SOME ESTIMATES OF LOCATION

BY
MELVILLE R. KLAUBER

TECHNICAL REPORT NO. 6
April 15, 1964

PREPARED UNDER THE AUSPICES
OF
PUBLIC HEALTH SERVICE GRANT 2G-25

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
SOME ESTIMATES OF LOCATION

by

Melville R. Klauber

TECHNICAL REPORT NO. 6

April 15, 1964

PREPARED UNDER THE AUSPICES

OF

PUBLIC HEALTH SERVICE GRANT 2G-25

DEPARTMENT OF STATISTICS

STANFORD UNIVERSITY

STANFORD, CALIFORNIA
SOME ESTIMATES OF LOCATION

by

Melville R. Klauber

SECTION 1

INTRODUCTION

1.1. Location Problems.

A common and important statistical problem is the estimation of a location parameter of a distribution. A useful criterion of performance of an estimate is the mean square error (m.s.e.); estimates will be compared throughout this paper on the basis of their m.s.e.'s. The m.s.e.'s of most commonly used estimates are asymptotically of the order, \( N^{-1} \), where \( N \) represents the sample size. Favorable exceptions to this are unusual. A notable exception is the sample mid-range used as the estimate of the mean of the uniform distribution, the low asymptotic variance of which was given by R. A. Fisher [1]. E. H. Lloyd [2] has shown that the same estimate is the best linear function of the order statistics. It should be noted that this and other examples of low asymptotic m.s.e., a number of which are given for the first time below, take advantage of edge effects.

The case of the triangular distribution has been treated by A. E. Sarhan [3], in which the coefficients for the best linear functions of the order statistics have been computed up to \( N = 5 \).

In the sequel these problems are extended in a number of ways. Cases are discussed where the edges are not so sharp as in the cases
above, i.e., observations can occur with relatively low probability
outside the region to be located, and conditions within the region to
be located are greatly relaxed. Of considerable interest is the extension
of the results to two dimensions.

A one-dimensional problem related to those mentioned in the last
paragraph has been discussed by H. Chernoff and H. Rubin [4]. They
treat the problem of locating a discontinuity in a density, and give an
heuristic argument to the effect that the maximum likelihood estimate is
one of the observations occurring near the point of discontinuity.

1.2 An Application.

Consider now a physical problem that led to the theoretical statisti-
cal problems discussed above. The usual photographs of planets taken
through telescopes are very unclear, and this has led to controversy
about the exact nature of the markings which have been seen and photo-
grahed. The blurring of photographs is due to a great number of causes,
including: turbulence, wind, thermal inhomogeneities, or water vapor in
the terrestrial atmosphere, causing fluctuations in the direction of the
light wave front [5]; diffraction in the optical instruments [6], photo-
graphic fog in development, and scattering of light in the plane of the
photographic emulsions [7]. Even when "good seeing" conditions exist
these factors prevent clear pictures of the planets. Much of the con-
troversy with respect to the nature of the surface of the planet Mars
is derived from the fact that the human eye compensates for some of the
atmospheric disturbance, whereas during the period of exposure the
photographic plate cannot. Optical systems have been designed with the
intent of reducing this effect as much as possible; still astronomers
may see more things than they can photograph.

Dr. Elliot Levinthal of Stanford University has proposed a method for materially reducing the distortion due to disturbances during the period of exposure. In addition to translation, this distortion may involve rotation, expansion, and contraction. Assume that translation causes some major part of the distortion. This can be stopped in an exceedingly short exposure, but then the image will comprise only a number of isolated activated silver halide grains on the plate, and no picture. If a large number of such short exposures are taken, superimposing without alignment of the images reintroduces the blurring due to translation.

How can one best align the images? At the outset one might be tempted to align them by superimposing their centroids. On the other hand, consider the case of the uniform distribution and the favorable use of edge effects. The question here is whether better alignment can be accomplished by using the latter approach.

1.3 The General Model.

We now discuss in general terms a model having the elements of the theoretical and applied problems already mentioned. Since exposure time is variable and the number of observation points depends on the number of quanta that happen to arrive at the photographic plate, a Poisson process is the relevant model. We will confine our attention to a known bounded, connected set, $T$, on the real line or in the plane. $T$ is analogous to the photographic plate mentioned in the application. The mean value of this Poisson process has the following form:
(1.1) \( I(x) = 0 \) if \( x \in T^c \)

\[ = \lambda(x) \text{ if } x \in S, \text{ a closed interval (or disc) contained in } T, \]

\[ = \mu(x) \text{ if } x \in U, \exists x \in T-S \text{ and includes the boundary of } T. \]

In the remaining set, \( T - U - S, \lambda' \geq I(x) \geq \mu' \),

where

\[ \mu' = \sup_{x \in U} \mu(x) \]

and

\[ \lambda' = \inf_{x \in S} \lambda(x). \]

Diagram: Showing Partitioning of the Rectangle \( T \).

Figure 1.1

In Figure 1.1 we show in a Venn diagram the sets mentioned above. \( T \) is represented by the rectangle and its entire interior. We make the following additional identifications: \( S, U, \) and \( T - U - S \) are analogous to portions of a photograph showing a planet, space, and scattering of light at the edge of the planet, respectively. We wish to estimate the location of \( S \); in most cases we will estimate the center for convenience.
Some further comments are in order on an assumption made in the theorems to follow and how it relates to the physical situation. The treatment, where \( \mu \) is a constant and \( \lambda = K t \) proportional to exposure time \( t \), leads to error terms in powers of \( \lambda^{-1} \) but not in powers of \( \mu/\lambda \). Physically it is somewhat artificial to take \( \mu \) as a constant since it must increase as the exposure time, \( t \), does. On the other hand, it is mathematically convenient and for the following (physically reasonable) circumstances it is realistic:

\[
(1.2) \quad \mu(t) = \mu_0 + \beta t
\]

\[
(1.3) \quad \lambda(t) = \mu_0 + \gamma t,
\]

where \( \beta << \gamma \) and \( \mu_0 \) represents the photographic noise in an unexposed developed plate. In (1.2) and 1.3) we have suppressed the location, \( x \), which was the argument in the function (1.1). When \( t \) is small enough that the "noise" from the exposure is only a fraction, \( c < 1 \), of the "noise" produced in the photographic processing we have \( t = c \mu_0/\beta \) and

\[
\frac{\mu(t)}{\lambda(t)} = \frac{\mu_0 (1+c)}{\mu_0 + \lambda t},
\]

which for \( \sqrt{\beta/\gamma} < c < 1 \) implies

\[
(1.4) \quad \frac{\mu_0}{\gamma} \cdot \frac{1}{t} < \frac{\mu(t)}{\lambda(t)} < \frac{2\mu_0}{\gamma} \cdot \frac{1}{t}.
\]

When \( t \) is "large", but is still small enough that \( t < c \mu_0/\beta \) the ratio \( \mu(t)/\lambda(t) \) is of the same order of magnitude as \( t^{-1} \). Of course for \( t \to \infty \),
(1.5) \[ \frac{\mu(t)}{\lambda(t)} = \beta/\gamma, \] a constant

but the situation (1.4) is more realistic that (1.5), since the ideal of exposing a plate so that \( \mu(t) \to \infty \) is simply not physically realistic.

In this paper a number of models, with differing complexity and approximation of the physical situation described above, will be considered. In the most complex case, the boundary of a disc is assumed to be a circle, (although in fact all planets, to greater and lesser degrees, go through phases and are very slightly ellipsoidal). In this case the following is assumed: a high varying intensity over the disc, a low background intensity, and a decreasing intensity along any radius in the neighborhood of the edge.\(^1\)

\(^1\)Professor Rudolf Minkowski of the University of California Department of Astronomy has communicated, that due to the great variety in optical and photographic procedures and their differing effects on the distortions mentioned above, we should use as an approximation, either a two-dimensional normal error distribution for each observation point, or assume, as we have, that the intensity decreases linearly along any radius in the neighborhood of the edge.
2.1 Uniform Intensity on S.

Let $\psi$ be the center of a closed interval of length two on the real line contained in an interval, $T$, of known location and length. For the purpose of determining the mean square error of an estimate $\hat{\psi}$ of $\psi$, we may, without loss of generality, assume $\psi = 0$. We assume observations occur in $S$ and $T - S$ according to the Poisson distribution with parameters $\lambda$ and $\mu (\mu < \lambda)$ per unit distance, respectively, and 0 elsewhere. We have assumed that the length of $S$ is scaled to be two, but if will be treated as if it were unknown. Consider the estimate, $\hat{\psi}$, of the type, $\hat{\psi} = (X_u + X_v)/2$, where $X_1, X_2, \ldots$ are observation points, and

$$X_v = \min \left\{ X_i : [X_i, X_i + c] \text{ contains } N \text{ or more observations} \right\}, \quad X_i < 0$$

$$= 0, \text{ otherwise, and}$$

$$X_u = \max \left\{ X_i : [X_i - c, X_i] \text{ contains } N \text{ or more observations} \right\}, \quad X_i > 0$$

$$= 0, \text{ otherwise.}$$

Note that if $\mu = 0, N = 1$, and $c = 1$, we have,

$$\text{m.s.e. } \hat{\psi} = E(X_u + X_v)^2 / 4,$$

where $-X_u$ and $X_v$ are two independent random variables distributed according to the cumulative distribution function (c.d.f.),
\[ P_X(x) = 1 - e^{-\lambda(x+1)} \quad \text{for} \quad -1 \leq x < 0 \]
\[ = 0 \quad \text{for} \quad x < -1 \]
\[ = 1 \quad \text{for} \quad x \geq 0 . \]

\[
\text{m.s.e.} \hat{\Psi} = \frac{(\mathbb{E}_u^2 + 2 \mathbb{E}_u \mathbb{E}_v + \mathbb{E}_v^2)}{4} \\
= \frac{(\mathbb{E}_u^2 - \mathbb{E}_v^2)}{2} = \frac{\text{Var } X_u}{2} \\
= \frac{(1/\lambda^2 - 2e^{-\lambda}/\lambda - e^{-2\lambda}/\lambda^2)}{2} \\
= (1/2) \lambda^{-2} + o(e^{-\lambda}).
\]

We now show that by a suitable selection of \( c \), we may attain essentially the same m.s.e. \( \hat{\Psi} \) when \( 0 < \mu < < \lambda \).

**Theorem 2.1:** Let \( T \supset -1 - c, T \supset 1 + c, c = (k \log \lambda) / \lambda, \lambda > > \mu > 0 \), \( \mu \) is a constant, \( N \geq 4 \), and \( k \geq 3 \), then

\[ \text{m.s.e.} \hat{\Psi} \leq (1/2) \lambda^{-2} + o(\lambda^{5-3}) \quad \text{for any} \quad \delta > 0 . \]

**Lemma 2.1:** \( P(X_u > 1 + c) = o(\lambda^{5+1-N}) \).

\( P(X_v < -1 - c) = o(\lambda^{5+1-N}) \), for any \( \delta > 0 \).

**Proof:** We prove the first assertion; the second follows by the same argument.

\[ P(\text{there are exactly } m \text{ observations} > 1) = \\
\quad e^{-\mu D} (\mu D)^m / m! , \]

where \( D \) is the length of \((1, \infty) \cap T\).

Given that there are \( m \) observations \( > 1 \), the probability that any given one has \( N-1 \) or more observations in the interval of length \( c \) to its left is
\[
\leq \sum_{n=N-1}^{m-1} \binom{m-1}{n} (c/D)^n (1-c/D)^{m-1-n}.
\]

The probability that this event occurs for any of the \( m \) observations must be less than \( m \) times this quantity, hence

\[
P(X_u > 1 + c) \leq \sum_{m=N}^{\infty} e^{-\mu D} (\mu D)^m / m!
\]

\[
\cdot \sum_{n=N-1}^{m-1} \binom{m-1}{n} (c/D)^n (1-c/D)^{m-1-n}
\]

\[
\leq \sum_{n=N-1}^{\infty} \left[ (c/D)^n / n! \right]
\]

\[
\cdot \sum_{m=1}^{\infty} e^{-\mu D (\mu D)^m} m(m-1) \cdots (m-n)/m!
\]

\[
= \sum_{n=N-1}^{\infty} M_{n+1} (c/D)^n / n!, \quad \text{where}
\]

\( M_{n+1} \) is the \((n+1)\)th factorial moment of the Poisson distribution.

Hence we have

\[
P(X_u > 1 + c) \leq \sum_{n=N-1}^{\infty} \left[ k \log (\lambda D) (\mu D)^{n+1} / n! \right]
\]

\[
= O(\lambda^{\delta+1-N}), \quad \text{for any } \delta > 0.
\]

**Lemma 2.2:**

\[
P(1 \leq X_u \leq 1 + c) = O(\lambda^{\delta-1})
\]

\[
P(-1 - c \leq X_v \leq -1) = O(\lambda^{\delta-1}),
\]

for any arbitrarily small \( \delta > 0 \).
Proof: Again both statements follow from the same argument.

\[ P(1 < X_u < 1+c) \leq 1 - e^{-\mu c} \]
\[ = 1 - 1 + O(\mu c) \]
\[ = O(\mu \log \lambda / \lambda) = O(\lambda^{5-1}) , \]

for any \( \delta > 0. \)

We now return to the proof of the theorem. For \( N \) fixed, \( X_u \) and \(-X_v\) are asymptotically (as \( \lambda \to \infty \)) independently distributed, and we have

\[ (2.3) \quad \text{m.s.e.} \; \hat{\psi} = \frac{E(X_u + X_v)^2}{4} = \frac{\text{Var } X_u}{2} + O(\lambda^{5+1-N}) , \]

since \( X_u \) and \(-X_v\) are bounded and their c.d.f.'s differ by \( O(\lambda^{5+1-N}) \), by symmetry and Lemma 2.1.

If we transform all the observations by \( g(X) = 1 - X \), we have

\[ (2.4) \quad \text{Var } X_u = \text{Var } Z, \; \text{where } Z = 1 - X_u \]
\[ (2.5) \quad \text{Var } Z = E Z^2 - E^2 Z. \]

Order the observations contained in \([0, 1]\) downward, \(X_1, X_2, \ldots\), i.e., \(X_1\) is the largest, and let \(Z_1 = 1 - X_1\).

\[ (2.6) \quad E Z^2 \leq D^2 \; P(Z < -c) + c^2 \; P( -c \leq Z \leq 0) \]
\[ + E(\zeta^2 |Z = Z_1) + P(\zeta = Z_2 \text{ or } Z_3 \text{ or } \ldots) . \]
\[ (2.7) \quad E Z \geq E Z_1 \; P(Z = Z_1) - D \; P(Z < -c) \]
\[ - c P( -c \leq Z \leq c) . \]
In the sequel we shall bound various terms in (2.6) and (2.7), and to facilitate the argument we shall denote \( P(Z < -c) \) by \( A_1 \), \( \sigma^2 \) by \( A_2 \), \( E(Z^2 | Z = Z_1) \) by \( A_3 \), \( P(Z = Z_2 \text{ or } Z_3 \text{ or } ... ) \) by \( A_4 \), \( E[Z_1 | Z = Z_1] \) by \( B_1 \), \( -DP(Z < -c) \) by \( B_2 \), and \( -cP(-c \leq Z \leq 0) \) by \( B_3 \).

\[
P(Z < -c) = P(X_u > 1 + c) = O(\lambda^{\delta+1-N}) \text{ by Lemma 2.1, hence}
\]

(2.8) \( A_1 = O(\lambda^{\delta-3}) \), if \( N \geq 4 \).

(2.9) \( A_2 = O(\lambda^{\delta-3}) \), since \( c^2 \sigma^2 \leq Z \leq 0 = c^2 \sigma^2 (1 \leq X_u \leq 1 + c) \)

\[
= (k \log \lambda / \lambda)^2 \sigma^2 (1 \leq X_u \leq 1 + c) = O(\lambda^{\delta-3}) \text{ by Lemma 2.2.}
\]

\( A_3 = E(Z^2 | Z = Z_1) = E[Z_1^2] \), since \( Z = Z_1 \)

if and only if \( Z_N - Z_1 < c \) (given \( Z \) is not less than 0), and the latter event is independent of \( Z_1 \).

Let \( m(\cdot, \cdot) \) be the number of observations in the interval designated.

(2.10) \( A_4 = P(Z = Z_2 \text{ or } Z_3 \text{ or } ... ) = P(m(Z_1, Z_1 + c) < N - 1) \)

\[
= \sum_{i=0}^{N-1} e^{-\lambda} c (\lambda c)^i / i! = \lambda^{-k} \sum_{i=0}^{N-1} (k \log \lambda)^i / i! \\
= O(\lambda^{\delta-k}) = O(\lambda^{\delta-3}), \text{ for any } \delta > 0 \text{ and } k \geq 3 .
\]

\( B_1 = EZ_1 P(Z = Z_1) = EZ_1 [1 - A_4 - P(Z < 0)] \)

\[
= E[Z_1] [1 + O(\lambda^{\delta-1})] \text{ for any } \delta > 0 \text{ by Lemmas 2.1, 2.2, and (2.10).}
\]

\( B_2 = O(\lambda^{\delta+1-N}) \) by (2.8) and since \( B_2 = -A_1 / D, \)

\[
= O(\lambda^{\delta-3}), \text{ for any } \delta > 0, \text{ if } N \geq 4 .
\]
\[ B_3 = \sigma^2 \text{ for any } \sigma > 0 \text{ by (2.9) and since } B_3 = -A_2 / c. \]

\[ E(Z_1) = 1/\lambda, \text{ the mean of an exponentially distributed random variable. Below we bound the terms in } \text{Var } Z. \]

\[ \text{Var } Z \leq (A_1 + A_2 + A_3 + A_4) - (R_1 + R_2 + R_3)^2 \]

\[ \leq E Z^2 + O(\lambda^{-3}) - [E Z_1 [1 + O(\lambda^{-1})] + O(\lambda^{-2})]^2 \]

\[ = E Z^2 + O(\lambda^{-3}) - [E Z_1 + O(\lambda^{-2})]^2, \]

since \[ E Z_1 = \lambda^{-1}, \]

\[ = \text{Var } Z_1 + O(\lambda^{-3}), \text{ again because } E Z_1 = \lambda^{-1}, \]

(2.11) \[ = 1/\lambda^2 + O(\lambda^{-3}), \text{ since } \text{Var } Z_1 = 1/\lambda^2 + o(e^{-\lambda}) \text{ by (2.2)}. \]

This completes the proof of Theorem 2.1, since for \( N \geq 4 \) in (2.3)

\[ \text{m.s.e. } \hat{\psi} = \text{Var } X_u / 2 + O(\lambda^{8+1-N}) \]

\[ = \text{Var } Z / 2 + O(\lambda^{8+1-N}) \text{ by (2.4)} \]

\[ \leq (1/2) \lambda^{-2} + O(\lambda^{-3}) \text{ by (2.11)}. \]

It should be observed that the m.s.e. \( \hat{\psi} \) obtained above is analogous to the m.s.e. of the maximum likelihood estimate of the median of the rectangular distribution. The maximum likelihood estimate is the mid-range and has variance, \( 1/ [2(m+1)(m+2)] \), for sample size \( m \).

For comparison we compute the order of the variance of the sample mean \( \bar{X} \). The conditional variance of the sample mean given \( m \) observations,

\[ \text{Var } \bar{X}_m = \sigma^2 / m, \]

12.
where \( \sigma_u^2 \) is clearly some constant between \( (\text{length of } T)^2 / 12 \) and \( 1/3 \), corresponding to the variance of the uniform distribution over \( T \) (see beginning of this section) and the variance of the uniform distribution over \([-1, 1]\). Since \( m \) is a random variable of order \( \lambda \),

\[
\text{Var } X = \sigma_u^2 O(\lambda^{-1}).
\]

2.2 **Trapezoidal Intensity on \( S \).**

We will call "trapezoidal" the case where the intensity function of the Poisson process increases linearly from \( \mu \) to \( \lambda \) in the region \( 1 - \xi \leq |x| \leq 1, 0 < \xi \leq 1 \). The m.s.e. of an estimate \( \hat{\psi} \) will now be derived for the case where observations are distributed per unit distance as follows: \( I(x) = \)

\[
\begin{align*}
(2.12a) & \quad \lambda, & \text{for } |x| \leq 1 - \xi, \ 0 \leq |\xi| \leq 1 \\
(2.12b) & \quad \lambda - (|x| - 1 + \xi)(\lambda - \mu) / \xi, & \text{for } 1 - \xi \leq |x| \leq 1 \\
(2.12c) & \quad \mu, & \text{for } |x| > 1, \ x \in T = [A, B] \\
(2.12d) & \quad 0, & \text{otherwise.}
\end{align*}
\]

![Intensity Function, I(x), Trapezoidal Case](image)

**Figure 2.1**
Since $\lambda > \mu$, we may for mathematical simplicity let $I(x) = (1 - |x|) \lambda / \xi$ in (2.12b); the effect of this is a negligible overstatement of $\text{Var} Z_M$ given below (2.17). We define $Z_M$ in the same way as in the case of the "uniform" intensity function and obtain the density for any fixed $M$.

\begin{equation}
I_{Z_M}(t) = (\lambda \ t^2/2 \ \xi) \ (\lambda \ t/\xi) \ e^{\lambda t^2/2 \xi} / (M-1)!
\end{equation}

\begin{align*}
&\quad + O(\lambda^{M-1} e^{-\theta \lambda}), \text{ for } 0 \leq t < \xi \leq 1, \ 1/2 \leq \theta \leq 1; \\
&= \lambda^M (t - \xi/2)^{M-1} e^{-\lambda(t - \xi/2)} / (M-1)!
\end{align*}

\begin{equation}
+ O(\lambda^{M-1} e^{-\theta \lambda}), \text{ for } \xi \leq t \leq 1,
\end{equation}

\begin{equation}
= 0, \text{ otherwise [8]. Note that } \lambda t^2/(2 \xi)
\end{equation}

is the area in a triangular portion of the intensity function to the right of $1 - t$, and \( \lambda \theta = \int_0^1 I(x)dx \).

In the derivation of the mean and variance of $Z_M$ and in the sequel we designate the incomplete gamma function with parameters $\alpha, M$, by

\[ \Gamma(\xi; \alpha, M) = \int_{0}^{\xi} e^{-\alpha x} (\alpha x)^{M-1} \alpha^{-1} (M) \ dx. \]

\begin{equation}
E Z_M^k = \int_{0}^{\xi} t^k e^{[-\lambda t^2/(2 \xi)] \ [\lambda t/\xi]^{M-1}} \exp[-\lambda t^2/(2 \xi)] [(m-1)!]^{-1} \ dt
\end{equation}

\begin{align*}
&\quad + \int_{\xi}^{1} t^k \lambda^m (t - \xi/2)^{M-1} \exp[-\lambda(t - \xi/2)] [(m-1)!]^{-1} \ dt \\
&\quad + O(R), \text{ where } R \text{ decreases exponentially in } \lambda.
\end{align*}

We may change the upper limit of integration in the second integral to $\infty$ below, since it changes the result by an amount only of order $R$. Let $u = t^2$ and $v = t - \xi/2$ in the first and second integrals, respectively, of (2.14).
(2.15) \( E_{Z_M}^k = \int_0^{\xi^2} u^{k/2} \left[ \frac{\lambda u}{(2\xi)} \right]^{M-1} \exp \left[ -\frac{\lambda u}{(2\xi)} \right] (m-1)!^{-1} du \)
\[ + \int_{\xi/2}^{\infty} (v+\xi/2)^k \lambda(v)^{M-1} \exp \left[ -\lambda v \right] (m-1)!^{-1} dv + R. \]

Letting \( k = 1 \) in (2.15) we obtain

(2.16) \( E_{Z_M}^1 = \left( \frac{2\xi}{\lambda} \right)^{1/2} \left[ \Gamma(M+1/2)/\Gamma(M) \right] \Gamma(\xi^2; \lambda/(2\xi), M+1/2) \)
\[ + (M/\lambda)[1-\Gamma(\xi/2; \lambda, M+1)] + (\xi/2)[1-\Gamma(\xi/2; \lambda, M)] + R. \]

\( E_{Z_M}^2 = \int_0^{\xi^2} u e^{-\lambda u/2\xi}(\lambda u/2\xi)^{M-1}(\lambda/2\xi)\Gamma^{-1}(M) du \)
\[ + \int_{\xi/2}^{\infty} (v^2+\xi^2+\xi^2/4)e^{-\lambda v}(\lambda v)^{M-1}\Gamma^{-1}(M) dv + R \]
\[ = (2\xi/\lambda)\Gamma(\xi^2; \lambda/2\xi, M+1) + [M(M+1)/\lambda^2][1-\Gamma(\xi/2; \lambda, M+2)] \]
\[ + (M/\lambda)[1-\Gamma(\xi/2; \lambda, M+1)] + (\xi^2/4)[1-\Gamma(\xi/2; \lambda, M)]. \]

\( E_{Z_M}^2 = (2\xi/\lambda)[\Gamma(M+1/2)/\Gamma(M)]^2\Gamma^2(\xi^2; \lambda/2\xi, M+1/2) \)
\[ + (M^2/\lambda^2)[1-\Gamma(\xi/2; \lambda, M+1)]^2 + (\xi^2/4)[1-\Gamma(\xi/2; \lambda, M)]^2 \]
\[ + (2^{3/2}\xi^{1/2}/M[\lambda^{3/2}][\Gamma(M+1/2)/\Gamma(M)]\Gamma(\xi^2; \lambda/2\xi, M+1/2)[1-\Gamma(\xi/2; \lambda, M+1)] \]
\[ + (2^{3/2}\xi^{3/2}/\lambda^{1/2})[\Gamma(M+1/2)/\Gamma(M)]\Gamma(\xi^2; \lambda/2\xi, M+1/2)[1-\Gamma(\xi/2; \lambda, M)] \]
\[ + (M/\lambda)[1-\Gamma(\xi/2; \lambda, M+1)][1-\Gamma(\xi/2; \lambda, M)] + R. \]

(2.17) \( \text{Var } Z_M = (\xi^2/4)\Gamma(\xi/2; \lambda, M)[1-\Gamma(\xi/2; \lambda, M)] \)
\[ - (2^{1/2}\xi^{3/2}/\lambda^{1/2})[\Gamma(M+1/2)/\Gamma(M)]\Gamma(\xi^2; \lambda/2\xi, M+1/2)[1-\Gamma(\xi/2; \lambda, M)]. \]
\[ + \left( \frac{\xi}{\lambda} \right) \partial \Gamma(\xi^2; \lambda/2 \xi, M+1) - 2^2(M+1/2) \Gamma^{-2}(M) \Gamma^2(\xi^2; \lambda/2 \xi, M+1/2) \\
+ \Gamma(\xi/2; \lambda, M) [1 - \Gamma(\xi/2; \lambda, M+1)] \\
- (2^{3/2} \xi^{1/2} M/\lambda^{3/2}) \Gamma(M+1/2) \Gamma(M) \Gamma(\xi^2; \lambda/2 \xi, M+1/2) [1 - \Gamma(\xi/2; \lambda, M+1)] \\
+ (1/\lambda^2) \left( M(M+1) [1 - \Gamma(\xi/2; \lambda, M+2) - \Gamma^2(1-\Gamma(\xi/2; \lambda, M+1))] \right) + R. \]

In order to obtain a better bound on the m.s.e. than can be obtained by an exact repetition of the arguments of the preceding section, a minor change in the estimation procedure is made. Consider estimates \( \psi^* \) of the type, \( \psi^* = (X_u + X_v)/2 \), where \( X_1, X_2, \ldots \) are the observation points, and

\[ X_v = \min_{X_i < 0} \left\{ X_i : [X_i, -c, X_i] \text{ contains } N \text{ or more observations} \right\}, \]

if such an \( X_i \) exists, = 0, otherwise, and

\[ X_u = \max_{X_i > 0} \left\{ X_i : [X_i, X_i + c] \text{ contains } N \text{ or more observations} \right\}, \]

if such an \( X_i \) exists, = 0, otherwise.

**Theorem 2.2:** Given the Poisson intensity function (2.12), then there exists a \( c \) and an integer \( N \) such that,

\[ \text{m.s.e. } \psi^* \leq \text{Var } Z_N/2 + O(\lambda^\delta), \text{ for any } \delta > 0. \]

**Proof:** Since \( \xi \) is fixed and \( \lambda \) is a parameter which may be varied, we choose \( \lambda \) and \( K \) such that

\[ c = (K \log \lambda/\lambda)^{1/2} < \xi, \text{ where } K/(2\xi) = 3, \text{ i.e.,} \]

\[ \xi > 6 \log \lambda/\lambda. \]
We need the following lemma.

Lemma 2.3: \[ P(X_u > 1) = O(\lambda^{8+(1-N)/2}) \]
\[ P(X_v < 1) = O(\lambda^{8+(1-N)/2}) \quad \text{for any } \delta > 0. \]

Proof: Note that the conditions here are somewhat analogous to those of Lemma 2.1. If we change "left" to "right" and "1 + c" to "1" and use the new value of \( c \) in the proof of Lemma 2.1, the result,

\[
P(X_u > 1) = \sum_{n=N-1}^{\infty} (K \log \lambda/\lambda D)^{n/2}(\mu D)^{n+1}/n! = O(\lambda^{8+(1-N)/2}),
\]
is obtained.

We return to the theorem. Likewise, m.s.e. \( \psi^* = \text{Var } X_u/\hat{\epsilon} + O(\lambda^{8+(1-N)/2}) \) is also obtained. Again transform the observations contained in \([0,1]\) by \( g(X) = 1 - X \), and let \( Z = 1 - X_u \).

\[ E Z^2 \leq D^2 P(Z < 0) + E(Z^2|0 \leq c) P(0 \leq Z \leq c) + P(c < Z \leq 1). \]

Since \( 0 \leq Z \leq Z_N \), given \( Z \in [0,c] \), we have

\[ E(Z^2|0 \leq Z \leq c) P(0 \leq Z \leq c) \leq E Z_N^2 \]

and hence

\[ \text{Var } X_u = \text{Var } Z \leq E Z_N^2 + D^2 P(Z < 0) + P(c < Z \leq 1) - E^2 Z. \]

\[ E Z \geq E(Z|Z = Z_N) P(Z = Z_N) - D P(Z < 0) \]

\[ E(Z|Z = Z_N) \geq E(Z|Z = Z_N, Z_N \leq c) \geq E(Z_N|Z_N \leq c) \]

\[ \geq \int_{-\infty}^c z dF_{Z_N}(z) = E Z_N - \int_c^1 z dF_{Z_N}(z) \geq E Z_N - P(Z_N \geq c). \]
(2.20) \( \text{Var } Z \leq \text{Var } Z_N + D^2 P(Z < 0) + P(c < Z \leq 1) + 2P(Z_N \geq c) \), since \( E Z_N < 1 \).

Lemma 2.4: \[ P(c < Z \leq 1) \leq P(Z_N \geq c) \leq O(\lambda^{\delta-3}) \]

Proof: The first inequality follows from the fact that

\[ Z > c \text{ implies } Z_N > c \text{ or } Z = 1. \]

\[
P(Z_N > c) = P(m \{0, c\} \leq N-1)
= \sum_{i=0}^{N-1} (c^2 \lambda/(2\xi))^i \exp \left[-(c^2 \lambda/(2\xi))/i! \right] + R
= \lambda^{-3} \sum_{i=0}^{N-1} (3 \log \lambda)^i/i = O(\lambda^{\delta-3}), \text{ for any } \delta > 0.
\]

We note that \( \xi > c \) implies the computation of \( P \{m \{0, c\} \leq N-1\} \) involves only the triangular portion of the intensity function. Note that \( c^2 \lambda/(2\xi) \) is the mean number of observations in \([0, c]\). \( R \) decreases exponentially in \( \lambda \). Hence by (2.20), Lemma 2.4 and letting \( N \geq 7 \) in Lemma 2.3 we have

\[ \text{Var } Z \leq \text{Var } Z_N + O(\lambda^{\delta-3}), \text{ for any } \delta > 0, \]

and the theorem follows.

In applications it may be impractical to take \( \lambda \) so large that (2.8) holds, if \( \xi \) is very small. It should be noted that by adjusting \( \lambda \) other values of \( c \) and \( N \) can be found so that the result of Theorem 2.2 holds. We let \( c = (K \log \lambda)/\lambda, N \geq 4 \) and \( K = 6 \) (except for \( K \) these are the conditions of Theorem 2.1), and adjust \( \lambda \) so that \( c > \xi \). Following the proof of Lemma 2.3 we obtain
\[ P(Z < 0) = P(X_u > 1) = O(\lambda^{8-3}). \]

Lemma 2.4 holds since, if \( c \geq \zeta \),

\[
P(m \{0, c\} \leq N-1) = \sum_{i=0}^{N-1} \left[ \lambda(c - \zeta/2) \right]^i \exp \left[ -\lambda(c - \zeta/2) \right] / i! + R
\]

\[
\leq \sum_{i=0}^{N-1} \left( \lambda c/2 \right)^i \exp \left[ -\lambda c/2 \right] / i! + R
\]

\[
= \exp \left[ -3 \log \lambda \right] \sum_{i=0}^{N-1} \left( 3 \log \lambda \right)^i / i! + R
\]

\[
= O(\lambda^{8-3}) \text{ for any } \delta > 0.
\]

By (2.20) and the proof of the lemmas under the new condition, Theorem 2.2 again follows.

Note that for \( \zeta = O(1) \) all but the \( \lambda^{-2} \) term in the expression for \( \text{Var}(Z_N) \), (2.17), decreases exponentially in \( \lambda \), hence \( \text{Var} Z_N \leq O(\lambda^{-2}) \), and we obtain superefficient estimates in the trapezoidal case, as in the uniform case.

We would like to compare the estimate proposed in this paper with some estimate with known optimal properties, e.g., the maximum likelihood estimate. Even if we assume \( \mu = 0 \), and the range of the distribution and \( \zeta \) known, we encounter difficulties in computing the variance of the maximum likelihood estimate. The maximum likelihood estimate for a given sample could be determined by numerical methods. First determine the region of the maximum graphically and then use the gradient method of maximization to converge on the maximum [9]. Monte Carlo methods could be used to generate observations, and the variance could be estimated from a large number of such samples. A few
hypothetical examples indicated that even for $\zeta \leq 1/2$ the mid-range is close to the maximum likelihood estimate.

It should be noted for the case $\mu = 0$ the asymptotic variance of the mid-range is given by $\text{Var} \frac{Z_n}{2}$. We now derive the variance of the sample mean, $\bar{Y}_m$, of $m$ independent observations which are distributed,

$$f_y(y) = \frac{1}{2 - \zeta}, \quad \text{for } 0 \leq |y| \leq 1 - \zeta$$

$$= \frac{1 - |y|}{(2 - \zeta)}, \quad \text{for } 1 - \zeta \leq |y| \leq 1.$$  

![Density of Y](image)

**Figure 2.2**

Since $f_Y(y) = 0$ for all $y \notin [-1, 1]$ and $f(y)$ is symmetric about 0,

$$\text{Var } Y = EY^2 = \frac{2}{2 - \zeta} \left[ \int_0^{1-\zeta} y^2 \, dy + \zeta^{-1} \int_{1-\zeta}^1 y^2(1-y) \, dy \right]$$

(2.21) \hspace{1cm} \frac{4-6\zeta + 4\zeta^2 - \zeta^3}{6(2 - \zeta)}.$$
\[(2.22) \quad \text{Var } \bar{Y}_m = \text{Var } Y/m.\]

We note that for the Poisson case with \( \mu = 0 \), the variance of the sample mean, \( \text{Var } \bar{Y}^* \), is obtained by considering \( m \) in (2.22) a Poisson random variable in the sense that

\[(2.23) \quad P(m) = \frac{e^{-\lambda(2-\xi)}[\lambda(2-\xi)]^m}{[1-e^{-\lambda(2-\xi)}]m!}, \quad m = 1, 2 \ldots \]

\[= 0, \quad \text{otherwise}.\]

Hence the variance of the sample mean, when the size of the sample itself is a random variable with distribution (2.23), is given by

\[(2.24) \quad \text{Var } \bar{Y} = \sum_{m=1}^{\infty} P(m) \text{ Var } \bar{Y}_m = \text{Var } Y \sum_{m=1}^{\infty} P(m)/m \text{ by (2.22)}.\]

\[(2.25) \quad \sum_{m=1}^{\infty} P(m)/m = \sum_{m=1}^{\infty} \frac{\exp[-\lambda(2-\xi)][\lambda(2-\xi)]^m}{(1 - \exp[-\lambda(2-\xi)])m!} (m)\]

\[(2.26) \quad = \sum_{m=1}^{\infty} \frac{\exp[-\lambda(2-\xi)][\lambda(2-\xi)]^m [1 + 1/m]}{(1 - \exp[-\lambda(2-\xi)]) (m + 1)!},\]

by multiplying the numerator and denominator in (2.25) by \((m + 1)/m\),

\[(2.27) \quad = [\lambda(2-\xi)]^{-1}[1 - P(1)] + \sum_{m=1}^{\infty} P(m)/[m(m + 1)]\]

\[(2.28) \quad = [\lambda(2-\xi)]^{-1}[1 - P(1)] + [\lambda(2-\xi)]^{-2}[1 - P(1) - P(2)]
+ 2[\lambda(2-\xi)]^{-3}[1 - \sum_{m=1}^{3} P(m)]
+ \ldots + n[\lambda(2-\xi)]^{-(n+1)}[1 - \sum_{m=1}^{n+1} P(m)]\]
+ . . . , by repeating the same process in (2.26) and (2.27).

Hence by (2.21), (2.24), and (2.28)

\[ \text{Var } \bar{Y}^* = \frac{(4 - 6\xi + 4\xi^2 - \xi^3)}{[6(2 - \xi)^2][1/\lambda + o(\lambda^{-2})]}, \]

and we observe that the estimation procedure, \( \psi^* \), of Theorem 2.2 compares favorably by a whole order in \( \lambda \) with the sample mean.
SECTION 3

TWO-DIMENSIONAL CASE

3.1 Uniform Intensity On S.

Define the center, $\psi$, of any closed bounded strictly convex set, $S$, in the plane in the following way: move $m$ lines, each one perpendicular to one of the angles $\theta_1, \theta_2, \ldots, \theta_m$, from "infinity" until they each first intersect $S$. Call these intersection points $(x_1, y_1)$, $(x_2, y_2)$, $\ldots$, $(x_m, y_m)$. These points are unique and are unchanged by the translation of $S$.

Let $\psi = \left( \frac{1}{m} \sum_{i=1}^{m} x_i, \frac{1}{m} \sum_{i=1}^{m} y_i \right)$.

Note, that if $(x_i, y_i)$, $i = 1, 2, \ldots, m$ are unique for a given $\Theta = (\theta_1, \theta_2, \ldots, \theta_m)$, we may remove the requirement of strict convexity on $S$. We wish to calculate the mean square error of a procedure to estimate $\psi$, so we may, without loss of generality, assume the center of $S$ is located at the origin.

We now wish to consider the special case where $S$ is a circle of radius one. If the angles $\theta_1, \theta_2, \ldots, \theta_m$ are symmetric with respect to the origin, in the sense that $\theta \in \Theta$ if and only if $\theta + \pi \in \Theta$, then $\psi = (0, 0)$. Let $n = m/4$ be a positive integer and

$$\Theta = (\alpha_1, \alpha_1 + \pi/2, \alpha_1 + \pi, \alpha_1 + 3\pi/2, \alpha_2, \ldots, \alpha_j, \alpha_j + \pi/2, \alpha_j + \pi, \alpha_j + 3\pi/2, \ldots, \alpha_n + 3\pi/2),$$

where $0 \leq \alpha_1 < \alpha_2 < \ldots < \alpha_n < \pi/2$. For each $j$, $j = 1, 2, \ldots, n,$
i.e., for each quadruple of right angles, we consider the coordinate system with positive x-axis directed along $\alpha_j$. Let $(\hat{x}_{j_1}, \hat{x}_{j_2}, \hat{y}_{j_3}, \hat{y}_{j_4})$, \ldots, $(\hat{x}_{j_1}, \hat{x}_{j_2}, \hat{y}_{j_3}, \hat{y}_{j_4})$ be estimates in the \( j \)th coordinate system of $(x_{4j-3}, y_{4j-3})$, \ldots, $(x_{4j}, y_{4j})$, \( j = 1, 2, \ldots, n \). Estimate $\Psi$ by

\[
\hat{\Psi} = \left( \frac{\sum_{j=1}^{n} \hat{\xi}_j}{n}, \frac{\sum_{j=1}^{n} \hat{\eta}_j}{n} \right),
\]

where

\[
\hat{\xi}_j = (1/2)[(\hat{x}_{j_1} + \hat{x}_{j_2})\cos \alpha_j + (\hat{y}_{j_3} + \hat{y}_{j_4})\sin \alpha_j]
\]

\[
\hat{\eta}_j = (1/2)[(\hat{y}_{j_2} + \hat{y}_{j_4})\cos \alpha_j - (\hat{x}_{j_1} + \hat{x}_{j_3})\sin \alpha_j].
\]

We have the following theorem concerning the mean square error of the above stated estimation procedure.

Theorem 3.1: Suppose points are distributed according to the Poisson distribution with parameters per unit area as follows:

i) $\lambda$ in $S$, a disc of radius one, centered at the origin,

ii) $\mu$ in $T$, a bounded region containing $S$,

iii) 0, elsewhere.

Assume $V = (\hat{x}_{1_1}, -\hat{x}_{1_3}, \hat{y}_{1_2}, -\hat{y}_{1_4}, \ldots, \hat{x}_{j_1}, -\hat{x}_{j_3}, \hat{y}_{j_2}, -\hat{y}_{j_4}, \ldots, \hat{x}_{n_1}, -\hat{x}_{n_3}, \hat{y}_{n_2}, -\hat{y}_{n_4}) = (v_1, v_2, \ldots, v_m)$ has elements which are: bounded functions of the observations, asymptotically ($\mu$ and $n$ fixed, $\lambda \to \infty$) uncorrelated and identically distributed.

\footnote{The author is indebted to Professor Herman Chernoff for the ideas which lead to this estimation procedure.}
Letting $\sigma^2$ be the common variance of the $x_{ij}$, $i = 1, 2, \ldots, m$, enumerated above, we have the asymptotic mean square error,

$$\text{m.s.e. } \psi = \sigma^2/n.$$ 

Proof: Let $X_j = (\hat{x}_{j1} + \hat{x}_{j2})/2$ and $Y_j = (\hat{y}_{j1} + \hat{y}_{j4})/2$, $j = 1, 2, \ldots, n$.

$$\text{EX}_j = \text{EY}_j = 0.$$ 

Since boundedness implies the existence of the expectations, this equation follows from the symmetry assumed.

$$\text{m.s.e. } \hat{\psi} = \text{E|\hat{\psi} - \psi|^2} = \text{E|\hat{\psi}|^2},$$

since for the present case $\psi = (0,0)$.

$$\text{m.s.e. } \hat{\psi} = [1/n^2]\text{E}(\sum x_i \cos \alpha_i + \sum y_i \sin \alpha_i)^2$$
$$+ \text{E}(\sum y_i \cos \alpha_i - \sum x_i \sin \alpha_i)^2]$$
$$= (1/n^2)\sum_{i=1}^{n} (\text{EX}_i^2 + \text{EY}_i^2),$$

since $\hat{\psi}$ is obtained from the $X$'s and $Y$'s by an orthogonal transformation. We now have

$$n^2(\text{m.s.e. } \hat{\psi}) = [n/4][\text{E}(x_{11} + x_{12})^2 + \text{E}(y_{12} + y_{14})^2], \text{ by symmetry},$$
$$= [n/4][\text{Var}(x_{11} + x_{12}) + \text{Var}(y_{12} + y_{14})]$$
$$= n\sigma^2.$$ 

We now have the asymptotic result,

$$\text{m.s.e. } \hat{\psi} = \sigma^2/n,$$

which completes the proof the Theorem 3.1.
From our observation points we define the random variable $X$ in the following way: we move a line perpendicular to the $x$-axis to the right, from the most negative point in the $\mu$-region. At each observation point in succession we observe the number of points in the rectangle, $R(x', y') = \{ (x, y) : x' - c \leq x \leq x', y' - 2^\frac{3}{2} c^{1/2} \leq y \leq y' + 2^\frac{3}{2} c^{1/2} \}$, where $(x', y')$ represents the coordinates of some observation point and $c$ is a positive constant to be determined. We define $X$ as the $x$-coordinate of the first observation point having $N$ or more points in $R$. If no point satisfies this condition we let $X = 1$. Without loss of generality, for the discussion of the variance of $X$, we may center the $\lambda$-region at $(1, 0)$, and then the edge will cross the $x$-axis at the origin.

![Diagram showing $R(x', y')$ and $\lambda$-region]

**Figure 3.1**

1 The region within the large rectangle, not including the $\lambda$-region, represents the $\mu$-region.
Theorem 3.2: If the \( \mu \)-region is contained in 
\[ ((x,y): x \geq -T, 0 < y < \infty), T > 0, \mu \text{ is a constant}, \quad 3 \leq N \leq \lambda, \] 
and
\[ c = K(\log \lambda / \lambda)^{2/3}, \quad K \geq 3^{2/3}/2 = 1.04, \]
we have
\[ \text{Var } X \leq \frac{1}{4} \left( \frac{3}{2} \right)^{4/3} \frac{\Gamma(N+4/3)}{\Gamma(N)} \left[ \frac{\Gamma(N+4/3)}{\Gamma(N)} - \left( \frac{\Gamma(N+2/3)}{\Gamma(N)} \right)^2 \right]^{-4/3} \]
\[ + o(\lambda^{\delta-2}), \quad \text{for any } \delta > 0. \]

Proof:
\[ \text{Var } X = \mathbb{E} X^2 - \mathbb{E}^2 X. \]
\[ \mathbb{E} X \leq T^2 \mathbb{P}(X < 0) + \mathbb{E}(X^2 | 0 \leq X) \mathbb{P}(0 \leq X \leq c) \]
\[ + \mathbb{P}(c < X \leq 1). \]

We order the \( \lambda \)-region observations by their \( x \)-coordinates thus:
\[ X_1, X_2, \ldots, X_N, \ldots \]
and let the corresponding \( y \)-coordinates be
\[ Y_1, Y_2, \ldots, Y_N, \ldots. \]
Since \( 0 \leq X \leq X_N \), given that \( X \), the stopping value, lies in the interval \([0,c]\), we have
\[ \mathbb{E}(X^2 | 0 \leq X \leq c) \mathbb{P}(0 \leq X \leq c) \leq \mathbb{E} X_N^2; \]
hence we have
\[ \text{Var } X \leq \mathbb{E} X_N^2 + T^2 \mathbb{P}(X < 0) + \mathbb{P}(c < X \leq 1) - \mathbb{E}^2 X. \]

(3.3) \[ \mathbb{E} X \geq \mathbb{E}(X | X = X_N) \mathbb{P}(X = X_N) + \mathbb{E}(X | X \geq 0, X \neq X_N) \mathbb{P}(X \geq 0, X \neq X_N) \]
\[ - \mathbb{P}(X < 0) \]

(3.4) \[ \geq \mathbb{E}(X | X = X_N) \mathbb{P}(X = X_N) - \mathbb{P}(X < 0). \]
Now,

\[(3.5) \quad E(X|X = X_N) \geq E(X_N|X_N \leq c),\]

since \(X = X_N\) implies \(X_N \leq c\) or \(X_N > c\) and the number of points in the rectangle at \(X_N\) exceeds \(N\).

\[
E(X_N|X_N \leq c) \geq \int_{-\infty}^{c} x \, dF_{X_N}(x) = E \, N - \int_{c}^{1} x \, dF_{X_N}(x)
\]

\[
\geq E \, N - P(X_N \geq c).
\]

\[(3.6) \quad \text{Var } X \leq E \, N^2 + T^2 \, P(X < 0) + P(c \leq X \leq 1)
\]

\[
- (P(X = X_N)[E(X_N) - P(X_N \geq c)] - TP(X < 0))^2.
\]

We now compute the asymptotic mean and variance of \(X_N^\circ\). Let the random variable \(A\) be the area of the region \(H(X_N) = \{(x, y): x \leq X_N, -\infty < y < \infty\} \cap S\), i.e., the segment of the \(\lambda\)-region to the left of \(X_N\).

\[\text{Diagram Showing The Disc, } S, \text{ And A Subset, } H(X_N)\]

\[\text{Figure 3.2}\]

28
Let \( Z = 1 - X_N \) and \( \omega^2 = 1 - (1 - X_N)^2 = 1 - Z^2 \).

\[
A = \sin^{-1} \sqrt{1 - Z^2} - Z \sqrt{1 - Z^2}
\]

\[
= \omega + \omega^3/6 + 3\omega^5/40 + \ldots - \omega(1 - \omega^2/2 - \omega^4/8 - \ldots)
\]

\[
= 2\omega^3/3 - 0(\omega^5).
\]

\[(3.7) \quad A \left[ \frac{2^{5/2}}{3} X_N^{3/2} \right] = 1 + O(X_N).
\]

\[
A^{2/3} \left[ \left( \frac{2^{5/3}}{3} \right)^{3/2} X_N \right] = 1 + O(X_N),
\]

by taking the \( 2/3 \) root in (3.7).

\[
X_N = (3A/2)^{2/3}/2[1 + O(X_N)].
\]

\[
X_N = (3A/2)^{2/3}/2[1 + O(\omega^2)].
\]

We note that \( A \) has the gamma distribution with parameters \( \lambda \) and \( N \), i.e.,

\[
\Gamma_A(A) = \lambda(\lambda A)^{N-1} e^{-\lambda A} / \Gamma(N), \quad A > 0,
\]

\[
= 0, \quad \text{otherwise}.
\]

\[(3.8) \quad E X_N = (3/2)^{2/3}(1/2) E A^{2/3} + O(\omega^4)
\]

\[
= (3/2)^{2/3}(1/2) \left[ \Gamma(N + 2/3)/\Gamma(N) \right] \lambda^{-2/3} + O(\lambda^{-4/3})
\]

\[(3.9) \quad E X_N^2 = (1/4)(3/2)^{4/3} E A^{4/3} + O(\omega^6)
\]

\[
= (1/4)(3/2)^{4/3} \left[ \Gamma(N + 4/3)/\Gamma(N) \right] \lambda^{-4/3} + O(\lambda^{-2})
\]

29
To continue the proof of the theorem we use the following lemmas.

Lemma 3.1: If \( N \geq 3 \), and \( c = K(\log \lambda/\lambda)^{2/3}, K \) a constant, then
\[
P(X < 0) = O(\delta^{5/2})
\] for any \( \delta > 0 \).

Proof: Let \( D \) represent the area of the \( \mu \)-region, \( U \). The probability that there are exactly \( m \) \( \mu \)-observations in \( U \) is
\[
e^{-\mu D} (\mu D)^m / m! .
\]

Given that there are \( m \) \( \mu \)-observations, the probability that any given one has \( N-1 \) or more other \( \mu \)-observations in the rectangle \( R \), defined above, adjoining it, is
\[
\leq \sum_{n=N-1}^{m-1} \binom{m-1}{n} (2^{5/2} c^{3/2} / D)^n \left[ 1 - (2^{5/2} c^{3/2} / D) \right]^{m-1-n} .
\]

We note that \( D \) is a constant and for sufficiently large \( \lambda \) this expression is less than 1.

The probability that this event occurs for any of the \( m \) \( \mu \)-observations must be less than \( m \) times this quantity. Let \( P_U = P( \text{There exists some } R \subset U \text{ containing } N-1 \text{ or more } \mu \)-observations. )

Hence we have,
\[
P_U \leq \sum_{m=N}^{\infty} \sum_{n=N-1}^{m-1} \binom{m-1}{n} (2^{5/2} c^{3/2} / D)^n \left[ 1 - (2^{5/2} c^{3/2} / D) \right]^{m-1-n} .
\]

\[
\leq \sum_{n=N-1}^{\infty} \left[ (2^{5/2} c^{3/2} / D)^n / n! \right] \sum_{m=1}^{\infty} e^{-\mu D} (\mu D)^m m(m-1) \ldots (m-n)/m! .
\]

\[
= \sum_{n=N-1}^{\infty} M_{n+1} (2^{5/2} c^{3/2})^n / n! ,
\]

30
where $M_{n+1}$ is the $n+1$ factorial moment of the Poisson distribution with parameter $\mu D$. Hence we have

$$P_U \leq \sum_{n=N-1}^{\infty} \left(\frac{\lambda}{2} \right)^{3/2} \frac{e^{-\lambda}}{n!} \frac{D^n}{D^n},$$

and letting $U = \{(x, y): x < 0, -\infty < y < \infty\}$, $c = K(\log \lambda/\lambda)^{2/3}$, and $N \geq 3$ we have,

$$P(X < 0) = O(\lambda^{\delta-2}), \text{ for any } \delta > 0.$$

**Lemma 3.2**: If $N > 0$ (fixed), $c = K(\log \lambda/\lambda)^{2/3}$, $K \geq \frac{2^3}{3^2}$ a constant, then $P(c < X \leq 1) = O(\lambda^{\delta-2})$ for any $\delta > 0$.

**Proof**: If there are $N$ or more $\lambda$-observations less than or equal $c$ in $x$-coordinate, then $X \leq c$; hence

$$P(c < X \leq 1) \leq P(c < X) \leq$$

$$P(\text{there are fewer than } N \lambda\text{-observations} \leq c)$$

$$= P[n(c) < N] = e^{-\lambda} A_c \sum_{i=0}^{N-1} (\lambda A_c)^i / i!,$$

where $A_c$ is the area of the $\lambda$-region to the left of the line $x = c$.

By the series expansion used earlier we have

$$A_c = \left(\frac{2}{3}\right)(2c-\frac{c^2}{2})^{3/2} - |0(\lambda^{5/3})|,$$

$$A_c = \left(\frac{2}{3}\right)(2c)^{3/2} (1-c/2)^{3/2} - |0(\lambda^{5/3})|$$

$$= \left(\frac{2}{3}\right)(2K)^{3/2} \log \lambda |1 - |0(\lambda^{\delta-2}/3)|].$$

$$\lambda A_c = \log \lambda \left(\frac{2}{3}\right)(2K)^{3/2} \left[1 - |0(\lambda^{\delta-2}/3)|]\right].$$

Hence, since $K \geq \frac{2^3}{3^2}$, 2,
$$P(c < X \leq 1) \leq \lambda^{-2} \sum_{i=0}^{N-1} (\log \lambda^{K'}) \frac{1}{i!} = O(\lambda^{5-2})$$

where

$$K' = \frac{2}{3}(2K)^{3/2}.$$

**Corollary:** $P(X_N \geq c) = O(\lambda^{5-2})$, if $N > 0$ (fixed), $c = K(\log \lambda/\lambda)^{2/3}$, $K \geq 3^{2/3}/2 \approx 1.04$.

**Proof:** Since $X_N$ has a continuous distribution function and by the argument of Lemma 3.2 we have

$$P(X_N \geq c) = P[n(c) < N] = O(\lambda^{5-2})$$

for any $\delta > 0$, which completes the proof of the corollary.

Returning to the proof of the theorem, now consider the expression (3.6). We dispose of most of the terms using Lemma 3.1 and Lemma 3.2.

Lemma 3.1 proves that the $T^2P(X < 0)$ and $TP(X < 0)$ terms are $O(\lambda^{5-2})$ for any $\delta > 0$. Lemma 3.2 and its corollary prove that the $P(c \leq X \leq 1)$ and $P(X_N \geq c)$ terms, respectively, are $O(\lambda^{5-2})$. We now have

$$\text{Var } X \leq E X_N^2 - [P(X = X_N)]^2 E^2(X_N) + O(\lambda^{5-2}) \quad (3.10)$$

$$= E X_N^2 - [1 - P(X \neq X_N)]^2 E^2(X_N) + O(\lambda^{5-2}) \quad (3.11)$$

$$= \text{Var } X_N + E^2(X_N)0[(P(X \neq X_N)] + O(\lambda^{5-2}) \quad (3.12)$$

We need the following lemma to dispose of the second term in (3.12).

**Lemma 3.3:**

$$E^2X_N(P(X \neq X_N)) = O(\lambda^{8-7/3}).$$

32
Proof:

\[ E^2X_N = O(\lambda^{-4/3}) \] by (3.9)

\[ P(X \neq X_N) = P(X < 0) + P(X = X_i, i = 1 \text{ or } 2 \text{ or } \ldots \text{ or } N-1) \]
\[ + P(X = X_i, i = N+1 \text{ or } N+2 \text{ or } \ldots ) \] .

By Lemma 3.1, \( P(X < 0) = O(\lambda^{5/2}) \). The event \( X = X_i, i = N + 1 \text{ or } N + 2 \text{ or } \ldots \), implies \( X_N > c \), hence

\[ P(X = X_i, i = N+1 \text{ or } N+2 \text{ or } \ldots ) \leq P(X_N > c) = O(\lambda^{5/2}) \]

by the corollary to Lemma 3.2. For \( X \) to be \( X_i, i = 1 \text{ or } 2 \text{ or } \ldots \text{ or } N-1 \), the rectangle at \( X_i \) must contain at least one \( \mu \)-observation. Following an argument similar to Lemma 3.1 we obtain

\[ P'_U = P(R(X_1) \text{ or } R(X_2) \text{ or } \ldots \text{ or } R(X_{N-1}) \text{ contains at least one } \mu \text{-observation} \]
\[ \leq (N - 1) P(R(0, 0) \text{ contains at least one observation}) \]

This inequality follows from the fact that the probability that the rectangle at any \( X_i \) should contain \( \mu \)-observations is less that that for the rectangle at the origin, and the probability of the union of events does not exceed the sum of the probabilities.

\[ P'_U \leq (N - 1)[1 - P(R(0, 0) \text{ contains no observations})] \]
\[ = (N - 1)[1 - \exp(-2^{5/2} c^{3/2})] \]
\[ = (N - 1)[1 - (1 - 2^{5/2} c^{3/2} + O(c^3))] \]
\[ = O(\lambda^{5-1}) \text{ for any } \delta > 0 \text{ which } \]

33
completes the proof of Lemma 3.3. Returning to the proof of the theorem, we note that by Lemma 3.3 and (3.11) we have

\begin{equation}
(3.15) \quad \text{Var } X \leq \text{Var } X_N + O(\lambda^{8-2}) \text{ for any } \delta > 0
\end{equation}

\begin{equation}
\leq (1/4)(3/2)^{4/3} \left[ \Gamma(N + 4/3)/\Gamma(N) - [\Gamma(N + 2/3)/\Gamma(N)]^2 \right]^{4/3}
\end{equation}

\begin{equation}
+ O(\lambda^{8-2}), \text{ for any } \delta > 0, \text{ by (3.8), (3.9) and (3.15)},
\end{equation}

which completes the proof of Theorem 3.2.

To complete the proof that the asymptotic m.s.e. of \( \psi \) is \( \sigma^2/n \), when we use the stopping rule using the rectangle \( R \) we need the lemma below.

**Lemma 3.4**: Let the elements \( V = (v_1, v_2, \ldots, v_m) \) be the \( m \) random variables defined as \( X \) in Theorem 3.2 after successive rotations of the axes of \( \theta_1, \theta_2, \ldots, \theta_m \), respectively; then \( v_1, v_2, \ldots, v_m \) are bounded, asymptotically (\( \mu \) and \( m \) fixed, \( \lambda \to \infty \)) uncorrelated and identically distributed random variables.

**Proof**: We show zero asymptotic correlation by showing asymptotic pairwise independence. As \( \lambda \) grows large, the size of the rectangle, \( c \) by \( 2^{5/2} c^{1/2} \), \( c = (3^{2/3}/2)(\log \lambda/\lambda)^{2/3} \), grows small. Any two distinct rays \( \theta \) and \( \theta' \) are separated by some angle \( \varphi = |\theta - \theta'| > 0 \). With probability \( 1 - O(\lambda^{8-2}) \), the elements of \( V \) all lie within a distance of \( c \) of the edge of the disc (within the \( \lambda \)-region) by Lemma 3.2. For any \( \varphi > 0 \) there must exist \( \lambda \) large enough so that the rectangles tangent to the disc at \( \theta \) and \( \theta' \), respectively, do not overlap; hence \( v \) and

34
v' are asymptotically independent.

The distribution of the v's, given the v's lie within the disc S, are identically distributed by symmetry. The probability that any lie outside S is \(O(\lambda^{1/2})\) by Lemma 3.1; hence they are asymptotically identically distributed. This completes the proof of the Lemma.

It should be observed that the procedures of Theorems 3.1 and 3.2 provide, when \(N = 3\) and \(\lambda\) is large,

\[
(3.16) \quad \text{m.s.e. } \hat{\psi} = \left[1/(4n)\right] \lambda^{-1/3},
\]

where \(n\) is the number of quadruples ("frames") used. This procedure is superefficient in the sense that m.s.e. \(\hat{\psi} = o(\lambda^{-1})\). For comparison we obtain the m.s.e. using the centroid estimate,

\[
(3.17) \quad \psi_m' = \frac{\bar{w}}{w} = \sum \frac{w_i}{m}, \quad \bar{z} = \sum \frac{z_i}{m},
\]

where \((w_i, z_i), i = 1, \ldots, m\) are the coordinates of the observations, and \(m\) is a random variable with the truncated Poisson distribution. We assume the intensity outside the disc, \(\mu = 0\); hence \(m\) is distributed with Poisson parameter, \(\lambda n\). The m.s.e. of the centroid estimate,

\[
\text{m.s.e. } \psi' = \sum_{m=1}^{\infty} P(M) E_m |\psi_m'|^2,
\]

\[
E_m |\psi_m'|^2 = E_m (\bar{w}^2 + \bar{z}^2) = E_m (\bar{w}^2) + E_m (\bar{z}^2).
\]

The \(w_1, w_2, \ldots, w_m\) are independent and identically distributed with mean 0,

\[
E w^2 = E w_1^2 / m,
\]
and likewise
\[ E_z^2 = E_{z_1}^2 / m; \]
hence
\[ (3.18) \]
\[ E_m |\psi_m'|^2 = (1/m) E (w_1^2 + z_1^2). \]

Let
\[ \rho = \sqrt{w_1^2 + z_1^2}. \]

Since \((w_1, z_1)\) is uniformly distributed on
\[ \{(w, z) : \sqrt{w^2 + z^2} \leq 1\}, \]
the density of \(\rho\) is given by
\[ f(\rho) = 2\rho, \text{ for } 0 \leq \rho \leq 1 \]
\[ = 0, \text{ otherwise.} \]

\[ E_m |\psi_m'|^2 = (1/m) E \rho^2 = 1/(2m) \]
\[ P (m) = (1-e^{-\lambda \pi})^{-1} e^{-\lambda \pi (\lambda \pi)^m/m!}; \text{ for } m = 1, 2, \ldots \]
\[ = 0, \text{ otherwise.} \]

m.s.e. \(\psi' = [2(1-e^{-\lambda \pi})]^{-1} \sum_{m=1}^{\infty} e^{-\lambda \pi (\lambda \pi)^m/(m\cdot m!)} \]
\[ = (2\pi \lambda)^{-1} e^{-\lambda \pi} \sum_{m=1}^{\infty} [(\lambda \pi)^{m+1}/(m+1)!] (m+1)/m + o(\theta) \]
where \(\theta\) decreases exponentially in \(\lambda,\)
\[ = (2\pi \lambda)^{-1} \left[ \sum_{m=2}^{\infty} \exp [-\lambda \pi] (\lambda \pi)^m/m! \right] + \sum_{m=1}^{\infty} \exp [-\lambda \pi] (\lambda \pi)^{m+1}/ [m\cdot (m+1)!] + o(\theta) \]

36
\begin{equation}
(2\pi\lambda)^{-1} [1 + O(\text{m.s.e. } \psi')]
\end{equation}

(3.19) \quad = (2\pi\lambda)^{-1} + O(\lambda^{-2}).

It should be noted that the relative efficiency in terms of reduction in m.s.e. of the procedures of Theorems 3.1 and 3.2 to the centroid estimate is \(O(\lambda^{1/3})\). For the case \(\mu = 0\), and letting \(N = 3\) in the "frame" procedure, by the first term approximations of (3.16) and (3.19) we obtain the number of "frames" \(n\) to be competitive with the centroid estimate, i.e., \(n\) such that

\begin{equation}
\left[\frac{1}{(4n)}\right]^{\lambda^{-4/3}} < (2\pi\lambda)^{-1}
\end{equation}

(3.20) \quad n > (\pi/2)^{\lambda^{-1/3}}.

Hence for all \(\lambda > \frac{1}{8} \pi^3 \approx 3.88\), one "frame", and hence any number of "frames", is better than the centroid estimate.

Clearly if \(\mu > 0\) the m.s.e. of the centroid estimate is

\[ \text{m.s.e. } \psi' \geq (2\pi\lambda)^{-1}; \]

then (3.20) is sufficient for the superiority of the "frame" procedure.
3.2 Trapezoidal Intensity on $S$.

We have designated the case discussed in this section, "trapezoidal", since any plane perpendicular to $T$ (see Figure 1.1), passing through $S$ intersects the intensity function in a trapezoid or a triangle, i.e., this case is the two-dimensional analog to the one-dimensional trapezoidal case discussed above. Let $(\rho, \theta)$, represent the polar coordinates. The Poisson parameter (intensity function) per unit area is:

$$z(\rho, \theta) = \begin{cases} 
\lambda, & \text{for } \rho \leq 1 - \xi, \ 0 \leq \theta \leq 2\pi \\
\lambda - (\rho - 1 + \xi)(\lambda - \mu)/\xi, & \text{for } 1 - \xi \leq \rho \leq 1, \ 0 \leq \theta \leq 2\pi \\
\mu, & \text{for } \rho > 1, \ (\rho, \theta) \in T \\
0, & \text{otherwise}.
\end{cases}$$

(3.21)

![Diagram of the trapezoidal intensity function](image-url)

Figure 3.3
"Trapezoidal" Intensity Function
In the sequel we investigate the m.s.e. of $\hat{\psi}$ and $\psi'$ for the intensity function (3.21). In order to determine bounds on m.s.e. $\hat{\psi}$ we may follow most of the argument used for the uniform case. Note in Theorem 3.2 that the expression (3.6)

\begin{equation}
\text{Var} \ X \leq E X_N^2 + \sum T \ P(X < 0) + P(c \leq x \leq 1) \\
- \left\{ P(X = X_N) [E(X_N) - P(X_N \geq c) - T(X < 0)]^2 \right\}
\end{equation}

was derived without reference to the distribution functions of $X$ and $X_N$, respectively; hence we may evaluate (3.6) for the present case, (3.21). To this end, we first determine the distribution function of $X_N$, and bound $\text{Var} \ X_N$.

Let $\Delta$ be the right circular cone of height $\omega = \lambda/\xi$, truncated at $z = \lambda$ and base centered at $(1, 0, 0)$, with radius one (a translation of one unit along the $x$-axis in Figure 3.3). Let $x_0 \leq 1$, and

$\omega v(x_0) = \text{volume of } \Delta \cap \{(x, y, z): x < x_0\}$, i.e., the volume bounded by the planes $z = 0$, $z = \lambda$, $x = x_0$, and the right circular cone with base $(x - 1)^2 + y^2 = 1$, and height $\lambda/\xi$. The c.d.f. of $X_N$ (again we truncate the distribution at 1) is given by

\begin{equation}
F_{X_N}(x) = 1 - \sum_{i=0}^{N-1} \exp \left[ -\omega v(x) \right] [\omega v(x)]^i / i!, \text{ for } 0 \leq x < 1 \\
= 1, \text{ for } x \geq 1 \text{ and } \\
= 0, \text{ for } x < 0.
\end{equation}
In the continuous region of the probability distribution function, 
0 ≤ x < 1, we have the density,

\[ f_{X^{N}}(x) = 0, \text{ for } x \leq 0, \ x > 1 \]

\[ = \omega \exp[-\omega x]/(\lambda - 1)], \ \text{for} \ 0 \leq x < 1. \]

\[ P(X^{N} = 1) = 1 - F_{X^{N}}(1^{-}). \]

As we did in previous cases, we obtain (3.24) by considering the "density" of waiting time for a Poisson process with known mean value function [8].

We now evaluate \( \omega \theta(x_{0}) \) for \( x_{0} < \xi \). The cone \( \Delta \) can be represented by the equation,

\[ Z(x, y) = \omega(1 - [(x - 1)^{2} + y^{2}]^{1/2}), \ \omega = \lambda/\xi. \]

The path on the cone formed by the plane, \( x = \xi \), is

\[ Z = \omega(1 - [(\xi - 1)^{2} + y^{2}]^{1/2}). \]

The extreme values of \( y \) for \( x = \xi \) are the solutions in \( y \) of

\[ 1 = (\xi - 1)^{2} + y^{2} \]

\[ y = \pm (2\xi - \xi^{2})^{1/2}. \]

The area of the vertical section \( x = \xi \) through the cone

\[ A(\xi) = \omega \int_{-(2\xi - \xi^{2})^{1/2}}^{(2\xi - \xi^{2})^{1/2}} (1 - [y^{2} + (\xi - 1)^{2}]^{1/2}) \ dy \]

\[ = 2\omega \int_{0}^{(2\xi - \xi^{2})^{1/2}} (1 - [y^{2} + (\xi - 1)^{2}]) \ dy. \]
\[ \omega(x_0) = \int_0^{X_0} A(\xi) \, d\xi = 2\omega \int_0^{X_0} \left[ 1 - \left( y^2 + (\xi - 1)^2 \right)^{1/2} \right] \, dy \, d\xi \]

\[ = 2\omega \int_{\xi=0}^{X_0} (2\xi - \xi^2)^{1/2} \, d\xi - 2\omega \int_{\xi=0}^{X_0} \left\{ \int_0^{\sin^{-1}(2\xi - \xi^2)^{1/2}} \frac{1}{(1-\xi)^2 \sec^2 \phi} \, d\phi \right\} \]

setting \( y = (1 - \xi) \tan \varphi \),

\[ = 2\omega \int_{\xi=0}^{X_0} (2\xi - \xi^2)^{1/2} \, d\xi - 2\omega \int_{\xi=0}^{X_0} \frac{1}{(1-\xi)^2 \sec^3 \varphi} \, d\varphi \]

Now \[ \int \sec^3 \varphi \, d\varphi = \sin \varphi/(2 \cos^2 \varphi) + (1/2) \int \sec \varphi \, d\varphi = (1/2) \tan \varphi \sec \varphi \]

\[ + (1/4) \log [(1 + \sin \varphi)/(1 - \sin \varphi)]; \text{ hence} \]

\[ \omega(x_0) = 2\omega \int_{\xi=0}^{X_0} (2\xi - \xi^2)^{1/2} \, d\xi - \frac{\omega}{2} \int_{\xi=0}^{X_0} (1-\xi)^2 \left[ 2 \tan \varphi \sec \varphi \right. \]

\[ \left. + \log [(1 + \sin \varphi)/(1 - \sin \varphi)] \sin^{-1}(2\xi - \xi^2)^{1/2} \right|_{\varphi=0} \]

\[ = 2\omega \int_{\xi=0}^{X_0} (2\xi - \xi^2) \, d\xi - \frac{\omega}{2} \int_{\xi=0}^{X_0} (1-\xi)^2 \left( 2(2\xi - \xi^2)^{1/2}/(1-\xi)^2 \right. \]

\[ \left. + \log \left[ (1 + (2\xi - \xi^2)^{1/2})/[1 - (2\xi - \xi^2)^{1/2}] \right] \right) d\xi \]

\[ = \omega \int_0^{X_0} (2\xi - \xi^2)^{1/2} \, d\xi - \frac{\omega}{2} \int_0^{X_0} (2(1-\xi)^2 \sum_{j=0}^{\infty} \frac{(\sqrt{2\xi - \xi^2})^j}{(2j+1)!}) d\xi \]

41
\[= 2\omega \int_0^{x_0} \xi \sqrt{2\xi - \xi^2} \, d\xi - \omega \int_0^{x_0} \xi^2 \sqrt{2\xi - \xi^2} \, d\xi\]
\[-\omega \int_0^{x_0} \sum_{j=1}^{\infty} \left( \frac{\sqrt{2\xi - \xi^2}}{(2j + 1)} \right)^{2j+1} \, d\xi\]
\[= 2\omega \int_0^{x_0} \left[ (2^{1/2})^{3/2} + o(\xi^{5/2}) \right] \, d\xi + \omega o(x_0^{7/2})\]
\[-\omega \int_0^{x_0} \left[ (2^{3/2})^{3/2} + o(\xi^{5/2}) \right] \, d\xi\]
\[= \omega(2^{5/2}) \int_0^{x_0} \xi^{3/2} \, d\xi + \omega o(x_0^{7/2})\).

(3.26) \(\omega o(x_0) = \omega[(2^{7/2}/15) x_0^{5/2} + o(x_0^{7/2})] \). Conversely,

(3.27) \(x_0 = (15\omega)^{2/5} x_0^{7/5} + o(\omega^{4/5})\).

Returning to the derivation of \(\text{Var} X_N\), we find,

\[\text{Ex}_N^2 = \int_0^\xi x^2 \, dF_N(x) + \int_0^1 x^2 \, dF_N(x)\]
\[(3.28) \leq \int_0^\xi x^2 \, dF_N(x) + P(X_N > \xi),\]

We shall denote \(P(X_N > \xi)\) by \(R_0\).

\[\text{Ex}_N^2 \leq \int_0^\xi \{x^2 \omega^1(x)[\omega(x)]^{N-1} \exp[-\omega(x)]/(N-1)!\} \, dx + R_0,\]

by (3.24) and (3.28),
\[
\int_0^\nu(\xi) \frac{((15^{2/5}/2^{7/5})\xi^{2/5} + O(\xi^{4/5}))^2}{\omega^N \exp[-\omega \xi]/(N-1)!} dy + R_0,
\]
transforming by \( y = \nu(x) \), and using (3.27),
\[
(3.29) = \omega^{-4/5} (15^{4/5}/2^{4/5}) \Gamma(\nu(\xi);\omega, n+1/5) \frac{\Gamma(n+1/5)}{\Gamma(n)} + O(\omega^{-6/5}) + R_0
\]

\[
EX_N = \int_0^\xi x dF_{X_N}(x) + \int_\xi^1 x dF_{X_N}(x) \geq \int_0^\xi x dF_{X_N}(x)
\]

\[
= \int_0^\xi (x \omega v(x) [\omega v(x)]^{N-1} \exp[-\omega v(x)]/(N-1)! \) dx
\]

\[
= \int_0^\nu(\xi) \frac{((15^{2/5}/2^{7/5})\xi^{2/5} + O(\xi^{4/5}))^2}{\omega^N \exp[-\omega \xi]/(N-1)!} dy.
\]

Thus
\[
(3.30) EX_N \geq \omega^{-2/5} (15^{2/5}/2^{7/5}) \Gamma(\nu(\xi);\omega, n+2/5) \frac{\Gamma(n+2/5)}{\Gamma(n)} + O(\omega^{-4/5}),
\]
and from (3.30) and (3.29), we find:
\[
(3.31) \quad \text{Var} X_N \leq \omega^{-4/5} (15^{4/5}/2^{4/5}) \frac{\Gamma(\nu(\xi);\omega, n+4/5) \Gamma(n+4/5)}{\Gamma(n)}
\]

\[
- \Gamma^2(\nu(\xi);\omega, n+2/5) \frac{\Gamma^2(n+2/5)}{\Gamma^2(n)} + O(\omega^{-6/5}) + R_0,
\]

\[
= (\text{constant})(\xi/\lambda)^{4/5} + O[(\xi/\lambda)^{-6/5}] + R_0, \quad \text{since } \omega = \lambda/\xi \text{ by definition}.
\]

In Theorem 3.3 below we ignore the fact that letting \( \lambda \to \infty \) may be unrealistic.

Theorem 3.3: Given the intensity function (3.21) and the estimation procedure of Theorem 3.2, then there exists a \( c \) and an integer \( N \) such that
m.s.e. $\hat{\Psi} \leq \text{Var} X_N / 2 + \xi^{7/5} 0(\lambda^{6-7/5})$, for any $\xi > 0$.

Proof: Since $\xi$ is fixed and $\lambda$ is a parameter which may be varied, we may choose $\lambda$ so that

$$c = [(15^5 / 2) / \theta]^2 \leq \xi, \text{ i.e., so that,}$$

$$\log \lambda \lambda \leq (5/2)^2 \xi^{3/2} / 15.$$

We now proceed to evaluate the terms of (3.22). For the $P(X < 0)$ terms we follow the argument of Lemma 3.1.

$$P(X < 0) \leq \sum_{n=N-1}^{\infty} (2^{5/2} c^{3/2} / 2D)^n (\mu D)^{n+1} / n!$$

By (3.32), (3.33), and letting $N \geq 5$ we have

$$P(X < 0) = O(\lambda^{8-12/5}), \text{ for any } \xi > 0, \text{ since } c = O(\lambda^{8-3/5}).$$

We now dispose of the $P(c \leq X \leq 1)$ and $P(X_N > c)$ terms.

$$P(c \leq X \leq 1) \leq P(n(c) < N) = P(X_N > c).$$

$$P(X_N > c) = \sum_{i=0}^{N-1} \exp [-\omega(c)] \omega(c)^i / i !, \text{ by (3.23).}$$

Now

$$\omega(c) = (\lambda / \xi) [2^{7/2} / 15] c^{5/2} + O(c^{7/2}), \text{ by (3.26)}$$

$$= (\lambda / \xi) [2^{7/2} / 15] (15^5 / 25^2) (\log \lambda) / \lambda + O(\lambda^{6-7/5}),$$

substituting for $c$ from (3.32),

$$= (2 \log \lambda) + O(\lambda^{8-2/5}), \text{ for any } \xi > 0.$$

44
\[(3.36) \quad P(X_N > c) = \exp^{-2\log \lambda [1 + O(\delta^{-2/5})] \sum_{i=0}^{N-1} [2\log \lambda + O(\delta^{-2/5})] \frac{i}{i!}}, \]

substituting in (3.35),
\[
= \lambda^{-2}[1 + O(\delta^{-2/5})] \sum_{i=0}^{N-1} [2\log \lambda + O(\delta^{-2/5})] \frac{i}{i!} \\
= O(\delta^{-2}), \text{ for any } \delta > 0.
\]

Note also that \( R_0 = P(X_N > \xi) < P(X_N > c) \) by (3.32), and hence,
\[(3.37) \quad R_0 < O(\delta^{-2}) \text{ by (3.36)} .\]

\[(3.38) \quad \text{Var } X \leq \mathbb{E}_N^2 - [P(X=X_N)]^2 \mathbb{E}_N^2 X_N + O(\delta^{-2}) \quad \text{by (3.22), (3.34),}
(3.35), \text{ and (3.36)} ,
\]
\[
= \mathbb{E}_N^2 - [1 - P(X \neq X_N)]^2 \mathbb{E}_N^2 X_N + O(\delta^{-2}) \\
= \text{Var } X_N + \mathbb{E}_N^2 X_N 0[P(X \neq X_N)] + O(\delta^{-2}) .
\]

We now dispose of the term \( \mathbb{E}_N^2 X_N 0[P(X \neq X_N)] \) in (3.38). Since \( c \leq \xi \), and since we have just seen at (3.36) that \( P(X_N > c) = O(\delta^{-2}) \), the inequality (3.30) is actually an equality (to the order shown) and from (3.10) we may then write:
\[(3.39) \quad \mathbb{E}_N^2 X_N = O(\delta^{-4/5}) .\]

\[(3.40) \quad P(X \neq X_N) = P(X < 0) + P(X = X_i, \ i = 1 \text{ or } 2 \text{ or } \ldots \text{ or } N - 1)
+ P(X = X_i, \ i = N + 1 \text{ or } N + 2 \text{ or } \ldots \).
\]

\[(3.41) \quad P(X < 0) = O(\delta^{-12/5}) \text{ by (3.34)} . \]

\[(3.42) \quad P(X = X_i, \ i = N + 1 \text{ or } N + 2 \text{ or } \ldots ) \leq P(X_N > c) . \]

45
= o(\lambda^{6-2}) \text{ by (3.36)}.

Following the same argument found at the end of Lemma 3.3 we obtain

\[ P(X = X_i, \ i = 1 \text{ or } 2 \text{ or } \ldots \text{ or } N - 1) \]
\[ \leq (N - 1)[1 - P(R(0, 0) \text{ contains no observations})] \]
\[ = (N - 1)[1 - \exp(-2^{5/2} c^{3/2})] \]
\[ = (N - 1)[1 - [1 - 2^{5/2} c^{3/2} + o(c^3)]] \]
\[ = (N - 1)2^{5/2}[\log \lambda/\lambda]^{3/5} + o(\lambda^{6-6/5}) \]
\[ = 2(N - 1)(15\xi)^{3/5} o(\lambda^{6-3/5}). \]

The expressions (3.41-3.43) dispose of the three terms of (3.40), and show

\[ (3.44) \quad P(X \neq X_N) = \frac{3}{5} o(\lambda^{6-3/5}). \]

By (3.39) and (3.44),

\[ (3.45) \quad \mathbb{E}_{X_N}^2 [P(X \neq X_N)] = \frac{3}{5} o(\lambda^{6-3/5} \xi^{4/5}/\lambda^{4/5}), \]
\[ = \xi^{7/5} o(\lambda^{6-7/5}). \]

(3.46) \ Var X \leq \ Var X_N + \xi^{7/5} o(\lambda^{6-7/5}), \text{ by (3.38) and (3.45)},

which completes the proof of Theorem 3.3.

**Corollary:** If \( c < \xi \), then

\[ (3.47) \quad \text{m.s.e. } \hat{\psi} \leq o[(\xi/\lambda)^{4/5}] . \]

**Proof:** Theorem 3.3. shows that for suitable \( c \) and \( N \),
\[ \text{m.s.e. } \hat{\psi} \leq \text{Var } X_N^2 + \xi^{7/5} 0(\lambda^{-7/5}) \]
\[ \leq 0[(\xi/\lambda)^{4/5}] + R_0, \text{ by (3.31)} \]
\[ \leq 0[(\xi/\lambda)^{4/5}], \text{ since } R_0 < 0(\lambda^{8-2}) \text{ for any } \delta > 0 \]
by (3.37).

This completes the proof of the corollary.

We should note that in general Theorem 3.3 and its corollary do not yield efficient results. Assuming for the moment that any value of \( \lambda \) is attainable, let us find values of \( \lambda \) for which we have efficiency in the sense that \( \text{m.s.e. } \hat{\psi} = 0(\lambda^{-1}) \). If we have \( \lambda \) satisfying
\[ (3.48) \quad \xi \geq (15^{2/3}/2^{5/3})(\log \lambda/\lambda)^{2/3} \text{ by (3.32), and also} \]
\[ \lambda < \xi^{-4}, \text{ i.e., so that in (3.47) } \xi^{4/5} < \lambda^{-1/5}, \]
then
\[ \text{m.s.e. } \hat{\psi} \leq 0(\lambda^{-1}). \]

A brief computation indicates that for reasonably small \( \xi \), (3.48) does not place too stringent bounds on \( \lambda \), e.g., if \( \xi = .1, .01 \) then roughly, \( 500 < \lambda < 1000 \), and \( 2000 < \lambda < 10^8 \), respectively, yield at least efficient results. It should be noted that the upper bounds may not be realistic in terms of the physical situation in the photographic process.

In the case where \( \xi \) is so small that taking \( \lambda \) so large that \( \xi \geq (15^{2/3}/2^{5/3})(\log \lambda/\lambda)^{2/3} \) is physically unrealistic, we will show the two dimensional uniform model to be an adequate approximation and the m.s.e. of the estimation procedure to be at worst \( 0(\lambda^{8-4/3}) \) for any
\( \delta > 0 \). In this case we let \( K > (15^{2/3}25^{2/3}) = 1.916 \) in Theorem 3.2, so that \( \zeta < c = K(\log \lambda / \lambda)^{2/3} \). Again let \( X \) be defined as it was for Theorem 3.2.

\[
(3.49) \quad \text{Var} X \leq \text{EX}^2 \leq T^2 P(X < 0) + \zeta^2 P(X < \zeta) + E(X^2 | X > \zeta).
\]

We now dispose of the last two terms of (3.49). Lemma 3.1 again proves \( T^2 P(X < 0) = O(\lambda^{6-2}) \), if we allow \( 0 < \mu < \lambda \).

\[
(3.50) \quad \zeta^2 P(X < \zeta) \leq O(\lambda^{8-4/3}), \text{since } \zeta < c = O(\lambda^{5-2/3}).
\]

We now dispose of the \( E(X^2 | X > \zeta) \) terms in (3.49) by showing that it is \( O(\lambda^{8-4/3}) \). If \( X > \zeta \) our searching rectangle must stop before moving in the x-direction more than \( c \) beyond \( \zeta \) with probability \( 1 - O(\lambda^{6-2}) \). This follows because within \( 1 - \zeta \) radius of the center we have constant intensity, \( \lambda \), and by the argument of Lemma 3.2 our rectangle must stop with high probability before moving another \( c \).

Therefore,

\[
(3.51) \quad E(X^2 | X > \zeta) \leq [\zeta + c]^2 + O(\lambda^{6-2})
\]

\[
\leq [2c]^2 + O(\lambda^{6-2}) \text{ since } \zeta < c,
\]

\[
\leq O(\lambda^{5-4/3}) \text{ since } c = O(\lambda^{2/3}).
\]

Again we obtain \( \text{Var} X \leq O(\lambda^{8-4/3}) \) and the desired superefficient result is obtained.
In the same manner as in the uniform case, the asymptotic
m.s.e. $\hat{\psi} = \sigma^2/n$ by Theorem 3.1,
where $\sigma^2 = \text{Var} \ X$ of Theorem 3.3, and $n$ is the number of "frames" used.

Again for comparison we assume $\mu = 0$ and compute m.s.e. $\hat{\psi}'$ for the
centroid estimate (13.7). For the present case, (3.2l), $m$ has the
truncated Poisson distribution with parameter

$$V = (\pi \sqrt{3\xi})[1 - (1 - \xi)^3] = \pi \lambda (1 - \xi + \xi^2/3), \text{ i.e.,}$$

$$P(m) = (1 - e^{-V})^{-1} e^{-V} V^m / m!, \text{ for } m = 1, 2, \ldots$$

$$= 0, \text{ otherwise (V here is not to be confused with the V}
of Lemma 3.3 and before).$$

Again let $\rho$ be the polar distance of a single observation.

$$\text{m.s.e. } \psi' = \sum_{m=1}^{\infty} P(m) E(\rho^2)/m, \text{ by (3.18).}$$

Here the density of $\rho$ is

$$f(\rho) = 2\pi \lambda \rho / V$$

$$= 2\pi \lambda (1 - \rho)/(\xi V)$$

$$= 0$$

otherwise.

$$E\rho^2 = (2\pi \lambda / V) [\int_0^{1-\xi} \rho^3 d\rho + (1/\xi) \int_{1-\xi}^{1} \rho^3 d\rho - (1/\xi) \int_{1-\xi}^{1} \rho^4 d\rho]$$

$$= (2\pi \lambda / V)[(1 - \xi)^4/4 + [1 - (1 - \xi)^4]/(4\xi) - [1 - (1 - \xi)^5]/(5\xi)]$$

$$= (1/2 - \xi + \xi^2 - \xi^3/2 + \xi^4/5)/(1 - \xi + \xi^2/3).$$

49
In the manner as (3.19) was obtained we have,

\[
(3.52) \quad \text{m.s.e. } \psi' = E \hat{p}^2 / V + O(\lambda^{-2})
\]

\[
= \left( \frac{1}{2-\xi^2-\xi^3/2+\xi^4/5} \right) / \left[ \pi \lambda(1-\xi^2/3)^2 \right] + O(\lambda^{-2}) .
\]

Note that for \( \xi = 0 \), (3.19) is obtained. Expressions (3.47), (3.48) and (3.52) show that the "frame" procedure admits the possibility of super-efficient estimates.
SECTION 4
UNEQUAL BOUNDARY INTENSITIES

In actual applications the intensity function may not be constant over $S$, the $\lambda$-region. In fact, in the case of the planet Mars there may be a great deal of variation due to the ice caps and the "seasonal effects" on the landscape. To deal with such applications some of the earlier results will now be generalized. First, the one-dimensional case will be extended to the case where the intensities in $S$ near the boundaries, the edges of the $\lambda$-region, differ by some constant of proportionality, $L$. For example, let the intensity function be $\lambda$ and $L\lambda$ in the $2c$ neighborhood of the upper and lower boundaries of $S$, respectively. The intensity function in the region $(2c-1, 1-2c)$ may be left undefined, since the probability that observations in this region enter the estimation procedure is negligible, $O(\lambda^{6-\delta})$ for any $\delta > 0$. The estimates $X_u$ and $X_v$ are determined in the same way as above, depending on the model (see (2.1) and (2.19)). The estimate of location becomes

\begin{equation}
\hat{\Psi}_\alpha = \alpha X_u + (1 - \alpha)X_v,
\end{equation}

where $\alpha$, $0 < \alpha < 1$ is chosen to minimize the asymptotic $\text{Var} \hat{\Psi}_\alpha$. If instead of some estimate of location we wish an estimate of a particular point, e.g., the center of $S$, a correction for the bias may be applied, assuming the other parameters of the model are known.

\begin{equation}
\text{Var} \hat{\Psi}_\alpha = \alpha^2 \text{Var} X_u + (1 - \alpha)^2 \text{Var} X_v, \text{ since } X_u \text{ and } X_v \text{ are asymptotically independent.}
\end{equation}

By setting the first derivative with respect to $\alpha$ equal to 0, the minimum is attained by
\[ (3.55) \quad \alpha = \frac{\text{Var } X_v}{(\text{Var } X_u + \text{Var } X_v)}. \]

If we ignore higher order terms in \( \lambda^{-1} \), we have

\[ (3.56) \quad \text{Var } X_v \approx L^{-2} \text{Var } X_u, \]

since with high probability each of \( X_u \) and \( X_v \) will be the \( N \)th observation from an edge of the \( \lambda \)-region (\( N \) will depend on whether we are in the uniform case or not), and \( \text{Var } X_N \approx (\text{constant})\lambda^{-2} \).

\[ (3.57) \quad \alpha = \frac{L^{-2}}{1+L^{-2}} = \frac{1}{1+L^2} = \alpha^*. \]

Note that the estimation procedure, \((3.53)\), using an approximate optimal \( \alpha, \alpha^* \) depends only on the knowledge of \( L \).

\[ (3.58) \quad \text{Var } \frac{\hat{\lambda}}{\alpha^*} = \alpha^* \text{Var } X_u + (1-\alpha^*)^2 \text{Var } X_v \leq (\text{constant}) \left[ \frac{1}{1+L^2} \right] \frac{1}{\lambda^2} + \frac{h}{1/(L^2\lambda^2)} \],

by the bounds determined in earlier sections on \( \text{Var } X_u \)

\[ \leq O(\lambda^{-2+\delta}). \]

The use of \( \alpha^* \) instead of \( \alpha \) may not provide the minimum variance estimate, but it still gives a superefficient estimate with variance of the same order in \( \lambda \) as in the case of equal boundary intensities. Even if a poor estimate of \( \alpha \) were used we would still have super-efficiency in terms of variance. Such an estimate might be badly biased as an estimate of the center, but this is of no harm in the application, since an estimate of location with constant bias and small variance will suffice.
The two-dimensional procedures can be modified in a similar way for the case where the edge intensities of the disc vary.

Let

\[
\hat{y} = \left( \sum_{j=1}^{n} \beta_j \hat{x}_j, \sum_{j=1}^{n} \gamma_j \hat{y}_j \right),
\]

where \( 0 \leq \beta_j \leq 1, 0 \leq \gamma_j \leq 1, \) and \( \sum \beta_j = \sum \gamma_j = 1 \),

\[
\hat{x}_j = x_j \cos \alpha_j + y_j \sin \alpha_j,
\]

\[
\hat{y}_j = y_j \cos \alpha_j - x_j \sin \alpha_j,
\]

and

\[
x_j = \varepsilon_j \left( \hat{x}_{j_1} - E \hat{x}_{j_1} \right) + (1 - \varepsilon_j)(\hat{x}_{j_3} - E \hat{x}_{j_3}),
\]

\[
y_j = \kappa_j \left( \hat{y}_{j_2} - E \hat{y}_{j_2} \right) + (1 - \kappa_j)(\hat{y}_{j_4} - E \hat{y}_{j_4}),
\]

where \( 0 < \varepsilon_j < 1, 0 < \kappa_j < 1, j = 1, 2, \ldots, n \) (c.f. (2.22) and Theorem 3.1).

The \( \beta \)'s, \( \gamma \)'s, \( \varepsilon \)'s and \( \kappa \)'s are constants to be chosen so that m.s.e. \( \hat{Y} \) is minimized. Note that \( E \hat{Y} = \psi = 0 \).

m.s.e. \( \hat{Y} = E \left( \sum \beta_j \hat{x}_j \right)^2 + E \left( \sum \gamma_j \hat{y}_j \right)^2 \)

\[
= E \left( \sum X_j \beta_j \cos \alpha_j + \sum Y_j \beta_j \sin \alpha_j \right)^2
\]

\[
+ E \left( \sum Y_j \gamma_j \cos \alpha_j - \sum X_j \gamma_j \sin \alpha_j \right)^2
\]

\[
= E \left( \sum X_j \beta_j \cos \alpha_j \right)^2 + E \left( \sum Y_j \beta_j \sin \alpha_j \right)^2
\]

\[
+ E \left( \sum Y_j \gamma_j \cos \alpha_j \right)^2 + E \left( \sum X_j \gamma_j \sin \alpha_j \right)^2
\]

\[
= \sum (\beta_j^2 \cos^2 \alpha_j + \gamma_j^2 \sin^2 \alpha_j) (\text{Var } X_j + \text{Var } Y_j).
\]

(3.60)
The m.s.e. \( \hat{V} \) may be minimized in two stages; first we find the \( \epsilon_j \)'s and \( \kappa_j \)'s to individually minimize the \( 2n \) variances in (3.60); then the \( \beta_j \)'s and \( \gamma_j \)'s are computed. The minimizing values are

\[
(3.61) \quad \epsilon_j = \frac{\text{Var} \hat{x}_j}{\text{Var} \hat{x}_{j1} + \text{Var} \hat{x}_{j3}} 
\]

\[
(3.62) \quad \kappa_j = \frac{\text{Var} \hat{y}_j}{\text{Var} \hat{y}_{j2} + \text{Var} \hat{y}_{j4}}, \quad j = 1, 2, \ldots, n, \]

since the problem here is the same as that solved for the one-dimensional case. We now minimize m.s.e. \( \hat{V} = \sum \beta_j^2 Q_j + \sum \gamma_j^2 R_j \), where

\[
Q_j = \cos^2 \alpha_j (\text{Var} X_j + \text{Var} Y_j) 
\]

\[
R_j = \sin^2 \alpha_j (\text{Var} X_j + \text{Var} Y_j), \quad \text{subject to the constraints} 
\sum \beta_j = 1, \quad 0 \leq \beta_j \leq 1, \quad \sum \gamma_j = 1, \text{and } 0 \leq \gamma_j \leq 1, \quad j = 1, 2, \ldots, n. 
\]

Letting \( \theta_1 \) and \( \theta_2 \) be undetermined Lagrange's multipliers we obtain:

\[
F = \sum \beta_j^2 Q_j + \sum \gamma_j^2 R_j + \theta_1 (\sum \beta_j - 1) + \theta_2 (\sum \gamma_j - 1) 
\]

\[
\frac{\partial F}{\partial \beta_j} = 2 \beta_j Q_j + \theta_1, \quad \frac{\partial F}{\partial \gamma_j} = 2 \gamma_j R_j + \theta_2, \quad j = 1, 2, \ldots, n. 
\]

Setting the partial derivatives equal to zero we obtain,

\[
\beta_j = -\theta_1/(2Q_j), \quad \gamma_j = -\theta_2/(2R_j), 
\]

which with the constraints yields,

\[
\beta_j = Q_j^{-1} / \sum Q_j^{-1} \text{ and } \gamma_j = R_j^{-1} / \sum R_j^{-1}, \quad j = 1, 2, \ldots, n. 
\]

54
Assume that in the $2^{5/2}c^{1/2}$ neighborhood of $(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)$, (see the introductory paragraph for the two-dimensional case), the intensities are $\lambda L_1, \lambda L_2, \ldots, \lambda L_m$, respectively. Since (3.60) is linear in the variances of the $X_j$'s and $Y_j$'s, m.s.e. $\hat{y}$ must be of the same orders in $\lambda$ as was computed before for the uniform and "trapezoidal" cases.

Here an exact application of the estimation procedure depends on the knowledge of the variance of the edge estimate, and again as an approximation we could use first term approximations (giving a procedure dependent only on the $L$'s). Bias depends critically on this knowledge; less than superefficient estimates of the $L$'s lead to poorer estimates of the geometric center than does the equally weighted estimate. When only variance is of concern, then approximately correct weights based on approximately correct values for the $L$'s can provide most of the precision obtainable by the correctly weighted estimate.
SECTION 5
SUMMARY

In this paper various methods of locating a line segment or disc, over which occur observation points having the Poisson distribution with parameter of order $\lambda$ per unit measure, were discussed. Outside the region to be located, the Poisson parameter is assumed small compared to $\lambda$. All the procedures involve linear combinations of two or more estimates of the location of points on the edge of the disc, and the estimates of the edge involve moving in a region of given dimensions until a given number of observations is included. These procedures yield estimates which are asymptotically superefficient in the sense that they have m.s.e. of order $o(\lambda^{-1})$. In the one-dimensional and two-dimensional cases the m.s.e. is of order $\lambda^{-2}$ and $\lambda^{-4/3}$, respectively. In another model the Poisson parameter is assumed to decrease linearly (along any radius in the two-dimensional case) in the neighborhood of the edge. In this case superefficiency is obtainable if the region of decrease is sufficiently narrow. All bounds on the m.s.e. of the procedures are given. Also the case was handled where the Poisson parameter was allowed to vary over the disc up to an order in $\lambda$. The superefficient procedures given in this paper, combined with new photographic techniques, may greatly enhance the clarity of the photographs of planets.
REFERENCES


