Conceitos Básicos de Séries Temporais para Modelagem Macroeconômica

Material de apoio à aula de RBC

Referencia bibliográfica: *Introduction to Econometrics*

G S Maddala e Kajal Lahiri

4a. Edição, John Wiley & Sons Ltd, UK, 2009.

Séries Temporais

- O próposito da análise de séries temporais é estudar a **dinâmica, ie,** a **estrutura temporal** dos dados. A análise de uma única sequência de dados é chamada de análise de série temporal **univariada**, enquanto que a análise de várias coleções de dados para a mesma sequência de períodos de tempo é chamada de análise **multivariada** (***multivariate time-series analysis***).

- Desde a década de 60 que as novas técnicas em econometria (particularmente as que tentam corrigir correlações seriais nos resíduos) estão concentradas em modelar e entender dados de séries temporais.

Séries Temporais

- Uma série temporal é uma coleção de observações indexadas pela data de cada observação. Na prática temos a possibilidade de infinitas observações no tempo e pegamos apenas uma amostra de tamanho T, $\{y_1, y_2, \cdots, y_T\}$, isto é:

$$\{y_t\}_{t=-\infty}^{\infty} = \{\cdots, y_{-1}, y_0, \underbrace{y_1, y_2 \cdots, y_T}_{\text{amostra de tamanho } T}, \ldots\}$$

- A série temporal $\{y_t\}_{t=-\infty}^{\infty}$ é identificada pela descrição de seu t-ésimo elemento.

Exemplos: $y_t = \alpha$ para uma série onde todos os elementos são constantes e iguais a α; $y_t = t$ para um *time trend* que é uma série cujo valor de cada elemento é simplesmente a data da observação;

ou $y_t = \varepsilon_t$ para uma série cujos elementos são valores aleatórios e independentes com uma distribuição gaussiana.

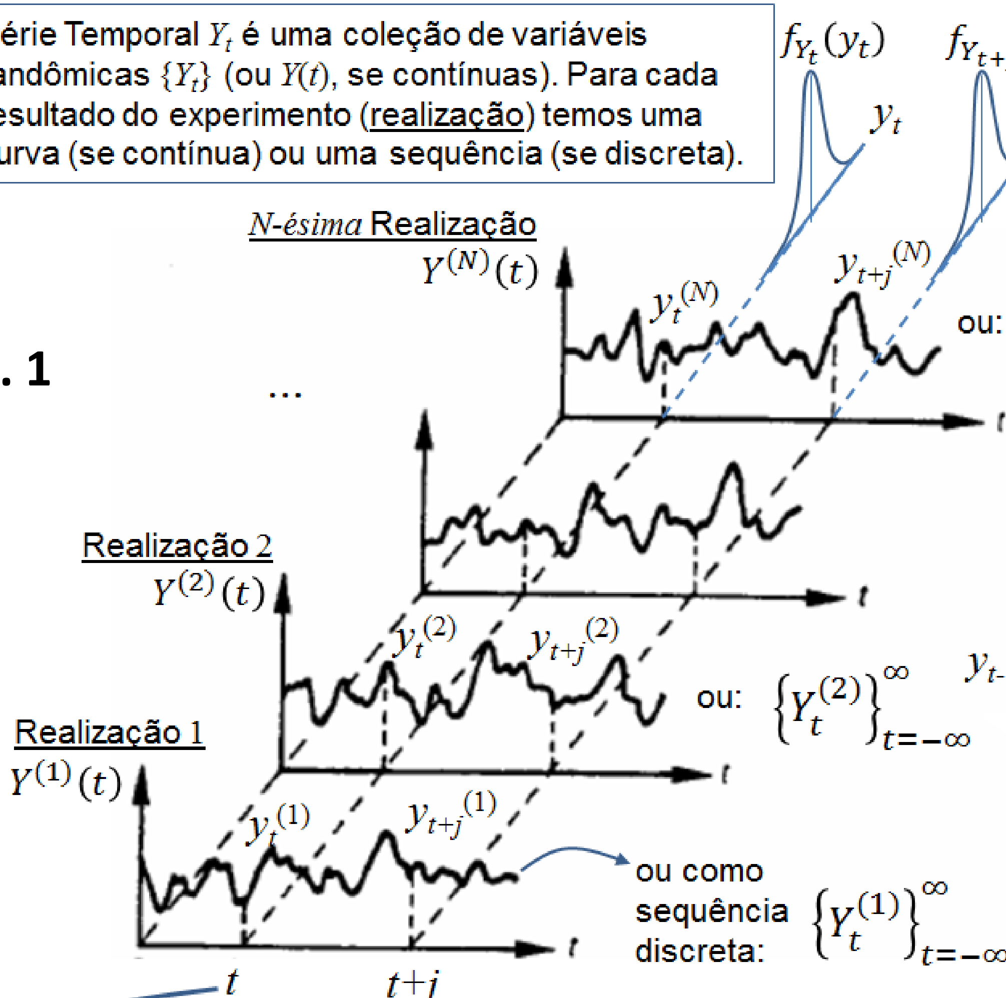
Processo Estocástico e Séries Temporais

Série Temporal Y_t é uma coleção de variáveis randômicas $\{Y_t\}$ (ou $Y(t)$, se contínuas). Para cada resultado do experimento (realização) temos uma curva (se contínua) ou uma sequência (se discreta).

O conjunto de curvas (ou sequências) é chamada de Processo Estocástico

A continuidade é tanto no conjunto de realizações como no tempo.

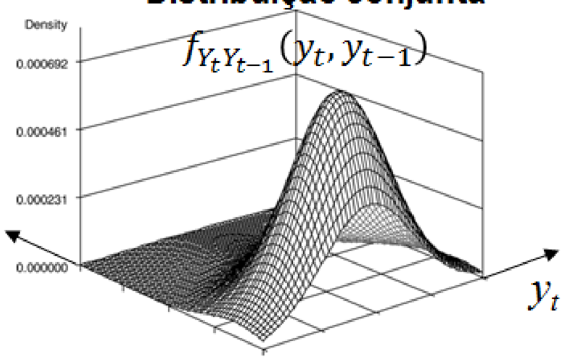
Fig. 1



ou: $\left\{Y_t^{(2)}\right\}_{t=-\infty}^{\infty}$

ou: $\left\{Y_t^{(2)}\right\}_{t=-\infty}^{\infty}$

Distribuição conjunta



ou como sequência discreta: $\left\{Y_t^{(1)}\right\}_{t=-\infty}^{\infty} = \left\{\cdots, y_{-1}^{(1)}, y_0^{(1)}, \underbrace{y_1^{(1)}, \cdots, y_T^{(1)}}, \ldots\right\}$

amostra de tamanho T de uma única realização

$$E(Y_t) = \underset{N\to\infty}{\text{plim}}(1/N)\sum_{i=1}^{N} y_t^{(i)} = \int_{-\infty}^{\infty} y_t f_{Y_t}(y_t)dy_t$$

Sob determinadas condições, se T $\to \infty$ então $\bar{y}$ converge para $E(Y_t)$

$$\bar{y} = (1/T)\sum_{t=1}^{T} y_t^{(1)}$$

Características Estatísticas das Séries Temporais

- A expectativa da t-ésima observação de uma série temporal se refere à média (estatística de primeira ordem, ou primeiro momento):

$$E(Y_t) = \mu_t = \int_{-\infty}^{\infty} y_t f_{Y_t}(y_t) dy_t$$

- Variância (estatística de segunda ordem):

$$\sigma_t^2 = var(y_t) = \gamma_{0t} = E[(Y_t - \mu_t)^2] = E(y_t^2) - \mu_t^2 = \int_{-\infty}^{\infty} (y_t - \mu_t)^2 \, f_{Y_t}(y_t) dy_t$$

- Um aspecto fundamental em séries temporais é como observações estão relacionadas entre si no tempo. A covariância é definida em termos de valores defasados (lagged), denominada de autocovariância, e mede o grau de variação de segunda ordem (segundo momento) entre dois elementos em dois tempos diferentes. A **autocovariância entre Y_t e Y_{t-1}** para o processo estocático $\{Y_t\}$ é:

$$cov(Y_t, Y_{t-1}) = \gamma_{1t} = E[(Y_t - \mu_t)(Y_{t-1} - \mu_{t-1})] = E(Y_t Y_{t-1}) - \mu_t \mu_{t-1}$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y_t - \mu_t)(y_{t-1} - \mu_{t-1}) f_{Y_t Y_{t-1}}(Y_t, Y_{t-1}) dy_t dy_{t-1}$$

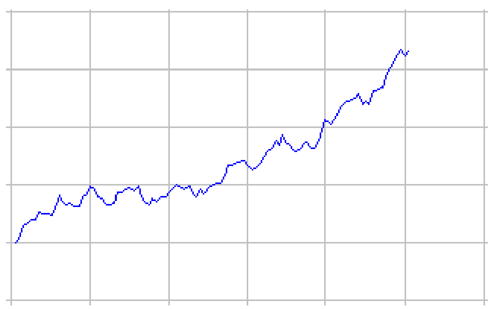
- Para considerar a relação de Y_t com as j mais recentes observações $\{Y_{t-1}, Y_{t-2}, \cdots, Y_{t-j}\}$, definimos a **$j$–ésima autocorrelação de Y_t**:

$$\gamma_{jt} = E\left[(Y_t - \mu_t)\left(Y_{t-j} - \mu_{t-j}\right)\right] = E\left(Y_t Y_{t-j}\right) - \mu_t \mu_{t-j}$$

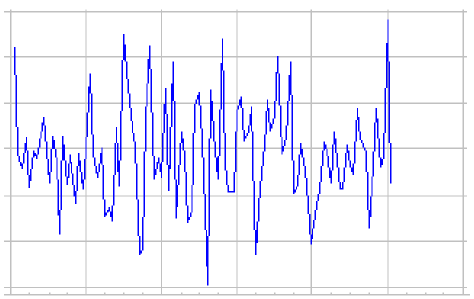
$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (y_t - \mu_t)\left(y_{t-j} - \mu_{t-j}\right) f_{Y_t Y_{t-1} \cdots Y_{t-j}}\left(Y_t, Y_{t-1}, \cdots, Y_{t-j}\right) dy_t dy_{t-1} \cdots dy_{t-j}$$

Estacionariedade (*Stationarity*)

- Uma série temporal é estacionária quando suas características estatísticas (média, variância, autocorrelação, …) são constantes ao longo do tempo. É uma série que se desenvolve aleatoriamente no tempo, em torno de uma média constante, refletindo alguma forma de equilíbrio estatístico estável (i.e. as leis de probabilidade que atuam no processo não mudam com o tempo).
- Muitas séries são inerentemente não-estacionárias e exibem *trends*, ciclos, padrões sazonais e outros comportamentos não-estacionários.
- Métodos de previsão (*forecasting*) usam transformações matemáticas para "estacionarizar" uma série e fazer previsões nesta série mais bem comportada, para depois inverter as transformações e obter as previsões para a série original. Transformações comuns são: tomar diferenças sucessivamente, deflacionar, aplicar log (p/estabilizar variâncias), fazer ajustes sazonais, *de-trending* (*e.g.* ajustar um "trend" e subtraí-lo), … .
- A inspeção visual de uma série já pode indicar não-estacionariedade (basta ver se não há uma média constante e uma variância constante, ou se existe um *trend*):



Série não-estacionária
(pois existe um *trend*)



Série estacionária

White Noise

- **Processo de Ruído Branco** (***White Noise***), também chamado de **Processo Puramente Randômico**, é um processo estocástico que consiste de uma sequência de variáveis randômicas <u>independente e identicamente distribuídas</u> (i.i.d.).

- W*hite noise* $\{\varepsilon_t\}_{t=-\infty}^{\infty}$ tem média e variância constantes e seus elementos são não-correlacionados.

$$E(\varepsilon_t) = 0$$
$$var(\varepsilon_t) = E(\varepsilon_t^2) - 0^2 = \sigma^2$$
$$E(\varepsilon_t \varepsilon_s) = 0; \quad t \neq s$$

- Indica-se *white noise* como: $\varepsilon_t \sim IID(0, \sigma^2)$ ou $\varepsilon_t \sim WN(0, \sigma^2)$.
- A condição de ε_t e ε_s serem independentes (para todo $t \neq$ s) é muito forte. Portanto, é comum definir white noise apenas com a condição de não-correlação, $E(\varepsilon_t \varepsilon_s) = 0$, e considerar <u>Processo de White Noise Independente</u> para a condição mais forte de independência.
- Se o white noise é independente e a distribuição é normal, denominamos de **Processo de *White Noise* Gaussiano** e indica-se como $\varepsilon_t \sim IN(0, \sigma^2)$..

Estacionariedade Fraca (Weak *Stationarity*)

- Um processo Y_t é **fracamente estacionário**, ou **estacionário**, se nem a média nem as j–ésimas autocovariâncias dependem da data t, isto é:

estacionariedade fraca também é chamada de **estacionariedade de covariância**.

$$E(Y_t) = \mu \quad \forall t$$
$$\gamma_{jt} = \gamma_j \quad \forall t\, \forall j$$
$$E(Y_t^2) < \infty \quad \forall t$$

A ***j*–ésima autocorrelação** de um processo estacionário, denotado por ρ_j é definido por: $\rho_j = \gamma_j / \gamma_0$ (o que vem da definição de $corr$)

- Exemplo: $\boxed{Y_t = c + \varepsilon_t}$, onde c é uma constante e ε_t é um *white noise* Gaussiano, é uma série estacionária (*white noise* sozinho também é estacionário).

$$E(Y_t) = E(c) + E(\varepsilon_t) = c + 0 = c$$
$$\gamma_{jt} = E\big[(Y_t - \mu_t)\big(Y_{t-j} - \mu_{t-j}\big)\big] = E\big[(c + \varepsilon_t - c)\big(c + \varepsilon_{t-j} - c\big)\big] = E\big(\varepsilon_t \varepsilon_{t-j}\big)$$
$$= \begin{cases} 0 & j \neq 0 \\ \sigma^2 & j = 0 \end{cases}$$

- Exemplo: Y_t é um <u>*time trend* linear</u> mais um *white noise* Gaussiano, *i.e.* $\boxed{Y_t = \beta t + \varepsilon_t}$, onde β é uma constante. Esta série <u>não é estacionária</u> porque a média é uma função do tempo:

$$E(Y_t) = E(\beta t) + E(\varepsilon_t) = \beta t + 0 = \beta t$$

- No conceito de estacionariedade está implícito que a <u>variância é finita</u> (*i.e.* não é uma estacionariedade explosiva). Nos dois exemplos acima, a variância é $var(y_t) = E[(Y_t - \mu_t)^2] = E(\varepsilon_t^2) = \sigma^2 < \infty$.

- Da definição $\gamma_{jt} = E\big[(Y_t - \mu)\big(Y_{t-j} - \mu\big)\big] = \gamma_j$ pode-se mostrar que se um processo é estacionário, então a covariância entre Y_t e Y_{t-j} dependerá somente da distância de tempo j que separa as observações e que: $\gamma_j = cov\big(Y_t, Y_{t+j}\big) = cov\big(Y_{t-j}, Y_t\big) \quad \forall t \forall j$

$$\gamma_j = \gamma_{-j} \quad \forall j$$

Estacionariedade Estrita (*Strict Stationarity*)

- Na literatura, geralmente estacionariedade significa estacionariedade fraca.
- Há uma condição bem mais forte chamada de estacionariedade estrita.
- Um processo é **estritamente estacionário** se, para quaisquer valores de $t_1, t_2, \ldots, t_n$, a distribuição conjunta de $\left(Y_{t_1}, Y_{t_2}, \cdots, Y_{t_n}\right)$ é a mesma distribuição conjunta de $\left(Y_{t_{1-j}}, Y_{t_{2-j}}, \cdots, Y_{t_{n-j}}\right)$ para todo o j. Isto é, a distribuição conjunta não depende das datas $t_1, t_2, \ldots, t_n$ mas apenas da distância separando as datas.
- Variância finita não é assumida na definição de estacionariedade estrita. Portanto, estacionariedade estrita não necessariamente implica em estacionariedade fraca.

- Na definição de estacionariedade fraca há apenas restrições com relação aos segundos momentos serem finitos. Portanto, uma série pode ser fracamente estacionária, mesmo tendo algum momento de alta ordem dependendo do tempo (e.g. $E(Y_t^3)$). Neste caso a série é fracamente estacionária, mas não é estritamente.

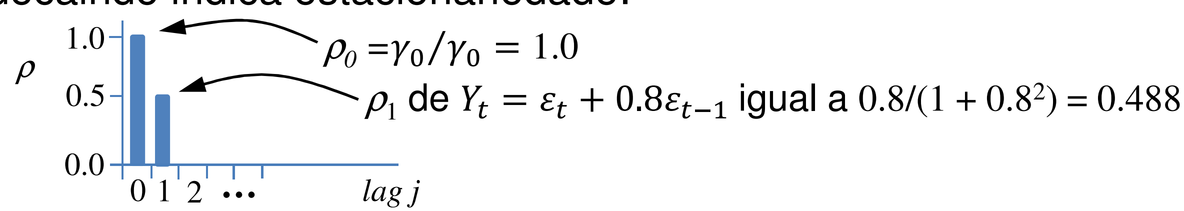
- *White noise* é estacionário mas pode não ser estritamente estacionário. Porém *white noise* Gaussiano é estacionário e estritamente estacionário.

Ergodicidade (*Ergodicity*)

- Na prática, geralmente só se tem uma única realização de um processo estocástico (i.e. uma única amostra de tamanho T). Séries temporais econômicas são únicas. Portanto, como mostra a Fig. 1 (slide 5), gostaríamos que a média no tempo ($\bar{y}$) fosse uma boa aproximação da média $E(Y_t)$ das inúmeras realizações. Se o tamanho T de uma única amostra tende a ∞, então $\bar{y}$ converge para $E(Y_t)$.
- *Def.*: Um processo fracamente estacionário é **ergódico para a média** , ou seja, **ergódico para o primeiro momento**, se $\bar{y}$ converge em probabilidade para $E(Y_t)$ quando $T \to \infty$.
- Um processo será ergódico para a média desde que a autocovariância γ_j vá a zero suficientemente rápido à medida que j cresce.
- É possível demonstrar que se $\sum_{j=0}^{\infty}|\gamma_j| < \infty$, então $\{Y_t\}$ é ergódico para a média.
- De maneira similar, também é possível definir ergodicidade para segundos momentos (lembrando que queremos trabalhar com uma única amostra e ainda assim ter as propriedades dos processos aleatórios garantidas).
- *White noise* Gaussiano é ergódico para todos os momentos.
- Estacionariedade não implica em ergodicidade. É fácil mostrar que a série estacionária $Y_t = \mu + \varepsilon_t$ não é ergódica para a média se μ é gerado a partir de uma distribuição normal $N(0, \lambda^2)$, mesmo que ε_t seja um *white noise* Gaussiano.

Processo MA – *Moving Average*

- *Def.*: um processo de médias móveis (***moving average***) **de primeira ordem** $MA(1)$ é uma série temporal do tipo $Y_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1}$, onde $\{\varepsilon_t\}$ é um *white noise* (sem a condição forte de independência) e μ e θ são constantes.
- O termo "médias móveis" vem do fato de que Y_t é construído a partir de uma soma ponderada (que é uma espécie de média) dos valores mais recentes de ε.
- A expectativa de Y_t é $E(Y_t) = \mu$. Usa-se o símbolo μ para o termo constante para enfatizar que o termo constante é a média do processo.
- Pode-se mostrar que a variância é finita e as autocovarianças são nulas, ou seja, média e autocovariâncias não são funções de t. Em outras palavras, $MA(1)$ é um processo estacionário (*i.e.* fracamente estacionário). E mais: $MA(1)$ é estacionário independentemente do valor de θ.
- Pode-se também mostrar que $\sum_{j=0}^{\infty}|\gamma_j| = (1+\theta^2)\sigma^2 + |\theta\sigma^2|$, ou seja $\sum_{j=0}^{\infty}|\gamma_j| < \infty$. Portanto, se $\{\varepsilon_t\}$ é um *white noise* Gaussiano, então $MA(1)$ é ergódico para todos os momentos. A primeira autocorrelação de um $MA(1)$ é $\rho_1 = \theta/(1+\theta^2)$; as outras são 0. Um correlograma que vai decaindo indica estacionariedade:

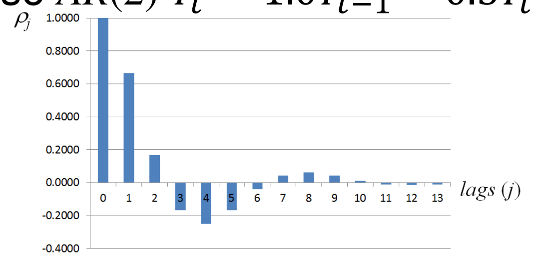


- Pode-se estender estes resultados para um processo de médias móveis **de q-ésima ordem** $MA(q)$ e mostrar que também se trata de um processo estacionário e ergódico para todos os momentos:

$$Y_t = \mu + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \cdots + \theta_q\varepsilon_{t-q}$$

Processos Autoregressivos AR

- *Def.*: um processo **autoregressivo de primeira ordem** $AR(1)$ é uma série temporal do tipo $Y_t = c + \varphi Y_{t-1} + \varepsilon_t$ onde $\{\varepsilon_t\}$ é um *white noise* (sem a condição forte de independência) e φ é uma constante.
- Pode ser mostrado que se $|\varphi| \geq 1$ então não existe uma solução estável, pois os ε vão se acumulando e a série explode. Por exemplo: $Y_t = \varphi Y_{t-1} + \varepsilon_t = \varphi(\varphi Y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \cdots = \varphi^t\ Y_0 + \varphi^{t-1}\varepsilon_1 + \cdots + \varphi^2\ \varepsilon_{t-2} + \varphi\varepsilon_{t-1} + \varepsilon_t$. Isto equivale a dizer que não existe um processo estacionário para Y_t quando $|\varphi| \geq 1$.
- *Def.*: um processo **autoregressivo de segunda ordem** $AR(2)$ é uma série temporal do tipo
$$Y_t = c + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \varepsilon_t$$
- Podemos mostrar que a condição para a solução desta série ser estável (e da série ser estacionária) é que as raízes z_1 e z_2 da equação característica $z^2 - \varphi_1 z - \varphi_2 = 0$ estejam dentro do círculo unitário, *i.e.* $|z_1| < 1$. Isto leva às seguintes três condições para estacionariedade: $\varphi_1 + \varphi_2 < 1$, $\varphi_1 - \varphi_2 > -1$, $|\varphi_2| < 1$.
- Podemos também calcular autocorrelações ρ_j (obtidos recursivamente usando as equações de Yule-Walker) para plotar o correlograma e evidenciar a estacionareidade. Para o processo $AR(2)$ $Y_t = 1.0Y_{t-1} - 0.5Y_{t-2} + \varepsilon_t$ temos $\rho_j = \rho_{j-1} - 0.5\rho_{j-2}$ e o correlograma:



Correlograma de série estacionária AR sempre amortece. Dependendo dos sinais dos coeficientes, há valores positivos e negativos no correlograma.

- Podemos estender estes resultados para uma **autoregressão de ordem p**, denominada $AR(p)$, que é estacionária se as p raízes da sua equação característica estiverem dentro do círculo unitário.
- A estimativa de modelos AR é bem simples: OLS minimizando $\sum \varepsilon_t^2$.

Processos de Médias Móveis Autoregressivos - *ARMA*

- O processo **autoregressive moving average** $ARMA(p,q)$ vem da combinação dos modelos $AR(p)$ e $MA(q)$:

$$Y_t = c + \underbrace{\varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \cdots \varphi_p Y_{t-p}}_{AR(p)} + \underbrace{\varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}}_{MA(q)}$$

- Exemplo $ARMA(1,1)$: $Y_t = c + \varphi Y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}$
- A estacionariedade de um processo $ARMA$ depende inteiramente dos parâmetros de autoregressão φ_i e não dos parâmetros de médias móveis θ_i: se todas as raízes da equação característica $1 - \varphi_1 z - \varphi_2 z^2 - \cdots \varphi_p z^p = 0$ estiverem dentro do círculo unitário, *i.e.* $|z_i| < 1$, então o modelo $ARMA$ é <u>estacionário</u> e a média $E(Y_t) = c/\left(1 - \varphi_1 - \cdots - \varphi_p\right)$.

- O modelo $ARMA$ combina a questão básica do AR (quão fortemente o passado influencia o presente) com a do MA (modelar o presente usando erros passados de predição). Há várias técnicas para determinar a ordem (p e q) de um modelo $ARMA$.

Processo *Random Walk*

- O modelo $AR(1)$ é um processo ***Random Walk*** quando $\varphi = 1$ e $c = 0$: $\boxed{y_t = y_{t-1} + \varepsilon_t}$
- Como $\varphi = 1$, então o processo não é estacionário. Este tipo de série é chamada de **série não-estacionária de raiz unitária**.
- A implicação de um processo deste tipo é que a melhor predição de y para o próximo período é o valor corrente. Este processo não permite predizer a mudança $\Delta y_t = y_t - y_{t-1}$. Em outras palavras, a mudança de y é absolutamente aleatória. A média é constante, mas a variância cresce com t.
- Na prática, a presença de um processo *random walk* torna o processo de *forecast* muito simples, porque todos os valores futuros de y_{t+s} é simplesmente y_t.
- O impacto de qualquer choque passado ε_{t-j} em y_t não decai ao longo do tempo. Consequentemente, a série *random walk* tem uma forte memória, porque ela se lembra dos choques passados; *i.e.* os choques têm um efeito permanente na série.
- No modelo $AR(1)$, quando $\varphi = 1$ e $c = 0$, temos um ***Random Walk* com *Drift***: $\boxed{y_t = y_{t-1} + \delta + \varepsilon_t}$, onde a constante δ é a média, $\delta = E(y_t - y_{t-1})$, representa um *time trend* e é referida como o *drift* do modelo. Este *trend* fica mais claro quando desenvolvemos a sequência a partir do valor inicial y_0 e apresentamos o modelo de outra forma:

$$y_1 = y_0 + \delta + \varepsilon_1$$
$$y_2 = y_1 + \delta + \varepsilon_2 = y_0 + 2\delta + \varepsilon_2 + \varepsilon_1$$
$$\dots$$
$$y_t = y_0 + t\delta + \varepsilon_t + \varepsilon_{t-1} + \dots + \varepsilon_1$$
$$\therefore \boxed{y_t = \alpha + \delta t + u_t}\ , u_t = \varepsilon_t + \varepsilon_{t-1} + \dots + \varepsilon_1$$
$$\alpha = constante$$

onde u_t é um ***Random Walk* Gaussiano**, *i.e.* um *random walk* com um tamanho aleatório de passo que varia de acordo com uma distribuição normal, e que tem $var(u_t) = t\sigma_\varepsilon^2$, onde σ_ε^2 é a variância de ε_t.

Imagine ir de um ponto numa reta para outro, dando passos para trás e para frente aleatoriamente.

ARIMA

- Se o modelo $ARMA$ permitir que o polinômio AR tenha 1 como uma das raízes da equação característica, tornamos o modelo $ARMA$ no modelo **autoregressive integrated moving average** ($ARIMA$). Um modelo $ARIMA$ é dito ser não-estacionário de raiz unitária porque seu polinômio AR tem uma raiz unitária. Assim como um modelo *random walk*, um modelo ARIMA tem forte memória porque os coeficientes θ_i da sua representação MA não decaem no tempo para zero, implicando que o choque passado ε_{t-i} do modelo tem um efeito permanente na série. Um tratamento usual para lidar com não-estacionaridade de raiz unitária é usar as diferenças sucessivas.

- Na prática, séries temporais são não-estacionárias. Um procedimento comum para converter uma série não-estacionária em uma série estacionária é através de diferenças sucessivas, aplicando o operador Δ. Suponha que fazemos primeiro $\Delta y_t = y_t - y_{t-1}$, depois $\Delta^2 y_t = \Delta y_t - \Delta y_{t-1} = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2})$, depois $\Delta^3 y_t = \cdots$, até que $\Delta^d y_t$ seja uma série estacionária que possa ser representada por um modelo $ARMA(p,q)$. Neste caso, dizemos que y_t pode ser representado por um modelo **autoregressive integrated moving average** $ARIMA(p,\mathbf{d},q)$ e que y_t é integrado de ordem $\mathbf{d}$. Neste caso denotamos $y_t \sim I(d)$. Uma série estacionária é dita ser $I(0)$. *White noise* é $I(0)$. *Random Walk* com *drift* é $I(1)$.

Trends e *Random Walk*

- Com os modelos dinâmicos das séries temporais, a questão do *time trend* ganha uma dimensão muito maior do que no caso clássico de modelo de regressão linear. Neste último, quando temos a reta que estima $y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$, o coeficiente β_1 pode assumir qualquer valor fixo. Em contraste, o modelo $AR(1)$ $y_t = \varphi_0 + \varphi_1 y_{t-1} + \varepsilon_t$ para ter significado deve ter o coeficiente $|\varphi_1| \leq 1$.
- Um modelo muito próximo do *random walk* com *drift* é o processo ***trend-stationary*** (TSP): $\boxed{y_t = \alpha + \delta t + u_t}$, onde u_t é um processo estacionário, *e.g.* um processo $AR(p)$ ou $MA(q)$. Por exemplo, um MA (∞): $u_t = \varepsilon_t + \varphi_1 \varepsilon_{t-1} + \varphi_2 \varepsilon_{t-2} + \cdots$.
- Nesta série, y_t cresce linearmente no tempo com uma taxa δ e portanto pode exibir comportamento similar ao de um modelo *random walk* com *drift*. Entretanto, há uma grande diferença. Supondo que y_0 é fixo, o modelo *random walk* com *drift* assume a média $E(y_t) = y_0 + \delta t$ e a variância $var(y_t) = t\sigma_\varepsilon^2$, ambas dependentes do tempo (i.e. é não estacionário). O modelo TSP também tem média $E(y_t) = y_0 + \delta t$ que depende do tempo, mas tem a variância $var(y_t) = var(u_t)$ que é finita e independente do tempo (pois u_t é um processo estacionário). Portanto, a série TSP pode ser transformada em uma série estacionária através da remoção do trend linear através de uma simples regressão linear. Já o modelo de *random walk* com *drift*, com o seu termo de perturbação não estacionário u_t (que cresce com o tempo), requer um procedimento de diferenças sucessivas para eliminar o *time trend* e se tornar estacionário. Por esta razão, o modelo de *random walk* com *drift* também é denominado de processo ***difference-stationary*** (DSP).
- Diz-se que TSP representa um *trend* determinístico, enquanto que o *random walk* com *drift*, *i.e.* o processo DSP, é um *trend* estocástico. Uma série pode ter componentes determinísticos (polinômio T para *trend* e polinômio S para sazonalidade) e estocásticos (processo u_t para *trend* e processo η_t para sazonalidade) juntos: $y_t = T_t + S_t + u_t + \eta_t + \varepsilon_t$

TSP vs DSP e RBC

- Usa-se o **teste Dickey-Fuller** (que é um tipo de **teste da raiz unitária**) para testar a hipótese de que uma série temporal pertence à classe *TSP* contra a alternativa de que ela pertence à classe *DSP* (*i.e. random walk* com *drift* – *trend* estocástico).

- Se, no *random walk* com *drift* $y_t = y_{t-1} + \delta + \varepsilon_t$, tem-se $\delta = 0$, então o modelo é chamado de *trendless random walk*. Mas deve-se notar que mesmo sem haver *trend* na média, existe *trend* na variância !

- No *TSP*, um choque não tem efeito nos futuros valores de y_t. No *DSP*, os efeitos de um choque são permanentes!; e não há um *trend* de longo prazo em torno do qual o sistema pode ser estabilizado.Este tema ressurge quando discute-se "**real business cycle**". Neste modelo, *trends* e ciclos têm uma fonte comum e é a acumulação de choques tecnológicos que produzem um *trend* de longo prazo, bem como produz flutuações de curto prazo em torno deste *trend*.

- Mudanças estruturais (por causa de eventos impactantes, *e.g.* crise 2007) falseiam testes de raízes unitárias. Quebrar as sequências (com *drifts* diferentes) seria uma solução, mas isto pode significar uma interferência na análise. Testes de raízes unitárias têm recebido mais atenção do que estimativas e predições.