Chapter 2 Probabilistic Thinking

"*In any decision problem, the only thing that is certain is uncertainty*"

2.1 Representing Uncertainty

2.1.1 Probability Distributions and Trees

- Suppose we are doing a survey on the beer-drinking habits of customers at a shopping mall.

- We are concerned about whether the next person we meet is a beer drinker or not.

- We denote this uncertain event by B.

- Let b_1 and b_2 be two possible outcomes of B, where b_1 denotes the outcome that the next person we meet is a beer drinker, and b_2 denotes otherwise.

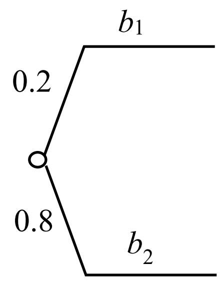
- Suppose we assess that the probability of outcome b_1 is 0.2 and that of b_2 is 0.8, we denote this information by a probability distribution denoted as follow:

  - $p(b_1 \mid \xi) = 0.2$
  - $p(b_2 \mid \xi) = 0.8$

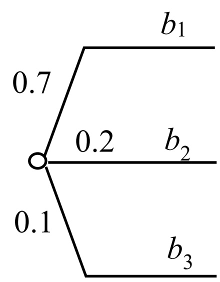
  where ξ represents all the **background information** brought to bear in assessing the probability values.

- Note that we must always have $p(b_1 \mid \xi) + p(b_2 \mid \xi) = 1$ exactly. Otherwise, the information is not valid.

- We may also represent this information using a **Probability Tree** as follow:



- For an event with three possible outcomes we use a probability tree with three branches:



- The probability distribution is denoted by:

  - $p(b_1 \mid \xi) = 0.7$
  - $p(b_2 \mid \xi) = 0.2$
  - $p(b_3 \mid \xi) = 0.1$

- Again, we must have $p(b_1 \mid \xi) + p(b_2 \mid \xi) + p(b_3 \mid \xi) = 1$.

2.1.2 Interpretations of Probability

- There are two main contending views about how to interpret and understand the meaning of probability:

  1. The Frequentist view
  2. The Subjective view

The Frequentist View

- Probabilities are fundamentally dispositional properties of non-deterministic physical systems,

- Probabilities are viewed as long-run frequencies of events. If an event is repeated, then the frequency of occurrence, as the number of trials approaches infinity, is the probability of the event.

- This is the standard interpretation used in classical statistics where experiments are conducted and data collected for inference.

Examples

1. The probability of kill of a surface-to-air missile is the proportion of missiles that hit the target if a very large number of identical missiles are fired under identical conditions.

2. The probability that a light bulb lasts for at least 1,000 hours is the proportion of identical bulbs that achieved that life in a very large sample.

The Subjective (Bayesian) View

- Probabilities are representations of our subjective degree of belief, according to Thomas Bayes and Pierre Simon de Laplace.

- Probabilities in general are not necessarily tied to any physical or process which can be repeated indefinitely.

Thomas Bayes (1702-1761)

Examples

1. A company may believe that a new product to be launched will be successful with a certain probability based on its market research.

2. You believe that you will score at least an A- for this course with a certain probability because you have reviewed its content and made an assessment based on your ability.

- Subjective probabilities are based on individual personal judgments based on the person's knowledge, expertise, past data, or any other relevant information.

- It is likely that no two individuals will always give the same numerical value for the same event. Hence, there is no such thing as "The Probability" of an event.

- Ultimately, all probabilities are subjectively assessed based on some information.

- We call the information brought to bear in assessing a probability the "**background information"** for the probability.

2.1.3 Marginal and Conditional Probabilities

- Probabilities that were assessed solely based on a specific set of background information are called Marginal Probabilities.

- For example, the following probabilities for outcomes of event B are marginal probabilities:

  - $p(b_1 \mid \xi) = 0.2$
  - $p(b_2 \mid \xi) = 0.8$

- In practice, we may not be able to assess the marginal probability distribution of an event directly.

- However, we may be able to assess its probability conditioned on the outcomes of another event. This results in Conditional Probabilities.

- The **Conditional Probability** of event X given event Y and background information ξ, denoted by $p(X \mid Y, \xi)$, is

$$p(X \mid Y, \xi) = \frac{p(X \text{ and } Y \mid \xi)}{p(Y \mid \xi)} \qquad \text{for } p(Y \mid \xi) \neq 0.$$

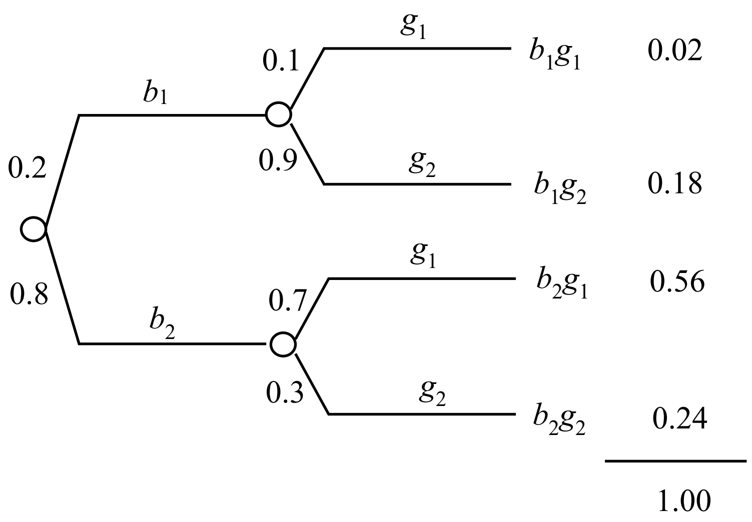
- Note that the numerator $p(X \text{ and } Y \mid \xi)$ is the probability that both X and Y occur. It is called the **Joint Probability** for X and Y.

- Also, the denominator $p(Y \mid \xi)$ cannot be zero because it is not possible or meaningless to talk about the probability of an event conditioned on another event that definitely will not happen.

Example (Conditional Probability)

- Suppose we are concern with whether the next person we meet is a beer-drinker or not (denoted by event B), and also whether he/she is a graduate or not (denote by event G).

- We could represent events B and G with the following probability tree:



- The probabilities of g_1 and g_2 are based on whether b_1 or b_2 occurs in addition to the general background information ξ.

- These **Conditional Probabilities** are denoted by:

$$p(g_1 | b_1, \xi) = 0.1 \qquad p(g_2 | b_1, \xi) = 0.9$$
$$p(g_1 | b_2, \xi) = 0.7 \qquad p(g_2 | b_2, \xi) = 0.3$$

- The above set of numbers is also called a **Conditional Probabilities Table** (CPT).

2.1.4 Probabilistic Dependence and Independence

- Let A and B be two uncertain binary-outcome events.
- Let $p(A \mid B, \xi)$ be the conditional probability of A given B.

Probabilistic Dependence

- Consider the following conditional probability table for A and B:

	$p(A{=}a_1 \mid B, \xi)$	$p(A{=}a_2 \mid B, \xi)$
$B = b_1$	0.7	0.3
$B = b_2$	0.2	0.8

- We observe that the conditional probability distribution for A **changes** when the outcome of B switches from b_1 to b_2. We say that A is probabilistically dependent on B (given background information ξ).
- Events A and B are said to be **Probabilistic Dependent** if $p(A \mid B{=}b_1, \xi) \neq p(A \mid B{=}b_2, \xi)$ for $A{=}a_1$ and a_2. That is, the conditional probability of A given B depends on the specific outcome of B.

Probabilistic Independence

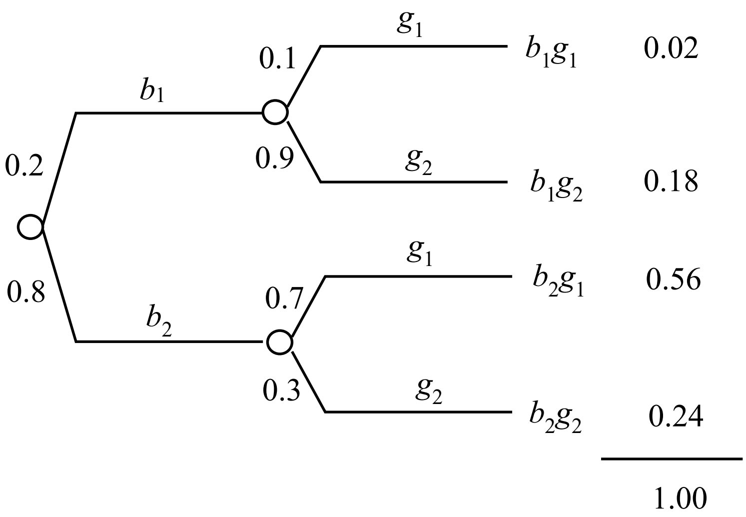
- Consider the following conditional probability table for A and B:

	$p(A{=}a_1 \mid B, \xi)$	$p(A{=}a_2 \mid B, \xi)$
$B = b_1$	0.7	0.3
$B = b_2$	0.7	0.3

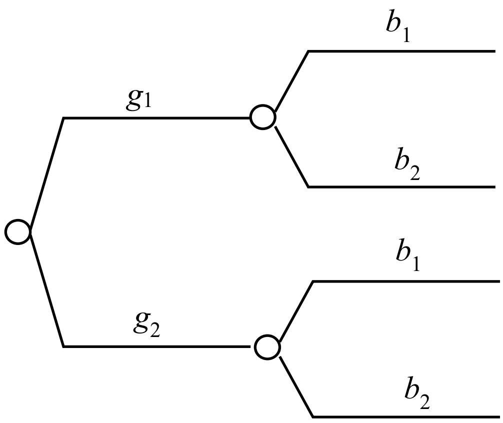
- We observe that the probability distribution for A **does not depend** on whether the outcome of B is b_1 or b_2. We say that A is probabilistically independent of B (given background information ξ).
- Events A and B are said to be **Probabilistic Independent** if $p(A \mid B{=}b_1, \xi) = p(A \mid B{=}b_2, \xi)$ for $A{=}a_1$ and a_2. That is, the conditional probability of A given B does not depend on the specific outcome of B.
- In fact, if A is independent of B, then $p(A \mid B, \xi) = p(A \mid \xi)$.
- Intuitively, probabilistic independence means knowing the outcome of one event does not provide any information on the probability of outcomes of the other event.

2.1.5 Changing the Order of Conditioning

- In the above example, we have drawn the tree (reproduced below) in the order B to G and conditioned the probability of G to the possible outcomes of B.



- In decision making, we may need to have the probability tree drawn in the order G to B:



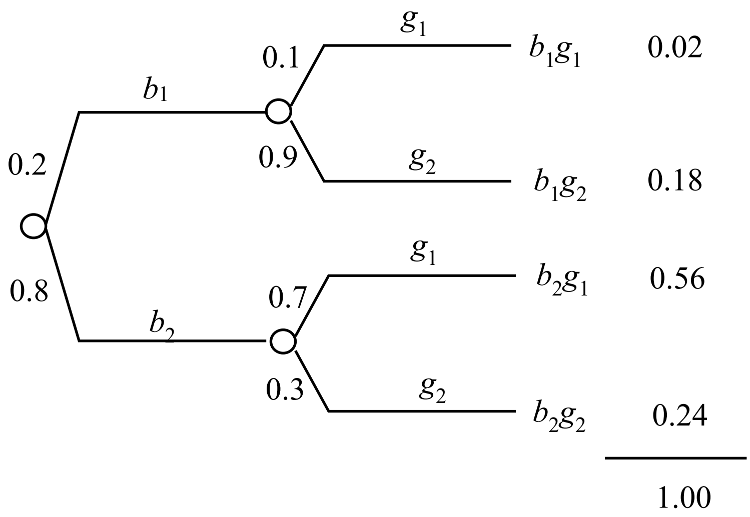
- We will need the following probabilities to complete the tree:

  - $p(g_1 \mid \xi)$, $\quad p(g_2 \mid \xi)$
  - $p(b_1 \mid g_1, \xi)$, $p(b_2 \mid g_1, \xi)$
  - $p(b_1 \mid g_2, \xi)$, $p(b_2 \mid g_2, \xi)$

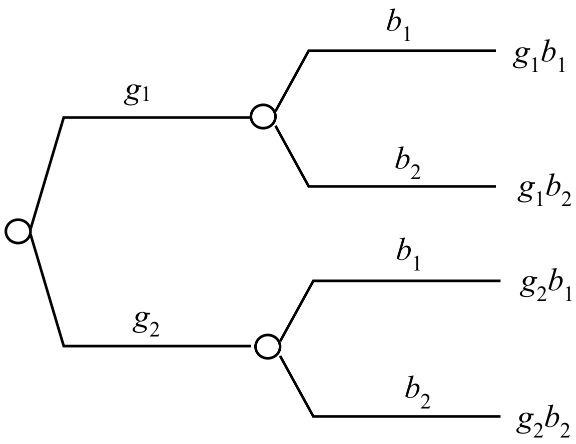
- These values may be obtained by "Flipping the Tree".

Method 1: Graphical Approach

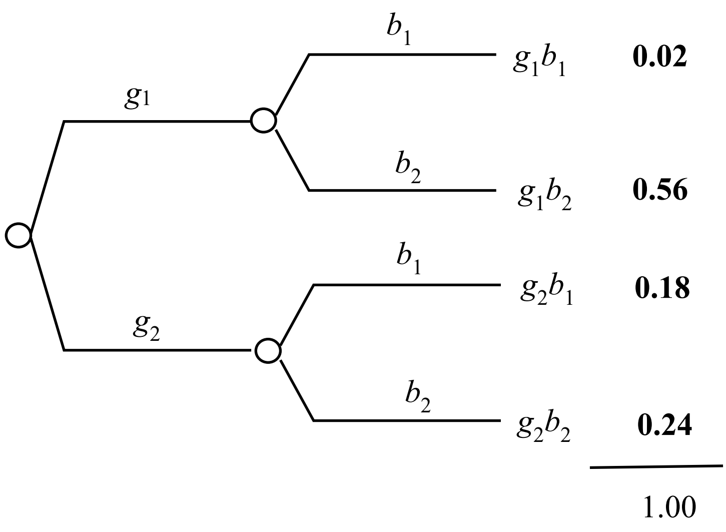
- Given the probability tree:



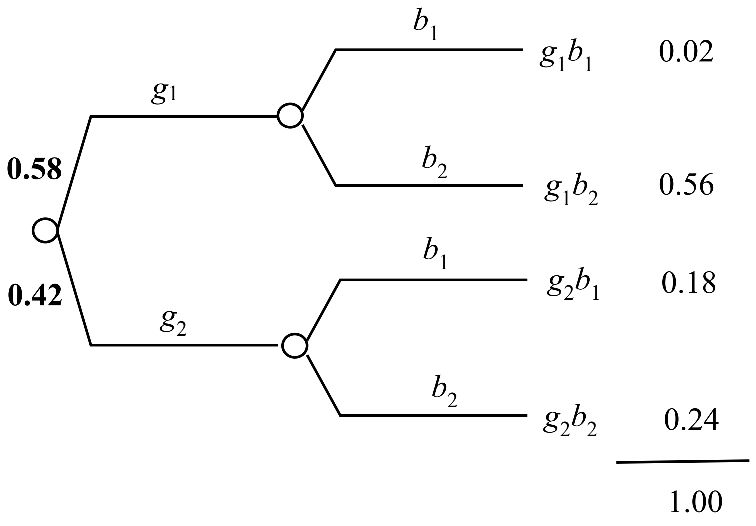
- Step 1: Change the ordering of the underlying tree:



- Step 2: Transfer the end-point joint probabilities from the original tree to the new tree:



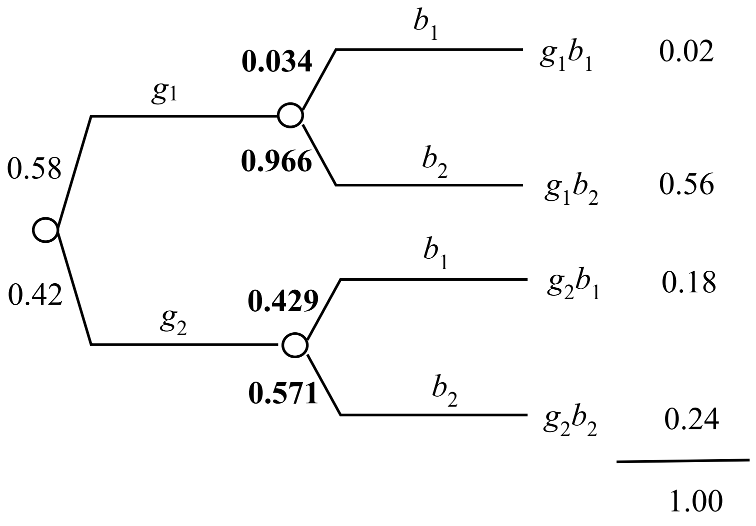
- Step 3: Compute the marginal probability for the first variable in the new tree, i.e., G. We add the joint probabilities that are related to g_1 and g_2 respectively.



$$p(g_1 | \xi) = 0.02 + 0.56 = 0.58$$

$$p(g_2 | \xi) = 0.18 + 0.24 = 0.42$$

- Step 4: Compute conditional probabilities for B given G.



$p(b_1 | g_1, \xi) = 0.02/0.58 = 0.034$

$p(b_2 | g_1, \xi) = 0.56/0.58 = 0.966$

$p(b_1 | g_2, \xi) = 0.18/0.42 = 0.429$

$p(b_2 | g_2, \xi) = 0.24/0.42 = 0.571$

Method 2: Using Bayes' Theorem

- If you can do the above tree-flipping to change the order of conditioning of two events, you are already applying Bayes' Theorem.

Bayes' Theorem

- Given two uncertain events X and Y. Suppose the probabilities $p(X \mid \xi)$ and $p(Y \mid X, \xi)$ are known, then

$$p(X \mid Y, \xi) = \frac{p(X \mid \xi)\, p(Y \mid X, \xi)}{p(Y \mid \xi)}$$

where the denominator may be computed as:

$$p(Y \mid \xi) = \sum_X p(X \mid \xi)\, p(Y \mid X \mid \xi)$$

Example

- In the previous tree-flipping example, we were given:

$p(b_1 \mid \xi) = 0.2$	$p(g_1 \mid b_1, \xi) = 0.1$	$p(g_1 \mid b_2, \xi) = 0.7$
$p(b_2 \mid \xi) = 0.8$	$p(g_2 \mid b_1, \xi) = 0.9$	$p(g_2 \mid b_2, \xi) = 0.3$

- We may compute the values of $p(b_1 \mid g_1, \xi)$ and $p(b_1 \mid g_2, \xi)$ as follows:

$$p(b_1 \mid g_1, \xi) = \frac{p(b_1 \mid \xi)\, p(g_1 \mid b_1, \xi)}{p(b_1 \mid \xi)\, p(g_1 \mid b_1, \xi) + p(b_2 \mid \xi)\, p(g_1 \mid b_2, \xi)} = \frac{0.2 \times 0.1}{(0.2 \times 0.1) + (0.8 \times 0.7)} = 0.034$$

$$p(b_1 \mid g_2, \xi) = \frac{p(b_1 \mid \xi)\, p(g_2 \mid b_1, \xi)}{p(b_1 \mid \xi)\, p(g_2 \mid b_1, \xi) + p(b_2 \mid \xi)\, p(g_2 \mid b_2, \xi)} = \frac{0.2 \times 0.9}{(0.2 \times 0.9) + (0.8 \times 0.3)} = 0.429$$

2.2 Application of Conditional Probabilities

Let's Make a Deal Game Show (Monte Hall Problem)

- Consider the TV game show where the contestant is shown on stage three doors, behind one of which contains a valuable prize; the other two nothing. The host knows where the prize is at the start of the game.

Rules of the Game

- The contestant is asked to choose one of the doors.
- The host will open one of the unchosen doors making sure that it has nothing.
- The contestant is then asked to decide if he wishes to keep his original selection or switch to the other unopen door.
- The contestant receives the "prize" behind the door that he finally chooses.

Questions

- Assuming that the prize was equally likely to be behind any of the three doors at the start of the game, and suppose the contestant chose door A and the host opened door B and was shown to be empty, what advice would you give the contestant?

  I. Stick to original choice door A

  II. Switch to the unopen door C

  III. It does not matter whether I choose A or C, as the chances of winning the prize are equal.

  Answer: ____________________

- Many people chose option III. One possible line of reasoning for choosing III is as follows:

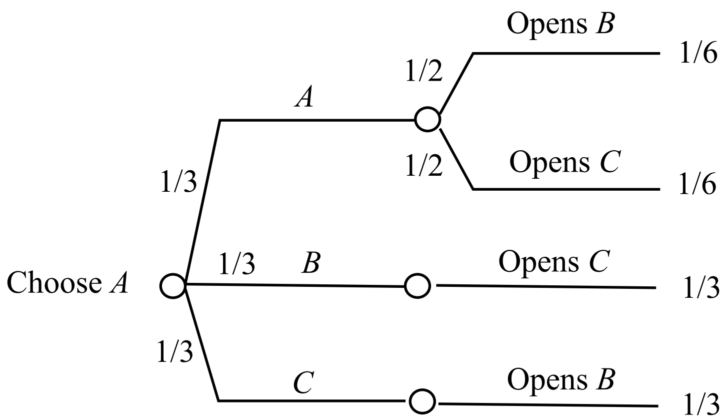
  Since the prize is equally likely to be behind any door, and since door B is confirmed to be empty, the prize is now equally likely to be behind either A or C. Hence, it does not matter whether I switch from A to C or not.

- Do you agree with the above reasoning?

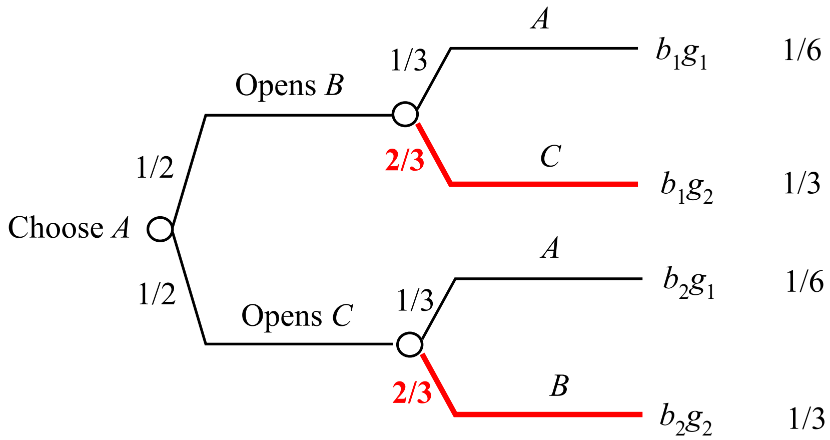
- Let us do a proper probabilistic analysis to select the door.

Probabilistic Analysis

- Assume that the doors are labeled A, B, and C, which also denotes the event that the prize is behind it.
- At the beginning of the game, we have $p(A) = p(B) = p(C) = 1/3$.
- Suppose the contestant chose A, then the conditional probabilities that the host will open door B or C are given by the following probability tree:

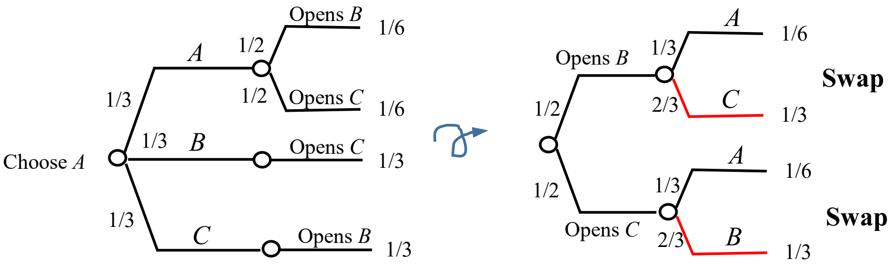


- If B is open (confirming that the prize is not there), we are interested in determining the following conditional probabilities:
  - p(Prize is at A | B is open, i.e., is empty)
  - p(Prize is at C | B is open, i.e. is empty)
- These probabilities can be obtained by flipping the above tree:

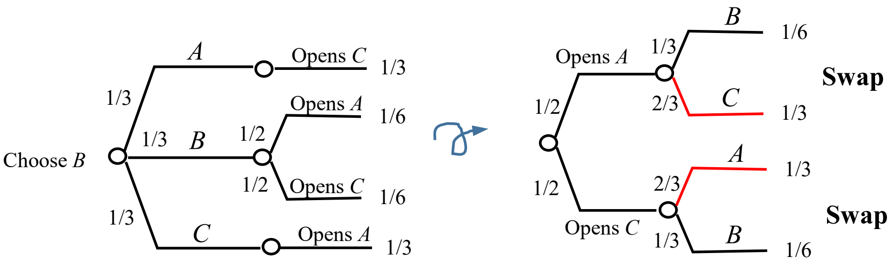


- From the flipped tree, we have
  - p(Prize is at A | B is open) = 1/3
  - p(Prize is at C | B is open) = 2/3
- The probability that the prize is at C given that B has been open is twice the probability that the prize is at A.
- Hence, the contestant should switch from A to C to double his chance of winning.
- In general, if you have chosen any door, just switch to the unopened door after the host has opened a door.

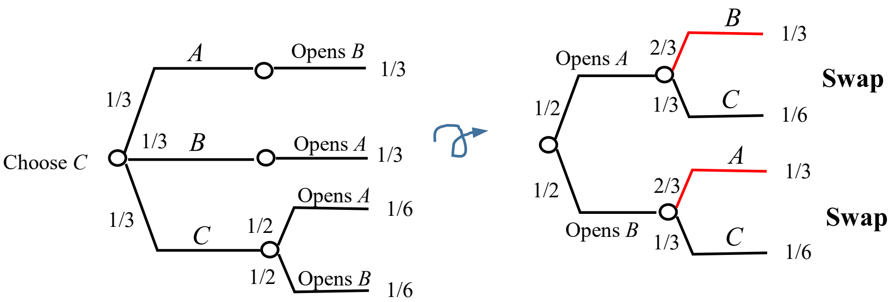
Complete Probabilistic Analysis



- If A is chosen, then
  If the host opens B, decision is swap to C to increase the probability from 1/3 to 2/3.
  If the host opens C, decision is swap to B to increase the probability from 1/3 to 2/3.



- If B is chosen, then
  If the host opens A, decision is swap to C to increase the probability from 1/3 to 2/3.
  If the host opens C, decision is swap to A to increase the probability from 1/3 to 2/3.



- If C is chosen, then
  If the host opens A, decision is swap to B to increase the probability from 1/3 to 2/3.
  If the host opens B, decision is swap to A to increase the probability from 1/3 to 2/3.

General Optimal Strategy

- Choose any door, and after the host opens a door, always swap the door you have chosen earlier with the unopened door.

- In all cases, your probability of winning will increase from 1/3 to 2/3.

Extensions to the Problem:

1. 3-door Problem with unequal prior probabilities
2. 4-door Problem with equal and unequal priors
3. N-door Problem with equal and unequal priors

2.3 Dealing with New Information using Bayes Theorem

2.3.1 Updating of Probabilities Based on New Evidence or Information

- Recall that we assign a probability $p(A \mid \xi)$ to the outcome of an event A based on our assessment of the likelihood of the event using whatever background information ξ we may have.

- $p(A \mid \xi)$ is known as our **Prior Probability** for event A.

- Suppose that later, a piece of new information or evidence E has arrived. How do we incorporate the new information or evidence into our assessment of the probability of A?

- The updated probability for A is given by $p(A \mid E, \xi)$ which is known as the **Posterior Probability** of A in the light of evidence E.

- It can be computed via Bayes' Theorem:

$$p(A \mid E, \xi) = \frac{p(A \mid \xi)\, p(E \mid A, \xi)}{p(E \mid \xi)}$$

where $p(E \mid A, \xi)$ is called the **Likelihood Function** for the evidence E and ξ, and

$$p(E \mid \xi) = \sum_j p(A_j \mid \xi)\, p(E \mid A_j, \xi).$$

Example: Weather Forecasting

- We are interested in the weather tonight. Define the uncertain event R as the weather condition tonight with outcome r_1 = "Rains tonight" and outcome r_2 = "Does not rain tonight".

- Based on our judgment and experience, we assess the following prior probabilities:

  $p(r_1 \mid \xi) = 0.6$
  $p(r_2 \mid \xi) = 0.4$

- We can improve our understanding of the weather tonight if we have more information.

- A weather forecast service is available free of charge. Unfortunately, its forecast is not always correct.

- How do we combine the information we already have with new information from the forecast which is subject to error?

- Let S represents the imperfect weather forecast with two possible outcomes:

  s_1 = "Weather forecast = rain"
  s_2 = "Weather forecast = no rain"

- Based on past performance of the forecast, we assess the **correctness** of the weather station by the following matrix:

Actual weather condition	**Weather forecast**	
	It will rain tonight (s_1)	It will not rain tonight (s_2)
Rains tonight (r_1)	0.9	0.1
Does not rain tonight (r_2)	0.3	0.7

- Interpretation of the data: Given that it will rain tonight, the station will forecast it correctly with a 90% chance, and given it will not rain tonight, the station will forecast it correctly with a 70% chance.

- Stated in terms of conditional probabilities
  - $p(s_1 | r_1, \xi) = 0.9$
  - $p(s_2 | r_2, \xi) = 0.7$

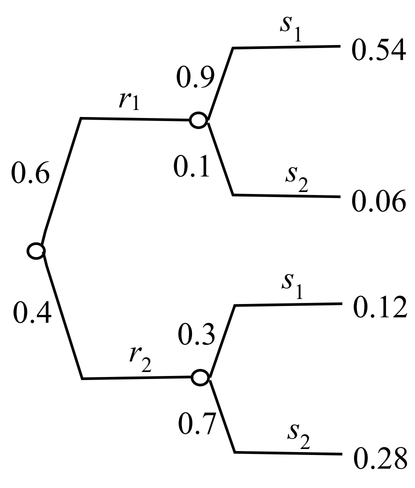
- Suppose you tune in to the weather station now and it says "It will rain tonight" i.e., s_1 is observed to be true, what probability should you now assign to the outcome it will indeed rain tonight?

- Remember that your prior assignment is $p(r_1 | \xi) = 0.6$.

- Given that s_1 is true, you should now be interested in the probability $p(r_1 \mid s_1, \xi)$ which is the probability that it will rain tonight given that the forecast is "It will rain tonight".

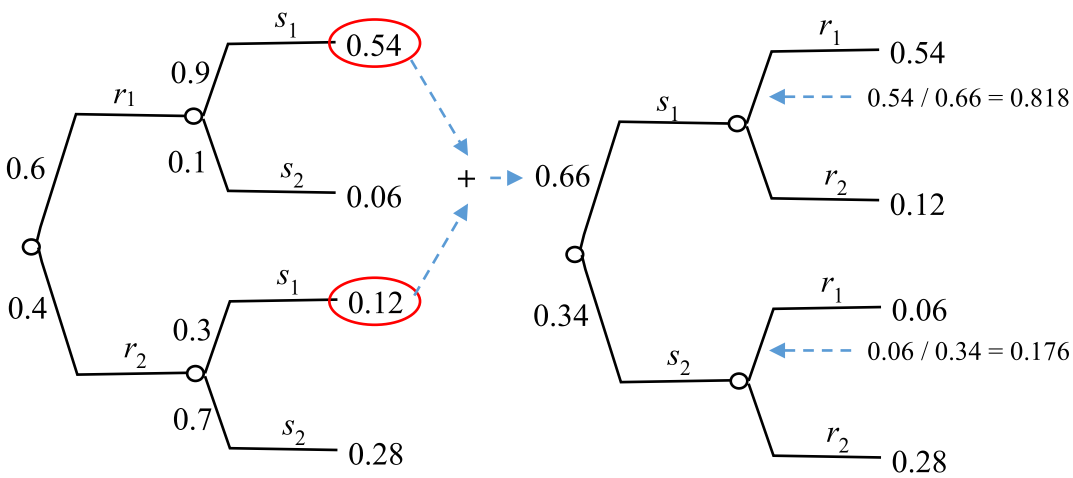
- Unfortunately, this information is not directly available and you have to flip the tree or apply Bayes' Theorem to get it.

- Your prior knowledge about the weather together with your knowledge about the **correctness** of the weather station can be represented by the probability tree.

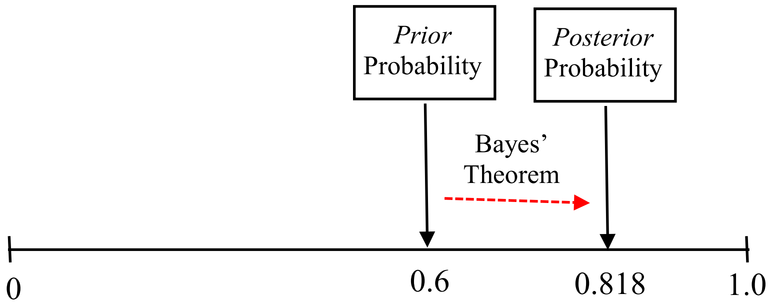


- Notice that the ordering of conditioning is not what we want.

- Hence we flip the probability tree:



- Hence given that the weather station says "it will rain tonight", the probability we should now assign to the outcome that it will actually rain tonight is 0.818, up from the prior value of 0.6.



- Numerical computations using Bayes' Theorem are as follows.

$p(r_1 \mid \xi) = 0.6 \qquad p(r_2 \mid \xi) = 0.4$

$p(s_1 \mid r_1, \xi) = 0.9 \qquad p(s_2 \mid r_1, \xi) = 0.1$

$p(s_1 \mid r_2, \xi) = 0.3 \qquad p(s_2 \mid r_2, \xi) = 0.7$

$$p(r_1 \mid s_1, \xi) = K\, p(r_1 \mid \xi)\, p(s_1 \mid r_1, \xi) = (0.6)(0.9)\, K = 0.54\, K$$
$$p(r_2 \mid s_1, \xi) = K\, p(r_2 \mid \xi)\, p(s_1 \mid r_2, \xi) = (0.4)(0.3)\, K = 0.12\, K$$

where K is a constant which can be eliminated by using the constraint
$$p(r_1 \mid s_1, \xi) + p(r_2 \mid s_1, \xi) = 1.$$

- Equivalently, we can normalize the numbers 0.54 and 0.12 (to add up to 1) and obtain:

$$p(r_1 \mid s_1, \xi) = \frac{0.54}{0.54 + 0.12} = 0.818$$

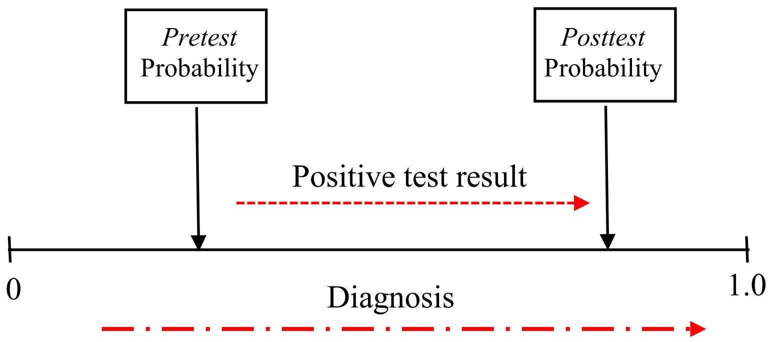
$$p(r_2 \mid s_1, \xi) = \frac{0.12}{0.54 + 0.12} = 0.182.$$

2.3.2 Application of Bayes' Theorem in Diagnosis and Testing

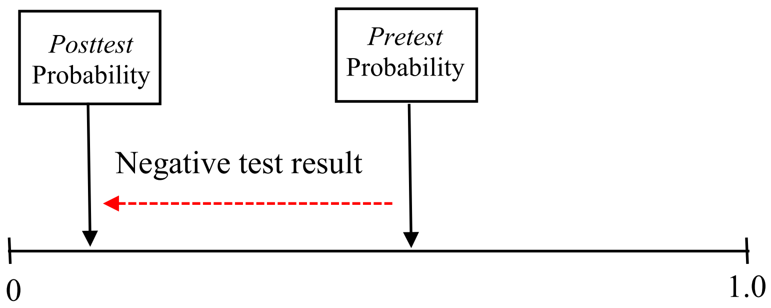
- In diagnosis, we are interested in the probability of an event that can be either "supported" or "ruled out" based on one or more tests.

- Examples of such events are: A person suffering from disease *X*, the transmission system of a vehicle being faulty, etc.

- A diagnosis process where positive test results or findings "supported" the initial diagnosis as shown below:



- It is also possible that the test produces negative results that reduce the pretest probability:



- For brevity, we will be omitting the ξ symbol in our probability notations from now on, but they are understood to be there.

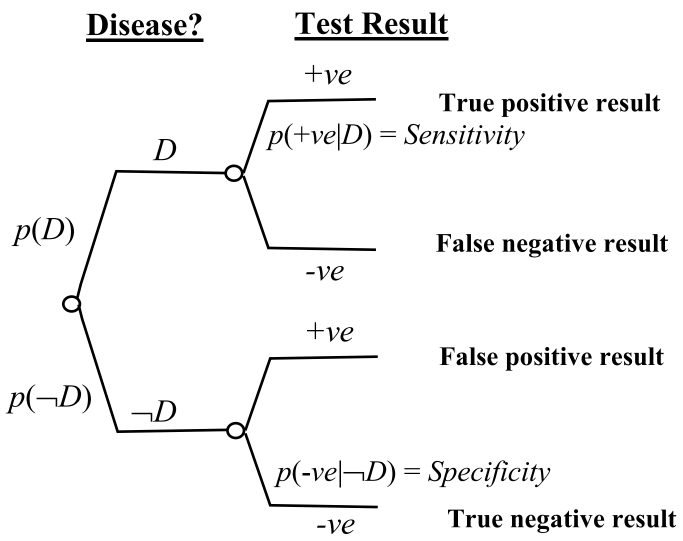
- Let D be the event that is being diagnosed, and let $p(D)$ = pretest probability = probability that a disease or fault is present before conducting any test.

Example

- In trying to determine if a person is suffering from TB, the prevailing risk of contracting TB in a certain community can be used as the pre-test probability.

- Let T be a test whose outcome can be either positive ($+ve$) or negative ($-ve$).

Four possible situations arising out of the test



True-Positive Result

- Given that the disease or fault is present, the probability of the test detecting it (i.e., correctly gives a positive test result) = $p(+ve \mid D)$.
- $p(+ve \mid D)$ is called the **True-Positive rate** or **Sensitivity** of the test.

True-Negative Result

- Given that the disease or fault is not present, the probability of the test indicating the absence of it (i.e., correctly gives a negative test result) = $p(-ve \mid \neg D)$.
- $p(-ve \mid \neg D)$ is called the **True-Negative rate** or **Specificity** of the test.

False-Positive Result

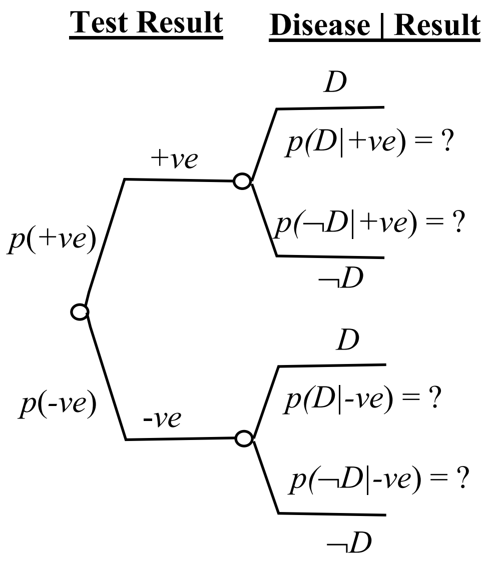
- Given that the disease or fault is not present, the probability of the test wrongly indicating a positive test result = $p(+ve \mid \neg D) = 1 - p(-ve \mid \neg D) = 1$ – specificity.
- $p(+ve \mid \neg D)$ is called the **False-Positive rate** of the test

False-Negative Result

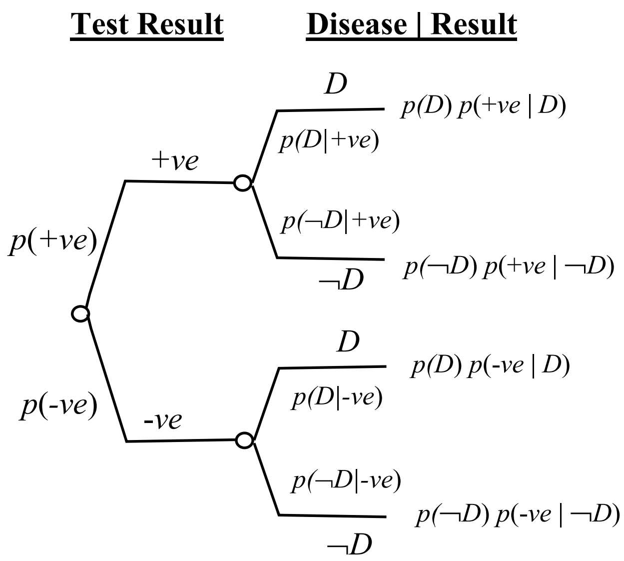
- Given that the disease or fault is present, the probability of the test wrongly indicating a negative test result = $p(-ve \mid D) = 1 - p(+ve \mid D) = 1$ – sensitivity.
- $p(-v\ e \mid D)$ is called the **False-Negative rate** of the test.

- In practice, we would like to determine what is the probability of the suspected cause D given a positive or a negative test result.

- We are interested in the information shown in the following tree:



- We can obtain the required conditional probabilities by flipping the original tree:



$$p(+ve) = p(D)p(+ve \mid D) + p(\neg D)p(+ve \mid \neg D)$$

$$p(-ve) = p(D)p(-ve \mid D) + p(\neg D)p(-ve \mid \neg D)$$

Case 1: When the test result is positive

- When the test result is positive, the post-test probability is

$$p(D \mid +ve) = \frac{p(D)\, p(+ve \mid D)}{p(D)\, p(+ve \mid D) + p(\neg D)\, p(+ve \mid \neg D)}$$
$$= \frac{p(D)\, p(+ve \mid D)}{p(D)\, p(+ve \mid D) + (1 - p(D))(1 - p(-ve \mid \neg D))}$$

  where

  - $p(D)$ is the prior probability for D.
  - $p(+ve \mid D)$ is the *sensitivity* (*true-positive rate*) of the test.
  - $p(-ve \mid \neg D)$ is the *specificity* (*true-negative rate)* of the test.

Case 2: When the test result is negative

- When the test result is negative, the post-test probability is

$$p(D \mid -ve) = \frac{p(D)\, p(-ve \mid D)}{p(D)\, p(-ve \mid D) + p(\neg D)\, p(-ve \mid \neg D)}$$
$$= \frac{p(D)(1 - p(+ve \mid D))}{p(D)(1 - p(+ve \mid D)) + (1 - p(D))\, p(-ve \mid \neg D)}$$

  where

  - $p(D)$ is the prior probability for D.
  - $p(+ve \mid D)$ is the *sensitivity* (*true-positive rate*) of the test.
  - $p(-ve \mid \neg D)$ is the *specificity* (*true-negative rate)* of the test.

Example: Diagnosis of lung cancer using X-ray

- An elderly man with hemoptysis and a long history of cigarette smoking is suspected to have lung cancer and the estimated pretest probability is 0.4.

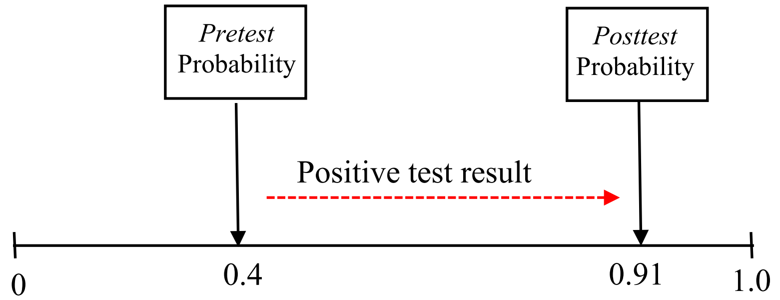
- The interpretation of the chest X-ray is "mass lesion in the right upper lobe".

- The effectiveness of the X-ray in detecting lung cancer is
  - Sensitivity or true-positive rate = 60%
  - Specificity or true-negative rate = 96%

- How should the doctor interpret the finding?

- The probability of lung cancer given a positive X-ray result is

$$p(D\,|+ve) = \frac{(0.4)(0.6)}{(0.4)(0.6)+(1-0.4)(1-0.96)} = 0.91$$

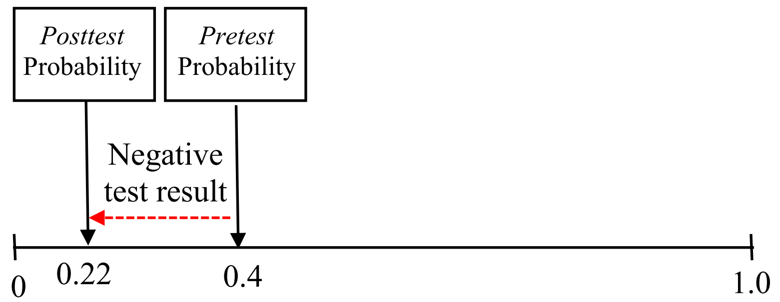


- Conclusion: The patient probably has lung cancer, although further confirmation may be required before starting treatment.

What if the chest X-ray had not shown a mass lesion?

- The probability of lung cancer given a negative X-ray result is

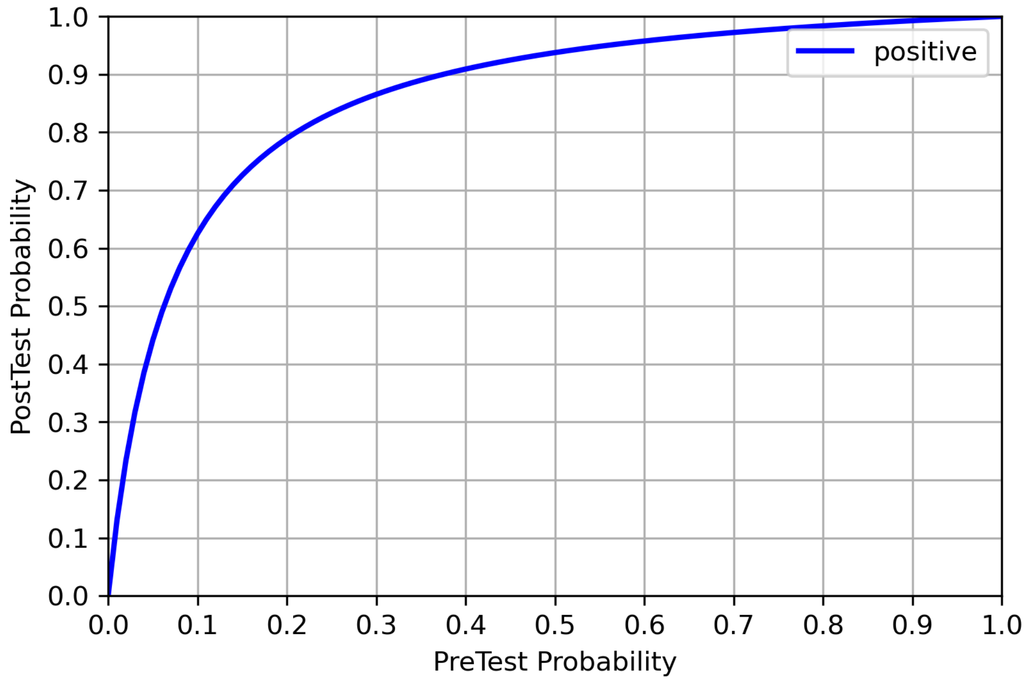
$$p(D\,|-ve) = \frac{(0.4)(1-0.6)}{(0.4)(1-0.6)+(1-0.4)(0.96)} = 0.22$$



The Effect of Pretest Probability on Posttest Probability

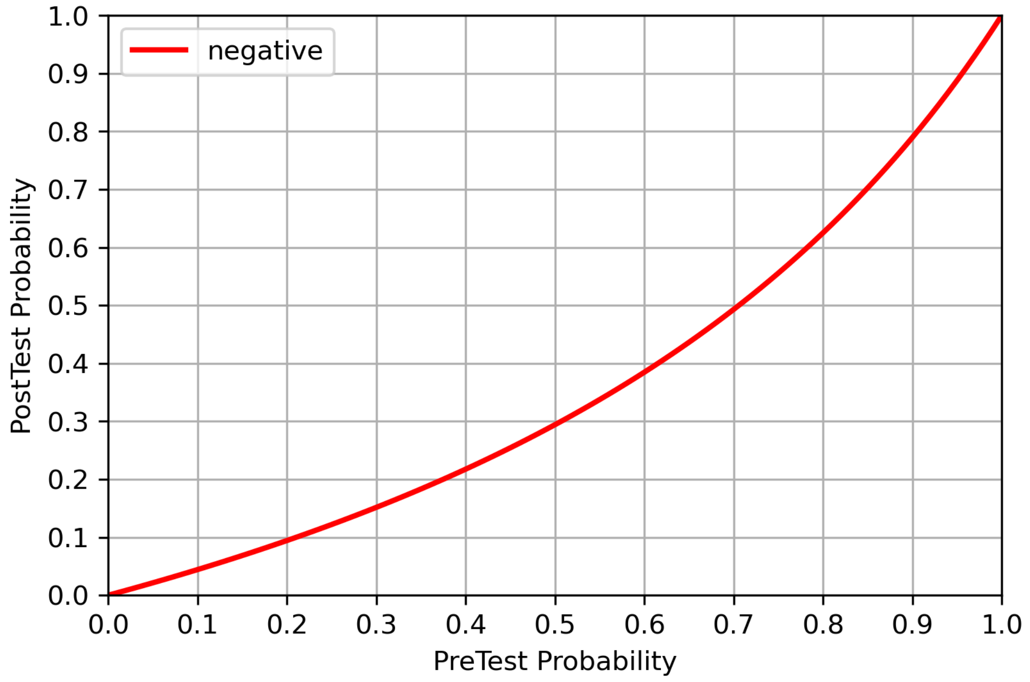
- How would the pretest probability affect the final conclusion?

I. When the result is positive:



- The response between posttest probability and pretest probability is usually an increasing concave function.

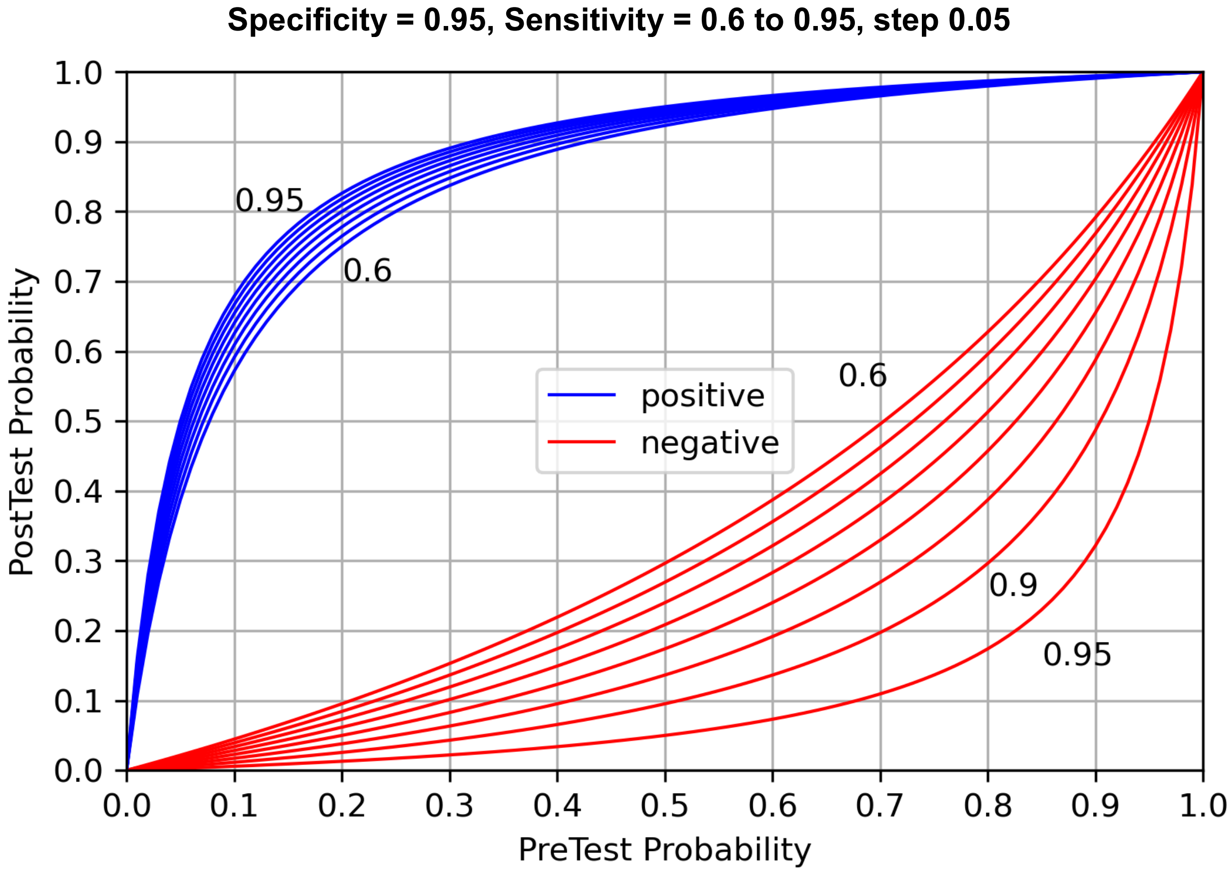
II. When the result is negative:



- The response between posttest probability and pretest probability is usually an increasing convex function.

Effect of Test Sensitivity

- Suppose we fix the Test Specificity at 0.95 and allow the Test Sensitivity to vary from 0.60 to 0.95 in steps of 0.05.

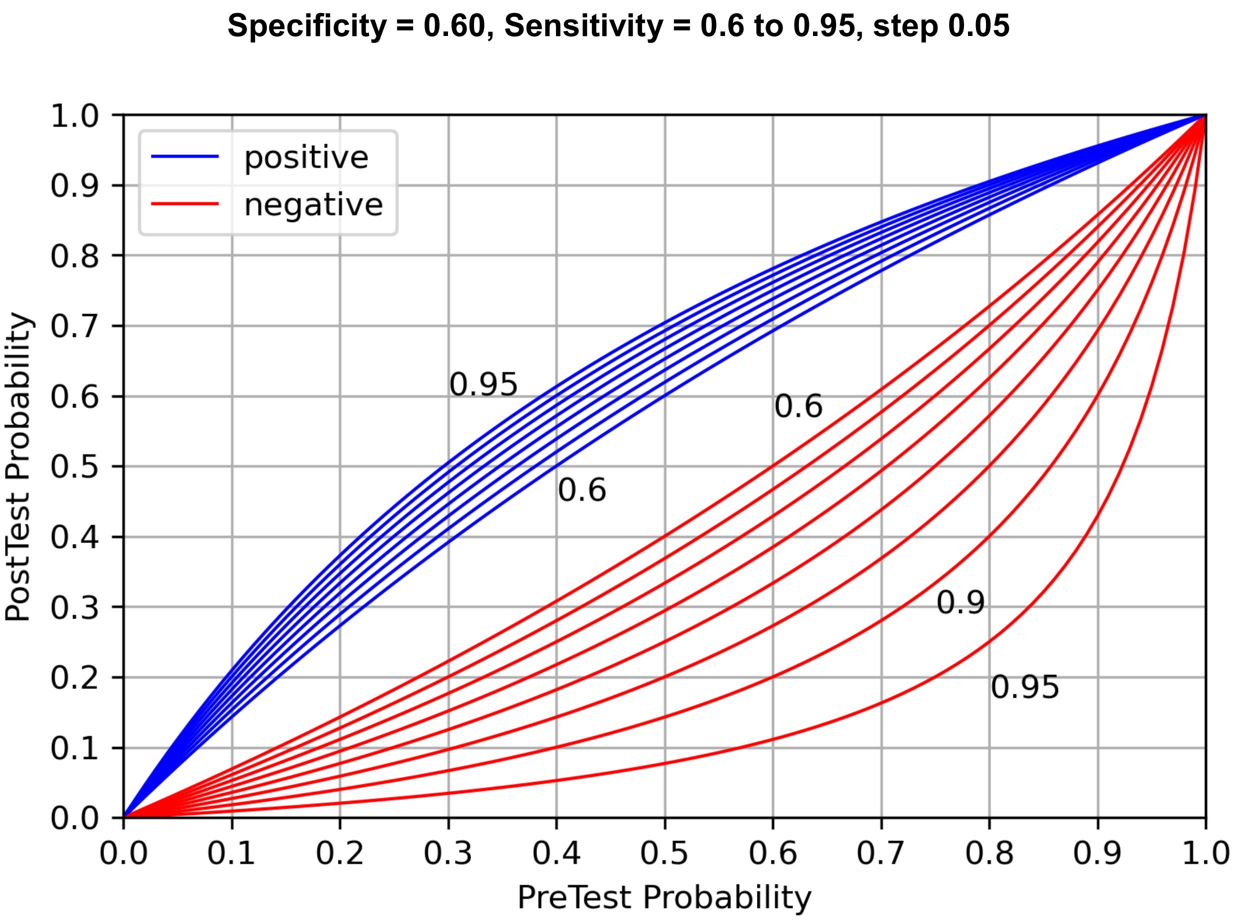


Observations:

- When the test result is positive, the set of response curves in the upper region are all near to each other. Hence for any pretest probability, the posttest probability *will not vary very much* when the test sensitivity varies from 0.6 to 0.95.

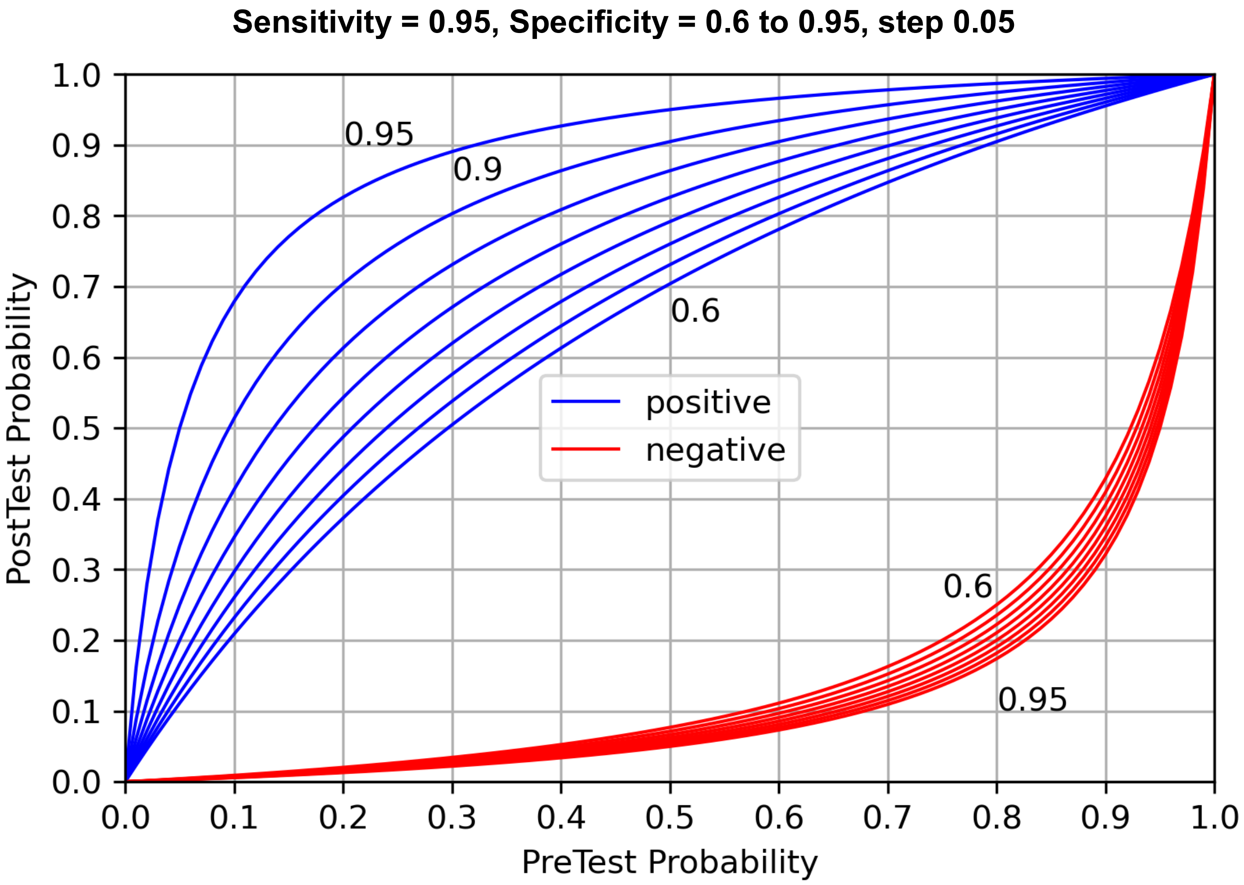
- On the other hand, when the test result is negative, the set of response curves in the lower region are very far apart. Hence for any pretest probability, the posttest probability *will vary* a lot when the test sensitivity varies from 0.6 to 0.95.

- We observe similar results when the Test Specificity is fixed at 0.6 and the Test Sensitivity varies from 0.60 to 0.95 in steps of 0.05.



Effect of Test Specificity

- Suppose we fix the Test Sensitivity at 0.95 and allow the Test Specificity to vary from 0.6 to 0.95 in steps of 0.05.

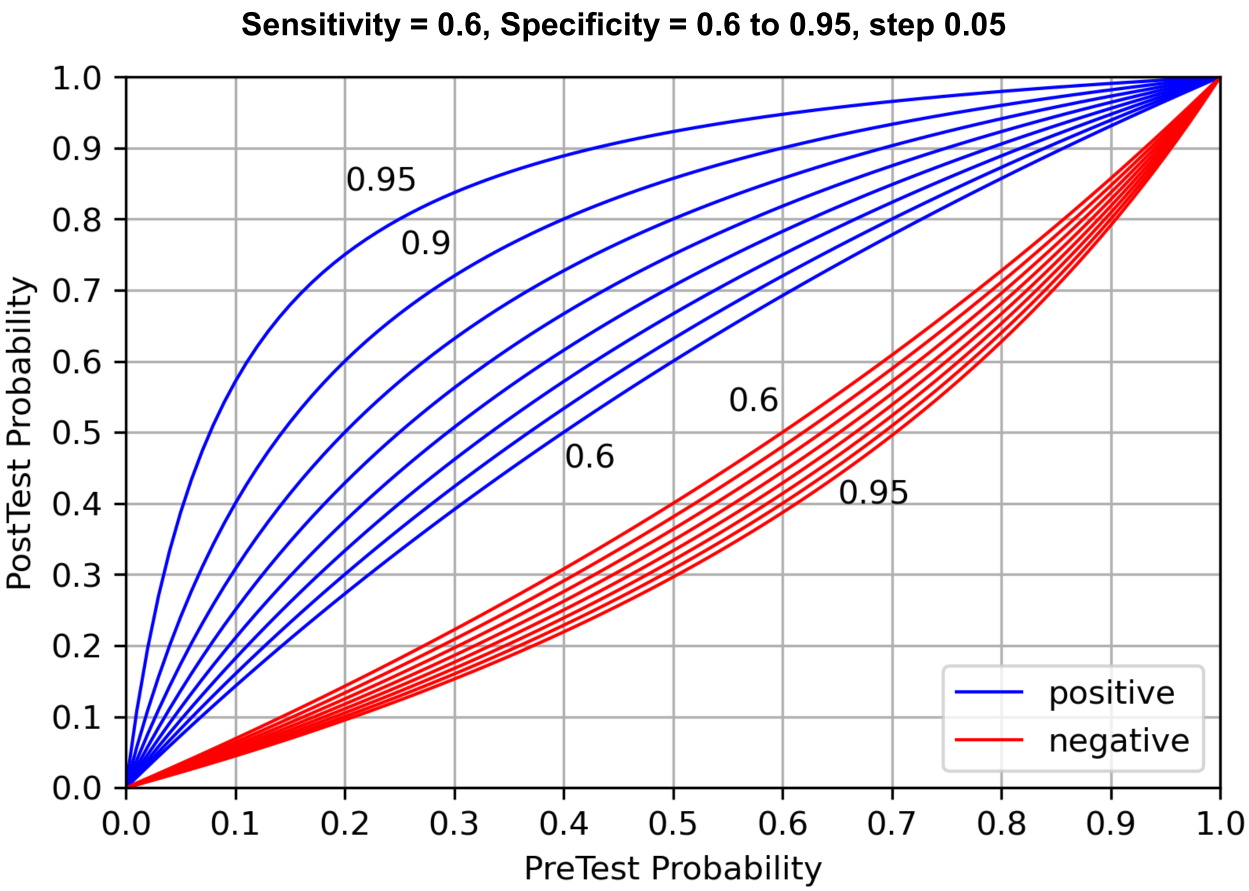


Observations:

- When the test result is positive, the set of response curves in the upper region are all very far apart. Hence for any pretest probability, the posttest probability will *vary a lot* when the test specificity varies from 0.60 to 0.95.

- On the other hand, when the test result is negative, the set of response curves in the lower region are all very near to each other. Hence for any pretest probability, the posttest probability *will not vary* very much when the test specificity varies from 0.60 to 0.95.

- We observe similar results when Test Sensitivity is fixed at 0.6 and the Test Specificity varies from 0.6 to 0.95 in steps of 0.05.

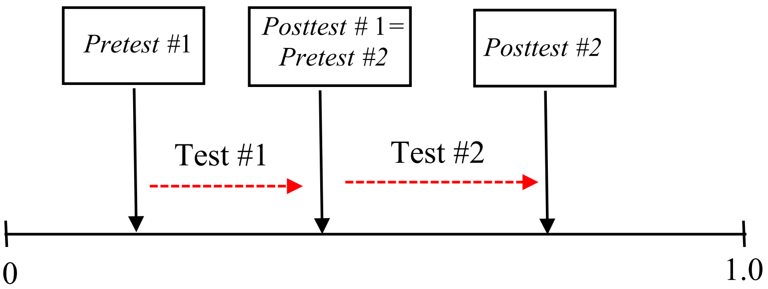


Interpreting a Sequence of Tests

- Diagnostic tests are often used in sequence.
- An abnormal finding on one test may raise concerns that only can be resolved by another test.

- What method should be used to interpret the results of the second test?
  1. Use the posttest probability of the first test as the pretest probability for the second test.
  2. Use the true-positive (sensitivity) and true-negative rate (specificity) of the second test and Bayes' Theorem to calculate the posttest probability for the second test.

- Assumption: The two tests are independent.

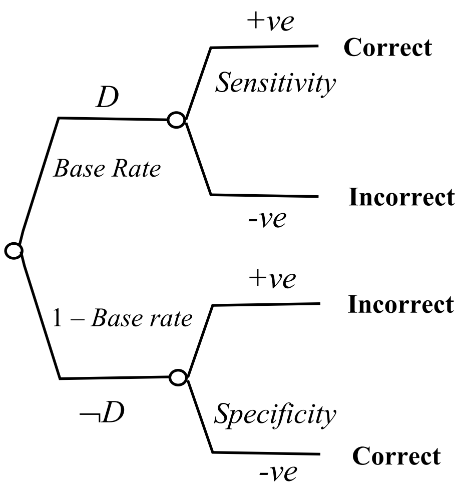


Accuracy of a Test

- The term **Accuracy** should not be confused with the **Sensitivity** (true +*ve* rate) and **Specificity** (true –*ve* rate) of a test.

- **Accuracy = Proportion of cases that are classified correctly.**

- It depends on the base rate of the population being tested, sensitivity and specificity.



- Accuracy = Base rate × Sensitivity + (1 – Base rate) × Specificity

References

1. S.M. Ross, *A first course in Probability*, 10th edition, Pearson Education, 2019.
2. S.M. Ross, *Introduction to Probability Models*, 12th edition, Elsevier, AP 2019.

Exercises

P2.1 (Clement and Reilly 2001, Exercise 7.15 p 283)
Julie Myers, a graduating senior in accounting is preparing for an interview with a Big Eight accounting firm. Before the interview, she sets her chances of eventually getting an offer from this firm at 50%. Then on thinking about her friends who have interviewed and gotten offers from this firm, she realizes that of the people who received offers, 95% had good interviews. On the other hand, of those who did not receive offers, 75% said they had good interviews. If Julie Myers has a good interview, what are her chances of receiving an offer?

P2.2 In the city, there are only two taxicab companies, the Blue and the Green. As you may suppose, the Blue cabs are blue and the Green cabs are green. The Blue Company operates 90% of all cabs in the city and the Green Company operates the rest. One dark evening, a pedestrian is killed by a hit-and-run taxicab.

There is one witness to the accident. In court, the witness's ability to distinguish cab colors in the dark is questioned, so he is tested under conditions similar to those in which the accident occurred. If he is shown a green cab, he says it is green 80% of the time and blue 20% of the time. If he is shown a blue cab, he says it is blue 80% of the time and green 20% of the time.

The judge believes that the test accurately represents the witness's performance at the time of the accident, so the probabilities he assigns to the vents of the accident agree with the figures reported by the test.

(***a***) Construct the probability tree representing the judge's state of information. Label all end-points, supply all branch probabilities, and calculate and label all endpoint probabilities.

(***b***) Flip the tree. Label all endpoints, supply all branch probabilities, and calculate and label all end-point probabilities.

(***c***) If the witness says "The cab involved in the accident was green," what probability should the judge assign to the cab involved in the accident being green?

(***d***) How does the answer to part *c* compare to the witness' accuracy on the test? Does this result seem surprising? Why or why not?

P2.3 Tommy is a contestant on the game show "Let's Make a Deal." Up on stage, there are three boxes, one of which contains a valuable prize; the other two are empty. The rules of the game are that Tommy first chooses one of the boxes. Then Paul, the game show host, opens one of the remaining two boxes, making sure to open an empty one. If both the remaining boxes are empty, then he opens either one at random. Tommy then gets to decide if he wants to stick with his initial selection or switch to the remaining unopened box. If the prize is in the box that he chooses, he wins the prize.

Suppose Tommy has been watching every show for the entire season and he believes that there is a 0.6 chance that the prize will be in Box A, a 0.1 chance that it will be in Box B, and a 0.3 chance that it will be in Box C. What is Tommy's best strategy?

P2.4 Your friend Ella is unsure about her plans for Friday night, although she knows that either John or Peter will ask her to see a movie (either "Cinderella Man" or "Red Eye"). John and Peter are friends, so only one of them will ask her to see a movie. You believe that Peter is twice as likely as John to ask her. If John asks, you believe that they will see "Cinderella Man" with probability 1/4. If Peter asks, you believe they will see "Cinderella Man" with probability 5/8. Suppose you found out later that Ella saw the movie "Red Eye", which of the following statements must be true?

I. Ella has a higher probability of seeing the movie with John than with Peter

II. Ella has a higher probability of seeing the movie with Peter than with John

III. Ella is equally likely to have seen the movie with either John or Peter

IV. Ella ate popcorn at the movies

This page is blank.