

# Tamanho de amostra para a estimação do coeficiente de correlação linear de Pearson entre caracteres de mamoneira

## Sample size to estimate the Pearson correlation coefficient among characters of castor bean

Alberto Cargnelutti Filho<sup>1\*</sup>; Sidinei José Lopes<sup>1</sup>; Betânia Brum<sup>2</sup>;  
Marcos Toebe<sup>3</sup>; Tatiani Reis da Silveira<sup>3</sup>; Gabriele Casarotto<sup>3</sup>

### Resumo

No estudo de relações lineares é importante dimensionar adequadamente a amostra para a estimação do coeficiente de correlação linear de Pearson entre pares de caracteres, com confiabilidade aceitável. O objetivo deste trabalho foi determinar o tamanho de amostra (número de plantas) para a estimação do coeficiente de correlação linear de Pearson entre 21 caracteres de mamoneira. Foram avaliadas 41 e 55 plantas dos híbridos Sara e Lyra, respectivamente, quanto aos caracteres de semente, de plântula, de planta adulta e de produção, em Santa Maria, Estado do Rio Grande do Sul, no ano agrícola de 2007/2008. Calculou-se o tamanho de amostra por meio de reamostragem “bootstrap”, para a estimação do coeficiente de correlação linear de Pearson, para as amplitudes do intervalo de confiança de “bootstrap” de 95% iguais a 0,20, 0,35 e 0,50, para cada um dos 210 pares de caracteres. Independentemente do híbrido de mamoneira, 96 plantas são suficientes para a estimação do coeficiente de correlação linear de Pearson, para a amplitude do intervalo de confiança de “bootstrap” de 95%, máxima de 0,52.

**Palavras-chave:** *Ricinus communis* L., planejamento experimental, reamostragem “bootstrap”, dimensionamento de amostra, precisão experimental

### Abstract

In the study of linear relationships, it is important to define correctly the sample size, to estimate the Pearson correlation coefficient among pairs of characters, with acceptable reliability. The aim of this research was to determine the sample size (number of plants) to estimate the Pearson correlation coefficient among 21 characters of castor bean. It was evaluated 41 and 55 plants of the Sara and Lyra hybrids, respectively, regarding to the characters of seed, seedling, adult plant and yield in Santa Maria, Rio Grande do Sul State, Brazil, in the agriculture year of 2007/2008. It was calculated the sample size through bootstrap resampling, to estimate the Pearson correlation coefficient, for the amplitudes of the bootstrap confidence interval 95% equal to 0.20, 0.35 and 0.50, for each of 210 pairs of characters. Regardless of the castor bean hybrids, 96 plants are sufficient to estimate the Pearson correlation coefficient, to the amplitude of the bootstrap confidence interval 95%, maximum of the 0.52.

**Key words:** *Ricinus communis* L., experimental planning, bootstrap resampling, sample dimensioning, experimental precision

<sup>1</sup> Engenheiros Agrônomos, Profs. Drs. do Deptº de Fitotecnia/CCR, Universidade Federal de Santa Maria, UFSM, Santa Maria, RS. E-mail: cargnelutti@pq.cnpq.br; sjlopes@pq.cnpq.br

<sup>2</sup> Engº Agrº, Doutoranda do Programa de Pós-Graduação em Agronomia da UFSM. E-mail: bbufsm@gmail.com

<sup>3</sup> Engenheiros Agrônomos, Discentes de Mestrado do Programa de Pós-Graduação em Agronomia da UFSM. E-mail: m.toebe@gmail.com; tatianisilveira@yahoo.com.br; gabrielecasarotto@gmail.com

\* Autor para correspondência

## Introdução

Nas Ciências Agrárias e, mais especificamente, em programas de melhoramento de plantas, o estudo de relações lineares entre os caracteres por meio do coeficiente de correlação linear de Pearson e por procedimentos complementares, tais como correlações parciais, análise de trilha e correlação canônica (CRUZ; REGAZZI, 1997; CRUZ; CARNEIRO, 2003; HAIR et al., 2005), pode fornecer resultados importantes, especialmente na identificação de caracteres para seleção indireta.

O coeficiente de correlação linear de Pearson ( $r$ ) é uma estatística utilizada para medir a força, a intensidade ou o grau de relação linear entre duas variáveis aleatórias (STEVENSON, 2001; BARBETTA; REIS; BORNIA, 2004; KAZMIER, 2007; FERREIRA, 2009). O sinal do coeficiente de correlação linear de Pearson expressa o sentido da correlação e a intensidade é representada por um valor numérico que oscila entre  $-1$  e  $1$ . Em situações extremas, dois caracteres podem apresentar correlação linear negativa perfeita ( $r = -1$ ) ou positiva perfeita ( $r = 1$ ), ou ainda ausência de relação linear ( $r = 0$ ).

Quando o tamanho de amostra é pequeno, o valor do coeficiente de correlação linear de Pearson, para ser significativo precisa ter elevada magnitude (próximo de  $|1|$ ). Neste caso, embora o valor do coeficiente apresente significância estatística, a amostra pode não ser representativa da população (STEVENSON, 2001; HAIR et al., 2005) e, conseqüentemente, o coeficiente também pode não representar a verdadeira relação entre os caracteres. Por outro lado, quando o tamanho de amostra for excessivamente grande, um pequeno valor do coeficiente de correlação linear de Pearson (próximo de zero) pode ser considerado significativo. Porém não necessariamente seja uma relação importante entre os caracteres do ponto de vista prático (STEVENSON, 2001; KAZMIER, 2007; HAIR et al., 2005).

Estudos do dimensionamento de amostra têm

sido realizados para a estimação da média de caracteres nas culturas de milho (STORCK et al., 2007), feijão (CARGNELUTTI FILHO et al., 2008), soja (CARGNELUTTI FILHO et al., 2009) e mamona (CARGNELUTTI FILHO et al., 2010a) e, de maneira geral, têm evidenciado variabilidade do tamanho de amostra entre genótipos, caracteres e experimentos. A qualidade fisiológica de sementes de mamona (*Ricinus communis* L.) também tem sido estudada (MACHADO et al., 2010).

No entanto, estudos sobre o dimensionamento de amostra para a estimação do coeficiente de correlação linear de Pearson entre pares de caracteres de mamoneira (*Ricinus communis* L.), não foram encontrados na literatura. Assim, o objetivo deste trabalho foi determinar o tamanho de amostra (número de plantas) para a estimação do coeficiente de correlação linear de Pearson, entre 21 caracteres de híbridos de mamoneira.

## Material e Métodos

Dois híbridos de mamoneira, Sara e Lyra, foram avaliados em um campo experimental de Santa Maria, Rio Grande do Sul (latitude  $29^{\circ}42'S$ , longitude  $53^{\circ}49'W$  e  $95m$  de altitude), no ano agrícola 2007/2008. Foram avaliadas, individualmente, 90 sementes do híbrido Sara e 90 sementes do híbrido Lyra quanto aos caracteres de semente (Tabela 1). As sementes foram identificadas com números de um a 180 e semeadas em área experimental de  $15m$  de comprimento x  $16m$  de largura, em 19 de outubro de 2007, conforme a recomendação para o cultivo de mamoneira (WREGÉ et al., 2007), juntamente com as demais sementes (não identificadas) que compuseram as seis parcelas (três parcelas do híbrido Sara e três parcelas do híbrido Lyra). O espaçamento foi de  $0,6m$  entre plantas e  $1,2m$  entre linhas. Entre as sementes identificadas, emergiram 41 e 55 plantas dos híbridos Sara e Lyra, respectivamente, as quais foram avaliadas individualmente quanto aos caracteres de plântula, de planta adulta e de produção (Tabela 1).

**Tabela 1.** Sigla, unidade e forma de coleta de caracteres de semente, de plântula, de planta adulta e de produção de híbridos de mamoneira (*Ricinus communis* L).

Caractere	Sigla	Unidade	Forma de coleta
----- Caracteres de semente -----			
Peso	PESO	g	Pesagem individual realizada em balança analítica de precisão de 0,0001g.
Comprimento	COMP	mm	Medido com paquímetro digital da carúncula até a outra extremidade da semente, no sentido da rafe.
Largura	LARG	mm	Medida com paquímetro digital na parte central da semente, perpendicular à medida do comprimento.
Área	AREA	mm <sup>2</sup>	Comprimento x largura
Espessura	ESPE	mm	Medida com paquímetro digital na região mediana da semente.
----- Caracteres de plântula -----			
Índice SPAD aos sete dias após a emergência	SP7	-	Medida direta com clorofilômetro. Média de três medidas realizada em três pontos da plântula.
Índice SPAD aos 14 dias após a emergência	SP14	-	Medida direta com clorofilômetro. Média de três medidas realizada em três pontos da plântula.
----- Caracteres de planta adulta -----			
Comprimento do caule com folhas verdadeiras	CC	cm	No aparecimento da 1ª inflorescência. Medida da distância entre a inserção da 1ª folha verdadeira e a inserção da última folha verdadeira.
Diâmetro do caule a 10 cm do solo	DC	cm	No aparecimento da 1ª inflorescência, mediu-se com paquímetro digital o diâmetro do caule a 10 cm do solo.
Área Foliar média – método comprimento e largura	AF	cm <sup>2</sup>	No aparecimento da 1ª inflorescência, mediu-se o máximo comprimento e largura da folha.
Altura de inserção do primeiro ráculo	APR	cm	No aparecimento da 1ª inflorescência. Distância entre o nível do solo até o ponto de inserção do 1º ráculo, com régua graduada em milímetros.
Altura do ráculo mais alto	AR	cm	No fim do florescimento. Distância entre o nível do solo e a extremidade do ráculo mais alto da planta.
----- Caracteres de produção -----			
Comprimento médio dos ráculos por planta	CRP	cm	Após a colheita, foi medido o comprimento do ráculo com presença de cápsulas, em todos os ráculos da planta e obteve-se a média.
Número de ráculos por planta	NRP	-	Contagem realizada no momento da colheita.
Número médio de cápsulas por ráculo	NCR	-	Após a colheita, contou-se o número de cápsulas em todos os cachos da planta, para obtenção do número médio de cápsulas por ráculo por planta.
Peso médio de cápsulas por ráculo	PCR	g	Após a colheita e retirada das cápsulas, pesaram-se as cápsulas de cada planta para obtenção da média em balança analítica de precisão de 0,0001 g.
Peso total de cápsula por planta	PCP	g	Após a colheita e retirada das cápsulas, pesaram-se todas as cápsulas de cada planta em balança analítica de precisão de 0,0001g.

continua

continuação

Peso médio de grãos por rácemo	PGR	g	Após a debulha das cápsulas, os grãos de todos os ráceros da planta foram pesados em balança analítica de precisão de 0,0001g.
Produtividade de grãos por planta	PGP	g	Após a debulha das cápsulas, foi pesada a produção de grãos por planta em balança analítica de precisão de 0,0001g.
Epicarpo (casca)	EPI	g	Obtida pela diferença entre o peso de cápsulas da planta e o peso de grãos.
Massa de cem grãos	MCG	g	Média obtida pela pesagem de duas repetições de 100 sementes de cada planta em balança analítica de precisão de 0,0001g. Umidade de 9%.

**Fonte:** Elaboração dos autores.

Inicialmente, com os dados individuais e em conjunto dos híbridos Sara e Lyra, foi investigada a linearidade entre os caracteres por meio de diagramas de dispersão (HAIR et al., 2005) e verificada a normalidade dos dados de cada caractere, por meio do teste de Kolmogorov-Smirnov (CAMPOS, 1983). Após, foram estimadas três matrizes de coeficientes de correlação linear de Pearson ( $r$ ) entre os 21 caracteres avaliados. As matrizes foram obtidas a partir das 41 plantas do híbrido Sara (matriz 1), das 55 plantas do híbrido Lyra (matriz 2) e das 96 plantas dos dois híbridos (matriz 3). Afim de verificar se as inferências poderiam ser realizadas com a matriz 3, foi estimado o coeficiente de correlação de Pearson entre cada par de matrizes (matriz 1 vs matriz 2, matriz 1 vs matriz 3 e matriz 2 vs matriz 3) e sua significância foi verificada pelo teste  $t$  de Student, a 5% de probabilidade, com  $n-2=208$  graus de liberdade, em que  $n$  são os 210 pares de caracteres (combinação de 21 caracteres, tomados dois a dois). Quanto mais próximo de 1,0 for a estimativa obtida, maior concordância haverá entre as matrizes.

Tomando-se por base os dados da matriz formada por 96 plantas e 21 caracteres, foram definidos 691 tamanhos de amostra (10, 11, 12, ..., 700 plantas). Após, para cada tamanho de amostra, foram realizadas 1.000 simulações, por meio de reamostragem “bootstrap”, com reposição. Para cada amostra simulada foi estimado o coeficiente de correlação linear de Pearson ( $r$ ) para cada um dos 210 pares de caracteres. Assim, para cada tamanho

de amostra, de cada par de caracteres, foram obtidas 1.000 estimativas de “bootstrap” de  $r$  (FERREIRA, 2009) e determinados os valores dos percentis 2,5% e 97,5%. Depois, calculou-se a amplitude do intervalo de confiança de “bootstrap” de 95% pela diferença entre o percentil 97,5% e o percentil 2,5%.

Para cada um dos 210 pares de caracteres, determinou-se o tamanho de amostra ( $n$  = número de plantas) para a estimação do coeficiente de correlação linear de Pearson ( $r$ ), para as amplitudes do intervalo de confiança de “bootstrap” de 95% iguais a 0,20, 0,35 e 0,50. Para essas determinações, partiu-se do tamanho inicial (10 plantas), e considerou-se como tamanho de amostra ( $n$ ), o número de plantas a partir do qual as amplitudes do intervalo de confiança de “bootstrap” de 95% foram menores ou iguais a 0,20, 0,35 e 0,50, respectivamente.

Construiu-se um diagrama de dispersão para visualizar as relações entre as 210 estimativas de  $r$ , obtidas a partir das 96 plantas, e o tamanho de amostra ( $n$ ) para as amplitudes do intervalo de confiança de “bootstrap” de 95% iguais a 0,20, 0,35 e 0,50. As análises estatísticas foram realizadas no programa R (R DEVELOPMENT CORE TEAM, 2011) e na planilha eletrônica Office Excel.

## Resultados e Discussão

Examinando-se os diagramas de dispersão entre os 21 caracteres, feitos com as 41 plantas do híbrido Sara, as 55 plantas do híbrido Lyra e as 96 plantas dos dois híbridos, não foram constatados padrões

não-lineares nos dados. Ainda, em apenas três casos (comprimento médio dos ramos por planta do híbrido Sara, número de ramos por planta do híbrido Lyra e número de ramos por planta das 96 plantas) o teste de Kolmogorov-Smirnov (CAMPOS, 1983) revelou não ajuste à distribuição normal ( $p \leq 0,05$ ) (cálculos não apresentados). Então, diante dessas considerações, e ainda do fato da determinação do tamanho de amostra, por meio de reamostragem “bootstrap”, ser independente da distribuição de probabilidade dos dados, pode-se inferir que esse conjunto de dados oferece credibilidade ao estudo proposto.

De maneira geral, os coeficientes de correlação linear de Pearson para cada par de caracteres, foram similares entre os híbridos Sara e Lyra de mamoneira,

o que mostra semelhança nas tendências, diferindo quanto à magnitude das associações. Sustentam essa inferência, o fato de que o coeficiente de correlação de Pearson ( $r$ ) entre os pares de matrizes obtidas a partir das 41 plantas do híbrido Sara, das 55 plantas do híbrido Lyra e das 96 plantas dos dois híbridos foram significativos e de alta magnitude ( $r \geq 0,81$ ,  $p \leq 0,05$ ). Portanto, a boa concordância existente entre essas matrizes possibilita que as inferências, de um modo geral, possam ser realizadas a partir das 96 plantas dos dois híbridos (Tabela 2), embora esses híbridos tenham sido considerados contrastantes quanto aos caracteres relacionados à plântula, à planta adulta e à produção (BRUM, 2009; CARGNELUTTI FILHO et al., 2010a), mas semelhantes quanto aos caracteres de semente (CARGNELUTTI FILHO et al., 2010a).

**Tabela 2.** Matriz de coeficientes de correlação linear de Pearson<sup>(1)</sup> entre 21 caracteres dos híbridos Sara e Lyra de mamoneira (*Ricinus communis* L.), obtida a partir de 96 observações.

Caractere <sup>(2)</sup>	COMP	LARG	AREA	ESPE	SP7	SP14	CC	DC	AF	APR
PESO	0,83	0,80	0,86	0,82	0,00	0,20	0,05	0,03	0,05	0,27
COMP		0,78	0,95	0,79	0,18	0,36	0,08	0,12	0,10	0,27
LARG			0,93	0,85	0,18	0,31	0,01	0,08	0,05	0,19
AREA				0,87	0,19	0,36	0,05	0,11	0,08	0,24
ESPE					0,13	0,29	0,07	0,15	0,07	0,23
SP7						0,61	0,03	0,29	0,27	-0,15
SP14							0,00	0,30	0,21	-0,04
CC								0,64	0,66	0,80
DC									0,77	0,37
AF										0,43
	AR	CRP	NRP	NCR	PCR	PCP	PGR	PGP	EPI	MCG
PESO	0,16	-0,14	0,13	-0,13	-0,02	0,10	-0,04	0,11	0,05	-0,03
COMP	0,09	-0,14	0,03	-0,05	0,00	0,11	-0,04	0,11	0,10	0,02
LARG	0,07	-0,06	0,12	-0,07	0,01	-0,01	-0,01	0,01	-0,05	0,02
AREA	0,08	-0,10	0,07	-0,06	0,00	0,06	-0,03	0,06	0,03	0,02
ESPE	0,10	-0,08	0,11	-0,02	0,04	0,08	0,04	0,09	0,05	0,13
SP7	-0,24	0,29	-0,35	0,22	-0,02	-0,15	0,02	-0,15	-0,15	0,08
SP14	-0,19	0,18	-0,32	0,18	0,03	-0,03	0,06	0,00	-0,09	0,15
CC	0,14	0,30	-0,18	0,35	0,12	0,16	0,09	0,15	0,18	0,21
DC	-0,28	0,55	-0,39	0,62	0,22	0,04	0,22	0,02	0,07	0,39
AF	-0,20	0,54	-0,42	0,60	0,17	0,02	0,17	0,03	0,02	0,27
	AR	CRP	NRP	NCR	PCR	PCP	PGR	PGP	EPI	MCG

continua



continuação

APR	0,38	0,00	0,08	0,07	0,07	0,16	0,06	0,16	0,15	0,04
AR		-0,37	0,33	-0,35	-0,17	0,02	-0,20	0,05	-0,04	-0,10
CRP			-0,51	0,80	0,03	-0,01	0,03	0,00	-0,01	0,14
NRP				-0,43	0,04	0,10	-0,03	0,08	0,12	-0,22
NCR					0,13	-0,01	0,10	-0,01	-0,01	0,34
PCR						0,24	0,89	0,17	0,39	0,02
PCP							0,00	0,99	0,94	-0,11
PGR								-0,06	0,15	0,05
PGP									0,88	-0,09
EPI										-0,14

<sup>(1)</sup> Valor superior a |0,21| é significativo a 5% de probabilidade de erro, pelo teste t, com 94 graus de liberdade. <sup>(2)</sup> Siglas definidas na tabela 1.

**Fonte:** Elaboração dos autores.

Com base nas 96 plantas avaliadas, entre os 210 pares de caracteres, houve associação linear negativa em 11 pares ( $r \leq -0,22$ ,  $p \leq 0,05$ ) e positiva em 50 pares ( $r \geq 0,22$ ,  $p \leq 0,05$ ). Nos demais 149 pares de caracteres (71%) não houve associação linear ( $-0,20 \leq r \leq 0,21$ ,  $p > 0,05$ ) (Tabela 2). O elevado número de plantas mensuradas (96 plantas) sugere que as estimativas de  $r$  apresentam elevada precisão e, com isso, associações lineares de baixa magnitude ( $r > |0,21|$ ) são significantes. É importante examinar, além da significância estatística, a magnitude do  $r$ , que fornece uma significância prática de determinada associação linear (HAIR et al., 2005). Portanto, esses resultados revelam que é importante determinar um tamanho de amostra para a estimação do  $r$ , com uma determinada precisão e, após inferir em relação à sua significância estatística.

O tamanho de amostra (número de plantas) para a estimação do coeficiente de correlação linear de Pearson, para a amplitude do intervalo de confiança de “bootstrap” de 95% igual a 0,20, oscilou entre 10 plantas para três pares de caracteres [comprimento da semente (COMP) e área da semente (AREA) ( $r = 0,95$ ), peso total de cápsula por planta (PCP) e produtividade de grãos por planta (PGP) ( $r = 0,99$ ) e peso total de cápsula por planta (PCP) e peso do epicarpo (EPI) ( $r = 0,94$ )], e 661 plantas para o par largura de semente (LARG) e peso do epicarpo (EPI) ( $r = -0,05$ ) (Tabela 3). Variabilidade

de tamanho de amostra para a estimação da média entre genótipos, caracteres e experimentos, também 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., 2010a).

A média dos coeficientes de correlação linear de Pearson entre os pares de caracteres de semente (10 pares), de plântula (1 par), de planta adulta (10 pares) e de produção (36 pares) foi, respectivamente, 0,85, 0,61, 0,37 e 0,14, e a média do tamanho de amostra para a estimação do coeficiente de correlação linear de Pearson, para a amplitude do intervalo de confiança de “bootstrap” de 95% igual a 0,20 foi, respectivamente, 44, 133, 198 e 300 plantas (Tabelas 2 e 3). Esses resultados revelam, de maneira geral, que para estudos de relações lineares, específicos por grupo de caracteres, o tamanho de amostra é crescente na seguinte ordem: caracteres de semente, de plântula, de planta adulta e de produção. Portanto, caracteres que permanecem por mais tempo sob atuação do ambiente, necessitam maior tamanho de amostra. Essa mesma ordem de acréscimo foi observada em relação ao tamanho de amostra para a estimação da média de 38 caracteres dos híbridos Sara e Lyra de mamoneira (CARGNELUTTI FILHO et al., 2010a).

Para os 210 pares de caracteres estudados, 661 plantas de mamoneira são suficientes para a

estimação do coeficiente de correlação linear de Pearson, para a amplitude do intervalo de confiança de “bootstrap” de 95% igual a 0,20 (Tabela 3). Do ponto de vista prático, a avaliação de 661 plantas é difícil, o que sugere que as estimativas de  $r$  sejam obtidas com menor número de plantas e com amplitude do intervalo de confiança de “bootstrap” de 95% superior a 0,20. Assim, menores tamanhos de

amostra foram estimados com base em amplitudes do intervalo de confiança de “bootstrap” iguais a 0,35 e 0,50 (Figura 1). Não cabe aqui o julgamento do erro de estimação (amplitude do intervalo de confiança) máximo aceitável, ficando isso a critério do pesquisador que usufruir dessas informações para o seu planejamento experimental.

**Tabela 3.** Tamanho de amostra (número de plantas) para a estimação do coeficiente de correlação linear de Pearson ( $r$ ), entre 21 caracteres dos híbridos Sara e Lyra de mamoneira (*Ricinus communis* L.), para a amplitude do intervalo de confiança de “bootstrap” de 95% igual a 0,20.

Caractere <sup>(1)</sup>	COMP	LARG	AREA	ESPE	SP7	SP14	CC	DC	AF	APR
PESO	61	47	41	43	453	341	343	338	345	285
COMP		61	10	64	344	302	353	302	291	334
LARG			11	49	314	321	310	265	228	337
AREA				47	329	279	334	275	264	343
ESPE					314	337	323	291	351	341
SP7						133	234	246	322	231
SP14							261	308	273	285
CC								123	94	62
DC									163	254
AF										179
	AR	CRP	NRP	NCR	PCR	PCP	PGR	PGP	EPI	MCG
PESO	349	328	297	331	359	390	312	343	437	323
COMP	416	242	313	328	297	336	296	337	353	348
LARG	414	282	242	313	302	457	275	392	661	399
AREA	419	276	288	328	309	392	275	358	515	348
ESPE	362	309	288	378	327	413	330	374	560	351
SP7	262	273	144	293	339	244	382	242	304	356
SP14	270	281	295	298	297	397	339	349	414	239
CC	395	276	286	212	385	340	331	341	340	348
DC	184	182	172	140	355	379	355	379	340	243
AF	199	163	166	149	394	336	361	336	319	382
	AR	CRP	NRP	NCR	PCR	PCP	PGR	PGP	EPI	MCG
APR	326	327	434	241	325	361	304	377	356	396
AR		245	218	271	335	270	296	274	229	352
CRP			153	52	364	295	391	287	310	346
NRP				129	273	587	308	560	570	311
NCR					346	327	361	340	297	277
PCR						280	37	309	251	447
PCP							312	10	10	399
PGR								297	336	406
PGP									22	385
EPI										416

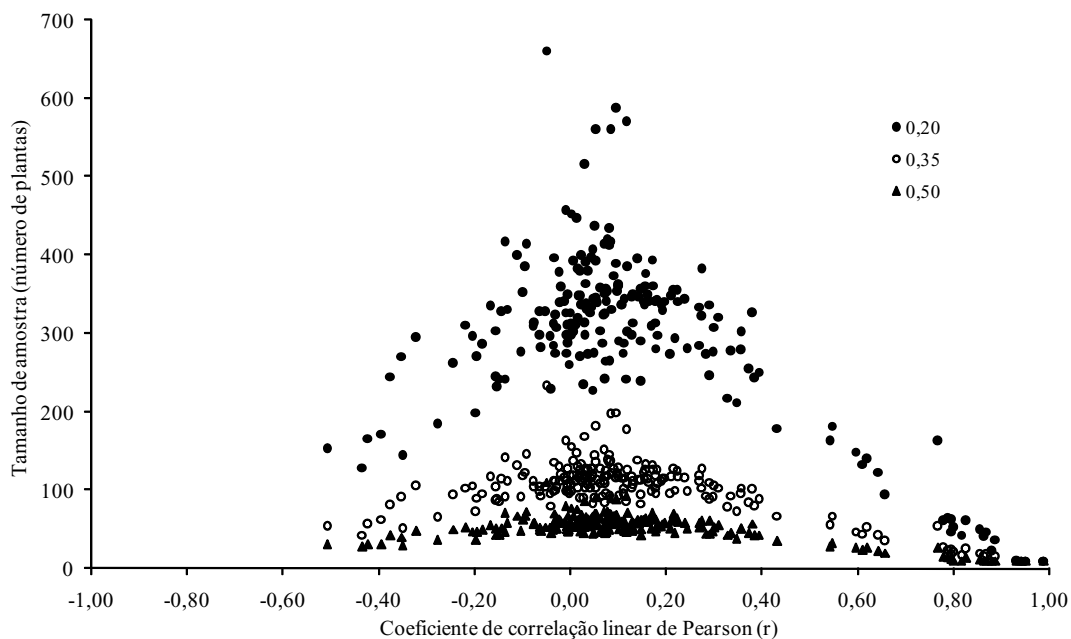
<sup>(1)</sup> Siglas definidas na tabela 1.

**Fonte:** Elaboração dos autores.

Independentemente da amplitude do intervalo de confiança de “bootstrap” (0,20, 0,35 ou 0,50) para uma mesma precisão, há aumento do tamanho de amostra (número de plantas) com o decréscimo da magnitude da associação linear entre caracteres. Esses resultados revelam que para estudos de relações lineares entre caracteres, com base no mesmo número de observações para todos os caracteres, estimativas mais precisas do coeficiente

de correlação de Pearson são obtidas para pares de caracteres com maior intensidade de associação linear e vice-versa. Portanto, esses resultados reforçam os resultados discutidos em Cargnelutti Filho et al. (2010b) de que, especialmente, para os caracteres com menor intensidade de associação linear, a insuficiência amostral associada a essa menor precisão poderia ser causa de resultados divergentes em publicações científicas.

**Figura 1.** Diagrama de dispersão entre as 210 estimativas do coeficiente de correlação linear de Pearson ( $r$ ), obtido a partir das 96 plantas dos híbridos Sara e Lyra de mamoneira (*Ricinus communis* L.) e do tamanho de amostra (número de plantas), para as amplitudes do intervalo de confiança de “bootstrap” de 95% iguais a 0,20, 0,35 e 0,50.



Fonte: Elaboração dos autores.

Tomando-se por base as 96 plantas avaliadas nesse estudo, a amplitude do intervalo de confiança de “bootstrap” de 95% do coeficiente de correlação linear de Pearson oscilou entre 0,01 para o par de caracteres peso total de cápsula por planta (PCP) e produtividade de grãos por planta (PGP) ( $r = 0,99$ ) e 0,52 para o par largura de semente (LARG) e peso do epicarpo (EPI) ( $r = -0,05$ ) (Tabela 4). Portanto, os resultados evidenciam que, do ponto de vista de técnicas experimentais, os pesquisadores da cultura

de mamoneira devem considerar os caracteres, ao dimensionar o tamanho de amostra para avaliação mais precisa das relações lineares entre os mesmos. Ainda, pode-se inferir que é difícil obter estimativas do coeficiente de correlação linear de Pearson com elevada precisão, especialmente para pares de caracteres com baixa intensidade de relação linear, em função do elevado número de plantas que devem ser mensuradas.



**Tabela 4.** Amplitude do intervalo de confiança de “bootstrap” de 95% do coeficiente de correlação linear de Pearson (r), entre 21 caracteres dos híbridos Sara e Lyra de mamoneira (*Ricinus communis* L.), com 96 observações.

Caractere <sup>(1)</sup>	COMP	LARG	AREA	ESPE	SP7	SP14	CC	DC	AF	APR
PESO	0,17	0,14	0,13	0,13	0,46	0,40	0,40	0,38	0,38	0,39
COMP		0,16	0,04	0,16	0,40	0,38	0,42	0,36	0,35	0,40
LARG			0,05	0,13	0,38	0,38	0,38	0,33	0,33	0,41
AREA				0,13	0,41	0,36	0,40	0,35	0,34	0,40
ESPE					0,37	0,39	0,42	0,37	0,40	0,40
SP7						0,24	0,35	0,34	0,39	0,34
SP14							0,36	0,37	0,36	0,36
CC								0,24	0,20	0,16
DC									0,26	0,35
AF										0,31
	AR	CRP	NRP	NCR	PCR	PCP	PGR	PGP	EPI	MCG
PESO	0,39	0,40	0,37	0,41	0,38	0,40	0,40	0,40	0,46	0,37
COMP	0,42	0,35	0,39	0,39	0,36	0,38	0,35	0,38	0,39	0,40
LARG	0,44	0,36	0,32	0,39	0,38	0,45	0,37	0,43	0,52	0,45
AREA	0,43	0,37	0,35	0,38	0,36	0,42	0,36	0,40	0,47	0,40
ESPE	0,42	0,39	0,36	0,43	0,38	0,41	0,40	0,40	0,46	0,38
SP7	0,36	0,36	0,26	0,38	0,40	0,35	0,39	0,32	0,38	0,39
SP14	0,35	0,35	0,36	0,38	0,38	0,43	0,40	0,40	0,44	0,33
CC	0,44	0,35	0,35	0,32	0,45	0,39	0,40	0,39	0,42	0,40
DC	0,29	0,28	0,28	0,25	0,43	0,41	0,42	0,40	0,40	0,34
AF	0,30	0,26	0,27	0,26	0,44	0,38	0,42	0,39	0,37	0,42
	AR	CRP	NRP	NCR	PCR	PCP	PGR	PGP	EPI	MCG
APR	0,39	0,40	0,44	0,34	0,40	0,40	0,40	0,41	0,41	0,44
AR		0,32	0,31	0,34	0,39	0,34	0,36	0,35	0,33	0,43
CRP			0,26	0,14	0,41	0,39	0,42	0,39	0,37	0,40
NRP				0,24	0,36	0,50	0,38	0,49	0,49	0,38
NCR					0,41	0,38	0,41	0,39	0,36	0,36
PCR						0,36	0,11	0,37	0,33	0,46
PCP							0,38	0,01	0,05	0,44
PGR								0,37	0,39	0,44
PGP									0,09	0,43
EPI										0,44

<sup>(1)</sup> Siglas definidas na tabela 1.

Fonte: Elaboração dos autores.

## Conclusão

Independentemente do híbrido (Sara e Lyra) de mamoneira, para os 210 pares de caracteres estudados, 96 plantas são suficientes para a estimação do coeficiente de correlação linear de Pearson, para a amplitude do intervalo de confiança de “bootstrap” de 95%, máxima de 0,52.

## Agradecimentos

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico e a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, pelas bolsas concedidas.

## Referências

- BARBETTA, P. A.; REIS, M. M.; BORNIA, A. C. *Estatística para cursos de engenharia e informática*. São Paulo: Atlas, 2004. 410 p.
- BRUM, B. *Relações entre variáveis de: sementes, plântulas, plantas, produção e óleo em mamoneira*. 2009. Dissertação (Mestrado em Agronomia) – Curso de Pós-graduação em Agronomia. Universidade Federal de Santa Maria, Santa Maria.
- CAMPOS, H. de. *Estatística experimental não-paramétrica*. 4. ed. Piracicaba: Departamento de Matemática e Estatística, ESALQ, 1983. 349 p.
- CARGNELUTTI FILHO, A.; EVANGELISTA, D. H. R.; GONÇALVES, E. C. P.; STORCK, L. Tamanho de amostra de caracteres de genótipos de soja. *Ciência Rural*, Santa Maria, v. 39, n. 4, p. 983-991, 2009.
- CARGNELUTTI FILHO, A.; LOPES, S. J.; BRUM, B.; SILVEIRA, T. R. da; TOEBE, M.; STORCK, L. Tamanho de amostra de caracteres em híbridos de mamoneira. *Ciência Rural*, Santa Maria, v. 40, n. 2, p. 250-257, 2010a.
- CARGNELUTTI FILHO, A.; RIBEIRO, N. D.; STORCK, L.; JOST, E.; POERSCH, N. L. Tamanho de amostra de caracteres de cultivares de feijão. *Ciência Rural*, Santa Maria, v. 38, n. 3, p. 635-642, 2008.
- CARGNELUTTI FILHO, A.; TOEBE, M.; BURIN, C.; SILVEIRA, T. R. da; CASAROTTO, G. Tamanho de amostra para estimação do coeficiente de correlação linear de Pearson entre caracteres de milho. *Pesquisa Agropecuária Brasileira*, Brasília, v. 45, n. 12, p. 1363-1371, 2010b.
- CRUZ, C. D.; CARNEIRO, P. C. S. *Modelos biométricos aplicados ao melhoramento genético*. Viçosa: UFV, 2003. 585 p.
- CRUZ, C. D.; REGAZZI, A. J. *Modelos biométricos aplicados ao melhoramento genético*. 2. ed. Viçosa: UFV, 1997. 390 p.
- FERREIRA, D. F. *Estatística básica*. 2. ed. Lavras: UFLA, 2009. 664 p.
- HAIR, J. F.; ANDERSON, R. E.; TATHAM, R. L.; BLACK, W. C. *Análise multivariada de dados*. 5. ed. Porto Alegre: Bookman, 2005. 593 p.
- KAZMIER, L. J. *Estatística aplicada à administração e economia*. 4. ed. Porto Alegre: Bookman, 2007. 392 p.
- MACHADO, C. G.; MARTINS, C. C.; CRUZ, S. C. S.; NAKAGAWA, J.; PEREIRA, F. R. S. Posição do racemo e do fruto na qualidade fisiológica de sementes de mamona durante o armazenamento. *Semina: Ciências Agrárias*, Londrina, v. 31, n. 2, p. 301-312, 2010.
- R DEVELOPMENT CORE TEAM. *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing, 2011. Disponível em: <<http://www.R-project.org>>. Acesso em: 08 jun. 2011.
- STEVENSON, W. J. *Estatística aplicada à administração*. São Paulo: Harbra, 2001. 495 p.
- STORCK, L.; LOPES, S. J.; CARGNELUTTI FILHO, A.; MARTINI, L. F. D.; CARVALHO, M. P. de. Sample size for single, double and triple hybrid corn ear traits. *Scientia Agricola*, Piracicaba, v. 64, n. 1, p. 30-35, 2007.
- WREGE, M. S.; SILVA, S. D. A.; STEINMETZ, S.; REISSER JUNIOR, C.; HERTER, F. G.; GARRASTAZU, M. C.; MATZENAUER, R. *Zoneamento agroclimático para mamona no Rio Grande do Sul*. Pelotas: Embrapa Clima Temperado, 2007. 30 p. (Embrapa Clima Temperado. Documentos, 192).