

MI813 A - Tópicos em Estatística Genética

1ª Lista de Exercícios

Profa.: Hildete Pinheiro

1. Suponha que dois alelos num locus dentro de um indivíduo são indexados por $j, j = 1, 2$, e os indivíduos amostrados são indexados por $i, i = 1, 2, \dots, n$. A variável indicadora X_{ij} é então definida por

$$X_{ij} = \begin{cases} 1 & \text{se o alelo } j \text{ no indivíduo } i \text{ é do tipo } A \\ 0 & \text{caso contrário} \end{cases}$$

A proporção amostral para o alelo A pode ser expressa por

$$\tilde{p}_A = \frac{1}{2n} \sum_{i=1}^n \sum_{j=1}^2 X_{ij}$$

Prove que:

$$E(\tilde{p}_A^2) = p_A^2 + \frac{p_A + P_{AA} - 2p_A^2}{2n}$$

e

$$\text{Var}(\tilde{p}_A) = \frac{p_A + P_{AA} - 2p_A^2}{2n}$$

2. (a) Use ambos Newton-Raphson e o algoritmo EM para estimar as frequências alélicas dos quatro alelos S, M, F e O se as frequências fenotípicas são:

$$\begin{array}{llll} n_{SS} + n_{SO} = 1149 & n_{MM} + n_{MO} = 36 & n_{FF} + n_{OF} = 17 & n_{OO} = 20 \\ n_{SM} = 336 & n_{MF} = 25 & n_{SF} = 203 & n = 1786 \end{array}$$

(b) Estime as variâncias e covariâncias dos estimadores. Assuma equilíbrio de H-W.

3. Faça 10 amostras bootstraps da seguinte amostra:

1	2	3	4	5	6	7	8	9	10
A_2A_2	A_1A_2	A_1A_1	A_1A_1	A_1A_1	A_1A_2	A_1A_2	A_1A_1	A_1A_2	A_1A_1

Numere as amostras originais de 1 a 10 e depois amostramos com reposição 10 números aleatórios. Por exemplo, o primeiro bloco de 10 seria:

3	0	2	4	6	8	6	1	4	9
A_1A_1	A_1A_1	A_1A_2	A_1A_1	A_1A_2	A_1A_1	A_1A_2	A_2A_2	A_1A_1	A_1A_2

A primeira proporção bootstrap para o alelo A_1 é então 0,70. Para este exercício, será necessário um ponto inicial arbitrário na tabela de números aleatórios e este pode ser o seu aniversário: A linha corresponde ao mês e a coluna ao dia.

Ordene as proporções bootstraps e tome a nona para definir um intervalo unilateral de 90% de confiança e compare com o valor encontrado pela teoria normal (utilizando $z = 1,28$).

4. Ache o estimador de máxima verossimilhança do coeficiente de desequilíbrio $D_A = P_{AA} - p_A^2$ e mostre que sua variância pode ser aproximada por

$$\frac{1}{n} \left[p_A^2(1 - p_A)^2 + (1 - 2p_A)^2 D_A - D_A^2 \right]$$

Faça o mesmo para D_{uv} , o desequilíbrio entre alelos u e v quando há mais de 2 alelos.

5. Estime as frequências dos gametas AB para os seguintes conjuntos de frequências genotípicas:

(a)	BB	Bb	bb	(b)	BB	Bb	bb	(c)	BB	Bb	bb
AA	154	81	8	AA	12	3	3	AA	12	3	3
Aa	37	14	3	Aa	3	54	3	Aa	3	51	3
aa	1	1	1	aa	12	3	3	aa	12	6	3