

BIOS 600: Principles of Statistical Inference

Probability

Fall 2012

Reading

- ▶ Pagano and Gauvreau, Chapter 6, Section 6.1
- ▶ Interactive BBC Game on Probability

Probability Theory

- ▶ Last time: talked about how *descriptive statistics* can be used to summarize data
- ▶ Goal: make *inferences* about a *population* based on data from a sample of that population
- ▶ *Statistical inference* is built on the foundation of *probability theory*

Basic Probability

An *event* is the basic element to which probability can be applied, e.g. result of an observation or experiment.

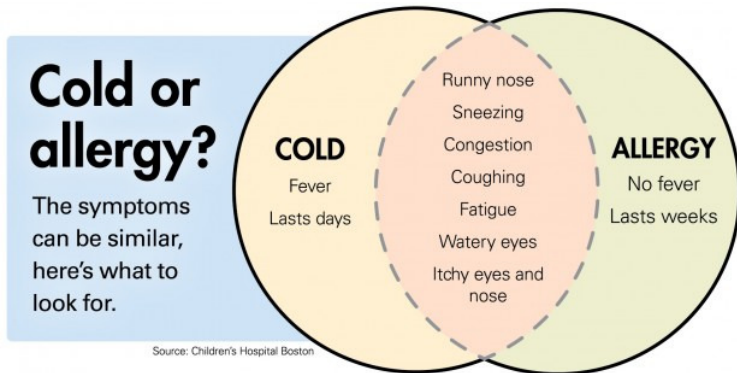
- ▶ An *event* can occur or not occur.
- ▶ Often upper case Latin letters used to define events, e.g. A, B, C, ...
- ▶ A is the event that a person is exposed to high levels of carbon monoxide
- ▶ B is the event that a person has blood type A+
- ▶ C is the event that a person gives birth to twin boys
- ▶ D is the event that a baby is female

Operations on Events

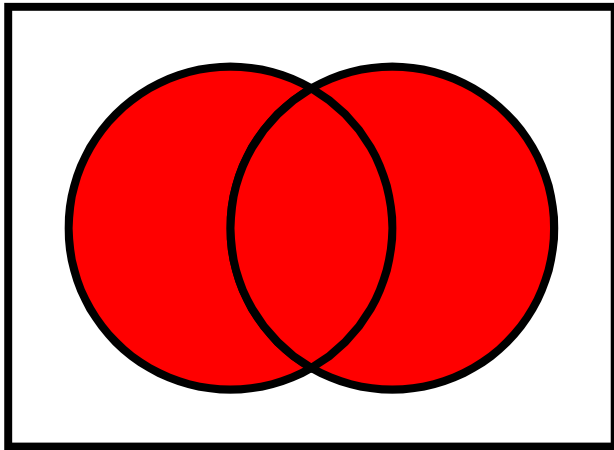
- ▶ A is the event that a person is exposed to high levels of carbon monoxide
- ▶ B is the event that a person has blood type A+
- ▶ C is the event that a person gives birth to twin boys
- ▶ D is the event that a baby is female
- ▶ The *union* of A and B , denoted $A \cup B$, is the event that either A or B , or both A and B , occur.
- ▶ Thus $A \cup B$ is the event that a person is exposed to high levels of CO, has blood type A+, or both.

Venn Diagram: Symptoms

A *Venn diagram* is a useful way to show possible relationships among sets.

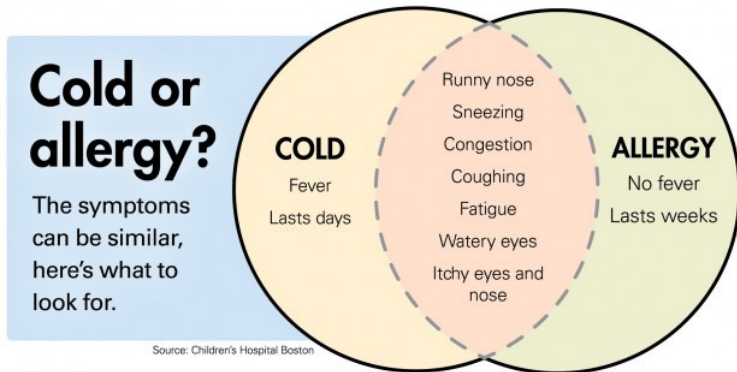


Venn Diagram: Union



Venn Diagram: Symptoms

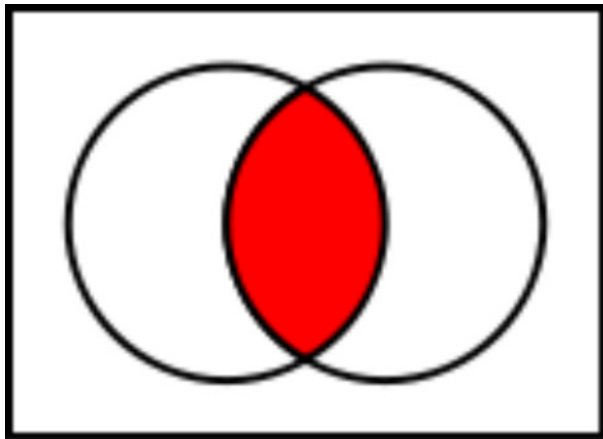
What is the union of cold and allergy symptoms?



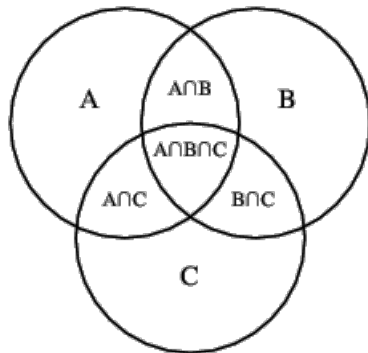
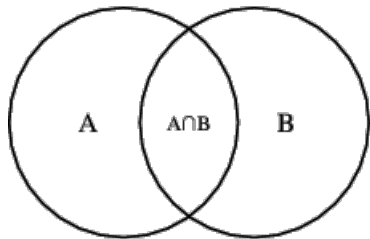
Operations on Events

- ▶ A is the event that a person is exposed to high levels of CO
- ▶ B is the event that a person has blood type A+
- ▶ C is the event that a person gives birth to twin boys
- ▶ D is the event that a baby is female
- ▶ The *intersection* of A and B , denoted $A \cap B$, is the event that both A and B occur.
- ▶ Thus $A \cap B$ is the event that a person is exposed to high levels of CO and has blood type A+.

Venn Diagram: Intersection

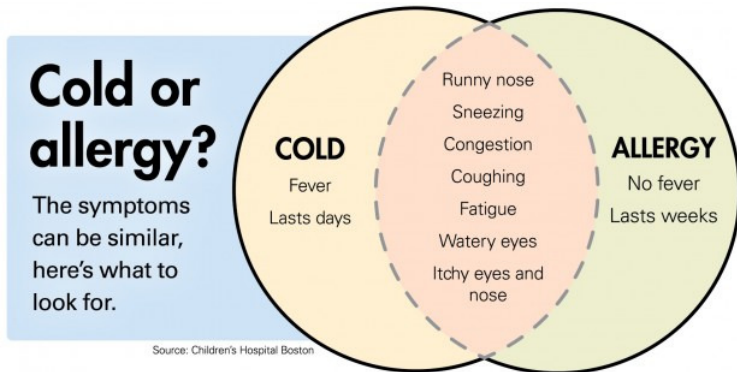


Venn Diagram: Intersection



Venn Diagram: Symptoms

What is the intersection of cold and allergy symptoms?



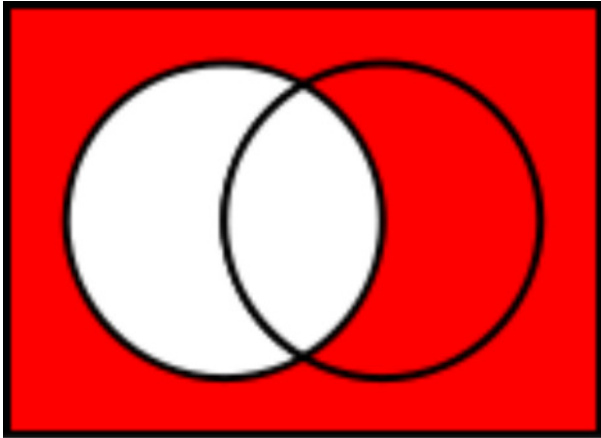
Venn Diagram



Operations on Events

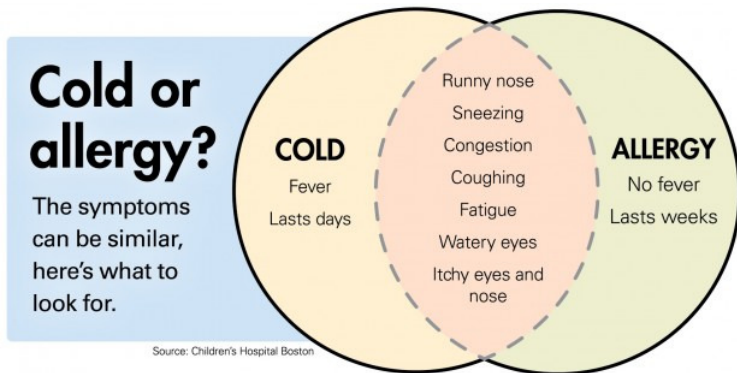
- ▶ A is the event that a person is exposed to high levels of CO
- ▶ B is the event that a person has blood type A+
- ▶ C is the event that a person gives birth to twin boys
- ▶ D is the event that a baby is female
- ▶ The *complement* of A , denoted A^c or \bar{A} , is the event that A does not occur.
- ▶ Thus \bar{A} is the event that a person is not exposed to high levels of CO
- ▶ A and \bar{A} are *mutually exclusive* events, which means they cannot happen together

Venn Diagram: Complement



Venn Diagram: Symptoms

What is the complement of cold symptoms?



Frequentist Definition of Probability

- ▶ If an experiment is repeated n times under essentially identical conditions, and if the event A occurs m times, then as n grows large, the ratio $\frac{m}{n}$ approaches a fixed limit that is the *probability* of A :

$$\Pr(A) = \frac{m}{n}.$$

What is the Probability of a Kiss?

This will be an in-class activity...prepare yourself!

Hands-on: Hospital-acquired Infections

Suppose you want to estimate the probability a patient will develop an infection while hospitalized at Dupe Hospital. In the past year, Dupe Hospital had 6450 patients, and 712 developed an infection. Estimate the probability that a patient at Dupe will develop an infection.



Hands-on: Come Home for Cardiac Arrest

An article in the *New York Times* reported that people who suffer a cardiac arrest in NYC have only a 1 in 100 chance of survival (attributed to factors such as traffic congestion and the difficulty of finding victims in large buildings).

- ▶ Express this in probability notation
- ▶ The basis of the article was a scientific paper reporting results of a study of 2329 consecutive cardiac arrests. To justify the “1 in 100 chance of survival” statement, how many of the 2329 do you think survived?



Subjective Definition of Probability

The *probability* of A is one's subjective degree of belief that the event will occur.

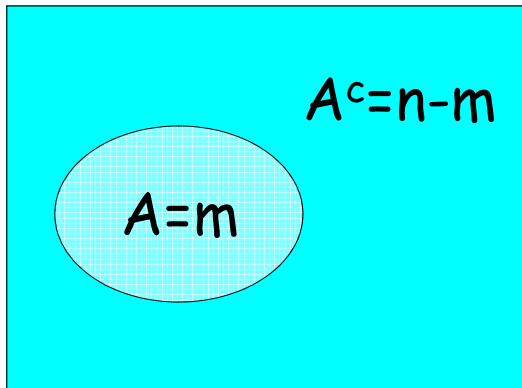
- ▶ The probability it will rain on a given day in Chapel Hill is 0.32
- ▶ The probability it will rain on a day upon which I have to walk across campus more than twice is 0.80

Rules of Probability

- ▶ Probabilities range from 0 to 1 ($m \leq n$)
- ▶ If an event always happens, it has probability 1 ($m = n$)
 - ▶ $\Pr(D \cup \bar{D}) = \Pr(\text{female, or not, or both}) = 1$
- ▶ If an event never happens, it has probability 0 ($m = 0$)
 - ▶ $\Pr(D \cap \bar{D}) = \Pr(\text{female AND male}) = 0$
 - ▶ An event that can never occur is called the *null event* and is sometimes represented by \emptyset
- ▶ $D \cap \bar{D} = \emptyset$
- ▶ $\Pr(D) + \Pr(\bar{D}) = 1$
 - ▶ The probability an event will not occur is equal to 1 minus the probability it will occur

Rules of Probability

$$S = n$$



Rules of Probability: Births

The probability of giving birth to a boy in the US is currently around 0.512. (Men then proceed to perish at a higher rate, and the gender ratio evens out in the 30's...by age 85, there are about twice as many women as men.)

- ▶ So at birth, $\Pr(\text{male}) = \Pr(\bar{D}) = 0.512$, so that 51.2% of babies are boys
- ▶ How do you find $\Pr(\text{female}) = \Pr(D)$?
 - ▶ Assuming all babies are either male or female, $\Pr(D \cup \bar{D}) = 1$
 - ▶ Babies cannot be both male and female, so $\Pr(D \cap \bar{D}) = 0$
 - ▶ The probability a baby is female is
 $\Pr(D) = 1 - \Pr(\bar{D}) = 1 - 0.512 = 0.488 \rightarrow 48.8\%$ of babies are girls

Rules of Probability: Asthma Patients

A 2008 study asked asthma patients two questions

1. Do conventional asthma medications usually help your symptoms?
2. Do you use complementary therapies (e.g., herbs, acupuncture) in treatment of your asthma?

and obtained the following results.

Use complementary therapy	Meds Help?		Total
	Yes	No	
Yes	122	24	146
No	757	97	854
Total	879	121	1000

Rules of Probability: Asthma Patients

Use complementary therapy	Meds Help?		Total
	Yes	No	
Yes	122	24	146
No	757	97	854
Total	879	121	1000

- ▶ What is the probability that conventional medicines help?
- ▶ What is the probability a patient uses complementary therapy?
- ▶ What is the probability conventional medicines do not help and a patient uses complementary therapy?

Mutually Exclusive Events

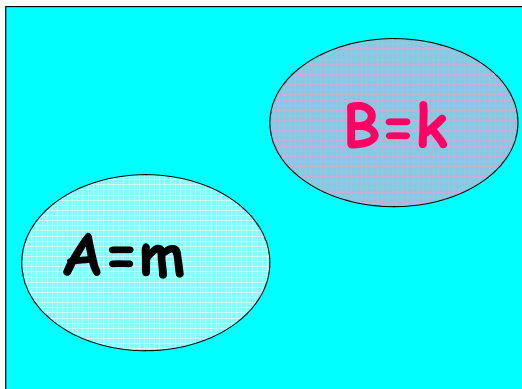
If two events A and B are mutually exclusive, then

$$\begin{aligned}\Pr(A \cup B) &= \Pr(A) + \Pr(B) \\ \frac{m+k}{n} &= \frac{m}{n} + \frac{k}{n}\end{aligned}$$

Additive Rule of Probability

$$S=n$$

$$A \cap B = \emptyset$$



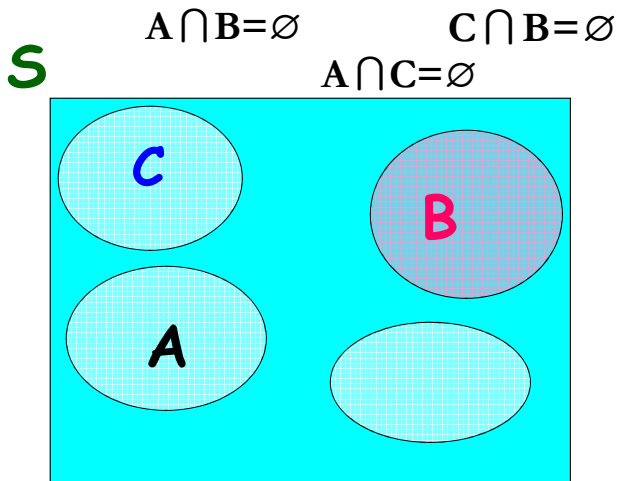
Additive Rule of Probability

More generally, if events A , B , C , \dots , are *mutually exclusive* – so that at most one of them may occur at any one time – then

$$\Pr(A \cup B \cup C \dots) = \Pr(A) + \Pr(B) + \Pr(C) + \dots$$

Interpretation: If events cannot happen together, the probability that any one of them will occur is the sum of their individual probabilities.

Additive Rule of Probability



Case Study: US Births

Consider these hypothetical data on births in the US over a given time period.

Type of Birth	# of Births
Singleton	41,500,000
Twins	500,000
Triplets	5,000
Quadruplets	100

- ▶ Are these events mutually exclusive?
- ▶ Let A be the event of a singleton birth. What is \bar{A} ?
- ▶ What is the probability of \bar{A} ?
- ▶ What is the probability a randomly selected pregnant woman will give birth to quadruplets? (Maybe you should watch out for kisses!)