

The question of the number of zeroes at the end of $N!$ is a chestnut dating back at least several decades. The well-known answer is

$$L_5(N) = \sum_{i=1}^{\infty} \lfloor N/5^i \rfloor.$$

Justification: There is exactly one trailing zero for each factor of 10 in $N!$, and each factor of 10 arises from a factor of 2 and a factor of 5. There are obviously many more of the former, so we need only count factors of 5. Every fifth integer contributes a factor of 5, total $\lfloor N/5 \rfloor$. Every twenty-fifth integer contributes another, for an additional $\lfloor N/5^2 \rfloor$, and so forth.

Now let's ask what happens when $N!$ is written in a base other than 10. Let $Z_b(N)$ denote the number of trailing zeroes when $N!$ is written in integral base $b \geq 2$. Once again, each such zero arises from a factor of b . Writing b as its unique prime factorization $b = 2^{b_2} 3^{b_3} \cdots p_i^{b_{p_i}} \cdots$ where p_i is the i^{th} prime, we see that to get a trailing zero we need b_2 factors of 2, b_3 factors of 3, and so forth. The number of factors of p in $N!$ is $L_p(N)$, so each p with $b_p > 0$ constrains the number of trailing zeros to be at most $\lfloor L_p(N)/b_p \rfloor$. The number we want is the minimum of all these constraints:

$$Z_b(N) = \min\{\lfloor L_{p_i}(N)/b_{p_i} \rfloor\} \tag{1}$$

where the minimum is taken over all i such that p_i divides b .

For $b = 10$ this formula becomes $\min\{L_2(N), L_5(N)\}$. But we've said that the answer is simply $L_5(N)$ because there are "obviously" many more factors of 2 than of 5. That is, 5 is the "limiting factor"; the L_2 term can be ignored because it's always greater than the L_5 term. In the rest of this note we examine this question more generally, asking when exactly can we ignore terms of (1) as we can ignore the L_2 term when $b = 10$.

We start by generalizing the observation that worked for $b = 10$: If p and q are positive integers with $p < q$, then clearly $L_p(N) \geq L_q(N)$. Hence when b has prime factors p and q with $p < q$ and $b_p \leq b_q$ there will always be enough factors of p to go around, that is, $\lfloor L_p(N)/b_p \rfloor$ is necessarily greater than $\lfloor L_q(N)/b_q \rfloor$ and the L_p term in (1) can be ignored.

So, for example, with $b = 13500 = 2^2 3^3 5^3$ we have $Z_b(N) = \lfloor Z_5(N)/3 \rfloor$ since the L_2 and L_3 terms in the minimum are necessarily larger than the L_5 term. Similarly, with $b = 1389150 = 2^1 3^4 5^2 7^3$ we can ignore the L_2 and L_5 terms, but not the others, giving $Z_b(N) = \min\{\lfloor L_3(N)/4 \rfloor, \lfloor L_7(N)/3 \rfloor\}$.

Copyright 2011 Larry Denenberg

But this can't be the whole story. Consider base $b = 3072 = 2^{10}3^1$. Clearly 2 is now the limiting factor: we need so many 2s for each 3 that there will always be 3s and to spare. Hence we need only count 2s, that is, we can ignore the L_3 term in (1). The rule above doesn't capture this case.

Let's estimate the value of $\lfloor Z_p(N)/b_p \rfloor$ by dropping all use of the floor function. The result is $N/(p-1)b_p$ since Z_p becomes simply the sum of a geometric sequence. So $1/(p-1)b_p$ is a multiplier that approximates the proportion of trailing zeros that p can contribute to $N!$. Now when b has factors p^{b_p} and q^{b_q} , the smaller of $1/(p-1)b_p$ and $1/(q-1)b_q$ indicates the limiting factor, and we can drop the term in (1) corresponding to the larger. It's easier to work with reciprocals, retaining the term corresponding to the *larger* of $(p-1)b_p$ and $(q-1)b_q$. In the $b = 3072$ example, $(2-1)10$ is greater than $(3-1)1$, hence we can discard the L_3 term of (1), as we had already concluded. Note that this rule subsumes the previous rule, since if $p < q$ and $b_p < b_q$ then necessarily $(q-1)b_q > (p-1)b_p$.

The correctness of this new rule depends on the following lemma: For primes p and q and positive integers b_p and b_q such that $(q-1)b_q > (p-1)b_p$, we have $\lfloor Z_p(N)/b_p \rfloor \geq \lfloor Z_q(N)/b_q \rfloor$ for all N . Is this lemma true?

It's easy to see that it's true for all sufficiently large N , that is, for any such p, q, b_p, b_q there exists N_0 such that the lemma is true for all $N > N_0$. Proof: Dropping a single use of the floor function increases a value by at most 1, so dropping all floor functions in $L_p(N)$ increases its value by at most $\log_p N$ plus a constant. Hence our estimate is too high by at most $(\log_p N)/b_p$ plus a constant. But the difference between two such estimates is N times a fixed constant, and so for all sufficiently large N exceeds the maximum possible error. Hence using the estimate yields the correct answer for large N . Based on numerical experimentation we conjecture that the lemma is in fact true for all N .

Finally, what happens when $(p-1)b_p$ is exactly equal to $(q-1)b_q$? Base $b = 12 = 2^23^1$ is a simple example: $(2-1) * 2 = (3-1) * 1 = 2$. Is there an additional criterion we can use to eliminate one term or the other? We conjecture that the answer is no, that is, we believe that for primes p and q and positive integers b_p and b_q such that $(p-1)b_p = (q-1)b_q$, we have $\lfloor Z_p(N)/b_p \rfloor > \lfloor Z_q(N)/b_q \rfloor$ for infinitely many N . (This conjecture is borne out by experimental evidence, but we have no other justification for it.) It follows that if there are two terms with maximal $(p-1)b_p$, both must be retained. Assuming the truth of the conjecture generalized in the obvious way to multiple primes, all terms with maximal $(p-1)b_p$ must be retained.

In summary, if our conjectures are true, then the terms of (1) that may be discarded are exactly those whose value of $(p_i - 1)b_{p_i}$ is not maximal.