Monstrous Menagerie with Vandehey 7/11

Jason Schuchardt

July 11, 2019

We're proving the lemma from the end of last time.

Proof. By Stirling's formula we have:

$$\begin{aligned}\binom{n}{cn} &= \frac{n!}{(cn)!((1-c)n)!} \\ &= \frac{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n\left(1+O\left(\frac{1}{n}\right)\right)}{\sqrt{2\pi cn}\left(\frac{cn}{e}\right)^{cn}\left(1+O\left(\frac{1}{cn}\right)\right)\sqrt{2\pi(1-c)n}\left(\frac{(1-c)n}{e}\right)^{(1-c)n}\left(1+O\left(\frac{1}{(1-c)n}\right)\right)} \\ &= \frac{1}{\sqrt{2\pi c(1-c)n}}\left[c^{-c}(1-c)^{-(1-c)}\right]^n \frac{1+O\left(\frac{1}{n}\right)}{\left(1+O\left(\frac{1}{cn}\right)\right)\left(1+O\left(\frac{1}{(1-c)n}\right)\right)}\end{aligned}$$

Then $O\left(c\frac{1}{n}\right) = O\left(\frac{1}{n}\right)$ (note that equals signs don't represent equality for asymptotic notation, so this is an asymmetric statement). Also

$$\frac{1}{1+O\left(\frac{1}{n}\right)} = 1 + O\left(\frac{1}{n}\right),$$

and

$$\left(1+O\left(\frac{1}{n}\right)\right)\left(1+O\left(\frac{1}{n}\right)\right) = 1 + O\left(\frac{1}{n}\right),$$

so we can reduce the last term in the product to just $1 + O(1/n)$, as desired. ■

Proposition 0.1. *Let $k \in \mathbb{N}$ be fixed, and let $\epsilon > 0$ be sufficiently small in terms of k and b. Then there exists $\delta > 0$, $\delta = \delta(\epsilon, k)$ such that the number of integers in $[1, m]$ that are not (ϵ, k)-normal is $O(m^{1-\delta})$. Also there eixsts $\delta' = \delta'(\epsilon, k) > 0$ such that the number of base-b strings of length ℓ (including those starting with 0) that are not not (ϵ, k)-normal is $O\left(b^{\ell(1-\delta')}\right)$.*

Proof. The first statement follows from the second (exercise). For the second half, assume $k = 1$ and $\epsilon > 0$ is sufficiently small. Let $d \in \{0, 1, \ldots, b-1\}$ be a digit, and consider how many strings of length ℓ have *exactly* m appearances of d.

There are $\binom{\ell}{m}$ ways to pick the positions for the d. Then there are $(b-1)^{\ell-m}$ possibilities for the remaining digits.

So there are $\binom{\ell}{m}(b-1)^{\ell-m}$ possible strings containing exactly m appearances of d.

So the number of $(\epsilon, 1)$-nonnormal strings of length ℓ must be bounded by

$$b\left(\sum_{m<\ell\left(\frac{1}{b}-\epsilon\right)}\binom{\ell}{m}(b-1)^{\ell-m} + \sum_{m>\ell\left(\frac{1}{b}+\epsilon\right)}\binom{\ell}{m}(b-1)^{\ell-m}\right)$$

We will show how to get the desired bound on the first sum, and the second sum will be similarly bounded.

Consider
$$\binom{\ell}{m}(b-1)^{\ell-m} = \frac{\ell!}{m!(\ell-m)!}(b-1)^{\ell-m}$$

In going from the mth term to the $(m+1)$th term, we multiply by
$$\frac{\ell-m}{(m+1)(b-1)}.$$

Thus the terms are increasing provided
$$\frac{\ell-m}{(m+1)(b-1)} > 1,$$

or
$$\begin{aligned}
\ell - m &> (m+1)(b-1) \\
\ell - m &> m(b-1) + (b-1) \\
\ell &> bm + (b-1) \\
\frac{\ell-(b-1)}{b} &> m \\
\ell\left(\frac{1}{b} - \frac{b-1}{b\ell}\right) &> m
\end{aligned}$$

Providided ℓ is sufficiently large, then this is true for all terms in our sum.

Thus we can bound all the terms in our sum by the last term, since all of the terms are increasing. This gives us
$$b \sum_{m<\ell\left(\frac{1}{b}-\epsilon\right)} \binom{\ell}{m}(b-1)^{\ell-m} \le b\ell^3\left(\frac{1}{b}-\epsilon\right)\cdot\binom{\ell}{\ell\left(\frac{1}{b}-\epsilon\right)}(b-1)^{\ell-\ell\left(\frac{1}{b}-\epsilon\right)}$$

The ℓ^3 comes from replacing $\ell\left(\frac{1}{b}-\epsilon\right)$ inside the binomial.
$$\le b\ell^3\binom{\ell}{\left(\frac{1}{b}-\epsilon\right)\ell}(b-1)^{\ell\left(\frac{b-1}{b}+\epsilon\right)}$$
$$= O\left(\ell^{5/2}\left[\left(\frac{b}{1-b\epsilon}\right)^{\frac{1}{b}-\epsilon}\left(\frac{b}{b-1+b\epsilon}\right)^{\frac{b-1}{b}+\epsilon}(b-1)^{\frac{b-1}{b}+\epsilon}\right]^{\ell}\right)$$

Doing some algebra, we get
$$= O\left(\ell^{5/2}b^{\ell}\left[(1-b\epsilon)^{-\left(\frac{1}{b}-\epsilon\right)}\left(1+\frac{b}{b-1}\epsilon\right)^{-\left(\frac{b-1}{b}+\epsilon\right)}\right]^{\ell}\right)$$

I claim it suffices to show what's in brackets is $b^{-2\delta'}$, because
$$\ell^{5/2} = O(b^{\delta'\ell}),$$

which will give the first sum being bounded by $O\left(b^{\ell(1-\delta')}\right)$. So consider
$$(1-b\epsilon)^{-\left(\frac{1}{b}-\epsilon\right)}\left(1+\frac{b}{b-1}\epsilon\right)^{-\left(\frac{b-1}{b}+\epsilon\right)}$$

$$= \exp\left(-\left(\frac{1}{b}-\epsilon\right)\log(1-b\epsilon)-\left(\frac{b-1}{b}+\epsilon\right)\log\left(1+\frac{b}{b-1}\epsilon\right)\right)$$

Then we can apply the Taylor series for $\log(1+x)$, which gives $\log(1+x) = x + \frac{x^2}{2} + \frac{x^3}{3} = x + O(x^2)$. Applying it to our case, we get

$$\exp\left(-\frac{b^2}{2(b-1)}\epsilon^2 + O(\epsilon^3)\right)$$

So if ϵ is small enough, this is less than 1, so we can find δ'.

This proves the $k = 1$ case.

For the general case, we have a trick! (ϵ, k)-normality for base-b acts like $(\epsilon, 1)$-normality for base-b^k.

For example, consider 1205287, for $(\epsilon, 3)$-normality, then we would really look at $(120)(528)*$ and $*(205)(287)$, and $**(052)**$. There are three shifts we need to care about.

If a string of base-b digits is *not* (ϵ, k)-normal, then at least one of these base b^k "shifts" is not $(\epsilon, 1)$-normal. (Perhaps $(\epsilon/k$,1)-normal? Exercise.)

So the number of non-(ϵ, k)-normal base-b strings of length ℓ is $\leq$ the number of non-$(\epsilon, 1)$-normal base b^k strings of length $\lfloor \frac{\ell}{k} \rfloor$ or $\lfloor \frac{\ell}{k} \rfloor - 1$ times k.

This equals

$$kO\left(\left(b^k\right)^{\left\lfloor \frac{\ell}{k} \right\rfloor (1-\delta')} + \left(b^k\right)^{\left\lfloor \frac{\ell}{k} - 1 \right\rfloor (1-\delta')}\right) = O\left(b^{\ell(1-\delta')}\right).$$

■

1 The techniques in the proof

Remember the sum

$$\sum \binom{\ell}{m}(b-1)^{\ell-m}?$$

Often functions increase for a while, reach a maximum, and then decrease. It's often easy to bound the sums on the tails.

Sometimes you can make it look like

$$e^{-\lambda n^2},$$

Then you can say that the sum is like the integral, and that this integral is easy. (This is one place we get $\sqrt{\pi}$s popping out).

Another useful trick is

$$(1+a)^b = e^{b\log(1+a)} \approx e^{ab}$$

if a is small.

Normal number construction theorem. We want to show $0.f(1)f(2)f(3)\cdots$ is base-b normal. We must show that for any string, s,

$$\lim_{N\to\infty}\frac{\nu_s(x,N)}{N} = \frac{1}{b^{|s|}},$$

where $\nu_s(x, N)$ counts the number of appearances of s in the first N digits of x.

For a given N, let $m(N)$ be such that the Nth digit of x lies in $f(m)$. In particular,

$$\sum_{n=1}^{m-1}\nu(f(n)) < N \leq \sum_{n=1}^{m}\nu(f(n)).$$

By our assumptions

$$\nu(f(m)) = o\left(\sum_{n=1}^{m} \nu(f(n))\right).$$

So therefore,

$$N \sim \sum_{n=1}^{m} \nu(f(n)).$$

Also $m = o(N)$, because of our assumption that the lengths of $\nu(f(n))$ are mostly large. And

$$\nu(f(m)) = o(N).$$

Therefore

$$\nu_s(x, N) = \nu_s(f(1)f(2)\cdots f(m)) + O(\nu(f(m))) = \nu_s(f(1)f(2)\cdots f(m)) + o(N).$$

■