DOCUMENT RESUME

ED 078 063 TM 002 882

AUTHOR	Boldt, Robert F.
TITLE	The Inverted Student Density and Test Scores.
INSTITUTION	Educational Testing Service, Princeton, N.J.
REPORT NO	ETS-RB-73-33
PUB DATE	Apr 73
NOTE	19p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Analysis of Variance; *Scores; *Statistical Analysis; Technical Reports; *Test Results
IDENTIFIERS	*Student Density

ABSTRACT

The inverted density is one whose contour lines are spheroidal as in the normal distribution, but whose moments differ from those of the normal in that its conditional arrays are not homoscedastic, being quadratic functions of the values of the linear regression functions. It is also platykurtic, its measure of kurtosis ranging from that of the normal to that of the uniform depending on the value of a parameter: as that parameter increases the inverted student distribution approaches normality. Measures of kurtosis are given for distributions of scores on a number of cognitive tests, and they are almost all seen to be platykurtic. Data are presented showing that a quadratic terms contributes substantially to the regression of conditional variances on test scores in a bivariate distribution. These data suggest that the inverted student distribution may provide a better description of distributions of test scores than does the normal. (Author)

RESEARCH BULLETIN

RB-73-33

THE INVERTED STUDENT DENSITY AND TEST SCORES

Robert F. Boldt

This Bulletin is a draft for interoffice circulation. Corrections and suggestions for revision are solicited. The Bulletin should not be cited as a reference without the specific permission of the author. It is automatically superseded upon formal publication of the material.

Educational Testing Service
Princeton, New Jersey
April 1973

THE INVERTED STUDENT DENSITY AND TEST SCORES

Robert F. Boldt
Educational Testing Service

Abstract

The inverted student density is one whose contour lines are spheroidal as in the normal distribution, but whose moments differ from those of the normal in that its conditional arrays are not homoscedastic, being quadratic functions of the values of the linear regression functions. It is also platykurtic, its measure of kurtosis ranging from that of the normal to that of the uniform depending on the value of a parameter: as that parameter increases the inverted student distribution approaches normality. Measures of kurtosis are given for distributions of scores on a number of cognitive tests, and they are almost all seen to be platykurtic. Data are presented showing that a quadratic term contributes substantially to the regression of conditional variances on test scores in a bivariate distribution. These data suggest that the inverted student distribution may provide a better description of distributions of test scores than does the normal.

THE INVERTED STUDENT DENSITY AND TEST SCORES

Robert F. Boldt
Educational Testing Service

Because of difficulties in handling certain problems in multivariate statistics, a polynomial distribution similar to the normal seemed to the author to be a promising alternative distribution (Boldt, 1962). However, the use of the distribution in analysis proved tedious, and the notion was dropped. More recently there has been renewed interest in the distribution, known as the Inverted Student Distribution, because of findings indicating that variances of conditional test score distributions are, in some cases at least, not homoscedastic. Rather, the conditional variances seem to decrease at extreme values of the independent variables (Boldt, 1968, 1972). Also, the platykurtic character of test score distributions has come to the author's attention. Both of these characteristics are in contrast to the normal distribution and are, indeed, properties of the Inverted Student Distribution (Press, 1972; Raiffa & Schlaifer, 1961) as will be shown below. Thus the interest in the distribution, which was originally based on its possible value as an approximation to the normal, may arise specifically because of the characteristics it has which are in distinction to the normal but which characterize test scores.

In the material that follows the distribution is introduced, its marginal distributions and the constant which makes the distribution a density function are developed. Following this, the relationship of certain parameters to the population means and to the conditional means is given. Then the relationship between certain other parameters and the population variances and covariances is given, followed by a development of the conditional variances which exhibits the nonhomoscedastic character of the conditional variances. Subsequently it is shown that the distribution tends to normality under certain conditions, followed by a section concerning higher moments of the univariate marginal distribution. Finally, data are presented which show that measures of kurtosis of test data and the conditional variances of some tests have properties in agreement with those of the distribution.

Let

X be a column vector of P random variables;

N, K, C be constants of the distribution;

μ be a column vector of P parameters;

A be a $P \times P$ symmetric positive definite matrix of parameters.

Then the density is

$$C[K - (X - \mu)'A(X - \mu)]^N \tag{1}$$

if the quantity in brackets is nonnegative, zero otherwise. Transform X to Y as follows:

$$X = \sqrt{K}\,TY + \mu \tag{2}$$

where

$$T'AT = I \ .$$

Then the joint distribution of the Y 's is

$$C\,\frac{K^{N + \frac{P}{2}}}{|A|^{\frac{1}{2}}}\,(1 - Y'Y)^N \ . \tag{3}$$

Note in (3) that the Y's are symmetrically distributed and hence have mean zero. Using (2) it follows that μ is the mean of X. To find C, define $V_i = Y_i^2$, when Y_i is the ith element of Y. Then the V's have the Dirichlet distribution.

$$\frac{CK^{N+\frac{P}{2}}}{|A|^{\frac{1}{2}}} \left(\prod_i^P V_i\right)^{-\frac{1}{2}} \left(1 - \sum_j V_i\right)^N . \tag{4}$$

Equating the constant of (4) to the constant of the Dirichlet distribution (Press, 1972, p. 131) yields

$$C = \frac{\Gamma(N + \frac{P+2}{2})}{\Gamma(N+1)\,\pi^{\frac{P}{2}}} \; \frac{|A|^{\frac{1}{2}}}{K^{N+\frac{P}{2}}} . \tag{5}$$

The interpretation of A and K can be developed as follows. The expected value of a V_i is $\frac{1}{2N + P + 2}$ (Press, 1972, p. 134) and is the expected value of a corresponding Y_i^2. It can be seen on examination of (3) that the covariance of Y's is zero so that the variance-covariance matrix of Y's is $(2N + P + 2)^{-1}I$. Then

$$E(X - \mu)(X - \mu)' = \sqrt{K}\, T[E(YY')]T' \sqrt{K}$$

$$= \frac{KTT'}{2N + P + 2} ,$$

where E is the expectation operator. But since $T'AT = I$, it follows that $(TT')^{-1} = A$ or

$$E(X - \mu)(X - \mu)' = \frac{K}{2N + P + 2} A^{-1} . \tag{6}$$

One would not want the variance-covariance matrix of X to depend on the choice of N and P, and hence we choose $K = 2N + P + 2$ and note that A is the inverse of the variance-covariance matrix of X's as in the normal distribution.

Marginal Distributions

Suppose the variables in X and μ are partitioned into subvectors that

$$X = \| X_1 | X_2 \| , \quad \mu = \| \mu_1 | \mu_2 \| ,$$

and A is partitioned as follows

$$A = \left\| \begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{12}' & A_{22} \end{array} \right\| ,$$

where the sizes of the submatrices of A are consistent with the partitioning of X. Then, using the transformation

$$X_1 - \mu_1 = U - A_{11}^{-1}A_{12}(X_2 - \mu_2) , \tag{7}$$

the joint distribution of U and X_2 becomes

$$C[K - U'A_{11}U - (X_2 - \mu_2)'(A_{22} - A_{12}'A_{11}^{-1}A_{12})(X_2 - \mu_2)]^N . \tag{8}$$

Suppose A^{-1} is partitioned in a fashion corresponding to the partitioning of A given above (this is by (6) a partitioning of the variance-covariance matrix of X), yielding say B_{11} , B_{12} , B_{22} with the subscripts corresponding to sections of the partitioning of A , then it can be shown that

$$A_{22} - A_{12}'A_{11}^{-1}A_{12} = B_{22}^{-1} .$$

Transforming U to $\tilde{Y}$ as follows:

$$U = \sqrt{K - (X_2 - \mu_2)B_{22}^{-1}(X_2 - \mu_2)}\ \tilde{T}\tilde{Y}$$

where $\tilde{T}'A_{11}\tilde{T} = I$ and using (5) yields

$$\frac{\Gamma(N + \frac{P+2}{2})}{\Gamma(N+1)\ \pi^{\frac{P}{2}}} \frac{|A|^{\frac{1}{2}}}{K^{N+\frac{P}{2}}} \frac{[K - (X_2 - \mu_2)'B_{22}^{-1}(X_2 - \mu_2)]^{N+\frac{Q}{2}}}{|A_{11}|^{\frac{1}{2}}} [1 - \tilde{Y}'\tilde{Y}]^N , \quad (9)$$

when there are Q variables $\tilde{Y}$, (e.g., A_{11} and B_{11} are $Q \times Q$; A_{12} and B_{12} are $Q \times (P - Q)$). Then using the transformation

$$V_i = Y_i^2$$

and the Dirichlet constant yields

$$\frac{\Gamma(N + \frac{P+2}{2})\ |B_{22}|^{-1}}{\Gamma(N + \frac{Q+2}{2})\ \pi^{\frac{P-Q}{2}}\ K^{N+\frac{P}{2}}} [K - (X_2 - \mu_2)'B_{22}^{-1}(X_2 - \mu_2)]^{N+\frac{Q}{2}} \quad (10)$$

as the marginal distribution of the remaining variables. Expression (10) is written with B_{22} using the relation

$$|A| = |A_{11}||A_{22} - A_{12}'A_{11}^{-1}A_{12}| \qquad \text{(Graybill, 1961, pp. 8-9).}$$

From (10) it can be seen that the marginal distributions are of the same form as (1), i.e., a quadratic form subtracted from a constant and raised to power. Thus any distribution for which the form (1) is assumed might in some contexts be regarded as a marginal distribution.

Conditional Means and Variances

From the expression (8) it can be seen that U is symmetrically distributed for any value of X_2 and hence has mean zero for that value of X_2. Therefore, from (7), the mean of X_1 given X_2 is

$$\mu_1 - A_{11}^{-1}A_{12}(X_2 - \mu_2) \quad .$$

It can be shown that $-A_{11}^{-1}A_{12} = B_{12}B_{22}^{-1}$ and hence, $\mu_1 + B_{12}B_{22}^{-1}(X_2 - \mu_2)$ is the mean of X_1 given X_2. This is the familiar regression formula.

Conditional variances are not homoscedastic in this distribution. That this is true can be seen as follows. Note that if a multivariate distribution function is CF, where C is the density constant and U is one of its arguments with mean zero, then the conditional variance of U is

$$\frac{\int_{R_U} CF\, U^2}{\int_{R_U} CF} = \frac{\int_{R_U} F\, U^2}{\int_{R_U} F} \quad . \tag{11}$$

In the present distribution, F is of the form in brackets in (8) and we are taking U and A_{11} as scalars, i.e., Q equals unity. Then if we use transformation

$$U^2A_{11} = V(K - (X_2 - \mu_2)B_{22}^{-1}(X_2 - \mu_2)) \quad , \tag{12}$$

the quantity in parentheses in (12) above by which V is multiplied, factors out of the brackets in (8) and, including the differential, is raised to the $N + \frac{3}{2}$ power in the numerator of (11), and to the $N + \frac{1}{2}$

power in the denominator. The conditional variance is of the form

$$W[K - (X_2 - \mu_2)'B_{22}^{-1}(X_2 - \mu_2)] \quad , \tag{13}$$

and W turns out to be $[A_{11}(K - P + 1)]^{-1}$.

Tendency Toward Normality

In this section it is shown that all marginals of the distribution approach normality as N approaches infinity. To do this factor K , which was defined as equal to $2N + P + 2$, out of the brackets in (1) and use the expression for C in (5) to get

$$\frac{\sqrt{|A|}}{(2\pi)^{\frac{P}{2}}} \left(\frac{\Gamma(N + \frac{P+2}{2})}{\Gamma(N + 1)(N + \frac{P+2}{2})^{\frac{P}{2}}} \right) \left(1 - \frac{Q}{K}\right)^N \quad ,$$

where Q is the quadratic form in the brackets of (1). Using Stirling's approximation in the quantity in braces above and letting $N \to \infty$ yields

$$\frac{\sqrt{|A|}}{(2\pi)^{\frac{P}{2}}} e^{-\frac{Q}{2}} \quad ,$$

which is the form of the multivariate normal distribution function.

Higher Moments

From expression (10), taking $Q = P - 1$ obtain the univariate marginal distribution as

$$\frac{\Gamma(N + \frac{P+2}{2})}{\Gamma(N + \frac{P+1}{2})\sigma \sqrt{\Pi}\, K^{N+\frac{P}{2}}} \left[K - \frac{(X - \mu)^2}{\sigma^2} \right]^{N+\frac{P-1}{2}} \quad . \tag{14}$$

Then

$$V = \frac{(X - \mu)^2}{K\sigma^2}$$

has the Dirichlet distribution,

$$\frac{\Gamma(N + \frac{P + 2}{2})}{\sqrt{\pi}\,\Gamma(N + \frac{P + 1}{2})} V^{-\frac{1}{2}}[1 - V]^{N+\frac{P-1}{2}} .$$

The r*th* moment of V about the origin, which is the 2r*th* moment of

$$\frac{(X - \mu)}{\sigma\sqrt{K}} ,$$

is (Press, 1972, p.133)

$$\frac{\Gamma(r + \frac{1}{2})\Gamma(N + \frac{P + 2}{2})}{\sqrt{\pi}\,\Gamma(N + \frac{P + 2}{2} + r)} .$$

Moments of V can be used to get the even moments of X , and from the symmetry, all odd moments are zero.

$$E(V) = E\left[\frac{(X - \mu)^2}{K\sigma^2}\right] = \frac{1}{2}\,\frac{\Gamma(\frac{1}{2})}{\Gamma(\frac{1}{2})}\;\frac{1}{N + \frac{P + 2}{2}} = \frac{1}{2N + P + 2} .$$

Since $K = 2N + P + 2$, it follows that $E(X - \mu)^2 = \sigma^2$ is in agreement with an earlier result.

For kurtosis take $r = 2$, to get

$$E\left[\frac{(X - \mu)^2}{K\sigma^2}\right]^2 = \frac{\Gamma(\frac{5}{2})}{\Gamma(\frac{1}{2})}\;\frac{\Gamma(N + \frac{P + 2}{2})}{\Gamma(N + \frac{P + 2}{2} + 2)} = \frac{3}{2}\cdot\frac{1}{2}\cdot\frac{1}{(N + \frac{P + 2}{2} + 1)(N + \frac{P + 2}{2} + 2)}$$

and

$$\tilde{\mu}_4 = E\left[\frac{(X - \mu)^4}{\sigma^4}\right] = 3\,\frac{2N + P + 2}{2N + P + 4} \quad . \qquad (15)$$

Note that $\lim_{N \to \infty} \tilde{\mu}_4 = 3$ as in the normal distribution. Further, since $\tilde{\mu}_4$ is monotonic in $2N + P$, it is at a minimum when the exponent in (14) equals zero, i.e., $2N + P - 1$ equals zero. Hence, a minimum for $\tilde{\mu}_4$ is $\frac{3}{5} \times 3 = 1.8$, the kurtosis of a rectangular distribution.

At the outset it was mentioned that this polynomial distribution function fits test data, at least in some respects, better than the normal. That this is true at least with respect to kurtosis can be seen by examination of Table 1 (for which the author is indebted to F. Swineford). The first data line in Table 1 indicates that for a certain form of SAT-V, which is scored R-W/4, the skewness and kurtosis measures based on 900 cases were .3417 and 2.5155 respectively. Further, by looking down the column headed Kurtosis, one can see that in almost all cases the measure of kurtosis, which is the ratio of the fourth moment around the mean to the squared variance, lies between 1.8 and 3.0 as indicated in equation (15) and the following discussion. If measures of kurtosis were symmetrically distributed around three, the probability of observing three or fewer kurtosis measures in excess of three is less than ten to the minus eight for thirty-two such measures. Clearly, three is certainly too large a number to be the median of the kurtosis measures in Table 1. Those three for which the normal might be better, on the basis of the kurtosis measures, are the most skewed and hence still might not be very well described by the normal distribution. In most cases, the polynomial distribution seems more appropriate than the normal on the basis of the measure of kurtosis, and one may note that this holds across test programs (SAT, GRE, LSAT, ATGSB, and SSAT) and hence over different scoring formulae.

Insert Table 1 about here

Table 2 contains data similar to those in Table 1. However, the data in Table 1 come from test programs in which many parallel tests have been constructed over the years for essentially similar populations, but those in Table 2 (for which the author is indebted to F. Lord) were constructed in a somewhat less familiar situation. Also, though the examinees for the testing program whose data are presented in Table 2 were volunteers, they did include some 80 to 90% of the students in the grades listed who were selected on a statewide basis, and who certainly extend the grade range of students beyond that provided in Table 1. In short, no attempt is made to choose data which are selective against the normal distribution.

Insert Table 2 about here

Another, and perhaps more important, kind of agreement between the polynomial distribution and data is in the matter of homoscedasticity. Table 3 contains data which indicate agreement between the polynomial distribution and observation. This table contains two types of evidence. The entries for the table come from scatterplots of reported SAT scores and shorter tests which are built to the same specifications except for length. The data in Table 3 concern reported score variances observed at each level of the shorter tests. The first column in Table 3 gives the correlations of the variances with the associated short test score level, and the second column shows the multiple correlation of the variances with the score level and its square. Note that in some cases the gain by using the quadratic term is substantial, but of course the correlation with two independent variables must be better than with one. Therefore the data in the last two columns are of special interest. These data give the standard score regression coefficients, and the important thing to note is that in every case the pattern of signs is identical, being negative for the quadratic term. The probability that all thirty-one examples would have the same sign is two to the minus thirtieth, and thus the trend noted seems quite reliable. In fact, the negative sign for the coefficient of the quadratic term is predictable on the basis of the polynomial distribution since in the discussion following the expression (8) it is pointed out that the coefficients of variables remaining after integration are the elements of the inverse of a matrix B_{22} which is a principal submatrix of a population variance-covariance matrix that will be positive definite. Therefore, the quadratic form in (13) must be positive definite, and it follows that the coefficient of the squared term is positive in the quadratic form or negative in (13). Thus, one would expect the signs of the quadratic terms in Table 3 to be negative, which they were. Under a null hypothesis of $p = .5$ (either sign is equally likely) the probability of all signs being negative is only two to the minus thirty-one (as opposed to the two to the minus thirty mentioned above).

Insert Table 3 about here

Thus, in two ways the data observed give better agreement with the polynomial distribution than with the normal. Of course, it is recognized that one must not assume that because the polynomial distribution is better in some respects it is better in all respects. There may be other deductions for which the normal is better. However, the agreement of the polynomial distribution with the lack of homoscedasticity seems fairly important in some applications, particularly in their relation to range restriction formulae.

References

Boldt, R. F. A multivariate function useful in personnel management models. Proceedings of First Army Operations Research Symposium, 1962.

Boldt, R. F. A study of linearity and homoscedasticity of test scores in the chance range. Educational and Psychological Measurement, 1968, 28, 47-60.

Boldt, R. F. Anchored scaling and equating: Old conceptual problems and new methods. College Board Research and Development Reports, RDR-72-73, No. 2, and ETS Research Bulletin 72-28. Princeton, N. J.: Educational Testing Service, 1972.

Graybill, F. A. An introduction to linear statistical models, Vol. I. New York: McGraw-Hill, 1961.

Mood, A. M. Introduction to the theory of statistics. New York: McGraw-Hill, 1950.

Press, S. J. Applied multivariate analysis. New York: Holt, Rinehart, & Winston, 1972.

Raiffa, H., & Schlaifer, R. Applied statistical decision theory. Boston: Harvard University, 1961. Pp. 129, 259-260.

Table 1

Skewness and Kurtosis of Various Operational Test Score Distributions

Test	Scoring Formula	No. Cases	Kurtosis	Skewness
SAT[a]-V[b]	R-W/4	900	2.5155	.3417
		865	2.7244	.4210
		900	2.6078	.1360
		2345	2.5168	.2492
		2000	2.6509	.3016
SAT-M[c]	R-W/4	900	2.3701	.2342
		865	2.4790	.2146
		900	2.3074	.0700
		2345	2.5009	.0245
		2000	2.5236	.1156
GRE[d]-V	R-W/4	1660	2.3863	-.0212
		1485	2.4395	-.0299
		700	2.4566	-.0341
		1495	2.4255	-.0788
GRE-Q[e]	R-W/4	1660	2.3154	-.0674
		1485	2.2742	-.1715
		700	2.4956	-.1244
		1495	2.4627	-.0863
LSAT[f]-Error Recognition	R	350	2.8700	-.4102
		1810	3.4599	-.5097
LSAT-Sentence Correction	R	350	2.8873	-.2638
		1810	2.9505	-.2344
LSAT-Reading Comprehension	R	1810	2.5161	-.1293
LSAT-Data Interpretation	R	1810	2.8832	.0481
LSAT-Reading Recall	R	1810	2.8627	-.2669
LSAT-Principles and Cases	R	1810	3.5500	-.4215
ATGSB[g]-Reading Recall	R-W/4	1090	2.7166	-.3354
		1260	3.1475	-.4912
		1260	2.8303	-.3187
ATGSB-M	R-W/4	1090	2.5628	-.1408
		1260	2.6103	.2223
		1260	2.5943	.1147

Table 1 (Continued)

Test	Scoring Formula	No. Cases	Kurtosis	Skewness
ATGSB-V	R-W/4	1090	2.4796	.1597
		1260	2.4084	.0629
		1260	2.4298	.0763
SSAT[h]-V	R-W/4	860	2.4076	.1659
		1820	2.4063	.0453
SSAT-M	R-W/4	860	2.7983	.4039
		1820	2.7536	.4367
SSAT-Reading Comprehension	R-W/4	860	2.7667	.4494
		1820	2.3904	.1878

[a] Scholastic Aptitude Test (SAT).

[b] Verbal (V).

[c] Mathematics (M).

[d] Graduate Record Examination (GRE).

[e] Quantitative.

[f] Law School Aptitude Test.

[g] Admissions Test for Graduate School in Business.

[h] Secondary School Aptitude Test.

Table 2

Survey of Students in a Midwestern State, September 1965

Kurtosis and Skewness (in Parentheses) of Test Score Distributions[a]

Test and No. of Items	Grade 4	Grade 6	Grade 8	Grade 10
English Achievement 20 items	2.284 (-.113)	2.586 (-.092)	2.372 (-.016)	
Math Ability 50 items	2.336 (.079)	2.481 (.018)	2.426 (.034)	2.438 (-.011)
Math Achievement 50 items	2.611 (.132)	2.900 (.138)	3.001 (.276)	2.864 (.200)
Reading 30 items	2.629 (-.479)	5.837 (-1.955)	2.893 (-.287)	
Verbal Ability 50 items	2.975 (.394)	2.491 (.112)	2.238 (.087)	2.426 (-.001)

[a]No entry is based on less than 19,500 cases, all volunteers.

Table 3

Correlation between Linear and Quadratic Functions of Equating Test Scores and the Conditional Variances and Standard Score Regression Coefficient for the Quadratic Function

Test	Correlations Linear	Correlations Quadratic	Regression Terms Linear	Regression Terms Quadratic
SAT-M	.3671	.6759	1.5920	-2.0396
	.0100	.8823	2.4882	-2.6494
	.3311	.8950	2.7800	-2.5862
	.3009	.5079	.8909	-1.2601
	.0158	.1810	.4856	-.5033
	.0126	.8184	2.1283	-2.2919
	.0278	.9179	2.7297	-2.8525
	.0322	.8518	2.2592	-2.3842
	.2287	.9558	3.2376	-3.5884
	.2563	.7507	2.8917	-2.7232
	.4807	.7687	1.7599	-2.3196
	.3864	.7027	1.8381	-2.3005
	.3037	.4656	1.0143	-1.3645
	.1179	.7061	2.4824	-2.6919
	.5800	.7334	1.0378	-1.6850
	.6531	.8880	1.5945	-2.3267
SAT-V	.2615	.5233	1.2869	-1.6134
	.0469	.7340	2.3549	-2.5110
	.1849	.4899	1.1059	-1.3632
	.0985	.8230	2.4892	-2.5350
	.0521	.8230	2.4892	-2.5350
	.0521	.7814	2.2954	-2.4736
	.3161	.7213	1.6617	-2.0813
	.1708	.7695	2.4980	-2.4452
	.3795	.7766	2.1826	-2.6801
	.2380	.3280	.6155	-.8828
	.4377	.7744	1.9791	-2.4998
	.4190	.7142	1.7678	-2.2619
	.2794	.4157	.8013	-1.1364
	.0534	.6771	2.5506	-2.6901
	.0978	.1020	.2072	-.1132
	.0170	.0665	.2598	-.2511