Processos Evolutivos - BIO 0208

Exercício em sala 1: diversidade genética e equilíbrio de Hardy-Weinberg

Diogo Meyer

Esses exercícios têm o objetivo de fixar conceitos básicos sobre o equilíbrio de Hardy-Weinberg, contagem de alelos e genótipos, e desenvolver a intuição sobre como medidas de diversidade podem ser informativas sobre processos evolutivos.

1. **Diversidade genética e tamanho populacional.** Para sete populações foram feitas amostras de 100 indivíduos, e as frequências genotípicas encontradas num lócus autossômico estão indicadas na tabela. Para cada população: (a) estime as frequências alélicas, (b) avalie se ela está em equilíbrio de Hardy-Weinberg, e (c) calcule a *taxa de heterozigose, H* (ou H_{esp}).

genótipo	pop1	pop2	pop3	pop4	pop5	pop6	pop7
AA	36	81	100	40	0	50	25
Aa	48	18	0	40	100	0	50
aa	16	1	0	20	0	50	25

2. **Seleção Natural sobre proporções Hardy-Weinberg.** Considere os genótipos da população 1, e assuma 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 essa 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.** Imagine que a população 1 experimente uma geração de auto-fecundação, que é a forma mais extrema de endogamia (acasalamento preferencial com parentes). Calcule as frequências genotípicas que serão observadas. Para comparar a distância dessas frequências daquelas 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.
$f = \frac{H_{esp} - H_{obs}}{H_{esp}}$.

4. **Análise de dados e teste estatístico.** Numa amostra populacional de 99 indivíduos da planta Arabidopsis thaliana, descobriu-se que as frequências genotípicas para um gene que codifica uma enzima eram as seguintes: 45 FF, 52 SS, 2 FS.

   (a) Calcule as frequências genotípicas e alélicas observadas, e as frequências genotípicas esperadas sob Hardy-Weinberg.

   (b) Usando a informação no box sobre como implementar o teste de qui-quadrado, realize um teste para avaliar se essa amostra vem de uma população em equilíbrio de Hardy-Weinberg.

   (c) Estime o coeficiente de endocruzamento, f, desta população.

   (d) Refaça os itens (a) e (b) com as seguintes frequências genotípicas: 24 FF, 25 SS, 45 FS. As frequências genotípicas são numericamente iguais às esperadas sob Hardy-Weinberg? E segundo o teste, esses valores correspondem ao esperado?

Teste de Qui-quadrado (χ^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 χ^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, χ^2) e valores críticos apropriados de acordo com o nível de significância (α) 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^2N	$2pqN$	q^2N	N
Contribuição para χ^2	$\frac{(F_{AA}-p^2N)^2}{p^2N}$	$\frac{(F_{Aa}-2pqN)^2}{2pqN}$	$\frac{(F_{aa}-q^2N)^2}{q^2N}$	χ^2

Após calcular o valor de χ^2, este é comparado com o valor crítico para o número de graus de liberdade (g.l.) apropriado e nível de significância (α) desejado. Caso o valor encontrado para χ^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 χ^2 para diferentes níveis de significância são:

α	10%	5%	1%
χ^2 crítico	2,71	3,84	6,63

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