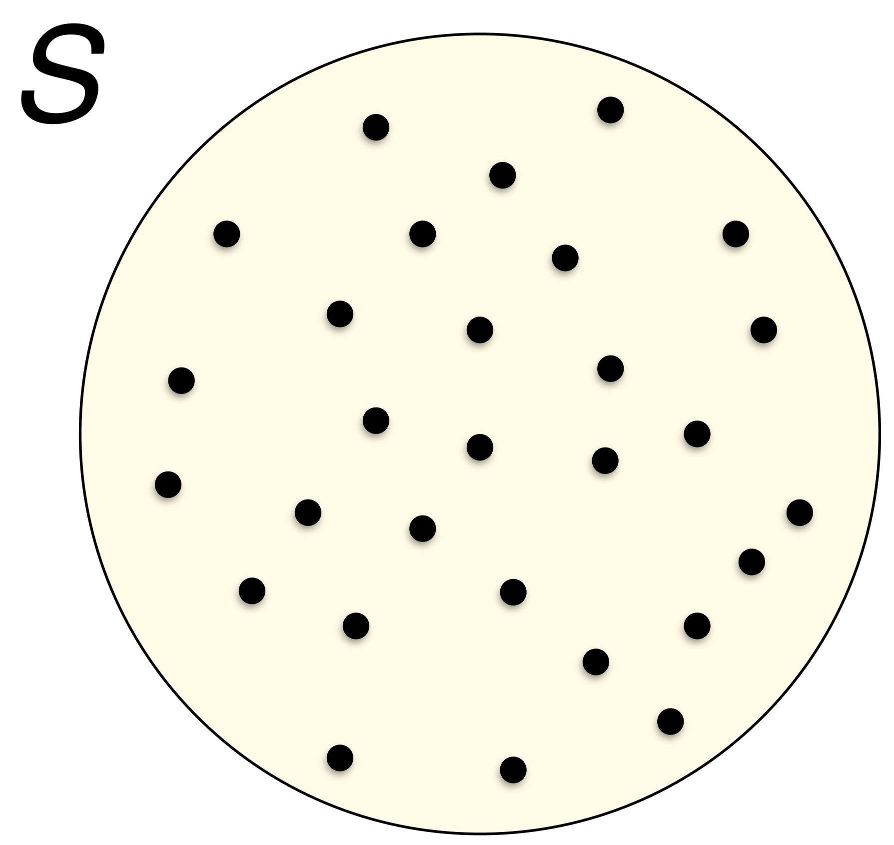
Lecture 2
31 Jan 2017

Logistics:
see piazza site for
bootcamps, ps0, bashprob

Discrete Probability and Counting

A *finite probability space* is a set S and a real function $p(s)$ on S such that:

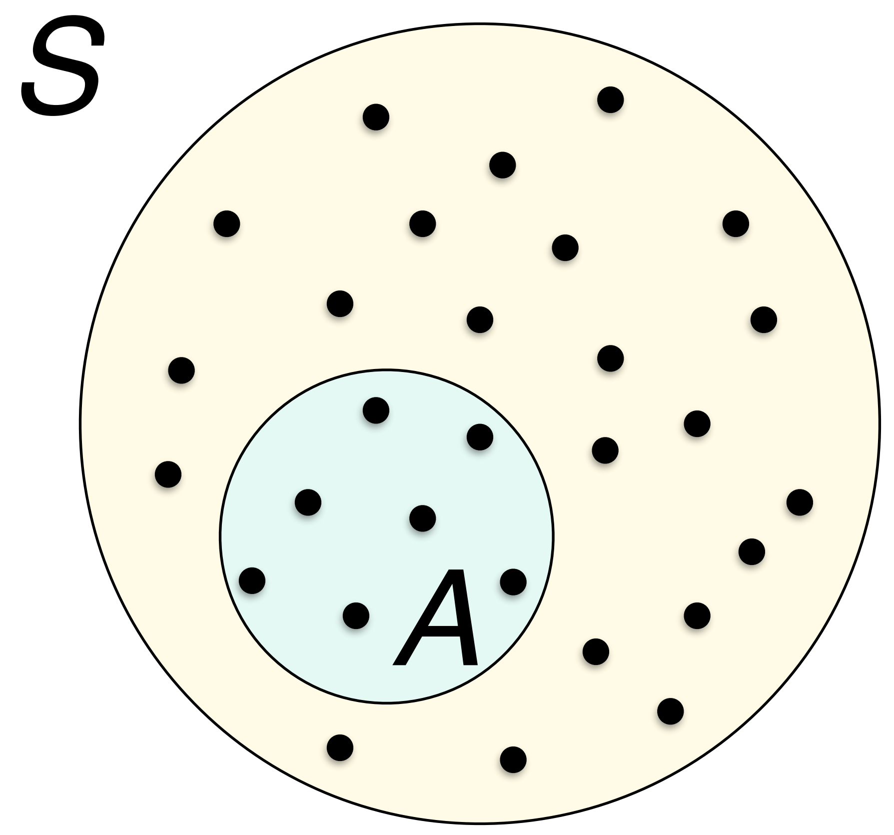
- $p(s) \geq 0\ ,\ \forall s \in S$, and
- $\sum_{s \in S} p(s) = 1\ .$



We refer to S as the *sample space*, subsets of S as *events*, and p as the *probability distribution.*

The probability of an event $A \subseteq S$ is $p(A) = \sum_{a \in A} p(a)$.

(Note that $p(\emptyset) = 0$.)



Example: Flip a fair coin. $S = \{H, T\}$

"Fair" implies that it is equally likely to come up H (heads) or T (tails), and therefore $p(H) = p(T) = 1/2$.

Note: when all elements of S have same probability, then p is the *uniform distribution.*

Example: Flip a biased coin where the probability of H is twice the probability of T.

Since $p(H) + p(T) = 1$, this implies $p(H) = 2/3$ and $p(T) = 1/3$.

Example: Flip a fair coin twice.

What is the probability of getting one H and one T?

Possible outcomes are $S = \{HH, HT, TH, TT\}$.

Two out of the possible 4 outcomes give one H and one T,

each outcome has probability $1/4$, so the total probability is $1/4 + 1/4 = 1/2$

Example: flip a fair coin 4 times.

What is the probability of getting exactly two heads?

The set of all possibilities

$$S = \{HHHH, HHHT, HHTH, HHTT, HTHH, HTHT, HTTH, HTTT \\ THHH, THHT, THTH, THTT, TTHH, TTHT, TTTH, TTTT\}$$

has size $|S| = 2^4 = 16$

There are six ways of getting exactly two heads,

$$E = \{HHTT,\ HTHT,\ HTTH,\ THHT,\ THTH,\ TTHH\},$$

so $p(E) = \sum_{a \in E} p(a) = |E|/|S| =6/16=3/8$.

Example: Suppose we flip a fair coin 10 times.

What is the probability of getting exactly 4 Hs?

Wait, this is getting tedious ...

Learn to Count I

The number of ways of rearranging k objects is called "k factorial":

$$k! = k(k-1)(k-2)\cdots 2\cdot 1$$

because there are k choices for the first, $k-1$ for the second, and so on (and where by definition $0! = 1$)

Example: There are $4! = 4\cdot 3\cdot 2\cdot 1 = 24$ ways of rearranging four objects

abcd abdc acbd acdb adbc adcb bacd badc bcad bcda bdac bdca

cabd cadb cbad cbda cdab cdba dabc dacb dbac dbca dcab dcba

Learn to Count II

Permutations: The number of ways to choose k objects from a set of n objects is given by

$$n(n-1)\cdots(n-k+1) = \frac{n!}{(n-k)!}$$

since there are n choices for the first object down to $n-k+1$ choices for the k^{th} object (after having chosen the first $k-1$ objects).

Example: There are $4 \cdot 3 = 4!/2! = 12$ ways of choosing two objects from four.

ab ac ad   ba bc bd   ca cb cd   da db dc

Learn to Count III

Combinations: If the order in which the objects are chosen does not matter, then the number of ways (combinations) to choose k objects from a set of n is given by dividing the above by $k!$ (the number of ways of rearranging those k objects).

The number of ways to choose k objects from n, independent of order, is thus given by:

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}$$

because we now divide by the k choices for the first object, $k-1$ choices for the second, and so on, down to two choices for the last two, and a single choice for the last remaining.

$\binom{n}{k}$ is called the "binomial coefficient", and pronounced "n choose k"

Example: There are $4 \cdot 3/2 = 4!/2!2! = 6$ ways of choosing two objects from four, if order doesn't matter (ab = ba):

ab ac ad   ba bc bd   ca cb cd   da db dc

Natural example of the distinction between permutations and combinations:

Deal from a standard 52 card deck

If one card to each of four players, then the number of possibilities is $52 \cdot 51 \cdot 50 \cdot 49 = 52!/48! = 6497400$.

If instead four cards to a single player, then "the hand" doesn't depend what *order* they're dealt, and the number of distinct possibilities is $52 \cdot 51 \cdot 50 \cdot 49/4 \cdot 3 \cdot 2 \cdot 1 = 52!/48!4! = \binom{52}{4} = 270725$.

Binomial Coefficients

$\binom{n}{k} = \frac{n!}{(n-k)!k!}$ is the number of ways of choosing k objects from n, independent of order

Note that $\binom{n}{k} = \binom{n}{n-k}$ (same count if we're choosing the $n-k$ to exclude)

These numbers are called *binomial coefficients*, because they appear as coefficients in the expansion of binomials (expressions of the form $(x+y)^n$).

Consider $(x+y)^2 = x^2 + 2xy + y^2$.

The coefficients of this polynomial are $\{1, 2, 1\}$, i.e., the numbers $\binom{2}{0}$, $\binom{2}{1}$, $\binom{2}{2}$.

In general, $(x+y)^n = \binom{n}{0}x^n + \binom{n}{1}x^{n-1}y + \binom{n}{2}x^{n-2}y^2 + \ldots + \binom{n}{n-1}xy^{n-1} + \binom{n}{n}y^n$

(each term contains a total of n x's and y's, and the number of times the term $x^k y^{n-k}$ occurs in the expansion is given by the number of combinations of n x's and y's with exactly k x's)

Repeat Example: flip a fair coin 4 times.

What is the probability of getting exactly two heads?

The set of all possibilities has size $|S| = 2^4 = 16$

There are $\binom{4}{2} = 4!/2!2! = 4 \cdot 3/2 = 6$ ways of getting exactly two heads,

$E = \{HHTT,\ HTHT,\ HTTH,\ THHT,\ THTH,\ TTHH\}$,

so $p(E) = \sum_{a \in E} p(a) = |E|/|S| = 6/16 = 3/8$.

Example: Flip a fair coin 10 times.

What is the probability of getting exactly four Hs?

First compute $\binom{10}{4} = \frac{10 \cdot 9 \cdot 8}{4 \cdot 3 \cdot 2} = 210$

(counts all $E = \{TTTTTTHHHH, ..., HTTHTHTTTH, ...HHHHTTTTTT\}$ places that the four Hs can occur)

The total number of outcomes is $2^{10} = 1024$.

The probability of getting exactly four Hs is $|E|/|S| = 210/1024 \approx .205$.

In General: Flip a fair coin n times.

How many possible outcomes?

Two choices for each flip of the coin, so 2^n possible outcomes.

Each coin flip an independent event (notion soon to be made precise), so the probability of any one of these is $1/2^n$.

Now suppose we want the probability of getting exactly k Hs.

How many of the 2^n strings have exactly k Hs?

$\binom{n}{k} = \frac{n!}{k!(n-k)!}$ counts number of ways the k H's can be distributed among the n tosses.

So the probability of k H's in n flips is $\binom{n}{k} \Big/ 2^n$.

Slightly more generality:

If the coin is biased, with probability p for H (and hence probability $1 - p$ for T), then

$$\text{probability}(k\ H\text{s in } N \text{ flips}) = \binom{N}{k} p^k (1-p)^{N-k}$$

where

p^k is the probability of k Hs,

$(1-p)^{N-k}$ is the probability of $N - k$ Ts,

and $\binom{N}{k}$ counts the number of ways that k Hs can be distributed among the N flips.

(or if $p = 1/6$, counts the number of ways that k 6s can be distributed among N rolls of a die)

Independent Events: joint probability is equal to the product of independent probabilities

Suppose event A in sample space S_1 has $|A| = m$ equal probability elements, and $|S_1| = M$, so $p(A) = m/M$

(examples: roll a fair die, flip a fair coin)

Suppose event B in sample space S_2 has $|B| = n$ equal probability elements, and $|S_1| = N$, so $p(B) = n/N$

Then the combined sample space $S_1 \times S_2$ has MN elements.

[Recall $S_1 \times S_2$ is the set of all pairs (s_1, s_2) for $s_1 \in S_1$ and $s_1 \in S_2$]

Of those, there are mn elements with $s_1 \in A$ and $s_2 \in B$, so

$$p(A \cap B) = \frac{mn}{MN} = \frac{m}{M} \cdot \frac{n}{N} = p(A) \cdot p(B)$$

Example: Roll four dice, what is the probability of at least one six?

a) Consider the complement problem:

there is a $5/6$ probability of not rolling a six for any given die, and since the four dice are independent,

the probability of not rolling a six is $(5/6)^4 = 5^4/6^4 = 625/1296$.

The probability of rolling at least one six is therefore $1 - 625/1296 = 671/1296 \approx .517$

Example: Roll four dice, what is the probability of at least one six?

b) Alternatively, recall that the number of ways of choosing r objects from a collection of N is $\binom{N}{r} = N!/r!(N-r)!$.

Any of the four dice can be the one that comes up six, and the other three don't, so the number of ways that

exactly one of the four dice is six is $\binom{4}{1} \cdot 5^3 = 4 \cdot 5^3 = 500$

exactly two sixes: $\binom{4}{2} \cdot 5^2 = (4 \cdot 3/2) \cdot 5^2 = 150$

exactly three sixes: $\binom{4}{3} \cdot 5 = 4 \cdot 5 = 20$

exactly four sixes: $\binom{4}{4} = 1$

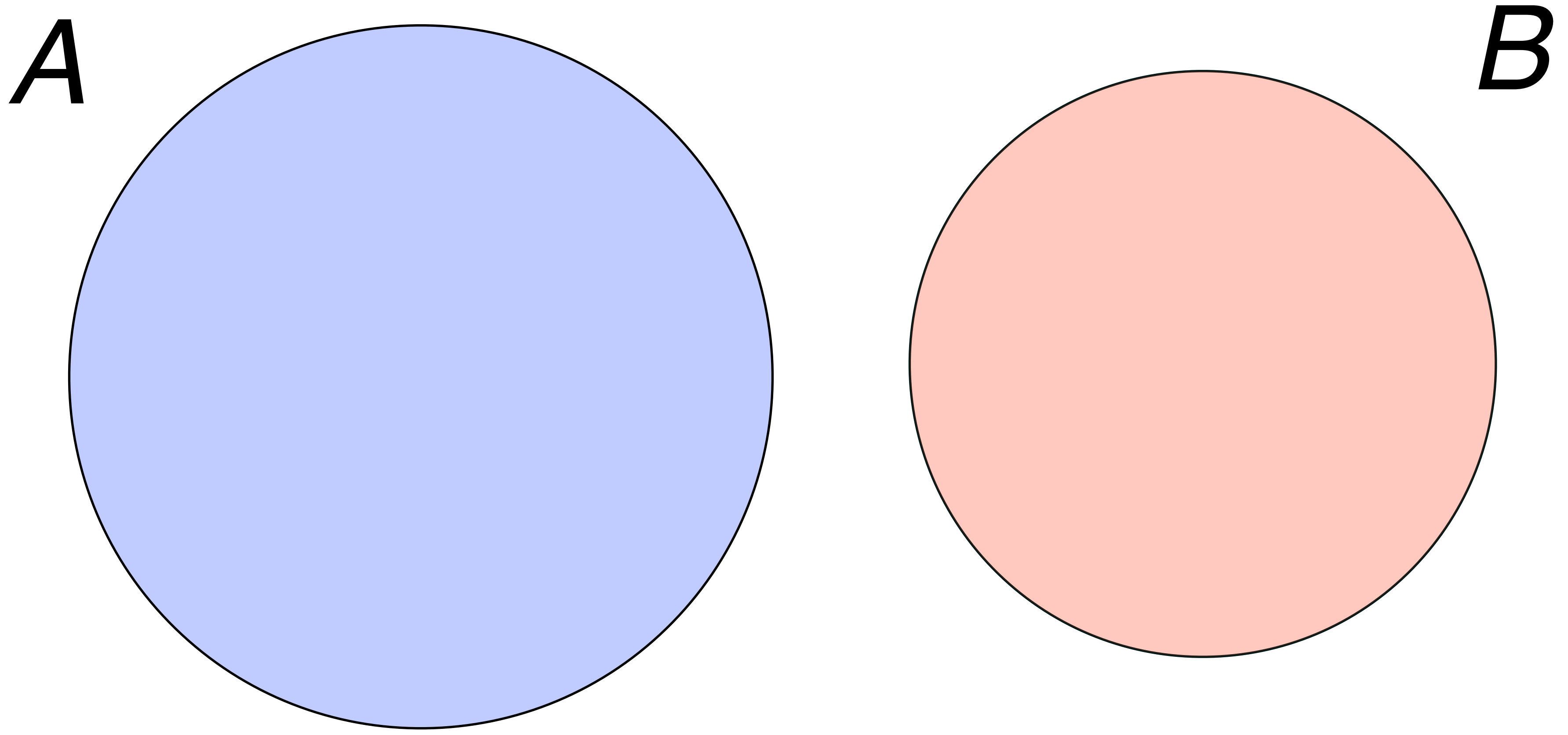
The total number of possibilities is $500 + 150 + 20 + 1 = 671$, and hence the probability is $671/6^4 = 671/1296$, as on previous slide.

Recall that events are sets. Two events are *disjoint* if their intersection is empty.

Example: In the example of flipping 2 coins, the events

$A =$ 'getting exactly one H'

$B =$ 'getting exactly $2H$s'



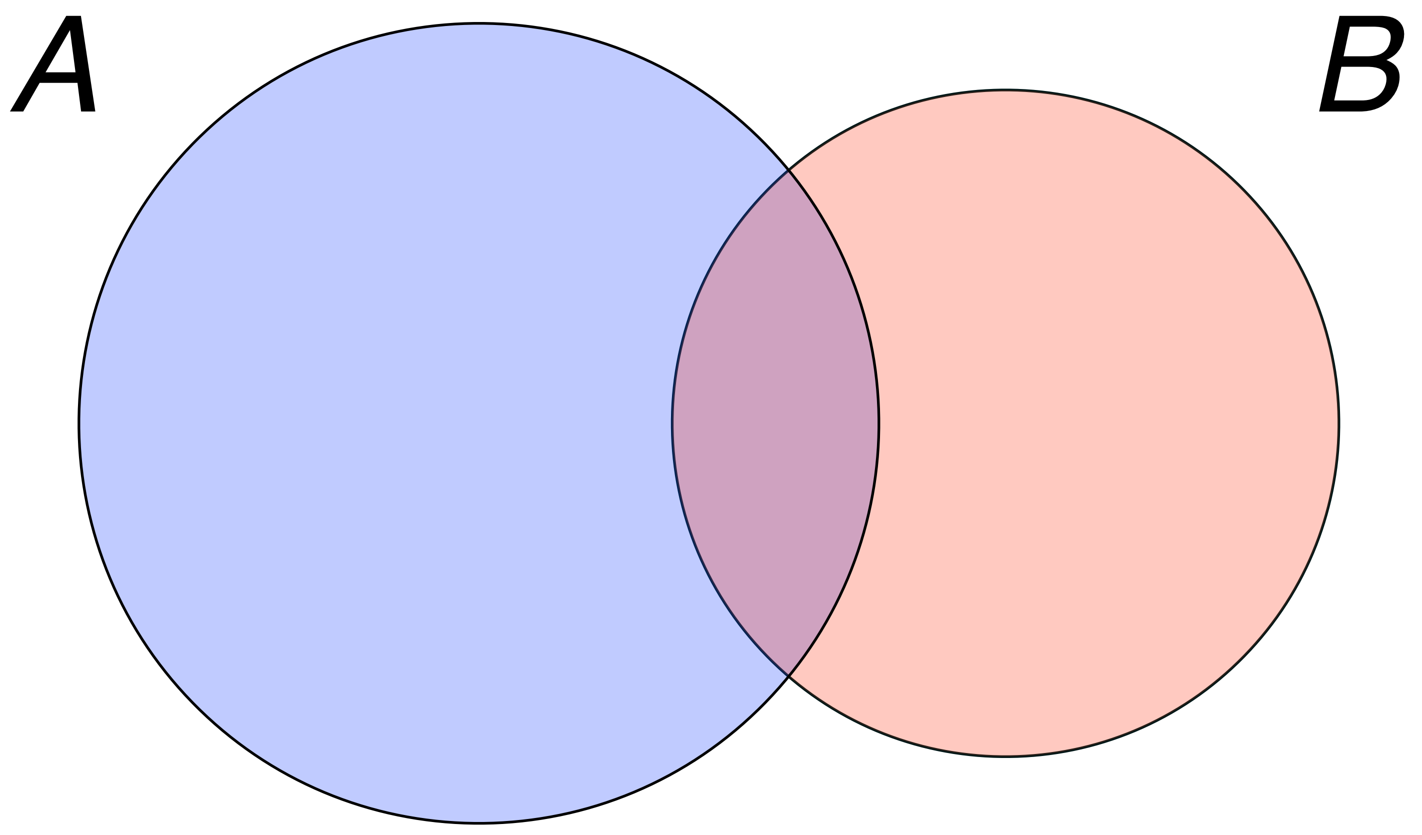
are disjoint. But, A is not disjoint from the event

$C =$ 'getting exactly one T'.

In fact, events A and C are the same in this case.

Principle of *inclusion - exclusion*:

$$|A \cup B| = |A| + |B| - |A \cap B|.$$



This gives

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

because

$$\sum_{a \in A \cup B} p(a) = \sum_{a \in A} p(a) + \sum_{a \in B} p(a) - \sum_{a \in A \cap B} p(a)$$

Therefore, for (and only for) disjoint events we have: $p(A \cup B) = p(A) + p(B)$

Example: Flip a coin 10 times.

What is the probability that the first flip is a T or the last flip is a T?

The number of outcomes with the first flip T is 2^9.

The number of outcomes where the last flip is a T is 2^9.

The number of strings with both properties is 2^8.

Hence, the number of strings with either property is $2^9 + 2^9 - 2^8 = 768$,

and the probability of first or last T is $768/1024 = .75$.