Scale Invariant Means and the First Digit Problem

Peter Renaud

Department of Mathematics & Statistics
University of Canterbury,
Private Bag 4800
Christchurch, New Zealand

No. 166 October, 1998

Scale Invariant Means and the First Digit Problem

Peter Renaud

Department of Mathematics and Statistics
University of Canterbury
Private Bag 4800
Christchurch, New Zealand

ABSTRACT

It is a well established fact that in some circumstances, lists of what appear to be random numbers, show a striking non-uniform distribution of digits. In many instances, these numbers arise relative to a system of units. In such cases there is an underlying assumption of scale invariance, by which is meant that the choice of units may well be arbitrary. In this paper we consider the general problem of scale invariance from the point of view of certain means on suitable function spaces. This is then applied to give a simple explanation for the distribution of first significant digits.

1. Introduction

It has long been realized that lists of numerical data, which could include physical constants or statistical data in a very general sense, exhibit the surprising property that the distribution of the first significant digit is not uniform over the set $\{1, 2, \ldots, 9\}$. In fact, a much better approximation is given by the probabilities $p_k = \log_{10}\left(1 + \frac{1}{k}\right)$ for $k = 1, \ldots, 9$.

Many attempts have been made to resolve what some see as a paradox (see e.g. Scozzafavva [4] and the references therein for a more detailed history and analysis of this problem). The more successful of these were based around two ideas. Firstly that a *countably additive* probability measure was a root cause of this "paradox" and secondly, that whatever probability idea was used, it should reflect the fact that many numbers, such as physical or chemical constants, are given relative to a system of units and to a large extent these systems are arbitrary. (As an example, the speed of light might be one of the numerical constants but could be expressed in miles per hour or metres per second. Obviously the first significant digit of that constant depends on the system of units chosen.) This second notion leads to the idea of invariance under rescaling.

The idea that countable additivity can be dispensed with, is not really so alarming. Number theorists are well used to using probabilities on the integers, which probabilities are only finitely additive and which nonetheless allow them to give perfectly rigorous interpretations of propositions such as "the probability that two randomly chosen positive integers are co-prime is $\frac{6}{\pi^2}$." It is an inescapable fact that, to require a probability measure on the reals (or integers) to also have certain invariance properties (such as shift or scale invariance), then countable additivity is the price which it is necessary to pay.

In this paper we take the view that scaling is most easily described in terms of a group action on the set of reals. This leads to the problem of finding finitely additive set functions (called means rather than probabilities because they will not be assumed countably additive) which are invariant under the group action. It turns out that even showing that such means exist is not trivial. Fortunately, we can appeal to results in amenable group theory to quickly obtain the required existence results.

Such means will assign to every bounded, measurable function, an "average" value. In particular they will give a density to measurable subsets of the reals. The main aim of this paper is to show that the set of reals with first significant digit k, has density $p_k = \log_{10}\left(1 + \frac{1}{k}\right)$ *no matter which scale invariant mean we use.* We contend that this result is necessary if we are to assign any meaning to scale invariance. Without it, the density of this set would depend on which mean is chosen and there seems no clear reason to choose one rather than another.

2. *Existence of scale invariant means on* $\mathbf{R}$

As mentioned above, scaling is best thought of in terms of a group action. Here, the underlying set is $\mathbf{R}$, the set of reals, taken with the usual Lebesgue measure dx. Scaling is achieved by multiplication with positive reals, so that we also need to consider the group $\mathbf{R}^+$ under multiplication. On this set therefore we need to take not the usual Lebesgue measure, but the group Haar measure $\frac{dt}{t}$. Note therefore that if $\phi \in L_1(\mathbf{R})$ then $\|\phi\|_1 = \int\limits_{-\infty}^{\infty} |\phi_n(x)|\, dx$, but for $\psi \in L_1(\mathbf{R}^+)$ we must define $\|\psi\|_1 = \int\limits_{0}^{\infty} |\psi(t)| \frac{dt}{t}$. The context will always make it clear which case we are considering.

Although much of what follows applies to integers rather than the reals, it seems profitable to deal with the more general real case. Wherever possible we emphasize the group action of $\mathbf{R}^+$ on $\mathbf{R}$ in the belief that much of this paper could be generalized to the more general setting that this implies.

DEFINITION 2.1. A *mean* m on $L_\infty(\mathbf{R})$ is a bounded linear functional satisfying

(i) $\langle m, f \rangle \geq 0$ if $f \geq 0$ and
(ii) $\langle m, 1 \rangle = 1$.

Clearly a mean defines a "finitely additive" probability measure, i.e. a *density*, on Borel subsets of $\mathbf{R}$.

The group $G = (\mathbf{R}^+, \cdot)$ acts on $\mathbf{R}$ via $a : x \to ax$.
This group action extends to an isometric action on $L_\infty(\mathbf{R})$ via $a : f \to f_a$ where $(f_a)(x) = f(ax)$ and also on $L_1(\mathbf{R})$, $a : \phi \to \phi_a$ where now $\phi_a(x) = a\phi(ax)$.

Using the same notation in both cases should not cause any confusion as it will always be clear which space is being acted upon. While the $L_\infty(\mathbf{R})$ definition is natural, the different $L_1(\mathbf{R})$ definition is forced on us if we want an isometry as well as the following duality result which we would expect

$$\langle \phi, f_a \rangle = \langle \phi_{a^{-1}}, f \rangle \quad \phi \in L_1(\mathbf{R}), f \in L_\infty(\mathbf{R}). \tag{2.1}$$

For the same reason, if $\psi \in L_1(\mathbf{R}^+)$ then we have to define ψ_a by $\psi_a(t) = \psi(at)$.

This action of $\mathbf{R}^+$ on $L_1(\mathbf{R})$ and $L_\infty(\mathbf{R})$ extends to an action of $L_1(\mathbf{R}^+)$ on these two spaces as follows:

DEFINITION 2.2.
(i). For $\psi \in L_1(\mathbf{R}^+), \phi \in L_1(\mathbf{R})$ define the convolution type operation $T_\psi : L_1(\mathbf{R}) \to L_1(\mathbf{R})$ by

$$(T_\psi \phi)(x) = (\psi \circ \phi)(x) = \int_0^\infty t\psi(t)\phi(tx)\frac{dt}{t}. \tag{2.2}$$

It is easy to check that $\psi \circ \phi \in L_1(\mathbf{R})$ with $\|\psi \circ \phi\|_1 \leq \|\psi\|_1 \|\phi\|_1$.

(ii). For $\psi \in L_1(\mathbf{R}^+)$, the dual map $T_\psi^* : L_\infty(\mathbf{R}) \to L_\infty(\mathbf{R})$ is then given by

$$(T_\psi^* f)(x) = \int_0^\infty \psi\left(\frac{1}{t}\right) f(tx)\frac{dt}{t}.$$

or (using the definition of convolution above)

$$T_\psi^* f = \tilde{\psi} \circ f \tag{2.3}$$

where $\tilde{\psi}(t) = \frac{1}{t}\psi\left(\frac{1}{t}\right)$.

DEFINITION 2.3. A *weight* in $L_1(\mathbf{R}^+)$ is a non-negative function ψ such that $\|\psi\|_1 = 1$. We denote by $W(R^+)$, the set of all weights in $L_1(\mathbf{R}^+)$.

DEFINITION 2.4. A mean m on $L_\infty(\mathbf{R})$ is *(topologically) scale invariant* if $\left\langle m, \tilde{\psi} \circ f \right\rangle = \langle m, f\rangle$ for all $f \in L_\infty(\mathbf{R})$ and weights $\psi \in W(\mathbf{R}^+)$.

The connection between this definition and what we would usually consider as scale invariance is provided by the following lemma.

LEMMA 2.5. *If m is scale invariant on $L_\infty(\mathbf{R})$ then*

$$\langle m, f_a\rangle = \langle m, f\rangle$$

for all $f \in L_\infty(\mathbf{R}), a \in \mathbf{R}^+$.

Proof. As is easily verified,

$$\tilde{\psi} \circ (f_a) = \left(\tilde{\psi}_{a^{-1}}\right) \circ f.$$

if $\psi \in W(\mathbf{R}^+)$.

Then if m is a scale invariant mean on $L_\infty(\mathbf{R})$ and $f \in L_\infty(\mathbf{R})$ we have

$$\langle m, f_a\rangle = \left\langle m, \tilde{\psi} \circ f_a \right\rangle = \left\langle m, \tilde{\psi}_{a^{-1}} \circ f \right\rangle = \langle m, f\rangle.$$

REMARK. That scale invariant means exist on $L_\infty(\mathbf{R})$ is not trivial. We show existence in two stages. Firstly we adapt a general result in amenable group theory (see e.g. Greenleaf [1], §1.3) to obtain a mean which is invariant under the group action. Secondly we use a technique of Namioka [2] which allows us to extend the construction to obtain scale invariant means, i.e. means invariant under the convolution product.

Let G be the "$ax + b$" group, i.e. defined on $\mathbf{R}^+ \times \mathbf{R}$ with group operation $(a, b)(c, d) = (ac, ad + b)$. Then G is a semidirect product of two Abelian groups and so ([1], theorem 1.2.6) there is a mean μ on $L_\infty(G)$ satisfying $\langle \mu, f_g\rangle = \langle \mu, f\rangle$ where $g \in G$.

Now if $f \in L_\infty(\mathbf{R})$, define $\widehat{f} \in L_\infty(G)$ by

$$\widehat{f}(t, x) = f(x), \ t \in \mathbf{R}^+, x \in \mathbf{R}.$$

Then if $a \in \mathbf{R}^+$ it is easily seen that

$$\widehat{(f_a)} = \left(\widehat{f}\right)_a.$$

Define ν on $L_\infty(\mathbf{R})$ by $\langle \nu, f\rangle = \left\langle \mu, \widehat{f}\right\rangle$. Clearly ν is a mean. Further

$$\langle \nu, f_a\rangle = \left\langle \mu, \widehat{(f_a)}\right\rangle = \left\langle \mu, \left(\widehat{f}\right)_a\right\rangle = \left\langle \mu, \widehat{f}\right\rangle = \langle \nu, f\rangle$$

and ν is invariant under the action of $\mathbf{R}^+$ on $L_\infty(\mathbf{R})$.

Fix $h \geq 0$ in $L_\infty(\mathbf{R})$. Define a linear functional λ on $L_1(\mathbf{R}^+)$ by

$$\langle \lambda, \psi\rangle = \left\langle \nu, \tilde{\psi} \circ h\right\rangle.$$

Clearly, λ is bounded and $\psi \geq 0 \implies \langle \lambda, \psi\rangle \geq 0$.

Furthermore because $\widetilde{\psi_t} \circ h = \left(\tilde{\psi} \circ h\right)_t$, it follows that

$$\langle \lambda, \psi_t\rangle = \langle \lambda, \psi\rangle.$$

By uniqueness of the Haar integral, there exists $m(h)$ such that

$$\left\langle \nu, \tilde{\psi} \circ h\right\rangle = m(h) \int_0^\infty \psi(t)\frac{dt}{t}. \tag{2.4}$$

Clearly $m(1) = 1$, $m(\alpha h) = \alpha m(h)$ if $\alpha \geq 0$ and $m(h_1 + h_2) = m(h_1) + m(h_2)$ for $h_1, h_2 \geq 0$.
So m extends to a linear functional (which we can again write as m) on $L_\infty(\mathbf{R})$ satisfying

$$\left\langle \nu, \tilde{\psi} \circ h\right\rangle = \langle m, h\rangle \int_0^\infty \psi(t)\frac{dt}{t} \text{ for all } \psi \in L_1(\mathbf{R}^+), h \in L_\infty(\mathbf{R}).$$

Then for a weight $\psi \in L_1(\mathbf{R}^+)$ and $f \in L_\infty(\mathbf{R})$,

$$\begin{aligned}
\left\langle m, \tilde{\psi} \circ f\right\rangle &= \left\langle m, \tilde{\psi} \circ f\right\rangle \int_0^\infty \psi(t)\frac{dt}{t} \\
&= \left\langle \nu, \tilde{\psi} \circ \left(\tilde{\psi} \circ f\right)\right\rangle \\
&= \left\langle \nu, \widetilde{(\psi * \psi)} \circ f\right\rangle \text{ where } (\psi * \psi)(t) = \int_0^\infty \psi(s)\psi\left(\frac{t}{s}\right)\frac{ds}{s} \in W(\mathbf{R}^+) \\
&= \langle m, f\rangle \int_0^\infty (\psi * \psi)(t)\frac{dt}{t} = \langle m, f\rangle.
\end{aligned}$$

So m is the required scale invariant mean.

3. *Scale invariant functions*

We now consider those functions which have the same "average value" or mean, independent of the choice of scale invariant mean. In this section we give a characterisation along similar lines to Wong ([5], theorem 7.3) in the case of amenable groups.

DEFINITION 3.1. A function $f \in L_\infty(\mathbf{R})$ is called scale invariant with mean α if $\langle m, f\rangle = \alpha$ for all scale invariant means $m \in L^*_\infty(\mathbf{R})$.

It is useful to characterise such functions in a way which makes no reference to means. This leads to

THEOREM 3.2. *Let N be the closure in $L_\infty(\mathbf{R})$ of functions of the form*

$$\sum_{k=1}^{n} f_k - \tilde{\psi}_k \circ f_k \text{ where } f_k \in L_\infty(\mathbf{R}) \text{ and } \psi_k \in W(\mathbf{R}^+).$$

Then f is scale invariant with mean α iff $f - \alpha 1 \in N$.

Proof. Replacing f_k by $f_k - \alpha 1$, we may assume that $\alpha = 0$. Choose $f \in N$ and $\epsilon > 0$.

Choose $f_k \in L_\infty(\mathbf{R}), \psi_k \in W(\mathbf{R}^+)$ such that

$$\left\| f - \sum_{k=1}^{n} \left(f_k - \tilde{\psi}_k \circ f_k\right) \right\|_\infty < \epsilon.$$

Then if m is a scale invariant mean,

$$\left|\left\langle m, f - \sum_{k=1}^{n}\left(f_k - \tilde{\psi}_k \circ f_k\right)\right\rangle\right| = \left|\langle m, f\rangle - \sum_{k=1}^{n}\left(\langle m, f_k\rangle - \left\langle m, \tilde{\psi}_k \circ f_k\right\rangle\right)\right| = |\langle m, f\rangle| < \epsilon$$

and we deduce that $\langle m, f\rangle = 0$ i.e. that f is scale invariant with mean 0.

Conversely suppose that f is scale invariant with mean 0 but that $f \notin N$. Let M be the subspace generated by N and f. Then an element of M can be (uniquely) written as $g + \alpha f$ with $g \in N$ and we can define a linear functional $\overline{m}$ on M by $\langle \overline{m}, g + \alpha f\rangle = \alpha$. It is a standard result that $\overline{m}$ is bounded on M so that by the Hahn-Banach theorem, $\overline{m}$ can be extended to a bounded linear functional m on $L_\infty(\mathbf{R})$. Note that $\langle m, f\rangle = 1$.

Since for $f \in L_\infty(\mathbf{R})$ and $\psi \in W(\mathbf{R}^+), f - \tilde{\psi} \circ f \in N$ we have

$$\left\langle m, \tilde{\psi} \circ f\right\rangle = \langle m, f\rangle \text{ for all } f \in L_\infty(\mathbf{R}), \psi \in W(\mathbf{R}^+).$$

Now, for $f \geq 0$, define

$$\left\langle m^{+}, f\right\rangle = \sup_{0 \leq h \leq f} \langle m, h\rangle$$

and similarly

$$\left\langle m^{-}, f\right\rangle = -\inf_{0 \leq h \leq f} \langle m, h\rangle.$$

Then (Shaeffer [3], p 72, Corollary 1), $m^{\pm}$ extend to bounded, positive linear functionals on $L_{\infty}(\mathbf{R})$ with

$$m = m^{+} - m^{-}.$$

Further for $f \geq 0$,

$$\begin{aligned}
\left\langle m^{+}, f_a\right\rangle &= \sup_{0 \leq h \leq f_a} \langle m, h\rangle \\
&= \sup_{0 \leq h_{a^{-1}} \leq f} \langle m, h\rangle \\
&= \sup_{0 \leq h_{a^{-1}} \leq f} \langle m, h_{a^{-1}}\rangle \text{ by lemma 2.5} \\
&= \left\langle m^{+} f\right\rangle
\end{aligned}$$

with a similar result for m^{-}.

So we can write $m = \alpha_1 m_1 - \alpha_2 m_2$ where α_1, α_2 are positive and m_1, m_2 are means invariant under the action of $\mathbf{R}^{+}$.

Using the technique that led to equation 2.4, there exist *scale invariant* means n_1, n_2 such that

$$\left\langle m_i, \tilde{\psi} \circ f\right\rangle = \langle n_i, f\rangle \ i = 1, 2, \ f \in L_{\infty}(\mathbf{R}), \ \psi \in W\left(\mathbf{R}^{+}\right).$$

But then

$$\begin{aligned}
1 = \langle m, f\rangle &= \left\langle m, \tilde{\psi} \circ f\right\rangle \\
&= \alpha_1\left\langle m_1, \tilde{\psi} \circ f\right\rangle - \alpha_2\left\langle m_2, \tilde{\psi} \circ f\right\rangle \\
&= \alpha_1\langle n_1, f\rangle - \alpha_2\langle n_2, f\rangle = 0
\end{aligned}$$

(as f is scale invariant with mean 0). This is a contradiction.

4. *Weights converging to scale invariance*

Scale invariant means in $L_\infty^*(\mathbf{R})$ are highly nonconstructable and therefore difficult to deal with directly. We will find it much more convenient to look for sequences of weights in $L_1(\mathbf{R}^+)$ which can be used to describe scale invariance. It is then possible to characterize, in terms of such sequences, those functions in $L_\infty(\mathbf{R})$ which are scale invariant. In practice this characterisation is straightforward to apply and as an example, we will use it to illustrate the first digit problem.

DEFINITION 4.1. Let (ψ_n) be a sequence in $W(\mathbf{R}^+)$.
We say that (ψ_n) is *convergent to scale invariance* if

$$\lim_n \|\psi_n(t) - \chi * \psi_n(t)\|_1 = 0 \text{ for all weights } \psi \in W(\mathbf{R}^+).$$

(Here again the operator $*$ is the convolution operation in the group algebra $L_1(\mathbf{R}^+)$.)

EXAMPLE 4.2. Let

$$\psi_n(t) = \begin{cases} \frac{1}{2\ln n} & \text{on } \left[\frac{1}{n}, n\right] \\ 0 & \text{otherwise.} \end{cases}$$

Then (ψ_n) is convergent to scale invariance.

Proof. Clearly ψ_n is a weight in $W(\mathbf{R}^+)$.

Since functions in $L_1(\mathbf{R}^+)$ can be approximated in norm by functions with support inside a compact interval $[\epsilon, K]$ with $0 < \epsilon < K$, it suffices to show that $\lim_n \|\psi_n(t) - \chi * \psi_n(t)\|_1 = 0$ for weight functions $\chi \in W(\mathbf{R}^+)$ with support in $[\epsilon, K]$.

We have

$$\psi_n(t) - \chi * \psi_n(t) = \int_\epsilon^K \chi(s)\left(\psi_n(t) - \psi_n\left(\frac{t}{s}\right)\right)\frac{ds}{s}$$

so that

$$\|\psi_n(t) - \chi * \psi_n(t)\|_1 \le \int_\epsilon^K \chi(s)\left[\int_0^\infty \left|\psi_n(t) - \psi_n\left(\frac{t}{s}\right)\right|\frac{dt}{t}\right]\frac{ds}{s}$$

For $n > M = \max\left(\sqrt{\frac{1}{\epsilon}}, \sqrt{K}\right)$ it is readily seen (though tedious to calculate) that

$$\int_0^\infty \left|\psi_n(t) - \psi_n\left(\frac{t}{s}\right)\right| \frac{dt}{t} \leq \frac{|\ln s|}{\ln n}$$

for all $s \in [\epsilon, K]$. So

$$\begin{aligned} \|\psi_n(t) - \chi * \psi_n(t)\|_1 &\leq \int_\epsilon^K \chi(s) \frac{|\ln s|}{\ln n} \frac{ds}{s} \\ &\leq \frac{2 \ln M}{\ln n} \|\chi\|_1 \\ &\to 0. \end{aligned}$$

This sequence (ψ_n) will be particularly useful when we consider the first digit problem.

THEOREM 4.3. *Let (ψ_n) be a sequence of weights in $W(\mathbf{R}^+)$ which is convergent to scale invariance and let $f \in L_\infty$. Then*

$$f \text{ is scale invariant with mean } \alpha \iff \lim_n \left\|\tilde{\psi}_n \circ f - \alpha 1\right\|_\infty = 0$$

Proof. Replacing f by $f - \alpha 1$, we may assume that $\alpha = 0$.

Suppose first that f is scale invariant with mean 0. Fix $\epsilon > 0$ and by theorem 3.2, choose $g_1, g_2, \ldots, g_N \in L_\infty(\mathbf{R})$ and $\chi_1, \chi_2, \ldots, \chi_N \in W(\mathbf{R}^+)$ such that

$$\left\|f - \sum_{i=1}^N (g_i - \tilde{\chi}_i \circ g_i)\right\|_\infty < \frac{\epsilon}{2}.$$

Let $M = \max_{1 \leq i \leq N} \|g_i\|_\infty$ and choose K such that

$$\text{if } n > K \text{ then } \|\psi_n(t) - (\chi_i * \psi_n)(t)\|_1 < \frac{\epsilon}{2MN} \text{ for } i = 1 \ldots N.$$

Then for $n > K$,

$$\begin{aligned}
\left\|\tilde{\psi}_n \circ f\right\|_\infty &\le \left\|\tilde{\psi}_n \circ \left(f - \sum_{i=1}^{N} (g_i - \tilde{\chi}_i \circ g_i)\right)\right\|_\infty + \left\|\tilde{\psi}_n \circ \sum_{i=1}^{N} (g_i - \tilde{\chi}_i \circ g_i)\right\|_\infty \\
&\le \frac{\epsilon}{2} + \left|\sum_{i=1}^{N} \left(\int_0^\infty \psi_n(t) g_i\left(\frac{ax}{t}\right)\frac{dt}{t} - \int_0^\infty \psi_n(t) g_i\left(\frac{a_i ax}{t}\right)\right)\frac{dt}{t}\right| \\
&\le \frac{\epsilon}{2} + \sum_{i=1}^{N}\left|\int_0^\infty (\psi_n(t) - \psi_n(a_i t)) g_i\left(\frac{ax}{t}\right)\frac{dt}{t}\right| \\
&\le \frac{\epsilon}{2} + \sum_{i=1}^{N} \|g_i\|_\infty \int_0^\infty \left|(\psi_n(t) - \psi_n(a_i t))\frac{dt}{t}\right| \\
&\le \frac{\epsilon}{2} + M \sum_{i=1}^{N} \frac{\epsilon}{2MN} = \epsilon.
\end{aligned}$$

Conversely, suppose that $\lim_n \left\|\tilde{\psi}_n \circ f\right\|_\infty = 0$.
If m is a scale invariant mean, then

$$|\langle m, f\rangle| = \left|\left\langle m, \tilde{\psi}_n \circ f\right\rangle\right| \le \left\|\tilde{\psi}_n \circ f\right\|_\infty \to 0$$

and we deduce that $\langle m, f\rangle = 0$.

5. *The first digit problem*

We are now in the position of giving a concise statement of the first digit problem. Theorem 4.3 in particular allows us to derive the result very simply.

THEOREM 5.1. *Let A_k be the set of reals whose first significant decimal digit is k ($k = 1, 2, \ldots, 9$) and let f_k be the characteristic function of A_k.*

Then for every scale invariant mean m,

$$\langle m, f_k\rangle = \log_{10} \frac{k+1}{k}.$$

To make the proof more transparent, we require the foillowiong simple lemma.

LEMMA 5.2. *Let A be a set of reals, consisting of a (doubly infinite) sequence of non overlapping intervals of constant length l, with successive intervals separated by gaps of constant length d.*

(So, e.g. we can write $A = \bigcup_{n=-\infty}^{\infty} [(l+d)n+b, (l+d)n+b+d]$.*)*

Then if I is an interval of length L and k is the number of sub-intervals of A contained in I,

$$\frac{L-2l-d}{l+d} \le k \le \frac{L+d}{l+d}.$$

Proof. If there are k sub-intervals in I, there are also (at least) $k-1$ gaps in I. Considering the total length of these gives $kl+(k-1)d \le L$, i.e. $k \le \frac{L+d}{l+d}$.

Furthermore, adding an extra sub-interval and a gap at each end, gives a set which now contains I. Therefore $(k+2)l+(k+1)d \ge L$ i.e. $k \ge \frac{L-2l-d}{l+d}$.

Proof of theorem 5.1 Fix $k = 1, 2, \dots, 9$. Choose (ψ_n) as defined in example 4.2. An easy calculation gives

$$\tilde{\psi}_n \circ f_k(x) = \frac{1}{2\ln n} \int_{-\ln n}^{\ln n} f_k(xe^u)du. \tag{5.1}$$

Note that f_k is the characteristic function of the set

$$\bigcup_{j=-\infty}^{j=\infty} (-10^j(k+1), -10^j k] \cup \bigcup_{j=-\infty}^{j=\infty} [10^j k, 10^j(k+1))$$

so that f_k is an even function. So we may assume that $x > 0$.

Then

$$f_k(xe^u) = 1 \iff u \in [j\ln 10 - \ln x + \ln k, j\ln 10 - \ln x + \ln(k+1))$$

for some $j \in Z$.

These are intervals of length $\ln\left(\frac{k+1}{k}\right)$ and gaps $\ln\left(\frac{10k}{k+1}\right)$. So by lemma 5.2

$$\frac{2\ln n - 2\ln\left(\frac{k+1}{k}\right) - \ln\left(\frac{10k}{k+1}\right)}{\ln\left(\frac{k+1}{k}\right) + \ln\left(\frac{10k}{k+1}\right)} \ln\left(\frac{k+1}{k}\right) \le \int_{-\ln n}^{\ln n} f_k(xe^u)du \le \frac{2\ln n + \ln\left(\frac{10k}{k+1}\right)}{\ln\left(\frac{k+1}{k}\right) + \ln\left(\frac{10k}{k+1}\right)} \ln\left(\frac{k+1}{k}\right)$$

i.e

$$\left(2\ln n - \ln\left(\frac{10(k+1)}{k}\right)\right)\log_{10}\left(\frac{k+1}{k}\right) \le \int_{-\ln n}^{\ln n} f_k(xe^u)du \le \left(2\ln n + \ln\left(\frac{10k}{k+1}\right)\right)\log_{10}\left(\frac{k+1}{k}\right)$$

and a little calculation gives

$$-\frac{\ln\left(\frac{10(k+1)}{k}\right)}{2\ln n}\log_{10}\left(\frac{10k}{k+1}\right) \le \frac{1}{2\ln n}\int_{-\ln n}^{\ln n} f_k(xe^u)du - \log_{10}\left(\frac{k+1}{k}\right) \le \frac{\ln\left(\frac{10k}{k+1}\right)}{2\ln n}\log_{10}\left(\frac{10k}{k+1}\right)$$

so that

$$\left|\frac{1}{2\ln n}\int_{-\ln n}^{\ln n} f_k(xe^u)du - \log_{10}\left(\frac{k+1}{k}\right)\right| \le \frac{\ln\left(\frac{10(k+1)}{k}\right)}{2\ln n}\log_{10}\left(\frac{10k}{k+1}\right)$$

for all x.

By equation 5.1,

$$\left\|\tilde{\psi}_n \circ f_k - \log_{10}\left(\frac{k+1}{k}\right)\right\| \to 0$$

and the result now follows from theorem 4.3

References

1. F. P. Greenleaf, *Invariant means on topological groups*, Van Nostrand Mathematical Studies #16, Van Nostrand, 1969.

2. I. Namioka, On a recent result of H. Reiter, Proc. Amer. Math. Soc. 17 (1966), 1101–1102.

3. H. H. Schaefer, *Banach lattices and positive operators*, Springer-Verlag, Berlin-Heidelberg-New York, 1974.

4. R. Scozzafava, A survey of some common misunderstandings concerning the role and meaning of finitely additive probabilities in statistical inference, Statistica, 44 No. 1, 1984, 21–45.

5. J. C. S. Wong, Topologically stationary locally compact groups and amenability, Trans. Amer. Math. Soc. 144 (1969), 351–363.