Computability of Probability Distributions and Distribution Functions

Takakazu Mori *, Yoshiki Tsujii **, and Mariko Yasugi * * *

Kyoto Sangyo University

Abstract. We define the computability of probability distributions on the real line as well as that of distribution functions. Mutual relationships between the computability notion of a probability distribution and that of the corresponding distribution function are discussed. It is carried out through attempts to effectivize some classical fundamental theorems concerning probability distributions. We then define the effective convergence of probability distributions as an effectivization of the classical vague convergence. For distribution functions, computability and effective convergence are naturally defined as real functions. A weaker effective convergence is also defined as an effectivization of pointwise convergence.

1 Introduction

In this article, we investigate computability aspects of probability distributions on the real line $\mathbb{R}$ in relation to their distribution functions. We will proceed as follows.

In Section 2, we briefly review some elementary notions of computability on the real line and some fundamentals of the classical theory of probability distributions on the real line.

In Section 3, we define the computability of probability distributions as well as that of distribution functions. Our central interest is the relation between those two computabilities. Meanwhile, we prove that the "vague sequential computability" is equivalent to the "weak sequential computability" for probability distributions.

In Section 4, we consider mutual relationships between effective convergence of probability distributions and that of distribution functions. If we restrict ourselves to the case where a probability distribution has a bounded density function, then the corresponding distribution function becomes effectively uniformly continuous, and we can prove the equivalence of the two effective convergences.

* Faculty of Science, Kyoto Sangyo University: morita@cc.kyoto-su.ac.jp. This work has been supported in part by Research Grant from KSU(2008, 339).

** Faculty of Science, Kyoto Sangyo University: tsujiiy@cc.kyoto-su.ac.jp. This work has been supported in part by Research Grant from KSU (2008, 282).

* * * Kyoto Sangyo University: yasugi@cc.kyoto-su.ac.jp. This work has been supported in part by JSPS Grant-in-Aid No. 20540143.

Andrej Bauer, Peter Hertling, Ker-I Ko (Eds.)
6th Int'l Conf. on Computability and Complexity in Analysis, 2009, pp. 185-196
http://drops.dagstuhl.de/opus/volltexte/2009/2270

In the general case, we need to define notions of computability and effective convergence for bounded monotonically increasing right continuous functions. Such a function may be discontinuous at most countably many points.

Computability of the probability distribution has been treated by many authors. For example, Weihrauch ([10]) and Schröder and Simpson ([9]) have treated computability of probability distributions on the unit interval from the stand point of the representation theory. We develop a theory along the Pour-El and Richards line.

2 Preliminaries

Here, we briefly review the introductory part of the computability theory on the real line developed by Pour-El and Richards [6] as well as some basics of probability distributions on the real line. A sequence of rational numbers $\{r_n\}$ is said to be *recursive* if there exist recursive functions α, β and γ such that $r_n = (-1)^{\gamma(n)}\frac{\beta(n)}{\alpha(n)}$. A sequence of real numbers $\{x_{m,n}\}$ is said to *converge effectively* to $\{x_m\}$ if there exists a recursive function $\alpha(m,k)$ such that $n \geqslant \alpha(m.k)$ implies $|x_{m,n} - x_m| < 2^{-k}$. A sequence of real numbers $\{x_m\}$ is said to be *computable* if there exists a recursive double sequence of rational numbers, which converges effectively to $\{x_m\}$.

We adopt the definition of computability of continuous real functions by Pour-El and Richards in Chapter 0 of [6].

A sequence of (real) functions $\{f_m\}$ is said to be *computable*, if it is (i) *sequentially computable*, that is, $\{f_m(x_n)\}$ is computable for all computable sequences of reals $\{x_n\}$, and (ii) *effectively continuous*, that is, there exists a recursive function $\alpha(m,n,k)$ such that $x, y \in [-n, n]$ and $|x - y| < 2^{-\alpha(m,n,k)}$ imply $|f_m(x) - f_m(y)| < 2^{-k}$. $\alpha(m,n,k)$ is called an effective modulus of continuity of $\{f_m\}$.

A sequence of (real) functions $\{f_m\}$ is said to be *uniformly computable*, if it is (i) sequentially computable and (ii) *effectively uniformly continuous*, that is, there exists a recursive function $\alpha(m,k)$ such that $|x - y| < 2^{-\alpha(m,k)}$ implies $|f_m(x) - f_m(y)| < 2^{-k}$.

For a probability distribution μ on the real line $\mathbb{R}$, its distribution function F is defined by $F(x) = \mu((-\infty, x])$. Such a distribution function is characterized by the following three properties: (i) monotonically increasing; (ii) right continuous; (iii) $F(\infty) = \lim_{x\to\infty} F(x) = 1$ and $F(-\infty) = \lim_{x\to-\infty} F(x) = 0$. A distribution function may be discontinuous, but the set of discontinuous points is at most countable.

It is well known that the above correspondence between probability distributions and distribution functions is one to one and onto.

In the following we denote the integral with respect to a probability distribution μ, $\int_{\mathbb{R}} f(x)\mu(dx)$, with $\mu(f)$.

Let $\{\mu_n\}$ be a sequence of probability distributions on $\mathbb{R}$ and let μ be a probability distribution on $\mathbb{R}$ with corresponding distribution functions $\{F_n\}$

and F respectively. Convergence of $\{\mu_n\}$ to μ is defined to be the convergence of $\{\mu_n(f)\}$ to $\mu(f)$ for all continuous functions with compact support. This convergence is called *vague convergence* and is equivalent to each of the following convergences.

Weak convergence: $\{\mu_n(f)\}$ converges to $\mu(f)$ for all bounded continuous functions f on $\mathbb{R}$.

Convergence of distribution functions: $\{F_n(x)\}$ converges to $F(x)$ at every continuous point x of $F(x)$.

We refer the reader to [1], [3], [4] and [7] for details of fundamentals of probability theory.

Since we adopt the notion of computability of functions by Pour-El and Richards, we will plan to confine ourselves to continuous distribution functions. A sufficient condition for continuity of a distribution function is the following.

Definition 1. (Absolute continuity of probability distributions) *A probability distribution μ is said to be absolutely continuous if there exists a nonnegative integrable function $\xi(x)$ which satisfies that $\mu(A) = \int_A \xi(x)dx$ for all measurable set $A \subset \mathbb{R}$.*

The function ξ is called a density (function) of μ. We also say that the corresponding distribution function F has a density ξ.

Remark 1. If μ is absolutely continuous, then the corresponding distribution function F is continuous, and equalities $\mu([a,b]) = \mu((a,b]) = \mu([a,b)) = \mu((a,b))$ $= F(b) - F(a)$ hold.

3 Computability of probability distributions

In this section, we define the computability of probability distributions on $\mathbb{R}$ and discuss its relation to the computability of distribution function.

Let $\{f_n\}$ be a sequence of continuous functions with compact support. We say that $\{f_n\}$ is a computable sequence of functions with compact support if it is a computable sequence of functions in the sense of Pour-El and Richards and furthermore there exists a recursive function $K(n)$ such that $f_n(x) = 0$ if $|x| \geqslant K(n)$.

We obtain the following lemma.

Lemma 1. *A computable sequence of functions with compact support is uniformly computable.*

Proof Let $\{f_m\}$ be a computable sequence of functions with compact support with respect to recursive functions $\alpha(m,n,k)$ and $K(m)$.

Define $\beta(m,k) = \alpha(m, K(m), k+1)$ and assume that $|x-y| < 2^{-\beta(m,k)}$.

If both x and y are in $[-K(m), K(m)]$, then it holds that $|f_m(x) - f_m(y)| < 2^{-(k+1)}$; otherwise, one of them, say, x is in $[-K(m), K(m)]$ and the other, say, y is not in $[-K(m), K(m)]$. Then $y < -K(m) \leqslant x$ or $y > K(m) \geqslant x$ and

$|x + K(m)| < 2^{-\beta(m,k)}$ or $|x - K(m)| < 2^{-\beta(m,k)}$ accordingly. So, $|f_m(x) - f_m(y)| = |f_m(x) - f_m(\pm K(m))| < 2^{-k}$, since $f_m(y) = f_m(\pm K(m)) = 0$. This proves that $\{f_m\}$ is effectively uniformly continuous with respect to $\beta(m,k)$. □

Definition 2. (Computability of probability distributions) *We say that a sequence of probability distributions $\{\mu_m\}$ is computable if it satisfies the following* vague sequential computability: *$\{\mu_m(f_n)\}$ is computable for all computable sequence of functions with compact support $\{f_n\}$.*

Remark 2. If we regard the integral $\mu(f)$ as a function on the set of all bounded continuous functions $\mathcal{C}_b(\mathbb{R})$ with sup-norm $|| \; ||$, Definition 2 only asserts sequential computability. For a probability distribution μ, it holds that $|\mu(f) - \mu(g)| \leqslant \mu(|f - g|) \leqslant ||f - g||$. This makes $\mu(f)$ effectively uniformly continuous as a function on $\mathcal{C}_b(\mathbb{R})$.

Let a and b with $a < b$ be computable numbers. For a computable function f on the interval $[a, b]$, its definite integral $\int_a^b f(x)dx$ is a computable number (cf. [6]). We can generalize this fact as follows.

Proposition 1. *Let $\{a_m\}$ and $\{b_m\}$ be computable sequences of reals with $a_m < b_m$ for each m, and let $\{f_n\}$ be a computable sequence of functions on $\mathbb{R}$. Then, $\{\int_{a_m}^{b_m} f_n(x)dx\}$ is a computable (double) sequence of real numbers.*

This proposition yields that, if a sequence of distributions has a computable sequence of density functions, then it is computable and the corresponding sequence of distribution functions is also computable.

We frequently use the following Lemma.

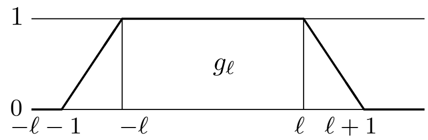
Lemma 2. (Monotone convergence [6]) *Let $\{x_{n,k}\}$ be a computable sequence of reals which converges monotonically to $\{x_n\}$ as k tends to infinity for each n. Then $\{x_n\}$ is computable if and only if the convergence is effective.*

We say that a sequence of functions $\{f_n\}$ is effectively bounded if there exists a recursive function $B(n)$ such that $|f_n(x)| \leqslant 2^{B(n)}$ for each $n, x \in \mathbb{R}$. We give some examples of probability distributions which have bounded density (Example 1).

Proposition 2. *If $\{\mu_m\}$ is vaguely sequentially computable, then it is* weakly sequentially computable*, that is, $\{\mu_m(f_n)\}$ is a computable sequence for all effectively bounded computable sequence of functions $\{f_n\}$.*

Proof. Let $\{f_n\}$ be an effectively bounded computable sequence of functions with an effective bound $B(n)$, and define $g_\ell(x)$ by:

$$g_\ell(x) = \begin{cases} 0 & \text{if } \; x \leqslant -\ell - 1 \\ (x+\ell)+1 & -\ell - 1 \leqslant x \leqslant -\ell \\ 1 & \text{if } \; -\ell \leqslant x \leqslant \ell \\ -(x-\ell)+1 & \text{if } \; \ell \leqslant x \leqslant \ell + 1 \\ 0 & \text{if } \; x \geqslant \ell + 1. \end{cases}$$



It is obvious that $\{g_\ell\}$ is a computable sequences of functions with compact support.

Since $g_\ell \uparrow 1$ pointwise, $\mu_m(g_\ell) \uparrow 1$ as ℓ tends to infinity by the bounded convergence theorem for each m, where $\uparrow$ means monotonically increasing convergence. Moreover, $\{\mu_m(g_\ell)\}$ is a computable sequence of reals by vague sequential computability of $\{\mu_m\}$ and the limit 1 is a computable number. So, the convergence of $\mu_m(g_\ell)$ to 1 is effective by Monotone Convergence Lemma 2. Therefore, we obtain a recursive function $N(m,k)$ such that $\mu_m([-\ell-1,\ell+1]^C) \leqslant 1-\mu_m(g_\ell) < 2^{-k}$ if $\ell \geqslant N(m,k)$, where A^C denotes the complement of the set A.

On the other hand, $\{\mu_m(f_n g_\ell)\}$ is a computable triple sequence of reals and

$$|\mu_m(f_n) - \mu_m(f_n g_{\ell+1})| = |\textstyle\int_{[-\ell-1,\ell+1]^C}(1-g_{\ell+1})\, f_n\, \mu_m(dx)|$$
$$\leqslant 2^{B(n)}\, \mu_m([-\ell-1,\ell+1]^C) < 2^{-k}$$

if $\ell \geqslant N(m, B(n)+k)$. This means that $\{\mu_m(f_n g_\ell)\}$ converges effectively to $\{\mu_m(f_n)\}$. Hence $\{\mu_m(f_n)\}$ is a computable sequence of reals. □

Proposition 3. *For a sequentially computable sequence of distribution functions $\{F_m\}$, effective continuity implies effective uniform continuity.*

Proof. By sequential computability of $\{F_m\}$, $\{F_m(n)\}$ and $\{F_m(-n)\}$ are computable sequences of reals. Since, F_m's are distribution functions, $F_m(n) \uparrow 1$ and $F_m(-n) \downarrow 0$ as n tends to infinity for each m. By Lemma 2, there exists a recursive function $N(m,k)$ such that $1 - F_m(x) \leqslant 1 - F_m(N(m,k)) < 2^{-k}$ for $x > N(m,k)$ and $F_m(x) \leqslant F_m(-N(m,k)) < 2^{-k}$ for $x < -N(m,k)$.

On the other hand, effective continuity of $\{F_m\}$ implies that there exists a recursive function $\alpha(m,n,k)$ such that $x, y \in [-n,n]$ and $|x-y| < 2^{-\alpha(m,n,k)}$ imply $|F_m(x) - F_m(y)| < 2^{-k}$.

If we put $\beta(m,k) = \alpha(m, N(m,k+2), k+2)$ and assume that $|x-y| < 2^{-\beta(m,k)}$, then the following four cases are possible.

The first case: Both x and y are in $[-N(m,k+2), N(m,k+2)]$. In this case, $|F_m(x) - F_m(y)| < 2^{-(k+2)}$.

The second case: Both x and y are in $(N(m,k+2),\infty)$. In this case, $|F_m(x) - F_m(y)| \leqslant |1-F_m(x)| + |1-F_m(y)| < 2^{-(k+1)}$.

The third case: Both x and y are in $(-\infty, -N(m,k+2))$. In this case, $|F_m(x) - F_m(y)| \leqslant |F_m(x)| + |F_m(y)| < 2^{-(k+1)}$.

The last case: One is in $[-N(k+2), N(k+2)]$ and the other is not. Suppose $x < -N(k+2) \leqslant y$, then

$$|F_m(x) - F_m(y)| \leqslant |F_m(x)| + |F_m(-N(k+2))| + |F_m(-N(k+2)) - F_m(y)|$$

$< 3 \cdot 2^{-(k+2)} < 2^{-k}$.

Therefore, we have shown that $\{F_m\}$ is effectively uniformly continuous with respect to $\beta(m,k)$. □

Theorem 1. If a sequence of distribution functions $\{F_m\}$ is sequentially computable, then the corresponding sequence of distributions $\{\mu_m\}$ is computable.

Proof. We prove that $\mu(f)$ is computable if f is a computable function with compact support. For such a function f, there exists an integer m such that $f(x) = 0$ if $|x| \geqslant m$.

Put

$$g_p(x) = f(-m + 2^{-p})\chi_{[-m,-m+2^{-p}]}(x) + \sum_{\ell=-m2^p+1}^{m2^p-1} f((\ell+1)2^{-p})\chi_{(\ell 2^{-p},(\ell+1)2^{-p}]}(x).$$

Then, $\mu(g_p) = \int_{[-m,m]} g_p \mu(dx)$

$$= \sum_{\ell=0}^{2m2^p-1} f(-m + (\ell+1)2^{-p})\big(F(-m + (\ell+1)2^{-p}) - F(-m + \ell 2^{-p})\big)$$

form a computable sequence of reals by sequential computability of F.

By Lemma 1, f is uniformly computable. So, there exists a recursive function $\alpha(k)$ such that $|f(x)-f(y)| < 2^{-k}$ if $|x-y| < 2^{-\alpha(k)}$. We note that $\{g_p\}$ converges effectively uniformly to f. More precisely, if $p \geqslant \alpha(k)$, then $||f - g_p|| \leqslant 2^{-k}$.

Therefore, for the above α, $p \geqslant \alpha(k+1)$ implies

$$|\mu(f) - \mu(g_p)| \leqslant ||f - g_p|| \leqslant 2^{-k}.$$

This proves the effective convergence of $\{\mu(g_p)\}$ to $\mu(f)$, and hence, $\mu(f)$ is computable. The proof goes through for a sequence $\{F_m\}$. □

If a probability distribution has a bounded density ξ with a bound M, then the corresponding distribution function F satisfies $|F(b)-F(a)| = |\int_a^b \xi(x)dx| \leqslant M|b-a|$. So, we obtain the following lemma.

Lemma 3. *If a sequence of densities of probability distributions is effectively bounded, then the corresponding sequence of distribution functions is effectively uniformly continuous.*

From the lemma above follows that, if a sequence of probability distributions has an effectively bounded sequence of densities, then uniform computability of the corresponding sequence of distribution functions is equivalent to sequential computability.

In the rest of this section, we assume the existence of bounded densities.

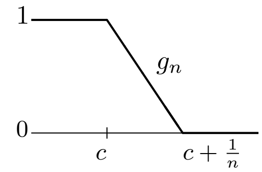
Proposition 4. *Let $\{\mu_m\}$ be a computable sequence of probability distributions which has effectively bounded densities. Then the corresponding sequence of distribution functions $\{F_m\}$ is sequentially computable.*

Proof. We prove that a single distribution function F is uniformly computable if the corresponding probability distribution μ is computable and there exists an integer M such that $|\xi(x)| \leqslant M$ for all x, where ξ is a density of μ.

By Lemma 3, F is effectively uniformly continuous.

We prove that $F(c)$ is computable if c is computable. First, we define the functions $\{g_n\}$ by

$$g_n(x) = \begin{cases} 1 & \text{if } \ x \leqslant c \\ -n(x-c)+1 & \text{if } \ c \leqslant x \leqslant c+\frac{1}{n} \\ 0 & \text{if } \ x \geqslant c+\frac{1}{n} \end{cases} .$$



Then, $\{g_n\}$ is a computable sequence and the following classical properties hold:

$\{g_n\}$ is monotonically decreasing, that is $m < n$ implies $g_m(x) \geqslant g_n(x)$ for all x.

$F(c) \leqslant \mu(g_n) \leqslant F(c + \frac{1}{n})$.

$F(c) = \lim_{n\to\infty} \mu(g_n)$ holds by the bounded convergence theorem.

On the other hand, $\{\mu(g_n)\}$ is a computable sequence of reals by the assumption and Proposition 2.

We obtain $0 \leqslant \mu(g_n) - F(c) = \int_c^{c+\frac{1}{n}} g_n(x)\xi(x)dx \leqslant \frac{M}{n}$.

Therefore, the convergence of $\mu(g_n)$ to $F(c)$ is effective, and hence $F(c)$ is computable.

This proof is also valid for a sequence $\{c_\ell\}$. The entire argument can be extended to a sequence $\{\mu_m\}$. □

We obtain the following theorem by Theorem 1, Lemma 3 and Proposition 4.

Theorem 2. *If a sequence of distributions* $\{\mu_n\}$ *has effectively bounded densities, then the computability of* $\{\mu_n\}$ *is equivalent to the uniform computability of the corresponding sequence of distribution functions.*

Example 1. In this example, μ denotes a probability distribution on $\mathbb{R}$, F denotes the corresponding distribution function and ξ denotes the corresponding density.

(1) Uniform distribution on $[0, 1]$:

$$\xi(x) = \chi_{[0,1]}(x); \quad F(x) = \begin{cases} 0 & \text{if } \ x \leqslant 0 \\ x & \text{if } \ 0 \leqslant x \leqslant 1 \\ 1 & \text{if } \ x \geqslant 1 \end{cases} .$$

$\xi(x)$ is bounded, but not continuous. On the other hand, $F(x)$ is continuous and indeed uniformly computable.

(2) Gaussian distribution: $\xi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$, $F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}y^2} dy$.

(3) Exponential distribution: $\xi(x) = e^{-x}$, $F(x) = 1 - e^{-x}$.

In (2) and (3), both ξ and F are computable. □

Example 2. (Translated Unit Distribution) The translated unit distribution δ_c is defined by

$$\delta_c(A) = \begin{cases} 1 & \text{if } \ c \in A \\ 0 & \text{otherwise} \end{cases} .$$

The corresponding distribution function is

$$F(x) = \chi_{[c,\infty)}(x) = \begin{cases} 0 \ (x < c) \\ 1 \ (x \geqslant c) \end{cases}.$$

The translated unit distribution is computable if c is a computable number. Its distribution function is not continuous. □

4 Convergence of probability distributions and distribution functions

We define effective convergence of probability distributions as an effectivization of classical vague convergence of probability distributions.

Definition 3. (Effective convergence of a sequence of probability distributions)
A sequence of probability distributions $\{\mu_m\}$ is said to effectively converge to a probability distribution μ if $\{\mu_m(f_n)\}$ converges effectively to $\{\mu(f_n)\}$ for any computable sequence of functions with compact support $\{f_n\}$.

It is well known that the set of all uniformly computable functions on a closed interval $[a, b]$ is dense in the set of all continuous functions on $[a, b]$ for any pair of computable numbers a and b with $a < b$. So, effective convergence of a sequence of probability distributions implies classical vague convergence.

The following proposition follows immediately.

Proposition 5. *If a computable sequence of probability distributions $\{\mu_n\}$ effectively converges to a probability distribution μ, then μ is computable.*

Proposition 6. *Let $\{\mu_m\}$ be a computable sequence of probability distributions and let μ be a computable probability distribution. If $\{\mu_m\}$ converges effectively to μ, then $\{\mu_m\}$ effectively weakly converges to μ, that is, $\{\mu_m(f_n)\}$ converges effectively to $\{\mu(f_n)\}$ for all effectively bounded computable sequence of functions $\{f_n\}$.*

Proof. We prove that $\{\mu_m(f)\}$ converges effectively to $\{\mu(f)\}$ for a bounded computable function f. For such f, there exists an integer M which satisfies that $|f(x)| \leqslant 2^M$ for all x.

Let us take a computable sequence of functions $\{g_\ell\}$ with compact support which is defined in the proof of Proposition 2. Then, we obtain a recursive function $N(k)$ which satisfies that $\mu([-\ell, \ell]^C) \leqslant 1 - \mu(g_{\ell-1}) < 2^{-k}$ for $\ell \geqslant N(k)$. Moreover, by effective convergence of $\{\mu_m(g_\ell)\}$ to $\{\mu(g_\ell)\}$, there exists a recursive function $\alpha(\ell, k)$ such that $m \geqslant \alpha(\ell, k)$ implies $|\mu_m(g_\ell) - \mu(g_\ell)| < 2^{-k}$.

Therefore, we obtain $1 - \mu(g_{N(k)}) < 2^{-k}$ and $m \geqslant \alpha(N(k), k)$ implies $|1 - \mu_m(g_{N(k)})| \leqslant |\mu_m(g_{N(k)}) - \mu(g_{N(k)})| + |1 - \mu(g_{N(k)})| < 2 \cdot 2^{-k}$.

On the other hand, since $\{fg_\ell\}$ is a computable sequence of functions with compact support, $\{\mu_m(fg_\ell)\}$ converges effectively to $\{\mu(fg_\ell)\}$. So, there exists a recursive function $\beta(\ell, k)$ such that $|\mu_m(fg_\ell) - \mu(f\,g_\ell)| < 2^{-k}$ for $m \geqslant \beta(\ell, k)$.

Therefore, $m \geqslant \beta(N(k), k)$ implies $|\mu_m(fg_{N(k)}) - \mu(fg_{N(k)})| < 2^{-k}$.

If we take $j = k + M + 2$ and assume $m \geqslant \max\{\alpha(N(j), j), \beta(N(j), j)\}$, then

$$\begin{aligned} &|\mu_m(f) - \mu(f)| \\ &\leqslant |\mu_m(f) - \mu_m(fg_{N(j)})| + |\mu_m(fg_{N(j)}) - \mu(fg_{N(j)})| \\ &\qquad + |\mu(f) - \mu(fg_{N(j)})| \\ &\leqslant 2^M(1 - \mu_m(g_{N(j)})) + |\mu_m(fg_{N(j)}) - \mu(fg_{N(j)})| + 2^M(1 - \mu(g_{N(j)})) \\ &< 2 \cdot 2^{-(k+2)} + 2^{-(k+2)} + 2^{-(k+2)} = 2^k. \end{aligned}$$

This proves the effective convergence of $\{\mu_n(f)\}$ to $\mu(f)$. □

Definition 4. (Effective pointwise convergence of functions)
A sequence of functions $\{F_m\}$ is said to converge effectively pointwise to a function F if $\{F_m(x_n)\}$ converges effectively to $\{F(x_n)\}$ for all computable sequence $\{x_n\}$.

By definition, the following proposition holds.

Proposition 7. *For a computable sequence of functions $\{F_m\}$, if it converge effectively pointwise to a function F, then F is sequentially computable.*

By Lemma 3, the existence of density of F implies the effective uniform continuity of F. So, we obtain the following proposition.

Proposition 8. *Let us consider a sequentially computable sequence of distribution functions $\{F_m\}$ and a distribution function F. If $\{F_m\}$ converges effectively pointwise to F, then the sequence of corresponding probability distributions $\{\mu_m\}$ converges effectively to μ.*

Proof. We follow the classical proof and prove that $\mu_m(f)$ converges effectively to $\mu(f)$ for a computable function with compact support f. By Lemma 1, f is uniformly computable. So, there exists a recursive function $\alpha(k)$, which is an effective modulus of uniform continuity of f. We also obtain an integer N such that $f(x) = 0$ if $|x| > 2^N$ and an integer B such that $|f(x)| \leqslant 2^B$ for all x.

Define $f_n(x) = \sum_{j=-2^N 2^n+1}^{2^N 2^n} f(j2^{-n})\chi_{((j-1)2^{-n}, j2^{-n}]}(x)$.

Then, $\mu_m(f_n) = \sum_{j=-2^N 2^n+1}^{2^N 2^n} f(j2^{-n})(F_m(j2^{-n}) - F_m((j-1)2^{-n}))$

and $\mu(f_n) = \sum_{j=-2^N 2^n+1}^{2^N 2^n} f(j2^{-n})(F(j2^{-n}) - F((j-1)2^{-n}))$ hold.

We note that each of the right-hand sides of the last two equations forms a computable sequence of reals.

By the definitions of f_n and α, $|f(x) - f_{\alpha(k)}(x)| = |f(x) - f(j2^{-\alpha(k)})| < 2^{-k}$ if $x \in ((j-1)2^{-\alpha(k)}, j2^{-\alpha(k)}]$.

Hence, we obtain $|\mu_m(f_{\alpha(k)}) - \mu_m(f)| \leqslant 2^{-k}$ and $|\mu(f_{\alpha(k)}) - \mu(f)| \leqslant 2^{-k}$.

By effective pointwise convergence of $\{F_m\}$ to F, there exists a recursive function $\beta(k, n, j)$ such that $m \geqslant \beta(k, n, j)$ implies

$|F_m(j2^{-n}) - F(j2^{-n})| < 2^{-k}$.

Define $\tilde{k} = N + 1 + B + \alpha(k+3) + k + 3$ and

$\gamma(k) = \max\{\beta(\tilde{k}, \alpha(k+3), 0), \ldots, \beta(\tilde{k}, \alpha(k+3), 2^{N+1}2^{\alpha(k+3)})\}$.

Assume $m \geqslant \gamma(k)$. Then,

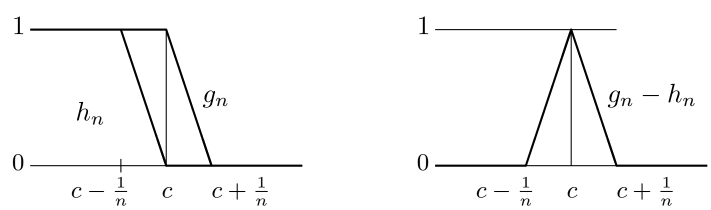
$$
\begin{aligned}
&|\mu_m(f) - \mu(f)| \\
\leqslant\ & |\mu_m(f_{\alpha(k+3)}) - \mu(f_{\alpha(k+3)})| + |\mu(f_{\alpha(k+3)}) - \mu(f)| \\
&\quad + |\mu_m(f_{\alpha(k+3)}) - \mu_m(f)| \\
\leqslant\ & \textstyle\sum_{j=-2^N 2^{\alpha(k+3)}+1}^{2^N 2^{\alpha(k+3)}} \big\{ |f(j2^{-\alpha(k+3)})|\, |F_m(j2^{-\alpha(k+3)}) - F(j2^{-\alpha(k+3)})| \\
&\quad + |f(j2^{-\alpha(k+3)})|\, |F_m((j-1)2^{-\alpha(k+3)}) - F((j-1)2^{-\alpha(k+3)})| \big\} \\
&\quad + 2 \cdot 2^{-(k+3)} \\
\leqslant\ & 2(2^{N+1} 2^{\alpha(k+3)}) 2^B 2^{-\tilde{k}} + 2 \cdot 2^{-(k+3)} < 2^{-k}.
\end{aligned}
$$

This proves the effective convergence of $\{\mu_m(f)\}$ to $\mu(f)$. □

Proposition 9. *Let us consider a computable sequence of probability distributions $\{\mu_m\}$ and a computable probability distribution μ with a bounded density. If $\{\mu_m\}$ converges effectively to μ, then the sequence of the corresponding distribution functions $\{F_m\}$ converges effectively pointwise to F, the distribution function corresponding to μ.*

Proof. We prove that $\{F_m(c)\}$ converges effectively to $F(c)$ if c is computable. Let us define $h_n(x)$ by

$$
h_n(x) = \begin{cases} 1 & \text{if } \ x \leqslant c - \frac{1}{n} \\ -n(x-c) & \text{if } \ c - \frac{1}{n} \leqslant x \leqslant c \\ 0 & \text{if } \ x \geqslant c \end{cases}.
$$



It holds that $h_n(x) \leqslant \chi_{(-\infty,c]}(x) \leqslant g_n(x)$, where g_n is the function defined in the proof of Proposition 4. Hence, we obtain $\mu(h_n) \leqslant F(c) \leqslant \mu(g_n)$ and $\mu_m(h_n) \leqslant F_m(c) \leqslant \mu_m(g_n)$.

Meanwhile, $\{g_n\}$ and $\{h_n\}$ are effectively bounded computable sequences of functions if c is a computable real. Hence, by Proposition 6, $\{\mu_m(h_n)\}$ and $\{\mu_m(g_n)\}$ converge effectively to $\mu(h_n)$ and $\mu(g_n)$ respectively as m tends to infinity. So, there exists a recursive function $\alpha(n,k)$ such that $m \geqslant \alpha(n,k)$ implies $|\mu_m(h_n) - \mu(h_n)| < 2^{-k}$ and $|\mu_m(g_n) - \mu(g_n)| < 2^{-k}$.
Hence, $m \geqslant \alpha(n,k)$ implies $\mu(h_n) - 2^{-k} \leqslant \mu_m(h_n) \leqslant F_m(c) \leqslant \mu_m(g_n) < \mu(g_n) + 2^{-k}$.

On the other hand,

$$g_n(x) - h_n(x) = \begin{cases} 0 & \text{if } x \leqslant c - \frac{1}{n} \\ n(x-c)+1 & \text{if } c - \frac{1}{n} \leqslant x \leqslant c \\ -n(x-c)+1 & \text{if } c \leqslant x \leqslant c + \frac{1}{n} \\ 0 & \text{if } x \geqslant c + \frac{1}{n} \end{cases}.$$

If we take an integer M such that 2^M is a bound of a density of μ, then $\mu(g_n - h_n) \leqslant \frac{2 \cdot 2^M}{n}$. If we put $N = 2^{k+M+2}$, then $\mu(g_N - h_N) \leqslant 2^{-(k+1)}$. Hence, we obtain $\mu(h_N) > F(c) - 2^{-(k+1)}$ and $\mu(g_N) < F(c) + 2^{-(k+1)}$. Therefore, $m \geqslant \alpha(N, k+1)$ implies $|F_m(c) - F(c)| < 2^{-k}$.

This proves the effective convergence of $\{F_m(c)\}$ to $F(c)$.

The argument above can be modified to a computable sequence of real numbers $\{c_n\}$. □

In the case where μ has a bounded density and $\{\mu_n\}$ has effectively bounded densities, we obtain the following theorem from Propositions 4, 8 and 9.

Theorem 3. *Let us consider a computable sequence of probability distributions $\{\mu_m\}$ with effectively bounded densities and a computable distribution μ with a bounded density. We denote their distribution functions with $\{F_m\}$ and F respectively. Then, $\{\mu_m\}$ converges effectively to μ if and only if $\{F_m\}$ converges effectively pointwise to F.*

In the following examples, μ_m and μ denote probability distributions, ξ_m and ξ denote the corresponding densities (if they exist) and F_m and F denote the corresponding distribution functions.

Example 3. Let μ_m be the Gaussian distribution with mean $\frac{1}{m}$ and variance $\frac{m}{m+1}$ and μ be the Gaussian distribution with mean 0 and variance 1, that is,

$$\xi_m(x) = \frac{\sqrt{m+1}}{\sqrt{2\pi m}} e^{-\frac{m+1}{2m}(x-\frac{1}{m})^2} \text{ and } \xi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}.$$

$\{\xi_m\}$ is computable and converges effectively to ξ. It also holds that $|\xi_m(x)|$, $|\xi(x)| \leqslant 1$. So, the assumption of Theorem 3 holds. By virtue of the properties of the densities, the effective convergence of $\{\mu_m\}$ to μ and that of $\{F_m\}$ to F are the consequences of Effective Dominated Convergence Theorem (see [5]). □

Example 4. Let ξ_m be defined as follows.

$$\xi_m(x) = \begin{cases} 0 & \text{if } x \leqslant -\frac{1}{m} \\ \frac{m}{2}x + \frac{1}{2} & \text{if } -\frac{1}{m} < x < \frac{1}{m} \\ 1 & \text{if } \frac{1}{m} \leqslant x \leqslant 1 - \frac{1}{m} \\ -\frac{m}{2}(x-1) + \frac{1}{2} & \text{if } 1 - \frac{1}{m} < x < 1 + \frac{1}{m} \\ 0 & \text{if } x \geqslant 1 + \frac{1}{m} \end{cases}.$$

$\{\xi_m\}$ is a computable sequence with compact support, and $\{\mu_m\}$ converges effectively to the uniform distribution on $[0,1]$ (cf. Example 1(1)). Although the density of the uniform distribution is not continuous, it is still bounded. So, the assumption of Theorem 3 holds.

By the inequality $|\int_{\mathbb{R}} f(x)\xi_m(x)dx - \int_{\mathbb{R}} f(x)\xi_{[0,1]}(x)dx| \leqslant ||f||$, we can prove the effective convergence of $\{\mu_m\}$, and hence of $\{F_m\}$. □

Example 5. Let ξ_n be defined as follows.

$$\xi_n(x) = \begin{cases} x \leqslant -\frac{1}{n} \\ n^2x + n & \text{if } -\frac{1}{n} \leqslant x \leqslant 0 \\ -n^2x + n & \text{if } 0 \leqslant x \leqslant \frac{1}{n} \\ 0 & \text{if } x \geqslant \frac{1}{n} \end{cases} .$$

$\{\mu_n\}$ converges effectively to the unit distribution δ_0, which does not have a density. This is a case to which Theorem 3 cannot be applied. Indeed, $F_n(0) = \frac{1}{2}$ but $F(0) = 1$. So, $\{F_n(0)\}$ does not converge to $F(0) = 1$. □

References

1. Billingsley, P. *Convergence of Probability Measures*, John Wiley & Sons, 1968.
2. Feller, W. *An Introduction to Probability Theory and its Applications, Vol. I, Third Edition*, John Wiley & Sons, 1968.
3. Feller, W. *An Introduction to Probability Theory and its Applications, Vol. II*, John Wiley & Sons, 1966.
4. Ito, K. *An Introduction to Probability Theory*, Cambridge University Press, 1984 (translated by the author from *Probability Theory* (in Japanese), Iwanami Shoten, 1978, 1991).
5. Mori, T., M. Yasugi and Y. Tsujii. *Effective Fine-convergence of Walsh-Fourier series,* MLQ 54, 519-534, 2008.
6. Pour-El, M.B. and J. I. Richards. *Computability in Analysis and Physics*, Springer, 1988.
7. Rényi, A. *Probability Theory*, Dover, 2007. (North-Holland, 1970.)
8. Schröder, M. Admissible Representations of Probability Measures, ENTCS 167(2007), 61-78.
9. Schröder, M. and A. Simpson. Representing Probability Measures using Probabilistic Processes, CCA 2005 (Kyoto), 211-226.
10. Weihrauch, K. Computability on the probability measures on the Borel sets of the unit interval, TCS 219, 421-437, 1999.