PROBABILITY SPACES

Definition 5.2.1 A sample space S is a non-empty countable set. Each element of S is called an outcome and each subset of S is called an event.


Your weather when it really matters™

	Mainly sunny	Mainly sunny	Cloudy with sunny breaks	Mainly sunny	Flurries	Mainly sunny	A mix of sun and clouds
	-7°	-10°	-14°	-14°	-13°	-12°	-9°
Feels like	-13	-16	-19	-19	-18	-18	-12
Night	-16°	-21°	-22°	-18°	-21°	-19°	-11°
POP	20 %	20 %	30 %	20 %	60 %	20 %	30 %
Wind (km/h)	16 W	14 NW	10 N	10 NW	11 NE	16 W	10 E
Wind gust (km/h)	24	21	15	15	17	24	15
Hrs Of Sun	8 h	8 h	2 h	6 h	3 h	6 h	4 h
24 Hr Snow	-	-	-	1-3 cm	1-3 cm	-	-

Here we see %. 20% POP

In probability theory, we want numbers in the interval [0,1].

So, 20% corresponds to 0.2

EXAMPLE: DIE ROLLING

A sample space S is a non-empty countable set. $S = \{1,2,3,4,5,6\}$

Each element of S is called an outcome e.g.: 3

and each subset of S is called an event. e.g., $\{2,4,6\}$ rolling an even number

PROBABILITY SPACES

Definition 5.2.2 Let S be a sample space. A probability function on S is a function $\Pr : S \to R$ such that for all $\omega \in S$, $0 \le \Pr(\omega) \le 1$, and $\sum_{\omega \in S} \Pr(\omega) = 1$.

For any outcome ω in the sample space S, we will refer to $\Pr(\omega)$ as the probability that the outcome is equal to ω.

Definition 5.2.3 A probability space is a pair (S, Pr), where S is a sample space and

$$\Pr : S \to R \text{ is a probability function on } S.$$

PROBABILITY SPACES DEFS CONT'D

A probability function $\Pr : S \to R$ maps each element of the sample space S (i.e., each outcome) to a real number in the interval [0, 1].

Extension: map any event to a real number in [0, 1].

If A is an event (i.e., $A \subseteq S$), then we define $\Pr(A) = \sum_{\omega \in A} \Pr(\omega)$ (5.1)

Pr(A) is the probability that the event A occurs.

PROBABILITY SPACES

Note: Note that since $S \subseteq S$, the entire sample space S is an event and

$$\Pr(S) = \sum_{\omega \in S} \Pr(\omega) = 1.$$

Why?

Time for some examples!

EXAMPLE 1: FLIPPING A COIN

It is useful to see this example to get the terminology clear on a familiar example.

Flipping an unbiased coin. So what is S? what is Pr? on what is Pr defined?

EXAMPLE: COIN FLIP – SAMPLE SPACE, EVENT, OUTCOME

Sample space $S = \{H, T\}$, where H = heads and T: tails are the outcomes.

$Pr : S \to R$ is defined as $Pr(H) = ½$ and $Pr(T) = ½$

Does Pr satisfy the Definition 5.2.2?

Sample space has two elements. there are four events, one event for each subset.

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

EXAMPLE: FLIPPING A COIN - PROBABILITY FUNCTION

1. $\Pr(\emptyset) = 0$,
2. $\Pr(\{H\})$
3. $\Pr(\{T\}) = \Pr(T) = 1/2, = \Pr(H) = 1/2$,
4. $\Pr(\{H, T\}) = \Pr(H) + \Pr(T) = 1/2 + 1/2 = 1$.

EXAMPLE 2: FLIPPING A COIN TWICE

We are now studying what happens if we flip a fair coin twice.

How many outcomes are there?

What is S?

What is Pr?

EXAMPLE 2: FLIPPING A COIN TWICE

Sample space: $S = \{HH, HT, TH, TT\}$

e.g. TH means that you first see tails then heads in the second toss

$Pr : S \to R$ is given by $Pr(HH) = Pr(HT) = Pr(TH) = Pr(TT) = 1/4$

Does it satisfy Definition 5.2.2?

EXAMPLE 2: FLIPPING A COIN TWICE

Note that the sample space consists of 4 elements,

The number of events is equal to $2^4 = 16$.

For example, $A = \{HT, TH\}$ is an event

By (5.1) $Pr(A) = \sum_{\omega \in A} Pr(\omega)$ we get that

$$Pr(A) = Pr(HT) + Pr(TH) = 1/4 + 1/4 = 1/2.$$

What does this mean now in words?

EXAMPLE 2: FLIPPING A COIN TWICE

Pr(A) = Pr(HT) + Pr(TH) = 1/4 + 1/4 = ½ means that the probability of getting one heads and one tails (in any order) is equal to ½.

EXAMPLE: ROLLING A DIE TWICE

Die rolled once: outcomes $\{1, 2, 3, 4, 5, 6\} = S$. $\Pr(i) = 1/6$ for $i=1,\ldots,6$

Die rolled twice: outcomes are now pairs with sample space

$S = \{(i,j) \mid i = 1,\ldots,6, \text{ and } j = 1,\ldots,6\}$, where i is the result of the first and j the result of the second roll. Assuming a fair die, each outcome has the same probability.

Since $\Pr(S) = \sum_{\omega \in S} \Pr(\omega) = 1$, we get that $\Pr(i,j) = 1/36$ for each outcome $(i,j) \in S$.

EXAMPLE: ROLLING A DIE TWICE

Now let us assume that we are interested in the sum of the results of the two rolls.

Event A_k is defined as "the sum of the results of the two rolls is equal to k"

$$A_k = \{(i,j) \in S \mid i+j = k\}.$$

E.g., $A_4 = \{(1,3), (2,2)\ (3,1)\}$.

EXAMPLE: SUM OF TWO ROLLS

Matrix view of sum of two rolls.

A_k is non-empty iff

$k \in \{2,3, ..., 12\}$

2nd roll

1st roll

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

recall: iff means if and only if

EXAMPLE: SUM OF TWO ROLLS

Now (5.1) $\Pr(A) = \sum_{\omega \in A} \Pr(\omega)$

$$A_k = \sum_{\omega \in A_k} \Pr(\omega) = \sum_{(i,j) \in A_k} \Pr(i,j) = \sum_{(i,j) \in A_k} 1/36 = |A_k|/36$$

For the previous case of k=4 we get:

$$A_4 = |A_4|/36 = 3/36 = 1/12$$

This can be calculated now for all k = 2,..., 12

EXAMPLE: SUM OF TWO ROLLS

$$
\begin{aligned}
\Pr(A_2) &= 1/36, \\
\Pr(A_3) &= 2/36 &&= 1/18, \\
\Pr(A_4) &= 3/36 &&= 1/12, \\
\Pr(A_5) &= 4/36 &&= 1/9, \\
\Pr(A_6) &= 5/36, \\
\Pr(A_7) &= 6/36 &&= 1/6, \\
\Pr(A_8) &= 5/36, \\
\Pr(A_9) &= 4/36 &&= 1/9, \\
\Pr(A_{10}) &= 3/36 &&= 1/12, \\
\Pr(A_{11}) &= 2/36 &&= 1/18, \\
\Pr(A_{12}) &= 1/36.
\end{aligned}
$$

NON-UNIQUENESS OF SAMPLE SPACES

A sample space is not necessarily uniquely defined.

Take the sum of two die example. We could have defined S'={2,3,..,12} as sample space.

Then defined a probability function Pr' for the sample space via:

$$Pr'(k) = Pr\ (A_k).$$

Why does this work? Pr'(k) is the probability of getting outcome k in sample space S'.

This is the same as the probability of event A_k to occur in sample space S.

Easy to verify that Pr satisfies the condition of Def 5.2.2. Therefore, Pr' is a (valid) probability function.

BASIC RULES OF PROBABILITIES

Assume that (S, Pr) is a probability space.

The empty set $\emptyset$ is an event. Intuitively, $Pr(\emptyset)$ must be zero. Formally,

<u>Lemma 5.3.1</u> $Pr(\emptyset) = 0$.

Proof. By (5.1), we have $Pr(\emptyset)) = \sum_{\omega \in \emptyset} Pr(\omega)$

Since there are zero terms in this summation, its value is equal to zero. q.e.d.

LEMMA: PAIRWISE DISJOINT EVENTS

Definition:

- Two events A and B are disjoint, if $A \cap B = \emptyset$.
- A sequence $A_1, A_2, \ldots, A_n$ of events is pairwise disjoint, if all pairs in this sequence are disjoint.

Lemma 5.3.2 If $A_1, A_2, \ldots, A_n$ is a sequence of pairwise disjoint events,

then $\Pr(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} \Pr(A_i)$.

PROOF: PAIRWISE DISJOINT EVENTS

Let $A = A_1 \cup A_2 \cup \cdot \cdot \cdot \cup A_n$

$\Pr(A) = \sum_{\omega \in A} \Pr(\omega) = \sum_{i=1}^{n} \sum_{\omega \in A_i} \Pr(\omega) = \sum_{i=1}^{n} \Pr(A_i)$ q.e.d.

Example: What is the probability that the sum of the results of two rolls of a die is even?

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

18 out of the 36 entries are even.
So, the probability of an even sum is ½.

EXAMPLE CONT'D

A more formal argument for this:

Consider the sample space $S = \{(i, j) : 1 \leq i \leq 6, 1 \leq j \leq 6\}$, where i is the result of the first roll and j is the result of the second roll. Recall that each of the 36 elements of S has the same probability, 1/36.

We are interested in the events: $A = \{(i, j) \in S \mid i + j \text{ is even}\}$.

When is a sum of two numbers even?

CONT'D

The sum of two numbers is even? if either both are odd or both are even,

$A_1 = \{(i, j) \in S \mid \text{both } i \text{ and } j \text{ are even}\}$ and $A_2 = \{(i, j) \in S \mid \text{both } i \text{ and } j \text{ are odd}\}$.

Since A_1 and A_2 are a partition of A, we get:

$$Pr(A) = Pr(A_1) + Pr(A_2).$$

We can easily verify that both A_1 and A_2 have 9 (3*3) elements each. Therefore,

$$Pr(A) = Pr(A_1) + Pr(A_2) = 9/36 + 9/36 = ½.$$

“COMPLEMENT” EVENT PROBABILITIES

If A is an event, then $\bar{A}$ denotes its complement, i.e., $\bar{A}$ = S \ A.

Either the event occurs or it does not, so the sum of Pr(A) and Pr($\bar{A}$) is equal to one.

Lemma 5.3.3 For any event A, Pr(A) = 1 – Pr ($\bar{A}$).

Proof We know that: 1 = Pr(S) = Pr(A ∪ $\bar{A}$) = Pr(A) + Pr ($\bar{A}$) by Lemma 5.3.2. q.e.d.

Application: the probability of getting an odd number if we roll a die twice is ½.

(using the previous result on even numbers and the Lemma)

"INCLUSION/EXCLUSION" FOR PROBABILITIES

Lemma 5.3.4 If A and B are events, then $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$.

Proof Note first that: $B \setminus A$ and $A \cap B$ are disjoint and $B = (B \setminus A) \cup (A \cap B)$ (see class)

Thus, again by Lemma 5.3.2 we obtain: $\Pr(B) = \Pr(B \setminus A) + \Pr(A \cap B)$.

Trivially, A and $B \setminus A$ are disjoint. Since $A \cup B = A \cup (B \setminus A)$, we yet again by Lemma 5.3.2 we obtain $\Pr(A \cup B) = \Pr(A) + \Pr(B \setminus A)$. Thus,

$\Pr(A \cup B) = \Pr(A) + \Pr(B \setminus A) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$ q.e.d.

EXAMPLE

Consider a sample space $S = \{1, 2, \ldots, 1000\}$, such that each element has the same probability, i.e., 1/1000, of being chosen. What is the probability that x is divisible by 2 or 3?

Again, like before, we consider two events

$A = \{i \in S \mid i \text{ is divisible by } 2\}$ and $B = \{i \in S \mid i \text{ is divisible by } 3\}$.

Now we need to compute $\Pr(A \cup B)$ by Lemma 5.3.4, is equal to

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B).$$

So, we need $\Pr(A)$, $\Pr(B)$, and $\Pr(A \cap B)$.

EXAMPLE

For Pr(A), A = {i ∈ S | i is divisible by 2} we need the even numbers of which there are 500/1000 so Pr(A) = 500/1000 = 1/2

Pr(B) B = {i ∈ S | i is divisible by 3}. $\lfloor 1000/3 \rfloor$ = 333 elements in S that are divisible by 3.

so Pr(B) = 333/1000

Now how to do Pr(A ∩ B) ?

EXAMPLE

Calculate: $\Pr(A \cap B)$

i belongs to $A \cap B$ iff i is divisible by 6, i.e.,

$A \cap B = \{i \in S : i \text{ is divisible by } 6\}$

$\lfloor 1000/6 \rfloor = 166$ elements in S that are divisible by 6, we therefore get

$$\Pr(A \cap B) = 166/1000.$$

Now we have everything we need.

EXAMPLE CONCLUDED

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

$$= 500/1000 + 333/1000 - 166/1000 = 667/1000.$$

Let us continue our lemmas; these will help solve problems.

UNION BOUND

- **<u>Lemma 5.3.5 (Union Bound</u>**) For any integer $n \geq 1$, if $A_1, A_2, \ldots, A_n$ is a sequence of events, then $\Pr(A_1 \cup A_2 \cup \cdots \cup A_n) \leq \sum_{i=1}^{n} \Pr(A_i)$.

- The proof is a simple induction proof using Lemma 5.3.4 $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$ and realizing that we would have to subtract the intersection probability and if we remove this, the number can't get smaller. The book has the detail.

- Let us do a few more lemmas.

LEMMAS

Lemma 5.3.6 If A and B are events with $A \subseteq B$, then $Pr(A) \leq Pr(B)$.

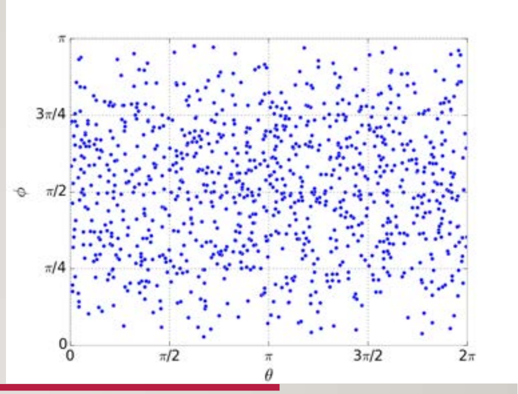
This makes intuitive a lot of senses. If you have all outcomes and possibly more in the second set, B, then this should have a higher probability. But this is not a proof.

Proof. Using (5.1) and the fact that $Pr(\omega) \geq 0$ for each ω in S, we have

$$Pr(A) = \sum_{\omega \in A} Pr(\omega) \leqq \sum_{\omega \in B} Pr(\omega) = Pr(B). \quad \text{q.e.d.}$$

because $A \subseteq B$

UNIFORM PROBABILITY SPACE



A finite sample spaces S in which each outcome has the same probability is called a uniform probability space. More formally,

<u>Definition</u>: 5.4.1 A uniform probability space is a pair (S, Pr), where S is a finite sample space and the probability function $Pr : S \to R$ satisfies $Pr(\omega) = 1/ |S|$, for each outcome ω in S.

So far, all but one probability space, we looked at were uniform. Which one was not?

UNIFORM PROBABILITY SPACE - EXAMPLE

Well the spaced defined by "sum of the results of two rolls of a die" was not a uniform probability space. Why? Verify yourself.

Here is another example of a uniform probability space:

Lotto 6/49, where you choose a 6-element subset of the set $A = \{1, 2, \ldots, 49\}$.

So, what has a higher chance of winning the ticket: 1, 2, 3, 4, 5, 6, or 4, 14, 22, 26, 45, 48?

LOTTO 6/49

S is consisting of all 6-element subsets of A. Since S has size $\binom{49}{6}$. Each subset is drawn with uniformly. Each outcome (6-element subset) has probability:

$$1/\binom{49}{6} = 1/\ 13,983,816 \approx 0.000000072.$$

So, your chances of winning are minimal.

Now, how about our two {1,2,3,4,5,6} and {4, 14, 22, 26, 45, 48} tickets? Same probability.

But then, 4, 14, 22, 26, 45, 48 was drawn on Feb 10^{th} , 2021.

The following Lemma is almost obvious.

UNIFORM PROBABILITY SPACES LEMMA

Lemma 5.4.2 If (S, Pr) is a uniform probability space and A is an event, then

$$\Pr(A) = |A| \,/\, |S|.$$

Proof $\Pr(A) = \sum_{\omega \in A} \Pr(\omega) = \sum_{\omega \in A} 1/|S| = 1/|S| \sum_{\omega \in A} 1 = |A| \,/\, |S|.$ q.e.d.

.

PROBABILITY OF GETTING A FULL HOUSE

In a standard deck of 52 cards, each card has a suit and a rank. There are four suits (spades , hearts , clubs , and diamonds), and 13 ranks (Ace, 2, 3, 4, 5, 6, 7, 8, 9, 10, Jack, Queen, and King).

Full house: three of the cards are of the same rank and the other two cards are also of the same (but necessarily different) rank. (5 cards in total)

So, how likely is it to get a full house?

PROBABILITY OF GETTING A FULL HOUSE

A poker hand is a five card subset of the set of all 52 cards.

S = {all poker hands} is the sample space. What is its size? $\binom{52}{5} = 2,598,960$

The sample space has a uniform distribution. 1/|S| per hand.

What is the probability of a random hand being a full house?

A = {all poker hands that are full houses}   Pr(A) = |A| / |S| (Lemma 5.4.2)

PROBABILITY OF GETTING A FULL HOUSE

To determine Pr(A) we need to know |A|.

Let us count the number of Full Houses. We know how to do this using the product rule from Chapter 3. We define our

Procedure "chose a full house'

1. Choose the rank of the three cards in the full house. $\binom{13}{1} = 13$ ways

2. Choose the suit of these three cards $\binom{4}{3}$ ways

Ace

, ,

PROBABILITY OF GETTING A FULL HOUSE

cont'd Procedure "chose a full house'

3, Chose the rank of the other two cards in the full house $\binom{12}{1} = 12$ King

4. Choose the suit of these two cards $\binom{4}{2}$,

So, by product rule we get: $|A| = 13 * \binom{4}{3} * 12 * \binom{4}{2} = 3,744.$

Result: $Pr(A) = {}^{|A|}/_{|S|} = {}^{3,744}/_{2,598,960} = 0.00144$

PROBABILITY OF GETTING A FULL HOUSE

So, $|A| = 13 * \binom{4}{3} * 12 * \binom{4}{2} = 3,744$.

We could also have first chosen the suits of these three cards $\binom{4}{3}$ then their rank 13 then, the suit $\binom{4}{2}$ for the pair followed by its rank 12.

Since the multiplication is commutative the result does not change.

SOME MORE EXAMPLES

Four of a kind

1. Choose the rank of four of a kind. $\binom{13}{1}$ = 13 ways
2. Choose the rank of the fifth card $\binom{12}{1}$ = 12 ways
3. Choose the suit of the fifth card $\binom{4}{1}$ = 4 ways

Note that after step 1. there is no need for a suit determination of that ranke as we pick all $\binom{4}{4}$ suites with that rank.

THREE OF A KIND - IS THIS RIGHT?

Three of a kind

1. Choose the rank of three of a kind. $\binom{13}{1}$ = 13 ways
2. Choose the suits of the three of a kind. $\binom{4}{3}$ = 4 ways
3. Choose the rank of the fourth card $\binom{12}{1}$ = 12 ways
4. Choose the suit of the fourth card $\binom{4}{1}$ = 4 ways
5. Choose the rank of the fifth card $\binom{11}{1}$ = 11 ways
6. Choose the suit of the fifth card $\binom{4}{1}$ = 4 ways

THREE OF A KIND - IS THIS RIGHT?

Three of a kind

1. Choose the rank of three of a kind $\binom{13}{1}$ = 13 ways
2. Choose the suits of the three of a kind $\binom{4}{3}$ = 4 ways
3. Choose the rank of the fourth&fifth card $\binom{12}{2}$ = 12 ways
4. Choose the suit of the fourth card $\binom{4}{1}$ = 4 ways
5. Choose the suit of the fifth card $\binom{4}{1}$ = 4 ways

THREE OF A KIND - WHICH ONE IS RIGHT?

The first method is not right. We are imposing an order for the last two cards.

so the order of the 4 and 7 matters.

THREE OF A KIND - WHICH ONE IS RIGHT?

We are imposing an order for the last two cards.

$$\frac{\binom{13}{1} * \binom{4}{3} * \binom{12}{1} * \binom{4}{1} * \binom{11}{1} * \binom{4}{1}}{\binom{52}{5}}$$

First

$$\frac{\binom{13}{1} * \binom{4}{3} * \binom{12}{2} * \binom{4}{1} * \binom{4}{1}}{\binom{52}{5}}$$

Second

$\binom{12}{1}\binom{11}{1}$ = 12*11 differ $\binom{12}{2}$ = 12*11/2

If we realize that the order is irrelevant we would need to divide 12*11 by 2

The number of distinct hands with three of a kind is: 54,912.

DO SOME EXERCISES YOURSELF

- e.g., one pair or two pairs
- or nothing

here is the solution why?

What are these subtractions

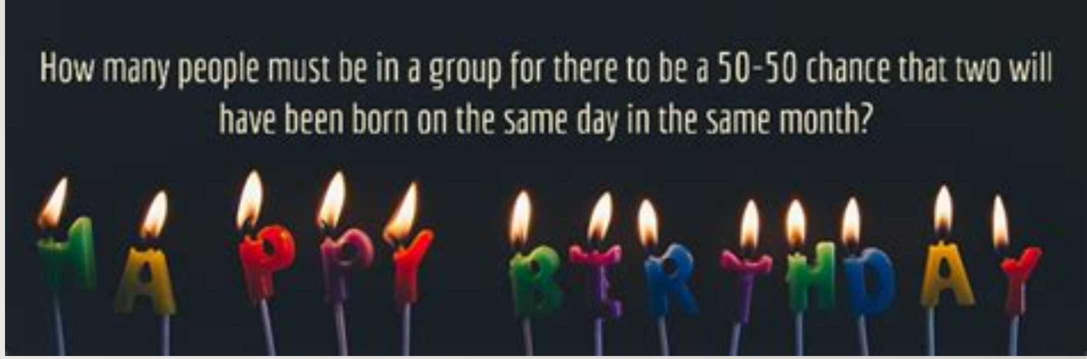
compensating for?

$$\left[\binom{13}{5} - 10\right]\left[\binom{4}{1}^5 - 4\right]$$

SOLUTIONS TO LOOK AT ONCE YOUR HAVE DONE THEM

Hand	Distinct hands	Frequency	Probability	Cumulative probability	Odds against	Mathematical expression of absolute frequency
Royal flush	1	4	0.000154%	0.000154%	649,739 : 1	$\binom{4}{1}$
Straight flush (excluding royal flush)	9	36	0.00139%	0.0015%	72,192 1/3 : 1	$\binom{10}{1}\binom{4}{1} - \binom{4}{1}$
Four of a kind	156	624	0.0240%	0.0256%	4,165 : 1	$\binom{13}{1}\binom{12}{1}\binom{4}{1}$
Full house	156	3,744	0.1441%	0.17%	693.17 : 1	$\binom{13}{1}\binom{4}{3}\binom{12}{1}\binom{4}{2}$
Flush (excluding royal flush and straight flush)	1,277	5,108	0.1965%	0.367%	508.8 : 1	$\binom{13}{5}\binom{4}{1} - \binom{10}{1}\binom{4}{1}$
Straight (excluding royal flush and straight flush)	10	10,200	0.3925%	0.76%	253.8 : 1	$\binom{10}{1}\binom{4}{1}^5 - \binom{10}{1}\binom{4}{1}$
Three of a kind	858	54,912	2.1128%	2.87%	46.33 : 1	$\binom{13}{1}\binom{4}{3}\binom{12}{2}\binom{4}{1}^2$
Two pair	858	123,552	4.7539%	7.62%	20.0 : 1	$\binom{13}{2}\binom{4}{2}^2\binom{11}{1}\binom{4}{1}$
One pair	2,860	1,098,240	42.2569%	49.9%	1.366 : 1	$\binom{13}{1}\binom{4}{2}\binom{12}{3}\binom{4}{1}^3$
No pair / High card	1,277	1,302,540	50.1177%	100%	0.995 : 1	$\left[\binom{13}{5} - 10\right]\left[\binom{4}{1}^5 - 4\right]$
Total	**7,462**	**2,598,960**	**100%**	**—**	**0 : 1**	$\binom{52}{5}$

BIRTHDAY PARADOX



You are in a group of n, $n \geqq 2$, persons. What is the probability that at least 2 of you have the same birthday? We denote by p_n that probability.

<u>Assumption</u>: no leap year (i.e., 365 days)

What is p_2? Below you will see $p_2 = 1/365$.

What is p_n for $n>365$? That is easy $p_n = 1$. Why? Pigeonhole

So the remaining, interesting cases are clearly for $2<n \leqq 365$.

Seems intuitive that p_i increases monotonically for i =, 2, 3, ..., n

BIRTHDAY PARADOX – INITIAL OBSERVATIONS

Seems intuitive that p_n increases monotonically for n =, 2, 3, ..., 365

p_2 = 1/365, ..., $p_n, p_{n+1}, .., p_{365}$

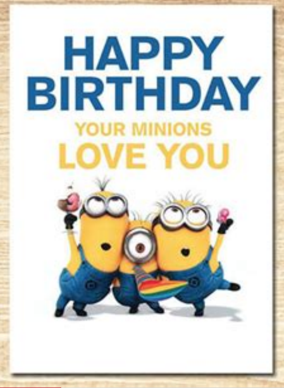
For which n do we get that p_n is the first time > ½?

Formally, determine n such that: $p_{n-1} \leqq ½ < p_n$.

Our counting strategies from Chapter 3 will help us determine n.

GUESS what the value of n might be!

BIRTHDAY PARADOX
– NOTATION & OBSERVATIONS



<u>Notation</u>: - we denote people by $P_1, P_2, \ldots, P_n$, and

- the number of days in one year by d, and
- we number the days in one year as $1, 2, \ldots, d$.
- sample space is set $S_n = \{(b_1, b_2, \ldots, b_n) : b_i \in \{1, 2, \ldots, d\}$ for each $1 \leq i \leq n\}$,

  where b_i denotes the birthday of P_i.

  Note that $|S_n| = d^n$

BIRTHDAY PARADOX – NOTATION & OBSERVATIONS

<u>Observation</u> Consider the <u>uniform</u> probability space:

for each element $(b_1, b_2, \ldots, b_n)$ in S_n, we therefore have

$$\Pr(b_1, b_2, \ldots, b_n) = 1/\,|S_n| = 1/\,d^n \quad .$$

We are interested in events where two have the same birthday, i.e.,

A_n = " at least two of the numbers $b_1, b_2, \ldots, b_n$ are equal"

in set notation: $A_n = \{\, (b_1, b_2, \ldots, b_n) \in S_n \mid b_1, b_2, \ldots, b_n \text{ contains duplicates}\}$.

BIRTHDAY PARADOX – $P_2 = 1/365$.

Recall, as introduced above, p_n is the probability that at least 2 have the same birthday

Then, $p_n = Pr(A_n)$ and $p_n = 1$, for $n > d$. Thus, we focus on $n \le d$.

So, what is p_2 formally derived as? Since we are considering the uniform probability space:

(by Lemma 5.4.2) $p_2 = Pr(A_2) = |A_2| / |S_2|$.

$|S_2| = d^2$

$A_2 = \{(1, 1), (2, 2), \ldots, (d, d)\}$ therefore $|A_2| = d$.

$p_2 = Pr(A_2) = |A_2| / |S_2| = d / d^2 = 1/d$ since $d = 365$ we get $p_2 = 1/365$.