

 **PIGEONHOLE PRINCIPLE** 


The Pigeonhole Principle	2
The Generalised Pigeonhole Principle	10
Exercises	12



THE PIGEONHOLE PRINCIPLE

In 1834, German mathematician Peter Gustav Lejeune Dirichlet (1805-1859) stated a simple - but extremely powerful - mathematical principle which he called the *Schubfachprinzip* (drawer principle). Today it is known either as the pigeonhole principle, as Dirichlet's principle, or as the cubby-hole principle.

The Pigeonhole Principle: If more than N objects are to go into N boxes, then at least one box must contain more than one object.

For example, if 13 pigeons are to lodge into 12 cubbies, then at least one cubby must contain two or more pigeons.

Note: One might be tempted to say in this example that exactly one cubby will contain exactly two pigeons, but this need not be true: we are examining *all* possible distributions of 13 pigeons into 12 cubbies (13 in one cubby, none in the rest OR 3 in one cubby, 5 in another, 5 in a third, none in the rest, and so on.) The pigeonhole principle states that for each and every possible distribution of pigeons, there is sure to be at least one cubby with two or more pigeons in it.

The pigeonhole principle can be phrased in terms of labels.

If more than N objects are to be assigned labels from a set of N labels, then there is sure to be two objects with the same label.

This simple principle allows us to make some mighty surprising conclusions about the world.

EXAMPLE: There exist at least two, non-bald, people in New York city with exactly the same number of hairs on their heads.

REASON: "Label" each person by the number of hairs on his or her head. The typical number of hairs on a human head is 150,000, and the count is certainly never more than a million. As there are well over a million (non-bald) people living in New York, at least two people must have the same number of hairs. □

A COMMENT ON "PROOF"

The previous example exhibits a common practice of mathematicians: to establish the existence of a certain phenomenon without presenting the slightest clue as to how to find a specific example of it! Many people find this somewhat disturbing and might even argue that the logic presented in this argument - although iron clad - does not represent "proof." Where are these two people possessing the same number of hairs?

This concern addresses the very issue of what constitutes a proof. High-school students, in particular, are accustomed to seeing only tangible results - to obtain a specific numerical answer, to find a specific function with certain properties, or to list all cases of scenario and examine them one by one. Such brute-force methods do not appeal to a broader scope of analysis, to a "meta" type of thinking of a problem at hand. (Also, it does not appeal to desire to seek beauty and elegance in mathematical reasoning.)

The New York example is curious. It has the feel of "real world" problem and therefore "should be analyzed in the high-school way of listing tangible options," some might say. Yet those that say this tend not to feel the same about the following problem:

TRUE or FALSE: If 27 people are in a room, then we can be sure that two people have first names that begin with the same letter.

Most say "true" and can readily articulate why *without* listing or even wanting to list all 26^{27} possibilities for 27 first-name letters! Why is this different? The logic behind the New York example is no different than the logic behind this scenario. Is it because these 27 people don't actually exist, but New Yorkers do?

We might feel dissatisfied with our reasoning for the New York example because it offers absolutely no hint of any kind as to how to find two people with the same number of hairs. New Yorkers exist, therefore our proof has failed? No! We have actually accomplished something practical: If you feel game to hunt for two such special people, then we have proved that your search won't be fruitless. Two people like this are "out there." And it is good to know this for sure before you start out.

It often occurs in mathematics that one can prove certain entity does exist (or doesn't exist), with nothing more to say about it. These results are not invaluable. They tell the physicist, engineer, or applied mathematician who might be interested in these quantities whether or not a search for them is worth the while.

EXAMPLE: Twenty people in a room take part in handshakes. Each person shakes hands at least once and no one shakes the same person's hand more than once. Prove that two people took part in the same number of handshakes.

ANSWER: Label each person by the number of handshakes she took part in. There are twenty people but only nineteen labels: 1, 2, 3, ..., 18, 19. By the pigeonhole principle, at least two people have the same label. \square

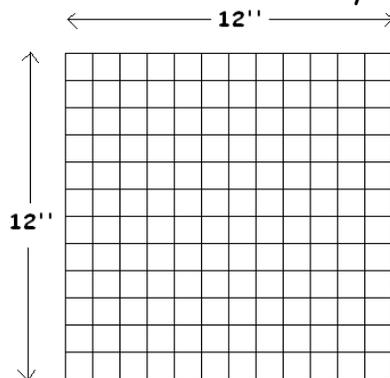
EXAMPLE: Eight positive numbers are chosen at random. Explain why two of them are sure to differ by a multiple of seven.

ANSWER: Label each number by the remainder it leaves when divided by seven. There are eight numbers and only seven labels: 0, 1, 2, 3, 4, 5, or 6. At least two numbers have the same label. The difference of these numbers is a multiple of seven. \square

COMMENT: Be sure to understand this idea: *If a and b leave the same remainder when divided by N , then $a - b$ is a multiple of N .* This is a standard trick used throughout this chapter.

EXAMPLE: 145 points are chosen at random in a one-foot by one-foot square. Prove that there exist two points no more than $\sqrt{2}$ inches apart.

ANSWER: Divide the square into 144 one-inch by one-inch squares.



At least one square contains two points. These two points are no more than $\sqrt{2}$ inches (the length of the diagonal of the square) apart. \square

EXERCISE:

- There are 27 students in a room. Explain why there are at least two students with first name starting with the same letter.
- My name is "James Tanton" and so my initials are JT. How many people need to be in a room to ensure that at least two people in that room have the same pair of initials?

EXAMPLE: Let $x_1, x_2, x_3, \dots, x_{20}$ be 20 consecutive integers. Choose any 11 of them at random. Then at least two chosen integers differ by 10.

Answer: Label each of the chosen numbers by its remainder upon division by 10. As there are 11 numbers and 10 labels, two must have the same label and hence differ by a multiple of 10.

Since the numbers $x_1, x_2, x_3, \dots, x_{20}$ are consecutive, two cannot differ by 20 or more. Thus the two numbers that differ by a multiple of 10 can only differ by exactly 10. □

COMMENT: This example generalizes

Choose $N + 1$ integers at random from a set of $2N$ consecutive integers. Two shall differ by N .

EXERCISE: $N + 1$ integers are chosen at random from a set of $2N$ consecutive integers. Prove that at least two chosen integers are consecutive.

HINT: One approach is to label each of the chosen integers by its result upon division by 2, ignoring remainders.

The following problem is surprisingly hard. It was a favourite of Hungarian mathematician Paul Erdős (1913-1996).

EXERCISE: If 26 integers are chosen at random from the set of consecutive integers 1, 2, 3, ..., 50, prove that there are sure to be two numbers so that one is a multiple of the other.

Answer: Each chosen integer can be written in the form $2^k m$ with m odd. (For example, $12 = 2^2 \cdot 3$, $40 = 2^3 \cdot 5$, $7 = 2^0 \cdot 7$ and $8 = 2^3 \cdot 1$.) As each chosen integer is ≤ 50 , the odd part m of each chosen number is no more than 49. That is, m can only be one of the following 25 odd values: 1, 3, 5, ..., 49.

But we have chosen 26 integers, and so two of them must have the same odd part. Call these integers a and b :

$$a = 2^k \cdot m$$

$$b = 2^r \cdot m$$

If $k < r$ then we have $a | b$. If $r < k$ then we have $b | a$. Either way, we have found a pair of integers with one a multiple of the other. \square

COMMENT: This problem generalizes: *If $N+1$ integers are chosen at random from a set of $2N$ consecutive integers, then one is sure to be a multiple of the other.*

This next problem has a hidden subtlety:

EXERCISE: In any group of six people, there are two who have an identical number of friends in the group.

HINT: Assign to each person the number of friends she has in the group. There are six possible labels - 0, 1, 2, 3, 4, and 5 - and so it seems that the pigeonhole principle does not apply. However ...

Let's ease down a bit on the difficulty level.

EXAMPLE: For any infinite sequence of integers, prove that there are sure to be two numbers in that sequence that differ by a multiple of 5243.

ANSWER: Label each number in the sequence by the remainder it leaves upon division by 5243. There are 5243 possible labels: 0, 1, 2, ..., 5242. As there are an infinite number of entries in the sequence, two entries must have the same label, that is, the same remainder upon division by 5243. This means that their difference is a multiple of 5243. \square

This result is interesting. It says that for each of the following sequences there are two numbers that differ by a multiple of 5243. Do you really believe this?

- 1, 10, 100, 1000, 10000, ...
- 1, 2, 3, 4, 5, ...
- 1, -2, 3, -4, 5, -6, 7, ...
- 1, 1, 1, 1, 1, 1, ...
- 2, 4, 8, 16, 32, 64, 128, ...
- 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, ...

EXERCISE:

- a) Five points with integer coordinates are chosen at random. Prove that the midpoint of the line segment connecting some two of those points also has integer coordinates.
- b) Show that this result need not be true if only four points with integer coordinates are selected.

HINT: For any point (a, b) with integer coordinates a is either odd or even, as is b . There are thus four possible states for the parity of the pair (a, b) .

EXAMPLE: Write down any 10 integers in a list. Then some sequence of consecutive terms in your list is sure to add to a multiple of 10.

e.g.

3 5 11 86 -4 7 24 3 -6 201

↖
add to 110

Try it!

REASON: Suppose the numbers are a_1, a_2, \dots, a_{10} . Consider the sums:

$$s_1 = a_1$$

$$s_2 = a_1 + a_2$$

$$s_3 = a_1 + a_2 + a_3$$

$$\vdots$$

$$s_{10} = a_1 + a_2 + a_3 + \dots + a_{10}$$

If any of these are a multiple of 10, then we're done.

Suppose, instead, we are not lucky and none of these sums leaves a remainder of zero when divided by 10.

In this case, label each sum by the remainder it leaves upon division by 10. There are 10 sums and nine labels: 1, 2, 3, 4, 5, 6, 7, 8, 9. Thus two of these sums have the same label: s_r and s_k , say, with $k > r$. This means that $s_k - s_r$ is a multiple of 10. That is, $a_{r+1} + a_{r+2} + \cdots + a_k$ is a multiple of 10. We're done! □

The following example is sneaky.

EXAMPLE: Show that there is a power of three that ends in 001.

ANSWER: Consider the powers of three 1, 3, 9, 27, ... and label them by the remainder they leave upon division by 1000. There are 1000 different labels. As there are an infinite number of powers of three, two must have the same label. That is, there are two powers of three, 3^m and 3^n , say, with $m > n$, that leave the same remainder upon division by 1000. This means $3^m - 3^n = 3^n(3^{m-n} - 1)$ is a multiple of 1000.

As 1000 has no primes in common with 3^n , we have that $3^{m-n} - 1$ is a multiple of 1000: $3^{m-n} - 1 = 1000k$. Thus $3^{m-n} = 1000k + 1$ ends 001. □

EXAMPLE: A circle has circumference 100 cm. Along this circumference, 103 points are chosen at random. Prove that two of those points are sure to be less than 1 cm apart.

ANSWER: Divide the circumference of the circle into 102 segments of equal length. Each of these segments is less than 1 cm long.

For any distribution of 103 points among these 102 segments, two are sure to land in the same segment. □



THE GENERALISED PIGEONHOLE PRINCIPLE

The pigeonhole principle can be generalized as follows:

Generalized Pigeonhole Principle

If each of N objects is to be assigned one of k labels, then there are sure to be at least $\frac{N}{k}$ objects with the same label.

Proof: Suppose for each of the k labels we have less than $\frac{N}{k}$ objects of that label. Then the number of objects is less than $k \cdot \frac{N}{k} = N$, which is not the case. Thus at least one label is used $\frac{N}{k}$ or more times. \square

[For example, if 55 objects are to be put in 6 boxes, at least one box will possess at least 10 objects. (If not, six boxes with 9 or less objects will constitute only a total of 54 objects). Notice that $\frac{55}{6} \approx 9.17$]

EXAMPLE: Among 2000 people, at least six were born on the same day of the year.

REASON: Label each person with the day on which she was born. There are 366 possible labels and 2000 people assigned those labels. At least $\frac{2000}{366} \approx 5.46$ people, that is, six people have the same label. \square

EXAMPLE: If 865 points are scattered in a one-foot by one-foot square, then at least seven points are clustered sufficiently close together as to lie in a one-inch by one-inch square.

REASON: Divide the square into 144 one-inch by one-inch squares. Notice that $\frac{865}{144} \approx 6.007$. \square

EXAMPLE: Fifty-one integers are chosen at random from the numbers 1 through 100. Prove that at least two of the chosen integers will differ by 10.

ANSWER: Label each of the chosen numbers by its remainder upon division by 10. As there are 51 numbers, some $\frac{51}{10} = 5.1$ of them, that is, at least 6 of them, must have the same label. Thus six numbers of the chosen numbers are spaced from each other by multiples of 10.

Since the six numbers are chosen from the range 1 to 100, it is not possible for all six to be 20 or more counts from each other. (Think about this.) Thus at least two of the six numbers differ by exactly 10. \square

Another variant of the pigeon-hole principle is used in this next example.

EXAMPLE: Six integers are chosen at random from the set 1, 2, 3, ..., 9, 10. Prove that two of the chosen integers are sure to have an odd sum.

ANSWER: Label each of the integers as either "even" or "odd." As there are at most five of each label we must have two integers of opposite label. These sum to an odd value. \square

We have: If there is an upper bound on the number of objects that may possess a certain label and one has more objects than that bound, then at least two objects will have different labels.

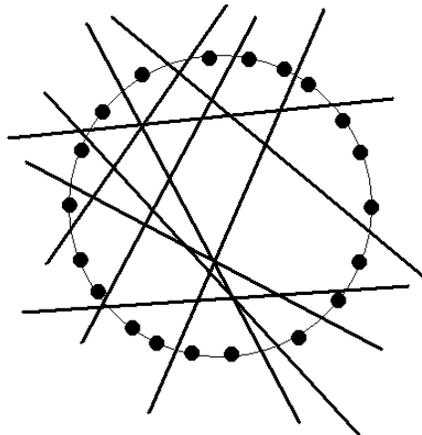


EXERCISES

Question 1: A drawer contains 14 red socks and 10 blue socks. It is dark. How many socks must you take out of the drawer to be sure of having a pair the same colour?

Question 2: There are more books in a library than there are pages in any one book the library possesses. Must there be two books in the library with the same number of pages?

Question 3: Twenty dots are drawn on the rim of a circle. Nine straight lines are drawn across the circle being sure to cross between dots. Prove that there must remain two neighbouring dots not separated by a line.



Question 4: A bag contains several red balls, several blue balls, and several yellow balls. Each day Jenny pulls out three balls one at a time, noting their colours in turn. She then returns the three balls to the bag. She does this each day for a month.

Prove that, in the month, there were at least two days with exactly the same outcome.

Question 5:

a) Five points are chosen at random inside an equilateral triangle of side-length 2 inches. Prove that there are sure to be two points one inch or less apart.

b) Seventeen points in the triangle are chosen at random. Prove at least two are no more than 0.5 inches apart.

Question 6:

a) Show that there is a power of 17 that ends 0001.

b) Explain why there is no power of 5 that ends 0001.

Question 7: Eleven numbers are chosen at random from the integers 1 through 20. Prove that two of the chosen integers have gcd equal to 1.

Question 8: We proved: *If 26 integers are chosen at random from the set of consecutive integers 1, 2, 3, ..., 50, there is sure to be two numbers so that one is a multiple of the other.*

Show that the result need not be true if only 25 integers are chosen from the set $\{1, 2, 3, \dots, 50\}$.

Question 9: Nine points in three-dimensional space with integer coordinates are chosen at random. Prove that the midpoint of the line segment connecting some two of them also has integer coordinates.

Question 10: Our goal is to answer the following classic problem:

During the month of April, Joey ate a total of 42 cupcakes. He ate at least one cake each day of the month. Prove that there was a string of consecutive days over which he ate a total of 17 cupcakes.

Let a_1, a_2, \dots, a_{30} be the number of cupcakes he ate each day of the month. Each value a_i is at least one.

Let $s_1 = a_1, s_2 = a_1 + a_2, s_3 = a_1 + a_2 + a_3, \dots, s_{30} = a_1 + a_2 + \dots + a_{30} = 42$.

- a) Explain why the following 60 numbers all lie between the values 1 and 59:

$$s_1, s_2, s_3, \dots, s_{30}, s_1 + 17, s_2 + 17, \dots, s_{30} + 17$$

- b) Explain why we must have $s_r = s_p + 17$ for some values r and p .
c) Solve the problem.

Question 11: Each day of the year Lashana takes at least one aspirin. In the year 2001 she took a total of 600 aspirins (one bottle). Prove, over a string of consecutive days of that year, she consumed exactly 129 aspirins.

Question 12:

a) Suppose we are given N integers a_1, a_2, \dots, a_N in a specific order, repetitions allowed. Prove that there is sure to be a set of consecutive numbers whose sum, $a_{r+1} + a_{r+2} + \dots + a_p$, is a multiple of N .

b) Write down six random integers. Find, in your example, a sum of consecutive terms divisible by 6. Is there also a sum of consecutive terms divisible by 5? By 4? By 3? By 2?

c) Suppose we are given N integers a_1, a_2, \dots, a_n in a specific order, repetitions allowed. Let k be a positive integer smaller than N . Prove that there is sure to be a set of consecutive numbers whose sum, $a_{r+1} + a_{r+2} + \dots + a_p$, is a multiple of k .

Question 13:

- a) Suppose $p(x)$ is a polynomial with integer coefficients. If $p(a) = 4$, $p(b) = 5$, and $p(c) = 4$ for some integers a , b and c , prove that a , b and c must be consecutive!

HINT: First establish the following algebraic identity:

$$x^n - y^n = (x - y)(x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \cdots + xy^{n-2} + y^{n-1})$$

It shows that for any n we have $(x - y) \mid (x^n - y^n)$. (We also proved this in chapter 13). Write the polynomial $p(x)$ as:

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0$$

and explain why $(x - y) \mid (p(x) - p(y))$

Now ... to return to the question ... explain why $(a - b) \mid 1$ and $(b - c) \mid 1$.

Deduce that $a = b \pm 1$ and $c = b \pm 1$ and explain the result.

- b) Suppose, instead, $p(x)$ takes the value 17 at three different integer inputs. Prove that it will never take the value 18 at a fourth integer input.
- c) Suppose $p(x)$, a polynomial with integer coefficients, takes the value 23 at four different integer inputs and the value 25 just once for another integer input. Prove that those five integers must be consecutive.

Question 14: Let α be an irrational number.

- Prove that there is no integer n such that $n\alpha$ is an integer.
- Use the pigeonhole principle to prove that one can find an integer n such that $n\alpha$ is within a distance 0.001 from being an integer.

HINT: Divide the unit interval $[0, 1]$ into 1000 subintervals. Consider the extent to which each of the numbers $\alpha, 2\alpha, 3\alpha, 4\alpha, \dots$ is larger than an integer.

Question 15: Some regard the following observation a variation of the pigeonhole principle:

If n numbers sum to S , then not all the numbers are larger than $\frac{S}{n}$. Also, not all the numbers are smaller than $\frac{S}{n}$.

- Explain why this principle is true.
- The salaries of five workers summed to \$100,000. Prove that at least one worker earned no more than \$20,000.

Question 16: For 105 people in a room how many people can we be sure have first name beginning with the same letter of the alphabet?

Question 17: In a state lottery one is required to submit a four-digit number composed of non-zero digits. (For example, 7823 and 8828 are valid entries, but 8906 is not.) If 10,000 people enter the lottery, how many people, for certain, submitted the same number?

Question 18: A university has seven curriculum committees. Albert, Bilbert, Cuthbert, Dilbert, Egbert, Filbert, and Gilbert each serve on three of those committees. Prove that one of the committees has at least three of these people on it.

Question 19: Twenty people are to sit in a row of 25 chairs. Prove that four consecutive chairs are sure to be occupied.

Question 20:

Fifty-five integers are chosen at random from the numbers 1 through 100.

- a) Prove that at least two of them differ by 9.
- b) Prove that at least two of them will differ by 10.
- c) (**HARD**) Prove that at least two of them will differ by 12.
- d) Prove that at least two of them will differ by 13.

- e) Surprisingly, there need not be two of them that differ by 11. Show this by listing 55 numbers from 1 through 100 that fail to have a pair that differ by 11.

CURRICULUM CONNECTIONS

Pigeonhole principle; Patterns; Counting techniques