**Introdução à Probabilidade e Estatística II
Exercícios para revisão e autoteste**“Estatística para Economistas”, Rodolfo Hoffmann
“Probabilidade e Estatística para engenharia e ciências”, Walpole e Myers
“Statics for Business & Economics”, Paul Newbold

**TESTE DE HIPÓTESE**

**1.** Uma nova cura foi desenvolvida para um tipo de cimento que resulta em uma força de compressão de 5.000 quilogramas por centímetro quadrado e desvio padrão de 120. Para testar a hipótese de que µ=5.000 contra a alternativa de que µ<5.000, uma amostra aleatória de 50 peças de cimento é testada. A região critica é definida como $\overbar{x}$ < 4.970.
a) Determine a probabilidade de cometer um erro tipo I quando H0 é verdadeira
b) Avalie ß para as alternativas µ=4.970 e µ=4.960.

**2.** Afirma-se que um automóvel é dirigido, em média, mais de 20.000 quilômetros por ano. Para testar essa afirmação, uma amostra aleatória de cem proprietários de automóveis registra os quilômetros viajados. Você concordaria com essa afirmação, se esta amostra mostrasse uma media de 23.500 quilômetros e desvio padrão de 3.900? Use um valor P em sua conclusão.

**3.** Teste a hipótese de que o conteúdo médio de recipientes de certo lubrificante é dez litros, se os conteúdos de uma amostra aleatória de dez recipientes são: 10,2; 9,7; 10,1; 10,3; 10,1; 9,8; 9,9; 10,4; 10,3 e 9,8 litros. Use o nível de significância de 0,01 e assuma que a distribuição dos conteúdos é normal.

**4.** Um especialista em marketing de uma fabrica de massas acredita que 40% dos amantes de massas preferem lasanha. Se nove de 20 amantes de massa escolhem lasanha em vez de outras massas, o que podemos concluir sobre a afirmação? Use um nível de significância de 5%.

**5.** Um novo equipamento de radar está sendo considerado para um sistema de defesa antimíssil. O sistema é experimentado em uma aeronave real, na qual uma morte ou não morte é simulada. Se, de 300 tentativas, ocorrem 250 mortes, aceite ou rejeite, no nível de significância 0,04, a afirmação de que a probabilidade de uma morte com o sistema não excede a probabilidade de 0,8 do equipamento já existente.

**6.** Uma máquina de refrigerantes é considerada fora de controle se a variância dos conteúdos exceder 1,15 decilitros. Se uma amostra aleatória de 25 copos de bebida dessa máquina tem variância de 2,03 decilitros, isso indica, no nível de significância de 0,05, que a máquina está fora de controle? Assuma que os conteúdos tem distribuição aproximadamente normal.

**7.** Uma máquina deveria misturar amendoins, avelãs, castanhas de caju e nozes em uma razão de 5:2:2:1. Uma lata com 500 dessas frutas secas misturadas contém 269 amendoins, 112 avelãs, 74 castanhas de caju, e 45 nozes. Num nível de significância de 0,05, teste a hipótese de que a máquina está misturando as frutas na razão adequada.

**8.** Pretende-se lançar uma moeda 5 vezes e rejeitar a hipótese de que a moeda é não tendenciosa, isto é, pretende-se rejeitar H0: p= ½ se em 5 jogadas ocorrerem 5 coroas ou 5 caras. Qual é a probabilidade de cometer erro tipo I?

**9.** Você suspeita que um dado é viciado, isto é, você suspeita que a probabilidade de obter ‘seis’ é maior do que 1/6. Você decide testar a hipótese de que o dado é não viciado, jogando-o 5 vezes e rejeitando essa hipótese se ocorrer a face ‘seis’ 4 ou 5 vezes. Qual é a probabilidade de ocorrer erro tipo I?

**10.** Uma moeda é lançada cinco vezes com o propósito de testar H0: p= ½ contra HA: p≠ ½. Se os lances resultarem em apenas uma cara, você concluiria, ao nível de significância de 6,25%, que a moeda é viciada?

**11.** Em cada uma das quatro faces de dois tetraedros regulares, aparentemente idênticos, estão marcados os valores 1,2,3 e 4. Entretanto, um dos tetraedros é feito de material homogêneo, enquanto o outro é ‘chumbado’, de forma que ao joga-lo, a face com o valor 4 tem probabilidade de 0,5 de ficar em contato com a mesa e cada uma das outras faces tem probabilidade 1/6. Suponha que um desses dois tetraedros foi lançado 48 vezes, para testar a hipótese H0 de que foi escolhido o tetraedro feito de material homogêneo, contra a hipótese HA de que foi escolhido o tetraedro chumbado. Vamos supor ainda que foi estabelecida a seguinte regra de decisão: se a face com o valor 4 for obtida 20 vezes ou mais, rejeita-se H0, caso contrário, rejeita-se HA. Determine as probabilidades de cometer erro tipo I e erro tipo II.

**12.** Para testar a hipótese de que a proporção (p) de analfabetos em certa população é igual a 0,5, vamos proceder a um teste bilateral, ao nível de significância de aproximadamente 6%. Suponha que dispomos de uma amostra aleatória de cinco pessoas. Delimite a região de rejeição, ou seja, se X é o número de analfabetos na amostra, determine para que valores de X rejeitamos H0.

**13.** Dados os valores 4,6,3,6 e 6 de uma amostra de 5 observações de uma variável aleatória X, estime a média e a variância de X. Admitindo que X tem distribuição normal, teste, ao nível de significância de 5%, a hipótese de que a média da população é 1, contra a hipótese alternativa de que é maior do que 1.

**14.** Uma amostra de quatro valores da variável aleatória X, com distribuição normal, apresentou os valores 8,3,5 e 12. Teste, ao nível de significância de 5%, a hipótese de que a média da população é igual a 13.

**15.** Uma amostra de nove elementos de uma população infinita forneceu os seguintes valores: 3,4,5,3,6,11,8,2 e 3.
a) Determine a média aritmética, a mediana e a(s) moda(s) da amostra.
b) Calcule a estimativa não tendenciosa da variância.
c) Admitindo que a variável observada tenha distribuição normal com média µ, teste a hipótese H0 : µ=3 contra a hipótese alternativa HA : µ>3. Ao nível de significância de 5%.

**16.** Uma amostra de 11 valores da variável aleatória X, que tem distribuição normal, apresentou os valores tabelados a seguir.

|  |  |
| --- | --- |
| **Valor de X** | **Frequência** |
| 4,5 | 4 |
| 4,0 | 4 |
| 3,5 | 2 |
| 3,0 | 1 |

a) Determine as estimativas da média e do desvio padrão de X.
b) Determine o intervalo de 95% de confiança para a média da população.
c) Determine o tamanho que deveria ter a amostra para que a média da amostra diferisse da média da população de 0,2 de unidade ou menos, ao nível de confiança de 99%.
d) Teste, ao nível de significância de 5%, a hipótese de que a média da população é 3,7 contra a hipótese de que a média da população é maior do que 3,7.

**17.** Uma amostra aleatória de 12 valores da variável X, com distribuição normal e população infinita, apresentou os valores 10,8,7,2,8,6,5,2,10,7,9 e 10.
a) Determine as estimativas não tendenciosas da média e da variância de X.
b) Determine o intervalo de confiança, ao nível de 90% de confiança, para µ=E(X).
c) Determine o tamanho que deveria ter a amostra para que a média da amostra diferisse da média da população de uma unidade ao menos, ao nível de confiança de 95%.
d) Teste ao nível de significância de 5%, a hipótese de que a média da população é igual a 5, contra a hipótese de que a média da população é maior que 5.

**18.** Refaça os itens b), c) e d) da questão anterior, considerando que a população tem apenas 1.296 elementos, com distribuição aproximadamente normal, e que a amostragem é feita sem reposição.

**19.** Suponha que pretendemos comparar as médias de duas populações, com base em duas amostras, mas a variável em análise é medida em escala ordinal. Podemos usar o teste t? Explique sumariamente.

**20.** Duas amostras aleatórias de tamanhos 20 e 30, obtidas de duas populações normais independentes, deram as seguintes estimativas: $\overbar{X}$1=10, $\overbar{X}$2=13 e s1=s2=6. Teste, ao nível de significância de 10%, a hipótese de que as médias das duas populações são iguais, supondo que as variâncias das duas populações são iguais.

**21.** Duas amostras aleatórias independentes, de tamanhos 10 e 12, obtidas de duas populações normais, forneceram as seguintes estimativas: $\overbar{X}$1=40, $\overbar{X}$2=48 e s1=10, s2=12. Teste, ao nível de significância de 5%, a hipótese de que µ1= µ2, supondo que σ1= σ2.

**22**. Uma amostra de 6 valores da variável X1 forneceu os valores $\overbar{X}$1=13,2 e $\sum\_{}^{}($X1i-$\overbar{X}$1)²=24 e uma amostra de 12 valores da variável X2 forneceu os valores $\overbar{X}$2=10,9 e $\sum\_{}^{}($X2i-$\overbar{X}$2)²=40. Que condições devem ser satisfeitas para que a aplicação do teste t, na comparação das médias das duas variáveis, seja rigorosamente válida? Supondo que essas condições são obedecidas e considerando um nível de significância de 5%, teste a hipótese de que as médias das duas variáveis são iguais contra a hipótese de que essas médias são diferentes.

**23.** Uma amostra da população A forneceu os valores 10,4,8,11,14,12,9,13 e 9, e uma amostra da população B forneceu os valores 14,16,16 e 10. Supondo que a variável em análise tem distribuição normal com variância σ², em ambas as populações, teste, ao nível de significância de a) 5% e b) 1%, a hipótese de que a média da população A é igual à média da população B.

**24.** Foram feitas entrevistas com duas amostras de 80 pessoas, uma de cada indústria, para testar a hipótese de que a proporção de empregados analfabetos é a mesma nas duas indústrias. Os resultados obtidos são os seguintes:

|  |  |  |
| --- | --- | --- |
| **Categoria** | **Indústria A** | **Indústria B** |
| Alfabetizado | 46 | 34 |
| Analfabeto | 34 | 46 |

Faça o teste considerando um nível de significância de a)5% e b)10%.

**25.** Suponha que você pretende verificar se a renda média dos advogados é igual à renda média dos economistas, no estado de São Paulo. Para isso, você obteve uma amostra de 12 advogados, que forneceu os valores 1,1,8,7,8,4,3,8,9,2,3 e 6, e uma amostra de 6 economistas, que forneceu os valores 9,13,6,7,10 e 9.. Supondo que as variâncias são iguais, teste a hipótese de que os economistas têm, em média, renda mais alta,
a) ao nível de significância de 2,5%
b) ao nível de significância de 0,5%

**26.** Suponha duas amostras independentes, de oito elementos cada uma, provenientes de duas populações distintas. Vamos admitir que a variável em analise tem distribuição normal e que σ12= σ22. Se a primeira amostra forneceu os valores 8,7,3,9,6,6,8 e 9 e a segunda amostra forneceu os valores 5,8,11,10,8,11,10 e 9, teste, ao nível de significância de 5%, a hipótese H0:µ1= µ2 contra a alternativa HA:µ1< µ2.

**27.** Duas amostras independentes, de quatro elementos cada uma, provenientes de duas populações com distribuições normais, variâncias iguais e médias µ1 e µ2 , respectivamente, fornecem os valores:

|  |  |
| --- | --- |
| **Amostra 1** | **Amostra 2** |
| 75 | 52 |
| 70 | 60 |
| 60 | 42 |
| 75 | 58 |

a) Teste, ao nível de significância de 5%, a hipótese H0:µ1= µ2b) Calcule o intervalo de confiança de 95% para µ1c) idem, para µ2
d) idem, para µ1- µ2

**28**. Um método rápido, porém impreciso, de determinar a concentração de uma solução é comparado com o método padrão. Tomam-se 12 amostras de uma solução. Em 8 dessas amostras a concentração é determinada pelo método rápido e em 4 amostras a concentração é determinada pelo método padrão. A tabela a seguir apresenta os resultados obtidos.

|  |  |
| --- | --- |
| **Método padrão** | **Método rápido** |
| 25 | 23 17 |
| 24 | 18 25 |
| 25 | 22 19 |
| 26 | 28 16 |

Admitindo que os resultados dos dois métodos de determinação da concentração sejam variáveis com distribuições normais, verifique se o método rápido, quando comparado com o método padrão, tende , em média, a superestimar ou subestimar o grau de concentração, adotando o nível de significância de 5%. Uma vez que o método rápido é sabidamente mais impreciso, não podemos pressupor que sejam iguais as variâncias dos resultados obtidos pelos dois métodos.

**29**. Na tabela a seguir são dadas as rendas reais anuais de nove diferentes indivíduos em dois anos, em milhares de reais. Teste a hipótese de que, em média, a renda real nesses dois anos se manteve a mesma, contra a hipótese de que houve variação na renda média, ao nível de significância de 5%.

|  |  |  |
| --- | --- | --- |
| **Indivíduo** | **Renda (ano 1)** | **Renda (ano 2)** |
| A | 6 | 7 |
| B | 5 | 8 |
| C | 7 | 8 |
| D | 3 | 5 |
| E | 3 | 9 |
| F | 2 | 2 |
| G | 6 | 2 |
| H | 2 | 7 |
| I | 2 | 6 |

**30**. Utilizou-se uma amostra de 112 alunos para analisar a relação entre inteligência (medida por QI) e nota em determinado curso. Os alunos foram classificados em dois estrator conforme o nível de QI e em dois estratos conforme o nível da nota no curso, obtendo-se a seguinte tabela:

|  |  |  |
| --- | --- | --- |
| **QI** | **Nota alta** | **Nota baixa** |
| Alto | 54 | 2 |
| Baixo | 38 | 18 |

Teste a hipótese de que a nota obtida independe do QI, considerando um nível de significância de 1%.

**31.** Suponhamos que a experiência tem mostrado que são reprovados 20% dos alunos, que se submetem a determinado tipo de exame. Se, de uma turma de 100 alunos, forem reprovados apenas 13, podemos concluir, em um teste unilateral, ao nível de significância de 5%, que os alunos dessa turma são, em média, melhores?

**32.** O fabricante de certa peça afirma que pelo menos 90% das peças que produz não tem defeitos. O exame de uma amostra de 200 peças revelou que 30 eram defeituosos. Teste a afirmativa do fabricante, considerando um nível de significância de 5%.

**33.** O fabricante de certe peça afirma que pelo menos 80% das peças que produz não tem defeitos. O exame de uma amostra de 400 pelas revelou que 100 peças eram defeituosas. Teste a afirmativa do fabricante, considerando um nível de significância de a)5% e b)1%.

**34.** Um exame é constituído por 400 testes tipo certo-errado.
a) Determine o número mínimo de testes que um aluno deve acertar para que o professor possa, ao nível de significância de 10%, rejeitar a hipótese de que o aluno nada sabe sobre a matéria e respondeu ao acaso, em favor da hipótese de que o aluno sabia alguma coisa sobre a matéria do exame.
b) Qual seria esse número mínimo, se o professor tivesse adotado um nível de significância de 5%?

**35.** Um exame é constituído por 100 testes tipo múltipla escolha, com cinco alternativas cada um, sendo correta apenas uma das alternativas.
a) Determine o número mínimo de testes que um aluno deve acertar para que o professor possa, ao nível de significância de 5%, rejeitar a hipótese de que o aluno respondeu ao acaso, em favor da hipótese de que o aluno sabia alguma coisa sobre a matéria do exame.
b) Qual seria esse número mínimo, se o professor tivesse adotado um nível de significância de 1%?

**36**. O rótulo de uma caixa de sementes informa que a taxa de germinação é de 90%. Entretanto, como a data de validade já foi ultrapassada, acredita-se que a taxa de germinação seja inferior a 90%. Faz-se um experimento e, de 400 sementes retiradas aleatoriamente, 350 germinam.
a) Ao nível de significância de 10%, rejeita-se a hipótese de que a taxa de germinação é de 90%?
b) Determine o intervalo de 95% de confiança para a taxa de germinação.

**37.** Em uma amostra de 30 pessoas do município A havia 20 católicos, e numa amostra de 50 pessoas do município B havia 44 católicos. Verifique se a proporção de católicos nos dois municípios é a mesma, considerando um nível de significância de 5%.

**38**. Em uma amostra de 300 pessoas do município A havia 200 católicos. Em uma amostra de 500 pessoas do município B havia 300 católicos. Verifique se a proporção de católicos nos dois municípios é a mesma, considerando um nível de significância de 5%.

**39.** Em dois anos consecutivos, foi feita uma pesquisa de mercado sobre preferência de donas de casa por determinada marca de certo produto. Para essa pesquisa foram utilizadas duas amostras independentes de 400 elementos. No primeiro ano, 33% das donas de casa preferiam a marca em estudo e, no ano seguinte, essa porcentagem era de 29%. Considerando um nível de significância de 5%, pode-se afirmar que houve mudança na preferencia das donas de casa?

**40.** Uma amostra de 100 empresários agrícolas cooperados mostrou que 80 usaram crédito e uma amostra de 400 não-cooperados mostrou que 300 usaram crédito. Verifique se a proporção de cooperados que usam crédito é estatisticamente superior à proporção de não-cooperados que usam crédito, ao nível de significância de a)5% e b)1%.

**41.** De uma amostra aleatória de 203 propagandas em revistas inglesas, 52 eram engraçados. Já de uma amostra aleatória de 270 propagandas em revistas americanas, 56 eram engraçados. Teste, contra uma hipótese alternativa bilateral, a hipótese nula de que a proporção de todos os anúncios engraçados nas revistas inglesas e americanas são iguais. *(Statics for Business & Economics, 4ª ed, pg 362)*

**SOLUÇÕES**

1. a) α=0,0384
 b) ß=0,2776

2. Sim, concordaria com a afirmação. z=8,97; µ>20.000
Valor P <0,001.

3. Não rejeitar H0. t=0,77.

4. A afirmação não é refutada (teste unicaudal com Valor P=0,4044)

5. Não rejeitar H0. z=1,44.

6. A máquina está fora de controle. X²=42,37 com valor P=0,8998.

7. Rejeita-se H0, a razão não é 5:2:2:1. X²=10,14.

8. 6,25%

9. 0,334%

10. Rejeita-se H0 para X=0 e X=5. Com X=1 não rejeitamos, ou seja, não concluímos que a moeda é viciada.

11. α=0,62% e ß=9,7%

12. X=0 e X=5 e (α=6,25%)

13. $\overbar{X}$=5; s²=2. Como t=6,32>t0 = 2,13, rejeita-se H0.

14. Como t=-3,06 e t0=3,18, não se rejeita H0.

15.a)$ \overbar{X}$=5, mediana =4 e moda=3
 b) s²=8,5
 c) Como t=2,06 e t0=1,86, rejeita-se H0.

16.a)$ \overbar{X}$=4 e s=0,5
 b) 3,66<µ<4,34
 c) n=63
 d) t=1,99, significativo (t0=1,81)

17.a)$ \overbar{X}$=7 e s²=8
 b) 5,53<µ<8,47
 c) n≥39
 d) t=2,449, significativo (t0=1,796)

18.b) 5,54<µ<8,46
 c) n=38
 d) t=2,461, significativo (t0=1,796)

19. A rigor, só podemos usar o teste t se a variável tiver distribuição normal. Isso implica que a variável é continua e é medida em escala de intervalos ou escala-razão (medida de terceiro ou quarto nível). Se a variável é ordinal, deve ser utilizado um teste não-paramétrico.

20. t=1,73, significativo (t0=1,68)

21. t=1,68, não significativo (t0=2,09)

22. As variáveis devem ter distribuições normais e variâncias iguais. As duas amostras devem ser aleatórias e independentes;
t=2,3 significativo (t0=2,12)

23.a) t=2,25, significativo (t0=2,20)
 b) t=2,25, não significativo (t0=3,11)

24.a) Z=1,74, não significativo (Z0=1,96)
 b) Z=1,74, significativo (Z0=1,64);
sem correção de continuidade, Z=1,90

25.a) t=2,83, significativo (t0=2,12)
 b) t=2,83, não significativo (t0=2,92)

26. t=2,00, significativo (t0=1,76)

27.a) t=3,17, significativo (t0=2,45)
 b) 58,75<µ1<81,25
 c) 40,14< µ2<65,86
 d) 3,86< µ1- µ2<30,14

28. t’=2,592, significativo (t’0=2,306)

29. t=2,00, não significativo (t’0=2,31)

30. Z=3,70, não significativo (Z0=2,58);
sem correção de continuidade, Z=3,95

31. Z=-1,625, não significativo. A região de rejeição é Z≤-1,645

32. Z=2,24, significativo (Z0=1,64)

33.a) Z=2,44, significativo (Z0=1,64)
 b) Z=2,44, significativo (Z0=2,33)

34.a) 214
 b) 217

35.a) 28
 b) 30

36.a) Z=-158, significativo. A região de rejeição é Z≤-1,28
 b) 0,843<p<0,907

37. Z=2,02, significativo (Z0=1,96);
sem correção de continuidade, Z=2,31

38. Z=1,81, não significativo (Z0=1,96);
sem correção de continuidade, Z=1,89

39. Z=1,15, não significativo (Z0=1,96);
sem correção de continuidade, Z=1,22

40.a) Z=0,92, não significativo (Z0=1,64)
 b) Z=0,92, não significativo (Z0=2,33);
sem correção de continuidade, Z=1,05

41. A hipótese nula pode ser rejeitada num nível de significância maior que 20.76% (p-value=0.2076).