

# A brief note on estimates of binomial coefficients

Shagnik Das

## Introductory remarks

Through the ages, humankind has struggled to come to terms with its own existence, a thorny issue that will no doubt keep philosophers tossing and turning in their beds for aeons to come<sup>1</sup>. The ‘how’ of our existence, though still fiercely debated in some corners, is now reasonably well-understood, but it is the ‘why’ that continues to plague us. Why are we here? What purpose are we meant to serve? Thus we continue to live our lives, with no end in sight on our quest of self-discovery.

However, this document, were it capable of intelligent introspection, would face no such quandary, for it has a very simple reason<sup>2</sup> for being. While presenting a proof in our Extremal Combinatorics course, in an ill-advised attempt to avoid Stirling’s Approximation, I made quite a mess of what should have been a routine calculation. I have thus written this note in atonement for my grievous error in judgement.

Within you shall find a brief survey of some useful bounds on binomial coefficients<sup>3</sup>, which hopefully covers what I tried to say in that ill-fated lecture. I claim no responsibility for any errors herein<sup>4</sup>, but if you do detect an error, I would be grateful if you would let me know, and I will issue a patch as soon as I am able.

## Binomial coefficients

Enough with the preamble, then; let us meet the main character of this essay — the binomial coefficient. While I suspect this object is one familiar to you all<sup>5</sup>, I shall define it here (quickly) for the sake of completeness.

The binomial coefficient,  $\binom{n}{k}$ , admits two parameters,  $n$  and  $k$ , and for our purposes we shall always have  $0 \leq k \leq n$ , with  $k$  and  $n$  both integers. The coefficient can be formulaically defined as below

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{\prod_{j=0}^{k-1} (n-j)}{k!}.$$

However, its prevalence in combinatorics is due to the fact that this humble expres-

---

<sup>1</sup>Until, that is, the Great Robot Uprising, when we humans are finally laid to rest, and the burden of this existential crisis is passed on to the victorious machines. These steel-plated philosophers will indubitably continue to toss and turn, but not in beds, because robots have no need of sleep.

<sup>2</sup>The ‘how’ is even more straightforward: it was typed on my laptop.

<sup>3</sup>“This much,” you protest, exasperatedly, “was evident from the title! Why did you not just get to the point?” I did consider opening with the definition of the binomial coefficient, and proceeding immediately to providing some bounds, but it felt improper, if not downright rude, to write a note without an introduction.

<sup>4</sup>Not because any such errors are due to a third party, but rather because I could do without additional responsibilities at this point.

<sup>5</sup>It is sometimes said (disparagingly) that combinatorics is naught but the study of binomial coefficients. “Haters gonna hate” [Swift, 2014].

sion appears when counting<sup>6</sup> items of natural interest. Different combinatorics<sup>7</sup> will have their own preferred examples of the coefficients appearing in everyday life<sup>8</sup>. Personally<sup>9</sup>, I prefer to interpret  $\binom{n}{k}$  as the number of  $k$ -element subsets of an  $n$ -element set. Among other viewpoints, one may also consider  $\binom{n}{k}$  to be the number of ways to get  $k$  heads in a sequence of  $n$  coin tosses. Whichever interpretation you choose to adopt, the binomial coefficient will follow your every step<sup>10</sup>, so you might as well befriend it.

## Bounds

Beautiful and versatile though the binomial coefficients may be, they come with a catch — we humans are not very good at working with them. Our brains can handle addition, multiplication, and, if really required, exponentiation. However, ask someone to compute, or even estimate,  $\binom{732}{32}\binom{32}{12} + \binom{1023}{94}$ , and they will most probably find some way to excuse themselves from your conversation. For this reason, we often seek some bounds on the binomial coefficients that are more convenient to work with<sup>11</sup>.

We begin with the simplest upper bound, which can often be useful when the binomial is a lower-order term.

$$\forall 0 \leq k \leq n : \binom{n}{k} \leq 2^n \quad (1)$$

To see why this is true, recall that  $\binom{n}{k}$  counts the number of subsets of  $[n]$  of size  $k$ , while  $2^n$  counts *all* subsets of  $[n]$ .

The upper bound in (1) is certainly easy to use, but is often a gross overestimate, and is thus not suitable when greater precision is needed. The following standard inequalities provide much better bounds on the size of the coefficients, and are used throughout the field.

$$\forall 1 \leq k \leq n : \left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{ne}{k}\right)^k \quad (2)$$

One can prove the lower bound in (2) by writing  $\binom{n}{k} = \frac{\prod_{j=0}^{k-1} (n-j)}{k!} = \prod_{j=0}^{k-1} \frac{n-j}{k-j}$ . Each factor in this product is at least  $\frac{n}{k}$ , and there are  $k$  factors, giving the lower bound  $\left(\frac{n}{k}\right)^k$ . For the upper bound, we need to use  $k! \geq \left(\frac{k}{e}\right)^k$ . Then  $\binom{n}{k} = \frac{\prod_{j=0}^{k-1} (n-j)}{k!} \leq \frac{n^k}{k!} \leq \left(\frac{ne}{k}\right)^k$ ,

---

<sup>6</sup>It is often said (more truthfully) that combinatorics is the art of counting.

<sup>7</sup>The correct term for one who studies combinatorics is still subject to discussion, with ‘combinatorialist’ and ‘combinatorist’ currently the leading expressions. While the polls remain open, though, I shall persist in my attempts to popularise my own preferred job title.

<sup>8</sup>Indeed, a combinator’s favourite application of the binomial coefficients may well be a more reliable means of identification than fingerprinting or DNA sequencing.

<sup>9</sup>I have made every attempt to keep this note secular and devoid of personal prejudices, but at this point I cannot help but offer you this window to my soul.

<sup>10</sup>Like a loyal, trustworthy dog, or that piece of toilet paper you can’t shake off your shoe, depending on your feelings towards binomials.

<sup>11</sup>In the sense that they are easier to manipulate; what we look for in our bounds is what con artists desire in their marks.

as desired. The lower bound on  $k!$  can be proven by induction on  $k$ , using the fact that for all  $k \geq 1$ ,  $(1 + \frac{1}{k})^k \leq (e^{\frac{1}{k}})^k = e$ .

When faced with a calculation involving binomial coefficients, these bounds offer much better control, and often suffice to determine the correct order of magnitude of the quantity in question. For instance, we applied them when proving our exponential lower bound on the diagonal Ramsey numbers  $R(t, t)$ .

## Asymptotics

Sometimes, though, we are not satisfied with just getting the order of magnitude correct, and ask for even more precision. The bounds in (2), useful though they may be, are still a factor of  $e^k$  apart. This begs the question: which bound better reflects the truth? It turns out that the answer depends on the relative sizes of  $k$  and  $n$ . If  $k$  is small compared to  $\sqrt{n}$ , then  $(\frac{ne}{k})^k$  is the better estimate for  $\binom{n}{k}$ . At the other extreme, if  $k = n$ , then  $(\frac{n}{k})^k$  is exactly equal to  $\binom{n}{k}$ , while  $(\frac{ne}{k})^k = e^n$  even exceeds our trivial upper bound of  $2^n$ .

For a more accurate answer, rather than seeking bounds on  $\binom{n}{k}$ , we can ask for the actual asymptotics — up to lower-order error terms, what is the binomial coefficient actually equal to? Unlike the bounds (1) and (2), which applied for all  $n$  and  $k$  with  $1 \leq k \leq n$ , here we will assume  $n = \omega(1)$ .<sup>12</sup> As you might suspect from the above discussion, the asymptotics will depend on how large  $k$  is. By symmetry, since  $\binom{n}{k} = \binom{n}{n-k}$ , we may assume  $0 \leq k \leq \frac{n}{2}$ , which in particular implies  $n - k \geq \frac{n}{2} = \omega(1)$ .

**Case I:**  $k = o(\sqrt{n})$

Here we use the representation

$$\binom{n}{k} = \frac{\prod_{j=0}^{k-1} (n-j)}{k!}.$$

Observe that

$$n^k \geq \prod_{j=0}^{k-1} (n-j) \geq (n-k)^k = \left(1 - \frac{k}{n}\right)^k n^k \geq \left(1 - \frac{k^2}{n}\right) n^k = (1 - o(1)) n^k,$$

and so, for  $0 \leq k = o(\sqrt{n})$ ,

$$\binom{n}{k} = (1 + o(1)) \frac{n^k}{k!}. \tag{3}$$

If  $k$  is in fact constant, then this is the best approximation one can hope for. However, when  $k = \omega(1)$  (but still  $k = o(\sqrt{n})$ ), the  $k!$  term is a little inconvenient. We can replace it with an exponential expression by making use of Stirling's Approximation.

---

<sup>12</sup>In other words,  $n$  tends to infinity.

**Theorem 1** (Stirling's<sup>13</sup> Approximation<sup>14</sup>). As  $m \rightarrow \infty$ ,

$$m! = (1 + o(1))\sqrt{2\pi m} \left(\frac{m}{e}\right)^m.$$

Applying Theorem 1 with  $m = k$ , we find that when  $\omega(1) = k = o(\sqrt{n})$ ,

$$\binom{n}{k} = (1 + o(1)) \frac{1}{\sqrt{2\pi k}} \left(\frac{ne}{k}\right)^k. \quad (4)$$

Thus in this range, the upper bound in (2) is tight up to a multiplicative  $O(\sqrt{k})$ -error.

**Case II:**  $k = \omega(1)$

If  $k \neq o(\sqrt{n})$ , then we cannot bound  $\prod_{j=0}^{k-1} (n - j)$  in the manner above. Instead, since now  $n$ ,  $k$ , and  $n - k$  are all tending to infinity, we can use Stirling's approximation to evaluate all three factorials. This gives

$$\begin{aligned} \binom{n}{k} &= \frac{n!}{k!(n-k)!} = (1 + o(1)) \frac{\sqrt{2\pi n} \left(\frac{n}{e}\right)^n}{\sqrt{2\pi k} \left(\frac{k}{e}\right)^k \cdot \sqrt{2\pi(n-k)} \left(\frac{n-k}{e}\right)^{n-k}} \\ &= (1 + o(1)) \sqrt{\frac{n}{2\pi k(n-k)}} \left(\frac{n}{k}\right)^k \left(\frac{n}{n-k}\right)^{n-k}. \end{aligned} \quad (5)$$

In practice, (5) is far too cumbersome to use. Thus, when  $k$  is large, we are happy to asymptotically determine  $\log \binom{n}{k}$ , rather than  $\binom{n}{k}$  itself. While this gives a less precise result, it is often much more useful. In this setting, there is a qualitative difference based on whether  $k$  is comparable to  $n$  or not. Note that all logarithms here are binary.

**Case IIa:**  $k = o(n)$

Recall that  $\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{ne}{k}\right)^k$ , which gives

$$k \log \frac{n}{k} \leq \log \binom{n}{k} \leq k \log \frac{ne}{k} = k \left( \log \frac{n}{k} + \log e \right).$$

Since  $k = o(n)$ ,  $\frac{n}{k}$  tends to infinity, and thus the  $\log e$  term is a lower-order error. Thus

$$\log \binom{n}{k} = (1 + o(1))k \log \frac{n}{k}. \quad (6)$$

In particular, the right-hand side of (6) is  $o(n)$ , which means the binomial coefficient  $\binom{n}{k}$  is subexponential (in  $n$ ).

---

<sup>13</sup>De Moivre was the first to discover this form of the approximation, while Stirling determined the constant  $\sqrt{2\pi}$ .

<sup>14</sup>In fact, the quantity on the right-hand side without the  $(1 + o(1))$  term is always a lower-bound for  $m!$ , while if the  $\sqrt{2\pi}$  is replaced by  $e$ , one obtains an upper bound for  $m!$  (these are valid for all  $m \geq 1$ ).

**Case IIb:**  $k = \Omega(n)$

The story changes when  $k$  is linear in  $n$ . Taking the logarithm of (5) gives

$$\log \binom{n}{k} = \log(1 + o(1)) + \log \sqrt{\frac{n}{2\pi k(n-k)}} + k \log \frac{n}{k} + (n-k) \log \frac{n}{n-k}.$$

Since both  $k$  and  $n-k$  are linear in  $n$ , they dwarf the other two terms. Thus we can simplify the above to

$$\log \binom{n}{k} = (1 + o(1)) \left( \frac{k}{n} \log \frac{n}{k} + \frac{n-k}{n} \log \frac{n}{n-k} \right) n = (1 + o(1)) H \left( \frac{k}{n} \right) n, \quad (7)$$

where  $H(p) = -p \log p - (1-p) \log(1-p)$  is the *binary entropy function*, defined for  $p \in [0, 1]$ . Hence, when  $k = cn$  for some fixed constant  $c \in (0, \frac{1}{2}]$ ,  $\binom{n}{k}$  is approximately  $2^{H(c)n}$ . Most importantly, we see that when  $k$  is linear,  $\binom{n}{k}$  grows exponentially in  $n$ .

## Real-world examples

We are now five pages into my brief note, and have encountered more than the recommended daily allowance of factorials and logarithms. You might<sup>15</sup> at this point stop and wonder why you have subjected yourself to this. They say knowledge is power, which, if true, implies you have become more powerful by virtue of now knowing more about the growth rate of the binomial coefficient (unless, of course, you had to forget some other knowledge to make space for this). If knowledge for knowledge's sake is not motivation enough, then in the remainder of this note I would like to draw upon some examples from the real world to show that these bounds truly are worth knowing.

**The middle coefficient** Of all the binomial coefficients, the case when  $n = 2k$  is arguably the most interesting. In this case,  $\frac{k}{n} = \frac{1}{2}$ , and since  $H(\frac{1}{2}) = 1$ , (7) gives  $\binom{2k}{k} = 2^{(1+o(1))2k}$ . Recalling that the trivial upper bound (1) gives  $\binom{2k}{k} = 2^{2k}$ , this shows that the middle binomial coefficient is close to being as large as possible. Upon further reflection, this should not be surprising. Indeed, if we were to choose a (uniformly) random subset of a set of  $2k$  elements, we would expect it to have size  $k$ . It is not too hard to show that not only is  $k$  the expected size, it is also the most popular size. Since there are  $2k+1$  possible sizes, that implies  $\frac{2^{2k}}{2k+1} \leq \binom{2k}{k} \leq 2^{2k}$ , a much stronger bound than that given by (7).

The truth lies halfway between these two bounds, as can be shown by our asymptotic results. If we substitute  $n = 2k$  into (5), we get

$$\binom{2k}{k} = (1 + o(1)) \sqrt{\frac{2k}{2\pi k(2k-k)}} \left( \frac{2k}{k} \right)^k \left( \frac{2k}{2k-k} \right)^{2k-k} = (1 + o(1)) \frac{2^{2k}}{\sqrt{\pi k}}. \quad (8)$$

<sup>15</sup>Then again, you might not, in which case you may<sup>16</sup> find it hard to relate to the contents of this paragraph.

<sup>16</sup>Then again, you may not, in which case the above footnote might not apply to you.

Hence we see that the number of sets is in fact a  $\Theta(\sqrt{k})$ -fraction, not just a  $O(k)$ -fraction, of all sets. Another way to reason why this would be true is to note that the binomial distribution with  $2k$  trials and probability  $\frac{1}{2}$  (which gives a uniform distribution over subsets of a set of  $2k$  elements) approximately follows a normal distribution with mean  $k$  and variance  $\frac{k}{2}$ , or standard deviation  $\Theta(\sqrt{k})$ . Thus, while there are  $2k+1$  possible sizes of sets, almost all sets have sizes ranging over an interval of only  $\Theta(\sqrt{k})$  sizes. As the most common size is  $k$ , we can expect  $\binom{2k}{k}$  to count a  $\Theta(\sqrt{k})$ -fraction of the  $2^{2k}$  sets.

**Random greedy two-colouring of hypergraphs** Recall<sup>17</sup> that in the Pluhár proof of the lower bound on  $m_B(k)$ , we saw that the bad events occurred with probability

$$\frac{(k-1)!(k-1)!}{(2k-1)!}. \quad (9)$$

To obtain an expression that is easier to work with, we shall first rewrite (9) in terms of a binomial coefficient, and then use our asymptotic estimates. We have

$$\frac{(k-1)!(k-1)!}{(2k-1)!} = \frac{(k-1)!(k-1)!}{(2k-1)(2k-2)!} = \frac{1}{(2k-1)\binom{2k-2}{k-1}},$$

which, in light of (8), shows

$$\frac{(k-1)!(k-1)!}{(2k-1)!} = (1+o(1)) \frac{\sqrt{\pi(k-1)}}{2k-1} 2^{2-2k} = O\left(k^{-\frac{1}{2}} 2^{-2k}\right).$$

We then used this to show  $m_B(k) = \Omega\left(k^{\frac{1}{4}} 2^k\right)$ . Observe that the same result could have been obtained by applying Stirling's Approximation to (9) directly. Theorem 1 gives

$$\begin{aligned} \frac{(k-1)!(k-1)!}{(2k-1)!} &= (1+o(1)) \frac{\left[\sqrt{2\pi(k-1)} \left(\frac{k-1}{e}\right)^{k-1}\right]^2}{\sqrt{2\pi(2k-1)} \left(\frac{2k-1}{e}\right)^{2k-1}} \\ &= (1+o(1)) \frac{e\sqrt{\pi(k-1)}}{2k-1} \left(\frac{k-1}{2k-1}\right)^{2k-2} \\ &= (1+o(1)) \frac{e\sqrt{\pi(k-1)}}{2k-1} \left(\frac{k-1}{2k-2}\right)^{2k-2} \left(\frac{2k-2}{2k-1}\right)^{2k-2} \\ &= (1+o(1)) \frac{e\sqrt{\pi(k-1)}}{2k-1} 2^{2-2k} \left(1 - \frac{1}{2k-1}\right)^{2k-2}. \end{aligned}$$

Since  $\left(1 - \frac{1}{2k-1}\right)^{2k-2}$  is asymptotically  $e^{-1}$ , this matches our earlier result. That being said, (8) is easier to remember than Theorem 1, and, in this case at least, easier to apply.

---

<sup>17</sup>If the context of this calculation is one you are unfamiliar with, then please believe me that this was an important calculation.

## Concluding remarks

As you progress through your combinatorial career, these inequalities and estimates will no doubt become trusted companions, always there in your time of need. While one could write more about how important the binomial coefficients are, and perhaps also include a section containing some useful binomial identities, I shall, in the interest of brevity, refrain from doing so. I hope you enjoyed reading this brief note, and, until next time, bid you farewell.<sup>18</sup>

---

<sup>18</sup>I only included this footnote so that this page could also have a footnote, as it might otherwise be ostracised and ridiculed by the other (footnoted) pages.