# 4
# Critérios para Avaliação dos Cenários

É desejável que um modelo de geração de séries sintéticas preserve as principais características da série histórica. Isto quer dizer que a utilidade de um modelo pode ser verificada pela sua habilidade de reproduzir distribuições de probabilidade de variáveis aleatórias relevantes ao processo.

Dessa maneira, com o intuito de verificar se os modelos de memória longa utilizados são capazes de reproduzir as propriedades estatísticas da série histórica, neste capítulo serão introduzidos os testes estatísticos baseados nos testes empregados por OLIVEIRA (2010).

## 4.1
## Teste de Médias

O teste de média é, sem dúvida, o teste mais importante a ser verificado. Este tem como objetivo comparar a média dos cenários com a média do histórico e foi realizado de duas maneiras.

Primeiramente, denomina-se teste de média global a comparação da média dos cenários com a média do histórico, isto é, nesta análise, verifica-se se a média dos cenários sintéticos é estatísticamente igual à média do histórico.

O segundo teste de médias a ser realizado diz respeito à comparação das médias período a período. Este tem como objetivo comparar as médias mensais dos cenários com as médias mensais do histórico, desta forma, compara-se os janeiros gerados com os meses de janeiro do histórico, e assim sucessivamente para todos os meses.

Suponha, segundo CASELLA & BERGER (2001), duas populações independentes e normal, sendo o primeiro com média $\mu_1$ e variância $\sigma_1^2$ e a segunda com média e variância igual a $\mu_2$ e $\sigma_2^2$, respectivamente. Inferências são realizadas em duas amostras aleatórias de tamanhos $n_1$ e $n_2$. Assim sendo, $X_{11}, X_{12}, \ldots, X_{1n_1}$ e $X_{21}, X_{22}, \ldots, X_{2n_2}$ são amostras aleatórias de tamanho $n_1$ e $n_2$

oriundas da população 1 e população 2, respectivamente.

Levando em consideração as suposições anteriores, as hipóteses nulas e alternativas são dadas por:

$$\begin{aligned} H_0 &: \mu_1 = \mu_2 \\ H_1 &: \mu_1 \neq \mu_2 \end{aligned} \tag{4.1}$$

Neste trabalho, a análise é apresentada através dos *p-valores* dos testes que devem estar acima do nível de significância adotado, para que a hipótese nula não seja rejeitada.

## 4.2 Teste de Variâncias

O teste para verificar a igualdade de variâncias é o teste de Levene. O motivo de tal escolha se dá pelo fato de que este teste é robusto à ausência de normalidade nos dados.

Segundo OLIVEIRA (2010), suponha que sejam tomadas $k \geq 2$ amostras aleatórias independentes entre si. A amostra $i$ representa uma coleção de $n_i$ variáveis aleatórias *i.i.d.*, com distribuição $G_i$, com média $\mu_i$ e variância $\sigma_i^2$ para $G_i$, $\mu_i$ e $\sigma_i^2$ desconhecidos.

As hipóteses nulas e alternativas são respectivamente:

$$\begin{aligned} H_0 &: \sigma_1^2 = \ldots = \sigma_k^2 \\ H_1 &: \sigma_i^2 \neq \sigma_j^2 \end{aligned} \tag{4.2}$$

para $i = 1,\ldots,k$ e $j = 1,\ldots,k$ e para alguma $i \neq j$

Como os resultados dos testes são apresentados através dos *p-valores*, estes devem ser comparados com o nível de significância adotado. Caso o *p-valor* seja maior do que o nível de significância, a hipótese nula não pode ser rejeitada, caso contrário a hipótese nula é rejeitada.

De maneira similar ao teste de média, o teste de igualdade de variâncias também foi realizado em dois momentos. No primeiro testou-se a igualdade de variâncias entre os cenários gerados e a série histórica. Em um segundo momento, testou-se a igualdade de variâncias entre os períodos dos cenários com os respectivos períodos do histórico.

## 4.3
## Análise de Aderência

Nesta seção, são apresentados os testes estatísticos não paramétricos utilizados neste trabalho. Esses testes correspondem a uma classe de testes de hipótese que têm como objetivo verificar a forma de uma distribuição de probabilidade. Em outras palavras, deseja-se verificar se os dados referentes a uma distribuição de probabilidade se adequam à curva de um modelo distributivo hipotético.

Os testes realizados foram os de Kolmogorov-Smirnov e Qui-Quadrado. O teste de Kolmogorov-Smirnov tem por finalidade determinar se duas distribuições de probabilidade são iguais ou se uma distribuição de probabilidade é igual a uma distribuição em hipótese. Neste caso, o objetivo é verificar se as distribuições de probabilidade dos cenários gerados são iguais à distribuição do histórico.

De forma intuitiva, este teste baseia-se na diferença entre as distribuições de probabilidades, empíricas e teóricas. Mais detalhes sobre esse teste podem ser encontrados em CONOVER (1971).

Adotando $F_1(x)$ e $F_2(x)$ como as distribuições de probabilidade das amostras 1 e 2, as hipóteses nulas e alternativas são definidas como:

$$\begin{aligned} H_o &: F_1(x) = F_2(x) \\ H_1 &: F_1(x) \neq F_2(x) \end{aligned} \tag{4.3}$$

Como nos casos anteriores, análise é realizada através dos *p-valores* e, desta forma, valores acima do nível de significância indicam que a hipótese nula não pode ser rejeitada.

Este teste também foi realizado em duas etapas. A primeira etapa possui caráter global e averigua se os cenários gerados seguem a mesma distribuição da série histórica. A segunda etapa verifica se os períodos dos cenários gerados possuem a mesma distribuição de probabilidade do histórico disponível.

O teste Qui-Quadrado é realizado nesta dissertação para avaliar a eficiência do ajuste da distribuição, isto é, o quanto a frequência observada está próxima da frequência esperada. Mais detalhes sobre este teste podem ser encontrados em SHESKIN (2003).

A hipótese nula e alternativa são, respectivamente:

$$\begin{aligned} &H_0 : \text{não existe diferença entre as frequências observadas e esperadas} \\ &H_1 : \text{existe diferença entre as frequências observadas e esperadas} \end{aligned} \tag{4.4}$$

O teste Qui-Quadrado foi empregado na variável aleatória *comprimento de sequência*, que será definida no próximo subitem, no qual apenas duas classes foram definidas.

## 4.4
## Análise de Sequência

Para a análise de sequência, novas varíaveis aleatórias foram criadas a partir do histórico e dos cenários sintéticos, com o objetivo de aferir a capacidade dos modelos em reproduzir as frequências observadas no histórico.

As variáveis aleatórias que serão introduzidas nesta subseção estão relacionadas com a representação de períodos críticos, como as secas registradas no histórico. Dessa forma, é utilizado então o conceito de sequência negativa. Uma sequência negativa pode ser definida por um período de tempo em que as ENAs estão continuamente abaixo de valores preestabelecidos, sucedidas por valores acima destes limites. Neste trabalho, os limites predeterminados foram as médias mensais.

O conceito de sequência negativa pode ser melhor entendido observando-se a figura 4.1 abaixo. Nesta, a linha contínua representa uma série qualquer de ENA e a linha pontilhada representa um limite predeterminado. Como pode ser observado, os intervalos $(t_1 - t_2)$ e $(t_3 - t_4)$ representam sequências negativas, pois as mesmas estão continuamente abaixo dos limites.

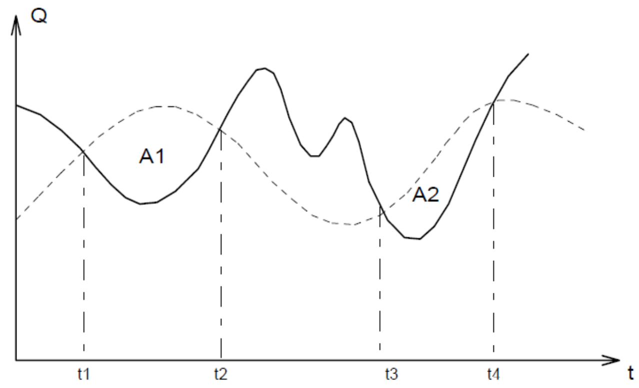


**Figura 4.1 – Representação de Sequência Negativa**
**Fonte:** Adaptado de PENNA (2009)

A partir de cada sequência negativa, três variáveis podem ser criadas: *comprimento, soma e intensidade de sequência*. Essas variáveis são definidas da seguinte forma:

**Tabela 4.1 – Variáveis para Análise de Sequência Negativa**

| Variável Aleatória | Descrição | Cálculo |
|---|---|---|
| Comprimento de Sequência | Corresponde ao comprimento de intervalos $(t_2 - t_1)$ e $(t_3 - t_4)$ | $C = (t_2 - t_1)$ |
| Soma de Sequência | Corresponde à área abaixo do limite durante a sequência. Na figura anterior, equivalem às áreas A1 e A2 | $S = \sum_{i=t1}^{t2} (Z_i - \mu_i)$ |
| Intensidade de Sequência | Corresponde ao valor médio abaixo do limite, isto é, à soma de sequência dividida pelo respectivo comprimento de sequência | $I = \frac{S}{C} = \frac{\sum_{i=t_1}^{t_2} (Z_i - \mu_i)}{(t_2 - t_1)}$ |

**Fonte:** Oliveira (2010)

Realizando os cálculos para cada sequência negativa encontrada, tanto para os cenários gerados, quanto para o histórico disponível, é possível obter assim amostras de cada variável. Assim sendo, tanto para o histórico quanto para os cenários sintéticos foram obtidas três variáveis aleatórias.

Isto posto, em posse de duas amostras de cada variável é possível testar hipóteses se as amostras são provenientes da mesma distribuição por meio dos testes estatísticos de aderência expostos no subitem 4.3. A variável *comprimento de sequência* é avaliada pelo teste Qui-Quadrado, enquanto que as variáveis *soma* e *intensidade de sequência* são avaliadas pelo teste de Kolmogorov-Sminorv.

PENNA (2009) salienta que os índices mais relevantes para o planejamento estão, de modo geral, associados a valores extremos das distribuições, como, por exemplo, o período crítico que corresponde à pior situação hidrológica em todo o histórico.

Desse modo, também foram utilizados os índices do tipo "máximo" sobre as três variáveis estabelecidas anteriormente.

Como a série histórica possui apenas um valor, faz mais sentido falar em "tipicidade" do valor histórico em relação à distribuição dos valores gerados, ao invés de se falar em aderência de distribuições. Isto é, deseja-se saber a probabilidade de a amostra histórica ser sorteada, uma vez que o modelo de geração adotado seja verdadeiro.

Em termos univariados, o desempenho do modelo pode ser medido pela proporção de índices gerados maiores ou menores do que o índice histórico. Caso esta proporção seja muito pequena, isso é um indício de que a observação histórica é atípica para o modelo considerado, PENNA (2009).

Nesta análise usamos as seguintes variáveis: *máximo comprimento de sequência*, *máxima soma de sequência e máxima intensidade de sequência*.