Pigeonhole Principle

Consider a flock of pigeons nestled in a set of n pigeonholes. If there are n pigeons, then it is possible for all of the pigeons to rest happily in separate pigeonholes. However, if at least one more pigeon arrives, making a total of *more than n* pigeons, then at least one of the pigeonholes, inevitably, will end up with more than one pigeon.

The idea that having more pigeons than pigeonholes requires a pigeonhole with more than one pigeon is seemingly trivial, but it turns out to be important enough that it has a name:

THEOREM

Pigeonhole Principle (naive form):

If more than n objects are placed into n boxes, then at least one box must contain more than one object.

At first glance, the **pigeonhole principle** (also known as **Dirichlet's principle** in honor of the eponymous German mathematician) might appear to be too obvious to be useful; indeed, the power of the principle comes from cleverly choosing the "boxes" and "objects." Even though the principle itself is quite simple, it is not always clear if it is useful and, if so, how. For instance, consider the following example:

EXAMPLE

A box contains three pairs of socks colored red, blue, and white, respectively. Suppose I take out the socks without looking at them. How many socks must I take out to be sure that they will include a matching pair?

If I take only 2 or 3 socks, it is possible that they are all different. For example, they may be one red and one blue; or one red, one blue, and one white. But if I take out 4 socks, these must include a matching pair. Here the 4 chosen socks are the "objects" and the 3 colors are the "boxes"; by PP1, it follows that at least two of the four chosen must have the same color and hence must be a matching pair. Thus the minimum number of socks to be taken out is 4. □

How many students do you need in a school to guarantee that there are at least 2 students who have the same first two initials?

There are $26 \times 26 = 676$ different possible sets of two initials that can be obtained using the 26 letters A, B, C, ..., Z, so the number of students should be greater than 676. Thus, the minimum number of students is 677. $\square$

Contents

Naive Pigeonhole Principle

In many situations, the "naive" form of the pigeonhole principle can be applied directly. In most problems, the "objects" and "boxes" are fairly obvious.

A box contains three pairs of socks colored red, blue, and green, respectively. If the socks are chosen without looking, how many socks must be drawn to guarantee at least one matching pair?

It is possible for any selection of three or fewer socks to consist of only distinct socks. Given a selection of three socks, it is less likely though to have one red, one blue, and one green sock. However, if **four socks** are chosen, the pigeonhole principle ensures that two socks must be the same. Each of the sock colors is a "box," with each sock drawn being an "object" to be sorted into the box. With three boxes, more than three items must be drawn to have two matching items. □

Note that, strictly speaking, one must always verify if n (or fewer) objects do not fit in n boxes. The pigeonhole principle guarantees that there will be a collision if *more* than n objects are placed in n boxes; however, it does not specify if n or *fewer* objects can fit in n boxes, which may certainly not be true depending on the restrictions given in the problem (imagine that some pigeons do not like each other or, perhaps, that some of the pigeonholes contain blockages).

EXAMPLE

A school has n students. What is the smallest n such that at least two of the students have matching first and last initials?

There are $26 \cdot 26 = 676$ distinct combinations of first and last initials. With 676 students, it is certainly possible to assign each student a distinct set of first and last initials. However, by the pigeonhole principle, having $n = 677$ students guarantees that at least two students must have matching initials. □

TRY IT YOURSELF

Calvin has 13 pairs of shoes thrown around in his closet. The light has gone out, and he is unable to determine which shoe is which. What is the minimum number of shoes he must pick in order to guarantee that he has a matching pair among them?

Submit your answer

Details and Assumptions:

- Each shoe that he removes is either a left or right shoe of one of the 13 pairs.

EXAMPLE

Show that given a set of n positive integers, there exists a non-empty subset whose sum is divisible by n.

Let the n integers be denoted by $a_1, \ldots, a_n$. Form the n sums

$$S_1 = a_1$$
$$S_2 = a_1 + a_2$$
$$\vdots$$
$$S_n = a_1 + \cdots + a_n.$$

If one of these sums is divisible by n, then we are done. Otherwise, by the pigeonhole principle, at least two of the sums must have the same remainder when divided by n (since only $n - 1$ distinct remainders are possible). Pick two such sums S_i and S_j, with $j > i$. Then it follows that

$$S_j - S_i = a_{i+1} + \cdots + a_j$$

must be divisible by n. $\square$

Note: If a and b have the same remainder when divided by n, then $a - b$ is divisible by n.

EXAMPLE

A group of statistics professors hold a meet-and-greet over dinner. No new guests enter the dinner after it has begun, and no one leaves until the dinner has ended. Show that at least two professors shook the same number of distinct individuals' hands. (*Hint:* A professor cannot shake his or her own hand.)

If there are n professors, then each individual could have shaken hands with at most $n - 1$ other individuals. Furthermore, if there exists an individual who shook hands with $n - 1$ other individuals, then there does not exist an individual who shook hands for no other individuals and vice versa. In other words, each of the n professors can be thought of as an object to be sorted into one of $n - 1$ boxes by the number of distinct individuals' hands shaken. By the pigeonhole principle, there exists

at least one integer between 1 and $n - 1$ or 0 and $n - 2$ that equals the number of distinct individuals' hands shaken by at least two professors. ☐

Suppose you randomly select k of the first 2016 positive integers. What is the smallest k that guarantees that at least one pair of the selected integers will sum to 2017?

Submit your answer

Missy and Mussy are very messy sisters. Their dresser drawer consists of 43 white socks, 2 black socks, 23 blue socks and 8 red socks. What is the minimum number of socks they must remove from the drawer, in order to be certain that they have removed four socks of the same color?

Submit your answer

Pigeonhole Principle on Continuous Spaces

The pigeonhole principle can also be applied to problems that take place on a "continuous" space if a clever partition of the space is chosen.

Given a line segment of length L that contains $n + 1$ points, let d be the length of the shortest segment between consecutive points. What is the maximum value of d over all possible configurations of points?

First, consider a trivial configuration of the points. Let all points be evenly spaced with one point at each end of the segment. In this case, the points divide up the line into n segments, each of length $\frac{L}{n}$.

$$d$$
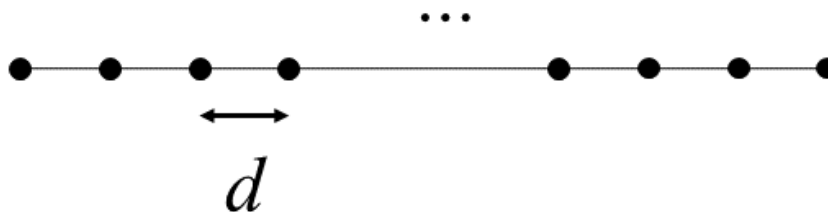
Clearly, we can do no worse than $\frac{L}{n}$, but is it possible to do better?

At first glance, it appears that the answer is no. Shifting any non-endpoint increases the length of one of the interior segments but always at the expense of shortening a neighboring segment.

Using the pigeonhole principle, we can approach the problem as follows: Consider each of the n evenly spaced segments as a "box" and each of the $n + 1$ points as an item to be placed into the boxes. The pigeonhole principle implies that at least one box (or segment) must have two items (or points), which guarantees that no two consecutive points can be farther apart than $\frac{L}{n}$. $\square$

EXAMPLE

Ten points are placed within a unit equilateral triangle. Show that there exists two points with distance at most $\frac{1}{3}$ apart.

Consider the partition of the triangle into nine congruent equilateral triangles as shown below:
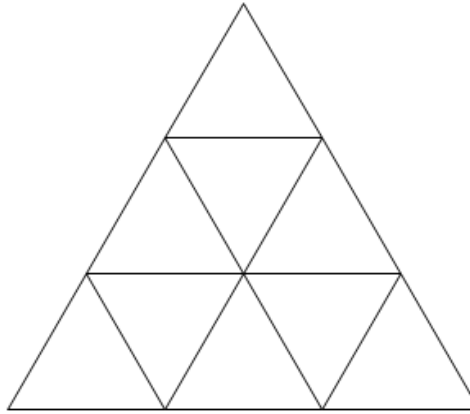
Each of the nine smaller triangles represents a box, with each of the ten points an item to be placed into the boxes. By the pigeonhole principle, at least one of the nine triangles must contain at least two points. Since the maximum distance between any two points in one of these triangles is $\frac{1}{3}$, no two such points can be separated by greater than distance $\frac{1}{3}$. □

EXAMPLE

Let the average of a set of positive numbers be μ. Use the pigeonhole principle to show that there exists at least one number less than or equal to μ.

Choosing a set of n numbers that average to μ is equivalent to placing $n + 1$ points on a line segment of length $n\mu$ with two points fixed as the endpoints. Each of the n segments formed between consecutive points represents one of the numbers in the set.

Furthermore, partition the entire segment into n sub-segments of length μ. The pigeonhole principle thus implies that at least one segment contains at least two points or, namely, that there must be at least one segment of length μ or less. It immediately follows that at least one number is less than or equal to μ. □

Generalized Pigeonhole Principle

A more general form of the pigeonhole principle is as follows:

> **THEOREM**
>
> **Pigeonhole Principle (general form):**
>
> If more than $k \cdot n$ objects are placed into n boxes then at least one box must contain more than k objects.

The case of $k = 1$ corresponds to the naive pigeonhole principle stated earlier.

Finally, let us prove the (generalized) pigeonhole principle. The argument is fairly straightforward.

> **PROOF**
>
> For the sake of contradiction, suppose there are no boxes that contain more than k pigeons. In that case, each pigeonhole may contain at most k pigeons, and all n pigeonholes may contain a total of at most $k \cdot n$ pigeons. But this contradicts the assumption that there were *more* than $k \cdot n$ objects. It follows that at least one box must contain more than k pigeons. $\square$

It should be noted that contradiction permits only the *weakest* claim possible regarding the contents of the boxes. If it is not the case that every box has at most k objects, all one can be certain of is that *at least one* box has *at least one more* than k objects.

> **EXAMPLE**
>
> What is the minimum number of cards that must be drawn from a standard deck to guarantee at least three cards all of the same suit?
>
> ---
>
> Let the four suits correspond to our "boxes." The pigeonhole principle implies that if we draw more than $2 \cdot 4$ cards from the 4 suits, then at least one suit must have more than 2 drawn cards.
>
> Lastly, we should note that, with eight cards drawn, it is possible to have exactly two cards of each suit, so the minimum number is indeed 9. $\square$

Pigeonhole Principle - Problem Solving

I have 7 pairs of red socks, 6 pairs of blue socks, and 5 pairs of green socks in my drawer. If I randomly pick out 4 socks, what is the probability that I will get (at least) a matching pair?

Details and Assumptions:

- Besides the color, there is no difference between a left sock and a right sock.

- ○ 0
- ○ $\frac{1}{4}$
- ○ $\frac{1}{3}$
- ○ $\frac{1}{2}$
- ○ $\frac{2}{3}$
- ○ $\frac{3}{4}$
- ○ 1

If we pick 77 numbers randomly from the set $\{1, 2, 3, 4, ..., 150\}$, we are guaranteed to have at least k pairs of numbers where the difference between each pair is 19. What is the maximum possible value of k?

Please treat a pair as a combination, not a permutation.

- ○ 0
- ○ 1
- ○ 2
- ○ Cannot be determined

Given any 10-element subset M of $\{1, 2, 3, \ldots, 117\}$, does there exists 2 non-empty disjoint subsets of M that have the same sum?

○ Yes, for all M

○ No, only for some M but not all

○ No, for no M

A lattice point is defined as a point in the two-dimensional plane with integral coordinates. We define the centroid of four points $(x_i, y_i), i = 1, 2, 3, 4$ as point $\left(\frac{x_1 + x_2 + x_3 + x_4}{4}, \frac{y_1 + y_2 + y_3 + y_4}{4} \right)$.

What is the largest number of distinct lattice points in the plane such that the centroid of any four of them is not a lattice point?

Submit your answer

Pigeonhole Principle (PP1)

If more than n objects are distributed into n boxes, then at least one box must have at least two objects.

This follows by noting that, if each of the n boxes contained at most 1 object, only then the n boxes will receive at most n objects contrary to the assumption that more than n objects were put into the boxes.

This principle was first stated formally by **Dirichlet** and is therefore also called **Dirichlet's box principle** or **Dirichlet's drawer principle**. Let us apply it to some worked-out examples:

A box contains three pairs of socks colored red, blue, and white, respectively. Suppose I take out the socks without looking at them. How many socks must I take out to be sure that they will include a matching pair?

If I take out only 2 or 3 socks, it is possible that they are all different. For example, they may be one red and one blue; or one red, one blue, and one white. But if I take out 4 socks, these must include a matching pair. Here the 4 chosen socks are the "objects" and the 3 colors are the "boxes"; by PP1, it follows that at least two of the four chosen must have the same color and hence must be a matching pair. Thus the minimum number of socks to be taken out is 4. □

EXAMPLE

How many students do you need in a school to guarantee that there are at least 2 students who have the same first two initials?

There are $26 \times 26 = 676$ different possible sets of two initials that can be obtained using the 26 letters A, B, C, ..., Z. So by PP1, the number of students should be grater than 676, and thus the minimum number of students is 677. □

Second Form of Pigeonhole Principle (PP2)

THEOREM

For any positive integers n and t, if $tn + 1$ or more objects are placed in n boxes, then at least one box will contain more than t objects.

PROOF

If each of the boxes contained at most t objects, then n boxes will contain at most tn objects. This contradicts the fact that $tn + 1$ or more objects are placed in the boxes. Hence at least one box must contain more than t objects. □

Third Form of Pigeonhole Principle (PP3)

THEOREM

If the average of n positive numbers is t, then at least one of the numbers is greater than or equal to t. Further, at least one of the numbers is less than or equal to t.

PROOF

Let $a_1, a_2, \ldots, a_n$ be the numbers. Then by data

$$\frac{\sum_{i=1}^{n} a_i}{n} = t \implies \sum_{i=1}^{n} a_i = nt. \qquad (1)$$

Hence, if each of the n numbers $a_1, a_2, \ldots, a_n$ is less than t, then the sum of these numbers will be less than nt, contradicting (1).

A similar argument shows that at least one of the numbers is less than or equal to t. $\square$

If the numbers $a_1, a_2, \ldots, a_n$ are integers, then PP3 says that at least one of them is $\geq \lceil t \rceil$, and at least one is $\leq \lfloor t \rfloor$, where $\lceil \; \rceil$ and $\lfloor \; \rfloor$ are the ceiling and floor functions, repectively.

Fourth Form of Pigeonhole Principle (PP4)

THEOREM

Let $q_1, q_2, \ldots, q_n$ be positive integers. If $\left(\sum_{i=1}^{n} q_i - n + 1 \right)$ objects are put into n boxes, either the first box contains at least q_1 objects, or the second box contains at least q_2 objects, ..., or the nth the box contains at least q_n objects.

Suppose we distribute $(q_1 + q_2 + \cdots + q_n - n + 1)$ objects in n boxes. If for each $i = 1, 2, 3, ..., n$ the ith box contains less than q_i objects, then the total number of objects in the n boxes is $\leq (q_1 - 1) + (q_2 - 1) + \cdots + (q_n - 1) = q_1 + q_2 + \cdots + q_n - n$.

But this number is less than the number of objects placed in the n boxes. For at least one i, the ith box must contain at least q_i objects. $\square$

Cite as: Pigeonhole Principle. *Brilliant.org*. Retrieved 16:40, June 19, 2021, from https://brilliant.org/wiki/pigeonhole-principle-definition/