

1 Donsker's Theorem for $(C[0, 1], \|\cdot\|_\infty)$

Proposition 1.1

- Let $\xi_1, \xi_2, \dots : \Omega \longrightarrow \mathbb{R}$ be a sequence of independent and identically distributed \mathbb{R} -valued random variables defined on the probability space $(\Omega, \mathcal{A}, \mu)$, with expectation value zero and common finite variance $\sigma^2 > 0$.
- Define the random variables:

$$\begin{cases} S_0 & : \Omega \longrightarrow \mathbb{R} : \omega \longmapsto 0, & \text{and} \\ S_n & : \Omega \longrightarrow \mathbb{R} : \omega \longmapsto \sum_{i=1}^n \xi_i(\omega), & \text{for each } n \in \mathbb{N}. \end{cases}$$

- For each $n \in \mathbb{N}$, define $X^{(n)} : \Omega \longrightarrow C[0, 1]$ as follows:

$$X^{(n)}(\omega)(t) := \frac{1}{\sigma \cdot \sqrt{n}} \left\{ S_{i-1}(\omega) + n \left(t - \frac{i-1}{n} \right) \xi_i(\omega) \right\}, \text{ for each } \omega \in \Omega, t \in \left[\frac{i-1}{n}, \frac{i}{n} \right], i = 1, 2, 3, \dots, n.$$

- For each $n \in \mathbb{N}$ and each $t \in [0, 1]$, define $X_t^{(n)} : \Omega \longrightarrow \mathbb{R}$ as follows:

$$X_t^{(n)}(\omega) := X^{(n)}(\omega)(t), \text{ for each } \omega \in \Omega.$$

Then, the following statements are true:

- (i) For each $\omega \in \Omega$ and each $n \in \mathbb{N}$,

$$X^{(n)}(\omega) \left(\frac{i}{n} \right) = \frac{1}{\sigma \cdot \sqrt{n}} \cdot S_i(\omega), \text{ for } i = 0, 1, 2, \dots, n.$$

- (ii) For each $\omega \in \Omega$ and each $n \in \mathbb{N}$,

$$X^{(n)}(\omega)(t) \text{ is the linear interpolation from } \frac{1}{\sigma \cdot \sqrt{n}} S_{i-1}(\omega) \text{ to } \frac{1}{\sigma \cdot \sqrt{n}} S_i(\omega) \text{ over } t \in \left[\frac{i-1}{n}, \frac{i}{n} \right],$$

where $i = 1, 2, \dots, n$.

- (iii) For any $0 \leq t_0 < t_1 < t_2 < \dots < t_k \leq 1$,

$$\left(X_{t_1}^{(n)} - X_{t_0}^{(n)}, \dots, X_{t_k}^{(n)} - X_{t_{k-1}}^{(n)} \right) \xrightarrow{d} N \left(\mu = \mathbf{0}, \Sigma = \text{diag}(t_1 - t_0, \dots, t_k - t_{k-1}) \right), \text{ as } n \longrightarrow \infty.$$

- (iv) For any $0 \leq t_1, t_2, \dots, t_k \leq 1$,

$$\left(X_{t_1}^{(n)}, X_{t_2}^{(n)}, \dots, X_{t_k}^{(n)} \right) \xrightarrow{d} N \left(\mu = \mathbf{0}, \Sigma = \left[\min\{t_i, t_j\} \right]_{1 \leq i, j \leq k} \right), \text{ as } n \longrightarrow \infty.$$

PROOF

Donsker's Theorems (Functional Central Limit Theorems)

Study Notes

November 1, 2015

Kenneth Chu

- (i) Obvious.
- (ii) Obvious.
- (iii) First, note that, for each $\omega \in \Omega$, $n \in \mathbb{N}$, and $t \in [0, 1]$, we have

$$X_t^{(n)}(\omega) = \frac{1}{\sigma \cdot \sqrt{n}} \left\{ S_{[nt]}(\omega) + (nt - [nt]) \cdot \xi_{[nt]+1}(\omega) \right\},$$

where $\lfloor \cdot \rfloor : \mathbb{R} \rightarrow \mathbb{Z}$, defined by

$$\lfloor x \rfloor := \max \left\{ k \in \mathbb{Z} \mid k \leq x \right\}, \quad \text{for each } x \in \mathbb{R},$$

is the round-down function. We next state three Claims, whose proofs will be given below. We note that the desired conclusion follows readily from Claim 3 and the Cramér-Wold Theorem (Theorem 1.9(iii), p.56, [4]); hence the present proof is complete once we establish the three Claims below.

Claim 1: If $\{a_n\}_{n \in \mathbb{N}}$ is a sequence of non-negative integers and $\{b_n\}_{n \in \mathbb{N}} \subset \mathbb{N}$ a sequence of positive integers satisfying:

$$a_n < b_n, \text{ for sufficiently large } n \in \mathbb{N}, \quad \text{and} \quad \lim_{n \rightarrow \infty} \frac{b_n - a_n}{n} = c > 0,$$

then

$$\frac{1}{\sigma \cdot \sqrt{n}} \cdot \sum_{i=1+a_n}^{b_n} \xi_i \xrightarrow{d} \sqrt{c} \cdot Z, \quad \text{where } Z \sim N(0, 1).$$

Claim 2: For each fixed $t \in [0, 1]$,

$$W(t)_n := \frac{1}{\sigma \cdot \sqrt{n}} \cdot (nt - [nt]) \cdot \xi_{[nt]+1} \xrightarrow{d} 0.$$

Claim 3: For $0 \leq t_0 < t_1 < t_2 < \dots < t_k \leq 1$, and arbitrary $c_1, c_2, \dots, c_k \in \mathbb{R}$,

$$\sum_{i=1}^k c_i \left(X_{t_i}^{(n)} - X_{t_{i-1}}^{(n)} \right) \xrightarrow{d} N \left(0, \sum_{i=1}^k c_i^2 (t_i - t_{i-1}) \right), \quad \text{as } n \rightarrow \infty.$$

Proof of Claim 1: Note that, for sufficiently large $n \in \mathbb{N}$, we may write

$$\frac{1}{\sigma \cdot \sqrt{n}} \cdot \sum_{i=1+a_n}^{b_n} \xi_i = \frac{\sqrt{b_n - a_n}}{\sqrt{n}} \cdot \left(\frac{1}{\sigma \cdot \sqrt{b_n - a_n}} \cdot \sum_{i=1+a_n}^{b_n} \xi_i \right).$$

Since, by hypothesis, that

$$\lim_{n \rightarrow \infty} \frac{b_n - a_n}{n} = c > 0,$$

Claim 1 follows by Slutsky's Theorem (Example 6, p.40, [3]), once we establish the following:

$$\frac{1}{\sigma \cdot \sqrt{b_n - a_n}} \cdot \sum_{i=1+a_n}^{b_n} \xi_i \xrightarrow{d} N(0, 1), \quad \text{as } n \rightarrow \infty.$$

Donsker's Theorems (Functional Central Limit Theorems)

We establish the above convergence by invoking the Lindeberg Central Limit Theorem (Theorem 1.15, §1.5.5, p.67, [4]). In the present context, the Lindeberg Condition is the following:

$$\lim_{n \rightarrow \infty} \frac{1}{B_n^2} \cdot E \left[\sum_{i=1+a_n}^{b_n} \xi_i^2 \cdot I_{\{|\xi_i| \geq \varepsilon S_n\}} \right] = 0, \quad \text{for each } \varepsilon > 0,$$

where

$$B_n^2 := \text{Var} \left[\sum_{i=1+a_n}^{b_n} \xi_i \right] = (b_n - a_n) \sigma^2 > 0.$$

The last equality used the hypothesis that ξ_1, ξ_2, \dots are independent and identically distributed with common finite variance $0 < \sigma^2 < \infty$. Hence, for each $\varepsilon > 0$,

$$\begin{aligned} \frac{1}{B_n^2} \cdot E \left[\sum_{i=1+a_n}^{b_n} \xi_i^2 \cdot I_{\{|\xi_i| \geq \varepsilon B_n\}} \right] &= \frac{1}{(b_n - a_n) \sigma^2} \cdot (b_n - a_n) \cdot E \left[\xi_1^2 \cdot I_{\{|\xi_1| \geq \varepsilon \sigma \sqrt{b_n - a_n}\}} \right] \\ &= \frac{1}{\sigma^2} \cdot E \left[\xi_1^2 \cdot I_{\{|\xi_1|/\sigma \geq \sqrt{b_n - a_n}\}} \right] \rightarrow 0, \quad \text{as } n \rightarrow \infty, \end{aligned}$$

since $\lim_{n \rightarrow \infty} \sqrt{b_n - a_n} = \infty$ and $\sigma^2 = E[\xi_1^2]$ is finite. This verifies that the Lindeberg Condition indeed holds in the present context, and completes the proof of Claim 1.

Proof of Claim 2: First, note that $E[W(t)_n] = 0$. We now argue that $W(t)_n \xrightarrow{p} 0$. To this end, let $\varepsilon > 0$ be given. Then,

$$\begin{aligned} \varepsilon^2 \cdot P(|W(t)_n| \geq \varepsilon) &\leq E[W(t)_n^2 \cdot I_{\{|W(t)_n| \geq \varepsilon\}}] \\ &\leq E[W(t)_n^2] = \text{Var}(W(t)_n) = \text{Var} \left[\frac{1}{\sigma \cdot \sqrt{n}} \cdot (nt - \lfloor nt \rfloor) \cdot \xi_{\lfloor nt \rfloor + 1} \right] \\ &= \frac{1}{n \cdot \sigma^2} \cdot (nt - \lfloor nt \rfloor)^2 \cdot \text{Var}(\xi_{\lfloor nt \rfloor + 1}) = \frac{1}{n \cdot \sigma^2} \cdot (nt - \lfloor nt \rfloor)^2 \cdot \sigma^2 \\ &\leq \frac{1}{n}, \end{aligned}$$

which implies

$$\lim_{n \rightarrow \infty} P(|W(t)_n| \geq \varepsilon) = 0, \quad \text{for each } \varepsilon > 0,$$

i.e. $W(t)_n \xrightarrow{p} 0$, as $n \rightarrow \infty$ (Definition 2, Chapter 1, [3]), which is equivalent to $W(t)_n \xrightarrow{d} 0$, as $n \rightarrow \infty$ (by Theorem 1, Chapter 1 and Theorem 2, Chapter 2, [3]). This proves Claim 2.

Proof of Claim 3: Let $0 \leq t_0 < t_1 < t_2 < \dots < t_k \leq 1$, and $c_1, c_2, \dots, c_k \in \mathbb{R}$ be arbitrary. Observe that:

$$\begin{aligned} &\sum_{i=1}^k c_i \left(X_{t_i}^{(n)} - X_{t_{i-1}}^{(n)} \right) \\ &= \sum_{i=1}^k \frac{c_i}{\sigma \cdot \sqrt{n}} \left\{ S_{\lfloor nt_i \rfloor} - S_{\lfloor nt_{i-1} \rfloor} \right\} + \sum_{i=1}^k \frac{c_i}{\sigma \cdot \sqrt{n}} \left\{ (nt_i - \lfloor nt_i \rfloor) \cdot \xi_{\lfloor nt_i \rfloor + 1} - (nt_{i-1} - \lfloor nt_{i-1} \rfloor) \cdot \xi_{\lfloor nt_{i-1} \rfloor + 1} \right\} \\ &= \sum_{i=1}^k \frac{c_i}{\sigma \cdot \sqrt{n}} \left\{ \sum_{j=1+\lfloor nt_{i-1} \rfloor}^{\lfloor nt_i \rfloor} \xi_j \right\} + \sum_{i=1}^k c_i \left\{ W(t_i)_n - W(t_{i-1})_n \right\} \\ &= \sum_{i=1}^k c_i Y_i^{(n)} + \sum_{i=1}^k c_i \left\{ W(t_i)_n - W(t_{i-1})_n \right\} \end{aligned}$$

Donsker's Theorems (Functional Central Limit Theorems)

By Claim 2 and Slutsky's Theorem (Corollary, p.40, [3]),

$$\sum_{i=1}^k c_i \left\{ W(t_i)_n - W(t_{i-1})_n \right\} \xrightarrow{d} 0, \text{ as } n \rightarrow \infty. \quad (1.1)$$

Next, note that since $\xi_1, \xi_2, \xi_3, \dots$ are independent, we see that, for each fixed $n \in \mathbb{N}$,

$$Y_i^{(n)} := \frac{1}{\sigma \cdot \sqrt{n}} \cdot \sum_{j=1+\lfloor nt_{i-1} \rfloor}^{\lfloor nt_i \rfloor} \xi_j, \quad i = 1, 2, 3, \dots, k,$$

are independent. Now, since $0 \leq t_{i-1} < t_i \leq 1$, it follows that $\lfloor nt_{i-1} \rfloor < \lfloor nt_i \rfloor$ for sufficiently large $n \in \mathbb{N}$. In addition,

$$\begin{aligned} \frac{\lfloor nt_i \rfloor - \lfloor nt_{i-1} \rfloor}{n} &= \frac{\lfloor nt_i \rfloor}{n} - \frac{\lfloor nt_{i-1} \rfloor}{n} = \left(\frac{nt_i}{n} + \frac{\lfloor nt_i \rfloor - nt_i}{n} \right) - \left(\frac{nt_{i-1}}{n} + \frac{\lfloor nt_{i-1} \rfloor - nt_{i-1}}{n} \right) \\ &= t_i - t_{i-1} + \frac{\lfloor nt_i \rfloor - nt_i}{n} - \frac{\lfloor nt_{i-1} \rfloor - nt_{i-1}}{n}, \end{aligned}$$

which implies

$$\left| \frac{\lfloor nt_i \rfloor - \lfloor nt_{i-1} \rfloor}{n} - (t_i - t_{i-1}) \right| = \left| \frac{\lfloor nt_i \rfloor - nt_i}{n} - \frac{\lfloor nt_{i-1} \rfloor - nt_{i-1}}{n} \right| \leq \frac{2}{n} \rightarrow 0, \text{ as } n \rightarrow \infty.$$

Thus,

$$\lim_{n \rightarrow \infty} \frac{\lfloor nt_i \rfloor - \lfloor nt_{i-1} \rfloor}{n} = t_i - t_{i-1} > 0.$$

Thus, by Claim 1, we see that, for each $i = 1, 2, \dots, k$,

$$Y_i^{(n)} := \frac{1}{\sigma \cdot \sqrt{n}} \cdot \sum_{j=1+\lfloor nt_{i-1} \rfloor}^{\lfloor nt_i \rfloor} \xi_j \xrightarrow{d} \sqrt{t_i - t_{i-1}} \cdot N(0, 1) = N\left(0, t_i - t_{i-1}\right), \text{ as } n \rightarrow \infty. \quad (1.2)$$

By (1.1), (1.2), Proposition A.1, and Slutsky's Theorem (Corollary, p.40, [3]), we now see that

$$\sum_{i=1}^k c_i \left(X_{t_i}^{(n)} - X_{t_{i-1}}^{(n)} \right) = \sum_{i=1}^k c_i Y_i^{(n)} + \sum_{i=1}^k c_i \left\{ W(t_i)_n - W(t_{i-1})_n \right\} \xrightarrow{d} N\left(0, \sum_{i=1}^k c_i^2 (t_i - t_{i-1})\right).$$

This completes the proof of Claim 3.

(iv) Let $t_0 := 0$, hence, $X_{t_0}^{(n)} \equiv 0$ for each $n \in \mathbb{N}$. We thus have, for each $n \in \mathbb{N}$,

$$\begin{pmatrix} X_{t_1}^{(n)} \\ X_{t_2}^{(n)} \\ \vdots \\ X_{t_k}^{(n)} \end{pmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & \cdots & \cdots & 0 & 0 \\ 1 & 1 & 0 & \cdots & \cdots & 0 & 0 \\ 1 & 1 & 1 & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 & 0 \\ 1 & 1 & 1 & \cdots & \cdots & 1 & 0 \\ 1 & 1 & 1 & \cdots & \cdots & 1 & 1 \end{bmatrix}}_T \cdot \begin{pmatrix} X_{t_1}^{(n)} \\ X_{t_2}^{(n)} - X_{t_1}^{(n)} \\ \vdots \\ X_{t_k}^{(n)} - X_{t_{k-1}}^{(n)} \end{pmatrix}.$$

Donsker's Theorems (Functional Central Limit Theorems)

By (iii), we know that

$$\begin{pmatrix} X_{t_1}^{(n)} \\ X_{t_2}^{(n)} - X_{t_1}^{(n)} \\ \vdots \\ X_{t_k}^{(n)} - X_{t_{k-1}}^{(n)} \end{pmatrix} \xrightarrow{d} \begin{pmatrix} Z_{t_1} \\ Z_{t_2-t_1} \\ \vdots \\ Z_{t_k-t_{k-1}} \end{pmatrix} \sim N\left(\mu = \mathbf{0}, \Sigma = \text{diag}(t_1, t_2 - t_1, \dots, t_k - t_{k-1})\right), \text{ as } n \rightarrow \infty.$$

Since the map $\mathbb{R}^k \rightarrow \mathbb{R}^k : x \mapsto T \cdot x$ is continuous, we see immediately by Slutsky's Theorem (Theorem 6(a), p.39, [3]) that

$$\begin{pmatrix} X_{t_1}^{(n)} \\ X_{t_2}^{(n)} \\ \vdots \\ X_{t_k}^{(n)} \end{pmatrix} \xrightarrow{d} T \cdot \begin{pmatrix} Z_{t_1} \\ Z_{t_2-t_1} \\ \vdots \\ Z_{t_k-t_{k-1}} \end{pmatrix}, \text{ as } n \rightarrow \infty.$$

Since the map $\mathbb{R}^k \rightarrow \mathbb{R}^k : x \mapsto T \cdot x$ is an invertible linear automorphism on \mathbb{R}^k , we see that

$$L = \begin{pmatrix} L_{t_1} \\ L_{t_2} \\ \vdots \\ L_{t_k} \end{pmatrix} := T \cdot \begin{pmatrix} Z_{t_1} \\ Z_{t_2-t_1} \\ \vdots \\ Z_{t_k-t_{k-1}} \end{pmatrix}$$

is still an \mathbb{R}^k -valued Gaussian random variable, and it clearly has expectation value $\mathbf{0} \in \mathbb{R}^k$, since each of $Z_{t_1}, Z_{t_2-t_1}, \dots, Z_{t_k-t_{k-1}}$ has expectation value $0 \in \mathbb{R}$. It remains only to compute the covariance matrix of the \mathbb{R}^k -valued Gaussian random variable L . To this end, consider $1 \leq i \leq j \leq k$, i.e. $t_i \leq t_j$. Then, using the alternative notation $Z_{t_i-t_0} := Z_{t_i}$, we have

$$\begin{aligned} \text{Cov}(L_{t_i}, L_{t_j}) &= \text{Cov}(Z_{t_1} + Z_{t_2-t_1} + \dots + Z_{t_i-t_{i-1}}, Z_{t_1} + Z_{t_2-t_1} + \dots + Z_{t_j-t_{j-1}}) \\ &= \text{Cov}\left(\sum_{a=1}^i Z_{t_a-t_{a-1}}, \sum_{b=1}^j Z_{t_b-t_{b-1}}\right) = \sum_{a=1}^i \sum_{b=1}^j \text{Cov}(Z_{t_a-t_{a-1}}, Z_{t_b-t_{b-1}}) \\ &= \sum_{a=1}^i \text{Cov}(Z_{t_a-t_{a-1}}, Z_{t_a-t_{a-1}}) = \sum_{a=1}^i \text{Var}(Z_{t_a-t_{a-1}}) = \sum_{a=1}^i (t_a - t_{a-1}) \\ &= (t_1 - t_0) + (t_2 - t_1) + \dots + (t_{i-1} - t_{i-2}) + (t_i - t_{i-1}) \\ &= t_i = \min\{t_i, t_j\}, \end{aligned}$$

as required. □

A Technical Lemmas

Note that $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{d} Y$ does NOT in general imply $X_n + Y_n \xrightarrow{d} X + Y$. But the implication does hold if X_n and Y_n are independent for each $n \in \mathbb{N}$, and both X and Y are Gaussian random variables, as the following Proposition shows.

Proposition A.1 *Let $k \in \mathbb{N}$ be fixed. Suppose:*

- For each $n \in \mathbb{N}$,

$$Y_1^{(n)}, Y_2^{(n)}, \dots, Y_k^{(n)} : \Omega^{(n)} \longrightarrow \mathbb{R}$$

are independent \mathbb{R} -valued random variables defined on the probability space $\Omega^{(n)}$.

- For each $i = 1, 2, \dots, k$,

$$Y_i^{(n)} \xrightarrow{d} N(\mu_i, \sigma_i^2), \quad \text{as } n \longrightarrow \infty.$$

Then, for any $c_1, c_2, \dots, c_k \in \mathbb{R}$,

$$\sum_{i=1}^k c_i Y_i^{(n)} \xrightarrow{d} N\left(\sum_{i=1}^k c_i \mu_i, \sum_{i=1}^k c_i^2 \sigma_i^2\right), \quad \text{as } n \longrightarrow \infty.$$

PROOF Let $Y^{(n)} := \sum_{i=1}^k c_i Y_i^{(n)}$. Let φ_X denote the characteristic function of a \mathbb{R} -valued random variable X . Then,

$$\begin{aligned} \varphi_{Y^{(n)}}(t) &= \varphi_{\sum_{i=1}^k c_i Y_i^{(n)}}(t) \\ &= \prod_{i=1}^k \varphi_{c_i Y_i^{(n)}}(t), \quad \text{since } Y_1^{(n)}, \dots, Y_k^{(n)} \text{ are independent} \\ &= \prod_{i=1}^k \varphi_{Y_i^{(n)}}(c_i t) \\ &\longrightarrow \prod_{i=1}^k \exp\left\{\sqrt{-1} \mu_i (c_i t) - \frac{1}{2} \sigma_i^2 (c_i t)^2\right\} \\ &= \exp\left\{\sqrt{-1} \left(\sum_{i=1}^k c_i \mu_i\right) t - \frac{1}{2} \left(\sum_{i=1}^k c_i^2 \sigma_i^2\right) t^2\right\}, \quad \text{as } n \longrightarrow \infty, \end{aligned}$$

where the second and third equalities follow from the properties of characteristic functions of random variables (see p.21, [3]), while the expression of the limit follows from the fact that the characteristic function φ_Z of a random variable Z with distribution $N(\mu, \sigma^2)$ is

$$\varphi_Z = \exp\left\{\sqrt{-1} \mu t - \frac{1}{2} \sigma^2 t^2\right\}.$$

The Proposition now follows immediately from the Lévy-Cramér Continuity Theorem (Theorem 1.9(ii), p.56, [4]). \square

References

- [1] BILLINGSLEY, P. *Probability and Measure*, third ed. John Wiley & Sons, 1995.
- [2] BILLINGSLEY, P. *Convergence of Probability Measures*, second ed. John Wiley & Sons, 1999.
- [3] FERGUSON, T. S. *A Course in Large Sample Theory*, first ed. Texts in Statistical Science. CRC Press, 1996.
- [4] SHAO, J. *Mathematical Statistics*, second ed. Springer, 2003.