The Annals of Statistics
1975, Vol. 3, No. 3, 680–687

ON THE MINIMAX ESTIMATION OF A RANDOM PROBABILITY WITH KNOWN FIRST N MOMENTS

BY V. M. JOSHI

Economics and Statistics Bureau, Maharashtra Government

X and Θ are random variables; for given $\Theta = \theta$, the conditional distribution of X is binomial with parameters N and θ; the first N moments of Θ are known. An estimate of Θ is made based on the observed value of X, the risk being defined in terms of squared error loss. It is shown that as conjectured by H. Robbins, the ratio of the Bayes risk to the minimax risk for all possible distributions of Θ uniformly tends to unity when $N \to \infty$.

1. Introduction. In this paper we prove a conjecture of H. Robbins (cf. Skibinsky (1968) page 501), relating to a minimax problem. The problem has been dealt with by Skibinsky (1968) and has the following structure. X and Θ are random variables defined on a measurable space $(\Omega, \mathscr{A})$; X is distributed on the set of integers $0, 1, 2, \cdots, N$ where N is fixed, and Θ on the unit interval $[0, 1]$; further for given $\Theta = \theta$, the conditional distribution of X is binomial with parameters N and θ; $\mathscr{P}$ is the class of all probability distributions on $\mathscr{A}$ which give rise to the above structure; any real function $t = t(X)$ of X is an estimate of Θ; the risk of an estimate t relative to a probability distribution P is defined as

$$R_N(t, P) = E_P(t(X) - \Theta)^2 . \tag{1}$$

The Bayes risk is obtained by minimizing the risk over the class of all estimates t, i.e.

$$R_N(P) = \inf_t E_P(t(X) - \Theta)^2 . \tag{2}$$

The first N moments of Θ under P (which for brevity we shall hereafter refer to as the 'first N moments of P'), have known values $c_1, c_2, \cdots, c_N$; i.e.

$$E_P \Theta^i = \int_\Omega \Theta^i \, dP = c_i , \qquad i = 1, 2, \cdots, N ; \tag{3}$$

$\mathbf{c}$ denotes the vector $(c_1, c_2, \cdots, c_N)$ and $\mathscr{M}_N(\mathbf{c})$ the subclass of all probability distributions $P \in \mathscr{P}$, for which (3) holds. The class $\mathscr{M}_N(\mathbf{c})$ is not empty if and only if there exists a probability measure Q on the Borel subsets of $[0, 1]$, such that

$$c_i = \int_{[0,1]} x^i \, dQ , \qquad i = 1, 2, \cdots, N . \tag{4}$$

We assume throughout the following that (4) is satisfied so that the class $\mathscr{M}_N(\mathbf{c})$ is not empty. Then by standard game-theoretic results (Remarks (6) and (7) in

Received August 1972; revised May 1974.
AMS 1970 *subject classification.* Primary 62C10.
Key words and phrases. Minimax estimation, random probability, first N moments known.

[2]), $\mathscr{M}_N(\mathbf{c})$ contains a (least favourable) distribution P^* for which

$$R_N(P^*) \geqq R_N(P) \qquad \text{for } P \in \mathscr{M}_N(\mathbf{c}) ,$$

and

$$R_N(P^*) \leqq \sup_{P \in \mathscr{M}_N(\mathbf{c})} R_N(t, P) \qquad \text{for all } t .$$

$R_N(P^*)$ is thus the minimax risk over the class $\mathscr{M}_N(\mathbf{c})$. It depends only on $\mathbf{c}$, i.e. the first N moments of P. Hence we put

$$W_N(P) = R_N(P^*) - R_N(P) .$$

$W_N(P)$ is the excess of the minimax risk over the Bayes risk. Skibinsky [2] has shown that

$$W_N(P) \leqq 2^N r_N \qquad \text{for all } P \in \mathscr{P}, \tag{5}$$

where r_N is a constant whose value is determined by the first N moments of P, and which satisfies $r_N \leqq 4^{-N}$. The conjecture of Robbins is

$$\lim_{N\to\infty} \sup_{P \in \mathscr{P}} (W_N(P)/R_N(P)) = 0 . \tag{6}$$

We note that the supremum in (6) has to be taken over only those $P \in \mathscr{P}$ for which $R_N(P)$ and $W_N(P)$ do not both vanish, as otherwise their ratio becomes indeterminate.

2. Lower bound for Bayes risk. We first outline the argument. It has been shown by Skibinsky [2] that the Bayes risk $R_N(P)$ attains its minimum either for $P = P^+$ or for $P = P^-$, these being distributions for which the $(N + 1)$st moment of P is respectively maximized and minimized for $P \in \mathscr{M}_N(\mathbf{c})$. Hence a lower bound for the Bayes risk is obtained by obtaining lower bounds for $R_N(P^+)$ and $R_N(P^-)$. Defining $\xi_{n,j}(P)$ as in (7), it follows from (13) that a lower bound for $R_N(P^-)$ is obtained by taking an upper bound for $\xi^2_{N+1,j+1}(P^-)/\xi_{N,j}(P^-)$ and correspondingly for $R_N(P^+)$. We next consider the $(N + 2)$nd moments of P^+ and P^- which depend only on $\mathbf{c}$. Denoting their difference by λ_N, we obtain mainly by using Schwarz' inequality, the upper bound in (21) for $\xi^2_{N+1,J+1}(P^-)/\xi_{N,j}(P^-)$ which involves λ_N. An upper bound for λ_N is obtained by defining the conjugate prior distribution in (22), and it is shown in (31) that either $\lambda_N \leqq \frac{1}{2}(N + 2)r_N$ or for the conjugate prior $\lambda_N' \leqq \frac{1}{2}(N + 2)r_N$ where r_N is the difference between the maximum and minimum values of the $(N + 1)$st moment. Using these upper bounds, we obtain an upper bound for $\xi^2_{N+1,j+1}(P^-)/\xi_{N,j}(P^-)$ by (21) and then a lower bound for $R_N(P^-)$ by (13). A similar calculation gives a lower bound for $R_N(P^+)$. Combining the two we get a lower bound for $R_N(P)$. This combined with the upper bound for $W_N(P)$ in (5) gives an upper bound for the supremum in (6). The final result is obtained by taking limits as $N \to \infty$ as in Section 3. Throughout the calculation $\mathbf{c}$ is assumed to be an interior point of $\mathscr{M}_N(\mathbf{c})$. If $\mathbf{c}$ is a boundary point, then $W_N(P) = 0$, so that for such points also the result holds.

We now give the detailed proof. For any integers $n, j, j \leqq n$, let

$$\xi_{n,j}(P) = E_P\{\Theta^j(1 - \Theta)^{n-j}\} . \tag{7}$$

If $m_0, m_1, m_2 \cdots$ denote the successive moments of P for $P \in \mathscr{P}$, i.e. $m_i = E_P \Theta^i = \int_\Omega \Theta^i \, dP$, $i = 1, 2, \cdots$, then

$$(8) \qquad \xi_{n,j}(P) = \sum_{r=0}^{n-j} \binom{n-j}{r} (-1)^r m_{r+j} .$$

We now introduce a restriction on the values of $\mathbf{c}$. Let $\mathscr{Q}$ denote the class of all probability measures on the Borel subsets of $[0, 1]$, and let

$$(9) \qquad M_n = \{(c_1, c_2, \cdots, c_N) : c_i = \int_{[0,1]} x^i \, dQ(x), \; i = 1, 2, \cdots, N, \; Q \in \mathscr{Q}\} .$$

If $\mathbf{c}$ is a boundary point of M_N, then by Remark 6 in [2], $W_N(P) = 0$. In the following therefore we make

ASSUMPTION 2.1. $\mathbf{c}$ is an interior point of the set M_N in (6). Then (see equation (3) in [2]),

$$(10) \qquad \xi_{N,j}(P) > 0 , \quad \text{all} \quad j \leqq N , \qquad \text{and all} \quad P \in \mathscr{M}_N(\mathbf{c}) ,$$

Now by (1) and putting $t(X) = t_j$ for $X = j$, we obtain as shown in Remark 4 of [2],

$$(11) \qquad R_N(t, P) = c_2 - \sum_{j=0}^{N} \binom{N}{j} \{2t_j \xi_{N+1,j+1}(P) - t_j^2 \xi_{N,j}(P)\} .$$

As $\xi_{Nj}(P) > 0$ by (10), the right-hand side of (11) is minimized by putting

$$(12) \qquad t_j = \xi_{N+1,j+1}(P)/\xi_{N,j}(P) ,$$

so that we obtain the Bayes risk

$$(13) \qquad R_N(P) = c_2 - \sum_{j=0}^{N} \binom{N}{j} \xi^2_{N+1,j+1}(P)/\xi_{N,j}(P) .$$

(12) and (13) are given in Remark 5 and formula (4) in [2].

We now use the following results given in [2], (equations (11) and (17) of [2]): When the first N moments of P have fixed values, its $(N+1)$st moment has a maximum value, say ν^+_{N+1} and a minimum value ν^-_{N+1} which are respectively attained for some probability distributions P^+ and P^-, and

$$(14) \qquad \nu^+_{N+1} - \nu^-_{N+1} = r_N$$

where r_N depends only on the first N moments of P and is the same constant as that which appears in (5). Hence for $P \in \mathscr{M}_N(\mathbf{c})$, P^+ and $P^- \in \mathscr{M}_N(\mathbf{c})$ and ν^+_{N+1}, ν^-_{N+1} and r_N depend on $\mathbf{c}$ only. Further, if $\mathbf{c}$ is an interior point of the set M_N in (9),

$$(15) \qquad r_N > 0 .$$

Also (Remark 6 in [2]) the distributions of Θ induced by P^+ and P^- are uniquely determined, so that the $(N+2)$nd moments of P^+ and P^- have unique values ν^{++}_{N+2} and ν^{--}_{N+2} say, which depend only on the first N moments of P, so that for $P \in \mathscr{M}_N(\mathbf{c})$, ν^{++}_{N+2} and ν^{--}_{N+2} depend on $\mathbf{c}$ only. Let

$$(16) \qquad \nu^{++}_{N+2} - \nu^{--}_{N+2} = \lambda_N .$$

We next obtain a lower bound for $R_N(P^-)$ by using (13). By the second equality in formula (23) in [2],

$$\xi_{N+1,j+1}(P^+) - \xi_{N+1,j+1}(P^-) = (-1)^{N+j} r_N .$$

Hence by (15), if $(N - j)$ is an even number:

$$\xi_{N+1,j+1}(P^-) < \xi_{N+1,j+1}(P^+) ; \tag{17}$$

also since P^- and $P^+ \in \mathscr{M}_N(\mathbf{c})$, using the expansion in terms of moments in (8), we have

$$\xi_{N,j}(P^-) = \xi_{N,j}(P^+) ; \tag{18}$$

next by an application of Schwarz' inequality and using (10), we have

$$\frac{\xi^2_{N+1,j+1}(P)}{\xi_{N,j}(P)} \leqq \xi_{N+2,j+2}(P) , \qquad \text{for all } P \in \mathscr{M}_N(\mathbf{c}) , \tag{19}$$

and in particular for $P = P^+$; lastly by (8), (14) and (16)

$$\xi_{N+2,j+2}(P^+) = \xi_{N+2,j+2}(P^-) - (N - j) r_N + \lambda_N . \tag{20}$$

Combining equations (17) to (20) we obtain that for even values of $(N - j)$

$$\xi^2_{N+1,j+1}(P^-)/\xi_{N,j}(P^-) \leqq \xi_{N+2,j+2}(P^-) - (N - j) r_N + \lambda_N . \tag{21}$$

We now obtain an upper bound for λ_N by using a device of 'conjugate' probability distributions. For any df $F(\theta)$ of Θ we define its conjugate distribution to be $F'(\theta)$ where

$$F'(\theta) = F(1 - \theta) . \tag{22}$$

If $P \in \mathscr{P}$ induces the df F and $P' \in \mathscr{P}$ the df F' in (22), then we say that P' is conjugate to P. From (22) it follows that for any function $\psi(\Theta)$

$$E_P \psi(\Theta) = E_{P'} \psi(1 - \Theta) . \tag{23}$$

Next consider the relationship between the Bayes risks of P and P'. We have from (1) and (23)

$$\begin{aligned} R_N(t, P) &= E_P[t(X) - \Theta]^2 = E_{P'}[t(X) - (1 - \Theta)]^2 \\ &= E_{P'}[-t'(X) + \Theta]^2 = R_N(t', P') \end{aligned} \tag{24}$$

where we have put $t'(X) = 1 - t(X)$.

From (24) and (2), it follows that

$$R_N(P) = R_N(P') \qquad \text{for all } P \in \mathscr{P}. \tag{25}$$

Next consider the variation of P' as P varies on $\mathscr{M}_N(\mathbf{c})$. Let c_i', $i = 1, 2, \cdots, N$ be the first N moments of P'. By (23),

$$c_i' = E_{P'} \Theta^i = E_P(1 - \theta)^i = \textstyle\sum_{r=0}^{i} (-1)^r \binom{i}{r} c_r , \tag{26}$$

where c_0 is taken to be unity.

Let $\mathbf{c}' = (c_1', c_2', \cdots, c_N')$. As P varies over $\mathscr{M}_N(\mathbf{c})$, P' varies on $\mathscr{M}_N(\mathbf{c}')$. If $\mathbf{c}$ is an interior point of the set M_N in (9), there exists more than one induced distribution of Θ for $P \in \mathscr{M}_N(\mathbf{c})$ (Remark (6) in [2]). Consequently by (22), there exists more than one induced distribution of Θ for $P' \in \mathscr{M}_N(\mathbf{c}')$; hence $\mathbf{c}'$ is also an interior point of $M_N(\mathbf{c}')$. Let

$$\nu'^{+}_{N+1}, \quad \nu'^{-}_{N+1}, \quad P'^{+}, \quad P'^{-}, \quad \nu'^{++}_{N+2}, \quad \nu'^{--}_{N+2}, \quad r_N', \quad \lambda_N'$$

have the same meanings for $P' \in \mathscr{M}_N(\mathbf{c}')$ as corresponding undashed symbols have for $P \in \mathscr{M}_N(\mathbf{c})$. For any $P \in \mathscr{M}_{N+1}(\mathbf{c})$ let the $(N+1)$st moments of P and P' be ν_{N+1}, ν'_{N+1} respectively. By (23),

$$\nu'_{N+1} = E_{P'}\Theta^{N+1} = E_P(1-\Theta)^{N+1} = \sum_{r=0}^{N}(-1)^r\binom{N+1}{r}c_r + (-1)^{N+1}\nu_{N+1}\,. \tag{27}$$

Hence if N is an odd number ν'_{N+1} is maximized (minimized) when ν_{N+1} is maximized (minimized). (This result is contained in Theorem 4 in [3].) Thus P'^{+}, P'^{-} are the conjugate probability distributions to P^{+} and P^{-} respectively. Similarly if N is even, P'^{+}, P'^{-} are respectively conjugate to P^{-} and P^{+}. Further, by (27) and (14),

$$r_N' = \nu'^{+}_{N+1} - \nu'^{-}_{N+1} = \nu^{+}_{N+1} - \nu^{-}_{N+1} = r_N\,. \tag{28}$$

(Note: (28) is an immediate consequence of Theorem 4 and formula (1.2) [3].)

Next consider λ_N'. Using the remarks below (27) we obtain that if N is an odd number

$$\begin{aligned}\lambda_N' &= E_{P'^{+}}\Theta^{N+2} - E_{P'^{-}}\Theta^{N+2}\\ &= E_{P^{+}}(1-\Theta)^{N+2} - E_{P^{-}}(1-\Theta)^{N+2}\\ &= (N+2)r_N - \lambda_N\,.\end{aligned} \tag{29}$$

If N is an even number then in the second step in (29), P^{+} and P^{-} are interchanged, so that the same final result holds. Hence for all N

$$\lambda_N + \lambda_N' = (N+2)r_N\,. \tag{30}$$

Thus one out of the following alternatives must hold:

$$\lambda_N \leqq \frac{N+2}{2}r_N\,, \tag{31-i}$$

$$\lambda_N' = \frac{N+2}{2}r_N\,. \tag{31-ii}$$

Suppose (31-i) holds. Substituting for λ_N in (21) by (31-i) and combining the resulting inequality which holds for even values of $(N-j)$ with (19) which holds for all j, we obtain for $j = 0, 1, \cdots, N$,

$$\xi^2_{N+1,j+1}(P^-)/\xi_{N,j}(P^-) \leqq \xi_{N+2,j+2}(P^-) - \tfrac{1}{2}(N-2j-2)^{+}\zeta_{N-j}r_N \tag{32}$$

where $(x)^{+} = x$ or 0, according as $x \geqq$ or $\leqq 0$ and $\zeta_k = 1$ or 0, according as k is an even integer or not.

Note further that for $P \in \mathscr{M}_N(\mathbf{c})$,

$$\sum_{j=0}^{N} \binom{N}{j} \xi_{N+2,j+2}(P) = E_P\{\sum_{j=0}^{N} \binom{N}{j} \Theta^{j+2}(1-\Theta)^{N-j}\} \tag{33}$$
$$= E_P \Theta^2 = c_2 .$$

Put

$$K_N^- = \tfrac{1}{2} \sum{}^1 (N - 2j - 2)\binom{N}{j} 2^{-N} , \tag{34}$$

where $\sum^1$ denotes summation over values of j for which (i) $(N - j)$ is an even integer and (ii) $0 \leqq j < \frac{1}{2}(N-2)$. Multiplying both sides of (32) by $\binom{N}{j}$ and summing over $j = 0, 1, \cdots, N$, we obtain

$$\sum_{j=0}^{N} \binom{N}{j} \xi_{N+1,j+1}^2(P^-)/\xi_{N,j}(P^-) \leqq c_2 - 2^N r_N K_N^- .$$

Hence by (13)

$$R_N(P^-) \geqq 2^N K_N^- r_N . \tag{35}$$

By an argument exactly similar to that from (16) onwards, except that we consider odd values of $(N - j)$ for which (20) holds with P^+ and P^- interchanged we obtain that

$$R_N(P^+) \geqq 2^N K_N^+ r_N , \tag{36}$$

where $K_N^+ = \frac{1}{2} \sum^2 (N - 2j - 2)\binom{N}{j} 2^{-N}$ and $\sum^2$ denotes summation over values of j for which (i) $(N - j)$ is an odd integer and (ii) $0 \leqq j < \frac{1}{2}(N - 2)$.

It is shown by Skibinsky [2] that $R_N(P)$ attains its minimum either for $P = P^+$ or for $P = P^-$. Hence putting

$$K_N = \min(K_N^-, K_N^+) , \tag{37}$$

we obtain from (35) and (36) that

$$R_N(P) \geqq 2^N r_N K_N \qquad \text{for } P \in \mathscr{M}_N(\mathbf{c}) . \tag{38}$$

We have derived (38) by assuming that the alternative (31-i) holds. If (31-i) does not hold then (31-ii) holds. Hence the whole of the argument from (32) to (38) applies to the conjugate probability distributions P', and using (28) we obtain that

$$R_N(P') \geqq 2^N r_N K_N . \tag{39}$$

But since $R_N(P') = R_N(P)$ by (25), it follows from (39) that (38) holds always, whether (31-i) holds or (31-ii).

We have thus obtained a lower bound for the Bayes risk. We next consider the behaviour of K_N^-, K_N^+ for large N.

3. Limits for K_N^-, K_N^+. We shall next show that

$$\lim_{N\to\infty} (N^{-\frac{1}{2}} K_N) = \tfrac{1}{4}(2\pi)^{-\frac{1}{2}} . \tag{40}$$

The limit is proved by using a standard result regarding the approximation of the binomial probabilities $\binom{N}{j} p^j q^{N-j}$ by the Normal probabilities for j such that $j^3 N^{-\frac{1}{2}} \to 0$ as $N \to \infty$ (cf. Feller [1], Theorem 1, page 170). In applying the theorem we put throughout $p = q = \frac{1}{2}$. We have from (34)

$$N^{-\frac{1}{2}}K_N^- = \tfrac{1}{2}N^{-\frac{1}{2}} \textstyle\sum^1 (N - 2j)\binom{N}{j}2^{-N} - \delta_N \tag{41}$$

where

$$0 < \delta_N \to 0 \quad \text{as} \quad N \to \infty\,. \tag{42}$$

Next as in Feller [1], put

$$h = 2N^{-\frac{1}{2}}\,, \tag{43-i}$$

$$x_j = \left(\frac{N}{2} - j\right) h = (N - 2j)N^{-\frac{1}{2}}\,. \tag{43-ii}$$

(Note that the definition of x_j in (43-ii) differs only by a change of sign from that in [1].) Consider values of x_j such that $x_j \leqq N^{\frac{1}{8}}$, so that $x_j^3N^{-\frac{1}{2}} \to 0$ as $N \to \infty$. Let ε be any given positive number which may be arbitrarily small. By the result (2.11) in [1], page 170, there exists $N_0(\varepsilon)$ such that for $N \geqq N_0(\varepsilon)$,

$$\left|\binom{N}{j}2^{-N} - 2N^{-\frac{1}{2}}\phi(x_j)\right| < 2\varepsilon N^{-\frac{1}{2}}\phi(x_j) \tag{44}$$

where $\phi(x) = (2\pi)^{-\frac{1}{2}} \exp(-x^2/2)$.

For convenience of notation put

$$C_N = N^{\frac{1}{8}}\,. \tag{45}$$

It follows from (44) and (45) that

$$\begin{aligned} &\left|\textstyle\sum^1_{x_j \leqq C_N} x_j\binom{N}{j}2^{-N} - 2N^{-\frac{1}{2}} \sum^1_{x_j \leqq C_N} x_j\phi(x_j)\right| \\ &\qquad \leqq \textstyle\sum^1_{x_j \leqq C_N} \left|x_j\binom{N}{j}2^{-N} - 2N^{-\frac{1}{2}}x_j\phi(x_j)\right| \\ &\qquad \leqq 2\varepsilon N^{-\frac{1}{2}} \textstyle\sum^1_{x_j \leqq C_N} x_j\phi(x_j)\,. \end{aligned} \tag{46}$$

Since in the summation $\sum^1 j$ assumes either even or odd values, the successive values of x_j in (43-ii) have a common difference of $4N^{-\frac{1}{2}}$. Hence $4N^{-\frac{1}{2}} \sum^1_{x_j \leqq C_N} x_j\phi(x_j)$ is a Riemann sum for the integral $\int_0^{C_N} x\phi(x)\,dx$. Hence as $N \to \infty$ in (46)

$$2N^{-\frac{1}{2}} \textstyle\sum^1_{x_j \leqq C_N} x_j\phi(x_j) \to \frac{1}{2} \int_0^\infty x\phi(x)\,dx = \frac{1}{2}(2\pi)^{-\frac{1}{2}}\,. \tag{47}$$

Thus the right-hand side of (46) can be made arbitrarily small by taking N sufficiently large. Hence the two terms in the left-hand side of (46) are asymptotically equal. It therefore follows from (47) that

$$\lim_{N\to\infty} \tfrac{1}{2}N^{-\frac{1}{2}} \textstyle\sum^1_{x_j \leqq C_N} (N - 2j)\binom{N}{j}2^{-N} = \frac{1}{4}(2\pi)^{-\frac{1}{2}}\,. \tag{48}$$

For $x_j > C_N$, we have from the theorem on page 178 in [1] (relation 5.2), using (45)

$$\textstyle\sum_{x_j > C_N} \binom{N}{j}2^{-N} \sim (2\pi)^{-\frac{1}{2}}N^{-\frac{1}{8}} \exp\left(-\frac{1}{2}N^{\frac{1}{4}}\right) \tag{49}$$

where $\sim$ denotes asymptotic equality. Since $N - 2j \leqq N$ it follows from (49) that for sufficiently large N

$$\textstyle\sum^1_{x_j > C_N} (N - 2j)\binom{N}{j}2^{-N} < 2(2\pi)^{-\frac{1}{2}}N^{\frac{7}{8}} \exp\left(-\frac{1}{2}N^{\frac{1}{4}}\right) \tag{50}$$

which $\to 0$ as $N \to \infty$. Combining (42), (48) and (50), we obtain (40).

A similar argument holds good for K_N^+ and hence $(N^{-\frac{1}{2}} K_N^+)$ and consequently by (37) $(N^{-\frac{1}{2}}K_N)$ have the same limiting value, i.e.

$$\lim_{N\to\infty} (N^{-\frac{1}{2}}K_N) = \tfrac{1}{4}(2\pi)^{-\frac{1}{2}} . \tag{51}$$

Let k be any arbitrary positive and fixed number less than unity, i.e. $0 < k < 1$. By (51), there exists N_k such that

$$K_N \geqq \frac{k}{4} (2\pi)^{-\frac{1}{2}} N^{\frac{1}{2}} \qquad \text{for } N \geqq N_K . \tag{52}$$

Substituting for K_N in (38) by (52) and then combining with (5) we obtain

$$W_N(P)/R_N(P) \leqq \frac{1}{k} 4(2\pi)^{\frac{1}{2}} N^{-\frac{1}{2}} \qquad \text{for } N \geqq N_k . \tag{53}$$

(53) is proved on the Assumption 2.1 that $\mathbf{c}$ is an interior point of the set M_N in (9). As observed in the remark following (9), if $\mathbf{c}$ is not an interior point of M_N then $W_N(P) = 0$.

Thus (53) holds for all $P \in \mathscr{P}$, excluding P for which $W_N(P)$ and $R_N(P)$ both vanish. The result in (6) follows from (53). This completes its proof.

4. Boundedness of $W_N(P)/R_N(P)$. It is observed in [2] that for $N = 1$, and $N = 2$, the ratio $W_N(P)/R_N(P)$ is unbounded on $\mathscr{P}$. We consider for what values of N the ratio is bounded. From (37) and (38) it follows that $R_N(P)$ is bounded away from zero if K_N^- and K_N^+ are both positive. One of these involves summation over even j, and the other over odd j. Hence in order that each summation should contain at least one term, the values $j = 0$ and $j = 1$ must at least be admissible. Since the summation is over j such that $j < N/2 - 1$, it follows that we must have $1 < N/2 - 1$, so that $N \geqq 5$. Hence for all $N \geqq 5$, the ratio $W_N(P)/R_N(P)$ is bounded on $\mathscr{P}$. Thus, for example, for $N = 5$, it is easily seen from (34) and (36) that $K_5^- = \frac{5}{2}$, $K_5^+ = \frac{3}{2}2^{-5}$, so that by (38) and (5) the upper bound for the ratio is $64/3$. Similarly for $N = 6$, the upper bound is 32. Thus the question of the boundedness of the ratio remains unsettled only for the two values $N = 3$ and 4.

REFERENCES

[1] FELLER, W. (1950). *An Introduction to Probability Theory and its Applications* **1** (2nd ed.). Wiley, New York.

[2] SKIBINSKY, M. (1968). Minimax estimation of a random probability whose first N moments are known. *Ann. Math. Statist.* **39** 492–501.

[3] SKIBINSKY, M. (1969). Some striking properties of binomial and beta moments. *Ann. Math. Statist.* **40** 1753–1764.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720