ISSN 1110-0419

Annals Of Agric. Sc., Moshtohor, Vol. 45(3): 1021-1032, (2007).

THE COEFFICIENT OF VARIATION AS A MEASURE OF FIELD TRIAL VALIDITY

BY

Tageldin, M.H.A.
Agronomy Dept., Faculty of Agriculture, Moshtohor, Banha University

ABSTRACT

The coefficient of variation (CV) as a statistical measure to validate field trials or to compare relative variability of many trials being used by most researchers has lately been subjected to further inquiry. The CV basis relies on the assumption that both the mean and the root mean square error are proportional. If data needs transformation in case of suspicion of any violations of the analysis of variance assumptions, the application of CV is invalid since one objective of transforming data is to minimize the dependency relationship between error variance and mean. The regression between the natural logarithms of both error mean square (EMS) and mean which results in a regression coefficient estimate =2.0 indicates a constant CV across trials. Our objectives here were to: (i) check for the relationship among CV, EMS, and mean based on calculating the regression coefficients resulted from the regression of ln (EMS) on ln (mean) (ii) study the effect of data transformation on error mean square (EMS), CV, and coefficient of determination (R^2). The regression of ln (EMS) on ln (mean) produced a significant slope (b)=2.798 for wheat (*Triticum aestivum* L.), and a non significant estimate of 1.139 for maize (*Zea mays* L.). Neither estimate was close to 2.0, the value that supports the assumption that the CV is a sound tool for judging a trial's validity. Six ways of data transformation-- square root, arcsine, logarithmic, inverse, addition of a constant, and linear—were performed on barley (*Hordeum vulgare* L.) biomass data that had been subjected to addition of constants to produce data that have problems with shifted means, and with variance heterogeneity. Data transformation lowered CV values depending on the mode of variance heterogeneity, and the impact caused by a specific transformation. Inverse transformation to either data with greater range of data points or with variance inversely proportional to the mean caused the CV values to inflate to reach higher values as of 12 and 72 %, respectively relative to non transformed data or other forms of transformation. The R^2 was less sensitive to transformation; so it is a reliable candidate to compare relative variability of trials whether some of which may or may not have been transformed, yet it is affected by the size of dataset. Adjusted R^2 may be considered another sound statistical measure of trial validity. It may be used to compare relative variability of experiments having different sizes. Assessment of the validity of a field trial, or comparing the relative variability of more than one trial requires not just one statistical tool to base decision on keeping or discarding data. Unfortunately, there

has not been so far a clear-cut way to enable researchers make a decision to validate his/her data. Researchers should apply the CV, R^2, and R^2_{adj} or EMS to determine the validity of a trial, and not just relying on one single measure.

INTRODUCTION

The coefficient of variation (CV) has ever been used, by agronomists, breeders, and extension personnel, as one important criterion for assessment of the validity of field experiments. It is frequently used to accept or reject the validity of trials (Bowman 2001). Snedecor and Cochran (1980) stated that "For data from different populations or sources, the mean and standard deviation often tend to change together so that the coefficient of variation is relatively stable or constant." Based on their argument, the CV can be used to compare the relative variation of any particular trait from different populations. Bowman (2001) held Snedecor and Cochran (1980) responsible for conveying on this particular CV usage. Rather, it was suggested by Karl Pearson in the late nineteenth century just as a measure of relative population variability (Bowman and Watson, 1997; and Bowman, 2001) since it expresses the relationship between the standard deviation as a fraction of the mean for any given dependent trait. Damon and Harvey (1987) pointed out its contribution for measuring the relative amount of variation between different characteristics or traits. The CV can be used, in addition, to compare phenotype stability among populations, and variability of plot sizes in uniformity trials (Bowman, 2001). Aflakpui (1995) argued against using the CV for evaluating a whole experiment, but just for individual variates.

The idea behind the CV is that the standard variation is proportional to the mean at a constant rate. Bowman and Watson (1997) indicated that this is the case particularly with data that are enumerated (counted), in which values on the lower end are bounded by zero and have less space to vary than values on the high end. Although many authors (Ludvwig and Bannerjee, 1960; Snedecor and Cochran, 1980) have indicated that the assumption of the proportionality between the standard deviation and the mean is quite appropriate for many biological variables, this contradicts one of the normal distribution properties that the mean and the variance are independent of each other. This leads to a contradictory situation as Bowman and Watson (1997) have argued. If the standard deviation is proportional to the mean in different experiments, therefore this kind of relationship may most likely exist among treatments within the same experiment. If this is the case for any given experiment and the premise behind the CV is true, this would lead to the conclusion that in experiments with varying treatment means, they probably posses heterogeneous variances. This violates one assumption of the analysis of variance (Draper and Smith, 1981).

If dataset violates variance homogeneity assumption, this warrants applying transformation as a remedy before performing ANOVA. This, however goes against the rationale behind using the CV that the standard deviation increases with the size of the mean. Transformations often have the effect of reducing the dependency of the standard deviation on the mean. In experiments where the data have been subjected to transformation, it is inappropriate to use the CV as a measure of relative variation since the standard deviation is not any more dependent on the size of the mean (Bowman and Watson, 1997).

Kendall and Stewart (1977) pointed out two serious problems concerning the CV. First, there is no upper bound to the CV, and second, it is very much affected by the mean. Moreover, the CV fails to account for explaining relative variation for those cases where standard deviation is not proportional to the mean.

A constant CV across field trials implies a relationship between the error variance and the mean such that the slope = 2.0 is resulted in the regression of natural logarithm of error on that of the mean (Bowman and Rawlings, 1995; Bowman and Watson, 1997). The use of the CV is only valid if the slope (the b value) of the line equals 2.0 (Bowman and Rawlings, 1995). To investigate this relationship, one must examine historical data for any particular crop to determine whether the relationship between the error variance and mean holds; hence the CV is appropriate to apply (Bowman, 2001).

Braun *et al.* (1992) analyzed 19 yr of data from the International Maize and Wheat Improvement Center (CIMMYT), heritability and error variance increased and the CV was inversely proportional to mean of the trial. They reported significant negative correlation of - 0.52 and - 0.61 between trial mean and CV for trials with and without abiotic stress, respectively. Standard deviation of CV increased with increasing mean yield of the trial. The ln error variance regressed on ln mean, for several crops from North Carolina crop performance trials, led to regression coefficients ranged from - 0.112 for full season maize (*Zea mays* L.) to 1.70 for oats (*Avena sativa* L.) (Bowman and Rawlings, 1995). Also, regression coefficient data from Mississippi crop performance trials for the abovementioned crops –calculated from 5 to 7 years of data—ranged from 0.272 for Maturity Group VI soybean [*Glycine max*(L.) Merr.] to 1.114 for oats (*Avena sativa* L.) as reported by Bowman and Watson (1997). Allen *et al.* (1978) reported values of 0.435 for soybean and of 1.306 for oats. No regression coefficient, in all these studies, approached 2.0, the value that supports the assumption that the CV is a valid statistical tool for comparing trials' relative variations. The practice of applying the CV to discard trials with questionable results should be abandoned (Bowman, 2001).

Data transformation (e.g., square root, logarithmic, angular, inverse, and addition of a constant) tend to lower CV values in most cases, but can cause dramatic increases, depending on the nature of the variance and the specific transformation (Bowman and Watson, 1997). To demonstrate the effects of various transformations on the error mean square (EMS), the CV, and the coefficient of determination (R^2), they used six hypothetical sample datasets. Transformations to normalize or stabilize the variances may result in improvement in both CV and R^2. The first four transformations listed above, which are often used to stabilize or normalize variances, resulted in large changes in the CV in some datasets.

The objectives here were to: (i) check for the relationship among CV, EMS, and mean based on calculating the regression coefficients resulted from the regression of ln (EMS) on ln (mean) of both wheat (*Triticum aestivum* L.), and maize grain yields (ii) study the effect of data transformation on EMS, CV, and R^2.

MATERIALS AND METHODS

Establishing the Relationship between Error Variance and Mean Wheat and Maize Grain Yield Data

To fulfill the first objective of this paper, and since there are no databases available in research centers for agronomic crop performance experiments, I had to refer to data from various published articles in Egyptian refereed journals as well as from both PhD and Master's theses that handled both wheat and maize crops spanning the period from early 1990s up to now. The CV values, grand yield mean, in addition to error mean square values were needed from ANOVA tables. Unfortunately, neither journal article nor thesis offered these pieces of information, except means. The error degrees of freedom were calculated based on the proposed experimental design, and then we referred to the values of the LSD at $\alpha = 0.05$, to calculate the EMS based on:

$$LSD_{\alpha} = t_{\alpha/2, df_e} \sqrt{\frac{2EMS}{K}}$$

Where K value depends on the experimental design. It represents a multiplicative factor represents replicates, and any independent factors not involved in the measured effect. To regress ln EMS on ln mean, 57 data points were used for wheat, and 58 for maize. The experimental design often used in these field trials was the randomized complete block for one-factor or factorial experiments.

The Relationship between Error Variance and Mean

The proposed model to relate the variance to the mean is based on the observed phenomenon that the variance is proportional to a power of the mean,

$$\sigma_i^2 = a\mu_i^{\beta}$$

Where σ_i^2 and μ_i are true error variance and true mean, respectively, for the ith trial. Thus a linear regression of $\ln(S_i^2)$ on $(\ln(\overline{X}_i)$ provides estimates of $\beta_0 = \ln(a)$ and β, and a test of significance of the relationship; S_i^2 and $\overline{X}_i$ are the observed error variance and mean, respectively, for the ith trial of a particular species in the historical data. If the regression is not significant, the error variance is declared to be independent of the mean (Bowman and Rawlings, 1995).

The following equation relates the error variance to the mean yield if the CV is assumed constant. The CV is calculated as the standard deviation (SD_i) or square root of the estimate of EMS divided by the mean $(\overline{X}_i)$. The error variance S_i^2 is proportional to the square of $(\overline{X}_i)$ as follows:

$$CV_i = S_i / \overline{X}_i = (EMS)^{1/2} / \overline{X}_i$$

$S_i^2 = CV_i^2 \overline{X}_i^2$ or $EMS = CV_i^2 \overline{X}_i^2$

$\ln(S_i^2) = 2\ln(CV) + 2\ln(\overline{X})$

Thus, the regression of $\ln(EMS)$ on $(\ln(\overline{X}_i)$ produces a regression coefficient (slope) of 2.0 if and only if the CV is proportional to the mean. The use of the CV for evaluating the validity of experiments is questionable if the average relationship between $\ln(EMS)$ and $(\ln(\overline{X}_i)$ is not roughly linear with a slope of approximately 2.0 (Bowman and Rawlings, 1995).

Data Transformation

Barley Cultivar Biomass Data

A data subset of biomass yield of eight barley (*Hordeum vulgare* L.) cultivars in four randomized complete blocks was used to assess effects of various data transformation methods on each of EMS, CV, and R^2. The initial field experiment was laid out in a 4x4 partially balanced lattice design in four replicates of 16 barley cultivars. A 2-year field experiment was carried out at Moshtohor Experiment Center, Kalubia in 2000 and 2001. The data used in the current research was taken from the Year 2000's data (Tageldin, 2004).

Data Transformation

Transformations are generally applied to biological data as a remedy for any violations of ANOVA assumptions –non normal distribution, error heterogeneity. To create such situations of any of these violations, barley data were subjected to the addition of constant values trying to create such situations of violations that may influence each of the EMS, CV, and R^2. Bowman and Watson's (1997) technique was used which they had applied to their artificial datasets.

Six datasets are initiated from the original barley cultivar biomass yield (Table 1). Datasets 1, 2, 3 have low, intermediate, and high means, respectively, and homogeneous variances. Dataset 4 has the same mean as Dataset 2 but represents a wider range of data points and therefore greater treatment variances. Datasets 5 and 6 have nearly equal means, but error variances are not homogeneous. Dataset 5 has a variance proportional to the mean and Dataset 6 has a variance inversely proportional to the mean. All six datasets were subjected to square root, angular (arcsine square root), logarithmic (see Bartlett, 1947), inverse, addition of a constant (change of origin), and linear transformations. A constant equals 10.0 was added to each data point, and for linear transformation data were multiplied by 2.222 to switch units from kg 4.5 m^{-2}(plot area) to t ha^{-1}. All datasets –none transformed and transformed—were analyzed as randomized complete block design.

The coefficient of multiple determination (R^2)

It is a measure of the amount of variation about the mean dependent variable explained by the fitted model (Draper and Smith, 1981). It can be calculated as:

$$R^2 = \frac{\underline{b}'\underline{X}'\underline{Y} - n\bar{Y}^2}{\underline{Y}'\underline{Y} - n\bar{Y}^2} = 1 - \frac{RSS_p}{CTSS}$$

Where $\underline{b}'\underline{X}'\underline{Y}$ is the uncorrected sum of squares due to fitted model, $\underline{Y}'\underline{Y}$ is the uncorrected total sum of squares, $n\bar{Y}^2$ is the correction factor, RSS_p is the error sum of squares, $CTSS$ is the corrected total sum of squares.

Table (1): Biomass yield of 4 barley cultivars and other 6 modified yield-based datasets for comparing effects of different transformations on CV and R²†.

Treatment	Reps	Biomass	١	٢	٣	٤	٥	٦
		Values before transformation						
		kg						
1	1	7.4	7.4	14.4	44.4	4.4	7.4	2.8
	2	7.9	7.9	17.9	50.9	4.9	7.9	4.2
	3	7.6	7.6	20.6	56.6	4.6	7.6	3.8
	4	8.5	8.5	24.5	63.5	5.5	8.5	13.8
2	1	7.3	9.2	24.3	64.3	15.2	27.3	20.4
	2	6.9	8.9	26.9	69.9	14.9	26.9	21.7
	3	7.4	9.4	30.4	76.4	15.4	27.4	24.8
	4	9.0	11.1	35.0	84.0	17.1	29.0	30.1
3	1	7.3	9.4	34.3	84.3	45.4	37.3	32.8
	2	7.4	9.6	37.4	90.4	45.6	37.4	35.1
	3	9.4	11.6	42.4	98.4	47.6	39.4	41.2
	4	9.5	11.8	45.5	104.5	47.8	39.5	39.5
4	1	6.9	10.0	43.9	103.9	70.0	54.9	58.0
	2	5.9	9.1	45.9	108.9	69.1	55.9	54.1
	3	7.6	10.9	50.6	116.6	70.9	57.6	53.7
	4	6.9	10.3	53.9	122.9	70.3	58.9	58.5
Mean		7.7	9.5	34.2	83.7	34.3	32.7	30.9

† All datasets 1-6 are linear functions of the original seed yield data. Datasets 1, 2, 3 have low, intermediate, and high means, respectively, and homogeneous variances. Dataset 4 has the same mean as Dataset 2 but represents a wider range of data points and therefore a greater treatment variances. Datasets 5 and 6 have nearly equal means, but error variances are not homogeneous. Dataset 5 has a variance proportional to the mean, and Dataset 6 has a variance inversely proportional to the mean.

RESULTS AND DISCUSSION

Establishing the Relationship between Error Variance and Mean

The regression coefficients for both wheat and maize were numerically different (Table 2). The regression of ln (EMS) on ln(mean) produced a significant $\hat{\beta} = 2.798$ for wheat, and a non significant estimate of 1.139 for maize (Table 2). Both slopes were quite away from a slope of 2.0 which means that error variance and mean do change together at a constant rate. As a result the CV is not a valid tool to measure relative variability (Bowman and Watson, 1997; Bowman, 2001).

Table (2): Regression parameter estimates for wheat and maize seed yield (tha^{-1}) field trials.

Crop	Parameter estimate		$p > \|T\|$		Data points
	$\hat{\beta}_0 \pm SE$	$\hat{\beta}_1 \pm SE$	$\hat{\beta}_0$	$\hat{\beta}_1$	
Wheat	-6.078 ± 0.8498	2.798 ± 0.5306	0.0001	0.0001	57
Maize	-3.966 ± 1.780	1.139 ± 0.941	0.0299	0.2314	58

However, both slope estimates indicate a positive relationship between the error variance and the mean. Both intercept estimates $\hat{\beta}_0$ were negative and significant (-6.078 for wheat, and -3.966 for maize) (Table 2). In the absence of the relationship between MSE and $\bar{X}$ that makes the CV a valid tool (i.e., the slope estimate should approach a value of 2.0), the crucial relationship between ln CV and mean ln EMS-- as indicated by the high negative intercept estimates (Table 3) for both crops-- should not be overlooked. In this study, correlation coefficients of 0.805 and 0.854 were found between MSE and CV for wheat and maize, respectively. By analyzing wheat an maize trials from CIMMYT, Braun *et al.* (1992) reported a reduction in CV as both trial mean and error variance went up. They also found that standard deviation of CV was positively affected by means getting great.

Regression coefficient or slope estimates considerably less than 2.0 indicate that the likeliness of obtaining lower CV values would increase as trial mean increases (Bowman, 2001). Applying this argument to our coefficient estimate of 1.139 for maize, although not significant, suggests that maize trials had higher means which led to lowering CV values in these trials. The mean and the CV had a correlation of -0.191 for maize. By analogy, we may say that if this β estimate, on the other hand, was much greater than this specific value of 2.0, this would most probably suggest that achieving higher CVs may be obtained with higher trial EMS. In case of wheat, we obtained a value of $\beta_1 = 2.798$ which is greater than 2.0; this would indicate how error variance played a crucial role in determining the values of CVs.

Using a hypothetical data to show the triangular relationship between error variance, mean and coefficient of variation, Bowman (2001) compared the differences in the change of ln(mean) with ln(EMS) to keep having a constant CV, e.g., =10.0 %. The range in ln(EMS) is twice the size of the change of ln(mean) (2.78 vs. 1.39), therefore resulting in a β_1 value of 2.0.

Hence, Accepting or rejecting a trial's data based on just the CV value, being greater than an arbitrary figure (e.g., >15%), considered a hasty decision. Snedecor and Cochran (1980) pointed out that the CV is more difficult to interpret unless both error variance and mean values are available. Not only trials with high CVs that cast some doubt on trial's validity, but also those with extremely low ones. Because Bowman (2001) argued that thè CV is highly influenced by the mean, checking trial's mean is a crucial step during the evaluation process especially when the CV is exceptionally low.

Bauman's argument is bases on an assumption of the validity of the CV as a measure of relative variability, i.e., mean is proportional to EMS. However, in real life situations this relationship has not so far been proven to be true for many studied crops as shown in the literature even under nearly homogeneous field experimental conditions. In Egypt, field trials face many inherent heterogeneity setbacks which cause error mean square to rise. Under the same region, it is likely to get quite different CVs in consequent years using the same independent variables. This requires researchers' attention to thoroughly check his/her trial's mean and error variance before deciding to invalidate results.

Data Transformation and Error Mean Square (EMS), Coefficients of Variation (CV) and of Determination (R^2)

Data transformation is a basic tool to handle problems associated with ANOVA assumption violations. A classical paper by Bartlett (1947) had a good discussion of the theoretical basis of various transformations. Researcher should choose one or more transformations if he/she thinks there is a good reason for that. Some researchers apply data transformation even though they may not be aware they do (Bowman and Watson, 1997). They further added that the CV as a measure of relative variation is sensitive to data transformation.

The first four modes of transformations listed above and in (Table 3) – square root, arcsine, logarithmic, and inverse – are used to stabilize variances and reduce the dependency of the variance on the mean (Bowman an Watson, 1997). Notice that reduction of error variance–mean dependency contradicts the basic foundation of the CV.

The CV value has greatly been affected by various transformations within and among the six datasets (Table 3). In the first three datasets, where they greatly differ in mean values (Table 1), but have homogeneous variances (Table 3); within each dataset, the CV value has dropped by using square root, arcsine, and log transformations (Table 3). This result was also true when heterogeneity of treatment variances exists as in dataset 4, and when variance is proportional to mean as in dataset 5. In contrast, the inverse transformation in all datasets dramatically caused a marked increase in CVs. In dataset 6, the CV was about 12-fold increase of that in dataset 1 (72.45 vs. 5.86 %). This would indicate how researcher's decision to use this kind of transformation, if he/she has had heterogeneous data variances, may seriously inflate the CV value. In their hypothetical datasets, Bowman and Watson (1997) obtained a CV of 197.2% for inverse transformation of dataset having serious variance heterogeneity problem.

Table (3): Error mean square (EMS), coefficient of variation (CV), and coefficient of determination (R^2) from ANOVA of six sample datasets.

Transfor-mation	Dataset 1			Dataset 2			Dataset 3		
	EMS	CV	R^2	EMS	1.62	99.8	EMS	CV	R^2
		---%---			---%---			---%---	
None	0.40784	6.69	86.5	0.31006			0.31006	0.66	99.9
Sq. root	0.00992	3.23	87.3	0.00644	1.39	99.6	0.00534	0.80	99.8
Arcsine.	0.03290	3.24	87.3	0.02070	1.37	99.6	0.01607	0.76	99.8
Log	0.00073	2.78	88.3	0.00064	1.68	98.6	0.00021	0.77	99.2
Inverse	0.00003	5.86	89.8	0.00002	13.70	93.8	<0.0001	7.57	96.7
+ Constant	0.40784	3.26	86.5	0.31006	1.25	99.8	0.31006	0.59	99.9
Linear	2.01525	6.70	86.3	1.53122	1.62	99.8	1.53136	0.66	99.9
Transfor-mation	Dataset 4			Dataset 5			Dataset 6		
	EMS	CV	R^2	EMS	CV	R^2	EMS	CV	R^2
		---%---			---%---			---%---	
None	0.40784	1.86	99.9	0.51840	2.20	99.9	8.25229	9.29	98.6
Sq. root	0.00444	1.24	99.9	0.00240	0.89	99.9	0.17613	8.05	97.3
Arcsine.	0.01485	1.25	99.9	0.00824	0.91	99.9	0.58379	8.04	97.4
Log	0.00041	1.51	99.8	0.00011	0.74	99.9	0.01954	10.36	93.7
Inverse	0.00008	12.31	99.1	0.00001	6.81	99.6	0.00335	72.45	82.0
+ Constant	0.40784	1.44	99.9	0.51840	1.68	99.9	8.25229	7.02	98.6
Linear	2.01557	1.86	99.9	2.56024	2.20	99.9	40.7546	9.29	*98.6*

Addition of a constant (change of the location of the mean) can also lead to changes in the CV values. Algebraically, addition of a constant changes the mean by the same value of the constant added to/subtracted from the original data points without altering the variance; the greater the constant, the smaller the CV. Thus, the CV is extremely sensitive to changing mean location. For all six datasets, adding a constant to all observations caused a drop in all CV values despite differences in the nature of the data of all datasets (Table 3).

All dependent variables that rely on calculating time duration, e.g., physiological maturity, harvesting or flowing dates, provide a common example of adding a constant. Such variables may be expressed using numerous ways depending on the origin as Bowman and Watson (2001) have indicated. Dates may often be expressed as days after planting (DAP), calendar date (day of the year) in a format such as YYMDD, where YY represents the 2-digit year, and M for month, and DD for day). Some researchers often analyze dates as SAS dates (days since 1 January 1960) (SAS Inst., 1988). These variations in methods of estimating dates can lead to inflation in CV values. Drawing sound inferences about relative variability of various experiments having different origins is not reliable using CVs.

In field experiments we usually transform our field data based on experimental unit area to kilogram or ton per unit area (e.g., kg fa^{-1} or t ha^{-1}). This linear transformation does not affect CV (Table 3) since multiplication by a constant changes the mean and the variance by multiplying by the constant for the first, and by squaring the constant for the latter. This results in the same origin when calculating the CV.

With all these problems which arise with applying the CV tool of judging trials, this calls for finding other candidates of validity tools. The coefficient of (multiple) determination (R^2) can be used. In datasets 1, 2, 3 of homogenous variances, but different in means, there was an improvement in R^2 values toward the higher means (Table 3). Bowman and Watson (1997) stated that "The R^2 generally erred on the conservative side, in that it almost always decreased with transformation when variances were homogeneous...." By looking at the R^2 values in their three datasets, with homogeneous errors, one can not approve what they have stated above. The R^2 values did not markedly decrease by applying various transformations to data of these three sets.

In the mean time, transformations within these three sets 1, 2, and 3 did not cause any noticeable change in their values relative to the severe change in the CVs. Values of R^2 for datasets 4 and 5 were quite high and that results when there is a wide range in data points (Cornell and Berger, 1987) as in dataset 4; however, in dataset 5– where variance is proportional to the mean-- the R^2 values were quite similar to those of dataset 4.

These high R^2 values in dataset 5, despite of the various transformations, were not much like what Bowman and Watson (1997) obtained for a set with variance proportional to the mean. They obtained extremely low R^2 values. One possible reason for this dissimilarity may refer to their using a small hypothetical data sets (3 treatments in 3 replicates) since any reduction in the size of data point may cause reduction in R^2 values. From 22 maize trials carried out in North Carolina in 1995, the R^2 values did go up in 12 trials when a single replicate out of five was randomly excluded (Bowman and Watson, 1997). According to them, and based on the definition formula of R^2 mentioned earlier, it appears that it is a function of the error sum of squares quite like the CV; hence, they both may change as the size of the dataset increases. Dislike the CV, the R^2 (Table 3) is not affected by adding a constant within any particular dataset since it is independent of the unit of measure. Linear transformation also has no effect on R^2.

Some researchers prefer to use the adjusted R^2 (R^2_{adj}) values instead of the R^2. It is calculated according to Draper and Smith (1981) as:

$$R^2 adj = 1 - \frac{(RSS_p)/(n-p)}{(CTSS)/(n-1)} = 1-(1-R^2)\frac{(n-1)}{(n-p)}$$

It can be used to relatively compare equations fitted not only to a specific set of data but also to two or more entirely different sets of data (Draper and Smith, 1981),i.e., to compare different experiments of different sizes (Rawlings, 1988). One shortcoming of the R^2_{adj} is that it is influenced by the size of the differences among treatment means since larger difference tend to result in larger R^2_{adj} values.

Assessment of the validity of a field trial, or comparing the relative variability of more than one trial requires not just one statistical tool to base our decision on keeping or discarding data. Unfortunately, there has not been so far a clear-cut way to enable researchers make a decision of keeping or discarding data . A researcher should apply the CV, R^2, and R^2_{adj} or EMS to determine the validity of a trial. Trials must possess both questionable high CV and quite low R^2 or R^2_{adj}, as suggested by Bowman and Watson (1997), before being discarded, and therefore not being published. These statistical measures are considered relative and should be subjected to a considerable thinking. They certainly help to identify problematic data; data that may have been subjected to unsuitable environmental conditions, have erroneously been handled and collected, or have incorrectly been analyzed.

REFERENCES

Aflakpui, G.K.S. (1995): Some uses/abuses of statistics in crop experimentation. Trop. Sci. 35:347-353.

Allen, F.L.; Comstock, R.E. and Rasmusson, D.C. (1978): Optimal environments for yield testing. Crop Sci. 18:747-751.

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

Bowman, D.T. (2001): Common use of the CV: A statistical aberration in crop performance trials. The Journal of Cotton Sciences 5:137-141.

Bowman, D.T. and Rawlings, J.O. (1995): Establishing a rejection procedure for crop performance data.

Bowman, D.T. and C.E. Watson. (1997): Measures of validity in cultivar performance trials. Agron. J. 80:860-866.

Braun, H.; Pfeiffer, W.H. and Pollmer, W.G. (1992): Environments for selecting widely adapted spring wheat. Crop Sci. 32:1420-1427.

Cornell, J.A., and Berger, R.D. (1987): Factors that influence the value of the coefficient of determination in simple linear and nonlinear regression models. Phytopathology 77:63-70.

Damon, R.A, and Harvey, W.R. (1987): Experimental design, ANOVA, and regression.

Draper, N.R., and Smith, H. (1981): Applied regression analysis. 2nd ed. John Wiley & Sons, New York.

Kendall, M., and Stewart, A. (1977): The advanced theory of statistics. Vol. I. Distribution theory. 4th ed. Macmillan, New York.

Ludvwig, W., and Bannerjee, V. (1960): On the empirical verification of Pearson's coefficient of variation. Biometrics 16:311-312.

Rawlings, J.O. (1988): Applied regression analysis: A research tool. Wadsworth & Brooks/Cole, Pacific Grove, CA, USA.

SAS Institute. (1988): SAS language guide: User's guide. Release 6.03 ed. SAS Inst., Cary NC.

Snedecor, G.W., and Cochran, W.G. (1980): Statistical methods. 7th ed. Iowa State Univ. Press, Ames, Iowa.

Tageldin, M.H.A. (2004): Spatial analysis of barley yield trial data : Different approaches. Proc. 4th Scientific Conference of Agricultural Sciences. 7-9 December, pp712-726. Assiut Univ.

معامل الاختلاف كوسيلة للحكم علي صحة التجارب الحقلية

محمد هانئ أحمد تاج الدين
قسم المحاصيل، كلية الزراعة بمشتهر ــ جامعة بنها

هناك أسئلة تدور حول فاعلية معامل الاختلاف كوسيلة إحصائية للحكم علي شرعية التجارب الحقلية، أو لمقارنة الاختلافات النسبية بين عدة تجارب. تعتمد شرعية معامل الاختلاف علي فرضية وجود تناسب بين الانحراف المعياري للخطأ التجريبي و المتوسط الحسابي. لو حدث تحويل للبيانات في حالة الشك بوجود أية انتهاكات لفروض تحليل التباين، يصير استخدام معامل الاختلاف غير شرعي، حيث أن أحد أهداف التحويل هو إلغاء ذلك الارتباط بين التباين و المتوسط الحسابي. يثمر إجراء الانحدار بين لوغاريتم تباين الخطأ التجريبي ولوغاريتم المتوسط الحسابي عن معامل انحدار=٢,٠ ، وذلك يدل علي ثبات العلاقة بينهما من تجربة لأخرى، ومن ثم شرعية استخدام معامل الاختلاف. أهداف هذا البحث هي:١) فحص تلك العلاقة بين معامل الاختلاف، تباين الخطأ التجريبي، و المتوسط الحسابي، وذلك بناء علي حساب معامل الانحدار الناتج من دراسة علاقة الانحدار بين لوغاريتم كل من المتوسط الحسابي و تباين الخطأ التجريبي، ٢) دراسة تأثير عدة تحويلات للبيانات علي تباين الخطأ التجريبي، معامل الاختلاف، ومعامل التحديد. أسفرت النتائج عن معامل انحدار=٢,٧٩٨ لمحصول القمح، و ١,١٣٩ لمحصول الذرة الشامية. من الواضح أن كل من التقديرين لم يقترب من القيمة السالفة الذكر لمعامل الانحدار=٢,٠ وهي القيمة التي تدعم شرعية استخدام معامل الاختلاف للحكم علي الاختلاف النسبي بين التجارب، وكذلك شرعية نتائج تجربة ما. ولدراسة هدف البحث الثاني تم تطبيق ستة أنواع مختلفة من التحويلات: الجذر التربيعي، معكوس جيب الزاوية، اللوغاريتم، المعكوس، إضافة قيمة ثابتة، وأخيرا تحويل خطي.، وذلك باستخدام المحصول البيولوجي من تجربة أجريت علي أصناف الشعير. انخفضت قيم معامل الاختلاف نتيجة استخدام التحويلات، وذلك تبعا لنوعية عدم تجانس التباين داخل البيانات وكذلك تبعا للأثر المحدث نتيجة استخدام تحويل معين. أحدث التحويل العكسي زيادة كبيرة في قيم معامل الاختلاف وصلت الي ١٢% في وجود بيانات بها مشكلة من المدى الواسع، وحوالي ٧٢% في حالة وجود ارتباط عكسي بين التباين والمتوسط ا لحسابي. كان معامل التحديد أقل حساسية نتيجة استخدام التحويلات المختلفة مقارنة بما حدث لمعامل الاختلاف مما يوضح أهمية تلك الوسيلة لاستخدامها للحكم علي شرعية التجارب سواء طبق علي تلك البيانات تحويل أم لا. تلك الوسيلة تتأثر سلبيا بحجم البيانات المستخدمة، ومن ثم يصعب استخدامها في تجارب تتباين في الحجم. يعتبر معامل التحديد المعدل أكثر فاعلية في حالة وجود اختلافات في أحجام التجارب. يتطلب تقييم شرعية تجربة حقلية ما، أو مقارنة الاختلافات النسبية لعدة تجارب استخدام أكثر من وسيلة إحصائية وذلك لاتخاذ قرار بالاحتفاظ أو التخلص من بيانات معينة. لابد لعدم نشر بيانات تجربة معينة أن يكون معامل الاختلاف كبير بدرجة كبيرة، ومعامل التحديد منخفض بدرجة ملحوظة. ليس متوفرا حتي الآن طريقة محددة المعالم تساعد الباحث علي اتخاذ مثل هذا القرار الحاسم، فالوسائل المذكورة آنفا وغيرها كثير جميعها وسائل نسبية و من ثم يجب أن تخضع قيمها للتفسير الدقيق من جانب الباحث فهي حتما تساعد في تحديد مشاكل البيانات، ولكنها ليست في حد ذاتها سمات تحدد التخلص من أو نشر بيانات ما.