

**TESTES MULTINOMIAIS OTIMIZADOS:  
UMA APLICAÇÃO NO EQUILÍBRIO GENÉTICO DE  
HARDY-WEINBERG**

**ANDRÉ JALLES MONTEIRO**

Tese apresentada à Escola Superior de Agricultura  
“Luiz de Queiroz”, Universidade de São Paulo, para  
obtenção do título de Doutor em Agronomia, Área  
de Concentração: Estatística e Experimentação  
Agrônômica.

**P I R A C I C A B A**  
Estado de São Paulo - Brasil  
Março - 2002

**TESTES MULTINOMIAIS OTIMIZADOS:  
UMA APLICAÇÃO NO EQUILÍBRIO GENÉTICO DE  
HARDY-WEINBERG**

**ANDRÉ JALLES MONTEIRO**

Estatístico

Orientador: Prof. Dr. **JOSÉ EDUARDO CORRENTE**

Tese apresentada à Escola Superior de Agricultura  
“Luiz de Queiroz”, Universidade de São Paulo, para  
obtenção do título de Doutor em Agronomia, Área  
de Concentração: Estatística e Experimentação  
Agrônômica.

**P I R A C I C A B A**

Estado de São Paulo - Brasil

Março - 2002

**Dados Internacionais de Catalogação na Publicação (CIP)  
DIVISÃO DE BIBLIOTECA E DOCUMENTAÇÃO - ESALQ/USP**

Monteiro, André Jalles

Testes multinomiais otimizados: uma aplicação no equilíbrio genético de Hardy-Weinberg / André Jalles Monteiro. - - 2002.  
57 p.

Tese (doutorado) - Escola Superior de Agricultura Luiz de Queiroz, 2002.  
Bibliografia.

1. Biometria 2. Genética estatística I. Título

CDD 575.10212

**“Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor”**

*A Ariel Azeredo Jalles*

DEDICO.

## **AGRADECIMENTOS**

Ao Professor Doutor José Eduardo Corrente, pela orientação.

Ao Departamento de Estatística e Matemática Aplicada, da Universidade Federal do Ceará, pelo apoio.

À CAPES/PICDT, pelo auxílio financeiro.

Aos amigos do programa de pós-graduação em Estatística e Experimentação Agronômica, pelos bons momentos compartilhados.

A todos que direta ou indiretamente, contribuíram para a realização deste trabalho.

## SUMÁRIO

	Página
RESUMO.....	vi
SUMMARY.....	vii
1 INTRODUÇÃO.....	1
2 REVISÃO DA LITERATURA.....	4
2.1 Testes de hipóteses.....	4
2.2 Testes multinomiais.....	5
2.3 Equilíbrio genético.....	10
2.4 Otimização inteira.....	13
3 METODOLOGIA.....	15
3.1 Definição do teste de máximo volume.....	15
3.2 Aplicação do máximo volume no teste de equilíbrio genético.....	17
3.3 Teste de equilíbrio genético com máximo número de pontos na região de rejeição.....	19
3.4 Definição do teste de equilíbrio genético não enviesado.....	22
4 RESULTADOS E DISCUSSÃO.....	26
4.1 Aplicação da estatística $X^2$ , no teste de equilíbrio genético.....	26
4.2 Aplicação da estatística $2nl^{2/3}$ , no teste de equilíbrio genético.....	31
4.3 Aplicação do teste exato, no equilíbrio genético.....	33
4.4 Comparação de procedimentos não aleatorizados.....	37
4.5 Comparação de procedimentos aleatorizados.....	39
4.6 Procedimento não enviesado.....	44
4.7 Solução otimizada com máximo volume de poder.....	47
5 CONCLUSÕES.....	52
REFERÊNCIAS BIBLIOGRÁFICAS.....	54

**TESTES MULTINOMIAIS OTIMIZADOS:  
UMA APLICAÇÃO NO EQUILÍBRIO GENÉTICO DE  
HARDY-WEINBERG**

Autor: ANDRÉ JALLES MONTEIRO

Orientador: Prof. Dr. JOSÉ EDUARDO CORRENTE

**RESUMO**

A estatística  $X^2$ , quando na aplicação do teste de equilíbrio genético de Hardy-Weinberg, apresenta baixa eficiência, principalmente quando a amostra é de pequeno porte. Alguns procedimentos alternativos foram apresentados, com excelentes propriedades estatísticas: nível de significância homogêneo e não-viés. Esses procedimentos apresentam uma grande desvantagem prática: muitos pontos na região de rejeição são aleatorizados. No presente trabalho é apresentada uma nova propriedade, o máximo volume da função poder. Na busca do teste com essa propriedade, é sugerida uma forma de construção da região de rejeição, que apresenta o maior número de pontos, sem aleatorizações. Esse procedimento surge como uma adaptação da construção da região de rejeição com a propriedade de nível de significância homogêneo, sem a desvantagem de muitos pontos aleatorizados, apresentando a maior quantidade de combinações genotípicas, associadas ao não-equilíbrio genético, qualquer que seja o nível de significância preestabelecido do teste. Tem-se, assim, uma alternativa prática que viabiliza propriedades teóricas desejáveis, a um teste de hipóteses.

**OPTIMIZED MULTINOMIAL TESTS:  
AN APLICATION TO HARDY-WEINBERG GENETIC  
EQUILIBRIUM**

Author: ANDRÉ JALLES MONTEIRO

Adviser: Prof. Dr. JOSÉ EDUARDO CORRENTE

**SUMMARY**

The chi-square statistic, on the application of Hardy-Weinberg genetic equilibrium test, has low efficiency, mostly if the sample is scarce. Some alternative procedures have been presented, have valuable statistic proprieties: homogeneous significance level and unbiasedness. Those procedures have a practical disadvantage: various points are parcially in the critical region. At the present work, it is shown a new approach, the maximum volume of the power function, as a that to construct a critical region, with the maximum number of points that not randomization. This approach is an adaptation of the critical region constructions with homogeneous significance level propriety, but it does not have the disadvantage of many points parcially in the critical region. Furthermore, this procedure shows the greater quantities of combinations associated with Hardy-Weinberg genetic disequilibrium, whichever is the significance level in the test. Therefore, it is a practical alternative which makes possible the propriety theory are desidered in a hypothesis test.



# 1 INTRODUÇÃO

Nas mais diversas áreas do conhecimento, muitos experimentos são conduzidos de modo a obter resultados, que podem ser definidos em categorias disjuntas, nas quais se supõe que as probabilidades de ocorrência de cada categoria são fixas e independentes. Nesse caso, a variável resposta observada é a frequência com que cada categoria ocorre. Tal variável possui distribuição multinomial e, dessa forma, esses experimentos são chamados de ensaios multinomiais. As aplicações desse tipo de ensaio são as mais diversas.

Uma das aplicações consiste em testar se as probabilidades de obtenção de cada categoria possuem alguma específica estrutura. Alguns exemplos de tais testes, intitulados testes multinomiais, são: teste de equilíbrio genético de Hardy-Weinberg, teste de independência em tabelas de contingência e teste de homogeneidade de proporções.

A forma mais tradicional de realização do teste multinomial tem como procedimento básico o cálculo da estatística qui-quadrado ( $X^2$ ), apresentada em Pearson (1900). Tal estatística apresenta uma forma de medir a divergência encontrada entre os valores observados e os esperados, supondo que a hipótese testada é verdadeira. Esse procedimento apresenta vários problemas, principalmente quando aplicado em amostras de pequeno porte. Alguns desses problemas são: a baixa eficiência, viés e nível de significância não homogêneo.

Ao longo dos últimos cem anos, muitos outros procedimentos foram sugeridos para aplicações em testes multinomiais, porém, devido à não-existência de um teste multinomial uniformemente mais poderoso, vários desses procedimentos são

tomados como melhores, sob princípios extremamente questionáveis. Alguns são analisados no presente trabalho.

Muitos desses procedimentos, empregados nos testes multinomiais, são baseados no cálculo de estatísticas, empregadas como medida de divergência em relação ao valor observado e um valor paramétrico de referência, que em geral é o valor esperado, comumente estimado pelo processo de máxima verossimilhança, tomando-se a hipótese em teste como verdadeira. Quanto maior essa divergência, maior será a descrença na veracidade da hipótese nula. Tais procedimentos apresentam problemas semelhantes aos encontrados na utilização da estatística  $X^2$ .

Devido às deficiências da estatística  $X^2$ , bem como dos procedimentos similares, quando na aplicação em amostras de pequeno porte, existe a necessidade de um procedimento que diminua essas deficiências.

Um procedimento alternativo, denotado por condicional, é apresentado em Fisher (1935), para o teste de independência em tabelas de contingência. Tal procedimento é abordado em Tocher (1950), que amplia sua utilização também para teste de homogeneidade de proporções. Nesse último trabalho é destacado que esses procedimentos condicionais necessitam de aleatorizações. Com esse procedimento aleatorizado, podem ser obtidas várias propriedades, que não são encontradas em testes baseados no simples cálculo de estatísticas, tais como: nível de significância homogêneo e não-viés. Porém, tal procedimento aleatorizado é visto como de difícil aplicação prática.

Um caso particular do teste multinomial, cuja aplicação da estatística  $X^2$  possui deficiências, quando empregado em amostras de pequeno porte, é encontrado no teste de equilíbrio genético de Hardy-Weinberg. Tal teste está baseado na observação das frequências genotípicas de um grupo de indivíduos, em que, estando a população em equilíbrio genético de Hardy-Weinberg, essas frequências são completamente definidas por intermédio das frequências alélicas.

Muitos trabalhos, tais como: Levene (1949), Haldane (1954), Chapco (1976), Elston & Forthofer (1977), Meulepas (1999) e Montoya-Delgado et al. (2001), foram apresentados com formas alternativas à utilização da estatística  $X^2$ , para testar o

equilíbrio genético de Hardy-Weinberg, porém não substituem por completo a utilização dessa estatística, pois contêm ainda algumas deficiências análogas.

No presente trabalho é mostrado que, quando na construção de um discriminante no espaço amostral em pontos de rejeição e pontos de não-rejeição, com um nível de significância preestabelecido sem a utilização de aleatorizações, os procedimentos apresentados como alternativos possuem regiões de rejeição muito semelhantes às encontradas na aplicação da estatística  $X^2$ . Então, muitos dos problemas apresentados pelos procedimentos que utilizam o simples cálculo de estatísticas, quando na aplicação de pequenas amostras, não são corrigidos.

Assim, o objetivo do presente trabalho é a busca por uma forma alternativa de procedimento, para aplicação no teste de equilíbrio genético de Hardy-Weinberg, em amostras de pequeno porte. Essa forma alternativa deve contemplar uma nova propriedade, que está definida no presente trabalho: máximo volume da função poder.

Na busca pelo teste com máximo volume da função poder serão utilizados princípios de otimização inteira, como melhor solução prática viável em tal aplicação. Essa solução combina algumas das vantagens da aplicação do simples cálculo de estatísticas, que possuem mecanismos de decisão não aleatorizados, com soluções aleatorizadas, que proporcionam algumas propriedades desejáveis a um teste, tais como homogeneidade do nível de significância e não-viés, compondo um mecanismo que apresenta a maior quantidade de pontos não aleatorizados na região de rejeição, gerando uma maior combinação de composições genotípicas, associadas ao não-equilíbrio, qualquer que seja o nível de significância preestabelecido.

## 2 REVISÃO DA LITERATURA

### 2.1 Testes de hipóteses

Segundo Lehmann (1959), um teste de hipóteses é um procedimento de decisão entre a veracidade ou não de uma questionada hipótese. Tal decisão está baseada na observação de uma variável aleatória  $X$ , cuja distribuição  $P_\theta$  pertence a uma família de distribuições  $\mathfrak{P} = \{ P_\theta : \theta \in \Theta \}$ . À hipótese em teste está associado um subconjunto dos possíveis valores de  $\theta$ ,  $\Theta_0 \subset \Theta$ .

Um procedimento de teste é dito não aleatorizado se, para cada possível valor da variável aleatória  $X$ , uma das duas decisões é tomada: rejeita-se ou não a hipótese em teste. A eficiência do teste é dada pela capacidade, medido por intermédio de probabilidades, de não realizar uma decisão incorreta: rejeitar a hipótese, quando esta é verdadeira ou não rejeitá-la, quando é falsa. Devido basicamente à restrição do tamanho amostral, é impossível controlar simultaneamente os dois possíveis erros de decisão. Assim, limita-se a incorreta rejeição da hipótese em teste a um pequeno valor preestabelecido  $\alpha$ , que é denotado por nível de significância do teste, buscando-se por um mínimo possível de erro da decisão, de não rejeição de tal hipótese, quando esta é falsa.

Ainda segundo Lehmann (1959), em uma situação cujo espaço paramétrico ( $\Theta$ ) possui somente dois elementos, um caracterizando a hipótese em teste e outro sua negação, é aplicável o lema fundamental de Neyman-Pearson, que consiste em construir um procedimento de decisão que relacione a razão das duas possíveis probabilidades, sob consideração da hipótese em teste como verdadeira e sob

consideração da hipótese em teste como falsa. Aos pontos de maior razão, associa-se a decisão de não-rejeição da hipótese em teste, e aos pontos de menor razão a decisão de rejeitar tal hipótese. Esse procedimento, sendo limitado o número de pontos de rejeição da hipótese em teste, para a obtenção de um específico nível de significância, é o de melhor eficiência.

No caso mais geral o conjunto  $\Theta$  possui bem mais que dois elementos e, de modo análogo, os subconjuntos associados à hipótese em teste. Assim, permanecendo o controle de rejeição errônea da hipótese em teste, busca-se por procedimentos de decisão que apresentem valores razoáveis de probabilidade para rejeitar a hipótese em teste quando esta realmente é falsa.

## 2.2 Testes multinomiais

Em ensaios multinomiais, cujos resultados são definidos em categorias disjuntas, em que se supõe que as probabilidades de ocorrência de cada categoria são fixas e independentes, o número amostral de observações em cada categoria é uma estatística suficiente e completa para tais probabilidades.

Essa estatística  $\mathbf{X} = (X_1, X_2, \dots, X_k)$ , sendo  $X_i$ ,  $i=1,2,\dots,k$ , o número de observações relacionado à  $i$ -ésima categoria, é um vetor aleatório com distribuição multinomial de parâmetros  $n$ , correspondendo ao número total de observações e  $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_k)$ , em que  $\theta_i$ ,  $i=1,2,\dots,k$ , representa a probabilidade de ser verificada a  $i$ -ésima categoria como resposta, em uma observação qualquer. O espaço paramétrico ( $\Theta$ ) e espaço amostral ( $\Omega$ ), relacionados a essa distribuição são, respectivamente, dados por:

$$\Theta = \left\{ (\theta_1, \theta_2, \dots, \theta_k) : \theta_i \in [0,1], i = 1,2,\dots,k \text{ e } \sum_{i=1}^k \theta_i = 1 \right\},$$

$$\Omega = \left\{ (x_1, x_2, \dots, x_k) : x_i \in \{0,1,\dots,n\}, i = 1,2,\dots,k \text{ e } \sum_{i=1}^k x_i = n \right\}.$$

A mais antiga e tradicional estatística utilizada em testes multinomiais é a  $X^2$ , originalmente citada em Pearson (1900). Tal estatística é dada por:

$$X^2 = \sum_{i=1}^k \frac{(X_i - E_i)^2}{E_i}, \quad (1)$$

em que  $X_i$  e  $E_i$  são, o número de observações e o valor esperado na  $i$ -ésima categoria. Havendo necessidade de estimações, para o cálculo dos valores esperados, estes são obtidos de forma a ser encontrado o mínimo valor para  $X^2$ , sob hipótese de interesse considerada como verdadeira. Esse critério de estimação é utilizado para proteger a hipótese em teste de uma falsa rejeição, já que a hipótese em teste é rejeitada para grandes valores de  $X^2$ . Tal estatística, quando a hipótese em teste não necessita de estimações, possui distribuição assintótica qui-quadrado com  $k-1$  graus de liberdade, sendo mantido, em Pearson (1900), o mesmo número de graus de liberdade nos casos em que existe necessidade de estimações. Essa manutenção acarreta a superestimação do nível de significância associado ao teste.

Uma aplicação da estatística  $X^2$ , em que se faz necessária a estimação do valor esperado sob consideração da hipótese em teste como verdadeira, pode ser encontrada no teste de independência em tabelas de contingência. Para tal aplicação, é discutido em trabalhos tais como Fisher (1922, 1924), o método de estimação bem como a alteração no parâmetro, graus de liberdade, da distribuição qui-quadrado para a qual converge a estatística  $X^2$ .

Em relação ao método de estimação, é mostrado em Fisher (1924) que qualquer método consistente satisfaz a convergência assintótica, sugerindo-se o critério de máxima verossimilhança como um bom substituto do critério de mínimo valor de  $X^2$ , que em muitos casos é de difícil cálculo. Na correção do número de graus de liberdade, é mostrado em Fisher (1922) que a necessidade de estimação reduz o número de graus de liberdade. Em uma tabela de contingência contendo  $r$  linhas e  $c$  colunas são sugeridos  $(c-1)(r-1)$  graus de liberdade. Tal mudança é formalmente provada em Cramér (1946).

Alternativamente à utilização da estatística  $X^2$ , que tendo distribuição assintótica, gera o nível de significância apenas aproximado ao que é pretendido, é apresentado em Fisher (1935) um mecanismo que gera o nível de significância exato, para testes de independência em tabelas de contingência. Nesse trabalho, os valores observados na tabela de contingência são comparados com todos os outros possíveis valores, que possuem iguais valores marginais. O nível de significância, associado à tabela observada, é dado pelo acúmulo das probabilidades associadas às tabelas de idêntica composição marginal, cujas probabilidades são iguais ou inferiores à da tabela observada, em que a distribuição marginal é tomada como condicionante. Devido a essa forma de cálculo do nível de significância, tal teste é classificado como condicional.

Um questionamento em relação ao mecanismo exato proposto em Fisher (1935) é encontrado em Tocher (1950). Nesse trabalho é citado que testes em variáveis discretas necessitam de aleatorizações, para que sejam obtidos níveis de significância exatos, sendo sugerido um algoritmo capaz de tornar o teste em tabelas de contingência exato, qualquer que seja o nível de significância preestabelecido. Tal algoritmo utiliza como variável condicionante a distribuição marginal. Essa variável, associada a tal distribuição, é suficiente em relação à região do espaço paramétrico relacionado à hipótese de independência. Essa variável, que pode ser denotada por estatística suficiente, é utilizada para construir uma partição do espaço amostral, e nesta, aplicar-se um procedimento aleatorizado que torna o teste com nível de significância exato, para qualquer valor paramétrico pertencente à região de independência. Além do teste de independência, também é citado em Tocher (1950), um teste exato, não enviesado, para testar homogeneidade em distribuições binomiais.

A aleatorização da decisão em testes discretos é vista em muitos trabalhos, tais como Lancaster (1961), como necessária para serem obtidos níveis de significância exatos. Nesse trabalho, a aleatorização está associado a um mecanismo auxiliar de decisão.

Em Cochran (1952), pode ser encontrada uma extensa discussão na utilização da estatística  $X^2$ , em diversas aplicações, sendo observado que a eficiência de tal estatística, em casos de necessidade de estimação de parâmetros, depende do método

de estimação, sendo considerado, como critério natural de estimação, o que gera o mínimo valor para a estatística  $X^2$ . Tal critério de estimação é atualmente esquecido em muitos trabalhos que utilizam essa estatística, sendo adotado, quase sempre, o critério de estimação por máxima verossimilhança. Em Cochran (1952), é sugerida a estatística do logaritmo da razão de verossimilhança,  $G^2$ , dada por:

$$G^2 = \sum_{i=1}^k X_i \text{Log} \left( \frac{X_i}{E_i} \right), \quad (2)$$

como melhor do que  $X^2$  em diversas situações, tais como no caso de testes multinomiais em amostras de pequeno porte.

Em Cressie & Read (1984) e mais detalhadamente em Read & Cressie (1988) é apresentada uma medida de divergência direcionada, que pode ser utilizada para mensurar distâncias entre pontos do espaço paramétrico multinomial. Tal medida, sendo adaptada ao espaço amostral, é utilizada para mensurar divergências entre valores observados e esperados, sob consideração de alguma hipótese de interesse, gerando, dessa forma, uma família de estatísticas com convergência assintótica qui-quadrado. Nessa família de estatísticas encontra-se presente, como casos particulares, muitas das estatísticas utilizadas em testes multinomiais. Tal família de estatísticas, parametrizada por  $\lambda$ , é dada por:

$$2nI^\lambda = \frac{2}{\lambda(\lambda+1)} \sum_{i=1}^k \left\{ X_i \left[ \left( \frac{X_i}{E_i} \right)^\lambda - 1 \right] \right\}. \quad (3)$$

Os casos particulares mais conhecidos são: para  $\lambda = 1$ , tem-se a estatística  $X^2$  e, definindo-se por continuidade quando  $\lambda \rightarrow 0$ , tem-se a estatística  $G^2$ . Outras estatísticas também são apresentadas como casos particulares, destacando-se: a estatística de Freeman-Tukey ( $\lambda = 1/2$ ), a estatística  $X^2$  modificada de Neyman ( $\lambda = -2$ )



e a estatística da razão do logaritmo da verossimilhança, modificada ( $\lambda \rightarrow -1$ ). Tais estatísticas são mostradas na Tabela 1.

Tabela 1. Casos particulares da família de estatísticas definida em Cressie & Read (1984) e Read e Cressie (1988)

$\lambda = 1/2$	$\lambda = -2$	$\lambda \rightarrow -1$
$T^2 = 4 \sum_{i=1}^k (\sqrt{X_i} - \sqrt{E_i})^2$	$NM^2 = \sum_{i=1}^k \frac{(X_i - E_i)^2}{X_i}$	$GM^2 = 2 \sum_{i=1}^k E_i \text{Log} \left( \frac{E_i}{X_i} \right)$

Quando se faz necessária a estimação dos valores esperados, o critério do mínimo valor para a estatística em questão é apresentado como natural, sendo, porém, admitido o critério da máxima verossimilhança no caso de grandes amostras. Nos referidos trabalhos, o caso particular  $\lambda = 2/3$  é apresentado como sugestão da melhor estatística para testes multinomiais.

Muitos questionamentos, em relação ao nível de significância pretendido e o realmente encontrado, na aplicação de estatísticas com distribuição assintótica qui-quadrado foram realizados, sendo sugeridos alguns procedimentos que tornam o nível de significância do teste com valores exatos, para aplicação em amostras de pequeno porte.

Em Gurian et al (1964) e, posteriormente, em Tate & Hyer (1973), um mecanismo de ordenação do espaço amostral é sugerido para testes multinomiais com hipótese em teste simples e hipótese alternativa complementar do espaço paramétrico: as probabilidades dos pontos pertencentes ao espaço amostral são dispostas em ordem de magnitude, tomando-se como verdadeira a hipótese em teste, em que o nível de significância associado ao valor observado ( $p$ -valor) é dado pelo acúmulo das probabilidades iguais ou inferiores à probabilidade associada ao valor observado.

Em Radlow & Alf (1975), para aplicação em testes semelhantes ao apresentado anteriormente, a ordenação do espaço amostral é construída através da distância  $X^2$  tomada como divergência entre o ponto observado e o valor esperado, tomando-se como verdadeira a hipótese em teste, em que o nível de significância

associado ao valor observado ( $p$ -valor) é dado pelo acúmulo das probabilidades dos pontos que possuem divergência igual ou superior à divergência associada ao valor observado.

Em Chapman (1976), uma comparação entre a ordenação do espaço amostral utilizando a estatística  $X^2$  e  $G^2$ , como medida de divergência entre os valores observados e esperados sob equi-probabilidade, mostra que a menor distância entre o tamanho exato do teste e o sugerido por tais estatísticas é obtida com a estatística  $X^2$ , sendo o teste exato obtido com a ordenação sugerida em Tate & Hyer (1973).

A busca por uma ordenação do espaço amostral, sob hipótese em teste, em relação ao que pode ser observado, bem como a aleatorização da decisão é vista em Tocher (1950) como sendo a extensão necessária para a aplicação da teoria de Neyman-Pearson em testes com a utilização de variáveis discretas.

### 2.3 Equilíbrio genético

Como pode ser visto em Weir (1996), em uma população de indivíduos, sendo observado um loco, cujos genes são formados por dois únicos alelos “A” e “a”, as frequências genotípicas “AA”, “Aa” e “aa”, quando em equilíbrio genético de Hardy-Weinberg, são respectivamente dadas por:  $\theta^2$ ,  $2\theta(1-\theta)$  e  $(1-\theta)^2$ , em que  $\theta \in [0, 1]$  representa a frequência do alelo “A”, e o complementar,  $1-\theta$ , a frequência do alelo “a”.

Associando-se  $X = (X_1, X_2, X_3)$  às frequências genotípicas, “AA”, “Aa” e “aa”, respectivamente, obtidas em uma amostra com  $n$  indivíduos selecionados aleatoriamente, tem-se o vetor aleatório  $X$ , em questão, com distribuição multinomial de parâmetros  $n$  e  $\theta = (\theta_1, \theta_2, \theta_3)$ , com espaço paramétrico dado por:

$$\Theta = \{(\theta_1, \theta_2, \theta_3): \theta_i \in [0,1] \text{ e } \theta_1 + \theta_2 + \theta_3 = 1\},$$

em que o subconjunto do espaço paramétrico, associado ao equilíbrio genético de Hardy-Weinberg, é dado por:

$$\Theta_0 = \{(\theta_1, \theta_2, \theta_3) \in \Theta : 4\theta_1\theta_3 = (\theta_2)^2\}.$$

Como no caso do teste de independência em tabelas de contingência, a estatística  $X^2$  é bastante utilizada, tendo, no caso particular de dois alelos, convergência assintótica qui-quadrado com 1 grau de liberdade. Porém, devido à baixa eficiência dessa estatística no caso de pequenas amostras, bem como outras estatísticas alternativas com convergência assintótica qui-quadrado, mecanismos alternativos foram apresentados.

Em Levene (1949), é apresentada a distribuição exata do número de homozigotos, condicionada à distribuição alélica, em um loco cujos genes são formados por múltiplos alelos. A estatística de teste, definida nesse trabalho, associa a distância entre o número de homozigotos encontrado na amostra e o número de homozigotos esperado sob hipótese de equilíbrio genético. Mostra-se que tal estatística possui distribuição assintótica normal.

Em Haldane (1954), é apresentada a distribuição do número de heterozigotos para aplicação a locos com somente dois alelos. A distribuição alélica é tomada como variável condicionante, observando-se, então, a distribuição do número de heterozigotos, supostamente tomada como recombinações das quantidades alélicas observadas. Nesse trabalho, é apresentada uma estatística de teste baseada na diferença entre o quadrado do número de heterozigotos e quatro vezes o produto do número de homozigotos, ajustado para que o valor esperado dessa distância seja nulo.

A distribuição do número de heterozigotos, condicionada à distribuição alélica, está generalizada para alelos múltiplos em Chapco (1976), em que é apresentado um procedimento de cálculo exato do nível de significância, para testar o equilíbrio genético. Tal procedimento é visto em Elston & Forthofer (1977) como análogo ao proposto em Fisher (1935) para testes de independência em tabelas de contingência. Em Elston & Forthofer (1977) são comparados o teste, cujo procedimento gera o nível de

significância exato, que doravante será tratado por teste exato, e quatro estatísticas com convergência assintótica qui-quadrado. A comparação é realizada por intermédio do cálculo do nível de significância para 20 possíveis valores observados em 15 possíveis tamanhos de amostra, concluindo-se que os níveis de significância apresentados nos testes com a utilização das estatísticas eram razoavelmente diferentes dos níveis de significância apresentados no teste exato.

Nos trabalhos de Louis & Dempster (1987) e Guo & Thompson (1992), são criados algoritmos para o cálculo do nível de significância do teste exato, utilizando a distribuição do número de heterozigotos, condicionada à distribuição alélica. Nesses trabalhos, tal procedimento exato é visto como inquestionavelmente superior a quaisquer outros, e, assim como em Elston & Forthofer (1977) e Chapco (1976), não é mencionado qualquer mecanismo de aleatorização para compor um teste exato, o que contraria o exposto em Tocher (1950) para testes discretos.

Alguns critérios para comparação de testes são citados em Emigh (1980), que em ordem de importância são: poder do teste, sensibilidade ao nível de significância e convergência assintótica. Nesse trabalho o teste exato é comparado com alguns testes que utilizam estatísticas com convergência assintótica qui-quadrado, em que se conclui serem essencialmente semelhantes na ordenação dos possíveis valores observados, em relação à rejeição da hipótese de equilíbrio, com exceção para certas específicas frequências genotípicas.

Em Maiste & Weir (1995), o teste de equilíbrio genético é estendido para dois locos com múltiplos alelos, em que são comparados o teste exato, um outro teste multinomial exato e três estatísticas:  $X^2$ ,  $G^2$  e a estatística  $2nI^{2/3}$  (Cressie & Read, 1984). Concluíram que os testes exatos são superiores aos que utilizam estatísticas com convergência assintótica, porém que, esses testes são semelhantes para grandes amostras.

Um procedimento de teste que utiliza conjuntamente a distância entre o número de heterozigotos observado na amostra e o esperado, sob o equilíbrio genético, bem como a probabilidade associada à obtenção de tal número de heterozigotos condicionada à distribuição alélica, é apresentado em Meulepas (1999). É mostrado, em

tal trabalho, que essa combinação proporciona uma minimização do viés encontrado na simples aplicação da distribuição exata do número de heterozigotos, condicionada à distribuição alélica, tal como é apresentado no teste exato.

Também, como forma alternativa do teste exato, é apresentado em Montoya-Delgado et al. (2001) uma forma de ordenação do espaço amostral baseado em uma razão da medida de verossimilhança em relação às hipóteses testadas. Para cada ponto do espaço amostral são calculadas a área de verossimilhança, correspondente à região do espaço paramétrico relacionado pela hipótese de equilíbrio genético e o volume de verossimilhança, correspondente à região do espaço paramétrico relacionado ao não equilíbrio genético. Os pontos do espaço amostral são ordenados pelo valor da razão entre tal área e volume de verossimilhança. Essa ordenação mostra-se superior no cálculo do nível de significância, em relação ao teste exato.

## **2.4 Otimização inteira**

Um caso particular na busca de pontos críticos em funções é bastante útil na aplicação em testes multinomiais, a otimização inteira ou combinatorial. Segundo Nemhauser & Wolsey (1988), a otimização inteira ou combinatorial é um caso particular na busca de máximos ou mínimos, em que as variáveis, sujeitas às restrições de equações ou inequações, somente podem apresentar valores inteiros como resultados. Busca-se, então, por valores inteiros, que cumpram as restrições a que o problema em questão esteja sujeito, e que ofereçam o melhor valor possível, que pode ser um máximo ou um mínimo, para uma certa função em análise.

O problema da mochila, caso particular de otimização combinatorial, é dado pela seguinte situação:  $n$  objetos úteis estão disponíveis para um aventureiro, que possui um única mochila para transportá-los em uma viagem. Para cada objeto existe um valor de utilidade bem como um valor de custo, que pode ser, por exemplo, o peso do objeto. Devido a uma restrição de peso, a mochila não comporta todos os objetos. O problema consiste na seleção de alguns dos objetos para alocar na mochila, de modo a

obter-se a melhor relação custo/benefício. Sejam  $u_i$  e  $c_i$ , respectivamente, a utilidade e o custo associados ao  $i$ -ésimo objeto.

Em Nemhauser & Wolsey (1988), a solução do problema é apresentada como a melhor combinação de objetos, que não ultrapasse a restrição de peso e que apresente a maior soma possível de utilidade. Tal solução é encontrada calculando-se a relação  $u_i/c_i$  para todos os objetos, ordenando-se tal relação e alocando-se na mochila, em ordem decrescente da relação até a obtenção do limite de peso da mochila.

Essa forma de solução tem analogia com a solução de problemas em testes de hipóteses discretos, como o teste multinomial. A restrição de peso é traduzida pelo nível de significância preestabelecido e a otimização buscada é o máximo poder para o teste. Assim, busca-se, dentro do espaço amostral, a melhor combinação de pontos, para compor a região de rejeição, que possua o máximo poder, porém não ultrapasse o nível de significância preestabelecido.

### 3 METODOLOGIA

#### 3.1 Definição do teste de máximo volume

Um teste de hipóteses  $\{ H_0, H_1 \}$ , em relação a uma variável aleatória  $X$  com distribuição  $P_\theta$  pertencente à família de distribuições  $\mathfrak{P} = \{ P_\theta : \theta \in \Theta \}$  e espaço amostral dado por  $\Omega$  (Lehmann, 1959), é especificado por uma função  $\varphi : \Omega \rightarrow [0, 1]$ , denotada por função crítica, que associa para cada ponto do espaço amostral a probabilidade deste pertencer à região de rejeição da hipótese  $H_0$ . Essas hipóteses estão associadas a uma partição do espaço paramétrico:

$$H_0 : \theta \in \Theta_0$$

$$H_1 : \theta \in \Theta_1, \quad \text{em que: } \Theta_0 \cup \Theta_1 = \Theta \text{ e } \Theta_0 \cap \Theta_1 = \emptyset.$$

Sendo as hipóteses tratadas de forma assimétrica, de tal modo que um erro de rejeição de  $H_1$  é melhor tolerado do que um erro de rejeição de  $H_0$ , a probabilidade de erro de rejeição de  $H_0$  é limitada a um pequeno valor  $\alpha$  ( $\alpha \in (0,1)$ ),

$$\sup_{\theta \in \Theta_0} E_\theta[\varphi(X)] \leq \alpha,$$

denotado por nível de significância do teste. Sendo  $\Theta_0$  um conjunto com mais de um elemento, pode ser definido a função probabilidade de rejeição de  $H_0$ , com domínio em  $\Theta_0$ , em que o valor máximo gera o nível de significância do teste.

Naturalmente, uma forma de julgar a eficiência do teste definido por  $\varphi(\cdot)$  é observar a capacidade de rejeitar  $H_0$  na região  $\Theta_1$ . Denota-se por poder do teste no ponto  $\theta \in \Theta_1$ , o valor:

$$\beta(\theta, \varphi) = E_{\theta}[\varphi(X)], \text{ em que } \theta \in \Theta_1.$$

Um teste  $\varphi'(\cdot)$  é considerado inadmissível, em relação a outro  $\varphi(\cdot)$ , ambos de mesmo nível de significância, se o poder do primeiro é sempre menor que o do segundo, isto é:

$$\beta(\theta, \varphi) > \beta(\theta, \varphi'), \forall \theta \in \Theta_1.$$

Existindo um teste  $\varphi^*(\cdot)$ , de tamanho  $\alpha$ , tal que, qualquer outro teste  $\varphi(\cdot)$  de tamanho  $\alpha$  é inadmissível em relação a  $\varphi^*(\cdot)$ ,  $\varphi^*(\cdot)$  é considerado o teste uniformemente mais poderoso de tamanho  $\alpha$  para testar as hipóteses  $\{ H_0, H_1 \}$ .

O teste uniformemente mais poderoso em casos multinomiais raramente existe. Sendo assim, outras propriedades de testes devem ser buscadas. A propriedade de não-viés é vista em Cohen e Sackrowitz (1992) como necessária para a admissibilidade de um teste. Tal propriedade define um valor mínimo de poder para um teste, dado pelo nível de significância, ou seja,

$$\beta(\theta, \varphi) \geq \alpha, \forall \theta \in \Theta_1.$$

No caso em que  $\Theta_1$  é limitado, pode-se definir uma propriedade necessária, porém não suficiente, para a existência do teste mais poderoso como uma qualidade desejável a um teste de hipóteses. Essa propriedade é a de máximo volume da função poder na região correspondente à hipótese alternativa.



Um teste  $\varphi^*(\cdot)$ , de tamanho  $\alpha$ , é considerado o teste de maior volume da função poder, dentre os testes de tamanho  $\alpha$ , se:

$$\int_{\Theta_1} E_{\theta}[\varphi^*(X)]d\theta \geq \int_{\Theta_1} E_{\theta}[\varphi(X)]d\theta, \forall \varphi : \sup_{\theta \in \Theta_0} E_{\theta}[\varphi(X)] \leq \alpha.$$

**Teorema 1.** O teste uniformemente mais poderoso apresenta o máximo volume de poder.

Prova: seja  $\varphi^*(\cdot)$  o teste uniformemente mais poderoso de tamanho  $\alpha$ , então:

$$E_{\theta}[\varphi^*(X)] \geq E_{\theta}[\varphi(X)], \forall \theta \in \Theta_1, \text{ em que } \varphi : E_{\theta}[\varphi(X)] \leq \alpha, \forall \theta \in \Theta_0$$

$$\text{assim, } \int_{\Theta_1} E_{\theta}[\varphi^*(X)]d\theta \geq \int_{\Theta_1} E_{\theta}[\varphi(X)]d\theta \blacksquare$$

### 3.2 Aplicação do máximo volume no teste de equilíbrio genético

Seja  $(X_1, X_2, X_3)$  o vetor aleatório associado ao número de indivíduos, selecionados aleatoriamente de uma específica população, portador dos genes “AA”, “Aa” e “aa”, respectivamente. Sendo  $n$  o tamanho amostral, o vetor aleatório em questão possui distribuição multinomial de parâmetros  $n$  e  $\theta = (\theta_1, \theta_2, \theta_3)$ , em que os espaços paramétrico e amostral são dados, respectivamente, por:

$$\Theta = \{(\theta_1, \theta_2, \theta_3) : \theta_i \in [0,1], i = 1,2,3 \text{ e } \theta_1 + \theta_2 + \theta_3 = 1\},$$

$$\Omega = \{(x_1, x_2, x_3) : x_i \in \{0,1,\dots, n\}, i = 1,2,3 \text{ e } x_1 + x_2 + x_3 = n\}.$$

O teste de equilíbrio genético de Hardy-Weinberg está associado às seguintes hipóteses:

$$H_0 : \theta \in \Theta_0 = \{(\theta_1, \theta_2, \theta_3) \in \Theta : 4\theta_1\theta_2 = (\theta_2)^2\},$$

$$H_1 : \theta \in \Theta_1 = \{(\theta_1, \theta_2, \theta_3) \in \Theta : 4\theta_1\theta_2 \neq (\theta_2)^2\}.$$

A função poder, portanto é dada por:

$$\beta(\boldsymbol{\theta}, \varphi) = E_{\boldsymbol{\theta}}[\varphi(X_1, X_2, X_3)] = \sum_{\Omega} \varphi(x_1, x_2, x_3) P_{\boldsymbol{\theta}}[X_1 = x_1, X_2 = x_2, X_3 = x_3],$$

tem-se, então, os seguintes teoremas associados ao teste de máximo volume para o equilíbrio genético de Hardy-Weinberg.

**Teorema 2.** O volume da função poder no teste de equilíbrio genético de Hardy-Weinberg, em relação ao espaço paramétrico associado a hipótese alternativa  $H_1$  ( $\Theta_1$ ) é o mesmo se associado a todo o espaço paramétrico  $\Theta$ .

Prova: em  $\Theta_0$ :  $\theta_1 = \theta^2$ ,  $\theta_2 = 2\theta(1-\theta)$  e  $\theta_3 = (1-\theta)^2$ , em que  $\theta \in [0,1]$ . Reparametrizando-se em função de  $\theta_1$ , tem-se:

$$\theta_2 = 2(\theta_1^{1/2} - \theta_1) \text{ e } \theta_3 = (1 - \theta_1^{1/2})^2, \text{ em que } \theta_1 \in [0,1].$$

Seja  $I_A(\boldsymbol{\theta})$  a função indicador, portanto:  $I_A(\boldsymbol{\theta}) = 1$ , se  $\boldsymbol{\theta} \in A$  e  $I_A(\boldsymbol{\theta}) = 0$ , se  $\boldsymbol{\theta} \notin A$ .

Calculando-se o volume da função poder na região associada a  $H_0$  ( $\Theta_0$ ), associando-se o valor máximo da função poder ( $E_{\boldsymbol{\theta}}[\varphi(\mathbf{X})] = 1$ ), tem-se:

$$\int_{\Theta} I(\boldsymbol{\theta}) E_{\boldsymbol{\theta}}[\varphi(X_1, X_2, X_3)] d\boldsymbol{\theta} = \int_{\Theta_0} I(\boldsymbol{\theta}) 1 d\boldsymbol{\theta} = \int_0^1 \int_{2(\theta_1^{1/2} - \theta_1)}^{2(\theta_1^{1/2} - \theta_1)} 1 d\theta_2 d\theta_1 = \int_0^1 0 d\theta_1 = 0.$$

Sendo o volume da função poder em  $\Theta_0$ , mesmo associando-se o máximo da função poder, igual a zero, o volume da função poder em  $\Theta_1$  é o mesmo que em  $\Theta$  ■

**Teorema 3.** A contribuição pontual de poder é fixa, para todos os pontos do espaço amostral no teste de equilíbrio genético de Hardy-Weinberg.

Prova: sendo a contribuição de volume de poder o volume proporcionado por cada ponto, esse valor é dado por:

$$\int_{\Theta_1} P_{\boldsymbol{\theta}}[X_1 = x_1, X_2 = x_2, X_3 = x_3] d\boldsymbol{\theta},$$

que pelo teorema 2 é igual a:

$$\begin{aligned}
& \int_{\Theta} P_{\theta}[X_1 = x_1, X_2 = x_2, X_3 = x_3] \partial \theta = \\
& = \int_0^1 \int_0^{1-\theta_1} \binom{n}{x_1 x_2 x_3} \theta_1^{x_1} \theta_2^{x_2} (1 - \theta_1 - \theta_2)^{x_3} \partial \theta_2 \partial \theta_1 = \\
& = \frac{n!}{x_1! x_2! x_3!} \int_0^1 \int_0^{1-\theta_1} \theta_1^{x_1} \theta_2^{x_2} (1 - \theta_1 - \theta_2)^{x_3} \partial \theta_2 \partial \theta_1
\end{aligned}$$

fazendo-se:  $\gamma = \theta_2 / (1 - \theta_1)$ , tem-se

$$\begin{aligned}
& \frac{n!}{x_1! x_2! x_3!} \int_0^1 \theta_1^{x_1} \int_0^1 [\gamma(1 - \theta_1)]^{x_2} [(1 - \gamma)(1 - \theta_1)]^{x_3} (1 - \theta_1) \partial \gamma \partial \theta_1 = \\
& = \frac{n!}{x_1! x_2! x_3!} \int_0^1 \theta_1^{x_1} (1 - \theta_1)^{x_2 + x_3 + 1} \partial \theta_1 \int_0^1 \gamma^{x_2} (1 - \gamma)^{x_3} \partial \gamma = \\
& = \frac{n!}{x_1! x_2! x_3!} \frac{x_2! x_3!}{(x_2 + x_3 + 1)!} \frac{x_1! (x_2 + x_3 + 1)!}{(x_1 + x_2 + x_3 + 2)!} = \frac{n!}{(n + 2)!} \blacksquare
\end{aligned}$$

Sendo a contribuição de volume de poder do ponto  $(x_1, x_2, x_3)$  um valor que depende somente de  $n$ , todos os pontos do espaço amostral têm a mesma contribuição de volume. Desse modo, o teste para o equilíbrio genético de Hardy-Weinberg com maior volume de poder é aquele que contém a maior quantidade de pontos na região de rejeição.

### 3.3 Teste de equilíbrio genético com máximo número de pontos na região de rejeição

O teste de equilíbrio genético de Hardy-Weinberg com máximo número de pontos na região de rejeição, com um nível de significância preestabelecido  $\alpha$ , está associado a uma função de decisão  $\varphi(\cdot)$  que satisfaz a seguinte otimização:

$$\begin{aligned}
& \sum_{\Omega} \varphi(x_1, x_2, x_3), \text{ é máximo e,} \\
& \sum_{\Omega} \varphi(x_1, x_2, x_3) P_{\theta}[X_1 = x_1, X_2 = x_2, X_3 = x_3] \leq \alpha, \forall \theta \in \Theta_0.
\end{aligned}$$

Em Tocher (1950), para aplicação do teste de independência em tabelas de contingência, a restrição do nível de significância do teste é mantido fixo, qualquer que seja o valor do parâmetro pertencente à região relacionada à hipótese de independência, por intermédio de otimizações em subconjuntos do espaço amostral, de tal modo que a probabilidade de rejeição não dependa do real valor do parâmetro.

Para o equilíbrio genético de Hardy-Weinberg, esses subconjuntos do espaço amostral podem ser definidos por intermédio da distribuição genotípica condicionada pela distribuição alélica, como é visto em Haldane (1954). Para tanto, sejam  $\Omega_i$ 's os subconjuntos do espaço amostral com igual distribuição alélica, ou seja, com igual número de alelos "A". Tem-se:

$$\Omega_i = \{(x_1, x_2, x_3) : 2x_1 + x_2 = i\}, i \in \{0, 1, \dots, 2n\}.$$

**Teorema 4.** O teste de equilíbrio genético de Hardy-Weinberg, com nível de significância pré-estabelecido  $\alpha$  para cada subconjunto do espaço amostral com igual distribuição alélica, apresenta nível de significância  $\alpha$ , qualquer que seja o valor do parâmetro associado ao equilíbrio  $((\theta_1, \theta_2, \theta_3) \in \Theta_0)$ .

Prova: como pode ser visto em Elston & Forthofer (1977), a distribuição genotípica, condicionada à distribuição alélica, independe do valor paramétrico sob equilíbrio genético, portanto:

$$P_{\theta}[X_1 = x_1, X_2 = x_2, X_3 = x_3 / 2X_1 + X_2 = 2x_1 + x_2] = c, \forall \theta \in \Theta_0,$$

seja  $\varphi(\cdot)$  definido, para cada subconjunto amostral  $\Omega_i$ , de tal modo que:

$$\sum_{\Omega_i} \varphi(x_1, x_2, x_3) P_{\theta}[X_1 = x_1, X_2 = x_2, X_3 = x_3 / 2X_1 + X_2 = 2x_1 + x_2] = \alpha, \forall \theta \in \Theta_0,$$

então o nível de significância definido pelo teste  $\varphi(\cdot)$  será dado por:

$$\begin{aligned} E_{\theta}[\varphi(\mathbf{X})] &= \sum_{\Omega} \varphi(x_1, x_2, x_3) P_{\theta}[X_1 = x_1, X_2 = x_2, X_3 = x_3] = \\ &= \sum_i \sum_{\Omega_i} \varphi(x_1, x_2, x_3) P_{\theta}[X_1 = x_1, X_2 = x_2, X_3 = x_3] = \\ &= \sum_i \sum_{\Omega_i} \varphi(x_1, x_2, x_3) P_{\theta}[2X_1 + X_2 = i] P_{\theta}[X_1 = x_1, X_2 = x_2, X_3 = x_3 / 2X_1 + X_2 = i] = \\ &= \sum_i P_{\theta}[2X_1 + X_2 = i] \sum_{\Omega_i} \varphi(x_1, x_2, x_3) P_{\theta}[X_1 = x_1, X_2 = x_2, X_3 = x_3 / 2X_1 + X_2 = i] = \\ &= \sum_i P_{\theta}[2X_1 + X_2 = i] \alpha = \alpha \sum_i P_{\theta}[2X_1 + X_2 = i] = \alpha \blacksquare \end{aligned}$$

Definindo-se  $\varphi(\cdot)$  de tal modo a conter o máximo número de pontos na região de rejeição para cada subconjunto do espaço amostral  $\Omega_i$ , por intermédio da ordenação de pontos com menor probabilidade condicional, em que se faz necessária a aleatorização para ser mantido o nível de significância preestabelecido  $\alpha$ , pode ser definido então o teste  $\varphi^*(\cdot)$  que, dentro da classe dos testes com uniformidade do nível de significância  $\alpha$ , contém a maior quantidade de pontos na região de rejeição.

O teste  $\varphi^*(\cdot)$  é definido então por intermédio da adaptação do algoritmo apresentado em Tocher (1950), para o teste de independência em tabelas de contingência, no caso do equilíbrio genético de Hary-Weinberg. Tal adaptação, para um teste com nível de significância preestabelecido  $\alpha$ , é dada por:

- i ) para cada subconjunto amostral da distribuição genotípica, com valores alélicos iguais ( $\Omega_i$ ) são calculadas as probabilidades genotípicas condicionada à distribuição alélica;
- ii ) os pontos são ordenados, por intermédio das probabilidades condicionais, para cada subconjunto amostral  $\Omega_i$ ;
- iii ) associa-se valor  $\varphi^*(\mathbf{x}) = 1$  para os valores correspondentes às menores probabilidades condicionas até a obtenção do valor limite  $\alpha$  de modo que, com a inclusão do próximo ponto da ordenação, o nível de significância ultrapasse o valor  $\alpha$ ;
- iv ) para esse ponto limite  $\mathbf{x}_i' = (x_{i,1}, x_{i,2}, x_{i,3})$  associa-se o valor aleatorizado  $\gamma_i$  e o restante dos pontos do subconjunto valor  $\varphi(\mathbf{x}) = 0$ , de modo a ser obtido, para cada subconjunto amostral um somatório acumulado  $\alpha$ , em que  $\gamma_i$  é dado por:

$$\gamma_i = \frac{\alpha - \sum_{\Omega_i; \varphi(x_1, x_2, x_3)=1} P[X_1 = x_1, X_2 = x_2, X_3 = x_3 / 2X_1 + X_2 = i]}{P[X_1 = x'_{i,1}, X_2 = x'_{i,2}, X_3 = x'_{i,3} / 2X_1 + X_2 = i]}.$$

Tal procedimento confunde-se com o teste exato apresentado em Elston & Forthofer (1977), quando é calculado o nível de significância associado a um ponto específico do espaço amostral. Porém, sem a aleatorização, o mesmo nível de

significância não é mantido, quando analisados outros pontos do espaço amostral, pois estes apresentam diferentes níveis de significância, quando calculados com o algoritmo apresentado no referido trabalho.

Sendo assim, o procedimento para testar o equilíbrio genético de Hardy-Weinberg, citado como adaptação do algoritmo apresentado em Tocher (1950), para testar independência em tabelas de contingência, é o que apresenta melhor solução para a inclusão da maior quantidade de pontos na região de rejeição, dentre os procedimentos que possuem a uniformidade do nível de significância.

### 3.4 Definição do teste de equilíbrio genético não enviesado

Um teste é dito ser não enviesado quando o poder, aplicado em qualquer ponto pertencente à região do espaço paramétrico relacionado com a hipótese alternativa, é igual ou superior ao nível de significância associado ao teste. No caso de uma distribuição binomial com hipótese nula simples tal propriedade é facilmente verificada.

Seja  $X$  uma variável aleatória com distribuição binomial de parâmetros  $n$  e  $\theta$ , em que  $n$  é um número inteiro positivo e  $\theta \in [0,1]$ . Deseja-se testar a hipótese de  $\theta$  ser um certo específico valor  $\theta_0$ .

$$H_0 : \theta = \theta_0$$

$$H_1 : \theta \neq \theta_0.$$

O espaço paramétrico e o espaço amostral do experimento são dados respectivamente por:

$$\Theta = [0, 1]$$

$$\Omega = \{0, 1, \dots, n\}.$$

Seja  $\varphi(\cdot)$  a função crítica do teste. O nível de significância associada ao teste e a função poder são dados respectivamente por:

$$E_{\theta_0}[\varphi(X)] = \sum_{x=0}^n \varphi(x) P_{\theta_0}[X = x] = \sum_{x=0}^n \varphi(x) \binom{n}{x} \theta_0^x (1 - \theta_0)^{n-x},$$

$$\beta(\theta, \varphi) = E_{\theta}[\varphi(X)] = \sum_{x=0}^n \varphi(x) P_{\theta}[X = x] = \sum_{x=0}^n \varphi(x) \binom{n}{x} \theta^x (1 - \theta)^{n-x}.$$

Um teste, com nível de significância preestabelecido  $\alpha$ , é não enviesado se a função poder assume um valor mínimo no ponto  $\theta_0$ . Para tanto a derivada primeira da função poder, aplicada no ponto  $\theta_0$  deve ser igual a zero.

$$\begin{aligned} \frac{\partial \beta(\theta, \varphi)}{\partial \theta} &= \frac{\partial \sum_{x=0}^n \varphi(x) \binom{n}{x} \theta^x (1 - \theta)^{n-x}}{\partial \theta} = \\ &= \sum_{x=0}^n \varphi(x) \binom{n}{x} \left( x \theta^{x-1} (1 - \theta)^{n-x} - (n - x) \theta^x (1 - \theta)^{n-x-1} \right) = \\ &= \sum_{x=0}^n \varphi(x) \binom{n}{x} \theta^{x-1} (1 - \theta)^{n-x-1} (x - n\theta). \end{aligned}$$

Aplicando-se a derivada primeira no ponto  $\theta_0$  e igualando-se a zero tem-se:

$$\frac{\partial \beta(\theta, \varphi)}{\partial \theta} \Big|_{\theta = \theta_0} = 0 \Rightarrow \sum_{x=0}^n \varphi(x) \binom{n}{x} \theta_0^{x-1} (1 - \theta_0)^{n-x-1} (x - n\theta_0) = 0.$$

Multiplicando-se ambos os lados da expressão por  $\theta_0(1-\theta_0)$  tem-se:

$$\sum_{x=0}^n \varphi(x) \binom{n}{x} \theta_0^x (1-\theta_0)^{n-x} (x - E_{\theta_0}[X]) = 0.$$

Nesse caso, a função crítica deve estabelecer um equilíbrio em relação ao valor esperado, sob hipótese nula considerada verdadeira.

No caso do teste de equilíbrio genético de Hardy-Weinberg, em cada subconjunto do espaço amostral com igual composição alélica ( $\Omega_i$ ), pode ser aplicado o equilíbrio em relação ao valor esperado na distribuição genotípica, baseada no número de heterozigotos.

O teste de equilíbrio genético de Hardy-Weinberg  $\varphi(\cdot)$ , que possui a propriedade do não-viés, é definido então por intermédio da adaptação do teste binomial, com hipótese nula simples, não enviesado, para a distribuição do número de heterozigotos em cada subconjunto do espaço amostral com igual composição alélica. Tal adaptação, para um teste com nível de significância preestabelecido  $\alpha$ , é dada por:

- i ) para cada subconjunto amostral da distribuição genotípica, com valores alélicos iguais ( $\Omega_i$ ) são calculadas as probabilidades genotípicas condicionada à distribuição alélica;
- ii ) em cada subconjunto amostral são calculados os valores esperados para a distribuição do número de heterozigotos, utilizando-se as probabilidades condicionais do item i;
- iii ) associa-se valor  $\varphi(x) = 1$  para os valores correspondentes às menores probabilidades condicionas, nos extremos da distribuição do número de heterozigotos, até a obtenção do valor limite  $\alpha/2$ , para cada lado dos extremos (valores maiores e menores que o valor esperado), de modo que, com a inclusão do próximo ponto da ordenação, o nível de significância ultrapassa o valor  $\alpha/2$ ;



iv ) para cada ponto limite associa-se um valor aleatorizado e o restante dos pontos do subconjunto valor  $\varphi(x) = 0$ , de modo a ser obtido, para cada subconjunto amostral um somatório acumulado  $\alpha$ , bem como um equilíbrio em torno do valor esperado.

Em Meulepas (1999), essa propriedade de equilíbrio, adaptada na distribuição do número de heterozigotos condicionada à distribuição alélica, é utilizada para o cálculo do nível de significância associado a pontos isolados do espaço amostral ( $p$ -valor), sendo citado, no referido trabalho, ter esse cálculo uma melhor forma do que o teste exato, reduzindo o viés apresentado neste último.

No presente trabalho, sendo aplicado o mecanismo de equilíbrio em torno do valor esperado para a distribuição do número de heterozigotos, foi encontrado um teste não enviesado para o equilíbrio genético de Hardy-Weinberg. A prova formal desse resultado não é encontrada em Meulepas (1999), bem como não é disponibilizada de forma algébrica no presente trabalho. Porém, após exaustiva construção de casos particulares, com diversos tamanhos de amostras (2, 5, 10, 16, 20, 25, 50), bem como vários níveis de significância (1%, 5% e 10%) não foram encontrados indícios de viés.

Devido à forma de construção da região de rejeição, do mencionado teste, com a inclusão dos pontos de maior probabilidade pode ser afirmado, sendo admitido o não-viés para esse procedimento, que tal teste, dentro da classe dos testes não enviesados, apresenta-se como o teste não enviesado uniformemente mais poderoso. Um exemplo do procedimento de construção desse teste encontra-se no capítulo 4.

## 4 RESULTADOS E DISCUSSÃO

### 4.1 Aplicação da estatística $X^2$ , no teste de equilíbrio genético

Em uma população de indivíduos, sendo observado um loco, cujos genes são formados por dois únicos alelos “A” e “a”, as frequências genotípicas “AA”, “Aa” e “aa”, quando em equilíbrio genético de Hardy-Weinberg, são respectivamente dadas por:  $\theta^2$ ,  $2\theta(1-\theta)$  e  $(1-\theta)^2$ , em que  $\theta \in [0, 1]$  representa a frequência do alelo “A”, e o complementar,  $1-\theta$ , a frequência do alelo “a”.

Associando-se  $X = (X_1, X_2, X_3)$  às frequências genotípicas, “AA”, “Aa” e “aa”, respectivamente, obtidas em uma amostra com  $n$  indivíduos selecionados aleatoriamente, tem-se o vetor aleatório  $X$ , em questão, com distribuição multinomial de parâmetros  $n$  e  $\theta = (\theta_1, \theta_2, \theta_3)$ , com espaço paramétrico dado por:

$$\Theta = \{(\theta_1, \theta_2, \theta_3): \theta_i \in [0,1] \text{ e } \theta_1 + \theta_2 + \theta_3 = 1\},$$

em que o subconjunto do espaço paramétrico, associado ao equilíbrio genético de Hardy-Weinberg, é dado por:

$$\Theta_0 = \{(\theta_1, \theta_2, \theta_3) \in \Theta : 4\theta_1\theta_3 = (\theta_2)^2\}.$$

O teste de equilíbrio genético de Hardy-Weinberg é definido por:

$$H_0 : (\theta_1, \theta_2, \theta_3) \in \Theta_0,$$

$$H_1 : (\theta_1, \theta_2, \theta_3) \notin \Theta_0.$$

Pode ser visto na Figura 1, o espaço amostral, para um exemplo particular adotando-se o tamanho amostral 16, projetado no plano das freqüências genóticas dos homozigotos, “AA” e “aa”.

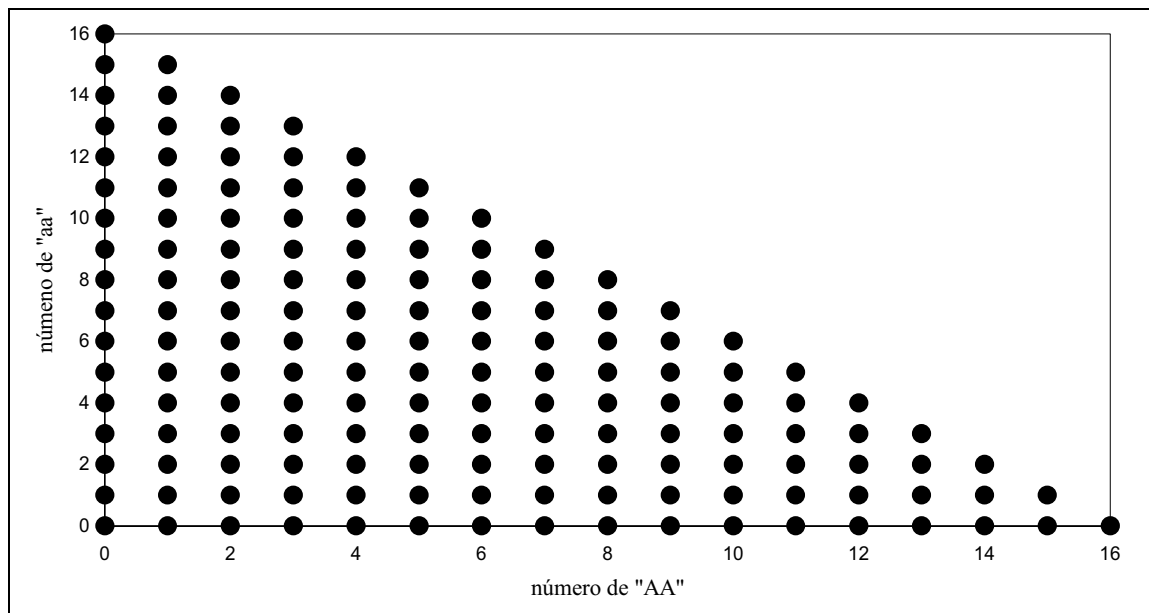


Figura 1 - Espaço amostral, projetado no plano das freqüências genóticas “AA” e “aa”, para o tamanho amostral 16.

Por ser  $X^2$  a mais antiga estatística aplicada em testes multinomiais, inicialmente será abordada sua utilização no teste de equilíbrio genético de Hardy-Weinberg. Como exemplo para análise, será tomada a amostra de 16 indivíduos, apresentada na Figura 1.

Desejando-se um teste com nível de significância 5%, a cota inferior da estatística  $X^2$ , para a composição da região de rejeição, é de 3,84. Essa cota está baseada

no quantil 5% da distribuição qui-quadrado, com 1 grau de liberdade. Para cada ponto do espaço amostral é calculado, então, o valor da estatística  $X^2$  e, sendo esse valor superior à 3,84, o ponto pertencerá a região de rejeição.

Como exemplo tem-se o ponto (7,2,7), que, sob hipótese de equilíbrio genético, possui valor esperado, estimado pelo processo de máxima verossimilhança, dado por: (4,8,4). O valor da estatística  $X^2$ , associado a esse ponto é 9, portanto tal ponto pertence a região de rejeição.

Na Figura 2, pode ser visto a região de rejeição, cujos pontos apresentam valor da estatística  $X^2$ , superior à 3,84, cuja função crítica está associado ao valor 1.

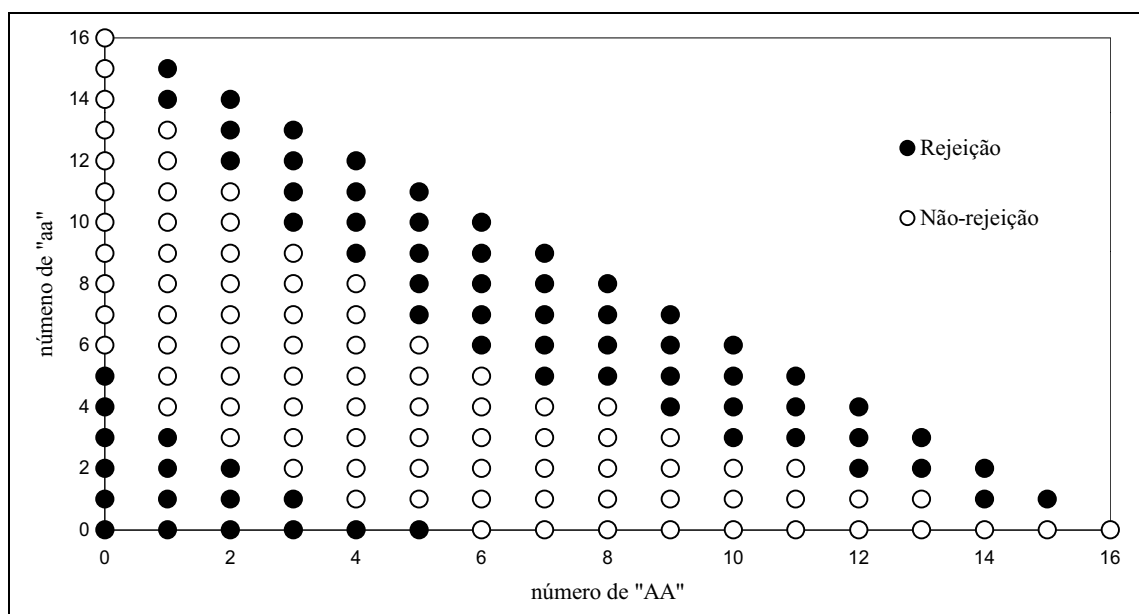


Figura 2 - Região de rejeição com a utilização de  $X^2$ , com valores superiores a 3,84.

Na Figura 3, tem-se o gráfico da função probabilidade de rejeição de  $H_0$ , que nesse caso, é o somatório das probabilidades dos pontos apresentados na Figura 2, para todos os possíveis valores paramétricos correspondentes ao equilíbrio genético de Hardy-Weinberg, em função da frequência do alelo "A".

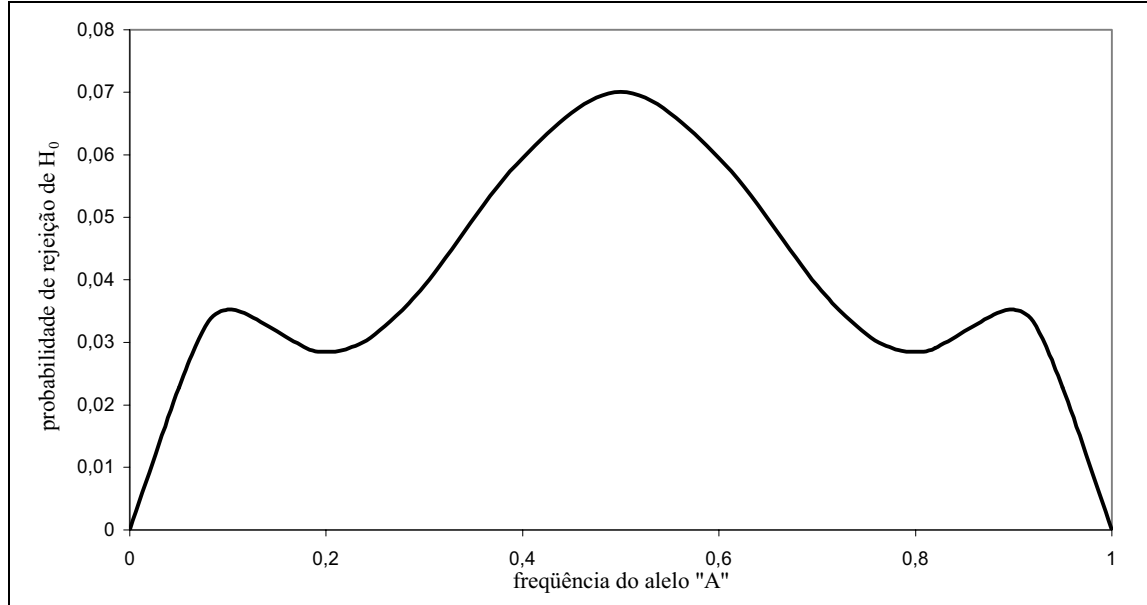


Figura 3 - Função probabilidade de rejeição de  $H_0$ , utilizando  $X^2$ , com valores superiores a 3,84.

Como pode ser visto na Figura 3, o máximo valor da probabilidade de rejeição de  $H_0$ , com domínio em  $\Theta_0$ , que corresponde ao nível de significância, é de aproximadamente 0,07. Isso torna o nível de significância do teste acima do que era pretendido.

Para obtenção de um teste com nível de significância preestabelecido, com a utilização da estatística  $X^2$ , será adaptado o mecanismo de ordenação do espaço amostral apresentado em Radlow & Alf (1975) para o caso de hipótese nula composta. Calcula-se para cada ponto do espaço amostral, da distribuição genotípica, o valor  $X^2$ , com valor esperado estimado pelo critério da máxima verossimilhança. Tal região de rejeição, nesse caso, é formada pelos pontos que estão associados aos maiores valores de  $X^2$  até a obtenção do maior número possível de pontos que, sendo acumuladas as probabilidades sob equilíbrio de Hardy-Weinberg, não exceda o nível de significância preestabelecido.

Fixando-se o nível de significância em 5%, os pontos  $\{(5,4,7),(7,4,5)\}$ , sendo 3,87 o valor da estatística  $X^2$  aplicada a esses pontos e  $\{(2,12,2), (6,4,6)\}$ , sendo 4

o valor da estatística  $X^2$  aplicada a esses pontos, não poderão ser incluídos na região de rejeição, que comportará todos os pontos que apresentam valor de  $X^2$  superior a 4.

Tem-se na Figura 4 a região de pontos que apresentam valor da estatística  $X^2$  superior a 4, cuja função crítica está associado ao valor 1, com relação a todos os possíveis valores genotípicos para 16 indivíduos. As freqüências genotípicas “AA”, “Aa” e “aa”, associadas à rejeição da hipótese de equilíbrio, estão projetadas no plano das freqüências genotípicas “AA” e “aa”.

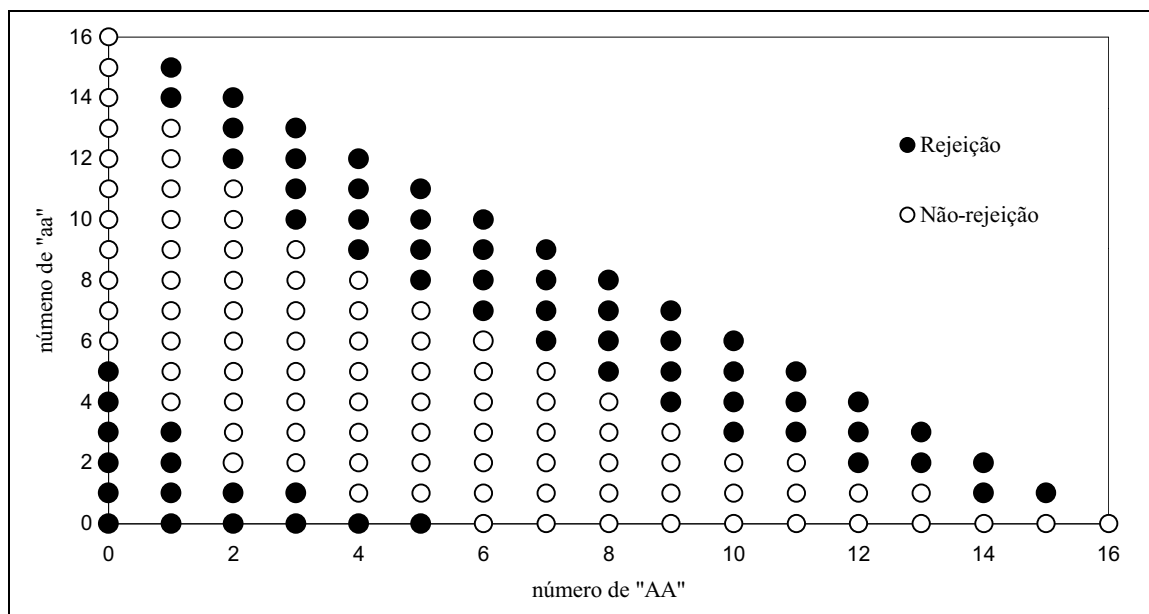


Figura 4 - Região de rejeição com a utilização de  $X^2$ , com valores superiores a 4.

Na Figura 5, tem-se o gráfico da função probabilidade de rejeição de  $H_0$ , com domínio em  $\Theta_0$ , que nesse caso, é o somatório das probabilidades dos pontos apresentados na Figura 4, para todos os possíveis valores paramétricos correspondentes ao equilíbrio genético de Hardy-Weinberg, em função da freqüência do alelo “A”.

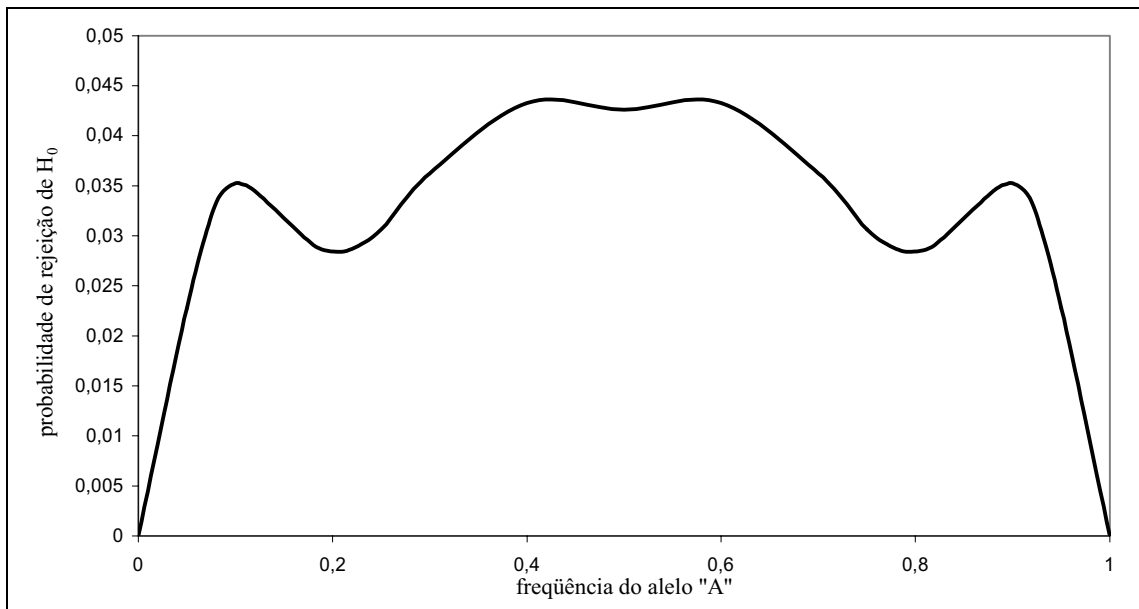


Figura 5 - Função probabilidade de rejeição de  $H_0$ , utilizando  $X^2$ , com valores superiores a 4.

Como pode ser notado na Figura 5, o real nível de significância do teste é inferior ao pretendido de 5%, bem como a probabilidade de rejeição da hipótese nula, no caso de valores de  $\theta$  próximos a zero ou um, é próxima a zero, acarretando uma dificuldade de rejeição dessa hipótese em regiões do espaço paramétrico próximas a estas, que correspondem à presença de alelos raros.

Um teste concorrente a este é o teste com a utilização da estatística definida por Cressie & Read (1984).

#### 4.2 Aplicação da estatística $2nI^{2/3}$ , no teste de equilíbrio genético

Como afirmado em Cressie & Read (1984), a estatística  $2nI^\lambda$ , para o caso particular  $\lambda = 2/3$ , apresenta-se como uma excelente candidata à substituição da estatística  $X^2$ . Para análise comparativa, também será aplicada tal estatística no teste de equilíbrio genético de Hardy-Weinberg, em uma amostra de tamanho 16.

Para obtenção de um teste com nível de significância preestabelecido,

com a utilização da estatística  $2nI^{2/3}$ , também será adaptado o mecanismo de ordenação do espaço amostral apresentado em Radlow & Alf (1975) para o caso de hipótese nula composta. Calcula-se para cada ponto do espaço amostral da distribuição genotípica o valor  $2nI^{2/3}$ , com valor esperado estimado pelo critério do mínimo valor de  $2nI^{2/3}$ , tal como é apresentado originalmente em Cressie & Read (1984).

Como exemplo de cálculo tem-se o ponto (1,1,14), que apresenta valor da estatística  $2nI^{2/3}$  de 4,01. No cálculo do mínimo valor da estatística é utilizado como valores esperados (0,29 , 3,75 , 11,96).

Fixando-se o nível de significância em 5%, os pontos  $\{(2,12,2), (6,4,6)\}$ , sendo 4,04 o valor da estatística  $2nI^{2/3}$ , não poderão ser incluídos na região de rejeição, que comportará todos os pontos que apresentam valor de  $2nI^{2/3}$  superior a 4,04.

A Figura 6 apresenta a região de pontos que apresentam valor da estatística  $2nI^{2/3}$  superior a 4,04, cuja função crítica está associado ao valor 1, com relação a todos os possíveis valores genotípicos para 16 indivíduos.

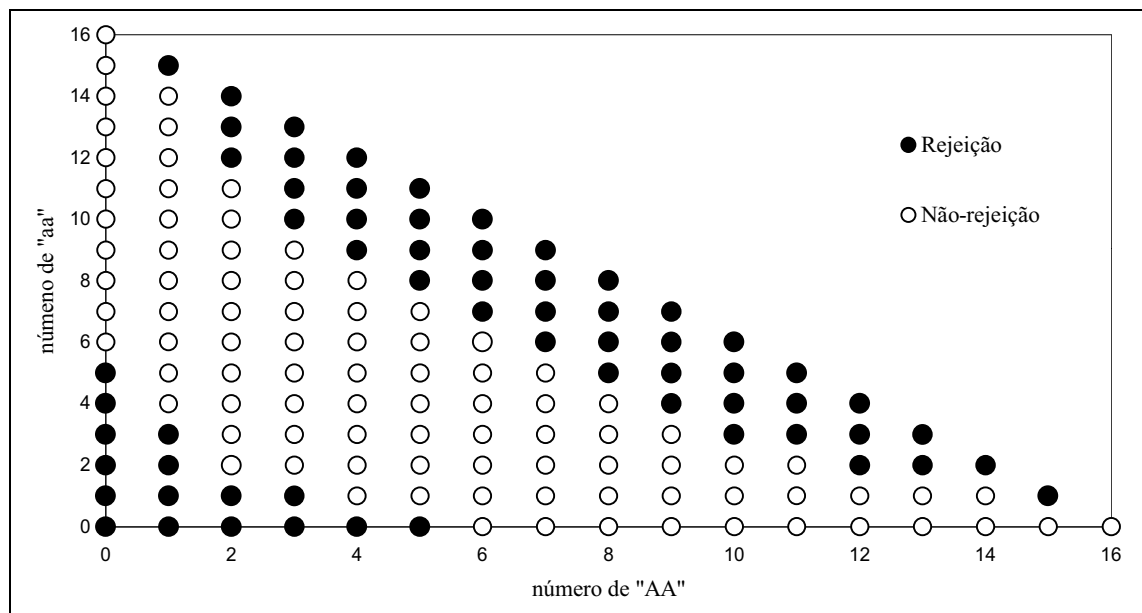


Figura 6 - Região de rejeição com a utilização de  $2nI^{2/3}$ , com valores superiores a 4,04.

Como pode ser visto, comparando-se as Figuras 4 e 6, todos os pontos que estão presentes na região de rejeição que utiliza a estatística  $2nI^{2/3}$ , também estão



presentes na região de rejeição que utiliza a estatística  $X^2$ , porém os pontos  $\{(1,1,14), (14,1,1)\}$ , estão presentes somente na região de rejeição que utiliza a estatística  $X^2$ . Assim, pode-se afirmar que a estatística  $X^2$  apresenta-se mais eficiente, pois a inclusão de um ponto a mais aumenta o poder do teste. Tal fato pode ser notado para outros tamanhos de amostras.

Na estatística  $2nI^{2/3}$  é utilizado o critério do mínimo valor, para a estimação dos valores esperados. Tal estimação reduz o valor da estatística, principalmente quando o ponto em questão apresenta uma concentração, do número total de indivíduos, em uma mesma categoria genotípica. Como exemplo o ponto (1,1,14) apresenta valor de  $2nI^{2/3}$  de 4,01, porém, caso seja utilizado a estimativa de máxima verossimilhança, o valor de  $2nI^{2/3}$  é de 5,07. Dessa forma, quando é estabelecida uma ordenação do espaço amostral, para que seja verificado um nível de significância preestabelecido, a estatística  $2nI^{2/3}$ , com mínimo valor utilizado como critério de estimação dos valores esperados, não possui a mesma eficiência que a estatística  $X^2$ , quando o critério da máxima verossimilhança é utilizado como critério de estimação dos valores esperados.

Adotando-se o critério de estimação da máxima verossimilhança, para a estatística  $2nI^{2/3}$ , no caso do teste de equilíbrio genético de Hardy-Weinberg, pode-se notar valores muito semelhantes aos valores de  $X^2$ . Desse modo não existem grandes vantagens na substituição da estatística  $X^2$ , por  $2nI^{2/3}$ .

### **4.3 Aplicação do teste exato, no equilíbrio genético**

Como pode ser visto em Elston & Forthofer (1977), o teste de equilíbrio genético de Hardy-Weinberg com procedimento exato está baseado nas probabilidades dos possíveis valores observados da distribuição genotípica, condicionada à distribuição alélica.

Para uma discussão inicial, sobre o tamanho exato do teste apresentado em Elston & Forthofer (1977) (teste exato), considerar-se-á uma amostra de 16 indivíduos em que é observado um loco, cujos genes são formados por dois únicos alelos

“A” e “a”, sendo tomado como um possível ponto observado 1 homozigoto “AA”, 12 heterozigotos “Aa” e 3 homozigotos “aa”.

Observando-se todos os outros pontos amostrais com igual distribuição alélica, 14 alelos “A” e 18 alelos “a”, tem-se:  $\{(0,14,2), (1,12,3), (2,10,4), (3,8,5), (4,6,6), (5,4,7), (6,2,8), (7,0,9)\}$ . Calculando-se as probabilidades condicionais, como apresentado em Elston & Forthofer (1977),

$$P[X_1 = x_1, X_2 = x_2, X_3 = x_3 / 2X_1 + X_2 = 2x_1 + x_2] = \frac{\binom{n}{x_1, x_2, x_3} 2^{x_2}}{\binom{2n}{2x_1 + x_2, 2x_3 + x_2}},$$

tem-se, respectivamente, os seguintes valores:  $\{0,00417041; 0,0632512; 0,260911; 0,391367; 0,228297; 0,0489209; 0,00305755; 0,0000242663\}$ . O nível de significância associado ao ponto observado (1,12,3), cuja probabilidade condicional é dada por: 0,0632512, acumulando-se as probabilidades igual ou inferior ao referido ponto, é dado por:  $0,00417041+0,0632512+0,0489209+0,00305755+0,0000242663 \cong 0,119424$ .

Assim, a função crítica, associada a esses pontos do espaço amostral, é dada por:

$$\varphi(\mathbf{x}) = \begin{cases} 1, & \text{se } \mathbf{x} \in \{(0,14,2), (1,12,3), (5,4,7), (6,2,8), (7,0,9)\}, \\ 0, & \text{se } \mathbf{x} \in \{(2,10,4), (3,8,5), (4,6,6)\}. \end{cases}$$

Ampliando-se a construção do teste de equilíbrio genético de Hardy-Weinberg, com um nível de significância preestabelecido de 0,119424 para o restante do espaço amostral, buscando a manutenção desse nível de significância para os pontos com diferente distribuição alélica, tem-se um problema: sem aleatorização da decisão não é possível manter o mesmo nível de significância.

Como exemplo têm-se os pontos, com o mesmo número de indivíduos (16), com uma outra composição alélica: 10 alelos “A” e 22 alelos “a”,  $\{(0,10,6),$

$(1,8,7), (2,6,8), (3,4,9), (4,2,10), (5,0,11)\}$ , que apresentam as seguintes probabilidades condicionais:  $\{0,127111; 0,40857; 0,357499; 0,099305; 0,007448; 0,0000677\}$ . Acumulando-se as probabilidades de menor valor, até atingir o limite de  $0,119424$ , tem-se:  $0,099305 + 0,007448 + 0,0000677 = 0,106821$ .

Assim, caso não haja aleatorização, a função crítica associada a esses pontos do espaço amostral é dada por:

$$\varphi(\mathbf{x}) = \begin{cases} 1, & \text{se } \mathbf{x} \in \{(3,4,9), (4,2,10), (5,0,11)\}, \\ 0, & \text{se } \mathbf{x} \in \{(0,10,6), (1,8,7), (2,6,8)\}. \end{cases}$$

Tem-se na Figura 7 a região de pontos cuja função crítica está associada ao valor 1, com relação a todos os possíveis valores genotípicos para 16 indivíduos, cujas probabilidades genotípicas acumuladas, condicionadas à distribuição alélica, não ultrapassam o valor  $0,119424$ .

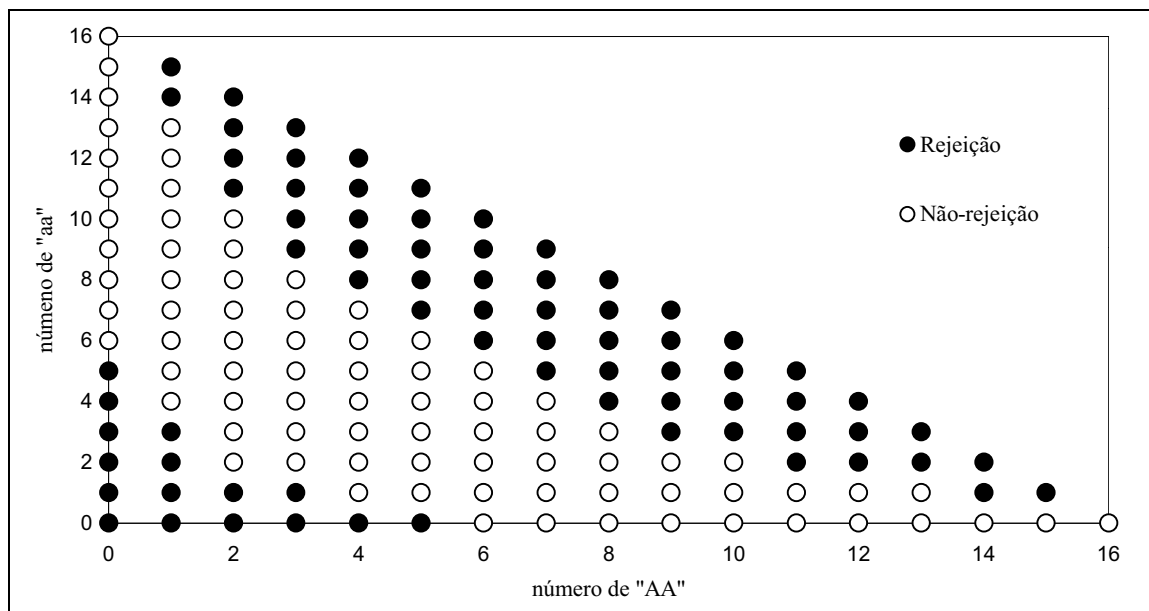


Figura 7 - Região de rejeição para o teste exato não aleatorizado, com nível de significância desejado de 11,94%.

Na Figura 8, tem-se o gráfico da função probabilidade de rejeição de  $H_0$ , com domínio em  $\Theta_0$ , ou seja, o somatório das probabilidades dos pontos apresentados na Figura 7, para todos os possíveis valores paramétricos correspondentes ao equilíbrio genético de Hardy-Weinberg, em função da frequência do alelo "A".

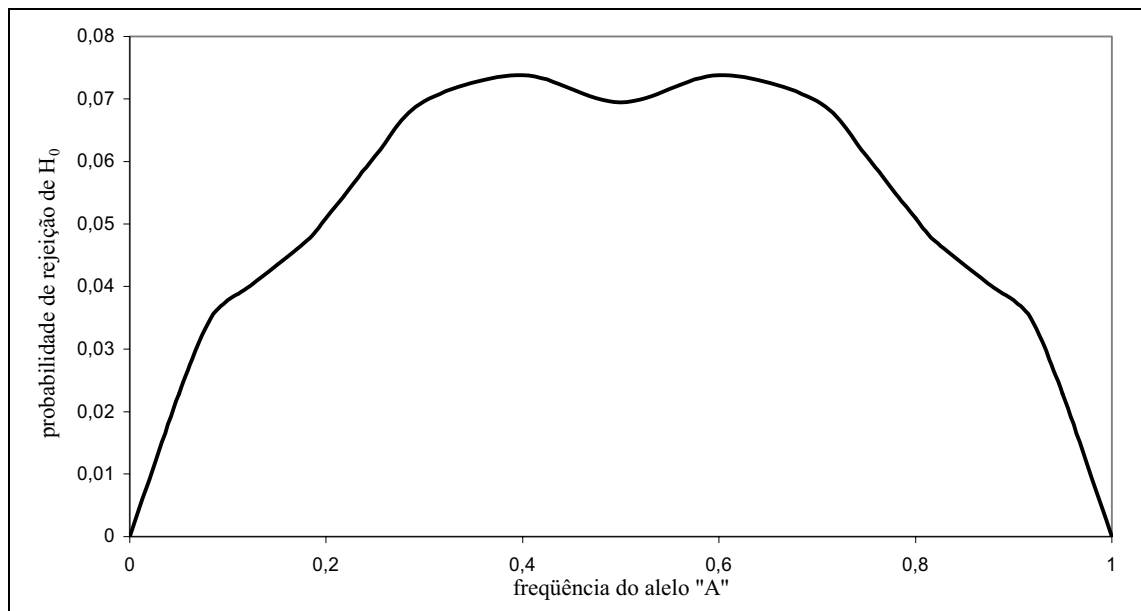


Figura 8 - Função tamanho do teste exato não aleatorizado, com nível de significância desejado de 11,94%.

Pode ser visto que o máximo valor da função tamanho do teste é de 0,0745 correspondente aos valores de  $\theta \cong 0,3692$  e  $\theta \cong 0,6308$ . Tal fato torna o nível de significância do teste menor do que 0,119424. Esse fato é observado para todos os pontos do espaço amostral, caso seja tomado o nível de significância do teste o valor de significância ( $p$ -valor) associado ao ponto. Como pode ser visto em Tocher (1950), o tamanho real do teste para variáveis discretas será exato somente quando houver aleatorização.

#### 4.4 Comparação de procedimentos não aleatorizados

Para que sejam comparados diferentes mecanismos de testes, é imprescindível que ambos tenham o mesmo nível de significância. Para tanto, a região de rejeição do teste exato deve ser construída de tal modo a não exceder um certo nível de significância preestabelecido. Inicialmente, será proposto um mecanismo que define um teste não aleatorizado, dividindo-se o espaço amostral em pontos favoráveis ou desfavoráveis ao equilíbrio.

No teste exato não aleatorizado, para cada subconjunto do espaço amostral da distribuição genotípica, com igual distribuição alélica, as probabilidades dos possíveis valores amostrais condicionados à respectiva distribuição alélica são calculadas. Tal região de rejeição é formada pelos pontos que estão associados às menores probabilidades condicionais, até a obtenção do maior valor possível acumulado que não exceda o nível de significância preestabelecido.

Na Figura 9 são apresentadas conjuntamente as regiões de rejeições associadas ao teste, utilizando a estatística  $X^2$  e o mecanismo exato, para um caso particular, de uma amostra com 16 indivíduos a um nível de significância 5%.

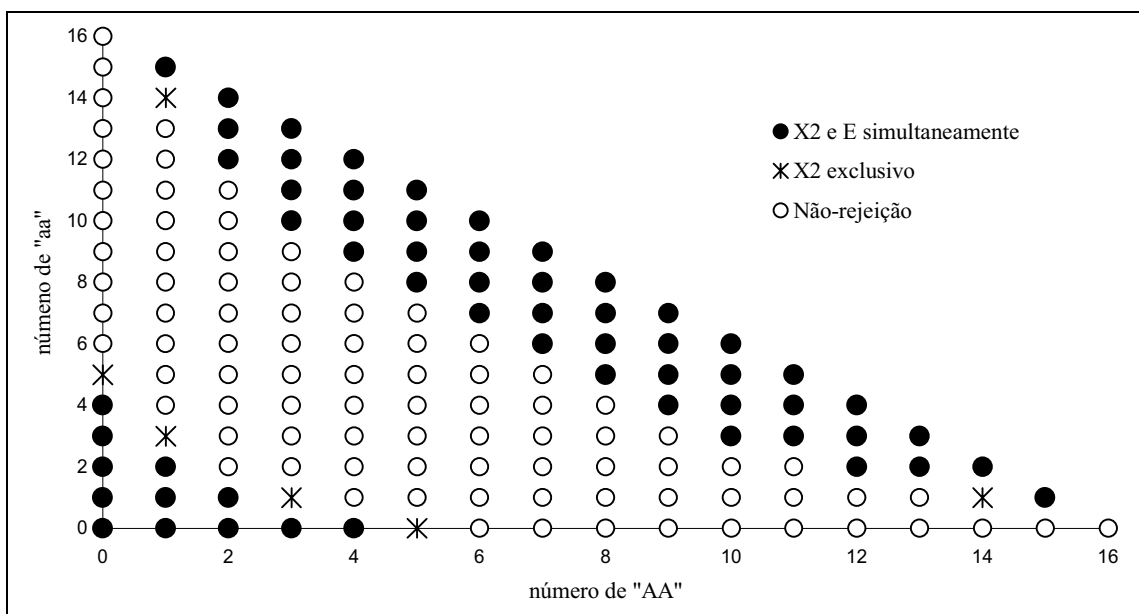


Figura 9 - Regiões de rejeição conjuntas projetadas no plano das frequências genóticas “AA” e “aa”.

Os pontos designados por “X2 e E simultaneamente” fazem parte das regiões de rejeição de ambos os testes, e os pontos designados por “X2 exclusivo” fazem parte somente da região de rejeição do teste que utiliza a estatística  $X^2$ . Como pode ser visto na Figura 9, o teste exato não aleatorizado, assim como a estatística  $2nI^{2/3}$ , apresenta um número de pontos inferior ao teste que utiliza  $X^2$ , de tal modo que todos os pontos que compõem a região de rejeição do primeiro estão presentes na região de rejeição do segundo, sendo que alguns pontos extras, que não pertencem à região de rejeição do teste exato não aleatorizado, pertencem à região de rejeição do teste que utiliza  $X^2$ . Assim, no exemplo particular, devido à presença de pontos exclusivos na região de rejeição, pode-se notar uma melhor eficiência no teste  $X^2$ , sendo o teste exato não aleatorizado considerado inadmissível em relação ao concorrente. Tal fato repete-se em muitos outros casos, com tamanhos de amostras diferentes, bem como outros níveis de significância.

Assim, caso não haja aleatorização na função crítica, o mecanismo defendido em trabalhos como: Chapco (1976), Elston & Fortholfer (1977) e Weir

(1996), dentre outros, não apresenta uma solução melhor que uma adaptação do tradicional teste  $X^2$  como citam tais trabalhos.

#### 4.5 Comparação de procedimentos aleatorizados

Sendo possível a aleatorização na função crítica, o teste exato, com nível de significância preestabelecido será obtido por uma adaptação do mecanismo proposto em Tocher (1950), para teste de independência em tabelas de contingência.

De modo análogo ao teste não aleatorizado, para cada subconjunto do espaço amostral da distribuição genotípica, com igual distribuição alélica, são calculadas as probabilidades dos possíveis valores amostrais condicionadas à respectiva distribuição alélica. Tal região de rejeição é formada pelos pontos que estão associados às menores probabilidades condicionais até a obtenção do maior valor possível acumulado que não exceda o nível de significância preestabelecido, aleatorizando-se o próximo ponto que seria incluído na região de rejeição, para a obtenção do exato nível de significância pretendido.

Como exemplo, pode ser citado, no caso particular analisado anteriormente, o seguinte subconjunto do espaço amostral:  $\{(0,8,8), (1,6,9), (2,4,10), (3,2,11), (4,0,12)\}$ , cujos elementos possuem a mesma distribuição alélica: 8 alelos “A” e 24 alelos “a”. A distribuição genotípica, condicionada à distribuição alélica, para os respectivos valores é dada por:  $\{0,3132; 0,4873; 0,1827; 0,0166; 0,0002\}$ . Desse modo, caso o nível de significância seja 5%, os pontos  $\{(3,2,11), (4,0,12)\}$  pertencem à região de rejeição, bem como o ponto  $(2,4,10)$  aleatorizado em 0,1818. Assim, a função crítica associada a esses pontos do espaço amostral é dada por:

$$\varphi(\mathbf{x}) = \begin{cases} 1, & \text{se } \mathbf{x} \in \{(3,2,11), (4,0,12)\}, \\ 0,1818, & \text{se } \mathbf{x} = (2,4,10), \\ 0, & \text{se } \mathbf{x} \in \{(0,8,8), (1,6,9)\}. \end{cases}$$

de tal modo que:  $0,05 \cong 0,1818 + 0,1827 + 0,0166 + 0,0002$ .

Procedendo-se de modo análogo a todos os outros subconjuntos do espaço amostral, com outros valores alélicos, a região de rejeição pode ser construída com pontos de rejeição, não rejeição e aleatorizados.

A Figura 10 mostra a aplicação do mecanismo de aleatorização para o teste exato no caso particular de uma amostra com 16 indivíduos, a um nível de significância 5%.

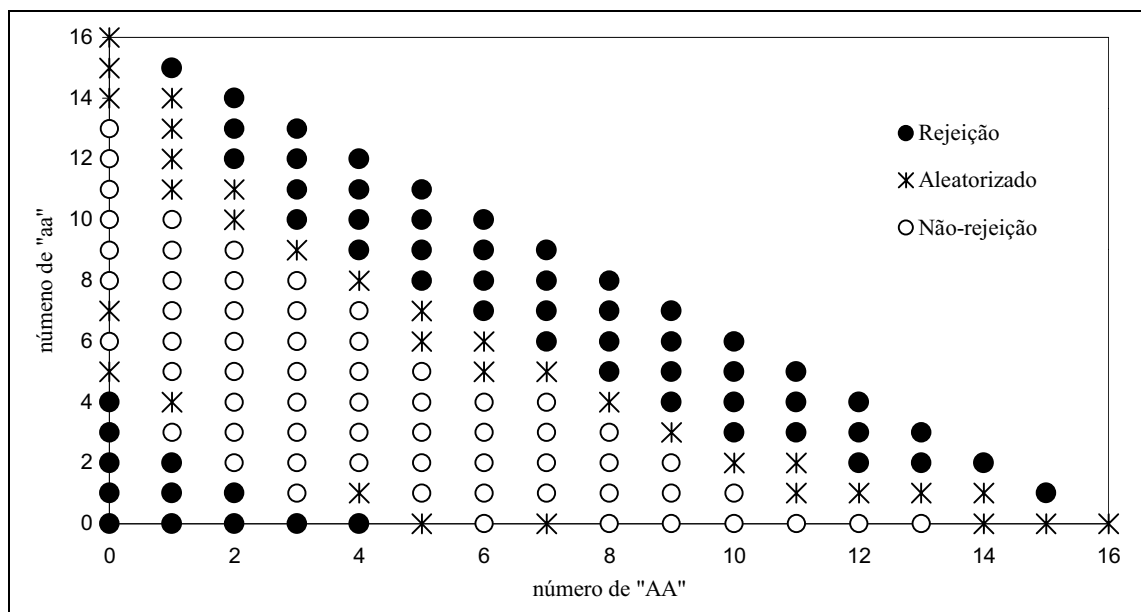


Figura 10 - Região de rejeição do teste exato, com nível de significância 5%.

No mecanismo alternativo em que é utilizada a estatística  $X^2$ , será aleatorizado o ponto seguinte, que seria incluído no mecanismo de ordenação do espaço amostral apresentado em Radlow & Alf (1975) para o caso de hipótese nula composta. Tal ponto, sendo incluído na região de rejeição, sem aleatorização, tornaria o real nível de significância do teste acima do pretendido. Sendo o nível de significância 5%, a região de rejeição será a mesma apresentada anteriormente, com a inclusão dos pontos:  $\{(6,4,6), (2,12,2)\}$ , que possuem a mesma divergência  $X^2$ , aleatorizados em 0,4438.

A Figura 11 mostra a aplicação do mecanismo de aleatorização para o teste  $X^2$  no caso particular de uma amostra com 16 indivíduos, a um nível de



significância 5%.

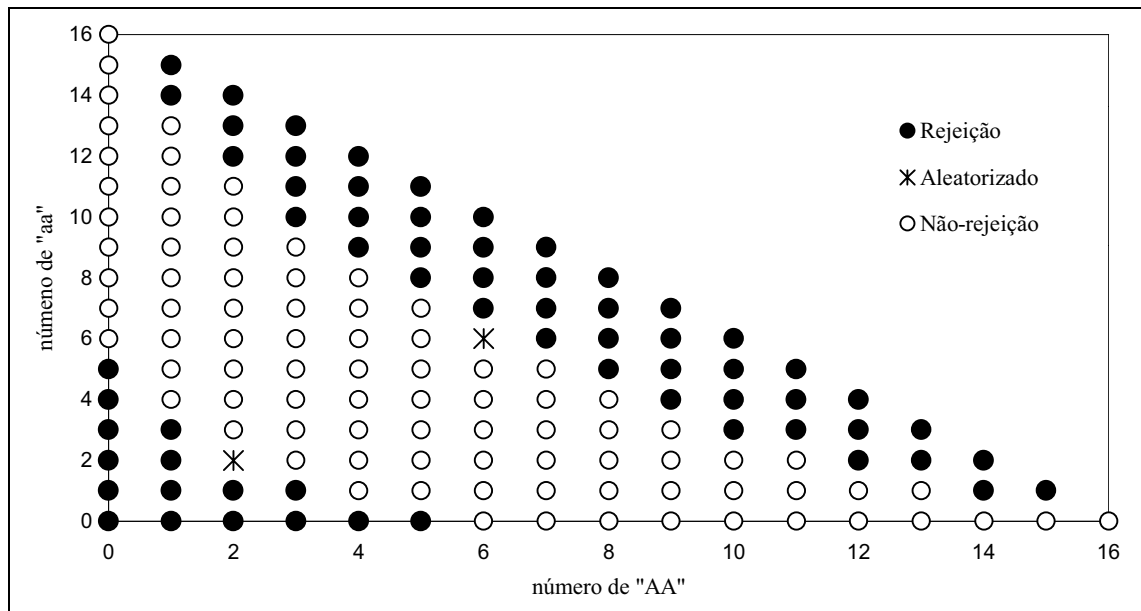


Figura 11 - Região de rejeição do teste com a utilização de  $X^2$ , aleatorizado, com nível de significância 5%.

Comparando-se as regiões de rejeição dos testes, pode-se notar que alguns pontos, que fazem parte da região de rejeição do teste  $X^2$ , fazem parte parcialmente (aleatorizado) do teste exato, que, por sua vez, inclui parcialmente alguns pontos que não fazem parte da região de rejeição do teste  $X^2$  aleatorizado. Desse modo, não será possível dizer que um dos testes é mais poderoso que o outro, pois, devido à constituição das regiões de rejeição, em alguns subconjuntos do espaço paramétrico, relacionado à hipótese de não-existência do equilíbrio genético de Hardy-Weinberg, tem-se alternadamente um dos testes apresentando um melhor poder de rejeição.

Procedendo-se a uma reparametrização natural (Weir, 1996), visto do princípio genético, tem-se:

$$\begin{cases} \theta_1 = \theta^2 + \theta(1-\theta)\gamma \\ \theta_2 = 2\theta(1-\theta)(1-\gamma) \\ \theta_3 = (1-\theta)^2 + \theta(1-\theta)\gamma \end{cases} \Leftrightarrow \begin{cases} \theta = \frac{2\theta_1 + \theta_2}{2} \\ \gamma = \frac{\theta_1 - (2\theta_1 + \theta_2)^2}{(2\theta_1 + \theta_2)(2\theta_3 + \theta_2)} \end{cases},$$

em que  $\theta$  é a frequência do alelo “A” e  $\gamma$  o índice de fixação, associado à autofecundação. O novo espaço paramétrico é dado por:

$$\Delta = \{(\theta, \gamma) : \theta \in [0; 0,5], \gamma \in [-\theta/(1-\theta), 1] ; \theta \in (0,5; 1], \gamma \in [-(1-\theta)/\theta, 1]\},$$

e o teste de equilíbrio genético de Hardy-Weinberg, nesse caso, é dado por:

$$H_0 : \gamma = 0$$

$$H_A : \gamma \neq 0.$$

Estudando-se a função poder para casos particulares de  $\gamma$ , pode-se notar um melhor poder no teste exato para valores positivos de  $\gamma$ . Isso se deve ao fato da região de rejeição definida por tal teste conter muitos pontos aleatorizados na borda superior da região de rejeição, como pode ser visto nas Figuras 10 e 11. Porém, para valores negativos de  $\gamma$ , o teste  $X^2$  aleatorizado torna-se mais poderoso para um grande intervalo de valores de  $\theta$  centrado em 0,5. Tal fato deve-se à presença de pontos na região de rejeição do teste  $X^2$  aleatorizado, que estão presentes apenas parcialmente na região de rejeição no teste exato. Uma comparação do poder para os testes, com valores fixos de  $\gamma$  é apresentada nas Figuras 12 e 13.

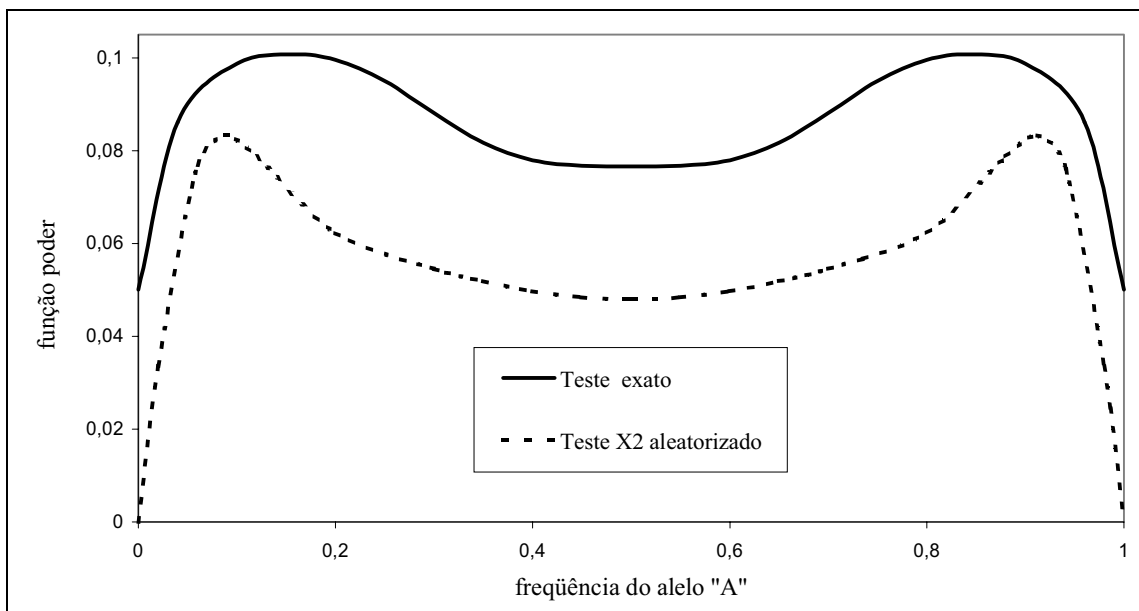


Figura 12 - Função poder dos testes aleatorizados no caso particular de  $\gamma = 0,1$ .

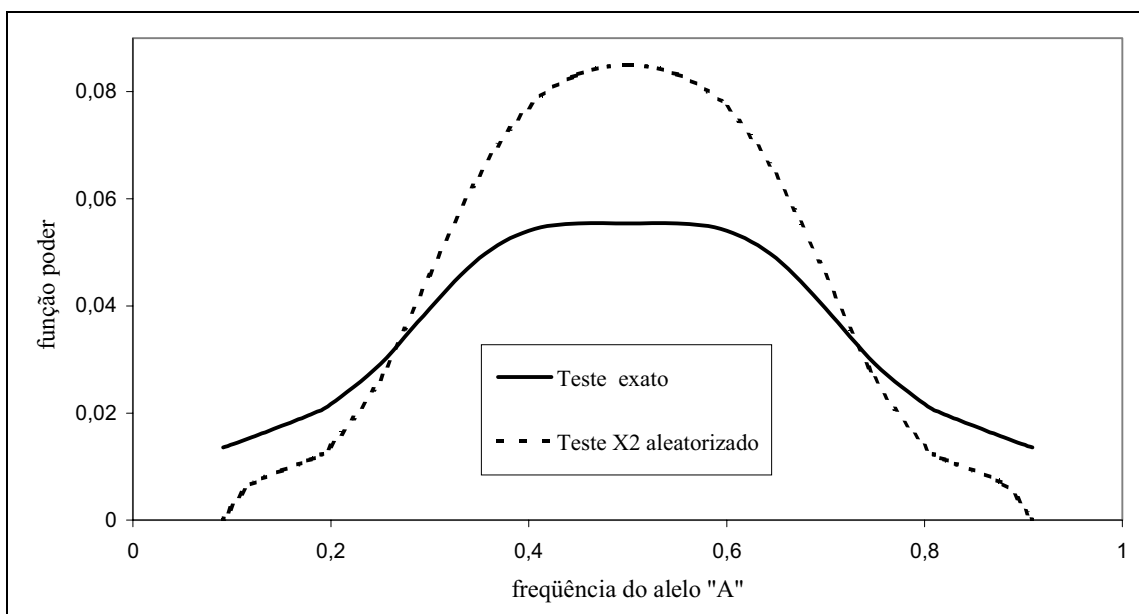


Figura 13 - Função poder dos testes aleatorizados no caso particular de  $\gamma = -0,1$ .

Contrariando o exposto em Weir (1996) de que os testes com a utilização da estatística  $X^2$  são ultrapassados e apresentados somente por considerações históricas,

pode-se notar, no caso específico do teste de equilíbrio genético de Hardy-Weinberg, nos casos de frequências alélicas dos homozigotos aproximadamente equilibradas, um melhor poder no teste com a utilização da estatística  $X^2$ .

#### 4.6 Procedimento não enviesado

Um outro procedimento aleatorizado pode ser encontrado em Meulepas (1999). Tal procedimento, como citado no capítulo 3, possui excelentes propriedades matemáticas, tais como ser não enviesado uniformemente mais poderoso. Como no procedimento do teste exato, a distribuição alélica encontra-se como variável condicionante.

Como exemplo de construção da região de rejeição do teste apresentado em Meulepas (1999), pode ser citado, no caso particular com 16 indivíduos, analisado anteriormente, o seguinte subconjunto do espaço amostral:  $\{(0,8,8), (1,6,9), (2,4,10), (3,2,11), (4,0,12)\}$ , cujos elementos possuem a mesma distribuição alélica: 8 alelos “A” e 24 alelos “a”. A distribuição do número de heterozigotos  $\{8, 6, 4, 2, 0\}$ , condicionada à distribuição alélica, para os respectivos valores é dada por:  $\{0,313237; 0,487257; 0,182722; 0,016611; 0,000173\}$ . Desse modo, caso o nível de significância seja 5%, o ponto  $(4,0,12)$  pertence à região de rejeição, bem como os pontos  $\{(0,8,8), (3,2,11)\}$  aleatorizados em 0,111749 e 0,892361, respectivamente. Assim, a função crítica associada a esses pontos do espaço amostral é dada por:

$$\varphi(\mathbf{x}) = \begin{cases} 1, & \text{se } \mathbf{x} = (4,0,12), \\ 0,892361, & \text{se } \mathbf{x} = (3,2,11), \\ 0,111749, & \text{se } \mathbf{x} = (0,8,8), \\ 0, & \text{se } \mathbf{x} \in \{(2,4,10), (1,6,9)\}. \end{cases}$$

de tal modo que:  $0,05 \cong 0,111749 \cdot 0,313237 + 0,892361 \cdot 0,016611 + 0,000173$ .

O valor esperado do número de heterozigotos, condicionado à distribuição alélica, nesse caso, é dado por aproximadamente 6,2. Portanto:

$$(8-6,2) 0,111749 0,313237 + (2-6,2) 0,892361 0,0116611 + (0-6,2) 0,000173 \cong 0$$

Procedendo-se de modo análogo a todos os outros subconjuntos do espaço amostral, com outros valores alélicos, a região de rejeição pode ser construída com pontos de rejeição, não rejeição e aleatorizados.

A Figura 14 mostra a aplicação do mecanismo de aleatorização para o teste apresentado em Meulepas (1999) no caso particular de uma amostra com 16 indivíduos, a um nível de significância 5%.

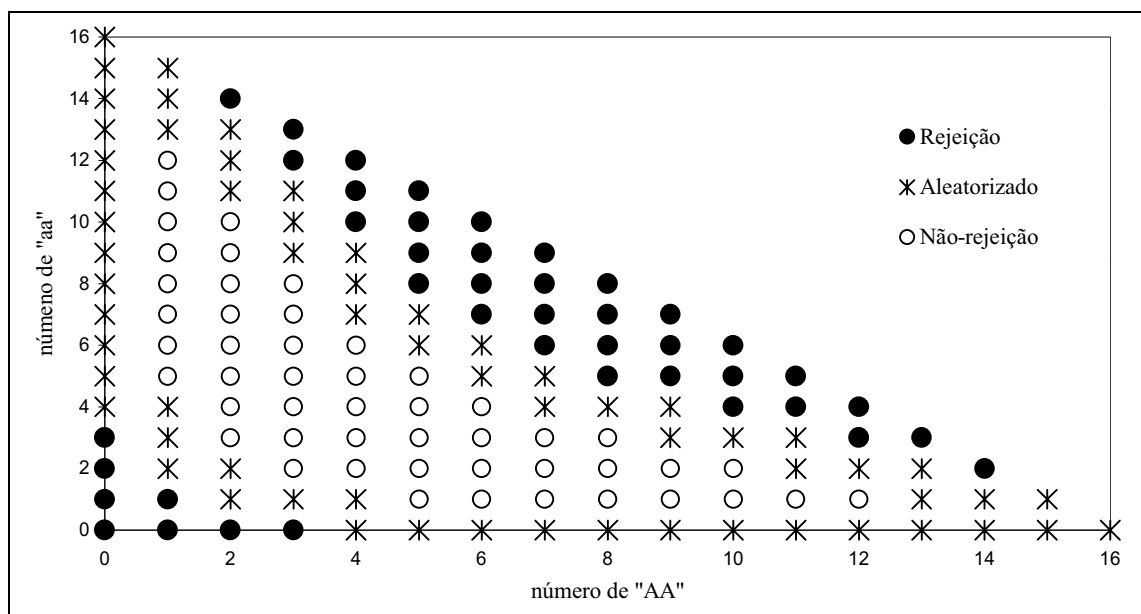


Figura 14 - Região de rejeição do teste não enviesado, com nível de significância 5%.

Uma comparação de poder para os testes, exato,  $X^2$  e não enviesado, já mencionados, com valores fixos de  $\gamma$  é apresentada nas Figuras 15 e 16.

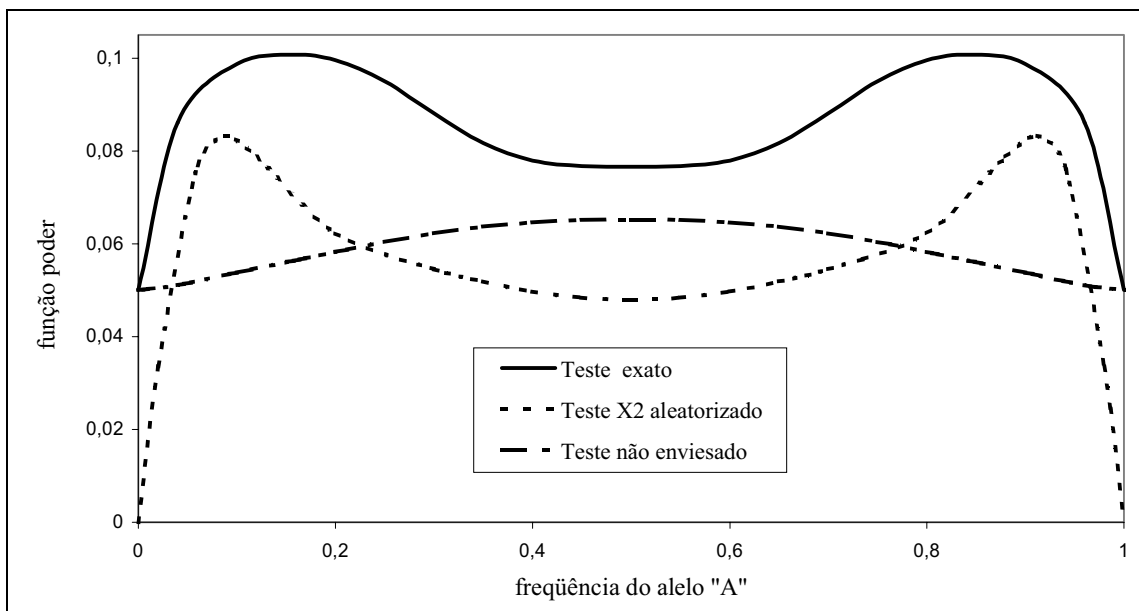


Figura 15 - Função poder dos testes aleatorizados no caso particular de  $\gamma = 0,1$ .

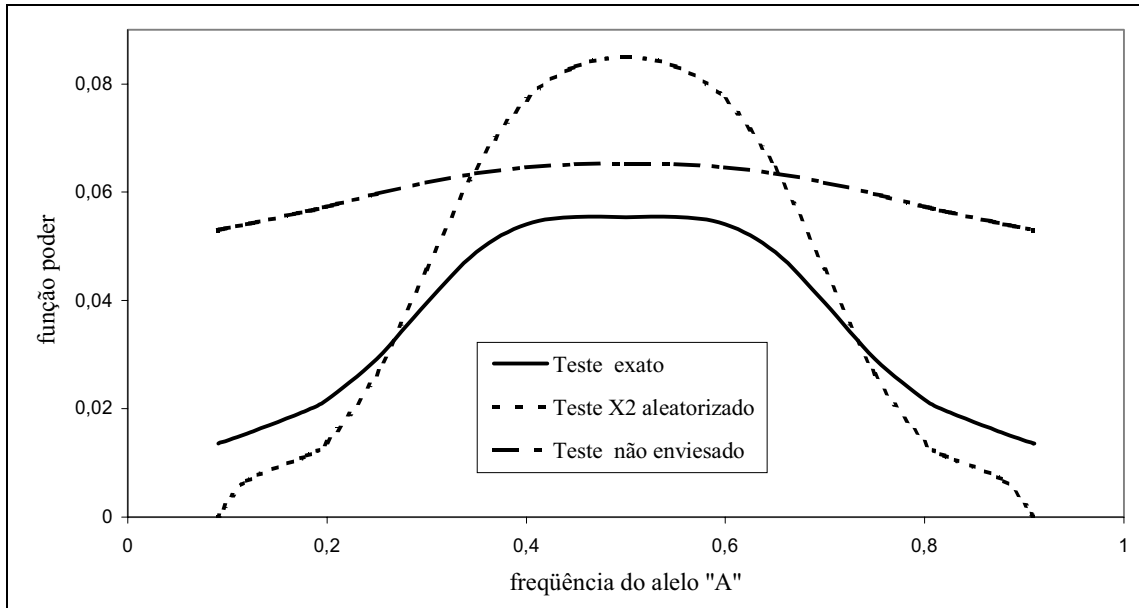


Figura 16 - Função poder dos testes aleatorizados no caso particular de  $\gamma = -0,1$ .

Na Figura 15, e mais evidenciado na Figura 16, pode ser notada a ausência de viés na proposição de teste apresentada em Meulepas (1999), o que não é

verificado nos outros testes. Porém, principalmente no caso de  $\gamma$  positivo, o poder apresentado no teste não enviesado é inferior ao poder apresentado no teste exato.

No teste não enviesado, no caso do nível de significância preestabelecido, tem-se muitos pontos do espaço amostral aleatorizados: sendo  $n$  o número de indivíduos selecionados aleatoriamente, o espaço amostral apresenta um total de  $(n+1)(n+2)/2$  pontos, dos quais  $4n$  aleatorizados. No teste exato, a quantidade de pontos aleatorizados diminui para  $2n + 1$ , porém ainda apresenta um número razoável de pontos aleatorizados. Assim, existe a necessidade de um procedimento que diminua a quantidade de pontos aleatorizados, sem contudo diminuir expressivamente a quantidade de pontos da região de rejeição encontrada no teste exato.

#### 4.7 Solução otimizada com máximo volume de poder

Para o teste de equilíbrio genético de Hardy-Weinberg com máximo número de pontos na região de rejeição, com um nível de significância preestabelecido  $\alpha$ , tem-se uma função crítica  $\varphi(\cdot)$ , satisfazendo a seguinte otimização:

$\sum_{\Omega} \varphi(x_1, x_2, x_3)$ , máximo,

$$\sum_{\Omega} \varphi(x_1, x_2, x_3) P_{\theta}[X_1 = x_1, X_2 = x_2, X_3 = x_3] \leq \alpha, \forall \theta \in \Theta_0.$$

Porém a solução  $\varphi^*(\cdot)$ , apresentada em 3.3, com a restrição do nível de significância homogêneo, tal como é apresentado no algoritmo adaptado de Tocher (1950), impõe a presença de muitos pontos aleatorizados na região de rejeição. Cabe então uma adaptação dessa solução que preserve o máximo possível a quantidade de pontos, bem como o nível de significância preestabelecido, ampliando-se a quantidade de pontos não aleatorizados ( $\varphi(x)=1$ ) na região de rejeição.

O teste  $\varphi(\cdot)$ , sugerido no presente trabalho, é definido então por intermédio da adaptação do algoritmo apresentado em Tocher (1950), com a inclusão de uma maior quantidade de pontos inteiros na região de rejeição. Tal adaptação, para o teste de equilíbrio genético de Hardy-Weinberg com nível de significância

preestabelecido  $\alpha$ , é dada por:

- i ) calcula-se o valor da função crítica  $\varphi(\cdot)$ , que inclui a maior quantidade de pontos na região de rejeição, associando-se  $\varphi(\mathbf{x}) = 1$  para os pontos em que  $\varphi^*(\mathbf{x}) = 1$ ;
- ii ) ordenam-se os pontos do espaço amostral, que estão associados a aleatorizações  $\Omega^*$  ( $\Omega^* = \{\mathbf{x} \in \Omega : 0 < \varphi^*(\mathbf{x}) < 1\}$ ), por intermédio do valor da função crítica;
- iii ) associa-se valor  $\varphi(\mathbf{x}) = 1$  para os pontos correspondentes aos maiores valores de  $\varphi^*(\mathbf{x})$ , em que  $\mathbf{x} \in \Omega^*$ , obedecendo à ordem decrescente, para todos os pontos cuja inclusão na região de rejeição não torna o teste com tamanho superior ao valor limite  $\alpha$ ;
- iv ) associa-se, ao ponto correspondente ao maior valor da função crítica que não foi incluído na região de rejeição por intermédio de iii, o maior valor possível, aleatorizado, da função crítica  $\varphi(\cdot)$ , de modo a ser obtido como tamanho do teste o nível de significância pretendido  $\alpha$ ;
- v ) associa-se aos pontos extremos  $(n,0,0)$  e  $(0,0,n)$  o mesmo valor da função crítica  $\varphi^*(\cdot)$ , dado por  $\alpha$ , para que seja preservado um poder mínimo aproximadamente  $\alpha$ , reduzindo o viés do teste  $\varphi(\cdot)$ , nas bordas do espaço paramétrico.

Como exemplo será aplicado esse mecanismo no caso particular, sendo considerado uma amostra com 16 indivíduos em que é observado um loco, cujos genes são formados por dois únicos alelos “A” e “a”.

O teste  $\varphi^*(\cdot)$  com a maior quantidade de pontos na região de rejeição apresenta a seguinte região de pontos aleatorizados, projetada no plano das freqüências genotípicas “AA” e “aa”,  $\Omega^* = \{(1, 11), (11, 1), (0, 7), (7, 0), (0, 14), (14, 0), (0, 15), (15, 0), (0, 16), (16, 0), (5, 6), (6, 5), (1, 12), (12, 1), (1, 4), (4, 1), (2, 10), (10, 2), (4, 8), (8, 4), (1, 13), (13, 1), (0, 5), (5, 0), (3, 9), (9, 3), (2, 11), (11, 2), (1, 14), (14, 1), (5, 7), (7, 5), (6, 6)\}$ , aos quais estão associados os seguintes valores da função crítica  $\varphi^*(\cdot)$ :  $\{0,0030609; 0,0030609; 0,013228; 0,013228; 0,0183333; 0,0183333; 0,05; 0,05; 0,05; 0,05; 0,0945497; 0,0945497; 0,115192; 0,115192; 0,153264; 0,153264; 0,181784; 0,181784; 0,198826; 0,198826; 0,249702; 0,249702; 0,352341; 0,352341; 0,427817;$



0,427817; 0,43994; 0,43994; 0,516667; 0,516667; 0,873815; 0,873815; 0,906686}, respectivamente.

Para o teste otimizado, com a maior quantidade de pontos não aleatorizados na região de rejeição, serão acrescidos na região de rejeição ( $\varphi(x) = 1$ ) os pontos:  $\{(0, 5), (5, 0), (2, 11), (11, 2), (1, 14), (14, 1), (5, 7), (7, 5), (6, 6)\}$ . Ainda serão inclusos, aleatorizados, os pontos:  $\{(3, 9), (9, 3), (0, 16), (16, 0)\}$ , com valores de  $\varphi(\cdot)$ , dados respectivamente por:  $\{0,25; 0,25; 0,05; 0,05\}$ .

A Figura 17 mostra os pontos, pertencentes à região de rejeição, do teste otimizado com a maior quantidade de pontos inteiramente na região de rejeição no caso particular de uma amostra com 16 indivíduos, com um nível de significância 5%.

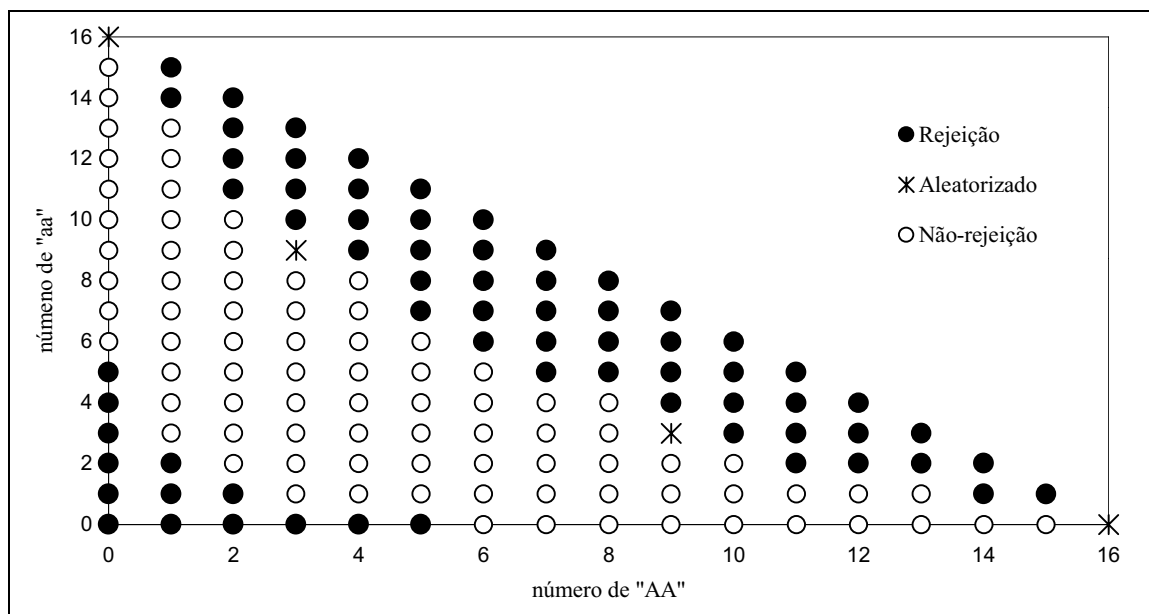


Figura 17 - Região de rejeição do teste teste otimizado, com nível de significância 5%.

Estudando-se a função poder para casos particulares de  $\gamma$ , pode-se notar uma semelhança de eficiência nos testes exato e otimizado. Isso se deve ao fato da região de rejeição definida por tais testes conter muitos pontos em comum, sendo compensada a ausência de alguns pontos aleatorizados no teste otimizado, pela presença completa de pontos que fazem parte apenas parcialmente do teste exato. Uma

comparação do poder para os testes, com valores fixos de  $\gamma$  é apresentada nas Figuras 18 e 19.

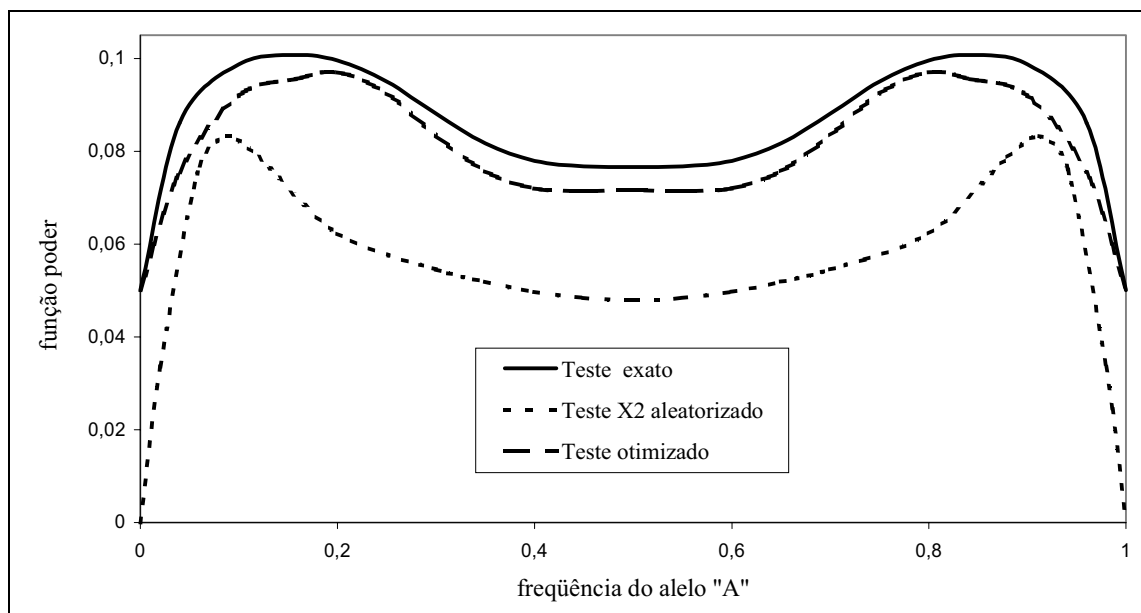


Figura 18 - Função poder dos testes aleatorizados no caso particular de  $\gamma = 0,1$ .

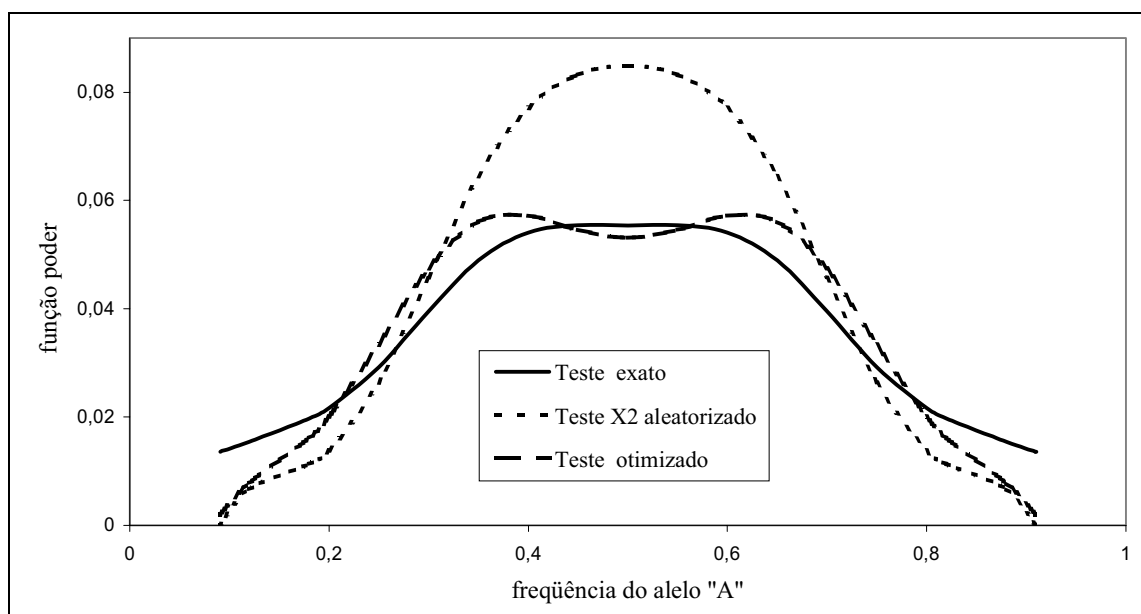


Figura 19 - Função poder dos testes aleatorizados no caso particular de  $\gamma = -0,1$ .

Assim, o teste exato aleatorizado, estabelecido como uma redefinição do teste apresentado em Elston & Forthofer (1977), por intermédio da adaptação do algoritmo de Tocher (1950) para o teste de equilíbrio genético de Hardy-Weinberg, que apresenta o máximo de pontos na região de rejeição, é visto, dentre os procedimentos com uniformidade de nível de significância, como a melhor opção de procedimento de construção da função crítica. Entretanto, por apresentar um grande número de pontos aleatorizados na região de rejeição, torna-se de difícil aplicação prática.

A sugestão que defende-se, no presente trabalho, mantém parcialmente a propriedade de nível de significância homogêneo, reduzindo o viés apresentado em testes com utilização de estatísticas, tal como a  $X^2$ , e contempla o número máximo de pontos na região de rejeição. Podendo, assim, ser sugerido, como uma nova opção viável de teste.

## 5. CONCLUSÕES

Muitos trabalhos, tais como: Levene (1949), Haldane (1954), Chapco (1976), Elston & Forthofer (1977), Meulepas (1999) e Montoya-Delgado et al. (2001), foram apresentados com formas alternativas à utilização da estatística  $X^2$ , para testar o equilíbrio genético de Hardy-Weinberg. Tais procedimentos são utilizados no cálculo do nível de significância ( $p$ -valor) para os pontos pertencentes ao espaço amostral, sendo os dois primeiros com cálculo assintótico. Quando tais procedimentos são tomados com um nível de significância preestabelecido, e o espaço amostral discriminado em pontos de rejeição e não-rejeição, existe uma grande semelhança na discriminação do espaço amostral por intermédio da utilização da estatística  $X^2$ .

No trabalho de Tocher (1950), a discrepância em relação à hipótese em teste é medida em subconjuntos do espaço amostral, sendo acrescida a idéia de aleatorização na função crítica. Tal procedimento, necessário para tornar o nível de significância homogêneo e exato, qualquer que seja o valor paramétrico correspondente à hipótese em teste, diminui o viés encontrado em testes dessa natureza. Esse procedimento é perfeitamente adaptável no teste de equilíbrio genético de Hardy-Weinberg, porém contém um número muito grande de pontos aleatorizados, o que dificulta sua aplicação em trabalhos práticos.

Sendo aplicado o procedimento proposto em Tocher (1950) em trabalhos como Elston & Forthofer (1977) e Meulepas (1999) submetidos a aleatorizações, estes apresentam algumas propriedades de grande interesse, tais como a uniformidade do nível de significância do teste, qualquer que seja o valor paramétrico pertencente à região relacionada ao equilíbrio, ou, no caso do procedimento apresentado em Meulepas (1999), o teste não enviesado.

No presente trabalho foi sugerida uma forma de diminuir a quantidade de pontos aleatorizados, quando na adaptação do procedimento de teste apresentado em Tocher (1950), na aplicação do teste de equilíbrio genético de Hardy-Weinberg, em que foi construída uma forma de obter-se a maior quantidade de pontos inteiros na região de rejeição, buscando a propriedade de maior volume da função poder. Tal procedimento contempla a maior quantidade de combinações genótípicas, associadas ao não-equilíbrio genético, qualquer que seja o nível de significância preestabelecido do teste.

Com este trabalho buscou-se uma adaptação de propriedades teóricas, desejáveis a um teste de hipóteses tais como a homogeneidade do nível de significância, bem como o não-viés, com características procuradas na formulação de trabalhos práticos, tal como a ausência de pontos sujeitos a aleatorizações, além de lançar um questionamento dos procedimentos tomados como exatos, que não estão associados a necessidades de aleatorizações.

Assim, sendo aceito o princípio de aleatorização como uma forma natural de procedimento em testes discretos, pode-se sugerir a adaptação da aleatorização exposta em Tocher (1950), na busca por um menor viés, ou a eliminação por completa do viés, no teste de equilíbrio genético de Hardy-Weinberg. Esse procedimento pode contornar, mesmo que parcialmente, problemas nas bordas do espaço paramétrico, que correspondem à presença de alelos raros.

Sendo, porém, a aleatorização descartada, a sugestão de otimização inteira, com a inclusão da maior quantidade de pontos na região de rejeição, foi apresentada como o melhor procedimento para aplicação prática, no teste de equilíbrio genético de Hardy-Weinberg, quando existe a necessidade do nível de significância preestabelecido.

A eficiência do teste otimizado foi basicamente a mesma que o teste exato, apresentado em Elston & Forthofer (1977), que é defendido em trabalhos tais como: Louis & Dempster (1987), Guo & Thompson (1992) e Weir (1996) como inquestionavelmente superior a qualquer outros testes, porém, sem a desvantagem de muitas aleatorizações, o que torna sua aplicação prática mais eficiente.

## REFERÊNCIAS BIBLIOGRÁFICAS

- CHAPCO, W. An exact test of the Hardy-Weinberg law. **Biometrics**, v.32, p.183-189, 1976.
- CHAPMAN, J.W. A comparison of the  $X^2$ ,  $-2\log R$ , and multinomial probability criteria for significance tests when expected frequencies are small. **Journal of the American Statistical Association**, v.71, p.854-863, 1976.
- COCHRAN, W.G. The  $\chi^2$  tests of goodness of fit. **Annals of Mathematical Statistics**, v.23, p.315-345, 1952.
- COHEN, A.; SACKROWITZ, H.B. An evaluation of some tests of trend in contingency tables. **Journal of the American Statistical Association**, v.87, p.470-475, 1992.
- CRESSIE, N.A.C.; READ, T.R.C. Multinomial goodness-of-fit tests. **Journal of the Royal Statistical Society B**, v.46, p.440-464, 1984.
- CRAMÉR, H. **Mathematical methods of statistics**. Princeton: Princeton University Press, 1946. 575p.
- ELSTON, R.C.; FORTHOFFER, R. Testing for Hardy-Weinberg equilibrium in small samples. **Biometrics**, v.33, p.536-542, 1977.

- EMIGH, T.H. A comparison of tests for Hardy-Weinberg equilibrium. **Biometrics**, v.36, p.627-642, 1980.
- FISHER, R.A. On the interpretation of  $\chi^2$  from contingency tables, and the calculation of P. **Journal of the Royal Statistical Society**, v.85, p.87-94, 1922.
- FISHER, R.A. The conditions under which chi square measures the discrepancy between observation and hypothesis. **Journal of the Royal Statistical Society**, v.87, p.442-450, 1924.
- FISHER, R.A. The logic of inductive inference. **Journal of the Royal Statistical Society**, v.98, p.34-54, 1935.
- GUO, S.W.; THOMPSON, E.A. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. **Biometrics**, v.48, p.361-372, 1992.
- GURIAN, J.M.; CORNFIELD, J.; MOSIMANN, J.E. Comparisons of power for some exact multinomial significance tests. **Psychometrika**, v.29, p.409-419, 1964.
- HALDANE, J.B.S. An exact test for randomness of mating. **Journal of Genetics**, v.52, p.631-635, 1954.
- LANCASTER, H.O. Significance tests in discrete distributions. **Journal of the American Statistical Association**, v.56, p.223-234, 1961.
- LEHMANN, E.L. **Testing Statistical Hypoteses**. New York: Jonh Wiley, 1959. 369p.
- LEVENE, H. On a matching pobleml arising in genetics. **Annals of Mathematical Statistics**, v.20, p.91-94, 1949.

- LOUIS, E.J.; DEMPSTER, E.R. An exact test for Hardy-Weinberg and multiple alleles. **Biometrics**, v.43, p.805-811, 1987.
- MAISTE, P.J.; WEIR, B.S. A comparison of tests for independence in the FBI RFLP data bases. **Genetica**, v.96, p.125-138, 1995.
- MEULEPAS, E. A p-value for an exact two-sided test of Hardy-Weinberg equilibrium. **Biometrical Journal**, v.41, p.499-505, 1999.
- MONTOYA-DELGADO, L.E.; IRONY, T.Z.; PEREIRA, C.A.B.; WHITTLE, M.R. An unconditional exact test for the Hardy-Weinberg equilibrium law: sample-space ordering using the bayes factor, **Genetics**, v.158, p.875-883, 2001.
- NEMHAUSER, G.L.; WOLSEY, L.A. **Integer and combinatorial optimization**. New York: Jonh Wiley, 1988. 763p.
- PEARSON, K. On the criterion that a given system of deviations from the probable in the case of a correlated system of a variables in such that in can be reasonably supposed to have arisen from random sampling. **Philosophy Magazine**, v.50, p.157-172, 1900.
- RADLOW, R.; ALF, E.F. An alternate multinomial assessment of the accuracy of the  $\chi^2$  test of goodness of fit. **Journal of the American Statistical Association**, v.70, p.811-813, 1975.
- READ, T.R.C.; CRESSIE, N.A.C. **Goodness-of-fit statistics for discrete multivariate data**. New York: Springer-Verlag, 1988. 211p.



TATE, M.W.; HYER, L.A. Inaccuracy of the  $\chi^2$  test of goodness of fit when expected frequencies are small. **Journal of the American Statistical Association**, v.68, p.503-511, 1973.

TOCHER, K.D. Extension of the Neyman-Pearson theory of tests to discontinuous variates. **Biometrika**, v.37, p.130-144, 1950.

WEIR, B.S. **Genetic Data Analysis II**. Sunderland: Sinauer, 1996. 445p.