SOME STOCHASTIC PROCESSES
ARISING IN NEUROBIOLOGY

Thesis submitted for the degree of
Doctor of Philosophy at the
Australian National University

IAN WILLIAM SAUNDERS

May 1978.
The work described in the main body of this thesis and in Appendix [A] is my own. Appendix [B] contains work carried out jointly with my supervisor, Professor P.A.P. Moran. It is difficult to separate our respective contributions to this work, but it is roughly true to say that section 1 is my own work, while in sections 2 and 3 Professor Moran devised the approach to be used and I worked out the final details.

J. W. Sarnader
Abstract

As an approach to modelling the "matching" of optical receptors in animals to the objects they are designed to see, we study the problem of locating regions of increased brightness in a random "noise" process. Two different models are considered. The first represents the incoming light by the points of a point process on the real line, and is appropriate for low levels of illumination when individual photons must be considered. The second, appropriate for high light levels, represents the incoming light by a Gaussian white noise process. In either case, we study the behaviour of a receptor which measures the total light input within a movable interval of fixed length and define performance measures for this receptor which are analogous to statistical size and power. These measures are then used to define "optimality" for such a receptor.

In the point process case, if the points form a renewal process, we can give conditions on the quantiles of the convolutions of the interpoint distribution which ensure that the optimal receptor has length close to that of the objects it is trying to detect. These conditions are satisfied for a Poisson process. Slightly different conditions ensure that the optimal receptor has length close to zero, and we give a class of distributions satisfying these conditions. We also consider extending the results to two-dimensional point processes, and to the case of more than one receptor.
The quantile properties of convolutions are themselves of interest and we investigate these in more detail. We further consider quantile properties of gamma and F distributions, and apply the gamma results to a problem arising in Bayesian reliability analysis.

The basic results found for the Poisson process model are true also for the Gaussian white noise model. However the proofs in the Gaussian case are generally simpler and the results more complete.
## SOME STOCHASTIC PROCESSES ARISING IN NEUROBIOLOGY

### Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>0. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>0.1 Statement of the problem.</td>
<td>1</td>
</tr>
<tr>
<td>0.2 Point process models.</td>
<td>2</td>
</tr>
<tr>
<td>0.3 Some results for Poisson processes; the scan statistic; Bouman and Van der Velden's model.</td>
<td>3</td>
</tr>
<tr>
<td>0.4 Quantile ranges and ratios; dispersive distributions.</td>
<td>7</td>
</tr>
<tr>
<td>0.5 Continuous state space models.</td>
<td>9</td>
</tr>
<tr>
<td>0.6 Summary.</td>
<td>10</td>
</tr>
<tr>
<td>1. Further results for point process models</td>
<td>11</td>
</tr>
<tr>
<td>1.1 Combining information from more than one receptor.</td>
<td>11</td>
</tr>
<tr>
<td>1.2 The waiting time to the first false detection.</td>
<td>13</td>
</tr>
<tr>
<td>2. Further results on quantile ranges and ratios</td>
<td>20</td>
</tr>
<tr>
<td>2.1 Asymptotic behaviour of quantile ranges and ratios.</td>
<td>20</td>
</tr>
<tr>
<td>2.2 Quantile ranges and ratios at small probability levels.</td>
<td>25</td>
</tr>
<tr>
<td>2.3 An application to Bayesian reliability analysis.</td>
<td>32</td>
</tr>
<tr>
<td>3. Locating an object in Gaussian white noise</td>
<td>35</td>
</tr>
<tr>
<td>3.1 Introduction.</td>
<td>35</td>
</tr>
<tr>
<td>3.2 Measures of long term performance.</td>
<td>39</td>
</tr>
<tr>
<td>3.3 The waiting time to the first false detection.</td>
<td>47</td>
</tr>
<tr>
<td>3.4 A &quot;natural noise&quot; model.</td>
<td>54</td>
</tr>
<tr>
<td>3.5 Combining information from several receptors.</td>
<td>56</td>
</tr>
<tr>
<td>4. References</td>
<td>58</td>
</tr>
</tbody>
</table>

### Appendices

[A] Locating bright spots in a point process. (Saunders (1978))

[B] On the quantiles of the gamma and F distributions. (Saunders and Moran (1978))
0. Introduction

0.1 Statement of the Problem

The problem which motivated the work described here is essentially one of signal detection. The detection of signals in random noise is central to many parts of statistics; it is most explicitly present in communication theory and time series filtering problems, but also occurs wherever the parameter values of a model may change with time or position. For example, the "signal" to be detected in a quality control procedure might be a drop in the quality of the product, whereas in the analysis of satellite land survey data it might be a land formation typical of ore-bearing rock. In the present case, the "signal" is the location of a bright object in a dark background.

First we shall describe in more detail the models that we wish to consider. Although the results have wider application, the original motivation came from a neurobiological problem and we shall retain the terminology suited to the original context.

Suppose that we are studying the physiology of an insect, and especially the structure of its eyes. We find that the eyes contain photoreceptors of a particular size and shape, and that these are connected together and to the insect's brain in a certain pattern. What implications does this have for the nature of the insect's vision? In particular, what sort of objects is the eye best suited to detect? Clearly this involves complex neurophysiological problems, and a purely mathematical approach cannot hope to include all the factors involved. However, by considering a much simplified model for the process of detection, we might hope to gain some insight into the behaviour of an actual biological system.
We shall consider models of two types. In the first, the subject of [A] and Section 1, the light entering the photoreceptor will be considered to consist of individual photons, which will be represented by the points of a point process. This is appropriate where the light level is low, so that the quantum nature of light becomes important. The second type of model considers the light level to be continuously variable, and represents the incoming light by a Gaussian process in continuous time. This can be considered as a limiting case of the point process model, and is appropriate for higher intensities of light. This model is described in Section 3.

0.2 Point Process Models

At low light intensities the quantum nature of light is important, and we consider the incoming light signal to consist of individual points or photons. An object in the visual field will cause an increase in the rate of arrival of the points, and the photoreceptor must attempt to detect such increases. In [A] we consider the following model for the detection process: we suppose that the photons can be represented by points on a line - which may be time or (one-dimensional) space - while the receptor is an interval which moves along the line, registering the number of points that it contains, and signalling the presence of an object whenever this number exceeds a prescribed "detection level" $\alpha$. The problem is then to determine the optimal length and detection level according to certain criteria described in detail in [A]: roughly speaking we wish to minimise the long term rates of false detections and of failures to detect objects.

* [A] and [B] refer to the papers
A. 'Locating bright spots in a point process' and B. 'On the quantiles of the gamma and F distributions' which form Appendices [A] and [B] of this thesis.
This model for the detection process is closely related to models for the detection of clustering in point processes. Here the "signal" to be detected is a deviation from a uniform distribution of points. A number of authors have considered this problem, including Maguire, Pearson and Wynn (1952), Birnbaum (1954), Bartholomew (1954, 1956, 1957), Cox (1955), Epstein (1960), Ederer, Myers and Mantel (1964), P. Lewis (1965), Naus (1966, 1967), Leslie (1969), Wallenstein and Naus (1973, 1974), Huntington and Naus (1975), and Cressie (1977a, 1977b). An example of the application of such models is given in Ederer, Myers and Mantel (1964) who investigate clustering of cases of leukemia in an attempt to discover whether the disease is spread by infection. Their approach to the problem differs from the present one in that they essentially assume a Poisson distribution of points and use counts in a finite number of disjoint intervals rather than the continuum of intervals represented by our receptor. A three dimensional version of the present method was used by Pinkel and Nefzeger (1959), who, however, neglected the fact that the intervals overlapped, so that the significance levels they used were incorrect (as noted by Ederer et al. (1964)). Formulae for the correct significance levels were found by Naus (1965, 1966a), Wallenstein and Naus (1973, 1974) and Huntington and Naus (1975). Since these results are closely related to the work described here, we shall consider them in more detail. We remark that these writers considered only Poisson processes, rather than more general point processes.

0.3 Some results for Poisson processes; the scan statistic; Bouman and Van der Velden's model

If \( \{\tau_i\} \) are the points of a Poisson process, and 

\[ 0 < \tau_1 < \cdots < \tau_N < T \]

are the points of the process falling in a fixed interval \([0,T]\), then the joint distribution of \((\tau_1', \cdots, \tau_N')\)
conditional on \(N\), the number of points in \([0,T]\), is that of (the order statistics of) \(N\) independent uniform random variables on \([0,T]\).

Thus if we are considering only data on a Poisson process in a finite interval, we can treat the points as coming from a uniform distribution. This has the advantage of removing the need to determine the rate of the Poisson process and so simplifies the analysis. Without loss of generality, we can take \(T = 1\).

For \(N\) points \(x_1, \ldots, x_N\) independently uniformly distributed on \([0,1]\), Naus (1965) defines \(P(n|N;p)\) to be the probability that some interval \(I \subseteq [0,1]\) of length \(p \leq 1\) contains \(n\) or more points. This is essentially the probability, conditional on \(N\), that our receptor signals an object somewhere in \([0,1]\) when in fact none are present. Naus (1965) gives formulae for \(P(n|N;p)\) when \(n > N/2\) for \(p \leq \frac{1}{2}\), and for \(n > (N+1)/2\) when \(p > \frac{1}{2}\), and gives the following application, which is directly related to the present work (cf. Section 1.3 of this thesis): A geiger counter receives impulses from a Poisson process of rate \(\lambda\) and registers whenever more than \(n\) pulses occur in a time interval shorter than \(t\). Then, conditioning on the number \(N\) of pulses in the interval \([0,T]\), it is easy to see that

\[
(1) \quad \text{Prob\{First registration occurs before time } T\text{\}} = \sum_{N=n}^{\infty} \frac{\lambda^N}{N!} e^{-\lambda T} (\lambda T)^N
\]

However, since Naus's formulae are valid only for \(N < 2n\), (1) could not be used to calculate this probability.

Naus (1966a), Wallenstein and Naus (1973, 1974) and Huntingdon and Naus (1975) extend this result, the most general formula being

\[
(2) \quad P(n|N;p) = 1 - \sum_{Q} \det(1/h_{ij}) \det(1/l_{ij})
\]

where \(Q\) is the set of all partitions of \(N\) into \(2[p^{-1}] + 1\) integers \(n_1 + \ldots + n_{2[p^{-1}]+1} = N\) satisfying \(n_i + n_{i+1} < n\), \(i=1,\ldots,2[p^{-1}]\)
\[ R = N! \cdot b^{M(p-b)^{N-M}} \]

\[ (M = \sum_{0}^{[p^{-1}]} n_{2i+1}, b = p^{-1} - [p^{-1}]) \]

\[ h_{ij} = \sum_{k=2j-1}^{2i-1} n_{k} - (i-j)n \]

\[ 1 \leq i < j \leq [p^{-1}] + 1 \]

\[ [p^{-1}] + 1 \geq i \geq j \geq 1 \]

\[ \ell_{ij} = \sum_{k=2j}^{2i} n_{k} - (i-j)n \]

\[ 1 \leq i < j \leq [p^{-1}] \]

\[ 1/v! = 0 \text{ if } v < 0 \text{ or } v > N \text{ and } [p^{-1}] \text{ is the integer part of } p^{-1}. \]

Combining (1) and (2) we immediately find

\[ \text{(3) } \text{Prob} \{ \text{First registration occurs before } T \} \]

\[ = 1 - \Sigma_{Q^*}^{R^*} \det(1/h_{ij}) \det(1/\ell_{ij}) \]

where \[ Q^* = \{ \{n_1, \ldots, n_K\} | n_{i} + n_{i+1} < n, \quad i = 1, \ldots, K-1 \} \]

\[ (K = 2[T/t] + 1) \]

\[ h \text{ and } \ell \text{ are defined as above} \]

and \[ R^* = Re^{-\lambda t/T} \lambda^{N/N!} (N=n_1 + \ldots + n_K). \]

These formulae are clearly not well suited for numerical calculation. In Section 1.3 we give simple approximations to the probability (3) and the expected time to first registration.

Naus (1966b) shows that a test based on the maximum number of points in an interval of length \( p \) is the generalised likelihood ratio test of the hypothesis of uniformity against certain clustering alternatives. Clearly the result of Huntingdon and Naus quoted above gives the distribution of this maximum value - called the "scan statistic" by Naus (1966b). Cressie (1977a) studies some properties of this statistic for uniformly distributed points on the unit interval and also on the unit circle. In the latter case the scan statistic gives the most
powerful invariant (under translation) test of uniformity. Cressie (1977b) investigates the asymptotic properties of the scan statistic as \( N \), the number of independent points, tends to infinity. He also considers the scan statistic defined for a Poisson process \( K(t) \) and shows that \( (K(t+h)-K(t)) \), suitably normalized, converges weakly to a certain Gaussian process studied by Slepian (1961) and Shepp (1966, 1971) which we shall consider in greater detail below. Using this limit result, and results of Esary, Proschan and Walkup (1967) on the properties of "associated" random variables, he obtains a bound on the distribution function of the first crossing time of the Gaussian process. This bound was obtained more directly by Lai (1973). In Section 1.2 we use essentially Cressie's methods to find bounds on the crossing time distribution for the (nonasymptotic) Poisson process.

Leslie (1966) considers a slightly different model for the detection process. He supposes that the receptor signals an object whenever \( k \) consecutive inter-point intervals are each shorter than a predetermined length \( \gamma \). This is perhaps less realistic than our model, but leads to simpler mathematics.

Bouman and Van der Velden (1947) propose an extended version of the present model for the response of the human eye to light quanta. They assume that quanta entering the eye remain active for a random (rather than fixed) length of time \( \tau \), and that a visual response occurs - that is, a flash of light is seen - whenever at least \( k \) quanta are simultaneously active. They define \( W_k(\mu,t) \) to be the probability that a response occurs before time \( t \) when the incoming quanta form a Poisson process of rate \( \mu \). The properties of \( W_k(\mu,t) \) when \( \mu \to 0 \) and \( t \to \infty \) have been studied by Yamamoto (1961), Ikeda (1961, 1962, 1965) and Isii (1963). Ikeda (1965) gives the result
\[ W_k(\mu t) \sim 1 - \exp\left(-\frac{\alpha^{k-1} \mu^k t}{(k-1)!}\right) \]

when \( t \to \infty, \mu \to 0 \) with \( \mu^k t \to \text{constant} \)

provided \( t \) has a finite mean \( \alpha \).

Runnenburg (1969) considers a further extension, allowing the incoming photons to form a general renewal process, rather than only a Poisson process. The intervals between arrivals are \( \{\alpha y_i\} \), where \( \alpha \) is a positive constant and \( \{y_1, y_2, \ldots\} \) are i.i.d. random variables on \( (0, \infty) \). He considers the asymptotic behaviour of the number of up-crossings of the level \( k \) by the number of active photons as \( \alpha \to \infty \), and shows that if \( t_\alpha \sim A \alpha^{\rho(k-1)+1} \), where \( A \) and \( \rho \) are constants depending on the distribution of \( y_{\infty} \), that

\[
\lim_{\alpha \to \infty} \Pr(\text{exactly } i \text{ upcrossings in } (0, t_\alpha)) = \frac{e^{-\mu} \mu^i}{i!},
\]

where \( \mu \) depends on the \( y_{\infty} \) distribution, and the upcrossings (with proper scaling) asymptotically form a Poisson process.

All these approaches differ from that which we shall adopt in that they consider the receptor to be operating on data from a finite time interval. We are interested here in the long term behaviour of the receptor and, as we shall see, this leads us to consider somewhat different aspects of the model.

0.4 Renewal processes; quantile ranges and ratios; dispersive distributions

In order to study the long term performance of the receptor we are interested in such quantities as the rate at which objects are mistakenly signalled when none are present, or the proportion of time spent mistakenly signalling. These quantities are analogous to the "size" of a statistical test; similar quantities analogous to "power"
are also defined in Section 3 of [A]. These measures of performance prove to be much more tractable than those considered by the authors referred to in (0.3) above. While they can be defined for arbitrary stationary point processes, they take particularly simple forms when we restrict our attention to renewal processes. In this case we are led to study the properties of the quantiles of convolutions of the inter-point distance distributions. In particular, we define the following classes of distribution functions. For a distribution function \( F \), let \( F_n \) denote the \( n \)-fold convolution of \( F \), and let \( \zeta_n \) denote \( F_n^{-1} \) so that \( \zeta_n(\alpha) \) is the \( \alpha \) quantile of \( F_n \). Then \( F \) is said to have increasing interquantile ranges if \( \zeta_n(\alpha) - \zeta_n(\beta) \) increases with \( n \) for \( \alpha > \beta \) while \( F \) has decreasing quantile ratios if \( \zeta_n(\alpha)/\zeta_n(\beta) \) decreases with \( n \). For a renewal process with such an \( F \) as inter-point distribution we show in [A] that the optimal receptor has length matching that of the object. If \( F \) has increasing quantile ratios as well as increasing interquantile ranges, then the optimal receptor is very short. Some asymptotic properties of quantile ranges and ratios are studied in Section 2.

T. Lewis (1977), prompted by the results in [A] on the interquantile ranges of the exponential distribution, has studied the class \( \mathcal{V} \) of distributions which he calls "dispersive distributions". (The exponential distribution belongs to \( \mathcal{V} \).) If two distributions \( G \) and \( H \) are such that

\[
(3) \quad G^{-1}(\alpha) - G^{-1}(\beta) \geq H^{-1}(\alpha) - H^{-1}(\beta) \quad \forall \alpha > \beta
\]

then we say that \( G \) and \( H \) are ordered in dispersion (o.d.) and write

\[
\text{disp.} \quad G \succ H.
\]

A dispersive distribution \( F \) is such that

\[
\text{disp.} \quad G \succ H \Rightarrow \text{disp.} \quad G \ast F \succ H \ast F
\]

so that convolution with \( F \) preserves ordering in dispersion.
Lewis (1977) shows that the class $\mathcal{D}$ consists essentially of all distributions $F$ with a differentiable density $f$ satisfying

\begin{equation}
\frac{d}{dx} (\ln f(x)) \text{ is non increasing in } x.
\end{equation}

(See Lewis' (1977) paper for a more precise statement).

It is easy to see that a dispersive distribution has increasing interquantile ranges (Lewis (1977, Section 3.1)), so that we have a class of distributions with this property. Unfortunately, we have been unable to construct a similar class of distributions with decreasing quantile ratios (see Lemma 8 of [A] and Section 2.2).

A natural extension of the above definitions is to consider the quantile ranges and ratios of a set $\{F_a\}$ of distributions where the parameter $a$ does not necessarily denote convolution. Lewis (1977) defines an o.d. class to be such a set $\{F_a\}$ with $a \in (-\infty, \infty)$ satisfying

$$a < b \Rightarrow F_a < F_b.$$ 

In [B] we show that the interquantile ranges of $F_a$ increase with $a$, while the quantile ratios decrease with $a$, when $F_a$ is either a gamma distribution with parameter $a$, or a scaled $F$-distribution with parameters $a$ and $m$ ($m$ fixed). In Section 2.3 we apply the gamma distribution result of [B] to solve a problem posed by Waller and Waterman (1977) arising in the determination of gamma priors for Bayesian reliability analysis.

### 0.5 Continuous state space models

As was noted above, when the rate $\lambda$ of a Poisson process tends to infinity, the process, suitably normalized, converges to a Gaussian white noise process. The approach used in [A] to study the long term behaviour of a receptor in a Poisson process can also be applied to Gaussian white noise, and gives similar results. This is the subject of Section 3. Although the problem of signal detection in white noise
has been much studied by workers in communications theory, this particular model has apparently received little attention. Zakai and Ziv (1969) consider essentially this model for a radar range estimation problem. The output of a receptor in this situation is the process studied by Slepian (1961), Shepp (1966, 1971) and Cressie (1977a) which was mentioned above as the limiting form of the scan statistic for a Poisson process. Lai (1973) derives approximations to Shepp's (1971) formulae which he applies to the problem of quickly detecting changes in the level of a white noise process. This is closely related to the present problem, and we shall use Lai's results and methods in Section 3.

0.6 Summary

The central part of this work is contained in the paper "Locating bright spots in a point process" (Saunders (1978)) which forms Appendix [A] to this thesis. In that paper we define the point process model for the receptor, and study its long term properties. This leads to the study of quantile properties of convolutions. In Section 1 we study further properties of the point process model, concentrating mainly on the Poisson process version, while in Sections 2.1–2.2 we study some asymptotic properties of quantile ranges and ratios. The paper "On the quantiles of the gamma and F distributions" (Saunders and Moran (1978)) forms Appendix [B] of the thesis. In this paper we establish properties analogous to the quantile range and ratio properties defined in [A] for two continuous families of distributions. In Section 2.3 we apply these results to a problem in Bayesian reliability analysis.

Finally, in Section 3, we set up and study Gaussian analogues of the receptor model of [A].
1. Further results for point process models

In this Section we consider two additional aspects of the point process model for object location described in [A]. In part 1 we examine the problem of combining information from several receptors to give the optimal combined performance. In part 2 we consider another possible criterion for the performance of a single receptor: the expected time elapsing from the origin to the first signalling of an "object" when in fact no objects are present.

1.1 Combining Information from More than One Receptor

In reality, insects have compound eyes, and must combine the information received from each component. Thus it is of some interest to investigate how the output of the receptors looking in the same direction should be combined to give optimal detection performance.

Suppose that the photons form a Poisson process of rate $\lambda$ outside the objects, and rate $(1+p)\lambda$ within the objects (cf. page 13 of [A]). Suppose also that we have $r \geq 2$ receptors of length $\ell$ each independently receiving photons at these rates. Write $N^{(j)}(\ell;t)$ for the number of photons contained by the $j$th receptor when it is occupying $(t-\ell,t]$. How should the receptors' output be interpreted? Three reasonable methods are:

1) Register an object whenever $\max_{j} N^{(j)}(\ell;t) \geq a_1$, so that an object is registered if any of the receptors individually registers one.

2) Register an object whenever $\min_{j} N^{(j)}(\ell;t) \geq a_2$, so that an object is registered only if all of the receptors simultaneously register one.

3) Register an object if $\sum_{i} N^{(i)}(\ell;t) \geq a_3$, combining the output of all the receptors.
As in Section 3 of [A] we can construct measures $\alpha_1$ and $\beta_2$ of size and power for each of these detection procedures. Again we write:

\begin{align}
(1.1) \quad p(x;n) &= e^{-x} x^n/n! \\
(1.2) \quad P(x;n) &= \sum_{r=0}^{n} p(x;r)
\end{align}

and it is easily shown that, for procedure 1), we have

\begin{align}
(1.3) \quad \alpha_1^{(1)} &= 1 - \{P(\lambda;\alpha_1-1)\}^r \\
(1.4) \quad \beta_2^{(1)} &= \{P(\lambda+\lambda\alpha\min(\ell,L);\alpha_1-1)\}^r
\end{align}

while for procedure 2) we have

\begin{align}
(1.5) \quad \alpha_1^{(2)} &= \{1 - P(\lambda;\alpha_2-1)\}^r \\
(1.6) \quad \beta_2^{(2)} &= 1 - \{1 - P(\lambda+\lambda\alpha\min(\ell,L);\alpha_2-1)\}^r
\end{align}

and, finally, for procedure 3) we have

\begin{align}
(1.7) \quad \alpha_1^{(3)} &= 1 - P(r;\alpha_3-1) \\
(1.8) \quad \beta_2^{(3)} &= P(r(\lambda+\lambda\alpha\min(\ell,L));\alpha_3-1)
\end{align}

It was shown in Section 5 of [A] that the exponential distribution has increasing interquantile ranges and decreasing quantile ratios, so that $P(\lambda;\alpha_1(a)+\lambda\alpha L;\alpha-1)$ increases with $\alpha$, while $P(\lambda(1+\lambda);\alpha(a);\alpha-1)$ decreases with $\alpha$, where $P(\lambda;\alpha(a),\alpha-1) = \alpha$. It is clear from (1.3)-(1.8) that this implies, as in the single-eye situation, that the optimal detector length is close to the object length $L$. In general, the exact length $\ell^{(i)}(a)$ giving the optimal size $\alpha$ detector for procedure $i$ ($i=1,2,3$) will depend on $i$, but since $\ell^{(i)}(a) = L$, we have, approximately, for the optimal detectors,

\begin{align}
(1.9) \quad \alpha_1^{(1)} &= 1 - \{P(\lambda L;\alpha_1-1)\}^r \\
(1.10) \quad \beta_2^{(1)} &= \{P(\lambda(1+\lambda)L;\alpha_1-1)\}^r \\
(1.11) \quad \beta_2^{(2)} &= \{1-P(\lambda L;\alpha_2-1)\}^r
\end{align}
Consider the problem of testing the null hypothesis $\sigma=\lambda L$ against the alternative $\sigma=\lambda(1+\rho)L$ where $\sigma$ is the parameter of a Poisson distribution. If we have $r$ independent observations $X_1, \ldots, X_r$, then it is easy to see that (1.9) and (1.10) give the size and power of a test based on $\max(X_i)$, while (1.11) and (1.12) give the size and power of a test based on $\min(X_i)$, and (1.13) and (1.14) give the size and power of a test based on $\Sigma X_i$. In this situation, $\Sigma X_i$ is a sufficient statistic for $\sigma$, and so the most powerful test of $\sigma=\lambda L$ vs $\sigma=\lambda(1+\rho)L$ of a given size must be based on $\Sigma X_i$. Thus for $a_1^{(1)} = a_1^{(2)} = a_1^{(3)} = a$, the optimal $\beta_2$ is given by procedure 3) with a detector length close to $L$. This suggests that the signals from different segments of a compound eye should be combined before being sent to the brain for interpretation. It would be interesting to investigate whether this is in fact the case.

Note that equations (1.9) to (1.13) are only approximations, so that our conclusion is also only approximate. However the computed values in Table 1.1 suggest that the approximation is sufficiently close for the conclusion to be valid.

1.2 Waiting Time to the First False Detection

A quantity of interest in examining the performance of a detector is the length of time from an arbitrary starting point to the next "false alarm". This is clearly related to the size measure $\alpha_2$ defined in Section 3(a) of [A], but differs in that $\alpha_2$ is based on the time intervals between spurious "objects" rather than the "waiting time" from any point to the next "object".
### Table 1.1a

Combining the Output from Two Independent Receotors

\[
\begin{array}{ccccccc}
\lambda = 1.0 & L = 1.0 & \rho = 1.0 \\
\hline
\alpha & \text{Procedure} & a_0 & a_0 & \beta_2 & a_{0+1} & a_{0+1} \\
\hline
0.05 & 1 & 3 & .622 & .76 & 4 & 1.094 & .71 \\
& 2 & 2 & .889 & .72 & 3 & 1.627 & .76 \\
& 3 & 5 & .985 & .64 & 6 & 1.307 & .68 \\
0.15 & 1 & 3 & .988 & .47 & 4 & 1.594 & .54 \\
& 2 & 1 & .490 & .61 & 2 & 1.340 & .54 \\
& 3 & 3 & .665 & .50 & 4 & 1.020 & .43 \\
0.25 & 1 & 2 & .636 & .41 & 3 & 1.261 & .37 \\
& 2 & 1 & .693 & .44 & 2 & 1.678 & .44 \\
& 3 & 3 & .864 & .33 & 4 & 1.268 & .34 \\
\hline
\end{array}
\]

To aid comparison, \( \beta_2(a^*) \), which is the smaller of \( \beta_2(a_0) \), \( \beta_2(a_{0+1}) \), has been underlined for each procedure.
Table 1.1b

<table>
<thead>
<tr>
<th>$\alpha_1$</th>
<th>Procedure</th>
<th>$a_0$</th>
<th>$\lambda_{a_0}$</th>
<th>$\beta_2$</th>
<th>$a_0+1$</th>
<th>$\lambda_{a_0+1}$</th>
<th>$\beta_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>1</td>
<td>17</td>
<td>0.992</td>
<td>0.054</td>
<td>18</td>
<td>1.068</td>
<td>0.061</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>12</td>
<td>0.927</td>
<td>0.084</td>
<td>13</td>
<td>1.016</td>
<td>0.071</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>28</td>
<td>0.995</td>
<td>0.021</td>
<td>29</td>
<td>1.037</td>
<td>0.022</td>
</tr>
<tr>
<td>0.10</td>
<td>1</td>
<td>15</td>
<td>0.928</td>
<td>0.030</td>
<td>16</td>
<td>1.007</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>11</td>
<td>0.919</td>
<td>0.050</td>
<td>12</td>
<td>1.011</td>
<td>0.040</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>26</td>
<td>0.986</td>
<td>0.010</td>
<td>27</td>
<td>1.030</td>
<td>0.010</td>
</tr>
<tr>
<td>0.15</td>
<td>1</td>
<td>15</td>
<td>0.989</td>
<td>0.013</td>
<td>16</td>
<td>1.071</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>11</td>
<td>0.976</td>
<td>0.028</td>
<td>12</td>
<td>1.072</td>
<td>0.030</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>25</td>
<td>0.994</td>
<td>0.005</td>
<td>26</td>
<td>1.039</td>
<td>0.005</td>
</tr>
<tr>
<td>0.20</td>
<td>1</td>
<td>14</td>
<td>0.956</td>
<td>0.0088</td>
<td>15</td>
<td>1.039</td>
<td>0.0081</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>10</td>
<td>0.926</td>
<td>0.0229</td>
<td>11</td>
<td>1.024</td>
<td>0.0189</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>24</td>
<td>0.991</td>
<td>0.0030</td>
<td>25</td>
<td>1.036</td>
<td>0.0033</td>
</tr>
<tr>
<td>0.25</td>
<td>1</td>
<td>14</td>
<td>0.998</td>
<td>0.0045</td>
<td>15</td>
<td>1.083</td>
<td>0.0058</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>10</td>
<td>0.967</td>
<td>0.0146</td>
<td>11</td>
<td>1.066</td>
<td>0.0150</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>23</td>
<td>0.980</td>
<td>0.0020</td>
<td>24</td>
<td>1.026</td>
<td>0.0020</td>
</tr>
</tbody>
</table>

To aid comparison, $\beta_2(a^*)$, which is the smaller of $\beta_2(a_0)$, $\beta_2(a_0+1)$, has been underlined for each procedure.
Since the process $I(t)$ of [A], Section 2, is stationary
in the absence of objects, we need only consider the waiting time
from zero to the first spurious object, since the time from other
points will have the same distribution. Thus we define

\[ T(a) = \inf\{t \geq 0 | I(t) = 1\}. \]

Clearly,

\[ \text{Prob}\{T(a) = 0\} = \text{Prob}\{I(0) = 1\} \]
\[ = \text{Prob}\{N(\ell; 0) \geq a\} \]
\[ = F_{\ell}^{(a-1)}(\ell) \]
\[ = a_1. \]

The following heuristic argument gives some idea of the behaviour
of $ET(a)$ when $a_1$ is small. As in [A] $\tau_k$ will denote the $k^{th}$ point of
the process after zero. Clearly,

\[ \text{Prob}\{T(a) > \tau_{na}\} = \text{Prob}\{N(\ell; \tau_{ka}) < a, 1 \leq k \leq na\} \]
\[ \leq \text{Prob}\{N(\ell; \tau_{ka}) < a, 1 \leq k \leq n\} \]
\[ = \text{Prob}\{\tau_{ka} - \tau_{k(a-1)+1} > \ell, 1 \leq k \leq n\} \]
\[ = (1 - a_1)^n \text{ since } \{\tau_{ka} - \tau_{k(a-1)+1}\} \text{ are independent} \]
\[ \text{random variables}. \]

Now when $a$ is large, so that $a_1$ is small, it is likely that the number
of points $\tau_k < T(a)$ will be large, so that values of $T(a)$ of the order
of $\tau_k$ will be very unlikely.

As $n \to \infty \tau_{na} \sim n a \mu$, by the strong law of large numbers, so that we
will have

\[ \text{Prob}(T(a) > na\mu) \approx \text{Prob}(T(a) > \tau_{na}) \]

for large $n$, so that

\[ \frac{t}{a_1\mu} \]

\[ \text{Prob}(T(a) > t) \leq (1 - a_1)^{\frac{t}{a_1\mu}} \]

for large values of $t/a_1\mu$.

Thus, since the bound (2.5) applies for large $t$, and small values of
$t$ make little contribution, we have approximately
(2.6) \[ E(T(a)) = \int_0^\infty \operatorname{Prob}(T(a) > t) \, dt \]
\[ \leq \int_0^\infty (1-a_1^t)^{-\lambda} dt \]
\[ = \frac{a_1^{\lambda}}{\ln(1-a_1^{-1})} \]

so that \( E(T(a)) = 0(\alpha) \) for fixed \( a_1 \).

More precise information can be obtained for a stationary Poisson process of rate \( \lambda \).

We first obtain an upper bound on \( E(T(a)) \) which improves on the very rough bound in (2.6).

Let \( B_i \) denote the event \( \{N(\ell; t) < a \text{ for all } t \in (i\ell, (i+1)\ell)\} \), and define \( A_1, A_2 \) and \( A \) by

(2.7)
\[
A_1(n) = \bigcap_{i=1}^n B_i \quad i \text{ odd}
\]
\[
A_2(n) = \bigcap_{i=0}^n B_i \quad i \text{ even}
\]
\[
A(n) = A_1 \cap A_2 = \{N(\ell; t) < a \text{ for all } t \in (0, (n+1)\ell)\}
\]

so that \( A \subseteq A_1 \) and \( A \subseteq A_2 \). Thus

(2.8) \[ \operatorname{Prob} A < \min(\operatorname{Prob} A_1, \operatorname{Prob} A_2) \]

Since Poisson counts in disjoint intervals are independent, it is clear that the events \( B_i \) and \( B_j \) are independent if \( |i-j| > 1 \). Also, since the process is stationary, \( \operatorname{Prob} B_i = \operatorname{Prob} B_j \) for any \( i, j \). Thus

(2.9) \[
\operatorname{Prob} A_1 = \operatorname{Prob} \bigcap_{i=1}^n B_i \quad i \text{ odd}
\]
\[
= \prod_{i=1}^n \operatorname{Prob} B_i \quad i \text{ odd}
\]
\[
= (\operatorname{Prob} B_1)^{n/2} \quad n \text{ even}
\]
\[
= (\operatorname{Prob} B_1)^{(n+1)/2} \quad n \text{ odd}
\]
Similarly,

\[(2.10) \quad \text{Prob } A = \left(\text{Prob } B_{1}\right)^{\frac{n}{2} + 1} \quad \text{n even} \]

\[= \left(\text{Prob } B_{1}\right)^{\frac{n+1}{2}} \quad \text{n odd} \]

which combined with (2.8) gives

\[(2.11) \quad \text{Prob } A < \left(\text{Prob } B_{1}\right)^{\frac{n}{2} + 1} \quad \text{n even} \]

\[< \left(\text{Prob } B_{1}\right)^{\frac{n+1}{2}} \quad \text{n odd} . \]

Now A is the event \(\{T(\alpha) > (n+1)e\}\) and \(\text{Prob } B_{1}\) can be evaluated using Karlin and McGregor's (1959) results as in section 3 of [A] to give

\[(2.12) \quad \text{Prob } B_{1} = r(\lambda;\alpha-1) \]

\[= \{p(\lambda;\alpha-1)\}^{2} - p(\lambda;\alpha) \sum_{m=0}^{\alpha-2} p(\lambda;\alpha). \]

Thus, writing \(F_{T}\) for the distribution function of \(T(\alpha)\), we see that

\[(2.13) \quad 1 - F_{T}(n+1)e) < r^{\frac{n}{2} + 1} \quad \text{n even} \]

\[< r^{\frac{n+1}{2}} \quad \text{n odd} . \]

Writing \(A_{i}^{c}\) for the complement of \(A_{i}\) \((i=1,2)\) we have also

\[(2.14) \quad \text{Prob } A = 1 - \text{Prob}(A_{1}^{c} \cup A_{2}^{c}) \]

\[= 1 - \text{Prob}(A_{1}^{c}) - \text{Prob}(A_{2}^{c}) + \text{Prob}(A_{1}^{c} \cap A_{2}^{c}) \]

\[> 1 - \text{Prob}(A_{1}^{c}) - \text{Prob}(A_{2}^{c}) \]

\[= \text{Prob } A_{1} + \text{Prob } A_{2} - 1 \]

\[= r^{\frac{n}{2} + 1} + r^{\frac{n}{2}} - 1 \quad \text{n even} \]

\[= 2r^{\frac{n+1}{2}} - 1 \quad \text{n odd} \]

giving a lower bound on \(1 - F_{T}\) which may be useful for small \(n\), but tends to -1 as \(n \to \infty\). We shall obtain a better lower bound below.
Now

\[ \text{ET}(a) = \int_0^\infty [1-F_T(x)]dx \]

\[ < \sum_{n=0}^{\infty} \lambda [1-F_T(n\lambda)] \]

\[ = \lambda + \sum_{n=0}^{\infty} \lambda [1-F_T(n+1)\lambda] \quad \text{since } F_T(0) = 0 \]

\[ < \lambda + 2\lambda \sum_{m=0}^{\infty} r^m, \quad \text{using } (2.14) \]

\[ = \lambda + 2\lambda \frac{r}{1-r} \]

\[ = \frac{\lambda(1+r)}{1-r} . \]

We can obtain a useful lower bound on ET(a) by the following method, which is based on a result of Cressie (1977b). Recall from (0.3) Huntingdon and Naus's (1975) formula

\[ \text{Prob}\{T(a) < t\} = 1 - \sum_{k=0}^{\infty} Q_k R^* \det(1/h_{ij}) \det(1/k_{ij}) \]

where \( Q_k^*, R^*, h_{ij}, k_{ij} \) are as defined in (0.3), with \( T \) replaced by \( t \) and \( t \) replaced by \( \lambda \). This gives us an exact expression for ET(a):

\[ \text{ET}(a) = \int_0^\infty [1-\text{Prob}(T(a) < t)]dt \]

\[ = \sum_{k=0}^{\infty} \int_{kl} \int_{kl} F_{Q_k^* R} e^{-\lambda/l} dt \lambda^N/N! \det(1/h_{ij}) \det(1/k_{ij}). \]

This formula, involving infinite sums of determinants, is clearly not well suited for numerical calculation.

Esary, Proschan and Walkup (1967) define associated random variables to be random variables \( T_1, \ldots, T_n \) such that
(2.18) \[ \text{cov}(f(T), g(T)) \geq 0 \quad \text{for all pairs of non-decreasing} \]
\[ f, g : \mathbb{R}^n \to \mathbb{R} \text{ for which } E[f(T), g(T)] \text{ and} \]
\[ E[f(T)]g(T) \text{ exist.} \]

They show that if \( T_1, \ldots, T_n \) are associated, then

(2.19) \[ \text{Prob}\{T_1 \leq t, \ldots, T_n \leq t\} \geq \prod_{j=1}^{n} \text{Prob}\{T_j \leq t\}. \]

Cressie (1975, 1977b) defines a stochastic process \( T(t) \) to be associated if \( (T(t_1), \ldots, T(t_n)) \) are associated for any time points \( t_1, \ldots, t_k \) and shows that \( \{N(\ell; t) = K(t+\ell) - K(t)\} \) is associated for a Poisson process \( K(t) \) of rate \( \lambda \).

Since the sample paths \( N(\ell; t) \) are a.s. right continuous, they are determined by their values on a dense set. Considering a countable dense set \( D \) and a sequence \( D^{+} \) of finite sets (cf. the derivation of (1.11) of Section 3), we can extend the result (2.19) to the interval \([0,n\ell]\) and deduce that

(2.20) \[ \text{Prob}\{\max\{N(\ell; t) | 0 < t \leq n\ell\} \leq a\} \]
\[ \geq \prod_{j=1}^{n} \text{Prob}\{\max\{N(\ell; t) | (j-1)\ell < t \leq j\ell\} \leq a\} \]
\[ = [\text{Prob}\{\max\{N(\ell; t) | 0 < t \leq \ell\} \leq a\}]^n \]
\[ \equiv r(\lambda \ell; a-1)^n \]

where, once again,

(2.21) \[ r(\lambda \ell; a-1) = P(\lambda \ell; a-1)^2 - P(\lambda \ell; a) \sum_{m=0}^{a-2} P(\lambda \ell; m). \]

The inequality (2.20) gives us a bound on the distribution function \( F_T \) of \( T(a) \), since

(2.22) \[ \{T(a) > n\ell\} = \{\max\{K(t+\ell) - K(t) | 0 < t \leq n\ell\} \leq a\} \]
Thus we can obtain a lower bound on $ET(a)$, since

$$ET(a) = \int_{0}^{\infty} (1 - F_T(x)) \, dx$$

$$> \sum_{n=1}^{\infty} \ell (1 - F_T(n\ell)) \quad \text{since } 1 - F_T(x) \text{ is decreasing in } x$$

$$> \ell \sum_{n=1}^{\infty} r^n$$

$$= \frac{\ell}{1-r}.$$ 

In Table 2.1 we compare the bounds (2.6), (2.15) and (2.23) with estimates of $ET(a)$ each based on 1000 computer simulation runs. The computed values suggest that the lower bound (2.23) may be close enough to $ET(a)$ to serve as an estimate.
### Expected Time to First False Detection for a Poisson Process

<table>
<thead>
<tr>
<th>Rate $\lambda$</th>
<th>Length $\ell$</th>
<th>Detection Level $\alpha$</th>
<th>( \text{ET}(\alpha) ) (2.23)</th>
<th>Simulation (2.15)</th>
<th>(2.6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>1.0</td>
<td>2</td>
<td>0.9</td>
<td>0.9</td>
<td>2.8</td>
</tr>
<tr>
<td>1.0</td>
<td>1.0</td>
<td>3</td>
<td>3.5</td>
<td>5.1</td>
<td>8.0</td>
</tr>
<tr>
<td>1.0</td>
<td>1.0</td>
<td>4</td>
<td>13.6</td>
<td>19.6</td>
<td>28.1</td>
</tr>
<tr>
<td>5.0</td>
<td>1.0</td>
<td>2</td>
<td>0.001</td>
<td>0.003</td>
<td>1.0</td>
</tr>
<tr>
<td>5.0</td>
<td>1.0</td>
<td>8</td>
<td>1.5</td>
<td>2.4</td>
<td>4.0</td>
</tr>
<tr>
<td>5.0</td>
<td>1.0</td>
<td>10</td>
<td>6.3</td>
<td>8.8</td>
<td>13.7</td>
</tr>
<tr>
<td>10.0</td>
<td>1.0</td>
<td>10</td>
<td>0.12</td>
<td>0.25</td>
<td>1.25</td>
</tr>
<tr>
<td>10.0</td>
<td>1.0</td>
<td>15</td>
<td>2.3</td>
<td>3.3</td>
<td>5.6</td>
</tr>
</tbody>
</table>
2. Further results on quantile ranges and ratios

In this Section we examine further the behaviour of the quantile ranges and ratios of convolutions and give an application of the results of [B] to a problem arising in the estimation of Bayesian priors.

In part 1 we consider the asymptotic behaviour of quantile ranges and ratios of the n-fold convolution $F_n$ as $n \to \infty$. We show that, for fixed $\alpha$ and $\beta$, $F_n^{-1}(\alpha) - F_n^{-1}(\beta)$ increases with n, while $F_n^{-1}(\alpha)/F_n^{-1}(\beta)$ decreases with n for all sufficiently large n, provided $F$ has a finite fourth moment.

In part 2 we consider the behaviour of the quantile ranges and ratios of $F_n$ at small probability levels, and obtain some conditions which ensure that $F_n$'s quantile ratios are not decreasing.

Finally in part 3 we give the application mentioned above.

2.1. Asymptotic behaviour of quantile ratios and ranges

If the distribution function $F$ satisfies the Central Limit Theorem, so that, writing $F_n$ for the n-fold convolution of $F$,

\begin{equation}
F_n(n\mu + \sqrt{n}\sigma x) \to \phi(x)
\end{equation}

where $\phi(x)$ is the standard normal distribution function, $\mu = \int_0^\infty x dF(x) < \infty$, $\sigma^2 = \int_0^\infty (x-\mu)^2 dF(x) < \infty$, then the quantiles of $F_n$ will be approximately

\begin{equation}
\gamma_n(\alpha) = n\mu + \sqrt{n}\sigma \phi^{-1}(\alpha).
\end{equation}

Since $\gamma_n(\alpha, \beta) = [n\mu + \sqrt{n}\sigma \phi^{-1}(\alpha)] - [n\mu + \sqrt{n}\sigma \phi^{-1}(\beta)]$ increases with n, while $[n\mu + \sqrt{n}\sigma \phi^{-1}(\alpha)]/[n\mu + \sqrt{n}\sigma \phi^{-1}(\beta)]$ decreases with n for $\alpha > \beta \in (0, 1)$, (1.2) suggests that $F$ will asymptotically have increasing interquantile ranges and decreasing quantile ratios as $n \to \infty$. In fact it is easy to see that the approximation (1.2) is not sufficiently accurate to ensure this, for the change $\Delta = \gamma_n(\alpha, \beta) - \gamma_{n+1}(\alpha, \beta)$ is $O(n^{-\frac{1}{2}})$, while the error
in approximating \( \zeta_n(\alpha) - \zeta_n(\beta) \) by \( \gamma_n(\alpha, \beta) \) is \( O(1) \), and so may outweigh \( \Delta \).

However, by taking additional terms in the expansion (1.1) we obtain the following result (See [A] Section 4 for a definition of the class \( S^+ \))

**Theorem 1:** If \( F \in S^+ \) has a finite fourth moment, then for any fixed \( \alpha > \beta \in (0,1) \), there exists \( N \) such that, for \( n > N \),

\[
(1.3) \quad \zeta_n(\alpha) - \zeta_n(\beta) \text{ increases with } n, \quad \text{while}
\]

\[
(1.4) \quad \zeta_n(\alpha)/\zeta_n(\beta) \text{ decreases with } n.
\]

**Proof:** We use the extension of (1.1) to the Edgeworth expansion (Feller (1971, p.541)). Since \( F \in S^+ \) is nonsingular, we have

\[
(1.5) \quad F_n(nu + \sqrt{n} \sigma x) = \phi(x) + \phi(x) \{ n^{-\frac{1}{2}} R_3(x) + n^{-\frac{1}{2}} R_4(x) \} + o(n^{-1})
\]

where \( \phi(x) = d\phi(x)/dx \) is the standard normal density, and \( R_3, R_4 \) are polynomials in \( x \) whose coefficients do not depend on \( n \).

Define

\[
(1.6) \quad F_n^*(x) = F_n(nu + \sqrt{n} \sigma x)
\]

and

\[
(1.7) \quad \psi_n(x) = \phi(x) + \phi(x) \{ n^{-\frac{1}{2}} R_3(x) + n^{-\frac{1}{2}} R_4(x) \}
\]

so that

\[
F_n^*(x) = \psi_n(x) + o(n^{-1}).
\]

Choose \( c_0 \in (0, \phi^{-1}(\beta)), c_1 \in (\phi^{-1}(\alpha), 1) \) and write

\[ d = \inf\{\phi(x) | c_0 \leq x \leq c_1 \}. \]

Differentiating (1.7), we have

\[
(1.8) \quad \psi_n'(x) = \phi(x) + \phi(x) \{ n^{-\frac{1}{2}} (R_3'(x) - xR_3(x)) + n^{-\frac{1}{2}} (R_4'(x) - xR_4(x)) \}
\]

so that \( \psi_n'(x) = \phi(x) + o(n^{-\frac{1}{2}}) \). Thus, for large enough \( n \),
\[(1.9) \quad d = \inf_n \{\psi'(x) | c_0 \leq x \leq c_1 \} > 0\]

so that \(\psi_n\) has an inverse \(\psi_n^{-1}\) in \([\psi_n(c_0), \psi_n(c_1)]\) and also, by (1.7), for large enough \(n\),

\[(1.10) \quad \psi_n(c_0) < \beta, \quad \alpha < \psi_n(c_1).\]

Write

\[(1.11) \quad e_n = \sup_n \{|F_n'(x) - \psi_n(x)| | c_0 \leq x \leq c_1 \} = o(n^{-1}) \text{ by (1.5)},\]

and so \(e_n/d_n = o(n^{-1})\). Hence, by (1.10)

\[(1.12) \quad c_0 < \psi_n^{-1}(\beta) - e_n/d_n, \quad c_1 > \psi_n^{-1}(\alpha) + e_n/d_n\]

for sufficiently large \(n\).

Thus, for any \(\delta \in [\alpha, \beta]\), since \(\psi_n > d_n\),

\[(1.13) \quad F_n(\psi_n^{-1}(\delta) - e_n/d_n) \leq \psi_n(\psi_n^{-1}(\delta) - e_n/d_n) + e_n < \psi_n(\psi_n^{-1}(\delta)) - e_n + e_n = \delta\]

and so, since \(F_n((\zeta_n(\delta) - n\mu)/\sqrt{n}\sigma) = \delta\) and \(F_n\) is increasing,

\[(1.14) \quad \frac{\zeta_n(\delta) - n\mu}{\sqrt{n}\sigma} > \psi_n^{-1}(\delta) - \frac{e_n}{d_n}.\]

Similarly,

\[(1.15) \quad \frac{\zeta_n(\delta) - n\mu}{\sqrt{n}\sigma} < \psi_n^{-1}(\delta) + \frac{e_n}{d_n}\]

and (1.14), (1.15) and (1.11) together imply

\[(1.16) \quad \zeta_n(\delta) = n\mu + \sqrt{n}\sigma \psi_n^{-1}(\delta) + o(\sqrt{n}).\]

We now invert (1.7), using the Lagrange inversion formula

(Whittaker and Watson (1927), p.133) to give \(\psi_n^{-1}\) in terms of \(\psi^{-1}\).
This, together with (1.16) gives a sufficiently accurate expansion for $\zeta_n$. The Lagrange formula states that, if $g$ and $h$ are analytic functions, and

$$w = v + g(w)$$

then

$$h(w) = h(v) + \sum_{r=1}^{\infty} \frac{D_v^{r-1} [h'(v) g(v)]^r}{r!}$$

where $D_v$ denotes $d/dv$, $h'(v) = D_v h(v)$.

Thus, taking

$$w = \phi(x), \delta = v = \psi_n(x), h(v) = \phi^{-1}(v),$$

$$g(v) = -\phi(\phi^{-1}(v)) [n^{-\delta} R_3(\phi^{-1}(v)) + n^{-\delta} R_4(\phi^{-1}(v))]$$

we find that

$$(1.17) \quad \psi_n^{-1}(v) = \phi^{-1}(v) + \sum_{r=1}^{\infty} \frac{D_v^{r-1} [\phi(\phi^{-1}(v))^{r-1} [n^{-\delta} R_3(\phi^{-1}(v)) + n^{-\delta} R_4(\phi^{-1}(v))]]}{r!}$$

where $R_3$ and $R_4$ are analytic functions, independent of $n$.

From (1.16) and (1.17) we see that, for $\delta \in [a, b]$

$$(1.18) \quad \zeta_n(\delta) = n\mu + n^{\delta} \sigma\phi^{-1}(\delta) + \sigma R_3(\phi^{-1}(\delta)) + n^{\delta} R_4(\phi^{-1}(\delta)) + o(n^{\delta})$$

Thus

$$\zeta_n(a) - \zeta_n(b) = n^b \sigma [\phi^{-1}(a) - \phi^{-1}(b)]$$

$$+ \sigma [R_3^*(\phi^{-1}(a)) - R_3^*(\phi^{-1}(b))]$$

$$+ o(n^{b-1}) [R_4^*(\phi^{-1}(a)) - R_4^*(\phi^{-1}(b))]$$

$$+ o(n^{b-1}),$$

so that
\[(\zeta_{n+1}(\alpha) - \zeta_{n+1}(\beta)) - (\zeta_{n}(\alpha) - \zeta_{n}(\beta))\]
\[= ((n+1)^{-\frac{1}{2}} - n^{-\frac{1}{2}}) \alpha[\phi^{-1}(\alpha) - \phi^{-1}(\beta)]\]
\[+ ((n+1)^{-\frac{1}{2}} - n^{-\frac{1}{2}}) \alpha[R^*_4(\phi^{-1}(\alpha)) - R^*_4(\phi^{-1}(\beta))]\]
\[+ o(n^{-\frac{1}{2}})\]
\[= \gamma_n^{-\frac{1}{2}} \alpha[\phi^{-1}(\alpha) - \phi^{-1}(\beta)] + o(n^{-\frac{1}{2}})\]
\[> 0 \text{ for large enough } n.\]

Thus \(\zeta_n(\alpha) - \zeta_n(\beta)\) is increasing with \(n\) for large enough \(n\). Similarly,

\[
\frac{\zeta_n(\alpha)}{\mu} = 1 + \frac{\alpha}{\mu} - \frac{n}{2} R^*_3(\phi^{-1}(\alpha)) + \frac{3}{2} R^*_4(\phi^{-1}(\alpha))
\]
\[+ o(n^{-\frac{1}{2}})\]
\[
n \mu / \zeta_n(\beta) = 1 - \frac{\alpha}{\mu} - \frac{n}{2} R^*_3(\phi^{-1}(\beta)) - \frac{3}{2} R^*_4(\phi^{-1}(\beta))
\]
\[+ \frac{\sigma^2}{\mu^2} n^{-1} [\phi^{-1}(\beta)]^2 + \frac{2\sigma^2}{\mu^2} R^*_3(\phi^{-1}(\beta)) + o(n^{-\frac{1}{2}}).\]

Hence

\[
\frac{\zeta_n(\alpha)}{\zeta_n(\beta)} = 1 + \frac{\alpha}{\mu} n^{-\frac{1}{2}} (\phi^{-1}(\alpha) - \phi^{-1}(\beta)) + n^{-\frac{3}{2}} A(\alpha, \beta)
\]
\[+ n^{-\frac{3}{2}} B(\alpha, \beta) + o(n^{-\frac{3}{2}})\]

where

\[
A(\alpha, \beta) \equiv -\frac{\sigma^2}{\mu^2} \phi^{-1}(\alpha) \phi^{-1}(\beta) + \frac{\sigma}{\mu} [R^*_4(\phi^{-1}(\alpha)) - R^*_4(\phi^{-1}(\beta))]
\]
\[+ \frac{\sigma^2}{\mu^2} [\phi^{-1}(\beta)]^2\]

and

\[
B(\alpha, \beta) \equiv \frac{\sigma}{\mu} [R^*_4(\phi^{-1}(\alpha)) - R^*_4(\phi^{-1}(\beta))] + \frac{2\sigma^2}{\mu^2} \phi^{-1}(\beta) R^*_3(\phi^{-1}(\beta))
\]
\[- \frac{\sigma^2}{\mu^2} [\phi^{-1}(\alpha) R^*_3(\phi^{-1}(\beta)) + \phi^{-1}(\beta) R^*_3(\phi^{-1}(\alpha))]
\]
\[+ \frac{\sigma^3}{\mu^3} \phi^{-1}(\alpha) [\phi^{-1}(\beta)]^2 ,\]

do not depend on \(n\).
Thus
\[
\frac{\zeta_n(\alpha)}{\zeta_n(\beta)} - \frac{\zeta_{n+1}(\alpha)}{\zeta_{n+1}(\beta)} = \frac{n^{-\frac{1}{2}} - (n+1)^{-\frac{1}{2}}}{\mu} \left[ \frac{-3}{\mu} \phi^{-1}(\alpha) - \phi^{-1}(\beta) \right] + \frac{3}{2} \left( \frac{n}{2} - \frac{n+1}{2} \right) B(\alpha, \beta) + o(n^{-2})
\]
> 0 for sufficiently large $n$,

and so $\zeta_n(\alpha)/\zeta_n(\beta)$ decreases with $n$ for large enough $n$.

2.2 Quantile Ranges and Ratios at Small Probability Levels

The results of Section 2.1 show that, for distribution $F$ with finite fourth moments, and any fixed $\alpha > \beta$, $\zeta_n(\alpha) - \zeta_n(\beta)$ increases with $n$, and $\zeta_n(\alpha)/\zeta_n(\beta)$ decreases with $n$ for all sufficiently large $n$.

In this section we shall examine the behaviour of the quantile ranges and ratios for small values of $\alpha$.

Recall from [A] the definition of the set $S^+$ of distributions supported on an interval $I$ contained in $(0, \infty)$ and having a nonvanishing density on $I$.

Let $F$ and $G$ be two distributions in $S^+$ and suppose that $F$ is supported on $(x^-, x^+)$ while $G$ is supported on $(y^-, y^+)$, where $x^+$ and $y^+$ may be $\infty$. Suppose further that $y^- > 0$, and that $F$ and $G$ can be expanded in Taylor series on the right of $x^-, y^-$ respectively. Write $H = F \ast G$, $\zeta(\alpha) = G^{-1}(\alpha)$, $\zeta^*(\alpha) = H^{-1}(\alpha)$.

Suppose that for some $j \geq 0$

(2.1) $F(x^-) = F'(x^-) = F''(x^-) = \ldots = F^{(j)}(x^-) = 0 < F^{(j+1)}(x^-)$

and for some $r \geq 0$

(2.2) $G(y^-) = G'(y^-) = G''(y^-) = \ldots = G^{(r)}(y^-) = 0 < G^{(r+1)}(y^-)$. 

Then, since

\[ H(x) = \int_{x^-}^{x^+} F(u) G'(x-u) \, du \]

\[ = 0 \quad \text{if } z < x^- + y^- \]

or if \( z > x^+ + y^+ \)

\( H \) is supported on \((x^- + y^-, x^+ + y^+)\), and, differentiating (2.3) we find

\[ H'(x^+ + y^-) = H'(x^+ + y^-) = \ldots = H^{(j+r+1)}(x^+ + y^-) = 0 \]

\[ H^{(j+r+2)}(x^- + y^-) = F^{(j+1)}(x^-) G^{(r+1)}(y^-) > 0. \]

Thus, using Taylor expansions of \( G \) and \( H \), we see that

\[ a = G(\zeta(a)) = G(y^-) + (\zeta(a) - y^-) G'(y^-) + \ldots \]

\[ = G^{(r+1)}(y^-) \frac{(\zeta(a) - y^-)^{x+r}}{(x+r+1)!} + \ldots \]

while

\[ a = H(\zeta*(a)) = H^{(j+r+2)}(x^- + y^-) \frac{(\zeta*(a) - x^- - y^-)^{j+r+2}}{(j+r+2)!} + \ldots \]

Combining (2.5) and (2.6) we obtain

\[ (\zeta(a) - y^-)^{x+r+1} \left\{ \frac{G^{(r+1)}(y^-)}{(r+1)!} + (\zeta(a) - y^-) \frac{G^{(r+2)}(y^-)}{(r+2)!} + \ldots \right\} \]

\[ = (\zeta*(a) - x^- - y^-)^{j+r+2} \left\{ \frac{H^{(j+r+2)}(x^- + y^-)}{(j+r+2)} + \ldots \right\} \]

or

\[ \left\{ \frac{\zeta(a)}{y^-} - 1 \right\}^{x+r+1} \left\{ \frac{(y^-)^{x+r+1} G^{(r+1)}(y^-)}{(r+1)!} + \ldots \right\} \]

\[ = \left\{ \frac{\zeta*(a)}{x^- + y^-} - 1 \right\}^{j+r+2} \left\{ \frac{(x^- + y^-)^{j+r+2} G^{(j+1)}(x^-) G^{(r+1)}(y^-)}{(j+r+2)!} + \ldots \right\} \]

As \( a \to 0 \), \( \zeta(a) \to y^- \) and \( \zeta*(a) \to x^- + y^- \). The terms omitted in the curly brackets in (2.8) tend to zero, while the included terms are independent of \( a \). Thus
\[
\frac{\left(\frac{\xi(a)}{y} - 1\right)}{\left(\frac{\xi^*(a)}{x+y} - 1\right)} \rightarrow C, \text{ a strictly positive constant,}
\]

and so, since \(\xi^*(a)/(x+y) \rightarrow 1\),

\[
\frac{\left(\frac{\xi(a)}{y} - 1\right)}{\left(\frac{\xi^*(a)}{x+y} - 1\right)} \rightarrow C \left(\frac{\xi^*(a)}{x+y} - 1\right) \rightarrow 0.
\]

Thus for small enough \(a\),

\[
\frac{\xi(a)}{\xi(0)} = \frac{\xi(a)}{y} < \frac{\xi^*(a)}{x+y} = \frac{\xi^*(a)}{\xi^*(0)}
\]

and clearly

\[
F \ast G \overset{q.r.}{=} G.
\]

Once again, let \(F_n\) denote the \(n\)-fold convolution of \(F\). Then if \(x^- > 0\), take \(G = F_n\), so that \(y^- = nx^- > 0\), and, by (2.12)

\[
F_{n+1} = F \ast F_n \overset{q.r.}{=} F_n
\]

so that \(F\) does not have decreasing quantile ratios.

Thus we have

**Lemma 1:** If 1) \(F \in S^+\) is supported on \((x^-, x^+)\) where \(x^- > 0\),

2) \(F\) is \((j+2)\) times differentiable for some \(j\)

3) \(F(x^-) = F'(x^-) = \ldots = F(j)(x^-) = 0, F(j+1)(x^-) > 0\)

4) \(F\) has a valid Taylor series expansion for \(x > x^-\),

\[
F(x) = \frac{(x-x^-)^{j+1}}{(j+1)!} F(j+1)(x^-) + \frac{(x-x^-)^{j+2}}{(j+2)!} F(j+2)(\theta)
\]

where \(x^- < \theta < x\)

then \(F\) does not have decreasing quantile ratios.
In particular, Lemma 1 implies that the shifted exponential distribution, and in general the shifted gamma distributions do not have decreasing quantile ratios.

Recall from Lemma 4 of [A] that if $F$ has decreasing quantile ratios and the quantiles $\ell_n(\alpha)$ of $F$ satisfy

\[
\frac{\ell_n(\alpha) - \ell_n(\beta)}{n} \geq \frac{\ell_{n+1}(\alpha) - \ell_{n+1}(\beta)}{n+1}
\]

then $F(x-c)$ has decreasing quantile ratios for arbitrary positive $c$.

If a distribution $F \in S^+$ satisfies conditions 2), 3) and 4) of Lemma 1 above, then for any positive $c$, $F(x-c)$ cannot have decreasing quantile ratios, and so the quantiles of $F_n$ cannot satisfy (2.14).

Consider in particular the case where $F$ is a strictly stable distribution on $(0,\infty)$ of exponent $\delta$, so that $\ell_n(\alpha) = n^{1/\delta} \ell_1(\alpha)$. Then (2.14) becomes

\[
(2.15) \quad n^{1/\delta - 1} (\ell_1(\alpha) - \ell_1(\beta)) \geq (n+1)^{1/\delta - 1} (\ell_1(\alpha) - \ell_1(\beta)).
\]

Since $F$ has a density and satisfies 2), 3) and 4) of Lemma 1, it follows from (2.15) that $\delta$ must be less than 1. (Feller (1971, p.448)).

The question whether any nontrivial distributions exists which satisfies (2.14) is still open. (The distributions $\delta_c$ concentrated at $c > 0$ trivially satisfy (2.14), since $\ell_n(\alpha) = nc$.)

The proof of (2.12) depends essentially on the results (2.4), which show that $H$'s derivatives at $x+y$ are zero to a higher order than $G$'s derivatives at $y$; it is not necessary that $H$ be of the form $F \ast G$. In fact we have the following result.

**Lemma 2.** If 1) $G,H \in S^+$ are supported on $(y^-,y^+)$, $(z^-,z^+)$ respectively, where $y^- > 0$, $z^- > 0$,

2) $G(y^-) = G'(y^-) = \ldots = G^{(r)}(y^-) = 0$, $G^{(r+1)}(y^-) > 0$

$H(z^-) = H'(z^-) = \ldots = H^{(r)}(z^-) = 0$, $H^{(r+1)}(z^-) > 0$
3) G and H have valid Taylor series expansions, and

$$4) \left( \frac{z}{y} \right)^{r+1} H^{(r+1)}(z) < \left( \frac{y}{z} \right)^{r+1} G^{(r+1)}(y),$$

then \( q.r. \)

\[ H \nmid G. \]

**Proof:** Write \( \zeta(a) \) for the \( a \) quantile of G, \( \zeta^*(a) \) for the \( a \) quantile of H. Then as in the derivation of (2.8) above, we have

$$\begin{align*}
(2.16) & \quad \frac{\zeta(a)}{y} - 1 = \left( \frac{y}{z} \right)^{r+1} \left( \frac{z}{y} \right)^{r+1} (r+1)! \\
& \quad \frac{\zeta^*(a)}{z} - 1 = \left( \frac{z}{y} \right)^{r+1} \left( \frac{y}{z} \right)^{r+1} (r+1)! \\
& \quad = \left( \frac{\zeta(a)}{y} - 1 \right)^{r+1} \left\{ \frac{\zeta^*(a)}{z} - 1 \right\}^{r+1} \left( \frac{z}{y} \right)^{r+1} (r+1)! \\
& \quad < 1 \text{ if condition 4) holds.}
\end{align*}$$

Thus, as \( a \to 0 \),

$$\begin{align*}
(2.17) & \quad \left( \frac{\zeta(a)}{y} - 1 \right)^{r+1} \left( \frac{\zeta^*(a)}{z} - 1 \right)^{r+1} \left( \frac{z}{y} \right)^{r+1} (r+1)! \\
& \quad \to C = \left( \frac{\zeta(a)}{y} - 1 \right)^{r+1} \left( \frac{\zeta^*(a)}{z} - 1 \right)^{r+1} \left( \frac{z}{y} \right)^{r+1} (r+1)! \\
& \quad < 1 \text{ if condition 4) holds.}
\end{align*}$$

Thus for small enough \( a \), \( \zeta(a)/y < \zeta^*(a)/z \) and \( H \nmid G. \)

For distributions supported in a neighbourhood of the origin, the situation is different. Consider distributions F, G, H supported on \((0,x^+)\), \((0,y^+)\), \((0,z^+)\) respectively. Let \( \zeta(a) \), \( \zeta^*(a) \) be the \( a \) quantiles of G, H respectively, and suppose that

$$\begin{align*}
(2.18) & \quad G(0) = G'(0) = \ldots = G^{(j)}(0) = 0, G^{(j+1)}(0) > 0 \\
& \quad H(0) = H'(0) = \ldots = H^{(r)}(0) = 0, H^{(r+1)}(0) > 0
\end{align*}$$

where \( r > j \), and that the Taylor series

$$\begin{align*}
(2.19) & \quad \alpha = G(\zeta(a)) = \frac{\zeta(a)^{j+1}}{(j+1)!} G^{(j+1)}(0) + \frac{\zeta(a)^{j+2}}{(j+2)!} G^{(j+2)}(0) + \ldots \\
& \quad \alpha = H(\zeta^*(a)) = \frac{\zeta^*(a)^{r+1}}{(r+1)!} H^{(r+1)}(0) + \frac{\zeta^*(a)^{r+2}}{(r+2)!} H^{(r+2)}(0) + \ldots
\end{align*}$$

are valid for \( 0 \leq a < 1 \).
Then for \( a > \beta \),

\[
\frac{a}{\beta} = \frac{G(\xi(a))}{G(\xi(\beta))} = \left( \frac{\xi(a)}{\xi(\beta)} \right)^{j+1} \frac{1 + \xi(a)}{\left( \frac{G(j+2)(0)}{(j+2) G(j+1)(0)} \right)^{j+1}} + \ldots
\]

\[
= \frac{H(\xi^*(a))}{H(\xi^*(\beta))} = \left( \frac{\xi^*(a)}{\xi^*(\beta)} \right)^{r+1} \frac{1 + \xi^*(a)}{\left( \frac{H(r+2)(0)}{(r+2) H(r+1)(0)} \right)^{r+1}} + \ldots
\]

If for some fixed \( k \in (0,1) \) we take \( \beta = k \alpha \) and let \( \alpha \to 0 \) we see from (2.20) that \( \frac{\xi(a)}{\xi(\beta)} > \frac{\xi^*(a)}{\xi^*(\beta)} \) for small enough \( \alpha \).

Applying this to the case \( G = F_n, H = F_{n+1}^* \) we see that the quantile ratios of \( F \) are decreasing for sufficiently small probability levels \( \alpha, \beta \).

A similar result holds for interquantile ranges. From (2.19) we have

\[
(2.21) \quad \frac{\xi(a)^{j+1} - \xi(\beta)^{j+1}}{(j+1)!} G(j+1)(0) + \ldots
\]

\[
= \xi^*(a)^{r+1} - \xi^*(\beta)^{r+1} \frac{H(r+2)(0)}{(r+2) H(r+1)(0)} + \ldots
\]

so that

\[
(2.22) \quad (\xi(a) - \xi(\beta)) \left\{ \frac{G(j+1)(0)}{(j+1)!} (\xi(a)^j + \xi(a)^{j-1} \xi(\beta) + \ldots + \xi(\beta)^j) + O(\xi(a)^{j+1}) \right\}
\]

\[
= (\xi^*(a) - \xi^*(\beta)) \left\{ \frac{H(r+1)(0)}{(r+1)!} (\xi^*(a)^r + \ldots + \xi^*(\beta)^r) + O(\xi^*(a)^{r+1}) \right\}
\]

Now from (2.19), \( \xi(a)^{j+1}/\xi^*(a)^{r+1} \to C > 0 \) as \( \alpha \to 0 \) so that

\[
(2.23) \quad \frac{\xi(a)}{\xi^*(a)} \sim C^{1/j+1}(\xi^*(a))^{3+1} \to 0 \quad \text{as} \quad \alpha \to 0
\]

and hence
\[(2.24) \quad \frac{\zeta(a)^j}{\zeta(a)^r} = \frac{\zeta*(a)}{\zeta(a)} \frac{\zeta(a)^{j+1}}{\zeta*(a)^{r+1}} \to \infty \text{ as } a \to 0.\]

Thus \(\zeta*(a)^r = o(\zeta(a)^j)\) as \(a \to 0\) and hence from (2.22), \(\zeta(a) - \zeta(\beta) < \zeta*(a) - \zeta*(\beta)\) for small enough \(a, \beta\).

Once again taking \(G = F_n, H = F_{n+1}\) we find that \(F\)'s interquantile ranges are increasing for small enough \(a, \beta\).
2.3 An Application to Bayesian Reliability Analysis

The results in [B] for the gamma distribution can be used to solve a problem posed by Waller and Waterman (1977). As part of a procedure for determining gamma priors for Bayesian reliability analysis, it is necessary to find values of the parameters $\lambda$ and $\phi$ satisfying

\[(3.1)\quad f(\lambda, \phi) = \int_0^\infty \frac{\lambda e^{-y} y^{\phi-1}}{\Gamma(\phi)} \, dy = \xi_1\]

and

\[(3.2)\quad g(\lambda, \phi) = \int_0^\infty \frac{b\lambda e^{-y} y^{\phi-1}}{\Gamma(\phi)} \, dy = \xi_2\]

where $\xi_1 < \xi_2 \in (0,1)$ and $b > 1$ are preassigned constants.

The problem posed by Waller and Waterman (1977) was to find whether a solution $(\lambda, \phi)$ satisfying (3.1) and (3.2) always exists, and if so, whether it is unique. We shall show that both are in fact true.

As in [B], we write $x_\phi(\xi)$ for the $\xi$ quantile of the gamma distribution with shape parameter $\phi$. Then (3.1) and (3.2) can be rewritten as

\[(3.3)\quad x_\phi(\xi_1) = \lambda\]

\[(3.4)\quad x_\phi(\xi_2) = b\lambda\]

or

\[(3.5)\quad \xi = \int_0^{x_\phi(\xi_1)} \frac{e^{-y} y^{\phi-1}}{\Gamma(\phi)} \, dy = \int_0^{x_\phi(\xi_2)} \frac{y^{\phi-1}}{\Gamma(\phi)} \, dy = \int_0^{x_\phi(\xi_2)} \frac{y^{\phi-1}}{\Gamma(\phi+1)} \, dy = \int_0^{x_\phi(\xi_1)} \frac{y^{\phi-1}}{\Gamma(\phi)} \, dy\]

Clearly, if we can find a value of $\phi$ satisfying (3.4), then $\phi$ and $\lambda = x_\phi(\xi_1)$ satisfy (3.3) and so also (3.1) and (3.2).

By Theorem 1 of [B], $r(\phi) = x_\phi(\xi_1)/x_\phi(\xi_2)$ decreases with $\phi$, since $\xi_1 < \xi_2$. Thus it will be sufficient to show that $r(\phi) \to \infty$ as $\phi \to 0$, $r(\phi) \to 1$ as $\phi \to \infty$ in order to establish (3.4).

First we note that
and also, integrating by parts,

\[(3.6)\]
\[
\xi = \int_0 \frac{e^{-y} y^{\phi-1}}{\Gamma(\phi)} \, dy
\]

\[
= \left[ \frac{e^{-y} y^{\phi}}{\Gamma(\phi+1)} \right]_0^{\xi} + \int_0^{\xi} \frac{e^{-y} y^{\phi}}{\Gamma(\phi+1)} \, dy
\]

\[
= -x_{\phi}(\xi) e^{-\phi} x_{\phi}(\xi) + \frac{e^{-\phi} x_{\phi}(\xi)}{\Gamma(\phi+1)}.
\]

Combining (3.5) for \(\xi = \xi_1\) with (3.6) for \(\xi = \xi_2\), we obtain

\[(3.7)\]
\[
\frac{\xi_1}{\xi_2} \geq \frac{-x_{\phi}(\xi_1) e^{1/\phi} x_{\phi}(\xi_1) \phi}{x_{\phi}(\xi_2)^{\phi}}
\]

or

\[(3.8)\]
\[
\frac{x_{\phi}(\xi_2)}{x_{\phi}(\xi_1)} \geq \left(\frac{\xi_2}{\xi_1}\right)^{1/\phi} \frac{-x_{\phi}(\xi_1)}{\phi}.
\]

Now, using the first result of Theorem 1 of [B], we see that
\[x_{\phi}(\xi) = x_{\phi}(\xi) - x_{\phi}(0)\]
increases with \(\phi\) so that \(x_{\phi}(\xi)\) is bounded for \(\phi < 1\). Thus it follows from (3.5) that \(x_{\phi}(\xi) \to 0\) as \(\phi \to 0\) for all \(\xi\), and so, for all small enough \(\phi\),

\[(3.9)\]
\[
\frac{\xi_2}{\xi_1} \geq \frac{-x_{\phi}(\xi_1)}{\phi} > 1, \text{ since } \xi_2 > \xi_1.
\]

Thus, by (3.8), \(r(\phi) = x_{\phi}(\xi_2)/x_{\phi}(\xi_1) \to \infty\) as \(\phi \to 0\).

To show that \(r(\phi) \to 1\) as \(\phi \to \infty\), we note that since \(r\) is monotone, it is sufficient to consider integer values of \(\phi\), and that a gamma distribution with parameter \(\phi = n\), an integer, is the convolution of \(n\) unit exponential distributions. Hence, by the central limit theorem,

\[(3.10)\]
\[
x_n(\xi) = n + \sqrt{n} \phi^{-1}(\xi) + o(\sqrt{n})
\]
so that
\[ (3.11) \quad \frac{x_n(\xi_2)}{x_n(\xi_1)} = \frac{1 + n^{-\frac{1}{2}} \phi^{-1}(\xi_2) + o(n^{-\frac{1}{2}})}{1 + n^{-\frac{1}{2}} \phi^{-1}(\xi_1) + o(n^{-\frac{1}{2}})} \]

\rightarrow 1 \text{ as } n \rightarrow \infty

and we have established (3.4).
3. Locating an Object in Gaussian White Noise

3.1 Introduction

A similar problem to that considered in [A] arises when it is required to locate an "object" in the presence of Gaussian white noise \( dW(t) \), where \( W \) is a standard Wiener process. Zakai and Ziv (1969) refer to this as a "range estimation problem" which they formulate as follows:

Suppose that a signal \( S(t), T_1 \leq t \leq T_2 \), is transmitted, having the form of a rectangular pulse

\[
S(t_0 + t) = s \text{ (constant)} \quad 0 \leq t \leq \ell \\
= 0 \quad \text{otherwise}
\]

and we wish to estimate \( t_0 \) from the received signal \( r(t) = s(t)dt + dW(t) \).

The maximum likelihood estimator \( \hat{t} \) is given by the value of \( t \) which maximizes

\[
\lambda(t) = \int_t^{t+\ell} r(u) \, du,
\]

which when \( T_2 - T_1 > \ell \) can be written as

\[
\lambda(t) = \int_{-\infty}^{\infty} S(u-t_0)S(u-t) \, du + \int_{-\infty}^{\infty} S(u-t) \, dW(u).
\]

Large errors in the estimate \( \hat{t} \) will occur when the noise component \( \int S(u-t_0) \, dW \) reaches a high value at a large distance from \( t_0 \). Thus we are interested in the behaviour of this component when \( |t-t_0| \) is large. Note that the value of the noise integral is

\[
\int_{-\infty}^{\infty} S(u-t_0) \, dW(u) = \mathbb{E}[W(t + \ell) - W(t)]
\]

and that the "signal" integral in (1.2) is zero when \( |t-t_0| > \ell \) and equal to \( s^2 \ell \) for \( t = t_0 \).

Zakai and Ziv define the "threshold probability" \( P \) as

\[
P = \text{Prob} \{ \lambda(t_0) < \max \{\lambda(t) \mid |t-t_0| > \ell, t \in [T_1, T_2]\} \}
\]

taking this probability as an approximation to the probability

\[
P' = \text{Prob} \{ |\hat{t} - t_0| > \ell \}.
\]

Bounds on \( P \) can be obtained by considering the events

\[
A_n = \{ \lambda(t_0) < \max \{\lambda(t) \mid n\ell < t < (n+1)\ell\} \}
\]

where \( T_1 \leq n\ell \leq T_2 - \alpha \). Zakai and Ziv use an argument based on the
independence of $A_n$ and $A_m$ for $|n-m| > 1$ (cf. the derivation of (2.11) of Section 2) to obtain the bounds

$$\begin{align*}
1 - \text{Prob}(A_1) & \leq P \leq 2[1 - \text{Prob}(A_1)] \\
\text{when } & \frac{k_1(t - T_1)}{2\ell} \text{ and } \frac{k_2(T_2 - t_0 - \ell)}{2\ell} \text{ are integers.}
\end{align*}$$

The probabilities $\text{Prob}(A_n)$ can be obtained using a result of Slepian (1961), who gives a formula for

$$\text{(1.6)} \quad \text{Prob}(W(t + \ell) - W(t) < a, \ 0 < t < T} \quad \text{where } 0 < T < \ell. \quad \text{Since } \lambda(t_0) \text{ is a normal } (s^2, b^2 s^2) \text{ random variable, the probability } \text{Prob}(A_n) \text{ can be obtained by integrating Slepian's formula.}

Shepp (1971) extended Slepian's (1961) result to the case where $T$ may be greater than $\ell$. The resulting expression can be used to give the exact value of the threshold probability $P$, although the formulae obtained are very complicated. Shepp's results can also be used to give closer bounds on $P$ than those in (1.5), as we now show.

The argument is based on Lai (1973). We shall again assume for simplicity that $k_1 = \frac{(t_0 - T_1)}{2\ell}$ and $k_2 = \frac{(T_2 - t_0 - \ell)}{2\ell}$ are integers. The results depend on noting that, if $X = (X_1, X_2, ..., X_k)$ is a multivariate normally distributed random variable with covariance matrix $(\lambda_{ij})$ and $Y$ is a random variable on $\mathbb{R}$ independent of $X$, then

$$\text{(1.7)} \quad \text{Prob}(X_1 < Y, ..., X_k < Y) \text{ is a nondecreasing function of } \lambda_{ij} \text{ for any } i, j.$$  

This can readily be verified by differentiation. In particular, if the $\lambda_{ij}$ are all positive, then reducing some of them to zero, introducing independence between the corresponding $\{X_i\}$, will reduce the value of the probability in (1.7).

Thus, by reducing $\lambda_{ij}$ to zero whenever $1 \leq i \leq r, r + 1 \leq j \leq k$, we see that
\[ \text{Prob}(X_1 < Y, \ldots, X_k < Y) \geq \text{Prob}(X_1 < Y, \ldots, X_r < Y) \times \text{Prob}(X_{r+1} < Y, \ldots, X_k < Y) \]

if \( \lambda_{ij} \geq 0 \ \forall \ i, j \). (Cf. equation (2.19) of Section 1.2.)

Then for any \( t_1, t_2, \ldots, t_k \in [T_i, t_0 - \varepsilon] \cup (t_0 + \varepsilon, T_i], \lambda(t_k) = W(t_k + \varepsilon) - W(t_k) \)

is independent of \( \lambda(t_0) = W(t_0 + \varepsilon) - W(t_0) + s^2 \), and
\[
\text{cov}(\lambda(t_k), \lambda(t_j)) = \max(b^2(1 - |t_k - t_j|/\varepsilon), 0) \geq 0. \text{ Hence}
\]

\[ \text{(1.8)} \quad \text{Prob}(\lambda(t_1) < \lambda(t_0), \ldots, \lambda(t_k) < \lambda(t_0)) \]
\[ \geq \text{Prob}(\lambda(t_1) < \lambda(t_0), \ldots, \lambda(t_r) < \lambda(t_0)) \]
\[ \times \text{Prob}(\lambda(t_{r+1}) < \lambda(t_0), \ldots, \lambda(t_k) < \lambda(t_0)) \]

i.e.

\[ \text{(1.9)} \quad \text{Prob}\left\{ \max_{j=1, \ldots, k} \{\lambda(t_j)\} < \lambda(t_0) \right\} \]
\[ \geq \text{Prob}\left\{ \max_{j=1, \ldots, r} \{\lambda(t_j)\} < \lambda(t_0) \right\} \]
\[ \times \text{Prob}\left\{ \max_{j=r+1, \ldots, k} \{\lambda(t_j)\} < \lambda(t_0) \right\}. \]

The sample paths \( \lambda(t) \) are a.s. continuous, and so a.s. determined by their values on a dense set \( D \subseteq \mathbb{R} \). Hence

\[ \text{(1.10)} \quad \text{Prob}\left\{ \sup_{t \in [a, b]} \lambda(t) < \lambda(t_0) \right\} = \text{Prob}\left\{ \sup_{t \in [a, b] \cap D} \lambda(t) < \lambda(t_0) \right\}. \]

Thus taking \( D \) to be a countable set dense in \( \mathbb{R} \), and considering finite sets \( D_n + D \), we can extend the result (1.9) to intervals, showing that

\[ \text{(1.11)} \quad \text{Prob}(\lambda(t) < \lambda(t_0) \ \forall \ t \in [0, 2k_1 \varepsilon] \cup [2(k_1 +1) \varepsilon, 2(k_1 + k_2 + 1) \varepsilon]) \]
\[ \geq \text{Prob}(\lambda(t) < \lambda(t_0) \ \forall \ t \in [0, 2k_1 \varepsilon] \cup [2(k_1 +1) \varepsilon, 2(k_1 + k_2 -1) + 1) \varepsilon]) \]
\[ \times \text{Prob}(\lambda(t) < \lambda(t_0) \ \forall \ t \in [0, \varepsilon]) \]
\[ = \text{Prob}(\lambda(t) < \lambda(t_0) \ \forall \ t \in [0, 2k_1 \varepsilon] \cup [2(k_1 +1) \varepsilon, 2(k_1 + k_2 -1) + 1) \varepsilon]) \]
\[ \times \text{Prob}(A_1). \]

Thus, using induction first on \( k_1 \) and then on \( k_2 \), we see that
so that

(1.12) \[ p \leq 1 - \text{Prob}(A_1) \]

giving an upper bound which is smaller than that in (1.5).

For \( t^* \in [(2k_1 + 2k_2 + 2)\ell, (2k_1 + 2k_2 + 3)\ell] \), \( \lambda(t^*) \) is independent of \( \{\lambda(t) | t \in [0, (2k_1 + 2k_2 + 1)\ell]\} \). Hence, by a similar argument to that used to derive (1.11), we see that

(1.13) \[ \text{Prob}\{\lambda(t) < \lambda(t_0) \forall t \in [0,2k_1\ell] \cup [(2k_1 + 2)\ell, (2k_1 + 2k_2 + 2)\ell], \lambda(t^*) > \lambda(t_0) \text{ for some } t^* \in [(2k_1 + 2k_2 + 2)\ell, (2k_1 + 2k_2 + 3)\ell]\} \]

\[ = p \text{ Prob}\{\lambda(t) < \lambda(t_0) \forall t \in [0,\ell] \cup [(2k_1 + 2)\ell, (2k_1 + 2k_2 + 1)\ell]\}
\]

where \( p = \text{Prob}\{\lambda(t) < \lambda(t_0) \forall t \in [0,\ell]\} \),

(1.14) \[ \lambda(t^*) > \lambda(t_0) \text{ for some } t^* \in [\ell, 2\ell]\} \).

Now, writing \( A_n = \{\lambda(t) < \lambda(t_0) \forall t \in [n\ell, (n+1)\ell]\} \), as above,

\( p_n = \text{Prob} \cap_{0}^{n} A_1 \)

and

(1.14) \[ B(k_1, k_2) = \cap_{n=0}^{2k_1+2k_2+2} A_n \]

we see that

(1.15) \[ p = \text{Prob}(A_1 - A_1 \cap A_2) \]

\[ = p_1 - p_2 \]

while (1.13) implies

(1.16) \[ \text{Prob}(B(k_1, k_2)) - \text{Prob}(B(k_1, k_2+1)) \]

\[ \geq p \text{ Prob}(B(k_1, k_2-1)) \].

Summing (1.16) over \( k_2 \), and a similar inequality over \( k_1 \), we obtain
\[
\begin{align*}
\text{Prob}(B(1,1)) - \text{Prob}(B(k_1,k_2)) & \geq p \sum_{r=0}^{2k_1-2} \text{Prob}(B(k_1,r)) + p \sum_{r=0}^{2k_1-1} \text{Prob}(B(r,1)) \\
& \geq p \sum_{r=0}^{2k_1-2} p_1^r + p \sum_{r=0}^{2k_1-1} p_1^r \quad \text{using (1.12)} \\
& = p \frac{1 - p_1^{2k_1 + 2k_2 - 1}}{1 - p_1}
\end{align*}
\]

so that

\[
1 - P = \text{Prob}(B(k_1,k_2)) \leq p_1^2 - p \frac{1 - p_1^{2k_1 + 2k_2 - 1}}{1 - p_1}
\]

giving a lower bound on \( P \).

### 3.2 Measures of Long Term Performance

Zakai and Ziv's approach to this problem differs from our approach to the point process problem principally in that they consider a single object on a finite interval \([T_1, T_2]\) rather than an infinite interval containing a potentially infinite number of objects. Our approach can also be used in the present situation and, as we shall see, gives similar results to those found for a renewal process.

Suppose that, as before, we set a detection level \( a \), and signal the presence of an object whenever the value of \( \lambda(t) \) exceeds \( a \). The "objects" will be, as in Zakai and Ziv's approach, intervals where the intensity is increased by an amount \( s > 0 \). Consider an object at \([T - L, T]\), and a detector of length \( L < L \). When the detector is occupying \([t-L, t]\), it will register

\[
\lambda(t) = \begin{cases} 
\beta(W(t) - W(t-L)) & \text{if } t \notin J \subseteq [T-L, T+L] \\
\beta(W(t) - W(t-L)) + (t-T) s & \text{if } t \in [T-L, T-L+L] \\
\beta(W(t) - W(t-L)) + L s & \text{if } t \in J_2 = [T-L+L, T] \\
\beta(W(t) - W(t-L)) + (T-t+L) s & \text{if } t \in [T, T+L] 
\end{cases}
\]

Similarly, if \( L > L \), when the detector occupies \([t-L, t]\), it will register
(2.2)  \( \lambda(t) = b(W(t) - W(t-\ell)) \quad \text{if} \quad t \not\in J_2 = [T-L, T+\ell] \)
\[= b(W(t) - W(t-\ell)) + (t-T+L)s \quad \text{if} \quad t \in [T-L, T] \]
\[= b(W(t) - W(t-\ell)) + Ls \quad \text{if} \quad t \in J_1 = [T-L-\ell, T-L] \]
\[= b(W(t) - W(t-\ell)) + (T-t+L)s \quad \text{if} \quad t \in [T-L+\ell, T+L]. \]

Again we wish to choose \( a \) and \( \ell \) to optimize the detector's long run performance. As performance criteria we take quantities analogous to those used in the point process case. Define

(2.3)  \( I(t) = 1 \quad \text{if} \quad \lambda(t) > a \)
\[= 0 \quad \text{otherwise} \]

(2.4)  \( \alpha_1(t) = \frac{1}{t-\ell} \int_\ell^t I(u)du. \)

Also, write

(2.5)  \( \phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \)
\[\Phi(x) = \int_\infty^x \phi(u)du \]
so that \( \phi \) and \( \Phi \) are respectively the standard normal density and distribution functions. Then, since \( W \) is a standard Wiener process,

(2.6)  \( \text{Prob}\{b(W(t) - W(t-\ell)) < x\} = \Phi\left(\frac{x}{b\sqrt{\ell}}\right). \)

Our measure of size will be \( \lim_{t\to\infty} \alpha_1(t) \), which is again a.s. constant, for in the absence of objects \( \{I(t)\} \) is a stationary \( \ell \)-dependent process, i.e. \( I(t) \) and \( I(t+x) \) are independent for \( |x|>\ell \), and so (Ibragimov and Linnik (1971, ch. 17)) \( I(t) \) is strong mixing which implies that it is regular, and so ergodic (or metrically transitive) so that, when no objects are present,
Thus if we require a size \( a \) detector with detection level \( a \), we must choose \( \ell = \ell_a(a) \) where

\[
\ell_a(a) = \frac{a^2}{\beta \phi^{-1}(1-a)^2}.
\]

Since \( \lambda(t) \) shares the "crinkly" nature of the Wiener process, we cannot define a size measure analogous to \( \alpha_2 \) simply as the number of upcrossings of the level \( a \). In fact, if \( W(t^*) = x \), then a.s. \( W(t) = x \) infinitely often in \( (t^*,t^*+\delta) \) for any \( \delta > 0 \) (Cox and Miller (1965, p.229)) and so if \( \lambda(t^*) = a \) then \( \lambda(t) \) crosses the level \( a \) infinitely often in an arbitrarily small interval following \( t^* \). Thus we would have \( \alpha_2 = \infty \) a.s. We shall consider a possible alternative to \( \alpha_2 \) later, but first we shall define a measure of power and find the optimal detector

length.

We define

\[
\beta_2 = 1 - \frac{\int_{J_1} I(t)dt}{\mathbb{E} \int_{J_1} dt} = \text{Prob}\{\lambda(T) \leq a\} = \Phi \left( \frac{a - s \min(\ell, L)}{b \sqrt{\ell}} \right)
\]

since \( \lambda(t) = b(W(t) - W(t-\ell)) + s \min(\ell, L) \) for \( t \in J_1 \).

Thus to obtain the optimal detector of size \( \alpha_1 = a \), we must choose \( a \) to minimize

\[
\Phi \left( \frac{a - s \min(\ell, L)}{b \sqrt{\ell} a} \right)
\]

where \( \ell = \frac{a^2 [\phi^{-1}(1-a)]^{-2}}{b^2} \).
Write \( a^* = \sqrt{L b^*} (1-a) \), so that \( \frac{l_a}{l} < L \) for \( a < a^* \), \( \frac{l_a}{l} > L \) for \( a > a^* \).

Thus for \( a < a^* \),

\[
(2.10) \quad \beta_2 = \phi \left( \frac{a-sa^-1(1-a) - 2b^-2}{a(\Phi^-1(1-a)^{-1})} \right)
\]

\[
= \phi \left( \frac{1-s\Phi^-1(1-a)^{-2b^-2}}{(\Phi^-1(1-a)^{-1})^2} \right)
\]

which decreases with \( a \), while for \( a > a^* \),

\[
(2.11) \quad \beta_2 = \phi \left( \frac{a-sL}{a(\Phi^-1(1-a)^{-1})} \right)
\]

\[
= \phi \left( \frac{1-sL/a}{(\Phi^-1(1-a)^{-1})} \right)
\]

which increases with \( a \).

Thus the minimum value of \( \beta_2 \) occurs when \( a = a^* \), so that \( \frac{l_a}{l} = L \), giving an optimal detector length equal to the object length.

If the length \( L \) is not known exactly, it will be of interest to investigate the effect on \( \beta_2 \) of using a detection level \( a \) and detector length \( l_a \) slightly different from the optimal values \( a^* \) and \( L \). From (2.10) above, we see that, for \( a < a^* \),

\[
(2.12) \quad \frac{d\beta_2}{da} = - \phi ((1-sa^-1(1-a) - 2b^-2) \Phi^-1(1-a)) s \Phi^{-1}(1-a)^{-1} b^{-2}
\]

\[
\frac{d^2\beta_2}{da^2} = \phi'((1-sa^-1(1-a) - 2b^-2) \Phi^-1(1-a)) s^2 \Phi^{-1}(1-a)^{-2} b^{-4}
\]

while for \( a > a^* \), from (2.11), we have

\[
(2.13) \quad \frac{d\beta_2}{da} = \phi ((1-sL/a) \Phi^-1(1-a)) s L a^{-2} \Phi^{-1}(1-a)
\]

\[
\frac{d^2\beta_2}{da^2} = - \phi'((1-sL/a) \Phi^-1(1-a)) s^2 L a^{-4} \Phi^{-1}(1-a)^2
\]

\[\quad - 2\phi((1-sL/a) \Phi^-1(1-a)) s L a^{-3} \Phi^{-1}(1-a)\].
Now, when $a = a^* = \sqrt{b\phi^{-1}(1-a)}$, we have

\begin{align*}
(2.14) \quad s a^* \phi^{-1}(1-a)^{-2} b^{-2} &= sL/a^* \\
&= s\phi^{-1}(1-a)^{-2} b^{-2} = sL(a^*)^{-2} \phi^{-1}(1-a)
\end{align*}

and, using Taylor expansions about $a^*$, we see that for $\delta > 0$

\begin{align*}
(2.15) \quad \beta_2(a^* - \delta) &= \beta_2(a^* + \delta) + \delta^2 \phi((1-sL/a) \phi^{-1}(1-a)) sL a^{-3} \phi^{-1}(1-a) + o(\delta^3) \\
&> \beta_2(a^* + \delta) \quad \text{for small enough } \delta.
\end{align*}

Thus for a given size of error $\delta$, the increase in $\beta_2$ will be less if $a$ is greater than $a^*$ than if $a$ is less than $a^*$. Thus we should prefer to overestimate, rather than underestimate the length of the objects to be detected. Note however that, to first order in $\delta$, $\beta_2(a^* - \delta)$ and $\beta_2(a^* + \delta)$ are identical. Some computed values of $\beta_2$ for various values of $\delta$ are given in Table 2.1.

From (2.12) and (2.13) it is also easy to see that the value $\delta'$ satisfying

\begin{align*}
(2.16) \quad \beta_2(a^* + \delta') &= \beta_2(a^* - \delta)
\end{align*}

is given approximately by

\begin{align*}
(2.17) \quad \delta' &= \delta + \delta^2 \phi((1-sL/a^*) \phi^{-1}(1-a)) sL a^{-3} \phi^{-1}(1-a) \\
&\quad / \phi((1-sL/a^*) \phi^{-1}(1-a)) sL(a^*)^{-2} \phi^{-1}(1-a)}
\end{align*}

so that underestimating the length $L$ by an amount $\delta$ has the same effect on power as overestimating $L$ by $\delta(1+\delta/a^*)$.

As in the point process case, we can evaluate the probability $\beta_1$ of missing an object completely:

\begin{align*}
(2.18) \quad \beta_1 &= \text{Prob}(\lambda(t) < a \text{ for all } t \in J_1).
\end{align*}
Table 2.1

Effect of wrong detector length on $\beta_2$

<table>
<thead>
<tr>
<th>$\alpha_1 = 0.05$</th>
<th>$L = 1.0$</th>
<th>$s/b = 1.0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta$</td>
<td>$\beta_2(L-\delta)$</td>
<td>$\beta_2(L+\delta)$</td>
</tr>
<tr>
<td>0.00</td>
<td>.7405</td>
<td>.7405</td>
</tr>
<tr>
<td>0.05</td>
<td>.7486</td>
<td>.7482</td>
</tr>
<tr>
<td>0.10</td>
<td>.7568</td>
<td>.7553</td>
</tr>
<tr>
<td>0.15</td>
<td>.7651</td>
<td>.7619</td>
</tr>
<tr>
<td>0.20</td>
<td>.7735</td>
<td>.7679</td>
</tr>
<tr>
<td>0.25</td>
<td>.7820</td>
<td>.7735</td>
</tr>
<tr>
<td>0.30</td>
<td>.7905</td>
<td>.7787</td>
</tr>
<tr>
<td>0.35</td>
<td>.7992</td>
<td>.7835</td>
</tr>
<tr>
<td>0.40</td>
<td>.8079</td>
<td>.7881</td>
</tr>
<tr>
<td>0.45</td>
<td>.8168</td>
<td>.7923</td>
</tr>
<tr>
<td>0.50</td>
<td>.8258</td>
<td>.7963</td>
</tr>
</tbody>
</table>
Again we use Karlin and McGregor's (1959) result for crossings of \( n \) independent processes. We shall use the result only for \( n = 2 \) processes, and so for simplicity we quote it only for \( n = 2 \).

**Theorem A** (Karlin and McGregor (1959))

If \( W_1(t), W_2(t) \) are independent standard Wiener processes on \([0,1]\), then

\[
\text{Prob}\{W_1(\tau) > W_2(\tau) \text{ for all } \tau \in [0,1] | W_1(1)=y_1, W_2(1)=y_2, W_1(0)=x_1, W_2(0)=x_2 \} \]

\[
= \frac{\phi(y_1-x_1)\phi(y_2-x_2) - \phi(y_1-x_1)\phi(y_2-x_1)}{\phi(y_1-x_1)\phi(y_2-x_2)}
\]

if \( x_1 > x_2, y_1 > y_2 \)

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

In the case \( \lambda \geq L \), we apply theorem A to \( W_1, W_2 \) defined by

\[
W_1(\tau) = \frac{1}{\sqrt{\lambda-L}} \left\{ W(T-(\lambda-L)\tau) + \frac{a-Ls}{b} \right\}
\]

\[
W_2(\tau) = \frac{1}{\sqrt{\lambda-L}} W(T+(\lambda-L)\tau),
\]

since then \( W_1(\tau) > W_2(\tau) \) for all \( \tau \) in \([0,1]\) if and only if

\( \lambda(t) < a \) for all \( t \) in \( J_1 \). Without loss of generality, we can take

\( W(T-\lambda) = 0 \). Let \( x = W(T-L), y = W(T), z = W(T+\lambda-L) \) so that

\[
W_1(0) = (\lambda-L)^{-\frac{1}{2}} (a-Ls)/b
\]

\[
W_1(1) = (\lambda-L)^{-\frac{1}{2}} (x+(a-Ls)/b)
\]

\[
W_2(0) = (\lambda-L)^{-\frac{1}{2}} y
\]

\[
W_2(1) = (\lambda-L)^{-\frac{1}{2}} z.
\]

Thus, using theorem A, and integrating over \( x, y \) and \( z \),
\[ \text{Prob}\{\lambda(t) < a \, \forall \, t \in J_1\} \]

\[ = E \left\{ \left[ 1 - \frac{\phi((\ell-L)^{-b}(x-y+(a-Ls)/b))\phi((\ell-L)^{-b}(z-(a-Ls)/b))}{\phi((\ell-L)^{-b}x)\phi((\ell-L)^{-b}(z-y))} \right] \times I\left\{ y < \frac{a-Ls}{b}, \, z-x < \frac{a-Ls}{b} \right\} \right\} \]

where \( I \) denotes the indicator function

\[ = \int_{-\infty}^{x} \int_{-\infty}^{y} \left[ \frac{\phi((\ell-L)^{-b}x)\phi((\ell-L)^{-b}(z-y)) - \phi((\ell-L)^{-b}(x-y+(a-Ls)/b))\phi((\ell-L)^{-b}(z-(a-Ls)/b))}{\phi((\ell-L)^{-b}x)\phi((\ell-L)^{-b}(z-y))} \right] \times (\ell-L)^{-b} \times L^{-b} \phi(L^{-b}(y-x)) dz \]

\[ = \int_{-\infty}^{x} \int_{-\infty}^{y} \left[ \phi((\ell-L)^{-b}(x-y + \frac{a-Ls}{b})) \phi((\ell-L)^{-b}(x-y + \frac{a-Ls}{b})) - \phi((\ell-L)^{-b}(x-y + \frac{a-Ls}{b})) \phi((\ell-L)^{-b}x) \right] \times (\ell-L)^{-b} \times L^{-b} \phi(L^{-b}(y-x)) dy . \]

So that, when \( \ell > L, \)

(2.22) \[ \beta_1 = \int_{-\infty}^{x} \left[ \phi((\ell-L)^{-b}(\frac{a-Ls}{b} - u))\phi((\ell-L)^{-b}(\frac{a-Ls}{b} - u)) - \phi((\ell-L)^{-b}(\frac{a-Ls}{b} - u))\psi((\ell-L)^{-b}(\frac{a-Ls}{b} - u)) \right] \times L^{-b} \phi(L^{-b}u) du \]

where we have taken \( u = y-x, \) changed the order of integration and written \( \psi(x) = \int_{-\infty}^{x} \phi(u) du. \)

By a similar argument we can evaluate \( \beta_1 \) for \( \ell \leq L \) as

(2.23) \[ \beta_1 = \int_{-\infty}^{x} \left[ \phi((\ell-L)^{-b}(\frac{a-Ls}{b} - u))^{2} - \phi((\ell-L)^{-b}(\frac{a-Ls}{b} - u))\psi((\ell-L)^{-b}(\frac{a-Ls}{b} - u)) \right] \times (2\ell-L)^{-b} \phi((2\ell-L)^{-b}u) du. \]
We note that the integral in (2.22) is the convolution of a normal (0,L) distribution function with the function \( \rho(x) \) defined by

\[ \rho(x) = \phi(x)^2 - \phi(x)\Psi(x). \]

In fact \( \rho(x) \) is a distribution function, for from (2.24) we have

\[ \rho'(x) = \phi(x)\{\phi(x) + x\Psi(x)\} \]

\[ > 0 \quad \text{for } x > 0. \]

If \( x < 0 \), write \( y = -x > 0 \), so that

\[ \rho'(x) = \phi(-y)\{-\phi(-y) - y\Psi(-y)\} \]

\[ = \phi(y)\{(1-\phi(y)) - y(\Psi(y) - y)\} \]

since \( \Psi(-y) = -y\phi(-y) + \phi(-y) = -y + \Psi(y) \).

Then it is sufficient to show that

\[ S(y) = 1 - \phi(y) - y\Psi(y) + y^2 > 0 \]

for all \( y > 0 \).

But

\[ S'(y) = -\phi(y) - \Psi(y) - y\phi(y) + 2y \]

\[ = 2(y - \Psi(y)) \]

\[ = -2\Psi(-y) \]

\[ < 0 \]

while

\[ \lim_{y \to \infty} S(y) = \lim_{y \to \infty} y(y - \Psi(y)) \]

\[ = \lim_{y \to \infty} y^2(1-\phi(y)) - y\phi(y) \]

\[ = 0 \quad \text{since } 1-\phi(y) \text{ and } \phi(y) \text{ are } o(y^{-k}) \]

for any \( k > 0 \) as \( y \to \infty \).

Thus \( S(y) \downarrow 0 \) as \( y \to \infty \), and so, since the inequality in (2.27) is strict, \( S(y) > 0 \) for all \( y > 0 \). Hence, combining this result with (2.25), we see that \( \rho'(x) > 0 \) for all \( x \), so that, since \( \rho(-\infty) = 0, \rho(\infty) = 1 \), \( \rho \) is a distribution function. In fact, as Shepp (1970) shows, \( \rho \) is the distribution function of the maximum of \([W(t) - W(t-1)]\) for \( 0 \leq t \leq 1 \).
3.3 Waiting time to the first false detection

Another measure of the performance of a detector is the time elapsing before the first false "object" is signalled, when in fact no objects are present. We can use Shepp's (1971) results and a result due to Lai (1973) to obtain a lower bound on the expected time to the first false detection in Zakai and Ziv's model, and use this to obtain a performance criterion analogous to $\alpha_2$.

Lai (1973) shows that, if $X(t) = W(t+1) - W(t)$, and

$$T^*(a) = \inf\{t^l | X(t) \geq a\},$$

then

$$E(T^*(a)) \geq \frac{1}{1 - \rho(a)} - 1.$$

Thus, if $T(a) = \inf\{t^l | A(t) \geq a\}$, then

$$E(T(a)) \geq \frac{\ell}{1 - \rho\left(\frac{a}{b\sqrt{\ell}}\right)}.$$

Lai (1973) also gives an upper bound on $E(T^*(a))$, but since we will be interested in ensuring that the detector does not give false alarms too frequently, we shall be principally interested in the lower bound (3.1).

Suppose that, after first registering a false object at time $T_1$, so that $\lambda(T_1) = a$, we ignore the output of the detector until time $T_1 + \ell$, and then again start watching for objects. For time intervals $\Delta > \ell$, $\lambda(t)$ and $\lambda(t+\Delta)$ are independent. Thus if we define $T_2$ by

$$T_2(a) = \inf\{t | \lambda(t) \geq a, t > T_1(a) + \ell\}$$

then $T_2 - T_1$ has the same distribution as $T_1(a)$, and is independent of $T_1(a)$. Defining $T_3, T_4, \ldots$ similarly, we see that the sequence $\{T_i(a)\}$ forms a renewal process. By excluding the intervals $(T_i, T_i + \ell)$ from consideration, we have removed the infinite number of crossings of the level $a$ which immediately follow $T_i$ and prevent us from considering $\alpha_2$. The natural analogue of $\alpha_2$ is thus the mean
number of renewals per unit length, which is given by the renewal theorem as
\[ \lim_{t \to \infty} \frac{N_t}{t} = \frac{1}{ET_1} \leq \frac{1 - \rho \left( \frac{a}{b/\ell} \right)}{\ell} \]

where \( N_t \) denotes \( \max\{n \mid T_n(a) < t\} \) in the standard notation for a renewal process. Thus we define

\[ \alpha_3(\ell, a) = \frac{1 - \rho(a / (b/\ell))}{\ell} \]

as a third size measure. Taking \( b \) to be fixed, we consider \( \alpha_3 \) as a function of \( a \) and \( \ell \) and, as a preliminary to examining its use as an optimality criterion, we establish some of its properties.

First, it is clear from (3.1), since \( \rho \) is a distribution function, that \( \alpha_3 < 1/\ell \) for all \( a, \ell \). This is also clear from the definition of \( \alpha_3 \), since the intervals \( (T_n - T_{n-1}) \) must be at least of length \( \ell \). Also, for fixed \( \ell \), \( \alpha_3 \) decreases with \( a \), from \( \alpha_3(\ell, -\infty) = 1/\ell \) to \( \alpha_3(\ell, \infty) = 0 \), so that \( a \) can be chosen uniquely to give any desired value of \( \alpha_3 \) between 0 and 1/\( \ell \). Define \( \tilde{\alpha}(\ell, a) \) to be this unique value, so that

\[ \tilde{\alpha}_3(\ell, a) = b/\ell \rho^{-1}(1-a/\ell). \]

The behaviour of \( \alpha_3 \) for fixed \( a \) as \( \ell \) varies is more involved, and we distinguish the cases \( a < 0 \), \( a = 0 \) and \( a > 0 \). For \( a = 0 \),

\( \ell \alpha_3 = 1 - \rho(0) = 1 - \left( \frac{3}{4} + \frac{1}{2\pi} \right) = \frac{3}{4} + \frac{1}{2\pi} \) so that

\[ \alpha_3(\ell, 0) = \left( \frac{3}{4} + \frac{1}{2\pi} \right)/\ell \to 0 \text{ as } \ell \to \infty \]

\[ \to \infty \text{ as } \ell \to 0 \]

and \( \alpha_3 \) takes each value between 0 and \( \infty \) exactly once. For \( a < 0 \)

\[ \ell \alpha_3 = 1 - \rho(a/b/\ell) \to 1 - \rho(0) = \frac{3}{4} + \frac{1}{2\pi} \text{ as } \ell \to \infty \]

\[ \to 1 - \rho(-\infty) = 1 \text{ as } \ell \to 0. \]
Thus for \( a < 0 \)

\[
(3.6) \quad \alpha_3(\ell, a) \to \infty \quad \text{as} \quad \ell \to 0 \\
\qquad \to 0 \quad \text{as} \quad \ell \to \infty
\]

and since \( 1 - p(-a/\sqrt{\ell}) \) is decreasing in \( \ell \), \( \alpha_3 \) again takes each value between 0 and \( \infty \) exactly once for \( \ell \in (0, \infty) \).

For \( a > 0 \) we must consider the properties of \( p \) in more detail.

Recall that

\[
p(x) = \phi(x)^2 - \phi(x)\psi(x)
\]

while

\[
p'(x) = \phi(x)\{\phi(x) + x\psi(x)\}.
\]

Thus, as \( \ell \to 0 \),

\[
(3.7) \quad \lim_{\ell \to 0} \frac{1-p(\frac{a}{\sqrt{\ell}})}{\ell} = \lim_{\ell \to 0} \frac{a}{\sqrt{\ell}^{3/2}} p'(\frac{a}{\sqrt{\ell}}) \\
\quad \text{by l'Hospital's rule,}
\]

\[
= \lim_{x \to \infty} \frac{\ell}{2} \frac{a^2}{\sqrt{\ell}^{3/2}} x^3 p'(x)
\]

where \( x = a/\sqrt{\ell} \)

\[
= 0
\]

since \( x^j \phi(x) \to 0 \) for any \( j \).

Also, as \( \ell \to \infty \), \( \ell \alpha_3 \to 1-p(0) \), and so \( \alpha_3 \to 0 \).

In fact we shall show that, for any \( a > 0 \), \( \alpha_3(\ell, a) \) is a unimodal function of \( \ell \), which is bounded above. Thus there is a limit on the \( \alpha_3 \) value attainable for a given \( a > 0 \), and any value below the limit is attained for exactly two distinct values of \( \ell \).

From the definition (3.2) it is clear that \( \alpha_3 \) is infinitely differentiable w.r.t. \( \ell \) on \( (0, \infty) \). Thus it is sufficient to show that

\[
\frac{\partial \alpha_3}{\partial \ell} = 0
\]

has a unique root in \( (0, \infty) \). Now
Thus it is sufficient to show that
\( (3.9) \quad H(x) = 1 \) has a unique root in \((0, \infty)\)
where \( H(x) = \frac{\phi'(x)}{\phi(x)} - \rho(x) \).

Note that \( H(x) \) is a function only of \( x \), and does not depend otherwise on \( a, b \) or \( \ell \). Now from (2.24) and (2.25) we have
\[
(3.9) \quad H(x) = \phi(x)^2 + \frac{1}{2} x (x^2 - 1) \phi(x) \phi(x) + (\frac{1}{2} x^2 - 1) \phi(x)^2
\]
since \( \psi(x) = \phi(x) + x \phi(x) \), and differentiating (3.9) gives
\[
(3.10) \quad H'(x) = \frac{1}{2} \phi(x) \{ x (5 - x^2) \phi(x) + (3 + 4 x^2 - x^4) \phi(x) \}.
\]

From (3.9) we see that \( H(0) = \frac{1}{2} - 1/2 \pi = 0.091 \), while \( H(\infty) = 1 \).
Also the polynomial \( 3 + 4x^2 - x^4 \) has the unique positive real root
\[ \sqrt{2 + \sqrt{7}} \approx 2.155 \] which is smaller than \( \sqrt{5} = 2.236 \). Thus for
\( x < \sqrt{2 + \sqrt{7}} \), both terms of the bracketed expression in (3.10) are positive, while for \( x > \sqrt{5} \), both are negative. Hence
\[
(3.11) \quad H'(x) > 0 \quad \text{for} \quad x < \sqrt{2 + \sqrt{7}}
\]
\[
(3.12) \quad H'(x) < 0 \quad \text{for} \quad x > \sqrt{5}.
\]

Since \( H(\infty) = 1 \), (3.12) implies that \( H(x) > 1 \) for \( x \geq \sqrt{5} \), and so, since
\( H(0) = \frac{1}{2} - 1/2 \pi < 1 \), \( H(x) = 1 \) for some \( x \in (0, \sqrt{5}) \). Also (3.11) implies that there is at most one root of \( H(x) = 1 \) in \([0, \sqrt{2 + \sqrt{7}}]\). Thus, if we show that \( H(x) > 1 \) in \((\sqrt{2 + \sqrt{7}}, \sqrt{5})\), we shall have established that \( H(x) = 1 \) has only one positive root. But, evaluating (3.9) we find that
\( H(2.16) \approx 1.122 \), while for \( x \in (\sqrt{2 + \sqrt{7}}, \sqrt{5}) \), from (3.10),
(3.13) \[ |H'(x)| < \frac{1}{2} \phi(\sqrt{5}) \{ \sqrt{5} (3-\sqrt{7}) \phi(\sqrt{5}) + (25-4(2+\sqrt{7})-3) \phi(\sqrt{5}) \} \]
\[ = 0.056 \]

so that, for \( x \in (\sqrt{2+\sqrt{7}}, \sqrt{5}) \)
\[ H(x) > 1.12 - 0.06(\sqrt{5} - \sqrt{2+\sqrt{7}}) \]
\[ = 1.115 \]
\[ > 1. \]

Thus \( H(x)=1 \) has a unique root \( x_0 \), which lies in \((0,\sqrt{2+\sqrt{7}})\). This root can be evaluated numerically, using Newton's method, giving

(3.14) \[ x_0 = 1.535. \]

Thus, for fixed \( \alpha \), \( \alpha_3(\ell,\alpha) \) attains its maximum value at \( \ell = \ell^* \), where

(3.15) \[ \ell^* = \left( \frac{a}{b x_0} \right)^2. \]

From (3.2) we see that the maximum value of \( \alpha_3 \) is

(3.16) \[ \alpha_3^*(\alpha) = \left( \frac{b}{a} \right)^2 x_0^2 (1-\rho(x_0)) = 0.736 \left( \frac{b}{a} \right)^2. \]

If we choose a particular value \( \alpha \) for \( \alpha_3 \) in advance, then the detection level \( \alpha \) must satisfy

(3.17) \[ \alpha^2 \leq \frac{b^2}{a^2} x_0^2 (1-\rho(x_0)) \Leftrightarrow 0.736 \left( \frac{b}{a} \right)^2 \]

and then, since \( \frac{\partial \alpha_3}{\partial \ell} \) is continuous, there are two detector lengths \( \ell \) giving \( \alpha_3(\ell,\alpha)=\alpha \), one shorter than \( \ell^* \) and one longer. For the longer detector, \( \lambda(t) \) has a large variance and so will spend much of the time above the level \( \alpha \). Thus a new "object" will be registered soon after the "dead period" so that the lengths of the intervals between detections will be dominated by the dead period. For the shorter detector, the interdetection intervals will be dominated by the "waiting time" for the next passage of \( \lambda(t) \) above \( \alpha \). Clearly the latter performance is preferable,
so that we shall be interested only in detectors shorter than
\( l^* = (a/bx_0)^2 = (1-\rho(x_0))/\alpha \approx 0.313/\alpha \). This implies that
\( a/b\sqrt{l} > x_0 \), so that
\( \alpha_1 = 1 - \Phi(a/b\sqrt{l}) < 1 - \Phi(x_0) \approx 0.062 \). Thus for
a detector to give reasonable \( \alpha_3 \)-performance we should choose
\( \alpha_1 < 0.062 \).

As in [A], we can link the behaviour of \( \alpha_1 \) and \( \alpha_3 \). In
fact, from (3.2) and (2.7) we see that
\[
\alpha_3 = \frac{1 - \rho \left( \frac{a}{b\sqrt{l}} \right) L}{l} = \frac{1 - \rho \left( \Phi^{-1}(1-\alpha_1) \right) L}{l}.
\]

If we vary \( a \) and \( l \), keeping \( \alpha_1 \) fixed, \( \alpha_3 \) decreases with \( l \), as did
\( \alpha_2 \) in the Poisson process case. Thus the "objects" detected tend to be
longer with the longer detector.

An alternative approach to optimality of a detector is to combine
\( \alpha_3 \) and \( \beta_2 \). If we choose a value of \( \alpha_3 \) which is greater than \( (1-\rho(x_0))/L \),
then \( l^* < L \), so that \( l < L \) for any detector satisfying \( l < l^* \). Recall that
for \( \alpha l < 1 \),
\[
\tilde{\alpha}(l,\alpha) = b\sqrt{l} \rho^{-1}(1-\alpha l)
\]
so that \( \alpha_3(l,\tilde{\alpha})=\alpha \). We consider the effect on \( \beta_2(a,l) \) of varying the
length \( l \), which will be kept less than \( l^* \). When \( L > l^* \), we have
\[
\beta_2 = \Phi(\Phi^{-1}(1-\alpha l) - s\sqrt{l}/b)
\]
which clearly decreases with \( l \), since \( \rho^{-1} \) and \( \Phi \) are both increasing.
Thus the optimal detector length will be \( l = l^* \).
When \( L < \ell^* \), the same argument shows that the optimal \( \ell \) is greater than \( L \). For \( \ell > L \), we have

\[
\beta_2 = \phi(\rho^{-1}(1-a\ell) - sL/b/\ell).
\]

Numerical calculations indicate that this expression is a unimodal function of \( \ell \), with the location of the mode depending on the values of \( a \) and \( sL/b \). Thus its minimum value on \([L, \ell^*]\) is attained for either \( \ell = L \) or \( \ell = \ell^* \), \( \ell^* \) being optimal for small values of \( sL/b \), while \( L \) is optimal when \( sL/b \) is large.
3.4 A "natural noise" model

An alternative model for the objects, which is more closely related to the model used in [A], is the following:

The background noise is again Gaussian white noise $dW(t)$. The objects consist of intervals of length $(1+p)L$ of this process which have been contracted to a length $L$ ($p > 0$). (cf. p.13 of [A]). Let $W(t)$ denote the uncontracted process and $W^*(t)$ denote the process after contraction. Then, for example, if there is only a single object located at $[T-L,T]$ $W^*$ will be given by

$$W^*(t) = W(t) \quad t \leq T-L$$
$$= W(t+p(t-T+L)) \quad T-L \leq t \leq T$$
$$= W(t+pL) \quad t \geq T.\tag{4.1}$$

The detector, of length $\ell$, registers $r(t) = b(W^*(t)-W^*(t-\ell))$, so that if $\ell \leq \ell$

$$r(t) = b(W(t)-W(t-\ell)) \sim N(0,b^2\ell) \quad t \leq T-L$$
$$= b(W(t+p(t-T+L))-W(t-\ell)) \sim N(0,b^2(1+p)(t-T+L)) \quad T-L \leq t \leq T-L+\ell$$
$$= b(W(t+p(t-T+L))-W(t-\ell+p(\ell-T+L))) \sim N(0,b^2(\ell+p(T-t+\ell))) \quad T-L+\ell \leq t \leq T$$
$$= b(W(t+pL)-W(t-\ell+p(\ell-T+L))) \sim N(0,b^2(\ell+p(T-t+\ell))) \quad T \leq t \leq T+\ell$$
$$= b(W(t+pL)-W(t-\ell+pL)) \sim N(0,b^2\ell) \quad t \geq T+\ell\tag{4.2}$$

where $\sim$ denotes "has the distribution", and $N(\mu,\sigma^2)$ represents the normal distribution with mean $\mu$ and variance $\sigma^2$.

Similarly, if $\ell \geq \ell$, the detector registers
In this model, it is the variance rather than the mean of the process which changes within the object. This can be considered as a limiting case of the point process model and might be more appropriate as a model for a "natural" object than Zakai and Ziv's (1969) model for a "man-made" signal. Since we are now interested in detecting regions of increased variability in the detector's output, the appropriate test method will be to register an object if the modulus of \( r(t) \) exceeds a predetermined level \( \alpha \). Thus we define

\[(4.4) \quad I(t) = 1 \text{ if } |r(t)| \geq \alpha \]
\[= 0 \text{ otherwise} .\]

The size measure \( a_1 \) could be redefined in terms of the two-tail probability, since

\[(4.5) \quad \lim_{t \to \infty} \frac{\int_t^\infty I(u)du}{t-L} = 2\left(1 - \Phi\left(\frac{\alpha}{b\sqrt{\lambda}}\right)\right) \text{ a.s. in the absence of objects,}
\]

but since this limit is simply twice that in (2.7), we shall continue to use \( a_1 = 1 - \Phi(a/b\sqrt{\lambda}) \) as our basic size measure. Again we define

\[\alpha_2(a) = a^2/(b^2 \Phi^{-1}(1-a)^2) \text{ as in (2.8).}\]

The power measure \( \beta_2 \) becomes \( \beta_2^* \) given by
and we wish to minimize

\[
\beta_2^\ast(a) = 2\Phi\left( \frac{a}{b\sqrt{1+p}l_a} \right) \quad a \leq a^\ast
\]

\[
= 2\Phi\left( \frac{a}{b\sqrt{l_a}+pL} \right) \quad a \geq a^\ast.
\]

But

\[
2\Phi\left( \frac{a}{b\sqrt{(1+p)l_a}} \right) = 2\Phi\left( \frac{1}{\sqrt{1+p}} \phi^{-1}(1-a) \right)
\]

is independent of \(a\), while

\[
2\Phi\left( \frac{a}{b\sqrt{l_a}+pL} \right) = 2\Phi\left( \frac{1}{\sqrt{\phi^{-1}(1-a)^2 + \frac{b^2pL}{a^2}}} \right)
\]

increases with \(a\).

Thus any value of \(a \leq a^\ast\), or \(l_a \leq L\) gives the same value of \(\beta_2^\ast\), and is optimal. This is analogous to the situation with a renewal process having a strictly stable interevent time distribution ([A] p.34).

3.5 Combining information from several receptors

As in Section 1.1 we can consider combining the information from a number of receptors for Gaussian objects. Recall from (2.7) and (2.9) that for the Zakai and Ziv "artificial" object, the size and power are given by

\[
a_1 = 1 - \phi(a/b\sqrt{l})
\]

\[
\beta_2 = \phi((a-s \min(l,L))/b\sqrt{l}).
\]
As in Section 1.1 we consider various ways of combining the outputs $N_k(\ell;t)$ of $r$ independent receptors. Again, the most natural combinations are 1) $\min_k N_k(\ell;t)$, 2) $\max_k N_k(\ell;t)$, 3) $\sum_k N_k(\ell;t)$ for which the size and power measures analogous to (5.1), (5.2) and (5.3) become the size and power of tests of $\mu = 0$ vs $\mu = -s\ell$ for normal random variables of mean $\mu$ and variance $\sigma^2$ based on a sample of $r$ independent observations $X_1, \ldots, X_r$ and $\min_k X_k$, $\max_k X_k$ and $\sum_k X_k$ as test statistics. Since $\sum_k X_k$ is sufficient for $\mu$, a test based on $\sum_k X_k$ will be most powerful of its size, and hence we see that the procedure based on $\sum_k N_k(\ell;t)$ will give the optimal detection performance. Note that this result is exact, in contrast to the approximate result obtained in Section 1.1.
4. References


APPENDIX A

LOCATING BRIGHT SPOTS IN A POINT PROCESS

I.W. Saunders

CSIRO Division of Mathematics and Statistics
Canberra
Abstract

As an approach to modelling the "matching" of optical receptors in animals to the objects they are designed to see, we study the problem of locating regions of high intensity in a point process on the real line, using the counts of points in a movable interval of fixed length. We define performance measures analogous to statistical size and power for this procedure and, for points forming a renewal process, give conditions on the quantiles of the convolutions of the interpoint distribution which ensure that the optimal length for the "detector" is close to that of the "object" to be detected. We show that these conditions are satisfied for a Poisson process. Similar conditions ensure that the optimal length is close to zero, and we give a class of distributions satisfying these conditions. Finally we show that the results can be extended to simple two-dimensional models.

Key words:

LOCATION; QUANTILES; RECEPTOR; RENEWAL PROCESS.
1. Introduction

The work described in this paper was suggested by the following biological problem: suppose an animal is interested in seeing a particular kind of object - perhaps its food, or a predator - what sort of optical receptors should its eye contain? Or given the output from an array of receptors, how should the animal's brain interpret the information?

It is intuitively reasonable that the best receptor would match the object's retinal image in size and shape, and in fact "feature extractors", which have been found in the visual systems of monkeys and cats, ignore uniform illumination, but "fire" when they "see" objects which match their shape. (Julesz (1975)).

The aim of this paper is to investigate a simple model for the detection process to see whether this intuitive idea of "matching" does give the best detector. We shall suppose that the photons registered by the receptor can be considered as an array of dots (Fig. 1). In practice the photons would be arriving at different times, but we shall neglect this, and suppose that the receptor is only registering the positions of the photons it sees, and not their time of arrival. An "object" is a region of increased photon intensity, and the receptor is to be used to locate the object as accurately as possible. Clearly this model does not include all the complexity of a real eye; in particular it does not allow for quantum effects. However, we may hope that it will give some indication of the behaviour that a real eye might exhibit.

Even the simplified problem illustrated in Fig. 1 allows for considerable complexity since the sizes and shapes of the object and receptor are capable of wide variation. As a first step, therefore,
we will simplify the problem still further and consider a one-dimensional version, where the "photons" lie along a line. Now the only possible shape for a connected object or receptor is an interval, and we have only to choose the best length for the receptor. We shall see later that solving this one-dimensional problem allows us to solve the two-dimensional version when the receptor and object shapes are suitably restricted.

The one-dimensional problem itself has applications in other fields. The dimension can be taken as time, rather than space, when the length of the interval becomes a "memory span". Thus the methods of this paper could be used to determine the optimal duration of memory for detecting a "signal" of a given duration. Another application, considered by Cressie (1977), is in the detection of radioactive ore bodies by aerial surveys using geiger counters. In Section 2 below we describe our model in more detail, and compare our approach with Cressie's.

In Section 3 we shall discuss the problem of choosing the optimal length for a detector or receptor. We consider various measures of "size" and "power" for the detection process; of these we choose a pair which have fairly simple mathematical properties, and which are also physically reasonable performance measures. We derive simple expressions for these measures in the case where the "photons" form a renewal process.

In Section 4, we give conditions on the quantiles of the interevent distribution of a renewal process which ensure that the optimal detector matches the length of the object (Theorem 1) or is close to zero (Theorem 2). We also show that any distribution satisfying the conditions of Theorem 2 must have infinite mean, and so is unlikely to arise in practice. In Section 5 we consider these
quantile properties in more detail, and show that the exponential
distribution satisfies the conditions of Theorem 1, so that the
optimal detector for an object in a Poisson process matches the
object length. Finally in Section 6 we show that our results
allow for some extension to two dimensions.


Our first task is to describe precisely what we are going
to mean by "photon", "object", "detection", etc. Having decided
on suitable definitions, we can proceed to consider questions of
measures of performance, or optimality, and attempt to devise
optimal procedures.

First, we shall assume that the photons (which in other applications
could be peaks on a graph, clicks of a geiger counter, etc.) can be
represented by points, neglecting any quantum effects. We shall
suppose further that the arrangement of these points can be described
by a stochastic point process on the real line \( \mathbb{R} \). (We shall be
principally concerned with points forming a renewal process or, as
a special case of this, a stationary Poisson process.)

We shall suppose that the detector or receptor acts as follows:
it occupies a line segment of length \( \ell \) and moves along the line \( \mathbb{R} \nach the positive direction starting from \((0,\ell]\) registering the number
of photons it contains: Thus we suppose that when it is occupying
\((t-\ell,t]\) it registers \(N(\ell;t) = \text{number of points in } (t-\ell,t]\). The detector
will signal the presence of an object when it "sees" a bright region. Thus
we assume that there is a threshold level \(a\), such that the detector signals
the presence of an object whenever it sees a number of photons $N(\ell;t) \geq a$. We define the indicator function $I(t)$ by

$$I(t) = \begin{cases} 1 & \text{if } N(\ell;t) \geq a \\ 0 & \text{if } N(\ell;t) < a , \end{cases}$$

so that $I(t)$ takes the value one when the detector is signalling an object, and is zero otherwise. We shall be interested in the properties of $I(t)$ and its usefulness for detecting regions of high photon intensity ("objects").

Cressie [1977] has considered a similar problem. He supposes that the "photons" form a Poisson process of constant rate $\lambda$ on the finite interval $[0,T]$, and that there may be an "object" consisting of an interval of higher intensity $\lambda + \mu$ somewhere within $[0,T]$. He wishes to test the hypothesis than an object is present against the null hypothesis that the rate is uniform. Instead of the indicator function $I(t)$, he uses the "scan statistic"

$$N(\ell) = \sup_{\ell \leq T \leq T} N(\ell;T)$$

where $N(\ell;T)$ is defined as above. This statistic is useful for the detection (rather than location) of a single object - the situation that Cressie is interested in. However, in our context, where the location of potentially many objects is of interest, the procedure using $I(t)$ is more appropriate.

In some applications it may be desirable to reduce the data storage requirements by measuring $I(t)$ only at the discrete points $t$ where photons are located. Thus if the points of the process are $\{\tau_k\}_{k=1}^{\infty}$, only $\{I(\tau_k)\}_{k=1}^{\infty}$ would be recorded. We shall consider this
situation, which we shall call "the discrete case", as well as the situation where I(t) is known for all t - "the continuous case". In the discrete case, we shall not need the stationarity assumption made for the continuous case.

3. Criteria for the Performance of a Detector.

For the one-dimensional detector described above, there are two parameters under our control. We can choose the length \( \ell \) of the interval, and also the "detection level" \( a \) at which the detector will register the presence of an object. We want to choose the values of \( \ell \) and \( a \) which will optimize, in some sense, the performance of the detector.

There are two aspects of the detector's behaviour that we must consider, which are analogous to Type 1 and Type 2 errors for a statistical test. The first is the frequency with which the detector registers objects when in fact none are present - we shall refer to this as the "size" of the detector. The second is the likelihood of correctly locating an object which is present - we shall call this the "power" of the detector. We shall not define size or power more precisely than this, but consider various different measures of them as criteria for optimality.

3a. Measures of Size.

The size of a detector refers to its performance in the absence of objects when only background noise is present. Thus in this Section we shall suppose that the photons received form a stationary point process. The points of this process will be denoted by \( \{ \tau_k \}_{k=0}^{\infty} \) where \( \tau_0 \) is the first point of the process to the right of the origin. We shall suppose that the detector starts from the position \( (0,\ell) \) so that we know \( I(t) \) for \( t \geq \ell \).
We first note that as \( t \) increases, \( I(t) \) changes from 0 to 1 at points \( \tau_k \) such that \( \tau_k - a \geq \tau_k - a' \) and changes from 1 to 0 at points \( \tau_k + a' \) such that \( \tau_k + a' \geq \tau_k - a' \). Thus, provided the process \( \{\tau_k\} \) has no finite accumulation points, the set \( \{t|I(t)=1\} \cap [\ell, T] \) is the union of a finite number of disjoint intervals (as in Fig. 2). We shall assume that this is the case.

There are two natural ways of measuring a detector's size. We can consider either the total length of the "objects" registered \( \int_{\ell}^{T} I(t) dt \), or the number of separate "objects", i.e. the number of times the value of \( I(t) \) changes from 0 to 1 (or this number plus 1 if \( I(\ell) = 1 \)). Thus we will consider the limiting behaviour as \( T \to \infty \) of

\[
\alpha_1(T) = \frac{1}{T} \int_{\ell}^{T} I(t) dt
\]

and

\[
\alpha_2(T) = \frac{1}{T} \times \text{number of 0-1 jumps of } I(t) \text{ in } [\ell, T].
\]

We shall show that when \( \{\tau_k\} \) is a renewal process, both \( \alpha_1(T) \) and \( \alpha_2(T) \) converge almost surely to constant limits, and we shall use these limits as our measures of size.

If the \( \{\tau_k\} \) form a renewal process, then we can write for \( k \geq 0 \)

\[
\tau_k = X_0 + X_1 + \ldots + X_k, \text{ where } X_0, X_1, \ldots \text{ are independent, and } X_1, X_2, \ldots
\]
Effect of A-value on Detector Output (1)

Figure 2
are identically distributed on \((0, \infty)\) with distribution function 
\(F(x)\), and mean \(\mu = E[X]\) which for the present we shall assume to 
be finite, while \(X_0\) has distribution function 
\(F_0(x) = \frac{1}{\mu} \int_0^x (1-F(u))du\) 
so that the sequence \(\{\tau_k\}\) is stationary. [Feller (1971)]. \(F_0(x)\) is 
then the distribution of the forward recurrence interval from any 
point \(t\), i.e. the distance between \(t\) and the next point of the process 
after \(t\).

Let \(F * G\) denote the convolution of the distribution functions 
\(F\) and \(G\), and let \(F^{(n)}\) denote the \(n\)-fold convolution of \(F\) with itself. 
Let \(F(X, Y, Z, \ldots)\) denote the \(\sigma\)-field generated by the random variables 
\(X, Y, Z, \ldots\).

**Lemma 1.** For a stationary renewal process \(\{\tau_k\}\), as \(T \to \infty\),

\[(3.3) \quad \alpha_1(T) \xrightarrow{a.s.} \alpha_1 = F_0 * F(a-1)(x),\]

\[(3.4) \quad \alpha_2(T) \xrightarrow{a.s.} \alpha_2 = (F(a-1)(x) - F(a)(x)) / \mu.\]

**Proof:** For any set \(C \subseteq [0, 1]\) and any integer \(n \geq 0\) the event 
\(A(C) = \{\alpha_1(T) + c \in C\}
\[
= \{\frac{1}{T} \int_T^T I(t)dt + c \in C\}
\]
is identical with the event 
\(A_n(C) = \{\frac{1}{T} \int_0^{\frac{T}{n+1}} I(t)dt + c \in C\}
\[
\leq \frac{T}{n+1} + \int_0^{T/n} I(t)dt + c \in C\}
\]
since \(\frac{1}{T} \int_0^{T/n} I(t)dt = c \to 0\) as \(T \to \infty\). But \(A_n(C)\) is in 
\(F(X, Y, Z, \ldots)\). Hence \(A(C)\) is in \(F(X, Y, Z, \ldots)\) for all \(n\), and so in the 
tail \(\sigma\)-field \(\mathcal{T}(X_0, X_1, \ldots) = \bigcap_{n=0}^{\infty} F(X_n, X_{n+1}, \ldots)\). But \(X_0, X_1, \ldots\) is a 
sequence of independent random variables, and so any tail event has 
probability zero or one (Breiman 1968, p. 40). Thus \(\text{Prob}(A(C)) = 0\) or 1.
But since \( I(t) \) is itself stationary we have

\[
\alpha_1(T) \overset{a.s.}{\to} E(I(t) \mid T^*)
\]

where \( T^* \) is the invariant \( \sigma \)-field of \( I(t) \) (Doob (1953) p. 515). Thus \( E(I(t) \mid T^*) \) is a.s. constant and so must equal \( EI(t) \) and

\[
EI(t) = \text{Prob} \left( I(t) = 1 \right)
\]

\[
= \text{Prob} \left( I(\ell) = 1 \right)
\]

\[
= \text{Prob} \left( \tau_{a-1} \leq \ell \right)
\]

\[
= \text{Prob} \left( X_0 + X_1 + \ldots + X_{a-1} \leq \ell \right)
\]

\[
= F_0 \ast F(a-1)(\ell)
\]

establishing (3.3).

Let \( N_T \) denote the number of points of the process in \([0,T]\). Then \( N_T / T \to [EX_1]^{-1} \) a.s. as \( T \to \infty \) (Breiman (1968) p. 220), and \( N_T \to \infty \) as \( T \to \infty \). Let \( M_T \) denote the number of points \( \tau_k \) in \([\ell,T]\) at which \( \tau_{k-a+1} > \tau_k - \ell \geq \tau_{k-a} \), i.e. the number of points at which \( I(t) \) changes from 0 to 1. Define

\[
I_k^* = \begin{cases} 
1 & \text{if } \tau_{k-a+1} > \tau_k - \ell \geq \tau_{k-a} \\
0 & \text{otherwise}
\end{cases}
\]

so that

\[
M_T = \sum_{N_T}^{\tau_k^*} I_k^*.
\]

The event \( \{I_k^* = 1\} \) is the same as the event

\[
\{X_{k-a+2} + \ldots + X_k < \ell \leq X_{k-a+1} + X_{k-a+2} + \ldots + X_k\}
\]

and so depends only on \( X_{k-a+1}, \ldots, X_k \). Thus the sequence \( \{I_k^*\} \)
is $a$-dependent, i.e. $I_k^*$ and $I_{k+a+1}^*$ are independent for all $k$, and (Ibragimov and Linnik (1971) Chapter 17) the sequence $\{I_k^*\}$ is strong mixing, and so metrically transitive. Hence as $T \to \infty$,

$$
\frac{M_T}{N_T} = \frac{\sum I_k^{*}}{N_k+1} \xrightarrow{a.s.} E[I_k^*]
$$

$$
= \text{Prob} \{X_{k-a+2}^{*} + \ldots + X_k^{*} < \ell < X_{k-a+1}^{*} + \ldots + X_k^{*}\}
$$

$$
= F(a-1)(\ell) - F(a)(\ell)
$$

and so

$$
\frac{M_T}{T} \xrightarrow{a.s.} \frac{F(a-1)(\ell) - F(a)(\ell)}{E[X_1^*]}
$$

Similar results can be obtained for the discrete case.

Define $\alpha_1'(n)$ and $\alpha_2'(n)$ by

\begin{equation}
\alpha_1'(n) = \frac{\sum I(\tau_k)}{N_k+1} 
\end{equation}

\begin{equation}
\alpha_2'(n) = \frac{\sum (I(\tau_k) - I(\tau_{k-1}))^+}{n}
\end{equation}

where $x^+ = \max(x, 0)$, so that $\alpha_1'(n)$ is the proportion of points $\tau_k$ of the process at which $I(\tau_k) = 1$, while $\alpha_2'(n)$ is the proportion of
points $\tau_k$ at which $I(\tau_k) = 1$ and $I(\tau_{k-1}) = 0$. Using methods similar to those of Lemma 1, we can show that, as $n \to \infty$,

$$\alpha_1(n) \xrightarrow{\text{a.s.}} \alpha_1 = EI(\tau_k)$$

$$\alpha_2(n) \xrightarrow{\text{a.s.}} \alpha_2 = E((I(\tau_k) - I(\tau_{k-1}))^+)$$

Also

$$EI(\tau_k) = \text{Prob} (X_k + x_{k+1} + \ldots + x_k \leq \ell)$$

$$= F(a-1)(\ell)$$

while

$$E((I(\tau_k) - I(\tau_{k-1}))^+) = \text{Prob} (X_k + x_{k+1} + \ldots + x_k \leq \ell, x_{k-a} + x_{k-a+1} + \ldots + x_k > \ell)$$

$$= \text{Prob} (X_1 + \ldots + x_{k-a} \leq \ell, x_2 + \ldots + x_a > \ell)$$

$$= \int_0^\ell \text{Prob} (x_1 \leq \ell - u, x_a > \ell - u | x_2 + \ldots + x_{a-1} = u) \, dF(a-2)(u)$$

$$= \int_0^\ell F(\ell-u)(1-F(\ell-u)) \, dF(a-2)(u)$$

so that we have proved

**Lemma 1'**: For a renewal process $\{\tau_k\}$ (not necessarily stationary), as $n \to \infty$,

$$\alpha_1(n) \xrightarrow{\text{a.s.}} \alpha_1 = F(a-1)(\ell), \quad \text{(3.7)}$$

$$\alpha_2(n) \xrightarrow{\text{a.s.}} \alpha_2 = \int_0^\ell F(\ell-u)(1-F(\ell-u)) \, dF(a-2)(u). \quad \text{(3.8)}$$

Writing $P(x,n)$ for the Poisson distribution function

$$P(x,n) = \sum_{j=0}^{n-x} \frac{e^{-\lambda} \lambda^j}{j!} \quad \text{(3.9)}$$

we find the following results for the special case where $\{\tau_k\}$ is a stationary Poisson process of rate $\lambda$. 


(3.10) \[ \alpha_1 = 1 - P(\lambda; a-1) \]

(3.11) \[ \alpha_2 = \lambda^{-1} p(\lambda; a-1) \]

(3.12) \[ \alpha_1' = 1 - P(\lambda; a-2) \]

(3.13) \[ \alpha_2' = (-1)^{a-2} e^{-2\lambda} \left\{ \sum_{j=0}^{a-2} \frac{(-\lambda)^j}{j!} e^{-\lambda} \right\} \]

(A simple inductive proof shows that the bracketed term in (3.13) has the same sign as \((-1)^{a-2}\), so that \(\alpha_2'\) is always positive.)

If the distribution function \(F\) has finite \((2+\delta)\)th moments, then for each of the quantities \(\alpha_1, \alpha_2, \alpha_1', \alpha_2'\) we can establish a central limit theorem (CLT), showing that the rates of convergence in Lemmata 1, 1' are \(O(n^{-\frac{1}{2}})\), \(O(n^{-\frac{1}{2}})\) respectively. The proofs all use Theorem 18.5.1 of Ibragimov and Linnik (1971) which implies that any a-dependent stationary process with finite \((2+\delta)\)th moments satisfies a CLT. We give details only for \(\alpha_1\):

Clearly,

(3.14) \[ \int_{T_k}^{T_{k+1}} I(t) dt = \sum_{k=N_k}^{N} \int_{T_k}^{T_{k+1}} I(t) dt + O(1) \]

while the sequence \(\int_{T_k}^{T_{k+1}} I(t) dt\) is a-dependent and

\[ E(\int_{T_k}^{T_{k+1}} I(t) dt)^{2+\delta} \leq E(\left| X_1 \right|^{2+\delta}) < \infty. \]

Thus the sum and hence the integral in (3.14) satisfies a CLT.

The measures \(\alpha_2, \alpha_2'\), while they represent an important aspect of the detector's performance, are somewhat ambiguous when considered on their own. Since they essentially count the number of times
N(\ell;t) or N(\ell;T_k) crosses the level \(a\), they do not distinguish between a process which spends most of its time above \(a\) and one which spends most of its time below \(a\). Thus the values of \(\alpha_2(T)\) for the two realizations depicted in Fig. 2 would be the same - \(3/T\) - although it is clear that \(2a\) represents a much more desirable performance than \(2b\). Therefore we shall concentrate on \(\alpha_1\) as a criterion for designing the detector, using \(\alpha_2\) as a description of the resulting performance.

**3b. Measures of Power.**

The power of a detector refers to its performance in the presence of objects. We shall suppose that objects consist of intervals of constant length \(L\) in which the intensity of the point process is greater than its intensity elsewhere. We shall also assume that the objects are positioned sufficiently far apart that we need consider only one object at a time.

Suppose that an object occupies the interval \([T-L,T]\). What should we mean by successfully "locating" the object with our detector? Clearly we must require that \(I(t) = 1\) in some region \(J\) near the object. A suitable region can be defined in two ways. The first requires that the object is completely included in the detector if \(L \leq \ell\) or vice versa if \(L \geq \ell\). The second requires only that the object and detector have a nonempty overlap. We thus define two "detection regions"

\[
(3.15) \quad J_1 = [T, T-L+\ell] \quad \text{for } L \leq \ell \\
= [T-L+\ell, T] \quad \text{for } L \geq \ell
\]

and

\[
(3.16) \quad J_2 = [T-L, T+\ell]
\]
and consider the object to have been successfully located if
\( I(t) = 1 \) for some \( t \in J \), where \( J \) is whichever of \( J_1, J_2 \) we have
decided to use.

As before we shall assume that the background noise forms a
stationary renewal process. The objects could be modelled in a
number of ways, but the following seems to be reasonably
realistic, and proves to be mathematically tractable:

We suppose that the object at \([T-L, T]\) is formed as follows.
The noise process is constructed along the whole line, after which
the interval \([T-L, T+\rho L]\) where \( \rho > 0 \), is contracted, together with
the points it contains, to a length \( L \), so that it now occupies
\([T-L, T]\). The points to the right of the interval are shifted back a
distance \( \rho L \). Thus the object is a scale-changed portion of the
background noise process.

In the special case where the photons form a Poisson process,
this rather complicated procedure is equivalent to changing the rate of the
process within \([T-L, T]\) from the background level \( \lambda \) to a level \((1+\rho)\lambda\).
Thus we can consider the object to be a section of a Poisson process
of rate \( \rho \lambda \) superimposed on the background process. In general,
however, the superposition of two renewal processes is not a renewal
process, and using a superimposed process to represent an object
complicates the formulae obtained.

The most natural definition of power for this detection method is

\[
\beta_1 = \text{Prob} \{ I(t) = 1 \text{ for some } t \in J \}
\]

There appears to be no simple expression for \( \beta_1 \) in general,
although for a Poisson process of rate \( \lambda \) and using \( J = J_1 \) we can
derive the formulae:
\begin{equation}
\beta_1 = 1 - \sum_{m=0}^{a-1} \left\{ p(\lambda(1+\rho)L;m) \left[ p(\lambda(L-L);a-m-1) \right]^2 - p(\lambda(L-L);a-m) \sum_{n=0}^{a-m-2} p(\lambda(L-L);n) \right\} \quad \text{for } \ell \geq L,
\end{equation}

\begin{equation}
\beta_1 = 1 - \sum_{m=0}^{a-1} \left\{ p(\lambda(1+\rho)(2\ell-L);m) \left[ p(\lambda(1+\rho)(L-\ell);a-m-1) \right]^2 - p(\lambda(1+\rho)(L-\ell);a-m) \sum_{n=0}^{a-m-2} p(\lambda(1+\rho)(L-\ell);n) \right\} \quad \text{for } L \leq \ell \leq L/2,
\end{equation}

where, as above, \( p(x;j) = e^{-X} x^j / j! \), and \( P(x;n) = \sum_{j=0}^n p(x;j) \).

The derivation of these expressions is deferred to the end of this Section.

Apart from its mathematical complexity, \( \beta_1 \) as a measure of power suffers from disadvantages similar to those of \( \alpha_2 \) as a measure of size. For example, a detector of length zero and detection level \( a = 1 \) will locate virtually every object, missing only those which "emit no photons", while its \( \alpha_1 \)-value will be zero. It would not, however, be a very suitable detector.

This unsatisfactory behaviour of \( \beta_1 \) is illustrated in Table 1, which compares the performance of various detectors for locating an object in a Poisson process. The background noise has rate \( \lambda = 10 \), while the object is of length \( L = 1 \), with \( \rho = 1 \), so that the rate of the process within the object is 20. The detectors have detection levels \( a = 2, 3, \ldots, 20 \) and lengths \( \ell_a \) chosen to give an \( \alpha_1 \)-value of 0.05 for each detector. Values of \( 1 - \beta_1 \) calculated from formulae (3.18), (3.19) are given where \( \ell_a > L/2 = 0.5 \), and values estimated from 1000 simulations for each detector are given where \( \ell_a < L = 1 \). We see that, as suggested by the above example, very short detectors give the highest \( \beta_1 \) values.
<table>
<thead>
<tr>
<th>section level $a$</th>
<th>Detector length $l$</th>
<th>$\alpha_1$</th>
<th>$1-\beta_1$ from (3.18), (3.19)</th>
<th>$1-\beta_1$ from simulations</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>.035536</td>
<td>.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.081769</td>
<td>.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>.136632</td>
<td>.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.197015</td>
<td>.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>.261301</td>
<td>.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>.328532</td>
<td>.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>.398082</td>
<td>.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>.469523</td>
<td>.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>.542541</td>
<td>.05</td>
<td>.080174</td>
<td>.079</td>
</tr>
<tr>
<td>11</td>
<td>.616901</td>
<td>.05</td>
<td>.089201</td>
<td>.082</td>
</tr>
<tr>
<td>12</td>
<td>.692421</td>
<td>.05</td>
<td>.096621</td>
<td>.094</td>
</tr>
<tr>
<td>13</td>
<td>.768958</td>
<td>.05</td>
<td>.103645</td>
<td>.101</td>
</tr>
<tr>
<td>14</td>
<td>.846394</td>
<td>.05</td>
<td>.111838</td>
<td>.114</td>
</tr>
<tr>
<td>15</td>
<td>.924633</td>
<td>.05</td>
<td>.124454</td>
<td>.148</td>
</tr>
<tr>
<td>16</td>
<td>1.003596</td>
<td>.05</td>
<td>.152853</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>1.083214</td>
<td>.05</td>
<td>.139368</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>1.163430</td>
<td>.05</td>
<td>.134520</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>1.244195</td>
<td>.05</td>
<td>.133081</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>1.325465</td>
<td>.05</td>
<td>.133363</td>
<td></td>
</tr>
</tbody>
</table>

Values of $1-\beta_1$ for various detectors when $\lambda=10$, $\rho=1$, $L=1$. 
To overcome the problems associated with $\beta_1$, we propose an alternative measure of power, $\beta_2$, which is analogous to $\alpha_1$ above:

\[
\beta_2 = \frac{\int_J E(t) \, dt}{\int_J dt} = \frac{\int_J E(t) I(t) \, dt}{\int_J dt}.
\]

To evaluate $\beta_2$ we need to find $E(t)$ for $t \in J$. When $J = J_1$, $I(t)$ is stationary within $J$ so that $E(t)$ is constant and $\beta_2 = E(t)$ for any $t \in J$. Thus for $L \geq l$, when $J_1$ corresponds to the detector's lying completely within the object, we see that

\[
\beta_2 = F_o \ast F_{(a-1)}((1+\rho)l).
\]

while for $L \leq l$, when $J_1$ corresponds to the object's being completely covered by the detector,

\[
\beta_2 = F_o \ast F_{(a-1)}(l+\rho L).
\]

Combining these two formulae, we see that, for detection interval $J_1$,

\[
\beta_2 = F_o \ast F_{(a-1)}(l+\rho \min(l, L)).
\]

For detection interval $J_2$, $I(t)$ is no longer stationary. In fact we have, for $L > l$

\[
E(t-L+\tau) = F_o \ast F_{(a-1)}((1+\rho)\tau + (l-\tau)) \quad 0 \leq \tau \leq l
\]

\[
= F_o \ast F_{(a-1)}((1+\rho)l) \quad l \leq \tau \leq L
\]

\[
= F_o \ast F_{(a-1)}((\tau-L) + (1+\rho)(l-L)) \quad L \leq \tau \leq L + l
\]
and from (3.25) and similar expressions for the case $L \leq t$, we can evaluate $\beta_2$ for detection interval $J_2$ as

\begin{equation}
\beta_2 = \frac{|L-t|}{L+t} \frac{F_0 * F(a-1)}{F_0 * F(a-1)} (t + \rho \min(t,L)) \\
+ \frac{2}{\rho (L+t)} \{ (t + \rho \min(t,L)) F_0 * F(a-1) (t + \rho \min(t,L)) \\
- t F_0 * F(a-1) t - \int_t t F_0 * F(a-1) v \, dF_0 * F(a-1) v \}.
\end{equation}

For the Poisson process of rate $\lambda$ described above we find, for $J = J_1$,

\begin{equation}
\beta_2 = 1 - P(\lambda t + \rho \min(t,L), a-1)
\end{equation}

while for $J = J_2$,

\begin{equation}
\beta_2 = 1 - (L+t)^{-1} \frac{F_0 * F(a-1)}{|L-t|} (\lambda t + \rho \min(t,L); a-1) \\
+ \frac{2}{\rho \lambda} \sum_{j=0}^{a-1} \{ P(\lambda t, j) - P(\lambda t + \rho \min(t,L), j) \}.
\end{equation}

In the discrete case, the measure of power $\beta'_1$ corresponding to $\beta_1$ above,

\begin{equation}
\beta'_1 = \text{Prob}(I(\tau_k) = 1 \text{ for some } \tau_k \in J)
\end{equation}

suffers from the same disadvantages as $\beta_1$. The natural analogue of $\beta_2$ is
where the denominator is the expected number of points \( \tau_k \) lying in \( J \).

For the detection interval \( J_1 \), we can evaluate \( \beta'_2 \) as follows. Consider the numerator in (3.30) as a function \( H(x) \) of the length \( x \) of \( J_1 \) for fixed \( \lambda \). Suppose \( \lambda < L \). Then \( H(x+\Delta) = H(x) + H(\Delta) \) so that \( H(x) \sim x \). Thus

\[
H(x) = x \lim_{\Delta \to 0} H(\Delta)/\Delta \\
= x \lim_{\Delta \to 0} \left[ F(\Delta) F^{(a-1)}((1+\rho)\lambda) + o(\Delta) \right]/\Delta \\
= \frac{x}{\mu} F^{(a-1)}((1+\rho)\lambda) \quad \text{since } F \text{ has no mass at the origin.}
\]

Similarly, we find that the denominator in (3.30) is \( x/\mu \), and using a similar argument for the case \( \lambda > L \) we find that, for detection interval \( J_1 \),

\[
\beta'_2 = F^{(a-1)}(\lambda + \rho \min(\lambda,L)) .
\]

Using similar methods we find that when detection interval \( J_2 \) is used:

\[
\beta'_2 = (L + \lambda + \rho L)^{-1} \left[ [L-\lambda] F^{(a-1)}(\lambda + \rho \min(\lambda,L)) \\
+ \frac{2}{\rho} \int_{\lambda}^{\lambda + \rho \min(\lambda,L)} F^{(a-1)}(v) \, dv \right] .
\]

For the Poisson process, with detection interval \( J_1 \) we find that
\[ \beta'_2 = 1 - P(\lambda l + \rho \min(l,L);a-2) \]

while for detection interval \( J_2 \),

\[
\beta'_2 = 1 - (L + l + \rho L)^{-1} \left[ |L-l| \ P(\lambda l + \rho \min(l,L);a-2) \right. \\
+ \left. 2 \ \frac{1+\rho}{\lambda \rho} \sum_{j=0}^{a-1} (P(\lambda l;j) - P(\lambda l + \rho \min(l,L);j)) \right]
\]

Comparing (3.24) and (3.32) we see that the results for \( \beta'_2 \) in the discrete and continuous cases are very similar, differing only in the term \( F_o \) in (3.24). In the case of the Poisson process, when \( F_o = F \), the values of \( \alpha'_1 \) and \( \beta'_2 \) are identical with those of \( \alpha'_1 \) and \( \beta_2 \) for a detection level of \( (a+1) \). For other distribution functions \( F \), the values will not be the same but it seems reasonable to suppose that, for moderately large values of \( a \), \( F_o \ast F^{(a-1)} \) and \( F^{(a)} \) will not be too different. The expressions for \( \alpha'_1 \) and \( \beta'_2 \) are more easily tractable than \( \alpha_1 \) and \( \beta_2 \), since they do not involve \( F_o \) and so we shall use these as our basic measures of size and power. Thus our conclusions will be exact for the discrete case and, when \( a \) is large, approximately true for the continuous case. For a Poisson process, the results will be exact in both cases.

We conclude this Section by giving the derivations of formulae (3.18), (3.19) for \( \alpha_1 \) in the case of a Poisson process. The argument used is similar to that of Shepp (1971) who was considering a Wiener process rather than a Poisson process.

We shall require the following result, due to Karlin and McGregor (1959).
Theorem A. If $N_1(t),\ldots,N_n(t)$ are $n$ independent birth-and-death processes, having identical birth rates and identical death rates, and $P_{ij}(t) = \text{Prob}(N_1(t) = j | N_1(0) = i)$ then

$$\text{Prob}(N_k(t) = j_k, k = 1,\ldots,n, \text{ and } N_r(T) \neq N_s(T) \text{ for } r \neq s, 0 \leq T \leq t | N_k(0) = i_k, k = 1,\ldots,n) = D$$

where $D$ is the determinant

$$\begin{vmatrix}
    P_{i1j1}(t) & \cdots & P_{i1jn}(t) \\
    \vdots & \ddots & \vdots \\
    P_{inj1}(t) & \cdots & P_{ijnj}(t)
\end{vmatrix}.$$ 

Thus $D$ gives the probability that, starting from states $\{i_k\}$, the $n$ processes reach states $\{j_k\}$ at time $t$ without any two of them having become coincident in the intervening time interval. We shall only need to apply Theorem A in the case $n = 2$ and where the processes are Poisson, so that the birth rate is constant and the death rate is zero.

It follows immediately from Theorem A that

$$(3.35) \quad \text{Prob}(\text{No coincidences in } [0,t] | N_k(0) = i_k, N_k(t) = j_k, k = 1,\ldots,n) = D/(P_{i1j1}(t) \cdots P_{ijnj}(t)).$$

Now consider the detection process for $\ell > L$ when the points form a Poisson process of rate $\lambda$. Suppose without loss of generality that the object occupies the interval $(\ell-L,\ell]$. For $T \in [0,\ell-L]$, let $N_1(T)$
be the number of points of the process in \((0, \tau]\) and let \(N_2(\tau)\) be the number of points of the process in \((\ell, \ell+\tau]\). Also let \(N_0\) be the number of points of the process in \((0, \ell]\). Then for \(t \in J_1 = [\ell, 2\ell-L]\), we have

\[(3.36) \quad N(\ell, t) = N_2(\tau) + N_0 - N_1(\tau) \quad \text{where} \quad \tau = t - \ell,
\]

and so

\[(3.37) \quad \text{Prob}(I(t) = 0 \text{ for all } t \in J_1) = \text{Prob}(N_2(\tau) + N_0 - N_1(\tau) < a \quad \text{for all } \tau \in [0, \ell-L])
\]

\[= \sum_{i,j,k} \text{Prob}(N_2(\tau) - N_1(\tau) < a-j, \forall \tau \in [0, \ell-L]| N_1(\ell-L) = i, N_0 = j, N_2(\ell-L) = k) \times \text{Prob}(N_1(\ell-L) = i, N_0 = j, N_2(\ell-L) = k)
\]

Now, using (3.35), and noting that \(N_1(0) = 0, N_2(0) = 0\), we see that

\[(3.38) \quad \text{Prob}(N_2(\tau) - N_1(\tau) < a-j \forall \tau | N_1(\ell-L) = i, N_0 = j, N_2(\ell-L) = k)
\]

\[= \begin{vmatrix}
\pi(0, i) & \pi(0, j+k-a) \\
\pi(j-a, i) & \pi(j-a, j+k-a)
\end{vmatrix}
\]

\[\pi(0, i) \pi(j-a, j+k-a)
\]

provided \(j < a, i > j+k-a
\]

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

where \(\pi(i, j) = p(\lambda(\ell-L); j-i)\).

Also, since \(N_1(\ell-L), N_0 - N_1(\ell-L), N_2(\ell-L)\) are independent Poisson random variables with means \(\lambda(\ell-L), (1+p)\lambda L, \lambda(\ell-L)\) respectively,

\[(3.39) \quad \text{Prob}(N_1(\ell-L) = i, N_0 = j, N_2(\ell-L) = k)
\]

\[= p(\lambda(\ell-L); i) \quad p((1+p)\lambda L; j-i) \quad p(\lambda(\ell-L); k).
\]
Substituting (3.38) and (3.39) into (3.37) we obtain

\begin{equation}
\text{Prob}(I(t) = 0 \text{ for all } t \in J_j) \quad (3.40)
\end{equation}

\[
\sum_{j=0}^{a-1} \sum_{i=0}^{a-j+1} \sum_{k=0}^{a-j+i-1} p((1+\rho)\lambda L; j-i) \times [p(\lambda(\ell-L); i) p(\lambda(\ell-L); k) - p(\lambda(\ell-L); j+k-a) p(\lambda(\ell-L); i-j+a)]
\]

\[
\sum_{m=0}^{a-1} \sum_{i=0}^{a-m-1} \sum_{k=0}^{a-m-1} p(\lambda(1+\rho)L; m) \times [p(\lambda(\ell-L); i) p(\lambda(\ell-L); k) - p(\lambda(\ell-L); m+i+k-a) p(\lambda(\ell-L); a-m)]
\]

where we have substituted \( m = j - i \)

\[
\sum_{m=0}^{a-1} p(\lambda(1+\rho)L; m) \times [p(\lambda(\ell-L); a-m-1) - p(\lambda(\ell-L); a-m)] \sum_{n=0}^{a-m-2} p(\lambda(\ell-L); n)
\]

from which (3.18) follows immediately.

When \( \ell \) lies between \( L \) and \( L/2 \) we can immediately deduce the corresponding formula for \( \beta_1 \) by noting that detecting an object of length \( L \in (\ell, 2\ell) \) and intensity \( \rho \) in a background noise level \( \lambda \) is equivalent to "detecting" an object of length \( 2\ell - L < \ell \) and intensity zero in a background noise level \( (1+\rho)\lambda \). (See Fig. 3)
FIGURE 3

EQUIVALENT DETECTION SITUATIONS
Replacing $\lambda(1+\rho)L$ by $\lambda(1+\rho)(2L-L)$, and $\lambda(L-L)$ by $\lambda(1+\rho)(L-L)$ in (3.18), we obtain (3.19).

Clearly, formulae for $\beta_1$ when $l < L/2$ could be obtained in a similar manner, using Theorem A for $n > 2$. These would, however, involve higher order determinants and be even more complicated than (3.18) and (3.19).

4. The optimal design of a detector.

Having constructed measures $\alpha'_1$, $\beta'_2$ of size and power, we can now combine them to determine optimal values of $a$ and $l$. Our approach to this is based on the standard statistical method: we specify the size $\alpha'_1$ required, and then, from detectors having that size, choose the one with the highest power. We shall consider only the case of a stationary renewal process.

We shall need to consider the quantiles of the convolutions $F(a)$, and to ensure the uniqueness of these quantiles, we restrict our attention to the class of distributions $F$ satisfying the following condition (*).

(*) $F(x)$ is supported on some interval $[x^-,x^+] \subseteq (0,\infty)$ where $x^+$ may be infinite, and is absolutely continuous with a density $f$ which does not vanish on any nonzero interval in $(x^-,x^+)$.  

Lewis (1977) has defined the class $S$ as those distributions which satisfy (*), but with support not necessarily restricted to $(0,\infty)$. We define $S^+ = \{F \mid F \text{ satisfies (*)} \} \subseteq S$.

For any value of $a = 1,2,\ldots$, and any $\alpha \in (0,1)$ define $l_{\alpha}(a)$ by

\[
F(a)(l_{\alpha}(a)) = \alpha.
\]

A detector of length $l_{\alpha}(a)$ will have an $\alpha'_1$-value equal to $\alpha$ when used with detection level $a+1$, and an $\alpha'_1$-value approximately equal...
to $\alpha$ (exactly equal in the Poisson case) with detection level $a$.

Note that $\ell_a(\alpha)$ is the $\alpha$ quantile of $F(a)$.

Using this length with detection level $a+1$ and detection interval $J_1$, the value of $\beta_2^{\prime}$ obtained is

\[(4.2) \quad F(a) \left( \ell_a(\alpha) + \rho \min (\ell_a(\alpha), L) \right) \]

and our aim is to choose the value $a^*$ of $a$ which maximizes this expression. Note that (4.2) is also approximately the value of $\beta_2$ obtained with detection level $a$.

We see that $\ell_a(\alpha)$ increases with $a$ since, using the sequence $X_0, X_1, \ldots$ of Section 3a,

\[(4.3) \quad \alpha = F(a) (\ell_a(\alpha)) = \text{Prob}(X_1 + \ldots + X_a \leq \ell_a) \]

\[= 1 - \text{Prob}(X_1 + \ldots + X_a > \ell_a) \]

\[> 1 - \text{Prob}(X_1 + \ldots + X_a + X_{a+1} > \ell_a) \]

\[= F(a+1) (\ell_a) \]

so that $\ell_{a+1}(\alpha) > \ell_a(\alpha)$. In fact, $\ell_a(\alpha) \uparrow \infty$, unless $F$ is concentrated at zero, for if $F(\ell) < 1$, then $F(n\ell) < 1$ and so for any $x$ there is a value of $n$ such that $F^{(n)}(x) < 1$. But then

\[F^{(na)}(x) = \text{Prob}(X_1 + \ldots + X_{na} \leq x) \]

\[< \text{Prob}(\max(X_1 + \ldots + X_a, X_{a+1} + \ldots + X_{2a}, \ldots, X_{(n-1)a+1} + \ldots + X_{na}) \leq x) \]

\[\leq [F^{(n)}(x)]^a \]

\[\to 0 \text{ as } a \to \infty. \]

Thus for large enough $a$, $F^{(a)}(x) < \alpha$ and so $\ell_a(\alpha) > x$. 
Thus, in particular, $l_a(\alpha) > L$ for large enough $a$. Define $a_o(\alpha, L)$ by

$$l_{a_0}(\alpha) < L < l_{a_0+1}(\alpha)$$

($a_0 = 0$ if $L < l_1(\alpha)$) so that $a_0$ and $a_0+1$ are the detection levels giving size $\alpha$ detectors with lengths as close as possible to the length of the object. We shall give two conditions on $F$ which together will ensure that either $a^* = a_0$ or $a^* = a_0+1$.

**Definition (4.5)** A distribution $F \in S^+$ has decreasing quantile ratios if for all $\alpha, \beta \in (0, 1)$ s.t. $\alpha < \beta$ we have

$$\frac{l_a(\beta)}{l_a(\alpha)}$$

decreases with $a$

where $l_a(\theta)$ is the $\theta$ quantile of $F^{(a)}$.

**Lemma 2** The detection level $a^*(\alpha, L)$ of the size $\alpha$ detector with greatest power $\beta_2'$ satisfies

$$a^*(\alpha, L) \geq a_o(\alpha, L)$$

for all $\alpha, L$ and $\rho$ if $F$ has decreasing quantile ratios.

**Proof:** Write $\phi(\rho, a, \alpha) = F^{(a)}((1+\rho)l_a(\alpha))$. Then if $a \leq a_0$,

so that $l_a \leq L$, $\beta_2$ is equal to $\phi$. If $\phi$ increases with $a$, then $\beta_2$ also increases with $a$ for $a \leq a_0$ and so we must have $a^* \geq a_0$. Thus it is enough to show that $\phi$ increases with $a$.

But, writing $\beta = F^{(a)}((1+\rho)l_a(\alpha))$ we have $\beta > \alpha$ and

$l_a(\beta) = (1+\rho)l_a(\alpha)$. Thus, if $F$ has decreasing quantile ratios,

$$\frac{l_a(\beta)}{l_a(\alpha)} > \frac{l_{a+1}(\beta)}{l_{a+1}(\alpha)}$$
so that

\[ F(a+1)((1+\rho)\ell_{a+1}(\alpha)) > F(a+1)(\ell_{a+1}(\beta)) \]

\[ = \beta \]

\[ = F(a)((1+\rho)\ell_{a}(\alpha)) \]

and hence \( \phi \) increases with \( a \).

**Definition (4.8)** A distribution \( F \in S \) has *increasing interquartile ranges* if for all \( \alpha, \beta \in (0,1) \) s.t. \( \alpha < \beta \) we have

\[ (4.9) \quad \ell_{a+1}(\beta) - \ell_{a+1}(\alpha) > \ell_{a}(\beta) - \ell_{a}(\alpha), \]

where \( \ell_{a}(\theta) \) is the \( \theta \) quantile of \( F(a) \).

**Lemma 3** Using the notation of Lemma 2, if \( F \) has increasing interquartile ranges, then

\[ (4.10) \quad a^*(\alpha,L) \leq a_0(\alpha,L) + 1 \quad \text{for all } \alpha, L, \rho. \]

**Proof:** For \( \ell_{a}(\alpha) \geq L \), i.e. \( a \geq a_0 + 1 \) we have

\[ \beta'_2 = F(a)(\ell_{a}(\alpha) + \rho L). \]

Writing \( \beta = F(a)(\ell_{a}(\alpha) + \rho L) \), and provided \( F \) has increasing interquartile ranges,

\[ \ell_{a}(\beta) - \ell_{a}(\alpha) = \rho L < \ell_{a+1}(\beta) - \ell_{a+1}(\alpha) \]

so that \( F(a+1)(\ell_{a+1}(\alpha) + \rho L) < \beta \), and hence \( \beta_2 \) decreases with \( a \) for \( a > a_0 + 1 \). Hence \( a^* \leq a_0 + 1 \).

Combining the results of Lemmata 2 and 3 we have the result

**Theorem 1:** If \( F \in S^+ \) has increasing interquartile ranges and decreasing quantile ratios, then
We shall show below that the exponential distribution has both decreasing quantile ratios, and increasing interquantile ranges, so that for the Poisson process the result of the theorem holds. It is clear that, if $F$ has increasing interquantile ranges or decreasing quantile ratios, then so does $F^{(n)}$ for any $n$. Thus the result of the theorem holds also for a renewal process with Erlangian interevent times. It is also clear that a scale change leaves both properties unaffected. However a location change may affect the quantile ratios.

If $X_i$ has distribution function $F(x)$, then $X_i + c$ has distribution function $F(x-c)$, and $\sum_{i=1}^{n}(X_i + c)$ has distribution function $F^{(n)}(x-nc)$; if $\bar{\lambda}_n^\alpha (\alpha)$ is the $\alpha$ quantile of $F^{(n)}(x-nc)$, then $\bar{\lambda}_n^\alpha (\alpha) = \lambda_n^\alpha (\alpha) + nc$. If $F(x)$ has increasing interquantile ranges, then so does $F(x-c)$ since $\bar{\lambda}_n^\alpha (\alpha) - \bar{\lambda}_n^\beta (\beta) = \lambda_n^\alpha (\alpha) - \lambda_n^\beta (\beta)$. However, to ensure that $F(x-c)$ has decreasing quantile ratios, we need another condition on $F$'s interquantile ranges.

**Lemma 4** If $F(x) \in S^+$ has decreasing quantile ratios and its interquantile ranges satisfy

\[
\frac{\lambda_n^\alpha (\alpha) - \lambda_n^\beta (\beta)}{n} \geq \frac{\lambda_{n+1}^\alpha (\alpha) - \lambda_{n+1}^\beta (\beta)}{n+1}
\]

$n = 1, 2, \ldots \quad 0 \geq \alpha > \beta \geq 1$

then $F(x-c)$ also has decreasing quantile ratios.

**Proof:** Since $F$ has decreasing quantile ratios, we have

\[
\lambda_n^\alpha (\alpha) \lambda_{n+1}^\beta (\beta) \geq \lambda_n^\beta (\beta) \lambda_{n+1}^\alpha (\alpha) \quad \text{for} \quad \alpha > \beta.
\]
By (4.12), 
\[(n+1) \left( \ell_n(\alpha) - \ell_n(\beta) \right) \geq n(\ell_{n+1}(\alpha) - \ell_{n+1}(\beta)) \]

and so, by (4.13),
\[
\ell_n(\alpha) \ell_{n+1}(\beta) + (n+1)c (\ell_n(\alpha) - \ell_n(\beta)) \\
\geq \ell_n(\beta) \ell_{n+1}(\alpha) + nc(\ell_{n+1}(\alpha) - \ell_{n+1}(\beta))
\]

and, after some rearrangement we have
\[
(\ell_n(\alpha) + nc)(\ell_{n+1}(\beta) + (n+1)c) \geq (\ell_n(\beta) + nc)(\ell_{n+1}(\alpha) + nc)
\]

so that \(F(x-c)\) has decreasing quantile ratios.

If the quantile ratios of \(F\) are increasing, so that \(\ell_a(\alpha)/\ell_a(\beta)\) increases with \(a\) for \(\alpha > \beta\), then a slight modification of the proof of Lemma 2 shows that \(F^{(a)}(\ell_a(\alpha) + \rho \ell_a(\alpha))\) decreases with \(a\). In fact we have the following result.

**Theorem 2:** If \(F \in S^+\) has increasing quantile ratios, and increasing interquartile ranges, then \(a^* = 1\).

**Proof:** By Lemma 3, \(\beta'_2\) is decreasing for \(a \geq a_0+1\), and by the above remark, \(\beta'_2\) is decreasing for \(a \leq a_0\). Thus we have only to show that \(\beta'_2(a_0) \geq \beta'_2(a_0+1)\) to establish the result. But

\[
(4.14) \quad \beta'_2(a_0) = F^{(a_0)}(\ell_{a_0}(\alpha) + \rho \ell_{a_0}(\alpha)) \\
\geq F^{(a_0+1)}(\ell_{a_0+1}(\alpha) + \rho \ell_{a_0+1}(\alpha))
\]

by the above remark,
\[
\geq F^{(a_0+1)}(\ell_{a_0+1}(\alpha) + \rho L)
\]

since \(\ell_{a_0+1} > L\)
\[
= \beta'_2(a_0+1)
\]

and the result follows.
In Section 5 we shall give a class of distributions satisfying the conditions of Theorem 2. It is intuitively clear that a detection level \( a = 1 \) will not give a satisfactory detector and so it seems that this approach to detector design is not appropriate for renewal process with these distributions. However, Lemma 5 shows that any distribution with increasing quantile ratios has infinite mean, so that these distributions are unlikely to arise in a practical context.

Lemma 5 \( \text{If } F \in S^+ \text{ has increasing quantile ratios, then } \int_0^\infty x dF(x) = \infty. \)

Proof: Suppose that \( F \) has a finite mean \( \mu \). By the law of large numbers, as \( a \to \infty \),

\[
F^{(a)}(ax) \to 0 \quad x < \mu
\]
\[
+ 1 \quad x > \mu
\]

and so, since \( F^{(a)}(\lambda^a(\alpha)) = \alpha \) for all \( a \),

\[
\lambda^a(\alpha)/a \to \mu \quad \text{for all } \alpha \in (0,1).
\]

Thus for any \( \alpha > \beta \in (0,1) \),

\[
\lambda^a(\alpha)/\lambda^a(\beta) \to 1
\]
\[
< \lambda^a(\alpha)/\lambda^a(\beta) \quad \text{since } \alpha > \beta,
\]

and (4.17) clearly implies that the quantile ratios of \( F \) are not increasing.
5. Interquantile ranges and quantile ratios.

We first show that the exponential distribution has increasing interquantile ranges and decreasing quantile ratios so that Theorem 1 holds for a Poisson process. If \( F \) is the distribution function of an exponential distribution with mean \( 1/\lambda \), then

\[
P^{(a)}(x) = 1 - P(\lambda x; a-1)
\]

where \( P(x,n) \) again denotes the Poisson distribution function. Thus \( l_a(\alpha) \) satisfies

\[
P(\lambda l_a; a-1) = 1 - \alpha
\]

We have to show that, for \( \alpha > \beta, l_a(\alpha) - l_a(\beta) < l_{a+1}(\alpha) - l_{a+1}(\beta) \) and \( l_a(\alpha)/l_a(\beta) > l_{a+1}(\alpha)/l_{a+1}(\beta) \). The first of these is easily seen to be equivalent to

\[
P(\lambda l_a + c; a-1) < P(\lambda l_{a+1} + c; a) \quad \text{for all } c > 0
\]

while the second is equivalent to

\[
P(\rho\lambda l_a; a-1) > P(\rho\lambda l_{a+1}; a) \quad \text{for all } \rho > 1.
\]

To prove (5.3) consider independent Poisson random variables \( X, Y, Z \) with means \( \lambda l_a, \lambda(\lambda - l_a), c \) respectively. Then (5.3) becomes, considering the complementary probabilities,

\[
\text{Prob}(X + Z \geq a) > \text{Prob}(X + Y + Z \geq a + 1).
\]

But we have
\[ \alpha = \text{Prob}(X \geq a) \]

\[ = \text{Prob}(X + Y \geq a + 1) \]

\[ = \sum_{j=0}^{a} \text{Prob}(X=j) \text{Prob}(Y \geq a - j + 1) \]

\[ = \alpha - \text{Prob}(X=a) \text{Prob}(Y \geq 1) \]

\[ + \sum_{j=0}^{a-1} \text{Prob}(X=j) \text{Prob}(Y \geq a - j + 1) \]

\[ = \alpha - \text{Prob}(X=a) \text{Prob}(Y=0) + \sum_{j=0}^{a-1} \text{Prob}(X=j) \text{Prob}(Y \geq a - j + 1) \]

so that

\[ (5.6) \quad \text{Prob}(Y=0) = \sum_{j=0}^{a-1} \frac{\text{Prob}(X=j)}{\text{Prob}(X=a)} \text{Prob}(Y \geq a - j + 1). \]

Similarly

\[ (5.7) \quad \text{Prob}(X + Y + Z \geq a + 1) = \text{Prob}(X + Z \geq a) - \text{Prob}(X+Z=a) \text{Prob}(Y=0) \]

\[ + \sum_{j=0}^{a-1} \text{Prob}(X+Z=j) \text{Prob}(Y \geq a - j + 1) \]

\[ = \text{Prob}(X + Z \geq a) \]

\[ + \sum_{j=0}^{a-1} \text{Prob}(X+Z=a)\left\{ \frac{\text{Prob}(X+Z=j)}{\text{Prob}(X+Z=a)} - \frac{\text{Prob}(X=j)}{\text{Prob}(X=a)} \right\} \]

\[ \times \text{Prob}(Y \geq a - j + 1). \]

But since \( j < a \) and \( c > 0 \)

\[ \frac{\text{Prob}(X+Z=j)}{\text{Prob}(X+Z=a)} = \frac{a!}{j!} \frac{1}{(a+c)^{a-j}} < \frac{a!}{j!} \frac{1}{(a)^{a-j}} = \frac{\text{Prob}(X=j)}{\text{Prob}(X=a)} \]

and (5.5) follows, so that we have shown that the exponential distribution has increasing interquantile ranges.
To prove (5.4) we again introduce independent Poisson random variables \(X, Y, V, W\), where \(X\) and \(Y\) are as above, \(V\) has mean \(\lambda (p-1)\ell a\), and \(W\) has mean \(\lambda (p-1)(\ell a + 1 - \ell a)\). Then (5.4) is equivalent to

\[
(5.8) \quad \text{Prob}(X + V \geq a) < \text{Prob}(X + Y + V + W \geq a + 1) .
\]

Arguing as above, and using (5.6), we find that

\[
(5.9) \quad \text{Prob}(X + Y + V + W \geq a + 1)
= \text{Prob}(X + V \geq a)
+ \sum_{j=0}^{a-1} \frac{\text{Prob}(X + V = a) \text{Prob}(W = 0)}{\text{Prob}(X + V = a)}
\times \left\{ \frac{\text{Prob}(X+V=j) \text{Prob}(Y+W \geq a-j+1)}{\text{Prob}(X+V=a) \text{Prob}(W=0)} - \frac{\text{Prob}(X=j) \text{Prob}(Y \geq a-j+1)}{\text{Prob}(X=a) \text{Prob}(Y \geq a-j+1)} \right\}.
\]

The bracketed term in the sum is

\[
(5.10) \quad \sum_{i=a-j+1}^{\infty} \frac{\text{Prob}(X+V=j)}{\text{Prob}(X+V=a)} \frac{\text{Prob}(Y+W=i)}{\text{Prob}(W=0)} - \frac{\text{Prob}(X=j)}{\text{Prob}(X=a)} \frac{\text{Prob}(Y=i)}{\text{Prob}(Y \geq a-j+1)}
= \sum_{i=a-j+1}^{\infty} \frac{a!}{i!} \frac{1}{\ell a} \frac{1}{a-j} \frac{\ell a + 1 - \ell a}{i!} e^{-\rho (\ell a + 1 - \ell a)} \rho i-a+j-1
= \sum_{i=a-j+1}^{\infty} \frac{a!}{i!} \frac{1}{\ell a} \frac{1}{a-j} \frac{(\ell a + 1 - \ell a)^i}{i!} e^{-\rho (\ell a + 1 - \ell a)} \rho i-a+j-1
> 0 \text{ since } \rho > 1
\]

and so the sum in (5.9) is positive, and (5.8) follows. Thus the exponential distribution has decreasing quantile ratios.

As was noted above, these results imply that Theorem 1 holds for a renewal process with any Erlangian distribution of interevent times.
As a corollary of (5.3) we note that we can deduce the following result for the behaviour of the $\alpha_2$-value of a detector.

**Lemma 6** For a Poisson process, the $\alpha_2$-value of detectors having detection levels, $a = 2, 3, \ldots$, and lengths $l_a(\alpha)$, where $\alpha$ is fixed, decreases with $a$.

**Proof:** Since $P(\lambda l_a; a-1) = \alpha$ is independent of $a$, and, from (5.3), $P(\lambda l_a + c; a-1)$ decreases with $a$ for fixed $\alpha$, it is clear that

\[(5.11) \quad \frac{\partial}{\partial c} P(\lambda l_a + c; a-1) \bigg|_{c=0} \text{ decreases with } a.\]

But

\[(5.12) \quad \frac{\partial}{\partial c} P(\lambda l_a + c; a-1) \bigg|_{c=0} = p(\lambda l_a; a-1)\]

so that $\alpha_2$ decreases with $a$ for fixed $\alpha'$. So that $\alpha_2$ decreases with $a$ for fixed $\alpha'$.

Since $\alpha_1$ measures the total length of objects registered, while $\alpha_2$ measures the number of objects registered, Lemma 6 implies that the average length of a registered "object" increases with the length of the detector.

A class of distributions for which Theorem 2 holds can be constructed from the strictly stable distributions on $[0, \infty)$. Let $F$ be a strictly stable distribution on $[0, \infty)$ of exponent $\gamma$. Feller (1971, p. 448) shows that $0 < \gamma < 1$. For any $s \in (-\infty, \infty)$ let $\delta_s$ denote the distribution concentrated at $s$.

The quantiles $l_a(\alpha)$ of $F^{(a)}$ are given by
\[(5.13) \quad \lambda_a(\alpha) = a^{1/\gamma} \lambda_1(\alpha) \]
since
\[
F(a)(x) = F(a^{-1/\gamma}x)\]
and so the quantiles \(\lambda_a(s,\alpha)\) of \(F_a = F(a)^s\) are given by
\[(5.14) \quad \lambda_a(s,\alpha) = as + a^{1/\gamma} \lambda_1(\alpha).\]

For the distributions \(F_s\) with \(s > 0\) we have the following result.

**Lemma 7** If \(F\) is a strictly stable distribution on \([0,\infty)\) of exponent \(\gamma \in (0,1)\), and \(s > 0\) then \(F_s\) has increasing interquantile ranges and increasing quantile ratios.

**Proof:** For \(\alpha > \beta\), by (5.14),
\[(5.15) \quad \lambda_a(s,\alpha) - \lambda_a(s,\beta) = a^{1/\gamma} (\lambda_1(\alpha) - \lambda_1(\beta))\]
which increases with \(a\), so that \(F_s\) has increasing interquantile ranges.

Also
\[(5.16) \quad \lambda_a(s,\alpha) \lambda_b(s,\beta) - \lambda_a(s,\beta) \lambda_b(s,\alpha)\]
\[= s(\lambda_1(\alpha) - \lambda_1(\beta))ab(a^{1/\gamma-1}b^{1/\gamma-1})\]
\[> 0\quad \text{for} \quad a > b, \quad \text{since} \quad 0 < \gamma < 1.\]

Thus \(F_s\) has increasing quantile ratios.

The strictly stable distributions themselves have increasing interquantile ranges, while their quantile ratios are given by
\[(5.17) \quad \lambda_a(\alpha)/\lambda_a(\beta) = (a^{1/\gamma} \lambda_1(\alpha))/(a^{1/\gamma} \lambda_1(\beta))\]
\[= \lambda_1(\alpha)/\lambda_1(\beta)\]
and so are independent of \(a\). Thus for a renewal process with strictly stable interevent time distribution, \(\beta_2'(a)\) is constant for \(a \leq a_0\).
For such a process, any value of $a \leq a_0$ is optimal. As is implied by Lemma 5, the strictly stable distributions on $(0,\infty)$ have infinite means. (Feller (1971)). The construction of the stationary process used in Section 3 is strictly speaking not possible for such distributions. However, the stationarity is not needed for the evaluation of $q'_1$ and $\beta'_2$ so that Theorem 2 is still valid.

Lewis (1977) has described a further class of distributions, called dispersive distributions, which satisfy (4.9). If two distributions $G, H$ in $S$ satisfy

\[ G^{-1}(\alpha) - G^{-1}(\beta) \leq H^{-1}(\alpha) - H^{-1}(\beta) \]

for all $\alpha, \beta \in (0,1)$, with strict inequality for some $\alpha, \beta$,

then $G$ and $H$ are said to be ordered in dispersion, denoted by

\[ G \overset{\text{disp}}{<} H. \]

A distribution $F$ is called dispersive if it satisfies

\[ \overset{\text{disp}}{G} < \overset{\text{disp}}{H} \Rightarrow F \ast G < F \ast H. \]

Lewis (1977) shows that any dispersive distribution satisfies (4.9), and so for any renewal process whose interevent time has a dispersive distribution, we see from Lemma 3 that the optimal detection level $a^*(\alpha, L)$ is not greater than $a_0(\alpha, L) + 1$.

We have not been able to carry out an analogous approach to the quantile ratio property (4.5). If we define "ordering in quantile ratio" $G < H$ for $G, H \in S^+$ to mean

\[ G^{-1}(\alpha)/G^{-1}(\beta) < H^{-1}(\alpha)/H^{-1}(\beta) \quad \forall \alpha > \beta \in (0,1), \]

then we would wish to find "ratio dispersive" distributions in $S^+$ such that
\[ (5.20) \quad G < H \Rightarrow F \ast G < F \ast H, \]

and also

\[ (5.21) \quad F \ast F < F. \]

However the following Lemma shows the search to be vain.

**Lemma 8** If \( F \) is ratio dispersive, then \( F \ast F \geq F \).

**Proof:** For any \( x > 0 \), and any \( G \in S^+ \),

\[ (5.22) \quad \delta_x < G \]

since every quantile of \( \delta_x \) is equal to \( x \).

Let \( F \) be a ratio dispersive distribution. Then

\[ F_x = F \ast \delta_x < F \ast G \]

and, in particular,

\[ F_x < F \ast F \]

so that, letting \( x \to 0 \), in the obvious notation

\[ F \leq F \ast F. \]

(Equality of quantile ratios could hold, e.g. if \( F \) were stable).

Lemma 8 shows that any ratio dispersive distribution must have increasing quantile ratios and so by Lemma 5, must have infinite mean.

It seems that the distributions with increasing interquantile ranges form a larger class than those having decreasing quantile ratios, so that Lemma 3 applies more widely than Lemma 2. Indeed it is easy to see that no distribution can have *decreasing* interquantile ranges, since, for any distribution \( F \),
(5.23) \[ \ell_a(\alpha) = \sum_{j=1}^{\alpha} \left[ \ell_{a+j}(\alpha) - \ell_a(\alpha+j+1) \right] \]

and since \( \ell_a(\alpha) \) increases with \( a \), as was noted in Section 4, not all of the terms in the sum can decrease with \( a \).

It is not however true that every distribution has increasing interquantile ranges, as might be supposed. For if \( F_{\varepsilon} \) is the distribution on \([0,2]\) whose density is

(5.24) \[ f_{\varepsilon}(x) = \begin{cases} 1 - \varepsilon & 0 \leq x < 1 \\ \varepsilon & 1 \leq x \leq 2 \end{cases} \]

then the density of \( F_{\varepsilon}^{(2)} \) is

(5.25) \[ f_{\varepsilon}^{(2)}(x) = \begin{cases} (1-\varepsilon)^2x & 0 \leq x \leq 1 \\ 2\varepsilon(1-\varepsilon)(x-1) + (1-\varepsilon)^2(2-x) & 1 \leq x \leq 2 \\ 2\varepsilon(1-\varepsilon)(3-x) + \varepsilon^2(x-2) & 2 \leq x \leq 3 \\ \varepsilon^2(4-x) & 3 \leq x \leq 4 \\ 0 & \text{elsewhere}. \end{cases} \]

The \((1-\varepsilon)\) quantiles of \( F_{\varepsilon} \) and \( F_{\varepsilon}^{(2)} \) are respectively 1 and 2. For \( \varepsilon < \frac{1}{4} \), the density \( f_{\varepsilon}(1+x) = \varepsilon \) is less than the corresponding density \( f_{\varepsilon}^{(2)}(2+x) \) for \( 0 \leq x < \frac{1}{4}(1-2\varepsilon)/(1-\varepsilon) \). Thus any quantile of \( F_{\varepsilon} \) lying between 1 and \( 1 + \frac{1}{4}(1-2\varepsilon)/(1-\varepsilon) \) will be further from 1 than the corresponding \( F_{\varepsilon}^{(2)} \) quantile is from 2. Hence \( F_{\varepsilon} \) does not have increasing interquantile ranges.

6. The two dimensional problem.

In sufficiently simple cases, the conclusion of Theorem 1 can be extended to the two dimensional problem.

Consider again the situation depicted in Fig. 1. Suppose the photons form a Poisson process of rate \( \lambda \) per unit area in the background,
and rate \((1 + \rho) \lambda\) within the object. Suppose further that the object is a circle of known area \(L\) and that we restrict our choice of detectors to be circles of area \(l\). Defining the position of the detector by the position of its centre, we can define a detection region \(J_1\) corresponding to the interval \(J_1\) of Section 3b, in which the object completely contains the detector, or vice versa. The measures \(\alpha_1, \beta_2\) of size and power have the forms (3.10) and (3.27), exactly as for the one dimensional case, and so the conclusions of Theorem 1 hold - i.e. the optimal detector will be a circle of the same size as the object.

Clearly the same conclusion holds if the object is e.g. a square, and the detector is a square having the same orientation.

Acknowledgements

This work forms part of my Ph.D. thesis supervised by Prof. P.A.P. Moran at the Australian National University. My thanks are due to Prof. T. Lewis for providing me with a preprint of his 1977 paper.
APPENDIX B

ON THE QUANTILES OF THE GAMMA
AND F DISTRIBUTIONS

I.W. SAUNDERS
CSIRO Division of Mathematics and Statistics
Canberra

and

P.A.P. MORAN
Australian National University
Canberra
1. Introduction

As an approach to the study of the 'matching' of optical receptors to the objects they are designed to detect, Saunders (1978) has proposed the following model:

The photons emitted by the object and background are represented by a point process on the real line $\mathbb{R}$, the 'object' being a region of higher 'photon' intensity, occupying an interval of length $L$. The receptor is an interval of length $\ell$ which moves along the line registering at each $t \in \mathbb{R}$ the number $N(\ell;t)$ of points of the process in the interval $(t-\ell,t]$. Saunders defines measures of size and power for a detection procedure based on $N(\ell;t)$, and shows that the most powerful detector of a given size has length $\ell$ close to $L$ provided the 'photons' form a renewal process whose interevent-time distribution $F$ satisfies the following conditions (*). Let $F^{(n)}$ denote the $n$-fold convolution of $F$ $(n=1,2,\ldots)$.

a) The quantiles $\zeta_n(a) = (F^{(n)})^{-1}(a)$ are uniquely defined, and continuous in $a$.

(*) b) The interquantile ranges $\zeta_n(a) - \zeta_n(\beta)$ increase with $n$ for fixed $a > \beta \in (0,1)$.

c) The quantile ratios $\zeta_n(a)/\zeta_n(\beta)$ decrease with $n$ for fixed $a > \beta \in (0,1)$.

He also shows that these conditions are satisfied when $F$ is an exponential distribution. This result may be stated as

Theorem A: If $x_n(a)$ denotes the $a$ quantile of the gamma distribution with shape parameter $n$, then for any integers $n > m \geq 1$, and for any $a > \beta \in (0,1)$,

(1.1) $x_n(a) - x_n(\beta) > x_m(a) - x_m(\beta),$

and

(1.2) $x_n(a)/x_n(\beta) < x_m(a)/x_m(\beta).$
The gamma distribution with integer shape parameter $n$ is an Erlangian distribution - the distribution of the sum of $n$ i.i.d. exponential random variables.

Lewis (1977) has defined the concept of 'ordering in dispersion'. Two distributions $F$ and $G$ are said to be ordered in dispersion (o.d.) denoted by

$$F \overset{\text{disp}}{\preceq} G$$

if their quantiles $x_F, x_G$ satisfy

$$x_F(a) - x_F(\beta) \preceq x_G(a) - x_G(\beta) \quad \text{for all } a > \beta.$$ 

Thus (1.1) implies that any two Erlangian distributions are o.d. so that the Erlangian distributions form an o.d. sequence. Note that $F$ has increasing interquantile ranges iff its convolutions $F^{(n)}$ form an o.d. sequence.

In this note we show that the results of theorem A are true also for noninteger values of the parameters $n, m$ so that from (1.1) the gamma distributions form an o.d. class. This implies that gamma distributions have increasing interquantile ranges and decreasing quantile ratios. Thus for a renewal process with gamma interevent times, the optimal detector has length matching the object length. We also establish analogous results for the distribution of the ratio of two gamma random variables which, apart from a scale change, is an F distribution.

We first establish some conditions on a one parameter class of distributions which are equivalent to the quantile properties (1.1), (1.2). Let $\{F_a \mid a \in \mathbb{R}\}$ be a class of distribution functions such that $F_a$ is supported on some interval $(x_-(a), x_+(a)) \subset (0, \infty)$ and has a density $f_a$ which does not vanish on any subinterval of $(x_-(a), x_+(a))$. Then $\zeta_a(a) = F_a^{-1}(a)$ is uniquely defined for any $a \in (0,1)$ and is continuous in $a$. The following condition is both necessary and sufficient to establish that $\zeta_a(a) - \zeta_a(\beta)$ increases with $a$ for $a > \beta$. 

2.
(1.3) \( F_a(\zeta_a(\alpha)+c) \) decreases with \( a \) for any constant \( c > 0 \).

For if (1.3) holds, then, writing \( \alpha = F_a(\zeta_a(\beta)+c) \), we see that for any \( d > 0 \)

\[
F_{a+d}(\zeta_{a+d}(\beta)+c) < F_a(\zeta_a(\beta)+c) = \alpha = F_{a+d}(\zeta_{a+d}(\alpha))
\]

and hence

(1.4) \( \zeta_{a+d}(\alpha) - \zeta_{a+d}(\beta) > c = \zeta_a(\alpha) - \zeta_a(\beta) \).

Thus (1.3) is a sufficient condition.

If (1.4) holds for all \( a \) and \( d \), then we can reverse the above argument to show that (1.3) is also necessary.

When \( F_a \) is differentiable with respect to \( a \), (1.3) is equivalent to

(1.5) \( \frac{d}{da} (F_a(\zeta_a(\beta)+c)) < 0 \) for any \( c > 0 \),

or

(1.6) \( F'_a(\zeta_a(\beta)+c) + f_a(\zeta_a(\beta)+c) \zeta'_a(\beta) > 0 \)

where \( F'_a(x) = \frac{\partial F_a(x)}{\partial a}, f_a(x) = \frac{\partial f_a(x)}{\partial x}, \zeta'_a(\beta) = \frac{\partial \zeta_a(\beta)}{\partial a} \).

Differentiating the defining equation \( F_a(\zeta_a(\beta)) = \beta \), and substituting into (1.5), we obtain the necessary and sufficient condition

(1.7) \( \frac{F'_a(\zeta_a(\beta)+c)}{F_a(\zeta_a(\beta)+c)} < \frac{F'_a(\zeta_a(\beta))}{F_a(\zeta_a(\beta))} \).

Now by suitably choosing \( \beta \) and \( c \) we can give \( \zeta_a(\beta) \) and \( \zeta_a(\beta)+c \) arbitrary positive values, and so (1.7) is equivalent to

(1.8) \( \frac{F'_a(x)}{f_a(x)} \) decreases with \( x \).

The following condition is necessary and sufficient to ensure that

\( \frac{\zeta_a(\alpha)}{\zeta_a(\beta)} \) decreases with \( a \) for \( a > \beta \):

(1.9) \( F_a(\rho \zeta_a(\beta)) \) increases with \( a \) for any \( \rho > 1 \).

for if (1.9) holds then taking \( \rho = \frac{\zeta_a(\alpha)}{\zeta_a(\beta)} \) so that \( F_a(\rho \zeta_a(\beta)) = \alpha \),
we see that, for any \( d > 0 \),

\[
F_{a+d}(\rho \xi_{a+d}(\beta)) > F_{a}(\rho \xi_{a}(\beta)) = \alpha = F_{a+d}(\xi_{a+d}(\alpha))
\]

so that

\[
(1.10) \quad \frac{\xi_{a+d}(\alpha)}{\xi_{a+d}(\beta)} < \rho = \frac{\xi_{a}(\alpha)}{\xi_{a}(\beta)}
\]

and (1.9) is sufficient. Once again, reversing the argument shows that (1.9) is also necessary.

When \( F_{a} \) is differentiable w.r.t. \( a \) the following equivalent conditions are necessary and sufficient; they are easily proved in the same way as (1.5), (1.6), and (1.8) to which they respectively correspond.

\[
(1.11) \quad \frac{d}{da} F_{a}(\rho \xi_{a}(\beta)) > 0 \quad \text{for any } \rho > 1, \beta \in (0,1).
\]

\[
(1.12) \quad F'_{a}(\rho \xi_{a}(\beta)) + \rho f_{a}(\rho \xi_{a}(\beta)) \frac{\xi'_{a}(\beta)}{\xi_{a}(\beta)} > 0 \quad \text{for any } \rho > 1.
\]

\[
(1.13) \quad F'_{a}(x) / x f_{a}(x) \text{ increases with } x.
\]

In section 2 below we show that (1.8) and (1.13) are true for the gamma distribution, for which

\[
(1.14) \quad f_{a}(x) = e^{-x} x^{a-1} / \Gamma(a),
\]

while in section 3 we establish (1.5) and (1.11) for the distribution of the ratio of two gamma random variables, when

\[
(1.15) \quad f_{\frac{1}{x}}(x) = \frac{\Gamma(\frac{1}{x}+m)}{\Gamma(\frac{1}{x}) \Gamma(m)} \frac{x^{\frac{1}{x}-1}}{(1+x)^{\frac{1}{x}+m}}.
\]

2. The Gamma Distribution

Theorem 1. If \( x_{a}(\alpha) \) is the \( \alpha \) quantile of the gamma distribution with shape parameter \( a \), then \( (x_{a}(\alpha)-x_{a}(\beta)) \) increases with \( a \), while \( (x_{a}(\alpha)/x_{a}(\beta)) \) decreases with \( a \), for \( \alpha > \beta \).

**Proof:** Write \( \Gamma(a) = \int_{0}^{\infty} e^{-u} u^{a-1} du \), \( p(a;x) = e^{-u} u^{a-1}/\Gamma(a) \),

\[
P(a;x) = \int_{0}^{x} p(a;u) du, \quad p'(a;x) = (\partial/\partial a)p(a;x), \quad p'(a;x) = (\partial/\partial a)p(a;x).
\]

Using the convolution property of the gamma distribution,
P(a+b;.) = P(a;.) * P(b;.), we see that

\[ \frac{P(a+b;x) - P(a;x)}{p(a;x)} = - \int_0^x \frac{(1-p(b;u)) P(\{x-u\})}{p(a;x)} \, du \]

\[ = - \int_0^x (1-u^a) du \]

\[ = - \int_0^1 Q(b;zx) e^{2zx} x(1-z)^{a-1} \, dz \]

where \( Q(b;y) = 1 - P(b;y) \)

\[ = \int_y^\infty p(b;u) \, du. \]

But

\[ \frac{\partial}{\partial x} \{ Q(b;zx) e^{2zx} x \} = e^{2zx} \left[ (1+zx)Q - zxp(b;zx) \right] \]

\[ = e^{2zx} \left[ Q + zx(Q-p) \right] \]

\[ = \Gamma(b)^{-1} e^{2zx} \int_0^x \left[ e^{-u} u^{b-1} + zx(b-1)e^{-u} u^{b-2} \right] \, du \]

since \( \int_0^\infty u^{b-1} \, du = \Gamma(y)^{-1}(b-1) \int_y^\infty e^{-u} u^{b-2} \, du \)

\[ = \Gamma(b)^{-1} e^{2zx} \int_0^x \left[ (u-zx) + bzx \right] e^{-u} u^{b-2} \, du \]

\[ < 0 \text{ for all } b > 0. \]

Thus \( \frac{P(a+b;x) - P(a;x)}{p(a;x)} \) decreases with \( x \) for \( b > 0 \), which is sufficient to establish (1.8).

Similarly,

\[ \frac{P(a+b;x) - P(a;x)}{xp(a;x)} = - \int_0^1 Q(b;zx) e^{2zx} (1-z)^{a-1} \, dz \]

and

\[ \frac{\partial}{\partial x} \{ Q(b;zx) e^{2zx} \} = ze^{2zx} (Q-p) \]

\[ = ze^{2zx} (b-1) \int_{zx}^\infty p(b-1;u) \, du \]

\[ < 0 \text{ for } 0 < b < 1 \]

which establishes (1.13)
3. The Ratio of Gamma Random Variables

Theorem 2. If \( y_\alpha(x) \) is the \( \alpha \) quantile of the distribution with density (1.15), then \( y_\alpha(x) - y_\beta(x) \) increases with \( \ell \), while \( y_\alpha(x) / y_\beta(x) \) decreases with \( \ell \), for \( \alpha > \beta \) and any fixed positive \( m \).

Remark. If \( X \) has the density (1.15) and \( 2\ell \) and \( 2m \) are integers, then \( \alpha X/\beta \) has an \( F \) distribution with \( 2\ell \) and \( 2m \) degrees of freedom.

Proof of Theorem 2: Write \( g(\ell;x) \) for the density (1.15), and \( G(\ell;x) \) for the corresponding distribution function. For any fixed \( \alpha, \ell > 0, k > 0 \) let \( A(x) = G(\ell;x) - G(\ell+k;x + y_{\ell+k}(\alpha) - y_{\ell}(\alpha)) \). We shall show that \( A(x) > 0 \) for \( x > y_{\ell}(\alpha) \) and so establish (1.5) for \( c = x-y_{\ell}(\alpha) > 0 \).

Assume for the moment that \( y_{\ell+k}(\alpha) - y_{\ell}(\alpha) > 0 \) for \( \alpha > 0 \). We shall prove this below. By its definition, \( A(x) = 0 \) for \( x = y_{\ell}(\alpha) \) and, by our assumption, \( A(0) < 0 \).

Also \( A(x) \to 0 \) as \( x \to \infty \), and

\[
A'(x) = \frac{\Gamma(\ell+m)}{\Gamma(\ell)\Gamma(m)} \frac{x^{\ell-1}}{(1+x)^{\ell+m}} - \frac{\Gamma(\ell+m+k)}{\Gamma(\ell+k)\Gamma(m)} \frac{(x+d)^{\ell+k-1}}{(1+x+d)^{\ell+k+m}}
\]

\[
\leq \frac{1}{\Gamma(m)x^{m+1}} \left[ \frac{\Gamma(\ell+m)}{\Gamma(\ell)} - \frac{\Gamma(\ell+m+k)}{\Gamma(\ell+k)} \right] < 0
\]

since \( \frac{\partial}{\partial \ell} \ln(\Gamma(\ell+m)/\Gamma(\ell)) = \psi(\ell+m) - \psi(\ell) \),

where \( \psi(\ell) \) is the digamma function, which is increasing for \( \ell > 0 \). (Abramowitz and Stegun (1965)).

Thus for sufficiently large \( x \), \( A(x) \) is positive. We shall show that \( A'(x) \) has only one zero in \( (y_{\ell}(\alpha),\infty) \), which implies that \( A(x) \) is positive in \( (y_{\ell}(\alpha),\infty) \).

From (3.1) we see that

\[
A'(x) = \frac{\Gamma(\ell+m+k)}{\Gamma(m)\Gamma(\ell+k)} \frac{x^{\ell-1}}{(1+x)^{\ell+m}} (c_0 - h(x))
\]

where \( h(x) = \frac{(x+d)^{\ell+k-1}(1+x)^{\ell+m}}{x^{\ell-1}(1+x+d)^{\ell+m+k}} \)

and \( c_0 = \frac{\Gamma(\ell+m)\Gamma(\ell+k)}{\Gamma(\ell)\Gamma(\ell+m+k)} < 1 \)

so that \( h(x) = c_0 \) at any zero of \( A'(x) \).
Consider

\[
(3.3) \quad \frac{d}{dx} \ln(h(x)) = \frac{\ell+m}{1+x} + \frac{\ell+k+1}{x+d} - \frac{\ell+k+m}{1+x+d} - \frac{\ell-1}{x}
\]

which is zero when

\[
(3.4) \quad (md+d+k)x^2 + [(1+m)d^2+k-2(\ell-1)d]x - (\ell-1)d(1+d) = 0
\]

If \( \ell > 1 \), then \((\ell-1)d(1+d) > 0\), and (3.4) has at most one root in \((0,\infty)\), so that \( h(x) = c_0 \) has at most two roots in \((0,\infty)\). If \( \ell < 1 \), (3.4) has no roots in \((0,\infty)\), since all the coefficients are positive, and \( h(x) = c_0 \) has at most one root in \((0,\infty)\).

Thus \( A'(x) = 0 \) at most twice in \((0,\infty)\). If \( A'(x) = 0 \) twice in the interval \((y^*_\ell(a),\infty)\) then since \( A'(x) < 0 \) for large \( x \), we must have \( A'(y^*_\ell) < 0 \). But \( A(0) < 0, A(y^*_\ell) = 0 \), and since \( A' \) cannot change sign in \((0,y^*_\ell)\) it is clear that \( A'(y^*_\ell) \) cannot be negative, which is a contradiction. Thus \( A' = 0 \) at most once in \((y^*_\ell(a),\infty)\).

We have still to prove that \( d \) is positive, or equivalently that \( G(\ell;x) \) decreases with \( \ell \) for fixed \( x \). But taking \( a = 0 \) in the above argument, so that \( y^*_\ell = d = 0 \) and \( A(x) = G(\ell;x) - G(\ell+k;x) \) we see from (3.3) that \( h'(x) \neq 0 \) for \( x > 0 \) while \( h(0) = 0 \) and \( h(\infty) = 1 \), so that \( h(x) = c_0 \) has exactly one solution in \((0,\infty)\). Since \( A(x) > 0 \) for large enough \( x \), we see that \( A(x) \) is positive for all \( x > 0 \), and \( G(\ell;x) \) decreases with \( \ell \).

A similar argument establishes (1.11). For any fixed \( a, \ell > 0, k > 0 \), write \( B(x) = G(\ell;x) - G(\ell+k;xy^*_\ell+k(a)/y^*_\ell(a)) \), and let \( r = y^*_\ell+k(a)/y^*_\ell(a) \). We shall show that \( B(x) < 0 \) for \( x > y^*_\ell(a) \), provided \( k \) is small enough.

\( B(x) = 0 \) for \( x = 0 \) and for \( x = y^*_\ell \), and \( B(x) \to 0 \) as \( x \to \infty \). Hence \( B'(x) = 0 \) at least once in \((0,y^*_\ell)\) and at least once in \((y^*_\ell,\infty)\). Considering the roots of

\[
\frac{r(1+x)(\ell+m)(rx)^{\ell+k+1}}{(1+rx)^{\ell+k+m}} x^{\ell-1} = c_0
\]

and using arguments similar to those above, we find that \( B' \) has at most two zeros in \((0,\infty)\) and so \( B \) has only one zero, which must be \( y^*_\ell(a) \).

For small enough \( k, c_0 > r^{-m} \) and hence, as \( x \to \infty \),
8.

\[ B'(x) \sim \frac{1}{\Gamma(m)x^{m-1}} \left[ \frac{\Gamma(\ell+m)}{\Gamma(\ell)} - r^{-m}\frac{\Gamma(\ell+k+m)}{\Gamma(\ell+k)} \right] > 0. \]

Thus \( B(x) < 0 \) for large enough \( x \), and so \( B(x) \) is negative for \( x > y_\ell^*(\alpha) \).

Acknowledgement

Our thanks are due to Professor T. Lewis for providing us with a prepublication copy of his 1977 paper and for his comments on an earlier draft of our paper.