A Constant-Time Formula for Counting Digit Occurrences in Structured Ranges

Rishav Vignesh

September 19, 2025

Abstract

We present a constant-time formula for counting the occurrences of a given digit in integer ranges of the form $[a \cdot 10^n,\ b \cdot 10^m]$, with $1 \leq a, b \leq 9$, $0 \leq n \leq m$ and digit $d \in \{1, \dots, 9\}$. The standard approach requires $O(\log N)$ time via digit-by-digit computation. Our result shows that in this restricted, yet common case, the count reduces to a simple arithmetic expression.

1 Introduction

Counting the occurrences of digits in integer ranges is a problem in elementary number theory, combinatorics, and computer science[1]. The standard method requires per-digit computation, yielding $O(\log N)$ complexity. In this note, we show that when the interval is made to align with powers of ten in the form $[a \cdot 10^n,\ b \cdot 10^m]$, the computation collapses into a form which is computable in $O(1)$ time.

2 Preliminaries

We consider digit $d \in \{1, \dots, 9\}$. Let $a, b \in \{1, \dots, 9\}$, and integers $n, m \geq 0$ with $n \leq m$. The problem is to compute:

$$C_d(a, n, b, m) = \#\{\text{occurrences of digit } d \text{ in } [a \cdot 10^n,\ b \cdot 10^m]\}.$$

3 Formula

Define the auxiliary function

$$E(n) = \begin{cases} n \cdot 10^{n-1}, & n \geq 1, \\ 0, & n = 0. \end{cases}$$

Then $\#\{$occurrences of d in integers of the range $[a{\cdot}10^n,\ b{\cdot}10^m]\}$ is given by

$$C_d(a, n, b, m) = b\,E(m) + \mathbf{1}_{\{b>d\}}\,10^m + \mathbf{1}_{\{b=d\}} - a\,E(n) - \mathbf{1}_{\{a>d\}}\,10^n,$$

where $\mathbf{1}_{\{P\}}$ is an indicator function (1 if P is true, 0 otherwise).

4 Proof

4.1 Correctness

Claim 1. For any digit $d \in \{1, \ldots, 9\}$ and integer $k \geq 1$,

$$S_k(d) = \#\{\text{occurrences of d in all integers } 0, ..., 10^k - 1\} = k \cdot 10^{k-1}$$

Proof. Consider each integer in the range $0, ..., 10^k - 1$ as a k-digit string by padding with leading zeros when necessary. There are 10^k such strings. This gives a one-to-one correspondence between the set $[0, 10^k - 1]$ and the set of all k-tuples $(a_{k-1}, a_{k-2}, ..., a_0)$ with each $a_i \in \{0, \ldots, 9\}$. There are 10^k such k-tuples. Fix a position j where $0 \leq j \leq k-1$. Count how many k-tuples have digit d in position j. The digit in position j is fixed as d; each of the remaining $k-1$ positions can be any of 10 choices independently. Hence there are exactly 10^{k-1} tuples with digit d in position j. Since there are k positions and for each fixed position j the digit d appears in exactly 10^{k-1} tuples, the total number of occurrences of d among all k-tuples is $k \cdot 10^{k-1}$.□

Claim 2. Let $m \geq 1$, and $b, d \in \{1, \ldots, 9\}$,

$$\#\{\text{occurrences of } d \text{ in } 0, ..., b \cdot 10^m - 1\} = b\,E(m) + \mathbf{1}_{\{b>d\}}10^m$$

and therefore

$$\#\{\text{occurrences of } d \text{ in } 0, ..., b \cdot 10^m\} = b\,E(m) + \mathbf{1}_{\{b>d\}}10^m + \mathbf{1}_{\{b=d\}}.$$

Proof. Partition $[0, b \cdot 10^m - 1]$ into b segments

$$S_i = [\, i \cdot 10^m,\ (i+1) \cdot 10^m - 1\,], \quad i = 0, 1, \ldots, b-1.$$

In each segment, the lower m digits range over all 10^m strings exactly once. By Claim 1, the lower m positions contribute $E(m) = m \cdot 10^{m-1}$ occurrences of digit d. Summing over b segments gives $b\,E(m)$ occurrences coming from the lower m positions. The most significant digit within S_i equals i, as per the definition of every number in S_i being of the form $i \cdot 10^m + x$ for some $x \in \mathbb{Z}$. It contributes an additional 10^m occurrences when $i = d$. Since $d \in \{1, \ldots, 9\}$, the index $i = d$ appears in the segment $\{0, \ldots, b-1\}$ if and only if $b - 1 \geq d$. This is equivalent to the term $\mathbf{1}_{\{b>d\}} 10^m$. Finally, the integer $b \cdot 10^m$ does not lie in the range $[0, b \cdot 10^m - 1]$, so we add it separately: it contributes one extra occurrence if and only if its most significant digit equals d, which can be stated in the formula $\mathbf{1}_{\{b=d\}}$. ☐

Claim 3. Let $n \geq 0$, and $a, d \in \{1, \ldots, 9\}$. Then

$$\#\{\text{occurrences of } d \text{ in } 0, \ldots, a \cdot 10^n - 1\} = a\,E(n) + \mathbf{1}_{\{a>d\}}\, 10^n.$$

Proof. If $n = 0$ then $a \cdot 10^0 - 1 = a - 1 < 9$ which reduces to $\mathbf{1}_{\{a>d\}}$ which is correct because d appears in $[0, a-1]$ if and only if $d \in \{1, \ldots, a-1\}$. For $n \geq 1$, partition $[0, a \cdot 10^n - 1]$ into a segments

$$[i \cdot 10^n, (i+1) \cdot 10^n - 1], \quad i = 0, 1, \ldots, a-1.$$

As stated in Claim 2, each segment contributes $E(n) = n \cdot 10^{n-1}$ occurrences in the lower n digits. Summing over all a segments gives a total $a\,E(n)$ occurrences. The most significant digit of each segment equals i, so it contributes 10^n occurrences when $i = d$. Since i is bound by the range $[0, a-1]$, this happens if and only if $d < a$, yielding the term $\mathbf{1}_{\{a>d\}}\, 10^n$. ☐

Remark. The formulas in the preceding claims exclude $d = 0$. This is because the segmentation argument relies on representing integers as strings with padded leading zeros, which introduces zeros that are not present in the original range. Handling $d = 0$ requires additional casework to account for leading-zero positions, so the stated formula is valid only for $d \in \{1, \ldots, 9\}$.

Theorem. For $d \in \{1, \ldots, 9\}$, $1 \leq a, b \leq 9$, and integers $n, m \geq 0$ with $n \leq m$, the number of occurrences of d in the range $[a \cdot 10^n,\ b \cdot 10^m]$ is

$$C_d(a, n, b, m) = b\,E(m) + \mathbf{1}_{\{b>d\}}\, 10^m + \mathbf{1}_{\{b=d\}} - a\,E(n) - \mathbf{1}_{\{a>d\}}\, 10^n.$$

Proof. By the inclusion and exclusion of terms,

$$C_d(a,n,b,m) = \#\{d \text{ in } 0,\ldots,b\cdot 10^m\} - \#\{d \text{ in } 0,\ldots,a\cdot 10^n - 1\}.$$

We can apply Claim 2 to the first term and Claim 3 to the second term to obtain the stated expression. □

4.2 Runtime

Proof. Under the unit-cost RAM model in which arithmetic and comparison operations on integers, regardless of their magnitude costs $O(1)$ and the word size is unbounded[2], the formula $C_d(a,n,b,m)$ evaluates in constant time. The expression requires only a fixed number of arithmetic and branching instructions independent of n and m, hence its run time is $O(1)$ in this model. □

Remark. In a bit-cost model where time depends on the bit-length of the integers manipulated. The integer 10^m has bit-length $\Theta(m)$, so constructing or operating on 10^m requires time proportional to that bit-length. Consequently, the bit-cost model would result in the evaluation of $C_d(a,n,b,m)$ taking time proportional to the maximum bit-length of the operands. Or more precisely $O(M(B))$ where B is the maximum bit-length of the operands and $M(B)$ denotes the cost of B-bit multiplication[3]. Hence, the claim of constant-time evaluation only holds under the unit-cost RAM model with unbounded word size; in practice, one must account for bit-level costs when m is large.

5 Tangible Example

To illustrate the formula in practice, consider the digit $d = 3$ in the interval $[10^2, 6\cdot 10^3]$, otherwise written as $[100, 6000]$. Applying the formula $C_d(a,n,b,m)$ with $a = 1$, $n = 2$, $b = 6$, and $m = 3$, we compute:

$$C_3(1,2,6,3) = 6\,E(3) + \mathbf{1}_{\{6>3\}}\,10^3 + \mathbf{1}_{\{6=3\}} - 1\,E(2) - \mathbf{1}_{\{1>3\}}\,10^2 = 2780$$

We can then evaluate this manually through per-digit summation to concretely verify our result:

- In the units place, 3 appears once every ten numbers; there are 5900 numbers in the range, giving 590 occurrences.

- In the tens place, 3 appears once every hundred numbers, yielding $59 \cdot 10 = 590$ occurrences.
- In the hundreds place, 3 appears once every thousand numbers, yielding $6 \cdot 100 = 600$ occurrences.
- In the thousands place, the digits 1 through 6 appear, so 3 appears in that position exactly 1000 times.

Summing, we obtain $590+590+600+1000 = 2780$. This matches the output of the formula, thereby confirming its correctness in this specific instance.

6 Real-World Applications

- Cryptography and randomness testing. If a certain digit appears more often in a range, bias is highlighted in the system.
- Statistical analysis. It is beneficial to have swift computation when counting digit occurrences in exceedingly large ranges.

7 Conclusion

In summary, we have presented a proof for the digit counting formula of the form $C_d(a, n, b, m)$ and analyzed its computational complexity under the unit-cost RAM model to be in $O(1)$ time complexity.

References

[1] D. E. Knuth, *The art of computer programming: Seminumerical algorithms, volume 2.* Addison-Wesley Professional, 2014.

[2] É. Grandjean and L. Jachiet, "Which arithmetic operations can be performed in constant time in the ram model with addition?" *arXiv preprint arXiv:2206.13851*, 2022.

[3] M. Fürer, "Faster integer multiplication," in *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, 2007, pp. 57–66.