

4. RESULTADOS E DISCUSSÃO

O conjunto de dados original apresentava alguns valores perdidos, uma vez que houve a morte de plantas nas parcelas antes da coleta dos dados, gerando assim um conjunto de dados desbalanceado, que impossibilitava a realização da reamostragem completa, conforme apresentado na metodologia.

Diante deste fato, optou-se por simular novos valores para estas plantas utilizando o método da inversão (LAW e KELTON, 2000), que consistiu basicamente em traçar um gráfico de distribuição acumulada empírica para cada variável dentro de cada família, gerar valores de uma distribuição uniforme no intervalo 0 a 1 e, então, identificar no gráfico qual valor do vetor de dados correspondia àquele valor da distribuição uniforme gerado. O vetor de dados utilizado para construção deste gráfico foi composto pelos valores de todos os seis blocos de cada família, uma vez que testes realizados mostraram não haver efeito significativo de blocos dentro das famílias.

Das 22 famílias avaliadas, quatro apresentaram um número elevado de plantas mortas, sendo, por este motivo, descartadas. Assim, todos os resultados a seguir foram obtidos avaliando-se apenas 18 famílias.

4.1. Caracterização da população

Na Tabela 1 são apresentados os valores de quadrado médio e média geral para todas as variáveis em estudo, obtidos através da análise de variância do conjunto de dados desbalanceado e deste mesmo conjunto balanceado pela recomposição, via simulação, dos valores perdidos. Comparando as estimativas do quadrado médio do erro dentro das parcelas, verifica-se que os valores encontrados foram os mesmos para diâmetro do colmo (DC) e comprimento do colmo (CC), e muito próximos para as demais variáveis avaliadas.

Tabela 1 – Resumo da análise de variância para o conjunto de dados desbalanceado e após ser balanceado via simulação dos valores perdidos para Brix, diâmetro do colmo (DC), comprimento do colmo (CC), número de colmos (NC), peso estimado de colmos (PE), tonelada de cana por hectare (TCH) e tonelada de brix por hectare (TBH)

	F.V.	G.L.	Quadrados Médios						
			Brix	DC	CC	NC	P	TCH	TBH
Experimento desbalanceado	Bloco	5	34,85	0,30	1,54	12,49	62,46	12746,72	367,81
	Família	17	22,95**	3,11**	3,38**	32,77**	87,95**	17949,12**	660,18**
	Entre	85	8,73**	0,24 ^{ns}	0,27**	7,24 ^{ns}	14,36*	2930,91*	117,63*
	Dentro	1924	3,16	0,20	0,16	7,67	11,77	2402,53	89,63
	Média	-	19,30	2,07	2,07	5,81	4,60	65,70	12,68
Experimento balanceado via simulação	Bloco	5	35,91	0,28	1,48	10,04	58,19	11876,24	337,70
	Família	17	23,79**	3,15**	3,60**	33,07**	91,53**	27677,88**	688,54**
	Entre	85	8,74**	0,22 ^{ns}	0,27**	7,58 ^{ns}	13,22*	2697,73*	111,32*
	Dentro	2052	3,19	0,21	0,16	7,66	11,63	2372,91	88,51
	Média	-	19,31	2,06	2,07	5,83	4,57	65,34	12,62

** significativo a 1% de probabilidade pelo teste F

* significativo a 5% de probabilidade pelo teste F

^{ns} não significativo a 1% e 5% de probabilidade pelo teste F

As estimativas de média geral também seguiram a mesma tendência, não sofrendo maiores alterações. Os níveis de significância das fontes de variação testadas também não foram alterados, indicando que mesmo havendo pequena diferença matemática entre as estimativas de quadrados médios, não houve influência no resultado dos testes de significância aplicados para as fontes de variação. Isso mostra que o método da inversão, utilizado na simulação dos valores perdidos, foi eficiente no sentido de não causar mudanças importantes nas características originais do conjunto de

dados coletado em campo, possibilitando a adaptação deste para atender às necessidades metodológicas do trabalho.

Assim, todos os resultados mostrados a seguir foram obtidos utilizando o conjunto de dados balanceado. Na Tabela 2 são apresentadas as estimativas dos parâmetros genéticos e fenotípicos para todas as variáveis avaliadas. Estas estimativas foram consideradas como os verdadeiros valores do parâmetro, uma vez que o conjunto de dados estudado foi considerado como sendo a população conhecida. Como as famílias utilizadas neste estudo são de irmãos completos, há apenas uma estimativa de variância genotípica, uma vez que para esta estrutura de família as estimativas da variância genotípica entre e dentro possuem o mesmo estimador, desconsiderando os efeitos atribuídos aos desvios de dominância.

Tabela 2 – Estimativas dos parâmetros genéticos e fenotípicos populacionais para Brix, diâmetro do colmo (DC), comprimento do colmo (CC), número de colmos (NC), peso estimado de colmos (PE), tonelada de cana por hectare (TCH) e tonelada de brix por hectare (TBH)

Parâmetros	Variáveis						
	Brix	DC	CC	NC	PE	TCH	TBH
σ_b^2	0,0755	0,0002	0,0039	0,0068	0,1249	25,4958	0,6288
σ_f^2	0,1982	0,0262	0,0300	0,2756	0,7627	155,6569	5,7378
σ_d^2	3,1902	0,2068	0,1622	7,6583	11,6272	2372,9067	88,5115
σ_e^2	0,2773	0,0004	0,0053	0,0000	0,0796	16,2410	1,1405
σ_g^2	0,1254	0,0045	0,0278	0,2124	0,6526	133,1759	4,8101
h_{entre}^2	0,63	0,93	0,93	0,77	0,85	0,85	0,84
h_{dentro}^2	0,04	0,12	0,17	0,03	0,06	0,06	0,05
%CVe	2,7267	1,0165	3,5163	-	4,8356	4,8356	8,3068
%CVg	1,8341	7,5771	8,0586	7,9010	13,8471	13,8471	17,0150
CVg/CVe	0,6726	7,4534	2,2917	-	2,8635	2,8635	2,0536

$\hat{\sigma}_b^2$: variância de blocos, $\hat{\sigma}_f^2$: variância fenotípica entre médias de famílias, $\hat{\sigma}_d^2$: variância fenotípica dentro da família, $\hat{\sigma}_e^2$: variância ambiental entre médias de famílias, $\hat{\sigma}_g^2$: variância genotípica entre e dentro das famílias, h_{entre}^2 : herdabilidade entre, h_{dentro}^2 : herdabilidade dentro, %CVe: coeficiente de variação experimental, %CVg: coeficiente de variação genético, CVg/CVe: relação CVg/CVe

Os valores de %CVe apresentados para todas as variáveis (Tabela 2) podem ser considerados baixos em comparação com alguns valores apresentados pela literatura. Bressiani et al. (2002), estudando a interação genótipo x ambiente também em famílias de cana-de-açúcar, avaliaram as mesmas variáveis aqui apresentadas e encontraram valores de 1,49% a 11,81% para o coeficiente de variação e concluíram que estes valores evidenciaram boa precisão experimental.

Bastos (2001) apresentou valores de 1,91% a 10,11% para o mesmo parâmetro em variáveis semelhantes às avaliadas neste estudo. Esta mesma autora, em 2005, mostrou ainda que os maiores valores de %CVe são estimados para TCH e TBH. De acordo com Jackson et al. (1995), os coeficientes de variação residual para TBH variaram de 14,2% a 23,1% e para TCH, entre 12,9% e 22,9%. Já Erazzú et al. (1996) obtiveram estimativas deste parâmetro que variaram entre 8,90% e 25,56% para produtividade de açúcar.

Em concordância com o exposto acima, os maiores valores encontrados neste trabalho para %CVe foram 4,83% e 8,31% correspondentes à TCH e TBH, respectivamente, e a característica de menor %CVe foi DC, com 1,02%, porém, para os autores citados acima, a variável que apresenta menores valores de %CVe é Brix. Estes resultados sugerem que o conjunto de dados utilizado apresenta boa precisão experimental, podendo assim ser considerado como base para um estudo de amostragem.

Os altos valores encontrados para herdabilidade entre as famílias indicam que a maior parte da variação fenotípica se deve à variação dos efeitos de família (Tabela 2). Semelhante ao encontrado por Bastos (2001), as variáveis que apresentaram maior e menor estimativa de herdabilidade entre as famílias foram DC e Brix, respectivamente. Já as estimativas para a herdabilidade dentro das famílias foram bem baixas, mostrando que a variância fenotípica dentro das famílias é superior à variância genotípica também dentro das famílias. Isso concorda com o fato de que apenas seleção individual seria uma estratégia pouco eficiente neste fase.

Como em experimentos com informação individual dentro das famílias o quadrado médio dentro representa a variância fenotípica dentro da família,

logo, é de se esperar que as estimativas deste quadrado médio sejam altas, principalmente nas gerações iniciais, devido a grande variabilidade nas famílias. Tal fato tem sido constatado em outros experimentos utilizando famílias (SOUZA et al., 2000).

4.2. Estudo das estimativas dos parâmetros

Foi observada a frequente estimação de componentes de variância negativos para a variância de blocos e variância ambiental entre as famílias, na maioria das variáveis, sendo que para DC e NC, estas estimativas se repetiram em todos os tamanhos de amostra propostos (Tabela 3). Como a variância de blocos não representa grande expressão, esta não será discutida. Valores negativos para variância ambiental ocorreram porque nestes casos as estimativas do QM dentro das famílias superou a estimativas do QM entre as famílias.

Como já exposto anteriormente, é de se esperar que em experimentos de famílias, sendo estas gerações iniciais, a variação dentro das famílias devido à segregação mostra-se tão grande que supera a variação entre as famílias, gerando estimativas negativas do componente de variância ambiental. Entretanto, problemas de amostragem também podem causar estimativas negativas.

De acordo com Barbin (1993), a presença de estimativas negativas de componentes de variância pode indicar uso de modelo matemático errado, pode ser devida ao método de estimação utilizado ou pode estar ligado ao tamanho da amostra. Este mesmo autor ainda mostra que, quando as estimativas de QM dentro superam as de QM entre, pode ter ocorrido um problema de amostragem, e uma amostra maior deve ser preferida.

Em estudo semelhante com trevo vermelho, Xie e Mosjidis (1997) também relataram a frequente estimação de componentes de variância negativos à medida que se diminuía o número de plantas amostradas para o cálculo da herdabilidade. Um pouco mais tarde, em 1999, estes mesmos autores citam o mesmo problema na definição do tamanho de amostra para estimação da correlação genética na mesma cultura e o atribuíram a problemas de amostragem.

Tabela 3 – Número de estimativas negativas encontradas para a variância de blocos (σ_b^2) e variância ambiental (σ_e^2) em cada tamanho de amostra para as variáveis Brix, diâmetro do colmo (DC), comprimento do colmo (CC), número de colmos (NC), peso estimado de colmos (PE), tonelada de cana por hectare (TCH) e tonelada de brix por hectare (TBH)

	Nº de plantas	σ_b^2	σ_e^2		Nº de plantas	σ_b^2	σ_e^2
Brix	4	31	2	PE	4	120	184
	6	9	0		6	53	169
	8	1	0		8	26	132
	10	1	0		10	8	95
	12	0	0		12	3	81
	14	0	0		14	0	21
	16	0	0		16	0	8
	18	0	0		18	0	0
DC	4	249	251	TCH	4	120	184
	6	246	243		6	53	169
	8	232	231		8	26	132
	10	200	252		10	8	95
	12	206	215		12	3	81
	14	168	202		14	0	21
	16	133	179		16	0	8
	18	66	150		18	0	0
CC	4	28	100	TBH	4	146	172
	6	3	36		6	84	145
	8	0	18		8	43	94
	10	0	0		10	27	58
	12	0	0		12	12	42
	14	0	0		14	1	6
	16	0	0		16	0	1
	18	0	0		18	0	0
NC	4	282	241				
	6	261	250				
	8	240	245				
	10	265	215				
	12	229	221				
	14	234	181				
	16	217	174				
	18	176	156				

Pode ser observado na Tabela 3 que amostras a partir de 6 plantas para a variável Brix não apresentaram mais estimativas negativas para σ_e^2 . Já para DC, amostras de 18 plantas não são suficientes para evitar tais estimativas. Este resultado pode ter ocorrido devido à maneira a qual a amostra é feita no campo, onde um colmo por planta é amostrado para DC.

Este fato pode mostrar que a medida de diâmetro de apenas um colmo não foi representativa do diâmetro de todos os colmos da touceira.

Por outro lado, como o número de 18 famílias que foi avaliado neste trabalho ainda é pequeno; para as variáveis que realmente apresentam grande variabilidade dentro das famílias, estimativas de QM dentro maiores que QM entre podem ser frequentes, sem haver problemas de amostragem.

É o que parece ocorrer com DC e NC, que obtiveram o maior número de estimativas negativas para σ_e^2 (Tabela 3). Verifica-se que estas duas variáveis possuem as menores estimativas de %CVe (Tabela 2), indicando boa precisão experimental; DC apresenta a maior estimativa de %CVg, o que indica grande variabilidade dentro das famílias, e DC, seguida de NC, atingiram os maiores valores de CVg/CVe acima de uma unidade (Tabela 2), mostrando que a maior parte da variação que está ocorrendo dentro das famílias, nestas duas variáveis, é de natureza genética e não ambiental.

Para auxiliar na tomada de decisão de qual o tamanho da amostra ideal a ser recomendado, estão apresentados nas figuras de 1 a 12, gráficos que mostram o comportamento das estimativas de alguns dos parâmetros mais importantes, dentre os estimados neste estudo, sendo eles: Variância Genotípica (σ_g^2), Variância Fenotípica entre as famílias (σ_f^2), Variância Fenotípica dentro das famílias (σ_d^2), Variância Ambiental (σ_e^2), Herdabilidade entre as famílias (h_{entre}^2) e Herdabilidade dentro das famílias (h_{dentro}^2).

Na Figura 1 é mostrada a dispersão das estimativas dos parâmetros para a variável Brix. As linhas tracejadas indicam os limites superior e inferior dos intervalos de confiança construídos para os valores dos parâmetros estimados na população. Estes valores estão evidenciados nos gráficos, nos pontos referentes às amostras de tamanho 20 plantas. Uma tendência geral mostra que a variação dentro das 500 estimativas dos parâmetros aumenta a medida que o tamanho da amostra diminui (Figura 1).

Para a variável Brix, a média das 500 estimativas dos parâmetros em cada tamanho de amostra tende a se situar próximo ao valor paramétrico, para a maioria dos parâmetros, exceto para herdabilidade entre, onde esta

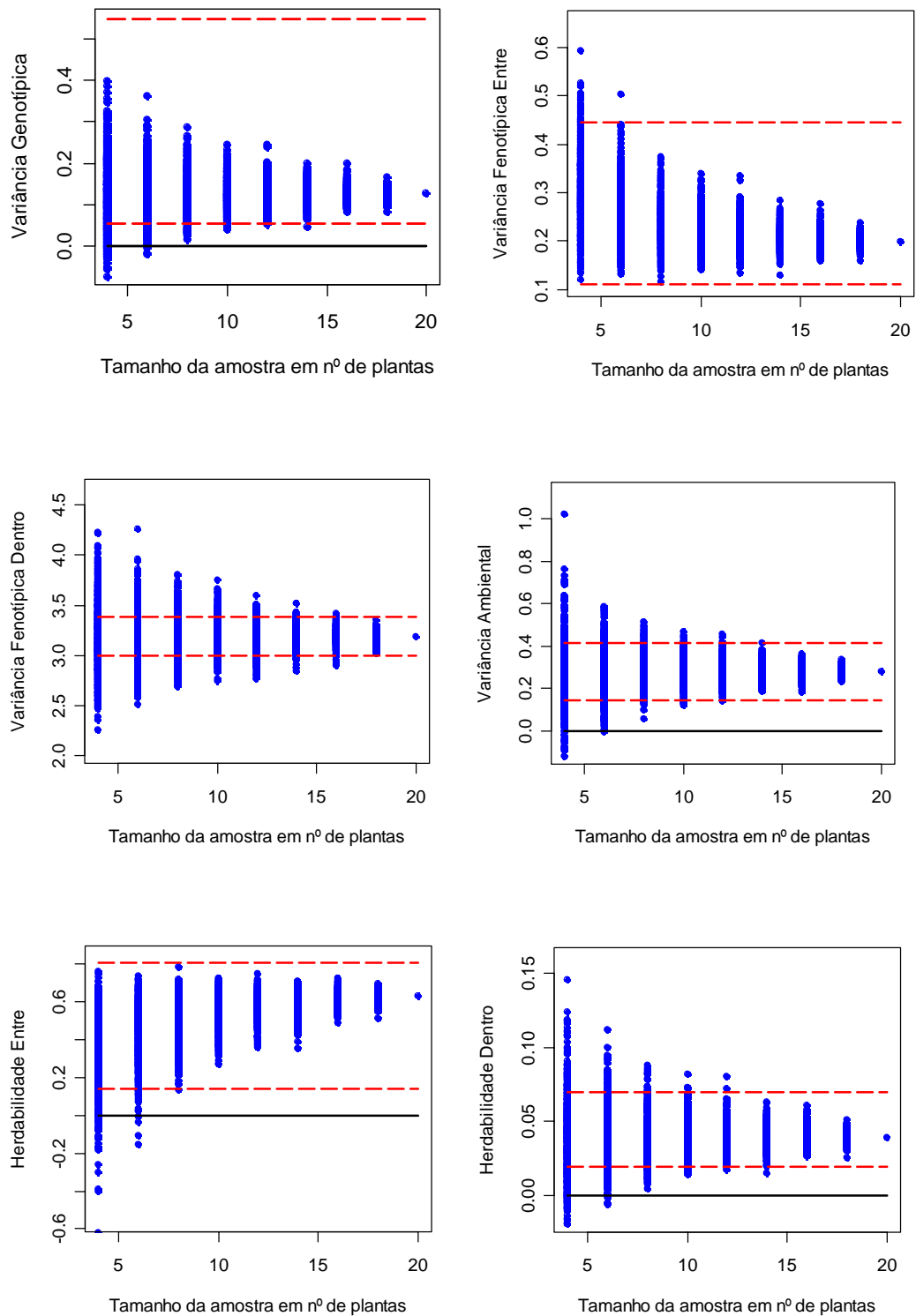


Figura 1 – Dispersão das 500 estimativas de Variância Genotípica, Variâncias Fenotípicas entre e dentro, Variância Ambiental, Herdabilidades entre e dentro, em cada tamanho de amostra, para a variável Brix.