THE INDEX OF DISPERSION

By

EDGAR G. AVELINO

B.Sc., The University of British Columbia, 1978

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

Statistics Department

The University of British Columbia

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

November 1984

ABSTRACT

The index of dispersion is a statistic commonly used to detect departures from randomness of count data. Under the hypothesis of randomness, the true distribution of this statistic is unknown. The accuracy of large sample approximations is assessed by a Monte Carlo simulation. Further approximations by Pearson curves and infinite series expansions are investigated. Finally, the powers of the individual tests based on the likelihood ratio, the index of dispersion and Pearson's goodness-of-fit statistic are compared.

TABLE OF CONTENTS

PAGE

TABLE OF CONTENTS

LIST OF TABLES

PAGE

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENT

I wish to express my sincere appreciation to Prof. A. John Petkau who devoted precious time to the supervision of this thesis.

Edgar G. Avelino

1. INTRODUCTION

1.1 HISTORY OF THE INDEX OF DISPERSION

The index of dispersion is a test statistic often used to detect spatial pattern, a term ecologists use to describe non-randomness of plant populations. This is equivalent to testing that the growth of plants over an area is purely random, or equivalently that the number of plants in any given area has the Poisson distribution.

Suppose then that we randomly partition some area by n disjoint equal-sized quadrats and make a count, x, of the number of plants in each quadrat. Under the hypothesis of randomness, $X_1, \ldots, X_n$ would have the Poisson distribution,

$$P(X=x) = e^{-\lambda}\lambda^x/x! \text{ , for } \lambda > 0 \text{ and } x = 0, 1, 2, \ldots ,$$

for which $E(X) = Var(X) = \lambda$.

For alternatives to complete randomness involving patches or clumping of plants, we would expect $Var(X) > E(X)$, while for more regular spacing of plants, we would expect $Var(X) < E(X)$ (see for example R.H. Green (1966)) .

These properties lead quite naturally to considering the variance-to-mean ratio as a population index to measure spatial pattern. An estimator of the variance-to-mean ratio is the index of dispersion, defined as

$$I = \begin{cases} 1 & \text{, if } \bar{X} = 0 \\ S^2/\bar{X} & \text{, if } \bar{X} > 0 \end{cases}$$

where $\bar{X} = (1/n) \sum_{i=1}^{n} X_i$ and $S^2 = \{1/(n-1)\} \sum_{i=1}^{n} (X_i - \bar{X})^2$, the unbiased estimators of E(X) and Var(X), respectively. (It is natural to define I to be 1 if $\bar{X} = 0$ because under the null hypothesis, the variance-to-mean ratio equals 1.)

Ever since G.E. Blackman (1935) used the Poisson model for counts of plants, the concept of randomness in a community of plants became a growing interest among ecologists. Although the index of dispersion was introduced by R.A. Fisher (Fisher, Thornton and MacKenzie, 1922), it was not until 1936 that it was first used by ecologists for the purpose of inference. A.R. Clapham (1936), using a χ^2 approximation for the distribution of the index of dispersion under the null hypothesis of randomness, found that among 44 plant species he studied, only four of these seemed to be distributed randomly, while over-dispersion (i.e. clumping) was clearly present for the remaining species. Student (1919) had already pointed out that the Poisson is not usually a good model for ecological data and in most cases, clumping occurs. This has been termed "contagious" by G. Polya (1930) and also by J. Neyman (1939).

Ever since Clapham's paper, the use of the index of dispersion as a test of significance of departures from randomness has been extensive, not just for field data, but also in other areas (for example, blood counts, insect counts and larvae counts).

Fisher et al (1922) showed that the distribution of the index of dispersion could be closely approximated by the χ^2 distribution with n-1 degrees of freedom. However, if the Poisson parameter, λ, is small, or if the estimated expectation, $\bar{X}$, is small, then the adequacy of the

χ^2 approximation becomes questionable. This is discussed by H.O. Lancaster (1952). Fisher (1950) and W.G. Cochran (1936) have pointed out that in this case, the test of randomness based on I should be done conditionally with given totals, ΣX_i. Since this sum is a sufficient statistic for the Poisson parameter λ, conditioning on the total will yield a distribution independent of λ. Hence, exact frequencies can be computed. The conditional moments of the index of dispersion are provided in Appendix A1.2. These moments are also given by J.B.S. Haldane (1937) (see also Haldane (1939)).

Several people have examined the power of the test based on the index of dispersion. G.I. Bateman (1950) considered Neyman's contagious distribution as an alternative to the Poisson and found that this test exhibits reasonably high power for $n \geq 50$ and $m_1 m_2 \geq 5$, where m_1 and m_2 are the parameters of Neyman's distribution. For $5 \leq n \leq 20$, she found that the power is also high, provided that $m_1 m_2$ is large (in particular, $m_1 m_2 \geq 20$). Proceeding along the same lines as Bateman, N. Kathirgamatamby (1953) and J.H. Darwin (1957) compared the power of this test when the alternatives are Thomas' double Poisson, Neyman's contagious distribution type A and the negative binomial. They found that this test attained about the same power in each of the three alternatives.

Finally, in a recent paper, J.N. Perry and R. Mead (1979) investigated the power of the index of dispersion test over a wide class of alternatives to complete randomness. They concluded that this test is very powerful particularly in detecting clumping, and they strongly recommend the use of this test. Examination of the power of this test relative to other tests of the null may also be important.

1.2 PURPOSE OF THE PAPER

The purpose of this paper is to examine the distribution of the index of disperion and compare its power to the power of other tests of randomness. We examine the properties of the index of dispersion and through these properties, attempt to answer such questions as:

> "How do we decide whether a given sample is significantly different from a Poisson sample?" and "How good is this test in detecting departures from randomness relative to other (perhaps reliable and well-studied) tests?"

Answers to the first question could be based on constructing a rejection region R, where if $I \in R$, we would tend to favor some other alternative. For example, if we wished to test the null hypothesis against alternatives involving clumping, then large values of I would provide evidence against the null hypothesis, and the rejection region would presumably be of the form $I > C$. For two sided alternatives, we would be interested in both large and small values of this statistic, say $I < C_1$ or $I > C_2$. We would also want to examine the chances of wrongly rejecting the null which in statistical terminology is called the probability of making a type I error or the significance level (or size) of the test. The constants C, C_1, and C_2 are called critical values, and it is through these critical values that the rejection region will be constructed.

We then rephrase the question as:

"Is there a method of determining the rejection region R at a given level of significance α?"

As the true probability distribution of I is unknown, we first attempt to solve the problem through large sample approximations which lead to asymptotic critical values. We will show that the asymptotic null distribution of I is normal with mean 1 and variance 2/n. We can then use the critical values from the normal and determine how accurate these critical values are. This study is done through a Monte Carlo simulation. Similarly, the χ^2 approximation to the distribution of I is also examined. We also examine critical values obtained from approximating the null distribution of I by Pearson curves and Gram-Charlier expansions.

To assess the "goodness" of the index of dispersion, we might be interested in determining how often we would correctly reject the null in repeated sampling. This is called the power of the test, the complement of this being the probability of making a type II error. With the negative binomial as an alternative to the Poisson, the power of I is then compared to the power of the Likelihood Ratio Test and Pearson's Goodness-of-fit test.

2. LARGE SAMPLE APPROXIMATIONS

2.1 THE JOINT DISTRIBUTION OF $\bar{X}$ AND S^2

Suppose we choose a random sample of n disjoint equal-sized quadrats and make a count, X_i, of the number of plants in the i^{th} quadrat. Let $X_1, \ldots, X_n$ be independent identically distributed random variables with mean μ and variance σ^2. Let $\mu_k = E[(X_i-\mu)^k]$ and suppose that $\mu_4<\infty$. In particular, $\mu_1 = 0$ and $\mu_2 = \sigma^2$. As a consequence of the Central Limit Theorem, we have

$$\sqrt{n}\ (\bar{X}-\mu) \xrightarrow{d} N(0,\mu_2), \tag{2.1}$$

$$\sqrt{n}\ (S^2-\mu_2) \xrightarrow{d} N(0,\mu_4-\mu_2{}^2). \tag{2.2}$$

These results can be found in Cramer (1946, pp. 345 - 348).

Similarly, the Multivariate Central Limit Theorem implies that $\sqrt{n}(\bar{X}-\mu)$ and $\sqrt{n}(S^2-\mu_2)$ converge jointly to a bivariate normal distribution with mean vector $\underline{0}$ and variance-covariance matrix $(1/n)\Sigma$, where

$$\Sigma = \begin{bmatrix} \mu_2 & \lim\limits_{n\to\infty} n\mathrm{COV}(\bar{X},S^2) \\ \lim\limits_{n\to\infty} n\mathrm{COV}(\bar{X},S^2) & \mu_4 - \mu_2{}^2 \end{bmatrix};$$

i.e. $$\begin{bmatrix} \sqrt{n}(\bar{X}-\mu) \\ \sqrt{n}(S^2-\mu_2) \end{bmatrix} \xrightarrow{d} N(\underline{0},\Sigma). \tag{2.3}$$

Assuming that $\mu = 0$, we have

$$\begin{aligned} COV(\bar{X},S^2) &= E(\bar{X}\cdot S^2) \\ &= (n/(n-1))\{E(\bar{X}U_2) - E(\bar{X}^3)\} \end{aligned}$$

where $U_2 = (1/n)\Sigma X_i^2$. Hence

$$\begin{aligned} E(\bar{X}U_2) &= (1/n^2)E\{\Sigma X_i{}^3 + \sum\sum_{i\neq j} X_i X_j{}^2\} \\ &= \mu_3/n \text{ , since by independence, the double} \end{aligned}$$

sum has zero expectation. Similarly,

$$E(\bar{X}^3) = \mu_3/n^2$$

from which it follows that

$$COV(\bar{X},S^2) = \mu_3/n + O(1/n^2). \tag{2.4}$$

From (2.3) and (2.4) we have for large n that

$$\begin{bmatrix} \bar{X} \\ \\ S^2 \end{bmatrix} \approx N\left(\begin{bmatrix} \mu \\ \\ \mu_2 \end{bmatrix}, (1/n)\begin{bmatrix} \mu_2 & \mu_3 \\ & \\ \mu_3 & \mu_4-\mu_2{}^2 \end{bmatrix}\right).$$

2.2 THE ASYMPTOTIC DISTRIBUTION OF I

We compute the asymptotic distribution of $S^2/\bar{X}$ using the "delta method". It will be seen that the asymptotic distribution of I is the same as that of $S^2/\bar{X}$.

Let $g(x,y) = y/x$, so that $S^2/\bar{X} = g(\bar{X},S^2)$. Assuming that $\partial g/\partial x$ and $\partial g/\partial y$ exist near the point (μ,σ^2), (note that this requires the assumption that $\mu > 0$), we can expand $g(\bar{X},S^2)$ in a Taylor series about (μ,σ^2) and have

$$g(\bar{X},S^2) = g(\mu,\sigma^2)+(\bar{X}-\mu)g_x(\mu,\sigma^2)+(S^2-\sigma^2)g_y(\mu,\sigma^2)+ \ldots$$

Let $\underline{U}(n)' = (\bar{X},S^2)$ and $\underline{b} = (\mu,\sigma^2)'$. Then

$$\text{(i)} \quad \underline{U}(n) \xrightarrow{p} \underline{b} \text{ ; (ii) } \sqrt{n}(\underline{U}(n)-\underline{b}) \xrightarrow{d} N(\underline{0},\Sigma).$$

The result of the delta method (see, for example, T.W. Anderson (1958, pp. 76 - 77)) is that

$$\sqrt{n}\,((S^2/\bar{X}) - (\sigma^2/\mu)) \xrightarrow{d} N(0\ ,\ \phi_b{}'\Sigma\phi_b),$$

where $\phi_b{}' = (\partial g/\partial x, \partial g/\partial y)$ evaluated at (μ, σ^2).

Under the stated assumptions, we have for large n,

$$S^2/\bar{X} \approx N(\sigma^2/\mu\ ,\ (1/n)\phi_b{}'\Sigma\phi_b). \tag{2.5}$$

After some matrix calculations, we get

$$\phi_b{}'\Sigma\phi_b = \mu_2{}^3/\mu^4 - 2\mu_2\mu_3/\mu^3 + (\mu_4 - \mu_2{}^2)/\mu^2. \tag{2.6}$$

So far, all of the results hold regardless of the underlying distribution of the X's. If we now assume that $X_1,\ldots,X_n \sim P(\lambda)$, then

$$\mu = E(X) = \lambda\ ,$$

$$\mu_2 = Var(X) = \lambda\ ,$$

$$\mu_3 = \lambda \text{ and}$$

$$\mu_4 = 3\lambda^2 + \lambda\ .$$

Substituting these into (2.6), we have that

$$\begin{aligned}\phi_b{}'\Sigma\phi_b &= (1/\lambda) - (2/\lambda) + [(3\lambda^2+\lambda) - \lambda^2]/\lambda^2 \\ &= 2,\end{aligned}$$

and hence from (2.5) that

$$S^2/\bar{X} \approx N(1\ ,\ 2/n)\ .$$

The asymptotic null distribution of I is easily seen to be the same as that of $S^2/\bar{X}$ since

$$P\{|I - (S^2/\bar{X})| > \varepsilon\} = P\{\bar{X}=0\} = e^{-n\lambda} \text{ for any } \varepsilon>0.$$

This probability approaches 0 as $n \rightarrow \infty$, and hence

$$I \approx N(\ 1\ ,\ 2/n)\ .$$

Note that the $O(1/n)$ approximation to the variance of I is independent of the parameter λ. This would be useful in practice because the source of error in estimating λ by the maximum likelihood estimator $\bar{X}$ would not have to be introduced. We should note however that the inclusion of higher order terms will introduce this dependence.

2.3 DESCRIPTION OF THE MONTE CARLO SIMULATION

To answer the question of how well the asymptotic critical values work, we perfromed a Monte Carlo simulation when the underlying distribution of the X's is Poisson. Fifteen thousand samples of n Poisson random variables were generated for $n = 10,20,50,100$ and for $\lambda = 1,3,5,8$ and fifteen thousand indices of dispersion were computed for each pair n and λ. The 1%,2.5%,5% and 10% quantiles for each pair n and λ are given in Table A2 . With such a large number of samples, these critical values may be regarded as exact and they assist in assessing the accuracy of the asymptotic critical values.

Given a nominal significance level α, two-sided rejection regions were constructed with $\alpha/2$ in each tail. Using the asymptotic normal critical values, the rejection regions used were the following:

$$R_{.01} = \{I^*: |I^*| > 2.58\}$$

$$R_{.05} = \{I^*: |I^*| > 1.96\}$$

$$R_{.10} = \{I^*: |I^*| > 1.64\}$$

$$R_{.20} = \{I^*: |I^*| > 1.28\}$$

where $I^* = (I-1)/\sqrt{(2/n)}$ and where R_α denotes the rejection region at the nominal significance level α. To test the accuracy of the normal critical values, we merely count the number of I^*'s that fall in R_α. This would then give us an estimate $\hat{p}$, of the true significance level p.

Now, $\hat{p}$ = (# of I*'s ϵ R_α)/15,000. Since the number of I*'s ϵ R_α is binomially distributed (with parameters N=15,000 and p), the standard error of $\hat{p}$ is

$$SE(\hat{p}) = \sqrt{p(1-p)/15{,}000} \; .$$

We then might conclude that the distribution of I is well approximated by the normal if $\hat{p}$ is within one standard error of the nominal significance level α.

To assist the reader in interpreting the results, we supply a list of how many I*'s would be expected in each tail of the rejection region if the true significance level corresponding to each tail was identically equal to $\alpha/2$, one-half the nominal significance level.

Table 1: Expected Number of I*'s

α	$\{\alpha/2 \pm SE(\hat{\alpha}/2)\} \cdot 15{,}000$
0.01	75 ± 9
0.05	375 ± 19
0.10	750 ± 27
0.20	1500 ± 37

The results, summarized in Tables 2(A-D), are shown in the following pages. The entries in the "I<L" and "I>U" columns are the number of I*'s that lie to the left and right of the lower and upper normal critical values, respectively.

We immediately notice that the normal approximation is very poor even for n as large as 50. The lower critical values are much

NORMAL APPROXIMATION

Table 2A ($\alpha = 0.01$)

	λ = 1		3		5		8	
n	I<L	I>U	I<L	I>U	I<L	I>U	I<L	I>U
10	0	342	0	293	0	286	0	276
20	0	280	0	243	0	227	0	213
50	9	220	9	197	13	175	12	182
100	19	184	27	145	22	134	26	136

Table 2B ($\alpha = 0.05$)

	λ = 1		3		5		8	
n	I<L	I>U	I<L	I>U	I<L	I>U	I<L	I>U
10	8	713	3	645	0	670	2	646
20	50	678	74	633	76	607	72	587
50	148	587	183	546	199	533	197	515
100	199	538	248	506	244	490	262	481

Table 2C (α = 0.10)

	λ = 1		3		5		8	
n	I<L	I>U	I<L	I>U	I<L	I>U	I<L	I>U
10	114	951	142	974	145	1011	149	1007
20	283	1010	368	964	345	983	359	986
50	457	962	519	909	527	938	541	903
100	554	943	570	872	570	859	592	862

Table 2D (α = 0.20)

	λ = 1		3		5		8	
n	I<L	I>U	I<L	I>U	I<L	I>U	I<L	I>U
10	596	1520	962	1571	953	1633	958	1607
20	1008	1447	1180	1627	1212	1643	1221	1648
50	1216	1612	1338	1589	1334	1587	1324	1584
100	1386	1528	1377	1522	1322	1576	1355	1564

too liberal. For the cases $n \geq 50$, the total number of I^*'s in the rejection region is close to the total number we would expect to be rejected, but the significance level in each tail is nowhere near $\alpha/2$. The probability of falsely rejcting the null would be too low in the lower tail and too high in the upper tail. There is obviously a problem of skewness in the distribution. Too many observations lie in the right tail implying that the distribution of I is positively skewed. Notice that for fixed λ and increasing n, the number of I^*'s rejected in each tail becomes more equal. However even for n=100, the lower critical values are still conservative in all significance levels while the upper critical values are too liberal. For fixed n and increasing λ on the other hand, no such pattern is obvious. Thus, it appears that the normal approximation is really only satisfactory for $n>100$ and this will not suffice for practical work.

An examination of the probability plots and histograms (see Appendix 2.1) provides more detail.

One might hope to find an improvement to this approximation and one approach taken to improve the approximation is through infinite series expansions (e.g. Edgeworth, Gram-Charlier, Fisher-Cornish expansions). The trade-off for having such an improvement is the requirement of higher order moments; and these higher order moments will surely have a dependence on λ. More of this will be seen in later chapters. For the moment, we abandon the normal approximation and move on to another simple large sample approximation.

2.4 THE χ^2 APPROXIMATION

As seen in section 2.2, the probability under the null that I and $S^2/\bar{X}$ differ by an amount bigger than $\varepsilon(\varepsilon>0)$ is $e^{-n\lambda}$ which approaches 0 as $n \longrightarrow \infty$. It is therefore sufficient to consider an approximation to the distribution of $S^2/\bar{X}$.

At a first glance, we might suspect that $S^2/\bar{X}$ has some relationship with the χ^2 distribution, for it is well known that if $X_1, \ldots, X_n$ is a random sample from the normal distribution with mean μ and variance σ^2, then

$$(n-1)S^2/\sigma^2 \sim \chi^2_{n-1} .$$

In our case, the X's are Poisson and Var(X) is only approximated by $\hat{\sigma}^2 = \bar{X}$. However, it would not be surprising that the null distribution of $(n-1)S^2/\bar{X}$ could be well approximated by χ^2_{n-1} for large n. A clearer motivation of this is outlined below.

Consider the following one-way contingency table:

O_j:	X_1	X_2	...	X_n	$X.$
E_j:	$\bar{X}$	$\bar{X}$	...	$\bar{X}$	$n\bar{X}$

The entries in the cells of the first row are just the observed counts themselves, having row total X., and the entries in the second row are the estimated expected counts, $\bar{X}$. (Note that this contingency table differs from the ordinary contingency table where observations are free to fall in any one cell. In our contingency table, we have one cell for each count. However, if we considered only those sampling experiments that produced the same order of experimental results in addition to the same marginal totals, the methods of the ordinary contingency tables still apply.) The goodness-of-fit statistic is formed by summing up over the columns, the square of the difference between the observed and the expected values and dividing this by the expected value. This gives us

$$\sum_{i=1}^{n} (X_i-\bar{X})^2/\bar{X} \ ,$$

which is precisely $(n-1)S^2/\bar{X}$. Providing the E_j's are not too small (for example, $E_j \geq 5$ for all j), the distribution of the goodness-of-fit statistic might be expected to be well approximated by χ^2_{n-1} for large n.

This motivation is due to P.G. Hoel (1943). In his paper, he approximated the moments of $S^2/\bar{X}$ under the null hypothesis by power series expansions, correct to $O(1/n^3)$, and showed that the first four moments of $(n-1)S^2/\bar{X}$ were in close agreement with those of the χ^2_{n-1} distribution.

2.5 DISCUSSION

Returning now to the simulation study, we recall that since the normal distribution is symmetric, it could not account for the skewness of the distribution of I. On the other hand, since the χ^2_{n-1} distribution is skewed, one might expect it to perform better than the normal approximation.

So as not to obscure the comparison of the two approximations, the same 15,000 samples generated for each case were used. The results are displayed in Tables 3(A-D).

χ^2 APPROXIMATION

Table 3A ($\alpha = 0.01$)

	λ							
	1		3		5		8	
n	I<L	I>U	I<L	I>U	I<L	I>U	I<L	I>U
10	44	72	59	81	77	83	77	60
20	39	100	69	93	63	76	66	78
50	50	93	69	93	73	90	75	82
100	50	102	72	72	72	67	77	69

Table 3B ($\alpha = 0.05$)

	λ							
	1		3		5		8	
n	I<L	I>U	I<L	I>U	I<L	I>U	I<L	I>U
10	182	342	356	364	379	356	385	360
20	213	408	336	385	340	365	359	360
50	285	408	350	405	363	380	373	365
100	309	419	349	396	359	373	370	367

Table 3C (α = 0.10)

λ

	1		3		5		8	
n	$I<L$	$I>U$	$I<L$	$I>U$	$I<L$	$I>U$	$I<L$	$I>U$
10	587	713	633	707	730	733	777	720
20	537	741	725	774	704	791	750	756
50	607	807	729	775	708	762	765	756
100	676	796	693	757	684	738	709	733

Table 3D (α = 0.20)

λ

	1		3		5		8	
n	$I<L$	$I>U$	$I<L$	$I>U$	$I<L$	$I>U$	$I<L$	$I>U$
10	968	1083	1383	1429	1464	1478	1489	1469
20	1225	1350	1478	1504	1468	1507	1517	1539
50	1463	1534	1448	1528	1457	1505	1468	1495
100	1440	1489	1432	1472	1411	1508	1454	1502

The χ^2 approximation clearly gives a better fit to the null distribution of I than the normal. Most of the entries in the cells fall within the range of values that one would expect to see. Notice that these tables display a similar pattern, namely that symmetry between the "I<L" and "I>U" columns becomes more apparent with increasing n and fixed λ (and with increasing λ and fixed n). This pattern appears with increasing α too. However, there seems to be more room for improvement for the cases $n \leq 20$ and $\lambda < 3$. In fact, even for $n=100$ and $\lambda=1$, the lower critical values tend to be too conservative while the upper critical values tend to lead to rejection too often. (In the ecological context however, this would cause no serious problems. One can simply take larger quadrats to ensure that the mean number of plants in each quadrat is larger than 1.) For the cases $n \geq 20$ and $\lambda \geq 3$, the χ^2 approximation gives a very reasonable approximation to the null distribution of I, and leads to a pleasantly simple method of constructing rejection regions.

As stated in the previous section, one way of improving these large sample approximations is through an infinite series expansion of the true density of I. Another technique commonly used is approximations by Pearson curves, which will require the first four moments of I.

3. PEARSON CURVES

3.1 THE THEORY OF PEARSON CURVES

The family of distributions that satisfies the differential equation

$$d(\log f)/dx = (x-a)/(b_0+b_1+b_2x^2) \qquad (3.1)$$

are known as Pearson Curves. Under regularity conditions, the constants a, b_0, b_1 and b_2 can be expressed in terms of the first four moments of the distribution f; see Kendall and Stuart(1958, vol. 1, p. 149). Karl Pearson (1901) identified 12 types of distributions each of which is completely determined by the first four moments of f.

It is convenient to rewrite the denominator as

$$B_0 + B_1(x-a) + B_2(x-a)^2$$

for suitably chosen constants B_0, B_1 and B_2, and hence (3.1) may be written as

$$d(\log f)/dx = (x-a)/\{B_0 + B_1(x-a) + B_2(x-a)^2 . \qquad (3.2)$$

As in (3.1), the constants in (3.2) are functions of the first four moments of f(x). By integrating the right-hand side of (3.2), an explicit expression can be obtained for f(x).

The criterion for determining which type of Pearson curve results is obtained from the discriminant of the denominator in (3.2). This criterion is given by:

$$\kappa = B_1^2/4B_0B_2 . \qquad (3.3)$$

Defining $\beta_1 = \mu_3^2/\mu_2^3$ and $\beta_2 = \mu_4/\mu_2^2$, where the μ_i's are the central moments for $f(x)$, the constants B_0, B_1 and B_2 can be expressed in terms of β_1 and β_2. The criterion κ then becomes

$$\kappa = \beta_1(\beta_2+3)^2/\{4(2\beta_2-3\beta_1-6)(4\beta_2-3\beta_1)\}. \tag{3.4}$$

For example, a value of $\kappa<0$ gives Pearson's type I curve, also called the Beta distribution of the first kind. In this case,

$$f(x) = kx^{p-1}(1-x)^{q-1}, \text{ for } 0 \le x \le 1,$$

where the constants k,p and q are functions of the first four moments.

If $1<\kappa<\infty$, then we get Pearson's type VI curve, also known as the Beta distribution of the second kind. Here,

$$f(x) = kx^{p-1}/(1+x)^{p+q}, \text{ for } 0 \le x < \infty.$$

The following is a summary of the steps one would take when approximating by Pearson curves.

Let $g(\underline{X})$ be a statistic whose null distribution we wish to approximate by Pearson curves. The first step is to compute the first four moments of $g(X)$ which will depend on the parameters of the null distribution of the X's (if the parameters are not specified, they may be estimated by the maximum likelihood). Then β_1 and β_2 can be computed from the moments.

From here, either one of two routes can be taken. If critical values are all that are required, then the Biometrika tables published by Pearson and Hartley (1966) can be used. The critical values are tabulated for a wide range of values of $\sqrt{\beta_1}$ and β_2, and if necessary, linear interpolation along rows and columns is sufficient. We should note that when using critical values from the Biometrika tables, one should keep in mind that those critical values are all standardized. So if X_α denotes the α level critical value from the tables, then the appropriate α level critical value to use for the test is

$$x_\alpha = (\sqrt{\mu_2})X_\alpha + \mu .$$

The other alternative is to compute κ from (3.4) and determine the type of Pearson curve to be used. If the resulting distribution is not too uncommon, the parameters of the distribution can be computed. The text by W.P. Elderton and N.L. Johnson (1964, pp.35-46) gives an excellent treatment of this situation. Once the Pearson curve is completely determined, critical values can usually be obtained from the computer. In particular, the IMSL library provides critical values for a wide class of distributions.

3.2 TWO EXAMPLES

Before applying Pearson curves as an approximation to the null distribution of I, we discuss briefly two of the examples from the paper by H. Solomon and M. Stephens (1978), where the accuracy of critical values obtained from a Pearson curve fit is examined.

Example 1: Let $Q_n(\underline{c},\underline{a}) = \sum_{i=1}^{n} c_i(X_i + a_i)^2$,

where $\underline{c} = (c_1,\ldots,c_n)'$ and

$\underline{a} = (a_1,\ldots,a_n)'$

are vectors with constant components, and the X_i's come from a standard normal distribution.

The exact moments of this statistic are known to any order. In fact, the rth cumulant κ_r, is

$$\kappa_r = 2^{r-1}\,(r-1)!\sum_{i=1}^{n} c_i^r(1+ra_i^2).$$

There is a long literature on obtaining the critical values for different combinations of n,$\underline{c}$ and $\underline{a}$. These critical values are tabulated in Grad and Solomon (1955) and in Solomon (1960). Much mathematical analysis was used to obtain these critical values and an extensive amount of numerical computations were made so accurately that all the critical values can be regarded as exact.

Pearson curve fits were obtained for different values of the constants and critical values were obtained by quadratic interpolation from Biometrika tables. The results were that the Pearson curve critical values agreed very closely with the exact critical values for the upper tail, but there was no close agreement at all between the two critical values in the lower tail of the distribution.

Now, Pearson curves can also be obtained when the first three moments and a left endpoint of the distribution are known. (See for example, R.H. Muller and H. Vahl (1976) and A.B. Hoadley (1968)). Solomon and Stephens proceeded to do this three-moment fit and found that the fit in the lower tail was improved considerably, but this approach made the fit in the upper tail less accurate. They point out however that whenever four moments are available, then these four moments should be used for the fit as it is the upper tail of the distribution that is usually of more importance in practice. Of course, in our case, depending on whether we are concerned with clumping or regular spacing of plants, either tail of the distribution might be of interest.

Example 2: Solomon and Stephens considered the statistic $U = R/S$, where R is the range and S the standard deviation of a sample from the standard normal distribution. For n=3, the density of U is known:

$$f(u) = (3/\pi)\{1-(u^2/4)\}^{-(1/2)}, \text{ for } \sqrt{3} \leq u \leq 2.$$

The Pearson curve turned out to be a Beta distribution of the first kind and had the form

$$g(u) = 0.9573/\{(u-1.7324)^{0.0101}(2.000-u)^{0.4970}\},$$

for $1.7324 \leq u \leq 2.000$.

First we notice that while the true distribution of U is bell-shaped, the Pearson curve is U-shaped. However, notice that the Pearson curve fit gave the correct left and right endpoints of the distribution, at least to three decimal places. Finally, Solomon and Stephens found that the Pearson curve crticial values agreed very well with the exact critical values in both the lower and upper tails of the distribution. Given that the Pearson curve is U-shaped, the accurate fit obtained in both tails of the distribution is extremely surprising!

These two examples illustrate the usefulness of Pearson curves as a means of approximating perhaps not so much the distribution, but the critical values.

3.3 THE FIRST FOUR MOMENTS OF I

Computing the first four moments of I is no easy task since I involves a random variable in its denominator, namely $\bar{X}$. However, we can express the expectation of I^k, for k = 1,2,3,4, as the expectation of the conditional expectation given the total $X. = \sum_{i=1}^{n} X_i$ (or given $\bar{X}$). Now

$$I^k = \begin{cases} 1 & \text{if } \bar{X} = 0 \\ (S^2/\bar{X})^k & \text{if } \bar{X} > 0 \end{cases}$$

If follows that

$$E(I^k|X.) = \begin{cases} 1 & \text{if } X. = 0 \\ E\{(S^2/\bar{X})^k|X.\} & \text{if } X. > 0 \end{cases}$$
$$= i\{X.=0\} + E\{(S^2/\bar{X})^k|X.\}\ i\ \{X.>0\},$$

where i{A} is an indicator function equalling 1 if A is true and 0 otherwise. Hence

$$\begin{aligned} \mu_k' &= E(I^k) \\ &= E\{E(I^k|X.)\} \\ &= P(X.=0) + \sum_{j=1}^{\infty} E\{(S^2/\bar{X})^k|\ X.=j\}\cdot P(X.=j) \\ &= P(X.=0) + E^+\{E[(S^2/\bar{X})^k|X.]\} \qquad (3.5) \end{aligned}$$

where E^+ denotes an expectation over the marginal distribution of X. restricted to the positive values of X. . Thus we may write the <u>kth</u>

moment of I as

$$\mu_k' = P(X.=0) + E^+\{(1/\bar{X})^k \cdot E[(S^2)^k \mid \bar{X}]\}. \qquad (3.6)$$

Hopefully, the conditional expectation, which will depend on $\bar{X}$, might cancel off the $\bar{X}^k$ in the denominator, and hence computation of the unconditional expectation will be relatively easy.

Another difficulty arises in computing the conditional expectation itself, which involves expanding

$$(S^2)^k = \{[1/(n-1)] \sum_{i=1}^{n} (X_i-\bar{X})^2\}^k, \text{ for } k = 1,2,3,4.$$

The conditional expectation of this random variable will involve moments and product moments of $(X_1,X_2,\ldots,X_n)|X.$ up to the eighth order, and hence if we choose to compute moments from the moment generating funtion, we would require mixed partial derivatives of the moment generating function of $(X_1,X_2,\ldots,X_n)|X.$ up to the eighth order. Fortunately, when the underlying distribution of the X_i's is Poisson, the distribution of $(X_1,X_2,\ldots,X_n)|X.$ has a farily simple form, reducing to a multinomial distribution with parameters X. and $p_i = 1/n$, for $i=1,2,\ldots,n$;

$$\text{i.e. } (X_1,X_2,\ldots,X_n)|X. \sim \text{Mult}_n(X.,1/n,\ldots,1/n).$$

This result, the derivation of which is provided in Appendix A1.1, facilitates the derivation of the conditional moments $E\{(S^2/\bar{X})^k|X.=x.\}$ for $x.>0$. These are provided in equations (A1.3)-(A1.6) of Appendix

A1.2 and lead via (3.6) to the following expressions for the first four raw moments of I, the index of dispersion:

$$\mu_1' = P(X.=0) + P(X.>0) = 1$$

$$\mu_2' = P(X.=0) + \{(n+1)/(n-1)\}P(X.>0) - \{2/n(n-1)\}E^+(1/\bar{X})$$

$$\mu_3' = P(X.=0) + \{(n+1)(n+3)/(n-1)^2\}P(X.>0) + \{1/(n-1)^2\}\cdot \{(-4/n)[1-(6/n)]E^+(1/\bar{X}^2) - 2[1+(13/n)]E^+(1/\bar{X})\}$$

$$\begin{aligned}\mu_4' = {} & P(X.=0) + \{(n+1)(n+3)(n+5)/(n-1)^3\}P(X.>0) \\ & + \{1/(n-1)^3\}\{4n[1+(5/n)][1-(17/n)]E^+(1/\bar{X}) \\ & - (4/n^2)(2n^2+53n-261)E^+(1/\bar{X}^2) \\ & - (8/n^3)(n^2-30n+90)E^+(1/\bar{X}^3)\}. \end{aligned} \tag{3.7}$$

From these expressions, we see that we have not overcome the problem of evaluating the expectation $E^+(1/\bar{X}^k)$ for k = 1,2,3. We have taken two approaches in evaluating these expectations. The practical approach is to express these expectations as integrals, and evaluate these integrals by asymptotic expansions. This is done in Appendix A1.2; the resulting expressions for the raw moments, correct to $O(1/n^4)$, are provided in equation (A1.7). The central moments, correct to

$O(1/n^4)$, are then immediate:

$$\mu = 1 \text{ (exact)},$$

$$\mu_2 \sim 2/n + (2/n^2)\{1-(1/\lambda)\} + (2/n^3)\{1-(1/\lambda)-(1/\lambda^2)\}$$
$$+ (2/n^4)\{1-(1/\lambda)-(1/\lambda^2)-(2/\lambda^3)\} + O(1/n^5),$$

$$\mu_3 \sim (1/n^2)\{3+(4/\lambda)\} + (1/n^3)\{16-(24/\lambda)\}$$
$$+ (1/n^4)\{24-(52/\lambda)-(8/\lambda^2)-(4/\lambda^3)\} + O(1/n^5),$$

$$\mu_4 \sim 12/n^2 + (1/n^3)\{72+(72/\lambda)+(8/\lambda^2)\}$$
$$+ (1/n^4)\{180-(240/\lambda)-(228/\lambda^2)+(16/\lambda^3)\} + O(1/n^5). \quad (3.8)$$

Using the definition of β_1 and β_2, we can also express these in a similar expansion:

$$\beta_1 \sim (2/n)\{2+(1/\lambda)\}^2 + (2/n^2)\{4-(16/\lambda)-(3/\lambda^2)+(3/\lambda^3)$$
$$+ (2/n^3)\{4-(16/\lambda)-(7/\lambda^2)-(17/\lambda^3)+(7/\lambda^4)+ \ldots\}. \quad (3.9)$$

$$\beta_2 \sim 3 + (2/n)\{6+(12/\lambda) + (1/\lambda^2)\}$$
$$+ (2/n^2)\{6-(36/\lambda)-(5/\lambda^2)+(4/\lambda^3)\} + \ldots \quad (3.10)$$

The accuracy of these approximations can be assessed by computing the moments "exactly". By this, we mean computing the moments to a reasonable degree of accuracy. To do this, we can approximate the infinite series in the expressions for the exact moments by N^{th} partial sums, S_N, where N, the number of terms in the partial sum, is chosen so that the difference between the true and approximated values is no bigger than 10^{-6}, say. A geometric bound on the error is shown below.

$$\text{Let } S = (1/n)E^{+}(1/\bar{X}) = \sum_{k=1}^{\infty} (1/k)e^{-\theta}\theta^{k}/k!$$

Where $\theta = n\lambda$. We want to determine N such that $S-S_N \leq 10^{-6}$. Now,

$$\begin{aligned}
S-S_N &= \sum_{k=N+1}^{\infty} (1/k)e^{-\theta}\theta^{k}/k! \\
&\leq \sum_{k=N+1}^{\infty} e^{-\theta}\theta^{k}/k! \\
&= \{e^{-\theta}\theta^{N+1}/(N+1)!\}\{1+[\theta/(N+2)] + [\theta^2/(N+2)(N+3)] \\
&\quad + [\theta^3/(N+2)(N+3)(N+4)] + \ldots\} \\
&\leq \{e^{-\theta}\theta^{N+1}/(N+1)!\}\{1+(\theta/N)+(\theta/N)^2+(\theta/N)^3+ \ldots\} \\
&= \{e^{-\theta}\theta^{N+1}/(N+1)!\}\{1/[1-(\theta/N)]\}, \text{ if } \theta<N.
\end{aligned}$$

We therefore want to choose N so that

(i) $N > n\lambda$ and

(ii) $S - S_N \leq 10^{-6}$

We note that this same value of N can be used for $E^+(1/\bar{X}^2)$ and $E^+(1/\bar{X}^3)$ since convergence is faster in these cases.

The importance of a good asymptotic expansion is clear; one would not want to compute partial sums when fitting by Pearson curves. While computation of "exact" moments may be relatively inexpensive for small values of θ, it can get quite expensive for larger values of θ, and furthermore, over-flow problems will occur in these cases. The accuracy of the asymptotic expansions is discussed in the next section.

3.4 DISCUSSION

Using the known values of λ, we can compute exact and asymptotic moments and hence obtain two Pearson curve fits for the simulated data. The Pearson curves obtained are the following:

i) Using asymptotic moments (up to the fourth order) a type IV fit was obtained for the case $\lambda=1$ (for all n) and a type VI for all other cases.

ii) Using exact moments, the same types were obtained except for the case $n=10$ and $\lambda=1$ where the fit turned out to be a type VI.

The type IV Pearson curve is not a common distribution. $f(x)$ has the form

$$f(x) = k\{1+(x^2/a^2)\}^{-m} \exp\{-b \arctan(x/a)\}.$$

Since the critical values from this distribution cannot be obtained from the IMSL library, all the type IV critical values were obtained from Biometrika tables (Pearson & Hartley, (1966)). However, the critical values from a type VI curve can be obtained from IMSL. The algorithm for determining the form of the density is outlined in Appendix A1.3.

Using the same 15,000 samples for a given n and λ , a Monte Carlo study was done to assess the Pearson curve fits. The reults of the study are presented in Tables 4(A-D) and Tables 5(A-D).

As seen from the tables, the number of rejections from a Pearson curve fit are close indeed to the ideal number of rejections listed in Table 1. The cases of main concern ($n \leq 20$ and $\lambda < 3$) seem to be satisfactory except for the case $n=10$ and $\lambda=1$ where the critical values tend to reject too often. However, a definite improvement from the χ^2 approximation is clearly present for these cases. While the lower χ^2 critical values tend to be too conservative, the Pearson curve fit has corrected for this - however, it has over-corrected, as now, the lower critical values of the Pearson curves tend to be too liberal! This is apparent in all the significance levels considered.

Note how similar the tables obtained using exact and asymptotic moments are. This indicates that the asymptotic expressions for the moments, when used up to the fourth order, are fairly accurate. The excellent performance of the asymptotic moments is indeed encouraging for its use in applications. Even for n as low as 10, the asymptotic values of μ_2,μ_3 and μ_4 were correct to the first 4,3 and 2 decimal places, respectively.

PEARSON CURVE FIT WITH EXACT MOMENTS

Table 4A (α = 0.01)

	λ = 1		3		5		8	
n	I<L	I>U	I<L	I>U	I<L	I>U	I<L	I>U
10	109	72	59	77	69	74	76	59
20	67	64	71	82	70	75	69	76
50	87	63	85	84	79	84	82	79
100	64	75	84	63	79	65	79	67

Table 4B (α = 0.05)

	λ = 1		3		5		8	
n	I<L	I>U	I<L	I>U	I<L	I>U	I<L	I>U
10	587	410	384	399	384	376	391	365
20	401	407	384	372	364	362	381	360
50	401	371	383	397	383	371	384	352
100	389	393	366	380	379	368	380	364

Table 4C (α = 0.10)

	λ = 1		λ = 3		λ = 5		λ = 8	
n	I<L	I>U	I<L	I>U	I<L	I>U	I<L	I>U
10	596	821	758	747	780	760	790	744
20	682	785	786	778	769	792	779	763
50	751	796	761	770	733	762	774	755
100	759	763	728	749	703	735	724	730

Table 4D (α = 0.20)

	λ = 1		λ = 3		λ = 5		λ = 8	
n	I<L	I>U	I<L	I>U	I<L	I>U	I<L	I>U
10	1636	1514	1578	1510	1549	1519	1519	1504
20	1424	1447	1520	1534	1547	1537	1535	1543
50	1576	1561	1531	1538	1476	1511	1489	1505
100	1502	1489	1470	1477	1426	1511	1467	1502

PEARSON CURVE FIT WITH ASYMPTOTIC MOMENTS

Table 5A (α = 0.01)

	λ = 1		3		5		8	
n	I<L	I>U	I<L	I>U	I<L	I>U	I<L	I>U
10	44	72	59	77	77	75	87	59
20	67	61	71	82	70	75	69	76
50	84	64	85	84	79	84	82	81
100	71	74	84	63	79	65	79	67

Table 5B (α = 0.05)

	λ = 1		3		5		8	
n	I<L	I>U	I<L	I>U	I<L	I>U	I<L	I>U
10	384	410	384	399	397	374	405	369
20	401	407	384	372	364	361	381	360
50	401	381	383	397	383	371	386	352
100	394	384	366	384	383	368	379	364

Table 5C (α = 0.10)

	λ = 1		λ = 3		λ = 5		λ = 8	
n	I<L	I>U	I<L	I>U	I<L	I>U	I<L	I>U
10	596	821	758	747	780	756	806	744
20	683	785	786	778	763	792	782	765
50	744	796	761	770	733	762	778	755
100	759	761	728	744	703	734	730	733

Table 5D (α = 0.20)

	λ = 1		λ = 3		λ = 5		λ = 8	
n	I<L	I>U	I<L	I>U	I<L	I>U	I<L	I>U
10	1636	2233	1578	1510	1547	1516	1519	1500
20	1424	1447	1520	1534	1547	1537	1533	1542
50	1576	1573	1531	1538	1478	1511	1489	1505
100	1511	1489	1470	1468	1422	1507	1468	1501

The critical values obtained using exact and asymptotic moments may be found in Tables A3 and A4 respectively. With Pearson curve critical values now available, two issues come to mind:

(i) While the approximate values obtained from Pearson curves clearly improve upon those obtained with the χ^2 approximations for $n \leq 20$ and/or $\lambda \leq 3$, is it worthwhile going through the Pearson curve algorithm, computing asymptotic moments, determining the Pearson curve and then obtaining the critical values, as opposed to simply going to χ^2 table and reading off the critical values?

(ii) Are the Pearson curves still better when we replace λ by the maximum likelihood estimator $\hat{\lambda}=\bar{X}$?

No attempt was made to examine the second question, although for large sample sizes, we would expect that Pearson curves would still be better. In answer to the first question, if accuracy of critical values is of primary importance, then we might favor Pearson curves. The asymptotic expressions for the moments are now known to be accurate, and once the moments are computed from these expressions (to the fourth order), β_1 and β_2 are determined and the Biometrika Tables (Pearson and Hartley (1966)) provide us with the critical values. If on the other hand, the criterion κ given in (3.4) results in a not too uncommon distribution, then the critical values may be obtained from the computer. We reiterate that in this case, the explicit form of the density has to be derived. Alternatively, given the values of n and $\hat{\lambda}$, one can obtain critical values through interpolation from the tables provided in the appendix. Although the accuracy of interpolating from these tables has not been assessed, the Pearson curve algorithm is smooth and presumably, a simple linear interpolation will suffice.

4. THE GRAM-CHARLIER SERIES OF TYPE A

4.1 THE THEORY OF GRAM-CHARLIER EXPANSIONS

In mathematics, a typical procedure for studying the properties of a function is to express the function as an infinite series. Two types of series that immediately come to mind are Taylor series (or power series) and Fourier series. While these two series express a function as a sum of powers of a variable or as a sum of trigonometric functions, we will instead consider expanding the true density of I as a sum of derivatives of the standard normal density. One can then think of such an expansion as a correction to the normal approximation that was examined in Chapter 2.

Let $\phi(x)$ be the standard normal density.

$$\phi(x) = (1/\sqrt{2\pi})\ \exp(-x^2/2).$$

Then

$$\phi'(x) = -x\phi(x)$$

$$\phi''(x) = (x^2-1)\phi(x)$$

$$\phi^{(3)}(x) = -(x^3-3x)\phi(x)$$

$$\phi^{(4)}(x) = (x^4-6x^2+3)\phi(x)$$

In general,

$$\phi^{(j)}(x) = (-1)^j H_j(x)\phi(x) \quad \text{for } j = 0,1,2,\ldots \tag{4.1}$$

The polynomials $H_j(x)$ of degree j are called the <u>Tchebycheff-Hermite polynomials</u>. By convention, $\phi^{(0)}(x) = \phi(x)$, i.e. $H_0 \equiv 1$.

Some important properties of these polynomials are that

(i) $H_j'(x) = jH_{j-1}(x)$

(ii) $$\int_{-\infty}^{\infty} H_j(x)H_k(x)\phi(x)dx = \begin{cases} 0 \ , & k \neq j \\ k! \ , & k=j \end{cases},$$

i.e. the Tchebycheff-Hermite poloynomials are orthogonal.

(iii) $$\int_{-\infty}^{x} H_j(x)\phi(x)dx = -H_{j-1}(x)\phi(x).$$

The proof of (i) involves expanding

$$\phi(x-t) = \phi(x)\exp\{tx-(t^2/2)\}$$

in a Taylor series about $t = 0$. This yields the equation

$$\exp\{tx-(t^2/2)\} = \sum_{j=0}^{\infty} (t^j/j!)H_j(x).$$

Substituting in the series for the exponential term and redefining the index of the summation gives the desired result.

(ii) follows from (i) by substituting in the expression for $H_k(x)$ from (4.1) in terms of $\phi^{(k)}(x)$ and performing successive integration by parts. (iii) follows immediately from (4.1).

Suppose then that a density function $f(x)$ can be expanded in an infinite series of derivatives of $\phi(x)$:

$$f(x) = \sum_{j=0}^{\infty} c_j H_j(x)\phi(x). \tag{4.2}$$

The conditions for this series expansion to be valid can be found in a theorem by Cramer (1926). The conditions are that:

(i) $\int_{-\infty}^{\infty} (df/dx)\exp(-x^2/2)dx$ converges, and

(ii) $f(x) \longrightarrow 0$ as $x \longrightarrow \pm\infty$.

To find the coefficients c_j, multiply equation (4.2) by $H_k(x)$ and integrate from $-\infty$ to ∞ and use the orthogonality property.

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)H_k(x)\,dx &= \int_{-\infty}^{\infty} \sum_{j=0}^{\infty} c_j H_j(x)H_k(x)\phi(x)dx \\ &= \sum_{j=0}^{\infty} \int_{-\infty}^{\infty} c_j H_j(x)H_k(x)\phi(x)dx \\ &= c_k \cdot k!\end{aligned}$$

(interchanging the sum and the integral is justified since $p(x)\cdot\phi(x)$ is always bounded, for any polynomial $p(x)$ of finite degree.) The c_j's can then be expressed in terms of the moments about the origin. We list the first five coefficients below.

$c_0 = 1$

$c_1 = \mu$

$c_2 = (1/2)(\mu_2'-1)$

$$c_3 = (1/6)(\mu_3{}'-\mu)$$

$$c_4 = (1/24)(\mu_4{}'-6\mu_2{}'+3).$$

For the purpose of computing critical values, it is convenient to express the cumulative distribution function (CDF) $F(x)$, in a similar series.

$$F(x) = \int_{-\infty}^{x} f(x)\,dx$$

$$= \int_{-\infty}^{x}\{\phi(x) + \sum_{j=1}^{\infty} c_j H_j(x)\phi(x)\}\,dx$$

$$= \Phi(x) - \sum_{j=1}^{\infty} c_j H_{j-1}(x)\phi(x),$$ where Φ is the standard normal CDF. This series is called the Gram-Charlier series of type A (Kendall and Stuart (1958), vol. 1, pp. 156-157).

Let $X = (I-1)/\sqrt{\mu_2}$. Then $c_1 = c_2 = 0$, and the Gram-Charlier series of type A for the CDF of X is given by

$$F(x) = \Phi(x) - \phi(x)\{c_3H_2(x) + c_4H_3(x) + \ldots\}$$

where
$$\begin{aligned} c_3 &= (1/6)\mu_3{}' \\ &= (1/6)E\{(I-1)^3/\mu_2{}^{3/2}\} \\ &= (1/6)\mu_3/\mu_2{}^{3/2} \\ &= (1/6)\sqrt{\beta_1}\ . \end{aligned}$$

Similarly,

$$c_4 = (1/24)(\beta_2-3).$$

Now consider using partial sums of the Gram-Charlier series as an approximation to $F(x)$. (Note that the one-term approximation is merely the normal approximation.) In particular, suppose we use the first two terms of the series to approximate $F(x)$.

$$F(x) \simeq \Phi(x) - (1/6)(\sqrt{\beta_1})(x^2-1)\phi(x) = G(x)$$

Then the corresponding approximate critical values can be computed for any significance level α, by solving the equation $G(x) = \alpha$. Of course, this equation has to be solved iteratively (by the Newton-Raphson method, say).

Since exact and asymptotic moments are available, c_3 and c_4 can be computed likewise. Note that from (3.9) and (3.10), $c_3 \sim O(1/\sqrt{n})$ and $c_4 \sim O(1/n)$. If we choose to do the two-term approximation, then we would be neglecting c_4, and hence neglecting terms of $O(1/n)$. Therefore, c_3 can be approximated by terms whose orders are less than $1/n$. In this case,

$$6c_3 \sim \sqrt{2/n}\ \{2+ (1/\lambda)\}\ ,$$

and the approximation becomes

$$F(x) \sim \Phi(x) - (1/6)\sqrt{(2/n)}\{2+(1/\lambda)\}(x^2-1)\phi(x).$$

In the case of a three-term approximation, we would be neglecting c_5. The fifth moment is unavailable, but we might anticipate that $c_5 \sim O(1/n^{3/2})$. If this is the case, then up to the order of neglected terms,

$$6c_3 \sim \sqrt{(2/n)}\ \{2+(1/\lambda)\}$$

$$24c_4 \sim (2/n)\{6+(12/\lambda) + (1/\lambda^2)\},$$

and the approximation becomes

$$F(x) \sim \Phi(x) - \{(1/6\sqrt{(2/n)}\{2+(1/\lambda)\}(x^2-1) + (1/24)\{(2/n)[6+(12/\lambda)+(1/\lambda^2)](x^3-3x)\}\phi(x).$$

4.2 DISCUSSION

Since the results with the asymptotic moments were virtually the same as those with the exact moments (as it was with Pearson curves), we only display the table for the two- and three- term fits with exact moments (see Tables 6(A-D) and Tables 7(A-D)).

We can immediately see that the series expansion has improved the normal approximation. We point out some interesting results arising from the comparison of the two series approximation. First, while the lower critical values from the three-moment fit tend to be too conservative, those from the four-moment fit are slightly liberal. This appears to be the case of $\alpha \geq 0.05$. In general, the four-moment fit seems to be adequate at the lower tail, except for the usual cases of concern $n \leq 20$ and $\lambda \leq 3$. On the other hand, the upper critical values from the three-moment fit tend to be adequate for most cases, but the inclusion of the fourth moment has made the upper critical values very conservative. The four-moment fit is only satisfactory for $n \geq 50$ and $\lambda \geq 3$.

Obviously, the Gram-Charlier approximation is not recommended since the much simpler χ^2 approximation is even better. However, it is interesting to note that a three-term partial sum approximation of the true density of I improves the normal approximation considerably. The critical values obtained from this approximation may be

GRAM-CHARLIER THREE-MOMENT FIT (EXACT)

Table 6A (α = 0.01)

	λ = 1		3		5		8	
n	I<L	I>U	I<L	I>U	I<L	I>U	I<L	I>U
10	109	140	32	125	22	132	26	106
20	105	123	79	115	66	101	66	106
50	101	93	96	98	86	98	90	94
100	85	96	89	74	88	69	84	76

Table 6B (α = 0.05)

	λ = 1		3		5		8	
n	I<L	I>U	I<L	I>U	I<L	I>U	I<L	I>U
10	190	381	176	364	192	352	199	347
20	305	370	304	354	287	344	284	330
50	354	363	358	384	345	363	352	345
100	384	377	356	370	359	362	368	354

Table 6C (α = 0.10)

	λ 1		3		5		8	
n	I<L	I>U	I<L	I>U	I<L	I>U	I<L	I>U
10	587	551	482	626	497	628	515	626
20	537	678	659	687	638	709	633	680
50	668	726	723	731	683	721	727	720
100	735	733	707	727	678	713	703	714

Table 6D (α = 0.20)

	λ 1		3		5		8	
n	I<L	I>U	I<L	I>U	I<L	I>U	I<L	I>U
10	974	1083	1104	1291	1240	1328	1248	1312
20	1297	1350	1352	1426	1402	1404	1389	1448
50	1517	1471	1436	1472	1430	1450	1439	1453
100	1468	1455	1432	1442	1398	1485	1438	1486

GRAM-CHARLIER FOUR-MOMENT FIT (EXACT)

Table 7A ($\alpha = 0.01$)

	λ							
	1		3		5		8	
n	I<L	I>U	I<L	I>U	I<L	I>U	I<L	I>U
10	109	81	16	89	0	89	5	70
20	39	74	36	82	30	75	32	77
50	54	69	62	83	59	83	12	115
100	59	74	72	62	71	63	70	67

Table 7B (α = .0.05)

	λ							
	1		3		5		8	
n	I<L	I>U	I<L	I>U	I<L	I>U	I<L	I>U
10	587	212	305	262	275	259	267	264
20	401	246	317	293	303	284	293	283
50	358	303	350	355	341	335	228	429
100	377	342	349	359	358	344	365	342

Table 7C ($\alpha = 0.10$)

	λ = 1		3		5		8	
n	I<L	I>U	I<L	I>U	I<L	I>U	I<L	I>U
10	968	426	774	504	738	502	732	515
20	839	532	761	655	707	663	7 7	639
50	744	687	746	725	713	713	747	718
100	759	733	720	725	695	712	709	713

Table 7D ($\alpha = 0.20$)

	λ = 1		3		5		8	
n	I<L	I>U	I<L	I>U	I<L	I>U	I<L	I>U
10	2147	2233	1824	1733	1667	1695	1665	1630
20	1672	1678	1619	1587	1600	1579	1580	1589
50	1628	1598	1538	1554	1520	1522	1508	1515
100	1536	1499	1499	1483	1429	1514	1469	1504

found in Tables A5 - A8.

Other types of series expansions could also be examined. Two of the more common ones are Edgeworth expansions, which are known to be equivalent to the Gram-Charlier series of type A, and Fisher-Cornish expansions, which are derived from Edgeworth expansions. Treatment of these can be found in Kendall and Stuart (1958, Vol. 1, pp. 157 - 167).

5. THE LIKELIHOOD RATIO AND GOODNESS-OF-FIT TESTS

In the previous chapters, we examined various approximations to the distribution of the index of dispersion for the case that the data is distributed as Poisson in order to obtain approximate critical values. We compared the performances of critical values obtained from large sample approximations, series expansions and Pearson curve fits, and found that Pearson curves seemed to give the most accurate critical values.

The one remaining question we will attempt to answer is:

> "How good is the test based on the index of dispersion relative to other tests of the null hypothesis that the data is distributed as Poisson?"

Two well-known methods of testing the adequacy of the model under the null are

(i) The Likelihood Ratio Test and

(ii) Pearson's Goodness-of-Fit (GOF).

To assess the performance of the test based on the index of dispersion, we can examine the power of these three tests against appropriate alternatives. In testing for over-dispersion, ecologists have used the negative binomial (Fisher, 1941), Neymann's contagious distribution Type A (Neymann, 1939) and Thomas' double Poisson (Thomas, 1949) as alternatives to the Poisson distribution. P. Robinson (1954) has pointed out that the Neymann distribution may have several modes (leading to non-unique estimates when estimating by maximum likelihood) and that a basic assumption of the double Poisson may not be satisfied by the distribution of plant populations. The negative binomial distribution is perhaps the most widely applied alternative to the Poisson. Letting the parameters of the negative binomial be k and θ ($k>0, \theta>0$), we may write:

$$f(x,k,\theta) = \binom{k+x-1}{x} \{1/(\theta+1)\}^k \{\theta/(\theta+1)\}^x,$$

where $x=0,1,2,\ldots$. From this, we have that

$$E(X) = k\theta \text{ and}$$

$$\mathrm{Var}(X) = k\theta(1+\theta) = E(X)(1+\theta) > E(X).$$

For alternatives involving under-dispersion, we can test the null against the positive binomial (although it will be noted that the maximum likelihood estimator of n, the number of Bernoulli trials, may not be unique).

5.1 THE LIKELIHOOD RATIO TEST

Let $\underline{\eta} = (k,\theta)$ be a two-dimensional vector of parameters and let $f(x,\underline{\eta})$ be the probability mass function of the negative binomial. It is shown in Appendix A1.4 that as $k \to \infty$ and $\theta \to 0$ in such a way that $k\theta = \lambda$, a constant, then the limiting distribution arrived at is the Poisson with parameter λ which has probability mass function $f(x,\lambda)$. Let Θ_0 and Θ be the space of values that the parameters λ and $\underline{\eta}$ may take on, respectively. The GOF problem then is to test

$$H_0: \ \underline{\eta} \in \Theta_0$$

$$H_1: \ \underline{\eta} \in \Theta-\Theta_0$$

The likelihood ratio statistic for testing H_0 against H_1 is

$$\Lambda = \sup L(\underline{x},\eta)/\sup' L(\underline{x},\eta),$$

where $L(\underline{x},\eta)$ is the likelihood function for a sample $X_1,\ldots,X_n$. Here, "sup" indicates a supremum taken over Θ_0 while "sup' " indicates a supremum taken over Θ. Note that this implies that $\Lambda \le 1$.

In general, the distribution of the likelihood ratio statistic is unknown. However, under regularity conditions (Kendall and Stuart (1958), vol. 1, pp. 230-231), as $n \to \infty$, it is known that asymptotically

$$-2 \ln \Lambda \approx \chi^2_{p-q} \ ,$$

where p and q are the dimensionalities of the parameter spaces under the alternative and the null, respectively.

We now compute the MLE's of λ,k and θ. The likelihood functions of the Poisson and negative binomial are respectively,

$$L(x,\lambda) = \prod_{i=1}^{n} \lambda^{x_i} e^{-\lambda}/x_i!$$

$$= \lambda^{x.} e^{-n\lambda}/\prod_{i=1}^{n} x_i! \quad , \text{ and}$$

$$L(\underline{x},k,\theta) = \prod_{i=1}^{n} \binom{k+x_i-1}{x_i} \{\theta/(1+\theta)\}^{x_i}\{1/(1+\theta)\}^{k} . \tag{5.1}$$

Let $l_0(\underline{x},\lambda)$ and $l_1(\underline{x},k,\theta)$ be the corresponding log-likelihood functions. Then

$$l_0(\underline{x},\lambda) = x.\ \ln\lambda \ -n\lambda - \sum_{i=1}^{n} \ln(x_i!) \tag{5.2}$$

$$l_1(\underline{x},k,\theta) = \sum_{i=1}^{n} \ln\{(k+x_i-1)! \,/\, [x_i!\,(k-1)!]\} + x.\ \ln\theta$$

$$- (x. + nk)\ \ln\,(1+\theta)$$

$$= \sum_{i=1}^{n} \ln\,\{(k+x_i-1)! \,/\, (k-1)!\} - \sum_{i=1}^{n} \ln\,(x_i!)$$

$$+ x.\ \ln\,\theta - (x.+nk)\ \ln\,(1+\theta).$$

But

$$\sum_{i=1}^{n} \ln\,\{(k+x_i-1)!/(k-1)!\}$$

$$= [\sum_{\{i:x_i=0\}} + \sum_{\{i:x_i>0\}}]\ \ln\,\{(k+x_i-1)! \,/\, (k-1)!\}.$$

Since the summation over the zero values of x_i is zero,

$$\sum_{i=1}^{n} \ln\{(k+x_i-1)!/(k-1)!\} = \sum_i^+ \ln\{(k+x_i-1)!/(k-1)!\},$$

where $\sum_i^+$ denotes a summation over i such that $x_i>0$. This sum can be

written as,

$$\sum_i^+ \sum_{j=1}^{x_i} \ln(k+j-1),$$

and hence,

$$l_1(\underline{x},k,\theta) = x.\ln\theta - (x.+nk)\ln(1+\theta) + \sum_i^+ \sum_{j=1}^{x_i} \ln(k+j-1)$$

$$- \sum_{i=1}^{n} \ln(x_i!). \tag{5.3}$$

To obtain the MLE of λ, (5.2) must be maximized with respect to λ while the MLE's of θ and k are obtained by maximizing (5.3) with respect to θ and k simultaneously. Thus,

$$\partial l_0/\partial\lambda = (x./\lambda) - n .$$

Setting this derivative to zero and solving for λ yields $\hat{\lambda}$, the MLE of λ as,

$$\hat{\lambda} = \bar{X}. \tag{5.4}$$

Similarly,

$$\partial l_1/\partial\theta = (x./\theta) - \{(x.+nk)/(1+\theta)\}$$

$$\partial l_1/\partial k = \sum_i^+ \sum_{j=1}^{x_i} \{1/(k+j-1)\} - n \ln(1+\theta),$$

Setting the derivatives to zero, the first equation can be explicitly solved for θ to yield

$$\theta = \bar{X}/k. \tag{5.5}$$

Substituting this value into the second equation leads to

$$\sum_{i}^{+} \sum_{j=1}^{X_i} \{1/(k+j-1)\} - n \ln\{1+(\bar{X}/k)\} = 0. \tag{5.6}$$

Levin and Reeds (1978) give a necessary and sufficient condition for the uniqueness of the MLE of k. This criterion can be stated as:

"$\hat{k}$, the MLE of k, exists uniquely in $[0,\infty)$ if and only if

$$\sum_{i=1}^{n} X_i^2 - \sum_{i=1}^{n} X_i > (\sum_{i=1}^{n} X_i)^2/n. \tag{5.7}$$

The right-hand side of this criterion is simply $n\bar{X}^2$, and so (5.7) can be rewritten as

$$\sum_{i=1}^{n} (X_i-\bar{X})^2 > \sum_{i=1}^{n} X_i \ , \quad \text{or}$$

$$(1/n) \sum_{i=1}^{n} (X_i-\bar{X})^2 > \bar{X}.$$

Since $S^2 = \{1/(n-1)\} \sum_{i=1}^{n} (X_i-\bar{X})^2 > (1/n) \sum_{i=1}^{n} (X_i-\bar{X})^2$

provided that $\sum_{i=1}^{n} (X_i-\bar{X})^2 > 0$, a consequence of Levin and Reed's criterion is that a unique $\hat{k}$ exists in $[0,\infty)$ if the index of dispersion is greater than 1.

Thus, subject to Levin and Reeds' criterion, the solution to (5.6) can be obtained numerically. Once $\hat{k}$, the MLE of k, is obtained, substitution into (5.5) yields $\hat{\theta}$, the MLE of θ. (The case where the criterion is not satisfied corresponds to $\hat{k}=\infty$, and is discussed in more detail in section 5.4.)

Continuing with the likelihood ratio test, we have

$$\begin{aligned}\ln \Lambda &= l_0(\underline{x}, \hat{\lambda}) - l_1(\underline{x},\hat{k},\hat{\theta}) \\ &= x.\{(\ln \hat{k})-1\} + (x.+n\hat{k})\ln\{1+(\bar{X}/\hat{k})\} - \sum_i^+ \sum_{j=1}^{x_i} \ln (\hat{k}+j-1).\end{aligned}$$

This is the form of the likelihood ratio test. The asymptotic result is that as $n \to \infty$

$$-2 \ln \Lambda \approx \chi_1^2.$$

5.2 PEARSON'S GOODNESS-OF-FIT TEST

A test that assesses goodness-of-fit is the well-known χ^2 test that was proposed by Karl Pearson (1900). The GOF statistic is:

$$X^2 = \sum_{j=0}^{\infty} (n_j-\lambda_j)^2/\lambda_j,$$

where n_j is the number of times that the integer j is observed in a sample and $\lambda_j = nP(X=j)$ where $X \sim P(\lambda)$, is the expected number of times the integer j will occur under the null.

The asymptotic result is that as $n \to \infty$,

$$X^2 \approx \chi_{\nu-1}^2$$

where ν is the number of cells. (Note that one degree of freedom is lost since the probabilities computed uner the null are subject to the constraint that they sum up to 1. Also, in the simulation that follows, the value of λ is specified and hence no further degree of freedom is lost.)

This approximation has been known to work well particularly if the expected number of observations, λ_j, in each cell is at least 5. Now, for sample sizes of about 10 to 20, this rule of thumb may not always be satisfied. The rule that has been implemented is that the expected number of observations in each cell is at least 3. As will be seen, the χ^2 approximation was still satisfactory in this case.

5.3 POWER COMPUTATIONS

We now have three tests whose power we wish to compare. Since the index of dispersion and the test based on X^2 do not depend on explicit alternative hypotheses, we might expect the likelihood ratio test to be superior of the three. Because of computational difficulties that may arise when using the likelihood ratio test (these are mentioned later on in this section), it is not recommended for use in practice. We use it here only to provide a baseline for the assessment of the power of the test based on the index of dispersion. Since the index of dispersion is devised to test for the variance being different from the mean, it will be geared towards alternative hypotheses which have this property and so we might expect the index of dispersion to perform better than the test based on X^2. Note that while a one-sided test was implemented for the test based on the index of dispersion (and necessarily for the likelihood ratio test), the test based on X^2 is necessarily two-sided. This should be taken into consideration when comparing the power of the tests.

Let us recall the hypotheses we are testing:

$$H_0: \quad X_1, X_2, \ldots, X_n \sim P(\lambda)$$
$$H_1: \quad X_1, X_2, \ldots, X_n \sim NB(k, \theta).$$

Through simulation studies, we are going to compare the power of the three tests. However, as the null hypothesis does not specify a particular value of λ, it is not clear how to choose k and θ for the

simulation. To do this, we argue as follows:
Since the Poisson distribution with parameter λ is a limiting case of the negative binomial with parameters k and θ, we can specify λ and choose k and θ so that $k\theta = \lambda$. Now we also want to choose k so that the tests exhibit reasonable power. For instance, we do not wish to generate data that yields power that is very close to 1. We would like to choose values of k so that the range of the power covers the unit interval [0,1].

To get a good idea of what k should roughly be, we can examine the asymptotic power of the index of dispersion test. To do this, we need the first four moments of the negative binomial which can be found in Kendall and Stuart (1958, vol. 1, p. 131). Letting ν, ν_2, ν_3, and ν_4 denote the central moments of the negative binomial, we have

$$\nu = k\theta,$$

$$\nu_2 = k\theta(\theta+1),$$

$$\nu_3 = k\theta(\theta+1)(2\theta+1), \text{ and}$$

$$\nu_4 = k\theta(\theta+1)(1+6\theta+6\theta^2+3k\theta+3k\theta^2).$$

Substituting these moments into equation (2.6) in Chapter 2, we have that

$$I \approx N(1+\theta\ ,\ (1/n)v^2)$$

where $v^2 = 2(1+\theta)^2 + (1+\theta)(2+3\theta)/k$. Notice that by setting $k=\lambda/\theta$ and letting $\theta \to 0$, we obtain the aysmptotic null distribution of the index of dispersion for the Poisson case, namely

$$I \approx N(1,2/n).$$

This result was seen in Chapter 2 where the performance of the asymptotic normal critical values was assessed. Hence the asypmtotic power of I can be computed from the set of hypotheses:

$$H_o: \theta = 0$$

$$H_1: \theta = \theta_1 > 0 .$$

Let $\mu(\theta) = 1+\theta$ and $\sigma^2(\theta) = (1/n)v^2$, where v^2 is defined as above. If we let $I_\alpha = 1+z_\alpha\sqrt{(2/n)}$, where z_α is the upper critical value of the standard normal distribution at significance level α, then the asymptotic power of I is

$$\text{Power} = \Phi(-w_\alpha), \text{ where } w_\alpha = \{I_\alpha - \mu(\theta_1)\}/\sqrt{\sigma^2(\theta_1)}.$$

The asymptotic power of I is presented in table 8 for the case $n = 20$ and $\alpha = 0.05$.

Table 8: ASYMPTOTIC POWER OF THE INDEX OF DISPERSION TEST ($n = 20$, $\alpha = 0.05$)

			k		
$\lambda = k\theta$	3	5	7	10	SIZE
1	.352	.224	.166	.125	.05
3	.739	.560	.425	.309	.05
5	.873	.752	.663	.484	.05

Thus, the values of $k = 3,5,7$ and 10 seem to be adequate. Notice the pattern in this table. The power decreases with increasing k and decreasing θ. This is not surprising because as k increases and θ decreases, the negative binomial approaches the Poisson, and hence it would be much harder to detect differences between the null and the alternative with k large and θ small.

We now proceed with the Monte Carlo simulation. A total of 500 samples of $n = 10$, 20 and 50 negative binomial random variables were generated for $k = 3$, 5, 7 and 10. The three statistics , $-2 \ln\Lambda$, I and X^2 were computed using the negative binomial data. While the computation of I and X^2 are very easy on the computer, some problems may occur in computing the likelihood ratio statistic, as mentioned previously. First, the computation of the double sum in (5.6) at each iteration will increase the cost of running the computer program. This will be more evident for large n and/or large λ. Second, for some of the samples, a negative value of $\hat{k}$ was obtained at some point in the iteration process. This may create a problem in computing $\ln \{1+(\bar{X}/k)\}$ in (5.6). Barring all difficulties however, the Newton-Raphson Algorithm achieved convergence in about 5 or 6 iterations.

Continuing with the simulation, an attempt is made to treat each test as equal as possible by using χ^2 critical values in each case. However, a problem may still occur in the power comparison. Since all these tests are based on asymptotic critical values, the asymptotic approximations may not treat each of the three tests exactly the same.

For example, for a given sample size, it may be that X^2 is better approximated by χ^2 than I, which in turn may be better approximated by χ^2 than $-2 \ln \Lambda$. This may make the conclusions on the power comparison unreliable.

Proceeding with the power computations, the number of statistics which fell in the rejection region were counted for each of the three tests (Recall that a one-sided rejection region was formed for the tests based on Λ and I while a two-sided rejection region is necessary for X^2).

The power of each test is displayed in tables 9-10 (A-D). Each cell in this table contains the power of the likelihood ratio test, the index of dispersion test and the GOF test, in that order. To provide a handle on the accuracy of the χ^2 approximation for each of the three tests, the estimated size of each test is also displayed in each table. If the approximation were good for a particular test, then the estimated size of that test should be close to the specified significance level.

As mentioned above, the χ^2 approximation may not treat these three tests equally. This in fact is the case when n = 10. The critical values for the likelihood ratio test are too conservative as can be seen from the estimated size of the test. For example, when $\alpha = 0.05$, the estimated size of the likelihood ratio test is slightly less than 0.01. On the other hand, the estimated size of the test based on X^2 is very close to the true significance level for all α, while the test based on the index of dispersion tends to be intermediate. Thus we could infer that if exact

POWER OF TESTS BASED ON Λ, I AND X^2 (n=10)

Table 9A: α = 0.01

	k				
$\lambda=k\theta$	3	5	7	10	SIZE
1	.028 .068 .010	.016 .044 .008	.012 .032 .006	.008 .028 .006	0 .008 .008
3	.148 .252 .156	.058 .138 .096	.026 .086 .086	.018 .056 .072	0 .002 .010
5	.356 .478 .290	.148 .276 .192	.086 .180 .128	.042 .114 .102	.002 .004 .024

Table 9B: α = 0.05

	k				
$\lambda=k\theta$	3	5	7	10	SIZE
1	.064 .162 .062	.048 .114 .068	.038 .092 .060	.032 .074 .054	.008 .034 .056
3	.270 .444 .242	.144 .272 .174	.098 .224 .136	.058 .152 .112	0 .038 .050
5	.486 .648 .388	.286 .464 .268	.182 .332 .208	.122 .252 .166	.006 .036 .068

Table 9C: α = 0.10

$\lambda=k\theta$	k = 3	5	7	10	SIZE
1	.118 .222 .086	.066 .174 .096	.052 .130 .092	.048 .110 .088	.020 .046 .094
3	.344 .566 .272	.198 .384 .204	.150 .310 .156	.098 .242 .136	.012 .070 .090
5	.570 .752 .516	.752 .588 .384	.250 .448 .312	.178 .342 .264	.006 .082 .136

Table 9D: α = 0.20

$\lambda=k\theta$	k = 3	5	7	10	SIZE
1	.166 .366 .224	.108 .294 .206	.086 .256 .200	.072 .236 .196	.036 .124 .182
3	.452 .688 .424	.286 .546 .332	.232 .466 .266	.178 .398 .244	.040 .158 .184
5	.656 .848 .582	.470 .718 .456	.356 .618 .382	.258 .504 .326	.040 .160 .216

POWER OF TESTS BASED ON Λ, I AND X^2 (n=20)

Table 10A: $\alpha = 0.01$

$\lambda=k\theta$	k: 3	5	7	10	SIZE
1	.044 .102 .048	.016 .050 .038	.012 .034 .030	.010 .024 .030	0 .010 .016
3	.314 .450 .266	.142 .228 .136	.076 .150 .096	.034 .098 .062	.002 .012 .024
5	.664 .770 .558	.334 .472 .296	.190 .312 .202	.112 .190 .140	0 .008 .022

Table 10B: $\alpha = 0.05$

$\lambda=k\theta$	k: 3	5	7	10	SIZE
1	.104 .214 .078	.050 .140 .056	.042 .100 .052	.032 .082 .050	.010 .046 .040
3	.502 .666 .410	.258 .416 .266	.164 .298 .206	.104 .218 .156	.012 .032 .056
5	.804 .898 .686	.534 .706 .408	.342 .526 .314	.208 .366 .224	.008 .042 .074

Table 10C: α = 0.10

	k				
λ=kθ	3	5	7	10	SIZE
1	.172 .316 .148	.104 .228 .128	.076 .174 .110	.054 .148 .104	.020 .086 .098
3	.594 .780 .498	.336 .544 .328	.232 .420 .268	.158 .320 .210	.020 .086 .110
5	.872 .936 .760	.620 .796 .514	.452 .650 .408	.290 .486 .298	.020 .086 .112

Table 10D: α = 0.20

	k				
λ=kθ	3	5	7	10	SIZE
1	.246 .492 .226	.164 .388 .202	.124 .328 .174	.104 .272 .174	.054 .172 .166
3	.704 .882 .614	.460 .714 .462	.326 .604 .382	.246 .476 .308	.032 .196 .194
5	.916 .966 .834	.746 .886 .652	.578 .792 .542	.396 .662 .422	.046 .192 .220

critical values were employed, the power of the likelihood ratio test would be considerably larger than indicated in the tables. Similarly, the critical values for the index of dispersion test are slightly conservative and hence we would expect the power of this test to increase if exact critical values were used. The power of the test based on X^2, however, would be pretty much what the tables indicate.

Turning to n = 20, the same problem still arises for the likelihood ratio -- very conservative critical values. On the other hand, while the asymptotic critical values used for the index of dispersion are still slightly conservative, the approximation has clearly improved and the estimated size of the index of dispersion test is closer to the true significance level. In fact, the estimated power of the index of dispersion is close indeed to its asymptotic power.

Although it is not clear that the likelihood ratio test is more powerful than the test based on the index of dispersion, we can make one additional observation if we compare tables with the same size (for instance, the 20% table for the likelihood ratio and the 5% table for the index of dispersion when n = 20), we see that in each cell, the estimated power of the likelihood ratio test is indeed higher than that of the index of dispersion - however, only marginally. This is an indication of what we might expect to see if the sample size were large enough so that the estimated size of the test is close to the true significance level.

Thus, the results displayed in tables 9 and 10 seem to suggest the following order in terms of the power of each test.

Likelihood Ratio, Index of Dispersion and GOF based on X^2.

A further attempt to compare the power of the likelihood ratio and the index of dispersion, tests are displayed in table 11 (A-D) for a sample size of n = 50. As before, we may compare the 20% table for the likelihood ratio with the 5% table for the index of dispersion to conclude that the likelihood ratio test is only slightly more powerful than the test based on the index of dispersion.

5.4 THE LIKELIHOOD RATIO TEST REVISITED

At the time when this thesis was first being written, no obvious explanation could be made about the conservatism of the critical values of the likelihood ratio test. Subsequently, the explanation became clear: For the situation under consideration, the null distribution of $-2\ln\Lambda$ does not converge to that of a χ^2_1, but rather to that of a mixture of a χ^2_1 and a zero random variable, each with probability 1/2. This is an example of the general results of Chernoff (1954). The reasoning goes as follows:

If the MLE $(\hat{\theta},\hat{k})$ for the negative binomial occurs at $\hat{k}=\infty$, then $\Lambda\equiv 1$ and $-2\ln\Lambda\equiv 0$. Levin and Reeds (1977) have established that this occurs if and only if $(n-1)S^2 \le n\bar{X}$. Thus, under the null hypothesis, we have

$$\begin{aligned}
P(-2\ln\Lambda \equiv 0) &= P\{(n-1)S^2 \le n\bar{X}\} \\
&= P\{n(S^2-\bar{X}) - S^2 \le 0\} \\
&= P\{\sqrt{n}[(S^2-\lambda) - (\bar{X}-\lambda)] - (1/\sqrt{n})S^2 \le 0\} \\
&\simeq P\{\sqrt{n}[(S^2-\lambda) - (\bar{X}-\lambda)] \le 0\}, \text{ for large } n.
\end{aligned}$$

POWER OF TESTS BASED ON Λ AND I (n=50)

Table 11A: α = 0.01

$\lambda=k\theta$	k: 3	5	7	10	SIZE
1	.094	.034	.016	.010	.004
	.166	.068	.050	.032	.012
3	.744	.376	.172	.082	.002
	.842	.510	.304	.152	.010
5	.984	.762	.500	.270	.002
	.990	.844	.636	.400	.008

Table 11B: α = 0.05

$\lambda=k\theta$	k: 3	5	7	10	SIZE
1	.224	.098	.060	.034	.014
	.354	.222	.156	.106	.046
3	.886	.572	.374	.206	.018
	.936	.724	.532	.354	.046
5	.996	.890	.714	.452	.014
	.998	.950	.818	.622	.040

Table 11C: α = 0.10

			k		
$\lambda=k\theta$	3	5	7	10	SIZE
1	.314	.178	.120	.080	.028
	.522	.318	.234	.178	.106
3	.922	.668	.482	.296	.032
	.960	.810	.642	.484	.096
5	.998	.942	.778	.564	.034
	.998	.978	.896	.722	.098

Table 11D: α = 0.20

			k		
$\lambda=k\theta$	3	5	7	10	SIZE
1	.430	.268	.198	.152	.066
	.664	.484	.380	.308	.216
3	.950	.770	.600	.422	.076
	.984	.910	.784	.638	.192
5	.998	.968	.876	.676	.070
	1.000	.988	.956	.848	.188

From sections 2.1 and 2.2, we have that

$$\begin{bmatrix} \sqrt{n}(\bar{X}-\lambda) \\ \sqrt{n}(S^2-\lambda) \end{bmatrix} \xrightarrow{d} N(\underline{0},\Sigma),$$

where $\Sigma = \begin{bmatrix} \lambda & \lambda \\ \lambda & \lambda + 2\lambda^2 \end{bmatrix}$.

Letting $f(x,y) = y-x$ so that $f(\bar{X},S^2) = S^2-\bar{X} = (S^2-\lambda) - (\bar{X}-\lambda)$, we have, as a consequence of the delta method, that

$$\sqrt{n}\{(S^2-\lambda) - (\bar{X}-\lambda)\} \xrightarrow{d} N(0,2\lambda^2).$$

Thus for large n, we have that

$$P(-2 \ln \Lambda \equiv 0) \simeq 1/2;$$

i.e. $-2 \ln \Lambda \equiv 0$ approximately half the time. Thus under the null, we have the following result:

$$-2 \ln \Lambda \xrightarrow{d} \begin{cases} 0 \text{ with probability } 1/2 \\ \chi^2_1 \text{ with probability } 1/2 \end{cases}. \tag{5.8}$$

As a supplement to (5.8), the first 500 Poisson samples from the 15,000 previously generated were again used in order to check if half of these 500 samples would give a value of $-2 \ln \Lambda = 0$. Table 12 displays the number of samples out of the 500 which led to $-2 \ln \Lambda = 0$.

Table 12: Number of Times $(n-1)S^2 \leq n\bar{X}$

	n		
λ	10	20	50
1	367	337	304
3	340	317	302
5	338	307	293

The fact that the entries in the table decrease as n gets large is indeed encouraging and re-affirms our position that the null distribution of -2 lnΛ converges to a mixture of distributions.

What effect then does (5.8) have on the power computations? In the previous computations, we have been assuming that

$$\alpha \simeq P(-2 \ln \Lambda \geq C_\alpha),$$

where C_α is the upper critical value corresponding to a χ_1^2 distribution. Letting Z be a standard normal random variable, we have instead that

$$\begin{aligned} P\{-2 \ln \Lambda \geq C_\alpha\} &\simeq (1/2)P\{I_0 \geq C_\alpha\} + (1/2)P\{Z^2 \geq C_\alpha\} \\ &= (1/2)\{1-P[Z^2 < C_\alpha]\} \\ &= 1-\Phi(\sqrt{C_\alpha}), \end{aligned}$$

where $I_0 = 0$ with probability 1 and Φ is the standard normal CDF. But,

$$\begin{aligned} \alpha &= P(Z^2 \geq C_\alpha) \\ &= 1 - P(-\sqrt{C_\alpha} < Z < \sqrt{C_\alpha}) \\ &= 2\{1-\Phi(\sqrt{C_\alpha}) , \end{aligned}$$

and hence,

$$P(-2 \ln \Lambda \geq C_\alpha) \simeq \alpha/2,$$

instead of the anticipated value of α!

Further simulations were not done as enough information can be gathered from the previous results. In particular, using the correct asymptotic critical values for the likelihood ratio test, for each fixed n, the previous results for $\alpha = 0.10$ are the appropriate results for $\alpha = 0.05$ and the previous results for $\alpha = 0.20$ are the appropriate results for $\alpha = 0.10$. These are displayed in Tables 13,14 and 15 (A-B).

POWER OF TESTS BASED ON Λ,I AND X^2 (n=10)

Table 13A: α = 0.05

			k		
λ=kθ	3	5	7	10	SIZE
1	.118 .162 .062	.066 .114 .068	.052 .092 .060	.048 .074 .054	.020 .034 .056
3	.344 .444 .242	.198 .272 .174	.150 .224 .136	.098 .152 .112	.012 .038 .050
5	.570 .648 .388	.752 .464 .268	.250 .332 .208	.178 .252 .166	.006 .036 .068

Table 13B: α = 0.10

			k		
λ=kθ	3	5	7	10	SIZE
1	.166 .222 .086	.108 .174 .096	.086 .130 .092	.072 .110 .088	.036 .046 .094
3	.452 .566 .272	.286 .384 .204	.232 .310 .156	.178 .242 .136	.040 .070 .090
3	.656 .752 .516	.470 .588 .384	.356 .448 .312	.258 .342 .264	.040 .082 .136

POWER OF TESTS BASED ON Λ,I AND X^2 (n=20)

Table 14A: $\alpha = 0.05$

	k				
$\lambda=k\theta$	3	5	7	10	SIZE
1	.172 .214 .078	.104 .140 .056	.076 .100 .052	.054 .082 .050	.020 .046 .040
3	.594 .666 .410	.336 .416 .266	.232 .298 .206	.158 .218 .156	.020 .032 .056
5	.872 .898 .686	.620 .706 .408	.452 .526 .314	.290 .366 .224	.020 .042 .074

Table 14B: $\alpha = 0.10$

	k				
$\lambda=k\theta$	3	5	7	10	SIZE
1	.246 .316 .148	.164 .228 .128	.124 .174 .110	.104 .148 .104	.054 .086 .098
3	.704 .780 .498	.460 .544 .328	.326 .420 .268	.246 .320 .210	.032 .086 .110
5	.916 .936 .760	.746 .796 .514	.578 .650 .408	.396 .486 .298	.046 .086 .112

POWER OF TESTS BASED ON Λ AND I (n=50)

Table 15A: $\alpha = 0.05$

	k				
$\lambda=k\theta$	3	5	7	10	SIZE
1	.314 .354	.178 .222	.120 .156	.080 .106	.028 .046
3	.922 .936	.668 .724	.482 .532	.296 .354	.032 .046
5	.998 .998	.942 .950	.778 .818	.564 .622	.034 .040

Table 15B: $\alpha = 0.10$

	k				
$\lambda=k\theta$	3	5	7	10	SIZE
1	.430 .522	.268 .318	.198 .234	.152 .178	.066 .106
3	.950 .960	.770 .810	.600 .642	.422 .484	.076 .096
5	.998 .998	.968 .978	.876 .896	.676 .722	.070 .098

The correction to the asymptotic null distribution of $-2 \ln \Lambda$ has certainly created a better picture. The estimated size of the likelihood ratio test is closer to the nominal significance level than it was when the χ^2_1 approximation was employed. However, the critical values for the likelihood ratio test are still very conservative. Hence, the power of the likelihood ratio test would be greater than that displayed in these tables.

As before, we may compare tables with approximately the same estimated size. For example the power of the test based on the index of dispersion from Table 13A might be compared to the power of the likelihood ratio test from Table 13B and similarly, Table 14A to 14B. We see that for n=10 and 20, the likelihood ratio test is only marginally better than the test based on the index of dispersion. For the case n=50, no reasonable comparison can be made, but we would expect the same behavior from both tests.

6. CONCLUSIONS

As mentioned in Chapter 1, the index of dispersion is a statistic often used to detect departures from randomness. As the null distribution of the index of dispersion is unknown, large sample approximations were used as a preliminary fit. The asymptotic null distribution of I was seen to be normal with mean 1 and variance 2/n. Asymptotic critical values from this distribution were then employed and assessed by a Monte Carlo simulation. The results were that the normal approximation was very poor for sample sizes typically encountered in practice and that this approximation only becomes satisfactory for a sample size of about 100 and $\lambda > 5$. A further attempt to improve the normal approximation was made by an infinite series expansion of of the true null distribution of I. We saw that a three-moment fit from the Gram-Charlier expansion improved the normal approximation enormously, but that this approximation was only satisfactory for $n \geq 50$.

The χ^2 approximation on the other hand seemed to be fairly accurate for $n>20$ and $\lambda>3$. This is certainly encouraging because of one important reason - the χ^2 approximation is simple to apply.

To further improve the χ^2 approximation (particularly for the cases $n \leq 20$ and $\lambda \leq 3$), Pearson curves were utilized. We found that except for the case $n=10$ and $\lambda=1$, Pearson curves definitely improved the approximation.

Two issues still remain unanswered:

(i) What should be done in the case $n=10$ and $\lambda=1$?

(ii) How well will the approximations remain when we replace λ by $\hat{\lambda}=\bar{X}$?

For the second question, we expect that the Pearson curve approximation will still perform well. As for the first question, let us keep in mind the suggestion put forth by Fisher (1950) and Cochran (1936) -- that the test based on the index of dispersion should be carried out conditionally, particularly when the Poisson parameter λ is small, for then exact frequencies can be computed.

Finally, the comparison of the powers of the tests based on the likelihood ratio, the index of dispersion and Pearson's X^2 statistic showed that the test based on the index of dispersion exhibits reasonable power when the hypothesis of randomness is tested against overdispersion. This supplements the results obtained by Perry and Mead (1979).

From the basis of accurate critical values and reasonably high power, we conclude that the index of dispersion is highly recommendable for its use in applications.

REFERENCES

ANDERSON, T.W. (1958). "An Introduction to Multivariate Statistical Analysis". Wiley, New York.

BATEMAN, G.I. (1950). "The Power of the χ^2 Index of Dispersion Test When Neyman's Contagious Distribution is the Alternative Hypothesis". Biometrika, 37, 59-63.

BLACKMAN, G.E. (1935). "A Study by Statistical Methods of the Distribution of Species in Grassland Communities". Annals of Botany, N.S., 49, 749-777.

CHERNOFF, H. (1954). "On the Distribution of the Likelihood Ratio". Annals of Mathematical Statistics, 25, 573-578.

CLAPHAM, A.R. (1936). "Over-dispersion in Grassland Communities and the Use of Statistical Methods in Ecology". Journal of Ecology, 24, 232-251.

COCHRAN, W.G. (1936). "The χ^2 Distribution for the Binomial and Poisson Series with Small Expectations". Annals of Eugenics, 7, 207-217.

CRAMER, H. (1926). "On Some Classes of Series Used in Mathematical Statistics". Skandinairske Matematikercongres, Copenhagen.

CRAMER, H. (1946). "Methods of Statistics". Princeton University Press, Princeton.

DARWIN, J.H. (1957). "The Power of the Poisson Index of Dispersion". Biomterika, 44, 286-289.

DAVID, F.N. and MOORE, P.G. (1954). "Notes on Contagious Distributions in Plant Populations". Annals of Botany, N.S., 28, 47-53.

ELDERTON, W.P. and JOHNSON, N.L. (1969). "Systems of Frequency Curves". Cambridge University.

FISHER, R.A. (1941). "The Negative Binomial Distribution". Annals of Eugenics , 11, 182-187.

FISHER, R.A. (1950). "The Significance of Deviations from Expectation in a Poisson Series". Biometrics, 6, 17-24.

FISHER, R.A., THORNTON, H.G. and MACKENZIE, W.A. (1922). "The Accuracy of the Plating Method of Estimating Bacterial Populations". Annals of Applied Biology, 9, 325.

GRAD, A. and SOLOMON, H. (1955). "Distribution of Quadratic Forms and Some Applications". Annals of Mathematical Statistics, 26, 464-477.

GREEN, R.H. (1966). "Measurement of Non-Randomness in Spatial Distributions". Res. Popul. Ecol., 8, 1-7.

HALDANE, J.B.S. (1937). "The Exact Value and Moments of the Distribution of χ^2, Used as a Test of Goodness-of-Fit, When Expectations are Small". Biometrika, 29, 133-143.

HALDANE, J.B.S. (1939). "The Mean and Variance of χ^2, When Used as a Test of Homogeneity, When Expectations are Small". Biometrika, 31, 419-355.

HOADLEY, A.B. (1968). "Use of the Pearson Densities for Approximating a Skew Density Whose Left Terminal and First Three Moments are Known". Biometrika, 55, 559-563.

HOEL, P.G. (1943). "On Indices of Dispersion". Annals of Mathematical Statistics, 14, 155-162.

KATHIRGAMATAMBY, N. (1953). "Notes on the Poisson Index of Dispersion". Biometrika, 40, 225-228.

KENDALL, M.G. and STUART, A. (1958). "The Advanced Theory of Statistics". Vol. 1. Griffin, London.

KENDALL, M.G. and STUART, A. (1958). "The Advanced Theory of Statistics", Vol. 2. Griffin, London.

LANCASTER, H.O. (1952). "Statistical Control of Counting Experiments". Biometrika, 39, 419-422.

LEVIN, B. and REEDS, J. (1977). "Compound Multinomial Likelihood Functions are Unimodal: Proof of a Conjecture of I.J. Good". Annals of Statistics, 5, 79-87.

MENDENHALL and SCHEAFFER (1973). "Mathematical Statistics with Applications". Duxbury Press, North Scituate, Massachusetts.

MULLER, P.H. and VAHL, H. (1976). "Pearson's System of Frequency Curves Whose Left Boundary and First Three Moments are Known". Biometrika, 54, 649-656.

NEYMAN, J. (1939). "On a New Class of Contagious Distributions Applicable in Entomology and Bacteriology". Annals of Mathematical Statistics, 10, 35-57.

PEARSON, E.S. and HARTLEY, H.O. (1966). "Biometrika Tables for Statisticians", Vol. 1, 3rd ed., Cambridge.

PEARSON, K. (1900). "On a Criterion that a Given System of Deviation from the Probable in the Case of a Correlated System of Variable is Such that it can be Reasonably Supposed to Have Risen in Random Sampling", Phil. Mag., (5), 50, 157

PEARSON, K. (1901). "Systematic Fitting of Curves to Observations". Biometrika, 1, 265.

PERRY, J.N. and MEAD, R. (1979). "On the Power of the Index of Dispersion Test to Detect Spatial Pattern". Biometrics, 35, 613-622.

POLYA, G. (1930). "Sur Quelques Points de la Theorie des Probabilites". Ann. de L'Inst. Henri Poincare, 1, 117-162.

ROBINSON, P. (1954). "The Distribution of Plant Populations". Annals of Botany, N.S. 18, 35-45.

SOLOMON, H. (1960). "Distributions of Quadratic Forms - Tables and Applications". Technical Report No. 45, Applied Mathematics and Statistics Laboratories, Stanford University, Stanford, California.

SOLOMON, H. and STEPHENS, M.A. (1978). "Approximations to Density Functions Using Pearson Curves". Journal of the American Statistical Association, 73, 153-160.

STUDENT (1919). "An Explanation of Deviation from Poisson's Law in Practice". Biometrika, 12, 211-215.

THOMAS, M. (1949). "A Generalization of Poisson's Binomial Limit for Use in Ecology". Biomtrika, 36, 18.

APPENDIX

A1.1 THE CONDITIONAL DISTRIBUTION OF A POISSON SAMPLE GIVEN THE TOTAL

Let $X_1,\ldots,X_n$ be independent identically distributed Poisson random variables with parameter λ. Then the sum of the X_i's

$$X. = \sum_{i=1}^{n} X_i ,$$

is distributed as Poisson with parameter $n\lambda$.

Consider the joint distribution of $X_1,\ldots,X_n$ given the total $X.$. Since

$$f_{\underline{X}|X.}(\underline{x}) = f_{\underline{X}}(\underline{x})/f_{X.}(x.)$$

$$= (\prod_{i=1}^{n} \lambda^{x_i} e^{-\lambda}/x_i!)/\{(n\lambda)^{x.} e^{-n\lambda}/(x.)!\}$$

$$= (x.!)(1/n)^{x.}/\prod_{i=1}^{n} x_i! ,$$

the desired result follows,

$$\text{i.e. } (X_1,\ldots,X_n|X.) \sim \text{Mult}_n(X.,1/n,1/n,1/n,\ldots 1/n).$$

The distribution of a vector of independent and identically distributed Poisson random variables conditioned on the total is a multinomial with parameter $m = X.$ and equal cell probabilities $1/n$. This conditional distribution is independent of the Poisson parameter λ since $X.$ is a sufficient statistic for λ.

The moment generating funtion of the multinomial is

$$\{p_1\exp(t_1) + \ldots + p_n\exp(t_n)\}^m.$$

In our case this becomes

$$M(t) = \{(1/n)[\exp(t_1) + \ldots + \exp(t_n)]\}^{X.}.$$

A1.2 THE FIRST FOUR MOMENTS OF I

From (3.5), we see that we require the evaluation for $x.>0$ of

$$E\{(S^2/\bar{X})^k \mid X.=x.\},$$

for $k=1,2,3$ and 4. Now for $k=1$ and $x.>0$, we have

$$\begin{aligned}(n-1)\bar{X}\cdot E\{S^2/\bar{X}\mid X.=x.\} &= E\{\sum_{i=1}^{n}(X_i-\bar{X})^2 \mid X.=x.\}\\ &= E\{\sum_{i=1}^{n} X_i^2 - n\bar{X}^2 \mid X.=x.\}\\ &= \sum_{i=1}^{n} E\{X_i^2 \mid X.=x.\} - n\bar{X}^2\\ &= n\{\mathrm{Var}(X_i\mid X.=x.) + E^2(X_i\mid X.=x.)\} - n\bar{X}^2\\ &= n\{x.(1/n)[1-(1/n)]+(x./n)^2\} - n\bar{X}^2\\ &= (n-1)\bar{X}\ .\end{aligned}$$

It follows that for $x.>0$, we have

$$E\{S^2/\bar{X} \mid X.=x.\} = 1. \tag{A1.3}$$

For $k=2$ and $x.>0$, we begin by noting that

$$\begin{aligned}\{(n-1)S^2\}^2 &= \{\sum_{i=1}^{n}(X_i-\bar{X})^2\}^2\\ &= \sum_{i=1}^{n}(X_i-\bar{X})^4 + \sum_{i\neq j}\sum (X_i-\bar{X})^2(X_j-\bar{X})^2 .\end{aligned}$$

Upon expanding these powers of $(X_i-\bar{X})$ and evaluating the required conditional expectations through the moment generating function of $(X_1, X_2, \ldots, X_n)\mid X.$, we have, for $x.>0$, that

$$E\{(S^2/\bar{X})^2 \mid X.=x.\} = (n+1)/(n-1) - \{2/n(n-1)\}(1/\bar{X}). \tag{A1.4}$$

It follows that for $x.>0$

$$\mathrm{Var}(S^2/\bar{X}\mid X.) = \{2/(n-1)\}\{1-(1/n\bar{X})\}.$$

Considerably more algebraic effort is required in the cases $k = 3$ and 4. For $x.>0$, we obtain

$$E\{(S^2/\bar{X})^3|X.=x.\} = (n+1)(n+3)/(n-1)^2 - 2\{1/(n-1)^2\}\cdot \{[1+(13/n)](1/\bar{X}) + (2/n)[1-(6/n)](1/\bar{X}^2)\} \tag{A1.5}$$

$$\begin{aligned} E\{(S^2/\bar{X})^4|X.=x.\} = {} & (n+1)(n+3)(n+5)/(n-1)^3 + \{2/(n-1)^3\}\cdot \\ & \{2n[1+(5/n)][1-(17/n)](1/\bar{X}) \\ & - (2/n^2)(2n^2+ 53n - 261)(1/\bar{X}^2) \\ & - (4/n^3)(n^2 - 30n +90)(1/\bar{X}^3)\}. \end{aligned} \tag{A1.6}$$

Equations (A1.3)-(A1.6) agree with those provided by Haldane (1937). Substitution of these conditional moments into (3.5) yields exact expressions for the first four raw moments of I which are given in (3.7). To obtain the central moments from these raw moments is a matter of using the formulas given in Kendall and Stuart (1958, vol. 1, p. 56). We should mention that the algebra involved in computing these conditional expectation was checked by UBC's symbolic manipulator, documented in "UBC REDUCE". Expansions of powers and the computation of the partial derivatives of the moment generating function were all checked on the computer.

It remains to evaluate $E^+(1/\bar{X}^j)$, for $j = 1,2$ and 3. Now,

$$\begin{aligned} E^+(1/\bar{X}) &= nE^+(1/X.) \\ &= n \sum_{k=1}^{\infty} (1/k)e^{-\theta}\theta^k/k! \text{ , where } \theta=n\lambda. \end{aligned}$$

If we let

$$f(\theta) = \sum_{k=1}^{\infty} (1/k)e^{-\theta}\theta^{k}/k!,$$

then

$$f'(\theta) = -f(\theta) + \sum_{k=1}^{\infty} e^{-\theta}\theta^{k-1}/k!$$

or

$$f'(\theta) + f(\theta) = e^{-\theta}(e^{\theta}-1)/\theta.$$

Since the solution to this differential equation is

$$f(\theta) = e^{-\theta}\int_0^{\theta} \{e^{t}-1)/t\}dt,$$

it follows that

$$E^{+}(1/\bar{X}) = ne^{-\theta}\int_0^{\theta} \{(e^{t}-1)/t\}dt.$$

Similarly,

$$E^{+}(1/\bar{X}^2) = n^2\{e^{-\theta}\ln\theta\int_0^{\theta} [(e^{t}-1)/t]dt - e^{-\theta}\int_0^{\theta} (\ln t)[(e^{t}-1)/t]dt\},$$

and

$$E^{+}(1/\bar{X}^3) = n^3[(1/2)e^{-\theta}(\ln\theta)^2 \int_0^{\theta}\{(e^{t}-1)/t\}dt - e^{-\theta}\ln\theta \int_0^{\theta} (\ln t)\{(e^{t}-1)/t\}dt + (1/2)e^{-\theta} \int_0^{\theta} (\ln t)^2\{(e^{t}-1)/t\}dt].$$

None of the above integrals can be evaluated explicitly, and so μ_2', μ_3' and μ_4' would either have to be approximated by numerical integration or by an asymptotic expansion. We illustrate this by expanding the integrals for large θ. Let

$$(1/n)E^{+}(1/\bar{X}) = f(\theta),$$

$$(1/n^2)E^{+}(1/\bar{X}^2) = g(\theta) \text{ and}$$

$$(1/n^3)E^{+}(1/\bar{X}^3) = h(\theta).$$

Under the transformation $t=\theta x$, we have

$$f(\theta) = e^{-\theta} \int_0^1 \{(e^{\theta x}-1)/x\}dx$$
$$= e^{-\theta}\ln x\ (e^{\theta x}-1)\Big|_0^1 - e^{-\theta} \int_0^1 (\ln x)\theta e^{\theta x}dx$$
$$= -\theta e^{-\theta} \int_0^1 \ln(1-z)e^{\theta(1-z)}dz, \text{ where } z=1-x.$$

Now,

$$-\ln(1-z) = z + z^2/2 + z^3/3 + \ldots,$$

and so

$$f(\theta) = \int_0^1 (z + z^2/2 + z^3/3 + \ldots)\theta e^{-\theta z}dz.$$

For $j = 1,2,3,\ldots,$ let

$$I_j(\theta) = \int_0^1 z^j\theta e^{-\theta z}dz$$
$$= -z^j e^{-\theta z}\Big|_0^1 + j \int_0^1 z^{j-1}e^{-\theta z}dz.$$

We then have an expression for I_j in terms of I_k, $k<j$:

$$I_j(\theta) = -e^{-\theta} + (j/\theta)I_{j-1}(\theta), \text{ where } I_0(\theta) = 1 - e^{-\theta}.$$

For $j\geq1$, this recursive formula yields

$$I_j(\theta) = -e^{-\theta} + (j/\theta)\{-e^{-\theta} + [(j-1)/\theta]I_{j-2}(\theta)\}$$
$$= -e^{-\theta} - (j/\theta)e^{-\theta} + \{j(j-1)/\theta^2\}\{-e^{-\theta}+[(j-2)/\theta]I_{j-3}(\theta)\}$$
$$\vdots$$
$$= j!/\theta^j + O(e^{-\theta})$$
$$\sim j!/\theta^j$$

Therefore,

$$f(\theta) \sim \sum_{k=0}^{N} k!/\theta^{k+1} + O(1/\theta^{N+2}) \quad \text{as } N \longrightarrow \infty .$$

and the asymptotic expansion for $E^+(1/\bar{X})$ is

$$E^+(1/\bar{X}) \sim n(1/\theta + 1/\theta^2 + 2/\theta^3 + \ldots).$$

Notice that the first term approximation of $E^+(1/\bar{X})$ is $n/\theta = 1/E(\bar{X})$, which would be the naive approximation to this expectation.

The asymptotic expansions for $g(\theta)$ and $h(\theta)$ are obtained in a similar fashion, except instead of expanding $\ln(1-z)$, we would need to expand $\{\ln(1-z)\}^2$ and $\{\ln(1-z)\}^3$ for $g(\theta)$ and $h(\theta)$ respectively. The results are

$$E^+(1/\bar{X}^2) \sim n^2(1/\theta^2 + 3/\theta^3 + 11/\theta^4 + \dots),$$

$$E^+(1/\bar{X}^3) \sim n^3(1/\theta^3 + 6/\theta^4 + 35/\theta^5 + \dots).$$

Alternately, these same expansions could be obtained by repeated applications of L'Hospital's rule.

Substituting these expansions into (3.7) yields the raw moments, correct to $O(1/n^4)$:

$$\mu_1' \sim 1 ,$$

$$\mu_2' \sim 1 + (2/n) + (2/n^2)[\,1 - (1/\lambda)] + (2/n^3)[\,1 - (1/\lambda) - (1/\lambda^2)] + (2/n^4)[\,1 - (1/\lambda) - (1/\lambda^2) - (2/\lambda^3)] ,$$

$$\mu_3' \sim 1 + (6/n) + (2/n^2)[\,7 - (1/\lambda)] + (2/n^3)[\,11 - (15/\lambda) - (3/\lambda^2)] + (2/n^4)[\,15 - (29/\lambda) - (7/\lambda^2) - (8/\lambda^3)] ,$$

$$\mu_4' \sim 1 + (12/n) + (4/n^2)[\,14 + (1/\lambda)] + (4/n^3)[\,37 - (9/\lambda) - (1/\lambda^2)] + (4/n^4)[\,72 - (115/\lambda) - (68/\lambda^2) - (6/\lambda^3)] . \qquad \text{(A1.7)}$$

Proceeding in the same way as Hoel (1943), we can also assess the accuracy of the χ^2 approximation to the null distribution of I by examining the ratio of the asymptotic moments of I with the moments of $[1/(n-1)]\chi^2_{n-1}$. The behavior of these ratios as n and/or λ increases, will indicate when the χ^2 approximation is satisfactory. The first four moments of a random variable distributed as $[1/(n-1)]\chi^2_{n-1}$ are:

$$\omega_1' = 1 ,$$

$$\omega_2' = (n+1)/(n-1) ,$$

$$\omega_3' = (n+1)(n+3)/(n-1)^2 ,$$

$$\omega_4' = (n+1)(n+3)(n+5)/(n-1)^3 ,$$

(see Mendenhall and Scheaffer, 1973, p.138).

Notice that the moments of $[1/(n-1)]\chi^2_{n-1}$ approximate the moments of a $\mathrm{Mult}_n(x.,1/n\ldots,1/n)$, correct to $O(1/n)$.

Let $R_i = \mu_i'/\omega_i'$, for i = 1,2,3 and 4 (note that $R_1 \equiv 1$ for all n and λ). Using the asymptotic expressions in (A1.7), these ratios are computed for n = 10,20,50 and 100 and λ = 1,3,5 and 8, and entered in Table A1.

The asymptotic moments of the index of dispersion agree very well with those of the χ^2_{n-1} distribution for $n \geq 20$ and $\lambda \geq 1$. In fact, this is also apparent for $n \geq 10$ and $\lambda \geq 5$. As n and/or λ increases, R_2, R_3, and R_4 all approach the limiting value 1. This is indeed encouraging and compliments the results obtained in section 2.5.

TABLE A1: The Ratios of the Moments of I and χ^2_{n-1}
(for each cell, the ratios R_2,R_3 and R_4 are entered in that order)

	λ			
n	1	3	5	8
10	0.9797 0.9631 0.9724	0.9937 0.9887 0.9921	0.9962 0.9932 0.9948	0.9977 0.9957 0.9962
20	0.9950 0.9925 1.0003	0.9984 0.9976 1.0002	0.9990 0.9986 1.0001	0.9994 0.9991 1.0001
50	0.9992 0.9991 1.0009	0.9997 0.9997 1.0003	0.9998 0.9998 1.0002	0.9999 0.9998 1.0001
100	0.9998 0.9998 1.0003	0.9999 0.9999 1.0001	0.99996 0.99998 1.00004	0.99998 0.99998 0.99996

A1.3 THE TYPE VI PEARSON CURVE

We rewrite the differential equation given by (3.1) as

$$d\{\log f(x)\}/dx = (x-a)/\{b_2(x-A_1)(x-A_2)\}, \qquad (A1.8)$$

where A_1 and A_2 are the roots of the quadratic $b_0 + b_1x + b_2x^2$. Kendall and Stuart (1958, vol. 1, p.149) give the formulas for a, b_0, b_1 and b_2 as functions of β_1, β_2 and μ_2. When using these formulas, one should keep in mind that the formulas were obtained assuming the origin at the mean.

For the type VI case, both roots of the quadratic are real and have the same sign. Without loss of generality, assume that $A_2>A_1$. Then, by partial fractions, we can write

$$d\{\log f(x)\}/dx = (1/b_2)\{C_1/(x-A_1) + C_2/(x-A_2)\},$$

where $C_1 = (a-A_1)/(A_2-A_1) = (a-A_1)/\delta$,

$C_2 = (A_2-a)/(A_2-A_1) = (A_2-a)/\delta$ and

$\delta = A_2-A_1$.

For $x>A_2$, we can integrate equation (A1.8) with respect to x to get

$$\log f(x) = (C_1/b_2)\log(x-A_1) + (C_2/b_2)\log(x-A_2) + C$$

where C is the arbitrary constant of integration.

Transforming back to the true origin, i.e. replacing x by x-1, yields

$$\log f(x) = (C_1/b_2)\log(x-a_1) + (C_2/b_2)\log(x-a_2) + C$$

where $a_1 = 1+A_1$ and $a_2 = 1+A_2$, and hence

$$f(x) = k(x-a_1)^{-q_1}(x-a_2)^{q_2} \qquad (A1.9)$$

where $q_1 = -C_1/b_2$, $q_2 = C_2/b_2$ and k is a normalizing constant.

Since $A_2 > A_1$ (and hence $a_2 > a_1$) and q_1 and q_2 are real numbers, it follows that type VI Pearson curve defined in (A1.9) is a distribution defined on $[a_2,\infty)$. If we let $y = x-a_1$, then

$$f(y) = ky^{-q_1}(y+a_1-a_2)^{q_2}, \text{ for } y>a_2-a_1>0,$$

$$= ky^{-q_1}(y-\delta)^{q_2}, \text{ since } \delta = A_2 - A_1 = a_2 - a_1.$$

Now let $z = \delta/y$ so that $dy/dz = -\delta/z^2$. Then

$$f(y) = k(\delta/z)^{-q_1}[(\delta/z)-\delta]^{q_2}|dy/dz|$$

$$= k\delta^{q_2-q_1+1} z^{q_1-2}[(1-z)/z]^{q_2}$$

$$= k'z^{q_1-q_2-2}(1-z)^{q_2}, \text{ for } 0<z<1.$$

This last form of the density of the beta distribution is what is required when using the IMSL library to compute critical values.

A1.4 A LIMITING CASE OF THE NEGATIVE BINOMIAL

The negative binomial distribution with parameters k and θ approaches different distributions depending on the limiting operation. In particular, let $k \to \infty$ and $\theta \to 0$ in such a way that $k\theta = \lambda$, a constant. If $X \sim NB(k,\theta)$, then the moment generating function of X is

$$M_X(t) = \{p/(1-qe^t)\}^k,$$

where $p = 1/(\theta+1)$ and $q = \theta/(\theta+1)$. Hence,

$$M_x(t) = \{[1/(\theta+1)]/[1-(\theta/\theta+1)e^t]\}^k$$

$$= \{[k/(\lambda+k)][1-(\lambda/(\lambda+k))e^t]\}^k$$

$$= \{k/[\lambda(1-e^t) + k]\}^k$$

$$= \{1+[\lambda(1-e^t)/k]\}^{-k} \qquad \text{(A1.5)}$$

But as $k \to \infty$, the limit of the right-hand side of (A1.5) is $e^{\lambda(e^t-1)}$, which is precisely the moment of generating function of the Poisson distribution.

A2.1 HISTOGRAMS OF I

FIG. A1 HISTOGRAM OF I (1000 samples, n = 10, λ = 3)

SYMBOL	COUNT	MEAN	ST.DEV.
X	1000	0.980	0.452

EACH SYMBOL REPRESENTS 1 OBSERVATIONS

INTERVAL NAME	FREQUENCY INT.	FREQUENCY CUM.	PERCENTAGE INT.	PERCENTAGE CUM.
*.240000	10	10	1.0	1.0
*.360000	32	42	3.2	4.2
*.480000	74	116	7.4	11.6
*.600000	86	202	8.6	20.2
*.720000	106	308	10.6	30.8
*.840000	111	419	11.1	41.9
*.960000	111	530	11.1	53.0
*1.08000	113	643	11.3	64.3
*1.20000	94	737	9.4	73.7
*1.32000	77	814	7.7	81.4
*1.44000	54	868	5.4	86.8
*1.56000	39	907	3.9	90.7
*1.68000	23	930	2.3	93.0
*1.80000	18	948	1.8	94.8
*1.92000	13	961	1.3	96.1
*2.04000	14	975	1.4	97.5
*2.16000	1	976	0.1	97.6
*2.28000	6	982	0.6	98.2
*2.40000	9	991	0.9	99.1
*2.52000	3	994	0.3	99.4
*2.64000	3	997	0.3	99.7
*2.76000	1	998	0.1	99.8
*2.88000	0	998	0.0	99.8
*3.00000	1	999	0.1	99.9
*3.12000	1	1000	0.1	100.0
*3.24000	0	1000	0.0	100.0

FIG. A2 HISTOGRAM OF I (1000 samples, n = 10, λ = 5)

```
                         SYMBOL  COUNT      MEAN      ST.DEV.
                            X     1000      0.981      0.447
                   EACH SYMBOL REPRESENTS      1 OBSERVATIONS
INTERVAL                                                                                                 FREQUENCY PERCENTAGE
NAME           5    10   15   20   25   30   35   40   45   50   55   60   65   70   75   80  INT. CUM. INT.  CUM.
          +----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+
*.240000  +XXXXXXXXXXXXXX                                                                      14    14   1.4    1.4
*.360000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                                                 35    49   3.5    4.9
*.480000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                    62   111   6.2   11.1
*.600000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*  101   212  10.1   21.2
*.720000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*   94   306   9.4   30.6
*.840000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*  116   422  11.6   42.2
*.960000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*  103   525  10.3   52.5
*1.08000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*  103   628  10.3   62.8
*1.20000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*   96   724   9.6   72.4
*1.32000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*   83   807   8.3   80.7
*1.44000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                                 51   858   5.1   85.8
*1.56000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                                         43   901   4.3   90.1
*1.68000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                                                     31   932   3.1   93.2
*1.80000  +XXXXXXXXXXXXXXXXXXXX                                                                20   952   2.0   95.2
*1.92000  +XXXXXXXXXXX                                                                         11   963   1.1   96.3
*2.04000  +XXXXXXXXXXXXX                                                                       13   976   1.3   97.6
*2.16000  +XXXXX                                                                                5   981   0.5   98.1
*2.28000  +XXXXXXXX                                                                             8   989   0.8   98.9
*2.40000  +XXXXX                                                                                5   994   0.5   99.4
*2.52000  +XXX                                                                                  3   997   0.3   99.7
*2.64000  +                                                                                     0   997   0.0   99.7
*2.76000  +                                                                                     0   997   0.0   99.7
*2.88000  +XX                                                                                   2   999   0.2   99.9
*3.00000  +X                                                                                    1  1000   0.1  100.0
*3.12000  +                                                                                     0  1000   0.0  100.0
*3.24000  +                                                                                     0  1000   0.0  100.0
          +----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+
               5    10   15   20   25   30   35   40   45   50   55   60   65   70   75   80
```

FIG. A3 HISTOGRAM OF I (1000 samples, n = 20, λ = 3)

```
                                  SYMBOL   COUNT      MEAN      ST.DEV.
                                    X      1000      0.998       0.330
                           EACH SYMBOL REPRESENTS      1 OBSERVATIONS
INTERVAL                                                                                         FREQUENCY PERCENTAGE
NAME          5    10   15   20   25   30   35   40   45   50   55   60   65   70   75   80    INT. CUM. INT.  CUM.
         +----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+
*.300000 +                                                                                     0     0   0.0   0.0
*.400000 +XXXXXXXXXXX                                                                         11    11   1.1   1.1
*.500000 +XXXXXXXXXXXXXXXXXXXXXXXX                                                            24    35   2.4   3.5
*.600000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                    64    99   6.4   9.9
*.700000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*   91   190   9.1  19.0
*.800000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*  103   293  10.3  29.3
*.900000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*  114   407  11.4  40.7
*1.00000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*  142   549  14.2  54.9
*1.10000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*   99   648   9.9  64.8
*1.20000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*   98   746   9.8  74.6
*1.30000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*   84   830   8.4  83.0
*1.40000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                               53   883   5.3  88.3
*1.50000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                             55   938   5.5  93.8
*1.60000 +XXXXXXXXXXXXXXXXXXXXXXX                                                             23   961   2.3  96.1
*1.70000 +XXXXXXXXXXX                                                                         11   972   1.1  97.2
*1.80000 +XXXXXXXXXXX                                                                         11   983   1.1  98.3
*1.90000 +XXXX                                                                                 4   987   0.4  98.7
*2.00000 +XXX                                                                                  3   990   0.3  99.0
*2.10000 +XXXX                                                                                 4   994   0.4  99.4
*2.20000 +XXX                                                                                  3   997   0.3  99.7
*2.30000 +                                                                                     0   997   0.0  99.7
*2.40000 +XX                                                                                   2   999   0.2  99.9
*2.50000 +                                                                                     0   999   0.0  99.9
*2.60000 +X                                                                                    1  1000   0.1 100.0
*2.70000 +                                                                                     0  1000   0.0 100.0
*2.80000 +                                                                                     0  1000   0.0 100.0
         +----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+
              5    10   15   20   25   30   35   40   45   50   55   60   65   70   75   80
```

FIG. A4 HISTOGRAM OF I (1000 samples, n = 20, λ = 5)

```
                        SYMBOL  COUNT     MEAN      ST.DEV.
                          X      1000     0.993       0.328
                   EACH SYMBOL REPRESENTS      1 OBSERVATIONS
INTERVAL                                                                                   FREQUENCY PERCENTAGE
NAME          5    10   15   20   25   30   35   40   45   50   55   60   65   70   75   80  INT. CUM.  INT.  CUM.
         +----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+
*.200000 +                                                                                   0    0   0.0   0.0
*.300000 +XXX                                                                                3    3   0.3   0.3
*.400000 +XXXXXXXXXXX                                                                       11   14   1.1   1.4
*.500000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                                                   31   45   3.1   4.5
*.600000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                   63  108   6.3  10.8
*.700000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*  81  189   8.1  18.9
*.800000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX* 126  315  12.6  31.5
*.900000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX* 113  428  11.3  42.8
*1.00000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX* 102  530  10.2  53.0
*1.10000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX* 124  654  12.4  65.4
*1.20000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*  93  747   9.3  74.7
*1.30000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX          72  819   7.2  81.9
*1.40000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX          72  891   7.2  89.1
*1.50000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                                           39  930   3.9  93.0
*1.60000 +XXXXXXXXXXXXXXXXXXXXXXXXXX                                                        26  956   2.6  95.6
*1.70000 +XXXXXXXXXXXXXXXXXX                                                                18  974   1.8  97.4
*1.80000 +XXXXXXXXX                                                                          9  983   0.9  98.3
*1.90000 +XXXXXXX                                                                            7  990   0.7  99.0
*2.00000 +XXXXXX                                                                             6  996   0.6  99.6
*2.10000 +X                                                                                  1  997   0.1  99.7
*2.20000 +X                                                                                  1  998   0.1  99.8
*2.30000 +X                                                                                  1  999   0.1  99.9
*2.40000 +X                                                                                  1 1000   0.1 100.0
*2.50000 +                                                                                   0 1000   0.0 100.0
*2.60000 +                                                                                   0 1000   0.0 100.0
*2.70000 +                                                                                   0 1000   0.0 100.0
         +----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+
              5    10   15   20   25   30   35   40   45   50   55   60   65   70   75   80
```

FIG. A5 HISTOGRAM OF I (1000 samples, n = 50, λ = 3)

```
                                   SYMBOL  COUNT     MEAN     ST.DEV.
                                     X     1000      1.003      0.204
                              EACH SYMBOL REPRESENTS      1 OBSERVATIONS
INTERVAL                                                                                      FREQUENCY PERCENTAGE
NAME             5    10   15   20   25   30   35   40   45   50   55   60   65   70   75   80  INT. CUM.  INT.  CUM.
                +----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+
*.550000 +XXXX                                                                                     4    4   0.4   0.4
*.600000 +XXXXXX                                                                                   6   10   0.6   1.0
*.650000 +XXXXXXXXXXXXXXX                                                                         15   25   1.5   2.5
*.700000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                                                     35   60   3.5   6.0
*.750000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                                                     35   95   3.5   9.5
*.800000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                                        48  143   4.8  14.3
*.850000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                     67  210   6.7  21.0
*.900000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*     109  319  10.9  31.9
*.950000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*     113  432  11.3  43.2
*1.00000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*     105  537  10.5  53.7
*1.05000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*      94  631   9.4  63.1
*1.10000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX           77  708   7.7  70.8
*1.15000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX             75  783   7.5  78.3
*1.20000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                     67  850   6.7  85.0
*1.25000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                                              42  892   4.2  89.2
*1.30000 +XXXXXXXXXXXXXXXXXXXXXXXXXXXX                                                            28  920   2.8  92.0
*1.35000 +XXXXXXXXXXXXXXXXXXXXXXX                                                                 23  943   2.3  94.3
*1.40000 +XXXXXXXXXXXXXXXX                                                                        16  959   1.6  95.9
*1.45000 +XXXXXXXXXXXXX                                                                           13  972   1.3  97.2
*1.50000 +XXXXXXXXXX                                                                              10  982   1.0  98.2
*1.55000 +XXXX                                                                                     4  986   0.4  98.6
*1.60000 +XXXX                                                                                     4  990   0.4  99.0
*1.65000 +XXXX                                                                                     4  994   0.4  99.4
*1.70000 +XX                                                                                       2  996   0.2  99.6
*1.75000 +XXX                                                                                      3  999   0.3  99.9
*1.80000 +X                                                                                        1 1000   0.1 100.0
                +----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+
                     5    10   15   20   25   30   35   40   45   50   55   60   65   70   75   80
```

FIG. A6 HISTOGRAM OF I (1000 samples, n = 50, λ = 5)

```
                          SYMBOL  COUNT     MEAN      ST.DEV.
                            X      1000     1.001      0.203
                    EACH SYMBOL REPRESENTS     1 OBSERVATIONS
INTERVAL                                                                                      FREQUENCY PERCENTAGE
NAME         5    10   15   20   25   30   35   40   45   50   55   60   65   70   75   80   INT. CUM. INT.  CUM.
          +----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+
*.480000  +                                                                                  0     0   0.0   0.0
*.540000  +XX                                                                                2     2   0.2   0.2
*.600000  +XXXXXXXXXXXXXXX                                                                  15    17   1.5   1.7
*.660000  +XXXXXXXXXXXXXX                                                                   14    31   1.4   3.1
*.720000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                                            37    68   3.7   6.8
*.780000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                              51   119   5.1  11.9
*.840000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX* 96   215   9.6  21.5
*.900000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX* 111  326  11.1  32.6
*.960000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX* 122  448  12.2  44.8
*1.02000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX* 121  569  12.1  56.9
*1.08000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX* 130  699  13.0  69.9
*1.14000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX         73   772   7.3  77.2
*1.20000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                 65   837   6.5  83.7
*1.26000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                         56   893   5.6  89.3
*1.32000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                                            37   930   3.7  93.0
*1.38000  +XXXXXXXXXXXXXXXXXXXXXXXXX                                                        25   955   2.5  95.5
*1.44000  +XXXXXXXXXXXXXXXXX                                                                17   972   1.7  97.2
*1.50000  +XXXXXXXXXX                                                                       10   982   1.0  98.2
*1.56000  +XXXXXXXX                                                                          8   990   0.8  99.0
*1.62000  +XXXX                                                                              4   994   0.4  99.4
*1.68000  +XXXX                                                                              4   998   0.4  99.8
*1.74000  +X                                                                                 1   999   0.1  99.9
*1.80000  +                                                                                  0   999   0.0  99.9
*1.86000  +X                                                                                 1  1000   0.1 100.0
*1.92000  +                                                                                  0  1000   0.0 100.0
*1.98000  +                                                                                  0  1000   0.0 100.0
          +----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+
               5    10   15   20   25   30   35   40   45   50   55   60   65   70   75   80
```

FIG. A7 HISTOGRAM OF I (1000 samples, n = 100, λ = 3)

```
                          SYMBOL   COUNT      MEAN       ST.DEV.
                            X      1000       1.001        0.144
                    EACH SYMBOL REPRESENTS       1 OBSERVATIONS
INTERVAL                                                                                      FREQUENCY  PERCENTAGE
NAME         5    10   15   20   25   30   35   40   45   50   55   60   65   70   75   80  INT.  CUM.  INT.   CUM.
           +----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+
*.640000   +XXX                                                                                3     3   0.3    0.3
*.680000   +XX                                                                                 2     5   0.2    0.5
*.720000   +XXXXXXXXX                                                                          9    14   0.9    1.4
*.760000   +XXXXXXXXXXXXXXXXXXXX                                                              20    34   2.0    3.4
*.800000   +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                                           39    73   3.9    7.3
*.840000   +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                       59   132   5.9   13.2
*.880000   +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                66   198   6.6   19.8
*.920000   +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX* 105   303  10.5   30.3
*.960000   +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*  98   401   9.8   40.1
*1.00000   +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX* 124   525  12.4   52.5
*1.04000   +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX* 108   633  10.8   63.3
*1.08000   +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*  85   718   8.5   71.8
*1.12000   +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX    79   797   7.9   79.7
*1.16000   +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX        75   872   7.5   87.2
*1.20000   +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                                           39   911   3.9   91.1
*1.24000   +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                                                 33   944   3.3   94.4
*1.28000   +XXXXXXXXXXXXXXXXXXX                                                               19   963   1.9   96.3
*1.32000   +XXXXXXXXXXXXX                                                                     13   976   1.3   97.6
*1.36000   +XXXXXXXXXXXX                                                                      12   988   1.2   98.8
*1.40000   +XXXXXX                                                                             6   994   0.6   99.4
*1.44000   +XX                                                                                 2   996   0.2   99.6
*1.48000   +XXX                                                                                3   999   0.3   99.9
*1.52000   +                                                                                   0   999   0.0   99.9
*1.56000   +                                                                                   0   999   0.0   99.9
*1.60000   +X                                                                                  1  1000   0.1  100.0
*1.64000   +                                                                                   0  1000   0.0  100.0
           +----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+
             5    10   15   20   25   30   35   40   45   50   55   60   65   70   75   80
```

FIG. A8 HISTOGRAM OF I (1000 samples, n = 100, λ = 5)

```
                                      SYMBOL   COUNT       MEAN      ST.DEV.
                                         X      1000       1.000        0.141
                              EACH SYMBOL REPRESENTS          1 OBSERVATIONS
INTERVAL                                                                                                FREQUENCY PERCENTAGE
NAME             5    10    15    20    25    30    35    40    45    50    55    60    65    70    75    80  INT. CUM.  INT.  CUM.
          +----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+
*.500000  +                                                                                   0     0   0.0    0.0
*.550000  +                                                                                   0     0   0.0    0.0
*.600000  +X                                                                                  1     1   0.1    0.1
*.650000  +X                                                                                  1     2   0.1    0.2
*.700000  +XXXXXXXX                                                                           8    10   0.8    1.0
*.750000  +XXXXXXXXXXXXXXXX                                                                  16    26   1.6    2.6
*.800000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                                        42    68   4.2    6.8
*.850000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX       75   143   7.5   14.3
*.900000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX* 107   250  10.7   25.0
*.950000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX* 129   379  12.9   37.9
*1.00000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX* 152   531  15.2   53.1
*1.05000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX* 118   649  11.8   64.9
*1.10000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX* 123   772  12.3   77.2
*1.15000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*  82   854   8.2   85.4
*1.20000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                          57   911   5.7   91.1
*1.25000  +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX                                   48   959   4.8   95.9
*1.30000  +XXXXXXXXXXXXXXXXX                                                                 17   976   1.7   97.6
*1.35000  +XXXXXXXXXXXXXXXX                                                                  16   992   1.6   99.2
*1.40000  +XXXX                                                                               4   996   0.4   99.6
*1.45000  +X                                                                                  1   997   0.1   99.7
*1.50000  +XX                                                                                 2   999   0.2   99.9
*1.55000  +                                                                                   0   999   0.0   99.9
*1.60000  +X                                                                                  1  1000   0.1  100.0
*1.65000  +                                                                                   0  1000   0.0  100.0
*1.70000  +                                                                                   0  1000   0.0  100.0
*1.75000  +                                                                                   0  1000   0.0  100.0
          +----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+----+
                 5    10    15    20    25    30    35    40    45    50    55    60    65    70    75    80
```

FIG. A.9: NORMAL PROBABILITY PLOT FOR I

(1000 Samples, n = 10, λ = 3)

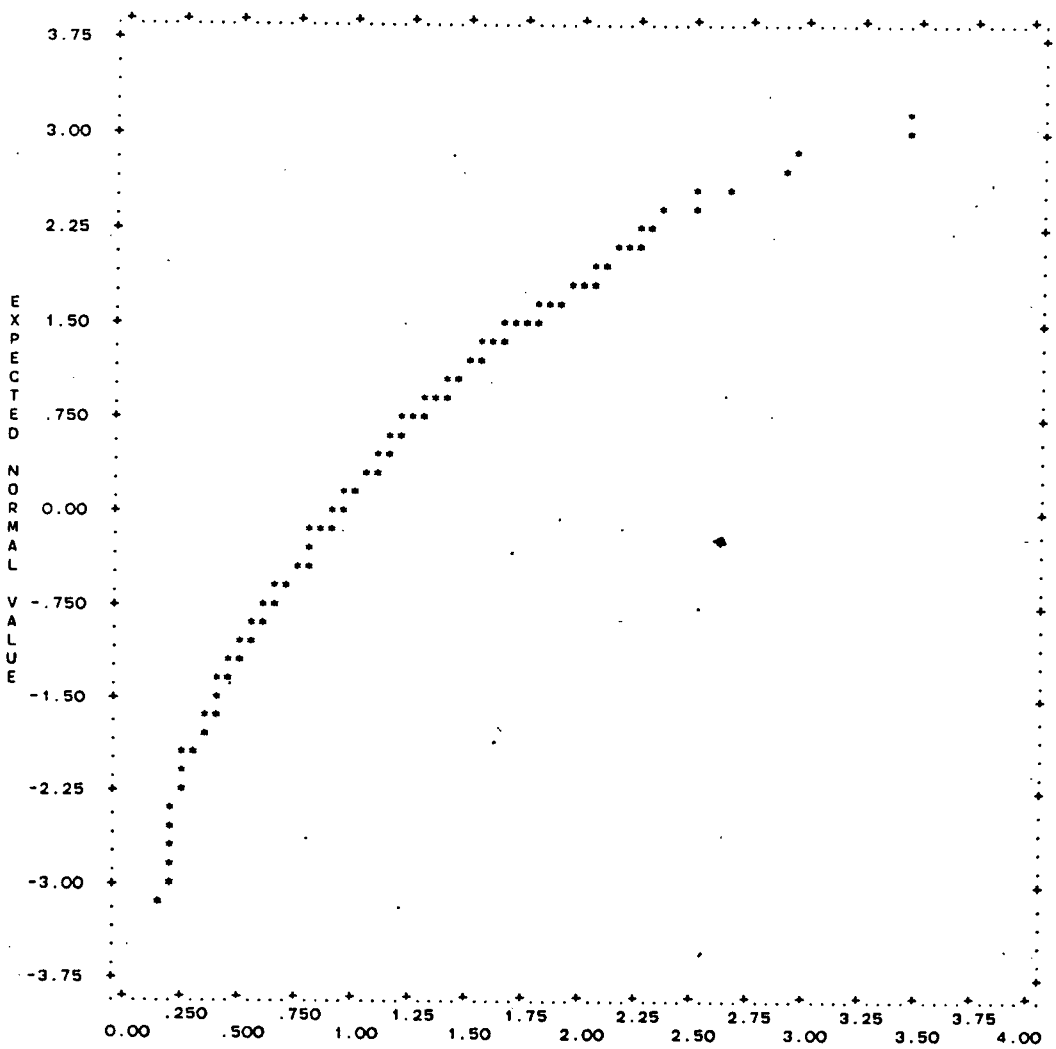


FIG. A.10: NORMAL PROBABILITY PLOT FOR $\hat{t}$

(1000 Samples, n = 10, λ = 5)

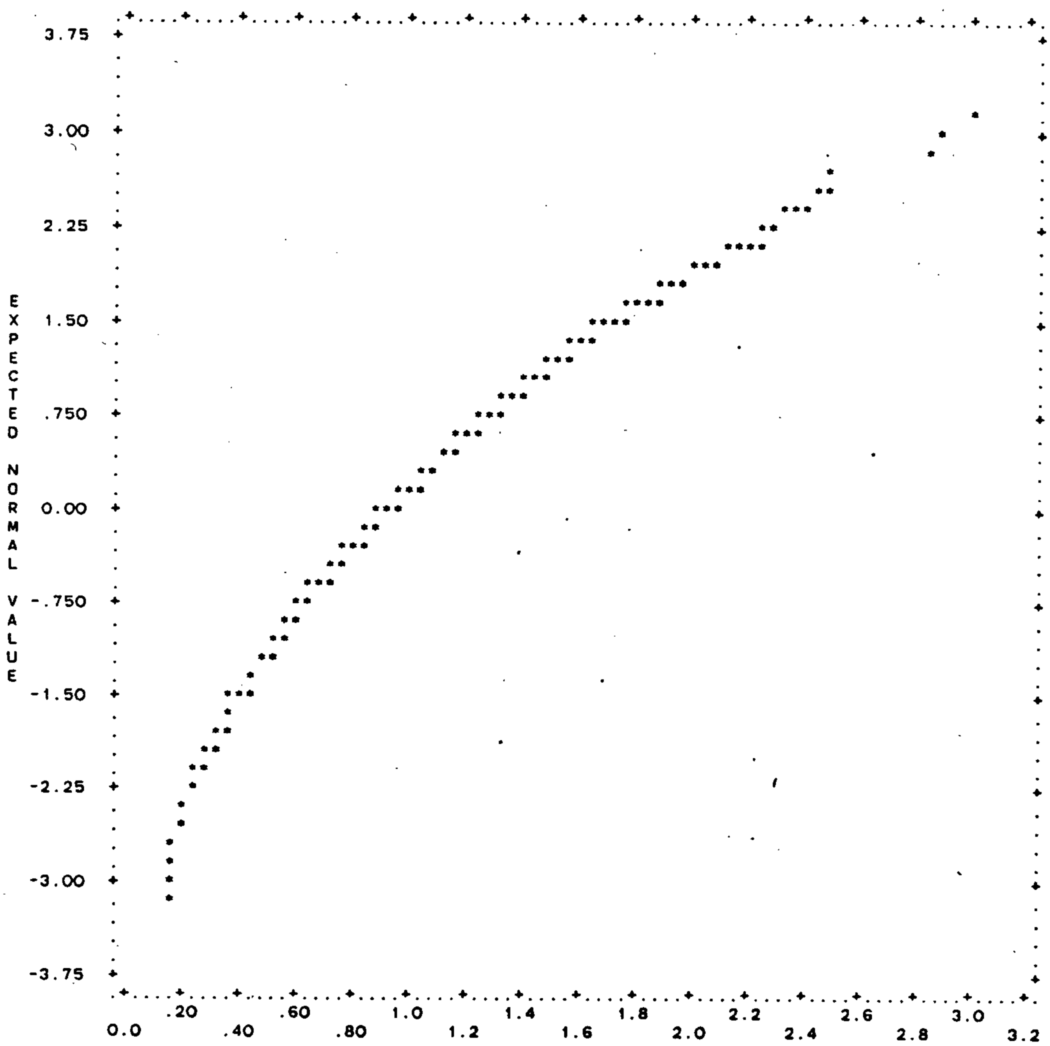


FIG. A.11: NORMAL PROBABILITY PLOT FOR I

(1000 Samples, n = 20, λ = 3)

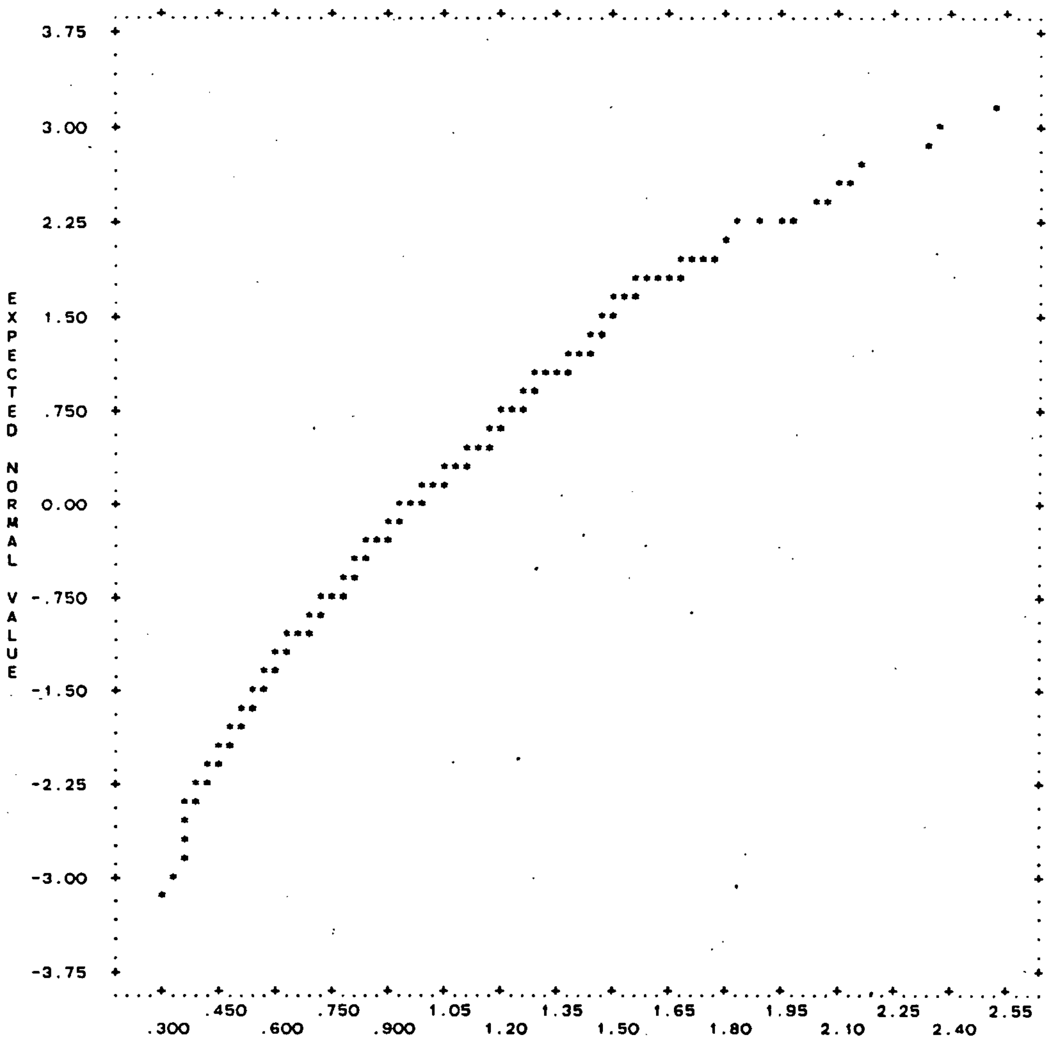


FIG. A.12: NORMAL PROBABILITY PLOT FOR I

(100C Samples, n = 20, λ = 5)

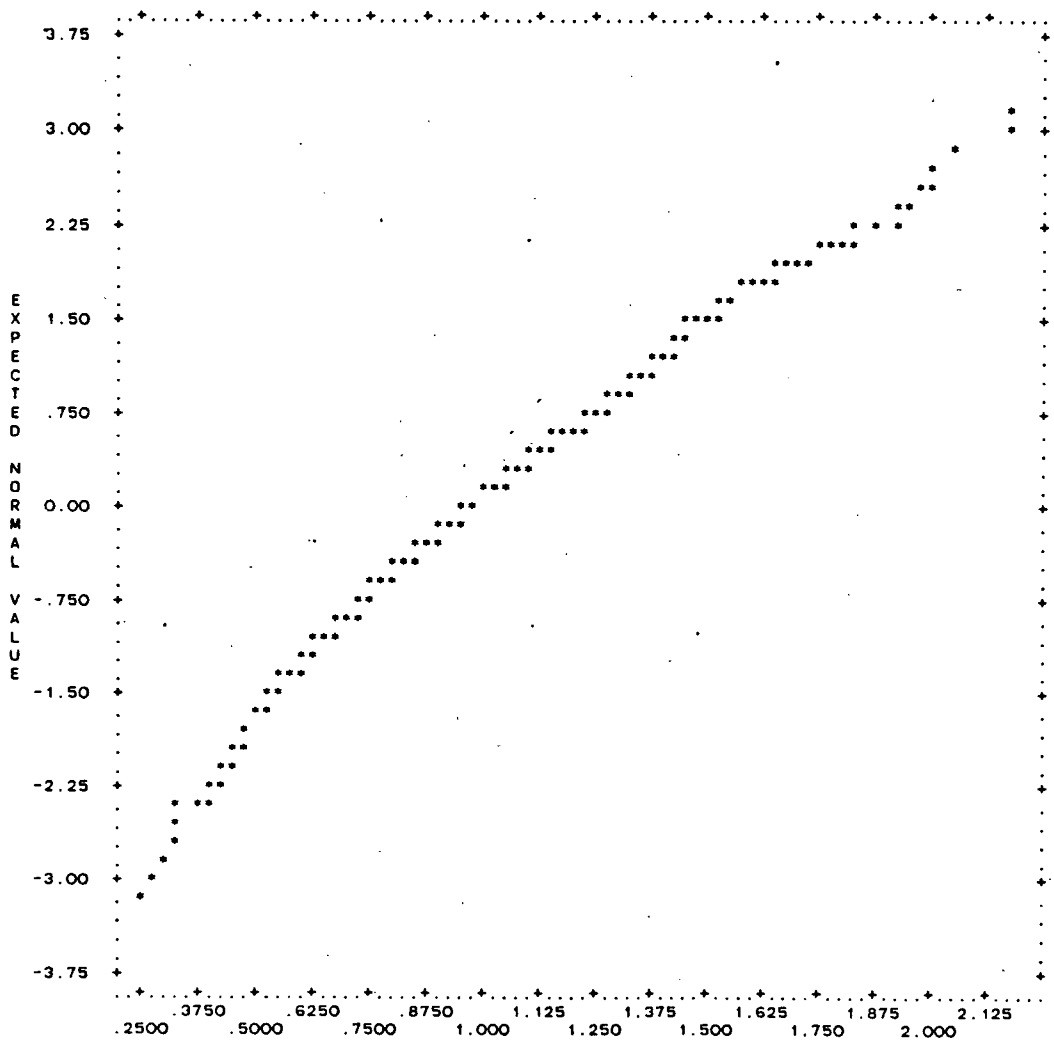


FIG. A.13: NORMAL PROBABILITY PLOT FOR I

(1000 Samples, n = 50, λ = 3)

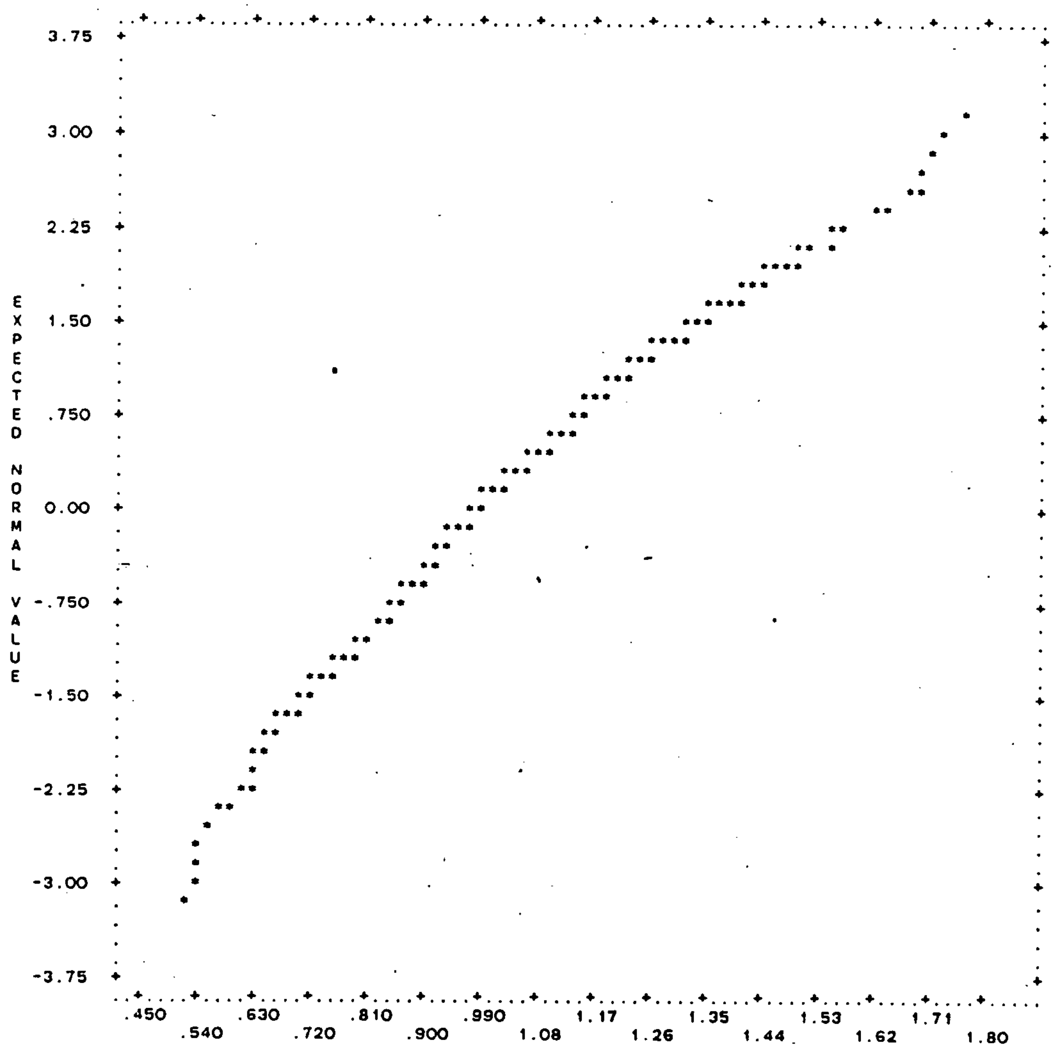


FIG. A.14: NORMAL PROBABILITY PLOT FOR I

(1000 Samples, n = 50, λ = 5)

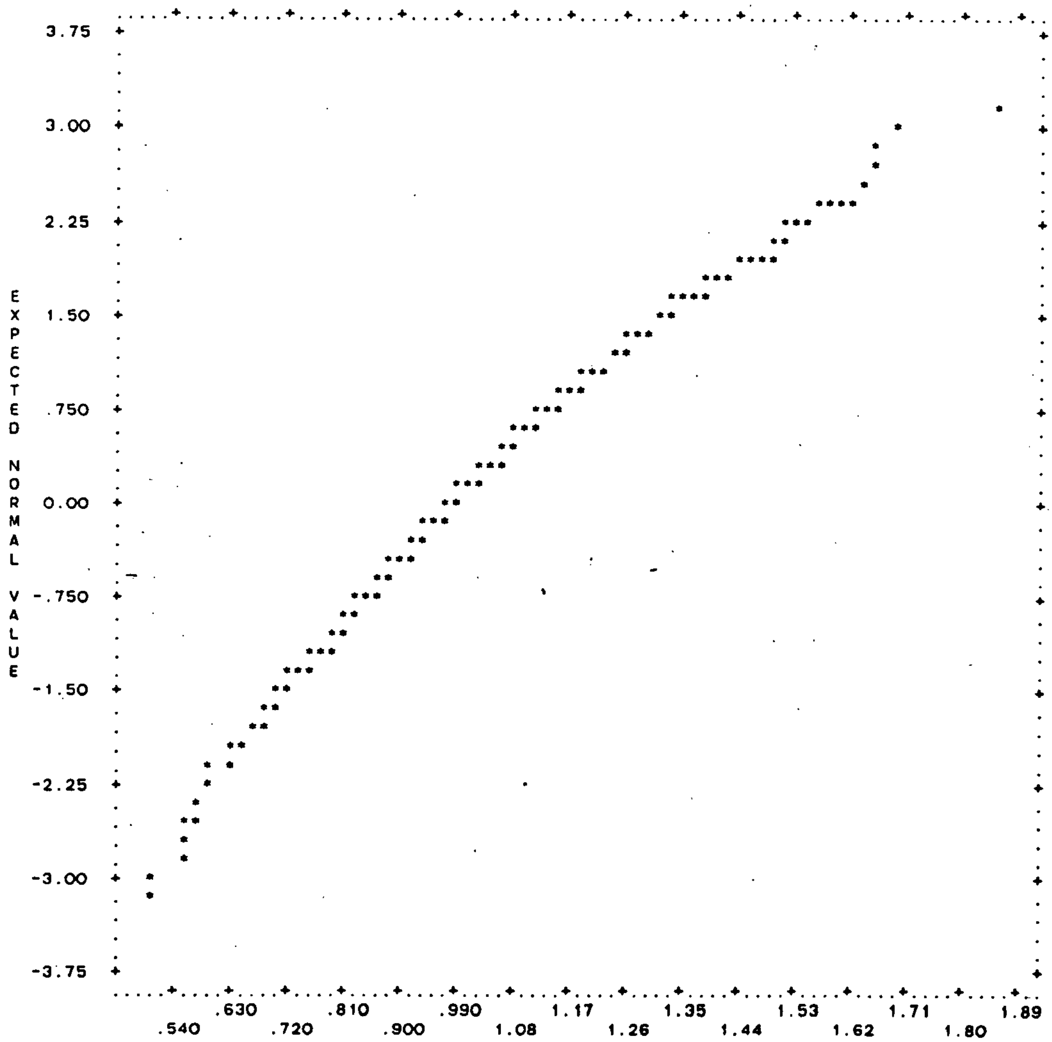


FIG. A.15: NORMAL PROBABILITY PLOT FOR I

(1000 Samples, n = 100, λ = 3)

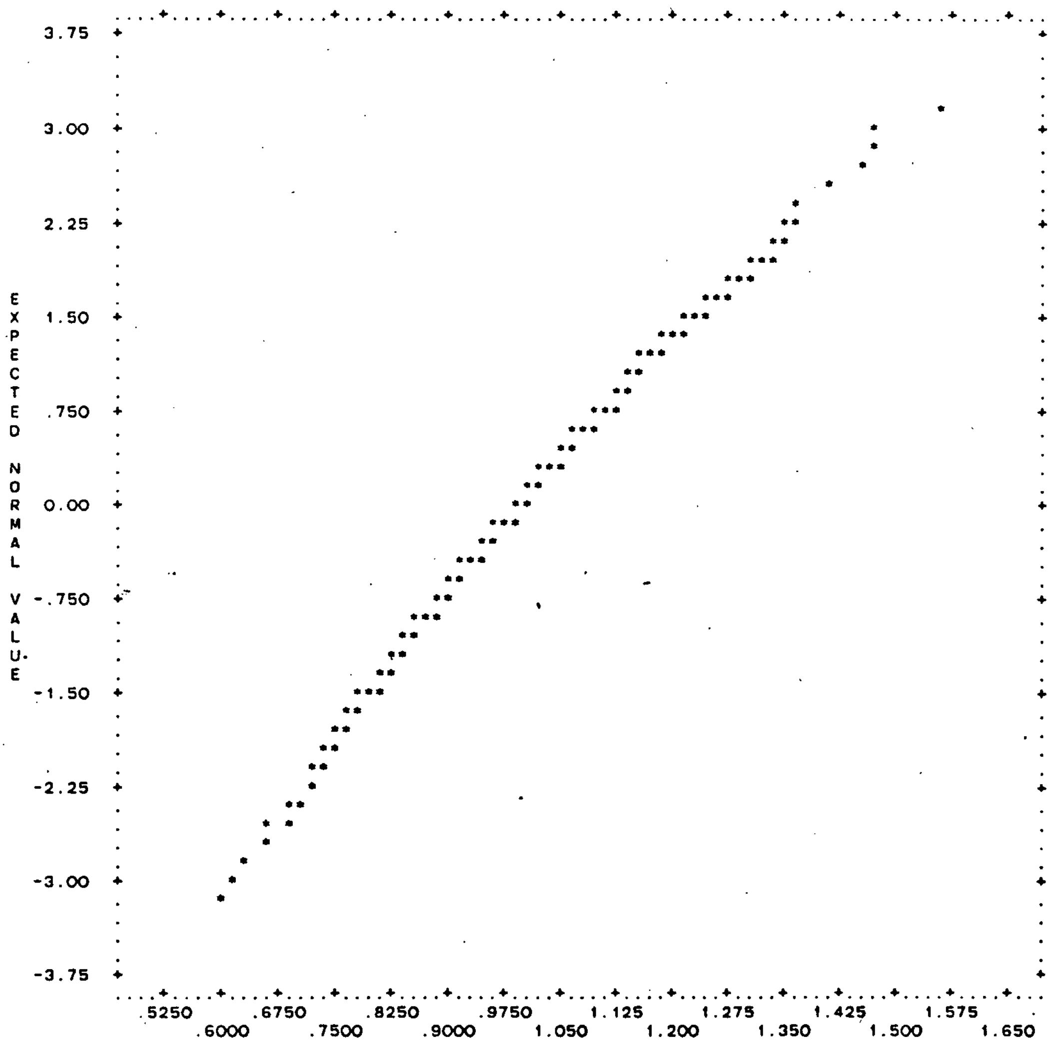
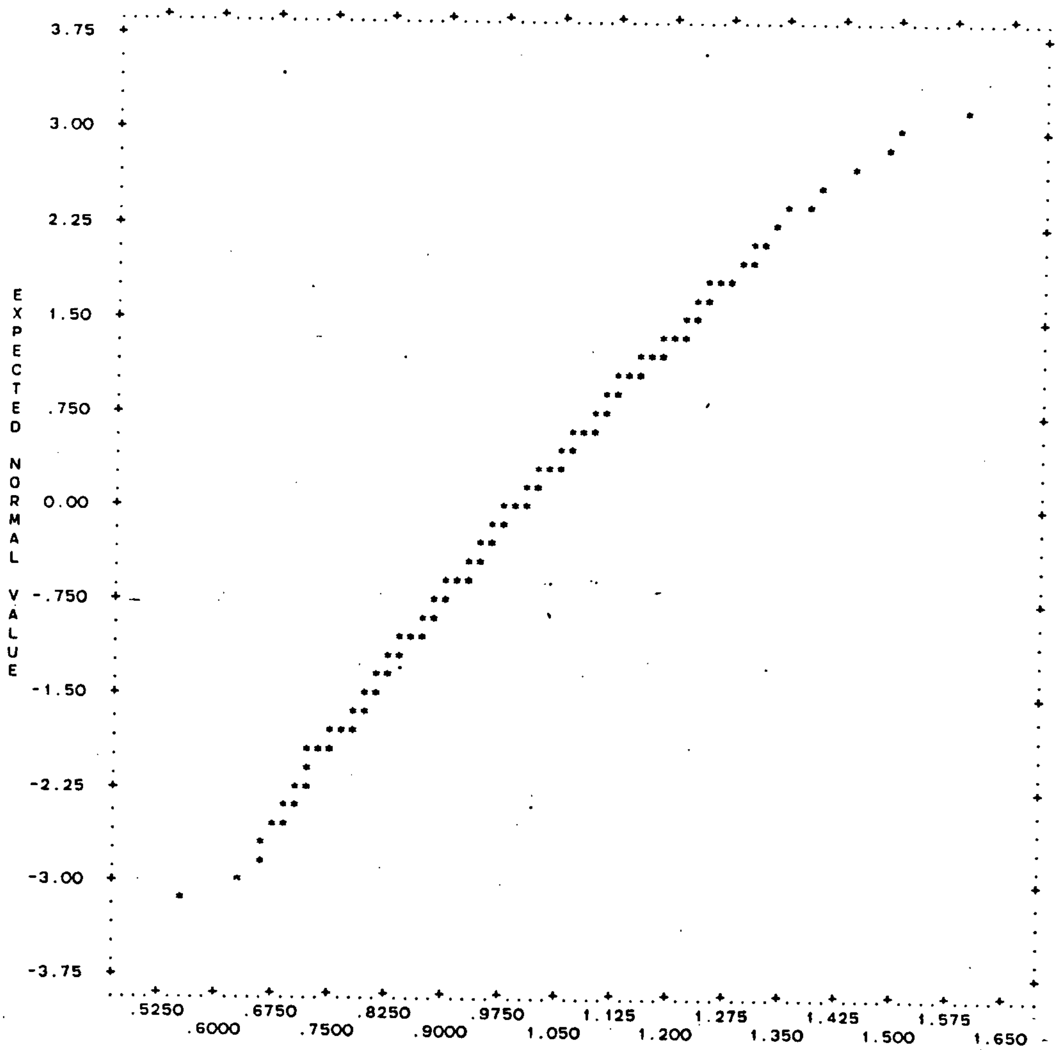


FIG. A.16: NORMAL PROBABILITY PLOT FOR $\hat{I}$

(1000 Samples, n = 100, λ = 5)



A2.2 EMPIRICAL CRITICAL VALUES

TABLE A2: EMPIRICAL CRITICAL VALUES OF I (BASED ON 15,000 SAMPLES)

n	λ	α = .005	.025	.05	.10	.90	.95	.975	.995
10	1	.2222	.3333	.4444	.5000	1.5926	1.8395	2.1111	2.5802
	3	.1981	.3016	.3827	.4709	1.6173	1.8596	2.1111	2.6667
	5	.1915	.2995	.3755	.4656	1.6257	1.8718	2.1058	2.6330
	8	.1888	.2963	.3668	.4632	1.6250	1.8650	2.0982	2.5679
20	1	.4035	.4958	.5673	.6491	1.3963	1.5812	1.7519	2.0702
	3	.3699	.4737	.5360	.6134	1.4332	1.5925	1.7303	2.0702
	5	.3661	.4777	.5359	.6157	1.4324	1.5961	1.7247	2.0360
	8	.3684	.4705	.5324	.6118	1.4368	1.5873	1.7224	2.0413
50	1	.5732	.6584	.7066	.7522	1.2699	1.3639	1.4439	1.6150
	3	.5578	.6468	.6957	.7533	1.2676	1.3558	1.4422	1.6192
	5	.5569	.6463	.6970	.7539	1.2662	1.3576	1.4341	1.6126
	8	.5547	.6445	.6919	.7531	1.2659	1.3553	1.4292	1.6046
100	1	.6885	.7467	.7835	.8263	1.1839	1.2496	1.3068	1.4220
	3	.6722	.7449	.7819	.8246	1.1839	1.2454	1.3004	1.4005
	5	.6718	.7426	.7836	.8267	1.1867	1.2430	1.2964	1.3981
	8	.6710	.7413	.7809	.8251	1.1863	1.2427	1.2954	1.4012

A2.3 PEARSON CURVE CRITICAL VALUES

TABLE A3: PEARSON CURVE CRITICAL VALUES (EXACT)

		α							
n	λ	.005	.025	.05	.10	.90	.95	.975	.995
10	1	.2460	.3594	.4270	.5143	1.5760	1.8252	2.0722	2.6560
	3	.1905	.3103	.3834	.4787	1.6108	1.8599	2.1000	2.6435
	5	.1889	.3050	.3772	.4723	1.6188	1.8677	2.1057	2.6361
	8	.1895	.3027	.3740	.4688	1.6235	1.8722	2.1087	2.6312
20	1	.3916	.5016	.5627	.6381	1.4113	1.5746	1.7327	2.0954
	3	.3675	.4788	.5423	.6218	1.4245	1.5828	1.7315	2.0570
	5	.3639	.4745	.5383	.6184	1.4273	1.5843	1.7306	2.0471
	8	.3623	.4723	.5360	.6165	1.4290	1.5851	1.7301	2.0411
50	1	.5813	.6624	.7068	.7609	1.2624	1.3565	1.4444	1.6349
	3	.5627	.6494	.6969	.7546	1.2647	1.3547	1.4369	1.6100
	5	.5600	.6473	.6951	.7533	1.2652	1.3544	1.4354	1.6047
	8	.5584	.6460	.6941	.7526	1.2656	1.3541	1.4346	1.6018
100	1	.6811	.7483	.7849	.8267	1.1846	1.2464	1.3029	1.4223
	3	.6761	.7440	.7805	.8241	1.1856	1.2456	1.2996	1.4108
	5	.6744	.7428	.7796	.8236	1.1857	1.2452	1.2987	1.4082
	8	.6734	.7421	.7791	.8232	1.1859	1.2450	1.2981	1.4067

TABLE A4: PEARSON CURVE CRITICAL VALUES (ASYMPTOTIC)

					α				
n	λ	.005	.025	.05	.10	.90	.95	.975	.995
10	1	.2169	.3477	.4216	.5140	1.4443	1.8188	2.0673	2.6656
	3	.1898	.3099	.3832	.4786	1.6108	1.8601	2.0998	2.6431
	5	.1928	.3070	.3775	.4720	1.6196	1.8692	2.1060	2.6340
	8	.1972	.3046	.3759	.4687	1.6259	1.8726	2.1084	2.6275
20	1	.3953	.5038	.5642	.6388	1.4114	1.5753	1.7342	2.0988
	3	.3674	.4789	.5424	.6219	1.4244	1.5828	1.7315	2.0573
	5	.3642	.4747	.5382	.6184	1.4271	1.5845	1.7307	2.0472
	8	.3630	.4728	.5366	.6163	1.4297	1.5848	1.7299	2.0398
50	1	.5786	.6612	.7062	.7608	1.2621	1.3560	1.4436	1.6341
	3	.5627	.6494	.6971	.7548	1.2649	1.3545	1.4370	1.6098
	5	.5599	.6473	.6951	.7534	1.2650	1.3545	1.4353	1.6044
	8	.5587	.6463	.6945	.7525	1.2653	1.3539	1.4346	1.6014
100	1	.6856	.7503	.7852	.8270	1.1852	1.2474	1.3042	1.4239
	3	.6763	.7440	.7804	.8240	1.1863	1.2459	1.2994	1.4110
	5	.6744	.7429	.7796	.8233	1.1860	1.2455	1.2987	1.4085
	8	.6733	.7420	.7794	.8234	1.1862	1.2447	1.2981	1.4069

TABLE A5: GRAM-CHARLIER CRITICAL VALUES (THREE EXACT MOMENTS)

n	λ	α = .005	.025	.05	.10	.90	.95	.975	.995
10	1	.2291	.3078	.3750	.4711	1.6649	1.9223	2.1094	2.4263
	3	.1618	.2581	.3357	.4433	1.6708	1.9292	2.1257	2.4618
	5	.1489	.2489	.3286	.4384	1.6720	1.9300	2.1279	2.4673
	8	.1418	.2438	.3247	.4357	1.6727	1.9304	2.1291	2.4703
20	1	.4127	.4855	.5419	.6184	1.4492	1.6200	1.7537	1.9847
	3	.3713	.4601	.5238	.6076	1.4459	1.6095	1.7437	1.9812
	5	.3626	.4550	.5202	.6055	1.4454	1.6075	1.7415	1.9800
	8	.3577	.4521	.5182	.6044	1.4452	1.6064	1.7402	1.9793
50	1	.5913	.6557	.6989	.7542	1.2722	1.3683	1.4496	1.5974
	3	.5694	.6446	.6917	.7505	1.2700	1.3616	1.4407	1.5884
	5	.5647	.6423	.6903	.7498	1.2696	1.3603	1.7415	1.9800
	8	.5621	.6411	.6894	.7494	1.2694	1.3595	1.4379	1.5852
100	1	.6924	.7482	.7822	.8243	1.1887	1.2517	1.3065	1.4097
	3	.6799	.7423	.7786	.8226	1.1875	1.2480	1.3010	1.4030
	5	.6773	.7411	.7778	.8223	1.1873	1.2473	1.3000	1.4015
	8	.6758	.7404	.7774	.8221	1.1872	1.2469	1.2994	1.4007

TABLE A6: GRAM-CHARLIER CRITICAL VALUES (THREE ASYMPTOTIC MOMENTS)

		α							
n	λ	.005	.025	.05	.10	.90	.95	.975	.995
10	1	.2406	.3131	.3770	.4698	1.6848	1.9478	2.1346	2.4506
	3	.1940	.2855	.3598	.4630	1.6496	1.9001	2.0898	2.4139
	5	.1828	.2792	.3559	.4615	1.6433	1.8901	2.0799	2.4054
	8	.1761	.2756	.3537	.4606	1.6399	1.8845	2.0741	2.4005
20	1	.4172	.4876	.5429	.6186	1.4524	1.6252	1.7593	1.9901
	3	.3830	.4698	.5322	.6144	1.4387	1.5999	1.7319	1.9653
	5	.3751	.4658	.5299	.6135	1.4362	1.5948	1.7260	1.9597
	8	.3705	.4635	.5285	.6130	1.4349	1.5920	1.7227	1.9565
50	1	.5924	.6562	.6991	.7542	1.2725	1.3690	1.4505	1.5983
	3	.5725	.6471	.6938	.7522	1.2683	1.3592	1.4379	1.5846
	5	.5681	.6451	.6927	.7518	1.2675	1.3573	1.4353	1.5816
	8	.5656	.6441	.6921	.7516	1.2670	1.3563	1.4338	1.5798
100	1	.6928	.7483	.7823	.8243	1.1888	1.2519	1.3067	1.4100
	3	.6810	.7432	.7793	.8232	1.1869	1.2472	1.3001	1.4017
	5	.6785	.7421	.7787	.8230	1.1865	1.2463	1.2988	1.3999
	8	.6771	.7415	.7784	.8229	1.1863	1.2458	1.2981	1.3989

TABLE A7: GRAM-CHARLIER CRITICAL VALUES (FOUR EXACT MOMENTS)

n	λ	α .005	.025	.05	.10	.90	.95	.975	.995
10	1	.2233	.3797	.4598	.5573	1.4398	2.0509	2.2601	2.5756
	3	.1123	.2918	.3868	.5010	1.5637	2.0050	2.2406	2.5873
	5	.1017	.2764	.3728	.4896	1.5831	1.9941	2.2324	2.5850
	8	.0965	.2681	.3651	.4832	1.5930	1.9881	2.2272	2.5829
20	1	.3518	.5031	.5733	.6553	1.3862	1.6631	1.8292	2.0707
	3	.3266	.4650	.5398	.6289	1.4162	1.6260	1.7896	2.0448
	5	.3223	.4584	.5336	.6238	1.4208	1.6202	1.7809	2.0375
	8	.3209	.4548	.5301	.6210	1.4232	1.6172	1.7759	2.0331
50	1	.5616	.6562	.7058	.7643	1.2592	1.3733	1.4709	1.6321
	3	.5502	.6439	.6951	.7561	1.2634	1.3630	1.4515	1.6110
	5	.5481	.6415	.6930	.7545	1.2641	1.3613	1.4478	1.6061
	8	.4823	.6138	.6919	.7536	1.2645	1.3604	1.4145	1.5670
100	1	.6785	.7474	.7843	.8280	1.1844	1.2524	1.3134	1.4252
	3	.6715	.7416	.7796	.8246	1.1853	1.2482	1.3045	1.4123
	5	.6701	.7405	.7787	.8240	1.1854	1.2474	1.3028	1.4095
	8	.6693	.7398	.7781	.8236	1.1855	1.2470	1.3018	1.4079

TABLE A8: GRAM-CHARLIER CRITICAL VALUES (FOUR ASYMPTOTIC MOMENTS)

n	λ	α .005	.025	.05	.10	.90	.95	.975	.995
10	1	.2461	.3181	.3815	.4737	1.6798	1.9409	2.1263	2.4400
	3	.1651	.2600	.3369	.4438	1.6728	1.9323	2.1288	2.4644
	5	.1474	.2481	.3281	.4382	1.6712	1.9287	2.1266	2.4662
	8	.1371	.2412	.3231	.4350	1.6703	1.9264	2.1250	2.4668
20	1	.4181	.4884	.5436	.6191	1.4517	1.6243	1.7582	1.9886
	3	.3724	.4606	.5241	.6078	1.4463	1.6102	1.7445	1.9820
	5	.3621	.4547	.5201	.6055	1.4453	1.6072	1.7411	1.9797
	8	.3561	.4513	.5178	.6042	1.4448	1.6055	1.7391	1.9782
50	1	.5924	.6563	.6992	.7543	1.2725	1.3689	1.4504	1.5982
	3	.5696	.6447	.6918	.7506	1.2701	1.3617	1.4408	1.5886
	5	.5646	.6423	.6902	.7498	1.2696	1.3602	1.4388	1.5863
	8	.5618	.6409	.6894	.7494	1.2694	1.3594	1.4377	1.5850
100	1	.6928	.7483	.7823	.8243	1.1888	1.2519	1.3067	1.4099
	3	.6799	.7423	.7786	.8226	1.1875	1.2481	1.3011	1.4030
	5	.6772	.7411	.7778	.8223	1.1873	1.2473	1.3000	1.4015
	8	.6757	.7404	.7774	.8221	1.1871	1.2469	1.2994	1.4006