SEQUENTIAL DESIGN OF EXPERIMENTS

Walter T. Federer, Cornell University

BU-237-M February, 1967

ABSTRACT

An exposition of analyses of variance procedures used in sequential experimentation with sequential hypothesis testing and with sequential estimation is presented. Several sequential test procedures, viz. a single F test in the analysis of variance, a set of F tests involving orthogonal single degree of freedom contrasts, and a set of t tests (or F tests) for all possible differences between pairs or contrasts or means, are presented for data from an experiment designed as a completely randomized design. Also, several error rate bases are considered. A numerical example is utilized to illustrate the procedures. Following a discussion of these procedures illustrating their extension to other experimental designs, analyses using a single F test are described for the randomized complete block design with blocks added sequentially, single v × v latin square added sequentially, v treatments designed in b blocks of size k (k usually but not necessarily less than v) with each treatment occurring r times and with sets of blocks being added sequentially, and k-row by b-column design for v treatments with r replicates on each treatment and with rows and/or columns added sequentially. The analyses presented were for fixed effects which was followed by a discussion of random and mixed effects cases for each of the designs presented.

Under sequential estimation a two stage sampling procedure is discussed along with results relating fixed sample size estimation procedures to sequential sampling estimation procedures. Some discussion is presented relative to unsolved problems in the sequential selection of an experimental design, the sequential selection of an analysis, and of the scale of measurement, or transformation. All of the discussion relates to sequential design and analysis of experiments with no presentation being made of sequential procedures for selecting treatments as, e.g. levels of the independent variate in regression, of levels for a dosage response, and related phenomena.

(Preliminary Draft)

SEQUENTIAL DESIGN OF EXPERIMENTS

Walter T. Federer, Cornell University

BU-237-M February, 1967

1. INTRODUCTION

Late in July, 1952, a number of pre- and post-doctoral students of statistics were gathered together with the late Sir Ronald A. Fisher to discuss research in and the future of statistics. I asked him what line of research he would pursue in life if he were one of us starting a career in statistics. He stroked his beard in a thoughtful manner and after a moment he replied, "This fellow Abraham Wald was a very ingenious person. His ideas on the sequential aspects of experimentation are very important. Yes, if I were a young man starting a career in statistics, I would work on the sequential design of experiments." This comment was made over 15 years ago; where do we stand today? Do we have anything like a coherent theory of sequential design of experiments which can be applied in practice? The purpose of this paper is to investigate these questions and to illustrate some available procedures.

We shall first take a look at sequential procedures for hypothesis testing under both fixed and random effects models for various experimental designs. In section 3 we shall consider sequential estimation procedures. In section 4 we shall discuss the relevance of some usual properties of sequential procedures from the experimenter's point of view.

The author undertook this subject not because he knows this field, but because he feels that it is very important for people to be thinking about sequential design of experiments and that it was important for him to learn

something more about this topic. The works of Johnson [1953,1961], Wetherill [1965], and Hall, Wijsman, and Ghosh [1965], were utilized extensively in preparing the following. The first two references are expository and attempt to cover the field. The last work is a comprehensive and sound theoretical justification of the work on hypothesis testing. These three references are in the must read class for anyone wishing knowledge in this area. No attempt was made to present a coverage of the topics in sequential design (see Jackson [1960] and Johnson [1961]), to assign priorities, etc. A few selected papers were utilized as reference material.

Furthermore, the discussion is confined to sequential design of experiments rather than to the sequential selection of treatments or levels of treatments. Selection of dosage levels, of the X values in regression, of varieties or drugs in screening experiments, and other treatment designs are not considered in the present discussion.

2. HYPOTHESIS TESTING

The relevance of hypothesis testing in experimentation is not being supported because a presentation is made of sequential procedures for doing this. To the contrary, the author is fairly well convinced that experimenters are often quite certain that the null hypothesis is untrue. Otherwise, they would not conduct the experiment. Then, since the null hypothesis is often not applicable, it is nonsense to test it. The same comments hold for a "region of indifference". However, hypothesis testing can be justified within the realm of the subject of Statistics and is required to complete the entire subject. The statistician's dilemma arises when he begins to believe that the Real World always conforms to procedures for which he has solutions.

We shall approach all analyses of data from experiments taken in a sequential manner via analysis of variance procedures under the assumption that the linear, additive model with independent effects is valid. This means that an error variance for linear contrasts of effects is available from the analysis of variance. No discussion appears to have been made of sequential procedures involving different error rate bases (e.g. see Tukey [1953], Hartley [1955], Federer [1961], etc.). Although it would appear that the straightforward extension of the fixed sample procedure as presented herein would be permissible, this requires justification.

Also, there is the problem of sequential stopping procedures when the v-1 treatment degrees of freedom have been partitioned into $1 < k \leq v-1$ contrasts. This is a simultaneous test of the k contrasts and a decision on composite hypotheses is required on each of the k contrasts. If the k contrasts have

unequal weights this could be compensated for by altering the error of selecting the wrong hypothesis. Presumably, the sampling would need to be continued until decisions were reached on each of the k contrasts. For unbounded sequential sampling procedures, the sample size would then be the maximum one for the contrast requiring the largest number of samples. For bounded procedures, the maximum sample size would be less than or equal to the largest sample size allowable. For k large the maximum number would often be achieved, and we would have essentially a fixed sample size procedure.

The simultaneous test of k hypotheses is in sharp contrast to the discussion of composite hypothesis testing as discussed by Hall, Wijsman, and Ghosh [1965], Hall [1965], Wetherill [1966], etc.; these authors consider composite hypotheses of the type $\theta = \theta_0$ and $\theta = \theta_1$ and $\theta < \theta_0$, $\theta_0 \leq \theta \leq \theta_1$, and $\theta > \theta_1$. A simultaneous test of k such composite hypothesis would be more to the point in experimentation.

2.1. Fixed Effects

The general regression procedure outlined by Johnson [1953] and Ray [1956] will be utilized in the following examples. Basically these and later writers use a series of sequential F tests with unbounded sampling. Some empirical results on sample size is available; relative to this the comment by Wetherill [1966], page 59, is interesting, "An outstanding feature of sequential t-tests is our state of ignorance concerning their properties.". The same comment would hold for sequential F-tests. Despite this, many experiments are naturally taken sequentially and the experimenter would like to reach a conclusion. Hence, a procedure may necessarily be used regardless of its statistical properties

and will continue to be used until one with more desirable statistical properties is available.

Example 2.1. Completely randomized design with fixed effects.

Random samples of 5 observations each were obtained from a near normal population with mean 30 and variance 100 (Snedecor [1956], table 2.3.1). Letters A,B,C,D, and E representing treatments were assigned to the samples as drawn. The sample drawing is sequential in nature. The data are given in table 2.1. A number of situations were examined with the experiment being terminated at various stages depending upon the treatment comparisons made. These will be discussed individually below.

The linear model is $Y_{ij} = \mu + \tau_i + \epsilon_{ij}$, given that the τ_i are fixed effects and the ϵ_{ij} are random, independent, normal variates with mean zero and variance σ_ϵ^2, $i=1,2,\cdots,v$, $j=1,2,\cdots,r$.

Accept either the H_0 or the H_1 hypothesis. We shall utilize the procedure given by Johnson [1953] and Ray [1956]. The first step in the procedure is to specify a value of $\delta = \sum_{i=1}^{v} \tau_i^2/v\sigma_\epsilon^2$. In most cases, $\Sigma\tau_i$ is set equal to zero. If this is so, then the average treatment effect in standard deviation units is the item of interest. Although the experimenter may not like to think in these terms this is the way the test procedure was constructed. Ray [1956] has constructed tables for $\delta = \frac{1}{2}$, 1, and 2 for the statistic G(vr=N) = Among treatments sum of squares divided by the within treatments sum of squares, i.e.,

$$s_t/s_e = G(N) = \left(\sum_{i=1}^{v} Y_{i.}^2/r - Y_{..}^2/vr \right) \Big/ \sum_{i=1}^{} \left(\sum_{j=1}^{r} Y_{ij}^2 - Y_{i.}^2/r \right) . \qquad (2.1)$$

Let $\lambda(N) = \lambda(vr) = vr\delta$; then the test procedure is:

"Accept H_1 if $e^{-\lambda(N)/2} \, M\left\{ \frac{vr-1}{2} , \frac{v-1}{2} ; \frac{\lambda(N)G(N)/2}{1+G(N)} \right\} \geq \frac{1-\beta}{\alpha}$;

accept H_0 if $e^{-\lambda(N)/2} \, M\left\{ \frac{vr-1}{2} , \frac{v-1}{2} ; \frac{\lambda(N)G(N)/2}{1+G(N)} \right\} \leq \frac{\beta}{1-\alpha}$.

Otherwise take a further set of v (or mv) observations, one (or m) on each treatment."

H_0 is that $\tau_i = 0$, H_1 is that τ_i = a specified value such that $\sum_{i=1}^{v} \tau_i^2 = v\sigma_\epsilon^2\delta$ for a specified δ, α and β are the (approximate) chances of erroneously rejecting H_0 and H_1, respectively, and $M(X,Y,;\mu)$ is the confluent hypergoemetric function which has been tabled by Rushton [1954], Rushton and Lang [1954], and Slater [1960]. Let the upper limit be $\bar{G}_\alpha(N)$ be the solution for $G(N)$ in the following equation:

$$\frac{1-\beta}{\alpha} = e^{-\lambda(N)/2} \, M\left\{ \frac{(vr-1)}{2} , \frac{v-1}{2} ; \frac{\lambda(N)G(N)/2}{1+G(N)} \right\} . \qquad (2.2)$$

Let the lower limit $\underline{G}_\alpha(N)$ be the solution for $G(N)$ obtained from the equation:

$$\frac{\beta}{1-\alpha} = e^{-\lambda(N)/2} \, M\left\{ \frac{(vr-1)}{2} , \frac{v-1}{2} ; \frac{\lambda(N)G(N)/2}{1+G(N)} \right\} \qquad (2.3)$$

The values in tables 2.2 and 2.3 which are reproduced from Ray's [1956] paper, may be obtained from tables of the confluent hypergeometric function, provided extensive tables are available. In these tables $\alpha = \beta = .05$ was used. Other

Table 2.1. Example of a completely randomized design with five treatments (A,B,C, D,E) and with one observation per treatment added in a sequential manner.

Stage	Treatment														
	A			B			C			D			E		
	Y_{Aj}	$Y_{A\cdot}$	ΣY^2_{Aj}	Y_{Bj}	$Y_{B\cdot}$	ΣY^2_{Bj}	Y_{Cj}	$Y_{C\cdot}$	ΣY^2_{Cj}	Y_{Dj}	$Y_{D\cdot}$	ΣY^2_{Dj}	Y_{Ej}	$Y_{E\cdot}$	ΣY^2_{Ej}
1	30	30	900	29	29	841	39	39	1521	17	17	289	12	12	144
2	19	49	1261	42	71	2605	27	66	2250	25	42	914	22	34	628
3	16	65	1517	41	112	4286	37	103	3619	31	73	1875	25	59	1253
4	17	82	1806	30	142	5186	24	127	4195	28	101	2659	35	94	2478
5	47	129	4015	33	175	6275	17	144	4484	33	134	3748	29	123	3319
6	17	146	4304	23	198	6804	31	175	5445	39	173	5269	30	153	4219
7	41	187	5985	26	224	7480	19	194	5806	32	205	6293	27	180	4948
8	20	207	6385	28	252	8264	39	233	7327	43	248	8142	30	210	5848

Source	d.f.	Sum of squares					
		r=3	r=4	r=5	r=6	r=7	r=8
Total	5r	12550	16324	21841	26041	30512	35966
CFM	1	11316	14906	19881	23801	28003	33062
Among treatments	4	747	608	336	280	169	219
Within treatments	5(r-1)	487	810	1624	1960	2340	2685
G(N)	-	1.53	0.75	0.21	0.14	-	-
$F\left(4,4(r-1)\right)$	-	3.83	2.81	1.03 ↑ terminates for $\delta=1$	0.89 ↑ terminates for $\delta=\frac{1}{2}$	-	-

levels of α and β could be obtained in the manner described by Ray [1956] for constructing tables 2.2 and 2.3.

Table 2.2. One-way classification by groups (true limits). v = number of treatments; r = number per treatment; $\lambda = rv\delta$; $\delta = 0.5$; $\alpha = \beta = 0.05$.

v=2				v=3				v=4			
r	λ	$\underline{G}$	$\bar{G}$	r	λ	$\underline{G}$	$\bar{G}$	r	λ	$\underline{G}$	$\bar{G}$
4	4	-	5.390	5	7.5	0.037	1.319	4	8	0.065	1.825
6	6	0.002	1.091	7	10.5	.072	0.696	6	12	.110	0.770
8	8	.025	0.639	9	13.5	.089	.497	8	16	.126	.521
10	10	.040	.471	11	16.5	.099	.400	10	20	.131	.411
12	12	.050	.385	13	19.5	.105	.344	12	24	.134	.350
14	14	0.059	0.333	15	22.5	0.108	0.306	14	28	0.136	0.310
16	16	.066	.300	17	25.5	.110	.280	16	32	.138	.282
18	18	.072	.276	19	28.5	.112	.261				
20	20	.076	.258	21	31.5	.115	.245				
30	30	0.089	0.205								

v=5				v=6				v=7			
r	λ	$\underline{G}$	$\bar{G}$	r	λ	$\underline{G}$	$\bar{G}$	r	λ	$\underline{G}$	$\bar{G}$
3	7.5	0.072	4.176	2	6	0.008	-	3	10.5	0.184	2.407
5	12.5	.142	0.927	4	12	.168	1.272	5	17.5	.211	0.784
7	17.5	.155	.568	6	18	.183	0.646	7	24.5	.207	.512
9	22.5	.158	.432	8	24	.183	.464	9	31.5	.199	.401
11	27.5	.159	.360	10	30	.180	.377				
13	32.5	0.159	0.317	12	36	0.176	0.327				
15	37.5	.157	.287								

Table 2.3. One-way classification by groups. v = number of treatments; r = number per treatment; $\lambda = rv\delta$; $\delta = 1.0$; $\alpha = \beta = 0.05$.

v=2				v=3				v=4			
r	λ	$\underline{G}$	$\bar{G}$	r	λ	$\underline{G}$	$\bar{G}$	r	λ	$\underline{G}$	$\bar{G}$
4	8	0.052	2.390	3	9	0.124	4.760	2	8	0.144	-
5	10	.082	1.380	4	12	.169	1.79	3	12	.231	3.13
6	12	.105	1.016	5	15	.195	1.174	4	16	.266	1.498
7	14	.121	0.826	6	18	.210	0.902	5	20	.276	1.040
8	16	.135	.710	7	21	.221	.762	6	24	.284	0.838
10	20	0.150	0.578	9	27	0.234	0.605	8	32	0.287	0.637
12	24	.166	.504	11	33	.240	.522	10	40	.286	.540
16	32	.187	.424	13	39	.244	.471	12	48	.284	.482
20	40	.199	.383	15	45	.247	.436	16	64	.280	.424
30	60	0.215	0.333	21	63	0.251	0.377	20	80	.277	.391
60	120	.235	.300	31	93	.251	.334	30	120	0.271	0.342
				51	153	.251	.306	50	200	.266	.306

v=5				v=6				v=7			
r	λ	$\underline{G}$	$\bar{G}$	r	λ	$\underline{G}$	$\bar{G}$	r	λ	$\underline{G}$	$\bar{G}$
3	15	0.331	2.469	2	12	0.381	24.042	3	21	0.469	1.925
4	20	.345	1.332	3	18	.405	2.133	4	28	.447	1.175
5	25	.340	0.973	4	24	.398	1.237	5	35	.420	0.888
6	30	.335	.792	5	30	.384	0.920	6	42	.405	.745
7	35	.331	.687	6	36	.373	.763	7	49	.387	.645
9	45	0.322	0.565	8	48	0.354	0.601	9	63	0.365	0.552
11	55	.314	.498	10	60	.340	.521	11	77	.349	.494
13	65	.308	.456	12	72	.330	.473	13	91	.338	.456
15	75	.303	.428	14	84	.322	.440	15	105	.328	.430
25	125	0.288	0.360	20	120	0.307	0.386	21	147	0.315	0.384

Table 2.3. (Continued)

v=8				v=9				v=10			
r	λ	$\underline{G}$	$\bar{G}$	r	λ	$\underline{G}$	$\bar{G}$	r	λ	$\underline{G}$	$\bar{G}$
2	16	0.566	7.500	3	27	0.565	1.684	2	20	0.709	5.066
3	24	.518	1.792	4	36	.510	1.085	4	40	.534	1.06
4	32	.479	1.120	5	45	.470	0.845	6	60	.457	0.705
5	40	.448	0.865					8	80	.415	.575
6	48	.424	.730	7	63	0.421	0.635	10	100	.389	.506
				9	81	.392	.540				
8	64	0.391	0.585	11	99	.372	.485	16	160	0.347	0.414
10	80	.371	.512					20	200	.332	.384
16	128	.336	.414	15	135	0.347	0.425				
20	160	.324	.387	21	189	.325	.380	30	300	0.312	0.348
				31	279	.306	.345	40	400	.300	.329
30	240	0.306	0.348	41	369	.296	.328				
40	400	.288	.318								

Utilizing the statistic in equation (2.1) we may now proceed with the sequential testing approach. We note from tables 2.2 and 2.3 that the minimum sample size for v=5 treatments is r=3. The "among groups" and "within groups" analysis of variance is computed at each state for r=3, r=4, r=5, etc. This is given in the lower part of table 2.1. Using table 2.3 for $\delta=1$, the sampling terminates at r=5 samples, and we accept H_0; using table 2.2 for $\delta=\frac{1}{2}$, the sampling terminates at r=6 samples. These results are indicated in the bottom part of table 2.1.

Single degree of freedom contrasts. Suppose that we partition the four treatment degrees of freedom into four orthogonal contrasts each with a single degree of

freedom. Contrasts 10, 11, 12, and 13 in table 2.4 form such a set, i.e., treatment D vs. E, D+E vs. C, C+D+E vs. B, and B+C+D+E vs. A. Now the sum of squares for any linear contrast of treatment means, say $\sum_{i=1}^{v} c_i \bar{y}_{i\cdot}$, on a per unit basis in the analysis of variance is $r\left(\sum_{i=1}^{v} c_i \bar{y}_{i\cdot}\right)^2 \Big/ \sum_{i=1}^{v} c_i^2$. Corresponding to this sum of squares, which would be the S_t of equation (2.1), we need a denominator sum of squares for two treatments. An average sum of squares for this procedure could be obtained by multiplying S_e = within groups sum of squares for v treatment, by the factor 2/v. This is justifiable if the within treatment variances are all estimates of the same parameter. Then, we would compute the statistic

$$G(2r) = vr\left(\sum_{i=1}^{v} c_i \bar{y}_{i\cdot}\right)^2 \Big/ 2\sum_{i=1}^{v} c_i^2 S_e \quad . \tag{2.4}$$

G(2r) would be compared with tabled values of upper and lower limit values of $G_\alpha(2r)$. The upper and lower limits for the contrast $\sum_{i=1}^{v} c_i \bar{y}_{i\cdot}$, ignoring sign, would be obtained as

$$\sqrt{\frac{2\Sigma c_i^2}{vr} S_e G_\alpha(2r)} \tag{2.5}$$

to yield a linewise error rate of size α. An experimentwise error rate would be obtained by using $\alpha/(v-1)$ where there are v-1 orthogonal single degree of freedom contrasts. One could also have an experimentwise error rate of the Scheffé [1953] type by utilizing $G_\alpha(vr)$ instead of $G_\alpha(2r)$ in equation (2.5) to compare each single degree of freedom contrast.

Table 2.4. Linear contrasts of means.

No.	Contrast	r=3	r=4	r=5	r=6	r=7	r=8
1	A-B	-15.7	-15.0	-9.2	-8.7	-5.3	-5.6
2	A-C	-12.7	-11.2	-3.0	-4.8	-1.0	-3.2
2	A-D	-2.7	-4.8	-1.0	-4.5	-2.6	-5.1
4	A-E	2.0	-3.0	1.2	-1.2	1.0	-0.4
5	B-C	3.0	3.8	6.2	3.8	4.3	2.4
6	B-D	13.0	10.2	8.2	4.2	2.7	0.5
7	B-E	17.7	12.0	10.4	7.5	6.3	5.2
8	C-D	10.0	6.5	2.0	0.3	-1.6	-1.9
9	C-E	14.7	8.2	4.2	3.7	2.0	2.9
10	D-E	4.7	1.8	2.2	3.3	3.6	4.8
	lower (2)	-	2.9	4.6	5.2	5.7	6.0
	upper (2)	-	19.7	18.9	16.3	14.9	13.8
11	2C-D-E	-	14.8	6.2	4.0	0.4	-
	lower (6)	-	5.0	8.0	9.1	9.9	-
	upper (6)	-	34.1	32.8	28.2	25.7	-
12	3B-C-D-E	-	26.0	24.8	15.5	13.3	-
	lower (12)	-	7.1	11.3	12.8	13.9	-
	upper (12)	-	48.2	46.4	39.9	36.4	-
13	B+C+D+E-4A	-	34.0	12.0	19.2	7.9	-
	lower (20)	-	9.2	14.6	16.6	18.0	-
	upper (20)	-	62.2	59.9	51.5	47.0	-
$\frac{q_{.05,5,(r-1)}}{q_{.05,2,5(r-1)}}$ (lower (2))			4.2	6.6	7.5	8.1	-
$\frac{q_{.05,5,5(r-1)}}{q_{.05,2,5(r-1)}}$ (upper (2))			28.5	27.2	23.2	21.1	-

In a similar manner, suppose that the v-1 treatment degrees of freedom are partitioned in sets of degrees of freedom (e.g. as in a factorial) greater than or equal to one. Any given contrast with p-1 degrees of freedom, say, could be compared in the same manner as described for single degree of freedom contrasts except that p/v would replace 2/v and $G_\alpha(pr)$ would replace $G_\alpha(2r)$ in equation (2.5).

To illustrate the procedure with the data in table 2.1, suppose that contrasts 10 to 13 in table 2.4 are the ones of interest, that $\delta = 1.0$, and that $\alpha = \beta = .05$. The upper and lower limits are obtained for each contrast using equation (2.5) and the tabulated values in table 2.3 for $G_\alpha(2r)$. For contrast 10 two means are involved and $\Sigma c_i^2 = 2$. The lower and upper limits for this contrast are denoted by lower (2) and upper (2), respectively. For this contrast, sampling would have stopped at r=4 samples. For contrast 11, $\Sigma c_i^2 = 6$ and the limits are denoted as lower (6) and upper (6); 5 samples would have been sufficient to reach a decision in favor of H_0 on this contrast. For contrast 12, 3B-C-D-E, $\Sigma c_i^2 = 3^2 + (-1)^2 + (-1)^2 + (-1)^2 = 12$; the upper and lower limits are denoted as upper (12) and lower (12), respectively, and are obtained from equation (2.5). Seven samples would have been required to reach a decision. The fourth single degree of freedom contrast, number 13, is the one involving the mean of A versus the others. The $\Sigma c_i^2 = 20$ and the upper and lower limits are denoted as upper (20) and lower (20), respectively, in table 2.4. A decision in favor of H_0 given that $\delta=1$ would have been reached for n=5 samples. Thus, the largest number of samples required for any contrast was 7, and if the process is terminated at this point we are assured that α and β are less than the prescribed value, .05 in this case, for a single degree-of-freedomwise error rate base.

In connection with the above sequential procedure, we note that a decision in favor of H_0 was reached for contrast 13 for r=5 samples and that r=6 samples indicates no decision for this contrast. If the sampling had terminated when r=6, the experimenter might be in a dilemma. However, if one uses the rule that any contrast will not be reconsidered once a decision has been reached, this dilemma will not arise, and the properties for α and β mentioned in the preceding paragraph still hold.

Comparison of all possible differences among pairs of means. In certain cases it is desired to compare differences between all $v(v-1)/2$ pairs of means. If a comparisonwise error rate is desired then the use of the upper and lower values obtained from equation (2.5), e.g. lower (2) and upper (2) in table 2.4, would suffice. For our example, the $5(5-1)/2 = 10$ contrasts are those numbered 1 through 10 in table 2.4. In order to reach a decision in favor of either H_0 or H_1 for $\delta=1$ it would have been necessary to use 8 samples, and H_0 would have been accepted for all pairs.

If, on the other hand, an experimentwise error rate had been desirable, the upper (2) and lower (2) values would be multiplied by the factor $q_{\alpha,v,f}/q_{\alpha,2,f}$, where the $q_{\alpha,v,f}$ values are obtained from the extensively tabulated tables of studentized ranges and where f is the number of degrees of freedom associated with the within groups sum of squares. Proceeding in this fashion we obtain the values in the bottom part of table 2.4. E.g.,

$$\frac{q_{.05,5,15}}{q_{.05,2,15}} = \frac{4.37}{3.01} = 1.45 \quad \text{and} \quad 2.90(1.45) = 4.2 \ .$$

Similarly, we could use $\sqrt{(v-1)F_\alpha(v-1,f)/F_\alpha(1,f)}$ as a multiplier to obtain an experimentwise error rate of the type discussed by Scheffé [1953].

Example 2.2. Randomized complete block design with r blocks, v treatments, rv observations, and fixed treatment effects.

The yield equation for a randomized complete block design of the following form is considered here:

$$Y_{ij} = \mu + \tau_i + \rho_j + \epsilon_{ij}$$

where the τ_i, ϵ_{ij}, i and j are as defined in example 2.1 and the ρ_j may be either fixed or random effects as this does not affect the test procedure. The test statistic comparable to the one in equation (2.1) will be S_t/S_e equal to

$$G(rv) = \left(\sum_{i=1}^{v} Y_{i.}^2/r - Y_{..}^2/rv \right) \Big/ \sum_{i=1}^{v} \sum_{j=1}^{r} (Y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}+\bar{y})^2 \qquad (2.6)$$

where $\bar{y}_{i.}$, $\bar{y}_{.j}$, and $\bar{y}$ are the treatment, block, and over-all means, respectively, $Y_{i.}$ = treatment totals, and $Y_{..}$ = grand total. This ratio is the treatment to error sums of squares from the analysis of variance. In the same manner as in example 2.1, the sequential procedure is defined by:

i) accept H_1 if $G(rv) > \bar{G}_\alpha(rv)$;

ii) accept H_0 if $G(rv) < \underline{G}_\alpha(rv)$;

iii) otherwise add another (or m) block(s) to the experiment.

$\underset{\sim}{G}_{\alpha}(rv=N)$ is the solution for $G(N)$ from the equation:

$$\frac{1-\beta}{\alpha} = e^{-\lambda(N)/2} M\left(\frac{r(v-1)}{2} , \frac{v-1}{2} ; \frac{\lambda(N)G(N)/2}{1+G(N)} \right) ; \qquad (2.7)$$

$\underset{\sim}{G}_{\alpha}(rv=N)$ is the solution for $G(N)$ from the equation:

$$\frac{\beta}{1-\alpha} = e^{-\lambda(N)/2} M\left(\frac{r(v-1)}{2} , \frac{v-1}{2} ; \frac{\lambda(N)G(N)/2}{1+G(N)} \right) . \qquad (2.8)$$

These are of the same form as equations (2.2) and (2.3) for the completely randomized design; the parameters in the confluent hypergeometric function change, necessitating different tables. Tables for the randomized complete block design have been constructed by Ray [1956] for $\delta = 0.5$, 1.0, and 2.0 for limited values of r and v. Here, as for the one-way classification, additional tables could be constructed from tables of the confluent hypergeometric function.

Provided tables are available, no additional difficulties over a completely randomized design are encountered in the analysis of experiments deisgned as a randomized complete block design. The various procedures described in example 2.1 may be applied directly here.

Example 2.3. n sets of v × v latin squares with v treatments, r = number replicates, nv^2 observations, and fixed treatment effects.

The yield equation considered for an experiment designed as n sets of v × v latin squares is:

$$Y_{hij} = \mu + \rho_h + \tau_i + \gamma_j + \epsilon_{hij}$$

where $i=1,2,\cdots,v$, $h=1,2,\cdots,nv$, $j=1,2,\cdots,v$, τ_i are fixed effects, ϵ_{hij} are randcm independent, normal variates with mean zero and variance σ^2_ϵ, and the ρ_h and γ_j may be either fixed or randcm effects. At each stage $(1,2,\cdots,n)$ an experiment designed in a v x v latin square design is conducted with the v additional rows being added in the v columns. Thus, v observations on each treatment are added at each stage. Treatment effects are orthogonal to rows and to columns in this design. The same procedures as used in example 2.1 may be used here provided suitable tables are available. (They need to be computed.) The statistic required is:

$$G(nv^2=N) = \frac{\left(\sum_{i=1}^{v} Y^2_{\cdot i \cdot}/nv - Y^2_{\cdots}/nv^2 \right)}{\sum_{h=1}^{nv} \sum_{j=1}^{v} (Y_{hij} - \bar{y}_{h\cdot\cdot} - \bar{y}_{\cdot i \cdot} - \bar{y}_{\cdot\cdot j} + 2\bar{y})^2} \qquad (2.9)$$

which is the treatment sum of squares with v-1 degrees of freedcm divided by the residual sum of squares with (nv-2)(v-1) degrees of freedcm. The sequential procedure is as defined for the previous two examples. The upper limit, $\bar{G}_\alpha(rv=nv^2=N)$, is obtained as a solution to the following equation:

$$\frac{1-\beta}{\alpha} = e^{-\lambda(N)/2} M\left(\frac{(b-1)(k-1)}{2}, \frac{v-1}{2}; \frac{\lambda(N)G(N)/2}{1+G(N)} \right) \qquad (2.10)$$

The lcw limit $\underline{G}_\alpha(rv)$ is obtained as a solution to the equation:

$$\frac{\beta}{1-\alpha} = e^{-\lambda(N)/2} M\left(\frac{(b-1)(k-1)}{2}, \frac{v-1}{2}; \frac{\lambda(N)G(N)/2}{1+G(N)} \right) \qquad (2.11)$$

Example 2.4. v treatments designed in b blocks of size k each with r replicates of each treatment with fixed treatment effects.

If the v treatments are arranged in b blocks of size k (k usually but not necessarily less than v) with each treatment appearing r times in the experiment and if the treatments are properly randomized in their allocation to blocks and within blocks, we have a class of experimental designs which includes the balanced and partially balanced incomplete block designs as well as many others. The yield equation considered here is of the form

$$Y_{ij} = n_{ij}(\mu+\tau_i+\beta_j+\epsilon_{ij}) \quad ,$$

where the τ_i and β_j are fixed effects, $n_{ij} = 1$ if the i^{th} treatment appears in the j^{th} block and equals zero otherwise. (This could easily be made more general if desired.) ϵ_{ij} are random, normal, independent variates with zero mean and a common variance σ_ϵ^2, $i=1,2,\cdots,v$, and $j=1,2,\cdots,b$. Intrablock analysis is considered to be appropriate. If $\hat{\tau}_i$ are the solutions for the τ_i obtained from the normal equations given that, e.g. $\Sigma \hat{\tau}_i=0$, and if $Q_{i\cdot} = Y_{i\cdot} - \sum_{j=1}^{b} n_{ij}\bar{y}_{\cdot j}$ where $Y_{i\cdot}$ = treatment total and $\bar{y}_{\cdot j}$ = block mean, then the sum of squares due to treatments (eliminating block effects) is $S_t = \sum_{i=1}^{v} \hat{\tau}_i Q_{i\cdot}$. Likewise, the residual sum of squares is $S_e = \sum_{i=1}^{v}\sum_{j=1}^{b} n_{ij}(Y_{ij}-\hat{\tau}_i-\widehat{(\mu+\beta_j)})^2$, then we may proceed as in the previous sections except to use $M\left(\frac{b(k-1)}{2}, \frac{v-1}{2}; \frac{\lambda(N)G(N)/2}{1+G(N)}\right)$, for $N = rv = bk$, as the confluent hypergeometric function instead of the one in equations (2.7) and (2.8), e.g.

If it is desired to have single degree of freedom contrasts among the τ_i, say $\sum_{i=1}^{v} \ell_i \tau$ for $\sum_{i=1}^{v} \ell_i = 0$, then the estimate of this contrast is $\Sigma\, \ell_i \hat{\tau}_i = \ell' \hat{\tau}$, say, and its variance is $\ell' V(\hat{\tau}) \ell$ where $V(\hat{\tau})$ is the variance-covariance matrix of the $\hat{\tau}_i$. One could then proceed as described for example 2.1.

It should be noted that r need not be a constant and that n_{ij} could be the number of times, 0,1,2,$\cdots$, that the i^{th} treatment occurs in the j^{th} block. No new ideas are encountered but the arithmetic is more difficult. The recovery of interblock information introduces special difficulties, and is not considered here.

Example 2.5. k-row by b-column design for v treatments with r replicates on each treatment.

Whenever k and b are multiples of v we essentially have the situation discussed under example 2.3. Therefore, the interim steps obtained by adding rows (and/or columns) one at a time given that b is a multiple of v should be amenable to the same procedure as for the orthogonal case. If the test in section 2.3 terminates with probability one then this one would have to terminate also as the former is embedded within the framework of the present one. Thus solutions, $\hat{\tau}_i$ say, of the τ_i in the normal equations would be obtained and a sum of squares for treatments eliminating row and column effects would be computed as $S_t = \sum_{i=1}^{v} \hat{\tau}_i Q_{i\cdot\cdot}$ where $Q_{i\cdot\cdot} = Y_{i\cdot\cdot} - \sum_{j=1}^{k} n_{ij\cdot} \bar{y}_{\cdot j\cdot} - \sum_{j=1}^{b} n_{i\cdot h} \bar{y}_{\cdot\cdot h} + r\bar{y}$ for $\bar{y}_{\cdot j\cdot}$ = j^{th} row mean, $\bar{y}_{\cdot\cdot h}$ = h^{th} column mean, $\bar{y}$ = over-all mean, $n_{ij\cdot}$ = number of times treatment i occurs in row j, $n_{i\cdot h}$ = number of times treatment i occurs in column h, $Y_{i\cdot\cdot}$ = total for treatment i, and the yield observation is $Y_{ijh} = n_{ijh}(\mu+\tau_i+\rho_j+\gamma_h+\epsilon_{ijh})$. In the yield equation μ, τ_i, ρ_j, and γ_h are

fixed effects for over-all effect, treatment effect, row effect, and column effect, respectively. The ϵ_{ijh} are random, independent, normal variates with mean zero and variance σ_ϵ^2. n_{ijh} = one if the i^{th} treatment occurs in the j^{th} row and h^{th} column and equals zero otherwise.

The residual sum of squares is equal to $\sum_{j=1}^{k}\sum_{h=1}^{b}(Y_{ijh}-\hat{\mu}-\hat{\tau}_i-\hat{\rho}_j-\hat{\gamma}_h)^2 = S_e$. The resulting test statistic would be of the form $G(rv=bk=N) = S_t/S_e$. New tables would need to be computed using the confluent hypergeometric function $M\left(\frac{b\,(k-1)}{2}\,,\ \frac{v-1}{2}\,;\ \frac{\lambda(N)G(N)/2}{1+G(N)}\right)$. Then, using these tables the test procedures of the previous sections would apply.

2.2. Random Effects

Several procedures (e.g. Johnson [1953,1954], Wetherill [1965], Hall, Wijsman, and Ghosh [1965]) are available concerning test of hypothesis on the ratio of two variance components in the analysis of variance. We shall confine our discussion to one of the procedures described by Johnson [1953], and to the situation wherein H_0 is the null hypothesis that the treatment variance component is zero. For this procedure and for designs having treatment effects orthogonal to the blocking effects, the tables may easily be constructed following the method described by Johnson [1953].

Example 2.6. The completely randomized design with v treatments and r replicates on each treatment with random treatment effects.

Suppose that the yield equation for a completely randomized design with v treatments and r replicates is of the form $Y_{ij} = \mu+\tau_i+\epsilon_{ij}$ where μ is a constant common to all observations, τ_i and ϵ_{ij} are random, independent, normal

random variates with means zero and variances σ_τ^2 and σ_ϵ^2, respectively, i=1,2,$\cdots$,v, and j=1,2,$\cdots$,r. Let $\theta = \sigma_\tau^2/\sigma_\epsilon^2$ and let H_0 be $\sigma_\tau^2/\sigma_\epsilon^2 = \theta_0 = 0$ and H_1 be that $\sigma_\tau^2/\sigma_\epsilon^2 = \theta_1$ a specified value. In this formulation $\theta = \sigma_\tau^2/\sigma_\epsilon^2$ in the random effects case plays the role that $\sum_{i=1}^{v} \tau_i^2/v\sigma_\epsilon^2$ played in the fixed effects case. The upper, $\bar{G}_R(N)$, and lower, $\underline{G}_R(N)$, limits as given by Johnson [1953] are:

$$\underline{G}_R(N=vr) = r\theta_1\underline{\Pi}(1-\underline{\Pi})^{-1} - 1 \tag{2.12}$$

and

$$\bar{G}_R(N=vr) = r\theta_1\bar{\Pi}(1-\bar{\Pi})^{-1} - 1 \tag{2.13}$$

where

$$\underline{\Pi} = \left(\frac{\beta}{1-\alpha}\right)^{2/(rv-1)} \left(\frac{1}{1+r\theta_1}\right)^{v(r-1)/(rv-1)} \tag{2.14}$$

and

$$\bar{\Pi} = \left(\frac{1-\beta}{\alpha}\right)^{2/(rv-1)} \left(\frac{1}{1+r\theta_1}\right)^{v(r-1)/(rv-1)} \tag{2.15}$$

Johnson [1953] presents tables for various values of v and r for $\theta_1 = 1$ and $\alpha = \beta = .05$ and .01. These tables are easily computed, even on a desk calculator. Therefore, extensive tables for various values of θ_1, α, and β are readily available should they be desired.

For the test procedure compute

$$G_R(rv=N) = \left(\sum_{i=1}^{v} Y_{i.}^2/r - Y_{..}^2/rv\right) \Big/ \sum_{i=1}^{v}\sum_{j=1}^{r} (Y_{ij}-\bar{y}_{i.})^2$$

and proceed as follows:

i) Accept H_0 if $G_R(N) < \underline{G}_R(N)$

ii) Accept H_1 if $G_R(N) > \bar{G}_R(N)$

iii) Otherwise add another (or m) observation(s) to each treatment group.

Example 2.7. The randomized complete block design with v treatments, r complete blocks, rv observations and random treatment effects.

Suppose that the yield equation for a randomized complete block design for v treatments in r complete blocks is of the form $Y_{ij} = \mu+\tau_i+\rho_j+\epsilon_{ij}$ where $\mu+\rho_j$ is the mean of the j^{th} block, τ_i are random, independent, normal variates with zero mean and common variance σ^2_τ, ϵ_{ij} are random, independent normal variates with mean zero and common variance σ^2_ϵ, i=1,2,$\cdots$,v, and j=1,2,$\cdots$,r.

The test procedure is the same as for example 2.6 except that

$$G_R(N) = \Big(\sum_{i=1}^{v} Y_{i.}^2/r - Y_{..}^2/rv \Big) \Big/ \sum_{i=1}^{v} \sum_{j=1}^{r} (Y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}+\bar{y})^2 \quad , \qquad (2.16)$$

$$\underline{\Pi} = \Big(\frac{\beta}{1-\alpha}\Big)^{2/r(v-1)} \Big(\frac{1}{1+r\theta_1}\Big)^{(r-1)(v-1)/r(v-1)} \quad ; \qquad (2.17)$$

$$\bar{\Pi} = \Big(\frac{1-\beta}{\alpha}\Big)^{2/r(v-1)} \Big(\frac{1}{1+r\theta_1}\Big)^{(r-1)(v-1)/r(v-1)} \quad , \qquad (2.18)$$

and one (or m) complete block(s) would be added at each stage. Tables would need to be computed but this would be an easy task.

Example 2.8. n sets of v x v latin squares with v treatments, r = nv replicates, nv^2 observations, and random treatment effects.

Suppose that the yield equation is of the same form as in example 2.3 except that the treatment effects are random, independent, normal variates with mean zero and common variance σ_τ^2. The test procedure would be the same as for the previous two examples except that S_t/S_e would equal

$$G_R(N) = \left(\sum_{i=1}^{v} Y_{i..}^2/r - Y_{...}^2/rv \right) \sum_{j=1}^{nv} \sum_{h=1}^{v} (Y_{ijh} - \bar{y}_{i..} - \bar{y}_{.j.} - \bar{y}_{..h} + 2\bar{y})^2, \quad (2.19)$$

$$\underline{\Pi} = \left(\frac{\beta}{1-\alpha}\right)^{2/(nv-1)(v-1)} \left(\frac{1}{1+r\theta_1}\right)^{(nv-2)(v-1)/(nv-1)(v-1)}, \quad (2.20)$$

and

$$\bar{\Pi} = \left(\frac{1-\beta}{\alpha}\right)^{2/(nv-1)(v-1)} \left(\frac{1}{1+r\theta_1}\right)^{(nv-2)(v-1)/(nv-1)(v-1)}, \quad (2.21)$$

Example 2.9. Other situations.

A test statistic and procedures for other experimental designs may be computed in the same manner prescribed for examples 2.4 and 2.5. Tables could be computed as described for the preceding three examples. However, before doing this the various properties of these procedures should be examined. Certainly, all designs with treatment effects orthogonal to the blocking effects could be treated in this manner.

2.3 Mixed Effects for Treatments in a Factorial Arrangement

Suppose that the treatments in examples 2.1, 2.2, and 2.3 were in a factorial arrangement and that the levels of one of the factors are random effects. For some hypotheses, the interaction mean square would be used to test hypotheses about main effects. An even more complicated situation arises when mean squares need to be added or subtracted to obtain a synthetic F test. Although one could construct a test procedure which is an analogue of the fixed sample case, there is no assurance that such a procedure would have desirable properties. If it is necessary to utilize a sequential test procedure and if none is available, then the experimenter's only recourse is to utilize the fixed sample procedure and use the easily constructed tables described by Johnson [1953].

3. SEQUENTIAL ESTIMATION

Sequential test procedures as described in section 2 are, as their name implies, specifically constructed for use in differentiating among a few (2 or 3) specific hypotheses. These procedures do not necessarily produce precise and low cost estimates of parameters. In fact (see Johnson [1961], section 5, and Wetherill [1965], chapter 8), the sequential estimation problem appears to be this:

i) Provided the sample size when the experiment is stopped is "not too small", fixed sample methods of estimation can be used, even though the sample was selected sequentially, for many estimation situations.

ii) Sequential estimation is no more (and perhaps less) efficient than fixed sample size estimation.

iii) Much theoretical work is required in this area.

As stated before many experiments are sequential by nature. Then, it would appear from the above that fixed sample size procedures may be utilized with relative efficiency in sequential estimation of confidence intervals if the sample size is "not too small" (whatever that means). Stopping rules need to be devised, and here (see Wetherill [1965], section 8.7), decision theory provides the framework for determining such rules.

Perhaps one of the simplest of sequential estimation procedures is the double sampling plan or two stage plan given by Stein [1945]. For this case, the random variables Y_i are independently and normally distributed with mean μ and variance σ_ϵ^2 and the problem is to estimate μ by a $(1-\alpha)\%$ confidence interval of length less than L, say. In the first stage a sample of n observations

is taken. Then an estimate of σ_ϵ^2 is computed as $s_n^2 = \sum_{i=1}^{n} (Y_i - \bar{y}_n)^2/(n-1)$ where $\bar{y}_n = \sum_{i=1}^{n} Y_i/n$. A second sample of size m is taken and the sample mean from all n + m = N observations is computed as $\sum_{i=1}^{N} Y_i/N = \bar{y}_N$. Then, a confidence interval $\bar{y}_N \pm t_{\alpha,n-1}\sqrt{s_1^2/N}$ where $t_{\alpha,n-1}$ is the two-sided α% point of Student's t distribution with n-1 degrees of freedom. The length of the confidence interval $L = 2t_{\alpha,n-1}\, s_1/\sqrt{N}$, and if we choose $N = [4t_{\alpha,n-1}^2\, s_1^2/L^2] + 1$ where the quantity in brackets, [x], is the greatest integer less than x, then the confidence interval has length less than L. Thus, if $N < n$ the confidence interval based on the first sample alone is already less than L. If $N > n$ then m more observations are required.

In this method no use is made of the last m observations in computing an estimate of σ_ϵ^2; if all N results are used, the confidence interval will not necessarily be less than L, although the proportion of times it would be less than L could be computed. In many situations there will not be such a rigid requirement on the length of the confidence interval, and hence all N observations would be utilized in computing the estimate of σ_ϵ^2. As Wetherill [1965] states the solution given by Stein [1965] is typical of many in sequential analysis. With ingenuity a solution to a mathematically precisely stated problem is obtained, "but neither the problem nor, still less, the solution corresponds to what the practicing statistician really wants to do".

In this connection in biological experimentation the error variance often changes from condition, or environment, to condition. (Perhaps this could be countered with an appropriate transformation.) Also, the cost from stage to

stage is much larger per additional sample than within stages. Therefore, it would appear more economical to select n large enough in the first stage to obtain a confidence interval of the desired length. This would essentially amount to a fixed sample one stage type of procedure.

If an observation arises one at a time as in rotation experiments then a transformation should be used to stabilize the error variances; a sequential estimation procedure with a stopping rule possessing desirable characteristics, should then be utilized. Such experiments as rotation experiments allow considerable time between observations, usually one year, to analyze the results and determine whether to stop the experiment or to proceed. It is felt that many such experiments are carried on because "it is a good idea", "we may learn something", etc., and that many of them should have been stopped many years ago! Summarization of data at the end of each stage should contribute greatly to the efficiency of experiments of this type. Thus, if the experimenter and statistician think in terms of sequential estimation and keep analyses current with data collected, experiments will be terminated on a rational basis and will not be continued indefinitely until "funds are no longer available" or "something appears which is more exciting to work on". Many medical trials are sequential in nature, but analyses are often not kept current with results.

4. DISCUSSION

In their paper, Hall, Wijsman, and Ghosh [1965] list the statistical properties used for assessing characteristics of a sequential test procedure as follows:

i) strength which refers to degree of correctness of the error probabilities α and β,

ii) termination which refers to the fact that the test does or does not terminate with probability one,

iii) ASN-function which is the average sample size for termination, and

iv) OC-function which is the operating characteristic function of the test.

A sequential test procedure has the correct strength if the error of rejecting H_0 when true is α and of rejecting H_1 when true is β. In practice, the procedure would be adequate even if α and β were within 2-3 percentage points of the stated values. In other words, the experimenter would desire that α and β be approximately correct. Several procedures are constructed to have the true errors less than or equal to α and β.

With regard to the property that a test will or will not terminate with probability one, in practice a sequential experiment will always terminate long before infinity. Therefore, a practicing statistician couldn't care less if the probability of termination before infinity is unity. He will terminate the experiment at some sample size, say N_0, which may not be exceptionally large, say $N_0 = 20$, 30, or 40. Under these circumstances he is interested in the proportion of the time that he will be able to reach a decision about the specified

hypotheses. A closed or bounded form of the test is required. Few closed form procedures appear to be available.

The average sample size property, ASN, is an important consideration in sequential test procedures but appears of little consequence in sequential estimation. But, even more important in test procedures, it would appear, is the operating characteristic or the distribution or stopping time sample size (DSN). Except for a few cases, knowledge of the distribution of DSN appears to be limited to a few special cases although a "large-sample" approximation has been found (Johnson [1961]). From the distribution, when available, the proportion of time a decision would have been reached before some fixed upper limit on sample size = N_0, can be computed.

It is suggested that the properties of tests be re-examined in light of application rather than in terms that are justifiable because of their mathematical simplicity or tractibility. Also, if the experiment is sequential in nature, and many are, and if hypothesis testing is desired, it will be necessary to utilize a test procedure. Any knowledge of the properties of such a procedure would be desirable. Therefore, much work needs to be done to obtain procedures meeting the requirements of the experimenter. If the mathematics is too difficult to investigate the procedure analytically, it may be necessary to utilize a high-speed computer to empirically investigate some of the properties of the test procedure.

In all the procedures discussed in this section nothing has been said about analysis carried out "to spot the winner" in which case it would appear desirable to eliminate non-contenders early in the testing procedure (see Wetherill [1965], p. 72). It appears that many procedures will need to be devised to meet all situations.

One problem that requires investigation is the use of Johnson's [1953] procedure to construct tables in the random effects case for use in the fixed effects case. In fact, no table is required since the computation is so simple. If this procedure can be used as a reasonably good approximation, the table construction problem would be solved.

Another problem that would definitely fall in the area of sequential design of experiments is the selection of the experimental design at each stage. For example, a latin square design could be utilized at the first stage, then based on the experimental results, a randomized complete block or a completely randomized design might be utilized in the remaining stages of the sequential experiment. Also, a sequential selection of the analysis might be utilized. For example, suppose that the design is a latin square and the standard row-column-treatment-residual analysis of variance is performed. Based on the results of the first stage, a differential gradient within columns-column-treatment-residual analysis of variance might be utilized in the next or all of the following stages. The sequential selection of blocking, of analyses, and of the function of the observations are pertinent unsolved problems facing the statistician and, of course, the experimenter who has these problems whether or not the statistician has solutions for them.

The suggestion by Sir Ronald A. Fisher that sequential design of experiments would be a fruitful field in which to work is just as true today as it was 15 years ago when the statement was made. It would be desirable to have analytic results or reasonable approximations thereof. If this is not forthcoming then high speed computers may be utilized to obtain some empirical evidence on the properties of a procedure. This may lead to ideas for analytic results.

5. SUMMARY

An exposition of analyses of variance procedures used in sequential experimentation with sequential hypothesis testing and with sequential estimation is presented. Several sequential test procedures, viz, a single F test in the analysis of variance, a set of F tests involving orthogonal single degree of freedom contrasts, and a set of t tests (or F tests) for all possible differences between pairs or contrasts or means, are presented for data from an experiment designed as a completely randomized design. Also, several error rate bases are considered. A numerical example is utilized to illustrate the procedures. Following a discussion of these procedures illustrating their extension to other experimental designs, analyses using a single F test are described for the randomized complete block design with blocks added sequentially, single v × v latin square added sequentially, v treatments designed in b blocks of size k (k usually but not necessarily less than v) with each treatment occurring r times and with sets of blocks being added sequentially, and k-row by b-column design for v treatments with r replicates on each treatment and with rows and/or columns added sequentially. The analyses presented were for fixed effects which was followed by a discussion of random and mixed effects cases for each of the designs presented.

Under sequential estimation a two stage sampling procedure is discussed along with results relating fixed sample size estimation procedures to sequential sampling estimation procedures. Some discussion is presented relative to unsolved problems in the sequential selection of an experimental design, the sequential selection of an analysis, and of the scale of measurement, or

transformation. All of the discussion relates to sequential design and analysis of experiments with no presentation being made of sequential procedures for selecting treatments as, e.g. levels of the independent variate in regression, of levels for a dosage response, and related phenomena.

6. LITERATURE CITED

Anscombe, F. J. [1953]. Sequential estimation. Journal Royal Statistical Society, Series B, 15:1-29.

Federer, W. T. [1961]. Experimental error rates. Proceedings of the American Society for Horticultural Science. 78:605-615.

Hall, W. J. [1965]. Methods of sequentially testing composite hypotheses with special reference to the two-sample problem. Institute of Statistics Mimeo Series No. 441, University of North Carolina, Chapel Hill, North Carolina.

Hall, W. J., Wigsman, R. A., and Ghosh, J. K. [1965]. The relation between sufficiency and invariance with applications in sequential analysis. Annals of Mathematical Statistics, 36:575-614.

Hartley, H. O. [1955]. Some recent developments in analysis of variance. Communications of Pure and Applied Mathematics, 8:47-

Jackson, J. E. [1960]. Bibliography on Sequential analysis. Journal of the American Statistical Association, 55:561-580.

Johnson, N. L. [1953]. Some notes on the application of sequential methods in the analysis of variance. Annals of Mathematical Statistics, 24:614-623.

Johnson, N. L. [1954]. Sequential procedures in certain component of variance problems. Annals of Mathematical Statistics, 25:357-366.

Johnson, N. L. [1961]. Sequential analysis: A survey. Journal Royal Statistical Society, Series A, 124:372-411.

Ray, W. D. [1956]. Sequential analysis applied to certain experimental designs in the analysis of variance. Biometrika 43:388-403.

Rushton, S. [1954]. On the confluent hypergeometric function $M(\alpha,\gamma,x)$. Sankhyā 13:369-376.

Rushton, S. and Lang, E. D. [1954]. Tables of the confluent hypergeometric function. Sankhya 13:377-411.

Scheffé, H. [1953]. A method for judging all contrasts in the analysis of variance. Biometrika 40:87-104.

Slater, L. J. [1960]. Confluent hypergeometric functions. Princeton University Press, Princeton, New Jersey.

Literature Cited, Continued.

Snedecor, G. W. [1956]. Statistical methods, 5th ed. The Iowa State University Press, Ames, Iowa, U.S.A.

Stein, C. [1945]. A two-sample test for a linear hypothesis whose power is independent of the variance. Annals of Mathematical Statistics 16:243-258.

Tukey, J. W. [1953]. The problem of multiple comparisons. Dittoed Notes, Princeton University, 396 pp.

Wetherill, G. B. [1965]. Sequential methods in statistics. Methuen and Co., Ltd., London. John Wiley and Sons, Inc., New York.