Lecture 1 March 21st August 28, 2014

- Syllabus
- Oct 9/16 with Sept 23 schedule

Roadmap

I prob theory
- elementary set theory
- counting: combinations & permutations
- axioms / philosophy
- probability calculations
- Bayes Rule conditional prob.
- disjoint, dependence, independence

II random variables (r.v.)
- definition
- prob. mass function / cum. mass function
- expectation / variance
- combining r.v.'s - convolution, covariance,
- prob. density function / cum. density function
- multi-dimensional density & mass functions
- conditional expectation
- normal density
- moment generating function
- characteristic functions ?

III Real Data
- summary statistics
- visualization

IV Statistical Inference
- confidence intervals
- hypothesis testing

Basic Set Theory & Vocabulary

All mathematics is built on the foundation of "sets".

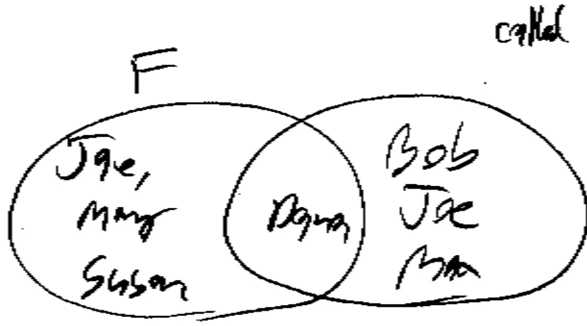
Since prob & stats is a branch of math, we need to be familiar with sets. Invented in the 1870's, formalized in the 20th century. Defn: a set is a well-defined collection of distinct objects (AKA "members" or "elements") for now

$F = \{\text{Jane, Mary, Susan, Dana}\}$

$M = \{\text{Bob, Joe, Max, Dana}\}$

they have no order and duplicates not allowed

We can draw it like a cartoon: called a "Venn diagram"



Inclusion is denoted with an $\in$-like symbol

Inclusion: Jane $\in F$ equivalent Jane $\in \{\text{Jane, Mary, Susan, Dana}\}$

Sets can have any # of elements, even infinite elements

define as $\mathbb{N} := \{1,2,3,\ldots\}$ or $\mathbb{N}_0 := \{0,1,2,3,\ldots\}$ disagreement $\mathbb{Z} = \{\ldots,-2,-1,0,1,2,\ldots\}$

Set equality is defined as both sets have the same members

There's a special set called the "natural #'s."

Subsets are denoted with a "$\subset$" or "$\subseteq$"

$\{$Jane, Mary$\} \subset F$

$\subset$ means "proper subset" which means the set on the l.h.s. is not equal to the set on the r.h.s.
$\subseteq$ means plain "subset" and it can be equal.

We combine sets by using unions "$\cup$" big cup symbol. e.g.

$\{$Jane$\} \cup \{$Mary, Susan, Dana$\} = F$

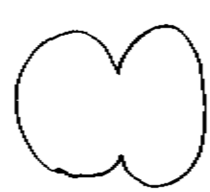
and what is

$\{$Jane$\} \cup \{$Jane$\} = ? \stackrel{?}{=}$ Jane

the union is like "addition" but doesn't "double-count."

it is also "non-exclusive or"

$F \cup M$ is female or male, none or both

$= \{$Jane, Mary, Susan, Dana, Bob, Joe, Max$\}$ pic 

We can also "intersect" sets use the $\cap$ symbol which take common elements

$$F \cap M = \{Dana\} \overset{?}{=} Dana \; ?$$

Intersection is AND. What's a female name AND a male name at the same time?

picture  ← erase

$$F \cap \{Bob, Joe\} \overset{?}{=} \{\}, \quad \phi := \{\}$$

null set special symbol

these two sets are called "mutually exclusive" or disjoint

We can also subtract sets

ODDS $\cap$ EVENS $= ?$

$$F \setminus M = \{Joe, Amy, Susan\}$$

picture: 

We can also expand a set into a larger set called the "powerset" which contains all subsets. $A = \{1,2,3\}$

$$2^A = \{\phi, \{1\}, \{2\}, \{3\}, \{1,2\}, \{2,3\}, \{1,3\}, \{1,2,3\}\}$$

We can also take the "size" of the set, or "measure" of, i.e. how many elements there are. This is called "cardinality" and denoted with absolute value signs

$|F| = 4$, $|F \cap M| = ?$, $|F \cup M| = ?$, $|2^F| = ?$

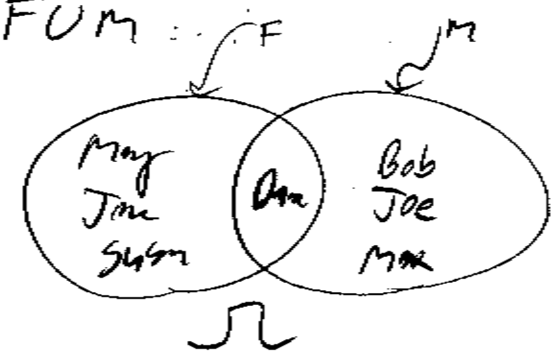
Why is $|2^F| = 8$? On/off bits $\{0/1, 0/1, 0/1\}$

$2 \times 2 \times 2 = 8$

The "most special" set is called "Ω" which denotes the "universe", the "universe of discourse" or the "sample space" (the book calls this "S") which is all elements ...

Our current investigation is restricted to.

e.g. $\Omega = F \cup M$



pieces

Note that $F \subseteq \Omega$, $M \subseteq \Omega$

notation same as $\sum_{i=1}^{n} i$

for all

In fact it's by definition $\Omega = \bigcup_{i=1}^{n} A_i \Rightarrow A_i \subseteq \Omega \;\forall i$

union set is usual term

$F \setminus M = \{$Mary, Jane, Susan$\}$
$M \setminus F = \{$Bob, Joe, Max$\}$
$F \cap M = \{$Dana$\}$

Claims: $F \setminus M \cup M \setminus F \cup (F \cap M) = \Omega$

but $F \setminus M \cap M \setminus F = \emptyset$

$M \setminus F \cap (F \cap M) = \emptyset$

$F \setminus M \cap (F \cap M) = \emptyset$

dice roll |1|2|3|4|5|6|

All sets "unioned" together are "collectively exhaustive"

$$\bigcup_{i=1}^{n} A_i = \Omega \quad \text{(definition)}$$

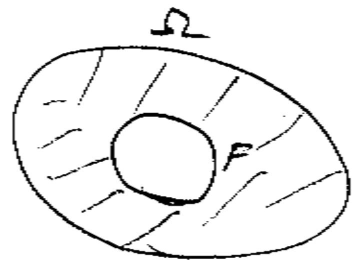
And no sets overlapped $A_i \cap A_j = \phi \quad \forall i \neq j$

once again this's mutually exclusive: being in one means you're not in the other.

One more set operation: "complement"

F^c means all elements that are not in F

$$F^c := \Omega \setminus F$$



$= \{$Bob, Joe, [illegible]$\}$

$\Omega^c = ?$

$F \cup F^c = ?$, $F \cap F^c = ?$

What's the "chance" or "probability" a name pulled out of Ω is female?

142857

$$\frac{|F|}{|\Omega|} = \frac{4}{7} \approx .57$$