Now let us examine the binary expansions of the elements $y \in B$. $\forall y \in B$, $y = 1/2^{p1} + 1/2^{p2} + \ldots + 1/2^{pk}$ for $p_1, \ldots, p_k \in N^*$, $p_1 < p_2 < \ldots < p_k$. The two binary expansions corresponding to y, $b_{y,1}$ and $b_{y,2}$ are of the form $0.x_1x_2 \ldots x_{pk-1}x_{pk}x_{pk+1}$ where $x_1, \ldots, x_{pk-1}$ are common to $b_{y,1}$ and $b_{y,2}$, whereas $x_{pk}, x_{pk+1}, \ldots$ differ.
Now x_j = {1 if $j=p_1, p_2, \ldots, p_{k-1}$, 0 otherwise for $1 \le j \le p_k$ is the common part corresponding to $1/2^{p1} + 1/2^{p2} + \ldots + 1/2^{pk-1}$ whereas the difference comes from the two possible ways of representing the last term in the sum $1/2^{pk}$ namely 10000… or 011111…. Therefore, $b_{y,1}$ has $x_{pk} = 1$ and $x_j = 0$ $\forall j > p_k$ (corresponding to 1000....) whereas $b_{y,2}$ has $x_{pk} = 0$ and $x_j = 1$ $\forall j > p_k$ (corresponding to 01111…).
Let $s_{y,1} \in A$ be the sequence corresponding to $b_{y,1}$ in A, the set of all sequences of 0s and 1s, i.e. if $b_{y,1} = 0.x_1x_2x_3\ldots$ $s_{y,1} = \{x_1, x_2, x_3, \ldots\}$.
Let $s_{y,2} \in A$ be the sequence corresponding to $b_{y,2}$.
We now define $B_1 = \{b_{y,1} \mid y \in B\}$, $B_2 = \{b_{y,2} \mid y \in B\}$, $A_1 = \{s_{y,1} \mid y \in B\}$, $A_2 = \{s_{y,2} \mid y \in B\}$.
B is in one-to-one correspondence to B_1, B_2, A_1, A_2 by construction, so $B \sim B_1$, $B \sim B_2$, $B \sim A_1$, $B \sim A_2$, but B is countably infinite $\Rightarrow A_1, A_2, B_1, B_2$ are all countably infinite.

We have just one more observation to make regarding the correspondence between sequences of 0s and 1s in A and elements of (0, 1), namely that the zero sequence {0, 0, …} corresponds to the binary expansion $0.000\ldots = 0 \notin (0, 1)$ since $(0, 1) = \{x \in R \mid 0 < x < 1\}$ and the sequence {1, 1, 1, …} corresponds to the binary expansion $0.111\ldots = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots$

$$= \sum_{n=1}^{\infty} 1/2^n = 1 \notin (0, 1).$$

Now we can finally prove that (0, 1) is uncountably infinite.

Proposition: (0, 1) is uncountably infinite.

Proof: We define a map $f:(0, 1) \to \{0.x_1x_2x_3\ldots \mid x_j \in \{0, 1\}\ \forall j \geq 1\}$ as follows:
$f(y) = \{\ b_{y,1}$ if $y \in B$ (The first of the two possible binary expansions) or
$0.x_1x_2x_3\ldots$ if $y \notin B$ (The unique binary expansion)$\}$
By our previous discussion, f is a bijection as defined $\Rightarrow (0, 1) \sim \{0.x_1x_2x_3\ldots \mid x_j \in \{0, 1\}\ \forall j \geq 1\}$
Also by our previous discussion $\{0.x_1x_2x_3\ldots \mid x_j \in \{0, 1\}\ \forall j \geq 1\} \sim A\backslash(A_2 \cup \{0, 0, \ldots\} \cup \{1, 1, \ldots\})$ where:

- A = set of all sequences of 1s and 0s
- $\{0, 0, \ldots\}$ = sequence of all 0s
- $\{1, 1, \ldots\}$ = sequence of all 1s

Therefore $(0, 1) \sim A \setminus (A_2 \cup \{0, 0, \ldots\} \cup \{1, 1, \ldots\})$ since ~ is transitive (it is an equivalence relation).
A_2 is countably infinite, so $A_2 \cup \{0, 0, \ldots\} \cup \{1, 1, \ldots\}$ is countably infinite (we've added two elements to A_2, so it stays countably infinite).

In a previous theorem we proved A is uncountably infinite.
Thus $A \setminus (A_2 \cup \{0, 0, \ldots\} \cup \{1, 1, \ldots\})$ is of the form {uncountably infinite set} \ {countably infinite set}. I claim $A \setminus (A_2 \cup \{0, 0, \ldots\} \cup \{1, 1, \ldots\})$ is uncountably infinite.
Indeed, let $\tilde{A} = A \setminus (A_2 \cup \{0, 0, \ldots\} \cup \{1, 1, \ldots\})$. Assume $\tilde{A}$ is countable, then A is the union of a countable set with a countably infinite set, hence A is countable $\Rightarrow\Leftarrow$
Therefore, $\tilde{A} = A \setminus (A_2 \cup \{0, 0, \ldots\} \cup \{1, 1, \ldots\})$ is uncountably infinite, but $\tilde{A} \sim (0, 1)$ (~ is asymmetric) $\Rightarrow$ (0, 1) is uncountably infinite. (*q.e.d*)

Theorem: R is uncountably infinite.

Proof:
By the previous proposition, (0, 1) is uncountably infinite. By the proposition before this one, $(0, 1) \sim R \Rightarrow R$ is uncountably infinite. (*q.e.d*)

Under the equivalence relation ~ of bijective correspondence, we have shown that
$N, N^*, N^n\ \forall n \geq 1, Z^n\ \forall n \geq 1, Q^n\ \forall n \geq 1 \in [N]$ are all countably infinite and
A (all sequences of 0s and 1s), P(N) and [R] are uncountably infinite.

Question: Is there some intermediate equivalence class in size between [N] and [R]?

Answer: The continuum hypothesis (CH) gives a negative answer to this question.

The continuum hypothesis:
There is no set whose cardinality is strictly between the cardinality of the integers and the cardinality of real numbers. Cardinality means size or number of elements.

Georg Cantor stated CH in 1878, believed it was true, but could not prove it. It became one of the crucial open problems in mathematics. Hilbert stated in 1900 first among the 23 problems that were supposed to hold the key for the advancement of mathematics. Everybody expected CH to be either true or false.

The answer is that CH is independent from the standard axiomatic system used in mathematics called ZFC (Zermelo–Fraenkel with The Axiom of Choice). In other words, CH **cannot** be proven either true or false when working within the axiomatic framework of ZFC. In 1940 Kurt Gödel showed CH <u>cannot be proven false</u> within ZFC. In 1963 Paul Cohen showed CH <u>cannot be proven true</u> within ZFC and won the Fields Medal (like the Nobel Prize for mathematics) for his work.

Applications of Countability of Sets to Formal Languages:

Task:
Figure out the size of the set of all languages over a finite alphabet and the size of all regular languages over a finite alphabet. Let A be a finite alphabet, i.e. $A = \{a_1, \ldots, a_n\}$.

Recall that $A^* = \bigcup_{j=0}^{\infty} A^i$ is the set of all possible words in the alphabet A.

A^j is the set of all words of length j in the alphabet A.

Question: What is $\#(A^j)$, the size (cardinality) of A^j?

Answer: If j=0, $A^0 = \{\varepsilon\}$ where ε is the empty word, so $\#(A^0) = 1$. In general, we have n choices of letters in the first position, n choices of letters $(a_1, \ldots, a_n)$ in the second position and so on up to the j^{th} position. In total we have n x n x … x n (j times) = n^j possibilities. Therefore, $\#(A^j) = n^j$. Note that when j=0 $n^0=1 = \#(A^0) = \#(\{\varepsilon\})$.

Theorem:
If A is a finite alphabet then the set of all words over A $A^* = \bigcup_{j=0}^{\infty} A^j$ is countably infinite.

Proof:
We showed A^j is a finite set for each j. In fact, $\#(A^j) = n^j$. $\bigcup_{j=0}^{\infty} A^j$ is therefore a countably infinite union of disjoint infinite sets.
Note that $A^j \cap A^k = \emptyset$ if $j \neq k$ as no words of length j can be of length k if $j \neq k$.