

APLICAÇÃO DA TRANSFORMAÇÃO RAIZ QUADRADA, NA ANÁLISE DA VARIÂNCIA DE DADOS EXPERIMENTAIS (*). HERMANO VAZ DE ARRUDA. Ao executar a análise de variância de uma experiência deve-se levar em conta que estão sendo devidamente respeitadas diversas hipóteses básicas, as quais proporcionam validade e nível de probabilidade corretos aos chamados testes de significância. A discordância dos dados experimentais com as hipóteses pré-estabelecidas pode conduzir o experimentador a inferências falsas.

As hipóteses são as seguintes:

1) Os efeitos de tratamentos e blocos devem ser aditivos. Esta condição é imposta por um modelo matemático, variável com o delineamento usado. Para um delineamento em blocos ao acaso o modelo seria:

$$Y_{ij} = m + t_i + b_j + e_{ij}$$

m que Y_{ij} é o valor observado relativo à parcela que recebeu o tratamento i no bloco j ; m é a média geral; t_i é o efeito do tratamento i ; b_j é o efeito do bloco j ; e_{ij} é o erro experimental associado àquele valor observado.

2) Os valores dos erros, e_{ij} , devem ser independentes uns dos outros, isto é, o erro de uma observação não deve estar associado ao erro de qualquer outra observação.

3) Os erros e_{ij} devem ter uma variância comum. É considerado que a variância do erro para as diversas comparações de tratamentos são homogêneas. Admitindo-se a existência desta hipótese, é possível usar um erro médio para todas as comparações de tratamentos, como é usualmente feito nas análises de rotina.

4) A variável e_{ij} deve ter uma distribuição normal. Quando os dados experimentais não se adaptam às hipóteses enumeradas recomenda-se fazer uma transformação, usando-se uma nova escala de medidas que seja capaz de assegurar a citada adaptação.

No presente trabalho é apresentada a aplicação de uma transformação dos dados originais, mostrando o seu efeito sobre a homogenização da variância do erro e sobre a aditividade dos efeitos dos tratamentos e blocos.

Material e método — Usaram-se os dados de uma experiência com inseticidas sistêmicos no combate aos pulgões do algodoeiro, instalada na Estação Experimental do Instituto Agrônomico, em Ribeirão Preto. A variável X , número médio de pulgões por folha, determinada em 15 folhas,

(*) Recebida para publicação em 29 de janeiro de 1959.

foi usada para medir a eficiência dos diversos tratamentos no contróle da praga. Estes dados referem-se a uma das contagens do número de pulgões feitas na experiência e, portanto, só servem como ilustração do método estatístico.

Dados como estes e outros referentes também a contagens, tais como número de ervas más por canteiro nas experiências com herbicidas, número de galhos secos em cafeeiros em função de diversos adubos usados etc., tendem a seguir a distribuição de Poisson. Nesta distribuição a variância é igual à média. Nos casos mais freqüentes os dados se aproximam apenas da distribuição teórica, apresentando a variância proporcional à média. Esta proporcionalidade, desde que as médias sejam bem diferentes, provoca a heterogeneidade na variância do erro e, às vezes, a não aditividade dos efeitos de blocos e tratamentos (1).

A transformação indicada para o caso da presente experiência é a raiz quadrada dos dados originais (2).

Assim, fêz-se a citada transformação, e mediante a comparação das duas análises, uma correspondente aos dados originais e outra aos transformados, foi possível julgar a eficiência da transformação usada.

Análise dos dados originais — Observando inicialmente o quadro 1, correspondente aos dados originais x , nota-se que existem diferenças acentuadas nas amplitudes de variação dentro dos tratamentos. Estas diferenças, indo de 37 para o tratamento I até 193 para o tratamento 8, fazem suspeitar da heterogeneidade da variância do erro.

QUADRO 1. — Dados originais, x , correspondentes ao número médio de pulgões por fôlha

Tratamentos	B L O C O S				Totais dos tratamentos	Amplitude de variação
	I	II	III	IV		
G1 1.....	25	7	8	44	84	37
2.....	13	30	19	74	136	61
3.....	54	36	10	16	116	44
G2 4.....	66	86	114	177	443	111
5.....	124	63	132	160	479	97
6.....	266	185	121	241	813	145
7.....	261	202	151	277	891	126
8.....	177	226	370	298	1071	193
Totais dos blocos.....	986	835	925	1287	4033	

Para estudar a heterogeneidade da variância decompôs-se a variância total do erro em três componentes: dentro do Grupo I (G1), que reúne

(1) SNEDECOR, G. W. Statistical methods. The Iowa State College Press, Iowa, 1956. p.314-328.

(2) BARTLETT, M. S. The use of transformations. Biometrics 3:39-52, 1947.

os inseticidas sistêmicos; dentro do Grupo 2 (G2), dos não sistêmicos, inclusive a testemunha; e a variância, para a comparação entre grupos. Esta decomposição aparece no quadro 2, juntamente com a análise da variância.

QUADRO 2. — Análise da variância dos dados originais, \bar{x} , número médio de pulgões por fôlha

Fontes de variação	Graus de liberdade	Soma de quadrados	Quadrados médios	F
Blocos	3	14 392		
Tratamentos	7	258 363	36 909	12,7
Erro	21	61 022	2 905	
Total	31	333 777		
Dentro G1	6	2 564	427	
Dentro G2	12	55 043	4 586	
G2 versus G 1	3	3 415	1 138	
Erro	21	61 022	2 905	

Coefficiente de variação = 42,8%

Observa-se que a variância dentro do G1 (427) é 11 vezes menor do que a variância dentro do G2 (4 586). Para a comparação G1 versus G2, tem-se uma função destas duas variâncias. Conclui-se pela heterogeneidade entre as variâncias dos dois grupos, pois o valor 11,7 para F com 12 a 6 graus de liberdade, é altamente significativo. O uso de um só erro médio (pooled error) para tôdas as comparações fica, pois, rejeitado (³).

O coeficiente de variação no valor de 42,8%, muito alto, faz suspeitar da não aditividade dos efeitos de blocos e tratamentos, embora não se possa executar um teste, no presente caso, em vista de se ter verificado a heterogeneidade na variância do erro.

Análise dos dados transformados — Os dados, após a transformação para \sqrt{x} , são apresentados no quadro 3.

QUADRO 3. — Dados transformados para \sqrt{x} , sendo \bar{x} o número médio de pulgões por fôlha.

Tratamentos	B L O C O S				Totais tratamentos
	I	II	III	IV	
G1 1	5,00	2,64	2,83	6,63	17,10
2	3,60	5,47	4,36	8,60	22,03
3	7,35	6,00	3,16	4,00	20,51
G2 4	8,12	9,27	10,68	13,30	41,37
5	11,13	7,97	11,49	12,65	43,21
6	16,31	13,60	11,00	15,52	56,43
7	16,15	14,21	12,29	16,64	59,29
8	13,30	15,03	19,23	17,26	64,82
Totais dos blocos	80,96	74,16	75,04	94,60	324,76

(³) EISENHART, C. The assumptions underlying the analysis of variance. Biometrics 3:1-21. 1947.

A análise da variância e a decomposição do erro estão incluídas no quadro 4.

QUADRO 4. — Análise da variância dos dados transformados para \sqrt{x}

Fontes de variação	Graus de liberdade	Soma de quadrados	Quadrados médios	F
Blocos	3	33,39		
Tratamentos	7	623,65	89,09	23,0
Não aditividade	1	1,20	1,20	0,3
Erro	20	77,45	3,87	
Total	31	735,69		
Dentro G1	6	22,47	3,74	
Dentro G2	12	50,72	4,23	
G2 versus G1	3	5,45	1,82	
Erro	21	78,64	3,74	

Coefficiente de variação = 19,1%.

Pode-se ver no quadro 4 que a transformação homogenizou a variância do erro: 3,74 dentro do G1, 4,23 dentro do G2 e 1,82 para a comparação G2 versus G1. Verificada a homogeneidade da variância do erro pode-se agora usar uma variância média, 3,74, para tôdas as comparações.

O coeficiente de variação reduziu de muito na análise dos dados transformados, passando de 42,8% para 19,1%. Houve, pois, uma grande redução no erro experimental e, portanto, um grande aumento na precisão dos resultados.

Fêz-se, segundo Tukey (4), um teste para verificar a aditividade dos efeitos de blocos e tratamento. Separou-se do erro um componente ao qual êle atribui a não aditividade, com 1 grau de liberdade, numa classificação de dupla entrada (blocos x tratamentos). O valor de F, 0,3, não significativo, para êste componente, apresentado no quadro 4, deixa pouca dúvida sôbre a falta de aditivade.

Conclusões — O uso da transformação raiz quadrada dos dados originais foi eficiente, assegurando a homogeneidade da variância do erro e a aditividade dos efeitos de blocos e tratamentos, na análise dos dados transformados.

A eficiência da transformação usada pôde ser verificada também através do coeficiente de variação. Após a transformação êste coeficiente reduziu-se para menos da metade, isto é, de 42,8% na análise dos dados origi-

(4) TUKEY, J. W. One degree of freedom for non-additivity. Biometrics 5:232-242. 1949.

nais para 19,1% na análise dos dados transformados. A precisão nas estimativas das médias dos tratamentos dobrou, após o uso da transformação. ESTAÇÃO EXPERIMENTAL DE RIBEIRÃO PRÊTO, INSTITUTO AGRONÔMICO DO ESTADO DE SÃO PAULO.

SQUARE ROOT TRANSFORMATION IN THE ANALYSIS OF
CERTAIN FIELD EXPERIMENTAL DATA

SUMMARY

This transformation was applied to the analysis of data dealing with the determination of field populations of the cotton aphid, *Aphis gossypii* Glov.

The data thus transformed were analysed in comparison with the original ones. It was verified that the transformation increased the homogeneity of the error variance and secured the additivity of the effects of block and treatment. The advantage of the use of transformed data was likewise confirmed by a reduction in the coefficient of variation from 42.8% for the original data to 19.1% after transformation.