We can view the Binomial Theorem as a statement about particular linear equations:
We can view the Binomial Theorem as a statement about particular linear equations: let $x_1, x_2, \ldots, x_n$ be variables that take on the values 0 and 1. Then the number of solutions to

$$x_1 + \cdots + x_n = k$$

is $C(n, k)$ for $0 \leq k \leq n$. 
To put this in the framework of the BT, we note that

$$(1 + x)^n = (1 + x) \cdot (1 + x) \cdot \cdots (1 + x),$$

and we associate $x_i$ with the exponent of $x$ in the $i$th term in the product.
To put this in the framework of the BT, we note that

\[(1 + x)^n = (1 + x) \cdot (1 + x) \cdots (1 + x),\]

and we associate \(x_i\) with the exponent of \(x\) in the \(i\)th term in the product.

Since \(1 + x = x^0 + x^1\), we see that \(x_i\) can be 0 or 1.
To put this in the framework of the BT, we note that

\[(1 + x)^n = (1 + x) \cdot (1 + x) \cdot \ldots \cdot (1 + x),\]

and we associate \(x_i\) with the exponent of \(x\) in the \(i\)th term in the product.

Since \(1 + x = x^0 + x^1\), we see that \(x_i\) can be 0 or 1. Furthermore, the coefficient of \(x^k\) is precisely the number of ways we can choose \(k\) terms provide an \(x^1\) and \(n - k\) to provide an \(x^0\).
To put this in the framework of the BT, we note that

$$(1 + x)^n = (1 + x) \cdot (1 + x) \cdots (1 + x),$$

and we associate $x_i$ with the exponent of $x$ in the $i$th term in the product.

Since $1 + x = x^0 + x^1$, we see that $x_i$ can be 0 or 1. Furthermore, the coefficient of $x^k$ is precisely the number of ways we can choose $k$ terms provide an $x^1$ and $n - k$ to provide an $x^0$. In other words, we're choosing which $k$ values of $i$ have $x_i = 1$, and which $n - k$ have $x_i = 0$. 
But why should we restrict ourselves to just solutions with variables 0 or 1? What if we want to solve \( x + y + z = 10 \), with \( x, y, z \) all nonnegative integers?
But why should we restrict ourselves to just solutions with variables 0 or 1? What if we want to solve $x + y + z = 10$, with $x, y, z$ all nonnegative integers?

The idea is the same: we associate each variable with a polynomial whose exponents correspond to the domain of the variable, and so the coefficient of $x^{10}$ in the product of them is the number we want.
But if we want $x, y, z$ to be nonnegative integers, don’t we need an infinite polynomial (i.e. power series)?

Through the power of a CAS, we can actually see that the coefficient of $x^{10}$ is 66, and therefore there are 66 solutions.
But if we want $x, y, z$ to be nonnegative integers, don’t we need an infinite polynomial (i.e. power series)?

No, if we cleverly realize that $x, y, z$ are all nonnegative and sum to 10, so $x, y, z \leq 10$.
But if we want $x, y, z$ to be nonnegative integers, don’t we need an infinite polynomial (i.e. power series)?

No, if we cleverly realize that $x, y, z$ are all nonnegative and sum to 10, so $x, y, z \leq 10$. So each can be associated with $1 + x + x^2 + \cdots + x^{10}$, and therefore we want the coefficient of $x^{10}$ in

$$(1 + x + \cdots + x^{10})^3.$$
But if we want $x, y, z$ to be nonnegative integers, don’t we need an infinite polynomial (i.e. power series)?

No, if we cleverly realize that $x, y, z$ are all nonnegative and sum to 10, so $x, y, z \leq 10$. So each can be associated with $1 + x + x^2 + \cdots + x^{10}$, and therefore we want the coefficient of $x^{10}$ in

$$(1 + x + \cdots + x^{10})^3.$$

Through the power of a CAS, we can actually see that the coefficient of $x^{10}$ is 66, and therefore there are 66 solutions.
This method clearly has some advantages: it lets us reduce many number-theoretic questions about solutions to integer-only linear equations to a silly problem of computing polynomial products.
Are You Serious?!?!

This method clearly has some advantages: it lets us reduce many number-theoretic questions about solutions to integer-only linear equations to a silly problem of computing polynomial products.

But what if we have more variables and a larger RHS? Do we really want to use this method all the time?
This method clearly has some advantages: it lets us reduce many number-theoretic questions about solutions to integer-only linear equations to a silly problem of computing polynomial products.

But what if we have more variables and a larger RHS? Do we really want to use this method all the time?

No!
For a much more entertaining way of solving this problem, consider the solution \((x, y, z) = (3, 3, 4)\).
For a much more entertaining way of solving this problem, consider the solution \((x, y, z) = (3, 3, 4)\). But, instead of writing numbers using our awesome notational system (thanks Fibonacci), we make the Romans look advanced by writing \(1 = 1\), \(2 = 11\), \(3 = 111\), etc., and a blank space for 0, but only on the LHS.
For a much more entertaining way of solving this problem, consider the solution \((x, y, z) = (3, 3, 4)\). But, instead of writing numbers using our awesome notational system (thanks Fibonacci), we make the Romans look advanced by writing \(1 = 1\), \(2 = 11\), \(3 = 111\), etc., and a blank space for 0, but only on the LHS. Then our equation \(3 + 3 + 4 = 10\) becomes

\[
111 + 111 + 1111 = 10.
\]
For a much more entertaining way of solving this problem, consider the solution \((x, y, z) = (3, 3, 4)\). But, instead of writing numbers using our awesome notational system (thanks Fibonacci), we make the Romans look advanced by writing \(1 = 1, 2 = 11, 3 = 111\), etc., and a blank space for 0, but only on the LHS. Then our equation \(3 + 3 + 4 = 10\) becomes

\[
111 + 111 + 1111 = 10.
\]

Another solution, \((2, 0, 8)\), becomes

\[
11 + +11111111 = 10.
\]
Now, look at the LHS in our two solutions:

\[ 111 + 111 + 1111 = 10, \]
\[ 11 + +11111111 = 10. \]
Now, look at the LHS in our two solutions:

\[ 111 + 111 + 1111 = 10, \]
\[ 11 + +11111111 = 10. \]

The LHS is just a sequence of ten 1’s and two +’s: \( x \) is just the number of 1s before the first +, \( y \) is the number between the first and second +’s, and \( z \) is the number after the second +.
Now, look at the LHS in our two solutions:

\[111 + 111 + 1111 = 10,\]
\[11 + +11111111 = 10.\]

The LHS is just a sequence of ten 1’s and two +’s: \(x\) is just the number of 1s before the first +, \(y\) is the number between the first and second +’s, and \(z\) is the number after the second +. Clearly for every solution of our problem, we can make this type of sequence, and for every sequence of this type, we get a different solution to our problem.
Therefore, we can find a combinatorial solution to our problem if we can only count the number of sequences of length 12 that contain 10 1’s and 2 +’s.
Therefore, we can find a combinatorial solution to our problem if we can only count the number of sequences of length 12 that contain 10 1’s and 2 +’s.

But of course we can count that: we either pick the 2 positions for the +’s or the 10 positions for the 1’s: in either case, we get

\[ C(12, 2) = C(12, 10) = \frac{12(11)}{2} = 66 \]

sequences, and therefore 66 solutions (just as we’d already found!).
What about the number of solutions to

$$x_1 + x_2 + \cdots + x_n = k,$$

where $k$ is a fixed nonnegative integer, and the $x_i$ are nonnegative integers?
Generalizations Are Never True.

What about the number of solutions to

$$x_1 + x_2 + \cdots + x_n = k,$$

where $k$ is a fixed nonnegative integer, and the $x_i$ are nonnegative integers?

By the same reasoning as before, we get that the number is

$$C(k + n - 1, k) = C(k + n - 1, n - 1).$$
As a reminder, Fibonacci numbers are numbers defined as follows: \( F_1 = F_2 = 1 \), and

\[
F_n = F_{n-1} + F_{n-2} \quad \text{for} \quad n \geq 3.
\]

(Additionally, we let \( F_0 = 0 \).)
Let’s look at the first couple numbers in the sequence:

0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, . . .
Let’s look at the first couple numbers in the sequence:

\[0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, \ldots\]

They seem to grow pretty quickly. In fact, we could (inductively) show that, for \(n \geq 3\),

\[
\left(\frac{5}{4}\right)^n \leq F_n \leq 2^n.
\]
Suppose $F_n = t^n$ for some $t$. Then, for $n \geq 3$, $t^n - t^{n-1} - t^{n-2} = 0$. Since obviously $t \neq 0$, we can divide through by $t^{n-2}$, giving us $t^2 - t - 1 = 0$. By the quadratic formula, we therefore have $t = t + \frac{\sqrt{5}}{2}$ or $t = t - \frac{\sqrt{5}}{2}$. Clearly, neither of them is right. (Though $t$ is the renowned golden ratio, of which much is written, little is true.)
Suppose $F_n = t^n$ for some $t$. Then, for $n \geq 3$, $t^n - t^{n-1} - t^{n-2} = 0$. Since obviously $t \neq 0$, we can divide through by $t^{n-2}$, giving us

$$t^2 - t - 1 = 0.$$
Suppose $F_n = t^n$ for some $t$. Then, for $n \geq 3$, $t^n - t^{n-1} - t^{n-2} = 0$. Since obviously $t \neq 0$, we can divide through by $t^{n-2}$, giving us

$$t^2 - t - 1 = 0.$$  

By the quadratic formula, we therefore have

$$t = t_+ = \frac{1 + \sqrt{5}}{2} \text{ or } t = t_- = \frac{1 - \sqrt{5}}{2}.$$
Suppose $F_n = t^n$ for some $t$. Then, for $n \geq 3$, 
$t^n - t^{n-1} - t^{n-2} = 0$. Since obviously $t \neq 0$, we can divide through by $t^{n-2}$, giving us

$$t^2 - t - 1 = 0.$$ 

By the quadratic formula, we therefore have

$$t = t_+ = \frac{1 + \sqrt{5}}{2} \text{ or } t = t_- = \frac{1 - \sqrt{5}}{2}.$$ 

Clearly, neither of them is right. (Though $t_+$ is the renowned golden ratio, of which much is written, little is true.)