Diagnostics and Transformation

Case studies from history

November 18, 2021

Overview

- We will review some classic references on the analysis of variance, tests for the assumptions of AOV, and what can be done when data are not suitable for AOV.

Some consequences when the assumptions for the analysis of variance are not satisfied

- W. G. Cochran. Biometrics, 3(1):22–38, March 1947 https://www.jstor.org/stable/3001535
    - Text quoted from sources will be typeset in san serif font

Purposes of the Analysis of Variance

- To estimate certain treatment differences that are of interest.
- To obtain some idea of the accuracy of our estimates, e.g., by attaching to them estimated standard errors, fiducial or confidence limits, etc.
- To perform tests of significance.

Purposes of the Analysis of Variance

- To estimate certain treatment differences that are of interest.
  - *Efficient* - the variance between the estimate and the true difference is as small as possible
- To obtain some idea of the accuracy of our estimates, e.g., by attaching to them estimated standard errors, fiducial or confidence limits, etc.
  - *Unbiased* - the estimated variance is the same as the expected value of the variance.
- To perform tests of significance.
  - *Valid* - If the reported p-value is, say 0.026, then the chance of a greater test statistic is close to 0.026
  - *Sensitive* - real treatment differences are detected as often as possible

Purposes of the Analysis of Variance

- To estimate certain treatment differences that are of interest.
  - *Efficient* - the variance between the estimate and the true difference is as small as possible
- To obtain some idea of the accuracy of our estimates, e.g., by attaching to them estimated standard errors, fiducial or confidence limits, etc.
  - *Unbiased* - the estimated variance is the same as the expected value of the variance.
- To perform tests of significance.
  - *Valid* - If the reported p-value is, say 0.026, then the chance of a greater test statistic is close to 0.026
  - *Sensitive* - real treatment differences are detected as often as possible
- **These desirable properties are met when the assumptions of AOV are met.**

Assumptions Required for the Analysis of Variance

- For this discussion, we assume three types of effects:
  - treatment effects
  - environmental effects
  - experimental errors

Assumptions Required for the Analysis of Variance

- The treatment effects and the environmental effects must be additive.
- The experimental errors must all be independent.
- The experimental errors must have a common variance.
- The experimental errors should be normally distributed.

Assumptions Required for the Analysis of Variance

- The treatment effects and the environmental effects must be additive.
  - *Tukey's 1 d.f.*
- The experimental errors must all be independent.
  - *Moran I*
- The experimental errors must have a common variance.
  - *Levene's, Bartlett's*
- The experimental errors should be normally distributed.
  - *Shapiro-Wilks, Skewness, Kurtosis*

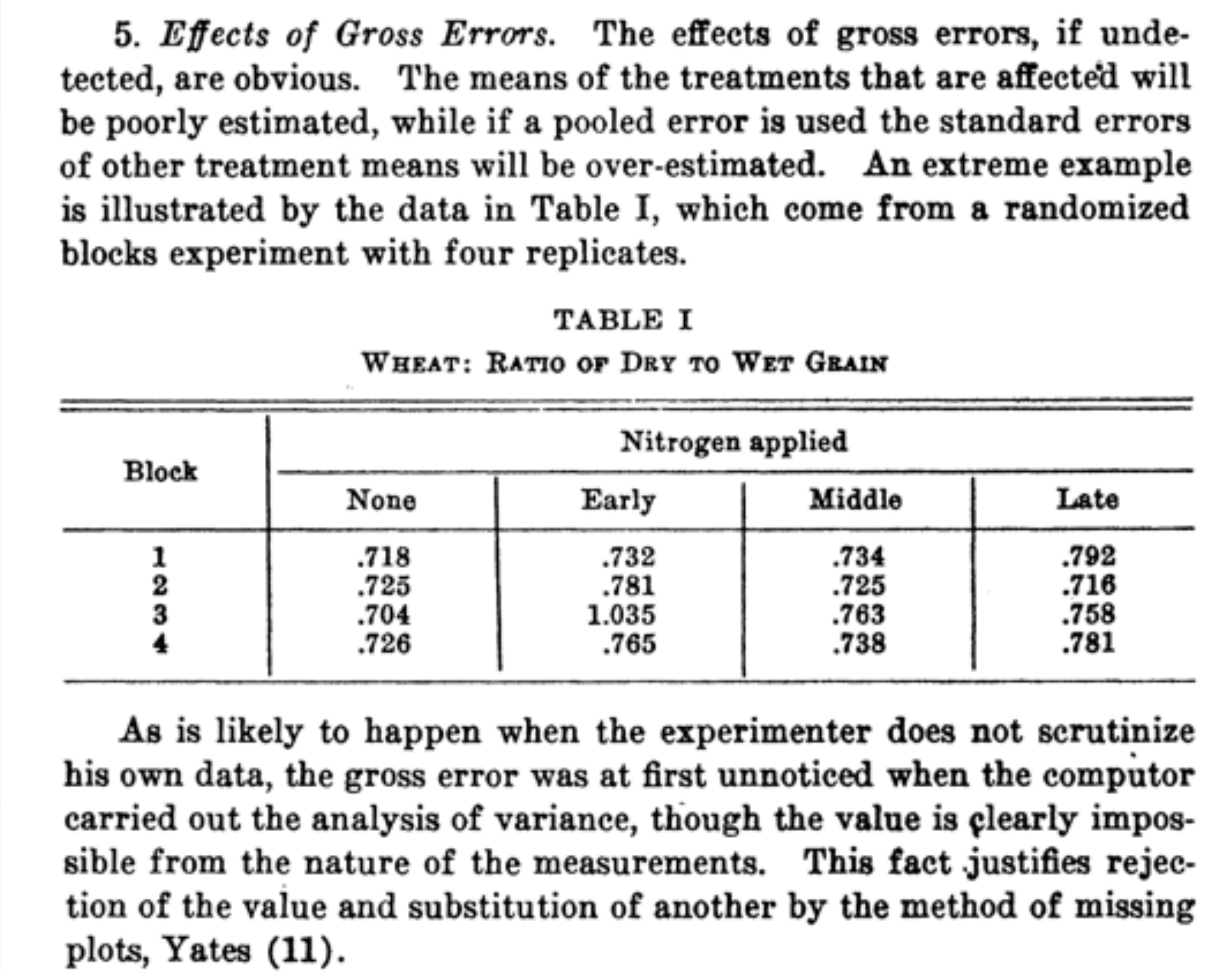

5. *Effects of Gross Errors.* The effects of gross errors, if undetected, are obvious. The means of the treatments that are affected will be poorly estimated, while if a pooled error is used the standard errors of other treatment means will be over-estimated. An extreme example is illustrated by the data in Table I, which come from a randomized blocks experiment with four replicates.

TABLE I

WHEAT: RATIO OF DRY TO WET GRAIN

Block	Nitrogen applied			
	None	Early	Middle	Late
1	.718	.732	.734	.792
2	.725	.781	.725	.716
3	.704	1.035	.763	.758
4	.726	.765	.738	.781

As is likely to happen when the experimenter does not scrutinize his own data, the gross error was at first unnoticed when the computor carried out the analysis of variance, though the value is clearly impossible from the nature of the measurements. This fact justifies rejection of the value and substitution of another by the method of missing plots, Yates (11).

Effects of Gross Errors

Cochran 1947 I.dat0

Cochran 1947 I

- Since these data are ratios, a value greater than 1 is unexpected
- This outlier distorts the distribution of the residuals, and no data transformation corrects with distortion.
- This example trial includes three duplicated columns so that we can compare the effects of outlier on analysis.

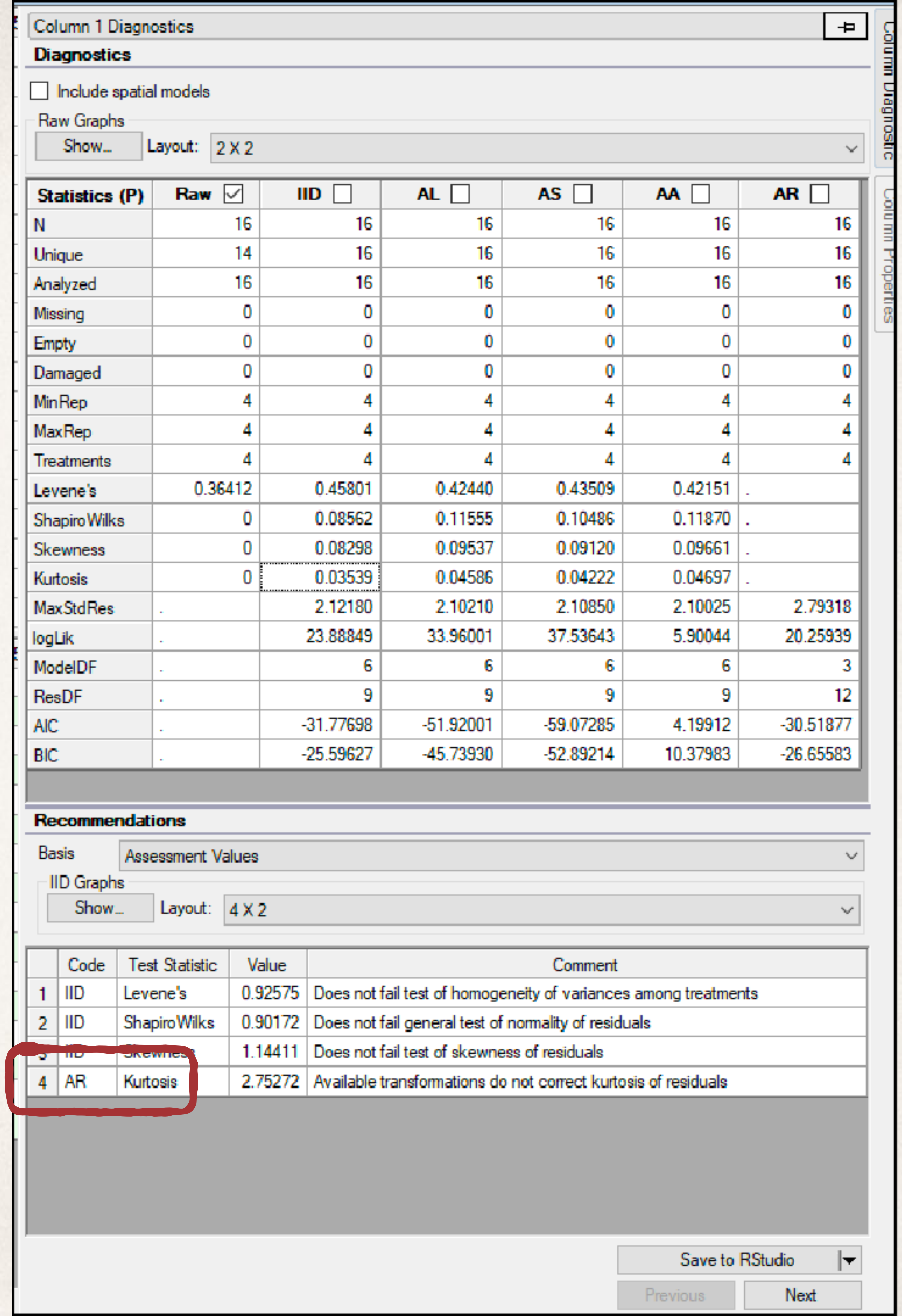
Assessment Data - Line 10

Column Number	1
Pest Type	
Pest Code	
Pest Name	
Crop Type, Code	C TRZAW
Crop Scientific Name	Triticum aestivum
Crop Name	Winter wheat
Rating Date	
Part Rated	SEED C
Rating Type	MOICON
Rating Unit/Min/Max	RATIO 0 1
Number of Subsamples	1
Assessed By	
Data Entry Date	*Aug-20-2019*
Rating Timing	
Days After First/Last Applic.	
Trt-Eval Interval	
Days After Emergence	
ARM Action Codes	
Number of Decimals	

Sub	*Rep*	*Blk*	*Col*	*Plot*	*Trt*	1
1	1	1	1	101	2	0.732
1	1	1	2	102	3	0.734
1	1	1	3	103	4	0.792
1	1	1	4	104	1	0.718
1	2	2	1	201	1	0.725
1	2	2	2	202	4	0.716
1	2	2	3	203	3	0.725
1	2	2	4	204	2	0.781
1	3	3	1	301	[illegible]	[illegible]
1	3	3	2	302	2	1.035
1	3	3	3	303	[illegible]	[illegible]
1	3	3	4	304	4	0.758
1	4	4	1	401	4	0.781
1	4	4	2	402	1	0.726
1	4	4	3	403	3	0.738
1	4	4	4	404	2	0.765

Cochran 1947 I

- ARM diagnostics does not detect the outlier, but the influence of the outlier is detectable by tests of normality
- The presence of the outlier is detectable as kurtosis in the AOV analysis. This kurtosis is not corrected by any known transformation, so analysis of ranks is advised.


Column 1 Diagnostics
Diagnostics
Include spatial models
Raw Graphs
Show... Layout: 2 X 2

Statistics (P)	Raw	IID	AL	AS	AA	AR
N	16	16	16	16	16	16
Unique	14	16	16	16	16	16
Analyzed	16	16	16	16	16	16
Missing	0	0	0	0	0	0
Empty	0	0	0	0	0	0
Damaged	0	0	0	0	0	0
Min Rep	4	4	4	4	4	4
Max Rep	4	4	4	4	4	4
Treatments	4	4	4	4	4	4
Levene's	0.36412	0.45801	0.42440	0.43509	0.42151	.
Shapiro Wilks	0	0.08562	0.11555	0.10486	0.11870	.
Skewness	0	0.08298	0.09537	0.09120	0.09661	.
Kurtosis	0	0.03539	0.04586	0.04222	0.04697	.
Max Std Res	.	2.12180	2.10210	2.10850	2.10025	2.79318
logLik	.	23.88849	33.96001	37.53643	5.90044	20.25939
ModelDF	.	6	6	6	6	3
ResDF	.	9	9	9	9	12
AIC	.	-31.77698	-51.92001	-59.07285	4.19912	-30.51877
BIC	.	-25.59627	-45.73930	-52.89214	10.37983	-26.65583

Recommendations
Basis: Assessment Values
IID Graphs
Show... Layout: 4 X 2

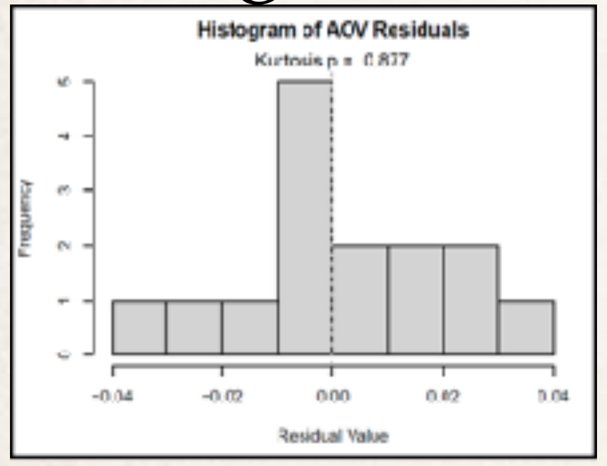
	Code	Test Statistic	Value	Comment
1	IID	Levene's	0.92575	Does not fail test of homogeneity of variances among treatments
2	IID	Shapiro Wilks	0.90172	Does not fail general test of normality of residuals
3	IID	Skewness	1.14411	Does not fail test of skewness of residuals
4	AR	Kurtosis	2.75272	Available transformations do not correct kurtosis of residuals

Save to RStudio
Previous Next

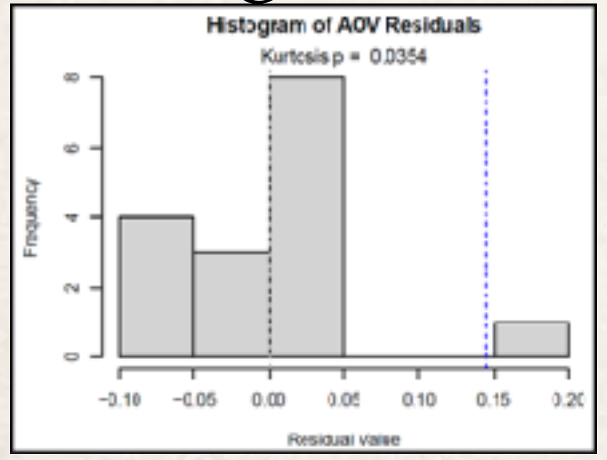
Cochran 1947 I

- Both marking the outlier as damaged followed by traditional AOV, and analysis of ranks suggest differences not detected in the AOV with the outlier.

- Histogram with outlier



- Histogram without outlier



Nov-10-2021 (Cochran 1947 I)

GDM Solutions, Inc.

Trial ID: Cochran 1947 I
Protocol ID: Cochran 1947 I
Project ID: Some Consequences

Location:
Investigator (Creator): Peter Claussen
Study Director:
Sponsor Contact:

Trial Year: 2019

Crop Type, Code	C, TRZAW	C, TRZAW	C, TRZAW
BBCH Scale	BCER	BCER	BCER
Crop Scientific Name	Triticum aestiv>	Triticum aestiv>	Triticum aestiv>
Crop Name	Winter wheat	Winter wheat	Winter wheat
Part Rated	SEED, C	SEED, C	SEED, C
Rating Type	MOICON	MOICON	MOICON
Rating Unit/Min/Max	RATIO, 0, 1	RATIO, 0, 1	RATIO, 0, 1
Number of Subsamples	1	1	1
Data Entry Date	Aug-20-2019	Dec-17-2019	Jan-17-2020
ARM Action Codes	IID	IID	AR
Trt No. / Treatment Name	1*	2*	3*
1 None	0.7183 -	0.7183 b	0.7183 b
2 Early	0.8283 -	0.7599 ab	0.8283 a
3 Middle	0.7400 -	0.7400 ab	0.7400 ab
4 Late	0.7618 -	0.7618 a	0.7618 ab
LSD P=.05	0.11596	0.04172	0.11596r
Standard Deviation	0.07249	0.02558	.
CV	9.51	3.44	.
Grand Mean	0.76206	0.74387	.
Levene's F^	0.926	1.048	0.926
Levene's Prob(F)	0.458	0.41	0.458
Friedman's X2	.	.	4.692
P(Friedman's X2)	.	.	0.196
Skewness^	1.1441*	-0.0856	1.1441*
Kurtosis^	2.7527*	-0.1935	2.7527*
Replicate F	0.980	0.259	
Replicate Prob(F)	0.4445	0.8531	
Treatment F	1.722	2.384	
Treatment Prob(F)	0.2317	0.1449	

TABLE II

MANHOLDS, PLANT NUMBERS PER PLOT

Block	Control		chalk			Lime			Total
	0	0	1	2	3	1	2	3	
I	140	49	98	135	117	81	147	330	897
II	142	37	132	151	137	129	131	112	971
III	36	114	130	143	137	135	103	130	928
IV	129	125	153	146	143	104	147	121	1068
Total	447	325	513	575	534	449	528	493	3864
Range	106	88	55	16	26	54	44	18	

Heterogeneity of errors may arise in several ways. It may be produced by mishaps or damage to some part of the experiment. It may be present in one or two replications through the use of less homogeneous material or of less carefully controlled conditions. The nature of the treatments may be such that some give more variable responses than others. An example of this type is given by the data in Table II.

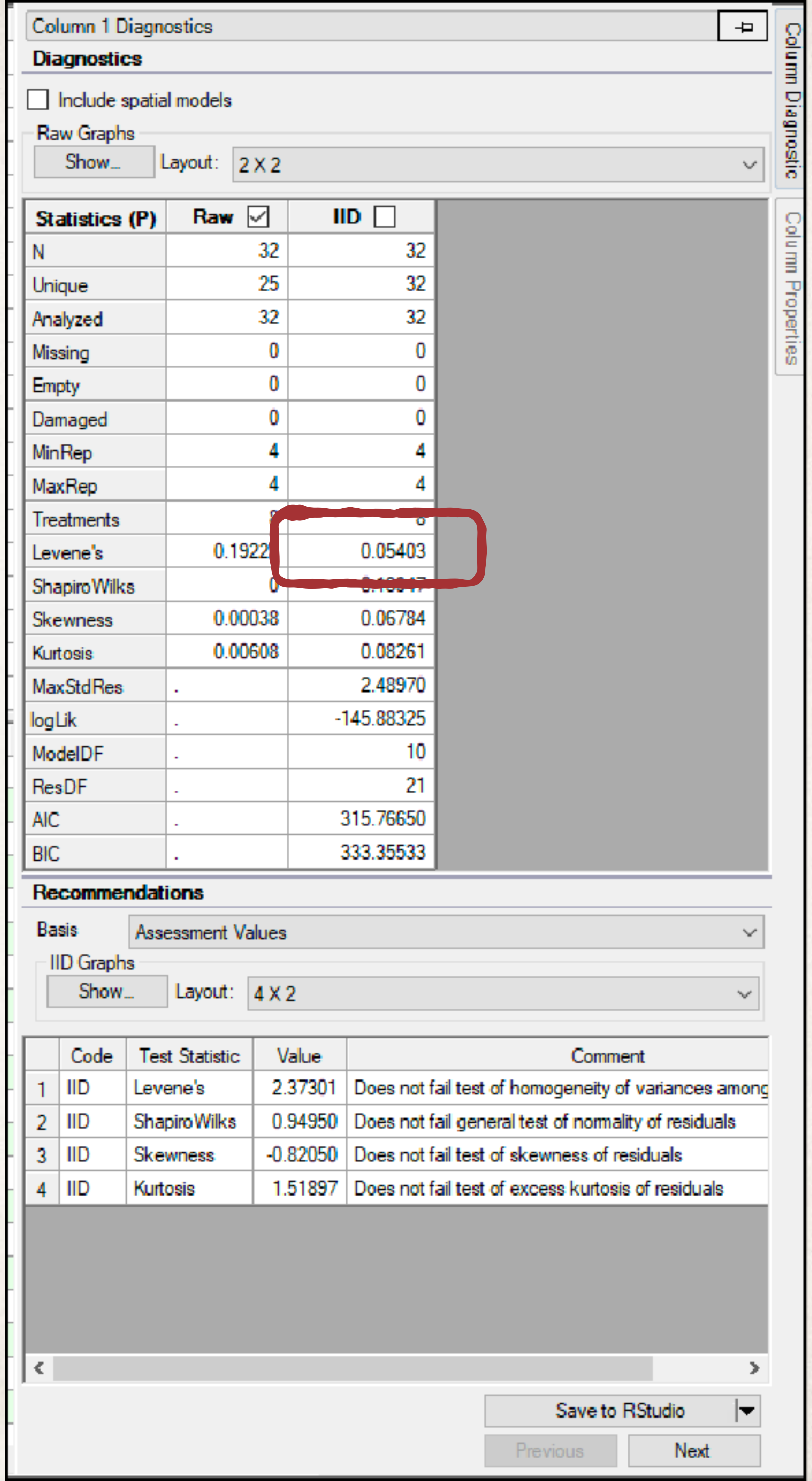
The experiment investigated the effects of three levels of chalk dressing and three of lime dressing on plant numbers of mangolds. There were four randomized blocks of eight plots each, the control plots being replicated twice within each block.[4]

Effects of Heterogeneity of Errors

Cochran 1947 II.dat0

Cochran 1947 II

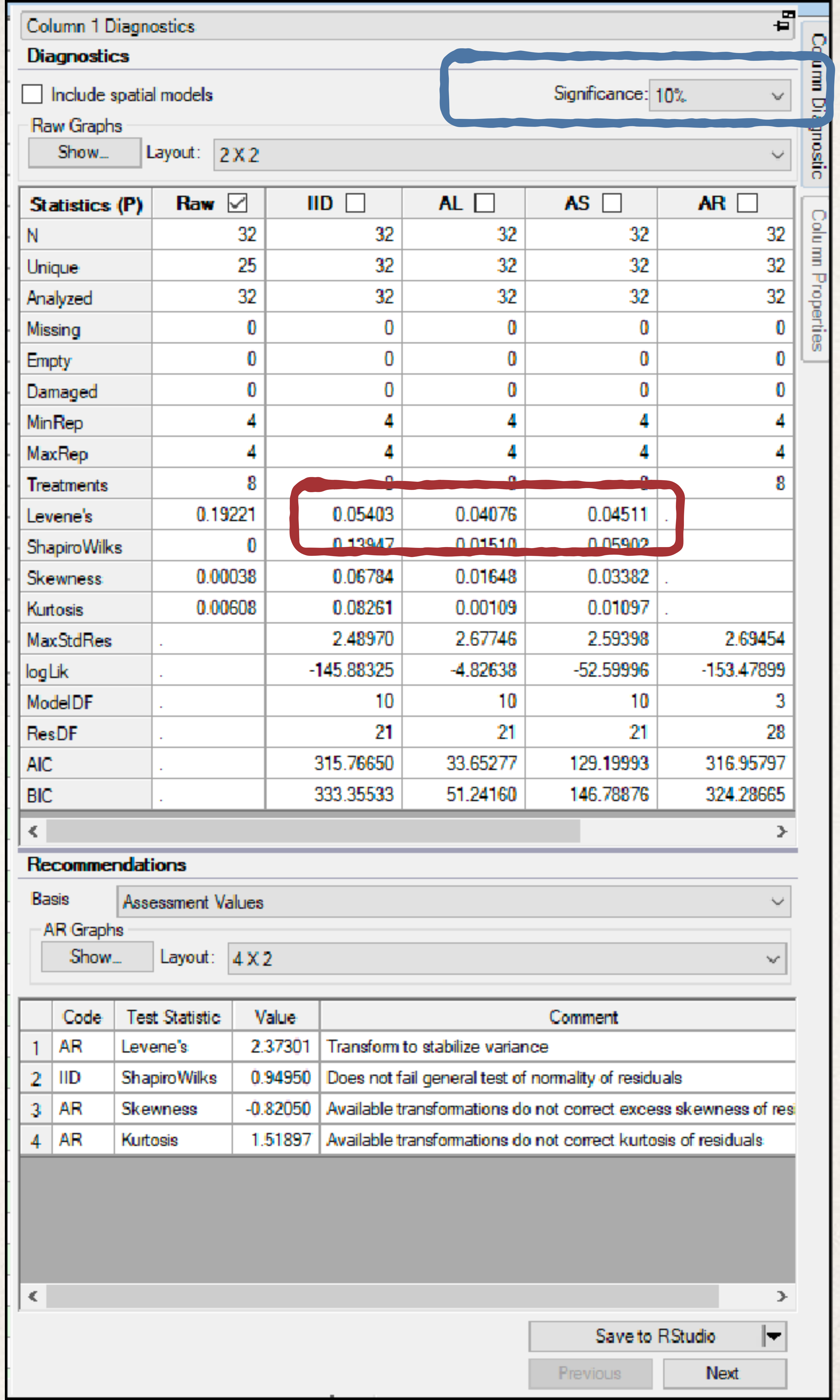
- These data do not fail Levene's test at a 5% significance level, but just barely.



New feature, not in ARM 2021.2

Cochran 1947 II

- If we change our threshold for significance, homogeneity of variance of these data are not improved by available transformations
- Data are outside the range for AA transformation


Column 1 Diagnostics

Diagnostics

Include spatial models — Significance: 10%

Raw Graphs — Show... — Layout: 2 X 2

Statistics (P)	Raw	IID	AL	AS	AR
N	32	32	32	32	32
Unique	25	32	32	32	32
Analyzed	32	32	32	32	32
Missing	0	0	0	0	0
Empty	0	0	0	0	0
Damaged	0	0	0	0	0
MinRep	4	4	4	4	4
MaxRep	4	4	4	4	4
Treatments	8	[illegible]	[illegible]	[illegible]	8
Levene's	0.19221	0.05403	0.04076	0.04511	.
ShapiroWilks	0	0.13947	0.01510	0.05902	
Skewness	0.00038	0.06784	0.01648	0.03382	.
Kurtosis	0.00608	0.08261	0.00109	0.01097	.
MaxStdRes	.	2.48970	2.67746	2.59398	2.69454
logLik	.	-145.88325	-4.82638	-52.59996	-153.47899
ModelDF	.	10	10	10	3
ResDF	.	21	21	21	28
AIC	.	315.76650	33.65277	129.19993	316.95797
BIC	.	333.35533	51.24160	146.78876	324.28665

Recommendations

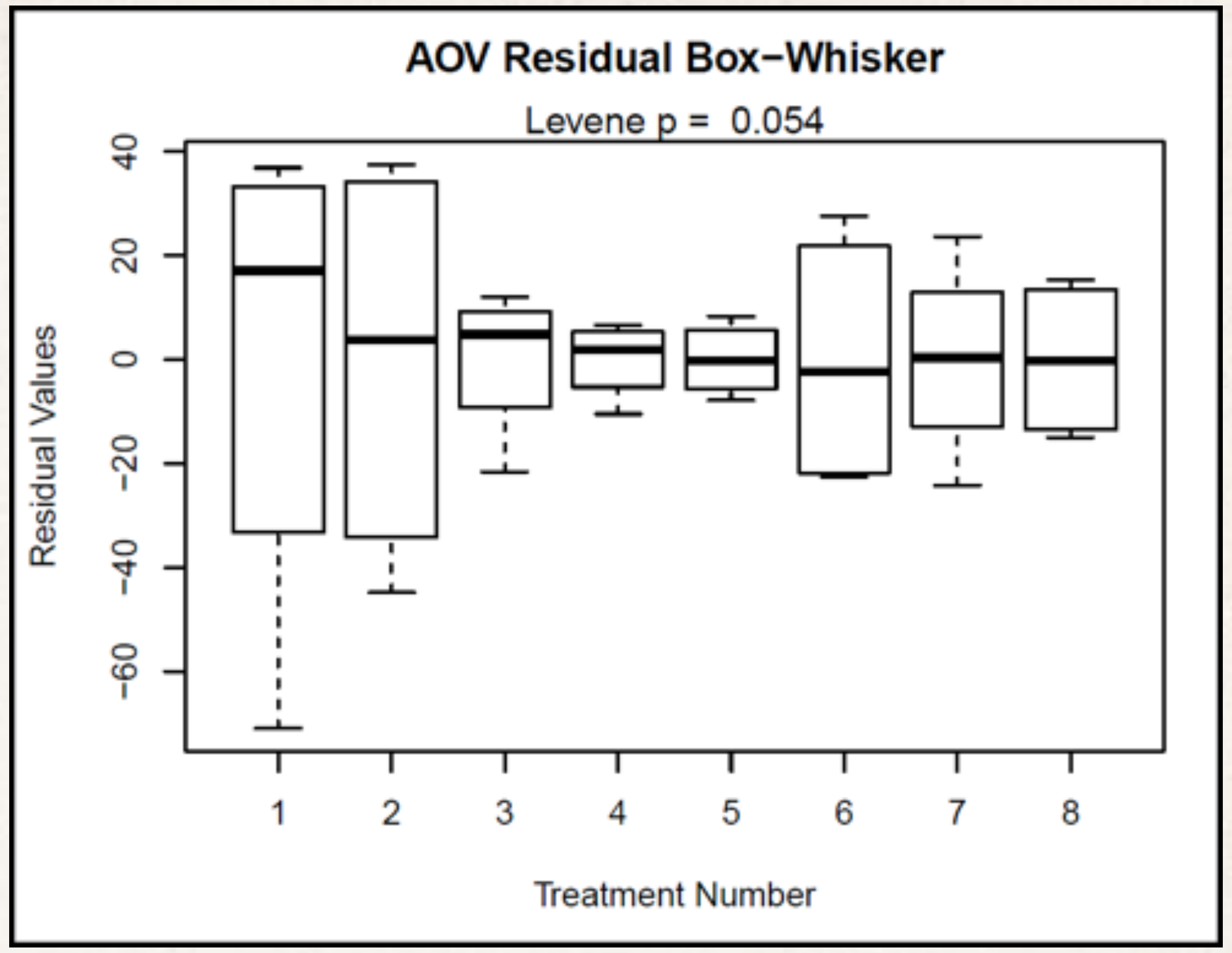
Basis: Assessment Values

AR Graphs — Show... — Layout: 4 X 2

	Code	Test Statistic	Value	Comment
1	AR	Levene's	2.37301	Transform to stabilize variance
2	IID	ShapiroWilks	0.94950	Does not fail general test of normality of residuals
3	AR	Skewness	-0.82050	Available transformations do not correct excess skewness of res
4	AR	Kurtosis	1.51897	Available transformations do not correct kurtosis of residuals

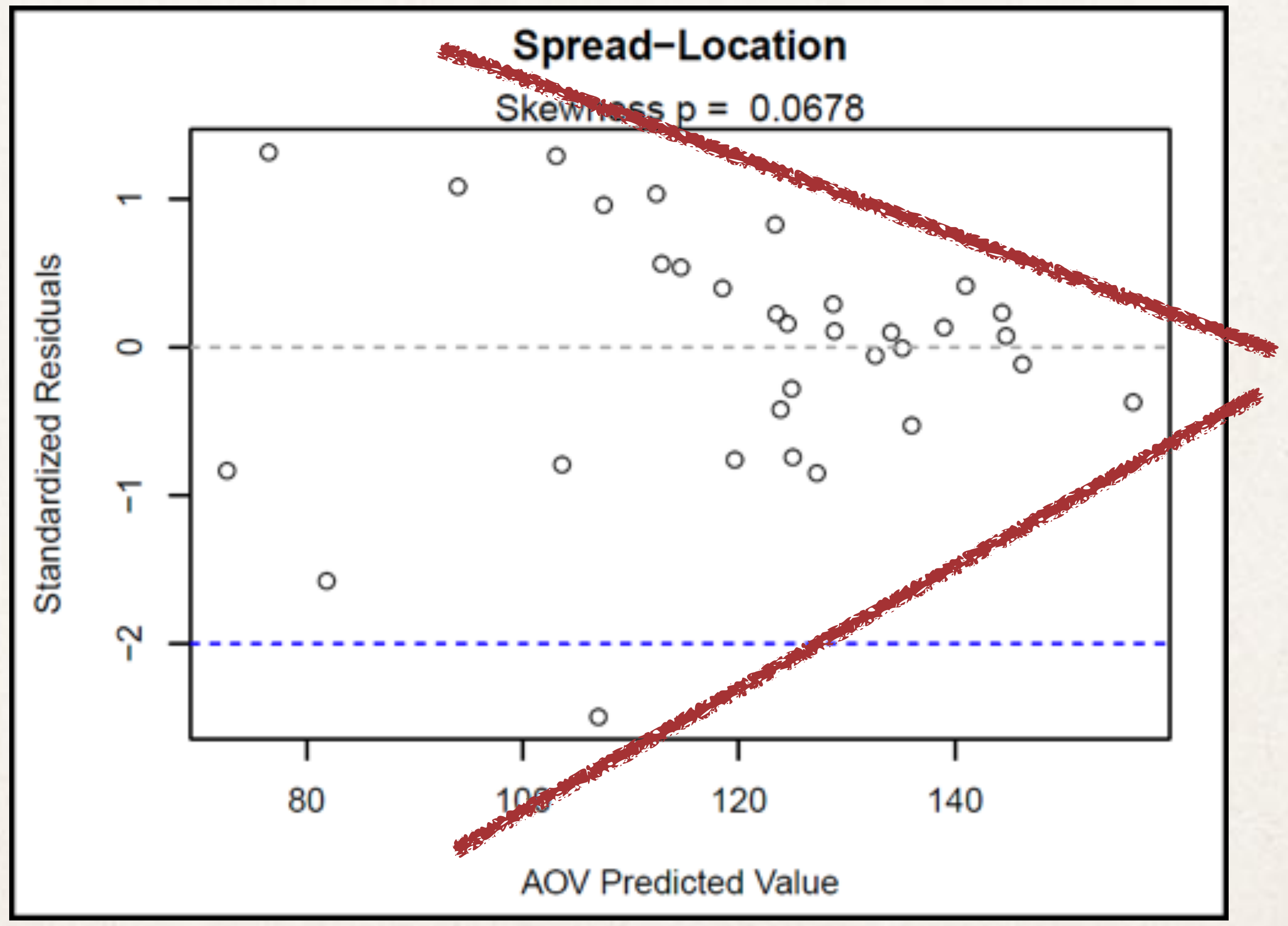
Save to RStudio — Previous — Next

Cochran 1947 II



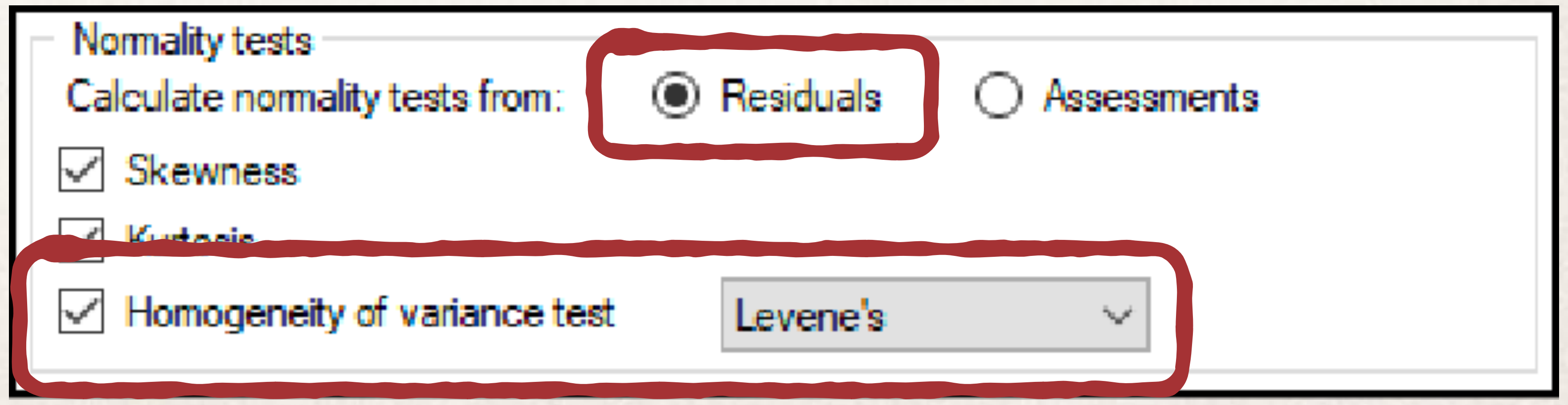
- Treatment variances tend to be larger with smaller treatment values, producing a wedge shape in the scale-location plot.

- This pattern of heterogeneity is not commonly corrected by transformations, so an analysis of ranks would be preferred.



The Partition of Error in Randomized Blocks

- O. Kempthorne and W. D. Barclay. The partition of error in randomized blocks. Journal of the American Statistical Association, 48(263):610–614, Sep 1953. https://www.jstor.org/stable/2281012
- When the error is heterogeneous, the usual procedure is to make a transformation which makes the error as homogeneous as possible. Additivity on the new scale is then assumed. …
- It was found that the verdict of heterogeneity of error based on Bartlett's test at the 5 per cent level, would be reached in 13.3 per cent of the samples. There is therefore a marked tendency to conclude that there is heterogeneity of error when in fact each of a complete set of normalized orthogonal comparisons is subject to the same error variance



- Take-away points:
  - Use Levene's test for heterogeneity
  - Use diagnostics on residuals, not on raw assessments.

The data in Table V provide an instance. The experiment was a 2^4 factorial, testing the effects of lime (L), fish manure (F) and artificial fertilizers (A). Lime was applied in the first year only; the other dressings were either applied in the first year only (1) or at a half rate every year (2). Two randomized blocks were laid out, the crop being pyrethrum, which forms an ingredient in many common insecti-

TABLE V

WEIGHTS OF DRY HEADS PER PLOT
(Unit, 10 grams)

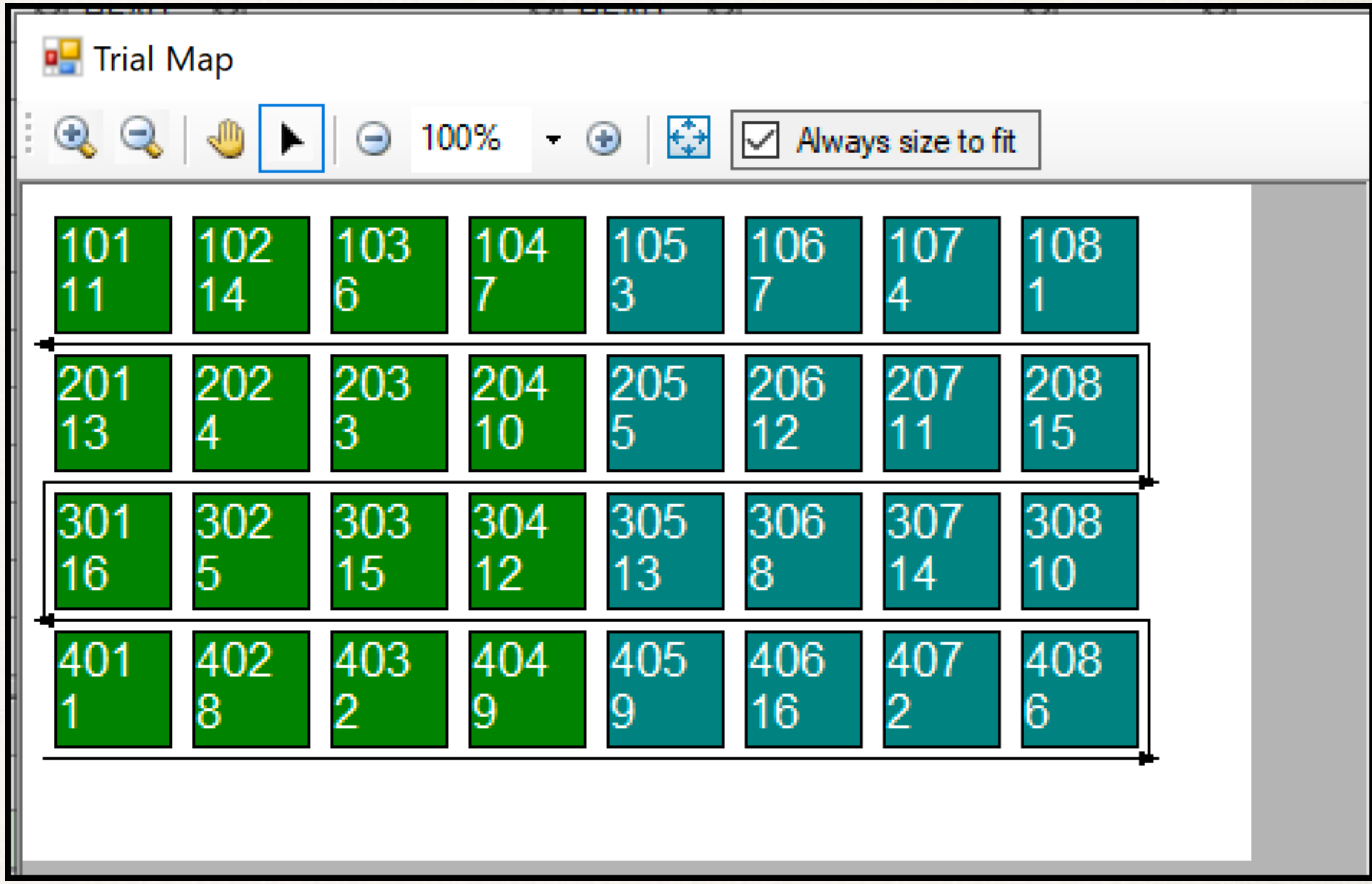
Block 1				Block 2			
LA1	LF2	F2	L1	A1	L1	A2	0
84	66	70	81	63	97	56	64
1	1	1	1	1	1	1	1
LF1	A2	A1	FA2	F1	LA2	LA1	LFA1
148	137	146	171	168	158	189	152
0	0	0	0	0	0	0	0
LFA2	F1	LFA1	LA2	LF1	L2	LF2	FA2
179	218	247	228	191	195	189	179
0	0	0	0	0	0	0	0
0	L2	0	FA1	FA1	LFA2	0	F2
124	166	177	153	133	145	141	130
0	0	0	0	0	0	0	0

Effects of Correlations Amongst the Errors

Cochran 1947 V.dat0

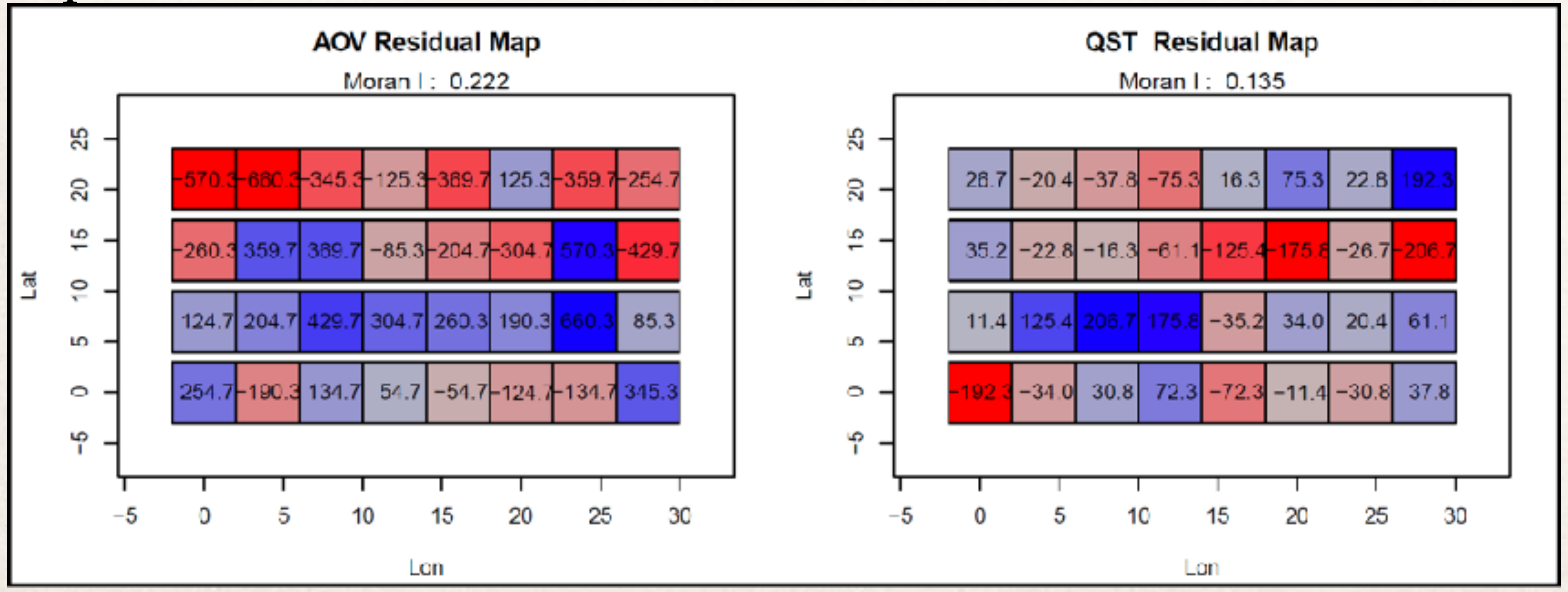
Cochran 1947 V

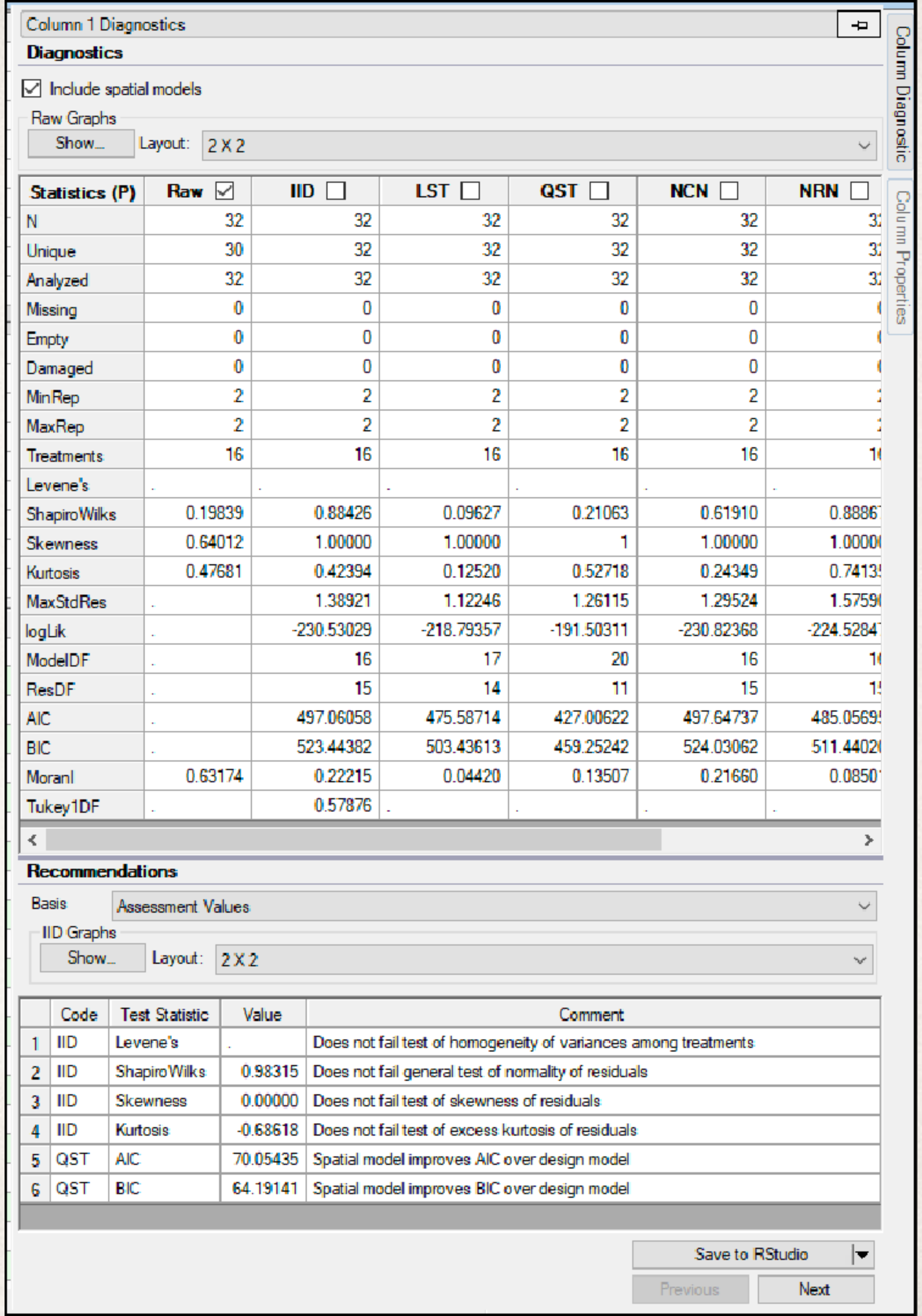
- The blocking pattern of this trial suggests spatial models may improve upon the RCB AOV.



Cochran 1947 V

- Spatial diagnostics includes Moran's I and Tukey's 1 d.f. test for non-additivity.
- These data do not fail the assumption of additivity, but Moran's I suggests spatial correlation among the experimental units
- A quadratic spatial model reduces the correlation among experimental units



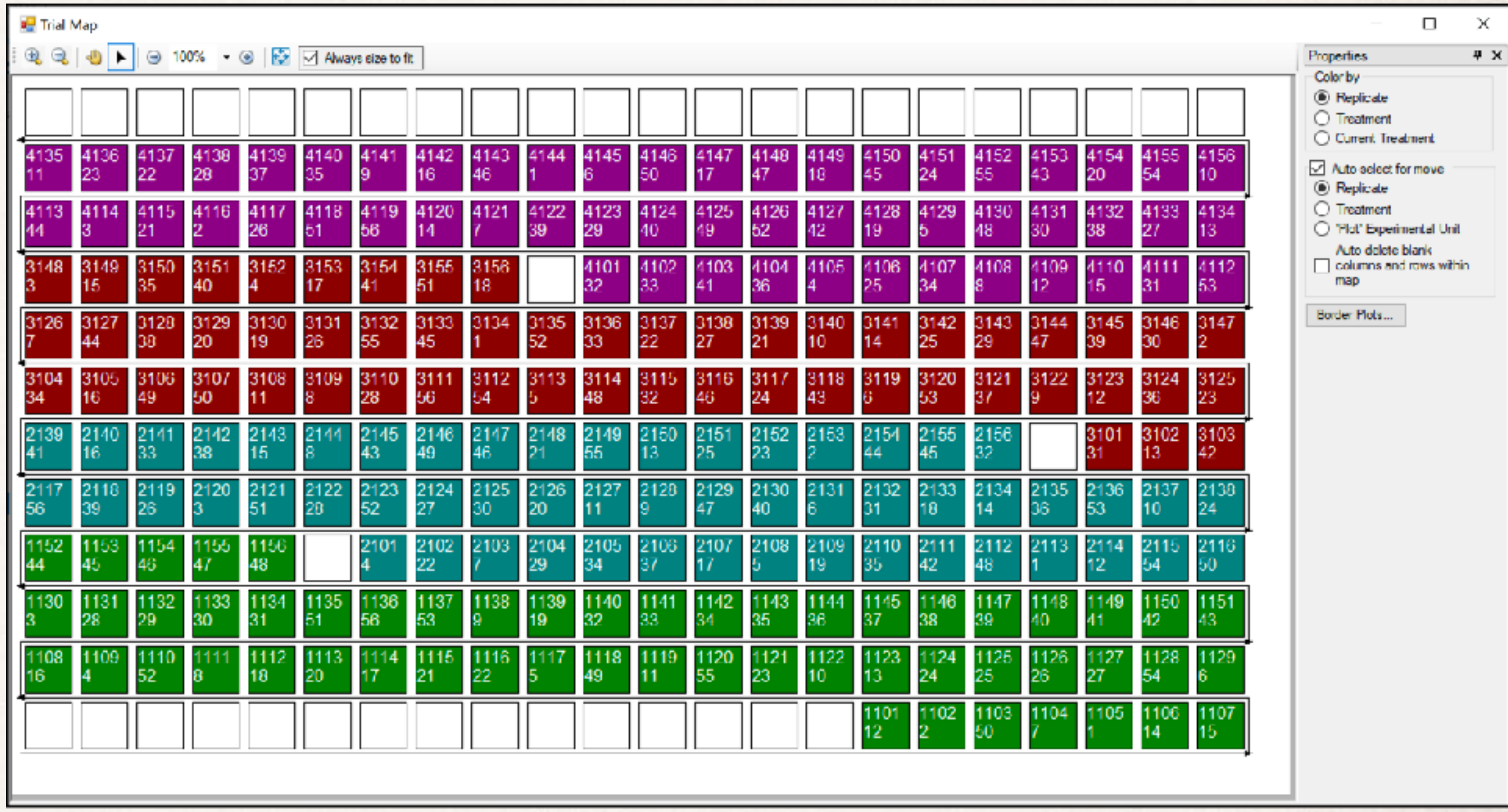


Statistics (P)	Raw	IID	LST	QST	NCN	NRN
N	32	32	32	32	32	3
Unique	30	32	32	32	32	3
Analyzed	32	32	32	32	32	3
Missing	0	0	0	0	0	
Empty	0	0	0	0	0	
Damaged	0	0	0	0	0	
MinRep	2	2	2	2	2	
MaxRep	2	2	2	2	2	
Treatments	16	16	16	16	16	1
Levene's	.	.	.	.	.	.
ShapiroWilks	0.19839	0.88426	0.09627	0.21063	0.61910	0.8886
Skewness	0.64012	1.00000	1.00000	1	1.00000	1.0000
Kurtosis	0.47681	0.42394	0.12520	0.52718	0.24349	0.7413
MaxStdRes	.	1.38921	1.12246	1.26115	1.29524	1.5759
logLik	.	-230.53029	-218.79357	-191.50311	-230.82368	-224.5284
ModelDF	.	16	17	20	16	1
ResDF	.	15	14	11	15	1
AIC	.	497.06058	475.58714	427.00622	497.64737	485.0569
BIC	.	523.44382	503.43613	459.25242	524.03062	511.4402
MoranI	0.63174	0.22215	0.04420	0.13507	0.21660	0.0850
Tukey1DF	.	0.57876	.	.	.	.

	Code	Test Statistic	Value	Comment
1	IID	Levene's	.	Does not fail test of homogeneity of variances among treatments
2	IID	ShapiroWilks	0.98315	Does not fail general test of normality of residuals
3	IID	Skewness	0.00000	Does not fail test of skewness of residuals
4	IID	Kurtosis	-0.68618	Does not fail test of excess kurtosis of residuals
5	QST	AIC	70.05435	Spatial model improves AIC over design model
6	QST	BIC	64.19141	Spatial model improves BIC over design model

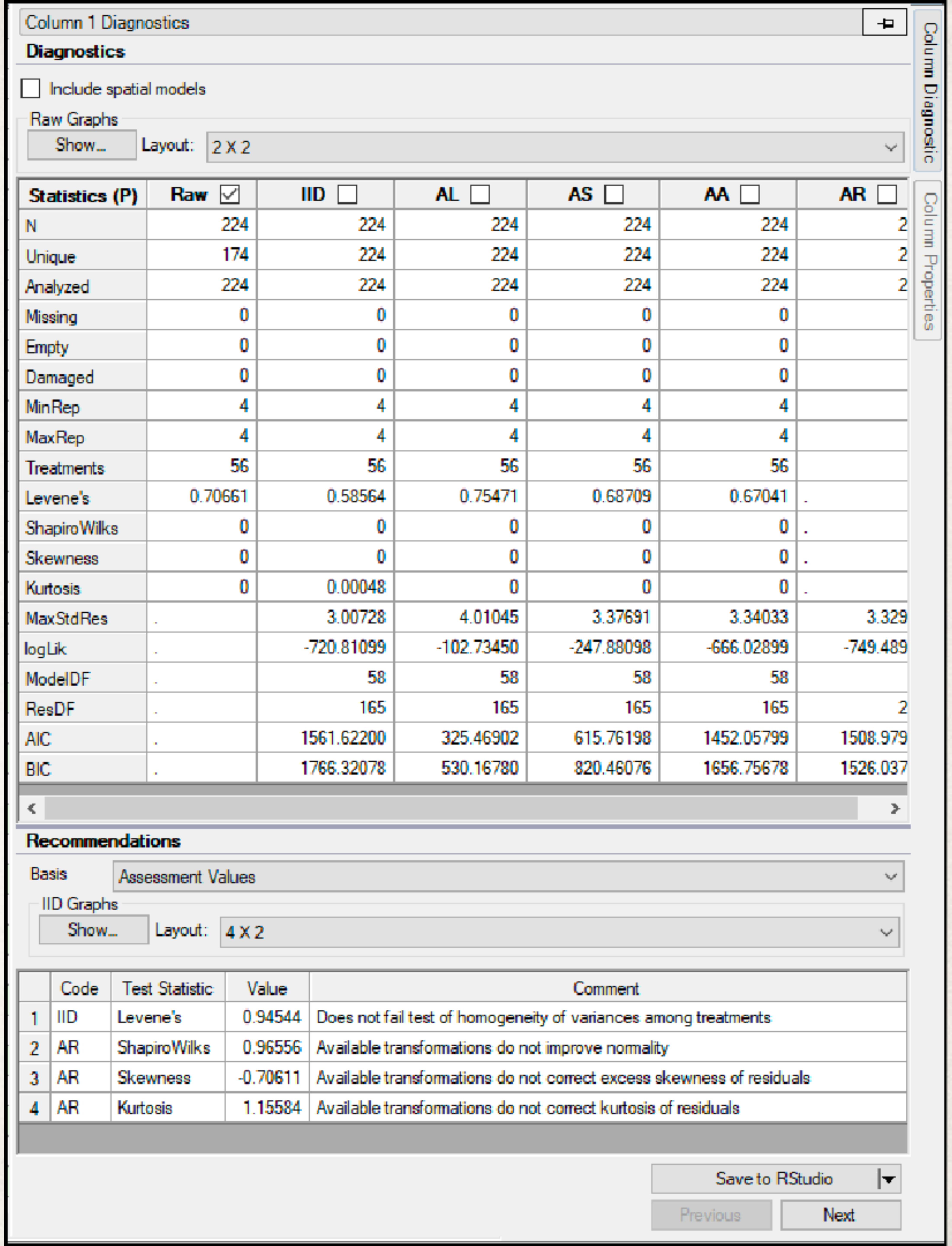
Little 1996

- Littell, R. C., Milliken, G. A., Stroup, W. W., & Wolfinger, R. D. (1996). SAS system for mixed models
- Data set 9.6.2, Wheat Yield
- This trial was not appropriately blocked, so is a common example for spatial analysis of agronomic trials.



Little 1996

- These data fail most tests of normality.



Statistics (P)	Raw ☑	IID ☐	AL ☐	AS ☐	AA ☐	AR ☐
N	224	224	224	224	224	2
Unique	174	224	224	224	224	2
Analyzed	224	224	224	224	224	2
Missing	0	0	0	0	0	
Empty	0	0	0	0	0	
Damaged	0	0	0	0	0	
MinRep	4	4	4	4	4	
MaxRep	4	4	4	4	4	
Treatments	56	56	56	56	56	
Levene's	0.70661	0.58564	0.75471	0.68709	0.67041	.
ShapiroWilks	0	0	0	0	0	.
Skewness	0	0	0	0	0	.
Kurtosis	0	0.00048	0	0	0	.
MaxStdRes	.	3.00728	4.01045	3.37691	3.34033	3.329
logLik	.	-720.81099	-102.73450	-247.88098	-666.02899	-749.489
ModelDF	.	58	58	58	58	
ResDF	.	165	165	165	165	2
AIC	.	1561.62200	325.46902	615.76198	1452.05799	1508.979
BIC	.	1766.32078	530.16780	820.46076	1656.75678	1526.037

	Code	Test Statistic	Value	Comment
1	IID	Levene's	0.94544	Does not fail test of homogeneity of variances among treatments
2	AR	ShapiroWilks	0.96556	Available transformations do not improve normality
3	AR	Skewness	-0.70611	Available transformations do not correct excess skewness of residuals
4	AR	Kurtosis	1.15584	Available transformations do not correct kurtosis of residuals

Little 1996

- Tukey's 1 d.f. test for non-additivity.
- Tests the assumption
  - The treatment effects and the environmental effects must be additive.

Statistics (P)	Raw	IID
N	224	224
Unique	174	224
Analyzed	224	224
Missing	0	0
Empty	0	0
Damaged	0	0
MinRep	4	4
MaxRep	4	4
Treatments	56	56
Levene's	0.70661	0.58564
ShapiroWilks	0	0
Skewness	0	0
Kurtosis	0	0.00048
MaxStdRes	.	3.00728
logLik	.	-720.81099
ModelDF	.	58
ResDF	.	165
AIC	.	1561.62200
BIC	.	1766.32078
MoranI	0.64278	0.44094
Tukey1DF	.	0.07086

Little 1996

- Normality is improved by spatial trend surface models. A cubic spatial trend seems to provide the best fit to the data.

Column 1 Diagnostics

Diagnostics

☑ Include spatial models

Raw Graphs

Show... Layout: 2 X 2

Statistics (P)	LST	QST	CST	NCN	NRN	NRCI
Missing	0	0	0	0	0	
Empty	0	0	0	0	0	
Damaged	0	0	0	0	0	
MinRep	4	4	4	4	4	
MaxRep	4	4	4	4	4	
Treatments	56	56	56	56	56	
Levene's	0.67807	0.16789	0.02880	0.34312	0.26274	(
ShapiroWilks	0.01906	0.85903	0.17134	0.04596	0.15261	(
Skewness	0.02573	0.43823	0.91259	0.03883	0.06488	(
Kurtosis	0.39468	0.84298	0.04285	0.00717	0.46147	(
MaxStdRes	2.47055	2.39379	2.72003	3.48727	2.40996	:
logLik	-677.31506	-638.74201	-622.19704	-672.94294	-667.14185	-65
ModelDF	57	60	64	56	56	
ResDF	166	163	159	167	167	
AIC	1472.63013	1401.48403	1376.39408	1461.88589	1450.28371	142
BIC	1673.91726	1613.00611	1601.56275	1659.76139	1648.15920	162
MoranI	0.40019	0.22366	0.14783	0.03022	-0.28259	-
Tukey1DF	.	.	.	.	.	.

Recommendations

Basis: Assessment Values

IID Graphs

Show... Layout: 4 X 2

	Code	Test Statistic	Value	Comment
1	IID	Levene's	0.94544	Does not fail test of homogeneity of variances among treatments
2	AR	ShapiroWilks	0.96556	Available transformations do not improve normality
3	QST	Skewness	-0.70611	Select action code to reduce skewness of residuals
4	LST	Kurtosis	1.15584	Select action code to reduce kurtosis of residuals
5	CST	AIC	185.22791	Spatial model improves AIC over design model
6	CST	BIC	164.75803	Spatial model improves BIC over design model

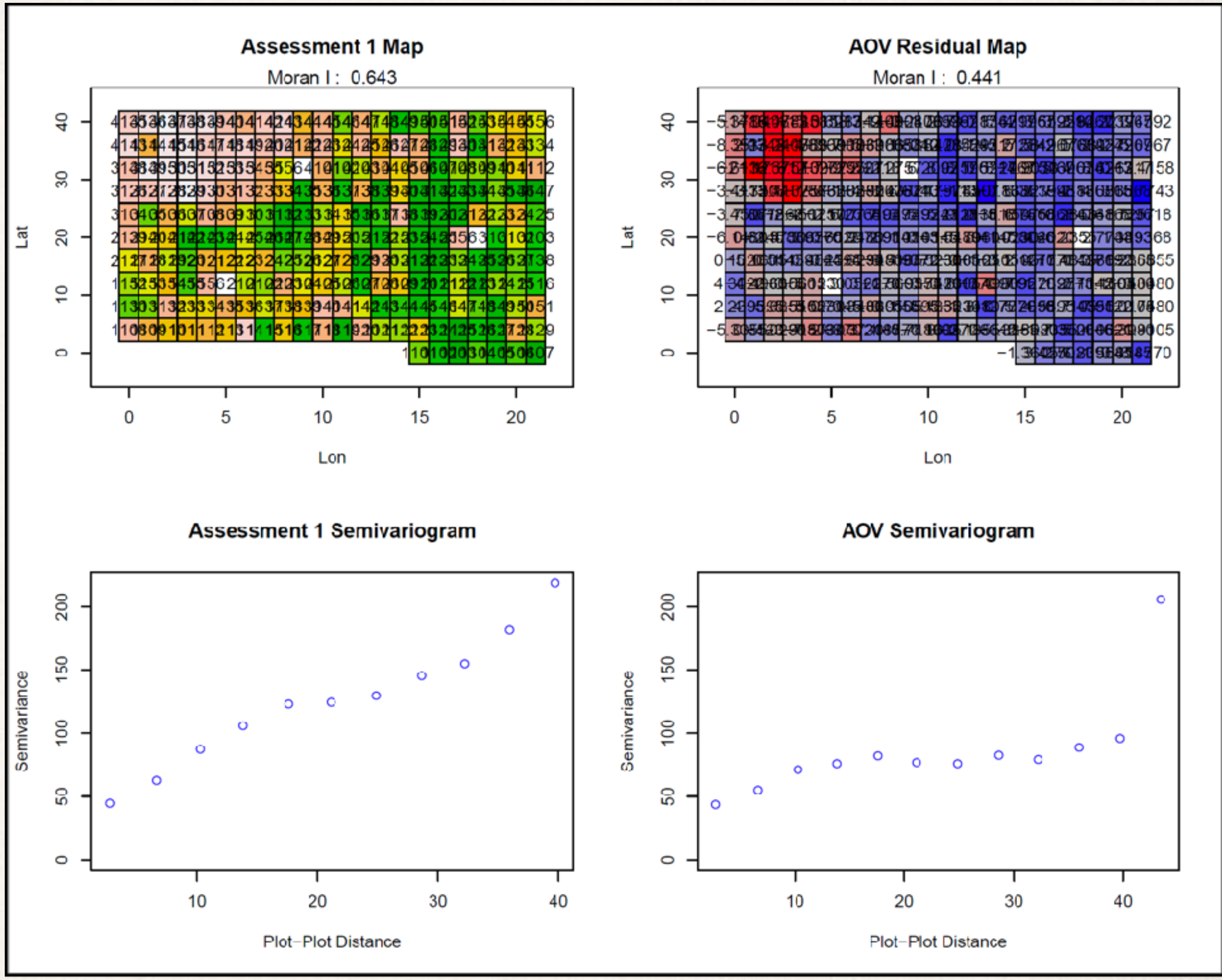
Save to RStudio

Previous Next

Column Diagnostic

Column Properties

Little 1996

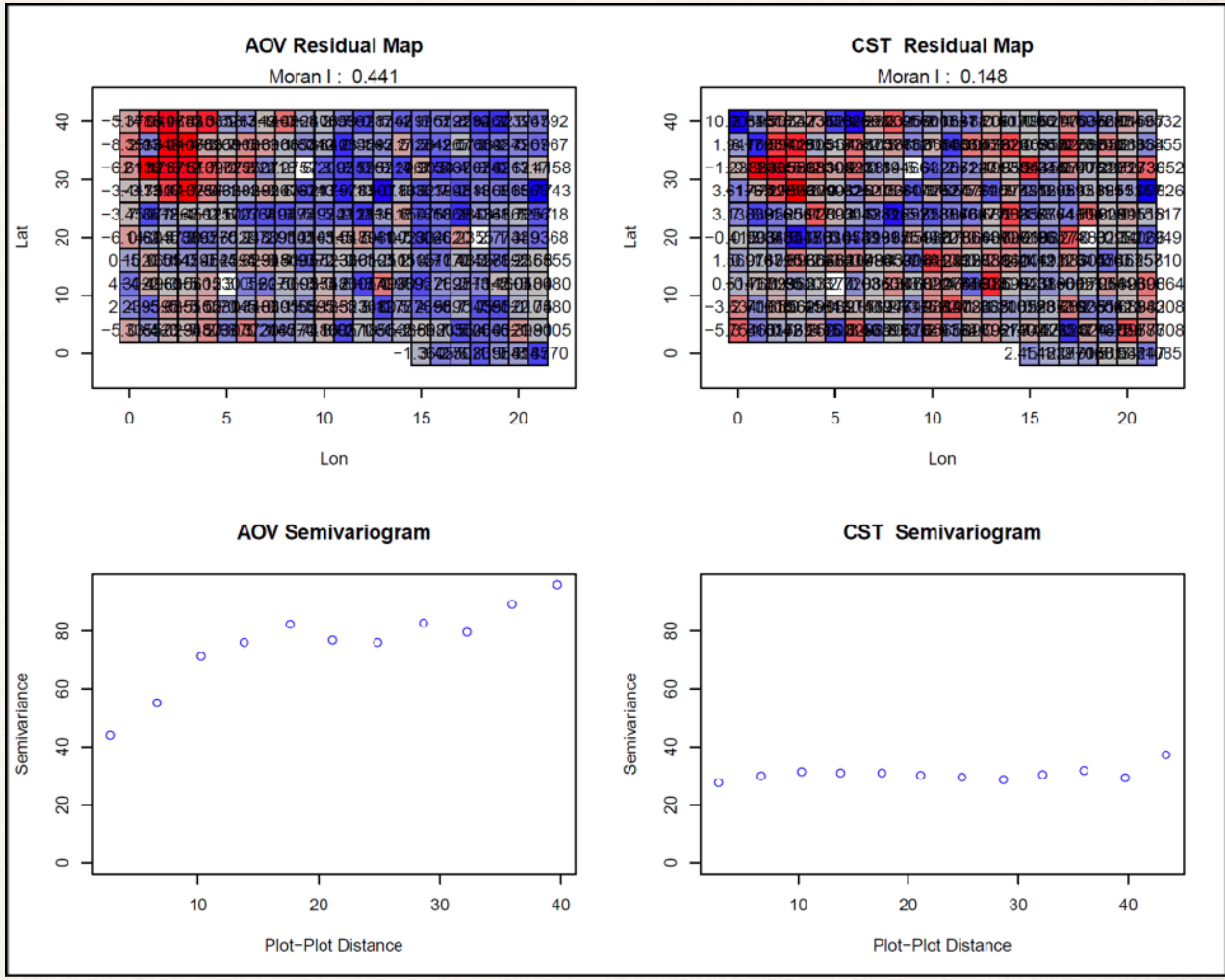


- The data have a spatial correlation that is not well captured by the blocks, so there is considerable spatial autocorrelation in the residuals, as measured by Moran's I
- Lack of independence in the experimental error is also suggested by the residual semivariogram.

Little 1996

- A cubic spatial trend model improves the tests of spatial dependence, although there is some degree of correlation remaining in the residuals.

- Ultimately, a non-parametric (rank based analysis) may be preferred.



The use of transformations

- M. S. Bartlett. Biometrics, 3(1):39–52, Mar 1947. https://www.jstor.org/stable/3001536
    - The Square Root Transformation
    - Logarithmic Transformations
    - The Inverse Sine or Angular Transformation

Theoretical Discussion.

- The purpose of this note is to summarize the transformations which have been used on raw statistical data, with particular reference to analysis of variance. For any such analysis the usual purpose of the transformation is to change the scale of the measurements in order to make the analysis more valid. Thus the conditions required for assessing accuracy in the ordinary unweighted analysis of variance include the important one of a constant residual or error variance, and if the variance tends to change with the mean level of the measurements, the variance will only be stabilized by a suitable change of scale.
    - *Valid* - If the reported p-value is, say 0.026, then the chance of a greater test statistic is close to 0.026

Theoretical Discussion

- If the form of the change of variance with mean level is known, this determines the type of transformation to use. Suppose we write
- $$\sigma_x^2 = f(m)$$
- where σ_x^2 is the variance on the original scale of measurements x with the mean of x equal to m

Theoretical Discussion

- Then for any function $g(x)$ we have approximately
- $$\sigma_g^2 = \left(\frac{dg}{dm}\right)^2 f(m)$$
- so that if σ_g^2 is to be constant, C^2 say, we must have
- $$g(m) = \int \frac{C dm}{\sqrt{f(m)}}$$
- For example, if the standard the mean level m, we have $f(m)$ proportional to m^2, and $g(m)$ proportional to $\log m$; i.e., we should use the logarithmic scale.

Theoretical Discussion.

- However, a constant variance is not the only condition we seek, and precautions are still necessary when using analysis of variance with the transformed variate. In the ideal case (cf. Reference 6 at the end of this paper),
  - (a) The variance of the transformed variate should be unaffected by changes in the mean level (this is taken to be the primary purpose of the transformations of sections 2, 3, and 4).
  - (b) The transformed variate should be normally distributed.
  - (c) The transformed scale should be one for which an arithmetic average is an efficient estimate of the true mean level for any particular group of measurements.
  - (d) The transformed scale should be one for which real effects are linear and additive.
- Although these conditions are to some extent related [for example, (a) and (b) and (d) together imply (c)], we obviously cannot necessarily expect to arrange for conditions (b), (c), and (d) to be satisfied if our scale has already been fixed by condition (a).

Square root transformation

Bartlett 1947 II

- When statistical data consist of integers, i.e., whole numbers, such as number of bacterial colonies in a plate count, or number of plants in a given area, homogeneous conditions will often lead to variation in these numbers x following the Poisson distribution

- Since for such a distribution the variance is exactly equal to the mean, we readily obtain from our general equation in section 1, that to stabilize the variance we must work on the square-root scale. In many cases, the variance of the observations (on scale y) is proportional to (that is, a function f of) the expected value or mean (m) of the observations.

TABLE II

PLAN SHOWING LAYOUT OF EXPERIMENT ON OATS, AND NUMBERS OF POPPIES (IN 3¾ SQ. FT. AREAS)

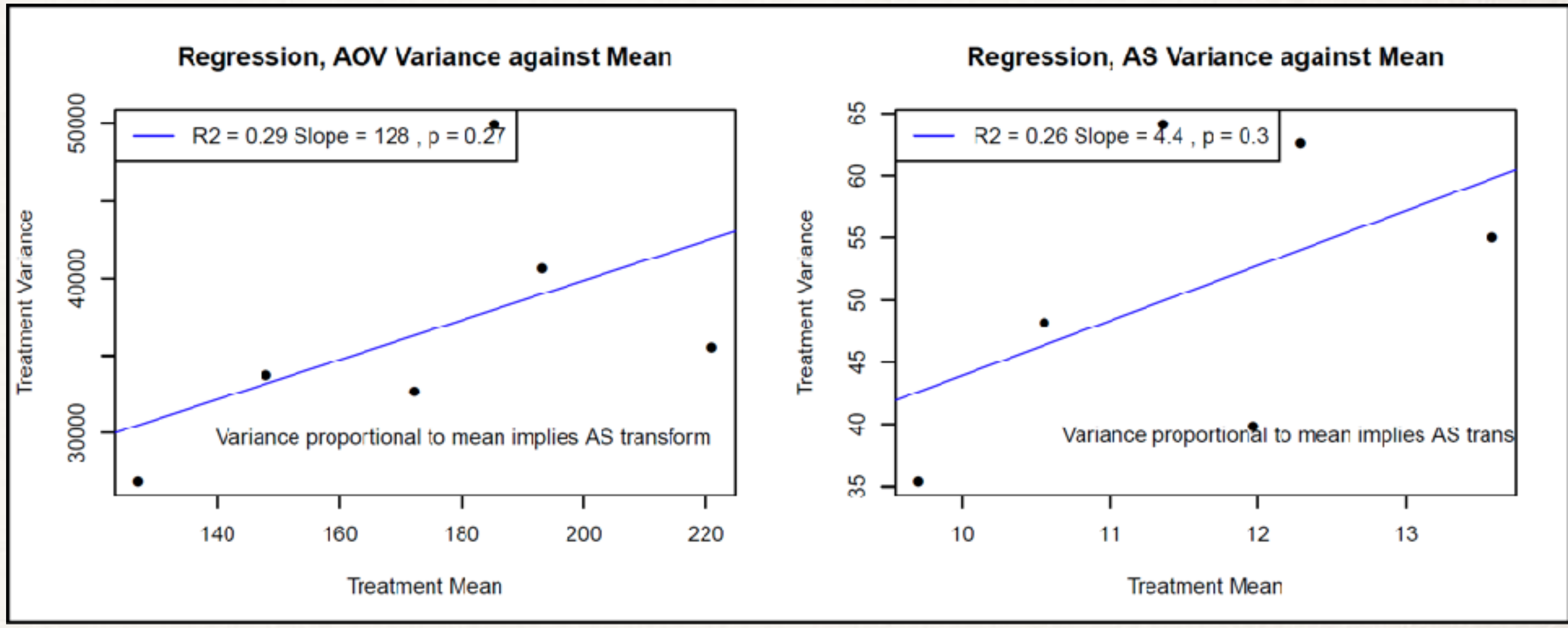
Block A	(1)* 438	(4) 17	(2) 538	(5) 18	(3) 77	(6) 115
B	(3) 61	(2) 422	(6) 57	(1)* 442	(5) 26	(4) 31
C	(5) 77	(3) 157	(4) 87	(6) 100	(2) 377	(1)* 319
D	(2) 315	(1)* 380	(5) 20	(3) 52	(4) 16	(6) 45

* Control.

Bartlett 1947 II

- In practice we often use an analysis of variance for data of the above integer type because we suspect heterogeneity of one kind or another to be present, especially if our data have been collected under field conditions. We do not then need to assume Poisson variation, but will still transform to the square-root scale if the variation of $\sqrt{x}$ appears stable.

- The following set of data (quoted from Reference 1; see also Reference 13), representing weed-infestation counts in one of a series of experiments on weed control in cereals, is an example of stability of variance on the square-root scale, even when the level of variability is far higher than expected on the assumptions of a Poisson distribution for x.



New feature, not in ARM 2021.2

Bartlett 1947 II

- As with many historical data, this is not the best example of an analysis that is best served by the theoretical transformation

Column 1 Diagnostics

Diagnostics

Include spatial models

Graphs

Show... Layout: 2 X 2

Statistics (P)	Raw	IID	AS	AR
N	24	24	24	24
Unique	23	24	24	23
Analyzed	24	24	24	24
Missing	0	0	0	0
Empty	0	0	0	0
Damaged	0	0	0	0
MinRep	4	4	4	4
MaxRep	4	4	4	4
Treatments	6	6	6	6
Levene's	0.95157	0.93732	0.89648	.
ShapiroWilks	0.00037	0.00189	0.01195	.
Skewness	0.11157	0.14293	0.25432	.
Kurtosis	0.33637	0.31351	0.25160	.
MaxStdRes	.	1.55935	1.45814	1.82924
logLik	.	-156.68965	-77.75767	-157.09059
ModelDF	.	8	8	3
ResDF	.	15	15	20
AIC	.	333.37930	175.51533	324.18118
BIC	.	345.15984	187.29587	330.07145

Recommendations

Basis Rating Type : COUNT

IID Graphs

Show... Layout: 2 X 2

	Code	Test Statistic	Value	Comment
1	IID	Levene's	0.24425	Does not fail test of homogeneity of variances among treatments
2	AR	ShapiroWilks	0.84659	Available transformations do not improve normality
3	IID	Skewness	0.76034	Does not fail test of skewness of residuals
4	IID	Kurtosis	-1.00383	Does not fail test of excess kurtosis of residuals

Save to RStudio

Previous Next

Bartlett 1936b VI

- M. S. Bartlett. Some notes on insecticide tests in the laboratory and in the field. Supplement to the Journal of the Royal Statistical Society, 3(2):185–194, 1936. https://www.jstor.org/stable/2983670
- A better example when theory matches practice.

TABLE VI.

Leatherjacket Counts.

		1 (Control).	2 (Control).	3.	4.	5.	6.
Block I	(i)	33	30	8	12	6	17
	(ii)	59	36	11	17	10	8
II	(i)	36	23	15	6	4	3
	(ii)	24	23	20	4	7	2
III	(i)	19	42	10	12	4	6
	(ii)	27	39	7	10	12	3
IV	(i)	71	39	17	5	5	1
	(ii)	49	20	26	8	5	1
V	(i)	22	42	14	12	2	2
	(ii)	27	22	11	12	6	5
VI	(i)	84	23	22	16	17	6
	(ii)	50	37	30	4	11	5

A square-root analysis of the totals for the two sample counts on the treated plots completes the following summary of results.

Bartlett 1936b VI

- Unconstrained diagnostics recommend AL

Recommendations

Basis: Assessment Values

AL Graphs

Show... Layout: 4 X 2

	Code	Test Statistic	Value	Comment
1	AL	Levene's	4.28247	Transform to stabilize variance
2	IID	Shapiro Wilks	0.96379	Does not fail general test of normality of residuals
3	IID	Skewness	0.29576	Does not fail test of skewness of residuals
4	IID	Kurtosis	0.91368	Does not fail test of excess kurtosis of residuals

- When Rating Type is selected, AS is recommended

Recommendations

Basis: Rating Type : COUINS

AS Graphs

Show... Layout: 4 X 2

	Code	Test Statistic	Value	Comment
1	AS	Levene's	4.28247	Transform to stabilize variance
2	IID	Shapiro Wilks	0.96379	Does not fail general test of normality of residuals
3	IID	Skewness	0.29576	Does not fail test of skewness of residuals
4	IID	Kurtosis	0.91368	Does not fail test of excess kurtosis of residuals

TABLE VI.

Leatherjacket Counts.

	1 (Control).	2 (Control).	3.	4.	5.	6.
Block I (i)	33	30	8	12	6	17
(ii)	59	36	11	17	10	8
II (i)	36	23	15	6	4	3
(ii)	24	23	20	4	7	2
III (i)	19	42	10	12	4	6
(ii)	27	39	7	10	12	3
IV (i)	71	39	17	5	5	1
(ii)	49	20	26	8	5	1
V (i)	22	42	14	12	2	2
(ii)	27	22	11	12	6	5
VI (i)	84	23	22	16	17	6
(ii)	50	37	30	4	11	5

A square-root analysis of the totals for the two sample counts on the treated plots completes the following summary of results.

Bartlett 1936 II

- M. S. Bartlett. The square root transformation in analysis of variance. Supplement to the Journal of the Royal Statistical Society, 3(1), 68-78 1936. https://www.jstor.org/stable/2983678

TABLE II.

Control of Cockchafer Larvæ.

	A		*B*		*C*		*D*		*E*		*F*		*G*		*H*	
	a	*b*	*a*	*b*	*a*	*b*	*a*	*b*	*a*	*b*	*a*	*b*	*a*	*b*	*a*	*b*
1	13	28	29	61	5	7	5	14	0	3	1	7	1	10	4	13
2	16	12	12	49	4	2	12	5	2	3	1	6	3	5	4	11
3	13	40	23	48	4	4	1	14	2	2	1	7	1	8	7	10
4	20	31	15	44	1	5	5	9	2	7	3	7	0	3	3	12
5	16	22	17	45	2	2	3	8	0	0	5	4	1	6	1	8

The experiment consisted of five treatments in eight randomized blocks; the division of the larvæ into two groups demonstrated the differential effects of the treatments according to age. Unfortunately the count of the larvæ in the whole plot as in blocks *A* and *B* was not continued for reasons of time, and sample quadrat counts over a quarter of the plot were carried out in blocks *C* to *H*. This fact was, however, ignored in the analysis (of $\sqrt{x + \frac{1}{2}}$), for reasons given below, the figures for the sample counts in blocks *C* to *H* being recorded directly in Table II without adjustment. The standard errors were 0·50 for *a*, and 0·59 for *b*, these figures suggesting fairly homogeneous random material, in spite of significant differences between treatments and between blocks for *a* (apart of course from the artificial block difference *A*, *B* versus *C*, *D*, *E*, *F*, *G*, *H*).

Bartlett 1936 II

- These data fail tests of normality; with two different corrections recommended

Recommendations

Basis: Assessment Values

IID Graphs

Show... Layout: 4 X 2

	Code	Test Statistic	Value	Comment
1	IID	Levene's	0.43195	Does not fail test of homogeneity of variances among treatments
2	AS	ShapiroWilks	0.92413	Analyze transformed assessments to improve normality of residuals
3	IID	Skewness	0.14185	Does not fail test of skewness of residuals
4	AA	Kurtosis	3.64067	Select action code to reduce kurtosis of residuals

- When Rating Type is selected, the AS transformation is recommended

Recommendations

Basis: Rating Type : COUINS

IID Graphs

Show... Layout: 4 X 2

	Code	Test Statistic	Value	Comment
1	IID	Levene's	0.43195	Does not fail test of homogeneity of variances among treatments
2	AS	ShapiroWilks	0.92413	Analyze transformed assessments to improve normality of residuals
3	IID	Skewness	0.14185	Does not fail test of skewness of residuals
4	AS	Kurtosis	3.64067	Select action code to reduce kurtosis of residuals

TABLE II.

Control of Cockchafer Larvæ.

	A		*B*		*C*		*D*		*E*		*F*		*G*		*H*	
	a	*b*	*a*	*b*	*a*	*b*	*a*	*b*	*a*	*b*	*a*	*b*	*a*	*b*	*a*	*b*
1	13	28	29	61	5	7	5	14	0	3	1	7	1	10	4	13
2	16	12	12	49	4	2	12	5	2	3	1	6	3	5	4	11
3	13	40	23	48	4	4	1	14	2	2	1	7	1	8	7	10
4	20	31	15	44	1	5	5	9	2	7	3	7	0	3	3	12
5	16	22	17	45	2	2	3	8	0	0	5	4	1	6	1	8

The experiment consisted of five treatments in eight randomized blocks; the division of the larvæ into two groups demonstrated the differential effects of the treatments according to age. Unfortunately the count of the larvæ in the whole plot as in blocks *A* and *B* was not continued for reasons of time, and sample quadrat counts over a quarter of the plot were carried out in blocks *C* to *H*. This fact was, however, ignored in the analysis (of $\sqrt{x+\frac{1}{2}}$), for reasons given below, the figures for the sample counts in blocks *C* to *H* being recorded directly in Table II without adjustment. The standard errors were 0·50 for *a*, and 0·59 for *b*, these figures suggesting fairly homogeneous random material, in spite of significant differences between treatments and between blocks for *a* (apart of course from the artificial block difference *A*, *B* versus *C*, *D*, *E*, *F*, *G*, *H*).

Logarithmic Transformation

Bartlett 1947 III

- The stability of variance on the square-root scale in the case of the series of weed-control experiments was rather unexpected; since, if considerable heterogeneity in numbers is present, the variance is often found still to be correlated with the mean level on a square-root scale, and may only be stabilized if transformation is made to the logarithmic scale. The natural explanation of a variance greater than the mean is that the mean level itself fluctuates, so that

$$\sigma_x^2 = m + \sigma_m^2$$

For biological populations, increases in numbers are often proportional to the numbers already present, giving rise to variations in mean from place to place themselves proportional to the mean. This illustrates how σ_m^2 might be proportional to m^2, so that we might expect

$$\sigma_x^2 = m + \lambda^2 m^2$$

For λ large, or m large, this variance law implies the logarithmic transformation.

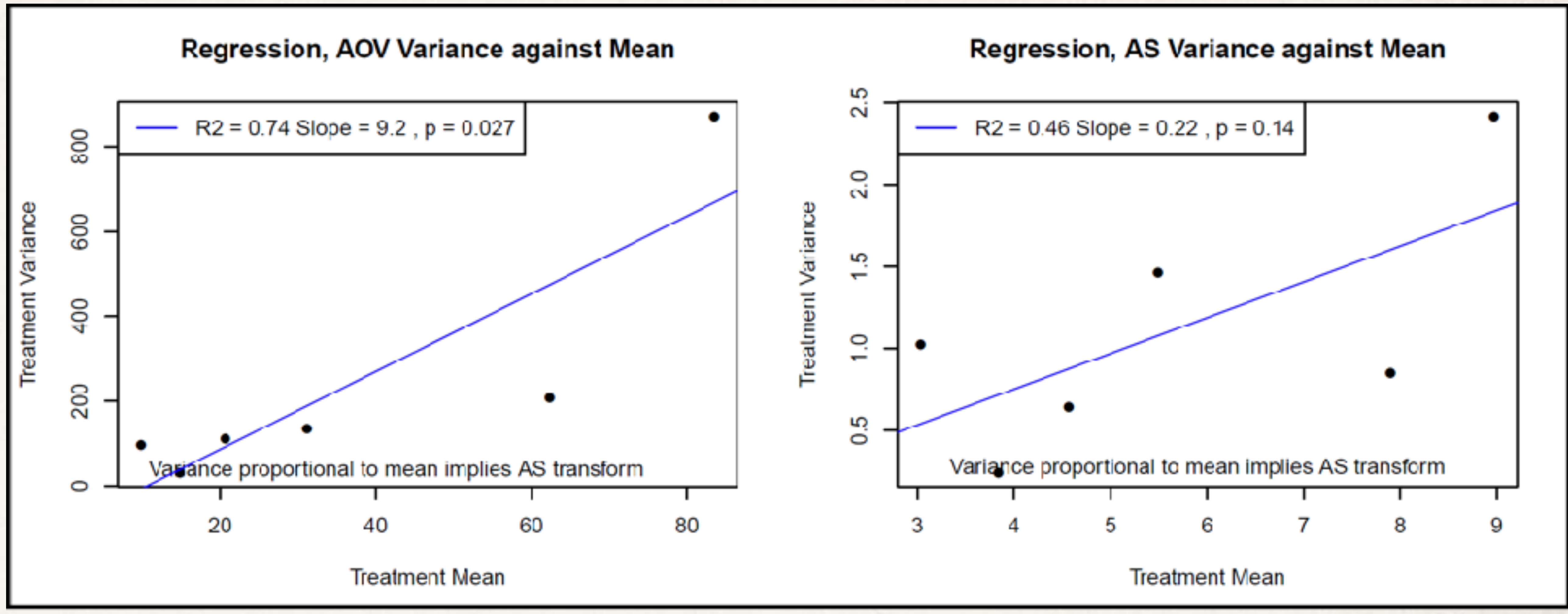
TABLE III

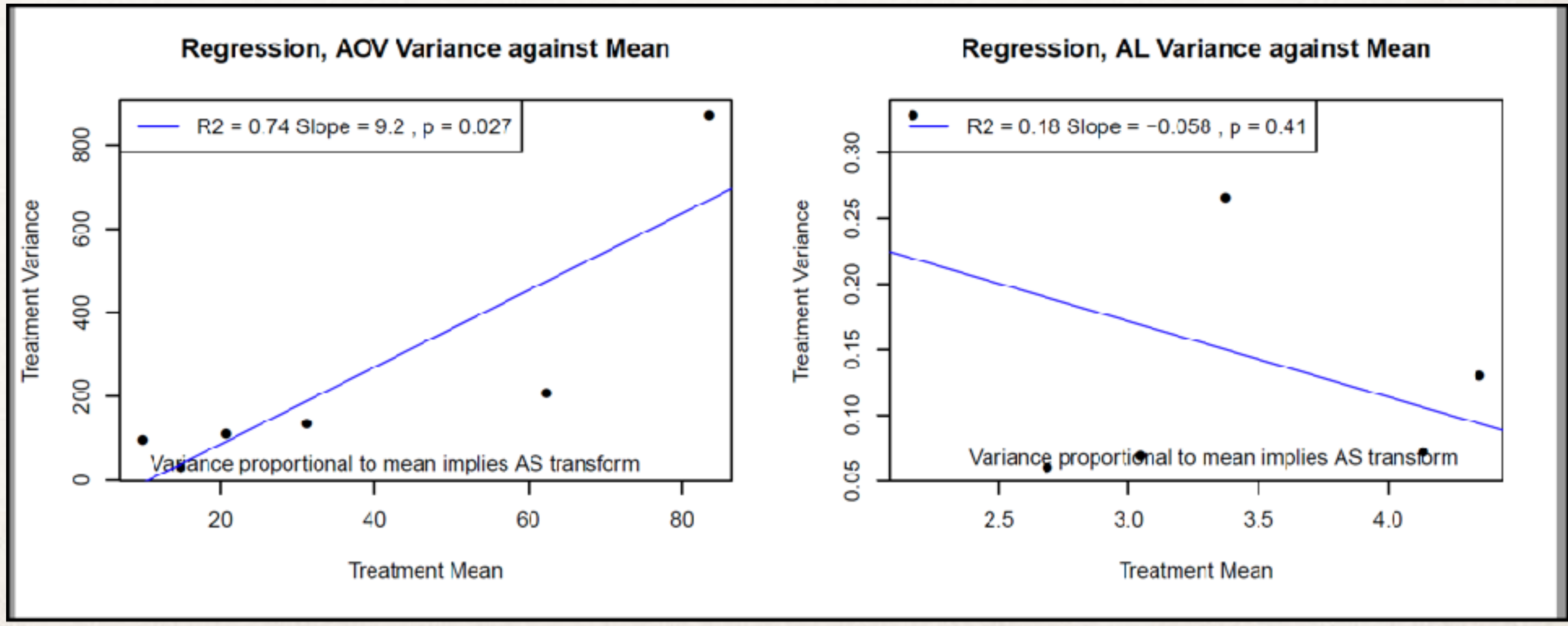
LEATHERJACKET COUNTS

Treatment	1 (Control)	2 (Control)	3	4	5	6
Block I	92	66	19	29	16	25
II	60	46	35	16	11	5
III	46	81	13	22	16	9
IV	120	59	43	13	10	2
V	49	62	25	24	8	7
VI	134	60	52	20	28	11

Bartlett 1947 III

- The stability of variance on the square-root scale in the case of the series of weed-control experiments was rather unexpected; since, **if considerable heterogeneity in numbers is present, the variance is often found still to be correlated with the mean level on a square-root scale, and may only be stabilized if transformation is made to the logarithmic scale.** The natural explanation of a variance greater than the mean is that the mean level itself fluctuates, so that
$$\sigma_x^2 = m + \sigma_m^2$$
For biological populations, increases in numbers are often proportional to the numbers already present, giving rise to variations in mean from place to place themselves proportional to the mean. This illustrates how σ_m^2 might be proportional to m^2, so that we might expect
$$\sigma_x^2 = m + \lambda^2 m^2$$
For λ large, or m large, this variance law implies the logarithmic transformation.





New feature, not in ARM 2021.2

Bartlett 1947 III

- AL and AS both improve Levene's test of heterogeneity
- AA is not included because these data contain values outside the limit of percentages (>100)
-

Recommendations

Basis: Assessment Values

AL Graphs — Show... Layout: 4 X 2

	Code	Test Statistic	Value	Comment
1	AL	Levene's	3.94649	Transform to stabilize variance
2	IID	ShapiroWilks	0.95693	Does not fail general test of normality of residuals
3	IID	Skewness	0.31298	Does not fail test of skewness of residuals
4	IID	Kurtosis	0.79954	Does not fail test of excess kurtosis of residuals

Column 1 Diagnostics

Diagnostics

☐ Include spatial models

Raw Graphs — Show... Layout: 2 X 2

Statistics (P) ☑		IID ☐	AL ☐	AS ☐	AR ☐
N	36	36	36	36	36
Unique	29	36	36	36	36
Analyzed	36	36	36	36	36
Missing	0	0	0	0	0
Empty	0	0	0	0	0
Damaged	0	0	0	0	0
MinRep	6	6	6	6	6
MaxRep	6	6	6	6	6
Treatments	6	6	6	6	6
Levene's	0	0.00721	0.51827	0.23840	.
ShapiroWilks	024	0.17256	0.99336	0.74259	.
Skewness	131	0.44893	0.69951	0.88794	.
Kurtosis	248	0.32427	0.59387	0.42456	.
MaxStdRes		2.15288	2.11882	1.74160	2.49704
logLik		-146.63502	-14.19384	-49.64171	-174.00674
ModelDF		10	10	10	5
ResDF		25	25	25	30
AIC		317.27004	52.38767	123.28342	362.01348
BIC		336.27227	71.38990	142.28565	373.09811

Column Diagnostic

Column Properties

Take-away

- On strictly theoretical basis, the square root transformation will frequently be preferred to logarithmic transformation, particularly when the data are counts
    - Count data may be expected to follow a Poisson distribution, and the square root transform is the natural complement to the Poisson distribution.
- However, it is common for biological data to show variances proportional to the mean, and these phenomena will be more likely corrected with logarithmic transformation.

Arc sine transformation

Bartlett 1947 V

- The inverse sine square-root transformation
$$g(x) = Sin^{-1}\sqrt{x}$$
bears the same relation to estimated probabilities or proportions x with binomial variance $p(1-p)/n$, where n is the number of individuals in the sample, as the square-root transformation does to a Poisson variate.

TABLE V

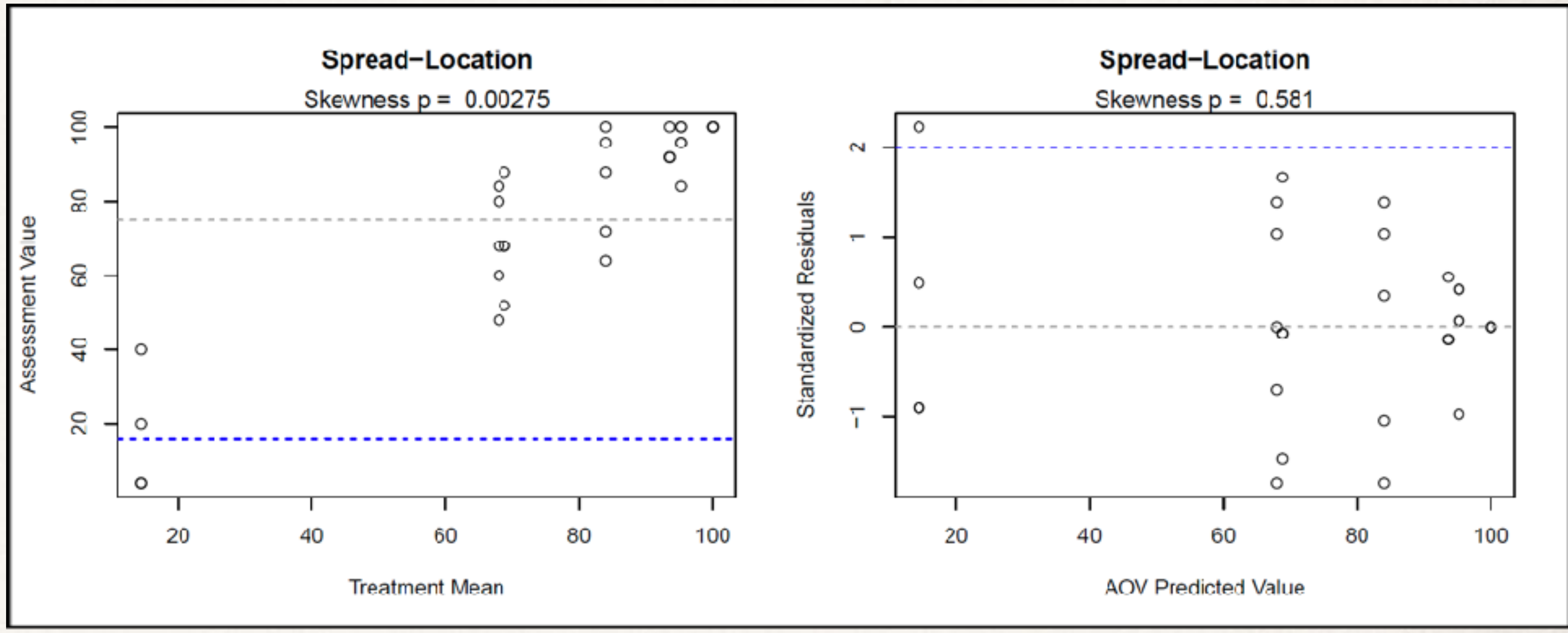
NUMBER OF DEAD FLIES

Treatments	A	B	C	D	E	F	G*
1	24	(25)	17	17	18	23	1
2	25	(25)	15	17	25	25	1
3	24	(25)	12	17	24	23	1
4	21	(25)	20	22	16	23	10
5	25	(25)	21	13	22	23	4, 6

* Control (with one extra replication).

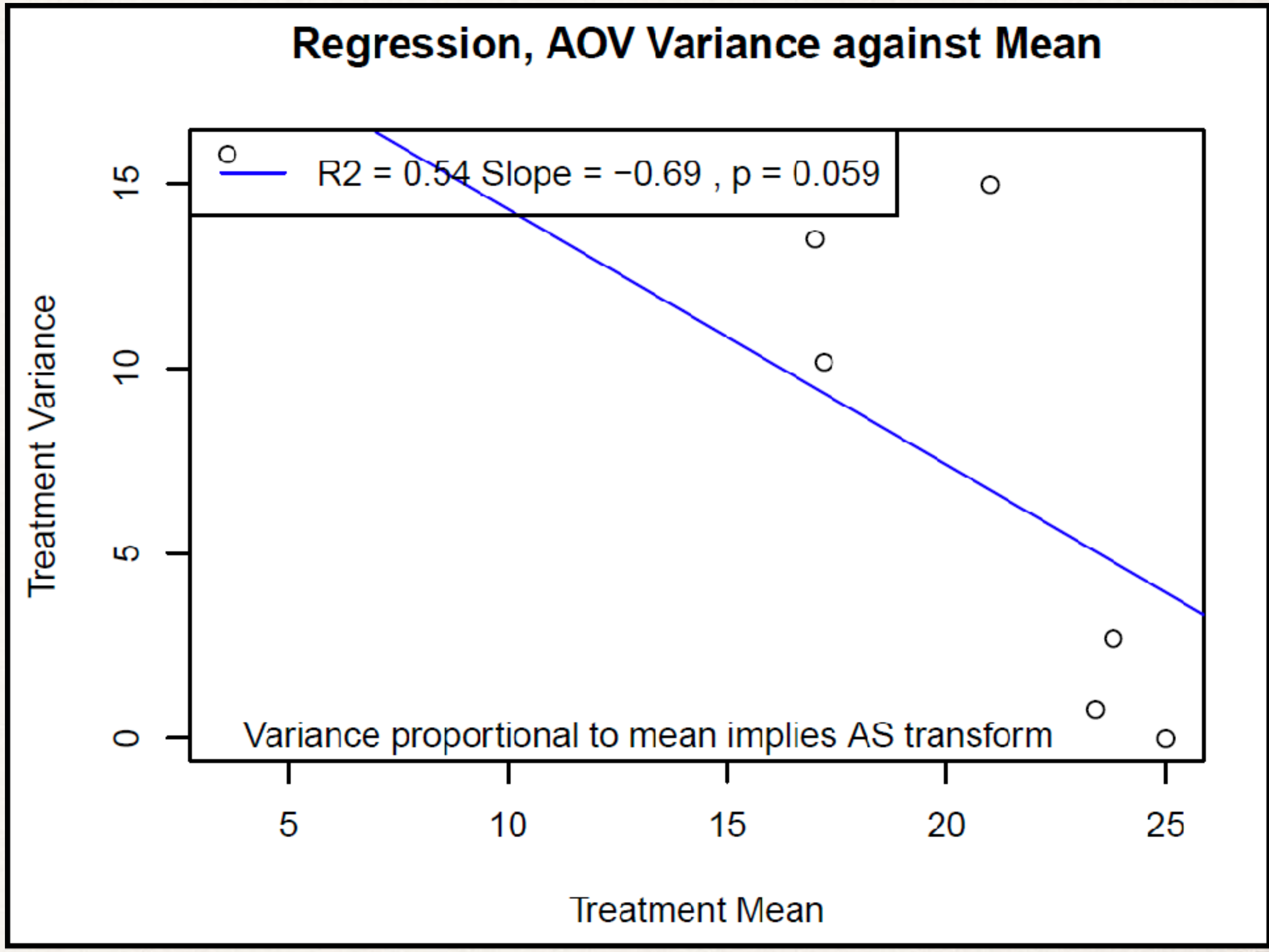
Bartlett 1947 V

- The inverse sine square-root transformation
$$g(x) = Sin^{-1}\sqrt{x}$$
bears the same relation to estimated probabilities or proportions x with binomial variance $p(1-p)/n$, where n is the number of individuals in the sample, as the square-root transformation does to a Poisson variate.



Bartlett 1947 V

- The inverse sine square-root transformation
$$g(x) = Sin^{-1}\sqrt{x}$$
bears the same relation to estimated probabilities or proportions x with binomial variance $p(1-p)/n$, where n is the number of individuals in the sample, as the square-root transformation does to a Poisson variate.

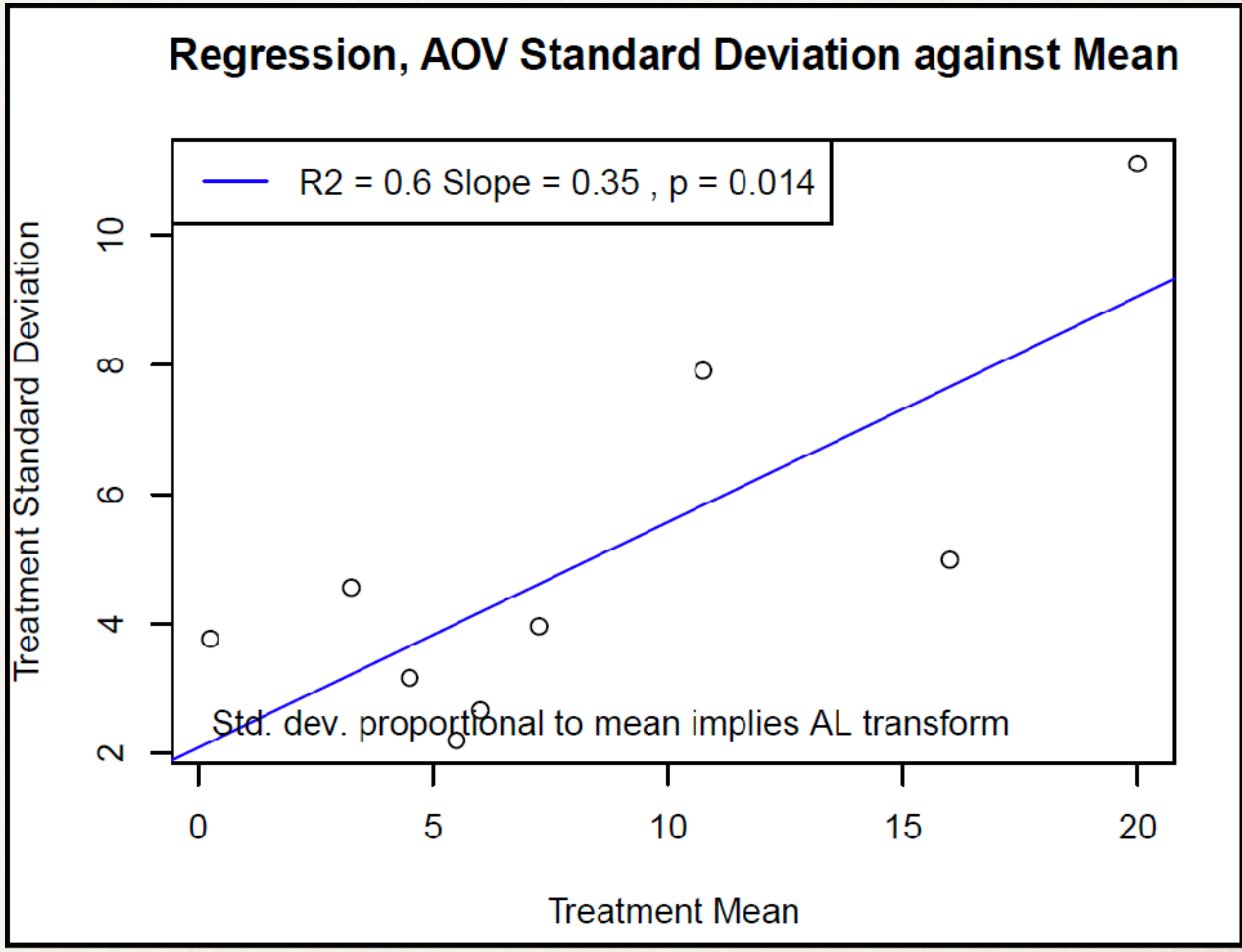


Extra Slides

Including some quotes from other sources.

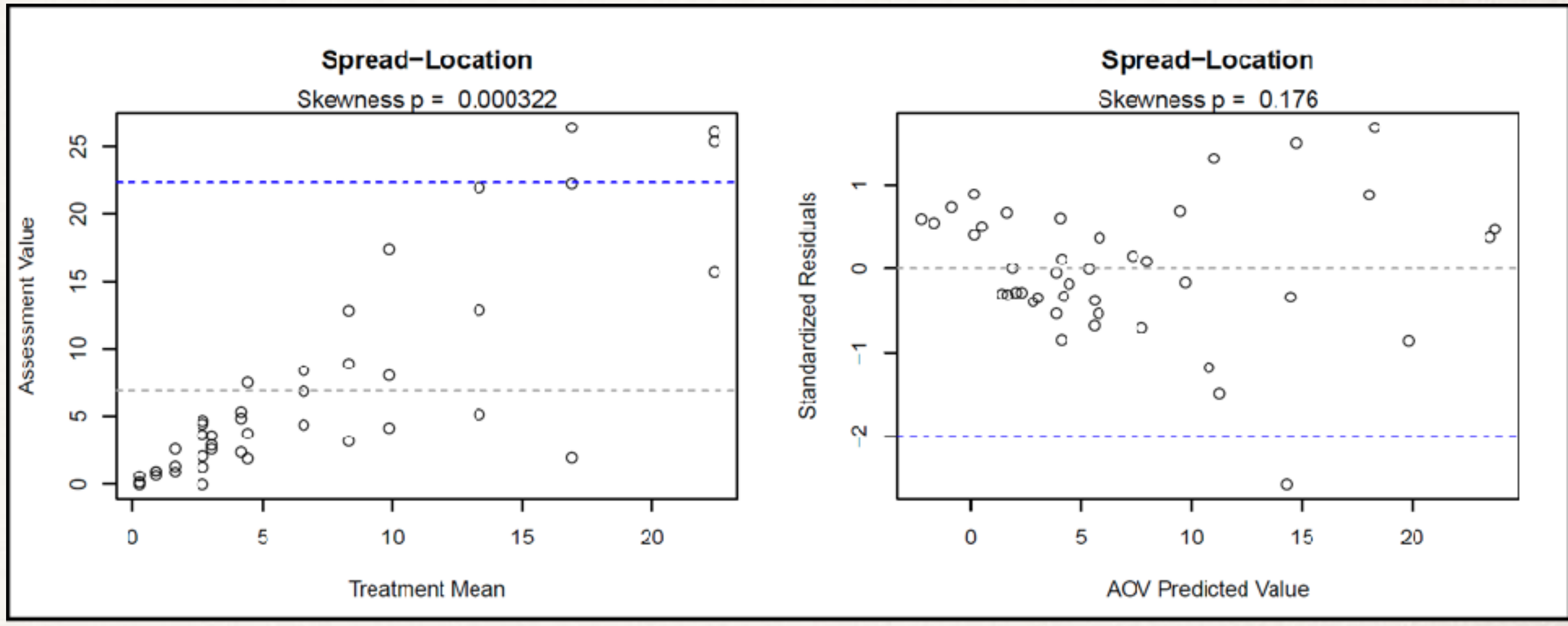
Gomez 7.14

- K. A. Gomez and A. A. Gomez. Statistical Procedures for Agricultural Research. Wiley-Interscience, 1984.
- The logarithmic transformation is most appropriate for data where the **standard deviation is proportional to the mean** or where the effects are multiplicative. These conditions are generally found in data that are whole numbers and cover a wide range of values. Data on the number of insects per plot or the number of egg masses per plant (or per unit area) are typical examples.



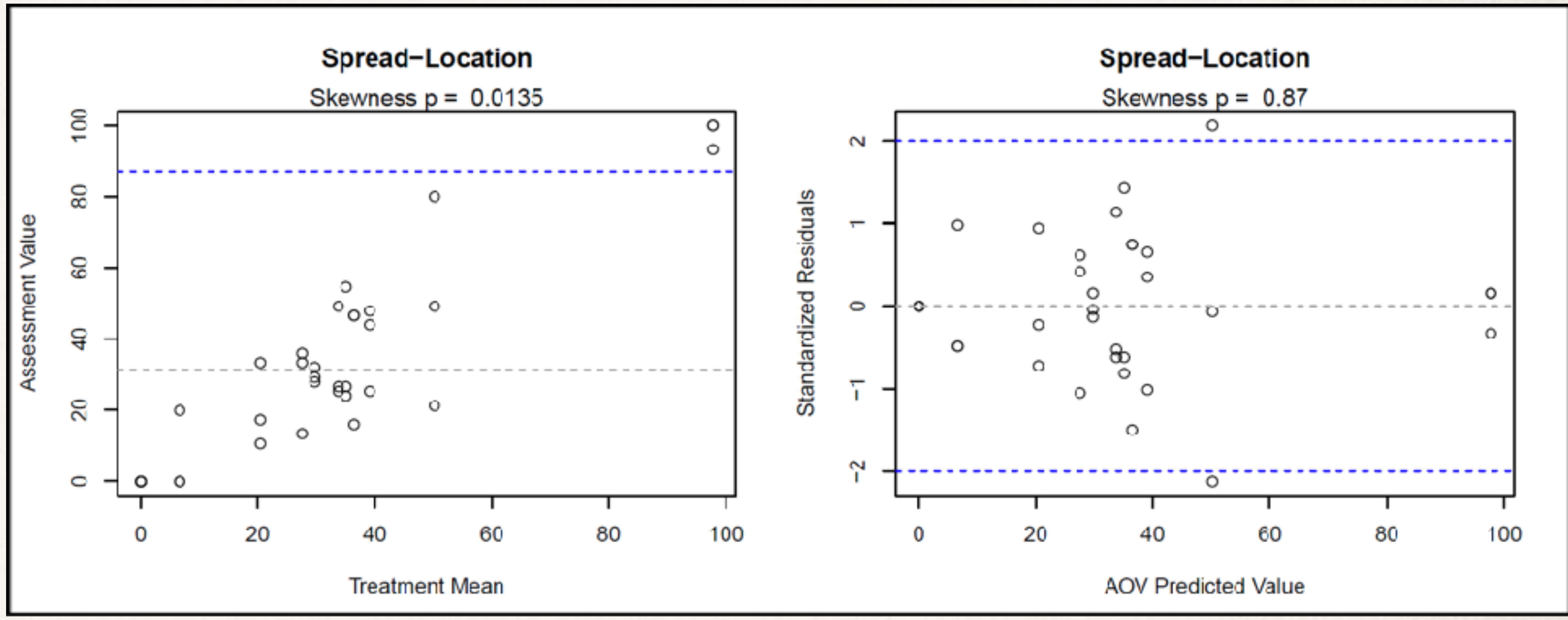
Gomez 7.19

- Square-root transformation is appropriate for data consisting of small whole numbers, for example, data obtained in counting rare events, such as the number of infested plants in a plot, the number of insects caught in traps, or the number of weeds per plot. For these data, the variance tends to be proportional to the mean The square root transformation is also appropriate for percentage data where the range is between 0 and 30% or between 70 and 100%. For other ranges of percentage data, see discussion on the use of the arc sine transformation in the next section.



Gomez 7.23

- The following rules may be useful in choosing the proper transformation scale for percentage data derived from count data
- Rule 1. For percentage data lying within the range of 30 to 70%, no transformation is needed
- Rule 2. For percentage data lying within th range of either 0 to 30% or 70 to 100%, but not both, the square root transformation should be used.
- Rule 3. For percentage data that do no follow the ranges specified in either rule 1 or rule 2, the arc sine transformation should be used.

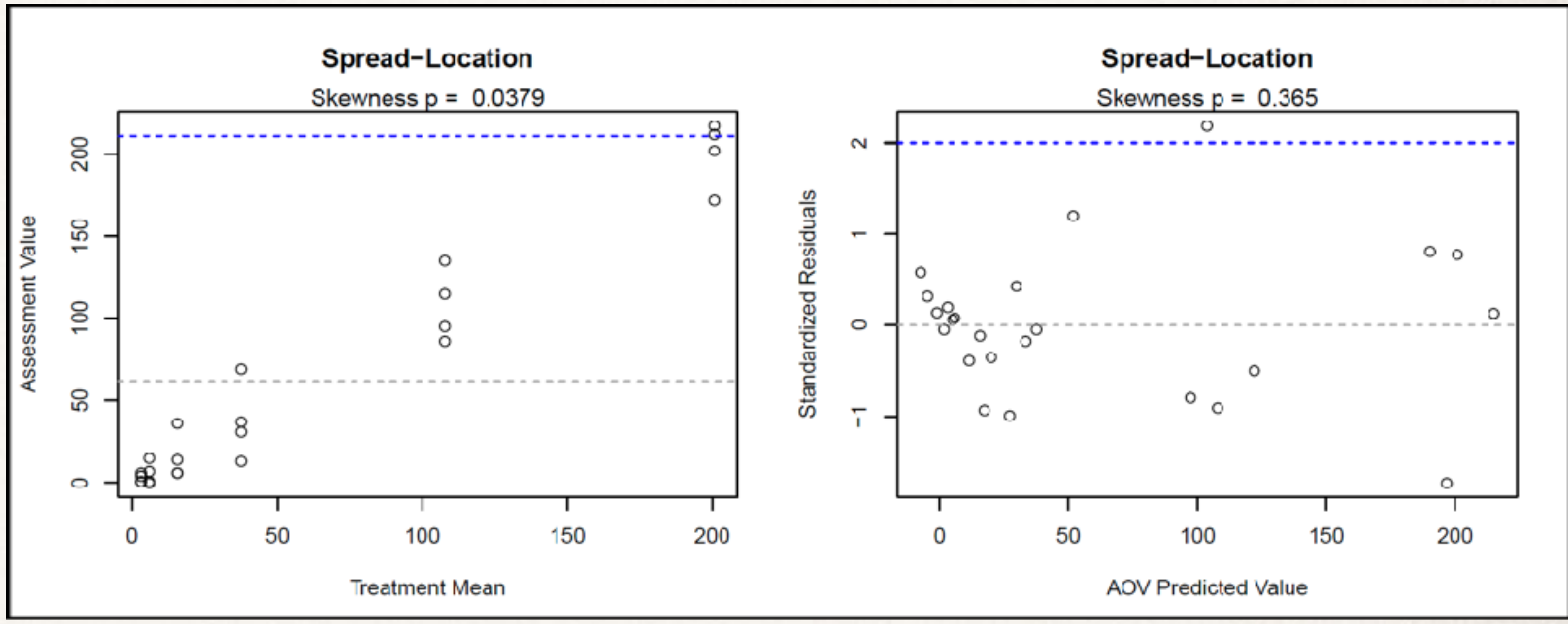


Steel 9.16.1

- R. G. D. Steel and J. H. Torrie. Principles and Procedures of Statistics, A Biometrical Approach. McGraw-Hill, second edition, 1960.

- Square Root Transformation

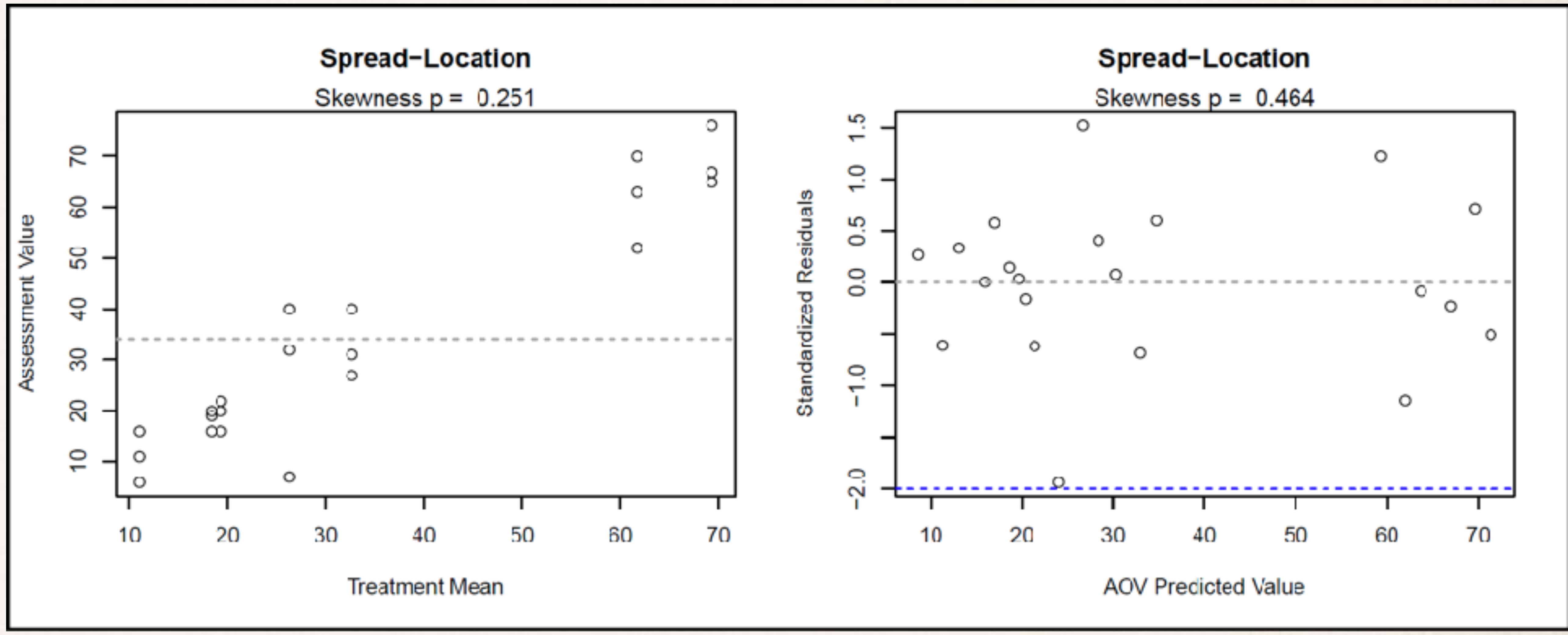
- When data consist of small whole numbers, for example, number of bacterial colonies in a plate count, number of plants or insects of a stated species in a given area, etc., they often follow the Poisson distribution, for which the mean and variance are equal.
  …
  Percentage data based on counts and a common denominator, where the range of percentages is 0 to 20 percent or 80 to 100 percent but not both, may also be analyzed using the square root transformation. Percentages between 80 and 100 should be subtracted from 100 before the transformation is made.



Steel 9.16.2

- Arcsine Square Root Transformation

- This transformation is applicable to binomial data expressed as decimal fractions or percentages, and is especially recommended when the percentages cover a wide range of values.
  …
  The square root transformation has already been recommended for percentages between 0 and 20 or 80 and 100, the latter being subtracted from 100 before transformation. If the range of percentages is 30 to 70, it is doubtful if any transformation is needed.



Steel 7.8

- When variances are proportional to the squares of the treatment means or standard deviations are proportional to means, the logarithmic transformation equalizes the variances. … Effects which are multiplicative on the original scale of measurement become additive on the logarithmic scale, for example, where one treatment gives a response 20 percent higher than another…

Column 1 Diagnostics

Diagnostics

☐ Include spatial models

Raw Graphs

Show... Layout: 2 X 2

Statistics (P)	Raw ☑	IID ☐	AL ☐	AS ☐	AA ☐	AR ☐
N	18	18	18	18	18	18
Unique	14	17	18	18	18	17
Analyzed	18	18	18	18	18	18
Missing	0	0	0	0	0	0
Empty	0	0	0	0	0	0
Damaged	0	0	0	0	0	0
MinRep	3	3	3	3	3	3
MaxRep	3	3	3	3	3	3
Treatments	6	6	6	6	6	6
Levene's	0.97906	0.99372	0.97430	0.99006	0.98946	.
ShapiroWilks	0.43789	0.02903	0.06026	0.04495	0.04628	.
Skewness	0.87551	0.40425	0.31599	0.36136	0.35930	.
Kurtosis	0.29609	0.25142	0.39523	0.29460	0.29871	.
MaxStdRes	.	1.19742	1.41538	1.31064	1.31780	1.62643
logLik	.	-17.74256	14.85532	10.49943	-22.30058	-35.41320
ModelDF	.	7	7	7	7	2
ResDF	.	10	10	10	10	15
AIC	.	53.48513	-11.71065	-2.99886	62.60116	78.82641
BIC	.	61.49847	-3.69730	5.01449	70.61451	82.38790

Recommendations

Basis: Assessment Values

IID Graphs

Show... Layout: 4 X 2

	Code	Test Statistic	Value	Comment
1	IID	Levene's	0.08249	Does not fail test of homogeneity of variances among treatments
2	AL	ShapiroWilks	0.88274	Analyze transformed assessments to improve normality of residuals
3	IID	Skewness	-0.49610	Does not fail test of skewness of residuals
4	IID	Kurtosis	-1.33266	Does not fail test of excess kurtosis of residuals

Save to RStudio

Previous Next

Column Diagnostic

Column Properties

Beall 1942

- G. Beall. The transformation of data from entomological field experiments so that the analysis of variance becomes applicable. Biometrika, 32(3/4):243–262, Apr 1942. https://www.jstor.org/stable/2332128

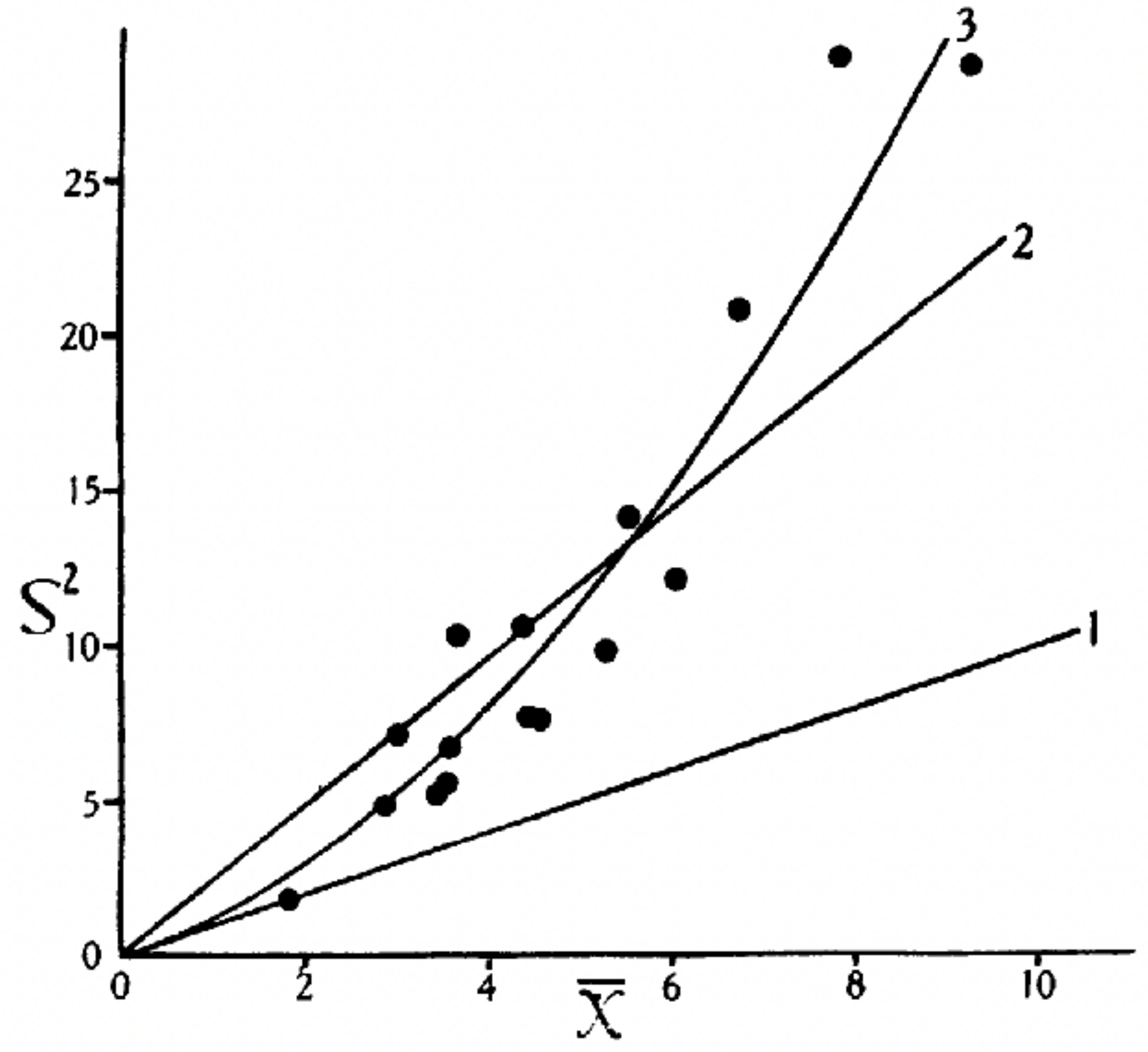


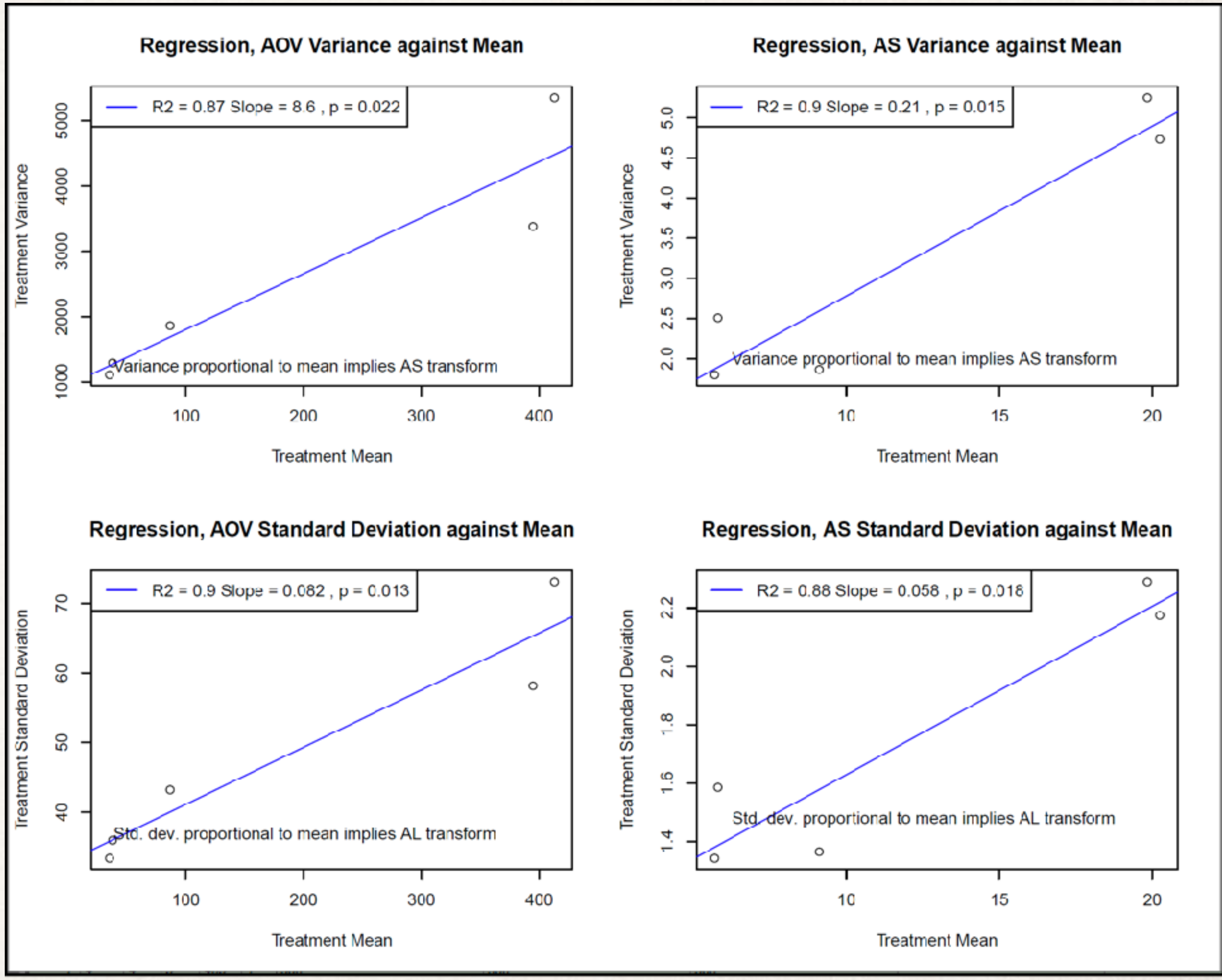
Fig. 1. The squared standard deviation plotted against the mean for 144 small areas within each of 16 large areas; line 1 is from equation (1), line 2 from (3) and line 3 from (6). The counts had been made on *Leptinotarsa decemlineata* Say.

Snedecor 15.11.1

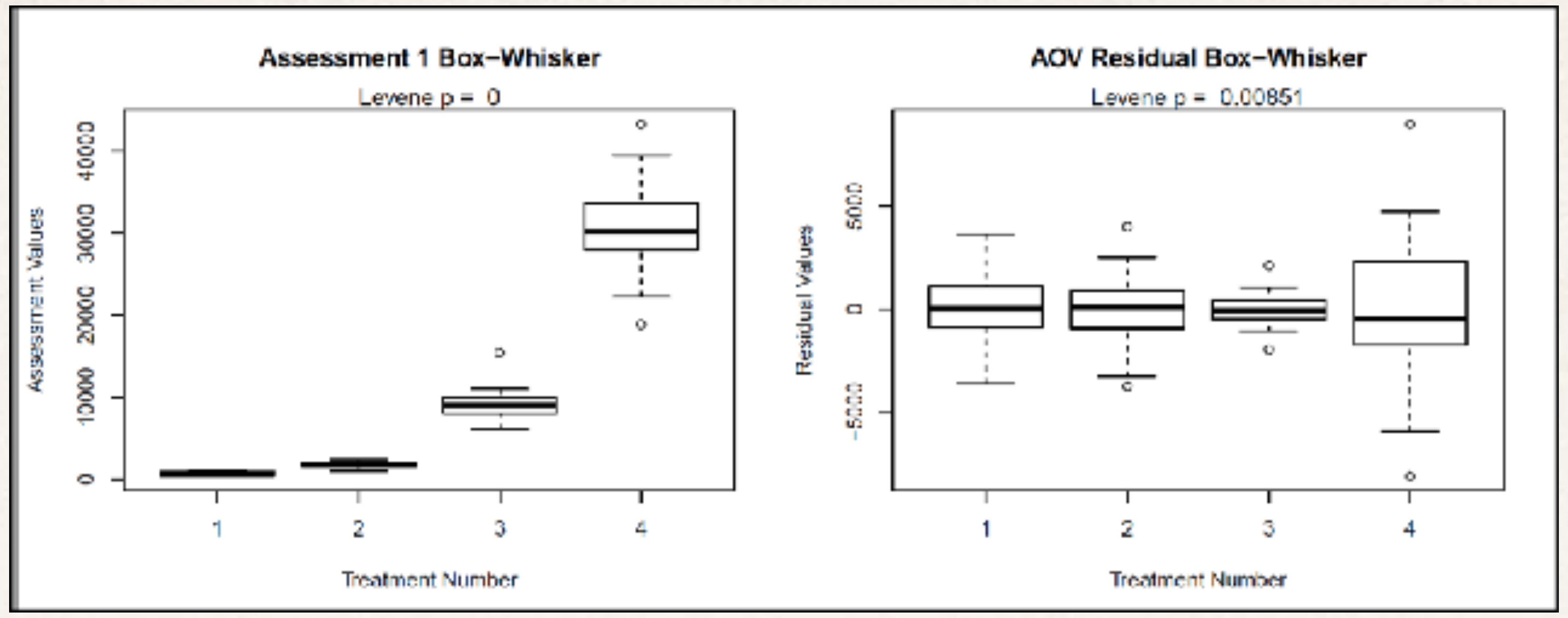
- G. Snecdecor and W. Cochran. Statistical Methods. Statistical Methods. Wiley, seventh edition, 1980.

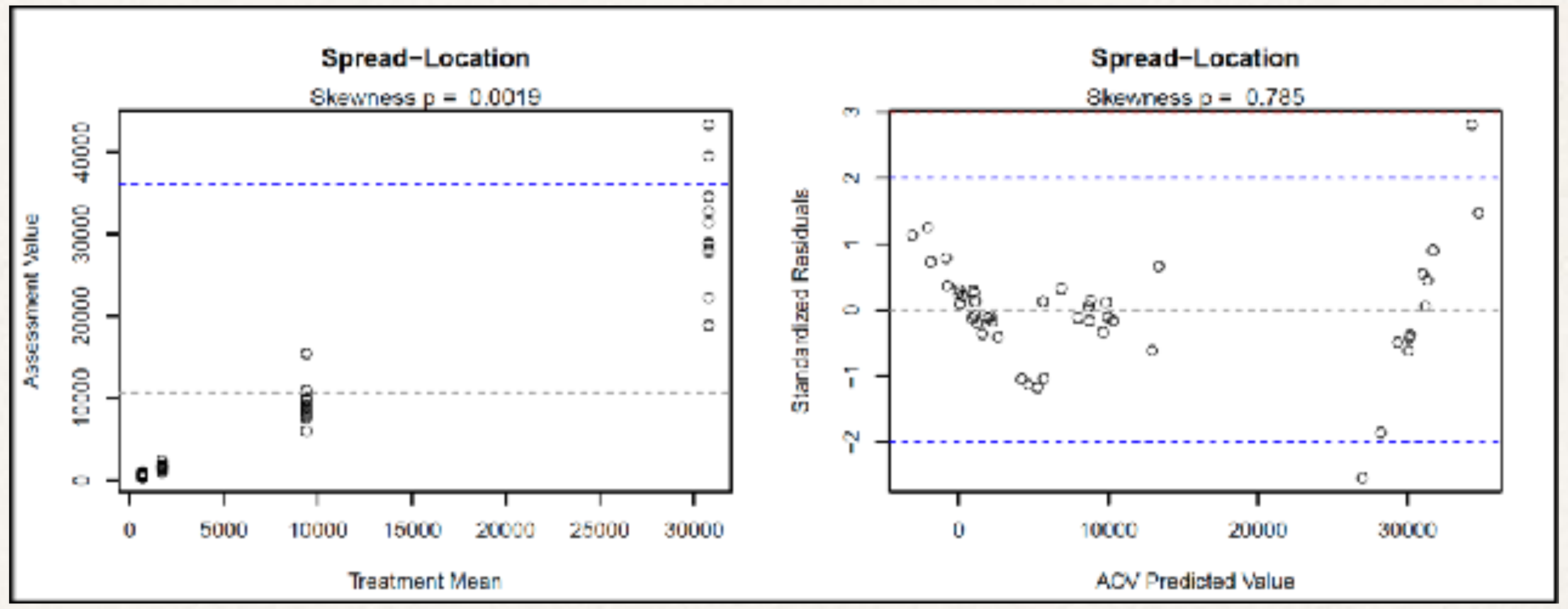
- Counts of rare events, such as numbers of defects or of accidents, tend to be distributed approximately in Poisson fashion. A transformation to $\sqrt{X}$ is often effective; the variance on the square root scale will be close to 0.25. …

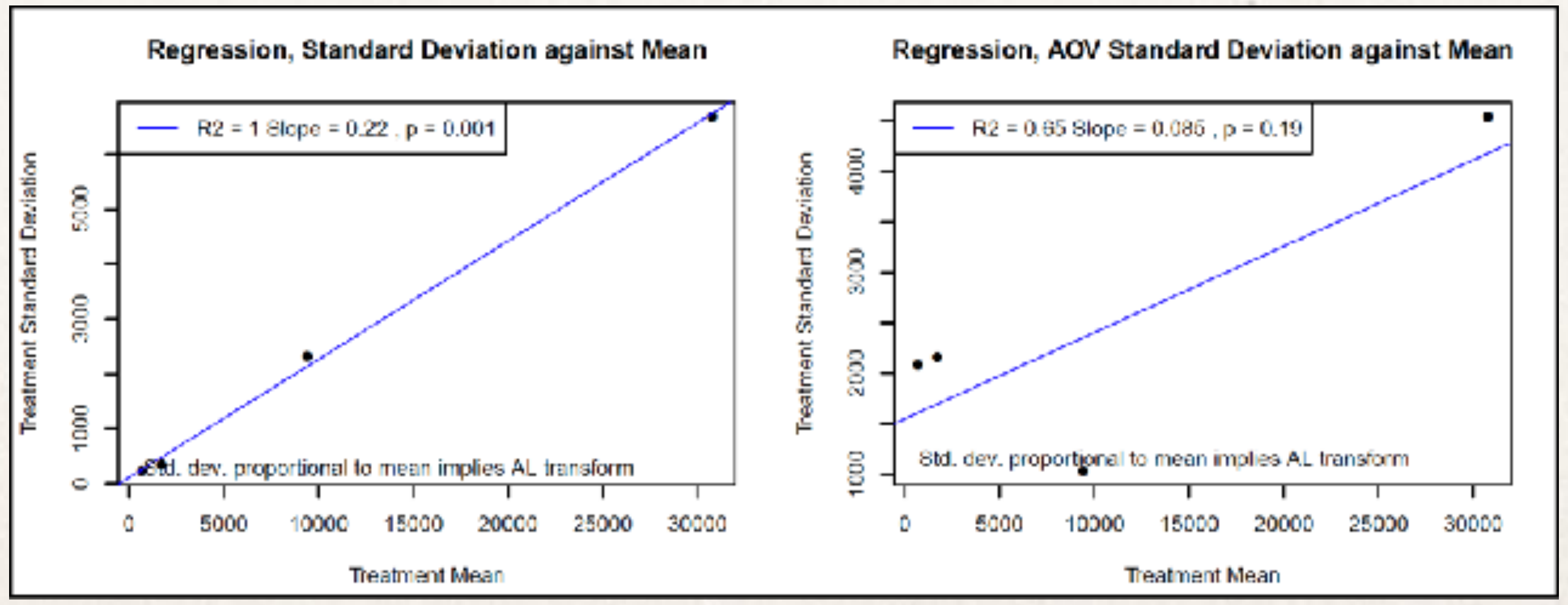
  The square root transformation can also be used with counts in which it appears that the variance of X is *proportional* to the mean of X, that is $\sigma_X^2 = k\bar{X}$



Curran-Everett 2018







- D. Curran-Everett. Explorations in statistics: the log transformation. Adv Physiol Educ, 42(2):343–347, Jun 2018.