ON A PROBLEM OF NON-IDENTIFIABILITY ARISING IN SIMPLE STOCHASTIC MODELS FOR STEREOLOGICAL COUNTS

by

PREM S. PURI* and JAMES W. YACKEL

PURDUE UNIVERSITY

Technical Report #83-25

Department of Statistics
Purdue University

July 1983

*These investigations were supported in part by the U. S. National Science Foundation Grant No. MCS-8102733.

1. INTRODUCTION.

The results presented here were inspired by a live problem about which the authors were consulted by Dr. Shirley Bayer of the Biology Dept. at Purdue University. The problem is concerned with the development of statistical methods for the estimation of the total number of cells of certain type in a specified region of the rat brain [1]. Reviewing the existing literature in geometric probability ([5], [8]) and the literature on closely related stereological problems ([3], [4], [9], [10]), it was felt appropriate to estimate the number of cells through a volumetric analysis approach.

Examination of sample data obtained from microtome sections of the rat brain revealed that the observed cell size sections were noticeably lacking of small values. This problem is of course well recognized in the literature ([4], [10]). However, to the best of our knowledge, no specific adjustment of the mathematical model itself has been considered to account for this phenomenon. Working closely with Dr. Bayer and her colleagues, our attempts to explain this lack of small observed cell sections led us to the two stochastic models considered here.

Again, the chosen statistical approach for estimating the number of rat brain cells, required us to estimate first the expected volume of a randomly selected cell. This naturally leads to the question: Is it possible to identify the underlying cell-size distribution uniquely from knowing only the distribution of the sizes of the cell-sections? As pointed out in ([6], [7]), the problem of nonidentifiability in the context of stochastic modeling is often more acute than is usually thought of or looked into or even reported. This appears to be so as well in the case of stereological models. The question of nonidentifiability becomes of particular concern in the present case, where the models considered do not afford identification of that part of the cell size distribution which concentrates on arbitrarily small cells. Also, in general, it is important to investigate this question ifrst, before the model is put to any practical use for the purposes of inferences. Otherwise,in the presence of nonidentifiability, as indicated by Clifford [2] through numerical examples in his case, one may arrive at quite conflicting predictions by using them. Finally, while we shall investigate elsewhere the statistical estimation problems relating to the two models considered here, our primary concern here will be to answer the identifiability question raised above, associated with each of these models.

2. Model Development

In this section we will describe the two mathematical models which determine the applicability of our results. Certain apsects of these models are specified by the physical process and methods which were of interest while other features fall into the category of mathematical expedients. We will identify and provide a rationale for the main features of these models labeling them as assumptions. While the first three assumptions are common for the two models, they differ in the treatment of small sections as assumptions 4 and 5 specify.

For convenience in making calculations, it is common to restrict the object of our observations to some class of geometric shapes. Spheres are commonly used for this purpose ([4], [9], [10]).

Assumption 1. Spheres of random radii are embedded in a medium. The radii, R_i, are mutually independent random variables with a common distribution function G(r).

A second mathematical condition allows us to compute the probability distribution for observations made on the spheres.

Assumption 2. Spheres are distributed uniformly throughout the medium.

The method of obtaining observations which we studied, involved staining tissue sections so that the objects of interest became opaque and the surrounding material remained translucent. These conditions led to certain features of this next statement.

Assumption 3. Observations are obtained when a section of thickness ϵ is removed from the medium and the projected image of any embedded sphere is observed on one face of the section. The measurements recorded are radii of circles projected onto a face of the medium. See Figure 1.

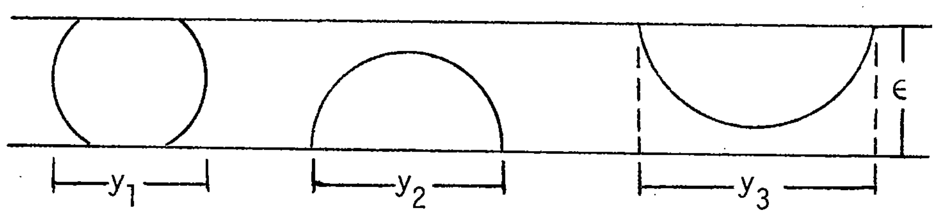

Figure 1.

When the model is modified to account for a lack of small circles observed in practice, there are a variety of possible accommodations. Here we consider two obvious ones. Model I results when the optical resolution reaches a minimum threshold below which there is no observation recorded.

Model I

Assumption 4. Δ is the smallest observable radius of a projected image and is independent of R, the radius of the sphere.

The observed circle has a radius Y, a random variable bounded below by Δ. If the embedded spheres have radii R_i which are mutually independent random variables with a common probability distribution G then the conditional probability distribution function of Y, given that $Y \geq \Delta$, is

$$F(y|Y \geq \Delta) = \frac{1}{C(\Delta,G)} \left[\epsilon \int_{\Delta}^{y} \frac{dG(r)}{2r + \epsilon} + \int_{\Delta}^{y} 2s \int_{s}^{\infty} \frac{dG(r)}{(2r + \epsilon)\sqrt{r^2 - s^2}} ds \right] \qquad (1)$$

where $C(\Delta, G) = P[Y \geq \Delta]$ is the normalizing factor.

Model II

Our second accommodation for small observed circles resulted from a conference with biologists working with stereological methods. The premise that the depth of material available to absorb staining chemicals falls below a threshold and reduces opacity led to our next assumption.

Assumption 5. When the maximum depth of a sphere contained in the section is less than ρ, the projected image is not observed.

Notice that the smallest observed radius Y is random and depends on the sphere radius R. Figure 2 shows the relationship of the projected radius Y to the cell radius r and the depth of material x.

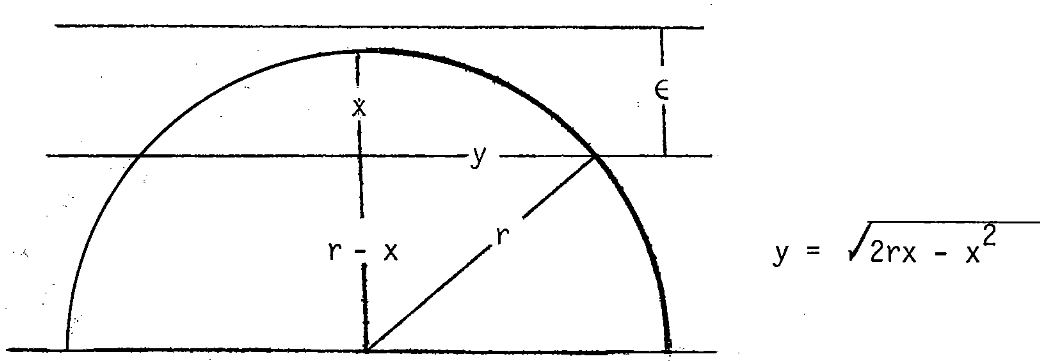

Figure 2.

The observed radius, Y, is a random variable bounded below by $\sqrt{2R\rho-\rho^2}$, where R is the radius of the embedded sphere. If the embedded spheres have radii R_i which are mutually independent random variables with a common probability distribution G then the conditional probability distribution function of Y, given that $Y \geq \sqrt{2R\rho-\rho^2}$, is

$$(2) \qquad F(y|Y \geq \sqrt{2R\rho-\rho^2}) = \frac{1}{D(\rho,G)} \left[\epsilon \int_{\rho/2}^{y} \frac{dG(r)}{2r+\epsilon} + \int_{\rho/2}^{y} 2s \int_{s}^{s+\frac{(s-\rho)^2}{2\rho}} \frac{dG(r)}{(2r+\epsilon)} \frac{ds}{\sqrt{r^2+s^2}} \right]$$

where $D(\rho,G) = P\left[Y \geq \sqrt{2R\rho-\rho^2}\right]$ is the normalizing factor.

3. THE IDENTIFIABILITY PROBLEM

Any researcher applying one of the models proposed in section 2 above will want to investigate properties of the distribution of sphere radii. Most probably moments of that distribution will be of interest. It is therefore important to know if there is a one to one correspondence between the distribution G and the distribution F which we observe from experimentation.

Since either model includes the possibility for information about G to be lost whenever G has support arbitrarily near 0 we surely cannot distinguish two distributions which differ on the interval $[0,\Delta]$ and we are using Model 1 or which differ on the internal $[0,\rho/2]$ and we are using Model 2. Even excluding these obvious problems the question of identifiability of G needs to be answered before inferences about G can be considered.

The identifiability question may be stated as follows for Model 1:

If G is not identifiable then there must exist two distinct distributions

G_1 and G_2 so that for all $y > \Delta$

$$(3)\qquad \frac{1}{C(\Delta,G_1)}\left[\epsilon\int_{\Delta}^{y}\frac{dG_1(r)}{2r+\epsilon}+\int_{\Delta}^{y}2s\int_{s}^{\infty}\frac{1}{2r+\epsilon}\,\frac{dG_1(r)}{\sqrt{r^2-s^2}}\,ds\right]=$$

$$\frac{1}{C(\Delta,G_2)}\left[\epsilon\int_{\Delta}^{y}\frac{dG_2(r)}{2r+\epsilon}+\int_{\Delta}^{y}2s\int_{s}^{\infty}\frac{1}{2r+\epsilon}\,\frac{dG_2(r)}{\sqrt{r^2-s^2}}\,ds\right].$$

For Model II, a slightly different equation results:

$$(4)\qquad \frac{1}{D(\rho,G_1)}\left[\epsilon\int_{1/2\rho}^{y}\frac{dG_1(r)}{2r+\epsilon}+\int_{1/2\rho}^{y}2s\int_{s}^{s+\frac{(s-\rho)^2}{2\rho}}\frac{1}{2r+\epsilon}\,\frac{dG_1(r)}{\sqrt{r^2-s^2}}\,ds\right]=$$

$$\frac{1}{D(\rho,G_2)}\left[\epsilon\int_{1/2\rho}^{y}\frac{dG_2(r)}{2r+\epsilon}+\int_{1/2\rho}^{y}2s\int_{s}^{s+\frac{(s-\rho)^2}{2\rho}}\frac{1}{2r+\epsilon}\,\frac{dG_2(r)}{\sqrt{r^2-s^2}}\,ds\right].$$

We cannot give a complete answer to the identifiability questions. Our results will be given under varying conditions to cover the cases where G can be shown to be identifiable.

3.1 G HAS BOUNDED SUPPORT.

When the distribution has bounded support, i.e., there exists some number B so that $G(B) = 1$, then we can show that G is identifiable.

<u>THEOREM 1</u>. (i) <u>If</u> G <u>has bounded support and</u> $G(\Delta) = 0$ <u>then</u> G <u>is identifiable for Model I</u>.

(ii) <u>If</u> G <u>has bounded support and</u> $G(\rho/2) = 0$ <u>then</u> G <u>is identifiable for Model II</u>.

<u>PROOF</u>:

To simplify notation we only give the proof for the case when the distributions have density functions. We will write G_1 for G and g_1 for the density of G and consider the density function equivalents of equations (3) and (4).

Suppose that there exists a distribution G_2 with density function g_2 so that G_1 and G_2 satisfy equation (3). By setting $\phi(r) = g_1(r)/c(\Delta,G_1)-g_2(r)/c(\Delta,G_2)$ and differentiating equation (3) we obtain

$$(5) \qquad \frac{\epsilon\,\phi\,(y)}{2y+\epsilon} + 2y\int_y^\infty \frac{\phi(r)}{(2r+\epsilon)\sqrt{r^2-y^2}}\,dr = 0 \quad .$$

Our assumption of bounded support implies that $\phi(r) = 0$ if $r > B$. We will next show that there is some positive number K so that $\phi(r) = 0$ for $r > B$ implies that $\phi(r) = 0$ for $r > B - K$.

Let $0 < K < \epsilon^2/32B$, $K < B$ be given. Let $M = \max_{B-K\le r\le B} |\phi(r)|$ and choose $b_0 \in [B-K,B]$ so that $|\phi(b_0)| > M/2$. From (5),

$$\phi(b_0) = \frac{-2b_0+\epsilon}{\epsilon} \cdot 2\,b_0 \int_{b_0}^{B} \frac{\phi(r)\qquad dr}{(2r+\epsilon)\sqrt{r^2-b_0^2}}$$

so that

$$(6) \qquad |\phi(b_0)| \leq \frac{2b_0+\epsilon}{\epsilon} \cdot 2b_0 \int_{b_0}^{B} \frac{|\phi(r)|}{(2r+\epsilon)\sqrt{r^2-b_0^2}} \ .$$

Since $\frac{2b_0 + \epsilon}{2r + \epsilon} \leq 1$ when $b_0 \leq r$ we can obtain a bound from (6).

$$(7) \qquad |\phi(b_0)| \leq \frac{2}{\epsilon} \int_{b_0}^{B} \frac{r|\phi(r)|}{\sqrt{r^2 - b_0^2}} dr \ .$$

The bound of M on $|\phi(r)|$ and integration of the remaining expression results in the estimate

$$(8) \qquad |\phi(b_0)| \leq \frac{2M}{\epsilon} \int_{b_0}^{B} \frac{r}{\sqrt{r^2 - b_0^2}} dr = \frac{2M}{\epsilon} \sqrt{B^2 - b_0^2} \ ,$$

By our choice of K, $B^2 - b_0^2 = (B + b_0)(B - b_0) \leq 2BK$ and finally we obtain the result

$$(9) \qquad |\phi(b_0)| \leq \frac{2M}{\epsilon} \sqrt{2BK} \leq \frac{2M}{\epsilon} \sqrt{\frac{2B\epsilon^2}{32B}} = \frac{M}{2}$$

which contradicts the coince of b_0 unless $M = 0$. Since K does not depend on M and K decreases as $1/B$, this argument shows that $\phi(r) = 0$ for all r. The same argument is valid if we use equation (4) rather than equation (3) as our starting point.

To conclue the proof we need only show that $C(\Delta,G_1) = C(\Delta,G_2)$ but that fact follows from the condition that $\int_{\Delta}^{\infty} dG_1(r) = \int_{\Delta}^{\infty} dG_2(r) = 1$.

Note that when $G(\Delta) > 0$ we can still conclude that if some G_2 exists so that equation (3) is satisfied then $\phi(r) = 0$ for $r > \Delta$, that is $g_1(r)/C(\Delta,G_1) = g_2(r)/C(\Delta,G_2)$ for $r > \Delta$.

3.2 G IS A DISCRETE DISTRIBUTION.

When the distribution G has a discrete part the distribution F, derived from G, will also have a discrete part. In fact, for Model I,

(10) $$F_G(r) - F_G(r^-) = \epsilon \frac{[G(r)-G(r^-)]}{2r+\epsilon} \cdot \frac{1}{C(\Delta,G)}$$

or the same quantity with $D(\Delta,G)$ replaced by $D(\rho,G)$ in the case of Model II.

THEOREM 2. *If G is a discrete distribution and*

(i) $G(\Delta) = 0$, *then G is identifiable for Model I*,

(ii) $G(\rho/2) = 0$, *then G is identifiable for Model II*.

PROOF: If G_1 and G_2 are two discrete distributions which satisfy equation (3) or (4) then they must have the same support since equation (10) identifies the discontinuity points of F and of G_1 and G_2. Furthermore we can equate the right sides of equation (10) for the two distributions to obtain, for Model I,

(11) $$\frac{G_1(r) - G_1(r^-)}{C(\Delta,G_1)} = \frac{G_2(r) - G_2(r^-)}{C(\Delta,G_2)} \quad .$$

Summing (11) over all points of support and using the condition that $G_1(\Delta) = G_2(\Delta) = 0$ implies that $\Sigma[G_i(r) - G_i(r^-)] = 1$ and so $C(\Delta,G_1) = C(\Delta,G_2)$.

The exact same reasoning and steps applied to Model II furnishes the remaining proof to the theorem.

3.3 IDENTIFIABILITY FOR MODEL I.

We will show that the assumption of boundedness is not needed for the identifiability of G when working with Model I. We are not able to prove a comparable result for Model II.

THEOREM 3. If G is an absolutely continuous probability distribution and $G(\Delta) = 0$, then G is identifiable for Model I.

Proof: For the proof we use the following lemma.

Lemma. Let

$$\varepsilon Q(u) + \int_u^\infty \frac{Q(s)}{\sqrt{s-u}}\, ds = 0 \text{ for all } u \geq B.$$

Then $Q(u)$ is identically zero for $u \geq B$.

PROOF: Define

$$M(u) = \varepsilon Q(u) + \int_u^\infty \frac{Q(s)}{\sqrt{s-u}}\, ds.$$

Since $M(u) \equiv 0$ by hypothesis,

$$\text{(12)} \qquad \varepsilon M(u) - \int_u^\infty \frac{M(s)}{\sqrt{s-u}}\, ds = 0$$

and the left side of (12) by definition is

$$\varepsilon[\varepsilon Q(u) + \int_u^\infty \frac{Q(s)}{\sqrt{s-u}} ds] - \int_u^\infty \int_v^\infty \frac{Q(s)}{\sqrt{v-u} \quad \sqrt{s-v}} ds\, dv - \varepsilon \int_u^\infty \frac{Q(s)}{\sqrt{s-u}} du$$

which simplifies to become

$$\varepsilon^2 Q(u) - \int_u^\infty \int_v^\infty \frac{Q(s) \quad ds\, dv}{\sqrt{v-u} \quad \sqrt{s-v}} .$$

Interchanging the order of integration above and computing the resultant integral gives us the equation

$$\varepsilon^2 Q(u) - \pi \int_u^\infty Q(s) ds = 0 .$$

Differentiation produces the differential equation

$$(13) \qquad \varepsilon^2 Q'(u) + \pi\, Q(u) = 0$$

The solution to (13) is given by

$$Q(u) = C \exp\, (-\pi u/\varepsilon^2)$$

Upon substitution into the original equation (12), we obtain

$$C \exp\, (-\pi u/\varepsilon^2) + \int_u^\infty \frac{C \exp\, (-\pi s/\varepsilon^2)}{\sqrt{s-u}} ds = 0 \quad ,$$

which can only hold when C = 0. Thus $Q(u) \equiv 0$.

PROOF OF THEOREM 3.

To reduce the statement of Theorem 3 to the hypothesis of the lemma we proceed as follows: Since each distribution is absolutely continuous, $F(y|Y>\Delta)$ has a density and differentiating equation (3) we obtain

$$f(y|Y\geq\Delta) = \frac{1}{C(\Delta,G_i)}\left[\varepsilon \frac{g_i(y)}{2y+\varepsilon} + 2y\int_y^\infty \frac{1}{(2r+\epsilon)} \frac{g_i(r)\,dr}{\sqrt{r^2-y^2}}\right] .$$

We make the substitution $y = \sqrt{u}$, $r = \sqrt{s}$ and $h_i(y) = \dfrac{g_i(y)}{C(\Delta,G_i)(2y+\varepsilon)}$

to obtain

$$f(\sqrt{u}|Y\geq\Delta) = \sqrt{u}\left[\varepsilon \frac{h_i(\sqrt{u})}{\sqrt{u}} + \int_u^\infty \frac{h_i(\sqrt{s})}{\sqrt{s}\ \sqrt{s-u}}\,ds\right]$$

Next consider this equation for i = 1 and i = 2, then equate the two expressions for f. This yields

$$\frac{\varepsilon h_i(\sqrt{u})}{\sqrt{u}} + \int_u^\infty \frac{h_i(\sqrt{s})}{\sqrt{s}\ \sqrt{s-u}}\,ds = \frac{\varepsilon h_2(\sqrt{u})}{\sqrt{u}} + \int_u^\infty \frac{h_2(\sqrt{s})}{\sqrt{s}\ \sqrt{s-u}}\,ds$$

and by setting $Q(u) = \dfrac{h_1(\sqrt{u})}{\sqrt{u}} - \dfrac{h_2(\sqrt{u})}{\sqrt{u}}$ the

resulting equation is given by

$$\varepsilon Q(u) + \int_u^\infty \frac{Q(s)}{\sqrt{s-u}}\,ds = 0, \text{ for all } u.$$

The Lemma tells us that $Q(u) = 0$ for all u and so $h_1(\sqrt{u}) = h_2(\sqrt{u})$

and in turn $g_1(y)/C(\Delta,G_1) = g_2(y)/C(\Delta,G_2)$.

As before, $G(\Delta) = 0$ implies that $\int_\Delta^\infty g_i(y)dy = 1$ so that $C(\Delta,G_1) = C(\Delta,G_2)$ and the proof is complete.

4. SOLVING FOR THE DISTIRUBTION, G, OF R.

In those instances for which there is a unique G for a given distribution F, i.e. when G is discrete or when F is obtained from Model 1, we can hope to invert equation (1) or equation (2) in order to find G. Determining G or some of its moments will be our objective in this section.

4.1 THE CASE OF DISCRETE DISTRIBUTIONS.

Recall equation (11) from section 3.2. We are able to solve for the discontinuities of G to obtain the expression:

$$G(r_i) - G(r_i^-) = \frac{(2r_i + \varepsilon)}{\varepsilon}(F(r_i) - F(r_i^-))D(G) \tag{14}$$

in the case of Model I. When this is summed over i and if $G(\Delta) = 0$ then

$$D(G) = \epsilon/\{\sum_{i=1}^{\infty} (2r_i + \epsilon)(F(r_i) - F(r_i^-))\}$$

and the explicit distribution G is given by

$$G(r_i) - G(r_i^-) = \frac{(2r_i + \epsilon)(F(r_i) - F(r_i^-))}{\sum_{j=1}^{\infty}{}' (2r_j + \epsilon)(F(r_j) - F(r_j^-))}.$$

In particular,

$$E_G(R^k) = \frac{\sum_{i=1}^{\infty} r_i^k \, (2r_i + \epsilon)(F(r_i) - F(r_i^-))}{\sum_{j=1}^{\infty} (2r_j + \epsilon)(F(r_j) - F(r_j^-))}$$

4.2 MODEL I: THE CONTINUOUS CASE.

Similar attempts to solve the functional equation (1) for the underlying distribution G produce less satisfactory results. We consider only the case when G has a density function and we take the density version of (1),

$$(15) \qquad f(y|Y\geq\Delta) = \frac{1}{D(G)} \left[\frac{\epsilon g(y)}{2y + \epsilon} + 2y \int_y^{\infty} \frac{g(r) \qquad dr}{(2r + \epsilon)\sqrt{r^2 - y^2}} \right] .$$

For the purpose of solving this equation for g we first substitute

$$(16) \qquad h(y) = \frac{g(y)}{D(G)(2y+\varepsilon)}$$

Then (15) is written as

$$(17) \qquad F(y|Y>\Delta) = \varepsilon h(y) + 2y \int_y^{\infty} \frac{h(r)}{\sqrt{r^2 - y^2}} \, dr$$

Next, replacing h(r) by the expression given by equation (17) produces

(18) $$f(y|Y\geq\Delta) = \epsilon\, h(y) + 2y \int_y^\infty \left\{ \frac{f(r|y\geq\Delta)}{\epsilon} - \frac{2r}{\epsilon} \int_r^\infty \frac{h(s)}{\sqrt{s^2-r^2}}\, ds \right\} \frac{dr}{\sqrt{r^2-y^2}}$$

$$= \epsilon h(y) + \frac{2y}{\epsilon} \int_y^\infty \frac{f(r|y\geq\Delta)}{\sqrt{r^2-y^2}}\, dr - \frac{2y}{\epsilon} \int_y^\infty \frac{r}{\sqrt{r^2-y^2}} \int_r^\infty \frac{h(s)}{\sqrt{s^2-r^2}}\, ds\, dr$$

changing the order of integration in the last term gives the result

(19) $$f(y|Y\geq\Delta) = \epsilon\, h(y) + \frac{2y}{\epsilon} + \int_y^\infty \frac{f(r|y\geq\Delta)}{\sqrt{r^2 - y^2}}\, dr - \frac{\pi y}{\epsilon} \int_y^\infty h(s)\, ds.$$

Finally, we substitute $h(y)/y = Q(y)$, rearrange terms and find:

(20) $$Q(y) - \frac{\pi}{\epsilon^2} \int_y^\infty sQ(s)\, ds = \frac{f(y)}{\epsilon y} - \frac{2}{\epsilon^2} \int_y^\infty \frac{f(r)}{\sqrt{r^2-y^2}}\, dr$$

Write $N(y)$ for the r.h.s. of (20), differentiate w.r.t. y and the resulting differential equation is

(21) $$Q'(y) + \frac{\pi}{\epsilon^2}\, yQ(y) = N'(y)$$

The solution to (21) is given next.

$$Q(y) = e^{-\frac{\pi y^2}{2\epsilon^2}} \left[\int e^{\frac{\pi r^2}{2\epsilon^2}} N'(r)\, dr + C \right]$$

Substitution for $Q(y)$ and $N(u)$ then produces an expression for $g(y)$.

(22) $$\frac{g(y)}{D(G)(2y+\epsilon)y} = \frac{f(y)}{\epsilon y} - \frac{2}{\epsilon^2}\int_y^\infty \frac{f(r)}{\sqrt{r^2-y^2}}\,dr$$

$$- \frac{\pi}{\epsilon^2} e^{-\frac{\pi y^2}{2\epsilon^2}} \int u e^{\frac{\pi u^2}{2\epsilon^2}} \left\{ \frac{f(u)}{\epsilon u} - \frac{2}{\epsilon^2}\int_u^\infty \frac{f(r)}{\sqrt{r^2-u^2}}\,dr \right\} du$$

$$+ c' e^{-\frac{\pi y^2}{2\epsilon^2}}$$

4.3 EXAMPLES. In some simple cases we can compute f to illustrate the relationship of f to g.

Example 1. Take $g(r) = \frac{2\Delta^2}{4\Delta+\epsilon} \cdot \frac{(2r+\epsilon)}{r^3}$ for $r > \Delta$ and 0 otherwise.

Then we find

$$f(y|Y \geq \Delta) = \frac{2\Delta^2}{\epsilon+\pi\Delta}\left[\frac{\epsilon}{y^3} + \frac{\pi}{2y^2}\right] \quad \text{for } y > \Delta$$

and 0 otherwise. Graphically, when $\Delta = 1$ and $\epsilon = 3$ this gives figure 3.

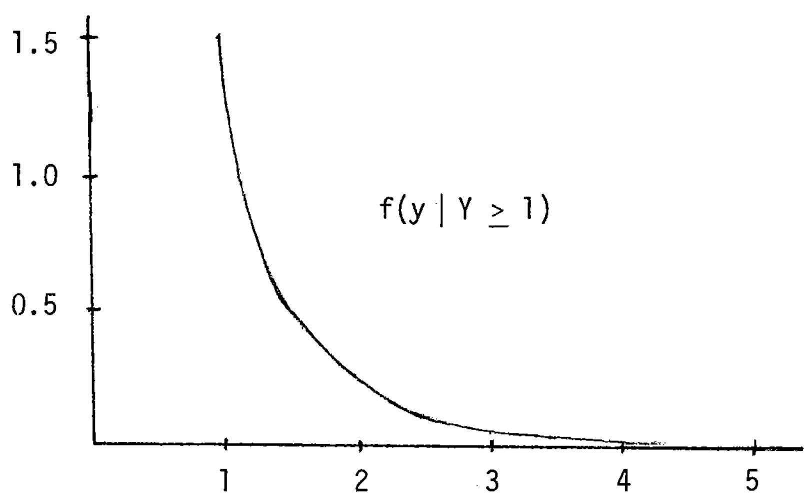

Figure 3.

Example 2. Taking $g_{(a,b)}(r) = \alpha(2r + \epsilon)r$ on an interval [a,b] and 0 otherwise gives

$$f(y) = \begin{cases} \frac{2y\alpha}{D}\left[\sqrt{b^2 - y^2} - \sqrt{a^2 - y^2}\right] & \text{for } \Delta \le y < a \\ \frac{y\alpha}{D}\left[\epsilon + 2\sqrt{b^2 - y^2}\right] & \text{for } a \le y \le b \\ 0 & \text{otherwise .} \end{cases}$$

When we take a mixture of these, such as,

$$g(r) = (\cdot 3)\, g_{(4,5)}(r) + (\cdot 4)\, g_{(5,6)}(r) + (\cdot 3)\, g_{(6,7)}(r)$$

the graph is given in figure 4.

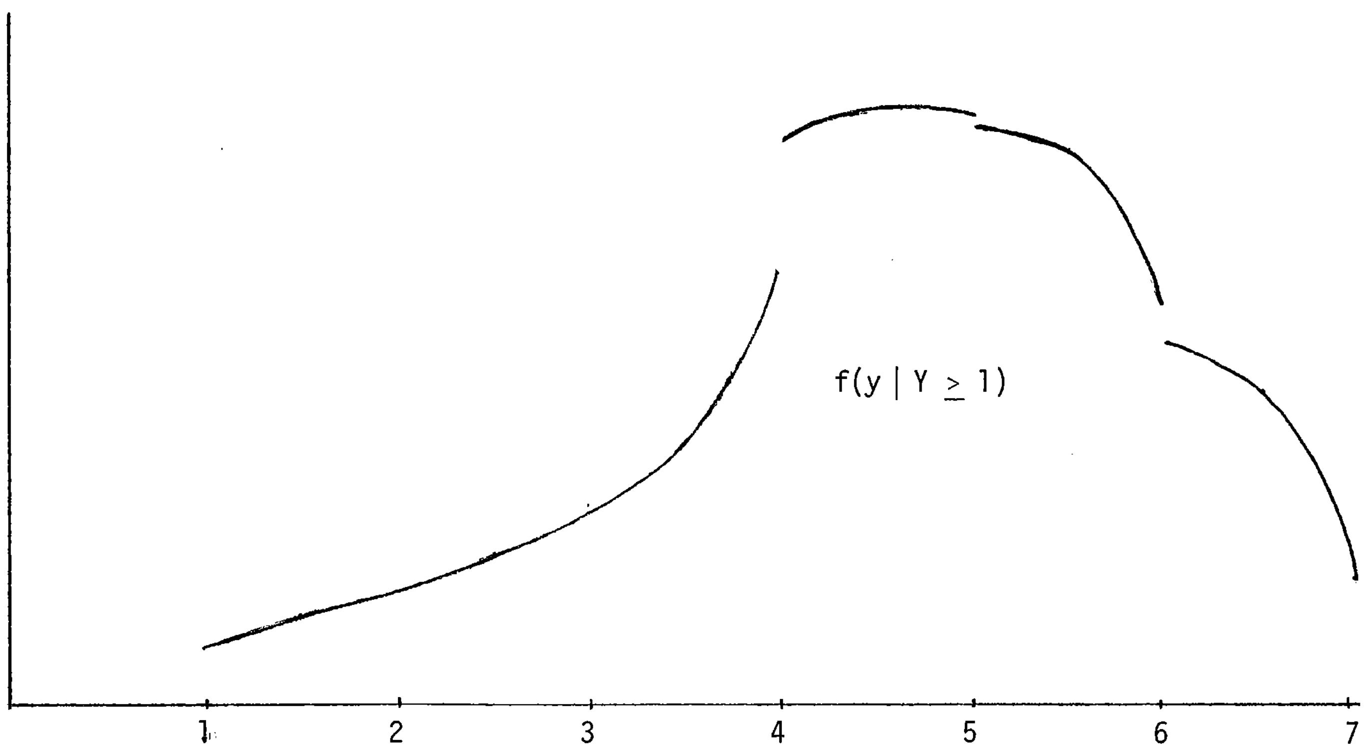

Figure 4.

5. CONCLUDING REMARKS.

(a) Conjecture. We believe that the distribution G should turn out to be identifiable for Model II even when G does not have bounded support. Limitations of our technique prevent proving this result. Any counter example would exhibit pathological properties which in applications are uninteresting.

(b) Practical Implications. In any application, when the support of the distribution G includes arbitrarily small values it is evident that G will not be identifiable with either Model I or Model II. For example, observing the cells of an immature laboratory speciman poses a greater risk for statistical inferences, since presumably, developing cells would be much smaller than mature cells. If the cells are sufficiently small to not be observed, then any statistical inferences concerning cell sizes would be in error.

ACKNOWLEDGMENT. The authors thank Dr. Shirley Bayer for her discussions of the laboratory techniques used for microtome sections and for bringing this problem to our attention.

REFERENCES

[1] Bayer, Shirley A., Puri, Prem S. and Yackel, James W., "Neurons in the Rat Dentate Gyrus Granular Layer Substantially Increase During Juvenile and Adult Life." Science, 216, 890-892 (1982).

[2] Clifford, P., "Nonidentifiability in Stochastic Models of Illness and Death," Proc. Natl. Acad. Sci. USA, 74, 1338-1340 (1977).

[3] Cruz-Orive, L. M. and Myking, A. O., "Stereological estimation of volume ratios by systematic sections," J of Microscopy, 122, 143-157 (1981).

[4] Hendry, A., "A Method to correct adequately for the change in neuronal size when estimating neuronal numbers after nerve growth factor treatment", J. of Neurocytology 5, 337-349 (1976).

[5] Kendall, M. G. and Moran, P. A. P., Geometrical Probability, Charles Griffin, London (1963).

[6] Puri, P. S., "On certain Problems Involving Non-Identifiability of distributions arising in Stochastic Modeling", Optimizing Methods in Statistics, Academic Press, New York (1979).

[7] Puri, P. S., "On Nonidentifiability Problems Among Some Stochastic Models in Reliability Theory", To appear in the proceedings of the Neyman-Kiefer Conference held during June 20-July 1, 1983 at the University of California Berkeley.

[8] Solomon, Herbert, Geometric Probability, Society for Industrial and Applied Mathematics, Philadelphia (1978).

[9] Underwood, E. E., Quantitative Stereology, Addison-Wesley, Reading Mass. (1970).

[10] Weibel, E. R., Stereological Methods Volume I Practical Methods for Biological Morphometry, Academic Press, New York (1979).