9/10
#4 is up
1/10
#4 is down
9⁴/10⁴ = 0.6561 4up, 0down
9³/10⁴ = 0.0729 3up, 1down
9³/10⁴ = 0.0729 3up, 1down
9²/10⁴ = 0.0081 2up, 2down
9³/10⁴ = 0.0729 3up, 1down
9²/10⁴ = 0.0081 2up, 2down
9²/10⁴ = 0.0081 2up, 2down
9¹/10⁴ = 0.0009 1up, 3down
9³/10⁴ = 0.0729 3up, 1down
9²/10⁴ = 0.0081 2up, 2down
9²/10⁴ = 0.0081 2up, 2down
9¹/10⁴ = 0.0009 1up, 3down
9²/10⁴ = 0.0081 2up, 2down
9¹/10⁴ = 0.0009 1up, 3down
9¹/10⁴ = 0.0009 1up, 3down
9⁰/10⁴ = 1/10⁴ = 0.0001 0up, 4down

Answers

1. Here are the answers to the first question about HIV.

   (a) What is the probability that a random person (from this particular population) tests HIV positive?

   $$[\text{Answer: } 0.049000 + 0.019000 = 6.80000\%.]$$

   (b) What is the probability that a random person (from this particular population) tests HIV negative?

   $$[\text{Answer: } 0.931000 + 0.001000 = 93.2000\%.]$$

   (c) What is the probability that a random person (from this particular population) who is holding a positive result in their hands is actually HIV negative?

   $$[\text{Answer: } \frac{0.019000}{0.049000 + 0.019000} = 27.9411\%.]$$

   (d) What is the probability that a random person (from this particular population) who is holding a positive result in their hands is actually HIV positive?

   $$[\text{Answer: } \frac{0.049000}{0.049000 + 0.019000} = 72.0588\%.]$$

   (e) What is the probability that a random person (from this particular population) who is holding a negative result in their hands is actually HIV negative?

   $$[\text{Answer: } \frac{0.931000}{0.931000 + 0.001000} = 99.8927\%.]$$

   (f) What is the probability that a random person (from this particular population) who is holding a negative result in their hands is actually HIV positive?

   $$[\text{Answer: } \frac{0.001000}{0.931000 + 0.001000} = 0.107296\%.]$$

2. Here are the answers to the second question about HIV.

   (a) What is the probability that a random person (from this particular population) tests HIV positive? [Answer: 2.96000%.]

   (b) What is the probability that a random person (from this particular population) tests HIV negative? [Answer: 97.0400%.]

   (c) What is the probability that a random person (from this particular population) who is holding a positive result in their hands is actually HIV negative? [Answer: 66.8918%.]

   (d) What is the probability that a random person (from this particular population) who is holding a positive result in their hands is actually HIV positive? [Answer: 33.1081%.]

   (e) What is the probability that a random person (from this particular population) who is holding a negative result in their hands is actually HIV negative? [Answer: 99.9793%.]

   (f) What is the probability that a random person (from this particular population) who is holding a negative result in their hands is actually HIV positive? [Answer: 0.0206100%.]

   Note: Isn't this shocking? If you are holding a positive result in your hands, there is a 2/3rds chance that you're clean, and only a 1/3rd chance that you've got HIV.

3. Here are the answers to the question about tuberculosis.

(a) What is the probability that a random person (from this particular population) tests positive for tuberculosis? [Answer: 2.24250%.]

(b) What is the probability that a random person (from this particular population) tests negative for tuberculosis? [Answer: 97.7575%.]

(c) What is the probability that a random person (from this particular population) who is holding a positive result in their hands is actually not infected with tuberculosis? [Answer: 88.9632%.]

(d) What is the probability that a random person (from this particular population) who is holding a negative result in their hands is indeed not infected with tuberculosis? [Answer: 99.9974%.]

(e) What is the probability that a random person (from this particular population) who is holding a positive result in their hands is indeed infected with tuberculosis? [Answer: 11.0367%.]

(f) What is the probability that a random person (from this particular population) who is holding a negative result in their hands is actually infected with tuberculosis? [Answer: 0.00255734%.]

Note: This is spectacularly shocking. However, that's what happens with rare diseases. The fact that only 1/400 people have tuberculosis, and 399/400 people do not have tuberculosis, completely overshadows the accuracy/inaccuracy of the test. In practice, positive test results are simply treated as follows: "I'm sorry, we have to run the test a second time." Only if both come back positive, do they inform the patient. Sometimes, they even run a third test, and then inform the patient.

Recall: The actual rate of tuberculosis today is 3 in 100,000. So this problem would be even more shocking with the more modern numbers.

4. Here are the answers to the question about the older machine and the newer machine making defective parts.

(a) What is the probability that one of these particular parts happens to be defective?
[Answer: 7/180]

(b) How many of the defective parts were from the older machine? [Answer: 3/7.]

(c) How many were from the newer machine? [Answer: 4/7.]

Note: Isn't this very misleading? The boss would imagine that the newer machine is worse than the older machine, since a majority of the defective parts are from the newer machine. However, we know that the older machine has a higher defect rate than the newer machine.

5. Here are the answers to the question about polygraphs, as suggested by Prof. Dutter.

(a) What percentage of the firm's employees will fail the polygraph examination? [Answer: 10.8000%.]

(b) What is the probability that an employee who failed the polygraph exam is innocent?
[Answer: 91.6666%.]

(c) What is the probability that an employee who failed the polygraph exam is guilty?
[Answer: 8.33333%.]

Note: To put this in perspective, let's suppose the firm has 5000 employees. We expect that 540 employees will fail the polygraph examination. Of those 540, there will be 495 who are completely innocent, and 45 who are actually thieves. Moreover, out of the 50 employees who are stealing, 5 of them will pass the polygraph examination.

Note: Another way to look at it is from a legal point of view. Even if only 20% of the innocent-but-investigated employees sue, then that's 99 lawsuits. Because lawyers are very expensive, this is a serious problem.

Note: In defense of polygraphers, it is much easier to conduct an investigation of 540 employees than of 5000 employees. So long as the failure of the polygraph examination results only in a further investigation or audit, and not in firing someone, then this might actually be a good way of catching people. For example, the FBI did use this technique, at times, to find out who is collaborating with the mafia. However, managers are managers, and it is hard to imagine how the 495 innocent people who get investigated will ever be trusted again at that firm—even though they've done nothing wrong. That's why this practice is illegal in almost all US states.

6. Here are the answers to the problem about credit cards and cash transactions at a fast-food restaurant.
    - What is the probability that a random transaction uses cash? [Answer: 26.0000%.]
    - What is the probability that a random transaction uses credit cards? [Answer: 74.0000%.]
    - What is the probability that a random credit-card transaction was at the drive-through window? [Answer: 56.7567%.]
    - What is the probability that a random cash transaction was at the drive-through window? [Answer: 69.2307%.]
    - What is the probability that a random credit-card transaction was at the counter? [Answer: 43.2432%.]
    - What is the probability that a random cash transaction was at the counter? [Answer: 30.7692%.]
7. Here are the answers to the question about exit polls during the Bush/Kerry election.
    - What is the probability that a random voter ignores the reporters conducting the exit polls? [Answer: 12.8000%.]
    - What is the probability that a random voter speaks to the reporters conducting the exit poll? [Answer: 87.2000%.]
    - What is the probability that a voter who speaks to the reporters is a Kerry voter? [Answer: 52.2935%.]
    - What is the probability that a voter who speaks to the reporters is a Bush voter? [Answer: 47.7064%.]
    - What is the probability that a voter who ignores the reporters is a Kerry voter? [Answer: 18.7500%.]
    - What is the probability that a voter who ignores the reporters is a Bush voter? [Answer: 81.2500%.]
    - What does the exit poll indicate? [Answer: Kerry has 52.3% of the vote and Bush has 47.7% of the vote. Kerry wins by 4.6%?! Oh dear....]
8. Here are the answers to the question about marijuana-using students at a university.
    (a) What percentage of the students at this college are marijuana users? [Answer: 43.7500%.]
    (b) What percentage of the students at this college are marijuana abstainers? [Answer: 56.2500%.]
    (c) If you pick a random marijuana-using student, what is the probability that they are majoring in the Arts & Humanities? [Answer: 52.0000%.]
    (d) If you pick a random marijuana-using student, what is the probability that they are majoring in the Social/Behavioral Sciences? [Answer: 41.1428%.]
    (e) If you pick a random marijuana-using student, what is the probability that they are majoring in S.T.E.M.? [Answer: 6.85714%.]
    (f) If you pick a random marijuana-abstaining student, what is the probability that they are majoring in the Arts & Humanities? [Answer: 21.7777%.]

(g) If you pick a random marijuana-abstaining student, what is the probability that they are majoring in the Social/Behavioral Sciences? [Answer: 48.0000%.]

(h) If you pick a random marijuana-abstaining student, what is the probability that they are majoring in S.T.E.M.? [Answer: 30.2222%.]

9. Here are the answers to question about the smartphone factory getting chips from Companies A, B, and C.

- What percentage of the chips, overall, are defective? [Answer: 1.35000%.]
- If I select a random defective chip, what is the probability that it is from Company A? [Answer: 33.3333%.]
- If I select a random defective chip, what is the probability that it is from Company B? [Answer: 51.8518%.]
- If I select a random defective chip, what is the probability that it is from Company C? [Answer: 14.8148%.]

Note: The above is very different from the overall supply distribution: 45%, 35%, 20%.

Note: While this analysis correctly reveals that Company B has lower quality than the other two, it does misleadingly suggest that Company C is significantly better than Company A, which is completely false.

- If I select a random non-defective chip, what is the probability that it is from Company A? [Answer: 45.1596%.]
- If I select a random non-defective chip, what is the probability that it is from Company B? [Answer: 34.7693%.]
- If I select a random non-defective chip, what is the probability that it is from Company C? [Answer: 20.0709%.]

Note: The above is very similar to the overall supply distribution: 45%, 35%, 20%. The difference is that Company A and Company C are slightly elevated (because they have 1% defective) and Company B is slightly lower (because it has 2% defective).

10. Here are the answers to the circuit-board diagnostic system.

(a) What is the probability that the fault is not located at all? [Answer: 6%.]

(b) If the fault is located, what is the probability that it was found by Procedure 1? [Answer: 74.4680%.]

(c) If the fault is located, what is the probability that it was found by Procedure 2? [Answer: 19.1489%.]

(d) If the fault is located, what is the probability that it was found by Procedure 3? [Answer: 6.38297%.]

(e) For a fault that is not located by Procedure 1, what is the probability that it is found by Procedure 2? [Answer: 60.0000%.]

(f) For a fault that is not located by Procedure 1, what is the probability that it is found by Procedure 3? [Answer: 20.0000%.]

(g) For a fault that is not located by Procedure 1, what is the probability that it ends up never being found? [Answer: 20.0000%.]

11. Here are the answers to the engineering students problem.

(a) What is the probability that a student who starts the program will successfully finish it? [Answer: 28.2240%.]

Note: Of course, at a real university, students can repeat failed classes or withdraw-and-retry a class, so the failure rate is never *that* bad.

(b) Given that a student flunks out, what is the probability that they failed Calculus I? [Answer: 13.9322%.]

(c) Given that a student flunks out, what is the probability that they failed Calculus II? [Answer: 37.6170%.]

(d) Given that a student flunks out, what is the probability that they failed Physics I? [Answer: 17.5546%.]

(e) Given that a student flunks out, what is the probability that they failed Calculus III? [Answer: 14.0436%.]

(f) Given that a student flunks out, what is the probability that they failed Physics II? [Answer: 16.8524%.]

(g) As you can see, Calculus II is the primary "cause of death," by a large margin. Most engineering students that I have spoken to would agree.

12. Here are the answers to the two-step hiring program.

(a) What is the probability that a randomly chosen hiree will be assigned to management position at the end of the trainee period? [Answer: 28.5000%.]

(b) What is the probability that a randomly chosen hiree will be fired at the end of the trainee period? [Answer: 29.5000%.]

(c) What is the probability that a randomly chosen hiree will be designated a management trainee (in Step I), but not be appointed to a management position (after Step II)? [Answer: 7.50000%.]

(d) After all is done, if you select someone appointed to a managerial position, what is the probability that they were assigned to the management track during Step II? [Answer: 78.9473%.]

(e) After all is done, if you select someone appointed to a non-managerial position, what is the probability that they were assigned to the non-management track during Step II? [Answer: 85.7142%.]

(f) After all is done, if you select someone who has been fired, what is the probability that they were fired during Step II? [Answer: 66.1016%.]

13. For the "confirmation" of Bernoulli's Distribution Formula, we are asked to compute the following probabilities:

Note: Before we start, feel free to check that the leaf probabilities all add to one.

(a) 4 up and 0 down: 0.6561

(b) 3 up and 1 down: $\underbrace{0.0729 + 0.0729 + 0.0729 + 0.0729}_{4 \text{ times}} = 0.2916$

(c) 2 up and 2 down: $\underbrace{0.0081 + 0.0081 + 0.0081 + 0.0081 + 0.0081 + 0.0081}_{6 \text{ times}} = 0.0486$

(d) 1 up and 3 down: $\underbrace{0.0009 + 0.0009 + 0.0009 + 0.0009}_{4 \text{ times}} = 0.0036$

(e) 0 up and 4 down: 0.0001

Note: It is extremely worthwhile to also compute these five probabilities with Bernoulli's Distribution Formula, so that you can see with your own eyes that indeed, all the numbers match.