TAMANHO E FORMA ÓTIMOS DE PARCELAS EM
DELINEAMENTOS EXPERIMENTAIS

OLAF ANDREAS BAKKE

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: Estatística e Experimentação Agronômica.

P I R A C I C A B A

Estado de São Paulo - Brasil

Maio - 1988

Bakke, Olaf Andreas
B167t Tamanho e forma ótimos de parcelas em delineamentos
experimentais. Piracicaba, 1988.
142p.

Diss.(Mestre) - ESALQ
Bibliografia.

1. Delineamento de experimento - Parcela
2. Estatística agrícola I. Escola Superior de
Agricultura Luiz de Queiroz, Piracicaba

CDD 630.219

TAMANHO E FORMA ÓTIMOS DE PARCELAS EM
DELINEAMENTOS EXPERIMENTAIS

OLAF ANDREAS BAKKE

Aprovada em: 10.06.1988

Comissão Julgadora:

Prof. Dr. Humberto de Campos	ESALQ/USP
Prof. Dr. Frederico Pimentel Gomes	ESALQ/USP
Prof. ^a Dr. ^a Clarice Garcia Borges Demétrio	ESALQ/USP
Prof. Dr. David Ariovaldo Banzatto	FMVAJ/UNESP
Prof. Dr. Sérgio do Nascimento Kronka	FMVAJ/UNESP

Prof. Dr. HUMBERTO DE CAMPOS
Orientador

Aos meus pais,

\`a minha esposa,

e aos meus filhos,

DEDICO.

AGRADECIMENTOS

Agradeço ao meu orientador, Prof. Dr. Humberto de Campos, a todos os professores e funcionários do Departamento de Matemática e Estatística da ESALQ/USP, e aos meus colegas e amigos do curso de Mestrado e também do Curso de Doutorado, pela colaboração ao longo do curso.

Agradeço também à Universidade Federal da Paraíba (UFPb) e à CAPES, pelo salário e pela bolsa, respectivamente.

Agradeço a meus pais e a minha esposa pelo auxílio, que não vem de hoje e nem hoje se encerrará, imprescindível a todos que apresentam as virtudes e os defeitos da pessoa humana.

SUMÁRIO

	Página
RESUMO	vii
SUMMARY	xi
1. INTRODUÇÃO	1
2. REVISÃO DE LITERATURA	4
2.1. Método da Máxima Curvatura	4
2.1.1. Exemplo ilustrativo	6
2.2. Método Empírico de H. FAIRFIELD SMITH.....	11
2.2.1. Método original	11
2.2.2. Exemplo ilustrativo 1	20
2.2.3. Método modificado	24
2.2.4. Exemplo ilustrativo 2	39
2.3. Método da Máxima Curvatura Modificado	49
2.3.1. Exemplo ilustrativo	54
2.4. Método de E.J.KOCH & H.J.RIGNEY	59
2.4.1. Dados de experimentos em parcelas subdivididas	61
2.4.2. Exemplo ilustrativo 1	69
2.4.3. Dados de experimentos em reticulado quadrado ..	75
2.4.4. Exemplo ilustrativo 2	80
2.5. Método de W.H.HATHEWAY	86
2.5.1. Exemplo ilustrativo 1	89
2.5.2. Exemplo ilustrativo 2	93

	Página
2.6. Método de PIMENTEL GOMES	97
2.6.1. Tamanho ótimo de parcela, para o caso de parcelas sem bordadura	101
2.6.2. Tamanho ótimo de parcela, para o caso de parcelas com meia bordadura	102
2.6.2.1. Parcelas com meia bordadura e <u>u</u> ma linha de árvores úteis	102
2.6.2.2. Parcelas com meia bordadura e com duas linhas de árvores úteis	104
2.6.3. Tamanho ótimo de parcela, para o caso de parcelas com bordadura completa	107
2.6.3.1. Parcelas com bordadura completa e uma linha de árvores úteis ..	107
2.6.3.2. Parcelas com bordadura completa e duas linhas de árvores úteis.	108
2.6.4. Exemplo ilustrativo 1	110
2.6.5. Tamanho ótimo de parcela para o caso ge- ral, considerando genericamente o número de linhas e observações úteis por parce- las	112
2.6.5.1. Parcelas com meia bordadura ...	112
2.6.5.2. Parcelas com bordadura completa	116
2.6.6. Exemplo ilustrativo 2	120
2.6.7. O caso de ρ negativo	122
3. RESULTADOS E DISCUSSÃO	126

Página

4. CONCLUSÕES	134
REFERÊNCIAS BIBLIOGRÁFICAS	136
APÊNDICE	140

TAMANHO E FORMA ÓTIMOS DE PARCELAS EM
DELINEAMENTOS EXPERIMENTAIS

Autor: OLAF ANDREAS BAKKE

Orientador: Prof. Dr. Humberto de Campos

RESUMO

Há situações onde se constata que os artigos e publicações sobre determinado assunto apresentam embasamento teórico tão resumido que não permite, à primeira vista, uma decisão segura a respeito dos seus méritos teóricos.

No caso específico da bibliografia sobre métodos de determinação de tamanho e forma ótimos de parcelas experimentais, este problema é bastante comum.

Seis dentre os métodos mais usados foram abordados, com o objetivo de esclarecer e tornar mais acessível a teoria na qual se baseiam.

A abordagem de cada método consistiu numa descrição geral, ressaltando as vantagens e desvantagens, explicando os seus aspectos teóricos, seguida da resolução de exemplo ilustrativo, finalizando com comentários adicionais a seu respeito e comparação com os demais métodos.

O método da máxima curvatura, de origem mais remota, necessita de um ensaio em branco, sendo indicado para os casos em que a unidade básica do ensaio é algo natural, não convencionalmente arbitrariamente (FEDERER, 1955), servindo mais como auxiliar aos outros métodos.

O método de SMITH (1938), o mais usado, se baseia na relação empírica da variância (V_x) entre parcelas de tamanho x reduzida à área unitária e o tamanho (x) de parcela, através da fórmula:

$$V_x = \frac{V_1}{x^b}, \text{ onde:}$$

$$V_1 \text{ é } V_x \text{ para } x = 1$$

b = coeficiente de heterogeneidade do solo.

O estimador original para b foi melhorado, sendo preterido por aquele da metodologia proposta por HATHEWAY e WILLIAMS (1958).

Este método aproveita dados provenientes de ensaios em branco e de experimentos em parcelas subdivididas ou em reticulado quadrado (KOCH & RIGNEY, 1951).

O tamanho ótimo de parcela é obtido através da fórmula:

$$x_{\text{ótimo}} = \frac{b}{1 - b} \cdot \frac{K_1}{K_2}, \text{ onde}$$

K_1 e K_2 se originam de uma função de custos do tipo:

$$C = K_1 + K_2 x$$

O método da máxima curvatura modificado se baseia na função:

$$CV_x = \frac{a'}{x^{b'}}, \text{ onde}$$

CV_x = coeficiente de variação relativo a parcelas de tamanho x .

a' e b' são constantes a serem determinadas pelo método dos mínimos quadrados ponderado (pelos graus de liberdade associados a cada CV_x).

O ponto de máxima curvatura dessa função é determinado e dado por:

$$x_{\text{crítico}} = \left[a'^2 b'^2 (2b' + 1) / (b' + 2) \right]^{1/(2b'+2)}$$

O $x_{\text{crítico}}$ assim obtido indica o tamanho de parcela que corresponde à maior taxa de variação direcional do coeficiente de variação, servindo de informação auxiliar aos diversos outros métodos de determinação de tamanho e forma ótimos de parcelas experimentais.

O método de HATHEWAY (1961) conjuga a fórmula de COCHRAN & COX (1957), de determinação do número de repetições, e a fórmula de SMITH (1938), resultando numa terceira fórmula, de fácil uso, relacionando tamanho de parcela, número de repetições e precisão desejada, permitindo a escolha da combinação destes três fatores mais conveniente à pesquisa.

O método de PIMENTEL GOMES (1984) relaciona o coeficiente de correlação intra-classe (ρ) e a variância da média

de cada tratamento, determinando a forma da parcela que a minimiza, em função do número de linhas e observações por parcela.

Fica claro que um método complementa o outro, que a informação de cada um deles se baseia ora em custos, ora na minimização da variância da média de tratamento etc. , de modo que a aplicação simultânea de mais de um método, se couber, na determinação do tamanho ótimo da parcela experimental, é altamente recomendável, a fim de que o tamanho realmente adotado atenda, na medida do possível, aos diversos requisitos presentes nos métodos empregados.

OPTIMUM PLOT SIZE AND SHAPE IN
EXPERIMENTAL DESIGNS

Author: OLAF ANDREAS BAKKE

Adviser: Prof. Dr. Humberto de Campos

SUMMARY

There are situations where one finds out that the papers about a particular subject, display their fundamental principles so briefly that make impossible, at first view, to take a precise decision about them.

On the papers about the methods of determination of size and shape of plots this problem is usual.

Six among the most used of those methods are discussed here.

Each method approach consists of a general description, pointing out its advantages and disadvantages, explaining the theory the method is based on, followed by a step-by-step resolution of one or more exercises, finishing with comments about it and comparisons with the others.

The maximum curvature method needs a blank experiment, and is advised to the cases where the basic units

of the blank experiment are natural, do not stipulated in an arbitrary manner (FEDERER, 1955), serving more as an auxiliary tool for the other methods.

SMITH's (1938) method, the most employed of them, is based on the empirical law, between the variance (V_x) of yield per unit area for plots of x units of size, and the size (x) of plot, expressed by the formula:

$$V_x = \frac{V_1}{x^b}, \text{ where:}$$

$$V_1 = V_x \text{ for } x = 1,$$

b = coefficient of soil heterogeneity.

The b coefficient is estimated by the HATHEWAY & WILLIAMS' (1958) methodology.

SMITH's (1938) method utilizes data from blank experiments or data from split-plot or lattice designs (KOCH & RIGNEY, 1951).

SMITH's (1938) optimum plot size is given by:

$$x_{opt} = \frac{b \cdot K_1}{(1 - b) K_2}, \text{ where:}$$

K_1 and K_2 are cost factors of the cost function:

$$C = K_1 + K_2 x$$

The modified maximum curvature method is based on the equation:

$$CV_x = \frac{a'}{x^b}, \text{ where:}$$

CV_x = coefficient of variability of plots of x basic units of size.

a' and b' are constants to be determined by the weighted least square method.

The point of maximum curvature is given by:

$$x_{\text{critical}} = \left[a'^2 b'^2 (2b' + 1) / (b' + 2) \right]^{1/(2b'+2)}$$

This critical point reveals the plot size that corresponds to the greatest directional change rate at CV_x , being useful as an auxiliary information to the other methods.

HATHEWAY's (1961) method combines SMITH's (1938) and COCHRAN & COX's (1957) formulas, resulting in a third one, associating plot size, number of replication and desired precision, allowing the choice of a combination of this three factors proper for each case.

PIMENTEL GOMES' (1984) method associates the intra-class correlation coefficient and the variance $[V(\hat{m})]$ of the treatment means, determining the plot size and shape that minimize $V(\hat{m})$, in terms of the number of lines and observations per plot.

It's clear that one method complements the others, and the application of more than one of them, if possible, is recomendable, so that the plot size actually used attend the requirements of the various methods simultaneously.

1. INTRODUÇÃO

Um problema de ordem prática que surge no planejamento de experimentos de um modo geral, reside na determinação da unidade padrão de coleta de dados, sendo necessária a sua caracterização no que se refere ao tamanho, e muitas vezes à sua forma.

Em muitos casos a própria experiência e capacidade de discernimento do pesquisador levaria a decisões corretas a este respeito, porém é inegável a contribuição que métodos objetivos, baseados em regras claras e procedimentos teóricos aceitáveis, podem levar à minimização dos custos e maximização da quantidade de informação de um experimento.

De uma maneira geral, a recomendação mais frequente em relação ao tamanho e forma de parcelas experimentais indica o uso de parcelas retangulares e pequenas, em detrimento das parcelas quadradas e grandes. Esta recomendação pode ser confirmada ao se utilizarem os diversos métodos de determinação de tamanho e forma ótimos de parcelas experimentais, quando se observa, geralmente, um menor coeficiente de variação ao se fazer uso de alguns tipos de parcelas retangulares.

Para tornar a determinação do tamanho e forma ótimos de parcelas experimentais algo mais acessível e assimilável, o que nem sempre ocorre nas publicações sobre este assunto, foram abordados seis métodos de determinação de tamanho e forma ótimos de parcelas experimentais, dentre os mais usados ou de maior valor teórico.

Esta abordagem é feita ao longo do Capítulo 2, consistindo, para cada método, de uma descrição geral, ressaltando as suas vantagens e desvantagens, explicando seus aspectos teóricos, seguida da resolução de exemplo ilustrativo acompanhado de comentários adicionais a seu respeito e comparações com os demais métodos.

Os seis métodos apresentados neste trabalho são:

- a) Método da máxima curvatura
- b) Método empírico de H. Fairfield Smith
- c) Método da máxima curvatura modificado
- d) Método de E.J. Koch e H.J. Rigney
- e) Método de W.H. Hatheway
- f) Método de Frederico Pimentel Gomes.

Há vários outros métodos de determinação de tamanho e forma ótimos de parcelas experimentais, porém os seis aqui abordados cobrem um amplo e diversificado conjunto de situações práticas, que darão ao pesquisador ou responsável pelo planejamento do experimento ferramentas suficientes para a resolução do problema em questão, ora focalizando aspectos

de custos, ora se preocupando com a minimização da variância da média e precisão desejada, dentre outros fatores.

Vale ainda ressaltar que mesmo ao se fixar uma determinada espécie vegetal, tem-se que o tamanho e forma ótimos da parcela experimental podem variar de acordo com as diversas características coletadas, o local do ensaio, época de coleta etc.

Por exemplo, ao se pesquisar com milho, o tamanho ótimo para a característica altura da planta possivelmente será diferente daquele obtido para a característica produção de grãos ou outra qualquer.

Neste trabalho, a determinação do tamanho e forma ótimos de parcela experimental, considera apenas uma única característica, um único local, uma época de coleta etc., extrapolando-se a estimativa do tamanho ótimo obtido para vários locais e épocas.

2. REVISÃO DE LITERATURA

O presente capítulo inclui, além da revisão de literatura propriamente dita, alguns outros tópicos, talvez não muito comuns numa revisão usual.

A configuração aqui adotada aproveita as citações bibliográficas do capítulo na descrição e desenvolvimento teórico dos métodos, permitindo um encadeamento de idéias natural e de maior entendimento.

Ao estruturar cada método, é apresentado um exemplo ilustrativo permitindo ao leitor um melhor acompanhamento da sua metodologia.

2.1. Método da Máxima Curvatura

A origem e autoria deste método são desconhecidas. CHACIN LUGO (1977) chega a caracterizá-lo como o primeiro método usado na obtenção de tamanho ótimo de parcelas experimentais.

Explica que sua metodologia consiste em se utilizar de ensaios em branco, calculando-se então os coefi-

cientes de variação (CV) para cada tamanho (x) de parcela, sem considerar a sua forma, obtendo-se um conjunto de pontos do tipo (x, CV), que são relacionados num sistema de eixos coordenados.

Uma curva à mão livre é então traçada através das coordenadas resultantes, e o seu ponto de máxima curvatura é localizado por inspeção visual, adotando-se como tamanho ótimo o valor correspondente à abscissa do ponto de máxima curvatura (FEDERER, 1955).

FEDERER (1955) aponta dois pontos restritivos a este método:

- a) Não considera os custos envolvidos.
- b) O tamanho adotado para a unidade básica do ensaio em branco influencia no ponto de máxima curvatura.

O mesmo autor parte em sua defesa para os casos em que a escolha da unidade básica do ensaio é algo natural, não convencionalmente arbitrariamente, como por exemplo um animal, uma árvore, uma leitura ou determinação de algum aparelho, etc., desaconselhando sua aplicação nos demais casos, pois o tamanho arbitrário da unidade básica influencia no ponto de máxima curvatura.

2.1.1. Exemplo ilustrativo

GOMEZ & GOMEZ (1984) apresentam um ensaio em branco com arroz, originalmente com $18 \times 36 = 648$ unidades básicas de $1 \times 1 = 1$ metro quadrado de área, do qual aproveitando apenas as dez primeiras linhas e colunas, reduzindo-o a $10 \times 10 = 100$ unidades básicas de $1 \times 1 = 1$ metro quadrado, tem-se um exemplo numérico válido para efeito ilustrativo, apesar das unidades básicas não caracterizarem uma unidade natural, não convencionada arbitrariamente. Os dados deste ensaio em branco estão reproduzidos na Tabela 1:

Tabela 1. Produções de grãos (g/m^2), de arroz, variedade IR-8, de um ensaio em branco.

842	844	808	822	979	954	965	906	898	856
803	841	870	970	943	914	916	836	858	926
773	782	860	822	932	971	765	875	853	936
912	887	815	937	844	661	841	844	809	778
874	792	803	793	818	799	767	855	792	858
908	875	899	788	867	790	831	757	751	774
875	907	921	963	875	880	898	802	874	928
891	928	871	875	865	777	738	796	855	901
823	784	754	873	764	775	752	753	820	798
785	794	764	822	714	748	724	717	736	724

Fonte: GOMEZ & GOMEZ (1984).

Pode-se agrupar estas cem unidades básicas de quinze modos distintos:

a) Agrupar as unidades básicas uma a uma, resultando em cem parcelas de tamanho 1×1 , e então $x = 1$ unidade básica = 1 ub.

b) Dividir cada linha a cada duas colunas, formando cinquenta parcelas de tamanho 1×2 , e então $x = 2$ ub.

c) Dividir cada coluna a cada duas linhas, formando cinquenta parcelas de tamanho 2×1 , e então $x = 2$ ub.

d) Vinte e cinco parcelas de tamanho 2×2 , e então $x = 4$ ub.

e) Vinte parcelas de tamanho 1×5 , e então $x = 5$ ub.

f) Vinte parcelas de tamanho 5×1 , e então $x = 5$ ub.

g) Dez parcelas de tamanho 1×10 , e então $x = 10$ ub.

h) Dez parcelas de tamanho 10×1 , e então $x = 10$ ub.

i) Dez parcelas de tamanho 2×5 , e então $x = 10$ ub.

j) Dez parcelas de tamanho 5×2 , e então $x = 10$ ub.

l) Cinco parcelas de tamanho 2×10 , e então $x = 20$ ub.

8.

m) Cinco parcelas de tamanho 10 x 2, e então
x = 20 ub.

n) Quatro parcelas de tamanho 5 x 5, e então
x = 25 ub.

o) Duas parcelas de tamanho 5 x 10, e então
x = 50 ub.

p) Duas parcelas de tamanho 10 x 5, e então
x = 50 ub.

Observe-se que parcelas de dimensões L x C têm mesmo tamanho que parcelas de dimensões C x L, pois as unidades básicas são quadradas. Logo, apesar de se observar quinze formas de parcelas, observa-se apenas oito tamanhos diferentes, a saber x = 1, 2, 4, 5, 10, 20, 25 ou 50.

Apenas para ilustrar, seja o cálculo do coeficiente de variação referente às parcelas de dimensões 2 x 1

$$CV_{2 \times 1} = \frac{s_{y_{2 \times 1}}}{\bar{y}_{2 \times 1}} \cdot 100$$

$$= \frac{\left[\frac{1645^2 + 1685^2 + \dots + 1522^2 - \frac{(1645 + 1685 + \dots + 1522)^2}{50}}{49} \right]^{1/2}}{\frac{1645 + 1685 + \dots + 1522}{50}} \cdot 100$$

$$= 6,5724\%.$$

Analogamente para as demais formas chega-se aos resultados da Tabela 2 e à Figura 1:

Tabela 2. Dimensões (L x C) e tamanho (x) de parcelas de um ensaio em branco com arroz, e respectivos coeficientes de variação (por dimensão: $CV_{L \times C}$, e por número de unidades básicas: CV_x^1).

Dimensões L x C	Número de unidades básicas (x)	$CV_{L \times C}$ (%)	CV_x (%)
1 x 1	1	8,1942	8,1942
1 x 2	2	7,3192	6,9458
2 x 1	2	6,5724	
2 x 2	4	6,0975	6,0975
1 x 5	5	6,2676	5,0855
5 x 1	5	3,9035	
1 x 10	10	5,4334	4,2196
10 x 1	10	2,0818	
2 x 5	10	5,6767	
5 x 2	10	3,6863	
2 x 10	20	5,4599	3,5304
10 x 2	20	1,6009	
5 x 5	25	3,4435	3,4435
5 x 10	50	2,8914	2,4733
10 x 5	50	2,0552	

¹ O coeficiente de variação de parcelas de tamanho x é considerado como a média aritmética de todos os $CV_{L \times C}$ para um determinado tamanho x, facilitando o desenho da curva à mão livre. Procedimento semelhante se encontra em SMITH (1938).

Figura 1. Relação entre CV_x e x de um ensaio em branco com arroz e determinação do tamanho ótimo de parcela através de método da máxima curvatura.

Por inspeção visual determina-se o ponto de máxima curvatura, que pela Figura 1 indica um valor de x ao redor de 10,5 unidades básicas, ou seja parcelas em torno de 10,5 metros quadrados. Quanto à forma, podemos afirmar, observando a Tabela 2, que parcelas de dimensões 10 x 1 ou 5 x 2 são mais recomendadas, pois apresentam CV_{LxC} menores para este tamanho.

2.2. Método Empírico de H. FAIRFIELD SMITH

2.2.1. Método original

SMITH (1938) apresenta a equação:

$$V_x = \frac{V_1}{x^b},$$

para relacionar a variância da produção por área unitária (V_x) de parcelas de x unidades básicas de tamanho e o tamanho (x) de parcela (V_1 é a variância da produção por área unitária de parcelas de $x = 1$ unidade básica de tamanho).

O autor explica que, a princípio, o coeficiente b pode ser considerado como uma medida de correlação entre áreas adjacentes, porém acrescenta que erros nas técnicas de semeadura, colheita, debulha, pesagem, etc, além da própria variabilidade genética das plantas, tenderão a aumentar o valor estimado de b , o que, em termos teóricos, não é desejável, mas que na prática é até correto que sejam incluídos no valor de b , e assim o coeficiente b mede o grau de heterogeneidade do solo e das plantas (entre outros fatores) de um determinado ensaio.

Quanto maior for o valor de b maior a heterogeneidade do solo, onde se instalou o ensaio em branco, e das plantas, das quais foram obtidos os dados de produção.

Uma vez que as variâncias são estimadas de números variáveis de parcelas, o que implica em diversos números de graus de liberdade para as diversas variâncias, é aconselhável que ao se ajustar uma regressão, cada ponto seja ponderado pelo número de graus de liberdade a ele associado (SMITH, 1938).

O mesmo autor lembra que os métodos estatísticos de ajustamento de curvas se baseiam na hipótese de independência entre as observações. É visível que esta condição não é satisfeita nos ensaios em branco, pois as V_x são obtidas a partir de um mesmo conjunto de dados, o que leva à dependência das mesmas, porém, explica o autor, tudo o que se quer é uma medida de uma relação empírica, e testes de significância são de importância secundária, descartando a necessidade de um ajuste no método de estimação de b pela não independência das observações.

FEDERER (1955) apresenta a fórmula de estimação de b , obtida pelo uso do método dos mínimos quadrados ponderado [pelo número de graus de liberdade (w_i) associado a cada ponto], após a linearização da equação de SMITH (1938):

$$a) V_{x_i} = \frac{V_1}{x_i^b}$$

$$b) \log V_{x_i} = \log V_1 - b \log x_i$$

c) $y_i' = b_0 + b_1 x_i'$, resultando no seguinte modelo de regressão:

$$y_i' = b_0 + b_1 x_i' + e_i'$$

de onde se obtêm as estimativas de b_0 e b_1 dadas por:

$$\hat{b}_1 = \frac{\Sigma w_i x_i' y_i' - \frac{\Sigma w_i x_i' \Sigma w_i y_i'}{\Sigma w_i}}{\Sigma w_i x_i'^2 - \frac{[\Sigma w_i x_i']^2}{\Sigma w_i}} = -\hat{b}$$

$$\hat{b}_0 = \frac{\Sigma w_i y_i'}{\Sigma w_i} - \hat{b}_1 \frac{\Sigma w_i x_i'}{\Sigma w_i} = \bar{y}'_{\text{pond}} - \hat{b}_1 \bar{x}'_{\text{pond}} = \widehat{\log V_1}$$

HATHEWAY (1961), ao desenvolver o seu próprio método de determinação de tamanho ótimo de parcelas, tece alguns comentários a respeito da lei empírica de SMITH (1938), esclarecendo, entre outras coisas, como se calcular V_x da fórmula de SMITH (1938).

Seja S_x^2 a variância entre parcelas de x unidades básicas de tamanho. Reduzindo-a a uma base unitária, tem-se:

$$V_x' = \frac{S_x^2}{x}$$

Reduzindo-a para área unitária, vem:

$$V_x = \frac{V_x'}{x} = \frac{S_x^2 \div x}{x} = \frac{S_x^2}{x^2},$$

no que concordam diversos outros autores (LESSMAN & ATKINS, 1963; AFONJA, 1967 e 1968).

SILVA (1972) deduz o tamanho ótimo de parcela deste método de uma maneira bem simples:

Sabe-se que:

$$V_x = \frac{V_1}{x^b}$$

Considerando por exemplo, parcelas com dois tamanhos distintos (x_1 e x_2) com as respectivas variâncias $S_{x_1}^2$ e $S_{x_2}^2$.

Sejam r_1 e r_2 os números de repetições referentes aos tamanhos x_1 e x_2 , de modo que façam ocupar, nos dois conjuntos de parcelas, a mesma área, tal que:

$$r_1 x_1 = k = r_2 x_2 ,$$

onde k é uma constante (área).

Logo:

$$r_1 = \frac{k}{x_1} \quad e \quad r_2 = \frac{k}{x_2}$$

As variâncias das produções totais nas duas áreas iguais são:

$$r_1 S_{x_1}^2 = k \frac{S_{x_1}^2}{x_1} \quad e \quad r_2 S_{x_2}^2 = k \frac{S_{x_2}^2}{x_2}$$

Mas:

$$V_x = \frac{V_1}{x^b} ,$$

$$e: \quad V_x = \frac{S_x^2}{x^b} \Rightarrow V_{x_1} = \frac{S_{x_1}^2}{x_1^2} \quad e \quad V_{x_2} = \frac{S_{x_2}^2}{x_2^2} ,$$

então:
$$\frac{S_{x_1}^2}{x_1^2} = \frac{V_1}{x_1^b} \quad e \quad \frac{S_{x_2}^2}{x_2^2} = \frac{V_1}{x_2^b} .$$

Das igualdades anteriores, obtém-se:

$$S_{x_1}^2 = V_1 x_1^{2-b} \quad e \quad S_{x_2}^2 = V_1 x_2^{2-b}$$

Logo, as variâncias das produções totais, em áreas iguais, são:

$$r_1 S_{x_1}^2 = \frac{k}{x_1} \cdot V_1 x_1^{2-b} = k V_1 x_1^{1-b}$$

$$r_2 S_{x_2}^2 = \frac{k}{x_2} \cdot V_1 x_2^{2-b} = k V_1 x_2^{1-b}$$

As invariâncias (informação) respectivas, são:

$$\frac{1}{k V_1 x_1^{1-b}} = \frac{x_1^{b-1}}{k V_1} = I_{x_1} \quad e \quad \frac{1}{k V_1 x_2^{1-b}} = \frac{x_2^{b-1}}{k V_1} = I_{x_2}$$

Mas os custos por parcela de tamanho x são dados por (SMITH, 1938):

$$C_x = K_1 + K_2 x, \quad \text{onde}$$

K_1 = custo da unidade experimental, independente do número de indivíduos na unidade experimental, ou seja custos fixos por parcela (DAGNELIE, 1981).

K_2 = custo de um item adicional na parcela experimental, independente do custo das unidades experimentais, ou seja custos variáveis por parcela (DAGNELIE, 1981),

e em r_1 e r_2 repetições, resulta:

$$r_1 (K_1 + K_2 x_1) = \frac{k}{x_1} (K_1 + K_2 x) = k (K_1 x_1^{-1} + K_2) = C_{x1}$$

$$r_2 (K_1 + K_2 x_2) = k (K_1 x_2^{-1} + K_2) = C_{x2}$$

O custo por unidade de informação $\left| \frac{C_x}{I_x} \right|$ é obtido dividindo o custo total pela invariância (informação):

$$\frac{C_{x1}}{I_{x1}} = k^2 V_1 \left[K_1 x_1^{-b} + K_2 x_1^{1-b} \right]$$

$$\frac{C_{x2}}{I_{x2}} = k^2 V_1 \left[K_1 x_2^{-b} + K_2 x_2^{1-b} \right]$$

donde se conclui que o custo por unidade de informação, em igual área, com parcelas de x unidades básicas, é proporcional a:

$$y = K_1 x^{-b} + K_2 x^{1-b}$$

Derivando em relação a x e anulando esta derivada:

$$\frac{\partial y}{\partial x} = K_1 (-b) x^{-b-1} + K_2 (1-b) x^{-b} = 0$$

$$K_2 (1-b) x^{-b} = b K_1 x^{-b-1}$$

$$x = \frac{b K_1}{(1-b) K_2}$$

Logo, o x anteriormente obtido é considerado como o tamanho ótimo, pois maximiza a quantidade de informação por unidade de custo, e então:

$$\frac{bK_1}{(1-b)K_2} = x_{\text{ótimo}}$$

Segundo SMITH (1938), esta fórmula estima o tamanho de parcela mais eficiente em termos de custos por unidade de informação, para qualquer experimento, não sendo afeta-do pelo número de repetições. Este aspecto deve ser determi-nado em função da área disponível ou precisão desejada.

Procedimentos semelhantes são utilizados para se determinar o tamanho ótimo de parcelas ao se considerar os custos de bordadura, resultando numa função de custos do ti-po:

$$C = K_1 + K_2x + K_g (A + Bx) ,$$

e num $x_{\text{ótimo}}$ dado por:

$$x_{\text{ótimo}} = \frac{b(K_1 + K_g A)}{(1-b)(K_2 + K_g B)} , \quad \text{onde:}$$

K_g = custo por unidade de área para manusear a
bordadura

A = área de bordadura

B = razão entre o material de bordadura e o material de prova (área útil).

Na prática, porém, tem sido considerado difí-

cil avaliar devidamente esses custos, o que prejudica o uso deste método (PIMENTEL GOMES, 1984).

Dificuldades à parte, o importante é que admitindo que a função de custos seja da forma apresentada por SMITH (1938), o que é razoável supor, tem-se que os dois valores de x ótimos que ele determina serão tanto mais precisos quanto mais precisa for essa função de custos.

MARANI (1963) alerta para o uso indevido dos fatores K_1 e K_2 , na medida em que são expressos em porcentagem do custo total, pois então a relação K_1/K_2 , onde K_1 e K_2 são porcentagens do custo total ($K_1 + K_2 = 1$), na verdade é proporcional a $K_1/K_2 A$, onde A é a área da unidade básica, e o seu uso na obtenção de $x_{\text{ótimo}}$ é incorreto.

A estimativa de b , dada por:

$$\hat{b} = -\hat{b}_1 = - \frac{\sum w_i x_i' y_i' - \frac{\sum w_i x_i' \sum w_i y_i'}{\sum w_i}}{\sum w_i x_i'^2 - \frac{[\sum w_i x_i']^2}{\sum w_i}},$$

deve ser corrigida para o que SMITH (1938) chamou de b equivalente, em função da área finita do ensaio em branco. Esta correção leva em conta a relação entre a área de uma unidade básica e a área total do ensaio (x/n).

A maneira mais fácil de se proceder a essa correção é através da Tabela 3, adaptada de SMITH (1938).

Tabela 3. Valores estimados de $\hat{b}$ e seus respectivos valores equivalentes, considerando a relação entre a área de uma unidade básica e a área total de ensaio (x/n).

$\hat{b}$	1,0	0,800	0,700	0,600	0,500	0,400	0,350	0,300	0,250	0,200	0,150	0,100
$b \left\{ \begin{array}{l} 0,001 < \frac{x}{n} < 0,01 \\ 0,01 < \frac{x}{n} < 0,1 \end{array} \right.$	1,0	0,804	0,710	0,617	0,528	0,443	0,403	0,364	0,326	0,291	0,257	0,226
equivalente	1,0	0,822	0,738	0,656	0,578	0,504	0,469	0,434	0,402	0,371	0,343	0,312

Fonte: SMITH (1938).

Por exemplo, para $\hat{b} = 0,80$ e $\frac{x}{n} = 0,002$, temos: $\hat{b}_{\text{equiv.}} = 0,804$, e este valor é o que deve ser utilizado na fórmula do $x_{\text{ótimo}}$:

$$x_{\text{ótimo}} = \frac{b}{(1-b)} \cdot \frac{K_1}{K_2} \quad \text{ou} \quad x_{\text{ótimo}} = \frac{b}{(1-b)} \cdot \frac{(K_1 + K_g A)}{(K_2 + K_g B)}$$

2.2.2. Exemplo ilustrativo 1

Seja então o exemplo anterior, apresentado em 2.1.1..

Apenas para ilustrar, seja a variância da produção por área unitária de parcelas de tamanho $x = 2$, e forma $L \times C = 2 \times 1$, dada por:

$$V_{2 \times 1} = \frac{s_{2 \times 1}^2}{2^2} = \frac{1645^2 + 1685^2 + \dots + 1522^2 - \frac{(83.883)^2}{50}}{4} = 3.039,476$$

Procedendo de maneira análoga para os diversos outros tamanhos e formas, obtemos a Tabela 4:

Tabela 4. Tamanho (x) de parcela, em termos de unidades básicas, variância da produção por área unitária (V_x) de parcela de tamanho x, número de graus de liberdade (w) associados às V_x e transformação logarítmica de x e V_x , de dados provenientes de um ensaio em branco com arroz, adaptado de GOMEZ e GOMEZ (1984).

L x C	x	V_x	w	$x' = \log_e x$	$y' = \log_e V_x$
1 x 1	1	4.724,526	99	0,000000	8,460523
1 x 2	2	3.769,435	49	0,693147	8,234681
2 x 1	2	3.039,476	49	0,693147	8,019441
2 x 2	4	2.616,118	24	1,386295	7,869446
1 x 5	5	2.764,102	19	1,609438	7,924470
5 x 1	5	1.072,157	19	1,609438	6,977427
1 x 10	10	2.077,245	9	2,302585	7,638797
10 x 1	10	304,936	9	2,302585	5,720102
2 x 5	10	2.267,429	9	2,302585	7,726401
5 x 2	10	956,136	9	2,302585	6,862900
2 x 10	20	2.097,581	4	2,995732	7,648539
10 x 2	20	180,332	4	2,995732	5,194799
5 x 5	25	834,373	3	3,218875	6,726681
5 x 10	50	588,245	1	3,912023	6,377144
10 x 5	50	297,192	1	3,912023	5,694377

Da Tabela 4 obtemos:

$$\sum w_i x_i' y_i' = 2.114,833027$$

$$\sum w_i x_i' = 286,697717$$

$$\sum w_i y_i' = 2.441,205218$$

$$\sum w_i x_i'^2 = 515,993863$$

$$\sum w_i = 308$$

$$\hat{b} = -\hat{b}_1 = - \frac{\sum w_i x_i' y_i' - \frac{\sum w_i x_i' \sum w_i y_i'}{\sum w_i}}{\sum w_i x_i'^2 - \frac{[\sum w_i x_i']^2}{\sum w_i}} = 0,6323$$

Aplicando a correção para população finita, através da Tabela 3, temos, por interpolação linear simples, segundo GOMEZ & GOMEZ (1984):

$$\hat{b}_{\text{equiv.}} = 0,617 + \frac{(0,0323) (0,710 - 0,617)}{0,1} \approx 0,647$$

Usando uma relação de custos K_1/K_2 igual a 4, ou seja, $K_1/K_2 = 4$, e substituindo em:

$$x_{\text{ótimo}} = \frac{\hat{b}}{1 - \hat{b}} \cdot \frac{K_1}{K_2}, \text{ vem:}$$

$$x_{\text{ótimo}} = \frac{0,647}{1 - 0,647} \cdot 4 = 7,33 \text{ unidades básicas}$$

Assim, o tamanho de parcela que minimiza os custos por unidade de informação, ou em outras palavras, o tamanho de parcela que maximiza a quantidade de informação por

unidade de custo é $x = 7,33$ unidades básicas, o que em termos de área equivale a parcelas de $7,33 \text{ m}^2$.

De acordo com as Tabelas 2 e 4, percebe-se que é desejável usar a maior dimensão da parcela no sentido das colunas do ensaio em branco ($L > C$), pois para um dado tamanho, o CV_x e a V_x são menores quando a parcela apresenta maior dimensão no sentido das colunas do ensaio.

Note-se que o método de SMITH (1938) não determina a forma ótima de parcela, porém o artifício lembrado no parágrafo anterior é totalmente válido, e deste modo o pesquisador deverá tender a escolher parcelas de dimensões $7,0 \text{ m} \times 1,0 \text{ m}$; $6,0 \text{ m} \times 1,2 \text{ m}$; $5,0 \text{ m} \times 1,5 \text{ m}$; $4,0 \text{ m} \times 1,8 \text{ m}$ ou semelhantes, segundo as conveniências práticas da pesquisa.

Há casos em que a estimativa de b resulta superior à unidade, o que não é o esperado. O uso dessa estimativa para a obtenção do tamanho ótimo da parcela, através da fórmula de $x_{\text{ótimo}}$, leva a valores negativos. AFONJA (1968) afirma que essa possibilidade (estimativa de b superior à unidade) é freqüentemente considerada inadmissível, pois implica em correlação negativa entre as unidades básicas, embora o método de obtenção da estimativa de b e do coeficiente de correlação (r) admita valor superior à unidade e negativo, respectivamente, finalizando que, teoricamente, $0 < b < +\infty$ e $-1 \leq r \leq +1$.

FEDERER (1955) comenta que o coeficiente b as

sume valores entre 0 e 1, a menos que ocorra competição entre as parcelas, e que fora este caso, em trabalhos de campo, existe provavelmente correlação entre parcelas adjacentes, resultando em estimativas de b menores do que a unidade.

BRIM & MASON (1959) afirmam que quando $b = 1$, o tamanho ótimo de parcela independe dos custos, e nos casos em que o solo for extremamente heterogêneo, o tamanho da parcela é determinado apenas pelo número de repetições por tratamento que se pode e que se deseja colocar na área disponível para o experimento.

Os mesmos autores acrescentam que para $b = 0$, a área de ensaio é completamente uniforme em termos de fertilidade, e para $b = 1$, a área de ensaio apresenta um padrão de fertilidade do solo totalmente aleatório (manchas de fertilidade) ou de extrema heterogeneidade (acentuado gradiente de fertilidade)

2.2.3. Método modificado

HATHEWAY & WILLIAMS (1958), apesar de reconhecerem que, talvez, a mais útil medida da heterogeneidade do solo já desenvolvida seja aquela devida a SMITH (1938), consideram insuficiente, para a obtenção da sua estimativa (estimativa de b), a ponderação pelo número de graus de liberdade (w) associado a cada V_x , por ele proposta.

Mostram uma ponderação que leva a uma estimativa não tendenciosa e de variância mínima assintótica para b , sendo aplicável tanto a dados oriundos de ensaios em branco, como a dados experimentais comuns, como sugerido por KOCH & RIGNEY (1951), e que se baseia nos elementos da matriz de variâncias e covariâncias das n observações do tipo $y_i' = \log_a V_{x_i}$ ($\forall a > 0$ e $a \neq 1$).

ROSSETTI (1979) estuda o estimador proposto por HATHEWAY & WILLIAMS (1951) de uma maneira bastante elegante e objetiva. Segundo esse autor, ressaltando os resultados das suas deduções, temos:

"Linearizando a equação de SMITH (1938) resulta:

$$\log V_{x_i} = \log V_1 - b \log x_i$$

$$\text{Seja } \log V_{x_i} = y_i',$$

$$\log V_1 = a,$$

$$\log x_i = x_i' \quad \text{e}$$

$$-b = B,$$

temos o seguinte modelo de regressão linear:

$$y_i' = a + Bx_i' + e_i,$$

onde os y_i' são correlacionados e apresentam variâncias desiguais.

Assim, para os n valores observados, a forma matricial será:

$$\underline{Y} = \underline{X}\underline{\beta} + \underline{\varepsilon} ,$$

onde:

$$E(\underline{Y}) = \underline{X}\underline{\beta} ,$$

$$V(\underline{Y}) = \begin{bmatrix} V(y_1') & \text{COV}(y_1', y_2') & \text{COV}(y_1', y_3') & \dots & \text{COV}(y_1', y_n') \\ \text{COV}(y_2', y_1') & V(y_2') & \text{COV}(y_2', y_3') & \dots & \text{COV}(y_2', y_n') \\ \text{COV}(y_3', y_1') & \text{COV}(y_3', y_2') & V(y_3') & \dots & \text{COV}(y_3', y_n') \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \text{COV}(y_n', y_1') & \text{COV}(y_n', y_2') & \text{COV}(y_n', y_3') & \dots & V(y_n') \end{bmatrix}$$

$$V(\underline{Y}) = W = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1n} \\ w_{21} & w_{22} & \dots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \dots & w_{nn} \end{bmatrix} = \text{matriz de dispersão dos } y_i' ,$$

onde W é simétrica e positiva definida, e $w_{ij} = c_{ij}\sigma^2$, onde c_{ij} é uma constante associada ao elemento w_{ij} , e σ^2 é a variância da parcela base (unidade básica).

Assim, supondo que o valor $\underline{Y}$ seja normalmente distribuído, tem-se:

$$\underline{Y} \sim N(\underline{X}\underline{\beta}, W)$$

Para minimizar a soma dos quadrados dos erros, no sistema linear $\underline{Y} = \underline{X}\underline{\beta} + \underline{\varepsilon}$, através do método dos míni-

mos quadrados generalizados, procedendo da maneira indicada por Rao (1973)*, ROSSETTI (1979) demonstra a não tendenciosidade do vetor $\hat{\beta}$, caracterizando $\hat{a}$ e $\hat{b}$ da maneira a seguir:

$$\hat{a} = \widehat{\log V_1} = \bar{y}' - \hat{b}\bar{x}'$$

$$\hat{b} = -B = - \frac{\sum_{ij} w^{ij} y_i' (x_j' - \bar{x}')}{\sum_{ij} w^{ij} x_i' (x_j' - \bar{x}')} = \frac{-U}{T},$$

onde:

w^{ij} = são os elementos da inversa da matriz de dispersão (W) das observações y_i' , são estimativas das verdadeiros pesos populacionais, e por isto mesmo aproximados. HATHEWAY & WILLIAMS (1958) consideram desprezíveis os erros dessa aproximação e que a estimativa de b obtida com esses pesos continua não tendenciosa e de variância mínima (assintótica).

$\bar{y}'$ = média ponderada das observações y_i'

$$= \frac{\sum_{ij} w^{ij} y_i'}{\sum_{ij} w^{ij}}$$

$\bar{x}'$ = média ponderada da variável independente x'

$$= \frac{\sum_{ij} w^{ij} x_j'}{\sum_{ij} w^{ij}}$$

* RAO, C.R. Linear Statistical Inference and its applications. 2. ed, New York, John Wiley. 1973. 625 p. (Wiley series in probability and mathematical statistics).

HATHEWAY & WILLIAMS (1958) propõem desconsiderar a correção para população finita, da estimativa de b , pois "as estimativas de variância que esta correção mais afeta são aquelas menos acuradamente estimadas...".

ROSSETTI (1979) determina a variância do vetor $\hat{\beta}$, demonstra que é mínima (assintótica) dentre todos os estimadores lineares não tendenciosos para β , explicitando a estimativa da variância de $\hat{b}$, dada por:

$$\hat{V}(\hat{b}) = \frac{1}{\sum_{ij} w^{ij} x_i' (x_j' - \bar{x}')} = \frac{1}{T}$$

Freqüentemente, existem mais de dois pontos ($n > 2$) dos quais se estima a regressão linear, e neste caso o desvio da regressão pode ser testado de uma maneira bem simples.

Numa análise de variância para esta regressão, a Soma de Quadrados Total (SQTotal), a Soma de Quadrados devidos a $\hat{b}$ (SQReg) e a Soma de Quadrados dos Desvios da Regressão (SQDR) são dados, respectivamente, segundo ROSSETTI (1979), por:

$$SQTotal = \sum_{ij} w^{ij} y_i' (y_j' - \bar{y}') = V,$$

com $n - 1$ graus de liberdade,

$$SQReg = \frac{\left[\sum_{ij} w^{ij} y_i' (x_j' - \bar{x}') \right]^2}{\sum_{ij} w^{ij} x_i' (x_j' - \bar{x}')} = \frac{U^2}{T},$$

com 1 grau de liberdade, e

$$SQDR = SQ_{Total} - SQ_{Reg} = V - \frac{U^2}{T}$$

ROSSETTI (1979) demonstra que a SQDR assim caracterizada, apresenta distribuição de qui-quadrado, com $n - 2$ graus de liberdade, ou seja:

$$SQDR \sim \chi^2_{n-2},$$

e desta forma o desvio da regressão pode ser testado.

É possível subdividir um ensaio em branco de modo a simular um experimento no delineamento em parcelas subdivididas ou em *lattice* (reticulado quadrado), o qual poderia ser analisado da seguinte maneira (KOCH & RIGNEY, 1951):

FONTE DE VARIAÇÃO	GL	QM	E [QM]
Repetições	$d - 1$	V_1	$S + aP + abB + abcR$
Blocos dentro de repetições	$d (c - 1)$	V_2	$S + aP + abB$
Parcelas dentro de blocos	$cd (b - 1)$	V_3	$S + aP$
Subparcelas dentro de parcelas	$bcd (a - 1)$	V_4	S

Nas condições do quadro de análise anterior, temos quatro tamanhos de parcelas:

- a) Parcelas do tamanho das repetições
- b) Parcelas do tamanho dos blocos

c) Tamanho das parcelas comuns

d) Parcelas do tamanho das subparcelas.

A variância da produção entre parcelas dos diversos tamanhos deve ser estimada, de modo a reduzi-las à base unitária.

Relacionando os Quadrados Médios da análise de variância (V_1 , V_2 , V_3 e V_4) com as variâncias da produção, reduzidas a uma base unitária, de acordo com 2.2.1., e considerando a subparcela como a unidade básica, temos:

a) $V'_{x=abc} = V_1$, pois cada repetição tem abc subparcelas

$$b) V'_{x=ab} = \frac{\text{SQ entre blocos}}{cd - 1} = \frac{(d-1)V_1 + d(c-1)V_2}{(d-1) + d(c-1)}$$

$$= \frac{(d-1)[\hat{S} + a\hat{P} + ab\hat{B} + abc\hat{R}] + d(c-1)[\hat{S} + a\hat{P} + ab\hat{B}]}{cd - 1}$$

$$V'_{x=ab} = \hat{S} + a\hat{P} + ab\hat{B} + \frac{(d-1)}{cd - 1} abc\hat{R}$$

$$c) V'_{x=a} = \frac{\text{SQ entre parcelas}}{bcd - 1} =$$

$$= \frac{(d-1)V_1 + d(c-1)V_2 + cd(b-1)V_3}{(d-1) + d(c-1) + cd(b-1)}$$

$$= \frac{(d-1)[\hat{S} + a\hat{P} + ab\hat{B} + abc\hat{R}] + d(c-1)[\hat{S} + a\hat{P} + ab\hat{B}] + dc(b-1)[\hat{S} + a\hat{P}]}{bcd - 1}$$

$$= \hat{S} + a\hat{P} + \frac{(cd-1)ab\hat{B}}{bcd - 1} + \frac{(d-1)}{bcd - 1} abc\hat{R}, \quad e$$

$$\begin{aligned}
 d) \quad V'_{x=1} &= \frac{\text{SQ entre subparcelas}}{abcd - 1} = \\
 &= \frac{(d-1)V_1 + d(c-1)V_2 + cd(b-1)V_3 + bcd(a-1)V_4}{(d-1) + d(c-1) + cd(b-1) + bcd(a-1)} = \\
 &= \frac{(d-1)[\hat{S} + a\hat{P} + ab\hat{B} + abc\hat{R}] + d(c-1)[\hat{S} + a\hat{P} + ab\hat{B}] + cd(b-1)[\hat{S} + a\hat{P}] + bcd(a-1)\hat{S}}{abcd - 1} \\
 &= \hat{S} + \frac{(bcd-1)a\hat{P}}{abcd-1} + \frac{(cd-1)ab\hat{B}}{abcd-1} + \frac{(d-1)abc\hat{R}}{abcd-1}
 \end{aligned}$$

A partir das V' obtêm-se as variâncias da produção por área unitária, e então:

$$a) \quad V_{x=abc} = \frac{V'_{x=abc}}{abc},$$

$$b) \quad V_{x=ab} = \frac{V'_{x=ab}}{ab},$$

$$c) \quad V_{x=a} = \frac{V'_{x=a}}{a} \quad e$$

$$d) \quad V_{x=1} = V'_{x=1}$$

Apenas para esclarecimento dos procedimentos anteriores, seja o exemplo ilustrativo a seguir, a nível de repetições e blocos:

Repetição 1					Repetição 2				
Bloco 1	2	3	4	9	Bloco 3	4	6	4	14
Bloco 2	2	5	3	10	Bloco 4	4	3	5	12

FONTE DE VARIAÇÃO	GL	SQ	QM
Repetições	$2 - 1 = 1$	4,0833...	$4,0833... = V_1$
Blocos dentro de repetições	$2(2 - 1) = 2$	0,833...	$0,4166... = V_2$
Blocos	$4 - 1 = 3$	4,9166...	

$$SQ_{\text{Repetições}} = \frac{1}{6} [19^2 + 26^2] - \frac{45^2}{12} = 4,0833 = V_1,$$

$$e \quad V'_{x=abc} = V_1 = 4,0833...$$

$$SQ_{\text{Blocos dentro de repetições}} = \frac{1}{3} [9^2 + 10^2] - \frac{19^2}{6} + \frac{1}{3} [14^2 + 12^2] - \frac{26^2}{6}$$

$$= 0,833... \Rightarrow V_2 = \frac{0,833...}{2} = 0,4166...$$

$$SQ_{\text{Blocos}} = \frac{1}{3} [9^2 + 10^2 + 14^2 + 12^2] - \frac{45^2}{12} = 4,9166...$$

$$= SQ_{\text{Repetições}} + SQ_{\text{Blocos dentro de repetições}}.$$

$V'_{x=ab}$ = Variância estimada entre blocos, reduzida a uma base unitária

$$= \frac{SQ_{\text{Bloco}}}{(2-1) + 2(2-1)} = \frac{(2-1)V_1 + 2(2-1)V_2}{3} = \frac{4,9166...}{3} = 1,6388...$$

Da teoria matricial, envolvendo as formas quadráticas V_1 , V_2 , V_3 e V_4 do quadro de análise, temos que as mesmas são independentes.

Sejam f , SQ e QM o número de graus de liberdade, soma de quadrados e quadrado médio, respectivamente, tal que:

$$QM = \frac{SQ}{f},$$

obtidos de uma análise de variância de dados balanceados, onde os erros têm distribuição normal de média zero e variância σ^2 , todos os outros i fatores aleatórios do modelo também têm distribuição normal de média zero e variância σ_i^2 , e todos os fatores aleatórios são independentes entre si.

Da caracterização acima, resulta, segundo SEARLE (1971):

$$\frac{SQ}{E(QM)} \sim \chi_f^2 \quad \text{ou} \quad \frac{fQM}{E(QM)} \sim \chi_f^2$$

Logo:

$$a) \frac{(d-1)V_1}{E(V_1)} \sim \chi_{d-1}^2$$

$$b) \frac{d(c-1)V_2}{E(V_2)} \sim \chi_{d(c-1)}^2$$

$$c) \frac{cd(b-1)V_3}{E(V_3)} \sim \chi_{cd(b-1)}^2$$

$$d) \frac{bcd(a-1)V_4}{E(V_4)} \sim \chi_{bcd(a-1)}^2$$

Mas a variância de uma variável aleatória X com distribuição de χ_f^2 é (SEARLE, 1971):

$$V(X) = 2 \cdot f,$$

assim:

$$a) \quad V \left[\frac{(d-1)V_1}{E(V_1)} \right] = 2(d-1)$$

$$\frac{(d-1)^2}{[E(V_1)]^2} V(V_1) = 2(d-1)$$

$$V(V_1) = \frac{2[E(V_1)]^2}{d-1} \Rightarrow \hat{V}(V_1) = \frac{2V_1^2}{d-1}$$

Analogamente para V_2 , V_3 e V_4 , obtêm-se:

$$b) \quad V(V_2) = \frac{2[E(V_2)]^2}{d(c-1)} \Rightarrow \hat{V}(V_2) = \frac{2V_2^2}{d(c-1)}$$

$$c) \quad V(V_3) = \frac{2[E(V_3)]^2}{cd(b-1)} \Rightarrow \hat{V}(V_3) = \frac{2V_3^2}{cd(b-1)}$$

$$d) \quad V(V_4) = \frac{2[E(V_4)]^2}{bcd(a-1)} \Rightarrow \hat{V}(V_4) = \frac{2V_4^2}{bcd(a-1)}$$

SEARLE (1971) comenta que os estimadores:

$$\hat{V}(V_i) = \frac{2V_i^2}{f_i},$$

não apresentam propriedades desejadas conhecidas, não sendo nem ao menos não tendenciosos, ou seja:

$$E[\hat{V}(V_i)] \neq V(V_i).$$

O autor apresenta um estimador de $V(V_i)$ não tendencioso $[\hat{V}^*(V_i)]$, dado por:

$$\hat{V}^*(V_i) = \frac{2V_i^2}{f_i + 2},$$

porém HATHEWAY & WILLIAMS (1958) não consideram este estimador, preferindo aquele dado por:

$$\hat{V}(V_i) = \frac{2V_i^2}{f_i}.$$

Os erros médios quadráticos de $\hat{V}(V_i)$ e $\hat{V}^*(V_i)$ não são comentados por SEARLE (1971), de modo que uma decisão a respeito da escolha de HATHEWAY & WILLIAMS (1958) pelo estimador $\hat{V}(V_i)$, e sua influência na determinação da matriz dos pesos a serem usados na estimação de b , carece de maiores estudos.

Retomando, porém, o raciocínio de HATHEWAY & WILLIAMS (1958), tem-se:

Como as V' são funções lineares de V_1 , V_2 , V_3 e V_4 , e estas são independentes, fica fácil determinar as variâncias e covariâncias das V' :

$$a) \hat{V}(V'_{x=abc}) = \hat{V}(V_1) = \frac{2V_1^2}{d - 1}$$

$$\begin{aligned} b) \hat{V}(V'_{x=ab}) &= \hat{V} \left[\frac{(d - 1)V_1 + d(c - 1)V_2}{cd - 1} \right] \\ &= \frac{2}{(cd - 1)^2} \left[(d - 1)V_1^2 + d(c - 1)V_2^2 \right] \end{aligned}$$

$$\begin{aligned} \text{c) } \hat{V}(V'_{x=a}) &= \\ &= \frac{2}{(bcd-1)^2} [(d-1)V_1^2 + d(c-1)V_2^2 + cd(b-1)V_3^2] \end{aligned}$$

$$\begin{aligned} \text{d) } \hat{V}(V'_{x=1}) &= \\ &= \frac{2}{(abcd-1)^2} [(d-1)V_1^2 + d(c-1)V_2^2 + cd(b-1)V_3^2 + bcd(a-1)V_4^2] \end{aligned}$$

$$\begin{aligned} \text{e) } \text{COV}(V'_{x=abc}, V'_{x=ab}) &= \text{COV}(V'_{x=ab}, V'_{x=abc}) \\ &= E\{[V'_{x=abc} - E(V'_{x=abc})][V'_{x=ab} - E(V'_{x=ab})]\} \end{aligned}$$

Mas:

$$\begin{aligned} 1) V'_{x=abc} &= V_1 \quad \text{e} \quad E(V'_{x=abc}) = E(V_1), \quad \text{e} \\ V'_{x=abc} - E(V'_{x=abc}) &= V_1 - E(V_1) \end{aligned}$$

$$2) V'_{x=ab} = \frac{(d-1)V_1 + d(c-1)V_2}{cd-1},$$

$$E(V'_{x=ab}) = \frac{(d-1)}{cd-1} E(V_1) + \frac{d(c-1)}{cd-1} E(V_2), \quad \text{e}$$

$$V'_{x=ab} - E(V'_{x=ab}) = \frac{d-1}{cd-1} [V_1 - E(V_1)] + \frac{d(c-1)}{cd-1} [V_2 - E(V_2)]$$

Assim:

$$\begin{aligned} \text{COV}(V'_{x=abc}, V'_{x=ab}) &= E \left[\left\{ V_1 - E(V_1) \right\} \left\{ \frac{(d-1)}{cd-1} [V_1 - E(V_1)] + \frac{d(c-1)}{cd-1} [V_2 - E(V_2)] \right\} \right] \\ &= E \left\{ \frac{d-1}{cd-1} [V_1 - E(V_1)]^2 + \frac{d(c-1)}{cd-1} [V_1 - E(V_1)] [V_2 - E(V_2)] \right\} \\ &= \frac{d-1}{cd-1} V(V_1) + \frac{d(c-1)}{cd-1} \text{COV}(V_1, V_2) \end{aligned}$$

Mas V_1 e V_2 são independentes, logo
 $\text{COV}(V_1, V_2) = 0$.

Assim:

$$\text{COV}(V'_{x=abc}, V'_{x=ab}) = \frac{d-1}{cd-1} V(V_1)$$

Daqui obtêm-se:

$$\begin{aligned} \widehat{\text{COV}}(V'_{x=abc}, V'_{x=ab}) &= \frac{d-1}{cd-1} \widehat{V}(V_1) \\ &= \frac{d-1}{cd-1} \frac{2V_1^2}{d-1} = \frac{2V_1^2}{cd-1} \end{aligned}$$

Procedendo de maneira análoga, pode-se determinar as demais covariâncias, e a matriz de variâncias e covariâncias das V' fica determinada e caracterizada, segundo HATHEWAY & WILLIAMS (1958), por:

$$\widehat{V}(V') = \begin{bmatrix} \frac{D}{(d-1)^2} & \frac{D}{(d-1)(cd-1)} & \frac{D}{(d-1)(bcd-1)} & \frac{D}{(d-1)(abcd-1)} \\ \frac{D}{(d-1)(cd-1)} & \frac{C+D}{(cd-1)^2} & \frac{C+D}{(cd-1)(bcd-1)} & \frac{C+D}{(cd-1)(abcd-1)} \\ \frac{D}{(d-1)(bcd-1)} & \frac{C+D}{(cd-1)(bcd-1)} & \frac{B+C+D}{(bcd-1)^2} & \frac{B+C+D}{(bcd-1)(abcd-1)} \\ \frac{D}{(d-1)(abcd-1)} & \frac{C+D}{(cd-1)(abcd-1)} & \frac{B+C+D}{(bcd-1)(abcd-1)} & \frac{A+B+C+D}{(abcd-1)^2} \end{bmatrix},$$

onde:

$$\begin{aligned} A &= 2bcd(a-1)V_4^2 & B &= 2cd(b-1)V_3^2 \\ C &= 2d(c-1)V_2^2 & D &= 2(d-1)V_1^2 \end{aligned}$$

Segundo HATHEWAY & WILLIAMS (1958) a matriz dos pesos W^{-1} é obtida a partir da inversa da matriz $\hat{V}(V')$, multiplicando-se cada linha e cada coluna desta inversa pelas V' correspondentes (Apêndice).

Verifica-se que a inversa de $\hat{V}(V')$ é dada por:

$$[\hat{V}(V')]^{-1} = \begin{bmatrix} (d-1)^2 \left[\frac{1}{C} + \frac{1}{D} \right] & \frac{-(d-1)(cd-1)}{C} & 0 & 0 \\ \frac{-(d-1)(cd-1)}{C} & (cd-1)^2 \left[\frac{1}{B} + \frac{1}{C} \right] & \frac{-(cd-1)(bcd-1)}{B} & 0 \\ 0 & \frac{-(cd-1)(bcd-1)}{B} & (bcd-1)^2 \left[\frac{1}{A} + \frac{1}{B} \right] & \frac{-(bcd-1)(abcd-1)}{A} \\ 0 & 0 & \frac{-(bcd-1)(abcd-1)}{A} & \frac{(abcd-1)^2}{A} \end{bmatrix}$$

De $\hat{V}(V')^{-1}$ determina-se a matriz dos pesos w^{ij} , segundo HATHEWAY & WILLIAMS (1958):

$$W^{-1} = \begin{bmatrix} (d-1)^2 \left[\frac{1}{C} + \frac{1}{D} \right] (V'_{x=a\hat{c}})^2 & \frac{-(d-1)(cd-1)V'_{x=a\hat{c}} V'_{x=a\hat{b}}}{C} & 0 & 0 \\ \frac{-(d-1)(cd-1)V'_{x=a\hat{b}} V'_{x=a\hat{c}}}{C} & (cd-1)^2 \left[\frac{1}{B} + \frac{1}{C} \right] (V'_{x=a\hat{b}})^2 & \frac{-(cd-1)(bcd-1)V'_{x=a\hat{b}} V'_{x=a}}{B} & 0 \\ 0 & \frac{-(cd-1)(bcd-1)V'_{x=a} V'_{x=a\hat{b}}}{B} & (bcd-1)^2 \left[\frac{1}{A} + \frac{1}{B} \right] (V'_{x=a})^2 & \frac{-(bcd-1)(abcd-1)V'_{x=a\hat{b}} V'_{x=a}}{A} \\ 0 & 0 & \frac{-(bcd-1)(abcd-1)V'_{x=1} V'_{x=a}}{A} & \frac{(abcd-1)^2 (V'_{x=1})^2}{A} \end{bmatrix}$$

HATHEWAY & WILLIAMS (1958) estendem o uso desta metodologia, de obtenção da estimativa ponderada do coeficiente b , a dados experimentais com efeitos de tratamentos, aperfeiçoando a proposta de KOCH & RIGNEY (1951). A extensão dessa metodologia será, então, abordada quando da descrição

do método de determinação do tamanho ótimo de parcela de KOCH & RIGNEY (1951), mais adiante.

2.2.4. Exemplo ilustrativo 2

Seja o exemplo apresentado em 2.1.1., subdividido à semelhança do proposto por KOCH & RIGNEY (1951) e HATHEWAY & WILLIAMS (1958), para o caso em que:

$$d = 2, \quad c = 5, \quad b = 2 \quad e \quad a = 5,$$

e caracterizado a seguir:

Bloco	Totais de parcelas						Totais de parcelas						Totais de blocos
1	842	844	808	822	979	4295	954	965	906	898	856	4579	8874
2	803	841	870	970	943	4427	914	916	836	858	926	4450	8877
3	773	782	860	822	932	4169	971	765	875	853	936	4400	8569
4	912	887	815	937	844	4395	661	841	844	809	778	3933	8328
5	874	792	803	793	818	4080	799	767	855	792	858	4071	8151
TOTAL DE REPETIÇÃO 1 : 42799													
6	908	875	899	788	867	4337	790	831	757	751	774	3903	8240
7	875	907	921	963	875	4541	880	898	802	874	928	4382	8923
8	891	928	871	875	865	4430	777	738	796	855	901	4067	8497
9	823	784	754	873	764	3998	775	752	753	820	798	3898	7896
10	785	794	764	822	714	3879	748	724	717	736	724	3649	7528
TOTAL DE REPETIÇÃO 2 : 41084													
TOTAL GERAL : 83883													

Sua análise de variância é a seguinte:

FONTE DE VARIAÇÃO	GL	SQ	QM
Repetições	$2 - 1 = 1$	29.412,25	$29.412,25 = v_1$
Blocos dentro de repetições	$2 \cdot (5 - 1) = 8$	157.539,76	$19.692,47 = v_2$
Parcelas dentro de blocos	$2 \cdot 5 \cdot (2 - 1) = 10$	75.637,70	$7.563,77 = v_3$
Subparcelas dentro de parcelas	$2 \cdot 5 \cdot 2 \cdot (5 - 1) = 80$	205.138,40	$2.564,23 = v_4$
Subparcelas	$2 \cdot 5 \cdot 2 \cdot 5 - 1 = 99$	467.728,11	

A partir da análise de variância, obtem-se:

$$v'_{x=abc} = v'_{x=50} = v_1 = 29.412,250000$$

$$v'_{x=ab} = v'_{x=10} = \frac{(d - 1)v_1 + d(c - 1)v_2}{cd - 1} = 20.772,445556$$

$$v'_{x=a} = v'_{x=5} = \frac{(d - 1)v_1 + d(c - 1)v_2 + cd(b - 1)v_3}{bcd - 1} = 13.820,511053$$

$$v'_{x=1} = \frac{(d - 1)v_1 + d(c - 1)v_2 + cd(b - 1)v_3 + bcd(a - 1)v_4}{abcd - 1} = 4.724,526364$$

e ainda,

$$v_{x=50} = \frac{v'_{x=50}}{50} = 558,245000$$

$$v_{x=10} = \frac{v'_{x=10}}{10} = 2.077,244556$$

$$v_{x=5} = \frac{v'_{x=5}}{5} = 2.764,102211$$

$$v_{x=1} = v'_{x=1} = 4.724,526364$$

Antes de calcular, diretamente, a matriz dos pesos, é preciso obter:

$$A = 2bcd(a - 1)V_4^2 = 1.052.044.078,864$$

$$B = 2cd(b - 1)V_3^2 = 1.144.212.332,258$$

$$C = 2d(c - 1)V_2^2 = 6.204.693.995,2144$$

$$D = 2(d - 1)V_1^2 = 1.730.160.900,125$$

Assim, pela seção 2.2.3., onde se observa a forma algébrica de W^{-1} , esta inversa é determinada numericamente por:

$$W^{-1} = \begin{bmatrix} 0,639424 & -0,886213 & 0 & 0 \\ -0,886213 & 36,178974 & -42,904413 & 0 \\ 0 & -42,904413 & 125,805251 & -116,744265 \\ 0 & 0 & -116,744265 & 207,947817 \end{bmatrix}$$

Note-se que $\Sigma w^{ij} = 49,5 = 1/2 (abcd - 1)$, ou seja: Σw^{ij} é igual à metade do total dos números de graus de liberdade das somas de quadrados das quais as V_x são estimadas, pois, explicam HATHEWAY & WILLIAMS (1958), se as variâncias não são afetadas pelo tamanho da parcela, então todas as somas de quadrados servem para estimar a variância básica (V_1). As diferentes estimativas do logaritmo natural das variâncias obtidas das diferentes linhas da análise de variância são independentes e têm variância assintótica igual a duaz vezes a recíproca dos correspondentes números de graus de liberdade. Consequentemente, a informação de cada uma delas é a metade do seu número de graus de liberdade e, portanto, a informação

(de todas as estimativas do logaritmo natural das variâncias) é a metade do total dos números de graus de liberdade envolvidos.

Na verdade, quando $\Sigma w^{ij} = 1/2(abcd - 1)$, está implícito que a transformação logarítmica aplicada aos dados foi de logaritmo natural. Ao se usar logaritmos decimais e ao se calcular a matriz W^{-1} da maneira que se procede aqui, embora o valor de $\hat{b}$ continue com a mesma caracterização, ou seja:

$$\hat{b} = \frac{-U}{T},$$

temos que a sua variância passa a ser:

$$\hat{V}(\hat{b}) = \frac{1}{5,302 T} = 0,1886 T^{-1},$$

bem como também os verdadeiros pesos serão os w^{ij} multiplicados por 5,302, uma vez que:

$$\begin{aligned} \log_e V'_i \cdot \log_e V'_j &= \frac{\log_{10} V'_i \log_{10} V'_j}{(\log_{10} e)^2} \\ &= \log_{10} V'_i \cdot \log_{10} V'_j \cdot (\log_{10} e)^{-2} \\ &= 5,302 \cdot \log_{10} V'_i \cdot \log_{10} V'_j \end{aligned}$$

Considere agora a Tabela 5, que auxiliará no cálculo da estimativa 6.

Tabela 5. Tamanho de parcela e variância por área unitária (x e V_x), respectivos logaritmos e número de graus de liberdade associados (w).

x	V_x	$\log_{10} x = x'$	$\log_{10} V_x = y'$	w
1	4.724,526364	0,	3,674358	99
5	2.764,102211	0,698970	3,441554	19
10	2.077,244556	1,000000	3,317488	9
50	588,245000	1,698970	2,769558	1

Pela tabela anterior, obtêm-se:

a) Usando ponderação pelos GL associados:

$$\hat{b} = - \frac{\sum w_i x_i' y_i' - \frac{\sum w_i x_i' \cdot \sum w_i y_i'}{\sum w_i}}{\sum w_i x_i'^2 - \frac{[\sum w_i x_i']^2}{\sum w_i}} = 0,374228$$

$$R^2 = \frac{SQReg}{SQTotal} = \frac{(\hat{b}_1)^2 \left[\sum w_i x_i'^2 - \frac{(\sum w_i x_i')^2}{\sum w_i} \right]}{\sum w_i y_i'^2 - \frac{(\sum w_i y_i')^2}{\sum w_i}} = \frac{2,335542}{2,424893} = 0,963153$$

b) Usando a ponderação proposta por HATHEWAY & WILLIAMS (1958), e seguindo a sua marcha de cálculo:

$$b.1) x_j = \sum_i w^{ij} x_i'$$

$$x_1 = 0,639424(1,698970) + (-0,886213)(1,000000)$$

$$= 0,200149$$

$$x_2 = (-0,886213)(1,698970) + 36,178974 (1,000000) +$$

$$+ (-42,904413)(0,698970) = 4,684427$$

$$x_3 = (-42,904413)(1,000000) + (125,805251)(0,698970)$$

$$= 45,029683$$

$$x_4 = (-116,744265) \cdot 0,698970 = -81,600739$$

$$\sum x_j = -31,686480$$

$$b.2) Y_j = \sum_i w^{ij} y_i$$

$$Y_1 = (0,639424)(2,769558) + (-0,886213)(3,317488)$$

$$= -1,169079$$

$$Y_2 = (-0,886213)(2,769558) + 36,178974 \cdot (3,317488)$$

$$+ (-42,904413)(3,441554) = -30,088960$$

$$Y_3 = (-42,904413)(3,317488) + (125,805251)(3,441554)$$

$$+ (-116,744265)(3,674358) = -138,329534$$

$$Y_4 = (-116,744265)(3,441554) + 207,947817 (3,674358)$$

$$= 362,293033$$

$$\sum Y_j = 192,705460$$

$$b.3) T = \sum_i x_i x_i' - \frac{(\sum_i x_i)^2}{\sum_{ij} w^{ij}}$$

$$= 0,200149 (1,698970) + 4,684427 (1,000000) +$$

$$+ 45,029683 (0,698970) - \frac{(-31,686480)^2}{49,5}$$

$$= 16,215377$$

$$\begin{aligned}
 \text{b.4) } U &= \sum_{ij} \sum w^{ij} y_i' (x_j' - \bar{x}') = \\
 &= \sum x_i y_i' - \frac{\sum x_i \sum y_i}{\sum \sum w^{ij}} = \sum y_i x_i' - \frac{\sum x_i \sum y_i}{\sum \sum w^{ij}}
 \end{aligned}$$

Usando a última igualdade, vem:

$$\begin{aligned}
 U &= (-1,169079)(1,698970) + (-30,088960)(1,000000) + \\
 &\quad + (-138,329534)(0,698970) - \\
 &\quad - \frac{(-31,686480)(192,705460)}{49,5} \\
 &= -5,406663
 \end{aligned}$$

$$\text{b.5) } \hat{b} = \frac{-U}{T} = 0,333424$$

$$\text{b.6) } s(\hat{b}) = \left[\frac{0,1886}{T} \right]^{1/2} = 0,107847$$

b.7) O desvio da regressão é testado da maneira a seguir:

$$5,302 \left(v - \frac{U^2}{T} \right) \sim \chi_{n-2}^2,$$

onde:

$$v = \sum_i y_i y_i' - \frac{(\sum y_i)^2}{\sum \sum w^{ij}} = 1,881118$$

Assim:

$$5,302 \left(v - \frac{U^2}{T} \right) = 0,078384 = \chi_{\text{calculado}}^2 - 2GL$$

Comparando o $\chi^2_{\text{calculado-2GL}}$ com o $\chi^2_{\text{tabelado-2GL}}$, ao nível $\alpha = 0,05$ ($\chi^2_{\text{tabelado}} = 5,99$), conclui-se que os dados se ajustam bem à relação linear e, por extensão, à lei empírica de SMITH (1938).

$$\text{b.8) } R^2 = \frac{\text{SQReg}}{\text{SQTotal}} = \frac{5,302(U^2/T)}{5,302(V)} = 0,959710.$$

Adotando a mesma relação de custos $K_1/K_2 = 4$, de 2.2.2., vem:

$$x_{\text{ótimo}} = \frac{\hat{b}}{1 - \hat{b}} \cdot \frac{K_1}{K_2} \approx 2,$$

e assim, o tamanho ótimo de parcela, para este caso, é de 2 m^2 .

Comparando os R^2 obtidos com as duas ponderações, observa-se que aquele obtido com a ponderação através dos GL associados a cada ponto é levemente superior àquele obtido pela ponderação proposta por HATHEWAY & WILLIAMS (1958), porém as propriedades deste último ($\hat{b} = -U/T$) são superiores, principalmente na variância mínima assintótica.

Comparando a estimativa de b obtida pela metodologia de HATHEWAY & WILLIAMS (1958) com aquela obtida pela metodologia de SMITH (1938), que resultaram em $\hat{b} = 0,33366$ e $\hat{b} = 0,6323$, calculadas em 2.2.4. e 2.2.2., respectivamente, nota-se uma grande diferença entre elas, em razão dos diferentes números de pontos (4 e 15 pontos, respectivamente) que

serviram de base para os cálculos das estimativas de b , porém a metodologia de HATHEWAY & WILLIAMS (1958) pode ser aperfeiçoada, de modo a conservar a maior quantidade de pontos possíveis e não prejudicar a estimativa de b resultante da sua aplicação.

Outra razão para a diferença entre as estimativas de b seria a forma das parcelas utilizadas nas duas metodologias: enquanto na metodologia de SMITH (1938) as formas variam de quadradas a retangulares, umas com a maior dimensão no sentido das colunas e outras com a maior dimensão no sentido das linhas, na metodologia de HATHEWAY & WILLIAMS (1958), no exemplo em questão, as quatro formas de parcelas são em sua maioria retangulares (1 x 1, 1 x 5, 1 x 10, 5 x 10), com a maior dimensão no sentido das linhas, uniformizando, para mais, os valores dos CV_x e V_x , diminuindo consideravelmente a estimativa de b . Uma solução seria subdividir o ensaio em branco em parcelas de formas variadas, observando as limitações da disposição das unidades básicas do ensaio em branco e do delineamento simulado.

Observa-se que os quatro pontos da Tabela 5, utilizados na obtenção de estimativa de b pela metodologia de HATHEWAY & WILLIAMS (1958), fazem parte, como não poderia deixar de ser, do conjunto maior dos quinze pontos da Tabela 4. Ao se usar apenas esses quatro pontos e a metodologia de SMITH (1938), obtém-se uma estimativa de b igual a:

$$\hat{b} = 0,3742 \quad ,$$

bem semelhante àquela obtida pela metodologia de HATHEWAY & WILLIAMS (1958), e igual a:

$$\hat{b} = 0,3334$$

Vale ressaltar que a metodologia de HATHEWAY & WILLIAMS (1958) tende a diminuir o valor da estimativa de b , em relação à metodologia de SMITH (1938), porém não é verdade que ao se utilizar dessa metodologia obtenha-se, necessariamente, estimativas de b menores que a unidade, como levam a crer os autores, pois BRIM & MASON (1959) aplicaram essa metodologia e obtiveram estimativas de b superiores à unidade.

Há casos em que o desvio da regressão é significativo. Nestes casos, explicam HATHEWAY & WILLIAMS (1958), uma vez que dados relativos aos custos estejam disponíveis, o tamanho ótimo de parcela pode ser estimado com razoável precisão.

Para tanto, seja o custo de uma parcela de tamanho x , dado por:

$$C = K_1 + K_2 x \quad ,$$

onde K_1 e K_2 são os fatores de custos já definidos anteriormente.

Seja $1/V_x$ a informação fornecida ao se utilizarem parcelas de tamanho x , então, o custo por parcela e por unidade de informação é dado por:

$$\frac{C}{V_x} = (K_1 + K_2 x) V_x = F(x)$$

A partir dos dados experimentais, $F(x)$ pode ser determinada para alguns valores de x , e daí o seu mínimo é facilmente observável. Há um exemplo numérico a este respeito em 2.4.2. adiante.

2.3. Método da Máxima Curvatura Modificado

Com o intuito de livrar a determinação do ponto de máxima curvatura da dependência da escala dos eixos coordenados, LESSMAN & ATKINS (1963) propõem uma modificação no método da máxima curvatura, baseados no método empírico de SMITH (1938), mais especificamente na equação geral daquele método.

Com as constantes apropriadas, a equação geral:

$$y = \frac{a}{x^b},$$

define a relação entre a variância da produção por área unitária e tamanho de parcela em termos de unidades básicas (SMITH, 1938), ou seja:

$$V_x = \frac{V_1}{x^b}.$$

Esta mesma equação geral define, também, a relação entre o coeficiente de variação e o tamanho de parcela, quando as constantes apropriadas, digamos a' e b' , são conhecidas. Isto é explicável, porque os coeficientes de variação computados nos ensaios em branco são dados por:

$$CV_x = \frac{(S_x^2)^{1/2}}{\bar{x}} \cdot 100$$

onde S_x^2 já foi definido em 2.2.1., e:

$\bar{x}$ = média da produção das parcelas de x unidades básicas de tamanho, o que não deixa de ser uma medida do tamanho da parcela, em termos de produção.

As fórmulas de obtenção de V_x e CV_x são dadas por:

$$V_x = \frac{S_x^2}{x^2}, \quad CV_x = \frac{S_x^2}{\bar{x}} \cdot 100 = \left[\frac{S_x^2}{\bar{x}^2} \right]^{1/2} \cdot 100,$$

mas x é uma medida do tamanho da parcela, em termos de unidades básicas, e $\bar{x}$ também é uma medida do tamanho da parcela, em termos de produção, o que leva a uma forte relação entre x e $\bar{x}$, e entre V_x e CV_x , justificando o uso da equação geral, dada anteriormente, para relacionar o coeficiente de variação e o tamanho da parcela, ou seja:

$$CV_x = \frac{a'}{b'x}.$$

Estimados a' e b' pelo método dos mínimos quadrados ponderados (GL associados) determina-se a derivada

de primeira ordem da função, e dela calcula-se a inclinação (M) da reta tangente para cada valor de x .

O ângulo de intersecção entre duas retas tangentes aos pontos x_1 e x_2 pode ser determinado através de:

$$\operatorname{tg} \theta = \frac{M_2 - M_1}{1 + M_1 M_2}.$$

As duas retas tangentes com o maior ângulo de intersecção devem indicar a região de uma curva onde a taxa direcional de mudança na função é maior em relação a incrementos de x . Esta região é determinada através de cálculos sequenciais, onde os incrementos de x são fixos e, no caso específico do artigo de LESSMAN & ATKINS (1963), os incrementos de x foram de um décimo de unidade básica.

Com este ponto de máxima curvatura, LESSMAN & ATKINS (1963) determinam o número ótimo de unidades básicas (z), através de:

$$z = (\text{ponto de máxima curvatura}) \cdot \frac{K_1}{K_2},$$

onde K_1 e K_2 são os fatores de custos do método de SMITH (1938).

Os autores afirmam que ao se proceder desta maneira, chegar-se-ia a resultados comparáveis àqueles obtidos pelo método de SMITH (1938) explicando que:

a) Os coeficientes de variação e as variâncias entre parcelas foram calculadas de um mesmo grupo de dados,

b) A relação entre coeficiente de variação e tamanho de parcela é similar à relação entre variância da produção por área unitária e tamanho de parcela.

Então, o ponto de máxima curvatura, em termos de x , da expressão:

$$CV_x = \frac{a'}{b'x},$$

não deve ser independente do coeficiente de regressão b da expressão:

$$V_x = \frac{a}{bx}.$$

Supondo que esta relação seja "direta e razoavelmente precisa", a melhor estimativa da região de máxima curvatura, em termos de x , da relação $CV_x = a'/x^{b'}$, deve variar igualmente a:

$$\frac{b}{1-b} \quad (0 < b < 1),$$

não dando, porém, maiores explicações.

Segundo os autores, se a relação for consistente, tal procedimento pode ser usado para detectar a ocorrência de erros no cálculo de b e como uma medida confirmatória do tamanho ótimo de parcela.

A respeito deste método, MEIER & LESSMAN (1971) escrevem que ao se determinar a região de máxima curvatura através de cálculos seqüenciais, com incrementos fixos

na variável x , estar-se-ia cometendo um erro, pois, a incrementos fixos e sucessivos de x correspondem comprimentos variáveis da curva, ou seja, LESSMAN & ATKINS (1963) ao determinarem a região de máxima curvatura através da fórmula:

$$\operatorname{tg} \theta = \frac{M_2 - M_1}{1 + M_1 M_2},$$

e incrementos fixos de x , não consideraram que a curvatura também é função do comprimento do arco considerado.

MEIER & LESSMAN (1971) aconselham tomar incrementos iguais ao longo da curva, e a curvatura em qualquer ponto da linha é, então, determinada por:

$$K = \frac{y''}{(1 + y'^2)^{3/2}},$$

onde:

K = curvatura da linha

y' e y'' são a primeira e segunda derivadas da função que define a linha.

A curvatura é máxima onde a derivada primeira de K em relação a x for igual a zero.

Para a função:

$$CV_x = \frac{a'}{x^{b'}},$$

o ponto de máxima curvatura, x_c , é dado por:

$$x_c = \left[\frac{a'^2 b'^2 (2b' + 1)}{(b' + 2)} \right]^{1/(2b' + 2)}$$

MEIER & LESSMAN (1971), após obterem x_c , propõem uma fórmula para $x'_{\text{ótimo}}$, dada por:

$$x'_{\text{ótimo}} = \left(\frac{x_c}{A} \cdot \frac{K_1}{K_2} \right)^{1/2},$$

onde K_1 e K_2 são os fatores de custos propostos por SMITH (1938), e A é a área da unidade básica. Porém, a argumentação apresentada para justificar esta fórmula de $x'_{\text{ótimo}}$ foi, aparentemente, insuficiente para sustentá-la, de modo que não lhe será dada maior atenção.

2.3.1. Exemplo ilustrativo

Do exemplo das Tabelas 1 e 2, obtém-se a Tabela 6.

Tabela 6. Tamanho (x) de parcelas e coeficientes de variação (CV), e respectivos logaritmos e número de graus de liberdade associados, de um ensaio em branco com arroz, adaptado de GOMEZ & GOMEZ (1984).

x	CV _x	log ₁₀ x = x'	log ₁₀ CV _x = y'	w
1	8,1942	0,000000	0,913507	99
2	7,3192	0,301030	0,864464	49
2	6,5724	0,301030	0,817724	49
4	6,0975	0,602060	0,785152	24
5	6,2676	0,698970	0,797102	19
5	3,9035	0,698970	0,591454	19
10	5,4334	1,000000	0,735072	9
10	2,0818	1,000000	0,318439	9
10	5,6767	1,000000	0,754096	9
10	3,6863	1,000000	0,566591	9
20	5,4599	1,301030	0,737185	4
20	1,6009	1,301030	0,204364	4
25	3,4435	1,397940	0,537000	3
50	2,8914	1,698970	0,461108	1
50	2,0552	1,698970	0,312854	1

Da Tabela 6, obtém-se:

$$\sum w_i = 308$$

$$\sum w_i x_i' = 124,511240 \Rightarrow \bar{x}_{\text{pond}} = 0,404257$$

$$\Sigma w_i x_i'^2 = 97,322466$$

$$\Sigma w_i x_i' y_i' = 84,433432$$

$$\Sigma w_i y_i' = 245,609557 \Rightarrow \bar{y}_{pond} = 0,797434$$

$$\hat{b}' = -\hat{b}_1 = \frac{\Sigma w_i x_i' y_i' - \frac{\Sigma w_i x_i' \Sigma w_i y_i'}{\Sigma w_i}}{\Sigma w_i x_i'^2 - \frac{(\Sigma w_i x_i')^2}{\Sigma w_i}} = 0,316167$$

$$\hat{b}_0 = \bar{y}_{pond} - \hat{b}_1 \cdot \bar{x}_{pond} = 0,925247$$

$$\log_{10} \hat{a}' = \hat{b}_0 \Rightarrow \hat{a}' = 8,418738$$

Assim:

$$CV_x = \frac{8,418738}{x^{0,316167}}$$

$$\begin{aligned} \frac{\partial CV_x}{\partial x} &= \frac{8,418738 \cdot (-0,316167)}{x^{1,316167}} = \\ &= \frac{-2,661727}{x^{1,316167}} \end{aligned}$$

Tabela 7. Determinação do ponto de máxima curvatura da função $CV_x = 8,418738 \cdot x^{-0,316167}$, através de cálculos sequenciais (LESSMAN & ATKINS, 1963).

x	M	$1 + M_1 M_2$	$M_2 - M_1$	$\text{tg } \theta = \frac{M_2 - M_1}{1 + M_1 M_2}$
1,0	-2,661727			
1,1	-2,347925	7,249535	0,313802	0,043286
1,2	-2,093864	5,916236	0,254061	0,042943
1,3	-1,884497	4,945880	0,209367	0,042332
1,4	-1,709366	4,221295	0,175131	0,041488
1,5	-1,560983	3,668291	0,148383	0,040450

Pela Tabela 7 percebe-se que a região de "máxima curvatura" está situada entre $1,0 < x < 1,1$.

Apenas para fixar um valor, tomando-se $x = 1,1$ como ponto de máxima curvatura e $K_1/K_2 = 4$, obtêm-se

$$z = (\text{ponto de máxima curvatura}) \cdot \frac{K_1}{K_2} = 4,4 ,$$

ou seja, o tamanho ótimo de parcela, por este método, fica em torno de $4,4 \text{ m}^2$.

Em 2.2.2., tem-se que:

$$\hat{b}_{\text{equiv}} = 0,647$$

e então:

$$\frac{\hat{b}_{\text{equiv}}}{1 - \hat{b}_{\text{equiv}}} \approx 1,83 ,$$

e comparando este valor com o ponto de máxima curvatura ($x = 1,1$), temos que a aproximação entre o ponto de máxima curvatura e $b/(1-b)$ não é tão boa quanto admitem LESSMAN & ATKINS (1963).

Analisando melhor a relação $b/(1-b)$, onde $0 < b < 1$, verifica-se que a mesma assume valores entre 0 e infinito, enquanto que o ponto de máxima curvatura varia entre a unidade e o maior tamanho de parcela possível de se obter no ensaio em branco.

Pelo método de MEIER & LESSMAN (1971) o

$x_{\text{crítico}}$ e o $x_{\text{ótimo}}'$ são obtidos da maneira a seguir:

$$x_{\text{crítico}} = [a'^2 b'^2 (2b' + 1) / (b' + 2)]^{1/(2b'+2)} \approx 1,84.$$

$$x'_{\text{ótimo}} = \sqrt{\frac{x_c}{A} \cdot \frac{K_1}{K_2}} = 2,71 ,$$

assumindo $\frac{K_1}{K_2} = 4$ e $A = 1 \text{ m}^2$.

Seja a Tabela 8, a seguir, com os resultados dos três métodos já abordados até o momento.

Tabela 8. Resultados dos métodos da máxima curvatura, de SMITH e máxima curvatura modificado.

	Máxima curvatura	SMITH		Máxima curvatura modificado	
		Original	Modificado	LESSMAN & ATKINS	MEIER & LESSMAN
$x_{\text{crítico}}$				1,10	1,84
$x'_{\text{ótimo}}$ 15 pontos	10,50	7,33		4,40	2,71
$\hat{b}_{15 \text{ pontos}}$		0,6323			
$\hat{b}_{\text{equiv}}$ 15 pontos		0,6470			
$\hat{b}_4 \text{ pontos}$		0,3742	0,3336		
$\hat{b}_{\text{equiv (interp)}}$ 4 pontos		0,4224			
$x'_{\text{ótimo}}$ 4 pontos		2,93	2,00		

Pela Tabela 8 observa-se que, ao se utilizar 15 pontos, o $x_{\text{ótimo}}$ varia de $2,71 \text{ m}^2$ a $10,5 \text{ m}^2$, não havendo, aparentemente, uma equivalência entre os métodos a este respeito, mesmo comparando o método de SMITH (1938) com o de LESSMAN & ATKINS (1963), que, segundo estes autores, deveria servir de elemento confirmatório dos resultados do método de SMITH (1938).

Porém, fica a informação do ponto crítico do método de MEIER & LESSMAN (1971), $x_c = 1,84$ unidades básicas, que representa o ponto da curva

$$CV_x = \frac{a'}{x b'}$$

onde a taxa de diminuição do CV tem maior desaceleração.

Observa-se, também, que a metodologia proposta por HATHEWAY & WILLIAMS (1958) reduziu o número de pontos de 15 para 4, alterando substancialmente a estimativa de b e o valor do $x_{\text{ótimo}}$, já comentado em 2.2.4..

2.4. Método de E.J.KOCH & H.J.RIGNEY

KOCH & RIGNEY (1951) percebem que as esperanças matemáticas dos quadrados médios da análise de variância de alguns experimentos comuns (com efeitos de tratamentos) podem ser relacionadas com as esperanças matemáticas da análise de variância de dados provenientes de um ensaio em branco, per

mitindo a determinação do coeficiente b de SMITH (1938) a partir dos dados de determinados delineamentos experimentais, não necessitando, então, de ensaios em branco.

Segundo KOCH & RIGNEY (1951) a estimativa do coeficiente b é obtida pelo método dos mínimos quadrados não ponderado, descartando a ponderação proposta por SMITH (1938), pois as V_x obtidas de dados experimentais são calculadas a partir de várias estimativas (quadrados médios), não havendo técnica desenvolvida (1951) que estime satisfatoriamente os números de graus de liberdade associados a cada V_x .

A técnica de SATTERTHWAITE (1946) serve para determinar, aproximadamente, os números de graus de liberdade associados a funções lineares de quadrados médios independentes, bem caracterizadas pelas funções lineares que expressam as diversas V_x , podendo ser utilizada na aplicação da ponderação proposta por SMITH (1938).

Porém, a metodologia de ponderação proposta por HATHEWAY & WILLIAMS (1958), abordada na seção 2.2.3., pode ser estendida a dados de experimentos em parcelas subdivididas e em reticulado quadrado, sem o inconveniente da redução no número de pontos que servirá de base à determinação da estimativa de b , comentada na seção 2.2.4., uma vez que este método fornece informações sobre um número, geralmente, já limitado e bem reduzido de formas de parcelas.

Assim, a obtenção das V_x , a partir de dados

experimentais, obedecerá à metodologia proposta por KOCH & RIGNEY (1951), e a obtenção da estimativa de b seguirá as orientações de HATHEWAY & WILLIAMS (1958).

2.4.1. Dados de experimentos em parcelas subdivididas

KOCH & RIGNEY (1951) explicam que a análise de variância de um experimento em parcelas subdivididas pode ser "reconstruída" para dar uma estimativa de variância entre parcelas:

- a) Do tamanho de uma repetição,
- b) Usuais do experimento (*main plots*),
- c) Do tamanho das subparcelas.

O problema consiste em "reconstruir" a análise de variância do experimento, de modo que simule dados de um ensaio em branco. A principal diferença entre os dados de um ensaio em branco e dados de um experimento em parcelas subdivididas (ou em reticulado quadrado) é a presença de efeitos de tratamentos.

A análise de variância de um ensaio em branco subdividido de modo a simular um experimento em parcelas subdivididas (ou em reticulado quadrado) é o seguinte:

FONTE DE VARIAÇÃO	GL	QM	E(QM)
Repetições	$d - 1$	V_1	$S + aP + abB + abcR$
Blocos dentro de repetições	$d(c - 1)$	V_2	$S + aP + abB$
Parcelas dentro de blocos	$cd(b - 1)$	V_3	$S + aP$
Subparcelas dentro de parcelas	$bcd(a - 1)$	V_4	S

Um experimento em parcelas subdivididas, com "a" observações por subparcela, apresenta a seguinte análise de variância:

FONTE DE VARIAÇÃO	GL	QM	E(QM)
Entre blocos	$d - 1$	V_1	$S + aP + abB + abcR$
Entre parcelas dentro de blocos	$d(c - 1)$	V_2	$S + aP + abB$
Entre parcelas	$cd - 1$		
Subparcelas dentro de parcelas	$cd(b - 1)$	V_3	$S + aP$
Entre subparcelas	$bcd - 1$		
Sub-subparcelas dentro de subparcelas	$bcd(a - 1)$	V_4	S
Entre sub-subparcelas	$abcd - 1$		

Do quadro de análise pode-se determinar as V' .

Considerando a unidade básica como sendo a sub-subparcela, obtém-se:

$$a) V'_{x=abc} = \frac{\text{SQ entre blocos}}{d - 1} = V_1 =$$

$$= \hat{S} + a\hat{P} + ab\hat{B} + abc\hat{R}$$

$$b) V'_{x=ab} = \frac{\text{SQ entre parcelas, corrigida para Trat(1)}}{cd - 1}$$

$$= \frac{(d - 1)V_1 + d(c - 1)V_2}{cd - 1} = \hat{S} + a\hat{P} + ab\hat{B} + \frac{d - 1}{cd - 1} abc\hat{R}$$

$$c) V'_{x=a} =$$

$$= \frac{\text{SQ entre subparcelas, corrigida para Trat(1), Trat(2) e Int.Trat(1) x Trat(2)}}{bcd - 1}$$

$$= \frac{(d - 1)V_1 + d(c - 1)V_2 + cd(b - 1)V_3}{bcd - 1} =$$

$$= \hat{S} + a\hat{P} + \frac{dc - 1}{bcd - 1} ab\hat{B} + \frac{d - 1}{bcd - 1} abc\hat{R}$$

$$d) V'_{x=1} =$$

$$= \frac{\text{SQ entre sub-subparcelas, corrigida para Trat(1), Trat(2) e Int.Trat(1) x Trat(2)}}{abcd - 1}$$

$$= \frac{(d - 1)V_1 + d(c - 1)V_2 + cd(b - 1)V_3 + bcd(a - 1)V_4}{abcd - 1}$$

$$= \hat{S} + \frac{(bcd - 1)}{abcd - 1} a\hat{P} + \frac{(cd - 1)}{abcd - 1} ab\hat{B} + \frac{(d - 1)}{abcd - 1} abc\hat{R}$$

Comparando estes resultados com aqueles apresentados em 2.2.3., nota-se que são idênticos em termos de componentes de variância.

Reduzindo as V' a área unitária, vem:

$$a) V_{x=abc} = \frac{V'_{x=abc}}{abc} = \frac{V_1}{abc}$$

$$b) V_{x=ab} = \frac{V'_{x=ab}}{ab}$$

$$c) V_{x=a} = \frac{V'_{x=a}}{a}$$

$$d) V_{x=1} = V'_{x=1}$$

Sabe-se, de 2.2.3., que:

a) V_1, V_2, V_3 e V_4 são independentes

$$b) \hat{V}(V_i) = \frac{2V_i^2}{n-1}, \quad i = 1, 2, 3, 4, \text{ e } n-1$$

representa o número de graus de liberdade associado a V_i .

As V' são obtidas por combinação linear das V_i . Associando este fato às duas assertivas (a e b) anteriores, pode-se determinar as variâncias de cada V' e as suas covariâncias:

$$a) \hat{V}(V'_{x=abc}) = \hat{V}(V_1) = \frac{2V_1^2}{d-1}$$

$$b) \hat{V}(V'_{x=ab}) = \hat{V} \left[\frac{(d-1)V_1 + d(c-1)V_2}{cd-1} \right] =$$

$$= 2 \frac{[(d-1)^2 V_1^2 + d^2(c-1)V_2^2]}{(d-1)(cd-1)^2}$$

$$\begin{aligned} \text{c) } \hat{V}(V'_{x=a}) &= \hat{V} \left[\frac{(d-1)V_1 + d(c-1)V_2 + cd(b-1)V_3}{bcd-1} \right] \\ &= 2 \left[\frac{(d-1)^2 V_1^2 + d^2(c-1)V_2^2 + cd^2(b-1)V_3^2}{(bcd-1)^2(d-1)} \right] \end{aligned}$$

$$\begin{aligned} \text{d) } \hat{V}(V'_{x=1}) &= \\ &= \hat{V} \left[\frac{(d-1)V_1 + d(c-1)V_2 + cd(b-1)V_3 + bcd(a-1)V_4}{abcd-1} \right] = \\ &= 2 \left[\frac{(d-1)^2 V_1^2 + d^2(c-1)V_2^2 + cd^2(b-1)V_3^2 + bcd(a-1)(d-1)V_4^2}{(abcd-1)^2(d-1)} \right] \end{aligned}$$

$$\begin{aligned} \text{e) } \text{COV}(V'_{x=abc}, V'_{x=ab}) &= \text{COV}(V'_{x=ab}, V'_{x=abc}) = \\ &= E \left\{ [V'_{x=abc} - E(V'_{x=abc})] [V'_{x=ab} - E(V'_{x=ab})] \right\} \end{aligned}$$

Mas:

$$\text{i) } V'_{x=abc} = V_1$$

$$\text{ii) } E(V'_{x=abc}) = E(V_1)$$

$$\text{iii) } V'_{x=abc} - E(V'_{x=abc}) = V_1 - E(V_1)$$

$$\text{iv) } E(V'_{x=ab}) = \frac{d-1}{cd-1} E(V_1) + \frac{d(c-1)}{cd-1} E(V_2)$$

$$\begin{aligned} \text{v) } V'_{x=ab} - E(V'_{x=ab}) &= \\ &= \frac{d-1}{cd-1} [V_1 - E(V_1)] + \frac{d(c-1)}{cd-1} [V_2 - E(V_2)] \end{aligned}$$

Assim:

$$\begin{aligned} \text{COV}(V'_{x=abc}, V'_{x=ab}) &= E \left\{ \left[V_1 - E(V_1) \right] \left\{ \frac{d-1}{cd-1} [V_1 - E(V_1)] + \frac{d(c-1)}{cd-1} [V_2 - E(V_2)] \right\} \right\} \\ &= \frac{d-1}{cd-1} V(V_1) + \frac{d(c-1)}{cd-1} \text{COV}(\hat{V}_1, V_2) \end{aligned}$$

Mas V_1 e V_2 são independentes, então:

$$\begin{aligned} \widehat{\text{COV}}(V'_{x=abc}, V'_{x=ab}) &= \frac{d-1}{cd-1} \hat{V}(V_1) \\ &= \frac{2V_1^2}{cd-1} \end{aligned}$$

As demais covariâncias são obtidas por processo semelhante, e a matriz de variâncias e covariâncias das V' fica determinada e caracterizada, segundo HATHEWAY & WILLIAMS (1958), por:

$$\hat{V}(V') = \begin{bmatrix} \frac{D}{(d-1)^2} & \frac{D}{(d-1)(cd-1)} & \frac{D}{(d-1)(bcd-1)} & \frac{D}{(d-1)(abcd-1)} \\ \frac{D}{(d-1)(cd-1)} & \frac{C+D}{(cd-1)^2} & \frac{C+D}{(cd-1)(bcd-1)} & \frac{C+D}{(cd-1)(abcd-1)} \\ \frac{D}{(d-1)(bcd-1)} & \frac{C+D}{(cd-1)(bcd-1)} & \frac{B+C+D}{(bcd-1)^2} & \frac{B+C+D}{(bcd-1)(abcd-1)} \\ \frac{D}{(d-1)(abcd-1)} & \frac{C+D}{(cd-1)(abcd-1)} & \frac{B+C+D}{(bcd-1)(abcd-1)} & \frac{A+B+C+D}{(abcd-1)^2} \end{bmatrix}$$

onde:

$$A = 2bcd(a-1)V_4^2$$

$$B = \frac{2cd^2(b-1)}{d-1} V_3^2$$

$$C = \frac{2d^2(c-1)}{d-1} V_2^2$$

$$D = 2(d - 1)v_1^2$$

A matriz W^{-1} dos pesos w^{ij} , é obtida a partir da inversa da matriz $\hat{V}(V')$, multiplicando-se cada linha e coluna de $\hat{V}(V')^{-1}$ pelas v' correspondentes, caracterizada algebricamente, segundo HATHEWAY & WILLIAMS (1958), por:

$$I = \begin{bmatrix} (d-1)^2 \left[\frac{1}{c} + \frac{1}{d} \right] (v'_{x=abc})^2 & \frac{-(d-1)(cd-1)}{c} v'_{x=abc} v'_{x=ab} & 0 & 0 \\ \frac{-(d-1)(cd-1)}{c} v'_{x=ab} v'_{x=abc} & (cd-1)^2 \left[\frac{1}{b} + \frac{1}{c} \right] (v'_{x=ab})^2 & \frac{-(cd-1)(cd-1)}{b} v'_{x=ab} v'_{x=a} & 0 \\ 0 & \frac{-(cd-1)(bcd-1)}{b} v'_{x=a} v'_{x=ab} & (bcd-1)^2 \left[\frac{1}{a} + \frac{1}{b} \right] (v'_{x=a})^2 & \frac{-(bcd-1)(abcd-1)}{a} v'_{x=a} v'_{x=1} \\ 0 & 0 & \frac{-(bcd-1)(abcd-1)}{a} v'_{x=1} v'_{x=a} & \frac{(abcd-1)^2}{a} (v'_{x=1})^2 \end{bmatrix}$$

Verifica-se que

$$\begin{aligned} \Sigma w^{ij} &= \frac{1}{2} \left[(d-1) + (c-1)(d-1) + c(b-1)(d-1) + bcd(a-1) \right] \\ &= \frac{1}{2} bc(ad-1) \end{aligned}$$

(HATHEWAY & WILLIAMS, 1958; ROSSETTI, 1979).

Obtidos os pesos w^{ij} , os procedimentos, nos demais passos da ponderação, são idênticos aos apresentados em 2.2.4.

Como uma parte dos números de graus de liberdade é desviada para a estimação dos efeitos de tratamentos, as variâncias entre parcelas e subparcelas são estimadas com um número de graus de liberdade menor, e, portanto, com menor preci-

são do que poderiam ter se os dados fossem provenientes de um ensaio em branco (sem efeitos de tratamentos) (HATHEWAY & WILLIAMS, 1958).

O uso de experimentos de campo para estimar o coeficiente b tem duas vantagens principais (GOMEZ & GOMEZ, 1984):

a) Este procedimento é menos dispendioso do que conduzir um ensaio em branco, porque o tempo e o dinheiro necessário ao ensaio em branco são economizados.

b) Vários experimentos podem ser usados, sem ocasionar maiores custos, e assim mais áreas e épocas são abrangidas, ao contrário da área e época limitadas de um ensaio em branco.

Há, também, as desvantagens (GOMEZ & GOMEZ, 1984):

a) As estimativas das variâncias só podem ser obtidas para apenas uns poucos tamanhos de parcelas, em cada experimento, e, então, o ajuste da regressão tem pouca precisão.

b) Os blocos, na maioria dos experimentos, são formados de modo que uma grande parte da variabilidade do solo permaneça entre blocos, o que leva a um valor superestimado da variância de parcelas do tamanho de um bloco, introduzindo uma tendenciosidade na estimativa de b .

GOMEZ & GOMEZ (1984) aconselham ao pesquisador

conduzir pelo menos um ensaio em branco para avaliar e confirmar a validade da relação entre a variância e o tamanho, antes de tentar usar dados existentes de experimentos de campo na estimação do coeficiente b .

2.4.2. Exemplo ilustrativo 1.

Seja o exemplo adaptado de HATHEWAY & WILLIAMS (1958), referente a um ensaio de competição de variedades de soja, cuja análise de variância é reproduzida a seguir:

FONTE DE VARIAÇÃO	GL	QM
Blocos	$d - 1 = 2$	$452 = v_1$
Trat(1) (Variedades)	$c - 1 = 11$	$d(c - 1) = 33$ $10.589 = v_2$
Resíduo (a)	$(c - 1)(d - 1) = 22$	
Parcelas	$cd - 1 = 35$	$5.938 = v_3$
Trat(2)	$b - 1 = 1$	
Int.Trat(1) x Trat(2)	$(b - 1)(c - 1) = 11$	
Resíduo (b)	$c(b - 1)(d - 1) = 24$	
Subparcelas	$bcd - 1 = 71$	$2.862 = v_4$
Sub-subparcelas dentro de subparcelas	$bcd(a - 1) = 72$	
Sub-subparcelas	$abcd - 1 = 143$	

Do quadro de análise observa-se que:

$$d = 3, \quad c = 12, \quad b = 2 = a.$$

Assim:

$$a) \quad V'_{x=48} = V_1 = 452 \implies V_{x=48} = \frac{452}{48} = 9,4167$$

$$b) \quad V'_{x=4} = \frac{(d-1)V_1 + d(c-1)V_2}{cd-1} \approx 10.010 \implies$$

$$V_{x=4} = \frac{10.010}{4} = 2.502,5$$

$$c) \quad V'_{x=2} = \frac{(d-1)V_1 + d(c-1)V_2 + cd(b-1)V_3}{bcd-1} \approx$$

$$\approx 7,945 \implies V_{x=2} = \frac{7.945}{2} = 3.972,5$$

$$d) \quad V'_{x=1} = \frac{(d-1)V_1 + d(c-1)V_2 + cd(b-1)V_3 + bcd(a-1)V_4}{abcd-1} \approx$$

$$\approx 5.386 \implies V_{x=1} = 5.386.$$

É possível assim estruturar a Tabela 9, a seguir:

Tabela 9. Logaritmos decimais do tamanho (x) de parcelas e da variância V_x , a um ensaio de competição de variedades de soja.

x	V_x	$\log_{10} x = x'$	$\log_{10} V_x = y'$
48	9,4167	1,6812	0,9739
4	2.502,5	0,6021	3,3984
2	3.972,5	0,3010	3,5991
1	5.386,0	0,0000	3,7313

Fonte: HATHEWAY & WILLIAMS (1958)

Segundo os autores, o coeficiente de regressão não ponderado ($\hat{b}_{NP}$), é:

$$\hat{b}_{NP} = \frac{-\left[\sum x'y' - \frac{\sum x' \sum y'}{n}\right]}{\left[\sum (x')^2 - \frac{(\sum x')^2}{n}\right]} = 1,7334$$

Segundo a marcha de cálculo seguida pelos autores para a obtenção da matriz dos pesos e estimativa de b , obtiveram:

$$a) A = 2bcd(a - 1)v_4^2 = 1.179,5 \times 10^6$$

$$b) B = \frac{2cd^2(b - 1)v_3^2}{(d - 1)} = 3.808,1 \times 10^6$$

$$c) C = \frac{2d^2(c - 1)v_2^2}{(d - 1)} = 11.100,6 \times 10^6$$

$$d) D = 2(d - 1)v_1^2 = 0,8 \times 10^6$$

$$e) w^{11} = (d - 1)^2 (v'_{x=abc})^2 \left[\frac{1}{C} + \frac{1}{D} \right] = 1,00$$

$$f) w^{12} = w^{21} = \frac{-(d - 1)(cd - 1)}{C} \cdot v'_{x=abc} v'_{x=ab} = -0,03$$

$$g) w^{13} = w^{31} = w^{14} = w^{41} = w^{24} = w^{42} = 0$$

$$h) w^{22} = (cd - 1)^2 \left[\frac{1}{B} + \frac{1}{C} \right] (v'_{x=ab})^2 = 43,29$$

$$\begin{aligned} \text{i) } w^{23} = w^{32} &= \frac{-(cd - 1)(bcd - 1)}{B} V'_{x=ab} V'_{x=a} = \\ &= -51,90 \end{aligned}$$

$$\text{j) } w^{33} = (bcd - 1)^2 \left[\frac{1}{A} + \frac{1}{B} \right] (V'_{x=a})^2 = 353,36$$

$$\begin{aligned} \text{l) } w^{34} = w^{43} &= \frac{-(bcd - 1)(abcd - 1)}{A} V'_{x=a} V'_{x=1} = \\ &= -368,34 \end{aligned}$$

$$\text{m) } w^{44} = \frac{(abcd - 1)^2}{A} (V'_{x=1})^2 = 502,89.$$

Assim a matriz dos pesos (W^{-1}) fica:

$$W^{-1} = \begin{bmatrix} 1,00 & -0,03 & 0,00 & 0,00 \\ -0,03 & 43,29 & -51,90 & 0,00 \\ 0,00 & -51,90 & 353,36 & -368,34 \\ 0,00 & 0,00 & -368,34 & 502,89 \end{bmatrix}, \quad e$$

$$\Sigma w^{ij} = \frac{1}{2} bc(ad - 1) = 60.$$

$$\text{n) } Y_j = \sum_i w^{ij} y_i'$$

$$Y_1 = 1,00 \cdot 0,9739 + (-0,03) \cdot 3,3984 = 0,87$$

$$Y_2 = -39,19$$

$$Y_3 = -282,52$$

$$Y_4 = 554,42$$

$$\sum_j Y_j = 233,58$$

$$o) X_j = \sum_i w_{ij} x_i'$$

$$X_1 = 1,66$$

$$X_2 = 10,39$$

$$X_3 = 75,11$$

$$X_4 = -110,87$$

$$\sum_j X_j = -23,71$$

$$p) T = \sum_{ij} X_i x_i' - \frac{(\sum_i X_i)^2}{\sum_{ij} w_{ij}} = 31,65 - \frac{(-23,71)^2}{60} = 22,28$$

$$q) U = \sum_{ij} Y_i x_i' - \frac{\sum_i X_i \sum_i Y_i}{\sum_{ij} w_{ij}} = -14,87$$

$$r) V = \sum_{ij} Y_i y_i' - \frac{(\sum_i Y_i)^2}{\sum_{ij} w_{ij}} = 13,05$$

$$s) \hat{b} = \frac{-U}{T} = 0,667$$

$$t) s(\hat{b}) = \left[\frac{0,1886}{T} \right]^{1/2} = 0,092$$

$$u) SQDR = 5,302 \left[V - \frac{U^2}{T} \right] = 16,57 \sim \chi_2^2$$

$$\text{Mas } \chi_2^2 = 9,21, \text{ logo } \chi_2^2 = 16,57^{**}, \text{ indican } 0,99$$

do que os dados fogem significativamente da relação linear as sumida (lei empírica de SMITH, 1938).

O resultado do ítem "u" pode causar preocupação, mas, uma vez que estejam disponíveis dados relativos aos custos do experimento, o tamanho ótimo pode ser estimado com razoável precisão (HATHEWAY & WILLIAMS, 1958), conforme se mostrou no final da seção 2.2.4..

Para ilustrar, seja o exemplo adaptado de HATHEWAY & WILLIAMS (1958), no qual a função de custos foi determinada como sendo:

$$\hat{C} = K_1 + K_2x = 4,9 + 1,43x ,$$

e os desvios da regressão são significativos, resumido na Ta bela 10, a seguir:

Tabela 10. Dados referentes a um experimento que apresentou desvios significativos da lei empírica de SMITH (1938), para o qual foi determinada uma função de custos do tipo $C = K_1 + K_2x$ e aproveitada para a determinação do tamanho ótimo de parcela, de acordo com HATHEWAY & WILLIAMS (1958).

x	V_x	$\hat{C} = 4,9 + 1,43x$	$F(x) = (4,9 + 1,43x) V_x$
1	97.869	6,33	619.500
4	22.306	10,62	236.900
16	8.647	27,28	240.200
64	4.330	96,42	417.500

Pela Tabela 10 se observa que $F(x)$ tem mínimo para $4 < x < 16$. Aceitando que nesta região não há desvios consideráveis de linearidade $[\log V_x = \log V_1 - \log x]$, a estimativa de b será aproximadamente dada por:

$$\hat{b} = -\frac{[\log \hat{V}_{16} - \log \hat{V}_4]}{\log 16 - \log 4} = 0,68$$

Logo, o tamanho ótimo de parcela será:

$$x_{\text{ótimo}} = \frac{\hat{b}}{(1 - \hat{b})} \cdot \frac{K_1}{K_2} = 7,28 \text{ unidades básicas}$$

2.4.3. Dados de experimentos em reticulado quadrado

KOCH & RIGNEY (1951) comparam também os componentes de variância de uma análise de variância de dados provenientes de um ensaio em branco subdividido de modo a simular um experimento em reticulado quadrado, com os componentes de variância de uma análise deste delineamento, com "a" observações por parcela. O quadro de análise e as esperanças dos quadrados médios deste experimento estão reproduzidas a seguir:

FONTE DE VARIAÇÃO	GL	QM	E(QM)
Repetições	$d - 1$	V_1	$S + aP + a\bar{b}B + a\bar{c}C$
Blocos ajustados dentro de repetições	$d(c - 1)$	V_2	$S + aP + a\bar{b}B$
Tratamentos	$b \cdot c - 1$		$S + aP + a\bar{d}D^1$
Resíduo	$cd(b - 1) - (bc - 1)$	V_3	$S + aP$
Erro amostral	$bcd(a - 1)$	V_4	S
TOTAL	$abcd - 1$		

¹ Esta esperança contém efeitos de blocos.

As estimativas dos componentes de variância podem ser obtidas facilmente da análise anterior, com a diferença de que é preciso determinar o valor de λ , que varia de acordo com o tipo do delineamento em *lattice*. A fórmula geral de λ para *lattice* $K \times K$ (reticulados quadrados) é dada por:

$$\lambda = \frac{b(d-1)}{d}$$

As V' são obtidas substituindo as estimativas de S , P , B e R nas suas respectivas fórmulas apresentadas em 2.2.3. (KOCH & RIGNEY, 1951).

Mas, ao expressar as V' em função dos componentes de variância, a determinação da matriz de variância e covariância das V' é dificultada pela presença de correlação entre os componentes de variância (HATHEWAY & WILLIAMS, 1958; SEARLE, 1971). Entretanto, não notaram aqueles autores que as V' podem ser escritas em função de V_1 , V_2 , V_3 e V_4 , bastando para tanto introduzir fatores de correção, F e f , da seguinte maneira:

Após a obtenção das estimativas de S , P , B , R e V_2 , relaciona-se V_2 com um fator F_2 , e o seu valor em termos de estimativas de componentes de variância com um fator $f_2 = d/(d-1)$, ou seja:

$$F_2 V_2 = \hat{S} + a\hat{P} + af_2 \lambda \hat{B} = \hat{S} + a\hat{P} + a \frac{d}{d-1} \frac{b(d-1)}{d} \hat{B} = \hat{S} + a\hat{P} + ab\hat{B},$$

e então:

$$F_2 = \frac{\hat{S} + a\hat{P} + ab\hat{B}}{V_2}$$

Pode-se escrever, então, que:

$$a) V'_{x=abc} = V_1 = \hat{S} + a\hat{P} + ab\hat{B} + abc\hat{R}$$

$$b) V'_{x=ab} = \frac{(d-1)V_1 + d(c-1)F_2V_2}{cd-1} =$$

$$\begin{aligned} & \frac{(d-1)(\hat{S} + a\hat{P} + ab\hat{B} + abc\hat{R}) + d(c-1) \frac{(\hat{S} + a\hat{P} + ab\hat{B})}{V_2} V_2}{cd-1} = \\ & = \hat{S} + a\hat{P} + ab\hat{B} + \frac{d-1}{cd-1} abc\hat{R}. \end{aligned}$$

Analogamente, obtêm-se:

$$c) V'_{x=a} = \hat{S} + a\hat{P} + \frac{(cd-1)}{bcd-1} ab\hat{B} + \frac{(d-1)}{bcd-1} abc\hat{R}$$

$$\begin{aligned} d) V'_{x=1} &= \hat{S} + \frac{(bcd-1)a\hat{P}}{abcd-1} + \frac{cd-1}{abcd-1} ab\hat{B} + \\ &+ \frac{(d-1)}{abcd-1} abc\hat{R} \end{aligned}$$

Comparando estas fórmulas com as caracterizações das V' em 2.2.3., nota-se que são idênticas em termos de componentes de variâncias.

Considerando F_2 como constante, podemos determinar a matriz de variâncias e covariâncias das V' , ou seja:

$$a) \hat{V}(V'_{x=abc}) = \hat{V}(V_1) = \frac{2V_1^2}{d-1}$$

$$b) \hat{V}(V'_{x=ab}) = \hat{V} \left[\frac{(d-1)V_1 + d(c-1)F_2V_2}{cd-1} \right] =$$

$$= \frac{2}{(cd-1)^2} \left[(d-1)V_1^2 + F_2^2 d(c-1)V_2^2 \right]$$

$$c) \hat{V}(V'_{x=a}) =$$

$$= \hat{V} \left[\frac{(d-1)V_1 + d(c-1)F_2V_2 + cd(b-1)V_3}{bcd-1} \right] =$$

$$= \frac{2}{(bcd-1)^2} \left[(d-1)V_1^2 + F_2^2 d(c-1)V_2^2 + \frac{c^2 d^2 (b-1)^2 V_3^2}{cd(b-1) - (bc-1)} \right]$$

$$d) \hat{V}(V'_{x=1}) =$$

$$= \hat{V} \left[\frac{(d-1)V_1 + d(c-1)F_2V_2 + cd(b-1)V_3 + bcd(a-1)V_4}{abcd-1} \right] =$$

$$= \frac{2}{(abcd-1)^2} \left[(d-1)V_1^2 + F_2^2 d(c-1)V_2^2 + \frac{c^2 d^2 (b-1)^2 V_3^2}{cd(b-1) - (b-1)} + bcd(a-1)V_4^2 \right]$$

$$e) \text{COV}(V'_{x=abc}, V'_{x=ab}) = \text{COV}(V'_{x=ab}, V'_{x=abc}) =$$

$$= E[V'_{x=abc} - E(V'_{x=abc})][V'_{x=ab} - E(V'_{x=ab})]$$

Mas:

$$i) V'_{x=abc} = V_1$$

$$ii) E(V'_{x=abc}) = E(V_1)$$

$$iii) V'_{x=abc} - E(V'_{x=abc}) = V_1 - E(V_1)$$

$$iv) V'_{x=ab} = \frac{(d-1)V_1 + d(c-1)F_2V_2}{cd-1}$$

$$v) E(V'_{x=ab}) = \frac{d-1}{cd-1} E(V_1) + \frac{F_2 d(c-1)}{cd-1} E(V_2)$$

$$vi) V'_{x=ab} - E(V'_{x=ab}) = \frac{d-1}{cd-1} [V_1 - E(V_1)] + \\ + \frac{F_2 d(c-1)}{cd-1} [V_2 - E(V_2)]$$

Então:

$$COV(V'_{x=abc}, V'_{x=ab}) = E \left\{ V_1 - E(V_1) \right\} \left\{ \frac{d-1}{cd-1} [V_1 - E(V_1)] + \right. \\ \left. + \frac{F_2 d(c-1)}{cd-1} [V_2 - E(V_2)] \right\} \\ = \frac{d-1}{cd-1} V(V_1) + \frac{F_2 d(c-1)}{cd-1} COV(V_1, V_2)$$

Mas V_1 e V_2 são independentes, então:

$$COV(V'_{x=abc}, V'_{x=ab}) = \frac{d-1}{cd-1} \cdot 2 \cdot \frac{V_1^2}{d-1} = \frac{2V_1^2}{cd-1}$$

Analogamente, obtemos as demais covariâncias, explicitadas a seguir:

$$f) \widehat{COV}(V'_{x=abc}, V'_{x=a}) = \frac{2V_1^2}{bcd-1}$$

$$g) \widehat{COV}(V'_{x=abc}, V'_{x=1}) = \frac{2V_1^2}{abcd-1}$$

$$h) \widehat{COV}(V'_{x=ab}, V'_{x=a}) =$$

$$= \frac{2}{(cd-1)(bcd-1)} \left[(d-1)V_1^2 + F_2^2 d(c-1)V_2^2 \right]$$

$$\begin{aligned}
 \text{i) } \widehat{\text{COV}}(V'_{x=ab}, V'_{x=1}) &= \\
 &= \frac{2}{(cd - 1)(abcd - 1)} \left[(d - 1)V_1^2 + F_2^2 d(c - 1)V_2^2 \right] \\
 \text{j) } \widehat{\text{COV}}(V'_{x=a}, V'_{x=1}) &= \\
 &= \frac{2}{(bcd - 1)(abcd - 1)} \left[(d - 1)V_1^2 + F_2^2 d(c - 1)V_2^2 + \right. \\
 &\quad \left. + \frac{c^2 d^2 (b - 1)^2 V_3^2}{cd(b - 1) - (bc - 1)} \right]
 \end{aligned}$$

A partir das variâncias e covariâncias estimadas, pode-se construir a matriz de variâncias e covariâncias das $V' [\widehat{V}(V')]$ e obter a matriz dos pesos (W^{-1}) , pela multiplicação de cada linha e cada coluna da inversa de $\widehat{V}(V')$ pelas V' correspondentes.

Após a obtenção da matriz W^{-1} , os procedimentos para a determinação da estimativa do coeficiente b são idênticos àqueles reproduzidos em 2.2.4.

2.4.4. Exemplo ilustrativo 2

KOCH & RIGNEY (1951) apresentam os resultados de um experimento em reticulado quadrado 4×4 simples, no qual se testaram dezesseis variedades em quatro repetições, em parcelas de 15×40 pés = 600 pés quadrados, apresentados no quadro de análise de variância a seguir:

FONTE DE VARIAÇÃO	GL	QM	E (QM)
Repetições	$d - 1 = 3$	$36.588 = V_1$	$S + aP + abB + abcR$
Blocos ajustados dentro de repetições	$d(c - 1) = 12$	$12.856 = V_2$	$S + aP + a\lambda B$
Variedades	$bc - 1 = 15$	89.580	$S + aP + adV^1$
Resíduo	$cd(b - 1) - (bc - 1) = 33$	$6.529 = V_3$	$S + aP$
Total	$abcd - 1 = 63$		

¹ Esta esperança contém efeitos de blocos.

Observa-se que:

a) $a = 1$

b) $b = c = d = 4$

c) $\lambda = \frac{b(d - 1)}{d} = 3$

d) $\hat{S} + \hat{P} = 6.529$

e) $\hat{B} = 2.109$

f) $\hat{R} = 1.351,4375$

g) $F_2 = \frac{\hat{S} + a\hat{P} + ab\hat{B}}{V_2} = 1,164047915370255 *$

Determinando as V' e as V_x , considerando a parcela como a unidade básica, obtém-se:

a) $V'_{x=abc} = V'_{x=16} = V_1 = 36.588$ e

$$V_{x=16} = \frac{36.588}{16} = 2.286,75$$

b) $V'_{x=ab} = V'_{x=4} = \frac{(d - 1)V_1 + d(c - 1)F_2V_2}{cd - 1} =$

$$= 19.289,6 \quad e \quad V_{x=4} = \frac{19.289,6}{4} = 4.822,4$$

* O uso de uma aproximação maior é necessário, pela própria característica numérica dos dados do exemplo, para a convergência de $\sum w_{ij}$.

$$c) V'_{x=a} = V'_{x=1} =$$

$$= \frac{(d-1)V_1 + d(c-1)F_2V_2 + cd(b-1)V_3}{bcd-1} = 9.567,238095238093 \quad e$$

$$V_{x=1} = 9.567,238095238093$$

$$d) \hat{V}(V'_{x=16}) = \hat{V}(V_1) = 892.454.496$$

$$e) \hat{V}(V'_{x=4}) = \frac{2}{(cd-1)^2} \left[(d-1)V_1^2 + F_2^2 d(c-1)V_2^2 \right]$$

$$= 59.586.310,50666662$$

Analogamente para $\hat{V}(V'_{x=1})$ e covariâncias, obtem-se:

$$f) \hat{V}(V'_{x=1}) = 4.877.630,781282206$$

$$g) \widehat{COV}(V'_{x=16}, V'_{x=4}) = 178.490.899,2$$

$$h) \widehat{COV}(V'_{x=16}, V'_{x=1}) = 42.497.833,14285714$$

$$i) \widehat{COV}(V'_{x=4}, V'_{x=1}) = 14.187.216,78730157$$

A matriz $\hat{V}(V')$ é dada por:

$$\hat{V}(V') = \begin{bmatrix} 892.454.496,000000000 & 178.490.899,200000000 & 42.497.833,14285714 \\ 178.490.899,200000000 & 59.586.310,506666620 & 14.187.216,78730157 \\ 42.497.833,14285714 & 14.187.216,78730157 & 4.877.630,781282206 \end{bmatrix},$$

e a matriz W^{-1} é dada por:

$$W^{-1} = \begin{bmatrix} 3,8 & -5,9 & 0 \\ -5,9 & 29,6 & -29,3 \\ 0 & -29,3 & 61,0 \end{bmatrix}$$

A soma dos pesos, $\sum_{ij} w^{ij}$, é igual à metade

dos graus de liberdade associados a V_1 , V_2 e V_3 , ou seja:

$$\sum_{ij} w^{ij} = \frac{1}{2} [(d-1) + d(c-1) + cd(b-1) - (bc-1)] = \frac{1}{2} bc(d-1) = 24 .$$

Do exposto, obtêm-se a Tabela 11 que se segue:

Tabela 11. Logarítmo do tamanho de parcela ($\log_e x = x'$) e da variância por área unitária ($\log_e V_x = y'$), relativos a um ensaio de competição de dezesseis variedades.

x	V_x	$\log_e x = x'$	$\log_e V_x = y'$
16	2.286,7500	2,772589	7,734886
4	4.822,4000	1,386294	8,481027
1	9.567,2381	0,000000	9,166100

Com os dados da Tabela 11 e pela marcha de cálculo proposta por HATHEWAY & WILLIAMS (1958) para a obtenção da estimativa de b , obtêm-se:

$$a) X_j = \sum_i w^{ij} x_i'$$

$$\begin{aligned} X_1 &= 3,8 \cdot 2,772589 + (-5,9) \cdot 1,386294 = \\ &= 2,356703 \end{aligned}$$

$$\begin{aligned} X_2 &= (-5,9) \cdot 2,772589 + 29,6 \cdot 1,386294 = \\ &= 24,676027 \end{aligned}$$

$$X_3 = (-29,3) \cdot 1,386294 = -40,618414$$

$$\Sigma X_j = -13,585684$$

$$b) Y_j = \sum_i w^{ij} y_i'$$

$$Y_1 = -20,645492$$

$$Y_2 = -63,164158$$

$$Y_3 = 310,638009$$

$$\Sigma Y_j = 226,828359$$

$$\begin{aligned} c) T &= \sum_{ij} \sum w^{ij} x_i' (x_j' - \bar{x}') = \sum_i X_i x_i' - \frac{(\Sigma X_i)^2}{\sum_{ij} w^{ij}} = \\ &= 40,742397 - \frac{(-13,585684)^2}{24} = 33,051947 \end{aligned}$$

$$\begin{aligned} d) U &= \sum_{ij} \sum w^{ij} y_i' (x_j' - \bar{x}') = \sum_i X_i y_i' - \frac{\Sigma X_i \Sigma Y_i}{\sum_{ij} w^{ij}} = \\ &= \sum_i Y_i x_i' - \frac{\Sigma X_i \Sigma Y_i}{\sum_{ij} w^{ij}} \end{aligned}$$

Usando da última igualdade, obtém-se:

$$\begin{aligned} U &= -20,645492(2,772589) + (-63,164158)(1,386294) - \\ &\quad - \frac{(-13,585684)(226,828359)}{24} \\ &= -16,404790 \end{aligned}$$

$$e) V = \sum_i Y_i y_i' - \frac{(\Sigma Y_i)^2}{\sum_{ij} w^{ij}} =$$

$$= (-20,645492) 7,734886 + (-63,164158) 8,481027 + \\ + 310,638009(9,166100) - \frac{(226,828359)^2}{24}$$

$$= 8,155579$$

$$f) \hat{b} = \frac{-U}{T} \approx 0,496334$$

$$g) s(\hat{b}) = \left(\frac{1}{T} \right)^{1/2} = 0,173940$$

$$h) SQDR = V - \frac{U^2}{T} = 0,013331 = \chi_1^2$$

$\chi_1^2 = 0,013331$ não excede o valor tabelado da distribuição de qui-quadrado ao nível $\alpha = 5\%$ de probabilidade, e os dados se ajustam bem à lei empírica de SMITH (1938).

$$i) x_{\text{ótimo}} = \frac{\hat{b}}{1 - \hat{b}} \cdot \frac{K_1}{K_2} \quad \text{Supondo } \frac{K_1}{K_2} = 2, \text{ por}$$

exemplo, então:

$$x_{\text{ótimo}} = \frac{0,496334}{1 - 0,496334} \cdot 2 \approx 1,97 \text{ unidades bá}$$

sicas. Logo, o tamanho ótimo corresponde a aproximadamente duas vezes o tamanho da parcela experimental usada no experimento, ou seja:

$$\text{tamanho ótimo} \approx 2 [15 \text{ pés} \times 40 \text{ pés}] = 1.200$$

pés quadrados.

Assim, nos próximos experimentos o pesquisador deve aumentar as suas parcelas, porém, quando $b = 0,5$ (e aqui

$\hat{b} = 0,496334$), a eficiência não é muito afetada quando se usa parcelas com o dobro ou a metade do tamanho ótimo, reduzindo-a para 96% em relação à eficiência obtida com o tamanho ótimo (SMITH, 1938).

KOCH & RIGNEY (1951) lembram que o uso do maior número possível de experimentos para a determinação da estimativa de b para determinada área, é altamente desejável (para se obter uma estimativa média de b), lembrando também que devem ser incluídas, sempre que possível, subparcelas, e mesmo, sub-subparcelas nos experimentos, de modo a caracterizar melhor a relação entre o tamanho (x) da parcela e a sua variância (V_x).

2.5. Método de W.H.HATHEWAY

HATHEWAY (1961) desenvolveu seu próprio método de determinação do tamanho conveniente de parcelas experimentais, através da junção de duas fórmulas bem conhecidas:

$$a) V_x = \frac{V}{x}, \text{ é a fórmula de SMITH (1938), para}$$

a sua lei empírica, já definida em 2.2..

$$b) r = \frac{2C^2(t_1 + t_2)^2}{d^2}, \text{ é a fórmula de COCHRAN}$$

& COX (1957) para determinação do número de repetições, onde:

r = número de repetições necessárias para de-

testar uma diferença verdadeira d

d = diferença verdadeira entre dois tratamentos, medida em porcentagem da média

C = verdadeiro erro padrão por parcela, medido em porcentagem da média, usualmente estimado pelo coeficiente de variação

t_1 = valor crítico da distribuição t de Student, ao nível α_1 de probabilidade, de acordo com as hipóteses experimentais

t_2 = valor crítico bilateral da distribuição t de Student, ao nível $\alpha_2 = 2(1 - P)$ de probabilidade, onde P é a probabilidade de se obter um resultado significativo.

Destas duas fórmulas HATHEWAY (1961) desenvolve o seu método da seguinte maneira:

Seja S_x^2 a variância da produção entre parcelas de x unidades básicas de tamanho. A variância da produção entre parcelas de x unidades básicas de tamanho reduzida à base unitária (V'_x) e à área unitária (V_x) são dadas, respectivamente, por:

$$V'_x = \frac{S_x^2}{x} \quad \text{e} \quad V_x = \frac{S_x^2}{x^2} .$$

O coeficiente de variação para parcelas de x unidades básicas de tamanho (CV_x), é dado por:

$$CV_x = \frac{S_x}{m_x} \cdot 100, \quad \text{onde}$$

m_x = média das parcelas de x unidades básicas de tamanho.

$$\text{Mas } m_1 = \frac{m_x}{x}, \text{ e então:}$$

$$m_x = m_1 \cdot x.$$

Substituindo na fórmula de CV_x , resulta:

$$CV_x = \frac{S_x}{m_1 \cdot x}$$

$$(CV_x)^2 = \frac{(S_x)^2}{x^2} \cdot \frac{1}{m_1^2} = \frac{V_x}{m_1^2}.$$

Introduzindo esta expressão na fórmula de COCHRAN & COX (1957), obtém-se:

$$r = \frac{2C^2(t_1 + t_2)^2}{d^2} = 2 \frac{V_x}{m_1^2} \cdot \frac{(t_1 + t_2)^2}{d^2}$$

$$\text{Mas } V_x = \frac{V_1}{x^b}, \text{ então}$$

$$\begin{aligned} r &= \frac{2(t_1 + t_2)^2}{d^2 m_1^2} \cdot \frac{V_1}{x^b} = \frac{2(t_1 + t_2)^2}{x^b d^2} \cdot \frac{V_1}{m_1} \\ &= \frac{2(t_1 + t_2)^2}{x^b d^2} \cdot CV_1^2, \text{ onde:} \end{aligned}$$

CV_1^2 = quadrado do coeficiente de variação de parcelas de uma unidade básica de tamanho, o qual pode, na prática, ser tomado como o tamanho correntemente em uso.

Isolando o fator x^b da expressão anterior:

$$x^b = \frac{2(t_1 + t_2)^2 CV_1^2}{rd^2},$$

que determina o denominado tamanho conveniente de parcela (x), em termos de múltiplos do tamanho básico (associado ao CV_1).

Neste método o pesquisador apenas especifica o número de repetições (r) que deseja usar, a precisão requerida (α , d , P), e o prévio conhecimento do coeficiente b de SMITH (1938), daí a denominação tamanho conveniente de parcela, não sendo necessariamente o tamanho ótimo, pois atende, apenas, a condições convenientes à pesquisa.

2.5.1. Exemplo ilustrativo 1

HATHEWAY (1961) ilustra o seu método através do seguinte exemplo:

Um melhorista de milho deseja reduzir o número de repetições de $r = 4$ para $r = 3$. Planeja usar o delineamento em blocos ao acaso, com dez variedades, e requer uma precisão para detectar, ao nível $\alpha_1 = 0,05$, uma diferença $d = 25\%$ em $P = 80\%$ dos seus experimentos. O seu coeficiente de variação é de 12%, usando parcelas de dez covas (duas filas de cinco covas). De experimentos anteriores, estimou $\hat{b} = 0,72$. Que tamanho de parcela deverá usar?

Do exposto, resulta:

FONTE DE VARIACÃO	GL	
Tratamentos	9	$t_1 = t_{[1-(\alpha/2)]} = 0,975 = 2,101$ com 18 GL
Blocos	2	$t_2 = t_{[1-(\alpha/2)]} = 0,80 = 0,862$ com 18 GL
Resíduo	18	
Total	29	

Assim:

$$x^b = \frac{2(t_1 + t_2)^2 CV_1^2}{d^2 r}$$

$$x = 1,52$$

Logo, o tamanho conveniente de parcelas, para estas condições, é $1,52 \cdot 10$ covas ≈ 16 covas em cada parcela.

Um outro enfoque abordado por HATHEWAY (1961), é o que se segue:

A relação entre o tamanho de parcela, número de repetições e a verdadeira diferença que se deseja detectar, pode ser escrita da maneira a seguir:

$$d^2 = \frac{2(t_1 + t_2)^2 CV_1^2}{r x^b}$$

Se CV_1 e b forem conhecidos, e se forem especificados níveis de significância de modo a permitir a obtenção de t_1 e t_2 , então d depende apenas de r e x , e pode ser relacionado como uma função dessas variáveis.

Assim, no exemplo anterior, para

$$a) \quad r = 2 \quad e \quad 0,5 \leq x \leq 3,0$$

$$t_1 = 2,26$$

$$t_2 = 0,883$$

$$d_{x=0,5} = 48,405667$$

$$d_{x=1,0} = 37,715999$$

$$d_{x=1,5} = 32,593640$$

$$d_{x=2,0} = 29,386979$$

$$d_{x=2,5} = 27,118608$$

$$d_{x=3,0} = 25,395806$$

$$b) \quad r = 3 \quad e \quad 0,5 \leq x \leq 3,0$$

$$t_1 = 2,101$$

$$t_2 = 0,862$$

$$d_{\hat{x}=0,5} = 37,259571$$

$$d_{x=1,0} = 29,031352$$

$$d_{x=1,5} = 25,088591$$

$$d_{x=2,0} = 22,620207$$

$$d_{x=2,5} = 20,874161$$

$$d_{x=3,0} = 19,548061$$

$$c) \quad r = 4 \quad e \quad 0,5 \leq x \leq 3,0$$

$$t_1 = 2,05$$

$$t_2 = 0,855$$

$$d_{x=0,5} = 31,636102$$

$$d_{x=1,0} = 24,649742$$

$$d_{x=1,5} = 21,301964$$

$$d_{x=2,0} = 19,206211$$

$$d_{x=2,5} = 17,723690$$

$$d_{x=3,0} = 16,597732$$

Seja então a Figura 2 a seguir, relacionando r e x com a variável d :

Figura 2 - Relação entre o tamanho (x) de parcela ($lx=10$ covas), número(r) de repetições, e diferença (d) verdadeira entre dois tratamentos, referentes aos dados do exemplo apresentado em 2.5.1.

Assim, podem-se escolher pares de número de repetições e tamanho x , de acordo com um d desejável.

Por exemplo, para $d = 30\%$ podemos escolher $x \approx 0,6$ e $r = 4$, ou $x \approx 0,9$ e $r = 3$ ou $x \approx 1,9$ e

$r = 2$, através de visualização direta no gráfico anterior, ou se quiser ser mais preciso, utilizando a fórmula de HATHEWAY (1961).

2.5.2. Exemplo ilustrativo 2

Aproveitando a estimativa de b do exemplo ilustrativo 1, em 2.2.2., onde $\hat{b} \approx 0,647$, resultando em $x_{\text{ótimo}} \approx 7 \text{ m}^2$, observa-se pela Tabela 2, que o CV_x para $x \approx 7$ deve situar-se em torno de 3% e 4%. Então, seja a seguinte condição hipotética:

Um pesquisador deseja estudar a relação entre o número de repetições, tamanho de parcela e a verdadeira diferença que deseja detectar. Para experimentos com dez variedades de arroz, no delineamento em blocos casualizados, para o qual requer uma precisão para detectar, ao nível $\alpha_1 = 0,05$, uma diferença d em $P = 80\%$ dos seus experimentos, e adotando $CV = 4,0\%$ (ao se usar parcelas de 7 m^2) e $b = 0,647$, como se comportam r , x e d ?

Pela fórmula:

$$d^2 = \frac{2(t_1 + t_2)^2 CV^2}{rx^b}, \text{ e}$$

$$a) \ r = 2 \quad \text{e} \quad 0,2 \leq x \leq 2,6$$

$$t_1 = 2,26$$

$$t_2 = 0,883$$

$$d_{0,2} = 21,160275$$

$$d_{0,4} = 16,909766$$

$$d_{0,6} = 14,831061$$

$$d_{1,0} = 12,571999$$

$$d_{1,4} = 11,275403$$

$$d_{1,8} = 10,394985$$

$$d_{2,2} = 9,741613$$

$$d_{2,6} = 9,229130$$

$$b) \quad r = 4 \quad e \quad 0,2 \leq x \leq 2,6$$

$$t_1 = 2,05$$

$$t_2 = 0,855$$

$$d_{0,2} = 13,829550$$

$$d_{0,6} = 9,693017$$

$$d_{1,0} = 8,216581$$

$$d_{1,4} = 7,369175$$

$$d_{1,8} = 6,793767$$

$$d_{2,2} = 6,366747$$

$$d_{2,6} = 6,031808$$

$$c) \quad r = 6 \quad e \quad 0,2 \leq x \leq 2,6$$

$$t_1 = 2,015$$

$$t_2 = 0,85325$$

$$d_{0,2} = 11,148932$$

$$d_{0,6} = 7,814199$$

$$d_{1,0} = 6,623939$$

$$d_{1,4} = 5,940788$$

$$d_{1,8} = 5,476914$$

$$d_{2,2} = 4,862647$$

$$d_{2,6} = 4,862647$$

Seja então a Figura 3 a seguir, relacionando x e r com a variável d :

Figura 3 - Relação entre o tamanho (x) de parcela ($1x = 7m^2$), número (r) de repetições, e diferença (d) verdadeira entre dois tratamentos, referente aos dados do exemplo apresentado em 2.5.2.

Para $d = 10\%$, por exemplo, podemos escolher $r = 6$ e $x \approx 0,3$ (parcelas de $7 \times 0,3 \approx 2,0 m^2$), ou $r = 4$

e $x \approx 0,54$ (parcelas de $7 \cdot 0,54 \approx 3,8 \text{ m}^2$) ou $r = 2$ e $x \approx 2,0$ (parcelas de $7 \cdot 2 = 14,0 \text{ m}^2$).

Para $x = 1$ (parcelas de 7 m^2), pode-se detectar como significativas diferenças (entre médias) de $d \approx 6,6\%$, ao se usar $r = 6$, de $d \approx 8,2\%$ ao se usar $r = 4$ e de $d \approx 12,6\%$ ao se usar $r = 2$.

Cabe ao pesquisador decidir qual combinação (r, x, d) é a mais conveniente. Esta decisão é função do tamanho considerado prático e que seja válido em termos de amostragem, número de graus de liberdade, área e recursos disponíveis, etc.

Uma alternativa seria conciliar mais de um método. Por exemplo, se o pesquisador deseja $d = 10\%$, e conciliando este método com o método de SMITH (1938), que resultou em $x_{\text{ótimo}} \approx 7 \text{ m}^2$, deve-se procurar a dupla (r, x) que satisfaça $d = 10\%$ e que apresente o x o mais próximo de 1 possível ($x = 1$ equivale a parcelas de 7 m^2).

Assim para $r = 3$ e $d = 10\%$, teríamos:

$$t_1 = 2,101$$

$$t_2 = 0,862$$

$$x_{0,647} = \frac{2(t_1 + t_2)^2 CV_1^2}{rd^2} = 0,936466 \implies$$

$$x = 0,9085$$

Para $d = 10$ e $r = 3$, obtém-se $x \approx 1$, de modo que conciliando estes dois métodos, pode-se decidir pela u

tilização de parcelas de 7 m² e três repetições, atendendo às considerações de custos (método de SMITH, 1938) e de precisão (método de HATHEWAY, 1961, $d = 10\%$, $P = 80\%$ e $\alpha_1 = 0,05$).

2.6. Método de PIMENTEL GOMES

Este método se aplica especialmente ao caso de experimentos com árvores, muito embora toda a teoria se aplique a espécies menores, desde que se considere linhas ou subparcelas unitárias de maneira geral (BUENO & PIMENTEL GOMES, 1983).

PIMENTEL GOMES (1984) admite um ensaio com parcelas de k árvores úteis, sendo os dados obtidos planta por planta (k subparcelas por parcela), e no caso do delineamento em blocos casualizados (DBC), com " c " cultivares de macieira (tratamentos), " b " blocos e " k " plantas por parcela, o quadro da análise de variância é dado a seguir:

FONTE DE VARIAÇÃO	GL	QM	E(QM)
Blocos	$b - 1$		
Cultivares	$c - 1$		
Resíduo A	$(b - 1)(c - 1)$	V_1	$\sigma_1^2 + k\sigma_2^2$
Resíduo B	$bc(k - 1)$	V_2	σ_1^2

Esta análise envolve o seguinte modelo matemático:

$$y_{ijk} = m + c_i + b_j + e_{ij} + e_{ijk}, \text{ onde:}$$

m = média inerente a cada observação y_{ijk}

c_i = efeito da cultivar i

b_j = efeito do bloco j

e_{ij} e e_{ijk} são erros aleatórios com distribuição normal, e variâncias σ_2^2 e σ_1^2 respectivamente, e $E(e_{ij}, e_{ijk}) = 0$.

Por este modelo deve-se ter sempre:

$$E(V_1) = \sigma_1^2 + k\sigma_2^2 \geq \sigma_1^2 = E(V_2)$$

Um modelo mais geral é:

$$y_{ijk} = m + c_i + b_j + e_{ijk}, \text{ onde:}$$

$$a) E(e_{ijk}) = 0$$

$$b) E(e_{ijk}^2) = \sigma^2 = \text{variância residual (de parcelas) quando } k = 1$$

$$c) E(e_{ijk}, e_{i'j'k'}) = 0 \text{ para } (i, j) \neq (i', j') \text{ e}$$

$$d) E(e_{ijk}, e_{ijk'}) = \rho\sigma^2, \text{ onde } \rho \text{ é o coeficiente de correlação intraclasse.}$$

A análise de variância deste modelo é:

FONTE DE VARIAÇÃO	GL	QM	E(QM)
Blocos	$b - 1$		
Cultivares	$c - 1$		
Resíduo A	$(b - 1)(c - 1)$	V_1	$\sigma^2 [1 + (k - 1)\rho]$
Resíduo B	$bc(k - 1)$	V_2	$\sigma^2(1 - \rho)$

Para $\rho \geq 0$, estes resultados equivalem aos anteriores, mas para $\rho < 0$ pode-se ter, agora, $E(V_1) < E(V_2)$.

Deve-se ter sempre:

$$\sigma^2 [1 + (k - 1)\rho] \geq 0 \implies \rho \geq \frac{-1}{k - 1} \text{ e } k > 1.$$

Da análise de variância obtém-se um estimador para ρ (e para σ^2):

$$\hat{\sigma}^2 [1 + (k - 1)\hat{\rho}] = V_1 \quad (I)$$

$$\hat{\sigma}^2(1 - \hat{\rho}) = V_2 \implies \hat{\sigma}^2 = \frac{V_2}{1 - \rho} \quad (II)$$

Substituindo-se (II) em (I):

$$\frac{V_2}{1 - \hat{\rho}} [1 + (k - 1)\hat{\rho}] = V_1 \implies V_2 + (k - 1)V_2\hat{\rho} = V_1 - V_1\hat{\rho}$$

$$V_1\hat{\rho} + (k - 1)V_2\hat{\rho} = V_1 - V_2 \implies \hat{\rho} = \frac{V_1 - V_2}{V_1 + (k - 1)V_2}$$

Desse estimador se conclui que, se $V_1 \geq V_2$, então $\hat{\rho} \geq 0$, e se $V_1 < V_2$, então $\hat{\rho} < 0$.

Sendo V_1 e V_2 ambos positivos, tem-se necessariamente:

$$a) \text{ para } V_1 = 0 \implies \hat{\rho} = \frac{-1}{k-1}$$

$$b) \text{ para } V_2 = 0 \implies \hat{\rho} = 1, \text{ e então:}$$

$$\frac{-1}{k-1} \leq \hat{\rho} \leq 1.$$

A variância de cada parcela tem a expressão:

$$V_1 = \sigma^2 [1 + (k-1)\rho],$$

e sendo "m" a média geral das subparcelas, então o coeficiente de variação (CV), é dado por:

$$\begin{aligned} CV &= \frac{100}{km} \cdot \sqrt{\sigma^2 [1 + (k-1)\rho]} = \frac{100}{m} \sigma \left(\frac{1 + k\rho - \rho}{k^2} \right)^{1/2} \\ &= \frac{100}{m} \sigma \left(\frac{1-\rho}{k^2} + \frac{\rho}{k} \right)^{1/2} \end{aligned}$$

Para $1 + (k-1)\rho \geq 0$ e $\rho \geq 0$, o máximo de CV se dá para $k=1$, pois os termos do radicando:

$$\frac{1-\rho}{k^2} \quad \text{e} \quad \frac{\rho}{k}$$

são funções decrescentes de k . Para $\rho < 0$, o numerador

$$[(1-\rho) + k\rho] \text{ do radicando } \left[\frac{(1-\rho) + k\rho}{k^2} \right]$$

decrece quando k cresce, mas seu valor é limitado pela condição necessária:

$$(1-\rho) + k\rho \geq 0.$$

Ao mesmo tempo cresce o denominador k^2 , de sorte que o radicando é monotonicamente decrescente, para $k > 1$. Conclui-se então que o CV é função decrescente do número k de árvores úteis por parcela, o que favoreceria o uso de parcelas grandes.

Porém, continua o autor, o mais importante é reduzir a variância da média de cada tratamento, sem aumentar o número de árvores do experimento.

Considerando o modelo geral:

$$y_{ijk} = m + c_i + b_j + e_{ijk},$$

já definido anteriormente, a variância da média de r repetições de um tratamento é, a nível de subparcela, dada por:

$$V(\hat{m}) = \frac{\sigma^2 [1 + (k - 1)\rho]}{kr}$$

2.6.1. Tamanho ótimo de parcela, para o caso de parcelas sem bordadura

Seja N o número total de árvores por tratamento, então:

$$N = kr$$

$$V(\hat{m}) = \frac{\sigma^2 [1 + (k - 1)\rho]}{N}$$

Caso $\rho > 0$, essa variância é mínima para

$k = 1$ (parcelas com uma árvore apenas), e para $\rho < 0$, $V(\hat{m})$ é função decrescente de k , e o mínimo se daria para $k = N$, o que implica $r = 1$, mas com duas restrições:

$$a) 1 + (k - 1)\rho \geq 0$$

b) Deve haver um número razoável de repetições ($r \geq 2$) que permita obter um número razoável de GL para o resíduo (em geral pelo menos 10 GL).

2.6.2. Tamanho ótimo de parcela, para o caso de parcelas com meia bordadura

2.6.2.1. Parcelas com meia bordadura e uma linha de árvores úteis

O número total (K) de árvores (úteis e de bordadura) por parcela é dado por:

$$K = 2(k + 1) ,$$

e o número total (N) de árvores por tratamento é dado por:

$$N = rK = r[2(k + 1)]$$

$$r = \frac{N}{2(k + 1)}$$

Vale ressaltar que as quantidades K e N , segundo as fórmulas anteriores, consideram que todas as parcelas estão circundadas por outras, em todos os seus lados, o que na prática não ocorre, pois há as parcelas que se situam nos limites do experimento, comprometendo a determinação precisa de K e N .

A utilização de K e N, da maneira como caracterizados anteriormente, na minimização de $V(\hat{m})$, nos casos de meia bordadura, acarreta prejuízos, os quais não foram abordados no presente trabalho, porém os resultados a que chega PIMENTEL GOMES (1984) não precisam necessariamente ser desprezados, apenas devem ser considerados com maior cautela.

A variância da média de tratamento é dada por:

$$V(\hat{m}) = \frac{\sigma^2 [1 + (k - 1)\rho]}{rk} = \frac{\sigma^2 [1 + (k - 1)\rho]}{\frac{N}{2(k + 1)} \cdot k}$$

Desenvolvendo e simplificando, obtêm-se:

$$V(\hat{m}) = \frac{2\sigma^2}{N} \left[\left(1 + \frac{1}{k} \right) + \left(k - \frac{1}{k} \right) \rho \right]$$

Para $\rho > 0$, esta função tem mínimo para k igual a:

$$k = \left[\frac{1 - \rho}{\rho} \right]^{1/2}, \quad \rho > 0, \quad \text{pois:}$$

$$\begin{aligned} \frac{\partial V(\hat{m})}{\partial k} &= \frac{2\sigma^2}{N} \left[\frac{-1}{k^2} + \frac{2k^2 - (k^2 - 1)}{k^2} \rho \right] \\ &= \frac{2\sigma^2}{N} \left[\frac{(k^2 + 1)\rho - 1}{k^2} \right] \end{aligned}$$

$$\text{Igualando } \frac{\partial V(\hat{m})}{\partial k} \text{ a zero:}$$

$$\frac{2\sigma^2}{N} \left[\frac{(k^2 + 1)\rho - 1}{k^2} \right] = 0$$

$$(k^2 + 1)\rho - 1 = 0$$

$$k = \left[\frac{1 - \rho}{\rho} \right]^{1/2}, \quad k \geq 1 \quad \text{e} \quad \rho > 0.$$

Para $\rho \leq 0$, $\frac{\partial V(\hat{m})}{\partial k}$ é sempre negativa, para

todo k . Para reduzir o valor de $V(\hat{m})$ deve-se aumentar o valor de k o mais possível, obedecendo às condições "a" e "b" de 2.6.1..

2.6.2.2. Parcelas com meia bordadura e com duas linhas de árvores úteis

O número total (K) de árvores por parcela é da do por:

$$K = \frac{3}{2} (k + 2),$$

e o número total (N) de árvores por tratamento é dado por:

$$N = rK = r \frac{3}{2} (k + 2) \quad \text{e} \quad r = \frac{2N}{3(k + 2)}$$

Assim:

$$V(\hat{m}) = \frac{\sigma^2}{rk} [1 + (k - 1)\rho] = \frac{\sigma^2}{2N} [1 + (k + 1)\rho] \cdot k$$

Desenvolvendo e simplificando, obtêm-se:

$$V(\hat{m}) = \frac{3\sigma^2}{2N} \left[1 + \frac{2}{k} + \left(k + 1 - \frac{2}{k} \right) \rho \right],$$

para k par e $k > 2$, pois para $k = 2$ temos duas linhas de uma árvore, ou uma linha de duas árvores, e que recai no caso anterior.

Para $\rho > 0$, $V(\hat{m})$ é mínima para:

$$k = \left[\frac{2(1 - \rho)}{\rho} \right]^{1/2},$$

uma vez que:

$$\begin{aligned} \frac{\partial V(\hat{m})}{\partial k} &= \frac{3\sigma^2}{2N} \left[\frac{k - (k + 2)}{k^2} + \frac{k(2k + 1) - (k^2 + k - 2)}{k^2} \rho \right] \\ &= \frac{3\sigma^2}{2N} \left[\frac{-2 + (k^2 + 2)\rho}{k^2} \right] \end{aligned}$$

Igualando a zero:

$$-2 + (k^2 + 2)\rho = 0$$

$$k = \left[\frac{2(1 - \rho)}{\rho} \right]^{1/2}, \rho > 0.$$

Mas como $k > 2$, então:

$$k = \left[\frac{2(1 - \rho)}{\rho} \right]^{1/2} > 2$$

$$\frac{2(1 - \rho)}{\rho} > 4$$

$$\rho < 1/3,$$

ou seja, na prática e para esta situação (meia bordadura e duas linhas úteis) é preciso que $\rho < 1/3$, pois se $\rho > 1/3$ resulta na situação anterior (meia bordadura e uma linha útil).

Seja $\rho = 0,1$, logo:

$$k = \left[\frac{2(1 - 0,1)}{0,1} \right]^{1/2} \approx 4,24$$

Substituindo os valores pares mais próximos de $k = 4,24$, na fórmula de $V(\hat{m})$, obtém-se:

$$\begin{aligned} \text{a) } k = 4 \implies V(\hat{m}) &= \frac{3\sigma^2}{N} \left[1 + \frac{2}{4} + \left(4 + 1 - \frac{2}{4} \right) 0,1 \right] \\ &= \frac{\sigma^2}{N} 2,925 \end{aligned}$$

$$\begin{aligned} \text{b) } k = 6 \implies V(\hat{m}) &= \frac{3}{2} \frac{\sigma^2}{N} \left[1 + \frac{2}{6} + \left(6 + 1 - \frac{2}{6} \right) 0,1 \right] \\ &= \frac{\sigma^2}{N} 3,000 \end{aligned}$$

Logo, escolhe-se $k = 4$ árvores úteis, em duas linhas de duas árvores, ou seja:

$$K = \frac{3}{2} (k + 2) = 9 \text{ árvores por parcela (úteis}$$

e de bordadura), ou ainda, parcelas de meia bordadura de 4 x 4 árvores.

Para $\rho \leq 0$, $\frac{\partial V(\hat{m})}{\partial k}$ é sempre negativa, para todo $k > 1$.

Para reduzir o valor de $V(\hat{m})$ deve-se aumentar k o mais possível, obedecendo às condições "a" e "b" de 2.6.1.

2.6.3. Tamanho ótimo de parcela, para o caso de parcelas com bordadura completa

2.6.3.1. Parcelas com bordadura completa e uma linha de árvores úteis

Aqui K e N são dados por:

$$K = 3(k + 2)$$

$$N = rK = r \cdot 3(k + 2) \implies r = \frac{N}{3(k + 2)},$$

e então:

$$V(\hat{m}) = \frac{\sigma^2}{rk} [1 + (k - 1)\rho] = \frac{\sigma^2}{\frac{N}{3(k + 2)} k} [1 + (k - 1)\rho]$$

Desenvolvendo e simplificando, obtêm-se:

$$V(\hat{m}) = \frac{3\sigma^2}{N} \left[\frac{k + 2}{k} + \frac{(k^2 + k - 2)\rho}{k} \right]$$

Para $\rho > 0$, esta variância é mínima para:

$$k = \left[\frac{2(1 - \rho)}{\rho} \right]^{1/2}, \quad \rho > 0 \quad \text{e} \quad k > 1$$

Para $\rho \geq 2/3$, então $k = 1$.

Para $0 < \rho < 2/3$, a expressão de k sugere os valores mais próximos (números naturais) a serem testados.

Para $\rho \leq 0$, $\frac{\partial v(\hat{m})}{\partial k}$ é negativa para todo

$k > 1$.

Para reduzir o valor de $V(\hat{m})$ 'deve-se' aumentar k o mais possível, obedecendo às condições "a" e "b" de 2.6.1...

2.6.3.2. Parcelas com bordadura completa e duas linhas de árvores úteis

Aqui K e N são dados por:

$$K = 2(k + 4)$$

$$N = rK = r \cdot 2(k + 4) \implies r = \frac{N}{2(k + 4)},$$

e então:

$$V(\hat{m}) = \frac{\sigma^2}{rk} [1 + (k - 1)\rho] = \frac{\sigma^2}{\frac{N}{2(k + 4)} k} [1 + (k - 1)\rho]$$

Desenvolvendo e simplificando, obtém-se:

$$V(\hat{m}) = \frac{2\sigma^2}{N} \left[1 + \frac{4}{k} + \left(k + 3 - \frac{4}{k} \right) \rho \right],$$

$k > 2$ e par

$$\begin{aligned} \frac{\partial V(\hat{m})}{\partial k} &= \frac{2\sigma^2}{N} \left[0 - \frac{4}{k^2} + \left(1 + \frac{4}{k^2} \right) \rho \right] \\ &= \frac{2\sigma^2}{N} \left[\frac{-4 + k^2\rho + 4\rho}{k^2} \right] \end{aligned}$$

Igualando a zero:

$$-4 + k^2\rho + 4\rho = 0 \implies k = 2 \left(\frac{1 - \rho}{\rho} \right)^{1/2}$$

Com as condições $\rho > 0$, $k > 2$ e par:

$$k = 2 \left(\frac{1 - \rho}{\rho} \right)^{1/2} > 2$$

$$\frac{1 - \rho}{\rho} > 1$$

$$\rho < 1/2,$$

mas $\rho > 0$, então:

$$0 < \rho < 1/2$$

Para $\rho \geq 1/2$ resulta no caso de bordadura completa e uma linha útil.

Para $\rho \leq 0$, $\frac{\partial V(\hat{m})}{\partial k}$ é negativa para todo k

natural par.

Para reduzir o valor de $V(\hat{m})$ deve-se aumen-

tar k o mais possível, obedecendo às condições "a" e "b" de 2.6.1...

2.6.4. Exemplo ilustrativo 1

Seja a análise a seguir (dados hipotéticos), referente a um experimento com algarobeira, no delineamento em blocos casualizados, com cinco tratamentos, cinco repetições, parcelas sem bordadura e com quatro árvores por parcela, das quais foi coletado a CAP (circunferência a altura do peito), em centímetros, aos 10 anos de plantadas:

FONTE DE VARIAÇÃO	GL	SQ	QM
Blocos	4	55,8765	
Tratamentos	4	4.257,8246	
Resíduo (A)	16	270,6884	16,918025 = V_1
Entre parcelas	24	4.584,3896	
Árvores dentro de parcelas	75	790,3900	10,538533 = V_2
Árvores	99	5.374,7796	

Da análise acima, obtém-se:

$$\hat{p} = \frac{V_1 - V_2}{V_1 + (k - 1)V_2} \approx 0,131445$$

Por 2.6.1., $V(\hat{m})$ será mínima para $k = 1$.

Logo, o tamanho ótimo é uma árvore por parcela (parcela sem bordadura).

Considerando a necessidade de bordadura, e o desejo de instalar os novos experimentos com duas linhas úteis, então $V(\hat{m})$ é mínima para:

$$k = 2 \left(\frac{1 - \rho}{\rho} \right)^{1/2} \approx 5,14$$

Pela fórmula de $V(\hat{m})$, de 2.6.3.2., dada por

$$V(\hat{m}) = \frac{2\sigma^2}{N} \left[1 + \frac{4}{k} + \left(k + 3 - \frac{4}{k} \right) \rho \right],$$

observa-se que, para:

$$a) k = 4, \text{ então } V(\hat{m}) = 5,577340 \frac{\sigma^2}{N}$$

$$b) k = 6, \text{ então } V(\hat{m}) = 5,524084 \frac{\sigma^2}{N}$$

$$c) k = 8, \text{ então } V(\hat{m}) = 5,760346 \frac{\sigma^2}{N}$$

Logo, deve-se escolher parcelas de duas linhas úteis de três árvores cada uma, ou seja, parcelas com bordadura completa, de dimensões 4 x 5 árvores por parcela.

2.6.5. Tamanho ótimo de parcela para o caso geral, considerando genericamente o número de linhas e observações úteis por parcela.

PIMENTEL GOMES (1988), complementa as considerações do seu artigo de 1984, pois quando ρ , suposto positivo, fica próximo de zero (digamos $0 < \rho < 0,15$), o número de árvores úteis (k) da parcela de tamanho ótimo pode crescer, e então, convém estudar também o número de linhas úteis (n) a ser usado.

Esta complementação só tem razão de ser quando as parcelas apresentam bordadura.

Sejam então, os casos de meia bordadura e, principalmente, o de bordadura completa, considerando $0 < \rho \leq 1$.

2.6.5.1. Parcelas com meia bordadura

Sejam n linhas úteis por parcela, então K , N e r serão dados por:

$$K = (n + 1) \left(1 + \frac{k}{n} \right) = (n + 1) \left(\frac{n + k}{n} \right) = \frac{1}{n} (n + 1) (n + k)$$

$$= \left(1 + \frac{1}{n} \right) (n + k)$$

$$N = rK = r \left(1 + \frac{1}{n} \right) (n + k) ,$$

$$r = \frac{N}{\left(1 + \frac{1}{n}\right)(n + k)}$$

Assim:

$$\begin{aligned} V(\hat{m}) &= \frac{\sigma^2}{kr} [1 + (k - 1)\rho] \\ &= \frac{\sigma^2}{k \cdot \frac{N}{\left(1 + \frac{1}{n}\right)(n + k)}} [1 + (k - 1)\rho] \end{aligned}$$

Desenvolvendo, obtém-se:

$$V(\hat{m}) = \frac{\sigma^2}{N} \left[\frac{n}{k} + \frac{n(k - 1)\rho}{k} + 1 + (k - 1)\rho + \frac{1}{k} + \frac{(k - 1)\rho}{k} + \frac{1}{n} + \frac{(k - 1)\rho}{n} \right]$$

$$\frac{\partial V(\hat{m})}{\partial k} = \frac{\sigma^2}{N} \left[\frac{-n}{k^2} + \frac{n\rho}{k^2} + \rho - \frac{1}{k^2} + \frac{\rho}{k^2} + \frac{\rho}{n} \right]$$

Igualando a zero:

$$\frac{-n^2 + n^2\rho + nk^2\rho - n + n\rho + k^2\rho}{nk^2} = 0$$

$$-n^2 - n + nk^2\rho + k^2\rho + n^2\rho + n\rho = 0$$

$$-n(1 + n) + k^2\rho(n + 1) + n\rho(n + 1) = 0$$

$$(n + 1) [-n + k^2\rho + n\rho] = 0$$

Mas $n + 1 > 0$, logo:

$$-n + k^2 \rho + n\rho = 0 \quad (I)$$

$$\frac{\partial V(\hat{m})}{\partial n} = \frac{\sigma^2}{N} \left[\frac{1}{k} + \frac{(k-1)\rho}{k} - \frac{1}{n^2} - \frac{(k-1)\rho}{n^2} \right]$$

Igualando a zero:

$$\frac{n^2 + n^2(k-1)\rho - k - k(k-1)\rho}{n^2 k} = 0$$

$$n^2 - k + n^2(k-1)\rho - k(k-1)\rho = 0$$

$$n^2 - k + (k-1)\rho(n^2 - k)\rho = 0$$

$$(n^2 - k) [1 + (k-1)\rho] = 0$$

Mas, como interessa apenas valores positivos para ρ , então:

$$1 + (k-1)\rho > 0,$$

e então:

$$n^2 - k = 0 \quad (II)$$

$$n^2 = k \quad (III)$$

Substituindo (III) em (I), vem:

$$-n + n\rho + n^4 \rho = 0$$

$$n[-1 + \rho + n^3 \rho] = 0$$

$$-1 + \rho + n^3 \rho = 0$$

$$n = \left(\frac{1 - \rho}{\rho} \right)^{1/3} \quad (IV)$$

Logo:

$$k = n^2 \quad \text{e} \quad n = \left(\frac{1 - \rho}{\rho} \right)^{1/3}$$

é solução para o sistema de equações (I) e (II)..

Seja a matriz Hessiana de $\hat{V}(\hat{m})$, dada por:

$$H = \begin{bmatrix} \frac{\partial^2 V(\hat{m})}{\partial k^2} & \frac{\partial^2 V(\hat{m})}{\partial k \partial n} \\ \frac{\partial^2 V(\hat{m})}{\partial n \partial k} & \frac{\partial^2 V(\hat{m})}{\partial n^2} \end{bmatrix},$$

os seus elementos no ponto crítico são:

$$\begin{aligned} \frac{\partial^2 V(\hat{m})}{\partial k^2} &= \frac{\sigma^2}{Nk^3} [2n - 2n\rho + 2 - 2\rho] \\ &= \frac{\sigma^2}{Nk^3} 2[n + 1 - \rho(n + 1)] = \frac{\sigma^2}{Nn^6} 2(n + 1)(1 - \rho) = \\ &= A \cdot 2(n + 1)(1 - \rho) \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 V(\hat{m})}{\partial n^2} &= \frac{\sigma^2}{N} \left[\frac{2}{n^3} + \frac{2(k - 1)}{n^3} \rho \right] = \frac{\sigma^2}{N} \left[\frac{2(1 + k\rho - \rho)}{n^3} \right] \\ &= \frac{\sigma^2}{N} \left[\frac{2(1 - \rho + n^2\rho)}{n^3} \right] = \frac{\sigma^2}{N} \cdot \frac{n^3}{n^3} \left[\frac{2(1 - \rho + n^2\rho)}{n^3} \right] = \\ &= \frac{\sigma^2}{N} \frac{2n^3}{n^6} (1 - \rho + n^2\rho) = A \cdot 2n^3(1 - \rho + n^2\rho) \end{aligned}$$

$$\frac{\partial^2 V(\hat{m})}{\partial k \partial n} = \frac{\sigma^2}{N} \left[\frac{-1}{k^2} + \frac{\rho}{k^2} - \frac{\rho}{n^2} \right]$$

Mas, $k = n^2 \implies k^2 = n^4$, e então:

$$\begin{aligned}
\frac{\partial^2 V(\hat{m})}{\partial k \partial n} &= \frac{\sigma^2}{N} \left[\frac{-1}{n^4} + \frac{\rho}{n^4} - \frac{\rho}{n^2} \right] = \frac{\sigma^2}{N} \left[\frac{-1 + \rho - \rho n^2}{n^4} \right] = \\
&= \frac{\sigma^2}{N n^6} n^2 (-1 + \rho - \rho n^2) = \frac{\sigma^2}{N n^6} [-n^2 (1 - \rho + \rho n^2)] = \\
&= A (-n^2) (1 - \rho + \rho n^2).
\end{aligned}$$

Assim, a matriz Hessiana, no ponto crítico, é dada por:

$$H = A \begin{bmatrix} 2(n+1)(1-\rho) & -n^2(1-\rho+n^2\rho) \\ -n^2(1-\rho+n^2\rho) & 2n^3(1-\rho+n^2\rho) \end{bmatrix},$$

onde $A = \frac{\sigma^2}{N n^6}$ = constante positiva e $n = \left(\frac{1-\rho}{\rho} \right)^{1/3}$.

O determinante dessa matriz é:

$$|H| = A^2 [4n^3(n+1)(1-\rho)(1-\rho+n^2\rho) - n^4(1-\rho+n^2\rho)^2]$$

Desenvolvendo esta expressão, chega-se ao seguinte resultado:

$$|H| = 3A^2 n^8 \rho^2 (n+1)^2 > 0, \text{ logo o ponto crítico}$$

$$(k, n) = \left[n^2, \left(\frac{1-\rho}{\rho} \right)^{1/3} \right]$$

é o ponto onde $V(\hat{m})$ é mínima.

2.6.5.2. Parcelas com bordadura completa

Sejam n linhas úteis por parcela, então K ,

N e r são dados por:

$$K = (n + 2) \left(2 + \frac{k}{n} \right) = \frac{1}{n} (n + 2) (2n + k) = \left(1 + \frac{2}{n} \right) (2n + k)$$

$$N = rK = r \left(1 + \frac{2}{n} \right) (2n + k)$$

$$r = \frac{N}{K} = \frac{N}{\left(1 + \frac{2}{n} \right) (2n + k)}$$

Assim:

$$V(\hat{m}) = \frac{\sigma^2}{rk} [1 + (k - 1)\rho] = \frac{\sigma^2}{\frac{N}{\left(1 + \frac{2}{n} \right) (2n + k)} \cdot k} [1 + (k - 1)\rho]$$

Desenvolvendo, chega-se a:

$$V(\hat{m}) = \frac{\sigma^2}{N} \left[1 + (k - 1)\rho + \frac{2n}{k} + \frac{2n(k - 1)\rho}{k} + \frac{2}{n} + \frac{2(k - 1)\rho}{n} + \frac{4}{k} + \frac{4(k - 1)\rho}{k} \right]$$

$$\begin{aligned} \frac{\partial V(\hat{m})}{\partial k} &= \frac{\sigma^2}{N} \left[\rho - \frac{2n}{k^2} + \frac{2n\rho}{k^2} + \frac{2\rho}{n} - \frac{4}{k^2} + \frac{4\rho}{k^2} \right] \\ &= \frac{\sigma^2}{N} \left[\frac{nk^2\rho - 2n^2 + 2n^2\rho + 2k^2\rho - 4n + 4n\rho}{nk^2} \right] \end{aligned}$$

Igualando esta derivada a zero:

$$nk^2\rho - 2n^2 + 2n^2\rho + 2k^2\rho - 4n + 4n\rho = 0$$

$$k^2\rho(n+2) + 2n^2(\rho-1) + 4n(\rho-1) = 0$$

$$k^2\rho(n+2) + 2n(\rho-1)(n+2) = 0$$

$$(n+2) \left| k^2\rho + 2n(\rho-1) \right| = 0$$

Mas $n+2 > 0$, então:

$$k^2\rho + 2n(\rho-1) = 0$$

$$k^2\rho - 2n(1-\rho) = 0 \quad (\text{I})$$

$$\begin{aligned} \frac{\partial V(\hat{m})}{\partial n} &= \frac{\sigma^2}{N} \left[\frac{2}{k} + \frac{2(k-1)\rho}{k} - \frac{2}{n^2} - \frac{2(k-1)\rho}{n^2} \right] \\ &= \frac{\sigma^2}{N} \left[\frac{2n^2 + 2n^2(k-1)\rho - 2k - 2k(k-1)\rho}{n^2k} \right] \end{aligned}$$

Igualando a zero:

$$2n^2 + 2n^2(k-1)\rho - 2k - 2k(k-1)\rho = 0$$

$$2n^2[1 + (k-1)\rho] - 2k[1 + (k-1)\rho] = 0$$

$$2[1 + (k-1)\rho][n^2 - k] = 0$$

Mas $2[1 + (k-1)\rho] > 0$ para $\rho > 0$, então:

$$n^2 - k = 0 \quad (\text{II})$$

$$n^2 = k \quad (\text{III})$$

Substituindo (III) em (I), resulta:

$$n^4\rho + 2n(\rho-1) = 0$$

$$n[n^3\rho + 2(\rho-1)] = 0$$

$$n^3\rho + 2(\rho-1) = 0$$

$$n = \left[\frac{2(1 - \rho)}{\rho} \right]^{1/3}$$

$$\text{Logo o ponto } (k, n) = \left\{ n^2, \left[\frac{2(1 - \rho)}{\rho} \right]^{1/3} \right\} \quad \bar{e}$$

solução para o sistema de equações (I) e (II).

Semelhantemente ao caso anterior, obtém-se o determinante da matriz hessiana no ponto crítico, solução do sistema de equações (I) e (II), verificando-se que é positivo neste ponto, e então $V(\hat{m})$ é mínima para $k = n^2$ e

$$n = \left[\frac{2(1 - \rho)}{\rho} \right]^{1/3}.$$

PIMENTEL GOMES (1988) observa ainda que, quando se fixa "n" *a priori*, temos que a variância é mínima para

$$a) k = \left[\frac{n(1 - \rho)}{\rho} \right]^{1/2}, \quad \text{no caso de meia borda-}$$

dura. Note-se que este resultado é uma generalização para os valores de k apresentados em 2.6.2.1 e 2.6.2.2, sendo obtido da igualdade:

$$\frac{\partial V(\hat{m})}{\partial k} = 0,$$

apresentada em 2.6.5.1.

$$b) k = \left[\frac{2n(1 - \rho)}{\rho} \right]^{1/2}, \quad \text{no caso de bordadura}$$

completa, que generaliza os resultados de k apresentados em 2.6.3.1. e 2.6.3.2., sendo obtido da igualdade:

$$\frac{\partial V(\hat{m})}{\partial k} = 0, \text{ apresentada em 2.6.5.2.}$$

2.6.6. Exemplo ilustrativo 2

Seja $\rho = 0,070$, para o caso de bordadura com pleta.

a) Neste caso:

$$n = \left[\frac{2(1 - \rho)}{\rho} \right]^{1/3} = 2,984 \approx 3$$

$$k = n^2 = 3^2 = 9.$$

b) Se fixássemos antecipadamente, $n = 1, 2, 3$ ou 4:

$$\text{b.1) } n = 1 \implies k = \left[\frac{2(1 - \rho)}{\rho} \right]^{1/2} = 5,15 \approx 5$$

$$\text{b.2) } n = 2 \implies k = \left[\frac{4(1 - \rho)}{\rho} \right]^{1/2} = 7,29, \text{ o}$$

que sugere os números pares mais próximos (6 e 8).

$$\text{b.3) } n = 3 \implies k = \left[\frac{6(1 - \rho)}{\rho} \right]^{1/2} = 8,93 \approx 9$$

$$\text{b.4) } n = 4 \implies k = \left[\frac{8(1 - \rho)}{\rho} \right]^{1/2} = 10,31, \text{ o}$$

que sugere os múltiplos de 4 mais próximos (8 e 12).

Substituindo os pares (k, n) na fórmula de $V(\hat{m})$, de 2.6.5.2., obtem-se os resultados da Tabela 12.

Tabela 12. Variância da média de tratamento $[V(\hat{m})]$ em função de diversos valores do número de linhas úteis (n) por parcela, e número de árvores úteis (k) por parcela, relativa ao exemplo ilustrativo 2.

n	k	$V(\hat{m}) = \frac{\sigma^2}{N} \left[1 + \frac{(k-1)\rho}{k} + \frac{2n}{k} + \frac{2n(k-1)\rho}{k} + \frac{2}{n} + \frac{2(k-1)\rho}{n} + \frac{4}{k} + \frac{4(k-1)\rho}{k} \right]$
1	5	$(\sigma^2/N) \cdot 5,376$
2	6	$(\sigma^2/N) \cdot 4,500$
2	8	$(\sigma^2/N) \cdot 4,470 \longrightarrow$ Variância mínima para $n = 2$
3	9	$(\sigma^2/N) \cdot 4,333 \longrightarrow$ Variância mínima
4	8	$(\sigma^2/N) \cdot 4,470$
4	12	$(\sigma^2/N) \cdot 4,425 \longrightarrow$ Variância mínima para $n = 4$

O autor comenta que, no presente caso, a função $V(\hat{m})$ varia pouco nas vizinhanças de $n = 3$ e $k = 9$, sendo quase indiferente tomar parcelas de seis plantas úteis em duas linhas, ou de nove plantas úteis em três linhas. Mas, parcelas menores, que permitem maior número de repetições para um determinado N , levam a maior número de graus de liberdade para o resíduo A , de modo que o uso de parcelas do tipo $n = 2$ e $k = 6$, poderia compensar o aumento de $V(\hat{m})$, pela estimação mais precisa do erro entre parcelas.

2.6.7. O caso de ρ negativo

PIMENTEL GOMES (1984) mostra que para ρ negativo, a variância $V(\hat{m})$ é função decrescente de k . Nestas condições, o ideal é tomar k o maior possível, obedecendo às duas condições apresentadas em 2.6.2.

Entretanto, uma vez fixado o valor de k , o número de linhas úteis deve ser tão próximo de $(k)^{1/2}$ quanto possível, pois o mínimo de $V(\hat{m})$ se dá para:

$$n = (k)^{1/2};$$

aliás, este resultado já foi demonstrado, pois, fixando k :

$$\frac{\partial V(\hat{m})}{\partial n} = 0 \implies n^2 = k.$$

O autor considera este resultado bastante intuitivo, pois em parcelas de k plantas úteis em n linhas

úteis, o número total de plantas por parcela $\bar{e}$, para o caso de bordadura completa, dado por:

$$K = \left(1 + \frac{2}{n}\right)(2n + k).$$

A relação $g(n)$ entre o número total de árvores e o número de árvores úteis é:

$$g(n) = \frac{rK}{rk} = \frac{K}{k} = \left(1 + \frac{2}{n}\right)\left(1 + \frac{2n}{k}\right).$$

Esta função passa por um mínimo, em $n^2 = k$, ou seja $n = (k)^{1/2}$, pois fixando k :

$$\frac{\partial g(n)}{\partial n} = \frac{2}{k} - \frac{2}{n^2} = \frac{2n^2 - 2k}{n^2 \cdot k}$$

Igualando a zero:

$$2n^2 - 2k = 0 \implies n^2 = k$$

$$n = (k)^{1/2}.$$

Ou seja, para $n = (k)^{1/2}$, a relação entre o número total de árvores e o número de árvores úteis:

$$\left| \frac{cN}{crk} \right|$$

é mínima, resultando em um número de graus de liberdade total o maior possível, para um dado cN , influenciando nos graus de liberdade dos resíduos A e B, aumentando-os.

Por exemplo, sejam dois tratamentos, no delineamento em blocos casualizado, parcelas com bordadura comple

ta. Para $k = 16$ e $cN = 720$, correspondem os seguintes esquemas de análise para $n = 2$ e $n = (16)^{1/2} = 4$, por exemplo, respectivamente.

a)

FONTE DE VARIACÃO	GL
Blocos	8
Tratamentos	1
Resíduo (A)	8
Resíduo (B)	270
Total	287

b)

FONTE DE VARIACÃO	GL
Blocos	9
Tratamentos	1
Resíduo (A)	9
Resíduo (B)	300
Total	319

Para $n = 4$, observa-se que para um dado cN ($cN = 720$), o GL do resíduo (A) e do resíduo (B) aumentam em relação ao caso de $n = 2$, passando de 8 para 9 GL, e de 270 para 300 GL, respectivamente, aumentando a precisão das estimativas de ρ e σ^2 .

Pela fórmula de $V(\hat{m})$, em 2.6.5.2., tem-se que, para k e ρ fixos, $V(\hat{m})$ é mínima para $n = (k)^{1/2}$.

Sejam $k = 16$ e $\rho = -0,05$. Para $n = 2$ e $n = 4$ obtêm-se os seguintes valores para $V(\hat{m})$

$$a) \ n = 2 \implies V(\hat{m}) = 0,6250 \frac{\sigma^2}{N}$$

$$b) \quad n = 4 \quad V(\hat{m}) = 0,5625 \frac{\sigma^2}{N},$$

mínima entre $n = 4$ e $n = 2$, mantendo k e ρ fixos.

O mesmo acontece para o caso de meia bordadura, onde:

$$g(n) = \frac{K}{k} = \frac{\left(1 + \frac{1}{n}\right)(n + k)}{k} = \left(1 + \frac{1}{n}\right)\left(1 + \frac{n}{k}\right),$$

que apresenta mínimo para:

$$n = (k)^{1/2}.$$

3. RESULTADOS E DISCUSSÃO

O tamanho da parcela experimental é um detalhe de grande importância na experimentação, provocando o surgimento de métodos que auxiliam na determinação do tamanho ótimo da parcela experimental.

O método da máxima curvatura, o mais antigo e simples dos métodos abordados, se baseia na relação entre o coeficiente de variação obtido de parcelas de diferentes tamanhos (de um ensaio em branco) e o tamanho da parcela.

Na sua simplicidade, recebeu diversas interpretações no que refere ao termo "máxima curvatura". Uma delas é: maior distância entre um segmento de linha curva e a reta que une as suas extremidades. Já, LESSMAN & ATKINS (1963) consideram ponto de máxima curvatura como sendo o ponto na curva onde a taxa de variação direcional para o CV, por incremento fixo de x , for máximo. MEIER & LESSMAN (1971) observam que este procedimento desconsidera os comprimentos dos arcos, e propuseram, acertadamente, que se considerasse a fórmula da curvatura de uma linha, dada por:

$$K = \frac{y''}{[1 + y'^2]^{3/2}},$$

onde: y' e y'' são as derivadas de primeira e segunda ordens da função que define a linha, dada, neste caso, por:

$$CV_x = \frac{a'}{x^{b'}}.$$

A função K apresenta máximo para

$$\frac{\partial K}{\partial x} = 0,$$

o que equivale a $x_{\text{crítico}}$ dado por:

$$x_{\text{crítico}} = [a'^2 b'^2 (2b' + 1) / (b' + 2)]^{1/(2b' + 2)}$$

Assim, das três interpretações para máxima curvatura, esta última é a correta do ponto de vista algébrico, porém, do ponto de vista prático, o local de máxima curvatura representa apenas o ponto onde a taxa de diminuição do CV tem maior desaceleração, mas nada indica que seja um ponto ótimo. Ocorre que, por exemplo, ao ponto de máxima curvatura pode corresponder ainda um CV alto, que desaconselhe o uso do tamanho obtido por esse método. Mas o $x_{\text{crítico}}$ tem o seu valor, intrínseco ao seu próprio significado, ou seja, é o valor de x a partir do qual o CV diminuirá a taxas inferiores às anteriores, de modo que, este método, na sua terceira interpretação, pode e deve ser usado como auxiliar a um ou mais métodos de determinação do tamanho ótimo de parcelas experimentais, acrescentando-lhes informações.

No que se refere à utilização do $x_{\text{crítico}}$ na

fórmula de $x_{\text{ótimo}}$ proposta por SMITH (1938), das maneiras como propuseram LESSMAN & ATKINS (1963) e MEIER & LESSMAN (1971), depreende-se não serem satisfatórias, não servindo de instrumento corroborativo nem confirmatório do método de SMITH (1938), como admitem LESSMAN & ATKINS (1963), uma vez que no exemplo utilizado no presente trabalho, a substituição não corroborou resultado algum, e, além disto, os próprios LESSMAN & ATKINS (1963) utilizam um $x_{\text{crítico}}$ médio, obtido de duas funções:

$$CV_x = \frac{a'}{x^{b'}} \quad \text{e} \quad CV_x = \frac{a^*}{(1 + \log x)^{b^*}},$$

chegando apenas desta maneira a resultados semelhantes aos obtidos pelo método de SMITH (1938).

O método de SMITH (1938) serve de referência a muitos outros métodos de determinação de tamanho de parcela. Recebe elogios de diversos autores, entre eles PEARCE (1976), que conclui: "Parece, então, que a lei empírica de Fairfield Smith é quase inevitável como uma aproximação em relação a pequenas mudanças no tamanho das parcelas, tal como ocorre na prática".

Este método tem a vantagem de considerar os custos envolvidos no experimento, de modo que o tamanho ótimo, obtido por este método, maximiza a quantidade de informação por unidade de custo.

Alguns autores, por outro lado, consideram di-

fácil a determinação dos fatores de custos K_1 e K_2 , o que dificultaria o uso deste método.

Realmente a presença de K_1 e K_2 dificulta a aplicação deste método, tanto é que muitos autores preferem atribuir valores para a relação K_1/K_2 , sem se preocuparem em individualizá-los, porém, é inegável a necessidade de se considerar os custos, qualquer que seja a atividade desenvolvida, tentando reduzi-los ao mínimo possível.

Espera-se que um economista, ligado a qualquer estação experimental, tenha condições de superar essas dificuldades, obtendo-se, isto sim, vantagens na utilização do método de SMITH (1938) na determinação do tamanho ótimo de parcelas experimentais.

Com a possibilidade da aplicação deste método a dados provenientes de alguns delineamentos experimentais com efeitos de tratamentos, levantada por KOCH & RIGNEY (1951), e não apenas a dados de ensaios em branco, como originalmente proposto por SMITH (1938), este método ganhou um forte argumento adicional a seu favor, pela economia de tempo e dinheiro que esta possibilidade propicia.

Por fim, vale ressaltar que HATHEWAY & WILLIAMS (1958) apresentam procedimentos para a obtenção da estimativa de b , com características de não tendenciosidade e variância mínima assintótica, o que influencia beneficemente na determinação do tamanho ótimo, dado por:

$$x_{\text{ótimo}} = \frac{b}{1-b} \cdot \frac{K_1}{K_2}.$$

O método de HATHEWAY (1961) utiliza-se da fórmula empírica de SMITH (1938) e da fórmula de COCHRAN & COX (1957) que determina o número de repetições necessárias para detectar, a uma certa probabilidade, uma determinada diferença entre médias de tratamentos, percebendo uma relação entre elas, que permitiu unir o coeficiente b de SMITH, à fórmula de COCHRAN & COX (1957), resultando em:

$$x^b = \frac{2(t_1 + t_2)^2 CV_1^2}{rd^2}.$$

Por esta fórmula, se forem especificados níveis de significância, de modo a permitir a obtenção de t_1 e t_2 , o número (r) de repetições que se vai utilizar e a diferença " d " entre médias que se deseja detectar, determina-se o tamanho conveniente de parcela.

Por outro lado, esta fórmula possibilita, ao se fixar " d ", determinar várias combinações de número (r) de repetições e tamanho (x) de parcelas que propiciem a detecção de uma diferença " d " entre médias de tratamento, sem considerar os custos. É válido, também, observar nessas várias combinações (r , x), qual ou quais delas apresentam valores de x próximos do $x_{\text{ótimo}}$ obtido pelo método de SMITH (1938), método este que considera os custos, combinando, desta maneira, os

dois métodos, obtendo-se, então, um tamanho de parcela que considera número de repetições, grau de precisão e custos.

O método de PIMENTEL GOMES (1984) é outra opção para o aproveitamento de dados experimentais na determinação do tamanho ótimo de parcela.

Este método se baseia na minimização da variância da média de tratamentos, provenientes de ensaios com k observações por parcela. Essa variância é função do número (k) de observações e do número (n) de linhas úteis, por parcela, de modo que o tamanho ótimo de parcela, por este método, corresponde à combinação de valores k e n que torna mínima a variância da média de tratamento.

PIMENTEL GOMES (1988) deduziu valores genéricos ótimos para k e n , nos casos de parcelas sem bordadura, com bordadura (completa e meia bordadura), e coeficiente de correlação intraclassa (ρ) positivo ou negativo.

Este método apresenta princípios teóricos simples e objetivos, porém, ao considerar os casos com meia bordadura e ao determinar os números (K e N) de árvores por parcela e por tratamento, pressupõe que todas as parcelas estão circundadas por outras parcelas, o que na prática não ocorre, pois as parcelas limítrofes do experimento não satisfazem essa pressuposição. E, ao se considerar como é bastante comum, blocos separados uns dos outros e as parcelas dispostas em fila, a fórmula de K em função de k e n fica alterada,

assumindo uma caracterização diferente conforme o caso de disposição dos blocos, das parcelas no bloco, e mesmo de acordo com o tratamento considerado.

Assim, ao se expressar K em função de k e n de uma maneira genérica (casos de parcelas com meia bordadura), determinando N através da expressão $N = rK$, onde r é o número de repetições, temos que N/K pode assumir diversas configurações em termos de n e k , conforme o caso, de modo que deve-se proceder um estudo mais detalhado deste aspecto, e sua influência no valor da variância da média de tratamento, observando se ocorre alteração significativa no ponto de ótimo, em termos de número de observações por parcela e número de linhas por parcela.

No que se refere à forma ótima de parcelas, observa-se que os métodos de SMITH (1938) e HATHEWAY (1961) não fornecem informações explícitas a esse respeito, mas é verdade que os dois se utilizam de ensaios em branco, dos quais é possível obter algum subsídio através da observação do coeficiente de variação de cada tamanho e forma de parcela.

No ensaio em branco com arroz, de 2.1.1., observa-se, pela Tabela 2, que o CV , para um determinado tamanho, é sempre menor quando a parcela apresenta maior dimensão no sentido das colunas do ensaio ($L > C$). Por exemplo, o $CV_{10 \times 1}$ é menor que o $CV_{1 \times 10}$, o $CV_{5 \times 1}$ é menor que o $CV_{1 \times 5}$, de modo que, quando se determinou o tamanho ótimo, pelo método de

SMITH (1938), igual a $x_{\text{ótimo}} \approx 7 \text{ m}^2$, é válido pensar em parcelas de dimensões 7 m x 1 m; 4,5 m x 1,6 m ou 3,5 m x 2,0 m, e a escolha da forma ficaria na dependência dos fatores operacionais da área e equipamentos disponíveis e do bom senso do pesquisador.

De um modo geral, os resultados presentes na literatura indicam que parcelas retangulares, com a sua maior dimensão disposta na mesma direção de gradiente de fertilidade do solo, leva a variabilidade menor do que a proporcionada por parcelas quadradas, ou parcelas retangulares perpendiculares às primeiras, considerando-se um determinado tamanho x.

4. CONCLUSÕES

Da presente exposição é possível concluir:

4.1. O método da máxima curvatura modificado determina um $x_{\text{crítico}}$, que pode e deve ser usado como uma informação adicional aos outros métodos de determinação do tamanho e forma ótimos de parcelas experimentais.

4.2. A utilização do $x_{\text{crítico}}$, acima mencionado, na fórmula do $x_{\text{ótimo}}$ proposta por SMITH (1938), das maneiras como propuseram LESSMAN & ATKINS (1963) e MEIER & LESSMAN (1971), deve ser evitada.

4.3. O método de SMITH (1938) apresenta propriedades que o tornam recomendável sempre que se disponha de dados provenientes de ensaio em branco ou de delineamentos experimentais do tipo parcelas subdivididas ou reticulado quadrado.

4.4. A função de custos envolvida neste método deve ser determinada com precisão, pois dela depende, em grande parte, a confiabilidade do tamanho ótimo obtido.

4.5. O método de HATHEWAY (1961) deve ser usado quando se deseja estudar a relação entre tamanho de parcela, número de repetições e precisão desejada. Depende do método de SMITH

(1938), no que se refere ao coeficiente b , da fórmula:

$$V_x = \frac{V_1}{x^b}.$$

4.6. O método de PIMENTEL GOMES (1984) deve ser usado quando se dispõe de dados provenientes de experimento com k observações por parcela. Determina o tamanho e a forma ótimos de parcela.

4.7. Os resultados dos ensaios em branco devem auxiliar na escolha da forma ótima de parcela, nos casos de se utilizar os métodos de SMITH (1938) e HATHEWAY (1961).

4.8. O uso de mais de um método de determinação de tamanho e forma ótimos de parcelas experimentais é recomendável, sempre que couber, de modo que o tamanho de parcela adotado, atenda, na medida do possível, aos diversos fatores considerados em cada método.

REFERÊNCIAS BIBLIOGRÁFICAS

- AFONJA, B. The analysis of an experiment with plots of different sizes. Tropical Agriculture, Trinidad, 44: 39-44, 1967.
- AFONJA, B. Analysis of a uniformity trial on Cassava. Experimental Agriculture, London, 4: 135-41, 1968.
- BRIM, C.A. & MASON, D.D. Estimates of optimum plot size for soybean yield trials. Agronomy Journal, Madison, 51: 331-4, 1959.
- BUENO, A. & PIMENTEL GOMES, F. Estimativa do tamanho de parcelas em experimentos de mandioca. Revista Brasileira de Mandioca, Cruz das Almas, 2(2): 39-44, 1983.
- CHACIN LUGO, F. Tamaño de parcela experimental y su forma. Revista de la Facultad de Agronomía, Maracay, 9(3): 55-74, 1977.
- COCHRAN, W. & COX, G.M. Experimental designs. 2. ed. New York, John Wiley, 1957. 611 p.
- DAGNELIE, P. Principes d'experimentation. Gembloux, Les Presses Agronomiques de Gembloux, 1981. 182 p.

FEDERER, W.T. Experimental design. New York, MacMillan, 1955.
544 p.

GOMEZ, K.A. & GOMEZ, A.A. Statistical procedures for agricultural research. 2 ed. New York, John Wiley, 1984. 680 p.

HATHEWAY, W.H. Convenient plot size. Agronomy Journal, Madison, 53: 279-80, 1961.

HATHEWAY, W.H. & WILLIAMS, E.J. Efficient estimation of the relationship between plot size and the variability of crop yields. Biometrics, Washington, 14(2): 207-22, 1958.

KOCH, E.J. & RIGNEY, H.J. A method of estimating optimum plot size from experimental data. Agronomy Journal, Madison, 43: 17-21, 1951.

LESSMAN, K.J. & ATKINS, R.E. Optimum plot size and relative efficiency of lattice designs for grain sorghum yield tests. Crop Science, Madison, 3(6): 477-81, 1963.

MARANI, A. Estimation of optimum plot size using Smith's procedure. Agronomy Journal, Madison, 55: 503, 1963.

MEIER, V.D. & LESSMAN, K.J. Estimation of optimum field plot shape and size for testing yield in *Crambe abyssinica* Hochst. Crop Science, Madison, 11(5): 648-50, 1971.

MEYER, P.L. Probabilidade: aplicações à estatística. Trad. de R. de C.B.L.F. Rio de Janeiro, Livros Técnicos e Científicos, 1976. 391 p.

- MOOD, A.McFarlane; GRAYBILL, F.A.; BOES, D.C. Introduction to the theory of statistics. 3.ed. Tokyo, McGraw-Hill, 1974. 564 p.
- PEARCE, S.C. An examination of Fairfield Smith's law of environmental variation. J. Agric. Science, Cambridge, 87(1): 21-4, 1976.
- PIMENTEL GOMES, F. O problema de tamanho das parcelas em experimentos com plantas arbóreas. Pesquisa Agropecuária Brasileira, Brasília, 19(12): 1507-12, 1984.
- PIMENTEL GOMES, F. Novos aspectos do problema do tamanho ótimo das parcelas em experimentos com plantas arbóreas. Pesquisa Agropecuária Brasileira, Brasília, 23(1): 59-62, 1988.
- ROSSETTI, A.G. Determinação do tamanho ótimo de parcelas em ensaios agrícolas. Piracicaba, 1979. 70 p. [Mestrado - Escola Superior de Agricultura "Luiz de Queiroz"/USP].
- SATTERTHWAITE, F.E. An approximate distribution of estimates of variance components. Biometrics, Washington, 2: 110-4, 1946.
- SEARLE, S.R. Linear models. New York, John Wiley, 1971. 532 p.
- SILVA, E.C. Estudo do tamanho e forma de parcelas para experimentos com soja. Piracicaba, 1972. 71 p. [Mestrado - Escola Superior de Agricultura "Luiz de Queiroz"/USP].
- SMITH, H.F. An empirical law describing heterogeneity in the yields of agricultural crops. Journal of Agricultural Science, Cambridge, 28: 1-23, 1938.

APÊNDICE

APÊNDICE

De acordo com a teoria da "propagação do erro" (MOOD *et alii*, 1974, p. 181; MEYER, 1976, p. 148), pode-se mostrar que a matriz (W^{-1}) dos pesos de V_x , é obtida da inversa da matriz de variâncias e covariâncias das V' .

Seja, então, a função:

$$h^* = \log V'_x.$$

Expandindo h^* em série de Taylor, no ponto $E(V'_x) = \mu_x$ vem:

$$h^* = \log \mu_x + \frac{1}{\mu_x} (V'_x - \mu_x) + R.$$

$$\text{Sejam } h^*_{x_i} = \log V'_{x_i} \text{ e } h^*_{x_j} = \log V'_{x_j},$$

então $h^*_{x_i}$ e $h^*_{x_j}$ serão aproximados pela expansão em série de Taylor pelas seguintes equações:

$$h^*_{x_i} = \log \mu_{x_i} + \frac{1}{\mu_{x_i}} (V'_{x_i} - \mu_{x_i}) + R$$

$$\approx \log \mu_{x_i} + \frac{1}{\mu_{x_i}} (V'_{x_i} - \mu_{x_i})$$

$$h^*_{x_j} \approx \log \mu_{x_j} + \frac{1}{\mu_{x_j}} (V'_{x_j} - \mu_{x_j})$$

Assim:

$$\text{COV}(h_{x_i}^*, h_{x_j}^*) = \text{COV}(\log V_{x_i}', \log V_{x_j}')$$

$$\begin{aligned} &\approx \text{COV} \left[\log \mu_{x_i} + \frac{(V_{x_i}' - \mu_{x_i})}{\mu_{x_i}}, \log \mu_{x_j} + \frac{(V_{x_j}' - \mu_{x_j})}{\mu_{x_j}} \right] \\ &\approx \text{COV} \left[\frac{V_{x_i}' - \mu_{x_i}}{\mu_{x_i}}, \frac{V_{x_j}' - \mu_{x_j}}{\mu_{x_j}} \right] \\ &\approx \frac{1}{\mu_{x_i} \mu_{x_j}} \text{COV}(V_{x_i}', V_{x_j}') \end{aligned}$$

Em termos de estimativas, temos:

$$\widehat{\text{COV}}(\log V_{x_i}', \log V_{x_j}') \approx \frac{1}{V_{x_i}' V_{x_j}'} \widehat{\text{COV}}(V_{x_i}', V_{x_j}')$$

Da expressão anterior pode-se escrever genericamente, em termos matriciais, que:

$$W^* = \widehat{V}(\log V') \approx \begin{bmatrix} \frac{1}{V_{x_1}'} & 0 & \dots & 0 \\ 0 & \frac{1}{V_{x_2}'} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{V_{x_n}'} \end{bmatrix} \widehat{V}(V')$$

$$\approx A B A.$$

Mas ABA é invertível, $\exists A^{-1}$, B^{-1} e $(ABA)^{-1}$;

e então:

$$W^{*-1} \approx [ABA]^{-1} = A^{-1}B^{-1}A^{-1} =$$

$$\approx \begin{bmatrix} V'_{x_1} & 0 & \dots & 0 \\ 0 & V'_{x_2} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & V'_{x_n} \end{bmatrix} [\hat{V}(V')]^{-1} \begin{bmatrix} V'_{x_1} & 0 & \dots & 0 \\ 0 & V'_{x_2} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & V'_{x_n} \end{bmatrix},$$

onde cada termo de $[\hat{V}(V')]^{-1}$ é multiplicado pelas V' correspondentes à linha e à coluna que ele ocupa. Seja, por exemplo, o elemento v_{23} de $[\hat{V}(V')]^{-1}$, então $w_{23}^{*-1} \approx v_{23} V'_{x_2} V'_{x_3}$. Genêricamente tem-se $w_{ij}^{*-1} \approx v_{ij} V'_{x_i} V'_{x_j}$ = pesos dos $\log V'$.

Seja agora a função:

$$h_{x_i} = V_{x_i} = \frac{V'_{x_i}}{x_i} \implies h'_{x_i} = \log V_{x_i} = \log \frac{V'_{x_i}}{x_i} = \log V'_{x_i} - \log x_i$$

$$\text{COV}(\log V_{x_i}, \log V_{x_j}) = \text{COV}(h'_{x_i}, h'_{x_j}) =$$

$$= \text{COV} \left[\log \frac{V'_{x_i}}{x_i}, \log \frac{V'_{x_j}}{x_j} \right]$$

$$= \text{COV} [\log V'_{x_i} - \log x_i, \log V'_{x_j} - \log x_j]$$

$$= \text{COV} [\log V'_{x_i}, \log V'_{x_j}]$$

Logo:

$$\hat{V}(\log V'_x) = \hat{V}(\log V_x)$$

$$W^* = W \implies W^{*-1} = W^{-1}$$