FACULDADE FIAP



Conceitos Estatísticos

|  |  |
| --- | --- |
| NOME | RM |
| Rafael Favretti da Silva | 335332 |
| Mauricio Gallo Fausto Jr | 336168 |

**Avaliação 01 - Conceitos Estatísticos para IA**

**Exercício 1**

Uma amostra de 500 pessoas adultas de certa população foi sorteada e foram observdas duas variáveis para cada pessoa: dependência de álcool(sim/não) e número de progenitores com dependência (0,1 ou 2). Entre os 150 portadores de dependência, 43 não tinham qualquer progenitor dependente e 42 tinham 1 progenitor dependente. Das 500 pessoas entrevistadas, 149 tinham progenitor dependente e 108 tinham 2 progenitores dependentes.

1. Classifique as variáveis do estudo
2. Coloque as informações em uma tabela de contingência
3. Qual é a proporção de pessoas com 1 progenitor dependente?
4. Qual é a proporção de pessoas com 1 progenitor dependente dentre as que são dependentes de álcool?
5. Com base nessa amostra, você diria que existe um componente familiar que influi no alcoolismo? Justifique. Qual o teste que você aplicaria? Com alpha=5%. Formule o teste a decisão tomada.

**Exercício 2**

Em uma cidade, estima-se que a temperatura em setembro seja normalmente distribuída com uma média de 28º C e um desvio padrão de 5º C, enquanto em junho a temperatura ( em o C) tem distribuição N(32,16). Uma pessoa deseja viajar para essa cidade.

1. Qual é o melhor mês para viajar (junho ou setembro) tendo em vista que o viajante prefere temperaturas menores que 34º C ?
2. Considerando que no mês de setembro as passagens áreas são mais baratas, o viajante deseja saber qual é a probabilidade de que nesse mês a temperatura esteja entre 24º C e 30º C?

**Exercício 3**

Sabe-se que as pontuações de um teste de quociente de inteligência (QI), para indivíduos do mesmo grupo etário, segue uma distribuição N(65,18).

1. É desejável organizar os participantes em três grupos: Baixa, Média e Alta Inteligência. Quais são as pontuações que diferenciam cada grupo de inteligência se um pesquisador deseja que 20% dos indivíduos sejam considerados no primeiro grupo, 65% no segundo e 15% no terceiro?
2. Determine a porcentagem da população que tem uma pontuação entre 59 e 70.
3. Encontre um intervalo simétrico em torno da média que contenha 95% dos possíveis valores de QI.
4. De acordo com um pesquisador, uma pessoa é considerada “atípica” se o valor de QI for superior a 78 ou inferior a 53 pontos. Sorteadas 4 pessoas ao acaso, qual a probabilidade de encontrarmos ao menos 3 atípicas?
5. Para uma população de 7.500, qual é o número esperado de indivíduos com pontuaçãp superior a 78?

**Exercício 4**

Uma loja de utensílios vende em média 320 peças por dia, com desvio de 40. Após exaustiva campanha publicitária a administração pede para calcular a “nova média” de vendas para avaliar se ocorreu alguma melhora.

1. Quais as hipóteses nula e alternativa?
2. Em um teste clássico com α=5%. Qual deve ser a nova média para que Ho seja rejeitado.

**Exercício 5**

Uma amostra de 1250 graus universitários conferidos em uma certa universidade, acusou a seguinte tabela.

|  |  |  |  |
| --- | --- | --- | --- |
| Sexo \ Grau | Bacharel | Mestre | Doutor |
| Masculino | 505 | 160 | 25 |
| Feminino | 405 | 145 | 10 |

Teste a hipótese de que as variáveis Grau e Sexo são independentes com α=5%.

**Exercício 6**

Testar se há alguma relação entre as notas escolares e o salário de 300 pessoas entrevistadas aleatoriamente. Adote α=5%.

|  |  |  |  |
| --- | --- | --- | --- |
| Salários \ Notas | ALTA | MÉDIA | BAIXA |
| ALTO | 53 | 32 | 10 |
| MÉDIO | 42 | 71 | 30 |
| BAIXO | 17 | 29 | 16 |

**Exercício 7**

Um piloto de Fórmula Um tem 50% de probabilidade de vencer determinada corrida, quando esta se realiza sob chuva. Caso não chova durante a corrida, sua probabilidade de vitória é de 25%. Se o serviço de Meteorologia estimar em 30% a probabilidade de que chova durante a corrida, qual é a probabilidade deste piloto ganhar a corrida?

**Exercício 8**

Uma companhia multinacional tem três fábricas que produzem o mesmo tipo de produto. A fábrica I é responsável por 30% do total produzido, a fábrica II produz 45% do total, e o restante vem da fábrica III. Cada uma das fábricas, no entanto, produz uma proporção de produtos que não atendem aos padrões estabelecidos pelas normas internacionais. Tais produtos são considerados “defeituosos” e correspondem a 1%, 2% e 1,5%, respectivamente, dos totais produzidos por fábrica. No centro de distribuição, é feito o controle de qualidade da produção combinada das fábricas.

1. Qual é a probabilidade de encontrar um produto defeituoso durante a inspeção de qualidade?
2. Se durante a inspeção, encontramos um produto defeituoso, qual é a probabilidade que ele tenha sido produzido na fábrica II?

**Exercício 9**

Um técnico em aparelhos elétricos faz consertos em domicílio e deve consertar um ferro elétrico na casa de um cliente. Ele avalia que o defeito deve estar na tomada de força da área de serviço, no cabo de força de alimentação ou na resistência do ferro. Por experiência, ele sabe que as probabilidades: do defeito estar na tomada, no cabo ou na resistência são de 20%, 50% e 30%, respectivamente. Pensando em termos de ferramentas e peças de reposição do estoque que ele carrega, ele imagina que se o defeito for na tomada a probabilidade de conserto é de 95%. Se for no cabo de força é de 70% e se for na resistência é de 20%.

a. Qual a probabilidade de o técnico consertar o ferro no local com os seus recursos?

b. Qual a probabilidade do defeito ter sido no cabo de força, se o técnico conseguiu realizar o conserto?

c. O técnico chama o cliente e apresenta o ferro consertado. Perguntado do defeito, ele diz que teve que trocar a resistência (conserto mais caro). Qual a probabilidade de ele estar sendo sincero.

**Respostas**

**Exercício 1**

1. Dependência de álcool(sim/não) e número de progenitores com dependência.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Sem Progenitor Dependente** | **1 Progenitor Dependente** | **2 Progenitor Dependente** | **Total** |
| Dependente | 43 | 42 | 65 | 150 |
| Não Dependente | 106 | 201 | 43 | 350 |
| **Total** | **149** | **243** | **108** | **500** |

1. 48,6%
2. 28%
3. Graus de liberdade: 2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Observado** | | | | |
|  | **Sem Progenitor Dependente** | **1 Progenitor Dependente** | **2 Progenitor Dependente** | **Total** |
| Dependente | 43 | 42 | 65 | 150 |
| Não Dependente | 106 | 201 | 43 | 350 |
| **Total** | **149** | **243** | **108** | **500** |
|  | 29,8% | 48,6% | 21,6% |  |
|  |  |  |  |  |
| **Esperado** | | | | |
|  | **Sem Progenitor Dependente** | **1 Progenitor Dependente** | **2 Progenitor Dependente** | **Total** |
| Dependente | 44,7 | 72,9 | 32,4 | 150 |
| Não Dependente | 104,3 | 170,1 | 75,6 | 350 |
|  | **149** | **243** | **108** | **500** |
|  |  |  |  |  |
|  |  |  |  |  |
| Calculo Qui-quadrado | 0,06 | 13,10 | 32,80 |  |
|  | 0,03 | 5,61 | 14,06 | 65,66 |

Graus de liberdade: 2

Rejeitamos Ho ao nível de 5%, isto é os dados trazem evidência de uma forte dependência entre o fatores: Componente familiar e dependência.

**Exercício 2**

|  |  |
| --- | --- |
| **Setembro** | **Junho** |
|  |  |
| Z = 34-28 / 5 | Z = 34-32 / 4 |
| Z = 6/5 | Z = 2/4 |
| Z = 1,2 | Z = 0,5 |
| P = 0,8849 | P = 0,6915 |
| P = 88% | P = 69% |

O melhor mês para viajar será setembro.

|  |  |
| --- | --- |
| **24ºC** | **30ºC** |
|  |  |
| Z = 24-28 / 5 | Z = 30-28 / 5 |
| Z = -4/5 | Z = 2/5 |
| Z = -0,8 | Z = 0,4 |
| P = 1 – 0,2119 | P = 1 – 0,6554 |
| P = 78% | P = 34% |

A probabilidade em setembro de que a temperatura esteja entre 24º C e 30º C é de 44%.

**Exercício 3**

a.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **QI** | **Porcentagem** | **Z-Score** | **Média** | **Desvio** | **Total** |
| BAIXO | 20% | -0,84 | 65 | 4,242 | 49,88 |
| MEDIO | 65% | 0,39 | 65 | 4,242 | 72,02 |
| ALTO | 15% | -1,04 | 65 | 4,242 | 46,28 |

b.

|  |  |
| --- | --- |
| **QI 59** | **QI 70** |
|  |  |
| Z = 59-65 / 4,242 | Z = 70-65 / 4,242 |
| Z = -6/4,242 | Z = 5/4,242 |
| Z = -1,41 | Z = 1,17 |
| P = 0,0793 | P = 0,8790 |
| P = 7% | P = 87% |

87 – 7 = 80

A chance é de 80% para que o QI esteja entre a população de 59 e 70.

1. 65 + 1,96 \* 4,242 = 73,31

65 – 1,96 \* 4,242 = 56,68

O intervalo médio é de 56,68 a 73,31.

|  |  |
| --- | --- |
| **QI 53** | **QI 78** |
|  |  |
| Z = 53-65 / 4,242 | Z = 78-65 / 4,242 |
| Z = -12/4,242 | Z = 13/4,242 |
| Z = -2,828 | Z = 3,064 |
| P = 1 – 0,9976 | P = 1 – 0,9989 |
| P = 0,0024 | P = 0,0011 |

(0,0024 + 0,0011) \* 100 = 0,35% de chances de uma pessoa ser atípica.

P(3)=4!/3!(4−3)!⋅0.0035³⋅(1−0.0035)³ˉ⁴ = 0,00000017089975 ou aproximadamente 0,0000002% de chances de 3 pessoas atípicas.

1. 7500 \* 0,2358‬ = 1.768,5‬

**Exercício 4**

1. H0 <= 320

HAlternativa > 320

1. 1,64 = (X-320)/40 = 385,6

Caso a amostra tenha uma média maior que 385,6 a campanha teve efeito.

**Exercício 5**

Rejeitamos Ho ao nível de 5%, isto é os dados trazem evidência de uma forte dependência entre o fatores: Sexo\Grau

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Observado** | | | | |
| **Sexo\Grau** | **Bacharel** | **Mestre** | **Doutor** | **Total** |
| Masculino | 505 | 160 | 25 | 690 |
| Feminino | 405 | 145 | 10 | 560 |
| **Total** | **910** | **305** | **35** | **1250** |
|  | 72,8% | 24,4% | 2,8% |  |
|  |  |  |  |  |
| **Esperado** | | | | |
| **Sexo\Grau** | **Bacharel** | **Mestre** | **Doutor** | **Total** |
| Masculino | 502 | 168 | 19 | 690 |
| Feminino | 408 | 137 | 16 | 560 |
|  | 910 | 305 | 35 | 1.250 |
|  |  |  |  |  |
|  |  |  |  |  |
| Calculo Chiquadrado | 0,01 | 0,42 | 1,67 |  |
|  | 0,02 | 0,51 | 2,06 | 4,69 |

**Exercício 6**

Rejeitamos Ho ao nível de 5%, isto é os dados trazem evidência de uma forte dependência entre o fatores: Salários/Notas

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Observado** | | | | |
| **Salários/Notas** | **ALTA** | **MEDIA** | **BAIXA** | **Total** |
| ALTO | 53 | 32 | 10 | 95 |
| MÉDIO | 42 | 71 | 30 | 143 |
| BAIXO | 17 | 29 | 16 | 62 |
| **Total** | **112** | **132** | **56** | **300** |
|  | 37,3% | 44,0% | 18,7% |  |
|  |  |  |  |  |
| **Esperado** | | | | |
| **Salários/Notas** | **ALTA** | **MEDIA** | **BAIXA** | **Total** |
| ALTO | 35 | 42 | 18 | 95 |
| MÉDIO | 53 | 63 | 27 | 143 |
| BAIXO | 23 | 27 | 12 | 62 |
| **Total** | **112** | **132** | **56** | **300** |
|  |  |  |  |  |
|  |  |  |  |  |
| Calculo Chiquadrado | 8,67 | 2,30 | 3,37 | 21,65 |
| 2,43 | 1,04 | 0,41 |
| 1,63 | 0,11 | 1,69 |

**Exercício 7**

Probabilidade de vitória na chuva = 0,50

Probabilidade de vitória sem chuva = 0,25

Probabilidade de chuva = 0,30

Probabilidade de não chover = 0,70

Probabilidade de vitória = 0,50\*0,30 + 0,25\*0,70 = 0,325 ou 32%

**Exercício 8**

P(F1) = 0,30

P(F2) = 0,45

P(F3) = 0,25

P(DF1) = 0,01

P(DF2) = 0,02

P(DF3) = 0,015

a.

P(D) = P(DF1) \* P(F1) + P(DF2) \* P(DF2) + P(DF3) \* P(F3)

P(D) = 0,01 \* 0,30 + 0,02 \* 0,45 + 0,015 \* 0,25 = 0,01575 ou 1%

b.

P(DF2) \* P(F2) / P(A)

0,02 \* 0,45 / 0,1575 = 0,5714

A probabilidade de que o produto tenha sido produzido na fábrica 2 é de 57%.

**Exercício 9**

CN = Probabilidade de consertar o ferro

CT = Probabilidade de conserto na tomada

CC = Probabilidade de conserto no cabo

CR = Probabilidade de conserto na resistência

PC = Problema no cabo

PT = Problema na tomada

PR = Problema na resistência

a.

CT = 95%

CC = 70%

CR = 20%

CN = CT \* PT + CC \* PC + CR \* PR

CN = 0,95 \* 0,20 + 0,70 \* 0,50 + 0,20 \* 0,30 = 0,60

A probabilidade de conserto do ferro é de 60%.

b.

0,70 \* 0,50 / 0,60 = 0,58

A probabilidade de o defeito ter ocorrido no cabo caso o ferro tenha sido consertado é de 58%

c.

0,20 \* 0,30 / 0,60 = 0,10

A probabilidade de o defeito ter ocorrido na resistência caso o ferro tenha sido consertado e o técnico estar sendo sincero é de 10%