

Bioinformática – 2018/2019

Exercícios

Revisões de Estatística com recurso ao Testes de hipóteses. Tabelas de contingência.

- Gere aleatoriamente uma sequência de nucleótidos de dimensão 1000, s_1 , supondo que a ocorrência de "A", "C", "G" e "T", é feita ao acaso.
 - Gere aleatoriamente outra sequência de nucleótidos de dimensão 1000, s_2 , supondo que a ocorrência de "A", "C", "G" e "T" se verifica com uma probabilidade de 0.20, 0.30, 0.18 e 0.32, respectivamente.
 - Conte o número de "A", "C", "G" e "T" em cada uma das sequências.
 - Alinhe as duas sequências s_1 e s_2 e calcule o número total de posições em que há emparelhamento nas duas sequências.
 - Crie uma nova sequência s_3 , a partir de s_2 , que resulte de substituir a adenina (A) por timina (T).
- Admita-se uma molécula de DNA em que cada base tem a mesma probabilidade de ocorrer. Numa sequência de comprimento $n = 1000$, observa-se 280 ocorrências de A. Qual a probabilidade de se observar aquela ocorrência (ou um número mais extremo) sob a hipótese formulada?
- Considere uma sequência de DNA de comprimento $n=2000$, na qual se observaram 520 ocorrências de A, 460 de C, 560 de G e 460 de T.
Sob a hipótese da independência serão estes valores compatíveis com a hipótese de igual probabilidade de ocorrência de qualquer das bases?
- De um grupo de indivíduos vacinados contra a cólera foram escolhidos ao acaso 279 e de outro grupo de indivíduos não vacinados escolheu-se também ao acaso 539. Em cada um dos grupos registou-se posteriormente os que foram atacados e os que não foram atacados pela bactéria, tendo-se organizado os resultados no seguinte quadro:

	Atacado	Não atacado
Vacinados	3	276
Não vacinados	66	473

(a) Com base nestes dados, será que se pode afirmar que a população dos vacinados difere da dos não vacinados no que refere ao ataque pela bactéria? Justifique convenientemente, elaborando uma resposta completa.

(b) Alguém afirmou que **na população dos não vacinados** se podia admitir que 10% são atacados. O que pode dizer sobre esta afirmação? Justifique convenientemente.

5. É realizado um teste genético para averiguar qual o alelo que os indivíduos em teste possuem e um teste de diagnóstico para averiguar se o indivíduo possui uma dada doença. Os dados encontram-se no seguinte quadro

	Doença	
	Sim	Não
Alelo 1	45	122
Alelo 2	67	38

Pretende-se averiguar se o tipo de Alelo e a existência ou não de doença são independentes. Responda à questão de forma completa.

6. Para estudar o efeito do tipo de solo no crescimento de uma dada planta, plantaram-se amostras em três tipos de solo e classificou-se o crescimento das plantas em três categorias, tendo-se obtido os resultados seguintes:

	Tipo de Solo		
Crescimento	Barrento	Arenoso	Orgânico
Pobre	16	8	14
Médio	31	16	21
Bom	18	36	25

Com base nestes dados, será que se pode afirmar que o crescimento da planta difere significativamente consoante o tipo de solo ?

7. Pretende-se verificar se duas vacinas contra a brucelose (uma vacina padrão e uma outra nova) são igualmente eficazes. Realizou-se a seguinte experiência: um grupo de 24 bezerras tomou a vacina padrão e outro grupo de 26 bezerras tomou a vacina nova, considerando que os dois grupos estavam igualmente expostos ao risco de contrair a doença. Após algum tempo verificou-se quantos animais, em cada grupo, tinham contraído a doença. Os resultados estão no seguinte quadro:

	Brucelose	
Vacina	Sim	Não
Padrão	15	9
Nova	10	16

Poder-se-á considerar que há diferença estatisticamente significativa entre as proporções de bezerras que contraíram brucelose usando a nova vacina e a vacina padrão? Justifique convenientemente.