DOCUMENT RESUME

ED 064 343 TM 001 508

AUTHOR	Draper, John F., Jr.
TITLE	A Monte Carlo Investigation of the Analysis of Variance Applied to Non-Independent Bernoulli Variates.
PUB DATE	Apr 72
NOTE	40p.; Paper presented at the annual meeting of the American Educational Research Association (Chicago, Illinois, April 1972)
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Analysis of Variance; *Factor Analysis; *Mathematical Models; Research Methodology; *Statistical Studies; *Tables (Data)

ABSTRACT

The applicability of the Analysis of Variance, ANOVA, procedures to the analysis of dichotomous repeated measure data is described. The design models for which data were simulated in this investigation were chosen to represent simple cases of two experimental situations: situation one, in which subjects' responses to a single randomly selected set of three stimuli or items were evaluated dichotomously on four successive occasions, and situation two, in which subjects' responses are evaluated dichotomously on four successive occasions with respect to four separate randomly selected set of three stimuli or items. The two experimental situations are represented by "three way factorial" designs with random factors for subjects and items and a fixed factor for occasions; the first with subjects, items and occasions completely crossed and the second with items and occasions crossed with subjects, but with items nested within occasions. Results of the study are presented in 11 tables and 7 figures. (Author/DB)

A MONTE CARLO INVESTIGATION OF THE ANALYSIS OF VARIANCE APPLIED TO NON-INDEPENDENT BERNOULLI VARIATES

By

John F. Draper, Jr.
CTB/McGraw-Hill

Paper presented at the annual
meeting of the American Educational Research Association
April, 1972

The problem with which this paper will be concerned is the applicability of the Analysis of Variance, ANOVA, procedures to the analysis of dichotomous repeated measure data. Recall, that in the employment of ANOVA procedures it is assumed that the data can be reasonably modeled by a vector variate with multi-variate normal density and diagonal or patterned "compound symmetric" covariance matrix, that is with equal variances and equal covariances in diagonal blocks for subjects and zeros elsewhere, (Greenhouse and Geisser, 1959). Can these familiar procedures still be employed when the data can more reasonably be modeled by a vector variate with non-independent Bernoulli variates as elements? Other investigators, Mandeville (1970) and Seegar and Gabrielson (1968) have employed Monte Carlo methods in order to achieve a partial answer to the above question. Their results tend toward a cautious but affirmative answer with respect to the design models and model parameter configurations with which their simulations were concerned. The results reported in this paper are less sanguine. That is, however, not to say that the results reported in this paper are in disagreement with those of previous investigators, since the design models and model parameter configurations with which this paper is concerned are somewhat different. Bradley (1968) has indicated that the robustness of a parametric statistical test is idiosyncratic rather than general. For example, the F-test has been shown to be robust to heteroscedasiticity for balanced designs (equal cell n's), but generally not robust to heteroscedasiticity for unbalanced designs.

The design models for which data were simulated in this investigation were chosen to represent simple cases of two experimental situations, situation one, in which subjects responses to a single randomly selected set of three stimuli or items were evaluated dichotomously on four successive occasions and situation two, in which subjects' responses are evaluated dichotomously on four successive occasions with respect to four separate randomly selected sets of three stimuli or items. Table 1 contrasts these two experimental situations.

Table 1. Contrast of experimental Situation 1 with Experimental Situation 2.

Situation 1	Occasion	O_1	O_2	O_3	O_4
	Item	$I_1I_2I_3$	$I_1I_2I_3$	$I_1I_2I_3$	$I_1I_2I_3$
Situation 2	Occasion	O_1	O_2	O_3	O_4
	Item	$I_1I_2I_3$	$I_4I_5I_6$	$I_7I_8I_9$	$I_{10}I_{11}I_{12}$

In terms of experimental design the two experimental situations can be represented by "three way factorial" designs with random factors for subjects and items[1] and a fixed factor for occasions; the first with subjects, items and occasions completely crossed and the second with items and occasions crossed with subjects, but with items nested within occasions.

[1]Items were considered to represent a random source of variation, since experimenters often wish to generalize their results to some domain of items or stimuli, rather than to restrict their findings to only those items they actually employed.

These two experimental situations were considered to be of interest because experimenters often form occasion measures for subjects as the sum of dichotomous response evaluations to a series of items and then ignore items as a factor in their experimental design. And because it may be shown, (Draper, 1971), that for such experimental situations, there is a source of variation associated with items[2], which if non-null will be confounded with the source of variation for occasions. A situation which holds whether or not, items are included as a factor in the experimental design and an ANOVA consistent with the design is employed for analysis and which also holds whether the dependent measures, or item scores, are dichotomous or have distributions which may be reasonably approximated by a normal probability density.

For situation one an examination of expected mean squares, suggested possible variance ratio tests for sources of variation due to subjects, occasions, items, subjects by occasions interaction, and items by occasions interaction. Further, since the items by occasions interaction could be non-null and would thus be confounded with occasions, a Quasi-F variance ratio test for occasions (Satterthwaite, 1941) suggested itself. Similarly for situation two possible variance ratio tests were suggested for the sources of variation, subjects, occasions, items, and subjects by occasions interaction, where both a "regular" variance ratio test and a "Quasi-F" test were suggested to test for occasions effects. Calculations with respect to all the suggested tests were made on simulated data, but of major interest were those tests for occasions and the subjects by occasions interaction.

---

[2]An item by occasion interaction effect in situation one or a main effect due to items in situation two.

It was considered of interest to vary the following model parameters over examples of the experimental situations to be investigated.

1) The base probability of one[3], on item difficulty, in the data,

2) The number of subjects,

3) The degree of relative heterogenity of the subjects[4],

4) The effect of items[4],

5) The effect of occasions[4],

and

6) The effect of subject by occasion interaction[4].

Three different values of the base probability of a one, .5, .2, and .1, suggested by results due to Lunney (1969), were investigated as were five values of the number of subjects 4, 6, 8, 10, and 12. The four levels of relative subject heterogenity, Levels 1, 2, 3, and 4, which were employed, correspond to four ratios of subject to error variance, 2:8, 3:7, 4:6 and 5:5 respectively. Combinations of these parameter values in combination with the values for effects due items, occasions and subject by occasion interaction resulted in 720 different model parameter configurations for which data were simulated. The data for the model parameter configurations were simulated as pseudo-random numbers with distributions which could reasonably be approximated by a normal probability density function and all ANOVA statistics calculated, then the data were dichotomized and the ANOVA calculations repeated. Determinations were then made as to which statistics fell to various α rejection regions and counters which had been initialized at zero were incremented by one where appropriate.

---

[3]The base probability of a one will refer to the general probability of a one in the absence of possible main and interaction effects.

[4]Whether null or non-null.

This process was repeated 999 times for each parameter configuration for a total of 1,000 simulations.

As an example of the type of analysis which was done on the frequency data obtained from the simulation runs consider the following. Table 2 contains frequencies in α = .05, .025 and .01 rejection regions of variance ratio tests of occasion effects under conditions in which the occasion effects were null, items were crossed with occasions and there was no interaction between subjects and occasions. Table 2 is laid out as a six-dimensional array with three left margins and three upper margins. The left-most margin indicates the number of subjects employed with respect to a given simulation of data. Proceeding from left to right the next margin indicates the nature of the items, whether null in effect or non-null, and the right-most indicates the level of subject heterogeneity. The upper margins from top to bottom indicate: (1) the probability of a one with respect to a given simulation of data, (2) one thousand times nominal α or the expected frequency given a true F-variate, and (3) an indication of whether the 1,000 variance ratios with respect to a given cell of the table were calculated from the dependent variables when they were variates with approximate normal density (N) or from the subsequently dichotomized "normals" (D).

In order to test for trends in the data reported in tables, the margins were considered as fixed sources of variation and a multivariate analysis of variance, MANOVA, was performed on the frequency data three-tuples[5] within the table, employing the highest order interaction mean products as an estimate of error under the assumption that the highest order interaction is truly null[6].

---

[5]For α = .05, .025, and .01.

[6]An assumption checked by X^2 tests separately on each interaction variance estimate.

An initial analysis of the data in each of the tables indicated that the frequencies for tests based on the dichotomous variates were in each case significantly different from the frequencies for tests based on the normal variates.

A more detailed account of the results may be found in the hand out. However, without extensive reference to the results Figures 1. and 2. tell most of the story with respect to the "regular" variance ratio test of occasions under null occasion effect conditions. Note the interaction of the effects due to the number of subjects and the probability of a one in both figures. Note also the strong effect due to non-null items in Figure 2. The effect of the non-null items condition confirms the previously indicated confounding of a non-null items effect with occasions.

An examination of Tables 4. and 5. indicate something about the behavior of the Quasi-F test under null occasions effects. Note that its behavior is not particularly good, even when based on the "normal" data. Tables 6. through 9. contain the most startling results. Note that the Quasi-F tests based on normal variates responded in an appropriate manner to non-null occasion effects. However, the Quasi-F tests based on dichotomous data had significantly fewer frequencies in the α rejection regions when the effect being tested was non-null than they did when the effect being tested was null! Thus it would appear that Quasi-F tests based on dichotomous data, such as that simulated in this study, are biased tests!

Tables 10. and 11. and Figures 6. and 7. are with respect to the variance ratio tests of subject by occasion interactions. Although the results concerning the Quasi-F were the most startling of the results of this study, the most disappointing results were those concerning the tests of subject by occasion interactions. Examine in Figure 6. how the test becomes increasingly liberal with increases in the number of subjects under conditions where the probability of a one is not equal to .5.

In order to briefly examine the implications of this study, consider what the results might suggest to an experimenter who would like to do a repeated measures type of experiment and analyze his results with hypothesis testing in mind. First an experimenter who is considering doing a repeated measures type of experiment should consider the nature of the items or response evokers that will be employed in his experiment, examine them for possible confounding, and then employ a design and analysis which includes the items as a factor in it.

If an experimenter must have confounding in the analysis consistent with the design of his repeated measures experiment, he is in a somewhat difficult position with regard to analysis of variance testing of the source of variation with which confounding is present. For even if he can expect his dependent variables to have an approximately normal distribution, the results of this study suggest that the Quasi-F test will not have particularly good properties. If, on the other hand, the experimenter must evaluate responses in such a manner that his dependent variables are dichotomous, the Quasi-F test appears to be completely unacceptable.

On the other hand, if an experimenter finds that he can expect to have no confounding such as that mentioned above, but must have dichotomous dependent variables the results suggest, he can expect to appropriately employ the ordinary analysis of variance, variance ratio test for the occasion effect, if he employs enough subjects. The results also suggest that an experimenter should expect the power of analysis of variance tests based on dichotomous data to be between one third and one half the power he could expect if his dependent variables could be considered normal in distribution. The practical suggestion implied by the results of this study is, that if the above experimenter must have dichotomous dependent variables he should employ a still larger number of subjects in order to obtain a reasonable power situation.

The results of this study with respect to the test of the subject by occasion interaction when based on dichotomous data, suggest that even a very large variance ratio statistic may not indicate that the null hypothesis is false if there are a large number of subjects and the probability of a one is not close to .5 (see again Figure 6). For a probability of a one close to .5, however, it would appear that the probability of a Type I error is only approximately 1.2 times the nominal α level. Since the power of the test of the subject by occasion interaction based on dichotomous data with a .5 probability of a one was greater than half the power of the test based on normal data, the results suggest that when the probability of a one is close to .5 the test may be appropriately employed. If the probability of a one is not close to .5, however, the results suggest the above test may not be appropriately employed. As a practical suggestion, an experimenter who could expect possible subject by occasion interaction and who must have dichotomous dependent variables should endeavor to employ items or stimuli which will give him in general a .5 probability of a one in this data.

BIBLIOGRAPHY

Bradley, J. V. *Distribution-Free Statistical Tests.* New York: Prentice-Hall, 1968.

Donaldson, T. S. Power of the F-test for non-normal distributions and unequal error variances. Rand Corporation research memorandum RM-5072-PR, September, 1966.

Draper, J. F. A Monte Carlo investigation of the analysis of variance applied to non-independent Bernoulli variates. Unpublished dissertation, Michigan State University, 1971.

Greenberger, M. Notes on a new pseudo-random number generator. "Journal Assoc. Comp. Mach.," 8, 163-67.

Greenhouse, S. W. and Geisser, S. On methods in the analysis of profile data. "Psychometrika," 1959, 24, 95-112.

Hammersley, J. M. and Handscomb, D. C. *Monte Carlo Method.* New York: John Wiley, 1964.

Hsu, T. C. and Feldt, L. S. The effect of limitations on the number of criterion score values on the significance level of the F-test. "American Educational Research Journal," 1969, 6, No. 4, 515-28.

Hudson, M. D. and Krutchkoff, R. G. A Monte Carlo investigation of the size and power of tests employing Satterthwaite's synthetic mean squares. "Biometrika," 1968, 55, 431-33.

Lunney, G. H. A Monte Carlo investigation of basic analysis of variance models when the dependent variable is a Bernoulli variable. Unpublished dissertation, Univ. of Minnesota, 1968.

Mandeville, G. K. An empirical investigation of repeated measures analysis of variance for binary data. Paper presented at the annual meeting of the American Educational Research Association, 1970.

Marsaglia, G. and Bray, T. A. One line random number generators and their use in combinations. "Communications of ACM," 1968, 11, 757-59.

Satterthwaite, F. E. Synthesis of variance. "Psychometrica," 1941, 6, 309-16.

Seeger, P. and Gabrielsson. Applicability of the Cochran Q test and the F test for statistical analysis of dichotomous data for dependent samples. "Psychol. Bull.," 1968, 69, 269-77.

TABLES AND FIGURES TO ACCOMPANY THE PAPER ENTITLED
"A MONTE CARLO INVESTIGATION OF THE ANALYSIS OF
VARIANCE APPLIED TO NON-INDEPENDENT BERNOULLI VARIATES"

By

John F. Draper, Jr.
CTB/McGraw-Hill

Paper presented at the annual
meeting of the American Educational Research Association
April, 1972

INDEX

FIGURES PAGE

Table 1. Contrast of Experimental Situation 1 with Experimental Situation 2.

Situation 1	Occasion	O_1	O_2	O_3	O_4
	Item	$I_1 I_2 I_3$	$I_1 I_2 I_3$	$I_1 I_2 I_3$	$I_1 I_2 I_3$
Situation 2	Occasion	O_1	O_2	O_3	O_4
	Item	$I_1 I_2 I_3$	$I_4 I_5 I_6$	$I_7 I_8 I_9$	$I_{10} I_{11} I_{12}$

"Regular" Variance Ratio Test of Occasion Results

With respect only to the frequencies based on dichotomous variates in Table 2, it was concluded that there were significant main effects due to the probability of a one, Pr (1), and the number of subjects as well as an interaction between the two significant main effects. The significant interaction is represented in Figure 1. In the figure the two horizontal lines represent .95 probability limits[1] for mean frequencies such as those graphed, given an expected value of 50. Observing Figure 1, it appears that a favorable comparison of nominal and empirical probabilities of a Type I error occurred when the probability of a one was .5 and there were six or more subjects. Also a favorable comparison occurred when the probability of a one was .2 and there were ten or more subjects. However, when the probability of a one was .1 no favorable comparisons occurred, although the graph suggests that a favorable comparison might occur given more subjects.

Table 3 differs from Table 2 in only two aspects: (1) the values in the table were obtained by simulating data which could have arisen from situation 2 rather than situation 1, and (2) no frequencies appear with respect to tests based on normal variates.

From the analysis of the frequencies in Table 3, it was concluded that there were significant main effects due to: (1) the probability of a one, (2) the number of subjects, and (3) items null in effect versus items non-null in effect. There were also significant interactions between

[1]See Appendix A

Table 2. The Empirical Frequencies, in α = .1000 Rejection Regions, of the Variance Ratio Test for Occasions, Based on Normal and Dichotomous Data, Items Crossed with Occasions, Interaction and Occasion Effects Both Null

Probability of a One			.5						.2						.1					
n · 1000			50		25		10		50		25		10		50		25		10	
Dichotomous or Normal			D	N	D	N	D	N	D	N	D	N	D	N	D	N	D	N	D	N
No. of Sub.s.	Items	Level of Sub. Het.																		
4	null	1	51	44	26	20	7	8	41	41	21	24	12	11	41	67	15	42	7	15
		2	47	55	28	27	9	9	22	33	9	17	7	6	10	53	5	25	4	14
		3	39	46	18	27	4	16	14	56	6	33	2	15	12	50	2	21	2	7
		4	45	45	22	23	8	10	12	55	6	37	3	24	9	49	5	15	5	11
	non-null	1	47	50	28	25	3	11	27	56	9	30	1	19	26	44	14	20	7	12
		2	40	52	20	33	2	14	29	44	7	15	4	10	18	44	10	15	4	9
		3	27	50	18	29	6	11	27	51	11	25	3	14	13	36	1	14	1	6
		4	38	45	15	30	7	17	20	40	6	24	0	17	12	49	4	19	0	9
6	null	1	51	40	25	23	8	8	50	32	16	15	7	6	42	30	13	17	9	7
		2	66	50	31	22	6	6	30	32	19	14	6	6	37	24	6	12	6	3
		3	52	51	34	28	10	14	35	34	20	12	10	5	22	29	3	14	0	5
		4	61	57	38	27	14	12	39	32	21	15	2	8	16	25	6	12	0	5
	non-null	1	48	58	22	28	5	12	41	25	23	14	15	5	22	23	14	8	7	3
		2	46	49	21	27	6	10	30	29	12	16	5	7	18	29	9	12	7	4
		3	53	34	26	14	9	9	30	31	17	19	6	10	28	24	6	11	3	4
		4	51	38	15	13	5	6	38	28	20	13	3	5	32	22	11	4	0	1
8	null	1	48	36	30	19	11	8	34	36	16	25	7	4	31	29	19	16	1	4
		2	47	39	22	27	8	9	42	41	18	31	10	4	30	45	22	30	0	4
		3	59	47	29	26	14	13	35	41	17	32	12	6	40	38	8	30	0	0
		4	61	46	38	23	14	11	29	44	13	26	6	8	16	48	0	34	0	5
	non-null	1	53	52	26	30	10	11	52	34	15	24	7	4	39	45	19	31	9	7
		2	47	62	16	35	5	9	32	43	16	28	3	2	16	54	16	38	4	6
		3	66	51	40	37	19	9	55	32	7	27	3	3	26	54	9	37	2	4
		4	57	38	30	29	17	2	39	44	22	33	6	7	28	45	8	27	1	5
10	null	1	46	55	30	29	9	15	36	49	21	26	5	16	28	49	22	25	13	10
		2	42	46	16	21	5	10	45	59	19	28	4	13	42	57	18	30	6	8
		3	55	63	32	28	5	16	44	65	15	36	3	14	25	51	3	23	0	11
		4	53	52	16	28	7	13	64	53	16	29	4	14	20	65	1	40	1	13
	non-null	1	51	60	21	40	12	22	32	46	19	22	1	9	41	58	33	35	12	18
		2	57	55	30	29	12	17	47	62	20	35	8	16	36	43	12	25	7	15
		3	53	55	29	34	8	17	59	64	23	27	6	13	23	55	3	20	0	9
		4	52	59	29	37	12	21	30	59	18	40	5	14	23	43	0	35	0	12
12	null	1	46	53	28	27	11	13	44	62	23	35	2	8	60	53	19	36	4	9
		2	45	68	25	39	15	16	70	49	28	29	7	4	39	47	16	35	2	4
		3	65	55	23	27	8	8	37	50	15	26	4	11	36	46	14	36	8	12
		4	63	53	24	33	10	12	58	46	21	29	9	5	17	50	4	34	4	9
	non-null	1	59	57	29	33	7	18	47	56	35	37	13	14	31	45	25	28	10	6
		2	46	50	18	22	4	6	71	65	50	39	25	10	39	41	28	29	18	3
		3	53	80	28	43	12	18	60	54	33	34	22	18	45	31	25	22	11	5
		4	49	63	33	40	14	8	48	48	21	30	10	7	64	50	19	30	4	4

Table 3. The Empirical Frequencies, in α•1000 Rejection Regions, of the Variance Ratio Test for Occasions, Based on the Dichotomous Data, Items Nested within Occasions, Interaction and Occasion Effects Both Null.

Probability of a One			.5			.2			.1		
α • 1000			50	25	10	50	25	10	50	25	10
No. of Sub.s	Items	Level of Sub. Het.									
4	null	1	51	26	7	41	21	12	41	15	7
		2	47	28	9	22	9	7	10	5	4
		3	39	18	4	14	6	12	12	2	2
		4	45	22	5	12	6	3	9	5	5
	non-null	1	73	38	14	115	61	22	44	17	6
		2	56	30	15	72	47	14	19	5	0
		3	74	23	14	59	26	6	54	32	8
		4	62	26	5	69	28	9	29	10	0
6	null	1	51	25	8	50	16	7	42	13	9
		2	66	31	6	30	19	6	37	6	6
		3	52	34	10	35	20	10	22	3	0
		4	61	38	14	39	24	2	18	6	0
	non-null	1	100	68	37	120	63	26	69	36	21
		2	99	56	20	119	41	18	87	51	56
		3	125	67	24	119	48	24	104	56	45
		4	134	64	19	119	63	31	78	34	19
8	null	1	48	30	11	34	16	7	31	19	1
		2	47	22	8	42	18	10	30	22	0
		3	59	29	14	35	17	12	40	8	0
		4	61	38	14	29	13	6	16	0	0
	non-null	1	127	63	39	131	86	38	133	64	33
		2	125	78	30	153	105	64	82	23	11
		3	122	76	36	171	115	56	90	47	24
		4	115	56	31	156	101	57	105	71	21
10	null	1	46	30	9	38	21	5	28	22	13
		2	42	16	5	45	19	4	42	18	8
		3	55	32	5	44	15	3	25	3	0
		4	53	16	7	64	16	4	20	1	1
	non-null	1	145	83	42	155	110	77	200	122	60
		2	172	106	58	204	135	69	153	96	48
		3	162	89	46	207	141	67	111	50	30
		4	173	104	49	211	148	94	108	42	24
12	null	1	46	28	11	44	23	2	60	19	4
		2	45	25	15	70	28	7	39	16	2
		3	65	23	8	37	15	4	36	14	8
		4	63	24	10	56	21	9	17	4	4
	non-null	1	206	143	80	258	160	101	177	129	50
		2	188	126	61	253	185	123	213	108	54
		3	202	148	81	206	159	121	174	97	48
		4	174	128	84	268	164	107	160	90	55

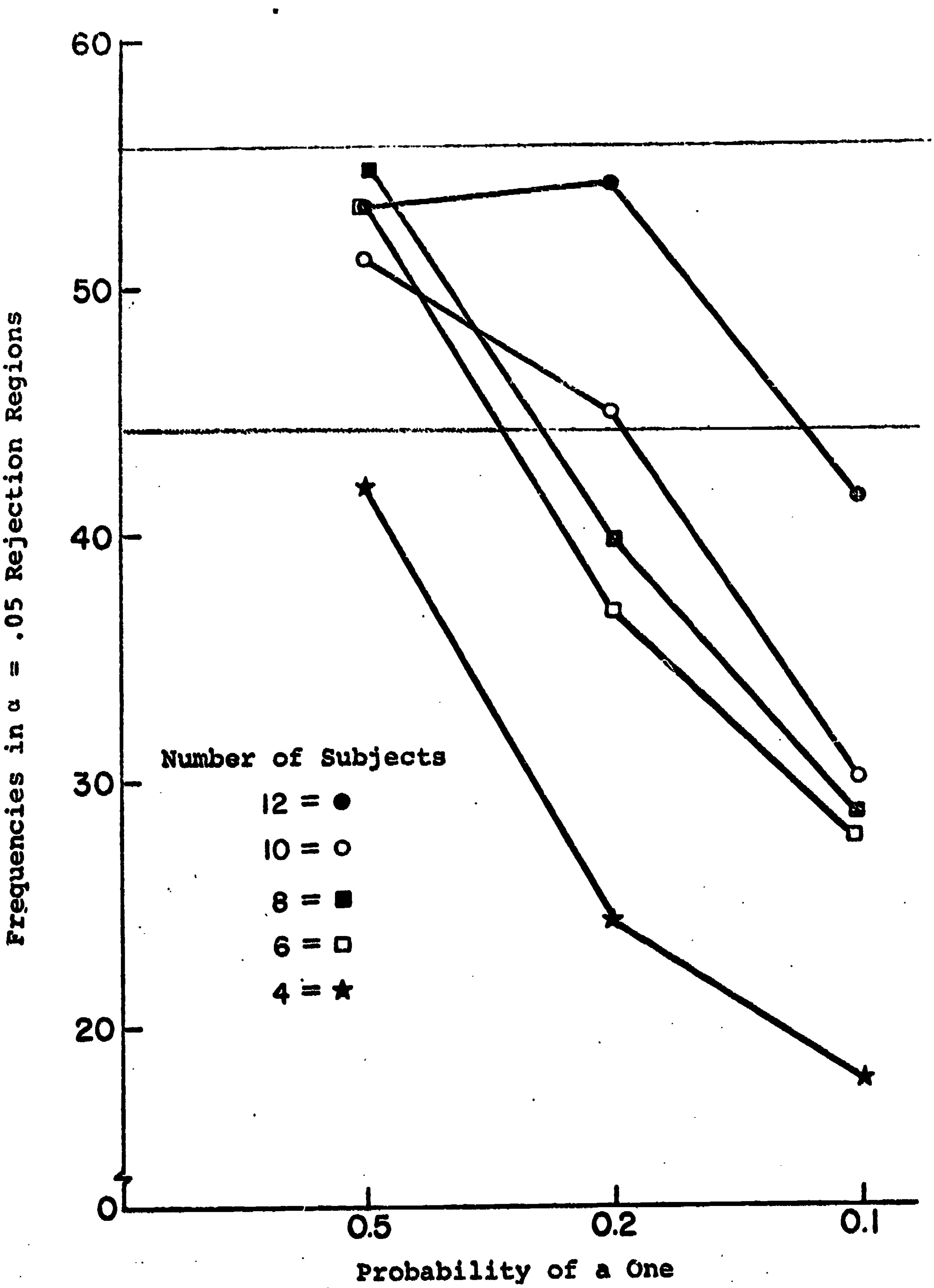

Figure 1. The interaction of the probability of a one and the number of subjects, with respect to the data in Table 2., data for the "regular" occasions test.

the probability of a one and items null vs. non-null and between the number of subjects and items null vs. non-null. Both of these interactions were represented on one graph, Figure 2. The horizontal lines in Figure 2 are .95 probability limits about 50, for means such as those graphed. An inspection of Figure 2 indicates that a favorable comparison of nominal α and empirical probabilities of a Type I error occurred when items were null in effect and the probability of a one was .5. Also a favorable comparison occurred when the items were null, the probability of a one was .2 and there were 10 or more subjects. In the absence of the above conditions, however, only unfavorable comparisons resulted.

The marginal mean vectors[2] for items null in effect was [40, 18, 6] and for items non-null it was [133, 78, 42]. The large mean frequencies for items under non-null conditions confirm the contention that non-null item effects are confounded with occasions and will cause the regular variance ratio test for occasions to have too many Type I errors.

In order to examine the behavior of the "regular" variance ratio test for occasions effects under non-null occasion conditions, an index of relative power was formed. The relative power of the test statistic[3] based on dichomous data was defined to be, *the frequency of the test statistic based on normal non-null data which fell into an α rejection region divided into the frequency in the same rejection region of that same statistic based on the same data after it had been subsequently dichotomized.* Relative power values were obtained for the non-null occasion effect analogs of all

[2]Ordered for frequencies in α = [.05, .025, .01] rejection regions.

[3]For each design model and parameter configuration.

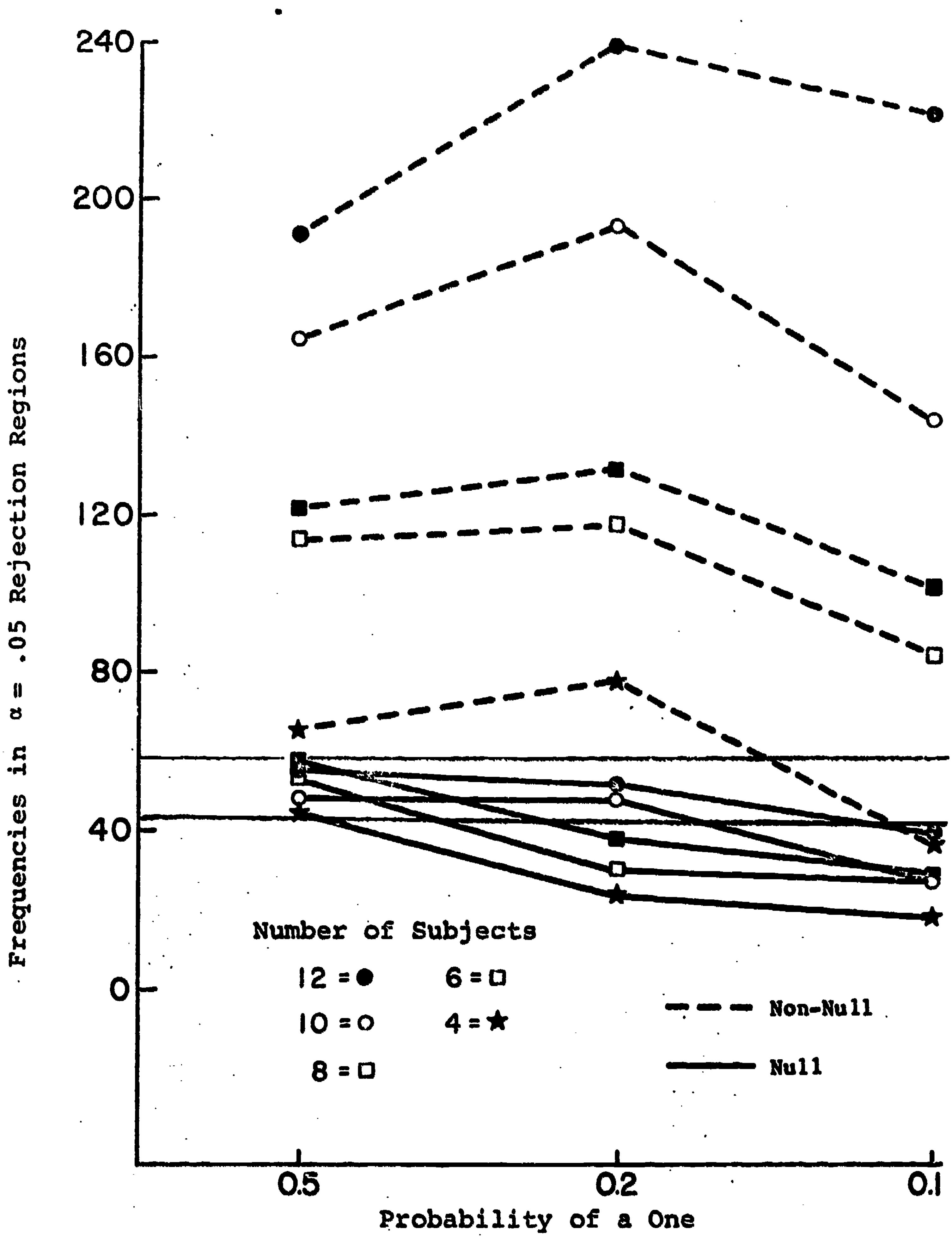

Figure 2. The interaction of the probability of a one, the number of subjects, and items null vs. non-null, with respect to the data in Table 3., data for the "regular" occasions test.

of the parameter configurations indicated in Table 2. These data were put to an arc-sin transformation and then analyzed by means of a MANOVA. This analysis indicated three significant main effects and no interactions. The significant effects were (1) the probability of a one, Pr (1), (2) the number of subjects, and (3) the level of subject heterogeneity. The overall mean vector of relative power variables was [.45, .38, .32], which indicates that the relative power of the variance ratio test for occasions effects decreased as the nominal α level decreased from .05, to .025, to .01. An ANOVA revealed no interaction between nominal α level and other sources of variation.

For probabilities of a one equal to .5, .2, and .1 the mean relative powers for $\alpha = .05$ were .58, .49, and .28 respectively. For numbers of subjects 4, 6, 8, 10, and 12, the means were .39, .41, .43, .50, and .54. For the four levels of subject heterogeneity 1, 2, 3, and 4 the means were .52, .46, .43, and .41 respectively.

The differences in mean relative power show clear trends. As the probability of a one becomes smaller so does the relative power, which is the opposite of the trend which might have been expected, since the degree of non-null effect in the simulated data was selected to counter the effect of decreased variance corresponding to a decreased probability of a one. Thus the power of tests based on the dichotomous variates should not have changed across levels of a probability of a one, whereas the power of the test based on normals should have and did decrease across levels of a probability of a one (and decreasing non-null effects). The results indicate however, that the power of tests based on the dichotomous variates fell off more rapidly across levels of the probability of a one than did the power of tests based on the normals. The explanation of the above may be that as the probability of a one becomes smaller the point of dichotomization is such that more of the "information" carried in the normals is lost. Another clear trend is that as the number of subjects increases, the relative power does so as well. The trend with respect to the number of subjects is most likely a function of the effect of the central limit theorem. The third trend indicates a loss in relative power with an increase in subject heterogeneity, a trend for which this investigator has at present no explanation.

Quasi-F Test of Occasion Results

The most startling result of this study concerned the empirical testing the application of the Quasi-F test of occasion on dichotomous data.

Tables 4 and 5 contain the frequencies in α = .05, .025, and .01 rejection regions of the Quasi-F ratio test statistics for tests of occasion effects when the data were simulated under null occasion effect conditions and null subject by occasion interaction effect conditions. That is Tables 4 and 5 are the Quasi-F analogs to Tables 2 and 3.

Again the data in which were with respect to normal variates were significantly different from the data in Table 4 which were with respect to dichotomous variates.

The analysis of the data with respect to the dichotomous variates indicated significant effects due to: (1) the probability of a one, Pr (i) (2) the number of subjects, the interaction of (1) and (2), and the interaction of (1) and items null vs non-null respectively. The above two significant interactions are represented in Figures 3 and 4 respectively.

Table 4. The Empirical Frequencies, in $\alpha \cdot 1000$ Rejection Regions, of the Quasi-F for Occasions, Based on Normal and Dichotomous Data, Items Crossed with Occasions, Interaction and Occasion Effects Both Null.

Probability of a One			.5						.2						.1					
$\alpha \cdot 1000$			50		25		10		50		25		10		50		25		10	
Dichotomous or Normal			D	N	D	N	D	N	D	N	D	N	D	N	D	N	D	N	D	N
No. of Sub.s	Items	Level of Sub. Het.																		
4	null	1	75	45	37	25	15	12	45	59	23	29	9	15	20	60	6	33	2	22
		2	63	47	33	26	19	12	61	42	39	17	15	6	16	54	4	29	2	14
		3	68	47	36	30	12	8	30	66	20	27	8	14	6	46	3	20	0	8
		4	51	64	26	32	10	17	27	56	19	33	6	18	6	61	3	33	0	14
	non-null	1	68	57	43	25	17	9	61	71	28	33	9	11	3	55	0	23	0	14
		2	55	57	28	27	15	12	34	47	26	23	10	10	8	59	5	33	2	17
		3	76	39	25	12	12	8	31	52	23	22	13	10	22	52	11	27	9	12
		4	57	44	26	25	12	12	23	62	12	38	3	19	16	50	14	23	4	13
6	null	1	78	39	45	18	18	8	74	35	36	17	21	3	42	55	19	27	11	5
		2	81	34	47	13	24	5	80	37	50	17	29	5	43	38	18	21	3	3
		3	73	45	46	25	18	9	71	32	31	19	16	5	26	52	8	29	1	4
		4	87	56	45	26	19	7	49	48	25	30	14	5	12	48	7	21	0	4
	non-null	1	77	39	41	20	13	7	66	43	28	19	14	7	43	41	7	21	0	2
		2	60	44	31	21	10	2	40	46	33	27	16	6	50	45	22	16	4	1
		3	83	43	49	21	26	6	60	49	44	26	13	7	38	37	15	13	0	0
		4	67	50	45	27	27	6	55	48	30	19	12	3	24	39	16	10	2	0
8	null	1	102	24	57	12	29	5	57	17	40	11	15	0	48	14	30	6	11	0
		2	81	30	52	17	30	7	71	16	29	9	13	0	41	26	21	14	9	0
		3	73	38	44	19	14	8	72	25	48	15	14	0	30	27	24	5	15	0
		4	67	23	47	17	23	7	80	25	51	16	17	0	37	25	26	13	6	0
	non-null	1	62	32	30	18	16	8	51	18	34	8	19	0	82	27	53	9	42	0
		2	72	36	43	21	26	10	54	19	38	5	16	0	96	33	58	12	35	0
		3	77	27	42	10	17	2	58	13	34	3	9	0	46	34	38	15	12	0
		4	64	14	40	5	16	0	60	20	38	9	19	0	53	21	30	5	0	0
10	null	1	98	37	59	19	25	10	94	29	76	6	31	6	43	24	14	3	7	2
		2	93	27	63	10	33	7	94	36	53	8	32	8	110	38	81	8	53	3
		3	92	38	53	15	27	8	104	42	51	12	29	9	58	21	29	6	12	3
		4	80	36	46	14	23	4	78	34	41	5	31	5	53	39	35	11	8	4
	non-null	1	92	34	62	12	26	7	73	29	39	4	5	2	62	31	33	10	18	7
		2	89	27	52	11	25	7	75	36	32	7	15	4	72	24	42	6	10	3
		3	84	29	44	12	24	6	98	27	63	2	39	2	92	30	51	8	15	4
		4	82	37	50	13	17	9	112	24	76	9	53	5	81	41	43	12	21	7
12	null	1	102	25	63	12	35	7	90	25	59	10	27	1	87	30	54	16	41	5
		2	83	46	50	20	27	5	74	24	51	10	26	2	89	27	63	22	27	2
		3	94	20	50	9	23	2	91	25	44	10	23	2	92	21	35	14	17	6
		4	80	28	59	6	35	0	79	13	39	7	11	0	74	31	53	19	27	5
	non-null	1	90	39	56	19	34	3	66	35	37	24	28	4	121	33	74	15	40	6
		2	84	22	41	10	24	1	67	35	49	20	19	1	101	17	49	10	30	3
		3	78	45	33	23	21	4	87	37	47	21	26	7	95	19	54	14	33	5
		4	71	21	47	10	27	2	96	34	57	12	40	3	68	34	38	15	22	4

Table 5. The Empirical Frequencies, in α·1000 Rejection Regions, of the Quasi-F for Occasions, Based on Dichotomous Data, Items Nested within Occasions, Interaction and Occasion Effects Both Null.

Probability of a One			.5			.2			.1		
α · 1000			50	25	10	50	25	10	50	25	10
No. of Sub.s	Items	Level of Sub. Het.									
4	null	1	61	37	14	27	15	9	8	1	0
		2	59	35	17	50	32	13	6	4	0
		3	62	29	11	25	12	6	4	4	3
		4	47	25	4	24	9	5	6	3	0
	non-null	1	59	38	14	45	25	16	18	4	4
		2	68	40	19	50	28	13	27	19	7
		3	64	37	13	42	13	5	32	18	11
		4	58	32	17	35	14	9	24	17	4
6	null	1	77	41	17	61	36	13	51	24	3
		2	62	39	22	68	38	20	47	20	10
		3	68	39	15	70	31	23	35	13	5
		4	85	43	16	52	21	11	16	10	2
	non-null	1	108	61	28	67	44	21	46	24	11
		2	86	55	28	61	35	23	35	17	8
		3	84	30	13	61	42	23	34	12	9
		4	72	46	16	64	37	21	47	25	11
8	null	1	92	52	23	54	19	16	70	25	11
		2	73	44	18	51	29	9	54	20	15
		3	66	38	14	65	35	14	53	30	22
		4	67	41	16	83	47	24	37	18	1
	non-null	1	102	65	33	74	45	17	44	23	8
		2	92	55	34	85	49	33	46	32	11
		3	140	101	51	113	62	38	85	36	25
		4	119	85	54	86	46	27	64	38	20
10	null	1	82	49	22	90	57	35	71	27	13
		2	100	51	23	78	49	28	93	66	41
		3	87	48	25	80	50	31	66	29	12
		4	88	53	21	73	50	35	67	38	11
	non-null	1	106	59	32	96	63	52	50	21	12
		2	102	65	48	134	82	52	56	34	24
		3	113	70	28	114	85	50	61	41	24
		4	87	46	28	113	75	44	45	34	10
12	null	1	101	67	39	102	54	29	94	46	31
		2	78	48	19	92	53	28	96	59	33
		3	77	49	27	82	47	26	105	58	30
		4	97	69	25	85	48	12	77	52	16
	non-null	1	123	77	40	85	63	29	103	82	41
		2	122	87	68	72	45	28	213	108	54
		3	128	89	58	81	53	37	102	68	44
		4	174	128	84	115	60	31	101	43	31

Both figures indicate that the empirical probability of a Type I error is not in general close to .05. Figure 3 indicates that although the frequency in the rejection region is less affected by the probability of a one as the number of subjects increases, the tests become rather liberal. Figure 4 is self explanatory.

The analysis of the data in Table 4 with respect to the normal variates was interesting in that it tends to contradict earlier findings. The analysis indicated a strong significant effect due to the number of subjects and indicates a general downward trend in the empirical probability of a Type I error with an increase in number of subjects.

The data in Table 5 were analyzed by means of a multivariate analysis of variance and significant effects were found with respect to: (1) the probability of a one, (2) the number of subjects, (3) items null vs. non-null, and the interaction of (1) and (2). The significant interaction is represented in Figure 5 which is somewhat similar to Figure 3 and which in general lends itself to the same interpretation.

Tables 6 and 7 contain frequencies in α = .05, .025, and .01 rejection regions of the Quasi-F test statistics for tests of occasion effects when the data were simulated under non-null occasion effect conditions. The results in Tables 6 and 7 are most startling for it is apparent that although the Quasi-F tests based on normal variates responded in an appropriate manner to non-null effects that the Quasi-F tests based on dichotomous variates did not. The Quasi-F test based on dichotomous variates had significantly fewer frequencies in rejection regions when the effect it was testing was non-null than it did when the effect it was testing was null, and in all of the cases investigated the empirical power of the Quasi-F test was consistently less than the nominal α level!

Table 6. The Empirical Frequencies, in the α=.1000 Rejection Region, of the Quasi-F for Occasions, Based on Normal and Dichotomous Data, Items Crossed with Occasions, Occasion Effects Non-null, Interaction Null.

Probability of a One			.5						.2						.1					
α=.1000			50		25		10		50		25		10		50		25		10	
Dichotomous or Normal			D	N	D	N	D	N	D	N	D	N	D	N	D	N	D	N	D	N
No. of Sub.s	Items	Level of Sub. Het.																		
4	null	1	28	298	13	189	6	107	20	186	16	91	7	43	29	111	17	67	12	34
		2	20	328	9	225	3	129	26	208	16	126	3	65	8	109	8	66	6	36
		3	11	384	7	258	4	155	18	244	15	133	8	62	7	140	6	68	2	35
		4	6	490	2	358	1	216	15	289	7	184	3	82	5	151	5	88	0	44
	non-null	1	24	288	10	205	5	122	33	183	19	101	10	56	14	113	4	55	1	32
		2	17	342	11	237	3	125	14	193	2	109	1	55	12	118	3	61	2	32
		3	12	346	4	232	1	126	25	244	11	141	5	78	13	136	9	67	5	25
		4	13	463	5	345	1	216	11	280	1	194	0	94	5	151	1	100	1	47
6	null	1	14	412	6	275	2	150	22	232	11	168	4	61	43	142	25	65	10	30
		2	5	474	2	343	1	222	23	266	8	183	5	93	45	166	11	79	3	44
		3	11	545	6	406	3	277	28	289	7	202	7	123	25	177	10	91	4	51
		4	6	656	3	516	2	354	13	386	6	264	2	149	14	219	9	118	1	64
	non-null	1	9	403	3	286	0	182	22	236	15	155	8	73	51	127	19	64	10	39
		2	8	439	4	309	3	202	18	257	8	189	7	84	48	142	20	67	0	47
		3	7	544	4	415	2	276	29	265	24	197	19	98	38	178	18	89	10	48
		4	7	602	1	472	1	319	23	401	19	300	11	193	34	221	27	126	14	69
8	null	1	10	544	9	389	4	233	21	308	10	219	8	112	48	218	19	113	1	48
		2	5	629	1	497	0	299	32	342	17	241	6	145	42	209	22	109	13	58
		3	4	697	2	567	0	393	20	420	9	276	9	173	30	253	12	176	5	92
		4	2	792	1	690	1	529	22	515	14	376	7	228	31	295	25	238	20	131
	non-null	1	5	582	4	443	1	273	34	315	23	218	14	118	41	207	24	112	16	53
		2	11	614	5	487	3	314	18	353	4	241	3	147	57	233	34	133	24	76
		3	15	672	9	579	7	399	24	438	12	298	0	194	57	245	20	151	20	91
		4	8	790	2	653	0	512	16	520	14	388	0	237	37	304	14	223	0	109
10	null	1	9	686	6	559	4	371	9	379	5	291	2	207	56	234	16	163	16	77
		2	5	732	4	582	4	436	18	441	10	307	6	229	39	249	21	189	14	107
		3	2	915	2	702	1	525	13	525	6	377	2	269	7	294	1	203	0	122
		4	2	910	0	804	0	658	11	620	2	475	0	299	20	343	7	231	5	141
	non-null	1	3	689	1	534	1	347	24	356	18	287	2	197	41	223	18	160	11	82
		2	5	760	3	620	0	440	10	449	10	314	1	230	43	266	30	196	13	118
		3	2	827	0	726	0	532	21	526	12	357	4	255	30	330	14	225	0	142
		4	2	911	2	784	0	644	22	611	5	465	0	311	23	343	13	231	5	151
12	null	1	3	734	0	622	0	444	10	432	6	313	3	173	46	353	32	119	24	53
		2	1	802	1	679	0	497	12	462	6	331	3	208	36	254	25	139	22	82
		3	0	883	0	777	0	625	4	569	4	404	2	274	40	319	31	221	18	115
		4	0	937	0	878	0	728	0	666	0	519	0	367	47	330	31	257	17	141
	non-null	1	0	514	0	387	0	278	9	402	4	308	0	163	52	261	38	131	21	63
		2	1	528	0	420	0	279	13	494	6	360	0	221	45	267	22	146	6	71
		3	1	591	1	447	0	331	6	548	5	391	0	246	33	318	9	239	7	110
		4	2	603	2	479	2	364	6	616	2	497	2	345	35	329	35	271	24	150

Table 7. The Empirical Frequencies, in α·1000 Rejection Regions, of the Quasi-F for Occasions, Based on Dichotomous Data, Items Nested within Occasions, Occasion Effects Non-null, Interaction Null.

Probability of a One			.5			.2			.1		
α·1000			50	25	10	50	25	10	50	25	10
No. of Sub.s	Items	Level of Sub. Het.									
4	null	1	25	14	8	17	8	4	15	10	4
		2	12	4	2	28	13	2	5	3	0
		3	8	4	2	16	13	5	13	1	0
		4	4	2	0	8	4	0	7	7	0
	non-null	1	30	17	8	38	25	12	16	3	1
		2	36	13	6	47	26	17	15	9	5
		3	30	12	5	24	9	9	21	8	5
		4	16	9	5	23	13	6	23	7	3
6	null	1	13	6	4	26	13	4	45	32	8
		2	10	4	1	25	6	2	39	31	8
		3	15	5	3	29	16	3	28	23	6
		4	10	3	3	11	4	2	18	12	1
	non-null	1	27	17	7	44	25	12	40	23	13
		2	13	4	2	38	31	10	25	8	4
		3	13	6	2	44	33	16	36	24	8
		4	14	6	3	26	22	11	50	20	4
8	null	1	9	7	2	17	8	3	59	20	7
		2	3	1	0	32	9	3	49	17	11
		3	4	4	2	22	6	1	33	11	4
		4	2	1	0	32	10	6	37	24	10
	non-null	1	24	14	2	60	43	32	72	34	12
		2	28	12	4	39	34	16	54	30	10
		3	17	14	4	45	29	16	53	28	15
		4	22	17	9	46	37	24	79	26	15
10	null	1	5	5	5	18	9	5	47	21	14
		2	4	3	3	9	5	2	32	21	9
		3	3	2	0	17	11	1	20	6	0
		4	0	0	0	10	3	0	18	14	12
	non-null	1	19	14	5	38	29	14	51	32	11
		2	9	8	3	52	39	23	49	33	22
		3	4	2	1	42	31	13	35	26	17
		4	6	3	1	14	10	6	7	7	3
12	null	1	3	0	0	8	3	3	37	32	11
		2	2	1	0	6	3	3	62	29	4
		3	1	1	0	8	3	2	55	31	6
		4	0	0	0	0	0	0	43	24	17
	non-null	1	8	4	1	41	20	11	51	39	27
		2	8	5	3	38	19	12	91	56	40
		3	10	10	6	33	20	5	85	57	37
		4	10	8	4	69	48	8	99	42	24

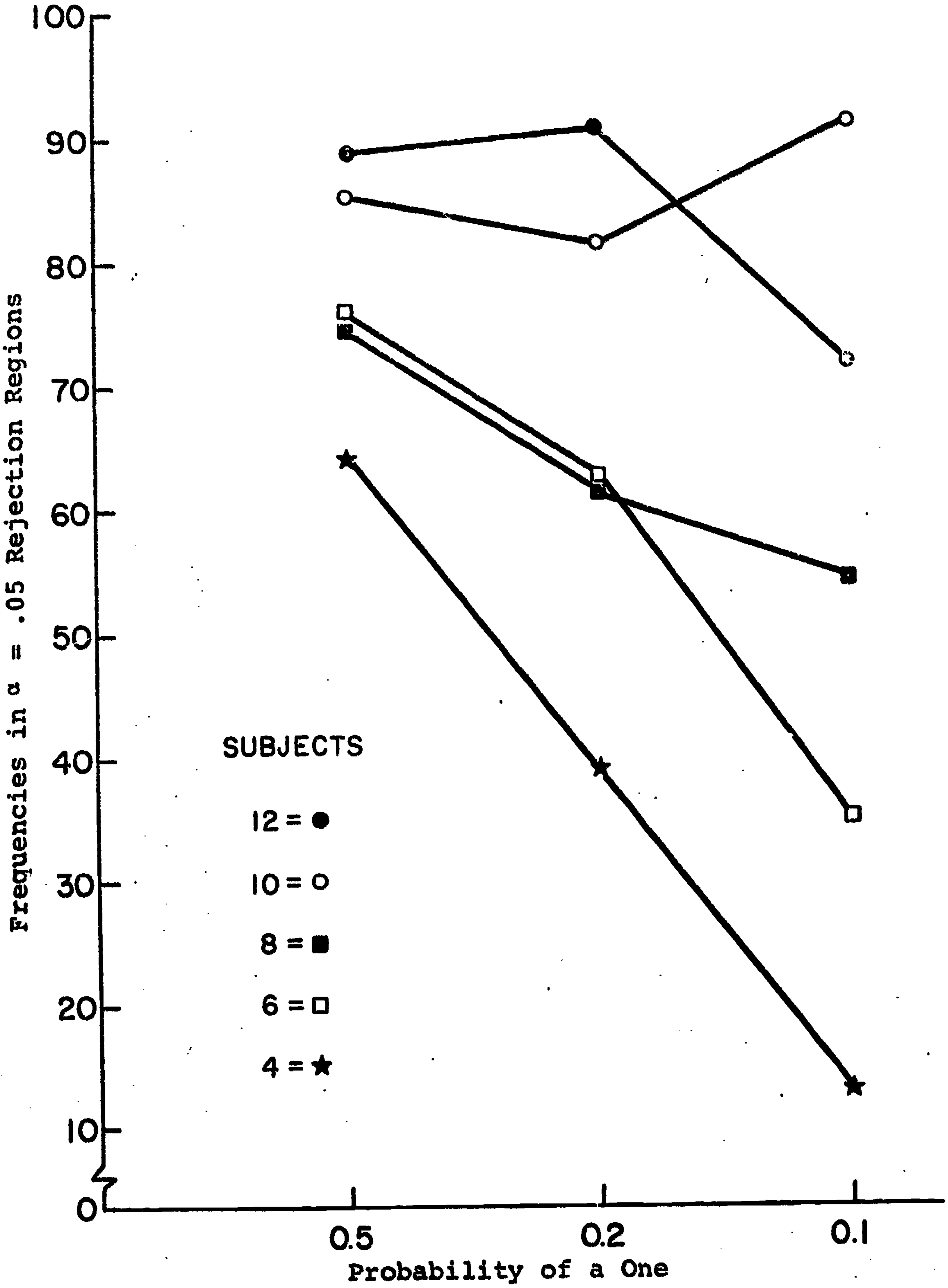

Figure 3. The interaction of the probability of a one and the number of subjects, with respect to the data in Table 4., data for the quasi-F test of occasions.

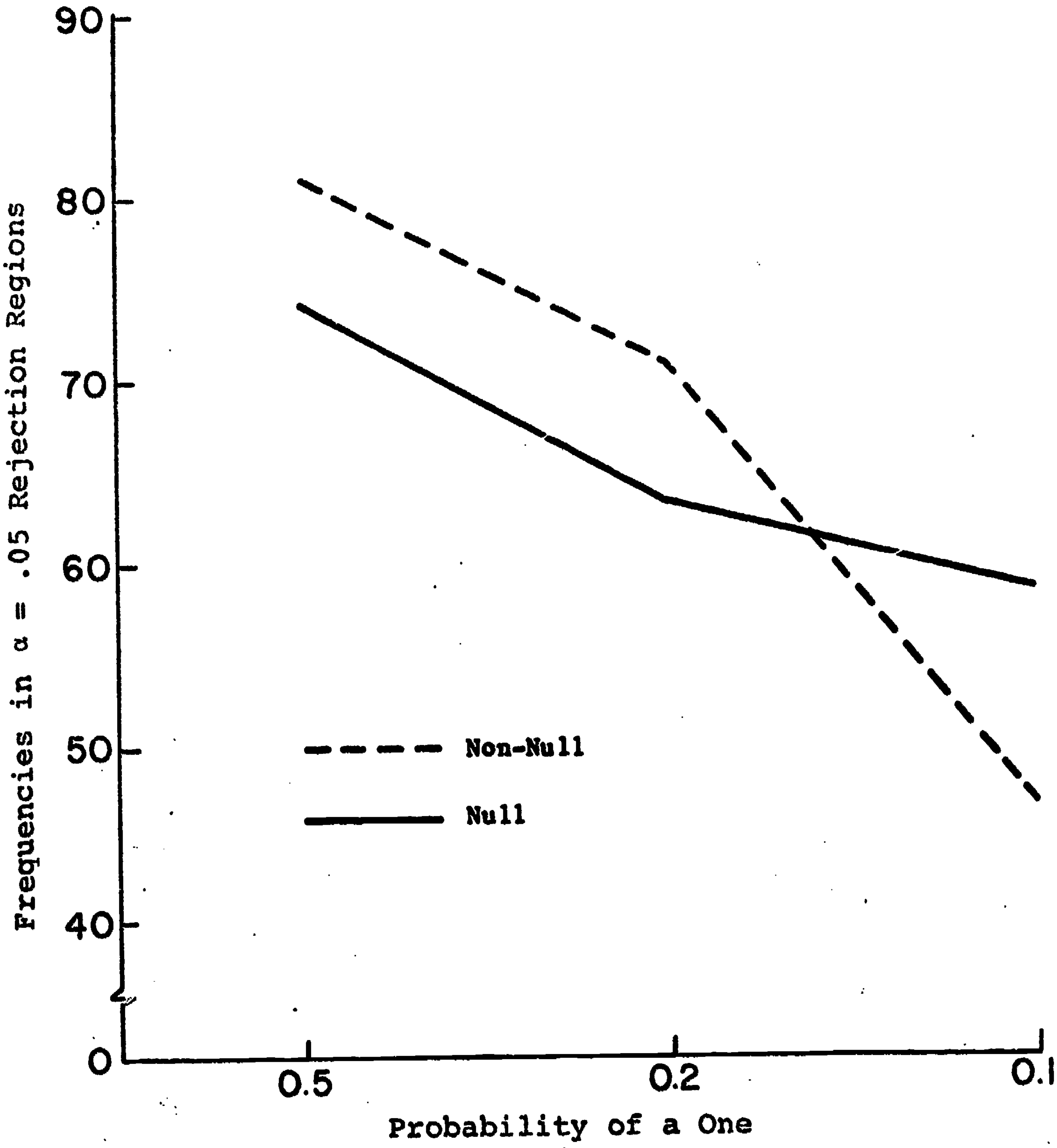

Figure 4. The interaction of the probability of a one and items null vs. non-null, with respect to the data in Table 4., data for the quasi-F test of occasions.

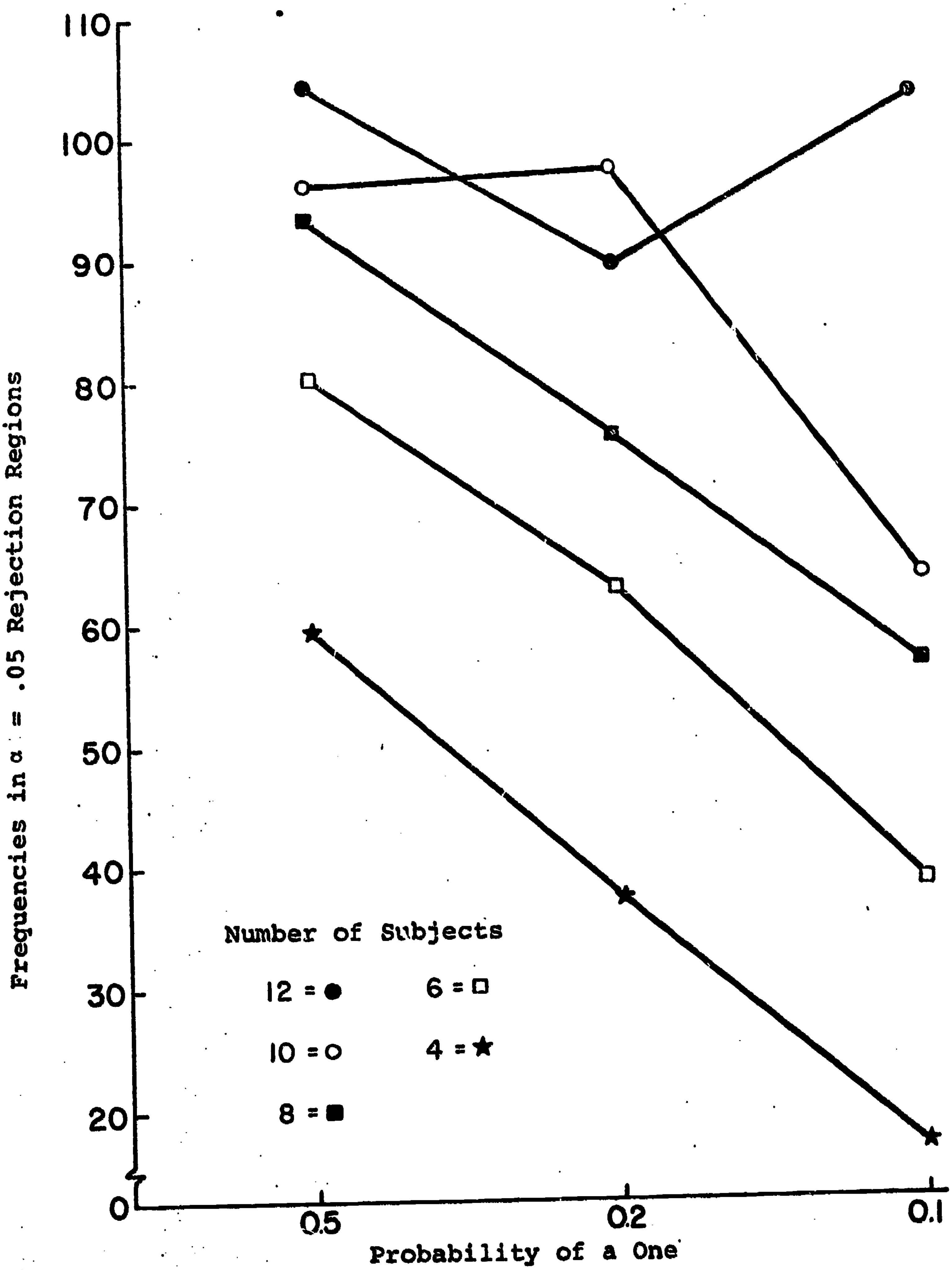

Figure 5. The interaction of the probability of a one and the number of subjects, with respect to the data in Table 5., data on the quasi-F test of occasions.

The data in Table 6 with respect to normal variates has an overall mean vector of [407, 300, 195] which is significantly less than the mean vector of analogous power data based on the regular variate ratio test.

Table 8 is a rearrangement of data which has been previously presented. The arrangement of data in Table 8 was established to allow easy contrast of data with respect to regular F and Quasi-F tests based on normal and dichotomous variates under null and non-null repeated measures conditions.

Table 9 is laid out in the same manner as Table 8 and includes some new data, that with respect to normal variates under situation 2.

Results of the Test of the Subject by Occasion Interactions

Tables 10 and 11 present frequencies in α = .05, .025, and .01 rejection regions for variance ratio tests of subject by occasion interaction effects, when the data were simulated under null occasion and null subject by occasion interaction effect conditions. The frequency data in Table 10 are with respect to situation 1 and the frequency data in Table 11 with respect to situation 2.

Again MANOVA indicated a significant difference between the data in Table 10 based on normal variates and the data in Table 10 based on dichotomous data.

The multivariate analysis of the data in Table 10 based on dichotomous variates indicated significant effects due to: (1) the probability of a one, (2) the number of subjects, (3) items fixed vs. random, (4) the level of subject heterogeneity, the interaction (1) and (2) and the interaction of (1) and (3). The two interactions are displayed graphically in Figures 6 and 7.

Inspection of Figure 6 indicates that a favorable comparison of empirical probability of a Type I error to nominal occurred only for 6 subjects and a

Table 8. The Empirical Frequencies, in α=.1000 Rejection Regions, of the Variance Ratio Test and Quasi-F for Occasions, Based on Normal and Dichotomous Data, Items Crossed with Occasions, Twelve Subjects.

Probability of a One				.5						.2						.1					
α = .1000				50		25		10		50		25		10		50		25		10	
Normal or Dichotomous				D	N	D	N	D	N	D	N	D	N	D	N	D	N	D	N	D	N
Occasions	F	Items	Level of Sub. Het.																		
Central	Regular	null	1	45	53	28	27	11	13	44	62	23	35	2	8	60	53	15	30	4	9
			2	45	68	25	39	15	16	70	49	28	29	7	4	39	47	16	35	2	4
			3	65	55	23	27	8	8	37	50	15	26	4	11	36	48	14	36	8	12
			4	63	53	24	33	10	12	58	46	21	29	9	5	17	50	4	34	4	9
		non-null	1	59	57	29	33	7	18	47	55	35	37	13	14	31	45	25	28	10	6
			2	46	50	18	22	4	6	71	65	50	39	25	13	39	41	28	29	18	3
			3	53	80	28	43	12	18	60	54	33	34	22	18	45	31	25	22	11	5
			4	49	63	33	40	14	8	48	48	21	30	10	7	64	50	19	30	4	4
	Quasi-	null	1	102	25	63	12	35	7	90	25	59	10	27	1	87	30	54	16	41	5
			2	83	46	50	20	27	5	74	24	51	10	26	2	69	27	63	22	27	2
			3	94	20	50	9	23	2	91	25	44	10	23	2	92	21	35	14	17	6
			4	80	28	59	6	35	0	79	18	39	7	11	0	74	31	53	19	27	5
		non-null	1	90	39	56	19	34	3	66	35	37	24	28	4	121	33	74	15	40	6
			2	84	22	41	10	24	1	67	35	49	20	19	1	101	17	49	10	30	3
			3	78	45	33	23	21	4	87	37	47	21	26	7	95	19	54	14	33	5
			4	71	21	47	10	27	2	96	34	57	12	40	3	68	34	38	15	22	4
Non-Central	Regular	null	1	615	898	494	820	351	699	365	609	263	507	174	349	158	343	85	244	45	142
			2	623	938	560	878	356	746	343	653	250	556	133	407	125	379	94	252	44	158
			3	661	977	532	929	397	852	396	771	269	656	191	512	147	464	75	357	36	227
			4	751	988	636	969	443	936	454	854	320	754	186	614	176	489	102	405	54	292
		non-null	1	611	915	472	835	333	719	365	558	255	462	157	349	150	350	117	255	54	171
			2	563	933	495	878	349	766	376	664	287	558	185	409	166	388	115	294	65	157
			3	666	966	538	930	387	858	405	752	298	642	176	495	157	487	119	340	41	220
			4	696	992	582	970	428	941	402	839	308	712	201	615	160	491	92	435	45	294
	Quasi-	null	1	3	734	0	622	0	444	10	432	6	313	3	173	46	353	32	119	24	53
			2	1	802	1	679	0	497	12	462	6	331	3	208	38	254	25	139	22	82
			3	0	890	0	777	0	525	4	569	4	404	2	274	40	319	31	221	18	115
			4	0	937	0	878	0	728	0	666	0	519	0	367	47	330	31	267	17	141
		non-null	1	0	514	0	387	0	273	9	402	4	308	0	163	52	261	30	131	21	63
			2	1	528	0	420	0	279	13	494	8	360	0	221	45	267	22	146	6	71
			3	1	591	1	447	0	331	6	548	5	391	0	246	33	318	9	239	7	110
			4	2	603	1	479	2	364	6	616	2	497	2	345	35	329	35	271	24	150

Table 9. The Empirical Frequencies, in ∝·1000 Rejection Regions, of the Variance Ratio Test and Quasi-F for Occasions, Based on the Normal and Dichotomous Data, Items Nested within Occasions, Twelve Subjects.

Probability of a One				.5						.2						.1					
∝·1000				50		25		10		50		25		10		50		25		10	
Normal or Dichotomous				D	N	D	N	D	N	D	N	D	N	D	N	D	N	D	N	D	N
Occasions	F	Items	Level of Sub. Het.																		
Central	Regular	null	1	46	53	28	27	11	13	44	62	23	35	2	8	60	53	19	36	4	9
			2	45	68	25	39	15	16	70	49	28	29	7	4	39	47	16	35	2	4
			3	65	55	23	27	8	8	37	50	15	26	4	11	36	48	14	36	8	12
			4	63	53	24	33	10	12	58	46	21	29	9	5	17	50	4	34	4	9
		non-null	1	206	276	143	207	80	127	258	445	160	339	101	266	177	486	129	369	50	291
			2	188	309	126	217	61	154	253	484	185	410	123	313	213	515	106	454	54	353
			3	202	334	143	245	81	181	206	534	159	476	121	352	174	546	97	459	48	380
			4	174	393	128	304	84	204	268	571	164	495	107	413	160	652	90	550	55	405
	Quasi-	null	1	101	40	67	14	39	8	102	34	54	14	29	4	94	46	46	20	31	9
			2	78	55	48	27	19	13	92	39	53	14	28	4	96	41	59	17	33	12
			3	77	29	49	7	27	3	82	38	47	13	26	4	105	40	59	14	30	9
			4	97	40	69	11	25	2	85	31	48	9	12	3	77	46	52	20	16	10
		non-null	1	123	84	77	44	40	20	85	104	63	89	29	51	103	93	82	71	41	50
			2	122	75	87	52	68	17	72	113	45	82	26	48	213	119	109	160	54	66
			3	126	89	89	66	58	28	81	109	53	77	37	47	102	127	66	107	44	68
			4	174	95	128	73	94	36	115	109	60	88	31	49	101	144	43	101	31	59
Non-Central	Regular	null	1	615	898	494	820	351	659	365	609	263	507	174	349	158	343	85	244	45	142
			2	623	938	500	878	356	746	363	653	250	556	133	407	125	379	94	292	44	158
			3	661	977	632	929	397	852	396	771	269	656	191	512	147	464	75	357	36	227
			4	751	988	636	969	448	936	452	654	320	754	186	614	176	489	102	405	54	292
		non-null	1	655	885	550	827	447	744	492	764	404	681	293	593	311	650	265	584	199	480
			2	686	909	590	969	479	795	499	818	404	720	317	604	272	707	214	593	182	479
			3	734	928	646	901	508	849	516	838	422	792	313	705	317	774	226	699	170	574
			4	750	933	667	907	541	876	515	907	445	871	333	791	284	834	221	749	136	676
	Quasi-	null	1	3	791	0	639	0	541	8	499	3	360	3	250	37	259	32	206	11	99
			2	2	649	1	745	0	600	6	525	3	413	3	273	62	262	29	175	4	91
			3	1	904	1	828	0	708	8	610	3	513	2	341	55	317	31	254	6	153
			4	0	949	0	914	3	842	0	698	0	591	0	406	43	366	24	306	17	197
		non-null	1	8	553	4	446	1	319	41	245	20	173	11	130	51	167	39	134	27	110
			2	9	577	5	469	3	354	38	278	19	189	12	137	91	179	56	124	40	86
			3	10	593	10	524	6	391	33	321	20	217	5	144	85	209	57	153	37	111
			4	10	632	8	523	4	416	69	321	48	229	8	143	99	224	42	167	24	132

Table 10. The Empirical Frequencies, in α=.1000 Rejection Regions, of the Variance Ratio Test for the Subjects by Occasions Interaction, Based on the Normal and Dichotomous Data, Items Crossed with Occasions, Interaction and Occasions Effects Both Null.

Probability of a One			.5						.2						.1					
α=.1000			50		25		10		50		25		10		50		25		10	
Dichotomous or Normal			D	N	D	N	D	N	D	N	D	N	D	N	D	N	D	N	D	N
No. of Sub.s	Items	Level of Sub. Het.																		
4	null	1	54	62	31	30	16	10	47	50	24	28	11	13	38	40	25	27	17	14
		2	54	40	26	25	7	10	62	65	37	33	24	20	46	57	30	20	27	13
		3	71	48	40	22	25	7	71	35	56	16	31	13	64	58	47	19	28	9
		4	54	41	26	20	12	8	65	35	50	14	36	10	47	35	32	13	21	7
	non-null	1	55	68	35	35	14	12	53	45	30	21	15	9	51	54	33	18	16	10
		2	49	62	29	29	22	13	49	43	24	17	13	8	57	46	36	19	29	12
		3	62	71	37	36	26	21	44	45	25	18	13	9	61	48	31	18	27	11
		4	75	51	46	32	26	18	47	52	24	25	21	15	64	53	37	27	32	18
6	null	1	58	43	29	20	10	7	46	46	50	23	12	5	56	66	31	39	12	18
		2	63	53	30	20	18	5	65	49	37	22	12	11	101	51	70	27	57	14
		3	65	34	41	19	17	3	81	62	57	29	23	8	116	55	83	27	52	10
		4	68	57	39	28	18	15	87	52	59	31	23	12	113	60	72	34	35	16
	non-null	1	39	37	23	15	10	10	61	48	36	24	25	8	52	45	33	29	17	20
		2	54	43	27	14	14	6	48	58	31	34	19	21	85	49	51	27	28	12
		3	63	45	43	26	15	9	51	51	41	31	34	11	81	56	49	33	37	20
		4	69	50	44	27	11	7	43	51	37	32	22	14	75	39	47	23	37	13
8	null	1	52	52	28	24	8	5	53	36	39	20	15	7	78	33	35	18	18	4
		2	58	58	31	31	14	11	66	40	41	16	28	7	111	43	79	27	67	7
		3	44	45	15	25	5	12	85	55	52	22	22	11	130	29	76	19	51	2
		4	73	47	36	19	22	10	90	53	49	30	28	13	122	43	83	22	49	7
	non-null	1	55	38	36	21	11	11	60	39	44	23	25	6	67	31	46	22	29	10
		2	49	42	23	18	10	10	70	33	45	19	19	11	74	38	70	25	36	4
		3	56	52	31	33	14	12	83	42	45	28	33	13	75	33	47	14	26	3
		4	60	49	35	22	17	13	84	38	39	26	25	8	76	36	49	19	15	4
10	null	1	68	60	41	29	15	15	64	49	45	16	15	15	83	51	54	26	25	22
		2	56	41	23	17	13	10	93	49	47	21	22	19	115	35	67	11	51	9
		3	65	45	31	23	13	11	112	52	57	20	38	17	142	47	109	25	63	20
		4	67	42	35	18	18	10	103	62	67	24	43	18	150	55	118	27	81	20
	non-null	1	48	41	18	21	8	11	66	61	28	17	20	14	112	60	47	29	30	26
		2	47	46	22	15	9	8	59	58	24	24	18	19	77	55	42	26	39	18
		3	49	49	21	18	10	12	88	52	58	26	24	21	130	52	64	23	37	21
		4	53	43	22	14	12	8	99	48	60	30	52	26	124	54	64	21	51	17
12	null	1	45	51	20	28	6	9	84	42	47	25	15	11	130	46	92	30	50	12
		2	40	52	23	28	11	16	110	52	72	35	38	19	121	58	92	35	55	20
		3	74	56	34	28	20	11	80	49	45	30	20	12	76	51	72	22	57	11
		4	66	46	42	31	12	7	108	50	80	35	24	9	153	50	129	34	91	16
	non-null	1	54	45	32	29	10	12	56	55	30	38	17	16	60	56	36	28	27	12
		2	62	59	27	40	15	19	61	44	51	34	26	8	104	53	76	26	50	10
		3	46	48	19	36	11	16	62	47	32	32	4	15	92	40	55	29	35	10
		4	65	45	29	24	14	12	121	47	62	26	42	14	100	47	83	26	45	15

Table 11. The Empirical Frequencies, in α·1000 Rejection Regions, of the Variance Ratio Test for the Subjects by Occasions Interaction, Based on the Normal and Dichotomous Data, Items Nested within Occasions, Interaction and Occasions Effects Both Null.

Probability of a One			.5						.2						.1					
α · 1000			50		25		10		50		25		10		50		25		10	
Dichotomous or Normal			D	N	D	N	D	N	D	N	D	N	D	N	D	N	D	N	D	N
No. of Sub.s.	Items	Level of Sub. Het.																		
4	null	1	59	57	32	22	16	10	50	53	29	20	9	12	46	38	31	26	25	14
		2	55	44	27	16	8	7	63	65	45	28	28	14	52	51	41	29	28	9
		3	64	52	34	24	18	6	61	42	47	18	24	7	71	46	55	26	30	12
		4	59	51	33	21	15	10	68	43	48	11	32	6	60	29	39	17	18	6
	non-null	1	55	49	26	28	11	14	50	42	13	19	9	8	54	49	42	22	23	10
		2	65	63	32	26	9	10	58	42	20	18	8	11	68	34	44	25	32	11
		3	54	55	28	29	9	13	76	52	39	19	14	4	61	39	50	26	21	14
		4	57	60	37	31	17	17	52	47	42	15	25	6	70	29	51	20	25	7
6	null	1	55	49	24	23	12	8	61	43	44	18	14	8	65	65	36	35	19	18
		2	69	50	42	28	22	8	70	47	41	23	21	16	99	59	64	29	57	16
		3	66	45	39	22	21	9	73	52	59	23	26	11	108	57	79	38	63	17
		4	85	64	42	32	21	17	81	48	69	30	36	17	106	54	61	31	58	14
	non-null	1	64	44	30	27	13	14	84	38	40	19	21	11	55	52	37	24	17	14
		2	69	47	39	25	16	10	80	44	56	18	35	10	73	42	51	23	51	9
		3	79	50	44	28	18	11	81	44	53	27	40	12	80	50	58	37	48	17
		4	72	57	37	30	25	14	102	52	70	34	52	15	86	52	60	33	39	13
8	null	1	61	49	36	30	12	8	51	44	31	14	10	2	55	52	40	18	21	6
		2	52	55	35	31	17	6	70	44	38	15	20	6	114	62	66	16	56	10
		3	57	45	26	27	9	10	60	58	51	21	26	5	123	48	65	17	56	8
		4	85	49	55	28	31	4	106	53	64	24	29	3	126	51	71	20	50	9
	non-null	1	61	43	30	22	9	7	69	43	37	22	18	7	77	68	44	17	35	6
		2	61	50	31	17	10	7	81	47	41	15	25	3	53	55	61	16	28	2
		3	51	59	36	23	18	2	76	48	46	14	22	4	92	62	51	13	35	0
		4	62	43	30	21	17	7	95	49	56	17	29	5	81	43	52	6	43	0
10	null	1	69	52	38	30	12	14	70	56	55	28	21	12	88	55	59	32	29	16
		2	67	50	24	25	7	12	85	49	42	30	28	15	112	45	98	17	59	7
		3	68	60	32	25	13	13	119	59	70	37	36	16	138	55	110	31	76	17
		4	59	46	36	23	18	13	104	56	67	34	46	17	146	69	109	36	73	18
	non-null	1	63	54	34	26	14	17	64	65	45	45	19	15	71	76	51	37	26	12
		2	76	56	51	33	22	14	88	62	61	34	35	15	106	73	71	40	45	16
		3	67	46	36	23	6	10	100	67	52	32	34	17	128	62	76	35	61	13
		4	99	54	62	25	24	13	95	65	77	37	28	19	141	61	118	38	92	18
12	null	1	50	53	21	28	10	14	82	47	49	28	19	17	119	40	57	33	41	17
		2	44	59	19	28	10	15	105	49	74	39	54	24	108	47	79	37	56	29
		3	84	72	47	35	17	12	90	50	61	37	21	18	93	49	73	26	41	14
		4	79	52	45	23	18	12	106	39	74	34	41	14	143	52	127	33	86	21
	non-null	1	62	79	42	42	25	16	91	39	41	34	17	20	70	49	51	36	46	13
		2	44	52	25	34	13	14	76	48	46	37	16	20	103	35	64	26	51	12
		3	46	50	33	30	12	8	56	56	26	40	3	23	110	42	100	25	64	14
		4	60	47	35	29	19	18	76	28	41	16	19	14	101	33	66	20	34	7

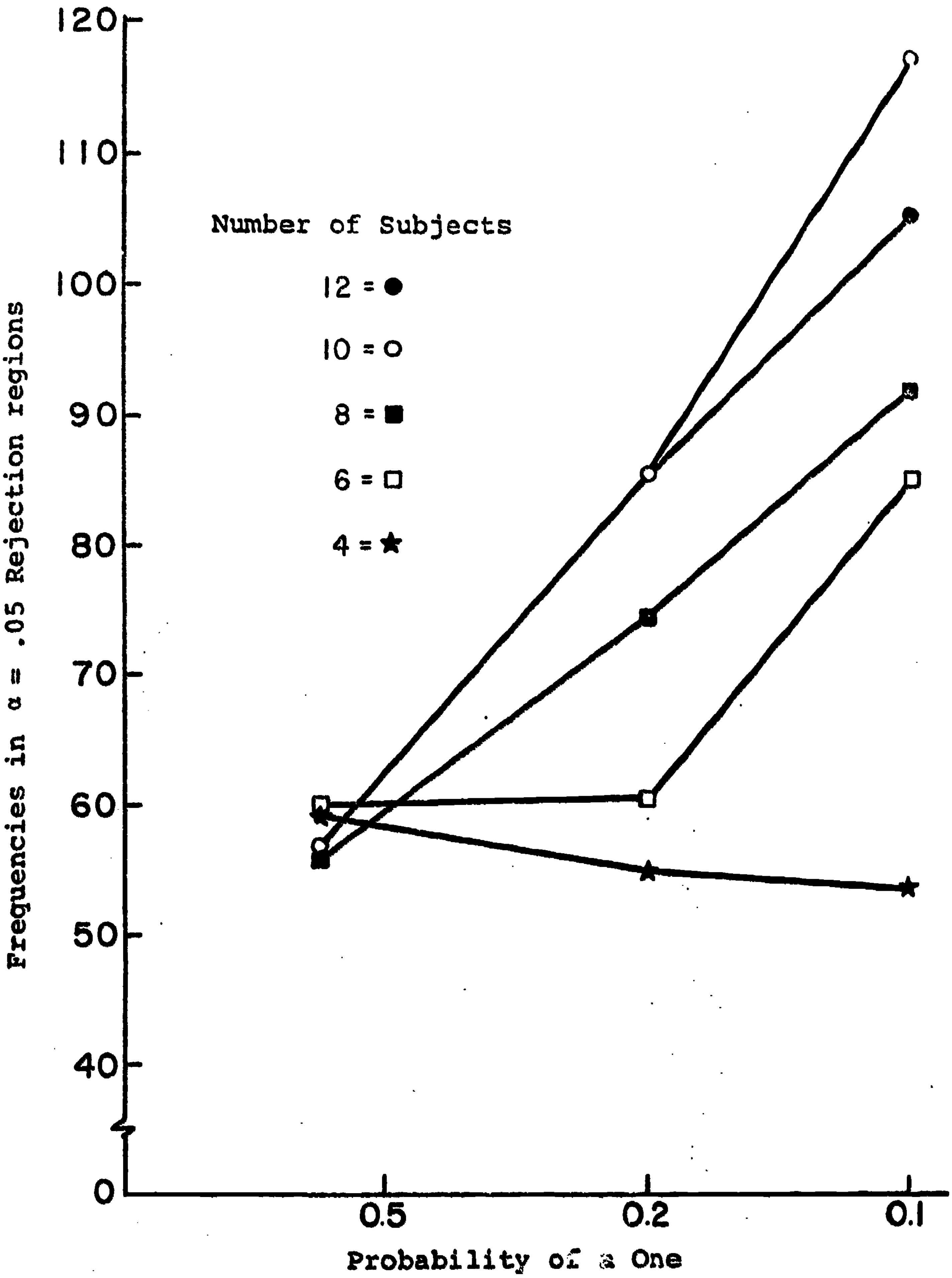

Figure 6. The interaction of the probability of a one and the number of subjects, with respect to the data in Table 10., data on the test of subjects by occasions interaction.

probability of a one equal to .5, and for 4 subjects and a probability of a one equal to .2 or .1. For all other conditions the test is too liberal. Figure 7 is self explanatory.

A multivariate analysis of the data in Table 11 indicated the same trends and significant effects that were found in Table 10.

Relative power variables were formed for both Table 10 and Table 11 and separate multivariate analyses of variance were performed on the relative power variables for both tables. In both analyses significant effects were found due to: (1) the probability of a one, (2) the number of subjects, the interaction of (1) and (2), the interaction of (1) and items fixed vs. random, and the interaction of (1) and the level of subject heterogeneity. In addition the analysis of the relative power variables from Table 11 disclosed a significant main effect due to items fixed vs. random.

In general it may be observed that higher relative powers correspond to greater "liberalness" in the test of a true null hypothesis. Thus the general increases in relative power observed across levels of a probability of a one .5 to .2 to .1 are in some sense specious.

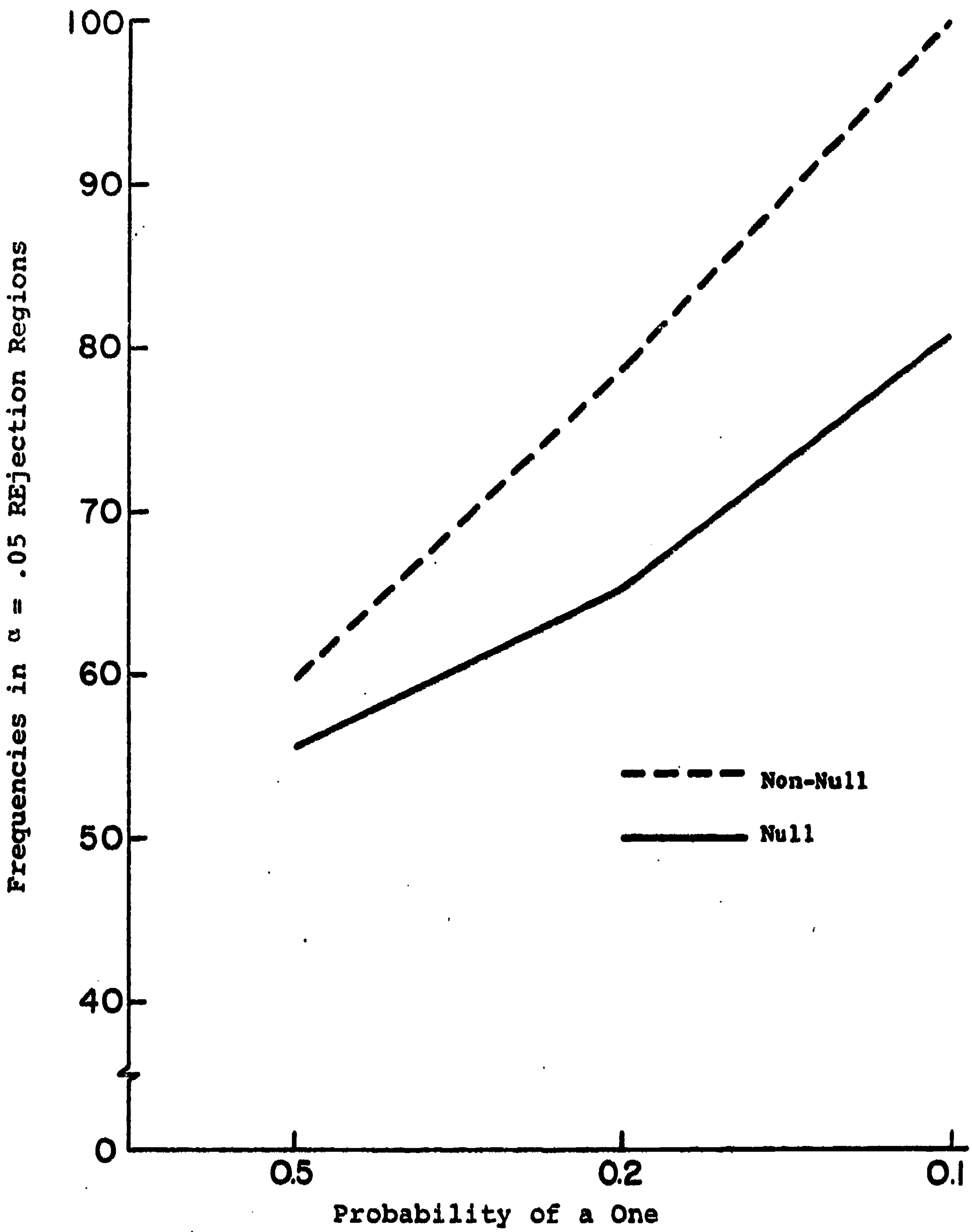

Figure 7. The interaction of the probability of a one and items null vs. non-null, with respect to the data in Table 10., data on the test of subjects by occasions interaction.

BIBLIOGRAPHY

Bradley, J. V. *Distribution-Free Statistical Tests*. New York: Prentice-Hall, 1968.

Donaldson, T. S. Power of the F-test for non-normal distributions and unequal error variances. Rand Corporation research memorandum RM-5072-PR, September, 1966.

Draper, J. F. A Monte Carlo investigation of the analysis of variance applied to non-independent Bernoulli variates. Unpublished dissertation, Michigan State University, 1971.

Greenberger, M. Notes on a new pseudo-random number generator. "Journal Assoc. Comp. Mach.," 8, 163-67.

Greenhouse, S. W. and Geisser, S. On methods in the analysis of profile data. "Psychometrika," 1959, 24, 95-112.

Hammersley, J. M. and Handscomb, D. C. *Monte Carlo Method*. New York: John Wiley, 1964.

Hsu, T. C. and Feldt, L. S. The effect of limitations on the number of criterion score values on the significance level of the F-test. "American Educational Research Journal," 1969, 6, No. 4, 515-28.

Hudson, M. D. and Krutchkoff, R. G. A Monte Carlo investigation of the size and power of tests employing Satterthwaite's synthetic mean squares. "Biometrika," 1968, 55, 431-33.

Lunney, G. H. A Monte Carlo investigation of basic analysis of variance models when the dependent variable is a Bernoulli variable. Unpublished dissertation, Univ. of Minnesota, 1968.

Mandeville, G. K. An empirical investigation of repeated measures analysis of variance for binary data. Paper presented at the annual meeting of the American Educational Research Association, 1970.

Marsaglia, G. and Bray, T. A. One line random number generators and their use in combinations. "Communications of ACM," 1968, 11, 757-59.

Satterthwaite, F. E. Synthesis of variance. "Psychometrica," 1941, 6, 309-16.

Seeger, P. and Gabrielsson. Applicability of the Cochran Q test and the F test for statistical analysis of dichotomous data for dependent samples. "Psychol. Bull.," 1968, 69, 269-77.

APPENDIX A

If 1,000 samples are simulated so that a null hypothesis is true and so that the assumptions of an ANOVA are met, the number of F-statistics testing the above hypothesis which have values which exceed the $F_{1-\alpha}$ quantile of a corresponding F-distribution[4] will be approximately 1,000α. If 1,000 samples of data are simulated so that a null hypothesis is not true and so that the assumptions of an ANOVA are met, the number of F-statistics testing the above hypothesis which have values which exceed the $F_{1-\alpha}$ quantile of a corresponding F-distribution will be approximately 1,000 times the nominal power (1-β) for the situation simulated.

Let X_i be defined according to the rule

$$X_i = \begin{cases} 1, & F_i > F_{1-\alpha} \\ 0, & \text{otherwise} \end{cases}$$

where F_i is the F-variate calculated on the i^{th} sample, i = 1, 2, ..., 1,000, and $F_{1-\alpha}$ is the 1-α quantile of the corresponding F-distribution. The variate X_i is then an indicator variable, which takes on the value one when F_i is in the α rejection region and which takes on the value zero when F_i does not occur in the rejection region. Note that the frequency of F_i's which fall in the α rejection region,

f, may be represented by the expression $f = \sum_{i=1}^{1000} X_i$.

Observe that f is a binominal variate with parameters p = α and n = 1000. The variance of f is then np(1-p) or for n = 1000 and α = .05, the var (f) = 47.5. Therefore by employing the normal approximation a, .95 probability interval may be formed about the expected value of f, E(f) = 1000 α, which for α = .05 is Pr(36.5 < f < 63.5) = .95.

---

[4]An F-distribution with the same degrees of freedom as the variance ratio for the test.