# MÉTODOS ESTATÍSTICOS DE PREVISÃO

Profa. Glaura Franco

# SÉRIE TEMPORAL

**Definição:** Uma Série Temporal é um conjunto de observações geradas sequencialmente no tempo.

- Uma série temporal pode ser observada como uma realização parcial de um processo estocástico.

**Característica principal:** as variáveis são dependentes.

- Denotaremos a série temporal por $Y_1, Y_2, \ldots, Y_T$ onde $T$ é o tamanho da série.

- Trabalharemos com séries temporais a tempo discreto, onde os dados são coletados diariamente, semanalmente, mensalmente ou anualmente.

# ANÁLISE NO DOMÍNIO DO TEMPO

- Mede a magnitude do evento que ocorre em determinado instante de tempo (ano, trimestre, mês, semana, dia, etc.)

- Utiliza as funções de autocovariância e autocorrelação.

- **Objetivos:**

  - Modelagem;

  - Previsão e controle de qualidade;

  - Periodicidade.

# SÉRIES REAIS

- Neste curso serão utilizadas 4 séries reais para ilustrar os métodos apresentados:

1. **IPCA:** Série do Índice de Preços ao Consumidor Amplo de Belo Horizonte (jan/97 a out/05).
2. **TEMPERATURA:** Série de temperatura global Media, de 1900 a 1997. Os dados foram calculados como um desvio da temperatura global média anual do período 1961-1990.
3. **FORTAL:** Série de precipitação pluviométrica mensal na cidade de Fortaleza (jan/24 a dez/99).
4. **CEP:** Série do consumo de energia elétrica das Centrais Elétricas do Paraná (jan/80 a dez/94).

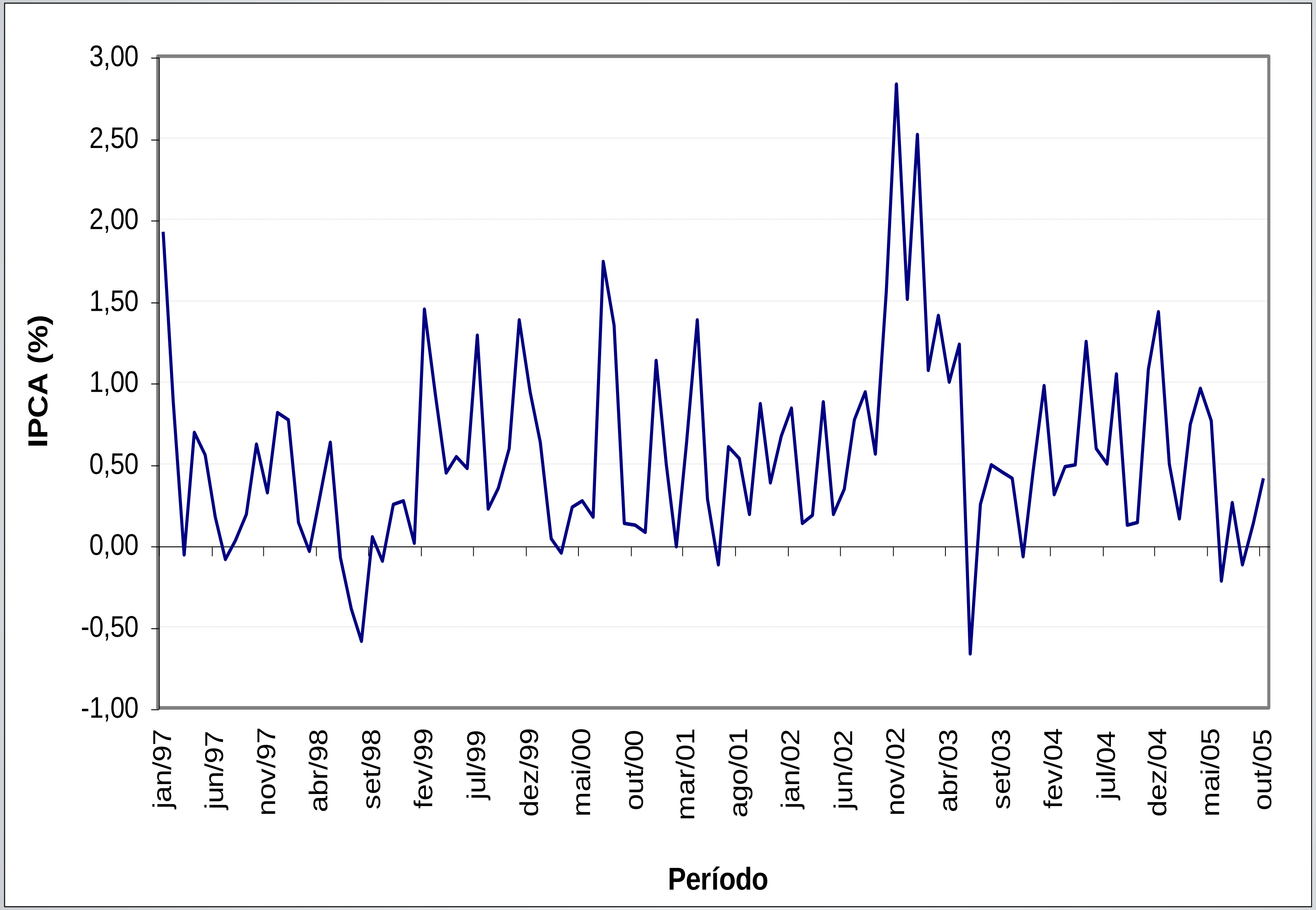



**FIGURA 1 :** SÉRIE IPCA

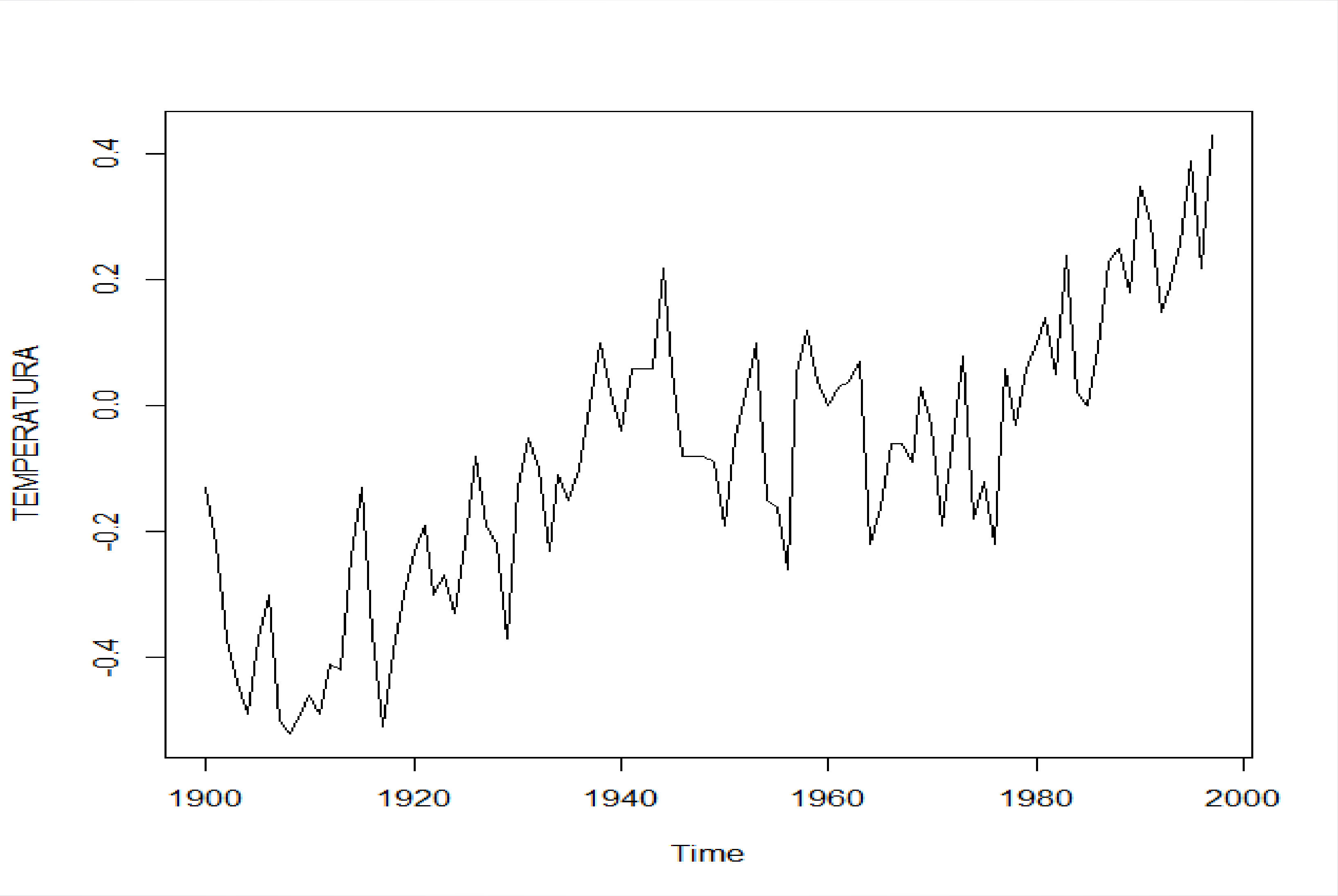



**FIGURA 2 :** SÉRIE TEMPERATURA GLOBAL

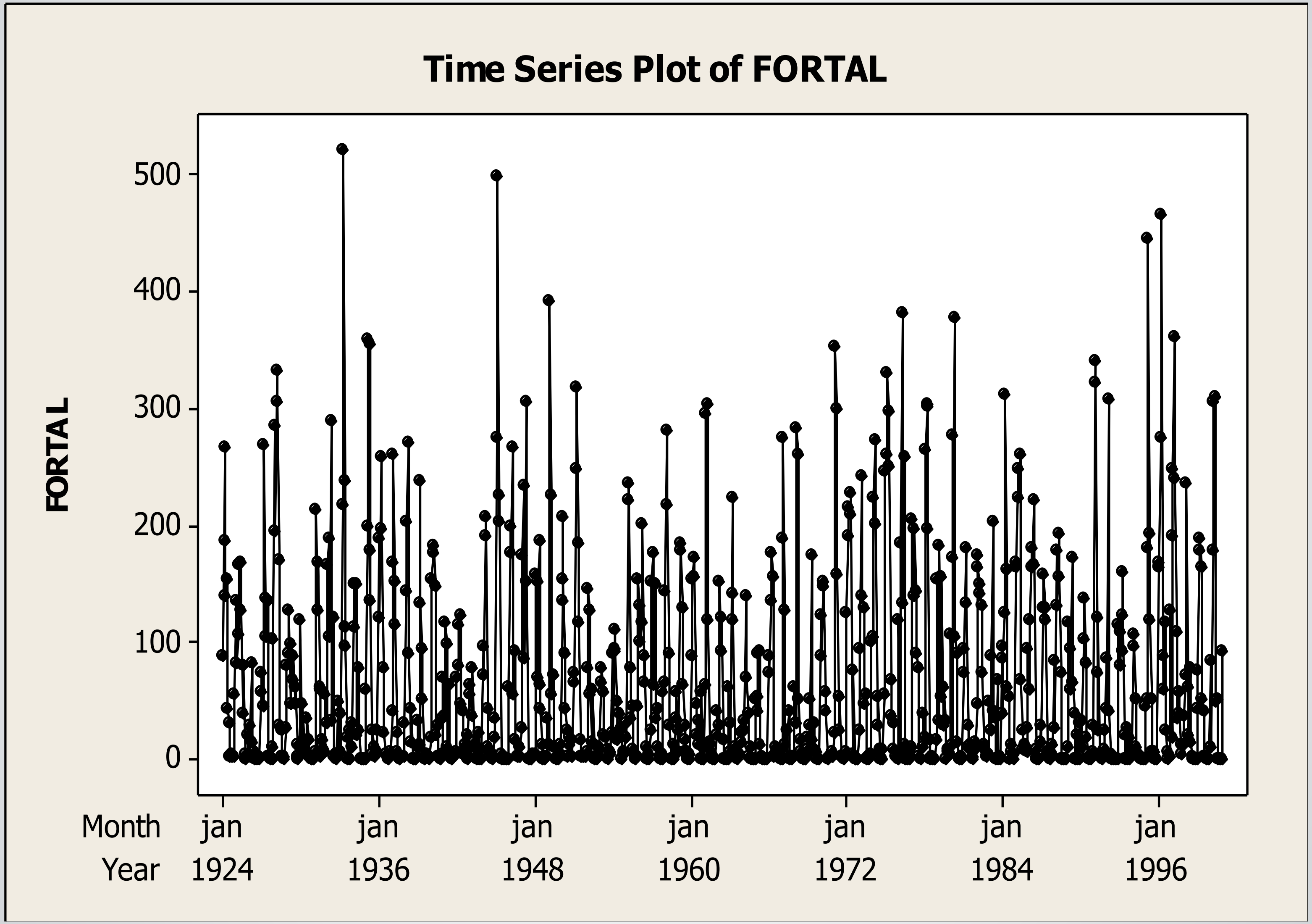



**FIGURA 3 :** SÉRIE FORTAL

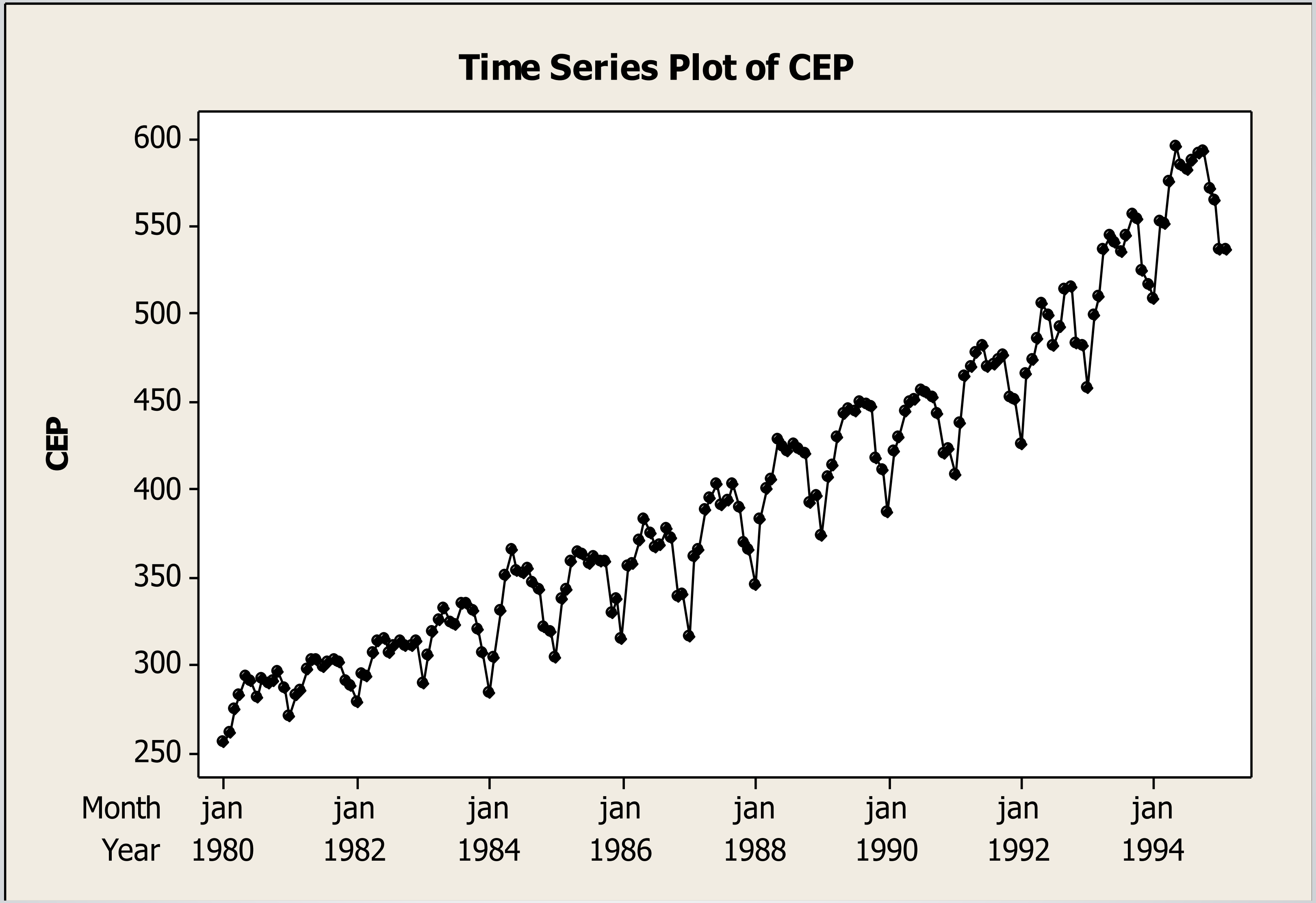



**FIGURA 4:** SÉRIE CEP

# OPERADORES

- **Operador de retardo** ou de translação para o passado: representa uma defasagem de *k* períodos de tempo para trás. É denotado por *B* e definido por:

$$B^k Y_t = Y_{t-k}$$

- **Operador de avanço** ou de translação para o futuro: representa uma defasagem de *k* períodos de tempo para frente. É denotado por *F* e definido por:

$$F^k Y_t = Y_{t+k}$$

- **Operador diferença:** é uma transformação nos dados que consiste em tomar diferenças sucessivas da série original. É denotado por $\nabla$ e definido por:

$$\nabla^d Y_t = (1 - B)^d Y_t$$

# ESTACIONARIEDADE

Se o processo estocástico que gerou a série de observações é invariante com respeito ao tempo, diz-se que o mesmo é estacionário. Um processo estacionário pode ser classificado em:

- **Estritamente estacionário:** Quando a distribuição de $Y_t$ é a mesma de $Y_{t+k}$.

- **Fracamente estacionário** (ou estacionário de segunda ordem): quando os dois primeiros momentos da distribuição de $Y_t$ forem constantes ao longo do tempo, o que implica em:

$$E(Y_t) = \mu;$$

$$Var(Y_t) = \sigma^2.$$

Em termos práticos, uma série é fracamente estacionária quando ela se desenvolve no tempo ao redor de uma média e variância constantes, indicando um padrão de equilíbrio.

# PROCESSO NÃO ESTACIONÁRIO HOMOGÊNEO

Um processo é não estacionário homogêneo se, ao tomarmos suas diferenças sucessivas, encontramos um processo estacionário. Por exemplo, seja a série temporal não estacionária abaixo:

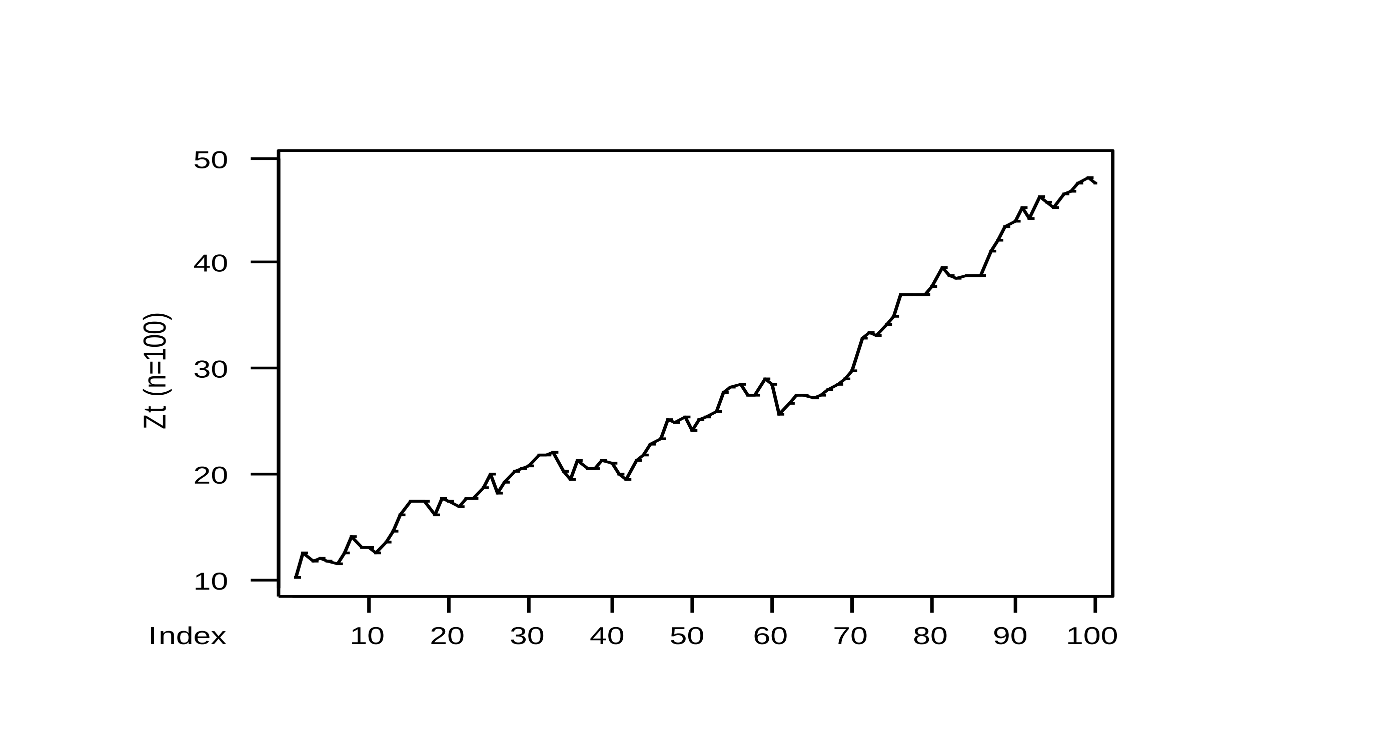



**Figura 5:** Série não estacionária

Se calcularmos a primeira diferença desta série, teremos:

$$W_t = \nabla Y_t = Y_t - Y_{t-1}$$

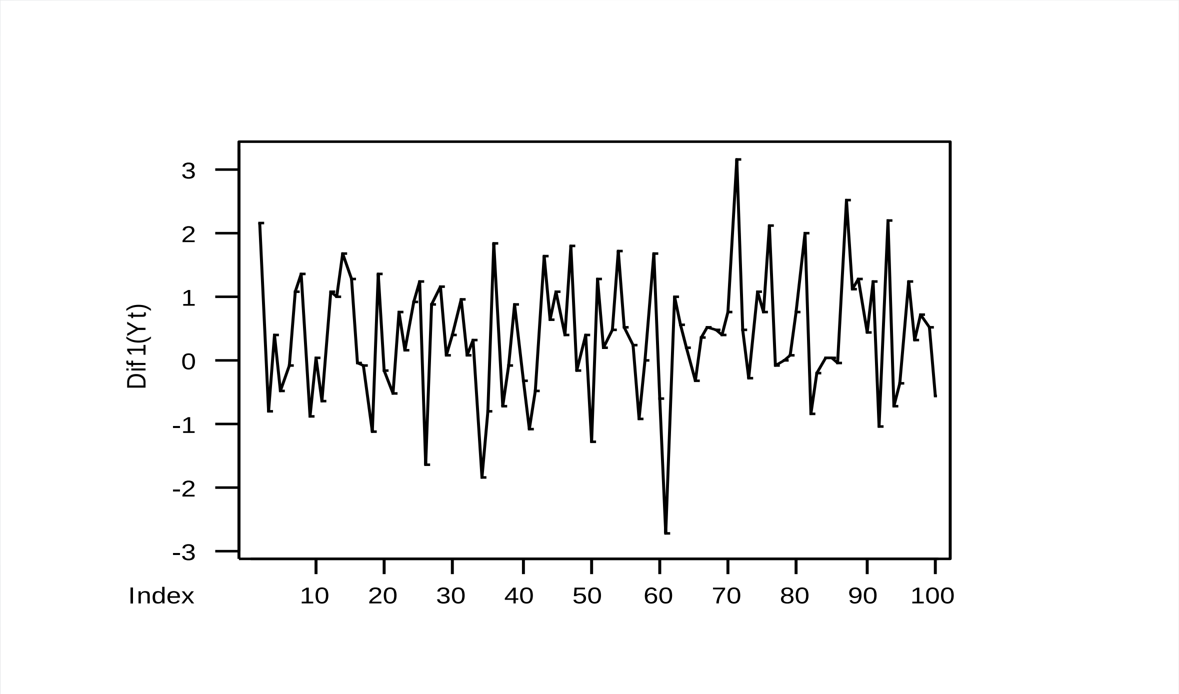



**Figura 6**: Série da primeira diferença

Se a série diferenciada é estacionária, dizemos que a série original é integrada de ordem 1, ou I(1).

# AUTOCOVARIÂNCIA

É a covariância entre $Y_t$ e $Y_{t-k}$ separados por $k$ intervalos de tempo:

$$\gamma_k = Cov[Y_t, Y_{t-k}] = E[(Y_t - \mu)(Y_{t-k} - \mu)], \qquad k = 0, \pm 1, \pm 2, \dots$$

Se temos uma série real, o estimador amostral aproximadamente não-tendencioso (para grandes amostras) da autocovariância é dado por:

$$\hat{\gamma}_k = \frac{1}{T} \sum_{t=k+1}^{T} (Y_t - \bar{Y})(Y_{t-k} - \bar{Y}).$$

Como a autocovariância é uma função par, temos que para todo inteiro $k$, $\gamma_k = \gamma_{-k}$. Portanto, é necessário determinar $\gamma_k$ apenas para $k \geq 0$.

# AUTOCORRELAÇÃO

- A autocorrelação é a autocovariância padronizada. Serve para medirmos o comprimento e a memória de um processo, ou seja, a extensão para a qual o valor tomado no tempo $t$ depende daquele tomado no tempo $t-k$,

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{Cov[Y_t, Y_{t-k}]}{\sqrt{Var(Y_t)Var(Y_{t-k})}}$$

- Claramente $\rho_0 = 1$ e $\rho_k = \rho_{-k}$. Um estimador amostral da autocorrelação de defasagem $k$ é dado por:

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0}, \qquad k = 0,1,2,\ldots$$

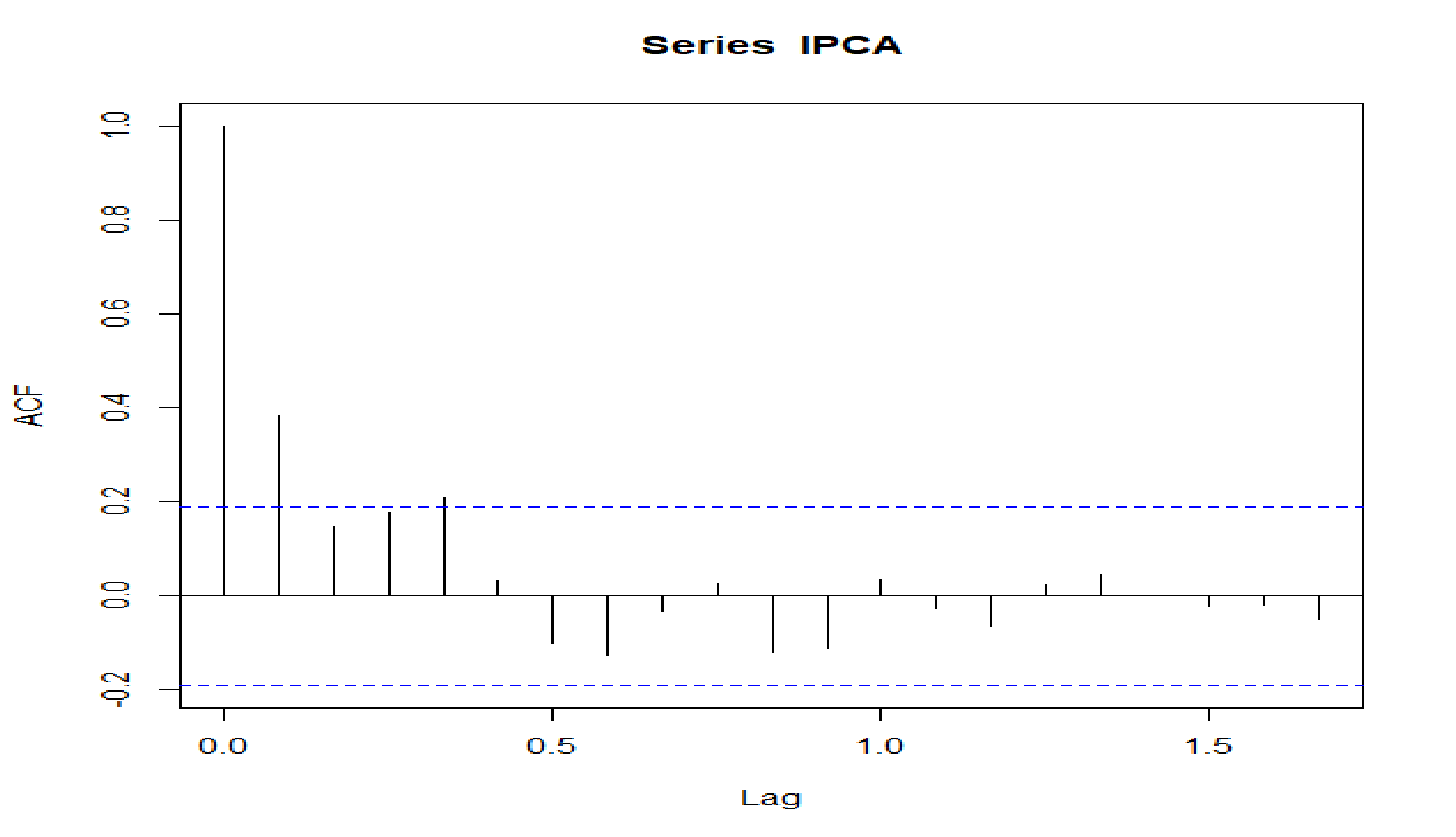
# FUNÇÃO DE AUTOCORRELAÇÃO AMOSTRAL (ACF)

- A ACF pode ser construída no R usando o comando:

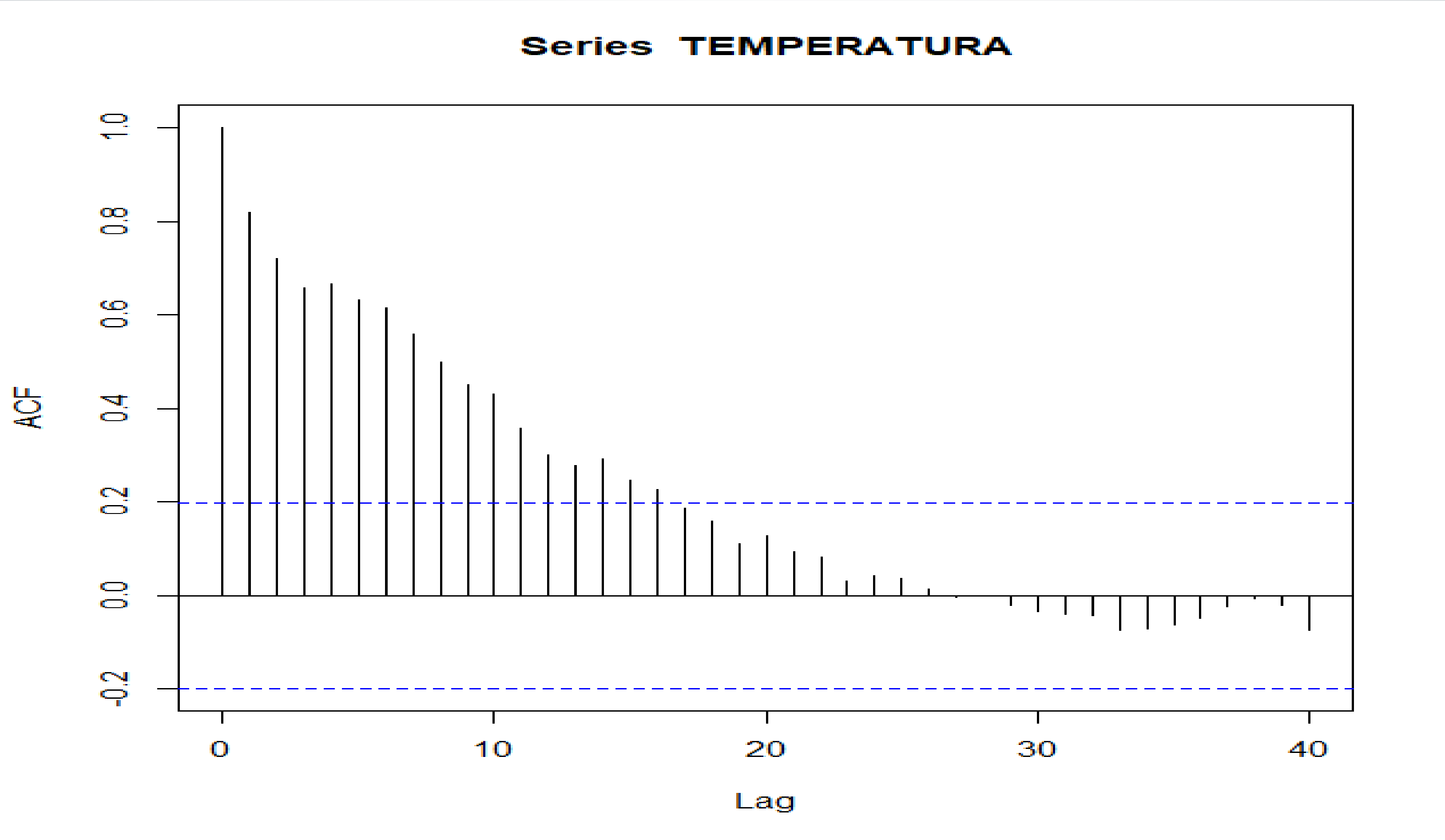
**acf(x, lag.max = NULL)**

onde x é a série e lag.max é o número de *lags* que se quer utilizar no cálculo da ACF. Se não for especificado um número, como no caso acima, o R usa o *default* de $10 \times log_{10}(T/m)$ onde $T$ é o número de observações e $m$ é o número de séries

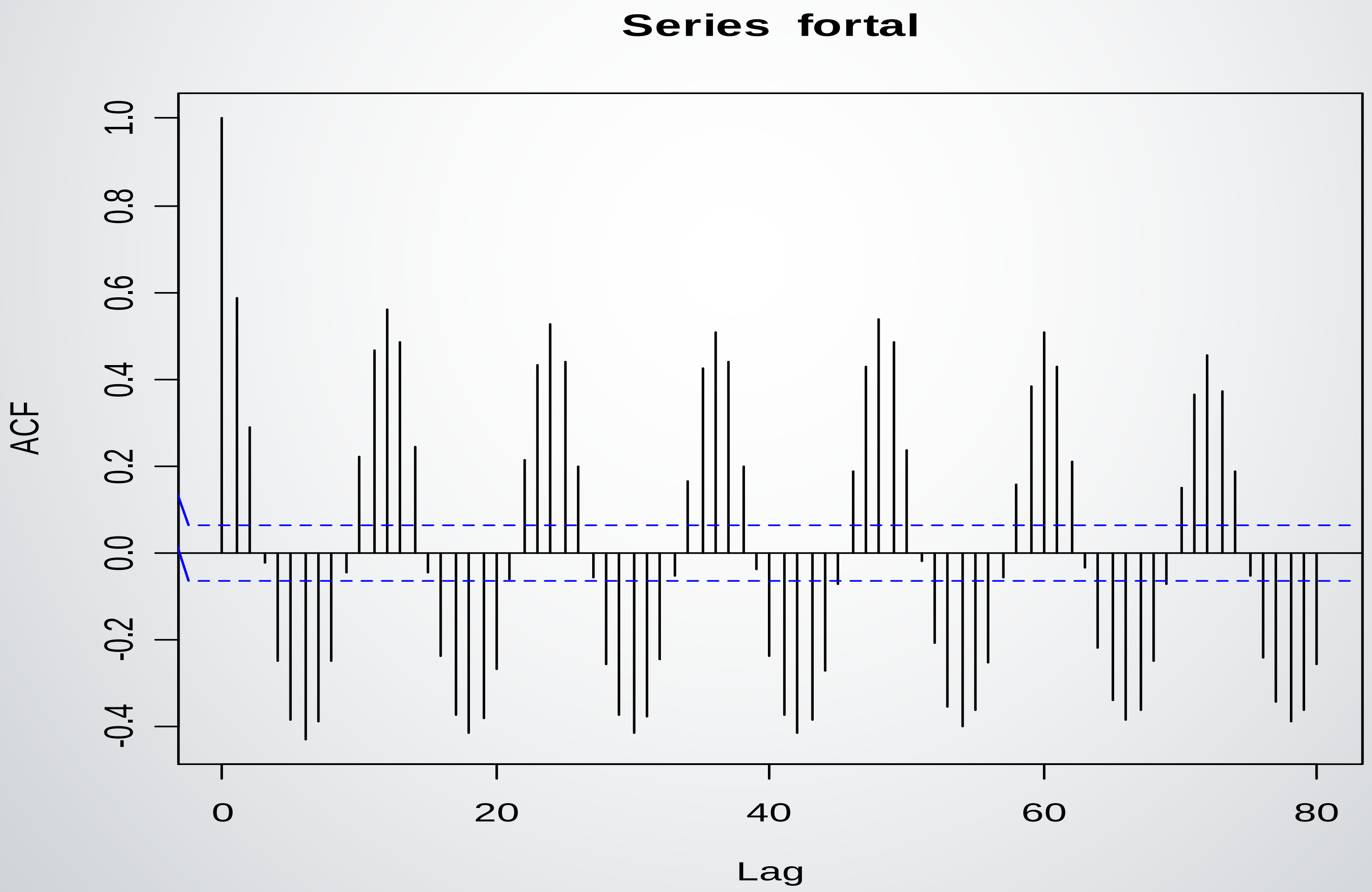
# ACF da série IPCA



# ACF da série TEMPERATURA GLOBAL



# ACF da série FORTAL



# RUÍDO BRANCO

- Seja $u_t$ um processo estocástico.

$u_t$ é um ruído branco se:

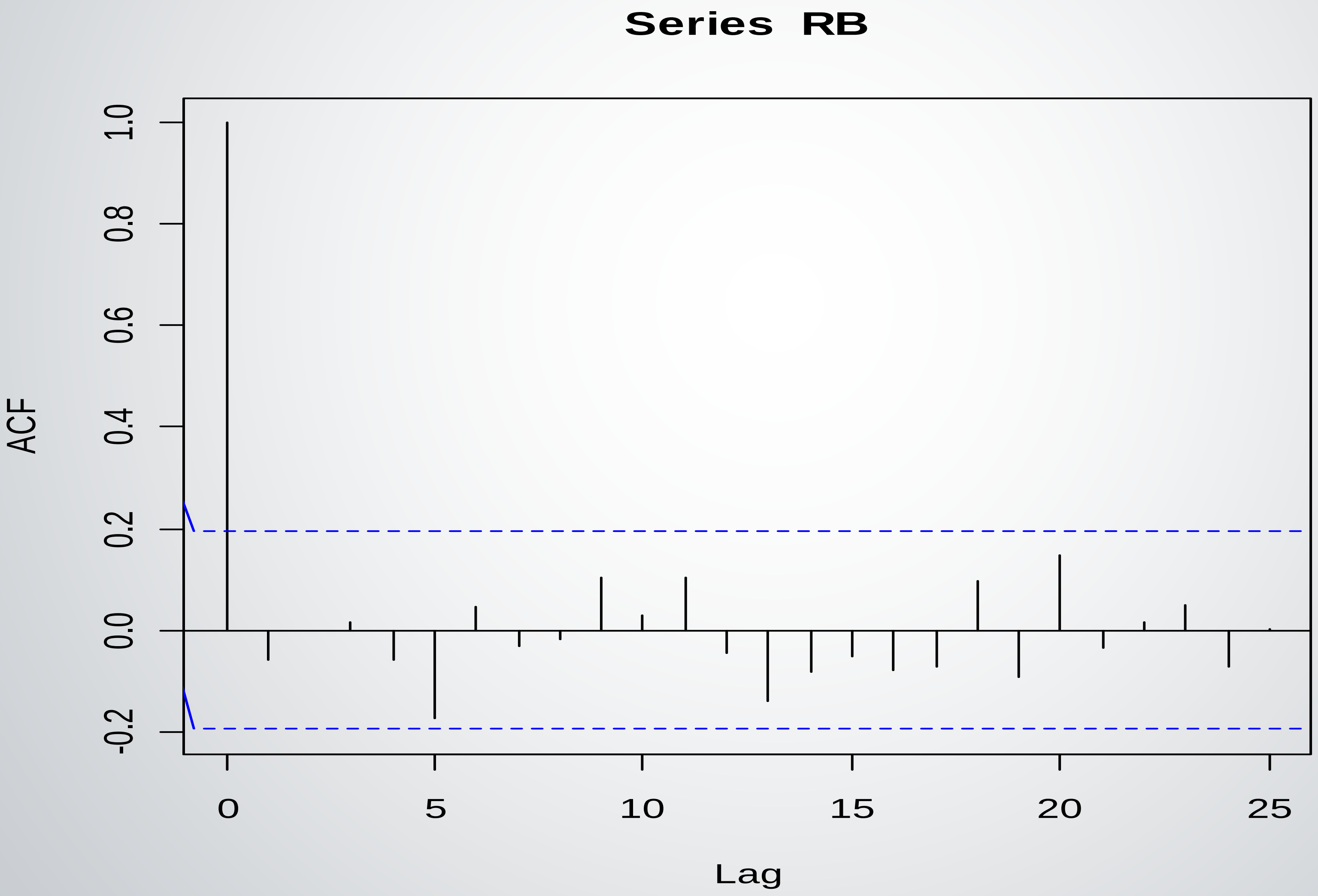
*i.* $E[u_t] = 0,$

*ii.* $E[u_t^2] = \sigma^2 < \infty$ e

*iii.* $E[u_t\ ,\ u_{t+k}] = 0 \qquad k = 0, \pm 1, \pm 2, \ldots$

- Além disto, $\rho_k = 1$ para $k = 0,$ e $\rho_k = 0$ para $k = \pm 1, \pm 2, \ldots$

# ACF de um ruído branco



# TRANSFORMAÇÃO BOX - COX

- Para séries temporais que não apresentam variância constante, é necessário transformar a série original para estabilizá-la. Através da transformação Box-Cox, pode-se obter uma série mais simétrica e com variância constante, ou seja, próxima da distribuição normal, quando valores apropriados de $\lambda$ na equação abaixo são encontrados,

$$Y_t^{(\lambda)} = \begin{cases} \dfrac{Y_t^{\lambda} - c}{\lambda}, & se \ \ \lambda \neq 0 \\ logY_t, & se \ \ \lambda = 0 \end{cases}$$

- **Obs.:**
  - $\lambda$ e c são parâmetros a serem estimados.
  - $\lambda$ é obtido a partir de uma varredura sobre um intervalo de valores possíveis, no qual procura-se identificar aquele que apresente a menor soma dos quadrados dos resíduos.