Testing the finiteness of the support of a distribution: a statistical look at Tsirelson’s equation

Sylvain Delattre
LPMA, Université Paris Diderot
sylvain.delattre@univ-paris-diderot.fr

Mathieu Rosenbaum
LPMA, Université Pierre et Marie Curie
mathieu.rosenbaum@upmc.fr

25 February 2012

Abstract
We consider the following statistical problem: based on an i.i.d. sample of size $n$ of integer valued random variables with common law $\mu$, is it possible to test whether or not the support of $\mu$ is finite as $n$ goes to infinity? This question is in particular connected to a simple case of Tsirelson’s equation, for which it is natural to distinguish between two main configurations, the first one leading only to laws with finite support, and the second one including laws with infinite support. We show that it is in fact not possible to discriminate between the two situations, even using a very weak notion of statistical test.

1 Introduction
In this paper, we consider an i.i.d. sample of integer valued random variables with common law $\mu$ and address the following problem: is it possible to build an asymptotic test for the finiteness (or for the non finiteness) of the support of $\mu$ as the sample size goes to infinity? This question is in fact motivated by Tsirelson’s equation in discrete time, see [1, 2]. Let $(\nu_k)_{k \in \mathbb{N}}$ be a sequence of probability laws on the torus $T = \mathbb{R}/\mathbb{Z}$. A process $(\eta_k)_{k \in \mathbb{N}}$ taking value in $T$ is a weak solution of Tsirelson’s equation associated to $(\nu_k)_{k \in \mathbb{N}}$ if for all $k \in \mathbb{N}$, the random variable $\xi_k = \eta_k - \eta_{k-1}$ has law $\nu_k$ and is independent of $\mathcal{F}_{k-1}^\eta = \sigma(\eta_n, n \leq k-1)$. If moreover for all $k \in \mathbb{N}$, $\mathcal{F}_k = \mathcal{F}_k^\xi$, then the process $(\eta_k)_{k \in \mathbb{N}}$ is said to be a strong solution. It is proved by Yor in [6] that there is always existence of a weak solution to the equation: if $U$ and $\zeta_k$, $k \in \mathbb{N}$, are
independent random variables taking values in $\mathbb{T}$ such that $U$ is uniform and, for all $k$, $\zeta_k$ has distribution $\nu_k$, then the process $(\eta^*_k)_{k \in -\mathbb{N}}$ defined by
\[
\eta_0^* = U, \quad \eta_{-n}^* = U - \zeta_0 - \cdots - \zeta_{-n+1} \quad \text{for all } n \geq 1,
\] (1)
is a weak solution of Tsirelson’s equation associated to $(\nu_k)_{k \in -\mathbb{N}}$.

In this work we consider only the simple case where $\nu_k = \nu$ for all $k$. It can be shown that in terms of existence and uniqueness of strong or weak solutions to the equation, depending on the law $\nu$, three very different situations have to be considered, see [6] for details:

- **Case 1**: If $\nu$ is a Dirac measure: existence of a strong solution, no uniqueness in law.

- **Case 2**: If the support of $\nu$ contains at least two points and there exists an integer $p \geq 2$ and $x \in \mathbb{T}$ such that it is included in $x + \{k/p, k = 0, \ldots, p-1\}$: no strong solution, no uniqueness in law.

- **Case 3**: If $\nu$ is a law which does not satisfy the conditions of Case 1 or Case 2: no strong solution, uniqueness in law.

In [4, 5], Tsirelson’s equation is seen as a cosmological caricature: $\eta_{k-1}$ represents the state of the universe at time $k - 1$ and $\xi_k$ the action of the evolution process at time $k$. In particular, the authors raise the issue of the “initial magma” that is the description of the $\sigma$-algebra
\[
\mathcal{F}^-_{-\infty} = \bigcap_{n \geq 1} \sigma(\eta_k, k \leq -n).
\]
It turns out that it is necessarily a trivial $\sigma$-algebra in Case 3.

Our goal here is to give a statistical look at this cosmological caricature. More precisely, based on the historical observations of the state of the universe $\eta_0, \eta_{-1}, \ldots, \eta_{-n}$, can we (at least partially) decide whether we are in Case 1, Case 2 or Case 3? Of course we can already remark that under Case 1, the $\xi_k$ are constant and under Case 2 the $p_k = \xi_k - \xi_{k-1}$ are all rational. Therefore, if we observe two different values for the $\xi_k$ in the sample, we can conclude that we are not in Case 1 and if we observe one irrational $p_k$, we can discard Case 2. Consequently, if there are at least two different $\xi_k$ and one irrational $p_k$, we can conclude that we are in Case 3. However, in practice, we have only access to rounded values so it does not really make sense to build a procedure based on the fact that some quantity is irrational or not.

In order to rigorously answer the question, we use the notion of statistical test. We first consider an auxiliary problem, which is interesting on its own, namely the possibility of building a consistent test for the finiteness (or for the non-finiteness) of the support of an integer valued distribution $\mu$. Based on an i.i.d.
sample \((X_1, \ldots, X_n)\) with law \(\mu\). We prove that designing such a test procedure is in fact impossible, even using a very weak notion of consistent test. Then we show that a corollary of this result is the impossibility to build a consistent test in order to separate on the one hand the union of Case 1 and Case 2 and on the other hand Case 3.

The paper is organized as follows. The framework we use in order to test for the finiteness (or for the non finiteness) of the support of a distribution based on i.i.d. data is detailed in Section 2. In Section 3, we suggest an a priori natural functional in order to build such a test. The impossibility of building a consistent test is stated in Section 4 together with the connection with Tsirelson’s equation. The proofs are relegated to Section 5.

2 Testing framework

In this section we consider an i.i.d. sample \((X_1, \ldots, X_n)\) of integer valued random variables with common law \(\mu\). Our goal is to investigate the possibility of testing for the finiteness (or for the non finiteness) of the support of \(\mu\) as \(n\) goes to infinity. We refer to [3] for a detailed presentation of the notion of asymptotic statistical test. In the following, we write \(P_\mu\) for the probability measure on the canonical space \(\mathbb{N}^\mathbb{N}\) under which the canonical process \((X_i)_{i \geq 1}\) is a family of i.i.d. random variables with law \(\mu\). Moreover, we denote by \(\mathcal{P}_f\) the set of laws on \(\mathbb{N}\) with finite support and by \(\mathcal{P}_\infty\) the set of laws on \(\mathbb{N}\) with infinite support. To fix idea, let us consider in this section the null hypothesis of the finiteness of the support against the alternative of the non finiteness of it (the following definitions can be adapted in an obvious way if the null and alternative hypotheses are switched).

2.1 Uniform test

In order to discriminate between the two situations, the first idea is to try to build a consistent test with asymptotic uniform level \(\alpha\), with \(\alpha\) given in \([0, 1)\). This means designing a rejection area \(W_n\) in \(\mathbb{N}^n\) such that

\[
\limsup_n \sup_{\mu \in \mathcal{P}_f} \mathbb{P}_\mu [(X_1, \ldots, X_n) \in W_n] \leq \alpha \tag{2}
\]

and for all \(\mu \in \mathcal{P}_\infty\),

\[
\lim_n \mathbb{P}_\mu [(X_1, \ldots, X_n) \in W_n] = 1.
\]

This is clearly impossible. Indeed, for fixed \(n\), any law \(\mu_1^{\otimes n}\) with \(\mu_1 \in \mathcal{P}_f\) can be approached with arbitrary accuracy (in total variation norm for example) by a law \(\mu_2^{\otimes n}\) with \(\mu_2 \in \mathcal{P}_\infty\).
2.2 Pointwise test

We now consider a much weaker notion of test: we replace (2) by

\[ \sup_{\mu \in \mathcal{P}_f} \limsup_n \mathbb{P}(X_1, \ldots, X_n) \in W_n \leq \alpha. \]

The statistical meaning of this notion of test is of course very arguable but our point of view is the following: we want to show that it is impossible to build a consistent test with given level \( \alpha \) even in this very weak setting.

Remark 1. If we do not consider all distributions with finite support but only distributions with support in \([0, N]\) for some known \(N\), a test statistic can of course be easily designed. Indeed, it is then enough to consider the statistics \(\max(X_1, \ldots, X_n) \wedge (N + 1)\): when \(X_1\) follows \(\mu_1 \in \mathcal{P}_f\) it converges to the right endpoint of \(\mu_1\) which is smaller than \(N\), when \(X_1\) follows \(\mu_2 \in \mathcal{P}_\infty\) it converges to \(N + 1\).

3 Candidate test statistics for pointwise test

We want to investigate potential test statistics for discriminating between our two situations. Since the laws in \(\mathcal{P}_f\) and \(\mathcal{P}_\infty\) are different because they have different support, a natural idea is to design a test statistics based on the empirical support. For example, we can split the sample of size \(2n\), into two parts and compare the maximum over each subsample, that is we consider

\[ S_n = \max_{1 \leq i \leq n} X_i, \quad \tilde{S}_n = \max_{n+1 \leq i \leq 2n} X_i, \quad \text{and} \quad T_{2n} = 1\{S_n = \tilde{S}_n\}. \]

If \(\mu\) has finite support, it is clear that for \(n\) large enough, the maxima of the two subsamples coincide and in particular \(T_{2n}\) converges in probability to 1. In the infinite support case, one may expect an opposite behavior. In fact, we have the following result whose proof is given in Section 5.

Proposition 1. If \(\mu\) has infinite support, then \(T_{2n}\) does not converge in \(\mathbb{P}_\mu\)-probability to 1.

Therefore for any \(\mu_1\) with finite support and \(\mu_2\) with infinite support the statistics \(T_{2n}\) shows different asymptotic behaviors under \(\mu_1\) and \(\mu_2\). However, this is not enough to build a consistent test since we need to investigate the asymptotic probability that \(T_{2n}\) is different from 1. In fact, the first part of Theorem 1 in the next section implies that there exist a distribution \(\mu_2\) with infinite support and a subsequence of \(T_{2n}\) which converges to 1 in probability under \(\mathbb{P}_{\mu_2}\).

4 Impossibility of testing

The following theorem shows that it is not possible to build a test for the finiteness (or for the non finiteness) of the support. Indeed, it states that if one finds
a set in which any i.i.d. sample of size \(n\) from a law with finite support \(\mu_1\) is very unlikely to be, there is also one distribution \(\mu_2\) with infinite support so that an i.i.d. sample of it is very unlikely to be in this set, and conversely.

We write \(E_{\mu}\) for the expectation with respect to \(P_\mu\). We have the following result.

**Theorem 1.** Let \(\alpha \in [0,1)\), let \(\varphi : \mathbb{N}^* \to \mathbb{N}^*\) be an increasing map and let \(A_n : \mathbb{T}^{\varphi(n)} \to [0,1], n \geq 1\), be a sequence of measurable functions.

1. Assume that for any distribution \(\mu_1\) on \(\mathbb{N}\) with finite support,
   \[
   \limsup_n E_{\mu_1}[A_n(X_1,\ldots,X_{\varphi(n)})] \leq \alpha.
   \]
   Then, there exists a distribution \(\mu_2\) on \(\mathbb{N}\) with infinite support such that
   \[
   \liminf_n E_{\mu_2}[A_n(X_1,\ldots,X_{\varphi(n)})] \leq \alpha.
   \]

2. Assume that for any distribution \(\mu_2\) on \(\mathbb{N}\) with infinite support,
   \[
   \limsup_n E_{\mu_2}[A_n(X_1,\ldots,X_{\varphi(n)})] \leq \alpha.
   \]
   Then, for all \(\alpha' \in (\alpha,1)\) there exists a distribution \(\mu_1\) on \(\mathbb{N}\) with finite support such that
   \[
   \liminf_n E_{\mu_1}[A_n(X_1,\ldots,X_{\varphi(n)})] \leq \alpha'.
   \]

**Remark 2.** In Theorem 1, we do not only consider the usual testing framework where the \(A_n\) are indicators of a set and \(\varphi(n) = n\). Indeed, we allow for randomized test procedures and subsequences in the sample size. This is slightly more general and will be useful for the proof of Corollary 1.

We now come back to Tsirelson’s equation and state the result showing the impossibility of testing the hypothesis that \(\nu\) belongs to the union of Case 1 and Case 2 against the one that \(\nu\) belongs to Case 3, and conversely. Denote by \(P_\nu^*\) the law on \(\mathbb{T}^{-N}\) of the “uniform solution” \((\eta_k^*)_{k \in -N}\) given by (1) (when \(\nu_k = \nu\) for all \(k\)) and by \((\eta_k)_{k \in -N}\) the canonical process on \(\mathbb{T}^{-N}\). We have the following corollary.

**Corollary 1.** Let \(\alpha \in [0,1)\), let \(\varphi : \mathbb{N} \to \mathbb{N}\) be an increasing map and let \(B_n \subset \mathbb{T}^{\varphi(n)+1}\) be a sequence of measurable sets.

1. Assume that for any distribution \(\nu_1\) on \(\mathbb{T}\) belonging to Case 1 or Case 2 of Section 1,
   \[
   \limsup_n P_{\nu_1}^*[(\eta_0,\ldots,\eta_{-\varphi(n)}) \in B_n] \leq \alpha.
   \]
   Then, there exists a distribution \(\nu_2\) belonging to Case 3 such that
   \[
   \liminf_n P_{\nu_2}^*[(\eta_0,\ldots,\eta_{-\varphi(n)}) \in B_n] \leq \alpha.
   \]
2. Assume that for any distribution \( \nu_2 \) on \( \mathbb{T} \) belonging to Case 3 of Section 1,

\[
\limsup_n P^*_n[(\eta_0, \ldots, \eta_{-\varphi(n)}) \in B_n] \leq \alpha.
\]

Then, for all \( \alpha' \in (\alpha, 1) \) there exists a distribution \( \nu_1 \) belonging to Case 1 or Case 2 of Section 1 such that

\[
\liminf_n P^*_n[(\eta_0, \ldots, \eta_{-\varphi(n)}) \in B_n] \leq \alpha'.
\]

5 Proofs

5.1 Proof of Proposition 1

Let \( \mu \) be a probability measure on \( \mathbb{N} \) and assume that \( \mathbb{P}_\mu[T_{2n} = 1] \to 1 \). Since \( S_n \) and \( \tilde{S}_n \) are independent with the same law, it implies that there exists a sequence of integers \( k_n \) such that \( \mathbb{P}_\mu[S_n = k_n] \to 1 \).

Let us show that \( k_n \) is a bounded sequence, which will imply that the support of \( \mu \) is finite. If \( k_n \) is not bounded, then there exist a subsequence \( n_\ell \) such that \( k_{1 + n_\ell} > k_{n_\ell} \) for all \( \ell \) and \( k_{n_\ell} \to \infty \). Therefore

\[
\mathbb{P}_\mu[S_{n_\ell} = k_{n_\ell}] \to 1
\]

and

\[
\mathbb{P}_\mu[S_{1 + n_\ell} = k_{n_\ell}] \to 0
\]

because \( \mathbb{P}(S_{1 + n_\ell} = k_{1 + n_\ell}) \to 1 \) and \( k_{1 + n_\ell} \neq k_{n_\ell} \). Moreover, we have that

\[
\mathbb{P}_\mu[S_{1 + n_\ell} = k_{n_\ell}] - \mathbb{P}_\mu[S_{n_\ell} = k_{n_\ell}]
\]

is equal to

\[
\mathbb{P}_\mu[S_{n_\ell} = k_{n_\ell}] \mu([0, k_{n_\ell}]) + \mathbb{P}_\mu[S_{n_\ell} < k_{n_\ell}] \mu(k_{n_\ell}) - \mathbb{P}_\mu[S_{n_\ell} = k_{n_\ell}].
\]

This absolute value of this last quantity is smaller than \( \mu([k_{n_\ell}, +\infty)) \) which goes to zero as \( n \) goes to infinity. This shows the contradiction.

5.2 Proof of Theorem 1

5.2.1 Proof of Part 1 in Theorem 1

We recursively define a sequence of distributions \( \mu_2^n, n \geq 1 \), and a sequence of integers \( \psi(n), n \geq 0 \).

- At rank \( n = 1 \), we consider \( \mu_2^1 \) the Dirac measure at point 0, \( \psi(0) = 1 \) and \( \psi(1) = 1 \).

- At rank \( n > 1 \), we define \( \mu_2^n \) and \( \psi(n) \). The law \( \mu_2^n \) is the distribution with discrete support \( \{0, \ldots, n - 1\} \) defined by

\[
\mu_2^n(k) = \frac{c_n}{(\varphi \circ \psi(k))^2}, \quad k \in \{0, \ldots, n - 1\},
\]
with $c_n$ such that
\[ \sum_{k=0}^{n-1} \frac{c_n}{(\varphi \circ \psi(k))^2} = 1. \]
The integer $\psi(n)$ is taken so that
\[ \psi(n) > \max(\psi(n-1), n^2) \]
and for all $m \geq \psi(n)$,
\[ \mathbb{E}_{\mu_2} [A_m(X_1, \ldots, X_{\varphi(m)})] \leq \alpha + \frac{1}{n}. \]
Note that finding such a $\psi(n)$ is always possible thanks to the assumption on the sequence $(A_m)_m$. In particular, the sequence $\psi(n)$ is increasing and satisfies
\[ \sum_{k=0}^{+\infty} \frac{1}{(\varphi \circ \psi(k))^2} < \infty. \]
Now define the distribution $\mu_2$ on $\mathbb{N}$ with infinite support by
\[ \mu_2(k) = \frac{c}{(\varphi \circ \psi(k))^2}, \quad k \in \mathbb{N}, \]
with $c$ such that
\[ \sum_{k=0}^{+\infty} \frac{c}{(\varphi \circ \psi(k))^2} = 1. \]
We have that
\[ \mathbb{E}_{\mu_2} [A_{\psi(n)}(X_1, \ldots, X_{\varphi\psi(n)})] \]
is smaller than
\[ \mathbb{E}_{\mu_2} [A_{\psi(n)}(X_1, \ldots, X_{\varphi\psi(n)}) | \bigcap_{i=1}^{\varphi\psi(n)} \{X_i \leq n - 1\}] + \mathbb{P}_{\mu_2} [\bigcup_{i=1}^{\varphi\psi(n)} \{X_i \geq n\}] \]
\[ = \mathbb{E}_{\mu_2} [A_{\psi(n)}(X_1, \ldots, X_{\varphi\psi(n)})] + \mathbb{P}_{\mu_2} [\bigcup_{i=1}^{\varphi\psi(n)} \{X_i \geq n\}] \]
\[ \leq \alpha + \frac{1}{n} + \varphi \circ \psi(n) \sum_{k=n}^{+\infty} \frac{1}{(\varphi \circ \psi(k))^2}. \]
Using the fact that $\psi(n)$ is increasing and the inequality $\psi(n) > n^2$, we obtain
\[ \alpha + \frac{1}{n} + \varphi \circ \psi(n) \sum_{k=n}^{+\infty} \frac{1}{(\varphi \circ \psi(k))^2} \leq \alpha + \frac{1}{n} + \sum_{k=n}^{+\infty} \frac{1}{\varphi \circ \psi(k)} \leq \alpha + \frac{1}{n} + \sum_{k=n}^{+\infty} \frac{1}{k^2}. \]
This quantity goes to $\alpha$ as $n$ goes to infinity, which gives the result.
5.2.2 Proof of Part 2 in Theorem 1

Let $\alpha' \in (\alpha, 1)$. Assume that for any distribution $\mu_1$ on $\mathbb{N}$ with finite support
\[
\liminf_n \mathbb{E}_{\mu_1} \left[ A_n(X_1, \ldots, X_{\varphi(n)}) \right] > \alpha'.
\]
This last inequality is equivalent to
\[
\limsup_n \mathbb{E}_{\mu_1} \left[ 1 - A_n(X_1, \ldots, X_{\varphi(n)}) \right] < 1 - \alpha'.
\]
According to point 1. of Theorem 1, there exists a distribution $\mu_2$ on $\mathbb{N}$ with infinite support such that
\[
\liminf_n \mathbb{E}_{\mu_2} \left[ 1 - A_n(X_1, \ldots, X_{\varphi(n)}) \right] \leq 1 - \alpha',
\]
that is
\[
\limsup_n \mathbb{E}_{\mu_2} \left[ A_n(X_1, \ldots, X_{\varphi(n)}) \right] \geq \alpha'
\]
which gives the contradiction since $\alpha' > \alpha$.

5.3 Proof of Corollary 1

Let $f$ be an injection from $\mathbb{N}$ into $\mathbb{T} \cap \mathbb{Q}/\mathbb{Z}$. If $\mu$ is a probability measure on $\mathbb{N}$, denote by $f(\mu) = \mu \circ f^{-1}$ the image of $\mu$ by $f$. On the one hand, one has:
- if $\mu$ has finite support, $f(\mu)$ belongs to Case 1 or Case 2.
- if $\mu$ has infinite support, $f(\mu)$ belongs to Case 3 since $f(\mu)$ has infinite support.

On the other hand, because of the definition of $\mathbb{P}_\nu^*$, one has
\[
\mathbb{P}_{f(\mu)}^* \left[ (\eta_0, \ldots, \eta_{\varphi(n)}) \in B_n \right] = \mathbb{E}_\mu \left[ A_n(X_1, \ldots, X_{\varphi(n)}) \right]
\]
where $A_n(X_1, \ldots, X_{\varphi(n)})$ is equal to
\[
\int_0^1 1_{B_n}(u, u-f(X_1), u-f(X_1)-f(X_2), \ldots, u-f(X_1)-f(X_2)-\cdots-f(X_{\varphi(n)})) \, du.
\]
Using these two facts, together with Theorem 1, the proof of Corollary 1 is easily completed.

Acknowledgments

We are very grateful to Kouji Yano and Marc Yor for introducing us to the statistical question linked to Tsirelson’s equation.
References

[1] Cirel’son, B.S., *An example of a stochastic differential equation that has no strong solution*, Teor. Verojatnost. i Primenen., 20(2):427-430, 1975.

[2] Tsirelson, B.S., *My drift, Citing works.*, http://www.math.tau.ac.il/~tsirel/Research/mydrift/citing.html

[3] Van der Vaart, A., *Asymptotic Statistics*, Cambridge University Press, 1998.

[4] Yano, K. and Yor, M., *Around Tsirelson’s equation, or: The evolution process may not explain everything*, Preprint, 2010.

[5] Yano, K. and Yor, M., *A cosmological caricature derived from Tsirelson type equations*, Preprint, 2011.

[6] Yor, M., *Tsirel’son’s equation in discrete time*, Probability Theory and Related Fields, 91(2):135-152, 1992.