



CAPÍTULO 7



TESTES DE HIPÓTESES





CAPÍTULO 7 TESTES DE HIPÓTESES

Além dos métodos de estimação de parâmetros e de construção de intervalos de confiança, os *testes de hipóteses* são procedimentos usuais da inferência estatística, úteis na tomada de decisões que concernem à forma, ou ao valor de um certo parâmetro, de uma distribuição de probabilidades, da qual se conhece apenas uma amostra de observações. Tais testes envolvem a formulação de uma hipótese, na forma de uma declaração conjectural sobre o comportamento probabilístico da população. Essa hipótese pode se materializar, por exemplo, em uma premissa, formulada *a priori*, a respeito de um certo parâmetro populacional de uma variável aleatória. Não rejeitar ou rejeitar uma tal hipótese irá depender do confronto entre a conjectura e a realidade física, essa concretizada pelas observações que compõem a amostra. A *rejeição* da hipótese implica na necessidade de eventual revisão da conjectura inicial, em decorrência de seu desacordo com a realidade imposta pelos dados amostrais. Por outro lado, a *não rejeição* da hipótese significa que, com base nos dados amostrais, não há elementos suficientes para descartar a plausibilidade da premissa inicial sobre o comportamento da variável aleatória; observe que ‘não rejeitar’ não significa ‘aceitar’ a hipótese.

Por tratar-se de uma inferência a respeito de uma variável aleatória, a decisão de não rejeitar (ou de rejeitar) uma hipótese, é tomada com base em uma certa probabilidade ou *nível de significância* α . Pode-se, por exemplo, não rejeitar a hipótese de que houve um decréscimo *significativo* da vazão média dos últimos trinta anos, em uma certa seção fluvial. Contrariamente, a eventual variação da vazão média do período, pode ser uma mera decorrência das flutuações amostrais, sem conseqüências para a média populacional em questão; nesse caso, a variação é dita *não significativa*. A especificação prévia de um nível de significância α , cumpre o papel de remover o grau de subjetividade associado à tomada de decisão intrínseca a um teste de hipótese. De fato, para um mesmo nível de significância, dois analistas diferentes, ao realizarem o teste de uma certa hipótese, sob condições idênticas, tomariam uma única e igual decisão. O nível de significância α de um teste de hipótese é complementar à probabilidade $(1 - \alpha)$ com que um certo intervalo de confiança $[I, S]$ contém o valor populacional de um parâmetro θ . De fato, o intervalo $[I, S]$ estabelece os limites de variação da chamada *estatística de teste*, dentro dos quais a hipótese sobre θ não pode ser rejeitada. Contrariamente, se os valores da estatística de teste localizarem-se fora dos limites impostos por $[I, S]$, a hipótese sobre θ deve ser rejeitada, a um nível de significância α . Portanto, segundo essa interpretação, a construção de intervalos de confiança representa a operação inversa à de testar uma certa hipótese sobre um parâmetro populacional θ .

Em essência, testar uma hipótese é recolher evidências nos dados amostrais, que justifiquem a rejeição ou a não rejeição de uma certa afirmação (i) sobre um parâmetro populacional ou (ii) sobre a forma de um modelo distributivo, tendo-se em conta as probabilidades de serem tomadas decisões incorretas. Os testes de hipóteses podem ser classificados em *paramétricos* ou *não paramétricos*. Eles são ditos paramétricos se os dados amostrais, por premissa, tiverem sido extraídos de uma população Normal ou de qualquer outra população, cujo modelo distributivo é conhecido ou previamente especificado. Ao contrário, os testes não paramétricos não necessitam da especificação prévia do modelo distributivo da população, da qual foram extraídos os dados amostrais. De fato, em geral, os testes não paramétricos não são formulados com base nas observações amostrais, propriamente ditas, e, sim, em algumas de suas características ou atributos, tais como, ordens de classificação ou número de diferenças positivas ou negativas entre dados. Do ponto de vista da hipótese a ser testada, os testes mais freqüentes são aqueles que se referem a afirmações sobre um parâmetro populacional. Quando a hipótese, a ser testada, diz respeito à forma do modelo distributivo da população de onde a amostra foi extraída, os testes são denominados de aderência.

No presente capítulo, abordaremos, nos itens iniciais, as linhas gerais, segundo as quais os testes de hipóteses são construídos. Em seguida, ilustraremos esses procedimentos gerais, com a formulação dos testes de hipóteses paramétricos mais conhecidos, para populações normais. Na seqüência, descreveremos a lógica inerente aos testes não paramétricos, concentrando-nos naqueles de maior aplicação às variáveis hidrológicas. Nos itens finais, abordaremos os testes de aderência, enfatizando os testes do Qui-Quadrado, de Kolmogorov-Smirnov, de Anderson-Darling e de Filliben, bem como o teste de Grubbs e Beck, para a detecção de pontos amostrais atípicos, os quais são de grande utilidade na análise de frequência de variáveis hidrológicas.

7.1 – Os Elementos de um Teste de Hipótese

Os procedimentos gerais para a realização de um teste de hipótese são:

- Formule a hipótese a ser testada, denotando-a por H_0 e denominando-a hipótese nula. Essa pode ser, por exemplo, a declaração conjectural de que não houve, nos últimos trinta anos, uma alteração da vazão média anual μ_1 , de uma certa seção fluvial, quando comparada à média μ_0 , do período anterior. Se a hipótese nula é verdadeira, qualquer diferença entre as médias populacionais μ_1 e μ_0 é devida meramente a flutuações das amostras extraídas de uma única população. A hipótese nula é expressa por $H_0: \mu_1 - \mu_0 = 0$.

- Formule a hipótese alternativa e denote-a por H_1 . De acordo com o exemplo da etapa anterior, a hipótese alternativa, e contrária a H_0 , é expressa por $H_1: \mu_1 - \mu_0 \neq 0$.

- Especifique uma estatística de teste T , que esteja em acordo com as hipóteses nula e alternativa, anteriormente formuladas. No exemplo em foco, a estatística de teste deve ter como base a diferença $T = \bar{X}_1 - \bar{X}_0$, entre as médias observadas nos períodos correspondentes às médias populacionais a serem testadas.

- Especifique a distribuição de amostragem da estatística de teste, de acordo com a hipótese nula, bem como com a distribuição de probabilidades da população de onde as observações foram extraídas. No exemplo em foco, caso as vazões médias anuais tenham sido extraídas de uma população Normal, sabe-se que é possível deduzir a distribuição de amostragem da estatística de teste T .

- Especifique a região de rejeição R , ou região crítica R , para a estatística de teste. A especificação da região crítica depende da definição prévia do nível de significância α , o qual, conforme menção anterior, cumpre o papel de remover o grau de subjetividade associado à tomada de decisão. No exemplo em foco, o nível de significância 100α poderia ser arbitrado, por exemplo, em 5 %, o que resultaria na fixação dos limites $[T_{0,025}, T_{0,975}]$, respectivamente abaixo e acima dos quais inicia-se a região de rejeição R .

- Verifique se a estatística de teste \hat{T} , estimada a partir das observações amostrais, está dentro ou fora dos limites estabelecidos para a região de rejeição R . No exemplo, se $\hat{T} < T_{0,025}$ ou se $\hat{T} > T_{0,975}$, a hipótese nula H_0 deve ser rejeitada; nesse caso, interpreta-se que a diferença $\mu_1 - \mu_0$ é significativa, a um nível $\alpha = 0,05$. Caso contrário, se \hat{T} estiver dentro dos limites $[T_{0,025}, T_{0,975}]$, a decisão é a de não rejeitar a hipótese H_0 , implicando que não há diferença significativa entre as médias populacionais μ_1 e μ_0 .

Nos procedimentos gerais, anteriormente delineados, o exemplo citado refere-se a diferenças positivas ou negativas entre μ_1 e μ_0 , o que implica que a região crítica R estende-se pelas duas caudas da distribuição de amostragem da estatística de teste T . Nesse caso, diz-se que o teste é bilateral. Se a hipótese nula tivesse sido formulada de modo diferente, tal como $H_0: \mu_1 > \mu_0$ ou $H_0: \mu_1 < \mu_0$, o teste seria unilateral porque a região crítica se estenderia por apenas uma das caudas da distribuição de amostragem da estatística de teste.

Depreende-se, dos procedimentos gerais, que há uma relação estreita entre os testes de hipóteses e a construção de intervalos de confiança. Para melhor esclarecer esse fato, considere a hipótese nula $H_0: \mu = \mu_0$, a respeito da média de uma população Normal de variância conhecida e igual a σ^2 . Sob essas circunstâncias, sabe-se que, para uma amostra de tamanho N , a estatística de teste é

$T = (\bar{X} - \mu_0) / \sigma / \sqrt{N}$ e que a distribuição de probabilidades dessa estatística de teste é a Normal padrão. Nesse caso, se fixarmos o nível de significância em $\alpha = 0,05$, o teste bilateral estaria definido para a região crítica abaixo de $T_{\alpha/2=0,025} = z_{0,025} = -1,96$ e acima de $T_{1-\alpha/2=0,975} = z_{0,975} = +1,96$. Se, a esse nível de significância, H_0 não foi rejeitada, verifica-se que tal decisão teve como argumento os fatos que $\hat{T} > T_{0,025}$ ou $\hat{T} < T_{0,975}$, os quais são equivalentes a $\bar{X} > \mu_0 - 1,96\sigma/\sqrt{N}$ ou $\bar{X} < \mu_0 + 1,96\sigma/\sqrt{N}$. Manipulando essas desigualdades, é possível colocá-las sob a forma $\bar{X} - 1,96\sigma/\sqrt{N} < \mu_0 < \bar{X} + 1,96\sigma/\sqrt{N}$, a qual é a expressão do intervalo a 100 (1- α) = 95% de confiança para a média μ_0 . Por meio desse exemplo, verifica-se a estreita ligação, no sentido matemático, entre a construção de intervalos de confiança e os testes de hipóteses. A despeito dessa ligação, entretanto, as duas técnicas servem a propósitos diferentes: enquanto o intervalo de confiança estabelece o quão acurado é o conhecimento de μ , o teste de hipótese indica se é plausível assumir o valor μ_0 para μ .

De acordo com o exposto, a rejeição da hipótese nula acontece quando a estimativa da estatística de teste encontrar-se dentro da região crítica. A decisão de rejeitar a hipótese nula é o mesmo que declarar que a estatística de teste é estatisticamente significativa. Em outros termos, no contexto de $H_0: \mu = \mu_0$ e de $\alpha = 0,05$, se as diferenças observadas ocorrem, de modo aleatório, em menos de 5 de 100 testes idênticos, então, os resultados são considerados estatisticamente significativos e a hipótese nula deve ser rejeitada. Por outro lado, a falta de evidências empíricas para rejeitar a hipótese nula, não implica na imediata aceitação de H_0 e, sim, em sua eventual reformulação, seguida de verificações suplementares.

Supondo que a hipótese nula é, de fato, *verdadeira*, a probabilidade de que H_0 seja rejeitada é dada por

$$P(T \in R | H_0 \text{ verdadeira}) = P(T \in R | H_0) = \alpha \quad (7.1)$$

É evidente que se uma hipótese verdadeira é rejeitada, tomou-se uma decisão incorreta. O erro decorrente dessa decisão é denominado *erro do tipo I*. Da equação 7.1, resulta que a probabilidade de ocorrer o erro do tipo I é expressa por

$$P(\text{Erro do Tipo I}) = P(T \in R | H_0) = \alpha \quad (7.2)$$

Na ausência de erro, ou seja, se uma hipótese verdadeira H_0 não é rejeitada, a probabilidade dessa decisão é complementar à probabilidade do erro do tipo I. Em termos formais,

$$P(T \notin R|H_0) = 1 - \alpha \quad (7.3)$$

Contrariamente, não rejeitar a hipótese nula quando ela é, de fato, *falsa*, é outra possível decisão incorreta. O erro decorrente dessa decisão é denominado *erro do tipo II*. A probabilidade de ocorrer o erro do tipo II é expressa por

$$P(\text{Erro do Tipo II}) = P(T \notin R|H_1) = \beta \quad (7.4)$$

Na ausência de erro, ou seja, se uma hipótese falsa H_0 é rejeitada, a probabilidade dessa decisão é complementar à probabilidade do erro do tipo II. Em termos formais,

$$P(T \in R|H_1) = 1 - \beta \quad (7.5)$$

A probabilidade complementar a β , expressa pela equação 7.5, é denominada *poder do teste* e, como se verá mais adiante, é um importante critério de comparação entre diferentes testes de hipóteses.

Os erros dos tipos I e II estão fortemente relacionados. Para demonstrar essa relação, considere que a Figura 7.1 ilustra um teste unilateral de uma hipótese nula $H_0: \mu = \mu_0$ contra a hipótese alternativa $H_1: \mu = \mu_1$, onde μ representa a média de uma população Normal e $\mu_1 > \mu_0$.

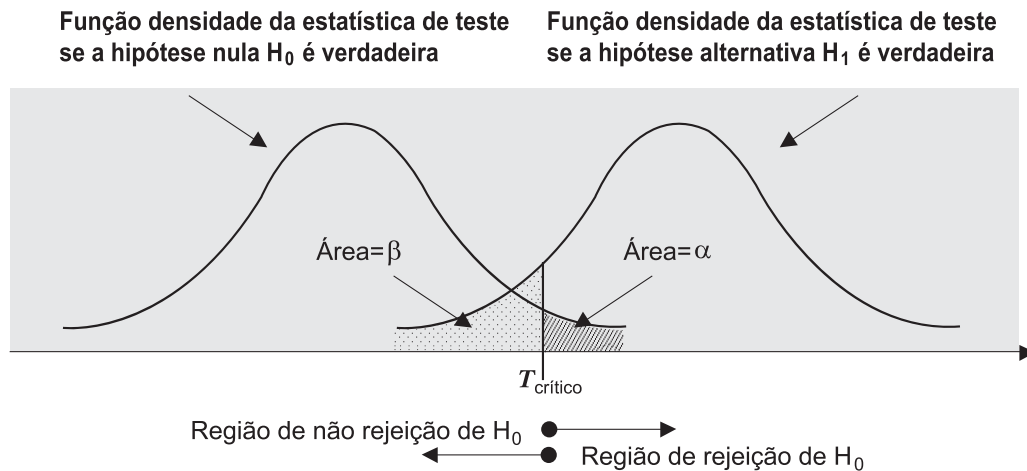


Figura 7.1 – Ilustração dos erros dos tipos I e II em um teste de hipótese unilateral

Se a estatística de teste T é superior ao valor $T_{\text{crítico}}$, a hipótese nula é rejeitada, a um nível de significância α . Nesse caso, supondo que H_0 é verdadeira, a decisão de rejeitá-la é incorreta e a probabilidade de se cometer esse erro é α . Contrariamente, se a estatística de teste T é inferior ao valor $T_{\text{crítico}}$, a hipótese

nula não é rejeitada, a um nível de significância α . Nesse outro caso e supondo que, desta feita, H_1 é verdadeira, a decisão de não rejeitar a hipótese falsa H_0 também é incorreta e a probabilidade de se cometer esse erro é β . Pela ilustração da Figura 7.1, é evidente que a diminuição de α irá levar o valor de $T_{\text{crítico}}$ mais para a direita, causando um aumento de β . Conclui-se, portanto, que diminuir a probabilidade de se cometer um erro do tipo I provoca o aumento da probabilidade de se cometer o erro do tipo II. A situação inversa é igualmente verdadeira.

É óbvio que, ao realizar um teste de hipótese, não se quer tomar uma decisão incorreta e que, portanto, a situação desejável é a de minimizar as probabilidades de se cometer erros de ambos os tipos. Entretanto, em função da dependência entre α e β , ilustrada pela Figura 7.1, bem como das diferentes características dos erros dos tipos I e II, é forçosa uma solução de compromisso no planejamento das regras de decisão de um teste de hipóteses. Em geral, essa solução de compromisso passa pela prescrição prévia de um determinado nível de significância α , tal que ele seja suficientemente pequeno para que β encontre-se em uma faixa aceitável de variação. Essa estratégia de ação advém do fato que, em geral, é possível prescrever antecipadamente o nível α , enquanto tal possibilidade não existe para a probabilidade β . Essa afirmação é justificada pela constatação de que a hipótese alternativa é mais genérica do que a hipótese nula; por exemplo, a hipótese alternativa $H_1: \mu_1 - \mu_0 \neq 0$ compreende a união de diversas outras hipóteses alternativas (e.g.: $H_1: \mu_1 - \mu_0 < 0$ ou $H_1: \mu_1 - \mu_0 > 0$), enquanto a hipótese nula $H_0: \mu_1 - \mu_0 = 0$ é completamente definida. Em outras palavras, enquanto α depende apenas da hipótese nula, β irá depender de qual das hipóteses alternativas é de fato verdadeira, o que, evidentemente, não se sabe *a priori*. Na prática, é considerado razoável prescrever, antecipadamente, o nível de significância α em 0,05, o que implica em uma média de 5 rejeições incorretas de H_0 , em 100 decisões possíveis. Se as conseqüências de um erro do tipo I forem muito graves, pode-se escolher um nível de significância ainda menor, como $\alpha = 0,01$ ou $\alpha = 0,001$.

Embora β dependa de qual hipótese alternativa H_1 é, de fato, verdadeira e, portanto, não possa ser antecipadamente prescrito, é útil o estudo do comportamento de β , sob diferentes possibilidades para H_1 . Essa investigação é feita por meio da quantidade $1 - \beta$, a qual, conforme menção anterior, é denominada poder do teste. Na Figura 7.1, o poder do teste, para a hipótese alternativa específica $H_1: \mu = \mu_1$, pode ser visualizado pela área da função densidade da estatística de teste, sob H_1 , à direita da abscissa $T_{\text{crítico}}$. Para outra hipótese alternativa, por exemplo $H_1: \mu = \mu_2$, é claro que o poder do teste teria outro valor. As relações entre β , ou $(1 - \beta)$, e uma seqüência contínua de hipóteses alternativas específicas, definem, respectivamente, a *curva característica operacional*, ou a *função poder de teste*, as quais permitem distinguir e comparar testes diferentes.

Para exemplificar a construção da curva característica operacional e da função poder do teste, considere o seguinte teste bilateral da média de uma população Normal de parâmetros μ e σ : $H_0: \mu = \mu_0$ contra o conjunto de hipóteses alternativas $H_1: \mu \neq \mu_0$. Mais uma vez, a estatística de teste, nesse caso, é $T = (\bar{X} - \mu_0) / (\sigma / \sqrt{N})$, a qual segue uma distribuição $N(0,1)$. O numerador da estatística de teste pode ser alterado para expressar deslocamentos $\mu_0 + k$, em relação a μ_0 , onde k denota uma constante positiva ou negativa. Desse modo, com $T = k\sqrt{N}/\sigma$, o teste refere-se a $H_0: \mu = \mu_0$ contra um conjunto de deslocamentos padronizados $k\sqrt{N}/\sigma$, em relação a zero, ou, equivalentemente, contra um conjunto de deslocamentos $\mu_0 + k$, em relação a μ_0 . O erro do tipo II corresponde a não rejeitar H_0 , quando H_1 é verdadeira, o que irá acontecer quando a estatística de teste T satisfizer $-z_{\alpha/2} \leq T \leq +z_{\alpha/2}$, onde $-z_{\alpha/2}$ e $z_{\alpha/2}$ representam os limites de definição da região crítica. A probabilidade β de se cometer o erro do tipo II pode ser escrita como $\beta = \Phi(z_{\alpha/2} - k\sqrt{N}/\sigma) - \Phi(-z_{\alpha/2} - k\sqrt{N}/\sigma)$, onde $\Phi(\cdot)$ denota a FAP da distribuição Normal padrão. Portanto, percebe-se que β depende de α , de N e das diferentes hipóteses alternativas dadas por k/σ . Essa dependência pode ser expressa graficamente por meio da curva característica operacional, ilustrada na Figura 7.2, para $\alpha = 0,10$ ($z_{0,05} = 1,645$), amostras de tamanho N variável entre 1 e 50, e $k/\sigma = 0,25, 0,50, 0,75$ e 1.

O exame da curva característica operacional mostra que, para uma amostra de tamanho N fixo, a probabilidade de se cometer o erro do tipo II decresce, quando

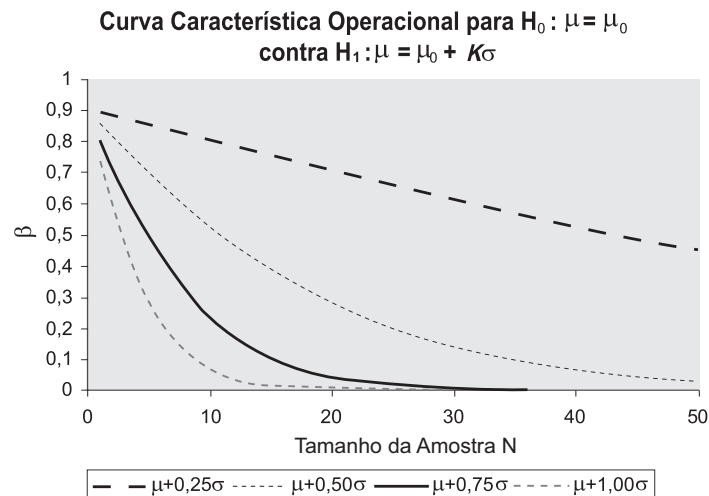


Figura 7.2 – Exemplos da curva característica operacional de um teste de hipóteses

k/σ aumenta. Isso equivale a dizer que pequenas diferenças na média são mais difíceis de detectar, o que conduz a maiores probabilidades de se tomar a decisão incorreta de não rejeitar uma falsa hipótese nula. Observa-se também um decréscimo de β , com o aumento de N , demonstrando a menor probabilidade de se cometer um erro do tipo II, quando o teste tem, como base, amostras de maior tamanho.

A função poder de teste é dada pelo complemento de β , em relação a 1, e encontra-se ilustrada na Figura 7.3, para o exemplo em foco. O poder do teste, conforme definição anterior, representa a probabilidade de se tomar a decisão correta de rejeitar uma falsa hipótese nula, em favor de uma hipótese alternativa. A Figura 7.3 mostra que, para amostras de mesmo tamanho, a probabilidade de não se cometer o erro do tipo II cresce, quando k/σ aumenta. Do mesmo modo, o poder do teste aumenta quando o tamanho da amostra cresce.

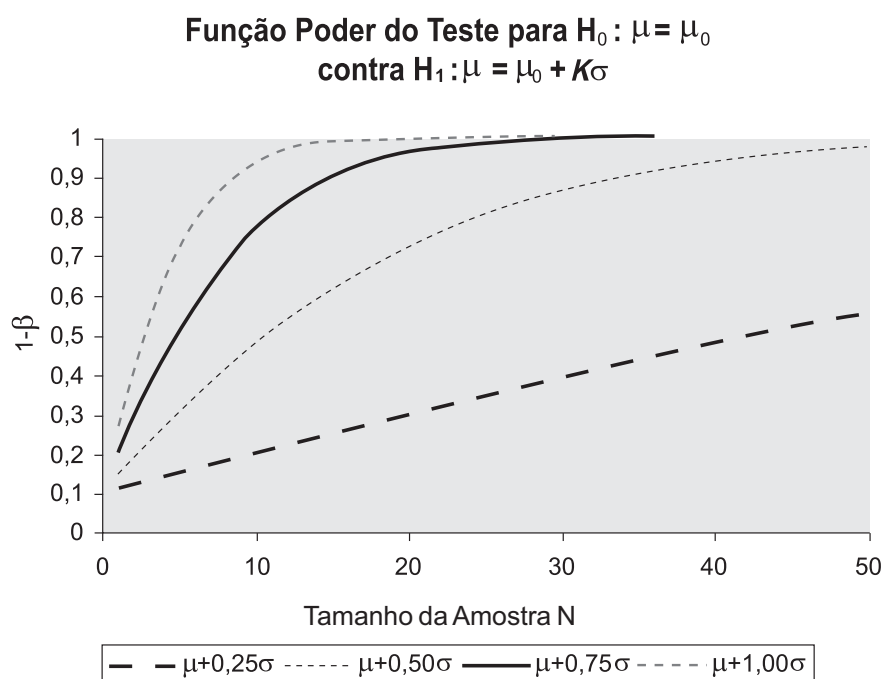


Figura 7.3 - Exemplos de função poder de um teste de hipóteses

As Figuras 7.2 e 7.3 mostram que se, por exemplo, desejarmos manter simultaneamente as respectivas probabilidades de se cometer os erros dos tipos I e II em $100\alpha = 10\%$ e $100\beta = 20\%$, e se estivermos testando a hipótese nula $H_0 : \mu = \mu_0$, contra a hipótese alternativa $H_1 : \mu = \mu_0 + 0,5\sigma$, necessitaríamos de uma amostra de tamanho pelo menos igual a 26. Para esse exemplo, se uma amostra

de tamanho 26 não estiver disponível ou se a obtenção de observações adicionais for excessivamente onerosa, o analista deve buscar alguma solução de compromisso entre a confiabilidade do teste, imposta pela especificação de α e β , e a disponibilidade e/ou ônus de amostragem suplementar. No restante desse capítulo, nos restringiremos a testes de hipóteses que têm como base uma amostra de tamanho fixo, sob a especificação prévia de um nível de significância usual, digamos $100\alpha = 5\%$ ou 10% , aceitando implicitamente os níveis de β decorrentes dessa decisão.

7.2 – Alguns Testes Paramétricos Usuais para Populações Normais

Grande parte da construção matemática em torno dos testes paramétricos de hipóteses refere-se a populações normais. Essa constatação pode ser explicada, primeiramente, pela possibilidade de dedução das distribuições de amostragem de variáveis normais, e, em segundo lugar, pela ampla extensão de aplicações do teorema do limite central. O que se apresenta a seguir é uma descrição dos principais testes paramétricos para populações normais, incluindo as premissas e as estatísticas de teste, sob as quais são construídos. Para que tais testes produzam resultados rigorosos, as premissas devem ser rigorosamente observadas. Em alguns casos práticos e como decorrência do teorema do limite central, pode-se cogitar a extensão desses testes paramétricos para populações não-normais. Deve-se ressaltar, entretanto, que os resultados dessa extensão serão apenas aproximados. Em geral, o grau de aproximação, nesses casos, é dado pela diferença entre o verdadeiro nível de significância do teste, o qual, pode ser avaliado por meio de simulações de Monte Carlo, e o nível previamente estabelecido.

7.2.1 – Testes Paramétricos sobre a Média de uma Única População Normal

A premissa básica dos testes, descritos a seguir, é a de que as variáveis aleatórias independentes $\{X_1, X_2, \dots, X_N\}$, componentes de uma certa amostra aleatória simples, foram todas extraídas de uma única população normal, de média μ desconhecida. O conhecimento ou o desconhecimento da variância populacional σ^2 determina a estatística de teste a ser usada.

• $H_0: \mu = \mu_1$ contra $H_1: \mu = \mu_2$. Atributo de σ^2 : conhecida.

$$\text{Estatística de teste: } Z = \frac{\bar{X} - \mu_1}{\sigma / \sqrt{N}}$$

Distribuição de probabilidades da estatística de teste: Normal $N(0,1)$

Tipo de Teste: unilateral a um nível de significância α

Decisão:

Se $\mu_1 > \mu_2$, rejeitar H_0 se $\frac{\bar{X} - \mu_1}{\sigma/\sqrt{N}} < -z_{1-\alpha}$

Se $\mu_1 < \mu_2$, rejeitar H_0 se $\frac{\bar{X} - \mu_1}{\sigma/\sqrt{N}} > +z_{1-\alpha}$

• $H_0: \mu = \mu_1$ contra $H_1: \mu = \mu_2$. Atributo de σ^2 : desconhecida e estimada por s_X^2 .

Estatística de teste: $T = \frac{\bar{X} - \mu_1}{s_X/\sqrt{N}}$

Distribuição de probabilidades da estatística de teste: t de Student com $v = N-1$ ou t_{N-1}

Tipo de Teste: unilateral a um nível de significância α

Decisão:

Se $\mu_1 > \mu_2$, rejeitar H_0 se $\frac{\bar{X} - \mu_1}{\sigma/\sqrt{N}} < -t_{1-\alpha, v=N-1}$

Se $\mu_1 < \mu_2$, rejeitar H_0 se $\frac{\bar{X} - \mu_1}{\sigma/\sqrt{N}} > +t_{1-\alpha, v=N-1}$

• $H_0: \mu = \mu_0$ contra $H_1: \mu \neq \mu_0$. Atributo de σ^2 : conhecida.

Estatística de teste: $Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{N}}$

Distribuição de probabilidades da estatística de teste: Normal $N(0,1)$

Tipo de Teste: bilateral a um nível de significância α

Decisão:

Rejeitar H_0 se $\left| \frac{\bar{X} - \mu_0}{\sigma/\sqrt{N}} \right| > z_{1-\alpha/2}$

• $H_0: \mu = \mu_0$ contra $H_1: \mu \neq \mu_0$. Atributo de σ^2 : desconhecida e estimada por s_X^2 .

Estatística de teste: $T = \frac{\bar{X} - \mu_0}{s_X/\sqrt{N}}$

Distribuição de probabilidades da estatística de teste: t de Student com $v = N-1$ ou t_{N-1}

Tipo de Teste: bilateral a um nível de significância α

Decisão:

$$\text{Rejeitar } H_0 \text{ se } \left| \frac{\bar{X} - \mu_0}{s_x / \sqrt{N}} \right| > t_{1-\alpha/2, v=N-1}$$

Exemplo 7.1 – Considere as vazões médias do mês de Julho do Rio Paraopeba em Ponte Nova do Paraopeba, listadas no Anexo 1, para o período de 1938 a 1999. Teste a hipótese de que a média populacional do mês de Julho é 47,65 m³/s, a um nível de significância $100\alpha = 5\%$.

Solução: A premissa básica é a que as vazões médias do mês de Julho, em Ponte Nova do Paraopeba, seguem uma distribuição Normal. A amostra de 62 observações fornece $\bar{X} = 44,526$ e $s_x = 12,406$ m³/s, não havendo nenhuma informação adicional sobre a variância populacional. Nesse caso, a hipótese nula é $H_0: \mu = 47,65$ contra a hipótese alternativa $H_1: \mu \neq 47,65$. Trata-se, portanto, de um teste bilateral ao nível

$100\alpha = 5\%$, com a estatística de teste dada por $T = \frac{\bar{X} - 47,65}{s_x / \sqrt{N}}$, a qual

possui uma distribuição t de Student com 61 graus de liberdade. Substituindo os valores amostrais, resulta que o valor absoluto da estimativa de T é igual a 1,9828. A tabela de t de Student, do Anexo 7, fornece $t_{0,975, v=61} = 1,9996$. Como $1,9828 < 1,9996$, a hipótese H_0 não deve ser rejeitada, em favor de H_1 . Em outras palavras, com base na amostra disponível, não há evidências de que a média populacional difira significativamente de 47,65 m³/s, ou seja, que a diferença existente entre a média amostral $\bar{X} = 44,526$ e a média hipotética $\mu = 47,65$ deve-se unicamente a flutuações aleatórias das observações.

Exemplo 7.2 – Repita o exemplo 7.1, supondo que a variância populacional σ^2 seja conhecida e igual a 153,9183 (m³/s)².

Solução: A premissa básica continua sendo a de que as vazões médias do mês de Julho, em Ponte Nova do Paraopeba, seguem uma distribuição Normal. O fato de que a variância populacional é conhecida altera a estatística de teste. Nesse caso, trata-se de um teste bilateral ao nível $100\alpha = 5\%$, com a estatística

de teste dada por $Z = \frac{\bar{X} - 47,65}{\sigma / \sqrt{N}}$, a qual possui uma distribuição $N(0,1)$.

Substituindo os valores amostrais, resulta que o valor absoluto da estimativa de Z é igual a 1,9828. A tabela 5.1, do capítulo 5, fornece $z_{0,975} = 1,96$.

Como $1,9828 > 1,96$, a hipótese H_0 deve ser rejeitada, em favor de H_1 . Portanto, sob as condições estipuladas para esse caso, é significativa a diferença entre a média amostral $\bar{X} = 44,526$ e a média hipotética $\mu = 47,65$.

7.2.2 – Testes Paramétricos sobre as Médias de Duas Populações Normais

A premissa básica dos testes, descritos a seguir, é a de que as variáveis aleatórias independentes $\{X_1, X_2, \dots, X_N\}$ e $\{Y_1, Y_2, \dots, Y_M\}$, componentes de duas amostras aleatórias simples de tamanhos iguais a N e M , foram extraídas de duas populações normais, de respectivas médias μ_X e μ_Y desconhecidas. O conhecimento ou o desconhecimento das variâncias populacionais σ_X^2 e σ_Y^2 , assim como a condição de igualdade entre elas, determinam a estatística de teste a ser usada. Os testes descritos a seguir são tomados como bilaterais, podendo ser transformados em unilaterais pela modificação de H_1 e de α .

• $H_0: \mu_X - \mu_Y = \delta$ contra $H_1: \mu_X - \mu_Y \neq \delta$.

Atributos de σ_X^2 e σ_Y^2 : conhecidas

$$\text{Estatística de teste: } Z = \frac{(\bar{X} - \bar{Y}) - \delta}{\sqrt{\frac{\sigma_X^2}{N} + \frac{\sigma_Y^2}{M}}}$$

Distribuição de probabilidades da estatística de teste: Normal $N(0,1)$

Tipo de Teste: bilateral a um nível de significância α

Decisão:

$$\text{Rejeitar } H_0 \text{ se } \left| \frac{(\bar{X} - \bar{Y}) - \delta}{\sqrt{\frac{\sigma_X^2}{N} + \frac{\sigma_Y^2}{M}}} \right| > z_{1-\frac{\alpha}{2}}$$

• $H_0: \mu_X - \mu_Y = \delta$ contra $H_1: \mu_X - \mu_Y \neq \delta$.

Atributos de σ_X^2 e σ_Y^2 : supostamente iguais, mas desconhecidas.

Estimadas por s_X^2 e s_Y^2 .

$$\text{Estatística de teste: } T = \frac{(\bar{X} - \bar{Y}) - \delta}{\sqrt{(N-1)s_X^2 + (M-1)s_Y^2}} \sqrt{\frac{NM(N+M-2)}{N+M}}$$

Distribuição de probabilidades da estatística de teste: t de Student com $v = N + M - 2$

Tipo de Teste: bilateral a um nível de significância α

Decisão:

Rejeitar H_0 se

$$\left| \frac{(\bar{X} - \bar{Y}) - \delta}{\sqrt{(N-1)s_X^2 + (M-1)s_Y^2}} \sqrt{\frac{NM(N+M-2)}{N+M}} \right| > t_{1-\frac{\alpha}{2}, v=N+M-2}$$

• $H_0: \mu_X - \mu_Y = \delta$ contra $H_1: \mu_X - \mu_Y \neq \delta$.

Atributos de σ_X^2 e σ_Y^2 : supostas desiguais, mas desconhecidas.

Estimadas por s_X^2 e s_Y^2 .

$$\text{Estatística de teste: } T = \frac{(\bar{X} - \bar{Y}) - \delta}{\sqrt{(s_X^2/N) + (s_Y^2/M)}}$$

Distribuição de probabilidades da estatística de teste: segundo Casella e Berger (1990), a distribuição de T pode ser *aproximada* por uma

$$\text{distribuição } t \text{ de Student com } v = \frac{[(s_X^2/N) + (s_Y^2/M)]^2}{\frac{(s_X^2/N)^2}{N-1} + \frac{(s_Y^2/M)^2}{M-1}}$$

Tipo de Teste: bilateral a um nível de significância α

Decisão:

$$\text{Rejeitar } H_0 \text{ se } \left| \frac{(\bar{X} - \bar{Y}) - \delta}{\sqrt{(s_X^2/N) + (s_Y^2/M)}} \right| > t_{1-\frac{\alpha}{2}, v}$$

Exemplo 7.3 – Considere as vazões médias do mês de Julho do Rio Paraopeba em Ponte Nova do Paraopeba, listadas no Anexo 1, separando-as em duas amostras iguais de mesmo tamanho: a amostra denotada por X , para o período de 1938 a 1968, e a amostra Y , para o período de 1969 a 1999. Teste a hipótese de que, considerados os períodos de 1938-1968 e de 1969-1999, as médias populacionais do mês de Julho não sofreram alterações importantes, a um nível de significância $100\alpha = 5\%$.

Solução: A premissa básica é a que, considerados os períodos de 1938-1968 e de 1969-1999, as vazões médias do mês de Julho, em Ponte Nova do Paraopeba, seguem duas distribuições normais de médias μ_X e μ_Y , com variâncias σ_X^2 e σ_Y^2 supostamente desiguais e desconhecidas. A amostra de 31 observações, para o período de 1938 a 1968, fornece $\bar{X} = 45,08$ e $s_X = 11,505 \text{ m}^3/\text{s}$, enquanto, para o período restante, esses valores resultam ser $\bar{Y} = 43,97$ e $s_Y = 13,415 \text{ m}^3/\text{s}$. Nesse caso, a

hipótese nula é $H_0: \mu_x - \mu_y = \delta = 0$ contra a hipótese alternativa $H_1: \mu_x - \mu_y = \delta \neq 0$. Como as variâncias são supostamente desiguais e devem ser estimadas pelas variâncias amostrais, a estatística de teste é

$$T = \frac{(\bar{X} - \bar{Y})}{\sqrt{\left(\frac{s_x^2}{31}\right) + \left(\frac{s_y^2}{31}\right)}}, \text{ a distribuição de probabilidades da qual pode ser aproximada por uma } t \text{ de Student com } v = \frac{\left[\left(\frac{s_x^2}{31}\right) + \left(\frac{s_y^2}{31}\right)\right]^2}{\left[\frac{\left(\frac{s_x^2}{31}\right)^2}{30} + \frac{\left(\frac{s_y^2}{31}\right)^2}{30}\right]} = 58 \text{ graus}$$

de liberdade. Substituindo os valores amostrais, resulta que o valor absoluto da estimativa de T é igual a 0,3476. A tabela de t de Student, do Anexo 7, fornece $t_{0,975,v=58} = 2,00$. Como $0,3476 < 2,00$, a hipótese H_0 não deve ser rejeitada, em favor de H_1 . Em outras palavras, com base nas amostras disponíveis, não há evidências de que as médias populacionais, dos períodos considerados, difiram significativamente entre si, ao nível de $100\alpha = 5\%$.

7.2.3 – Testes Paramétricos sobre a Variância de uma Única População Normal

A premissa básica dos testes, descritos a seguir, é a de que as variáveis aleatórias independentes $\{X_1, X_2, \dots, X_N\}$, componentes de uma certa amostra aleatória simples, foram todas extraídas de uma única população normal, de variância σ^2 desconhecida. O conhecimento ou o desconhecimento da média populacional μ determina a estatística de teste a ser usada. Os testes são tomados como bilaterais, podendo ser transformados em unilaterais pela modificação de H_1 e de α .

$$\bullet H_0: \sigma^2 = \sigma_0^2 \text{ contra } H_1: \sigma^2 \neq \sigma_0^2.$$

Atributo de μ : conhecida.

$$\text{Estatística de teste: } Q = \frac{\sum_{i=1}^N (X_i - \mu)^2}{\sigma_0^2} = N \frac{s_x^2}{\sigma_0^2}$$

Distribuição de probabilidades da estatística de teste: χ^2 com $v = N$, ou χ_N^2

Tipo de Teste: bilateral a um nível de significância α

Decisão:

$$\text{Rejeitar } H_0 \text{ se } N \frac{s_x^2}{\sigma_0^2} < \chi_{\frac{\alpha}{2}, N}^2 \text{ ou se } N \frac{s_x^2}{\sigma_0^2} > \chi_{1-\frac{\alpha}{2}, N}^2$$

• $H_0: \sigma^2 = \sigma_0^2$ contra $H_1: \sigma^2 \neq \sigma_0^2$.

Atributo de μ : desconhecida, estimada por \bar{X} .

$$\text{Estatística de teste: } K = \frac{\sum_{i=1}^N (X_i - \bar{X})^2}{\sigma_0^2} = (N-1) \frac{s_x^2}{\sigma_0^2}$$

Distribuição de probabilidades da estatística de teste: χ^2 com $v = N-1$, ou χ_{N-1}^2

Tipo de Teste: bilateral a um nível de significância

Decisão:

$$\text{Rejeitar } H_0 \text{ se } (N-1) \frac{s_x^2}{\sigma_0^2} < \chi_{\frac{\alpha}{2}, N-1}^2 \text{ ou se } (N-1) \frac{s_x^2}{\sigma_0^2} > \chi_{1-\frac{\alpha}{2}, N-1}^2$$

Exemplo 7.4 – Considere novamente as vazões médias do mês de Julho do Rio Paraopeba em Ponte Nova do Paraopeba, listadas no Anexo 1, para o período de 1938 a 1999. Teste a hipótese nula de que a variância populacional σ_0^2 , das vazões médias do mês de Julho, é de $150 \text{ (m}^3/\text{s)}^2$ contra a hipótese alternativa $H_1: \sigma_0^2 > 150 \text{ (m}^3/\text{s)}^2$, a um nível de significância $100\alpha = 5\%$.

Solução: Novamente, a premissa básica é a que as vazões médias do mês de Julho, em Ponte Nova do Paraopeba, seguem uma distribuição Normal. A amostra de 62 observações fornece $\bar{X} = 44,526$ e $s_x = 12,406 \text{ m}^3/\text{s}$, não havendo nenhuma informação adicional sobre a média populacional. Nesse caso, a hipótese nula é $H_0: \sigma_0^2 = 150$ contra a hipótese alternativa $H_1: \sigma_0^2 > 150$. Trata-se, portanto, de um teste unilateral ao nível

$100\alpha = 5\%$, com a estatística de teste dada por $K = (N-1) \frac{s_x^2}{\sigma_0^2}$, a qual

possui uma distribuição χ^2 com 61 graus de liberdade. Substituindo os valores amostrais, resulta que o valor de K é igual a 62,593. A tabela de χ^2 , do Anexo 6, fornece $\chi_{0,95,61}^2 = 80,232$. Como $62,593 < 80,232$, a hipótese H_0 não deve ser rejeitada, em favor de H_1 . Em outras palavras, com base na amostra disponível, não há evidências de que a variância populacional supere significativamente o valor de $150 \text{ (m}^3/\text{s)}^2$, ou seja, que a diferença existente entre a variância amostral $s_x^2 = 153,918$ e a variância $\sigma_0^2 = 150$ deve-se unicamente a flutuações aleatórias das observações.

7.2.4 – Testes Paramétricos sobre as Variâncias de Duas Populações Normais

A premissa básica dos testes, descritos a seguir, é a de que as variáveis aleatórias independentes $\{X_1, X_2, \dots, X_N\}$ e $\{Y_1, Y_2, \dots, Y_M\}$, componentes de duas amostras aleatórias simples de tamanhos iguais a N e M , foram extraídas de duas populações normais, de respectivas variâncias σ_X^2 e σ_Y^2 desconhecidas. O conhecimento ou o desconhecimento das médias populacionais μ_X e μ_Y determina a estatística de teste a ser usada. Os testes são tomados como bilaterais, podendo ser transformados em unilaterais pela modificação de H_1 e de α .

$$\bullet H_0: \frac{\sigma_X^2}{\sigma_Y^2} = 1 \text{ contra } H_1: \frac{\sigma_X^2}{\sigma_Y^2} \neq 1$$

Atributos de μ_X e μ_Y : conhecidas

$$\text{Estatística de teste: } \varphi = \frac{s_X^2 / \sigma_X^2}{s_Y^2 / \sigma_Y^2}$$

Distribuição de probabilidades da estatística de teste: F de Snedecor com $v_1 = N$ e $v_2 = M$, ou $F_{N,M}$

Tipo de Teste: bilateral a um nível de significância α

Decisão:

Rejeitar H_0 se $\varphi < F_{N,M,\alpha/2}$ ou se $\varphi > F_{N,M,1-\alpha/2}$

$$\bullet H_0: \frac{\sigma_X^2}{\sigma_Y^2} = 1 \text{ contra } H_1: \frac{\sigma_X^2}{\sigma_Y^2} \neq 1$$

Atributos de μ_X e μ_Y : desconhecidas, estimadas por \bar{X} e \bar{Y}

$$\text{Estatística de teste: } f = \frac{s_X^2 / \sigma_X^2}{s_Y^2 / \sigma_Y^2}$$

Distribuição de probabilidades da estatística de teste: F de Snedecor com $v_1 = N - 1$ e $v_2 = M - 1$, ou $F_{N-1,M-1}$

Tipo de Teste: bilateral a um nível de significância α

Decisão:

Rejeitar H_0 se $f < F_{N-1,M-1,\alpha/2}$ ou se $f > F_{N-1,M-1,1-\alpha/2}$

Exemplo 7.5 – Um certo constituinte de um efluente foi analisado 7 e 9 vezes por meio dos procedimentos X e Y, respectivamente. Os resultados das análises apresentaram os seguintes desvios-padrão: $s_X = 1,9$ e $s_Y = 0,8$ mg/l. Teste a hipótese de que o procedimento Y é mais preciso do que o procedimento X, ao nível de significância $100\alpha = 5\%$. (adap. de Kottegoda e Rosso, 1997)

Solução: Supondo tratar-se de duas populações normais, a hipótese nula

a ser testada é $H_0: \frac{\sigma_X^2}{\sigma_Y^2} = 1$ contra a hipótese alternativa $H_1: \frac{\sigma_X^2}{\sigma_Y^2} > 1$ ou $\sigma_X^2 > \sigma_Y^2$.

Trata-se, portanto, de um teste unilateral com $\alpha = 0,05$. A estatística de

teste é $f = \frac{s_X^2/\sigma_X^2}{s_Y^2/\sigma_Y^2}$, a qual segue uma distribuição F de Snedecor com

$v = 7-1 = 6$ e $v_2 = 9-1 = 8$ graus de liberdade para o numerador e denominador, respectivamente. Substituindo os valores amostrais, resulta que $f=5,64$. Da tabela de F, do Anexo 8, lê-se que $F_{6,8,0,05} = 3,58$. Como $5,64 > 3,58$, a decisão é de rejeitar a hipótese nula em favor da hipótese alternativa, ao nível de significância $\alpha = 0,05$. Em outras palavras, conclui-se que a variância dos resultados do procedimento Y é menor do que a de seu concorrente, tratando-se, portanto, de um método de análise mais preciso.

7.3 – Alguns Testes Não-Paramétricos Usuais em Hidrologia

Os testes paramétricos de hipóteses, anteriormente descritos, requerem que a distribuição da variável aleatória, ou das variáveis aleatórias de origem, seja a distribuição Normal. De fato, se a distribuição dos dados originais é a Normal, é possível deduzir as distribuições das estatísticas de testes, em razão, principalmente, da propriedade reprodutiva das variáveis Gaussianas e do teorema do limite central. Entretanto, se a distribuição dos dados originais não é Gaussiana, o uso das distribuições das estatísticas de testes conhecidas terá como consequência, a violação do nível de significância previamente estabelecido. Por exemplo, se T denota a estatística de teste $T = (\bar{X} - \mu_0)/s_X/\sqrt{N}$ para uma variável aleatória X , cujo comportamento se afasta da normalidade, a verdadeira probabilidade de se cometer o erro do tipo I não será necessariamente igual ao nível nominal α . Em outros termos, nesse caso, pode-se escrever que

$$\int_{-\infty}^{-t_{\alpha/2}} f_T(t|H_0)dt + \int_{t_{\alpha/2}}^{\infty} f_T(t|H_0)dt \neq \alpha \quad (7.6)$$

onde $f_T(t)$ é a função densidade desconhecida de $T = (\bar{X} - \mu_0)/s_X/\sqrt{N}$, sob a premissa que X não seja normalmente distribuída.

A estatística matemática apresenta duas soluções possíveis para o problema identificado pela equação 7.6. A primeira solução diz respeito à tentativa de mostrar, via simulações de Monte Carlo, que, mesmo que uma certa variável aleatória X não seja Gaussiana e que, portanto, a estatística de teste T não tenha uma distribuição de probabilidades conhecida, a verdadeira densidade $f_T(t|H_0)$ irá se comportar, em muitos casos, de modo suficientemente similar à distribuição usual, caso X fosse, de fato, normal. Por exemplo, Larsen e Marx (1986) mostram alguns exemplos nos quais, se a distribuição de X não é exageradamente assimétrica ou se o tamanho da amostra não é excessivamente pequeno, a distribuição t de Student pode aproximar satisfatoriamente a distribuição $f_T(t|H_0)$, para testes de hipóteses relativas à média populacional de X . Nesses casos, afirma-se que o teste de t de Student é *robusto*, em relação a desvios moderados da normalidade. Dadas as características marcadamente assimétricas das distribuições de probabilidades de grande parte das variáveis hidrológicas, essa primeira possível solução, para o problema identificado pela equação 7.6, tem aplicações muito limitadas na hidrologia estatística.

A segunda solução possível, para o problema posto pela equação 7.6, é a de substituir as estatísticas de teste convencionais por outras, cujas distribuições de probabilidades permanecem invariáveis, sob a veracidade da hipótese H_0 e a despeito das características distributivas populacionais da variável aleatória de origem X . Os procedimentos de inferência estatística e, particularmente, os testes de hipóteses, que possuem tais propriedades, são denominados *não-paramétricos*. Os procedimentos gerais para a construção de testes paramétricos de hipóteses, alinhados no item 7.1, permanecem os mesmos para os testes não-paramétricos. A diferença fundamental entre eles é que os testes não-paramétricos são formulados com base em estatísticas invariáveis com a distribuição dos dados originais. De fato, as estatísticas de testes não-paramétricos, em sua grande maioria, baseiam-se em características que podem ser deduzidas dos dados amostrais, mas que não os incluem diretamente em seu cálculo. São exemplos dessas características: o número de diferenças positivas (ou negativas) entre uma certa mediana hipotética e os valores amostrais, ou o coeficiente de correlação entre as ordens de classificação dos elementos de duas amostras, ou, ainda, o número de inflexões dos valores amostrais ao longo de uma seqüência de índices de tempo, entre outras.

A variedade e o número de testes não-paramétricos têm crescido enormemente desde que foram introduzidos, no início da década de 1940. Não se tem aqui o objetivo de abordar a formulação e a construção dos inúmeros testes não-paramétricos de hipóteses; o leitor interessado nesses detalhes deve remeter-se a textos especializados, tais como Siegel (1956), Gibbons (1971) e Hollander e

Wolfe (1973). O que se segue é uma descrição, acompanhada de exemplos de aplicação, dos principais testes não-paramétricos de hipóteses empregados em hidrologia. Os testes aqui descritos têm, como objeto principal, a verificação das hipóteses fundamentais da análise de frequência de uma variável hidrológica. De fato, a premissa de base para a aplicação dos métodos estatísticos a um conjunto de observações de uma variável hidrológica, é que se trata de uma amostra aleatória simples, extraída de uma população única, cuja função de distribuição de probabilidades não é conhecida *a priori*. Nessa premissa de base estão implícitas as hipóteses de *aleatoriedade*, *independência*, *homogeneidade* e *estacionariedade*, as quais, pelas características distributivas das variáveis hidrológicas e pelo tamanho típico de suas amostras, podem ser testadas apenas com o emprego dos testes não-paramétricos. Os testes, apresentados a seguir, estão entre os procedimentos não-paramétricos de maior utilidade na hidrologia estatística.

7.3.1 – Teste da Hipótese de Aleatoriedade

No contexto dos fenômenos do ciclo da água, o termo ‘aleatoriedade’ significa, essencialmente, que as flutuações de uma certa variável hidrológica decorrem de causas naturais. Nesse sentido, as vazões de um curso d’água regularizadas pela operação de um reservatório, a montante, constituiriam um exemplo de uma série não-aleatória. A aleatoriedade de uma série hidrológica não pode ser demonstrada, mas pode ser rejeitada pela presença de uma estrutura ou por alguma intervenção de natureza não-aleatória. NERC (1975) sugere que a rejeição/não-rejeição da hipótese de aleatoriedade de uma série hidrológica possa ser decidida por meio do *teste não-paramétrico do número de inflexões*. Particularizando para uma série de vazões máximas anuais Q_t , ou seja, considerando um gráfico entre Q_t versus o ano de ocorrência t , uma inflexão poderia ser tanto um pico, quanto um ‘vale’, nesse diagrama. Um número excessivamente pequeno, ou excessivamente grande, de inflexões é um indicador de não-aleatoriedade. Por outro lado, se uma amostra de N observações é aleatória, pode-se mostrar que o valor esperado do número de inflexões, denotado por p , é dado por

$$E[p] = \frac{2(N-2)}{3} \quad (7.7)$$

com variância que pode ser aproximada por

$$\text{Var}[p] = \frac{16N-29}{90} \quad (7.8)$$

Para amostras de tamanho $N > 30$, é possível provar que a variável p segue aproximadamente uma distribuição Normal. Portanto, se a hipótese nula é H_0 : (a amostra é aleatória), a estatística do teste não-paramétrico pode ser formulada como

$$T = \frac{p - E[p]}{\sqrt{\text{Var}[p]}} \quad (7.9)$$

onde p representa o número de picos, e/ou de ‘vales’, no gráfico da variável aleatória, ao longo do tempo. Por tratar-se de um teste bilateral, a um nível de significância α , a decisão deve ser a de rejeitar a hipótese nula se $|T| > z_{1-\alpha/2}$.

7.3.2 – Teste da Hipótese de Independência

O termo ‘independência’ significa, essencialmente, que nenhuma observação presente na amostra pode influenciar a ocorrência, ou a não ocorrência, de qualquer outra observação seguinte. Mesmo que uma série seja considerada aleatória, as observações que a constituem podem não ser independentes. No contexto de variáveis hidrológicas, os armazenamentos naturais de uma bacia hidrográfica, por exemplo, podem determinar a ocorrência de vazões de maior porte, na sequência de vazões elevadas, ou, contrariamente, de vazões de menor porte, na sequência de vazões reduzidas. A dependência, nesse caso, varia com o intervalo de tempo que separa as observações consecutivas da série hidrológica: forte, para vazões médias diárias, e fraca ou nenhuma, para vazões médias (ou máximas, ou mínimas) anuais. A rejeição ou não-rejeição da hipótese de independência de uma série hidrológica é freqüentemente decidida por meio do *teste não-paramétrico* proposto por Wald e Wolfowitz (1943), o qual encontra-se descrito a seguir.

Dada uma amostra $\{X_1, X_2, \dots, X_N\}$, de tamanho N , e as diferenças $\{X'_1, X'_2, \dots, X'_N\}$, entre as observações X_i e a média amostral \bar{X} , a estatística do teste de Wald-Wolfowitz é dada por

$$R = \sum_{i=1}^{N-1} X'_i X'_{i+1} + X'_1 X'_N \quad (7.10)$$

Sob a hipótese de que as observações são independentes, pode-se demonstrar que a estatística R segue uma distribuição Normal de média igual a

$$E[R] = -\frac{S_2}{N-1} \quad (7.11)$$

e variância dada por

$$\text{Var}[R] = \frac{s_2^2 - s_4}{N-1} + \frac{s_2^2 - 2s_4}{(N-1)(N-2)} - \frac{s_2^2}{(N-1)^2} \quad (7.12)$$

onde r denota a ordem dos momentos amostrais em relação à origem, $s_r = Nm'_r$

e $m'_r = \sum_{i=1}^N (X'_i)^r / N$. Portanto, se a hipótese nula é H_0 : (os elementos da amostra

são independentes), a estatística do teste não-paramétrico de Wald-Wolfowitz pode ser formulada como

$$T = \frac{R - E[R]}{\sqrt{\text{Var}[R]}} \quad (7.13)$$

a qual segue uma distribuição Normal padrão. Por tratar-se de um teste bilateral, a um nível de significância α , a decisão deve ser a de rejeitar a hipótese nula se $|T| > z_{1-\alpha/2}$.

7.3.3 – Teste da Hipótese de Homogeneidade

O termo “homogeneidade” implica que todos os elementos de uma certa amostra provêm de uma única e idêntica população. Em uma série de vazões máximas anuais, por exemplo, muitos de seus valores podem decorrer de enchentes provocadas por precipitações ordinárias ou comuns, enquanto outros advêm de precipitações extraordinariamente elevadas, resultantes de condições hidrometeorológicas muito especiais, como por exemplo, a ocorrência do fenômeno El Niño. Nesse exemplo, temos duas populações de enchentes, diferenciadas pelo seu mecanismo de formação, e, certamente, a série hidrológica deveria ser considerada heterogênea. Entretanto, as amostras hidrológicas, geralmente de tamanhos pequenos, tornam difícil a detecção da heterogeneidade eventualmente presente na série completa. Em geral, é mais fácil identificar heterogeneidades em séries de valores médios ou totais anuais, do que em séries de valores extremos, tomados em intervalos de tempo mais curtos. A rejeição ou não-rejeição da hipótese de homogeneidade de uma série hidrológica é freqüentemente decidida por meio do *teste não-paramétrico* proposto por Mann e Whitney (1947), o qual encontra-se descrito a seguir.

Dada uma amostra $\{X_1, X_2, \dots, X_N\}$, de tamanho N , separe-a em duas sub-amostras $\{X_1, X_2, \dots, X_{N_1}\}$, de tamanho N_1 , e $\{X_{N_1+1}, X_{N_1+2}, \dots, X_N\}$, de tamanho N_2 , de modo que $N_1 + N_2 = N$ e que N_1 e N_2 sejam aproximadamente iguais, com $N_1 \leq N_2$. Em seguida, classifique, em ordem crescente, a amostra completa de tamanho N , indicando a ordem de classificação m de cada observação e se ela provem da

primeira ou da segunda sub-amostra. A idéia intuitiva do teste de Mann-Whitney é se as duas sub-amostras não forem homogêneas, os elementos da primeira apresentarão ordens de classificação consistentemente mais baixas (ou mais altas), em relação às ordens de classificação correspondentes à segunda sub-amostra. A estatística do teste V de Mann-Whitney é dada pelo *menor valor* entre as quantidades

$$V_1 = N_1 N_2 + \frac{N_1(N_1 + 1)}{2} - R_1 \quad (7.14)$$

$$V_2 = N_1 N_2 - V_1 \quad (7.15)$$

onde R_1 denota a *soma das ordens de classificação dos elementos da primeira sub-amostra*. Se $N_1, N_2 > 20$, e sob a hipótese de que se trata de uma amostra homogênea, demonstra-se que V segue uma distribuição Normal de média igual a

$$E[V] = \frac{N_1 N_2}{2} \quad (7.16)$$

e variância dada por

$$\text{Var}[V] = \frac{N_1 N_2 (N_1 + N_2 + 1)}{12} \quad (7.17)$$

Portanto, se a hipótese nula é H_0 : (a amostra é homogênea), a estatística do teste não-paramétrico de Mann-Whitney pode ser formulada como

$$T = \frac{V - E[V]}{\sqrt{\text{Var}[V]}} \quad (7.18)$$

a qual segue uma distribuição Normal padrão. Por tratar-se de um teste bilateral, a um nível de significância α , a decisão deve ser a de rejeitar a hipótese nula se $|T| > z_{1-\alpha/2}$

7.3.4 – Teste da Hipótese de Estacionariedade

O termo “estacionariedade” refere-se ao fato que, excluídas as flutuações aleatórias, as observações amostrais são invariantes, com relação à cronologia de suas ocorrências. Os tipos de não-estacionariedades incluem tendências, ‘saltos’ e ciclos, ao longo do tempo. Em um contexto hidrológico, os “saltos” estão relacionados a alterações bruscas em uma bacia ou trecho fluvial, tais como, por exemplo, a construção de barragens. Os ciclos, por sua vez, podem estar relacionados a flutuações climáticas de longo período, sendo de difícil detecção. As tendências temporais, em geral, estão associadas a alterações graduais que se processam na bacia, tais como, por exemplo, uma evolução temporal lenta da

urbanização de uma certa área geográfica. Uma tendência temporal, eventualmente presente em uma série hidrológica X_t , ao longo do tempo t , pode ser detectada pela correlação entre a série e o índice de tempo. Essa é a idéia essencial do teste não-paramétrico de Spearman, conforme descrito por NERC (1975), cuja base é o coeficiente de correlação entre as ordens de classificação m_t , da sequência X_t , e os índices de tempo T_t , esses iguais a $1, 2, \dots, N$.

A estatística do teste de Spearman tem, como base, o seguinte coeficiente:

$$r_s = 1 - \frac{6 \sum_{t=1}^N (m_t - T_t)^2}{N^3 - N} \quad (7.19)$$

Para $N > 10$ e sob a hipótese nula de que não há correlação entre m_t e T_t , demonstra-se que a distribuição de r_s pode ser aproximada por uma Normal de média igual a

$$E[r_s] = 0 \quad (7.20)$$

e variância dada por

$$\text{Var}[r_s] = \frac{1}{N-1} \quad (7.21)$$

Portanto, se a hipótese nula é H_0 : (a amostra não apresenta tendência temporal), a estatística do teste não-paramétrico de Spearman pode ser formulada como

$$T = \frac{r_s}{\sqrt{\text{Var}[r_s]}} \quad (7.22)$$

a qual segue uma distribuição Normal padrão. Por tratar-se de um teste bilateral, a um nível de significância α , a decisão deve ser a de rejeitar a hipótese nula se $|T| > z_{1-\alpha/2}$.

Exemplo 7.6 – Considere a série de vazões médias anuais do Rio Paraopeba em Ponte Nova do Paraopeba, listadas na Tabela 7.1, e teste as hipóteses de (a) aleatoriedade, (b) independência, (c) homogeneidade e (d) estacionariedade, a um nível de significância $\alpha = 0,05$.

Solução: (a) Teste da hipótese de aleatoriedade. A variação temporal da vazões é mostrada na Figura 7.4. Nessa figura, observa-se que o número total de inflexões importantes é $p = 34$. Para $N = 62$, as equações 7.7 e 7.8 resultam em $E[p] = 40$ e $\text{Var}[p] = 10,7$. Com esses valores, a estatística de teste, da equação 7.9, é $T = -1,8340$. Para o nível de significância $\alpha = 0,05$, $z_{0,975} = 1,96$. Como $|T| < z_{0,975}$, a decisão é a de não rejeitar a hipótese H_0 de que as observações são aleatórias.

(b) Teste da hipótese de independência. A sexta coluna da tabela 7.1 apresenta as diferenças entre as vazões médias anuais e o valor médio global de 86,105 m³/s. São esses os valores necessários para o cálculo da estatística do teste de Wald-Wolfowitz, pela equação 7.10. O resultado desse cálculo é $R = 8253,759$. As diferenças, listadas na tabela 7.1, também fornecem os valores $s_2 = 38003,47$ e $s_4 = 87362890,7$, cuja substituição nas equações 7.11 e 7.12 resultam em $E[R] = -623,01$ e $\text{Var}[R] = 22203003,87$. Com esses valores, a estatística de teste, da equação 7.13, é $T = 1,8839$. Para o nível de significância $\alpha = 0,05$, $z_{0,975} = 1,96$. Como $|T| < z_{0,975}$, a decisão é a de não rejeitar a hipótese H_0 de que as observações são independentes.

(c) Teste da hipótese de homogeneidade. A quarta coluna da tabela 7.1 apresenta as ordens de classificação das vazões médias anuais, denotadas por m_i . São esses os valores necessários para o cálculo da estatística do teste de Mann-Whitney, pelas equações 7.14 e 7.15, lembrando que a soma das ordens de classificação da primeira sub-amostra de 31 elementos, também anotada na tabela 7.1, é $R_1 = 1004$. A estatística de teste é o menor valor entre V_1 e V_2 , ou seja $V = 453$. A substituição de R_1 e V nas equações 7.16 e 7.17 resulta em $E[V] = 480,5$ e $\text{Var}[V] = 71,0299$. Com esses valores, a estatística de teste, da equação 7.18, é $T = 0,3872$. Para o nível de significância $\alpha = 0,05$, $z_{0,975} = 1,96$. Como $|T| < z_{0,975}$, a decisão é a de não rejeitar a hipótese H_0 de que as observações são homogêneas.

(d) Teste da hipótese de estacionariedade. A quarta coluna da tabela 7.1 apresenta as ordens de classificação das vazões médias anuais e a segunda coluna lista o índice de tempo cronológico T_i . São esses os valores necessários para o cálculo da estatística do teste de Spearman, pela equação 7.19. A estatística de teste calculada é $r_s = -0,07618$. As equações 7.20 e 7.21 resultam em $E[r_s] = 0$ e $\text{Var}[r_s] = 0,0164$. Com esses valores, a estatística de teste, da equação 7.22, é $T = 0,5949$. Para o nível de significância $\alpha = 0,05$, $z_{0,975} = 1,96$. Como $|T| < z_{0,975}$, a decisão é a de não rejeitar a hipótese H_0 de que as observações são estacionárias.

Tabela 7.1 – Vazões médias anuais do Rio Paraopeba em Ponte Nova do Paraopeba (m³/s) e grandezas auxiliares para a realização dos testes de hipóteses de Wald-Wolfowitz, Mann-Whitney e Spearman.

Ano Civil	T_i	X_i	m_i	Sub-Amostra	$X'_i = X_i - \bar{X}$	X_i classificados
1938	1	104,3	51	1	18,20	43,6
1939	2	97,9	45	1	11,80	46,8
1940	3	89,2	38	1	3,10	49,4
1941	4	92,7	40	1	6,60	50,1
1942	5	98	46	1	11,90	53,1
1943	6	141,7	60	1	55,60	57

Tabela 7.1 – Continuação

Ano Civil	T_i	X_i	m_i	Sub-Amostra	$X'_i = X_i - \bar{X}$	X_i classificados
1944	7	81,1	30	1	-5,00	57,3
1945	8	97,3	43	1	11,20	59,9
1946	9	72	20	1	-14,10	60,6
1947	10	93,9	41	1	7,80	61,2
1948	11	83,8	33	1	-2,30	62,6
1949	12	122,8	58	1	36,70	63,6
1950	13	87,6	36	1	1,50	64,2
1951	14	101	50	1	14,90	66,8
1952	15	97,8	44	1	11,70	67,2
1953	16	59,9	8	1	-26,20	68,2
1954	17	49,4	3	1	-36,70	68,7
1955	18	57	6	1	-29,10	69,3
1956	19	68,2	16	1	-17,90	71,6
1957	20	83,2	32	1	-2,90	72
1958	21	60,6	9	1	-25,50	72,4
1959	22	50,1	4	1	-36,00	74,8
1960	23	68,7	17	1	-17,40	76,4
1961	24	117,1	56	1	31,00	77,6
1962	25	80,2	28	1	-5,90	78
1963	26	43,6	1	1	-42,50	78,9
1964	27	66,8	14	1	-19,30	79
1965	28	118,4	57	1	32,30	80,2
1966	29	110,4	52	1	24,30	80,9
1967	30	99,1	47	1	13,00	81,1
1968	31	71,6	19	1	-14,50	82,2
			Soma=1004			
1969	32	62,6	11	2	-23,50	83,2
1970	33	61,2	10	2	-24,90	83,8
1971	34	46,8	2	2	-39,30	85,1
1972	35	79	27	2	-7,10	87,4
1973	36	96,3	42	2	10,20	87,6
1974	37	77,6	24	2	-8,50	88,1
1975	38	69,3	18	2	-16,80	89,2
1976	39	67,2	15	2	-18,90	89,8
1977	40	72,4	21	2	-13,70	92,7
1978	41	78	25	2	-8,10	93,9
1979	42	141,8	61	2	55,70	96,3
1980	43	100,7	49	2	14,60	97,3
1981	44	87,4	35	2	1,30	97,8
1982	45	100,2	48	2	14,10	97,9
1983	46	166,9	62	2	80,80	98
1984	47	74,8	22	2	-11,30	99,1
1985	48	133,4	59	2	47,30	100,2
1986	49	85,1	34	2	-1,00	100,7
1987	50	78,9	26	2	-7,10	101
1988	51	76,4	23	2	-9,70	104,3
1989	52	64,2	13	2	-21,90	110,4
1990	53	53,1	5	2	-33,00	110,8
1991	54	112,2	54	2	26,10	112,2
1992	55	110,8	53	2	24,70	114,9
1993	56	82,2	31	2	-3,90	117,1
1994	57	88,1	37	2	2,00	118,4
1995	58	80,9	29	2	-5,20	122,8
1996	59	89,8	39	2	3,70	133,4
1997	60	114,9	55	2	28,80	141,7
1998	61	63,6	12	2	-22,50	141,8
1999	62	57,3	7	2	-28,80	166,9
			Soma=949			

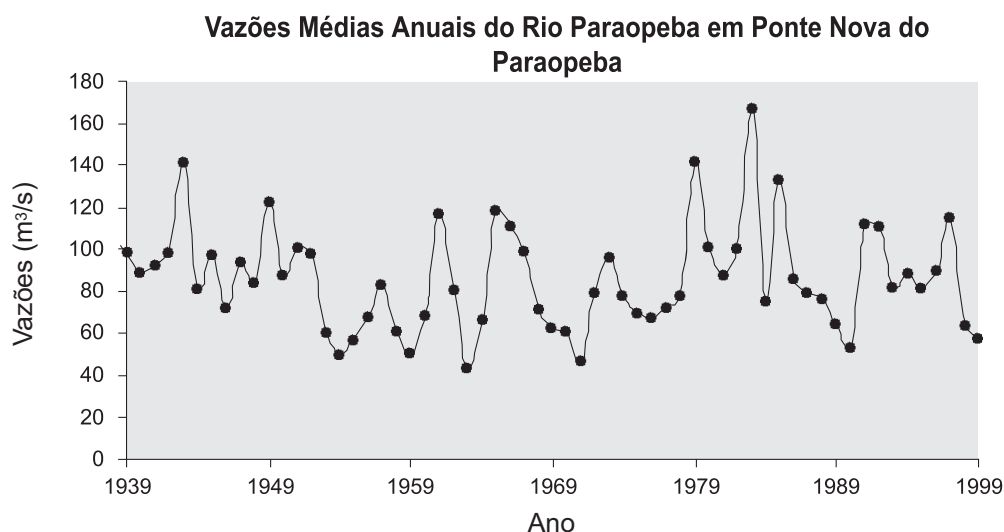


Figura 7.4 - Variação temporal das vazões médias anuais do Rio Paraopeba em Ponte Nova do Paraopeba

7.4 – Alguns Testes de Aderência Usuais em Hidrologia

Nos itens anteriores, foram descritos alguns testes de hipóteses referentes aos parâmetros de uma certa população ou referentes a atributos necessários a uma amostra aleatória simples. Outra classe importante de testes de hipóteses refere-se à *verificação da forma* de uma distribuição de probabilidades. Essa classe é constituída pelos chamados *testes de aderência*, por meio dos quais, verifica-se a eventual adequação entre as probabilidades ou frequências, tal como calculadas por um certo modelo distributivo hipotético, e as correspondentes frequências com que, determinados valores amostrais são observados. Os testes de aderência permitem, por exemplo, verificar se uma variável aleatória discreta segue uma distribuição de Poisson ou se uma variável aleatória contínua é distribuída segundo um modelo de Gumbel.

No contexto das variáveis aleatórias hidrológicas, é muito freqüente a situação em que não se conhece *a priori* a distribuição de probabilidades que descreve a população da qual se extraiu um certo conjunto de observações. Nessas circunstâncias, a seleção das distribuições de probabilidades aptas à modelação de uma determinada variável hidrológica é realizada com base (i) nas características físicas do fenômeno em foco; (ii) em possíveis deduções teóricas quanto às propriedades distributivas da variável em questão; e (iii) na aderência da distribuição

proposta à distribuição empírica dos valores amostrais. No que concerne ao item (i), a dinâmica do mecanismo de formação de cheias, por exemplo, é um fator que indica que as distribuições positivamente assimétricas sejam mais aptas à modelação de vazões máximas anuais, enquanto que a capacidade máxima de dissolução de um gás, em um meio líquido, é outro fato que determina que as distribuições limitadas, à esquerda e à direita, sejam mais adequadas à descrição do comportamento probabilístico das concentrações de oxigênio dissolvido, em um trecho fluvial.

No que se refere ao item (ii), é possível conceber algumas poucas variáveis hidrológicas, tais como as alturas anuais de precipitação em locais de sazonalidade pouco marcada, como decorrentes da aplicação do teorema do limite central às alturas pluviométricas diárias. Entretanto, para a grande maioria das variáveis aleatórias hidrológicas, é patente a inexistência de leis dedutivas teóricas que amparem a escolha do modelo que descreve o seu comportamento probabilístico.

Em relação ao item (iii), embora não se prestem à seleção de uma dentre várias distribuições, os testes de aderência são instrumentos da estatística matemática que auxiliam a tomada de decisão quanto à adequação, ou inadequação, de um certo modelo distributivo a uma dada amostra. Os principais testes de aderência, empregados na hidrologia estatística, são o do Qui-Quadrado, o de Kolmogorov-Smirnov, o de Anderson-Darling e o de Filliben. A descrição e aplicação de tais testes são objetos dos itens que se seguem.

7.4.1 – O Teste de Aderência do Qui-Quadrado (χ^2)

Considere que A_1, A_2, \dots, A_r representem um conjunto de eventos mútua e coletivamente disjuntos, de modo que o espaço amostral seja definido pela união desses eventos. Considere também a hipótese nula $H_0: P(A_i) = p_i$, para $i = 1, 2,$

\dots, r , tal que $\sum_{i=1}^r p_i = 1$. Sob tais condições, suponha que, de um certo número de

experimentos N , as frequências absolutas dos elementos pertencentes aos eventos A_1, A_2, \dots, A_r sejam dadas, respectivamente, pelas quantidades $\rho_1, \rho_2, \dots, \rho_r$. Se a hipótese nula é verdadeira, a distribuição conjunta das variáveis $\rho_1, \rho_2, \dots, \rho_r$ é a multinomial (ver item 4.3.1, do capítulo 4), cuja função massa é dada por

$$P(\rho_1 = O_1, \rho_2 = O_2, \dots, \rho_r = O_r | H_0) = \frac{N!}{O_1! O_2! \dots O_r!} p_1^{O_1} p_2^{O_2} \dots p_r^{O_r} \quad (7.23)$$

onde $\sum_{i=1}^r O_i = N$

Em seguida, considere a seguinte estatística :

$$\chi^2 = \sum_{i=1}^r \frac{(O_i - Np_i)^2}{O_i} = \sum_{i=1}^r \frac{(O_i - E_i)^2}{O_i} \quad (7.24)$$

formada pelas realizações O_i , das variáveis p_i , e pelos seus respectivos valores esperados $E_i = E[p_i]$, os quais, *sob a veracidade da hipótese nula*, são iguais a Np_i . A estatística χ^2 expressa, portanto, a soma das diferenças quadráticas entre as realizações das variáveis aleatórias p_i e suas respectivas médias populacionais.

No item 5.9.1, do capítulo 5, viu-se que a soma das diferenças quadráticas entre N variáveis normais e independentes, e sua média comum μ , possui uma distribuição do χ^2 com $v = N$ graus de liberdade. Embora seja evidente a semelhança entre a definição da variável da distribuição do Qui-Quadrado com a estatística χ^2 , a equação 7.24 contém a soma de r variáveis não necessariamente independentes ou normais. Entretanto, é possível demonstrar que, quando N tende para o infinito, a estatística χ^2 , tal como expressa pela equação 7.24, segue uma distribuição do Qui-Quadrado, com $v = (r-1)$ graus de liberdade. Em outros termos,

$$\lim_{N \rightarrow \infty} P(\chi^2 < x | H_0) = \int_0^x \frac{x^{(r-3)/2} e^{-x/2}}{2^{(r-1)/2} \Gamma[(r-1)/2]} dx \quad (7.25)$$

Para grandes valores de N , pode-se, portanto, empregar esse resultado para testar a hipótese nula H_0 de que as frequências relativas esperadas de p_i sejam dadas por Np_i , com p_i calculadas pela distribuição de probabilidades proposta. Um valor elevado da estatística de teste revela grandes diferenças entre as frequências observadas e esperadas, sendo um indicador da pouca aderência da distribuição especificada, sob H_0 , à amostra.

Observe que a distribuição limite da estatística de teste, dada pela equação 7.25, não depende de p_i , contido em H_0 . De fato, a distribuição limite de χ^2 depende apenas do número de partições r do espaço amostral. Isso faz com que o teste possa ser aplicado para diferentes hipóteses nulas, desde que r seja adequadamente especificado. Na prática, o teste de aderência do χ^2 fornece resultados satisfatórios para $N > 50$ e para $Np_i \geq 5$, com $i = 1, 2, \dots, r$. Se as probabilidades p_i forem calculadas a partir de uma distribuição de k parâmetros, estimados pelas observações amostrais, perde-se k graus de liberdade adicionais. Em outras palavras, o parâmetro v , da distribuição da estatística de teste χ^2 , será $v = (r - k - 1)$. Os exemplos 7.7 e 7.8 ilustram a aplicação do teste de aderência do χ^2 para variáveis aleatórias discretas e contínuas.

Exemplo 7.7 - Considere que uma ETA recebe água bruta de um manancial de superfície, captada por uma tomada d'água simples, instalada em determinada cota. Suponha que a variável aleatória discreta X represente o número anual de dias em que o nível d'água, medido na estação fluviométrica local, é inferior à cota da tomada d'água de projeto. Com base em 50 anos de observações, determinou-se a distribuição empírica das freqüências de X , a qual é dada pela Tabela 7.2. Use o método dos momentos para ajustar uma distribuição de Poisson à variável X , calcule as freqüências esperadas por esse modelo e teste sua aderência aos dados empíricos, a um nível de significância $\alpha = 0,05$.

Tabela 7.2 - Número anual de dias em que o nível d'água é inferior à cota da tomada d'água de projeto

$x_i \rightarrow$	0	1	2	3	4	5	6	7	8	>8
$f(x_i)$	0,0	0,06	0,18	0,2	0,26	0,12	0,09	0,06	0,03	0,0

Solução: A função massa de Poisson é $p_x(x) = \frac{v^x}{x!} e^{-v}$, para $x = 0, 1, \dots$ e $v > 0$, com valor esperado $E[X] = v$. A média amostral pode ser calculada pela ponderação de x por suas freqüências observadas e resulta ser $\bar{x} = 3,86$. Portanto, pelo método dos momentos, a estimativa do parâmetro \hat{v} é 3,86. Os valores E_i , da Tabela 7.3 representam as freqüências esperadas, tal como calculadas pela produto de $N = 50$, pela função massa de Poisson.

Tabela 7.3 – Freqüências observadas e empíricas.

x_i	$O_i = 50 f(x_i)$	$E_i = 50 p_x(x_i)$	$O_i - E_i$	$\frac{(O_i - E_i)^2}{E_i}$
0	0	1,0534	-1,0534	1,0534
1	3	4,0661	-1,0661	0,2795
2	9	7,8476	1,1524	0,1692
3	10	10,0973	-0,0973	0,0009
4	13	9,7439	3,2561	1,0880
5	6	7,5223	-1,5223	0,3080
6	4,5	4,8393	-0,3393	0,0238
7	3	2,6685	0,3315	0,0412
8	1,5	1,2876	0,2124	0,0350
>8	0	0,8740	-0,8740	0,8740
Soma	50	50	—	3,8733

A Tabela 7.3 também mostra os outros elementos necessários para o cálculo da estatística de teste χ^2 , quais sejam, as diferenças simples e quadráticas padronizadas, entre as frequências empíricas e esperadas pelo modelo de Poisson. A soma da última coluna da tabela fornece o valor da estatística de teste $\chi^2 = 3,8733$. O número total de partições do espaço amostral, nesse caso, é $r = 10$. Como foi estimado um parâmetro a partir da amostra, $k = 1$, o que resulta em $v = (r - k - 1) = 8$ graus de liberdade para a distribuição da estatística de teste. Trata-se de *um teste unilateral*, no qual, a região crítica, para $\alpha = 0,05$, é definida por $\chi^2_{0,95, v=8} = 15,5$ (Anexo 6). Como $\chi^2 < \chi^2_{0,95, v=8}$, a decisão é a de não rejeitar a hipótese H_0 de que o comportamento probabilístico da variável aleatória, em questão, possa ser modelado pela distribuição de Poisson. Nesse exemplo, embora $N = 50$, algumas frequências esperadas pelo modelo de Poisson foram inferiores a 5, o que pode vir a comprometer o poder do teste de aderência. Essa situação pode ser resolvida satisfatoriamente pela agregação de algumas partições do espaço amostral; por exemplo, as frequências esperadas para $x = 0$ e $x = 1$ podem ser agrupadas em uma nova partição, correspondente a $x \leq 1$, cuja nova frequência seria 5,1195. Da mesma forma, outras partições poderiam ser agrupadas para constituir a nova classe $x \geq 6$. Evidentemente, esse rearranjo das partições implicaria em novos valores de r e da estatística de teste χ^2 .

Exemplo 7.8 – Considere as vazões médias anuais do Rio Paraopeba em Ponte Nova do Paraopeba, listadas na Tabela 7.1, e faça um teste de aderência da distribuição Normal a esses dados, por meio do teste do χ^2 , a um nível de significância $\alpha = 0,05$.

Solução: No caso de variáveis aleatórias contínuas, as partições do espaço amostral são feitas por meio da divisão em classes, com o cálculo das frequências observadas e esperadas, dentro dos limites dos intervalos de classe. Para a amostra em questão, já foram mostradas anteriormente a tabela de frequências absolutas e o histograma correspondente, com 7 classes de largura fixa. A tabela de frequências absolutas, anteriormente elaborada, é a Tabela 2.3, do capítulo 2. Entretanto, nessa tabela, observa-se uma frequência muito pequena em alguns intervalos, o que força o rearranjo das classes, cujas larguras não necessitam ser fixas. A Tabela 7.4 mostra esse rearranjo e outros elementos auxiliares para a construção da estatística de teste.

Tabela 7.4 – Freqüências observadas e empíricas.

Classe	Intervalos	O_i	E_i	$O_i - E_i$	$\frac{(O_i - E_i)^2}{E_i}$
1	(0,60]	8	9,1468	-1,1468	0,1438
2	(60,70]	10	6,9179	3,0822	1,3732
3	(70,90]	21	18,7621	2,2379	0,2669
4	(90,105]	12	13,2355	-1,2355	0,1153
5	(105,120]	6	8,5117	-2,5117	0,7412
6	(120,200]	5	5,4085	-0,4085	0,0309
Soma		62	61,9824	—	2,6712

Com $r = 6$ classes de largura variável, as freqüências observadas O_i variam em torno de valores aceitáveis. Para o cálculo das freqüências esperadas pela distribuição Normal, é preciso estimar os seus parâmetros μ e σ . A amostra fornece a média $\bar{x} = 86,105$ e o desvio-padrão $s_x = 24,960$, os quais, pelo método dos momentos, resultam nas estimativas $\hat{\mu} = 86,105$ e $\hat{\sigma} = 24,960$. Desse modo, a freqüência relativa esperada na classe i é dada por $p_i = \Phi[(LS - \hat{\mu})/\hat{\sigma}] - \Phi[(LI - \hat{\mu})/\hat{\sigma}]$, onde LS e LI representam, respectivamente, os limites superior e inferior de cada classe, e $\Phi(\cdot)$ denota a FAP da distribuição Normal de parâmetros $\hat{\mu}$ e $\hat{\sigma}$. A freqüência absoluta E_i , da classe i , é dada pelo produto de p_i pelo tamanho da amostra $N = 62$. Em seguida, calcula-se as diferenças simples e quadráticas padronizadas, entre as freqüências empíricas e esperadas pelo modelo Normal. A soma da última coluna da Tabela 7.4 fornece o valor da estatística de teste $\chi^2 = 2,6712$. O número total de partições do espaço amostral, nesse caso, é $r = 6$. Como foram estimados dois parâmetros a partir da amostra, $k = 2$, o que resulta em $v = (r - k - 1) = 3$ graus de liberdade para a distribuição da estatística de teste. Trata-se de *um teste unilateral*, no qual, a região crítica, para $\alpha = 0,05$, é definida por $\chi_{0,95,v=3}^2 = 7,81$ (Anexo 6). Como $\chi^2 < \chi_{0,95,v=3}^2$, a decisão é a de não rejeitar a hipótese H_0 de que o comportamento probabilístico da variável aleatória, em questão, possa ser modelado pela distribuição Normal.

7.4.2 – O Teste de Aderência de Kolmogorov-Smirnov (KS)

O teste de aderência de Kolmogorov-Smirnov (KS) é um teste não paramétrico, cuja estatística de teste tem como base a *diferença máxima* entre as funções de probabilidades acumuladas, empírica e teórica, de variáveis aleatórias contínuas. O teste não é aplicável a variáveis aleatórias discretas.

Considere que X represente uma variável aleatória contínua, de cuja população extraiu-se a amostra $\{X_1, X_2, \dots, X_N\}$. A hipótese nula a ser testada é $H_0: P(X < x) = F_X(x)$, onde $F_X(x)$ é suposta *completamente conhecida*, ou seja, seus parâmetros não são estimados a partir da amostra. Para implementar o teste KS, inicialmente, classifique os elementos da amostra $\{X_1, X_2, \dots, X_N\}$ em ordem crescente, de modo a constituir a sequência $\{x_{(1)}, x_{(2)}, \dots, x_{(m)}, \dots, x_{(N)}\}$, na qual $1 \leq m \leq N$ denota a ordem de classificação. Para cada elemento $x_{(m)}$, a distribuição empírica $F_N(x_{(m)})$ é calculada pela proporção de valores amostrais que não excedem $x_{(m)}$, ou seja,

$$F_N(x_{(m)}) = \frac{m}{N} \quad (7.26)$$

Em seguida, calcule as probabilidades teóricas, segundo $F_X(x)$, tendo como argumento os valores $x_{(m)}$. A estatística do teste KS é dada por

$$D_N = \sup_{-\infty < x < \infty} |F_N(x) - F_X(x)| \quad (7.27)$$

e corresponde, portanto, à maior diferença entre as probabilidades empírica e teórica.

Se H_0 é verdadeira e quando $N \rightarrow \infty$, a estatística D_N irá tender a zero. Por outro lado, se N é um valor finito, a estatística D_N deverá ser da ordem de grandeza de $1/\sqrt{N}$ e, portanto, a quantidade $\sqrt{N}D_N$ não irá tender a zero, mesmo para valores muito elevados de N . Smirnov (1948) determinou a distribuição limite da variável aleatória $\sqrt{N}D_N$, a qual, sob a premissa de veracidade da hipótese H_0 , é expressa por

$$\lim_{N \rightarrow \infty} P(\sqrt{N}D_N \leq z) = \frac{\sqrt{2\pi}}{z} \sum_{k=1}^{\infty} \exp\left[-\frac{(2k-1)^2 \pi^2}{8z^2}\right] \quad (7.28)$$

Portanto, para amostras de tamanho superior a 40, os valores críticos da estatística de teste D_N serão $1,3581/\sqrt{N}$, para o nível de significância $\alpha = 0,05$, e $1,6276/\sqrt{N}$, para $\alpha = 0,01$; esses valores resultam da soma dos cinco primeiros termos da somatória da equação 7.28, e permanecem praticamente inalterados a partir do sexto termo. Para amostras de tamanho inferior a 40, os valores críticos de D_N devem ser obtidos da Tabela 7.5. O exemplo 7.9 ilustra a aplicação do teste de aderência de Kolmogorov-Smirnov.

A construção da estatística do teste KS parte da premissa que $F_X(x)$ é completamente conhecida e, portanto, que seus parâmetros são especificados e, portanto, não são estimados a partir da amostra. Entretanto, quando as estimativas dos parâmetros são obtidas dos elementos da amostra, simulações de Monte

Carlo demonstram que o teste KS é conservador quanto à magnitude do erro do tipo I, podendo ocorrer rejeições indevidas da hipótese nula. Com o objetivo de corrigir tal situação, Crutcher (1975) *apud* Haan (1977), apresenta novas tabelas de valores críticos da estatística $D_{N,\alpha}$ para amostras de tamanhos variáveis, considerando, sob H_0 , as distribuições exponencial, gama, normal e Gumbel.

Tabela 7.5 – Valores críticos da estatística $D_{N,\alpha}$ do teste de aderência KS

N	$D_{N,0,10}$	$D_{N,0,05}$	$D_{N,0,02}$	$D_{N,0,01}$	N	$D_{N,0,10}$	$D_{N,0,05}$	$D_{N,0,02}$	$D_{N,0,01}$
10	0,369	0,409	0,457	0,489	26	0,233	0,259	0,290	0,311
11	0,352	0,391	0,437	0,468	27	0,229	0,254	0,284	0,305
12	0,338	0,375	0,419	0,449	28	0,225	0,250	0,279	0,300
13	0,325	0,361	0,404	0,432	29	0,221	0,246	0,275	0,295
14	0,314	0,349	0,390	0,418	30	0,218	0,242	0,270	0,290
15	0,304	0,338	0,377	0,404	31	0,214	0,238	0,266	0,285
16	0,295	0,327	0,366	0,392	32	0,211	0,234	0,262	0,281
17	0,286	0,318	0,355	0,381	33	0,208	0,231	0,258	0,277
18	0,279	0,309	0,346	0,371	34	0,205	0,227	0,254	0,273
19	0,271	0,301	0,337	0,361	35	0,202	0,224	0,251	0,269
20	0,265	0,294	0,329	0,352	36	0,199	0,221	0,247	0,265
21	0,259	0,287	0,321	0,344	37	0,196	0,218	0,244	0,262
22	0,253	0,281	0,314	0,337	38	0,194	0,215	0,241	0,258
23	0,247	0,275	0,307	0,330	39	0,191	0,213	0,238	0,255
24	0,242	0,269	0,301	0,323	40	0,189	0,210	0,235	0,252
25	0,238	0,264	0,295	0,317	>40	$1,22/\sqrt{N}$	$1,36/\sqrt{N}$	$1,52/\sqrt{N}$	$1,63/\sqrt{N}$

Exemplo 7.9 – Refaça o exemplo 7.8, com o teste de aderência de Kolmogorov-Smirnov.

Solução: A última coluna da Tabela 7.1 apresenta as vazões médias anuais do Rio Paraopeba, em Ponte Nova do Paraopeba, já classificadas em ordem crescente. As freqüências empíricas correspondentes às vazões classificadas podem ser calculadas pela equação 7.26. As freqüências teóricas, correspondentes à distribuição Normal, podem ser calculadas por $F_X(X_i^*) = \Phi[(X_i^* - \mu)/\sigma]$, onde X_i^* representa a vazão classificada, μ e σ denotam os parâmetros populacionais, supostamente iguais a $\bar{X} = 86,105$ e $s_X = 24,960$, respectivamente, e $\Phi(\cdot)$ é a FAP da distribuição Normal. A Figura 7.5 apresenta o gráfico das freqüências empíricas e teóricas, versus as vazões médias anuais classificadas.

No gráfico da Figura 7.5, também está indicada a máxima diferença absoluta entre as freqüências empíricas e teóricas, a qual foi calculada em $D_N^{calc} = 0,08179$. Consultando a Tabela 7.5, para $\alpha = 0,05$ (teste unilateral) e

$N=62$, vê-se que o valor crítico da estatística de teste é $D_{N,0,05} = 1,36 / \sqrt{N} = 1,36 / \sqrt{62} = 0,1727$, o qual define o limite inferior da região de rejeição da hipótese nula. Portanto, como $D_N^{calc} < D_{n,0,05}$, a decisão é a de não rejeitar a hipótese H_0 de que o comportamento probabilístico da variável aleatória, em questão, possa ser modelado pela distribuição Normal.

Diagrama de Frequências Empíricas e Teóricas para o Teste de Aderência KS

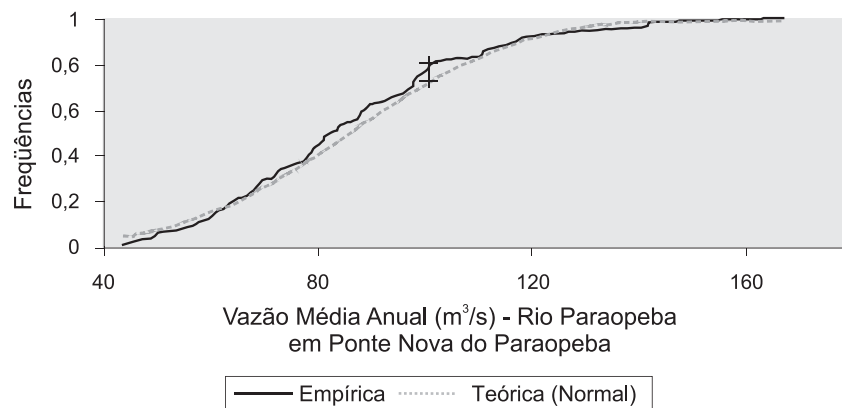


Figura 7.5 – Frequências empíricas e teóricas para o teste de aderência de Kolmogorov-Smirnov

7.4.3 – O Teste de Aderência de Anderson-Darling (AD)

O poder dos testes de aderência do Qui-Quadrado e de Kolmogorov-Smirnov, de discriminar entre hipóteses falsas e verdadeiras, é bastante diminuído nas caudas inferior e superior, tanto em função do reduzido número de observações amostrais, quanto em decorrência dos maiores erros de estimação, nessas partições do espaço amostral. Alternativamente, o teste de aderência de Anderson-Darling é um teste não-paramétrico que procura ponderar mais fortemente as caudas das distribuições, nas quais, as maiores (ou as menores) observações da amostra podem alterar sobremaneira a qualidade do ajuste. O teste de aderência de Anderson-Darling, tal como o de Kolmogorov-Smirnov, baseia-se na diferença entre as funções de probabilidades acumuladas, empírica, $F_N(x)$, e teórica, $F_X(x)$, de variáveis aleatórias contínuas. Entretanto, o teste AD dá mais peso às caudas, por meio da divisão das diferenças entre $F_N(x)$ e $F_X(x)$ por $\sqrt{F_X(x)[1 - F_X(x)]}$. Desse modo, a estatística do teste de Anderson-Darling torna-se

$$A^2 = \int_{-\infty}^{\infty} \frac{[F_N(x) - F_X(x)]^2}{F_X(x)[1 - F_X(x)]} f_X(x) dx \quad (7.29)$$

onde $f_X(x)$ é a função densidade, segundo a hipótese nula. Anderson e Darling (1954) demonstraram que a equação 7.29 é equivalente a

$$A^2 = -N - \sum_{i=1}^N \frac{(2i-1) \{ \ln F_X(x_{(i)}) + \ln [1 - F_X(x_{(N-i+1)})] \}}{N} \quad (7.30)$$

onde $\{x_{(1)}, x_{(2)}, \dots, x_{(m)}, \dots, x_{(N)}\}$ representam as observações ordenadas em modo crescente.

Se a estatística A^2 resulta ser um valor elevado, as distribuições empírica, $F_N(x)$, e teórica, $F_X(x)$, diferem muito entre si e, em consequência, a hipótese nula deve ser rejeitada. A distribuição de probabilidades da estatística do teste AD depende da distribuição de probabilidades hipotética $F_X(x)$. Se a distribuição de probabilidades, sob H_0 , é a Normal ou a Log-Normal, os valores críticos de A^2 são os apresentados na Tabela 7.6.

Tabela 7.6 - Valores críticos da estatística A^2_α do teste de aderência AD, se a distribuição hipotética é Normal ou Log-Normal (Fonte: D'Agostino e Stephens, 1986).

α	0,1	0,05	0,025	0,01
$A^2_{crit,\alpha}$	0,631	0,752	0,873	1,035

Para esse caso, a estatística de teste, calculada pela equação 7.30, deve ser multiplicada pelo fator de correção $(1 + 0,75/N + 2,25/N^2)$.

Se a distribuição de probabilidades, sob H_0 , é a Weibull de dois parâmetros, para mínimos, ou a Gumbel, para máximos, os valores críticos de A^2 são os apresentados na Tabela 7.7.

Tabela 7.7 - Valores críticos da estatística A^2_α do teste de aderência AD, se a distribuição hipotética é Weibull (mínimos, 2p) ou Gumbel (máximos) (Fonte: D'Agostino e Stephens, 1986).

α	0,1	0,05	0,025	0,01
$A^2_{crit,\alpha}$	0,637	0,757	0,877	1,038

Para esse caso, a estatística de teste, calculada pela equação 7.30, deve ser multiplicada pelo fator de correção $(1 + 0,2/\sqrt{N})$. A Tabela 7.7 pode ser usada também para a distribuição exponencial.

Não existem tabulações dos valores críticos da estatística A^2_a , para outras distribuições de probabilidades passíveis de serem incluídas na hipótese H_0 . Para essas, as alternativas são (i) utilizar os outros testes de aderência ou (ii) obter

resultados aproximados e independentes de N , com os valores da Tabela 7.6. O exemplo 7.10 ilustra a aplicação do teste de Anderson-Darling para as vazões médias anuais do Rio Paraopeba em Ponte Nova do Paraopeba.

Exemplo 7.10 - Refaça o exemplo 7.8, com o teste de aderência de Anderson-Darling.

Solução: A Tabela 7.8 apresenta um resumo dos cálculos necessários para a determinação da estatística A^2 . Na Tabela 7.8, a distribuição hipotética $F_X(x)$ é a Normal, calculada por $\Phi[(x - 86,105)/24,960]$. A estatística de teste A^2 , sem correção, é calculada por

$$A^2 = -N - \sum_{i=1}^N S_i / N = -62 - (-3876,63)/62 = 0,5262$$
. Nesse caso, o fator de correção é $(1 + 0,75/N + 2,25/N^2) = 1,0127$. Logo, a estatística de teste, já corrigida, é $A^2 = 0,5329$. Consultando a Tabela 7.6, para $\alpha = 0,05$ (teste unilateral), vê-se que o valor crítico da estatística de teste é $A_{crit,0,05}^2 = 0,752$, o qual define o limite inferior da região de rejeição da hipótese nula. Portanto, como $A^2 < A_{crit,0,05}^2$, a decisão é a de não rejeitar a hipótese H_0 de que o comportamento probabilístico da variável aleatória, em questão, possa ser modelado pela distribuição Normal.

Tabela 7.8 – Cálculo da estatística do teste de aderência AD – Vazões médias anuais em Ponte Nova do Paraopeba

Ano Civil	X_i	i	$x_{(i)}$	$x_{(N-i+1)}$	$w = F_X(x_{(i)})$	$\ln w$	$t = 1 - F_X(x_{(N-i+1)})$	$\ln t$	S_i^*
1938	104,3	1	43,6	166,9	0,0443	-3,1170	0,0006	-7,4118	-10,5288
1939	97,9	2	46,8	141,8	0,0577	-2,8532	0,0128	-4,3561	-21,6279
1940	89,2	3	49,4	141,7	0,0707	-2,6492	0,0130	-4,3458	-34,9750
1941	92,7	4	50,1	133,4	0,0746	-2,5959	0,0291	-3,5385	-42,9406
1942	98	5	53,1	122,8	0,0930	-2,3748	0,0708	-2,6485	-45,2095
1943	141,7	6	57	118,4	0,1218	-2,1054	0,0979	-2,3243	-48,7266
1944	81,1	7	57,3	117,1	0,1242	-2,0855	0,1072	-2,2335	-56,1469
1945	97,3	8	59,9	114,9	0,1469	-1,9181	0,1243	-2,0849	-60,0445
1946	72	9	60,6	112,2	0,1534	-1,8745	0,1479	-1,9112	-64,3572
1947	93,9	10	61,2	110,8	0,1592	-1,8377	0,1612	-1,8249	-69,5882
1948	83,8	11	62,6	110,4	0,1732	-1,7535	0,1652	-1,8007	-74,6371
1949	122,8	12	63,6	104,3	0,1836	-1,6949	0,2330	-1,4567	-72,4854
1950	87,6	13	64,2	101	0,1901	-1,6603	0,2753	-1,2898	-73,7520
1951	101	14	66,8	100,7	0,2196	-1,5158	0,2794	-1,2752	-75,3583
1952	97,8	15	67,2	100,2	0,2244	-1,4943	0,2861	-1,2513	-79,6224
1953	59,9	16	68,2	99,1	0,2366	-1,4415	0,3013	-1,1996	-81,8735
1954	49,4	17	68,7	98	0,2428	-1,4155	0,3168	-1,1494	-84,6409
1955	57	18	69,3	97,9	0,2504	-1,3848	0,3183	-1,1449	-88,5369
1956	68,2	19	71,6	97,8	0,2806	-1,2709	0,3197	-1,1404	-89,2179
1957	83,2	20	72	97,3	0,2860	-1,2518	0,3269	-1,1181	-92,4257
1958	60,6	21	72,4	96,3	0,2915	-1,2328	0,3415	-1,0745	-94,5990
1959	50,1	22	74,8	93,9	0,3253	-1,1230	0,3774	-0,9744	-90,1897
1960	68,7	23	76,4	92,7	0,3487	-1,0535	0,3958	-0,9268	-89,1165
1961	117,1	24	77,6	89,8	0,3666	-1,0034	0,4412	-0,8184	-85,6203
1962	80,2	25	78	89,2	0,3727	-0,9870	0,4507	-0,7970	-87,4178
1963	43,6	26	78,9	88,1	0,3864	-0,9508	0,4681	-0,7590	-87,1999
1964	66,8	27	79	87,6	0,3880	-0,9469	0,4761	-0,7421	-89,5146
1965	118,4	28	80,2	87,4	0,4065	-0,9002	0,4793	-0,7354	-89,9578
1966	110,4	29	80,9	85,1	0,4174	-0,8737	0,5161	-0,6615	-87,5081
1967	99,1	30	81,1	83,8	0,4205	-0,8662	0,5368	-0,6221	-87,8140
1968	71,6	31	82,2	83,2	0,4378	-0,8259	0,5463	-0,6045	-87,2570

Tabela 7.8 – Continuação

Ano Civil	X_i	i	$x_{(i)}$	$x_{(N-i+1)}$	$w = F_X(x_{(i)})$	$\ln w$	$t = 1 - F_X(x_{(N-i+1)})$	$\ln t$	S_i^*
1969	62,6	32	83,2	82,2	0,4537	-0,7904	0,5622	-0,5760	-86,0798
1970	61,2	33	83,8	81,1	0,4632	-0,7696	0,5795	-0,5457	-85,4897
1971	46,8	34	85,1	80,9	0,4839	-0,7258	0,5826	-0,5403	-84,8258
1972	79	35	87,4	80,2	0,5207	-0,6526	0,5935	-0,5217	-81,0272
1973	96,3	36	87,6	79	0,5239	-0,6465	0,6120	-0,4909	-80,7582
1974	77,6	37	88,1	78,9	0,5319	-0,6314	0,6136	-0,4884	-81,7478
1975	69,3	38	89,2	78	0,5493	-0,5990	0,6273	-0,4663	-79,9020
1976	67,2	39	89,8	77,6	0,5588	-0,5819	0,6334	-0,4567	-79,9734
1977	72,4	40	92,7	76,4	0,6042	-0,5039	0,6513	-0,4288	-73,6792
1978	78	41	93,9	74,8	0,6226	-0,4739	0,6747	-0,3935	-70,2554
1979	141,8	42	96,3	72,4	0,6585	-0,4177	0,7085	-0,3446	-63,2723
1980	100,7	43	97,3	72	0,6731	-0,3958	0,7140	-0,3369	-62,2813
1981	87,4	44	97,8	71,6	0,6803	-0,3852	0,7194	-0,3293	-62,1633
1982	100,2	45	97,9	69,3	0,6817	-0,3831	0,7496	-0,2882	-59,7466
1983	166,9	46	98	68,7	0,6832	-0,3810	0,7572	-0,2781	-59,9829
1984	74,8	47	99,1	68,2	0,6987	-0,3585	0,7634	-0,2699	-58,4502
1985	133,4	48	100,2	67,2	0,7139	-0,3371	0,7756	-0,2541	-56,1626
1986	85,1	49	100,7	66,8	0,7206	-0,3276	0,7804	-0,2480	-55,8340
1987	78,9	50	101	64,2	0,7247	-0,3220	0,8099	-0,2108	-52,7537
1988	76,4	51	104,3	63,6	0,7670	-0,2653	0,8164	-0,2029	-47,2844
1989	64,2	52	110,4	62,6	0,8348	-0,1805	0,8268	-0,1902	-38,1828
1990	53,1	53	110,8	61,2	0,8388	-0,1758	0,8408	-0,1734	-36,6677
1991	112,2	54	112,2	60,6	0,8521	-0,1601	0,8466	-0,1666	-34,9479
1992	110,8	55	114,9	59,9	0,8757	-0,1328	0,8531	-0,1589	-31,7866
1993	82,2	56	117,1	57,3	0,8928	-0,1133	0,8758	-0,1327	-27,3070
1994	88,1	57	118,4	57	0,9021	-0,1030	0,8782	-0,1299	-26,3125
1995	80,9	58	122,8	53,1	0,9292	-0,0734	0,9070	-0,0976	-19,6692
1996	89,8	59	133,4	50,1	0,9709	-0,0295	0,9254	-0,0775	-12,5184
1997	114,9	60	141,7	49,4	0,9870	-0,0130	0,9293	-0,0733	-10,2788
1998	63,6	61	141,8	46,8	0,9872	-0,0129	0,9423	-0,0594	-8,7484
1999	57,3	62	166,9	43,6	0,9994	-0,0006	0,9557	-0,0453	-5,6465
Soma	-	-	-	-	-	-	-	-	-3876,63

$$* S_i = (2i - 1) \{ \ln w + \ln t \}$$

7.4.4 – O Teste de Aderência de Filliben

O teste de aderência de Filliben foi introduzido por Filliben (1975), como um teste de verificação da hipótese nula de normalidade. Posteriormente, o teste de Filliben foi adaptado, para contemplar diversas outras distribuições de probabilidades, sob H_0 . Dada uma amostra $\{X_1, X_2, \dots, X_N\}$, de uma variável aleatória X , e posta a hipótese nula de que a amostra foi extraída de uma população cuja distribuição de probabilidades é $F_X(x)$, a estatística do teste de aderência de Filliben é construída com base no coeficiente de correlação linear r , entre as observações ordenadas em modo crescente $\{x_{(1)}, x_{(2)}, \dots, x_{(i)}, \dots, x_{(N)}\}$ e os quantis teóricos $\{w_1, w_2, \dots, w_i, \dots, w_N\}$, os quais são calculados por $w_i = F_X^{-1}(1 - q_i)$, onde q_i representa a probabilidade empírica correspondente à ordem de classificação i . Formalmente, a estatística do teste de Filliben é expressa por

$$r = \frac{\sum_{i=1}^N (x_{(i)} - \bar{x})(w_i - \bar{w})}{\sqrt{\sum_{i=1}^N (x_{(i)} - \bar{x})^2 \sum_{i=1}^N (w_i - \bar{w})^2}} \quad (7.31)$$

onde $\bar{x} = \sum_{i=1}^N x_{(i)} / N$ e $\bar{w} = \sum_{i=1}^N w_i / N$.

A idéia essencial do teste de aderência de Filliben é que a eventual existência de uma forte associação linear entre $x_{(i)}$ e w_i , é um indicador de que as observações podem, de fato, ter sido extraídas de uma população cuja distribuição de probabilidades é $F_X(x)$. Portanto, a hipótese nula é $H_0: r = 1$, a qual deve ser testada contra a hipótese alternativa $H_1: r < 1$, tratando-se de um teste unilateral. Nesse caso, a região de rejeição de H_0 , a um nível de significância α , é formada pelos valores de r inferiores ao valor crítico r_{crit} , dado pela distribuição de probabilidades da estatística de teste. Assim, se $r < r_{\text{crit}}$, a hipótese nula deve ser rejeitada em favor de H_1 .

Na construção da estatística de teste, expressa pela equação 7.31, é implícita a especificação de $F_X(x)$, na forma de $w_i = F_X^{-1}(1 - q_i)$. As probabilidades empíricas q_i , correspondentes às ordens de classificação i , são também denominadas posições de plotagem e variam em conformidade à especificação de $F_X(x)$. Em geral, cada uma das diferentes fórmulas para a posição de plotagem q_i procura obter quantis quase não-enviesados, em relação a cada uma das diferentes distribuições de probabilidade $F_X(x)$. A maioria dessas fórmulas pode ser expressa pela seguinte expressão geral:

$$q_i = \frac{i - a}{N + 1 - 2a} \quad (7.32)$$

onde a varia conforme a especificação de $F_X(x)$. A Tabela 7.9 apresenta um sumário das diferentes fórmulas para a posição de plotagem, indicando também os valores de a correspondentes, bem como as principais motivações de sua proposição, em conformidade com a especificação de $F_X(x)$.

Tabela 7.9 – Fórmulas para o cálculo da posição de plotagem q_i

Denominação	Fórmula	a	Motivação
Weibull	$q_i = \frac{i}{N + 1}$	0	Probabilidades de superação não-enviesadas para todas as distribuições.
Blom	$q_i = \frac{i - 3/8}{N + 1/4}$	0,375	Quantis não-enviesados para a distribuição Normal.
Cunnane	$q_i = \frac{i - 0,40}{N + 0,2}$	0,40	Quantis aproximadamente não-enviesados para quase todas as distribuições.
Gringorten	$q_i = \frac{i - 0,44}{N + 0,12}$	0,44	Otimizada para a distribuição de Gumbel.

Fonte: adaptada de tabela original de Stedinger et al. (1993).

Uma vez que os quantis w_i variam conforme $F_X(x)$, é evidente que a distribuição de probabilidades da estatística do teste também irá variar, de acordo com a especificação da distribuição $F_X(x)$, sob a hipótese H_0 . A Tabela 7.10 apresenta os valores críticos $r_{\text{crit}, \alpha}$ para o caso em que $F_X(x)$ é especificada como a distribuição Normal, com as probabilidades empíricas q_i calculadas pela fórmula de Blom. Os valores da Tabela 7.10 permanecem válidos para os logaritmos de uma variável Log-Normal.

Tabela 7.10 – Valores críticos $r_{\text{crit}, \alpha}$ para a distribuição Normal, com $\alpha = 0,375$ na equação 7.32.			
N	$\alpha = 0,10$	$\alpha = 0,05$	$\alpha = 0,01$
10	0,9347	0,9180	0,8804
15	0,9506	0,9383	0,9110
20	0,9600	0,9503	0,9290
30	0,9707	0,9639	0,9490
40	0,9767	0,9715	0,9597
50	0,9807	0,9764	0,9664
60	0,9835	0,9799	0,9710
75	0,9865	0,9835	0,9757
100	0,9893	0,9870	0,9812

Fonte: adaptada de tabela original de Stedinger et al. (1993).

A Tabela 7.11 apresenta os valores críticos $r_{\text{crit}, \alpha}$ para o caso em que $F_X(x)$ é especificada como a distribuição de Gumbel, para máximos, com as probabilidades empíricas q_i calculadas pela fórmula de Gringorten. Os valores da Tabela 7.11 permanecem válidos para o caso em que $F_X(x)$ é especificada como a distribuição de Weibull de 2 parâmetros.

Tabela 7.11 – Valores críticos $r_{\text{crit}, \alpha}$ para a distribuição Gumbel, com $\alpha = 0,44$ na equação 7.32.			
N	$\alpha = 0,10$	$\alpha = 0,05$	$\alpha = 0,01$
10	0,9260	0,9084	0,8630
20	0,9517	0,9390	0,9060
30	0,9622	0,9526	0,9191
40	0,9689	0,9594	0,9286
50	0,9729	0,9646	0,9389
60	0,9760	0,9685	0,9467
70	0,9787	0,9720	0,9506
80	0,9804	0,9747	0,9525
100	0,9831	0,9779	0,9596

Fonte: adaptada de tabela original de Stedinger et al. (1993).

A Tabela 7.12 apresenta os valores críticos $r_{crit,\alpha}$ para o caso em que $F_X(x)$ é especificada como a distribuição Generalizada de Valores Extremos - GEV, com as probabilidades empíricas q_i calculadas pela fórmula de Cunnane. Os valores críticos da Tabela 7.12 foram obtidos por Chowdhury et al. (1991), mediante simulações de amostras de diferentes tamanhos, extraídas da população de uma variável aleatória GEV, com parâmetro de forma especificado por κ .

Tabela 7.12 – Valores críticos $r_{crit,\alpha}$ para a distribuição GEV, com $\alpha=0,40$ na equação 7.32							
α	N	$\kappa=-0,30$	$\kappa=-0,20$	$\kappa=-0,10$	$\kappa=0$	$\kappa=0,10$	$\kappa=0,20$
0,01	5	0,777	0,791	0,805	0,817	0,823	0,825
0,01	10	0,836	0,845	0,856	0,866	0,876	0,882
0,01	20	0,839	0,855	0,878	0,903	0,923	0,932
0,01	30	0,834	0,858	0,89	0,92	0,942	0,953
0,01	50	0,825	0,859	0,902	0,939	0,961	0,97
0,01	100	0,815	0,866	0,92	0,959	0,978	0,985
0,05	5	0,853	0,863	0,869	0,874	0,877	0,88
0,05	10	0,881	0,89	0,9	0,909	0,916	0,92
0,05	20	0,898	0,912	0,926	0,938	0,948	0,953
0,05	30	0,903	0,92	0,937	0,952	0,961	0,967
0,05	50	0,908	0,929	0,95	0,965	0,974	0,979
0,05	100	0,914	0,94	0,963	0,978	0,985	0,989
0,10	5	0,888	0,892	0,896	0,899	0,901	0,903
0,10	10	0,904	0,912	0,92	0,927	0,932	0,936
0,10	20	0,92	0,932	0,943	0,952	0,958	0,962
0,10	30	0,928	0,941	0,953	0,962	0,969	0,973
0,10	50	0,935	0,95	0,963	0,973	0,979	0,982
0,10	100	0,944	0,961	0,974	0,983	0,988	0,991

Vogel e McMartin (1991) empregaram simulações de Monte Carlo para encontrar os valores críticos $r_{crit,\alpha}$ válidos para variáveis aleatórias distribuídas segundo o modelo Pearson Tipo III. De acordo com esses autores, o valor crítico da estatística do teste de Filliben, a um nível de significância $\alpha = 0,05$, pode ser aproximado pela seguinte expressão:

$$r_{crit,\alpha=0,05} \cong \exp(3,77 - 0,0290\gamma^2 - 0,000670N)N^{0,105\gamma-0,758} \quad (7.33)$$

onde γ denota o coeficiente de assimetria populacional da distribuição de Pearson Tipo III, com posição de plotagem calculada pela fórmula de Blom ($\alpha=0,375$). A equação 7.33 é válida para $|\gamma| \leq 5$, podendo ser empregada também para os logaritmos de variáveis aleatórias distribuídas segundo um modelo Log-Pearson Tipo III.

Exemplo 7.11 - Refaça o exemplo 7.8, com o teste de aderência de Filliben.
Solução: A Tabela 7.8 apresenta os quantis observados $x_{(i)}$, já ordenados em modo crescente. Os quantis teóricos da distribuição Normal, de média 86,105 e desvio-padrão 24,960, devem ser calculados pela função inversa $\Phi^{-1}(q_i)$, onde q_i denota a probabilidade empírica, tal como calculada pela fórmula de Blom, para a ordem de classificação i . Para exemplificar esse cálculo, considere que $i = 1$, resultando, portanto, em $q_1 = 0,01004$, para $N = 62$ e $a = 0,375$ na equação 7.32. A inversa $\Phi^{-1}(q_i)$ pode ser facilmente calculada pela função estatística INV.NORM do software Microsoft Excel, com argumentos q_i , μ e σ ; para o exemplo $q_1 = 0,01004$, com $\mu = 86,105$ e $\sigma = 24,960$, a função INV.NORM retorna o valor $w_1 = 28,0769$. Esse cálculo deve ser efetuado para todas as ordens de classificação até $i = N = 62$. A Figura 7.6 apresenta o gráfico entre os quantis teóricos w_i e os observados $x_{(i)}$, assim como a linha correspondente à associação linear entre ambos.

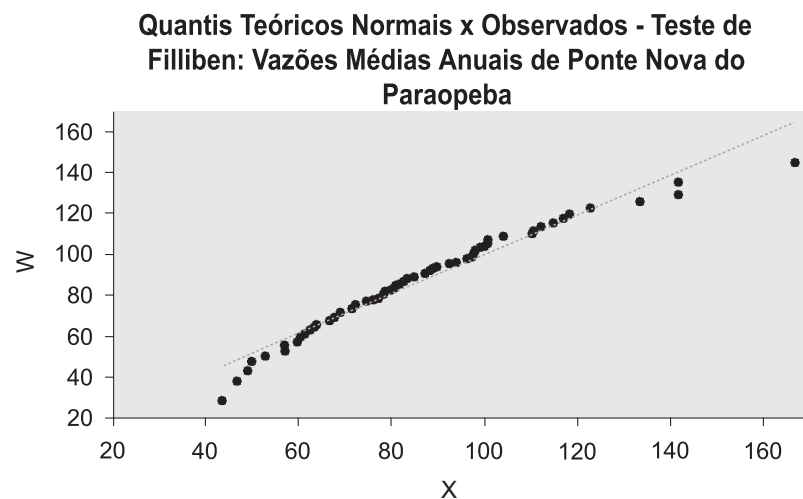


Figura 7.6 – Associação entre os quantis teóricos Normais e os observados no Rio Paraopeba em Ponte Nova do Paraopeba

Em seguida aos cálculos mencionados anteriormente, determina-se a estatística do teste de Filliben pela aplicação da equação 7.31, cujo resultado é $r = 0,9798$. Consultando a Tabela 7.10, para $\alpha = 0,05$ (teste unilateral), e usando interpolação linear entre os valores de N iguais a 60 e 75, vê-se que o valor crítico da estatística de teste é $r_{\text{crit},0,05} = 0,9803$, o qual define o limite superior da região de rejeição da hipótese nula. Portanto, como $r < r_{\text{crit},0,05}$, a decisão é a de rejeitar a hipótese H_0 de que o comportamento probabilístico da variável aleatória, em questão, possa ser modelado pela distribuição Normal.

7.4.5 – Comentários a Respeito dos Testes de Aderência

Em geral, os testes de aderência são deficientes em discernir as diferenças entre as frequências teóricas e empíricas (e/ou quantis teóricos e empíricos), nas caudas inferior e superior das distribuições em análise. No contexto da análise de frequência de variáveis hidrológicas, essa limitação dos testes de aderência é crítica, uma vez que as amostras são de tamanho relativamente pequeno e que, em geral, o interesse é o de inferir sobre o comportamento da variável aleatória, justamente, nas caudas de sua distribuição de probabilidades. Por exemplo, o teste do χ^2 , quando aplicado a variáveis aleatórias contínuas, está sujeito à prescrição de classes, cujo número e amplitude podem interferir profundamente na decisão do teste. No caso do teste de Kolmogorov-Smirnov, a mera observação de sua tabela de valores críticos (Tabela 7.5), revela o conservadorismo do teste no que se refere à decisão de rejeição da hipótese nula. O teste de Anderson-Darling, apesar de constituir uma interessante alternativa aos testes KS e do χ^2 , apresenta a limitação de que a distribuição de sua estatística de teste é conhecida apenas para algumas distribuições hipotéticas $F_X(x)$. O teste de Filliben, como alternativa restante, apresenta, como principais vantagens, a simplicidade de construção de sua estatística de teste e algumas comparações favoráveis de seu poder de teste, em relação aos demais, tais como aquelas apontadas por Chowdhury et al. (1991) e Vogel e McMartin (1991). Entretanto, esses mesmos autores demonstram o baixo poder do teste de Filliben, quando se trata de análise de frequência local, com base em amostras de tamanho relativamente pequeno.

Os testes de aderência, como quaisquer testes de hipóteses, têm o objetivo de verificar se há uma diferença estatisticamente significativa entre as observações e as supostas realizações, caso essas proviessem de uma população hipotética incluída em H_0 . Portanto, a eventual decisão de não rejeitar a hipótese nula, a um nível de significância previamente estabelecido, não implica em estabelecer a idéia de que os dados foram, de fato, amostrados a partir da população hipotética. Essa é, por princípio, desconhecida e pode ser uma, entre tantas outras populações incluídas na hipótese alternativa H_1 . Por outro lado, as estatísticas dos testes de aderência têm distribuições de probabilidades e, portanto, valores críticos que dependem da distribuição $F_X(x)$, sob H_0 , assim como, implicitamente, de suas estimativas paramétricas e dos respectivos erros de estimativa. Com essas considerações em mente, é possível concluir que os resultados de diferentes testes de aderência *não são comparáveis entre si* e, portanto, *não se prestam à seleção do modelo distributivo mais adequado* para uma certa amostra de observações.

7.5 – Teste para Detecção e Identificação de Pontos Atípicos (*outliers*)

Em uma certa amostra de observações, um elemento ou ponto amostral é considerado atípico, ou um *outlier*, do ponto de vista estatístico, quando ele se desvia *significativamente* do conjunto dos outros pontos. Esse desvio pode ter origem em erros de medição ou de processamento, mas também pode ser o produto de causas naturais indeterminadas. Em qualquer caso, a presença de pontos atípicos em uma dada amostra, pode afetar drasticamente o ajuste de uma certa distribuição de probabilidades àqueles dados. No item 2.1.4 do capítulo 2, foi descrito um procedimento de identificação de pontos atípicos, por meio dos quartis amostrais e da amplitude inter-quartis. Este procedimento, embora bastante útil, é meramente exploratório e não constitui, do ponto de vista estatístico, um teste de hipótese, com um nível de significância previamente estabelecido.

Entre os diversos testes de hipóteses para detecção e identificação de pontos atípicos, o teste de Grubbs e Beck, descrito por Grubbs (1950, 1969) e estendido por Grubbs e Beck (1972), encontra-se entre os mais freqüentemente empregados. De acordo com esse teste, as quantidades x_S e x_I definem, respectivamente, os limites superior e inferior, acima e abaixo dos quais, os pontos atípicos, eventualmente presentes em uma amostra, são detectados e identificados. Essas quantidades são definidas pelas seguintes expressões:

$$x_S = \exp(\bar{x} + k_{N,\alpha} s_X) \quad (7.34)$$

e

$$x_I = \exp(\bar{x} - k_{N,\alpha} s_X) \quad (7.35)$$

onde \bar{x} e s_X representam, respectivamente, a média aritmética e o desvio-padrão de uma amostra de tamanho N , de uma variável aleatória X , e $k_{N,\alpha}$ denota o valor crítico da estatística de Grubbs e Beck, para um nível de significância α . Para $100\alpha = 10\%$, Pilon et al. (1985) propõem a seguinte aproximação para o valor crítico da estatística de Grubbs e Beck:

$$k_{N,\alpha=0,10} = -3,62201 + 6,28446 N^{1/4} - 2,49835 N^{1/2} + 0,491436 N^{3/4} - 0,037911 N \quad (7.36)$$

De acordo com o teste de Grubbs e Beck, a um nível $\alpha = 0,10$ e $k_{N,0,10}$ dado pela equação 7.36, as observações eventualmente superiores a x_S , e/ou inferiores a x_I , estariam se desviando significativamente do conjunto dos dados e deveriam ser consideradas como *outliers*.

Uma vez detectados e identificados os pontos atípicos presentes em uma amostra, a decisão de mantê-los ou expurgá-los da análise de frequência é matéria de investigação suplementar. Se o exame detalhado de uma certa observação atípica for conclusivo, quanto a caracterizá-la como uma medição incorreta ou sujeita a erros de processamento, ela deve ser certamente expurgada da análise. Entretanto, se a observação atípica resultar de causas naturais, tais como a manifestação de fenômenos extraordinários e diferenciados, em relação ao conjunto dos outros pontos amostrais, a melhor decisão é certamente a de manter os *outliers* na análise de frequência, buscando encontrar o modelo probabilístico, ou os modelos probabilísticos, que melhor descrevam aquele comportamento observado.

Exercícios

- 1) Considere o teste da hipótese nula $H_0: p = 0,5$, contra $H_1: p > 0,5$, onde p representa a probabilidade de sucesso em 18 tentativas independentes de um processo de Bernoulli. A decisão é arbitrada como a de rejeitar a hipótese nula, caso a variável aleatória discreta Y , dada pelo número de sucessos em 18 tentativas, seja maior ou igual a 13. Calcule a função poder do teste, denotada por $[1-\beta(p)]$, e ilustre-a graficamente, para diferentes valores de $p > 0,5$.
- 2) Repita o exercício 1, para a hipótese alternativa $H_1: p \neq 0,5$.
- 3) Considere as vazões médias anuais do Rio Paraopeba em Ponte Nova do Paraopeba, listadas na Tabela 7.1. Suponha que essa amostra tenha sido extraída de uma população Normal, de desvio-padrão populacional conhecido e igual a $\sigma = 24,960 \text{ m}^3/\text{s}$. Teste a hipótese $H_0: \mu_1 = 85 \text{ m}^3/\text{s}$, contra a alternativa $H_1: \mu_2 = 90 \text{ m}^3/\text{s}$, para $\alpha = 0,05$.
- 4) Repita o exercício 3, supondo que, desta feita, o desvio-padrão populacional não é conhecido.
- 5) Refaça o exercício 3 para a hipótese alternativa $H_1: \mu_1 \neq 85 \text{ m}^3/\text{s}$.
- 6) Repita o exercício 5, supondo que, desta feita, o desvio-padrão populacional não é conhecido.
- 7) Considerando, novamente, as vazões médias anuais do Rio Paraopeba em Ponte Nova do Paraopeba, listadas na Tabela 7.1, separe-as em duas amostras de igual tamanho, uma para o período de 1938 a 1968, e a outra para o período de 1969 a 1999. Supondo tratarem-se de variáveis normais, teste a hipótese de

que, considerados os períodos de 1938-1968 e de 1969-1999, as médias populacionais correspondentes não diferem entre si, em *mais* de $5 \text{ m}^3/\text{s}$, para $\alpha = 0,05$.

8) De volta às vazões médias anuais do Rio Paraopeba em Ponte Nova do Paraopeba, listadas na Tabela 7.1, suponha que essa amostra tenha sido extraída de uma população Normal, de média populacional conhecida e igual a $\mu = 86,105 \text{ m}^3/\text{s}$. Teste a hipótese $H_0: \sigma_1 = 25 \text{ m}^3/\text{s}$, contra a alternativa $H_1: \sigma_1 < 25 \text{ m}^3/\text{s}$, para $\alpha = 0,05$.

9) Repita o exercício 8, supondo que, desta feita, a média populacional não é conhecida.

10) Considerando, novamente, as vazões médias anuais do Rio Paraopeba em Ponte Nova do Paraopeba, listadas na Tabela 7.1, separe-as em duas amostras de igual tamanho, uma para o período de 1938 a 1968, e a outra para o período de 1969 a 1999. Supondo tratem-se de variáveis normais, teste a hipótese de que, considerados os períodos de 1938-1968 e de 1969-1999, as variâncias populacionais correspondentes não diferem entre si, para $\alpha = 0,05$.

11) Repita o exercício 10, considerando que a hipótese nula, desta feita, é a de que a variância do período de 1938 a 1968, é 10% maior do que a correspondente ao período de 1969-1999.

12) Considere a amostra de alturas diárias de precipitação máxima anual da estação pluviométrica de Ponte Nova do Paraopeba, listadas no Anexo 3. Teste a hipótese nula de que as observações são aleatórias, para $\alpha = 0,05$.

13) Com os dados do exercício 12, teste a hipótese nula de que as observações são independentes, para $\alpha = 0,05$.

14) Com os dados do exercício 12, teste a hipótese nula de que as observações são homogêneas, para $\alpha = 0,05$.

15) Com os dados do exercício 12, teste a hipótese nula de que as observações são estacionárias, para $\alpha = 0,05$.

16) Fez-se a contagem de *E. Coli* em 10 amostras de água. As contagens positivas, expressas em centenas de organismos por 100 ml de água ($10^2/100\text{ml}$), são 17, 21, 25, 23, 17, 26, 24, 19, 21 e 17, com média e a variância amostrais iguais a 21 e 10,6 respectivamente. Suponha que N represente o número total dos diferentes

organismos presentes em cada amostra e que p represente a fração correspondente ao organismo *E. Coli*. Ajuste uma distribuição Binomial à variável Y =centenas de organismos *E. Coli* por 100 ml de água. Verifique a aderência da distribuição Binomial aos dados amostrais, por meio do teste do χ^2 , a um nível de significância $\alpha = 0,10$.

17) Para os dados do exercício 12, teste a aderência da distribuição Log-Normal, de 2 parâmetros, por meio do teste do χ^2 , a um nível de significância $\alpha = 0,05$.

18) Para os dados do exercício 12, teste a aderência da distribuição Gumbel (máximos), por meio do teste do χ^2 , a um nível de significância $\alpha = 0,05$.

19) Para os dados do exercício 12, teste a aderência da distribuição GEV, por meio do teste do χ^2 , a um nível de significância $\alpha = 0,05$.

20) Para os dados do exercício 12, teste a aderência da distribuição Exponencial, por meio do teste do χ^2 , a um nível de significância $\alpha = 0,05$.

21) Para os dados do exercício 12, teste a aderência da distribuição Pearson Tipo III, por meio do teste do χ^2 , a um nível de significância $\alpha = 0,05$.

22) Para os dados do exercício 12, teste a aderência da distribuição Log-Pearson Tipo III, por meio do teste do χ^2 , a um nível de significância $\alpha = 0,05$.

23) Para os dados do exercício 12, teste a aderência da distribuição Log-Normal, de 2 parâmetros, por meio do teste de Kolmogorov-Smirnov, a um nível de significância $\alpha = 0,05$.

24) Para os dados do exercício 12, teste a aderência da distribuição Gumbel (máximos), por meio do teste de Kolmogorov-Smirnov, a um nível de significância $\alpha = 0,05$.

25) Para os dados do exercício 12, teste a aderência da distribuição GEV, por meio do teste de Kolmogorov-Smirnov, a um nível de significância $\alpha = 0,05$.

26) Para os dados do exercício 12, teste a aderência da distribuição Exponencial, por meio do teste de Kolmogorov-Smirnov, a um nível de significância $\alpha = 0,05$.

27) Para os dados do exercício 12, teste a aderência da distribuição Pearson Tipo III, por meio do teste de Kolmogorov-Smirnov, a um nível de significância $\alpha = 0,05$.

- 28) Para os dados do exercício 12, teste a aderência da distribuição Log-Pearson Tipo III, por meio do teste de Kolmogorov-Smirnov, a um nível de significância $\alpha = 0,05$.
- 29) Para os dados do exercício 12, teste a aderência da distribuição Log-Normal, de 2 parâmetros, por meio do teste de Anderson-Darling, a um nível de significância $\alpha = 0,05$.
- 30) Para os dados do exercício 12, teste a aderência da distribuição Gumbel (máximos), por meio do teste de Anderson-Darling, a um nível de significância $\alpha = 0,05$.
- 31) Para os dados do exercício 12, teste a aderência da distribuição Exponencial, por meio do teste de Anderson-Darling, a um nível de significância $\alpha = 0,05$.
- 32) Para os dados do exercício 12, teste a aderência da distribuição Log-Normal, de 2 parâmetros, por meio do teste de Filliben, a um nível de significância $\alpha = 0,05$.
- 33) Para os dados do exercício 12, teste a aderência da distribuição Gumbel (máximos), por meio do teste de Filliben, a um nível de significância $\alpha = 0,05$.
- 34) Para os dados do exercício 12, teste a aderência da distribuição GEV, por meio do teste de Filliben, a um nível de significância $\alpha = 0,05$.
- 35) Para os dados do exercício 12, teste a aderência da distribuição Pearson Tipo III, por meio do teste de Filliben, a um nível de significância $\alpha = 0,05$.
- 36) Para os dados do exercício 12, teste a aderência da distribuição Log-Pearson Tipo III, por meio do teste de Filliben, a um nível de significância $\alpha = 0,05$.
- 37) Para os dados do exercício 12, use o teste de Grubbs e Beck, com $\alpha = 0,10$, para detectar e identificar a presença de pontos atípicos. Compare os resultados com aqueles encontrados por meio do critério da amplitude inter-quartis. Lembre-se que, segundo tal critério, é considerado um ponto atípico superior todo elemento da amostra superior a $(Q_3 + 1,5AIQ)$ e, analogamente, um ponto atípico inferior é todo e qualquer elemento menor do que $(Q_1 - 1,5AIQ)$, onde Q_1 e Q_3 representam, respectivamente, o primeiro e o terceiro quartis, e $AIQ = Q_3 - Q_1$.

