DETERMINAÇÃO DE UMA DISTRIBUIÇÃO ESTATÍSTICA
ADEQUADA A DADOS DE CONTAGEM DE LAGARTAS
DE *Scrobipalpula absoluta* (Meyrick), EM TOMATE
(*Lycopersicum esculentum*, Mill.), E UM PLANO
DE AMOSTRAGEM APROPRIADO

HUGO OSVALDO QUINTEROS

Engenheiro Agrônomo

Orientador: Prof. Dr. HUMBERTO DE CAMPOS

Dissertação apresentada à Escola
Superior de Agricultura "Luiz de
Queiroz", da Universidade de São
Paulo, para obtenção do título de
Mestre em Agronomia. Área de
Concentração em "Estatística e
Experimentação Agronômica".

PIRACICABA
Estado de São Paulo - Brasil
Novembro, 1984

*A minha esposa, colega e amiga
Sarita,*

OFEREÇO e DEDICO.

AGRADECIMENTOS

A Deus, por tudo.

Ao Dr. Humberto de Campos, Coordenador do Curso de Pós-Graduação em Estatística e Experimentação Agronômica, por sua generosidade na orientação deste trabalho e durante o Curso, e muito especialmente pela honra de me permitir ser seu amigo.

Ao Superior Governo da Provincia de Jujuy - República Argentina, que me favoreceu com bolsas nos anos de 1982 e 1983, fazendo possível, em parte, a obtenção do título de Mestre. Também, porque através do Ex-Secretário de Agricultura, M.S. Eng^o-Agr^o Eduardo Santos, obtiveram-se os dados deste trabalho, e ao Eng^o-Agr^o Luis Pablo Emilio Picchetti, que com dedicação e honestidade fez o trabalho de campo pertinente.

À Universidade Nacional de Jujuy, R.A., por ter-me favorecido, através da Faculdade de Ciências Agrárias, com bolsas de Pós-Graduação, possibilitando meus estudos e estadia neste País.

À ESALQ/USP, por ter colocado à minha disposição suas eficientes instalações.

- Ao "Consejo Profesional de Agrimensores, Arquitectos e Ingenieros de Jujuy" - R.A., por ser sensível aos reiterados pedidos de apoio econômico, solicitados oportunamente, contribuindo ao desenvolvimento e conclusão do Curso de Mestrado.
- À "Cámara del Tabaco de Jujuy", R.A., por ter me apoiado econômica e moralmente.
- À "Comisión de Apoyo a la U.N.Ju.", por sua contribuição econômica.
- Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela bolsa de estudos recebida.
- Ao Ex-Secretário Acadêmico da U.N.Ju., Eng^o-Agr^o Alberto Renê Vigiani, por seu constante e desinteressado apoio moral, profissional e amizade.
- Ao Ex-Decano da Faculdade de Ciências Agrárias da U.N.Ju., Eng^o-Agr^o Nestor José Alcoba, por constante preocupação para que fosse concretizado meu Curso de Pós-Graduação.
- À Decana da Faculdade de Ciências Agrárias da U.N.Ju., Prof^a Dora Vignale, pelo seu interesse no feliz acabamento do Curso.

- À Cátedra de "Biometria y Diseño Experimental" da F.C.A. da U.N.Ju., pelos conhecimentos transmitidos oportunamente.
- Ao colega, Prof. José Carlos Barbosa, da UNESP, "Campus" de Jaboticabal, por seu companheirismo, amizade e pelas valiosas sugestões feitas para este trabalho.
- Ao Corpo Docente do Departamento de Matemática e Estatística da ESALQ, pelos conhecimentos colocados à minha disposição.
- Aos colegas do Curso de Pós-Graduação, por terem me distinguido a honra de ser mais um da turma de 1982.
- A meus pais, pelo apoio permanente.
- Aos funcionários do Departamento de Matemática e Estatística da ESALQ, pela compreensão.
- À Secretária do Departamento de Matemática e Estatística da ESALQ, Maria Izalina Ferreira Alves, pelo excelente trabalho de datilografia efetuado.
- Ao Brasil e ao seu Povo, pois, no tempo de minha estadia neste país acolhedor, senti-me como um hóspede de honra.
- A todos que, de uma ou outra maneira, contribuíram para a concretização deste trabalho.

I N D I C E

	Pág.
RESUMO	x
SUMMARY	xiii
1. INTRODUÇÃO	1
2. REVISÃO DE LITERATURA	5
3. MATERIAL E MÉTODOS	12
3.1 - Material	12
3.2 - Métodos	13
3.2.1 - Distribuições Probabilísticas ...	13
3.2.1.1 - Distribuição de Poisson	14
3.2.1.2 - Distribuição Binomial Ge neralizada	14
3.2.1.3 - Distribuição Binomial Ne gativa	15
3.2.1.3.1 - Fôrmula re- corrente pa- ra a obten- ção de $P\{X=x\}$	17
3.2.1.4 - Distribuição de Neyman Tipo A, com dois parâme- tros	18
3.2.1.5 - Distribuição Poisson Du- pla de Thomas	19
3.2.2 - Escolha de uma Unidade Amostral ..	20
3.2.3 - Aplicação e Comparação de Três Mē- todos de Amostragem	26

	Pág.
4. RESULTADOS E DISCUSSÕES	32
4.1 - Ajuste às Distribuições Propostas	32
4.1.1 - Distribuição de Poisson	33
4.1.2 - Distribuição Binomial Generalizada	35
4.1.3 - Distribuição Binomial Negativa	37
4.1.3.1 - Obtenção do parâmetro $\bar{k}$	37
4.1.3.2 - Cálculo das probabilidades e teste de χ^2	40
4.1.4 - Distribuição de Neyman Tipo A, com Dois parâmetros	41
4.1.5 - Distribuição Poisson Dupla de Thomas	44
4.1.6 - Discussão	47
4.2 - Escolha de Uma Unidade Amostral	49
4.2.1 - Discussão	56
4.3 - Aplicação e Comparação de Três Métodos de Amostragem	59
4.3.1 - Amostragem Simples ao Acaso (A.S.A.)	60
4.3.1.1 - Prê-amostragem	60
4.3.1.1.1 - Aleatorização	60
4.3.1.1.2 - Dimensionamento.. ..	61
4.3.1.2 - Amostragem propriamente dita	63
4.3.2 - Amostragem Estratificada	64
4.3.2.1 - Prê-amostragem	64
4.3.2.1.1 - Aleatorização	65
4.3.2.1.2 - Dimensionamento.. ..	65

	Pág.
4.3.2.1 - Prê amostragem	64
4.3.2.1.1 - Aleatorização ...	65
4.3.2.1.2 - Dimensionamento..	65
4.3.2.2 - Amostragem propriamente dita	65
4.3.3 - Amostragem por Conglomerados em Dois Estágios	70
4.3.3.1 - Prê-amostragem	70
4.3.3.1.1 - Aleatorização ...	71
4.3.3.1.2 - Dimensionamento..	71
4.3.3.2 - Amostragem propriamente dita	75
4.3.4 - Discussão	78
4.4 - Considerações Finais	81
5. CONCLUSÕES	84
6. BIBLIOGRAFIA	86
7. APÊNDICE	91

DETERMINAÇÃO DE UMA DISTRIBUIÇÃO ESTATÍSTICA ADEQUADA A
DADOS DE CONTAGEM DE LAGARTAS DE *Scrobipalpula absoluta*
(Meyrick), EM TOMATE (*Lycopersicum esculentum*, Mill.),
E UM PLANO DE AMOSTRAGEM APROPRIADO

Candidato: Hugo Osvaldo Quinteros
Orientador: Prof. Dr. Humberto de Campos

R E S U M O

Numa cultura de tomate (*Lycopersicum esculentum*, Mill.) de aproximadamente 1/4 de hectare, na Estação Experimental "El Remate", sob administração da Direção Provincial de Estações Experimentais da Província de Jujuy, República Argentina, foi levantado um censo de lagartas de *Scrobipalpula absoluta* (Meyrick).

Foi extraída do referido censo a correspondente distribuição de frequências do citado microlepidoptero por planta viva, que foi submetida ao ajuste das seguintes distribuições probabilísticas discretas:

- Distribuição de Poisson;
- Distribuição Binomial Generalizada;
- Distribuição Binomial Negativa;

- Distribuição de Neyman Tipo A, com dois parâmetros;
- Distribuição Poisson Dupla de Thomas.

Com o objetivo de se aplicar métodos de amostragem, foram ensaiados 17 tipos de unidades amostrais de diferentes formas e tamanhos, visando à escolha do mais apropriado ao estudo em questão. Três métodos de amostragem foram confrontados:

- Amostragem Simples ao Acaso;
- Amostragem Estratificada;
- Amostragem por Conglomerados em Dois Estágios.

Concluiu-se que a distribuição probabilística discreta que melhor descreveu o comportamento dos dados foi a Binomial Generalizada, ou a Binomial Negativa, que para o caso foi coincidente.

Também determinou-se que a unidade amostral de formato quadrado, contendo 4 plantas (2 numa linha da cultura e 2 na linha ao lado) foi a mais eficiente e prática para a aplicação dos métodos de amostragem citados.

Dos métodos de amostragem, optou-se pelo da Amostragem Estratificada, pois as estimativas obtidas através dele foram as de maior precisão, e também por fornecer outras informações adicionais sob o ponto de vista agrônomo,

que são de fundamental importância para o produtor e para o pesquisador interessado em aprofundar os conhecimentos sobre a praga em questão.

DETERMINATION OF A SUITABLE DISTRIBUTION OF CATERPILLAR
COUNTING DATA OF *Scrobipalpula absoluta* (Meyrick),
IN TOMATO (*Lycopersicum esculentum*, Mill.) AND AN
APPROPRIATE SAMPLING

Candidate: Hugo Osvaldo Quinteros
Adviser: Prof. Dr. Humberto de Campos

S U M M A R Y

In a tomato (*Lycopersicum esculentum*, Mill.) crop, measuring approximately a quarter of hectare, a survey of caterpillar of *Scrobipalpula absoluta* (Meyrick) was made, in the "El Remate" Experimental Station, under the management of the Provincial Direction of Experimental Station of Province of Jujuy, Republic of Argentina.

The correspondent frequency distribution of that microlepidoptero for each living plant was extracted from the referred survey, and the following discrete probability distribution were fitted:

- Poisson Distribution;
- Generalized Bionomial Distribution;
- Negative Binomial Distribution;

- Neyman's Distribution type A, with two parameters;
- Double Poisson Distribution by Thomas.

Objecting the application of sampling methods, 17 types of experimental unit was essayed with different forms and sizes, in order to select the most appropriate for this study. Three sampling methods were compared:

- Simple Random Sampling;
- Stratified Sampling;
- Two-Stages Cluster Sampling.

We concluded that the probability distribution that best fitted the data was either the Generalized Binomial or the Negative Binomial, with in both cases were coincident.

It was also determined that the sample unit was square form, with four plants (2 in one crop line and 2 in the adjacent line) was the most efficient and practical for the sampling methods above referred.

Among the sampling methods, we choose the stratified one because the estimates obtained through it had the greatest precision, and also because it provided other additional information from the agronomic view point that have fundamental importance for the farmer and

for the investigator interested in to deepen his knowledges about the plaque studied.

1. INTRODUÇÃO

O homem necessita produzir alimentos e consequentemente controlar as pragas, que se enquadram entre os principais inimigos de uma produção agrícola. Para conseguir esse objetivo, é necessário estudar, dentre outros aspectos, os insetos que a provocam, seu tipo de disposição no campo, a distribuição de frequências decorrentes de sua contagem é uma distribuição estatística que descreva apropriadamente o comportamento dos dados (no caso presente, número de insetos por planta).

Existem muitos tipos de disposições dos insetos numa cultura, sendo das mais comuns as esquematizadas a seguir.

(a) É o tipo de disposição chamada "Regular" ou "Uniforme", pois os insetos distribuem-se ordenadamente. É raro encontrá-la no campo; um dos exemplos é o de alguns insetos sociais como os da Ordem Isoptera (família Termitidae). Esta distribuição espacial, cujos dados podem ser estudados através da Distribuição Probabilística Discreta Binomial, se caracteriza por:

$$\sigma^2 < m ,$$

onde σ^2 é a variância dos dados e m a sua média.

(b) É o tipo de distribuição espacial "Aleatória", que se encontra com bastante frequência na natureza. Neste caso, os dados (número de insetos por planta) apresentam

$$\sigma^2 \cong m .$$

Assim, pode-se estudar a distribuição de frequências resultante da contagem dos insetos, através da Distribuição Probabilística Discreta de Poisson quando

a ocorrência dos insetos é pouco frequente. Quando esta ocorrência é muito elevada, pode-se trabalhar com a Distribuição Normal, feitos os ajustes e transformações necessários.

(c) A disposição "Em Grupos" é a mais comum em contagem de indivíduos de populações em cultivo e se caracteriza por formar "Grupos" ou "Focos" no campo, onde acumulam-se. Um exemplo típico é o de ácaros em citros. Este tipo de disposição se caracteriza por:

$$\sigma^2 > m ,$$

que define uma família de distribuições probabilísticas discretas, às quais PÓLYA (1931) chamou de "Contágio", onde a presença de uma lagarta numa planta aumenta a probabilidade de existência de outra ou outras lagartas.

Estes estudos podem ser levados a efeito em todas as culturas atacadas por pragas, dentre elas, a de tomate (*Lycopersicum esculentum*, Mill.) que, em certas ocasiões, torna-se de fundamental importância para a economia de uma região agrícola, como é o caso das Províncias de Jujuy e Salta - República Argentina, onde a maior parte da produção tomateira é obtida numa época do ano (maio-junho-julho), na qual os preços do mercado nacional são os mais elevados.

Um dos insetos que mais atacam esta valiosa cultura é a "Polilla¹ del Tomate" (*Scrobipalpula absoluta*), que na fase de lagarta permite a contagem de indivíduos.

Assim, com o objetivo de achar-se um método apropriado de amostragem para levantamento de dados experimentais desta praga, este trabalho tem em meta as seguintes etapas:

- 1º) Determinação de uma distribuição de contágio que os descreva adequadamente;
- 2º) Seleção de um tipo de unidade amostral para a coleta dos dados;
- 3º) A escolha de um método de amostragem apropriado.

¹/ Polilla = Traça.

2. REVISÃO DE LITERATURA .

GREENWOOD e YULE (1920) estudaram "acidentes de trabalho", determinando uma forma de Distribuição Binomial Negativa, ou seja, uma distribuição composta por uma variável aleatória com Distribuição Gama, e esta, por sua vez, com uma distribuição de Poisson, representando o grau de propensão a acidentes de cada indivíduo, variável para cada trabalhador.

A Distribuição de Pólya-Aeppli (PÓLYA, 1931) surgiu da colonização aleatória e simultânea de um habitat por adultos individuais. O autor afirmou que os adultos produzem grupos de prole que se distribuem segundo uma distribuição de Poisson, e o número de indivíduos por grupo têm uma distribuição Geométrica. Este modelo

pode ter duas modas em condições particulares.

NEYMAN (1939) apresentou uma fórmula para uma família de distribuições de onde derivam as Distribuições Tipos A, com 2 e 3 parâmetros, e os Tipos B e C. No mesmo trabalho aplicou a Distribuição Tipo A com 2 parâmetros a 12 distribuições de frequências de lagartas e 3 de leveduras em quadrados dum hematocitômetro. O ajuste obtido em todos os casos por ele apresentados foi excelente.

Embora o ajuste à Distribuição Tipo A seja muito bom, deve-se notar que uma das pressuposições de Neyman foi a de que a fase do desenvolvimento das lagartas deve ser a mesma, o que é muito difícil de acontecer na prática. Também Neyman pressupôs que as lagartas devem ter pouca ou nenhuma mobilidade. Isto restringe as distribuições de Neyman a casos particulares. Estas distribuições são, no mínimo, bi-modais.

BEALL (1940) ajustou diversas distribuições de frequências de contagens de lagartas às distribuições de Neyman e de Pólya. Chamou a atenção para o fato de que as distribuições de Neyman são geralmente bi-modais, e mostrou um melhor ajuste em todos os casos em que as comparou com as de Pólya. Quando houve lagartas do mesmo instar, as de Neyman ajustaram-se muito bem, o que não ocorreu nos casos em que os insetos tiveram diferente fase de desenvolvimento. BEALL sugere, neste caso, fazer-se conta-

gens por tamanho ou instar.

COLE (1946) analisou populações que apresentavam "Distribuições de Contágio" ($\sigma^2 > m$), através da Distribuição Exponencial de Poisson. Neste caso, se N é o número total de amostras, e m é o número médio de lagartas por amostra, o número de amostras contendo exatamente x indivíduos é dado por:

$$N \cdot e^{-m} \frac{m^x}{x!}$$

Assim sendo, ele apresentou uma mistura de distribuições de Poisson, cada uma delas correspondendo a um tamanho de grupo diferente, pois pressupôs que cada tamanho de grupo distribui-se aleatoriamente no campo. Desta maneira, explicou alguns fenômenos biológicos.

Em suas próprias conclusões, o autor enfatizou que este tipo de análise só foi preciso no caso de distribuições de frequências com poucas classes, sendo muito difícil concluir acertadamente em outros casos. Provavelmente, a independência que ele pressupôs entre os diferentes tamanhos de grupos seja a causa desse problema.

THOMAS (1949) apresentou a Distribuição de Poisson Dupla gerada através de amostragem em quadrados, em plantas. A sua distribuição é muito parecida com a de Neyman Tipo A com 2 parâmetros, com a característica de

poder descrever distribuições de frequências uni e poli-modais.

ASCOMBE (1950) exibiu a Distribuição Binomial Negativa em forma teórica e mostrou métodos para a determinação dos parâmetros.

Em vários trabalhos, o mesmo autor aplicou a Distribuição Binomial Negativa em dados de contagem de insetos e, em 1949, sugeriu um tamanho de amostra de 10 unidades para se obter o ajuste a esta "Distribuição de Contágio". No mesmo trabalho o autor concluiu que esse tamanho ainda poderia ser insuficiente.

EVANS (1953) comparou as distribuições de Neyman Tipo A, de Pólya-Aeppli e a Binomial Negativa em populações de plantas e insetos. Nas suas conclusões afirmou ter encontrado um melhor ajuste da última distribuição comparativamente às outras estudadas.

HARTENSTEIN (1961), trabalhando com dados experimentais reais de micro-artrópodos de várias espécies, estudou em 31 amostras o ajuste às distribuições de Poisson, de Neyman Tipo A e Binomial Negativa. Obteve ajuste à Distribuição de Poisson só numa delas, 28 ajustaram-se à Binomial Negativa e 25 à de Neyman Tipo A, sendo que esta última superou, no ajuste, em apenas uma das amostras a Distribuição Binomial Negativa.

PIEDRABUENA (1968) estudou a Distribuição Binomial Negativa como uma distribuição "Interferida" com expressões estacásticas, à qual chamou Distribuição Binomial Generalizada, apresentando interessantes exemplos extraídos da prática.

ROJAS (1970) aplicou a Distribuição Binomial Negativa à contagem de insetos no solo, recomendando a amostragem deles através das características dessa distribuição.

CAMPOS GALVEZ (1977) informou sobre ensaios de controle efetuados com a "Polilla del Tomate" na Região Agrícola de Cañete - Peru, apresentando conclusões biológicas muito importantes para o conhecimento da praga.

ELLIOT (1977) mostrou e comparou diferentes métodos de amostragem aplicados a invertebrados "Benthic". Verificou, também, ser a Distribuição Binomial Negativa a mais apropriada aos dados de seu estudo.

ESTEFANEL (1977) aplicou amostragem sequencial, baseado no Teste Sequencial da Razão de Probabilidades, e mostrou a sua utilidade no controle de lagartas que atacam soja no Estado do Rio Grande do Sul, Brasil. Para isso, usou as distribuições: Normal, Binomial, Poisson e Binomial Negativa, concluindo que esta última foi a mais

apropriada no caso.

PERECIN (1978) estudou a aplicação de distribuições Binomiais Negativas com expoente k comum (k : fator de dispersão) em dados pedológicos e entomológicos, mostrando aplicações do Teste Sequencial da Razão de Verossimilhança. Concluiu que todas as distribuições de frequências, obtidas com dados de contagem na prática, podem se ajustar aproximadamente a distribuições Binomiais Negativas com expoente k comum.

BARBOSA (1978) estudou um modelo probabilístico para as distribuições de lagartas de *Spodoptera frugiperda* (J.E. Smith, 1797) em milho, concluindo que o "número de lagartas por planta" não foi uma boa variável para estudo, pois não se ajustou a uma distribuição definida. Por outro lado, achou, em alguns dos casos apresentados, um correto ajuste à Distribuição Binomial Negativa com expoente k comum. Deve-se salientar, no entanto, que trabalhou só com as distribuições Binomial, Poisson e Binomial Negativa.

CUCCHI (1980) deu sua contribuição ao estudo da "Polilla del Tomate", propondo um método de controle desta praga na Região Agrícola de Mendoza e San Juan - Argentina, onde se produz a maior porcentagem de "enlatados" do País.

BENAVENT (1981) desenvolveu estudos muito amplos sobre *Scrobipalpus absoluta* (Meyrick). Analisou danos, biologia, métodos de controle, parasitas, predadores, condições predisponentes, etc., mas não estudou a distribuição da praga no campo, e nem um método de amostragem adequado.

3. MATERIAL E MÉTODOS

3.1 - Material

Os dados utilizados neste estudo foram obtidos de uma cultura de tomate (*Lycopersicum esculentum*, Mill.), na Estação Experimental "El Remate", sob a administração da Direção Provincial de Estações Experimentais da Província de Jujuy - República Argentina.

Foi feito um censo de lagartas de "Polilla del Tomate" (*Scrobipalpula absoluta*), com o objetivo de se determinar a distribuição estatística do que considerar-se-á a população, obtendo-se, assim, a distribuição de frequências do número de lagartas do citado microlepidoptero por planta da cultura.

A área foi de aproximadamente 1/4 de hectare, com uma distância entre plantas, na linha, de 0,40 m, e entre linhas de 1,40 m, que são as medidas de prática na região.

Assim, foram possíveis 36 linhas com 132 plantas por linha, e que deu um total teórico de 4.752 plantas censadas na contagem de lagartas em toda a cultura, onde não se aplicou qualquer tipo de controle de insetos.

No que se refere às práticas de manejo da cultura, estas foram as tradicionais da região.

Os dados coletados são apresentados no Apêndice, ao final deste trabalho.

3.2 - Métodos

3.2.1 - Distribuições Probabilísticas

Para desenvolver um método de amostragem adequado ao estudo da praga, deve-se antes estudar uma distribuição estatística que descreva apropriadamente o comportamento dos dados, especialmente se o método é baseado

nela. Com este intuito, serão testadas algumas distribuições, apresentadas a seguir, e que, segundo a bibliografia, seriam as mais apropriadas para o presente caso (contagem de lagartas por planta viva).

A fim de se eleger, dentre as distribuições em estudo, aquela que seja mais adequada ao caso, será aplicado a todas elas o teste de aderência de qui-quadrado (χ^2).

3.2.1.1 - Distribuição de Poisson

Definida pela seguinte função de probabilidade:

$$f(x) = \frac{e^{-m} m^x}{x!} \quad (\text{para } x=0,1,2,\dots)$$

onde, $f(x)$ traduz a probabilidade de ocorrência de x lagartas por planta viva, ou seja, $P\{X=x\}$.

3.2.1.2 - Distribuição Binomial Generalizada

Definida pela função de probabilidade:

$$f(x) = \binom{v}{x} \chi^{v-x} \psi^x \quad (\text{para } x=0,1,2,\dots)$$

onde $f(x) = P\{X=x\}$, isto é, a probabilidade de ocorrência de x lagartas por planta viva.

Os parâmetros desta distribuição são:

ψ que representa a expressão estocástica de uma lagarta "contagiada" ou influenciada pela presença de outras lagartas em suas proximidades;

$$\chi = 1 - \psi, \text{ pois } \chi + \psi = 1 \text{ ou } \chi = \frac{\sigma^2}{m}; \text{ e}$$

v que substitui o expoente n da Distribuição Binomial e cuja fórmula de cálculo é:

$$v = \frac{m}{1 - \chi} \quad (\text{para } 1 - \chi \neq 0)$$

ou ainda:

$$v = \frac{m^2}{m - \sigma^2} \quad (\text{para } m - \sigma^2 \neq 0)$$

Observa-se que v é um fator de dispersão, onde: $\underline{m}$ é o número médio de lagartas por planta viva, e σ^2 é a variância dos dados da contagem dessas lagartas.

3.2.1.3 - Distribuição Binomial Negativa

A sua expressão é:

$$P\{X=x\} = \frac{(k+x-1)!}{x! (k-1)!} \left(1 + \frac{m}{k}\right)^{-k} \left(\frac{m}{m+k}\right)^x \quad (\text{para } x=0,1,2,\dots)$$

onde $P\{X=x\}$ é a probabilidade de ocorrência de $\underline{x}$ lagartas numa planta viva.

Os parâmetros desta distribuição são: $\underline{m}$ e $\underline{k}$, sendo estimados por $\hat{m}$ e $\hat{k}$.

O parâmetro $\underline{m}$ é o número médio de lagartas por planta viva, e $\underline{k}$ é o fator de dispersão que pode ser estimado através de vários métodos, todos eles aproximados. O método escolhido pela sua simplicidade, foi o derivado da equação da variância da Distribuição Binomial Negativa:

$$\sigma^2 = m + \frac{m^2}{k} ,$$

e portanto,

$$k = \frac{m^2}{\sigma^2 - m} ,$$

sendo m e σ^2 estimados por $\hat{m}$ e $\hat{\sigma}^2$, obtidos da forma usual.

Assim:

$$\hat{k} = \frac{\hat{m}^2}{\hat{\sigma}^2 - \hat{m}} .$$

ASCOMBE (1950) chamou a atenção para o fato de que este método não é muito eficiente para valores de $\hat{k}$ menores que 4, a menos que $\hat{m}$ seja também menor que 4. (O autor dá um diagrama mostrando a eficiência da estimação de k desta forma).

Para melhorar a aproximação de $\hat{k}$, o valor obtido é substituído na equação de máxima verossimilhança:

$$n \ln(1 + \frac{\hat{m}}{\hat{k}}) = \sum \left(\frac{A(x)}{\hat{k} + x} \right) ,$$

onde: n é o número de plantas;

$\ln$ é o logaritmo neperiano;

$A(x)$ é o número total de contagens que excedem $\underline{x}$.

Esta distribuição é um caso particular da Distribuição Binomial Generalizada, onde $m > \sigma^2$; o que caracteriza as "Distribuições de Contágio".

3.2.1.3.1 - Fórmula recorrente para a obtenção de $P\{X=x\}$

Na prática, a obtenção das probabilidades através da fórmula apresentada depara com um inconveniente, ou seja, geralmente o valor de $\bar{k}$ é fracionário, o que impossibilita o cálculo do seu número fatorial.

Pode-se contornar esse problema com o uso da fórmula recorrente que facilita a obtenção dos valores das $P\{X=x\}$. Conhecido:

$$P\{X=x\} = \left(1 + \frac{m}{k}\right)^{-k}, \quad (\text{para } x=0),$$

a expressão da fórmula recorrente é a seguinte:

$$P\{X=x\} = \frac{k+x-1}{x!} \left(\frac{m}{m+k}\right) P\{X=x-1\}, \quad (\text{para } x=1,2,\dots)$$

que, conforme se observa, elimina o cálculo do fatorial de k .

3.2.1.4 - Distribuição de Neyman Tipo A, com dois parâmetros

É outra das "Distribuições de Contágio". A sua expressão geradora de probabilidade, simplificada por BEALL (1939), é:

$$P\{X=0\} = e^{-m_1(1-e^{-m_2})},$$

que é a frequência esperada de zeros, e

$$P\{X=n+1\} = \frac{m_1 m_2 e^{-m_2}}{n+1} \sum_{t=0}^n \frac{m_2^t}{t!} P\{X=n-t\}, \quad p/n=0,1,2,\dots$$

que traduz a probabilidade de se observar um valor $n+1$, inteiro e não negativo, de lagartas por planta viva, onde:

m_1 é o número médio de colônias de lagartas no campo;

m_2 é o número médio de lagartas dentro de colônias.

Os estimadores de m_1 e m_2 são, respectivamente:

$$\hat{m}_1 = \frac{\hat{m}^2}{\hat{\sigma}^2 - \hat{m}}, \quad \hat{m}_2 = \frac{\hat{\sigma}^2 - \hat{m}}{\hat{m}},$$

sendo $\hat{m}$ e $\hat{\sigma}^2$ estimativas da média e da variância da distribuição de frequências, respectivamente.

Assim,

$$m = m_1 \cdot m_2$$

e $\sigma^2 = m_1 \cdot m_2(1 + m_2)$, logo:

$$\sigma^2 = m(1 + m_2).$$

3.2.1.5 - Distribuição de Poisson Dupla de Thomas

É uma "Distribuição de Contágio" semelhante à Distribuição de Neyman Tipo A.

A frequência esperada de zeros é dada por:

$$P\{X=0\} = e^{-m_1}, \quad \text{para } m_1 > 0, .$$

e a probabilidade de se observar um valor x , inteiro, não negativo, de lagartas por planta viva, é a que se segue:

$$P\{X=x\} = \sum_{r=1}^x \frac{m_1^r e^{-m_1}}{r!} \cdot \frac{(r \cdot m_2)^{x-r} e^{-r \cdot m_2}}{(x-r)!}$$

para $x=1,2,\dots$, onde:

m_1 é o número médio de colônias de lagartas no campo;

m_2 é o número médio de lagartas dentro de colônias.

A estima de m_1 e m_2 é feita através do método dos momentos, proposto por THOMAS (1949), onde:

$$\mu_1' = m_1(1 + m_2)$$

$$\mu_2 = m_1(1 + 3m_2 + m_2^2),$$

sendo, então, o primeiro momento em relação à origem $\mu_1' = m$, e o segundo momento em relação à média $\mu_2 = \sigma^2$, ou seja, a média e a variância da população, respectivamente.

Uma vez selecionadas as distribuições já anteriormente descritas, proceder-se-ã aos testes de aderência (qui-quadrado) dos dados de cada uma delas, a fim de se determinar a que melhor descreve o comportamento dos dados em estudo.

3.2.2 - Escolha de Uma Unidade Amostral

Com o objetivo de se comparar métodos de amostragem, será necessário definir uma unidade amostral para a coleta das amostras.

Assim, comparar-se-ã a eficiência relativa dos diferentes tipos de unidades amostrais apresentados na Figura 1, onde cada quadrinho representa uma planta.

A fim de se estudar a eficiência relativa dos diferentes tipos de unidades amostrais propostos, a população das plantas foi dividida em 33 blocos.

Para a constituição dos blocos, as 36 linhas da população de plantas foram divididas em 3 grupos de 12 linhas. Cada um dos grupos foi, por sua vez, subdividido em subgrupos constituídos de 12 plantas por linha.

Figura 1 - Tipos de unidades amostrais.

Geraram-se, assim, 33 blocos de 12x12 plantas, enumerados de 1 a 33, conforme ilustrado na Figura 2.

Nº DE LINHAS										
	1	2	...	12	13	...	24	25	...	36
Nº DE P L A N T A S	1 2 : : 12 13	1			2			3		
	13 : : 24	4			5			6		
	24 : : :	7			8			9		
		10			11			12		
		13			14			15		
		16			17			18		
		19			20			21		
		22			23			24		
		25			26			27		
		28			29			30		
132 : : :	31			32			33			

Figura 2 - Blocos de 12x12 plantas, enumerados de 1 a 33.

As 144 plantas de cada bloco foram agrupadas independentemente para cada tipo de unidade amostral considerada, gerando, assim, no bloco, os seguintes números de unidades amostrais:

Tipo 1 :	144 unidades amostrais
Tipo 2 :	72 unidades amostrais
Tipo 3 :	72 unidades amostrais
Tipo 4 :	48 unidades amostrais
Tipo 5 :	48 unidades amostrais
Tipo 6 :	48 unidades amostrais
Tipo 7 :	48 unidades amostrais
Tipo 8 :	36 unidades amostrais
Tipo 9 :	36 unidades amostrais
Tipo 10:	36 unidades amostrais
Tipo 11:	36 unidades amostrais
Tipo 12:	36 unidades amostrais
Tipo 13:	12 unidades amostrais
Tipo 14:	12 unidades amostrais
Tipo 15:	12 unidades amostrais
Tipo 16:	12 unidades amostrais
Tipo 17:	12 unidades amostrais

Devido ao formato em diagonal das unidades amostrais dos Tipos 5, 6, 10, 11, 15 e 16, a sua localiza

ção nos blocos será feita segundo o esquema preconizado pela Regra de Sarrus, isto é, agregando as colunas necessárias para se formar as respectivas diagonais completas.

Assim sendo, calcular-se-á a variância referente a cada tipo de unidade amostral, para cada um dos 33 blocos, que será dividida pelo número de plantas abrangidas nela.

Dessa maneira, a variável usada para a comparação será:

$$Q_{hj} = \frac{S_{hj}^2}{v_{hj}},$$

onde:

$$S_{hj}^2 = \frac{\sum_{i=1}^{n_{hj}} (y_{hji} - \bar{y}_{hj})^2}{n_{hj} - 1}, \text{ é a variância dos dados}$$

associados às unidades amostrais do tipo j, no bloco h;

y_{hji} é o número de lagartas na i -ésima unidade amostral do j -ésimo tipo, no h -ésimo bloco;

$\bar{y}_{hj}$ é o número médio de lagartas por unidade amostral do tipo j, no bloco h;

n_{hj} é o número de unidades amostrais do tipo j, no bloco h, constante para j fixo;

v_{hj} é o número de plantas por unidade amostral do tipo j, no bloco h.

Observa-se, assim, que a variável Q_{hj} mede a variância das unidades amostrais do tipo $\underline{j}$, no bloco $\underline{h}$, reduzida a nível de planta. Na sua obtenção, admitiu-se constante o número de plantas no bloco, isto é, $n_{hj} \cdot v_{hj} = 144$. Desprezou-se, conseqüentemente, a possibilidade de falhas oriundas de morte ou inexistência de plantas.

No caso de se levar em conta as falhas referidas anteriormente, isto é, uma população de plantas variável por bloco, pode-se definir uma outra variável:

$$Q'_{hj} = \frac{S_{hj}^2}{v_{hj}}$$

onde:

$$S_{hj}^2 = \frac{\sum_{i=1}^{n_{hj}} y_{hji}^2 - 2R_{hj} \sum_{i=1}^{n_{hj}} x_{hji} \cdot y_{hji} + R_{hj}^2 \sum_{i=1}^{n_{hj}} x_{hji}^2}{n_{hj} - 1},$$

é a variância dos dados associados às unidades amostrais do tipo $\underline{j}$, no bloco $\underline{h}$;

$$R_{hj} = \frac{\sum_{i=1}^{n_{hj}} y_{hji}}{\sum_{i=1}^{n_{hj}} x_{hji}}, \text{ define uma razão associada às unidades amostrais do tipo } \underline{j}, \text{ no bloco } \underline{h};$$

x_{hji} é o número de plantas vivas na i -ésima unidade amostral do j -ésimo tipo, no h -ésimo bloco;

e as demais variáveis já definidas anteriormente.

Assim sendo, será comparada a eficiência relativa dos 17 tipos de unidades amostrais apresentados na Figura 1, por ambos os processos, isto é, considerando-se as variáveis Q_{hj} e Q'_{hj} , através do Teste Não-Paramétrico de Friedman e de Comparações Múltiplas, segundo CAMPOS (1984).

No caso de não ocorrência de lagartas ou de ocorrência de empates para todas as observações dentro de um bloco, este será descartado para aplicação dos métodos não paramétricos em apreço.

Dessa maneira, poder-se-á optar por um dos processos anteriormente detalhados, e por um dos tipos de unidade amostral, ou seja, aquele que apresentar menor valor nas variáveis (Q_{hj} ou Q'_{hj}).

3.2.3 - Aplicação e Comparação de Três Métodos de Amostragem

Uma vez selecionada a unidade amostral mais eficiente dentre as comparadas, proceder-se-á à aplicação e comparação dos métodos de amostragem a seguir:

- A: Amostragem Simples ao Acaso;
- B: Amostragem Estratificada;
- C: Amostragem por Conglomerados.

Para cada método de amostragem será usada a metodologia detalhada por COCHRAN (1965), com uma amostragem prévia para dimensionamento da amostra definitiva onde optar-se-á por um d (semi-amplitude do intervalo de confiança da média) igual a 30% da média obtida na amostra prévia para os Métodos A e B.

No Método C será adotada a Amostragem por Conglomerados em dois estágios, com mesmo número de unidades secundárias por unidade primária. Nestas condições, o dimensionamento da amostra será feito conforme exposto a seguir.

É sabido que:

$$V(\bar{y}) = \frac{N-n}{N} \frac{S_1^2}{n} + \frac{1}{n^2} \sum_{i=1}^N \frac{n}{N} \frac{M-m}{M} \frac{S_{2i}^2}{m} \quad (1)$$

onde:

$V(\bar{y})$ é a variância da média amostral das unidades secundárias (U.S.);

N é o número de unidades primárias (U.P.) na população;

n é o número de U.P. na amostra;

$S_1^2 = \frac{\sum_{i=1}^N (\bar{y}_i - \bar{\bar{y}})^2}{N - 1}$, é a variância dos valores médios das U.P. com $\bar{y}_i$ sendo a média das U.S. dentro do conglomerado i e $\bar{\bar{y}}$ sendo a média geral a nível de U.S.;

M é o número de U.S. em cada U.P.;

m é o número de U.S. amostradas em cada U.P.;

$$S_{2i}^2 = \frac{\sum_{j=1}^M (y_{ij} - \bar{y}_i)^2}{M - 1}, \text{ é a variância das U.S. dentro das U.P., com } y_{ij} \text{ sendo o valor da unidade amostral } \underline{j} \text{ do conglomerado } \underline{i}.$$

Definindo-se:

$$S_2^2 = \frac{1}{N} \sum_{i=1}^N S_{2i}^2,$$

pode-se expressar (1) na forma:

$$V(\bar{y}) = \frac{1}{n} \left[S_1^2 - \frac{S_2^2}{M} \right] + \frac{S_2^2}{n m} - \frac{S_1^2}{N}. \quad (2)$$

Considerando-se a seguinte função custo:

$$C = c_1 n + c_2 n.m, \quad (3)$$

onde:

C é o custo total do levantamento;

c_1 é o custo do acesso às U.P.;

c_2 é o custo do levantamento de uma U.S. dentro das U.P.

Assim, para se achar os valores de $\underline{m}$ e $\underline{n}$ que minimizam a variância $V(\bar{y})$ para um dado custo (fixo) recorre-se a Mínimos Condicionados com o uso dos Multipl*ic*adores de Lagrange, sendo:

$$Z = V(\bar{y}) + \lambda g \quad (4)$$

$$\text{com } g = C - c_1 n - c_2 nm. \quad (5)$$

Substituindo-se em (4) os valores (2) e (5) fica:

$$Z = \frac{1}{n} \left[S_1^2 - \frac{S_2^2}{M} \right] + \frac{S_2^2}{nm} - \frac{S_1^2}{N} + \lambda [C - c_1 n - c_2 nm]. \quad (6)$$

Derivando-se parcialmente (6), virã:

$$\frac{\partial Z}{\partial n} = - \frac{1}{n^2} \left[S_1^2 - \frac{S_2^2}{M} \right] + \frac{1}{n^2} \frac{S_2^2}{m} - \lambda c_1 - \lambda c_2 m \quad (7)$$

$$\frac{\partial Z}{\partial m} = - \frac{1}{m^2} \frac{S_2^2}{n} - \lambda c_2 n \quad (8)$$

$$g = C - c_1 n - c_2 nm \quad (9)$$

Igualando-se as derivadas a zero e adicionando a função de restrição (5), configura-se um sistema de equações com três incógnitas. Daí obtêm-se:

$$m(\text{ótimo}) = \frac{S_2}{\sqrt{S_1^2 - \frac{S_2^2}{M}}} \cdot \sqrt{\frac{c_1}{c_2}} \quad (10)$$

Note-se que a estrutura da fórmula (10) é a que seria de se esperar, pois, se os elementos foram distribuídos aleatoriamente pelas unidades, isto é, se as U.P. tiveram a mesma eficiência que um elemento, a variância do valor médio de uma U.P. será S^2/M , de modo que $[S_1^2 - (S_2^2/M)]$ terá valor zero. Isso dá para $m(\bar{o}timo)$ o valor infinito, ou seja, a enumeração integral de todas as U.P.

Ao contrário, quanto maior for a variância S_1^2 entre os valores médios das U.P., relativamente à variância dentro das U.P., menor será o valor de $m(\bar{o}timo)$. E quanto maior forem as despesas c_1 , do acesso à U.P., em relação às despesas c_2 , da obtenção dos dados de levantamento de um elemento qualquer da U.P., maior será o $m(\bar{o}timo)$.

A seguir, pode-se obter o valor de $n(\bar{o}timo)$ (número de U.P.), utilizando a equação (9), que igualada a zero é:

$$C - c_1 n - c_2 n m = 0 ,$$

ou seja:

$$n(\bar{o}timo) = \frac{C}{c_1 + c_2 m(\bar{o}timo)} \quad (11)$$

Na prática geralmente não se conhecem os valores de S_1^2 e S_2^2 , porém pode-se obter suas estimativas através de uma amostra prévia, na qual selecionam-se n' U.P., tomando-se m' elementos de cada U.P. selecionada.

Assim, o valor de $m(\bar{o}timo)$ da fórmula (10) é estimado através de uma amostra prévia com:

$$\hat{m}(\bar{o}timo) = \frac{s_2}{\sqrt{s_1^2 - \frac{s_2^2}{m'}}} \sqrt{\frac{c_1}{c_2}}, \quad (12)$$

onde:

s_1^2 é a variância entre os valores médios das U.P., obtida na amostra prévia;

s_2^2 é a variância entre os elementos dentro das U.P.;

e supondo-se a mesma função custo (3), a estimativa para o número de U.P. necessárias no levantamento é:

$$\hat{n}(\bar{o}timo) = \frac{C}{c_1 + c_2 \hat{m}(\bar{o}timo)}. \quad (13)$$

Uma vez dimensionadas as amostras para os três métodos considerados, tornar-se viável a estruturação das amostras propriamente ditas, permitindo, consequentemente, o confronto entre eles.

4. RESULTADOS E DISCUSSÕES

4.1 - Ajuste às Distribuições Propostas

A distribuição de frequências da população em estudo foi extraída do censo apresentado no Apêndice, no final deste trabalho.

Em todas as distribuições estudadas, embora os parâmetros integrantes dos cálculos das probabilidades sejam com um número reduzido de casas decimais, foram considerados para efeito de cálculos com 9 casas.

4.1.1 - Distribuição de Poisson

Através dos dados obteve-se:

$$\hat{m} = \frac{\sum_i x_i f_i}{\sum_i f_i} = \frac{368}{4.174} = 0,088$$

$$p_{(0)} = P\{X=0\} = e^{-\hat{m}} = 0,9156$$

$$p_{(1)} = P\{X=1\} = e^{-\hat{m}} \frac{\hat{m}}{1!} = 0,0807$$

$$p_{(2)} = P\{X \geq 2\} = 1 - [p_{(0)} + p_{(1)}] = 0,0037$$

O teste $\bar{e}$ apresentado na Tabela 1.1.

Tabela 1 - Teste de χ^2 para a Distribuição de Poisson.

x_i (nº de lagartas/ planta viva)	Frequência Observada (fr.O) (nº de plantas vivas com x_i lagartas)	Frequência Esperada (fr.E) $n \cdot p_i$	$\frac{(fr.O - fr.E)^2}{fr.E}$
0	4.043	3.821,756	12,808
1	54	336,944	237,599
≥ 2	77	15,300	248,826
			$\chi^2_c = 499,232^{**}$

Assim, o nível mínimo de significância (n.m.s.), de acordo com a tabela apropriada ao teste, é:

$$\text{n.m.s.} = P_0 (\chi_0^2 \geq 499,232) < 0,01.$$

Portanto, rejeita-se, ao nível de 1% de probabilidade, a hipótese de que os dados podem ser estudados através da Distribuição de Poisson.

4.1.2 - Distribuição Binomial Generalizada

Sendo:

$$\hat{m} = 0,088 \quad , \quad \hat{\sigma}^2 = 0,5189$$

$$\hat{\chi} = 5,8847 \quad , \quad \hat{\psi} = -4,8847 \quad , \quad \hat{v} = -0,0181 \quad ,$$

$$p_{(0)} = P\{X=0\} = \hat{\chi}^{\hat{v}} = 0,9685$$

$$p_{(1)} = P\{X=1\} = \hat{v} \hat{\chi}^{\hat{v}-1} \hat{\psi} = 0,0145$$

$$p_{(2)} = P\{X=2\} = \frac{\hat{v}(\hat{v}-1)}{2!} \hat{\chi}^{\hat{v}-2} \hat{\psi}^2 = 0,0061$$

$$p_{(3)} = P\{X=3\} = \frac{\hat{v}(\hat{v}-1)(\hat{v}-2)}{3!} \hat{\chi}^{\hat{v}-3} \hat{\psi}^3 = 0,0034$$

$$p_{(4)} = P\{X=4\} = \frac{\hat{v}(\hat{v}-1)(\hat{v}-2)(\hat{v}-3)}{4!} \hat{\chi}^{\hat{v}-4} \hat{\psi}^4 = 0,0021$$

$$P_{(5)} = P\{X=5\} = \frac{\hat{v}(\hat{v}-1)(\hat{v}-2)(\hat{v}-3)(\hat{v}-4)}{5!} \hat{x}^{\hat{v}-5} \hat{\psi}^5 = 0,0014$$

$$P_{(6)} = P\{X=6\} = \frac{\hat{v}(\hat{v}-1)(\hat{v}-2)(\hat{v}-3)(\hat{v}-4)(\hat{v}-5)}{6!} \hat{x}^{\hat{v}-6} \hat{\psi}^6 = 0,0010$$

$$P_{(7)} = P\{X \geq 7\} = 1 - \sum_{i=1}^6 P_{(i)} = 0,0028 .$$

O teste é apresentado na Tabela 2.

Tabela 2 - Teste de χ^2 para a Distribuição Binomial Generalizada.

x_i	fr.O	fr.E	$(\text{fr.O}-\text{fr.E})^2/\text{fr.E}$
0	4.043	4.042,588	0,000
1	54	60,567	0,712
2	31	25,591	1,143
3	20	14,368	2,207
4	5	8,943	1,744
5	6	5,970	0,000
6	6	4,144	0,831
≥ 7	9	11,823	0,674
	4.174	4.174	$\chi^2_{\text{C}}=7,312$

$$\text{n.m.s.} = P_0(\chi_0^2 \geq 7,312) > 0,10 .$$

Não se rejeita H_0 ao nível de 10% de probabilidade. Logo, os dados podem ser estudados através da Distribuição Binomial Generalizada.

4.1.3 - Distribuição Binomial Negativa

4.1.3.1 - Obtenção do parâmetro $\underline{k}$

Conforme já visto:

$$\hat{m} = 0,088$$

$$\hat{\sigma}^2 = 0,5189$$

um primeiro valor de $\underline{k}$ foi obtido da forma a seguir:

$$\hat{k} = \frac{\hat{m}^2}{\hat{\sigma}^2 - \hat{m}^2} = 0,01805$$

A fim de melhorar a aproximação de $\hat{k}$, e segundo ASCOMBE (1950), foram obtidos os valores de $A(x)$ apresentados na Tabela 3, juntamente com a distribuição de frequência, a serem aplicados na expressão de máxima verossimilhança:

$$n \ln \left(1 + \frac{\hat{m}}{\hat{k}} \right) = \sum_{x=0}^{19} \left[\frac{A(x)}{\hat{k} + x} \right]$$

Pelo método iterativo e tomando como valor inicial o anteriormente determinado, ou seja, $\hat{k} = 0,01805$, tem-se:

$$n \ln \left(1 + \frac{\hat{m}}{\hat{k}} \right) = 7.397,78556$$

$$\sum_{x=0}^{19} \left[\frac{A(x)}{\hat{k} + x} \right] = 7.377,90727$$

Tabela 3 - Obtenção de $A(x)$.

x_i	f_i	$A(x)$
0	4.043	131
1	54	77
2	31	46
3	20	26
4	5	21
5	6	15
6	6	9
7	1	8
8	2	6
9	0	6
10	0	6
11	3	3
12	0	3
13	1	2
14	0	2
15	0	2
16	0	2
17	0	2
18	1	1
19	1	0

Nesta primeira tentativa, tem-se:

$$\begin{aligned} \text{diferença} &= 7.397,78556 - 7.377,90727 \\ &= 19,87829 \quad ; \end{aligned}$$

para $\bar{k} = 0,017$, na segunda tentativa:

$$n \ln \left(1 + \frac{\hat{m}}{\bar{k}} \right) = 7.606,34531$$

$$\sum_{x=0}^{19} \left[\frac{A(x)}{\bar{k} + x} \right] = 7.827,13484$$

$$\begin{aligned} \text{diferença} &= 7.606,34531 - 7.827,13484 \\ &= -220,78952 \end{aligned}$$

$\bar{k}$	diferença
0,01805	19,87829
$\bar{k}$	0
0,017	-220,78952

Assim, interpolando-se, obtem-se:

$$\bar{k} = 0,01796 \quad .$$

4.1.3.2 - Cálculo das probabilidades e teste de χ^2

Com os valores dos parâmetros $\hat{m}$, $\hat{\sigma}^2$ e $\hat{k}$ obte-se:

$$p_{(0)} = P\{X=0\} = \left(1 + \frac{\hat{m}}{\hat{k}}\right)^{-\hat{k}} = 0,9686$$

$$p_{(1)} = P\{X=1\} = \hat{k} \left(\frac{\hat{m}}{\hat{m} + \hat{k}}\right) P\{X=0\} = 0,0145$$

$$p_{(2)} = P\{X=2\} = \frac{(\hat{k} + 1)}{2} \left(\frac{\hat{m}}{\hat{m} + \hat{k}}\right) P\{X=1\} = 0,0061$$

$$p_{(3)} = P\{X=3\} = \frac{(\hat{k} + 2)}{3} \left(\frac{\hat{m}}{\hat{m} + \hat{k}}\right) P\{X=2\} = 0,0034$$

$$p_{(4)} = P\{X=4\} = \frac{(\hat{k} + 3)}{4} \left(\frac{\hat{m}}{\hat{m} + \hat{k}}\right) P\{X=3\} = 0,0021$$

$$p_{(5)} = P\{X=5\} = \frac{(\hat{k} + 4)}{5} \left(\frac{\hat{m}}{\hat{m} + \hat{k}}\right) P\{X=4\} = 0,0014$$

$$p_{(6)} = P\{X=6\} = \frac{(\hat{k} + 5)}{6} \left(\frac{\hat{m}}{\hat{m} + \hat{k}}\right) P\{X=5\} = 0,0010$$

$$p_{(7)} = P\{X \geq 7\} = 1 - \sum_{i=0}^6 p_{(i)} = 0,0029$$

Através destes valores foi estruturado o teste apresentado na Tabela 4.

Tabela 4 - Teste de χ^2 para a Distribuição Binomial Negativa.

x_i	fr.O	fr.E	$(\text{fr.O}-\text{fr.E})^2/\text{fr.E}$
0	4.043	4.042,919	0,000
1	54	60,330	0,664
2	31	25,510	1,182
3	20	14,255	2,316
4	5	8,935	1,733
5	6	5,965	0,000
6	6	4,144	0,831
≥ 7	9	11,943	0,725
	4.174	4.174	$\chi^2_c = 7,451$

$$\text{n.m.s.} = P_0(\chi_0^2 \geq 7,451) > 0,10 \quad .$$

Não se rejeita H_0 . Portanto, os dados podem ser estudados através da Distribuição Binomial Negativa.

4.1.4 - Distribuição de Neyman Tipo A com Dois Parâmetros

Sendo conhecidos $\hat{\sigma}^2 > \hat{m}$, obteve-se:

$$\hat{m}_1 = \frac{\hat{m}^2}{\hat{\sigma}^2 - \hat{m}} = 0,018$$

$$\hat{m}_2 = \frac{\hat{\sigma}^2 - \hat{m}}{\hat{m}} = 4,885$$

Com os valores de $\hat{m}_1$ e $\hat{m}_2$, foram calculadas as probabilidades respectivas:

$$p_{(0)} = P\{X=0\} = e^{-\hat{m}_1(1-e^{-\hat{m}_2})} = 0,9822$$

e, com o auxílio da fórmula de Beall, onde:

$$P\{X=n+1\} = \frac{\hat{m}_1 \hat{m}_2 e^{-\hat{m}_2}}{n+1} \sum_{t=0}^n \frac{\hat{m}_2^t}{t!} P\{X=n-t\},$$

obteve-se:

- para $n=0$:

$$p_{(1)} = P\{X=1\} = \hat{m}_1 \hat{m}_2 e^{-\hat{m}_2} P\{X=0\} = 0,0007$$

- para $n=1$:

$$p_{(2)} = P\{X=2\} = \frac{\hat{m}_1 \hat{m}_2 e^{-\hat{m}_2}}{2} [\hat{m}_2 P\{X=0\} + P\{X=1\}] = 0,0060$$

- para $n=2$:

$$p_{(3)} = P\{X=3\} = \frac{\hat{m}_1 \hat{m}_2 e^{-\hat{m}_2}}{3} \left[\frac{\hat{m}_2^2}{2!} P\{X=0\} + \hat{m}_2 P\{X=1\} + P\{X=2\} \right] = 0,0026$$

- para $n=3$:

$$p_{(4)} = P\{X=4\} = \frac{\hat{m}_1 \hat{m}_2 e^{-\hat{m}_2}}{4} \left[\frac{\hat{m}_2^3}{3!} P\{X=0\} + \frac{\hat{m}_2^2}{2!} P\{X=1\} + \hat{m}_2 P\{X=2\} + P\{X=3\} \right] = 0,0032$$

- para $n=4$:

$$p_{(5)} = P\{X=5\} = \frac{\hat{m}_1 \hat{m}_2 e^{-\hat{m}_2}}{5} \left[\frac{\hat{m}_2^4}{4!} P\{X=0\} + \frac{\hat{m}_2^3}{3!} P\{X=1\} + \frac{\hat{m}_2^2}{2!} P\{X=2\} + \hat{m}_2 P\{X=3\} + P\{X=4\} \right] = 0,0031$$

- para $n=5$:

$$p_{(6)} = P\{X=6\} = \frac{\hat{m}_1 \hat{m}_2 e^{-\hat{m}_2}}{6} \left[\frac{\hat{m}_2^5}{5!} P\{X=0\} + \frac{\hat{m}_2^4}{4!} P\{X=1\} + \frac{\hat{m}_2^3}{3!} P\{X=2\} + \frac{\hat{m}_2^2}{2!} P\{X=3\} + \hat{m}_2 P\{X=4\} + P\{X=5\} \right] = 0,0025$$

- para $n \geq 6$:

$$p_{(7)} = P\{X \geq 7\} = 1 - \sum_{i=0}^6 p_{(i)} = 0,0041$$

que são apresentados na Tabela 5.

Tabela 5 - Teste de χ^2 para a Distribuição de Neyman Tipo A, com dois parâmetros.

x_i	fr.0	fr.E	$(\text{fr.0}-\text{fr.E})^2/\text{fr.E}$
0	4.043	4.099,897	0,790
1	54	2,733	961,577
2	31	6,677	88,614
3	20	10,874	7,659
4	5	13,286	5,168
5	6	12,994	3,764
6	6	10,601	1,997
≥ 7	9	16,938	3,720
	4.174	4.174	$\chi^2_C=1.073,288^{**}$

$$\text{n.m.s.} = P_0(\chi_0^2 \geq 1.073,288) < 0,01$$

Rejeita-se H_0 . Os dados não podem ser estudados através da Distribuição de Neyman Tipo A.

4.1.5 - Distribuição Poisson Dupla de Thomas

Considerando-se:

$$\mu_1' = m_1 (1 + m_2)$$

$$\mu_2 = m_1 (1 + 3m_2 + m_2^3)$$

e sendo conhecidos os valores de $\mu_1' = \hat{m}$ e $\mu_2 = \hat{\sigma}^2$, obtiveram-se:

$$\hat{m}_1 = 0,017 \quad \text{e} \quad \hat{m}_2 = 4,081 .$$

Com os valores de $\hat{m}_1$ e $\hat{m}_2$, foram calculadas as probabilidades respectivas.

$$p_{(0)} = P\{X=0\} = e^{-\hat{m}_1} = 0,9828$$

e aplicando-se:

$$P\{X=x\} = \sum_{r=1}^x \frac{\hat{m}_1^r e^{-\hat{m}_1}}{r!} \frac{(r\hat{m}_2)^{x-r} e^{-r\hat{m}_2}}{(x-r)!}$$

obteve-se:

$$p_{(1)} = P\{X=1\} = \hat{m}_1 e^{-\hat{m}_1} e^{-\hat{m}_2} = 0,0003$$

$$p_{(2)} = P\{X=2\} = \frac{\hat{m}_1 e^{-(\hat{m}_1+\hat{m}_2)}}{2!} \left\langle 2\hat{m}_2 + \hat{m}_1 e^{-\hat{m}_2} \right\rangle = 0,0012$$

$$p_{(3)} = P\{X=3\} = \hat{m}_1 e^{-(\hat{m}_1+\hat{m}_2)} \left\langle \frac{\hat{m}_1^2 e^{-2\hat{m}_2}}{3!} + 2\hat{m}_1 \hat{m}_2 \frac{e^{-\hat{m}_2}}{2!} + \frac{\hat{m}_2^2}{2!} \right\rangle = 0,0024$$

e assim por diante, cujos valores são:

$$p_{(4)} = P\{X=4\} = 0,0033$$

$$p_{(5)} = P\{X=5\} = 0,0033$$

$$p_{(6)} = P\{X=6\} = 0,0027$$

$$p_{(7)} = P\{X \geq 7\} = 1 - \sum_{i=0}^6 p(i) = 0,0040 ,$$

apresentadas na Tabela 6.

Tabela 6 - Teste de χ^2 para a Distribuição Poisson Dupla de Thomas.

x_i	fr.O	fr.E	$(\text{fr.O}-\text{fr.E})^2/\text{fr.E}$
0	4.043	4.102,204	0,854
1	54	1,202	2.319,904
2	31	4,905	138,844
3	20	10,010	9,970
4	5	13,622	5,458
5	6	13,910	4,498
6	6	11,374	2,539
>7	9	16,773	3,602
	4.174	4.174	$\chi^2_c=2.485,669^{**}$

$$n.m.s. = P_0(\chi_0^2 \geq 2.485,669) < 0,01$$

Rejeita-se H_0 e, portanto, os dados não podem ser estudados através da Distribuição Poisson Dupla de Thomas.

4.1.6 - Discussão

Os dados obtidos no censo apresentado no Apêndice ajustaram-se apropriadamente às distribuições Binomial Generalizada e Binomial Negativa, sendo esta um caso particular da primeira, que mostra assim seu caráter geral, pois aplica-se num grande número de situações. Note-se que se os dados da distribuição de frequências observadas tivessem se ajustado a uma distribuição de Poisson, a Distribuição Binomial Generalizada também teria mostrado bom ajuste, pois Poisson é um outro caso particular dela.

As distribuições de Neyman Tipo A e Poisson Dupla de Thomas não se ajustaram aos dados com tanta precisão, demonstrando que são formas matemáticas de descrever situações biológicas muito particulares.

Assim, levando-se em conta que está se tentando descrever matematicamente um equilíbrio ecológico dinâmico, no qual interagem inúmeros fatores, conhecidos e desconhecidos, é preciso optar por uma distribuição que recorra a expressões estocásticas, não determinísticos, de maneira a tornar a sua aplicabilidade mais geral.

É evidente, então, que procura-se, com este trabalho, estabelecer conclusões de valor relativo, pois

o objetivo final é achar generalizações. Não obstante, com os dados da experimentação descritos, não se poderá chegar a uma conclusão definitiva, reconhecendo-se a necessidade de repetir o ensaio em anos sucessivos, e várias vezes, durante o ciclo da cultura e da praga.

Assim, observa-se que o presente trabalho tende a estabelecer os primeiros estudos para se desenvolver um conhecimento apropriado de *Scrobipalpula absoluta* (Meyrick) no campo, e definir uma mecânica que deverá se repetir para cada experimentação.

Desta forma, poder-se-á repetir a mesma sistemática de trabalho para condições análogas em outras culturas e pragas de importância.

Tentar descrever matematicamente dados biológicos é uma tarefa complexa. Assim, pode-se explicá-los em cada situação, baseando-se numa série de pressuposições impostas, e de valor momentâneo, conseqüentemente, sem grande valor prático.

O mais recomendável, então, é achar expressões gerais que abranjam um grande número de situações particulares.

Por isso, opta-se por considerar que os dados da experimentação em estudo, que de agora em diante serão chamados de "população", distribuem-se segundo uma distribuição Binomial Generalizada.

4.2 - Escolha de Uma Unidade Amostral

Para se selecionar uma unidade amostral entre as 17 propostas na Figura 1, procedeu-se segundo já foi descrito, ao cálculo da variável $Q_{hj} = S_{hj}^2/v_{hj}$, cujos valores são apresentados na Tabela 7, e nos quais, a efeito de comparação, foi aplicado o Teste de Friedman.

Os blocos 14, 16, 17 e 24 foram descartados por não ocorrência de lagartas ($Q_{hj} = 0$), e os blocos 11, 20 e 29 também o foram por apresentarem empates em todos os tratamentos (Q_{hj} constante em todas as unidades amostrais).

Assim sendo, os valores entre parênteses, ao lado de cada valor de Q_{hj} , representam a ordem, crescente, dentro do bloco.

Os valores dos R_j foram:

$R_1 = 369,0$	$R_7 = 256,0$	$R_{13} = 182,5$
$R_2 = 328,5$	$R_8 = 204,0$	$R_{14} = 234,5$
$R_3 = 332,5$	$R_9 = 235,0$	$R_{15} = 170,0$
$R_4 = 279,0$	$R_{10} = 176,0$	$R_{16} = 130,0$
$R_5 = 291,0$	$R_{11} = 170,0$	$R_{17} = 154,0$
$R_6 = 280,0$	$R_{12} = 186,0$	

e sendo $k=17$ e $n=26$, aplicou-se:

$$\chi_r^2 = \frac{12}{nk(k+1)} \sum_{j=1}^k R_j^2 - 3n(k+1) \quad ,$$

onde, R_j é a soma das ordens atribuídas aos dados do tratamento ou Tipo j , nos n blocos;

k é o número de tratamentos ou Tipos de unidades amostrais comparadas.

Desta forma, obteve-se a estatística de Friedman:

$$\chi_r^2 = 119,700$$

e, corrigindo para casos de empates, com:

$$C = 1 - \frac{\sum_h T_h}{nk(k^2 - 1)}$$

onde: $T_h = t_{ih}^3 - k$;

t_{ih} é o número de observações empatadas no grupo i do bloco h ;

obteve-se:

$$C = 0,969 \quad .$$

Logo, a estatística de Friedman fica:

$$\chi_r^{2'} = \frac{\chi_r^2}{C} = 123,519 \quad ,$$

significativo ao nível $\alpha < 0,005$.

Conclui-se, então, que pelo menos um dos 17 tipos de unidades amostrais difere dos demais no referente à sua eficiência.

Para estudar essa diferença, procedeu-se às Comparações Múltiplas, envolvendo todos os pares de tratamentos (tipos de unidades amostrais), segundo a seguinte expressão para a diferença mínima significativa (d.m.s.):

$$\text{d.m.s.} = Q^+ \sqrt{\frac{nk(k+1)}{12}},$$

sendo Q^+ o valor extraído de tabelas apropriadas.

Obteve-se, assim:

$$\text{d.m.s.}(5\%) = 125,94$$

e

$$\text{d.m.s.}(1\%) = 142,52.$$

Os resultados obtidos são apresentados na Tabela 8.

A seguir foram efetuados os mesmos cálculos e comparações para a variável $Q'_{hj} = \frac{S_{hj}}{v_{hj}}$, já descrita anteriormente, e cujos resultados são apresentados nas Tabelas 9 e 10, respectivamente.

Na Tabela 9 foram eliminados os blocos 14, 16, 17 e 24 por inexistência de lagartas ($Q'_{hj} = 0$).

Tabela 8 - Diferença entre os R_j para Comparações Múltiplas envolvendo todos os tratamentos - variável Q_{hj} .

	R_1	R_2	R_3	R_5	R_6	R_4	R_7	R_9	R_{14}	R_8	R_{12}	R_{13}	R_{10}	R_{11}	R_{15}	R_{17}	R_{16}
R_1	-																
R_3	36,5	-															
R_2	40,5	4,0	-														
R_5	78,0	41,5	37,5	-													
R_6	89,0	52,5	48,5	11,0	-												
R_4	90,0	53,5	49,5	12,0	1,0	-											
R_7	113,0	76,5	72,5	35,0	24,0	23,0	-										
R_9	134,0*	97,5	93,5	56,0	45,0	44,0	21,0	-									
R_{14}	134,5*	98,0	94,0	56,5	45,5	44,5	21,5	0,5	-								
R_8	165,0**	128,5*	124,5	87,0	76,0	75,0	52,0	31,0	30,5	-							
R_{12}	183,0**	146,5**	142,5*	105,0	94,0	93,0	70,0	49,0	48,5	18,0	-						
R_{13}	186,5**	150,0**	146,0**	108,5	97,5	96,5	73,5	52,5	52,0	21,5	3,5	-					
R_{10}	193,0**	156,5**	152,5**	115,0	104,0	103,0	80,0	59,0	58,5	28,0	10,0	6,5	-				
R_{11}	199,0**	162,5**	158,5**	121,0	110,0	109,0	86,0	65,0	64,5	34,0	16,0	12,5	6,0	-			
R_{15}	199,0**	162,5**	158,5**	121,0	110,0	109,0	86,0	65,0	64,5	34,0	16,0	12,5	6,0	0,0	-		
R_{17}	215,0**	178,5**	174,5**	137,0*	126,0*	125,0	102,0	81,0	80,5	50,0	32,0	28,5	22,0	16,0	16,0	-	
R_{16}	239,0**	202,5**	198,5**	161,0**	150,0**	149,0**	126,0**	105,0*	104,5	74,0	56,0	52,5	46,0	40,0	40,0	24,0	-

Para o Teste de Friedman obtiveram-se os seguintes valores dos R_j 's.

$R_1 = 387,0$	$R_7 = 279,5$	$R_{13} = 193,5$
$R_2 = 337,5$	$R_8 = 239,0$	$R_{14} = 251,0$
$R_3 = 349,5$	$R_9 = 278,0$	$R_{15} = 209,0$
$R_4 = 298,0$	$R_{10} = 239,0$	$R_{16} = 136,0$
$R_5 = 326,5$	$R_{11} = 204,0$	$R_{17} = 180,5$
$R_6 = 301,0$	$R_{12} = 228,0$	

e sendo $k=17$ e $n=29$, obteve-se:

$$\chi_r^2 = 97,664 ,$$

e corrigindo para caso de empates, onde:

$$C = 0,998 ,$$

resultou:

$$\chi_r^{2'} = \frac{\chi_r^2}{C} = 97,867 ,$$

significativo ao nível $\alpha < 0,005$.

Conclui-se, analogamente ao caso anterior, que pelo menos um dos 17 tipos de unidades amostrais difere dos demais, no que se refere à sua eficiência.

Nas Comparações Múltiplas envolvendo todos os pares de tratamentos (tipos de unidades amostrais), pa

Tabela 9 - Valores de Q'_{hj} para os 17 diferentes tipos de unidades amostrais, nos 29 blocos.

BLOCOS	TIPOS DE UNIDADES AMOSTRAIS																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	1,3774 (14,0)	1,3156 (9,0)	1,3208 (12,0)	1,3202 (11,0)	1,2661 (6,5)	1,4079 (16,0)	1,2661 (6,5)	1,2601 (6,0)	1,2370 (4,0)	1,3539 (13,0)	1,2417 (5,0)	1,3160 (10,0)	1,3036 (15,0)	2,1472 (17,0)	1,0769 (2,0)	0,8145 (1,0)	1,1364 (3,0)
2	0,2721 (13,0)	0,2659 (12,0)	0,2650 (11,0)	0,2597 (8,0)	0,2594 (6,0)	0,2658 (9,5)	0,2658 (9,5)	0,2532 (2,0)	0,2541 (3,0)	0,2594 (7,0)	0,2800 (14,0)	0,2543 (4,0)	0,4375 (17,0)	0,3547 (16,0)	0,2566 (5,0)	0,2917 (15,0)	0,2031 (1,0)
3	3,2894 (17,0)	3,1663 (14,0)	3,2095 (16,0)	3,1628 (13,0)	3,0968 (12,0)	3,0886 (10,0)	3,0906 (11,0)	3,0519 (6,0)	3,0011 (4,0)	3,0701 (7,0)	3,1695 (15,0)	2,9746 (5,0)	3,0569 (9,0)	2,7067 (3,0)	2,5944 (2,0)	2,7352 (4,0)	2,3956 (1,0)
4	0,5212 (15,0)	0,5138 (10,5)	0,5138 (10,5)	0,5095 (9,0)	0,5152 (12,5)	0,5152 (12,5)	0,5066 (7,0)	0,5019 (6,0)	0,5077 (8,0)	0,5171 (14,0)	0,4992 (5,0)	0,4909 (4,0)	0,4607 (16,0)	0,6507 (16,0)	0,6603 (17,0)	0,4533 (1,0)	0,4658 (3,0)
5	2,7667 (11,0)	2,7934 (13,0)	2,7462 (10,0)	2,7255 (8,0)	2,7645 (12,0)	2,7255 (8,0)	2,7254 (8,0)	2,7044 (4,5)	2,7044 (4,5)	2,7044 (4,5)	2,7044 (4,5)	2,7044 (4,5)	2,5821 (2,0)	3,1193 (15,0)	2,7873 (12,0)	2,5255 (1,0)	3,1026 (14,0)
6	0,1504 (12,0)	0,1478 (10,0)	0,1487 (11,0)	0,2594 (17,0)	0,1469 (9,0)	0,1451 (7,0)	0,1451 (7,0)	0,2575 (16,0)	0,1452 (8,0)	0,1424 (4,0)	0,1423 (3,0)	0,1433 (5,0)	0,2543 (15,0)	0,1270 (1,0)	0,1834 (13,0)	0,1764 (2,0)	0,2432 (14,0)
7	0,0680 (11,0)	0,0665 (16,0)	0,0665 (15,0)	0,0649 (14,0)	0,0648 (12,0)	0,0648 (12,0)	0,0648 (12,0)	0,0632 (6,5)	0,0639 (8,0)	0,0645 (9,0)	0,0646 (10,0)	0,0632 (6,5)	0,0509 (1,0)	0,0542 (4,5)	0,0528 (3,0)	0,0542 (4,5)	0,0514 (2,0)
8	0,1487 (11,0)	0,1444 (9,5)	0,1444 (9,5)	0,1400 (7,0)	0,1400 (7,0)	0,1400 (7,0)	0,2019 (17,0)	0,1355 (3,0)	0,1386 (4,0)	0,1387 (5,0)	0,1937 (15,5)	0,1937 (15,5)	0,1320 (2,0)	0,1882 (14,0)	0,1658 (12,0)	0,1667 (13,0)	0,1060 (1,0)
9	0,3753 (14,0)	0,3718 (13,0)	0,3803 (15,0)	0,3594 (8,0)	0,3816 (16,0)	0,3595 (9,5)	0,3595 (9,5)	0,3557 (6,5)	0,3687 (11,0)	0,3568 (5,0)	0,3557 (6,5)	0,3689 (12,0)	0,5655 (17,0)	0,3305 (4,0)	0,2921 (1,0)	0,2944 (2,0)	0,2104 (3,0)
10	0,5241 (11,0)	0,5045 (8,0)	0,5195 (10,0)	0,4889 (5,0)	0,4973 (6,0)	0,7569 (17,0)	0,4976 (7,0)	0,4641 (2,0)	0,7148 (15,0)	0,5104 (7,0)	0,4862 (4,0)	0,4840 (3,0)	0,6655 (14,0)	0,5939 (12,0)	0,7314 (16,0)	0,3516 (1,0)	0,5100 (13,0)
11	0,1728 (6,0)	0,2121 (5,0)	0,1734 (16,0)	0,1732 (12,5)	0,1732 (12,5)	0,1732 (12,5)	0,1732 (12,5)	0,1730 (7,5)	0,1730 (10,0)	0,1731 (10,0)	0,1730 (7,5)	0,1730 (10,0)	0,1716 (1,0)	0,1717 (4,0)	0,1716 (2,5)	0,1744 (17,0)	0,1716 (2,5)
12	0,2183 (15,0)	0,2158 (12,0)	0,2137 (10,0)	0,2167 (13,0)	0,2168 (14,0)	0,2090 (8,0)	0,2156 (11,0)	0,2084 (7,0)	0,2084 (7,0)	0,2084 (7,0)	0,2084 (7,0)	0,2084 (7,0)	0,2084 (7,0)	0,2084 (7,0)	0,2084 (7,0)	0,2084 (7,0)	0,2084 (7,0)
13	0,6688 (10,0)	0,6474 (8,0)	0,6478 (9,0)	0,9666 (17,0)	0,7302 (12,0)	0,6326 (7,0)	0,6258 (6,0)	0,9473 (16,0)	0,6910 (11,0)	0,6221 (5,0)	0,6179 (4,0)	0,6042 (3,0)	0,8019 (14,0)	0,7802 (13,0)	0,4642 (1,0)	0,6471 (2,0)	0,9074 (15,0)
15	0,2355 (13,0)	0,2297 (12,0)	0,2283 (11,0)	0,2223 (8,0)	0,2227 (9,0)	0,2179 (6,0)	0,2240 (10,0)	0,2549 (16,0)	0,2138 (4,0)	0,2376 (14,0)	0,2167 (5,0)	0,2194 (7,0)	0,1965 (2,0)	0,1463 (1,0)	0,3855 (17,0)	0,2392 (15,0)	0,2666 (3,0)
18	0,2711 (12,0)	0,2855 (16,0)	0,0714 (9,0)	0,0735 (11,0)	0,0696 (7,0)	0,0686 (5,0)	0,0726 (10,0)	0,0850 (15,0)	0,0675 (3,0)	0,0674 (2,0)	0,0714 (8,0)	0,0801 (13,0)	0,0690 (6,0)	0,1096 (17,0)	0,0495 (1,0)	0,0830 (14,0)	0,0677 (4,0)
19	0,2723 (15,0)	0,6543 (13,0)	0,6731 (17,0)	0,6354 (7,0)	0,6553 (14,0)	0,6531 (12,0)	0,6403 (10,0)	0,6255 (5,0)	0,6405 (11,0)	0,6382 (9,0)	0,6268 (8,0)	0,6069 (4,5)	0,0668 (1,5)	0,0669 (12,0)	0,0889 (13,0)	0,5565 (3,0)	0,5123 (1,0)
20	0,0049 (14,0)	0,0069 (10,0)	0,0069 (10,0)	0,0069 (9,0)	0,0070 (16,0)	0,0069 (7,0)	0,0069 (8,0)	0,0069 (8,0)	0,0070 (15,0)	0,0069 (3,0)	0,0069 (6,0)	0,0069 (6,0)	0,0069 (4,5)	0,0069 (12,0)	0,0069 (3,0)	0,0069 (1,5)	0,0076 (17,0)
21	0,0630 (14,0)	0,0604 (9,5)	0,0604 (9,5)	0,0603 (7,0)	0,0608 (6,0)	0,0602 (16,0)	0,0603 (8,0)	0,0603 (8,0)	0,0603 (15,0)	0,0603 (15,0)	0,0603 (15,0)	0,0603 (15,0)	0,0603 (15,0)	0,0603 (15,0)	0,0603 (15,0)	0,0603 (15,0)	0,0603 (15,0)
22	0,0679 (14,0)	0,0660 (11,0)	0,0667 (12,0)	0,0783 (15,0)	0,0647 (9,0)	0,0655 (10,0)	0,0669 (13,0)	0,0636 (5,0)	0,0636 (5,0)	0,0636 (5,0)	0,0644 (4,0)	0,0629 (4,0)	0,0629 (4,0)	0,0629 (4,0)	0,0629 (4,0)	0,0629 (4,0)	0,0629 (4,0)
23	0,1239 (17,0)	0,1252 (13,0)	0,1226 (11,0)	0,1212 (7,0)	0,1213 (7,0)	0,1213 (7,0)	0,1218 (8,0)	0,1225 (9,0)	0,1218 (8,0)	0,1225 (9,0)	0,1205 (5,0)	0,1226 (10,0)	0,1181 (3,0)	0,1709 (14,0)	0,1146 (2,0)	0,1131 (1,0)	0,1737 (15,0)
25	2,0162 (11,0)	1,9847 (10,0)	2,2919 (17,0)	1,9514 (8,0)	2,2614 (16,0)	1,9041 (9,0)	1,9493 (6,0)	1,9512 (7,0)	2,2238 (14,0)	1,9238 (5,0)	1,9211 (4,0)	2,2305 (15,0)	1,8359 (3,0)	2,1757 (13,0)	1,7134 (1,0)	1,8141 (2,0)	2,1277 (12,0)
26	0,0204 (15,0)	0,0203 (13,0)	0,0203 (14,0)	0,0205 (11,0)	0,0202 (12,0)	0,0202 (13,0)	0,0198 (10,0)	0,0204 (15,0)	0,0193 (7,0)	0,0197 (9,0)	0,0189 (6,0)	0,0193 (8,0)	0,0173 (4,0)	0,0166 (2,0)	0,0169 (3,0)	0,0158 (1,0)	0,0173 (5,0)
27	0,0204 (15,0)	0,0203 (13,0)	0,0203 (14,0)	0,0205 (11,0)	0,0202 (12,0)	0,0202 (13,0)	0,0198 (10,0)	0,0204 (15,0)	0,0193 (7,0)	0,0197 (9,0)	0,0189 (6,0)	0,0193 (8,0)	0,0173 (4,0)	0,0166 (2,0)	0,0169 (3,0)	0,0158 (1,0)	0,0173 (5,0)
28	0,0344 (16,0)	0,0344 (15,0)	0,0341 (14,0)	0,0340 (12,0)	0,0340 (11,0)	0,0336 (8,0)	0,0340 (8,0)	0,0336 (7,0)	0,0332 (6,0)	0,0347 (17,0)	0,0340 (9,0)	0,0340 (10,0)	0,0325 (5,0)	0,0325 (3,5)	0,0327 (2,0)	0,0324 (2,0)	0,0329 (1,0)
29	0,0069 (14,0)	0,0069 (12,0)	0,0069 (13,0)	0,0069 (11,0)	0,0069 (10,0)	0,0070 (17,0)	0,0069 (9,0)	0,0069 (9,0)	0,0069 (8,0)	0,0070 (16,0)	0,0069 (5,0)	0,0069 (5,0)	0,0069 (4,0)	0,0069 (3,0)	0,0069 (2,0)	0,0069 (1,0)	0,0069 (1,0)
30	0,2819 (11,0)	0,2594 (8,5)	0,2594 (8,5)	0,2567 (6,0)	0,2624 (12,0)	0,2624 (12,0)	0,2624 (12,0)	0,2538 (4,0)	0,2666 (14,0)	0,2541 (5,0)	0,2597 (10,0)	0,2541 (5,0)	0,2326 (1,0)	0,2592 (7,0)	0,4522 (17,0)	0,2521 (3,0)	0,2439 (2,0)
31	0,4788 (13,0)	0,4581 (11,5)	0,4581 (11,5)	0,4476 (9,0)	0,4677 (14,0)	0,4389 (7,0)	0,4497 (10,0)	0,4260 (5,0)	0,5160 (16,0)	0,4196 (4,0)	0,4405 (8,0)	0,4309 (6,0)	0,3739 (1,0)	0,3868 (2,0)	0,5084 (15,0)	0,6915 (17,0)	0,4540 (3,0)
32	0,0896 (15,0)	0,0887 (14,0)	0,0887 (13,0)	0,0878 (9,5)	0,0878 (9,5)	0,0878 (9,5)	0,0878 (9,5)	0,0869 (8,0)	0,0869 (8,0)	0,0869 (8,0)	0,0869 (8,0)	0,0869 (8,0)	0,0869 (8,0)	0,0869 (8,0)	0,0869 (8,0)	0,0869 (8,0)	0,0869 (8,0)
33	0,2044 (15,0)	0,2019 (14,0)	0,1987 (13,0)	0,1892 (9,0)	0,1913 (10,0)	0,1822 (5,5)	0,1822 (5,5)	0,2653 (16,0)	0,1834 (7,0)	0,1746 (3,0)	0,1725 (2,0)	0,1918 (11,0)	0,3065 (17,0)	0,1793 (4,0)	0,1868 (8,0)	0,1350 (1,0)	0,1978 (12,0)

$R_1 = 387$ $R_2 = 337,5$ $R_3 = 340,5$ $R_4 = 298$ $R_5 = 326,5$ $R_6 = 30$ $R_7 = 279,5$ $R_8 = 239$ $R_9 = 278$ $R_{10} = 239$ $R_{11} = 204$ $R_{12} = 228$ $R_{13} = 193,5$ $R_{14} = 251$ $R_{15} = 209$ $R_{16} = 136$ $R_{17} = 180,5$

ra grandes amostras, obtiveram-se os valores apresentados na Tabela 10, e comparados com a d.m.s. a seguir:

$$d.m.s.(5\%) = 133,00$$

e

$$d.m.s.(1\%) = 150,52 .$$

4.2.1 - Discussão

Pode-se observar nas Tabelas 8 e 10, que a tendência, para ambas as formas de estimativas, é a de diminuição dos valores das variáveis (Q_{hj} e Q'_{hj} , respectivamente), dentro do mesmo v_{hj} .

Os maiores valores foram observados nas unidades amostrais cujo formato acompanhou a direção das linhas da cultura (Tipos 3, 5, 9 e 14). Os menores valores foram constatados nas unidades amostrais de formato diagonalizado (Tipos 6, 7, 10, 11, 15 e 16). Os tipos agrupados de unidades amostrais (12 e 17) se situaram entre os de menores valores.

Por outro lado, ao aumentar o valor de v_{hj} , isto é, o tamanho da unidade amostral, a tendência também é de diminuição das variâncias, ou seja, melhorar a eficiência.

Tabela 10 - Diferenças entre os R_j para Comparações Múltiplas envolvendo todos os tratamentos - variável Q'_{hj} .

	R_1	R_3	R_2	R_5	R_6	R_4	R_7	R_9	R_{14}	R_8	R_{10}	R_{12}	R_{15}	R_{11}	R_{13}	R_{17}	R_{16}
R_1	-																
R_3	37,5	-															
R_2	49,5	12,0	-														
R_5	60,5	23,0	11,0	-													
R_6	86,0	48,5	36,5	25,5	-												
R_4	89,0	51,5	39,5	28,5	3,0	-											
R_7	107,5	70,0	58,0	47,0	21,5	18,5	-										
R_9	109,0	71,5	59,5	48,5	23,0	20,0	1,5	-									
R_{14}	136,0*	98,5	86,5	75,5	50,0	47,0	28,5	27,0	-								
R_8	148,0*	110,5	98,5	87,5	62,0	59,0	40,5	39,0	12,0	-							
R_{10}	148,0*	110,5	98,5	87,5	62,0	59,0	40,5	39,0	12,0	0,0	-						
R_{12}	159,0**	121,5	109,5	98,5	73,0	70,0	51,5	50,0	23,0	11,0	11,0	-					
R_{15}	178,0**	140,5*	128,5	117,5	92,0	89,0	70,5	69,0	42,0	30,0	30,0	19,0	-				
R_{11}	183,0**	145,5**	133,5*	122,5	97,0	94,0	75,5	74,0	47,0	35,0	35,0	24,0	5,0	-			
R_{13}	193,5**	156,0**	144,0**	133,0*	107,5	104,5	86,0	84,5	57,5	45,5	45,5	34,5	15,5	10,5	-		
R_{17}	206,5**	169,0**	157,0**	146,0**	120,5	117,5	99,0	97,5	70,5	58,5	58,5	47,5	28,5	23,5	13,0	-	
R_{16}	251,0**	213,5**	201,5**	190,5**	165,0**	162,0**	143,5**	142,0*	115,0	103,0	103,0	92,0	73,0	68,0	57,5	44,5	-

Assim sendo, observa-se que as unidades amostrais com maior número de plantas e em formas diagonalizadas seriam as mais convenientes, pois apresentam as menores variâncias. Porém, na prática, esses tipos de unidades amostrais são muito difíceis de aplicação, pois não é possível individualizar as diagonais no campo.

Em consequência, optou-se pelos tipos agrupados (12 e 17) e, dentre estes, o Tipo 12 (quadrado com $v_{hj} = 4$ plantas) pela sua maior simplicidade, uma vez que entre eles não existe diferença estatisticamente significativa.

A mesma análise foi desenvolvida para dois processos de obtenção das estimativas, o baseado diretamente na variável em estudo (Q_{hj}) e o baseado na estimativa por razão (Q'_{hj}), levando ambas às mesmas conclusões. Desta forma, optou-se pela forma de trabalho mais simples, neste caso a de estimativa pelo primeiro processo (Q_{hj}), que como já foi dito, considera as plantas inexistentes ou mortas como plantas vivas, sem lagartas.

4.3 - Aplicação e Comparação de Três Métodos de Amostragem

Para a aplicação dos três métodos de amostragem propostos em 3.2.3, foi primeiramente realizada, para cada um deles, uma amostragem prévia, com o objetivo de obtenção de subsídios para o dimensionamento da amostra definitiva.

Uma vez obtidas as amostras definitivas associadas aos três métodos propostos, através de cada uma delas foram estimados os seguintes parâmetros:

- Número médio de lagartas por unidade amostral ($\bar{y}$);
- Número total de lagartas na cultura ($\bar{Y}$);
- Variância do número médio de lagartas por unidade amostral [$V(\bar{y})$];
- Intervalo de confiança do número médio de lagartas por unidade amostral.

É importante salientar que em todos os casos foi usado o processo de obtenção das estimativas, baseado diretamente na variável em estudo, ou seja, número de lagartas por planta, considerando as plantas mortas ou inexistentes como sendo vivas e livres de lagartas.

Uma pressuposição foi a de que o ataque da praga à cultura é leve, o que surgiu em decorrência da

inspeção visual da área em estudo.

4.3.1 - Amostragem Simples ao Acaso (A.S.A.)

4.3.1.1 - Pré-amostragem

Supondo-se um ataque leve, considerou-se apropriado o levantamento de $n=50$ unidades amostrais do Tipo 12, segundo a decisão adotada no item 4.2.1.

4.3.1.1.1 - Aleatorização

O critério adotado para a localização das unidades amostrais a serem levantadas foi o da escolha aleatória das coordenadas da planta ocupante do vértice superior esquerdo da unidade amostral, considerando-se a população disposta na forma apresentada no Apêndice. Assim, supondo-se, por exemplo, que foram sorteadas as coordenadas 12, 5, a unidade amostral será constituída pelas plantas das

linha 12	,	planta 5
linha 12	,	planta 6
linha 13	,	planta 5
linha 13	,	planta 6

Segundo o critério adotado, foram sorteados, numa primeira etapa, 50 números de 1 a 35 (correspondentes às linhas da cultura), com reposição. A seguir sortearam-se outros 50 números de 1 a 131 (correspondentes à ordem da planta dentro de cada linha), também com reposição. Formaram-se, assim, 50 pares de coordenadas, de acordo com a ordem de sorteio nas duas etapas.

A partir de cada planta sorteada (correspondente a cada par de coordenadas), foi estruturada uma unidade amostral do Tipo 12 (4 plantas), conforme critério já citado anteriormente, tomando-se o cuidado de que nenhuma planta fosse levantada mais de uma vez.

4.3.1.1.2 - Dimensionamento

Uma vez estruturada a amostra prévia, procedeu-se ao levantamento, obtendo-se os seguintes resultados:

- Número de lagartas encontradas:

$$\sum_{i=1}^n y_i = 15 ;$$

- Variância dos dados decorrentes do levantamento:

$$s^2 = 0,6224 ;$$

Têm-se ainda os seguintes valores:

- Número total de unidades amostrais possíveis:

$$N = 1.188 ;$$

- Número de unidades amostrais levantadas:

$$n = 50 ;$$

A partir dos resultados obtidos na amostra prévia, foi dimensionada a amostra definitiva, conforme es tabelecido em 3.2.3, obtendo-se:

$$d = 0,30 \quad \bar{y} = 0,09$$

$$n_0 = \frac{t^2 s^2}{d^2} = 307,4$$

e corrigido para população finita:

$$n = \frac{n_0}{1 + \frac{n_0}{N}} = 245 ,$$

desta forma, o tamanho da amostra definitiva é de 245 un dades amostrais.

4.3.1.2 - Amostragem propriamente dita

Sendo necessárias $n=245$ unidades amostrais e considerando-se que já haviam sido levantadas 50 na amostra prévia, foram extraídas mais 195 depois de se proceder à aleatorização da forma descrita.

Obtiveram-se os seguintes valores:

$$N = 1.188$$

$$n = 245$$

$$\sum_{i=1}^n y_i = 92$$

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n} = 0,376$$

$$\bar{Y} = N \cdot \bar{y} \cong 447$$

$$s^2 = 2,6699$$

$$\hat{V}(\bar{y}) = \frac{N - n}{N} \frac{s^2}{n} = 0,0087$$

$$\hat{V}(\bar{Y}) = N^2 \hat{V}(\bar{y}) = 12.208,3045$$

- I.C.($\bar{Y}$) com 95% de nível de confiança:

$$0,1932 < \bar{Y} < 0,5578$$

- I.C.(Y) com 95% de nível de confiança:

$$229,5434 \leq Y \leq 662,6689$$

4.3.2 - Amostragem Estratificada

Partindo do pressuposto de que a população em estudo (Apêndice) é totalmente desconhecida, optou-se por dividi-la em 33 partes iguais de 12x12 plantas, cada uma delas sendo considerada um estrato, coincidentemente com os blocos definidos no item 3.2.2, usados para a obtenção da unidade amostral mais eficiente dentre as propostas.

4.3.2.1 - Pré-amostragem

Para efeito de estruturação da amostra prévia, foram levantadas 5 unidades amostrais por estrato, que constitui o mínimo recomendado por CAMPOS (1982)^{1/}.

^{1/} Informação pessoal.

4.3.2.1.1 - Aleatorização

A aleatorização foi efetuada da mesma forma enunciada em 4.3.1.1.1 para A.S.A., sendo agora aplicada separadamente em cada um dos 33 estratos.

4.3.2.1.2 - Dimensionamento

Procedeu-se ao levantamento das 165 unidades amostrais constituintes da amostra prévia, cujos resultados estão resumidos na Tabela 11.

A partir dos dados da Tabela 11, obteve-se

$$\bar{y}_{est.} = \sum_h W_h \bar{y}_h = 0,570$$

$$d = 0,30 \bar{y}_{est.} = 0,171$$

$$V = \frac{d^2}{t^2} = 0,0073$$

e procedeu-se ao dimensionamento através da Partilha de Neyman, obtendo-se:

$$n_0 = \frac{\left(\sum_{h=1}^{33} N_h \cdot s_h \right)^2}{N^2 \cdot V} = 159,2$$

Tabela 11 - Valores obtidos da amostra prévia para dimensionamento na Amostragem Estratificada.

ESTRATO (h)	N_h	n_h	$\bar{y}_h$	s_h^2	$N_h s_h$
1	36	5	0,8	1,699999	46,938257
2	36	5	0,6	1,799999	48,299068
3	36	5	1,4	3,799999	70,176919
4	36	5	0,4	0,799999	32,199378
5	36	5	1,2	7,199999	96,598136
6	36	5	0,2	0,199999	16,099689
7	36	5	0,6	0,799999	32,199378
8	36	5	0,4	0,799999	32,199378
9	36	5	0,2	0,199999	16,099689
10	36	5	1,0	4,999999	80,498447
11	36	5	1,0	4,999999	80,498447
12	36	5	0,4	0,799999	32,199378
13	36	5	0,4	0,799999	32,199378
14	36	5	0	0	0
15	36	5	0,2	0,199999	16,099689
16	36	5	0	0	0
17	36	5	0	0	0
18	36	5	0,2	0,199999	16,099689
19	36	5	2,4	12,799999	128,797515
20	36	5	0	0	0
21	36	5	0,6	1,799999	48,299068
22	36	5	0	0	0
23	36	5	0,2	0,199999	16,099689
24	36	5	0	0	0
25	36	5	2,6	33,799999	209,295962
26	36	5	0,2	0,199999	16,099689
27	36	5	0	0	0
28	36	5	0,2	0,199999	16,099689
29	36	5	0	0	0
30	36	5	0,2	0,199999	16,099689
31	36	5	2,0	7,499999	98,590060
32	36	5	0,4	0,799999	32,199378
33	36	5	1,0	1,999999	50,911688
	1.188	165		88,799999	1.280,897355

Corrigindo para população finita:

$$n = \frac{n_0}{1 + \frac{\sum_h N_h s_h^2}{N^2 v}} = 121,5$$

e pela Partilha propriamente dita:

$$n_h = n \frac{N_h s_h}{\sum_h N_h s_h},$$

obtiveram-se:

$n_1 = 4,453 \Rightarrow n_1 = 5$	$n_2 = 4,582 \Rightarrow n_2 = 5$
$n_3 = 6,657 \Rightarrow n_3 = 7$	$n_4 = 3,055 \Rightarrow n_4 = 5$
$n_5 = 9,163 \Rightarrow n_5 = 10$	$n_6 = 1,527 \Rightarrow n_6 = 5$
$n_7 = 3,055 \Rightarrow n_7 = 5$	$n_8 = 3,055 \Rightarrow n_8 = 5$
$n_9 = 1,527 \Rightarrow n_9 = 5$	$n_{10} = 7,636 \Rightarrow n_{10} = 8$
$n_{11} = 7,636 \Rightarrow n_{11} = 8$	$n_{12} = 3,055 \Rightarrow n_{12} = 5$
$n_{13} = 3,055 \Rightarrow n_{13} = 5$	$n_{14} = 0 \Rightarrow n_{14} = 5$
$n_{15} = 1,527 \Rightarrow n_{15} = 5$	$n_{16} = 0 \Rightarrow n_{16} = 5$
$n_{17} = 0 \Rightarrow n_{17} = 5$	$n_{18} = 1,527 \Rightarrow n_{18} = 5$
$n_{19} = 12,218 \Rightarrow n_{19} = 13$	$n_{20} = 0 \Rightarrow n_{20} = 5$
$n_{21} = 4,582 \Rightarrow n_{21} = 5$	$n_{22} = 0 \Rightarrow n_{22} = 5$
$n_{23} = 1,527 \Rightarrow n_{23} = 5$	$n_{24} = 0 \Rightarrow n_{24} = 5$
$n_{25} = 19,855 \Rightarrow n_{25} = 20$	$n_{26} = 1,527 \Rightarrow n_{26} = 5$
$n_{27} = 0 \Rightarrow n_{27} = 5$	$n_{28} = 1,527 \Rightarrow n_{28} = 5$
$n_{29} = 0 \Rightarrow n_{29} = 5$	$n_{30} = 1,527 \Rightarrow n_{30} = 5$
$n_{31} = 9,353 \Rightarrow n_{31} = 10$	$n_{32} = 3,055 \Rightarrow n_{32} = 5$
$n_{33} = 4,829 \Rightarrow n_{33} = 5$	

Em todos os casos onde $n_h < 5$ adotou-se o critério de se tomar o mínimo de 5 unidades amostrais por estrato, conforme já enunciado anteriormente.

4.3.2.2 - Amostragem propriamente dita

Foram levantadas as unidades amostrais complementares para integralizar os n_h 's dimensionados, tomando-se o cuidado de não levantar as plantas já amostradas anteriormente. Dessa maneira, ficou constituída a amostra definitiva, da qual obtiveram-se os valores apresentados na Tabela 12 e, a partir dela, calcularam-se as estimativas dos parâmetros da população em estudo.

$$\bar{y}_{est.} = \frac{1}{N} \sum_h N_h \bar{y}_h = 0,386$$

$$\bar{Y} = N \cdot \bar{y}_{est.} \cong 459$$

$$\hat{V}(\bar{y}_{est.}) = \frac{1}{N^2} \sum_h \frac{N_h(N_h - n_h)}{n_h} s_h^2 = 0,0040$$

$$\hat{V}(\bar{Y}) = N^2 \hat{V}(\bar{y}_{est.}) = 5.643,4710$$

- I.C. ($\bar{Y}$) com 95% de nível de confiança:

$$0,2616 \leq \bar{Y} \leq 0,5095$$

Tabela 12 - Valores obtidos da amostra definitiva para a determinação dos parâmetros na Amostragem Estratificada.

ESTRATO (h)	N_h	n_h	$\bar{y}_h$	s_h^2	$N_h s_h$
1	36	5	0,800	1,699999	28,80
2	36	5	0,600	1,799999	21,60
3	36	7	1,000	2,999999	36,00
4	36	5	0,400	0,799999	14,40
5	36	10	0,600	3,599999	21,60
6	36	5	0,200	0,199999	7,20
7	36	5	0,600	0,799999	21,60
8	36	5	0,400	0,799999	14,40
9	36	5	0,200	0,199999	7,20
10	36	8	0,625	3,124999	22,50
11	36	8	0,625	3,124999	22,50
12	36	5	0,400	0,799999	14,40
13	36	5	0,400	0,799999	14,40
14	36	5	0	0	0
15	36	5	0,200	0,199999	7,20
16	36	5	0	0	0
17	36	5	0	0	0
18	36	5	0,200	0,199999	7,20
19	36	13	0,923	5,743589	33,23
20	36	5	0	0	0
21	36	5	0,600	1,799999	21,60
22	36	5	0	0	0
23	36	5	0,200	0,199999	7,20
24	36	5	0	0	0
25	36	20	0,650	8,449999	23,40
26	36	5	0,200	0,199999	7,20
27	36	5	0	0	0
28	36	5	0,200	0,199999	7,20
29	36	5	0	0	0
30	36	5	0,200	0,199999	7,20
31	36	10	1,100	4,322222	39,60
32	36	5	0,400	0,799999	14,40
33	36	5	1,000	1,999999	36,00
	1.188	206			458,03

- I.C.(Y) com 95% de nível de confiança:

$$310,7889 \leq Y \leq 605,2712$$

4.3.3 - Amostragem por Conglomerados em Dois Estágios

Para este método de Amostragem, também dividiu-se a população em estudo em 33 partes iguais, analogamente ao procedimento adotado em 3.2.2 e 4.3.2, admitindo-se, neste caso, que cada bloco ou estrato anteriormente caracterizado passou a constituir uma Unidade Primária (U.P.) ou Conglomerado. As unidades amostrais do Tipo 12, já definidas, constituem as Unidades Secundárias (U.S.) ou Sub-Unidades.

4.3.3.1 - Pré-amostragem

Procedeu-se à constituição da amostra prévia, considerando-se as $N=33$ Unidades Primárias e as $M=36$ Unidades Secundárias dentro de cada Unidade Primária, optando-se, para o levantamento prévio, por $n'=10$ U.P. e $m'=6$ U.S., em cada U.P.

4.3.3.1.1 - Aleatorização

Sortearam-se, sem reposição, 10 Conglomerados e, em cada um deles, procedeu-se à aleatorização de 6 U.S. a serem levantadas, obedecendo a mesma sistemática já descrita no item 4.3.1.1.1 para A.S.A.

4.3.3.1.2 - Dimensionamento

Foram levantadas as $n'm'=60$ unidades amostrais, a partir das quais calcularam-se os valores apresentados na Tabela 13.

Com os valores da amostra prévia detalhados na Tabela 13, e conhecendo-se:

$$N = 33 \quad , \quad n' = 10$$

e

$$M = 36 \quad , \quad m' = 6 \quad ,$$

calcularam-se:

$$\bar{y} = \frac{\sum_{i=1}^{n'} y_{i.}}{n'} = 3,3$$

$$\bar{\bar{y}} = \frac{1}{n' m'} \sum_{i=1}^{n'} \sum_{j=1}^{m'} y_{ij} = 0,55$$

$$\hat{Y} = N M \bar{\bar{y}} = 653,4$$

Tabela 13 - Valores obtidos da amostra prévia para dimensionamento na Amostragem por Conglomerado em Dois Estágios.

CONGLOMERADO (U.P.)	UNIDADES AMOSTRAIS (U.S.)		$y_{i.}$	$\bar{y}_i$	s^2_{2i}
	y_{ij}				
1	0 0 3	1 2 1	7	1,167	1,3667
5	0 0 0	0 2 0	2	0,333	0,6667
6	0 0 0	4 0 0	4	0,667	2,6667
7	0 0 0	0 1 0	1	0,167	0,1667
8	0 0 0	0 0 2	2	0,333	0,6667
9	0 0 3	0 0 0	3	0,500	1,5000
15	1 1 2	0 0 0	4	0,667	0,6667
16	0 0 0	0 0 0	0	0	0
30	0 0 1	0 0 0	1	0,167	0,1667
33	1 3 0	2 0 3	9	1,500	1,9000
			33	5,500	9,7667

$$s_1^2 = \frac{\sum_{i=1}^{n'} \bar{y}_i^2 - \frac{(\sum_{i=1}^{n'} \bar{y}_i)^2}{n'}}{n' - 1} = 0,2225$$

$$s_2^2 = \frac{\sum_{i=1}^{n'} s_{2i}^2}{n'} = 0,9767$$

$$f_1 = \frac{n'}{N} = 0,303$$

$$f_2 = \frac{m'}{M} = 0,167$$

$$\hat{V}(\bar{y}) = \frac{(1 - f_1)}{n'} s_1^2 + \frac{f_1(1 - f_2)}{n'm'} s_2^2 = 0,0213$$

Para se obter um dimensionamento através desta amostra prévia, necessário se torna fixar os valores da equação (3), apresentada em 3.2.3, isto é,

$$C = c_1 n' + c_2 n' m' ,$$

onde c_1 é o custo do acesso a cada U.P. (Conglomerado) e c_2 é o custo do levantamento de uma U.S. (unidade amostral Tipo 12) dentro das U.P.

Assim, se o local da cultura está a 20 km da cidade "ponto de partida" (no presente caso, San Salvador de Jujuy - R.A.), e supondo-se que o veículo de que se dispõe tem um consumo de 10 km/litro, o gasto em gaso-

lina será de 4 litros por viagem (ida e volta). Considerando-se apenas isso como custo referente às U.P., têm-se $4 \times \text{Cr\$ } 890,00$ (preço de agosto de 1984) = $\text{Cr\$ } 3.560,00$, mas, devido as 10 U.P. estarem no mesmo local, deve-se dividir o custo total entre elas. Logo, $c_1 = \text{Cr\$ } 356,00$ por U.P.

Para o cálculo de c_2 , considerou-se o custo de um Técnico Agrícola, com um salário médio diário de $\text{Cr\$ } 15.000,00$, para levantar as 60 unidades amostrais (U.S.) da amostra prévia. Assim, para uma U.S. tem-se: $\text{Cr\$ } 15.000,00/60$, ou seja, $c_2 = \text{Cr\$ } 250,00$.

Dessa forma, a equação custo fica:

$$C = \text{Cr\$ } 18.560,00 .$$

Uma vez calculados c_1 e c_2 , tem-se, conforme 3.2.3:

$$\hat{m}(\bar{o}timo) = \frac{s_2}{\sqrt{s_1^2 - (s_2^2/m')}} \cdot \sqrt{\frac{c_1}{c_2}} = 4,8$$

ou seja:

$$\hat{m}(\bar{o}timo) = 5 \text{ U.S.}$$

e

$$\hat{n}(\bar{o}timo) = \frac{C}{c_1 + c_2 \hat{m}(\bar{o}timo)} = 11,9$$

isto é:

$$\hat{n}(\bar{o}t\acute{i}mo) = 12 \text{ U.P.}$$

Embora o dimensionamento sugira levantar $\hat{m}(\bar{o}t\acute{i}mo) = 5$ U.S. por U.P., optou-se por manter as 6 U.S. por U.P., já consideradas na prē-amostragem, pois, desprezā-las seria uma perda inútil de informação.

Assim, a amostra definitiva foi tomada com as dimensões $m=6$ e $n=12$. Justifica-se este procedimento pela estrutura já existente na amostra prēvia, e pelo fato de que o número de U.S. complementares a serem levantadas seria acrescido apenas de 2 delas, isto é, 12 ao invés de 10, para as 2 U.P. ainda não levantadas.

4.3.3.2 - Amostragem propriamente dita

Conforme visto no item anterior, a amostra propriamente dita é constituída de $n=12$ Conglomerados (U. P.) e $m=6$ unidades amostrais Tipo 12 (U.S.) dentro de cada Conglomerado.

Assim sendo, foi acrescentada a amostra prēvia de dois Conglomerados com 6 U.S. cada um deles.

A amostra assim constituída é a amostra definitiva apresentada na Tabela 14, e da qual obtiveram-se as estimativas seguintes.

Tabela 14 - Valores obtidos da amostra definitiva para a determinação dos parâmetros na Amostragem por Conglomerados em Dois Estágios.

CONGLOMERADO (U.P.)	UNIDADES AMOSTRAIS (U.S.)		$y_{i.}$	$\bar{y}_i$	s^2_{2i}
	y_{ij}				
1	0 0 3	1 2 1	7	1,167	1,3667
5	0 0 0	0 2 0	2	0,333	0,6667
6	0 0 0	4 0 0	4	0,667	2,6667
7	0 0 0	0 1 0	1	0,167	0,1667
8	0 0 0	0 0 2	2	0,333	0,6667
9	0 0 3	0 0 0	3	0,500	1,5000
14	0 0 0	0 0 0	0	0	0
15	1 1 2	0 0 0	4	0,667	0,6667
16	0 0 0	0 0 0	0	0	0
24	0 0 0	0 0 0	0	0	0
30	0 0 1	0 0 0	1	0,167	0,1667
33	1 3 0	2 0 3	9	1,500	1,9000
			33	5,500	9,7667

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n} = 2,75$$

$$\bar{\bar{y}} = \frac{1}{n \cdot m} \sum_{i=1}^n \sum_{j=1}^m y_{ij} = 0,458$$

$$\hat{Y} = N \cdot M \cdot \bar{\bar{y}} \cong 545$$

$$s_1^2 = \frac{\sum_{i=1}^n \bar{y}_i - \frac{(\sum_{i=1}^n \bar{y}_i)^2}{n}}{n-1} = 0,2279$$

$$s_2^2 = \frac{\sum_{i=1}^n s_2^2 i}{n} = 0,8139$$

$$f_1 = \frac{n}{N} = 0,364$$

$$f_2 = \frac{m}{M} = 0,167$$

$$\hat{V}(\bar{\bar{y}}) = \frac{(1 - f_1)}{n} s_1^2 + \frac{f_1(1 - f_2)}{n \cdot m} s_2^2 = 0,0155$$

- I.C. ($\bar{Y}$) com 95% de nível de confiança:

$$0,2142 \leq \bar{Y} \leq 0,7024$$

- I.C. (Y) com 95% de nível de confiança:

$$254,5011 \leq Y \leq 834,4989 .$$

4.3.4 - Discussão

Com o objetivo de se comparar os valores estimados dos parâmetros obtidos através dos três métodos de Amostragem utilizados neste trabalho, foram calculados os valores verdadeiros dos parâmetros da população, e que são apresentados na Tabela 15.

Tabela 15 - Estimativas e valores verdadeiros dos parâmetros da população.

CARACTERÍSTICAS	MÉTODOS DE AMOSTRAGEM			Valores Verdadeiros
	A.S.A.	Estratificada	Conglomerados	
nº médio de lagartas por unidade amostral	0,376	0,386	0,458	0,310
Total de lagartas na cultura	447	459	545	368
nº de unidades amostrais usadas	245	206	72	1.188
Variância da média amostral	0,0087	0,0040	0,0155	-

Pode-se observar, na Tabela 15, que os três métodos de Amostragem superestimaram o número médio de lagartas por unidade amostral na cultura, sendo que a dife-

rença entre estas estimativas não foi tão marcante, e também não são muito discrepantes dos valores populacionais. Por outro lado, todos os intervalos de confiança incluem os valores verdadeiros, razão pela qual pode-se dizer que os três métodos tiveram um desempenho relativamente satisfatório.

No tocante às variâncias das estimativas das médias, a amostra estratificada apresentou um valor correspondente a 46% do apresentado pela A.S.A., e 26% em relação ao da amostra por Conglomerados.

Um outro aspecto de fundamental importância a se analisar em casos em que o fator econômico é decisivo, é a dimensão da amostra. Neste caso destaca-se claramente a menor exigência do Método por Conglomerados, dado que por este método foram necessárias apenas 72 unidades amostrais, comparativamente as 245 da A.S.A., e as 206 da amostragem estratificada.

Embora não associado à precisão das amostras, também deve ser levado em conta, na escolha do método mais apropriado, que na amostra estratificada é possível identificar, através dos estratos, a distribuição geográfica da praga na cultura, isto é, os locais de maior incidência.

Isto se viabiliza pela existência de uma

associação estreita entre o tamanho de amostra (n_h) em cada estrato, e a incidência da praga. Verifica-se que os baixos valores de n_h se situam nos estratos onde a incidência é baixa, e em contrapartida os altos valores de n_h coincidem com as altas incidências da praga, no entanto, distribuída de forma irregular (Em Grupos) entre as plantas, ocasionando altos valores das variâncias.

Sob o ponto de vista agrônomo, interpreta-se que os estratos com altos valores de n_h reúnem os pontos mais adequados para o estudo das condições predisponentes para o desenvolvimento da praga, no tocante a solo, ventos, radiação solar, umidade, etc. Outra consequência imediata é que a localização desses pontos é de fundamental importância para o produtor em termos de orientação para o controle da praga, refletindo, consequentemente, no fator econômico.

Ainda com referência à localização das maiores incidências nos estratos, pode-se destacar a possibilidade de indicar a direção de uma provável migração de *Scrobipalpula absoluta* (Meyrick), proveniente de outras culturas vizinhas ou de plantas hospedeiras, onde a praga poderia se desenvolver nos períodos de entre-safras. Esta hipótese se evidencia pelo fato de que o microlepidóptero em estudo apresenta baixa capacidade de deslocamento (BE-

NAVENT, 1981).

4.4 - Considerações Finais

Segundo já foi observado, o presente trabalho não apresenta conclusões definitivas, mas tenta contribuir para um mais profundo conhecimento da praga em questão.

Pode-se afirmar, inicialmente, que *Scrobipalpula absoluta* (Meyrick) distribui-se segundo uma distribuição de contágio, a qual eventualmente poderá variar entre as muitas distribuições que descrevem este tipo de dados de contagem de insetos agrupados ou em focos, como é o presente caso.

Optou-se pela Distribuição Binomial Generalizada, devido a que ela abrange um grande número de casos, pois os seus parâmetros são estocásticos, e que eventualmente poderiam acompanhar as variações da intensidade desses "contágios" ou tendências a agrupamentos. Isto não significa que ocasionalmente uma mesma população não possa ser melhor descrita por alguma outra Distribuição de Contágio, como a de Neyman, a Poisson Dupla de Thomas, etc.

Admitida a Distribuição de Contágio, são sugeridas as unidades amostrais do tipo agrupadas, apro-

priadas a estes casos.

Ao se comparar os dois processos de estimativas (o baseado diretamente na variável em estudo - Q_{hj} , e o da estimativa por razão, com a variável Q'_{hj}), é lógico supor a equivalência de ambos, pois numa planta morta ou inexistente não existem lagartas, dado que o inseto em estudo ataca brotos apicais de crescimento ou gemas florais, ocasionando apenas a debilidade e não a sua morte. Apenas em ataques graves destroi folhas e frutos, o que não foi analisado, por não constituir objeto deste trabalho, dado que não foi enfocado o caso de uma cultura muito atacada, sem possibilidade de defesa, e, sim, aquela cultura capaz de produzir resultados rentáveis.

Pelo fato de se ter constatado a maior eficiência nas unidades amostrais em forma de diagonal, pode-se sugerir que, no ato da coleta da amostra, o pesquisador caminhe em diagonal na cultura.

Pode-se admitir que a superestimação obtida nos métodos de amostragem aplicados foi devida à aparição de alguns dados que poderiam ter sido considerados discrepantes. Não se os consideraram como tais por se supor um total desconhecimento "a priori" da população, o que, eventualmente, não daria condições de caracterizá-los como discrepantes.

É evidente, assim, que se faz necessário continuar com estudos deste tipo, repetindo-o também para outros métodos de amostragem, como, por exemplo, o Sequencial (baseado na Razão de Probabilidades, extraída de uma distribuição pressuposta), que poderiam contribuir para generalizações mais precisas.

5. CONCLUSÕES

Do presente trabalho, é possível concluir:

- 1º) A distribuição que melhor descreveu os dados de contagem de lagartas de *Scrobipalpula absoluta* (Meyrick) na cultura de tomate (*Lycopersicum esculentum*, Mill.) foi a Binomial Generalizada, ou o que é o mesmo, no caso, a Distribuição Binomial Negativa.
- 2º) Das unidades amostrais comparadas, optou-se pelos tipos agrupados (12 ou 17), sendo a mais simples a de Tipo 12, que "a posteriori" foi aplicada nos métodos de amostragem.

- 39) Dos três métodos de amostragem comparados, apresentou-se como o melhor, para estes casos, a Amostragem Estratificada, com um levantamento das amostras, caminhando-se em diagonal dentro de cada estrato, e evitando-se passar mais de uma vez pelo mesmo local, para não amostrar a mesma planta repetidamente.
- 40) As recomendações oriundas do presente trabalho não podem ser tomadas como definitivas. Torna-se necessária a continuidade da pesquisa em outros locais e em outros estágios de desenvolvimento da praga e da cultura.

6. BIBLIOGRAFIA

ASCOMBE, F.J., 1949. The statistical analysis of insects counts based on the Negative Binomial Distributions. Biometrics. 5: 165-173.

ASCOMBE, F.J., 1950. Sampling theory of the Negative Binomial and Logarithmic Serie Distributions. Biometrics. 37: 352-382.

BARBOSA, J.C., 1978. Modelos probabilísticos para distribuições de lagartas de *Spodoptera frugiperda* (J.E. Smith, 1797) na cultura de milho. Jaboticabal. UNESP, 56 pp. [Dissertação de Mestrado].

BEALL, G., 1939. Methods of estimating the population of insects in a field. Biometrics, 30: 422-439.

- BEALL, G., 1940. The fit and significance of contagious distributions when applied to observations on larval insects. Ecology, 21: 460-474.
- BENAVENT, J.M., 1981. "La Polilla del Tomate" - Su expansión, gravedad y métodos de control. Rev. Panorama Agropecuario, nºs 8, 9, 10, 12 y 28. [I.N.T.A.].
- CALLIL, F., 1980. Estudo da distribuição frequencial de quatro pragas na cultura da soja (*Glycine max*, M.) coletados no Município de Goiânia - Goiás. Goiânia, UFG/IPT, 101 pp. [Dissertação de Mestrado].
- CAMPOS, H., 1983. Estatística Experimental Não-Paramétrica, 4ª ed. Piracicaba, SP, ESALQ/USP. 349 pp.
- CAMPOS GALVEZ, R., 1977. Control químico del complejo "Gusano Ejército y el Minador de hojas y frutos" (*Scrobipalpus absoluta*, Meyrick) en el cultivo del tomate. Estación Experimental Agrícola de Cañete, Circular nº 1-77, DEF-EEAON, Peru.
- COLE, L.C., 1946. A theory for analyzing contagiously distributed populations. Ecology, 27: 329-341.
- COCHRAN, W.G., 1965. Técnicas de Amostragem. Ed.Fundo de Cultura, São Paulo, 555 pp.
- CUCCHI, N.J.A., 1980. Propuesta de un método de control para "La Polilla del Tomate". Estación Experimental Regional Agropecuaria, Mendoza, República Argentina. [I.N.T.A.].

- ELLIOT, J.M., 1977. Some Methods for the Statistical Analysis of Samples of Benthic Invertebrates. 2ª ed., Freshwater Biological Association.
- ENGELHARDT, M., 1975. On simple estimation of the parameters of the Weibull or Extreme-Value Distribution. Technometrics, Vol. 17, nº 3.
- ESTEFANEL, V., 1977. A amostragem seqüencial baseada no teste seqüencial da razão de probabilidade e seu uso no controle das lagartas da soja, no Estado do Rio Grande do Sul. Piracicaba, SP, ESALQ/USP. 117 pp. [Dissertação de Mestrado].
- EVANS, D.A., 1953. Experimental evidence concerning contagious distributions in Ecology. Biometrics, 40: 186-211.
- GREENWOOD, M. e G.U. YULE, 1920. An enquiry into the nature of frequency distributions representative to multiple happenings with particular reference repeated accidents. J. Roy. Stat. Soc., 83: 255-279.
- GUENTHER, W.C., 1969. Introducción a la Inferencia Estadística. Ed. "Abril", Buenos Aires. p. 42-49.
- HARTENSTEIN, R., 1961. On the distributions of forest soil micro-arthropods and their fit to "contagious" distributions functions. Ecology. San Diego, 42: 190-194.
- JOHNSON, N.L. e S. KOTZ, 1977. Discrete Distributions. Houghton Mifflin Co., Boston. Vol. 1. 327 pp.

- KOGAN, M. e D.C. HERZOG, 1980. Sampling Methods in Soybean Entomology. Spriger-Verlag. New York. 587 pp.
- MADOW, W.G., 1951. Teoria dos Levantamentos por Amostragem. I.B.G.M. Conselho Nacional de Estatística. Brasil. 256 pp.
- NAKANO, O.; S. SILVEIRA NETO e R.A. ZUCCHI, 1981. Entomologia Econômica. Piracicaba, SP, Livroceres Ltda. 314 pp.
- NEYMAN, J., 1939. On a new class of "Contagious" distributions, applicable in Entomology and Bacteriology. Ann. Math. Statistics, 10: 35-47.
- PERECIN, D., 1978. Distribuições binomiais negativas com expoente k comum e algumas aplicações em Pedologia e Entomologia. Jaboticabal, UNESP, 129 pp. [Tese de Livre Docência].
- PIEDRABUENA, A.E., 1968. Distribuições interferidas. Estudo sobre uma Distribuição Generalizada. Piracicaba, SP, ESALQ/USP. [Dissertação de Mestrado].
- PIMENTEL GOMES, F., 1982. Curso de Estatística Experimental. 10ª ed. Piracicaba, SP, Livraria Nobel, 430 pp.
- POLYA, G., 1931. Sur quelques points de la théorie des probabilités. Annls. Inst. Henri Poincaré, 1: 117-161.
- RABINOVICH, J.E., 1980. Introducción a la Ecología de Poblaciones Animales. México, Compañía Editorial Continental, S.A., 313 pp.

- ROJAS, B.A., 1970. "La Binomial Negativa" y la estimación de intensidad de plagas en el suelo. Colegio de Pos-Graduados, E.N.A., Chapingo, México.
- SAMPFORD, M.R., 1962. An Introduction to Sampling Theory. Oliver and Boyd. Great Britain, Edimburgh, 292 pp.
- SUKHATME, P.V., 1953. Sampling Theory of Surveys with Applications. The Indian Society of Agricultural Statistics, Nueva Delhi (versión en castellano, Teoría de Encuestas por Muestro con Aplicaciones, Fondo de Cultura Económica, México, 1956).
- THOMAS, M., 1949. A generalization of Poisson's Binomial Limit for use in Ecology. Biometrics, 36: 18-25.
- YA-LUN CHOU, 1977. Análisis Estadístico. 2ª ed. México, Nova Editorial Interamericana S.A. de C.V. 808 pp.
- YAMAME, T., 1967. Elementary Sampling Theory. Prentice-Hall, Inc., Englewood Cliffs, N.J., 405 pp.

7. A P Ê N D I C E

POPULAÇÃO DE LAGARTAS DE *Scrobipalpula absoluta* (Meyrick), NUMA CULTURA DE TOMATE (*Lycopersicum esculentum*, Mill.) NA ESTAÇÃO EXPERIMENTAL "EL REMATE", JUJUY, REPÚBLICA ARGENTINA.