Review

A probability model is a description of a random phenomenon in the language of mathematics

A probability model consists of

1) the set of possible outcomes of the phenomenon (sample space)

2) the probability of each outcome, more precisely each event (subset of sample space) (we'll see today why probabilities are assigned to events not outcomes)

Set Theory

A set is a collection of objects

Eg {Head, Tails}

{People in this room}

{Numbers between 1 and 6 (e.g. die faces}

~~III~~

Notions

Is an element of $\in$
Is not an element of $\notin$
Is a subset of $\subset$
Is not a subset of $\not\subset$

e.g. give examples

If A and B are sets
A is a subset of B, $A \subset B$, if every element of A is also an element of B

Empty set $\emptyset$, set of no elements.

$\emptyset \subset A$ for every set A
$A \subset A$ for every set A

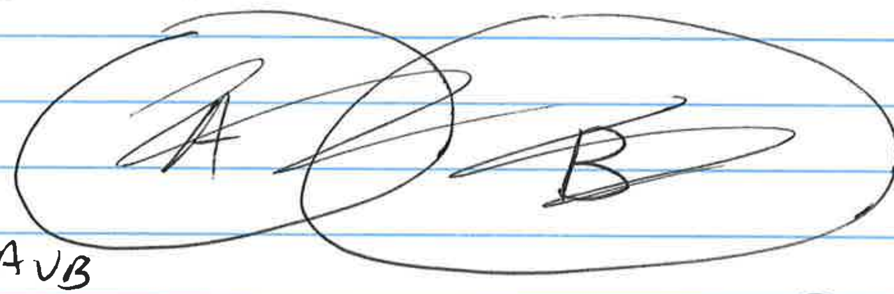
The sample space S of a random phenonomenon is the set of all possible outcomes of the phenomenon

{Heads, Tails} Coin
{1, 2, 3, 4, 5, 6} Die

An event is a subset of sample space

{H}, {T}, {H, T}, $\emptyset$

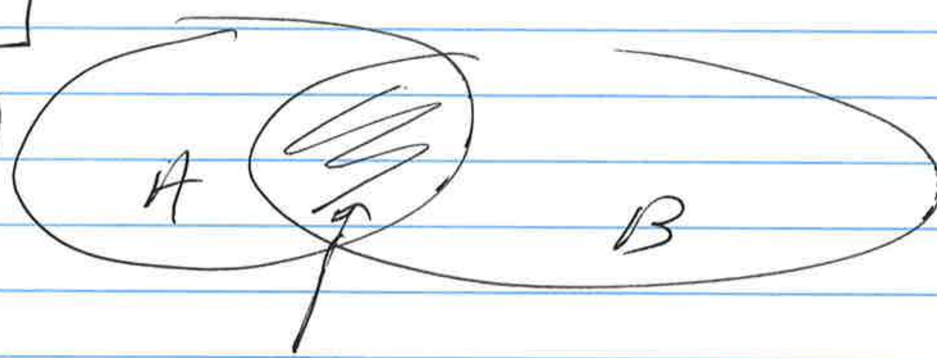
$\boxed{A \cup B}$ (A or B)



Elements of $A \cup B$
are in A or in B or Both

$\boxed{A \cap B}$

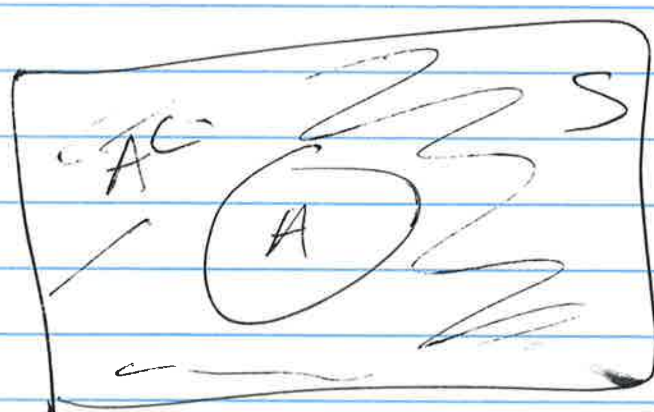
(A and B)



$A \cap B$

Elements of $A \cap B$ are in both A and B

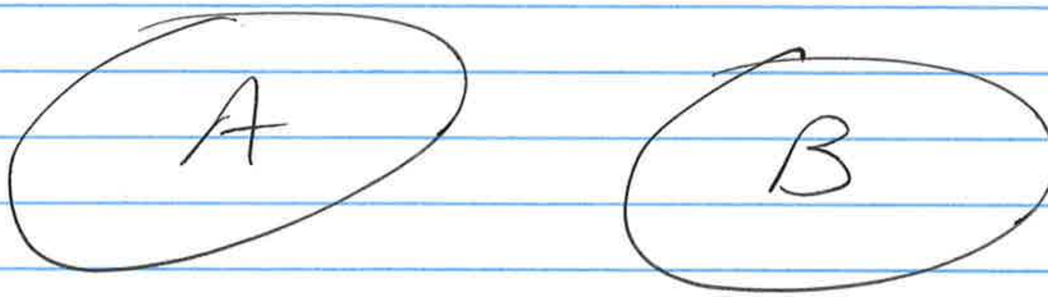
$\boxed{A^c}$



elements of A^c are in the sample space S
but not in A.

Disjoint Sets

HAVE no elements in common



A Probability is a number assigned to an event.

When probabilities are assigned to all events, they must obey certain rules

Rule 1: Each probability must be a number between 0 and 1

For all events A, $0 \leq P(A) \leq 1$

Rule 2: All possible outcomes together have probabity 1

$P(S) = 1$

↑ sample space

Rule 3: If two events have no outcomes in common, the probability that either one occurs or the other is the sum of the individual probabilities

If A and B are disjoint events
$P(A \cup B) = P(A) + P(B)$

Rule 4: The probability that an event does not occur is one minus the probability that it does occur.

For all events A, $P(A^c) = 1 - P(A)$

Rule 5: For Rule 5 we must define independent events

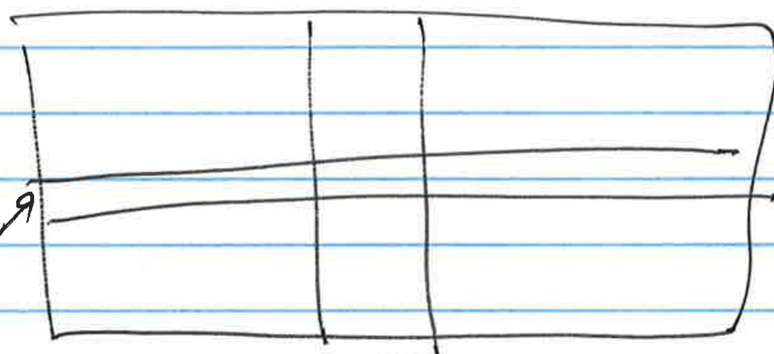
Two events A and B are <u>independent</u> if knowing that one occurs does not change the probability that the other occurs (sucessive coin tosses, sex of children)

first Heads can't predict if second Heads or tails

(First born a boy doesn't tell you if second is a boy or girl.

The picture that goes with this definition is

B



A

rectangular Sample Space

S

Areas of sets are there probabilities

A is a vertical bar B is a horizontal bar

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

area of intersection   base × height

Rule 5: If A and B are independent

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

If all outcomes are equally likely then

$$\text{Probability(each outcome)} = \frac{1}{\text{total number of outcomes}}$$

And if A is an event in a probability model where all outcomes are equally likely

Then

$$P(A) = \frac{\text{total number of outcomes in } A}{\text{total number of outcomes in } S}$$

Entire Sample Space.

§4.3 Random Variables

A random variable is a function that assigns a number to each outcome of a random phenomenon

EG: Flip a coin three times in a row.

TTT $\to$ 0
HTT $\to$ 1
THT $\to$ 1
TTH $\to$ 1
HHT $\to$ 2
HTH $\to$ 2
THH $\to$ 2
HHH $\to$ 3

outcomes

these are numbers assigned to outcomes

In this case the number assigned happens to be the number of heads seen.

However any assignment of numbers to outcomes is a random variable.

Not all of these assignments (RVs) would be useful.

A discrete random variable X is one with a finite number of possible values,

The probability distribution of X (discrete R.V.) lists ~~~~ all the possible values of X and their probabilities

Table.

Value of X	x_1	x_2	x_3	...	x_k
Probability	p_1	p_2	p_3	...	p_k

In the coin toss example there were 8 possible outcomes, each equally likely, for the random phenomenon, eg TTH, HTH, etc

There were only 4 possible values for the random variable,

Value of X	0	1	2	3
Probability	1/8	3/8	3/8	1/8

Here's another example

Lets say that every point on the earth is equally likely to be struck by a meteor (~~I~~ don't know if this is true)

Then the probability of an area being struck next is equal to

$$\frac{\text{its area}}{\text{area of earth's whole surface}}$$

Let (x,y) be the lattitude and longitude of the next meteor strike.

These are the outcomes.

The random variable is going to a function that assigns a number to each (x,y).

My random variable is going to assign

1 if Northern hemisphere
0 if equator
-1 if Southern hemisphere.

Is this a discrete random variable?

Yes! Finite number of values
even though sample space is infinite
(assuming infinite precision in measurement
of latitude/longitude.

What is distribution

Value of X	-1	0	1
Probability	1/2	0	1/2

Equator has zero probability (zero area)
(even though it isn't impossible to land there)

↳ Kind of like a coin landing on edge.

For a discrete random variable

1. Every probability p_i is a number between 0 and 1

2. $p_1 + p_2 + p_3 + \cdots + p_k = 1$

If the sample space is discrete you can find each p_i by adding the probabilities of the events that give x_i

Continuous Random Variable

A continuous random variable X takes all values in an interval of numbers. (lattitude of meteor strike)

The probability distribution of X is described by a density curve.

The probability of any event is the area under the density curve and above the values of X that make up the event.

All continuous probability distributions assign probability 0 to every individual outcome. (every lattitude probability 0 like equator)

Only intervals of values have positive probability

Binomial Random variable

Concerns a sequence of observations.

(Think: Coin tosses)

1. There must be a fixed number of observations n (in other words the number of observations can't depend on the results.

2. The results fall into two categories for convenience called "success" and "failure"
(Must decide if heads is success or heads is failure, tails is whatever heads isn't)

3. Results of each observation is independent (knowing results of some doesn't affect probabilities of the others)

4. For each observation, the probability of success is p.

The ~~Bernoulli~~ Binomial Random variable is the count of number of successes.

If (1) – (4) Hold can use Binomial calculator to get probabilities for number of sucesses } easy way to do problems.

This is easy way to do problems 1–3 on Homework #10 (ie 4.34, 4.35, 4.36)

There is a hard way without StatCrunch A challenge is to figure it out.

(Answers on back).