Séries Temporais

Neste texto introduzimos e iniciamos o estudo das séries temporais, um tipo particular de sequência de observações de um fenómeno aleatório, com características específicas que importam realçar.

- Uma *série temporal* pode ser encarada como um conjunto de observações registadas de forma sequencial no tempo.
- A dependência da evolução no tempo faz com que muitos dos métodos estatísticos convencionais, que se baseiam em amostras aleatórias, não possam ser utilizados.
- Uma característica intrinseca das séries temporais, é que observações adjacentes sejam dependentes e a natureza dessa dependência será de interesse em qualquer estudo sobre séries temporais.

1. Exemplos de Séries Temporais

Alguns exemplos tradicionais de séries temporais são apresentados seguidamente.

Exemplo 1. Um exemplo clássico de série temporal é o número de mortes acidentais ocorridos por mês, nos EUA, entre 1973 e 1978, que podem ser consultados na Tabela 1[1] ou observados na Figura 1, sob a forma de gráfico.

Exemplo 2. No campo da medicina, temos o exemplo do número de novos casos (diários) de infetados por SARS-COV-2 em Portugal e durante o ano de 2021, observados na forma de gráfico na Figura 2.

[1]Source: Time Series Data Library (Brockwell and Davis (1991))

TABELA 1. Número de mortes acidentais por mês USA (1973-1978)

	1973	1974	1975	1976	1977	1978
Janeiro	9007	7750	8162	7717	7792	7836
Fevereiro	8106	6981	7306	7461	6957	6892
Março	8928	8038	8124	7776	7726	7791
Abril	9137	8422	7870	7925	8106	8129
Maio	10017	8714	9387	8634	8890	9115
Junho	10826	9512	9556	8945	9299	9434
Julho	11317	10120	10093	10078	10625	10484
Agosto	10744	9823	9620	9179	9302	9827
Setembro	9713	8743	8285	8037	8314	9110
Outubro	9938	9129	8433	8488	8850	9070
Novembro	9161	8710	8160	7874	8265	8633
Dezembro	8927	8680	8034	8647	8796	9240

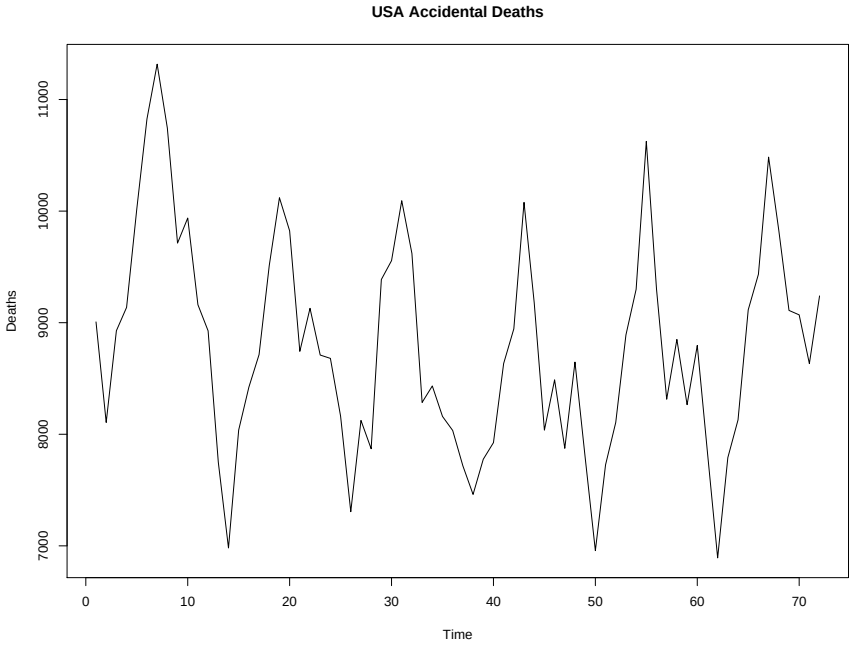

FIGURA 1. Número de mortes acidentais por mês USA (1973-1978)

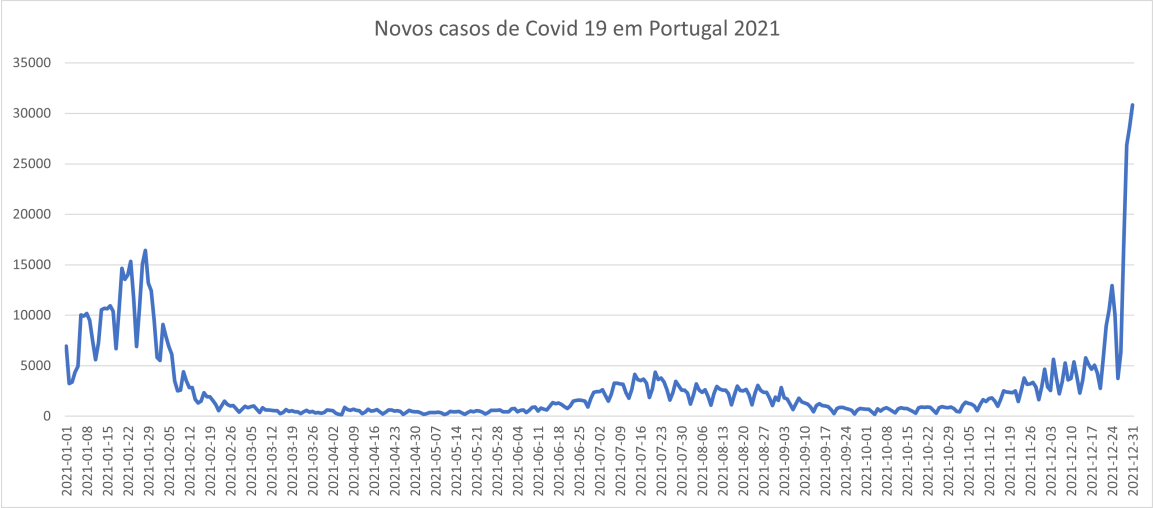

FIGURA 2. Novos casos diários de COVID 19

2. Processos Estocásticos

DEFINIÇÃO 3 (Definição formal). Um **processo estocástico** é uma família de variáveis aleatórias (v.a.), $\mathbb{X} = (X_t)_{t \in T} = \{X_t : t \in T\}$, definidas num espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$.

- Se por exemplo temos, $T = \mathbb{N}$ ou $T = \mathbb{Z}$ temos um processo estocástico a tempo discreto e é nesta classe que encontramos as séries temporais.
- Se por exemplo temos, $T = [0, +\infty[$ ou $T = \mathbb{R}$ temos um processo estocástico a tempo contínuo.

Repare-se que em rigor $X_t \equiv X_t(\omega)$ com $t \in T$ e $\omega \in \Omega$, ou seja, estamos perante uma função de duas variáveis.

- Se fixarmos t então a função $X_t(.) : \Omega \to \mathbb{R}$ é uma variável aleatória;
- Se fixarmos ω então a função $X_.(\omega) : T \to \mathbb{R}$ é uma trajectória (ou realização) do processo estocástico $\mathbb{X}$.

EXEMPLO 4. Exemplos de processos estocásticos são os passeios aleatórios.

- Um **passeio aleatório** pode ser definido por:

$$X_t = X_{t-1} + \varepsilon_t = \sum_{j=1}^{t} \varepsilon_j, \quad X_0 = 0, t = 1, 2, ... \tag{1}$$

onde $\varepsilon_t, t = 1, 2, ...$ são v.a. independentes com distribuição normal de média zero e variância σ^2.

- Um **passeio aleatório com tendência** pode ser definido por:

$$Y_t = c + Y_{t-1} + \varepsilon_t = ct + \sum_{j=1}^{t} \varepsilon_j, \quad X_0 = 0, t = 1, 2, ... \tag{2}$$

onde c é uma constante real e os ε_t definidos como anteriormente.

Na Figura 3 apresentamos duas trajectórias simuladas de passeio aleatório e passeio aleatório com tendência, com $c = 0.1$ e $\varepsilon_t \sim N(0, 1), t = 1, 2, ..., 365$.

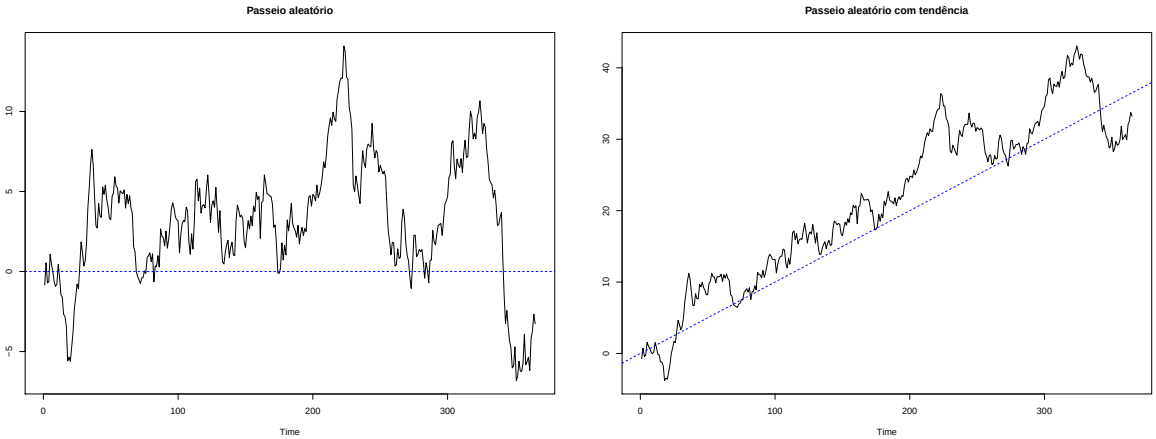

FIGURA 3. Passeios aleatórios

OBSERVAÇÃO 5. Os gráficos apresentados na Figura 3 foram gerados no R, usando:

3. Média, Autocovariância e Autocorrelação

DEFINIÇÃO 6. A função **média** ou **valor médio** de uma série temporal é definida como,

$$\mu_X(t) = \mathbb{E}[X_t], \quad t \in T, \tag{3}$$

desde que o valor médio $\mathbb{E}[X_t]$ exista.

Recorde que o valor médio de uma variável aleatória é uma medida de localização permitindo localizar no espaço a variável aleatória e note também que o valor médio (desde que exista) é um número real e portanto constante para cada variável aleatória X_t, no entanto o resultado pode ser diferente para diferentes valores de t, pelo que de facto μ_t deve ser considerada como uma função de t.

No caso dos passeios aleatórios apresentados no exemplo 4 e usando as propriedades usuais do valor esperado, temos:

- Passeio aleatório:

$$\mu_X(t) = \mathbb{E}[X_t] = \mathbb{E}\left[\sum_{j=1}^{t} \varepsilon_j\right] = \sum_{j=1}^{t} \mathbb{E}\left[\varepsilon_j\right] = 0 \tag{4}$$

- Passeio aleatório com tendência:

$$\mu_Y(t) = \mathbb{E}[Y_t] = \mathbb{E}\left[ct + \sum_{j=1}^{t} \varepsilon_j\right] = ct + \mathbb{E}\left[\sum_{j=1}^{t} \varepsilon_j\right] = ct \tag{5}$$

Portanto, o passeio aleatório e o passeio aleatório com tendência têm funções valor médio distintas.

Definição 7. A função **variância** de uma série temporal é definida de forma análoga ao valor médio, como sendo (para cada t) a variância da variável aleatória X_t,

$$var_X(t) = Var(X_t) = \mathbb{E}[(X_t - \mu_X(t))^2], \quad t \in T, \tag{6}$$

desde que o valor médio do lado direito da igualdade exista.

Recorde que a variância de uma variável aleatória é uma medida de dispersão dando-nos informação sobre a variabilidade da variável aleatória e a função variância, neste contexto, será também uma função (não negativa) de t.

Novamente, no caso do exemplo 4 e usando agora as propriedades da variância conjuntamente com o facto de as variáveis aleatórias $(\varepsilon_j)_{j\geq 0}$ serem independentes, teremos que:

- Passeio aleatório:

$$var_X(t) = Var(X_t) = Var\left(\sum_{j=1}^{t} \varepsilon_j\right) = \sum_{j=1}^{t} Var\left(\varepsilon_j\right) = \sigma^2 t \tag{7}$$

- Passeio aleatório com tendência:

$$var_Y(t) = Var(Y_t) = Var\left(ct + \sum_{j=1}^{t} \varepsilon_j\right) = \sum_{j=1}^{t} Var\left(\varepsilon_j\right) = \sigma^2 t \tag{8}$$

Ou seja, o passeio aleatório e o passeio aleatório com tendência têm a mesma função de variância.

Definição 8. A função **autocovariância** de uma série temporal $(X_t)_{t\in T}$, é a função que traduz a covariância entre todos os pares de variáveis aleatórias que formam a série temporal, sendo por isso definida como,

$$\gamma_X(s,t) = Cov(X_s, X_t) = \mathbb{E}[(X_s - \mu_X(s))(X_t - \mu_X(t))], \quad s,t \in T. \tag{9}$$

Recorde que a covariância entre duas variáveis aleatórias é uma medida da relação linear entre as variáveis e repare que:

- Pela própria definição da função de autocovariância,

$$\gamma_X(s,t) = \gamma_X(t,s), \quad \forall s,t \in T. \tag{10}$$

- Também teremos,

$$\gamma_X(t,t) = Cov(X_t, X_t) = var_X(t), \quad \forall t \in T. \tag{11}$$

- Pela definição de covariância,

$$\gamma_X(s,t) = Cov(X_s, X_t) = \mathbb{E}[X_s X_t] - \mu_X(s)\mu_X(t), \quad \forall s,t \in T. \tag{12}$$

No caso do passeio aleatório ou do passeio aleatório com tendência (exemplo 4), temos que para $s,t \in T$:

(13)

$$\gamma_Y(s,t) = \mathbb{E}[(Y_s - \mu_Y(s))(Y_t - \mu_Y(t))] = \mathbb{E}[(Y_s - cs)(Y_t - ct)] = \mathbb{E}[X_s X_t] = \gamma_X(s,t).$$

Como

$$\gamma_X(s,t) = Cov(X_s, X_t) = Cov\left(\sum_{i=1}^{s} \varepsilon_i, \sum_{j=1}^{t} \varepsilon_j\right) \tag{14}$$

como os $\varepsilon_i, i = 1, 2, ...$ são independentes ($Cov(\varepsilon_i, \varepsilon_j) = 0, i \neq j$) e supondo que $s < t$, teremos pelas propriedades da covariância que:

$$\gamma_X(s,t) = \sum_{i=1}^{s}\sum_{j=1}^{t} Cov\left(\varepsilon_i, \varepsilon_j\right) = \sum_{i=1}^{s} Cov\left(\varepsilon_i, \varepsilon_i\right) = s\sigma^2 \tag{15}$$

se $t < s$, o resultado viria $t\sigma^2$, pelo que:

$$\gamma_X(s,t) = \gamma_Y(s,t) = \min(s,t)\sigma^2. \tag{16}$$

Definição 9. A função de **autocorrelação** (ACF) de uma série temporal $(X_t)_{t\in T}$, é definida como,

$$\rho_X(s,t) = \frac{\gamma_X(s,t)}{\sqrt{\gamma_X(s,s)\gamma_X(t,t)}} = \frac{Cov(X_s, X_t)}{\sqrt{Var(X_s)Var(X_t)}}, \quad s,t \in T. \tag{17}$$

A autocorrelação tem uma interpretação idêntica à de autocovariância com uma diferença de escala, temos sempre que, $|\rho_X(s,t)| \leq 1$.

Novamente, no exemplo dos passeios aleatórios (exemplo 4), temos que para $s, t \in T$:

$$\rho_X(s,t) = \frac{\gamma_X(s,t)}{\sqrt{\gamma_X(s,s)\gamma_X(t,t)}} = \frac{\min(s,t)\sigma^2}{\sqrt{\min(s,s)\sigma^2 \min(t,t)\sigma^2}}$$

$$= \frac{\min(s,t)\sigma^2}{\sqrt{st}\sigma^2} = \begin{cases} \sqrt{\frac{s}{t}}, & s \leq t \\ \\ \sqrt{\frac{t}{s}}, & s > t \end{cases} \tag{18}$$

Exercício 1. Considere o processo (de média móvel[2]),

$$M_t = \frac{\varepsilon_t + \varepsilon_{t-1}}{2}, t = 1, 2, ...$$

onde $\varepsilon_t, t = 0, 1, ...$ são v.a. independentes com média zero e variância σ^2.

(1) Calcule o valor médio de M_t, $\mu_M(t)$.

(2) Calcule a variância de M_t, $var_M(t)$.

[2]Os processos de média móvel serão introduzidos e estudados em maior detalhe numa outra e-atividade

(3) Mostre que a função de autocovariância de M_t, $\gamma_M(s,t)$, é dada por:

$$\gamma_M(s,t) = \begin{cases} 0.5\sigma^2, & s = t \\ 0.25\sigma^2, & |t-s| = 1 \\ 0, & |t-s| > 1 \end{cases}$$

(4) Conclua que:

$$\rho_M(s,t) = \begin{cases} 1, & s = t \\ 0.5, & |t-s| = 1 \\ 0, & |t-s| > 1 \end{cases}$$