The Canadian Journal of Statistics
Vol. 21, No. 1, 1993, Pages 97–106
La Revue Canadienne de Statistique

Detecting the interaction of quantity and quality: A study of Fisher's analysis

Gordon Hilton FICK

University of Calgary

Key words and phrases: Experimental design, orthogonal polynomials, interaction
AMS 1985 subject classifications: Primary 62K15; secondary 05B05.

ABSTRACT

In Sections 49 and 50 of the *Design of Experiments*, Fisher discusses an experiment designed to compare the effects of several types of manure on yield. Each type of manure is applied at three dosage levels: zero, single, and double doses. Fisher points out that the usual contrasts constructed for a factorial experiment are unsatisfactory in this setting. In particular, since the response curves necessarily meet at the zero dose, the usual notion of interaction as a lack of parallelism cannot apply. Fisher then gives an appropriate definition for interaction in this setting. This paper is concerned with a class of orthogonal polynomials that can be used as an aid in the detection of this modified definition of interaction.

RÉSUMÉ

Dans les sections 49 et 50 de *Design of Experiments*, Fisher discute d'une expérience conçue afin de comparer les effets de plusieurs types d'engrais sur le rendement. Chaque type d'engrais est appliqué à trois différents dosages: pas d'engrais, dose simple et dose double. Fisher indique que les contrastes habituels pour une expérience factorielle ne sont pas satisfaisants dans ce cadre. En particulier, puisque les courbes de réponse se coupent forcément à la dose zéro, la notion habituelle d'intéraction comme absence de parallélisme ne peut pas être utilisée. Fisher présente alors une définition d'intéraction appropriée à ce cadre. Cet article examine une classe de polynômes orthogonaux qui peuvent aider à détecter cette définition d'intéraction modifiée.

1. INTRODUCTION

Consider the administration of a graded series of doses of some preparation and the measured responses to such a series. Suppose that this preparation is diluted with some inert material, so that a given dose now contains less of the active ingredient than the same dose of the original preparation. The dose-response curve of this diluted preparation will be the same as that of the original preparation except that it is plotted on a different dose scale.

More generally, let us suppose that we wish to compare an unknown (or test) preparation with a standard one. Suppose the doses of the standard are denoted x_1, and of the unknown, x_2. If a transformation of the form $x_2 = kx_1$, with k properly chosen, brings the response curves into coincidence, then one preparation can be viewed as a diluted version of the other. The value of k that accomplishes this is called the relative potency of the unknown (relative to the standard).

If, in any instance, the curves cannot be brought into coincidence by the transformation $x_2 = kx_1$, the preparations are qualitatively different and we have, in Fisher's words, an

TABLE 1: Fisher's analysis.

	y_{10}	y_{20}	y_{11}	y_{21}	y_{12}	y_{22}
z_1	1	1	1	1	1	1
z_2	−1	−1	0	0	1	1
z_3	−1	−1	2	2	−1	−1
z_4	0	0	−1	1	−2	2
z_5	0	0	2	−2	−1	1
z_6	−1	1	0	0	0	0

interaction of quantity and quality. Fisher gives an illustration of this notion in Section 50 of Fisher (1935). Yates gives illustrations in Section 15 of Yates (1937) and Section 6 of Yates (1933).

With Fisher's illustration, the preparations are different types of nitrogeneous manure applied at three dosage levels: zero, single, and double doses. For discussion here, suppose there are two types of manure and the yields of sugar beets are denoted y_{ij}, corresponding to the ith quality and the jth quantity. In this setting, there will be six combinations of type and dose applied. The structure of the experiment appears to be factorial. However, as is pointed out by Fisher, the usual contrasts constructed for a factorial experiment are unsatisfactory. In particular, since the response curves necessarily radiate from a point on the response axis, the interaction described above cannot be interpreted as a lack of parallelism.

Fisher suggests that to check for this interaction one can consider whether the differences $d_j = y_{2j} - y_{1j}$ are proportional to the quantities of nitrogen, i.e., d_2 should be close to $2d_1$. If the comparison $d_2 - 2d_1$ contains more than error, the curve in the differences has a nonlinear component.

The appropriate contrasts in Fisher's analysis can be placed in an orthogonal transformation. Subject to suitable normalizing, the z's in Table 1 display the implied comparisons. Notice that the component z_4 displays the linear component of the trend in the curve relating the differences to the dose, while the component z_5 displays the quadratic (or nonlinear) component of the trend of the curve relating the differences to the dose. The curve in the differences is constrained to be zero, corresponding to a zero dose. The components have values which are proportional to the values of the orthogonal polynomials x and $ax^2 + bx$ for $x = 1, 2$, where a and b satisfy $9a + 5b = 0$.

The next section describes a model for Fisher's analysis and suggests how his analysis can be extended to more than two nonzero doses.

2. MORE THAN TWO DOSES

Let $x_{1\alpha}$, $\alpha = 1, 2, \ldots, n_1$, be doses of the standard, yielding responses $y_{1\alpha}$, and let $x_{2\alpha}$, $\alpha = 1, 2, \ldots, n_2$, be doses of the unknown, yielding responses $y_{2\alpha}$. To describe two lines radiating from a point on the response axis, a regression equation will have

$$Y = b_0 + b_1 x + b_2 \delta x,$$

where $\delta = 0$ in sample 1 and $\delta = 1$ in sample 2. See Finney (1978, Ch. 7) for an example and for an outline of Fieller's method applied to give confidence limits for the relative potency.

To use the same doses for both preparations, $x_{1\alpha} = x_{2\alpha} = x_\alpha$ (say) and $n_1 = n_1 = n$ (say), ensures that the arrangement is completely orthogonal. A rearrangement of the

TABLE 2: More than two doses: transformation.

	y_{11}	y_{21}	$\cdots$	$y_{1\alpha}$	$y_{2\alpha}$	$\cdots$	y_{1n}	y_{1n}
z_1	1	1		1	1		1	1
z_2	$x_1 - \bar{x}$	$x_1 - \bar{x}$		$x_\alpha - \bar{x}$	$x_\alpha - \bar{x}$		$x_n - \bar{x}$	$x_n - \bar{x}$
$\vdots$				Orthogonal				
z_n	a_1	a_1		a_α	a_α		a_n	a_n
z_{n+1}	$-x_1$	x_1		$-x_\alpha$	x_α		$-x_n$	x_n
z_{n+2}								
$\vdots$				Orthogonal				
z_{2n}	$-b_1$	b_1		$-b_\alpha$	b_α		$-b_n$	b_n

TABLE 3: More than two doses: analysis of variance.

	df	ss		df	ss
Among levels	$n-1$	$z_2^2 + \cdots + z_n^2$	Average slope	1	z_2^2
			Residual	$n-2$	
Within levels	n	$z_{n+1}^2 + \cdots + z_{2n}^2$	Preparations	1	z_{n+1}^2
			Levels × preparations	$n-1$	

regression equation to take advantage of the orthogonality is to replace

$$Y = b_0 + b_1 x + b_2 \delta x$$

by

$$Y = B_0 + B_1(x - \bar{x}) + B_2 \delta' x, \qquad \delta' = \delta - \tfrac{1}{2}.$$

The normal equations are then

$$2nB_0 = \sum (y_{1\alpha} + y_{2\alpha}),$$

$$2\sum (x_\alpha - \bar{x})^2 B_1 = \sum (x_\alpha - \bar{x})(y_{1\alpha} + y_{2\alpha}),$$

$$\tfrac{1}{2} \sum x_\alpha^2 B_2 = \tfrac{1}{2} \sum x_\alpha (y_{2\alpha} - y_{1\alpha}).$$

The first two equations represent an analysis of the sums (or averages), and indeed they constitute the fitting of a regression line to the average of the pairs of y's on the same dose. The third equation is a fitting of a regression to the differences of the same pair of y's, constrained to pass through the origin. The two questions about the straightness of the two response curves have been changed to one about the straightness of the average curve and one about the straightness of the curve relating differences to dose.

The analysis thus far can be embedded in the orthogonal transformation, displayed in Table 2. The analysis-of-variance table corresponding to Table 2 is displayed in Table 3. The components $z_3, \ldots, z_n$ would be chosen to test the nature of the departures from linearity of the average points. Values of polynomials chosen to be orthogonal over the

doses used could be constructed. If we define

$$P_0 = \lambda_{00},$$
$$P_1 = \lambda_{10} + \lambda_{11}x,$$
$$P_2 = \lambda_{20} + \lambda_{21}x + \lambda_{22}x^2,$$
$$\vdots$$

with the λ's chosen so that $\sum P_i P_j = 0$, $i \neq j$, we could choose $\lambda_{00} = 1$, $P_1 = x - x$, etc. Then $z_3 = \sum P_2 y$ is the quadratic component, and so on. If the doses are arranged to be equally spaced, Fisher's ξ'-polynomials can be used.

In the same way, we may wish to choose the coefficients in $z_{n+2}, \cdots, z_{2n}$ to display the nature of the trend in the curve relating the differences to the dose. We could define polynomials

$$Q_1 = \mu_{11}x,$$
$$Q_2 = \mu_{21}x + \mu_{22}x^2,$$
$$\vdots$$

with the μ's chosen so that $\sum Q_i Q_j = 0$, $i \neq j$. These polynomials would be suitable for fitting polynomial regressions constrained to pass through the origin.

If we could arrange to have the doses not only equally spaced, but reducible to $1, 2, 3, \cdots$ by a change of scale, then a fixed set of polynomials could be used in a similar way to Fisher's ξ'-polynomials. These Q-polynomials are exhibited and illustrated in the next section.

3. THE FIRST FIVE POLYNOMIALS AND THEIR SUMS OF SQUARES

Now suppose that there are n doses, reducible to $1, 2, 3, \cdots, n$ by change of scale. The polynomials up to degree five are displayed in Table 4. Most standard references such as Snedecor and Cochran (1980) give Fisher's ξ-polynomials to this degree. Tables of the values of these polynomials for $n = 2$ to 12 are included in the appendix. Following Fisher, the Q-polynomials are defined with the coefficient of the highest power being one. The sums of squares of the values of the polynomials in Table 4 are given in Table 5.

4. ILLUSTRATIONS AND COMMENTS

To illustrate the use of these polynomials, suppose we have doses $1, 2, 3$, so that the polynomials with $n = 3$ can be used; then the transformation will be as displayed in Table 6. For this example, z_5 and z_6 display the interactions of doses and preparations. The example given in Section 1 has doses $0, 1, 2$, so that the polynomials with $n = 2$ have been used.

To compare three or more preparations, a familiar multiplicative rule can be used to obtain the interaction components. For example, suppose there is a standard preparation and two unknown preparations, each applied at three doses: $1, 2, 3$. The transformation in Table 7 will suffice. The components z_4 and z_5 display the quality components, i.e., the differences between preparations. The components z_6 to z_9 record the interactions between doses and preparations.

TABLE 4: The Q-polynomials.

$$
\begin{aligned}
Q_1(x) &= x \\
c_{22}Q_2(x) &= c_{22}x^2 + c_{21}x \\
c_{33}Q_3(x) &= c_{33}x^3 + c_{32}x^2 + c_{31}x \\
c_{44}Q_4(x) &= c_{44}x^4 + c_{43}x^3 + c_{42}x^2 + c_{41}x \\
c_{55}Q_5(x) &= c_{55}x^5 + c_{54}x^4 + c_{53}x^3 + c_{52}x^2 + c_{51}x \\
c_{22} &= 2(2n+1) \\
c_{21} &= -3n(n+1) \\
c_{33} &= 5(3n^2 + 3n + 2) \\
c_{32} &= -10n(n+1)(2n+1) \\
c_{31} &= 6n^4 + 12n^3 + 3n^2 - 3n + 2 \\
c_{44} &= 28(2n+1)(n^2 + n + 3) \\
c_{43} &= -105n(n+1)(n^2 + n + 2) \\
c_{42} &= 10(2n+1)(3n^4 + 6n^3 + 5n^2 + 2n + 6) \\
c_{41} &= -5n(n+1)(2n^4 + 4n^3 + n^2 - n + 18) \\
c_{55} &= 42(5n^4 + 10n^3 + 55n^2 + 50n + 24) \\
c_{54} &= -252n(n+1)(2n+1)(n^2 + n + 8) \\
c_{53} &= 70(6n^6 + 18n^5 + 55n^4 + 80n^3 + 71n^2 + 34n + 24) \\
c_{52} &= -70n(n+1)(2n+1)(n^4 + 2n^3 + 5n^2 + 4n + 24) \\
c_{51} &= 15n^8 + 60n^7 + 120n^6 + 150n^5 + 1145n^4 + 2110n^3 + 640n^2 - 400n + 192
\end{aligned}
$$

TABLE 5: The sums of the squares of the Q-polynomials.

$$
\begin{aligned}
\Sigma Q_1^2 &= \prod_{k=0}^{1} (n+k)(2n+1)/6 \\
\Sigma c_{22}^2 Q_2^2 &= \prod_{k=-1}^{2} (n+k)(2n+1)(3n^2 + 3n + 2)/30 \\
\Sigma c_{33}^2 Q_3^2 &= \prod_{k=-2}^{3} (n+k)(2n+1)(n^2 + n + 3)(3n^2 + 3n + 2)/42 \\
\Sigma c_{44}^2 Q_4^2 &= \prod_{k=-3}^{4} (n+k)(2n+1)(n^2 + n + 3)(5n^4 + 10n^3 + 55n^2 + 50n + 24)/90 \\
\Sigma c_{55}^2 Q_5^2 &= \prod_{k=-4}^{5} (n+k)(2n+1)(3n^4 + 6n^3 + 77n^2 + 74n + 120)(5n^4 + 10n^3 + 55n^2 + 50n + 24)/330
\end{aligned}
$$

If *a priori* comparisons between preparations are not available, then incomplete subdivisions of sums of squares can be obtained. See, for example, Cochran and Cox (1957). Notice that if

$$
\begin{aligned}
r_i &= -11y_{i1} - 8y_{i2} + 9y_{i3}, \\
s_i &= 3y_{i1} - 3y_{i2} + y_{i3},
\end{aligned}
$$

TABLE 6: Example with $n = 3$.

	y_{11}	y_{21}	y_{12}	y_{22}	y_{13}	y_{23}
z_1	1	1	1	1	1	1
z_2	−1	−1	0	0	1	1
z_3	−1	−1	2	2	−1	−1
z_4	−1	1	−2	2	−3	3
z_5	−11	11	−8	8	9	−9
z_6	−3	3	3	−3	−1	1

TABLE 7: Three doses.

	y_{11}	y_{21}	y_{31}	y_{12}	y_{22}	y_{32}	y_{13}	y_{23}	y_{33}
z_1	1	1	1	1	1	1	1	1	1
z_2	−1	−1	−1	0	0	0	1	1	1
z_3	1	1	1	−2	−2	−2	1	1	1
z_4	0	−1	1	0	−2	2	0	−3	3
z_5	−2	1	1	−4	2	2	−6	3	3
z_6	0	−11	11	0	−8	8	0	9	−9
z_7	0	−3	3	0	3	−3	0	−1	1
z_8	−22	11	11	−16	8	8	18	−9	−9
z_9	−6	3	3	6	−3	−3	−2	1	1

then $z_6^2 + z_8^2$ is proportional to

$$\sum (r_i - \bar{r})^2,$$

and $z_7^2 + z_9^2$ is proportional to

$$\sum (s_i - \bar{s})^2.$$

We are supposing that the experiment furnishes a suitable estimate of experimental error. Indeed, each of the illustrations discussed here requires that each y_{ij} be a sum (or average) of the same number of observations. If the number of doses is large, it may not be acceptable to carry out a repetition (or replication) of the experiment. In such a case, the polynomials (P's and Q's) could be fitted to a degree high enough to remove all trends. With caution and qualification, the remaining components could then be used to estimate error.

Suppose that each of the nonzero doses is repeated the same number of times, but the zero dose is not. Fisher's ξ'-polynomials are no longer suitable for the components between levels, but the Q-polynomials will still be appropriate for the components within levels. See Section III.3 of Claringbold (1959) for an illustration with two nonzero doses and the zero dose repeated with half the frequency of the nonzero doses. In this reference, the terms slope and blank are used for the linear and quadratic components between levels. In addition, this paper gives a collection of orthogonal contrasts derived using a technique due to Lorraine (1952). These contrasts could be useful in determining the doses where linearity might be an acceptable assumption.

The concepts associated with the fitting of response curves radiating from a point have also been studied by Bliss (1952, Chs. 7, 8), Williams (1959, Ch. 8), and Kempthorne (1952, Ch. 18).

TABLE 8: $n = 2$.

	Q_1	Q_2
	1	−2
	2	1
λ	1	$\frac{5}{2}$
S	5	5

TABLE 9: $n = 3$.

	Q_1	Q_2	Q_3
	1	−11	3
	2	−8	−3
	3	9	1
λ	1	7	$\frac{15}{6}$
S	14	266	19

TABLE 10: $n = 4$.

	Q_1	Q_2	Q_3	Q_4
	1	−7	71	−4
	2	−8	−3	6
	3	−3	−67	−4
	4	8	34	1
λ	1	3	$\frac{155}{6}$	$\frac{23}{8}$
S	30	186	10,695	69

APPENDIX: THE VALUES OF THE POLYNOMIALS

The values in Tables 8 through 18 are given as integers with no common factors. The necessary proportionality constants are given as λ's, along with sums of squares of the tabulated values, given as S's.

ACKNOWLEDGEMENT

The author would like to thank Professor Daniel B. DeLury for many valuable conversations and correspondence.

REFERENCES

Bliss, C.I. (1952). *The Statistics of Bioassay*. Academic Press, New York.
Claringbold, P.J. (1959). Orthogonal contrasts in slope ratio investigations. *Biometrics*, 15, 307–322.
Cochran, W.G., and Cox, G.M. (1957). *Experimental Designs*. Wiley, New York.
Finney, D.J. (1978). *Statistical Method in Biological Assay*. Griffen, London.
Fisher, R.A. (1935). *The Design of Experiments*. Hafner, New York.
Kempthorne, O. (1952). *The Design and Analysis of Experiments*. New York, Wiley.
Lorraine, P.K. (1952). On a useful set of orthogonal comparisons. *J. Roy. Statist. Soc. Ser. B*, 14, 234–237.

TABLE 11: $n = 5$.

	Q_1	Q_2	Q_3	Q_4	Q_5
	1	−17	103	−379	5
	2	−23	46	256	−10
	3	−18	−56	246	10
	4	−2	−88	−374	−5
	5	25	65	125	1
λ	1	$\frac{11}{2}$	$\frac{115}{6}$	$\frac{847}{12}$	$\frac{251}{120}$
S	55	1771	27,830	425,194	251

TABLE 12: $n = 6$.

	Q_1	Q_2	Q_3	Q_4	Q_5
	1	−25	47	−99	1,849
	2	−37	35	17	−2,315
	3	−36	−4	82	10
	4	−22	−38	12	2,300
	5	5	−35	−95	−1,843
	6	45	37	41	461
λ	1	$\frac{13}{2}$	$\frac{16}{3}$	$\frac{91}{12}$	$\frac{798}{5}$
S	91	5824	7488	27,664	17,677,296

TABLE 13: $n = 7$.

	Q_1	Q_2	Q_3	Q_4	Q_5
	1	−23	37	−879	3,043
	2	−36	35	−149	−2,272
	3	−39	11	587	−1,945
	4	−32	−18	552	1,980
	5	−15	−35	−205	2,209
	6	12	−23	−809	−3,008
	7	49	35	441	917
λ	1	5	$\frac{17}{6}$	$\frac{413}{12}$	$\frac{1799}{20}$
S	140	7140	6018	2,335,102	36,893,892

TABLE 14: $n = 8$.

	Q_1	Q_2	Q_3	Q_4	Q_5
	1	−91	329	−2,177	4,747
	2	−148	359	−907	−1,808
	3	−171	199	933	−3,793
	4	−160	−42	1,656	60
	5	−115	−255	765	3,801
	6	−36	−331	−1,047	1,648
	7	77	−161	−1,897	−4,627
	8	224	364	1,288	1,688
λ	1	17	$\frac{109}{6}$	$\frac{595}{12}$	$\frac{1231}{20}$
S	204	155,652	611,490	16,113,790	81,615,300

TABLE 15: $n = 9$.

	Q_1	Q_2	Q_3	Q_4	Q_5
	1	−58	1,358	−3,398	2,026
	2	−97	1,631	−2,053	−209
	3	−117	1,159	627	−1,549
	4	−118	282	2,412	−804
	5	−100	−660	2,250	858
	6	−63	−1,327	267	1,519
	7	−7	−1,379	−2,233	119
	8	68	−476	2,768	−1,936
	9	162	1,722	2,322	834
λ	1	$\frac{19}{2}$	$\frac{170}{3}$	$\frac{589}{12}$	$\frac{67}{5}$
S	285	85,272	13,217,160	45,145,672	14,695,512

TABLE 16: $n = 10$.

	Q_1	Q_2	Q_3	Q_4	Q_5
	1	−24	2,118	−966	876
	2	−41	2,726	−726	101
	3	−51	2,239	−41	−569
	4	−54	1,072	554	−539
	5	−50	−360	750	18
	6	−39	−1,642	464	554
	7	−21	−2,359	−161	539
	8	4	−2,096	−756	−146
	9	36	−438	−726	−816
	10	75	3,030	750	411
λ	1	$\frac{7}{2}$	$\frac{415}{6}$	$\frac{113}{12}$	$\frac{131}{40}$
S	385	19,173	40,235,910	4,233,658	2,845,713

TABLE 17: $n = 11$.

	Q_1	Q_2	Q_3	Q_4	Q_5
	1	−175	1,263	−417	3,063
	2	−304	1,713	−363	894
	3	−387	1,549	−103	−1,541
	4	−424	970	167	−2,156
	5	−415	175	320	−908
	6	−360	−637	298	1,040
	7	−259	−1,267	112	2,156
	8	−112	−1,516	−158	1,376
	9	81	−1,185	−363	−1,059
	10	320	−75	−285	−2,766
	11	605	2,013	363	1,617
λ	1	23	$\frac{199}{6}$	$\frac{23}{8}$	$\frac{279}{40}$
S	506	1,309,022	17,671,797	917,631	37,038,924

TABLE 18: $n = 12$.

	Q_1	Q_2	Q_3	Q_4	Q_5
	1	−209	363	−15,279	7,656
	2	−368	513	−14,799	3,351
	3	−477	497	−6,561	−2,662
	4	−536	362	3,289	−5,427
	5	−545	155	10,460	−3,936
	6	−504	−77	12,516	220
	7	−413	−287	8,876	4,200
	8	−272	−428	814	5,262
	9	−81	−453	−8,541	2,112
	10	160	−315	−14,205	−3,747
	11	451	33	−9,339	−6,666
	12	792	638	14,751	4,499
λ	1	25	$\frac{47}{6}$	$\frac{1855}{24}$	$\frac{1349}{120}$
S	650	2,352,350	1,781,065	1,431,369,936	250,428,360

Snedecor, G.W., and Cochran, W.G. (1980). *Statistical Methods*. Iowa State Univ., Ames, Iowa.

Williams, E.J. (1959). *Regression Analysis*. Wiley, New York.

Yates, F. (1933). Principles of orthogonality and confounding in replicated experiments. *J. Agricultural Sci.*, 23, 108–145.

Yates, F. (1937). *The Design and Analysis of Factorial Experiments*. Commonwealth Agricultural Bureaux, Harpenden, England.

Received 28 May 1990
Revised 8 October 1992
Accepted 28 October 1992

Department of Community Health Sciences
University of Calgary
Calgary, Alberta
T2N 4N1