

# METODOLOGIA E TÉCNICAS EXPERIMENTAIS

## EXCESSO DE ZEROS NAS VARIÁVEIS OBSERVADAS: ESTUDO DE CASO EM EXPERIMENTO COM BRÓCOLIS <sup>(1)</sup>

ALESSANDRO DAL'COL LÚCIO <sup>(2)</sup>; MARA RÚBIA MACHADO COUTO <sup>(3)</sup>;  
JORGE NADIR TREVISAN <sup>(2)</sup>; GUSTAVO ADOLFO KLIPPEL MARTINS <sup>(2)</sup>; SIDINEI JOSÉ LOPES <sup>(2)</sup>

### RESUMO

Em olericultura o interesse é a produção de produtos comercializáveis. Nesses experimentos, comumente, são avaliadas variáveis tais como a fitomassa das cabeças e ramificações comerciais e não comerciais da planta. No entanto, essas variáveis possuem uma característica particular que está relacionada com a presença ou ausência da estrutura comercial da planta no momento da colheita, o que resulta em valores zeros para essas variáveis, gerando certa variabilidade e afetando a análise de variância dos dados. Assim, este trabalho teve por objetivo definir uma transformação adequada para as variáveis observadas em experimentos com a cultura do brócolis, realizados em diferentes épocas do ano, visando estabilizar a variabilidade gerada pela presença de valores zeros nas múltiplas colheitas. Analisaram-se quatro experimentos com a cultura do brócolis utilizando a família de transformações de Box-Cox, com uma adaptação de Yamamura apresentada em 1999, devido a presença de valores nulos. Concluiu-se que para experimentos com brócolis a transformação mais adequada para as variáveis cabeças comerciais e não comerciais, e ramificações comerciais e não comerciais foi logarítmica para cultivos no inverno, na primavera e no verão. Para cultivos no outono, a transformação indicada para as variáveis cabeças comerciais e não comerciais e ramificações não comerciais foi “inverso da raiz quarta”, enquanto para ramificações comerciais, “inverso da raiz quadrada”.

**Palavras-chave:** Transformação de dados; Box-Cox; variabilidade; análise de variância.

### ABSTRACT

#### EXCESS OF ZERO VALUE IN VARIABLES: STUDY OF CASE IN BROCCOLI FIELD EXPERIMENTS

The major purpose of vegetable crops production systems is the profitable and economically viable yield of marketable products. Numerous field experiments with vegetable crops are carried out to determine biomass of weekly and heads production of marketable and non-marketable variables. However these variables possess the particular characteristic related to their presence or absence of the commercial structure of the plant at harvest time, what implicates in zero values for those variables, generate variability and affect variance analyses. This paper deals with the search of applicable transformations for variables studied in broccoli field experiments conducted in different planting times attempting to stabilize the variability resulting from the presence of zero values in multiple harvests. Due to the presence of zero values, broccoli experiments were analyzed using an adaptation of the Box-Cox transformation family found in Yamamura in 1999. The most adequate transformation found for marketable and non-marketable heads and side shoots of broccoli grown in winter, spring and summer was the logarithmical. For marketable and non-marketable heads and non-marketable side shoots for fall-grown broccoli the inverse fourth root and to marketable side shoots is a inverse square root was the indicated transformation.

**Key words:** Transformation of data; Box-Cox; variability; analysis of variance.

---

<sup>(1)</sup> Recebido para publicação em 9 de outubro de 2008 e aceito em 13 de abril de 2010.

<sup>(2)</sup> Departamento de Fitotecnia, Centro de Ciências Rurais, Universidade Federal de Santa Maria (UFSM), 97105-900 Santa Maria (RS). E-mail: adlucio@smail.ufsm.br (\*) Autor correspondente; trevisan@smail.ufsm.br; klipmart@smail.ufsm.br; sjlopes@smail.ufsm.br

<sup>(3)</sup> Departamento de Estatística, Centro de Ciências Naturais e Exatas, Universidade Federal de Santa Maria (UFSM), 97105-900 Santa Maria (RS). E-mail: mararubiacouto@hotmail.com.

## 1. INTRODUÇÃO

A variabilidade da produção entre plantas pode acarretar em heterogeneidade das variâncias residuais que, para realização da análise de variância, deverá ser rejeitada para atender à pressuposição do modelo matemático da homogeneidade dos erros. Assim é assegurado o nível de significância dos testes F e os de comparação múltipla de médias. Em análises de dados experimentais com variâncias heterogêneas, o nível de significância passa a ficar acima do especificado (CONAGIN et al., 1993). O estudo do comportamento da média e da variância da produção ao longo de diferentes colheitas se faz necessário, bem como a interferência direta que diferentes tratamentos culturais realizados nos experimentos exercem nessas estatísticas, pois são fatores que podem alterar as conclusões de trabalhos científicos.

No cultivo de oleráceas, a produção das partes comercializáveis da planta é de grande interesse. No entanto, a ocorrência de valores zeros é bastante frequente, pois nem sempre existem estruturas comerciais a serem colhidas ou de tamanho e formato adequado para sua colheita e comercialização (SOUZA et al., 2002; LORENTZ et al., 2005; LÚCIO et al., 2008). Esta frequência de zeros gera variabilidade, o que pode provocar distorções nos procedimentos de inferência derivados da análise de variância. Um exemplo de heterocedasticidade das variâncias refere-se a dados provindos de contagens, pois algumas amostras resultam em uma variabilidade maior que a variabilidade nominal esperada pelo modelo Poisson a ser ajustado. Segundo BORGATTO et al. (2006), esse fenômeno é conhecido como superdispersão, podendo ocorrer devido: à variabilidade da média, ao excesso de zero, à correlação entre indivíduos e/ou omissão de variáveis não observadas.

Nos casos da presença de superdispersão na amostra, o uso de transformações de variáveis se faz necessário para que se possa atingir a homocedasticidade na análise de variância. Em experimentos cuja resposta é observada com alta taxa de valor zero, como é o caso de experimentos com oleráceas com múltiplas colheitas, a eficácia das transformações fica prejudicada, especialmente aquelas que empregam a função logarítmica, pois o zero não pertence ao domínio dessa função. Estudos que descrevem a variabilidade de populações utilizam frequentemente a transformação logarítmica para alcançar a homocedasticidade ou a estabilidade da variância para testar os efeitos na análise de variância.

Ao se dispor de uma amostra com alta taxa de valores zero é comum o acréscimo de uma constante ( $c$ ) a cada um dos dados (BANZATTO e KRONKA, 2006; STORCK et al., 2000). A determinação do valor constante  $c$  que deve ser adicionado a variável original ( $Y$ ) tem sido bastante discutida na literatura. Conforme

YAMAMURA (1999), não há base concreta quanto às razões pelas quais a unidade é adicionada ao invés de outra constante, tais como 0,5 ou dois. Outros princípios alternativos foram propostos por HILL (1963); GRIFFITHS (1980) e BERRY (1987) para determinar o valor de  $c$ . Os procedimentos propostos por esses autores são aplicáveis em um determinado conjunto de pressupostos, os quais requerem cálculos complexos. Devido a este fato, um princípio prático foi proposto por YAMAMURA (1999) sugerindo como uma escolha razoável  $c = 0,5$  se a forma da distribuição não for conhecida. O autor defende o uso de  $c$  igual 0,5 pelo motivo de que, se uma distribuição discreta definida em  $(0, 1, 2, \dots)$  pode ser aproximadamente descrita por uma distribuição contínua definida em  $(0, \infty)$ , quando se usa  $c = 0,5$ .

Este trabalho teve por objetivo definir uma transformação adequada para as variáveis observadas em experimentos com a cultura de brócolis, realizados em diferentes épocas do ano, visando estabilizar a variabilidade gerada pela presença de valores zeros nas múltiplas colheitas.

## 2. MATERIAL E MÉTODOS

Foram realizados quatro experimentos com a cultura do brócolis nas estações inverno, primavera, verão e outono, em Santa Maria (latitude: 29°43'21.17"S, longitude: 53°43'26.64"O e altitude: 95 m), em solo classificado como Planossolo Gley Pouco Húmico (DALMOLIN e KLAMT, 1997).

Cada parcela foi constituída por 14 plantas, arranjadas em linhas simples, em área total de 9,80 m<sup>2</sup>, com espaçamento de 1,00 m entre filas e 0,70 m entre plantas e os dados foram obtidos de dez plantas úteis por parcela. Utilizou-se o delineamento experimental de blocos ao acaso bifatorial com parcela subdividida no tempo e amostragem na subparcela em quatro repetições. Nas parcelas principais foram casualizadas as cultivares, sendo 15 no verão, 11 no outono e 14 no inverno e primavera e, nas subparcelas, foram casualizadas as colheitas, sendo seis no inverno, oito no outono e nove na primavera e no verão.

As aplicações de adubos, fungicidas e inseticidas foram realizadas sempre que necessário, conforme recomendação da cultura (FILGUEIRA, 2000) e as colheitas, em dois a três repasses semanais, coletando os valores das variáveis cabeças comerciais e não comerciais, e ramificações comerciais e não comerciais. Os critérios para determinação da maturação comercial foram flores fechadas, botões florais compactos e de coloração verde intensa e pedúnculos florais sem alongamento excessivo (CORREIA, 1983; FILGUEIRA, 2000). Aquelas cabeças e ramificações com flores abertas, alongamentos

no pedúnculo, danos de pragas, doenças e geadas e ramificações com comprimento inferior a 15 cm foram consideradas não comerciais.

Utilizou-se a transformação Box-Cox (Box e Cox, 1964) para o conjunto de funções potências, para encontrar a transformação para estabilizar ou reduzir a variabilidade existente e normalizar os resíduos. Como para valores nulos a família de transformações de Box-Cox fica restrita, utilizou-se a variável somada a uma constante. Adotou-se o procedimento proposto por YAMAMURA (1999) com  $c=0,5$ , gerando assim as expressões

$$f(y) = \frac{(y + 0,5)^\lambda - 1}{\lambda}, \lambda \neq 0 \text{ e } f(y) = \ln(y + 0,5), \lambda = 0$$

Para a escolha da melhor potência  $\lambda$ , considerou-se para  $\lambda$  valores no intervalo  $[-2, 2]$ , conforme descrevem DRAPPER e SMITH (1981), sendo substituídos nas

$$\text{expressões } Z(\lambda) = \frac{y^\lambda - 1}{\lambda y^{\lambda-1}}, \lambda \neq 0 \text{ ou } Z(\lambda) = y \ln y,$$

$\lambda=0$ , em que o jacobiano  $J(\lambda, z)=1$ , gerando valores transformados ( $Z(\lambda)$ ) para cada variável sendo, após, realizada a análise de variância para cada variável transformada. Para determinar o estimador de máxima verossimilhança ( $L_{\max}$ ), substituiu-se a soma de quadrados residual ( $SSE_{(y)}$ ), usado, para cada valor de  $\lambda$  na expressão  $L_{\max}(\lambda) = -\frac{n}{2} \ln\left(\frac{\hat{\sigma}_\lambda^2(z)}{n}\right) = -\frac{n}{2} \ln\left(\frac{SSE(\lambda)}{n}\right)$ , e o melhor  $\lambda$  foi escolhido de maneira que maximizasse  $L_{\max}$  e minimizasse o resíduo.

Foram obtidas as medidas descritivas e a análise de variância para as variáveis, antes e após o uso de transformação, aplicando o teste de Bartlett, conforme STEEL et al. (1997), para verificar se a transformação foi suficiente para que a homocedasticidade dos erros tenha sido satisfeita e, finalmente, realizada a análise dos resíduos antes e após a transformação. Para as análises estatísticas dos dados, utilizaram-se os *softwares* Statistica 7.0 e SAS, em nível de 5% de significância.

### 3. RESULTADOS E DISCUSSÃO

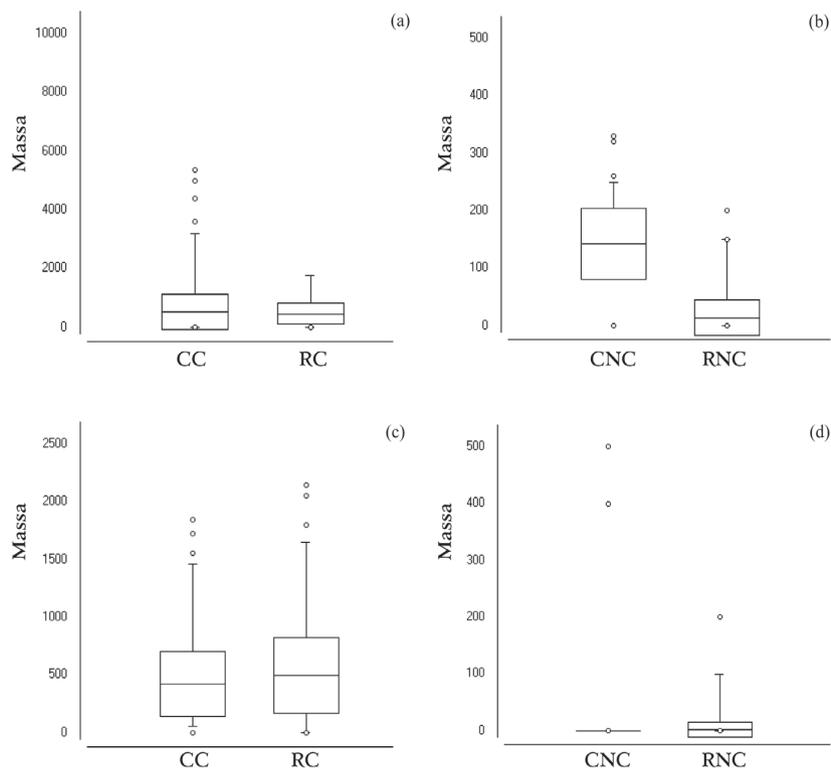
Nas figuras 1 e 2, pode-se verificar o efeito na média e no desvio-padrão de uma taxa elevada de valores zero para rendimento de cabeças comerciais (CC) e não comerciais (CNC) e para ramificações comerciais (RC) e não comerciais (RNC) nas quatro épocas. A frequência de zeros para CNC e RNC foi muito alta, acima de 90%, exceto para RNC no experimento no verão (60%). Abordando esse resultado com enfoque agrícola é o que se espera de uma cultivar de brócolis, com pouca colheita de cabeças e ramificações não comerciais, mas do ponto de vista da estatística experimental, esse comportamento

gera alta taxa de valores zero nas amostras, para essas variáveis, o que inflaciona a variabilidade já existente entre as repetições do mesmo tratamento, alterando o erro experimental e sua estimativa (QMe).

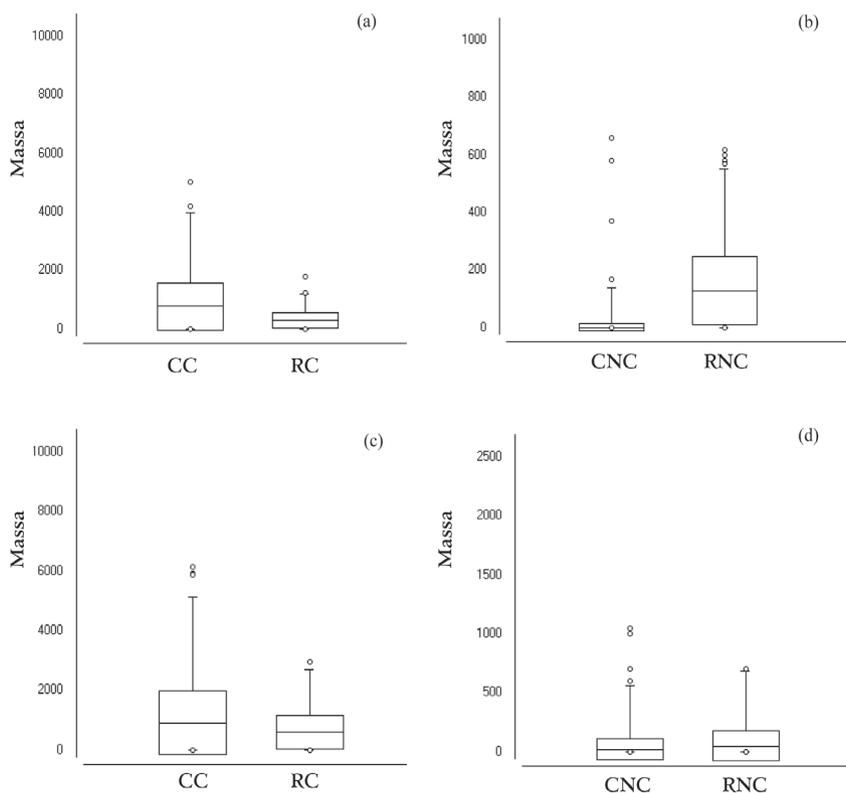
Nas tabelas 1, 2, 3 e 4 são apresentados os vários valores de  $\lambda$ , a raiz do quadrado médio do erro (RQME) e a função de máxima verossimilhança ( $L_{\max}$ ) para as quatro épocas de cultivo, respectivamente para as variáveis CC, RC, CNC e RNC. Analisando-se essas tabelas, o valor da potência  $\lambda$  que minimiza a RQME e maximiza a  $L_{\max}$  foi 0 (zero), para os experimentos no inverno, primavera e verão, independentemente da variável observada, indicando que a melhor transformação seria a função logarítmica ( $y^0 = \log(y)$ ). No outono, o valor de  $\lambda$  para RQME mínimo e  $L_{\max}$  máximo foi de -0,25 para todas as variáveis, exceto para ramificações comerciais em que  $\lambda = -0,50$ . Desta forma, para o outono, a transformação indicada para CC, CNC e RNC foi inverso da raiz quarta ( $y^{0,25}$ ) e para RC, a transformação inverso da raiz quadrada ( $y^{0,5}$ ). No outono, foram observados os maiores valores da média para CC e RC e uma das menores para CNC e RNC e os maiores valores QMe em todas as variáveis, quando comparados com os experimentos no inverno, na primavera e no verão. Comprova-se ser o outono a época preferencial de cultivo, mas com maior variabilidade, quando comparada com as demais épocas, concordando com os resultados observados por TREVISAN et al. (2003). Esse comportamento no outono pode ser justificado pelo fato da cultura de brócolis requerer, conforme SONNENBERG (1985) e FILGUEIRA (2000), clima ameno ou temperado, predominando no Brasil o cultivo de cultivares de outono-inverno, como o caso das cultivares avaliadas no presente trabalho.

Após a aplicação da transformação, a redução no valor do coeficiente de variação foi evidente, variando de 35% para RC no inverno até 99% para as CNC e RNC no verão (Tabela 5), indicando o efeito na redução da variabilidade existente na amostra com excesso de zeros. Com a aplicação da transformação, houve maior evidência em rejeitar a hipótese  $H_0$  para as fontes de variação avaliadas na análise de variância, principalmente referente às cultivares avaliadas, visto que em todas as variáveis, em todas as épocas, o efeito da interação entre cultivares e colheita não foi significativo, nem com a variável original mais a constante, nem com a transformação da variável mais a constante. A melhoria na precisão da inferência é uma das consequências esperadas após a aplicação de qualquer transformação de dados, fato comprovado no trabalho.

Mesmo com a redução nos valores do CV% e da maior significância no teste F na análise de variância, após a aplicação do teste de Bartlett para as variáveis transformadas, as transformações realizadas nos dados não foram suficientes para o atendimento à pressuposição



**Figura 1.** Box-plot para massa (g) das cabeças comerciais (CC), ramificações comerciais (RC), cabeças não comerciais (CNC) e ramificações não comerciais (RNC) em experimentos com brócolis no inverno (a; b) e primavera (c; d). Santa Maria (RS), 2008.



**Figura 2.** Box-plot para massa (g) das cabeças comerciais (CC), ramificações comerciais (RC), cabeças não comerciais (CNC) e ramificações não comerciais (RNC) em experimentos com brócolis no verão (a; b) e outono (c; d). Santa Maria (RS), 2008.

**Tabela 1.** Potência  $\lambda$ , raiz do quadrado médio do erro (RQME) e função de máxima verossimilhança ( $L_{\max}$ ) nas transformações Box-Cox para a variável cabeças comerciais em experimentos de brócolis no inverno, na primavera, no verão e outono. Santa Maria (RS), 2008

$\lambda$	Inverno		Primavera		Verão		Outono	
	RQME	$L_{\max}$	RQME	$L_{\max}$	RQME	$L_{\max}$	RQME	$L_{\max}$
-2,00	-4768,30	13,7746	-3479,54	17,9499	-5826,61	18,1525	-2128,42	6,1859
-1,75	-4231,35	10,2520	-3102,04	13,1222	-5196,31	13,2663	-1856,06	4,8993
-1,50	-3727,72	7,7714	-2747,50	9,7775	-4602,25	9,8717	-1606,95	3,9583
-1,25	-3265,51	6,0268	-2421,43	7,4595	-4054,06	7,5153	-1387,59	3,2805
-1,00	-2854,97	4,8085	-2130,78	5,8608	-3564,52	5,8908	-1206,19	2,8086
-0,75	-2508,79	3,9748	-1883,98	4,7754	-3150,16	4,7934	-1072,61	2,5051
-0,50	-2241,84	3,4320	-1690,76	4,0679	-2831,32	4,0903	-997,75	2,3496
-0,25	-2069,69	3,1219	-1561,07	3,6528	-2630,42	3,7012	-992,11	2,3383
0,00	-2005,71	3,0140	-1503,20	3,4815	-2567,42	3,5870	-1063,45	2,4855
0,25	-2057,16	3,1005	-1521,44	3,5346	-2653,10	3,7432	-1214,49	2,8287
0,50	-2222,49	3,3956	-1614,65	3,8189	-2884,44	4,1998	-1441,94	3,4368
0,75	-2491,99	3,9382	-1776,77	4,3688	-3245,96	5,0274	-1737,79	4,4275
1,00	-2851,19	4,7985	-1998,93	5,2533	-3715,91	6,3516	-2091,72	5,9946
1,25	-3284,68	6,0907	-2271,66	6,5876	-4272,47	8,3780	-2493,44	8,4553
1,50	-3778,57	7,9919	-2586,28	8,5531	-4897,11	11,4315	-2933,89	12,3282
1,75	-4321,42	10,7727	-2935,45	11,4279	-5575,37	16,0196	-3405,69	18,4641
2,00	-4904,18	14,8436	-3313,17	15,6350	-6296,33	22,9313	-3902,97	28,2638

**Tabela 2.** Potência  $\lambda$ , raiz do quadrado médio do erro (RQME) e função de máxima verossimilhança ( $L_{\max}$ ) nas transformações Box-Cox para a variável ramificações comerciais em experimentos de brócolis no inverno, na primavera, no verão e outono. Santa Maria (RS), 2008

$\lambda$	Inverno		Primavera		Verão		Outono	
	RQME	$L_{\max}$	RQME	$L_{\max}$	RQME	$L_{\max}$	RQME	$L_{\max}$
-2,00	-3191,24	14,0994	-3191,24	14,0994	-5399,94	14,7596	-1934,30	5,2387
-1,75	-2835,23	10,4954	-2835,23	10,4954	-4818,49	11,0457	-1678,87	4,2097
-1,50	-2507,71	7,9994	-2507,71	7,9994	-4274,88	8,4237	-1451,75	3,4658
-1,25	-2213,98	6,2702	-2213,98	6,2702	-3778,08	6,5758	-1258,97	2,9385
-1,00	-1960,06	5,0797	-1960,06	5,0797	-3339,67	5,2848	-1107,43	2,5809
-0,75	-1752,58	4,2768	-1752,58	4,2768	-2974,37	4,4050	-1004,76	2,3637
-0,50	-1598,37	3,7635	-1598,37	3,7635	-2700,13	3,8421	-958,98	2,2729
-0,25	-1503,73	3,4794	-1503,73	3,4794	-2536,93	3,5419	-977,47	2,3091
0,00	-1473,28	3,3927	-1473,28	3,3927	-2503,27	3,4830	-1065,57	2,4900
0,25	-1508,85	3,4943	-1508,85	3,4943	-2610,60	3,6744	-1224,93	2,8541
0,50	-1608,90	3,7965	-1608,90	3,7965	-2858,53	4,1578	-1452,83	3,4690
0,75	-1768,88	4,3351	-1768,88	4,3351	-3234,76	5,0156	-1742,87	4,4468
1,00	-1982,36	5,1745	-1982,36	5,1745	-3719,74	6,3873	-2086,66	5,9687
1,25	-2242,19	6,4185	-2242,19	6,4185	-4292,56	8,4983	-2475,52	8,3266
1,50	-2541,45	8,2263	-2541,45	8,2263	-4934,76	11,7048	-2901,57	11,9917
1,75	-2873,93	10,8376	-2873,93	10,8376	-5631,58	16,5662	-3358,12	17,7273
2,00	-3234,25	14,6113	-3234,25	14,6113	-6371,80	23,9596	-3839,75	26,7747

**Tabela 3.** Potência  $\lambda$ , raiz do quadrado médio do erro (RQME) e função de máxima verossimilhança ( $L_{\max}$ ) nas transformações Box-Cox para a variável cabeças não comerciais em experimentos de brócolis no inverno, na primavera, no verão e outono. Santa Maria (RS), 2008

$\lambda$	Inverno		Primavera		Verão		Outono	
	RQME	$L_{\max}$	RQME	$L_{\max}$	RQME	$L_{\max}$	RQME	$L_{\max}$
-2,00	-4695,66	13,2351	-3466,03	17,7074	-5451,59	15,1446	-1994,18	5,5143
-1,75	-4185,03	9,9941	-3100,89	13,0818	-4878,26	11,3797	-1741,60	4,4420
-1,50	-3703,00	7,6665	-2755,34	9,8227	-4337,38	8,6903	-1509,16	3,6404
-1,25	-3256,60	5,9973	-2434,28	7,5268	-3837,26	6,7726	-1302,87	3,0510
-1,00	-2855,08	4,8088	-2144,16	5,9175	-3389,04	5,4165	-1130,68	2,6328
-0,75	-2510,57	3,9787	-1893,34	4,8063	-3007,61	4,4786	-1002,77	2,3597
-0,50	-2238,42	3,4255	-1692,22	4,0681	-2712,34	3,8656	-931,27	2,2196
-0,25	-2056,55	3,0994	-1552,66	3,6235	-2526,59	3,5237	-928,82	2,2149
0,00	-1982,86	2,9763	-1486,04	3,4288	-2474,31	3,4331	-1005,78	2,3658
0,25	-2030,18	3,0548	-1500,06	3,4689	-2572,95	3,6061	-1166,90	2,7157
0,50	-2201,32	3,3563	-1595,93	3,7559	-2825,85	4,0906	-1409,61	3,3430
0,75	-2487,79	3,9291	-1767,95	4,3317	-3220,34	4,9796	-1725,36	4,3806
1,00	-2873,53	4,8579	-2006,06	5,2772	-3733,69	6,4318	-2102,72	6,0513
1,25	-3340,40	6,2802	-2299,01	6,7282	-4341,28	8,7072	-2530,25	8,7261
1,50	-3871,99	8,4133	-2636,53	8,9011	-5021,83	12,2240	-2998,07	13,0246
1,75	-4455,06	11,5945	-3010,17	12,1338	-5758,84	17,6513	-3498,19	19,9858
2,00	-5079,45	16,3459	-3413,21	16,9487	-6540,17	26,0574	-4024,39	31,3602

**Tabela 4.** Potência  $\lambda$ , raiz do quadrado médio do erro (RQME) e função de máxima verossimilhança ( $L_{\max}$ ) nas transformações Box-Cox para a variável ramificações não comerciais em experimentos de brócolis no inverno, na primavera, no verão e outono. Santa Maria (RS), 2008

$\lambda$	Inverno		Primavera		Verão		Outono	
	RQME	$L_{\max}$	RQME	$L_{\max}$	RQME	$L_{\max}$	RQME	$L_{\max}$
-2,00	-4701,68	13,2601	-3462,10	17,6499	-5385,79	14,6559	-1992,14	5,5047
-1,75	-4191,23	10,0156	-3097,35	13,0435	-4814,46	11,0235	-1738,91	4,4318
-1,50	-3709,42	7,6849	-2752,24	9,7975	-4278,14	8,4374	-1506,11	3,6309
-1,25	-3263,25	6,0133	-2431,67	7,5106	-3785,31	6,5995	-1299,78	3,0429
-1,00	-2861,96	4,8229	-2142,07	5,9072	-3347,13	5,3045	-1127,93	2,6266
-0,75	-2517,66	3,9912	-1891,77	4,8001	-2978,17	4,4133	-1000,73	2,3556
-0,50	-2245,61	3,4368	-1691,14	4,0644	-2696,90	3,8360	-930,21	2,2176
-0,25	-2063,68	3,1097	-1552,00	3,6215	-2524,92	3,5208	-928,80	2,2149
0,00	-1989,68	2,9857	-1485,69	3,4278	-2483,62	3,4490	-1006,59	2,3674
0,25	-2036,48	3,0635	-1499,90	3,4684	-2587,87	3,6330	-1168,08	2,7185
0,50	-2207,01	3,3646	-1595,87	3,7557	-2839,66	4,1189	-1410,63	3,3459
0,75	-2493,00	3,9374	-1767,96	4,3318	-3226,85	4,9958	-1725,76	4,3821
1,00	-2878,51	4,8670	-2006,12	5,2775	-3728,43	6,4150	-2102,15	6,0484
1,25	-3345,44	6,2913	-2299,12	6,7288	-4321,53	8,6219	-2528,51	8,7131
1,50	-3877,35	8,4282	-2636,71	8,9024	-4986,14	12,0085	-2995,03	12,9908
1,75	-4460,91	11,6162	-3010,42	12,1363	-5706,51	17,1967	-3493,79	19,9107
2,00	-5085,89	16,3787	-3413,53	16,9532	-6470,84	25,1722	-4018,59	31,2048

**Tabela 5.** Média ( $\bar{X}$ ) em kg por parcela, quadrado médio do erro (QMe), coeficiente de variação (CV%) para a variável mais a constante e a transformação da variável mais constante e o percentual de redução nos valores do CV% após a transformação, em cabeças (CC) e ramificações comerciais (RC) e em cabeças (CNC) e ramificações não comerciais (RNC) em experimentos com brócolis no inverno, na primavera, no verão e outono. Santa Maria (RS), 2008

Variáveis	$\bar{X}$	QMe	CV (%)	Redução do CV	$\bar{X}$	QMe	CV (%)	Redução do CV
Inverno								
CC + 0,5	0,41	0,34	142,22		0,26	0,06	94,21	
Transf (CC+0,5)	0,27	0,02	52,38	63%	0,36	0,01	27,78	70%
RC + 0,5	0,39	0,08	72,52		0,20	0,07	132,29	
Transf (RC+0,5)	0,21	0,01	47,62	35%	0,42	0,01	23,81	82%
CNC + 0,5	0,45	0,84	203,67		0,20	0,007	41,83	
Transf (CNC+0,5)	0,11	0,01	90,91	55%	0,18	0,002	24,84	40%
RNC + 0,5	0,52	0,005	13,59		0,56	0,008	15,97	
Transf (RNC+0,5)	1,08	0,001	2,93	79%	1,06	0,002	4,21	74%
Verão								
CC + 0,5	0,44	0,41	145,52		0,47	0,82	192,66	
Transf (CC+0,5)	0,27	0,03	64,15	56%	1,08	0,01	9,26	95%
RC + 0,5	0,45	0,04	44,44		0,55	0,28	96,21	
Transf (RC+0,5)	0,12	0,01	21,45	52%	1,07	0,05	20,89	78%
CNC + 0,5	0,005	0,002	872,13		0,07	0,002	63,89	
Transf (CNC+0,5)	0,30	0,0006	8,16	99%	0,25	0,0003	6,93	89%
RNC + 0,5	0,01	0,006	774,59		0,03	0,02	471,40	
Transf (RNC+0,5)	1,18	0,003	4,65	99%	1,40	0,01	7,14	98%
Outono								

da homogeneidade dos resíduos, pois se rejeitou a hipótese de homocedasticidade para todas as variáveis em todas as épocas de cultivo, apesar do aumento observado no valor do  $\chi^2$  calculado para todas as situações estudadas.

As variáveis estudadas houve melhoras significativas ao atendimento à normalidade dos erros, verificado nas figuras 3 e 4, para CC e RC com os dados transformados, independentemente da época de experimentação. Mesmo se rejeitando a hipótese de homocedasticidade dos resíduos, o valor-p apresentado no teste, após a transformação, foi muito próximo do nível de significância nominal no estudo ( $p < 0,05$ ), além da transformação indicada em cada situação para essas duas variáveis, indicando a melhor condição possível para transformação dos dados e homogeneidade dos erros. Portanto, é uma possibilidade que permite estabilizar a variância ou, pelo menos, reduzir esta variabilidade, auxiliando nas análises estatísticas e interpretações dos resultados deste tipo de experimento, sem comprometer a precisão experimental.

Já para as cabeças e ramificações não comerciais, os valores residuais obtidos continuaram ainda muito dispersos após a transformação dos dados (Figuras 5 e 6), indicando que esta não controlou a

variação de forma efetiva. Uma explicação para este resultado é que para essas variáveis ocorreu uma frequência de valores zero muito elevada (Figuras 1 e 2) e, provavelmente, esse seja o motivo que mesmo com o emprego da transformação os resíduos ainda não estejam próximos à normalidade.

#### 4. CONCLUSÕES

1. A transformação mais adequada para cabeças comerciais e não comerciais e ramificações comerciais e não comerciais de brócolis é a logarítmica para experimentos no inverno, na primavera e no verão.

2. Para experimentos no outono, a transformação indicada para cabeças comerciais e não comerciais e ramificações não comerciais é inverso da raiz quarta, enquanto para ramificações comerciais é inverso da raiz quadrada.

#### AGRADECIMENTOS

Ao CNPq pelo auxílio financeiro para a realização do trabalho e pela concessão de bolsas de produtividade em pesquisa aos autores Alessandro Dal'Col Lúcio e Sidinei José Lopes.

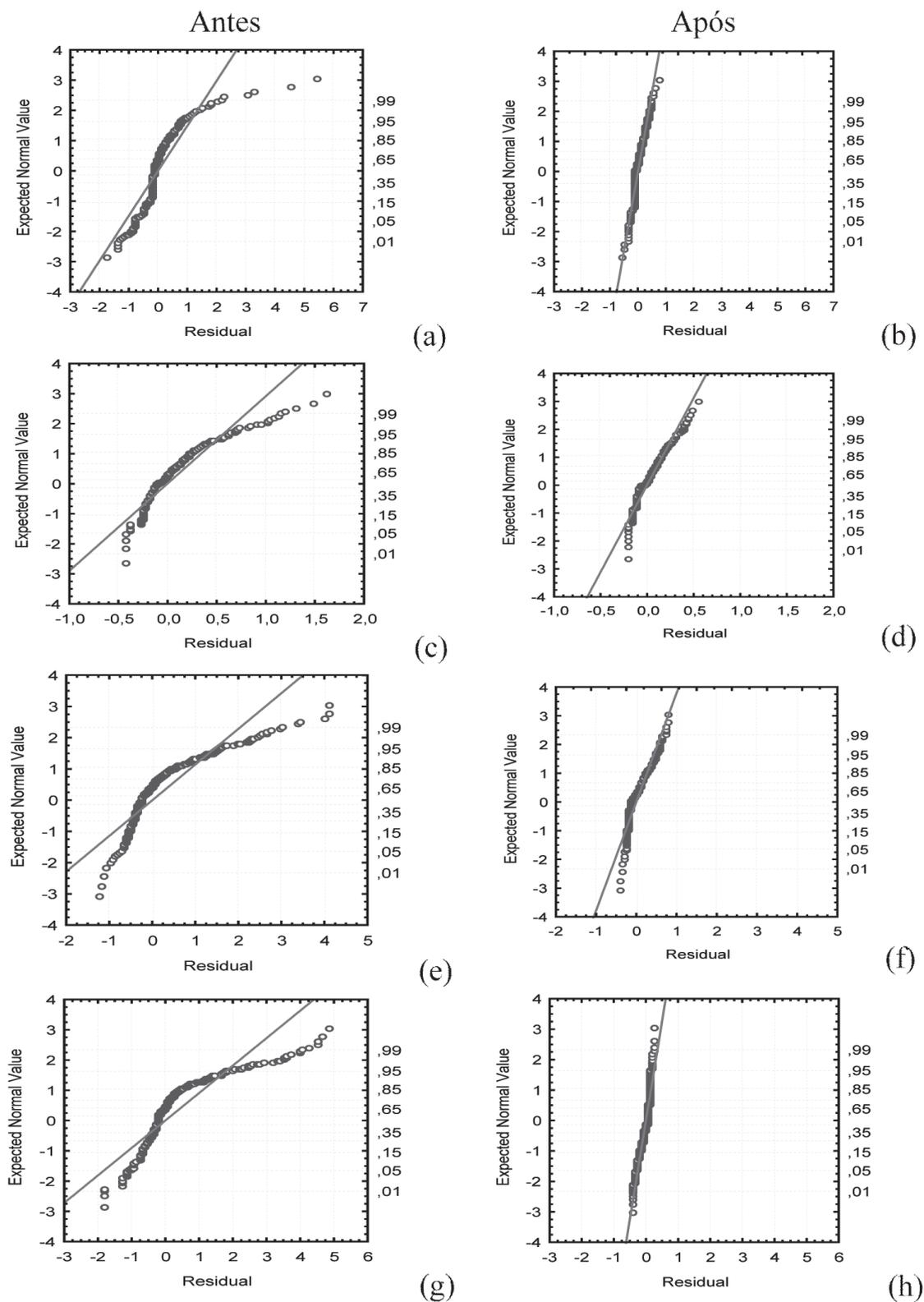


Figura 3. Resíduos (Residual) em função do valor normal (Expected normal value) de cabeças comerciais, antes e após a transformação, em experimentos com brócolis no inverno (a; b), primavera (c; d), verão (e; f) e outono (g; h). Santa Maria (RS), 2008.

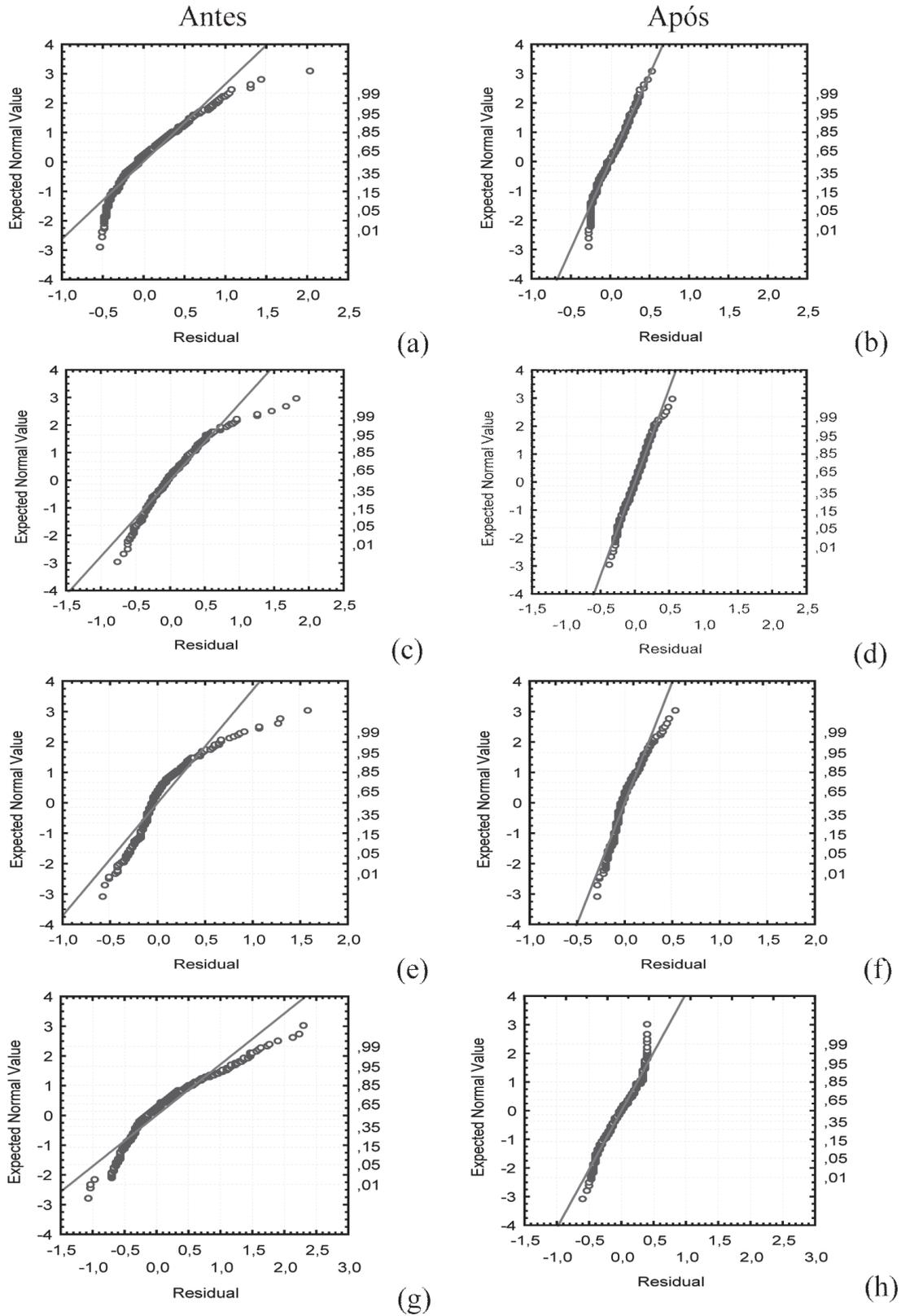


Figura 4. Resíduos (Residual) em função do valor normal (Expected normal value) de ramificações comerciais, antes e após a transformação, em experimentos com brócolis no inverno (a; b), primavera (c; d), verão (e; f) e outono g; h). Santa Maria (RS), 2008.

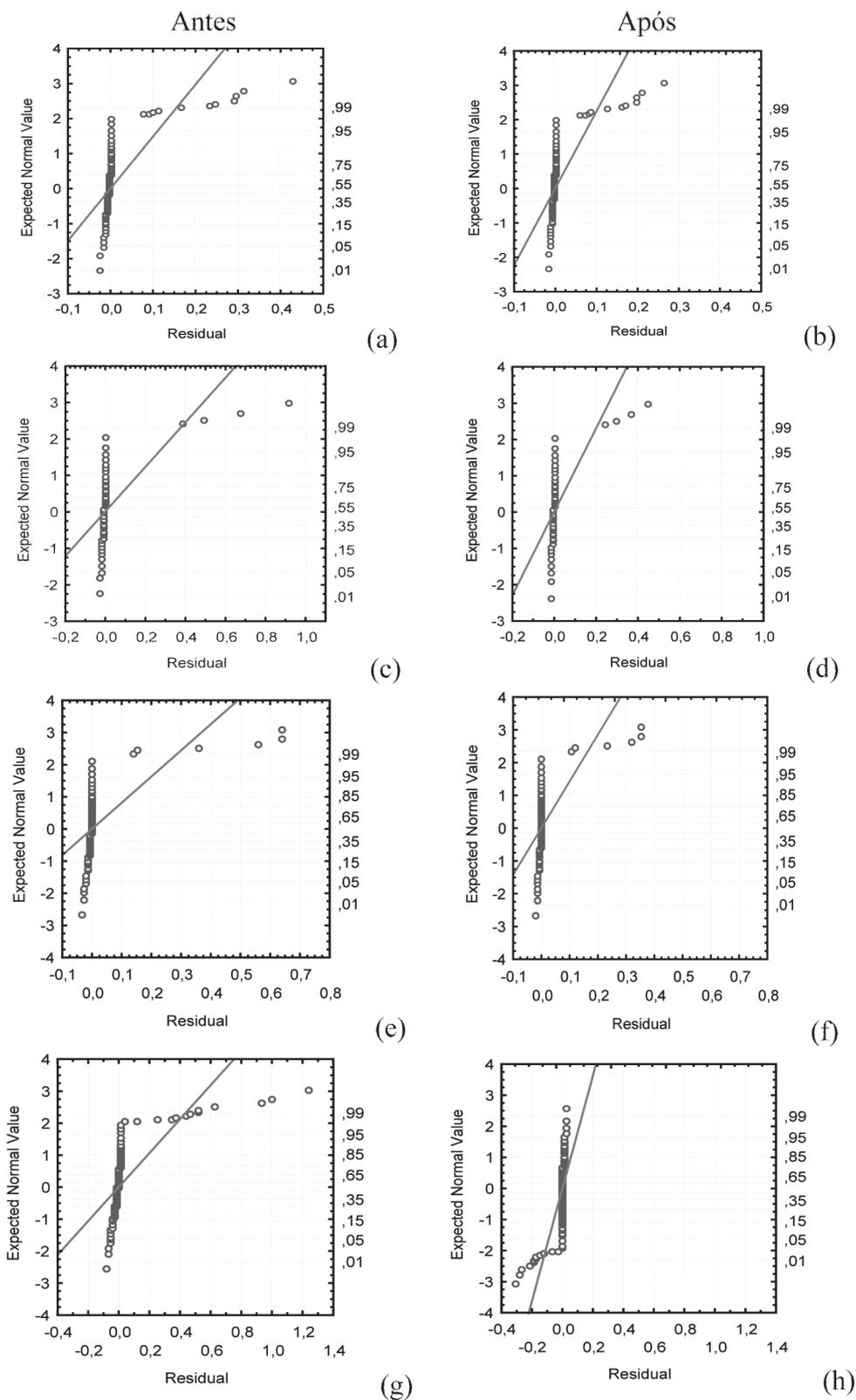


Figura 5. Resíduos (Residual) em função do valor normal (Expected normal value) de cabeças não comerciais, antes e após a transformação, em experimentos com brócolis no inverno (a; b), primavera (c; d), verão (e; f) e outono (g; h). Santa Maria (RS), 2008.

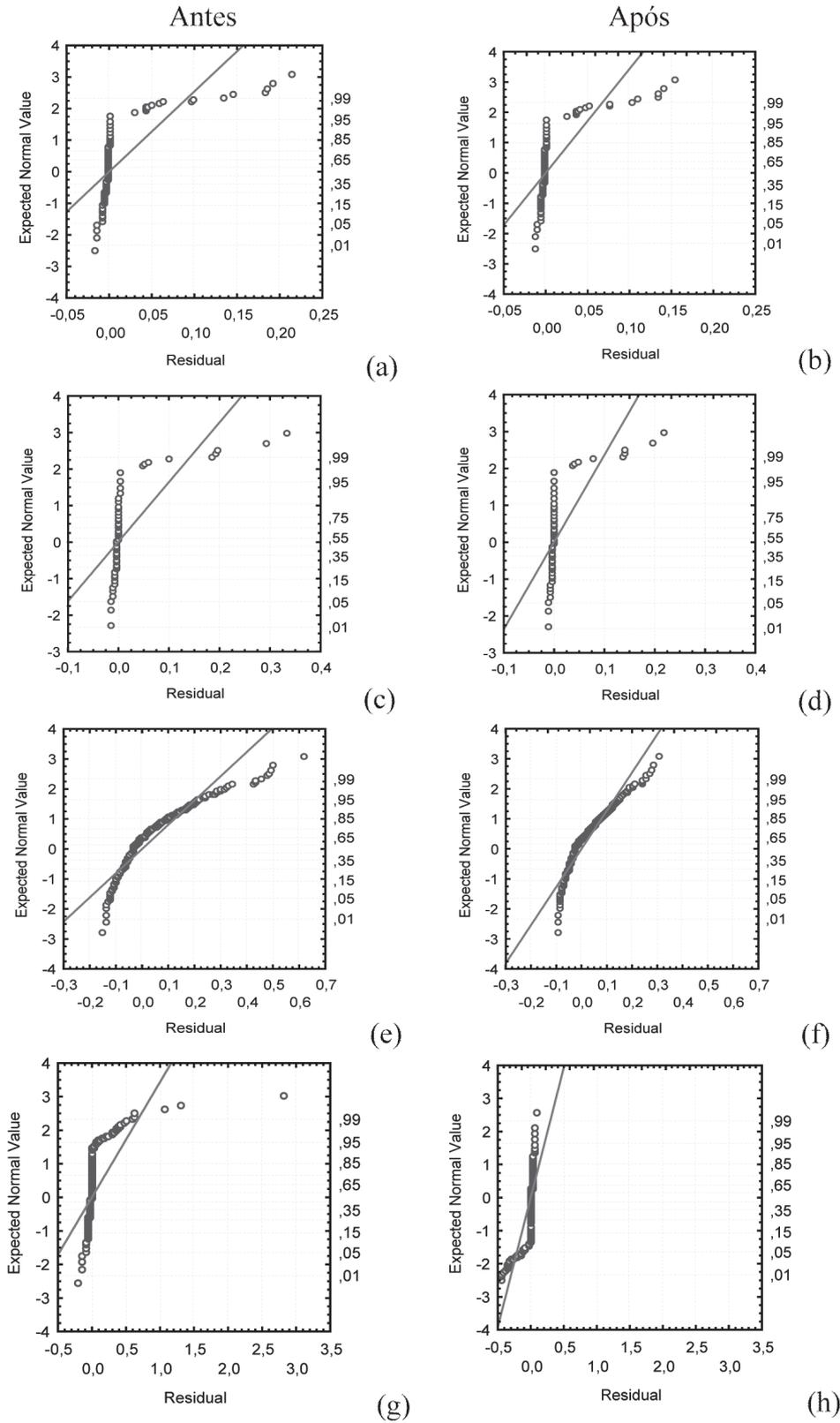


Figura 6. Resíduos (Residual) em função do valor normal (Expected normal value) de ramificações não comerciais, antes e após a transformação, em experimentos com brócolis no inverno (a; b), primavera (c; d), verão (e; f) e outono (g; h). Santa Maria (RS), 2008.

## REFERÊNCIAS

- BANZATTO, D.A.; KRONKA, S.N. **Experimentação agrícola**. 4.ed. Jaboticabal: FUNEP, 2006. 246p.
- BERRY, D.A. Logarithmic transformations in ANOVA. **Biometrics**, v.43, p.439-456, 1987.
- BORGATTO, A.F.; DEMÉTRIO, C.G.B.; LEANDRO, R.A. Modelos para proporções com superdispersão e excesso de zeros - Um procedimento bayesiano. **Revista de Matemática e Estatística**, v.24, p.125-135, 2006.
- BOX, G.E.P.; COX, D.R. An analysis of transformations. **Journal of the Royal Statistical Society**, v.26, p.211-252, 1964. (Series B - Statistical Methodological)
- CONAGIN, A.; NAGAI, V.; IGUE, T. Efeito da falta de normalidade em testes de homogeneidade das variâncias. **Bragantia**, v.57, p.203-214, 1993.
- CORREIA, L.G. Colheita, classificação, embalagem e comercialização de brássicas. **Informe Agropecuário**, v.9, p.52-54, 1983.
- DALMOLIN, R.S.D.; KLAMT, E. Solos de textura superficial média e arenosa do RS: características, classificação e aptidão de uso agrícola. In: FRIES, M.R.; DALMOLIN, R.S.D. (Ed.). **Atualização em recomendação de adubação e calagem: ênfase em plantio direto**. Santa Maria: UFSM/SBCS, 1997. p.9-22.
- DRAPER, N. R.; SMITH, H. **Applied regression analysis**. New York: John Wiley, 1981. 709p.
- FILGUEIRA, F.A.R. **Novo manual de olericultura: tecnologia moderna na produção e comercialização de hortaliças**. Viçosa: UFV, 2000. 402p.
- GRIFFITHS, D.A. Interval estimation for the three-parameter lognormal distribution via the likelihood function. **Journal of the Royal Statistical Society**, v.29, p.58-68, 1980. (Series C - Applied Statistics)
- HILL, B.M. The three-parameter lognormal distribution and bayesian analysis of a point-source epidemic. **Journal of the American Statistical Association**, v.58, p.72-84, 1963.
- LORENTZ, L.H.; LÚCIO, AD.; BOLIGON, A.A.; LOPES, S.J.; STORCK, L. Variabilidade da produção de frutos de pimentão em estufa plástica. **Ciência Rural**, v.35, p.316-323, 2005.
- LÚCIO, A.D.; CARPES, R.H.; STORCK, L.; LOPES, S.J.; LORENTZ, L.H.; PALUDO, A.L. Variância e média da massa de frutos de abobrinha-italiana em múltiplas colheitas. **Horticultura Brasileira**, v.26, p.333-339, 2008.
- SONNENBERG, P.E. **Olericultura especial - II**. 3.ed. Goiânia: Líder, 1985. 149p.
- SOUZA, M.F.; LÚCIO, A.D.; STORCK, L.; CARPES, R.H.; SANTOS, P.M.; SIQUEIRA, L.F.F. Tamanho da amostra para massa da massa de frutos, na cultura da abóbora italiana em estufa plástica. **Revista Brasileira de Agrociência**, v.8, p.123-128, 2002.
- STEEL, R.G.D.; TORRIE, J.H.; DICKY, D.A. **Principles and procedures of statistics: a biometrical approach**. New York: McGraw-Hill, 1997. 666p.
- STORCK, L.; LOPES, S.J., ESTEFANEL, V.; GARCIA, D.C. **Experimentação vegetal**. 1.ed. Santa Maria: UFSM, 2000. 198p.
- TREVISAN, J.N.; MARTINS, G.A.K.; LÚCIO, A.D.; CASTAMAN, C.; MARION, R.R.; TREVISAN, B.G. Rendimento de cultivares de brócolis semeadas em outubro na região centro do Rio Grande do Sul. **Ciência Rural**, v.33, p.233-239, 2003.
- YAMAMURA, K. Transformation using  $(x + 0.5)$  to stabilize the variance of populations. **Journal of Researches on Population Ecology**, v.42, p.229-234, 1999.