

# Exercício em sala 1: Frequências alélicas e genotípicas e equilíbrio de Hardy-Weinberg

Genômica Evolutiva – BIO 5789

*Diogo Meyer*

Esses exercícios têm o objetivo de fixar conceitos básicos sobre o equilíbrio de Hardy-Weinberg, contagem de alelos e genótipos e o uso de modelos em genética e evolução.

## 1. Diversidade genética e equilíbrio de Hardy-Weinberg

Foram colhidas amostras para sete populações, e as frequências genotípicas encontradas para um locus bialélico estão indicadas na tabela. Para cada população:

- Estime as frequências alélicas e as taxas de heterozigose.
- Avalie se a população está em equilíbrio de Hardy-Weinberg. Isso requer comparar as frequências genotípicas observadas com aquelas esperadas de acordo com o modelo de Hardy-Weinberg. Sua resposta pode ser simples e qualitativa, bastando dizer se as frequências observadas e esperadas são iguais ou diferentes.
- A **taxa de heterozigose** ( $H$ ) é usada para medir o grau de variabilidade genética numa população. Ela pode ser intuitivamente definida como a probabilidade de se amostrar dois alelos e eles serem diferentes. A forma de calcular o valor de  $H$  é exatamente a mesma usado para calcular a proporção de indivíduos heterozigotos esperados sob Hardy-Weinberg (afinal, a chance de haver heterozigotos é dada pela chance de dois alelos diferentes unirem-se). Compare as populações 1, 2, e 3 quanto à sua taxa de heterozigose.

genótipo	pop1	pop2	pop3
AA	0,36	0,81	0,4
Aa	0,48	0,18	0,4
aa	0,16	0,01	0,2

## 2. Seleção Natural sobre proporções Hardy-Weinberg

Considere os genótipos da população 1, e assumo que as frequências refletem o que é observado ao nascimento. Imagine que os genótipos “aa” sejam suscetíveis a uma infecção viral, de modo que todos os indivíduos com esse genótipo morram antes de se tornarem adultos, e portanto não contribuam para a amostra. Recalcule frequências genotípicas e alélicas para esse cenário e avalie se a amostra está em equilíbrio de Hardy-Weinberg (não é necessário fazer um teste estatístico).

## 3. Efeito de endogamia sobre Hardy-Weinberg

Um criador de plantas possuía em sua estufa uma população com frequências genotípicas conforme apresentadas para a população 1 acima. Ele submeteu todos os indivíduos dessa população a uma rodada de autofecundação, originando uma nova geração.

Dessa forma ele obteve uma população com as seguintes frequências genotípicas: 0,48; 0,24; 0,28 (para AA, Aa e aa, respectivamente).

- Essas frequências genotípicas são aquelas esperadas em função do regime de autofecundação ao qual as plantas foram submetidas?
- Compare as frequências alélicas e genotípicas antes e depois da endogamia. Elas diferem?

- (c) Para comparar a distância das frequências genotípicas observadas após a endogamia em relação às esperadas sob Hardy-Weinberg (isto é, caso não houvesse endogamia), use a seguinte fórmula, que nos dá o valor de  $f$ , o coeficiente de endogamia. Em que situação “f” seria zero?

$$f = \frac{H_{exp} - H_{obs}}{H_{exp}}$$

#### 4. Hardy-Weinberg numa população subdividida

Considere duas populações de uma espécie de primata, para a qual duas populações foram amostradas, ambas com tamanho de 100 indivíduos.

População 1:  $f_{AA} = 0.81$ ;  $f_{Aa} = 0.18$ ;  $f_{aa} = 0.01$  População 2:  $f_{AA} = 0.01$ ;  $f_{Aa} = 0.18$ ;  $f_{aa} = 0.81$

Para cada uma individualmente, avalie se está em Hardy-Weinberg.

Agora considere que o pesquisador ignore o fato de que essa espécie está subdividida em duas espécies, e simplesmente amalgame todos os indivíduos numa única amostra de 200 indivíduos. Para essa nova amostra, referente à espécie como um todo, calcule quais as frequências genotípicas e alélicas observadas. Essa espécie como um todo está em Hardy-Weinberg?

#### 5. Estudo de caso: usando HW para estudar evolução em populações humanas

Hoje em dia é bem estabelecido que populações humanas têm frequências genotípicas muito próximas ao esperado sob Hardy-Weinberg. Isso não significa que as populações atendem *exatamente* a todos os pressupostos do modelo, mas apenas que se há desvios em relação aos pressupostos, eles não são suficientemente intensos para causar grandes desvios. Para uma população não estar em Hardy-Weinberg, um dos pressupostos precisa estar sendo violado de modo drástico.

Por essa razão, quando há um desvio em relação ao esperado sob o equilíbrio de Hardy-Weinberg, é importante nos atermos e considerarmos o que pode estar ocorrendo para explicá-lo. O exemplo a seguir, que descreve dados reais, usa as previsões do modelo de Hardy-Weinberg para fazer inferências sobre o processo evolutivo.

##### A história dos Fore

Em populações humanas, há variabilidade não-sinônima (isto é, que resulta numa mudança no amino-ácido) no SNP rs1799990, que corresponde ao códon 129 do gene da *human prion protein gene* (PRNP). Esse polimorfismo resulta em dois possíveis amino-ácidos nessa posição: metionina ou valina. Ele é particularmente interessante pois indivíduos heterozigotos têm um fenótipo especial: são resistentes a uma doença neurodegenerativa adquirida chamada kuru.

Kuru é causada e transmitida por um prion codificado pela PRNP, e tornou-se conhecida para a medicina na década de 1950, quando havia presença colonial na Papua-Nova Guiné. Entre os povos dessa região estão os Fore, nos quais havia uma alta incidência de kuru, com mortalidade chegando a 2% por ano. Descobriu-se que o kuru é transmitido ao se ingerir os cérebros de indivíduos infectados com kuru, e que os Fore comiam parentes mortos em rituais funerários. Os homens tinham direito de escolher as partes do corpo comer antes dos demais, e deixavam as partes menos atraentes – como o cérebro, que é enriquecido em prions – para as mulheres. Essa prática funerária deixou de ser praticada pelos Fore, e os jovens Fore jovens já não a praticam.

Quando se descobriu que a PRNP com a mutação no SNP rs1799990 conferia resistência ao kuru, estudos populacionais foram realizados. Os resultados para as frequências genotípicas entre mulheres nascidas antes de 1950, entre homens nascidos antes de 1950, ou entre os Fore nascidos recentemente estão na Tabela abaixo.

A sua tarefa é usar a previsão do modelo de HW e fazer um teste estatístico (usando teste do Chi-quadrado, descrito no final deste documento) para identificar se em algum grupo populacional há desvio significativo. Com base nos resultados obtidos, proponha uma hipótese evolutiva capaz de explicar as frequências genotípicas entre os Fore.

Grupo Fore	Homozigotos para metionina	Heterozigotos	Homozigotos para valina
Mulheres nascidas antes de 1950	16	86	23

Grupo Fore	Homozigotos para metionina	Heterozigotos	Homozigotos para valina
Homens nascidos antes de 1950	34	111	60
Indivíduos Fore jovens	52	136	94

## Teste de Qui-quadrado ( $\chi^2$ ) para hipótese de equilíbrio de Hardy-Weinberg

O teste qui-quadrado é frequentemente utilizado para verificar se valores obtidos para dados reais correspondem aos esperados por uma previsão teórica. No nosso caso, testaremos se o número de indivíduos em cada classe genotípica corresponde ao esperado sob a hipótese da população estar em equilíbrio de Hardy-Weinberg.

O teste de qui-quadrado quantifica o quão “próximos” ou “distantes” os dados reais estão dos esperados pela teoria. Essa quantificação é feita através da estatística de qui-quadrado, definida abaixo:

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i}$$

Onde  $n$  é o número de classes.

Quanto maior o valor de  $\chi^2$ , mais distantes estão os dados reais dos observados. Para exprimir essa distância num contexto estatístico, o teste de qui-quadrado se baseia na comparação entre o valor de uma estatística obtida para os dados (neste caso,  $\chi^2$ ) e valores críticos apropriados de acordo com o nível de significância ( $\alpha$ ) e o número de graus de liberdade (g.l.) do teste.

No caso do teste da hipótese de equilíbrio de Hardy-Weinberg, a previsão teórica testada (Hipótese nula, ou  $H_0$ ) para os três genótipos (classes) é de que as frequências genotípicas  $F_{AA}$ ,  $F_{Aa}$  e  $F_{aa}$  (valores observados) estejam nas proporções esperadas  $p^2$ ,  $2pq$  e  $q^2$  (ocorrendo, portanto, com frequências esperadas  $p^2 * N$ ,  $2pq * N$  e  $q^2 * N$ ).

	AA	Aa	aa	Total
Observado	$F_{AA}$	$F_{Aa}$	$F_{aa}$	$N = F_{AA} + F_{Aa} + F_{aa}$
Esperado	$p^2 N$	$2pq N$	$q^2 N$	$N$
Contribuição para $\chi^2$	$\frac{(F_{AA} - p^2 N)^2}{p^2 N}$	$\frac{(F_{Aa} - 2pq N)^2}{2pq N}$	$\frac{(F_{aa} - q^2 N)^2}{q^2 N}$	$\chi^2$

Após calcular o valor de  $\chi^2$ , este é comparado com o valor crítico para o número de graus de liberdade (g.l.) apropriado e nível de significância ( $\alpha$ ) desejado. Caso o valor encontrado para  $\chi^2$  seja maior que o valor crítico, rejeita-se a hipótese.

Para o caso do teste de que a população encontra-se em equilíbrio de Hardy-Weinberg para um locus bialélico, em que o número de graus de liberdade é igual a 1, os valores críticos de  $\chi^2$  para diferentes níveis de significância são:

$\alpha$	10%	5%	1%
$\chi^2$ crítico	2,71	3,84	6,63

Se o valor de  $\chi^2$  encontrado for maior que o valor crítico para o  $\alpha$  selecionado, rejeita-se a hipótese de que a população está em equilíbrio de Hardy-Weinberg. Usualmente, adotamos um nível de significância de 5%.



Figure 1: Um grupo de indivíduos Fore, na Papua-Nova Guiné