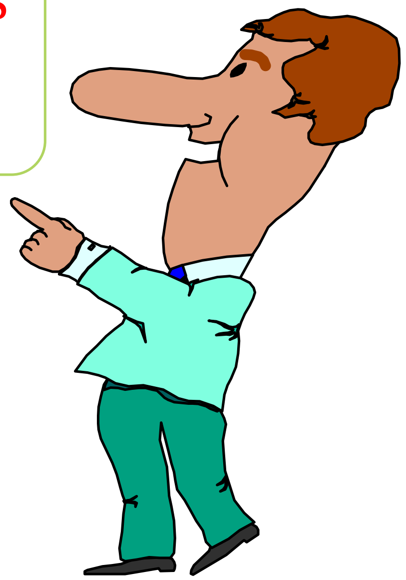
Week 2
Probability

Events, Sample Spaces, and Probability

Experiments & Sample Spaces

- **Experiment**
  - Process of obtaining an observation, outcome or simple event
- **Sample point**
  - Most basic outcome of an experiment
- **Sample space (S)**
  - Collection of ***all*** possible outcomes

Sample space depends on experimenter!



Sample Space Properties

Experiment: Observe gender

- **Mutually Exclusive**

  2 outcomes cannot occur at the same time

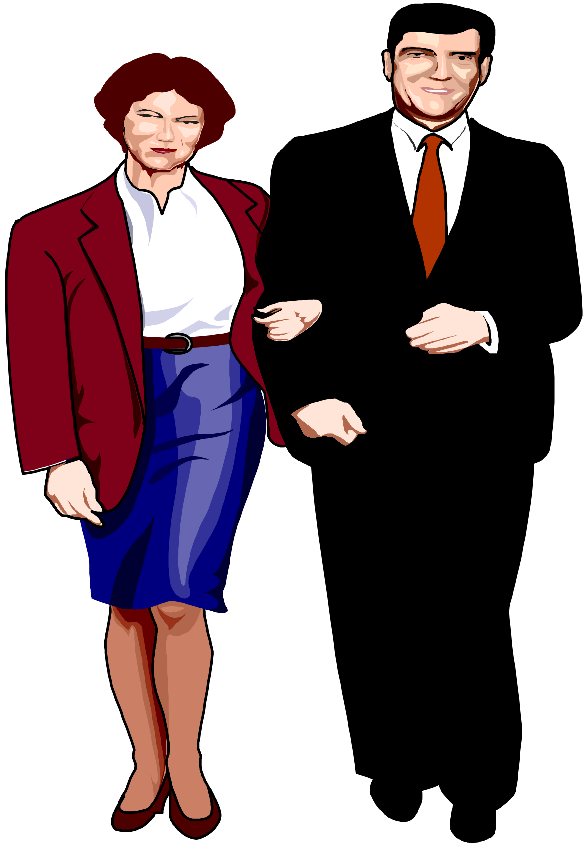
  - Cannot be male and female in same person

- **Collectively Exhaustive**

  One outcome in sample space must occur.

  The set of events covers the entire sample space

  - Male or female



Sample Space Examples

Experiment	Sample Space
• Toss a coin, note face	{Head, Tail}
• Toss 2 coins, note faces	{HH, HT, TH, TT}
• Play a football game	{Win, Lose, Tie}
• Inspect a part, note quality	{Defective, Good}
• Observe gender	{Male, Female}

Events

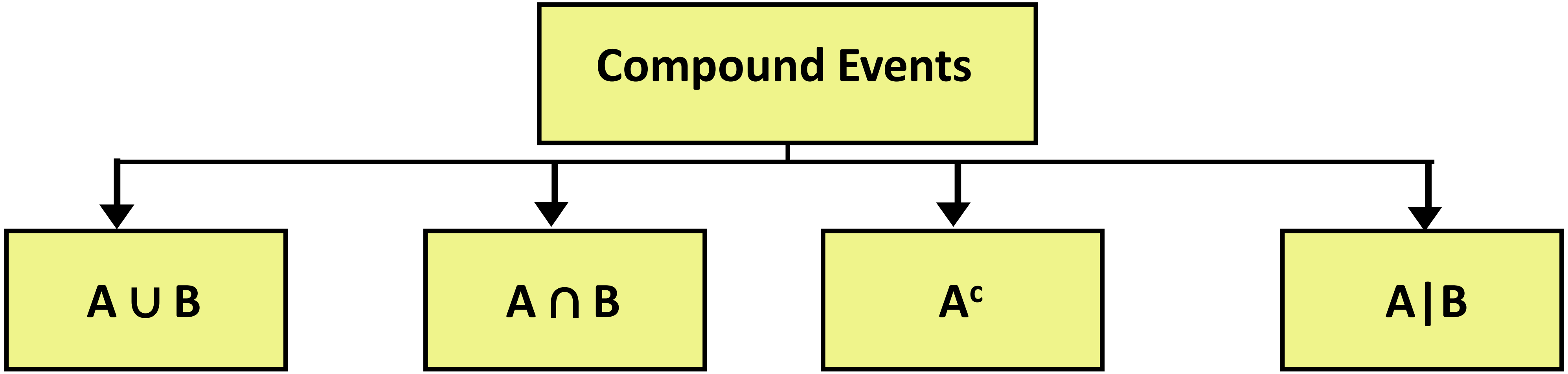
- Any collection of sample points
- **Simple event**:
  - outcome with one characteristic
- **Compound event**:
  - collection of outcomes or simple events
  - 2 or more characteristics
  - joint event is a special case: two events occurring simultaneously

Experiment: Toss 2 coins. Note faces.

Sample space: {HH, HT, TH, TT}

Event	**Outcomes in Event**
1 Head & 1 Tail	HT, TH
Head on 1st Coin	HH, HT
At Least 1 Head	HH, HT, TH
Heads on Both	HH

Compound (or multiple) Events
- occur when 2 or more experiments are conducted together.



Union

- Outcomes in either events A ***or*** B ***or*** both
- ‘OR’ statement
- ∪ symbol (i.e., **$A \cup B$**)

Intersection

- Outcomes in both events A ***and*** B
- ‘AND’ statement
- ∩ symbol (i.e., **$A \cap B$**)

Complement of event A

- All events that are not part of event A.
- All events not in A: **A^C**

Conditional event

- Event A occurs **given that** event B (on which it depends) has occurred; i.e., $A \mid B$.

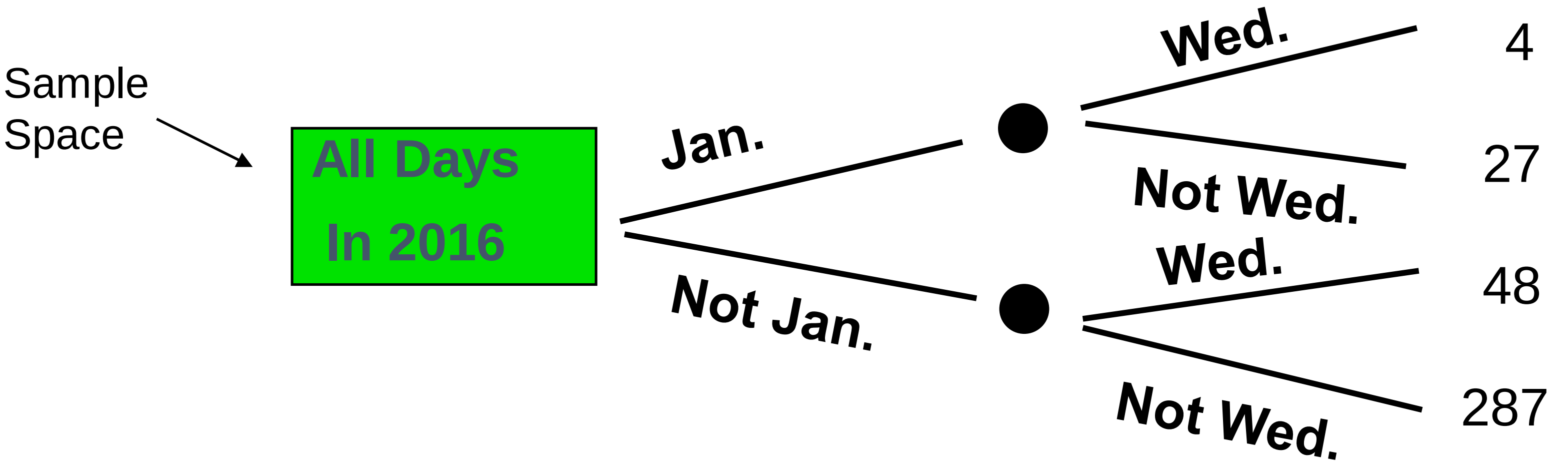
Organizing & Visualizing Events

- Contingency Tables – For all days in 2016.

	Jan.	Not Jan.	Total
Wed.	**4**	**48**	**52**
Not Wed.	**27**	**287**	**314**
Total	**31**	**335**	**366**

Total number of sample space outcomes.

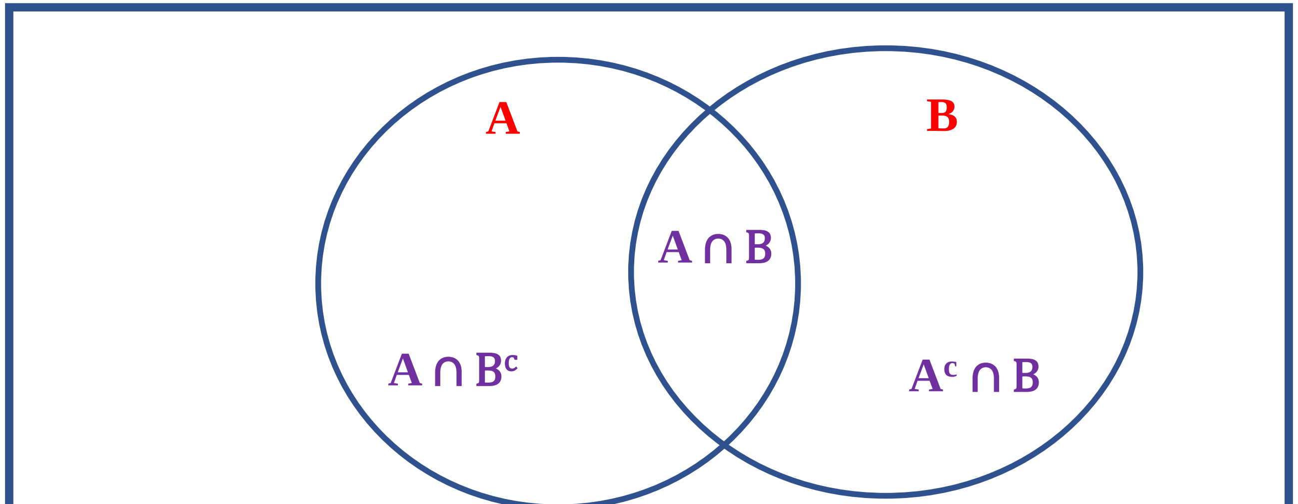
- Decision Trees.



Organizing & Visualizing Events – cont'd

- Venn Diagram

A Venn diagram is a picture that represents the universe of all possible outcomes (i.e., sample space) as a rectangle with events being indicated inside, often as circles or ovals. It provides an effective display of the meanings of the operations **not** (^{c}), or ($\cup$), & **and** ($\cap$).



$A = (A \cap B) + P(A \cap B^c)$ $B = (B \cap A) + P(B \cap A^c)$

DeMorgan's Law:

$P(A^c \cup B^c) = P(A \cap B)^c = 1 - P(A \cap B)$

$P(A^c \cap B^c) = P(A \cup B)^c = 1 - P(A \cup B)$

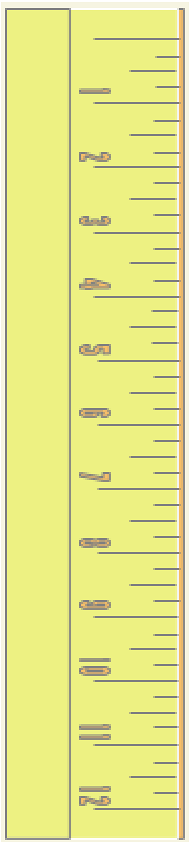
What is Probability?

- Numerical measure of the likelihood that event will occur
  - P(Event)
  - P(A)
  - Prob(A)
- Lies between 0 & 1
- Sum of sample points is 1

$$\text{Probability of an event happening} = \frac{\text{Number of ways it can happen}}{\text{Total number of outcomes}}$$

1 Certain

0.5



0 Impossible

Two types of Probability

There are two main types of probabilities that we will examine in this course:

- 1. Data-based probability
- 2. Model-based probability

Data-based (frequentist) probability

- This is the classical interpretation of probability.
- The probability of an event is the proportion of times that event would occur in a large number of repeated experiments (simulation).
- In some ways, subjective probability is an informal version of frequentist probability, using one's own history and experience ('personal data') as the basis
- E.g. Tossing a coin 10,000 times and calculate the frequency of Head.

Model-based (theoretical) probability

- The probability of an event is based on a model of the context. This is a mathematical construction, assigning a number to every possible event.
- 2 models for coin tossing:
  - Deterministic physics model: the side that the coin lands on is determined by a number of complicated factors such as which way up it started, the degree of spin, the speed and angle with which it left the thumb and how far it has to fall.
  - Statistical probability model: toss a fair coin results in a head and a tail half the time, in an unpredictable order (random).
- E.g. Tossing a fair coin: given a coin toss has 2 outcomes, the probability of tossing a fair head = 0.5

Basic Rules of Probability

- **Complement rule**

  Let A^c (or A') denote the event which consists of all the elements of the sample space that are **not contained in** A.

  $P(A) = 1 - P(A^c)$

- **Addition rule – to find the probability of the union of events**

  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

- **Multiplication rule – to find the probability of the intersection events**

  $P(A \cap B) = P(A|\mathbf{B}) * P(\mathbf{B})$ **Joint probability = conditional probability * P(condition)**

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

  where the conditional probability $P(A|\mathbf{B}) = \frac{P(A \cap B)}{P(\mathbf{B})}$ stands for the probability that A will occur **given that** B has already occurred.

  Similiary, $P(B|\mathbf{A}) = \frac{P(A \cap B)}{P(\mathbf{A})}$ stands for the probability that B will occur **given that** A has already occurred.

  Conditional probability = Joint probability ÷ P(condition)

Checking Mutually Exclusive and Independent Events

- **Mutually exclusive events**

  Events A and B are **mutually exclusive** if one of the following conditions holds:

  - $P(A \cap B) = 0$
  - $P(A \cup B) = P(A) + P(B)$

- **Independent events**

  Events A and B are **independent** if one of the following conditions holds:

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

  conditional probability = unconditional probability

Marginal & Joint Probabilities in a Contingency Table

Event	Event B_1	B_2	Total
A_1	$P(A_1 \cap B_1)$	$P(A_1 \cap B_2)$	$P(A_1)$
A_2	$P(A_2 \cap B_1)$	$P(A_2 \cap B_2)$	$P(A_2)$
Total	$P(B_1)$	$P(B_2)$	1

Joint Probabilities

Marginal (Simple) Probabilities

Example

Of the cars on a used car lot, **70%** have air conditioning (A) and **40%** have a GPS (G), and **20%** of the cars have both. Find $P(A \cup G)$, $P(A^c \cap G^c)$, and $P(G \mid A)$.

	G	G^c	
A	0.2	0.5	0.7
A^c	0.2	0.1	0.3
	0.4	0.6	1.0

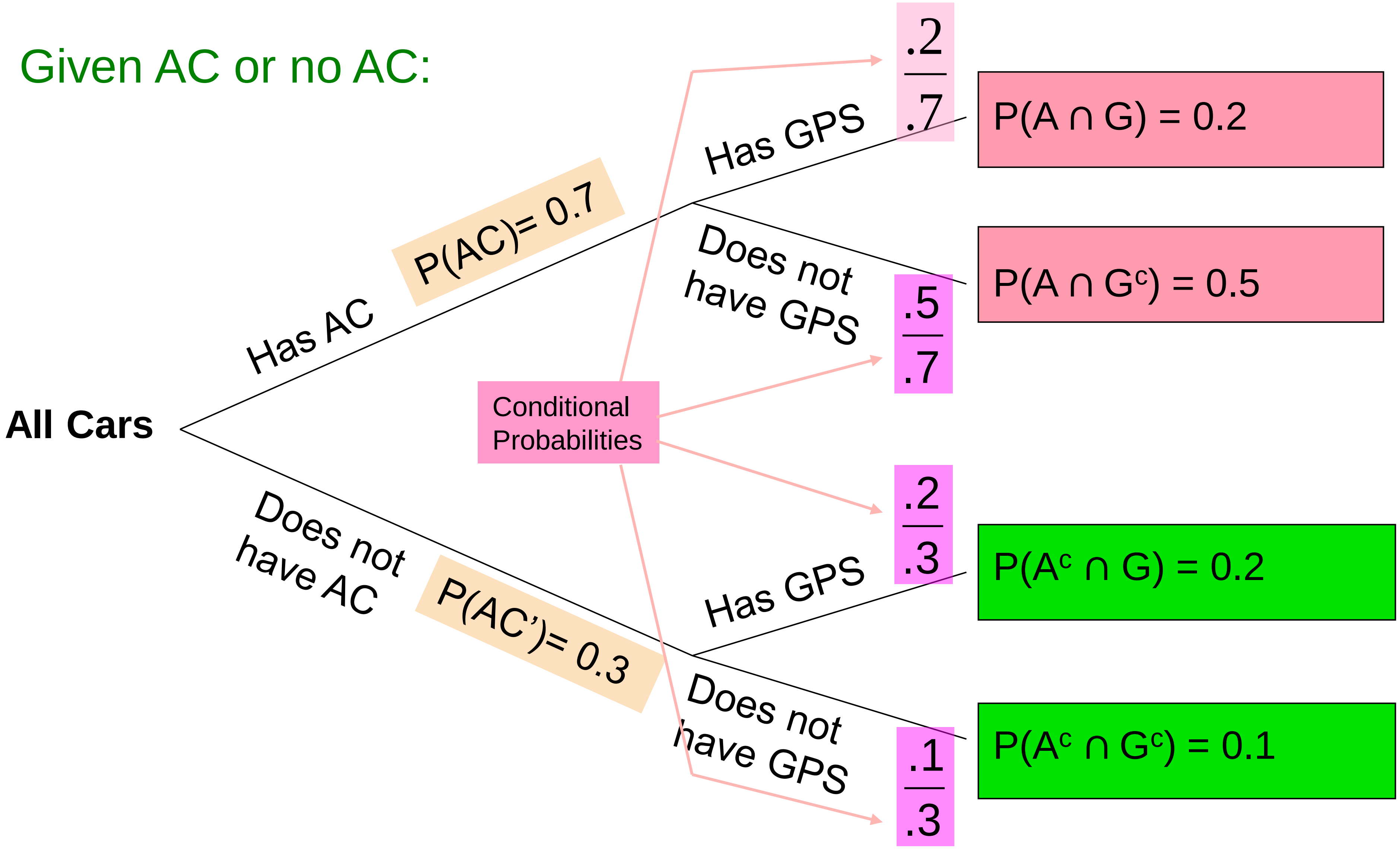
$$P(A \cup G) = P(A) + P(G) - P(A \cap G) = 0.7 + 0.4 - 0.2 = 0.9$$

$$P(A^c \cap G^c) = P(A \cup G)^c = 1 - P(A \cup G) = 1 - 0.9 = 0.1$$

$$P(G \mid A) = \frac{P(G \cap A)}{P(A)} = \frac{0.2}{0.7} = 0.2857$$

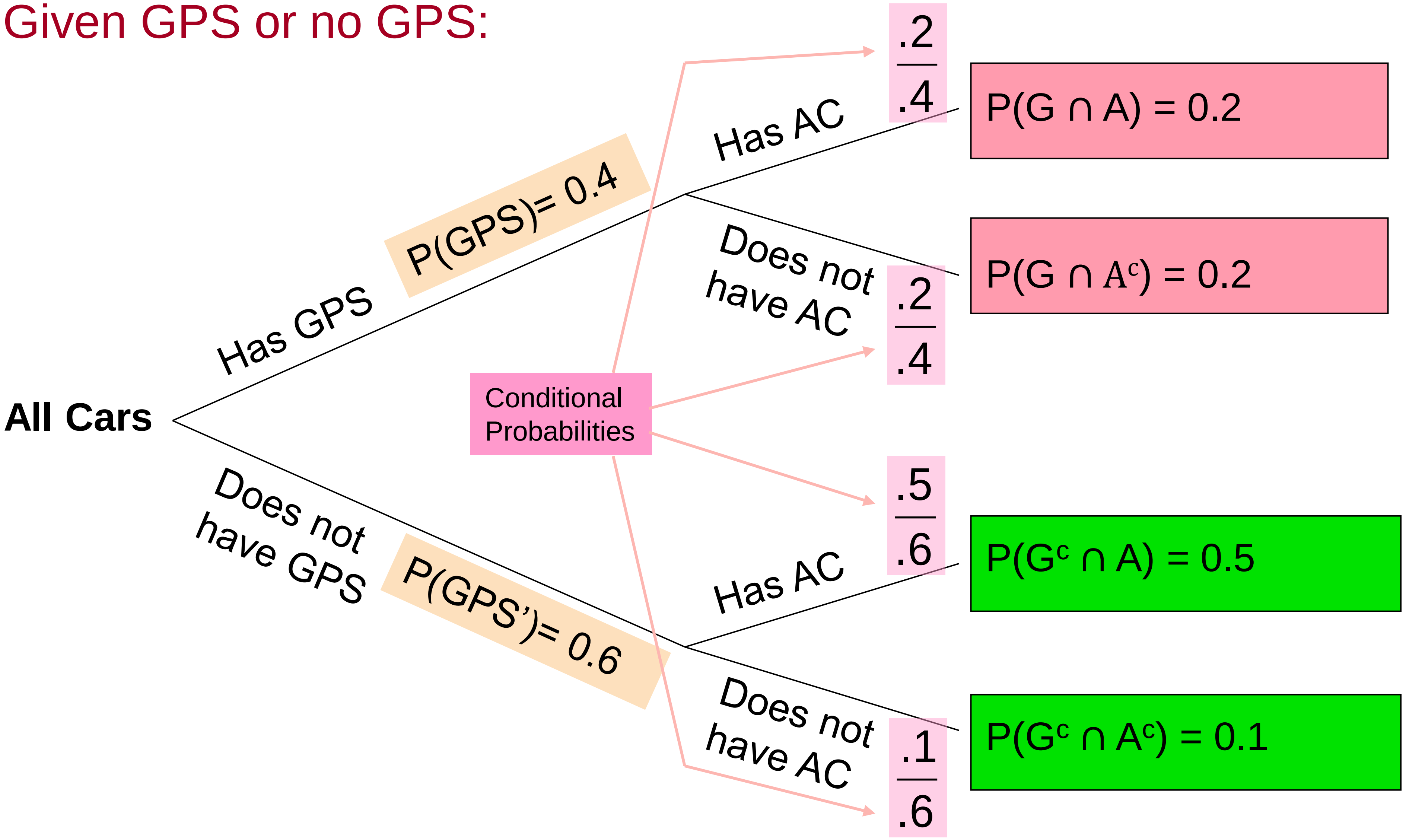
Using Decision Trees

Given AC or no AC:



Using Decision Trees

Given GPS or no GPS:



Bayes' Theorem

- Bayes' Theorem is used to revise previously calculated probabilities based on new information.
- Developed by Thomas Bayes in the 18th Century.
- It is an extension of conditional probability.

$$P(B_i \mid A) = \frac{P(A \mid B_i)P(B_i)}{P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2) + \cdots + P(A \mid B_k)P(B_k)}$$

where

B_i = i^{th} event of k mutually exclusive and collectively exhaustive events

A = new event that might impact $P(B_i)$

The Law of Total Probability: $\boldsymbol{\sum_i P(A|B_i)P(B_i)}$

Example

A drilling company has estimated a 40% chance of striking oil for their new well. A detailed test has been scheduled for more information. Historically, 60% of successful wells have had detailed tests, and 20% of unsuccessful wells have had detailed tests. Given that this well has been scheduled for a detailed test, what is the probability that the well will be successful?

Let S = successful well and D = detailed test

$P(S) = 0.4$, $P(S^c) = 0.6$ (prior probabilities)

$P(D \mid S) = 0.6$, $P(D \mid S^c) = 0.2$

$$\begin{aligned} P(S \mid D) &= \frac{P(D \mid S)P(S)}{P(D \mid S)P(S) + P(D \mid S^c)P(S^c)} \\ &= \frac{(0.6)(0.4)}{(0.6)(0.4) + (0.2)(0.6)} \\ &= \frac{0.24}{0.24 + 0.12} = 0.667 \end{aligned}$$

So, the revised probability of success, given that this well has been scheduled for a detailed test, is 0.667.

Example

Given the detailed test, the revised probability of a successful well has risen to 0.667 from the original estimate of 0.4.

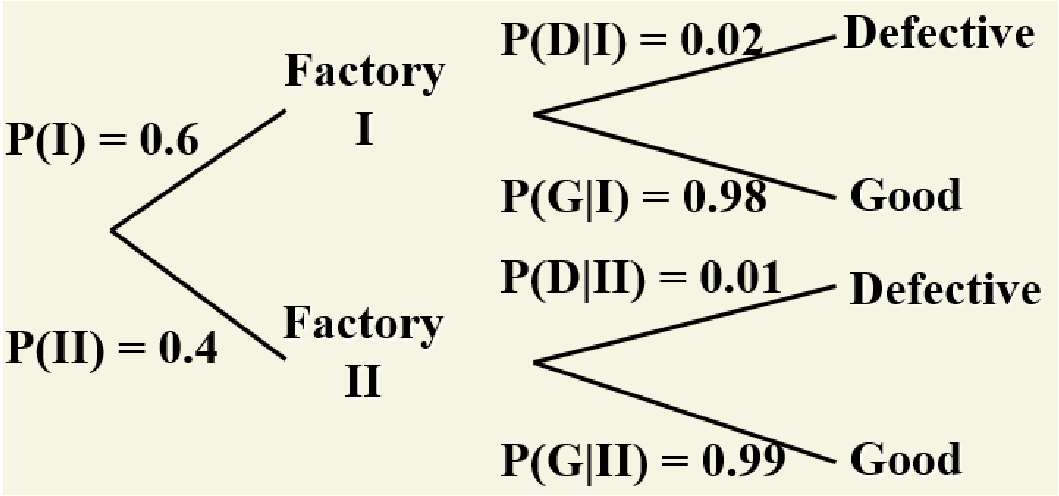
Event	Prior Prob.	Conditional Prob.	Joint Prob.	Revised Prob.
S (successful)	0.4	0.6	(0.4)(0.6) = 0.24	0.24/0.36 = 0.667
S^c (unsuccessful)	0.6	0.2	(0.6)(0.2) = 0.12	0.12/0.36 = 0.333

P(D) = 0.36

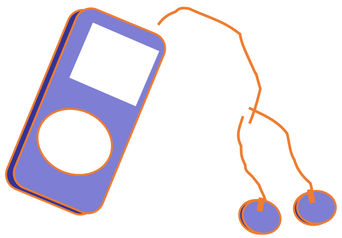
Example

A company manufactures mp3 players at two factories. Factory I produces 60% of the mp3 players and Factory II produces 40%. Two percent of the mp3 players produced at Factory I are defective, while 1% of Factory II's are defective. An mp3 player is selected at random and found to be defective. What is the probability it came from Factory I?

$P(I) = 0.6$, $P(II) = 0.4$, $P(D \mid I) = 0.02$, $P(D \mid II) = 0.01$



$$P(I \mid D) = \frac{P(D \mid I)P(I)}{P(D \mid I)P(I) + P(D \mid II)P(II)} = \frac{0.02*0.6}{0.02*0.6 + 0.01*0.4} = 0.75$$



Definitions

- **Random variable**

  A random variable is a variable whose numerical value is determined by the outcome of a random trial.

  - We regard the random variable as a name for the number associated with the outcome of the experiment **before** the experiment is performed. Carrying out the experiment converts the random variable into a specific number.
  - A random variable has either an associated **probability mass function** (for discrete random variable) or **probability density function** (for continuous random variable).

- **Types of random variable:**

  - **Discrete** variables produce outcomes that **come from a counting process** (e.g. number of classes you are taking).
  - **Continuous** variables produce outcomes that **come from a measurement** (e.g. your annual salary, or your weight).

What is a Distribution?

- A distribution defines the behaviour of a situation modelled by a variable X. We need at least a pair of information:
  - The set of all possible values {x}; i.e., “what can happen”
  - The probability of each value (discrete) or each range of values (continuous) occurring. $P(X = x)$; i.e., “how often something happens”

Types of Distributions

We can characterise distributions in 2 ways:

1. By context:
   - Population distribution: it is the distribution of values for a property we are interested in examining.
   - Sample distribution: we rarely know the population distribution, but we take a subset of samples from this distribution. The distribution of these values is the sample distribution.
   - Sampling distribution (for a statistic like $\bar{X}$): We might be interested in estimating the mean from this sample. We could perform another study where we take another set of samples and calculate the sample mean. If we did this repeatedly we would get the sampling distribution of the sample mean.
2. By underlying nature:
   - Discrete distribution
   - Continuous distribution

Counting Techniques – Multiplication rule

Multiplication rule

If there are m ways of doing one thing and n ways of doing another thing, there are **m*n** ways of doing both.

Example

If a man has 2 shirts and 3 ties, how many ways of choosing a shirt and then a tie?

2*3 = 6

The rule can be extended to more than 2 events. For 3 events, p, q, and r, the total number of arrangements = **p*q*r**

Independent or dependent?

- However, it only works when all choices are **independent** of each other.
- If one choice affects another choice (i.e. **depends** on another choice), then a simple multiplication is not right.

Counting Techniques – Multiplication rule

Example

You are buying a new car. There are 2 body styles (sedan or hatchback), 5 colours available (black, red, green, blue, white), and there are 3 models (standard model, sports model with bigger engine, luxury model with leather seats).

- How many total choices are there?

Total choices = 2*5*3 = 30

- If the salesman says that you cannot choose black for the hatchback, how many total choices are there?

Total choices = (5*3) + (4*3) = 27

Counting Techniques – Permutation

A **permutation** is an arrangement of a set of objects in which there is a first, a second, and a third order through n; i.e., the **order is important**. {a, b} differs from {b, a}.

There are two types of permutation as follows:

- **Repetition is not allowed**.

The number of permutations of *n* objects taken *r* at a time is $_{n}P_{r} = \frac{n!}{(n-r)!}$ where

$_{n}P_{r}$ is read as "*n* permute *r*."

P is the number of permutations (or ways) the objects can be arranged.

n is the total number of objects.

r is the number of objects to be used at one time.

n! = n*(n – 1)*(n – 2)*(n – 3)*4*3*2*1

If r = n, then the number of permutations of *n* objects taken *n* at a time is $_{n}P_{n} = n!$

- **Repetition is allowed.**

The number of permutations of *n* objects taken *r* at a time when repetition is allowed is $_{n}P_{r} = n^{r}$

Counting Techniques – Permutation

<u>Example</u>

Janet has 6 business suits and will travel to Melbourne for a business convention the first 4 days of next week. How many ways can she choose suits for Monday, Tuesday, Wednesday, and Thursday if

- she wears a different suit each day?

$${}_6P_4 = \frac{6!}{(6-4)!} = 360$$

- she is willing to wear the same suit repeatedly?

$${}_6P_4 = 6^4 = 1296$$

<u>Example</u>

An amusement park has 28 different rides. In how many different ways can a person try 4 of these rides, assuming that order matters and that he/she does not want to try any ride more than once?

$${}_{28}P_4 = \frac{28!}{(28-4)!} = 491400$$

Counting Techniques – Combination

A **combination** of a set of objects is a subset of the objects disregarding their order; i.e., the **order is not important**. {a, b} is the **same** as {b, a}.

There are two types of combination as follows:

- **Repetition is not allowed.**

  The number of combinations of *n* distinct taken *r* at a time is $\binom{n}{r} = \frac{n!}{r!(n-r)!}$ where

  $\binom{n}{r}$ is read as "*n* choose *r*".

  n is the total number of objects.

  r is the number of objects to be used at one time.

  n! = n*(n – 1)*(n – 2)*(n – 3)*4*3*2*1

If a set has *n* elements, a total of 2^n subsets can be formed from those elements; i.e., $\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n} = 2^n$

- **Repetition is allowed.**

  The number of combinations of *n* distinct taken *r* at a time when repetition is allowed is

  $$\binom{n+r-1}{r} = \binom{n+r-1}{n-1} = \frac{(n+r-1)!}{r!(n-1)!}$$

Counting Techniques – Combination

Example

In how many different ways can a person choose 3 books from a list of 10 best-sellers, assuming that the order in which the books are chosen is of no consequence?

$$\binom{10}{3} = \frac{10!}{3!(10-3)!} = 120$$

Example

There are five flavours of ice-cream: banana, chocolate, lemon, strawberry and vanilla. We can have three scoops. Assuming that you can repeat the flavours, how many variations will there be?

$$\binom{7}{3} = \frac{7!}{3!(7-3)!} = 35$$

Example

A pizzeria offers 10 toppings for its pizzas. The customer can order any combinations of these toppings, including all of them and none of them. How many different types of pizza are possible?

$2^{10} = 1024$