ECONOMICS 150: Quantitative Methods for Economics

Class 4

- Conditional Probability
- Independence
- Probability Trees

Announcement

- Online Test 4 is due before next class

Practice Problem Set 1, #4 (stereo equipment)

The owner of a stereo equipment shop has kept careful records on sales. He has found that once a person enters the store and asks about stereo equipment, there is a .15 probability that a receiver will be purchased, a .10 probability that speakers will be purchased, and a .20 probability that either speakers or a receiver, or both, will be purchased.

UNION

- If a customer enters the store and asks about stereo equipment, what is the probability that the customer does not purchase a receiver and purchases speakers?

INTERSECTION

- Always start by making a list of what you know.

Practice Problem Set 1, #4 (stereo equipment)

buy Receiver $\to R$ $\quad P(R)=0.15$

buy Speakers $\to S$ $\quad P(S)=0.10$

buy either $\to R\cup S$ $\quad \boxed{P(R\cup S)=0.2}$

WANT $P(R^c\cap S)$

$R^c\cap S^c$ is complement of $R\cup S$

$P(R^c\cap S^c)=0.8$

$P(R^c\cap S)=0.85-0.8=0.05$

	R	R^c	
S		0.05	0.1
S^c		0.80	0.9
	0.15	0.85	1

Practice Problem Set 1, #4 (stereo equipment)

a) If speakers are purchased, what is the probability that a receiver will not be purchased?

$R \quad = receiver\ purchased$ $\qquad$ $\mathrm{P}(R) \quad = 0.15$

$R^C = receiver\ not\ purchased$ $\qquad$ $\mathrm{P}(R^C) = 0.85$

$S \quad = speakers\ purchased$ $\qquad$ $\mathrm{P}(S) \quad = 0.10$

$S^C = speakers\ not\ purchased$ $\qquad$ $\mathrm{P}(S^C) = 0.90$

The probability that either a receive or a speaker is purchased is $\mathrm{P}(R \cup S) = 0.20$

We want to find $\mathrm{P}(R^C \cap S)$

	R	R^C	
S	$\mathrm{P}(R \cap S)$	$\mathrm{P}(R^C \cap S)$	0.10
S^C	$\mathrm{P}(R \cap S^C)$	$\mathrm{P}(R^C \cap S^C)$	$\mathrm{P}(S^C)$
	0.15	$\mathrm{P}(R^C)$	1

The shaded areas represent $\mathrm{P}(R \cup S) = 0.20$

Hence $\mathrm{P}(R^C \cap S^C) = 1 - \mathrm{P}(R \cup S) = 0.80$

Since: $\mathrm{P}(R^C \cap S) + \mathrm{P}(R^C \cap S^C) = \mathrm{P}(R^C)$

We can solve $\mathrm{P}(R^C \cap S) = 0.85 - 0.80 = 0.05$

Top Hat Attendance Code for Today

7808

Conditional Probability

Probability valuations might change when new information becomes available

- This is one of the main ideas in all of statistics
  - Same flavor of the Walmart example from Class 1, where what we thought changed because new data became available

An example:

- What is the probability that a roll is even if you know it is high? Is it the same as the probability you assigned to even before having that information?
  - A={even roll} ={2,4,6} thus P(A)= ½
  - B={high roll}={4,5,6}
- We need a new symbol for this new way of "combining" events

Conditional Probability

Probability valuations might change when new information becomes available

- This is one of the main ideas in all of statistics
  - Same flavor of the Walmart example from Class 1, where what we thought changed because new data became available

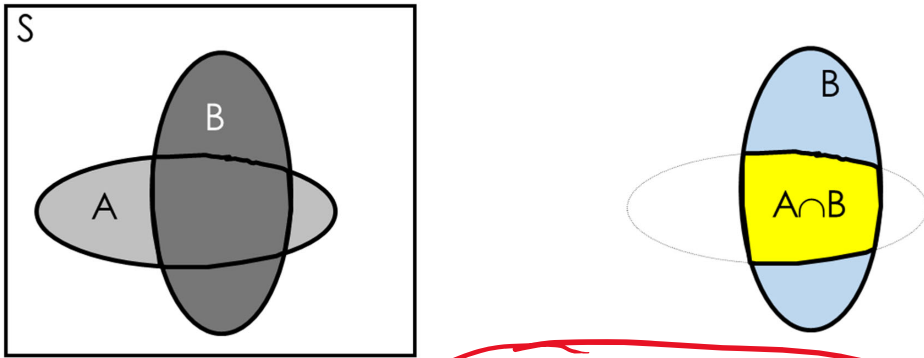
Notation:

$$P(A|B) = probability\ of\ A\ given\ that\ B\ occurred$$

We also say "probability of A if B happened" or some similar wording

Conditional Probability Formula

$P(A|B) = probability\ of\ A\ given\ that\ B\ occurred$



Conditional Probability Formula

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

What is the probability that a roll is even if you know it is high?

Is it the same as the probability you assigned to even before having that information? NO

even roll = {2,4,6} thus P(even roll)= ½

high roll = {4,5,6}

$$P(Even \mid High) = \frac{P(Even \cap High)}{P(High)} = \frac{1/3}{1/2} = \frac{2}{3}$$

$$Even \cap High = \{4,6\} \rightarrow P(Even \cap High) = \frac{2}{6} = \frac{1}{3}$$

From Online Quiz 3

The weather on Christmas day in Milan (Italy) has been has follows

Clear and dry with probability 0.20 | Cloudy and dry with probability 0.30

Rain with probability 0.40 | Snow with probability 0.10

1. What is the probability that next Christmas will be dry?

$P(DRY) = P(DRY \cap Clear) + P(Dry \cap Cloudy) = 0.2 + 0.3 = 0.5$

2. What is the probability that next Christmas will be rainy or cloudy and dry?

$P(Rainy \cup (Cloudy \cap Dry)) = 0.4 + 0.3 + 0 = 0.7$

$Rainy \cap (Cloudy \cap Dry) = \emptyset$

3. Suppose next Christmas is dry, determine the probability that it will also be cloudy.

$$P(Cloudy \mid Dry) = \frac{P(Cloudy \cap Dry)}{P(Dry)} = \frac{0.3}{0.5} = \frac{3}{5} = 0.6$$

Conditional Probability Is Not Intuitive

- First, the event one "starts from" matters
  - The conditional probability will reflect that
- Second, conditional probability does not commute
  - The probability of A given B is not the probability of B given A

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \neq P(B|A) = \frac{P(A \cap B)}{P(A)}$$

Numerator is the same, but denominator is different

(Back to) Walmart ads on FB

Walmart contracts 1 billion additional impressions.

UNTARGETED ADVERTISING

- Provide a simple forecast for the number of clicks generated by them.

Remember, from last class, P(Click) = 0.02

$0.02 \cdot 1{,}000{,}000{,}000 = 20{,}000{,}000$

- If Walmart's campaign were to target only women above 24, would you revise the above forecast?

need a conditional probability : $P(\text{Click} \mid F \geq 24)$

	Click	No Click	
Any gender, age ≤ 24	0.0034	0.3366	0.34
Female > 24	0.0153	0.3847	0.4
Not Female >24	0.0013	0.2587	0.26
	0.02	0.98	1

$$P(\text{click} \mid F > 24) = \frac{P(\text{click} \cap F > 24)}{P(F > 24)} = \frac{0.0153}{0.4}$$

$$= 0.038$$

(Back to) Walmart ads on FB

Walmart contracts 1 billion additional impressions.

- Provide a simple forecast for the number of clicks generated by them.

P(Click) = 0.02

20,000,000

Forecast: 0.02*1,000,000,000 = ~~2,000,000~~

- If Walmart's campaign were to target only women above 24, would you revise the above forecast? YES

38,000,000

Yes! New Foreacst: 0.038*1,000,000,000 = ~~3,800,000~~

- Based on these calculations, would you recommend Walmart to implement such a targeted campaign instead?

(Statistical) Independence

Two events are (statistically) **independent** if the information that one of them has happened does **not** change the probability of the other

- This is a strange definition: it only refers to probabilities, so it is not about "logical" independence
  - In math: A and B **are statistically independent** whenever $P(A|B)=P(A)$
    - For example: with die roll, are $A=\{1,3,5\}$ and $B=\{3,6\}$ independent?
    - There is only one way to find out: do the math (next slide)

Are A={1,3,5} and B={3,6} independent?

$P(A) = \frac{1}{2}$

$P(B) = \frac{2}{6} = \frac{1}{3}$

$A \cap B = \{3\}$

$P(A \cap B) = 1/6$

$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/6}{1/3} = \frac{1}{2}$

INDEPENDENT!

Recognizing Independence

There are three ways to recognize two independent events:

- Events A and B are independent if any of the following holds:

$$\boldsymbol{P(A|B) = P(A)}$$
$$\boldsymbol{P(B|A) = P(B)}$$
$$\boldsymbol{P(A \cap B) = P(A) * P(B)}$$

  - If any of the above equalities is true, then all three are true (can you show this with math?)

Statistical Independence

- Statistical independence is not necessarily intuitive, and it is a very hard concept to fully grasp (one of the hardest in this class)

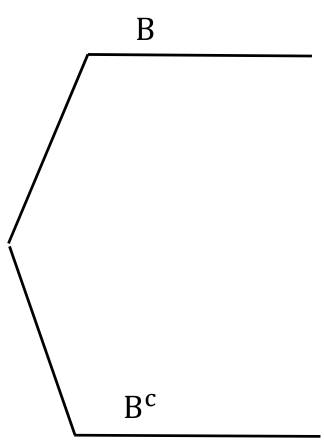
The main idea is about learning from information:

- When events are not independent, learning one of them has happened changes the likelihood that the other may happen;
- When events are independent, learning one of them has happened does not change the likelihood that the other may happen.

- Independence is about processing new information, and how that information may or may not change what one thinks.

Probability Trees

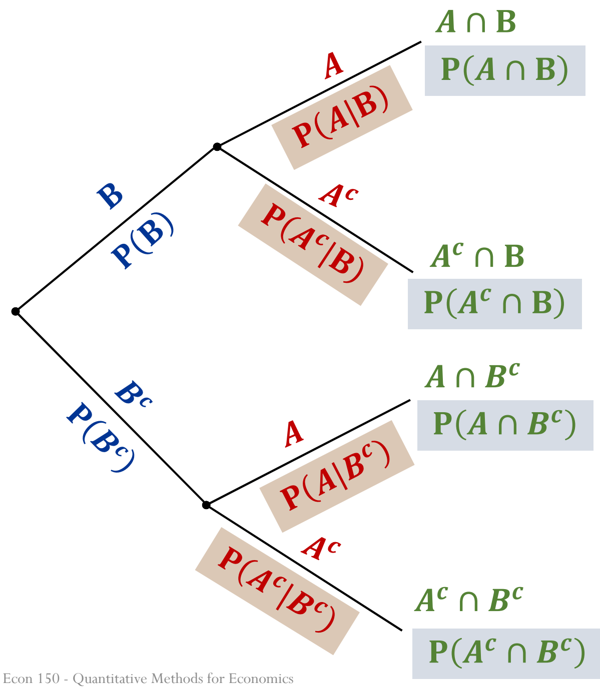
Probability trees are a way to organize information about probabilities that is useful when one has information about conditional probabilities



- Start by branching over a set of mutually exclusive and collectively exhaustive events.
- Each initial branch corresponds to one such event (there could be more than two branches in total)
- A good rule of thumb: You should know the probabilities of these events (otherwise, the probability tree might not help you much)

Continue by branching each of the formed branches…

Probability Tree



In a probability tree, the probabilities "multiply through" because of the conditional probability formula

$$P(A \cap B) = P(B)P(A|B)$$

because $P(A|B) = \frac{P(A \cap B)}{P(B)}$

Practice Problem Set 2, #3 (anonymous drug use questionnaire)

Flip a coin: If the coin comes up heads, answer the first question.
If the coin comes up tails, answer the second question.

1. Do you carpool to work? 2. Have you used illegal drugs within the last month?

Out of 8000 responses, 1420 are marked "YES". The firm knows that 35% of its employees carpool

Assuming everyone tells the truth, and the coin is fair, what is the probability that an employee used illegal drugs within the last month?

Start by making a list of what you know:

I is the event "answered question 1"

II is the event "answered question 2"

Y is the event "answered YES"

We know: $P(Y) = 1420/8000 = 0.1775$

$P(Y|I) = 0.35$

$P(I) = 0.5$

We want P(Y|I)

Practice Problem Set 2, #3 (anonymous drug use questionnaire)

What is the probability that an employee used illegal drugs within the last month?

We want P(Y|II)

We know: P(Y) = 1420/8000 = 0.1775
P(Y|I) = 0.35
P(I) = 0.5   P(II) = 0.5

From last class we know: $P(Y) = P(Y \cap I) + P(Y \cap II)$

Using the conditional probability formula, we can rewrite this expression as

$$P(Y) = P(Y|I)\,P(I) + P(Y|II)\,P(II)$$

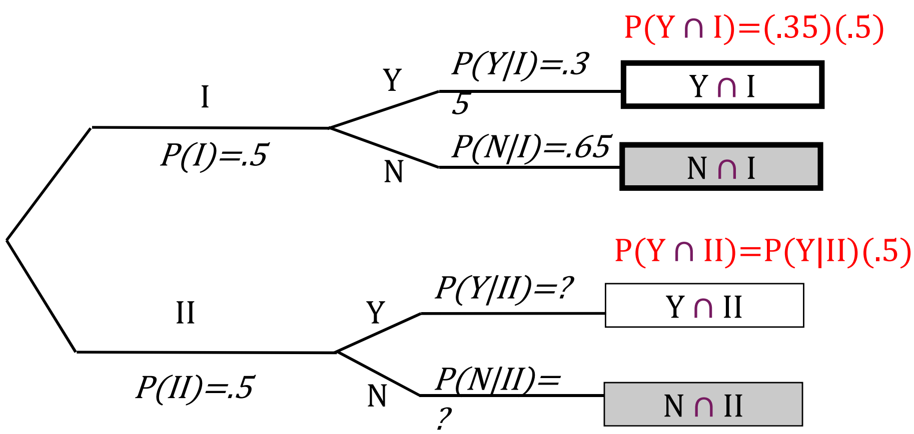
$$0.1775 = (.35)\ (.5) + P(Y|II)\ (.5)$$

Solving, we have: $P(Y|II) = .005$

Practice Problem Set 2, #3 (anonymous drug use questionnaire)

What is the probability that an employee used illegal drugs within the last month?

Want P(Y|II)



$$P(Y) = P(Y|I)\,P(I) + P(Y|II)\,P(II)$$

Disease Testing Preview

- Test an individual from a "low risk" group for a disease: 1 in 500 chances they are sick
- TEST outcomes: POSITIVE or NEGATIVE
- Tests parameters: the test is correct 99% of the time on a sick person
  (1% false negative)
  the test is correct 95% of the time on an healthy person
  (5% false alarm)
- Suppose the TEST outcome is POSITIVE

Question: ***what are the chances the individual is sick?***

Class 3: useful facts

- Conditional probability formula:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

- Two events A and B are independent if any of the following holds:

$$P(A|B) = P(A)$$
$$P(B|A) = P(B)$$
$$P(A \cap B) = P(A)P(B)$$

- Problem solving:
  - Probability trees (useful when you have conditional probability information)

Next time

- Elisa test reading discussion
- Random variables