SOFTWARE RELIABILITY MODELING AND EXPONENTIAL ORDER STATISTICS

by

Giovanni Andreatta and Gordon M. Kaufman

MIT Sloan School Working Paper 3114-90MS
January 1990

SOFTWARE RELIABILITY MODELING AND EXPONENTIAL ORDER STATISTICS

by

Giovanni Andreatta and Gordon M. Kaufman*

ABSTRACT: Properties of software failure times modelled as realizations of order statistics generated by independent but non-identically distributed exponential random variables are developed. Edgeworth and saddle point approximations to central order statistic densities so generated are developed using an exact integral representation of these densities. A comparison of Edgeworth and saddle point approximation with exact densities for two different population types is given. The accuracy of the saddle point approximation, even for very small population sizes ($N = 6$) and small samples ($n = 2$) is excellent.

The same technique is used to provide an exact integral representation of the probability that a particular fault appears in a sample of a given size. Some numerical comparisons of Rosén's (1972) approximation of inclusion probabilities with exact values are provided. His simple approximation appears to give excellent results as well.

The intimate connection between successive sampling theory and EOS models for software reliability is documented.

KEY WORDS: SOFTWARE RELIABILITY, SUCCESSIVE SAMPLING, EDGEWORTH APPROXIMATION, SADDLE POINT APPROXIMATIONS, INCLUSION PROBABILITY, ORDER STATISTICS

* Supported by AFOSR Contract #AFOSR-89-0371

1. INTRODUCTION

Goel (1985) has defined software reliability as the probability that during a prespecified testing or operational time interval, software faults do not cause a program to fail:

> "Let F be a class of faults, defined arbitrarily, and T be measure of relevant time, the units of which are dictated by the application at hand. Then the reliability of the software with respect to the class of faults F and with respect to the metric T, is the probability that no fault of the class occurs during the execution of the program for a prespecified period of relevant time."

Several classes of models have been proposed to capture this definition of reliability; among the most prominent are models built on the assumptions that waiting times between software failures are exponentially distributed and in addition are, conditional on knowledge of the appropriate parameter set, mutually independent. Such models have been called Exponential Order Statistics (**EOS**) models by Miller (1986) in his investigation of similarities of and differences between models based on the aforementioned assumptions. Littlewood (1981) was perhaps the first to challenge the assumption adopted by many authors that each fault "...contributes the same amount to the overall failure rate..." She posits a model in which (a) each fault possesses a parameter (occurrence rate) individual to that fault and (b) the collection of fault parameters is generated by a superpopulation process. This approach has the decisive advantage of avoiding some analytical and computational complexities that arise when assumption (b) is dropped. It is empirical Bayes in spirit and so is in formal correspondence with the Bayesian approach to reliability modeling adopted by Singpurwalla and his co-authors (Langberg and Singpurilla). However, Miller argues that Littlewood's model minus the assumption (b), a model that he calls a deterministic **EOS** model, "...has a certain physical motivation: the individual failure rates are physical quantities in the sense that they can be estimated to any desired degree of accuracy. The **IDOS** [empirical Bayes] and **NHPP** [non-homogeneous Poisson process] models are attractive because of mathematical tractibility and successful application experience; however, they are more difficult to motivate and verify in a physical sense." (Miller (1986), p. 12). In sum, some researchers view the **EOS** model as a first principles model that captures the physics of fault occurrence more accurately than the alternatives explored in the literature. This led Miller (1986) and Scholz (1986) to explore properties of order statistics generated by mutually independent but non-identically distributed random variables – the analytical concommitant of the **EOS** model.

The connection of this line of research with a sampling scheme well known to sample survey statisticians – successive sampling or sampling proportional to magnitude and without replacement from a finite population of magnitudes – has passed unnoticed until now. One of the purposes of this paper is to establish the nature of this connection. The problem of making inferences about unobserved finite population parameters of the **EOS** model based on observation of waiting times between failures and possibly the magnitude of observed faults is a dual of the problem of inference based on observation of fault magnitudes alone. The later problem has been investigated in detail by several authors (Andreatta and Kaufman (1986); Gordon (1989); Wang and Nair (1986); Bickel, Nair, and Wang (1989)). Other features of the link between software reliability models and successive sampling appear in a companion paper (Kaufman (1989b)).

Another purpose is to provide tools for the computation of the distribution of central order statistics for the **EOS** model and for the probability that a fault possessing a prespecified magnitude will be included in a sample of faults of a given size. Both play an important role in theories of inference for **EOS** models. The distribution of the waiting time to occurrence of the nth fault is an analytical benchmark for understanding properties of the **EOS** model and for a theory of unbiased estimation of the empirical distribution of magnitudes of unobserved faults and of the number of faults remaining in the software system.

Gordon (1982) has shown that the distribution of permutations of the order in which successively sampled elements of a finite population are observed can be characterized in terms of exponential waiting times with expectations inversely proportional to magnitudes of the finite population elements. This leads naturally to a corollary interpretation of the probability that a particular element of the population will be included in a sample as the expectation of an exponential function of an order statistic generated by independent but non–identically distributed exponential random variables (***rvs***).

In Section 3 we present an exact integral representation of the marginal density of an order statistic so generated. The integrand is interpretable as a probability mixture of characteristic functions of sums of conditionally independent Bernoulli ***rvs***, an interpretation that suggests a first approximation of the density, and the form that leading terms in Edgeworth and saddle–point approximations will take.

An Edgeworth type approximation is presented in Section 4. While this expansion could in principle be derived by first computing a saddle–point approximation and then using the idea of recentering a conjugate distribution as suggested by Daniels (1954), we have chosen to compute it directly.

As the "large" parameter N appears in the integrand of this representation, both as the number of terms in a product and in a sum, the integral representation (Lemma 3.1) of this density is not of "standard" form in which, the integrand is expressible as $\exp\{Ng(t)\}$, $g(t)$ functionally independent of N. Nevertheless, conditions for application of Watson's lemma hold and the steepest descent method produces valid results. A saddle–point approximation is presented in Section 5. The form of the order $1/N^2$ correction was checked using MACSYMA (Project MAC Symbolic Manipulation system), a large computer program designed to manipulate algebraic expressions, symbolically integrate and differentiate, as well as carry out manifold other mathematical operations. The $1/N^2$ term computed via MACSYMA is in correspondence with (6.2) in Good (1956) who made the prescient statement:

> "... we have calculated the third term $[0(N^{-2})]$ asymptotic series. More terms could be worked out on an electronic computer programmed to do algebra."

When magnitudes of finite population elements are identical, the leading term of the steepest descent approximation (cf. (5.15)), upon renormalization, reproduces the exact density of the nth smallest order statistic generated by $N \geq n$ mutually independent and identically distributed exponential *rvs*.

Numerical examples appear in Section 6. The accuracy displayed by use of $0(1/N)$ corrections to the leading term of the saddle point approximation, even for small finite population sizes ($N = 6, 10$), suggests that $0(1/N^2)$ corrections are only of curiousity value in these examples. Field and Hampel (1982) call saddle point type approximations accurate for very small samples "small sample asymptotic" approximations. In the comparisons made here, no renormalization to unity of the approximations is done. This additional step would further improve the already excellent accuracy of the saddle point approximations.

2. SUCCESSIVE SAMPLING

We consider a finite population consisting of a collection of N uniquely labelled units. Let k denote the label of the kth unit and define $U = \{1, 2, \ldots, k, \ldots, N\}$. Associated with the unit labelled k is an attribute – magnitude –that takes on a bounded value $y_k > 0$; $\underline{y}_N \equiv (y_1, \ldots, y_N)$ is a parameter of U. An ordered sample of size $n \leq N$ is a sequence $\underline{s}_n = (k_1, \ldots, k_n)$ of labels $k_i \epsilon U$. Successive sampling of U is sampling without replacement and proportional to magnitude, and is generated by the following sampling scheme: for $n = 1, 2, \ldots, N$, the probability that the *rv* $\underline{\tilde{s}}_n$ assumes value $\underline{s}_n$ in the set $\{k_1, \ldots, k_n) \mid k_j \epsilon U,\ k_i \neq k_j \text{ if } i \neq j\}$ of all possible distinct sequences with n elements is, setting $R_N = y_1 + \cdots + y_N$,

$$P\{\underline{\tilde{s}}_n = \underline{s}_n \mid \underline{y}_N\} = \prod_{j=1}^{n} y_{k_j} / [R_N - (y_{k_1} + \cdots + y_{k_{j-1}})] \tag{2.1}$$

with $y_{k_0} = 0$.

Let $\tilde{X}_1, \ldots, \tilde{X}_N$ be mutually independent exponential *rvs* with common mean equal to one. Then (Gordon (1982)),

$$P\{\underline{\tilde{s}}_N = (1, 2, \ldots, N) \mid \underline{y}_N\} = P\left\{\frac{\tilde{X}_1}{y_1} < \frac{\tilde{X}_2}{y_2} < \cdots < \frac{\tilde{X}_N}{y_N}\right\}. \tag{2.2}$$

Upon defining $\tilde{Z}_k = \tilde{X}_k / y_k$ and $\tilde{Z}_{(k)}$ such that $\tilde{Z}_{(1)} \leq \tilde{Z}_{(2)} \leq \cdots \leq \tilde{Z}_{(k)} \leq \cdots \leq \tilde{Z}_{(N)}$, the kth element of U will appear in a sample of size n if and only if $\tilde{Z}_k \leq \tilde{Z}_{(n)}$. Defining $I_{\{\tilde{X}_k > \tilde{Z}_{(n)} y_k\}}$ as the indicator function assuming value one if $\tilde{X}_k > \tilde{Z}_{(n)} y_k$ and zero otherwise, the probability $\pi_k(n)$ that element $k \epsilon \underline{\tilde{s}}_n$ is

$$\pi_k(n) = 1 - E\left[I_{\{\tilde{X}_k > \tilde{Z}_{(n)} y_k\}}\right] = 1 - E\left(e^{-y_k \tilde{Z}_{(n)}}\right). \tag{2.3}$$

Together with the identity $\sum_{k=1}^{N} \pi_k(n) = n$, (2.3) affords a simple motivation for Rosén's (1972) approximation to $\pi_k(n)$: for $x \epsilon (0, \infty)$, $C(x) = \sum_{k=1}^{N} \exp\{-y_k x\}$ decreases monotonically as x increases. Consequently there is a unique value $Z_{n,N}$ of $\tilde{Z}_{(n)}$ for which $C(Z_{n,N}) = N - n$. Rosén's approximation to $\pi_k(n)$ is $1 - \exp\{-y_k Z_{n,N}\}$. Hájek (1981) presents some numerical examples illustrating the accuracy of this approximation.

3. AN INTEGRAL REPRESENTATION OF THE MARGINAL DENSITY OF $\widetilde{Z}_{(n)}$

The marginal density $f_{Z_{(n)}}(\lambda)$ of $\widetilde{Z}_{(n)}$ is concentrated on $(0,\infty)$ and possesses the following integral representation:

Lemma 3.1: For arbitrary positive values of $y_1,\ldots,y_N$ and $\lambda\epsilon(0,\infty)$ the marginal density $f_{Z_{(n)}}(\lambda)$ of $\widetilde{Z}_{(n)}$, $n = 1,2,\ldots,N$, is equal to

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i(n-1)u}\prod_{j=1}^{n}\left[e^{-\lambda y_j}+\left(1-e^{-\lambda y_j}\right)e^{iu}\right]$$
$$\times\sum_{k=1}^{N} y_k e^{-\lambda y_k}\Big/\left[e^{-\lambda y_k}+\left(1-e^{-\lambda y_k}\right)e^{iu}\right]du \tag{3.1}$$

Proof: For $k = 1,2,\ldots,N$, the probability that $\widetilde{Z}_{(n)} = \widetilde{Z}_k$ is

$$\sum{}' P\left(\max\{\widetilde{Z}_{i_1},\ldots,\widetilde{Z}_{i_{n-1}}\} < \widetilde{Z}_k < \min\{\widetilde{Z}_{i_{n+1}},\ldots,\widetilde{Z}_{i_N}\}\right) \tag{3.2}$$

where $\sum{}'$ denotes summation over $\binom{N-1}{n-1}$ distinct partitions of $\{1,2,\ldots,k-1,k+1,\ldots,N\}$ into two subsets with $n-1$ and $N-n$ elements respectively. Given $\widetilde{Z}_k = \lambda$, a generic term is

$$P\left(\max\{\widetilde{Z}_{i_1},\ldots,\widetilde{Z}_{i_{n-1}}\}< \lambda < \min\{\widetilde{Z}_{i_{n+1}}\ldots,\widetilde{Z}_{i_N}\}\right)$$
$$=\left(\prod_{j=1}^{n-1}\left[1-e^{-\lambda y_{i_j}}\right]\right)\left(\prod_{\ell=n+1}^{N} e^{-\lambda y_{i_\ell}}\right). \tag{3.3}$$

Consequently, given $\widetilde{Z}_k = \lambda$ the probability (3.2) is

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i(n-1)u}\prod_{\substack{j=1\\ j\neq k}}^{N}\left[e^{-\lambda y_j}+\left(a-e^{-\lambda y_j}\right)e^{iu}\right]du. \tag{3.4}$$

As the marginal density of $\widetilde{Z}_k$ is $y_k\exp\{-\lambda y_k\}$, multiplying (3.3) by this density and summing over the N possibilities for the nth smallest among $\widetilde{Z}_1,\ldots,\widetilde{Z}_N$, the density of $\widetilde{Z}_{(n)}$ is as shown in (3.1). ▯

The integral (3.1) is the principal vehicle for computation of approximations to $f_{Z_{(n)}}$. To motivate these approximations we begin with an interpretation of the integrand in random variable (rv) terminology. The integrand of (3.1) is the characteristic function of a mixture of characteristic functions of sums of conditionally independent rvs. Appropriately scaled and properly centered, a non-equal components version of the central limit theorem applies. This interpretation suggests a "normal-like" approximation to $f_{Z_{(n)}}$.

In what follows the infinite sequence $y_1, \ldots, y_k, \ldots$ shall be regarded as a fixed sequence of positive bounded numbers. In our setting there is no loss in generality in rescaling the $y_k s$. For each finite sequence $y_1, \ldots, y_N$, define $p_{kN} = y_k/(y_1 + \cdots + y_N)$ so that $\sum_{k=1}^{N} p_{kN} = 1$. We assume throughout that $\max_k p_{kN} \to 0$ as $N \to \infty$, a condition that asserts itself in statements about the order of functions such as $\frac{1}{N} \sum_{k=1}^{N} \exp\{-\lambda p_{kN}\}[1 - \exp\{-\lambda p_{kN}\}]$ for N large. In order to simplify notation we shall suppress explicit display of the triangular array p_{kN}, $k = 1, 2, \ldots, N$ for $N = 1, 2, \ldots$ and let it be understood that for given N, $y_k \equiv p_{kN}$ is scaled as stated. A statement that, for example, the aforementioned function is of order one as $N \to \infty$ implies an appropriate balance between the rates at which p_{kN}, $k = 1, 2, \ldots, N$ approach zero as $N \to \infty$ and the value of λ. This avoids a cataloguing of special cases but exercises the sin of omitting precise details. To illustrate details in one case, assume that $n/N = f$ is fixed as $N \to \infty$ and that there exists a constant $\epsilon > 0$ independent of N such that $\epsilon \leq \min y_k / \max y_k$ (cf. Hájek (1981), for example). Then it is easy to show that the solution λ_N to $\sum_{k=1}^{N} \exp\{-\lambda y_k\} = N - n$ is $0(N)$ and that $\sum_{k=1}^{N} \exp\{-\lambda_N y_k\}[1 - \exp\{-\lambda_N y_k\}] = 0(N)$.

To facilitate discussion, at $\widetilde{Z}_{(n)} = \lambda$ define $a_k(\lambda) = \exp\{-\lambda y_k\}$; at times we regard λ as fixed and write a_k in place of $a_k(\lambda)$ for notational convenience when doing so.

The integrand of (3.1) may be interpreted as a probability mixture of characteristic functions times the characteristic function of a point mass at $n - 1 \equiv Np \equiv N(1-q)$. So doing leads to approximations that mimic the leading term of Edgeworth type expansions of the density $f_{Z_{(n)}}(\lambda)$. With $\theta_k = a_k y_k / \sum_{k=1}^{N} a_k y_k$, this integrand is

$$\left(\sum_{k=1}^{N} a_k y_k\right) \times e^{-iNpu} \sum_{k=1}^{N} \theta_k \prod_{\substack{j=1 \\ j \neq k}}^{N} (a_j + (1 - a_j)e^{iu}). \tag{3.5}$$

For fixed $u, a_j + (1-a_j)e^{iu}$ is the characteristic function of a *rv* $\widetilde{W}_j$ taking value 1 with probability $1-a_j$ and 0 with probability a_j. Consequently, $\zeta_{kN}(u) = \prod_{\substack{j=1 \\ j\neq k}}^{N} (a_j + (1-a_j)e^{iu})$ is the characteristic function of a sum $\widetilde{\Delta}_{kn} = (\widetilde{W}_1 + \cdots \widetilde{W}_N) - \widetilde{W}_k$ of $N-1$ independent *rvs* that can assume values $0, 1, \ldots, N-1$.

As $\widetilde{W}_j$ has mean $(1-a_j)$ and variance $a_j(1-a_j)$, $\widetilde{\Delta}_{kN}$ has mean $\overline{\Delta}_{kN} = \left[\sum_{j=1}^{N}(1-a_j)\right] - (1-a_k)$ and variance $v_{kN} = \left[\sum_{j=1}^{N} a_j(1-a_j)\right] - a_k(1-a_k)$. If $v_{kN} \longrightarrow \infty$ as $N \longrightarrow \infty$, the sequence of *rvs* composing $\widetilde{\Delta}_{kN} - \overline{\Delta}_{kN}$ fulfill the Lindeberg condition, so at atoms of the distribution of $\widetilde{\Delta}_{kN}$, $P\{\widetilde{\Delta}_{kN} = x\}$ can be approximated by a normal density with mean $\overline{\Delta}_{kN}$ and variance v_{kN}.

Consider a discrete valued *rv* $\widetilde{B}_N$ with range $\{1, 2, \ldots, N\}$ and probability function $P\{\widetilde{B}_N = k\} = \theta_k$, $k = 1, 2, \ldots, N$. In terms of $\widetilde{B}_N$ and $\widetilde{\Delta}_{kN}$, $k = 1, 2, \ldots, N$, the mixture $\sum_{k=1}^{N} \theta_k \zeta_{kN}(u)$ represents a *rv* $\widetilde{T}_N$ such that $\widetilde{T}_N \mid (\widetilde{B}_N = k) = \widetilde{\Delta}_{kN}$ for $k = 1, 2, \ldots, N$ so upon approximating $P\{\widetilde{\Delta}_{kN} = x\}$ at its atoms as stated, at atoms of $\widetilde{T}_N$, $P\{\widetilde{T}_N = x\}$ is approximable by a probability mixture of normal densities with means $\overline{\Delta}_{kN}$ and variances v_{kN}, $k = 1, 2, \ldots, N$. Since v_{kN} and v_{jN}, $j \neq k$ differ by at most $1/4$ and $\overline{\Delta}_{kN}$ and $\overline{\Delta}_{jN}$ differ by at most one, when $v_{kN} \longrightarrow \infty$, $k = 1, 2, \ldots, N$, the probability function of $\widetilde{T}_N - (n-1)$ is in turn approximable to the same order of accuracy by $\sum_{k=1}^{N} a_k y_k$ times a single normal density with mean $E(\widetilde{T}_N) - (n-1)$ and variance $\mathrm{Var}(\widetilde{T}_N)$. The expectation of $\widetilde{T}_N$ is

$$E(\widetilde{T}_N) = E_{B_N} E(\widetilde{T}_N \mid \widetilde{B}_N) = \sum_{k=1}^{N} \theta_k E(\widetilde{\Delta}_{kN}) = \sum_{k=1}^{N} (1-\theta_k)(1-a_k) \tag{3.6}$$

and its variance

$$\begin{aligned} \mathrm{Var}(\widetilde{T}_N) &= E_{B_N} \mathrm{Var}(\widetilde{T}_N) \mid \widetilde{B}_N) + \mathrm{Var}_{B_N} E(\widetilde{T}_N \mid \widetilde{B}_N) \\ &= \sum_{k=1}^{N} \left[(1-\theta_k) a_k (1-a_k) + \theta_k \left(a_k - \sum_{k=1}^{N} \theta_j a_j \right)^2 \right]. \end{aligned} \tag{3.7}$$

Upon accounting for the point mass at $n-1$, an approximation to the integral (3.1) emerges: with $a_k \equiv \exp\{-\lambda y_k\}$

$$f_{Z_{(n)}}(\lambda) \simeq \frac{\sum_{k=1}^{N} a_k y_k}{\sqrt{2\pi \mathrm{Var}(\widetilde{T}_N)}} e^{-\frac{1}{2}(n-1-E(\widetilde{T}_N))^2/\mathrm{Var}(\widetilde{T}_N)}. \tag{3.8}$$

The approximation (3.8) turns out to be identical to the leading term of the Edgeworth type approximation studied next.

The above approximation links up with Sen's (1968) study of sample quantiles for m–dependent processes when $m = 0$ (Theorem 2.1 and (2.3) of Sen) in the following fashion: if there exists an $\epsilon > 0$ independent of N such that $\epsilon \le \min y_k / \max y_k$, then using Lemma 4.1 of the next section, the solution λ_N to $N - n = \sum_{k=1}^{N} \exp\{-\lambda y_k\} \equiv N g_N(\lambda)$ is $0(N)$ and for $\mid \lambda - \lambda_N \mid = o(N)$,

$$\frac{1}{N} E(\widetilde{T}_N) = (\lambda - \lambda_N) g'_N(\lambda_N) + o(1) \tag{3.9}$$

and

$$\mathrm{Var}(\widetilde{T}_n) = g_N(\lambda_N) - g_N(2\lambda_N) + o(1) \tag{3.10}$$

as $N \longrightarrow \infty$ with $n/N \longrightarrow f$ fixed. Consequently as $N \longrightarrow \infty$, the density f_{U_N} of $\widetilde{U}_N = N(\widetilde{\lambda} - \lambda_N) \mid g'_N(\lambda_N) \mid / [g_N(\lambda_N) - g_N(2\lambda_N)]^{\frac{1}{2}}$ approaches $\exp\{-\frac{1}{2} U_N^2\} / \sqrt{2\pi}$.

4. AN EDGEWORTH APPROXIMATION OF $f_{Z_{(n)}}$

The preceding discussion provided an heuristic approximation for $f_{Z_{(n)}}$. We next compute an Edgeworth type expansion of (3.1) and show that the leading term can be presented in the form (3.8).

Since Edgeworth expansions exhibit notoriously bad behavior in the tails, we restrict the expansion to an interval in λ for which $\left(N-n-\sum_{k=1}^{N} a_k\right)/\left[\sum_{j=1}^{N} a_j(1-a_j)\right]^{1/2} = 0(1)$. Conditions defining such intervals are given in the following

Lemma 4.1: Let λ_N be a solution to $\frac{1}{N}\sum_{k=1}^{N} \exp\{-\lambda y_k\} = 1-(n/N)$. If $n/N = f$ is fixed as $N \longrightarrow \infty$, and there exists a constant $\epsilon > 0$ independent of N such that $\epsilon \le \min y_k/\max y_k$, then defining $M_N(\lambda) = \frac{1}{N}E(\widetilde{T}_N) - p = 1 - f - \frac{1}{N}\sum_{k=1}^{N}(1-\theta_k)a_k$, and $V_N(\lambda) = \operatorname{Var}(\widetilde{T}_N)$, $NM_N(\lambda)/V_N^{1/2}(\lambda) = 0(1)$ implies that there exists a positive constant $c = 0(1)$ independent of N such that

$$\lambda_N - c\sqrt{N} < \lambda < \lambda_N + c\sqrt{N}. \tag{4.1}$$

Proof: $0 < \epsilon \le \min y_k/\max y_k$ implies that $\epsilon/N \le y_k \le 1/N\epsilon$, $k = 1,2,\ldots,N$. As $\frac{1}{N}\sum_{k=1}^{N}\exp\{-\lambda_N y_k\} = 1-f$, $\exp\{-\lambda_N/N\epsilon\} \le 1-f \le \exp\{-\lambda_N\epsilon/N\}$, so for N large, $\lambda_N = 0(N)$. In addition

$$Ne^{-\lambda/\epsilon N}(1-e^{-\lambda\epsilon/N}) \le \sum_{k=1}^{N} e^{-\lambda y_k}(1-e^{-\lambda y_k}) \le Ne^{-\lambda\epsilon/N}(1-e^{-\lambda/N\epsilon}).$$

Letting $\delta = \lambda - \lambda_N$,

$$\left|(1-f) - \frac{1}{N}\sum_{k=1}^{N} e^{-\lambda y_k}\right| = \left|(1-f) - \frac{1}{N}\sum_{k=1}^{N}\left(e^{-\lambda_N y_k}e^{-\delta y_k}\right)\right|$$

$$\le \begin{cases} (1-f)\,|\,1-e^{-\delta/N\epsilon}\,| & \text{if} \quad \delta > 0 \\ (1-f)\,|\,1-e^{-\delta/\epsilon N}\,| & \text{if} \quad \delta \le 0. \end{cases}$$

Define $\eta(\lambda) = M_N(\lambda)/V_N^{1/2}(\lambda)$ and consider $\delta = \lambda - \lambda_N > 0$. Then since

$| M_N(\lambda) \, |\leq (1-f)(1-e^{-\delta/N\epsilon})$ if $\delta > 0$, and $V_N(\lambda) \leq Ne^{-\lambda/\epsilon N}(1-e^{-\lambda\epsilon/N})$, $| \eta(\lambda) \, |\leq \sqrt{N}(1-f)(1-e^{-\delta/N\epsilon})/[e^{-\lambda/\epsilon N}(1-e^{-\lambda\epsilon/N})]^{1/2}$. Consider $\delta = o(N)$: then

$$| \eta(\lambda) \, | \leq (1-f)e^{\lambda_N/2N\epsilon}\left[\frac{\delta}{\sqrt{N\epsilon}} + 0\left(\frac{\delta^2}{N^{3/2}}\right)\right] \Big/ \left[e^{-\delta/N\epsilon}\left(1 - e^{-\lambda_N\epsilon/N} \cdot e^{-\delta\epsilon/N}\right)\right]^{1/2}$$

or

$$| \eta(\lambda) \, | \leq (1-f)e^{\lambda_N/2\epsilon N}\left[\frac{\delta}{\sqrt{N\epsilon}} + 0\left(\frac{\delta^2}{N^{3/2}}\right)\right]\left[1 - e^{-\lambda_N\epsilon/N}\right]^{1/2}[1+o(1)]^{1/2}.$$

Since $\lambda_N = 0(N)$,

$$| \eta(\lambda) \, | \leq C_N\left[\frac{\delta}{\sqrt{N\epsilon}} + 0\left(\frac{\delta^2}{N^{3/2}}\right)\right], \quad C_N = 0(1).$$

A similar argument for $\delta \leq 0$ gives

$$| \eta(\lambda) \, | \leq C'_N\left[\frac{\delta_\epsilon}{\sqrt{N}} + 0\left(\frac{\delta^2}{N^{3/2}}\right)\right], \quad C'_N = 0(1).$$

Hence $| \eta(\lambda) \, |= 0(1)$ iff $\delta = 0(\sqrt{N})$. []

We next present an Edgeworth–type approximation to $f_{z_{(n)}}(\lambda)$. Writing $a_j(\lambda)$ as a_j for notational convenience, define the cumulant functions

$$\kappa_{3N}(\lambda) = -\sum_{j=1}^{N} a_j(1-a_j)(1-2a_j), \tag{4.1a}$$

$$\kappa_{4N}(\lambda) = \sum_{j=1}^{N} a_j(1-a_j)(1-6a_j+6a_j^2), \tag{4.1b}$$

and

$$d_{3N}(\lambda) = \sum_{j=1}^{N} \theta_j(1-a_j)(1-6a_j+6a_j^2), \tag{4.1c}$$

$$d_{4N}(\lambda) = \sum_{j=1}^{N} \theta_j(1-a_j)(1+6a_j-24a_j^2+16a_j^3). \tag{4.1d}$$

Let $H_\ell(x)$ be Hermite polynomials; e.g. $H_3(x) = x^3 - 3x$, $H_4(x) = x^4 - 6x^2 + 3$, and $H_6(x) = x^6 - 15x^4 + 45x^2 - 15$.

Theorem 4.1: *For $f = n/N$ fixed as $N \longrightarrow \infty$, when $NM_N(\lambda)/V_N^{1/2}(\lambda) = 0(1)$ or smaller.*

$$f_{Z_{(n)}}(\lambda) = \frac{\sum_{k=1}^{N} a_k y_k}{\sqrt{2\pi V_N(\lambda)}} e^{-\frac{1}{2}N^2 M_N^2(\lambda)/V_N(\lambda)}$$

$$\times \left\{1 - \frac{1}{6}\left[\frac{\kappa_{3N}(\lambda) + d_{3N}(\lambda)}{V_N^{3/2}(\lambda)}\right] H_3(NM_N(\lambda)/V_N^{1/2}(\lambda))\right.$$

$$+ \frac{1}{72}\left[\frac{\kappa_{4N}(\lambda) + d_{4N}(\lambda)}{V_N^2(\lambda)} H_4(NM_N(\lambda)/V_N^{1/2}(\lambda))\right.$$

$$\left.\left. - \frac{(\kappa_{3N}(\lambda) + d_{3N}(\lambda))^2}{V_N^3(\lambda)} H_6(NM_N(\lambda)/V_N^{1/2}(\lambda))\right] + 0(N^{-3/2})\right\}. \quad (4.2)$$

Before turning to the proof, observe that $NM_N(\lambda)/V_N^{1/2} = 0(1)$ maintains the order of the argument of H_3, H_4, and H_6 at $0(1)$; $\kappa_{3N}(\lambda)$ and $\kappa_{4N}(\lambda)$ are of order N at most, and $d_{3N}(\lambda)$ and $d_{4N}(\lambda)$ are of order one at most. Thus the coefficients of correction terms in (4.3) are of orders $N^{-1/2}$ and N^{-1} respectively.

The magnitudes of coefficients are more clearly revealed by reexpressing them in a form suggested by Hájek (1981): Let $v_N(\lambda) = \sum_{k=1}^{N} a_k(1 - a_k)$. Then

$$-\kappa_{3N}(\lambda) = v_N(\lambda)\left\{1 - \frac{\sum_{j=1}^{N} a_j^2(1 - a_j)}{2v_N(\lambda)}\right\}$$

and

$$\kappa_{4N}(\lambda) = v_N(\lambda)\left\{1 - \frac{\sum_{j=1}^{N} a_j^2(1 - a_j)^2}{6v_N(\lambda)}\right\}$$

from which it is apparent $\kappa_{3N}(\lambda)/V_N^{3/2}(\lambda) = 0(1/v_N^{1/2}(\lambda))$ and $\kappa_4 N(\lambda)/V_N^2(\lambda) = 0(1/v_N(\lambda))$, since $V_N(\lambda)$ and $v_N(\lambda)$ are of the same order of magnitude.

Proof: In (3.1) let $\overline{a}_N = \frac{1}{N}\sum_{j=1}^{N} a_j$ and $M_{kN} = 1 - f - \overline{a}_N + (a_k/N)$. Then with

$$\zeta_N(s) = \sum_{k=1}^{N} a_k y_k e^{iNM_{kN}u} \prod_{\substack{j=1 \\ j\neq k}}^{N} \left[a_j e^{-i(1-a_j)u} + (1 - a_j)e^{ia_j u}\right],$$

$$= e^{iN(1-f-\overline{a}_N)u} \prod_{j=1}^{N} \left[a_j e^{-i(1-a_j)u} + (1-a_j)e^{ia_j u}\right], \tag{4.3}$$

$$\times \sum_{k=1}^{N} \frac{a_k y_k}{a_k e^{-iu} + (1-a_k)}$$

(3.1) is

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \zeta_N(u) du. \tag{4.4}$$

To account for contributions from the tails of (4.3), observe that for $-\pi < u < \pi$,

$$\begin{aligned} &\mid a_k e^{-i(1-a_k)u} + (1-a_k)e^{ia_k u} \mid^2 \\ &= \{1 - 2a_k(1-a_k)(1-\cos u)\} \leq 1 - \frac{a_k(1-a_k)u^2}{\pi^2} \\ &\leq e^{-a_k(1-a_k)u^2/\pi^2}. \end{aligned} \tag{4.5}$$

As a consequence,

$$\mid \zeta_N(u) \mid \leq \sum_{k=1}^{N} a_k y_k e^{-\frac{1}{2}[v_N(\lambda) - a_k(1-a_k)]u^2/\pi^2} \tag{4.6}$$

for $-\pi < u < \pi$. Since $\sum_{k=1}^{N} a_k y_k \leq 1$ by virtue of our scaling assumption, $\mid \zeta_N(u) \mid \leq e^{-vu^2/\pi}$ for some $v > 0$, and when $v_N(\lambda) = 0(N)$, $\zeta_N(u) \longrightarrow 0$ outside the origin faster than any power of $1/N$. This property of $\zeta_N(u)$ permits an Edgeworth type expansion of the integral (3.1).

Albers et al. (1976), p. 115, justify Taylor expansion of product terms like those in (4.3) and a corresponding Edgeworth expansion as follows: for $-\pi/2 \leq x \leq \pi/2$, the real part of $a_j \exp\{-i(1-a_j)x\} + (1-a_j)\exp\{ia_j x\}$ is $\geq \frac{1}{2}$, so

$$\begin{aligned} &\log\left[a_j e^{-i(1-a_j)x} + (1-a_j)e^{ia_j x}\right] \\ &= -a_j(1-a_j)(x^2/2) + a_j(1-a_j)(1-2a_j)(ix^3/6) \\ &\quad + a_j(1-a_j)(1-6a_j+6a_j^2)(x^4/24) \\ &\quad + \beta_j(x)a_j(1-a_j)(1+6a_j-24a_j^2+16a_k^3)(x^5/120) \end{aligned} \tag{4.7}$$

where $|\beta_j(x)| \leq 1$ for $-\frac{1}{2}\pi < x < \frac{1}{2}\pi$. Letting $c_{j\ell}$ denote the ℓth cumulant arising from a single Bernoulli trial with probability a_j, (4.7) can be displayed as

$$\log\left[a_j e^{-i(1-a_j)x} + (1-a_j)e^{ia_j x}\right]$$

$$= -\, c_{j2}(x^2/2) + c_{j3}(ix^3/6) + c_{j4}(x^4/24) + c_{j5}\beta_j(x)(x^5/120) \tag{4.8}$$

with $|\beta_j(x)| \leq 1$. Provided that $|x| \leq \frac{\pi}{2}$, each term in (4.3) of this form can be so exapnded and

$$\prod_{j=1}^{N}\left[a_j e^{-i(1-a_j)u} + (1-a_j)e^{ia_j u}\right]$$

$$= \exp\{-v_N(\lambda)(u^2/2) + \kappa_{3N}(\lambda)(iu^3/6) \tag{4.9}$$

$$+ \kappa_{4N}(\lambda)(u^4/24) + \beta(u)\kappa_{5N}(\lambda)(iu^5/120)\}.$$

We next expand in Taylor series

$$B_N(iu) \equiv \sum_{k=1}^{N} \frac{a_k y_k}{a_k e^{-iu} + (1-a_k)} \tag{4.10}$$

and combine this expansion with (4.9) so that the resulting approximation to $\zeta_N(u)$ is in a form leading to (4.2).

The function $B_N(iu)$ has a useful property:

$$-ie^{-iu}B_N(iu) = \sum_{k=1}^{N} y_k \frac{d}{du}\log\left[a_k e^{-iu} + (1-a_k)\right]. \tag{4.11}$$

For $u\epsilon(-\pi,\pi)$, $a_k \exp\{-iu\} + (1-a_k)$ is analytic and possesses no singularities. As a result, asymptotic expansions of B_N can be differentiated.

Expand each logarithmic term in (4.11) using

$$\log\left(1 + \sum_{\ell=1}^{\infty} \frac{b_\ell}{\ell!}(iu)^\ell\right) \equiv \sum_{\ell=1}^{\infty} \frac{d_\ell}{\ell!}(iu)^\ell \tag{4.12a}$$

with

$$d_1 = b_1, \quad d_2 = b_2 - b_1^2, \quad d_3 = b_3 - 3b_1b_2 + 2b_1^3, \tag{4.12b}$$

$$d_4 = b_4 - 3b_2^2 - 4b_1b_3 + 12b_1^2b_2 - 6b_1^4, \tag{4.12c}$$

$$d_5 = b_5 - 10b_2b_3 - 5b_1b_4 + 30b_1b_2 + 20b_1^2b_3 - 60b_1^3b_2 + 24b_1^5.$$

(cf. Kendall and Stuart (1969) Vol. I p. 70.)

Differentiation with respect to u yields

$$
\begin{aligned}
B_N(iu) = \left(\sum_{k=1}^{N} a_k y_k\right) e^{iu} \Big\{ 1 - & \sum_{k=1}^{N} \theta_k (1-a_k)(iu) \\
& + \sum_{k=1}^{N} \theta_k (1-a_k)(1-2a_k)(iu)^2/2 \\
& - \sum_{k=1}^{N} \theta_k (1-a_k)(1-6a_k+6a_k^2)(iu)^3/6 \\
& + \sum_{k=1}^{N} \theta_k (1-a_k)(1-14a_k+36a_k^2-24a_k^3)(iu)^4/24 \\
& + \ldots \Big\}.
\end{aligned}
\tag{4.13}
$$

Next, exponentiate the term in curly brackets in (4.13) using (4.12) again with

$$
b_1 = -\sum_{k=1}^{N} \theta_k (1-a_k), \quad b_2 = \sum_{k=1}^{N} \theta_k (1-a_k)(1-2a_k), \tag{4.14a}
$$

$$
b_3 = -\sum_{k=1}^{N} \theta_k (1-a_k)(1-6a_k+6a_k^2), \tag{4.14b}
$$

and

$$
b_4 = \sum_{k=1}^{N} \theta_k (1-a_k)(1-14a_k+36a_k^2-24a_k^3). \tag{4.14}
$$

Then b_1, b_2, b_3, b_4 are probability mixtures of terms of order one or smaller, so

$$
\left| \log\left(1 + \sum_{\ell=1}^{\infty} \frac{b_\ell}{\ell!}(iu)^\ell\right) - \sum_{k=1}^{4} \frac{d_k}{k!}(iu)^k \right| \leq C \mid u \mid^5 /120 \tag{4.16}
$$

where C is constant of order one or smaller.

Upon assembling terms of the same order in the expansions (4.13) of $B_N(iu)$ and (4.8) of the product terms and changing variables to $z = [V_N(\lambda)]^{1/2} u$,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\zeta_N(u)du = \frac{\sum_{k=1}^{N} a_k y_k}{2\pi\sqrt{V_N(\lambda)}}\int_{-\pi\sqrt{V_N(\lambda)}}^{\pi\sqrt{V_N(\lambda)}} \exp\Big\{i[NM_N(\lambda)/\sqrt{V_N(\lambda)}]z - \frac{1}{2}z^2$$

$$+\frac{1}{6}\left[\frac{\kappa_{3N}(\lambda)+d_{3N}(\lambda)}{V_N^{3/2}(\lambda)}\right]iz^3 - \frac{1}{24}\left[\frac{\kappa_{4N}(\lambda)+d_{4N}(\lambda)}{V_N^{2}(\lambda)}\right]z^4$$

$$+\frac{\nu(z)}{120}\left[\frac{\kappa_{5N}(\lambda)+d_{5N}(\lambda)}{V_N^{3/2}(\lambda)}\right]iz^5\Big\}dz \tag{4.17}$$

for some $|\nu(z)| \leq 1$.

Since $c_{2k} = a_k(1-a_k)$, $c_{3k} = a_k(1-a_k)(2a_k-1)$, $c_{4k} = a_k(1-a_k)(1-6a_k+6a_k^2)$, and $c_{5k} = a_k(1-a_k)(1+6a_k-24a_k^2+16a_k^3)$ are cumulants arising from a single binomial trial with probability a_k, $\kappa_{3N}(\lambda)/V_N^{3/2}(\lambda) \leq 1/V_N^{1/2}(\lambda)$, $\kappa_{4N}(\lambda)/V_N^2(\lambda) \leq 1/V_N(\lambda)$, and $\kappa_{5N}(\lambda)/V_N^{5/2}(\lambda) \leq 1/V_N^{3/2}(\lambda)$. Consequently, for $-\frac{1}{2}\pi V_N^{1/2}(\lambda) < z < \frac{1}{2}\pi V_N^{1/2}(\lambda)$,

$$\exp\Big\{-\frac{1}{2}z^2 + \frac{1}{6}\left|\frac{\kappa_{3N}(\lambda)+d_{3N}(\lambda)}{V_N^{3/2}(\lambda)}\right| |z|^3 + \frac{1}{24}\left|\frac{\kappa_{4N}(\lambda)+d_{4N}(\lambda)}{V_N^{2}(\lambda)}\right| z^4$$

$$+\frac{1}{120}\left|\frac{\kappa_{5N}(\lambda)+d_{5N}(\lambda)}{V_N^{5/2}(\lambda)}\right| |z|^5\Big\}$$

$$\leq \exp\{-\frac{1}{2}z^2 + \frac{1}{6}(|z^3|/V_N^{1/2}(\lambda)) + \frac{1}{24}(z^4/V_N(\lambda)) + \frac{1}{120}(|z^5|/V_N^{3/2}(\lambda)) + 0(1/N)\}$$

$$\leq \exp\{-\frac{1}{2}z^2[1 - \frac{\pi}{6} - \frac{\pi^2}{48} - \frac{\pi^3}{480} + 0(1/N)\}$$

$$\leq \exp\{-\frac{1}{8}z^2[1+0(1/N)]\}$$

and the exponential term involving z^3, z^4, and z^5 may be expanded in Taylor series. As the contributions from the tails are exponentially small by virtue of (4.6), they may be ignored, and this last expansion followed by integration over $\infty < z < \infty$ yields (4.2). []

5. SADDLE–POINT APPROXIMATION TO $f_{Z_{(n)}(\lambda)}$

Daniels (1954) develops saddle–point approximations and associated asymptotic expansions for the density of a mean of mutually independent and identically distributed *rvs*, and establishes the relation between this form of approximation and Edgeworth type expansions using Khinchin's (1949) concept of conjugate distributions (cf. Cox and Bardnorff–Nielsen (1979) for a more recent discussion).

In the representation (3.1) for $f_{Z_{(n)}}(\lambda)$, the large parameter N appears both as the number of terms in a product and the number of terms in a sum. This is not the standard case treated by Daniels. The nature of the problem is this: let $K_N(iu)$ be the cumulant function

$$K_N(iu) = \frac{1}{N}\sum_{k=1}^{N} \log\left[a_k e^{-ipu} + (1-a_k)e^{iqu}\right]. \tag{5.1}$$

Then with $h_N(iu) \equiv \exp\{-iu\}B_N(iu)$ and $B_N(iu)$ as in (4.10), (3.1) is

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{NK_N(iu)}h_N(iu)du = \frac{1}{2\pi i}\int_{-i\pi}^{i\pi} e^{NK_N(v)}h_N(v)dv. \tag{5.2}$$

Since $h_N(v)$ depends on N, we expand about a stationary point that is a solution to

$$\frac{d}{dv}\left[K_N(v) + \frac{1}{N}\log h_N(v)\right] = 0 \tag{5.3}$$

rather than a stationary point satisfying

$$\frac{d}{dv}K_N(v) = 0. \tag{5.4}$$

A solution to (5.3) is a value of v satisfying

$$q - \frac{1}{N}\sum_{k=1}^{N}\frac{a_k}{a_k + (1-a_k)e^v} + \frac{1}{Nh_N(v)}\frac{d}{dv}h_N(v) = 0. \tag{5.5}$$

For fixed $a_k\epsilon(0,1)$, $k = 1,2,\ldots,N$ the second term on the *LHS* of (5.5) provides a correction of magnitude at most $1/N$ as

$$\left|\frac{1}{Nh_N(v)}\frac{d}{dv}h_N(v)\right| \tag{5.6}$$

$$\leq \frac{1}{N} - \frac{1}{N}\Bigg\}\sum_{k=1}^{N} y_k\left(\frac{a_k}{a_k+(1-a_k)e^v}\right)^2 \Bigg/ \sum_{k=1}^{N} y_k\left(\frac{a_k}{a_k+(1-a_k)e^v}\right)\Bigg\} \leq \frac{1}{N},$$

for all $v\epsilon(-\infty,\infty)$.

That (5.5) possesses a unique solution v_0 in $(-\infty,\infty)$ and that the corresponding saddle–point approximation to (5.2) generated by expanding the integrand about v_0 when $v_0\epsilon(-\pi,\pi)$ leads to a valid asymptotic expansion must be verified. To this end we list needed properties of $D_N(v) = K_N(v) + \frac{1}{N}\log h_N(v)$, and its derivatives $D_N^{(j)}(v)$, $j = 1,2,\ldots$.

For $y_1,\ldots,y_N$ and λ fixed, so that $a_1,\ldots,a_N$ are fixed positive numbers, higher order derivatives $D_N^{(2)}(v),\ldots,D_N^{(j)}(v)$, are conveniently representable in terms of $\xi_k(v) = a_k/[a_k+(1-a_k)e^v]$, the N point probability function $\theta_k(v) = y_k\xi_k(v)/h_N(v)$, $k = 1,2,\ldots,N$, and the averages $\overline{\xi}_N(v) = \frac{1}{N}\sum_{k=1}^{N}\xi_k(v)$ and $\dot{\xi}_N = \sum_{k=1}^{N}\theta_k(v)\xi_k(v)$. Here, $\xi_k(v)$ plays the role of a_k in the Edgeworth type expansion of $f_{Z_{(n)}}$.

<u>**Assertion**</u>: The function $D_N(v)$ has these properties:

(i) $D_N^{(1)}(v) = q - \overline{\xi}_N(v) + \frac{h_N^{(1)}(v)}{Nh_N(v)} = q - \overline{\xi}_N(v) - \frac{1}{N}(1-\dot{\xi}_N(v))$ (5.7)

(ii) $D_N^{(1)}(v) = 0$ has a unique solution in the interval $(-\infty,\infty)$ provided that $q > 1/N$.

(iii) At $v = v_0$, (at $D_N^{(1)}(v_0) = 0$), $q - \overline{\xi}_N(v_0) = \frac{1}{N}(1-\dot{\xi}_N(v_0)) > 0$ since $\dot{\xi}_N(v)\epsilon(0,1)$ for $v\epsilon(-\infty,\infty)$.

(iv) $D_N^{(2)}(v) = \frac{1}{N}\sum_{k=1}^{N}[1-\theta_k(v)]\xi_k(v)(1-\xi_k(v)) + \frac{1}{N}\sum_{k=1}^{N}\theta_k(v)(\xi_k(v)-\dot{\xi}_N(v))^2 > 0$ and at $v = v_0$, when $\xi_k = \xi$, all k, $D_N^{(2)}(v_0) = (\frac{n-1}{N-1})(\frac{N-n}{N})$.

<u>**Proof**</u>: (i) follows from the definitions of $\xi_k(v)$, $\overline{\xi}(v)$, and $\dot{\xi}(v)$.

As $v \to -\infty$, $D_N^{(1)}(v) \to q-1 < 0$; at $v = 0$, $D_N^{(1)}(0) = q - \frac{1}{N} - \frac{1}{N}\sum_{k=1}^{N}a_k + \frac{1}{N}\left\{\sum_{k=1}^{N}y_ka_k^2\Big/\sum_{j=1}^{N}y_ka_k\right\}$ which may be less than, equal to or greater than zero; as $v \to +\infty$, $D_N^{(1)}(v) \to 1-f > 0$. Thus $D_N^{(1)}(v) = 0$ for at least one $v\epsilon(-\infty,\infty)$. As will be shown, $D_N^{(2)}(v) > 0$, so the solution to this equation is unique.

That $D_N^{(2)}(v) > 0$ can be shown via

$$D_N^{(2)}(v) = \frac{1}{N}\sum_{k=1}^{N}\xi_k(v)[1-\xi_k(v)] + \frac{1}{N}\left\{\frac{h_N^{(2)}(v)}{h_N(v)} - \left[\frac{h_N^{(1)}(v)}{h_N(v)}\right]^2\right\}. \tag{5.8}$$

Differentiation of $\log h_N(v)$ yields

$$\frac{h_N^{(1)}(v)}{h_N(v)} = -\sum_{k=1}^{N} \theta_k(v)(1-\xi_k(v)) \tag{5.9}$$

and

$$\frac{h_N^{(2)}(v)}{h_N(v)} = -2\sum_{i=1}^{N} \theta_k(v)\xi_k(v)[1-\xi_k(v)] + \sum_{k=1}^{N} \theta_k(v)[1-\xi_k(v)] \tag{5.10}$$

and (iv) follows directly. []

For notational convenience, we have in places suppressed explicit display of λ and $y_1, \ldots, y_N$. However, as a solution v_0 to $D_N^{(1)}(v) = 0$ depends on λ and we wish to approximate the density $f_{Z_{(n)}}(\lambda)$ over an interval for λ, we now write v_0 as an explicit function $v(\lambda)$ of λ and D_N as an explicit function of λ and v. Given positive numbers $y_1, \ldots, y_N$, there is a set $S_\lambda^{(N)} = \{\lambda \mid D_N^{(1)}(\lambda, v(\lambda)) = 0 \text{ and } -\pi < v(\lambda) < \pi\}$ of λ values corresponding to the restriction $-\pi < v < \pi$ imposed by the range of integration of (3.1). For $\lambda \epsilon S_\lambda^{(N)}$, stationary points of D_N lie in $(-\pi, \pi)$; for $\lambda \epsilon (0, \infty)$ but $\lambda \not\epsilon S_\lambda$, the integrand has no stationary point in $(-\pi, \pi)$ and the principal contribution to the value of the integral comes from an endpoint.

<u>**Theorem 5.1**</u>: *Let $p = n - 1/N = 1 - q$, $\lambda \epsilon S_\lambda^{(N)}$ and $v(\lambda)$ be a solution to $D_N^{(1)}(\lambda, v) = 0$ for given λ. Define*

$$L_j(\lambda, v(\lambda)) = D_N^{(j)}(\lambda, v(\lambda))/[D_N^{(2)}(\lambda, v(\lambda))]^{j/2} \tag{5.11}$$

for $j = 2, 3, \ldots$. Then for n/N fixed and N large,

$$f_{Z_{(n)}}(\lambda) = \left[\frac{1}{2\pi N D_N^{(2)}(\lambda, v(\lambda))}\right]^{1/2} \exp\{N D_N(\lambda, v(\lambda))\}$$

$$\times\{1 + \frac{1}{24N}[3L_4(\lambda, v(\lambda)) - 5L_3^2(\lambda, v(\lambda))] \tag{5.12}$$

$$+ \frac{1}{1152N^2}[-24L_6(\lambda, v(\lambda)) + 168L_3(\lambda, v(\lambda))L_5(\lambda, v(\lambda))$$

$$+ 105L_4^2(\lambda, v(\lambda)) - 630L_3^2(\lambda, v(\lambda))L_4(\lambda, v(\lambda))$$

$$+ 38L_3^4(\lambda, v(\lambda))] + 0(N^{-3})\}$$

Proof: The formal development of (5.12) follows the pattern of analysis of Daniels (1954) or Good (1956) and is not repeated in detail. The only task is to show that the function $dz/d\omega$ as defined below is analytic in a neighborhood of zero and bounded in an interval on the steepest descent contour.

Define $a = (v - v(\lambda))/[D_N^{(2)}(\lambda, v(\lambda))]^{1/2}$ and ω as a function satisfying

$$-\frac{1}{2}\omega^2 = D_N(\lambda, v) - D_N(\lambda, v(\lambda)) = \frac{1}{2}z^2 + \frac{1}{6}L_3(\lambda, v(\lambda))z^3 + \frac{1}{24}L_4(\lambda, v(\lambda))z^4 + \dots \tag{5.13}$$

with the same sign as the imaginary part of z on the steepest descent contour. For some $\alpha, \beta > 0$ the contribution on this contour in a neighborhood of $v(\lambda)$ is

$$\frac{e^{ND_N(\lambda, v(\lambda))}}{[2\pi N D_N^{(2)}(\lambda, v(\lambda))]^{1/2}} \int_{-\alpha}^{\beta} e^{-1/2N\omega^2} \frac{dz}{d\omega} d\omega \tag{5.14}$$

That $D_N(\lambda, v)$ is bounded and analytic for $v\epsilon(-\pi, \pi)$ is effectively established in the course of computing the Edgeworth type approximation to $f_{Z_{(n)}}$ (cf. (4.6) and $ff\cdot$). By the inversion theorem for analytic functions (cf. Levinson and Redheffer (1970) for example) z is analytic in $(-\alpha, \beta)$ hence $dz/d\omega$ is also. An application of Watson's lemma to (5.14) yields (5.12). []

With $a_j(\lambda) = \exp\{-\lambda y_j\}$, the leading term of (5.12) is

$$\left[\frac{1}{2\pi N D_N^{(2)}(\lambda, v(\lambda))}\right]^{1/2} e^{-(n-1)v(\lambda)} \sum_{j=1}^{N} [a_j(\lambda) + (1 - a_j(\lambda))e^{v(\lambda)}] \times \sum_{k=1}^{N} \frac{y_k a_k(\lambda)}{a_k(\lambda) + (1 - a_k(\lambda))e^{v(\lambda)}}. \tag{5.15}$$

When $y_k = y$, $k = 1, 2, \dots, N$, $D_N^{(1)}(\lambda, v(\lambda)) = 0$ becomes

$$\frac{a(\lambda)}{a(\lambda) + (1 - a(\lambda))e^{v(\lambda)}} = \frac{N-n}{N-1} = \frac{N}{N-1}\left(q - \frac{1}{N}\right)$$

and

$$D_N^{(2)}(\lambda, v(\lambda)) = \left(q - \frac{1}{N}\right)\left(1 - \frac{N}{N-1}\left(q - \frac{1}{N}\right)\right) = \left(\frac{n-1}{N}\right)\left(\frac{N-n}{N-1}\right).$$

The leading term is then, with $Q = (N-n)/(N-1) = 1-P$,

$$\left\{N\left[\frac{1}{2\pi(N-1)}\right]^{1/2}\left(\frac{1}{P}\right)^{n-1/2}\left(\frac{1}{Q}\right)^{N-n+1/2}\right\} \times ye^{-(N-n+1)\lambda y}\left(1-e^{-\lambda y}\right)^{n-1}. \quad (5.16)$$

Aside from the normalizing constant, this is the exact form of the density of the nth order statistic generated by N independent exponential *rvs* with common means $1/y$. To the order of the first term of the Stirling approximation the term in curly brackets in (5.16) is $n\binom{N}{n}$.

An Edgeworth–type approximation of $f_{z_n}(\lambda)$ can be computed directly from the leading term (5.15) of the saddle–point approximation by expanding the latter in Taylor series. For example, for values of λ such that $v(\lambda) = 0(1/\sqrt{N})$,

$$\exp\{-(n-1)v(\lambda)\}\prod_{k=1}^{N}\left[a_k + (1-a_k)e^{v(\lambda)}\right] \quad (5.17)$$

$$= \exp\left\{-N\left[\left(\frac{1}{N}\sum_{k=1}^{N}a_k - q\right)v(\lambda) + \frac{1}{2}D_N^{(2)}(\lambda,0)v^2(\lambda)\right]\right\}(1+0(1/\sqrt{N}))$$

with $ND_N^{(2)}(\lambda,0) = \sum_{k=1}^{N} a_k(1-a_k)$. Upon completing the square in (5.17), (5.15) becomes

$$\rho_N(\lambda, v(\lambda)) \quad (5.18)$$

$$\times\left[2\pi D_N^{(2)}(\lambda,0)\right]^{-1/2}\exp\left\{-\frac{1}{2}N\left(q-\frac{1}{N}\sum_{k=1}^{N}a_k\right)^2/D_N^{(2)}(\lambda,0)\right\}\left(\sum_{k=1}^{N}a_k y_k\right)$$

with

$$\rho_N(\lambda, v(\lambda)) = \left[D_N^{(2)}(\lambda,0)/D_N^{(2)}(\lambda,v(\lambda))\right]^{1/2}\left[h_N(\lambda,v(\lambda))/h_N(\lambda,0)\right]$$

$$\times\exp\left\{\frac{1}{2}ND_N^{(2)}(\lambda,0)\left[v(\lambda)-\left(\left(\frac{1}{N}\sum_{k=1}^{N}a_k-q\right)/D_N^{(2)}(\lambda,0)\right)\right]^2\right\}$$

$$\times(1+0(1/\sqrt{N})), \quad (5.19)$$

and $h_N(\lambda, v(\lambda)) \equiv h_N(v)$. For $v(\lambda) = 0(1/\sqrt{N})$, $\rho_N(\lambda, v(\lambda)) = 1 + 0(1/\sqrt{N})$, so (5.18) is asymptotically equivalent to (3.8).

6. NUMERICAL EXAMPLES

This section provides numerical comparisons of Edgeworth and saddle–point approximations with the exact density of $\widetilde{\lambda}$ at .01, .05, .10, .25, .50, .75, .95, and .99 fractile values. Integration of the exact density was done using a Rhomberg–type integration routine, CADRE, allowing prespecified error tolerance.

Two finite population magnitude shapes — exponential and lognormal — are examined for $(N, n) = (6, 2), (10, 3), (30, 10)$ and $(150, 50)$. Given N, y_k is the $=(k/N + 1)$st fractile of an exponential distribution with mean one if shape is exponential, or the $(k/N + 1)$st fractile of a lognormal distribution with parameter $(\mu, \sigma^2) = (0, .5)$ if shape is lognormal. Here the y_k values are <u>not</u> normalized by scaling so that $\sum_{k=1}^{N} y_k = 1$.

Figures 6.1 to 6.16 provide visual comparisons of the approximations to the exact density. For N =10 and n/N =.3, the leading terms of both Edgeworth and saddle-point approximations behave sufficiently well to obviate need for visual display of the fit of leading term plus $0(1N)$ corrections.

Principal features of these examples are:

(1) As expected, saddle–point approximations outperform Edgeworth type approximations in all cases. With $0(1/N)$ corrections the former works very well for N as small as 6, providing almost uniform error of .74 – –1.0% over a .01 to .99 fractile range (without renormalization). See Tables 6.1 and 6.2.

(2) Edgeworth type approximations with $0(1/N^{1/2})$ corrections are drastically bad for small values of N.

(3) In the fractile range .01 to .99, $-\pi < v(\lambda) < \pi$ in all examples, suggesting that a LaPlace approximation to (5.2) at $v(\lambda) = \pm\pi$ need be employed only when extreme tail values of the density of $\widetilde{\lambda}$ are desired.

Table 6.1

COMPARISONS OF EDGEWORTH AND SADDLE-POINT APPROXIMATIONS TO EXACT DENSITY (EXPONENTIAL MAGNITUDES)

N=6, n=2

			Leading Term		Order 1/N	
Fractile	λ	Exact	Edgeworth	Saddle-point	Edgeworth	Saddle-point
.01	.024018	.419235	.061133	.452944	4.506486	.416150
.05	.070789	.995708	1.207501	1.075392	-3.064328	.988296
.10	.110246	1.306700	1.783591	1.411114	-.976734	1.297162
.25	.208603	1.594300	1.928963	1.719877	1.974146	1.581960
.50	.374516	1.372730	1.397845	1.478411	1.476357	1.361506
.75	.612565	.792829	.749612	.851366	.763020	.785703
.90	.888184	.351786	.332229	.376120	.355551	.348188
.95	1.096582	.180270	.170609	.191963	.194361	.178220
.99	1.546093	.040050	.035761	.042185	.043265	.039481

Value of Integral .9998717
Error Tolerance .0000029

N=10, n=3

			Leading Term		Order 1/N	
Fractile	λ	Exact	Edgeworth	Saddle-point	Edgeworth	Saddle-point
.01	.053154	.383824	.142332	.398898	1.527169	.383156
.05	.105215	1.016052	1.098824	1.055634	.276358	1.014246
.10	.144096	1.423442	1.704973	1.478544	.768202	1.420872
.25	.230052	1.909645	2.169207	1.982440	1.996582	1.906059
.50	.358080	1.794180	1.816430	1.860817	1.862227	1.790607
.75	.531728	1.113901	1.060479	1.153562	1.094682	1.111495
.90	.733447	.499164	.474913	.515902	.495855	.497989
.95	.876189	.260651	.250723	.268963	.265011	.260007
.99	1.193754	.054578	.052538	.056071	.156349	.054409

Value of Integral .9999983
Error Tolerance .0000017

N=30, n=10

			Leading Term		Order 1/N	
Fractile	λ	Exact	Edgeworth	Saddle-point	Edgeworth	Saddle-point
.01	.192970	.276352	.208041	.278583	.309282	.276326
.05	.256436	.924684	.923148	.932018	.906719	.924597
.10	.296020	1.458765	1.522675	1.470208	1.420540	1.458628
.25	.371413	2.320714	2.420816	2.338526	2.323860	2.320498
.50	.470351	2.525443	2.541794	2.544266	2.534905	2.525212
.75	.586948	1.746635	1.714022	1.759206	1.741558	1.746484
.90	.708751	.850230	.832892	.866131	.848056	.850162
.95	.784187	.485144	.477188	.488434	.485371	.485108
.99	.954164	.110534	.110771	.111250	.111273	.110527

Value of Integral .9999999
Error Tolerance .0000013

N=150, n=50

			Leading Term		Order 1/N	
Fractile	λ	Exact	Edgeworth	Saddle-point	Edgeworth	Saddle-point
.01	.340765	.418278	.38920	.418867	.420089	.418277
.05	.378084	1.441982	1.438072	1.443992	1.440523	1.441977
.10	.400909	2.447910	2.478349	2.451304	2.443853	2.447903
.25	.441198	4.386328	4.451661	4.392349	4.385808	4.386316
.50	.489443	5.408805	5.422773	5.416141	5.412618	5.408791
.75	.541386	4.202887	4.165031	4.208515	4.202032	4.202876
.90	.594501	2.154998	2.132952	2.157847	2.153343	2.154992
.95	.626052	1.231495	1.225315	1.233112	1.230967	1.231493
.99	.688725	.300758	.305691	.301147	.301147	.300758

Value of Integral 1.0000000
Error Tolerance .0000059

Table 6.2
COMPARISONS OF EDGEWORTH AND SADDLE-POINT APPROXIMATIONS TO EXACT DENSITY (LOGNORMAL MAGNITUDES)

N=6, n=2

			Leading Term		Order 1/N	
Fractile	λ	Exact	Edgeworth	Saddle-point	Edgeworth	Saddle-point
.01	.017639	.577943	.0994731	.626857	6.240404	.573936
.05	.051422	1.370941	1.730503	1.486703	-4.142351	1.361381
.10	.080750	1.900438	2.498308	1.952162	-1.016195	1.787826
.25	.152217	2.197556	2.651076	2.381746	2.869058	2.181954
.50	.272280	1.900095	1.911543	2.057711	2.110507	1.886252
.75	.443512	1.105528	1.032229	1.195591	1.070097	1.097101
.90	.640027	.495043	.462681	.534307	4.873857	.491008
.95	.787235	.255540	.237720	.275307	.266645	.253329
.99	1.103846	.056798	.045859	.060883	.062162	.056224

Value of Integral 1.000000
Error Tolerance .000003

N=10, n=3

			Leading Term		Order 1/N	
Fractile	λ	Exact	Edgeworth	Saddle-point	Edgeworth	Saddle-point
.01	.040351	.553575	.233514	.576717	2.097992	.552688
.05	.078319	1.416704	1.584679	1.475687	.254551	1.414405
.10	.106721	1.970996	2.405085	2.052795	1.020777	1.967768
.25	.169326	2.630971	2.997934	2.739330	2.806427	2.626561
.50	.261764	2.476106	2.487440	2.576811	2.615171	2.471804
.75	.388349	1.535930	1.444273	1.797143	1.512724	1.533109
.90	.532447	.692709	.653094	.719554	.682349	.691344
.95	.633461	.363634	.346830	.377406	.365492	.362881
.99	.837582	.0879810	.082196	.091131	.092637	.087780

Value of Integral 1.000000
Error Tolerance .000002

N=30, n=10

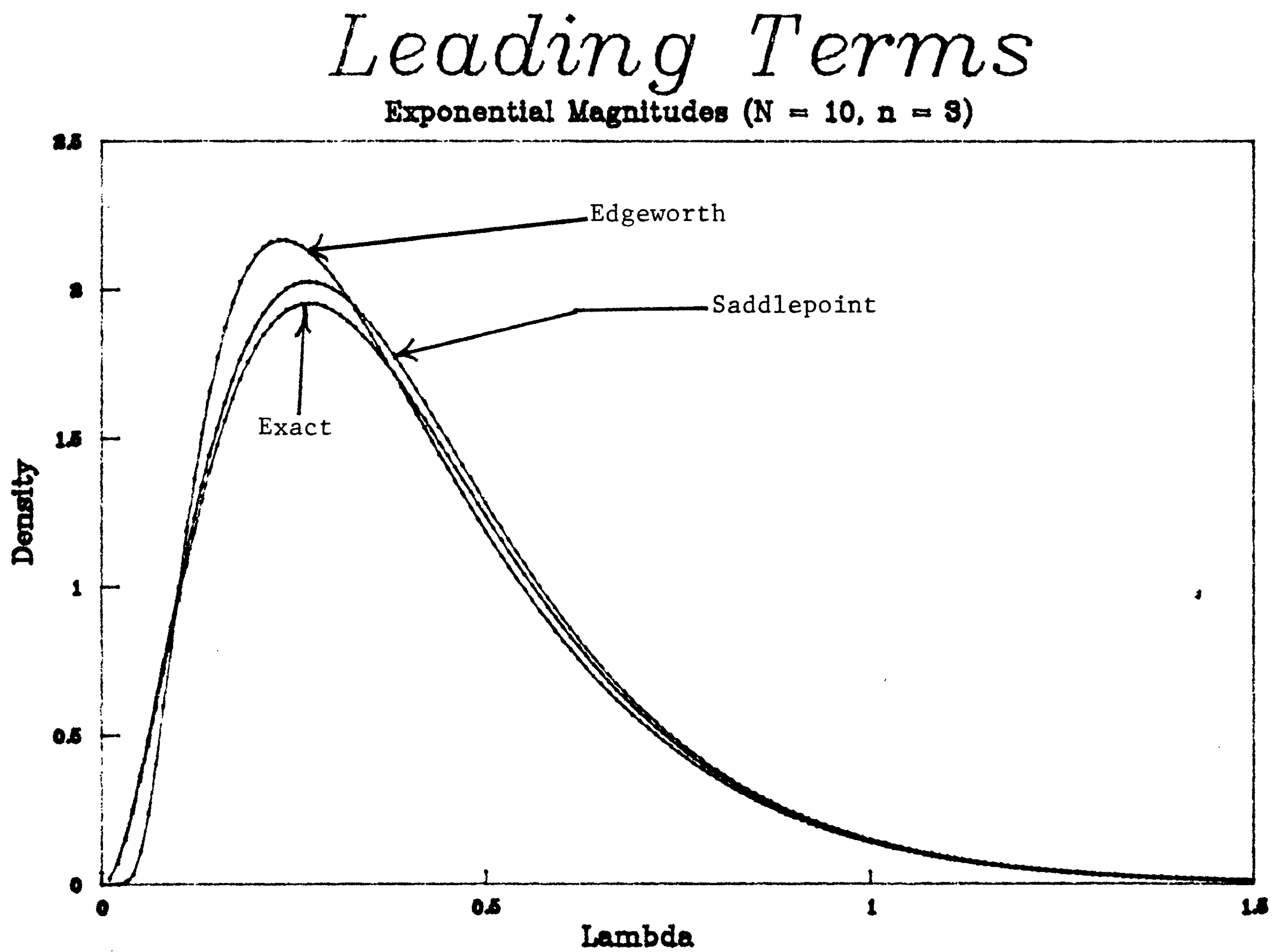
			Leading Term		Order 1/N	
Fractile	λ	Exact	Edgeworth	Saddle-point	Edgeworth	Saddle-point
.01	.139756	.3856811	.283097	.389169	.446834	.385646
.05	.185325	1.296510	1.301642	1.308228	.125824	1.296390
.10	.213515	2.047494	2.158589	2.065711	1.981476	2.047305
.25	.263297	3..210598	3.381253	3.2388233	3.205778	3.210952
.50	.333126	3.613187	3.642386	3.644420	3.639233	3.612847
.75	.414752	2.5570232	2.508919	5.591988	2.564500	2.569991
.90	.498863	1.283466	1.2500149	1.294088	1.277916	1.289947
.95	.554812	.707727	.695470	.7133495	7.067966	.707662
.99	.670414	.163706	.163987	.164999	.165021	.163692

Value of Integral .9999999
Error Tolerance .0000014

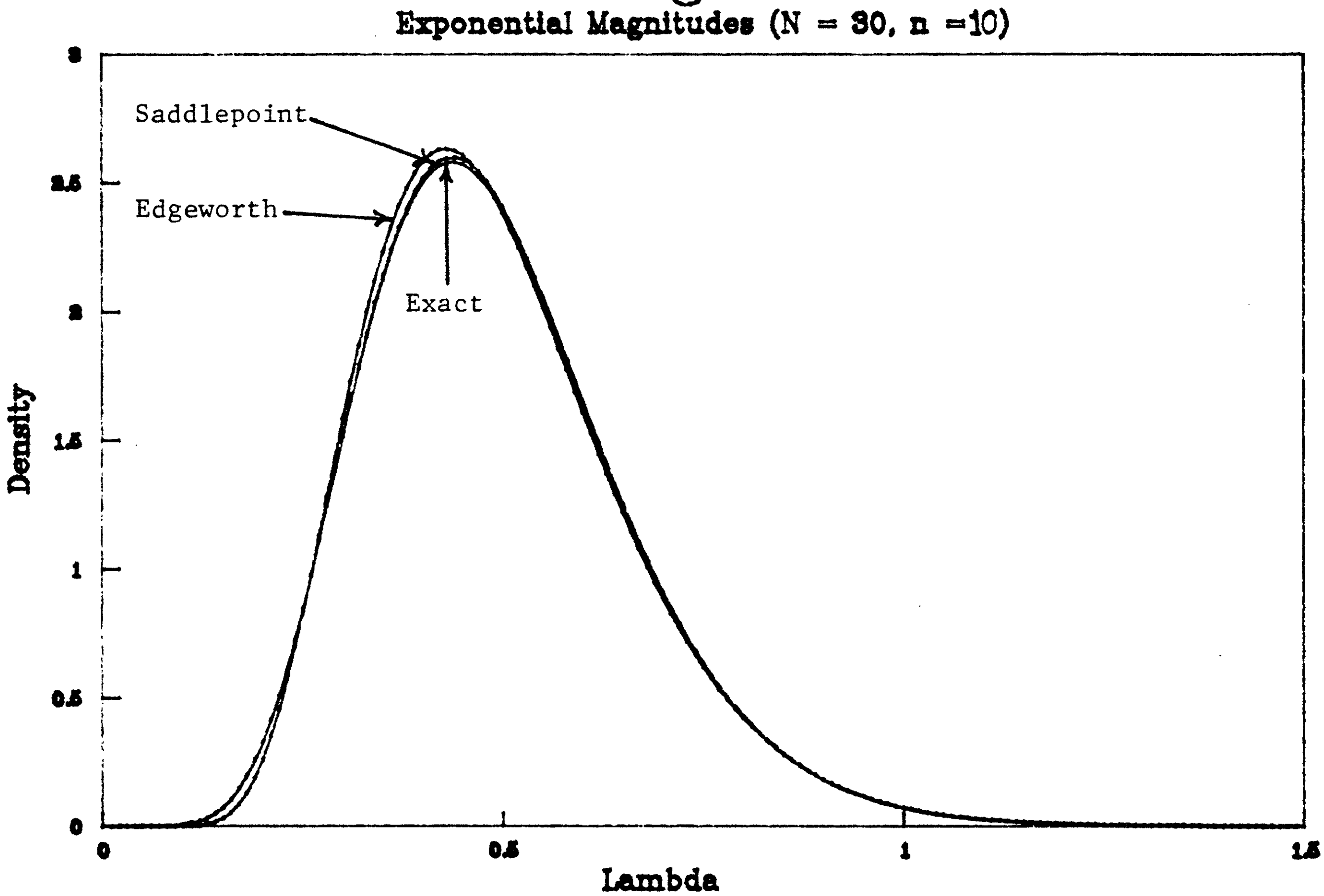
N=150, n=50

			Leading Term		Order 1/N	
Fractile	λ	Exact	Edgeworth	Saddle-point	Edgeworth	Saddle-point
.01	.243575	.559731	.511935	.560630	.565706	.559729
.05	.271998	2.135209	2.133936	2.139615	.213118	2.135202
.10	.288112	3.598506	3.658174	3.604226	3.588553	3.598496
.25	.316433	6.367083	6.482507	6.377138	6.365301	6.367063
.50	.350075	7.751501	7.769003	7.763652	7.760408	7.751478
.75	.387676	5.850960	5.778543	5.860052	5.849384	5.850939
.90	.421207	3.164488	3.122881	3.169373	3.160891	3.164481
.95	.442698	1.832762	1.820915	1.835577	1.831383	1.832757
.99	.485595	.455719	.466149	.456512	4.565742	.455817

Value of Integral 1.0000000
Error Tolerance .0000120

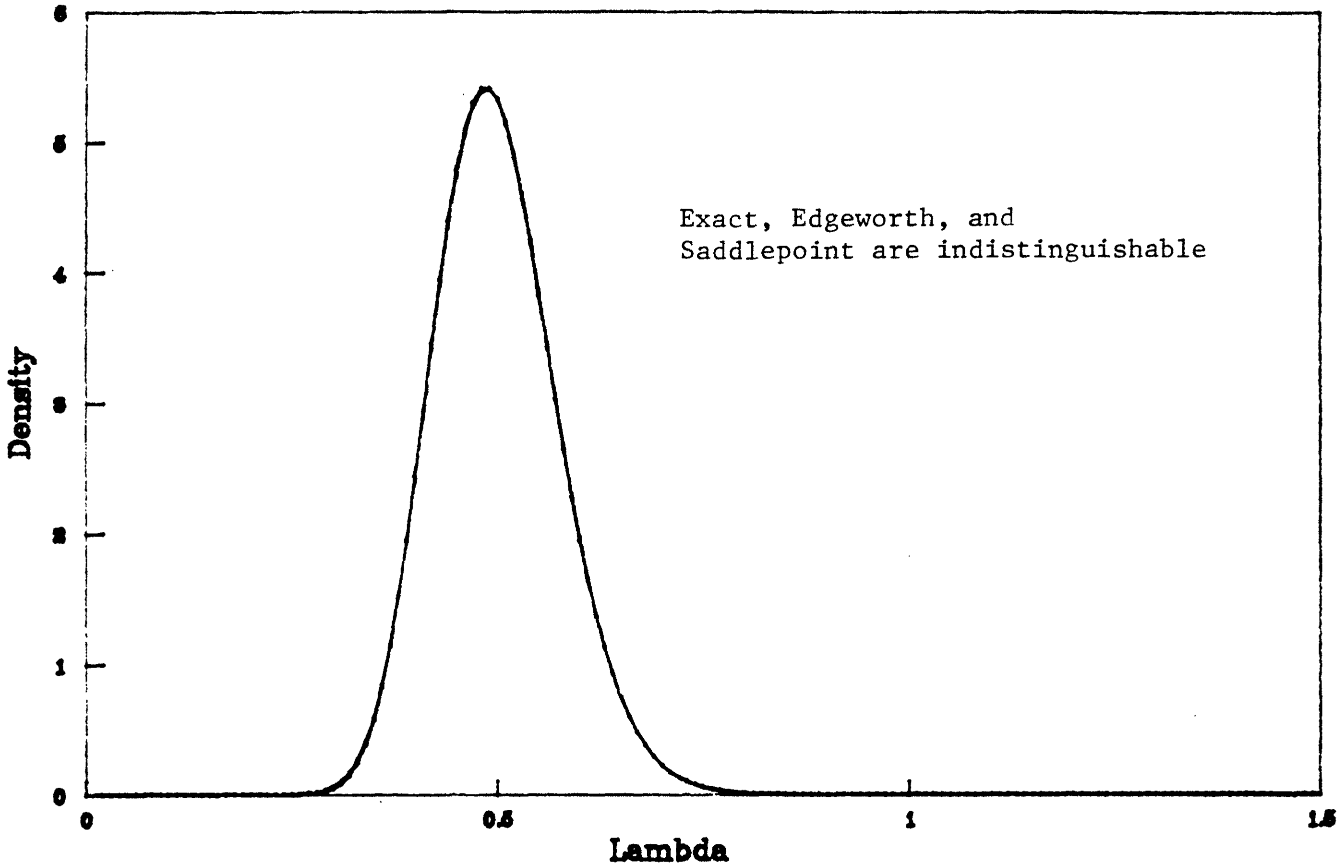

Leading Terms
Exponential Magnitudes (N = 10, n = 3)
Edgeworth
Saddlepoint
Exact
Density
Lambda
2.5
2
1.5
1
0.5
0
0
0.5
1
1.5

Leading Terms

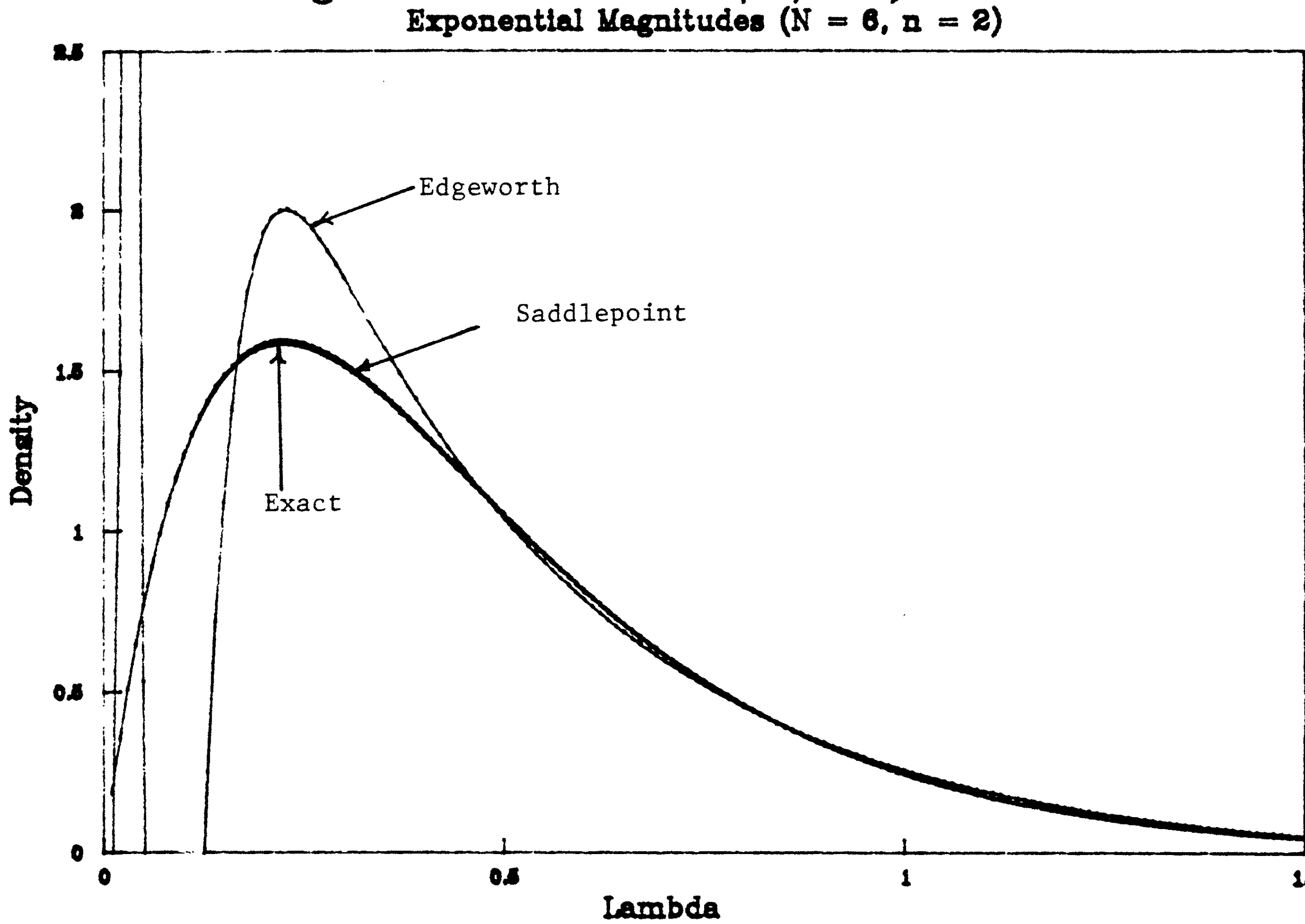


Leading Terms

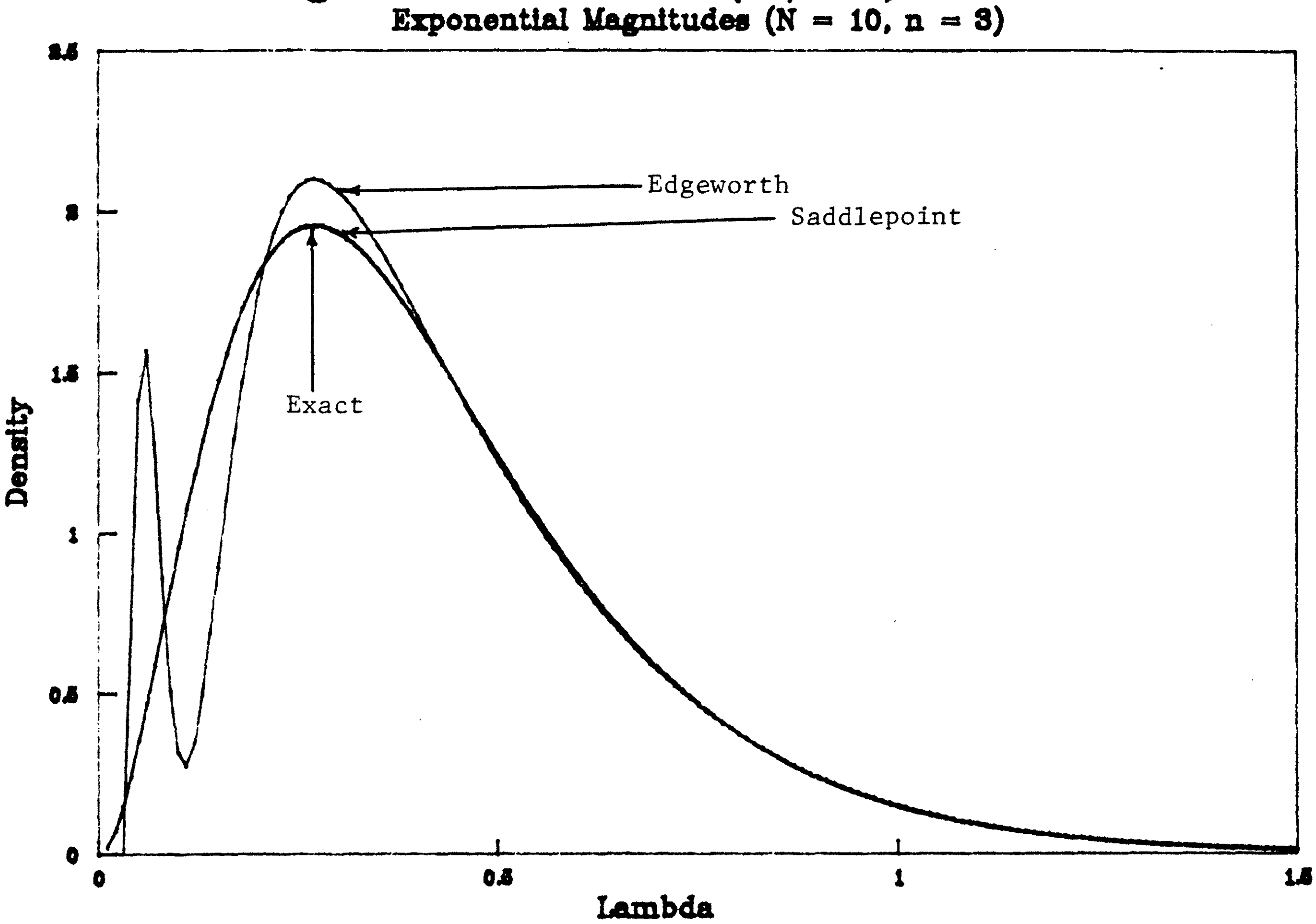
Exponential Magnitudes (N =150, n =50)



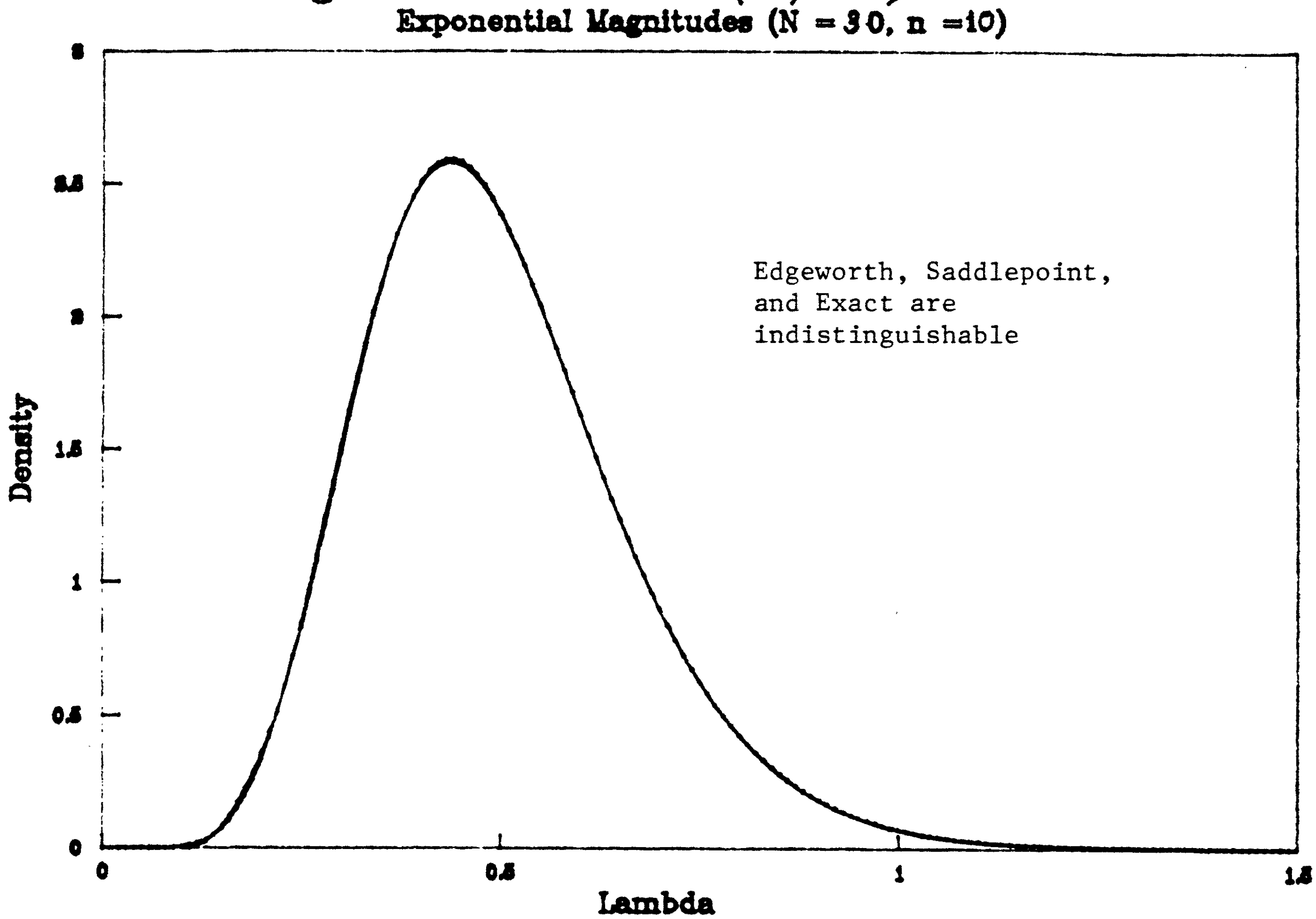
Leading Terms + O(1/N) Corrections



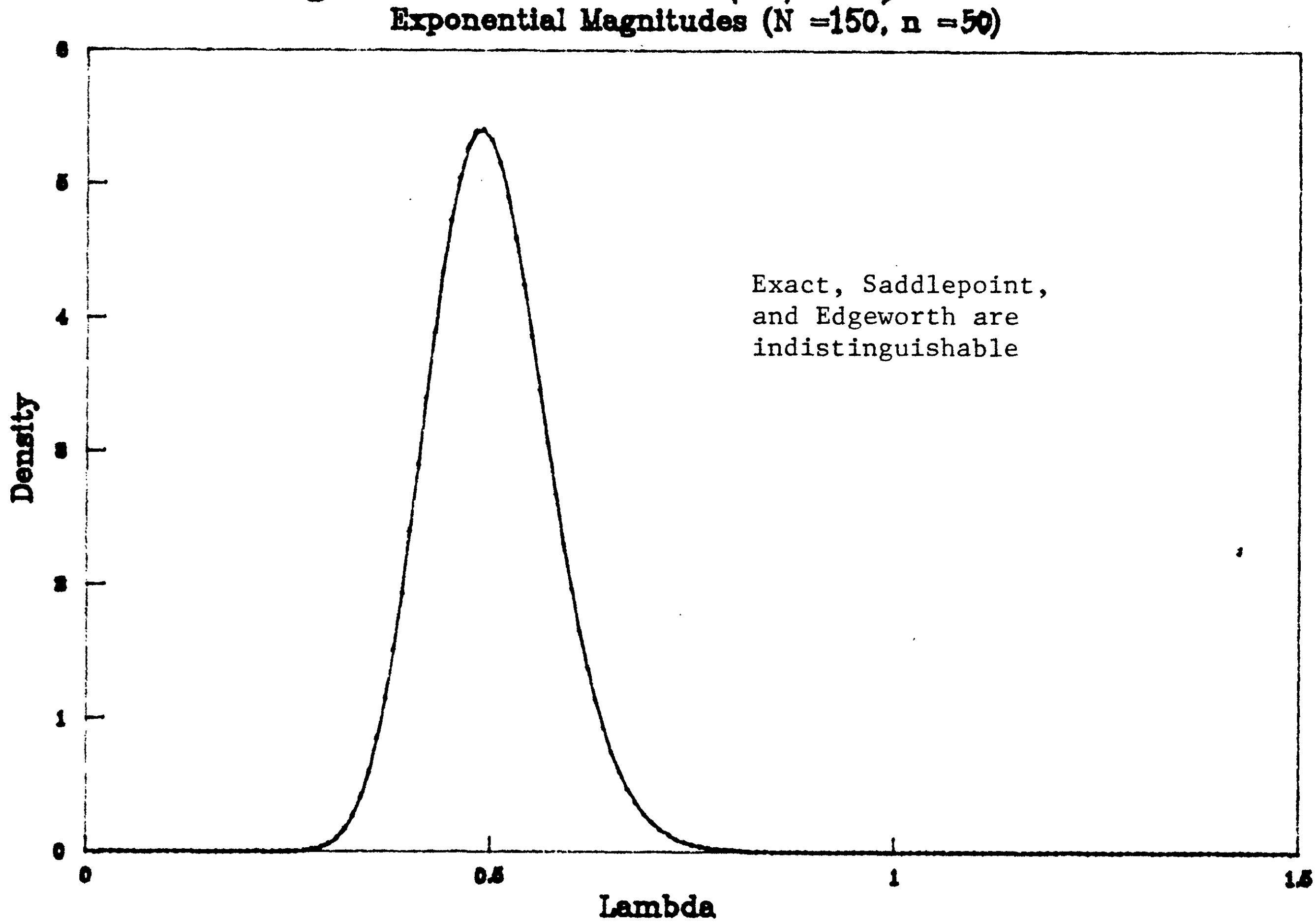
Leading Terms + O(1/N) Corrections



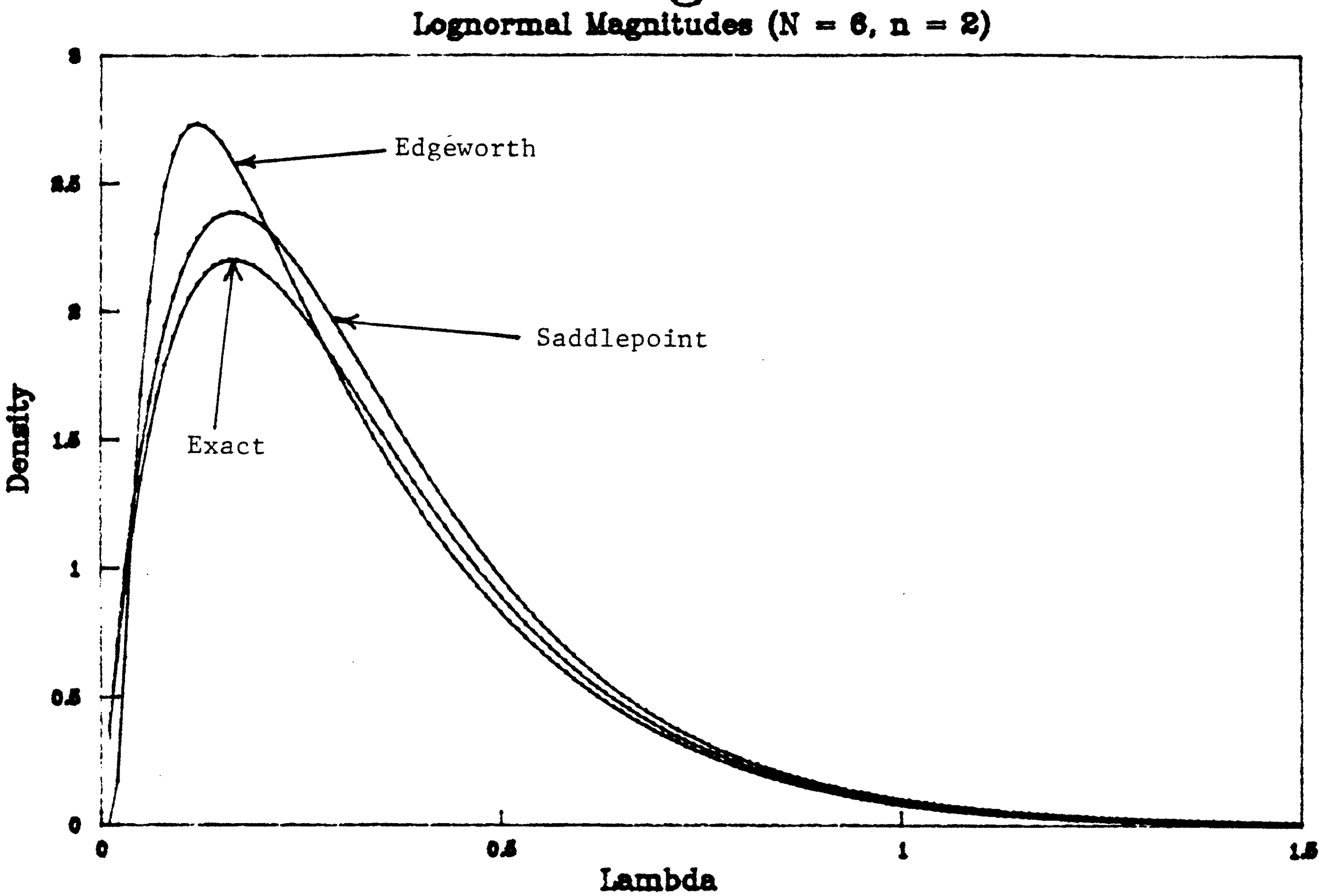
Leading Terms + O(1/N) Corrections



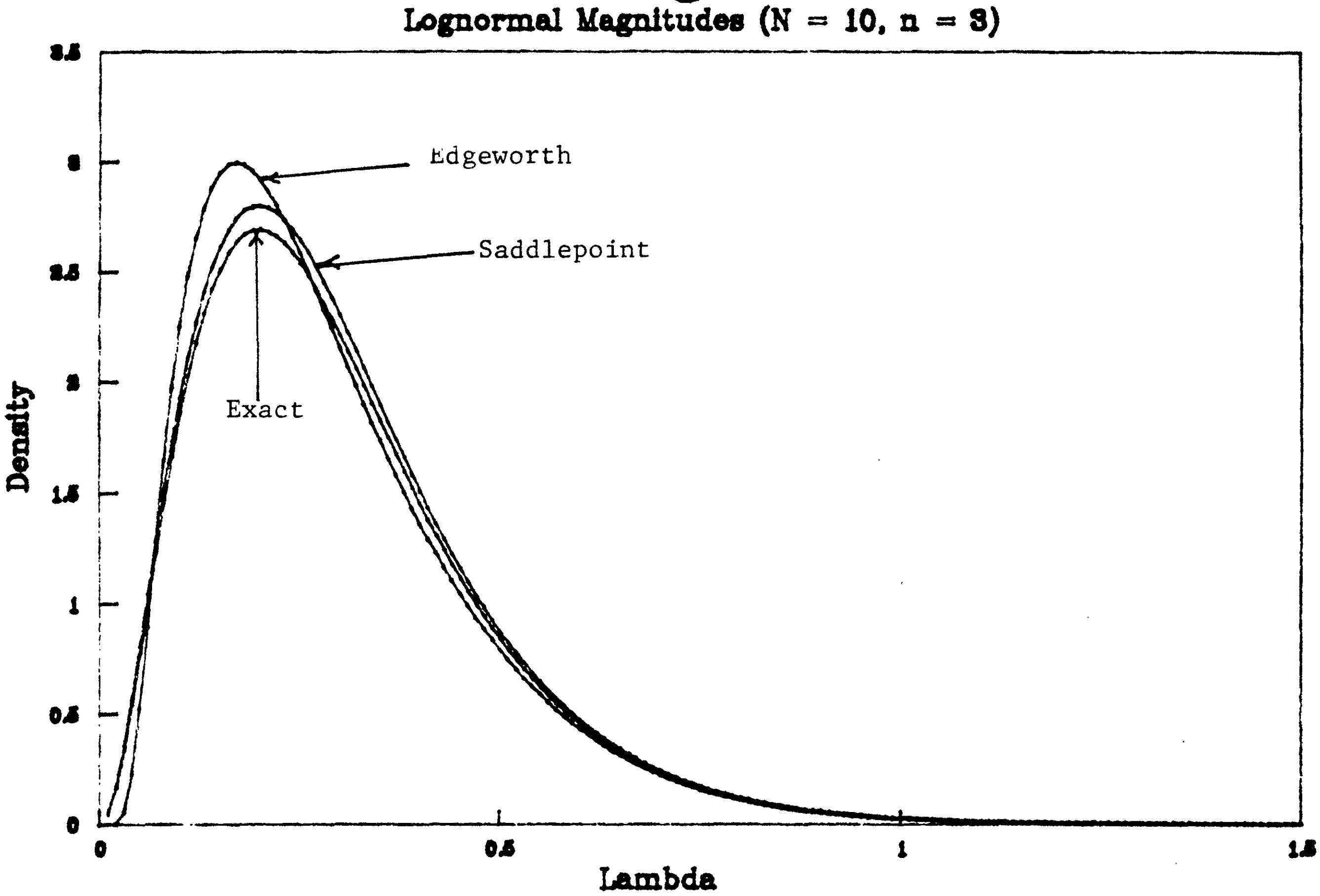
Leading Terms + O(1/N) Corrections

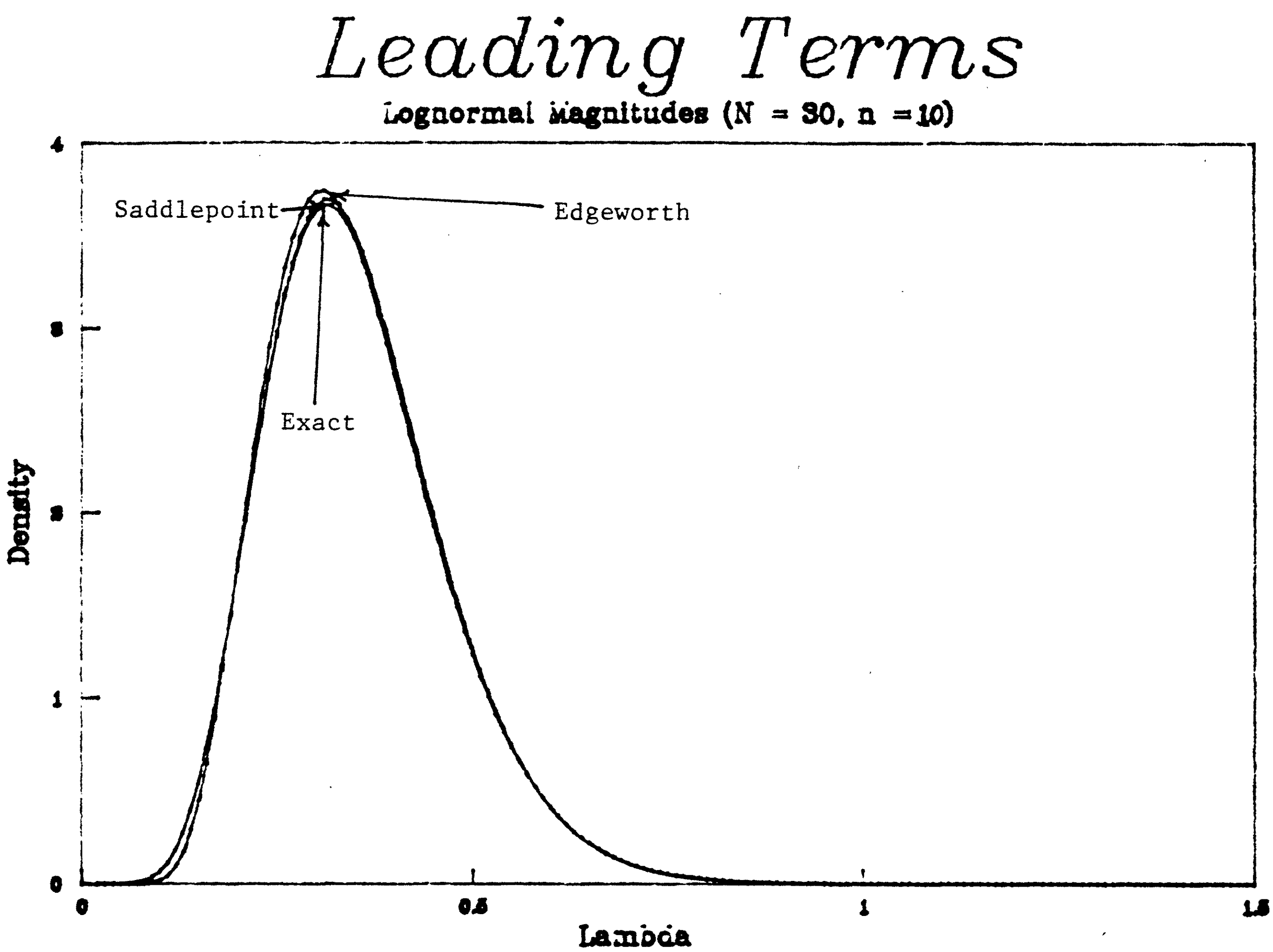


Leading Terms

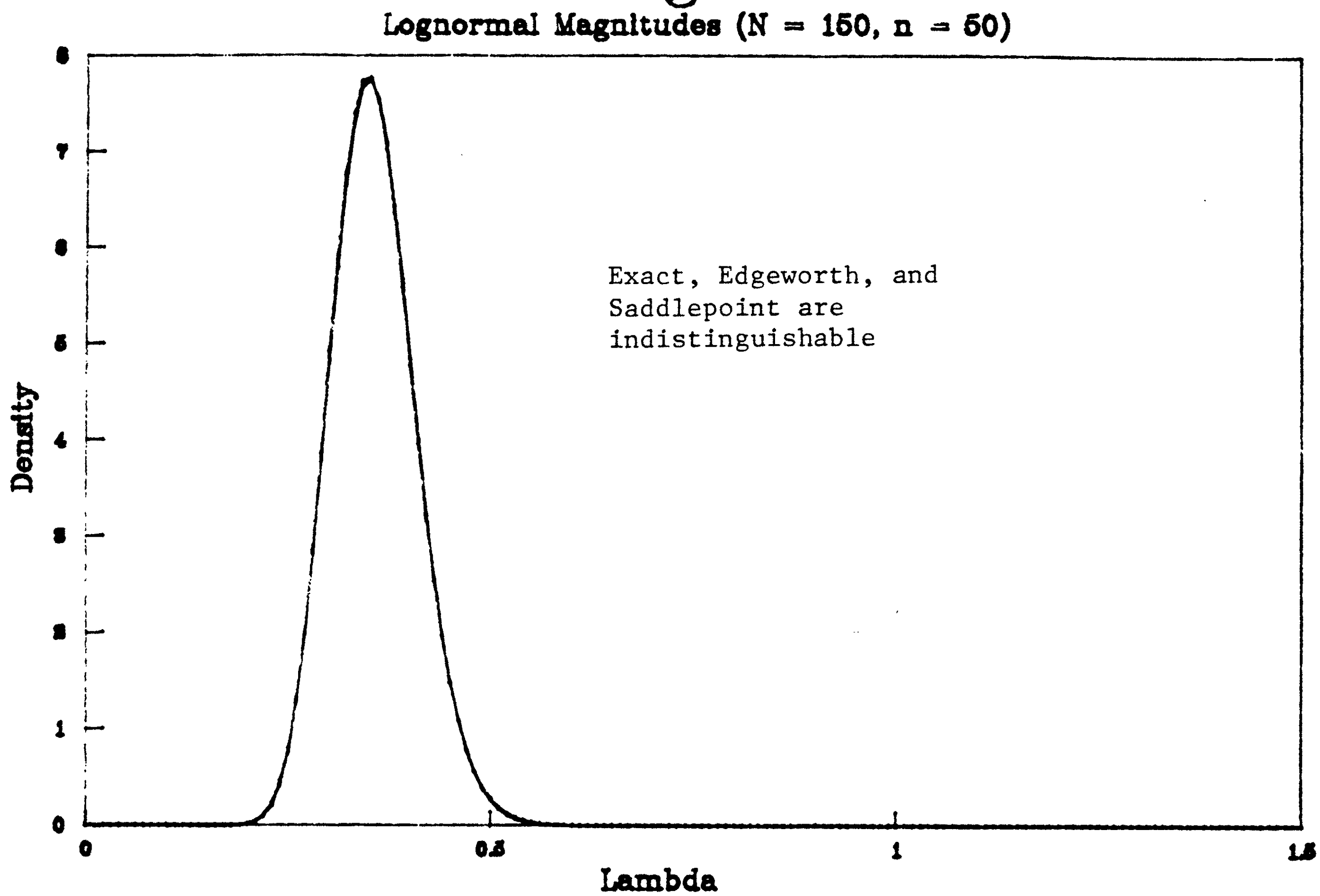


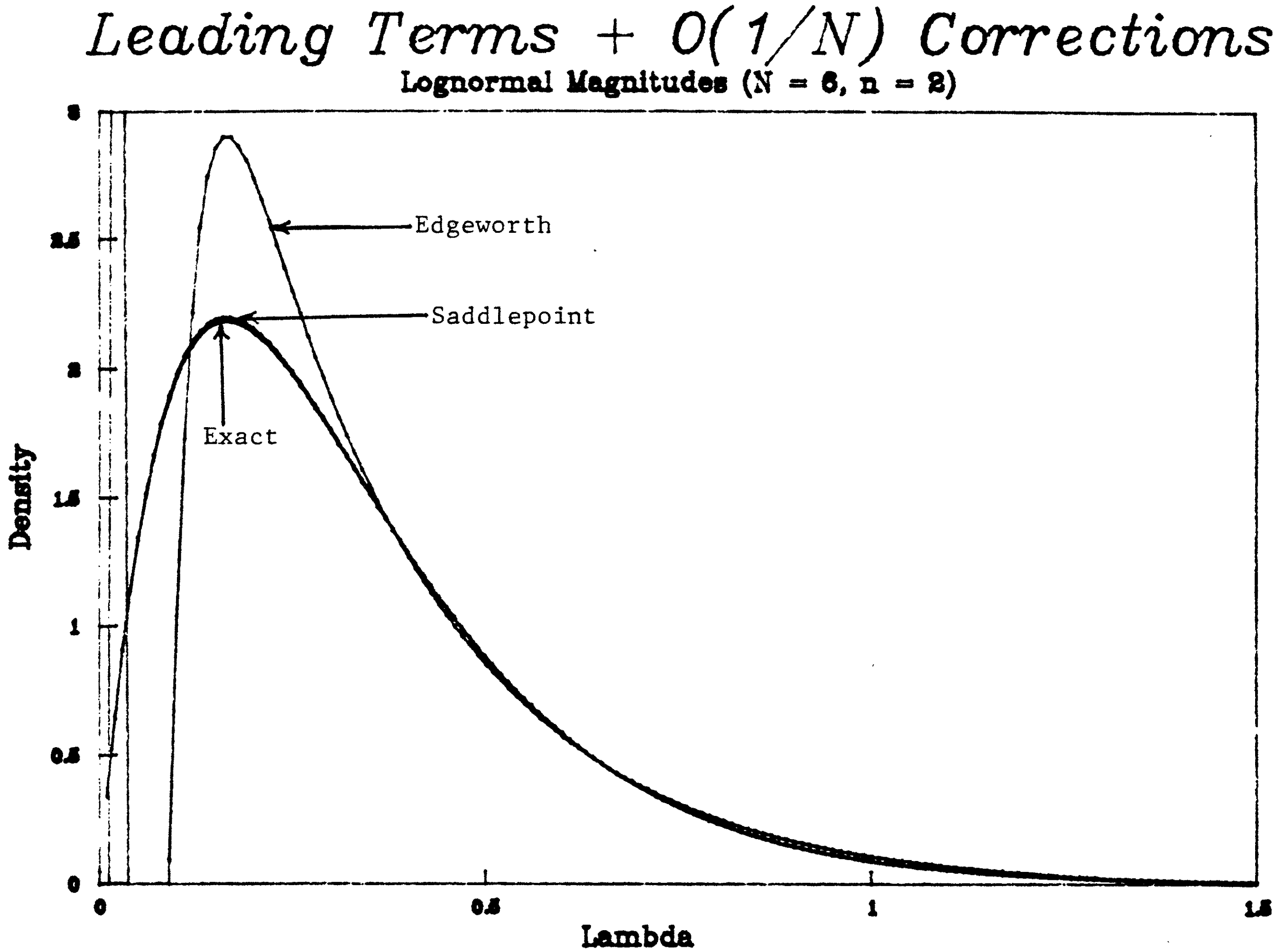
Leading Terms




Leading Terms
Lognormal Magnitudes (N = 30, n = 10)
Saddlepoint
Edgeworth
Exact
Density
4
3
2
1
0
0
0.5
1
1.5
Lambda

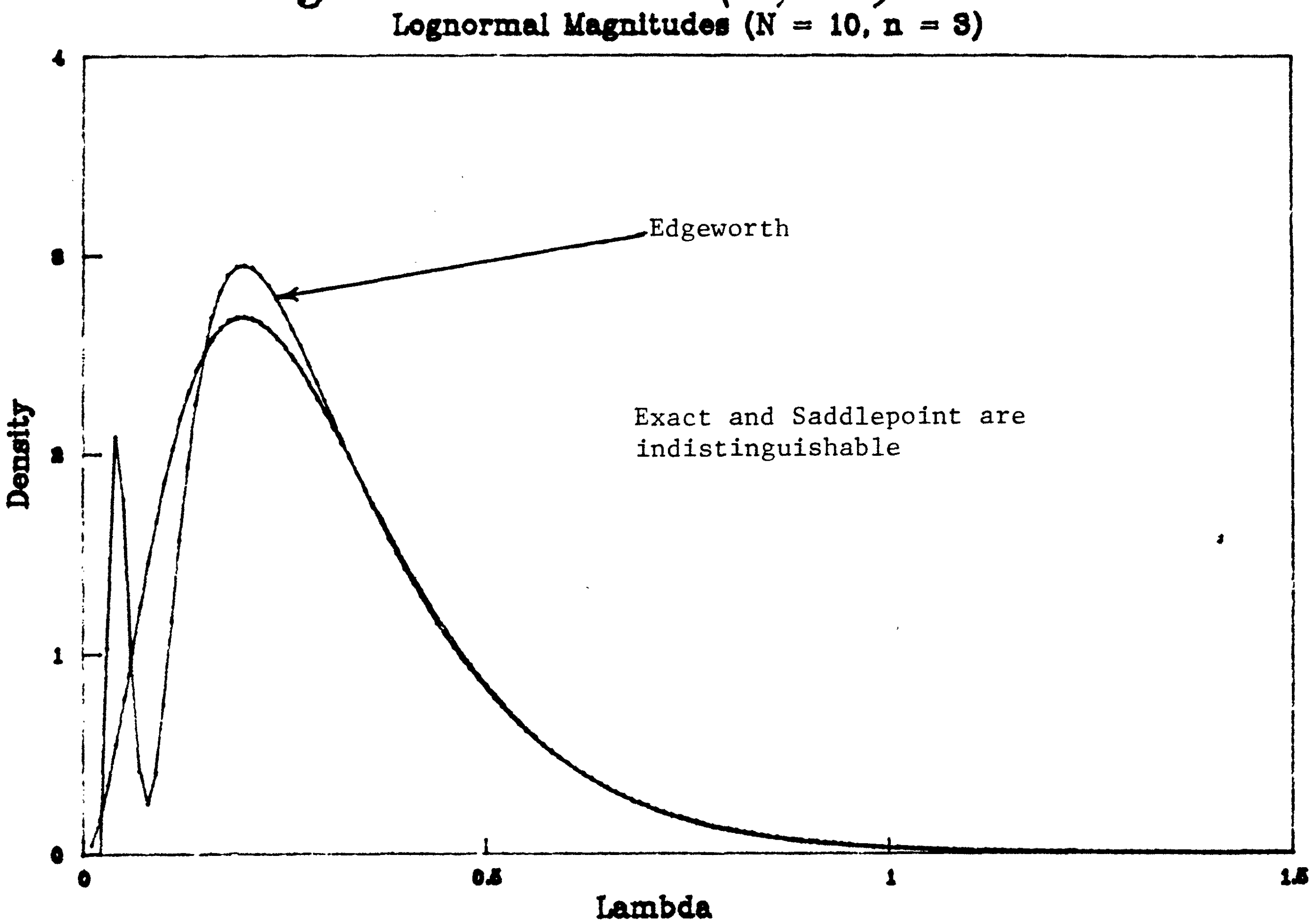
Leading Terms



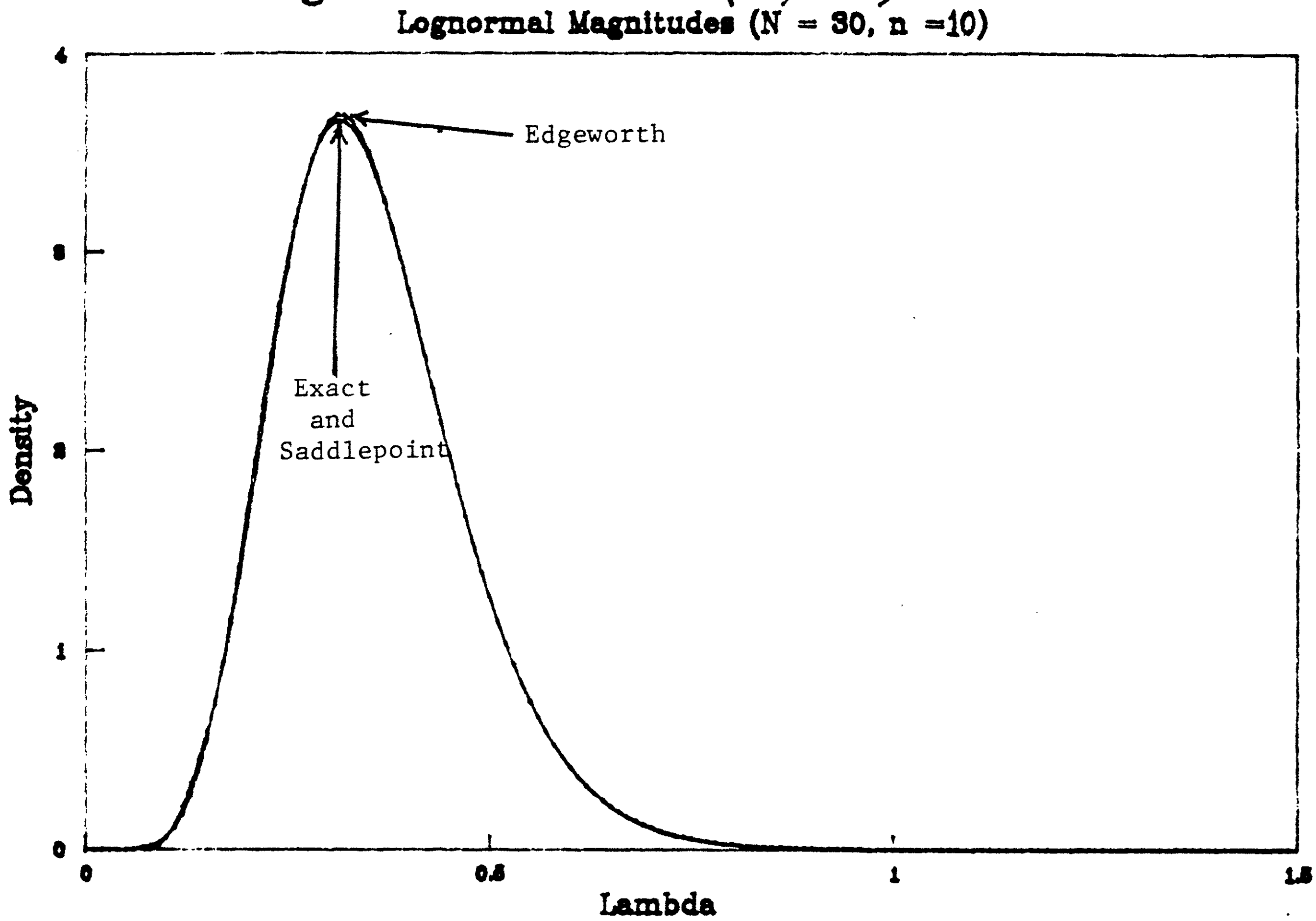


Leading Terms + O(1/N) Corrections
Lognormal Magnitudes (N = 6, n = 2)
Edgeworth
Saddlepoint
Exact
Density
Lambda

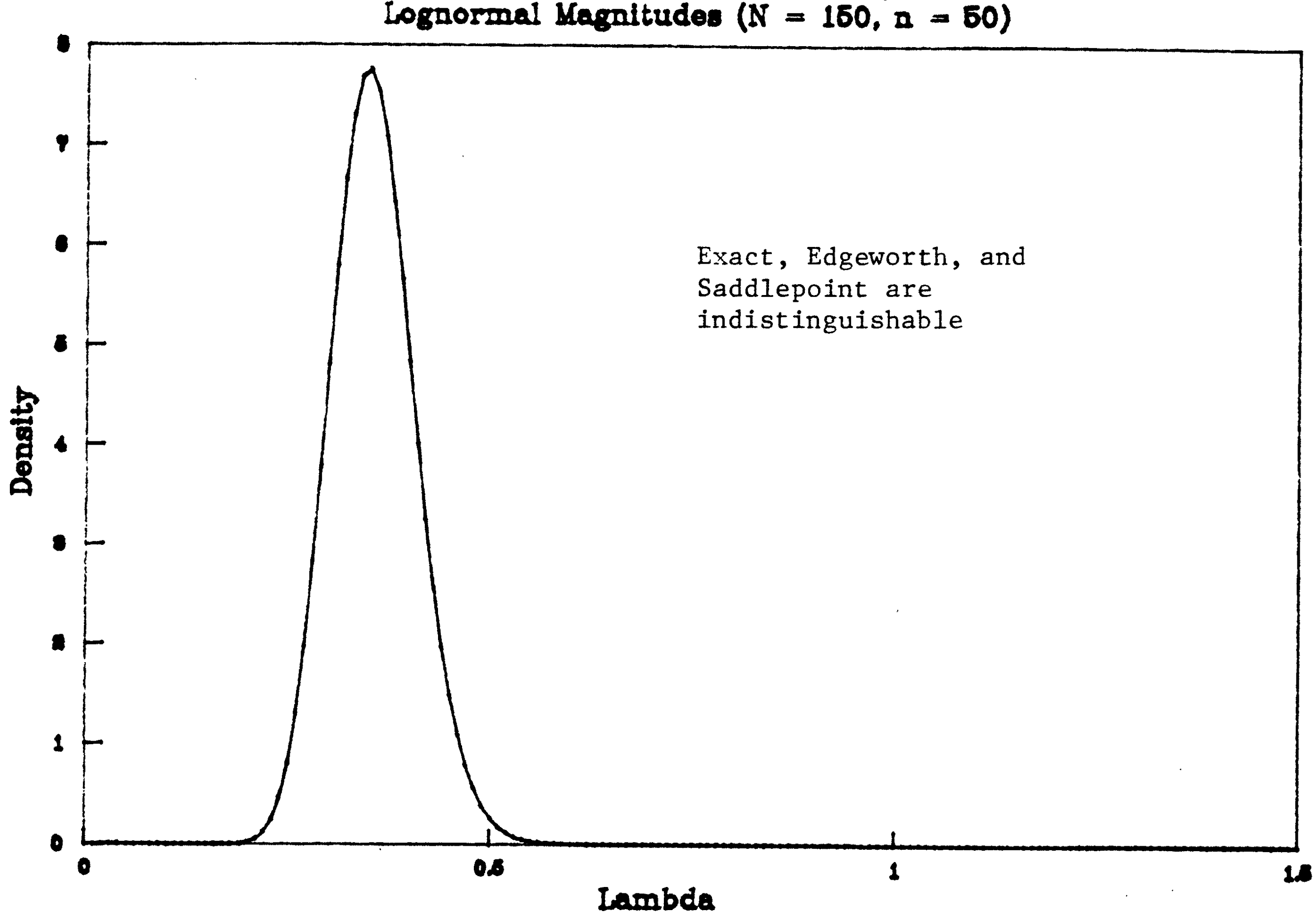
Leading Terms + O(1/N) Corrections



Leading Terms + O(1/N) Corrections



Leading Terms + O(1/N) Corrections



7. INCLUSION PROBABILITIES

An exact integral representation of $P\{k\epsilon\tilde{s}_n\} \equiv \pi_k(n)$ can be constructed in the same fashion as (2.1). The probability that the kth element of U is not in a sample of size n is

$$1 - \pi_k(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \int_0^{\infty} e^{-i(n-1)u} a_k \prod_{\substack{j=1 \\ j \neq k}}^{N} (a_j + (1 - a_j)e^{iu}) \times \sum_{\substack{\ell=1 \\ \ell \neq k}}^{N} \frac{y_\ell a_\ell}{(a_\ell + (1 - a_\ell)e^{iu})} d\lambda du \tag{7.1}$$

where as before, $a_j \equiv \exp\{-\lambda y_j\}$. Defining $p_{kj} = y_j/1-y_k$, $j = 1, 2, \ldots, k-1, k+1, \ldots, N$, transforming from λ to $w = (1 - y_k)\lambda$, and integrating over u,

$$1 - \pi_k(n) = \int_0^{\infty} e^{-wy_k/1-y_k} D_{N-1,n}(w)dw \tag{7.2}$$

with $D_{N-1,n}(w)$ interpretable as the density of the nth smallest statistic $\widetilde{w}_{(n)}$ generated by $\{\widetilde{X}_j/p_{kj}, j = 1, 2, \ldots, k-1, k+1, \ldots, N\}$. Thus $1 - \pi_k(n)$ is interpretable as the expectation of $\exp\{-\widetilde{w}_{(n)} y_k/1 - y_k\}$ with respect to $D_{N-1,n}$.

Numerical computation of inclusive probabilities can be done efficiently using a generating function for $\pi_k(n)$, $k = 1, 2, \ldots, N$.

Lemma 7.1: Given U with N elements and the corresponding parameter $\underline{y}_N$, a generating function for $\pi_k(n)$, $n = 1, 2, \ldots, N-1$ is

$$\sum_{n=1}^{N-1} \xi^{n-1} \pi_k(n) \equiv \phi_k(\xi) \tag{7.3}$$

$$= \int_0^{\infty} \left(1 - e^{-\lambda y_k}\right) \prod_{\substack{j=1 \\ j \neq k}}^{N} \left[\left(1 - e^{-\lambda a_j}\right)\xi + e^{-\lambda a_j}\right] \times \sum_{\substack{\ell=1 \\ \ell \neq k}}^{N} \frac{y_\ell e^{\lambda y_\ell}}{\left[\left(1 - e^{-\lambda y_\ell}\right)\xi + e^{-\lambda y_\ell}\right]} d\lambda$$

Proof: Introducing an auxiliary parameter θ, it is evident that

$$\phi_k(\xi) = -\frac{\partial}{\partial\theta} \int_0^{\infty} \left(1 - e^{-\lambda y_k}\right) \prod_{\substack{j=1 \\ j \neq k}}^{N} \left[\left(1 - e^{-\lambda a_j}\right)\xi + e^{-\lambda\theta a_j}\right] \frac{d\lambda}{\lambda} \tag{7.4}$$

at $\theta = 1$. Differentiating under the integral sign yields (7.4). []

An alternative representation of $\pi_k(n)$ is

$$\pi_k(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(n-1)u} \phi_k(e^{iu}) du. \tag{7.5}$$

Using the transformation $x = e^{-\lambda R}$ and defining $p_k = y_k/R$, (7.3) can be written as:

$$\int_0^1 (x^{-p_k} - 1) \prod_{\substack{i=1 \\ i \neq k}}^{N} [(x^{-p_i} - 1)\xi + 1] \sum_{\substack{\ell=1 \\ \ell \neq k}}^{N} \frac{p_\ell}{[(x^{-p_\ell} - 1)\xi + 1]} dx. \tag{7.6}$$

In terms of the elementary symmetric functions $\sigma_{k,\ell}^n(x)$, defined implicitly by:

$$\prod_{\substack{i=1 \\ i \neq k \\ i \neq \ell}}^{N} [(x^{-p_i} - 1)\xi + 1] = \sum_{n=1}^{N-1} \sigma_{k,\ell}^n(x) \xi^{n-1}, \tag{7.7}$$

(7.3) is representable as

$$\sum_{n=1}^{N-1} \left[\int_0^1 (x^{-p_k} - 1) \sum_{\substack{\ell=1 \\ \ell \neq k}}^{N} p_\ell \sigma_{k,\ell}^n(x) dx \right] \xi^{n-1}, \tag{7.8}$$

so,

$$\pi_k(n) = \int_0^1 (x^{-p_k} - 1) \sum_{\substack{\ell=1 \\ \ell \neq k}}^{N} p_\ell \sigma_{k,\ell}^n(x) dx. \tag{7.9}$$

Numerical computation of this integral presents two kinds of difficulties: first, the integrand has a singularity at 0, thus the integral is improper. Second, the number of steps required to compute $\sigma_{k,\ell}^n(x)$ grows exponentially with N.

Transforming from x to $x = t^\alpha$ with $\alpha = \max_k\{p_k^{-1}, k = 1, 2, \ldots, N\}$ changes the improper integral into a proper one. To overcome the second difficulty a recursive scheme to compute $\sigma_{k,\ell}^n(x)$ has been devised.

Consider the identity

$$\prod_{n=1}^{N} (a_n \xi + 1) = \sum_{n=0}^{N} \sigma_n \xi^n, \quad \sigma_0 \equiv 1. \tag{7.10}$$

For $m \geq n \geq 1$, define

$$S_{n,m} = \sum_{(m)} a_{i_1} \cdots a_{i_n} \tag{7.11}$$

where $S_{n,m} = \sum_{(m)} a_{i_1} \cdots a_{i_n}$ is the sum over all possible products of n different a_i's such that $i_1 \leq m$, $i_2 \leq m, \ldots, i_n \leq m$. Then

$$\begin{aligned} S_{1,m} &= a_1 + a_2 + \ldots a_m & m &= 1, 2, \ldots N \\ S_{n,n} &= a_1 \cdot a_2 \ldots a_n & n &= 1, 2, \cdots, N \\ S_{n,m} &= S_{n,m-1} + a_m S_{n-1,m-1} & 1 &\leq n < m \leq N \end{aligned} \tag{7.12}$$

This computation of $\sigma_n = S_{n,N}$ can be done recursively in less than N^2 steps.

Hájek (1981) worked out an example in which $N = 10$, $n = 4$, and $y_k = k/55$, $k = 1, 2, \ldots, 10$. He found good agreement between exact $\pi_k(4)$ values and Rosén's approximation (less than .7% relative difference for all $k = 1, 2, \ldots, 10$).

Table 7.1 presents similar comparisons for another example: $N = 30$, $n = 10$, and for $k = 1, 2, \ldots, 30$ the value of y_k is that of the $(k/31)$st fractile of an exponential distribution with mean one.

Table 7.1. Comparison of exact inclusion probabilities (1) for N=30, and n=10 (2) with Rosén's approximation.

Label k	yk	(1)* $P(k \varepsilon \tilde{s}_n \mid \underline{y}_N)$	(2)** $(1 - e^{-Z_{n,N}yk})$	(2)/(1)
1	3.433987	.825690	.826268	1.000700
2	2.740840	.754816	.752651	.997132
3	2.335375	.699221	.695869	.995206
4	2.047693	.651615	.647840	.994207
5	1.824549	.609188	.605422	.993818
6	1.642228	.570499	.566998	.993864
7	1.488077	.534689	.531606	.994234
8	1.354546	.501196	.498618	.994855
9	1.236763	.469629	.467597	.995673
10	1.131402	.439699	.438225	.996648
11	1.036092	.411186	.410262	.997751
12	.949081	.383920	.383519	.998958
13	.869038	.357760	.357849	1.000250
14	.794930	.332595	.333130	1.001611
15	.725937	.308329	.309263	1.003030
16	.661398	.284883	.286164	1.004496
17	.600774	.262189	.263763	1.006001
18	.543615	.240188	.241999	1.007539
19	.489548	.218828	.220820	1.009102
20	.438255	.198065	.200181	1.010685
21	.389465	.177857	.180043	1.012289
22	.342945	.158169	.160369	1.013905
23	.298493	.138970	.141129	1.015531
24	.255933	.120231	.122295	1.017165
25	.215111	.101925	.103842	1.018805
26	.175891	.084029	.085747	1.020448
27	.138150	.066521	.067991	1.022094
28	.101783	.049382	.050554	1.023740
29	.066691	.032593	.033420	1.025386
30	.032790	.016137	.016573	1.027031

Bibliography

Albers, W., Bickel., P. J., and van Zwet., W. R. (1976) "Asymptotic Expansion for the Power of Distribution Free Tests in the One-Sample Problem," *Annals of Statistics*, Vol. 4, No. 1, pp. 108-156.

Andreatta, G., and Kaufman, G.M. (1986), "Estimation of Finite Population Properties When Sampling is Without Replacement and Proportional to Magnitude," *Journal of the American Statistical Association*, Vol. 81, pp. 657-666.

Barndorff-Nielsen, O.E., and D.R. Cox (1989), *Asymptotic Techniques for Use in Statistics* (New York: Chapman and Hall, 1989).

Cox, D. R. and Barndorff-Nielsen, O. (1979) "Edgeworth and Saddle-point Approximations with Statistical Applications," *Journal of the Royal Statistics Society (B)*, Vol. 25, pp. 631-650.

Daniels, H. E. (1954), "Saddle-point Approximations in Statistics," *Annals of Mathematical Statistics*, Vol. 25, pp. 631-650.

Field, C. A., and Hampel, F. R., "Small-Sample Asymptotic distributions of M-estimators of location," *Biometrika* (1982), 69, pp. 29-46.

Goel, A., "Software Reliability Models: Assumptions, Limitations, and Applicability," *IEEE Trans. Software Eng.* Vol. SE-11, No. 12 (April 1985), pp. 375-386.

Good, I. J. (1957) "Saddle-point Methods for the Multinomial Distribution," *Annals of Mathematics Statistics*, Vol. 28, pp. 861-881.

Gordon, L. (1983) "Successive Sampling in Large Finite Populations," *Annals of Statistics*, Vol. 11, No. 2, pp. 702-706.

Gordon, L. (1989), "Estimation for Large Successive Samples of Unkonwn Inclusion Probabilities," *Advances in Applied Mathematics*, (to appear).

Hájek, J. (1981), *Sampling from a Finite Population*, (New York: Marcel Decker, Inc.), 247 pp.

Kaufman, G. M. (1989), "Conditional Maximum Likelihood Estimation for Successively Sampled Finite Populations," M.I.T. Sloan School of Management Working Paper.

Kendall, M. G., and Stuart, A. (1969) *The Advanced Theory of Statistics*, Vol. I, (London: Griffin).

Khinchin, A.I. (1949) *Mathematical Foundations of Statistical Mechanics* (New York: Dover Publications, 1949).

Langberg, N. and Singpurwalla, N. (1985), " A Unification of Some Software Reliability Models,: *SIAM Journal on Scientific and Statistical Computing*, 6(3), pp. 781-790.

Levinson, N. and Redheffer, R. M. (1970), *Complex Variables*, (San Francisco: Holden-Day).

Miller, D.R., "Exponential Order Statistic Models of Software Reliability Growth," *IEEE*, Trans. Software Eng. Vol. SE-12 No. 1, Jan. 1986, pp. 12-24.

Rosén, B. (1972), "Asymptotic Theory for Successive Sampling with Varying Probabilities without Replacement," I and II, *Annals of Statistics*, Vol. 43, pp. 373-397, 748-776.

Ross, s. M., "Software Reliability: The Stopping Rule Problem," *IEEE Trans. Software Eng.* Vol. SE-11, No. 12, (Dec. 1985), pp. 1472-1472.

Scholz, F. W. (1986), "Software Reliability Modeling and Analysis," *IEEE Transactions on Software Engineering*, Vol. SE-12, No. 1, Jan 1986.

Sen, P. K. (1968), "Asymptotic Normality of Sample Quantiles for m-Dependent Processes," *Annals of Mathematical Statistics*, Vol. 39, No. 5, pp. 1725-1730.