Topic 2: Introduction to Probability

Optional Reading: Chapter 4

Xiner Zhou

Department of Statistics

University of California, Davis

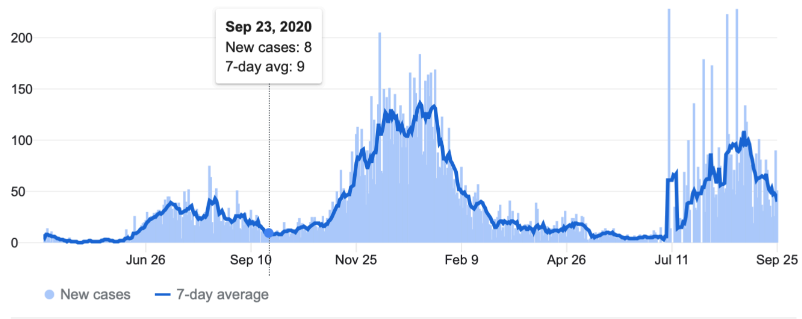
Part I: Probability

- Basic definitions in probability
- Counting rules for calculating probabilities
- Event Relations and Probability Rules

Why Study Probability?

Math is the logic of certainty, Probability is the logic of uncertainty.

casual conversation: “Luck, Coincidence, Random, Risk, Doubt, Chance……” — sense of Uncertainty



Despite its ubiquity, “probability” can be deeply counterintuitive.

Just reply on intuition, run serious risk of making inaccurate predictions or overconfident decisions.

That's why probability is extremely useful in a wide variety of fields:

- Foundation for Statistics:
  - Powerful methods of using data to learn about the world:
    - Explaining variation: why some people live longer, why some students have higher performance
    - separating signal from noise: blood tests for the same person in the morning and afternoon
    - Modeling complex phenomenon: how to build a model that predict the evolution of COVID-19
- Applications in:
  - Physics, Biology, Computer Science, Gambling, Finance, Economics,
  - Political Science, Medicine, Astronomy….

An experiment is the process by which an observation (or measurement) is obtained.

- It is basically a method of data collection.
- Data are obtained by observing either uncontrolled events in nature or by observing events in controlled situations. We use the term experiment to describe either method of data collection.
- Examples:
    - Recording a test grade
    - Measuring daily rainfall
    - Interviewing a householder to obtain his or her opinion on …
    - Tossing a coin and observing the face that appears

The sample space S of an experiment is: the collection of all possible outcomes of the experiment.

- Example:
  - Flip a coin S={H,T}

A simple event (also called a sample point) is the outcome that is observed on a single repetition of the experiment.

- Example:
  - Flip a coin: H, T both are a simple event or a sample point

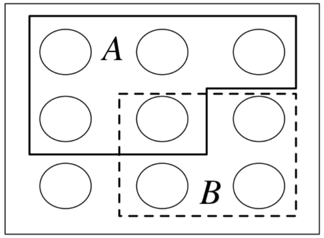
An event A is a collection of simple events.

- Example:
  - Getting a head: A={H}

We say that event A occurred or happened if: the actual outcome is in A

Historically, the earliest definition of probability of an event A was:

$$\frac{\text{number of ways the event could happen}}{\text{Total number of possible outcomes for the experiment}}$$



- Reply on assumptions that
  - every outcome is equally likely to happen:
    - when is/isn't appropriate?
    - You want to ask a friend for help, is everyone you know equally likely to be called?
  - Only have finite number of possible outcomes of an experiment:
    - when is/isn't appropriate?
    - Throw a coin into a fountain, probability inside a small circle
- Nevertheless, it's important to understand the first definition of probability and its many applications

Let A be an event for an experiment, the probability of A is:

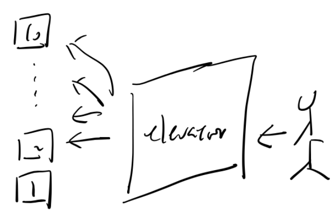
$$P(A) = \frac{|A|}{|S|}$$

|A|: number of simple events in A

|S|: number of simple events in S

Elevator: people are equal likely to go to floor 2-10

Calculating probability of an event = Counting!



S={2,3,…,10}

A: go to top floor

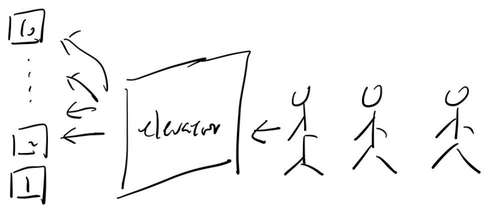
A={10} => P(A)=1/9

B: go to even floor

B={2,4,6,8,10) => P(B)=5/9

Elevator: 3 people are equal likely to go floor 2-10

What is the probability that buttons for 3 consecutive floors are pressed?



S={222,223,224,225,……….}

A={234,324,423,432,…………}

Wait… Can you imagine and count the sample space and the event A?

How to Count? When the sample space or the event is not so simple

Counting Rules for Probability

- Multiplication rule (the fundamental counting rule)

  If an experiment is performed in k stages,

  with n_1 ways to accomplish the first stage,

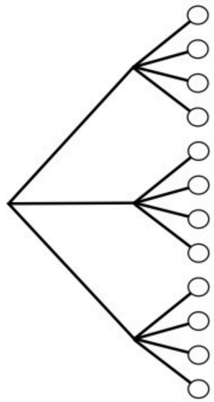
  n_2 ways to accomplish the second stage, ...,

  and n_k ways to accomplish the kth stage,

  then the number of ways to accomplish the experiment is

  $n_1 n_2 n_3 \cdots n_k$

Why:



If an experiment can be performed in 2 stages,

There is 3 ways to do the first stage, and 3 ways to do the second stage,

Then whichever you choose to do the first stage,

there are 3 multiple ways to do the second stage,

So, in total, if we count the number of circles in the end which represents unique ways to do the experiment

3+3+3=3 * 3 = n1 * n2

Note: The multiple rules says that an experiment can be break down into a sequence of sub-tasks,

It seems to suggest that the sub-tasks are taken in a specific order:

Sub-task 1-> sub-task 2 -> sub-task 3 ……

But: There is no requirement that sub-tasks are taken in a particular time order!

It is often easier to think about experiments in chronological order,

just to help you clarify mental barriers and think in a logical way.

A truck driver can take three routes from city Ato city B, four from city B to city C, and three from city C to city D.

If, when traveling from A to D, the driver must drive from A to B to C to D, how many possible A-to- D routes are available?

$$A \longrightarrow B \longrightarrow C \longrightarrow D$$

Solution: To go from A -> D, we have to complete the following sub-tasks:

Sub-task 1: A to B

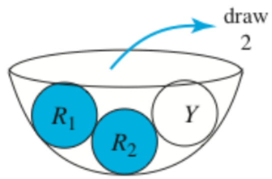
Sub-task 2: B to C

Sub-task 3: C to D

=> the total number of possible ways A-D = n1*n2*n3=3*4*3=36

(Think about how long it takes you to enumerate all possibility by hand…)

A candy dish contains one yellow and two red candies. Two candies are selected one at a time from the dish, and their colors are recorded. How many simple events are in the sample space S?



Solution: subtask1: choose 1st candy in n1=3 ways

Subtask2: choose the 2nd candy in n2=3-1=2 ways, since one candy

is now gone

=> total number of ways two candies can be chosen =n1*n2=3*2=6

- How many simple events are in the sample space when three coins are tossed?

  Solution:

  subtask 1: flip the 1st coin in n1=2

  subtask 2: flip the 1st coin in n2=2

  subtask 3: flip the 1st coin in n3=2

  => the total number of possible results we can get from flipping 3 coins is n1*n2*n3=2*2*2=8

You own 4 pairs of jeans, 12 clean T-shirts, and 4 wearable pairs of sneakers.

How many outfits (jeans, T-shirt, and sneakers) can you create?

Solution:

subtask 1: choose jeans in n1=4 ways

Subtask 2: choose T-shirt in n2=12 ways

Subtask 3: choose sneakers in n3=4 ways

=> how many outfits you could create? n1*n2*n3=4*12*4= 192!

No repeat for half of the year…

Your family vacation involves a cross-country air flight, a rental car, and a hotel stay in Boston. If you can choose from four major air carriers, five car rental agencies, and three major hotel chains, how many options are available for your vacation accommodations?

Solution:

Subtask1: choose airlines in n1=4 ways

Substask2: choose car rental in n2=5 ways

Subtask 3: choose hotel in n3=3 ways

=> You have to choose from n1*n2*n3=4*5*3=60 available vacation accommodations!

Planning is exhausting…

There are 3 extremely common situations that cover most problems,

and can be dealt with by multiplication rule.

So it's a good idea to summarize them so that don't have to redo the same mental exercise every time.

We have: n distinct objects

Making k times of choices from them

1. When order matters & Sampling with replacement
    - Order matters: object 1 -> object 2 is different from object 2 -> object 1
    - Sampling with replacement: whichever object is taken, it will be returned to the mix so that it can be chosen again and again

- How many simple events are in the sample space when three coins are tossed?

  Solution:

  subtask 1: flip the 1st coin in n1=2

  subtask 2: flip the 1st coin in n2=2

  subtask 3: flip the 1st coin in n3=2

  => the total number of possible results we can get from flipping 3 coins is n1*n2*n3=2*2*2=8

We have seen this example of coin flip before:

There are 2 objects : Head or Tail to choose from

We make 3 times of choices

We care about the order we see heads/tails: order matters

No matter we saw tail/head in previous flip, we can still see tail/head in next flip: sampling with replacement

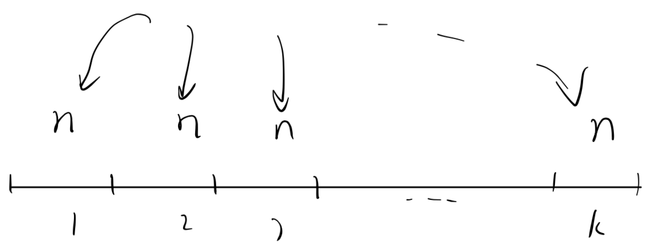
We have: n distinct objects

Making k times of choices from them

1. When order matters & Sampling with replacement

- Order matters: object 1 -> object 2 is different from object 2 -> object 1
- Sampling with replacement: whichever object is taken, it will be returned to the mix so that it can be chosen again and again

=> There are $n^k = n \times \ldots \times n$ possible outcomes



Due to basic multiplication rule!

We have: n distinct objects

Making k times of choices from them

2. When order matters & Sampling without replacement

- Order matters: object 1 -> object 2 is different from object 2 -> object 1
- Sampling without replacement: whichever object is taken, it will not be returned to the mix so that it can’t be chosen again

Subtask1: n1=n ways

Subtask2: n2=n-1 ways because 1 way has been chosen

Subtask 3:n3=n-2 ways

…….

Subtask k: n-(k-1)=n-k+1 ways because k-1 ways has been chosen

=> There are $n \times (n-1) \times \ldots \times (n-k+1)$ possible outcomes

$P_k^n = n \times (n-1) \times \ldots \times (n-k+1)$ is called Permutation

Read as "n permute k"

$P_k^n = \dfrac{k!}{(n-k)!}$ where $n! = n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1$ and $0! = 1$

Three lottery tickets are drawn from a total of 50 .

If the tickets will be distributed to each of three employees in the order in which they are drawn, the order will be important. How many simple events are associated with the experiment?

Solution: There are 50 distinct employee,

Choose 3 from them

Order matters & one employee can’t be chosen twice

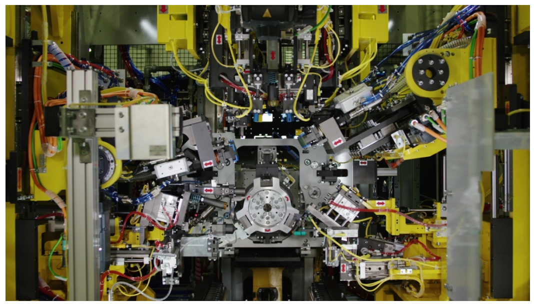
=> the Toal number of possible outcomes of this experiment is

$$P_3^{50} = \frac{50!}{47!} = 50(49)(48) = 117{,}600$$

(Think about how long it would take you to enumerate ……)

- A piece of equipment is composed of five parts that can be assembled in any order.

  A test is to be conducted to determine the time necessary for each order of assembly. If each order is to be tested once, how many tests must be conducted?



Solution: There are 5 distinct parts,

Choose 5 from them

Order matters & one part can't be chosen twice

=> the Toal number of possible configuration is

$$P_5^5 = \frac{5!}{0!} = 5(4)(3)(2)(1) = 120$$

A businessman in New York is preparing an itinerary for a visit to six major cities. The distance traveled, and hence the cost of the trip, will depend on the order in which he plans his route.

How many different itineraries (and trip costs) are possible?

Solution: There are 6 distinct cities,

Choose 6 from them

Order matters & one city must be chosen once and only once, no repeat

=> the Toal number of possible itineraries is

$$P_6^6 = \frac{6!}{0!} = 6(5)(4)(3)(2)(1) = 720$$

We have: n distinct objects

Making k times of choices from them

3. When order DOES NOT matters & Sampling without replacement

- Order does not matters: object 1 -> object 2 = object 2 -> object 1
- Sampling without replacement: whichever object is taken, it will be not returned so that it can't be chosen again

Compare with previous situation

2. When order matters & Sampling without replacement

Just one difference: whether order matters, (1,2) ?= (2,1)

Let's use the result from situation 2:

If order matters, then there are $P_k^n = \dfrac{k!}{(n-k)!}$ possible outcomes,

Where we count the following outcomes as different ones:

(1,2,3,…,k)   (2,1,3,…,k) …….. (k,k-1,k-2,…,1)

But in situation 3, we don't care about the order

So these different configurations should just count as 1!

=> we have over-counted the same outcome many many times

=> so we divide $P_k^n = \dfrac{k!}{(n-k)!}$ by how many times we have over-counted

=> There are $k!$ different configurations of (1,2,3,…,k)   (2,1,3,…,k) …….. (k,k-1,k-2,…,1)

=> meaning, we have over-counted 1 outcome for $k!$ times

=> adjust for over counting by dividing $k!$:

$$\frac{k!}{(n-k)!k!} := C_k^n$$

The number of k-combinations from n objects, also called binomial coefficient

Read as "n choose k"

A printed circuit board may be purchased from five suppliers.

In how many ways can three suppliers be chosen from the five?

Solution:

Choose 3 from 5, where order does not matter, and each supplier at most chosen once (sampling without replacement)

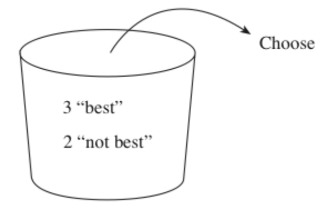
=> this is a problem of combination

=> $C_3^5 = \dfrac{5!}{3!2!} = \dfrac{(5)(4)}{2} = 10$

Five manufacturers produce a certain electronic device, whose quality varies from manufacturer to manufacturer.

If you were to select three manufacturers at random, what is the chance that the selection would contain exactly two of the best three?

Solution:



Let A denote the event that selection contains exactly two of the best three

$$P(A) = \frac{|A|}{|S|}$$

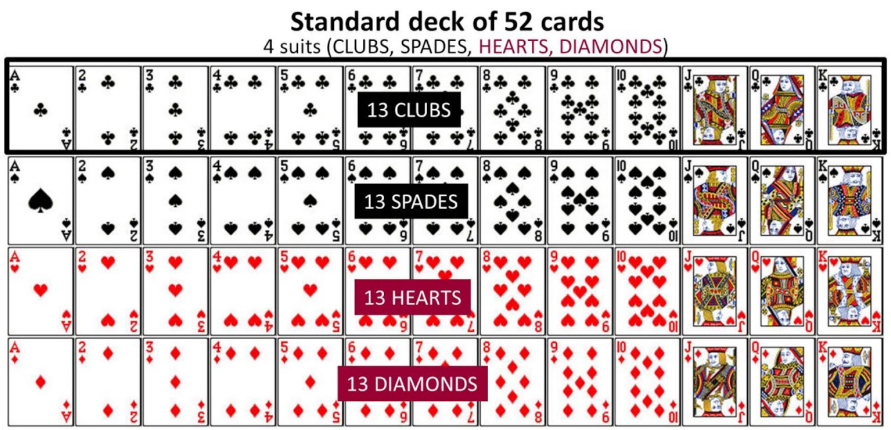
How to count |S| and |A|?

|S|=number of ways to choose 3 from 5= $C_3^5 = \frac{5!}{3!2!} = \frac{5 \times 4}{2} = 10$

|A|=|choose 2 best from 3|*|choose 1 not best from 2|=$C_3^2 C_2^1 = \frac{3!}{2!1!} \times \frac{2!}{1!1!} = 3 \times 2 = 6$

=> $P(A) = \frac{|A|}{|S|} = 0.6$

(Full House in Poker)



A 5-card hand is dealt from a standard, well-shuffled 52-card deck.

The hand is called a full house in poker if:

It consists of 3 cards of same rank + 2 cards of another rank

i.e. 3 7's and 2 10's in any order (order by which we draw cards does not matter)

What is the probability of a full house?

Solution:

Let A denote the event of a full house

$$P(A) = \frac{|A|}{|S|}$$

How to count ISI and IAI?

IAI: how can we have a full house?

Subtask 1: choose a rank from 13 ranks {A,2,3…K}

Subtask 2: choose 3 cards from 4 cards in the chosen rank

Subtask 3: choose a rank from the remaining 12 ranks

Subtask 4: choose 2 cards from 4 cards in the chosen rank

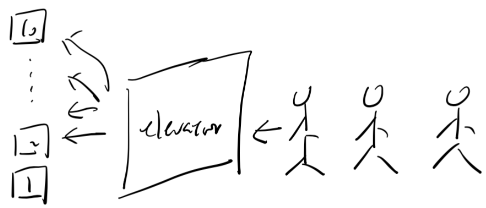
=> IAI=n1*n2*n3*n4= $13 \times \binom{4}{3} \times 12 \times \binom{4}{2} = 13 \times \frac{4!}{3!1!} \times 12 \times \frac{4!}{2!2!} = 3744$

ISI=choose 5 from 52=$\binom{52}{5} = \frac{52!}{5!47!} = 2598960$

=> $P(A) = \frac{|A|}{|S|} = 3744/2598960 = 0.001440576 = 0.144\,\%$

Elevator: 3 people are equal likely to go floor 2-10

What is the probability that buttons for 3 consecutive floors are pressed?



S={222,223,224,225,……….}

A={234,324,423,432,…………}

Wait… Can you imagine and count the sample space and the event A?

How to Count? When the sample space or the event is not so simple

General Definition of Probability and Properties

We've seen probabilities if naive definition of probability applies:

1. Finite sample space
2. Equal likely outcomes

To generalize probability to account for all situations.

General definition of Probability

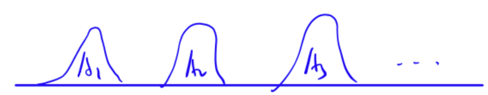
A probability P assigns an event A contained in the sample space S, a number between 0 and 1, and probability P must satisfy the following axioms:

1. $P(\varnothing) = 0, P(S) = 1$

mass of no sand = 0

total mass of the whole pile of sand = 1

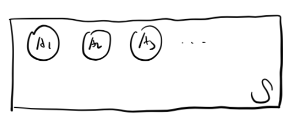
2. If $A_1, A_2, \ldots$ are disjoint events, then $P\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} P\left(A_j\right)$



total mass of non-overlapping piles of sand

= sum of individual mass

- Disjoint means: $A_i \cap A_j = \varnothing$ for $i \neq j$

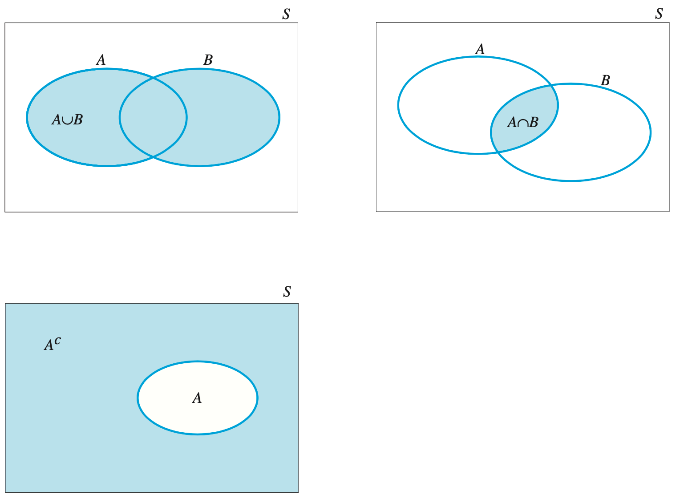


Event Relations and Probability Rules

- The union of events A and B, denoted by $A \cup B$, is the event that either A or B or both occur.

  The intersection of events A and B, denoted by $A \cap B$, is the event that both A and B occur.

  The complement of an event A, denoted by A^C, is the event that A does not occur.

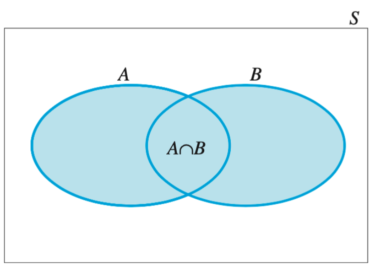


The Addition Rule for calculating probability of unions

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

If A and B are mutually exclusive, $P(A \cap B) = 0$. Then

$$P(A \cup B) = P(A) + P(B)$$



Rule for Complements

$$P\left(A^{c}\right) = 1 - P(A)$$

In a telephone survey of 1000 adults, respondents were asked their opinion about the cost of a college education. The respondents were classified according to whether they currently had a child in college and whether they thought the loan burden for most college students is too high, the right amount, or too little. The proportions responding in each category are shown as:

	Too High (*A*)	Right Amount (*B*)	Too Little (*C*)
Child in College (*D*)	.35	.08	.01
No Child in College (*E*)	.25	.20	.11

Suppose one respondent is chosen at random from this group.

Think each cell as a simple event, i.e. one possible outcome in a experiment

1. What is the probability that the respondent has a child in college?

Solution: respondent has a child in college={D&A, D&B, D&C}

P(respondent has a child in college) = sum of probabilities of simple events that satisfy “respondent has a child in college”=P(D&A)+P(D&B)+P(D&C)= .35+.08+.01=.44

2. What is the probability that the respondent does not have a child in college?

Solution: the event that the respondent does not have a child in college is the

Complement of A, denote as A^C

$$P(A^C) = 1 - P(A) = 1 - .44 = .56$$

3.What is the probability that the respondent has a child in college or thinks that the loan burden is too high or both?

Solution:

By the additive rule:

$$P(D \cup A) = P(D) + P(A) - P(D \cap A) = .44 + .6 - .35 = .69$$

1. Birthday problem There are 10 people in a room. Assume each person's birthday is equally likely to be any of the 365 days of the year (we exclude February 29), and that people's birthdays are independent (we will define independence formally later, but intuitively it means that knowing some people's birthdays gives us no information about other people's birthdays; this would not hold if, e.g., we knew that two of the people were twins).

    1. How many ways are there to assign birthdays to all the people in the room?
    2. Can you easily count how many ways to assign birthdays so that at least one pair of people in the group have the same birthday?
    3. Can you easily count how many ways to assign birthdays so that no two people share a birthday?
    4. What is the probability that at least one pair of people in the group have the same birthday?

Solution:

1. By the counting rule for sampling with replacement, there are 365^{10} ways to assign birthdays to the people in the room, since we can imagine the 365 days of the year being sampled k times, with replacement. By assumption, all of these possibilities are equally likely, so the naive definition of probability applies.

2. Used directly, the naive definition says we just need to count the number of ways to assign birthdays to 10 people such that there are two people who share a birthday.

   But this counting problem is hard, since it could be Emma and Steve who share a birthday, or Steve and Naomi, or all three of them, or the three of them could share a birthday while two others in the group share a different birthday, or various other possibilities.

3. Instead, let's count the complement: the number of ways to assign birthdays to 10 people such that no two people share a birthday. This amounts to sampling the 365 days of the year without replacement, using the counting rule for sampling without replacement, so the number of possibilities is $365 \cdot 364 \cdot 363 \cdots (365 - 10 + 1)$

4. Therefore the probability of no birthday matches in a group of k people is

$$P(\text{ no birthday match }) = \frac{365 \cdot 364 \cdots (365 - 10 + 1)}{365^{10}}$$

and the probability of at least one birthday match is

$$P(\text{ at least 1 birthday match }) = 1 - \frac{365 \cdot 364 \cdots (365 - 10 + 1)}{365^{10}} = 0.1169$$

The first value of k for which the probability of a match exceeds 0.5 is $k = 23$. Thus, in a group of 23 people, there is a better than 50% chance that there is at least one birthday match. At $k = 57$, the probability of a match already exceeds 99%.

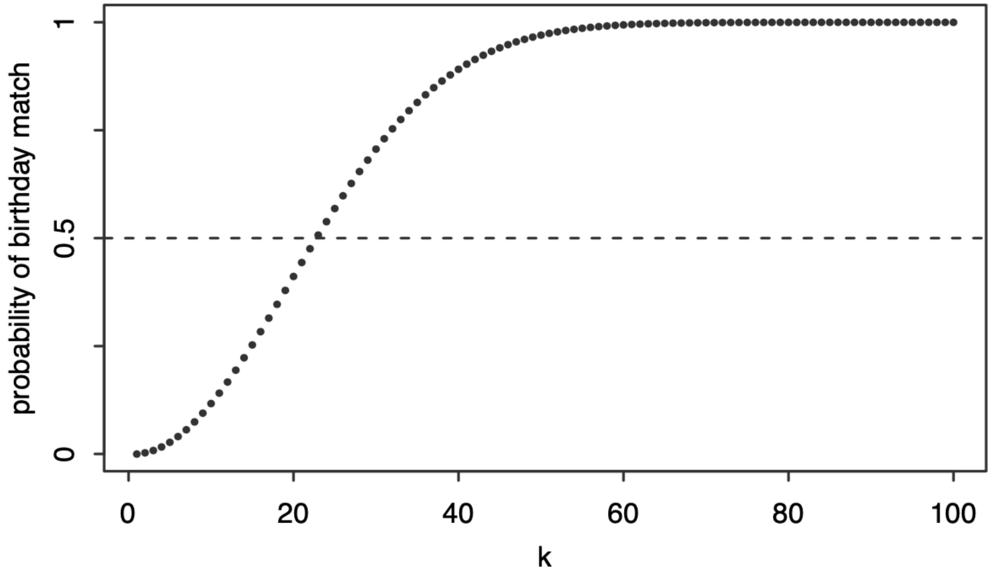


FIGURE 1.5
Probability that in a room of k people, at least two were born on the same day. This probability first exceeds 0.5 when $k = 23$.