The Annals of Statistics
1987, Vol. 15, No. 2, 875–879

MONOTONE EMPIRICAL BAYES TEST FOR UNIFORM DISTRIBUTIONS USING THE MAXIMUM LIKELIHOOD ESTIMATOR OF A DECREASING DENSITY

BY J. C. VAN HOUWELINGEN

University of Leiden

An empirical Bayes test for testing $\vartheta \leq \vartheta_0$ against $\vartheta > \vartheta_0$ for the uniform distribution on $[0, \vartheta)$ is discussed. The relation is shown with the estimation of a decreasing density on $[0, \infty)$ and a monotone empirical Bayes test is derived based on the least-concave majorant of the empirical distribution function. The asymptotic distribution of the Bayes risk is obtained and some Monte Carlo results are given.

1. Construction. Suppose that the random variable X has $U[0, \vartheta]$-distribution with pdf $f(x|\vartheta) = \vartheta^{-1} I_{[0, \vartheta]}(x)$, $\vartheta > 0$, and that θ has prior distribution with cdf G. For testing H_0: $\vartheta \leq \vartheta_0$ against H_1: $\vartheta > \vartheta_0$ with loss function $L(\vartheta, a_0) = (\vartheta - \vartheta_0)^+$, $L(\vartheta, a_1) = (\vartheta_0 - \vartheta)^+$, where a_i is the decision to accept H_i, a Bayes test w.r.t. G is given by $\varphi(x) = a_0$ if $d(x) < \vartheta_0$, $\varphi(x) = a_1$ if $d(x) > \vartheta_0$, $\varphi(x) =$ arbitrary if $d(x) = \vartheta_0$, where $d(x) = E(\theta | X = x)$.

Since the class $\{f(x|\vartheta) | \vartheta > 0\}$ has monotone likelihood ratio in x, $d(x)$ is nondecreasing and $\varphi(x) = a_0$ if $x \leq c_0$, $\varphi(x) = a_1$ if $x > c_0$ is Bayes w.r.t. G provided $d(x) \leq \vartheta_0$ for $x < c_0$ and $d(x) \geq \vartheta_0$ for $x > c_0$. Recall that the class of all monotone tests is essentially complete; see Ferguson (1967).

In the empirical Bayes set up the prior G is unknown but an iid sample $X_1, \dots, X_n$, independent of X is available from $f(x) = \int f(x|\vartheta)\, dG(\vartheta)$ the marginal density of X. A monotone empirical Bayes test (MEBT) can be constructed by estimating c_0 by $c_n(X_1, \dots, X_n)$, say, and defining

$$\varphi_n(x) = a_0, \quad \text{if } x \leq c_n, \qquad \varphi_n(x) = a_1, \quad \text{if } x > c_n.$$

The increase in Bayes risk due to replacing of c_0 by c_n is given by

$$\Delta_n = E[L(\theta, \varphi_n(X)) - L(\theta, \varphi(X))] = \int_{c_0}^{c_n} (d(x) - \vartheta_0) f(x)\, dx.$$

The MEBT φ_n is called asymptotically optimal [Robbins (1955)] if $\Delta_n \to 0$ (P).

MEBT's for the one-parameter exponential family were presented by van Houwelingen (1976) and Stijnen (1982). In this paper an MEBT will be constructed for the uniform distribution. Interest in this problem was raised by Gupta and Hsiao (1983), where it was suggested that such a construction is difficult.

An estimator c_n of c_0 can be constructed by using the results of Fox (1978) and Grenander (1956).

Received September 1984; revised November 1986.

AMS 1980 *subject classifications*. Primary 62C12; secondary 62F12.

Key words and phrases. Empirical Bayes, monotone test, uniform distribution, decreasing density, maximum likelihood estimator, concave majorant.

LEMMA 1 (Fox).

$$d(x) = E[\Theta|X = x] = [1 - F(x)]/f(x) + x,$$

where $f(x) = \int_x^\infty \vartheta^{-1}\, dG(\vartheta)$ *is the marginal density of* X *and* $F(x)$ *is the corresponding cdf.*

Observe that f is a nonincreasing density on $(0, \infty)$ and that the continuity points of f and G coincide. The monotonicity of $d(x)$ could be derived directly from Lemma 1.

In order to avoid degeneracy of $\varphi(X)$, it is assumed that

$$\lim_{x \downarrow 0} d(x) < \vartheta_0 \quad \text{and} \quad \lim_{x \uparrow \vartheta_0} d(x) > \vartheta_0.$$

This is equivalent with $E(\Theta^{-1}) > \vartheta_0^{-1}$ and $P(\Theta > \vartheta_0) > 0$. It is easy to verify that the assumption implies $0 < c_0 < \vartheta_0$.

The importance of Fox's lemma is that it gives an explicit expression for $d(x)$ in terms of the marginal distribution of X, which enables the estimation of $d(x)$ from $X_1, \ldots, X_n$.

Grenander (1956) showed that the MLE of f in the class of all nonincreasing densities on $(0, \infty)$ is given by the derivative f_n of the least-concave majorant $\hat{F}_n$ on $[0, \infty)$ of the empirical cdf F_n. It is straightforward to check that $d_n(x) = (1 - \hat{F}_n(x))/f_n(x) + x$ is nondecreasing on $[0, X^{(n)})$. The definition of d_n can be extended to $[0, \infty)$ by setting $d_n(x) = x$ for $x \geq X^{(n)}$. An estimator c_n of c_0 is obtained by defining

$$c_n = \sup\{x | x \geq 0,\ d_n(x) < \vartheta_0\}.$$

There exists an equivalent expression for c_n that facilitates both the computation and the derivation of asymptotic results.

LEMMA 2.

$$c_n = \vartheta_0, \quad \textit{if } X^{(n)} < \vartheta_0$$
$$= \textit{smallest minimizer of } (1 - F_n(x))/(\vartheta_0 - x) \textit{ on } [0, \vartheta_0), \textit{ if } X^{(n)} \geq \vartheta_0.$$

PROOF. The result is obvious if $X^{(n)} < \vartheta_0$. In case $X^{(n)} \geq \vartheta_0$, let c_n^* be the smallest minimizer of $(1 - F_n(x))/(\vartheta_0 - x)$ on $[0, \vartheta_0)$. Since the line through $(c_n^*, F_n(c_n^*))$ and $(\vartheta_0, 1)$ is a tangent to the graph of $F_n(x)$, $F_n(c_n^*) = \hat{F}_n(c_n^*)$ and c_n^* is also the smallest minimizer of $(1 - \hat{F}_n(x))/(\vartheta_0 - x)$ on $[0, \vartheta_0)$.

Let $y < c_n^*$. Then $(1 - \hat{F}_n(c_n^*))/(1 - \hat{F}_n(y)) < (\vartheta_0 - c_n^*)/(\vartheta_0 - y)$. Since $\hat{F}_n$ is concave,

$$\begin{aligned} f_n(y) &\geq (\hat{F}_n(c_n^*) - \hat{F}_n(y))/(c_n^* - y) \\ &\geq (1 - \hat{F}_n(y))(1 - (\vartheta_0 - c_n^*)/(\vartheta_0 - y))/(c_n^* - y) \\ &= (1 - \hat{F}_n(y))/(\vartheta_0 - y). \end{aligned}$$

Hence, $d_n(y) < \vartheta_0$. A similar argument shows that $d_n(y) \geq \vartheta_0$ for $y > c_n^*$. Combination gives $c_n^* = c_n$. □

Observe that c_0 minimizes $(1 - F(x))/(\vartheta_0 - x)$ on $[0, \vartheta_0)$, which gives a direct motivation of c_n as defined in Lemma 2 as an estimator of c_0. The advantage of referring to Grenander is that not only an MEBT is obtained but also a monotone estimator of $d(x)$ that could serve as an MEBE (monotone empirical Bayes estimator) of ϑ under squared error loss.

It is easy to show that actually $c_n = X^{(i)}$ for some i. Therefore, it suffices to find the smallest $X^{(i)}$ that minimizes $(1 - i/n)/(\vartheta_0 - X^{(i)})$.

2. Asymptotics. In this section it will be shown that φ_n is asymptotically optimal. Moreover, the asymptotic distribution of $c_n - c_0$ and Δ_n will be derived under the additional assumption that $f'(c_0)$ exists and $-\infty < f'(c_0) < 0$, which is equivalent to $0 < G'(c_0) < \infty$. Observe that in that case c_0 is unique. The asymptotic distribution obtained below is closely related with the asymptotic distribution of the mode estimator of Chernoff (1964) and the asymptotic distribution of $f_n(x)$ in Prakasa-Rao (1969).

LEMMA 3.

(i) $\Delta_n = O_P(n^{-1/2})$.

(ii) *Under the assumption* $-\infty < f'(c_0) < 0$, $n^{1/3}(c_n - c_0) \to_{\mathscr{L}} AZ$ *and* $n^{2/3}\Delta_n \to_{\mathscr{L}} BZ^2$, *where*

$$A = \left[4f(c_0)/f'(c_0)^2\right]^{1/3},$$

$$B = [1 - F(c_0)][-f(c_0)f'(c_0)/2]^{-1/3}$$

and Z is the almost-sure unique minimizer of $W^*(x) + x^2$, where $W^*(x)$ is standard two-sided Brownian motion.

PROOF. Recall that

$$0 \le \Delta_n = \int_{c_0}^{c_n} (d(x) - \vartheta_0) f(x)\, dx$$

$$= \int_{c_0}^{c_n} [1 - F(x) - (\vartheta_0 - x) f(x)]\, dx$$

by Lemma 1. Using the concavity of F for $x \ge c_0$ and the inequality $(1 - F(x))/(\vartheta_0 - x) \ge (1 - F(c_0))/(\vartheta_0 - c_0)$ for $x \le c_0$, it can be shown that

$$\Delta_n \le [\vartheta_0/(\vartheta_0 - c_0)] \int_{c_0}^{c_n} [1 - F(c_0) - (\vartheta_0 - c_0) f(x)]\, dx.$$

Define $R_n(x) = F_n(x) - F(x)$. After working out the integral, it follows from Lemma 2 that

$$(*) \qquad 0 \le \Delta_n \le [\vartheta_0/(\vartheta_0 - c_0)][(\vartheta_0 - c_0)R_n(c_n) - (\vartheta_0 - c_n)R_n(c_0)].$$

Since $\sup|R_n(x)| = O_P(n^{-1/2})$, this proves part (i).

The proof of part (ii) resembles the proof in Prakasa-Rao (1969). First, it is established that $c_n - c_0 = O_P(n^{-1/3})$ and then the limit distribution of $n^{1/3}(c_n - c_0)$ is derived.

Under the assumption $-\infty < f'(c_0) < 0$, $(*)$ implies that

$$(\vartheta_0 - c_0)R_n(c_n) - (\vartheta_0 - c_n)R_n(c_0) \ge C(c_n - c_0)^2,$$

for some $C > 0$. From this it follows that

$$P(|c_n - c_0| \geq z)$$
$$\leq P\Big(\sup_{|x - c_0| \geq z} \big[((\vartheta_0 - c_0)R_n(x) - (\vartheta_0 - x)R_n(c_0))/(x - c_0)^2\big] \geq C \Big).$$

Using the embedding theorem of Komlós, Major and Tusnády (1975), the relation between Brownian bridge and Brownian motion, the scaling property of Brownian motion and the law of the iterated logarithm, it can now be shown that

$$\lim_{y \to \infty} \limsup_{n \to \infty} P\big(|c_n - c_0| \geq yn^{-1/3}\big) = 0, \quad \text{i.e., } (c_n - c_0) = O_P(n^{-1/3}).$$

Define $T_n = n^{1/3}(c_n - c_0)$. By Lemma 2, T_n is the minimizer of

$$U_n(t) = n^{2/3}\big[(\vartheta_0 - c_0)\big(1 - F_n\big(c_0 + tn^{-1/3}\big)\big)$$
$$/\big(\vartheta_0 - c_0 - tn^{-1/3}\big) - \big(1 - F_n(c_0)\big)\big].$$

The process $U_n(t)$ can be split in a nonstochastic term that converges to $-1/2 f'(c_0)t^2$, uniformly on $[-K, K]$ and a stochastic part

$$n^{2/3}\big[R_n\big(c_0 + tn^{-1/3}\big) - R_n(c_0)\big] + o_P(1)$$

uniformly on $[-K, K]$. By using the same techniques mentioned above, it can be shown that the stochastic part converges in distribution to $W^*(f(c_0)t)$ on $[-K, K]$.

Since $W^*(at) + bt^2$, $b > 0$, has an almost sure unique minimizer $Z_{a,b}$ on $(-\infty, \infty)$ [see Section 4 of Chernoff (1964)], it follows that $n^{1/3}(c_n - c_0) \to_{\mathscr{L}} Z_{a,b}$ with $a = f(c_0)$ and $b = -1/2 f'(c_0)$.

Finally, $Z_{a,b} =_{\mathscr{L}} (ab^{-2})^{1/3} Z_{1,1}$ by the scaling property of Brownian motion. The result about Δ_n follows from

$$\Delta_n = -\tfrac{1}{2} f'(c_0)(\vartheta_0 - c_0)(c_n - c_0)^2 + o\big((c_n - c_0)^2\big)$$

and $\vartheta_0 - c_0 = (1 - F(c_0))/f(c_0)$. □

Chernoff (1964) also establishes the relation between the density of Z and the heat equation. More details about the distribution of Z can be found in a forthcoming paper by Groeneboom (1986). In order to get an impression of the validity of the asymptotic distributions, a Monte Carlo simulation based on 4000 repetitions for each sample size n was performed for the case $dG(\vartheta) = \vartheta e^{-\vartheta}$, $\vartheta > 0$, $f(x) = e^{-x}$, $d(x) = x + 1$, $\vartheta_0 = 2$ and $c_0 = 1$. The following result was obtained:

n (sample size)	10	20	30	40	50	75	100
$\bar{\Delta}_n$ (average value of Δ_n)	0.045	0.034	0.026	0.022	0.019	0.014	0.012
$n^{2/3}\bar{\Delta}_n$	0.21	0.25	0.25	0.26	0.25	0.25	0.25

It seems that the limiting distribution is valid for rather small values of n.

3. Extensions. Gupta and Hsiao (1983) consider the case where $f(x|\vartheta) = p(x)c(\vartheta)I_{[0,\vartheta]}(x)$ and $L(\vartheta, a_0) - L(\vartheta, a) = \vartheta - \vartheta_0$. It is not hard to show that in that case

$$d(x) = x + \int_x (f(y)/p(y))\,dy/(f(x)/p(x)),$$

and that c_0 is the value of x that minimizes $\int_x (f(y)/p(y))\,dy/(\vartheta_0 - x)$. Hence, c_0 can be estimated by the minimizer of $\sum_{X_i > x}(1/p(X_i))/(\vartheta_0 - x)$. It is just a matter of technique to obtain the asymptotic distribution of $c_n - c_0$ along the lines of Section 2.

Another extension is to generalize the loss function to the case where

$$L(\vartheta) = L(\vartheta, a_0) - L(\vartheta, a_1) = \int_{\vartheta_0}^{\vartheta} l(t)\,dt,$$

with $l(\vartheta) > 0$. In the case of the standard uniform distribution on $[0, \vartheta]$, the Bayes test is again a monotone test with cut-off point c_0 = minimizer of $\int_x^\infty l(y)f(y)\,dy/\int_x^{\vartheta_0} l(y)\,dy$ on $[0, \vartheta_0)$. The construction of an MEBT for this problem is obvious.

REFERENCES

CHERNOFF, H. (1964). Estimation of the mode. *Ann. Inst. Statist. Math.* **16** 31–41.

FERGUSON, T. S. (1967). *Mathematical Statistics: A Decision Theoretic Approach.* Academic, New York.

FOX, R. J. (1978). Solution to empirical Bayes squared error loss estimation problems. *Ann. Statist.* **6** 846–854.

GRENANDER, U. (1956). On the theory of mortality measurement. II. *Skand. Aktuarietidskr.* **39** 125–153.

GROENEBOOM, P. (1986). Brownian motion with a parabolic drift and Airy functions. *Probab. Theory Rel. Fields.* To appear.

GUPTA, S. S. and HSIAO, P. (1983). Empirical Bayes rules for selecting good populations. *J. Statist. Plann. Inference* **8** 87–101.

KOMLÓS, J., MAJOR, P. and TUSNÁDY, G. (1975). An approximation of partial sums of independent r.v.'s and the sample d.f. *Z. Wahrsch. verw. Gebiete* **32** 111–131.

PRAKASA-RAO, B. L. S. (1969). Estimation of a unimodal density. *Sankyā Ser. A* **31** 23–36.

ROBBINS, H. (1955). An empirical Bayes approach to statistics. *Proc. Third Berkeley Symp. Math. Statist. Probab.* **1** 157–164. Univ. California Press.

STIJNEN, TH. (1982). A monotone empirical Bayes estimator and test for the one-parameter exponential family based on spacings. *Scand. J. Statist.* **9** 153–158.

VAN HOUWELINGEN, J. C. (1976). Monotone empirical Bayes tests for the continuous one-parameter exponential family. *Ann. Statist.* **4** 981–989.

DEPARTMENT OF MEDICAL STATISTICS
UNIVERSITY OF LEIDEN
P.O. BOX 9512
2300 RA LEIDEN
THE NETHERLANDS