Chapter 3

Generalised AR and MA Models and Applications

3.1 Generalised Autoregressive Processes

Consider an AR(1) process given by

$$(1 - \alpha B)X_t = Z_t; \quad |\alpha| < 1.$$

In this case, the acf is, $\rho_k = \alpha^k$ for $k \geq 0$ and the pacf is:

$$\phi_k = \begin{cases} 1 & k = 0 \\ \alpha & k = 1 \\ 0 & k \geq 2 \end{cases}$$

with spectral density function:

$$f_X(\omega) = \frac{\sigma^2}{2\pi(1 - 2\alpha \cos\omega + \alpha^2)} \quad -\pi \leq \omega \leq \pi$$

If you consider some simulations of time series generated by

$$(1 - \alpha B)^\delta X_t = Z_t; \quad |\alpha| < 1, \ \delta > 0$$

you notice that the ACF, PACF and spectrum are similar for various valued of δ. Therefore it is difficult to identify these cases from the standard AR(1) case with $\delta = 1$.

Therefore we call the class of time series generated by $(1 - \alpha B)^\delta X_t = Z_t$ "Generalised first order autoregression" or GAR(1).

It can be seen that this parameter δ may increase the accuracy of forecast values. We now discuss some properties of this new class of GAR(1).

3.1.1 Some basic results

1. **Gamma Function**

$$\Gamma(x) = \begin{cases} \int_0^\infty t^{x-1}e^{-t}dt & x > 0 \\ \text{undefined} & x = 0 \\ x^{-1}\Gamma(x+1) & x < 0 \end{cases}$$

   It can be seen that $\Gamma(x) = (x-1)\Gamma(x-1)$ so when x is an integer, $\Gamma(x) = (x-1)!$. Furthermore, $\Gamma(\frac{1}{2}) = \pi$.

2. **Beta Function**

$$B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}dt$$

   It can be shown that

$$B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$$

3. **Binomial Coefficients** for any $n \in \mathbb{R}$:

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 1} = \frac{\Gamma(n+1)}{\Gamma(k+1)\Gamma(n-k+1)}$$

   When n is a positive integer,

$$\binom{n}{k} = {}^nC_k = \frac{n!}{(n-k)!k!}$$

   Note that you can't define factorials when n is not an integer.

4. **Hypergeometric Function**

$$F(a,b;c;x) = 1 + \frac{ab}{1!c}x + \frac{a(a+1)b(b+1)}{2!c(c+1)}x^2 + \frac{a(a+1)(a+2)b(b+1)(b+2)}{3!c(c+1)(c+2)}x^3 + \ldots$$

   We can re-express this using Pochhammer notation:

$$(a)_j = a(a+1)\cdots(a+j-1) \quad \text{and} \quad (a_0) = 1$$

   then

$$F(a,b;c;x) = \sum_{j=0}^{\infty} \frac{(a)_j(b)_j}{j!(c)_j}x^j$$

   Further, since $(a)_j = \frac{\Gamma(a+j)}{\Gamma(a)}$ we have

$$F(a,b;c;x) = \sum_{j=0}^{\infty} \frac{\Gamma(a+j)\Gamma(b+j)\Gamma(c)}{\Gamma(j+1)\Gamma(a)\Gamma(b)\Gamma(c+j)}x^j$$

Now, our model is $(1-\alpha B)^\delta X_t = Z_t$ or equivalently, $X_t = (1-\alpha B)^{-\delta} Z_t$. So we can look at the general binomial expansion:

$$(1-x)^{-1} = 1 + x + x^2 + x^3 + \dots$$

and

$$(1-x)^n = \sum_{r=0}^{\infty} \binom{n}{r} (-x)^r$$

when n is an integer, the above expression cuts off at a particular point. For our purposes we have:

$$\begin{aligned}(1-\alpha B)^\delta &= \sum_{j=1}^{\infty} \binom{\delta}{j} (-\alpha B)^j, \quad \delta \in \mathcal{R} \\ &= \sum_{j=0}^{\infty} (-1)^j \binom{\delta}{j} (\alpha B)^j\end{aligned}$$

Let

$$\begin{aligned}\pi_j &= (-1)^j \binom{\delta}{j} \\ &= \frac{(-1)^j \delta(\delta-1)\cdots(\delta-j+1)}{j!} \\ &= \frac{(-\delta)(-\delta+1)\cdots(-\delta+j-1)}{j!} \\ &= \frac{\Gamma(j-\delta)}{\Gamma(-\delta)\Gamma(j+1)}\end{aligned}$$

Thus, $(1-\alpha B)^\delta X_t = Z_t$ is equivalent to the AR(∞) process:

$$\sum_{j=0}^{\infty} \pi_j \alpha^j X_{t-j} = Z_t$$

which can be expressed in equivalent MA(∞) form:

$$X_t = (1-\alpha B)^{-\delta} Z_t$$

we can derive the form of the weights as above, just changing $+\delta$ to $-\delta$:

$$X_t = \sum_{j=1}^{\infty} \psi_j \alpha^j Z_{t-j}$$

where

$$\psi_j = \frac{\Gamma(j+\delta)}{\Gamma(\delta)\Gamma(j+1)}$$

(i.e. π_j with $-\delta$ changed to $+\delta$).

3.1.2 Convergence

Theorem 3.1.1.

$$X_t = \sum_{j=0}^{\infty} \psi_j \alpha^j Z_{t-j}$$

converges in the mean square sense.

Proof.

$$\mathbb{E}(X_t^2) = \sigma^2 \sum_{j=0}^{\infty} (\psi_j \alpha^j)^2$$

Recall the ratio test (of convergence): $u_1, u_t, \ldots$, let

$$\lim_{j\to\infty} \frac{u_{j+1}}{u_j} = L$$

if $L < 1$ then $\sum u_j$ converges; $L = 1$ then there's no decision; and if $L > 1$ then $\sum u_j$ diverges.

Letting $u_j = \psi_j \alpha^j$, then

$$\begin{aligned}
\frac{u_{j+1}}{u_j} &= \frac{\psi_{j+1}}{\psi_j}\alpha \\
&= \frac{\Gamma(j+1+\delta)}{\Gamma(j+2)\Gamma(\delta)} \frac{\Gamma(j+1)\Gamma(\delta)}{\Gamma(j+\delta)}\alpha \\
&= \frac{j+\delta}{j+1}\alpha \\
&= \left(\frac{1+\frac{\delta}{j}}{1+\frac{1}{j}} \right) \alpha
\end{aligned}$$

So, we have,

$$\lim_{j\to\infty} \frac{u_{j+1}}{u_j} = \lim_{j\to\infty} \left(\frac{1+\frac{\delta}{j}}{1+\frac{1}{j}} \right) \alpha = \alpha$$

Thus, as $|\alpha| < 1$, the result follows. □

3.1.3 Autocovariance function

$$\begin{aligned}
\gamma_k &= \operatorname{cov}(X_t, X_{t-k}) \\
&= \mathbb{E}(X_t X_{t-k}) \\
&= \mathbb{E}\left[\left(\sum_{j=0}^{\infty} \psi_j \alpha^j Z_{t-j} \right) \left(\sum_{l=0}^{\infty} \psi_l \alpha^l Z_{t-k-l} \right) \right] \\
&= \sigma^2 \left(\psi_k \alpha^k + \psi_{k+1}\psi_1 \alpha^{k+2} + \psi_{k+2}\psi_2 \alpha^{k+4} + \ldots \right) \\
&= \sigma^2 \alpha^k \sum_{j=0}^{\infty} \psi_{j+k}\psi_j \alpha^{2j}
\end{aligned}$$

When $k = 0$, $\gamma_0 = \text{var}(X_t) = \sigma^2 \sum_{j=0}^{\infty} \psi_j^2 \alpha^{2j}$. Note that when $\delta = 1$, $\psi_j = \frac{\Gamma(j+\delta)}{\Gamma(j+1)\Gamma(\delta)} = 1$ and we collapse back to our standard case:

$$\gamma_0 = \text{var}(X_t) = \sigma^2 \sum_{j=0}^{\infty} \alpha^{2j} = \frac{\sigma^2}{1-\alpha^2}.$$

Recall the Hypergeometric function

$$F(\theta_1, \theta_2; \theta_3; \theta) = \sum_{j=0}^{\infty} \frac{\Gamma(\theta_1 + j)\Gamma(\theta_2 + j)\Gamma(\theta_3)\theta^j}{\Gamma(\theta_1)\Gamma(\theta_2)\Gamma(\theta_3 + j)\Gamma(j+1)}$$

Now, we have

$$\begin{aligned}
\gamma_0 &= \sigma^2 \sum_{j=0}^{\infty} (\psi_j)^2 \alpha^{2j} \\
&= \sigma^2 \sum_{j=0}^{\infty} \frac{\Gamma(j+\delta)}{\Gamma(j+1)\Gamma(\delta)} \frac{\Gamma(j+\delta)}{\Gamma(j+1)\Gamma(\delta)} \alpha^{2j} \\
&= \sigma^2 F(\delta, \delta; 1; \alpha^2)
\end{aligned}$$

I.e. we have a hypergeometric function with $\theta_1 = \theta_2 = \delta$, $\theta_3 = 1$ and $\theta = \alpha^2$.

Similarly, we can re-express the autocovariance function as:

$$\begin{aligned}
\gamma_k &= \sigma^2 \alpha^k \sum_{j=0}^{\infty} \psi_{k+j} \psi_j \alpha^{2j} \\
&= \sigma^2 \alpha^k \frac{\Gamma(k+\delta)}{\Gamma(\delta)\Gamma(k+1)} F(\delta, k+\delta; k+1, \alpha^2)
\end{aligned}$$

Combining these results we have the autocorrelation function:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\alpha^k \Gamma(k+\delta) F(\delta, k+\delta; k+1; \alpha^2)}{\Gamma(\delta)\Gamma(k+1) F(\delta, \delta; 1; \alpha^2)}$$

3.1.4 Spectral Density Function

Recall when we have $X_t = G(B)Z_t$,

$$f_X(\omega) = \left|G(e^{-i\omega})\right|^2 f_Z(\omega)$$

AR(1) case: $(1 - \alpha B)X_t = Z_t$ or $X_t = (1 - \alpha B)^{-1} Z_t$:

$$\begin{aligned}
f_X(\omega) &= \left|1 - \alpha e^{-i\omega}\right|^{-2} f_Z(\omega) \\
&= (1 - 2\alpha \cos\omega + \alpha^2)^{-1} f_Z(\omega)
\end{aligned}$$

I $\{Z_t\} \sim WN(0, \sigma^2)$ then

$$f_Z(\omega) = \frac{\sigma^2}{2\pi} \quad -\pi < \omega < \pi.$$

GAR(1) case: $(1-\alpha B)^\delta X_t = Z_t$ and we have:

$$f_X(\omega) = (1-2\alpha\cos\omega+\alpha^2)^{-\delta}\frac{\sigma^2}{2\pi} \qquad -\pi < \omega < \pi.$$

Thus, the basic properties (turning points, etc.) don't change. Therefore, it's hard to identify a δ class from an integer class as they have very similar properties.

3.2 Parameter Estimation

3.2.1 Maximum Likelihood Estimation

Need to give a distribution to $\{Z_t\}$ rather than just white noise, therefore we assume $\{Z_t\} \sim NID(0,\sigma^2)$.

$$L = (2\pi)^{-\frac{n}{2}}\big|\Sigma\big|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\mathbf{X}^T\Sigma^{-1}\mathbf{X}\right\}$$

where $\mathbf{X} = (X_1,\dots,X_n)^T$, $\Sigma = \mathbb{E}(\mathbf{X}\mathbf{X}^T)$. We can do MLE by numerical maximisation of L which gives ML estimates.

3.2.2 Peridogram Estimation

An alternative method of estimating α, δ and σ^2 is to take

$$\log f_X(\omega) = -\delta\log(-2\alpha\cos\omega+\alpha^2)+C$$

and let $I_X(\omega)$ be an estimate (based on the periodogram) of $f_X(\omega)$. Then we have

$$\log f_X(\omega)+\log I_X(\omega) = -\delta\log(-2\alpha\cos\omega+\alpha^2)+C+\log I_X(\omega)$$

so:

$$\log I_X(\omega) = -\delta\log(-2\alpha\cos\omega+\alpha^2)+\log\left(\frac{I_X(\omega)}{f_X(\omega)}\right)+C.$$

Now, because $I_X(\omega)$ are uncorrelated random variables, we can write this as a regression type equation by letting

$$\begin{aligned} y_j &= \log I_X(\omega_j) \\ x_j &= \log(1-2\alpha\cos\omega+\alpha^2) \\ b &= -\delta \\ e_j &= \log\left(\frac{I_X(\omega)}{f_X(\omega)}\right) \\ a &= c \end{aligned}$$

Putting the above together we have: $y_j = a + bx_j + e_j$.

We can then estimate parameters by minimising the sum of squared errors:

$$\sum e_j^2 = \sum (y_j - a - bx_j)^2$$

but we can differentiate as out parameters are embedded, therefore we use numerical minimisation. In other words, since there is no explicit solution, we need to use the numerical method.

3.2.3 Whittle Estimation Procedure

See Shitan and Peris (2008).

3.3 Generalised Moving Average Models

Standard MA(1) models are defined as: $X_t = Z_t + \theta Z_{t-1} = (1 - \theta B)Z_t$. We didn't go into much detail, but we can write the Generalised Moving Average process of order 1, GMA(1) as:

$$X_t = (1 - \theta B)^\delta Z_t$$

We can get all of our results, as with the GAR(1) by changing the sign of δ and letting $\theta = -\alpha$.

Chapter 4

Analysis and Applications of Long Memory Time Series

Let $\{X_t\}$ be a stationary time series and let ρ_k be it's ACF at lag k. $\{X_t\}$ is said to be

1. **No memory** if $\rho_k = 0$ for all $k \neq 0$.
2. **Short memory** if $\rho_k \sim \phi^k$; $|\phi| < 1$. That is, ρ_k decays to zero at an exponential rate. In this case, it is clear that there exists a constant $A > 1$ such that $A^k \rho_k \to c$ as $k \to \infty$. Equivalently, $|\rho_k| \leq c\phi^k$ and hence $\sum |\rho_k| < \infty$.
3. **Long memory** if $\rho \sim k^{-\delta}$, $0 < \delta < 1$. That is, ρ_k decays slowly at the hyperbolic rate. In this case, $k^\delta \rho_k \to c$ as $k \to \infty$.

4.1 Properties of Long Memory Time Series

With an acf, $\rho_k \sim k^{-\delta}$, $0 < \delta < 1$, we can check if $\sum |\rho_k|$ is convergent by recalling that

$$\sum \frac{1}{r^p} \text{ is convergent for } p > 1 \text{ and divergent for } p \leq 1.$$

equivalently:

$$\sum r^{-p} \text{ is convergent for } p > 1 \text{ and divergent for } p \leq 1.$$

Thus

$$\sum |\rho_k| \sim \sum k^{-\delta} \text{ is divergent as } 0 < \delta < 1.$$

Thus $\sum |\rho_k| = \infty$, however, $\rho_k \to 0$ as $k \to \infty$ so we still have a stationary process.

Note that the acf gives an indication of the memory type of the underlying process. We can also look at the spectral properties. The sdf of a stationary process $\{X_t\}$ is:

$$\begin{aligned} f_X(\omega) &= \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\rho_k e^{i\omega k}; \quad -\pi < \omega < \pi \\ &= \frac{1}{2\pi}\left[1 + 2\sum_{k=1}^{\infty}\rho_k \cos\omega k\right] \end{aligned}$$

1. **No memory** processes have $\rho_k = 0$ for $k \neq 0$ which implies that

$$f_X(\omega) = \frac{1}{2\pi}; \quad -\pi < \omega < \pi$$

   Thus the spectrum is uniformly distributed over $(-\pi, \pi)$.

2. **Short memory** processes such as ARMA type processes have an sdf for all ω. If we look at the sdf at $\omega = 0$,

$$f_X(0) = \frac{1}{2\pi}\left[1 + 2\sum_{k=1}^{\infty}\rho_k\right] = c < \infty$$

   So for short memory type processes, we get a bounded point on the $f_X(\omega)$ axis.

3. **Long memory** type processes have a discontinuity at $\omega = 0$:

$$f_X(0) = \frac{1}{2\pi}\left[1 + 2\sum_{k=1}^{\infty}\rho_k\right] = \infty$$

   This the sdf is unbounded at $\omega = 0$.

Note that typically we plot with ω on the x axis and $f_X(\omega)$ on the y axis, so a short memory process will be continuous when crossing the y axis, but long memory processes will only approach the y axis asymptotically.

4.2 Identification

1. Calculate the sample acf, r_k.

2. Calculate the sample sdf or periodogram using

$$I_n(\omega) = \frac{1}{2}\left|\sum_{t=1}^{n} x_t e^{-it\omega_j}\right|^2; \quad -\pi < \omega_j = \frac{2\pi j}{n} < \pi$$

3. **Hurst coefficient.** As famous hydrologist, Hurst (1951), notices that there is a relationship between the rescaled adjusted range, R, and long memory, or persistent time series.

The rescaled adjusted range is defined as

$$R = \max_{1\le i\le n}\left[Y_i - Y_1 - \frac{i}{k}(Y_n - Y_1)\right] - \min_{1\le i\le n}\left[Y_i - Y_1 - \frac{i}{k}(Y_n - Y_1)\right]$$

where $Y_i = \sum_{t=1}^{i} X_t$. We further define the adjusted standard deviation as

$$S = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

The plots of $\log\left(\frac{R}{S}\right)$ and $\log n$ were observed for many time series and it was noticed that for persistent time series

(a) the points were scattered around a straight line

(b) the slope was always greater than a half.

Therefore we can write a regression line:

$$\log\left(\frac{R}{S}\right) = a + H\log(n)$$

where the slope coefficient H is called the Hurst coefficient. An estimate is:

$$\hat{H} = \frac{\log\left(\frac{R}{S}\right)}{\log n}$$

4.3 Fractional Integration and Long Memory Time Series Modelling

In many time series problems, we usually consider the order of differencing, d, as either 0, 1 or 2. That is the dth difference of $\{X_t\}$ is stationary. Let

$$Y_t = (1-B)^d X_t$$

A notation: $X_t \sim I(d)$. If $\{X_t\} \sim I(0)$ then no differencing is required to get a stationary series and the acf decays exponentially. If $X_t \sim I(1)$ then X_t is non stationary and

$$Y_t = a + bt + Z_t$$

so $\mathbb{E}(Y_t - Y_{t-1}) = b$. Further if $X_t \sim I(2)$ this implies that

$$Y_t = a + bt + ct^2 + Z_t$$

Let $u_t = Y_t - Y_{t-1}$ then $v_t = u_t - u_{t-1}$ is a constant.

Thus, we cannot accommodate the hyperbolic decay of the acf by integer differencing. It has been noticed that when $-1 < d < 1$ and $d \neq 0$ or equivalently, $0 < |d| < 1$ one can accommodate this hyperbolic decay of the acf.

In particular we will show that when $0 < |d| < \frac{1}{2}$, $Y_t = (1-B)^d X_t$ represents a stationary, long memory time series.

Note that in general then $-1 < d < 1$, $(1-B)^d X_t$ is called fractional differencing (or integration).

4.3.1 FARIMA(p, d, q)

If we fit an ARMA(p, q) model to $\{Y_t\}$ in the form: $Y_t = (1 - B)^d X_t$ we get:

$$\Phi(B)Y_t = \Theta(B)Z_t$$

where $\{Z_t\} \sim WN(0, 1)$ and $-1 < d < 1$ we have what's called fractionally integrated ARIMA or FARIMA(p, d, q).

To develop the theory, we first look at FARIMA(0,d,0) or FDWN (fractionally differenced white noise) given by:

$$(1 - B)^d X_t = Z_t \quad \text{or} \quad \nabla^d X_t = Z_t$$

Note that this $\nabla^d X_t = Z_t$ is known as fractionally differenced white noise. Note that from here on take d as a fractional real number or $-1 < d < 1$.

Theorem 4.3.1. *1. When $d < \frac{1}{2}$, $\{X_t\}$ is stationary and equivalent to*

$$X_t = \nabla^{-d} Z_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$$

where

$$\psi_j = \frac{\Gamma(j + d)}{\Gamma(j + 1)\Gamma(d)}$$

2. When $d > -\frac{1}{2}$, $\{X_t\}$ is invertible and equivalent to

$$\sum_{j=0}^{\infty} \pi_j X_{t-j} = Z_t$$

where

$$\pi_j = \frac{\Gamma(j - d)}{\Gamma(j + 1)\Gamma(-d)}$$

Thus you need $-\frac{1}{2} < d < \frac{1}{2}$ otherwise, you still have fractional integration, but you'll have stationarity and not invertibility or invertibility without stationarity.

Proof. 1.

$$\text{var}(X_t) = \sigma^2 \sum_{j=0}^{\infty} \psi_j^2$$

Recall that

$$\frac{\Gamma(x + a)}{\Gamma(x + b)} \sim x^{a-b} \quad \text{for large } x$$

so $\psi_j \sim j^{d-1}$ for large j and therefore

$$\text{var}(X_t) \sim \sigma^2 \sum j^{2(d-1)}.$$

Now if we recall that

$$\sum r^{-p} \text{ is convergent for } p > 1 \text{ and divergent for } p \leq 1$$

we require $\text{var}(X_t) < \infty$ which will only occur if $-2(d - 1) > 1$ or $d < \frac{1}{2}$.

2. Similarly for invertibility, we need $\sum \pi_j^2 < \infty$ which only occurs when $d > -\frac{1}{2}$. $\square$

Exercise 4.3.1. *Show that the sdf of fractionally differenced white noise is given by*

$$f_X(\omega) = \frac{2^{-2d}}{\pi}\left[\sin\left(\frac{\omega}{2}\right)\right]^{-2d}$$

Recall that if you have X_t and Y_t stationary, with $Y_t = G(B)X_t$ then

$$f_Y(\omega) = \left|G(e^{-i\omega})\right|^2 f_X(\omega)$$

so with $X_t = (1-B)^{-d}Z_t = G(B)Z_t$ we have

$$\begin{aligned}
f_X(\omega) &= \left|1 - e^{-i\omega}\right|^{-2d} f_Z(\omega) \\
&= \left|(1-\cos\omega) + i\sin\omega\right|^{-2d} f_Z(\omega) \\
&= \left(\left|(1-\cos\omega) + i\sin\omega\right|^2\right)^{-d} f_Z(\omega) \\
&= \left((1-\cos\omega)^2 + \sin^2\omega\right)^{-d} f_Z(\omega) \\
&= (2 - 2\cos\omega)^{-d} f_Z(\omega) \\
&= 2^{-d}\left(2\sin^2\frac{\omega}{2}\right)^{-d} f_Z(\omega) \\
&= 2^{-2d}\left(\sin^2\frac{\omega}{2}\right)^{-d} f_Z(\omega)
\end{aligned}$$

Note that we've used the trigonometric identity: $\cos 2A = 1 - 2\sin^2 A$.

Note that for $0 < d < \frac{1}{2}$, $f_X(\omega) \to \infty$ as $\omega \to 0$. We can also see this by noting that for small ω, $\sin\frac{\omega}{2} \approx \frac{\omega}{2}$ so near zero, $f_X(\omega) \sim \omega^{-2d} f_Z(\omega)$ so as $\omega \to 0$ we have a discontinuity.

Note that when $-\frac{1}{2} < d < 0$, $f_X(\omega)$ exists near $\omega = 0$, therefore the series has short memory behaviour.

It is clear that, a FARIMA process has long memory properties when $0 < d < \frac{1}{2}$.

Exercise: Let γ_k be the acf at lag k of a FDWN. Show that

1. $\gamma_k = \frac{(-1)^k\Gamma(1-2d)}{\Gamma(-d+k+1)\Gamma(-d-k+1)}$.
2. $\rho_k = \frac{\Gamma(k+d)\Gamma(1-d)}{\Gamma(k-d+1)\Gamma(d)}$.

Exercise: Find the pacf at lags 1 and 2 of of a FDWN.

Ans: $\frac{d}{1-d}$, $\frac{d}{2-d}$.

4.3.2 Estimation of d

First note that

$$f_X(\omega) = \left(4\sin^2\frac{\omega}{2}\right)^{-d} f_Z(\omega)$$

therefore taking logs we have

$$\log\big(f_X(\omega)\big) = -d\log\left(4\sin^2\frac{\omega}{2}\right) + \log\big(f_Z(\omega)\big)$$

Let $I_X(\omega)$ be the periodogram based on n observations. Now,

$$\begin{aligned}\log\big(I_X(\omega)\big) &= -d\log\left(4\sin^2\frac{\omega}{2}\right) + \log I_X(\omega) - \log f_X(\omega) + \log f_Z(\omega)\\ &= \log f_Z(0) - d\log\left(4\sin^2\frac{\omega}{2}\right) + \log\left(\frac{I_X(\omega)}{f_X(\omega)}\right) + \log\left(\frac{f_Z(\omega)}{f_Z(0)}\right)\end{aligned}$$

Since $Z_t \sim WN(0,\sigma^2)$, $f_Z(\omega) = \frac{\sigma^2}{\pi}$ for $0 < \omega < \pi$. So in this case, $\frac{f_Z(\omega)}{f_Z(0)} = 1$ and $\log(1) = 0$, so we can drop this term.

In general this will hold for ω near 0. So, letting

$$\begin{aligned}y_j &= \log I_X(\omega_j)\\ x_j &= \log\left(4\sin^2\frac{\omega_j}{2}\right)\\ b &= -d\\ \varepsilon_j &= \log\left(\frac{I_X(\omega_j)}{f_X(\omega_j)}\right)\\ a &= \log f_Z(0)\end{aligned}$$

where $\omega_j = \frac{2\pi j}{n}$ for $j = -(n-1), \ldots, (n-1)$. Thus we can write

$$y_j = a + bx_j + \varepsilon_j$$

where the distribution of ε_j is unknown. In practice we set $j = 1, 2, \ldots, m < n$.

4.3.3 Distribution of ε_j

We've seen that for **short memory** time series,

$$I_X(\omega_j) \overset{asy}{\sim} \frac{1}{2} f_X(\omega_j)\chi_2^2; \quad j \neq 0, \left[\frac{n}{2}\right]$$

Theorem 4.3.2. *Let X_t be a process generated by $(1-B)^d X_t = Z_t$ with $-\frac{1}{2} < d < 0$. Then for large n,*

$$\log\left(\frac{I_X(\omega_j)}{f_X(\omega_j)}\right); \quad j \neq 0, \left[\frac{n}{2}\right]$$

follows an independent Gumble distribution with mean $\delta = 0.577216$ *(Eulers constant) and variance* $\frac{\pi^2}{6}$*. Note that* Y *is said to be a Gumble distribution with parameters* α *and* β *if the cdf of* Y *is:*

$$F_Y(y) = \exp\left\{-e^{-\left(\frac{y-\alpha}{\beta}\right)}\right\}$$

In this case, $\mathbb{E}(Y) = \alpha + \beta\delta$ *and* $var(Y) = \frac{\pi^2\beta}{6}$*.*

Proof. When $d < 0$,

$$\frac{2I_X(\omega_j)}{f_X(\omega_j)} \sim \chi_2^2 \sim \exp\left(\frac{1}{2}\right)$$

with pdf $L_X(x) = \frac{1}{2}e^{-\frac{x}{2}}$ and $\mathbb{E}(X) = \frac{1}{2}$ and $\mathrm{var}(X) = \frac{1}{4}$. Let

$$U = -\log\left(\frac{I_X(\omega_j)}{f_X(\omega_j)}\right)$$

then,

$$\begin{aligned} F(u) = P(U \leq u) &= P\left(-\log\left(\frac{I_X(\omega_j)}{f_X(\omega_j)}\right) \leq u\right) \\ &= P\left(\frac{I_X(\omega_j)}{f_X(\omega_j)} \geq e^{-u}\right) \\ &= P\left(\frac{2I_X(\omega_j)}{f_X(\omega_j)} \geq 2e^{-u}\right) \\ &= \exp\left\{-e^{-u}\right\} \end{aligned}$$

Noting that if $X \sim \exp(\lambda)$ then $P(X > x) = e^{-\lambda x}$.

Thus we have a Gumble type distribution with $\alpha = 0$ and $\beta = 1$, therefore, $\mathbb{E}(U) = \delta$ and $\mathrm{var}(U) = \frac{\pi^2}{6}$, neither of which depend on j therefore we have an iid sequence and we can use standard regression techniques. □

So our model is:

$$y_j = a + bx_j + \varepsilon_j$$

Our least squares estimate of b_j is:

$$\hat{b} = \frac{\sum_{j=1}^{m}(x_j - \bar{x})y_j}{\sum_{j=1}^{m}(x_j - \bar{x})^2} \quad m < n$$

and

$$\hat{a} = \bar{y} - \hat{b}\bar{x}.$$

It is easy to see that

$$\mathrm{var}(\hat{b}) = \frac{\sigma^2}{\sum(x_j - \bar{x})^2} = \frac{\pi^2}{6\sum(x_j - \bar{x})^2}$$

Let $\hat{d}_p$ be the corresponding estimate of d based on the periodogram approach: $\hat{d}_p = -\hat{b}$.

Under certain conditions and large n, we have the usual result:

$$\frac{\hat{d}_p - d}{SE(\hat{d}_p)} \sim \mathcal{N}(0,1)$$

Now in the **general case** we have $X_t = (1-B)^{-d} Y_t$ with

$$f_X(\omega) = \left(4\sin^2\frac{\omega}{2}\right)^{-d} f_Y(\omega)$$

and

$$\log f_X(\omega) = -d\log\left(4\sin^2\frac{\omega}{2}\right) + \log f_Y(\omega)$$

thus

$$\log I_X(\omega) = -d\log\left(4\sin^2\frac{\omega}{2}\right) + \log\left(\frac{I_X(\omega)}{f_X(\omega)}\right) + \log f_Y(\omega)$$

Note that at this stage, we don't have a regression as we $\log f_Y(\omega)$ is not a constant, therefore we add and subtract $\log f_Y(0)$:

$$\log I_X(\omega) = \log f_Y(0) - d\log\left(4\sin^2\frac{\omega}{2}\right) + \log\left(\frac{I_X(\omega)}{f_X(\omega)}\right) + \log\left(\frac{f_Y(\omega)}{f_Y(0)}\right)$$

Now, $\log\left(\frac{f_Y(\omega)}{f_Y(0)}\right)$ is not a constant for all ω_j (as it was previously) but it is a constant when $\omega_j \approx 0$. Thus, this can be considered as a simple linear regression equation if we consider the values of ω near the origin.

In this case,

$$y_j = a + bx_j + \varepsilon_j; \quad \omega_j = \frac{2\pi j}{n}; \quad j = 1, 2, \ldots, m \ll n$$

In practice we usually take $m = n^\alpha$ with $0 < \alpha < 1$, a popular choice is $\alpha = 0.5$.

4.3.4 Estimation of d using a smoothed periodogram

We can also estimate d using a smoothed, or weighted, periodogram. Let

$$\hat{f}_s(\omega) = \frac{1}{2\pi}\sum_{|s|<m} \lambda(s)\hat{R}(s)e^{-i\omega s}$$

be a smoothed periodogram based on a suitable lag window. In this case,

$$\frac{\hat{f}_s(\omega)}{f(\omega)} \sim a\chi^2_\nu$$

where $\nu = \frac{2n}{\sum_{|s|<n}\lambda(s)}$ and $a = \frac{1}{\nu}$.

References:

V. Reisen (1994) J. Time Series Analysis 15(3) 335-350

Chen, Abraham and Peiris (1994) J. Time Series Analysis 15(5) 473-487

Generalized Fractional Processes