A GENERALIZATION OF CHU-VANDERMONDE’S IDENTITY

SEYED SAEED NAGHIBI AND MOHSEN HOOSHMAND

Abstract. We present and prove a general form of Vandermonde’s identity and use it as an alternative solution to a classic probability problem.

1. Introduction

Vandermonde’s identity, its diverse proofs, and applications have been the spot of research throughout the centuries. Nowadays, it is a classic identity whose proof is available in most combinatorics-related books [2]. However, there are still some efforts to give new proofs for this well-established identity, e.g., [3]. In addition to its original form proof, some specific types of its generalization have been proposed. Some works concentrated on its higher dimension [8]. Some others put their efforts into its complex [9] or coefficients of its numbers [2].

However, our focus was not generalizing this identity directly. We were doing some exercises in Stochastic Process and reached a classic problem during its teaching, which we put in this paper with an alternative solution in section 3. To have a formal proof of it, we reached an equation that we later realized could be considered a generalization of Vandermonde’s identity.

This article gives a general form of Vandermonde’s identity, including first order and higher orders of this identity. Then, we proof the mentioned classic problem, borrowed from Ross book [1] using the first order of general Vandermonde’s identity. Lastly, we briefly discuss some of our thoughts about this proof.

2. The Vandermonde’s identity General form.

At first, we encountered the first order of Vandermonde’s identity. We prove it as follows.

Theorem 2.1 (First Order of Vandermonde’s Identity General Form.). Let assume $k \leq m$ and $r-k \leq n$, then $\sum_{k=0}^{r} k^m \binom{m}{k} \binom{n}{r-k} = m^{m+n-1}$, where $\binom{x}{y} = \frac{x!}{(x-y)!y!}$, and $x, y \in \mathbb{N}$.

Proof. Let start with

$$m (x + 1)^{m+n-1} = m (x + 1)^{m-1} (x + 1)^n$$

By $m (1 + x)^{m-1} = \sum_{i=0}^{m} i^m \binom{m}{i} x^{i-1}$ [3] we have,

$$= \sum_{i=0}^{m} i^m \binom{m}{i} x^{i-1} \sum_{j=0}^{n} \binom{n}{j} x^j.$$
We get factor of $\frac{1}{x}$ from $x^{i-1}$ in the above equation; thus, we have
\[ \frac{1}{x} \sum_{i=0}^{m} i \binom{m}{i} x^i \sum_{j=0}^{n} \binom{n}{j} x^j. \]

Using polynomial ring, $\sum_{i=0}^{m} a_i x^i \sum_{j=0}^{n} b_j x^j = \sum_{r=0}^{m+n} \left( \sum_{k=0}^{r} a_k b_{r-k} \right) x^r$, we convert the above formula as follows.
\[ \frac{1}{x} \sum_{r=0}^{m+n} \left( \sum_{k=0}^{r} k \binom{m}{k} \binom{n}{r-k} \right) x^r = \sum_{r=0}^{m+n} \left( \sum_{k=0}^{r} k \binom{m}{k} \binom{n}{r-k} \right) x^r \]
\[ = 0 + \sum_{r=1}^{m+n} \left( \sum_{k=0}^{r} k \binom{m}{k} \binom{n}{r-k} \right) x^{r-1}. \]
or
\[ m (x+1)^{m+n-1} = \sum_{r=1}^{m+n} \left( \sum_{k=0}^{r} k \binom{m}{k} \binom{n}{r-k} \right) x^{r-1}. \]
The equation 2.1 is our first expansion. On the other hand, we consider binomial expansion,
\[ (x + 1)^{m+n-1} = \sum_{r=0}^{m+n-1} \binom{m+n-1}{r} x^r. \]
By changing the index variable, the equation is
\[ = \sum_{r=1}^{m+n} \binom{m+n-1}{r-1} x^{r-1}. \]
Multiplying both sides with $m$
\[ m (x+1)^{m+n-1} = \sum_{r=1}^{m+n} m \binom{m+n-1}{r-1} x^{r-1}. \]
The left sides of equation 2.1 and equation 2.2 are equal. Thus,
\[ \sum_{r=1}^{m+n} \left( \sum_{k=0}^{r} k \binom{m}{k} \binom{n}{r-k} \right) x^{r-1} = \sum_{r=1}^{m+n} m \binom{m+n-1}{r-1} x^{r-1}, \]
and results in
\[ \sum_{k=0}^{r} k \binom{m}{k} \binom{n}{r-k} = m \binom{m+n-1}{r-1}. \]

\[ \boxed{\text{Theorem 2.1 is a leading part of an alternative solution to the problem in section 3.}} \]
However, going further, we realized that the Theorem 2.1 is a particular case of a more general form. Thus, we introduce the general form of Vandermonde’s identity as the below theorem.
Theorem 2.2 (General Form of Vandermonde’s Identity). Let assume \( l \leq k,\) 
\( k \leq m \) and \( r - k \leq n \) then,

\[
\sum_{k=0}^{r} \left[ \binom{m}{k} \binom{n}{r-k} \right] = \left[ \binom{m+n-l}{r-l} \right],
\]

where \( \binom{x}{y} = \binom{x}{x-y} \) and \( \binom{y}{x} = \binom{x}{x-y}^{-1} \), and \( x, y \in \mathbb{N} \).

Proof. Let \( f(x) = (1 + x)^m \). its binomial expansion is

\[
f(x) = \sum_{i=0}^{m} \binom{m}{i} x^i.
\]

Then, utilizing induction, its \( l \)-th derivative is

\[
f^{(l)}(x) = \left[ \binom{m}{l} (1 + x)^{m-l} \right] = \sum_{i=0}^{m} \binom{m}{l} \binom{m}{i} x^{i-l}.
\]

Now, we start with \( \left[ \binom{m}{l} (x+1)^{m+n-l} \right]. \) We have

\[
\left[ \binom{m}{l} \right] (x+1)^{m+n-l} = \left[ \binom{m}{l} \right] (x+1)^{m-l} (x+1)^n
\]

using equation 2.5 and in a similar way to the corresponding steps of Theorem 2.1, we utilize the polynomial ring and the equation is as follows.

\[
\text{or}
\]

\[
\left[ \binom{m}{l} \right] (x+1)^{m+n-l} = \sum_{r=l}^{m+n} \sum_{k=0}^{r} \left[ \binom{k}{l} \binom{m}{k} \binom{n}{r-k} \right] x^{r-l}
\]

On the other hand, we use binomial expansion and have

\[
(x+1)^{m+n-l} = \sum_{r=0}^{m+n-l} \binom{m+n-l}{r} x^r
\]

changing the index variable and multiplying both sides with \( \left[ \binom{m}{l} \right] \)

\[
\left[ \binom{m}{l} \right] (x+1)^{m+n-l} = \sum_{r=l}^{m+n-l} \left[ \binom{m}{l} \binom{m+n-l}{r-l} \right] x^{r-l}
\]

The left-hand sides of both equations 2.6 and equation 2.7 are the same, so the right-hand side of both equations are the same as well, or simply

\[
\sum_{k=0}^{r} \left[ \binom{k}{l} \binom{m}{k} \binom{n}{r-k} \right] = \left[ \binom{m}{l} \binom{m+n-l}{r-l} \right].
\]
Remark 2.3. In Theorem 2.2, if

1. $l = 0$, then it is equivalent with Vandermonde’s identity, or

   \begin{equation}
   \sum_{k=0}^{r} \binom{m}{k} \binom{n}{r-k} = \binom{m+n}{r};
   \end{equation}

2. $l = 1$, then it is equivalent with first order of general order of Vandermonde’s identity;

3. $l > 1$, then it is equivalent to higher order of general form of Vandermonde’s identity.

3. Application.

   In this section, we present the query which led us to face the first order of Vandermonde’s identity general form. We prove the problem using Theorem 2.1.

3.1. Statement of the Problem. “An urn has $r$ red and $w$ white balls that are randomly removed one at a time. Let $R_i$ be the event that the $i\text{th}$ ball removed is red. Find $P(R_i)$.”

Solution. The answer is \( \frac{r}{r+w} \). Also sprach Ross “… each of the $r + w$ balls is equally to be the $i\text{th}$ ball removed”. In addition to the intuition introduced by Sheldon Ross, we would like to have a direct alternative formal proof of the answer. Thus, we tried to prove it, and here it is.

Proof. Let assume $r_j$ denotes choosing of $j$ red balls in $i - 1$ previous steps. Then,

\[ P(R_i) = \sum_{j=0}^{i-1} P(R_i \cap r_j) \]

By applying Law of total probability we have,

\[ = \sum_{j=0}^{i-1} P(R_i|r_j) P(r_j) \]

$P(r_j)$, as mentioned above, is equal to choosing $j$ red balls among the $i-1$ balls that we have already picked. Thus, its value is equal to \( \frac{\binom{j}{r} \binom{i-r}{i-1}}{\binom{i}{i-1}} \). In addition, $P(R_i|r_j)$ denotes choosing a red ball in the $i\text{th}$ step. Because, we have already picked $j$ red balls in the previous steps, its value is equal to \( \frac{\binom{j}{r}}{\binom{i-j}{i-1}} \). Consequently, we have,
A GENERALIZATION OF CHU-VANDERMONDE’S IDENTITY

\[\sum_{j=0}^{i-1} \binom{r-j}{i-1-j} \binom{w}{i-1-j} \binom{r+w}{i-1} = \sum_{j=0}^{i-1} \frac{(r-j)}{(r+w-i-1)} \binom{w}{i-1-j} \binom{r+w}{i-1} \]

\[= \frac{1}{(r+w-i+1)(r+w)} \sum_{j=0}^{i-1} (r-j)(j)(i-j) \]

\[= \frac{1}{(r+w-i+1)(r+w)} \sum_{j=0}^{i-1} r \binom{r+j}{i-j} - \sum_{j=0}^{i-1} j \binom{r+j}{i-j} \]

The \(\sum_{j=0}^{i-1} j \binom{r+j}{i-j}\) term in the above formula is in accordance with Theorem 2.1. Thus, we proceed with it as follows,

\[= \frac{1}{(r+w-i+1)(r+w)} \left( r \binom{r+w}{i-1} - r \binom{r+w-1}{i-2} \right) \]

\[= \frac{r}{(r+w-i+1)(r+w)} \left( \binom{r+w}{i-1} - \binom{r+w-1}{i-2} \right) \]

\[= \frac{r}{(r+w-i+1)(r+w-1)} \left( \binom{r+w}{i-1} - \frac{i-1}{r+w} \binom{r+w}{i-1} \right) \]

\[= \frac{r}{(r+w-i+1)(r+w-1)} \left( \binom{r+w}{i-1} - \frac{i-1}{r+w} \binom{r+w}{i-1} \right) \]

\[= \frac{r}{r+w} \]

\[= \frac{1}{(r+w-i+1)(r+w)} \left( r \binom{r+w}{i-1} - r \binom{r+w-1}{i-2} \right) \]

\[= \frac{r}{(r+w-i+1)(r+w-1)} \left( \binom{r+w}{i-1} - \frac{i-1}{r+w} \binom{r+w}{i-1} \right) \]

\[= \frac{r}{r+w} \]

4. Final Thoughts

There have been many efforts to expand or extend Vandermonde’s identity or Vandermonde’s convolution. For example, Graham et al. [2] defined some general forms of identities, including Vandermonde’s identity, in chapter 5. But, that generalization is different from our point of view. They are a generalization of Vandermonde’s identity, discussing the coefficients of \(m\) or \(n\) in the equation. The same happens in works such as [3, 7]. Yaccov [8] investigated the determinant of the Vandermonde in a higher dimension. However, to our knowledge, we proposed this type of generalization of Vandermonde’s identity for the first time, or at least we have given a new way of this generalization proof. In our proposal, we consider the general form of Vandermonde’s identity from the point of view of Combinatorics’ coefficients. It is worth mentioning that it is possible to observe our proposed generalization as a higher derivative of Vandermonde’s identity.
References

1. Ross, S.M. (2010). *A First Course in Probability*. Pearson Hall.
2. Graham, R., Knuth, D. E., Patashnik, O. (1994). *Concrete Mathematics: A Foundation for Computer Science*. Addison-Wesley Longman Publishing Co., Inc.
3. Koh, K.-M., Chen, C. C. (1992). *Principles and Techniques in Combinatorics*. World Scientific Publishing.
4. Whitelaw, T. A. (1991). *An Introduction to Linear Algebra*. Springer US.
5. Gasper, G., Rahman, M. (2004). *Basic Hypergeometric Series*. Cambridge.
6. Spivey, M. (2017). The Chu–Vandermonde Identity via Leibniz’s Identity for Derivatives. *The College Mathematics Journal*. 47(3):219. doi. 10.4169/college.math.j.47.3.219
7. Fang, J.P. (2007). Extensions of q-Chu–Vandermonde’s identity. *Journal of Mathematical Analysis and Applications* 339(2):845–852. doi.10.1016/j.jmaa.2007.07.029
8. Yaacov, I. B. (2017). The Vandermonde determinant identity in higher dimension. hal-00983584v4. doi.org/10.48550/arXiv.1405.0993
9. Mestrović, R. (2018). Several generalizations and variations of Chu-Vandermonde iden-
   [arxiv.org/abs/1807.10604]ty.

Dept. of Computer Science and Information Technology, Institute for Advanced Studies in Basic Sciences (IASBS), Zanjan, Iran

Email address: ngsaeed@iasbs.ac.ir

Dept. of Computer Science and Information Technology, Institute for Advanced Studies in Basic Sciences (IASBS), Zanjan, Iran

Email address: mohsen.hooshmand@iasbs.ac.ir