

**PLANO DE AMOSTRAGEM SEQUENCIAL
PARA A FERRUGEM DO CAFEIEIRO
(*HEMILEIA VASTATRIX* BERK. & BR.)
CONSIDERANDO A DISTRIBUIÇÃO
BETA-BINOMIAL**

LÍVIA TERRA VIEIRA PENIDO

2006

LÍVIA TERRA VIEIRA PENIDO

**PLANO DE AMOSTRAGEM SEQUENCIAL PARA A FERRUGEM
DO CAFEIEIRO (*HEMILEIA VASTATRIX* BERK. & BR.)
CONSIDERANDO A DISTRIBUIÇÃO BETA-BINOMIAL**

Dissertação apresentada à Universidade Federal de Lavras como parte das exigências do Programa de Pós-Graduação em Agronomia, Área de Concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de "Mestre".

Orientador

Prof. Dr. Paulo César Lima

LAVRAS
MINAS GERAIS - BRASIL
2006

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da
Biblioteca Central da UFLA**

Penido, Livia Terra Vieira

Plano de amostragem seqüencial para a ferrugem do cafeeiro (*Hemileia vastatrix* Berk. & Br.) considerando a distribuição beta-binomial / Livia Terra Vieira Penido. - Lavras : UFLA, 2006.

75 p. : il.

Orientador: Paulo César Lima
Dissertação (Mestrado) – UFLA.
Bibliografia.

1. Padrão espacial. 2. Distribuições de probabilidade. 3. Amostragem seqüencial. I. Universidade Federal de Lavras. II. Título.

CDD-519.52

LÍVIA TERRA VIEIRA PENIDO

**PLANO DE AMOSTRAGEM SEQUENCIAL PARA A
FERRUGEM DO CAFEIEIRO (*HEMILEIA*
VASTATRIX BERK. & BR.) CONSIDERANDO A
DISTRIBUIÇÃO BETA-BINOMIAL**

Dissertação apresentada à Universidade Federal de Lavras como parte das exigências do Programa de Pós-Graduação em Agronomia, Área de Concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de "Mestre".

Prof. Dr. Renato Ribeiro de Lima	UFLA
Prof. Dr. Julio Neil Cassa Louzada	UFLA
Prof.Dr. Rodrigo Luz da Cunha	EPAMIG

Prof. Dr. Paulo César Lima

UFLA

(Orientador)

LAVRAS
MINAS GERAIS - BRASIL

À Deus pela grandeza de seu nome e
poder de sua misericórdia
Aos amigos, mestres e familiares pelo estímulo
que alimentou a energia necessária à
conclusão deste trabalho,

OFEREÇO

À minha querida mãe, Constança Terra, pelo incentivo, sem o qual nada seria possível.
Aos meus queridos irmãos Maria Alice e Mozar (in memorian) pelo amor e apoio,
À minha adorada filha Débora e meu querido esposo Epifânio pelo amor e compreensão,

DEDICO.

AGRADECIMENTOS

À Universidade Federal de Lavras (UFLA), em especial ao Departamento de Ciências Exatas (DEX), pelo curso oferecido.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa de estudo.

A todos os professores do DEX-UFLA pela contribuição à minha formação acadêmica, em especial aos professores Paulo César Lima pela orientação, e Renato Ribeiro de Lima pela colaboração carinhosa e espontânea.

Às minhas professoras da graduação Elise Menicucci Nogueira e Rosângela Moura Cortez pelo estímulo e apoio.

A todos meus colegas de curso, em especial, Nádia e Verônica pela amizade e colaboração no convívio durante o curso.

Aos pesquisadores da Empresa de Pesquisa Agropecuária de Minas Gerais (EPAMIG), Julio César de Souza, Sara Maria Chalfoun, e aos amigos Devanil, Lúcia, Maria Imaculada e Carla pelo apoio na condução e desenvolvimento da pesquisa.

Ao meu amigo Carlos Antônio Bortollus, pela lição profissional ensinada.

A família Mendes pelo carinho e apoio.

Aos meus pais, João Vieira (in memorian) e Constança Terra, pelo amor, estímulo, estrutura financeira e emocional fornecidos ao longo de minha vida.

Aos meus queridos irmãos Moises, Maria Alice pelo amor e apoio recebidos e em especial ao meu irmãozinho Mozar (in memoriam) que tendo nos deixado tão cedo, encheu meu coração de prazer e alegria com a sua nobre presença e com as notas musicais de suas composições.

A todos os meus sobrinhos, em especial ao Felipe e Bruno pela disposição e amizade.

Ao meu marido Epifânio e minha adorada filha Débora pela compreensão e amor

A Deus, que na verdade é o Princípio, pois sem Ele nada do que foi dito anteriormente teria sido possível.

SUMÁRIO

LISTA DE TABELAS

LISTA DE FIGURAS

RESUMO.....	i
ABSTRACT.....	ii
1 INTRODUÇÃO.....	1
2 REFERENCIAL TEÓRICO.....	3
2.1 Padrão espacial de doenças em plantas.....	3
2.2 A distribuição beta-binomial.....	14
2.3 Aplicações da distribuição beta-binomial.....	24
2.4 Métodos de amostragem.....	26
2.4.1 Amostragem seqüencial	26
2.4.2 Amostragem para a ferrugem do cafeeiro.....	27
2.4.2.1 Fatores que determinam a severidade da doença.....	28
2.4.2.2 Monitoramento.....	30
2.5 Planos de amostragem seqüencial	31
2.5.1 Plano de amostragem seqüencial baseado no teste seqüencial da razão de verossimilhança.....	32
2.5.2 Lei de potência binária.....	39
2.5.3 Aproximação para as linhas de decisão da distribuição beta- binomial.....	41
2.5.4 Curva característica de operação.....	44
3 MATERIAL E MÉTODOS.....	46
3.1 Dados reais.....	46
3.2 Dados simulados.....	49
3.3 Planos de amostragem seqüencial.....	49

4 RESULTADOS E DISCUSSÃO.....	51
4.1 Padrão espacial da ferrugem do cafeeiro para dados reais.....	51
4.2 Padrão espacial da ferrugem do cafeeiro para dados simulados.....	57
4.3 Planos de amostragem seqüencial.....	59
4.3.1 Plano de amostragem seqüencial ano de baixa carga.....	59
4.3.2 Plano de amostragem seqüencial ano de alta carga.....	63
4.4 Curva característica de Operação.....	67
5 CONCLUSÕES.....	70
6 REFERÊNCIAS BIBLIOGRÁFICAS.....	71

LISTA DE TABELAS

TABELA		Página
TABELA 1	Estimativa dos parâmetros π e θ em duas épocas amostradas.....	51
TABELA 2	Resultados do primeiro teste de homogeneidade de variâncias e índices de dispersão, nas duas épocas amostradas.....	52
TABELA 3	Testes de aderência qui-quadrado para as distribuições beta-binomial e binomial, nas duas épocas amostradas.....	53
TABELA 4	Estimativas dos valores de $\ln(A_x)$, A_x , A e b da lei de potência binária.....	57
TABELA 5	Caracterização do padrão espacial dos conjuntos simulados utilizando π e θ nas duas épocas amostradas.....	58
TABELA 6	Probabilidade de aceitação (Pac) de H_o em função da proporção de folhas infectadas.....	67

LISTA DE FIGURAS

FIGURA	Página
FIGURA 1 Disposição espacial esquemática de plantas infectadas por uma doença fictícia.....	4
FIGURA 2 Número de plantas infectadas em cada unidade amostral da Figura 1.....	4
FIGURA 3 Representação esquemática dos padrões de distribuição, em que cada círculo representa uma planta doente: (a) Padrão regular; (b) Padrão Aleatório e (c) Padrão Agregado.....	5
FIGURA 4 Representação de uma floresta Poisson (a) e de uma distribuição Neyman tipo A (b).....	9
FIGURA 5 Gráfico da função de densidade beta para vários valores dos parâmetros α e β	15
FIGURA 6 Imagem de satélite QuickBird® 27/05/2003 referente à área estudada. UFLA, Lavras, MG, 2006. Composição colorida 321 e infravermelha falsa cor 432.....	46
FIGURA 7 Representação bidimensional dos pontos amostrais.....	47
FIGURA 8 Distribuição de frequência observadas e esperadas das folhas com ferrugem segundo a distribuição beta-binomial. Amostra julho/2005.....	54
FIGURA 9 Distribuição de frequência observadas e esperadas das folhas com ferrugem segundo a distribuição binomial. Amostra julho/2005.....	54
FIGURA 10 Distribuição de frequência observadas e esperadas das folhas com ferrugem segundo a distribuição beta-binomial. Amostra julho/2006.....	56
FIGURA 11 Distribuição de frequência observadas e esperadas das folhas com ferrugem segundo a distribuição binomial. Amostra julho/2006.....	56
FIGURA 12 Linhas de decisão – número acumulado de folhas infectadas T_N , versus número de unidades amostrais N para a distribuição beta-binomial (equações 23 á 25).....	60
FIGURA 13 Linhas de decisão – número acumulado de folhas infectadas T_N , versus número de unidades amostrais N para a	

distribuição beta-binomial (equações 23 a 25) $\alpha = \beta = 0,05$	60
FIGURA 14 Ficha de campo para amostragem seqüencial da ferrugem do cafeeiro, para ano de baixa carga pendente.....	63
FIGURA 15 Linhas de decisão – número acumulado de folhas infectadas T_N , versus número de unidades amostrais N para a distribuição beta-binomial (equações 23 a 25) $\alpha = \beta = 0,05$	64
FIGURA 16 Linhas de decisão utilizando-se a distribuição beta-binomial e distribuição binomial para $\pi_0 = 0.025$ e $\pi_1 = 0.075$	65
FIGURA 17 Ficha de campo para amostragem seqüencial da ferrugem do cafeeiro, para o ano de alta carga pendente.....	66
FIGURA 18 Curva característica da operação simulado pelo método de Monte Carlo utilizando as estimativas de π e θ época amostrada julho/2005 e julho/2006	69

RESUMO

PENIDO, L.T.V. **Plano de amostragem seqüencial para a ferrugem do cafeeiro (*Hemileia vastatrix* Berk. & Br.) considerando a distribuição beta-binomial.** 2006. 75 p. Dissertação (Mestrado em Agronomia/Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras, MG. *

Planos de amostragem seqüencial possibilitam, na maioria das vezes, decisões mais rápidas, baratas e precisas sobre a aplicação de medidas de controle de doenças de plantas. A construção destes planos tem como requisito básico a caracterização do padrão espacial da doença. Na avaliação do padrão espacial, diferentes métodos têm sido utilizados dentre os quais o ajuste de distribuição, como por exemplo a distribuição beta-binomial, o estudo da relação variância média e o estudo da lei de potência binária. O Brasil é o maior produtor e também o segundo maior consumidor de café mundial. A ferrugem causada por *Hemileia vastatrix* ataca as plantas de café em todas as regiões do mundo onde é cultivado, sendo que no Brasil ela representa em média um prejuízo na ordem de 35%. Para ilustração das teorias apresentadas foi proposta a caracterização do padrão espacial e a construção de planos de amostragem seqüencial baseada na metodologia do teste seqüencial da razão de probabilidade da ferrugem do cafeeiro utilizando a distribuição beta-binomial. Foram utilizados dados de incidência da ferrugem do cafeeiro de uma lavoura da cultivar Mundo Novo, com quinze anos de idade, localizada no município de Ijací, sul de Minas Gerais. A incidência da doença foi avaliada em duas épocas, sendo que a distribuição beta-binomial se ajustou bem aos dados, indicando um padrão espacial agregado. Para a obtenção das estimativas dos parâmetros da lei de potência binária, foram simulados conjuntos com valores das estimativas dos parâmetros dos dados reais, no qual verificou-se padrão agregado em uma data amostrada e padrão aleatório em outra. As estimativas dos parâmetros da lei de potência binária foram utilizados na construção das linhas de decisão da beta-binomial do plano de amostragem seqüencial. Verificou-se nos planos desenvolvidos ser necessário um pequeno número de unidades amostrais na tomada de decisão, pelo motivo da infecção estar muito acima do nível de controle. Através de simulações da curva característica de operação, avaliou-se o plano de amostragem seqüencial, verificando-se as probabilidades da correta aceitação da hipótese de nulidade, em função de diferentes valores de níveis de infecção, sendo que quando estes se afastam do nível de controle as decisões são tomadas com menor risco.

* Comitê Orientador: Paulo César Lima – UFLA (orientador) e Renato Ribeiro de Lima UFLA.

ABSTRACT

PENIDO, L.T.V. **Sequential sampling for coffee plantation rust (*Hemileia vastatrix* Berk. & Br.) by taking into account the beta-binomial distribution.** 2006. 75p. Dissertation. (Master in Agronomy/Statistics and Agricultural Experimentation)- Federal University of Lavras, Lavras, Minas Gerais, Brazil.

Sequential sampling plans make it possible, most of the time, to make decisions more precisely, faster and at a lower cost regarding the application of control measures on plant diseases. The elaboration of these plans has as basic requirement the disease's space pattern characterization. For the space pattern evaluation, various methods have been applied, namely the variance mean relation study, the binary potency law study and the distribution adjustment such as the beta-binomial distribution. Brazil is the world's biggest producer as well as the second largest coffee consumer. The rust caused by *Hemileia vastatrix* afflicts coffee crops worldwide, and it represents an average loss of 35% in Brazil. A characterization of the space pattern and the making of a sequential sampling plan based on the methodology of the sequential probability ratio test for coffee rust using beta-binomial distribution, was applied to illustrate the theories presented in this paper. Rust incidence data on coffee plantations were taken from fifteen-years old "Mundo Novo" crops located in Ijaci, south of the state of Minas Gerais, Brazil. Rust incidence was evaluated in July/2005 and in July/2006, periods during which the beta-binomial distribution adjusted to the data which characterized an aggregated space pattern. In order to obtain the binary power law's parameter estimates, sets containing estimate values from the original data parameters, were simulated, in which a clustered pattern was observed for the first period (July/2005) and a random patterns for the second period (July/2006). The binary power law's parameter estimates were used in the making of the sequential sampling plan's beta-binomial decision lines. A small sampling unit number was needed on the decision-making, as the existing infection level was above control level. The sequential sampling plan was evaluated by means of operation characteristic curves simulation. The null hypothesis correct acceptance was observed according to different infection levels, besides, as the infection levels depart from the control levels, the decision-makings are made with less risk.

* Advising Committee: Orientador: Paulo César Lima – UFLA (Adviser) and Renato Ribeiro de Lima UFLA.

1 INTRODUÇÃO

A distribuição beta-binomial tem sido utilizada na área de fitopatologia para descrever o comportamento espacial de determinadas doenças quando se tem um padrão espacial agregado. A característica principal das distribuições agregadas é que são observados focos da doença fazendo com que a probabilidade de incidência não seja constante na lavoura.

O conhecimento da distribuição de probabilidade que descreva o comportamento espacial de uma doença permite sua utilização na amostragem seqüencial que tem, como principal vantagem, propiciar que a tomada de decisão sobre o momento de se efetuar o controle da doença seja feita com maior rapidez e com conseqüente redução de custos.

A amostragem seqüencial é indicada principalmente para doenças que requerem monitoramento constante de seus níveis populacionais principalmente em culturas cujo valor econômico é alto, como a cafeicultura no Brasil.

O Brasil é o maior produtor e também o segundo maior consumidor, de café no mundo e existem várias doenças que causam prejuízos a esta cultura. A ferrugem do cafeeiro é considerada uma das principais doenças e os prejuízos causados nas regiões cafeeiras do Brasil, onde as condições climáticas são favoráveis à doença, chegam a atingir em média 35% a 50%.

Desta forma, considerando a importância da cafeicultura no cenário econômico internacional e nacional, os prejuízos decorrentes da ferrugem causada pelo fungo *Hemileia vastatrix* Berk. & Br. nas lavouras de café e o excesso de fungicidas muitas vezes utilizados para o controle desta doença, justifica-se plenamente a necessidade de se estabelecer um plano de amostragem seqüencial para o seu monitoramento e controle, de forma a reduzir e aplicar racionalmente os fungicidas nas lavouras de café.

A água é o principal agente de disseminação local do fungo da ferrugem do cafeeiro, pois os respingos de chuva proporcionam a disseminação para outras plantas formando focos de doença, o que pode caracterizar um padrão agregado.

Este trabalho teve por objetivo utilizar a distribuição beta-binomial para elaborar um plano de amostragem seqüencial, pelo método do teste seqüencial da razão de probabilidade de Wald, para a ferrugem do cafeeiro.

2 REFERENCIAL TEÓRICO

Nesta seção é apresentado o embasamento teórico de algumas distribuições de probabilidade freqüentemente usadas para descrever o padrão espacial de doenças de plantas destacando-se, neste trabalho, a distribuição beta-binomial. É revisada a técnica da amostragem seqüencial e sua possível utilização para a determinação de linhas de decisão para a ferrugem do cafeeiro.

2.1 Padrão espacial de doenças em plantas

Segundo Campbell & Madden (1990), costuma-se classificar o padrão espacial de doenças de plantas em três grupos básicos denominados regular, aleatório e agregado. Uma das formas apropriadas para caracterizar esses padrões, através da Estatística, é a utilização de distribuições de probabilidade. Basicamente esta metodologia consiste na comparação das freqüências observadas com as freqüências esperadas, obtidas com base em alguma distribuição de probabilidade utilizando-se um teste de aderência. A grande dificuldade está em se determinar a distribuição de probabilidade mais adequada.

Os padrões espaciais de uma doença em plantas podem ser descritos através de um exemplo bem simples, tomando como base a Figura 1. Nesta figura está representado um talhão ou uma lavoura em que a área total foi dividida em unidades amostrais, representadas pelos retângulos menores podendo, cada unidade amostral, referir-se a uma planta, uma folha, um grupo de folhas ou de plantas, dentre outras.

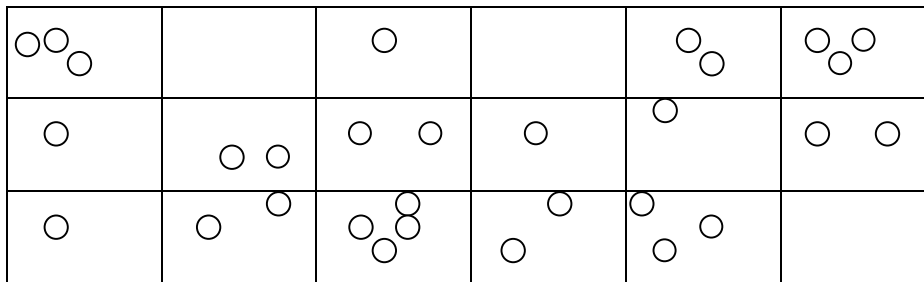


FIGURA 1 Disposição espacial esquemática de plantas infectadas por uma doença fictícia.

Os círculos representam as plantas infectadas em cada unidade amostral. Como o número de plantas observadas em cada uma dessas unidades amostrais é variável e considerando-se que a cada unidade corresponde uma variável aleatória, para as quais admite-se a mesma distribuição o que permite visualizar um tratamento matemático que descreva o comportamento espacial da doença.

Para a situação fictícia apresentada na Figura 1, têm-se os números de plantas infectadas em cada unidade amostral apresentados na Figura 2.

3	0	1	0	2	3
1	2	2	1	1	2
1	2	4	2	3	0

FIGURA 2 Número de plantas infectadas em cada unidade amostral da Figura 1

Segundo Campbell & Madden (1990), no padrão de distribuição espacial uniforme (Figura 3a) existe a tendência das plantas doentes e sadias ocorrerem de forma sistemática. Este padrão pode ser provocado com o plantio alternado de variedades resistentes e suscetíveis, por exemplo. Neste padrão, a variância é menor do que a média, o que indica uma sub-dispersão espacial.

Segundo Bergamin Filho et al. (2002), o padrão aleatório (Figura 3b), está relacionado a iguais oportunidade de infecção sendo que a presença de uma planta doente não afeta a ocorrência de outra planta doente na vizinhança, não havendo, portanto, dependência entre as ocorrências. Madden & Hughes (1995) afirmam que muitos pesquisadores assumem geralmente que este padrão pode ser descrito, apropriadamente, pela distribuição de Poisson, em que a variância é igual à média.

Segundo Lima (2005), a associação da distribuição de Poisson com o padrão espacial aleatório pode ser explicada considerando-se um raciocínio semelhante ao utilizado na obtenção da distribuição de Poisson a partir da distribuição binomial, conforme apresentado por Upton & Fingleton (1985).

Seja uma região com área A e com x pontos ou x folhas doentes, distribuídas aleatoriamente. A intensidade de pontos por área ou de folhas

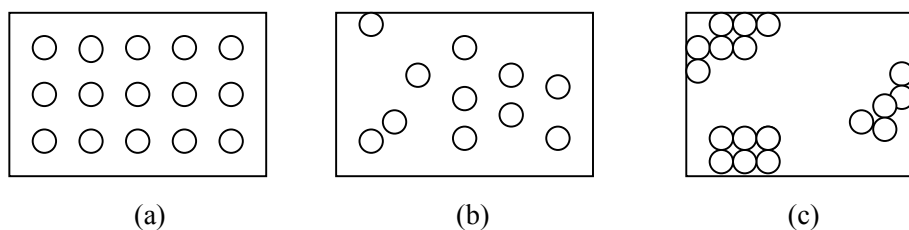


FIGURA 3 Representação esquemática dos padrões de distribuição, em que cada círculo representa uma planta doente: (a) Padrão Regular; (b) Padrão Aleatório e (c) Padrão Agregado.

doentes por região será dada por $\lambda = \frac{x}{A}$. Admitindo-se que a área A seja composta por um grande número n de sub-regiões extremamente pequenas, de maneira que a probabilidade p de que cada uma dessas sub-regiões contenha mais de uma folha doente seja desprezível, tem-se que $n \rightarrow \infty$ e $p \rightarrow 0$. Considerando-se que cada uma dessas x folhas doentes tenha a mesma probabilidade p de estar localizada em qualquer uma das n sub-regiões, então o padrão aleatório implica em que as folhas doentes possam existir em qualquer lugar dentro da região A de forma que cada uma das sub-regiões tenha uma probabilidade igual de contê-las.

Pode-se assumir que, obtidos x sucessos em n realizações de um ensaio e sendo X a variável aleatória que representa esse número de sucessos, X terá uma distribuição binomial ou seja $X \sim B(n, p)$. Pode-se escrever:

$$\begin{aligned}
 P[X = x] &= \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} I_{\{0,1,\dots,n\}}(x) \\
 &= \frac{n(n-1)(n-2)\dots(n-x+1)(n-x)!}{x!(n-x)!} p^x (1-p)^{n-x} \\
 P[X = x] &= \frac{n(n-1)\dots(n-x+1)}{x!} p^x (1-p)^{n-x} . \tag{1}
 \end{aligned}$$

Seja $\lambda = np$, o número esperado de sucessos por unidade de espaço, ou seja, a intensidade ou o número de pontos por unidade de área. Fazendo $p = \frac{\lambda}{n}$,

$1-p = \frac{n-\lambda}{n}$ e substituindo na equação 1, tem-se:

$$P[X = x] = \frac{n(n-1)\dots(n-x+1)}{x!} \left[\frac{\lambda}{n} \right]^x \left[\frac{n-\lambda}{n} \right]^{n-x}$$

$$P[X = x] = \frac{n(n-1)\dots(n-x+1)}{x!} \left[\frac{\lambda}{n} \right]^x \left[\frac{n-\lambda}{n} \right]^n \left[\frac{n-\lambda}{n} \right]^{-x}$$

$$P[X = x] = \frac{\lambda^x n(n-1)\dots(n-x+1)}{x! n^x} \left[1 - \frac{\lambda}{n} \right]^x \left[1 - \frac{\lambda}{n} \right]^{-x}. \quad (2)$$

Observa-se que:

$$\begin{aligned} \frac{n(n-1)\dots(n-x+1)}{n^x} &= \frac{(n-1)\dots(n-x+1)}{n^{-1}n^x} \\ &= \frac{(n-1)\dots[n-(x-1)]}{n^{x-1}} \\ &= \frac{(n-1)}{n} \frac{(n-2)}{n} \dots \frac{[n-(x-1)]}{n} \\ &= \left[1 - \frac{1}{n} \right] \left[1 - \frac{2}{n} \right] \dots \left[1 - \frac{(x-1)}{n} \right]. \end{aligned} \quad (3)$$

Usando 3 em 2, tem-se:

$$P[X = x] = \frac{\lambda^x}{x!} \left[1 - \frac{1}{n} \right] \left[1 - \frac{2}{n} \right] \dots \left[1 - \frac{(x-1)}{n} \right] \left[1 - \frac{\lambda}{n} \right]^n \left[1 - \frac{\lambda}{n} \right]^{-x}.$$

O termo $\left[1 - \frac{\lambda}{n} \right]^n$, pela própria definição do número neperiano, tende a

$e^{-\lambda}$ quando $n \rightarrow \infty$. Os demais termos, do tipo $\left[1 - \frac{a}{n} \right]$ tendem a 1 quando

$n \rightarrow \infty$, inclusive $\left[1 - \frac{\lambda}{n} \right]^{-x}$. Portanto,

$$f(x; \lambda) = P[X = x] = \frac{\lambda^x e^{-\lambda}}{x!} I_{\{0,1,2,\dots\}}(x)$$

ou seja, tem-se a distribuição de Poisson na qual a média é igual à variância.

No padrão agregado (Figura 3c), a presença de plantas doentes é afetada pela maior ou menor ocorrência de plantas doentes na vizinhança e apresenta tendência à formação de grupos. De acordo com Bergamin Filho et al. (2002), este padrão ocorre quando, em condições naturais, o patógeno se dispersa apenas a curtas distâncias. Assim, a probabilidade de que uma planta situada próxima à fonte de inóculo se torne infectada é maior do que a de plantas distantes dessa fonte de infecção.

Uma das distribuições que se adapta a este padrão é a binomial negativa, que pode ser usada para descrever dados de contagem, pois apresenta variância maior que a da distribuição de Poisson, com a mesma média, o que caracteriza um padrão de agrupamento ou de agregação nas unidades amostrais.

Para melhor compreensão da relação de algumas distribuições com o padrão agregado, pode-se recorrer à descrição apresentada por Upton & Fingleton (1985) na qual consideram-se generalizações da distribuição de Poisson.

Na Figura 4(a), pode-se observar a distribuição aleatória de pontos no espaço, chamado de processo de Poisson ou floresta Poisson. Cada ponto representa um único indivíduo (ou uma única folha doente).

Partindo-se da Figura 4(a), obtém-se uma nova situação, apresentada na Figura 4(b), a qual consiste na substituição de cada ponto por um agrupamento de pontos. Para estes agrupamentos, o número de pontos é definido considerando-se uma distribuição de Poisson. Este padrão resultante não parece ser aleatório e sugere algum tipo de agrupamento ou de um padrão espacial agregado.

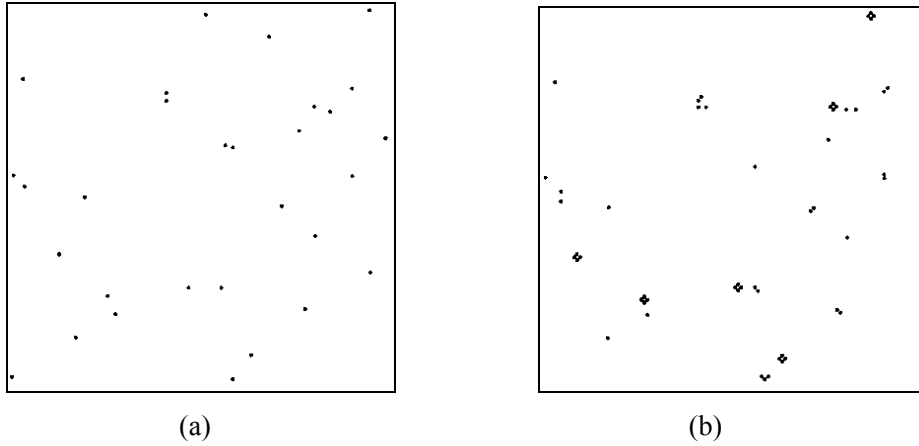


FIGURA 4. Representação de uma floresta Poisson (a) e de uma distribuição Neyman tipo A (b).

Segundo Upton & Fingleton (1985), esta combinação de agrupamentos foi sugerida por Neyman e é conhecida por distribuição Neyman tipo A. Nesta distribuição, considera-se que os grupos distribuem-se na área conforme a distribuição de Poisson de parâmetro λ e os elementos dentro de cada grupo distribuem-se como uma outra distribuição de Poisson de parâmetro ϕ . Embora a distribuição Neyman tipo A apresente uma motivação teórica bastante interessante, sua forma matemática é bastante complexa. Em função desta complexidade, uma generalização mais favorável da distribuição de Poisson pode ser obtida através da distribuição binomial negativa.

Admitindo-se a distribuição de Poisson, $f(x; \lambda) = \frac{e^{-\lambda} \lambda^x}{x!} I_{\{0,1,\dots\}}(x)$ e $g(\lambda) = \frac{\theta^r}{\Gamma(r)} \lambda^{r-1} e^{-\lambda\theta} I_{(0,\infty)}(\lambda)$, a função de densidade distribuição gama.

Combinando-se, estas duas distribuições e utilizando-se dos conceitos de distribuição conjunta, condicional e marginal, tem-se:

$$\int_0^{\infty} f(x; \lambda) \cdot g(\lambda) d\lambda = \int_0^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} \cdot \frac{\theta^r}{\Gamma(r)} \lambda^{r-1} e^{-\lambda\theta} d(\lambda).$$

Obtendo-se, assim, a distribuição binomial negativa com parâmetros r e $p = \frac{\theta}{\theta + 1}$, a qual é dada por:

$$P(X = x) = \binom{r + x - 1}{x} \left(\frac{\theta}{\theta + 1} \right)^r \left(\frac{1}{\theta + 1} \right)^x. \quad (5)$$

As distribuições de Poisson e binomial negativa podem ser usadas em situações de contagem quando não há um limite superior para a contagem nas unidades amostrais. Por exemplo, quando são contados os números de folhas infectadas em uma planta e o número total de folhas que a planta contém é desconhecido. Segundo Madden & Hughes (1995), estas distribuições são inapropriadas para análise de dados de proporção de plantas doentes, por causa da natureza binária da incidência de plantas doentes ou plantas sadias.

Ao contrário da situação de dados de contagem, quando se conhece o número total de indivíduos por unidade amostral, a incidência da doença é avaliada, dentro da unidade amostral, como se segue: coleta-se uma amostra composta por folhas das plantas, conta-se o número de folhas infectadas e não infectadas obtendo-se a proporção de folhas infectadas por unidade amostral. Nesta situação, quando se tem um padrão aleatório, a distribuição adequada é a binomial e, para um padrão agrupado, a distribuição adequada é a beta-binomial (Madden & Hughes, 1995).

Em muitos ensaios, quando se estuda a presença ou ausência de alguma característica numa unidade amostral, deve-se usar a distribuição de Bernoulli, na qual se atribui a uma variável aleatória Y apenas os valores de 0 ou 1, sendo

que 1 representa o sucesso (presença da doença) e 0 o fracasso (ausência). No contexto de incidência de doenças, sendo π a probabilidade que a folha esteja infectada, $1 - \pi$ a probabilidade de que a folha não esteja infectada, a distribuição de probabilidades da variável aleatória Y é dada por:

$$P(Y = y) = \begin{cases} \pi, & \text{se } y = 1 \\ 1 - \pi, & \text{se } y = 0 \\ 0 & \text{em outros casos.} \end{cases}$$

Estes dados são de natureza binária e a distribuição de Bernoulli é a única distribuição de probabilidade possível para esta situação. Pode-se demonstrar que a média e a variância de Y são dadas por: $E(Y) = \pi$ e $V(Y) = \pi(1 - \pi)$.

Considerando a distribuição de Bernoulli, pode-se definir uma variável aleatória binária Y_i para indicar a presença da doença em uma folha i , localizada em uma determinada unidade amostral, em que $i = 1, \dots, n$. Assim, o número total de folhas doentes na unidade amostral será $X = \sum_{i=1}^n Y_i$ e:

$$E(X) = E\left(\sum_{i=1}^n Y_i\right) = \sum_{i=1}^n E(Y_i) = n\pi$$

Se a localização de uma folha doente é independente da localização das outras folhas doentes e toda folha tem a probabilidade de sucesso constante π , então a variável aleatória X , que representa o número de folhas infectadas em uma unidade amostral de tamanho n , tem distribuição binomial, ou seja:

$$P(X = x) = \binom{n}{x} \pi^x (1 - \pi)^{n-x} \cdot I_{\{0,1,\dots,n\}}(x), \quad (6)$$

em que $Prob(.)$ representa a probabilidade de x folhas estarem infectadas. Nesta distribuição tem-se:

$$V(X) = V\left(\sum_{i=1}^n Y_i\right) = \sum_{i=1}^n V(Y_i) = n\pi(1-\pi).$$

Pode-se também considerar a proporção de folhas doentes $\frac{X}{n}$ no lugar do número de folhas doentes. Neste caso, a média esperada de $\frac{X}{n}$ será:

$$E\left(\frac{X}{n}\right) = \frac{1}{n}E(X) = \frac{1}{n}n\pi = \pi$$

e a variância:

$$V\left(\frac{X}{n}\right) = \frac{1}{n^2}V(X) = \frac{1}{n^2}n\pi(1-\pi) = \frac{\pi(1-\pi)}{n}.$$

No caso de uma lavoura, pode-se coletar um número (n) de folhas em cada planta e defini-lo como uma unidade amostral, tal que todas as n folhas sejam inspecionadas. Assim determina-se o número de folhas infectadas x_j , em cada unidade amostral, $j = 1, 2, \dots, N$ e determina-se a estimativa de π , por:

$$\hat{\pi} = \frac{\sum_{j=1}^N x_j}{nN} \quad (7)$$

que representa a incidência média da doença (Hughes & Madden, 1993;1995).

Um dos primeiros trabalhos na área de fitopatologia usando a distribuição binomial para caracterizar o padrão espacial de plantas doentes foi desenvolvido por Cochran (1936), que estudou a incidência de um vírus no tomate.

Uma outra técnica utilizada para descrever os padrões espaciais são os índices de dispersão. Os índices de dispersão são medidas do grau de agregação espacial de uma determinada população, e tem como princípio básico a característica da distribuição de Poisson, de que a média igual à variância, ou

seja, $E(X) = V(X) = \lambda$, em que X é uma variável aleatória referente a algum dado de contagem, como por exemplo, o número de plantas doentes em uma unidade amostral e λ é único termo dessa distribuição e representa a intensidade.

Esta relação variância/média (VM) pode ser dada por: $VM = \frac{s^2}{m}$ sendo

que a média pode ser entendida como a variância esperada (teórica) para a distribuição de Poisson.

Assim de acordo com, Van Maanen & Xu (2003) e Lima (2005), VM é essencialmente uma relação entre a variância observada e a variância esperada. Partindo-se dessa consideração, pode-se definir um índice de dispersão D como a relação entre a variância observada e a teórica.

Para o caso de avaliação da incidência de doenças em plantas e assumindo que a distribuição binomial seja adequada para descrever os dados na qual a variância esperada é dada por $n\pi(1-\pi)$, tem-se a equação de Fisher (1970):

$$D = \frac{s^2}{n\pi(1-\pi)},$$

em que s^2 representa a estimativa da variância do número observado de plantas doentes por unidade amostral. Se a incidência média da doença π é pequena, $1-\pi \approx 1$, a variância teórica será, aproximadamente, $n\pi$. Quando $D = 1$ os dados são aleatórios, se $D > 1$ os dados apresentam uma superdispersão e se $D < 1$ tem-se uma subdispersão.

Pode-se utilizar o teste χ^2 para decidir sobre a aceitação ou rejeição da hipótese de nulidade de que o padrão espacial seja aleatório ($D = 1$). Se a localização das plantas doentes é independente ou aleatória, o que implica em π

constante, segundo Madden e Hughes 1995, $(N-1)D$ tem distribuição χ^2_{N-1} e para valores altos de $(N-1)D$, a hipótese alternativa de superdispersão é aceita.

Cochran (1936), observando plantas de tomate, verificou a ocorrência de frequências elevadas para plantas sadias e para mais de três plantas doentes e baixas frequências para uma, duas ou três plantas doentes, concluindo que esta série observada não teria uma distribuição binomial. Para este caso, em que em uma unidade amostral há uma grande concentração de indivíduos apresentando sucesso enquanto que, em outra, a concentração é baixa ou nula, foi proposta a distribuição beta-binomial.

2.2 A distribuição beta-binomial

Em muitas situações, quando a probabilidade de sucesso π não é constante, pode-se utilizar a distribuição beta-binomial assumindo que esta variação de π corresponde às realizações de uma variável aleatória limitada por 0 e 1, com distribuição beta e função densidade probabilidade $f(\pi)$, dada por:

$$f(\pi) = \left[\frac{\pi^{\alpha-1} (1-\pi)^{\beta-1}}{B(\alpha, \beta)} \right] I_{(0,1)}(\pi) \quad (8)$$

em que $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}$ é a função beta, α e β são parâmetros positivos

constantes e $\Gamma(\cdot)$ é a função gama, $\Gamma(\alpha) = \int_0^{\infty} e^{-x} x^{\alpha-1} dx$.

A função densidade de probabilidade da distribuição beta pode assumir uma série de diferentes formas e, conseqüentemente, pode ser usada como modelo em experimentos para os quais uma das formas seja apropriada.

Analisando-se o gráfico da função de probabilidade beta (Figura 5) verifica-se que: se seus parâmetros α e β forem ambos maiores que um, o gráfico será unimodal; se $\alpha = \beta = 1$, uma distribuição uniforme é obtida. Para $\alpha = \beta$ porém diferentes de um, o gráfico é simétrico em $\pi = 0,5$.

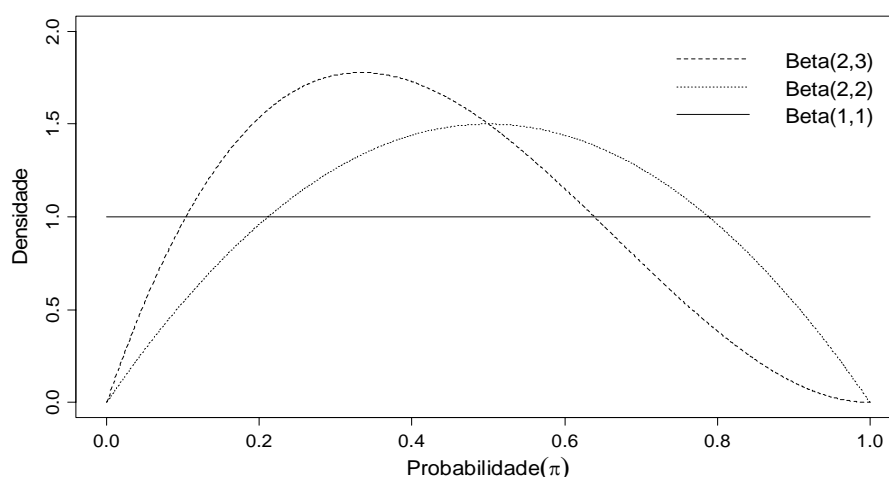


FIGURA 5. Gráfico da função de densidade beta para vários valores dos parâmetros α e β

A distribuição beta-binomial é obtida combinando-se a distribuição binomial com a distribuição beta, conforme desenvolvido por Skellam (1948) e adaptada por Madden & Hughes (1995). Para combinar estas duas distribuições utilizam-se os conceitos de distribuição conjunta, distribuição condicional e distribuição marginal.

Sempre que duas ou mais variáveis são estudadas, pode-se descrever o comportamento de todas elas simultaneamente ou isoladamente. Caso o interesse seja conhecer o comportamento de uma variável aleatória apenas, desconsideram-se as demais variáveis e tem-se a distribuição marginal. Já, se o interesse for conhecer o comportamento de uma variável aleatória isoladamente, dado que as outras variáveis assumem determinados valores, teremos a distribuição condicional.

Sabendo-se que, para um determinado valor fixo de π , X segue uma distribuição binomial então, condicionado ao parâmetro π , a distribuição de X é binomial, ou seja, $(X|\pi) \sim B(n, \pi)$. A função de probabilidade da condicional de X dado π , representada por $f(x|\pi)$, é definida por:

$$f(x|\pi) = \frac{f(x, \pi)}{f(\pi)}.$$

em que $f(\pi)$ é a função de densidade de probabilidade da distribuição marginal de π . Assim, tem-se que a função densidade de probabilidade conjunta das variáveis aleatórias X e π é dada por:

$$f(x, \pi) = f(x|\pi)f(\pi) .$$

A partir de $f(x, \pi)$, pode-se obter $f(x)$, dada por:

$$f(x) = \int_0^1 f(x|\pi)f(\pi)d\pi.$$

Então, utilizando-se destes conceitos, a função de probabilidade desta distribuição combinada é obtida através de:

$$\begin{aligned} P(X = x) &= \int_0^1 \binom{n}{x} \pi^x (1-\pi)^{n-x} \frac{\pi^{\alpha-1} (1-\pi)^{\beta-1}}{B(\alpha, \beta)} d\pi. \\ &= \binom{n}{x} \frac{1}{B(\alpha, \beta)} \int_0^1 \pi^{x+\alpha-1} (1-\pi)^{n-x+\beta-1} d\pi. \end{aligned}$$

Fazendo $x + \alpha = a$ $n - x + \beta = b$ tem-se

$$P(X = x) = \binom{n}{x} \frac{1}{B(\alpha, \beta)} \int_0^1 \pi^{\alpha-1} (1-\pi)^{\beta-1} d\pi.$$

Multiplicando-se por $\frac{B(a, b)}{B(\alpha, \beta)}$, tem-se:

$$P(X = x) = \binom{n}{x} \frac{1}{B(\alpha, \beta)} B(a, b) \int_0^1 \frac{\pi^{\alpha-1} (1-\pi)^{\beta-1}}{B(a, b)} d\pi.$$

Como $\frac{\pi^{\alpha-1} (1-\pi)^{\beta-1}}{B(a, b)}$ corresponde à função densidade de probabilidade

da distribuição beta, com parâmetros a e b tem-se que:

$$\int_0^1 \frac{\pi^{\alpha-1} (1-\pi)^{\beta-1}}{B(a, b)} d\pi = 1.$$

Assim,

$$P(X = x) = \binom{n}{x} \frac{B(a, b)}{B(\alpha, \beta)}.$$

Como $a = \alpha + x$ e $b = \beta + n - x$, tem-se :

$$P(X = x) = \binom{n}{x} \frac{B(\alpha + x, \beta + n - x)}{B(\alpha, \beta)} I_{\{0,1,2,\dots,n\}}(x), \quad (9)$$

que corresponde à função de densidade de probabilidade da distribuição beta-binomial.

A fórmula obtida para a distribuição beta-binomial é a de representação mais simples, mas Griffiths (1973), sugeriu uma reparametrização em que

$\pi = \frac{\alpha}{\alpha + \beta}$ e $\theta = \frac{1}{\alpha + \beta}$. De acordo com Madden & Hughes (1999), em

estudos de doenças de plantas π representa a incidência média da doença, ou seja, o valor esperado do parâmetro π ao invés do π constante da distribuição

binomial e θ representa uma medida da heterogeneidade espacial de X . O parâmetro θ é chamado de índice de agregação, pois indica o grau de agregação das folhas infectadas quando a distribuição beta-binomial é adequada.

Neste contexto, o coeficiente de correlação intra-classe ρ é um indicador da similaridade que existe entre indivíduos dentro de um mesmo grupo e pode ser expresso como (Ridout, et al., 1999): $\rho = \frac{\theta}{1+\theta}$, do qual tem-se:

$$\rho = \frac{\frac{1}{\alpha + \beta}}{1 + \frac{1}{\alpha + \beta}} = \frac{\frac{1}{\alpha + \beta}}{\frac{\alpha + \beta + 1}{\alpha + \beta}} = \frac{1}{\alpha + \beta} \cdot \frac{\alpha + \beta}{\alpha + \beta + 1} = \frac{1}{(\alpha + \beta + 1)}$$

Segund Collett (1991), os valores ρ sempre estão entre 0 e 1. Se $\rho = 0$, então o estado da doença de uma folha dentro de uma unidade amostral não influencia o estado da doença de outras folhas dentro da mesma unidade amostral. Como θ na distribuição beta-binomial é sempre maior que zero conseqüentemente ρ também será positivo .

De acordo com Mood et al (1974), dadas duas variáveis aleatórias quaisquer, X e Y , tem-se que $E(X) = E[E(X | Y)]$ e $V(X) = E[V(X | Y)] + V[E(X | Y)]$.

Assim, no caso da distribuição beta-binomial, tem-se:

$$\begin{aligned} E(X) &= E[E(X | \pi)] \\ &= E(n\pi) \\ E(X) &= nE(\pi) . \end{aligned}$$

Como $E(\pi) = \frac{\alpha}{\alpha + \beta}$, uma vez que $\pi \sim \text{Beta}(\alpha, \beta)$

$$E(X) = n \frac{\alpha}{\alpha + \beta}$$

Seguindo raciocínio semelhante tem-se

$$\begin{aligned} V(X) &= E[V(X | \pi)] + V[E(X | \pi)] \\ &= E[n\pi(1 - \pi)] + V[n\pi] \\ &= nE[\pi(1 - \pi)] + n^2 V[\pi] \end{aligned}$$

No primeiro termo da soma tem-se

$$\begin{aligned} E[\pi(1 - \pi)] &= \int_0^1 \pi(1 - \pi) f(\pi) d\pi \\ &= \frac{1}{B(\alpha, \beta)} \int_0^1 \pi(1 - \pi) \pi^{\alpha-1} (1 - \pi)^{\beta-1} d\pi \\ &= \frac{1}{B(\alpha, \beta)} \int_0^1 \pi^{\alpha+1-1} (1 - \pi)^{\beta+1-1} d\pi. \end{aligned}$$

Sabendo-se que

$$\int_0^1 \pi^{\alpha+1-1} (1 - \pi)^{\beta+1-1} d\pi = B(\alpha + 1, \beta + 1), \text{ e que } B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a + b)},$$

$$\begin{aligned} E[\pi(1 - \pi)] &= \frac{1}{B(\alpha, \beta)} B(\alpha + 1, \beta + 1) \\ &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha + 1)\Gamma(\beta + 1)}{\Gamma(\alpha + \beta + 2)} \\ &= \frac{\Gamma(\alpha + 1)\Gamma(\beta + 1)}{\Gamma(\alpha + \beta + 2)} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \\ &= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha + \beta + 2)} \frac{\Gamma(\alpha + 1)}{\Gamma(\alpha)} \frac{\Gamma(\beta + 1)}{\Gamma(\beta)}. \end{aligned}$$

Como $\Gamma(a) = (a-1)\Gamma(a-1)$,

$$\begin{aligned} E[\pi(1-\pi)] &= \frac{\Gamma(\alpha+\beta)}{(\alpha+\beta+1)\Gamma(\alpha+\beta+1)} \frac{\alpha\Gamma(\alpha)}{\Gamma(\alpha)} \frac{\beta\Gamma(\beta)}{\Gamma(\beta)} \\ &= \frac{\Gamma(\alpha+\beta)}{(\alpha+\beta+1)(\alpha+\beta)\Gamma(\alpha+\beta)} \alpha\beta \\ &= \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)}. \end{aligned}$$

Assim,

$$V(X) = n \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)} + n^2 V[\pi],$$

$$\text{em que } V(\pi) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

Portanto,

$$\begin{aligned} V(X) &= n \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)} + n^2 \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \\ &= \frac{n\alpha\beta}{(\alpha+\beta)^2} \left[\frac{(\alpha+\beta)}{(\alpha+\beta+1)} + \frac{n}{(\alpha+\beta+1)} \right] \\ V(X) &= n \frac{\alpha}{(\alpha+\beta)} \frac{\beta}{(\alpha+\beta)} \left[\frac{(\alpha+\beta+1)}{(\alpha+\beta+1)} + \frac{(n-1)}{(\alpha+\beta+1)} \right] \end{aligned}$$

Conforme definido por Griffiths (1973), $\pi = \frac{\alpha}{\alpha+\beta}$ e, conseqüentemente:

$$1-\pi = \frac{\alpha+\beta-\alpha}{\alpha+\beta} = \frac{\beta}{\alpha+\beta}. \text{ Ent\~{a}o,}$$

$$V(X) = n\pi(1-\pi) \left[1 + \frac{n-1}{(\alpha + \beta + 1)} \right].$$

$$V(X) = n\pi(1-\pi) \left[1 + \frac{1}{(\alpha + \beta + 1)}(n-1) \right].$$

De acordo com Ridout et al., 1999, $\rho = \frac{1}{(\alpha + \beta + 1)}$ e assim

$$V(X) = n\pi(1-\pi)[1 + \rho(n-1)]. \quad (10)$$

Para os dados de proporções, ou seja, $\left(\frac{X}{n}\right)$, a variância será:

$$V\left(\frac{X}{n}\right) = \frac{\pi(1-\pi)}{n} [1 + \rho(n-1)].$$

Reescrevendo a equação 10 tem-se $V(X) = n\pi(1-\pi) + n\pi(1-\pi)\rho(n-1)$.

Observa-se que a média da distribuição beta binomial é igual à média da distribuição binomial e as variâncias diferem entre si por um termo maior que 1, dado por $n\pi(1-\pi)\rho(n-1)$, em que $\rho = \frac{\theta}{1+\theta}$.

Pode-se dizer, então, que variância da distribuição beta binomial é a variância da distribuição binomial somada a um “fator de heterogeneidade”. Portanto, o parâmetro θ é um indicador do quanto os dados se afastam de uma distribuição binomial. O parâmetro θ varia de 0 a ∞ mas, geralmente, está entre 0 e 1. Como $\theta > 0$, a variância da distribuição beta-binomial é maior do que a binomial. Se o parâmetro $\theta \rightarrow 0$, a variância será a mesma para as duas distribuições. Assim, substituindo-se $\rho=1$ na equação da variância, a distribuição beta-binomial se reduz à uma distribuição binomial.

Segundo Madden & Hughes (1995), o estimador de π para a distribuição beta binomial, obtido pelo método dos momentos, é o mesmo da distribuição binomial, dado pela equação 7 e o estimador de θ , obtido pelo método dos momentos, é dado por:

$$\hat{\theta} = \frac{s^2 - n\pi(1-\pi)}{n^2\pi(1-\pi) - s^2} \quad (11)$$

em que s^2 é o estimador da variância do número de folhas infectadas por unidade amostral :

$$s^2 = \hat{V}(X) = \frac{\sum_{j=1}^N (x_j - m)^2}{N-1} \quad (12)$$

e $m = \sum_{j=1}^N \frac{x_j}{N}$, na qual x_j é o número de folhas infectadas na unidade amostral.

No caso de se considerar a proporção de folhas doentes na unidade amostral tem-se:

$$s^2 = V\left(\frac{X}{n}\right) = \frac{1}{n^2} V(X) = \frac{\sum_{j=1}^N (x_j - m)^2}{n^2(N-1)} . \quad (13)$$

Uma vez definidos $\hat{\pi}$ e $\hat{\theta}$, pode-se obter $\hat{\alpha}$ e $\hat{\beta}$, pois:

$$\hat{\pi} = \frac{\hat{\alpha}}{\hat{\alpha} + \hat{\beta}} \quad \text{e} \quad \hat{\theta} = \frac{1}{\hat{\alpha} + \hat{\beta}}$$

Madden & Hughes (1999) dizem que a melhor maneira de se estimar π e θ é pelo método da máxima verossimilhança, apesar de as estimativas serem mais facilmente obtidas pelo método dos momentos e serem, muitas vezes, próximas.

O método da máxima verossimilhança consiste na obtenção de valores para $\hat{\theta}$ e $\hat{\pi}$ que maximizem a probabilidade da amostra X_1, X_2, \dots, X_n ter ocorrido, ou seja, dentre os infinitos valores que $\hat{\theta}$ e $\hat{\pi}$ podem assumir, são tomados aqueles mais prováveis de serem os geradores dos dados.

Para a obtenção das estimativas dos parâmetros da distribuição beta-binomial pelo método da verossimilhança, exige-se que a distribuição beta-binomial seja reparametrizada para depois se obter a função de verossimilhança, adotando-se os procedimentos indicados em Griffiths (1973).

Para verificar se os dados observados estão de acordo com a distribuição binomial ou de acordo com a distribuição beta-binomial, indicando um padrão espacial aleatório ou agrupado, respectivamente, basta aplicar um teste de aderência sobre as frequências observadas e esperadas.

Madden & Hughes (1994), apresentaram um programa computacional para verificar o ajuste da distribuição beta-binomial a dados binários denominado BBD. Neste programa primeiro são calculadas as estimativas dos parâmetros π e θ pelos métodos dos momentos, e executados dois testes de homogeneidade de variâncias: o primeiro teste de homogeneidade de variâncias onde se considera o índice de dispersão D , sendo as hipóteses testadas: H_0 : os dados seguem uma distribuição binomial e H_1 : os dados não seguem uma distribuição binomial. A estatística χ^2 é calculada multiplicando-se o valor de D por $N-1$ (graus de liberdade associado ao teste), sendo que probabilidade menor que 0,05 indica que a hipótese de nulidade deve ser rejeitada.

O segundo teste de homogeneidade de variâncias é semelhante ao primeiro teste mas a hipótese alternativa é descrita pela distribuição beta-binomial . A estatística utilizada neste teste é dada por:

$$Z = \frac{[n(N-1)D - nN]}{[2N(n^2 - n)]^{\frac{1}{2}}}$$

Probabilidade menor que 0,05 indica que a hipótese de nulidade deve ser rejeitada, em favor da hipótese alternativa.

As estimativas de π e θ , obtidas pelo método dos momentos, são usadas como valores iniciais para estimar os parâmetros pelo método da máxima verossimilhança (Smith, 1983). As estimativas de máxima verossimilhança e seus erros padrões são calculados usando a técnica de Newton Raphson.

As frequências esperadas da beta-binomial são calculadas pelo método de Skellam (1948). Para o propósito de comparação o programa calcula também os valores esperados da distribuição binomial como o mesmo estimador π . As classes de frequências são agrupadas de modo que sejam maiores que cinco. É então calculado o teste χ^2 para ambas as distribuições binomial e beta-binomial.

Um bom ajuste da distribuição binomial sugere um padrão aleatório de incidência de doenças enquanto um bom ajuste da distribuição beta-binomial sugere um padrão de agregação (Madden & Hughes, 1995).

2.3 Aplicações da distribuição beta-binomial

Na literatura encontram-se várias aplicações da distribuição beta-binomial (DBB).

Uma aplicação da DBB foi apresentada por Griffiths (1973), que obteve um bom ajuste de dados referente à transmissão de resfriado comum em várias famílias. O exemplo pode ser ligeiramente descrito como segue: foram coletados dados da incidência de resfriado comum para remover possível fonte de heterogeneidade; as famílias foram classificadas de acordo com o número de

indivíduos optando-se por famílias com quatro indivíduos e classificando-as como i) somente adultos, ii) adultos e crianças em fase escolar, iii) adultos e crianças menores, iv) adultos, crianças em fase escolar e pré-escolar. Após observar os dados ele constatou que o número de indivíduos infectados variava bastante de família para família indicando que a probabilidade de infecção não era constante entre as famílias. Essa investigação serviu de base para consideráveis pesquisas relacionadas ao estudo dos efeitos de tratamentos, especialmente em toxicologia, onde é comum a obtenção de dados binários que apresentam superdispersão.

Tripathi et al. (1994) relatam que Williams (1975) pesquisou em laboratório o efeito de uma substância química em filhotes de roedores de diferentes fêmeas e registrou o número de filhotes que conseguiram sobreviver ao período de lactação sob efeito da substância. Neste estudo foi verificado que a sobrevivência dos filhotes variava de ninhada para ninhada e que a DBB se aplicava aos dados.

Estudos sobre o efeito teratogênico em animais foram desenvolvidos por Paul (1982) para verificar que tratamentos seriam responsáveis pelo aparecimento de anormalidades em fetos. Tais efeitos foram analisados como se seguisse uma DBB.

Crowder (1978) descreveu a germinação de sementes com a DBB onde considerou uma certa quantidade de sementes como unidade amostral na qual observou a probabilidade de germinação das sementes. Repetindo este processo ele verificou que a probabilidade de germinação entre as unidades amostrais eram diferentes e que a DBB se ajustou aos dados.

Existem muitos trabalhos onde é utilizado o parâmetro de agregação da DBB para descrever o padrão espacial de danos epidemiológicos e entomológicos, dentre os quais pode-se destacar: mofo da uva causado pelo fungo *Plasmopara viticola* (Madden & Hughes, 1995), fungo nas folhas do

morango (Turechek et al. 2001), plantas de milho infestadas pela broca do colmo do milho *Ostrinia nubilalis* (Hoffmann et al,1996), danos causados pelas larvas de *Ostrinia nubilalis* na batata (Nault & Kennedy, 1996), além de Hughes e Madden (1993) que aplicaram a DBB para avaliar a incidência de um vírus nas plantas de quatro patossistemas.

Através da observação dos dados da ferrugem no cafeeiro, verifica-se que o número de folhas infectadas entre as unidades amostrais varia, sendo que, em algumas unidades amostrais, ocorrem várias folhas infectadas e, em outras unidades, poucas folhas doentes ou nenhuma. Como estes dados são de natureza binária e a probabilidade de sucesso das unidades amostrais não é constante, isto sugere que a DBB deve ser adequada para descrevê-los.

2.4 Métodos de amostragem

Basicamente duas metodologias de amostragens, convencional e seqüencial, são utilizadas quando se tem interesse em avaliar o nível de infestação de pragas e a incidência de doenças em culturas, com objetivo de se tomar decisão sobre a necessidade do controle.

As decisões se baseiam em hipóteses sobre o nível de ocorrência da doença que justifique o seu controle e podem ser formuladas com base em experiências anteriores e de acordo com o interesse do pesquisador.

Na amostragem convencional o tamanho da amostra é fixo, procura-se estimar os parâmetros populacionais para depois, eventualmente, testar uma hipótese a respeito dos mesmos.

Muitas vezes a amostragem convencional exige muito tempo para ser realizada e a decisão de controlar ou não a praga ou a doença é tomada de forma empírica. A consequência disto é o uso indiscriminado de produtos químicos, causando desequilíbrio ecológico, dentre outros problemas.

2.4.1 Amostragem seqüencial

Na amostragem seqüencial, a amostra é de tamanho variável, sendo testada uma hipótese a respeito dos parâmetros de interesse, sem a necessidade de estimá-los (Barbosa, 1992). A amostragem seqüencial caracteriza-se em examinar em seqüência as unidades amostrais, tomando-se a primeira unidade amostral e estabelecendo-se linhas de decisão que definem três possibilidades de decisão: aceitar H_o , rejeitar H_o ou continuar amostrando. Se for tomada a decisão de continuar a amostragem, examina-se a segunda unidade amostral e o processo é repetido até que a decisão de aceitar ou rejeitar H_o seja tomada.

A razão principal do interesse pelo método da amostragem seqüencial é que, quando o nível de infestação de uma praga ou a incidência de uma doença for muito baixo ou muito alto, o número de amostras necessário será pequeno. Porém, de acordo com Estefanel (1977), o número de amostras necessário para a tomada de decisão aumenta se o índice de ataque estiver próximo do nível de controle da praga ou da doença. Segundo Pieters e Sterling (1975), Binns (1994), esta técnica de amostragem pode gerar uma economia de tempo com conseqüente diminuição de custo.

2.4.2 Amostragem para a ferrugem do cafeeiro

A ferrugem causada pelo fungo *Hemileia vastatrix* Berk.& Br. ocorre nas lavouras de café em todas as regiões onde ele é cultivado.

Os prejuízos no Brasil, nas regiões cafeeiras onde as condições climáticas são favoráveis à doença, atingem, em média, cerca de 35%. O cafeeiro é uma planta tipicamente bianual, isto é, de dois em dois anos as plantas apresentam alta produção. Nestes anos agrícolas de alta produção, a doença inicia-se em dezembro-janeiro e, de março a abril aumenta sua incidência até atingir o pico em torno de junho-julho. A partir deste período começa a

decrecer por causa das baixas temperaturas, da queda de folhas provocada pela colheita, da senescência natural e da própria severidade da doença que provoca intensa desfolha das plantas. Sob condições de estiagem prolongada nos períodos de maior severidade da doença, as perdas na produção podem chegar a mais de 50% . No ano agrícola de baixa produção a doença não ultrapassa 30% a 35% de incidência, mesmo sob condições favoráveis do clima. Portanto, é justamente nos anos de alta produção que a doença atinge alta severidade (Zambolim, 1999).

Os principais danos causados pela ferrugem ao cafeeiro são a queda precoce das folhas e a seca dos ramos e, em consequência, as plantas não produzem frutos no ano seguinte. A queda na produção, entretanto, varia de região para região e, até mesmo, de uma lavoura para outra (Zambolim, 1999).

Segundo Godoy et. al (1997), a disseminação ocorre pelo vento, pelas gotas de chuva, pelo homem durante a execução de tratamentos culturais, por insetos e por outros animais que entram em contato com plantas infectadas. A água é o mais eficiente agente de disseminação local, e se dá pelo respingo das gotas de chuva de uma folha para outra. Este fato pode explicar a formação de focos da doença nas lavouras para, depois, se disseminar de planta para planta (Zambolim, 1999).

2.4.2.1 Fatores que determinam a severidade da doença

Segundo Carvalho & Chalfoun (1998), um ponto importante para que ocorra uma epidemia da ferrugem nos cafezais é a alta densidade de hospedeiros susceptíveis. Nas grandes áreas, ou mesmo em regiões inteiras com plantios de café, em que usam-se variedades susceptíveis como o “ Catuí” e “ Mundo Novo”, o desenvolvimento da doença é favorecido, pois o potencial de inoculo será maior e fácil a disseminação dele entre as lavouras. Dentro deste aspecto, o

uso de variedades resistentes ou tolerantes é sumamente importante no conceito de manejo integrado de doenças.

Um outro aspecto que determina a severidade da doença, são as condições de alto enfolhamento, devido ao crescimento vegetativo ocorrido após o início das chuvas. Segundo Carvalho & Chalfoun (1998), tanto as folhas saudáveis como as infectadas, que são mantidas da estação seca até o período chuvoso, proporcionam maior quantidade de inoculo residual e maior área foliar para ser atingida pelas gotículas carregadas de uredosporos, promovendo a disseminação entre folhas e plantas. Com um bom pegamento da florada, ocorrerá uma carga pendente maior.

Várias pesquisas têm mostrado que o nível de produção das plantas é mais um fator que determina um maior ou menor grau de infecção do cafeeiro pela ferrugem. Assim sendo, os cuidados devem ser redobrados em relação às plantas enfolhadas e/ou com cargas altas pendentes .

Quanto aos fatores relacionados com o ambiente segundo Almeida, 1986, o fungo da ferrugem necessita de água livre, temperatura na faixa de 21 à 23 °C e ausência de luz direta para germinação e penetração dos uredosporos pelo estômatos da folha.

Os uredosporos só germinam na presença de água em estado líquido. Portanto a chuva, além de promover a distribuição deles entre folhas e plantas, sua quantidade assim como a hora de seu início tem grande influência na germinação dos esporos.

A temperatura pode atuar diretamente sobre a fisiologia dos esporos, ou indiretamente como fator de evaporação de água.. Temperaturas menores que 15 °C , assim como temperaturas diurnas maiores que 30 °C c podem limitar significativamente as epidemias.

Um outro fator climático que tem papel importante no desenvolvimento da doença é a luminosidade, que pode atuar diretamente sobre a germinação dos

uredosporos, ou atuar indiretamente em relação às condições microclimáticas de umidade e temperatura.

Nos últimos anos, com a necessidade de melhor aproveitar as áreas em busca de maiores rendimentos e retornos mais rápidos na cafeicultura, o sistema de plantio adensado vem sendo largamente praticado. Segundo Carvalho & Chalfoun (1998), o plantio denso compreende o uso de espaçamentos que resultam numa maior população cafeeira, que varia de 5.000 a 10.000 plantas por hectare. O aumento do número de plantas por unidade de área muda o ambiente ao redor da planta, formando um novo microclima, que interfere, entre outros fatores, na luminosidade, taxa de evaporação e velocidade dos ventos.

Uma adubação equilibrada poderá também diminuir os efeitos das desfolhas causadas pela ferrugem. Segundo Krüger (1978), em muitas relações hospedeiro-patógeno, a variação da resistência ou da suscetibilidade a doenças está ligada ao suprimento adequado ou equilíbrio de nutrientes das folhas.

Zambolim et al. 1985, sugere que a carga pendente alta promove um desequilíbrio de nutrientes das folhas por causa do dreno desses nutrientes para os frutos, tornando os cafeeiros mais susceptíveis à ferrugem. Cafeeiros com carga pendente alta podem ter até 60% mais folhas com ferrugem que cafeeiros sem produção.

Todos estes fatores ligados ao patógeno, à planta e ao meio ambiente podem influenciar em maior ou menor ocorrência da ferrugem. Se as lavouras já estão implantadas e com todos os fatores influenciando no sistema de ocorrência e evolução da doença, pode-se também empregar medidas culturais, tratamentos preventivos e controle químico procurando reduzir a doença.

2.4.2.2 Monitoramento

O controle da ferrugem do cafeeiro envolve, principalmente, o uso de fungicidas protetores e sistêmicos. Segundo Zambolim et al.(2005), pode-se

controlar a doença, preventivamente, com fungicidas de contato ou com o emprego de fungicidas sistêmicos, via foliar ou via solo.

Se a opção de controlar a doença for via foliar, esta poderá ser vinculada à incidência da ferrugem nas plantas e implicará em amostragens regulares.

A recomendação técnica é de coletar, ao acaso, folhas situadas no meio dos ramos nas posições dos pontos cardeais num total de 10 folhas por planta, no terço inferior das plantas, o que normalmente corresponde a uma altura de 70 cm a 90 cm do solo. A amostragem deve ser bem representativa, devendo caminhar em zigue-zague no talhão para efetuar a coleta de folhas. Pelo menos 100 folhas devem ser coletadas por talhão. Proceder em seguida a contagem do número total de folhas coletadas e do número de folhas com pústula, calculando-se a porcentagem de folhas doentes. Se a porcentagem de folhas infectadas estiver em torno de 5%, recomenda-se iniciar o controle da doença com fungicidas de contato; se ultrapassar 5% chegando até 12%, recomenda-se a aplicação de fungicidas sistêmicos (Zambolim, 1999).

Segundo Zambolim (1999), as pulverizações indiscriminadas de fungicidas e inseticidas via solo, na mesma área e ano após ano podem, a partir do quinto ou sexto ano, modificar a fisiologia da planta a tal ponto que a lavoura manifeste reflexos negativos na produção. Portanto, torna-se importante planejar, racionalmente, o uso destes produtos, o que pode ser obtido através de planos de amostragem seqüencial.

2.5 Planos de amostragem seqüencial

Os planos de amostragem seqüencial comumente usados para dados entomológicos e de doenças de plantas, podem se basear em dois métodos: o Teste Seqüencial da Razão de Verossimilhança e o Intervalo de Confiança de (Iwao 1975).

O método do intervalo de confiança de Iwao (1975), independe de modelo probabilístico associado ao padrão de distribuição da praga ou da doença. Esta metodologia apoia-se no teorema do limite central e consiste em adaptar um intervalo de confiança em torno da média (que é o nível de controle), gerando 2 curvas divergentes chamadas de linhas de decisão. Já no teste seqüencial da razão de verossimilhança inicialmente deve-se conhecer o modelo probabilístico que descreve o padrão espacial da praga ou doença

Conforme citado por Marcelino (1996), em estudos realizados para se comparar a eficiência dos métodos de Iwao (1975) e o Teste Seqüencial da razão de verossimilhança (Wald, 1945, 1947), foi concluído que o método de Wald é mais preciso e eficiente, e pelo fato de apresentar teoria estatística mais conhecida deve ter utilização preferencial ao método de Iwao.

2.5.1 Plano de amostragem seqüencial baseado no Teste Seqüencial da Razão de Verossimilhança

A técnica de amostragem seqüencial baseada no teste seqüencial da razão de probabilidade foi desenvolvida por Abraham Wald em 1943, e divulgada posteriormente em 1945 e 1947 (Wald 1945,1947). Wald, trabalhando para a Marinha dos Estados Unidos, iniciou uma pesquisa visando o controle de qualidade de material bélico. A partir daí, muitas indústrias começaram a utilizar a técnica no controle de qualidade de seus produtos.

Posteriormente, o método foi aplicado por Oakland (1950) no estudo de populações de peixes, por Stark (1952), Morris (1954), Waters (1955), dentre outros para controle de insetos. Estes planos demonstraram ser seguros e rápidos, tendo aumentado a confiança de decisão reduzindo o tamanho das amostras (Sterling, 1983).

Vários autores construíram planos de amostragem seqüencial visando o manejo de pragas, destacando-se: plano seqüencial para o levantamento da

lagarta da soja -*Anticarsia gemmatalis*- nas quais as distribuições de probabilidade utilizadas foram a binomial, a normal, a Poisson e a binomial Negativa (Estefanel, 1977); plano para o controle dos danos causados pelo bicho-mineiro do café -*Perileuoptera coffeella*- utilizando a distribuição binomial negativa (Villacorta & Tornero, 1982); plano para o controle da lagarta do cartucho do milho -*Spodoptera frugiperda*- (Farias, 1996); plano para controle do tripses dos folíolos -*Enneothrips flavens*- que causa o prateamento nas culturas do amendoim (Marcelino, 1996); plano para controle da broca européia do colmo -*Ostrinia nubilalis*- nas plantações de batata (Naut & Kennedy, 1996) e plano para controle da broca européia do colmo -*Ostrinia nubilalis*- na cultura do milho (Hoffmann et al., 1996).

No caso de doenças de plantas, Turecheck et al (2001) elaboraram um plano de amostragem seqüencial para avaliar a incidência de manchas de dendrofoma, causada por *Phomopsis obscurans*, nas folhas de morangueiro, onde foram utilizadas fórmulas aproximadas para as linhas de decisão considerando a distribuição beta- binomial.

Para elaborar um plano de amostragem seqüencial pelo método do teste seqüencial da razão de verossimilhança, inicialmente deve-se conhecer o modelo probabilístico que descreve a distribuição da praga ou doença. Barbosa (1985), cita que, além do conhecimento de uma função de probabilidades que descreva a padrão espacial dos indivíduos da população, os requisitos básicos para o desenvolvimento de um plano de amostragem seqüencial, são:

a) o estabelecimento prévio do nível de dano econômico, na forma de duas densidades críticas, de modo que o dano ocorrerá se a população do agente causador ou o nível do sintoma avaliado ultrapassar o limite superior e não ocorrerá se permanecer abaixo do limite inferior;

b) a seleção de níveis máximos de probabilidade de cometer erros na decisão sobre densidades populacionais, ou seja, a probabilidade α de estimar

uma densidade populacional não prejudicial como sendo prejudicial (erro tipo I) e a probabilidade β de estimar uma densidade populacional prejudicial como sendo não prejudicial (erro tipo II).

Os níveis de α e β mais comumente usados em pesquisas agrônômicas são $\alpha = \beta = 0,05$ e $\alpha = \beta = 0,10$.

A teoria para o teste seqüencial da razão de verossimilhança apresentada é baseada em Wald (1947) e nas adaptações propostas por Binns (1994) e Madden & Hughes (1999).

Seja μ o nível da incidência média das folhas infectadas π no caso da distribuição beta-binomial. Considerando-se ainda que μ_t seja o nível de controle, de modo que seja economicamente vantajoso iniciar alguma ação de controle como aplicar fungicida quando $\mu > \mu_t$ e que não seja vantajoso aplicar o fungicida quando $\mu \leq \mu_t$.

Se $\mu = \mu_t$ o dano causado pela doença é igual ao custo de aplicação do fungicida e é indiferente qual decisão tomar.

Se μ for pouco superior a μ_t , o dano causado pela doença é levemente superior ao custo da aplicação do fungicida, neste caso a não indicação do controle não implicará em erro de conseqüências graves.

Da mesma forma, se μ for pouco inferior a μ_t , o dano causado pela doença será levemente inferior ao custo da aplicação de fungicida e se utilizarmos fungicida não estaremos constituindo erro grave.

Assim é possível escolher dois valores: μ_0 e μ_1 , sendo μ_0 inferior a μ_t e μ_1 superior a μ_t , tais que, preconizar a aplicação do fungicida seja considerado erro de conseqüências graves se (e somente se) $\mu \leq \mu_0$, e desaconselhar a aplicação seja considerado erro de conseqüências graves se (e

somente se) $\mu > \mu_1$. Se $\mu_o < \mu < \mu_1$ nenhuma das decisões possíveis de serem tomadas conduz a erro de conseqüências práticas.

Escolhidos os valores de μ_o e μ_1 , a probabilidade de aconselhar a aplicação de fungicidas quando não for necessária é pré-fixada pelo valor α e a probabilidade de desaconselhar a aplicação quando necessária é dada pelo valor β , também pré-fixado.

Assim os riscos tolerados de tomar decisões erradas são fixadas pelos quatro valores: μ_o, μ_1, α e β . A escolha dessas quatro constantes não é uma questão estatística. Elas são escolhidas com bases em considerações práticas para cada caso particular.

Desta forma, as hipóteses a serem testadas são:

$$H_0 : \mu = \mu_o$$

$$H_1 : \mu = \mu_1$$

sendo $\mu_t = \frac{(\mu_1 + \mu_o)}{2}$ em que μ_o e μ_1 são constantes tais que $\mu_1 > \mu_o$. Se a hipótese de nulidade H_0 for aceita, conclui-se que $\mu < \mu_t$, caso contrário $\mu > \mu_t$. Pode-se usar o teste da razão de verossimilhança, para a tomada de decisão.

Por exemplo, se uma amostra aleatória é coletada em uma população com distribuição com parâmetros μ e σ^2 , a verossimilhança, é:

$$L(x_1, x_2, \dots, x_n | \mu, \sigma^2) = \prod_{i=1}^n \left\{ \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{1}{2} \left(\frac{x_i - \mu}{\sigma} \right)^2} \right\}$$

e o seu logaritmo é:

$$\begin{aligned}
\ln(L) = l(x_1, \dots, x_n | \mu, \sigma^2) &= \sum_{i=1}^n \ln \left\{ \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2} \left(\frac{x_i - \mu}{\sigma} \right)^2} \right\} \\
&= \sum_{i=1}^n \left\{ \ln \frac{1}{\sqrt{2\pi\sigma}} \right\} + \sum_{i=1}^n \left\{ -\frac{1}{2} \left(\frac{x_i - \mu}{\sigma} \right)^2 \right\} \\
&= n \ln \left(\frac{1}{\sqrt{2\pi\sigma}} \right) - \frac{1}{2} \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma} \right)^2.
\end{aligned}$$

A razão da verossimilhança, comumente chamada de z_i , é definida como:

$$\begin{aligned}
z_i &= \ln \left(\frac{L(x_1, x_2, \dots, x_n | \mu_1, \sigma^2)}{L(x_1, x_2, \dots, x_n | \mu_o, \sigma^2)} \right) \\
&= l(x_1, x_2, \dots, x_n | \mu_1, \sigma^2) - l(x_1, x_2, \dots, x_n | \mu_o, \sigma^2) \\
&= n \ln \frac{1}{\sqrt{2\pi \cdot \sigma}} - \frac{1}{2} \sum_{i=1}^n \left(\frac{x_i - \mu_1}{\sigma} \right)^2 - n \ln \frac{1}{\sqrt{2\pi \cdot \sigma}} + \frac{1}{2} \sum_{i=1}^n \left(\frac{x_i - \mu_o}{\sigma} \right)^2 \\
&= -\frac{1}{2} \sum_{i=1}^n \left(\frac{x_i - \mu_1}{\sigma} \right)^2 + \frac{1}{2} \sum_{i=1}^n \left(\frac{x_i - \mu_o}{\sigma} \right)^2 \\
&= -\frac{1}{2} \left[\sum_{i=1}^n \left(\frac{x_i - \mu_1}{\sigma} \right)^2 - \sum_{i=1}^n \left(\frac{x_i - \mu_o}{\sigma} \right)^2 \right] \\
&= -\frac{1}{2\sigma^2} \left(\sum_{i=1}^n x_i^2 - 2 \sum_{i=1}^n x_i \mu_1 + \sum_{i=1}^n \mu_1^2 - \sum_{i=1}^n x_i^2 + 2 \sum_{i=1}^n x_i \mu_o - \sum_{i=1}^n \mu_o^2 \right)
\end{aligned}$$

$$\begin{aligned}
&= -\frac{1}{2\sigma^2} \left(-2\sum_{i=1}^n x_i \mu_1 + 2\sum_{i=1}^n x_i \mu_o + \sum \mu_1^2 - \sum \mu_o^2 \right) \\
&= -\frac{1}{2\sigma^2} \left[-2(\mu_1 - \mu_o) \sum_{i=1}^n x_i + n\mu_1^2 - n\mu_o^2 \right] \\
&= -\frac{1}{2\sigma^2} \left[-2(\mu_1 - \mu_o) \sum_{i=1}^n x_i + n(\mu_1^2 - \mu_o^2) \right].
\end{aligned}$$

Definido $T_N = \sum_{i=1}^n x_i$, a razão da verossimilhança para a distribuição normal se torna:

$$z_i = \frac{1}{\sigma^2} \left[(\mu_1 - \mu_o) T_n - \frac{n}{2} (\mu_1^2 - \mu_o^2) \right].$$

Madden & Hughes (1999), baseados em Wald (1947) e considerando a *i*-ésima unidade amostral, sugeriram comparar a razão de verossimilhança com dois limites que dependem da probabilidade dos erros α e β , que podem ser escritas como:

$$\ln\left(\frac{\beta}{1-\alpha}\right) < \sum_{i=1}^n z_i < \ln\left(\frac{1-\beta}{\alpha}\right). \quad (14)$$

Para este teste seqüencial da razão de verossimilhança, depois que cada unidade amostral é coletada, determinam-se os valores de z_i , baseados na função de verossimilhança. Se $\sum_{i=1}^n z_i$ for menor que $\ln\left(\frac{\beta}{1-\alpha}\right)$ então a hipótese de que $\mu = \mu_o$ será aceita e, se $\sum_{i=1}^n z_i$ for maior que $\ln\left(\frac{1-\beta}{\alpha}\right)$, a

hipótese $\mu = \mu_1$ será aceita. Se $\sum_{i=1}^n z_i$ estiver entre estes dois valores, continua-se a observação de mais unidades amostrais.

Quando se deseja usar a equação (14), em funções complexas, como a da distribuição beta-binomial, pode-se recorrer a algumas alternativas tais como aproximações obtidas através de outras distribuições, como a Poisson, a binomial e a normal.

Quando a razão de verossimilhança da distribuição normal é assumida, a equação (14) se torna:

$$\ln\left(\frac{\beta}{1-\alpha}\right) < \frac{1}{\sigma^2} \left[(\mu_1 - \mu_o) T_n - \frac{n}{2} (\mu_1^2 - \mu_o^2) \right] < \ln\left(\frac{1-\beta}{\alpha}\right)$$

$$\frac{\sigma^2 \ln\left(\frac{\beta}{1-\alpha}\right) + \frac{n}{2} (\mu_1^2 - \mu_o^2)}{\mu_1 - \mu_o} < T_n < \frac{\sigma^2 \ln\left(\frac{1-\beta}{\alpha}\right) + \frac{n}{2} (\mu_1^2 - \mu_o^2)}{\mu_1 - \mu_o}.$$

Sendo $\mu_1^2 - \mu_o^2 = (\mu_1 - \mu_o)(\mu_1 + \mu_o)$, então,

$$\frac{\sigma^2 \ln\left(\frac{\beta}{1-\alpha}\right)}{\mu_1 - \mu_o} + \frac{n}{2} (\mu_1 + \mu_o) < T_n < \frac{\sigma^2 \ln\left(\frac{1-\beta}{\alpha}\right)}{\mu_1 - \mu_o} + \frac{n}{2} (\mu_1 + \mu_o),$$

que geralmente é representada na forma:

$$i_o + sn < T_n < i_1 + sn. \quad (15)$$

Os dois lados da inequação dada em (15) correspondem às linhas de decisão, tendo s como coeficiente angular e interceptos i_o e i_1 . No caso da distribuição normal tem-se:

$$s = \frac{\mu_1 + \mu_o}{2}$$

$$i_o = \frac{\sigma^2 \ln\left(\frac{\beta}{1-\alpha}\right)}{\mu_1 - \mu_o} \qquad i_1 = \frac{\sigma^2 \ln\left(\frac{1-\beta}{\alpha}\right)}{\mu_1 - \mu_o} .$$

Segundo Madden & Hughes (1999), pode-se obter linhas de decisão aproximadas para a distribuição beta-binomial assumindo que $\hat{\pi}$ tenha uma distribuição normal, ou pelo menos, aproximadamente normal. Sugeriu-se então usar o teste seqüencial da razão de verossimilhança com a distribuição normal, substituindo o termo do parâmetro da variância da verossimilhança $\left(\frac{\sigma^2}{\mu_1 - \mu_o}\right)$ da normal pela lei de potência binária e rearranjar para resolver T_n .

2.5.2 Lei de potência binária

A lei de potência binária é um outro método utilizado para descrever padrões espaciais de doenças em plantas, que será utilizado na elaboração das linhas de decisão com a distribuição beta-binomial.

Ao avaliar dados de contagem de animais de espécies diferentes, em vários locais diferentes, Taylor (1961), observou que as variâncias das populações geralmente eram proporcionais a uma potência da média:

$$v = a.m^b \tag{16}$$

em que v representa a variância, m representa a média, a e b representam parâmetros a serem estimados. Essa observação deu origem ao modelo da lei de potência de Taylor (1961):

$$\ln(v) = \ln(a) + b \ln(m) \tag{17}$$

sendo que b foi definido como índice de agregação ou um indicativo do comportamento espacial das espécies animais e a , um parâmetro de menor interesse. Se $a = b = 1$ na equação (17), tem-se que os dados de contagem

seguem uma distribuição Poisson, uma vez que a variância é igual a média. Pode-se escrever a equação (17) da seguinte maneira:

$$\ln(v) = \ln(a) + b \ln(v_r) \quad (18)$$

em que v é a variância observada, v_r é a variância teórica para uma distribuição aleatória, a e b são parâmetros.

Seguindo o mesmo raciocínio, Hughes & Madden (1992) adaptaram a lei da potência Taylor para dados de incidência de doenças. Para isto é necessário que a incidência da doença seja avaliada em diferentes áreas independentes como, por exemplo, talhões. Os talhões são divididos em N unidades amostrais, nas quais, ou conta-se o número de plantas doentes ou determina-se a proporção de plantas doentes. Isto implica em definições diferentes para cada situação.

Com N unidades amostrais e n plantas em cada unidade amostral e, se $\frac{X}{n}$ representa um dado de proporção, sendo $X \sim B(n, \pi)$, a variância esperada é definida por $v_r = V\left(\frac{X}{n}\right) = \frac{1}{n^2} V(X) = \frac{\pi(1-\pi)}{n}$ e a equação (18) é representada como:

$$\ln(v) = \ln(A_x) + b \ln\left[\frac{\pi(1-\pi)}{n}\right]. \quad (19)$$

Se x representa dados de contagem, tem-se que $v_r = V(X) = n\pi(1-\pi)$ e a equação (18) será:

$$\ln(v) = \ln(A_x) + b \ln[n\pi(1-\pi)]. \quad (20)$$

De acordo com Madden e Hughes (1995) e Lima (2005), quando $A_x = b = 1$, tem-se um padrão aleatório, pois, $\ln(A_x) = 0$. Então, $\ln(v) = \ln(v_r)$ ou seja $v = v_r$. Se $A_x > 1$ e $b = 1$, tem-se um padrão de agregação, o que implica

em superdispersão, ou seja, a variância observada é maior do que a variância teórica, porém, o grau de superdispersão em cada talhão não depende de π . Se A_x e b forem maiores do que 1, também existe superdispersão, mas esta varia sistematicamente com os valores de π .

Para a estimação dos parâmetros A_x e b através da análise de regressão, primeiro é necessário que se tenham as estimativas de π e v , as quais podem ser calculadas pelas expressões (7) e (12) ou (13). Segundo Madden & Hughes (1995), o parâmetro θ da distribuição beta-binomial pode ser representado em função dos parâmetros A_x e b da lei de potência binária, ou seja:

$$\theta = \frac{a - \frac{f(\pi)}{n}}{f(\pi) - a} \quad (21)$$

em que $f(\pi) = [\pi(1 - \pi)]^{1-b}$ e $a = A_x n^{-b}$ quando a análise for de proporção de plantas doentes. Se for utilizado a contagem de plantas doentes na unidade amostral então $a = A_x n^{b-2}$.

Segundo Bergamin Filho et al. (2002), a vantagem de se utilizar a lei de potência binária na caracterização de padrões espaciais é que a agregação de todo o conjunto de dados é descrita por um número menor de parâmetros, ou seja, A_x e b .

2.5.3 Aproximação para as linhas de decisão da distribuição beta-binomial.

Conforme Binns (1994), os interceptos das linhas de decisão podem ser considerados como um produto do “fator de intercepto” (f) multiplicado por

$\ln\left[\frac{\beta}{1-\alpha}\right]$ para i_o e por $\ln\left[\frac{(1-\beta)}{\alpha}\right]$ para i_i . Segundo Madden & Hughes

(1999), no caso da distribuição normal, o fator de intercepto é igual à $\frac{\sigma^2}{\mu_1 - \mu_0}$,

sendo σ^2 a variância de X e $\mu_0 = n\pi_0$, $\mu_1 = n\pi_1$ são dois parâmetros representando valores hipotéticos da média incidente. Então

$$\frac{\sigma^2}{\mu_1 - \mu_0} = \frac{\sigma^2}{n\pi_1 - n\pi_0} = \frac{\sigma^2}{n(\pi_1 - \pi_0)}.$$

Assumindo que a lei de potência binária seja apropriada, pode-se substituir a variância de X , σ^2 por $A_x[n\pi_t(1-\pi_t)]^b$, pois

$$\ln(v) = \ln(A_x) + b \ln[n\pi_t(1-\pi_t)]$$

$$\ln(v) = \ln\{A_x[n\pi_t(1-\pi_t)]^b\} \Rightarrow v = A_x[n\pi_t(1-\pi_t)]^b, \text{ obtendo-se:}$$

$$f = \frac{A_x[n\pi_t(1-\pi_t)]^b}{n(\pi_1 - \pi_0)} = \frac{A_x n^b [\pi_t(1-\pi_t)]^b}{n(\pi_1 - \pi_0)}. \quad (22)$$

Substituindo A_x por An^{2-2b} tem-se:

$$\begin{aligned} f &= \frac{An^{2-2b} n^b [\pi_t(1-\pi_t)]^b}{n(\pi_1 - \pi_0)} = \frac{An^{2-2b+b} [\pi_t(1-\pi_t)]^b}{n(\pi_1 - \pi_0)} \\ &= \frac{An^{2-b} [\pi_t(1-\pi_t)]^b}{n(\pi_1 - \pi_0)} \\ f &= \frac{An^{1-b} [\pi_t(1-\pi_t)]^b}{(\pi_1 - \pi_0)}. \end{aligned}$$

Sendo que π representa a incidência média da doença na reparametrização da distribuição beta-binomial, as hipóteses a serem testadas serão $H_0 : \pi = \pi_0$ e $H_1 : \pi = \pi_1$. Como o coeficiente angular s da linha de

decisão para distribuição normal é: $\mu_t = s = \frac{(\mu_1 + \mu_0)}{2}$, na distribuição beta-

binomial $\pi_t = \frac{(\pi_1 + \pi_0)}{2}$. Os interceptos e o coeficiente angular da reta podem ser dados por:

$$i_o = \frac{\ln\left[\frac{\beta}{1-\alpha}\right] An^{1-b} [\pi_t(1-\pi_t)]^b}{\pi_1 - \pi_0} . \quad (23)$$

$$i_i = \frac{\ln\left[\frac{(1-\beta)}{\alpha}\right] An^{1-b} [\pi_t(1-\pi_t)]^b}{\pi_1 - \pi_0} . \quad (24)$$

$$s = \pi_t = \frac{\pi_1 + \pi_0}{2} . \quad (25)$$

Para as condições de amostragem em que se tem N unidades amostrais com n elementos em cada, define-se a variável T_N como sendo o número acumulado de folhas infectadas que pode ser escrita como:

$$T_N = \sum_{j=1}^N x_j$$

em que x_j é o número de folhas infectadas na j -ésima unidade amostral ($j=1,2,\dots,N$). E assim a expressão (15) se torna:

$$i_o + snN < T_N < i_i + snN . \quad (26)$$

Os dois lados desta inequação correspondem as linhas de decisão chamadas por Madden & Hughes (1999), de linhas de decisão aproximadas para a beta-binomial, onde sn representa o coeficiente angular da reta, i_o e i_i são os interceptos.

Na utilização das linhas de decisão, calcula-se T_N depois que cada unidade amostral é observada plota-se o valor no gráfico. A amostragem é interrompida se T_N estiver acima da linha superior (decide-se a favor de π_1) e

se T_N estiver abaixo da linha inferior (decide-se a favor de π_o). Caso T_N esteja entre as linhas superior e inferior segue-se com a amostragem.

2.5.4 Curva característica de operação

Para que se possa visualizar, em um plano de amostragem seqüencial, a probabilidade de tomar uma decisão correta ou errada para qualquer grau de infecção, utiliza-se a curva característica de operação (CCO).

A CCO é um dos métodos primários utilizados para avaliar um teste de hipótese, que é definida como a probabilidade de aceitação da hipótese de nulidade dado um valor de π . Valores de π acima de π_t , a CCO nos fornece a probabilidade da correta aceitação da hipótese de nulidade ($1-\alpha$), e valores de π abaixo de π_t , a CCO nos fornece a probabilidade da incorreta aceitação da hipótese de nulidade (erro tipo II). A CCO ideal é 1 quando $\pi \leq \pi_t$ e 0 quando $\pi > \pi_t$, mas isso quase nunca acontece. Existem fórmulas conhecidas para algumas distribuições estatísticas que permitem calcular a CCO, mas segundo Turechek et al, (2001) utilizar simulação de Monte Carlo é uma boa solução.

Pode-se obter a CCO por simulações de conjunto de variáveis aleatórias com uma distribuição beta-binomial usando, em cada simulação, diferentes valores de π e θ , sendo que, este último, pode ser estimado pela fórmula (21), pelo método dos momentos ou o método da máxima verossimilhança.

Para cada conjunto simulado, estima-se $\hat{\pi}$, registrando-se o número de vezes em que a hipótese H_o é aceita. Este número é dividido pelo número de vezes em que o processo foi repetido e assim, pode-se dispor no gráfico a probabilidade de aceitação da hipótese de nulidade em função de um determinado valor de π .

A função da curva característica de operação é importante para se avaliar o plano elaborado indicando, sob determinadas circunstâncias, as probabilidades de aceitação da hipótese de nulidade e a probabilidade de cometer erros para a tomada de decisão.

3 MATERIAL E MÉTODOS

3.1 Dados reais

Os dados utilizados neste estudo foram de um experimento conduzido em 2005, na fazenda Cafua, localizada no município de Ijací, sul de Minas Gerais. A área corresponde a aproximadamente 6,5 ha de lavoura da cultivar Mundo Novo, com quinze anos de idade, plantada no espaçamento de 3,8 m entre linhas e 1 m entre plantas, totalizando 2632 plantas.ha⁻¹ conforme pode ser visto da Figura 6. As coordenadas geográficas são de 21° 10' 11" de latitude sul e 44° 58' 37" de longitude oeste de Greenwich, com altitude média de 934 metros e declividade de 0,84 % no sentido norte-sul e de 12 % no sentido leste-oeste.

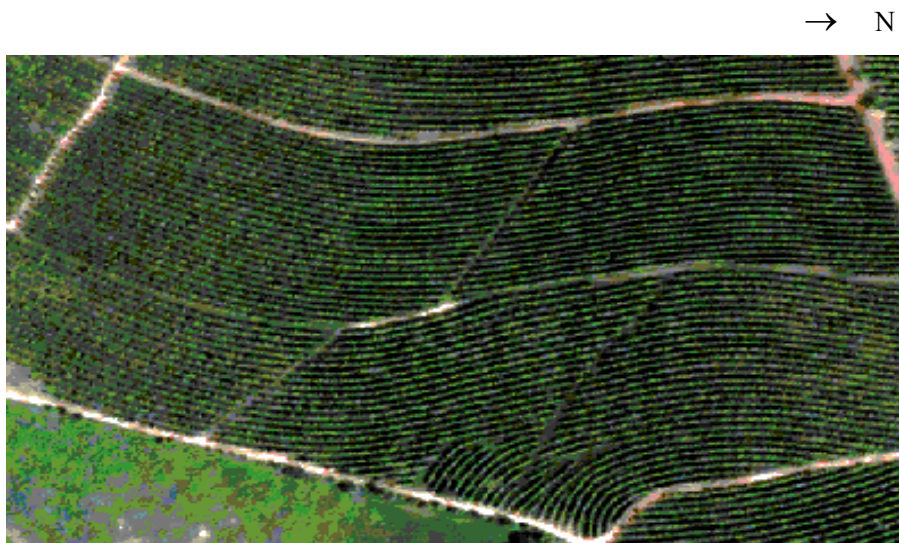


FIGURA 6. (Adaptado de Alves, 2006) Imagem de satélite QuickBird[®] 27/05/2003 referente à área estudada. UFLA, Lavras, MG, 2006. Composição colorida 321 e infravermelha falsa cor 432 .

Foram demarcados na área transectos com distâncias de 25x25m e 50x50m, num total de 67 pontos amostrais, distribuídos de forma irregular (Figura 7). As avaliações de incidência de ferrugem nas plantas foram realizadas em duas épocas: junho/ 2005 e julho/2006. Os pontos amostrados foram constituídos por cinco plantas. As amostras foram coletadas, aleatoriamente, de cada lado da linha de plantio, correspondentes aos pontos cardeais oeste e leste, no terço médio das plantas, a partir do terceiro e do quarto par de folhas, sendo dez folhas de cada lado da rua, por planta amostrada, totalizando 100 folhas por ponto amostral, num total de 6700 folhas avaliadas. As folhas foram destacadas e analisadas em laboratório.

Assim, tem-se um total de 67 unidades amostras ($N = 67$), sendo cada unidade amostral constituída por 100 folhas ($n = 100$), nas quais eram contadas as folhas com presença de ferrugem.

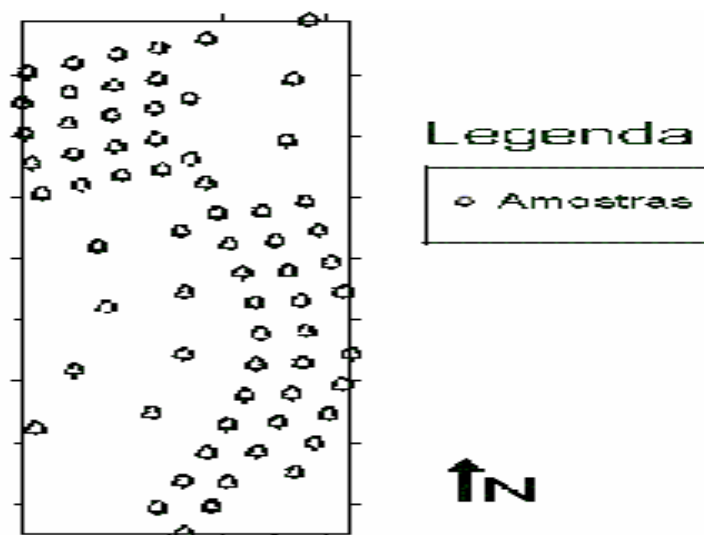


FIGURA 7. Representação bidimensional dos pontos amostrais.

A adubação foi realizada em novembro de 2004, janeiro, novembro, dezembro de 2005 e janeiro de 2006 aplicando-se 400g.cova⁻¹ N, P, K (20 05 15), por cova. Realizaram-se aplicações de Zn, B (2,5 – 1,5) em novembro de 2004, março, abril de 2005, novembro de 2005, janeiro e fevereiro de 2006, utilizando-se bomba de 400 L.ha. O controle químico de pragas, doenças e plantas daninhas foi realizado em novembro de 2004 (0,6 L.ha⁻¹ Opus[®]), abril de 2005 (2 L.ha⁻¹ de Nimbus[®], 2 L.ha⁻¹ de Endosulfan AG[®], 100 g.ha⁻¹ de Amistar[®], 0,4 L.ha⁻¹ Opus[®]), maio de 2005 (3 kg.ha⁻¹ de Roundup WG[®]), novembro de 2005 (1 kg.ha⁻¹ Verdadero 600 WG[®]), janeiro de 2006 (2 L.ha⁻¹ de Nimbus[®], 2 L.ha⁻¹ de Endosulfan AG[®], 100 g.ha⁻¹ de Amistar[®] e 2 kg.ha⁻¹ de Roundup WG[®]). Realizou-se capina manual, mecânica e retirada de cipó sempre que necessário.

Para determinar se o padrão espacial era agregado ou aleatório sugerindo uma distribuição de probabilidade foi utilizado o programa BBD de Madden e Hughes (1994).

Neste programa, primeiro foram calculadas as estimativas dos parâmetros π e θ pelos métodos dos momentos, e executados dois testes de homogeneidade de variâncias: no primeiro teste de homogeneidade de variâncias as hipóteses testadas foram: H_0 : os dados seguem uma distribuição binomial e H_1 : os dados não seguem uma distribuição binomial; para o segundo teste de homogeneidade de variâncias a hipótese de nulidade é a mesma do primeiro teste e a hipótese alternativa H_1 : os dados seguem uma distribuição beta-binomial.

As estimativas dos parâmetros π e θ foram obtidas pelo método da máxima verossimilhança. As frequências esperadas das distribuições binomial e beta-binomial foram calculadas, as classes foram agrupadas de modo que fossem

maiores que cinco conduzindo a um teste de aderência de χ^2 para ambas as distribuições.

Como só se pode observar um conjunto de dados da ferrugem do café para cada época amostrada, e para que se possa utilizar as linhas de decisão da beta-binomial são necessárias estimativas de $\ln(A_x)$ e b da lei de potência binária que requerem diversas variâncias, foram simulados dados com uma distribuição beta-binomial.

3.2 Dados Simulados

Foi usada uma simulação Monte-Carlo, utilizando-se o programa R De Development Core Team (2005) para gerar observações aleatórias e independentes X com distribuição beta-binomial, sendo que o número de indivíduos dentro de cada grupo era sempre igual a 100 ($n=100$) e total de unidades amostrais de 67 ($N=67$).

Com os valores estimados π e θ de cada data amostrada foram obtidos a variância observada e esperada em 200 conjuntos e obtida uma regressão linear estimado-se parâmetros $\ln(A_x)$ e b da lei de potência binária, os quais foram usados no cálculo das linhas de decisão, sabendo-se que $A_x = e^{\ln A_x}$ e

$$A = \frac{A_x}{n^{2-2b}}.$$

Na simulação 1 utilizou-se os valores de π e θ de 07/2005 e na simulação 2 os valores de π e θ de 07/2006.

3.3 Planos de amostragem seqüencial

Os planos de amostragem seqüencial foram construídos utilizando-se a metodologia baseada no teste seqüencial da razão de verossimilhança. Foram

comparados os valores acumulados das folhas infectadas em cada data amostrada com as linhas de decisão da beta-binomial.

Sendo o cafeeiro uma planta bianual e o nível de produção das plantas ser um fator que determina um maior ou menor grau de infecção da ferrugem, decidiu-se elaborar os planos considerando-se níveis de infecção para ano de baixa carga pendente e ano de alta carga.

As linhas de decisão foram obtidas utilizando-se sempre o nível de controle $\pi_i (=0,05)$ e com variações para π_o e π_1 sendo $\pi_o = 0,025$ e $\pi_1 = 0,075$; $\pi_o = 0,03$ e $\pi_1 = 0,07$; $\pi_o = 0,04$ e $\pi_1 = 0,06$, de modo a controlar o tamanho do intervalo das linhas de decisão. Foram utilizados $\alpha = \beta = 0,05$ para cada variação de π_o e π_1 . Para a variação $\pi_o = 0,025$ e $\pi_1 = 0,075$, foi utilizado $\alpha = \beta = 0,10$ apenas para verificar se havia algum efeito na mudança de $\alpha = \beta = 0,05$ para $\alpha = \beta = 0,10$. Foram construídos 4 planos de amostragem seqüencial considerando-se um ano de baixa carga e 6 planos considerando-se ano de alta carga. .

O número acumulado de folhas infectadas da ferrugem do cafeeiro de cada amostrada foi comparado com as linhas de decisão até que alguma decisão pudesse ser tomada.

As curvas característica de operação foram obtidas por simulação através do método de Monte Carlo, utilizando o programa R Development Core Team (2005), e considerando a distribuição beta-binomial. Os parâmetros utilizados na simulação foram $\pi = 0,1609$ e $\theta = 0,0622$; $\pi = 0,5523$ e $\theta = 0,0895$, que correspondem às estimativas obtidas nas duas épocas de avaliação da ferrugem do cafeeiro.

4 RESULTADOS E DISCUSSÃO

4.1 Padrão espacial da ferrugem do cafeeiro para dados reais

As estimativas dos parâmetros π e θ , obtidas pelo BBD utilizando os métodos dos momentos e da máxima verossimilhança estão apresentadas na Tabela 1.

Nota-se na Tabela 1 que as estimativas do nível de folhas infectadas aumentou no decorrer das épocas amostradas. Pode-se observar também que as estimativas obtidas pelo método dos momentos e pelo método da máxima verossimilhança foram semelhantes para as duas épocas amostradas, de acordo com os resultados obtidos por Madden & Hughes (1999).

Comparando as estimativas da época 07/2005 e 07/2006, houve um aumento significativo do nível de infecção das folhas com ferrugem pois 2005 foi um ano de baixa carga pendente e 2006 de alta carga. De acordo com o que as pesquisas têm mostrado, o nível de produção das plantas é um fator que determina um maior ou menor grau de infecção do cafeeiro pela ferrugem.

Segundo Zambolim et al.(1985), carga pendente alta promove um desequilíbrio de nutrientes das folhas por causa do dreno desses nutrientes para os frutos, tornando os cafeeiros mais susceptíveis a ferrugem. Cafeeiros com

TABELA 1. Estimativas dos parâmetros π e θ , em duas épocas amostradas.

Épocas	Métodos	Parâmetro	
		π	θ
07/2005	Momentos	0,1622	0,0538
	Verossimilhança	0,1609	0,0622
07/2006	Momentos	0,5510	0,0891
	Verossimilhança	0,5523	0,0895

carga pendente alta podem ter até 60% mais folhas com ferrugem que cafeeiros sem produção.

No primeiro teste de homogeneidade de variâncias, para as duas épocas amostradas foram encontrados altos valores da estatística χ^2 e baixa probabilidade de significância, indicando que a distribuição binomial não foi apropriada para descrever o padrão espacial conforme pode ser visto na Tabela 2. Neste teste a estatística χ^2 foi calculada multiplicando-se o valor de D por N-1 (=GL do teste), considerando que a hipótese de nulidade deve ser rejeitada para probabilidades menores que 0.01.

No segundo teste de homogeneidade de variâncias, foram obtidos altos valores para z_c e baixa probabilidade de significância, indicando que os dados seguem uma distribuição beta-binomial.

Na Tabela 3 observa-se que para o teste de aderência χ^2 aplicado, em que hipótese de nulidade representa que os dados seguem uma distribuição binomial, nota-se altos valores χ_c^2 para a distribuição binomial e baixa probabilidade de significância, indicando que a hipótese de nulidade deve ser rejeitada ou seja, os dados não estão de acordo com a distribuição binomial. No teste de aderência considerando que a hipótese alternativa seja que os dados

TABELA 2. Resultados do primeiro teste de homogeneidade de variâncias e índices de dispersão nas duas épocas amostradas.

Épocas	χ_c^2	Probabilidade	z_c	Probabilidade	D
07/2005	405,66	<0,01	29,40	<0,01	6,15
07/2006	609,40	<0,01	47,09	<0,01	9,23

TABELA 3. Testes de aderência χ^2 para as distribuições beta-binomial e binomial, nas duas épocas amostradas.

Épocas	Binomial			Beta-binomial		
	χ_c^2	GL	p-valor	χ_c^2	GL	p-valor
07/2005	51,29	7	<0,01	17,73	8	0,023
07/2006	65,42	8	<0,01	7,22	8	0,512

seguem uma distribuição beta-binomial, nota-se probabilidade de significância maiores que 0,01, indicando que a hipótese alternativa deve ser aceita ao nível de significância de 1%. Então nas duas épocas amostradas os dados de incidência de ferrugem nas folhas do café se ajustou melhor à distribuição beta-binomial indicando agregação.

As Figuras 8 e 9 mostram a distribuição de frequência do número de folhas com ferrugem em julho/2005. Na Figura 8 são mostradas as frequências observadas e as frequências esperadas segundo a distribuição beta-binomial, e na Figura 9 são mostradas as frequências observadas e as frequências esperadas segundo a distribuição binomial.

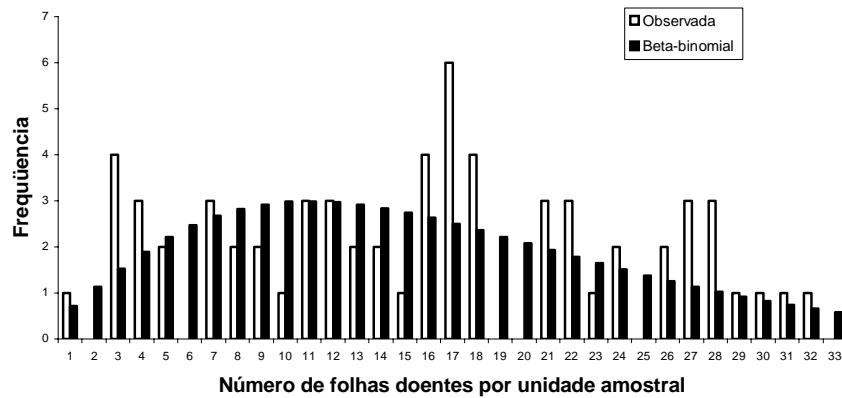


FIGURA 8. Distribuição de frequência observadas e esperadas das folhas com ferrugem segundo a distribuição beta-binomial. Amostra julho/2005.

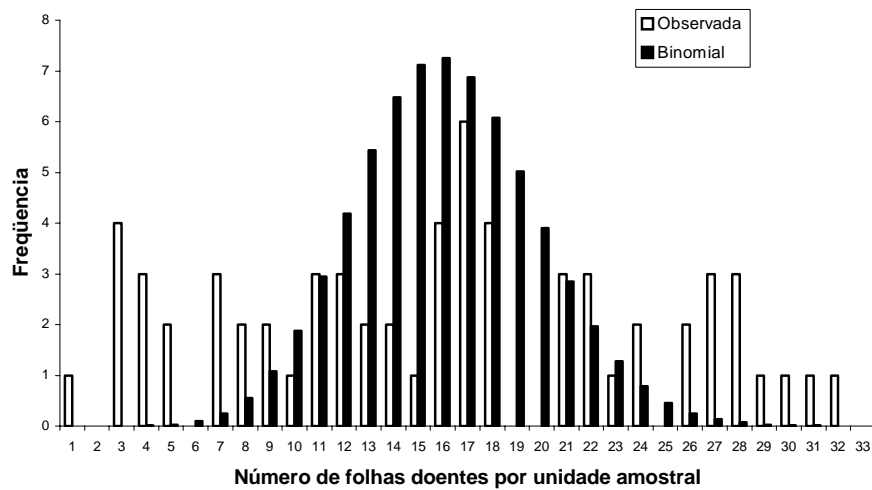


FIGURA 9. Distribuição de frequência observadas e esperadas das folhas com ferrugem segundo a distribuição binomial. Amostra julho/2005.

Nota-se no caso da distribuição binomial uma superestimação do número de folhas doentes no centro do gráfico e uma subestimação nas extremidades indicando que a mesma não apresentou um bom ajuste. A distribuição beta-binomial se ajustou melhor aos dados, ou seja, esta superestimação no centro e subestimação nas extremidades não é tão evidente quanto no caso anterior.

As Figuras 10 e 11 mostram a distribuição de frequência do número de folhas com ferrugem em julho/2006. Na Figura 10 são mostradas as frequências observadas e as frequências esperadas segundo a distribuição beta-binomial, e na Figura 11 são mostradas as frequências observadas e as frequências esperadas segundo a distribuição binomial.

Com o propósito de melhorar a apresentação das frequências na Figura 10 e 11, como o número observado de folhas infectadas por unidade amostral foram todos maiores que 23, os valores menores que este foram omitidos.

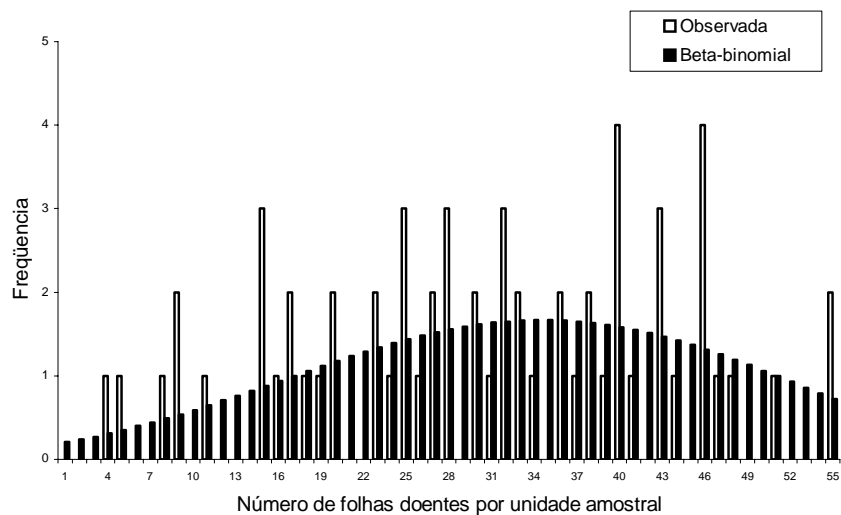


FIGURA 10. Distribuição de frequência observadas e esperadas das folhas com ferrugem segundo a distribuição beta-binomial. Amostra julho/2006.

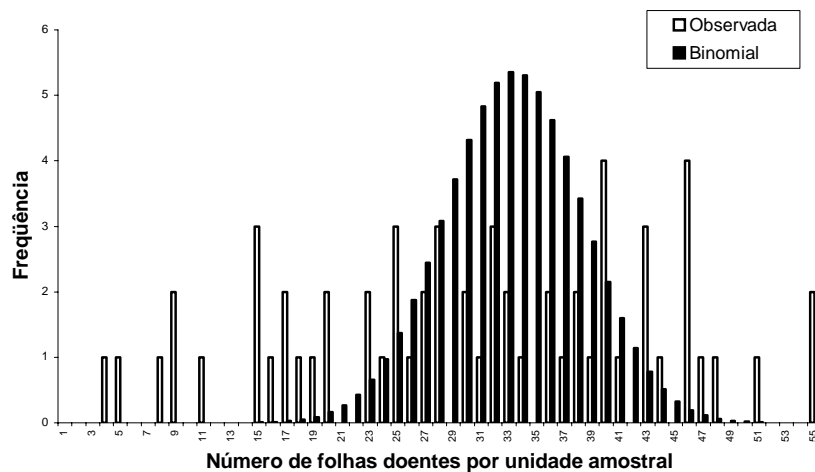


FIGURA 11. Distribuição de frequência observadas e esperadas das folhas com ferrugem segundo a distribuição binomial. Amostra julho/2006.

Nesta amostra de julho/2006 no caso da distribuição binomial também houve uma superestimação no centro e subestimação nas extremidades, indicando que a distribuição beta-binomial se ajustou melhor aos dados que a distribuição binomial.

4.2 Padrão espacial da ferrugem do cafeeiro para os dados simulados

As estimativas dos parâmetros $\ln(A_x)$, e b da lei de potência binária e os valores estimados de A usados para caracterização do padrão e nas linhas de decisão podem ser vistos na Tabela 4.

Na Tabela 5 observa-se que os testes aplicados a respeito dos parâmetros A_x e b da lei de potência binária resultou em padrão agregado para a simulação 1 (π e θ de 07/2005) e padrão aleatório para a simulação 2 (π e θ de 07/2006).

TABELA 4. Estimativas dos valores de $\ln(A_x)$, A_x , A e b da lei de potência binária.

Estimativas	1	2
$\ln(A_x)$	-0,0191	-1,2369
A_x	0,9810	0,2903
A	891,81	5692,15
b	1,7368	20,716

TABELA 5. Caracterização do padrão espacial dos conjuntos simulados utilizando π e θ nas duas épocas amostradas.

Simulação	1	2
	(π e θ 07/2005)	(π e θ 07/2006)
	$A_x = 1$ e $b > 1$	$A_x = 1$ e $b = 1$
Padrão	Agregado	Aleatório

Na simulação com mais conjuntos, comparando as simulação 1 com a simulação 2 e sabendo-se que o ano de 2005 era um ano de baixa carga, a doença manteve-se restrita a alguns pontos da lavoura, já no caso de carga elevada (2006), este padrão passou a ser aleatório devido a presença da produção constituir um fator de elevada suscetibilidade à doença.

Como a elevada incidência da doença sugere a presença de múltiplos focos de doença nas áreas de plantio, se estes focos estiverem disponíveis em grande quantidade dentro da área, podem apresentar um padrão aleatório, não permitindo a distinção entre o efeito de disseminação a curtas distâncias e longas distâncias.

Quando o nível de infestação aumentou, o padrão espacial se alterou passando de agregado para aleatório, ficando evidente não haver somente um padrão espacial da doença, o que está de acordo com Campbell & Madden (1990), mudanças no arranjo espacial de plantas doentes parecem ser comuns em vários patossistemas .

4.3 Planos de amostragem seqüencial

4.3.1 Plano de amostragem seqüencial ano de baixa carga

Utilizou-se $\alpha = \beta = 0,05$ que é o nível de erro mais usado em pesquisas agronômicas. As linhas de decisão foram comparadas com os dados da amostra 07/2005 que representa um ano de baixa carga.

Segundo Zambolim, (1999) se a percentagem de folhas infectadas ultrapassar 5% é recomendável a aplicação de fungicidas sistêmicos e decidiu-se então fixar o nível de controle π_t em 0,05. Os valores de π_o e π_1 foram escolhidos de acordo com a literatura citada no referencial teórico, de modo que

$$\pi_o \text{ seja inferior a } \pi_t \text{ e } \pi_1 \text{ superior, além de } \pi_t = \frac{\pi_1 - \pi_o}{2}.$$

A inequação (26) é plotada na Figura 12 versus N para os dados com $n=100$, $\pi_o = 0.025$, $\pi_1 = 0.075$ e $\pi_t = 0.05$, $A= 891,81$, $b=1,7368$ $\alpha = \beta = 0,05$. Baseado na equação (23) a (25) $s=0,05$, $i_0 = -8.82$ e $i_1 = +8.82$. A decisão pela hipótese alternativa foi tomada na 6ª unidade amostral. Apenas para o propósito de comparação a inequação (26) também foi calculada, para as mesmas condições, utilizando-se $\alpha = \beta = 0,10$, sendo $i_0 = -6,58$ e $i_1 = +6,58$.

A alteração de $\alpha = \beta = 0,05$ para $\alpha = \beta = 0,10$ não trouxe nenhum efeito pois a decisão também foi tomada na 6ª unidade amostral.

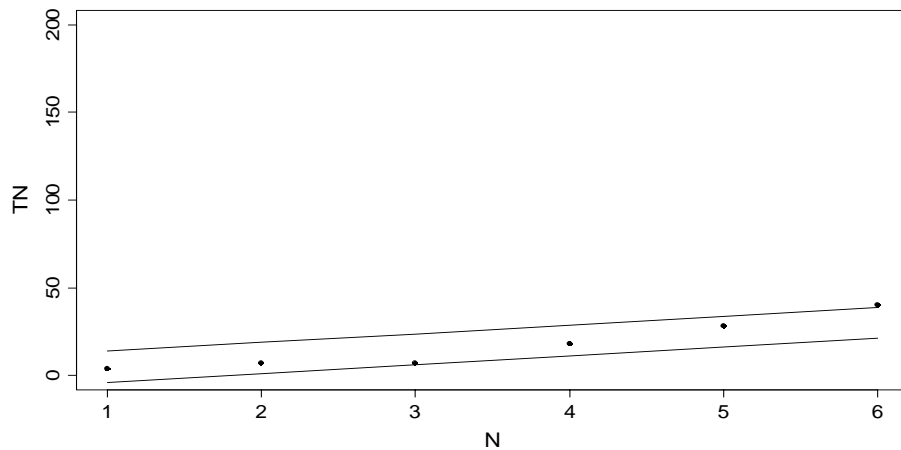


FIGURA 12 Linhas de decisão - número acumulado de folhas infectadas T_N , versus número de unidades amostrais N para a distribuição beta-binomial (equações 23 a 25).

No plano de amostral seqüencial para $\pi_o = 0,03$, $\pi_1 = 0,07$, $\alpha = \beta = 0,05$, baseado nas equações (23) a (25) $s=0,05$, $i_0 = -11,03$ e $i_1 = 11,03$ a decisão foi tomada na 11ª unidade amostral.

No plano de amostral seqüencial para $\pi_o = 0,04$, $\pi_1 = 0,06$, $\alpha = \beta = 0,05$, baseado na equação (23) a (25) $s=0,05$, $i_0 = -22,06$ e $i_1 = 22,06$ a decisão foi tomada na 13ª unidade amostral.

Na Figura 13 pode-se observar as linhas de decisão com as variações de π_o e π_1 , permanecendo $\pi_t = 0,05$.

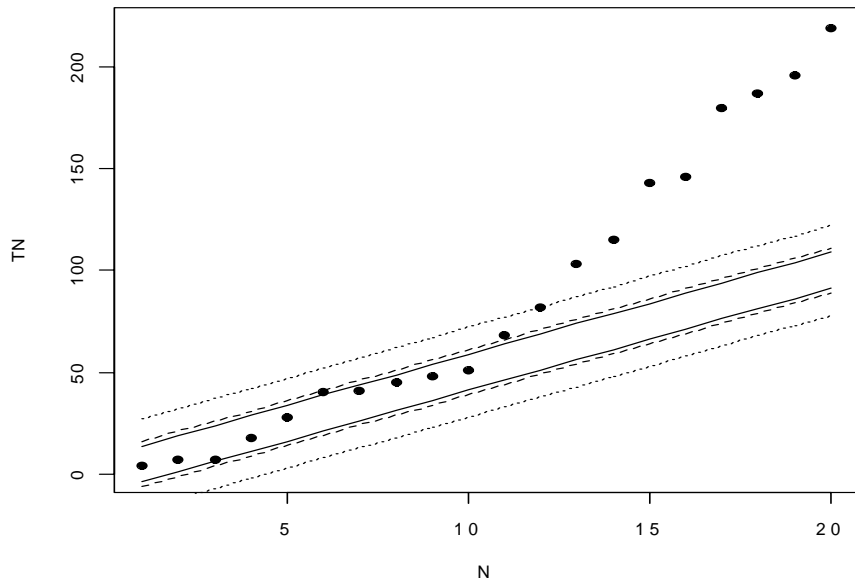


FIGURA 13 Linhas de decisão - número acumulado de folhas infectadas T_N , versus número de unidades amostrais N para a distribuição beta-binomial (equações 23 a 25). $\alpha = \beta = 0,05$.

Linhas contínuas $\pi_o = 0,025$, $\pi_1 = 0,075$, $i_0 = -8,82$ e $i_1 = +8,82$

Linhas tracejadas $\pi_o = 0,03$, $\pi_1 = 0,07$, $i_0 = -11,03$ e $i_1 = 11,03$

Linhas pontilhadas $\pi_o = 0,04$, $\pi_1 = 0,06$, $i_0 = -22,06$ e $i_1 = 22,06$.

Valores de π_o e π_1 mais distantes do nível de controle ($\pi_t=0,05$) produzem um menor intervalo, sendo necessário um menor número de unidades amostrais para a tomada de decisão. Quando os valores de π_o e π_1 se afastam do valor do nível de controle o intervalo se alarga e é necessário um maior número de unidades amostrais para a tomada de decisão.

A escolha de qual variação de π_0 e π_1 usar não é uma questão estatística e poderá ser escolhida com base em considerações práticas para cada caso em particular.

Uma ficha de campo foi confeccionada a partir das equações das retas $i_0 = -8.82$ e $i_1 = +8.82$, utilizando-se $\alpha = \beta = 0,05$ Figura 14. O procedimento a ser seguido é o seguinte: observa-se as 100 folhas da primeira unidade amostral, conta-se o número de folhas infeccionadas presentes, anotando-se na primeira coluna em branco: em seguida, vão se fazendo as observações e acumulando os resultados na segunda coluna em branco. A cada amostra compara-se o valor acumulado com os valores da tabela. Se o valor acumulado foi inferior ao número da coluna de decisão da esquerda (limite inferior) deve-se parar a amostragem e aceitar H_0 , optando-se por não controlar a doença, caso o valor acumulado for superior ao número da coluna de decisão da direita (limite superior), deve-se parar a amostragem e rejeitar H_0 (aceitar H_1), optando-se pela adoção do controle. Quando o valor acumulado estiver compreendido entre os dois valores, deve-se continuar amostrando.

Amostragem seqüencial da ferrugem do cafeeiro ano de baixa carga				
Unidades Amostrais	Limite inferior (não controlar)	Número de folhas com ferrugem	Número de folhas (acumulado)	Limite inferior (não controlar)
1				14
2	1			19
3	6			24
4	11			29
5	16			34
6	21			39
7	26			44
8	31			49
9	36			54
10	41			59
11	46			64
12	51			69
13	56			74
14	61			79

FIGURA 14 Ficha de campo para amostragem seqüencial da ferrugem do cafeeiro para ano de baixa carga pendente.

4.3.2 Plano de amostragem seqüencial ano de alta carga

As linhas de decisão foram comparadas com os dados da amostra 07/2006 que representa um ano de alta carga.

Para esta data amostrada foram construídos planos de amostragem seqüencial utilizando-se das linhas de decisão da beta-binomial e binomial, pois

na simulação dos parâmetros da lei de potência binária, os mesmos caracterizaram um padrão aleatório. Nestes planos foram utilizados $n=100$, $\pi_o = 0.03$, $\pi_1 = 0.07$; $\pi_o = 0.04$, $\pi_1 = 0.06$; $\pi_o = 0.025$, $\pi_1 = 0.075$, $\pi_t = 0.05$ com $\alpha = \beta = 0,05$.

Na Figura 15 pode-se observar as linhas de decisão utilizando-se a distribuição beta-binomial com as variações de π_o e π_1 .

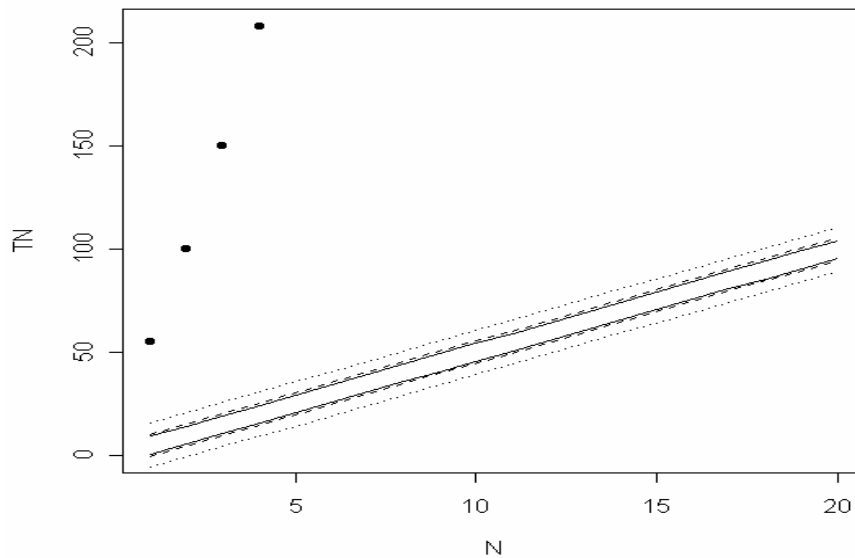


FIGURA 15 Linhas de decisão - número acumulado de folhas infectadas T_N , versus número de unidades amostrais N para a distribuição beta-binomial (equações 23 a 25). $\alpha = \beta = 0,05$.

Linhas contínuas $\pi_o = 0.025$, $\pi_1 = 0.075$, $i_0 = -4.28$ e $i_1 = +4,28$

Linhas tracejadas $\pi_o = 0,03$, $\pi_1 = 0,07$, $i_0 = -5,35$ e $i_1 = 5,35$

Linhas pontilhadas $\pi_o = 0.04$, $\pi_1 = 0.06$, $i_0 = -10,71$ e $i_1 = 10,71$.

Valores de π_o e π_1 mais distantes do nível de controle ($\pi_t=0,05$) produzem um menor intervalo, a medida que os valores de π_o e π_1 se afastam do valor do nível de controle o intervalo se alarga

Na Figura 16 pode-se observar as linhas de decisão utilizando-se a distribuição beta-binomial e distribuição binomial para $\pi_o = 0.025$, $\pi_1 = 0.075$.

Para ambas as distribuições e variações de π_o e π_1 a decisão foi tomada na primeira unidade amostral, pois a infestação estava muito acima do

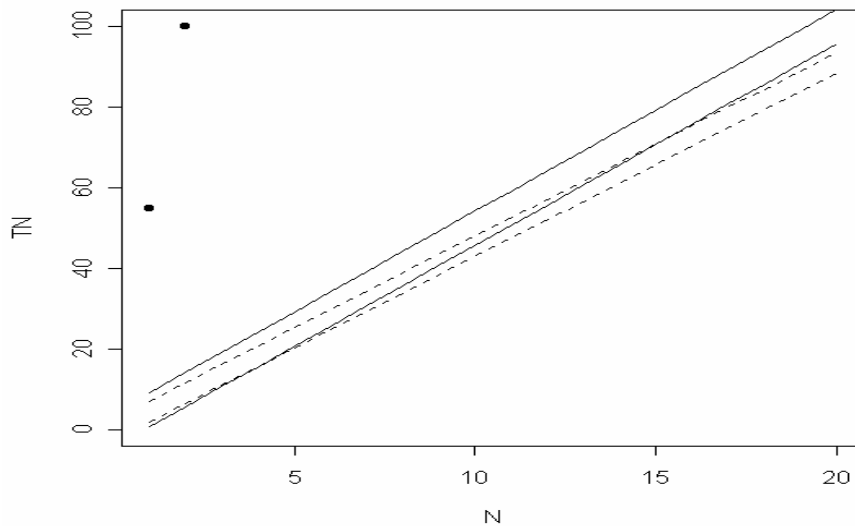


FIGURA 16 Linhas de decisão utilizando-se a distribuição beta-binomial e distribuição binomial para $\pi_o = 0.025$, $\pi_1 = 0.075$.

Reta contínuas: Linhas de decisão utilizando-se a distribuição beta-binomial e
 Reta tracejadas: Linhas de decisão utilizando-se a distribuição binomial.

nível de controle ($\hat{\pi} = 0,5523$).

Também foi confeccionada uma ficha de campo para ano de alta carga a partir das equações (23) a (25) $i_0 = -4,28$ e $i_1 = +4,28$ utilizando-se $\alpha = \beta = 0,05$, sendo que o procedimento a ser seguido é o mesmo apresentado para a ficha do ano de baixa carga.

Amostragem seqüencial da ferrugem do cafeeiro ano de alta carga				
Unidades Amostrais	Limite inferior (não controlar)	Número de folhas com ferrugem	Número de folhas (acumulado)	Limite inferior (não controlar)
1				10
2	5			15
3	10			20
4	15			25
5	20			30
6	25			35
7	30			40
8	35			45
9	40			50
10	45			55
11	50			60
12	55			65
13	60			70
14	65			75

FIGURA 17 Ficha de campo para amostragem seqüencial da ferrugem do cafeeiro para ano de alta carga pendente.

4.4 Curva característica de operação

Verifica-se na Tabela 6 e na Figura 18 a probabilidade de aceitação de H_o em função da proporção das folhas infectadas.

TABELA 6. Probabilidade de aceitação (Pac) de H_o em função da proporção das folhas infectadas.

julho/05		julho/06	
π	Pac	π	Pac
0,030	0,9988	0,030	0,9962
0,031	0,9986	0,031	0,996
0,032	0,9968	0,032	0,9948
0,033	0,9964	0,033	0,9882
0,034	0,9954	0,034	0,9838
0,035	0,9898	0,035	0,9814
0,036	0,9866	0,036	0,9732
0,037	0,9800	0,037	0,9616
0,038	0,9682	0,038	0,9474
0,039	0,9606	0,039	0,9312
0,040	0,9378	0,040	0,9136
0,041	0,9216	0,041	0,8950
0,042	0,8858	0,042	0,8620
0,043	0,8706	0,043	0,8378
0,044	0,8264	0,044	0,7962
0,045	0,7914	0,045	0,7552
0,046	0,7236	0,046	0,7074
0,047	0,6632	0,047	0,6688
0,048	0,6404	0,048	0,6118
0,049	0,5768	0,049	0,5632
0,050	0,5136	0,050	0,5156
0,051	0,4648	0,051	0,4608
0,052	0,4062	0,052	0,4236
0,053	0,3564	0,053	0,3806
0,054	0,2968	0,054	0,3338
0,055	0,2470	0,055	0,2836

TABELA 6, Continuação.

Julho/05		Julho/06	
π	Pac	π	Pac
0,056	0,2060	0,056	0,2456
0,057	0,1712	0,057	0,2126
0,058	0,1418	0,058	0,184
0,062	0,0548	0,062	0,0806
0,063	0,0398	0,063	0,0604
0,064	0,0294	0,064	0,0652
0,065	0,0212	0,065	0,0424
0,066	0,0138	0,066	0,0300
0,067	0,0116	0,067	0,0234
0,068	0,0090	0,068	0,0218
0,069	0,0078	0,069	0,0164
0,070	0,0042	0,070	0,0108
0,071	0,0020	0,071	0,0084
0,072	0,0014	0,072	0,0060
0,073	0,0016	0,073	0,0050
0,074	0,0014	0,074	0,0040
0,075	0,0004	0,075	0,0014
0,076	0	0,076	0,0018
0,077	0	0,077	0,0010
0,078	0	0,078	0,0010
0,079	0	0,079	0,0004
0,080	0	0,080	0,0004

Para valores de π mais distantes do nível de controle (0,05), como por exemplo $\pi = 0,04$ da amostra de julho/2005 nota-se que existe uma probabilidade de 0,9378 de aceitar a hipótese de nulidade, sendo ela verdadeira, e para $\pi = 0,07$ existe a probabilidade de 0,0042 de aceitar a hipótese de nulidade sendo ela falsa (erro tipo II).

Já para valores bem próximos do nível de controle, como por exemplo $\pi = 0,047$ da amostra de julho/2005 nota-se que a probabilidade aceitar a hipótese

de nulidade é 0,6632 e para $\pi=0,57$ a probabilidade de cometer erro tipo II é 0,1712.

Como era esperado, quando π se afasta do nível de controle as decisões são tomadas com menor risco, e, ao contrário, quando π se aproxima do nível de controle os riscos de tomar decisões aumentam.

Para índices de infecção muito acima do nível de controle as decisões são tomadas quase sem nenhum risco.

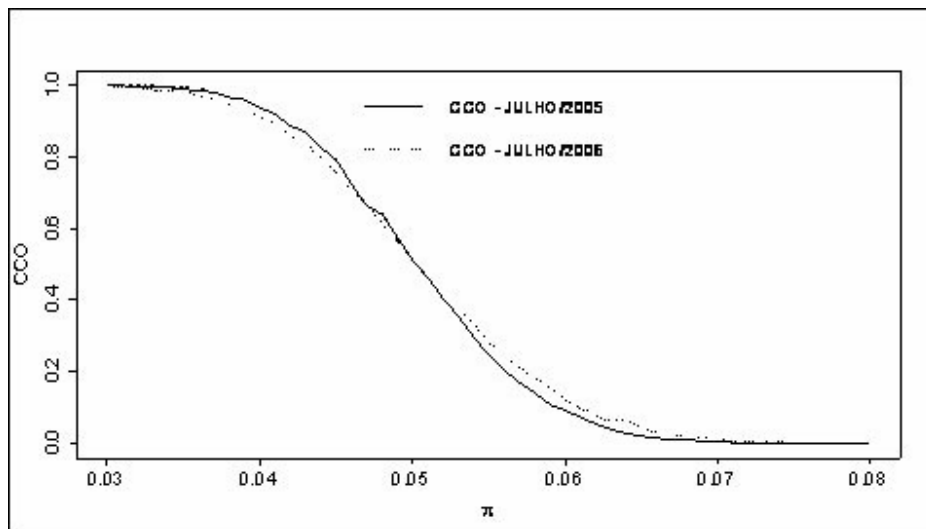


FIGURA 18 Curva característica da operação simulado pelo método de Monte Carlo utilizando as estimativas de π e θ época amostrada julho/2005 e julho/2006.

5 CONCLUSÕES

A utilização da distribuição beta-binomial na elaboração do plano de amostragem seqüencial pelo método do teste seqüencial da razão de probabilidade de Wald para a ferrugem do cafeeiro demonstrou ser uma eficiente ferramenta na tomada de decisão quanto controlar ou não a doença..

Foi necessário observar um pequeno número de unidades amostrais para a tomada de decisão, pois nos dados utilizados neste estudo, a infecção da ferrugem nas folhas de café estavam muito acima do nível de controle.

Uma indicação para reafirmar a caracterização do padrão espacial é que como a obtenção de dados consistentes sobre o padrão espacial de plantas doentes requer, via de regra, vários anos de coleta, estudos adicionais sobre a caracterização do padrão espacial da ferrugem do cafeeiro necessitam ser realizados.

A adequação do plano de amostragem seqüencial deve levar em consideração diferenças em áreas geográficas, práticas culturais, adubação, tratamentos preventivos, controle químico, ano de baixa produção e alta produção, os quais tornam necessários alterar as hipóteses referentes aos níveis de controle do plano de amostragem seqüencial.

6 REFERÊNCIAS BIBLIOGRÁFICAS

ALMEIDA, S. R. Doenças do cafeeiro. In: RENA, A.B. (Ed.) et al. **Cultura do cafeeiro**: fatores que afetam a produtividade. Piracicaba: POTAFOS, 1986. p.391-399.

ALVES, M. C. **Geoestatística e sistema 'fuzzi' na proteção de plantas**. 2006. 186p. Tese (Doutorado em Fitotecnia) - Universidade Federal de Lavras, Lavras.

BARBOSA, J. C. A Amostragem seqüencial. In: FERNANDES, O. A.; CORREIA, A. C. B.; BORTOLI, S. A. de. **Manejo integrado de pragas e nematóides**. Jabotical: FUNEP, 1992, p.205-211.

BARBOSA, J. C. **Distribuições de probabilidades como base para análises estatísticas, amostragem e estratégias de manejo de Diatraea saccharalis (Fabr. 1794) na cultura da cana-de-açúcar**. 1985. 131p. Tese (Doutorado em Estatística e Experimentação Agrônômica)- Escola Superior de Agricultura Luiz de Queiroz, Piracicaba.

BERGAMIN FILHO, A. et al. Análise espacial de epidemias. **Revisão anual de patologia de plantas**, Passo Fundo, v.10, p.155-218, 2002.

BINNS, M. R. Sequential sampling for classifying pest status. In: PEDIGO, L. P. ; BUTIN, G. D. (Ed.). **Handbook of sampling methods for arthropods in agriculture**. Boca Raton, Florida: CRC Press, 1994. p.137-174.

CAMPBELL, C. L.; MADDEN, L. V. **Introduction to plant disease epidemiology**. New York: John Wiley, 1990. 560p.

CARVALHO, V. L.; CHALFOUN, S. M. Manejo integrado das principais doenças do cafeeiro. **Informe agropecuário-EPAMIG**, Belo Horizonte, v.19, n.193, p.27-35, 1998.

COCHRAN, W. G. The statistical analysis of field counts of diseased plants. **Journal of the royal statistical**, New York, v.3, n.1, p.49-67, 1936.

COLLET, D. **Modeling binary data**. London: Chapman & Hall, 1991. 369p.

CROWDER, M. J. Beta binomial anova for proportions. **Applied Statistics**, Britain, v.27, n.1, p.34-37, May, 1978.

ESTEFANEL, V. **A amostragem seqüencial baseada no teste seqüencial da razão de probabilidades e seu uso no controle de lagartas da soja no estado do Rio grande do Sul**. 1977. 117p. Tese (Mestrado em Experimentação e Estatística). – Escola Superior de Agricultura Luiz de Queiroz, Piracicaba.

FARIAS, P. R. S. **Distribuição espacial e amostragem seqüencial de *Spodoptera frugiperda* (J. E. Smith, 1797) na cultura do milho**. 129p.1996. Dissertação (Mestrado em Entomologia Agrícola). – Universidade Estadual Paulista Julio de Mesquita Filho, Jaboticabal.

FISCHER, R. A. **Statistical methods for research workers**. 14. ed. New York: Hafner, 1970.

GODOY, C. V.; BERGAMIN FILHO, A.; SALGADO, C. L. Doenças do cafeeiro (*Coffea arábica* L.). In: KIMATI, H. **Manual de fitopatologia: doenças de plantas e seu controle**. 3. ed. São Paulo: Agronômica Ceres, 1997. v.2, p.184-189.

GRIFFITHS, D.A. Maximum likelihood estimation for the beta-binomial distribution and an application to household distribution of total number of a disease. **Biometrics**, Washington, v.29, n.4, p.637-648, Mar.1973.

HOFFMANN, J.P. et al. Sequential sampling plan for scheduling control of lepdopteran pests of fresh market sweet corn. **Journal of Economic Entomology**, Lanham, v.89, n.2, p.386-395, Apr.1996.

HUGHES, G.; MADDEN, L. V. Aggregation and incidence of disease. **Plant Pathology**, Londres, v.41, n.4, p.657-660, 1992.

HUGHES, G.; MADDEN, L. V. Using the beta-binomial distribution to describe aggregated patterns of disease incidence. **Phytopathology**, [s.l.], v.83, n.7, p. 759-763, Feb., 1993.

HUGHES, G.; MADDEN, L.V.; MUNKVOLD, G. P. Cluster Sampling for disease incidence data. **Phytopathology**, [s.l.], v.86, n.2, p. 132-137, Oct., 1996.

IWAO, S. A new method of sequential sampling to classify populations relative to a critical density. **Research on population ecology**, Kyoto, v.16, p.281-288, 1975.

KRÜGNER, T. L. Ação do ambiente sobre doenças de plantas. In: GALLI, F. (Coord.). **Manual de fitopatologia**. 2. ed. São Paulo: Agronômica Ceres, 1978. v.1, p.215-226.

LIMA, R. R. de. **Modelagem espaço-temporal para dados de incidência de doenças em plantas**. 2005. 149p. Tese (Doutorado em Estatística e Experimentação Agronômica) - Escola Superior de Agricultura Luiz de Queiroz, Piracicaba.

MAANEN, A van. ; XU, X. M. Modelling plant disease epidemics. **European Journal of plant pathology**, [s.l.], v.109, n.7 p.669-682, 2003.

MADDEN, L.V; HUGHES, G. BBD – Computer software for fitting the beta-binomial distribution to disease incidence data. **Plant disease**, St. Paul, v.78, n.5, p.536-540, Jan. 1994.

MADDEN, L.V; HUGHES, G. Plant disease incidence: distributions, heterogeneity, and temporal analysis. **Annual review of phytopathology**, Palo Alto, v.33, p.529-564, 1995.

MADDEN, L.V; HUGHES, G. Sampling for plant disease incidence. **Phytopathology**, [s.l.]v.89, n.11 p.1088-1103, June 1999.

MARCELINO, M. C. S. **Distribuição espacial e amostragem seqüencial de *Enneothrips flavens* (Moulton,1941) (Thysanoptera, Thripidae) na cultura do amendoim**. 1996. 169p. Tese (Mestrado em Produção Vegetal) – Universidade Estadual Paulista Julio de Mesquita Filho, Jabotical.

MOOD, A. M.; GRAYBILL, F. A.; BOES, D. C. **Introduction to the theory of statistics**. 3. ed. São Paulo: McGraw-Hill, 1974. 564p.

MORRIS, R. F. A. Sequential sampling technique for spruce budworm egg surveys. **The canadian journal zoology**, Ottawa, v.32 ,p.302-313, 1954.

NAULT, B. A.; KENNEDY, G. G. Sequential sampling plans for use in timing insecticide applications for control of European corn borer (Lepidoptera: pyralidae) in potato. **Journal of economic entomology**, Lanham, v.89, n.6, p.1468-1476, Dec. 1996.

OAKLAND, G.B. An application of sequential analysis to whitefish. **Biometrics**, Raleigh, v.6, p59-67, 1950.

PAUL, S. R. Analysis of proportions of affected fetuses in teratological experiments. **Biometrics**, Washington, v.38, n.2, p.361-370, June 1982.

PIETERS, E. P.; STERLING, W. L. Sequential sampling cotton squares damaged by boll weevils and *Heliothis spp.* in the coastal bend of Texas. **Journal of economic entomology**, Lanhan, v.68, p.543-545, Aug. 1975.

R DEVELOPMENT CORE TEAM (2005). **R: A language and environment for statistical computing**.. Disponível em: <<http://www.R-project.org>>. Acesso em: 20 out. 2006.

RIDOUT, M. S.; DEMÉTRIO, C. G. B; FIRTH, D. Estimating intraclass correlation for binary data. **Biometrics**, Washington, v.55, n.1, p.137-148, Mar.1999.

SKELLAM, J. G. A probability distribution derived from the binomial distribution by regarding the probability of success as variable between the sets of trials. **Journal royal statistical society: serie B**, Cambridge, v.10, p.257-261, 1948.

SMITH, D. M. Maximum likelihood estimation of the parameters of beta binomial distribution. **Applied statistics**, [s.l.],v.32, n.2, p.192-204, 1983.

STARK, R.W. Sequential sampling of the lodgepole needle miner. **Forestry Chronicle**, Toronto, v. 28, p.57-60.1952.

STERLING, W. L.; BLEICHER, E.; JESUS, F. M. M. Um programa de manejo integrado para insetos do algodoeiro no Nordeste do Brasil usando amostragem seqüencial. **Anais da sociedade entomológica do Brasil**, Jaboticabal, v.12, n.1, p.85-98, jun.1983.

TAYLOR, L. R. Aggregation, variance and the mean. **Nature**, [s.l.], v.189, n.476, p.732-735,1961.

TRIPATHI, R. C.;GUPTA, R. C.; GURLAND, J. Estimation of parameters in the beta binomial model. **Annals of the Institute of Statistical Mathematics**. v.46, n.2, p.317-331, July 1994.

TURECHEK, W. W.; ELLIS, M. A.; MADDEN L. V. Sequential Sampling for incidence of *Phomopsis* leaf blight of strawberry. **Phytopathology**, [s.l.],v.91,n.4, p.336-347, December, 2001 .

UPTON, G.J.G.; FINGLETON, B. **Spatial data analysis by example**. New York: John Wiley, 1985.

VILLACORTA, A; TORNERO, M.T.T. Plano de Amostragem seqüencial de dano causado por *Perileuoptera coffeella* no Paraná. **Pesquisa agropecuária brasileira**, Brasília, v.17, n.9. p.1249-1260, set.1982.

WALD, A. **Sequential analysis**. New York: J. Wiley, 1947. 212p.

WALD, A. Sequential tests of statistical hypothesis. **The annals of mathematical statistics**, Baltimore, v.16, p.117-186, 1945.

WATERS, W. E. Sequential sampling in forest insect surveys. **Forest science**, Washington, v.1, n.1, p.68-79, Mar.1955.

WILLIAMS, D. A . The analysis of binary responses frim toxicological experiments involving reproduction and teratogenicity. **Biometrics**, [s.l.], v.31, p.949-952,1975.

ZAMBOLIM, L. **I Encontro sobre produção de café com qualidade**. Viçosa: UFV/ DFP, 1999.p. 159-169.

ZAMBOLIM, L.; VALE, F. X. R.; ZAMBOLIM, E. M. Doenças do cafeeiro (*Coffea arábica* e *C. canephora*). In: KIMATI, H.(Ed.) et al. **Manual de Fitopatologia: doenças das plantas cultivadas**. 4. ed. São Paulo: Agronômica Ceres, 2005. v. 2, p.167-169.

ZAMBOLIM, L.et al. Influência da carga pendente sobre o desenvolvimento da ferrugem do cafeeiro. In: CONGRESSO BRASILEIRO DE PESQUISAS CAFEEIRAS, 12, 1985, Caxambu. **Trabalhos apresentados...** Rio de Janeiro: IBC, 1985a. p.123-125.