ANALYSES FOR DIALLEL CROSSING AND MATCHED PAIRS SCHEMES WHEN INDIVIDUAL CONTRIBUTIONS ARE MEASURABLE

BU-246-M

W. T. Federer

July, 1967

Abstract

Several statistical analyses were proposed for diallel crossing or mixing schemes when the individual responses are available. For example, in an equal mixture of two varieties (or drugs) the yields for individual varieties are often obtainable. Statistical analyses were required to utilize the information in the individual yields of a mixture in addition to the already available analyses for the total yields of mixtures. Some unsolved statistical problems in this area were also discussed.

Biometrics Unit, Department of Plant Breeding and Biometry, Cornell University, Ithaca, New York.

ANALYSES FOR DIALLEL CROSSING AND MATCHED PAIRS SCHEMES WHEN INDIVIDUAL CONTRIBUTIONS ARE MEASURABLE

BU-246-M W. T. Federer July, 1967

In a diallel crossing scheme involving crosses among k lines, the yield of the combination or cross is available and statistical analyses have been devised to handle this situation. Now, suppose that the individual contributions of each line in a cross were measurable as in matched pair experiments. What statistical analyses should now be performed? Analyses for matched pair experiments are available, but these are not appropriate for genetic experiments, say (see H. A. David, Griffin's Statistical Monogram, No. 12, 1963). This situation actually does arise in treatment of illnesses where combinations of two drugs are given in two periods and response is measured in each period, in mixtures of two varieties where the individual yields are available, or in other comparable situations.

First consider the following treatment design which we shall presume to be arranged in a randomized complete blocks design of r replicates:

Drug given in first period	Drug given in second period: 1		2		3		...	k		Totals	
	Period 1	Period 2	Period 1	Period 2	Period 1	Period 2		Period 1	Period 2	Period 1	Period 2
1	$Y_{1111\cdot}$	$Y_{1112\cdot}$	$Y_{1211\cdot}$	$Y_{1222\cdot}$	$Y_{1311\cdot}$	$Y_{1332\cdot}$		$Y_{1k11\cdot}$	$Y_{1kk2\cdot}$	$Y_{1\cdot 11\cdot}$	$Y_{1\cdot\cdot 2\cdot}$
2	$Y_{2121\cdot}$	$Y_{2112\cdot}$	$Y_{2221\cdot}$	$Y_{2222\cdot}$	$Y_{2321\cdot}$	$Y_{2332\cdot}$		$Y_{2k21\cdot}$	$Y_{2kk2\cdot}$	$Y_{2\cdot 21\cdot}$	$Y_{2\cdot\cdot 2\cdot}$
3	$Y_{3131\cdot}$	$Y_{3112\cdot}$	$Y_{3231\cdot}$	$Y_{3222\cdot}$	$Y_{3331\cdot}$	$Y_{3332\cdot}$		$Y_{3k31\cdot}$	$Y_{3kk2\cdot}$	$Y_{3\cdot 31\cdot}$	$Y_{3\cdot\cdot 2\cdot}$
⋮											
k	$Y_{k1k1\cdot}$	$Y_{k112\cdot}$	$Y_{k2k1\cdot}$	$Y_{k222\cdot}$	$Y_{k3k1\cdot}$	$Y_{k332\cdot}$		$Y_{kkk1\cdot}$	$Y_{kkk2\cdot}$	$Y_{k\cdot k1\cdot}$	$Y_{k\cdot\cdot 2\cdot}$
Totals	$Y_{\cdot 1\cdot 1\cdot}$	$Y_{\cdot 112\cdot}$	$Y_{\cdot 2\cdot 1\cdot}$	$Y_{\cdot 222\cdot}$	$Y_{\cdot 3\cdot 1\cdot}$	$Y_{\cdot 332\cdot}$		$Y_{\cdot k\cdot 1\cdot}$	$Y_{\cdot kk2\cdot}$	$Y_{\cdot\cdot\cdot 1\cdot}$	$Y_{\cdot\cdot\cdot 2\cdot}$

A table of totals over periods and blocks would be:

Drug given in first period	Drug given in second period 1	2	3	...	k	Totals
1	$Y_{11\cdots}$	$Y_{12\cdots}$	$Y_{13\cdots}$		$Y_{1k\cdots}$	$Y_{1\cdots\cdot}$
2	$Y_{21\cdots}$	$Y_{22\cdots}$	$Y_{23\cdots}$		$Y_{2k\cdots}$	$Y_{2\cdots\cdot}$
3	$Y_{31\cdots}$	$Y_{32\cdots}$	$Y_{33\cdots}$		$Y_{3k\cdots}$	$Y_{3\cdots\cdot}$
⋮						
k	$Y_{k1\cdots}$	$Y_{k2\cdots}$	$Y_{k3\cdots}$		$Y_{kk\cdots}$	$Y_{k\cdots\cdot}$
Totals	$Y_{\cdot 1\cdots}$	$Y_{\cdot 2\cdots}$	$Y_{\cdot 3\cdots}$		$Y_{\cdot k\cdots}$	$Y_{\cdots\cdot\cdot}$

A breakdown of degrees of freedom for the above would be (The part above the dashed line utilizes totals over periods while that below utilizes differences between periods.):

Source of variation	Degrees of freedom
Total	$2rk^2$
CFM	1
Blocks	r-1
Treatments	k^2-1
(Partitioning here could follow various forms depending upon the nature of the treatment design. One form could be the one given by Bruce Griffing for selfs, crosses, and reciprocals.)	
Treatments x blocks	$(r-1)(k^2-1)$
Within combinations and blocks (on differences between period yields)	rk^2
Differences among yields for drug 1 in period one	k
Differences among yields for drug 2 in period one	k
Differences among yields for drug 3 in period one	k
⋮	⋮
Differences among yields for drug k in period one	k
Differences x blocks	$k^2(r-1)$
Differences for drug 1 x blocks	k(r-1)
⋮	⋮
Differences for drug k x blocks	k(r-1)

Alternative partitionings for degrees of freedom on differences between periods totalled over blocks (i.e. $Y_{fghi.}$) are:

Source of variation	Degrees of freedom
Differences for drug 1 in period two	k
Differences for drug 2 in period two	k
⋮	
Differences for drug k in period two	k

Source of variation	Degrees of freedom
Periods one vs. two	1
Homogeneity among row total differences	k-1
Homogeneity among column total differences	k-1
Interaction	$(k-1)^2$

Source of variation		Degrees of freedom
Between period totals in row one	k	
periods one vs. two		1
periods x drugs in columns of row one		k-1
Between period totals in row two	k	
period one vs. two		1
periods x drugs in columns of row two		k-1
⋮		
Between period totals in row k	k	
period one vs. two		1
period x drugs in columns of row k		k-1

Source of variation	Degrees of freedom
(Same as preceding except use columns instead of rows)	

Source of variation	Degrees of freedom
period one vs. two	1
period x drug in first period (rows)	k-1
period x drug in second period within drugs given in first period (rows)	k(k-1)

etc.

Now let us suppose that there is no first or second period, but that the two results for each combination are still available. An example where this would be the case would be in equal mixture of seed for pairs of varieties of beans or wheat. Here the yield of each variety would be available and the information contained in the individual yields should be utilized as well as the total yield of a mixture. In this situation the mixture of variety 1 with variety 2 is identical (except for sampling variation) to the mixture of variety 2 with variety 1 if the mixture is 50-50. In the event that the mixture were not 50-50 but some other combination, say 60-40 of variety 1 with variety 2 and 60-40 of variety 2 with variety 1, then the two mixtures would not be identical and would need to be considered as for the drug experiment with two periods. If the mixtures of variety f with variety g and variety g with variety f are identical apart from sampling variation, then the $k(k-1)/2$ combinations (the reciprocals) below the diagonal in the first table would be ommited in many experiments. With the omission of reciprocals a breakdown in the degrees of freedom in the analysis of variance would be:

Source of variation	Degrees of freedom	
Total	$2rk(k+1)/2$	
CFM	1	
Blocks	$r-1$	
Treatments	$(v-1)=(k^2+k-2)/2$	
General mixing effect		$k-1$
Specific mixing effect		$k(k-1)/2$
Blocks x treatments	$(r-1)(v-1)$	
Within combinations or treatments	$k(k+1)/2$	
Variety f yields vs. yield all var. with f (Σ variety f) vs. (Σ yields all var. with f)=A		k
		1
Interaction of general mixing effects and A		$k-1$
A x specific mixing effects		$k(k-1)/2$
Within combinations x blocks	$v(r-1)$	

If the selfs were omitted the above analysis would be appropriate with an appropriate change in degrees of freedom.

The above analyses represent a start on the problem of partitioning degrees of freedom for pairs of mixtures. Other analyses are needed to compare, for example, totals like $Y_{f....}$ with totals $Y_{.g...}$, $Y_{f....}$ with $Y_{f.h..}$, $Y_{f.h..}$ with $Y_{.gh..}$, etc. Also, the problem of using p drugs or p varieties in a mixture where the p responses are measurable for each combination, requires investigation. This corresponds to triallel, quadrallel, etc. crosses or mixtures. The development of analyses for 4 or higher way crosses and construction of treatment designs utilizing fractional replication or other procedures to construct the designs and their analyses require investigation.