On the Chacon–Walsh construction in the Skorokhod Embedding Problem

Alexander Gushchin

Steklov Mathematical Institute

Asymptotic Statistics of Stochastic Processes and Applications,
Peterhof, July 18, 2017

Outline

1. **History**
    - The Skorokhod Embedding Problem
    - The Chacon–Walsh construction
    - Integrable starting and target distributions
2. **Non-integrable starting and target distributions**
    - Setting of the problem
    - Integrated distribution and quantile functions
    - Main results

Outline

1. **History**
   - The Skorokhod Embedding Problem
   - The Chacon–Walsh construction
   - Integrable starting and target distributions
2. **Non-integrable starting and target distributions**
   - Setting of the problem
   - Integrated distribution and quantile functions
   - Main results

Skorokhod Embedding Problem

The Skorokhod Embedding Problem (SEP) was originally formulated and solved in Skorokhod (1961) (English translation in Skorokhod (1965)), and gave rise to a huge amount of literature. In Obłój (2004) one finds a comprehensive survey of the state of the art to 2004, in particular, more than twenty different approaches to solve the SEP with the relations between them, different settings and generalisations as well as some other offsprings.

Original formulation of the SEP

Given: a real-valued random variable Y with

$$\mathrm{E}\,Y = 0 \quad \text{and} \quad \mathrm{E}\,Y^2 < \infty.$$

To find: a filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathrm{P})$, an $(\mathscr{F}_t, \mathrm{P})$-Brownian motion $B = (B_t)$ and an $(\mathscr{F}_t)$-stopping time τ with $\mathrm{E}\,\tau < \infty$ such that

$$B_\tau \sim Y.$$

Remark

Let f be a function such that

$$f(B_1) \sim Y,$$

where Y is an arbitrary random variable. Put

$$\tau := \inf\{t \geq 1 \colon B_t = f(B_1)\}.$$

Due to recurrence of a Brownian motion, we have $B_\tau = f(B_1) \sim Y$. But the requirement $\mathsf{E}\,\tau < \infty$ is not satisfied for this stopping time τ (unless f is the identity function a.e.). This solution is attributed to Doob and is intended to show that without additional requirements the problem is trivial.

Note, however, that the requirement $\mathsf{E}\,\tau < \infty$ implies $\mathsf{E}\,B_\tau = 0$ and $\mathsf{E}\,B_\tau^2 = \mathsf{E}\,\tau < \infty$.

Outline

1. **History**
   - The Skorokhod Embedding Problem
   - The Chacon–Walsh construction
   - Integrable starting and target distributions
2. **Non-integrable starting and target distributions**
   - Setting of the problem
   - Integrated distribution and quantile functions
   - Main results

Balayage

Let us recall the definition of the balayage.
For a probability measure μ on $\mathbb{R}$ and an interval $I = (a, b)$, $-\infty < a < b < +\infty$, the balayage μ_I of μ on I is defined as the measure which coincides with μ outside $[a, b]$, vanishes on (a, b), and such that

$$\mu_I(\{a\}) = \int_{[a,b]} \frac{b-x}{b-a}\,\mu(dx), \qquad \mu_I(\{b\}) = \int_{[a,b]} \frac{x-a}{b-a}\,\mu(dx).$$

Since

$$\int_{[a,b]} \mu_I(dx) = \int_{[a,b]} \mu(dx) \quad \text{and} \quad \int_{[a,b]} x\,\mu_I(dx) = \int_{[a,b]} x\,\mu(dx),$$

the balayage μ_I is a probability measure and has the same mean as μ (if defined).

Chacon & Walsh (1976)

Chacon & Walsh (1976) proposed to construct a solution τ to the SEP as the limit of an increasing sequence of stopping times τ_n, each being the first exit time of B after the previous one from a bounded interval:

$$\tau_n := \inf\{t \geq \tau_{n-1} \colon B_t \notin (a_n, b_n)\}.$$

The essence of the construction is a graphical picture explaining a right choice of intervals (a_n, b_n). It includes a more special construction of Dubins (1968).

It is easy to check that the distribution of B_{τ_n} is the balayage of the distribution of $B_{\tau_{n-1}}$ on (a_n, b_n).

Convex orders

For integrable random variables X and Y, we say that (the distribution of) X is less than or equal to (the distribution of) Y with respect to convex ordering ($X \leqslant_{cx} Y$), if $\mathsf{E} f(X) \leq \mathsf{E} f(Y)$ for any convex $f \colon \mathbb{R} \to \mathbb{R}$. It is necessary and sufficient for $X \leqslant_{cx} Y$ that one can find random variables X' and Y' on a common probability space such that $X' \overset{\text{law}}{=} X$, $Y' \overset{\text{law}}{=} Y$, and $\mathsf{E}(Y'|X') = X'$.

X is less than Y in increasing convex order ($X \leq_{icx} Y$) if $\mathsf{E}\, X^+ < \infty$, $\mathsf{E}\, Y^+ < \infty$, and $\mathsf{E}[\varphi(X)] \leq \mathsf{E}[\varphi(Y)]$ for any bounded from below increasing convex function.

X is less than Y in decreasing convex order ($X \leq_{decx} Y$) if $\mathsf{E}\, X^- < \infty$, $\mathsf{E}\, Y^- < \infty$, and $\mathsf{E}[\varphi(X)] \leq \mathsf{E}[\varphi(Y)]$ for any bounded from above decreasing convex function.

Potential functions

The potential function of an integrable random variable X is defined by

$$\mathsf{U}_X(x) := -\mathsf{E}\,|X - x|.$$

It is easy to check that $\mathsf{U}_X(x)$ is a concave function on $\mathbb{R}$ whose right derivative equals

$$[\mathsf{U}_X(x)]'_+ = 1 - 2\mathsf{F}_X(x)$$

for every x. Here and below $\mathsf{F}_X(x) := \mathsf{P}(X \leq x)$ is the distribution function of X.

Properties of potential functions

Obviously, if $X \leqslant_{cx} Y$, then

$$\mathrm{U}_Y(x) \le \mathrm{U}_X(x), \quad x \in \mathbb{R}.$$

In particular,

$$\mathrm{U}_X(x) \le -|\mathrm{E}\, X - x|, \quad x \in \mathbb{R}.$$

If a sequence X_n is uniformly integrable and weakly converges to X, then $\mathrm{U}_{X_n}(x)$ converges pointwise to $\mathrm{U}_X(x)$.
If $\mathrm{E}\, X = 0$, then $\mathrm{E}\, X^2$ is the area between the graphs of $\mathrm{U}_X(x)$ and $-|x|$, the potential function of 0.

Balayage: picture

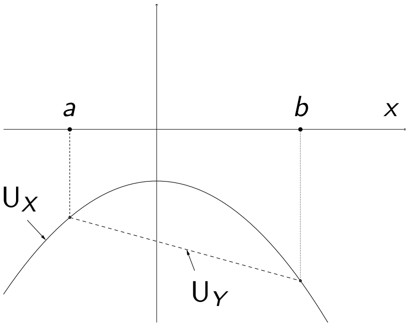


Figure: The solid line is the graph of the potential function $\mathrm{U}_X(x)$. If the distribution of Y is the balayage of the distribution of X on (a, b), then the graph of the potential function $\mathrm{U}_Y(x)$ coincides with $\mathrm{U}_X(x)$ outside (a, b) and is the dashed straight line on (a, b).

The Chacon–Walsh construction: picture

Let $X \leqslant_{cx} Y$ and $\mathrm{E}\, Y^2 < \infty$. Represent $\mathrm{U}_Y(x)$ as the countable infimum of its tangent lines. Take the first tangent line and construct $\mathrm{U}_{X_1}(x)$ as the minimum of the corresponding affine function and $\mathrm{U}_X(x)$. Then take the second tangent line and so on. . . .

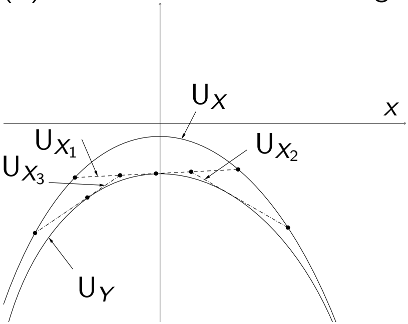


Figure: Plots of $\mathrm{U}_X(x)$, $\mathrm{U}_{X_1}(x)$, $\mathrm{U}_{X_2}(x)$, $\mathrm{U}_{X_3}(x)$, $\mathrm{U}_Y(x)$.

Outline

1. **History**
   - The Skorokhod Embedding Problem
   - The Chacon–Walsh construction
   - Integrable starting and target distributions
2. **Non-integrable starting and target distributions**
   - Setting of the problem
   - Integrated distribution and quantile functions
   - Main results

Minimal stopping time

If we do not assume that the target distribution has a finite variance, it is no longer clear how to select "good" stopping times instead of requiring $\mathsf{E}\,\tau < \infty$. A very natural way to select "good" stopping times is to require them to be minimal in the following sense.

> *A finite stopping time τ is said to be minimal (for a process B) if, for a stopping time σ, $\sigma \leqslant \tau$ and $B_\sigma \sim B_\tau$ imply $\sigma = \tau$ a.s.*

In the context of the SEP this definition was suggested in Monroe (1972). In particular, for centered target distributions minimality is characterized in that paper as follows: Let $B_0 = 0$ and τ be a finite stopping time such that $\mathsf{E}\,B_\tau = 0$. Then τ is minimal if and only if the process $(B_{t\wedge\tau})_{t\geqslant 0}$ is uniformly integrable.

Centered target distributions

This allows us to extend the Chacon–Walsh construction to the case where the target distribution is centered (without requiring finiteness of the variance) and to show that the obtained solution of the SEP is a minimal stopping time. Moreover, the same conclusion holds if we are given a starting distribution μ_0 and a target distribution μ such that $\mu_0 \leqslant_{\mathsf{cx}} \mu$ (in particular, starting and target distributions have the same finite means).

Noncentered target distributions

It was shown in Monroe (1972) that a minimal stopping time exists for any target distribution. However, we loose uniform integrability property if a target distribution is noncentered. Nevertheless, if a target distribution is integrable, there are convenient characterization of minimal stopping times proved in Cox and Hobson (2006). If a target distribution is non-integrable, a general convenient characterization of minimal stopping times is not known.

Integrable starting and target distributions

Using the characterization of minimal stopping times in Cox and Hobson (2006), it is rather simply to modify the Chacon–Walsh construction to the case where a target distribution has a finite mean, or, more generally, if we are given an integrable starting distribution μ_0 and an integrable target distribution μ such that $\mu_0 \leqslant_{\mathsf{icx}} \mu$ or $\mu_0 \leqslant_{\mathsf{decx}} \mu$ and to show that the corresponding solution is a minimal stopping time.
Finally, a nontrivial result by Cox (2008) says that the same conclusion holds for any integrable starting and target distributions.

Outline

1. History
   - The Skorokhod Embedding Problem
   - The Chacon–Walsh construction
   - Integrable starting and target distributions
2. **Non-integrable starting and target distributions**
   - Setting of the problem
   - Integrated distribution and quantile functions
   - Main results

Setting of the problem

Given: a starting distribution μ_0 and a target distribution μ.

To construct: a sequence of pairs (a_n, b_n) such that, for a Brownian motion $B = (B_t)$ with $B_0 \sim \mu_0$, $\tau := \lim \tau_n$, where $\tau_0 := 0$ and

$$\tau_n := \inf\{t \geq \tau_{n-1} \colon B_t \notin (a_n, b_n)\},$$

the stopping time τ is minimal and $B_\tau \sim \mu$.

We do not know the full answer to this problem. We only provide a sufficient condition on the pair (μ_0, μ) for such construction to exist.

Problems

The first problem is to provide an adequate replacement of potential functions in the construction. Then we need to construct a sequence (a_n, b_n) such that

- B_{τ_n} weakly converges to μ.
- τ is a.s. finite.
- τ is minimal.

Outline

1. History
   - The Skorokhod Embedding Problem
   - The Chacon–Walsh construction
   - Integrable starting and target distributions
2. **Non-integrable starting and target distributions**
   - Setting of the problem
   - Integrated distribution and quantile functions
   - Main results

Integrated distribution function (IDF)

This part is based on the joint work with Dmitriy Borzykh (Higher School of Economics)

The **integrated distribution function of a random variable** X is defined by

$$\mathsf{J}_X(x) := \int_0^x \mathsf{F}_X(t)\,dt, \quad x \in \mathbb{R},$$

with convention: $\int_a^b f(x)\,dx := -\int_b^a f(x)\,dx$, if $b < a$.

Properties of IDF

An integrated distribution function J_X has the following properties:

(i) $\mathrm{J}_X(0) = 0$,

(ii) J_X is convex, increasing and finite everywhere on $\mathbb{R}$,

(iii) $\lim\limits_{x\to-\infty} \mathrm{J}_X(x) = -\mathrm{E}[X^-]$ and $\lim\limits_{x\to+\infty}(x - \mathrm{J}_X(x)) = \mathrm{E}[X^+]$,

(iv) $\lim\limits_{x\to-\infty} \frac{\mathrm{J}_X(x)}{x} = 0$ and $\lim\limits_{x\to+\infty} \frac{\mathrm{J}_X(x)}{x} = 1$,

(v) $(\mathrm{J}_X)'_+(x) = \mathrm{F}_X(x)$,

(vI) If $\mathrm{E}\,|X| < \infty$, then $\mathrm{U}_X(x) = x - \mathrm{E}\,|X| - 2\mathrm{J}_X(x)$.

A typical plot of IDF

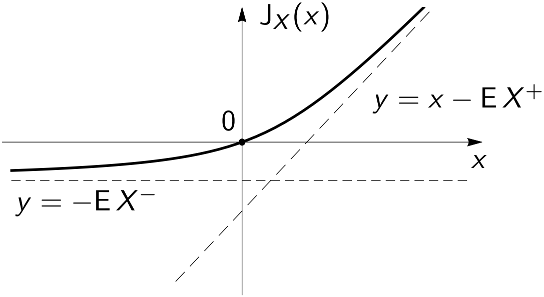


Figure: A typical plot of IDF under assumption $\mathsf{E}\,X^- < \infty$, $\mathsf{E}\,X^+ < \infty$.

Quantile functions

For a random variable X, its lower quantile function Q_X is defined by

$$\mathsf{Q}_X(u) := \inf\{x \colon \mathsf{F}_X(x) \geqslant u\}, \quad u \in (0,1).$$

The letter U always stands for a random variable with the uniform distribution on $(0,1)$. As a rule, U is just the identity function on $(0,1)$ interpreted as a random variable on the probability space, which consists of the interval $(0,1)$ with the Borel σ-field and the Lebesgue measure. Recall that

$$\mathsf{Q}_X(U) \overset{\text{law}}{=} X.$$

Integrated quantile function (IQF)

The Fenchel transform of the integrated distribution function of a random variable X

$$\mathsf{K}_X(u) := \sup_{x\in\mathbb{R}}\{xu - \mathsf{J}_X(x)\}, \quad u \in \mathbb{R},$$

is called the **integrated quantile function** of X.

The statement (v) on the next slide explains the meaning of the term 'integrated quantile function'.

Properties of IQF

An integrated quantile function K_X has the following properties:

(i) the function K_X is convex and lower semicontinuous,

(ii) it takes finite values on $(0,1)$ and equals $+\infty$ outside $[0,1]$,

(iii) the Fenchel transform of K_X is J_X, i. e. for any $x \in \mathbb{R}$,

$$\sup_{u\in\mathbb{R}}\{xu - \mathsf{K}_X(u)\} = \mathsf{J}_X(x),$$

(iv) $\min\limits_{u\in\mathbb{R}} \mathsf{K}_X(u) = 0$ and $\mathsf{K}_X(u) = 0 \Leftrightarrow \mathsf{F}_X(0-) \leq u \leq \mathsf{F}_X(0)$.

(v) for every $u \in [0,1]$,

$$\mathsf{K}_X(u) = \int_{u_0}^{u} \mathsf{Q}_X(s)\,ds,$$

where u_0 is any zero of K_X,

(vi) $\mathsf{K}_X(0) = \mathsf{E}\,X^-$ and $\mathsf{K}_X(1) = \mathsf{E}\,X^+$.

A typical plot of IQF

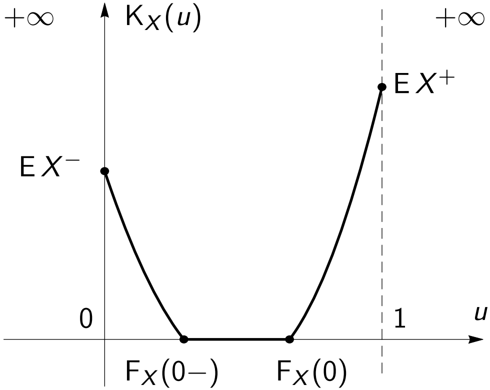


Figure: A typical plot of IQF.

IQF and the convex orders

Assuming that $\mathsf{K}_X(1) = \mathsf{E}\, X^+ < \infty$ let us introduce a **shifted integrated quantile function**:

$$\mathsf{K}_X^{[1]}(u) := \mathsf{K}_X(u) - \mathsf{K}_X(1), \qquad u \in [0,1].$$

Let X and Y be random variables with finite means.

(i) $X \leq_{cx} Y$ if and only if $\mathsf{K}_X^{[1]}(u) \geq \mathsf{K}_Y^{[1]}(u)$ for all $u \in [0,1]$ and $\mathsf{K}_X^{[1]}(0) = \mathsf{K}_Y^{[1]}(0)$.

(ii) $X \leq_{icx} Y$ if and only if $\mathsf{K}_X^{[1]}(u) \geq \mathsf{K}_Y^{[1]}(u)$ for all $u \in [0,1]$.

IQF and weak convergence

Theorem

The following statements are equivalent:

(i) *The sequence* (X_n) *weakly converges.*

(ii) *There is a sequence* (c_n) *of numbers such that, for every* $u \in (0,1)$ *, the sequence* $(\mathsf{K}_{X_n}(u) - c_n)$ *converges to a finite limit.*

(iii) *The sequence* (K_{X_n}) *converges uniformly on every* $[\alpha, \beta] \subseteq (0,1)$*.*

Moreover, in this case if X *is a weak limit of* (X_n)*, then* $\mathsf{K}_X(u) = \lim_{n\to\infty} \mathsf{K}_{X_n}(u)$ *for all* $u \in (0,1)$*.*

Balayage and IQF: picture

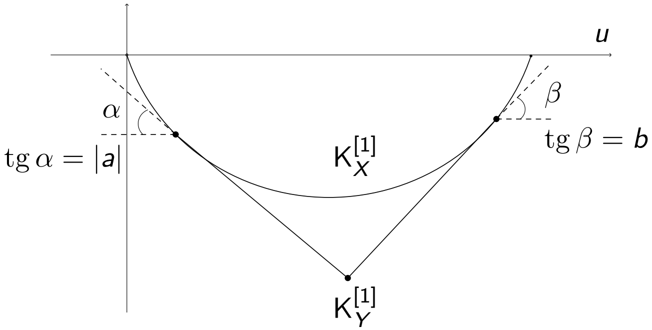
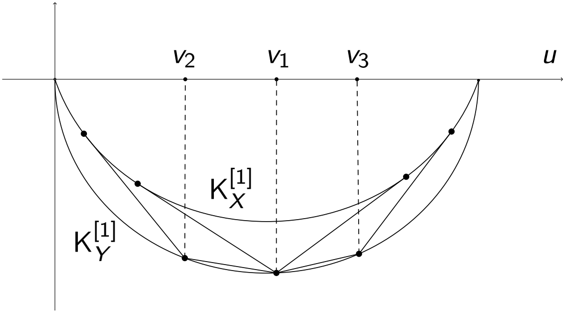


Figure: If the distribution of Y is the balayage of the distribution of X on (a, b), then, to construct $\mathsf{K}_Y^{[1]}(u)$, pass the tangent lines with the slopes a and b to the graph of K_X, replace the curve on this graph between points where the graph meets the lines by the corresponding segments of these lines.

The Chacon–Walsh construction and IQF: picture

Let $X \leqslant_{\mathrm{cx}} Y$. Take an arbitrary sequence $\{v_n\}$ of distinct points in $(0,1)$ such that $\{v_n\colon n = 1, 2, \dots\}$ is dense in $[0,1]$. Draw tangent lines to the curve $\mathsf{K}_X^{[1]}$ through the point $(v_1, \mathsf{K}_Y^{[1]}(v_1))$, their slopes are a_1 and b_1. Construct $\mathsf{K}_{X_1}^{[1]}$ as on the previous slide. Continue recursively.



Outline

1. History
   - The Skorokhod Embedding Problem
   - The Chacon–Walsh construction
   - Integrable starting and target distributions
2. **Non-integrable starting and target distributions**
   - Setting of the problem
   - Integrated distribution and quantile functions
   - Main results

Two cases with non-integrable starting distributions

The above considerations show that we may easily extend the Chacon–Walsh construction to the following two cases:

- ("submartingale") A starting distribution μ_0 and a target distribution μ satisfy
$$\mu_0 \leqslant_{\mathsf{icx}} \mu.$$
- ("supermartingale") A starting distribution μ_0 and a target distribution μ satisfy
$$\mu_0 \leqslant_{\mathsf{decx}} \mu.$$

What can be non-integrable?

In particular, in the first ("submartingale") case

$$\int_{(0,\infty)} x\,\mu_0(dx) \le \int_{(0,\infty)} x\,\mu(dx) < \infty$$

and

$$-\infty \le \int_{(-\infty,\infty)} x\,\mu_0(dx) \le \int_{(-\infty,\infty)} x\,\mu(dx).$$

Theorem

Theorem

Let μ_0 and μ be distributions on $\mathbb{R}$ such that $\int_{\mathbb{R}} x^+ \, \mu(dx) < \infty$ and

$$\int_{\mathbb{R}} (x-y)^+ \, \mu_0(dx) \le \int_{\mathbb{R}} (x-y)^+ \, \mu(dx)$$

for all $y \in \mathbb{R}$. Let B be a Brownian motion with the initial distribution $B_0 \sim \mu_0$. Then there is an increasing sequence of stopping times $0 = \tau_0 \le \tau_1 \le \dots \le \tau_n \le \dots$ such that $\tau := \lim_{n\to\infty} \tau_n$ is a minimal a. s. finite stopping time, the distribution of B_{τ_n} is a balayage of the distribution of $B_{\tau_{n-1}}$ for each $n = 1, 2, \dots,$ and

$$B_\tau \sim \mu.$$

Proof

Let $X \sim \mu_0$ and $Y \sim \mu$. Then $\mathrm{K}_Y^{[1]}(u) \le \mathrm{K}_X^{[1]}(u)$ for all $u \in [0,1]$. Take an arbitrary sequence $\{v_n\}$ of distinct points in $(0,1)$ such that $\{v_n \colon n = 1,\,2,\,\dots\}$ is dense in $[0,1]$. Recursively define X_n as on the previous picture. Then we obtain a sequence $\{X_n\}$ of random variables such that

$$\mathrm{K}_Y^{[1]}(u) \le \mathrm{K}_{X_n}^{[1]}(u) \le \mathrm{K}_{X_{n-1}}^{[1]}(u) \le \mathrm{K}_X^{[1]}(u), \quad u \in (0,1]$$

and $\mathrm{K}_Y^{[1]}(v_n) = \mathrm{K}_{X_n}^{[1]}(v_n)$, which implies $\mathrm{K}_Y^{[1]}(v_n) = \mathrm{K}_{X_m}^{[1]}(v_n)$ for all n and $m \ge n$. Then $\lim_{n\to\infty} \mathrm{K}_{X_n}^{[1]}(u)$ exists, is finite for all $u \in (0,1)$, and coincides with $\mathrm{K}_Y^{[1]}(u)$ on the set $\{v_n \colon n = 1,\,2,\,\dots\}$. Being a convex function in u, this limit coincides with $\mathrm{K}_Y^{[1]}(u)$ everywhere on $(0,1)$. It follows from the theorem on "weak convergence and IQF" that X_n weakly converges to Y. One can also show that the sequence $\{X_n^+\}$ is uniformly integrable.

Proof (continued)

Moreover, our construction provides an interval (a_n, b_n) such that the distribution of X_n is the balayage of the distribution of X_{n-1} on (a_n, b_n). Now recursively define

$$\tau_n := \inf\{t \geq \tau_{n-1} \colon B_t \notin (a_n, b_n)\}.$$

Then B_{τ_n} has the same distribution as X_n.

The next step is to show that $\tau = \lim \tau_n < \infty$ a.s.

Proof (continued)

For $c > 0$, let

$$H_c = \inf\{t \geq 0\colon B_t \geq c\}.$$

We prove, by induction on n, that, for any $c > 0$ and n,

$$c\mathsf{P}(\tau_n \geq H_c) \leq \mathsf{E}[B_{\tau_n}\mathbb{1}_{\{\tau_n \geq H_c\}}] \leq \mathsf{E}[B^+_{\tau_n}] \leq \mathsf{E}[Y^+].$$

If $\mathsf{P}(\tau = \infty) = \delta > 0$, then the limit of the expression on the left in the last inequality is greater than or is equal to $c\delta$, which is greater than the right-hand side if c is large enough. This contradiction proves that $\tau < \infty$ a.s. This implies that B_{τ_n} converges a.s. to B_τ and, hence, $B_\tau \sim \mu$.

Proof (continued)

It remains to prove that τ is a minimal stopping time. Here we apply a result from G. and Urusov (2016), which says that it is enough to find a one-to-one function G such that $G(B)^\tau = (G(B_{t\wedge\tau})_{t\geq 0})$ is a closed submartingale. This is easy: take a convex G which is strictly increasing and bounded from below and such that $G(x) = x$ for $x > 0$.

A general result

Combining the previous result with considerations in Cox (2008), we obtain that the following condition on $X \sim \mu_0$ and $Y \sim \mu$ is sufficient for the conclusion of the previous theorem to be true:

$$\sup_{u \in (0,1)} \left[\mathsf{K}_Y(u) - \mathsf{K}_X(u)\right] < \infty,$$

or, equivalently,

$$\sup_{x \in \mathbb{R}} \left[\mathsf{J}_X(x) - \mathsf{J}_Y(x)\right] < \infty.$$

In particular, there exists a minimal embedding with the Chacon–Walsh construction for any integrable target distribution and any starting distribution.

Thank you for the attention!