

## CE008 - Introdução à Bioestatística - Prova 1

1. O que acontece com a média e com o desvio-padrão:

(a) Se um mesmo número for somado a todos os elementos de um conjunto de dados?

Resp: Seja  $x$  o conjunto de dados original,  $\bar{x}$  a média de  $x$  e  $s$  o desvio-padrão de  $x$ . Denotando por  $k$  o valor adicionado a todos os elementos de  $x$  então: a média passará a ser  $\bar{x} + k$  e o desvio-padrão continuará a ser  $s$ .

(b) Se cada elemento de um conjunto de dados for multiplicado por um mesmo valor?

Resp: Seja  $x$  o conjunto de dados original,  $\bar{x}$  a média de  $x$  e  $s$  o desvio-padrão de  $x$ . Denotando por  $k$  o valor multiplicado a cada elemento de  $x$  então: a média passará a ser  $\bar{x} \times k$  e o desvio-padrão  $s \times k$ .

2. Numa grande população humana, a distribuição de daltonismo segundo sexo foi:

Daltonismo	Masculino	Feminino	Total
Presente	4,23%	0,65%	4,88%
Ausente	48,48%	46,64%	95,12%
Total	52,71%	47,29%	100%

(a) Uma pessoa é escolhida ao acaso dessa população. Qual é a probabilidade dela ser daltônica?

Resp: Seja o evento  $D$ : pessoa daltônica, então pela tabela acima,  $P(D) = 0,0488$ .

(b) Uma pessoa é escolhida ao acaso dessa população. Qual é a probabilidade dela ser daltônica sabendo-se que é do sexo masculino?

Resp: Seja o evento  $M$ : pessoa do sexo masculino, então:  $P(D|M) = 0,0423/0,5271 = 0,0802$ .

(c) Os eventos ser daltônico e ser do sexo masculino são independentes? O que isso significa na prática?

Resp: Não, pois  $P(D|M) \neq P(D)$ .

3. Os valores de sensibilidade e especificidade de um teste para a detecção de diabetes são  $s = 0,95$  e  $e = 0,95$ .

(a) Um paciente assintomático submete-se ao teste. Se o resultado for positivo, qual a probabilidade dele ter realmente a doença? (Considere que a prevalência de diabetes na população é 0,05.)

[1] "VPP=0.5"

(b) Um outro paciente que apresenta alguns sintomas da doença (como sede anormal e perda de peso) e possui história de diabetes na família também se submeteu ao teste. Se o resultado for positivo, o que você espera que aconteça com o VPP? Por quê?

Resp: Espera-se que o VPP aumente. Após o exame clínico e histórico familiar do paciente o médico pode assumir que a probabilidade de diabetes deste paciente seja 0,70. O VPP neste caso sobe para:

[1] "VPP=0.98"

4. Suponha que a prevalência de uma dada doença é  $1/100$  e que um teste diagnóstico para esta doença esteja sendo proposto. Se numa verificação preliminar deste teste 99% das pessoas realmente doentes e 1% das pessoas realmente sãs produziram resultados positivos, quais são o VPP e o VPN? Na sua opinião o teste é confiável? Justifique.

[1] "VPP=0.5"

[1] "VPN=1"

Resp: O teste só é bom para descartar a doença pois o VPN é 100%.

5. Suponha que a sensibilidade e a especificidade de dois testes (A e B) para o diagnóstico de uma determinada doença sejam:

Teste	sensibilidade	especificidade
A	0,9	0,7
B	0,7	0,9

- (a) Assumindo que os testes fossem realizados de forma independente, qual a ordem que produziria as melhores características gerais do teste conjunto em série (resultado positivo se ambos forem positivos)?

Resp: A ordem não afeta a sensibilidade e especificidade do teste em série.

- (b) Baseado na resposta de (a) e na informação de que os dois testes são igualmente invasivos e os custos de aplicação são os mesmos, em que ordem eles devem ser realizados para uma população cuja prevalência é  $p = 1000/11000$ ? (Sugestão: para as duas ordens, simule o que seria esperado para uma amostra de 11000 indivíduos.)

Resp: Com esta prevalência, espera-se que de 11000 indivíduos testados, 1000 sejam de fato doentes. Se o teste A for realizado primeiramente, espera-se que destes 1000 doentes, 900 ( $= 1000 \times 0,9$ ) tenham resultado positivo no teste e portanto tenham que fazer o teste B.

Se por outro lado, o teste B for realizado primeiramente, espera-se que dos 1000 doentes, 700 ( $= 1000 \times 0,7$ ) tenham que realizar o teste A.

Então o teste combinado de menor custo seria o segundo.

6. O tempo em meses entre a remissão de uma doença e a recidiva de 49 pacientes (24 homens e 25 mulheres) de uma determinada clínica médica foi registrado. A tabela abaixo apresenta a média e desvio-padrão desta variável, segundo o sexo do paciente:

Grupo	Média	Desvio-padrão
Homens	10,67	6,74
Mulheres	6,12	4,75

A frase: "As mulheres deste estudo são mais homogêneas no tempo entre remissão e recidiva da doença do que os homens pois o desvio-padrão no grupo das mulheres foi menor." Você concorda ou discorda desta frase? Justifique.

Resp: Discordo pois o grupo dos homens apresenta um  $CV = 6,74/10,67 = 0,63$  (63%) enquanto que no grupo das mulheres esta medida é  $CV = 4,75/6,12 = 0,78$  (78%).

7. Considere os três conjuntos de dados abaixo.

Amostra 1	0, 1, 2, 3, 4, 5
Amostra 2	0, 0, 0, 5, 5, 5
Amostra 3	0, 2.5, 2.5, 2.5, 2.5, 5

(a) Qual das 3 amostras tem a maior variabilidade? Por quê?

Resp: A Amostra 2 tem a maior variabilidade pois é aquela que tem mais valores individuais distantes da média.

```
[1] "sd(Amostra 1)= 1.87"
```

```
[1] "sd(Amostra 2)= 2.74"
```

```
[1] "sd(Amostra 3)= 1.58"
```

(b) Qual amostra tem a menor variabilidade? Por quê?

Resp: A Amostra 3 tem a menor variabilidade pois a maior parte dos valores individuais são iguais à média.