Encouraging Undergraduate Students to Explore Multiple Proofs of the Multinomial Theorem

Christian Farkash, Michael Storm, Thomas Palmeri, and Chunhui Yu

Mathematics Department, Farmingdale State College, NY USA

chunhui.yu@farmingdale.edu

Abstract: *Several studies indicate that exploring mathematical ideas by using more than one approach to prove the same statement is an important matter in mathematics education. In this work, we have collected a few different methods of proving the multinomial theorem. The goal is to help further the understanding of this theorem for those who may not be familiar with it. These proofs can also be used by undergraduate college instructors in a calculus, a discrete mathematics or a probability course.*

Keywords: Multinomial distribution; Differential calculus; Probability; Combinatorics

INTRODUCTION

The multinomial theorem is used to expand any sum to an integer power and is an extension of the binomial theorem. The binomial theorem only deals with the addition of two variables to an integer power, whereas the multinomial theorem deals with more than two variables. The binomial and multinomial theorems are important results in elementary mathematics, and aside from the straightforward application of expanding polynomials of high degree, they also have applications in probability, combinatorics, number theory, and several other fields of mathematics. The multinomial theorem is written as follows:

$$(x_1 + x_2 + \cdots + x_m)^n = \sum_{\sum_{i=1}^{m} k_i = n} \frac{n!}{k_1!\, k_2! \cdots k_m!} {x_1}^{k_1} {x_2}^{k_2} \cdots {x_m}^{k_m}$$

Here, $k_1, k_2, \cdots, k_m \geq 0$ and the multinomial coefficient $\frac{n!}{k_1! k_2! \cdots k_m!} = \binom{n}{k_1, k_2, \cdots, k_m}$ is the number of possible ways to put n balls into m boxes.

When introducing the binomial theorem, most instructors often employ various methods to engage students and deepen their understanding (Flusser & Francia, 2000). Two typical approaches are:

1. Algebraic derivation: The instructor may start by introducing the binomial theorem and its formula, then proceed to prove it algebraically using mathematical induction or combinatorial arguments.
2. Pascal's triangle: The instructor may start by introducing Pascal's triangle and its connection to binomial coefficients, then show how each row corresponds to the coefficients in the binomial expansion.

The Multinomial theorem serves as a generalization of the binomial theorem, extending its principles from binomials to multinomials. We will present several approaches to prove the multinomial theorem in the following section.

Mathematics educators agree that exploring mathematical ideas by using more than one approach to solving the same problem (e.g., proving the same statement) is an essential element in the development of mathematical reasoning (NCTM, 2000; Polya, 2004; Schoenfeld, 2014; Dreyfus, Nardi & Leikin, 2012; Stupel & Ben-Chaim, 2013; Stupel & Ben-Chaim, 2017). Dreyfus, Nardi & Leikin (2012) discusses the pedagogical importance of multiple proof tasks and of taking into account the mathematical, pedagogical, and cognitive structures related to the effective teaching of proof and proving. Leikin (2009) indicates that the differences between the proofs are based on using: (1) different representations of a mathematical concept; (2) different properties (definitions or theorems) of mathematical concepts from a particular mathematical topic; (3) different mathematics tools and theorems from different branches of mathematics; or (4) different tools and theorems from different subjects (not necessarily mathematics). In our case, we apply the third type of differences between the proofs; we shall present various proofs using the tools and theorems of combinatorics, induction, probability, and differential calculus.

Proofs of the Multinomial Theorem

Combinatorial proof and induction proof are two classical methods which can be easily found in a standard textbook or with an online search. For readers' convenience, we state them here first.

Combinatorial Proof

Given variables $x_1, x_2, \cdots x_m$, we look to expand

$$(x_1 + x_2 + \cdots + x_m)^n \tag{1.1}$$

By definition, we know that this can be expressed as

$$\underbrace{(x_1 + x_2 + \cdots + x_m)(x_1 + x_2 + \cdots + x_m)\cdots(x_1 + x_2 + \cdots + x_m)}_{\text{n times}} \tag{1.2}$$

Each term of this expression when expanded will be of the form

$$Cx_1^{k_1}x_2^{k_2}\cdots x_m^{k_m} \tag{1.3}$$

where $\sum_{i=1}^{m} k_i = n$ and C is a numerical coefficient. To determine the value of C, consider the following method. Suppose we choose x_1 from k_1 sets of parentheses; there are $\binom{n}{k_1}$ ways this can be done because there is no more than one of x_1 in each set. Now when we choose x_2 from k_2 sets of parentheses, we cannot choose the same set that we have already chosen x_1 from. This means that we are left with $n-k_1$ sets from which we can choose x_2. So the number of ways to choose x_1 from k_1 sets and x_2 from k_2 sets is as follows:

$$\binom{n}{k_1}\binom{n-k_1}{k_2} \tag{1.4}$$

Following the same approach for the remaining x values, we can see that the coefficient for each term can be represented as such:

$$\binom{n}{k_1}\binom{n-k_1}{k_2}\binom{n-k_1-k_2}{k_3}\cdots\binom{n-k_1-k_2-\cdots-k_{m-1}}{k_m} \tag{1.5}$$

By definition,

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{1.6}$$

When we expand our coefficient by the definition above, many of the terms will cancel, leaving the following value for determining the coefficient:

$$\frac{n!}{k_1!k_2!k_3!\cdots k_m!} \tag{1.7}$$

Finally, to obtain every term from the expansion of $(x_1 + x_2 + \cdots + x_m)^n$, we add together every possible combination of $k_1 + k_2 + \cdots + k_m = n$.

$$\sum_{\sum_{i=1}^{m} k_i = n} \frac{n!}{k_1!k_2!\cdots k_m!} x_1^{k_1}x_2^{k_2}\cdots x_m^{k_m} \tag{1.8}$$

Induction Proof

Proof. We will prove this with induction on m. To start, we show that this holds for $m = 1$.

$$(x_1)^n = \sum_{\sum_{i=1}^{1} k_i = n} \binom{n}{k_1} x_1^{k_1} = x_1^n \tag{2.1}$$

Next, suppose the multinomial theorem holds for m. Then

$$(x_1 + x_2 + \cdots + (x_m + x_{m+1}))^n = \sum_{\sum_{i=1}^{m-1} k_i + K = n} \binom{n}{k_1 k_2 \cdots k_{m-1} K} x_1^{k_1}x_2^{k_2}\cdots x_{m-1}^{k_{m-1}}(x_m + x_{m+1})^K$$

(2.2)

Applying the binomial theorem to the right-hand side gives us

$$(x_1 + x_2 + \cdots + (x_m + x_{m+1}))^n$$
$$= \sum_{\sum_{i=1}^{m-1} k_i + K = n} \left[\binom{n}{k_1 k_2 \cdots k_{m-1} K} x_1^{k_1} x_2^{k_2} \cdots x_{m-1}^{k_{m-1}} \sum_{k_m + k_{m+1} = K} \binom{K}{k_m k_{m+1}} x_m^{k_m} x_{m+1}^{k_{m+1}} \right]$$

(2.3)

Then since

$$\binom{n}{k_1 k_2 \cdots k_{m-1} K} \binom{K}{k_m k_{m+1}} = \frac{n!}{k_1! k_2! \cdots k_{m-1}! K!} \cdot \frac{K!}{k_m! k_{m+1}!} = \frac{n!}{k_1! k_2! \cdots k_m! k_{m+1}!} \quad (2.4)$$

it follows that

$$(x_1 + x_2 + \cdots + (x_m + x_{m+1}))^n = \sum_{\sum_{i=1}^{m} k_i = n} \binom{n}{k_1 k_2 \cdots k_{m+1}} x_1^{k_1} x_2^{k_2} \cdots x_{m+1}^{k_{m+1}}$$

(2.5)

Since we now have shown that $m \Rightarrow m + 1$, we can conclude by the principle of induction that this statement holds for all integers m greater than or equal to 1.

Probability Proof

The following is a proof in Kataria (2016), which is an extension of the proof in Rosalsky (2007). Consider an experiment with n independent trials. The outcome of each trial results in the occurrence of one of the m mutually exclusive and exhaustive events $E_1, E_2, \cdots, E_m$. For each $i = 1,2,\cdots,m$, let p_i be the constant probability of the occurrence of the event E_i and X_i be the random variable that denotes the number of times event E_i has occurred. Then, the joint probability mass function of the random variables $X_1, X_2, \cdots, X_m$ is

$$P\{X_1 = k_1, X_2 = k_2, \cdots, X_m = k_m\} = n! \prod_{j=1}^{m} \frac{p_j^{k_j}}{k_j!} \quad (3.1)$$

where $\sum_{i=1}^{m} k_i = n$. Also, since (3.1) is a valid statistical distribution, we have

$$1 = \sum_{\sum_{i=1}^{m} k_i = n} n! \prod_{j=1}^{m} \frac{p_j^{k_j}}{k_j!} \quad (3.2)$$

By using the distributive property in (1.2), it follows that for all real number x_i's,

$$(x_1 + x_2 + \cdots + x_m)^n = \sum_{\sum_{i=1}^m k_i = n} C(n, k_1, k_2, \cdots, k_m) {x_1}^{k_1} {x_2}^{k_2} \cdots {x_m}^{k_m} \quad (3.3)$$

where $C(n,k_1,k_2,\cdots,k_m)$ are positive integers and k_i's are nonnegative integers which satisfy $\sum_{i=1}^m k_i = n$. Now we need to show that

$$C(n, k_1, k_2, \cdots, k_m) = \frac{n!}{k_1! k_2! \cdots k_m!} \quad (3.4)$$

Assume $x_i > 0$ for all $i = 1,2,\cdots,m$ and define

$$p_i = \frac{x_i}{x_1 + x_2 + \cdots + x_m} \quad (3.5)$$

It follows that $0 < p_i < 1$ and $\sum_{i=1}^m p_i = 1$. Substituting (3.5) into (3.2), we obtain for positive reals

$$(x_1 + x_2 + \cdots + x_m)^n = \sum_{\sum_{i=1}^m k_i = n} \frac{n!}{k_1! k_2! \cdots k_m!} x_1^{k_1} x_2^{k_2} \cdots x_m^{k_m} \quad (3.6)$$

Finally, subtracting (3.6) from (3.3),

$$\sum_{\sum_{i=1}^m k_i = n} \left(C(n, k_1, k_2, \cdots, k_m) - \frac{n!}{k_1!\, k_2! \cdots k_m!} \right) {x_1}^{k_1} {x_2}^{k_2} \cdots {x_m}^{k_m} = 0, x_i > 0 \quad (3.7)$$

Since (3.7) shows the left-hand side equals zero when subtracting (3.6) from (3.3), it follows that (3.4) is true.

Proof with Differential Calculus

The following proof extends Hwang's proof of the binomial theorem in Hwang (2009) using differential calculus into the multinomial theorem.

Again, it follows that upon distribution that for any integer n

$$(x_1 + x_2 + \cdots + x_m)^n = \sum_{\sum_{i=1}^m k_i = n} C(n, k_1, k_2, \cdots, k_m) \prod_{i=1}^m {x_i}^{k_i} \quad (4.1)$$

where the k_i's are nonnegative integers and $C(n, k_1, k_2, \cdots, k_m)$ are positive integers.
Given any set of nonnegative integers $c_1, c_2, \cdots, c_m$ such that $\sum_{i=1}^m c_i = n$, we calculate the partial derivatives of both sides of (4.1) with respect to each x_i c_i times for $i = 1, 2, \cdots, m$.

For the left side of (4.1), since $\sum_{i=1}^{m} c_i = n$,

$$\frac{\partial^n}{\partial x_1^{c_1}\partial x_2^{c_2}\cdots\partial x_m^{c_m}}(x_1 + x_2 + \cdots + x_m)^n = n! \tag{4.2}$$

For the right side of (4.1), if and only if $c_i = k_i$ for all $i = 1,2,\cdots,m$, then

$$\frac{\partial^n}{\partial x_1^{c_1}\partial x_2^{c_2}\cdots\partial x_m^{c_m}}\prod_{i=1}^{m} x_i^{k_i} = \frac{\partial^n}{\partial x_1^{k_1}\partial x_2^{k_2}\cdots\partial x_m^{k_m}}\prod_{i=1}^{m} x_i^{k_i} = k_1!\,k_2!\cdots k_m! \tag{4.3}$$

Otherwise,

$$\frac{\partial^n}{\partial x_1^{c_1}\partial x_2^{c_2}\cdots\partial x_m^{c_m}}\prod_{i=1}^{m} x_i^{k_i} = 0 \tag{4.4}$$

Therefore,

$$\frac{\partial^n}{\partial x_1^{c_1}\partial x_2^{c_2}\cdots\partial x_m^{c_m}}\left[\sum_{\sum_{i=1}^{m} k_i=n} C(n,k_1,k_2,\cdots,k_m)\prod_{i=1}^{m} x_i^{k_i}\right] = C(n,k_1,k_2,\cdots,k_m)k_1!\,k_2!\cdots k_m! \tag{4.5}$$

Then from (4.1), (4.2), and (4.5), we have for nonnegative integers $k_i$satisfying$\sum_{i=1}^{m} k_i = n$,

$$C(n,k_1,k_2,\cdots,k_m)k_1!\,k_2!\cdots k_m! = n!$$

Which means that

$$C(n,k_1,k_2,\cdots,k_m) = \frac{n!}{k_1!\,k_2!\cdots k_m!}$$

$$(x_1 + x_2 + \cdots + x_m)^n = \sum_{\sum_{i=1}^{m} k_i=n} \frac{n!}{k_1!\,k_2!\cdots k_m!} x_1^{k_1} x_2^{k_2} \cdots x_m^{k_m}$$

Thus, we have collected multiple methods of proving the multinomial theorem.

1. A combinatorial proof.
2. A proof by induction.
3. A probability proof.
4. A proof with differential calculus.

METHOD

A case study was conducted in an undergraduate math major junior and senior level topic class at Farmingdale State College, with a total of 12 students. The aim of this survey includes:

1. Test the ability of junior and senior level college math major students to prove the multinomial theorem;
2. Examine math major undergraduate students' attitudes toward presenting multiple proof approaches for the multinomial theorem;

Step 1: The following question was asked in class.

Do you think it is valuable to present multiple proof approaches for a same mathematics statement? Can you give one or more examples that can be proved in different methods?

Step 2: The multinomial theorem was presented in class.

Step 3: The combinatory and induction methods were presented in class.

Step 4: The probability and differential calculus method were presented in class.

Step 5: The following question was asked in class:

Assuming you are the instructor, will you present multiple proofs of the multinomial theorem? Will you require your students to know all of them?

Results

In step 1, all 12 students agreed that presenting multiple proof approaches for a same mathematics statement is important and valuable. However, only 6 students could provide meaningful examples. With instructor's hints, they recalled the proofs of Pythagorean theorem and some geometry and combinatory properties.

In step 2, although 8 students claimed familiarity of the multinomial theorem, initially none of the students felt confident in proving it completely.

In step 3, all 12 students followed the combinatory and induction proofs comfortably, with some remembering their use in proving the binomial theorem.

In step 4, none of the students had learned the probability method or the differential calculus method before. After a brief review of the same proof methods for the binomial theorem, they all gained better understanding of the same proof methods applied to the multinomial theorem and appeared to be impressed with these two additional proofs, especially the probability one.

In step 5, students were allowed to discuss this question. Their consensus was with a suggestion that at least present the combinatorial and induction methods. If class time permits, consider introducing the probability proof as well. The differential calculus method can be left as optional homework for interested students. This approach allows flexibility and caters to varying levels of interest and readiness. Some students did worry that presenting more than two methods simultaneously might overwhelm and confuse the class.

DISCUSSION

While some students struggled initially, exposure to various methods surely enhanced their understanding. It can develop students' divergent reasoning (Kwon et al. 2006), as well as their mental flexibility and fluency (Dreyfus, Nardi & Leikin, 2012; Leikin, 2009; Silver, 1997; Sriraman, 2003). As an instructor, presenting multiple proofs can enrich students' mathematical experience and foster deeper comprehension of theorems. From an educational viewpoint, such a comparison provides teachers and students with interesting connections between different viewpoints. Of course, the perspective presented requires a good level of epistemological skill on the part of teachers (Bagni, 2008).

CONCLUSION

For the multinomial theorem, the classroom study indicates that students find the induction approach most rigorous. Connecting it to the simpler binomial theorem, which they are already familiar with, makes it more accessible. Additionally, the combinatorial proof by counting also provides a concrete interpretation and students who enjoying combinatorial structure tend to find this approach appealing. Some students also like the probability proof, especially after gaining a clear understanding of the ideas presented in Rosalsky (2007). However, some students are not accustomed to the differential calculus method, as it can feel quite abstract. Only students with a strong background in multivariable calculus tend to follow it through well.

In general, comparison of different proofs can be an appropriate method to make the nature of proof visible to the students (Pfeiffer, 2010). Educators should consider using alternative methods for proof, which can provide students with alternative strategies to approach complex problems and enhance their understanding of underlying concepts (Mowahed & Mayar, 2023). Exposure to diverse proofs also hones students' problem-solving skills. They learn to adapt, generalize, and apply techniques across different scenarios (Stylianides & Ball, 2008). In an undergraduate-level mathematics course, when introducing the multinomial theorem, instructors can cover the induction and combinatorial methods with students during classroom lectures. For the probability and differential calculus approaches, instructors can provide students with materials related to proofs for binomial theorem, such as those found in reference Rosalsky (2007) and Hwang (2009), and encourage them to extend these concepts to the multinomial theorem case. This approach allows students to explore this topic from different angles and deepen their understanding. Further interested instructors and advanced students can even refer to Noble (2022) for a detailed historical background review.

References

[1] Dreyfus, T., Nardi, E., & Leikin, R. (2012). Forms of proof and proving in the classroom. Proof and proving in mathematics education: The 19th ICMI study, 191-213. https://doi.org/10.1007/978-94-007-2129-6_20

[2] Polya, G. (2004). How to solve it: A new aspect of mathematical method (No. 246). Princeton university press.

[3] Schoenfeld, A. H. (2014). Mathematical problem solving. Elsevier.

[4] Bagni, G. T. (2008). Theorem and Its Different Proofs: History, Mathematics Education, and the Semiotic-Cultural Perspective. Canadian Journal of Science, Mathematics and Technology Education, 8(3), 217-232. https://doi.org/10.1080/14926150802169297

[5] Stupel, M., & Ben-Chaim, D. (2017). Using multiple solutions to mathematical problems to develop pedagogical and mathematical thinking: A case study in a teacher education program. Investigations in Mathematics Learning, 9(2), 86–108. https://doi.org/10.1080/19477503.2017.1283179

[6] Kataria, K. K. (2016). A probabilistic proof of the multinomial theorem. The American Mathematical Monthly, 123(1), 94-96.

[7] Rosalsky, A. (2007). A simple and probabilistic proof of the binomial theorem. The American Statistician, 61(2), 161-162. https://doi.org/10.1198/000313007X188397

[8] Hwang, L. C. (2009). A simple proof of the binomial theorem using differential calculus. The American Statistician, 63(1), 43-44. https://doi.org/10.1198/tast.2009.0009

[9] Noble, E. (2022). A History of the Binomial and Multinomial Theorems. In The Rise and Fall of the German Combinatorial Analysis (pp. 17-66). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-93820-8_2

[10] Pfeiffer, K. (2010). How do students evaluate and compare mathematical proofs? Research in Mathematics Education, 12(2), 161–162. https://doi.org/10.1080/14794802.2010.496986

[11] Mowahed, A. K., & Mayar, J. A. (2023). Problematic and Supportive Aspects of Indirect Proof in Afghan Undergraduate Students' Proofs of the Irrationality of $\sqrt{3}$ and $\sqrt{5/8}$. Mathematics Teaching Research Journal, 15(4), 124-135. https://eric.ed.gov/?id=EJ1409364

[12] Stylianides, A. J., & Ball, D. L. (2008). Understanding and describing mathematical knowledge for teaching: Knowledge about proof for engaging students in the activity of proving. Journal of mathematics teacher education, 11, 307-332. https://doi.org/10.1007/s10857-008-9077-9

[13] Flusser, P., & Francia, G. A. (2000). Derivation and visualization of the binomial theorem. International Journal of Computers for Mathematical Learning, 5, 3-24. https://doi.org/10.1023/A:1009873212702

[14] Leikin, R. (2009). Multiple proof tasks: Teacher practice and teacher education. ICME Study, 19(2), 31-36.

[15] Stupel, M., & Ben-Chaim, D. (2013). One problem, multiple solutions: How multiple proofs can connect several areas of mathematics. Far East Journal of Mathematical Education, 11(2), 129.

[16] National Council of Teachers of Mathematics (2000), Principles and Standards for School Mathematics, NCTM, Reston, VA, 2000.

[17] Kwon, O. N., Park, J. S., & Park, J. H. (2006). Cultivating divergent thinking in mathematics through an open-ended approach. Asia Pacific Education Review, 7, 51–61. https://doi.org/10.1007/BF03036784

[18] Leikin, R. (2009b). Exploring mathematical creativity using multiple solution tasks. In R. Leikin, A. Berman, & B. Koichu (Eds.), Creativity in mathematics and the education of gifted students (pp. 129–145). Rotterdam: Sense Publishers.

[19] Silver, E. A. (1997). Fostering creativity through instruction rich in mathematical problem solving and problem posing. ZDM – The International Journal on Mathematics Education, 3, 75–80. https://10.1007/s11858-997-0003-x

[20] Sriraman, B. (2003). Mathematical giftedness, problem solving, and the ability to formulate generalizations. Journal of Secondary Gifted Education, 14, 151–165. https://doi.org/10.4219/jsge-2003-425