Appendix A

Bojan Basrak
University of Zagreb

APPENDIX A :

CONVERGENCE IN PROBABILITY & DISTRIBUTION

$\mathbb{X}_n = (X_{n,1}, \ldots, X_{n,k})$ sequence of random vectors in $\mathbb{R}^k$

Recall :

$$\mathbb{X}_n \xrightarrow{d} \mathbb{X}$$

means $P(\mathbb{X}_n \le x) \to P(\mathbb{X} \le x)$

for all $x \in \mathbb{R}^k$ s.t. x is not discontin. point of the function

$$x \mapsto P(\mathbb{X} \le x).$$

Recall

$$
\begin{array}{ccccc}
\xrightarrow{a.s.} & & & & \xrightarrow{L^p} \\
& \searrow\!\!\searrow & & \swarrow\!\!\swarrow & \\
& & \xrightarrow{P} & & \\
& & \Downarrow & & \\
& & \xrightarrow{d} & &
\end{array}
$$

THEOREM 1 (Portmanteau)

The following are equivalent

i) $X_n \xrightarrow{d} X$

ii) $Ef(X_n) \to Ef(X)$ for all bound. cont. $f: \mathbb{R}^k \to \mathbb{R}$

iii) $P(X_n \in B) \to P(X \in B)$ for all Borel sets B s.t. $P(X \in \partial B) = 0$ where $\partial B = \overline{B} \setminus \operatorname{Int} B$

⋮

THEOREM 2 (continuous mapping thm)

Suppose $g : \mathbb{R}^k \to \mathbb{R}^m$ is measurable & continuous on C s.t. $P(X \in C) = 1$

$$X_n \xrightarrow{\cdot} X \Rightarrow g(X_n) \xrightarrow{\cdot} g(X)$$

where $\xrightarrow{\cdot}$ can stand for:

$\xrightarrow{d}$, $\xrightarrow{P}$, $\xrightarrow{a.s.}$.

$\boxed{\text{DEF}}$ A family of random vectors $\{X_\alpha : \alpha \in T\}$ is tight if $\forall \varepsilon > 0 \ \exists M = M(\varepsilon)$ s.t.

$$\sup_\alpha P(\|X_\alpha\| > M) < \varepsilon$$

We also say that this family is stochastically bounded.

EXE 1> Any finite family of r.v.'s is tight.

EXAMPLE 1

Suppose (X_n) satisfy $E\|X_n\| < C \quad \forall n$

then

$$P(\|X_n\| > M) \le \frac{E\|X_n\|}{M} \le \frac{C}{M} \quad \forall n.$$

$\Rightarrow$ the family (X_n) is tight.

EXE<> Show $E\|X_n\|^p < C \quad \forall n$ & some $p > 0 \Rightarrow (X_n)$ is tight

THEOREM 3 (Prokhorov)

i) If $X_n \xrightarrow{d} X$ for some X, then (X_n) is tight

ii) If (X_n) is tight, then there is a r. vector X & subsequence (n_k) s.t.

$$X_{n_k} \xrightarrow{d} X.$$

EXE 3) Suppose (X_n) is a sequence of random variables s.t. $|X_n| < 1$ a.s.

$\Rightarrow$ there is a subseq. n_k & r.v. X s.t.

$$X_{n_k} \to X$$

THEOREM 4

i) $X_n \xrightarrow{d} c \iff X_n \xrightarrow{P} c$ for a constant c

ii) $X_n \xrightarrow{d} X$ & $\|X_n - Y_n\| \xrightarrow{P} 0 \Rightarrow Y_n \xrightarrow{d} X$

iii) $X_n \xrightarrow{d} X$ & $Y_n \xrightarrow{P} c \Rightarrow (X_n, Y_n) \xrightarrow{d} (X, c)$

iv) $X_n \xrightarrow{P} X$ & $Y_n \xrightarrow{P} Y \Rightarrow (X_n, Y_n) \xrightarrow{P} (X, Y)$

Thm 4 iii) $\Rightarrow$ a very useful lemma

LEMMA 1 (Slutsky)

Suppose $X_n \xrightarrow{d} X$, $Y_n \xrightarrow{d} c$, then

i) $X_n + Y_n \xrightarrow{d} X + c$

ii) $Y_n \cdot X_n \xrightarrow{d} c \cdot X$

iii) for $c \neq 0$, $Y_n \in \mathbb{R}$

$X_n / Y_n \xrightarrow{d} X / c$.

EXAMPLE 2

Suppose estimators T_n, S_n satisfy

$$\sqrt{n}(T_n - \vartheta) \xrightarrow{d} N(0, \sigma^2); \quad S_n^2 \xrightarrow{P} \sigma^2$$

for some ϑ, $\sigma^2 > 0$ then

$$\sqrt{n}\, \frac{T_n - \vartheta}{S_n} \xrightarrow{d} N(0,1)$$

$\Downarrow$

$(1-\alpha)\,100\%$ – confidence interval for the parameter ϑ is

$$\left(T_n - \frac{S_n}{\sqrt{n}} z_{\alpha/2},\ T_n + \frac{S_n}{\sqrt{n}} z_{\alpha/2}\right)$$

THEOREM 5 (Lévy)

i) $X_n \xrightarrow{d} X \iff E e^{it'X_n} \to E e^{it'X}$ for all $t \in \mathbb{R}^k$

ii) If $E e^{it'X_n} \to \varphi(t)$ for $t \in \mathbb{R}^k$ & φ is contin. at $0 \Rightarrow \exists$ r. vector X st.

$$X_n \xrightarrow{d} X \quad \& \quad \varphi(t) = E e^{it'X}.$$

COROLLARY 1 (Cramer – Wold device)

$$\mathbb{X}_n \xrightarrow{d} \mathbb{X} \iff t'\mathbb{X}_n \to t'\mathbb{X} \quad \forall t \in \mathbb{R}^k$$

Suppose

$$r_n (T_n - \vartheta) \xrightarrow{d} T$$

φ is a function from $\mathbb{R}^k$ to $\mathbb{R}^m$ differentiable at ϑ with differential $D\varphi_\vartheta$;

what can we say about the limiting behavior of

$$r_n \left(\varphi(T_n) - \varphi(\vartheta) \right) \quad ?$$

THEOREM 6 (Delta method)

For random vectors $(T_n), T$ in $\mathbb{R}^k$ & a real sequence (r_n), $r_n \to \infty$ suppose

$$r_n(T_n - \vartheta) \xrightarrow{d} T, \qquad \vartheta \in \mathbb{R}^k$$

If $f : \mathbb{R}^k \to \mathbb{R}^m$ is differentiable at ϑ, then

$$r_n\big(f(T_n) - f(\vartheta)\big) \xrightarrow{d} Df_\vartheta(T)$$

EXAMPLE 3

Suppose $X_n \overset{iid}{\sim} \mathrm{Exp}(\lambda)$, $\lambda > 0$

by the CLT

$$\sqrt{n}\left(\overline{X}_n - \frac{1}{\lambda}\right) \xrightarrow{d} N\left(0, \frac{1}{\lambda^2}\right) \sim X$$

take now $\varphi(x) = x^{-1}$, which is clearly differ. at $1/\lambda$ & $\varphi'(x) = -1/x^2$

$\Downarrow$

$$\sqrt{n}\left(\frac{1}{\overline{X}_n} - \lambda\right) \xrightarrow{d} \underbrace{\varphi'\left(\frac{1}{\lambda}\right)}_{= -\lambda^2} X \sim N(0, \lambda^2)$$

EXAMPLE 4

More generally if

$$\sqrt{n}\,(T_n - \vartheta) \xrightarrow{d} N_k(\mu, \Sigma)$$

$$\Downarrow$$

$$\sqrt{n}\,\big(\varphi(T_n) - \varphi(\vartheta)\big) \longrightarrow N_m\big(\varphi'_\vartheta \mu,\ \varphi'_\vartheta \Sigma \varphi'^{\,T}_\vartheta\big)$$

THEOREM 7 (Lévy C.L.T.)

X_i iid, $\sigma^2 = \operatorname{Var} X_i < \infty$ & $\mu = E X_i$

$$\Rightarrow \quad \sqrt{n}\left(\overline{X}_n - \mu\right) \xrightarrow{d} N(0, \sigma^2)$$

THEOREM 8 (Lindeberg C.L.T.)

Suppose $Y_{n,1}, \ldots, Y_{n,n}$ are indep. for each n & have covariance matrices s.t.

$$\frac{1}{n}\sum_{i=1}^{n} \mathrm{Cov}\, Y_{n,i} \to \Sigma$$

& $$\frac{1}{n}\sum_{1}^{n} E\left(\|Y_{n,i}\|^2 \mathbf{1}_{\|Y_{n,i}\| > \varepsilon\sqrt{n}}\right) \to 0$$

$\forall \varepsilon > 0$

Then

$$\frac{1}{\sqrt{n}}\left(\sum_{1}^{n} Y_{n,i} - EY_{n,i}\right) \xrightarrow{d} N(0, \Sigma)$$

To estimate autocov. & autocorr fuction on the mean of a stationary sequence (X_n) we need

$\boxed{\text{DEF}}$ The sample mean

$$\overline{X}_n = \frac{1}{n} \sum_{1}^{n} X_j ,$$

the sample autocovariance function $\hat{\gamma}$:

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{j=1}^{n-h} (X_{j+h} - \overline{X}_n)(X_j - \overline{X}_n)$$

$$0 \le h \le n-1$$

& the sample autocorrelation function $\hat{\rho}$:

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)} \qquad 0 \le h \le n-1$$

Clearly

$$E\overline{X}_n = \mu \qquad \text{(if } (X_i) \text{ is weakly stationary)}$$

$\Downarrow$

the sample mean is unbiased estimator

$$\mathrm{Var}(\sqrt{n}\,\overline{X}_n) = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \mathrm{Cov}(X_i, X_j)$$

$$= \frac{1}{n} \sum_{h=-n}^{n} (n-|h|)\, \gamma_X(h)$$

$$= \sum_{h=-n}^{n} \frac{n-|h|}{n}\, \gamma_X(h)$$

Thus if $\sum \gamma_X(h) < \infty$, the domin. conv. thm $\Rightarrow$

$$\mathrm{Var}(\sqrt{n}\,\overline{X}_n) \longrightarrow \sum_{-\infty}^{\infty} \gamma_X(h)$$

$\Downarrow$ Čebišev. ineq.

$\overline{X}_n \longrightarrow \mu$ in P & L_2 ; note: $\sqrt{n}(\overline{X}_n - \mu)$ is tight !!!

$\boxed{\text{DEF}}$ Time series $(X_t)_{t\in\mathbb{Z}}$ is m-dependent if for all $t \in \mathbb{Z}$ families of r.v.'s

$$(\ldots, X_{t-1}, X_t) \ \& \ (X_{t+m+1}, X_{t+m+2}, \ldots)$$

are independent

EXE 4) If (Z_t) is iid, and $X_t = Z_t + \vartheta Z_{t-1}$ is MA(1) process, show that it is 1-dependent.

EXE 5) Show: (X_t) is 0-dependent $\Leftrightarrow$ X_t are independent r.v.'s

LEMMA 2 (extends Slutsky)

Suppose

i) $Y_{nj} \xrightarrow{d} Y_j \qquad n \to \infty$

ii) $Y_j \xrightarrow{d} Y \qquad j \to \infty$

iii) $\lim_{j\to\infty} \limsup_{n\to\infty} P(|X_n - Y_{nj}| > \varepsilon) = 0 \qquad \forall \varepsilon > 0$

Then $\quad X_n \xrightarrow{d} Y$.

THEOREM 9 (C.L.T. for m-depend. seq.)

Suppose (X_t) is strongly stationary m-dependent time series with mean zero & finite variance.

Then

$$\sqrt{n}\,\overline{X}_n \xrightarrow{d} N\Big(0, \sum_{-m}^{m} \gamma_X(h)\Big)$$

The Large Sample Theory

III LARGE SAMPLE THEORY

(FOR THE ESTIMATORS OF μ, α, γ & ρ)

EXAMPLE 1 (MA(1) process)

Suppose $Z_i \overset{iid}{\sim} (0, \sigma^2)$ & $X_t = Z_t + \vartheta Z_{t-1}$,

$\vartheta \in \mathbb{R} \Rightarrow (X_t)$ is strictly stationary

& 1-dependent

We showed (see Ch 1) that

$$\gamma_X(h) = \begin{cases} \sigma^2(1+\vartheta^2) & h = 0 \\ \sigma^2 \cdot \vartheta & h = \pm 1 \\ 0 & \text{otherwise} \end{cases}$$

By the c.l.t. for m-depend. sequences

$$\sqrt{n}\,\overline{X}_n \longrightarrow N\Big(0, \sum_{-1}^{1} \gamma_X(h)\Big)$$

$$= N\big(0, \sigma^2 (1+\vartheta^2+2\vartheta)\big)$$

In particular, for $\vartheta = -1$

$$\sqrt{n}\,\overline{X}_n \xrightarrow{d} N(0,0) \equiv 0$$

(check this directly).

A natural extension of $MA(1)$ process is $MA(q)$ process

$\boxed{\text{DEF}}$ Weakly stationary sequence (X_t) is called MA(q) process (moving average of order q) if

$$X_t = Z_t + \vartheta_1 Z_{t-1} + \cdots + \vartheta_q Z_{t-q}$$

for a W.N. sequence $Z_t \sim WN(0, \sigma^2)$ & some real parameters $\vartheta_1, \ldots, \vartheta_q$,

(we typically ask $\vartheta_q \neq 0$)

EXE 1) Show that MA(q) process (X_t) has expectation 0 & autocov. function

$$\gamma_X(h) = \begin{cases} \sigma^2 \cdot \sum_{j=0}^{q-|h|} \vartheta_j \cdot \vartheta_{j+|h|} & \text{for } |h| \le q \\ 0 & \text{for } |h| > q \end{cases}$$

(here we write $\vartheta_0 = 1$)

EXE2) Apply Thm 9, Appendix A to show that if $Z_t \sim \mathrm{IID}(0,\sigma^2)$, MA(1) process also satisfies

$$\sum_{-1}^{1} \gamma(j) = \sigma^2 \cdot \Big(\sum_{j=0}^{1} \vartheta_j\Big)^2$$

& therefore

$$\sqrt{n}\,\overline{X}_n \xrightarrow{d} N\Big(0,\sigma^2\Big(\sum_0^1 \vartheta_j\Big)^2\Big)$$

ESTIMATION OF THE MEAN

Suppose (X_t) is a stationary time series. Then $\overline{X}_n$ estimates $\mu = EX_1$ but the quality of the estimation changes with dependence

THEOREM 1

Suppose (X_t) is weakly stationary with mean μ & autocov. funct. γ. Then

$$\operatorname{Var} \overline{X}_n = E(\overline{X}_n - \mu)^2 \to 0 \quad \text{if } \gamma(h) \to 0$$

$$n E(\overline{X}_n - \mu)^2 \to \sum_{-\infty}^{\infty} \gamma(h) \quad \text{if } \sum |\gamma(h)| < \infty$$

Theorem suggests that for short range dependent (X_t) i.e. if $\sum |\gamma(h)| < \infty$

$$n \operatorname{Var} \overline{X}_n \approx \sum_{-\infty}^{\infty} \gamma(h) =: v$$

& therefore, one might expect

$$\sqrt{n}\,(\overline{X}_n - \mu) \xrightarrow{d} N(0, v)$$

THEOREM 2

If (X_t) is weakly stationary & s.t.

$$X_t = \mu + \sum_{-\infty}^{\infty} \psi_j Z_{t-j}, \quad Z_t \sim \mathrm{IID}(0, \sigma^2)$$

where $\sum_{-\infty}^{\infty} |\psi_j| < \infty$ & $\sum_{-\infty}^{\infty} \psi_j \neq 0$.

Then

$$\sqrt{n}(\overline{X}_n - \mu) \xrightarrow{d} N(0, v)$$

Thm 2 allows one to build asymptotic confidence intervals for μ.

Moreover, for Gaussian sequences it holds exactly that

$$\sqrt{n}\,(\overline{X}_n - \mu) \sim N\Big(0, \sum_{|h|<n} \Big(1 - \frac{|h|}{n}\Big)\gamma(h)\Big)$$

Observe:

95% confidence interval for μ is approx.

$$\left(\overline{X}_n - \frac{\nu}{\sqrt{n}} 1.96 \,,\ \overline{X}_n + \frac{\nu}{\sqrt{n}} 1.96 \right)$$

In practice

$$\nu = \sum_{-\infty}^{\infty} \gamma(h) \quad \text{is unknown}$$

so it is frequentely estimated by

$$\hat{\nu} = \sum_{|h|<\sqrt{n}} \left(1 - \frac{|h|}{n}\right) \hat{\gamma}(h)$$

EXE 3) Find the asymptotic variance of $\overline{X}_n$ in the case of AR(1) sequence (X_t) s.t.

$$X_t = \varphi X_{t-1} + Z_t, \quad Z_t \overset{iid}{\sim} N(0, \sigma^2)$$

& $|\varphi| < 1$.

STRONG / α - MIXING

For strongly stationary (X_t) consider

$$\mathcal{F}_{-\infty}^{s} = \sigma(\ldots, X_{s-1}, X_s) \quad \&$$

$$\mathcal{F}_{t}^{\infty} = \sigma(X_t, X_{t+1}, \ldots) \quad \& \text{ coefficients}$$

$$\alpha(n) = \sup_{\substack{A \in \mathcal{F}_{-\infty}^{0} \\ B \in \mathcal{F}_{n}^{\infty}}} \left| P(A \cap B) - P(A)P(B) \right|$$

$\leftarrow$ mixing coefficients

$\boxed{\text{DEF}}$ (X_t) is strongly / α-mixing if

$$\alpha(n) \to \infty, \quad n \to \infty.$$

THEOREM 3

Suppose (X_t) is strongly stationary & one of the following holds

i) $E|X_t|^\delta < \infty$ & $\sum_{j=1}^{\infty} \alpha(j)^{1-2/\delta}$ for some $\delta > 2$

ii) $P(|X_t| < C) = 1$ for some $C > 0$ & $\sum_1^{\infty} \alpha(j) < \infty$

Then $\sum |\gamma(j)| < \infty$ &

$$\sqrt{n}(\overline{X}_n - \mu) \xrightarrow{d} N(0, v)$$

ESTIMATION OF FUNCTIONS γ, ρ & α

We will use the estimators

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{1}^{n-h} (X_t - \overline{X}_n)(X_{t+h} - \overline{X}_n)$$

&

$$\hat{\rho}(h) = \hat{\gamma}(h) / \hat{\gamma}(0) \qquad h = 0, \ldots, n-1$$

These estimators are biased.

However they make matrix

$$\hat{\Gamma}_k = \begin{bmatrix} \hat{\gamma}(0) & \hat{\gamma}(1) & \cdots & \hat{\gamma}(k-1) \\ \hat{\gamma}(1) & \hat{\gamma}(0) & \ddots & \\ & \ddots & \ddots & \\ \hat{\gamma}(k-1) & & & \hat{\gamma}(0) \end{bmatrix}$$

positive semi definite

for all k

& $\hat{R}_k = \hat{\Gamma}_k / \hat{\gamma}(0)$ too.

For linear processes these estimators are consistent & asymptotically normal.

THEOREM 4

For weakly stat. (X_t) s.t.

$$X_t = \mu + \sum_{-\infty}^{\infty} \psi_j \cdot Z_{t-j} \quad \text{for } Z_t \sim \text{IID}(0, \sigma^2)$$

where $E Z_t^4 < \infty$ & $\sum |\psi_j| < \infty$, then

$$\sqrt{n}\left(\hat{\rho}_n(h) - \rho_X(h)\right) \xrightarrow{d} N(0, V_{hh})$$

where

$$V_{hh} = \rho_X(h)^2 \cdot \kappa_4(Z) + \sum_{-\infty}^{\infty} \rho_X(h+j)\,\rho_X(j-h)$$

& $\kappa_4(Z) = \dfrac{E Z_1^4}{(E Z_1^2)^2} - 3$ ← kurtosis

THEOREM 5

Under conditions of thm 4.

$$\sqrt{n}\left[\begin{pmatrix}\hat{\rho}(1)\\ \vdots \\ \hat{\rho}(h)\end{pmatrix} - \begin{pmatrix}\rho(1)\\ \vdots \\ \rho(h)\end{pmatrix}\right] \xrightarrow{d} N(0, W)$$

where the covariance matrix has entries given by Bartlett's formula

$W = (w_{ij})$ in thm 5 satisfies

$$w_{ij} = \sum_{k=-\infty}^{\infty} \Big[\rho(k+i)\rho(k+j) + \rho(k-i)\rho(k+j) - 2\rho(i)\rho(j)\rho^2(k) - 2\rho(i)\rho(k)\rho(k+j) - 2\rho(j)\rho(k)\rho(k+i) \Big]$$

↳ Bartlett's formula

Note : heavier tail $\rightarrow$ higher kurtosis

$\rightarrow$ larger variance V_{hh}

EXAMPLE 2 (iid sequence & $\hat{\rho}$)

Let $X_t \sim IID(0, \sigma^2)$, clearly

$$\rho_X(h) = 0 \qquad h \neq 0$$

$\Rightarrow$ $w_{ij} = \delta_{ij}$ in Bartlett's formula $\Rightarrow$

$\hat{\rho}(1), \ldots, \hat{\rho}(h)$ are asympt. iid $N(0, \frac{1}{n})$.