

Tamanho de amostra e relações lineares de caracteres morfológicos e produtivos de crambe

Sample size and linear relations of the morphological characters and production of crambe

Alberto Cargnelutti Filho^{I*} Marcos Toebe^{II} Tatiani Reis da Silveira^{II} Gabriele Casarotto^{III}
Fernando Machado Haesbaert^{III} Sidinei José Lopes^I

RESUMO

Os objetivos deste trabalho foram determinar o tamanho de amostra para a estimação da média de caracteres morfológicos e produtivos de crambe (*Crambe abyssinica* Hochst) e avaliar as relações entre esses caracteres. Foi conduzido um experimento a campo em Santa Maria, Estado do Rio Grande do Sul (latitude 29°42'S, longitude 53°49'W e 95m de altitude), no ano de 2008. Na colheita, foram selecionadas, aleatoriamente, 75 plantas e foram mensurados os caracteres morfológicos (altura de planta, número de ráculos por planta e comprimento médio dos ráculos por planta) e os produtivos (número de frutos por planta, número de frutos por ráculo e produtividade de grãos). Foi calculado o tamanho de amostra de cada caráter e investigada a relação entre os caracteres por meio de diagramas de dispersão, análises de correlação, de trilha e regressão. Na cultura de crambe, para estimar a média com uma mesma precisão, o tamanho de amostra (número de plantas) dos caracteres produtivos é maior que o dos morfológicos. O número de frutos por planta tem relação linear positiva com a produtividade de grãos e pode ser utilizado para seleção indireta.

Palavras-chave: *Crambe abyssinica* Hochst, planejamento experimental, dimensionamento de amostra, modelagem, seleção indireta.

ABSTRACT

The aim of this research was to determine the sample size to estimate the average of the morphological and productive traits of crambe (*Crambe abyssinica* Hochst) and assess the relationships among these traits. An experiment was conducted in Santa Maria, Rio Grande do Sul State, Brazil (latitude 29°42'S, longitude 53°49'W and 95m altitude) in the

year 2008. At harvest, 75 plants were randomly selected and measured for morphological traits (plant height, number of racemes per plant and length average of racemes per plant) and productivity (number of fruits per plant, number of fruits per raceme and grain yield). It was calculated the sample size of each trait and studied the relationship among the traits by scatter plots, correlation, path analysis and regression. In the culture of crambe, the sample size (number of plants) of productive traits is greater than morphology to estimate average and it has the same precision. The number of fruits per plant has a positive linear relationship with grain yield and can be used for indirect selection.

Key words: *Crambe abyssinica* Hochst, experimental planning, sample dimensioning, modeling, indirect selection.

INTRODUÇÃO

Em experimentos agrícolas, a mensuração em todas as plantas (indivíduos) da área útil da unidade experimental é a técnica adequada para estimar a média do caráter em avaliação. Entretanto, a disponibilidade de tempo, mão-de-obra, recursos financeiros e humanos, muitas vezes, limita a avaliação de grande número de plantas, principalmente, quando muitos caracteres devem ser mensurados. Nesses casos, a amostragem de plantas dentro da unidade experimental é uma alternativa adequada.

O tamanho de amostra é diretamente proporcional à variabilidade dos dados e ao grau de

^IDepartamento de Fitotecnia, Centro de Ciências Rurais (CCR), Universidade Federal de Santa Maria (UFSM), 97105-900, Santa Maria, RS, Brasil. E-mail: cargnelutti@pq.cnpq.br. *Autor para correspondência.

^{II}Programa de Pós-graduação em Agronomia, UFSM, Santa Maria, RS, Brasil.

^{III}Curso de Agronomia, UFSM, Santa Maria, RS, Brasil.

confiança desejado na estimativa e inversamente proporcional ao erro de estimação permitido, fixado *a priori* pelo pesquisador (FONSECA & MARTINS, 1995; BARBETTA et al., 2004; BUSSAB & MORETTIN, 2004; SPIEGEL et al., 2004). Em relação à estimativa da média de caracteres morfológicos e produtivos de crambe (*Crambe abyssinica* Hochst), não há um valor que possa ser definido como erro de estimação máximo tolerável. Porém, deve ser considerada a possibilidade de obter uma estimativa da média com a máxima confiabilidade possível.

Em um programa de melhoramento de plantas, o estudo das relações entre os caracteres é importante, a fim de identificar caracteres que possam ser utilizados na seleção indireta. No caso das relações existentes serem lineares, as análises de correlação, de trilha e de regressão linear multivariada possibilitam a obtenção de informações substanciais quanto à relação entre os caracteres.

Na cultura do crambe, planejamentos experimentais quanto ao tamanho e forma de parcela (MEIER & LESSMAN, 1971) e relações entre caracteres (LESSMAN, 1975; MASTEBROEK et al., 1994) têm sido pesquisados. No entanto, informações sobre o tamanho de amostra e sobre relações entre caracteres morfológicos e produtivos de crambe ainda são escassos. Assim, os objetivos deste trabalho foram determinar o tamanho de amostra para a estimação da média de caracteres morfológicos e produtivos de crambe (*Crambe abyssinica* Hochst) e avaliar as relações entre esses caracteres.

MATERIAL E MÉTODOS

Foi realizado um experimento com a cultura de crambe (*Crambe abyssinica* Hochst) na Universidade Federal de Santa Maria, Santa Maria, Rio Grande do Sul (latitude 29°42'S, longitude 53°49'W e 95m de altitude), de julho a outubro de 2008. No mês de julho de 2008, foi efetuada a adubação de base na área experimental com 35kg ha⁻¹ de N, 140kg ha⁻¹ de P₂O₅ e 140kg ha⁻¹ de K₂O. Em seguida, foi feita a semeadura em 60 linhas de 20m de comprimento e foi utilizado o espaçamento de 0,058m entre plantas na linha e 0,17m entre linhas, o que resulta em 100 plantas m⁻².

Na área experimental, foram selecionadas, aleatoriamente, 75 plantas. Na colheita dessas plantas, foram mensurados os caracteres morfológicos: altura de planta em cm (AP), número de ráceros por planta

(NRP) e comprimento médio dos ráceros por planta em cm (CMRP). Também foram avaliados os caracteres produtivos: número de frutos por planta (NFP) e número de frutos por rácermo (NFR), obtidos dividindo-se NFP por NRP e produtividade de grãos, em gramas planta⁻¹ (PROD). Em relação aos dados de AP, NRP, CMRP, NFP, NFR e PROD, foram calculadas medidas de tendência central, de variabilidade, de assimetria e curtose e verificada a normalidade, por meio do teste de *Lilliefors* (CAMPOS, 1983).

Para cada caráter, tomando-se por base as 75 plantas, foi calculado o tamanho de amostra (η) para uma semi-amplitude do intervalo de confiança (erro de estimação) igual a 1, 2, ... e 40% da estimativa da média (m), com grau de confiança $(1-\alpha)$ de 95%, por meio da

$$\text{expressão } \eta = \frac{t_{\alpha/2}^2 \cdot s^2}{(\text{erro de estimação})^2} \text{ (FONSECA \& MARTINS$$

1995; BARBETTA et al., 2004; BUSSAB & MORETTIN, 2004; SPIEGEL et al., 2004), na qual $t_{\alpha/2}$ é o valor crítico da distribuição *t* de *Student*, cuja área à direita é igual a $\alpha/2$, isto é, o valor de t , tal que $P(t > t_{\alpha/2}) = \alpha/2$, com $(n-1)$ graus de liberdade, com $\alpha=5\%$ de probabilidade de erro, e s^2 é a estimativa de variância.

Em seguida, fixando-se η igual a 75 plantas, que foi o tamanho de amostra utilizado na amostragem, foi calculado o erro de estimação em percentagem da estimativa da média (m) para cada um dos caracteres,

$$\text{por meio da expressão } \text{erro de estimação} = 100 \frac{t_{\alpha/2} \cdot s}{\sqrt{\eta} \cdot m}, \text{ em}$$

que s é a estimativa do desvio padrão.

A fim de visualizar as relações entre os caracteres AP, NRP, CMRP, NFP, NFR e PROD, foi feito o diagrama de dispersão. Após, foi calculada a matriz de coeficientes de correlação linear de Pearson (r) entre os caracteres e por meio do teste *t* de *Student*, a 5% de probabilidade de erro, foi verificada a significância do r . A seguir, foi realizado o diagnóstico de multicolinearidade nessa matriz (CRUZ, 2006), conforme critério de MONTGOMERY & PECK (1982). Depois disso, foi realizada a análise de trilha (*path analysis*) das variáveis explicativas (AP, NRP, CMRP e NFP) sobre a variável PROD e, em seguida, ajustado um modelo de regressão linear múltipla da variável dependente (PROD) em função das independentes (AP, NRP, CMRP e NFP). Os parâmetros do modelo foram estimados pelo método passo-a-passo (*stepwise*), com 5% de significância para entrada e saída de variáveis

(CRUZ, 2006). O número de frutos por rácemo (NFR) foi excluído das análises de trilha e de regressão por causar multicolinearidade (MONTGOMERY & PECK, 1982). As análises estatísticas foram realizadas com o auxílio do programa GENES (CRUZ, 2006) e do aplicativo Office Excel®.

RESULTADOS E DISCUSSÃO

De maneira geral, as medidas de tendência central, de variabilidade, de assimetria e curtose e o teste de Lilliefors, em relação aos caracteres morfológicos: altura de planta (AP), número de ramos por planta (NRP) e comprimento médio dos ramos por planta (CMRP), e aos produtivos: número de frutos por planta (NFP), número de frutos por rácemo (NFR) e produtividade de grãos (PROD), revelaram boa aderência dos dados, das 75 plantas de crambe (*Crambe abyssinica* Hochst), à distribuição normal (Tabela 1). Particularmente, os caracteres NFP e PROD apresentaram um leve afastamento da normalidade, caracterizado, principalmente, pela assimetria positiva da distribuição dos dados (Tabela 1 e Figura 1). No entanto, de acordo com o teorema limite central, mesmo que a população básica seja não normal, a distribuição da média amostral será aproximadamente normal para amostras superiores a 30 observações (FONSECA & MARTINS, 1995; BUSSAB & MORETTIN, 2004). Então, diante dessas considerações, em relação à

normalidade, pode-se inferir que os dados desses caracteres oferecem credibilidade ao estudo do dimensionamento de amostra.

A magnitude do coeficiente de variação (CV) oscilou entre 19,54% para a altura de planta e 62,29% para o número de frutos por planta. De maneira geral, o CV dos caracteres produtivos (média dos três caracteres = 56,77%), foi, aproximadamente, 2,4 vezes maior que o CV dos caracteres morfológicos (média dos três caracteres = 23,78%). Isso sugere que para a obtenção da estimativa da média, com uma determinada precisão, o tamanho de amostra de caracteres produtivos é maior em relação aos morfológicos. Ainda, a variabilidade existente é importante e confere adequabilidade ao estudo das relações entre os caracteres, por meio das análises de correlação, de trilha e de regressão.

O tamanho de amostra (número de plantas) para a estimação da média de cada caráter, com semi-amplitude do intervalo de confiança igual a 1% da estimativa da média (m) e grau de confiança de 95%, oscilou entre 1.516 plantas, para o caráter altura de planta; a 15.405 plantas para o caráter número de frutos por planta (Tabela 2). Portanto, em relação altura de planta, pode-se inferir, com 95% de confiança, que o intervalo de confiança da média (μ), obtido com 1.516 plantas, é de $\mu \pm 0,01m$. Em outro extremo, a precisão de $\mu \pm 0,01m$ é obtida com 15.405 plantas, para o número de frutos por planta. Esses resultados confirmam a variabilidade existente entre os caracteres. A

Tabela 1 - Mínimo, máximo, média, mediana, desvio-padrão, erro-padrão da média, coeficiente de variação (CV), assimetria, curtose e teste de Lilliefors da altura de planta (AP), em cm; do número de ramos por planta (NRP), do comprimento médio dos ramos por planta (CMRP), em cm; do número de frutos por planta (NFP), do número de frutos por rácemo (NFR) e da produtividade de grãos (PROD), em gramas planta⁻¹ de 75 plantas de crambe (*Crambe abyssinica* Hochst).

Estatística	AP	NRP	CMRP	NFP	NFR	PROD
Mínimo	31,00000	4,00000	7,90000	26,00000	3,25000	0,13670
Máximo	92,30000	19,00000	30,48571	535,00000	29,87500	3,41640
Média	65,69467	11,17333	19,96338	186,32000	15,85630	1,18392
Mediana	65,50000	11,00000	19,36923	173,00000	15,90909	1,08870
Desvio-padrão	12,83450	2,79175	5,35658	116,05825	7,30529	0,73352
Erro-padrão	1,48200	0,32236	0,61852	13,40125	0,84354	0,08470
CV (%)	19,53659	24,98584	26,83203	62,28974	46,07185	61,95716
Assimetria ⁽¹⁾	-0,16476 ns	0,18039 ns	-0,01002 ns	0,98010 *	0,25064 ns	0,92429 *
Curtose + 3 ⁽²⁾	2,71832 ns	3,25082 ns	2,09478 ns	3,64077 ns	2,11606 ns	3,48343 ns
Lilliefors ⁽³⁾	0,06456 ns	0,10357 *	0,07374 ns	0,09928 ns	0,09527 ns	0,11887 *

⁽¹⁾ * Assimetria difere de zero pelo teste t, em nível de 5% de probabilidade. ns = não-significativo.

⁽²⁾ * Curtose difere de três pelo teste t, em nível de 5% de probabilidade. ns = não-significativo.

⁽³⁾ * Significativo a 5% de probabilidade. ns = não-significativo.

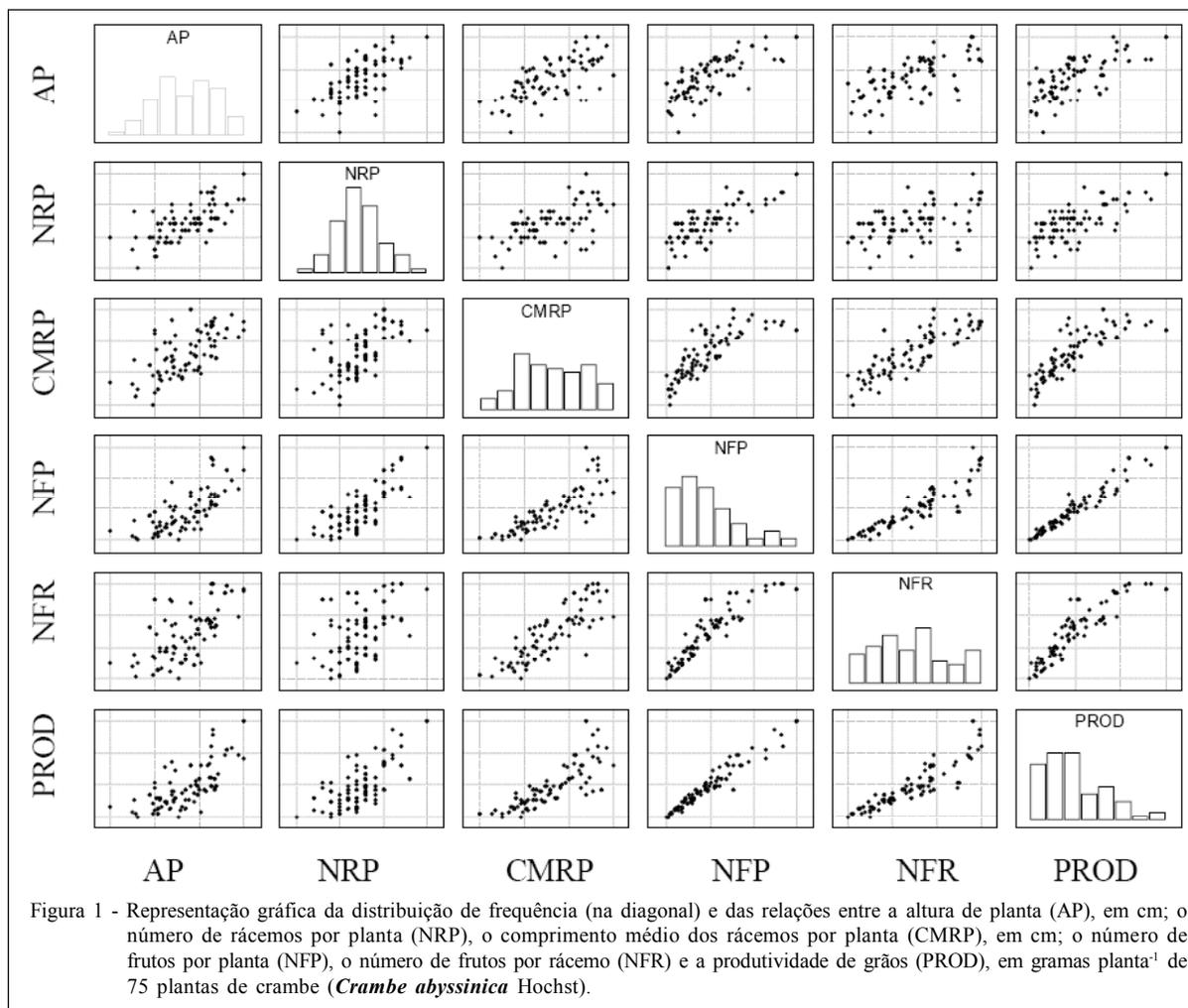


Figura 1 - Representação gráfica da distribuição de frequência (na diagonal) e das relações entre a altura de planta (AP), em cm; o número de ráculos por planta (NRP), o comprimento médio dos ráculos por planta (CMRP), em cm; o número de frutos por planta (NFP), o número de frutos por ráculo (NFR) e a produtividade de grãos (PROD), em gramas planta⁻¹ de 75 plantas de crambe (*Crambe abyssinica* Hochst).

variabilidade de tamanho de amostra entre genótipos, caracteres e experimentos foi constatada em milho (STORCK et al., 2007), em feijão (CARGNELUTTI FILHO et al., 2008), em soja (CARGNELUTTI FILHO et al., 2009) e em mamona (CARGNELUTTI FILHO et al., 2010).

Do ponto de vista prático, pode-se inferir que é difícil obter estimativas da média de caracteres morfológicos e produtivos de crambe, com erro de estimação de 1%, em função do elevado número de plantas a serem mensuradas. Por exemplo, para o caráter altura de planta, que apresenta menor variabilidade, seriam necessárias 1.516 plantas. Assim, menores tamanhos de amostra foram determinados com base em semi-amplitudes do intervalo de confiança iguais a 2, 3, ... e 40% da média (m) (Tabela 2). Não cabe aqui o julgamento do erro de estimação máximo aceitável, ficando isso a critério do pesquisador que usufruir dessas informações para o seu planejamento

experimental. Caso a opção seja amostrar 75 plantas, o erro de estimação seria de $\pm 4,49$, $\pm 5,75$, $\pm 6,17$, $\pm 14,33$, $\pm 10,60$ e $\pm 14,26\%$ da estimativa da média (m), respectivamente, para os caracteres AP, NRP, CMRP, NFP, NFR e PROD. Portanto, os resultados evidenciam que o tamanho de amostra de caracteres produtivos é maior que o de morfológicos. Resultados semelhantes foram obtidos na cultura da mamona, por CARGNELUTTI FILHO et al. (2010), que concluíram que há aumento do tamanho de amostra (número de plantas) para estimar caracteres de semente, de plântula, de planta adulta e de produção, nessa ordem.

O diagnóstico de multicolinearidade na matriz de coeficientes de correlação linear de Pearson, entre as variáveis explicativas altura de planta (AP), número de ráculos por planta (NRP), comprimento médio dos ráculos por planta (CMRP), número de frutos por planta (NFP) e número de frutos por ráculo (NFR), revelou número de condição (NC) igual a 228, o

Tabela 2 - Tamanho de amostra (número de plantas) para a estimação da média dos caracteres altura de planta (AP), número de ráculos por planta (NRP), comprimento médio dos ráculos por planta (CMRP), número de frutos por planta (NFP), número de frutos por ráculo (NFR) e produtividade de grãos (PROD) de crambe (*Crambe abyssinica* Hochst), para uma semi-amplitude do intervalo com 95% de confiança e erros iguais a 1, 2, ..., 40% da média.

Erro (%)	AP	NRP	CMRP	NFP	NFR	PROD
1	1516	2479	2859	15405	8428	15241
2	379	620	715	3852	2107	3811
3	169	276	318	1712	937	1694
4	95	155	179	963	527	953
5	61	100	115	617	338	610
6	43	69	80	428	235	424
7	31	51	59	315	172	312
8	24	39	45	241	132	239
9	19	31	36	191	105	189
10	16	25	29	155	85	153
11	13	21	24	128	70	126
12	11	18	20	107	59	106
13	9	15	17	92	50	91
14	8	13	15	79	43	78
15	7	12	13	69	38	68
16	6	10	12	61	33	60
17	6	9	10	54	30	53
18	5	8	9	48	27	48
19	5	7	8	43	24	43
20	4	7	8	39	22	39
21	4	6	7	35	20	35
22	4	6	6	32	18	32
23	3	5	6	30	16	29
24	3	5	5	27	15	27
25	3	4	5	25	14	25
26	3	4	5	23	13	23
27	3	4	4	22	12	21
28	2	4	4	20	11	20
29	2	3	4	19	11	19
30	2	3	4	18	10	17
31	2	3	3	17	9	16
32	2	3	3	16	9	15
33	2	3	3	15	8	14
34	2	3	3	14	8	14
35	2	3	3	13	7	13
36	2	2	3	12	7	12
37	2	2	3	12	7	12
38	2	2	2	11	6	11
39	1	2	2	11	6	11
40	1	2	2	10	6	10

que classifica a matriz como de colinearidade moderada a forte, conforme critério de MONTGOMERY & PECK (1982). Com a retirada do NFR, o NC reduziu para 24 e, assim, as análises de trilha e de regressão linear múltipla da produtividade de grãos (PROD) em função das

variáveis explicativas (AP, NRP, CMRP e NFP) podem ser realizadas adequadamente (CRUZ, 2006) (Tabela 3).

A produtividade de grãos (PROD) apresentou correlação linear positiva ($r=0,96523$, $P\leq 0,05$) e efeito direto (0,97729) com mesmo sinal e de magnitude semelhante com o número de frutos por planta (NFP), explicando a verdadeira associação existente. Assim, pode-se inferir que as plantas com maior número de frutos são aquelas com maior produtividade de grãos. Em menor magnitude, a PROD apresentou associação linear com a AP ($r=0,69596$, $P\leq 0,05$), o NRP ($r=0,69368$, $P\leq 0,05$) e o CMRP ($r=0,81797$, $P\leq 0,05$). No entanto, os efeitos diretos da AP (-0,00787), do NRP (-0,08449) e do CMRP (0,07073) sobre a PROD foram desprezíveis e, portanto, a associação existente é explicada pelos elevados efeitos indiretos via NFP (AP=0,71348; NRP=0,74314 e CMRP=0,80049) (Tabela 3). Esses resultados sugerem que a seleção indireta para a PROD deve ser realizada com base no NFP.

A estimativa da equação de regressão linear múltipla, obtida por meio do método *stepwise*, da PROD foi função somente do NFP (PROD=0,047262+0,006101 NFP) e apresentou coeficiente de determinação (R^2) de 0,9317. A não inclusão das variáveis explicativas AP, NRP e CMRP, nesse modelo, é explicada pelos efeitos diretos desprezíveis e elevados efeitos indiretos via NFP sobre a PROD (Tabela 3). Então, diante desses resultados, confirma-se a relação linear positiva existente entre o NFP e a PROD e, portanto, o NFP é um caráter adequado para a seleção indireta quanto a PROD. Do ponto de vista prático, é possível selecionar as plantas diretamente na área experimental, por meio da contagem do número de frutos, sem a necessidade de destruir as plantas para a colheita.

CONCLUSÃO

Na cultura de crambe (*Crambe abyssinica* Hochst), para estimar a média com uma mesma precisão, o tamanho de amostra (número de plantas) dos caracteres produtivos deve ser maior do que o tamanho dos morfológicos. O número de frutos por planta tem relação linear positiva com a produtividade de grãos e pode ser utilizado para seleção indireta.

AGRADECIMENTOS

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelas bolsas concedidas.

Tabela 3 - Estimativas dos coeficientes de correlação linear de Pearson e dos efeitos diretos e indiretos dos caracteres altura de planta (AP), número de ráceros por planta (NRP), comprimento médio dos ráceros por planta (CMRP) e número de frutos por planta (NFP) sobre a produtividade de grãos (PROD) de 75 plantas de crambe (*Crambe abyssinica* Hochst).

Efeito	AP	NRP	CMRP	NFP
Direto sobre PROD	-0,00787	-0,08449	0,07073	0,97729
Indireto via AP		-0,00522	-0,00516	-0,00574
Indireto via NRP	-0,05606		-0,04809	-0,06425
Indireto via CMRP	0,04639	0,04025		0,05793
Indireto via NFP	0,71348	0,74314	0,80049	
Total (Correlação de Pearson)	0,69596*	0,69368*	0,81797*	0,96523*
Coefficiente de determinação	0,93708			
Efeito da variável residual	0,25083			

* Significativo a 5% de probabilidade de erro pelo teste t, com 73 graus de liberdade.

REFERÊNCIAS

- BARBETTA, P.A. et al. **Estatística para cursos de engenharia e informática**. São Paulo: Atlas, 2004. 410p.
- BUSSAB, W.O.; MORETTIN, P.A. **Estatística básica**. 5.ed. São Paulo: Saraiva, 2004. 526p.
- CAMPOS, H. De. **Estatística experimental não-paramétrica**. 4.ed. Piracicaba: Departamento de Matemática e Estatística - ESALQ, 1983. 349p.
- CARGNELUTTI FILHO, A. et al. Tamanho de amostra de caracteres de genótipos de soja. **Ciência Rural**, v.39, p.983-991, 2009. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-84782009000400005&lng=pt&nrm=iso>. Acesso em: 18 mar. 2010. doi: 10.1590/S0103-84782009005000016.
- CARGNELUTTI FILHO, A. et al. Tamanho de amostra de caracteres de cultivares de feijão. **Ciência Rural**, v.38, p.635-642, 2008. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-84782008000300007&lng=pt&nrm=iso>. Acesso em: 18 mar. 2010. doi: 10.1590/S0103-84782008000300007.
- CARGNELUTTI FILHO, A. et al. Tamanho de amostra de caracteres em híbridos de mamoneira. **Ciência Rural**, v.40, p.250-257, 2010. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-84782010000200005&lng=pt&nrm=iso>. Acesso em: 18 mar. 2010.
- CRUZ, C.D. **Programa genes: estatística experimental e matrizes**. Viçosa: UFV, 2006. 285p.
- FONSECA, J.S.; MARTINS, GA. **Curso de estatística**. 5.ed. São Paulo: Atlas, 1995. 317p.
- LESSMAN, K.J. Variation in crambe, *Crambe abyssinica* Hochst. **Journal of the American Oil Chemists' Society**, v.52, p.386-389, 1975.
- MASTEBROEK, H.D. et al. Variation for agronomic characteristics in crambe (*Crambe abyssinica* Hochst. ex Fries). **Industrial Crops and Products**, v.2, p.129-136, 1994.
- MEIER, V.D.; LESSMAN, K.J. Estimation of optimum field plot shape and size for testing yield in *Crambe abyssinica* Hochst. **Crop Science**, v.11, p.648-650, 1971.
- MONTGOMERY, D.C.; PECK, E.A. **Introduction to linear regression analysis**. New York: John Wiley & Sons, 1982. 504p.
- SPIEGEL, R.A. et al. **Probabilidade e estatística**. 2.ed. Porto Alegre: Bookman, 2004. 398p.
- STORCK, L. et al. Sample size for single, double and triple hybrid com ear traits. **Scientia Agrícola**, v.64, p. 30-35, 2007. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-90162007000100005&lng=pt&nrm=iso>. Acesso em: 18 mar. 2010. doi: 10.1590/S0103-90162007000100005.