

Licenciatura em Engenharia Informática

Métodos Estatísticos

Exercícios

Testes de Hipóteses Paramétricos

Departamento de Matemática



6 Testes de Hipóteses Paramétricos

Exercício 6.1 Uma firma tem seguido a política de oferecer uma garantia de 2000 utilizações para determinado aparelho que comercializa. Este procedimento baseia-se em estudos levados a cabo no período inicial de produção, que indicavam um número médio de utilizações possíveis por aparelho de 2060, com uma variabilidade traduzida por $\sigma = 20$. Existindo indícios de que presentemente a situação pode ter mudado, pretende-se averiguar se continua a ser 2060 o número médio de utilizações por aparelho. Para o efeito foram selecionados ao acaso e testados pela firma 10 aparelhos, os quais forneceram os seguintes valores:

2100 2025 2071 2067 2150 2115 2064 2088 1995 2095

Suponha que o número de utilizações permitidas por aparelho comporta-se de forma aproximadamente normal.

1. Como define o teste de hipóteses a efetuar?
2. Proceda ao cálculo da região crítica para o teste definido anteriormente, considerando $\alpha = 0.05$.
3. Face à amostra recolhida, que decisão tomaria quanto a H_0 ? Porquê?

Exercício 6.2 Um fabricante de cabos indicava que os seus cabos apresentavam uma tensão média de rutura de 1800 toneladas, com um desvio padrão de 100 toneladas. Um outro fabricante dizia que tal valor não correspondia à verdade, indicando como média de rutura de 1750 toneladas, com o mesmo desvio padrão. Para decidir qual dos dois dizia a verdade foi recolhida uma amostra de 35 cabos, verificando-se que apresentavam uma tensão média de rutura de 1650 toneladas. Admita que a tensão de rutura segue uma distribuição aproximadamente normal.

1. Realize o teste de hipóteses: $H_0 : \mu = 1800$ contra $H_1 : \mu = 1750$, e diga qual dos dois fabricantes terá razão para um nível de significância 5% e recorrendo apenas à região crítica.
2. Realize o teste de hipóteses: $H_0 : \mu = 1750$ contra $H_1 : \mu = 1800$, e diga qual dos dois fabricantes terá razão para um nível de significância 10% e recorrendo apenas ao valor- p .

Exercício 6.3 Um comerciante recebe ovos de um determinado aviário, onde os ovos são classificados consoante o peso, em duas classes A e B . O peso dos ovos de classe A tem distribuição $N(50, 8)$ e o peso dos ovos de classe B tem distribuição $N(55, 8)$. Um comerciante recebe uma remessa de ovos com garantia de serem de classe B e tem um prazo de dois dias para reclamar, caso considere ter havido engano da parte do aviário. Para tomar uma decisão ele analisou 10 ovos cujo peso total foi de 530 gramas. Recorrendo a um teste de hipóteses para uma média e um nível de significância de 0.05, diga qual a atitude que o comerciante deve tomar.

Exercício 6.4 A especificação de produção de determinada liga, com comportamento normal, exige 23.2% de cobre. Uma amostra de 10 análises da liga acusou um conteúdo médio de 23.5% de cobre com um desvio padrão de 0.24%.

1. Considerando $\alpha = 0.01$, pode-se concluir que o produto satisfaz as especificações?
2. Através do cálculo do valor $-p$ e considerando $\alpha = 0.01$, teste a hipótese de que o conteúdo médio de cobre seja superior ao exigido pelas especificações.

Exercício 6.5 Da produção diária de um dado fertilizante colheram-se nove amostras que se analisaram para calcular a percentagem média de azoto. Os resultados obtidos foram os seguintes:

6.2; 5.7; 5.8; 5.8; 6.1; 5.9; 6.0; 5.7; 5.9

1. Sabendo que o processo de análise utilizado fornece valores com distribuição Normal, $\mathcal{N}(\mu, 0.24)$, determine o menor nível de significância que assegura, com base nestas observações que a percentagem média de azoto no fertilizante é inferior a 6%.
2. Sabendo que o processo de análise utilizado fornece valores com distribuição Normal, determine o menor nível de significância que assegura, com base nestas observações que a percentagem média de azoto no fertilizante é inferior a 6%.

Exercício 6.6 Com o objetivo de acelerar o tempo que um analgésico leva para penetrar na corrente sanguínea, um químico analista acrescentou certo componente à fórmula original, que acusava um tempo de médio de 43 minutos. Em 36 observações com a nova fórmula, obteve-se um tempo médio de 42 minutos e um desvio padrão de 6 minutos. O que podemos concluir, ao nível de 5% de significância, sobre a eficiência do novo componente? Qual seria a resposta ao nível de 10%?

Exercício 6.7 A dona de uma boutique pode escolher entre dois fornecedores para a aquisição de calças para senhora. Aparentemente, a única diferença reside no preço. A dona da boutique compraria ao primeiro fornecedor (que tem preços mais baixos) a não ser que haja razões para crer que o produto é de qualidade inferior (tenha, em média, menor resistência). Recolhidas duas amostras, obtiveram-se os seguintes resultados:

Fornecedor 1: $n_1 = 12$ $\bar{x}_1 = 88$ $s_1^2 = 9$

Fornecedor 2: $n_2 = 9$ $\bar{x}_2 = 92$ $s_2^2 = 11$

1. Admitindo que a característica em estudo (resistência média) segue distribuição normal (com variâncias iguais), diga, justificando, que decisão deve tomar a dona da boutique (considere $\alpha = 0.05$).
2. Comente quais as consequências para o teste anterior de ser utilizado um grau de significância de 0.01.
3. Suponha que conhece as variâncias das populações, sendo $\sigma_1^2 = 8$ e $\sigma_2^2 = 9$. Elabore novamente um teste de hipóteses, indicando que decisão deve tomar a dona da boutique (considere $\alpha = 0.10$).

Exercício 6.8 Numa experiência, dois grupos de ratos fêmeas foram alimentados com dietas apresentando alto e baixo conteúdo de proteína. A tabela seguinte fornece, para cada rato, o ganho de peso, em gramas, entre o 28.^o e o 84.^o dia de vida:

Conteúdo de proteína	Ganho de peso											
Alto	123	134	146	104	119	124	161	107	83	113	97	129
Baixo	70	118	101	85	107	132	94					

Suponha que o ganho de peso, com alto ou baixo conteúdo de proteína, segue uma distribuição normal com a mesma variância.

1. Ao nível de significância de 1%, há evidência estatística de que a dieta com alto conteúdo de proteína aumenta o ganho de peso?
2. Calcule o valor $-p$ associado ao teste realizado na alínea 1.

Exercício 6.9 Cinco operadores de máquinas são treinados em duas máquinas de diferentes fabricantes, para verificar qual delas apresentava maior facilidade de aprendizagem. Mediu-se o tempo que cada um dos operadores gastou na realização de uma mesma tarefa com cada um dos dois tipos de máquinas. Os resultados estão na tabela seguinte:

Operador	1	2	3	4	5
Fabricante 1	80	72	65	78	85
Fabricante 2	75	70	60	72	78

Suponha que a variável em estudo segue uma distribuição normal.

1. Estas amostras podem ser consideradas independentes?
2. Ao nível de significância de 10% é possível afirmar que a tarefa realizada na máquina do fabricante 1 demora, em média, mais do que na máquina do fabricante 2?

Exercício 6.10 Um laboratório possui dois equipamentos de precisão. O diretor do laboratório suspeita que existe uma pequena diferença de calibração entre os dois (ele não sabe em qual deles) de modo que um tende a dar leituras um pouco maiores do que o outro. Ele propõe testar os dois aparelhos através da leitura de 10 medidas em cada um dos aparelhos:

Medida	1	2	3	4	5	6	7	8	9	10
Aparelho A	12.2	12.1	10.55	13.33	11.42	10.3	12.32	13.27	11.93	12.5
Aparelho B	12.5	12.2	10.57	13.32	11.47	10.3	12.36	13.29	11.91	12.61

Suponha que a variável em estudo segue uma distribuição normal.

1. Estas amostras podem ser consideradas independentes?
2. Ao nível de significância de 1% acha que a suspeita do diretor faz sentido?

Exercício 6.11 Num exame de estatística, 12 estudantes da turma A tiveram nota média de 78% com desvio padrão de 6%, enquanto que 15 estudantes da turma B tiveram nota média de 74% com desvio padrão de 8%. Ao nível de significância de 0.05, verifique se, em média, a turma A é superior à turma B (considere as populações normais com variâncias iguais).

Exercício 6.12 Uma clínica pretende comparar dois tipos de dietas. Com esse objetivo, escolheu aleatoriamente e independentemente, uma amostra de 100 pacientes com excesso de peso e durante 10 semanas metade desses pacientes foram sujeitos à dieta 1 e os restantes à dieta 2. Após as 10 semanas, anotou-se o total de peso perdido (em kg) por cada paciente e obteve-se $\bar{x}_1 = 9.3$ e $s_1 = 2.4$ para a dieta 1 e $\bar{x}_2 = 8.2$ e $s_2 = 2.6$ para a dieta 2. Deve-se admitir que as duas dietas têm, em média, o mesmo efeito na perda de peso? Justifique a resposta considerando $\alpha = 0.05$.

Exercício 6.13 Após longos estudos, concluiu-se que o tempo (em segundos) de reação dos condutores a obstáculos na estrada, pode ser descrito por uma variável aleatória com distribuição normal de média 0.95 seg. Num circuito com diversos obstáculos foram feitos testes a 100 condutores, verificando-se

$$\sum_{i=1}^{100} (x_i - 0.95)^2 = 50 \text{ seg}^2.$$

Teste, ao nível de significância de 1%, se o desvio padrão do tempo de reação é no máximo 1 segundo.

Exercício 6.14 Num exame de Estatística efetuado na 2.^a época do ano letivo de 2016/2017 numa Escola Superior, foram avaliados 31 alunos. Considerando estes alunos como uma amostra representativa da população dos alunos matriculados na unidade curricular de Estatística e tendo em conta que as notas seguem um comportamento normal e que, para essa amostra, se obtiveram os seguintes resultados:

$$\sum_{i=1}^{31} x_i = 299 \quad \sum_{i=1}^{31} (x_i - \bar{x})^2 = 120 .$$

Teste a hipótese $\sigma^2 = 5$ contra a hipótese $\sigma^2 > 5$ para um $\alpha = 0.05$.

Exercício 6.15 Duas amostras (extraídas de duas populações normais) consistindo em 21 e 9 observações, têm variâncias $s_1^2 = 16$ e $s_2^2 = 8$ respetivamente. Para $\alpha = 0.05$, teste a hipótese da variância da 1.^a população ser superior à da 2.^a população.

Exercício 6.16 Suponha que as produções (em gramas) de comprimidos, em intervalos de tempo fixos, aleatoriamente selecionados de duas máquinas $M1$ e $M2$ de um laboratório se podem considerar normais (com variâncias iguais). Os pesos obtidos em duas amostras permitiram determinar as quantidades seguintes:

$$\begin{array}{cc} \mathbf{M1} & \mathbf{M2} \\ \sum_{i=1}^8 x_i = 80.8 & \sum_{j=1}^9 y_j = 96.3 \\ \sum_{i=1}^8 x_i^2 = 816.664 & \sum_{j=1}^9 (y_j - \bar{y})^2 = 0.549 \end{array}$$

1. Verifique se é plausível considerar que a variabilidade, em gramas, da produção das duas máquinas é idêntica ($\alpha = 0.01$).
2. O técnico responsável pelas duas máquinas garante que a máquina 2 produz, em média, mais quantidade que a máquina 1. Ao nível de 5%, esta afirmação é consistente com os dados?

Exercício 6.17 O diretor de produção de uma dada empresa suspeita que uma dada máquina da cadeia de produção está avariada em mais de 10% dos dias de laboração. Tendo recorrido ao histórico de produção nos últimos 180 dias verificou que a máquina esteve avariada em 20 desses dias. Através de um teste de hipóteses paramétrico apropriado diga, justificando, se as suspeitas do diretor de produção são estatisticamente aceitáveis para um nível de significância de 5%.

1. Tome a decisão recorrendo à região crítica.
2. Tome a decisão recorrendo ao valor-p.

Exercício 6.18 O consumidor de um certo produto acusou o fabricante, dizendo que mais de 20% das unidades fabricadas apresentam defeito. Para confirmar a sua acusação, o consumidor usou uma amostra de tamanho 50, onde 27% das peças eram defeituosas. Mostre como o fabricante poderia refutar a acusação. Utilize um nível de significância de 5%.

Exercício 6.19 Membros de uma associação patronal desejam demonstrar que pelo menos 60% dos seus associados apoiam a política de privatização do governo. Determine a região crítica do teste de hipóteses para essa situação, para um nível de significância de 5%, supondo que os dados são recolhidos a partir de uma amostra com 80 associados selecionados ao acaso. Qual deveria ser a estimativa da proporção para não rejeitar a hipótese nula?

Exercício 6.20 Pretende-se saber a quantidade de nicotina (medida em miligramas) existente numa determinada marca de cigarros. Examinaram-se 5 cigarros dessa marca, tendo-se obtido os seguintes resultados:

16; 16.5; 19; 15.4; 15.6;

Sabendo que a variável segue uma distribuição normal.

1. Teste ao nível de significância de 0.1 a hipótese: $H_0 : \mu = 13.5$ contra $H_1 : \mu \neq 13.5$.
2. Teste, ao nível de significância de 0.05, a hipótese da quantidade média de nicotina ser de 14 mg contra a hipótese alternativa de ser de 13.5 mg.
3. Numa amostra de 100 fumadores, 20 contraíram cancro do pulmão. Será que a percentagem de fumadores que contraem cancro é superior a 18%? Justifique a sua resposta utilizando um teste de hipóteses adequado, com um nível de significância de 0.05.
4. Calcule o valor-p associado ao teste de hipóteses realizado na alínea 3.

Exercício 6.21 A tabela seguinte contém os resultados de um inquérito realizado junto de pessoas de ambos os sexos, acerca da opinião sobre a proibição de fumar nos restaurantes::

	Deve ser proibido	Não deve ser proibido
Número de homens	50	50
Número de mulheres	60	40

1. Um jornal afirma que a percentagem de mulheres a favor desta proibição se situa acima dos 65%. Verifique a veracidade desta afirmação elaborando um teste de hipóteses adequado (com $\alpha = 0.025$).
2. Teste a hipótese da percentagem de pessoas favoráveis à não proibição ser igual para ambos os sexos (com $\alpha = 0.05$).

Exercício 6.22 Um estudo nutricional detetou numa amostra de 55 hipertensos, 24 com dietas pobres em sódio. Paralelamente, numa amostra de 149 não hipertensos detetaram-se 36 com dietas pobres em sódio. Poderá concluir-se, para um nível de significância de 0.05, que a proporção de indivíduos sujeitos a dietas pobres em sódio é maior entre hipertensos?

Exercício 6.23 Uma rutura na coligação que o elegeu levou o Presidente da Câmara Municipal de Vila do Rio a demitir-se e à consequente marcação de eleições intercalares. O Presidente da Câmara demissionário, uma figura muito popular na região, decidiu recandidatar-se ao cargo. Uma rádio local pretende anunciar logo após o fecho das urnas se o presidente demissionário foi ou não reeleito. Para isso decidiu realizar uma sondagem à boca das urnas (isto é, as intenções de voto serão recolhidas logo após os eleitores terem votado) entrevistando 500 eleitores da margem sul do rio selecionados aleatoriamente, tendo obtido que 220 eleitores votaram no presidente demissionário. Uma rádio local concorrente também realizou uma sondagem à boca das urnas, tendo entrevistado 300 eleitores da margem norte, tendo registado que 156 votaram no presidente demissionário. A seleção dos eleitores nesta segunda sondagem também foi aleatória e as duas sondagens foram realizadas de forma independente. Verifique ao nível de significância de 0.05 se as duas sondagens são ou não contraditórias.

Soluções

6.1 1) Teste bilateral **2)** $RC =] - \infty, -1.960] \cup [1.960, +\infty[$ **3)** Rejeitar H_0

6.2 1) Com base na amostra e para um nível de significância de 5%, o segundo fabricante parece ter razão. **2)** Com base na amostra e para um nível de significância de 10%, o segundo fabricante parece ter razão.

6.3 Com base na amostra e para um nível de significância de 5%, o comerciante não deve reclamar.

6.4 1) Com base na amostra e para um nível de significância de 1%, o produto não deve satisfazer as especificações. **2)** Com base na amostra e para um nível de significância de 1%, o conteúdo médio de cobre deve ser superior ao exigido.

6.5 1) $\alpha \geq 0.1056$ **2)** $\alpha > 0.05$

6.6 Com base na amostra e para um nível de significância de 5%, não há evidência estatística que o tempo que o analgésico leva a penetrar na corrente sanguínea tenha sido acelerado. Com base na amostra e para um nível de significância de 10%, não há evidência estatística que o tempo que o analgésico leva a penetrar na corrente sanguínea tenha sido acelerado.

6.7 1) Com base nas amostras e para um nível de significância de 5%, a dona da boutique deve comprar as calças ao fornecedor 2. **2)** Nenhuma, a decisão mantém-se. **3)** Com base nas amostras e para um nível de significância de 10%, a dona da boutique deve comprar as calças ao fornecedor 2.

6.8 1) Com base nas amostras e para um nível de significância de 1%, não há evidência estatística que a dieta com alto conteúdo de proteína aumente o ganho de peso. **2)** valor $-p \approx 0.05$

6.9 1) Amostras emparelhadas **2)** Com base nas amostras e para um nível de significância de 10%, há evidência estatística que a tarefa realizada na máquina do fabricante 1 demora, em média, mais do que na máquina do fabricante 2.

6.10 1) Amostras emparelhadas **2)** Com base nas amostras e para um nível de significância de 1%, não há evidência estatística que a suspeita do diretor faça sentido.

6.11 Com base nas amostras e para um nível de significância de 5%, há evidência estatística que, em média, a turma A é superior à turma B .

6.12 Com base nas amostras e para um nível de significância de 5%, as dietas parecem ter efeitos diferentes na perda de peso.

6.13 Com base nas amostras e para um nível de significância de 1%, há evidência estatística que o desvio padrão do tempo de reação é no máximo 1 segundo.

6.14 Não rejeitar H_0

6.15 Com base nas amostras e para um nível de significância de 5%, há evidência estatística que a variância da 1.^a população é superior à da 2.^a população.

6.16 1) Com base nas amostras e para um nível de significância de 1%, há evidência estatística que a variabilidade da produção das duas máquinas é idêntica. **2)** Com base nas amostras e para um nível de significância de 5%, a afirmação é consistente com os dados.

6.17 Com base na amostra recolhida e para um nível de significância de 5%, não há evidência estatística para corroborar as suspeitas do diretor de produção.

6.18 Com base na amostra recolhida e para um nível de significância de 5%, há evidência estatística para refutar a acusação.

6.19 $p^* > 0.5099$

6.20 1) Rejeitar H_0 **2)** Não rejeitar H_0 **3)** Com base na amostra recolhida e para um nível de significância de 5%, não há evidência estatística que a percentagem de fumadores que contraem cancro é superior a 18%. **4)** valor $-p = 0.3015$

6.21 1) Com base na amostra e para um nível de significância de 2.5%, a afirmação do jornal não parece ser verdadeira. **2)** Com base nas amostras e para um nível

de significância de 5%, há evidência estatística que a percentagem de pessoas favoráveis à não proibição é igual para ambos os sexos.

6.22 Com base nas amostras e para um nível de significância de 5%, há evidência estatística que a proporção de indivíduos sujeitos a dietas pobres em sódio é maior entre hipertensos.

6.23 Com base nas amostras e para um nível de significância de 5%, há evidência estatística que as duas sondagens são contraditórias.