INTEGRATED PROCESSES.

- RANOM WALK: X_t:

$$X_t - X_{t-1} = u_t \quad \forall t \in T$$

WITH $\{u_t, t \in Z\} \sim i.i.d.$ IF $u_0 = 0$, WE HAVE

$$X_t = X_0 + \sum_{j=1}^{t} u_j$$

IF $E(u_t) = \bar{\mu}$,

$$X_t - X_{t-1} = \bar{\mu} + v_t$$, $\bar{\mu}$ IS CALLED DRIFT

- X_t IS INTEGRATED OF ORDER d : ARIMA(p,d,q)

$$A(L)(1-L)^d X_t = B(L) u_t$$

$d = 0, 1, 2, \vdots$

MA(q) AND MA(∞)

$$X_t = B(L) \varepsilon_t$$

1. CONSIDER

$$\sum_{j=-\infty}^{\infty} b_j \varepsilon_{t-j} \quad , t \in Z :$$

$$\varepsilon_t \sim WN(0, \sigma^2)$$

CONVERGENCE CONDITION

1. X_n CONVERGES IN PROBABILITY TO X : $X_n \xrightarrow{P} X$ IFF

$$\lim_{n\to\infty} P\left[\,|X_n - X| > \varepsilon\,\right] = 0 \qquad \forall\, \varepsilon > 0$$

2. X_n CONVERGES A.S. TO X: $X_n \xrightarrow{a.s.} X$ IFF

$$P\left[\lim_{n\to\infty} X_n = X\right] = 1$$

3. $X_n \to X$ IN MEAN SQUARED : $X_n \xrightarrow{2} X$

$$\lim_{n\to\infty} E\left[\,|X_n - X|^2\,\right] = 0$$

4. IN DISTRIBUTION: $\{X_n\}$ WITH DISTRIBUTION $F_{X_n}(\ )$ CONVERGES IN DISTRIBUTION, IF $\exists$ A R.V. X :

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x) \qquad \forall\, x \in \text{continuity pts of distribution } F_X \text{ of } X$$

a.s $\Rightarrow$ P , m.q $\Rightarrow$ P

P $\Downarrow$ L

$$\sum_{j=-\infty}^{\infty} b_j \varepsilon_{t-j} = \sum_{j=-\infty}^{\infty} Y_j(t) = \sum_{j=-\infty}^{-1} Y_j(t) + \sum_{j=0}^{\infty} Y_j(t)$$

WHERE $Y_j(t) = b_j \varepsilon_{t-j}$ AND:

$$E\left[\,|Y_j(t)|\,\right] = |b_j|\, E\left[\,|\varepsilon_{t-j}|\,\right] \leq |b_j| \left[E\left(\varepsilon_{t-j}^2\right)\right]^{\frac{1}{2}} = |b_j|\, \sigma_\varepsilon^2 < \infty$$

$$E\left[Y_j(t)\,Y_K(t)\right] = E\left[Y_j(t)^2\right] = b_j^2\,\sigma_\varepsilon^2 \quad, \text{if } j=K$$
$$= 0 \quad, \text{if } j \neq K$$

TH.1

Let $\{X_t\}$ A STOCH PR. ON INTEGERS, $r \geq 1$ and $\{a_j, j \in \mathbb{Z}\}$ A SEQUENCE OF REAL NUMBERS. IF $\sum_{j=-\infty}^{\infty} |a_j|\, E\left[|X_{t-j}|^r\right)^{1/r} < \infty$, than FOR ALL t, THE SEQUENCE $\sum_{j=-\infty}^{\infty} a_j X_{t-j}$ CONVERGES A.S. AND IN THE MEAN OF ORDER r. TO A R.V. Y_t SUCH THAT $E\left(|Y_t|^r\right) < \infty$.

TH.2.

Let $\{X_t\}$ A 2^{nd} ORDER STATIONARY AND $\{a_j, j \in \mathbb{Z}\}$ A CONVERGENT SEQUENCE OF REAL NUMBERS, IE $\sum_{j=-\infty}^{\infty} |a_j| < \infty$. Than $\sum_{j=-\infty}^{\infty} a_j X_{t-j}$ CONVERGES A.S. AND IN THE SQUARED MEAN TO A R.V. $Y_t \in L_2$ AND $\{Y_t\}$ IS 2^{nd} ORDER STATIONARY

ASSUME $\sum_{j=-\infty}^{\infty} b_j^2 < \infty$.

$$Y_1^m(t) = \sum_{j=-m}^{-1} b_j\,\varepsilon_{t-j} \xrightarrow[m\to\infty]{2} Y_1(t) = \sum_{j=-\infty}^{-1} b_j\,\varepsilon_{t-j}$$

$$Y_2^n(t) = \sum_{j=0}^{n} b_j\,\varepsilon_{t-j} \xrightarrow[n\to\infty]{2} Y_2(t) = \sum_{j=0}^{\infty} b_j\,\varepsilon_{t-j}$$

hence $\forall t$

$$Y^{m,n}(t) = Y_1^m(t) + Y_2^n(t) \xrightarrow[\substack{m\to\infty \\ n\to\infty}]{2} X_t = Y_1(t) + Y_2(t) = \sum_{j=-\infty}^{\infty} b_j\,\varepsilon_{t-j}$$

also

$$X_n(t) = Y_1^n(t) + Y_2^n(t) = \sum_{j=-n}^{-1} b_j\,\varepsilon_{t-j} + \sum_{j=0}^{n} b_j\,\varepsilon_{t-j} \xrightarrow[n\to\infty]{2} X_t \equiv \sum_{j=-\infty}^{\infty} b_j\,\varepsilon_{t-j}$$

hence

- $\sum_{j=-\infty}^{\infty} b_j^2 < \infty \Rightarrow \sum_{j=-\infty}^{\infty} b_j \varepsilon_{t-j}$ CONVERGES IN SQUARED MEAN TO A R.V. X_t

- $\sum_{j=-\infty}^{+\infty} |b_j| < \infty \Rightarrow \sum_{j=-\infty}^{\infty} b_j \varepsilon_{t-j}$ CONVERGE A.S. TO X_t

$\sum_{j=-\infty}^{\infty} |b_j| < \infty \Rightarrow \sum_{j=-\infty}^{\infty} b_j^2 < \infty \Rightarrow \sum_{j=-\infty}^{\infty} b_j \varepsilon_{t-j}$ CONV. IN SQUARED MEAN TO X_t

IF ε_t INDEPENDENT:

$\sum_{j=-\infty}^{\infty} b_j^2 < \infty \Rightarrow \sum_{j=-\infty}^{\infty} b_j \varepsilon_{t-j}$ CONVERGE A.S. TO X_t

IF X_t IS THE LIMIT, $\Rightarrow$

$$X_t = \sum_{j=-\infty}^{\infty} b_j \varepsilon_{t-j}$$

OR IF $\tilde{X}_t = X_t + \mu$,

$$\tilde{X}_t = \mu + \sum_{j=-\infty}^{\infty} b_j \varepsilon_{t-j}$$

MEAN, VAR, COVAR OF DEMEANED MA(∞)

- **$E(X_t) = 0$**
- **$var(X_t) = E(X_t^2) = \lim_{n\to\infty} \sum_{j=-n}^{n} b_j^2 \sigma_\varepsilon^2 = \sigma_\varepsilon^2 \sum_{j=-\infty}^{\infty} b_j^2$**

$\mathrm{cov}(X_t, X_{t+K}) = E(X_t\, X_{t+K}) =$

$$= \lim_{n\to\infty} E\left\{\left[\sum_{i=-n}^{n} b_i\, \varepsilon_{t-i}\right]\left[\sum_{j=-n}^{n} b_j\, \varepsilon_{t+K-j}\right]\right\}$$

$$= \lim_{n\to\infty} \sum_{i=-n}^{n}\sum_{j=-n}^{n} b_i\, b_j\, E(\varepsilon_{t-i}\, \varepsilon_{t+K-j})$$

- $= \lim_{n\to\infty} \sum_{i=-n}^{n-K} b_i\, b_{i+K}\, \sigma_\varepsilon^2 = \sigma_\varepsilon^2 \sum_{i=-\infty}^{\infty} b_i\, b_{i+K}$ **, if $K \geq 1$**

- $= \lim_{n\to\infty} \sum_{j=-n}^{n} b_j\, b_{j+|K|}\, \sigma_\varepsilon^2 = \sigma_\varepsilon^2 \sum_{j=-\infty}^{\infty} b_j\, b_{j+|K|}$ **, if $K \leq -1$**

since $t-i = t+K-j \Rightarrow j = i+K$ and $i = j-K$, for all $K \in \mathbb{Z}$.

- $\mathrm{cov}(X_t, X_{t+K}) = \sigma_\varepsilon^2 \sum_{j=-\infty}^{\infty} b_j\, b_{j+|K|}$

- $\mathrm{correl}(X_t, X_{t+K}) = \sum_{j=-\infty}^{\infty} b_j\, b_{j+|K|} \Big/ \sum_{j=-\infty}^{\infty} b_j^2$

the series $\sum_{j=-\infty}^{\infty} b_j\, b_{j+K}$ CONVERGE IN ABSOLUTE VALUES, SINCE:

$$\left| \sum_{j=-\infty}^{\infty} b_j\, b_{j+K} \right| \leq \sum_{j=-\infty}^{\infty} \left| b_j\, b_{j+K} \right| \leq \left[\sum_{j=-\infty}^{\infty} b_j^2\right]^{\frac{1}{2}} \left[\sum_{j=-\infty}^{\infty} b_{j+K}^2\right]^{\frac{1}{2}} < \infty$$

• FOR A CAUSAL MA(∞): •

$$\tilde{X}_t = \mu + \sum_{j=0}^{\infty} b_j\, \varepsilon_{t-j} \qquad \varepsilon_t \sim WN(0, \sigma^2)$$

$\mathrm{cov}(X_t, X_{t+K}) = \sigma_\varepsilon^2 \sum_{j=0}^{\infty} b_j\, b_{j+|K|}$, $\mathrm{corr} = \sum_{j=0}^{\infty} b_j\, b_{j+|K|} \Big/ \sum_{j=0}^{\infty} b_j^2$

STATIONARITY

$$X_t = \mu + \sum_{j=-\infty}^{\infty} b_j \varepsilon_{t-j}$$

WHERE $\varepsilon_t \sim WN(0, \sigma^2)$ AND $\sum_{j=-\infty}^{\infty} b_j^2 < \infty$ IS 2^{nd} ORDER STATIONARY BECAUSE $cov(X_t, X_{t+k})$ DOES NOT DEPEND ON t.

IF WE FURTHER ASSUME $\varepsilon_t \sim$ I.I.D. WITH $E|\varepsilon_t| < \infty$, AND $\sum_{j=-\infty}^{\infty} b_j^2 < \infty$ THE X_t IS STRICTLY STATIONARY

MA(q)

$$\tilde{X}_t = \mu + B(L)\varepsilon_t$$

$$B(L) = 1 + b_1 L + \dots + b_q L^q$$

MA(q) IS ALWAYS STATIONARY

CONSIDER DEMEANED X_t:

- $$E(\tilde{X}_t) = \mu$$
  $$E(X_t) = E(\tilde{X}_t - \mu) = 0$$

- $var(X_t) = \sigma_\varepsilon^2(1 + b_1^2 + \dots + b_q^2) = \sigma_\varepsilon^2(1 + \sum_{j=1}^{q} b_j^2)$

$cov(X_t, X_{t+k}) = E(X_t X_{t+k}) = \gamma(k) = \sigma_\varepsilon^2 \sum_{j=0}^{q-k} b_j b_{j+k}$

- $\gamma(k) = \sigma_\varepsilon^2 \left[b_k + \sum_{j=1}^{q-k} b_j b_{j+k} \right] =$

$= \sigma_\varepsilon^2 [b_k + b_1 b_{k+1} + \dots + b_{q-k} b_q]$ FOR $1 \leq k \leq q$

- $\gamma(K) = 0$ FOR $K \geq q+1$!

$\gamma(K) = \gamma(-K)$ FOR $K<0$

$$\rho(K) = \left[b_K + \sum_{j=1}^{q-K} b_j b_{j+K}\right] \Big/ \left[1 + \sum_{j=1}^{q} b_j^2\right], \quad 1 \leq K \leq q$$

$$= 0 \qquad , \quad K \geq q+1$$

<u>$q=1$</u> MA(1)

$$X_t = (1 + b_1 L)\,\varepsilon_t$$

$$\rho(1) = \frac{b_1}{1+b_1^2} \qquad K=1$$

$$= 0 \qquad , K \geq 2$$

<u>$q=2$</u> MA(2)

$$\rho(K) = (b_1 + b_1 b_2) / (1 + b_1^2 + b_2^2) \quad , \quad K=1$$

$$= b_2/(1+b_1^2+b_2^2) \quad , \quad K=2$$

$$= 0 \quad , \quad K \geq 3$$

MA(q)

$$\rho(q) = \frac{b_q}{1+b_1^2 + \dots + b_q^2}$$

$|\rho(q)| \leq 0.5$. In general $|\rho(K)| \leq \cos\{\pi/([q/K]+2)\}$

$[x]$ = largest integer $\leq x$:

$q=1 \Rightarrow |\rho(1)| \leq \cos(\pi/3) = 0.5$

$q=2 \Rightarrow |\rho(1)| \leq \cos(\pi/4) = 0.7071$, $|\rho(2)| \leq \cos(\pi/3) = 0.5$

$q=3 \Rightarrow |\rho(1)| \leq \cos(\pi/5) = 0.809$, $|\rho(2)| \leq \cos(\pi/3) = 0.5$, $|\rho(3)| \leq \cos\frac{\pi}{3} = 0.5$

CONSIDER AGAIN MA(1). BY RECURSIVE SUBSTITUTION:

$$\begin{aligned} \varepsilon_t &= X_t + b_1 \varepsilon_{t-1} \\ &= X_t + b_1(X_{t-1} - b_1 \varepsilon_{t-2}) \\ &= \sum_{j=0}^{\infty} (b_1)^j X_{t-j} \end{aligned}$$

IT IS MEAN SQUARED CONSISTENT IF $|b_1| < 1$; I.E. IF THE ROOT OF $(B(L) = 0)$ LIES OUTSIDE THE UNIT CIRCLE. THE SAME PRINCIPLE CARRIES OVER TO HIGHER ORDER PROCESSES.

$$MA(q): \quad X_t = B(L)\,\varepsilon_t$$

IS INVERTIBLE IF ALL ROOTS OF $B(L)=0$ LIE OUTSIDE THE UNIT CIRCLE. SO THAT

$$\varepsilon_t = B^{-1}(L)\, X_t = \sum_{i=0}^{\infty} h_i X_{t-i}$$

I.E AN MA(q) INVERTIBLE CAN BE WRITTEN AS AR(∞) WITH COEFF h_i SUCH THAT $\sum_{i=0}^{\infty} |h_i| < \infty$. HENCE ENSURING THAT $\sum_{i=0}^{\infty} h_i X_{t-i}$ IS MEAN SQUARED CONVERGENT

MA(∞) REPRESENTATION OF AR(p)

Define a 2nd order STATIONARY PROCESS

$$A(L)X_t = \varepsilon_t$$

where $A(L) = 1 - a_1 L - \dots - a_p L^p$.

$$X_t = \Psi(L)\varepsilon_t$$

where $\Psi(L) = [A(L)]^{-1}$. To compute the Ψ_j coefficients note that

$$A(L)\Psi(L) = 1$$

Let $\Psi_j = 0$ FOR $j < 0$.

$$\left[1 - \sum_{K=1}^{p} a_K L^K\right]\left[\sum_{j=-\infty}^{\infty} \Psi_j L^j\right] = \sum_{j=-\infty}^{\infty} \Psi_j \left[L^j - \sum_{K=1}^{p} a_K L^{j+K}\right]$$

$$= \sum_{j=-\infty}^{\infty}\left[\Psi_j - \sum_{K=1}^{p} a_K \Psi_{j-K}\right]L^j = \sum_{j=-\infty}^{\infty} \tilde{\Psi}_j L^j = 1$$

hence: $\tilde{\Psi}_j = 1$ if $j=0$, and $\tilde{\Psi}_j = 0$ for $j \neq 0$, i.e.:

$$\Psi_j - \sum_{K=1}^{p} a_K \Psi_{j-K} = 1 \quad , \text{ if } j=0$$

$$= 0 \quad , \text{ if } j \neq 0$$

where $L^K \Psi_j = \Psi_{j-K}$

SINCE $\psi_j = 0$ for $j < 0$, we see that:

$$\psi_0 = 1$$

$$\psi_j = \sum_{K=1}^{p} a_K \psi_{j-K}, \quad j \geq 1$$

or:

$$\psi_0 = 1 \qquad a_1 \psi_{j-1} + a_2 \psi_{j-2} + a_3 \psi_{j-3}$$

$$\psi_1 = a_1 \psi_0 = a_1 \qquad a_1 \psi_{2-1} + a_2 \psi_{2-2} + a_3 \psi_{2-3}$$

$$\psi_2 = a_1 \psi_1 + a_2 \psi_0 = a_1^2 + a_2$$

$$\psi_3 = a_1 \psi_2 + a_2 \psi_1 + a_3 = a_1^3 + 2a_2 a_1 + a_3$$

$$\vdots$$

$$\psi_p = \sum_{K=1}^{p} a_K \psi_{j-K}$$

and

$$\psi_j = \sum_{K=1}^{p} a_K \psi_{j-K}, \text{ for } \underline{j \geq p}$$

UNDER STATIONARITY, ψ_j DIMINISH EXPONENTIALLY WHEN $j \to \infty$ POSSIBLY WITH OSCILLATIONS

GIVEN $X_t = \psi(L)\, \varepsilon_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$

WE CAN COMPUTE:

- $\mathrm{cov}(X_t, X_{t+K}) = \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+|K|}$
- $\mathrm{corr}(X_t, X_{t+K}) = \sum_{j=0}^{\infty} \psi_j \psi_{j+|K|} \Big/ \sum_{j=0}^{\infty} \psi_j^2$

ARMA(p,q)

THE IDENTIFICATION REQUIREMENT: STATIONARY AND INVERTIBLE
I.E. ROOTS OF $A(L)$ AND $B(L)$ ARE OUTSIDE THE UNIT CIRCLE.

IN CASE OF ARIMA(p,d,q) we have: $A(L)(1-L)^d \tilde{X}_t = B(L)\varepsilon_t$
SO THAT $\Delta^d \tilde{X}_t$ FOLLOWS A STATIONARY PROCESS

$$A(L)\, X_t = B(L)\, \varepsilon_t$$

$$\tilde{X}_t = \bar{\mu} + \sum_{j=1}^{p} a_j X_{t-j} + \varepsilon_t - \sum_{j=1}^{q} b_j \varepsilon_{t-j}$$

to remove
the constant: WE DEMEAN THE PROCESS:

$$X_t = \tilde{X}_t - \mu, \text{ where } \mu = \frac{\bar{\mu}}{A(L)} = \frac{\bar{\mu}}{1-\sum_{j=1}^{p} a_j}$$

$$X(t) = \frac{B(L)}{A(L)}\varepsilon_t$$

$$\frac{B(L)}{A(L)} = \Psi(L) = \sum_{j=0}^{\infty} \psi_j L^j$$

THE ψ_j COEFFICIENTS ARE OBTAINED FROM SOLVING:

$$A(L)\,\Psi(L) = B(L)$$

RECALL: $E(X_{t-j}\varepsilon_t) = 0 \quad \forall j \geq 1$

$$\left[1-\sum_{k=1}^{p} a_k L^k\right]\left[\sum_{j=0}^{\infty}\psi_j L^j\right] = 1 + \sum_{j=1}^{q} b_j L^j = -\sum_{j=0}^{q} b_j L^j; \quad b_0 = -1$$

$$A(L)\psi_j = -b_j \quad , \text{ for } j = 0, 1, \ldots, q$$

Let $\psi_j = 0, \forall j < 0$

- $\psi_j = \sum_{k=1}^{p} a_k \psi_{j-k} - b_j \quad ; \quad j = 0, 1, \dots, q$
- $\quad = \sum_{k=1}^{p} a_k \psi_{j-k} \quad ; \quad j \geq q+1$

$\psi_0 = 1$

$\psi_1 = a_1 \psi_0 - b_1 = a_1 - b_1$

$\psi_2 = a_1 \psi_1 + a_2 \psi_0 - b_2 = a_1 \psi_1 + a_2 - b_2 = a_1^2 - a_1 b_1 + a_2 - b_2$

$\vdots$

$\psi_j = \sum_{k=1}^{p} a_k \psi_{j-k} \quad , \quad j \geq q+1$

THE COEFF'S ψ_j ARE BEHAVED AS AUTOCORRELATIONS OF AN AR(p) EXCEPT FOR $\psi_1, \dots \psi_q$.

- AUTOCOVAR AND AUTOCORRELATIONS.

SUPPOSE: a) X_t 2nd ORDER STATIONARY WITH $\sum_{j=1}^{p} a_j \neq 1$

b) $E(X_{t-j}\, \varepsilon_t) = 0 \quad \forall j \geq 1$

BY STATIONARITY: $E(X_t) = \mu \;,\; \forall t \quad ; \mu = \bar{\mu} + \sum_{j=1}^{p} a_j \mu$:

$$\mu = \bar{\mu} \Big/ 1 - \sum_{j=1}^{p} a_j$$

ONLY THE AR PART DETERMINES THE MEAN

FOR DEMEANED PROCESSES $x_t = \tilde{x}_t - \mu$;

$$x_t = \sum_{j=1}^{p} a_j x_{t-j} + \varepsilon_t - \sum_{j=1}^{q} b_j \varepsilon_{t-j}$$

$$x_{t+k} = \sum_{j=1}^{p} a_j x_{t+k-j} + \varepsilon_{t+k} - \sum_{j=1}^{q} b_j \varepsilon_{t+k-j}$$

$$E(x_t x_{t+k}) = \sum_{j=1}^{p} a_j E(x_t x_{t+k-j}) + E(x_t \varepsilon_{t+k}) - $$

$$- \sum_{j=1}^{q} b_j E(x_t \varepsilon_{t+k-j})$$

$$\gamma(k) = \sum_{j=1}^{p} a_j \gamma(k-j) + \gamma_{x\varepsilon}(k) - \sum_{j=1}^{q} b_j \gamma_{x\varepsilon}(k-j)$$

$\rightarrow$ $$\gamma_{x\varepsilon}(k) = E(x_t \varepsilon_{t+k}) = 0 \text{ , if } k \geq 1$$
$$\neq 0 \text{ , if } k \leq 0$$

$\rightarrow$ $$\gamma_{x\varepsilon}(0) = E(x_t \varepsilon_t) = \sigma_\varepsilon^2$$

FOR $k \geq q+1$

$$\gamma(k) = \sum_{j=1}^{p} a_j \gamma(k-j)$$

$$\rho(k) = \sum_{j=1}^{p} a_j \rho(k-j)$$

THE VARIANCE

$$\gamma(0) = \sum_{j=1}^{p} a_j \gamma(j) + \sigma_\varepsilon^2 - \sum_{j=1}^{q} b_j \gamma_{x\varepsilon}(-j)$$

hence:

$$\gamma(0) = \left[\sigma_\varepsilon^2 - \sum_{j=1}^{q} b_j \gamma_{x\varepsilon}(-j) \right] \Big/ \left[1 - \sum_{j=1}^{p} a_j \rho(j) \right]$$

IN GENERAL:

$$A(L)\,\gamma(K) = B(L)\,\gamma_{x\varepsilon}(K) \quad , \quad K \geq 0$$

WHERE:

$$\gamma(-K) = \gamma(K) \quad , \quad L^j \gamma(K) = \gamma(K-j)$$

AND $L^j \gamma_{x\varepsilon}(K) = \gamma_{x\varepsilon}(K-j)$. IN PARTICULAR $A(L)\,\gamma(K) = 0$ FOR $K \geq q+1$ AND $A(L)\,\rho(K) = 0$ FOR $K \geq q+1$.

AUTOCORRELATIONS OF ARMA(p,q) ARE BEHAVED AS THOSE OF AR(p)

example: ARMA(1,1)

$$\tilde{x}_t = \bar{\mu} + a_1 \tilde{x}_{t-1} + \varepsilon_t - b_1 \varepsilon_{t-1} \qquad |a_1| < 1$$

$$x_t - a_1 x_{t-1} = \varepsilon_t - b_1 \varepsilon_{t-1} \quad , \quad x_t = \tilde{x}_t - \mu$$

$$\gamma(0) = a_1 \gamma(0) + \gamma_{x\varepsilon}(0) - b_1 \gamma_{x\varepsilon}(-1)$$

$$\gamma(1) = a_1 \gamma(0) + \gamma_{x\varepsilon}(1) - b_1 \gamma_{x\varepsilon}(0)$$

• $\gamma_{x\varepsilon}(1) = 0$, • $\gamma_{x\varepsilon}(0) = \sigma_\varepsilon^2$

• $\gamma_{x\varepsilon}(-1) = E(x_t \varepsilon_{t-1}) = a_1 E(x_{t-1}\varepsilon_{t-1}) + E(\varepsilon_t \varepsilon_{t-1}) - b_1 E(\varepsilon_{t-1}^2)$

$$= a_1 \gamma_{x\varepsilon}(0) - b_1 \sigma_\varepsilon^2 = (a_1 - b_1)\,\sigma_\varepsilon^2$$

$$\gamma(0) = a_1 \gamma(1) + \sigma_\varepsilon^2 - b_1 (a_1 - b_1)\,\sigma_\varepsilon^2 = a_1 \gamma(1) + \left[1 - b_1(a_1 - b_1)\right] \sigma_\varepsilon^2$$

$$\gamma(1) = a_1\gamma(0) - b_1\sigma_\varepsilon^2 = a_1\left[a_1\gamma(1) + [1 - b_1(a_1-b_1)\sigma_\varepsilon^2]\right] - b_1\sigma_\varepsilon^2$$

$$= \{a_1[1-b_1(a_1-b_1)-b_1\}\sigma_\varepsilon^2/(1-a_1^2) = \{a_1 - b_1a_1^2 + a_1b_1^2 - b_1\}\frac{\sigma_\varepsilon^2}{(1-a_1^2)}$$

$$= (1-b_1a_1)(a_1-b_1)\sigma_\varepsilon^2/(1-a_1^2)$$

$$\gamma(0) = a_1\gamma(1) + [1-b_1(a_1-b_1)]\sigma_\varepsilon^2 = a_1\frac{(1-b_1a_1)(a_1-b_1)\sigma_\varepsilon^2}{1-a_1^2} + [1-b_1(a_1-b_1)]\sigma_\varepsilon^2$$

$$= \frac{\sigma_\varepsilon^2}{1-a_1^2}\left[a_1(1-b_1a_1)(a_1-b_1) + (1-a_1^2)[1-b_1(a_1-b_1)]\right]$$

$$= \frac{\sigma_\varepsilon^2}{1-a_1^2}\left[a_1^2 - b_1a_1^3 + a_1^2b_1^2 - a_1b_1 + 1 - a_1^2 - b_1a_1 + b_1a_1^3 + b_1^2 - a_1^2b_1^2\right]$$

$$= \frac{\sigma_\varepsilon^2}{1-a_1^2}\left[1 - 2a_1b_1 + b_1^2\right]$$

$$\gamma(0) = (1-2a_1b_1+b_1^2)\frac{\sigma_\varepsilon^2}{1-a_1^2}$$

$$\gamma(1) = (1-a_1b_1)(a_1-b_1)\frac{\sigma_\varepsilon^2}{1-a_1^2}$$

$$\gamma(K) = a_1\gamma(K-1), \quad K \geq 2$$

INVERTIBILITY

ARMA(P,q) $\quad A(L)X_t = B(L)\varepsilon_t$

IS INVERTIBLE IF THE ROOTS OF THE $B(L)$ POLYNOMIAL ARE OUTSIDE THE UNIT ROOT. THE PROCESS CAN BE WRITTEN AS AR(∞)

WOLD'S THEOREM

ANY STATIONARY PROCESS X_t, LINEAR OR NONLINEAR CAN BE UNIQUELY REPRESENTED AS THE SUM OF TWO MUTUALLY UNCORRELATED PROCESSES

$$X_t = D_t + Y_t$$

WHERE D_t IS DETERMINISTIC AND Y_t IS PURELY NONDETERMINISTIC, AND CAN BE WRITTEN AS AN MA(∞) PROCESS

$$Y_t = \sum_{i=0}^{\infty} h_i \, \varepsilon_{t-j} \quad ; \quad \varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$$

IMPLICATION:

EVEN THOUGH THE PROCESS MAY BE NONLINEAR THE DECOMPOSITION IS LINEAR AND IS DETERMINED BY THE 2nd MOMENT OF THE PROCESS. THUS WE HOPE TO APPROXIMATE ANY PROCESS BY A LINEAR ONE. THIS DOES NOT MEAN THAT A NONLINEAR PROCESS WON'T BE A BETTER APPROXIMATION

ESTIMATION IN THE TIME DOMAIN

OBSERVATIONS $X_1 \dots X_T$ ON A 2nd ORDER STATIONARY.

- ESTIMATING THE MEAN

$$\bar{X}_T = \frac{1}{T} \sum_{t=1}^{T} X_t$$

Unbiased:

$$E(\bar{X}) = \frac{1}{T} \sum E(X_t) = \mu$$

CONSISTENT:

$$\operatorname{var}(\bar{X}) = \frac{1}{T^2} \sum_{s=1}^{T} \sum_{z=1}^{T} \operatorname{cov}(X_s, X_z) = \frac{1}{T} \sum_{K=-(T-1)}^{T-1} \left(1 - \frac{|K|}{T}\right) \gamma_X(K$$

AUTOCORRELATIONS

$$\hat{\rho}_h = \frac{\frac{1}{T-h} \sum_{t=h+1}^{T} (x_t - \bar{x})(x_{t-h} - \bar{x})}{\frac{1}{T} \sum_{t=1}^{T} (x_t - \bar{x})^2}$$

WE NEED THE ASYMPTOTIC DISTRIBUTION OF $\hat{\rho}_h$ to be able to test $H_0: \rho_h = 0$ AGAINST $H_A: \rho_h \neq 0$

ASY DISTRIBUTION OF

$$\begin{bmatrix} \sqrt{T}(\hat{\rho}_1 - \rho_1) \\ \sqrt{T}(\hat{\rho}_2 - \rho_2) \\ \vdots \\ \sqrt{T}(\hat{\rho}_m - \rho_m) \end{bmatrix} \overset{ASY}{\sim} N(0, W_m) \quad ; \quad W_m = [w_{jk}]_{j,k=1\ldots m}$$

$$\sqrt{T}(\hat{\rho}_h - \rho_h) \overset{ASY}{\sim} N[0, w_{kk}]$$

WE KNOW THAT IF ALL AUTOCORRELATIONS ARE ZERO,

$$\sqrt{T}\,\hat{\rho}_h \overset{ASY}{\sim} N(0,1), \text{ FOR } K \geq 1.$$

HENCE WE CAN BUILD A STATISTIC

$$\frac{\hat{\rho}_h - 0}{1/\sqrt{T}}$$

AND CONFIDENCE INTERVAL

$$\hat{\rho}_h \pm 1.96 \frac{1}{\sqrt{T}} \simeq \hat{\rho}_h \pm 2 \frac{1}{\sqrt{T}}$$

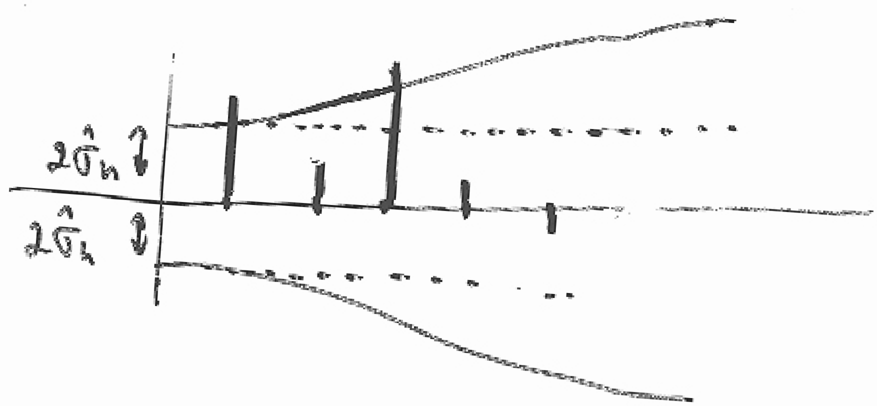
TO TEST $H_0: \rho_h = 0$ AGAINST $H_A: \rho_h \neq 0$, $\boxed{\forall h}$

≡ THAT THE h'th AUTOCORRELATION AND ALL FOLLOWING ARE ZERO AGAINST THE CONTRARY

$$var(\hat{\rho}_h) = \frac{1}{T}\left(1 + 2\sum_{j=1}^{K-1} \hat{\rho}_j^2\right), \quad K > 1$$

hence

h	1	2	3
ρ_h			
$\hat{\sigma}_h$	$\frac{1}{\sqrt{T}}$	$\frac{1}{\sqrt{T}}(1+2\hat{\rho}_1^2)^{\frac{1}{2}}$	$\frac{1}{\sqrt{T}}[1+2(\hat{\rho}_1^2+\hat{\rho}_2^2)]^{\frac{1}{2}}$



IN PRACTICE

IF $\hat{\rho}_1 > 2\frac{1}{\sqrt{T}}$ reject $H_0: \rho_1 = 0$ AND THE PROCESS IS MA(0) (W.N.)

$\Rightarrow$ IT STILL MAY BE MA(2) $\Rightarrow$ TRY $\hat{\rho}_2$

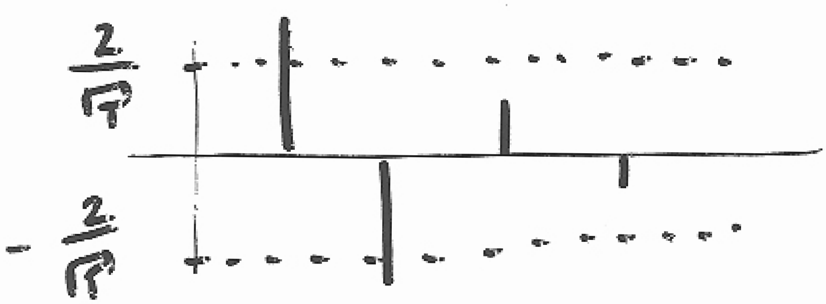
$var(\hat{\rho}_2) = \frac{1}{T}(1 + 2\rho_1^2)$ IF $\hat{\rho}_2 > 2 \cdot \hat{\sigma}(\rho_2)$ reject $H_0: \rho_2 = 0$

$\Rightarrow$ IT STILL CAN BE MA(3) $\Rightarrow$ CHECK ρ_3, ETC.

PACF

$$\sqrt{T}\,\hat{a}_{kk} \overset{ASY}{\sim} N(0,1) \quad , \; k > p$$

APPROXIMATE THE VARIANCE BY $\frac{1}{T}$



BOX - JENKINS APPROACH

1. FIND THE ORDER OF $\overset{A(L),\,B(L)}{X_t}$ USING ACF AND PACF
2. ESTIMATE
3. VERIFY IF THE RESIDUALS ARE W.N.
   (BOX-PIERCE, LJUNG-BOX)
4. GOODNESS OF FIT (AIC, BIC)

ESTIMATION IN TIME DOMAIN

• AR(p)

$$X_t = a_1 X_{t-1} + a_2 X_{t-2} + \dots + a_p X_{t-p} + e_t$$

$; t = p+1, \dots, T \qquad ; e_t \sim iid(0, \sigma^2)$

THE USUAL METHOD IS $\boxed{\text{OLS}}$

• equivalent to ML CONDITIONAL ON THE INITIAL OBSERVATIONS $X_1, \dots, X_K$ IF THE ERRORS ARE NORMALLY DISTRIBUTED

$$e_1 = X_1 - a_1 X_0 - a_2 X_{-1} - \dots - a_p X_{1-p}$$

and $X_0, \dots, X_{1-p}$ are not observed

• IF WE ASSUME NORMALITY AND THE DATA ARE NORMAL, WE HAVE EXACT MLE

drawbacks: – precision depends on the normality
– complex compared to OLS
⇒ OLS IS ADEQUATE IN MANY CASES.

• GIVEN STATIONARITY, i.e. ROOTS OF $A(L)=0$ OUTSIDE UNIT CIR THE ASYMPTOTIC DISTRIBUTION OF THE t-RATIO ARE ALL STAND NORMAL, THE PARAMETER ESTIMATES ARE ALSO ASYMPTOTICALLY NORMALLY DISTRIBUTED

• ADEQUACY OF THE NORMAL APPROXIMATION
THE NORMAL APPROXIMATION TO THE EXACT DISTRIBUTION BECOMES LESS ADEQUATE AS THE ROOTS GET CLOSER TO 1.

- **MA(q)**

$$X_t = \mu + B(L)\,\varepsilon_t \quad , \quad t = 1, \dots T$$

$$= \mu + \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$$

the roots of $B(L)$ are outside unit circle

$\varepsilon_t \sim i.i.d\ (0, \sigma^2)$

We cannot use OLS since $A(L) X_t = B(L)^{-1} X_t = \varepsilon_t$ involv[es] an ∞ number of parameters.

We assume normality of the errors and proceed with MLE

- **ARMA(p,q)**

Same approach as in the pure MA. Use normality of [ε] approximate the likelihood f. of $x_1, \dots, x_T$ to obtain: $\hat{a}$'s, $\hat{b}$'s and σ_ε^2. using numerical optimization.

CONDITIONS FOR PARAMETER IDENTIFICATION

(1) $A(L)$ and $B(L)$ have no common factor

(2) all roots of $A(L)$ and $B(L)$ are outside unit circle

(3) a_p and b_q are not both zero

SOME RESULTS THAT WILL REAPPEAR LATER....

(1) IDENTIFICATION :

LOOK AT ACF AND PACF

(2) DIAGNOSTIC CHECKING :

MLE ESTIMATORS

LET $\tilde{X}_t \sim ARIMA(p,d,q)$:

$$A(L)\nabla^d \tilde{X}_t = \bar{\mu} + B(L)\varepsilon_t$$

$$A(L) = 1 - a_1 L - \cdots - a_p L^p$$
$$B(L) = 1 - b_1 L - \cdots - b_q L^q$$

FIRST: DIFFERENCES FOLLOW A STATIONARY PROCESS WHICH CAN BE ESTIMATED

$$X_t = \nabla^d \tilde{X}_t$$

$$X_t \sim ARMA(p,q)$$

NEXT: ASSUME ε_t ARE IID NORMAL

$$A(L)X_t = \bar{\mu} + B(L)\varepsilon_t$$

$$X_t = a_1 X_{t-1} + \cdots + a_p X_{t-p} + \varepsilon_t - b_1\varepsilon_{t-1} - \cdots - b_q\varepsilon_{t-q} + \bar{\mu}$$

$$\varepsilon_t = X_t - a_1 X_{t-1} - \cdots - a_p X_{t-p} + b_1\varepsilon_{t-1} + \cdots + b_q\varepsilon_{t-q} - \bar{\mu}$$

AT THIS STAGE WE CAN DEMEAN THE DATA BY SUBSTACTING $\mu = \dfrac{\bar{\mu}}{1-a_1-\cdots-a_p}$. In that case: $\varepsilon_t = X_t - \mu_t(\theta)$

WHERE

$$\theta = \begin{bmatrix} a_1 \\ \vdots \\ a_p \\ b_1 \\ \vdots \\ b_q \\ \sigma^2 \end{bmatrix}$$

CONTAINS ALL UNKNOWN COEFFICIENTS

$\Rightarrow$ $\mu_t(\theta) = E(X_t \mid X_{t-1} \dots X_1, \varepsilon_{t-1} \dots \varepsilon_1)$

OTHERWISE $\theta = [\bar{\mu}, a_1 \dots a_p, b_1 \dots b_q, \sigma^2]'$

$$L(\theta) = \prod_{t=1}^{T} l_t(\varepsilon_t; \theta)$$

$$\hat{\theta}_{MLE} = \arg\max \log L(\varepsilon; \theta)$$

$$\log L = -\frac{T}{2}\log\sigma^2 - \frac{T}{2}\log 2\pi - \frac{1}{2\sigma^2}\sum_{t=1}^{T}\varepsilon_t^2$$

$$\sqrt{T}\,(\hat{\theta}_T - \theta_0) \underset{A}{\sim} N(0, I^{-1})$$

WHERE $I = E_0\left[-\frac{\partial^2 \log l_t(X_t;\theta)}{\partial\theta\,\partial\theta'}\right]$

$$\hat{V}_{as} \sqrt{T} (\hat{\theta}_T - \theta_0) = \hat{I}_T^{-1}$$

$$\hat{I}_T = -\frac{1}{T} \sum_{t=1}^{T} \frac{\partial^2 \log l_t(\hat{\theta}_T)}{\partial\theta\, \partial\theta'} \simeq \frac{1}{T} \sum_{t=1}^{T} \frac{\partial \log l_t(\hat{\theta}_T)}{\partial\theta} \frac{\partial \log l_t(\hat{\theta}_T)}{\partial\theta}$$

THE COMPONENTS OF $\hat{I}_T^{-1}$ APPROXIMATE VARIANCES OF ELEMENTS OF θ.

TECHNICAL COMMENTS

THERE ARE 2 METHODS : UNCONDITIONAL MLE AND BACKCASTING

1. UNCONDITIONAL

PROBLEM:

AT $t = 1$ $\quad x_{t-1} \dots x_{t-p} = X_* = x_0, x_{-1}, \dots, x_{-p+1}$

$$\varepsilon_{t-1} \dots \varepsilon_{t-q} = \varepsilon_* = \varepsilon_0, \varepsilon_{-1}, \dots, \varepsilon_{-q+1}$$

ARE UNKNOWN.

THE METHOD CONSISTS ON SETTING

$$E(\varepsilon_*) = 0$$

I.E. REPLACING BY THE UNCONDITIONAL MEAN

AND USING $(x_1, \dots, x_p)$ TO INITIALIZE THE ALGORITHM, I.E. REDUCING THE SAMPLE SIZE FROM T TO $T-p$.

let $\varepsilon_t = \varepsilon_t(A, B \mid X_*, \varepsilon_*)$, $t = 1 \dots T$ (FOR SIMPLICITY)

JOINT DENSITY

$$f(\varepsilon_1 \dots \varepsilon_T) = \frac{1}{(2\pi)^{T/2} \sigma^T} \exp\left(-\frac{1}{2\sigma^2} \sum_{t=1}^{T} \varepsilon_t^2\right)$$

$$f(\varepsilon_1, \dots \varepsilon_T \mid X_*, \varepsilon_*) = \frac{1}{(2\pi)^{T/2} \sigma^T} \exp\left(-\frac{1}{2\sigma^2} \sum_{t=1}^{T} \varepsilon_t^2(A, B, \bar{\mu} \mid X_*, \varepsilon_*)\right)$$

$$= L_*(A, B, \bar{\mu}, \sigma^2) = L_*(\theta)$$

- MAXIMIZE L_* W.R. TO $a_1 \dots a_q = A$; $b_1 \dots b_q = B$, $\bar{\mu}$, σ^2 IS EQUIVALENT TO MAX $\log L_*$

$$\log L_*(\theta) = -\frac{T}{2} \ln 2\pi - \frac{T}{2} \ln \sigma^2 - \frac{S_*(A, B, \bar{\mu})}{2\sigma^2}$$

WHERE $S_*(A, B, \bar{\mu}) = \sum_{t=1}^{T} \varepsilon_t^2(A, B, \bar{\mu} \mid X_*, \varepsilon_*)$

MAXIMIZE $\log L_*(\theta) \Longleftrightarrow$ MINIMIZE $S_*(A, B, \bar{\mu})$

$$\Longrightarrow \hat{A}, \hat{B}, \hat{\bar{\mu}}$$

NEXT FIND $\hat{\sigma}$: MAX $\log L_*$ W.R. TO σ^2

$$-\frac{T}{2} \frac{1}{\hat{\sigma}^2} + \frac{S_*(\hat{A}, \hat{B}, \hat{\bar{\mu}})}{2\hat{\sigma}^2} = 0$$

$$\Rightarrow \hat{\sigma}^2 = \frac{1}{T} S_*(\hat{A}, \hat{B}, \hat{\bar{\mu}})$$

SINCE WE REPLACED THE INITIAL CONDITIONS BY THE FIRST p OBS WE IN FACT HAVE

$$S_* = \sum_{t=p+1}^{T} \varepsilon_t^2 (A, B, \bar{\mu} \mid X_*, \varepsilon_*)$$

2. BACKCASTING (TSP)

$$A(L)\, X_t = B(L)\, \varepsilon_t + \bar{\mu}$$

IS EQUIVELENT TO

$$A\left(\tfrac{1}{L}\right) X_t = B\left(\tfrac{1}{L}\right) e_t + \bar{\mu}$$

$$A(F)\, X_t = B(F)\, e_t + \bar{\mu}$$

$$X_t = a_1 X_{t+1} + \dots + a_p X_{t+p} + e_t - b_1 e_{t+1} - \dots - b_q e_{t+q} + \bar{\mu}$$

$$e_t = X_t - a_1 X_{t+1} - \dots - a_p X_{t+p} + b_1 e_{t+1} + \dots + b_q \varepsilon_{t+q} - \bar{\mu}$$

$$\tilde{\varepsilon}_t = \tilde{X}_t - a_1 \tilde{X}_{t+1} - \dots - a_p \hat{X}_{t+p} + b_1 \hat{e}_{t+1} + \dots + b_q \hat{\varepsilon}_{t+q} - \bar{\mu}$$

START ITERATING AT THE "TOP" OF SAMPLE I.E AT $T \Rightarrow$ BACKWORD TOWARDS $t_1 = 1$. PREDICT THE ERRORS AND $\tilde{X}$ PRIOR TO THE SAMPLE.

$$\tilde{\varepsilon}_t = \tilde{X}_t - a_1 \tilde{X}_{t-1} - \dots - a_p \tilde{X}_{t-p} + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q} - \bar{\mu}$$

MLE BASED ON THE AUGMENTED SAMPLE

Newbold (Biometrika 1974, p 423–426)

TESTS :

$$\frac{\sqrt{T}\,(\hat{\theta}_K - \theta_{0K})}{\sqrt{\hat{var}\ \hat{\theta}_K}} \underset{A}{\sim} N(0,1)$$, where θ_K is the K^{th} component of θ.

WALD TEST

$$\xi_T^W = T\left[\hat{\theta}_T - \theta^0\right]\left[var\ \hat{\theta}_T\right]^{-1}\left[\hat{\theta}_T - \theta^0\right] \underset{A}{\sim} \chi^2(r)$$

FOR A SCALAR $\hat{\theta}_K$ WE TEST $H_0 : \hat{\theta}_K = 0$:

$$\frac{\hat{\theta}_K}{\sqrt{\hat{var}(\hat{\theta}_K)}}$$ ~ using a $N(0,1)$ or a t student approximation

CONFIDENCE INTERVALS:

$$P\left[-1.96 \leq \frac{\hat{\theta}_K - \theta_K^0}{\hat{var}\ \hat{\theta}_K} \leq 1.96\right] = 0.95$$

$$P\left[\hat{\theta}_K - 1.96\,\hat{\theta}_K \leq \theta_K^0 \leq \hat{\theta}_K + 1.96\,\hat{\theta}_K\right] = 0.95$$

LAGRANGE MULTIPLIER

$$\xi_T^S = \frac{1}{T}\,\frac{\partial \log L(\hat{\theta}_T^C)}{\partial \theta}\,I(\hat{\theta}_T)^{-1}\,\frac{\partial \log L(\hat{\theta}_T^C)}{\partial \theta} \underset{A}{\sim} \chi^2(r)$$

$\log L(\hat{\theta}_T^C)$ IS THE CONSTRAINED LIKELIHOOD EVALUATED AT $\hat{\theta}_T$.
WE HAVE r constraints.

LIKELIHOOD RATIO TEST:

$$\xi_T^2 = 2\left[\log L(\hat{\theta}_T) - \log L(\hat{\theta}_T^C)\right] \underset{A}{\sim} \chi^2(r)$$

RECALL

$$\log L(\hat{A}, \hat{B}, \hat{\mu}, \hat{\sigma}^2) \simeq -\frac{T}{2}\log\hat{\sigma}^2 - \frac{T}{2}$$

$$-2\left[L_C - L_{NC}\right] \to \chi^2(r)$$

$$-2\left[-\frac{T}{2}\log\hat{\sigma}_C^2 + \frac{T}{2}\log\hat{\sigma}_{NC}^2\right] = T\log\left[\frac{\hat{\sigma}_C^2}{\hat{\sigma}_{NC}^2}\right]$$

MODEL SELECTION.

$$A(L)\,X_t = \bar{\mu} + B(L)\,\varepsilon_t$$

Let $\hat{\sigma}_x^2 = \sum_{t=1}^{T}(x_t - \bar{x})^2/T$ the empirical variance of X

$$\hat{\sigma}_T^2 = \frac{\sum \hat{\varepsilon}_t^2}{T}$$

1. MAXIMIZE

$$R^2 = 1 - \frac{\hat{\sigma}_T^2}{\hat{\sigma}_X^2} \qquad \left(\frac{ESS}{TSS}\right)$$

The more parameters are introduced, the larger the maximized log likelihood L AND $R^2 \uparrow$. BUT FOR EACH INCREASE IN THE NUMBER OF PARAMETERS, THE INCREASE IN L IS LESS AND LESS IMPORTANT

$\Rightarrow$ PENALIZE FOR TOO MANY PARAMETERS (PARSIMONY)

1. MAXIMIZE

$$\bar{R}^2 = 1 - \frac{S_T^2}{S_x^2}$$

WHERE $S_T^2 = \frac{T}{T-p-q} \hat{\sigma}_T^2 = \frac{\sum \hat{\varepsilon}_t^2}{T-p-q}$

$$S_x^2 = \sum_{t=1}^{T} \frac{(x_t - \bar{x})^2}{T-1}$$

KULLBACK INFORMATION BASED CRITERIA

let $F(w)$ be the density of the selected model and $F_0(w)$ the TRUE DENSITY. WE RELY ON THE KULLBACK DISTANCE

$$I(F/F_0) = \int \log\left[F_0(w)/F(w)\right] F_0(w)\, dw$$

$$= \underset{F_0}{E}\left\{\log\left[F_0(w)/F(w)\right]\right\}$$

$$= \underset{F_0}{E}\left\{\log\left[F_0(w)\right]\right\} - \underset{F_0}{E}\left\{\log\left[F(w)\right]\right\}$$

MIMMIZE $I()$ W.R TO F IS EQUIVALENT TO MIMMIZE $-E\{\log F(W)\}$. THIS IS A CHOICE CRITERION OF THE TYPE:

$$IC = -\frac{1}{T}\log(F) + \alpha(T)(p+q)$$

WHERE $\alpha(T)$ IS A DECREASING FUNCTION OF T. DEPENDING ON THE CHOICE OF $\alpha(T)$ WE HAVE:

(1) $\alpha(T) = 2/T$ Akaike (1969)

(2) $\alpha(T) = \log(T)/T$ Schwartz (1978)

(3) $\alpha(T) = c\log[\log(T)]/T$, $c>2$ Hannan-QUINN (1979)

a) $AIC(p,q) = -2\log(L) + 2(p+q)$

$= \log(\hat{\sigma}^2_T) + \frac{2(p+q)}{T}$

or $(p+q+1)$ IF THE CONSTANT IS INCLUDED

b) $BIC(p,q) = \log(\hat{\sigma}_T^2) + (p+q)\frac{\log T}{T}$

c) $\phi(p,q) = \log(\hat{\sigma}_T^2) + c(p+q)\frac{\log[\log(T)]}{T}$, $c>2$

DIAGNOSTICS

IF THE FIT IS CORRECT, THE RESIDUALS ARE W.N.

IF $u_1 \ldots u_T \sim NN(0, \sigma_u^2)$, the correlation estimator is

$$\hat{\rho}_K(u) = \sum_{t=1}^{T-K} u_t u_{t+K} \Big/ \sum_{t=1}^{T} u_t^2$$

FOR LARGE T, ASYMPTOTICALLY UNDER H_0: WHITE NOISE

$$\hat{\rho}_K(u) \underset{A}{\sim} N\left(0, \frac{1}{T}\right)$$

$$\frac{\hat{\rho}_K(u)}{1/\sqrt{T}} \sim N(0,1)$$

USE FOR TESTING INDIVIDUAL ρ'S.

- FOR A GROUP $\rho_K(u)$, $K = 1, \ldots, K$ to show that X's ARE UNCORRELATED, $K < T$

$$Q(\hat{\rho}) = \sum_{K=1}^{K} \left(\frac{\hat{\rho}_K(u)}{1/\sqrt{T}} \right)^2 = T \sum_{K=1}^{K} \hat{\rho}_K^2(u) \underset{A}{\sim} \chi^2(K)$$

FOR ARMA RESIDUALS ADJUST THE DEGREES OF FREEDOM

SAY $k = p+q$ OR $k = p+q+1$ (IF A CONSTANT IS ADDED)

$$Q(\hat{\rho}) = T \sum_{K=1}^{K} \hat{\rho}_K^2(u) \underset{A}{\sim} \chi^2(K-l)$$

BOX-PIERCE

OTHERWISE LJUNG BOX IS BETTER PERFORMING IN SMALL SAMPLES.

LJUNG-BOX:

$$Q = T(T+2) \sum_{K=1}^{K} \frac{\hat{\rho}_K^2}{T-K} \sim \chi^2(K-l)$$

$l = p+q$
or $l = p+q+1$