COMPARISONS OF STOP RULE AND SUPREMUM EXPECTATIONS OF I.I.D. RANDOM VARIABLES

By T. P. HILL¹ AND ROBERT P. KERTZ

Georgia Institute of Technology

Implicitly defined (and easily approximated) universal constants $1.1 < a_n < 1.6$, $n = 2, 3, \cdots$, are found so that if X_1, X_2, \cdots are i.i.d. non-negative random variables and if T_n is the set of stop rules for X_1, \cdots, X_n , then $E\left(\max\{X_1, \cdots, X_n\}\right) \le a_n \sup\{EX_t: t \in T_n\}$, and the bound a_n is best possible. Similar universal constants $0 < b_n < \frac{1}{4}$ are found so that if the $\{X_i\}$ are i.i.d. random variables taking values only in [a, b], then $E\left(\max\{X_1, \cdots, X_n\}\right) \le \sup\{EX_t: t \in T_n\} + b_n(b-a)$, where again the bound b_n is best possible. In both situations, extremal distributions for which equality is attained (or nearly attained) are given in implicit form.

1. Introduction. Let T_n denote the set of stop rules for random variables X_1, \dots, X_n . If the $\{X_i\}$ are independent and non-negative, then it has been shown [4] that

$$(1) E(\max\{X_1, \dots, X_n\}) \le 2\sup\{EX_t : t \in T_n\}$$

and that 2 is the best possible bound, and [2] that in fact strict inequality holds in all but trivial cases. If the $\{X_i\}$ are independent and take values only in [a, b], then

(2)
$$E(\max\{X_1, \dots, X_n\}) \leq \sup\{EX_t : t \in T_n\} + (\frac{1}{4})(b - a),$$

and ¼ is the best possible bound [3]. Probabilistic interpretations have been given for these results: (1) says that the optimal return of a gambler (player using non-anticipating stop rules) is at least half that of the expected return of a prophet (player with complete foresight) playing the same game; and (2) says that a side payment of ¼ the game limits, paid by the prophet to the gambler, makes the game at least favorable for the gambler.

If the random variables in question are not only independent, but also identically distributed, then it turns out that the gambler's situation improves, and the constants "2" and "¼" in (1) and (2) respectively can be improved (lowered). The purpose of this paper is to determine these improvements. Probabilistically, the main results give the minimal odds and side payments, respectively, needed to achieve fairness for a gambler matched against a prophet playing the same game (in which the random variables are independent and identically distributed (i.i.d.)).

Implicitly defined (and easily approximated) universal constants $1.1 < a_n < 1.6$, n = 2, 3, ..., are found (e.g., $a_2 \cong 1.171$, $a_{100} \cong 1.337$, $a_{10,000} \cong 1.341$) satisfying the first main result, Theorem A.

THEOREM A. If n > 1 and X_1, X_2, \dots, X_n are i.i.d. non-negative random variables, then $E(\max\{X_1, \dots, X_n\}) \leq a_n \sup\{EX_t : t \in T_n\}$. Moreover, a_n is the best possible bound and is not attained except in the trivial cases where X_1 is almost surely 0 or has infinite expectation.

Similar universal constants $0 < b_n < \frac{1}{4}$ are found (e.g., $b_2 = .0625$, $b_{100} \cong .110$, $b_{10,000} \cong .111$) satisfying the second main result, Theorem B.

THEOREM B. If X_1, \dots, X_n are i.i.d. random variables taking values only in [a, b], then $E(\max\{X_1, \dots, X_n\}) \leq \sup\{EX_t: t \in T_n\} + b_n(b-a)$, (equivalently, $E(\min\{X_1, \dots, X_n\})$)

Received October 1980; revised March 1981.

¹ Partially supported by AFOSR Grant F49620-79-C-0123.

AMS 1980 subject classifications. Primary 60G40; secondary 62L15, 90C99.

Key words and phrases. Optimal stopping, extremal distributions, inequalities for stochastic processes.

 $\dots, X_n\}$ $\geq \inf\{EX_t: t \in T_n\} - b_n(b-a)$ and b_n is the best possible bound and is attained.

In Proposition 4.4 actual distributions are given implicitly (but again, in easily approximated form) for which equality in Theorem A nearly holds; Proposition 5.3 likewise gives extremal distributions for which equality in Theorem B holds.

2. Preliminaries. For random variables X and Y, $X \vee Y$ denotes the maximum of X and Y, $X^+ = X \vee 0$, and EX denotes the expectation of X. For $n = 1, 2, \dots, E_n(X) = E(X_1 \vee \dots \vee X_n)$, and $V_n(X) = \sup\{EX_t : t \in T_n\}$, where X_1, \dots, X_n are i.i.d. random variables each with distribution that of X. Throughout the remainder of this paper, all random variables will be assumed to have finite expectation.

The first lemma, a special case of [1, page 50], is included for ease of reference.

LEMMA 2.1. (i) $V_n(X) = E(X \vee V_{n-1}(X))$ for all n > 1; and (ii) if $t^* \in T_n$ is the stop rule defined for $j = 0, \dots, n-1$ by $t^* = j \Leftrightarrow \{t^* > j-1 \text{ and } X_j \geq V_{n-j}(X)\}$, then $EX_{t^*} = V_n(X)$.

Lemmas 2.4 and 2.5 are probabilistic results which will be used in the proofs of Theorems A and B to restrict attention to simple random variables of special form. In setting up this reduction, a definition and a special case of a result (Lemma 2.2) from [3] are useful.

DEFINITION 2.2. For random variable Y and constants $0 \le a < b < \infty$, let Y_a^b denote a random variable with $Y_a^b = Y$ if $Y \notin [a, b]$, = a with probability $(b - a)^{-1} \int_{Y \in [a, b]} (b - Y)$, and = b otherwise.

LEMMA 2.3. Let Y be any random variable and $0 \le a < b < \infty$. Then $EY = EY_a^b$, and if X is any random variable independent of both Y and Y_a^b , then $E(X \lor Y) \le E(X \lor Y_a^b)$.

It may be seen that Y_a^b is the distribution with maximum variance which both coincides with Y off [a, b] and has expectation EY.

LEMMA 2.4. Let n > 1 and X be any random variable taking values in [0, 1]. Then there exists a simple random variable Y, taking on only the values $0, V_1(X), V_2(X), \dots, V_{n-1}(X)$, and 1, and satisfying both $V_j(Y) = V_j(X)$ for $j = 1, 2, \dots, n$, and $E_n(Y) \ge E_n(X)$.

PROOF. If X is constant, the conclusion is trivial with Y = X. Otherwise, $P(X < V_1(X)) > 0$, $P(X > V_{n-1}(X)) > 0$, and $0 < V_1(X) < \cdots < V_{n-1}(X) < 1$.

Let $X_1 = X_0^{V_1(X)}$ be as in Definition 2.2, and independent of X. By Lemma 2.3, $V_1(X_1) = EX_1 = EX = V(X)$, and thus by Lemma 2.1 and induction, $V_j(X_1) = E(X_1 \vee V_{j-1}(X_1)) = E(X \vee V_{j-1}(X)) = V_j(X)$ for $j = 1, \dots, n$. By Lemma 2.3, $E_2(X) \leq E(X \vee X_1) \leq E_2(X_1)$, and similarly $E_n(X) \leq E_n(X_1)$.

Next define X_2, \dots, X_{n-1} inductively by $X_k = (X_{k-1}) \frac{V_k(X)}{V_{k-1}(X)}$ and X_n by $X_n = (X_{n-1})^1_{V_{n-1}(X)}$, and conclude as for X_1 that $V_j(X_k) = V_j(X_{k-1}) = \dots = V_j(X_1) = V_j(X)$ for $1 \le j, k \le n$, and that $E_n(X) \le E_n(X_1) \le \dots \le E_n(X_n)$. Letting $Y = X_n$ completes the proof.

Lemma 2.5. For n>1, let X be a simple random variable taking values $0< V_1(X)<\cdots< V_{n-1}(X)<1$ with probabilities p_0,p_1,\cdots,p_n respectively, and let $s_j=p_0+\cdots+p_j$ and $s_{-1}=1$. Then: (i) $V_j(X)=V_1(X)$ $[1+s_0+s_0s_1+\cdots+s_0s_1\cdots s_{j-2}], j=2,\cdots,n$; (ii) $V_1(X)=(1-s_{n-1})/[(1-s_{n-1})(1+s_0+s_0s_1+\cdots+s_0s_1\cdots s_{n-3})+s_0s_1\cdots s_{n-2}],$ and (iii) $E_n(X)=V_1(X)[(1+s_0+s_0s_1+\cdots+s_0s_1\cdots s_{n-3})+s_0s_1\cdots s_{n-2}(1+s_{n-1}+\cdots+s_{n-1}^{n-1})-(s_0^n+s_0s_1^n+\cdots+s_0\cdots s_{n-3}s_{n-2}^n)].$

PROOF. For (i), observe that by Lemma 2.1, $V_j(X) = V_j(X)p_j + V_{j+1}(X)p_{j+1} + \cdots + V_{n-1}(X)p_{n-1} + 1 \cdot p_n + s_{j-1}V_{j-1}(X)$. Since $V_j(X)p_j + \cdots + V_{n-1}(X)p_{n-1} + 1 \cdot p_n = V_1(X) - [V_1(X)p_1 + \cdots + V_{j-1}(X)p_{j-1}]$, the desired conclusion follows easily by induction on j.

Conclusion (ii) follows since $V_1(X) = V_1(X)(s_1 - s_0) + \cdots + V_{n-1}(X)(s_{n-1} - s_{n-2}) + (1 - s_{n-1})$ by solving the equations in (i) for V_1 in terms of s_0, s_1, \dots, s_{n-1} .

For (iii), note that $E_n(X) = \sum_{j=1}^{n-1} V_j(X)(s_j^n - s_{j-1}^n) + (s_n^n - s_{n-1}^n)$, and apply (i) and (ii).

For the proof of Theorem A the following complements to Definition 2.2. and Lemmas 2.3 and 2.4 are given.

DEFINITION 2.6. For random variable Y and constants $\alpha > a \ge 0$ satisfying $\alpha \cdot P(Y \ge \alpha) \ge \int_{Y \ge a} Y$, let $Y_{a,\alpha}$ denote a random variable with $Y_{a,\alpha} = Y$ if $Y \notin [a, \infty]$, = a with probability $(\alpha - a)^{-1} \int_{Y \ge a} (\alpha - Y)$, and $= \alpha$ otherwise.

LEMMA 2.7 Let Y be any integrable random variable and $0 \le a < \infty$. Then $EY = EY_{a,\alpha}$, and for all α sufficiently large, if X is any random variable independent of both Y and $Y_{a,\alpha}$, then $E(X \lor Y) \le E(X \lor Y_{a,\alpha})$. This last inequality is strict if and only if $P(X > a) \cdot P(Y > a) > 0$.

PROOF. That $EY = EY_{a,\alpha}$ is immediate. For the remainder assume $P(Y \ge a) > 0$ and fix any X independent of both Y and $\{Y_{a,\alpha}\}$. From the definition of $Y_{a,\alpha}$, the convexity of the function $\psi(y) = E(X \lor y)$, and the independence of X and Y, it follows that $E(X \lor Y_{a,\alpha})$ is a non-decreasing function of α and $\lim_{\alpha \to \infty} E(X \lor Y_{a,\alpha}) = \int_{Y < a} X \lor Y + E(X \lor a)P(Y \ge a) + E(Y - a)^+$, with the limit being attained if $P(X > a) \cdot P(Y > a) = 0$. The conclusion follows from these results and the dichotomy that $\int_{Y < a} X \lor Y + E(X \lor a)P(Y \ge a) + E(Y - a)^+ > E(X \lor Y)$ if $P(X > a) \cdot P(Y > a) > 0$, and $= E(X \lor Y)$ if $P(X > a) \cdot P(Y > a) = 0$. The strict inequality in this dichotomy follows since for $P(X > a) \cdot P(Y > a) > 0$,

$$\int_{X \ge a, Y \ge a} (X - a + Y - a) - \int_{X \ge a, Y \ge a} [(X - a) \lor (Y - a)]$$

$$= \int_{X \ge a, Y \ge a} [(X - a) \land (Y - a)] > 0. \quad \Box$$

If P(Y > a) > 0, then $\{Y_{a,\alpha}\}$ are random variables which coincide with Y off $[a, \infty)$, have expectation EY, and have variances which increase to infinity.

LEMMA 2.8. Let n > 1 and X be any non-negative unbounded (ess sup $X = +\infty$) random variable. Then there exists a non-negative bounded random variable Y satisfying both $V_j(Y) = V_j(X)$ for $j = 1, 2, \dots, n$, and $E_n(Y) > E_n(X)$.

PROOF. Define Y through Definition 2.6 by $Y = X_{V_{n-1}(X),\alpha}$. Then the conclusion follows from Lemmas 2.1 and 2.7 for α sufficiently large. \square

3. Definition of the constants $\{a_n\}$ and $\{b_n\}$. The purpose of this section, which is purely analytical (non-probabilistic) in nature, is to define the constants $\{a_n\}$ and $\{b_n\}$ appearing in Theorems A and B, respectively, and to concurrently develop results useful in the proofs of these theorems.

DEFINITION 3.1. For n > 1 and $w, x \in [0, \infty)$, let $\phi_n(w, x) = (n/(n-1))w^{(n-1)/n} + x/(n-1)$. For $\alpha \in [0, \infty)$, define inductively the functions $\eta_{j,n}, j = 0, 1, \dots, n$, by $\eta_{0,n}(\alpha) = \phi_n(0, \alpha)$, and $\eta_{j,n}(\alpha) = \phi_n(\eta_{j-1,n}(\alpha), \alpha)$.

LEMMA 3.2. $\eta_{j,n}$ is continuous, non-negative, strictly increasing and concave for all n > 1, and all $j = 0, 1, \dots, n$.

PROOF. Fix n > 1; proof will be by induction on j. First observe that $\eta_{0,n}$ is continuous, and for $\alpha > 0$, $\eta_{0,n}(\alpha) > 0$, $\eta'_{0,n}(\alpha) > 0$, and $\eta''_{0,n}(\alpha) = 0$ (where ()' denotes differentiation with respect to α). Assume $\eta_{j-1,n}$ is continuous and, for $\alpha > 0$, that $\eta_{j-1,n}(\alpha) > 0$, $\eta'_{j-1,n}(\alpha) > 0$, and $\eta''_{j-1,n}(\alpha) \le 0$. Then it is clear that $\eta_{j,n}$ is continuous, and for $\alpha > 0$,

$$\eta_{j,n}(\alpha) > 0, \qquad \eta'_{j,n}(\alpha) > 0, \qquad \text{and}$$

$$\eta''_{j,n}(\alpha) = [\eta_{j-1,n}(\alpha)]^{-1/n} [(-n^{-1})(\eta_{j-1,n}(\alpha))^{-1}(\eta'_{j-1,n}(\alpha))^2 + \eta''_{j-1,n}(\alpha)] \le 0. \square$$

DEFINITION 3.3. Let $G_n:[0,\infty)\to\mathbb{R}$ be the function $G_n(\alpha)=\eta_{n-1,n}(\alpha)$.

PROPOSITION 3.4. (a) For all $\alpha \in [0, 1]$, $G_n(\alpha) \leq \alpha [(n/(n-1))^n - 1] + [1 - ((n-1)/n)^{n-1}]$; and (b) there is a unique number $\alpha_n > 0$ for which $G_n(\alpha_n) = 1$. Moreover, $\alpha_n < 1$, and for $\alpha \in [0, \alpha_n]$, $\alpha [(n/(n-1))^n - 1] \leq G_n(\alpha)$.

PROOF. Let $\psi_n(w, x) = (n/(n-1))w + x/(n-1)$ for $w, x \in [0, \infty)$. For (a), define inductively the functions $\sigma_{j,n}(\alpha)$, $0 \le j \le n-1$, $\alpha \in [0, 1]$, by $\sigma_{0,n}(\alpha) = \psi_n(0, \alpha)$ and $\sigma_{j,n}(\alpha) = \psi_n(\sigma_{j-1,n}(\alpha) + c_n, \alpha)$, where $c_n = n^{-1}((n-1)/n)^{n-1}$. It will be shown that $\eta_{j,n}(\alpha) \le \sigma_{j,n}(\alpha)$ for all $\alpha \in [0, 1]$. First observe that $\eta_{0,n}(\alpha) = \alpha/(n-1)$, and assume $\eta_{j-1,n}(\alpha) \le \sigma_{j-1,n}(\alpha)$. Since $x^{(n-1)/n} \le x + c_n$, it follows that

$$\eta_{j,n}(\alpha) = (n/(n-1))(\eta_{j-1,n}(\alpha))^{(n-1)/n} + \alpha/(n-1)
\leq (n/(n-1)) (\sigma_{j-1,n}(\alpha))^{(n-1)/n} + \alpha/(n-1)
\leq (n/(n-1))(\sigma_{j-1,n}(\alpha) + c_n) + \alpha/(n-1) = \sigma_{j,n}(\alpha).$$

For j = n - 1, this yields

$$\eta_{n-1,n}(\alpha) = G_n(\alpha) \le \sigma_{n-1,n}(\alpha) = \alpha [(n/(n-1))^n - 1] + [1 - ((n-1)/n)^{n-1}],$$

completing the proof of (a).

For (b), define inductively the functions $\mu_{j,n}(\alpha)$, $0 \le j \le n-1$, $\alpha \in [0, 1]$, by $\mu_{0,n}(\alpha) = \psi_n(0, \alpha)$ and $\mu_{j,n}(\alpha) = \psi(\mu_{j-1,n}(\alpha), \alpha)$. It shall first be shown that

(3)
$$\mu_{j,n}(\alpha) \le \eta_{j,n}(\alpha)$$
, for $0 \le j \le n-1$ and all $\alpha \in [0,1]$ with $G_n(\alpha) \le 1$.

Given $\alpha \in [0, 1]$ with $G_n(\alpha) \le 1$, observe that $\mu_{0,n}(\alpha) = \eta_{0,n}(\alpha) = \alpha/(n-1)$, and assume that $\mu_{j-1,n}(\alpha) \le \eta_{j-1,n}(\alpha)$. Since $0 \le \eta_{0,n}(\alpha) \le \eta_{1,n}(\alpha) \le \cdots \le \eta_{n-1,n}(\alpha) = G_n(\alpha) \le 1$ and $x \le x^{(n-1)/n}$ for $x \in [0, 1]$, it follows that

$$\mu_{j,n}(\alpha) = (n/(n-1))\mu_{j-1,n}(\alpha) + \alpha/(n-1) \le (n/(n-1))\eta_{j-1,n}(\alpha) + \alpha/(n-1)$$

$$\le (n/(n-1))(\eta_{j-1,n}(\alpha))^{(n-1)/n} + \alpha/(n-1) = \eta_{j,n}(\alpha),$$

completing the proof of (3).

If $G_n(\alpha) \leq 1$ for all $\alpha \in [0, 1]$, then it would follow from (3) that $G_n(1) = \eta_{n-1,n}(1) \geq \mu_{n-1,n}(1) = (n/(n-1))^n - 1 > e - 1 > 1$, a contradiction. Thus there exists $\alpha_n \in (0, 1)$ with $G_n(\alpha_n) = 1$; the uniqueness of α_n follows from the strict monotonicity of $\eta_{n-1,n}$ proved in Lemma 3.2. \square

Example 3.5. (a) For n=2, $\phi_2(w,x)=2\sqrt{w}+x$, $G_2(\alpha)=2\sqrt{\alpha}+\alpha$, and $\alpha_2=3-2\sqrt{2}\cong .171$.

(b) $\alpha_3 \cong 0.221$, $\alpha_4 \cong 0.248$, $\alpha_5 \cong 0.264$, $\alpha_{10} \cong 0.301$, $\alpha_{100} \cong 0.337$, and $\alpha_{10,000} \cong 0.341$.

Although the authors believe that the α_n 's are strictly monotone increasing with limit e^{-1} , they have established only the general quantitative information about them given in the following proposition.

PROPOSITION 3.6. For all n > 1, (a) $((n-1)/n)^{n-1}[(n/(n-1))^n - 1]^{-1} \le \alpha_n \le [(n/(n-1))^n - 1]^{-1}$; and (b) $(3e)^{-1} \le \alpha_n \le (e-1)^{-1}$.

PROOF. Part (a) follows from Proposition 3.4 with $\alpha = \alpha_n$. Part (b) follows from (a) since $(n/(n-1))^n \searrow e$, and $(n/(n-1))^{n-1} \nearrow e$ imply that $((n-1)/n)^{n-1}[(n/(n-1))^n - 1]^{-1} \ge e^{-1}(2^2-1)^{-1} = (3e)^{-1}$ and that $[(n/(n-1))^n - 1]^{-1} \nearrow (e-1)^{-1}$. \square

DEFINITION 3.7. Let $H_n:[0, 1] \to \mathbb{R}$ be the function $H_n(\beta) = (n-1) \cdot [\eta_{n,n}(\beta) - \eta_{n-1,n}(\beta)]$.

PROPOSITION 3.8. For each n > 1 there is a unique number $\beta_n \in [0, 1]$ such that $H_n(\beta_n) = 1$. Moreover, $0 < \beta_n < 1$.

PROOF. Let $f(x) = (n/(n-1))x^{(n-1)/n} - x$, let $g(\beta) = f(\eta_{n-1,n}(\beta))$, and let u be the linear function $u(\beta) = (1-\beta)/(n-1)$. Then $H_n(\beta) = 1$ if and only if

$$g(\beta) = u(\beta).$$

Let α_n be as in Proposition 3.4(b). By Lemma 3.2, $\eta_{n-1,n}$ is strictly increasing from 0 to 1 on $[0, \alpha_n]$. Since f is strictly increasing on [0, 1], it follows that g is strictly increasing on $[0, \alpha_n]$, and since g(0) = 0 and $g(\alpha_n) = 1/(n-1)$, it follows that (4) has a unique solution in $[0, \alpha_n]$. It remains only to show that (4) has no solution on $[\alpha_n, 1]$. This will be accomplished by exhibiting a function f which lies between g and g on g and g on g and g on points in common with g.

Let $k_n = (n/(n-1))^n - 1$, let $d_n = 1 - ((n-1)/n)^{n-1}$, and let $t(\beta) = f(k_n\beta + d_n)$ for $\beta \in [0, 1]$. Since f is decreasing on $[1, \infty]$, in order to show that $g(\beta) = f(\eta_{n-1, n}(\beta)) \ge f(k_n\beta + d_n) = t(\beta)$ on $[\alpha_n, 1]$, it suffices to show that

(5)
$$1 \le \eta_{n-1,n}(\beta) \le \beta k_n + d_n \quad \text{for} \quad \beta \in [\alpha_n, 1].$$

The first inequality in (5) follows since $\eta_{n-1,n}$ is strictly increasing (Lemma 3.2) and since $\eta_{n-1,n}(\alpha_n) = 1$; and the second by Proposition 3.4(a).

In order to show that t > u on $[\alpha_n, 1]$, it is enough to show that t > u on $[b_n, 1]$, where $b_n = k_n^{-1}(1-d_n)$, since $b_n \le \alpha_n$ by Proposition 3.6(a). Since $1 < e - e^{-1} n/(n-1) < k_n + d_n < (n/(n-1))^n$ and f is decreasing on $[1, \infty)$, it follows that $t(1) = f(k_n + d_n) > f((n/(n-1))^n) = 0 = u(1)$. But since $u(b_n) < (n-1)^{-1} = t(b_n)$, and t is concave, it then follows that t > u on $[b_n, 1]$, completing the proof. \square

EXAMPLE 3.9. (a) For
$$n=2$$
, $H_2(\beta)=2(2\sqrt{\beta}+\beta)^{1/2}-2\beta^{1/2}$, and $\beta_2=\mathcal{N}_{16}$. (b) $\beta_3\cong .077$, $\beta_4\cong .085$, $\beta_5\cong .090$, $\beta_{10}\cong .100$, $\beta_{100}\cong .110$, $\beta_{10,000}\cong .111$.

DEFINITION 3.10. For n > 1, let $a_n = 1 + \alpha_n$, and $b_n = \beta_n$.

4. Proof of Theorem A and its extremal distributions.

DEFINITION 4.1. For a random variable X with $V_n(X) > 0$, let $R_n(X) = E_n(X)/V_n(X)$, $n = 1, 2, \cdots$.

Probabilistically, $R_n(X)$ is the odds which must be given a gambler playing against a prophet (faced with the same n i.i.d. random variables each with distribution that of X) in order to make the game fair for the gambler. (In terms of R_n , Theorem A simply states that $R_n(X) < a_n$ for all distributions X, and that the bound a_n is the best possible.)

PROOF OF THEOREM A. Fix n > 1. The case where X has infinite expectation is trivial, so assume $EX < \infty$. First, it shall be shown that it suffices to consider random variables taking values in [0, 1] by proving that

(6) for any random variable X, there exists a random variable Y taking values in [0, 1] for which $R_n(X) \le R_n(Y)$.

For random variable X, from Lemma 2.8 there exists a bounded random variable Z such that $R_n(X) \leq R_n(Z)$. Define Y = Z/(supremum of Z); then Y is a random variable taking its values in [0, 1] and $R_n(X) \leq R_n(Z) = R_n(Y)$. This establishes (6).

By Lemma 2.4, attention may be further restricted to simple random variables X taking on the values $0, V_1(X), \dots, V_{n-1}(X)$, and 1 (with probabilities p_0, p_1, \dots, p_n respectively). Let $s_j = p_0 + \dots + p_j$ for $j = 0, \dots, n-1$ and let $s_{-1} = 1$. Now, if $s_{n-1} = 0$ or 1, then X is constant and $R_n(X) = 1$; if $0 < s_{n-1} < 1$, then $0 < V_1(X) < \dots < V_{n-1}(X) < 1$ and from Lemma 2.5 $R_n(X) = R_n(s_0, \dots, s_{n-1})$ where $R_n(s_0, \dots, s_{n-1})$ is the function defined for $s_j \ge 0, j = 0, \dots, n-1$, by

(7)
$$R_n(s_0, s_1, \dots, s_{n-1}) = 1 + \frac{(\sum_{j=1}^{n-1} s_{n-1}^j) s_0 s_1 \dots s_{n-2} - s_0^n - s_0 s_1^n - \dots - s_0 \dots s_{n-3} s_{n-2}^n}{1 + s_0 + s_0 s_1 + \dots + s_0 s_1 \dots s_{n-2}}.$$

The conclusion of Theorem A follows once it is shown that

(8) there exists a unique point $(\hat{s}_0, \dots, \hat{s}_{n-1})$ with $0 < \hat{s}_0 < \dots < \hat{s}_{n-2} < \hat{s}_{n-1} = 1$ for which $R_n(s_0, \dots, s_{n-1}) < R_n(\hat{s}_0, \dots, \hat{s}_{n-1}) = a_n$ for all (s_0, \dots, s_{n-1}) with $0 \le s_0 \le \dots \le s_{n-2} \le s_{n-1} < 1$,

where a_n was given in Definition 3.10.

For each (s_0, \dots, s_{n-1}) with $0 = s_j \le \dots \le s_{n-1} < 1$, $R_n(s_0, \dots, s_{n-1}) = 1$, and for each (s_0, \dots, s_{n-1}) with $0 < s_0 \le s_1 \le \dots \le s_{n-1} < 1$, $R_n(s_0, \dots, s_{n-1}) < R_n(s_0, \dots, s_{n-2}, 1)$. If the function $r_n(s_0, \dots, s_{n-2})$ is defined for $s_j \ge 0$, $j = 0, \dots, n-2$, by $r_n(s_0, \dots, s_{n-2}) = R_n(s_0, \dots, s_{n-2}, 1)$, then the proof of (8) follows from showing that

(9) there is a unique point $(\hat{s}_0, \dots, \hat{s}_{n-2})$ with $0 < \hat{s}_0 < \dots < \hat{s}_{n-2} < 1$ for which $\max\{r_n(s_0, \dots, s_{n-2}); 0 \le s_0 \le \dots \le s_{n-2} < 1\} = r_n(\hat{s}_0, \dots, \hat{s}_{n-2}) = a_n$.

First, verify that the following four statements are equivalent for (s_0, \dots, s_{n-2}) with $s_j > 0$ for $j = 0, \dots, n-2$:

(10a)
$$\frac{\partial r_n}{\partial s_i}(s_0, \dots, s_{n-2}) = 0 \quad \text{for} \quad j = 0, \dots, n-2;$$

- (10b) $ns_{j+1} \cdots s_{n-2} ns_{j}^{n-1} + \left[(1 s_{j+1}^{n}) + s_{j+1} (1 s_{j+2}^{n}) + \cdots + s_{j+1} \cdots s_{n-3} (1 s_{n-2}^{n}) \right] r_{n}(s_{0}, \cdots, s_{n-2})(1 + s_{j+1} + s_{j+1}s_{j+2} + \cdots + s_{j+1} \cdots s_{n-2}) = 0 \text{ for } 0 \le j \le n 4, ns_{n-2} ns_{n-3}^{n-1} + (1 s_{n-2}^{n}) r_{n}(s_{0}, \cdots, s_{n-2}) \cdot (1 + s_{n-2}) = 0, \text{ and } n ns_{n-2}^{n-1} r_{n}(s_{0}, \cdots, s_{n-2}) = 0;$
- (10c) $(n-1)s_{j+1}^n = ns_j^{n-1} + (n-1)s_0^n$ for $0 \le j \le n-3$, and $n-1 = ns_{n-2}^{n-1} + (n-1)s_0^n$, and at (s_0, \dots, s_{n-2}) satisfying these n-2 equations, $r_n(s_0, \dots, s_{n-2}) = 1 + (n-1)s_0^n$; and
- (10d) letting $\alpha = (n-1)s_0^n$, $\eta_{j,n}(\alpha) = s_j^n$ for $0 \le j \le n-2$, $1 = \eta_{n-1,n}(\alpha) = G_n(\alpha)$, and at (s_0, \dots, s_{n-2}) satisfying these equations, $r_n(s_0, \dots, s_{n-2}) = 1 + (n-1)s_0^n = 1 + \alpha$.

Let $B \subset \mathbb{R}^{n-1}$ be the region $B = \{(s_0, \cdots, s_{n-2}); s_j \geq 0 \text{ for } j = 0, \cdots, n-2\}$. By (10a-d) and Proposition 3.4 there is a unique point $(\hat{s}_0, \cdots, \hat{s}_{n-2})$ in the interior of B at which $\partial r_n/\partial s_j = 0$ for $j = 0, \cdots, n-2$, and at this point $r_n(\hat{s}_0, \cdots, \hat{s}_{n-2}) = 1 + (n-1)\hat{s}_0^n > 1$. Thus the maxima and minima for r_n in B, if they exist, occur at $(\hat{s}_0, \cdots, \hat{s}_{n-2})$ or on the boundary of B. However, if $s_j = 0$ for some $j = 0, \cdots, n-2$, or if $s_j \to \infty$ for some or all $j = 0, \cdots, n-2$, then $r_n(s_0, \cdots, s_{n-2}) \leq 1$. Thus the maximum for r_n in B is at $(\hat{s}_0, \cdots, \hat{s}_{n-2})$. Since $0 < \hat{s}_0 < \cdots < \hat{s}_{n-2} < 1$ from (10d), Definition 3.1, and Lemma 3.2, and since $\{(s_0, \cdots, s_{n-2}); 0 \leq s_0 \leq \cdots \leq s_{n-2} < 1\} \subset B$, it follows that (10d), Proposition 3.4, and Definition 3.10 imply that (9) holds.

That the bound a_n is sharp is clear from the above reasoning (see also Proposition 4.4.).

EXAMPLE 4.2. Let X_1, X_2, \cdots be non-negative i.i.d. random variables (with positive finite expectations). Calculations of $\{a_n\}$ indicate that $E(X_1 \vee X_2) < 1.172 \sup\{EX_t: t \in T_2\}$; $E(X_1 \vee \cdots \vee X_{100}) < 1.338 \sup\{EX_t: t \in T_{100}\}$; and $E(X_1 \vee \cdots \vee X_{10,000}) < 1.342 \sup\{EX_t: t \in T_{10,000}\}$.

COROLLARY 4.3. Let X_1, X_2, \cdots be i.i.d. non-negative random variables and let T denote the stop rules for X_1, X_2, \cdots . Then $E(\sup X_i) \leq (1 + (e - 1)^{-1}) \sup \{EX_i : t \in T\}$.

PROOF. Apply Proposition 3.6(b) to Theorem A.

It is perhaps of some interest to identify distributions for which equality in Theorem A is nearly attained. For this purpose the following parameters are collected here. Fix n > 1. Let $\alpha_n \in (0, 1)$ be the unique solution of $G_n(\alpha_n) = 1$ from Proposition 3.4. For $j = 0, \dots, n-2$, \hat{s}_j is given by $\hat{s}_j = (\eta_{j,n}(\alpha_n))^{1/n}$ and \hat{p}_j by $\hat{p}_0 = \hat{s}_0$, $\hat{p}_j = \hat{s}_j - \hat{s}_{j-1}$ for $j = 1, \dots, n-2$, and $\hat{p}_{n-1} = 1 - \hat{s}_{n-2}$.

PROPOSITION 4.4. For each n > 1 and $\varepsilon > 0$ there exists a simple random variable $\hat{X} = \hat{X}(n, \varepsilon)$ with $P(\hat{X} = 0) = \hat{p}_0$, $P(\hat{X} = V_j(\hat{X})) = \hat{p}_j$ for $j = 0, \dots, n-2$, $P(\hat{X} = V_{n-1}(\hat{X})) \in (\hat{p}_{n-1} - \varepsilon, \hat{p}_{n-1})$ and $P(\hat{X} = 1) < \varepsilon$ satisfying $R_n(\hat{X}) > a_n - \varepsilon$ and hence $R_n(\hat{X}) > R_n(X) - \varepsilon$ for every non-negative random variable X.

PROOF. For $\varepsilon > 0$ sufficiently small consider the random variables $X = X(n, \varepsilon)$ taking values $0 < V_1(X) < \cdots < V_{n-1}(X) < 1$ with probabilities $\hat{p}_0, \cdots, \hat{p}_{n-2}, \hat{p}_{n-1} - \varepsilon, \varepsilon$ respectively; the values $V_j(X), j = 0, \cdots, n-1$, can be computed from Lemma 2.5 (i, ii). From the proof of Theorem A it is clear that $R_n(X(n, \varepsilon)) \to a_n$ as $\varepsilon \to 0$. \square

Example 4.5. (a) For n=2, $(\hat{p}_0, \hat{p}_1)\cong (0.414, 0.586)$. Calculations indicate that the random variable \hat{X} taking values 0, 2.41421 × 10⁻⁵, and 1 with probabilities 0.41421, 0.58578, and 10^{-5} respectively satisfies $R_2(\hat{X}) > R_2(X) - 10^{-4}$ for every non-negative random variable X. (b) For n=10, $(\hat{p}_0, \dots, \hat{p}_9)\cong (0.711925, 0.070190, 0.047863, 0.037426, 0.030936, 0.026304, 0.022730, 0.019837, 0.017423, 0.015367). For <math>\varepsilon > 0$ small consider the random variables $\hat{X} = \hat{X}(n, \varepsilon)$ taking values $0, \varepsilon \cdot v_1, \dots, \varepsilon \cdot v_9$, and 1 with probabilities $\hat{p}_0, \dots, \hat{p}_8, \hat{p}_9 - \varepsilon$, and ε respectively, where $(v_1, \dots, v_9)\cong (3.32872, 5.69852, 7.55198, 9.09031, 10.4247, 11.6234, 12.7317, 13.7818, 14.7974). For <math>\varepsilon > 0$ sufficiently small $R_{10}(\hat{X}) > R_{10}(X) - 10^{-3}$ for every random variable X.

The assumption of non-negativity in Theorem A is essential, as the following example shows.

EXAMPLE 4.6. Let X be uniformly distributed on [0, 1] (so $E_n(X) > V_n(X)$ for all n > 1). For $\epsilon > 0$, let $Y_{\epsilon} = X - V_n(X) + \epsilon$. Then $E_n(Y_{\epsilon}) = E_n(X) - V_n(X) + \epsilon$, and $V_n(Y_{\epsilon}) = \epsilon$, so $1 + (E_n(X) - V_n(X))/\epsilon = R_n(Y_{\epsilon}) \nearrow \infty$ as $\epsilon \searrow 0$.

5. Proof of Theorem B and its extremal distributions.

DEFINITION 5.1. For a random variable X taking values in [0, 1], let $D_n(X) = E_n(X) - V_n(X)$, $n = 1, 2, \dots$

Probabilistically, $D_n(X)$ is twice the side payment which must be paid to a gambler playing against a prophet (faced with the same n independent random variables each with distribution that of X) in order to make the game fair for the gambler. In terms of D_n , the conclusion of Theorem B is that $D_n(X) \leq b_n$ for all n, and that the bound b_n is the best possible and is attained.

PROOF OF THEOREM B. Without loss of generality (add, or multiply by, suitable constants) a = 0 and b = 1. By Lemma 2.4, it may be assumed that X is a simple random

variable taking on the values 0, $V_1(X)$, ..., $V_{n-1}(X)$, and 1 with probabilities p_0, p_1, \dots, p_n respectively. Let $s_j = p_0 + p_1 + \dots + p_j$ for $0 \le j \le n + 1$, and let $s_{-1} = 1$. By Lemma 2.5, for $0 < s_{n-1} < 1$ (otherwise X is constant and $D_n(X) = 0$), $D_n(X) = D_n(s_0, s_1, \dots, s_{n-1})$ where $D_n(s_0, \dots, s_{n-1})$ is the continuous function defined on $\{(s_0, \dots, s_{n-1}); 0 \le s_0 \le \dots \le s_{n-1} \le 1\} \cup \{(s_0, \dots, s_{n-1}); 0 < s_{n-1} < 1 \text{ and } s_j > 0 \text{ for } j = 0, \dots, n-2\}$ by

(11)
$$D_n(s_0, s_1, \dots, s_{n-1}) = 0 \quad \text{if} \quad 0 = s_0 = \dots = s_i \le \dots \le s_{n-1} = 1,$$

and

$$=\frac{(1-s_{n-1})\{(\sum_{j=1}^{n-1}s_{n-1}^{j})s_0s_1\cdots s_{n-2}-s_0^n-s_0s_1^n-\cdots-s_0s_1\cdots s_{n-3}s_{n-2}^n\}}{(1-s_{n-1})(1+s_0+s_0s_1+\cdots+s_0s_1\cdots s_{n-3})+s_0s_1\cdots s_{n-2}}$$
 otherwise.

It remains only to show that

(12)
$$\max\{D_n(s_0, s_1, \dots, s_{n-1}); 0 \le s_0 \le s_1 \le \dots \le s_{n-1} \le 1\} = b_n.$$

First observe that the following representations hold for (s_0, \dots, s_{n-1}) with $s_j > 0$ for $j = 0, \dots, n-2$ and $0 < s_{n-1} < 1$:

$$\frac{\partial D_n}{\partial s_0} = (\mu_1/s_0)[D_n - (n-1)s_0^n];$$

$$s_{j+1} \frac{\partial D_n}{\partial s_{j+1}} - s_j \frac{\partial D_n}{\partial s_j} = \mu_1 s_0 \cdots s_j [D_n + n s_j^{n-1} - (n-1)s_{j+1}^n] \quad \text{for} \quad 0 \le j \le n-3;$$

(13)
$$s_{n-1}(1-s_{n-1})\frac{\partial D_n}{\partial s_{n-1}} - s_{n-2}\frac{\partial D_n}{\partial s_{n-2}} = \mu_1 s_0 \cdots s_{n-2} [D_n + n s_{n-2}^{n-1} - (n-1) s_{n-1}^n]; \text{ and}$$

$$\frac{\partial D_n}{\partial s_{n-1}} = (-\mu_1 s_0 \cdots s_{n-2}/(1-s_{n-1})^2)[D_n - 1 + n s_{n-1}^{n-1} - (n-1) s_{n-1}^n];$$

where $\mu_1 = V_1(s_0, \dots, s_{n-1})$, the expression in Lemma 2.5(ii). From (13) it can be deduced that the following three statements are equivalent for (s_0, \dots, s_{n-1}) with $s_j > 0$ for $j = 0, \dots, n-2$ and $0 < s_{n-1} < 1$.

(14a)
$$\frac{\partial D_n}{\partial s_i}(s_0, \dots, s_{n-1}) = 0 \quad \text{for} \quad j = 0, \dots, n-1;$$

(14b)
$$-ns_{j}^{n-1} - (n-1)s_{0}^{n} + (n-1)s_{j+1}^{n} = 0 \quad \text{for } j = 0, \dots, n-2, -ns_{n-1}^{n-1} - (n-1)s_{0}^{n} + (n-1)s_{n-1}^{n} + 1 = 0, \text{ and at } (s_{0}, \dots, s_{n-1}) \text{ satisfying these } n \text{ equations, } D_{n}(s_{0}, \dots, s_{n-1}) = (n-1)s_{0}^{n}; \text{ and}$$

(14c) letting
$$\beta = (n-1)s_0^n$$
, then $\eta_{j,n}(\beta) = s_j^n$ for $0 \le j \le n-1$, $1 = (n-1)(\eta_{n,n}(\beta) - \eta_{n-1,n}(\beta)) = H_n(\beta)$, and at (s_0, \dots, s_{n-1}) satisfying these n equations, $D_n(s_0, \dots, s_{n-1}) = (n-1)s_0^n = \beta$.

Let C be the region $C=\{(s_0,\cdots,s_{n-1}); 0\leq s_0\leq\cdots\leq s_{n-1}\leq 1\}$. Over this region $C,D_n\leq 1$, as can be seen by considering D_n as the difference of $E_n(s_0,\cdots,s_{n-1})$ and $V_n(s_0,\cdots,s_{n-1})$, the expressions in Lemma 2.5 (iii) and (i) respectively. Hence, for (s_0,\cdots,s_{n-1}) in C satisfying (14c), $0\leq D_n(s_0,\cdots,s_{n-1})=\beta\leq 1$, and only solutions of $H_n(\beta)=1$ in [0,1] are of interest. From this fact, Proposition 3.8, and (14a-c), there is a unique point $(\tilde{s}_0,\cdots,\tilde{s}_{n-1})$ in the interior of C at which $\partial D_n/\partial s_j=0$ for $j=0,\cdots,n-1$, and at this point $D_n(s_0,\cdots,s_{n-1})=(n-1)\tilde{s}_0^n>0$. Thus the maxima and minima for D_n in C occur at $(\tilde{s}_0,\cdots,\tilde{s}_{n-1})$ or on the boundary of C.

Consider the behavior of D_n at and near the boundary of C. If $s_0 = 0$ or $s_{n-1} = 1$ (or both), then $D_n(s_0, \dots, s_{n-1}) = 0$. Let (s_0, \dots, s_{n-1}) be a boundary point of C satisfying $0 < s_0 \le \dots < s_j = \dots = s_k < \dots \le s_{n-1} < 1$ for some $0 \le j < k \le n-1$. It can be shown from (13) that one of the following three conditions must hold at (s_0, \dots, s_{n-1}) :

(I)
$$D_n(s_0, \dots, s_{n-1}) \leq 0;$$

(II)
$$k \le n-2$$
 and $(-1, 1) \cdot \left(\frac{\partial D_n}{\partial s_m}, \frac{\partial D_n}{\partial s_{m+1}}\right) > 0$ for $m = j, \dots, k-1$,
so that $(-1, 1) \left(\frac{\partial D_n}{\partial s_j}, \frac{\partial D_n}{\partial s_k}\right) > 0$; or

(III)
$$k = n - 1$$
 and $(-1, 1) \cdot \left(\frac{\partial D_n}{\partial s_m}, \frac{\partial D_n}{\partial s_{m+1}}\right) > 0$ for for $m = j, \dots, n-3$,

and
$$(-1, 1 - s_{n-1}) \cdot \left(\frac{\partial D_n}{\partial s_{n-2}}, \frac{\partial D_n}{\partial s_{n-1}}\right) > 0$$
, so that $(-1, 1 - s_{n-1}) \cdot \left(\frac{\partial D_n}{\partial s_j}, \frac{\partial D_n}{\partial s_{n-1}}\right) > 0$.

From these observations one can find a point $(\bar{s}_0, \dots, \bar{s}_{n-1})$ in the interior of C with $D_n(\bar{s}_0, \dots, \bar{s}_{n-1}) > D_n(s_0, \dots, s_{n-1})$. Thus the maximum for D_n in C is at $(\tilde{s}_0, \dots, \tilde{s}_{n-1})$, and (12) follows

That the bound b_n is best possible is clear from the above reasoning (see also Proposition 5.3). \square

EXAMPLE 5.2. Let X_1, X_2, \cdots be i.i.d. random variables taking values in [0, 1]. Calculations of $\{b_n\}$ indicate that

$$E(X_1 \lor X_2)$$
-sup $\{EX_t : t \in T_2\} \le 0.0625;$
 $E(X_1 \lor \cdots \lor X_{100})$ -sup $\{EX_t : t \in T_{100}\} \le 0.1101;$ and
 $E(X_1 \lor \cdots \lor X_{10,000})$ -sup $\{EX_t : t \in T_{10,000}\} \le 0.113.$

In the present (additive comparison) case, unique extremal distributions (for which equality in Theorem B holds) can be given explicitly. For this purpose the following parameters are collected here. Fix n > 1. Let $\beta_n \in (0, 1)$ satisfy $H_n(\beta_n) = 1$ as in Proposition 3.8. For $j = 0, \dots, n-1$, \tilde{s}_j is given by $\tilde{s}_j = (\eta_{j,n}(\beta_n))^{1/n}$ and \tilde{p}_j by $\tilde{p}_0 = \tilde{s}_0$, $\tilde{p}_j = \tilde{s}_j - \tilde{s}_{j-1}$ for $j = 1, \dots, n-1$, and $\tilde{p}_n = 1 - \tilde{s}_{n-1}$.

PROPOSITION 5.3. For each n > 1, let $\tilde{Y} = \tilde{Y}(n)$ be the simple random variable taking values 0, $V_1(\tilde{Y})$, ..., $V_{n-1}(\tilde{Y})$, and 1 with probabilities $\tilde{p}_0, \dots, \tilde{p}_n$ respectively. Then $D_n(\tilde{Y}) = b_n$.

Note that the values $V_1(\tilde{Y}_1), \dots, V_{n-1}(\tilde{Y})$ can be computed from Lemma 2.5 (i, ii) through $\tilde{s}_0, \dots, \tilde{s}_{n-1}$.

EXAMPLE 5.4. (a) $\tilde{Y}(2) = 0$, $\frac{1}{2}$, and 1 with probabilities $\frac{1}{4}$, $\frac{1}{2}$, and $\frac{1}{4}$ respectively, and $D_2(\tilde{Y}(2)) = b_2 = \frac{1}{16}$.

(b) $\widetilde{Y}(10) \cong 0$, .166, .272, .347, .404, .449, .486, .517, .545, .570, 1 with probabilities \cong .638, .067, .048, .039, .033, .029, .026, .023, .021, .019, .054 respectively and $D_{10}(\widetilde{Y}(10)) = b_{10} \cong$.100.

6. Remarks. It is easy to see that for any fixed distribution X, $R_n(X) \to 1$ and $D_n(X) \to 0$ as $n \to \infty$, that is, $\lim_{n \to \infty} E(X_1 \lor \cdots \lor X_n) = \lim_{n \to \infty} \sup\{EX_t : t \in T_n\}$ where X_1, X_2, \cdots are independent random variables each with distribution that of X.

The parenthetical conclusion in Theorem B that $E(\min\{X_1, \dots, X_n\}) \ge \inf\{EX_t : t \in T_n\} - b_n(b-a)$ is immediate by symmetry. In contrast, no corresponding universal constant exists for ratio comparisons of $E(\min\{X_1, \dots, X_n\})$ and $\inf\{EX_t : t \in T_n\}$. See example 4.1 in [3].

Although the authors believe that the constants $\{a_n\}$ and $\{b_n\}$ are monotonically increasing, and hence convergent, they have not been able to demonstrate this nor identify the limits.

Acknowledgement. The authors would like to thank Professor M. J. Christensen for assistance with the numerical approximations of the constants, and Professors I. Aharoni and L. Karlovitz for several useful conversations related to analysis of the functions G_n and H_n .

REFERENCES

- [1] CHOW, Y. S., ROBBINS, H., and SIEGMUND, D. (1971). Great Expectations: The Theory of Optimal Stopping. Houghton Mifflin, Boston.
- [2] Hill, T. P. and Kertz, R. P. (1981). Ratio comparisons of supremum and stop rule expectations. Z. Wahrsch. verw. Gebiete 56 283-285.
- [3] Hill, T. P., and Kertz, R. P. (1981) Additive comparisons of stop rule and supremum expectations of uniformly bounded independent random variables. *Proc. Amer. Math. Soc.* 83 582–585.
- [4] Krengel, U. and Sucheston, L. (1978). On semiamarts, amarts, and processes with finite value. Adv. in Probability 4 197–266.

School of Mathematics Georgia Institute of Technology Atlanta, GA 30332