Binomial coefficients, valuations, and words

Eric Rowland

Department of Mathematics
Hofstra University
Hempstead, NY 11549, USA
eric.rowland@hofstra.edu

Abstract. The study of arithmetic properties of binomial coefficients has a rich history. A recurring theme is that p-adic statistics reflect the base-p representations of integers. We discuss many results expressing the number of binomial coefficients $\binom{n}{m}$ with a given p-adic valuation in terms of the number of occurrences of a given word in the base-p representation of n, beginning with a result of Glaisher from 1899, up through recent results by Spiegelhofer–Wallner and Rowland.

Keywords: binomial coefficients, p-adic valuation, regular sequences

1 Valuations of binomial coefficients

In 1852, Kummer [10, pages 115–116] determined the exact power of a prime p that divides a binomial coefficient $\binom{n}{m}$. To state this result, let $\nu_p(n)$ denote the p-adic valuation of n, that is, the exponent of the highest power of p dividing n.

Theorem 1 (Kummer). *Let p be a prime, and let n and m be integers with $0 \leq m \leq n$. Then $\nu_p(\binom{n}{m})$ is the number of carries involved in adding m to $n-m$ in base p.*

Kummer's theorem is the first of many results to express arithmetic information about binomial coefficients in terms of base-p representations of integers.

Glaisher [6, §14] seems to have been the first to count binomial coefficients satisfying a given congruence condition. He showed that the number of integers m in the range $0 \leq m \leq n$ such that $\binom{n}{m}$ is odd is $2^{|n|_1}$. Here $|n|_1$ is the number of 1s in the standard base-2 representation of n.

Half a century later, Glaisher's result was generalized to an arbitrary prime by Fine [5, Theorem 2]. For a prime p and an integer $n \geq 0$, let $\theta_{p,0}(n)$ be the number of integers m in the range $0 \leq m \leq n$ such that $\binom{n}{m} \not\equiv 0 \mod p$. Let $|n|_w$ be the number of occurrences of the word w in the base-p representation of n. Fine showed that

$$\theta_{p,0}(n) = \prod_{d=0}^{p-1} (d+1)^{|n|_d}.$$

Since the publication of Fine's result, many authors have been interested in generalizations to higher powers of p. A natural quantity to study is the number $\theta_{p,\alpha}(n)$ of binomial coefficients $\binom{n}{m}$, for $0 \leq m \leq n$, with $\nu_p(\binom{n}{m}) = \alpha$.

Carlitz [3, Equations (1.7)–(1.9)] gave a recurrence involving $\theta_{p,\alpha}(n)$ and a secondary quantity $\psi_{p,\alpha}(n)$ defined as the number of integers m in the range $0 \leq m \leq n$ such that $\nu_p((m+1)\binom{n}{m}) = \alpha$. Namely,

$$\theta_{p,\alpha}(pn+d) = (d+1)\theta_{p,\alpha}(n) + (p-d-1)\psi_{p,\alpha-1}(n-1)$$
$$\psi_{p,\alpha}(pn+d) = \begin{cases} (d+1)\theta_{p,\alpha}(n) + (p-d-1)\psi_{p,\alpha-1}(n-1) & \text{if } 0 \leq d \leq p-2 \\ p\psi_{p,\alpha-1}(n) & \text{if } d = p-1. \end{cases}$$

As we will see in Section 3, this recurrence comes close to giving a matrix generalization of Fine's theorem, but it is not of the right form.

Nonetheless, Carlitz's recurrence can be used to obtain formulas for $\theta_{p,\alpha}(n)$. Let $n_\ell \cdots n_1 n_0$ be the standard base-p representation of n. For $\alpha = 1$, Carlitz [3, Equation (2.5)] showed

$$\theta_{p,1}(n) = \sum_{i=0}^{\ell-1} (n_\ell+1)(n_{\ell-1}+1)\cdots(n_{i+2}+1)n_{i+1}(p-n_i-1)(n_{i-1}+1)\cdots(n_0+1).$$

Dividing by $\theta_{p,0}(n) = (n_\ell+1)\cdots(n_0+1)$ gives

$$\frac{\theta_{p,1}(n)}{\theta_{p,0}(n)} = \sum_{i=0}^{\ell-1} \frac{n_{i+1}}{n_{i+1}+1} \cdot \frac{p-n_i-1}{n_i+1}.$$

This equation is our first indication that expressions for $\frac{\theta_{p,\alpha}(n)}{\theta_{p,0}(n)}$ can be simpler than expressions for $\theta_{p,\alpha}(n)$ alone. In particular, $\frac{\theta_{p,1}(n)}{\theta_{p,0}(n)}$ is a polynomial in the variables $|n|_w$ for words $w \in \{0,1,\ldots,p-1\}^*$ of length 2. For $p=2$ we obtain

$$\theta_{2,1}(n) = 2^{|n|_1} \cdot \frac{1}{2}|n|_{10},$$

which was obtained by Howard [7, Equation (2.4)] and by Davis and Webb [4, Theorem 7]. For $p=3$ we have

$$\theta_{3,1}(n) = 2^{|n|_1}3^{|n|_2}\left(|n|_{10} + \frac{1}{4}|n|_{11} + \frac{4}{3}|n|_{20} + \frac{1}{3}|n|_{21}\right),$$

which also follows from the work of Huard, Spearman, and Williams [8]. For $p=5$ we have

$$\begin{aligned}\theta_{5,1}(n) = 2^{|n|_1}3^{|n|_2}4^{|n|_3}5^{|n|_4}\biggl(&2|n|_{10} + \frac{3}{4}|n|_{11} + \frac{1}{3}|n|_{12} + \frac{1}{8}|n|_{13} \\ &+ \frac{8}{3}|n|_{20} + |n|_{21} + \frac{4}{9}|n|_{22} + \frac{1}{6}|n|_{23} \\ &+ 3|n|_{30} + \frac{9}{8}|n|_{31} + \frac{1}{2}|n|_{32} + \frac{3}{16}|n|_{33} \\ &+ \frac{16}{5}|n|_{40} + \frac{6}{5}|n|_{41} + \frac{8}{15}|n|_{42} + \frac{1}{5}|n|_{43}\biggr),\end{aligned}$$

and so on.

2 Formulas for arbitrary prime powers

We have seen that $\frac{\theta_{p,1}(n)}{\theta_{p,0}(n)}$ is a polynomial in $|n|_w$. In fact this is true for $\frac{\theta_{p,\alpha}(n)}{\theta_{p,0}(n)}$ in general. Barat and Grabner [2, §3] showed this implicitly while studying the asymptotic behavior of $\sum_{n=0}^{N} \theta_{p,\alpha}(n)$.

Rowland [12] gave an algorithm for computing a polynomial expression for $\frac{\theta_{p,\alpha}(n)}{\theta_{p,0}(n)}$. For $p = 2$ one computes

$$\theta_{2,2}(n) = 2^{|n|_1}\left(-\frac{1}{8}|n|_{10} + |n|_{100} + \frac{1}{4}|n|_{110} + \frac{1}{8}|n|_{10}^2\right)$$

(which was obtained by Howard [7, Equation (2.5)] and by Huard, Spearman, and Williams [9, Theorem C]),

$$\theta_{2,3}(n) = 2^{|n|_1}\left(\frac{1}{24}|n|_{10} - \frac{1}{2}|n|_{100} - \frac{1}{8}|n|_{110} + 2|n|_{1000} + \frac{1}{2}|n|_{1010} + \frac{1}{2}|n|_{1100}\right.$$
$$\left. + \frac{1}{8}|n|_{1110} - \frac{1}{16}|n|_{10}^2 + \frac{1}{2}|n|_{10}|n|_{100} + \frac{1}{8}|n|_{10}|n|_{110} + \frac{1}{48}|n|_{10}^3\right),$$

and so on. The number of nonzero terms in the polynomial $\frac{\theta_{2,\alpha}(n)}{2^{|n|_1}}$ for $\alpha = 0, 1, 2, \ldots$ is sequence A275012 [11]:

$1, 1, 4, 11, 29, 69, 174, 413, 995, 2364, 5581, 13082, 30600, 71111, 164660, 379682, \ldots$

The algorithm for computing these polynomials also establishes bounds on their total degree and on the length of words that appear.

Theorem 2 (Rowland [12]). *Let p be a prime, and let $\alpha \geq 0$. Then $\frac{\theta_{p,\alpha}(n)}{\theta_{p,0}(n)}$ is given by a polynomial of degree α in $|n|_w$ for words w satisfying $|w| \leq \alpha + 1$.*

However, the algorithm is not particularly fast, as it constructs a polynomial by summing over certain sets of integer partitions. Spiegelhofer and Wallner [14] produced a faster algorithm by developing a better understanding of the structure of this polynomial. In particular, they showed that the polynomial representation of $\frac{\theta_{p,\alpha}(n)}{\theta_{p,0}(n)}$ is unique, as long as words of the form $0w$ and $w(p-1)$ are not used. Therefore one can talk about *the* coefficient of a given monomial. Moreover, this coefficient can be read off from a certain power series. One can see evidence of this by looking at the coefficient of $|n|_{10}$ in the expressions computed for $\frac{\theta_{2,\alpha}(n)}{2^{|n|_1}}$. The sequence of coefficients is

$$0, \frac{1}{2}, -\frac{1}{8}, \frac{1}{24}, -\frac{1}{64}, \frac{1}{160}, -\frac{1}{384}, \ldots.$$

These are the coefficients in the power series for $\log(1 + \frac{x}{2})$ at $x = 0$.

The polynomial

$$T_p(n, x) := \sum_{m=0}^{n} x^{\nu_p(\binom{n}{m})} = \sum_{\alpha \geq 0} \theta_{p,\alpha}(n) x^\alpha$$

is a central object in Spiegelhofer and Wallner's result. Let

$$\overline{T}_p(w,x) := \frac{T_p(\mathrm{val}_p(w),x)}{\theta_{p,0}(\mathrm{val}_p(w))},$$

where $\mathrm{val}_p(w)$ is the integer obtained by reading w in base p. Define the rational function

$$r_p(w,x) := \frac{\overline{T}_p(w,x)\overline{T}_p(w_{LR},x)}{\overline{T}_p(w_R,x)\overline{T}_p(w_L,x)},$$

where the left and right truncations of a word are defined for $\ell \geq 0$, $c \in \{1,\dots,p-1\}$, and $d \in \{0,1,\dots,p-1\}$ by

$$\begin{array}{lll} \epsilon_L = \epsilon & (c0^\ell)_L = \epsilon & (c0^\ell w)_L = w \\ \epsilon_R = \epsilon & c_R = \epsilon & (wd)_R = w. \end{array}$$

Theorem 3 (Spiegelhofer–Wallner [14]). *Let p be a prime, and let $\alpha \geq 0$. Let $w_1,\dots,w_m$ be words of length ≥ 2 on the alphabet $\{0,1,\dots,p-1\}$ that do not begin with 0 or end with $p-1$. Then the coefficient of $|n|_{w_1}^{k_1}\cdots|n|_{w_m}^{k_m}$ in the polynomial $\frac{\theta_{p,\alpha}(n)}{\theta_{p,0}(n)}$ is the coefficient of x^α in the power series expansion for*

$$\frac{1}{k_1!}(\log r_p(w_1,x))^{k_1}\cdots\frac{1}{k_m!}(\log r_p(w_m,x))^{k_m}.$$

3 Matrix generalizations of Fine's theorem

Spiegelhofer and Wallner's polynomial $T_p(n,x)$ turns out to have a product formula which generalizes Fine's theorem. Note that the first equation in Carlitz's recurrence,

$$\theta_{p,\alpha}(pn+d) = (d+1)\theta_{p,\alpha}(n) + (p-d-1)\psi_{p,\alpha-1}(n-1),$$

can be rewritten in terms of $T_p(n,x)$ as

$$T_p(pn+d,x) = (d+1)\,T_p(n,x) + \begin{cases} 0 & \text{if } n=0 \\ (p-d-1)\,x^{\nu_p(n)+1}\,T_p(n-1,x) & \text{if } n \geq 1. \end{cases}$$

To simplify this equation, let us introduce a secondary polynomial

$$T'_p(n,x) := \begin{cases} 0 & \text{if } n=0 \\ x^{\nu_p(n)+1}\,T_p(n-1,x) & \text{if } n \geq 1, \end{cases}$$

so that $\psi_{p,\alpha-1}(n-1)$ is the coefficient of x^α in $T'_p(n,x)$. Then we have

$$T_p(pn+d,x) = (d+1)\,T_p(n,x) + (p-d-1)T'_p(n,x).$$

We are close to being able to write

$$\begin{bmatrix} T_p(pn+d,x) \\ T'_p(pn+d,x) \end{bmatrix} = M_p(d)\begin{bmatrix} T_p(n,x) \\ T'_p(n,x) \end{bmatrix} \tag{1}$$

for some 2×2 matrix $M_p(d)$, thereby expressing $T_p(pn+d,x)$ and $T'_p(pn+d,x)$ in terms of $T_p(n,x)$ and $T'_p(n,x)$. Carlitz's second equation,

$$\psi_{p,\alpha}(pn+d) = \begin{cases} (d+1)\theta_{p,\alpha}(n) + (p-d-1)\psi_{p,\alpha-1}(n-1) & \text{if } 0 \le d \le p-2 \\ p\psi_{p,\alpha-1}(n) & \text{if } d = p-1, \end{cases}$$

expresses $\psi_{p,\alpha}(pn+d)$ in terms of θ and ψ. But because the coefficient of x^α in $T'_p(n,x)$ is $\psi_{p,\alpha-1}(n-1)$ we instead need to express $\psi_{p,\alpha}(pn+d-1)$ in terms of θ and ψ. The desired equation is

$$\psi_{p,\alpha}(pn+d-1) = d\theta_{p,\alpha}(n) + (p-d)\psi_{p,\alpha-1}(n-1),$$

which is equivalent to

$$T'_p(pn+d,x) = d\,x\,T_p(n,x) + (p-d)\,x\,T'_p(n,x).$$

Therefore, the matrix we seek is

$$M_p(d) = \begin{bmatrix} d+1 & p-d-1 \\ d\,x & (p-d)\,x \end{bmatrix}.$$

Theorem 4 (Rowland [13]). *Let p be a prime, and let $n \ge 0$. Let $n_\ell \cdots n_1 n_0$ be the standard base-p representation of n. Then*

$$T_p(n,x) = \begin{bmatrix} 1 & 0 \end{bmatrix} M_p(n_0)\, M_p(n_1) \,\cdots\, M_p(n_\ell) \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Setting $x = 0$ gives Fine's result as a special case. Namely, the definition of $T'_p(n,x)$ implies $T'_p(n,0) = 0$, so Equation (1) becomes

$$\begin{bmatrix} \theta_{p,0}(pn+d) \\ 0 \end{bmatrix} = \begin{bmatrix} d+1 & p-d-1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \theta_{p,0}(n) \\ 0 \end{bmatrix},$$

or simply

$$\theta_{p,0}(pn+d) = (d+1)\,\theta_{p,0}(n).$$

Moreover, Theorem 4 generalizes naturally to multinomial coefficients. For a k-tuple $\mathbf{m} = (m_1, m_2, \ldots, m_k)$ of non-negative integers, define

$$\operatorname{total}\mathbf{m} := m_1 + m_2 + \cdots + m_k$$

and

$$\operatorname{mult}\mathbf{m} := \frac{(\operatorname{total}\mathbf{m})!}{m_1!\, m_2! \,\cdots\, m_k!}.$$

Let $c_{p,k}(n)$ be the coefficient of x^n in $(1 + x + x^2 + \cdots + x^{p-1})^k$. For each $d \in \{0, 1, \ldots, p-1\}$, let $M_{p,k}(d)$ be the $k \times k$ matrix whose (i,j) entry is

$c_{p,k}(p\,(j-1)+d-(i-1))\,x^{i-1}$. For example, let $p=5$ and $k=3$; the matrices $M_{5,3}(0),\ldots,M_{5,3}(4)$ are

$$\begin{bmatrix} 1 & 18 & 6 \\ 0 & 15x & 10x \\ 0 & 10x^2 & 15x^2 \end{bmatrix}, \quad \begin{bmatrix} 3 & 19 & 3 \\ x & 18x & 6x \\ 0 & 15x^2 & 10x^2 \end{bmatrix}, \quad \begin{bmatrix} 6 & 18 & 1 \\ 3x & 19x & 3x \\ x^2 & 18x^2 & 6x^2 \end{bmatrix},$$
$$\begin{bmatrix} 10 & 15 & 0 \\ 6x & 18x & x \\ 3x^2 & 19x^2 & 3x^2 \end{bmatrix}, \quad \begin{bmatrix} 15 & 10 & 0 \\ 10x & 15x & 0 \\ 6x^2 & 18x^2 & x^2 \end{bmatrix}.$$

Theorem 5 (Rowland [13]). *Let p be a prime, let $k \geq 1$, and let $n \geq 0$. Let $e = \begin{bmatrix} 1\,0\,0 \cdots 0 \end{bmatrix}$ be the first standard basis vector in $\mathbb{Z}^k$. Let $n_\ell \cdots n_1 n_0$ be the standard base-p representation of n. Then*

$$\sum_{\substack{\mathbf{m}\in\mathbb{N}^k \\ \operatorname{total}\mathbf{m}=n}} x^{\nu_p(\operatorname{mult}\mathbf{m})} = e\, M_{p,k}(n_0)\, M_{p,k}(n_1) \,\cdots\, M_{p,k}(n_\ell)\, e^\top.$$

The proof essentially amounts to showing that, for $d \in \{0,1,\ldots,p-1\}$, $0 \leq i \leq k-1$, and $\alpha \geq 0$, the map β defined by

$$\beta(\mathbf{m}) := (\lfloor \mathbf{m}/p \rfloor, \mathbf{m} \bmod p)$$

is a bijection from the set

$$A = \left\{\mathbf{m} \in \mathbb{N}^k : \operatorname{total}\mathbf{m} = pn+d-i \text{ and } \nu_p(\operatorname{mult}\mathbf{m}) = \alpha - \nu_p\!\left(\frac{(pn+d)!}{(pn+d-i)!}\right)\right\}$$

to the set

$$B = \bigcup_{j=0}^{k-1} \left(\left\{\mathbf{c} \in \mathbb{N}^k : \operatorname{total}\mathbf{c} = n-j \text{ and } \nu_p(\operatorname{mult}\mathbf{c}) = \alpha - \nu_p\!\left(\frac{n!}{(n-j)!}\right) - j\right\} \right.$$
$$\left. \times \left\{\mathbf{d} \in \{0,1,\ldots,p-1\}^k : \operatorname{total}\mathbf{d} = pj+d-i\right\} \right).$$

The following lemma implies that if $\mathbf{m} \in A$ then $\beta(\mathbf{m}) \in B$.

Lemma 6. *Let p be a prime, $k \geq 1$, $n \geq 0$, $d \in \{0,1,\ldots,p-1\}$, and $0 \leq i \leq k-1$. Let $\mathbf{m} \in \mathbb{N}^k$ with $\operatorname{total}\mathbf{m} = pn+d-i$. Let $j = n - \operatorname{total}\lfloor \mathbf{m}/p \rfloor$. Then $\operatorname{total}(\mathbf{m} \bmod p) = pj+d-i$, $0 \leq j \leq k-1$, and*

$$\nu_p\!\left(\frac{(pn+d)!}{(pn+d-i)!}\right) + \nu_p(\operatorname{mult}\mathbf{m}) = \nu_p\!\left(\frac{n!}{(n-j)!}\right) + \nu_p(\operatorname{mult}\lfloor \mathbf{m}/p \rfloor) + j.$$

We conclude by mentioning a connection to regular sequences. A sequence $s(n)_{n\geq 0}$, with entries in some field, is *p-regular* if the vector space generated by the set of subsequences $\{s(p^e n+i)_{n\geq 0} : e \geq 0 \text{ and } 0 \leq i \leq p^e-1\}$ is finite-dimensional. For example, the sequence $(\theta_{p,0}(n))_{n\geq 0}$ is a p-regular sequence of integers [1, Example 14]. It follows from Theorem 5 and [1, Theorem 2.2] that the sequence of polynomials

$$\left(\sum_{\substack{\mathbf{m} \in \mathbb{N}^k \\ \text{total}\, \mathbf{m} = n}} x^{\nu_p(\text{mult}\, \mathbf{m})} \right)_{n \geq 0}$$

is p-regular for each k.

References

1. Jean-Paul Allouche and Jeffrey Shallit, The ring of k-regular sequences, *Theoretical Computer Science* **98** (1992) 163–197.
2. Guy Barat and Peter J. Grabner, Distribution of binomial coefficients and digital functions, *Journal of the London Mathematical Society* **64** (2001) 523–547.
3. Leonard Carlitz, The number of binomial coefficients divisible by a fixed power of a prime, *Rendiconti del Circolo Matematico di Palermo* **16** (1967) 299–320.
4. Kenneth Davis and William Webb, Pascal's triangle modulo 4, *The Fibonacci Quarterly* **29** (1989) 79–83.
5. Nathan Fine, Binomial coefficients modulo a prime, *The American Mathematical Monthly* **54** (1947) 589–592.
6. James W. L. Glaisher, On the residue of a binomial-theorem coefficient with respect to a prime modulus, *Quarterly Journal of Pure and Applied Mathematics* **30** (1899) 150–156.
7. Fred T. Howard, The number of binomial coefficients divisible by a fixed power of 2, *Proceedings of the American Mathematical Society* **29** (1971) 236–242.
8. James Huard, Blair Spearman, and Kenneth Williams, Pascal's triangle $\pmod 9$, *Acta Arithmetica* **78** (1997) 331–349.
9. James Huard, Blair Spearman, and Kenneth Williams, Pascal's triangle $\pmod 8$, *European Journal of Combinatorics* **19** (1998) 45–62.
10. Ernst Kummer, Über die Ergänzungssätze zu den allgemeinen Reciprocitätsgesetzen, *Journal für die reine und angewandte Mathematik* **44** (1852) 93–146.
11. The OEIS Foundation, The On-Line Encyclopedia of Integer Sequences, `http://oeis.org`.
12. Eric Rowland, The number of nonzero binomial coefficients modulo p^α, *Journal of Combinatorics and Number Theory* **3** (2011) 15–25.
13. Eric Rowland, A matrix generalization of a theorem of Fine, `https://arxiv.org/abs/1704.05872`.
14. Lukas Spiegelhofer and Michael Wallner, An explicit generating function arising in counting binomial coefficients divisible by powers of primes, `https://arxiv.org/abs/1604.07089`.