TRANSMISSION RATIOS AT THE *T*-LOCUS IN THE MOUSE: INTER- AND INTRA-MALE HETEROGENEITY

By A. W. H. BRADEN* and H. WEILER†

[*Manuscript received May* 7, 1964]

Summary

Forty-nine litter series sired by Tt^0, Tt^1, Tt^3, Tt^9, Tt^{12}, and Tt^{27} males of the house mouse were examined for heterogeneity of the transmission ratios of the *t*-allele. It was found that heterogeneity often exists, not only between males with the same *t*-allele, but even between litters from the same male; the latter was demonstrated by means of a new statistical technique. To measure the degree of heterogeneity, the *mean square contingency* was used as an index. It was found that intra-male heterogeneity was generally lower with "late" mating than with "normal" mating.

It appears that the behaviour in the female tract of spermatozoa carrying a *t*-allele differs from that of spermatozoa that do not have such an allele, and, further, that this behaviour is sensitive to slight variations in the physiological conditions in the tract.

The statistical method used for examining the significance of the intra-male heterogeneity is described in an Appendix. It is a Monte Carlo method necessitating the use of an electronic computer. The method is useful in situations where cell frequencies are too low for the application of a χ^2 test.

I. INTRODUCTION

While *Mus musculus* females transmit alleles of the *T*-locus normally (in equal proportions), males heterozygous for a *t*-allele usually have very aberrant transmission ratios (Chesley and Dunn 1936; Dunn and Gluecksohn-Schoenheimer 1939). It has been found that with such males the transmission ratio observed depends on:

(1) the particular allele involved—t^0, t^1, t^3, etc. (Dunn and Gluecksohn-Waelsch 1953);

(2) the time of mating in relation to ovulation (Braden, 1958);

(3) the genotype of the egg with respect to the *T*-locus (Bateman 1960; Braden 1960).

With (1), (2), and (3) held constant, one would expect a constant transmission ratio. Instead, it was found that heterogeneity between males continued to exist when (1) and (3) were controlled (Dunn 1943), and when attempts were made to control (2) also (Yanagisawa, Dunn, and [Bennett] 1961). Moreover, heterogeneity appeared to exist even between litters sired by a single male.

A statistical proof that such "within-male heterogeneity" really exists is difficult because the low cell frequencies of the relevant contingency tables make the ordinary use of χ^2 tests unreliable. A valid method of evaluating probabilities in

* Division of Animal Physiology, CSIRO, Ian Clunies Ross Animal Research Laboratory, Prospect, N.S.W.

† Division of Mathematical Statistics, CSIRO, Ian Clunies Ross Animal Research Laboratory, Prospect, N.S.W.

such situations is described in the Appendix and has been applied in this paper to 49 litter series from *Tt* males mated with wild-type (++) females. The results are on the whole in agreement with those of earlier workers.

II. Materials and Methods

Much of the data analysed in this paper has been given in two earlier papers (Braden 1958, 1960) in the form of total numbers of normal-tailed and Brachy offspring. Details of colony management at Albert Einstein College of Medicine, New York, where part of the work was done (results from males A, B, C, D, E, F, and G), have been given elsewhere (Braden 1958; Braden and Gluecksohn-Waelsch 1958). The strain of female employed there was a randomly-bred Swiss albino strain (herein designated SA). At the Division of Animal Physiology, CSIRO, Prospect, N.S.W., three randomly bred strains have been used: SW, an albino strain maintained by the Department of Veterinary Physiology, University of Sydney; LG, the Laboratory Animals Bureau's grey stock; and P, a coloured stock originating at and maintained by the Prospect Laboratory. The mouse room at Prospect is kept at 73°F, with a 12 hr light/12 hr dark regime. The mutant alleles *T* (Brachy) t^0, t^1, t^3, t^9, t^{12}, and t^{p1} have been used. The allele t^{p1} was extracted from a wild house mouse caught at Parramatta, N.S.W. It is a recessive lethal and is evidently not identical with any of the other five alleles (i.e. t^0–t^{12}), for males carrying t^{p1} and t^0, or t^{p1} and t^1, etc. have been found to be sterile or very infertile (Braden and Gluecksohn-Waelsch 1958).

Mating cages normally contained one male and four females. Females were removed from the mating cages either when a copulation plug was seen or when they were obviously pregnant. They were allowed to drop their litters in separate cages, and the offspring were classified according to tail morphology as soon as possible after birth. The genotype $+T$ has a shortened tail, whereas $+t$ has a tail of normal length. The progeny were then usually discarded, but the dams were not returned to the mating cages for at least 2 days, in order to avoid mating at the oestrus that normally occurs soon after parturition. Care was also taken to introduce females into the mating cages only in the afternoon so as to avoid the possibility of mating taking place late in oestrus. These were the only precautions taken to ensure that the matings took place "early", i.e. before ovulation. This procedure is termed "normal mating". The "late-mating" procedure employed in New York has been described (Braden 1958). At Prospect the lights were on from 6·00 a.m. to 6·00 p.m.; the males were placed in the late-mating cages for $2\frac{1}{2}$ hr from 8·00 a.m. Females with copulation plugs were removed and kept in separate cages until after parturition.

III. Results

Litters sired by seven Tt^0, two Tt^1, four Tt^3, five Tt^9, eight Tt^{12}, and one Tt^{p1} males were analysed. In 10 instances, two strains (SW and LG) of female were mated with the same male, and with 12 males, "normal" and "late" matings were run concurrently. A total of 49 sets of litters from *Tt* males (representing altogether 435 litters and 3153 offspring) were obtained and analysed by the Monte Carlo method described in the Appendix.

TABLE 1

HETEROGENEITIES WITHIN LITTER SERIES AND TRANSMISSION RATIOS

Normal mating: Tt males

Male Genotype	Male No.	Strain of Female	No. of Litters	Total Offspring	Transmission Ratio	Probability*	N†	C^2/n‡
Tt^0	6	SW	7	46	0·85	0·0294	§	0·29
		LG	4	27	0·89	0·672	§	0·05
	7	SW	10	78	0·82	0·990	400	0·03
		LG	10	67	0·46	0·0300	400	0·28
	8	SW	11	70	0·74	0·370	200	0·15
		LG	8	63	0·78	0·0000	800	0·36
	19	SW	14	79	0·91	0·703	400	0·13
		LG	9	70	0·64	0·530	100	0·13
	23	SW	12	90	0·72	0·130	100	0·19
		LG	7	43	0·67	0·0688	800	0·25
	E	SA	11	109	0·80	0·285	200	0·12
	G	SA	8	68	0·85	0·0443	§	0·21
Tt^1	A	SA	10	68	0·77	0·900	200	0·07
	B	SA	11	71	0·75	0·00750	400	0·29
Tt^3	9	SW	15	102	0·30	0·0457	2100	0·23
		LG	15	113	0·32	0·0120	500	0·26
	10	LG	9	58	0·43	0·580	100	0·12
	21	SW	5	35	0·29	0·726	§	0·07
		LG	5	27	0·26	0·329	1000	0·16
	D	SA	12	120	0·51	0·604	800	0·08
Tt^9	13	SW	4	21	0·29	0·374	§	0·16
	18	SW	8	50	0·20	0·312	§	0·17
	25	SW	9	42	0·24	0·574	§	0·17
	32	SW	7	51	0·37	0·120	200	0·23
	33	SW	12	77	0·29	0·730	200	0·11
		LG	3	12	0·50	0·610	§	0·13
Tt^{12}	3	SW	14	96	0·94	0·0842	§	0·21
		LG	4	23	1·00	1·000	§	0
	4	SW	13	83	0·81	0·840	200	0·08
		LG	13	88	0·80	0·0800	200	0·22
	5	SW	4	24	0·71	0·861	§	0·04
	17	SW	5	35	0·77	0·905	400	0·05
	40	SW	11	70	0·79	0·0700	200	0·21
	C	SA	9	45	0·89	0·403	§	0·19
	I	SA	8	72	0·93	1·000	§	0·03
	F	SA	6	61	0·90	0·294	§	0·10
Tt^{p1}	77	P	13	118	0·88	0·0588	800	0·16

* The probability of obtaining a distribution as heterogeneous as, or more than, the observed distribution (estimated by a Monte Carlo method, see Appendix).

† N = number of random permutations generated by electronic computer for each series (see Appendix).

‡ Estimate of the *mean square contingency* (see Appendix, Section III).

§ These probabilities have been calculated by the exact method (see Appendix).

Table 1 deals with litters from normal matings. It shows the average transmission ratios of *t*-alleles in *Tt* males for each litter series, and the estimated probability (under the null-hypothesis of equal probabilities for all litters in a series) of obtaining a distribution of *t*-alleles as heterogeneous as, or more so than, the observed distribution. These probabilities were subsequently combined in a significance test, as explained in the Appendix, Section II. Statistical significance was found, suggesting that heterogeneity does, in fact, sometimes occur amongst the litters produced by females mated under normal conditions with a single *Tt* male.

TABLE 2

HETEROGENEITIES WITHIN LITTER SERIES AND TRANSMISSION RATIOS

Late mating: *Tt* males

Male Genotype	Male No.	Strain of Female	No. of Litters	Total Offspring	Transmission Ratio	Probability*	N†	C^2/n‡
Tt^0	E	SA	9	68	0·62	0·550	100	0·09
	G	SA	10	81	0·56	0·220	100	0·16
Tt^1	A	SA	9	74	0·53	0·480	100	0·09
	B	SA	9	72	0·53	0·560	100	0·13
Tt^3	D	SA	5	50	0·24	0·787	§	0·04
Tt^9	25	SW	10	70	0·26	0·705	200	0·10
	32	SW	11	82	0·33	0·330	100	0·15
	33	SW	12	87	0·20	0·845	200	0·08
Tt^{12}	C	SA	8	80	0·84	0·0612	800	0·16
	I	SA	5	41	0·71	0·805	400	0·04
	F	SA	3	16	0·63	0·790	§	0·08
Tt^{p1}	77	P	8	60	0·48	0·00167	2400	0·35

* The probability of obtaining a distribution as heterogeneous as, or more than, the observed distribution (estimated by a Monte Carlo method, see Appendix)

† N = number of random permutations generated by electronic computer for each series (see Appendix).

‡ Estimate of the *mean square contingency* (see Appendix, Section III).

§ These probabilities have been calculated by the exact method (see Appendix).

Table 2 deals with litters from late matings in a similar way. Here, the combination of probabilities yielded no statistical significance, although two males of the series (Nos. 77 and C) were by themselves statistically significant or close to it.

A non-significant result is, of course, no proof of homogeneity, nor does it necessarily indicate that such a sample is less heterogeneous than another statistically significant sample (see Cramér 1946, p. 443). To measure and compare degrees of heterogeneity, we propose the use of the coefficient C^2/n, which is an estimate of the so-called *mean square contingency*. Its properties are summarized in the Appendix, Section III. The values of C^2/n are also listed in Tables 1 and 3.

Comparing the C^2/n values for the 12 males used for both normal and late matings with females of the same strain, it is seen that nine out of the 12 males had higher C^2/n values for the normal mating, suggesting that greater heterogeneity resulted from normal than from late matings.

A similar comparison between SW and LG females, mated normally with the same male, shows that seven out of the nine males available produced a more heterogeneous litter series with the LG females than with the SW ones. In addition, with two males (Tt^0, Nos. 7 and 19) the average transmission ratios for SW and LG females differed significantly. Finally, comparing the C^2/n values for males bearing different t-alleles, we find no striking difference between the various alleles, though litters sired by Tt^0 males appeared to be somewhat more heterogeneous than those sired by Tt^{12} males. It may be noted also that males with high transmission ratio alleles did not produce more heterogeneous litter series than males with low ratio alleles.

TABLE 3

HETEROGENEITIES WITHIN LITTER SERIES AND TRANSMISSION RATIOS

Normal mating: $++$ and $+T$ males

Male Genotype	Female Genotype	Male No.	No. of Litters	Total Offspring	Transmission Ratio*	C^2/n†
$+T$	$++$	68	20	151	0·58	0·11
		71A	26	170	0·53	0·11
		74	28	200	0·54	0·13
		75	21	154	0·49	0·10
		76	18	118	0·48	0·11
$++$	Tt	67A	28	209	0·52	0·13
		68A	15	112	0·49	0·14

* The transmission ratio listed for the first five litter series represents the male segregation ratio, i.e. the proportion of offspring inheriting the $+$allele; for the remaining two series it represents the female segregation ratio, i.e. the proportion of offspring inheriting the t-allele.

† See Appendix, Section III.

Table 3 presents data from normal matings where the sire did not carry a t-allele. Here, the values of C^2/n turned out to be remarkably uniform, ranging only from 0·10 to 0·14 for the seven litter series available. The values of C^2/n for the Tt males in Table 1 were much more variable, ranging almost uniformly from 0 to 0·36.

It may be seen from Table 1 that the transmission ratios vary not only between males with different t-alleles but also between males with the same t-allele. But heterogeneity has now been shown to exist even between the litters of one and the same male, suggesting that the variation may be due to the females. To gain an impression of the variation of transmission ratios between males with the same t-allele, as compared with the variation attributable to the females, the variance components were analysed for the transmission ratios of Table 1, treating SW and LG females separately. Variation between the five Tt^0 males was found to be of the same order but somewhat less than the variation between females. The same was true for

the five Tt^{12} males. Variation between the other males carrying identical *t*-alleles was found to be negligible.

IV. Discussion

Dunn and co-workers (Dunn and Gluecksohn-Schoenheimer 1939; Dunn 1943, 1960; Yanagisawa, Dunn, and Bennett 1961) have reported that there is a tendency for litters sired by males with a *t*-allele to be heterogeneous in the proportion of offspring inheriting the *t*-allele. However, the statistical methods hitherto available for examining this type of problem were either too cumbersome or unreliable (e.g. a χ^2 test is unreliable because of the low cell frequencies of the contingency tables). The present re-investigation of the problem, using a new statistical technique, establishes more firmly the existence of genuine heterogeneity between litters sired by individual *Tt* males.

This "within-male" heterogeneity appears to be associated with males that have a *t*-allele: the heterogeneity between litters sired by $++$ and $+T$ males was relatively low, even when *Tt* females were used. Males heterozygous for a *t*-allele also exhibit a number of other aberrancies. These all relate to the male segregation (or transmission) ratio of *t* and *T*, or *t* and $+$. The transmission ratios of *Tt* and $+t$ males do not only in most cases depart significantly from the Mendelian expectation of 0·5, but are also affected by:

(1) the particular *t*-allele carried by the male (Dunn and Gluecksohn-Waelsch 1953; Braden 1958, 1960);

(2) the particular male, selected from males with the same *t*-allele ("inter-male" heterogeneity, cf. Dunn 1943; Braden 1960; and present results);

(3) the genotype of the egg with respect to the *T*-locus (Bateman 1960; Braden 1960);

(4) the time of mating in relation to ovulation (Braden 1958; Yanagisawa, Dunn, and Bennett 1961);

(5) the individual female or the conditions pertaining to individual matings ("intra-male" heterogeneity, cf. Dunn 1943, Yanagisawa, Dunn, and Bennett 1961; and present results).

It seems, therefore, that the *t*-allele has an effect on the function of the spermatozoon in which it is located, and that the type or magnitude of the effect can be influenced by relatively small changes in the physiological state of the female tract. Restriction of mating to a relatively short period (2–3 hr), as in late mating, might be expected to reduce variation between females in the physiological state of the tract and thus reduce variation between litters in the transmission ratio, and this, in fact, was observed. Further discussion of the effects of late mating will be reserved for a subsequent paper.

The observed heterogeneity in the transmission ratios of males carrying the same *t*-allele may indicate that loci other than the *T*-locus influence spermatozoan function: there is, in fact, evidence that the *dilute–short-ear* region of chromosome II in the mouse may affect spermatozoan function (Russell and Russell 1960). However,

the investigations of Dunn (1943) showed that variation in transmission ratios of sib $+t$ males was not reduced by five generations of selection from a single parent male, suggesting that much of the inter-male heterogeneity did not have a genetic basis. The present finding that intra-male heterogeneity was usually as great or greater than the inter-male heterogeneity also suggests a non-genetic basis for the variation between males with the same t-allele. Possibly individual variation in mating behaviour, which would be influenced by both genetic and environmental factors, may be sufficient to explain any residual inter-male heterogeneity.

V. Acknowledgments

The authors gratefully acknowledge the assistance of Dr. P. J. Claringbold, Division of Animal Genetics, CSIRO, who initiated the Monte Carlo method and designed its programming; Mr. N. H. Westwood, who programmed the method; and Mrs. Jean Williams, Division of Mathematical Statistics, CSIRO, for preliminary computational investigations.

VI. References

Bateman, N. (1960).—Selective fertilization at the T-locus of the mouse. *Genet. Res.* **1**: 226–38.

Braden, A. W. H. (1958). Influence of time of mating on the segregation ratio of alleles at the T-locus in the house mouse. *Nature* **181**: 786–7.

Braden, A. W. H. (1960).—Genetic influences on the morphology and function of the gametes. *J. Cell. Comp. Physiol.* **56** (suppl. 1): 17–29.

Braden, A. W. H., and Gluecksohn-Waelsch, S. (1958).—Further studies of the effect of the T-locus in the house mouse on male fertility. *J. Exp. Zool.* **138**: 431–52.

Chesley, P., and Dunn, L. C. (1936).—The inheritance of taillessness (anury) in the house mouse. *Genetics* **21**: 525–36.

Cochran, W. G. (1954).—Some methods for strengthening the common chi-square tests. *Biometrics* **10**: 417–51.

Cramér, H. (1946).—"Mathematical Methods of Statistics." (Princeton University Press.)

Dunn, L. C. (1943).—A test for genetic factors influencing abnormal segregation ratios in the house mouse. *Genetics* **28**: 187–92.

Dunn, L. C. (1960).—Variations in the transmission ratios of alleles through egg and sperm in *Mus musculus*. *Amer. Nat.* **94**: 385–93.

Dunn, L. C., and Gluecksohn-Schoenheimer, S. (1939).—The inheritance of taillessness (anury) in the house mouse. II. Taillessness in a second balanced lethal line. *Genetics* **24**: 587–609.

Dunn, L. C., and Gluecksohn-Waelsch, S. (1953).—Genetic analysis of seven newly discovered mutant alleles at locus T in the house mouse. *Genetics* **38**: 261–71.

Fisher, R. A. (1932).—"Statistical Methods for Research Workers." Rev. Ed. (Oliver and Boyd: London.)

Mood, A. M. (1950).—"Introduction to the Theory of Statistics." (McGraw-Hill Book Co. Inc.: New York.)

Patnaik, P. B. (1949).—The non-central χ^2- and F-distributions and their applications. *Biometrika* **36**: 202–32.

Pearson, E. S., and Hartley, H. O. (1958).—"Biometrika Tables for Statisticians." Vol. 1. (Cambridge Univ. Press.)

Russell, L. B., and Russell, W. L. (1960).—Genetic analysis of induced deletions and of spontaneous non-disjunction involving chromosome 2 of the mouse. *J. Cell. Comp. Physiol.* **56** (suppl. 1): 169–88.

Yanagisawa, K., Dunn, L. C., and Bennett, D. (1961).—On the mechanism of abnormal transmission ratios at the T-locus in the house mouse. *Genetics* **46**: 1635–44.

APPENDIX

TESTING FOR HETEROGENEITY IN $2 \times r$ CONTINGENCY TABLES WHEN CELL FREQUENCIES ARE SMALL

I. TEST OF HOMOGENEITY

Consider males with an abnormal allele which they transmit to a proportion of their progeny. When such males are mated with normal females, it seems reasonable to expect approximately the same proportion of abnormal animals in all litters produced by the same male. The assumption that this is true constitutes our null-hypothesis H_0, which may be stated in precise terms as follows:

H_0: *The probability of a heterozygous male producing an abnormal animal is the same for all normal females.*

However, if the observed proportion of abnormal animals varies considerably from litter to litter it may be justifiable to discard the null-hypothesis in favour of the alternative hypothesis:

H_1: *The probability of a heterozygous male producing an abnormal animal depends also on the mother of the litter (or the circumstances of the mating).*

To derive a procedure of testing H_0 against H_1, we consider the mathematical model of r classes of litters, $L_1, L_2, \ldots, L_r$, where L_i $(i = 1, 2, \ldots, r)$ is the class containing the litters (past and future) of the ith female, produced under specified conditions. Suppose now that one litter is drawn at random from each litter class and that the litter taken from the ith class $(i = 1, 2, \ldots, r)$ contains n_i animals of which a_i are abnormal and $b_i = n_i - a_i$ are normal. Altogether, there are then $n = n_1 + n_2 + \ldots + n_r$ animals of which $a = a_1 + a_2 + \ldots + a_r$ are abnormal and $b = b_1 + b_2 + \ldots + b_r = n - a$ are normal. The litters actually obtained when one male is mated with r females are identified with the above litters, drawn at random from the classes $L_1, L_2, \ldots, L_r$. Thus, for instance, male 32 (Table 1) produced $r = 7$ litters with a total of $n = 51$ offspring, consisting of $a = 19$ abnormal and $b = 32$ normal animals, as shown in the following tabulation:

Litter:	L_1	L_2	L_3	L_4	L_5	L_6	L_7	Total
Abnormal	0	1	2	5	1	5	5	19
Normal	4	4	6	3	8	3	4	32
Total	4	5	8	8	9	8	9	51

Now, the actually observed situation is only one of a (usually large) number of possible allocations of a abnormal and b normal animals to litters of fixed sizes

$n_1, n_2, \ldots, n_r$. The general case where the litters contain (say) $x_1, x_2, \ldots, x_r$ abnormal and $y_1, y_2, \ldots, y_r$ normal animals, is represented below:

Litter Class	L_1	$L_2 \ldots\ldots L_r$	Total
Abnormal	x_1	$x_2 \ldots\ldots x_r$	a
Normal	y_1	$y_2 \ldots\ldots y_r$	b
Total	n_1	$n_2 \ldots\ldots n_r$	n

If p_i is the (unknown) probability that an animal taken from any litter of the class L_i is abnormal, the null-hypothesis to be tested may be stated in the form

$$(H_0) \qquad p_1 = p_2 = \ldots = p_r = p \text{ (say)}. \tag{1}$$

Now, consider the expression

$$C^2 = \sum_{i=1}^{r} \left[\frac{(x_i - n_i\bar{p})^2}{n_i\bar{p}} + \frac{(y_i - n_i\bar{q})^2}{n_i\bar{q}} \right], \tag{2}$$

where

$$\bar{p} = a/n, \ \bar{q} = b/n. \tag{3}$$

It is clear that the expression for C^2 will tend to be the larger the more the p_i differ from each other. Hence, if the value, C_o^2, actually obtained from the observed litters, is large, the null-hypothesis is discarded. More precisely, H_0 is discarded in favour of H_1 whenever $\Pr\{C^2 \geqslant C_o^2 | H_0\}$ (i.e. the probability that C^2 is greater than C_o^2, given the null-hypothesis H_0) is less than a preassigned significance level α (e.g. $\alpha = 0{\cdot}05$).

It is well known (Cochran 1954) that the expression for C^2 is approximately a χ^2 variate with $r-1$ degrees of freedom, provided the $n_i\bar{p}$ and $n_i\bar{q}$ (the expected cell frequencies) are not too small. The above probability can then be obtained easily by consulting a set of χ^2 tables.

When, on the other hand, the litter sizes are small or when $\bar{p}$ or $\bar{q}$ is small, the exact distribution of (2) has to be considered. For this it is necessary to consider all possible allocations of a abnormal and b normal animals to the r litters in the above tabulation. Of these, some will yield a value $C^2 \geqslant C_o^2$. Assuming H_0 to be true, the probability for each of these allocations is given by the formula (see e.g. Mood 1950, §12·10)

$$P = \frac{a!\, b!\, n_1!\, n_2! \ldots n_r!}{n!\, x_1! \ldots x_r!\, y_1! \ldots y_r!}, \tag{4}$$

and the sum of all these probabilities is equal to the required probability $\Pr\{C^2 \geqslant C_o^2 | H_0\}$.

But it soon becomes apparent that the computations required to calculate all the probabilities are prohibitive even with an electronic computer, except for a small number of small litters. A Monte Carlo method has therefore been designed by which $\Pr\{C^2 \geqslant C_o^2 | H_0\}$ can be estimated without calculating the probabilities (4). The method used, and executed by electronic computer, is as follows.

Denoting each of the a abnormal animals by 1 and each of the b normal animals by 0, the allotment of the $n = a+b$ animals into litters of sizes $n_1, n_2, \ldots, n_r$ can be simulated by partitioning any permutation of the n symbols 11 . . . 100 . . . 0 (consisting of a ones and b zeros) into groups of lengths $n_1, n_2, \ldots, n_r$, respectively. Now, if H_0 is true, all permutations of these symbols are equivalent to equally-likely litter distributions, so that the probability P in (4) is equal to $N(x_1,x_2, \ldots ,x_r)/n!$, where $N(x_1,x_2, \ldots ,x_r)$ denotes the number of permutations producing the allocation of $x_1,x_2, \ldots ,x_r$ abnormal animals to the litters $L_1,L_2, \ldots ,L_r$, and $n!$ is the number of all possible permutations. [Incidentally, it can be shown easily that

$$N(x_1,x_2, \ldots ,x_r) = a!b!n_1! \ldots n_r!/(x_1! \ldots x_r!y_1! \ldots y_r!),$$

which provides a simple proof of formula (4).] Instead of using formula (4) to calculate $\Pr\{C^2 \geqslant C_o^2|H_0\}$ we now determine the value of C^2 for a sufficiently large number of random permutations and find out what proportion amongst them yields $C^2 \geqslant C_o^2$. If amongst N permutations, drawn at random, N_1 have $C^2 \geqslant C_o^2$, the probability $\Pr\{C^2 \geqslant C_o^2|H_0\}$ is estimated by N_1/N.

Note 1

The above method of determining the probability $\Pr\{C^2 \geqslant C_o^2|H_0\}$ is based on the use of C^2 to establish a ranking of heterogeneity of all possible litter combinations. In practice, C^2 need not be used directly for this purpose, because the same ranking is obtained by means of the simpler expression

$$D^2 = \Sigma x_i^2/n_i. \tag{5}$$

It can indeed be shown that

$$C^2 = (n^2/ab)[D^2-(a^2/n)], \tag{6}$$

so that C^2 increases monotonely with D^2, when a, b, n remain constant. The above probability is therefore identical with $\Pr\{D^2 \geqslant D_o^2|H_0\}$, where D_o^2 is the value of D^2 obtained with the actual litters.

Note 2

The precision of the estimate N_1/N of the probability $\Pr\{C^2 \geqslant C_o^2|H_0\}$ is shown in Figure 1, which gives the 95% confidence limits for the estimated probability. The graphs are similar to that in Pearson and Hartley (1958) but extend over smaller ranges, more suitable for the present purpose. In this paper, sample sizes of less than 1000 were deemed to be sufficiently large in most cases. To test the accuracy of the programming of the Monte Carlo methods, the exact probabilities [formula (4)] were calculated for a few small litter series. The agreement was satisfactory.

Note 3

The use of C^2 (or D^2) for ranking is only one of many possibilities. Another way of ranking for heterogeneity would be the one where, of two litter combinations, the one with the smaller probability (under H_0) is regarded as the more heterogeneous one. The hierarchy established in this way is not always the same as that established

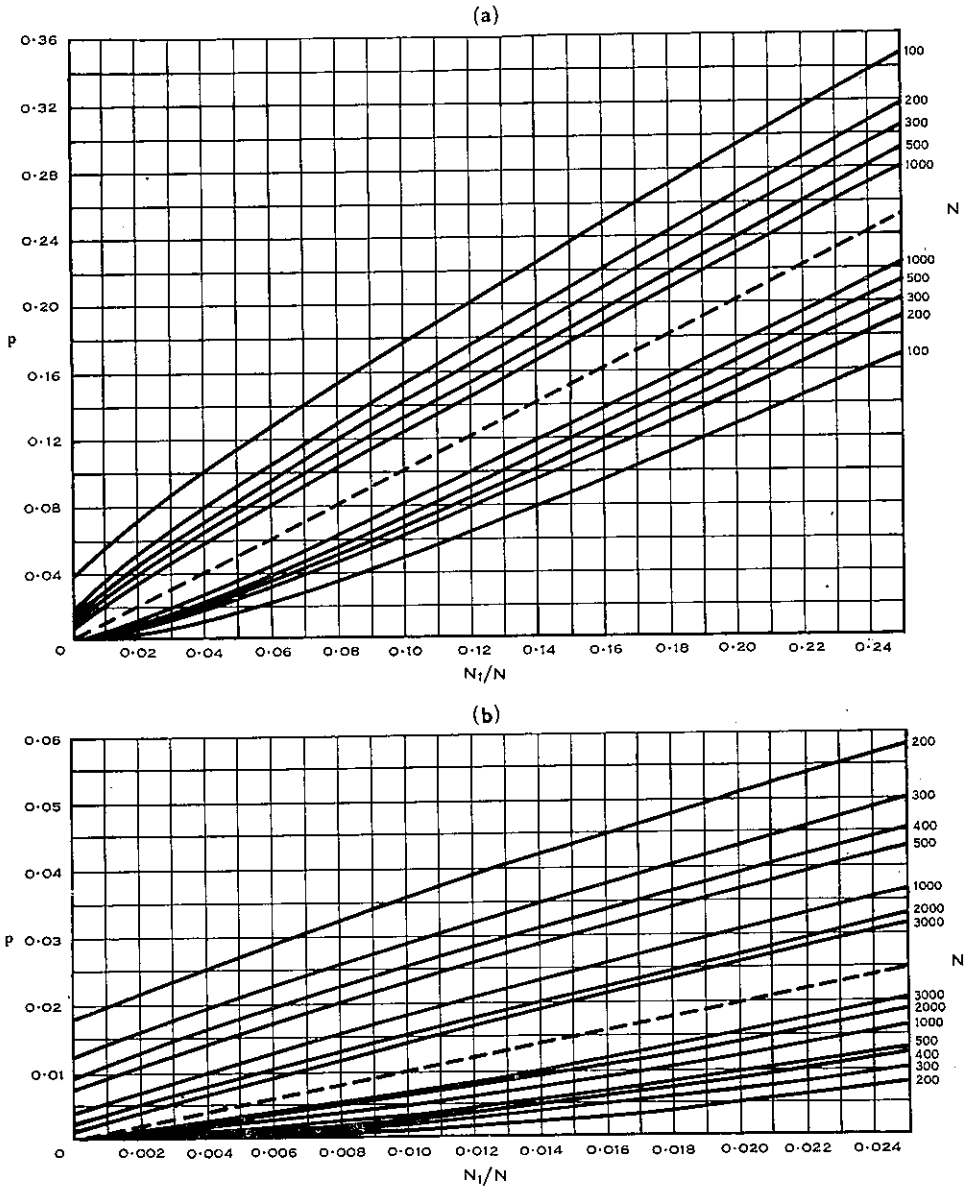

Fig. 1.—(*a*)95% confidence limits for p (= the true probability of success in a random trial) in binomial sampling in the range $0 \leq N_1/N \leq 0 \cdot 25$, and (*b*) in the range $0 \leq N_1/N \leq 0 \cdot 025$. N_1 = number of successes in a sample of N random trials. The confidence limits are given by the ordinates obtained for the abscissa N_1/N of the curves corresponding to the particular sample size N.

by C^2. It would be difficult to decide which of the various measures of heterogeneity would lead to the most powerful test, because it would depend on the type of deviations from H_0 most likely to occur. The use of C^2 has the advantage of simplicity and of being in common use.

II. Combination of Probabilities

When several independent experiments are conducted to test a single null-hypothesis H_0, the result is not necessarily statistically significant when one or more of the experiments yield low probabilities. For in a large number of independent experiments a certain percentage of them must be expected to yield low probabilities even when in fact H_0 is true. On the other hand, the weight of evidence of the combined total of all experimental results may well be sufficient to disprove the null-hypothesis even if none of the experiments is significant by itself.

To deal with such situations, Fisher (1932) considers the expression

$$\chi^2 = -2\Sigma \ln P_i, \tag{7}$$

where $\ln P_i$ is the natural logarithm of the probability P_i obtained in a significance test of the ith experiment. If there are k independent experiments ($i = 1, 2, \ldots, k$), the above expression is a χ^2 variate of $2k$ degrees of freedom. The null-hypothesis is rejected when χ^2 exceeds the significance value χ_0^2 obtainable from χ^2 tables.

Applying the theory to the situation in this paper, the null-hypothesis is that the group of litters produced by one heterozygous male is homogeneous, and that this is so for every one of the k heterozygous males used. Since there are k such males, we have k litter groups and k probabilities $\Pr\{C^2 \geqslant C_o^2 | H_0\}$, as defined above. To test the null-hypothesis, these k probabilities are combined by formula (7). If the result is significant, the null-hypothesis is rejected. The alternative hypothesis to be accepted is then that at least one of the k litter groups is not homogeneous.

In particular, for the $k = 37$ probabilities in Table 1, we would have to compute the sum

$$\chi^2 = -2[\ln 0{\cdot}0294 + \ln 0{\cdot}672 + \ldots + \ln 0{\cdot}0588].$$

However, since here the probabilities are only estimates of the true probabilities P_i (with a few exceptions), the value obtained is not exact. Errors are particularly serious when small P_i values are inaccurately estimated. Hence, the fact that the above sum yields a significant value does not allow the conclusion that the experimental results are themselves significant. To overcome this difficulty, the true P_i values were conservatively estimated by replacing the values in Table 1 by their upper 95% confidence limits. (For instance the fourth probability 0·0300 in Table 1 was replaced by 0·052, obtained from Figure 1 for $N = 400$). Since the value of χ^2 amended in this way continued to be significant, it was reasonably assumed that the experimental results were in fact statistically significant.

For the $k = 10$ probabilities listed in Table 2, the value of χ^2 turned out to be non-significant and continued to be so when the *lower* 95% confidence limits of the

probabilities were used instead. It was therefore reasonable to assume that the true probabilities, if available, would also have resulted in a non-significant result. Had the last result been significant, it could have been due to over-correction of the probabilities; a conclusive result, one way or the other, could then have been reached only by a more precise estimation of the probabilities.

The somewhat intuitive procedure adopted here should be satisfactory to overcome the difficulties of dealing with estimated probabilities instead of exact values. Generally, N will have to be large when P_i is small, which could be achieved by programming the computer to run until a fixed number N_1 of successes is obtained. A non-significant result means of course no more than that there is insufficient evidence of heterogeneity.

III. A General Coefficient of Heterogeneity

While C^2 is a suitable quantity for testing whether heterogeneity does or does not exist, it is, as it stands, quite unsuitable as a measure of the *degree* of heterogeneity. For C^2 and the probability that the value found for C^2 is exceeded both depend on the number and size of the litters. Take for instance the series of litters produced by male No. 32 (see tabulation p. 928). The value of C^2 turns out to be equal to $11 \cdot 7$, and the probability of exceeding this value by mere chance was found to be $0 \cdot 12$. Now consider the hypothetical litter series obtained by doubling all numbers in that tabulation. The new series should of course be regarded as equally heterogeneous. But now $C^2 = 2 \times 11 \cdot 7 = 23 \cdot 4$ and the probability of exceeding this value is less than $0 \cdot 01$, as may be seen at once by treating C^2 as a χ^2 variate of 6 degrees of freedom.

On the other hand, the quantity C^2/n is unaffected by the doubling of all numbers and has other properties that make it eminently suited to measure the degree of heterogeneity, viz:

(1) C^2/n always lies between 0 and 1. It is equal to 0 only when the abnormal animals are strictly distributed in proportion to the litter sizes and is equal to 1 only in the case of extreme heterogeneity when each litter contains either only normal or only abnormal animals. (In the very special case when there are only animals of one kind, C^2/n is put equal to 0 by definition, since it is obviously a case of perfect homogeneity.)

(2) C^2/n is independent of the number of litters in the sense that it remains unchanged when the r litters are replaced by kr litters with k litters similar to the original first litter, k litters similar to the original second litter, etc.

(3) C^2/n remains unchanged when all numbers in a table of litters are multiplied by the same constant.

Note 4

The quantity C^2/n represents an estimate of the so-called "mean-square contingency", introduced by K. Pearson for the general case of $r \times s$ contingency tables. Following Cramér (1946, p. 282), consider a discrete bivariate population in which every object has one of the characters $A_1, A_2, \ldots, A_r$ and one of the characters

$B_1, B_2, \ldots, B_s$. Let p_{ij} $(i = 1, \ldots, r;\ j = 1, \ldots, s)$ be the probability that a random object has the characters A_i, B_j. The (marginal) probability that a random object has the character A_i is then

$$p_{i.} = p_{i1} + \ldots + p_{is} \tag{8}$$

and similarly

$$p_{.j} = p_{1j} + \ldots + p_{rj} \tag{9}$$

for the character B_j. Complete independence (or homogeneity) between the A's and B's may then be expressed by the rs equations

$$p_{ij} - p_{i.}p_{.j} = 0. \tag{10}$$

When, on the other hand, the A's and B's are not completely independent, their degree of dependence (or heterogeneity) may be measured by K. Pearson's mean square contingency, defined by

$$\phi^2 = \sum_{i=1}^{r} \sum_{j=1}^{s} \frac{(p_{ij} - p_{i.}\, p_{.j})^2}{p_{i.}\, p_{.j}}. \tag{11}$$

Obviously, ϕ^2 is non-negative and is equal to zero only in the case of complete independence. Moreover, it can be shown that $\phi^2 \leqslant q - 1$, where q is the number of rows or columns in the contingency table, whichever is the smaller. Hence $\phi^2/(q-1)$ varies between 0 and 1 and may be taken as a standardized coefficient of heterogeneity.

To estimate the value of this coefficient from the cell frequencies n_{ij} when the true probabilities p_{ij} are unknown, it is natural (see Cramér, 1946, p. 443) to replace the probabilities p_{ij} by the relative frequencies n_{ij}/n. This leads to the estimate of the mean square contingency

$$\hat{\phi}^2 = \frac{1}{n} \sum_{i=1}^{r} \sum_{j=1}^{s} \frac{(n_{ij} - n_{i.}\, n_{.j}/n)^2}{n_{i.}\, n_{.j}/n},$$

where $n_{i.} = n_{i1} + \ldots + n_{is}$, $n_{.j} = n_{1j} + \ldots + n_{rj}$. We see that $\hat{\phi}^2$ is identical with C^2/n, where C^2 is the expression usually denoted by $\Sigma\,(o-e)^2/e$ and customarily used for a χ^2 test of contingency tables. To estimate the degree of heterogeneity in any contingency table we therefore propose the quantity $C^2/n(q-1)$, which in the special case of this paper reduces to C^2/n because $s = q = 2$.

In conclusion, we note that it can be shown (Patnaik 1949) that for large values of n (and subject to certain other conditions) the variate C^2 approaches a non-central χ^2 distribution with non-centrality parameter $n\phi^2$. Hence approximate confidence limits could be determined for the above coefficient of heterogeneity.