Expl Agric. (2001), volume 37, pp. 115–123
Printed in Great Britain
Copyright © 2001 Cambridge University Press

INVITED PAPER

PRESENTATION OF STATISTICAL ANALYSES

By JANET RILEY

Statistics Department, IACR-Rothamsted, Harpenden, AL5 2JQ, UK

(*Accepted 8 August 2000*)

SUMMARY

Instructions are given for the presentation of the results of statistical analyses in *Experimental Agriculture*. Statistical codes accepted by the journal are defined. Appropriate styles of presentation are illustrated for estimated parameters and variances for a selection of standard statistical analyses. The value of presentations of results of significance tests is discussed. Formats for acceptable tables and graphs are shown. References to further guiding texts are included for non-standard situations.

INTRODUCTION

This paper gives instructions for the presentation of the results of statistical analyses for authors wishing to submit typescripts to *Experimental Agriculture*. An earlier paper in this journal (Riley, 1994) provided guidance on both how to *describe* experimental objectives, designs, results and interpretation of results and how to *present* the numerical results of those experiments. Finney and Harper (1994) listed the statistical code acceptable to the Journal's Editor. In one concise summary this paper brings together the statistical code and the way to present the numerical results of experiments. It should be read in conjunction with the Editor's 'Instructions to Authors' included in this issue of the journal.

PRESENTATION OF RESULTS OF STATISTICAL ANALYSES

The presentation of the results of both initial and final analyses usually forms the major numerical section of a published paper or report. These analyses may range from displays of simple statistics to complex multidimensional graphs. The statistical codes and presentational formats acceptable to *Experimental Agriculture* are shown here for the most popular statistical analyses.

Basic statistical code

Most statistics textbooks introduce the subject of estimation using the terminology 'population' and 'sample' and 'mean' and 'variance'. The basic argument is that properties of a very large population can be estimated from the same properties of a representative sample of n values, $x_1, x_2, \ldots x_n$, from that population. In general terms, the average value of the population may be estimated by the

https://doi.org/10.1017/S0014479701001028 Published online by Cambridge University Press

average value of $x_1, x_2, \ldots x_n$ and the variability in the population may be estimated by the variability in $x_1, x_2, \ldots x_n$. The unknown average value of the population is usually noted as μ, the population mean, and the estimated average value of $x_1, x_2, \ldots x_n$ is usually noted as $\bar{x}$, the sample mean. The unknown variability in the population is usually noted as σ^2, the population variance, and the estimated variability in $x_1, x_2, \ldots x_n$ is usually noted as s^2, the sample variance. Written in the text of a paper these symbols should be in lower case italic as shown here.

Means and variances

The presentation of estimated mean values and the variability – or error – of these estimates demonstrates the efficiency of the completed analysis or model-fitting process and provides all the information needed to do significance tests, if required, upon the estimated values or meaningful functions of them.

The most efficient forms of variance to present for mean values are the standard deviation of the sample, *s.d.*, the standard error of the mean of the sample, *s.e.*, and the standard error of the difference between two sample mean values, *s.e.d.*. These are calculated as:

$$s.d. = \sqrt{s^2}$$

$$s.e. = \sqrt{\frac{s^2}{n}}$$

$s.e.d. = \sqrt{((s_1^2 + s_2^2)/n)}$ where s_1^2 and s_2^2 are the sample variances for two separate samples each of size n. Note that the code for these errors and for n should be presented in the text in italic.

Other estimates of error exist: the standard error per plot; the least significant difference (L.S.D.); critical difference (C.D.); and the coefficient of variation (C.V.). Ryder (1986) discusses these and concludes that all three forms of standard error – the *s.d.*, the *s.e.* and the *s.e.d.* – can be valuable when appropriately used, whereas the coefficient of variation and the least significant difference can be misleading and should be avoided. The critical difference is equally misleading.

The three forms of standard error whose formulae are presented above are acceptable to *Experimental Agriculture*. Whichever of the three is used, it should be located in the paper with its associated estimated mean values. Table 1 shows a clear presentation of three mean grass yields (kg plot^{-1}) with the associated standard error of the means (*s.e.*).

Altman and Gardner (1986) note that the use of $\pm$ to connect a unit of variability to a mean value should be avoided where the form of error is unclear. Thus a statement such as 13.25 ± 0.446 g min^{-1} does not specify whether 0.446 is a standard error or a standard deviation. They propose that the $\pm$ sign should be discontinued and that the use of the standard deviation or the standard error

https://doi.org/10.1017/S0014479701001028 Published online by Cambridge University Press

Table 1. Mean yield of grass (kg plot $^{-1}$).

N kg/ha		
0	20	40
9.64	10.32	11.54
	s.e. 0.916	

should be clarified. Thus the preferred form is 13.25 with *s.e.* 0.446; or 13.25, *s.e.* 0.446, with corresponding notation where *s.d.* is used.

An alternative way to present an estimated mean value and its variability is to use a confidence interval. This involves the concept of probability levels and distributions. For example, a confidence interval for a mean value can be found from its standard error and the tabulated *t*-value at a specified level of probability. Thus, if the sample mean value of 13.25 is calculated from $n = 16$ values, the 95% confidence interval for the population mean is $13.25 \pm (2.13 \times 0.446)$, i.e. 12.30 to 14.20. This can be presented as:

. . . . mean value 13.25 with confidence interval (12.30, 14.20); or
. . . . mean value 13.25 with *C.I.* (12.30, 14.20), maintaining italic in the abbreviated notation.

Usually, a set of mean values from the same dataset is calculated from samples of the same size and each will have the same value for its standard error of the mean. However, where the means are based upon different numbers of values or where adjusted mean values are produced from some model-fitting process, each mean will have a different standard error. Presentation must be done in such a way that the estimate of error is correctly associated with the corresponding mean value and that the replication is declared. Riley (1994, p. 389) gives a clear example of how this can be achieved.

Tests of significance

When the desired estimates are presented with their appropriate standard errors, whether in the text, in a table or in a graph, the reader has a compact set of information: the summarized dataset with salient features displayed and an estimate of error to demonstrate the precision with which the salient features have been estimated. *This gives the reader enough information to do whatever significance test at whatever level she or he chooses*. Nothing extra can be gained from the presentation of analyses that led to these means and standard errors, such as analysis of variance tables, although any unusual features of them may be discussed in the text.

It is not particularly helpful to provide the results of a significance test that you have chosen. To state that significant differences between means were achieved at

https://doi.org/10.1017/S0014479701001028 Published online by Cambridge University Press

the 5% level is only to present a significance test and a significance level chosen arbitrarily. It is particularly *unhelpful* to state that significant differences between means were achieved at the 5% level without stating (a) the mean values and the sample size, (b) the standard error and (c) the name of the test that you have chosen.

For the same reasons, the use of asterisks and N.S. should be avoided; these have the added disadvantage of cluttering tables and text. Associated with this is the practice of completing a multitude of comparisons between means and demonstrating all differences or similarities by attaching letters, such as a, b, c . . . , to the means. Such clutter results in difficulty in interpretation and, again, merely represents the results of an arbitrarily chosen test that the reader may not desire. Even where a global (say, F) test has shown no differences between treatments, comparisons between all pairs of means may still show significant differences. These will be due to chance and are unlikely to represent the true population features.

The use of multiple range tests should be limited to very specific situations where no prior knowledge is available about the performance of the treatments. Such a situation *may* occur if, say, a selection of new varieties are being compared (although in practice it is rare to have absolutely no knowledge about all the varieties).

Multiple range tests should **never** be used:

(a) in factorial experiments, which have an implicit structure to the treatments; or
(b) when known controls are to be compared with new treatments.

Much literature has appeared to explain the proper use of multiple range tests; Pearce (1993) should be consulted for a clear exposition of their value.

The best strategy is to specify, before an analysis is done, the biologically meaningful contrasts between mean values or groups of mean values according to the objectives of the study. Such contrasts include response curves – linear and quadratic polynomials being the most usual for quantitative treatments where, for example, increased amounts of fertilizer are applied. Pearce (1992) gives a clear exposition of some standard contrasts for quantitative treatments for both single and factorial treatment structures and for a non-standard set of contrasts where groups of mean values are compared.

Statistical distributions

N(0,1) refers to a Normal distribution with mean 0 and variance 1. $N(\mu, \sigma^2)$ refers to a Normal distribution with mean μ, and variance σ^2. Use of the italic form will distinguish it from N which represents the chemical element nitrogen.

The Student's t-statistic is used for comparing two different arithmetic means and the means must be clearly identified. The F statistic is formed from the ratio of two independent error mean squares usually, but not always, from an analysis of variance. The χ^2 statistic relates to tables of frequencies and is used in many tests of

https://doi.org/10.1017/S0014479701001028 Published online by Cambridge University Press

significance. Whenever they are referred to, or their results are presented, the associated degrees of freedom (*d.f.*) must also be quoted. Thus the presentation should be:

$$t_{12},\ F_{3,23} \text{ or } F_{[3,23]} \text{ or } \chi^2_{[17]}$$

When a test of significance is associated with these statistics the probability level should be stated as well. Thus $p = 0.001$ or $p < 0.005$. Note that the notation for probability is in lower case italic.

Transformations of data

To satisfy the assumptions of a statistical analysis, data may need transformation to a different scale before they are analysed. The most common transformations are to the logarithmic, angular or square root scales. Care should be taken that the base for the different types of logarithmic scale are noted. Thus, log_e indicates natural logarithms to base e, and log_{10} indicates logarithms to the base 10. Where no subscript is given – *log* – this represents the natural logarithm to base e.

The transformed scale should be noted in the text or in the table heading and the estimated values presented with their corresponding estimates of error on the same scale. Where the presentation of the results on the original scale is needed to assist with interpretation – and this is not always necessary – back-transformed means can be included, but note that this back transformation cannot be applied to the estimates of error. Confidence intervals can be back transformed, but this should be avoided as they can lead to excessively large tables. The minimum amount of necessary information is, the transformed means, their corresponding standard errors or standard errors of differences and, if deemed absolutely essential, the back transformed means. An example of a suitable form of presentation is in Table 4 of Riley (1994).

Correlation and regression analyses

Correlation between two variables should be noted by r, which is the notation that should be used only for the product moment or Pearson correlation coefficient and not for any other correlation coefficient. It may be modified to $r_{x,y}$ if the variables involved need to be specified. For ranked data, the Spearman rank correlation coefficient should be used and noted as r_s to distinguish it from r.

When linear regression or a generalized linear model has been used, interest will lie in the best form of model achieved together with its precision. This is best indicated by standard errors attached to the coefficients of the parameters, as follows: the estimated regression line between yield of sorghum (kg) and rainfall is

$$\text{yield} = a + b \text{ rainfall}$$

where $a = 14.56$, *s.e.* 1.382, and $b = 5.33$, *s.e.* 0.848

This presentation is adequate as it shows the estimated coefficients and their

https://doi.org/10.1017/S0014479701001028 Published online by Cambridge University Press

precision, reflecting the adequacy of the fit of the model to the data. A good example of this type of presentation is given by Peters, Tarawali and Schultz-Kraft (2000, p. 364).

An indication of the fit of the model to the data, particularly for more complex equations and for comparison of similar but different models, may **additionally** be given by r^2, the square of the product moment correlation coefficient. This measures the amount of total variation in the yield variable that is explained by its relationship with the rainfall variable. Note that only r^2 values of about 0.8 and above express strong relationships between two variables (Finney, 1980). However, an r^2 value should **never** be presented alone to express the strength of relationship between two variables without the corresponding estimated regression equation, its coefficients and their standard errors because, alone, it can be misleading.

Note also that r, r^2, r_s and the regression coefficients a and b should be presented in the text in lower case italics. Other codes $\bar{r}^2$, $\bar{r}^2$, R, R, R^2, R^2, are non-standard and are not acceptable.

Degree of precision

Computer software often prints results of analyses with large numbers of digits. But to give a mean weight as 44.3716 kg is unnecessary and, where several mean values are to be presented, confusing to the reader. A rounded value such as 44.4 kg is adequate and the conclusions of the study will not be biased because of this refinement of presentation. The retention of 3 significant figures aids calculation and loses nothing from the precision of the analysis. The distinction needs to be made between significant figures and decimal places. The number 0.0147 has been recorded to 4 decimal places and to 3 significant figures. The number 0.015 has been recorded to 3 decimal places and to only 2 significant figures. To obtain precision in analysis and in presentation the use of 0.0147, i.e. 3 significant figures, is preferred. However, when presenting a set of figures with decimal places and the numbers vary in size, strict rigidity to the significant figures rule may result in an unclear presentation. It is usual then to retain the same number of decimal places throughout, even if larger numbers are then presented with more than three significant figures.

Tabular presentation

A large data set cannot easily be presented in a paper and a statistical analysis will reduce it to a more compact form. Where a number of values are to be presented, a tabular display can assist clarity. Tables should not repeat information from the text or from any pictorial display. Ideally, each should provide a complete set of information that is comprehensible if the text is removed. Tables of treatment means and two-way tables of means from factorial experiments are common forms of presentation. Three-way tables can be confusing and greater clarity may be achieved by presenting a subset of the two-way tables for those

https://doi.org/10.1017/S0014479701001028 Published online by Cambridge University Press

Table 2. Mean pod and hay yields (t ha^{-1}) of five groundnut cultivars grown at Wad Medani, Sudan in 1984–86.

	Pod yield				Hay yield			
Cultivar	1984	1985	1986	Mean	1984	1985	1986	Mean
Early Bunch	2.47	3.34	2.14	2.65	3.56	2.40	2.61	2.86
Georgia 119–20	1.73	3.38	2.29	2.46	3.50	3.88	2.81	3.39
UF79–1499	1.61	2.05	2.34	2.00	5.23	1.82	3.26	3.44
Apollo	1.47	3.38	1.93	2.26	4.65	3.44	2.81	3.63
MH 383	1.75	3.70	2.14	2.53	4.78	4.58	2.99	4.12
Mean	1.81	3.17	2.17	2.38	4.34	3.22	2.88	3.49
s.e.	0.235	0.223	0.167	0.210	0.571	0.292	0.264	0.373

factors that interact. Ryder (1986) discusses the choice of these for a number of situations and gives extensive advice with regard to presentation.

A table needs to be identified clearly with a heading that states its purpose without need to consult the text. The variable(s) involved should be named, the units declared and any associated treatments named. The numbers of rows and columns in a table should be small – if in doubt, reduce them – and spacing and blank lines may usefully be inserted to aid its comprehension. This use of spacing can be particularly effective if used to segregate particular features of the data or to isolate different variables and is easier to view than divisions made with ruled lines. A good example of a clear table is given in Table 2 (which is Table 2 from Ishag, 2000, p. 306).

This table has a heading explaining what is in the table together with the units of measurement. The five cultivars, the two variables and the three years are clearly named, the mean values and standard errors are presented (the mean values to three significant figures and standard errors to one further decimal place). All the information is there in a clearly summarized form and the reader can do tests of significance if she or he wishes.

Graphical presentation

An alternative form of presentation is the graph. This may be a two-dimensional display such as a scatterplot, a pictorial presentation such as a pie-chart or a histogram or a more complex three-dimensional display. Whichever is used, it should not repeat information in the text or in a table and should be fully comprehensible even if removed from the paper.

A balance between clarity of presentation and adequate provision of information should be the aim. As with tables, too much information will confuse; the major features should be exposed and displayed to their full but true advantage. Modern software can provide sophisticated graphics that can be manipulated to such an extent that the data features to be illustrated can be hidden by the gloss of technology! An example of a graph plot (from Craufurd, 2000; Figure 2, p. 387) is given in Figure 1.

https://doi.org/10.1017/S0014479701001028 Published online by Cambridge University Press

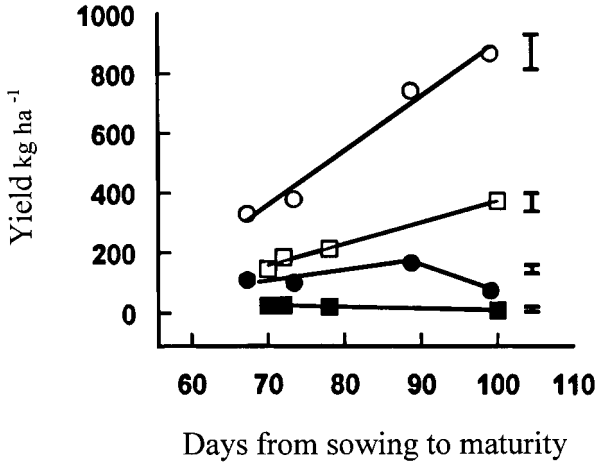


Fig. 2. Relationship between biomass (○●) and seed (□■) yield, and days from sowing to maturity, among four varieties of cowpea grown at Minjibir (open symbols) and Malam Madori (closed symbols) in Nigeria in 1992. Bars are s.e.d. between means. Fitted lines: ○Y = −919(156) + 18.2(1.88)X, r^2 = 0.98; □Y = −349(42) + 7.2(0.52)X, r^2 = 0.99; ■Y = 61(3.5) − 0.5(0.04)X, r^2 = 0.99; ● line fitted by eye.

Fig. 1. An example of a clear graph.

This has a clear legend explaining the two variables, the symbols used in the graph, and the regression relationship for each subset of data with the estimated coefficients, their standard errors and the corresponding values of r^2. Both axes are clearly labelled with the appropriate variables and units. The bars in the graphs indicating *s.e.d.*s between mean values for each subset of data are described in the legend. No further information is required.

Statistical packages

Much commercially-available statistics software is well-known. Genstat, SAS, SPSS, Minitab need no further reference, although it is helpful to provide the version number, as these packages are constantly being updated and features change. Note should be taken of the different styles of presentation of statistical results from the different packages, and the presentational requirements of this journal. Styles of presentation not in agreement with editorial requirements must not be inserted directly into a typescript but should be modified first.

Dimensions and irregularities

Dimensions of standard deviations and standard errors are those of the measurements to which they relate. A regression coefficient of y upon x has dimensions 'units of y per unit of x' and an explicit statement of these units is required if discussion of the regression coefficient ensues. The statistics t, r, F and χ^2 are dimensionless. All other variates and statistics have physical dimensions which must be stated clearly.

https://doi.org/10.1017/S0014479701001028 Published online by Cambridge University Press

Familiar and well-understood statistical methods generate terminology and symbols that should not cause difficulty for the reader provided each usage is fully supported by all desired information of the procedure. Any irregular usage of the statistical codes used here should be defined. Use of F, N or P (note the non-italic, capital forms) as chemical terms will be clear. Use of N or n should make clear whether these are sample sizes, degrees of freedom or something else. Use of t for time or r for a ratio should be clearly stated.

Use of general names for methods should be avoided. Reference to a generalized linear model without specification of the error distribution and link function or reference to multivariate analysis without specifying whether it is, for example, multiple regression or principal component analysis, is not acceptable.

FINAL REMARKS

Where the presentation of more complex analyses are required the assistance of a professional statistician should be sought. The material given here together with the cited references will provide guidance in presentation for the most popular statistical methods. Both authors and referees are encouraged to note these codes and presentational styles, as adherence to them will achieve quicker acceptance of submitted typescripts.

Acknowledgements. IACR-Rothamsted is grant aided by the Biotechnology and Biological Sciences Research Council of the United Kingdom.

REFERENCES

Altman, D. G. & Gardner, M. J. (1986). Presentation of variability. *The Lancet* September 13:639.

Craufurd, P. Q. (2000). Effect of plant density on the yield of sorghum-cowpea and pearl millet-cowpea intercrops in northern Nigeria. *Experimental Agriculture* 36:379–395.

Finney, D. J. (1980). *Statistics for Biologists*. London: Chapman and Hall.

Finney, D. J. & Harper, J. L. (1994). Editorial Code for presentation of statistical analyses. *Experimental Agriculture* 30:377–380.

Ishag, H. M. (2000). Phenotypic and yield responses of irrigated groundnut cultivars in a hot environment. *Experimental Agriculture* 36:303–312.

Pearce, S. C. (1992). Data analysis in agricultural experiments. I. Contrasts of interest. *Experimental Agriculture* 28:245–254.

Pearce, S. C. (1993). Data analysis in agricultural experiments. III. Multiple comparisons. *Experimental Agriculture* 29:1–8.

Peters, M., Tarawali, S. A. & Schultz-Kraft, R. (2000). Relative palatability and seasonal agronomic performance of selected pasture legumes for species mixtures in dry-subhumid West Africa. *Experimental Agriculture* 36:353–368.

Riley, J. (1994). Aspects of statistical technique and presentation. *Experimental Agriculture* 30:381–394.

Ryder, K. (1986). Numerical tables: guidelines for presentation. *Journal of Applied Statistics* 13:77–88.

https://doi.org/10.1017/S0014479701001028 Published online by Cambridge University Press