Computing Chernoff's Distribution

Piet GROENEBOOM and Jon A. WELLNER

A distribution that arises in problems of estimation of monotone functions is that of the location of the maximum of two-sided Brownian motion minus a parabola. Using results from the first author's earlier work, we present algorithms and programs for computation of this distribution and its quantiles. We also present some comparisons with earlier computations and simulations.

Key Words: Airy functions; Brownian motion; Density; Distribution function; Monotone; Parabolic drift; Quadratic drift; Quantiles.

1. INTRODUCTION

Our goal here is to compute, table, and plot the density, distribution function, moments, and quantiles of the location Z of the maximum of two-sided Brownian motion B minus the parabola t^2. We also provide several examples of the application of this distribution to problems including interval censoring, monotone density and hazard estimation, deconvolution, least median of squares estimation, and mode estimation.

To be explicit, let $B(t)$, $-\infty < t < \infty$, be two-sided standard Brownian motion with $B(0) = 0$. Then

$$Z \equiv \text{argmax}_t (B(t) - t^2) .$$

It follows from Lemma 2.6 of Kim and Pollard (1990) that Z is uniquely defined with probability 1. The distribution of Z apparently first arose in work of Chernoff (1964) on the estimation of the mode of a distribution function, and hence we refer to the distribution of Z as *Chernoff's distribution*.

Prakasa Rao (1969) showed that the distribution of the slope at zero of the greatest convex minorant of $B(t) + t^2$ is exactly $2Z$. This follows from the "switching relation"; see Groeneboom (1985, Eq. (2.2), p. 541) for the finite sample version of this relation. Groeneboom (1985; 1989) completely described the distribution of Z and characterized analytically the *process* $\{V(a) : a \in \mathbb{R}\}$, where

$$V(a) \equiv \sup\{t \in \mathbb{R} : B(t) - (t-a)^2 \quad \text{is maximal.}$$

Piet Groeneboom is Professor, Department of Mathematics, Delft University of Technology, 2628 CD Delft, The Netherlands (E-mail: P.Groeneboom@its.tudelft.nl). Jon A. Wellner is Professor, Department of Statistics, University of Washington, P.O. Box 354322, Seattle, WA 98195-4322 (E-mail: jaw@stat.washington.edu).

Journal of Computational and Graphical Statistics, Volume 10, Number 2, Pages 388–400

In particular, Z has a density f_Z with respect to Lebesgue measure on $\mathbb{R}$ which is symmetric about zero, and which satisfies

$$f_Z(z) \sim \frac{1}{2}\frac{4^{4/3}|z|}{Ai'(\tilde{a}_1)} \exp\left(-\frac{2}{3}|z|^3 + 2^{1/3}\tilde{a}_1|z|\right) \qquad \text{as} \qquad z \to \infty,$$

where $\tilde{a}_1 \approx -2.3381$ is the largest zero of the Airy function Ai and where $Ai'(\tilde{a}_1) \approx 0.7022$. The link of the distribution of Z with Airy functions is also given in Daniels and Skyrme (1985), but the process $\{V(a) : a \in \mathbb{R}\}$ is not discussed in that article. In unpublished notes, Groeneboom and Sommeijer (1984) numerically computed the absolute moments $E(|Z|^k)$, $k = 1, \ldots, 4$. The first of these was reported by Devroye and Gyorfi (1985, p. 214); and all four of them were reported by Keiding, Begtrup, Scheike, and Hasibeder (1996). Note that by symmetry of f_Z it follows that $E(Z^k) = 0$ for k odd.

2. APPLICATIONS AND EXAMPLES

Here we present several examples showing how the distribution of Z enters.

Example 1. (decreasing densities): A classical example of an application of the theory is the *Grenander estimator* of a decreasing density on $[0, \infty)$ Suppose that $X_1, \ldots, X_n$ is a sample, generated by a decreasing density f on $[0, \infty)$ that has a nonzero derivative $f'(x)$ at a point $x \in (0, \infty)$. Let $\hat{f}_n$ be the maximum likelihood estimator of f under the monotonicity restriction. Then $\hat{f}_n$ is the left-continuous derivative of the concave majorant of the empirical distribution function; see Grenander (1956) and Groeneboom and Lopuhaä (1993). This estimator has, after "cube root n" standardization, the following limiting distribution

$$n^{1/3}\left\{\frac{1}{2}f(x)f'(x)\right\}^{-1/3}\left\{\hat{f}_n(x) - f(x)\right\} \xrightarrow{\mathcal{D}} 2Z, \; n \to \infty,$$

where $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution, see Prakasa Rao (1969) and (for a different shorter proof) Groeneboom (1985).

Example 2. (monotone failure rates): For this situation a similar result holds as in Example 1. For example, let $X_1, \ldots, X_n$ is a sample, generated by a distribution with an increasing failure rate r on $[0, \infty)$ and let $\hat{r}_n$ be the NPMLE of r in the class of distributions with increasing failure rate. Then, under some regularity conditions:

$$n^{1/3}\left\{\frac{1}{2}r(x)^2r'(x)\right\}^{-1/3}\{\hat{r}_n(x) - r(x)\} \xrightarrow{\mathcal{D}} 2Z, \; n \to \infty,$$

see Prakasa Rao (1970).

Example 3. (least median of squares estimator): Let $X_1, \ldots, X_n$ be a sample from a density strongly unimodal distribution with density f_θ, given by

$$f_\theta(x) = f(x - \theta), \; x \in \mathbb{R},$$

where the density f is differentiable and symmetric around zero. Rousseeuw (1984) introduces the *least median of squares* estimator $\hat{\theta}_n$ for the shift parameter θ and gives a heuristic argument for the result that

$$n^{1/3}\left\{\hat{\theta}_n - \theta\right\},$$

converges in distribution to $c \cdot Z$, where $c > 0$ is a constant depending on the distribution function F corresponding to the density f, the density f itself and its derivative f'. A proof of this result is given in Kim and Pollard (1990). In Rousseeuw (1984) also recommendations for confidence intervals can be found.

Example 4. (interval censoring, case 1, also called "current status model"): Suppose that X is a "survival time" with distribution function F on $[0, \infty)$ and Y is an observation time which is independent of X and has distribution function G on $[0, \infty)$. However, we can observe only $(Y, 1_{\{X \le Y\}}) = (Y, \Delta)$, and want to estimate F, the distribution function of X based on iid replications $(Y_1, \Delta_1), \ldots, (Y_n, \Delta_n)$ of (Y, Δ). In this case the NPMLE $\widehat{\mathbb{F}}_n$ of F is known from Ayer, Brunk, Ewing, Reid, and Silverman (1955). It was proved in Groeneboom and Wellner (1992) that if F has density f and G has density g at $t_0 \in (0, \infty)$ with $g(t_0) > 0$, $f(t_0) > 0$, then

$$n^{1/3}(\widehat{\mathbb{F}}_n(t_0) - F(t_0)) \xrightarrow{\mathcal{D}} \left\{\frac{1}{2}F(t_0)(1 - F(t_0))f(t_0)/g(t_0)\right\}^{1/3} 2Z, \quad n \to \infty.$$

Thus, from Table 2 in Section 2, it follows that an asymptotic 95% confidence interval for $F(t_0)$ is given by

$$\widehat{\mathbb{F}}_n(t_0) \pm n^{-1/3}\left\{\frac{1}{2}\widehat{\mathbb{F}}_n(t_0)(1 - \widehat{\mathbb{F}}_n(t_0))\widehat{f}(t_0)/\widehat{g}(t_0)\right\}^{1/3} 2 \cdot (.99818),$$

where $\widehat{f}(t_0)$ and $\widehat{g}(t_0)$ are any consistent estimators of $f(t_0)$ and $g(t_0)$, respectively; for example, based on kernel smoothing of $\widehat{\mathbb{F}}_n$ and $\widehat{\mathbb{G}}_n(t) \equiv n^{-1}\sum_{i=1}^n 1_{\{Y_i \le t\}}$.

In the particular application discussed by Keiding et al. (1996), X_i represents "age of immunization" of individual i against rubella, Y_i represents "current age" of person i.

Example 5. (interval censoring, case 2): In this case the data consist of a sample of observations

$$(U_i, V_i, \Delta_i, \Gamma_i), \quad i = 1, \ldots, n,$$

where $U_i < V_i$ and $[U_i, V_i]$ is an "observation interval" for the (hidden and unobservable) variable X_i. The variables Δ_i and Γ_i are indicators, telling us whether X_i is left of U_i, between U_i and V_i, or right of V_i:

$$\Delta_i = 1_{\{X_i \le U_i\}}, \quad \Gamma_i = 1_{\{U_i < X_i \le V_i\}}.$$

The X_i are assumed to be independent of the (U_i, V_i). For analyses of this model; see, for example, Groeneboom (1996) and Geskus and Groeneboom (1999). In this situation

the NPMLE $\widehat{\mathbb{F}}_n$ has to be computed by an iterative method. A fast iterative algorithm is available, the so-called *iterative convex minorant algorithm*, proposed by Groeneboom and Wellner (1992) and further analyzed by Jongbloed (1998b).

Assume that (U_i, V_i) has a density h with respect to Lebesgue measure, with first and second marginal densities h_1 and h_2, respectively. Moreover, suppose that F is the distribution function of the variable X_i with a density f with respect to Lebesgue measure, and let k_i, $i = 1, 2$, and the function a be defined by

$$k_1(u) = \int_u^\infty \frac{h(u,v)}{F(v) - F(u)}\, dv, \quad k_2(v) = \int_0^v \frac{h(u,v)}{F(v) - F(u)}\, du,$$

and

$$a(t) = \frac{h_1(t_0)}{F(t_0)} + k_1(t_0) + k_2(t_0) + \frac{h_2(t_0)}{1 - F(t_0)}.$$

Then we have at a point t in the interior of the support of the distribution function F under some regularity conditions, in particular assuming that the observation points U_i and V_i are strictly separated (i.e., $P\{V_i - U_i < \epsilon\} = 0$ for some $\epsilon > 0$) and that $f(t) > 0$:

$$n^{1/3}\{2a(t)/f(t)\}^{1/3}\{\widehat{\mathbb{F}}_n(t) - F(t)\} \xrightarrow{\mathcal{D}} 2Z, \quad n \to \infty,$$

see Groeneboom (1996, theorem 4.4).

A long-standing conjecture is that in the situation where U_i and V_i are not strictly separated the rate of convergence increases to $(n \log n)^{1/3}$ and that the limiting distribution is again given by Z; see Groeneboom and Wellner (1992, p. 100), but this conjecture has at present still not been proved or disproved. Extensions to more than two observation points are possible (the situation is not very different from "case 2," since only the two observation points surrounding the hidden variable X_i will be relevant for the analysis), but we will not further discuss this here.

Example 6. (deconvolution): Let $X_1, \ldots, X_n$ be a sample from the convolution of an unknown distribution function F, concentrated on $[0, 1]$ and the uniform distribution and let $\widehat{\mathbb{F}}_n$ be the NPMLE of F. Then, if F has a positive density $f(x)$ at $x \in (0, 1)$:

$$n^{1/3}\left\{\widehat{\mathbb{F}}_n(x) - F(x)\right\} \Big/ \left\{\frac{1}{2}F(x)(1 - F(x))f(x)\right\}^{1/3} \xrightarrow{\mathcal{D}} 2Z, \quad n \to \infty,$$

see van Es (1991, theorem 4.5) or Groeneboom and Wellner (1992, p. 109). A similar result for deconvolution with the exponential distribution is given in Jongbloed (1998a), where it is shown that if $X_1, \ldots, X_n$ is a sample from the convolution of an unknown distribution function F, concentrated on $[0, \infty)$ with the standard exponential distribution, the NPMLE $\hat{F}_n$ of F satisfies

$$n^{1/3}\left\{\widehat{\mathbb{F}}_n(x) - F(x)\right\} \Big/ \left\{\frac{1}{2}e^{-x}f(x)\right\}^{1/3} \xrightarrow{\mathcal{D}} 2Z,$$

if F has a positive density $f(x)$ at $x > 0$. For related material, see van Es and van Zuijlen (1996) and van Es, Jongbloed, and van Zuijlen (1998).

Example 7. (mode Estimation; Venter's estimator): Suppose that $X_1, \ldots, X_n$ are iid with unimodal density f satisfying

$$f(x) = \gamma_0 - \frac{1}{2}\gamma(x-\theta)^2 + \frac{1}{6}\gamma_3(x-\theta)^3 + o(|x-\theta|^3)\,.$$

Then as shown by Venter (1967), his estimator $\widehat{\theta}_n$ of the mode θ satisfies

$$n^{1/5}(\widehat{\theta}_n - \theta) \to_d 2^{1/3}A^{-2/3}\gamma^{-2/3}\gamma_0 Z\,.$$

Thus, if $\hat{\gamma}_0$ and $\hat{\gamma}$ are consistent estimators of γ_0 and γ, respectively, then

$$\widehat{\theta}_n \pm \frac{2^{1/3}\hat{\gamma}_0}{n^{1/5}A^{2/3}\hat{\gamma}^{2/3}} \cdot (.99818)$$

yields an approximate 95% confidence interval for the mode θ.

Narayanan and Sager (1989) gave several nice examples of mode estimation via both Chernoff's estimators and Venter's estimators and their (simulated) quantiles of the distribution of Z to form confidence intervals; see especially pages 46–50.

Example 8. (panel count data): Wellner and Zhang (1998) showed that a pseudo-likelihood estimator $\widehat{\Lambda}_n$ of the mean function Λ of a counting process with "panel count" data satisfies

$$n^{1/3}(\widehat{\Lambda}_n^{ps}(t) - \Lambda(t)) \xrightarrow{\mathcal{D}} \left\{\frac{\sigma^2(t)\Lambda'(t)}{2G'(t)}\right\}^{1/3} 2Z,$$

where $G'(t) = \sum_{k=1}^{\infty} P(K=k)\sum_{j=1}^{k} G'_{k,j}(t)$. Thus, if $\hat{\sigma}^2(t)$, $\widehat{\Lambda}'$, and $\widehat{G}'(t)$ are consistent estimators of $\sigma^2(t)$, $\Lambda'(t)$, and $G'(t)$, respectively, then

$$\widehat{\Lambda}_n^{ps}(t) \pm \frac{1}{n^{1/3}}\left\{\frac{\hat{\sigma}^2(t)\widehat{\Lambda}'(t)}{2\widehat{G}'(t)}\right\}^{1/3} \cdot 2 \cdot (.99818)$$

yields an approximate 95% confidence interval for $\Lambda(t)$.

For a rather different approach to examples of the type presented here, see Politis and Romano (1994), especially their example 2.1.1, pages 2035–2036.

3. COMPUTATION OF THE DENSITY f_Z AND DISTRIBUTION FUNCTION F_Z

The density f_Z can in principle be found by solving the following partial differential equation (heat equation), given by Chernoff (1964):

$$\frac{\partial}{\partial t}u(t,x) = -\frac{1}{2}\frac{\partial^2}{\partial x^2}u(t,x), \tag{3.1}$$

for $x \le t^2$, under the boundary conditions:

$$u(t,t^2) \stackrel{\text{def}}{=} \lim_{x \uparrow t^2} u(t,x) = 1, \qquad \lim_{x \downarrow -\infty} u(t,x) = 0,\ t \in \mathbb{R}. \tag{3.2}$$

In terms of the (smooth) solution $u(t,x)$, the density f_Z is given by

$$f_Z(t) = \frac{1}{2}u_2(-t)u_2(t),\ t \in \mathbb{R}, \tag{3.3}$$

where (as in Groeneboom 1985), the function u_2 is defined by

$$u_2(t) = \lim_{x \uparrow t^2} \frac{\partial}{\partial t}u(t,x),\ t \in \mathbb{R}. \tag{3.4}$$

In fact, the original computations of the density were based on a numerical solution of this differential equation (this information is based on personal communications with Herman Chernoff and Willem van Zwet). The trouble with this approach is the behavior of the function u_2 for negative values of t. In fact, since, by (4.25) in Groeneboom (1985),

$$u_2(t) \sim c_1 \exp\left\{-\frac{2}{3}|t|^3 - c|t|\right\}, \quad t \to -\infty,$$

where $c \approx 2.9458\ldots$ and $c_1 \approx 2.2638\ldots$, the function u_2 tends to zero extremely rapidly, as t decreases away from zero. Some experiments with the numerical approach by the first author, in cooperation with B. Sommeyer, back in 1984, showed that this simple analytic fact invalidates any direct numerical approach, based on the partial differential equation: an error analysis showed that even with very fine grids the numerical solution was highly unstable. For this reason a more thorough analytic analysis of the problem was made, and the results of this analysis were given by Groeneboom (1985, 1989)

The following development is from Groeneboom (1985, sec. 4, pp. 548–553). Define a function $p : [0,\infty) \to R$ as follows:

$$p(y) = \begin{cases} -\sqrt{\frac{\pi}{2}}\sum_{k=0}^{\infty} a_k y^{3k} + \sum_{k=1}^{\infty} b_k y^{3(k-1/2)}\,, & \text{if} \quad y \in [0,1] \\ -y^{-3/2} + 2\sqrt{2\pi}\exp(-y^3/6)\sum_{k=1}^{\infty}\exp(2^{1/3}\tilde{a}_k y)\,, & \text{if} \quad y \in (1,\infty)\,. \end{cases} \tag{3.5}$$

Here the $\tilde{a_k}$'s are the zeros of the Airy function Ai, and a_k, b_k are defined recursively as follows: set $c_0 = 1$ and

$$c_n = -2^{-4}\frac{(2n-3)(2n+1)}{n^2(2n-1)}c_{n-1}\,, \qquad n = 1,2,\ldots\,.$$

The recursive relations for the coefficients a_k and b_k follow from the integral equation (4.14) in Groeneboom (1985). The integral equation leads to an accurate and useful analytic representation of the density in a neighborhood of zero, whereas the expansion on the second line of (3.5) does similar job away from zero.

Then with $a_0 = 1$, $b_1 = 2/3$, and $B(p,q) \equiv \Gamma(p)\Gamma(q)/\Gamma(p+q)$, the standard Beta function, set

$$a_n = c_n - \sum_{k=0}^{n-1}\frac{1}{\pi k!(-2)^k}b_{n-k}B(3n-2k-1/2, k+3/2)\,, \qquad n = 1,2,\ldots\,; \tag{3.6}$$

$$b_n = \sum_{k=0}^{n-1}\frac{1}{k!(-2)^{k+1}}a_{n-k-1}B(3n-2k-2, k+3/2)\,, \qquad n = 2,3,\ldots\,. \tag{3.7}$$

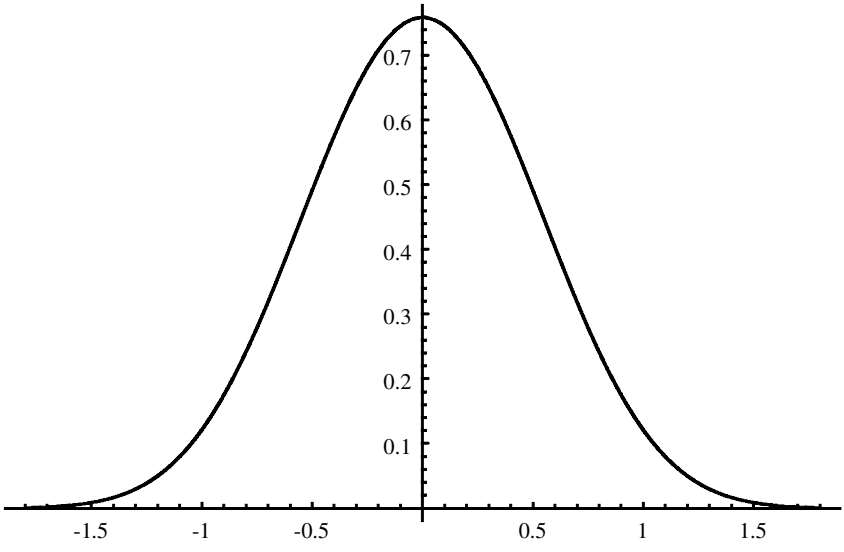

Figure 1. Density function of Z, f_Z.

The reason for treating the intervals $[0, 1]$ separately is that the series using the zeros of the Airy function diverges at zero and gives a bad approximation in neighborhoods of zero.

We then define $g : \mathbb{R} \to \mathbb{R}$ by

$$\begin{aligned} g(x) &= 2x - \frac{1}{\sqrt{2\pi}} \int_0^\infty p(y) \exp\left(-\frac{1}{2}y(2x+y)^2\right) dy \\ &\quad + 2\sqrt{\frac{2}{\pi}} \int_0^\infty \{(2x+y^2)y^2 + \frac{1}{2}(2x+y^2)^2\} \exp\left(-\frac{1}{2}y^2(2x+y^2)^2\right) dy \end{aligned} \tag{3.8}$$

if $x \in [-1, \infty)$, and

$$g(x) = \exp\left(\frac{2}{3}x^3\right) 4^{1/3} \sum_{k=1}^{\infty} \exp\left(-2^{1/3}\tilde{a}_k x\right) / Ai'(\tilde{a}_k) \tag{3.9}$$

if $x \in (-\infty, -1]$; here Ai' is the derivative of the Airy function Ai. The reason for using y^2 in the integrand of the first part of the definition of g instead of y as in Groeneboom (1985), is purely numerical: the present change of variables avoids a factor of $\sqrt{y}$ in the denominator of the integrand.

Finally, the density f_Z is expressed in terms of g as

$$f_Z(z) = \frac{1}{2} g(z) g(-z), \qquad z \in (-\infty, \infty)\,. \tag{3.10}$$

The distribution function F_Z of Z is simply

$$F_Z(z) = \int_{-\infty}^{z} f_Z(w) dw = \frac{1}{2} \int_{-\infty}^{z} g(w) g(-w) dw\,, \qquad z \in (-\infty, \infty)\,.$$

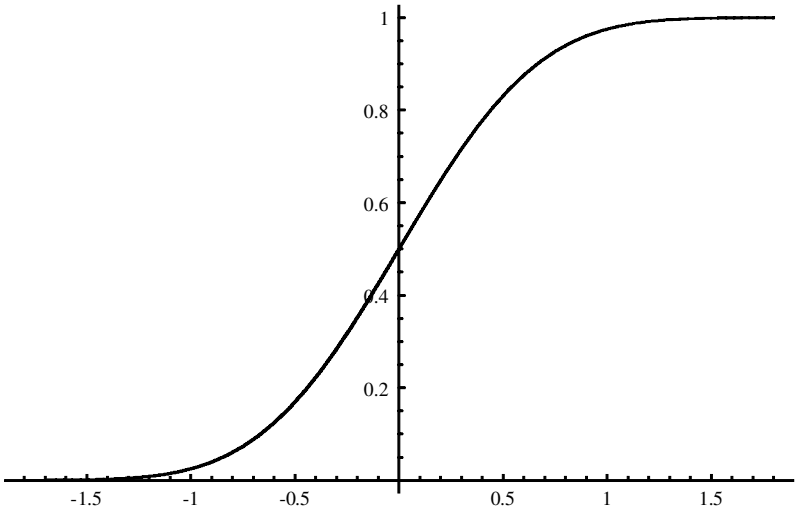

Figure 2. Distribution function of Z, F_Z.

Because of the symmetry of f_Z about 0, it suffices to calculate

$$F_Z(z) - F_Z(0) = \int_0^z f_Z(w)dw = \frac{1}{2}\int_0^z g(w)g(-w)dw\,, \qquad z \in [0,\infty)\,. \tag{3.11}$$

Figures 1 and 2 show plots of the density function f_Z and the distribution function F_Z, respectively; in these figures we used the first 20 terms of the series defining the function p, and also the first 20 terms of the series defining g in the region $(-\infty, -1)$. The figures shown here were produced by Mathematica; see Wolfram (1996).

The tables in this article were also first computed in Mathematica. Subsequently, a computer program, written in C, was developed. The program uses some routines for computing integrals and (zeros of) Airy functions by Lau (1995). These C routines are translations into C of the routines in the NUMAL library of the ALGOL 60 routines, developed at the Mathematical Centre, Amsterdam; see Hemker (1980). The results of the C program correspond in all decimals shown with the results obtained in Mathematica, except in a few cases where Mathematica could not reach sufficient accuracy (giving small differences in the last decimals). The C program was originally written on a Macintosh powerbook, using the Metrowerks Code Warrior C compiler and the sources could be compiled without any change on an HP Unix workstation by the standard C compiler available on this workstation. In the demonstration version of the program Table 1, of the this article is computed.

We also tried to provide a Microsoft Windows 98 executable, but ran into trouble here, since compilation by the Microsoft Visual C++ compiler, version 6.0, produced a so-called "release version" that either did not work or produced the wrong results (interestingly, depending on the machine on which it was run!). The so-called "debug version" worked better but was rather slow. Consulting the available information on differences between "debug" and "release" versions did not help us to solve the problem. Disabling the "maximize speed" option in compiling the release version also produced a correct but slowly working version!

Table 1. Values of the distribution function F_Z and density function f_Z

z	$F_Z(z)$	$f_Z(z)$	z	$F_Z(z)$	$f_Z(z)$	z	$F_Z(z)$	$f_Z(z)$
.00	.500000	.758345	.30	.716352	.649874	.60	.875858	.403594
.01	.507583	.758215	.31	.722817	.643059	.61	.879851	.394887
.02	.515163	.757828	.32	.729213	.636088	.62	.883756	.386214
.03	.522739	.757183	.33	.735538	.628967	.63	.887575	.377580
.04	.530306	.756281	.34	.741792	.621704	.64	.891308	.368989
.05	.537863	.755123	.35	.747972	.614303	.65	.894955	.360447
.06	.545408	.753709	.36	.754077	.606771	.66	.898517	.351960
.07	.552937	.752042	.37	.760107	.599115	.67	.901994	.343531
.08	.560448	.750122	.38	.766059	.591341	.68	.905388	.335166
.09	.567938	.747951	.39	.771933	.583455	.69	.908698	.326870
.10	.575406	.745532	.40	.777728	.575464	.70	.911925	.318646
.11	.582848	.742866	.41	.783442	.567374	.71	.915071	.310499
.12	.590263	.739957	.42	.789075	.559192	.72	.918136	.302433
.13	.597647	.736806	.43	.794626	.550925	.73	.921120	.294452
.14	.604998	.733416	.44	.800094	.542578	.74	.924025	.286560
.15	.612314	.729792	.45	.805477	.534159	.75	.926852	.278760
.16	.619593	.725935	.46	.810776	.525674	.76	.929601	.271057
.17	.626832	.721849	.47	.815991	.517130	.77	.932273	.263452
.18	.634029	.717539	.48	.821119	.508533	.78	.934870	.255950
.19	.641182	.713008	.49	.826161	.499890	.79	.937392	.248553
.20	.648289	.708260	.50	.831117	.491208	.80	.939841	.241264
.21	.655347	.703299	.51	.835985	.482492	.81	.942218	.234086
.22	.662354	.698131	.52	.840766	.473749	.82	.944523	.227020
.23	.669309	.692758	.53	.845460	.464986	.83	.946759	.220070
.24	.676208	.687187	.54	.850066	.456209	.84	.948925	.213237
.25	.683052	.681422	.55	.854584	.447424	.85	.951024	.206523
.26	.689836	.675469	.56	.859014	.438638	.86	.953056	.199931
.27	.696560	.669332	.57	.863357	.429855	.87	.955023	.193460
.28	.703222	.663017	.58	.867612	.421083	.88	.956926	.187113
.29	.709820	.656529	.59	.871779	.412327	.89	.958766	.180891

Some testing showed that Lau's integration routines derail the Microsoft Visual C++ release version (with "maximize speed" enabled), so replacing these by other integration routines may remove the problem. However, since the program worked fine on Unix and a Macintosh and even in a Microsoft Visual C++ "release" version with "maximize speed" disabled, we did not investigate this further. But, in view of these experiences, we recommend compiling the C sources either on a machine using the Unix operating system or on a Macintosh power PC, using Metrowerks Code Warrior. Compilation using Microsoft Visual C++ may produce unexpected results!

Dykstra and Carolan (1997) computed the density function f_Z by numerical Fourier inversion of formula (3.8) in Groeneboom (1989, p. 91). This section showed how f_Z is computable without numerical Fourier inversion.

Table 1 gives the distribution function $F_Z(z)$ and the density function $f_Z(z)$ for $z = 0.0(.01)2.0$. We took $n = 20$ in our computation of the power series; the demonstration program allows the user take a different number of terms before starting the computation. Our experience is that in going beyond 20 (in choosing the number of terms in the power series), the results did not change in 9 decimals. The C code also contains routines for computing moments and quantiles; these routines can all be found in the source "main.c". We used these routines in producing Tables 2–4.

Table 1. Continued

z	$F_Z(z)$	$f_Z(z)$	z	$F_Z(z)$	$f_Z(z)$	z	$F_Z(z)$	$f_Z(z)$
.90	.960544	.174795	1.27	.994448	.0346458	1.64	.999631	.0031329
.91	.962262	.168827	1.28	.994785	.0328273	1.65	.999662	.0028995
.92	.963921	.162985	1.29	.995105	.0310864	1.66	.999689	.0026817
.93	.965522	.157272	1.30	.995407	.0294208	1.67	.999715	.0024785
.94	.967067	.151687	1.31	.995693	.0278282	1.68	.999739	.0022891
.95	.968556	.146231	1.32	.995964	.0263065	1.69	.999761	.0021127
.96	.969992	.140904	1.33	.996220	.0248534	1.70	.999781	.0019485
.97	.971375	.135705	1.34	.996461	.0234666	1.71	.999800	.0017957
.98	.972706	.130635	1.35	.996689	.0221441	1.72	.999817	.0016537
.99	.973988	.125694	1.36	.996904	.0208837	1.73	.999833	.0015219
1.00	.975221	.120880	1.37	.997107	.0196831	1.74	.999848	.0013995
1.01	.976406	.116194	1.38	.997298	.0185404	1.75	.999861	.0012861
1.02	.977545	.111633	1.39	.997478	.0174534	1.76	.999873	.0011810
1.03	.978639	.107199	1.40	.997648	.0164201	1.77	.999885	.0010836
1.04	.979689	.102889	1.41	.997807	.0154385	1.78	.999895	.0009936
1.05	.980697	.0987031	1.42	.997957	.0145066	1.79	.999905	.0009103
1.06	.981664	.0946394	1.43	.998097	.0136226	1.80	.999913	.0008334
1.07	.982590	.0906969	1.44	.998229	.0127844	1.81	.999921	.0007625
1.08	.983478	.0868741	1.45	.998353	.0119902	1.82	.999929	.0006970
1.09	.984328	.0831697	1.46	.998469	.0112384	1.83	.999935	.0006367
1.10	.985142	.0795821	1.47	.998578	.0105269	1.84	.999941	.0005811
1.11	.985920	.0761096	1.48	.998680	.0098542	1.85	.999947	.0005300
1.12	.986665	.0727506	1.49	.998775	.0092187	1.86	.999952	.0004830
1.13	.987376	.0695033	1.50	.998864	.0086186	1.87	.999957	.0004399
1.14	.988055	.0663658	1.51	.998948	.0080523	1.88	.999961	.0004003
1.15	.988703	.0633362	1.52	.999025	.0075183	1.89	.999965	.0003639
1.16	.989322	.0604126	1.53	.999098	.0070151	1.90	.999968	.0003306
1.17	.989912	.0575930	1.54	.999166	.0065413	1.91	.999971	.0003001
1.18	.990474	.0548753	1.55	.999229	.0060955	1.92	.999974	.0002722
1.19	.991010	.0522574	1.56	.999288	.0056763	1.93	.999977	.0002468
1.20	.991520	.0497372	1.57	.999343	.0052824	1.94	.999979	.0002235
1.21	.992005	.0473125	1.58	.999394	.0049125	1.95	.999981	.0002022
1.22	.992466	.0449811	1.59	.999441	.0045655	1.96	.999983	.0001828
1.23	.992905	.0427408	1.60	.999485	.0042401	1.97	.999985	.0001652
1.24	.993321	.0405894	1.61	.999526	.0039352	1.98	.999986	.0001491
1.25	.993717	.0385246	1.62	.999564	.0036497	1.99	.999988	.0001345
1.26	.994092	.0365441	1.63	.999599	.0033826	2.00	.999989	.0001212

4. QUANTILES OF F_Z AND SOME COMPARISONS

Dykstra and Carolan (1997) suggested that f_Z and F_Z are closely approximated by the $N(0, (.52)^2)$ density and distribution functions respectively. Although this results in a simple approximation for the corresponding quantiles $F_Z^{-1}(p)$, the differences between the exact quantiles and the approximate quantiles, or exact distribution function and approximate distribution function based on the normal approximation can be substantial.

Table 2 compares a few quantiles computed directly by inverting the distribution function, computed in the preceding section, with analytically computed, approximate (by a normal distribution approximation), and Monte Carlo quantiles as computed by Dykstra and Carolan (1997), Narayanan and Sager (1989), and Keiding et al. (1996). The Dykstra and Carolan (1997) approach seems to fail in the tail, and we indeed believe that it is absolutely necessary to use different representations of the density in a neighborhood of zero and in the tail (which is a common phenomenon in the numerical evaluation of special functions), whereas Dykstra and Carolan (1997) essentially used the same representation for the

Table 2. Comparison of several computations and estimators of the quantiles $F_Z^{-1}(p)$ for certain p's

Percentile p	Exact $F_Z^{-1}(p)$	Dykstra and Carolan; Fourier inversion	$N(0, (.52)^2)$	Narayanan and Sager; Monte Carlo	Keiding et al. Monte Carlo
.9	.664235	.664	.666	.658	.66
.95	.845081	.846	.856	.838	.836
.975	.998181	.998	1.018	.986	1.009
.99	1.171530	1.156	1.205	1.156	1.176
.995	1.286659	1.270	1.314	1.281	1.278
.999	1.516664	1.452	1.515	1.510	NA

whole domain. The normal approximation clearly cannot be good for the whole domain; in this case it is reasonable for the outer values, but not so good for the intermediate values. The results of the Monte Carlo simulation of Narayanan and Sager seem pretty good for the values tabulated here and slightly better than the Monte Carlo simulation results reported by Keiding et al. (1996) Neverthless, there is no need for Monte Carlo simulations and these will generally produce results deteriorating rapidly if one goes further out in the tail.

Table 3 gives further quantiles of the distribution F_Z.

Table 3. Quantiles of $F_Z^{-1}(p)$ for $p = .5(.01).99$ and $p = .99(.001).999$

p	$F_Z^{-1}(p)$	p	$F_Z^{-1}(p)$
.50	.00	.80	.439828
.51	.013187	.81	.458525
.52	.026383	.82	.477804
.53	.039595	.83	.497731
.54	.052830	.84	.518383
.55	.066096	.85	.539855
.56	.079402	.86	.562252
.57	.092757	.87	.585706
.58	.106168	.88	.610378
.59	.119645	.89	.636468
.60	.133196	.90	.664235
.61	.146831	.91	.694004
.62	.160560	.92	.726216
.63	.174393	.93	.761477
.64	.188342	.94	.800658
.65	.202418	.95	.845081
.66	.216633	.96	.896904
.67	.230999	.97	.960057
.68	.24553	.98	1.043030
.69	.260242	.99	1.171530
.70	.275151	.991	1.189813
.71	.290274	.992	1.209897
.72	.305629	.993	1.232241
.73	.321238	.994	1.257496
.74	.337123	.995	1.286659
.75	.353308	.996	1.321370
.76	.369821	.997	1.364637
.77	.386694	.998	1.423026
.78	.403959	.999	1.516664
.79	.421656	.9999	1.784955

Table 4. Absolute Moments of Z; $E(|Z|^k)$, $k = 1(1)10$

k	$E(\|Z\|^k)$
1	.41273655
2	.26355964
3	.21135025
4	.19715702
5	.20573334
6	.23455025
7	.28760426
8	.37509901
9	.51604236
10	.74410271

5. MOMENTS OF Z

As remarked in the introduction, the first four moments were computed by Groeneboom and Sommeijer (1984). Table 4 shows the first 10 moments of Z.

ACKNOWLEDGMENTS

The first author's research was supported in part by National Science Foundation grant DMS-9743445. The second author's research was supported in part by National Science Foundation grant DMS-9532039 and NIAID grant 2R01 AI291968-04. AMS 1980 subject classifications. Primary: 62E17; secondary 65D20.

[Received July 1999. Revised November 1999.]

REFERENCES

Ayer, M., Brunk, H. D., Ewing, G. M., Reid, W. T., Silverman, E. (1955), "An Empirical Distribution Function for Sampling With Incomplete Information," *Annals of Mathematical Statistics*, 26, 641–647.

Chernoff, H. (1964), "Estimation of the Mode," *Annals of Statistical Mathematics*, 16, 31–41.

Daniels, H. E., and Skyrme, T. H. R. (1985), "The Maximum of a Random Walk Whose Mean Path has a Maximum," *Advances in Applied Probability*, 17, 85–99.

Devroye, L., and Györfi, L. (1985), *Nonparametric Density Estimation: the L_1 View*, New York: Wiley.

Dykstra, R., and Carolan, C. (1997), "The Distribution of the Argmax of Two-Sided Brownian Motion With Quadratic Drift," Technical Report 266, Department of Statistics, University of Iowa.

Geskus, R. B., and Groeneboom, P. (1999), "Asymptotically Optimal Estimation of Smooth Functionals for Interval Censoring, case 2," *The Annals of Statistics*, 27, 627–674.

Grenander, U. (1956), "On the Theory of Mortality Measurement. Part II," *Scandinavisk Aktuarietidskrift*, 39, 125–153.

Groeneboom, P. (1985), "Estimating a Monotone Density," in *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer, II*, eds. L. M Le Cam and R. A. Olshen, Belmont, CA: Wadsworth; New York: Springer-Verlag, pp. 535–555.

——— (1989), "Brownian Motion With a Parabolic Drift and Airy Functions," *Probability Theory and Related Fields*, 81, 79–109.

——— (1996), "Lectures on Inverse Problems," in *Lectures on Probability Theory, Ecole d'Eté de Probabilités de Saint-Flour XXIV-1994*, ed. P. Bernard, Berlin: Springer-Verlag.

Groeneboom, P., and Lopuhaä, H. P. (1993), "Isotonic Estimators of Monotone Densities and Distribution Functions: Basic Facts," *Statistica Neerlandica*, 47, 175–1833.

Groeneboom, P., and Sommeijer, B. (1984), "Computation of the Distribution of the Location of the Maximum of Brownian Motion Minus a Parabola," unpublished technical report, University of Amsterdam.

Groeneboom, P., and Wellner, J. A. (1992), *Information Bounds and Nonparametric Maximum Likelihood Estimation*, Basel: Birkhäuser.

Hemker, P. W. (ed.) (1980), *NUMAL, Numerical Procedures in ALGOL 60*, MC Syllabus, 47.1-47.7, Mathematisch Centrum (now CWI), Amsterdam.

Jongbloed, G. (1998a), "Exponential Deconvolution: Two Asymptotically Equivalent Estimators," *Statistica Neerlandica*, 52, 6–17.

——— (1998b), "The Iterative Convex Minorant Algorithm for Nonparametric Estimation," *Journal of Computational and Graphical Statistics*, 7, 310–321.

Keiding, N., Begtrup, K., Scheike, T. H., Hasibeder, G. (1996), "Estimation From Current-Status Data in Continuous Time," *Lifetime Data Analysis*, 2, 119–129.

Kim, J., and Pollard, D. (1990), "Cube Root Asymptotics," *The Annals of Statistics*, 18, 191–219.

Lau, H. T. (1995), *A Numerical Library in C for Scientists and Engineers*, Boca Raton, FL: CRC Press.

Narayanan, A., and Sager, T. W. (1989), "Table for the Asymptotic Distribution of Univariate Mode Estimators," *Journal of Statistical Computing and Simulation*, 33, 37–51.

Politis, D. N., and Romano, J. P. (1994), "Large Sample Confidence Regions Based on Subsamples Under Minimal Assumptions," *The Annals of Statistics*, 22, 2031–2050.

Prakasa Rao, B. L. S. (1969), "Estimation of a Unimodal Density," *Sankhya A*, 31, 23–36.

——— (1970), "Estimation for Distributions With Monotone Failure Rate," *The Annals of Mathematical Statistics*, 41, 507–519.

Rousseeuw, P. J. (1984), "Least Median of Squares Regression," *Journal of American Statistical Association*, 79, 871–880.

van Es, A. J. (1991), "Aspects of Nonparametric Density Estimation," CWI Tract 77, Amsterdam: Centre for Mathematics and Computer Science.

van Es, A. J., Jongbloed, G., and van Zuijlen, M. C. A. (1998), "Isotonic Inverse Estimators for Nonparametric Deconvolution," *The Annals of Statistics*, 26, 2395–2406.

van Es, A. J., and van Zuijlen, M. C. A. (1996), "Convex Minorant Estimators of Distributions in Nonparametric Deconvolution Problems," *Scandanavian Journal of Statistics*, 23, 85–104.

Venter, J. H. (1967), "On Estimation of the Mode," *Annals of Mathematical Statistics*, 38, 1446–1455.

Wellner, J. A., and Zhang, Y. (1998), "Two Estimators of the Mean of a Counting Process With Panel Count Data," Technical Report 341, Department of Statistics, University of Washington, Seattle; http://www.stat.washington.edu/www/research/reports/

Wolfram, S. (1996), *The Mathematica Book* (3rd ed.), Champaign, IL: Wolfram Research.