STAT 804 — Lecture 2

We are investigating assumptions on a discrete time process which will permit us to make reasonable estimates of the parameters. We will look for assumptions which guarantee at least the existence

Definition: A stochastic process $X_t; t = 0, \pm 1, \ldots$ is stationary if the joint distribution of $X_t, \cdots, X_{t+k}$ is the same as the joint distribution of $X_0, \cdots, X_k$ for all t and all k. (Often we call this **strictly** stationary.)

Definition: A stochastic process $X_t; t = 0, \pm 1, \ldots$ is **weakly** (or **second order**) stationary if

$$\mathrm{E}(X_t) \equiv \mu$$

for all t (that is the mean does not depend on t) and

$$\mathrm{Cov}(X_t, X_{t+h}) = \mathrm{Cov}(X_0, X_h) \equiv C_X(h)$$

is a function of h only (and does not depend on t).

Remark:

1. X finite variance, strictly stationary implies X weakly stationary.

2. X second order stationary and Gaussian implies X strictly stationary.

Definition: X is **Gaussian** if, for each $t_1, \ldots, t_k$ the vector $(X_{t_1}, \ldots, X_{t_k})$ has a Multivariate Normal Distribution.

Examples of Stationary Processes:

1) Strong Sense White Noise: A process ϵ_t is strong sense white noise if ϵ_t is iid with mean 0 and finite variance σ^2.

2) Weak Sense White Noise: ϵ_t is second order stationary with

$$\mathrm{E}(\epsilon_t) = 0$$

and

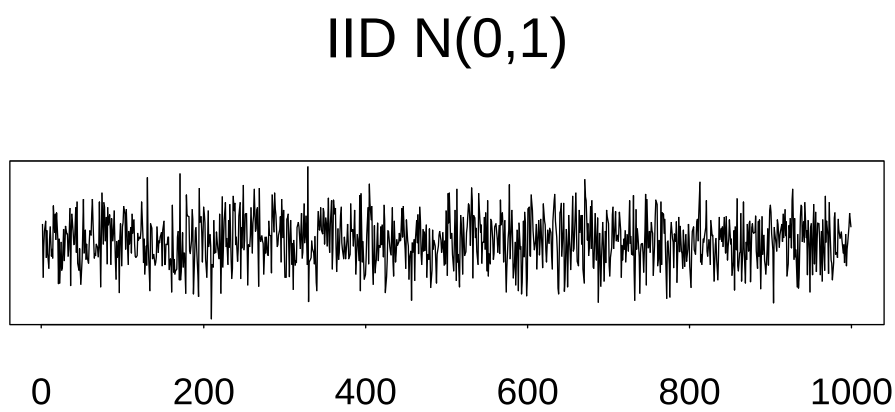
$$\mathrm{Cov}(\epsilon_t, \epsilon_s) = \begin{cases} \sigma^2 & s = t \\ 0 & s \neq t \end{cases}$$

Also called **Wide Sense** or **Second Order** white noise.

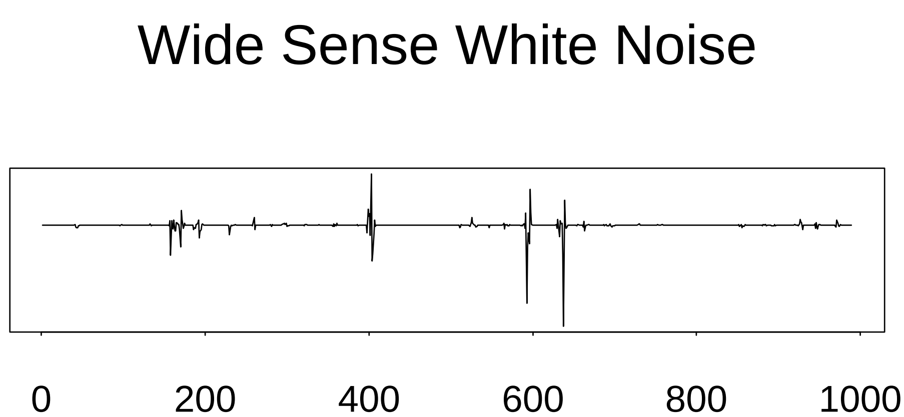
In this course we always use ϵ_t as notation for white noise and σ^2 as the variance of this white noise. We use subscripts to indicate variances of other things.

Example Graphics:

White noise: iid $N(0,1)$ data



White noise: $X_t = \epsilon_t \cdots \epsilon_{t+9}$



2) Moving Averages: if ϵ_t is white noise then $X_t = (\epsilon_t + \epsilon_{t-1})/2$ is stationary. (If you use second order white noise you get second order stationary. If the white noise is iid you get strict stationarity.)

Example proof: $\text{E}(X_t) = [\text{E}(\epsilon_t) + \text{E}(\epsilon_{t-1})]/2 = 0$ which is constant as required. Moreover:

$$\text{Cov}(X_t, X_s) = \begin{cases} \frac{\text{Var}(\epsilon_t)+\text{Var}(\epsilon_{t-1})}{4} & s = t \\ \frac{1}{4}\text{Cov}(\epsilon_t + \epsilon_{t-1}, \epsilon_{t+1} + \epsilon_t) & s = t+1 \\ \frac{1}{4}\text{Cov}(\epsilon_t + \epsilon_{t-1}, \epsilon_{t+2} + \epsilon_{t+1}) & s = t+2 \\ \vdots & \end{cases}$$

Most of these covariances are 0. For instance

$$\text{Cov}(\epsilon_t+\epsilon_{t-1}, \epsilon_{t+2}+\epsilon_{t+1}) = \text{Cov}(\epsilon_t, \epsilon_{t+2})+\text{Cov}(\epsilon_t, \epsilon_{t+1})+\text{Cov}(\epsilon_{t-1}, \epsilon_{t+2})+\text{Cov}(\epsilon_{t-1}, \epsilon_{t+1}) = 0$$

because the ϵs are uncorrelated by assumption. The only non-zero covariances occur for $s = t$ and $s = t \pm 1$. Since $\text{Cov}(\epsilon_t, \epsilon_t) = \sigma^2$ we get

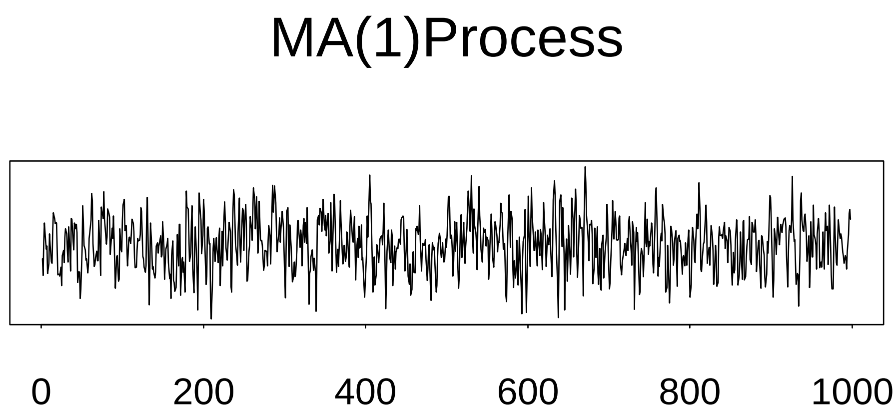
$$\text{Cov}(X_t, X_s) = \begin{cases} \sigma^2/2 & s = t \\ \frac{\sigma^2}{4} & |s-t| = 1 \\ 0 & \text{otherwise} \end{cases}$$

Notice that this depends only on $|s-t|$ so that the process is stationary.

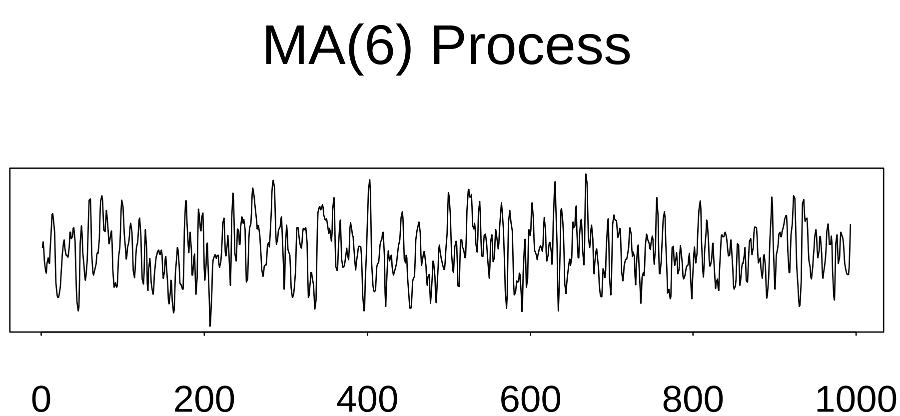
The proof that X is strictly stationary when the ϵs are iid is in your homework; it is quite different.

Example Graphics:

$$X_t = (\epsilon_t + \epsilon_{t-1})/2$$



$$X_t = \epsilon_t + 6\epsilon_{t-1} + 15\epsilon_{t-2} + 20\epsilon_{t-3} + 15\epsilon_{t-4} + 6\epsilon_{t-5} + \epsilon_{t-6}$$



The trajectory of X can be made quite smooth (compared to that of white noise) by averaging over many ϵs.

3) Autoregressive Processes:

An AR(1) process X is a process satisfying the equations:

$$X_t = \mu + \rho(X_{t-1} - \mu) + \epsilon_t \tag{1}$$

where ϵ is white noise. If X_t is second order stationary with $\mathrm{E}(X_t) = \theta$, say, then take expected values of (1) to get

$$\theta = \mu + \rho(\theta - \mu)$$

which we solve to get

$$\theta(1 - \rho) = \mu(1 - \rho)\,.$$

Thus either $\rho = 1$ (which will actually guarantee that X is not stationary) or $\theta = \mu$. Now we calculate variances:

$$\begin{aligned}\mathrm{Var}(X_t) &= \mathrm{Var}(\mu + \rho(X_{t-1} - \mu) + \epsilon_t)\\ &= \mathrm{Var}(\epsilon_t) + 2\rho\mathrm{Cov}(X_{t-1}, \epsilon_t) + \rho^2\mathrm{Var}(X_{t-1})\end{aligned}$$

We now **assume** that the meaning of (1) is that ϵ_t is uncorrelated with $X_{t-1}, X_{t-2}, \cdots$. In the strictly stationary case we are imagining that somehow X_{t-1} is built up out of past values of ϵ_s which are independent of ϵ_t. In the weakly stationary case we are imagining that X_{t-1} is actually a linear function of these past values. In either case this makes $\mathrm{Cov}(X_{t-1}, \epsilon_t) = 0$. If X is stationary so that $\mathrm{Var}(X_t) = \mathrm{Var}(X_{t-1}) \equiv \sigma_X^2$ then we find

$$\sigma_X^2 = \sigma^2 + \rho^2\sigma_X^2$$

whose solution is

$$\sigma_X^2 = \frac{\sigma^2}{1 - \rho^2}$$

Notice that this variance is negative or undefined unless $|\rho| < 1$. There is no stationary process satisfying (1) for $|\rho| \geq 1$.

Now for $|\rho| < 1$ how is X_t determined from the ϵ_s? (We want to solve the equations (1) to get an explicit formula for X_t.) The case $\mu = 0$ is notationally simpler. We get

$$\begin{aligned}X_t &= \epsilon_t + \rho X_{t-1}\\ &= \epsilon_t + \rho(\epsilon_{t-1} + \rho X_{t-2})\\ &\vdots\\ &= \epsilon_t + \rho\epsilon_{t-1} + \cdots + \rho^{k-1}X_{t-k+1} + \rho^k X_{t-k}\end{aligned}$$

Since $|\rho| < 1$ it seems reasonable to suppose that $\rho^k X_{t-k} \to 0$ and for a stationary series X this is true in the appropriate mathematical sense. This leads to taking the limit as $k \to \infty$ to get

$$X_t = \sum_{j=0}^{\infty} \rho^j \epsilon_{t-j} \, .$$

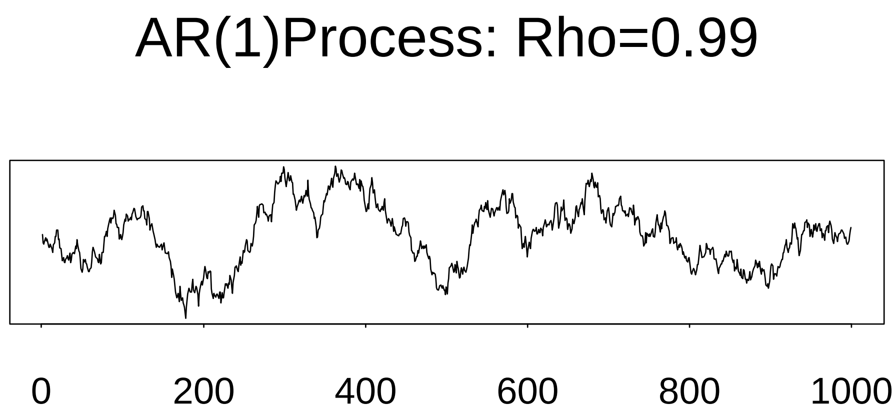
Claim: It is a theorem that if ϵ is a weakly stationary series then $X_t = \sum_{j=0}^{\infty} \rho^j \epsilon_{t-j}$ converges (technically it converges in mean square) and is a second order stationary solution to the equation (1). If ϵ is a strictly stationary process then under some weak assumptions about how heavy the tails of ϵ are $X_t = \sum_{j=0}^{\infty} \rho^j \epsilon_{t-j}$ converges almost surely and is a strongly stationary solution of (1).

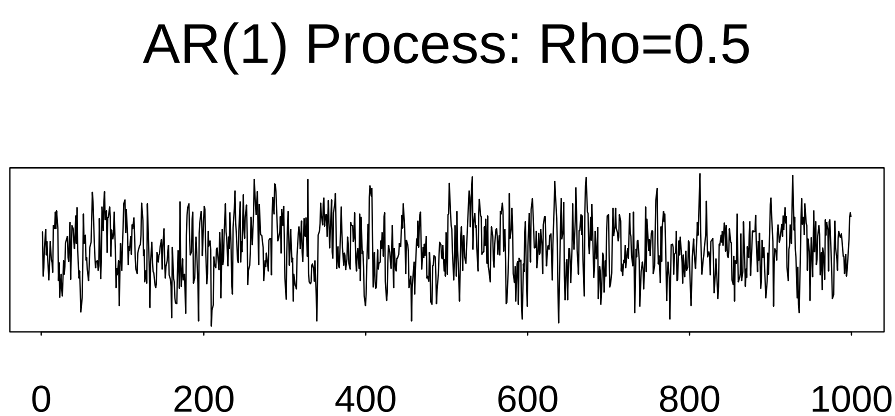
In fact if $\ldots, a_{-2}, a_{-1}, a_0, a_1, a_2, \ldots$ are constants such that $\sum a_j^2 < \infty$ and ϵ is weakly stationary (respectively strongly stationary with finite variance) then

$$X_t = \sum_{j=-\infty}^{\infty} a_j \epsilon_{t-j}$$

is weakly stationary (respectively strongly stationary with finite variance). In this case we call X a **linear filter** of ϵ.

Example Graphics:





Motivation of the jargon “filter” comes from physics. Consider an electric circuit with a resistance R in series with a capacitance C. We apply an “input” voltage $U(t)$ across the two elements and measure the voltage drop across the capacitor. We will call this voltage drop the “output” voltage and denote the output voltage by X_t. The relevant physical rules are these:

1. The total voltage drop around the circuit is 0. This drop is $-U(t)$ plus the voltage drop across the resistor plus $X(t)$. (The negative sign is a convention; the input voltage is not a “drop”.)

2. The voltage drop across the resistor is $Ri(t)$ where i is the current flowing in the circuit.

3. If the capacitor starts off with no charge on its plates then the voltage drop across its plates at time t is
$$X(t) = \frac{\int_0^t i(s)\,ds}{C}$$

These rules give
$$U(t) = Ri(t) + \frac{\int_0^t i(s)\,ds}{C}$$
Differentiate the definition of X to get
$$X'(t) = i(t)/C$$
so that
$$U(t) = RCX'(t) + X(t)\,.$$
Multiply by $e^{t/RC}/RC$ to see that
$$\frac{e^{t/RC}U(t)}{RC} = \left(e^{t/RC}X(t)\right)'$$
whose solution, remembering $X(0) = 0$, is obtained by integrating from 0 to s to get
$$e^{s/RC}X(s) = \frac{1}{RC}\int_0^s e^{t/RC}U(t)\,dt$$
leading to
$$\begin{aligned} X(s) &= \frac{1}{RC}\int_0^s e^{(t-s)/RC}U(t)\,dt \\ &= \frac{1}{RC}\int_0^s e^{u/RC}U(t-u)\,du \end{aligned}$$
This formula is the integral equivalent of our definition of filter and shows $X = \text{filter}(U)$.