On Some Estimation and Distribution Problems

by

Suresh R. Paranjape

Department of Statistics

Division of Mathematical Sciences

Mimeograph Series No. 194

July 1969

This research was supported in part by the Office of Naval Research Contract No. N 00014-67-0226-0008. Reproduction in whole or in part is permitted for any purposes of the United States Government.

CHAPTER I

GENERAL INTRODUCTION

In the following chapters four problems of estimation and distribution have been treated. Chapter II studies "Window-Estimators" for the location and the size of the discontinuity for a class of density functions with a single, simple discontinuity. Chapter III treats special distribution problems. Estimation of the location of the discontinuity is a known problem,but not the size of the jump.

The basic idea of a "Window-Estimator" originates from "Spectral Analysis". This was first noticed by Parzen (6). He used it to develop estimators for the density function at a point and for the mode of a density function. Chernoff (1) treated the case in which some of the regularity conditions, assumed by Parzen (6), were not satisfied. Parzen, while reviewing this paper in Mathematical Review, praised the techniques used by him in these words - "Using impressive stochastic process techniques the consistency and asymptotic distribution are obtained". In these papers, the conditions on the "weight-functions" are quite general, but the density function is still assumed to be continuous. By imposing more conditions on the weight function, whose choice is always up to the statistician, it then appears that the class of the density functions can be enlarged.

In posing the problem of estimating the size and the discontinuity, we have already relaxed the condition of continuity. As one may expect, the window-estimator for the size of the discontinuity is less efficient if it is to be compared with the windown-estimator of the density function at that point. But these two are not comparable. As far as the author knows, there is no other estimator suggested in the literature for the size of the discontinuity; so the merits or demerits of the window-estimator cannot be discussed. On the other hand, if the window-estimator for the location of a discontinuity is to have any appeal, it will have to compete with the estimator, treated by Chernoff and Rubin (2). It has been shown that they stand on the same level on the basis of the order of the biases of the estimators. In fact, it is shown that the window-estimator has, asymptotically, similar representation as does the estimator of Chernoff and Rubin.

Chapter III treats a similar problem. We are interested in making inferences about a non-stationarity point of a stochastic process instead of a discontuity of a density function. The estimation of the location of a single non-stationarity point was posed by Rubin (9). He showed in the paper that, under suitable conditions, for the estimation problem of discontinuities in multivariate densities, there do exist hyper-efficient estimators, among them the maximum likelihood estimator. He further showed that the estimation problem is asymptotically equivalent to that for a non-stationary process, with possibly multidimensional "time" and unknown center of non-stationarity. At the end of the paper, he compared the efficiency of the maximum likelihood estimator of the non-stationarity point of one dimensional pure Gaussian case with the best estimate with quadratic loss function. This paper, therefore, introduced some problems about distributions of well-known statistics. A

Bayesian philosopher will start by assuming a particular a-priori distribution on the location of the nonstationarity point and a particular form for the loss function. There are quite well known theorems, which give conditions under which the mean or the median of the a posteriori distribution is the best invariant estimator of the location parameter. Hence all these facts demand a further study on the distribution of the mean and the median.

Here we shall treat a Gaussian process whose likelihood ratio has a typical representation (1) of Chapter III. For justification of (1), we refer to a paper of Rao and Rubin (7), which gives a necessary and sufficient condition for a Gaussian process to be the log-likelihood ratio process of Gaussian process. Another reason can also be given that the logarithm of a likelihood ratio for the location of a discontinuity approaches a Wiener-process as the size of the discontinuity tends to zero. Normally one would believe that the distribution of the mean would be easier to obtain than that of the median, for any distribution. Here, on the contrary, the distribution of the median, at least as a computable expression, has been obtained. The distribution of the mean is quite an open problem. We have given some results obtained on this problem in 3.4. We have also outlined a possible way of solving the problem in the same section.

CHAPTER II

ESTIMATION OF THE SIZE AND THE LOCATION OF THE DISCONTINUITY

2.1 Introduction

The problem of estimating quantities related to the density function are quite delicate in nature. Several papers dealing with such problems assume the density function to be continuous and sometimes even uniformly continuous to prove special properties of the estimator. The problem would become more involved or the results less efficient if one modifies such assumptions. Chernoff and Rubin (2) studied the maximum likelihood estimator of the location of the discontinuity of a density function. They suggested an estimator which is close to the maximum likelihood estimator by the magnitude of the order $O_p(n^{-1})$. They showed, asymptotically, the distribution of their estimator was related to a certain random walk problem. Breakwell and Chernoff, in an unpublished memorandum, derived the asymptotic distribution of this estimator. These approaches need more rigorous techniques than the earlier studies estimating the quantities related to the continuous density function. In what follows we will give a Window-Estimator for the size and location of the discontinuity of a density function. The treatment is similar to that of Rosenblatt (8), Parzen (6) or Chernoff (1). It has been proved that the bias in the estimator of the size of the discontinuity is $O_p\ (n^{-1/3})$. The asymptotic normality and the consistency are also proved following Parzen's techniques (6). For the sake of completeness, we shall restate a theorem in the preliminaries which had been a basis for the Window-Estimators. This theorem would be applicable at every continuity point of the density function.

Section 3 treats a new estimator for the location of the discontinuity for a class of density function. It is shown that the order of the bias of this new estimator, viz., window-estimator is same as that of the maximum likelihood estimator which is $0_p(n^{-1})$. We will also obtain the distribution of the estimator.

2.2 Preliminaries and Earlier Results

Let X be a random variable with absolutely continuous distribution function F(x). Let f(x) be the denisty function with Θ as a location of the discontinuity. In what follows, we shall assume Θ to be the simple discontinuity of f(x), i.e.,

$$\lim_{x \downarrow \Theta} f(x) = A > \lim_{x \uparrow \Theta} f(x) = B \quad ,$$

where A and B are finite.

We wish to estimate g(Θ) = A - B as well as Θ by the class of functions defined by

$$(1) \qquad g_n(\Theta) = \int W_n(y - \Theta)\, dF_n(y) = \frac{1}{n} \sum_{i=1}^{n} W_n(x_i - \Theta) \quad ,$$

where (i) $x_1, \ldots, x_n$ are independent observations on X,

(ii) $W_n(.)$ is a suitable weight function.

Before we discuss the earlier work on the estimation of the density function at a point or the estimation of the mode, we shall restate a theorem This theorem had been a basis for spectral analysis as well as in the estimation of the density function, mode of the density function, treated by number of authors.

Theorem 1: Let $W(y)$ be a Borel function satisfying the following conditions

(2) $\sup |W(y)| < \infty$

(3) $\int W(y)\, dy < \infty$

(4) $\lim_{y\to\infty} |yW(y)| = 0$

Further, let $h(n)$ be a sequence of constants, depending on n, such that

(5) $\lim_{n\to\infty} h(n) = 0$.

If

(6) $g_n(x) = \int \frac{1}{h(n)} W(\frac{y}{h(n)}) f(x-y)\, dy$,

then

$$g_n(x) \to f(x) \int W(y)\, dy$$

at every continuity point x of $f(x)$.

From this theorem and from the definition of $g_n(\Theta)$ in (1), one can easily impose conditions on $W_n(\cdot)$ so as to get a good estimator of the density function at a point Θ. If the problem is of estimation of the mode of a density function, the intuition will suggest the estimator as the location of $\sup_{\Theta} g_n(\Theta)$ with proper conditions on $W_n(\cdot)$. For a class of density function for which q th order bounded derivatives exist, it has been shown that the bias of the estimators in the class (1) is of the order $0_p(n^{-q/2q+1})$. For the estimation of the mode or the density the weight function had been taken to be symmetric at zero. Since we are interested in the estimation of the location or the size of the discontinuity of a density function, we will restrict ourselves to the class of weight functions which will eventually make the estimator zero everywhere except at the discontinuity. For this to happen, we will need to restrict ourselves to the class of density functions, with one simple discontinuity, which are well behaved otherwise. Also we will consider the class of weight functions which are antisymmetric at zero.

We need one more theorem to obtain the distribution of the statistic studied. This result was proved by Breakwell and Chernoff in an unpublished memorandum.

Theorem 2: Let $\hat{\theta}_n$ be the estimator of the location of the density function, studied by Chernoff and Rubin (2). Let θ_o be the true location. Then the "asymptotic moment generating function" of $n(\hat{\theta}_n - \theta)$ is given by

(7) $$E\{\exp - c_o v n(\hat{\theta}_n - \theta)\}$$

$$= (1 - w_1)\left\{ \frac{s(v,w_1)}{v} - \frac{w_2 - w_1 - s(v,w_1)}{w_2 - w_1 + w_1 e^{-w_2+w_1} - v} \right.$$

$$\left. + \frac{(1 - e^{-w_2+w_1})(w_2 - w_1 + v + s(-v,w_1))}{(1 - e^{-w_2+w_1-s(-v,w_1)})(w_2 - 2w_1 + w_1 e^{-w_2+w_1-s(-v,w_1)} + s(-v,w_1))} \right\}$$

where $s(v,w_1)$ is that solution of the equation $s - w_1 - v + w_1 e^{-s} = 0$ i.e.,

$$s(v,w_1) = \frac{v}{1 - w_1} - \frac{w_1 v^2}{2(1 - w_1)^3} + \frac{w_1(1+2w_1)v^3}{6(1 - w_1)^5} + 0(v^4)$$

$$\to 0 \text{ as } v \to 0$$

while w_1 and w_2 are γ_o/c_o and β_o/c_o (for other notations refer to (2)).

2.3 Estimation of the Size of the Discontinuity

If the location of the discontinuity is known, it is then logical to think of an estimator which is the difference of the estimators of the density function from the right and from the left of Θ– the location of discontinuity. Hence one would expect that such an estimator should tend to zero at every continuity point of the density function and to the size of the discontinuity at Θ. We shall first formally introduce our estimator.

Let $W_n(\cdot)$ be a weight function satisfying

(8) (i) $W_n(y) = \frac{1}{h(n)} W\left(\frac{y}{h(n)}\right)$,

where $h(n)$ satisfies (5).

(ii) $W(\cdot)$ is anti-symmetric at the origin, i.e.,

$$W(-y) = -W(y) \quad y > 0$$

(iii) $\int_0^\infty W(y)\,dy = 1$

(iv) $W(\cdot)$ is a Borel function.

Hence (1) gives

$$(9) \qquad g_n(\Theta) = \sum_{i=1}^{n} \frac{1}{nh(n)} W\left(\frac{x_i - \Theta}{h(n)}\right)$$

$$= \sum_{i=1}^{n} \frac{1}{n} V_{ni}(\Theta) \qquad \text{, say.}$$

By (8·iv), $V_{ni}(\Theta) = (1/h(n))\, W\left(\frac{x_i - \Theta}{h(n)}\right)$, $i=1,2,\ldots,n$, are independent, identically distributed as a random variable

$$\xi_n(\Theta) = (1/h(n))\, W\left(\frac{X - \Theta}{h(n)}\right) .$$

Clearly

while w_1 and w_2 are γ_o/c_o and β_o/c_o (for other notations refer to (2)).

2.3 Estimation of the Size of the Discontinuity

If the location of the discontinuity is known, it is then logical to think of an estimator which is the difference of the estimators of the density function from the right and from the left of Θ - the location of discontinuity. Hence one would expect that such an estimator should tend to zero at every continuity point of the density function and to the size of the discontinuity at Θ. We shall first formally introduce our estimator.

Let $W_n(\cdot)$ be a weight function satisfying

(8) (i) $W_n(y) = \frac{1}{h(n)} W(\frac{y}{h(n)})$,

where $h(n)$ satisfies (5).

(ii) $W(\cdot)$ is anti-symmetric at the origin, i.e.,

$$W(-y) = -W(y) \quad y > 0$$

(iii) $\int_0^\infty W(y)\, dy = 1$

(iv) $W(\cdot)$ is a Borel function.

Hence (1) gives

$$(9) \qquad g_n(\Theta) = \sum_{i=1}^{n} \frac{1}{nh(n)} W(\frac{x_i - \Theta}{h(n)})$$

$$= \sum_{i=1}^{n} \frac{1}{n} V_{ni}(\Theta) \qquad \text{, say.}$$

By (8-iv), $V_{ni}(\Theta) = (1/h(n)) W(\frac{x_i - \Theta}{h(n)})$, i=1,2,...,n, are independent, identically distributed as a random variable

$$\xi_n(\Theta) = (1/h(n)) W (\frac{X - \Theta}{h(n)}) .$$

Clearly

$$(10)\quad (i)\quad E\, g_n(\Theta) = \frac{1}{h(n)} \int W\left(\frac{x-\Theta}{h(n)}\right) f(x)\, dx\ ;$$

$$(ii)\quad \text{Var. } g_n(\Theta) = \frac{1}{nh^2(n)} \left\{ \int W^2\left(\frac{x-\Theta}{h(n)}\right) f(x)dx - E^2 g_n(\Theta) \right\}$$

$$(iii)\quad \text{Cov.}(g_n(\Theta), g_n(\Theta+\delta))$$

$$= \frac{1}{nh(n)} \left\{ \int W(y)\cdot W\left(\frac{\delta}{h(n)} + y\right) f(h(n)y - \Theta)dy - E\, g_n(\Theta)\, E\, g_n(\Theta+\delta) \right\}$$

$$\text{for } \delta > 0$$

By using anti-symmetry of the weight function, it can be easily shown that

$$(11)\quad (i)\quad E\, g_n(\Theta) \sim g(\Theta) + \left(f'(\Theta+) + f'(\Theta-)\right)h(n) \int_0^\infty yW(y)dy + O(h^2(n))$$

where $g(\Theta) = f(\Theta+) - f(\Theta-)$

$$(ii)\quad \text{Var. } g_n(\Theta) \sim \frac{1}{nh(n)} \left(f(\Theta+) + f(\Theta-)\right) \int_0^\infty W^2(y)dy\ ,$$

For (11-i) to be valid we need the assumption of the existence of $f'(x)$ at Θ from left and right, otherwise we will have

$$E\, g_n(\Theta) \sim g(\Theta) + O(h(n))$$

From (11) we can get the asymptotic expression for the mean square error, i.e.,

$$E\left(g_n(\Theta) - g(\Theta)\right)^2 = \text{Var } g_n(\Theta) + (\text{Bias})^2$$

Now let us assume that $h(n) = k\, n^{-\alpha}$, $\alpha > 0$. Then, in order that the variance and the square of bias of $g_n(\Theta)$ are comparable, we must have $\alpha = 1/3$ and then the mean square error is of the order $n^{-2/3}$. Hence it is evident that the estimator $g_n(\Theta)$ of the size of the discontinuity, if the location is known, is consistent in the quadratic mean, i.e.,

(12) $$E\,|g_n(\Theta) - g(\Theta)|^2 \to 0 \quad \text{as } n \to \infty$$

At this point we stop for a moment to make an observation. If $f(x)$ is twice differentiable from the right as well as from the left, we can do better. We can use a "negative-window" to make $\int_0^\infty y\,W(y)dy=0$, which appears in the $E\,g_n(\Theta)$ and in this case, we can get the mean square error to the order $n^{-4/5}$. In general, if one can assume the q th order bounded derivatives of $f(x)$ exist from the left and from the right at Θ then the bias of the estimator can be made of the order $0_p(n^{-q/2q+1})$.

Also if Θ is the location of the discontinuity

(13) $$\operatorname{Cov}\left(g_n(\Theta),\, g_n(\Theta+\delta)\right) \sim \frac{1}{nh(n)}\int W(y)W\left(\frac{\delta}{h(n)}+y\right)f(h(n)y-\Theta)dy$$

Since $|W(y)| \to 0$ as $|y| \to \infty$; for any given $\varepsilon > 0$, we can find $y_o(\varepsilon)$ such that $|W(y)| > \varepsilon$ for $|y| > y_o(\varepsilon)$. Then we can choose n so large that either $|y| > y_o(\varepsilon)$ or $|y + \frac{\delta}{h(n)}| > y_o(\varepsilon)$. Hence

$$\operatorname{Cov}\left(n^{\frac{1}{2}}h^{\frac{1}{2}}(n)\, g_n(\Theta),\, n^{\frac{1}{2}}h^{\frac{1}{2}}(n)\, g_n(\Theta+\delta)\right) \to 0$$

as $n \to \infty$ for $\delta \gtrsim 0$. Similarly we can have a result for $\delta \lesssim 0$

Now we can prove the normality of the estimator, when Θ is known, following Parzen's techniques (6). If we use the representation in $V_{ni}(\Theta)$ for $g_n(\Theta)$ of (9), then $g_n(\Theta)$ is asymptotically normal if and only if

(14) $$nP\left[\frac{|\xi_n(\Theta) - E\xi_n(\Theta)|}{\sigma(\xi_n(\Theta))} \geq \varepsilon\, n^{\frac{1}{2}}\right] \to 0 \quad \text{as } n \to \infty$$

(See Loéve (5), p. 316).

A sufficient condition for (14) to hold is, for some $\delta > 0$,

$$(15)\qquad E\,\frac{|\xi_n(\Theta) - E\xi_n(\Theta)|^{2+\delta}}{n^{\frac{1}{2}\delta}\,\sigma^{2+\delta}(\xi_n(\Theta))} \to 0 \quad \text{as } n \to \infty$$

Now

$$E\,|\xi_n(\Theta)|^{2+\delta} \sim \frac{1}{h^{1+\delta}(n)}\left(f(\Theta+) + f(\Theta-)\right)\int_0^\infty W^{2+\delta}(y)dy$$

Using (11-ii), we have

$$\text{Var}\ \xi_n(\Theta) \sim \frac{1}{h(n)}\left(f(\Theta+)+f(\Theta-)\right)\int_0^\infty W^2(y)dy$$

Under the assumption that $\int_0^\infty W^{2+\delta}(y)\,dy < \infty$, the quantity in (15) tends to zero as $n \to \infty$ and hence the normality of $g_n(\Theta)$. Before we summarize the results of this section, for the case of known Θ, note that we can do better, from the order point of view of the bias if we restrict to the class of the density functions for which $f'(\Theta+) = - f'(\Theta-)$. This is obvious from (11-ii).

<u>Theorem 3</u>: Let $W_n(y)$ satisfy the conditions (8) with further restrictions

$$\text{(i)}\ \int_0^\infty W^{2+\delta}(y)\,dy < \infty$$

$$\text{(ii)}\ \int_0^\infty yW(y)\,dy < \infty$$

Let $f(x)$ possess bounded right and left derivatives at Θ. Let $h(n) = k\,n^{-1/3}$ where k is some constant. Let $g_n(\Theta)$ be defined by (9) with above $h(n)$ and $W(y)$. Then

(a) the mean square error of the estimator $g_n(\Theta)$ is of the order $n^{-2/3}$.

(b) the estimator $g_n(\Theta)$ is consistent estimator of $g(\Theta)$ in the quadratic mean.

(c) $(nh(n))^{\frac{1}{2}}(g_n(\Theta) - Eg_n(\Theta))$ is asymptotically normal with zero mean and variance; $(f(\Theta+)+f(\Theta-))\int_0^\infty W^2(y)\,dy$.

(d) $(nh(n))^{\frac{1}{2}} g_n(\Theta)$ and $(nh(n))^{\frac{1}{2}} g_n(\Theta+\delta)$, $\delta \geq 0$ are asymptotically uncorrolated.

In the case of unknown Θ, the results (a)-(c) will be still true if Θ is replaced by an estimator $\hat{\Theta}_n$ of Θ which is equal to $\Theta + O_p(n^{-\lambda})$, $\lambda > 1/3$. This follows because of the following heuristic argument. If $\hat{\Theta}_n$ satisfies the above condition then the rate of convergence of $\hat{\Theta}_n$ to Θ is faster than that of $g_n(\Theta)$ to $g(\Theta)$. Hence $g_n(\hat{\Theta}_n)$ will start behaving like $g_n(\Theta)$ and the results (a)-(c) will then hold for $g_n(\hat{\Theta}_n)$ as well. A natural question then arises is whether there exists any such estimator $\hat{\Theta}_n = \Theta + O_p(n^{-\lambda})$, $\lambda > 1/3$. There exist many estimators of Θ with $\lambda=1$, to mention one is the maximum likelihood estimator of Θ, studied by Chernoff and Rubin [2]. The proof of existence of such estimators was given by Rubin (9). In the next section we have suggested an estimator of Θ and studied its behavior.

2.4 "Window-Estimator" for the Location of the Discontinuity

Here we have suggested an estimator for the location of the discontinuity of a density function by the location of $\sup_{-\infty<\Theta<\infty} g_n(\Theta)$,

where $g_n(\Theta)$ is defined by (9). For the estimation of the location of the discontinuity, we will need different sets of conditions on the weight function $W_n(y)$. But if the aim is to estimate the location as well as the size of the discontinuity, earlier conditions on $W_n(y)$ will have to be retained. From the definition of $g_n(\Theta)$, it is easy to see that

$$\sup_{\infty > \theta < \infty} g_n(\theta) = \max_{1 \leq i \leq n} g_n(x_{(i)}+) ,$$

where $x_{(i)}$, i=1,2,...,n, are the ordered observations on X.

In section 2.3 we have proved that if θ is the true location of the discontinuity, then $g_n(\theta)$ and $g_n(\theta+\delta)$, for $n \geq n_o(\delta)$, are asymptotically independent. Hence the study of the limiting process would be complete if we determine the distribution of $g_n(\theta+\delta) - g_n(\theta)$, for $\delta > 0$ and $\delta < 0$. We shall treat the case for $\delta > 0$; $\delta < 0$ can be treated similarly. For convenience, let us assume that the true location of the discontinuity is the origin.

We assume all previous conditions on the weight function W(y), together with that W'(y) is absolutely integrable and continuous. We also assume that f(x) is bounded.

$$g_n(0) - g_n(\delta) = \int \left[W_n(x) - W_n(x-\delta)\right] dF_n(x)$$

$$= \left\{ \int_{-\infty}^{0} + \int_{0}^{\delta} + \int_{0}^{\infty} \right\} \left[W_n(x) - W_n(x-\delta)\right] dF_n(x)$$

$$= 2W(0+) \left[F_n(\delta) - F_n(0) \right] - J ,$$

where

$$J = \int_{-\infty}^{\infty} \left[W'_n(x) - W'_n(x-\delta)\right] F_n(x) \, dx$$

$$= \int_{-\infty}^{\infty} \left[W'_n(x) - W'_n(x-\delta)\right] F(x) \, dx$$

$$+ \int_{-\infty}^{\infty} \left[W'_n(x) - W'_n(x-\delta)\right] \left[F_n(x) - F(x)\right] dx$$

Now, under the assumptions made about the density function and the weight function, it can be shown that the second integral on the right hand side of J is $0_p(\delta)$.

First integral on the right hand side of J

$$= \int_{-\infty}^{\infty} W'(x) \; \left(F(x) - F(x+\delta)\right) dx$$

$$= \left(\int_{-\infty}^{-\delta} + \int_{-\delta}^{0} + \int_{0}^{\infty} \right) W_n'(x) \; \left(F(x) - F(x+\delta)\right) dx$$

$$= 0_p(\delta) + \delta \, W_n(0+) \; \left(f(0+) - f(0-) + o(1)\right)$$

Hence, we get

$$(21) \quad g_n(0) - g_n(\delta) \sim 2 \, W_n(0+) \; \left(F_n(\delta) - F_n(0)\right)$$
$$- \; \delta W_n(0+) \; \left(f(0+) - f(0-) + o(1)\right)$$

Note, for sufficiently large n, $n \left(F_n(\delta) - F_n(0)\right)$ is approximately distributed as a Poisson random variable with parameter $\left(F(\delta)-F(0)\right)$.

From (21) we can, therefore, conclude that the window estimator of the location of the discontinuity is the location of the maximum of

$$\left(F_n(\alpha) - F_n(\alpha_o)\right) - \frac{f(\alpha_o+) - f(\alpha_o-)}{2} (\alpha - \alpha_o)$$

This estimator is very similar to the maximum likelihood estimator studied by Chernoff and Rubin [2], the only difference being that, instead of $\frac{1}{2}\left(f(\alpha_o+) - f(\alpha_o-)\right)$, they have a coefficient

$\left(f(\alpha_o+) - f(\alpha_o-)\right) / \; \log f(\alpha_o+) - \log f(\alpha_o-)$. We will summarize the above results.

Theorem 4: Let $W_n(y)$ be a weight function satisfying the following conditions, in addition to conditions of Theorem 3.

(i) $W'(y)$ is absolutely integrable.

(ii) $W'(y)$ is continuous.

Then, for the class of bounded density functions, we have,

(a) the bias of the window estimator for the location of the discontinuity is $0_p(n^{-1})$.

(b) the asymptotic moment generating function of $n(\hat{\theta}_n-\theta)$, where θ is the true location and $\hat{\theta}_n$ is the window estimator, is given by Theorem 2 with $c_o = \frac{1}{2}(f(\theta+) - f(\theta-))$.

(c) the window estimator, $g_n(\hat{\theta}_n)$, for the size of the discontinuity, possesses the properties (a)-(c) of Theorem 3.

CHAPTER III

SPECIAL CASE OF THE DISTRIBUTIONS OF THE MEDIAN AND THE MEAN

3.1 Introduction

The problem of this chapter originated from the following consideration. Let us consider the problem of estimating the location of the discontinuity of a density function. It is then natural to expect that the asymptotic study of such a problem, under suitable regularity conditions, would lead to a stochastic process which behaves nicely on both sides of the true location of the discontinuity. This was first noticed by Rubin [9]. He then studied the maximum likelihood estimator of the non-stationarity point of a stochastic process. The study of section 4, Chapter II, also leads to a Gaussian process with independent increments and with a non-stationarity point. While concluding the above paper the author passed a remark that if one is interested in inference with an invariant loss function, there will exist an optimal invariant procedure which consists of assuming a uniform a-priori distribution on the parameter. For clarity, let $X(t)$, $-\infty < t < \infty$, be a stochastic process with t_o- as a non-stationarity point, i.e., the behavior of $X(t)$ for t $t < t_o$ and $t > t_o$ are identical and the process $X(t)$ is stationary in any interval contained in $(-\infty, t_o)$ or (t_o, ∞). Let the origin, for simplicity, be the true non-stationarity point of $X(t)$. Let $L(t)$ be the likelihood ratio of the process at t against $t=0$. Assume the Lebsgue measure on the real line as a-priori distribution on the non-stationarity point. Then $L(t)$ is proportional to the a posteriori distribution of t-the non-stationarity point.

Recently, Rao and Rubin (7) gave a necessary and sufficient condition for a Gaussian process to be the log-likelihood process of Gaussian process. We are going to consider a special case of it. We assume that $X(t)$, $-\infty < t < v$ is a Gaussian process such that the likelihood ratio $L(t)$ of it has following special form

(1) $$L(t) = \exp\left(W(t) - \frac{1}{2}|t|\right), \quad -\infty < t < \infty,$$

where $W(t)$ is a Gaussian process with independent increments and with,

(2) (i) $W(0) = 0$

(ii) $E\,W(t) = 0\,, \quad -\infty < t < \infty,$

(iii) $\text{Cov.}\left(W(t_1), W(t_2)\right) = \delta(t_1,t_2)\min(|t_1|,|t_2|),$

$$\text{where}\quad \delta(t_1,t_2) = 1 \quad \text{if} \quad t_1 t_2 \geq 0$$
$$= 0 \text{ otherwise};$$

i.e., $W(t)$ is a standard Wiener process. In the light of Rubin's comment, we consider in section 3.3, the invariant loss function - the absolute error and in section 3.4, the squared error loss function. Then for the stochastic process (1), the classical theorem will give the median and the mean of the distribution, which is proportional to (1), as the best invariant estimator of the parameter t. In (3.3), a computable expression for the distribution of the median of the distribution which is proportional to (1), is obtained. The distribution of the mean of the distribution, which is proportional to (1), has not been obtained yet. Section (3.4) includes some results on this problem which might lead to the solution of the problem.

3.2 Preliminaries

3.2.1 Notation

We introduce further notations in terms of L(t), where L(t) is defined by (1) and (2).

(3) (i) $X^{-} = \int_{-\infty}^{0} L(t)\,dt$

(ii) $X^{+} \equiv X = \int_{0}^{\infty} L(t)\,dt$

(iii) $Y = \int_{0}^{T} L(t)\,dt$

(iv) $Z = \int_{T}^{\infty} L(t)\,dt$

It then clearly follows that,

(4) $$P\left(\text{median} \geq T\right)$$

$$= \int_{-\infty}^{\infty} P\left(X^{-} + Y \leq Z \mid W(T) - \tfrac{1}{2}T = u\right)$$

$$\cdot \frac{1}{(2\pi T)^{\frac{1}{2}}} \exp -\left[\frac{1}{2T}\left(u + \tfrac{1}{2}T\right)^2\right] du.$$

Let,

(5) $$P\left(X^{-} + Y \leq Z \mid W(T) - \tfrac{1}{2}T = u\right) = g(T,u)$$

In 3.3 we shall derive a computable expression for g(T,u).

3.2.2 Some Well-known Properties of W(t)

We shall quote some well known properties of W(t), defined by (1) and (2), which are used in the derivation.

(6) Extrapolation Property:

$$W(t_1+t_2) = W(t_1) + W'(t_2), \text{ for } t_1 t_2 \geq 0,$$

where $W(t_1)$ and $W'(t_2)$ are independent Wiener processes with the same structure as $W(t)$.

(7) Interpolation Property:

The conditional density of $W(t)$, $t_1 < t < t_2$, $t_1 t_2 \geq 0$, given $W(t_1) = A$ and $W(t_2) = B$, is a normal density with mean

$$\text{(i)} \quad A + \frac{B - A}{t_2 - t_1}(t - t_1)$$

and the variance

$$\text{(ii)} \qquad \frac{(t_2-t)(t-t_1)}{(t_2-t_1)} .$$

We shall now prove a lemma which is not well known. Rubin (10) proved a related result with a different approach and Fox and Rubin (3) proved the same result for general stochastic processes.

Lemma 1: Let X be as defined in (3-ii). Then, the distribution of $1/X$ is an exponential distribution with scale factor two, that is, the Laplace transform of $1/X$ is given by

(8) Laplace transform of $1/X = E\left(\exp - (\lambda/X)\right)$

$$= 2/(\lambda + 2) , \qquad \lambda \geq 0,$$

where "E" stands for "Expectation" symbol.

Proof: Let $\varepsilon > 0$, then write

$$X = \int_0^\epsilon e^{W(t) - \frac{1}{2}t}dt + \int_\epsilon^\infty e^{W(t) - \frac{1}{2}t}dt$$

$$= \int_0^\epsilon e^{W(t) - \frac{1}{2}t}dt + \int_0^\infty e^{W(t+\varepsilon) - \frac{1}{2}(t+\varepsilon)}dt$$

$$= \int_0^\epsilon e^{W(t) - \frac{1}{2}t}dt + e^{W(\varepsilon) - \frac{1}{2}\varepsilon}\int_0^\infty e^{W'(t) - \frac{1}{2}t}\, dt, \text{ by (6)},$$

where $W(\varepsilon)$ and $W'(t)$ are independent.

Hence

(9) $$X \sim \varepsilon + L(\varepsilon)\, X',$$

where $L(\varepsilon) = e^{W(\varepsilon) - \frac{1}{2}\varepsilon}$ and X, X' are independent.

... If $\phi(\lambda) = E\left[\exp - (\lambda/X)\right]$, ϕ_λ and $\phi_{\lambda\lambda}$ are the first and the second derivatives of $\phi(\lambda)$ with respect to λ, then by substituting the expression (9) in $\phi(\lambda)$, by expanding the terms under the expectation sign without $\exp -(\lambda/X')$, by taking the expectation and omitting all the terms of order equal or greater than ε^2, it can be shown that $\phi(\lambda)$ satisfies the following differential equation

(10) $$\frac{1}{2}\lambda(\lambda+2)\, \phi_{\lambda\lambda} + \lambda\phi_\lambda = 0$$

with the obvioua boundary conditions such as

(i) $\phi(0) = 1$

(ii) $\phi(\lambda)$ is a non-negative, non-increasing function of λ, which tends to zero as λ approaches infinity.

Now, it can be easily verified that (8) is the required solution of (10), satisfying above conditions.

Remark: Rubin [10] showed that $X/(X^- + X)$ is uniformly distributed over $[0,1]$ by considering moment-relations. This result will follow immediately from the above lemma and the following well-known fact about exponential random variables, which we shall quote as a result.

Result 1: If X and Y are independent exponential random variables, then $X/(X + Y)$ (or $Y/(X + Y)$) is uniformly distributed over $[0,1]$ and is independent of $X + Y$ (The independence of $X/(X + X^-)$ from $X+X^-$ is not true for general processes.).

With these preliminaries we proceed to find $g(T,u)$ as defined by (5).

3.3 Distribution of the Median

Let us start by looking back to the definitions of the random variables Y and Z. Note, we could replace $W(t) - \frac{1}{2}t = V(t)$, where V(t) is a Wiener process with mean $-\frac{1}{2}t$, $t > 0$ and same variance-covariance structure as W(t). Then

$$Y = \int_0^T e^{V(t)}dt$$

$$Z = \int_T^\infty e^{V(t)}dt$$

$$= e^{V(T)}\, Z' \qquad \text{- by (6),}$$

where V(T) and Z' are independent.

Now, by using the interpolation property (7), we have,

$$\text{(11)} \qquad E\, V(t) = \frac{t\, u}{T}, \qquad 0 \le t \le T\, ;$$

$$\text{Var } V(t) = (T-t)\,\frac{t}{T}, \qquad 0 \le t \le T\, .$$

Let us introduce some notation.

$$\text{(12)} \quad F(a,T,u) = P\left(Y \le a \;\middle|\; T, V(T) = u \right), \qquad a \ge 0\, ,$$

$$f(a,T,u) = \frac{\partial}{\partial a} F(a,T,u).$$

Lemma 2: F(a,T,u), defined by (12), satisfies the following forward partial differential equation,

$$\text{(13)} \quad \left(-1-\frac{u}{T}a+\frac{1}{2}a\right)F_a - F_T - \frac{u}{T}F_u + \frac{a^2}{2}F_{aa} + a\,F_{au} + \frac{1}{2}F_{uu} = 0\, ,$$

where the subscripts stand for the corresponding partial differentials of F(a,T,u).

Proof: The techniques used here are the same as used in Lemma 1. We write,

$$Y = \int_0^{\epsilon} e^{V(t)}dt + \int_{\epsilon}^{T} e^{V(t)}dt$$

$$\sim \varepsilon + e^{V(\varepsilon)} V^*$$

where $V(\varepsilon)$ and V^* are independent.

$$\therefore F(a,T,u) \sim E\,P\,\left(V^* \leq ae^{-V(\varepsilon)} - \varepsilon \,\middle|\, T-\varepsilon,\ V(T-\varepsilon) = u - V(\varepsilon)\right)$$

This leads to equation (13).

Lemma 3: $F(a,T,u)$ also satisfies the following backward partial differential equation.

(14) $$- e^u F_a - F_T - \frac{u}{T} F_u + \frac{1}{2} F_{uu} = 0 .$$

Proof: Here the only difference in the proof is, that we write

$$Y = \int_0^{T-\epsilon} e^{V(t)}dt + \int_{T-\epsilon}^{T} e^{V(t)}dt$$

and proceed as before to get (14).

If (14) is subtracted from (13) we get the following corollory;

Corollory 1: $F(a,T,u)$ satisfies the following partial differential equation

(15) $$(e^u - 1 - \frac{u}{T} a + \frac{1}{2} a)\, F_a + \frac{1}{2} a^2 F_{aa} + a\, F_{au} = 0.$$

The above equation can be written in $f(a,T,u)$.

(16) $$(e^u - 1 - \frac{u}{T} a + \frac{1}{2} a) f + \frac{1}{2} a\, f_a + a\, f_u = 0.$$

Lemma 4: The general solution of the partial differential equation (16) is

(17) $$f(a,T,u) = \exp\left(\frac{1}{2}\frac{u^2}{T} - \frac{1}{2} u - \frac{4\cosh \frac{1}{2} u}{a\, e^{-\frac{1}{2}u}} + \psi(ae^{-\frac{1}{2}u})\right)$$

where ψ depends on a and u only through $ae^{-\frac{1}{2}u}$.

Proof: The factor, $\exp\left(\psi\ (ae^{-\frac{1}{2}u})\right)$, corresponds to the general solution of the homogeneous equation of (16). The remaining factors of (17) correspond to a particular solution of (16) which could be obtained by inspection as follows:

(a) $\exp\left(\frac{1}{2}\frac{u^2}{T} - \frac{1}{2}u\right)$ is the factor in the solution which corresponds to the term $(-\frac{u}{T}a + \frac{1}{2}a)$ in the coefficient of $f(a,T,u)$ in (16).

(b) $\exp\left[\frac{-4\cosh\frac{1}{2}u}{a\ e^{-\frac{1}{2}u}}\right]$ is included in the particular solution to correspond to $(e^u - 1)\ f(a,T,u)$. To obtain this, one makes the following simple observation: $\exp\left(\frac{K(u)}{a}\right)$ can correspond to $(e^u - 1)\ f(a,T,u)$, if $K(u)$ is a solution of the ordinary differential equation

$$(e^u - 1) + K'(u) - \frac{1}{2}K(u) = 0\ ,$$

where $K'(u)$ stands for the differential of $K(u)$.

$\frac{-4\cosh\frac{1}{2}u}{a\ e^{-\frac{1}{2}u}}$ is a solution of the above equation.
This completes the proof.

Corollery 2: The density function, $f(a,T,u)$, is given by (17), where ψ is a function of $ae^{-\frac{1}{2}u}$ only (possibly depending of T) such that

$$(18)\qquad \int_0^\infty \exp\left(\frac{1}{2}\frac{u^2}{T} - \frac{1}{2}u - \frac{4\cosh\frac{1}{2}u}{a\ e^{-\frac{1}{2}u}} + \psi(ae^{-\frac{1}{2}u})\right) da = 1.$$

Corollery 3: If $Y_1 = e^{\frac{1}{2}u}Y^{-1}$, then the Laplace transform of Y_1, viz., $E\left(\exp - \lambda Y_1\right) = \phi(\lambda)$, is given by

$$(19)\qquad \phi(\lambda) = \exp\left(\frac{1}{2}\frac{u^2}{T}\right)\ \exp - \frac{1}{2}T^{-1}\left(\text{arc}\cosh^2(\cosh\frac{1}{2}u + \frac{1}{4}\lambda)\right).$$

We are now in a position to find $g(T,u)$ as defined by (5).

Lemma 5:

$$(20)\quad g(T,u) = k - \int_0^k \exp\left(\frac{1}{2}\frac{u^2}{T}\right) \exp - \frac{1}{2T}\left(\text{arc cosh}^2 \frac{1}{2}u + \frac{1}{2}c\right)$$

$$\cdot \left[1 + \frac{1}{2T} \frac{\text{arc cosh}(\cosh \frac{1}{2}u + \frac{1}{2}c)}{\{(\cosh\frac{1}{2}u+\frac{1}{2}c)^2 - 1\}^{\frac{1}{2}}}\right] da ,$$

where $k = 1/(1+e^{-u})$, $c = \frac{e^{\frac{1}{2}u}}{a} - \frac{e^{-\frac{1}{2}u}}{1-a}$.

Proof: Note we can write $g(T,u)$ as

$$g(T,u) = P\left(X^- + Y \le e^u Z' \mid V(T) = u\right) ,$$

where X^- and Z' are identically distributed.

For the notational convenience, we shall omit the condition $V(T)=u$ and the prime of Z in the above definition of $g(T,u)$.

Let

$$W = \frac{1}{Z} + \frac{1}{X^-}$$

$$A = X^-/(Z + X^-) ,$$

then, from the result 1 and Lemma 1, we know that A is a uniform random variable on $(o,1)$, W is a Gamma random variable with parameter two and W,A are independent.

$$\therefore g(T,u) = P\left(e^{-\frac{1}{2}u}Y \le e^{\frac{1}{2}u} Z - e^{-\frac{1}{2}u} X^-\right)$$

$$= P\left(\frac{W}{Y_1} \le \frac{e^{\frac{1}{2}u}}{A} - \frac{e^{-\frac{1}{2}u}}{1-A}\right) ,$$

where Y_1 is defined in Corollary 3.

$$g(T,u) = \int P\left(\frac{W}{Y_1} \quad \frac{e^{\frac{1}{2}u}}{a} - \frac{e^{-\frac{1}{2}u}}{1-a}\right) da ,$$

where the range of integration is such that $\frac{e^{\frac{1}{2}u}}{a} - \frac{e^{-\frac{1}{2}u}}{1-a} \ge 0$, i.e., $k = 1/(1 + e^{-u}) \geq a \geq 0$

$$\therefore g(T,u) = \int_0^k P\left(\frac{W}{Y_1} \leq c\right) da \quad ,$$

where k is as defined above and

$$c = \frac{e^{\frac{1}{2}u}}{a} - \frac{e^{-\frac{1}{2}u}}{1-a}$$

$$= k - \int_0^k \int_0^\infty P\left(W \geq cy_1\right) \quad f(u_1)\, dy_1\, da \ ,$$

where $f(y_1)$ is the density function of the random variable Y_1, whose Laplace transform is given by (19).

Now, using

(i) $P\left(W \geq cy_1\right) = (1 + 2cy_1)\, e^{-2cy_1}$

and

(ii) (19) ,

one can arrive to (20).

We now summarize the results.

Theorem 1: Let t be the translation parameter of a Gaussian process $X(t)$, $-\infty < t < \infty$, with Lebsgue measure on the real line as the a-priori distribution for t. Let the process $X(t)$ be such that the likelihood ratio $L(t)$, at t against $t=0$, of it admits the following representation

(21) $L(t) = \exp\left(W(t) - \frac{1}{2}|t|\right)$, $-\infty < t < \infty$,

where $W(t)$ is a standard Wiener process. Then the distribution of the median of the a posteriori distribution, which is proportional to (21), is given by

$$P\left(\text{median} \geq T\right)$$

$$= 2\, e^{-T/8} \int_0^\infty e^{-\frac{1}{2} u^2/T} \operatorname{sech} \frac{1}{2} u \; du$$

$$- e^{-T/8} \int_{-\infty}^{\infty} e^{-\frac{1}{2}u} \int_0^k \exp - \frac{1}{2} \frac{1}{T} \left(\text{arc}\cosh^2 \left(\cosh \frac{1}{2}u + \frac{1}{2}c\right)\right)$$

$$\cdot \left\{1 + \frac{c}{2T} \; \frac{\text{arc}\cosh\left(\cosh \frac{1}{2}u + \frac{1}{2}c\right)}{\{(\cosh \frac{1}{2}u + \frac{1}{2}c)^2 - 1\}^{\frac{1}{2}}}\right\} da\; du \; ,$$

where $k = 1/(1 + e^{-u})$

$$c = \frac{e^{\frac{1}{2}u}}{a} - \frac{e^{-\frac{1}{2}u}}{1-a} \quad .$$

3.4 Some Results Connected with the Distribution of the Mean

In this section we shall give a few results connected with the distribution of the mean of the distribution, which is proportional to (1). We shall start with the definition of $L(t)$ given by (1). We further introduce some notation. Let,

$$Y^+ = Y = \int_0^\infty tL(t)\; dt \; ,$$

$$Y^- = \int_{-\infty}^0 |t| \; L(t)\; dt \; ,$$

$$X^+ = X = \int_0^\infty L(t)\; dt \; ,$$

$$X^- = \int_{-\infty}^0 L(t)\; dt \; ,$$

$$\mu = \frac{Y^+ - Y^-}{X^+ + X^-} \quad .$$

Y has a similar representation as given in (9).

(22) $\quad Y \sim (\epsilon X' + Y') \, L(\epsilon)$

One can easily see the difficulty in obtaining the distribution of Y, which is the presence of X' on the right hand side of (22).

3.4.1. Various Moment Relations

We start with the representations (9) and (22) for X and Y, respectively. Let,

$$\mu_{\alpha\beta} = E \, \frac{Y^{\alpha}}{X^{\beta}} \qquad E \, \frac{L^{\alpha-\beta}(\epsilon)}{X'^{\beta}} \left\{ Y'^{\alpha} + \alpha\epsilon X' \, Y'^{\alpha-1} \right\} \left(1 - \beta \frac{\epsilon}{X'}\right) ,$$

which leads to

$$\mu_{\alpha\beta} \sim \{E \, L^{\alpha-\beta}(\epsilon)\} \, \{\mu_{\alpha\beta} + \alpha\epsilon\mu_{\alpha-1\,\beta-1} - \beta\epsilon\mu_{\alpha\beta+1}\} .$$

Then, using $E \, L^{k}(\epsilon) = e^{\frac{1}{2}k(k-1)\epsilon}$, we can get

(23) $\quad \frac{1}{2}(\alpha - \beta)(\alpha - \beta - 1)\mu_{\alpha\beta} + \alpha\mu_{\alpha-1\beta-1} - \beta\mu_{\alpha\beta+1} = 0.$

If $\alpha < 0$, $\beta < 0$, i.e., if $\mu_{\alpha\beta} = E \, \dfrac{X^{\beta}}{Y^{\alpha}}$, then

(23) will become

(24) $\quad \frac{1}{2}(\beta - \alpha)(\beta - \alpha - 1)\mu_{\alpha\beta} - \alpha\mu_{\alpha+1\,\beta+1} + \beta\mu_{\alpha\beta-1} = 0 .$

It is necessary to remark here that $\mu_{\alpha\beta}$ exists if and only if $\alpha - \beta < 1$.

3.4.2 Some Partial Differential Equations Associated with the Problem

We have obtained the distribution of 1/X in Lemma 1 of this chapter. In the following, we have given a few partial differential equations which could be derived by the same techniques as used in Lemma 1 of this chapter. The derivations have been omitted to avoid repetitions.

Let $\phi(s,t) = E \quad \exp - (\frac{s}{X} + \frac{tY}{X})$, then $\phi(s,t)$ satisfies the following partial differential equation (p.d.e.)

(25) $$\frac{1}{2} s(s+2)\, \phi_{ss} + s\phi_s + t\phi_{st} = t\phi ,$$

where $\phi_{ss} = \frac{\partial^2}{\partial s^2}$, etc.

Here we know by Lemma 1 that $\phi(s,0) = 2/(s+2)$.

This equation can be solved theoretically. Let

(26) $$\phi(s,t) = \sum_{k=0}^{\infty} \psi_k(s) \frac{t^k}{t!} \qquad \text{with } \psi_0(s) = 2/(s+2) .$$

Using (26) in (25), we get

$$\frac{1}{2} s(s+2)\, \psi_k''(s) + (s+k)\, \psi_k'(s) = k\, \psi_{k-1}(s)$$

with the standard notation.

This equation can be rewritten in the following form by multiplying the requisite integrating factor.

(27) $$\frac{d}{ds} \frac{s^k}{(s+2)^{k-2}} \psi_k'(s) = 2k \frac{s^k}{(s+2)^{k-1}} \psi_{k-1}(s)$$

Since $\psi_k(s) \to 0$ as $s \to \infty$, we have

(28) $$\psi_k(s) = - \int_s^{\infty} \frac{(u+2)^{k-2}}{u^k} \int_0^u 2k\, t^{k-1}(t+2)^{-k+1} \psi_{k-1}(t)\, dt\, du .$$

Since we know $\psi_0(s)$, theoretically we should be able to evaluate all $\psi_k(s)$. But unfortunately this cannot be done in closed form. Even for $\psi_1(s)$, one gets the following complicated formula

(29) $$\psi_1(s) = -2 \qquad \frac{1 - (s/s+2)^n}{n^2}$$

(cont) $$= 4 \int_s^{\infty} \frac{\log(1+\frac{1}{2}t)}{t(t+2)}\, dt$$

With tedious evaluations of integrals one can get the following expression for $\psi_2(s)$

$$\psi_2(s) = -4\psi_1(s)\left(1 + \frac{s+2}{s} \log (2/s+2)\right)$$

$$-8 \frac{s}{(s+2)^2} \log^2(2/s+2) + \frac{32\, s}{(s+2)^2} \zeta(3)$$

$$-2 \frac{(s+2)}{s} \int_s^{\infty} \frac{\log^2(1+\frac{1}{2}t)}{t(t+2)}\, dt ,$$

where $\zeta(3) = \sum_{n=1}^{\infty} \frac{1}{n^3}$.

To obtain the rest of $\psi_k(s)$'s is very difficult. But it is possible to evaluate $\psi_1(0)$, $\psi_2(0)$ and $\psi_3(0)$ with tremendous amount of evaluations of integrals which run for ten to twelve pages. Also, we get, then, with less efforts, the quantities $\psi_1'(0)$, $\psi_2'(0)$, $\psi_3'(0)$. These are, with interpretations, as follows:

(30) (i) $-\psi_1(0) = E\,\frac{Y}{X} = \frac{\pi^2}{3}$

(ii) $\psi_2(0) = E\,\frac{Y^2}{X^2} = 16\,\zeta(3)$

(iii) $-\psi_3(0) = E\,(\frac{Y}{X})^3 = 16\,\frac{\pi^4}{5} + 48 - 8\pi^2$

and

(31) (i) $\psi_1'(0) = E\,\frac{Y}{X^2} = 1$

(ii) $-\psi_2'(0) = E\,\frac{Y^2}{X^3} = 2\,\frac{\pi^2}{3}$

(iii) $\psi_3'(0) = E\,\frac{Y^3}{X^4} = \frac{8}{3}\zeta(3)$

(iv) $-\psi_4'(0) = E\,\frac{Y^4}{X^5} = \frac{8}{5}\pi^4 + 24 - 4\pi^2$

One can obtain additional moments by using the moment relation (22). But the p.d.e. (25), otherwise, does not seem to be of any help.

It is possible to combine the moment relations (22)-(24) and the moments in (30) and (31) to obtain some more moments which were not otherwise obtainable.

In (23) put $\alpha = \beta - 1$, multiply by $\frac{(-t)^\alpha}{\alpha\,!}$ and sum for $\alpha = 1, 2, \ldots$. This will lead to the following equation.

(32) $$E\left[\frac{1}{X^2} e^{-t\frac{Y}{X}} - E\frac{1}{X^2}\right]$$

$$= E\left[\frac{1}{Y} e^{-t\frac{Y}{X}} - E\frac{1}{Y}\right] + \frac{1}{t}\left[E\frac{X}{Y^2} e^{-t\frac{Y}{X}} - tE\frac{1}{Y} - E\frac{X}{Y^2}\right]$$

Let $t \to \infty$ in (32) . This gives

(33) $$E\,\frac{1}{Y} = \tfrac{1}{2}E\,\frac{1}{X^2} = \tfrac{1}{4} \; .$$

Now put $\alpha = \beta = 1, 2$ in (23), then

$$E\,\frac{1}{Y} = E\,\frac{X^2}{Y^2} = \tfrac{1}{4} \; ,$$

$$E\,\frac{X}{Y^2} = E\left(\frac{X}{Y}\right)^3 \; .$$

Now put $\beta = 0$, $\alpha = 1$ in (23) to get $E\,\frac{1}{Y} = E\,\frac{X}{Y^2}$.

Therefore, combining all these results we get

(34) $$E\,\frac{1}{Y} = E\,\frac{X}{Y^2} = E\,(\frac{X}{Y})^2 = E(\frac{X}{Y})^3 = \frac{1}{4}\ !$$

Another interesting fact is - if $\alpha(u)$ is any differential function of u, it can be shown that

(35) $$E\,\alpha'(u)\,e^{-\alpha(u)} = E\,\frac{Y}{X^2}\,\alpha'(u)\,e^{-\alpha(u)} \quad ,$$

where $u = \frac{Y}{X}$.

From this we can deduce

(36') $$E\,e^{-tY/X} = E\,\frac{Y}{X^2}\,e^{-tY/X}$$

Also

(36") $$E\,e^{-tX/Y} = E\,\frac{Y}{X^2}\,e^{-tX/Y} \quad .$$

(36") gives an additional moment, viz.,

(37) $$E(X/Y) = E\,Y^{-1} = \frac{1}{4}\ .$$

Another possible p.d.e. is for $\phi(s,t) = E\,e^{-s/X - t\,X/Y}$,

(38) $$\frac{1}{2}\,s(s+2)\,\phi_{ss} + s\phi_s - t\phi_{st} + t\phi_{tt} = 0\ .$$

So again, if we write $\phi(s,t)$ in the series form of (25), the recurrence relation for $\psi_k(s)$ is

(39) $$\frac{1}{2}s(s+2)\,\psi_k''(s) + (s-k)\,\psi_k'(s) = -k(k+1)\,\psi_{k+1}(s)\ .$$

Therefore, theoretically, if $\psi_1(s) = E\,(X/Y)\,e^{-s/X}$ is known, the solution of (38) is known; and note that this (39) would be easier to handle than (27). Unfortunately $\psi_1(s)$ is as difficult to find as the solution of (25).

3.4.3 Another Possible Approach

The following approach may work. Introduce

(40) $$Z(\lambda) = \int_0^\infty \exp\left(W(t) - (\tfrac{1}{2} + \lambda)t \right) dt \quad , \quad \lambda \geq 0$$

with W(t) as defined earlier.

Then Z(0) = X and Z'(0) = Y

Clearly

(i) $Z(\lambda_1) \leq Z(\lambda_2)$, for $\lambda_1 \geq \lambda_2$,

(ii) $Z(\lambda)$ is ∞ - differentiable with respect to λ.

Hence,

(41) $$P\left(Z'(0) \leq a\right) = \lim_{\lambda \to 0} \left(P \; \frac{Z(0) - Z(\lambda)}{\lambda} \leq a\right) .$$

Therefore, it is sufficient to know the joint probability distribution of Z(0) and Z(λ). It can be easily proved that 2/Z(λ) is a gamma random variable with parameter (2λ+1). Hence, if

$\alpha(s,t) = E \exp - \left(\frac{s}{Z(0)} - \frac{t}{Z(\lambda)} \right)$, we obtain

(42) $$\frac{1}{2}\left(s(s+2)\alpha_{ss} + t(t+2)\alpha_{tt} + 2st\alpha_{st}\right) + s\alpha_s + (\lambda+t)\,\alpha_t = 0.$$

The advantage of this equation over (25) seems to be the knowledge of $\alpha(0,t) = 2^{2\lambda+1} / (2+t)^{2\lambda+1}$, whereas we have no idea about $\psi(0,t)$ in (25). Again this p.d.e. is solved if $E\frac{1}{Z(\lambda)} \exp\left(-s/Z(0)\right)$ is known.

Similar moment relations and other p.d.e. can be obtained for various combinations of Z(0) and Z(λ). But the problem remains as difficult as earlier.

3.4.4 General Outline of a Possible Method of Solving the Problem

In all the previous subsections (3.4.1)-(3.4.3), the aim was to find the joint characteristic function of Y and X, directly or indirectly. All the while the idea had been to use the quite overlooked formula for the distribution of the ratio of random variables, when the joint characteristic function is known. This formula was proved by Gurland (4). We shall first quote his formula. Let $X_1,\ldots,X_n$ have the joint distribution function $F(x_1,\ldots,x_n)$ with the characteristic function $\psi(t_1,\ldots,t_n)$. Let us consider a new random variable $Y = \sum_{i=1}^{n} a_i X_i / \sum_{n=1}^{n} b_i X_i$. If $P\left(\sum_{i=1}^{n} b_i X_i > 0\right) = 1$, and if $G(x)$ denotes the distribution function of Y, then

$$(43) \quad G(x+) + G(x-) = 1 - \frac{1}{\pi i} \oint \frac{(t(a_1-b_1x),\ldots,t(a_n-b_nx))}{t}\, dt$$

We are interested in the distribution of

$$\mu = \frac{Y_1-Y_2}{X_1+X_2}, \quad X_1 \geq 0, \quad X_2 \geq 0 ,$$

where (X_1,Y_1) and (X_2,Y_2) are independent, identically distributed. So, if $\phi(s,t)$ is the joint characteristic function of (X_1,Y_1), the characteristic function of (X_1,Y_1,X_2,Y_2) is ϕ^2. Then by (43) we can write down the distribution of μ. Let G stand for the required distribution, then

$$(44) \quad 1 - G(x) = \frac{1}{2\pi i} \oint \frac{\phi(-t,xt)\phi(-t,-xt)}{t}\, dt .$$

This should solve the problem.

LIST OF REFERENCES

(1) H. Chernoff (1964): Estimation of the Mode Ann. Stat. Math.; Vol. 16; pp 31-41.

(2) H. Chernoff and H. Rubin (1956): The Estimation of the Location of a Discontinuity in Density, Proc. 3rd Berk. Symp. on Math. Stat. and Prob.; Vol. 1, pp. 19-37.

(3) M. Fox and H. Rubin (1966): Some Properties of Quantile Estimation, Zeitschrift Fur Wahrscheinlichkeitstheorie Und Verwandte Gebeite, Band 4, pp. 353-353.

(4) J. Gurland (1948): Inversion Formulae for the Distribution of Ratios. Ann. Math. Stat.; Vol. 19, pp. 228-237.

(5) M. Loéve (1960) Probability Theory, 2nd Ed., Princeton, N.J., Van Nostrand.

(6) E. Parzen (1962): On Estimation of a Probability Density Function and Mode, Ann. Math. Stat.; Vol. 33, pp. 1056-1076.

(7) B.L.S.P. Rao and H. Rubin (1968): A Property of the Likelihood Ratio Process for Gaussian Processes. Ann. Stat. Math.; Vol. 20, pp. 311-314.

(8) M. Rosenblatt (1956): Remarks on Some Non-parametric Estimation of a Density Function, Ann. Math. Stat.; Vol. 16, pp. 31-41.

(9) H. Rubin (1961): The Estimation of Discontinuities in Multivariate Densities and Related Problems in Stochastic Processes. Proc. Berk. Symp. on Math. Stat. and Prob. ; Vol. 1; pp. 563-574.

(10) H. Rubin (1960): Application to Stochastic Processes of a Unique Property of Rectangular Distribution, Ann. Math. Stat; Vol. 31, Report No. 13, p. 528.

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Purdue University	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

On Some Estimation and Distribution Prbblems

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical Report

5. AUTHOR(S) (Last name, first name, initial)

Paranjape, Suresh R.

6. REPORT DATE July 1969	7a. TOTAL NO. OF PAGES 34	7b. NO. OF REFS 10

8a. CONTRACT OR GRANT NO. N 00014-67-0226-0008 b. PROJECT NO. c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) Mimeo Series No. 194 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. AVAILABILITY/LIMITATION NOTICES

Distribution of this document is unlimited

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Office of Naval Research Washington, D.C.

13. ABSTRACT

The report deals with two estimation and two distribution problems. The ''window-estimators'' for the size and the location of a single, simple discontinuity of a density function have been studied. The distribution of the median of a special type of Gaussian process has been obtained. Some partial results on the distribution of the mean of the same type of Gaussian process are included. The special type of Gaussian process which is treated here, is assumed to have the following representation for its likelihood-ratio, L(t);

$$L(t) = \text{esp}[W(t) - \frac{1}{2}|t|] \quad -\infty < t < \infty$$

where W(t) is a standard Wiener process.

DD FORM 1 JAN 64 1473

Unclassified

Security Classification

Unclassified

Security Classification

14. KEY WORDS	LINK A ROLE	LINK A WT	LINK B ROLE	LINK B WT	LINK C ROLE	LINK C WT
Discontinuity						
size						
location						
process						
mean						
median						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through ________________."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through ________________."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through ________________."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS), (S), (C), or (U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, roles, and weights is optional.