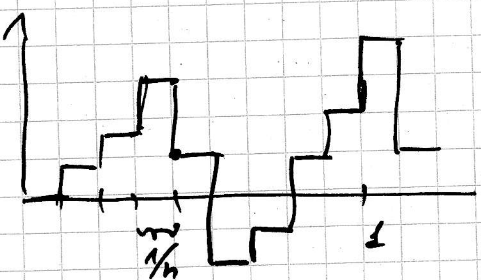
5. Weak convergence and the invariance principle

5.1 Maximum of the BM

$\xi_1, \xi_2, \dots$ iid $E\xi_i = 0$, $Var\, \xi_i = 1$

$$S_n = \xi_1 + \dots + \xi_n, \quad S_0 = 0$$

$$X_n(t) = \frac{S_{\lfloor nt \rfloor}}{\sqrt{n}}, \quad t \geq 0$$



Scaled random walk

$$\left(X_n(t_1), \dots, X_n(t_k)\right) \xrightarrow{d} \left(B(t_1), \dots, B(t_k)\right) \quad (*)$$

$0 \leq t_1 < \dots < t_k$ $\qquad$ $(B(t), t \geq 0)$ Brownian motion

$$\max_{t \in [0,1]} X_n(t) = \max_{0 \leq j \leq n} \frac{S_j}{\sqrt{n}} \xrightarrow[?]{d} \max_{t \in [0,1]} B(t)$$

Convergence $(*)$ of finite-dim. distributions is not enough, because every <u>version</u> of BM satisfies $(*)$

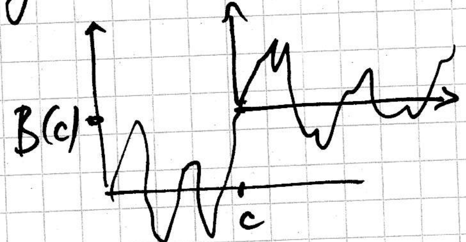
<u>Definition</u> Processes $(X(t), t \geq 0)$, $(Y(t), t \geq 0)$ are <u>versions</u> of one another if

$$X(t) = Y(t) \text{ a.s. for every fixed } t.$$

Example $B(t) + 1_{\{Y=t\}}$ $\qquad Y \sim$ Exponential

is a discontinuous version of the BM.

Easy to check: $B(t+c) - B(c)$ is a BM independent of the path on $[0,c]$



Strong Markov property: for τ a stopping time adapted to $\mathcal{F}_t = \sigma(B(s), s \le t)$, $t \ge 0$,

$$\hat{B}(t) = B(t+\tau) - B(\tau), \quad t \ge 0$$

is a BM independent of $\mathcal{F}_\tau$ (history before τ)

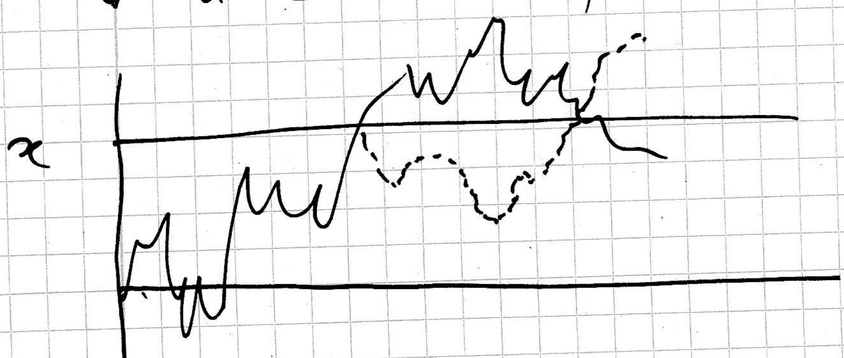
Def. $A \in \mathcal{F}_\tau \iff A \cap \{\tau \le t\} \in \mathcal{F}_t$, $t \ge 0$.

Also: $(-B(t), t \ge 0)$ is a BM.

Reflection principle: Let $\tau_x = \min\{t \ge 0, B(t) = x\}$, $x \ge 0$.

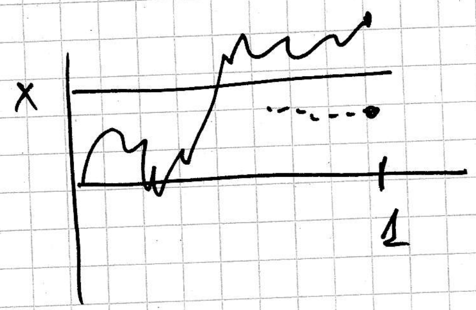
$$\tilde{B}(t) = B(t)\, 1(\tau_x \le t) + (x - (B(t) - x))\, 1(\tau_x > t)$$

is a BM independent of $\mathcal{F}_{\tau_x}$.



Let $M := \max\limits_{t \in [0,1]} B(t)$

③ Note that $\tau_x \le 1 \iff M \ge x$



By Reflection Principle

$$\mathbb{P}(M > x, B(1) > x) = \mathbb{P}(M > x, B(1) < x)$$

$$\Rightarrow \mathbb{P}(M > x) = \mathbb{P}(M > x, B(1) > x) + \mathbb{P}(M > x, B(1) \le x)$$

$$= 2\,\mathbb{P}(M > x, B(1) > x)$$

$$= 2\,\mathbb{P}(B(1) > x) = \mathbb{P}(|B(1)| > x)$$

$$= \sqrt{\frac{2}{\pi}} \int_x^{\infty} e^{-y^2/2}\, dy$$

$$\Longrightarrow M \overset{d}{=} |B(1)|$$

More generally, by self-similarity for every t

$M(t) = \max_{s \in [0,t]} B(s)$ satisfies $M(t) \overset{d}{=} |B(t)|$

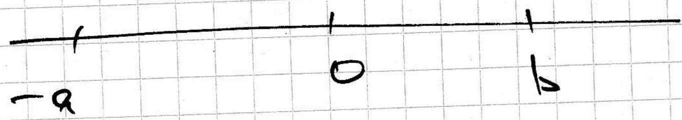
Also, $M(t) \overset{d}{=} \sqrt{t}\, M(1)$.

5.2 Skorokhod embedding

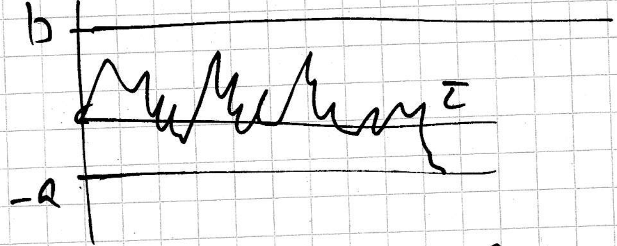
τ stopping time with $\mathbb{E}\tau < \infty$, adapted to BM

By optional sampling $\mathbb{E} B(\tau) = 0$, because $(B(t), t \geq 0)$ is a martingale. Which mean-zero distributions $B(\tau)$ may have?

Example



$\tau =$



$$\tau = \min \{ t : B(t) \in \{-a, b\} \}.$$

$$\mathbb{P}(B(\tau) = a) = \frac{b}{a+b} \; ; \; \mathbb{P}(B(\tau) = b) = \frac{a}{a+b}$$

$\Longrightarrow$ $B(\tau)$ may have any mean-0 two-point distribution.

Theorem (Skorokhod Embedding) For every random variable ξ with $\mathbb{E}\xi = 0$, $\mathrm{Var}\, \xi < \infty$ there exists a stopping time τ such that $\mathbb{E}\tau < \infty$ and

$$B(\tau) \overset{d}{=} \xi .$$

(5) For random walk $S_n = \xi_1 + \dots + \xi_n$

$\mathbb{E}\xi_1 = 0$; $\mathrm{Var}\,\xi_1 < \infty$

There exist i.i.d. $\tau_1, \tau_2, \dots$ such that

$S_n \stackrel{d}{=} B(\tau_1 + \dots + \tau_n)$ jointly for $n \geq 0$.

That is,

$\left(\tau_n, B(\tau_1 + \dots + \tau_n) - B(\tau_1 + \dots + \tau_{n-1})\right)$ are

iid copies of (τ_1, ξ_1).

5.3 Weak convergence of probability measures

Definition Let $(E, \mathcal{B}, \rho)$ be a complete metric space with distance function ρ.

For probability measures $P_n, n \in \mathbb{N}, P$ we say that P_n weakly converge to P if

$$\int_E f(x)\, dP_n(x) \longrightarrow \int_E f(x)\, dP(x)$$

for all continuous bounded functions

$f : E \to \mathbb{R}$.

Notation: $\quad P_n \Rightarrow P$

Example $\quad S_n = \xi_1 + \dots + \xi_n$ random walk

$$\frac{S_n}{n} \xrightarrow{d} m \iff \mathbb{E} f\left(\frac{S_n}{n}\right) \to \mathbb{E} f(m)$$

$$m = \mathbb{E}\xi_1, \qquad \mathbb{E} f(m) = f(m)$$

6 P_n distribution of $\frac{S_n}{n}$

P δ_m Dirac measure at m.

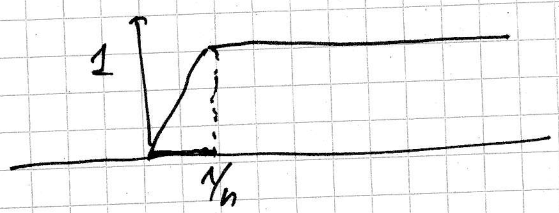
For $\mathbb{R}$-valued random variables X_n, X
weak convergence of distributions $\Leftrightarrow X_n \xrightarrow{d} X$
conv. in distr.

$\Leftrightarrow F_{X_n}(x) \to F_X(x)$ for every
continuity point x of F_X, where

$$F_{X_n}(x) = \mathbb{P}(X_n \le x), \quad F_X(x) = \mathbb{P}(X \le x)$$

<u>Example</u> Uniform $[0, \frac{1}{n}] \Rightarrow \delta_0$

$$F_n(x) = nx\mathbb{1}(x \in [0, 1/n]) \qquad F(x) = \mathbb{1}(x \ge 0)$$



<u>Example</u> CLT $\quad F_n(x) = \mathbb{P}\left(\frac{S_n - nm}{\sigma\sqrt{n}} \le x\right)$

$\sigma^2 = \mathrm{Var}\, S_1$ $\quad F(x) = \Phi(x) = \int_{-\infty}^{x} \frac{e^{-y^2/2}}{\sqrt{2\pi}} dy$

$F_n(x) \to \Phi(x)$ for every x, since Φ continuous

<u>Example</u> $X_n \sim$ Binomial$(n, 1/n)$

$X \sim$ Poisson(1)

Weak convergence of distributions means in the discrete case $\quad \mathbb{P}(X_n = k) \to \mathbb{P}(X = k)$ for every k.

(7) <u>Proposition</u> $P_n \Rightarrow P$ on E_1

$g : E_1 \to E_2$ continuous mapping

imply that also $g \circ P_n \Rightarrow g \circ P$,

where $g \circ P(A) = P(g^{-1}(A))$ for $A \in \mathcal{B}(E_2)$.

<u>Theorem</u> (The 'portmanteau' theorem).

Conditions are equivalent:

(i) $P_n \Rightarrow P$

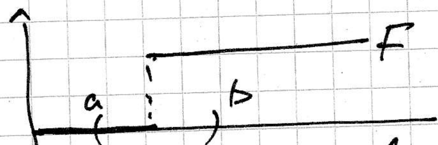
(ii) $\limsup_n P_n(A) \le P(A)$ for closed A

(iii) $\liminf_n P_n(A) \ge P(A)$ for open A

(iv) $\lim_{n\to\infty} P_n(A) = P(A)$ if $P(\partial A) = 0$

where $\partial A = \mathrm{cl}\, A \setminus \mathrm{int}\, A$

To (iv) in $\mathbb{R}$

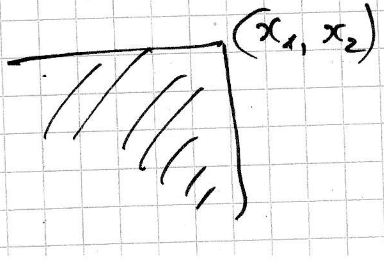


$F_n(b) - F_n(a)$ if a, b are not discontinuity points of F

<u>Example</u> $E = \mathbb{R}^k$ we have $P_n \Rightarrow P$ if

$F_n(x_1, \dots, x_k) \to F(x_1, \dots, x_k)$ for all continuity points $(x_1, \dots, x_k)$ of F, where

$F_n(x_1, \dots, x_k) = P_n((-\infty, x_1] \times \dots \times (-\infty, x_k])$



(8) <u>Example</u> $C[0,1]$, space of continuous functions $x : [0,1] \to \mathbb{R}$

$$\rho(x,y) = \sup_{t \in [0,1]} |x(t) - y(t)|$$

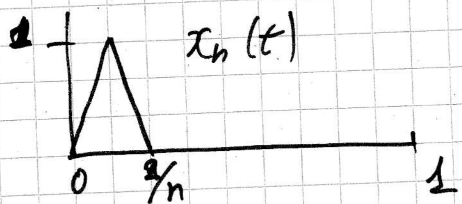
For $0 \le t_1 < \ldots < t_k \le 1$

$$\pi_{t_1,\ldots,t_k} : x \to (x(t_1), \ldots, x(t_k))$$

k- dim projection

$P_n \Rightarrow P$ implies $\pi_{t_1,\ldots,t_k} \circ P_n \Rightarrow \pi_{t_1,\ldots,t_k} \circ P$

but this is not sufficient!



P_n δ-measure on $x_n(t)$

$$\pi_{t_1,\ldots,t_k} \circ P_n \Rightarrow \delta_{(0,\ldots,0)}$$

But P_n has no weak limit, since x_n does not converge in $C[0,1]$.

<u>Definition</u> A family of probability measures $\mathcal{P} = (P_j, j \in J)$ on E is called ~~tigh~~ relatively compact if every sequence P_{j_n}, $j_n \in J$, contains a weakly convergent subsequence.

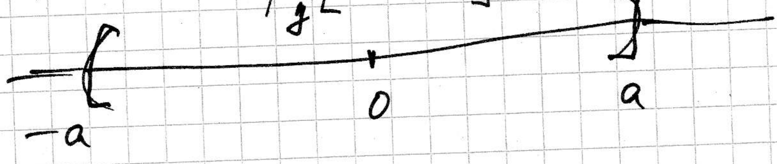
<u>Definition</u> $\mathcal{P}$ is tight if for every $\varepsilon > 0$ there exists compact $K \subset E$ such that

$$\sup_{j \in J} P_j(E \setminus K) \le \varepsilon$$

(9) (E, ρ) is Polish space if E is complete and separable (exists countable dense subset).

<u>Theorem</u> (Prokhorov) On Polish space (E, ρ) a family of probability measures $\mathcal{P} = (P_j, j \in J)$ is relatively compact if and only if it is tight.

<u>Example</u> P_n in $\mathbb{R}$ $\quad P_j[-a, a] > 1 - \varepsilon$



P_n has a weak limit only if (P_n) is tight.

<u>5.4</u> Weak convergence in $C[0,1]$

Note: $\pi_{t_1,\dots,t_k} \circ P$ uniquely determine P, but

$\pi_{t_1,\dots,t_k} \circ P_n \Rightarrow \pi_{t_1,\dots,t_k} P$ does not imply $P_n \Rightarrow P$.

<u>Proposition</u> Let P_n, P be distributions of random processes $(X_n(t), t \in [0,1])$ in $C[0,1]$
$(X(t), t \in [0,1])$.

Then $X_n \xrightarrow{d} X$ (or $P_n \Rightarrow P$) if

(i) $(X_n(t_1), \dots, X_n(t_k)) \xrightarrow{d} (X(t_1), \dots, X(t_k))$ for all $0 \le t_1 < \dots < t_k \le 1$, $k \in \mathbb{N}$.

(ii) The sequence (P_n) is tight.

(10) Need condition for tightness in $C[0,1]$.

<u>Definition</u> Modulus of continuity $x \in C[0,1]$

$$w(x;h) = \sup_{|t-s| \le h} |x(t) - x(s)|$$

By uniform continuity $w(x;h) \to 0$ as $h \to 0$

<u>Arzela–Ascoli Theorem</u>: $(x_j, j \in J) \subset C[0,1]$ is relatively compact if and only if

(i) $\sup_{j \in J} |x_j(0)| \le r$ for some $r > 0$

(ii) $\lim_{h \to 0} \sup_{j \in J} w(x_j; h) = 0$

(Family is uniformly bounded and equi-continuous)

<u>Theorem</u> (P_n) is tight on $C[0,1]$ if and only if

(i) $\lim_{r \to \infty} \limsup_n P_n\{x : |x_n(0)| > r\} = 0$

$(\pi_0 \circ P_n$ is tight$)$

(ii) $\lim_{h \to 0} \limsup_n P_n\{x : w(x;h) > \varepsilon\} = 0$

Condition (ii) can be replaced by more convenient condition

(ii') $\forall t \in [0,1]$, $\varepsilon > 0$

$$\lim_{h \to 0} \limsup_n \frac{1}{h} \max_{1 \le j \le \frac{1}{h}} P_n\{x : \sup_{(j-1)h \le s \le jh} |x(t) - x(s)| > \varepsilon\} = 0$$

This follows from estimate

$$P\{x : w(x; h) > 3\varepsilon\} \leq \sum_{1 \leq j \leq \frac{1}{h}} P\{x : \sup_{(j-1)h \leq s \leq jh} |x(t) - x(s)| > \varepsilon\}$$

5.5 Donsker's Invariance principle.

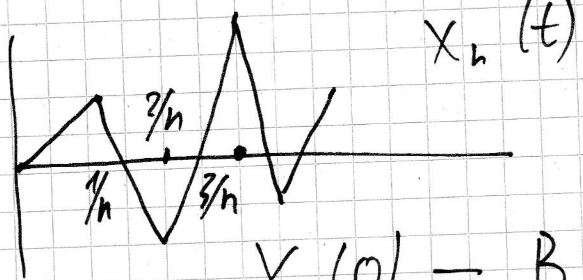
Want to show that

$$\left(\frac{S_{\lfloor nt \rfloor}}{\sqrt{n}}, t \geq 0\right) \xrightarrow{d} (B(t), t \in [0,1])$$

$S_n = \xi_1 + \ldots + \xi_n$ $\qquad \mathbb{E}\xi_1 = 0, \ \mathrm{Var}\, \xi_1 = 1$

Interpolate to have in $C[0,1]$

$$X_n(t) = \frac{1}{\sqrt{n}} \left(\sum_{n=1}^{\lfloor nt \rfloor} \xi_j + (nt - \lfloor nt \rfloor)\, \xi_{\lfloor nt \rfloor + 1} \right)$$



$X_n(0) = B(0) = 0$ "tightness at 0"

To check (ii') we use a maximal inequality

$$\mathbb{P}\left(\max_{1 \leq i \leq n} S_i > 2r\sqrt{n}\right) \leq \frac{\mathbb{P}(|S_n| > r\sqrt{n})}{1 - \frac{1}{r^2}}$$

(Ottaviani's inequality)

(12) $$\frac{1}{h}\mathbb{P}\Big(\sup_{s\le h}|X_n(s)|\ge\varepsilon\Big)=\frac{1}{h}\mathbb{P}\Big(\max_{1\le i\le nh}|S_i|\ge\varepsilon\sqrt{n}\Big)$$

$$\le\frac{1}{h}\,\frac{\mathbb{P}\big(S_{\lfloor nh\rfloor}\ge\frac{\varepsilon\sqrt{n}}{2}\big)}{1-4h/\varepsilon}$$

$$=\frac{1}{h}\,\frac{\mathbb{P}\Big(\frac{S_{\lfloor nh\rfloor}}{\sqrt{nh}}\ge\frac{\varepsilon}{2\sqrt{h}}\Big)}{1-4\frac{h}{\varepsilon}}$$

$$\xrightarrow[n\to\infty]{}\frac{1}{h}\,\frac{2\Big(1-\Phi\big(\frac{\varepsilon}{2\sqrt{h}}\big)\Big)}{1-4\frac{h}{\varepsilon}}\quad\text{by CLT}$$

By iid ~~assumpt~~ property of $\xi_1,\xi_2,\dots$
same estimate works for all $j=1,\dots,\lfloor\frac{1}{h}\rfloor$

From $$1-\Phi(x)<\frac{e^{-x^2/2}}{x\sqrt{2\pi}}\qquad x>0$$

follows that
$$\frac{1}{h}\mathbb{P}\Big(\sup_{0\le s\le h}|X_n(s)|\ge\varepsilon\Big)\xrightarrow[n\to\infty]{}0$$

Thus (X_n) is tight.

Theorem $\quad X_n\xrightarrow{d}B$

Also called functional CLT.

(13) <u>Corollary</u> $\max_{1\le j\le n} \frac{S_j}{\sqrt{n}} \xrightarrow{d} \max_{t\in[0,1]} B(t)$

That is

$$\mathbb{P}\left(\max_{1\le j\le n} \frac{S_j}{\sqrt{n}} \ge x\right) \to \sqrt{\frac{2}{\pi}} \int_x^\infty \frac{e^{-y^2/2}}{\sqrt{2\pi}}\, dy$$

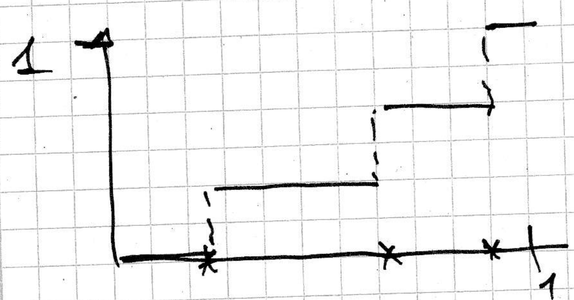
$$= 2\,(1-\Phi(x))$$

BM is continuous process. For convergence to discontinuous processes one needs larger space of functions.

<u>5.6</u> Functional limit for empirical distribution function.

$\xi_1, \xi_2, \dots$ iid Uniform $[0,1]$

$$F_n(t) := \frac{\#\{j\le n : \xi_j < t\}}{n}$$ empirical distr. function



$n=3$

E.g. $F_n(1/2)$ = proportion of points $\le 1/2$

(14) For every $t \in [0,1]$, $n F_n(t) \sim \text{Binomial}(n,t)$

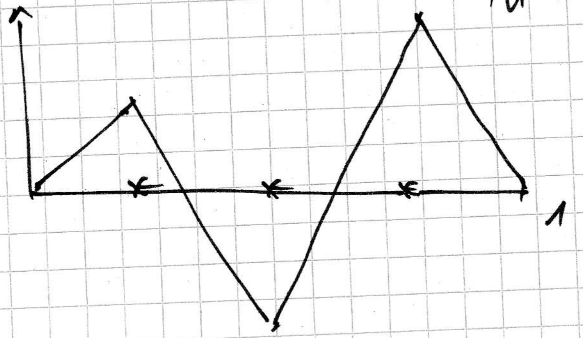
Glivenko-Cantelli Theorem:

$$\sup_{t\in[0,1]} |F_n(t) - t| \xrightarrow[n\to\infty]{} 0 \quad \text{a.s.}$$

$F(t) = t$ is Uniform $[0,1]$ c.d.f.

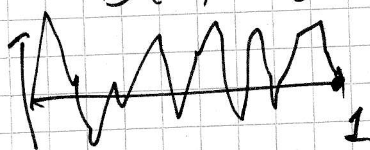
Let $(Y_n(t),\ t \in [0,1])$ be interpolation of

$$\left(\xi_j,\ \sqrt{n}\,(F_n(\xi_j) - \xi_j)\right)$$



Theorem $\quad Y_n \xrightarrow{d} B^\circ$

B° - Brownian Bridge $\quad B(t) - t\,B(1)$



$K = \max_{t\in[0,1]} |B^\circ(t)|$ has Kolmogorov-Smirnov distribution

Hence $\max_{t\in[0,1]} \left|\sqrt{n}\,(F_n(t) - t)\right| \xrightarrow{d} K$,

where

$$\mathbb{P}(K \le y) = \sum_{j=-\infty}^{\infty} (-1)^j e^{-2j^2y^2}, \quad y > 0$$

5.7 The arcsine law

Distribution of the ~~...~~ time within $[0,1]$ that the BM spends on positive side

$$\mathbb{P}\left(\int_0^1 \mathbb{1}(B(t)>0)\,dt \le x \right) = \frac{2}{\pi} \arcsin \sqrt{x}$$

p.d.f. $\quad x \mapsto \dfrac{1}{\pi\sqrt{x(1-x)}}$ $\qquad$ Beta $(1/2, 1/2)$

$x \in (0,1)$

$(S_n, n \ge 0)$ mean-zero, variance-one RW

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(S_i > 0) \le x \right) = \frac{2}{\pi} \arcsin \sqrt{x}$$

$x \in (0,1)$