chapter title: Quick on set theory

A *set* is a collection of things. I will use letters, often uppercase, to denote a set. Three examples are

$S \equiv \{\text{all squirrels who live on the University of Colorado campus}\}$

$A \equiv \{\text{all male students at C.U. Boulder who are math majors with BMWs}\}$

and

$$T \equiv \left\{ \begin{array}{c} \text{tennis sets in terms of who won each game: each element is a sequence of at least six games} \\ \text{between two individuals } (A \text{ and } B) \text{ with the last two won by the same person} \end{array} \right\}$$

{ } are called "braces." Some symbols are reserved for particular sets: $\emptyset$ denotes the empty set (a set with no members), my set of friends, and Ω denotes the universal set (the set that includes everything, or at least everything being considered).

Every squirrel on campus would be an element in (member of) the S set. In the T set, if A indicates that individual A won game, the element $AAAAAA$ is the element/outcome that individual A won the set by winning the first six games, and $ABABABABABABBB$ is an element indicating that A and B traded wins for six games, then B won the next two games to win the set 8 to 6.[1]

1 As statisticians we care about sets and set theory because?

Put simply (more later), we care about set theory because it is a foundation of probability theory. Consider the outcome of an experiment. Then consider all of the possible outcomes - the set of outcomes. A particular outcome is an element in the set of possible outcomes.

We are often interested in the likelihood (probability) that an outcome will have some property (e.g. the patient dies, interest rates go up, 10 cigarettes are consumed). For example, a treatment might characterized in terms of two random variables: patient dies within a year of treatment (or does not), and patient loses at least ten pounds (or not). We might be interested in the likelihood that someone who gets the treatment lives and loses weight. All the outcomes with the properties "patient dies" and "patient losese at least ten pounds" are a subset of all the possible outcomes of the experiment. We want to know the likelihood that the outcome will belong to that subset.

2 In economic theories constraints are often assumed, and these are sets, constraint sets

For example, conside the consumer's budget set in consumer theory and the input-requirement set in production theory

[1] Think about all the different possible elements in the set of possible tennis sets. Are the number of elements finite? Countable? Does the number of elements in this set depend on whether the match has the tie-breaker rule?

3 Some notation about the relationship between two sets

- $X \cup Y$ is the *union* of the sets X and Y; that is, the set that includes all the elements of X and Y.
- $X \cap Y$ is the *intersection* of the X and Y; that is, those elements that belong to both X and Y. Alternative notation is XY and "X and Y".
- $X \backslash Y$ is the elements that belong to X but not to Y. An alternative notation is $X - Y$, called a *set difference*.
- $\overline{X} \equiv \Omega \backslash X \equiv \Omega - X$. $\overline{X}$ is called the compliment of X. Sometimes you will see the notation X^c to denote the compliment of X. So $X \cup \overline{X} = \Omega$ and $X \cap \overline{X} = \emptyset$.

Some additional notation and concepts:

- $x \in A$ means x is an element in the set A. So, for example, $A \cap B = \{x : x \in A \text{ and } x \in B\}$
- $M \subset K$ means the set M is a subset of the set K.[2]
- Sets X and Y are said to be equivalent if $X \subset Y$ and $Y \subset X$.

3.1 Some examples:

Unions:

$\{1, 2, 3\} \cup \{a, b, c\} = \{1, 2, 3, a, b, c\}$

$\{1, 2, 3\} \cup \{3, 5\} \cup \{7\} = \{1, 2, 3, 5, 7\}$

$\left\{\sqrt{2}, \pi, 3.9, r\right\} \cup \{a, b, c\} = \left\{\pi, r, a, b, c, 3.9, \sqrt{2}\right\}$

Intersections:

$\{1, 2, 3\} \cap \{2, 4, 6\} = \{2\}$

$\{a, b, c, d\} \cap \{d, e, f\} = \{d\}$

$\{1, 2, 3\} \cap \{a, b, c\} = \emptyset$

$\{1, 2, 3\} \cap \{\} = \emptyset$

[2] Sometimes we distinguish between subsets and strict subsets, using $\subseteq$ to denote subset and $\subset$ to denote a strict subset. With this more precise notation $A \subset B$ mean that A is a subset of B but there are elements in B that are not in A, so A is a strict subset of B. Whereas, $A \subseteq B$ allows for the possibility that $A \equiv B$.

Note that $(A \subseteq B \text{ and } B \subseteq A) \iff (A \equiv B)$.

In these notes, we are using $\subset$ to mean $\subset$ or $\subseteq$.

Note that $\cap$ and $\cup$ are algebraic commands but that they apply to sets rather than to variables. Think of *algebra* for sets, *set algebra*. If two sets have no elements in common their intersection is the *empty set*, denoted by empty brackets $\{\}$ or the symbol $\emptyset$.

Combinations of union and intersections:

$\{1, 2, 3, c\} \cap (\{2, 4, 6\} \cup \{a, b, c\}) = \{2, c\}$

$(\{1, 2, 3, c\} \cap \{2, 4, 6\}) \cup (\{1, 2, 3, c\} \cap \{a, b, c\}) = \{2, c\}$

4 Venn diagrams

Venn diagrams, attributed to John Venn circa 1880, are a way to pictorially represent sets and the relationships between those sets. WolframMathWorld defines them as "A schematic diagram used in logic theory to depict collections of sets and represent their relationships." (http://mathworld.wolfram.com/VennDiagram.html).

The universal set, Ω, is often represented with a rectangle in two-dimentional space: the rectangle represents all possibilities, but sometimes Ω is simply implicit. The dimensions of the rectangle need not have cardinal or ordinal meaning, but can. The objective of the Venn diagram is typically to provide a visual representation of the **relationships** between **two** or more sets in terms of the set properties $\cup$ and $\cap$, $\setminus$, and . The sets to be considered (e.g. X,Y, and Z) are each represented an area in the rectangle.[3]

For example, one might represent set X , a strict subset of Ω, with

[3] "A Venn diagram is constructed with a collection of simple closed curves drawn in the plane. According to Cyndi Joyce Aguzar (1918), the 'principle of these diagrams is that classes [or sets] be represented by regions in such relation to one another that all the possible logical relations of these classes can be indicated in the same diagram. That is, the diagram initially leaves room for any possible relation of the classes, and the actual or given relation, can then be specified by indicating that some particular region is null or is not-null.'

Venn diagrams normally comprise overlapping circles. The interior of the circle symbolically represents the elements of the set, while the exterior represents elements which are not members of the set... Shapes other than circles can be employed, and this is necessary for more than three sets. Venn diagrams do not generally contain information on the relative or absolute sizes (cardinality) of sets; i.e. they are schematic diagrams." (http://en.wikipedia.org/wiki/Venn_diagram).

To be precise, most of the diagrams presented in this section are Euler diagrams, not Venn diagrams, but they are commonly called Venn diagrams–the expression Euler diagrams is not in common usage. Strictly speaking Venn diagrams must show all possible combinations of intersections between the sets, whether the intersection does or does not have members. In a strict Venn, empty intersections are indicated by shading. In an Euler diagram, only non-empty intersections are represented. For example, Figure 2_1 does not show all possible intersections between X and Y so is not strictly a Venn diagram.

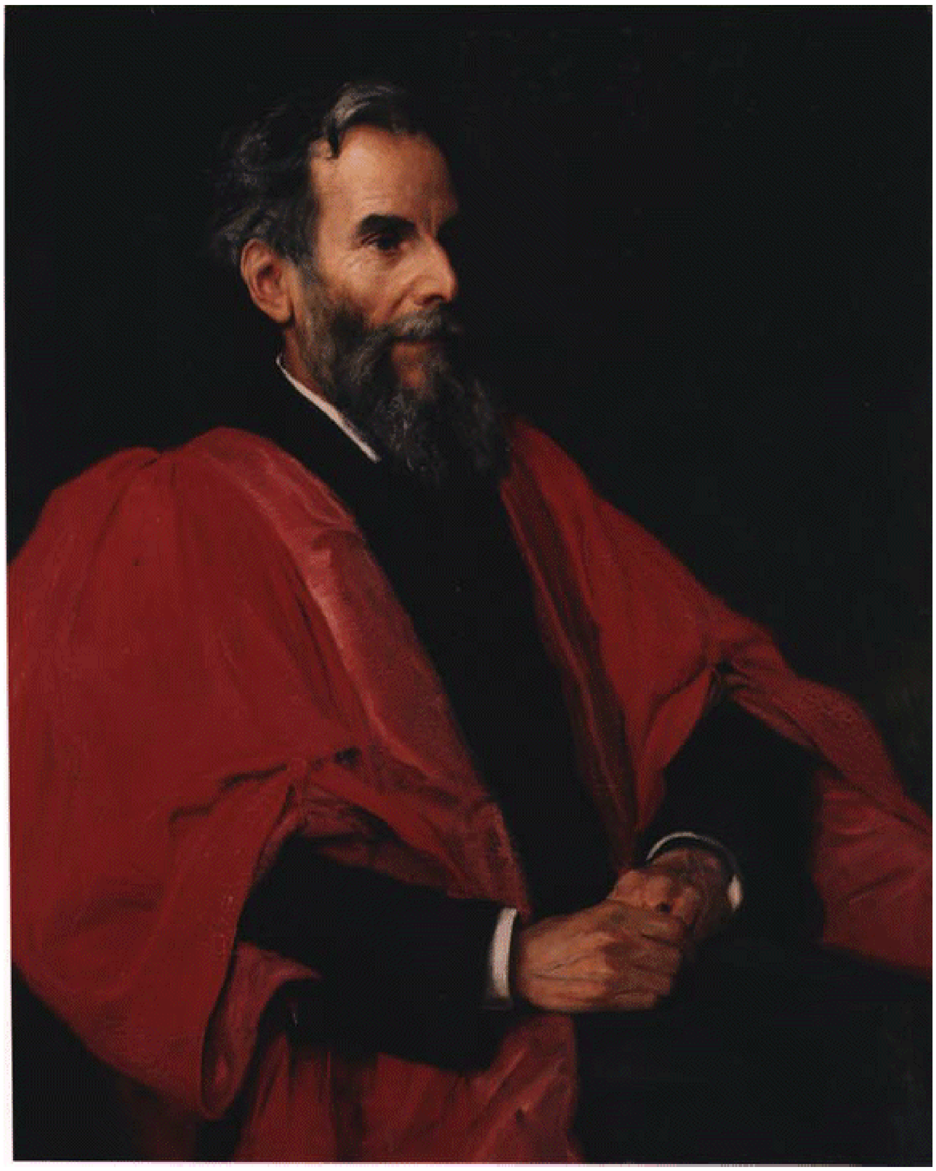

Figure 1: John Venn (1834-1923)

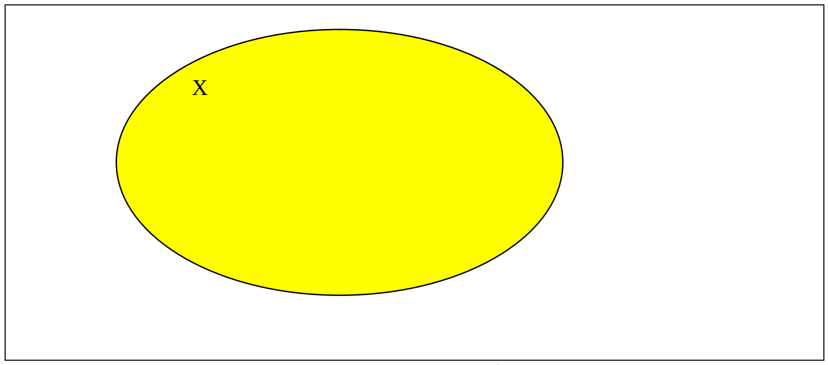

Figure 2_1: Venn of set X

The magnitude of the area used to represent X need not matter. For example, in the diagram, set X should be interpreted as a strict subset of Ω because it does not include all of Ω , but area (size) as a proportion of the rectangle need not reflect the proportion of the elements in Ω that are in X, but it could, depending on the intent of the diagram's creator. Consider now a Venn of two non-intersecting sets

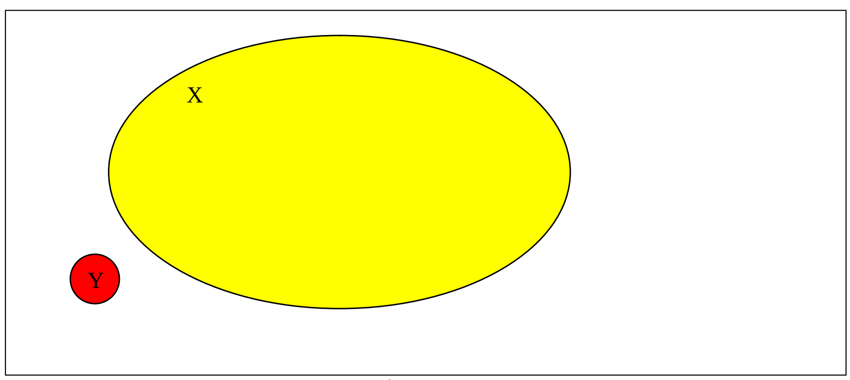

Figure 2_2: Venn of non-intersecting X and Y

If I were viewing Figure 2_2, but were not its creator, my tendency would be to assume set Y has fewer elements that set X.

While the shapes used to represent sets (circles, squares, rectangles, or more exotic shapes such as stars and donuts) typically have no significance, if the creator gives a set an exotic shape, the viewer will likely try to attach meaning to the shape. "Why does the shape used to represent all males in the U.S. look like the "Starship Enterprise?"". For example, Figure 2_3 is likely to confuse, unless Y represents Custer and his soldiers and Y represents the indian war-party right before the *Battle of Little Bighorn* commenced.

With these caveats and warnings in place, consider a Venn diagrams to represent $X \cup Y$, $(X \cup Y)\backslash Z$, and $(X \cup Y) \cap Z$. The intent is to represent the relationship between the sets. $X \cup Y$ is the bluish and greenish areas, including their overlap and their parts that are organgeish; $(X \cup Y)\backslash Z$ is the blue and

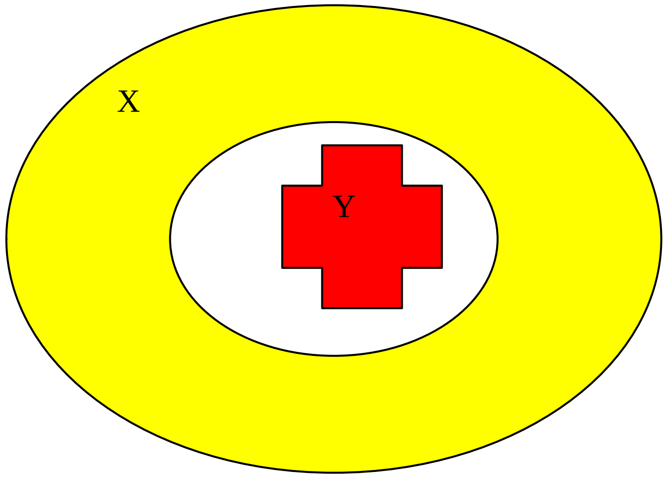

Figure 2: Figure 2_3

green excluding the orangeish, and $(X \cup Y) \cap Z$ is the area that is simaltaneously organeish and either bluish or greenish.

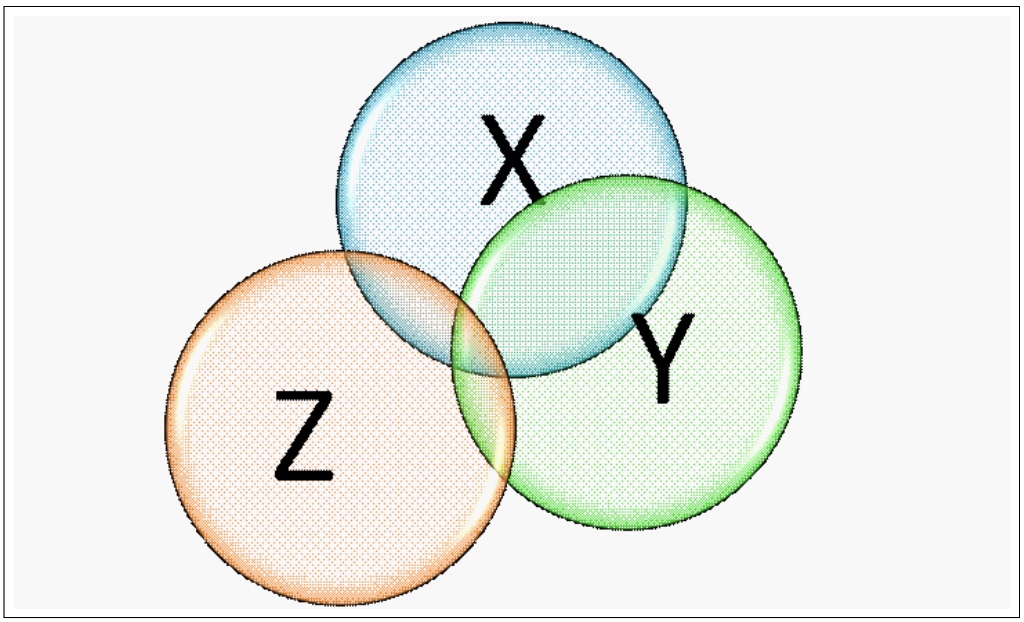

Figure 2_4: Venn of three overlapping sets

Venn diagrams prove useful for understanding the relationships between sets, for proving that certain relationships eliminate the possibility of other relationships, for intuiting equivalent ways to represent the same relationship, and for disproving conjectures about set relationships (by presenting a counter-example with a Venn diagram). For example, that $\overline{A \cup B} = \overline{A} \cap \overline{B}$ (De Morgan's Law) is, to me, visually intuitive in terms of a Venn diagram.

Is the following formula true? $X \backslash Y = Y \backslash X$. Demonstrate your answer with a Venn diagram.

4.1 An example: guys named Edward

Now consider the following sets: all homo sapiens who have ever lived, all male homo sapiens who ever lived, all homo sapiens currently alive, and all homo sapiens ever named Edward. Assume that females are never named Edward and that everyone is either a male or female. Consider a Venn diagram, Figure 2_4 that indicates that most humans with the name Edward are dead.[4] Not that the adjective "most" indicates that area sizes will have significance.

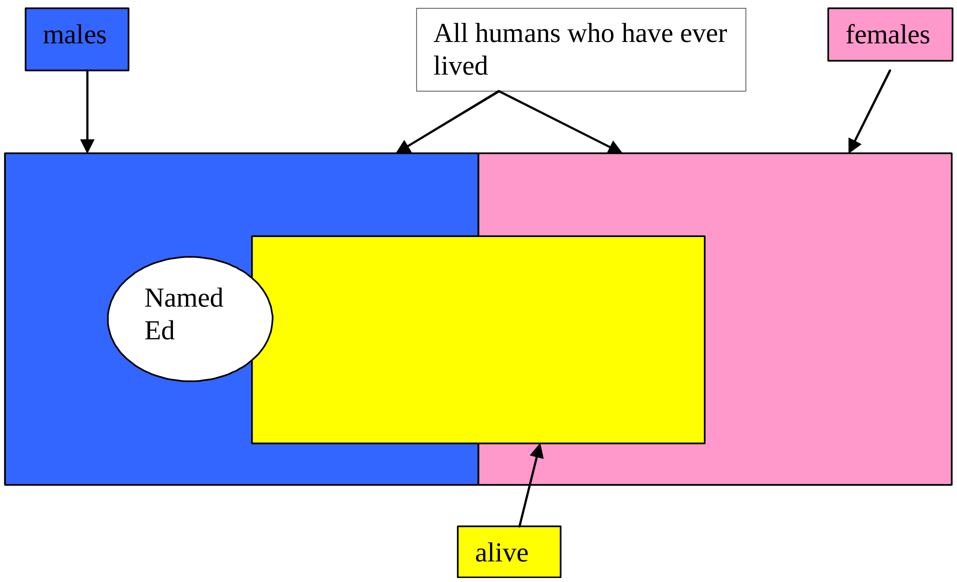

Figure 2_4: Venn of guys named Ed

I have assumed everyone was or is either a male or female–not quite true–and that half the births are female, so made the two gender sets of equal size. Those alive are a strict subset of those who have lived. I drew the "Ed set" so that most of them are in the dead-male category.

[4] Given the recent popularity of the Twilight series (http://en.wikipedia.org/wiki/Twilight_(series) with Edward, its hearthrob vampire, the number of living and dead males named Edward should see a dramatic rise.

4.2 An example: cows in the valley

Now consider the following bad joke about deduction. Assume three statisticianss: a statistics professor at CU, an undergraduate statistics major at CU, and a statistics professor at MIT. They are out hiking, come to a ridge and look down into a valley. They see a bunch of cows, all of which look black. The prof at CU concludes on the basis of this observation that "All cows are black." The undergraduate say "wrong", then goes on to say that the observation proves only that "All cows in this valley are black." The MIT prof then asserts that what has been proven is that "All cows in this valley are black on at least one side."

The Venn diagram, Figure 2_x, demonstrates that only the professor at MIT is necessarily correct.

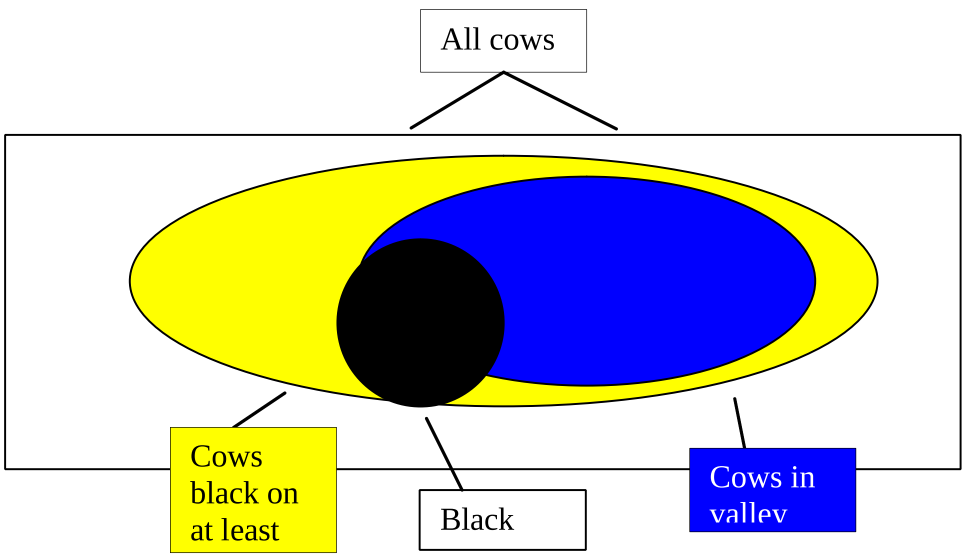

Figure 2_x: Venn of cows in the valley

The white area is all cows that are not black on at least one side.[5]

If all cows in the valley are black, one needs the set of "cows in the valley" to be a subset of "black cows." That is,

[5]Note here that the universal set is not everything; it is all cows.

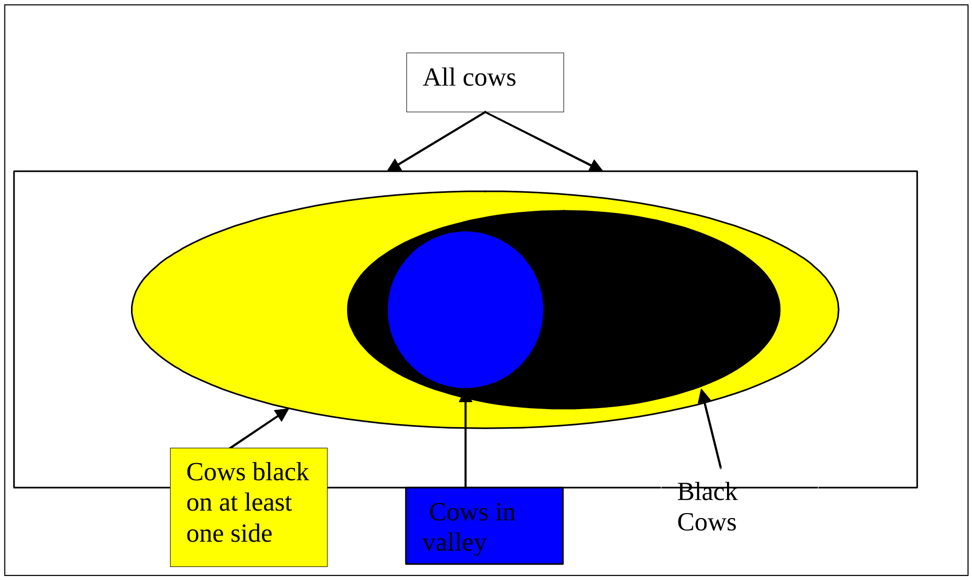

Figure 2_x: Venn if cows in the valley all black

In Figure 2_x, all cows in the valley are black but not all cows are black. Note that this graph implies something not implied by what the three saw. It requires the added assumption that if one is black, one is black on all sides; an assumption made by the CU professor and student. If, more restrictively, all cows are black, the Venn diagram, Figure 2_x, is

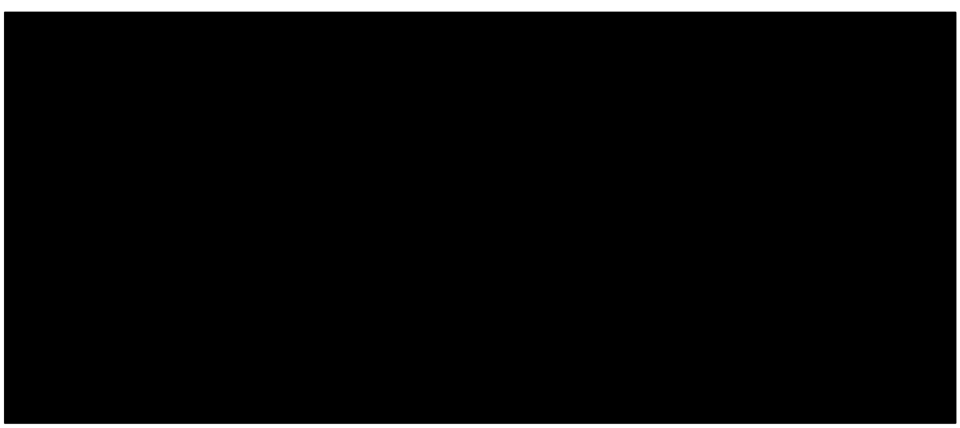

Figure 2_x: Venn of all cows black

A stupid thing to conclude if you only see a couple of cows who happen to be black on the side you are viewing.

5 Some theorems one can derive from the definitions and properties of sets

1. Communitve law: $A \cup B = B \cup A$ and $A \cap B = B \cap A$
2. Associate law: $A \cup (B \cup C) = (A \cup B) \cup C$ and $A \cap (B \cap C) = (A \cap B) \cap C$
3. Distributive law: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ and $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
4. $(\overline{\overline{A}}) = A$: the compliment of the compliment of A is A. That is $not(notA) = A$
5. $A \cap \Omega = A$, $A \cup \Omega = \Omega$, $A \cap \emptyset = \emptyset$, and $A \cup \emptyset = A$
6. $A \cap \overline{A} = \emptyset$, $A \cup \overline{A} = \Omega$, $A \cap A = A$ and $A \cup A = A$
7. $\overline{A \cup B} = \overline{A} \cap \overline{B}$ and $\overline{A \cap B} = \overline{A} \cup \overline{B}$ called *De Morgan's laws*[6]
8. $A/B = A \cap \overline{B}$
9. $A = AB \cup A\overline{B}$ and $AB \cap A\overline{B} = \emptyset$
10. If $A \subset B$, then $A \cap B = A$, and $A \cup B = B$

The above are algebraic relationships in terms of sets. You should be able to convince someone that each of these theorems follow from the basic properties of sets (what are those basic properties?).Venn diagrams might help you to be convincing, so would formal proofs, if your audience understands such things. Consider, for example theorem 10, use a Venn diagram to convince that this is true. It is pretty obvious. These ten theorems, along with the basic properties of sets are an *algebra*, an algebra of sets.

6 Repetition: so, what, if anything, does set theory have to do with probability theory and statistics?

Again, and put simply, set theory and the "algebra" of sets is a foundation of probability. In probability theory, we define the *sample space* as the set of all

[6]From Wikipedia (http://en.wikipedia.org/wiki/De_Morgan's_laws) "The law is named after Augustus De Morgan (1806–1871) who introduced a formal version of the laws to classical propositional logic. De Morgan's formulation was influenced by algebraization of logic undertaken by George Boole, which later cemented De Morgan's claim to the find. Although a similar observation was made by Aristotle and was known to Greek and Medieval logicians (in the 14th century William of Ockham wrote down the words that would result by reading the laws out). De Morgan is given credit for stating the laws formally and incorporating them in to the language of logic. De Morgan's Laws can be proved easily, and may even seem trivial. Nonetheless, these laws are helpful in making valid inferences in proofs and deductive arguments."

possible outcomes of an experiment or sampling. Uses Ω to denote this set. Each element in Ω , ω, is a specific outcome/sample. Note that Ω is typically defined as the universal set, here in the context of all possible experiments.

One can view sampling as the outcome of a data-generating process. Each time the process is run, out pops a sample/outcome. The goal of probability theory is to determine, or estimate, the probability of different events. Let A represent some *event* (all possible outcomes that share the property of interest): a set of outcomes.

For example, consider rolling a die. One possible event it that the number is 3, another event is the number is odd. Note that only one sample can produce the event 3, but 3 samples can produce the event odd. For a different experiments one event might be the outcome the individual is dead, another, the individual is fat. (Later, we will define events more precisely.)

Consider some experiment and denote the set of all possible events, $\mathcal{A}$. (My experience is that identifying all possible events is often a difficult task.) We are concerned with the probability that event A will occur where A is a subset of $\mathcal{A}$. A topic for the immediate future will be defining and studying the sets Ω and $\mathcal{A}$, and the relationship between them.