# Some Asymptotic Theory for the Bootstrap

Study Notes July 4, 2015 Kenneth Chu

This set of notes aims to provide complete proofs of a number of asymptotic results regarding the Bootstrap [3] contained in Bickel and Freedman [1].

## 1 Bootstrap asymptotics for sample mean

**Theorem 1.1 (Bootstrap Central Limit Theorem for I.I.D. sample mean, Theorem 2.1 [1])**
*Let $(\Omega, \mathcal{A}, \nu)$ be a probability space. Let $X, X_1, X_2, \ldots : \Omega \longrightarrow \mathbb{R}$ be a sequence of independent and identically distributed $\mathbb{R}$-valued random variables defined on $\Omega$ with finite expectation value $\mu_X \in \mathbb{R}$ and variance $\sigma_X^2 < \infty$. For each $n \in \mathbb{N}$, define:*

$$\overline{X}_n \;:\; \Omega \;\longrightarrow\; \mathbb{R} \;:\; \omega \;\longmapsto\; \frac{1}{n}\sum_{i=1}^{n} X_i(\omega).$$

*For $n, m \in \mathbb{N}$, define $\mathcal{S}_m^{(n)}$ to be the set of all functions from $\{1,2,\ldots,m\} \longrightarrow \{1,2,\ldots,n\}$. Thus, each*

$$s \;=\; (s(1), s(2), \ldots, s(m)) \;\in\; \mathcal{S}_m^{(n)}$$

*can be regarded as a length-m finite (ordered) sequence of positive integers between 1 and n, inclusive. Note that $\mathcal{S}_m^{(n)}$ is a finite set with $\mid \mathcal{S}_m^{(n)} \mid = n^m$. Endow $\mathcal{S}_m^{(n)}$ with the probability space structure induced by the uniform probability function:*

$$P_{\mathcal{S}_m^{(n)}}(s) \;:=\; \frac{1}{n^m}, \quad \text{for each } s \in \mathcal{S}_m^{(n)}.$$

*Let $\Omega \times \mathcal{S}_m^{(n)}$ be the product probability space of $\Omega$ and $\mathcal{S}_m^{(n)}$. Define:*

$$\overline{X}_m^{(n)} \;:\; \Omega \times \mathcal{S}_m^{(n)} \;\longrightarrow\; \mathbb{R} \;:\; (\omega, s) \;\longmapsto\; \frac{1}{m}\sum_{j=1}^{m} X_{s(j)}(\omega).$$

*For each $\omega \in \Omega$, define:*

$$\Phi_{m,\omega}^{(n)} \;:\; \mathcal{S}_m^{(n)} \;\longrightarrow\; \mathbb{R} \;:\; s \;\longmapsto\; \sqrt{m}\left(\overline{X}_m^{(n)}(\omega, s) - \overline{X}_n(\omega)\right)$$

*Then,*

$$P\Big(\,\Phi_{m,\omega}^{(n)} \xrightarrow{d} N(0,\sigma_X^2), \text{ as } n,m \to \infty\,\Big) \;=\; \nu\Big(\Big\{\,\omega \in \Omega \;\Big|\; \Phi_{m,\omega}^{(n)} \xrightarrow{d} N(0,\sigma_X^2), \text{ as } n,m \to \infty\,\Big\}\Big) \;=\; 1.$$

**Remark 1.2**
*For each fixed $\omega \in \Omega$, $\left\{\,\Phi_{m,\omega}^{(n)} : \mathcal{S}_m^{(n)} \longrightarrow \mathbb{R}\,\right\}_{n,m\in\mathbb{N}}$ is a doubly indexed sequence of $\mathbb{R}$-valued random variables. Note that their respective domains $\mathcal{S}_m^{(n)}$ are pairwise distinct probability spaces. The* ***Bootstrap Central Limit Theorem for I.I.D. sample mean*** *asserts that for almost every $\omega \in \Omega$, the doubly indexed sequence $\left\{\,\Phi_{m,\omega}^{(n)}\,\right\}$ of $\mathbb{R}$-valued random variables converges in distribution to $N(0,\sigma_X^2)$ as $n, m \longrightarrow \infty$.*

**Remark 1.3** *The following results are well known from classical asymptotic theory:*
*By the* ***Weak Law of Large Numbers****, $\overline{X}_n$ converges in probability to $\mu_X$, as $n \longrightarrow \infty$; in other words,*

$$\lim_{n\to\infty} P\big(\,\big|\,\overline{X}_n - \mu_X\,\big| > \varepsilon\,\big) \;=\; \lim_{n\to\infty} \nu\left(\left\{\,\omega \in \Omega \;:\; \big|\,\overline{X}_n(\omega) - \mu_X\,\big| > \varepsilon\,\right\}\right) \;=\; 0, \quad \text{for each } \varepsilon > 0.$$

*By the* ***Strong Law of Large Numbers****, $\overline{X}_n$ converges almost surely to $\mu_X$, as $n \longrightarrow \infty$; in other words,*

$$P\Big(\lim_{n\to\infty} \overline{X}_n = \mu_X\Big) \;=\; \nu\left(\Big\{\,\omega \in \Omega \;\Big|\; \lim_{n\to\infty} \overline{X}_n(\omega) = \mu_X\,\Big\}\right) \;=\; 1.$$

*By the* ***Central Limit Theorem****, $\sqrt{n}\left(\overline{X}_n - \mu_X\right)$ converges in distribution to $N(0,\sigma_X^2)$.*

PROOF Let $\mathcal{M}_1(\mathbb{R},\mathcal{B}(\mathbb{R}))$ denotes the collection of probability measures on $(\mathbb{R},\mathcal{B}(\mathbb{R}))$. Define

$$\Gamma_2 \;:=\; \left\{\; G \in \mathcal{M}_1(\mathbb{R},\mathcal{B}(\mathbb{R})) \;\;\middle|\;\; \int_{\mathbb{R}} x^2\,\mathrm{d}G(x) < \infty \;\right\}.$$

Define the **Wasserstein metric** on $\Gamma_2$:

$$d_2 \,:\, \Gamma_2 \times \Gamma_2 \longrightarrow \mathbb{R} \,:\, (G,G') \longmapsto \inf\left\{\; \sqrt{E[\rho(X,Y)^2]} \;\;\middle|\;\; (X,Y) \in C(G,G') \;\right\}$$

**Claim 1:** $d_2$ is indeed a metric on $\Gamma_2$.

**Claim 2:** For $G, G_1, G_2, \ldots \in \Gamma_2$,

$$G_n \overset{d_2}{\longrightarrow} G \qquad \text{if and only if} \qquad G_n \longrightarrow G \text{ weakly} \quad \text{and} \quad \int_{\mathbb{R}} x^2\,\mathrm{d}G_n(x) \longrightarrow \int_{\mathbb{R}} x^2\,\mathrm{d}G(x)$$

**Claim 3:** For $G \in \Gamma_2$ and $m \in \mathbb{N}$, let $G^{(m)}$ be the $m$**-fold empirical measure** of $G$, i.e. $G^{(m)}$ is the (empirical) measure of the random variable

$$S_m^{(G)} \;\;:=\;\; \frac{1}{m^{1/2}} \sum_{i=1}^{m} \left( Z_i^{(G)} - \mu_G \right),$$

where $\mu_G := \int_{\mathbb{R}} x\,\mathrm{d}G(x)$ is the expectation value of the measure $G$, and $Z_1^{(G)}, Z_2^{(G)}, \ldots, Z_m^{(G)}$ are independent and identically distributed random variables with distribution $G$. Then, for any $G, H \in \Gamma_2$, we have

$$d_2\Big(G^{(m)}, H^{(m)}\Big) \;\;\leq\;\; d_2(\,G, H\,)$$

**Claim 4:**

$$\nu\Big(\Big\{\; \omega \in \Omega \;\;\Big|\;\; F_n(\omega) \overset{\text{w}}{\longrightarrow} F \;\Big\}\Big) \;\;=\;\; 1$$

Claim 4 follows from the Glivenko-Cantelli Theorem, which states that:

$$\nu\left(\left\{\; \omega \in \Omega \;\;\middle|\;\; \lim_{n\to\infty} \sup_{t\in\mathbb{R}} |F_n(\omega)(t) - F(t)| \,=\, 0 \;\right\}\right) \;\;=\;\; 1,$$

which implies trivially

$$\nu\Big(\Big\{\; \omega \in \Omega \;\;\Big|\;\; \lim_{n\to\infty} F_n(\omega)(t) = F(t), \;\text{ for each } t \in \mathbb{R} \;\Big\}\Big) \;\;=\;\; 1,$$

which, in turn, is equivalent to Claim 4.

**Claim 5:**

$$\nu\left(\left\{\; \omega \in \Omega \;\;\middle|\;\; \int_{\mathbb{R}} x^2\,\mathrm{d}F_n(\omega)(x) \longrightarrow \int_{\mathbb{R}} x^2\,\mathrm{d}F(x) \;\right\}\right) \;\;=\;\; 1$$

By the Strong Law of Large Numbers, we have

$$\int_{\mathbb{R}} x^2\,\mathrm{d}F_n(\omega)(x) \;\;=\;\; \frac{1}{n}\sum_{i=1}^{n} X_i(\omega)^2 \;\;\overset{\text{a.e.}}{\longrightarrow}\;\; E\left[\,X^2\,\right] \;\;=\;\; \int_{\mathbb{R}} x^2\,\mathrm{d}F(x)$$

**Claim 6:**

$$\nu\Big(\Big\{\ \omega \in \Omega\ \Big|\ F_n(\omega) \overset{d_2}{\longrightarrow} F\ \Big\}\Big) \;=\; \nu(\{\ \omega \in \Omega\ |\ d_2(F_n(\omega), F\,) \longrightarrow 0\ \}) \;=\; 1$$

Immediate by Claims 2, 4, and 5.

Let $\omega \in \Omega$ be fixed.

$$\begin{aligned} d_2\Big(F_n^{(m)}(\omega), N(0,\sigma_X^2)\Big) &\leq d_2\Big(F_n^{(m)}(\omega), F^{(m)}\Big) + d_2\Big(F^{(m)}, N(0,\sigma_X^2)\Big) \\ &\leq d_2(\,F_n(\omega), F\,) + d_2\Big(F^{(m)}, N(0,\sigma_X^2)\Big) \end{aligned}$$

Now, $d_2\big(F^{(m)}, N(0,\sigma_X^2))\big) \longrightarrow 0$ by the classical Central Limit Theorem.

$$\begin{array}{rccll} & d_2(\,F_n(\omega), F\,) + d_2\big(F^{(m)}, N(0,\sigma_X^2)\big) & \longrightarrow & 0, & \text{as } n, m \longrightarrow \infty \\ \Longrightarrow & d_2\Big(F_n^{(m)}(\omega), N(0,\sigma_X^2)\Big) & \longrightarrow & 0, & \text{as } n, m \longrightarrow \infty \\ \Longrightarrow & F_n^{(m)}(\omega) & \overset{\text{w}}{\longrightarrow} & N(0,\sigma_X^2), & \text{as } n, m \longrightarrow \infty \end{array}$$

□

# A A stochastic process $\{\,X_t : \Omega \longrightarrow V\,\}_{t\in T}$ and its equivalent $V^T$-valued random variable $X : \Omega \longrightarrow V^T$

Let $\Omega$, $T$, and $V$ be non-empty sets. Let $\{\,X_t : \Omega \longrightarrow V\,\}_{t\in T}$ be a $T$-index family of maps, each of which maps from $\Omega$ into $V$. Note that this family of maps is set-theoretically equivalent (in the sense that either one completely determines the other) to the following $(V^T)$-valued map defined on $\Omega$:

$$X : \Omega \longrightarrow V^T : \omega \longmapsto (t \longmapsto X_t(\omega)),$$

where $V^T = \prod_{t\in T} V$ denotes the set of all (arbitrary) $V$-valued functions defined on $T$. In this section, we aim to establish the following two results:

- Suppose $(\Omega, \mathcal{A})$ and $(V, \mathcal{F})$ are measurable space structures on $\Omega$ and $V$, respectively. Then, $X : \Omega \longrightarrow V^T$ is $\big(\mathcal{A}, \sigma[(V,\mathcal{F})^T]\big)$-measurable if and only if $X_t : \Omega \longrightarrow V$ is $(\mathcal{A}, \mathcal{F})$-measurable for each $t \in T$. Here, $\sigma[(V,\mathcal{F})^T]$ denotes the product $\sigma$-algebra on $V^T$, which is by definition the smallest $\sigma$-algebra on $V^T$ such that, for each $t \in T$, the projection map (or evaluation map)

  $$\pi_t : V^T \longrightarrow V : x \longmapsto x(t)$$

  is $\big(\sigma[(V,\mathcal{F})^T], \mathcal{F}\big)$-measurable.

- An immediate corollary of the above result is that: Suppose $(\Omega, \mathcal{A}, \mu)$ is a probability space structure on $\Omega$, $(V, \mathcal{F})$ is a measurable space structure on $V$, and $\sigma[(V,\mathcal{F})^T]$ is the product $\sigma$-algebra on $V^T$. Then, $X : (\Omega, \mathcal{A}, \mu) \longrightarrow (V^T, \sigma[(V,\mathcal{F})^T])$ is $V^T$-valued random variable if and only if $\{\,X_t : (\Omega, \mathcal{A}, \mu) \longrightarrow (V, \mathcal{F})\,\}_{t\in T}$ is a stochastic process.

**Definition A.1 (The product $\sigma$-algebra of a Cartesian product of measurable spaces)**
*Let $T$ be an arbitrary non-empty set. For each $t \in T$, let $(V_t, \mathcal{F}_t)$ be a measurable space (in particular, $V_t \neq \varnothing$). Let $\prod_{t\in T} V_t$ be the Cartesian product of $\{ V_t \}_{t\in T}$. In other words,*

$$\prod_{t\in T} V_t \;:=\; \left\{ \; v : T \longrightarrow \bigsqcup_{t\in T} V_t \;\middle|\; v(t) \in V_t, \text{ for each } t \in T \; \right\}.$$

*That $\prod_{t\in T} V_t \neq \varnothing$ follows from the Axiom of Choice. For each $t \in T$, let*

$$\pi_t \;:\; \prod_{\tau\in T} V_\tau \;\longrightarrow\; V_t \;:\; v \;\longmapsto\; v(t)$$

*be the projection map from $\prod_{\tau\in T} V_\tau$ onto $V_t$. The **product $\sigma$-algebra** on $\prod_{t\in T} V_t$ is the following:*

$$\sigma\!\left(\left\{\; \pi_t^{-1}(F) \;\subset\; \prod_{\tau\in T} V_\tau \;\middle|\; F \in \mathcal{F}_t\,,\, t \in T \;\right\}\right) \;\subset\; \mathrm{PowerSet}\!\left(\prod_{t\in T} V_t\right).$$

*Clearly, it is the smallest $\sigma$-algebra on $\prod_{t\in T} V_t$ with respect to which each projection map $\pi_t : \prod_{t\in T} V_t \longrightarrow (V_t, \mathcal{F}_t)$ is measurable. We denote the product $\sigma$-algebra on $\prod_{t\in T} V_t$ by*

$$\sigma\!\left(\prod_{t\in T} (V_t, \mathcal{F}_t)\right).$$

**Theorem A.2**
*Suppose $\Omega$, $T$, and $V$ are non-empty sets. Let $\{ X_t : \Omega \longrightarrow V \}_{t\in T}$ be a $T$-indexed family of $V$-valued maps defined on $\Omega$. Then, the following statements are true:*

1. *The family $\{ X_t : \Omega \longrightarrow V \}_{t\in T}$ of maps is set-theoretically equivalent (in the sense that either completely determines the other) to the following $(V^T)$-valued map defined on $\Omega$:*

$$X \;:\; \Omega \;\longrightarrow\; V^T \;:\; \omega \;\longmapsto\; (t \longmapsto X_t(\omega))\,,$$

   *where $V^T = \prod_{t\in T} V$ denotes the set of all (arbitrary) $V$-valued functions defined on $T$.*

2. *Suppose:*
   - *$(\Omega, \mathcal{A})$ and $(V, \mathcal{F})$ are measurable space structures on $\Omega$ and $V$, respectively.*
   - *$W \subset V^T$ is a subset of $V^T$ such that $X(\Omega) = \bigcup_{t\in T} X_t(\Omega) \subset W$.*
   - *$(W, \mathcal{G})$ is a measurable space structure on $W$ such that, for each $t \in T$, the projection map*

$$\pi_t \;:\; W \;\longrightarrow\; V \;:\; w \;\longmapsto\; w(t)$$

     *is $(\mathcal{G}, \mathcal{F})$-measurable.*

   *Then, $(\mathcal{A}, \mathcal{G})$-measurability of $X : \Omega \longrightarrow W$ implies $(\mathcal{A}, \mathcal{F})$-measurability of $X_t : \Omega \longrightarrow V$ for each $t \in T$.*

3. *Suppose:*
   - *$(\Omega, \mathcal{A})$ and $(V, \mathcal{F})$ are measurable space structures on $\Omega$ and $V$, respectively.*
   - *$\sigma[(V, \mathcal{F})^T]$ is the product $\sigma$-algebra on $V^T = \prod_{t\in T} V$ generated by the collection of projection maps*

$$\left\{\; \pi_t \;:\; V^T = \prod_{\tau\in T} V \;\longrightarrow\; V \;:\; w \;\longmapsto\; w(t) \;\right\}_{t\in T}.$$

*Then,* $X : \Omega \longrightarrow V^T$ *is* $\left(\mathcal{A}, \sigma[(V,\mathcal{F})^T]\right)$*-measurable if and only if* $X_t : \Omega \longrightarrow V$ *is* $(\mathcal{A},\mathcal{F})$*-measurable for each* $t \in T$.

PROOF

1. The proof of this result is routine and we omit it.
2. Suppose $X : \Omega \longrightarrow W$ is $(\mathcal{A},\mathcal{G})$-measurable. Note that $X_t = \pi_t \circ X$, where

$$\pi_t \,:\, V^T = \prod_{t\in T} V \,\longrightarrow\, V \,:\, v \,\longrightarrow\, v(t)$$

is the projection from $V^T = \prod_{\tau\in T} V$ onto the $t$-th factor. By hypothesis, $\pi_t : W \longrightarrow V$ is $(\mathcal{G},\mathcal{F})$-measurable for each $t \in T$. This implies, for each $t \in T$, $X_t = \pi_t \circ X$ is $(\mathcal{A},\mathcal{F})$-measurable, being a composition of two measurable maps.

3. Since, for each $t \in T$, the projection map $\pi_t : V^T \longrightarrow V$ is $\left(\sigma[(V,\mathcal{F})^T], \mathcal{F}\right)$-measurable (by construction of the $\sigma$-algebra $\sigma[(V,\mathcal{F})^T]$ on $V^T$), the preceding result immediately implies the following implication:

$$(\mathcal{A}, \sigma[(V,\mathcal{F})^T])\text{-measurability of } X : \Omega \longrightarrow V^T \quad\Longrightarrow\quad (\mathcal{A},\mathcal{F})\text{-measurability of } X_t : \Omega \longrightarrow V, \text{ for each } t \in T.$$

Conversely, suppose $X_t$ is $(\mathcal{A},\mathcal{F})$-measurable for each $t \in T$. Recall that the product $\sigma$-algebra on $V^T$ is generated by sets of the form:

$$\pi_t^{-1}(F)\,, \quad \text{for some } t \in T \text{ and } F \in \mathcal{F}.$$

It follows that, for each $t \in T$ and each $F \in \mathcal{F}$, we have

$$X^{-1}\big(\pi_t^{-1}(F)\big) \;=\; (X^{-1}\circ\pi_t^{-1})(F) \;=\; (\pi_t \circ X)^{-1}(F) \;=\; X_t^{-1}(F) \;\subset\; \Omega$$

is $\mathcal{A}$-measurable, since $X_t : (\Omega,\mathcal{A}) \longrightarrow (V,\mathcal{F})$ is $(\mathcal{A},\mathcal{F})$-measurable by hypothesis. This proves that $X : \Omega \longrightarrow V^T$ is $(\mathcal{A}, \sigma[(V,\mathcal{F})^T])$-measurable. □

**Definition A.3 (Stochastic processes)**
*A **stochastic process** is a family, indexed by some non-empty set* $T$,

$$\{\; X_t \,:\, (\Omega,\mathcal{A},\mu) \,\longrightarrow\, (V,\mathcal{F}) \;\}_{t\in T}$$

*of* $(\mathcal{A},\mathcal{F})$*-measurable maps, where the common domain* $(\Omega,\mathcal{A},\mu)$ *is a probability space and the common codomain* $(V,\mathcal{F})$ *is a measurable space. The common codomain* $(V,\mathcal{F})$ *is called the **state space** of the stochastic process.*

**Corollary A.4**
*Suppose:*

- $(\Omega,\mathcal{A},\mu)$ *is a probability space and* $(V,\mathcal{F})$ *is a measurable space.*
- $T$ *is a non-empty set and* $W \subset V^T = \prod_{t\in T} V$.
- $(W,\mathcal{G})$ *is a measurable space structure on* $W$ *such that, for each* $t \in T$*, the projection map*

$$\pi_t \,:\, W \,\longrightarrow\, V \,:\, w \,\longmapsto\, w(t)$$

*is* $(\mathcal{G},\mathcal{F})$*-measurable.*

*If* $X : (\Omega, \mathcal{A}, \mu) \longrightarrow \left(V^T, \sigma[(V,\mathcal{F})^T]\right)$ *is a* $V^T$*-valued random variable (i.e.* $X$ *is* $\left(\mathcal{A}, \sigma[(V,\mathcal{F})^T]\right)$*-measurable), then its set-theoretically equivalent* $T$*-indexed family of* $V$*-valued maps defined on* $\Omega$

$$\left\{ \begin{array}{ccccc} X_t & : & (\Omega, \mathcal{A}, \mu) & \longrightarrow & (V, \mathcal{F}) \\ & & \omega & \longmapsto & (\pi_t \circ X)(\omega) = \pi_t(X(\omega)) = X(\omega)(t) \end{array} \right\}_{t \in T}$$

*is a stochastic process (i.e.* $X_t$ *is* $(\mathcal{A}, \mathcal{F})$*-measurable for each* $t \in T$*).*

**Corollary A.5**
*Suppose:*

- $T$*,* $\Omega$*,* $V$ *are non-empty sets.*
- $(\Omega, \mathcal{A}, \mu)$ *is a probability space structure on* $\Omega$*,* $(V, \mathcal{F})$ *is a measurable space structure on* $V$*.*
- $\sigma[(V,\mathcal{F})^T]$ *denotes the corresponding product* $\sigma$*-algebra on* $V^T = \prod_{t\in T} V$*.*

*Let* $\{ X_t : \Omega \longrightarrow V \}_{t \in T}$ *be a* $T$*-indexed family of* $V$*-valued maps defined on* $\Omega$*, and let*

$$X : \Omega \longrightarrow V^T : \omega \longmapsto (t \longmapsto X_t(\omega))$$

*be its set-theoretically equivalent* $(V^T)$*-valued map defined on* $\Omega$*. Then,*

$$\{ X_t : (\Omega, \mathcal{A}, \mu) \longrightarrow (V, \mathcal{F}) \}_{t \in T}$$

*is a stochastic process if and only if*

$$X : (\Omega, \mathcal{A}, \mu) \longrightarrow \left(V^T, \sigma[(V,\mathcal{F})^T]\right)$$

*is a* $(V^T)$*-valued random variable.*

# B Uniqueness of the "full distribution" of a stochastic process $\{ X_t : \Omega \longrightarrow V \}_{t \in T}$ given its finite-dimensional distributions

**Definition B.1 (Finite-dimensional distributions of a stochastic process)**
*Let* $\{ X_t : (\Omega, \mathcal{A}, \mu) \longrightarrow (V, \mathcal{F}) \}_{t \in T}$ *be a stochastic process. Let* $n \in \mathbb{N}$ *and* $t_1, t_2, \ldots, t_n \in T$ *be distinct elements of* $T$*. Let* $\mathcal{P}_{(X_{t_1}, \ldots, X_{t_n})} \in \mathcal{M}_1(V^n, \mathcal{F}^{\otimes n})$ *denote the probability measure induced on the product measurable space* $(V^n, \mathcal{F}^{\otimes n})$ *by the random variable*

$$(X_{t_1}, X_{t_2}, \ldots, X_{t_n}) : (\Omega, \mathcal{A}, \mu) \longrightarrow \left(V^n, \mathcal{F}^{\otimes n}\right)$$

$\mathcal{P}_{(X_{t_1}, \ldots, X_{t_n})}$ *is called a* ***finite-dimensional distribution*** *of the stochastic process.*

**Theorem B.2**
*Let* $(V, \mathcal{F})$ *be a measurable space, and* $\sigma\left[(V,\mathcal{F})^T\right]$ *the product* $\sigma$*-algebra on* $V^T = \prod_{t\in T} V$*. Let*

$$\{ X_t : (\Omega_X, \mathcal{A}_X, \mu_X) \longrightarrow (V, \mathcal{F}) \}_{t \in T} \quad \text{and} \quad \{ Y_t : (\Omega_Y, \mathcal{A}_Y, \mu_Y) \longrightarrow (V, \mathcal{F}) \}_{t \in T}$$

*be two stochastic processes with the same index set* $T$ *and the same state space* $(V, \mathcal{F})$*. Let*

$$X : (\Omega_X, \mathcal{A}_X, \mu_X) \longrightarrow \left(V^T, \sigma\left[(V,\mathcal{F})^T\right]\right) \quad \text{and} \quad Y : (\Omega_Y, \mathcal{A}_Y, \mu_Y) \longrightarrow \left(V^T, \sigma\left[(V,\mathcal{F})^T\right]\right)$$

*be their respective* $\left(V^T, \sigma\left[(V,\mathcal{F})^T\right]\right)$*-valued random variables. Let* $\mathcal{P}_X$*,* $\mathcal{P}_Y \in \mathcal{M}_1\left(V^T, \sigma\left[(V,\mathcal{F})^T\right]\right)$ *be the probability measures induced on* $\left(V^T, \sigma\left[(V,\mathcal{F})^T\right]\right)$ *by* $X$ *and* $Y$*, respectively. Then,*

$$\mathcal{P}_X = \mathcal{P}_Y \in \mathcal{M}_1\left(V^T, \sigma\left[(V,\mathcal{F})^T\right]\right)$$

*if and only if*

$$\mathcal{P}_{(X_{t_1},X_{t_2},\ldots,X_{t_n})} = \mathcal{P}_{(Y_{t_1},Y_{t_2},\ldots,Y_{t_n})} \in \mathcal{M}_1\big(V^n,\mathcal{F}^{\otimes n}\big), \text{ for each } n\in\mathbb{N} \text{ and pairwise distinct } t_1,t_2,\ldots,t_n\in T.$$

Proof

□

# C Existence of a stochastic process given its finite-dimensional distributions: Komolgorov's Existence Theorem

**Definition C.1 (Stochastic processes)**
*Suppose $(\Omega,\mathcal{A},\mu)$ is a probability space, $(V,\mathcal{F})$ is a measurable space, and $T$ is an arbitrary non-empty set. A* ***stochastic process*** *indexed by $T$ defined on $\Omega$ with codomain $V$ is a family $\{\,X_t:\Omega\longrightarrow V\,\}_{t\in T}$ indexed by $T$ of $V$-valued random variables defined on $\Omega$.*

**Definition C.2 (Finite-dimensional distributions of a stochastic processes)**
*Let $\{\,X_t:\Omega\longrightarrow(V,\mathcal{F})\,\}_{t\in T}$ be a stochastic process. Let $n\in\mathbb{N}$ and $t_1,t_2,\ldots,t_n\in T$ be distinct elements of $T$. The probability distribution induced on the product measurable space $(V^n,\mathcal{F}^{\otimes n})$ by $(X_{t_1},X_{t_2},\ldots,X_{t_n}):\Omega\longrightarrow V^n$ is called a* ***finite-dimensional distribution*** *of the stochastic process.*

**Definition C.3 (Komolgorov systems of finite-dimensional distributions & Komolgorov consistency)**
*Let $T$ be an arbitrary non-empty set, and $\mathcal{D}(T)$ the set of all finite ordered sequences of elements of $T$ whose elements are pairwise distinct; in other words,*

$$\mathcal{D}(T) := \left\{ (t_1,t_2,\ldots,t_n)\in\bigcup_{k=1}^{\infty}T^k \;\middle|\; n\in\mathbb{N},\ t_i\neq t_j,\ \text{whenever } i\neq j \right\}.$$

*For each $n\in\mathbb{N}$, let $\mathcal{M}_1(R^n,\mathcal{B}(\mathbb{R}^n))$ be the set of all probability measures defined on the product measurable space $(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n))$. A* ***Komolgorov system of finite-dimensional distributions*** *is a $\mathcal{D}(T)$-indexed family $\mathcal{P}$ of probability measures of the following form:*

$$\mathcal{P} = \left\{\, P_{(t_1,\ldots,t_n)}\in\mathcal{M}_1(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n)) \;\middle|\; (t_1,\ldots,t_n)\in\mathcal{D}(T)\,\right\}.$$

*Furthermore, $\mathcal{P}$ is said to be* ***Komolgorov consistent*** *if it satisfies both of the following conditions:*

- ***permutation invariance:*** *For any $n\in\mathbb{N}$, any $(t_1,\ldots,t_n)\in\mathcal{D}(T)$, any $B_1,\ldots,B_n\in\mathcal{B}(\mathbb{R})$, and any permutation $\pi:\{1,\ldots,n\}\longrightarrow\{1,\ldots,n\}$, the following equality holds:*

$$P_{(t_1,\ldots,t_n)}(B_1\times\cdots\times B_n) = P_{(t_{\pi(1)},\ldots,t_{\pi(n)})}\big(B_{\pi(1)}\times\cdots\times B_{\pi(n)}\big).$$

- ***projection invariance:*** *For any $n\in\mathbb{N}$, any $(t_1,\ldots,t_{n+1})\in\mathcal{D}(T)$, and any $B_1,\ldots,B_n\in\mathcal{B}(\mathbb{R})$, the following equality holds:*

$$P_{(t_1,\ldots,t_n,t_{n+1})}(B_1\times\cdots\times B_n\times\mathbb{R}) = P_{(t_1,\ldots,t_n)}(B_1\times\cdots\times B_n).$$

**Remark C.4**
*It is obvious that the collection of finite-dimensional distributions of any $\mathbb{R}$-valued stochastic process is a Komolgorov consistent Komolgorov system of finite-dimensional distributions.*

**Definition C.5**
*Let $\{\,X_t:\Omega\longrightarrow\mathbb{R}\,\}_{t\in T}$ be an $\mathbb{R}$-valued stochastic process, and*

$$\mathcal{P} = \left\{\, P_{(t_1,\ldots,t_n)}\in\mathcal{M}_1(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n)) \;\middle|\; (t_1,\ldots,t_n)\in\mathcal{D}(T)\,\right\}$$

*be a Komolgorov system of finite-dimensional distributions. We say that* ***the stochastic process*** $\{X_t\}$ ***admits*** $\mathcal{P}$ ***as its collection of finite-dimensional distributions*** *if, for each* $n \in \mathbb{N}$ *and any* $(t_1, t_2, \ldots, t_n) \in \mathcal{D}(T)$*, the probability distribution induced on* $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ *by the map*

$$(X_{t_1}, \ldots, X_{t_n}) \,:\, \Omega \longrightarrow \mathbb{R}^n$$

*equals* $P_{(t_1,\ldots,t_n)} \in \mathcal{P}$.

**Theorem C.6 (Komolgorov's Existence Theorem, Theorem 36.2, [2])**
*Let*

$$\mathcal{P} \;=\; \left\{\; P_{(t_1,\ldots,t_n)} \in \mathcal{M}_1(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)) \;\middle|\; (t_1, \ldots, t_n) \in \mathcal{D}(T) \;\right\}.$$

*be a Komolgorov system of finite-dimensional distributions. Then, there exists a stochastic process*

$$\{\, X_t : (\Omega, \mathcal{A}, \mu) \;\longrightarrow\; (\mathbb{R}, \mathcal{B}(\mathbb{R}))\,\}_{t\in T}$$

*which admits* $\mathcal{P}$ *as its collection of finite-dimensional distributions if and only if* $\mathcal{P}$ *is Komolgorov consistent.*

# D Gaussian Processes

**Definition D.1 (Gaussian processes)**
*An* $\mathbb{R}$*-valued stochastic process* $\{\, X_t : \Omega \longrightarrow \mathbb{R}\,\}_{t\in T}$ *is said to be* ***Gaussian*** *if each of its finite-dimensional distribution is a (univariate or multivariate) Gaussian distribution.*

**Definition D.2 (Mean and covariance functions of $\mathbb{R}$-valued stochastic processes)**
*Let* $\{\, X_t : \Omega \longrightarrow \mathbb{R}\,\}_{t\in T}$ *be an* $\mathbb{R}$*-valued stochastic process.*

- *If, for each* $t \in T$*, we have* $E(\,X_t\,) \in \mathbb{R}$*, then the function*

$$a_X \,:\, T \;\longrightarrow\; \mathbb{R} \,:\, t \;\longmapsto\; E(\,X_t\,)$$

  *is called the* ***mean*** *function of the* $\mathbb{R}$*-valued stochastic process* $\{\,X_t\,\}$.

- *In addition, if, for each* $t_1, t_2 \in T$*, we have* $0 \le \mathrm{Cov}(X_{t_1}, X_{t_2}) < \infty$*, then the function*

$$\Sigma_X \,:\, T \times T \;\longrightarrow\; \mathbb{R} \,:\, (t_1, t_2) \;\longmapsto\; \mathrm{Cov}(X_{t_1}, X_{t_2})$$

  *is called the* ***covariance*** *function of the* $\mathbb{R}$*-valued stochastic process* $\{\,X_t\,\}$.

**Theorem D.3**
*Let* $T$ *be an arbitrary non-empty set,* $a : T \longrightarrow \mathbb{R}$ *an arbitrary* $\mathbb{R}$*-valued function defined on* $T$*, and* $\Sigma : T \times T \longrightarrow [0, \infty)$ *a non-negative* $\mathbb{R}$*-valued function defined on* $T \times T$*. Then, there exists a Gaussian process whose mean and covariance functions are* $a$ *and* $\Sigma$*, respectively.*

**Theorem D.4**
*The mean and covariance functions of a Gaussian process together completely determine its collection of finite-dimensional distributions.*

**Definition D.5 (Brownian motion, a.k.a. Wiener process)**
*A* ***Brownian motion****, or* ***Wiener process****, is a stochastic process* $\{\, W_t : (\Omega, \mathcal{A}, \mu) \longrightarrow \mathbb{R}\,\}_{t\ge 0}$ *indexed by the non-negative real line satisfying the following conditions:*

- *At* $t = 0$*, the process takes value* $0$ *with probability* $1$*; more precisely:*

$$P(\,W_0 = 0\,) \;=\; \mu(\{\,\omega \in \Omega \mid W_0(\omega) = 0\,\}) \;=\; 1.$$

- *The process* $\{ W_t \}$ *has independent increments; more precisely: for any* $0 \leq t_1 \leq t_2 \leq \cdots \leq t_n < \infty$,

$$W_{t_n} - W_{t_{n-1}}, \quad W_{t_{n-1}} - W_{t_{n-2}}, \quad \ldots \;, \quad W_{t_2} - W_{t_1} \;:\; \Omega \longrightarrow \mathbb{R}$$

  *are independent random variables.*

- *For* $0 \leq t_1 < t_2 < \infty$, *the increment* $W_{t_2} - W_{t_1}$ *follows a Gaussian distribution with mean* $0$ *and variance* $t_2 - t_1$.

**Definition D.6 (Brownian bridge)**
*A* ***Brownian bridge*** *is a Gaussian process* $\{ W^\circ_t : (\Omega, \mathcal{A}, \mu) \longrightarrow \mathbb{R} \}_{t\in[0,1]}$ *indexed by the closed unit interval in* $\mathbb{R}$ *satisfying the following conditions:*

- *For each* $t \in [0, 1]$, *we have* $E( W^\circ_t ) = 0$.
- *For any* $t_1, t_2 \in [0, 1]$, *we have* $\mathrm{Cov}\big(W^0_{t_1}, W^\circ_{t_2}\big) = \min\{t_1, t_2\} - t_1 t_2$.

# References

[1] Bickel, P. J., and Freedman, D. A. Some asymptotic theory for the bootsrap. *The Annals of Statistics 9*, 6 (1981), 1196–1217.

[2] Billingsley, P. *Probability and Measure*, third ed. John Wiley & Sons, 1995.

[3] Efron, B. Bootsrap methods: another look at the jackknife. *The Annals of Statistics 7*, 1 (1979), 1–26.