

**UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE CIÊNCIAS RURAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM AGRONOMIA**

**TRANSFORMAÇÃO DE DADOS COM EXCESSO DE
ZERO EM EXPERIMENTOS COM CULTURAS
OLERÍCOLAS**

DISSERTAÇÃO

Mara Rubia Machado Couto

Santa Maria, RS, Brasil.

2008

TRANSFORMAÇÃO DE DADOS COM EXCESSO DE ZERO EM EXPERIMENTOS COM CULTURAS OLERÍCOLAS

por

Mara Rubia Machado Couto

Dissertação apresentada ao Programa de Pós-Graduação em
Agronomia, da Universidade Federal de Santa Maria (UFSM, RS), como
requisito parcial para obtenção do grau de
Mestre em Agronomia.

Orientador: Prof. Dr. Alessandro Dal'Col Lúcio

Santa Maria, RS, Brasil

2008

**Universidade Federal de Santa Maria
Centro de Ciências Rurais
Departamento de Fitotecnia
Programa de Pós-Graduação em Agronomia**

A Comissão Examinadora, abaixo assinada,
aprova a Dissertação

**TRANSFORMAÇÃO DE DADOS COM EXCESSO DE ZERO EM
EXPERIMENTOS COM CULTURAS OLERÍCOLAS**

elaborada por
Mara Rubia Machado Couto

como requisito parcial para obtenção do grau de
Mestre em Agronomia

COMISSÃO EXAMINADORA:

Alessandro Dal'Col Lúcio, Dr. (UFSM).
(Presidente/Orientador)

Luciane Flores Jacobi, Dra. (UFSM)

Sidinei José Lopes , Dr. (UFSM)

Santa Maria, 30 de abril de 2008.

©2008

Todos os direitos autorais reservados a Mara Rubia Machado Couto. A reprodução de partes ou do todo deste trabalho só poderá ser com autorização por escrito do autor.

Endereço: Rua Ary Vieira, n. 16, Camobi, Santa Maria, RS, 97105-630
Fone: (0XX)55 32268045; End. Eletr.: mararubiacouto@hotmail.com.br

Dedico com muito carinho este trabalho aos meus familiares pela dedicação, compreensão e principalmente pela paciência.

AGRADECIMENTOS

Ao meu orientador, Prof. Dr. Alessandro Dal'Col Lúcio pelo acompanhamento durante este trabalho, pela oportunidade deste aprendizado, por toda a valiosa contribuição nesta pesquisa, e sua amizade.

A todos os professores do Departamento de Fitotecnia, especialmente aos do setor de Experimentação, pela sabedoria que souberam transmitir.

A minha querida amiga, colega e professora Luciane Flores Jacobi, por toda sua dedicação e carinho em minha trajetória acadêmica e profissional.

Aos meus colegas pelas horas de estudos, por todas as suas contribuições e amizade.

“O covarde nunca tenta,
O fracassado nunca termina e
O vencedor nunca desiste.”

Norman Vicent Peale

RESUMO

Dissertação
Programa de Pós-Graduação em Agronomia
Universidade Federal de Santa Maria

TRANSFORMAÇÃO DE DADOS COM EXCESSO DE ZERO EM EXPERIMENTOS COM CULTURAS OLERÍCOLAS

AUTORA: MARA RUBIA MACHADO COUTO
ORIENTADOR: ALESSANDRO DAL'COL LÚCIO
Data e local da Defesa: Santa Maria, 30 de abril de 2008.

Em olericultura, o interesse maior é a produção de frutos comercializáveis. Nesses experimentos comumente são avaliadas variáveis como fitomassa e número de frutos. No entanto, essas variáveis apresentam uma característica particular que está relacionada com sua ausência ou presença, no momento da colheita. Este atributo está representado pela ocorrência de valores zeros, para essas variáveis, gerando variabilidade e afetando a análise de variação dos dados. Neste estudo, analisaram-se experimentos com culturas da abobrinha italiana, pimentão e brócolis, utilizando-se a família de transformações de Box-Cox. Como para valores nulos, esta fica restrita, utilizou-se uma adaptação sugerida por Yamamura (1999). Concluiu-se que, para cada cultura e cada variável estudada existe a necessidade de transformações diferenciadas. Sendo assim, tem-se que para o experimento com abobrinha italiana, a transformação indicada é a raiz quarta. Para pimentão, a transformação indicada é o inverso da raiz quarta. E para brócolis, a transformação mais adequada é a logarítmica. Apesar da redução na variabilidade, observada em todos os experimentos, o uso do método não foi suficiente para tornar as variâncias homocedásticas.

Palavras-chaves: Abobrinha italiana, pimentão, brócolis, Box-Cox, análise de variância.

ABSTRACT

Dissertation

Post-Graduation Program in Agronomy
Santa Maria Federal University

TRANSFORMATION OF DATA WITH EXCESS OF ZERO VALUES INTO OLERICULTURE EXPERIMENTS

AUTHOR: MARA RUBIA MACHADO COUTO

GUIDING: ALESSANDRO DAL'COL LÚCIO

Place and date of defense: Santa Maria, April 30th 2008

In olericulture, the highest interest is aimed at the production of commercial fruits. These experiments usually evaluate variables such as phytomass and number of fruits. However, these variables present a peculiar characteristic related to their presence or absence at the moment of harvesting. This attribute is represented by the occurrence of zero values for these variables, generating some variability and affecting the data variation analysis. In this study, experiments with summer squash, sweet pepper and broccoli were analyzed using the Box-Cox transformation family. Since it is restricted for null values, an adaptation suggested by Yamamura (1999) was used. It was concluded that for each culture and each variable studied, there is a need for differentiated transformations. In this context, in the experiment using summer squash, the indicated transformation is the square root. For sweet pepper, the indicated transformation is the inverted square root and for broccoli, the indicated transformation is the logarithmic function. Despite the reduced variability observed in all experiments, the use of the method was not sufficient to make these variances homoscedastic.

Key words: Summer squash, sweet pepper, broccoli, Box-Cox, analysis of variance.

LISTA DE FIGURAS

FIGURA 1 – Freqüência de valores zero para as variáveis número e fitomassa de frutos de abobrinha italiana, cultivada nas estações sazonais verão/outono (época 1) e inverno/primavera (época 2). Santa Maria – RS. 2008.	30
FIGURA 2 – Potência λ em função da raiz do quadrado médio do erro (RQME) pelo método de Box-Cox, com intervalo de confiança à 99%, para a fitomassa de frutos de abobrinha italiana cultivada nas estações sazonais verão/outono (a) e inverno/primavera (b). Santa Maria – RS. 2008.	31
FIGURA 3 – Potência λ em função da raiz do quadrado médio do erro (RQME) pelo método de Box-Cox, com intervalo de confiança à 99%, para o número de frutos de abobrinha italiana cultivada nas estações sazonais verão/outono (a) e inverno/primavera (b). Santa Maria – RS. 2008.	32
FIGURA 4 – Resíduos em função do valor normal esperado para o experimento com abobrinha italiana. FIGURAs 4.1(a), (b) e (c) para o número de frutos, número de frutos mais constante e transformação da variável número de frutos mais constante, respectivamente; para a estação verão/outono. Figuras 4.2(a), (b) e (c) para o número de frutos, número de frutos mais constante e transformação da variável número de frutos mais constante, respectivamente; para a estação inverno/primavera. Santa Maria – RS. 2008.	36
FIGURA 5 – Resíduos em função do valor normal esperado, para o experimento com abobrinha italiana. Figuras 5.1(a), (b) e (c) para a variável fitomassa de frutos, fitomassa mais constante e transformação da variável fitomassa de frutos mais constante, respectivamente; para a estação verão/outono. Figuras 6.2(a), (b) e (c) para a variável fitomassa de frutos, fitomassa mais constante e transformação da variável fitomassa de frutos mais constante, respectivamente; para a estação inverno/primavera. Santa Maria – RS. 2008.	37
FIGURA 6 – Freqüência de valores zero para as variáveis número e fitomassa de frutos de pimentão, cultivado nas estações sazonais verão/outono (época 1) e inverno/primavera (época 2). Santa Maria – RS. 2008.	38
FIGURA 7 – Potência λ em função da raiz do quadrado médio do erro (RQME) pelo método de Box-Cox, com intervalo de confiança à 99%, para a fitomassa de frutos de pimentão cultivado nas estações sazonais verão/outono (a) e inverno/primavera (b). Santa Maria – RS. 2008.	39

FIGURA 8 – Potência λ em função da raiz do quadrado médio do erro (RQME) pelo método de Box-Cox, com intervalo de confiança à 99%, para o número de frutos de pimentão cultivado nas estações sazonais verão/outono (a) e inverno/primavera (b). Santa Maria – RS. 2008.	40
FIGURA 9 – Resíduos em função do valor normal esperado, para o experimento com pimentão. Figuras 9.1(a), (b) e (c) para o número de frutos colhidos, número de frutos mais constante e transformação da variável número de frutos mais constante, respectivamente; para a estação verão/outono. Figuras 11.2(a), (b) e (c) para o número de frutos colhidos, número de frutos mais constante e transformação da variável número de frutos mais constante, respectivamente; para a estação inverno/primavera. Santa Maria – RS. 2008.	44
FIGURA 10 – Resíduos em função do valor normal esperado, para o experimento com pimentão. Figuras 10.1(a), (b) e (c) para a variável fitomassa de frutos, fitomassa mais constante e transformação da variável fitomassa de frutos mais constante, respectivamente; para a estação verão/outono. Figuras 10.2(a), (b) e (c) para a variável fitomassa de frutos, fitomassa mais constante e transformação da variável fitomassa de frutos mais constante, respectivamente; para a estação inverno/primavera. Santa Maria – RS. 2008.	45
FIGURA 11 – Frequência de valores zero para as variáveis Cabeças Comerciais (CC), Ramificações Comerciais (RC), Cabeças Não Comerciais (CNC) e Ramificações Não Comerciais (RNC) para a cultura de brócolis, em quatro épocas. Santa Maria – RS. 2008.	46
FIGURA 12 – Potência λ em função da raiz do quadrado médio do erro (RQME) pelo método de Box-Cox, com intervalo de confiança à 99%. Figura (a), (b), (c) e (d), época 1, 2, 3 e 4, respectivamente, para Cabeças Comerciais (CC) de brócolis. Santa Maria – RS. 2008.	47
FIGURA 13 – Potência λ em função da raiz do quadrado médio do erro (RQME) pelo método de Box-Cox, com intervalo de confiança à 99%. Figura (a), (b), (c) e (d), época 1, 2, 3 e 4, respectivamente, para Ramificações Comerciais (RC) de brócolis. Santa Maria – RS. 2008.	49
FIGURA 14 – Potência λ em função da raiz do quadrado médio do erro (RQME) pelo método de Box-Cox, com intervalo de confiança à 99%. Figura (a), (b), (c) e (d), época 1, 2, 3 e 4, respectivamente, para Cabeças Não Ramificações Comerciais (CNC) de brócolis. Santa Maria – RS. 2008.	51
FIGURA 15 – Potência λ em função da raiz do quadrado médio do erro (RQME) pelo método de Box-Cox, com intervalo de confiança à 99%. Figura (a), (b), (c) e (d), época 1, 2, 3 e 4, respectivamente, para Ramificações Não Comerciais (RNC) de brócolis. Santa Maria – RS. 2008.	52
FIGURA 16 – Resíduos em função do valor normal esperado para o experimento com brócolis. Figuras 16.1, 16.2, 16.3 e 16.4, (a), (b) e (c) para as variáveis cabeças comerciais (CC), cabeças comerciais (CC) mais constante e transformação da variável, cabeças comerciais (CC) mais constante, respectivamente; para as quatro épocas. Santa Maria – RS. 2008.	58

FIGURA 17 – Resíduos em função do valor normal esperado para o experimento com brócolis. Figuras 17.1, 17.2, 17.3 e 17.4, (a), (b) e (c) para as variáveis ramificações comerciais (RC), ramificações comerciais (RC) mais constante e transformação da variável ramificações comerciais (RC) mais constante, respectivamente; para as quatro épocas. Santa Maria – RS. 2008.	59
FIGURA 18 – Resíduos em função do valor normal esperado para o experimento com brócolis. Figuras 18.1, 18. 2, 18.3 e 18.4, (a), (b) e (c) para a variável cabeças não comerciais (CNC), cabeças não comerciais (CNC) mais constante e transformação da variável cabeças não comerciais (CNC) mais constante, respectivamente; para as quatro épocas. Santa Maria – RS. 2008.	60
FIGURA 19 – Resíduos em função do valor normal esperado para o experimento com brócolis. Figuras 19.1, 19.2, 19.3 e 19.4, (a), (b) e (c) para as variáveis ramificações não comerciais (RNC), ramificações não comerciais (RNC) mais constante e transformação da variável ramificações não comerciais (RNC) mais constante, respectivamente; para as quatro épocas. Santa Maria – RS. 2008.	61

LISTA DE TABELAS

TABELA 1 – Resultados do procedimento Box-Cox na escolha da potência λ , para a fitomassa de frutos de abobrinha italiana, cultivada nas estações sazonais verão/outono (época 1) e inverno/primavera (época 2). Santa Maria – RS. 2008.	31
TABELA 2 – Resultados do procedimento Box-Cox na escolha da potência λ , para o número de frutos de abobrinha italiana, cultivada nas estações sazonais verão/outono (época 1) e inverno/primavera (época 2). Santa Maria – RS. 2008.	32
TABELA 3 – Medidas Descritivas para as variáveis original, variável mais constante e transformação da variável mais constante para o número e fitomassa de frutos de abobrinha italiana cultivada nas estações sazonais verão/outono (a) e inverno/primavera (b). Santa Maria – RS. 2008.	33
TABELA 4 – Resumo da análise de variância para a variável original, variável acrescida da constante e transformação da variável acrescida da constante para o número de frutos de abobrinha italiana, cultivada nas estações sazonais verão/outono (época 1) e inverno/primavera (época 2). Santa Maria – RS. 2008.	34
TABELA 5 – Resumo da análise de variância para a variável original, variável acrescida da constante e transformação da variável acrescida da constante para a fitomassa de frutos de abobrinha italiana, cultivada nas estações sazonais verão/outono (época 1) e inverno/primavera (época 2). Santa Maria – RS. 2008.	35
TABELA 6 – Resultados do teste de Bartlett para homogeneidade de variâncias, para variável acrescida da constante e transformação da variável acrescida da constante para o número e fitomassa de frutos de abobrinha italiana, cultivada nas estações sazonais verão/outono (época 1) e inverno/primavera (época 2). Santa Maria – RS. 2008.	35
TABELA 7 – Resultados do procedimento Box-Cox na escolha da potência λ , para a fitomassa de frutos de pimentão, cultivado nas estações sazonais verão/outono (época 1) e inverno/primavera (época 2). Santa Maria – RS. 2008.	39
TABELA 8 – Resultados do procedimento Box-Cox na escolha da potência λ , para o variável número de frutos de pimentão, cultivado nas estações sazonais verão/outono (época 1) e inverno/primavera (época 2). Santa Maria – RS. 2008.	40

TABELA 9 – Medidas Descritivas para as variáveis original, variável mais constante e transformação da variável mais constante para o número e fitomassa de frutos de pimentão cultivado nas estações sazonais verão/outono (a) e inverno/primavera (b). Santa Maria – RS. 2008.	41
TABELA 10 – Resumo da análise de variância para a variável original, variável acrescida da constante e transformação da variável acrescida da constante para o número e fitomassa de frutos de pimentão, cultivado nas estações sazonais verão/outono (época 1) e inverno/primavera (época 2). Santa Maria – RS. 2008.	42
TABELA 11 – Resultados do teste de Bartlett para homogeneidade de variâncias, para variável acrescida da constante e transformação da variável acrescida da constante para o número e fitomassa de frutos de pimentão, cultivado nas estações sazonais verão/outono (época 1) e inverno/primavera (época 2). Santa Maria – RS. 2008.	43
TABELA 12 – Resultados do procedimento Box-Cox na escolha da potência λ , para a variável Cabeças Comerciais (CC) de brócolis, cultivado em quatro épocas. Santa Maria – RS. 2008.	47
TABELA 13 – Resultados do procedimento Box-Cox na escolha da potência λ , para a variável Ramificações Comerciais (RC) de brócolis, cultivado em quatro épocas. Santa Maria – RS. 2008.	48
TABELA 14 – Resultados do procedimento Box-Cox na escolha da potência λ , para a variável Cabeças Não Comerciais (CNC) de brócolis, cultivado em quatro épocas. Santa Maria – RS. 2008.	50
TABELA 15 – Resultados do procedimento Box-Cox na escolha da potência λ , para a variável Ramificações Não Comerciais (RNC) de brócolis, cultivado em quatro épocas. Santa Maria – RS. 2008.	51
TABELA 16 – Medidas descritivas para as variáveis original, variável mais constante e transformação da variável mais constante de Cabeças Comerciais (CC), de Ramificações Comerciais (RC), de Cabeças Não Comerciais (CNC) e de Ramificações Não Comerciais (RNC) para a cultura de brócolis, em quatro épocas.	53
TABELA 17 – Resumo da análise de variância para a variável original, variável acrescida da constante e transformação da variável acrescida da constante de Cabeças Comerciais (CC) em quatro épocas. Santa Maria – RS. 2008.	54
TABELA 18 – Resumo da análise de variância para a variável original, variável acrescida da constante e transformação da variável acrescida da constante de Ramificações Comerciais (RC) em quatro épocas. Santa Maria – RS. 2008.	55
TABELA 19 – Resumo da análise de variância para a variável original, variável acrescida da constante e transformação da variável acrescida da constante de Cabeças Não Comerciais (CNC) em quatro épocas. Santa Maria – RS. 2008.	56

TABELA 20 – Resumo da análise de variância para a variável original, variável acrescida da constante e transformação da variável acrescida da constante de Ramificações Não Comerciais (RNC) em quatro épocas. Santa Maria – RS. 2008. 56

TABELA 21 – Resultados do teste de Bartlett para homogeneidade de variâncias, para variável acrescida da constante e transformação da variável acrescida da constante das variáveis cabeças e ramificações comerciais e não comerciais, para a cultura de brócolis cultivado em quatro épocas. Santa Maria – RS. 2008. 57

LISTA DE APÊNDICES

APÊNDICE A – Diferença em termos percentuais (%) do coeficiente de variação antes e após transformação de fitomassa e número de frutos de abobrinha italiana cultivados nas estações sazonais verão/outono (época 1) e inverno/primavera (época 2). Santa Maria – RS.	70
APÊNDICE A – Diferença em termos percentuais (%) do coeficiente de variação antes e após transformação de fitomassa e número de frutos de pimentão, cultivados nas estações sazonais verão/outono (época 1) e inverno/primavera (época 2). Santa Maria – RS. 2008.	71
APÊNDICE B - Diferença em termos percentuais (%) do coeficiente de variação antes e após transformação de cabeças comerciais e não comerciais, e ramificações comerciais e não comerciais de brócolis, cultivado em quatro estações. Santa Maria – RS. 2008.	72

SUMÁRIO

1 INTRODUÇÃO	18
2 REVISÃO DE LITERATURA.....	20
2.1 Análise de variância	21
2.1.1 Pressuposições do modelo estatístico	22
2.1.2 Dados de contagens inflacionados de zero	23
2.2 Transformações Box-Cox	23
2.2.1 Família de transformações Box-Cox	24
2.2.2 Estimação da potência λ	24
2.2.3 Transformações Box-Cox aplicadas a culturas olerícolas .	25
2.3 Heterogeneidade de variância e excesso de zero em dados de culturas olerícolas	26
3 MATERIAL E MÉTODOS	28
4 RESULTADOS E DISCUSSÃO	30
4.1 Experimento com abobrinha italiana	30
4.2 Experimento com pimentão	38
4.3 Experimento com brócolis	45
4.4 Considerações Finais.....	62
5 CONCLUSÕES	64
6 REFERÊNCIAS BIBLIOGRÁFICAS	65
7 APÊNDICES	70

1 INTRODUÇÃO

Em experimentos agrícolas, o interesse é avaliar os efeitos dos tratamentos, no entanto, a própria aplicação destes pode causar certa variabilidade. Existem, casos em que outros tipos de variabilidade podem estar associados e esta pode não estar sob o domínio do pesquisador, isto é, ocorre devido à variabilidade inerente à variável resposta coletada no experimento. Em dados provindos de contagens uma variabilidade maior que a esperada pode ser decorrente de um excesso de zeros, pois nem sempre é possível observar a ocorrência de algum valor.

Em experimentos com olerícolas, o interesse maior está voltado para a produção de frutos comercializáveis, sendo esta variável obtida por contagem. A ocorrência de valores zeros é bastante freqüente nestes tipos de experimentos, pois nem sempre existem frutos a serem colhidos; ou nem sempre estes possuem tamanho adequado para a comercialização. Sendo assim, esta ocorrência de zeros gera variabilidade que interfere nos procedimentos de inferência derivados da análise de variação que se fundamentam num modelo estatístico e nas correspondentes pressuposições. As pressuposições do modelo estatístico visam assegurar a validade das inferências de interesse que serão derivadas do experimento. Entretanto, distorções ocorrem quando essas pressuposições são violadas, e a não satisfação de qualquer pressuposição pode afetar os níveis de significância e a sensibilidade dos testes.

Um procedimento de análise de dados, em que alguma pressuposição da técnica estatística tenha sido violada, é o uso de uma transformação, cuja idéia básica é encontrar alguma função conveniente que seja capaz de adequar a variável original às pressuposições do modelo. Além de corrigir os desvios das pressuposições, a transformação pode ter outras aplicações, tais como linearizar ou simplificar um modelo de regressão ou ainda eliminar a interação de um modelo de análise de variância.

Em estudos anteriores, verificou-se que a presença de zeros nas parcelas colhidas provoca conseqüências na variância, o que afeta diretamente o quadrado médio do erro (QME) e o teste F. A possibilidade de se encontrar uma transformação adequada que permita a redução do QME é justificável pelo fato que

quanto maior o erro, menor a probabilidade de se encontrarem diferenças significativas entre médias de tratamentos. Portanto, a possibilidade de se encontrar uma transformação adequada que permita estabilizar a variância, ou que, pelo menos reduza esta variabilidade, pode ser de grande auxílio na análise de variação deste tipo de experimento.

Embora algumas transformações de dados sejam conhecidas para algumas formas de relação entre variância e média, como por exemplo, a transformação arco-seno para dados provindos de proporção; o desenvolvimento de estudos se faz necessário, devido ao desconhecimento de uma transformação apropriada à dados de contagem com excesso de zeros.

Desta forma, os objetivos do trabalho foram apresentar o uso de transformações como um possível procedimento para satisfazer as pressuposições dos modelos; e verificar a necessidade de transformações diferenciadas para cada cultura; e, definir uma transformação adequada para os dados de cada cultura visando estabilizar a variabilidade gerada pelo excesso de zeros.

2 REVISÃO DE LITERATURA

Antes de se aplicar qualquer técnica estatística, os dados devem ser examinados cuidadosamente e, somente depois, um tratamento especial pode ser justificado. Aconselha-se que uma análise exploratória seja realizada para se obter informações adicionais e maior sensibilidade sobre a variável em estudo.

Em casos onde as pressuposições dos modelos estatísticos tenham sido violadas, as inferências sobre o experimento podem ser prejudicadas, pois essas alteram o nível de significância, a precisão e os testes de hipóteses perdem em sensibilidade (Montgomery, 1991).

Sabe-se que, na prática, as pressuposições de modelos nem sempre se encontram perfeitamente satisfeitas. Ao se analisarem dados que não estão de acordo com as pressuposições da técnica a ser utilizada, pode-se optar por dois procedimentos. Segundo Siqueira (1983), um deles seria a busca de novos métodos de análise que se ajusta melhor a estes dados; e, um segundo, seria adequar os dados às pressuposições através de uma transformação.

Detectada a necessidade do uso de uma transformação, o problema consiste na escolha da transformação adequada para os dados analisados. Campos (2000), diz que, o uso de transformações a dados amostrais é um recurso que sempre vale a pena tentar, porque a estatística paramétrica é evidentemente mais poderosa que a não-paramétrica. Esta forma de análise estatística dos dados somente foi desenvolvida como um recurso complementar destinado a suprir a necessidade de testes estatísticos nos casos em que alguma restrição desaconselha o uso da estatística paramétrica, ou quando a própria natureza dos dados, muitas vezes não exatamente numéricos, impossibilita a aplicação desta.

A escolha da transformação feita pelo método de tentativas, em muitos casos, pode ser bastante trabalhosa e dispendiosa. Segundo Siqueira (1983), existe na literatura muitas sugestões mais práticas de se escolher a transformação a ser utilizada, mas o procedimento adotado deve depender do objetivo da transformação, isto é, se o objetivo é obter a normalidade, a homogeneidade de variâncias dos erros ou a aditividade do modelo.

Um exemplo de heterocedasticidade das variâncias refere-se a dados provindos de contagens, pois alguns desses processos ocasionam uma variabilidade maior do que a esperada. Segundo Borgatto (2004), esse fenômeno é conhecido como superdispersão, podendo ocorrer devido: a variabilidade da média, excesso de zero, correlação entre indivíduos e/ou omissão de variáveis não observadas.

Uma metodologia geral para modelar dados de contagens, inflacionados de zeros (modelos ZIP – zero-inflated Poisson regression), é discutida por Ridout, Demétrio e Hinde (1998) em que é apresentada uma revisão de literatura com ênfase em horticultura. Alguns autores como Vieira (1998) e Vieira, Hinde, Demétrio (2000) aplicam esses modelos em experimento de controle biológico de pragas. Outros estudos sobre excessos de zeros foram também realizados; por Hall & Berenhaut (2002) na simulação de dados sobre raízes de maçã, apresentados por Hinde & Demétrio (2000) e, sobre ovos de jacaré; por Ghosh, Mukhopadhyay e Lu (2004) na indústria, em processos de fabricação modelando o número de defeitos; e por Lee et al. (2005); aplicam a metodologia no número de internações de uma equipe de limpeza, num estudo de ergonomia.

É importante observar que para dados com essas características o uso de algumas transformações pode ser matematicamente impedido, isto é, de nada adianta aplicar uma transformação raiz quadrada em um valor zero, ou ainda aplicar uma transformação logarítmica nesse valor, pois o zero não pertence ao domínio dessa função. Uma vez normalizada e homogeneizada a distribuição dos dados amostrais, por intermédio da transformação que se comprovar mais conveniente, o pesquisador poderá utilizar os testes paramétricos, sem prejuízos em sua interpretação e conclusão.

2.1 Análise de variância

Quando as pressuposições do modelo são violadas, a análise resultante pode levar a resultados equivocados. Segundo Borgatto (2004), este tipo de violação do modelo dá origem às chamadas falhas sistemáticas (não linearidade, não-normalidade, não homocedasticidade, não-independência, entre outros), que acarretam certo viés no modelo para predizer os valores, o que pode levar a super ou subestimação de variáveis importantes.

2.1.1 Pressuposições do modelo estatístico

A adequação do modelo estatístico é fundamental para a validade das inferências de interesse sobre o experimento, pois a falha de qualquer pressuposição afeta os níveis de significância e a sensibilidade do teste F e de outros testes derivados da análise da variância. Embora esses testes sejam tolerantes a graus moderados de desvio das pressuposições qualquer desvio aparentemente importante deve ser verificado e corrigido (MONTGOMERY, 1991).

A primeira pressuposição refere-se à descrição do modelo propriamente dito, isto é, a aditividade dos efeitos das fontes de variação da variável resposta. Os efeitos de duas fontes de variação de uma variável resposta são aditivos se o efeito de uma dessas fontes permanece constante entre os níveis da outra fonte.

As demais pressuposições referem-se aos termos do erro, sendo que estes devem ser: independentes, normalmente distribuídos e homocedásticos (MONTGOMERY, 1991).

A pressuposição de independência dos erros experimentais, em combinação com a pressuposição de distribuição normal destes, implica que os erros são mutuamente independentes, ou seja, que a probabilidade de que o erro de qualquer unidade experimental tenha um valor particular não depende dos valores dos erros das outras unidades. A ausência de correlação é usualmente assegurada pela atribuição aleatória das unidades experimentais aos tratamentos, e pela casualização.

A normalidade dos erros é a menos provável de ser válida, pois em situações reais, freqüentemente essa se verifica, apenas, aproximadamente. Esta pressuposição é essencial para a validade de sentenças probabilísticas referentes às decisões baseadas em testes de hipóteses e à confiabilidade de estimativas por intervalo (MONTGOMERY, 1991).

A homogeneidade de variâncias, também denominada homocedasticidade, é a pressuposição que usualmente requer maior cuidado. Em algumas situações, é possível que uma transformação da variável possibilite que as pressuposições sejam mais aproximadamente satisfeitas para as variáveis transformadas (MONTGOMERY, 1991).

De modo geral, em situações em que a normalidade e a homogeneidade de variâncias do modelo estatístico não são atendidas estas podem causar pequenas diferenças nos níveis de significância dos testes em relação aos níveis nominais

adotados, e também reduzir a potência do teste. Em particular, em estimativas por intervalo de componentes de variância, coeficientes de confiança podem diferir dos valores nominais (MONTGOMERY, 1991).

2.1.2 Dados de contagens inflacionados de zeros

Em diversas áreas do conhecimento, a presença de dados inflacionados de zeros é observada. Na literatura, encontram-se basicamente duas técnicas de investigação sobre dados com excessos de zero. A primeira discute modelos lineares generalizados que incorporam a superdispersão, tais como: Demétrio & Hinde (1997), Hinde & Demétrio (1998, a, b) e Hinde & Demétrio (2000), que aplicam esses modelos a alguns exemplos como germinação de *Orobanche*, sobrevivência de ratos em estudo toxicológico e sobrevivência de ovos de truta. A segunda defende o uso de modelos de dois estágios, isto é, modelos de probabilidade compostos, tais como: Poisson-normal, Beta-binomial e Logístico-normal. Conforme estudos desenvolvidos por Grogger & Carson (1991) sobre pesca desportiva, Lambert (1992) sobre processos de fabricação em indústria; Ridout & Demétrio (1992) sobre número de raízes de maçã em diferentes concentrações de citocinina.

No âmbito da agricultura e horticultura, alguns estudos sobre modelos para dados de proporções inflacionados de zero, foram desenvolvidos por Vieira (1998) e Vieira, Hinde, Demétrio (2000) no controle de pragas. Também, Hall (2000) usou modelos com distribuição Binomial e Poisson aplicados à horticultura.

2.2 Transformações Box-Cox

Na literatura, até por volta de 1950, a maioria das publicações tratava de transformações específicas como a logarítmica, a raiz quadrada, entre outras. No início da década de 50, surgiu a idéia do uso de famílias de transformações com Anscombe & Tukey (1952); Moore & Tukey (1954) e Tukey (1957).

Os primeiros estudos foram desenvolvidos por Tukey que definiu uma classe de funções potências. Mais tarde em 1964, Box-Cox, com a finalidade de evitar a descontinuidade para a potência igual à zero, que ocorria na família definida por Tukey, define uma nova família de transformações potência.

Uma informação importante é que o valor da estatística F da análise de variância é o mesmo para ambas as famílias, logo as conclusões não são alteradas, conforme Siqueira (1983). Sendo que a família de transformações Box-Cox é a que ganhou maior destaque no meio científico, e a qual se utilizou neste estudo.

2.2.1 Família de transformações Box-Cox

O procedimento desenvolvido por Box-Cox (1964) produz uma transformação da variável resposta (Y) tal que as pressuposições como homocedasticidade e normalidade do modelo estatístico da análise de variação sejam simultaneamente satisfeitas.

Box-Cox (1964) definiram a seguinte família de transformações:

$$y^{(\lambda)} = \begin{cases} \frac{y^\lambda - 1}{\lambda}, & \lambda \neq 0 \\ \ln y, & \lambda = 0 \end{cases} \quad (1)$$

A equação (1) pode ser expressa pela transformação normalizada a seguir, onde o jacobiano $J(\lambda, z) = 1$.

$$Z^{(\lambda)} = \begin{cases} \frac{y^\lambda - 1}{\lambda \bar{y}^{\lambda-1}}, & \lambda \neq 0 \\ \bar{y} \ln y, & \lambda = 0 \end{cases} \quad (2)$$

onde: λ é um número real e \bar{y} é a média geométrica das observações da variável y.

2.2.2 Estimação da potência λ

No procedimento Box-cox, λ é considerado como um parâmetro adicional e desconhecido no modelo e é estimado através dos métodos de inferência estatística como o de máxima-verossimilhança e o método Bayesiano. Em geral, esses dois métodos produzem estimativas muito próximas (SIQUEIRA, 1983).

Na estimativa pelo método da máxima verossimilhança (Siqueira, 1983), a estimação de λ consiste em determinar o estimador de máxima verossimilhança, o qual é expresso por:

$$L_{\max(\lambda)} = -\frac{n}{2} \ln\left(\frac{\hat{\sigma}_\lambda^2(z)}{n}\right) = -\frac{n}{2} \ln\left(\frac{SSE(\lambda)}{n}\right) \quad (3)$$

onde $SSE(\lambda)$ é a soma de quadrados residual, n é o número de observações e $\hat{\sigma}_\lambda^2(z)$ é a variância da variável (Z) transformada pela potência λ .

O problema é encontrar o valor de λ que maximiza $L_{\max(\lambda)}$ ou equivalentemente minimize $SSE(\lambda)$. Na prática, se constrói um intervalo de confiança para λ com $(1-\alpha)\%$ de confiança, baseado no fato de que $-2 \log \Lambda$ segue uma distribuição aproximadamente χ^2 , onde Λ é a razão de verossimilhança generalizada.

De modo prático, pode ser traçada uma linha que represente o intervalo de confiança para λ , no ponto $L_{(\max(\lambda))} - \frac{1}{2} \chi^2(\alpha)$ em um gráfico de $L_{\max(\lambda)}$ em função de λ (Draper & Smith, 1981).

2.2.3 Transformações Box-Cox aplicadas a culturas olerícolas

O uso de transformações de variáveis é necessário, em muitos casos, para que se possa atingir a homocedasticidade na análise de variância. Em experimentos onde há a incidência de variáveis que contêm zero, como é o caso de experimentos com culturas olerícolas, a eficácia das transformações fica prejudicada; especialmente transformações que empregam a função logarítmica, inclusive as transformações Box-Cox.

Estudos que descrevem a variabilidade de populações utilizam freqüentemente a transformação logarítmica para alcançar a homocedasticidade ou a estabilidade da variância para testar os efeitos na análise de variância (ANOVA). Em alguns destes trabalhos, se os dados contêm zero, é comum o acréscimo de uma constante a cada um dos dados, tradicionalmente é adicionado o valor um. 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 2, embora o acréscimo de uma constante influencie o resultado da ANOVA.

A determinação do valor constante c que deve ser adicionada a variável original (Y), tem sido bastante discutida na literatura. Box & Cox (1964) sugerem uma família de transformações modificadas através do método de máxima-verossimilhança. Outros princípios alternativos foram propostos por Hill (1963); Griffiths (1980) e Berry (1987) para determinar o parâmetro c .

Os procedimentos propostos por esses autores são aplicáveis em determinado conjunto de pressupostos, os quais requerem cálculos complexos. Devido a este fato, um princípio prático foi proposto por Yamamura (1999) tendo por finalidade estabilizar a variância de populações, em ecologia, o qual sugere como uma escolha razoável para o parâmetro c o valor 0,5, se a forma da distribuição não for conhecida. O autor defende o uso de c igual a 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 usa-se $c = 0,5$.

2.3 Considerações sobre heterogeneidade de variância e excesso de zero em culturas olerícolas

Na produção de olerícolas, o número de frutos comercializáveis é de grande interesse. No entanto, a ocorrência de valores zeros é bastante freqüente, pois nem sempre existem frutos a serem colhidos; ou nem sempre estes possuem tamanho adequado para sua colheita e comercialização.

Esta freqüência de zeros gera variabilidade, o que pode provocar distorções nos procedimentos de inferência derivados da análise de variância.

Ao estudar experimentos com olerícolas, Souza et al. (2002), Lorentz et al. (2005) e Carpes (2006), puderam verificar que a presença de zeros nas parcelas colhidas tem conseqüências na variância, isto é, o excesso de zeros afeta a homogeneidade de variâncias nos testes de hipóteses.

Muitas circunstâncias podem violar a pressuposição de homogeneidade de variância. Em geral, a heterogeneidade da variância pode ser associada com variáveis respostas cuja distribuição não é normal e cuja variância relaciona-se com a média em decorrência da própria forma da distribuição.

Em estudo sobre o efeito da média e heterogeneidade de variância em avaliações genéticas de ovinos, Nikolau *et al.* (2004), empregam a transformação logarítmica e raiz quadrada a dados de coeficientes de herdabilidade. Um dos resultados encontrados foi que o emprego de transformação (logarítmica ou raiz quadrada) para o ajuste da heterogeneidade não resultou em variâncias homogêneas.

3 MATERIAL E MÉTODOS

Neste trabalho foram utilizados dados de experimentos já desenvolvidos com as culturas de pimentão, abobrinha italiana e brócolis, os quais estão descritos a seguir.

Os dados provenientes do experimento com abobrinha italiana, cultivar *Caserta*, que foi conduzido em dois túneis plásticos no Departamento de Fitotecnia da Universidade Federal de Santa Maria em 2004/2005, em duas épocas verão/outono (época 1) e inverno/primavera (época 2). Utilizou-se um delineamento em blocos ao acaso em faixas para o fator A (planta) e o fator D (colheita), avaliando-se as variáveis respostas: fitomassa e número de frutos.

Para os dados de pimentão, cultivar *vidi*, dois experimentos desenvolvidos em duas estufas plásticas, foram realizados na área experimental do Departamento de Fitotecnia da Universidade Federal de Santa Maria, em duas épocas verão/outono (época 1) e inverno/primavera (época 2). Utilizou-se um delineamento em blocos ao acaso em faixas para o fator A (planta) e o fator D (colheita), avaliando-se as variáveis respostas: fitomassa e número de frutos.

Também foram utilizados os dados coletados do experimento com cultivares de brócolis, que foi realizado no ano agrícola de 1996, em condições de campo, no campus da Universidade Federal de Santa Maria, em quatro épocas. Utilizou-se o delineamento experimental de blocos ao acaso bifatorial com parcela sub-dividida no tempo e amostragem na sub-parcela. As variáveis respostas avaliadas foram: cabeças comerciais e não comerciais, e ramificações comerciais e não comerciais.

Para análise dos dados destes experimentos realizou-se uma análise exploratória a fim de se obter mais informações sobre o comportamento das distribuições de probabilidade das variáveis analisadas. Utilizou-se a metodologia Box-Cox, para o conjunto de funções potências, para encontrar uma transformação que estabilize a variância, ou que, pelo menos reduza esta variabilidade. Como para valores nulos, a família de transformações de Box-Cox fica restrita; utilizou-se a variável somada a uma constante. A falta de conhecimento sobre a distribuição de probabilidade das variáveis é a característica que leva a adotar-se o procedimento proposto por Yamamura (1999), o qual sugere como parâmetro c o valor 0,5. E está expresso a seguir:

$$\begin{aligned}
 f(y) &= \frac{(y+0,5)^\lambda - 1}{\lambda}, & \lambda \neq 0 \\
 f(y) &= \ln(y+0,5), & \lambda = 0
 \end{aligned}
 \tag{4}$$

Observa-se que a expressão (4) nada mais é do que a família de transformações Box-Cox acrescida do parâmetro $c = 0,5$.

Para a determinação da transformação mais adequada, consideraram-se (y) as variáveis repostas de interesse em cada experimento.

Para a escolha da melhor potência λ , considerou-se que λ assume valores no intervalo $[-2, 2]$, conforme Drapper & Smith (1981), que sugerem utilizar de 11 a 21 valores de λ pertencentes a esse intervalo. Cada valor λ , no intervalo $[-2, 2]$, foi substituído na equação (2), o que gerou valores transformados ($Z^{(\lambda)}$) para cada variável resposta (y) de interesse. Após, foi realizada a análise de variância para cada variável transformada, conforme o delineamento empregado.

Para determinar o estimador de máxima verossimilhança (L_{max}), substitui-se a soma de quadrados residual ($SSE_{(\lambda)}$), encontrada para cada valor de λ usado, na equação (3). Sendo que o melhor λ foi escolhido de maneira que maximizasse (L_{max}).

Após a escolha da transformação mais adequada, foram calculadas as medidas descritivas para as variáveis, antes e após o uso de transformação. Também se realizou a análise de variância para as variáveis, antes e após o uso de transformação, e aplicou-se o teste de Bartlett para conferir se o uso da transformação foi suficiente para que a suposição de homocedasticidade tenha sido satisfeita. Por fim, investigou-se visualmente o comportamento dos resíduos. Para a análise dos dados utilizaram-se os softwares *Statistica 7.0* e *SAS*.

4 RESULTADOS E DISCUSSÃO

4.1 Experimento com abobrinha italiana

Na Figura 1, observa-se a freqüência de valores zero para as variáveis, número e fitomassa de frutos de abobrinha italiana, nas estações verão/outono (época 1) e inverno/primavera (época 2).

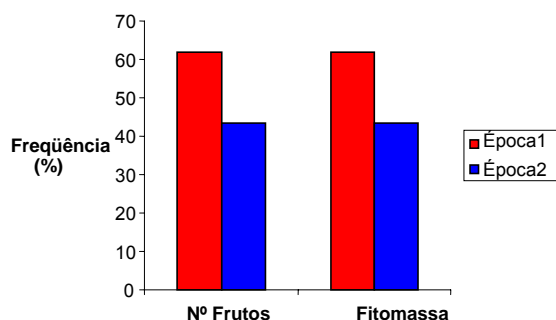


Figura 1 – Frequência de valores zero para as variáveis número e fitomassa de frutos de abobrinha italiana, cultivada nas estações sazonais verão/outono (época 1) e inverno/primavera (época 2). Santa Maria – RS. 2008.

Observa-se na Figura 1, frequência de valores zero varia em torno de 40%, na época 2; e 60%, na época 1. Observa-se também, que essas duas variáveis estão relacionadas, pois a ausência de frutos afeta o peso de frutos colhidos.

Na Tabela 1 encontram-se para vários valores de λ , a raiz do quadrado médio do erro (RQME) e a função de máxima verossimilhança (L_{max}) para a fitomassa de frutos colhidos para duas épocas. Analisando-se a Tabela 1, o valor da potencia λ que minimiza a RQME é 0,25, indicando que a melhor transformação para os dados seria a raiz quarta.

Tabela 1 – Resultados do procedimento Box-Cox na escolha da potência λ , para a fitomassa de frutos de abobrinha italiana, cultivada nas estações sazonais verão/outono (época 1) e inverno/primavera (época 2). Santa Maria – RS. 2008

Box-Cox (λ)	Época 1		Época 1	
	L_{\max}	RQME	L_{\max}	RQME
-2,00	19919,61	22,4867	-58751,00	39,3280
-1,75	-17839,87	16,2471	-52775,50	27,0711
-1,50	-15874,04	11,9497	-47127,26	19,0193
-1,25	-14047,69	8,9826	-41883,75	13,7046
-1,00	-12392,71	6,9356	-37143,21	10,1904
-0,75	-10948,35	5,5342	-33028,48	7,8796
-0,50	-9761,29	4,5972	-29687,65	6,3948
-0,25	-8883,44	4,0078	-27285,09	5,5031
0,00	-8365,66	3,6963	-25974,24	5,0702
0,25	-8246,36	3,6280	-25851,59	5,0315
0,50	-8538,66	3,7976	-26914,73	5,3772
0,75	-9224,06	4,2270	-29058,92	6,1483
1,00	-10257,88	4,9681	-32117,44	7,4435
1,25	-11583,01	6,1112	-35913,43	9,4365
1,50	-13142,91	7,7983	-40292,85	12,4075
1,75	-14889,02	10,2448	-45134,51	16,7921
2,00	-16782,76	13,7731	-50347,16	23,2591

* L_{\max} = função de máxima verossimilhança; RQME = raiz do quadrado médio do erro.

Na Figura 2, tem-se o gráfico de λ em função da RQME o qual apresenta o intervalo de confiança de 99% para λ .

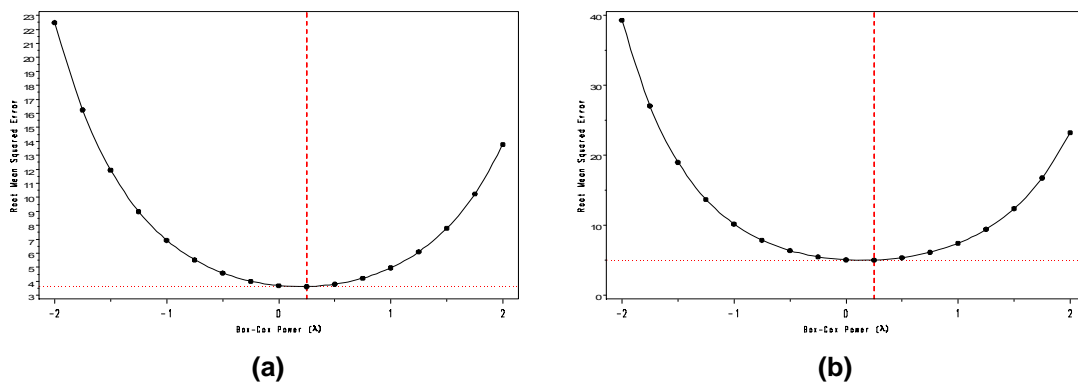


Figura 2 – Potência λ em função da raiz do quadrado médio do erro (RQME) pelo método de Box-Cox, com intervalo de confiança à 99%, para a fitomassa de frutos de abobrinha italiana cultivada nas estações sazonais verão/outono (a) e inverno/primavera (b). Santa Maria – RS. 2008.

A inspeção da Figura 2 mostra que o melhor estimador de máxima verossimilhança para λ é 0,25, o que também foi concluído a partir da Tabela 1.

Da mesma forma, a Tabela 2 e Figura 3, mostram os resultados para o melhor estimador de λ para o número de frutos colhidos.

Tabela 2 – Resultados do procedimento Box-Cox na escolha da potência λ , para o número de frutos de abobrinha italiana, cultivada nas estações sazonais verão/outono (época 1) e inverno/primavera (época 2). Santa Maria – RS. 2008.

Box-Cox (λ)	Época 1		Época 2	
	L_{\max}	RQME	L_{\max}	RQME
-2,00	-21147,13	27,2278	-63204,66	51,9505
-1,75	-18892,06	9,1420	-56550,28	34,2742
-1,50	-16746,51	13,6897	-50204,73	23,0530
-1,25	-14739,34	10,0044	-44264,04	15,9029
-1,00	-12908,27	7,5151	-38859,58	11,3444
-0,75	-11301,12	5,8462	-34162,63	8,4584
-0,50	-9974,77	4,7520	-30374,99	6,6755
-0,25	-8989,56	4,0740	-27693,79	5,6455
0,00	-8397,61	3,7141	-26252,50	5,1592
0,25	-8227,79	3,6168	-26070,74	5,1009
0,50	-8475,82	3,7597	-27053,03	5,4239
0,75	-9106,45	4,1491	-29035,05	6,1392
1,00	-10066,08	4,8202	-31838,01	7,3146
1,25	-11296,89	5,8424	-35302,12	9,0828
1,50	-12746,33	7,3273	-39297,64	11,6593
1,75	-14370,83	9,4446	-43723,58	15,3747
2,00	-16135,99	12,4441	-48502,26	20,7260

* L_{\max} = função de máxima verossimilhança; RQME = raiz do quadrado médio do erro.

Na Figura 3, tem-se o gráfico de λ em função da RQME para o número de frutos colhidos, o qual apresenta o intervalo de confiança de 99% para λ .

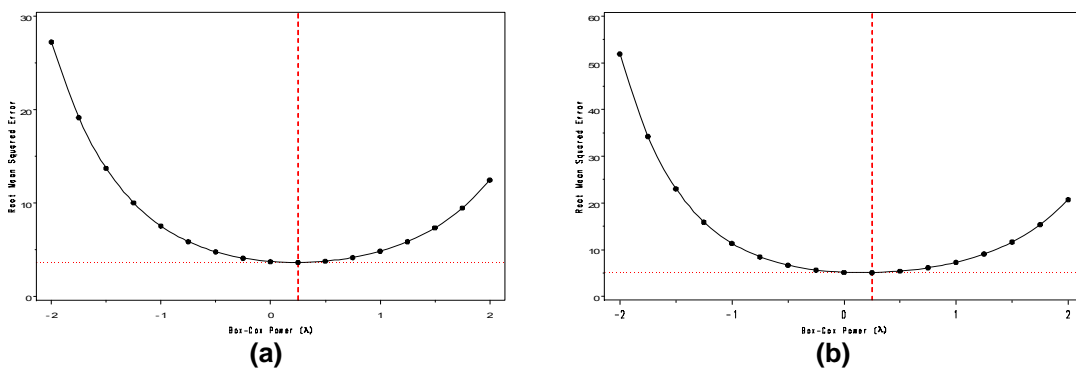


Figura 3 – Potência λ em função da raiz do quadrado médio do erro (RQME) pelo método de Box-Cox, com intervalo de confiança à 99%, para o número de frutos de abobrinha italiana cultivada nas estações sazonais verão/outono (a) e inverno/primavera (b). Santa Maria – RS. 2008.

Analisando-se a Tabela 2 e a Figura 3 tem-se que o melhor estimador de máxima verossimilhança para λ é 0,25. Também se observa a relação entre peso e número de frutos colhidos, o qual pode ter sido um dos motivos para o qual se tem o mesmo valor λ para essas duas variáveis.

Na Tabela 3 encontram-se as medidas descritivas para o peso e o número de frutos sem e com o acréscimo da constante, antes e após a transformação.

Tabela 3 – Medidas Descritivas para as variáveis original, variável mais constante e transformação da variável mais constante para o número e fitomassa de frutos de abobrinha italiana cultivada nas estações sazonais: verão/outono (a) e inverno/primavera (b). Santa Maria – RS. 2008.

	Túnel	Época 1			Época 2		
		\bar{X}	S	CV(%)	\bar{X}	S	CV(%)
Fitomassa (Kg)	1	0,09	0,15	16,67	0,19	0,23	121,05
	2	0,09	0,15	16,67	0,19	0,24	126,32
Transformada	1	0,87	0,05	5,75	0,91	0,07	7,69
	2	0,87	0,05	5,75	0,90	0,07	7,78
Nº Frutos (unidade)	1	0,53	0,79	149,06	0,81	0,82	101,23
	2	0,53	0,82	154,72	0,75	0,81	108,00
Transformado	1	0,96	0,16	16,67	1,03	0,17	16,50
	2	0,96	0,17	17,71	1,02	0,17	16,67

\bar{X} = média; S = desvio padrão, CV(%) = coeficiente de variação.

Observa-se que o acréscimo da constante reduz o coeficiente de variação devido ao aumento da média, logo não se pode atribuir uma melhora na variação. Ao se aplicar à transformação na variável acrescida da constante observa-se uma boa redução no coeficiente de variação o que pode indicar uma redução da variabilidade dos dados.

Na Tabela 4 tem-se o quadro de análise de variância para o número e o peso frutos para as duas épocas. Observa-se que há uma redução no quadrado médio do erro (QME) após o uso da transformação, tanto para o número quanto para o peso de frutos nas duas épocas. Este resultado é importante, pois quanto maior o erro, menor a probabilidade de se encontrarem diferenças significativas entre médias de tratamentos. Portanto, a transformação raiz quarta ($y^{0,25}$) para o número e o peso de frutos de abobrinha italiana permite uma redução de variabilidade.

Tabela 4 – Resumo da análise de variância para a variável original, variável acrescida da constante e transformação da variável acrescida da constante para o número de frutos de abobrinha italiana, cultivada nas estações sazonais verão/outono (época 1) e inverno/primavera (época 2). Santa Maria – RS. 2008.

Época 1						
Fonte de Variação	Túnel 1			Túnel 2		
	GL	Variável Original e Original +0,5 Quadrado Médio	Transformação (Original +0,5) Quadrado Médio	GL	Variável Original e Original +0,5 Quadrado Médio	Transformação (Original +0,5) Quadrado Médio
B	9	0,01*	0,002*	9	0,03*	0,008*
A	11	0,14*	0,02*	11	0,14*	519,10*
A*B	99	0,015	0,002	99	0,02	0,008
D	3	0,05	0,005	3	0,004	0,008
A*D	27	0,05	0,006	27	0,03	0,008
D*B	33	0,02	0,002	33	0,02	0,008
Total	182			182		

Época 2						
Fonte de Variação	Túnel 1			Túnel 2		
	GL	Variável Original e Original +0,5 Quadrado Médio	Transformação (Original +0,5) Quadrado Médio	GL	Variável Original e Original +0,5 Quadrado Médio	Transformação (Original +0,5) Quadrado Médio
B	13	4,73*	0,16*	13	4,74*	0,17*
A	29	9,76*	0,77*	29	9,80*	0,77*
A*B	116	0,43	0,03	116	0,51	0,03
D	7	0,61	0,04	7	0,66	0,04
A*D	63	0,89	0,05	63	0,92	0,05
D*B	195	0,51	0,04	195	0,51	0,04
Total	423			423		

* significativo a 5% de probabilidade de erro.

Tabela 6 – Resumo da análise de variância para a variável original, variável acrescida da constante e transformação da variável acrescida da constante para a fitomassa de frutos de abobrinha italiana, cultivada nas estações sazonais verão/outono (época 1) e inverno/primavera (época 2). Santa Maria – RS. 2008.

Época 1						
Túnel 1				Túnel 2		
Fonte de Variação	GL	Variável Original e Original +0,5 Quadrado Médio	Transformação (Original +0,5) Quadrado Médio	GL	Variável Original e Original +0,5 Quadrado Médio	Transformação (Original +0,5) Quadrado Médio
B	9	0,46*	0,02*	9	1,38*	0,08*
A	11	4,81*	0,20*	11	5,09*	0,24*
A*B	99	0,43	0,02	99	0,59	0,05
D	3	1,00	0,04	3	0,19	0,03
A*D	27	1,44	0,07	27	0,77	0,07
D*B	33	0,53	0,02	33	0,39	0,03
Total	182			182		

Época 2						
Túnel 1				Túnel 2		
Fonte de Variação	GL	Variável Original e Original +0,5 Quadrado Médio	Transformação (Original +0,5) Quadrado Médio	GL	Variável Original e Original +0,5 Quadrado Médio	Transformação (Original +0,5) Quadrado Médio
B	13	0,14*	0,008*	13	0,11*	0,01*
A	29	0,51*	0,02*	29	1,78*	0,14*
A*B	116	0,06	0,004	116	0,04	0,003
D	7	0,03	0,003	7	0,08	0,008
A*D	63	0,08	0,007	63	0,08	0,008
D*B	195	0,04	0,003	195	0,04	0,003
Total	423			423		

* significativo a 5% de probabilidade de erro.

Tabela 7 – Resultados do teste de Bartlett para homogeneidade de variâncias, para variável acrescida da constante e transformação da variável acrescida da constante para o número e fitomassa de frutos de abobrinha italiana, cultivada nas estações sazonais verão/outono (época 1) e inverno/primavera (época 2). Santa Maria – RS. 2008.

		Época 1			Época 2		
	Túnel	GL	Variável Original e Original +0,5 Transformação (Original +0,5)	GL	Variável Original e Original +0,5 Transformação (Original +0,5)	GL	Variável Original e Original +0,5 Transformação (Original +0,5)
Número de frutos	1	11	HET	HET	29	HET	HET
	2	11	HET	HET	29	HET	HET
Fitomassa	1	11	HET	HET	29	HET	HET
	2	11	HET	HET	29	HET	HET

HET: variâncias heterogêneas a 5% de significância.

Deste modo, a investigação quanto à homogeneidade de variâncias se faz necessária para verificar se o emprego da transformação raiz quarta ($y^{0,25}$) foi suficiente para adequar os dados. Na Tabela 5 têm-se os resultados do teste de Bartlett, mostrando que a transformação raiz quarta ($y^{0,25}$) não foi suficiente para adequar os dados quanto esta pressuposição, pois em nível de significância 0,05, rejeitou-se a hipótese de homocedasticidade, para o número e o peso de frutos nas duas épocas.

Uma análise residual também foi realizada a fim de verificar o comportamento dos erros após o uso da transformação raiz quarta ($y^{0,25}$). Nas Figuras 4 e 5, têm-se os gráficos dos resíduos versus o valor normal esperado, para as variáveis número e peso de frutos, respectivamente. Pode-se verificar que os resíduos apresentam uma menor variação após a transformação, 5.1 e 5.2 (c). Apesar disto, essa redução não foi aparentemente suficiente para tornar os dados homocedásticos.

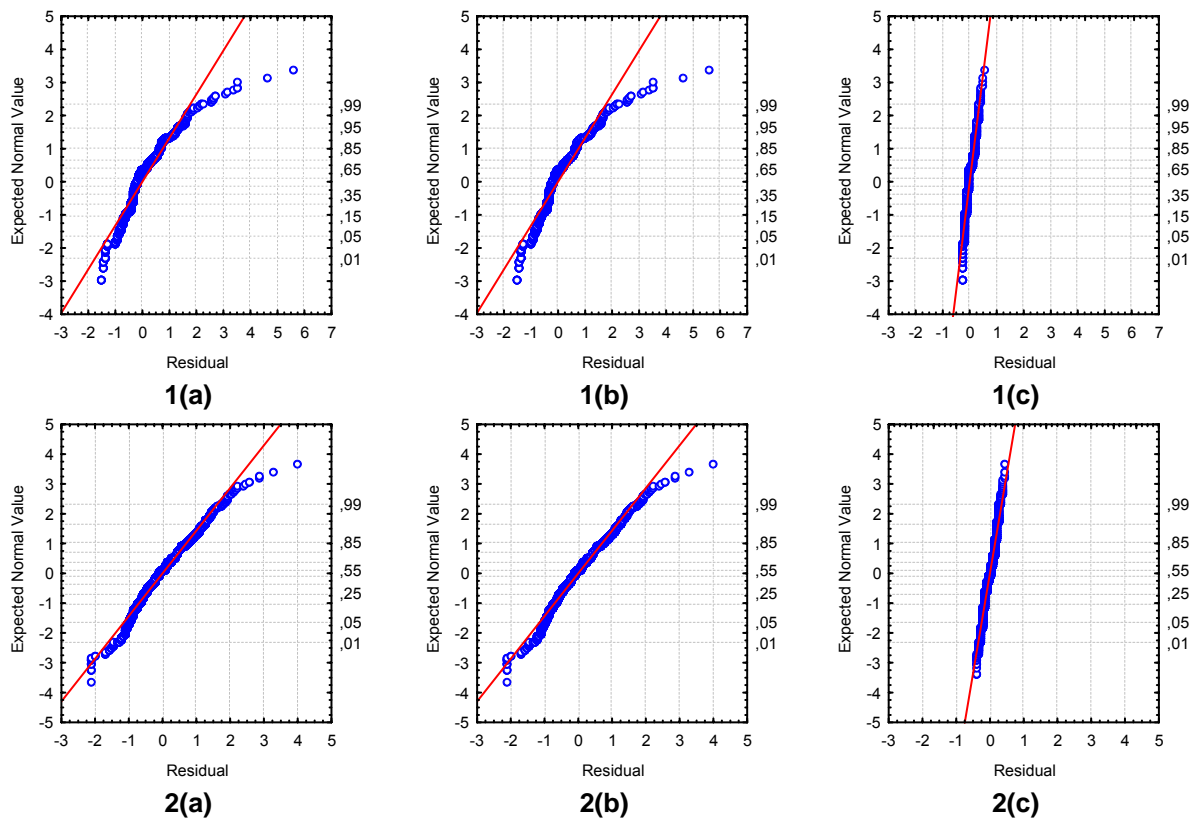


Figura 4 – Resíduos em função do valor normal esperado para o experimento com abobrinha italiana. Figuras 4.1(a), (b) e (c) para o número de frutos, número de frutos mais constante e transformação da variável número de frutos mais constante, respectivamente; para a estação verão/outono. Figuras 4.2(a), (b) e (c) para o número de frutos, número de frutos mais constante e transformação da variável número de frutos mais constante, respectivamente; para a estação inverno/primavera. Santa Maria – RS. 2008.

Na Figura 4 pode-se verificar ainda que os resíduos apresentam menor variação após a transformação, 4.1 e 4.2 (c). Apesar disto, essa redução não foi aparentemente suficiente para tornar os dados homocedásticos.

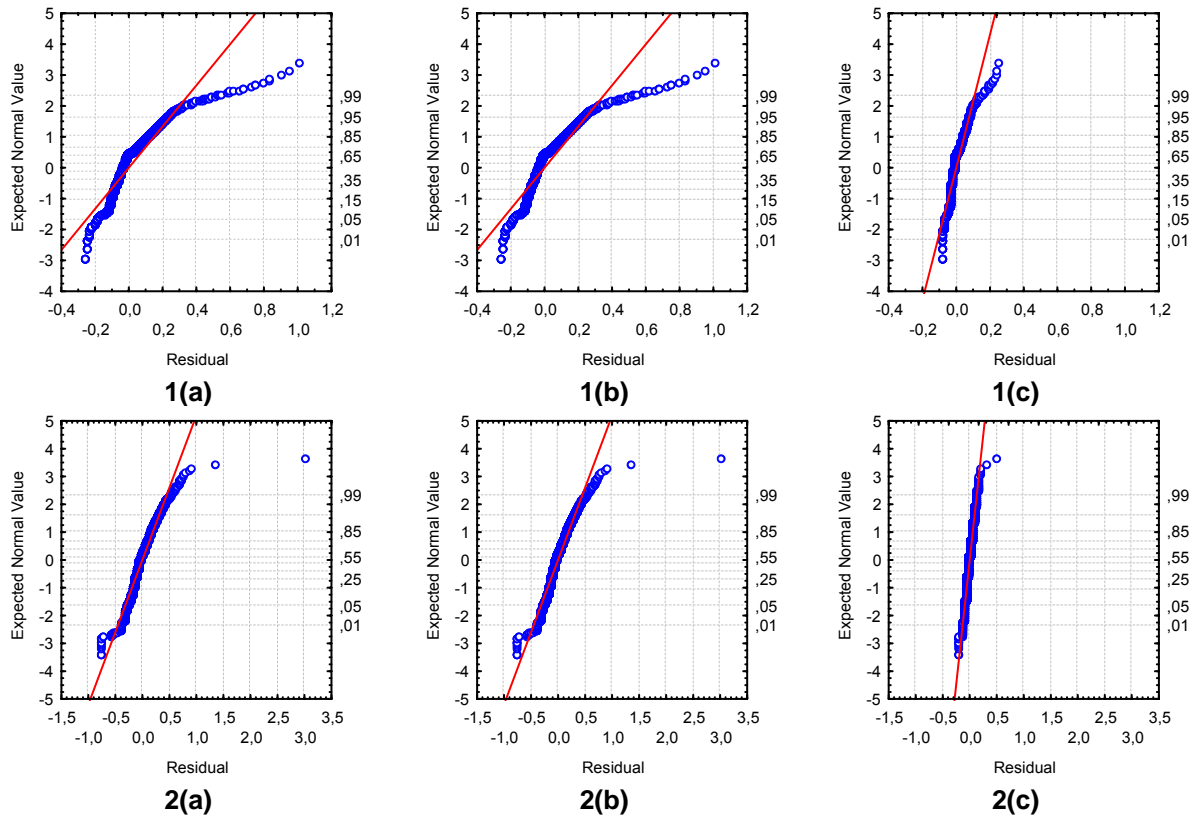


Figura 5 – Resíduos em função do valor normal esperado, para o experimento com abobrinha italiana. Figuras 5.1(a), (b) e (c) para a variável fitomassa de frutos, fitomassa mais constante e transformação da variável fitomassa de frutos mais constante, respectivamente; para a estação verão/outono. Figuras 5.2(a), (b) e (c) para a variável fitomassa de frutos, fitomassa mais constante e transformação da variável fitomassa de frutos mais constante, respectivamente; para a estação inverno/primavera. Santa Maria – RS. 2008.

Na Figura 5 pode-se verificar que os resíduos apresentam menor variação após a transformação, 5.1 e 5.2 (c). Apesar disto, essa redução não foi aparentemente suficiente para tornar os dados homocedásticos.

4.2 Experimento com pimentão

Na Figura 6, observar-se a freqüência de valores zero para o número e a fitomassa de frutos de pimentão. Para essas variáveis a freqüência de valores zero varia em torno de 35%, na época 1; e 40%, na época 2. Observa-se também, que essas duas variáveis estão relacionadas, pois a ausência de frutos afeta o peso de frutos colhidos; resultado também observado no experimento com abobrinha italiana.

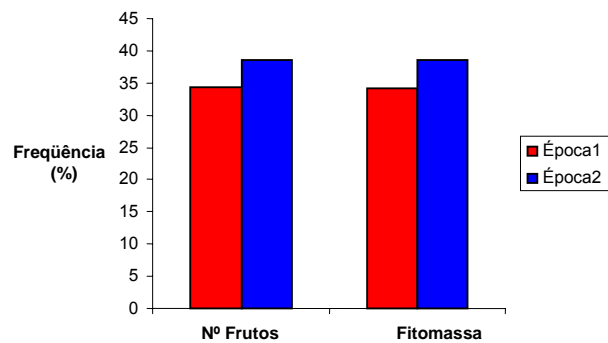


Figura 6 – Freqüência de valores zero para as variáveis número e fitomassa de frutos de pimentão, cultivado nas estações sazonais verão/outono (época 1) e inverno/primavera (época 2). Santa Maria – RS. 2008.

Comparando-se os dois experimentos, abobrinha italiana e pimentão, verifica-se que a freqüência de zeros no experimento com pimentão é menor. Porém, esta também não baixa de 35%, que é um valor ainda elevado para a ausência de frutos. E, considerando-se que, um dos objetivos em olericultura seria maximizar a produção destes.

Nas Tabelas 8 e 9 e nas Figuras 7 e 8, encontram-se os resultados procedimento Box-Cox.

Tabela 8 – Resultados do procedimento Box-Cox na escolha da potência λ , para a fitomassa de frutos de pimentão, cultivado nas estações sazonais verão/outono (época 1) e inverno/primavera (época 2). Santa Maria – RS. 2008.

Box-Cox (λ)	Época 1		Época 2	
	L_{\max}	RQME	L_{\max}	RQME
-2,00	-48140,55	34,369	-38005,58	32,805
-1,75	-43174,94	23,862	-33999,49	22,707
-1,50	-38496,20	16,920	-30223,48	16,052
-1,25	-34177,77	12,320	-26743,68	11,661
-1,00	-30322,84	9,281	-23653,52	8,780
-0,75	-27082,39	7,315	-21087,43	6,936
-0,50	-24678,93	6,131	-19235,47	5,851
-0,25	-23415,53	5,587	-18339,55	5,389
0,00	-23614,54	5,669	-18633,03	5,536
0,25	-25454,84	6,490	-20221,24	6,406
0,50	-28842,83	8,325	-23010,43	8,276
0,75	-33485,94	11,709	-26775,30	11,695
1,00	-39065,93	17,644	-31277,61	17,684
1,25	-45333,06	27,962	-36327,67	28,120
1,50	-52113,24	46,018	-41789,82	46,440
1,75	-59287,23	77,957	-47569,16	78,961
2,00	-66771,40	135,105	-53598,29	137,372

* L_{\max} = função de máxima verossimilhança; RQME = raiz do quadrado médio do erro.

Na Tabela 8 encontram-se para vários valores de λ a raiz do quadrado médio do erro (RQME) e a função de máxima verossimilhança (L_{\max}) para o peso de frutos colhidos para duas épocas. Analisando-se a Tabela 6, o valor da potencia λ que minimiza a RQME é -0,25, o que indica que a melhor transformação para os dados seria o inverso da raiz quarta ($y^{-0,25}$).

Na Figura 7, tem-se o gráfico de λ em função da RQME o qual apresenta o intervalo de confiança de 99% para λ .

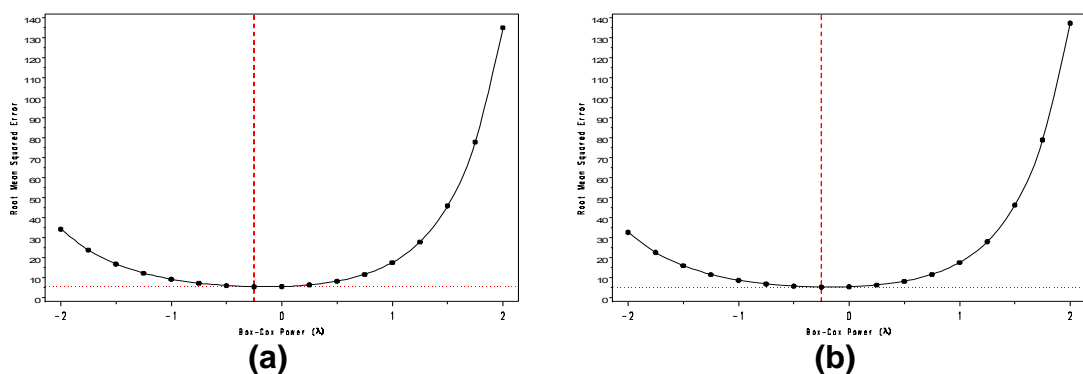


Figura 7 – Potência λ em função da raiz do quadrado médio do erro (RQME) pelo método de Box-Cox, com intervalo de confiança à 99%, para a fitomassa de frutos de pimentão cultivado nas estações sazonais verão/outono (a) e inverno/primavera (b). Santa Maria – RS. 2008.

A inspeção da Figura 8 mostra que o melhor estimador de máxima verossimilhança para λ é -0,25, o que também foi concluído a partir da Tabela 6.

Da mesma forma, a Tabela 7 e Figura 8, mostram os resultados para o melhor estimador de λ para o número de frutos colhidos.

Tabela 9 – Resultados do procedimento Box-Cox na escolha da potência λ , para o variável número de frutos de pimentão, cultivado nas estações sazonais verão/outono (época 1) e inverno/primavera (época 2). Santa Maria – RS. 2008.

Box-Cox (λ)	Época 1		Época 2	
	L_{\max}	RQME	L_{\max}	RQME
-2,00	-54278,43	53,954	-42780,22	50,862
-1,75	-48503,75	35,298	-38252,03	33,556
-1,50	-42995,88	23,550	-33932,69	22,568
-1,25	-37844,29	16,129	-29890,94	15,569
-1,00	-33184,60	11,453	-26231,36	11,125
-0,75	-29223,75	8,561	-23114,93	8,356
-0,50	-26260,76	6,886	-20777,93	6,742
-0,25	-24657,26	6,121	-19513,19	6,002
0,00	-24698,27	6,139	-19561,84	6,029
0,25	-26413,40	6,964	-20963,08	6,857
0,50	-29570,18	8,782	-23537,57	8,687
0,75	-33842,78	12,020	-27022,21	11,963
1,00	-38948,49	17,492	-31185,63	17,535
1,25	-44680,83	26,654	-35858,26	26,934
1,50	-50894,86	42,078	-40921,06	42,878
1,75	-57487,26	68,299	-46289,44	70,205
2,00	-64382,44	113,355	-51901,69	117,551

* L_{\max} = função de máxima verossimilhança; RQME = raiz do quadrado médio do erro.

Na Figura 8, tem-se o gráfico de λ em função da RQME para o número de frutos colhidos, o qual apresenta o intervalo de confiança de 99% para λ .

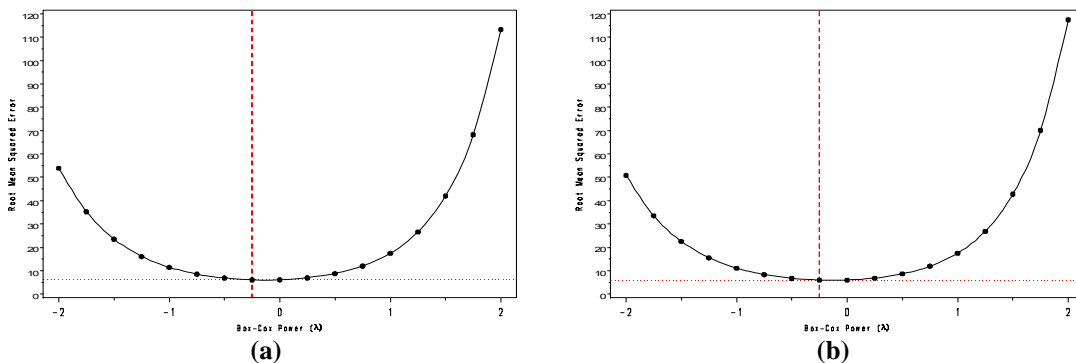


Figura 8 – Potência λ em função da raiz do quadrado médio do erro (RQME) pelo método de Box-Cox, com intervalo de confiança à 99%, para o número de frutos de pimentão cultivado nas estações sazonais verão/outono (a) e inverno/primavera (b). Santa Maria – RS. 2008.

Analisando-se a Tabela 9 e Figura 8, tem-se que o melhor estimador de máxima verossimilhança para λ é -0,25.

Aqui também se observa a relação entre peso e número de frutos colhidos, o qual pode ter sido um dos motivos para o qual se tem o mesmo valor λ para essas duas variáveis. Fato também observado para as mesmas variáveis no experimento com abobrinha italiana.

Após a determinação de λ , realizou-se uma análise descritiva das variáveis, antes e após a transformação, que segue na Tabela 9.

Tabela 9 – Medidas Descritivas para as variáveis original, variável mais constante e transformação da variável mais constante para o número e fitomassa de frutos de pimentão cultivado nas estações sazonais verão/outono (a) e inverno/primavera (b). Santa Maria – RS. 2008.

	Época 1			Época 2		
	\bar{X}	S	CV(%)	\bar{X}	S	CV(%)
Fitomassa (Kg)	0,22	0,23	106,75	0,31	0,34	109,00
Transformada	1,10	0,08	6,95	1,08	0,10	9,49
Nº Frutos (unidade)	1,31	1,31	99,95	1,43	1,47	102,68
Transformado	0,94	0,19	19,96	0,94	0,20	21,57

\bar{X} = média, S = desvio padrão, CV = coeficiente de variação.

Na Tabela 9 encontram-se as medidas descritivas para o peso e o número de frutos sem e com o acréscimo da constante, antes e após a transformação. Observa-se que com o acréscimo da constante há uma redução no coeficiente de variação devido ao aumento na média, logo não se pode atribuir uma melhora na variação. Ao se aplicar à transformação na variável acrescida da constante observa-se uma boa redução no coeficiente de variação (CV) o que pode indicar uma redução da variabilidade dos dados.

Fazendo um comparativo entre os dois experimentos, Tabelas 3 e 8, observa-se que o coeficiente de variação para a variável peso de frutos, após transformação, mantém-se na faixa de 6 a 9%. E que para a variável, número de frutos, após a transformação, o coeficiente de variação fica na faixa de 16 a 20%.

Ainda comparando-se os dois experimentos, tem-se que o número médio de frutos colhidos para o experimento com pimentão é superior ao número de frutos colhidos no experimento com abobrinha.

Na Tabela 10 tem-se o quadro de análise de variância para o número e o peso frutos para as duas épocas.

Tabela 10 – Resumo da análise de variância para a variável original, variável acrescida da constante e transformação da variável acrescida da constante para o número e fitomassa de frutos de pimentão, cultivado nas estações sazonais verão/outono (época 1) e inverno/primavera (época 2). Santa Maria – RS. 2008.

Número de Frutos						
Fonte de Variação	Época 1			Época 2		
	GL	Variável Original e Original +0,5	Transformação (Original +0,5)	GL	Variável Original e Original +0,5	Transformação (Original +0,5)
		Quadrado Médio	Quadrado Médio		Quadrado Médio	Quadrado Médio
B	34	0,05	0,02	34	0,06	0,007
A	4	0,74*	0,26*	3	4,94*	0,44*
A*B	136	0,05	0,03	102	0,10	0,01
D	9	0,12	0,10	9	0,14	0,008
D*B	306	0,04	0,03	306	0,08	0,007
A*D	36	0,18*	0,09	27	0,84*	0,07

Fitomassa de Frutos						
Fonte de Variação	Época 1			Época 2		
	GL	Variável Original e Original +0,5	Transformação (Original +0,5)	GL	Variável Original e Original +0,5	Transformação (Original +0,5)
		Quadrado Médio	Quadrado Médio		Quadrado Médio	Quadrado Médio
B	34	1,15*	0,03	34	1,68*	0,04
A	4	24,55*	869,05*	3	71,16*	1,19*
A*B	136	1,62	0,03	102	1,98	0,04
D	9	8,15	0,05	9	2,02	0,03
D*B	306	1,27	0,03	306	1,48	0,03
A*D	36	8,05*	0,04	27	14,13*	0,26

* significativo a 5% de probabilidade de erro.

Observa-se na Tabela 10 que há uma considerável redução no quadrado médio do erro (QME) após o uso da transformação, tanto para o número quanto para o peso de frutos nas duas épocas. Este resultado é importante, pois quanto maior o erro, menor a probabilidade de se encontrarem diferenças significativas entre médias de tratamentos. Portanto, o inverso da raiz quarta ($y^{-0,25}$) para o número e o peso de frutos de abobrinha italiana permite uma redução de variabilidade.

Deste modo, a investigação quanto à homogeneidade de variâncias se faz necessária para verificar se o emprego desta transformação ($y^{-0,25}$) foi suficiente para adequar os dados. Na Tabela 11, a seguir, têm-se os resultados do teste de Bartlett.

O teste Bartlett para homogeneidade de variâncias, mostra que a transformação ($y^{-0,25}$) não foi suficiente para adequar os dados quanto esta pressuposição, pois em nível de significância 0,05, rejeitou-se a hipótese de homocedasticidade, para o número e o peso de frutos nas duas épocas.

Tabela 11 – Resultados do teste de Bartlett para homogeneidade de variâncias, para variável acrescida da constante e transformação da variável acrescida da constante para o número e fitomassa de frutos de pimentão, cultivado nas estações sazonais verão/outono (época 1) e inverno/primavera (época 2). Santa Maria – RS. 2008.

	Época 1			Época 2		
	GL	Variável Original e Original +0,5	Transformação (Original +0,5)	GL	Variável Original e Original +0,5	Transformação (Original +0,5)
Número de frutos	4	HET	HET	3	HET	HET
Fitomassa	4	HET	HET	3	HET	HET

HET: variâncias heterogêneas a 5% de significância.

Uma análise residual também foi realizada a fim de verificar o comportamento dos erros após o uso da transformação. Nas Figuras 11 e 12, têm-se os gráficos dos resíduos versus o valor normal esperado, para as variáveis, número e peso de frutos, respectivamente.

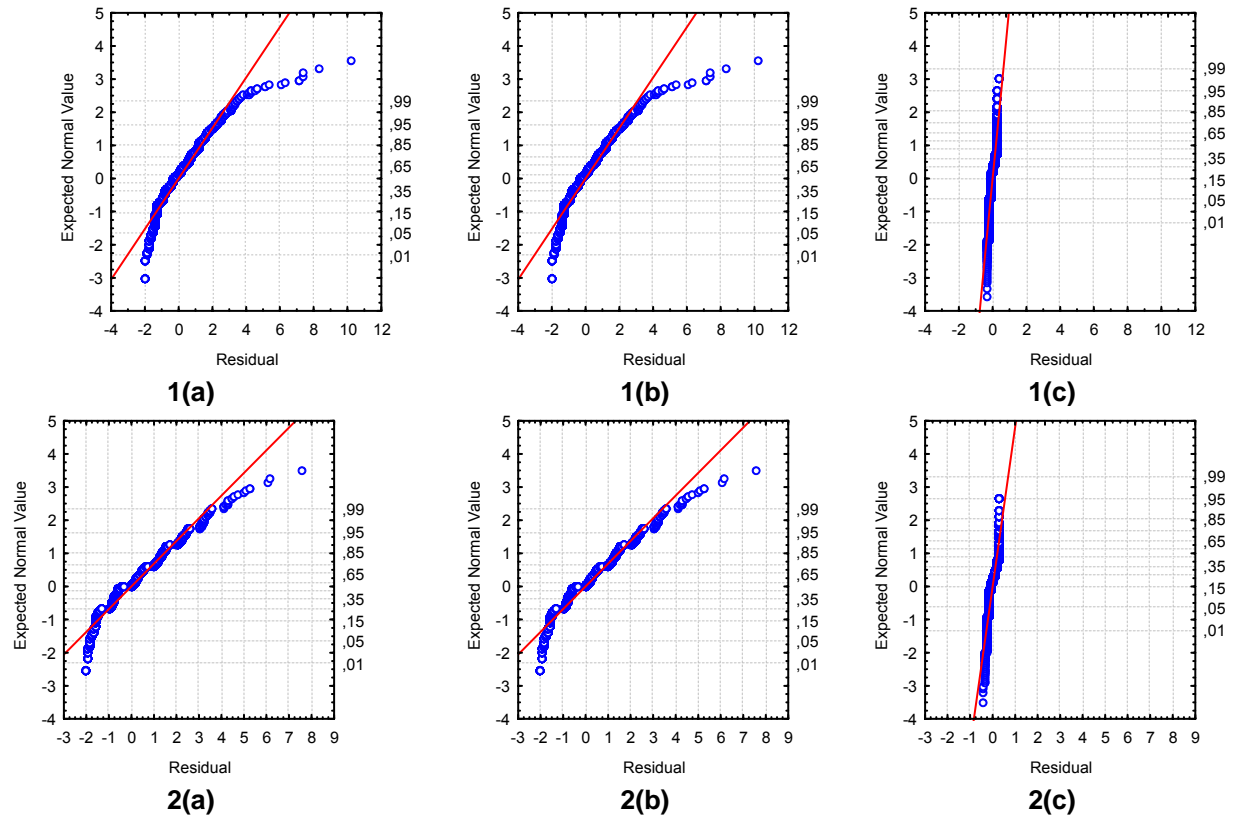


Figura 9 – Resíduos em função do valor normal esperado, para o experimento com pimentão. Figuras 9.1(a), (b) e (c) para o número de frutos colhidos, número de frutos mais constante e transformação da variável número de frutos mais constante, respectivamente; para a estação verão/outono. Figuras 9.2(a), (b) e (c) para o número de frutos colhidos, número de frutos mais constante e transformação da variável número de frutos mais constante, respectivamente; para a estação inverno/primavera. Santa Maria – RS. 2008.

Na Figura 9 pode-se verificar que os resíduos apresentam uma menor variação após a transformação, 9.1 e 9.2 (c). Apesar disto, essa redução não foi aparentemente suficiente para tornar os dados homocedásticos.

A seguir tem-se a Figura 10, com os gráficos dos resíduos para a variável, peso de frutos colhidos de pimentão.

Na Figura 10 pode-se verificar que os resíduos apresentam uma menor variação após a transformação, 10.1 e 10.2 (c). Apesar disto, essa redução não foi aparentemente suficiente para tornar os dados homocedásticos.

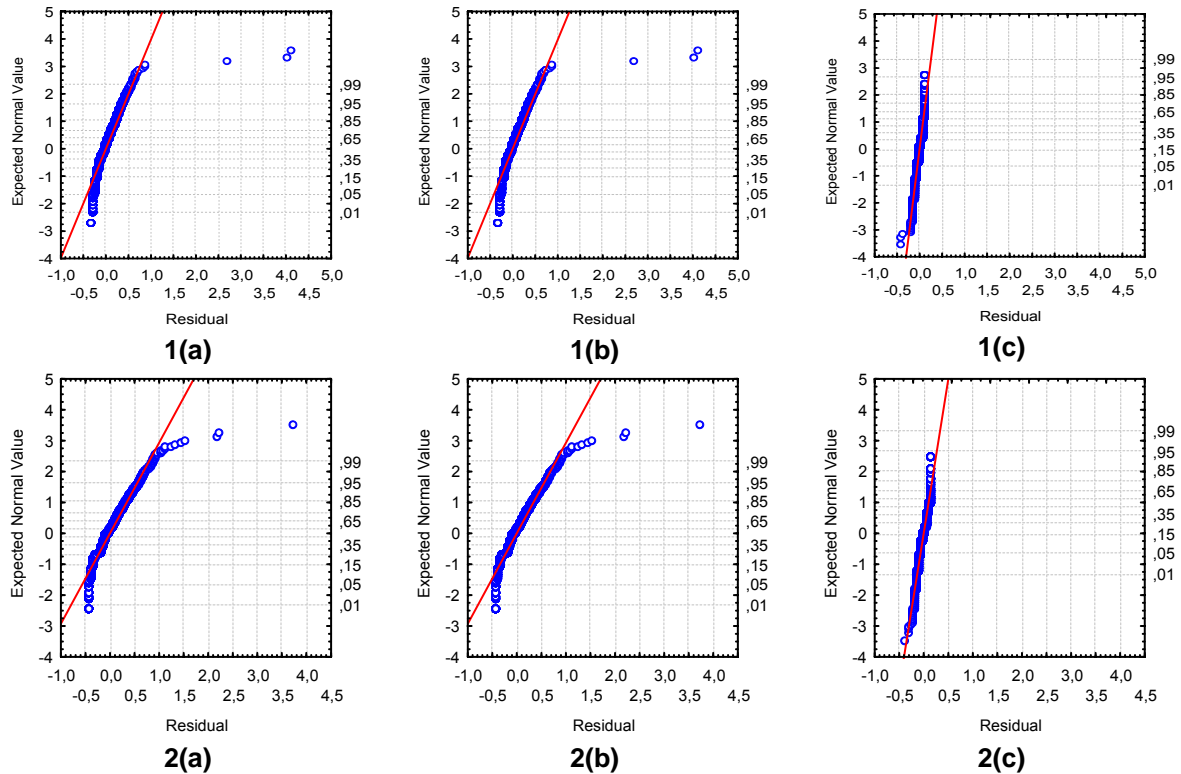


Figura 10 – Resíduos em função do valor normal esperado, para o experimento com pimentão. Figuras 10.1(a), (b) e (c) para a variável fitomassa de frutos, fitomassa mais constante e transformação da variável fitomassa de frutos mais constante, respectivamente; para a estação verão/outono. Figuras 10.2(a), (b) e (c) para a variável fitomassa de frutos, fitomassa mais constante e transformação da variável fitomassa de frutos mais constante, respectivamente; para a estação inverno/primavera. Santa Maria – RS. 2008.

4.3 Experimento com brócolis

Na Figura 11, pode-se observar a freqüência de valores zero para rendimento de cabeças comerciais e não-comerciais e, para ramificações comerciais e não-comerciais nas quatro épocas. Tem-se que o rendimento de cabeças e o número ramificações comerciais – (CC) e (RC), respectivamente - apresentaram percentuais de presença de zeros inferiores aos percentuais para essas variáveis não-comerciais – (CNC) e (RNC), respectivamente. Isto indica que apesar de valores nulos, o rendimento de cabeças e ramificações comerciais colhidas foi superior ao número de cabeças e ramificações não-comerciais.

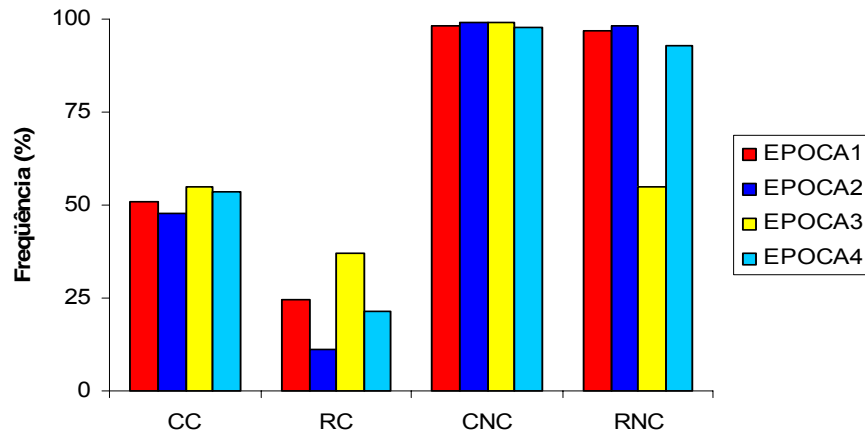


Figura 11 – Frequência de valores zero para as variáveis Cabeças Comerciais (CC), Ramificações Comerciais (RC), Cabeças Não Comerciais (CNC) e Ramificações Não Comerciais (RNC) para a cultura de brócolis, em quatro épocas. Santa Maria – RS. 2008.

Observa-se que a frequência de zeros para cabeças e ramificações não comerciais é muito alta, acima de 90%, exceto para ramificações não comerciais na época 3, que fica em torno de 60%. Este fato deve-se que, para a mesma época, a frequência de zeros para ramificações comerciais aumentou, logo gerou uma redução na frequência de zeros nas ramificações não comerciais.

Os resultados deste procedimento, para cabeças comerciais de brócolis, estão apresentados, na Tabela 11, onde se encontram para vários valores de λ a raiz do quadrado médio do erro (RQME) e a função de máxima verossimilhança (L_{\max}) para cabeças comerciais (CC) para as quatro épocas. Analisando-se a tabela, o valor da potência λ que minimiza a RQME é 0 (zero), para as épocas 1, 2 e 3, o que indica que a melhor transformação para os dados seria a função logarítmica ($y^0 = \log(y)$). Já para a época 4, o valor de λ que minimiza a RQME é -0,25, indicando que a transformação mais adequada é o inverso da raiz quarta ($y^{-0,25}$).

Tabela 12 – Resultados do procedimento Box-Cox na escolha da potência λ , para a variável Cabeças Comerciais (CC) de brócolis, cultivado em quatro épocas. Santa Maria – RS. 2008.

Box-Cox (λ)	Época 1		Época 2		Época 3		Época 4	
	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

* L_{\max} = função de máxima verossimilhança; RQME = raiz do quadrado médio do erro.

Na Figura 12, tem-se o gráfico de λ em função da RQME o qual apresenta o intervalo de confiança de 99% para λ , para cabeças comerciais (CC) nas quatro épocas.

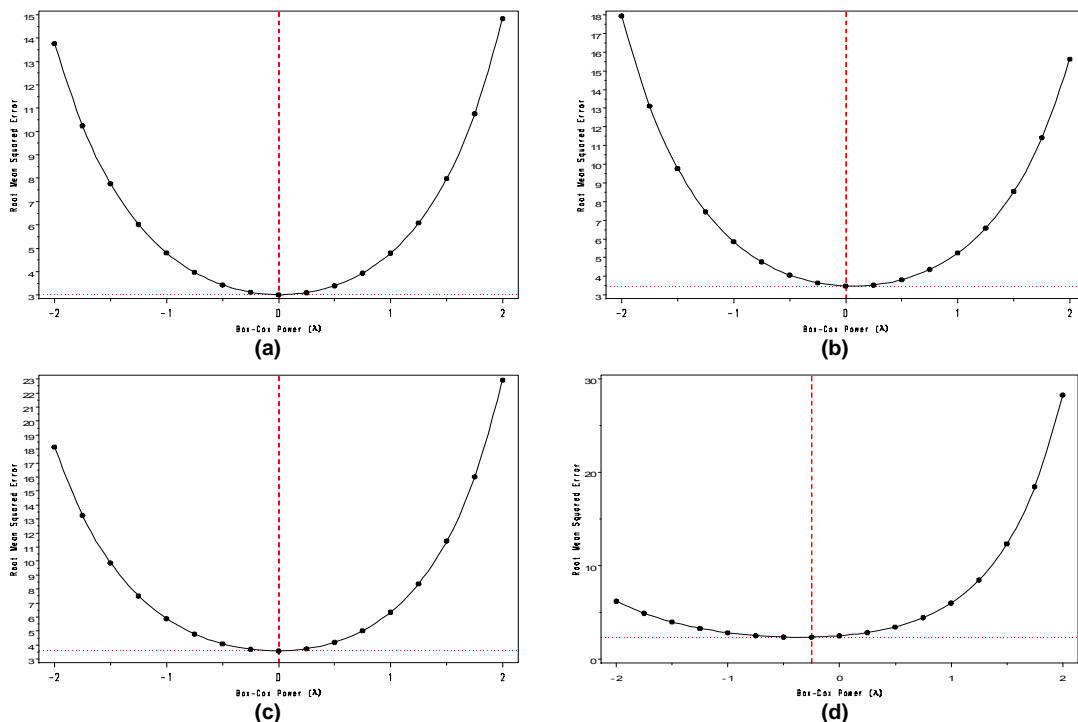


Figura 12 – Potência λ em função da raiz do quadrado médio do erro (RQME) pelo método de Box-Cox, com intervalo de confiança à 99%. Figura (a), (b), (c) e (d), época 1, 2, 3 e 4, respectivamente, para Cabeças Comerciais (CC) de brócolis. Santa Maria – RS. 2008.

A Figura 12 apresenta os gráficos com os intervalos de confiança para λ para cabeças comerciais (CC) de brócolis, nas quatro épocas, e vem confirmar os resultados apresentados e discutidos na Tabela 12.

Na Tabela 13, está os resultados do procedimento Box-Cox aplicado a variável, ramificações comerciais (RC) de brócolis.

Tabela 13 – Resultados do procedimento Box-Cox na escolha da potência λ , para a variável Ramificações Comerciais (RC) de brócolis, cultivado em quatro épocas. Santa Maria – RS. 2008.

Box-Cox (λ)	Época 1		Época 2		Época 3		Época 4	
	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

* L_{\max} = função de máxima verossimilhança; RQME = raiz do quadrado médio do erro.

Na Tabela 13 encontram-se para vários valores de λ a raiz do quadrado médio do erro (RQME) e a função de máxima verossimilhança (L_{\max}) para ramificações comerciais (RC) para as quatro épocas. Analisando-se a Tabela 12, o valor da potência λ que minimiza a RQME é 0 (zero), para as épocas 1, 2 e 3, o que indica que a melhor transformação para os dados seria a função logarítmica ($y^0 = \log(y)$). Já para a época 4, o valor de λ que minimiza a RQME é -0,50, indicando que a transformação mais adequada é o inverso da raiz quadrada ($y^{-0,5}$).

Na Figura 13, tem-se o gráfico de λ em função da RQME para ramificações comerciais (RC) de brócolis, nas quatro épocas, o qual apresenta o intervalo de confiança de 99% para λ .

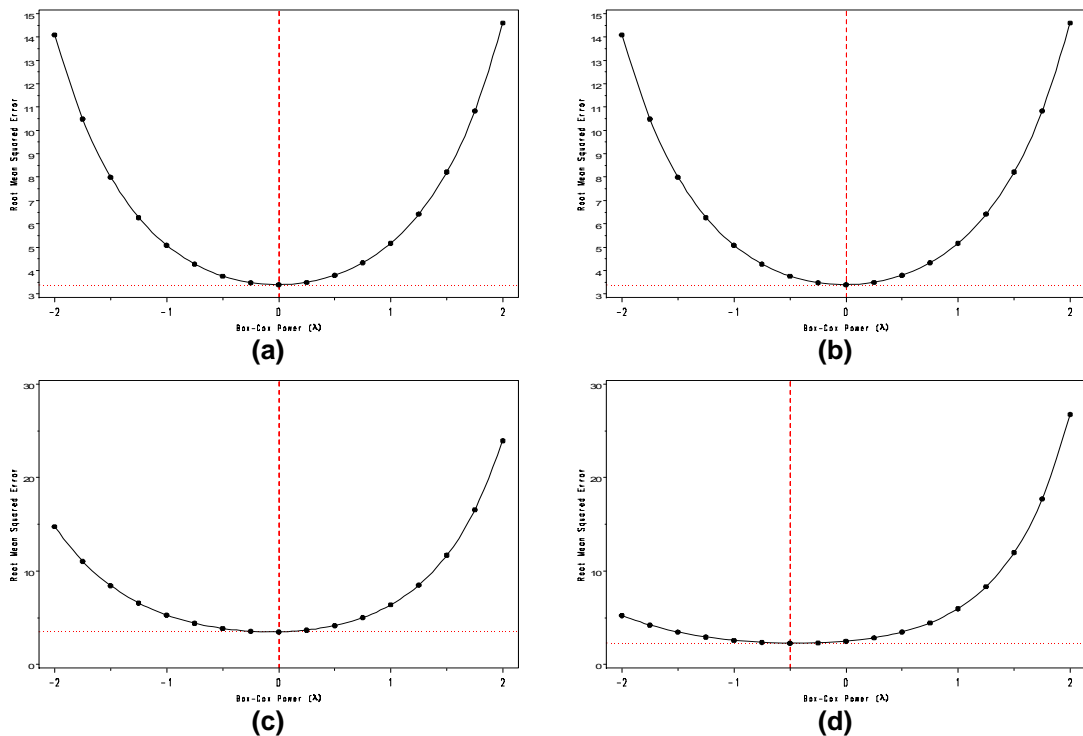


Figura 13 – Potência λ em função da raiz do quadrado médio do erro (RQME) pelo método de Box-Cox, com intervalo de confiança à 99%. Figura (a), (b), (c) e (d), época 1, 2, 3 e 4, respectivamente, para Ramificações Comerciais (RC) de brócolis. Santa Maria – RS. 2008.

Analisando-se a Tabela 13 e a Figura 13 tem-se que o melhor estimador de máxima verossimilhança para λ é 0, para as épocas 1, 2 e 3. E λ é -0,5, para época 4. Observa-se aqui, que da mesma forma que para cabeças comerciais as ramificações comerciais, na época 4, indicaram uma transformação distinta às demais épocas.

Do mesmo modo, aplicou-se o procedimento Box-Cox para cabeças não comerciais (CNC) de brócolis, para as quatro épocas. Os resultados deste procedimento estão apresentados a seguir, na Tabela 13.

Na Tabela 14 encontram-se para vários valores de λ a raiz do quadrado médio do erro (RQME) e a função de máxima verossimilhança (L_{\max}) para cabeças não comerciais (CNC) para as quatro épocas. Analisando-se a tabela, o valor da potencia λ que minimiza a RQME é 0 (zero), para as épocas 1, 2 e 3, o que indica que a melhor transformação para os dados seria a função logarítmica ($y^0 = \log(y)$). Já para a época 4, o valor de λ que minimiza a RQME é -0,25, indicando que a transformação mais adequada é o inverso da raiz quarta ($y^{-0,25}$).

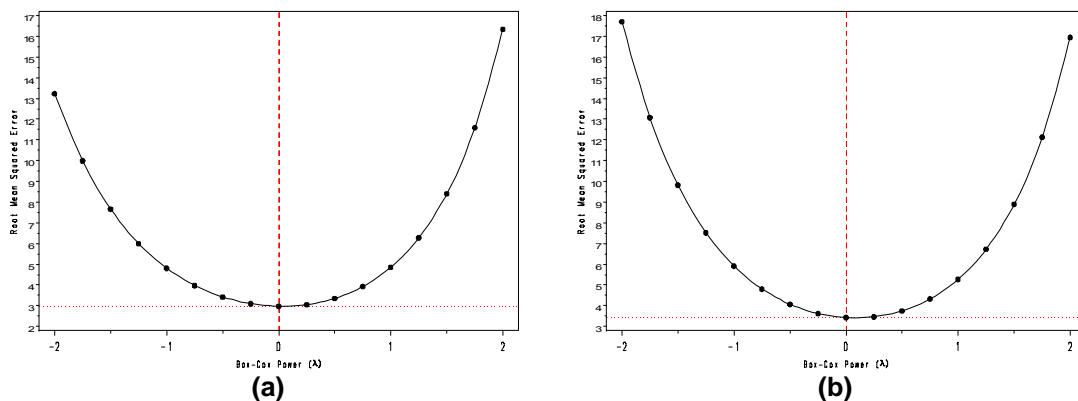
Tabela 14 – Resultados do procedimento Box-Cox na escolha da potência λ , para a variável Cabeças Não Comerciais (CNC) de brócolis, cultivado em quatro épocas. Santa Maria – RS. 2008.

Box-Cox (λ)	Época 1		Época 2		Época 3		Época 4	
	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

* L_{\max} = função de máxima verossimilhança; RQME = raiz do quadrado médio do erro.

Observando-se a Tabela 11, verifica-se que os resultados de λ , para cabeças comerciais, coincidem com os valores de λ encontrados para cabeças não comerciais de brócolis (Tabela 14).

Na Figura 14, tem-se o gráfico de λ em função da RQME para cabeças não comerciais (CNC) de brócolis, nas quatro épocas, o qual apresenta o intervalo de confiança de 99% para λ .



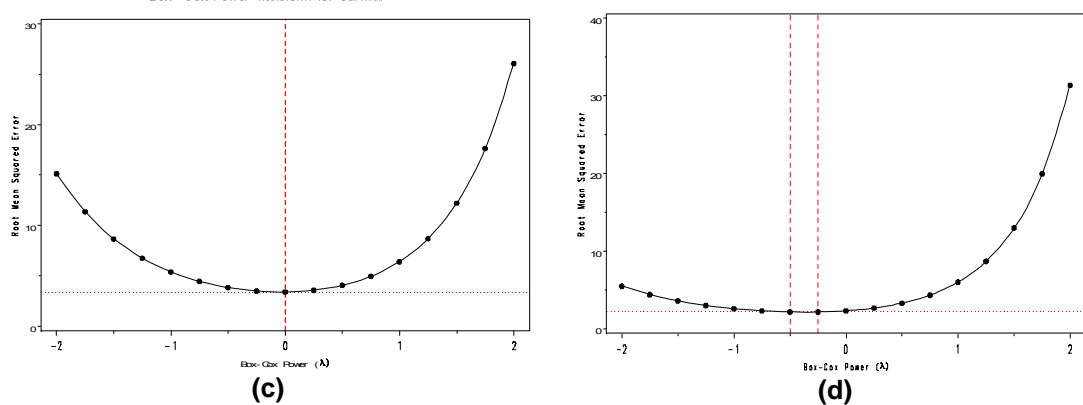


Figura 14 – Potência λ em função da raiz do quadrado médio do erro (RQME) pelo método de Box-Cox, com intervalo de confiança à 99%. Figura (a), (b), (c) e (d), época 1, 2, 3 e 4, respectivamente, para Cabeças Não-Comerciais (CNC) de brócolis. Santa Maria – RS. 2008.

Finalizando, aplicou-se, o procedimento Box-Cox, também às ramificações não comerciais (RNC) de brócolis, nas quatro épocas. Analogamente, às cabeças comerciais e não comerciais, espera-se que os resultados obtidos para λ sejam os mesmos obtidos para ramificações comerciais.

Os resultados para o procedimento Box-Cox estão apresentados na Tabela 15, a seguir:

Tabela 15 – Resultados do procedimento Box-Cox na escolha da potência λ , para a variável Ramificações Não Comerciais (RNC) de brócolis, cultivado em quatro épocas. Santa Maria – RS. 2008.

Box-Cox (λ)	Época 1		Época 2		Época 3		Época 4	
	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

* L_{\max} = função de máxima verossimilhança; RQME = raiz do quadrado médio do erro.

Na Tabela 15 encontram-se para vários valores de λ a raiz do quadrado médio do erro (RQME) e a função de máxima verossimilhança (L_{\max}) para ramificações não comerciais (RNC) de brócolis em quatro épocas. Analisando-se a Tabela 15, o valor da potência λ que minimiza a RQME é 0 (zero), para as épocas 1, 2 e 3, o que indica que a melhor transformação para os dados seria a função logarítmica ($y^0 = \log(y)$). Já para a época 4, o valor de λ que minimiza a RQME é -0,25, indicando que a transformação mais adequada é o inverso da raiz quarta ($y^{-0,25}$).

Na Figura 15, tem-se o gráfico de λ em função da RQME para ramificações não comerciais (RNC) de brócolis, nas quatro épocas, o qual apresenta o intervalo de confiança de 99% para λ .

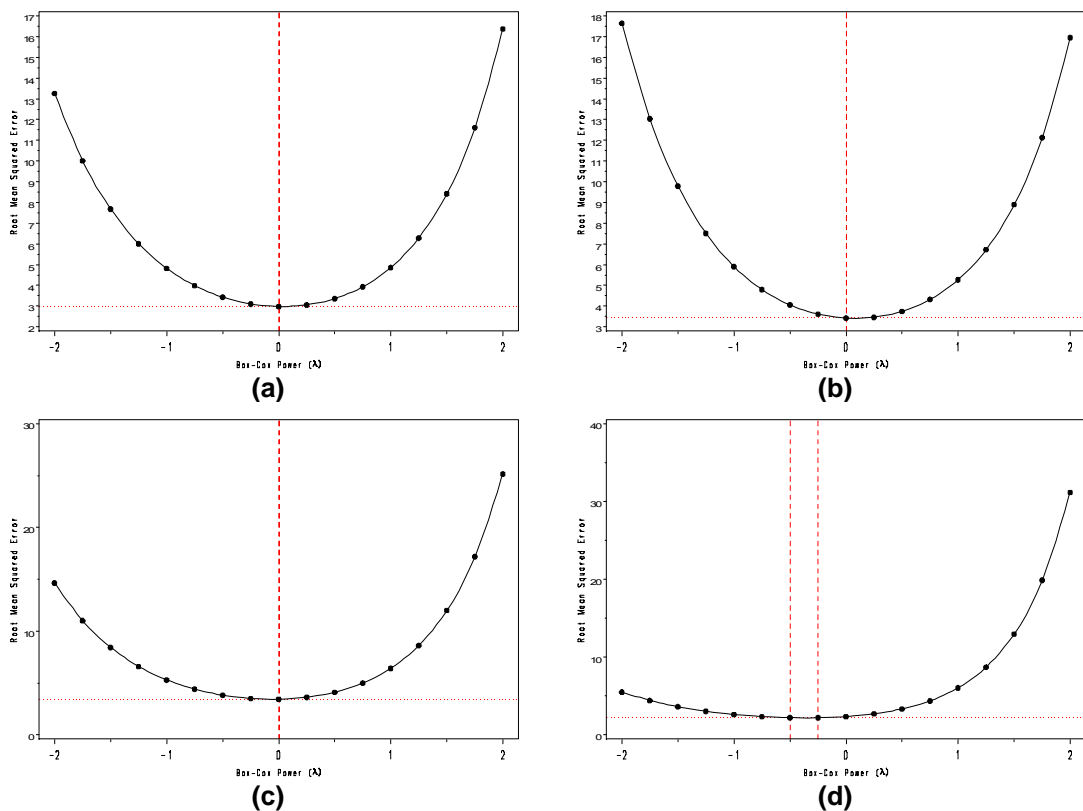


Figura 15 – Potência λ em função da raiz do quadrado médio do erro (RQME) pelo método de Box-Cox, com intervalo de confiança à 99%. Figura (a), (b), (c) e (d), época 1, 2, 3 e 4, respectivamente, para Ramificações Não Comerciais (RNC) de brócolis. Santa Maria – RS. 2008.

Analisando-se a Figura 15, podem-se visualizar os intervalos para λ para ramificações não comerciais (RNC), de brócolis nas quatro épocas. Na Figura 17(d),

observa-se que o intervalo de λ está em $(-0,5 \text{ à } -0,25)$, sendo que λ igual a $-0,25$ é o que apresenta a menor RQME. Desta forma, pode-se observar que as transformações indicadas pelo método para cabeças comerciais e não comerciais não foram às mesmas. Considerou-se λ igual a $-0,25$, para ramificações comerciais e não comerciais na época 4.

Deste modo, pode-se concluir que, tanto cabeças e ramificações comerciais e não comerciais, o método Box-Cox utilizado indicou mesmos valores de λ .

Após a determinação de λ , realizou-se uma análise descritiva das variáveis, antes e após a transformação, que segue na Tabela 15.

Tabela 16 – Medidas descritivas para as variáveis original, variável mais constante e transformação da variável mais constante de Cabeças Comerciais (CC), de Ramificações Comerciais (RC), de Cabeças Não Comerciais (CNC) e de Ramificações Não Comerciais (RNC) para a cultura de brócolis, em quatro épocas.

	ÉPOCA 1			ÉPOCA 2			ÉPOCA 3			ÉPOCA 4		
	\bar{X}	S	CV (%)	\bar{X}	S	CV (%)	\bar{X}	S	CV (%)	\bar{X}	S	CV (%)
CC	0,41	0,75	183,9	0,26	0,34	130,8	0,44	0,76	172,7	0,47	0,89	189,36
Trans(CC+0,5)	-0,27	0,53	196,3	-0,36	0,37	102,8	-0,27	0,57	211,1	1,08	0,14	12,96
RC	0,39	0,39	100,0	0,20	0,27	135,0	0,45	0,36	80,0	0,55	0,59	107,27
Trans(RC+0,5)	-0,21	0,42	200,0	-0,42	0,32	76,16	-0,12	0,37	107,0	1,07	0,02	23,36
CNC	0,45	0,80	177,4	0,20	0,28	138,3	0,00	0,05	1120,9	0,07	0,12	160,9
Trans(CNC+0,5)	-0,11	0,25	223,7	-0,18	0,14	79,5	-0,30	0,03	9,3	0,25	0,08	31,1
RNC	0,52	1,03	199,0	0,56	0,57	102,5	0,01	0,10	1000	0,03	0,17	601,4
Trans(RNC+0,5)	1,08	0,15	14,1	1,06	0,25	23,3	1,18	0,03	2,6	1,40	0,09	6,4

\bar{X} : média; S: desvio padrão; CV: coeficiente de variação.

Na Tabela 16 encontram-se as medidas descritivas para cabeças comerciais e não comerciais e para ramificações comerciais e não comerciais, sem e com o acréscimo da constante, antes e após a transformação. Observa-se que com o acréscimo da constante há uma redução no coeficiente de variação devido ao aumento na média, logo não se pode atribuir uma melhora na variação.

Analisando-se o coeficiente de variação para variável acrescida da constante, com e sem a transformação, não se consegue evidenciar um comportamento padrão

sobre a variabilidade dos dados. Desta forma, recorreu-se a um recurso gráfico para uma melhor visualização da variabilidade desses dados.

Também fica mais evidente o comportamento diferenciado destas variáveis, na época 4. Pois, para todas, foi observado um aumento no valor médio após a transformação. Resultado interessante foi que, embora este comportamento não regular, o procedimento Box-Cox aplicado sugere a transformação logarítmica como sendo a transformação mais adequada para essas variáveis, salvo a época 4.

Ao contrário das variáveis, peso e número de frutos, analisadas nos experimentos: com abobrinha italiana e pimentão, as variáveis cabeças comerciais e não comerciais e ramificações comerciais e não comerciais não apresentam um comportamento padrão o que pode ser observado de acordo com as medidas descritivas extraídas dessas variáveis (Tabela 16).

Este comportamento distinto das demais culturas pode estar associado ao ambiente no qual foi desenvolvido o experimento. Pois nos experimentos com abobrinha italiana e pimentão o cultivo foi em um ambiente controlado, e deste modo estas culturas tiveram uma necessidade diferenciada de adaptação às condições ambientais. Outros fatores como cultivares diferentes com características diferentes, também podem ter influenciado nas respostas.

A seguir, nas Tabelas 17, 18, 19 e 20 encontram-se os resumos das análises de variância para as variáveis cabeças comerciais e cabeças não-comerciais e, ramificações comerciais e não-comerciais.

Tabela 17 – Resumo da análise de variância para a variável original, variável acrescida da constante e transformação da variável acrescida da constante de Cabeças Comerciais (CC) em quatro épocas. Santa Maria – RS. 2008.

Fonte de Variação	Época 1			Época 2		
	GL	Variável Original e Original +0,5 Quadrado Médio	Transformação (Original +0,5) Quadrado Médio	GL	Variável Original e Original +0,5 Quadrado Médio	Transformação (Original +0,5) Quadrado Médio
B	4	2,14	0,50	4	1,32	0,34
A	14	4,03	0,36	10	0,32	0,06
A*B	56	0,62	0,06	40	0,19	0,04
D	3	0,16	0,01	3	0,03	0,004
A*D	42	0,12	0,02	30	0,04	0,009
D*B	12	0,29	0,02	12	0,07	0,013
A*D*B	167	0,34	0,02	117	0,06	0,013
Total						

Fonte de Variação	Época 3			Época 4		
	GL	Variável Original e Original +0,5	Transformação (Original +0,5)	GL	Variável Original e Original +0,5	Transformação (Original +0,5)
		Quadrado Médio	Quadrado Médio		Quadrado Médio	Quadrado Médio
B	4	11,29	1,95	4	6,10	0,26
A	13	0,57	0,05	13	2,32	0,04
A*B	52	0,54	0,04	52	0,82	0,02
D	3	0,39	0,01	2	0,34	0,002
A*D	39	0,35	0,03	26	0,39	0,006
D*B	12	0,31	0,02	3	0,19	0,0004
A*D*B	156	0,41	0,03	39	0,82	0,013
Total						

- significativo a 5% de probabilidade de erro

Tabela 18 – Resumo da análise de variância para a variável original, variável acrescida da constante e transformação da variável acrescida da constante de Ramificações Comerciais (RC) em quatro épocas. Santa Maria – RS. 2008.

Fonte de Variação	Época 1			Época 2		
	GL	Variável Original e Original +0,5	Transformação (Original +0,5)	GL	Variável Original e Original +0,5	Transformação (Original +0,5)
		Quadrado Médio	Quadrado Médio		Quadrado Médio	Quadrado Médio
B	4	4,21	1,10	4	2,27	0,54
A	14	0,30	0,08	10	0,17	0,04
A*B	56	0,10	0,016	40	0,12	0,02
D	3	0,08	0,016	3	0,25	0,04
A*D	42	0,09	0,017	30	0,09	0,02
D*B	12	0,12	0,021	12	0,08	0,01
A*D*B	167	0,08	0,015	117	0,07	0,01
Total						

Fonte de Variação	Época 3			Época 4		
	GL	Variável Original e Original +0,5	Transformação (Original +0,5)	GL	Variável Original e Original +0,5	Transformação (Original +0,5)
		Quadrado Médio	Quadrado Médio		Quadrado Médio	Quadrado Médio
B	4	0,37	0,12	4	1,64	0,56
A	13	0,40	0,11	13	0,38	0,07
A*B	52	0,09	0,02	52	0,34	0,05
D	3	0,07	0,02	2	0,90	0,05
A*D	39	0,04	0,01	26	0,14	0,02
D*B	12	0,04	0,01	3	0,08	0,02
A*D*B	156	0,04	0,01	39	0,28	0,05
Total						

- * significativo a 5% de probabilidade de erro

Tabela 19 – Resumo da análise de variância para a variável original, variável acrescida da constante e transformação da variável acrescida da constante de Cabeças Não Comerciais (CNC) em quatro épocas. Santa Maria – RS. 2008.

Fonte de Variação	Época 1			Época 2		
	GL	Variável Original e Original +0,5	Transformação (Original +0,5)	GL	Variável Original e Original +0,5	Transformação (Original +0,5)
		Quadrado Médio	Quadrado Médio		Quadrado Médio	Quadrado Médio
B	4	0,01	0,87	4	0,007	0,002
A	14	0,01	0,92	10	0,006	0,001
A*B	56	0,01	0,87	40	0,007	0,002
D	3	0,01	0,92	3	0,011	0,003
A*D	42	0,01	1,12	30	0,007	0,002
D*B	12	0,001	0,82	12	0,005	0,001
A*D*B	167	0,01	0,84	117	0,007	0,002
Total				451		

Fonte de Variação	Época 3			Época 4		
	GL	Variável Original e Original +0,5	Transformação (Original +0,5)	GL	Variável Original e Original +0,5	Transformação (Original +0,5)
		Quadrado Médio	Quadrado Médio		Quadrado Médio	Quadrado Médio
B	4	0,001	0,0004	4	0,002	0,0002
A	13	0,002	0,0005	13	0,002	0,0003
A*B	52	0,002	0,0006	52	0,002	0,0003
D	3	0,001	0,0004	2	0,001	0,0001
A*D	39	0,002	0,0006	26	0,001	0,0002
D*B	12	0,002	0,0006	3	0,002	0,0004
A*D*B	156	0,002	0,0006	39	0,002	0,0003
Total	601			583		

* significativo a 5% de probabilidade de erro

Tabela 20 – Resumo da análise de variância para a variável original, variável acrescida da constante e transformação da variável acrescida da constante de Ramificações Não Comerciais (RNC) em quatro épocas. Santa Maria – RS. 2008.

Fonte de Variação	Época 1			Época 2		
	GL	Variável Original e Original +0,5	Transformação (Original +0,5)	GL	Variável Original e Original +0,5	Transformação (Original +0,5)
		Quadrado Médio	Quadrado Médio		Quadrado Médio	Quadrado Médio
B	4	0,002	0,007	4	0,007	0,002
A	14	0,001	0,006	10	0,006	0,001
A*B	56	0,002	0,006	40	0,007	0,002
D	3	0,002	0,006	3	0,011	0,003
A*D	42	0,002	0,007	30	0,007	0,002
D*B	12	0,001	0,005	12	0,005	0,001
A*D*B	167	0,001	0,005	117	0,008	0,002
Total	605			451		

Fonte de Variação	Época 3			Época 4		
	GL	Variável Original e Original +0,5	Transformação (Original +0,5)	GL	Variável Original e Original +0,5	Transformação (Original +0,5)
		Quadrado Médio	Quadrado Médio		Quadrado Médio	Quadrado Médio
B	4	0,001	0,0007	4	0,01	0,005
A	10	0,001	0,0007	13	0,02	0,01
A*B	40	0,001	0,0007	52	0,02	0,008
D	3	0,0005	0,0002	2	0,03	0,02
A*D	30	0,0005	0,0003	26	0,01	0,005
D*B	12	0,0006	0,0003	3	0,01	0,006
A*D*B	117	0,0006	0,0003	39	0,02	0,01
Total	601			583		

* significativo a 5% de probabilidade de erro

Analisando as análises de variância, para cabeças comerciais e não comerciais e para ramificações comerciais e não comerciais, para todas houve uma redução no quadrado médio do erro (QME). Este resultado também observado para as outras duas culturas olerícolas neste estudo.

Seguindo o raciocínio, empregado para as culturas anteriores, realizou-se o teste de Bartlett, para cabeças comerciais e não comerciais e para ramificações comerciais e não comerciais. Que estão apresentados na Tabela 20, a seguir:

Tabela 21 – Resultados do teste de Bartlett para homogeneidade de variâncias, para variável acrescida da constante e transformação da variável acrescida da constante das variáveis cabeças e ramificações comerciais e não comerciais, para a cultura de brócolis cultivado em quatro épocas. Santa Maria – RS. 2008.

Época	GL	Cabeças Comerciais		Cabeças não comerciais		Ramificações comerciais		Ramificações não comerciais	
		VO e VO +0,5	Transf.	VO e VO +0,5	Transf.	VO e VO +0,5	Transf.	VO e VO +0,5	Transf.
		1	14	HET	HET	HET	HET	HET	HET
2	10	HET	HET	HET	HET	HET	HET	HET	HET
3	10	HET	HET	HET	HET	HET	HET	HET	HET
4	13	HET	HET	HET	HET	HET	HET	HET	HET

VO: variável original; VO+0,5: variável original mais constante; Transf.: Transformação da variável original mais constante; HET: variâncias heterogêneas a 5% de significância

O teste de Bartlett homogeneidade de variâncias na Tabela 20, mostram que a transformação logarítmica (y^0) e o inverso da raiz quarta ($y^{-0,25}$) não foram suficientes para adequar os dados quanto esta pressuposição, pois em nível de significância 0,05, rejeitou-se a hipótese de homocedasticidade, para cabeças

comerciais e não comerciais e para ramificações comerciais e não comerciais de brócolis nas quatro épocas.

Por fim, avaliaram-se os gráficos de resíduos para essas variáveis, Figuras 16, 17, 18 e 19, a seguir.

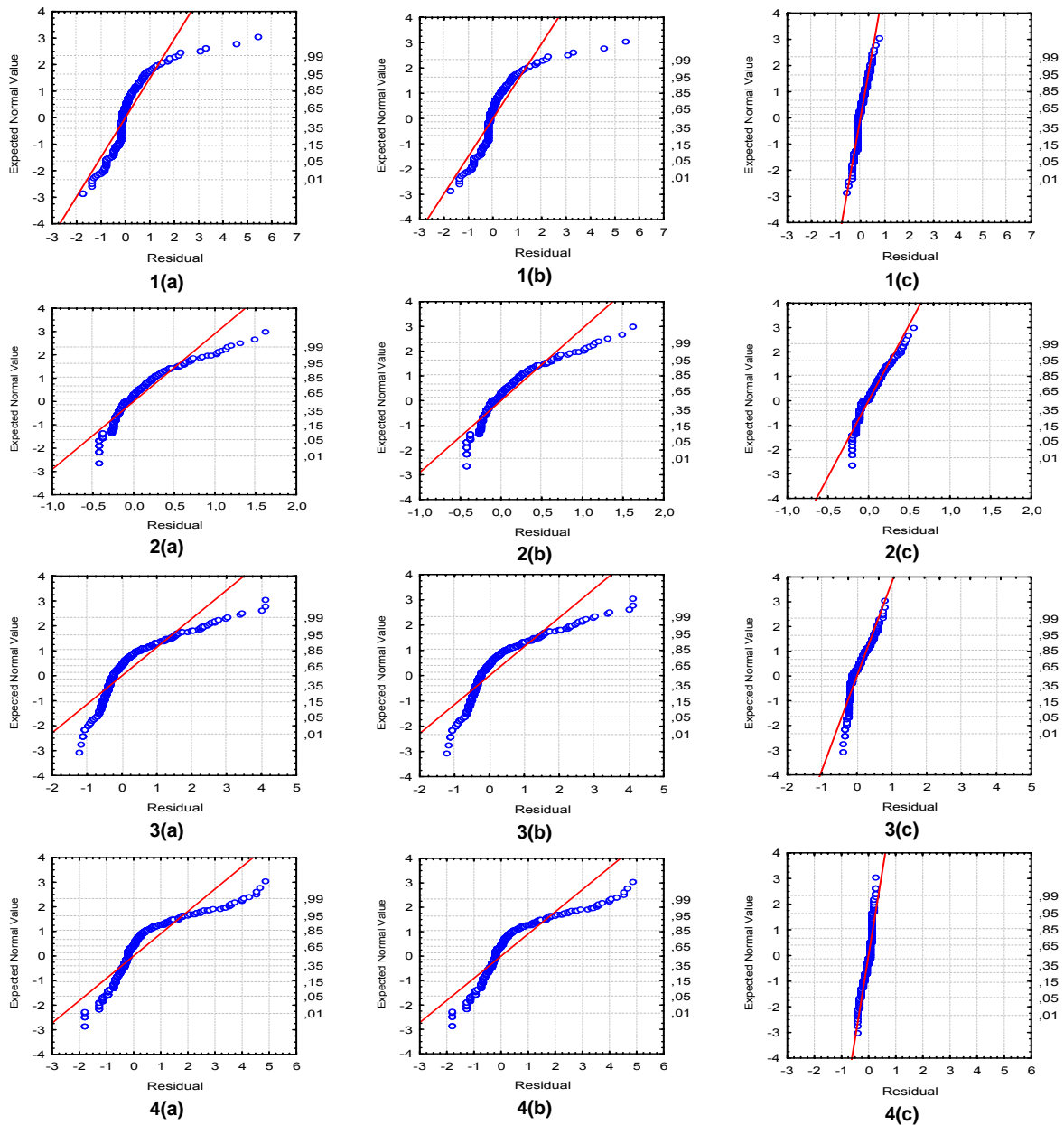


Figura 16 – Resíduos em função do valor normal esperado para o experimento com brócolis. Figuras 16.1, 16.2, 16.3 e 16.4, (a), (b) e (c) para as variáveis cabeças comerciais (CC), cabeças comerciais (CC) mais constante e transformação da variável, cabeças comerciais (CC) mais constante, respectivamente; para as quatro épocas. Santa Maria – RS. 2008.

Na Figura 17, têm-se os resíduos para ramificações comerciais de brócolis.

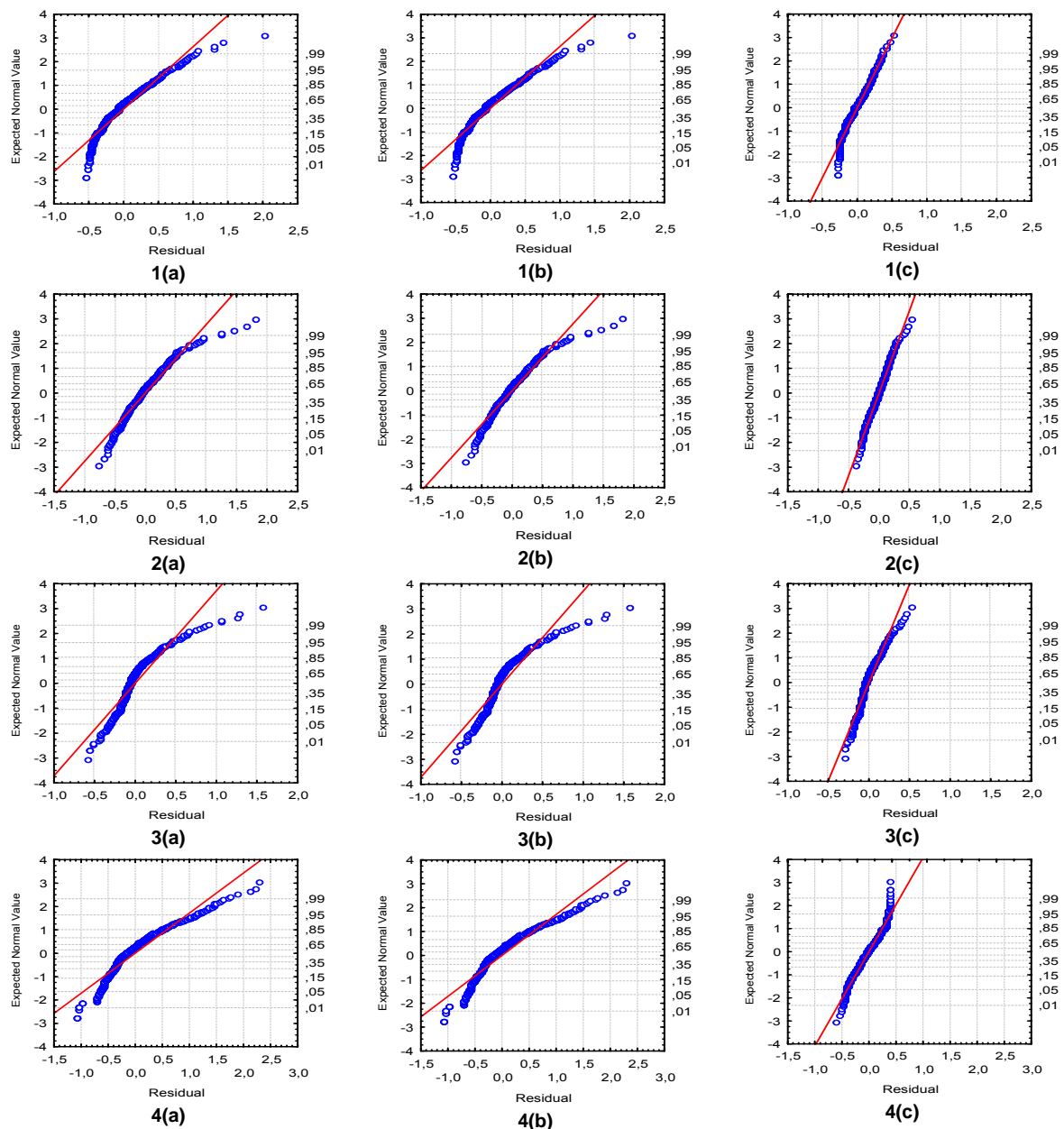


Figura 17 – Resíduos em função do valor normal esperado para o experimento com brócolis. Figuras 17.1, 17.2, 17.3 e 17.4, (a), (b) e (c) para as variáveis ramificações comerciais (RC), ramificações comerciais (RC) mais constante e transformação da variável ramificações comerciais (RC) mais constante, respectivamente; para as quatro épocas. Santa Maria – RS. 2008.

Na Figura 18, têm-se os resíduos para cabeças não comerciais de brócolis.

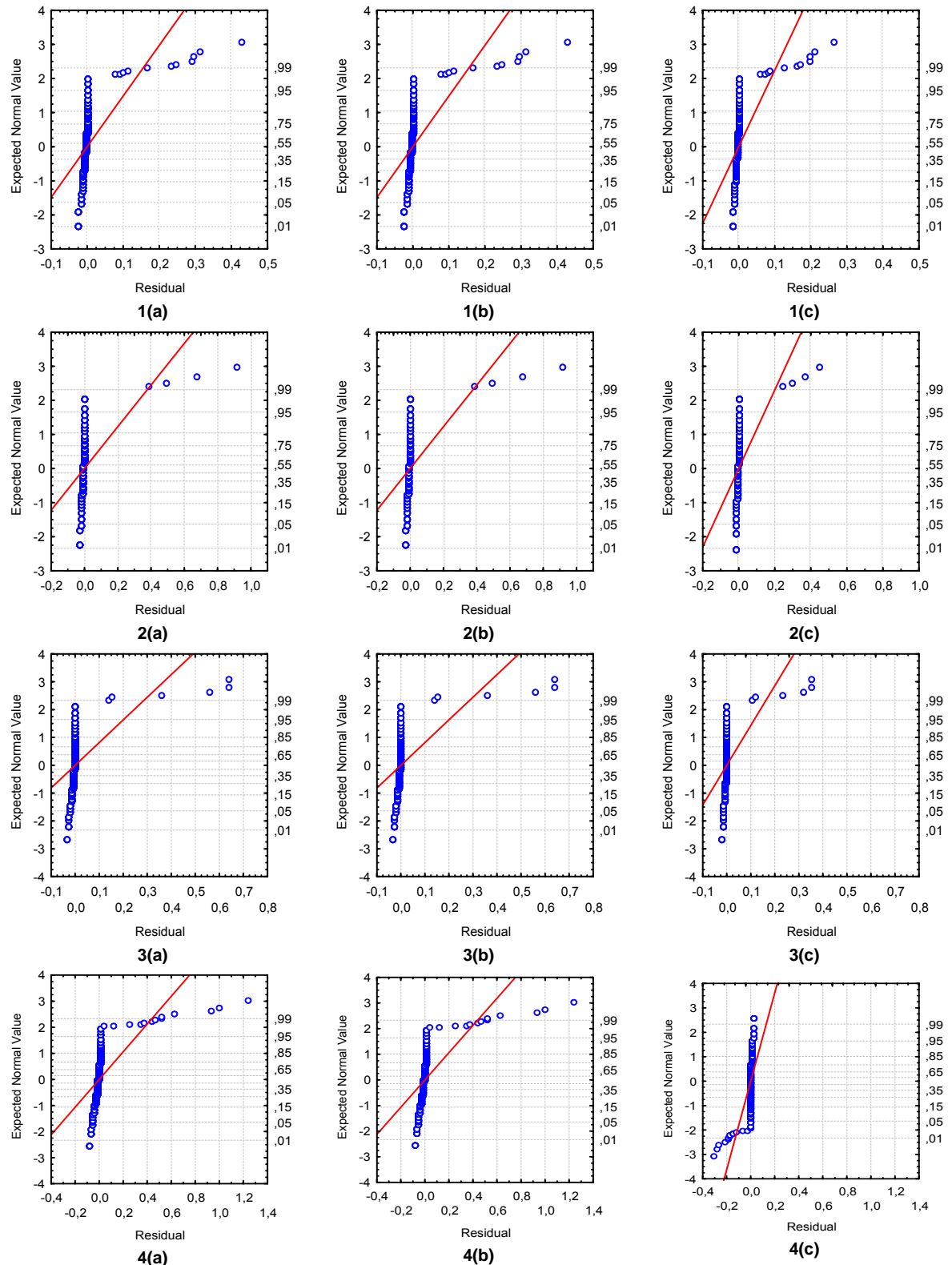


Figura 18 – Resíduos em função do valor normal esperado para o experimento com brócolis. Figuras 18.1, 18.2, 18.3 e 18.4, (a), (b) e (c) para a variável cabeças não comerciais (CNC), cabeças não comerciais (CNC) mais constante e transformação da variável cabeças não comerciais (CNC) mais constante, respectivamente; para as quatro épocas. Santa Maria – RS. 2008.

Na Figura 19, têm-se os resíduos para ramificações não comerciais de brócolis.

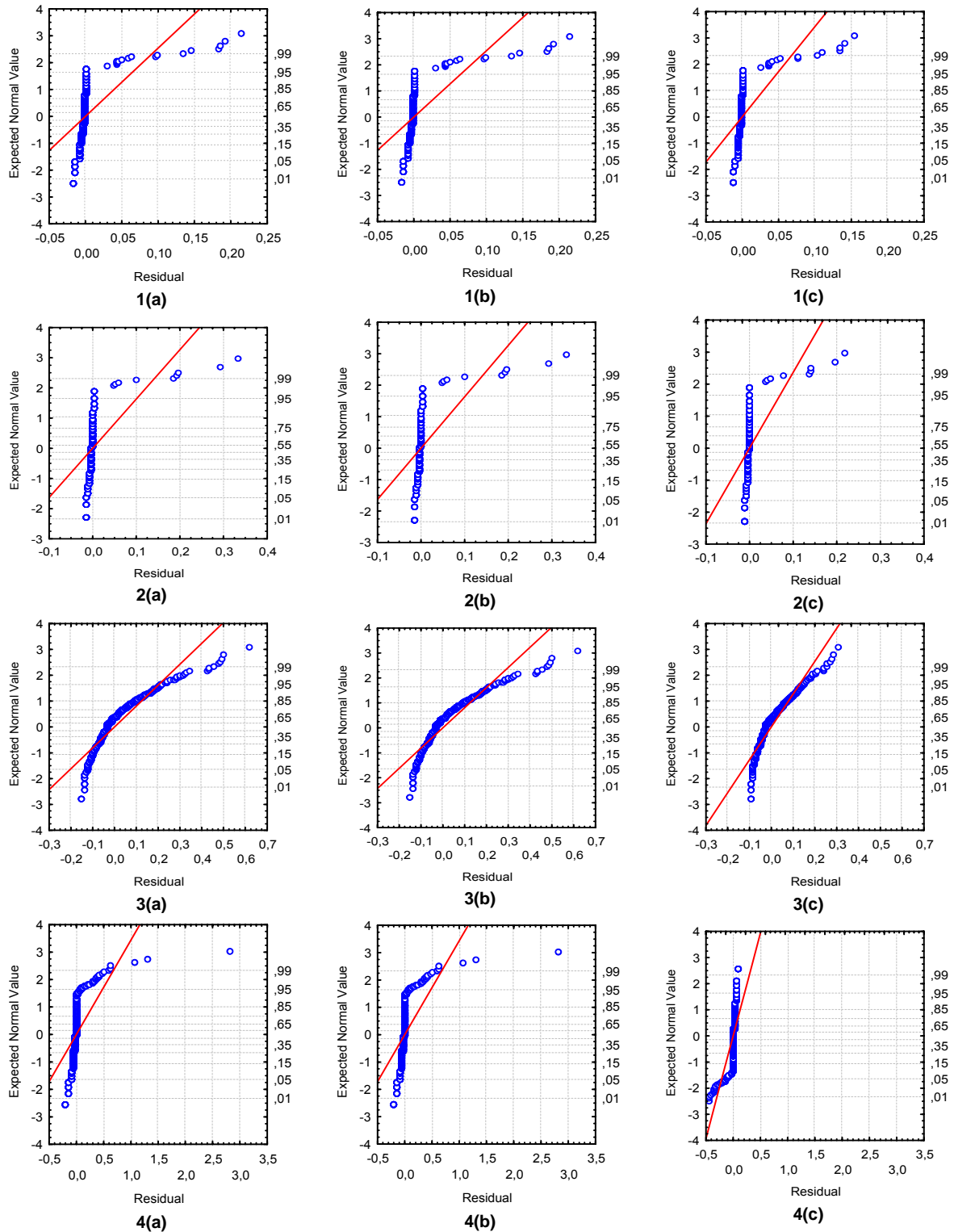


Figura 19 – Resíduos em função do valor normal esperado para o experimento com brócolis. Figuras 19.1, 19.2, 19.3 e 19.4, (a), (b) e (c) para as variáveis ramificações não comerciais (RNC), ramificações não comerciais (RNC) mais constante e transformação da variável: ramificações não comerciais (RNC) mais constante, respectivamente; para as quatro épocas. Santa Maria – RS. 2008.

Nas Figuras 16, 17, 18 e 19, observa-se que a análise residual, de cabeças e ramificações comerciais, sugere que o uso da variável transformada parece ter sido mais eficiente do que para cabeças e ramificações não comerciais. Nestas últimas, observam-se valores residuais muito dispersos, o que leva a concluir que a transformação não tenha controlado a variação de forma efetiva.

Novamente recorrendo à análise exploratória destes dados, uma possível explicação para este resultado seria que para cabeças e ramificações não comerciais ocorre uma frequência de valores zero elevada (Figura 11). E talvez, esse seja o motivo que mesmo com o emprego da transformação os resíduos ainda não estejam nem próximos da normalidade.

4.4 Considerações Finais

Verificou-se que a variabilidade em experimentos olerícolas gerada por dados com excesso de zero é extremamente difícil de contornar.

Nas culturas de abobrinha italiana e pimentão foram considerados a fitomassa e o número de frutos colhidos. Observou-se que apesar de culturas distintas, o comportamento destas variáveis foi bastante similar quanto à variabilidade. Considerando-se as diferenças em termos percentuais (%) do coeficiente de variação antes e após transformação da fitomassa e do número de frutos, tem-se que para a fitomassa, houve uma redução no coeficiente de variação acima de 90%, e para o número de frutos essa redução varia em torno de 80% (APÊNDICE A). Mas apesar disto, a transformação indicada para cada experimento foi distinta; pois para a cultura de abobrinha o valor da potência λ indica a raiz quarta e para a cultura de pimentão o inverso desta.

Para a cultura de brócolis analisaram-se as cabeças comerciais e não comerciais e as ramificações comerciais e não comerciais. Observou-se que para todas elas a transformação indicada pelo método foi à logarítmica, exceto para época 4, que apresentou resultados diferentes daqueles encontrados para as demais épocas. Considerando-se as diferenças em termos percentuais (%) do coeficiente de variação antes e após, para estas variáveis, verifica-se que não houve uma redução no coeficiente de variação na época 2 para cabeças e ramificações comerciais e também para cabeças não comerciais na época 1. Nas demais épocas todas as variáveis apresentaram uma redução no coeficiente de variação (APÊNDICE B).

Os resultados encontrados para todos os experimentos mostram uma redução na variabilidade dos dados após a transformação. Porém, os testes para a homogeneidade de variâncias e a análise residual, destes experimentos, demonstram que o uso da transformação não foi suficiente para tornar os dados homocedásticos.

Como sugestão para trabalhos futuros recomenda-se o estudo sobre a que níveis percentuais de zeros o uso de uma transformação seria eficaz. Pois, neste estudo verificou-se que a frequência elevada de valores iguais a zero influenciou nos resultados obtidos após a transformação, mesmo que esta tenha reduzido a variabilidade dos dados.

A partir desses resultados, observou-se que a não satisfação da pressuposição de homogeneidade de variâncias, que é a característica predominante neste tipo de estudo, pode ser controlada de forma eficaz com o uso de transformação. Porém esta estaria ligada ao nível percentual de ocorrências de valores nulos. Deste modo, a idéia de que, poder-se-ia limitar faixas percentuais para as frequências de zeros ajudaria na escolha de determinadas transformações que seriam indicadas para a adequação das pressuposições da análise de variância.

5 CONCLUSÕES

Dependendo da cultura e da variável analisada, existe a necessidade de transformações diferenciadas.

Para abobrinha italiana a transformação indicada foi a raiz quarta, para o número e fitomassa de frutos. Já para essas mesmas variáveis na cultura de pimentão a transformação indicada foi o inverso da raiz quarta.

Para o experimento com brócolis, a transformação mais adequada para as variáveis cabeças comerciais e não comerciais, e ramificações comerciais e não comerciais foi a logarítmica.

Mesmo com a redução dos valores dos coeficientes de variação, após a aplicação da transformação dos dados, a pressuposição da homogeneidade dos erros foi rejeitada.

6 REFERÊNCIAS BIBLIOGRÁFICAS

- ASCOMBE, F.J.; TUKEY, J.W. The criticism of transformations, paper presented before the **American Statistical Association and the Biometric Society**, Montreal, September 12, 1952.
- BERRY, D.A..Logarithmic transformations in ANOVA. **Biometrics**, v.43, p 439-456, 1987.
- BORGATTO, A. F. **Modelos para proporções com superdispersão e excesso de zeros** – um procedimento *bayesiano*. Piracicaba, 2004. 90 p. Tese (doutorado) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo.
- BOX, G. E. P.; COX, D.R. An analysis of transformations. **Journal of the Royal Society. Series B (Methodological)**, vol. 26, n 02, p. 211 – 252, 1964.
- CAMPOS, G. M., Estatística para docentes e pós-graduados. In: Programa de incentivo à produção de material didático do SIAE, Pró-reitoria de graduação e Pós-graduação da USP. São Paulo: USP, 2000. Disponível em: <<http://www.siae.usp.br/>>. Acesso em: 18 maio 2006.
- CARPES , R.H. **Variabilidade da produção de frutos de abobrinha italiana em função do manejo**. 2006. 68p. Dissertação (Mestrado em Agronomia) - Universidade Federal de Santa Maria, Santa Maria.
- DRAPER, N. R.; SMITH, H. **Applied regression analysis**. New York: John Wiley, 1981. 709p.
- DEMÉTRIO, C. G. B.; HINDE, J. Half-normal plots and overdispersion. **GLIM Newsletter**, v. 27, p.9-26, 1997.

- GHOSH, S. K.; MUKHOPADHYAY, P.; LU, JC. Bayesian analysis of zero-inflated regression models. **Journal of Statistical Planning and Inference**, v.136, p.1360 – 1375, 2004.
- GRIFFITHS, D. A. Interval Estimation for the Three-Parameter Lognormal Distribution via the Likelihood Function. **Applied Statistics**, vol. 29, n. 1, p. 58-68, 1980.
- GROGGER, J.T.; CARSON, R.T. Models for truncated counts. **Journal of Applied Econometrics**, v.6, p. 225-238, 1991.
- HALL, D. B. Zero-inflated Poisson and Binomial regression with random effects: a case study. **Biometrics**, v.56, p 1030-1039, 2000.
- HALL, D. B. & BERENHAUT, K.S. Score tests of Heterogeneity and overdispersion in zero-inflated Poisson and Binomial regression models. **The Canadian Journal of Statistics**, v. 30, n° 3, p.16, 2002.
- HILL, B. M. The three-parameter lognormal distribution and Bayesian analysis of a point-source epidemic. **Journal of the American Statistical Association**, vol. 58, n. 301, p. 72 – 84, 1963.
- HINDE, J.; DEMÉTRIO, C. G. B.; Overdispersion: models and estimation. **Computational Statistics and Data Analysis**, v. 27, p.151-170, 1998a.
- _____ Overdispersion: models and estimation. **ABE**: São Paulo, p. 73, 1998b.
- _____ Overdispersion: models and estimation. **Berkley**: University of California, p.110, 2000.
- LAMBERT, D. Zero-inflated Poisson regression, with application to defects in manufacturing. **Technometrics**, v.34, p 1-14. 1992.

- LEE, A. H.; et al. Modelling bivariate count series with excess zeros. **Mathematical Biosciences**, v. 196, p. 226 – 237, 2005.
- LORENTZ, L. H; et al. Variabilidade da produção de frutos de pimentão em estufa plástica. **Ciência Rural**. v.35, p.316 – 323, 2005.
- MONTGOMERY, D. C. **Design and Analysis of experiments**. Ed. 3rd. New York: John Wiley, 1991. 649p.
- MOORE, P.G.; TUKEY, J.W. Answer to query no 112. **Biometrics**, v. 10, p. 562-568, 1954.
- NIKOLAU, M.; et al. Effect of mean and variance heterogeneity on genetic evaluations of Lesbos dairy sheep. **Livestock Production Science**, v. 88, Issue 1-2, p. 107-115, 2004.
- RIDOUT, M.; DEMÉTRIO, C.G.B Generalized linear models for positive count data. **Revista de Matemática e Estatística**, v. 139, p. 139 – 148, 1992.
- RIDOUT, M.; DEMÉTRIO, C.G.B.; HINDE, J. Models for count data with many zeros. In: **INTERNATIONAL BIOMETRIC CONFERENCE**, Cape Town, 1998. Proceedings, Cape Town: IBC,. p.1 – 13, 1998.
- SIQUEIRA, A. L. **Uso de transformação em análise de variância e análise de regressão**. Piracicaba, 1983. 154p. Tese (doutorado) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo.
- SOUZA, M.F.; et al. Tamanho da amostra para peso da massa de frutos, na cultura da abóbora italiana em estufa plástica. **Revista Brasileira de Agrociência**, v.8, p.123 -128, 2002.
- TUKEY, J. W. On the Comparative Anatomy of Transformations. **The Annals of Mathematical Statistics**, Vol. 28, n. 3, pp. 602-632, 1957.

VIEIRA, A.M.C. **Modelos para dados de proporções com superdispersão aplicados ao controle biológico**. Piracicaba, 1998. 61p. Dissertação (mestrado) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo.

VIEIRA, A.M.C; HINDE, J.; DEMÉTRIO, C.G.B. Zero-inflated proportion data models applied to a biological control assay. **Journal of Applied Statistics**, v.27, p 373-389, 2000.

YAMAMURA, K. Transformation using $(x + 0.5)$ to stabilize the variance of populations. **Journal Researches on population ecology**. Publisher Springer Japan, v. 42, n.3, p. 229 – 234, 1999.

APÊNDICES

APÊNDICE A - Diferença em termos percentuais (%) do coeficiente de variação antes e após transformação de fitomassa e número de frutos de abobrinha italiana, cultivados nas estações sazonais verão/outono (época 1) e inverno/primavera (época 2). Santa Maria – RS. 2008.

	Época 1		Época 2	
	Túnel 1	Túnel 2	Túnel 1	Túnel 2
Fitomassa	65,51	65,61	93,65	93,84
Nº de Frutos	88,82	88,55	83,70	84,56

* valores negativos indicam aumento na variação.

APÊNDICE B - Diferença em termos percentuais (%) do coeficiente de variação antes e após transformação de fitomassa e número de frutos de pimentão, cultivados nas estações sazonais verão/outono (época 1) e inverno/primavera (época 2). Santa Maria – RS. 2008.

	Pimentão	
	Época 1	Época 2
Fitomassa	93,49	91,29
Nº de Frutos	80,03	78,99

* valores negativos indicam aumento na variação.

APÊNDICE C - Diferença em termos percentuais (%) do coeficiente de variação antes e após transformação de cabeças comerciais e não comerciais, e ramificações comerciais e não-comerciais, de brócolis, cultivado em quatro estações. Santa Maria – RS. 2008.

	Época 1	Época 2	Época 3	Época 4
Cabeças comerciais	-6,74	21,41	-22,24	93,16
Ramificações comerciais	-100,00	43,59	-33,75	78,22
Cabeças não comerciais	-26,10	42,52	99,17	80,67
Ramificações não comerciais	92,91	77,27	99,74	98,94

* valores negativos indicam aumento na variação.