MAT116 Project 1 Chapter 6: Bayes' Theorem

Introduction

- In the previous chapter, we found that we could compute the following values:
  - $P(Y \cap T \cap C|S) = P(Y|S) \times P(T|S) \times P(C|S)$
  - $P(Y \cap T \cap C|F) = P(Y|F) \times P(T|F) \times P(C|F)$
- These give you the probability of a specific loan matching all three of your borrower's criteria *given that* a loan workout succeeds or fails.
- However, we want the opposite.

Introduction

- Specifically, we want to know or be able to compute the probability that a loan workout will succeed or fail given that it meets all of our borrower's criteria. That way, we are comparing similar loans to each other during our analysis.

Introduction

- What we want, therefore, is a way to compute the following conditional probabilities.
  - P(S|Y∩T ∩C)
  - P(F|Y∩T ∩C)
- Notice that these are essentially the reverse of what we already know how to compute.
- ***In this chapter, we learn how to reverse a conditional probability. First, however, we some preliminary work to do.***

6-1: Partitions - Definition

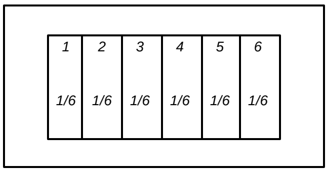
- A set of events form a ***partition*** if and only if ***both*** of the following conditions are true:
  - No two events can happen at the same time
  - All events combined make up the sample space (in other words, one of the events must occur)

6-1: Partitions - Definition

- Mathematically, a partition is defined this way:
- Events $A_1, A_2, A_3, \ldots, A_n$ form a **partition** if and only if:
  - $P(A_i \cap A_j) = 0$, for any i and j, as long as $i \neq j$
  - $P(A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_n) = 1$

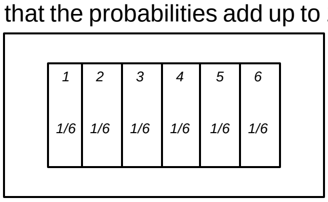
Example

- If you roll a single die, then the sample space has six possible outcomes.
- A diagram shows these outcomes and their probabilities



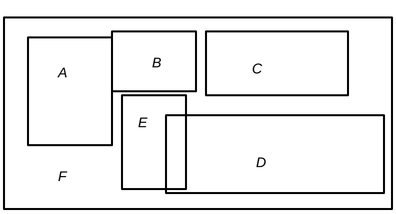
Example

- Notice that there is no overlap in the regions…each event is disjoint.
- Also note that the probabilities add up to 1.



Counter Example

- Why is this NOT a partition?



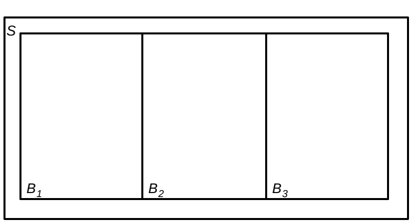
Project Example

- If the sample space is all our bank loans and events are defined by the highest level of education, then this forms a partition of our loans. The set of people whose highest level of education is high school is disjoint from the sets of people whose highest education level is a bachelor's degree or a graduate degree. Also, the probabilities (which ones?) will add up to 1.

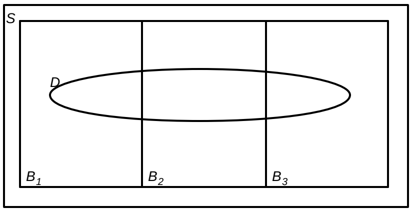
Example

- A company produces 1000 ovens per week at 3 plants. Plant B_1 produces 350 ovens, B_2 produces 250 ovens, and B_3 produces 400 ovens. Let D be the event that an oven is defective. 5% of ovens at B_1 are defective, 3% at B_2 and 7% at B_3.
- What is the sample space?
- Do the B's partition the sample space?
- What role does D play?
- What would a diagram look like?

Example



Example



Example: ?'s to Ponder

- All these 1000 ovens per week are sent to a central warehouse for inspection.
- If one of them is "randomly" chosen and inspected, what is the probability that it is defective ***and*** was made at B_1? (Why is this not 5%?)
- What about for B_2 or B_3?
- What is P(D)?

Example

- What is the ***overall*** probability that it is defective and was made at B_1?
- $P(D \cap B_1)/P(B_1) = P(D|B_1)$ So…
- $P(D \cap B_1) = P(D|B_1) \times P(B_1)$
  $= (0.05)\times(0.35)$
  $= 0.0175$

Example

- What is the ***overall*** probability that it is defective and was made at B_2?
- $P(D \cap B_2) = P(D|B_2) \times P(B_2)$
  $= (0.03) \times (0.25)$
  $= 0.0075$

Example

- What is the ***overall*** probability that it is defective and was made at B_3?
- $P(D \cap B_3) = P(D|B_3) \times P(B_3)$
  $= (0.07) \times (0.40)$
  $= 0.028$

Example

- $P(B_1 \cap D) + P(B_2 \cap D) + P(B_3 \cap D) =$
  $= 0.0175 + 0.0075 + 0.028$
  $= 0.053$
- These obviously do not add to 1.
- To get further, we first develop a visual tool to help us organize our data…

6-2: Tree Diagrams

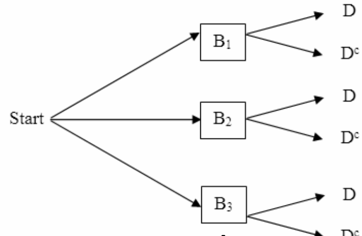
- **A tree diagram is a useful tool in probability because it allows you to "map out" all the possible events. If probabilities of each event are known, then the tree diagram becomes an even more powerful tool.**

Example

- Let's take the previous problem and show it on a diagram.
- A company produces 1000 ovens per week at 3 plants. Plant B_1 produces 350 ovens, B_2 produces 250 ovens, and B_3 produces 400 ovens. Let D be the event that an oven is defective. 5% of ovens at B_1 are defective, 3% at B_2 and 7% at B_3.

Example

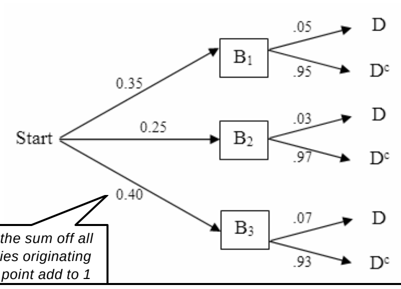
These events are related to conditional probabilities



These events partition the sample space in this example

Example

- We can fill in the appropriate probabilities:

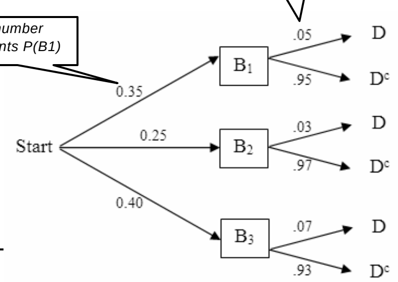


Note that the sum off all probabilities originating from one point add to 1

Example

This number represents P(D|B1).

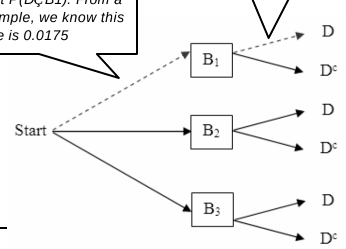
This number represents P(B1)



Example

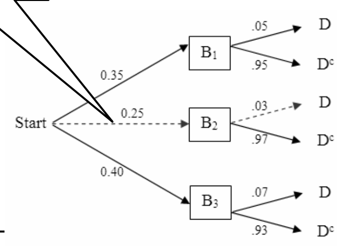
Hence, if you multiply the numbers along a path, you will get the probabilities of two intersecting events.

This path combines P(B1) and P(D|B1). When we multiply these, we get P(DÇB1). From a previous example, we know this value is 0.0175



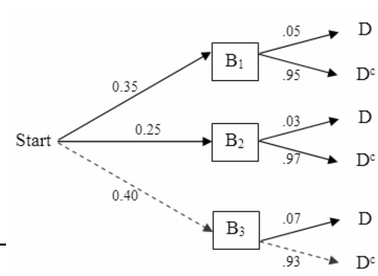
Example

This path give us:
P(DÇB2)=P(D|B2) × P(B2)
=(.03)(.25) = 0.0075



Example

What is the meaning/interpretation of the dotted path? What is the numerical value?



Example

- Cyber Video Games has been running a TV ad for its latest game. As the director of marketing, you want to know how effective the ad is, so you conduct a survey of Video Game players.

	Saw Ad	(Saw Ad)^c	Total
Purchased	1200	2000	3200
No Purchase	3800	43000	46800
Total	5000	45000	50000

Example

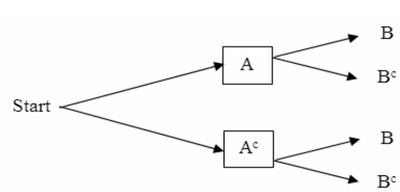
- Draw a tree diagram that gives the relevant probabilities.
  - What is the sample space?
  - What is the partition?
  - What other event(s) is/are there to consider?

	Saw Ad	(Saw Ad)^c	Total
Purchased	1200	2000	3200
No Purchase	3800	43000	46800
Total	5000	45000	50000

Example

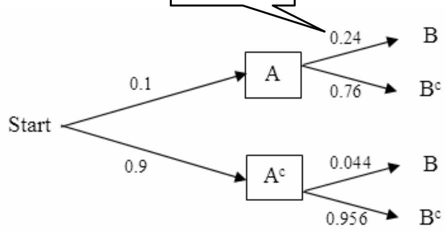
- Let B = player bought/purchased the game
- Let A = player saw the advertisement
- Our Sample Space is partitioned by all the people who did an did not see the ad, event A. (It's also partitioned by those who did and did not buy the game, so there is more than one way to partition this sample space.)
- Since A partitions the sample space, B is some other event. (What would a picture of the sample space look like?)

Example: The base diagram



Example: Filled-In Diagram

Where did these numbers come from???



Example

- Recall these figures:

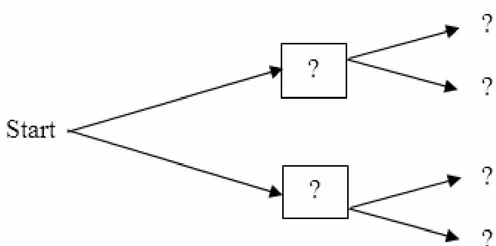
	Saw Ad	$(\text{Saw Ad})^c$	Total
Purchased	1200	2000	3200
No Purchase	3800	43000	46800
Total	5000	45000	50000

Example: Workspace

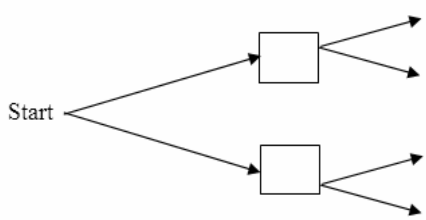
Example

- Gamma Chemicals claims its anabolic steroid drug test is 95% effective, meaning there is a 95% probability of the test being positive on a steroid user. It also states that its test has a low false positive rate of 6%. (This is the probability that a non-steroid user will test positive for use of these drugs.) It is estimated that 30% of all those who take the test actually use steroids. Draw a diagram that shows these events related to each other.

Example



Example: Fill it in



Example

- Which conditional probabilities are given in this situation/diagram?
- P(T|A): Known or unknown?
- What about P(A|T)? Known or unknown?

Computing P(A|T)

$$P(A|T)=\frac{P(A\cap T)}{P(T)}$$
$$=\frac{P(\text{Uses steroids and tests positive})}{P(\text{Tests positive})}$$
$$=\frac{P(\text{Using A and T branches})}{\text{Sum of P(Using branches ending in T)}}$$

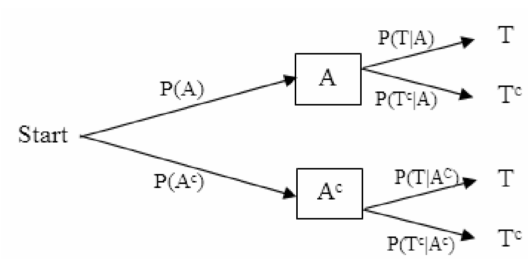
Computing P(A|T)

$$P(A|T)=\frac{(0.3)(0.95)}{0.285+0.042}=\frac{.285}{.327}\approx 0.87$$

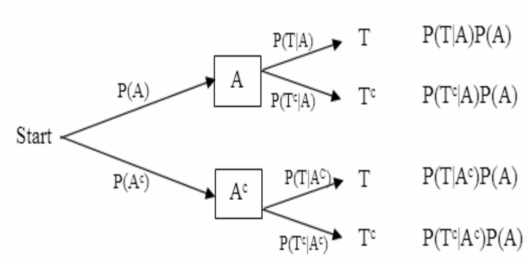
Bayes' Theorem

- The general idea behind the calculation just shown is known as Bayes' Theorem. Let's look at a tree diagram from a more general point of view:

Bayes' Theorem



Bayes and Probabilities



Bayes' Theorem

- We calculated P(A|T), ***the opposite of P(T|A),*** as follows:

$$P(A|T) = \frac{P(A \cap T)}{P(T)}$$
$$= \frac{P(T \mid A)P(A)}{P(T \mid A)P(A) + P(T \mid A^c)P(A^c)}$$

Bayes' Theorem

- 3.) Now we substitute

$$P(A|T) = \frac{P(A \cap T)}{P(T)} = \frac{P(A)P(T \mid A)}{P(T)}$$
$$= \frac{P(T \mid A)P(A)}{P(T \mid A)P(A) + P(T \mid A^c)P(A^c)}$$

Bayes' Theorem: Short Form

- We now have the following:

$$P(A|T) = \frac{P(T \mid A)P(A)}{P(T \mid A)P(A) + P(T \mid A^C)P(A^C)}$$

- This is known as the short form of Bayes' Theorem

Bayes' Theorem

- Using a Formula

$$P(A|T)=\frac{P(T\mid A)P(A)}{P(T\mid A)P(A)+P(T\mid A^C)P(A^C)}$$

Using a Tree

$$P(A|T)=\frac{P(\text{UsingA andT brances})}{\text{Sum of P(Branchesendingin T)}}$$

Bayes' Theorem: Short Form

- Note that P(A|T) and P{T|A) are both in this formula

$$\boxed{P(A|T)}=\frac{\boxed{P(T\mid A)}P(A)}{P(T\mid A)P(A)+P(T\mid A^C)P(A^C)}$$

- What does this mean???????

Bayes' Theorem: Short Form

$$P(A|T)=\frac{P(T\mid A)P(A)}{P(T\mid A)P(A)+P(T\mid A^C)P(A^C)}$$

Remember, the denominator is really just P(T)

Bayes’ Theorem: General Form

- Let $B_1, B_2, \ldots B_n$ be a partition of a sample space. Let A be some other event. Then:
- For any event B_k

$$P(B_K \mid A) = \frac{P(B_K)P(A \mid B_k)}{P(B_1)P(A \mid B_1) + P(B_2)P(A \mid B_2) + \ldots + P(B_n)P(A \mid B_n)}$$

Bayes’ Theorem: General Form

- Using summation notation, we have:

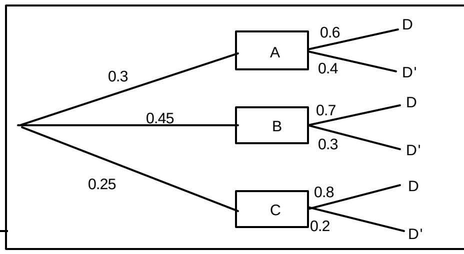
$$P(B_K \mid A) = \frac{P(B_K)P(A \mid B_k)}{\sum_{i=1}^{n} P(B_i)P(A \mid B_i)} = \frac{P(B_K)P(A \mid B_k)}{P(A)}$$

Baye’s Theorem

- Don’t let the convoluted formula trip you up! Just remember these important facts:
  - Baye’s theorem helps you to reverse conditional probabilities
  - If you draw a tree diagram, you can easily isolate only those branches that are relevant to your computation.

Example

- Find P(B|D), P(C|D^c), etc.



Example

- For a fixed length of time, the probability of worker error on a certain production line is 0.1. The probability that an accident will occur when there is a worker error is 0.3 and the probability that an accident will occur when there is no worker error is 0.2 What is the probability of a worker error if there is an accident?
- (See author notes…)

Example Workspace

Example Workspace

Example

- Professor X has divided her class into three categories before the final exam: those likely to pass(20%), those likely to fail (60%) and those of whom she is not sure (the rest). After grading the final, she finds that 80% of those likely to fail got an A, whereas only 10% of the students classifies as likely to pass got an A. 90% of students who she was unsure of also got A's. What fraction of the class who got A's on the final were in her "likely to pass" category?

Example Workspace

Example Workspace

Focus on the Project

- How can Bayes' Theorem help us with the decision on whether or not to attempt a loan work out?
- From part 4 of the project, you have found
- $P(Y \cap T \cap C|S)$ and $P(Y \cap T \cap C|F)$. However, in order to decide whether to work out or not, we need the reverse probability to get the following:
  $P(S|Y \cap T \cap C)$ and $P(F|Y \cap T \cap C)$.
- Bayes' Theorem can be used to calculate these.

Focus on the Project

- The events S and F form a partition. Given S, we already have computed $P(Y \cap T \cap C|S)$. The same is true for $P(Y \cap T \cap C|F)$.

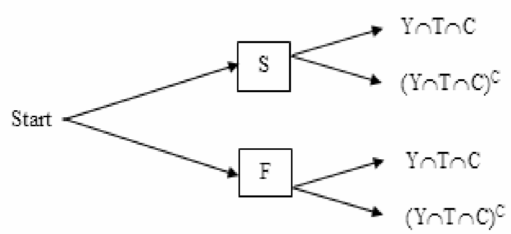
Focus on the Project

- Recall our Class Loan. John Sanders is our borrower. The loan was for $4,000,000. The foreclosure amount was r = $2,100,000, and the default value of the loan was d = $250,000.
- We know that P(S) = 0.464 and P(F) = 0.536. These were computed in Chapter 3
- In Part 4 of the Project Specifics, we found $P(Y \cap T \cap C|S) \approx 0.0220$. We can also compute and verify $P(Y \cap T \cap C|F) \approx 0.0209$.

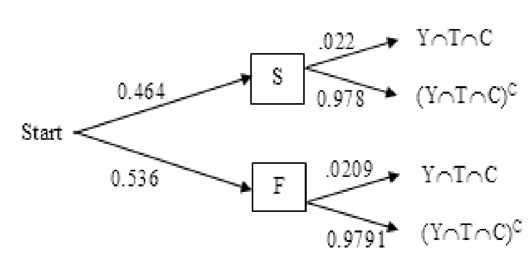
Focus on the Project

- We can now use Bayes' Theorem to reverse this and get $P(S|Y \cap T \cap C)$. Here's a copy of the tree diagram to help us out…

Focus on the Project



Focus on the Project



Focus on the Project

$$P(S \mid Y \cap T \cap C) = \frac{P(Y \cap T \cap C \mid S) \cdot P(S)}{P(Y \cap T \cap C \mid S) \cdot P(S) + P(Y \cap T \cap C \mid F) \cdot P(F)}$$

$$\cong \frac{(0.022) \cdot (0.464)}{(0.022) \cdot (0.464) + (0.021) \cdot (0.536)}$$

As shown in sheet ***Bayes*** of ***Loan Focus.xls***, this yields $\boldsymbol{P(S \mid Y \cap T \cap C)}$ @ **0.477**, when we retain more than three place precision in the component numbers. Likewise $\boldsymbol{P(F \mid Y \cap T \cap C)}$ @ **0.523**,

$$P(F \mid Y \cap T \cap C) = \frac{P(Y \cap T \cap C \mid F) \cdot P(F)}{P(Y \cap T \cap C \mid S) \cdot P(S) + P(Y \cap T \cap C \mid F) \cdot P(F)}$$

$$\cong \frac{(0.021) \cdot (0.536)}{(0.022) \cdot (0.464) + (0.021) \cdot (0.536)}.$$

Focus on the Project

- We are now ready to consider the work out question for the particular borrower, John Sanders. Recall that *Z* is the random variable giving the amount of money, in dollars, that ***Acadia Bank*** receives from a future loan work out attempt to borrowers with the same characteristics as Mr. Sanders, in normal times.

Focus on the Project

$$
\begin{aligned}
E(Z) &= \$4{,}000{,}000 \cdot P(Z = \$4{,}000{,}000) + \$250{,}000 \cdot P(Z = \$250{,}000) \\
&= \$4{,}000{,}000 \cdot P(S \mid Y \cap T \cap C) + \$250{,}000 \cdot P(F \mid Y \cap T \cap C) \\
&\approx \$4{,}000{,}000 \cdot (0.477) + \$250{,}000 \cdot (0.523) \\
&\approx \$2{,}040{,}000
\end{aligned}
$$

What should we do with the loan, based only on this information?

Focus on the Project

- This, however, might be a hasty decision. Let us see if there is any further information hidden in the bank records.
- The records contain only 239 loans where the borrower was known to have exactly 7 years experience in the given type of business. It seems unlikely that there is anything dramatically different about being in business for 7 years, rather than 6 years or 8 years. To make use of more of the bank's data, let Y' be the event that a borrower has 6, 7, or 8 years of experience in the business.

Focus on the Project

- Using the range

Former Bank	Years In Business	Years In Business	Education Level	State Of Economy	Loan Paid Back
BR	>=6	<=8			yes

- in ***DCOUNT*** allows us to find that 349 of the successful work outs at the BR Bank were in Y'. A similar range finds that 323 of the failed work outs at the BR Bank were in Y'. This uses a total of 673 loan records.
- Let Z' be the random variable giving the amount of money, in dollars, that ***Acadia Bank*** receives from a future loan work out attempt to borrowers with Y' and a Bachelor's Degree, in normal times. When all of the calculations are redone, with Y' replacing Y, we find that $P(Y' \cap T \cap C|S) \cong 0.073$ and $P(Y' \cap T \cap C|F) \cong 0.050$.

Focus on the Project

- Bayes' Theorem shows that $P(S|Y' \cap T \cap C) \cong 0.558$ and $P(F|Y' \cap T \cap C) \cong 0.442$. Finally, the expected value of $Z\ell$ is $E(Z\ell) \cong$ \$2,341,000. Since this is above the foreclosure value of \$2,100,000, a loan work out attempt is indicated. All of the relevant numbers and computations, are shown in the sheet ***Bayes*** of ***Loan Focus.xls***.
- ***Computations with exactly 7 years experience indicated foreclosure, while work with a range of 6, 7, or 8 years experience indicated a work out.*** Which one should we believe? The best plan is to gather more information from the bank records.
- A wider range of 5, 6, 7, 8, or 9 years experience produces an expected value of \$2,272,000 for a work out attempt. Finally, experience of either exactly 6 years or exactly 8 years leads to expected values that are greater than \$2,100,00.

Focus on the Project

- This would be a close call for ***Acadia Bank*** loan officers. A reasonable conclusion is that the scarcity of data for individual years of business experience, leads to artificial fluctuation in the expected value of a loan work out. It is quite likely that the low expected value for exactly 7 years is an anomaly of the data, rather than an indication of a poor loan risk for John Sanders. ***Based upon all of our calculations, we recommend that Acadia Bank enter into a work out arrangement with Mr. Sanders.***

Focus on the Project

- Now do the same calculations for your own Team Data.
- Draw a tree diagram with S and F as the main partition and fill in the appropriate probabilities. Include this in your final project's written report.
- CHALLENGE: Build a spreadsheet that will allow you to enter in Y, T, C and all the other loan information and which will then make a loan decision for you. This is optional and would require a LOT of work. I've done it, however, so you can stop by my office if you want to see how one might work.