

Tamanho de amostra para estimação do coeficiente de correlação linear de Pearson entre caracteres de milho

Alberto Cargnelutti Filho⁽¹⁾, Marcos Toebe⁽¹⁾, Cláudia Burin⁽¹⁾, Tatiani Reis da Silveira⁽¹⁾ e Gabriele Casarotto⁽¹⁾

⁽¹⁾Universidade Federal de Santa Maria, Departamento de Fitotecnia, Avenida Roraima, nº1.000, Bairro Camobi, CEP 97105-900 Santa Maria, RS. E-mail: cargnelutti@pq.cnpq.br, m.toebe@gmail.com, clauburin@yahoo.com.br, tatianisilveira@yahoo.com.br, gabrielecasarotto@yahoo.com.br

Resumo – O objetivo deste trabalho foi determinar o tamanho de amostra para a estimação do coeficiente de correlação linear de Pearson entre caracteres de três híbridos de milho. Para as análises, foram tomadas aleatoriamente 361, 373 e 416 plantas, respectivamente, de híbridos simples, triplo e duplo. Para cada planta, os seguintes caracteres foram mensurados: diâmetro maior e menor do colmo, altura da planta e altura, peso, comprimento e diâmetro da espiga, número de fileiras por espiga, peso e diâmetro de sabugo, massa de cem grãos, número de grãos por espiga, comprimento e produtividade de grãos. Para cada um dos 91 pares de caracteres e híbridos, foi determinado o tamanho de amostra a partir de "bootstrap", com reposição de 1.000 amostras, de cada tamanho de amostra simulado. Na estimação do coeficiente de correlação linear de Pearson com a mesma precisão, o tamanho de amostra (número de plantas) aumenta na direção de pares de caracteres com menor intensidade de relação linear, independentemente do tipo de híbrido. Para os 91 pares de caracteres estudados, 252 plantas são suficientes para a estimação do coeficiente de correlação linear de Pearson, no intervalo de confiança de "bootstrap" de 95%, igual a 0,30.

Termos para indexação: *Zea mays*, planejamento experimental, reamostragem "bootstrap", relação linear.

Sample size for estimating the Pearson correlation coefficient among corn characters

Abstract – The objective of this work was to determine the sample size for estimating the Pearson correlation coefficient among characters from three corn hybrids. For the analyses, 361, 373 and 416 plants were taken randomly from the hybrids single, triple and double, respectively. For each plant, the following characters were measured: major and minor stem diameters, plant and, heights, weight, length and diameter of ear, number of rows per ear, weight and diameter of cob, a hundred-grain weight, number of grain per ear, grain length and grain yield. For each of the 91 pairs of characters and hybrids, the sample size, based on bootstrap with replacement of 1,000 samples, of each simulated sample size, was determined. In the estimation of Pearson correlation coefficient with the same precision, the sample size (number of plants) increases in the direction of character pairs with lower intensity of the linear relationship, independently of the hybrid cross type. For the 91 pairs of characters studied, 252 plants are sufficient to estimate the Pearson correlation coefficient, in the bootstrap confidence interval 95%, equal to 0.30.

Index terms: *Zea mays*, experimental design, bootstrap resampling, linear relationship.

Introdução

Nas Ciências Agrárias, a mensuração de diversos caracteres de plantas é importante para maximizar as informações de um determinado experimento. Em programas de melhoramento de plantas, o estudo das relações lineares entre os caracteres pode fornecer resultados importantes, especialmente na identificação de caracteres que possam ser utilizados na 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; Bunchaft & Kellner, 2002; Barbeta et al., 2004; Kazmier, 2007; Ferreira, 2009).

O sinal do coeficiente de correlação linear de Pearson expressa o sentido da correlação, e sua 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 ($r = 0$). Estudos complementares,

como, por exemplo, os de correlações parciais, análise de trilha e correlação canônica, são importantes para avaliar as inter-relações em um conjunto de caracteres (Cruz & Regazzi, 1997; Cruz & Carneiro, 2003; Hair et al., 2005). Por dependerem do coeficiente de correlação linear de Pearson, esses procedimentos estatísticos serão tanto mais confiáveis, quanto mais precisa for sua estimativa.

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|). Nesse caso, embora o valor do coeficiente de correlação linear de Pearson apresente significância estatística, a amostra pode não ser representativa da população (Stevenson, 2001; Hair et al., 2005). Ainda, os sinais dos coeficientes de correlação fenotípica e genotípica podem, eventualmente, ser diferentes, e isso é, em geral, atribuído a erros de amostragem (Cruz & Regazzi, 1997). No entanto, quando o tamanho de amostra for excessivamente grande, um pequeno valor do coeficiente de correlação linear de Pearson (próximo de zero) ainda é considerado significativo, embora, do ponto de vista prático, não reflita, necessariamente uma importante relação linear entre os caracteres (Stevenson, 2001; Hair et al., 2005; Kazmier, 2007).

É importante dimensionar adequadamente a amostra, para estimar o coeficiente de correlação linear de Pearson de determinado par de caracteres com uma confiabilidade aceitável e, após isso, verificar sua significância estatística e prática. Embora o tamanho de amostra possa ser determinado a partir de intervalos de confiança apresentados por Ferreira (2009), a estimativa a partir de intervalos de confiança de "bootstrap" com reposição é um procedimento adequado, principalmente por ser independente da distribuição de probabilidade dos dados (Ferreira, 2009).

Estudos relacionados ao dimensionamento de amostra têm sido focalizados na estimação da média de caracteres de milho (Conceição et al., 1993; Silva et al., 1993; Fernandes & Silva, 1996; Resende & Souza Júnior, 1997; Silva et al., 1997, 1998; Palomino et al., 2000; Martin et al., 2005; Storck et al., 2007; Catapatti et al., 2008) e, de maneira geral, têm evidenciado a variabilidade do tamanho de amostra entre caracteres e genótipos. No entanto, não foram encontrados, na literatura, estudos sobre o dimensionamento de amostra, para a estimação do coeficiente de correlação linear de Pearson de pares de caracteres de milho.

O objetivo deste trabalho foi determinar o tamanho de amostra (número de plantas), para a estimação do coeficiente de correlação linear de Pearson, entre caracteres de híbridos simples, triplo e duplo de milho.

Material e Métodos

Em um experimento com a cultura de milho (*Zea mays* L.), foram realizados quatro ensaios de uniformidade com o híbrido simples P32R21, quatro com o híbrido triplo DKB 566 e quatro com o híbrido duplo DKB 747. Cada ensaio de uniformidade foi composto de quatro fileiras de 6 m de comprimento, espaçadas em 0,80 m. A semeadura foi realizada em 26/12/2008, e a densidade foi ajustada para cinco plantas por metro linear. Portanto, cada ensaio de uniformidade foi composto por 120 plantas, no total de 1.440 plantas (3 híbridos x 4 ensaios por híbrido x 120 plantas por ensaio). A adubação de base foi de 750 kg ha⁻¹ da fórmula 3-24-18 de NPK, e a adubação de cobertura foi de 300 kg ha⁻¹ de ureia. Os 12 ensaios de uniformidade foram instalados ao acaso, dentro da mesma área experimental do Departamento de Fitotecnia, da Universidade Federal de Santa Maria, Santa Maria, RS, a 29°42'S, 53°49'W, e a 95 m de altitude.

A avaliação foi realizada no máximo de plantas possível de cada ensaio, para se obter uma melhor representatividade das 1.440 plantas planejadas, 1.150 foram avaliadas, o que corresponde a 80% das plantas. Assim, foram tomadas aleatoriamente 361 plantas do híbrido simples P32R21, 373 plantas do híbrido triplo DKB 566 e 416 plantas do híbrido duplo DKB 747, e foram mensurados, em cada planta, os seguintes caracteres: diâmetros maior (DMAI) e menor (DMEN) do colmo, altura da planta na colheita (AP), altura da espiga na colheita (AE), peso da espiga (PE), número de fileiras por espiga (NF), comprimento (CE) e diâmetro (DE) de espiga, peso (PS) e diâmetro (DS) de sabugo, massa de cem grãos (MCG), número de grãos por espiga (NGR), comprimento de grãos (CGR) e produtividade de grãos (PROD).

Inicialmente, a partir de todas as plantas mensuradas, para cada híbrido, foi estimada a matriz de coeficientes de correlação linear de Pearson (r) entre os 14 caracteres, e a significância de r foi verificada por meio do teste t de Student, a 5 % de probabilidade. Assim, foram obtidas três matrizes (três híbridos) de dimensão 14. Em seguida, foi obtida a matriz média, por meio do

cálculo da média entre os r das três matrizes. Afim de verificar se as inferências poderiam ser realizadas com a matriz média dos r , foi calculado o coeficiente de correlação de Pearson (r) entre cada par de matrizes (três matrizes e a matriz média), e sua significância foi verificada pelo teste t de Student, a 5% de probabilidade, com $n-2 = 89$ graus de liberdade, em que n são os 91 pares de caracteres (combinação de 14 caracteres tomados dois a dois). Quanto mais próximo de um for a estimativa obtida, maior a concordância entre a matriz do híbrido e a matriz média.

A seguir, para cada um dos 91 pares de caracteres, foram definidos 71, 73 e 82 tamanhos de amostra, correspondentes aos híbridos simples, triplo e duplo. O tamanho de amostra inicial foi de dez plantas, e os demais foram obtidos com incrementos de cinco plantas até atingir 360, 370 e 415 plantas, para os híbridos simples, triplo e duplo, respectivamente. Em seguida, para cada tamanho de amostra, de cada par de caracteres e híbrido, 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. Assim, para cada tamanho de amostra, de cada par de caracteres e híbrido, foram obtidas 1.000 estimativas de "bootstrap" de r (Ferreira, 2009), e foram determinados os valores mínimo, percentil 2,5%, média, percentil 97,5% e máximo de r . Depois, calculou-se a amplitude do intervalo de confiança de "bootstrap" de 95%, pela diferença entre o percentil 97,5% e o 2,5%.

Para cada par de caracteres e híbrido, determinou-se o tamanho da amostra (n) para a estimação do coeficiente de correlação linear de Pearson (r), para uma amplitude do intervalo de confiança de "bootstrap" de 95%, igual a 0,30. Para essa determinação, partiu-se do tamanho inicial (10 plantas) e considerou-se como tamanho de amostra (n) o número de plantas a partir do qual a amplitude do intervalo de confiança de "bootstrap" de 95% foi menor ou igual a 0,30. Assim, foram obtidas três matrizes (três híbridos) de dimensão 14. Em seguida, foi obtida a matriz média, por meio do cálculo da média entre os n das três matrizes. Afim de verificar se as inferências poderiam ser realizadas com a matriz média dos n , foi calculado o coeficiente de correlação de Pearson entre cada par de matrizes (três matrizes e a matriz média), e sua significância foi verificada pelo teste t de Student, a 5% de probabilidade, com $n-2 = 89$ graus de liberdade, em que n são os 91 pares de caracteres.

A fim de visualizar as relações entre as 91 estimativas de (r) e de (n) dos híbridos simples, triplo e duplo e da média dos três híbridos, foi feito o diagrama de dispersão, e foi calculada a matriz de coeficientes de correlação linear de Pearson (r). As análises foram realizadas com o auxílio do programa R (R Development Core Team, 2006) e do aplicativo Office Excel.

Resultados e Discussão

Os coeficientes de correlação linear de Pearson (r), para cada par de caracteres, foram em geral similares entre os híbridos simples, triplo e duplo de milho, o que mostra as mesmas tendências, tendo diferido um pouco quanto à magnitude das associações (Tabela 1). Sustenta essas inferências o fato de que os coeficientes de correlação de Pearson, entre os pares de matrizes dos três híbridos, foram significativos e de alta magnitude ($r \geq 0,89$, $p \leq 0,05$), o que evidencia boa concordância entre os híbridos (Figura 1). Ainda, entre a matriz de cada híbrido e a matriz média dos três híbridos, houve associação linear significativa e de alta magnitude ($r \geq 0,96$, $p \leq 0,05$), o que confirma que as inferências, de um modo geral, podem ser realizadas em nível de média dos híbridos.

Entre os 91 pares de caracteres, a magnitude do coeficiente de correlação linear de Pearson (r) oscilou entre 0,06 ($p > 0,05$) (diâmetro de sabugo e comprimento de grãos) e 1,00 ($p \leq 0,05$) (peso de espiga e produtividade de grãos), o que revela, respectivamente, ausência de relação linear e relação linear positiva perfeita entre os caracteres. Entre quatro pares de caracteres não houve associação linear ($r \leq |0,10|$, $p > 0,05$) e entre os demais 87 pares de caracteres houve associação linear positiva ($r > 0,10$, $p \leq 0,05$) (Tabela 1). Do ponto de vista prático, o elevado número de plantas mensuradas (≥ 361) sugere que as estimativas de r apresentam elevada precisão, e com isso, associações lineares de baixa magnitude (r próximos de zero) são significantes. Nesses casos, 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.

Em relação aos híbridos simples, triplo e duplo, houve, em geral, um decréscimo gradativo da amplitude de variação (valor máximo menos valor mínimo) entre

as 1.000 estimativas de "bootstrap" de r , obtidas a partir de amostras de 10 plantas, desde os pares de caracteres com associação linear significativa, de baixa magnitude (r próximo a zero), até aos pares de caracteres com alta magnitude dessa associação linear (r próximo a $|1|$) (Tabela 2). Como exemplo, em relação ao diâmetro maior do colmo e altura da planta na colheita, r médio dos três híbridos igual a 0,11 (Tabela 1), a amplitude de variação entre as 1.000 estimativas de "bootstrap" de r oscilou entre -0,85 e 0,90, -0,83 e 0,91 e -0,88 e 0,91, respectivamente, para os híbridos simples, triplo e duplo. Em outro extremo, em relação aos caracteres

peso de espiga e produtividade de grãos, r médio dos três híbridos igual a 1,00 (Tabela 1), o valor de r variou entre 0,96 e 1,00, 0,99 e 1,00 e 0,94 e 1,00, respectivamente, para os híbridos simples, triplo e duplo (Tabela 2 e Figura 2). Especialmente para os caracteres com associação linear significativa de baixa magnitude, esses resultados evidenciam que amostras de 10 plantas podem fornecer resultados significativos e contraditórios em relação à associação linear entre eles. Portanto, é possível que resultados divergentes, em publicações científicas, estejam associados ao tamanho de amostra insuficiente. Para uma mesma

Tabela 1. Estimativas dos coeficientes de correlação linear de Pearson⁽¹⁾, entre 91 pares de caracteres dos híbridos simples (P32R21), triplo (DKB 566), duplo (DKB 747) de milho, e média dos híbridos.

Caráter ⁽²⁾	DMAI	DMEN	AP	AE	PE	NF	CE	DE	PS	DS	MCG	NGR	CGR	PROD
Híbrido simples acima da diagonal e híbrido triplo abaixo da diagonal														
DMAI		0,79	0,14	0,10	0,38	0,26	0,41	0,32	0,42	0,31	0,19	0,36	0,19	0,36
DMEN	0,77		0,34	0,31	0,44	0,32	0,41	0,40	0,39	0,37	0,27	0,41	0,24	0,43
AP	0,04	0,13		0,61	0,40	0,21	0,38	0,35	0,40	0,29	0,29	0,33	0,24	0,38
AE	0,03	0,16	0,74		0,26	0,09	0,24	0,24	0,20	0,14	0,22	0,23	0,20	0,26
PE	0,36	0,34	0,31	0,31		0,57	0,80	0,83	0,67	0,48	0,54	0,95	0,68	1,00
NF	0,17	0,18	0,09	0,17	0,42		0,38	0,70	0,23	0,41	0,11	0,65	0,57	0,59
CE	0,36	0,35	0,28	0,26	0,84	0,26		0,56	0,73	0,28	0,30	0,79	0,50	0,78
DE	0,36	0,39	0,26	0,31	0,83	0,57	0,66		0,47	0,55	0,53	0,79	0,84	0,83
PS	0,35	0,34	0,31	0,23	0,78	0,14	0,77	0,62		0,48	0,40	0,55	0,26	0,62
DS	0,35	0,38	0,19	0,16	0,54	0,43	0,44	0,68	0,61		0,40	0,40	0,02	0,46
MCG	0,24	0,28	0,27	0,18	0,48	-0,20	0,42	0,46	0,59	0,41		0,30	0,38	0,54
NGR	0,31	0,28	0,24	0,28	0,90	0,61	0,74	0,75	0,56	0,42	0,08		0,69	0,95
CGR	0,24	0,25	0,22	0,30	0,73	0,46	0,57	0,86	0,40	0,21	0,33	0,70		0,70
PROD	0,35	0,33	0,31	0,31	1,00	0,43	0,83	0,83	0,75	0,52	0,47	0,91	0,75	
Híbrido duplo acima da diagonal e média dos híbridos abaixo da diagonal														
DMAI		0,79	0,14	0,14	0,45	0,16	0,38	0,44	0,41	0,35	0,35	0,38	0,26	0,45
DMEN	0,78		0,31	0,26	0,48	0,13	0,42	0,46	0,43	0,35	0,37	0,41	0,28	0,47
AP	0,11	0,26		0,73	0,48	0,15	0,44	0,43	0,44	0,29	0,35	0,42	0,30	0,47
AE	0,09	0,24	0,69		0,37	0,04	0,30	0,35	0,26	0,15	0,29	0,34	0,32	0,38
PE	0,40	0,42	0,40	0,31		0,33	0,86	0,81	0,86	0,56	0,59	0,91	0,56	0,99
NF	0,19	0,21	0,15	0,10	0,44		0,20	0,54	0,20	0,46	-0,16	0,49	0,28	0,35
CE	0,38	0,39	0,37	0,27	0,83	0,28		0,61	0,80	0,45	0,47	0,79	0,39	0,84
DE	0,38	0,42	0,35	0,30	0,82	0,60	0,61		0,69	0,68	0,50	0,75	0,69	0,81
PS	0,40	0,39	0,38	0,23	0,77	0,19	0,77	0,59		0,65	0,57	0,69	0,29	0,80
DS	0,33	0,37	0,25	0,15	0,52	0,43	0,39	0,64	0,58		0,35	0,47	-0,06	0,52
MCG	0,26	0,30	0,30	0,23	0,54	-0,08	0,40	0,50	0,52	0,38		0,23	0,34	0,58
NGR	0,35	0,36	0,33	0,28	0,92	0,58	0,77	0,76	0,60	0,43	0,21		0,56	0,92
CGR	0,23	0,26	0,25	0,28	0,66	0,44	0,49	0,80	0,32	0,06	0,35	0,65		0,60
PROD	0,39	0,41	0,39	0,32	1,00	0,45	0,82	0,82	0,72	0,50	0,53	0,93	0,68	

⁽¹⁾Valor superior – a $[0,10]$ para o híbrido simples, $[0,10]$ para o triplo e $[0,09]$ para o duplo – é significativo a 5% de probabilidade pelo teste t , com 359, 371 e 414 graus de liberdade, respectivamente. ⁽²⁾DMAI, diâmetro maior do colmo; DMEN, diâmetro menor do colmo; AP, altura da planta na colheita; AE, altura da espiga na colheita; PE, peso da espiga; NF, número de fileiras por espiga; CE, comprimento da espiga; DE, diâmetro da espiga; PS, peso do sabugo; DS, diâmetro do sabugo; MCG, massa de cem grãos; NGR, número de grãos por espiga; CGR, comprimento de grãos; e PROD, produtividade de grãos.

precisão, maior tamanho de amostra será necessário para estimar o r de pares de caracteres com menor intensidade de relação linear e vice-versa.

Em geral, o diâmetro maior do colmo (DMAI) e a altura de planta na colheita (AP), e o peso de espiga (PE) e a produtividade de grãos (PROD) representam comportamentos extremos, em relação aos tamanhos de amostras simulados, entre os 91 pares de caracteres estudados (Figura 2). Para a estimação do coeficiente de correlação linear de Pearson com uma mesma precisão, pares de caracteres com menor intensidade de relação linear (DMAI e AP) necessitam maior tamanho de amostra, em comparação aos pares de caracteres com maior intensidade de relação linear (PE e PROD), independentemente do híbrido de milho.

Quanto à estimação do coeficiente de correlação linear de Pearson, para uma amplitude do intervalo de confiança "bootstrap" de 95% igual a 0,30, para cada par de caracteres, o tamanho de amostra (número de plantas) foi similar entre os três híbridos de milho (Tabela 3). Os elevados valores dos coeficientes de correlação de Pearson (r), entre os pares de matrizes dos três híbridos ($r \geq 0,78$, $p \leq 0,05$), e entre a matriz de cada híbrido e a matriz média dos três híbridos ($r \geq 0,93$, $p \leq 0,05$) (Figura 1) revelam, respectivamente, que há boa concordância entre os híbridos e que as inferências, em relação ao tamanho de amostra, de um modo geral, podem ser realizadas em nível de média dos híbridos.

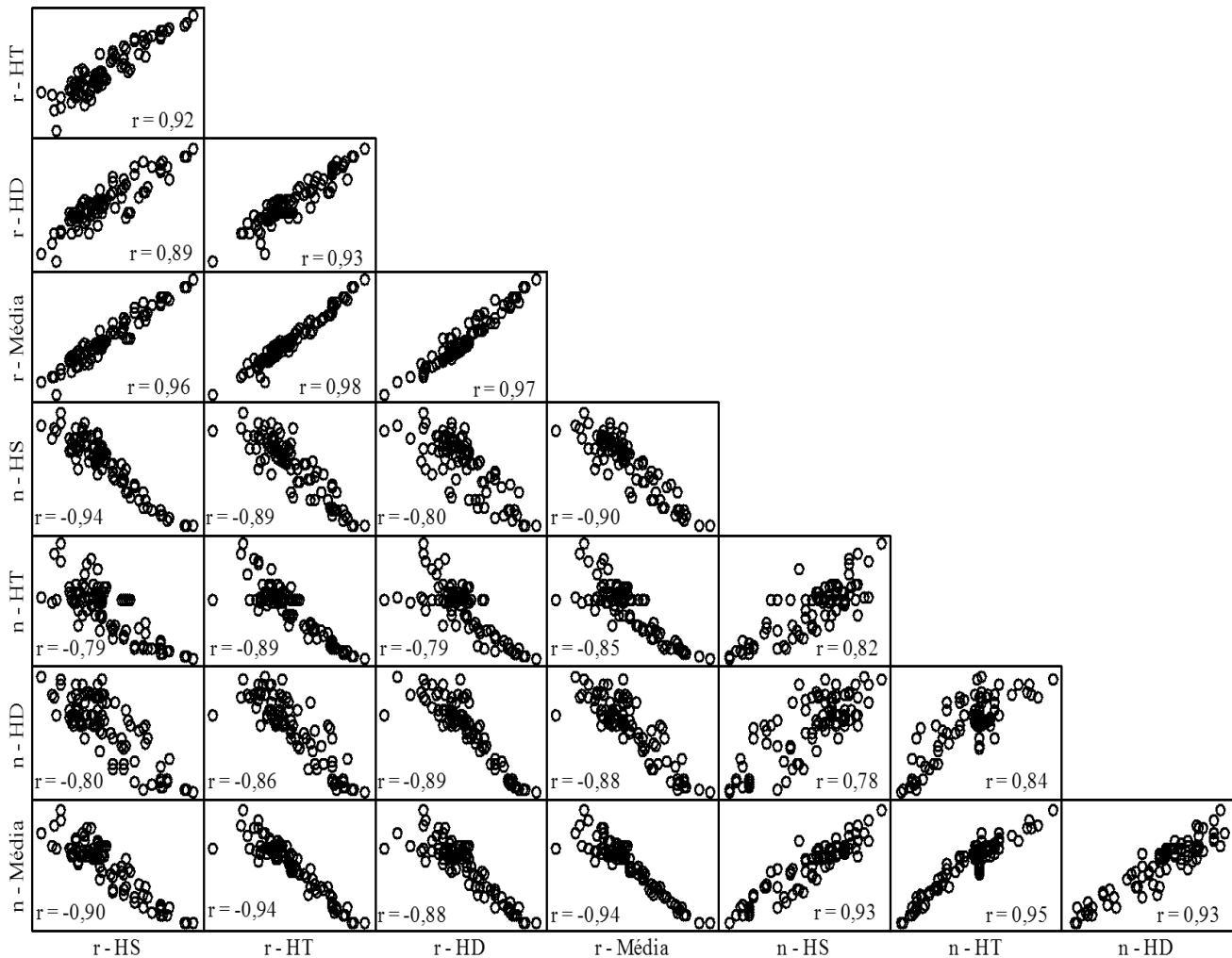


Figura 1. Diagramas de dispersão entre as 91 estimativas (14 caracteres tomados dois a dois) do coeficiente de correlação linear de Pearson (r) e do tamanho de amostra, em número de plantas (n), para a estimação de r para a amplitude do intervalo de confiança de "bootstrap" de 95%, igual a 0,30, referentes aos híbridos simples (HS), triplo (HT), duplo (HD) e média dos três híbridos de milho. Valor de r superior a $|0,20|$ é significativo a 5% de probabilidade, pelo teste t , com 89 graus de liberdade.

Tabela 2. Valores mínimo (acima da diagonal) e máximo (abaixo da diagonal) do coeficiente de correlação linear de Pearson⁽¹⁾, entre 91 pares de caracteres de híbridos simples, triplo e duplo de milho, obtidos em 1.000 reamostragens "bootstrap" com reposição de 10 plantas.

Caráter ⁽²⁾	DMAI	DMEN	AP	AE	PE	NF	CE	DE	PS	DS	MCG	NGR	CGR	PROD
Híbrido simples (P32R21)														
DMAI		-0,13	-0,85	-0,87	-0,70	-0,71	-0,77	-0,78	-0,78	-0,72	-0,84	-0,70	-0,87	-0,74
DMEN	0,99		-0,77	-0,83	-0,78	-0,73	-0,74	-0,74	-0,75	-0,76	-0,78	-0,71	-0,77	-0,74
AP	0,90	0,98		-0,56	-0,72	-0,80	-0,78	-0,67	-0,71	-0,66	-0,83	-0,73	-0,75	-0,75
AE	0,89	0,95	0,97		-0,75	-0,84	-0,85	-0,82	-0,81	-0,81	-0,78	-0,73	-0,72	-0,74
PE	0,92	0,99	0,95	0,92		-0,52	-0,17	-0,34	-0,62	-0,74	-0,76	0,59	-0,50	0,96
NF	0,92	0,92	0,90	0,92	0,98		-0,68	-0,32	-0,79	-0,69	-0,87	-0,53	-0,48	-0,53
CE	0,94	0,95	0,94	0,90	0,99	0,96		-0,51	-0,17	-0,86	-0,79	-0,38	-0,53	-0,19
DE	0,97	0,97	0,97	0,93	0,98	0,97	0,95		-0,64	-0,58	-0,59	-0,37	-0,34	-0,19
PS	0,95	0,91	0,95	0,95	0,99	0,95	0,98	0,95		-0,66	-0,85	-0,61	-0,73	-0,72
DS	0,89	0,92	0,94	0,92	0,93	0,94	0,97	0,96	0,95		-0,75	-0,71	-0,86	-0,75
MCG	0,93	0,94	0,94	0,92	0,96	0,86	0,96	0,98	0,91	0,97		-0,80	-0,86	-0,71
NGR	0,94	0,99	0,96	0,91	1,00	0,97	0,99	0,98	0,96	0,95	0,93		-0,51	0,63
CGR	0,94	0,92	0,90	0,92	0,98	0,98	0,97	0,99	0,95	0,87	0,95	0,99		-0,47
PROD	0,92	0,99	0,95	0,92	1,00	0,96	0,99	0,98	0,99	0,93	0,95	1,00	0,98	
Híbrido triplo (DKB 566)														
DMAI		-0,41	-0,83	-0,85	-0,74	-0,69	-0,72	-0,59	-0,72	-0,70	-0,86	-0,78	-0,65	-0,73
DMEN	0,98		-0,86	-0,80	-0,82	-0,81	-0,72	-0,73	-0,70	-0,73	-0,69	-0,80	-0,71	-0,82
AP	0,91	0,89		-0,22	-0,66	-0,89	-0,72	-0,72	-0,74	-0,92	-0,72	-0,81	-0,82	-0,66
AE	0,87	0,92	0,97		-0,67	-0,77	-0,71	-0,77	-0,82	-0,85	-0,94	-0,86	-0,84	-0,67
PE	0,94	0,94	0,94	0,94		-0,85	-0,29	-0,24	-0,27	-0,80	-0,76	0,20	-0,26	0,99
NF	0,86	0,87	0,85	0,96	0,95		-0,83	-0,48	-0,77	-0,67	-0,96	-0,54	-0,84	-0,85
CE	0,97	0,93	0,91	0,88	0,99	0,90		-0,53	0,07	-0,76	-0,66	-0,47	-0,72	-0,31
DE	0,90	0,94	0,97	0,93	0,99	0,97	0,98		-0,40	-0,47	-0,80	-0,56	0,03	-0,22
PS	0,95	0,92	0,95	0,88	0,99	0,91	0,98	0,98		-0,49	-0,42	-0,45	-0,58	-0,33
DS	0,90	0,94	0,93	0,93	0,95	0,93	0,95	0,98	0,96		-0,60	-0,76	-0,82	-0,80
MCG	0,94	0,94	0,91	0,92	0,97	0,76	0,97	0,96	0,97	0,93		-0,87	-0,91	-0,78
NGR	0,94	0,95	0,91	0,91	0,99	0,97	0,99	0,99	0,96	0,93	0,88		-0,48	0,19
CGR	0,94	0,92	0,95	0,92	0,98	0,94	0,97	0,99	0,93	0,93	0,97	0,98		-0,27
PROD	0,93	0,94	0,94	0,95	1,00	0,95	0,99	0,99	0,99	0,95	0,96	0,99	0,98	
Híbrido duplo (DKB 747)														
DMAI		-0,01	-0,88	-0,79	-0,70	-0,73	-0,80	-0,68	-0,84	-0,85	-0,81	-0,83	-0,80	-0,67
DMEN	0,98		-0,79	-0,89	-0,79	-0,83	-0,75	-0,79	-0,82	-0,80	-0,76	-0,83	-0,80	-0,78
AP	0,91	0,93		-0,18	-0,79	-0,88	-0,87	-0,74	-0,81	-0,85	-0,84	-0,75	-0,75	-0,77
AE	0,91	0,92	0,98		-0,76	-0,91	-0,82	-0,77	-0,85	-0,81	-0,66	-0,74	-0,76	-0,72
PE	0,94	0,97	0,98	0,96		-0,78	0,05	-0,37	0,28	-0,79	-0,35	0,35	-0,81	0,94
NF	0,91	0,90	0,91	0,90	0,92		-0,80	-0,60	-0,92	-0,65	-0,91	-0,74	-0,83	-0,79
CE	0,94	0,98	0,94	0,93	0,99	0,91		-0,70	-0,30	-0,68	-0,55	-0,12	-0,79	0,00
DE	0,94	0,97	0,96	0,93	0,98	0,95	0,98		-0,76	-0,51	-0,59	-0,43	-0,49	-0,32
PS	0,96	0,97	0,96	0,95	0,99	0,90	0,99	0,96		-0,52	-0,60	-0,17	-0,75	-0,04
DS	0,93	0,95	0,93	0,89	0,96	0,94	0,96	0,99	0,98		-0,78	-0,80	-0,86	-0,81
MCG	0,96	0,96	0,97	0,91	0,98	0,78	0,95	0,97	0,97	0,96		-0,72	-0,73	-0,28
NGR	0,95	0,96	0,92	0,93	0,99	0,96	0,99	0,99	0,98	0,96	0,95		-0,69	0,31
CGR	0,89	0,93	0,94	0,89	0,96	0,94	0,94	0,98	0,97	0,90	0,93	0,96		-0,79
PROD	0,93	0,99	0,97	0,96	1,00	0,93	0,99	0,98	0,99	0,96	0,99	1,00	0,95	

⁽¹⁾Valor superior a |0,64| é significativo a 5% de probabilidade pelo teste t, com 8 graus de liberdade. ⁽²⁾DMAI, diâmetro maior do colmo; DMEN, diâmetro menor do colmo; AP, altura da planta na colheita; AE, altura da espiga na colheita; PE, peso da espiga; NF, número de fileiras por espiga; CE, comprimento da espiga; DE, diâmetro da espiga; PS, peso do sabugo; DS, diâmetro do sabugo; MCG, massa de cem grãos; NGR, número de grãos por espiga; CGR, comprimento de grãos; e PROD, produtividade de grãos.

O número de plantas necessário para a estimação do coeficiente de correlação linear de Pearson (r), para uma amplitude do intervalo de confiança de "bootstrap" de 95% igual a 0,30, oscilou entre 252 plantas (média dos três híbridos), quanto ao par de caracteres diâmetro maior do

colmo e altura de planta na colheita, e 10 plantas quanto ao par de caracteres peso de espiga e produtividade de grãos (Tabela 3). A ocorrência de variabilidade do tamanho de amostra entre os pares de caracteres de híbridos de milho foram demonstradas com os resultados do presente estudo.

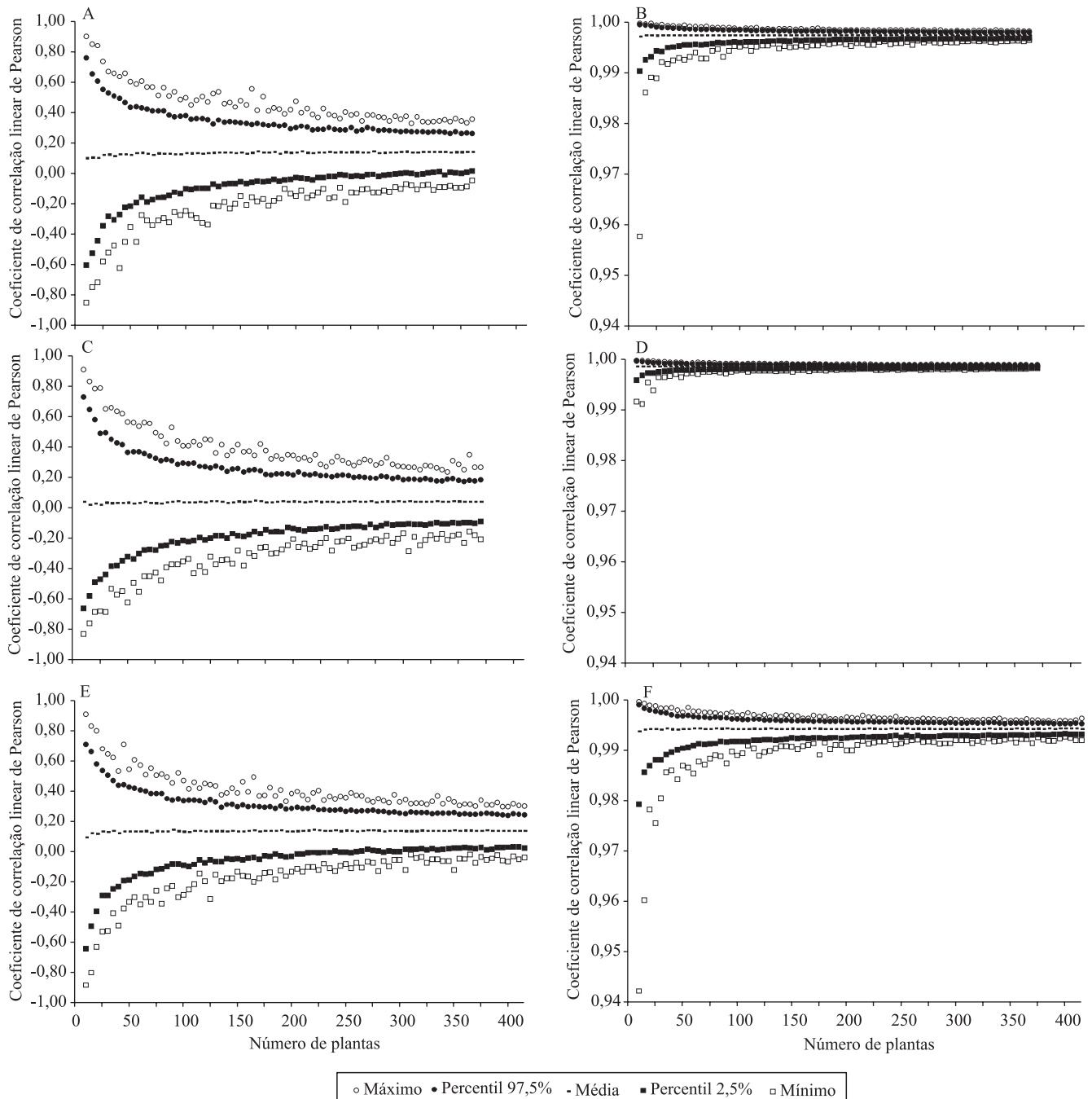


Figura 2. Valores mínimo, percentil 2,5%, média, percentil 97,5% e máximo das 1.000 estimativas de "bootstrap" do coeficiente de correlação linear de Pearson (r) de diferentes tamanhos de amostra em número de plantas (n), dos pares de caracteres diâmetro maior do colmo (DMAI) e altura da planta na colheita (AP) de híbridos simples (A), triplo (C) e duplo (E) de milho, e peso da espiga (PE) e produtividade de grãos (PROD) de híbridos simples (B), triplo (D) e duplo (F).

Variabilidade de tamanho de amostra, para a estimação da média, 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).

Houve correlação negativa ($r \leq -0,79$, $p \leq 0,05$) entre as 91 estimativas do coeficiente de correlação linear de Pearson (r) e do tamanho de amostra (n), para a estimação de r para uma amplitude do intervalo de confiança de "bootstrap" de 95% igual a 0,30, referente aos híbridos simples, triplo, duplo e média dos três híbridos de milho (Figura 1). Com esses resultados foram

confirmadas as inferências realizadas e demonstrado que, independentemente do híbrido de milho (simples, triplo ou duplo), para a estimação do coeficiente de correlação linear de Pearson, com uma mesma precisão, o tamanho de amostra (número de plantas) aumenta na direção de pares de caracteres com menor intensidade de relação linear.

Na prática, para os 91 pares de caracteres estudados, 252 plantas de milho são suficientes para estimar o coeficiente de correlação linear de Pearson, para uma amplitude do intervalo de confiança de "bootstrap" de 95%, igual a 0,30. Para uma mesma precisão, diferentes tamanhos de amostra deveriam ser utilizados. No entanto,

Tabela 3. Tamanho de amostra expressa em número de plantas (n) para a estimação do coeficiente de correlação linear de Pearson (r), de 91 pares de caracteres de híbridos simples (P32R21), triplo (DKB 566) e duplo (DKB 747) de milho, e média dos híbridos para a amplitude do intervalo de confiança de "bootstrap" de 95%, igual a 0,30.

Caráter ⁽¹⁾	DMAI	DMEN	AP	AE	PE	NF	CE	DE	PS	DS	MCG	NGR	CGR	PROD
Híbrido simples acima da diagonal e híbrido triplo abaixo da diagonal														
DMAI		40	240	205	140	125	150	155	140	170	185	150	160	140
DMEN	40		190	190	155	140	155	170	145	135	180	170	170	155
AP	300	260		90	160	165	190	165	145	200	220	185	145	160
AE	270	250	45		170	185	185	205	175	220	185	185	160	170
PE	180	190	150	155		80	30	20	80	135	100	10	45	10
NF	130	165	160	150	160		115	65	165	125	200	65	125	70
CE	150	155	170	150	25	235		80	40	205	175	40	105	40
DE	160	170	195	195	30	100	80		120	115	135	30	25	20
PS	180	165	155	155	40	160	35	65		110	140	105	160	95
DS	130	120	195	255	90	115	120	70	80		165	155	215	135
MCG	180	175	165	155	160	160	190	160	80	120		190	195	100
NGR	165	175	145	150	15	90	50	55	65	120	220		60	10
CGR	190	195	175	170	45	160	95	30	95	165	190	75		45
PROD	180	190	150	155	10	160	25	30	40	95	160	15	40	
Híbrido duplo acima da diagonal e média dos híbridos abaixo da diagonal														
DMAI		30	215	180	130	175	150	135	140	150	145	130	160	130
DMEN	37		185	190	170	190	215	180	180	180	140	185	205	185
AP	252	212		65	135	180	140	150	135	205	170	150	185	140
AE	218	210	67		135	185	160	150	150	205	130	140	135	130
PE	150	172	148	153		140	20	30	15	105	90	15	125	10
NF	143	165	168	173	127		205	100	165	115	150	120	160	130
CE	150	175	167	165	25	185		95	25	130	105	35	170	20
DE	150	173	170	183	27	88	85		50	60	105	55	70	35
PS	153	163	145	160	45	163	33	78		60	85	50	160	30
DS	150	145	200	227	110	118	152	82	83		170	120	220	115
MCG	170	165	185	157	117	170	157	133	102	152		185	160	95
NGR	148	177	160	158	13	92	42	47	73	132	198		130	15
CGR	170	190	168	155	72	148	123	42	138	200	182	88		105
PROD	150	177	150	152	10	120	28	28	55	115	118	13	63	

⁽¹⁾DMAI, diâmetro maior do colmo; DMEN, diâmetro menor do colmo; AP, altura da planta na colheita; AE, altura da espiga na colheita; PE, peso da espiga; NF, número de fileiras por espiga; CE, comprimento da espiga; DE, diâmetro da espiga; PS, peso do sabugo; DS, diâmetro do sabugo; MCG, massa de cem grãos, NGR, número de grãos por espiga; CGR, comprimento de grãos; e PROD, produtividade de grãos.

é comum o uso do mesmo tamanho de amostra, para mensurar todos os caracteres. Assim, caso a decisão seja a utilização de 252 plantas para mensuração de todos os caracteres, as estimativas de r entre os pares de caracteres terão precisão diferenciada, e serão maiores quanto maior for o valor absoluto de r . Portanto, 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 razão do elevado número de plantas a serem mensuradas. Não cabe aqui o julgamento da magnitude da amplitude do intervalo de confiança de "bootstrap" de 95%. Isto deve ficar a critério do pesquisador que usufruir dessas informações para o seu planejamento experimental.

Conclusões

1. Na estimação do coeficiente de correlação linear de Pearson com uma mesma precisão, o tamanho de amostra, expresso em número de plantas, aumenta na direção de pares de caracteres com menor intensidade da relação linear, independentemente de o híbrido de milho ser simples, triplo ou duplo.

2. Para os 91 pares de caracteres de plantas de milho estudados, 252 plantas 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 0,30.

Agradecimentos

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico e à 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**. 2.ed. São Paulo: Atlas, 2004. 410p.
- BUNCHAFT, G.; KELLNER, S.R. de O. **Estatística sem mistérios**. 4.ed. Petrópolis: Vozes, 2002. v.2, 303p.
- CARGNELUTTI FILHO, A.; EVANGELISTA, D.H.R.; GONÇALVES, E.C.P.; STORCK, L. Tamanho de amostra de caracteres de genótipos de soja. **Ciência Rural**, v.39, 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**, v.40, p.280-287, 2010.
- CARGNELUTTI FILHO, A.; RIBEIRO, N.D.; STORCK, L.; JOST, E.; POERSCH, N.L. Tamanho de amostra de caracteres de cultivares de feijão. **Ciência Rural**, v.38, p.635-642, 2008.
- CATAPATTI, T.R.; GONÇALVES, M.C.; SILVA NETO, M.R.; SOBROZA, R. Tamanho de amostra e número de repetições para avaliação de caracteres agrônômicos em milho-pipoca. **Ciência e Agrotecnologia**, v.32, p.855-862, 2008.
- CONCEIÇÃO, M.M. da; SILVA, P.S.L.; MACHADO, A.A. Efeitos de bordadura e amostragem em experimento de híbrido de milho. **Revista Ciência Agrônômica**, v.24, p.63-69, 1993.
- CRUZ, C.D.; CARNEIRO, P.C.S. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa: UFV, 2003. 585p.
- CRUZ, C.D.; REGAZZI, A.J. **Modelos biométricos aplicados ao melhoramento genético**. 2.ed. Viçosa: UFV, 1997. 390p.
- FERNANDES, E.N.; SILVA, P.S.L. Tamanho de amostra e método de amostragem para caracteres da espiga do milho. **Ciência e Agrotecnologia**, v.20, p.252-256, 1996.
- FERREIRA, D.F. **Estatística básica**. 2.ed. Lavras: UFLA, 2009. 664p.
- HAIR, J.F.; ANDERSON, R.E.; TATHAM, R.L.; BLACK, W.C. **Análise multivariada de dados**. 5.ed. Porto Alegre: Bookman, 2005. 593p.
- KAZMIER, L.J. **Estatística aplicada à administração e economia**. 4.ed. Porto Alegre: Bookman, 2007. 392p.
- MARTIN, T.N.; STORCK, L.; LÚCIO, A.D.; LORENTZ, L.H. Plano amostral em parcelas de milho para avaliação de atributos de espigas. **Ciência Rural**, v.35, p.1257-1262, 2005.
- PALOMINO, E.C.; RAMALHO, M.A.P.; FERREIRA, D.F. Tamanho da amostra para avaliação de famílias de meios-irmãos de milho. **Pesquisa Agropecuária Brasileira**, v.35, p.1433-1439, 2000.
- R DEVELOPMENT CORE TEAM. **R: a language and environment for statistical computing**. Vienna: R Foundation for Statistical Computing, 2006.
- RESENDE, M.D.V. de; SOUZA JÚNIOR, C.L. de. Número de repetições e tamanho da parcela para seleção de progênies de milho em solos sob cerrado e fértil. **Pesquisa Agropecuária Brasileira**, v.32, p.781-788, 1997.
- SILVA, J. da; SILVA, P.S.L. e; NUNES, R. de P. Amostragem e tamanho da amostra na estimação de caracteres da espiga do milho. **Pesquisa Agropecuária Brasileira**, v.33, p.1983-1988, 1998.
- SILVA, P.S.L. e; BEZERRA, A.R.; OLIVEIRA, F.H.T. de; SILVA, K.M.B. e. Efeitos da amostragem não-aleatória e do tamanho da amostra na estimação de dois caracteres de milho. **Revista Ceres**, v.44, p.72-82, 1997.
- SILVA, P.S.L. e; SILVA, G.S.; RODRIGUES, M.V.; MACHADO, A.A. Métodos de amostragem e tamanho da amostra para alguns caracteres do milho. **Ciência Agrônômica**, v.24, p.5-10, 1993.
- STEVENSON, W.J. **Estatística aplicada à administração**. São Paulo: Harbra, 2001. 495p.
- 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**, v.64, p.30-35, 2007.