Appendix to Notes 6

Chebychev's Theorem

Theorem 1 Chebychev's Theorem

Let f be a non-negative $\mathcal{F}$-measurable function. Then, for $c > 0$ we have

$$\mu\{x : f(x) > c\} \leq \frac{1}{c} \int_X f d\mu.$$

Proof Let $C = \{x : f(x) > c\} \in \mathcal{F}$. Then

$$\int_X f d\mu \geq \int_C f d\mu > \int_C c d\mu = c\mu(C).$$

■

This simple result has many applications, especially in probability. Here we give an application to number theory, using the simple functions defined in Appendix 5.

Let $I = [0, 1]$. For $x \in I$ we can use the notation of appendix 5 with $\ell = 2$ to write

$$x = \sum_{i=1}^{\infty} \frac{a_i(x)}{2^i}.$$

Set $S_N(x) = \sum_{1 \leq i \leq N} a_i(x)$. The question is, what is the average value of $S_N(x)$?

Note that

$$\int_0^1 a_i d\mu = \frac{1}{2} \quad \text{for all } i$$

and so

$$\int_0^1 S_N d\mu = \frac{N}{2} \quad \text{for all } N \geq 1.$$

Theorem 2

For all $\varepsilon > 0$ we have that

$$\mu \left\{ x \in I : \left| \frac{S_N(x)}{N} - \frac{1}{2} \right| > \varepsilon \right\} \to 0$$

as $N \to \infty$.

This shows that the values of $S_N(x)/N$ are concentrated about $1/2$ as N tends to infinity. This is a special case of the *weak law of large numbers.*

Proof

Consider the Radamacher function defined as $R_k(x) = 2a_k(x) - 1$ so

$$R_k(x) = \begin{cases} 1 & \text{if } a_k(x) = 1, \\ -1 & \text{if } a_k(x) = 0. \end{cases}$$

These are simple functions though not non-negative. Nontheless they can be integrated since the values and domains are finite. These functions have orthogonality properties, the first of which is trivial.

$$\int_I R_k^2 d\mu = 1,$$

and

$$\int_I R_i R_j d\mu = 0, \text{ when } i \neq j.$$

For the verification of the second of these we write

$$I_k(n) = \left(\frac{n}{2^k}, \frac{n+1}{2^k} \right],$$

so that

$$R_k(x) = \begin{cases} 1 & \text{if } x \in I_k(n) \text{ with } n \text{ odd} \\ -1 & \text{if } x \in I_k(n) \text{ with } n \text{ even.} \end{cases}$$

Now, without loss of generality, assume that $i < j$ so that, in fact, $i \leq j-1$. Then given an $n \leq 2^{j-1} - 1$, we can find an $n' \leq 2^i - 1$ such that $I_{j-1}(n) \subseteq I_i(n')$. Yet R_i is constant on $I_i(n')$ and so it must be constant on $I_{j-1}(n)$. Decompose

$$\begin{aligned} I_{j-1}(n) &= \left(\frac{n}{2^{j-1}}, \frac{n+1}{2^{j-1}} \right] = \left(\frac{2n}{2^j}, \frac{2n+1}{2^j} \right] \cup \left(\frac{2n+1}{2^j}, \frac{2n+2}{2^j} \right] \\ &= I_j(2n) \cup I_j(2n+1). \end{aligned}$$

Here, $R_j(x) = -1$ on $I_j(2n)$ and $R_j(x) = 1$ on $I_j(2n+1)$. Setting $\overline{R}_{i,,j-1}$ to be the value of $R_i(x)$ on $I_{j-1}(n)$ we get

$$\begin{aligned}\int_{I_{j-1}(n)} R_i R_j d\mu &= \overline{R}_{i,,j-1}\int_{I_{j-1}(n)} R_j d\mu \\ &= \overline{R}_{i,,j-1}\left(-\mu(I_j(2n)) + \mu(I_j(2n+1))\right) \\ &= \overline{R}_{i,,j-1}\left(-\frac{1}{2^j}+\frac{1}{2^j}\right) \\ &= 0.\end{aligned}$$

Thus

$$\int_I R_i R_j d\mu = \sum_{1\le n\le 2^{j-1}-1}\int_{I_{j-1}(n)} R_i R_j d\mu = 0.$$

Let $T_N(x) = \sum_{1\le i\le N} R_i(x)$. Then

$$T_N(x) = \sum_{1\le i\le N}(2a_i(x)-1) = 2S_N(x) - N.$$

So

$$\frac{S_N(x)}{N} - \frac{1}{2} = \frac{T_N(x)}{2N}$$

and it suffices to show that

$$\mu\left\{x\in I : \left|\frac{T_N(x)}{2N}\right| > \varepsilon\right\} \to 0$$

or, on replacing 2ε by ε,

$$\mu\left\{x\in I : |T_N(x)| > \varepsilon N\right\} \to 0.$$

But $|T_N(x)| > \varepsilon N$ if, and only if, $T_N(x)^2 > \varepsilon^2 N^2$. So it suffices to show that

$$\mu\left\{x\in I : T_N(x)^2 > \varepsilon^2 N^2\right\} \to 0.$$

Yet Chebychev's Theorem implies that

$$
\begin{aligned}
& \mu\left\{x \in I : T_N(x)^2 > \varepsilon^2 N^2\right\} \\
\leq\ & \frac{1}{\varepsilon^2 N^2} \int_I T_N^2 d\mu \\
=\ & \frac{1}{\varepsilon^2 N^2} \int_I \left(\sum_{1\leq i\leq N} R_i \right)^2 d\mu \\
=\ & \frac{1}{\varepsilon^2 N^2} \left(\sum_{1\leq i\leq N} \int_I R_i^2 d\mu + \sum_{\substack{1\leq i\leq N \\ i\neq j}} \sum_{1\leq j\leq N} \int_I R_i R_j d\mu \right) \\
=\ & \frac{1}{\varepsilon^2 N}
\end{aligned}
$$

which tends to zero as N tends to infinity. ■

It can be shown that if we set

$$A = \left\{ x \in I : \lim_{N\to\infty} \frac{S_N(x)}{N} = \frac{1}{2} \right\}$$

then $\mu(A^c) = 0$. So, for almost all (μ) of the $x \in I$ we have that

$$\lim_{N\to\infty} \frac{S_N(x)}{N} = \frac{1}{2}. \tag{1}$$

This would be an example of the *strong law of large numbers.* It says that for almost all x the digits 0 and 1 occur equally frequently in the binary expansion. What happens for general bases ℓ?

Suppose that $x \in [0, 1]$ is expressed as a non-terminating expansion in the base of ℓ. Suppose further that the digit b occurs n_b times in the first n places. If

$$\frac{n_b}{n} \to \beta$$

when $n \to \infty$ we say that b has *frequency* β. It is not a priori obvious that this limit exists, yet it does. We say that x is *simply normal in the scale of* ℓ if

$$\frac{n_b}{n} \to \frac{1}{\ell}$$

for all ℓ possible digits b. So (1) says that almost all numbers are simply normal in base 2.

Example 1 When $\ell = 2$,

$$x = 0.0101010101010... = \frac{0}{2} + \frac{1}{2^2} + \frac{0}{2^3} + \frac{1}{2^4} + \frac{0}{2^5} + \frac{1}{2^6} + ...$$

is simply normal. Yet

$$\begin{aligned} x &= \frac{0}{2} + \frac{1}{2^2} + \frac{0}{2^3} + \frac{1}{2^4} + \frac{0}{2^5} + \frac{1}{2^6} + ... \\ &= \frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^6} + ... \\ &= \frac{1}{4} + \frac{1}{4^2} + \frac{1}{4^3} + ... \\ &= 0.1111... \end{aligned}$$

in base 4. So x is not simply normal in base 4.

Definition We say that x is *normal in base* ℓ if all of (the fractional parts of) $x, \ell x, \ell^2 x, \ell^3 x, ...$ are simply normal in all bases $\ell, \ell^2, \ell^3, ...$.

Theorem 3

Almost all numbers are normal in any scale.

Proof Not given. ■