An Approximation to the Probablilities of the Multivariate Normal Distribution: Its Application to the Multifactorial Model of Qualitative Traits

Article · June 1979

CITATIONS
0

READS
79

1 author:

Robert Cloninger
Washington University in St. Louis
766 PUBLICATIONS **79,305** CITATIONS

SEE PROFILE

Reprinted from
BIOMETRICS COPYRIGHT © 1979
THE BIOMETRIC SOCIETY, Vol. 35, No. 2, June 1979

BIOMETRICS 35, 451–459
June, 1979

An Approximation to the Multivariate Normal Integral: Its Application to Multifactorial Qualitative Traits

J. RICE, T. REICH, C. R. CLONINGER

Department of Psychiatry, Washington University School of Medicine
and
The Jewish Hospital of St. Louis, 216 South Kingshighway, St. Louis, Missouri 63110, U.S.A.

R. WETTE

Division of Biostatistics, Washington University School of Medicine,
St. Louis, Missouri 63110, U.S.A.

Summary

Segregation analysis or pedigree analysis for quasicontinuous multifactorial traits requires probabilities under the multivariate normal distribution. Mendell and Elston (1974) describe a method for approximating recurrence risks which is used to provide an approximation to these probabilities. Three variants of this approximation are investigated and compared to results obtained using a method of Curnow (1972) in the equi-correlational case. These approximations have the advantage that no special covariance or threshold structure is required for their use. Although only genetic applications are given, the approximations are relevant to any problem which requires either the evaluation of the multivariate normal integral or estimation of the parameters of the multivariate normal distribution from polychotomized data.

The accuracy of each approximation is assessed over a range of truncation points for an s-variable normal, $s = 2, 3, 4, 6, 8, 10$, and specific recommendations are made as to which method can be used to achieve the best results. The sources of error in the approximations are discussed, and may serve as guidelines in implementing these procedures for specific applications.

1. The Multifactorial Model of Disease Transmission for Qualitative Traits

1.1. Description of the Model

The basic assumption of the multifactorial model for quasicontinuous traits is that a qualitative phenotype is determined by an underlying standard normal variate X, termed the liability. Manifestation of the forms of the trait under consideration is determined by one's liability score lying above or below one or more threshold values (Carter 1965, Falconer 1965, 1967, Reich, James and Morris 1972, and Curnow and Smith 1975). Multiple thresholds $T_1 < T_2 < \cdots < T_n$ may be used with a single liability distribution to model severity (Reich *et al.* 1972), with liability scores below T_1 corresponding to unaffected individuals, scores above T_n to severely affected individuals, and scores between T_i and T_{i+1}, $i = 1, \ldots, n-1$, to individuals affected with forms of intermediate severity.

Alternatively, multiple liability distributions may be postulated which reflect different disease processes yielding similar phenotypes. Kidd, Reich and Kessler (1973) have allowed for different liability distributions for males and females, and Reich, Rice, Cloninger, Wette and James (1979) have allowed for correlated liability distributions to resolve phenotypic heterogeneity. Rice, Cloninger and Reich (1978) and Cloninger, Rice and Reich (1979) have

proposed a general multifactorial model which allows for cultural as well as genetic transmission. In their model, different liability distributions, and the correlations between them, may depend upon non-genetic factors such as one's family structure of rearing.

Familial transmission of a multifactorial trait is determined by the correlations in liability between family members. Depending on the model chosen, these correlations may be functions simply of the degree of genetic relationship, or, as indicated above, of sex, severity, or designated environmental events. When individuals are measured for more than one phenotype, then the correlations in liability within an individual are also parameters of the model.

We consider here a general multifactorial model defined by standardized normal liability distributions $X_1, \ldots, X_N$, and by the correlations between them. Each X_i has K_i threshold values T_{ik}, $T_{i1} \leq T_{i2} \leq \cdots \leq T_{iK_i}$, with the phenotype of an individual of type i indicating whether his liability value is above or below the threshold T_{ik}. An individual whose value is above T_{ik} is said to be "affected" with form k of the trait.

1.2. Estimation of the Parameters of the Model

The threshold values T_{ik} may be estimated directly from population prevalence estimates. Let K_{Pik} denote the proportion of individuals of type i in the general population who are above the threshold T_{ik} and let $\Phi_c(y)$ denote the complement of the standard normal distribution function. Then

$$K_{Pik} = \text{Prob}(X_i \geq T_{ik}) = \Phi_c(T_{ik}) \qquad \text{and} \qquad T_{ik} = \Phi_c^{-1}(K_{Pik}). \tag{1}$$

Equation (1) may be used to yield the maximum likelihood estimate (MLE) of T_{ik} by replacing K_{Pik} by the observed proportion of individuals of type i who are affected with form k.

If pairs of individuals are sampled at random from (X_i, X_j), then the MLE of ρ_{ij} is given by the tetrachoric correlation coefficient (Kendall and Stuart 1973), determined by the 2×2 table of observed pairs of phenotypes. However, in human genetics the pairs (X_i, X_j) are often sampled using the proband method of selecting pairs through affected individuals. With either sampling scheme the MLE of ρ_{ij} is that correlation coefficient in the bivariate normal distribution which yields the observed proportions. This may be done by either integrating the bivariate normal density function or using the approximation described below.

In general, much more information about the transmission of a trait is available by studying families rather than independent pairs of individuals. Accordingly, family study data, usually ascertained through affected individuals, may be analyzed by segregation analysis (Elston and Yelverton 1975), i.e., by analyzing the multinominal distribution of ascertained families which is determined by the possible phenotypic configurations for families of a given sibship size.

The use of segregation analysis requires the computation of the probability of observing a family with specified phenotypes, where the joint distribution in liability of a family of size s is assumed to be an s-variate normal with variance-covariance matrix ϱ determined by the parameters of the multifactorial model. For a given family $F = (Y_1, \ldots, Y_s)$, with each Y_i a member of the set $\{X_1, \ldots, X_N\}$, and for $T = (T_1, \ldots, T_s)$, with each T_i a member of the set $\{T_{ik}\}$, we must evaluate integrals of the form

$$P_F = \int_{I_1} \cdots \int_{I_s} \phi(y_1, \ldots, y_s; \varrho) dy_1 \ldots dy_s, \tag{2}$$

where ϕ denotes the appropriate s-variate normal density function and each I_i is of the form $(-\infty, T_i)$ or (T_i, ∞). Unfortunately, except in a few special cases, a practical algorithm to evaluate integrals of the form of equation (2) is not known. We describe here an approximation which compares favorably to the values given by equation (2) in those cases where the integral can be evaluated.

2. *The Pearson Approximation to the Bivariate Normal*

Let (Y_1, Y_2) have a bivariate normal distribution with each Y_i having mean 0 and variance 1, and with the correlation between them equal to ρ_{12}, and let T_1 and T_2 be the threshold values for Y_1 and Y_2, respectively. Denoting the density function of the standardized normal by $\phi(y)$, and $K_{P1} = \text{Prob}(Y_1 \geq T_1)$ by P_1, we recall the following standard results:

$$a_1 = E[Y_1 | Y_1 \geq T_1] = \phi(T_1)/P_1, \qquad \sigma_1^2 = \text{Var}[Y_1 | Y_1 \geq T_1]$$
$$= 1 - a_1(a_1 - T_1), \qquad a_{2|1} = E[Y_2 | Y_1 \geq T_1] = \rho_{12} a_1,$$

and

$$\sigma_{2|1}^2 = \text{Var}[Y_2 | Y_1 \geq T_1] = 1 - \rho_{12}^2\, a_1(a_1 - T_1).$$

Thus, the mean and variance of relatives of "affected" individuals are $\rho_{12}a_1$ and $1 - \rho_{12}^2\, a_1(a_1 - T_1)$, respectively. The distribution of such relatives is not normal unless $\rho_{12} = 0$, but may be approximately so under certain conditions, so that we can approximate $P_2 = \text{Prob}(Y_1 \geq T_1, Y_2 \geq T_2)$ by $P_2 = P_1\, \Phi_c(Z_2)$, where the standardized threshold Z_2 is given by $Z_2 = (T_2 - a_{2|1})/\sigma_{2|1}$. This approximation is described by Pearson (1903), Reich *et al.* (1972), Mendell and Elston (1974), and Smith and Mendell (1974), and has been found to be quite good, especially for small ρ_{12}.

3. *The Mendell-Elston Approximation to the Multivariate Normal*

Let

$$(Y_1, \ldots, Y_s) \sim \mathbf{N}\left(\begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & & \rho_{ij} \\ \cdot & & \\ & & \cdot \\ \rho_{ij} & & 1 \end{bmatrix}\right)$$

have a multivariate normal distribution and let T_i, $i = 1, \ldots, s$, be a threshold point of Y_i. The bivariate approximation described above may be applied to each pair $(Y_i, Y_{i+1} | Y_1 \geq T_1, \ldots, Y_{i-1} \geq T_{i-1})$ by assuming their joint distribution is normal at each stage. Mendell and Elston (1974) use a result due to Aitken (1934), exact when a selected population is normally distributed, to approximate recurrence risks.

Standardized thresholds are defined recursively by

$$Z_{i|0,\ldots,j} = (Z_{i|0,\ldots,j-1} - a_{j|0,\ldots,j-1} r_{ji|0,\ldots,j-1})/\sigma_{j|0,\ldots,j-1},$$
$$\text{for } j = 1, \ldots, i-1, \text{ with } Z_{i|0} = T_i,\ r_{ji|0} = \rho_{ji},$$

and where

$$a_{j|0,\ldots,j-1} = \phi(Z_{j\,0,\ldots,j-1})/\Phi_c(Z_{j|0,\ldots,j-1}),$$
$$\sigma_{k|0}^2{}_{,\ldots,l} = 1 - r_{lk|0}^2{}_{,\ldots,l-1} a_{l|0,\ldots,l-1}(a_{l|0,\ldots,l-1} - Z_{l|0,\ldots,l-1}),$$

and

$$r_{mn\,|0,\ldots,j} = \frac{r_{mn\,|0,\ldots,j-1} - r_{jm\,|0,\ldots,j-1} r_{jn\,|0,\ldots,j-1} a_{j\,|0,\ldots,j-1}(a_{j\,|0,\ldots,j-1} - Z_{j\,|0,\ldots,j-1})}{\sigma_{m\,|0,\ldots,j}\sigma_{n\,|0,\ldots,j}}. \quad (3)$$

The recurrence risk to the *jth* individual conditional on the first $(j-1)$ individuals being affected is given approximately by

$$\text{RISK}(j|0, 1, 2, \ldots, j-1) = \Phi_c\,[Z_{j|0,1,\ldots,j-1}].$$

This approximation also follows by repeated applications of the bivariate approximation described above, where the numerator in equation (3) is the conditional covariance between Y_m and Y_n.

The joint probability $P_F = \text{Prob}(Y_1 \geq T_1, \ldots, Y_s \geq T_s)$ depends upon the ordering of $Y_1, \ldots, Y_s$, and is given by

$$P_F = \text{RISK}(1|0)\ \text{RISK}(2|0, 1) \ldots \text{RISK}(s|0, 1, 2, \ldots, s-1).$$

4. Variants on the Mendell-Elston Approximation and Other Approximations

We investigated three slightly different forms of the Mendell-Elston approximation:

1. *Derivative Mendell-Elston* (MED) This method applies the approximation only when all $T_i \geq 0$ as was suggested originally by Mendell and Elston (1974). Probabilities with one or more $T_i < 0$ are derived algebraically in terms of conditional probabilities which involve terms with all $T_i > 0$, and then estimated from these expressions. For most multifactorial models this corresponds to using the approximation only for "affected" individuals and deriving the probabilities for families with unaffected individuals.

2. *Mendell-Elston* (ME) In this method the approximation is applied directly even with some or all $T_i < 0$. If an individual of type i is unaffected with form k so that his liability is below T_i' on X_i', then X_i is set equal to $-X_i'$ and T_i to $-T_i'$, so that an unaffected individual is one whose liability score on X_i is above T_i. This method was suggested by Smith and Mendell (1974).

3. *Mendell-Elston with exact bivariate probabilities* (MEB) Error in the MED or ME approximation occurs from the assumption that pairs of selected individuals are distributed as a bivariate normal as well as from the bivariate approximation itself. In the MEB method an algorithm is used at each stage which would give exact probabilities if the assumption of bivariate normality were met. Thus, error results only from the joint normality assumption itself.

An evaluation of the accuracy of each method is described in the section that follows.

Curnow and Dunnett (1962) and Curnow (1972) have noted that when all $\rho_{ij} = \rho$, and all $T_i = T$, then the multiple integral of equation (2) reduces to the single integral

$$P_F = \int_{-\infty}^{\infty} \phi(t)\ \Phi_c^s\left[\frac{T + t\sqrt{\rho}}{\sqrt{(1-\rho)}}\right] dt.$$

This integral was evaluated using an adaptive Romberg extrapolation method (de Boor 1971) to give the Curnow (C) approximation.

Ruben [1954] has tabulated the value of P_F in the equal correlations case for $\rho = 1/j, j = 2, 3, \ldots, 12$ and $s = 1, 2, \ldots, 51 - j$ for the special case where all $T_i = 0$. His results are derived by computing the contents of regular hyperspherical simplices and relating these values to the above multiple integrals. Exact expressions for truncation at the mean of a

bivariate or trivariate normal are available for any covariance matrix, but for $s > 3$ an exact solution is not available, and power series solutions are slow to converge (Kendall and Stuart 1977). The results of Curnow and Dunnett and of Ruben enable us to assess the accuracy of the above approximations.

5. *Evaluation of Approximations to the Probabilities of the Multivariate Normal*

Although the Mendell-Elston approximation described above is valid for an arbitrary covariance matrix and an arbitrary number of thresholds, comparisons can only be made in the special cases outlined above. We feel, however, that such comparisons are valid also for the general case since the approximations are based only on the assumption of joint normality after selection and not on any particular covariance or threshold structure.

Three general observations may be made from our computations: Firstly, the approximation is always better for small ρ. This is to be expected since the liability of relatives of selected individuals is normally distributed when $\rho = 0$, and become less so as ρ increases. We

TABLE 1

Values of the Curnow and Mendell-Elston Approximations

(For the indicated ρ, K_p and s, the value of C is given first, then the value of ME)

		Dimension (s)					
ρ	K_p	2	3	4	6	8	10
0.7	0.1	0.047	0.030	0.022	0.014	0.011	0.009
		0.047	0.031	0.023	0.015	0.011	0.009
	0.2	0.113	0.080	0.063	0.045	0.035	0.030
		0.114	0.082	0.065	0.047	0.037	0.031
	0.3	0.191	0.144	0.118	0.089	0.073	0.063
		0.192	0.147	0.122	0.094	0.078	0.067
	0.4	0.278	0.221	0.188	0.148	0.126	0.110
		0.279	0.225	0.193	0.155	0.133	0.118
	0.5	0.373	0.310	0.271	0.223	0.194	0.174
		0.375	0.314	0.277	0.232	0.205	0.186
0.5	0.1	0.032	0.016	0.009	0.004	0.002	0.002
		0.032	0.016	0.009	0.004	0.002	0.001
	0.2	0.087	0.050	0.033	0.018	0.012	0.008
		0.087	0.050	0.033	0.018	0.012	0.008
	0.3	0.157	0.101	0.072	0.044	0.031	0.023
		0.157	0.102	0.073	0.045	0.031	0.023
	0.4	0.239	0.167	0.127	0.085	0.063	0.049
		0.240	0.169	0.129	0.086	0.064	0.050
	0.5	0.333	0.250	0.200	0.143	0.111	0.091
		0.334	0.252	0.202	0.146	0.114	0.093

have reported simulations for large ρ (0.7 and 0.5), and these together with other simulations have shown that relative error declines as ρ becomes smaller. In a model with two different correlations, we would expect the relative error to be intermediate to those found in the two equal correlation approximations. Secondly, for $T > 0$, the approximation is better for large T (i.e. for traits with low population prevalences). This in part may be explained by examination of equation (3) for the conditional r_{ij}'s. The decrement in r_{ij} increases monotonically as T increases, so that the approximation at each bivariate stage is applied with a smaller correlation when T is large. Finally, for $T < 0$, the approximation gets better as T decreases. The explanation for this is that the distribution of a selected relative is approaching normality as T decreases; thus, the ME approximation can be expected to be more accurate.

Table 1 gives the values of the C and ME approximations, respectively, for $\rho = 0.7$ and 0.5, for K_p (the population prevalence of the trait) between 0.1 and 0.5, and for $s = 2, 3, 4, 6, 8$ and 10. Unlike the estimation of recurrence risks, the absolute error is important for segregation analysis since families are sampled from a multinomial distribution. The ME approximation is quite good, especially for $\rho = 0.5$ and for $K_p \leq 0.3$. The MEB approximation (not displayed) was found to be better than the ME approximation, with an absolute error which is smaller, in general, by about 50%. In most applications, ρ would be smaller than 0.7, and K_p would be small, so that either the ME or MEB approximation would be appropriate. The ME approximation required considerably less computer time, although both were quite fast. The C, ME and MEB approximations were also compared for K_p of

TABLE 2
Values of the C, MED and ME Approximations

		Dimension (s)					
ρ	K_p	2	3	4	6	8	10
0.7	0.60	0.478	0.412	0.369	0.314	0.280	0.256
		0.479	0.413	0.369	0.312	0.275	0.249
		0.478	0.415	0.375	0.326	0.294	0.272
	0.80	0.713	0.658	0.620	0.566	0.528	0.501
		0.714	0.660	0.621	0.567	0.529	0.500
		0.711	0.657	0.621	0.572	0.540	0.516
	0.99	0.983	0.977	0.972	0.963	0.957	0.951
		0.983	0.977	0.972	0.963	0.957	0.951
		0.982	0.974	0.968	0.957	0.948	0.940
0.5	0.60	0.439	0.350	0.293	0.223	0.182	0.154
		0.440	0.351	0.294	0.223	0.181	0.152
		0.440	0.352	0.296	0.228	0.187	0.159
	0.80	0.687	0.611	0.556	0.478	0.424	0.385
		0.687	0.612	0.556	0.478	0.425	0.385
		0.687	0.612	0.558	0.483	0.432	0.394
	0.99	0.981	0.973	0.966	0.954	0.943	0.933
		0.981	0.974	0.966	0.954	0.943	0.933
		0.981	0.973	0.965	0.951	0.939	0.929

TABLE 3
Values of the R and ME Approximations for $K_p = 50\%$
(The number abc(−d) represents 0.abc × 10^−d)

	Dimension (s)						
ρ	15	20	25	30	35	40	45
1/2	625(−1)	476(−1)	385(−1)	323(−1)	278(−1)	244(−1)	217(−1)
	638(−1)	480(−1)	380(−1)	312(−1)	263(−1)	225(−1)	196(−1)
1/3	234(−1)	147(−1)	101(−1)	740(−2)	567(−1)	449(−2)	365(−2)
	224(−1)	140(−1)	928(−2)	654(−2)	481(−2)	365(−2)	284(−2)
1/4	115(−1)	612(−2)	366(−2)	238(−2)	164(−2)	118(−2)	882(−3)
	112(−1)	575(−2)	331(−2)	206(−2)	135(−2)	927(−3)	658(−3)

0.1% and 1% with relative errors for the ME and MEB approximations ranging from less than 1% for $s \leq 4$ to about 20% for $s = 10$.

The values for $\rho = 0.7$ and 0.5, and for $K_p \geq 0.6$ are given in Table 2 for the C, MED, and ME approximations, respectively. For intermediate K_p, the MED approximation gives considerably more accurate results than the ME approximation. For $\rho = 0.5$ or $\rho = 0.7$ and a large K_p, both approximations give good results. If desired, greater accuracy may be obtained by using the MEB approximation for $(1 - K_p)$ and applying the derivative method of the MED approximation.

The analysis of large pedigrees for a multifactorial trait would require probabilities for large s. Table 3 displays the exact values given by Ruben (1954), together with values given by the ME approximation for $T = 0$, $\rho = 1/2, 1/3, 1/4$, and $s = 15, \ldots, 45$. Table 1 would suggest that $T = 0$ is in the range where the greatest error would be expected for $0.1 \leq K_p \leq 0.5$, and Table 2 would suggest that the MED approximation would be significantly more accurate in the range $0.5 \leq K_p \leq 0.99$. Thus, the values in Table 3 suggest that the ME (or MED) approximation will give acceptable accuracy even for very large s. This, in part, may be explained by the small conditional correlations which are used in later steps of the approximation, so that error does not accumulate in an additive fashion. In practice, a large pedigree contains many ρ_{ij}'s which are themselves very small, so that by ordering the pedigree to begin with affected individuals (if their T_i's are large), the application of this method in extended pedigree analysis appears to be entirely feasible.

6. *Conclusions*

The ME approximation for the computation of multivariate probabilities is consistently accurate, especially in the parameter ranges which are relevant for the analysis of most multifactorial traits. Over certain parameter ranges the MED approximation may be used to provide greater accuracy and could be used in conjunction with the ME approximation. The MEB approximation indicates that the error in the ME approximation results from the assumption of the normality in selected relatives. The approximation is especially good for small K_p because of the rapid decrease in the conditional correlations, and for large K_p because the assumption of the normality of relatives of selected individuals results in less error.

The ME approximation is easy to program and requires little computer time. This is in

contrast to approaches which require either numerical integration, power series or Chebyshev-Hermite polynomial approximation which would converge very slowly for large s. The MED approximation requires many intermediate computations; therefore, it is considerably slower than the ME approximation.

Although the comparisons made are for special covariance and threshold structures, the basic observations should be valid in general. Indeed, the types of data to be expected in the application of this method, such as in either segregation or extended pedigree analysis, would indicate that the approximation will be acceptable. The method outlined above may be used as a basic module necessary for the application of any particular multifactorial model. A specific model which determines the T_i and ρ_{ij} will be needed for input to the module and the computed probabilities in turn will determine the likelihood needed for a maximum likelihood search.

Acknowledgments

This work was supported in part by USPHS grants AA-03539, MH-07081, MH-31302, and MH-25430, and Research Scientist Development Award MH-00048 (CRC). The authors express their gratitude to Dr. Kenneth Lange and to the referees for their useful comments.

Résumé

On a besoin des probabilités de la distribution gaussienne multivariate pour l'analyse de ségrégation ou l'analyse de pédigrée de caractères multifactoriels quasi-continus. Mendell et Elston (1974) décrivent une méthode pour approcher les risques de récurrence, et l'utilisent pour fournir une approximation et on les compare aux résultats obtenus par la méthode de Curnow (1972) dans le cas équicorrélé. Ces approximations ont l'avantage de ne requérir ni structure de covariance si seuil spéciaux. Bien qu'on ne donne que des applications génétiques, les approximations sont intéressantes pour tout problemme ou l'on a besoin soit de l'evaluation de l'intégrame gaussienne multivariable, soit de l'estimation des paramètres de la distribution gaussienne multivariable.

On vérifie la précision de l'approximation sur une gamme de points de troncation pour une s-variate gaussienne avec s = 2, 3, 4, 6, 8, 10, et on recommande les choix de méthode afin d'obtenir les meilleurs résultats. On discute les sources d'erreur dans les approximations, qui peuvent servir de guides pour utiliser ces procédures dans des applications spécifiques.

References

Aitken, A. C. (1934). Notes on selection from a multivariate normal population. *Proceedings of the Edinburgh Mathematical Society 4,* 106–110.

Carter, C. O. (1965). The inheritance of common congenital malformations. In *Progress in Medical Genetics IV,* 59–84.

Cloninger, C. R., Rice, J. and Reich, T. (1979). Multifactorial inheritance with cultural transmission and assortative mating. II. A general model of combined polygenic and cultural inheritance. *American Journal of Human Genetics,* in press.

Curnow, R. N. (1972). The multifactorial model for the inheritance of liability to disease and its implications for relatives at risk. *Biometrics 28,* 931–946.

Curnow, R. N. and Dunnett, C. W. (1962). The numerical evaluation of certain multivariate normal integrals. *Annals of Mathematical Statistics 33,* 571–579.

Curnow, R. N. and Smith, C. (1975). Multifactorial models for familial diseases in man. *Journal of the Royal Statistical Society, Series A 138,* 131–169.

de Boor, Carl (1971). Cadre: An algorithm for numerical quadrature. In *Mathematical Software.* John Rice, ed., Academic Press, New York.

Elston, R. C. and Yelverton, R. C. (1975). General models for segregation analysis. *American Journal of Human Genetics 27,* 31–45.

Falconer, D. S. (1965). The inheritance of liability to certain diseases estimated from the incidence among relatives. *Annlas of Human Genetics 29,* 51–76.

Falconer, D. S. (1967). The inheritance of liability to disease with variable age of onset with particular reference to diabetes mellitus. *Annals of Human Genetics 31,* 1–20.

Kendall, M. G. and Stuart, A. (1973). *The Advanced Theory of Statistics, Vol. 2, Inference and Relationship,* 3rd ed., Hafner Publishing, New York.

Kendall, M. G. and Stuart, A. (1977). *The Advanced Theory of Statistics, Vol 1, Distribution Theory,* 4th ed., Macmillan, New York.

Kidd, K. K., Reich, T. and Kessler, S. (1973). Sex effect and the single gene. The relevance of sex effect in discriminating between genetic hypotheses. *Genetics 74,* 137.

Mendell, N. R. and Elston, R. C. (1974). Multifactorial qualitative traits: genetic analysis and prediction of recurrence risks. *Biometrics 30,* 41–57.

Pearson, K. (1903). On the influence of natural selection on the variability and correlation of organs. *Philosophical Transactions of the Royal Society of London A 200,* 1–66.

Reich, T., James, J. W. and Morris, C. A. (1972). The use of multiple thresholds in determining the mode of transmission of semi-continuous traits *Annals of Human Genetics 36,* 163–184.

Reich, T., Rice, J., Cloninger, C. R., Wette, R. and James, J. (1979). The use of multiple thresholds and segregation analysis in analyzing the phenotypic heterogeneity of multifactorial traits. *Annals of Human Genetics 42,* 371–390.

Rice, J., Cloninger, C. R. and Reich, T. (1978). Multifactorial inheritance with cultural transmission and assortative mating. I. Description and basic properties of the unitary models. *American Journal of Human Genetics 30,* 618–643.

Ruben, H. (1954). On the moments of order statistics in samples from normal populations. *Biometrika 41,* 200–227.

Smith, C. and Mendell, N. R. (1974). Recurrence risks from family history and metric traits. *Annals of Human Genetics 37,* 275–286.

Received June 1978; *Revised November* 1978