

EFEITOS DAS TRANSFORMAÇÕES DE DADOS SOBRE O COEFICIENTE DE VARIAÇÃO¹

Débora Ferreira da Cunha²

José Carlos Seraphin³

Francisco José P. Zimmermann⁴

RESUMO

A transformação de dados é um processo auxiliar utilizado junto à análise de variância, para torná-la válida quando uma ou mais de suas hipóteses básicas não se verificam. No entanto, mesmo quando feita corretamente, pode trazer inconvenientes como alteração nos parâmetros analisados. O coeficiente de variação (CV) é um destes parâmetros e apresenta, como desvantagem, grande sensibilidade à transformações de dados. Neste trabalho foi estudado o efeito das transformações de potência presentes na Escala de Transformações de Tukey, sobre o CV, para 110 experimentos do arroz de sequeiro, classificados de acordo com tipo de ensaio e região de cultivo, considerando a variável produção de grãos e suas respectivas transformações. Foram usados os testes de Lilliefors e de Cochran para testar, respectivamente, a distribuição dos dados quanto à normalidade, e a homogeneidade de variância dos tratamentos. As transformações mais eficientes foram a raiz quadrada e a logarítmica, ambas reduzindo os valores do CV, corrigindo a falta de normalidade e/ou a heterogeneidade de variância quando havia necessidade de transforma-

1 Parte da Monografia de Especialização em Estatística do primeiro autor, apresentada na Universidade Federal de Goiás (UFG/GO). Jan/2000.

2 Na época, pós-graduanda em Estatística, Universidade Federal de Goiás – UFG, Goiânia-GO.

3 Instituto de Matemática e Estatística, Universidade Federal de Goiás – UFG, Caixa Postal 131, Goiânia-GO, Brasil.

4 Embrapa Arroz e Feijão (CNPAF), Caixa Postal 179, Goiânia-GO.

ção. Em geral os experimentos que apresentaram menores CV foram os do grupo Ensaio Comparativo Avançado, juntamente com aqueles conduzidos na região Centro-Oeste.

Palavras-chave: transformação de dados, coeficiente de variação, arroz de sequeiro, normalidade, heterogeneidade de variâncias.

ABSTRACT

EFFECTS OF TRANSFORMATIONS OF DATA ON THE COEFFICIENT OF VARIATION

Transformation of data is an auxiliary process used with the analysis of variance to turn it valid when one or more of the basic hypotheses could not be verified. Although, even if calculated correctly, it can bring the inconvenient of change in the analyzed parameters. The coefficient of variation (CV), is one of these parameters and presents as disadvantage the great sensibility to transformations of the data. In this work was studied the effect of power transformations from Tukey Transformations Scale, on CV, for 110 experiments of upland rice, classified by trial type and sow regions, through the variable yield and its transformations. The Lilliefors and Cochran tests were used to assess the normal distribution of the data and the variance homogeneity of the treatments. The most efficient transformations were the square root and the logarithms, both reducing the values of CV, correcting the lack of normality and/or variance heterogeneity when transformations were needed. In general, the experiments with small CV were those from the Advanced Comparative Trial group together with those from Central region.

Key words: transformation of data, coefficient of variation, upland rice, normality, heterogeneity of variances.

INTRODUÇÃO

A análise da variância é uma técnica estatística desenvolvida por Fisher (1925) para sistematizar e facilitar a análise e interpretação de dados experimentais. Para que a análise e os testes de significância aplicados tenham validade, necessita-se que as condições básicas da análise de variância sejam atendidas. Há um consenso generalizado de que quando isso não ocorre, a solução mais recomendada é a transformação dos dados originais (Cochran, 1947; Mead & Curnow, 1983; Gomez & Gomez, 1983).

Geralmente o pesquisador sabe historicamente quais variáveis precisam ser transformadas, por isso, muitas vezes aplicar transformações conhecidas antes de qualquer inspeção prévia dos dados. Fisher & Vanbelle (1983) recomendam que não se façam transformações desnecessárias e que se tenha muito cuidado, particularmente com estimações. Portanto, ao empregar uma transformação deve-se ter uma razão lógica, o que nem sempre ocorre, levando às vezes a utilizar transformações indiscriminadamente, sem nenhum critério ou preocupação de sua validade. Assim ao considerar-se uma transformação, deve-se verificar se os benefícios justificam o esforço necessário e a inconveniência que ela pode causar.

As transformações raiz quadrada, logarítmica, angular e inversa têm sido aplicadas a diferentes tipos de dados, para observação de seus efeitos sobre os resultados da análise de variância, e só devem ser empregadas se os dados violarem as suposições de normalidade e homogeneidade de variância (Bartlett, 1947; Demétrio, 1978 e Bowman & Watson, 1997). Quando não há razão teórica para se escolher uma transformação específica, Mead & Curnow (1983) sugerem ao experimentador selecionar uma transformação empiricamente e discutem como escolher a melhor transformação. Sokal & Rohlf (1995) recomendam as transformações Box-Cox em vez de tentar várias transformações para descobrir a melhor delas. Box & Cox (1964) desenvolveram um método para escolher a melhor transformação dentro da

família de transformações de potência. Uma discussão sobre esta família de transformações é apresentada em Emerson & Stoto (1983). Baseando-se também na família de transformações de potência, Tukey (1977) propôs uma Escala de Transformações, eficiente na maioria dos casos.

Após a análise dos dados transformados, devem-se verificar com precaução os resultados obtidos, pois, parâmetros, como o coeficiente de variação (CV), o coeficiente de determinação (R^2), o quadrado médio do resíduo (QMR), além de testes aplicados, podem ser influenciados conforme o tipo de transformação usada. Dentre estes, o coeficiente de variação, é uma medida de variação relativa e apresenta, como desvantagem, grande sensibilidade às transformações de dados. Por isso alguns autores o consideram de uso limitado, embora transformações possam resultar em sua melhoria. Segundo Sampaio (1998) a presença de CV's muito altos, pode ser percebida em situações onde a análise de variância não tem suas hipóteses garantidas e, na maioria das vezes, as variáveis que demandam transformação apresentam alterações, responsáveis pela elevação do CV. Em estudo recente sobre o desempenho de medidas de validade, Bowman & Watson (1997), examinaram os efeitos de transformações de dados sobre o CV e R^2 com dados simulados e concluíram que o R^2 pode ser uma medida estatística alternativa para avaliar o desempenho de experimentos.

Pouca ou quase nenhuma atenção tem sido reservada ao estudo do efeito das transformações sobre o coeficiente de variação. O que vem sendo estudado são critérios de classificação do CV para experimentos agrícolas, como pode ser visto em Steel & Torrie (1980), Pimentel-Gomes (1985), Estefanel *et al.* (1987), Garcia (1989) e mais recentemente em SCAPIM *et al.* (1995) e Amaral *et al.* (1997). O presente trabalho teve por objetivo verificar os efeitos das transformações sobre os resultados obtidos com o coeficiente de variação, com dados da produção de grãos do arroz de sequeiro.

MATERIAL E MÉTODOS

Material

Foram utilizados, para análise, dados de 110 experimentos relativos a produção de grãos (kg/ha) do arroz de sequeiro, cedidos pela Embrapa Arroz e Feijão. Esses experimentos foram conduzidos no período de 1991 a 1998, classificados de acordo com o tipo de ensaio (Ensaio Comparativo Avançado-ECA, Ensaio Comparativo Preliminar-ECP, Melhoramento Populacional, Ensaio de Densidade Populacional e Outros) e a região de cultivo, Norte (AC, AM, AP, RO); Nordeste (MA, PI, CE) e Centro-Oeste (GO, MT, MS, TO, MG). Os experimentos não foram realizados especialmente para atender os objetivos deste trabalho. Para efeito deste estudo, foram agrupados de acordo com a Tabela 1.

Tabela 1. Distribuição dos experimentos de acordo com o tipo de ensaio e região de cultivo

Grupos	Número de experimentos
Geral	110
- Regiões	
Centro-Oeste	78
Outras	32
- Ensaio	
ECA	56
ECP	30
Outros	24

Métodos

Considerando a natureza da variável (x) analisada (produção de grãos), foram aplicadas transformações de potência aos dados originais, usando uma modificação da Escala de Transformação de Tukey (1977). Obtiveram-se seis variáveis para análise; P1 – logaritmo de x ; P2 – \sqrt{x} ; P3 – x^2 ; P4 – x^{-1} ; P5 – $x^{-1/2}$; e P6 – x^{-2} (Figura 1). A partir destas, além de obter as estimativas do coeficiente de variação, avaliaram-se as condições de normalidade e

homogeneidade de variância dos dados, através dos testes de Lilliefors e Cochran, para interpretação dos efeitos das transformações de dados sobre o CV, independente da necessidade ou não de uma transformação (Figura 1).

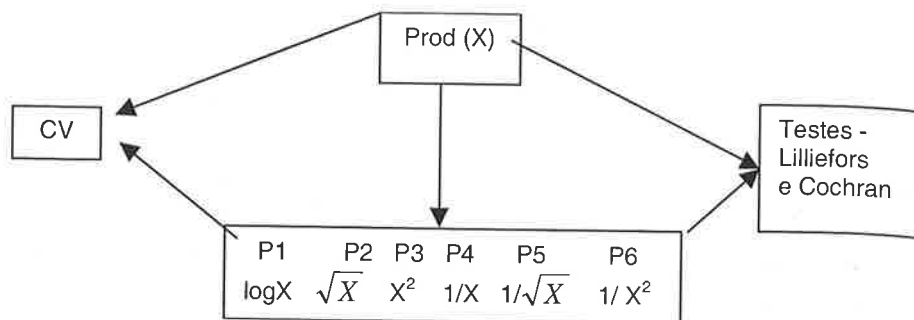


Figura 1. Esquema dos métodos utilizados. Prod = Produção.

RESULTADOS E DISCUSSÃO

Os resultados obtidos são apresentados, primeiro em termos gerais e depois classificados de acordo com a região de cultivo (Centro-Oeste e Outras) e o tipo de ensaio (ECA, ECP e Outros), para a variável produção de grãos e suas transformações.

As estimativas da média do CV são apresentadas na Tabela 2, e baseando-se no critério de classificação do CV de Pimentel-Gomes (1985), para a variável sem transformação x (produção), classificam-se em altas, e para as transformadas em baixas, médias e muito altas.

A acentuada variação entre as estimativas para as transformações e a variável original (Tabela 2), ocorreu devido aos tipos de transformações de potência (x^p) empregados. Percebeu-se uma diminuição nas estimativas do CV após as transformações com p variando entre $-1/2$ e $1/2$; estas transformações reduziram a amplitude e a variação entre os dados. Por outro lado, verificou-se uma elevação das estimativas do CV após as transformações com $p \leq -1$ ou $p = 2$, as transformações inversas reduziram a amplitude de variação

Tabela 2. Estatísticas descritivas do coeficiente de variação (CV), de acordo com o grupo, para cada uma das variáveis antes e após as transformações.

Grupo	Estatísticas do CV (%)	Variáveis						
		PROD	P1	P2	P3	P4	P5	P6
GERAL	Média	21,14	3,65	11,34	40,39	43,41	16,01	111,45
	Desvio Padrão	9,21	2,67	5,45	17,48	63,07	16,18	130,80
	Máximo	51,36	19,13	35,17	95,17	434,98	139,13	718,36
	Mínimo	3,78	0,52	1,90	7,75	4,21	2,00	9,47
Região C.O.	Média	20,05	3,62	10,88	38,08	46,60	16,36	115,56
	Desvio Padrão	9,30	3,01	5,72	17,21	73,33	18,63	143,90
	Máximo	51,36	19,13	35,17	95,17	434,98	139,13	718,36
	Mínimo	3,78	0,52	1,90	7,75	4,21	2,00	9,47
Outras Regiões	Média	23,80	3,74	12,47	46,04	35,63	15,16	101,43
	Desvio Padrão	8,40	1,56	4,54	16,83	21,97	7,26	90,52
	Máximo	43,30	7,49	23,03	87,91	103,25	36,57	449,38
	Mínimo	12,22	1,77	6,27	23,79	14,26	6,77	31,22
ECA	Média	19,53	2,85	10,15	37,92	27,92	11,99	78,73
	Desvio Padrão	7,30	1,24	3,81	14,77	17,11	5,57	73,66
	Máximo	40,60	6,26	19,60	88,13	101,37	29,99	449,38
	Mínimo	3,78	0,52	1,90	7,75	4,21	2,00	9,47
ECP	Média	22,20	3,54	11,57	43,49	33,10	14,07	94,29
	Desvio Padrão	10,12	1,89	5,45	20,09	21,73	7,92	80,66
	Máximo	46,90	7,99	24,27	95,17	103,25	36,57	346,04
	Mínimo	6,62	0,84	3,25	13,88	6,66	3,24	15,75
Outros Ensaios	Média	23,59	5,66	13,85	42,30	92,44	27,80	209,26
	Desvio Padrão	11,31	4,46	7,62	19,28	120,27	30,09	217,13
	Máximo	51,36	19,13	35,17	83,57	434,98	139,13	718,36
	Mínimo	11,07	1,37	5,31	19,89	11,34	5,54	23,90

PROD – produção; P1 – logaritmo da produção; P2 – raiz quadrada da produção; P3 – quadrado da produção; P4 – inverso da produção; P5 – inverso da raiz quadrada da produção; P6 – inverso do quadrado da produção.

dos dados, e tornaram as médias das variáveis muito pequenas, enquanto que a transformação x^2 aumentou a amplitude e a média dos dados, afetando muito as estimativas do CV nestes casos. Essa sensibilidade do CV a transformações de dados, podendo ser elevado ou reduzido, de acordo com a transformação aplicada, confirma os resultados de Mandel (1964) e Bowman & Watson (1997). As transformações que produziram a maior redução do CV

foram \sqrt{x} , $\log x$ e $x^{-1/2}$ (Tabela 2), concordando com Tukey (1977) que considerou estas transformações como as mais importantes de sua Escala.

Em relação aos grupos analisados, poucas diferenças foram observadas nas estimativas dos CV, entretanto, maiores diferenças foram percebidas entre as variáveis analisadas, como em Estefanel *et al.* (1987). Deve-se ressaltar que se obtiveram estimativas de CV menores, no grupo ECA e região Centro-Oeste (Tabela 2), o que pode ser explicado pelo fato do ECA ser a etapa final do processo de seleção de linhagens, e por serem os Estados da região Centro-Oeste mais apropriados para o cultivo do arroz, no sistema de sequeiro, principalmente pela condição climática favorável (Ferreira *et al.* 1983).

Maiores detalhes dos efeitos das transformações sobre o CV são observados, considerando as hipóteses de normalidade da distribuição dos dados e homogeneidade de variância. Desta forma, os resultados são apresentados para quando as duas condições foram atendidas, uma delas foi atendida, as duas não foram atendidas (Tabelas 3 e 4) e pelo menos uma não foi atendida (Tabela 5), em relação à variável original.

A Tabela 3 apresenta o número de experimentos em que as hipóteses de normalidade da distribuição dos dados e homogeneidade de variância são atendidas para variável PROD. Aplicadas as transformações os resultados tornaram-se desfavoráveis, mostrando com isso que pode haver mudanças nas conclusões se for feita uma transformação desnecessária. Verificaram-se menos influências sobre a normalidade dos dados e a homogeneidade de variância com aplicação das transformações raiz quadrada, quadrado e logarítmica respectivamente.

Como se pode notar na Tabela 3, foram poucos os casos de experimentos que atendiam à condição de normalidade e não atendiam à de homogeneidade de variância. Essa baixa frequência de experimentos pode ser devida ao emprego correto das técnicas experimentais e condições apropriadas para cultivo. Neste caso, ocorre heterogeneidade entre os tratamentos, que talvez possa ser denominada irregular (Gomez & Gomez, 1983; Banzatto & Kronka, 1989). Os métodos mais apropriados e indica-

Tabela 3- Distribuição do número de experimentos por grupo, para cada variável após a transformação, para os experimentos que inicialmente atenderem ou não às condições de normalidade e homogeneidade de variância, para variável produção.

Grupo	Condições		Normalidade e Homogeneidade						Não-Normalidade e Heterogeneidade																				
	N	H	Prod P1	P2	P3	P4	P5	P6	Prod P1	P2	P3	P4	P5	P6	Prod P1	P2	P3	P4	P5	P6									
Gera	NS	NS	74	42	64	48	16	22	2	-	1	3	-	-	-	9	10	3	3	5	1	-	3	4	1	-	1		
	NS	S	-	5	2	5	3	6	1	7	1	1	3	-	1	-	2	-	1	2	3	-	-	-	-	1	-	-	
	S	NS	-	22	8	18	26	29	20	-	2	-	3	2	1	23	10	12	16	6	9	4	-	2	-	2	2	1	1
	S	S	-	5	-	3	29	17	51	-	3	3	4	4	6	-	2	1	3	12	6	18	6	1	2	2	4	4	5
Região	NS	NS	53	31	45	36	11	18	1	-	-	-	-	-	-	8	9	1	3	4	1	-	2	3	1	-	1	-	
	NS	S	-	4	1	1	3	3	1	3	1	1	1	-	1	-	1	-	-	2	3	-	-	-	-	-	-	-	-
	S	NS	-	14	7	14	18	19	13	-	-	-	1	-	-	17	7	8	13	4	8	2	-	2	-	2	2	1	1
	S	S	-	4	-	2	21	13	38	-	2	2	2	2	3	-	1	-	3	8	2	14	5	1	2	2	3	3	4
Outras	NS	NS	21	11	19	12	5	4	1	-	1	3	-	-	-	1	1	2	-	1	-	-	1	1	-	-	-	-	
	NS	S	-	1	1	4	-	3	-	4	-	2	-	-	-	1	-	1	-	-	-	-	-	-	-	1	-	-	
	S	NS	-	8	1	4	8	10	7	-	2	-	2	2	1	6	3	4	3	2	1	2	-	-	-	-	-	-	-
	S	S	-	1	-	1	8	4	13	-	1	1	2	2	3	-	1	1	-	4	4	1	-	-	-	1	1	1	
ECA	NS	NS	41	26	36	25	11	14	1	-	1	-	-	-	-	6	6	2	2	4	1	-	1	1	-	-	-	-	
	NS	S	-	3	1	4	1	4	-	3	1	-	2	-	-	1	-	-	1	-	1	-	-	-	-	1	-	-	
	S	NS	-	10	4	10	13	15	12	-	-	-	2	1	-	10	2	3	6	3	2	4	-	-	-	-	1	-	
	S	S	-	2	-	2	16	8	28	-	2	2	1	1	2	3	-	1	1	2	5	3	5	2	1	1	1	2	2
ECP	NS	NS	24	12	21	17	3	5	-	-	2	-	-	-	-	1	1	1	1	1	-	-	1	1	-	-	-	-	
	NS	S	-	2	1	1	2	1	-	2	-	1	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	
	S	NS	-	9	2	5	10	12	6	-	2	-	1	1	1	3	2	2	1	1	1	-	-	-	-	1	1	-	
	S	S	-	1	-	1	9	6	18	-	-	1	1	1	1	-	-	-	-	1	1	3	1	-	-	-	-	-	
Outros	NS	NS	9	4	7	6	2	3	1	-	-	-	-	-	-	2	3	-	-	-	-	-	1	2	1	-	-		
	NS	S	-	-	-	-	1	1	-	2	1	1	-	1	-	1	-	-	2	2	-	-	-	-	-	-	-		
	S	NS	-	3	2	3	3	2	2	-	-	-	-	-	-	10	6	7	9	2	6	-	-	2	-	1	-		
	S	S	-	2	-	4	3	5	-	1	1	2	2	1	2	-	1	-	1	6	2	10	3	-	1	1	3	2	3

Prod - produção; P1 - logaritmo da produção; P2 - raiz quadrada da produção; P3 - quadrado da produção; P4 - inverso da produção; P5 - inverso da raiz quadrada da produção; P6 - inverso do quadrado da produção. NS - não significativo a 1% de probabilidade pelos Testes de Lilliefors ou Cochran. S - significativo a 1% de probabilidade pelos testes de Lilliefors ou Cochran. N - normalidade H - homogeneidade de variância.

dos para este tipo de falha são a subdivisão da variância residual ou a eliminação de parte do experimento (Cochran, 1947; Gomez & Gomez, 1983). Após as transformações ocorreu uma pequena melhora dos resultados, sendo que a melhor transformação foi a raiz quadrada (\sqrt{x}), e a pior, o inverso do quadrado (x^{-2}). Isso pode ser um indicativo de que a transformação de dados não é a técnica mais indicada, talvez um método não-paramétrico seja mais adequado nestes casos (Conover, 1980). Sabe-se que o valor do teste F da análise de variância, é afetado pela heterogeneidade de variância associada a dados provenientes da distribuição normal (Cochran, 1947; Zimmermann, 1987).

Em relação aos experimentos originalmente com não-normalidade e homogeneidade de variância (Tabela 3), considerando o controle das condições, as transformações apresentaram-se bem mais eficientes que o caso anterior; sendo a raiz quadrada e a logarítmica, as melhores, e o inverso do quadrado, a pior. Essa situação é menos importante, pois a robustez do teste F, é menos afetada se acompanhada de homogeneidade de variância Zimmermann (1987) mostrou que o tamanho e o poder do teste F, não são afetados quando os dados têm variância homogênea independentemente da distribuição de probabilidade.

Também foram analisados os experimentos que não atendiam às condições de normalidade e homogeneidade de variância (Tabela 3), denominada heterogeneidade regular (Gomez & Gomez, 1983; Banzatto & Kronka, 1989). Verificou-se nessa situação que as transformações logarítmica e raiz quadrada foram as mais eficientes, concordando com Cochran (1947) e Gomez & Gomez (1983), que afirmam que a aplicação de uma transformação adequada, quando a heterogeneidade é regular produz bons resultados.

Quando as condições de normalidade e homogeneidade de variância são atendidas, e aplica-se uma transformação indiscriminadamente, observam-se modificações desfavoráveis tanto na forma da distribuição dos dados, como em relação à homogeneidade de variância. Mudanças menores são observadas após a transformação raiz quadrada, mas

após a transformação inverso do quadrado, verificaram-se mudanças drásticas dos experimentos que atendiam às duas condições. Por outro lado, quando pelo menos uma das condições não é atendida, após as transformações percebe-se uma melhoria dos resultados, mas em todos os casos grande parte dos experimentos continuaram a não atender pelo menos a uma das condições, especialmente após as transformações inversas. As transformações mais eficientes foram a raiz quadrada e a logarítmica, a transformação raiz quadrada reduziu as estimativas do CV e aumentou o número de experimentos que atendiam às duas condições para todos os grupos (Tabela 5). Com isso, observou-se que transformações de dados só devem ser aplicadas, se uma ou mais das hipóteses da análise de variância não são satisfeitas. Isto confirma que, quando as condições da análise de variância estiverem satisfeitas, deve-se evitar a transformação dos dados, não só por questões práticas, mas também por inversões na ordenação das médias que ela pode trazer, como em Demétrio (1978).

Em relação às estimativas da média dos CV, os experimentos que atendiam às duas condições apresentaram médias inferiores aos que não atendiam à pelo menos uma condição, mesmo após as transformações (Tabelas 4 e 5). Situações onde a análise de variância fica inviabilizada, se denunciam por apresentarem coeficientes de variação muito elevados (Sampaio, 1998). Destaca-se uma grande diferença das estimativas das médias dos CV entre as duas situações, principalmente, após a transformação inverso do quadrado (Tabela 5). Isto confirma novamente que as estimativas dos CV apresentaram-se extremamente sensíveis à transformações.

Em torno de 65% a 80% dos experimentos analisados, de acordo com o grupo, tiveram as duas condições atendidas, exceto o grupo Outros Ensaio. Neste caso, a aplicação de qualquer das transformações de potência mostrou inadequada. Para os outros casos, onde pelo menos uma condição não é atendida, as transformações foram mais eficientes em experimentos que tinham heterogeneidade regular e não-normalidade associada a homogeneidade de variância, e menos eficientes em experimentos com normalidade e heterogeneidade de variância, recomendando-se um controle maior das cultivares e da região de cultivo.

Tabela 4 - Distribuição dos valores médios dos CV por grupo, para cada variável após a transformação, para os experimentos que inicialmente atendem ou não às condições de normalidade e homogeneidade de variância, para variável produção.

Grupo	Normalidade e Homogeneidade						Não-Normalidade e Heterogeneidade						Não-Normalidade e Heterogeneidade																			
	N	H	Prod	P1	P2	F6	F5	F4	F3	F2	F3	F4	F5	F6	Prod	P1	P2	F3	F4	F5	F6											
Gral	NS	NS	190	27	97	319	181	89	276	-	20	132	-	-	39	124	338	290	124	393	-	51	178	836	-	92						
	NS	S	-	26	148	426	184	10,7	288	287	191	362	339	-	139,1	-	-	39	-	621	304	148	-	-	-	-	51,3					
	S	NS	-	32	100	460	240	131	407	-	48	-	254	149	682	237	49	133	426	233	238	520	-	95	-	447	276	454	349			
	S	S	-	31	-	546	377	141	101,3	-	64	138	661	1774	240	287,7	-	50	139	644	106,8	239	165,5	287	32	150	391	75,2	248	155,8		
Feijão	NS	NS	173	27	90	234	161	84	239	-	-	-	-	-	32	123	139	290	122	393	-	54	193	836	-	92						
	NS	S	-	21	124	361	184	11,7	288	329	191	362	286	-	139,1	-	-	38	-	304	148	-	-	-	-	-	-	-	-			
	S	NS	-	23	91	398	202	11,7	382	-	-	-	168	-	-	235	53	134	436	233	255	559	-	95	-	447	276	454	349			
	S	S	-	32	-	646	383	135	99,0	-	76	142	661	3169	328	479,5	-	54	-	644	128,0	263	172,1	293	32	150	391	87,5	276	167,5		
Outras Feijões	NS	NS	232	28	115	393	226	114	314	-	20	132	-	-	-	-	36	135	438	-	131	-	-	-	-	-	-	-	-			
	NS	S	-	49	173	437	-	97	-	256	-	-	361	-	-	-	-	39	-	621	-	-	-	-	-	-	-	-	-	-		
	S	NS	-	47	167	677	326	156	455	-	48	-	237	149	682	243	38	132	38,7	234	100	481	-	-	-	-	-	-	-	-		
	S	S	-	27	-	348	360	160	108,1	-	41	130	660	379	152	96,0	-	47	139	-	64,5	227	232,3	260	-	-	-	-	-	-	-	
FCF	NS	NS	187	27	95	315	194	90	314	-	78	-	-	-	-	-	35	120	438	294	120	393	-	43	136	-	-	-	-	-		
	NS	S	-	32	173	455	115	118	-	180	20	-	287	-	-	-	-	39	-	-	157	-	-	-	-	-	-	-	-	-	-	
	S	NS	-	24	100	449	215	121	382	-	-	-	174	85	-	228	24	104	380	197	102	520	-	-	-	-	-	-	-	-	-	
	S	S	-	26	-	344	365	140	87,2	-	30	103	441	507	135	87,5	-	47	139	66,7	423	181	146,1	224	32	105	329	384	156	1122		
FCP	NS	NS	212	29	109	361	173	93	-	-	159	-	-	-	-	-	42	144	139	282	141	-	-	25	100	-	-	-	-	-	-	
	NS	S	-	19	124	309	219	71	-	31,8	-	-	424	-	-	-	-	-	-	621	-	-	-	-	-	-	-	-	-	-	-	
	S	NS	-	45	126	543	286	152	438	-	48	-	413	212	682	241	38	123	594	6,7	32	-	-	-	-	-	-	-	-	-	-	
	S	S	-	41	-	952	433	150	142,5	-	-	-	87,9	252	115	632	-	-	-	-	103,2	366	139,9	210	-	-	-	-	-	-	-	-
Outros Ensaios	NS	NS	144	23	76	244	122	78	239	-	-	-	-	-	-	-	48	124	-	-	-	-	-	-	-	-	-	-	-	-	-	
	NS	S	-	-	-	-	-	-	99	288	419	191	362	-	139,1	-	-	38	-	-	304	143	-	-	-	-	-	-	-	-	-	-
	S	NS	-	19	74	360	199	74	407	-	-	-	-	-	-	246	61	149	439	37,1	31,7	-	-	-	-	-	-	-	-	-	-	-
	S	S	-	30	-	-	340	124	148,7	-	133	207	661	3169	576	700,4	-	54	-	58,7	161,2	263	218,9	365	-	-	-	-	-	-	-	-

Prod - produção; P1 - logaritmo da produção; P2 - raiz quadrada da produção; P3 - quadrado da produção; P4 - inverso da produção; P5 - inverso da raiz quadrada da produção; P6 - inverso do quadrado da produção.
 NS - não significativo a 1% de probabilidade pelos testes de Lilliefors ou Cochran. S - significativo a 1% de probabilidade pelos testes de Lilliefors ou Cochran. N - normalidade H - homogeneidade de variância.

Tabela 5. Distribuição dos valores médios dos CV e do número de experimentos por grupo, para cada variável após a transformação, para os experimentos que inicialmente atendem as condições de normalidade e homogeneidade de variância e para os que não atendem a pelo menos uma condição, para variável produção.

Grupo	Condições e Resultados		Variáveis							
			PROD	P1	P2	P3	P4	P5	P6	
GERAL	Atende às 2 condições	NS	74	42	64	48	16	22	2	
		S	0	32	10	26	58	52	72	
	Média(%)	19,00	2,88	9,90	36,97	27,88	11,87	81,95		
GERAL	Não atende à pelo menos 1 condição	NS	0	13	17	4	3	6	1	
		S	36	23	19	32	33	30	35	
	Média(%)	25,55	5,24	14,32	47,43	75,32	24,51	172,09		
Região C.O.	Atende às 2 condições	NS	53	31	45	36	11	18	1	
		S	0	22	8	17	42	25	52	
	Média(%)	17,34	2,59	9,05	33,65	26,44	11,02	81,31		
Região C.O.	Não atende à pelo menos 1 condição	NS	0	10	12	2	3	5	1	
		S	25	15	13	23	22	20	24	
	Média(%)	25,82	5,80	14,77	47,46	89,33	27,67	188,18		
Outras Regiões	Atende às 2 condições	NS	21	11	19	12	5	4	1	
		S	0	10	2	9	16	17	20	
	Média(%)	23,21	3,62	12,04	45,34	31,52	14,03	83,58		
Outras Regiões	Não atende à pelo menos 1 condição	NS	0	3	5	2	-	1	-	
		S	11	8	6	9	11	10	11	
	Média(%)	24,93	3,97	13,29	47,38	43,49	17,34	135,52		
ECA	Atende às 2 condições	NS	41	26	36	25	11	14	1	
		S	0	15	5	16	30	27	40	
	Média(%)	18,71	2,68	9,72	36,27	26,17	11,38	71,80		
ECA	Não atende à pelo menos 1 condição	NS	0	7	8	2	2	4	1	
		S	15	8	7	13	13	11	14	
	Média(%)	21,77	3,31	11,33	42,43	32,70	13,67	97,65		
ECP	Atende às 2 condições	NS	24	12	21	17	3	5	-	
		S	0	12	3	7	21	19	24	
	Média(%)	21,22	3,43	11,09	41,42	32,11	13,59	93,45		
ECP	Não atende à pelo menos 1 condição	NS	0	2	4	1	1	2	-	
		S	6	4	2	5	5	4	6	
	Média(%)	26,12	3,98	13,48	51,76	37,07	15,97	97,66		
Outros Ensaios	Atende às 2 condições	NS	9	4	7	6	2	3	1	
		S	0	5	2	3	7	6	8	
	Média(%)	14,42	2,32	7,52	28,30	24,41	9,51	97,53		
Outros Ensaios	Não atende à pelo menos 1 condição	NS	0	3	5	1	-	-	-	
		S	15	12	10	14	15	15	15	
	Média(%)	29,09	7,67	17,65	50,71	133,25	38,78	276,30		

PROD – produção; P1 – logaritmo da produção; P2 – raiz quadrada da produção; P3 – quadrado da produção; P4 – inverso da produção; P5 – inverso da raiz quadrada da produção; P6 – inverso do quadrado da produção.

NS – não significativo a 1% de probabilidade pelos Testes de Lilliefors e Cochran, para as duas condições.

S – significativo a 1% de probabilidade pelos Testes de Lilliefors e Cochran, para pelo menos uma condição.

CONCLUSÕES

1. As transformações logarítmica ($\log x$), raiz quadrada (\sqrt{x}) e inverso da raiz ($x^{-1/2}$), reduzem os valores do CV, enquanto que as transformações quadrado (x^2), inverso (x^{-1}) e inverso do quadrado (x^{-2}) os elevaram.
2. A aplicação de qualquer uma das transformações de potência, para experimentos que atendem as condições de normalidade dos dados e homogeneidade de variância, pode tornar inválida uma dessas condições.
3. As transformações mais eficientes foram a raiz quadrada e a logarítmica, que além de reduzir os valores do CV, corrigiram a falta de normalidade dos dados e/ou a heterogeneidade de variância, quando havia necessidade de transformação.
4. A aplicação da transformação raiz quadrada é a que tem menos probabilidade de gerar resultados inadequados, quando se percebe, através de rápida inspeção dos dados, que uma das hipóteses da análise de variância não se verifica.
5. A transformação raiz quadrada é a que menos altera os resultados, quando se aplica uma transformação indiscriminadamente.

REFERÊNCIAS

- AMARAL, A.M.; J.A. MUNIZ & M. SOUZA., 1997. Avaliação do Coeficiente de Variação como Medida de Precisão na Experimentação com Citros. **Pesquisa Agropecuária Brasileira**, 32 (12): 1221-1225.
- BANZATO, D.A. & S.N. KRONKA., 1989. **Experimentação Agrícola**. Jaboticabal: Funep.
- BARTLETT, M.S., 1947. The Uses of Transformations. **Biometrics**, 3: 39-52.

- BOWMAN, D.T. & C. WATSON., 1997. Measures of Validity in Cultivar Performance Trials. **Agronomy Journal**, **89** (6): 860-866.
- BOX, G.E.P. & D.R. COX., 1964. An Analysis of Transformations. **Journal of the Royal Statistical Society. Séries B**, **26**: 211-252.
- COCHRAN, W.G., 1947. Some Consequences When the Assumptions for the Analysis of Variance are not Satisfied. **Biometrics**, **3**: 22-38.
- CONOVER, W.J., 1980. **Practical Nonparametric Statistics**. New York: John Wiley.
- DEMÉTRIO, C.G.B., 1978. **Transformação de Dados. Efeitos sobre a Análise de Variância**. Piracicaba, ESALQ/USP. (Dissertação de Mestrado).
- EMERSON, J.D. & M.A. STOTO., 1983. **Understanding Robust and Exploratory Data Analysis**. New York: John Wiley.
- ESTEFANEL, V.; I.A.B. PIGNATÁRO & L. STORCK., 1987. Avaliação do Coeficiente de Variação de Experimentos com Algumas Culturas Agrícolas. In: SIMPÓSIO DE ESTATÍSTICA APLICADA À EXPERIMENTAÇÃO AGRONÔMICA, 2. Londrina. **Anais ... Londrina: Univ. Est. de Londrina/Reg. Bras. da Soc. Intern. de Biometria**.
- FERREIRA, M. E.; T. YAMADA & E. MALAVOLTA., 1983. **Cultura do Arroz de Sequeiro**. Piracicaba, Inst. Potassa.
- FISHER, R.A., 1925. **Statistical Methods for Research Workers**. Edingurgh: Oliver & Boyd.
- FISHER, L.D. & G. VANBELLE., 1993. **Biostatistics. A Methodology for the Health Sciences**. Jonh Wiley.
- GARCIA, C.H., 1989. Tabelas para Classificação do Coeficiente de Variação. Piracicaba: Instituto de Pesquisas e Estudos Florestais. 12 p. **Circular Técnica**, 171.
- GOMEZ, K.A. & A.A. GOMEZ., 1983. **Statistical Procedures for Agricultural Research**. New York: John Wiley.
- MANDEL, J., 1964. **The Statistical Analysis of Experimental Data**. New York: Dover Publications.
- MEAD, R. & R.N. CURNOW., 1983. **Statistical Methods in Agriculture and Experimental Biology**. New York: Chapman and Hall.

- PIMENTEL-GOMES, F., 2000. **Curso de Estatística Experimental**. Piracicaba. 14^a ed., Piracicaba.
- SAMPAIO, I.B.M., 1998. **Estatística Aplicada à Experimentação Animal**. Fundação de Ensino e Pesq. em Medicina Veterinária e Zootecnia. Depto de Zootecnia, UFMG.
- SCAPIM, C.A.; C.G.P. CARVALHO. & C.D. CRUZ., 1995. Uma Proposta de Classificação dos Coeficientes de Variação para Cultura do Milho. **Pesquisa Agropecuária Brasileira**, **30** (5): 683-686.
- SOKAL, R.R. & F.J. ROHLF., 1995. **Biometry**. 3^a ed. New York: W. H. Freeman and Company.
- STEEL, R.G.D. & J.H. TORRIE., 1980. **Principles and Procedures of Statistics. A Biometrical Approach**. 2^a edição, New York: McGraw-Hill Book.
- TUKEY, J.W., 1977. **Exploratory Data Analysis**. Addison-Wesley Publ. Co., Reading. Mass.
- ZIMMERMANN, F.J.P., 1987. Efeito de Heterogeneidade de Variância e Distribuição de Probabilidade dos Dados sobre o Poder e Tamanho do Teste F. **Pesquisa Agropecuária Brasileira**, **22** (11/12): 1209-1213.