/ZU

GEORGIA INSTITUTE OF TECHNOLOGY
OFFICE OF CONTRACT ADMINISTRATION
SPONSORED PROJECT INITIATION

Date: June 6, 1979

Project Title: Goal and Average Cost Problems in Decision Processes with Finite State Spaces

Project No: G-37-612 Green Card

Project Director: Dr. Theodore P. Hill

Sponsor: Air Force Office of Scientific Research; Bolling AFB, DC 20332

Agreement Period: From 6/15/79 Until 9/30/80 (Performance Period)

Type Agreement: Contract No. F49620-79-C-0123

Amount: $18,303 AFOSR
4,578 GIT(G-37-322)
$22,881 TOTAL

NOTE: Partially funded through 9/30/79: $ 8,866 AFOSR
2,211 GIT
$11,077 TOTAL

Reports Required: Annaul Technical Report; Final Technical Report

Sponsor Contact Person (s):

Technical Matters
Program Manager: Dr. I. N. Shimi
Merle M. Andrew
Director of Mathematical & Information Sciences
Air Force Office of Scientific Research
ATTN: NM
Bldg. 410
Bolling AFB, DC 20332

Contractual Matters
(thru OCA)
Lt. Parsons
ATTN: PKZ
Air Force Office of Sci. Res.
Bldg. 410
Bolling AFB, DC 20332
202/767-4959

Marion R. Harrington, Major, USAF
Contracting Officer

Defense Priority Rating: DO-C9 under DMS Reg. 1

Assigned to: Mathematics (School/Laboratory)

COPIES TO:

- Project Director
- Division Chief (EES)
- School/Laboratory Director
- Dean/Director—EES
- Accounting Office
- Procurement Office
- Security Coordinator (OCA)
- Reports Coordinator (OCA)
- Library, Technical Reports Section
- EES Information Office
- EES Reports & Procedures
- Project File (OCA)
- Project Code (GTRI)
- Other

CA-3 (3/76)

GEORGIA INSTITUTE OF TECHNOLOGY
OFFICE OF CONTRACT ADMINISTRATION

SPONSORED PROJECT TERMINATION

Date: 1/12/81

Project Title: Goal and Average Cost Problems in Decision Processes with Finite State Spaces

Project No: G-37-612

Project Director: Dr. Theodore P. Hill

Sponsor: Air Force Office of Scientific Research (#F49620-79-C-0123)

Effective Termination Date: 9/30/80 (Perf. Period)

Clearance of Accounting Charges: Final Report due 11/30/80

Grant/Contract Closeout Actions Remaining:

- _ Final Invoice and Closing Documents
- x Final Fiscal Report
- x Final Report of Inventions
- _ Govt. Property Inventory & Related Certificate
- _ Classified Material Certificate
- _ Other __________

Assigned to: Mathematics (School/~~Laboratory~~)

COPIES TO:

Project Director
Division Chief (EES)
School/Laboratory Director
Dean/Director—EES
Accounting Office
Procurement Office
Security Coordinator (OCA)
✓ Reports Coordinator (OCA)

Library, Technical Reports Section
EES Information Office
Project File (OCA)
Project Code (GTRI)
Other C.E. Smith

CA-4 (1/79)

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

G-37-612

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Goal and average cost problems in decision processes with finite state spaces		5. TYPE OF REPORT & PERIOD COVERED Final technical 6/79-9/80
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Theodore P. Hill		8. CONTRACT OR GRANT NUMBER(s) F49620-79-C-0123
9. PERFORMING ORGANIZATION NAME AND ADDRESS T. P. Hill School of Mathematics, Georgia Inst. of Technology Atlanta, Georgia 30332		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research Air Force Systems Command, USAF Bolling A.F.B., Washington, D. C. 20332		12. REPORT DATE
		13. NUMBER OF PAGES
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report)
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Decision processes, gambling theory, dynamic programming, Markov processes, stationary strategy

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The principal investigator carried out research in probability theory, and, in particular, in the theory of decision processes. First, it was shown that in every finite state decision process (gambling problem) with single fixed goal, there always exsits a stationary strategy which not only (nearly) maximizes the probability of reaching the goal, but (nearly) minimizes the expected time to the goal. This result was considerably generalized to include decision processes with arbitrary state spaces and total cost criteria.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Investigations of these processes led to results in optimal stopping theory, and in classical probability theory. Universal, best possible constants were found which compared the optimal expected return of a decision maker with the expected supremum of a sequence of independent random variables. A generalization of the classical Borel-Cantelli Lemma was found, as was a very general conditioning principle for strong laws of several forms.

The question of existence of good Markov strategies in finite state decision processes with average reward criteria was addressed, and various partial results were obtained, although the general case was not settled.

I. Research Objectives

The objective of the supported research was a continuation of the principal investigator's analysis of decision processes with arbitrary decision sets, with special emphasis on two classical payoff functions. The first was a process with single fixed goal in which the objective is to maximize the probability of reaching the goal, and to minimize the expected time to the goal. A specific objective was to determine whether one can do as well with stationary strategies as he can with strategies which take the whole past into account. The second object of study was the average reward payoff in finite state decision processes, a specific objective being to determine whether or not strategies based only on the current time and state are as good as those which take the whole past into account.

II. Status of the Research

A complete, affirmative answer was found to the first question; in every dynamic programming or gambling problem with single fixed goal and finite state space, there exists a stationary strategy which not only uniformly (nearly) maximizes the probability of reaching the goal, but also uniformly (nearly) minimizes the expected time to the goal. Techniques in the proof of this result were considerably generalized to answer several questions about decision processes with arbitrary state spaces and total-cost criteria. It was shown that in a countable state decision process with non-negative costs depending on the current state, the action taken, and the following state, there is always available a Markov strategy which uniformly (nearly) minimizes the expected total cost. If the costs are strictly positive and depend only on the current state, there is even a stationary strategy with the same property.

Investigations of these results led peripherally to several results in optimal stopping theory, and in classical probability theory. Universal, best possible constants were found which compared the optimal expected return of a decision maker with the expected supremum of a sequence of random variables. For example, it was found that for every sequence of independent random variables taking only values between zero and one, the difference between the optimal stop rule expectation and the expected value of the supremum of the random variables is no more than one-fourth. Results in classical probability stemming from this research include a stronger form of the Borel-Cantelli Lemma, and a very general conditioning principle for strong laws which conclude the partial sums converge almost surely.

For the question of existence of good Markov strategies in decision processes with average reward criteria, various partial results have been obtained, including examples showing that the limit of good strategies for the discounted reward payoff is not necessarily average-reward good. This research is still in progress and it is hoped that a complete answer to the finite state case is not far away.

III. List of Publications

1. "On Reaching a Goal Quickly," Technical Report.

2. "Ratio Comparisons of Supremum and Stop Rule Expectations," (with R. Kertz), submitted to Z. Wahrscheinlichkeitstheorie.

3. "Decision Processes with Total-Cost Criteria," (with S. Demko), accepted for publication in Annals of Probability.

4. "Additive comparisons of Stop Rule and Supremum Expectations of Uniformly Bounded Independent Random Variables," (with R. Kertz), submitted to Proceedings of the A.M.S.

5. "A Stronger Form of the Borel-Cantelli Lemma," submitted to Pacific Journal of Mathematics.

6. "Comparisons of Stop Rule and Supremum Expectations of i.i.d. Random Variables," (with R. Kertz), submitted to Annals of Probability.

7. "Conditional Generalizations of Strong Laws Which Conclude the Partial Sums Coverage Almost Surely," submitted to Annals of Probability.

IV. Spoken Papers

A. Presented

1. "Betting Against a Prophet," seminar in Mathematics Department at University of California at Berkeley, August 1979.

2. "A Stronger Form of the Borel-Cantelli Lemma," seminar in the Mathematics Department at University of California at Berkeley, September 1979.

3. "On the Existence of Good Markov Strategies," Colloquium Presentation, Statistics Department, University of California at Berkeley, October 1979.

4. "Markov Decision Processes, Gambling, and Dynamic Programming," Colloquium Presentation, Mathematics Department, University of Hawaii, April 1980.

5. "Conditional Generalizations of Strong Laws," seminar in the Mathematics Department, University of California at Berkeley, June 1980.

6. ______________________________, Probability Seminar, Mathematics Department, Georgia Institute of Technology, October 1980.

7. "A Stronger Form of the Borel-Cantelli Lemma," Probability Seminar, Mathematics Department, Georgia Institute of Technology, November 1980.

B. Scheduled

1. "Conditional Generalizations of Strong Laws," Annual American Mathematical Society Meeting, San Francisco, January 1981.

2. "Finite State Decision Processes," Operations Research Seminar, Georgia Institute of Technology, February 1981.

G-31-612

Conditional Generalizations of Strong Laws Which Conclude the Partial Sums Converge Almost Surely

T.P. Hill[1]

Abstract

Suppose that for every independent sequence of random variables satisfying some hypothesis condition H, it follows that the partial sums converge almost surely. Then it is shown that for every arbitrarily-dependent sequence of random variables, the partial sums converge almost surely on the event where the conditional distributions (given the past) satisfy precisely the same condition H. Thus many strong laws for independent sequences may be immediately generalized into conditional results for arbitrarily-dependent sequences.

AMS(MOS) subject classifications (1970) Primary 60F15; Secondary 60G45.

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

CONDITIONAL GENERALIZATIONS OF STRONG LAWS

1. Introduction

If every sequence of independent random variables having property A has property B almost surely, does every arbitrarily-dependent sequence of random variables have property B almost surely on the set where the conditional distributions have property A?

Not in general, but comparisons of the conditional Borel-Cantelli Lemmas, the conditional three-series theorem, and many martingale results with their independent counterparts suggest that the answer is affirmative in fairly general situations. The purpose of this note is to prove Theorem 1, which states, in part, that if "property B" is "the partial sums converge", then the answer is *always* affirmative, *regardless* of "property A". Thus many strong laws for independent sequences (even laws yet undiscovered) may be immediately generalized into conditional results for arbitrarily-dependent sequences.

2. Main Theorem

In this note, $\mathbb{Y} = (Y_1, Y_2, \cdots)$ is a sequence of random variables on a probability triple $(\Omega, \mathcal{A}, P)$, $S_n = Y_1 + Y_2 + \cdots + Y_n$, and $\mathcal{F}_n$ is the sigma field generated by $Y_1, \cdots, Y_n$. Let $\pi_n(\cdot,\cdot)$ be a regular conditional distribution for Y_n given $\mathcal{F}_{n-1}$, and $\Pi = (\pi_1, \pi_2, \cdots)$. Let $\mathcal{B}$ denote the Borel σ-field on $\mathbb{R}$, and $\mathcal{B}^\infty$ the product Borel σ-field on $\mathbb{R}^\infty$; let $\mathcal{P}(\mathbb{R})$ denote the space of probability measures on $(\mathbb{R},\mathcal{B})$, and let $C = P(\mathbb{R}) \times P(\mathbb{R}) \times \cdots$. As a final convention, let $L(X)$ denote the distribution of the random variable X.

Let $B \in \mathcal{B}^{\infty}$. With the above notation, the question this note addresses is: when is the following statement (S) true?

(S) If $A \subset C$ is such that $(X_1, X_2, \cdots) \in B$ a.s. whenever $X_1, X_2, \cdots$ are independent and $(L(X_1), L(X_2), \cdots) \in A$, then for arbitrary $\mathbb{Y}$, $\mathbb{Y} \in B$ a.s. on the set where $\Pi \in A$.

A partial answer is given by

Theorem 1. (S) holds in the following three cases:

(i) $B = \{(r_1, r_2, \cdots) \in \mathbb{R}^{\infty} : \sum^{n} r_j \text{ converges}\}$;

(ii) $B = \liminf_{n\to\infty} \{(r_1, r_2, \cdots) : r_n \in A_n\}$; and

(iii) $B = \limsup_{n\to\infty} \{(r_1, r_2, \cdots) : r_n \in A_n\}$,

where $A_n \in \mathcal{B}$, $n = 1, 2, \cdots$).

3. Applications of Theorem 1

As a first application of Theorem 1, consider the two well-known conditional results: Levy's conditional form of the Borel-Cantelli Lemmas [4, p. 249],

(1) For any sequence of random variables $Y_1, Y_2, \cdots$ taking only the values 0 and 1, $\sum_1^{\infty} Y_n$ is finite (infinite) almost surely where $\sum_1^{\infty} E(Y_n | F_{n-1})$ is finite (infinite);

and the conditional three-series theorem [e.g., 5, p. 66],

(2) For any sequence of random variables $Y_1, Y_2, \cdots$, the partial sums S_n converge almost surely on the event where the three series

$$\sum_1^\infty P(|Y_n| \geq c | F_{n-1}), \quad \sum_1^\infty E[Y_n\ I(|Y_n| \leq c) | F_{n-1}], \text{ and}$$

$$\sum_1^\infty \ E[Y_n^2\ I(|Y_n| \leq c) | F_{n-1}] - E^2[Y_n\ I(|Y_n| \leq c) | F_{n-1}]\} \text{ all converge.}$$

Both results (1) and (2) follow immediately from Theorem 1 and their classical counterparts for independent sequences. Similarly, in many martingale theorems the independent case is also the extremal one. As a second application of Theorem 1, for example, note that the following martingale results of Doob [2, p. 320] and of Chow [1] follow immediately from (i) and the special case of independence:

(3) If $\{Y_n, F_n\ n \geq 1\}$ is a martingale difference sequence, then S_n converges a.s. where $\sum_1^\infty E[Y_n^2 | F_{n-1}] < \infty$; and

(4) If $\{Y_n, F_n\ n \geq 1\}$ is a martingale difference sequence and $\{b_n, n \geq 1\}$ is a sequence of positive constants such that $\sum_1^\infty b_n < \infty$, then S_n converges almost surely where $\sum_1^\infty b_n^{1-p/2}\ E[|Y_n|^p | F_{n-1}] > \infty$ for some $p > 2$.

Via Theorem 1(i) one may deduce immediately a conditional generalization of practically *any* result for sequences of independent random variables in which the conclusion is "S_n converges almost surely." Although the above applications all have hypotheses involving conditional moments, virtually any hypothesis conditions will carry over. As one final example, consider the well-known fact [e.g., 5, p. 102] that if $Y_1, Y_2, \cdots$ are independent and S_n converges in probability, then S_n converges almost surely. Theorem 1 allows the generalization of this fact given by Theorem 2 below.

Definition. A sequence of probability measures $(\mu_1, \mu_2, \cdots)$ *sums in probability* if, for any independent sequence of random variables $X_1, X_2, \cdots$ with $\mathcal{L}(X_i) = \mu_i$, it follows that $X_1 + \cdots + X_n$ converges in probability to some random variable X.

Theorem 2. Let $Y_1, Y_2, \cdots$ be an arbitrary sequence of random variables. Then S_n converges almost surely on the set where the conditional distributions Π sum in probability.

4. Proof of Theorem 1.

For fixed $B \in \mathcal{B}^\infty$, consider the statement

(S') $$P(\{\omega: P_{\Pi(\omega)}(B) = 1\} \cap \mathbf{Y} \notin B) = 0,$$

where $P_{\Pi(\omega)}$ is the product measure $\pi_1(\omega) \times \pi_2(\omega) \times \cdots$ on $(\mathbb{R}^\infty, \mathcal{B}^\infty)$.

Without loss of generality, assume $(\Omega, \mathcal{A}, P)$ is complete.

Lemma 1. (S) $\Leftrightarrow$ (S').

<u>Proof</u>. "$\Rightarrow$" Let $A = \{\vec{\mu} \in C: P_{\vec{\mu}}(B) = 1\}$. Then $\{\omega: P_{\Pi(\omega)}(B) = 1\} =$
$= \{\omega: \Pi(\omega) \in A\}$, so $P(\{\omega: P_{\Pi(\omega)}(B)=1\} \cap \mathbb{Y} \notin B) = P(\{\omega: \Pi(\omega) \in A\} \cap \mathbb{Y} \notin B) = 0$.
"$\Leftarrow$" Since $\{\omega: \Pi(\omega) \in A\} \subset \{\omega: P_{\Pi(\omega)}(B) = 1\} \in \mathcal{a}$, it follows (by completeness) that $P(\{\omega: \Pi(\omega) \in A\} \cap \mathbb{Y} \notin B) = P(\{\omega: P_{\Pi(\omega)}(B) = 1\} \cap \mathbb{Y} \notin B) = 0$. □

<u>Proof of Theorem 1</u>.

For (i), Lemma 1 implies it is enough to show that (S') holds for $B = \{(r_1, r_2, \cdots) \in \mathbb{R}^\infty: \sum^n r_j \text{ converges}\}$. Let $(\hat{\Omega}, \hat{\mathcal{a}}, \hat{P})$ be a copy of $(\Omega, \mathcal{a}, P)$, and (enlarging this new space if necessary) for each $\omega \in \Omega$, let $Z_1(\omega), Z_2(\omega), \cdots$ be a sequence of independent random variables on $(\hat{\Omega}, \hat{\mathcal{a}}, \hat{P})$ with $L(Z_n(\omega)) = \pi_n(\omega)$, (that is, $\hat{P}(Z_n(\omega) \in E) = \pi_n(\omega, E) =$
$= P(Y_n \in E \mid Y_1(\omega), \cdots, Y_{n-1}(\omega))$. Then

(5) $P(\{\omega: P_{\Pi(\omega)}(B)=1\} \cap \mathbb{Y} \notin B)$.

$=P(\{\omega: Z_1(\omega) + \cdots + Z_n(\omega) \text{ converges a.s. (in } \hat{\Omega}, \hat{\mathcal{a}}, \hat{P})\} \cap \mathbb{Y} \notin B)$

$=P(\{\omega: \text{the three series } \sum_1^\infty \hat{P}(|Z_n(\omega)| \geq 1), \sum_1^\infty \hat{E}(Z_n(\omega) \cdot I(|Z_n| \leq 1), \text{ and }$
$\sum_1^\infty \hat{V}\mathrm{ar}(Z_n(\omega) \cdot I(|Z_n| \leq 1)) \text{ all converge}\} \cap \mathbb{Y} \notin B)$

$=P(\{\text{event where the three series } \sum_1^\infty P(|Y_n| \geq 1 \mid F_{n-1}),$
$\sum_1^\infty E(Y_n \cdot I(|Y_n|\ 1) \mid F_{n-1}), \text{ and } \sum_1^\infty [E(Y_n^2 \cdot I(|Y_n| \leq 1) \mid F_{n-1}) -$
$- E^2(Y_n I(|Y_n| \leq 1) \mid F_{n-1})] \text{ all converge}\} \cap \mathbb{Y} \notin B) = 0$,

where the first equality in (5) follows by the definition of $Z_n(\omega)$, the second by Kolmogorov's Three-Series Theorem and independence of the $\{Z_i(\omega)\}$, the third by definition of $Z_n(\omega)$ and π_n, and the last by the conditional three series theorem (2). This completes the proof of (i).

For (ii) and (iii), application of the same technique using, in place of the three-series theorems, the classical and (Levy's) conditional form (1) of the Borel-Cantelli lemmas yields the desired conclusion. □

5. Remarks

The class of sets $B \in \mathcal{B}^\infty$ for which (S) and (S') hold is not closed under complementation; a counterexample to the converse of the conditional three-series theorem due to Dvoretzky and to Gilat [3] demonstrates that (S) does not hold in general for $B =$ "S_n does not converge". Whether (S) holds for such useful sets as "$S_n/n \to 0$", or "$\limsup S_n/a_n = 1$", is not known to the author.

Acknowledgement. The author is grateful to Professor Lester Dubins for several useful conversations.

References

1. Chow, Y.S. (1965) Local convergence of martingales and the law of large numbers. Ann. Math. Statist., 36, 552-558.

2. Doob, J.L. (1953) "Stochastic Processes". Wiley, New York.

3. Gilat, D. (1971) On the nonexistence of a three series condition for series of non-independent random variables. Ann. Math. Statist. 42, 409-410.

4. Lévy, P. (1937) Theorie de l'addition des variables aleatoires. Gauthier-Villars, Paris.

5. Stout, W.F. (1974) "Almost Sure Convergence". Academic Press, New York.

6-37-612

Comparisons of Stop Rule and Supremum Expectations of i.i.d. Random Variables

(Stop Rule and Supremum Expectations)

by

T.P. Hill[1]
Robert P. Kertz

Abstract

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 $E(\max\{X_1,\cdots,X_n\}) \le a_n \sup\{EX_t : t \in T_n\}$, and the bound a_n is best possible. Similar universal constants $0 < b_n < 1/4$ are found so that if the $\{X_i\}$ are i.i.d. random variables taking values only in $[a,b]$, then $E(\max\{X_1,\cdots,X_n\}) \le \sup\{EX_t : t \in T_n\} + b_n(b-a)$, 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. 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.

§1. Introduction.

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

$$(1) \qquad E(\max\{X_1,\cdots,X_n\}) \leq 2 \sup\{EX_t : t \in T_n\} \text{ and}$$

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

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

and 1/4 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 1/8 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 "1/4" 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 $1.1 < a_n < 1.6$, $n=2,3,\cdots$, 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. If $n > 1$ and $X_1, X_2, \cdots, X_n$ are i.i.d. non-negative random variables, then $E(\max\{X_1,\cdots,X_n\}) \le a_n \sup\{EX_t : t \in T_n\}$. Moreover, a_n is the best possible bound and is not attained except in the trivial cases X_1 is almost surely 0 or has infinite expectation.

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

Theorem B. If $X_1,\cdots,X_n$ are i.i.d. random variables taking values only in $[a,b]$, then $E(\max\{X_1,\cdots,X_n\}) \le \sup\{EX_t : t \in T_n\} + b_n(b-a)$, (equivalently, $E(\min\{X_1,\cdots,X_n\}) \ge \inf\{EX_t : t \in T_n\} - b_n(b-a)$) 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 $n=1,2,\cdots$, $E_n(X) = E(X_1 \vee \cdots \vee X_n)$, and $V_n(X) = \sup\{EX_t : t \in T_n\}$, where $X_1,\cdots,X_n$ 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, p. 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 by $t^* = j \Leftrightarrow \{t^* > j-1$ 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 $0 \leq a < b < \infty$, 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 $0 \leq a < b < \infty$. Then $EY = EY_a^b$, and if X is any random variable independent of both Y and Y_a^b, then $E(X \vee Y) \leq E(X \vee Y_a^b)$.

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, $V_1(X)$, $V_2(X), \cdots, V_{n-1}(X)$, and 1, and satisfying both

$V_j(Y) = V_j(X)$ for $j=1,2,\cdots,n$, and $E_n(Y) \geq E_n(X)$.

Proof. Define Y through Definition 2.2 by

$$Y = \left(\cdots\left(X_0^{V_1(X)}\right)_{V_1(X)}^{V_2(X)}\cdots\right)_{V_{n-1}(X)}^{1}.$$

The conclusion follows easily from Lemmas 2.1 and 2.3. □

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 $s_0,s_1,\cdots,s_{n-1}$.

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

Lemma 2.7. Let Y be any random variable and $0 \leq a < \infty$. Then for all α sufficiently large, $EY=EY_{a,\alpha}$ and if X is any random variable independent of both Y and $Y_{a,\alpha}$, then $E(X \vee Y) \leq E(X \vee Y_{a,\alpha})$. 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 $P(Y \geq a) > 0$ and fix any X independent of both Y and $\{Y_{a,\alpha}\}$. From the definition of $Y_{a,\alpha}$, the convexity of the function $\psi(y) = E(X \vee y)$, and the independence of X and Y, it follows that $E(X \vee Y_{a,\alpha})$ is a non-decreasing function of α and $\lim_{\alpha\to\infty} E(X \vee Y_{a,\alpha}) = \int_{Y<a} X \vee Y + E(X \vee a)P(Y \geq a) + E(Y-a)^{+}$, with the limit being attained if $P(X > a) \cdot P(Y > a) = 0$. The conclusion follows from these results and the dichotomy that $\int_{Y<a} X \vee Y + E(X \vee a)\, P(Y \geq a) + E(Y-a)^{+} > E(X \vee Y)$ if $P(X > a) \cdot P(Y > a) > 0$, and $=E(X \vee Y)$ if $P(X > a) \cdot P(Y > a) = 0$. The strict inequality in this dichotomy follows since for $P(X > a) \cdot P(Y > a) > 0$,

$$\int_{X\geq a, Y\geq a}(X-a+Y-a) - \int_{X\geq a, Y\geq a}[(X-a) \vee (Y-a)] = \int_{X\geq a, Y\geq a}[(X-a) \wedge (Y-a)] > 0. \quad \square$$

If $P(Y>a)>0$, then $\{Y,\alpha\}$ 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. Then there exists a non-negative bounded random variable Y satisfying both $V_j(Y) = V_j(X)$ for $j=1,2,\cdots,n$, 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 $\phi_n(w,x) = (n/(n-1))\cdot w^{(n-1)/n} + x/(n-1)$. For $\alpha \in [0,\infty)$, define inductively the functions $\eta_{j,n}$, $j=0,1,\cdots,n$, 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 $j=0,1,\cdots,n$.

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 $\eta_{j-1,n}(\alpha) > 0$, $\eta'_{j-1}(\alpha) > 0$, and $\eta''_{j-1,n}(\alpha) \leq 0$. Then it is clear that $\eta_{j,n}$ is continuous, and for

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

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

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 $\eta_{j,n}(\alpha) = (n/(n-1))(\eta_{j-1,n}(\alpha))^{(n-1)/n} + \alpha/(n-1) \le (n/(n-1))\cdot (\sigma_{j-1,n}(\alpha))^{(n-1)/n} + \alpha/(n-1) \le (n/(n-1))(\sigma_{j-1,n}(\alpha)+c_n)+\alpha/(n-1)= \sigma_{j,n}(\alpha)$. 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 $\mu_{j,n}(\alpha) = (n/(n-1))\, \mu_{j-1,n}(\alpha) + \alpha/(n-1) \le (n/(n-1))\, \eta_{j-1,n} + \alpha/(n-1) \le (n/(n-1))(\eta_{j-1,n}(\alpha))^{(n-1)/n} + \alpha/(n-1) = \eta_{j,n}(\alpha)$, completing the proof of (3).

If $G_n(\alpha) \le 1$ for all $\alpha\in[0,1]$, then it would follow from (3) that $G_n(1) = \eta_{n-1,n}(1) \ge \mu_{n-1,n}(1) = (n/(n-1))^n - 1 > e-1 > 1$, 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. □

<u>Example 3.5</u>. (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.

<u>Proposition 3.6</u>. For all $n > 1$,

(a) $[(n/(n-1))^{n-1}\,((n/(n-1))^n - 1)]^{-1} \le \alpha_n \le [(n/(n-1))^n - 1]^{-1}$; 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 $(n/(n-1))^{n-1} \nearrow e$. □

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

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

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 t which lies between g and u on $[\alpha_n,1]$, and which has no points in common with u.

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 $g(\beta) = f(\eta_{n-1,n}(\beta)) \geq f(k_n\beta + d_n) = t(\beta)$ on $[\alpha_n,1]$, it suffices to show that

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

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 $b_n \leq \alpha_n$ 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. □

<u>Example 3.9</u>. (a) For $n=2$, $H_2(\beta) = 2(2\sqrt{\beta} + \beta)^{1/2} - 2\beta^{1/2}$, and $\beta_2 = 1/16$.

(b) $\beta_3 \cong .077$, $\beta_4 \cong .085$, $\beta_5 \cong .090$, $\beta_{10} \cong .100$, $\beta_{100} \cong .110$, $\beta_{10,000} \cong .111$.

<u>Definition 3.10</u>. 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 $R_n(X) \le R_n(Z)$. Define Y=Z/(supremum of Z); then Y is a random variable taking its values in [0,1] and $R_n(X) \le R_n(Z) = R_n(Y)$. This establishes (6).

By Lemma 2.4, attention may be further restricted to simple random variables X taking on the values 0, $V_1(X),\cdots,V_{n-1}(X)$, and 1 (with probabilities $p_0,p_1,\cdots,p_n$ respectively). Let $s_j=p_0+\cdots+p_j$

for $j=0,\cdots,n-1$ 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 $0 < V_1(X)<\cdots<V_{n-1}(X) < 1$ and from Lemma 2.5 $R_n(X) = R_n(s_0,\cdots,s_{n-1})$ where $R_n(s_0,\cdots,s_{n-1})$ is the function defined for $s_j \geq 0$, $j=0,\cdots,n-1$, by

(7) $$R_n(s_0,s_1,\cdots,s_{n-1}) =$$

$$1 + \frac{\left(\sum_{j=1}^{n-1} s_{n-1}^j\right) s_0 s_1\cdots s_{n-2} - s_0^n - s_0 s_1^n - \cdots - s_0\cdots s_{n-3}s_{n-2}^n}{1+s_0+s_0s_1+\cdots+s_0s_1\cdots s_{n-2}}$$

The conclusion of Theorem A follows once it is shown that

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

where a_n was given in Definition 3.10.

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

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

$= r_n(\hat{s}_0,\cdots,\hat{s}_{n-2}) = a_n.$

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

(10a) $\frac{\partial r_n}{\partial s_j}(s_0,\cdots,s_{n-2}) = 0$ for $j=0,\cdots,n-2$;

(10b) $n\, s_{j+1}\cdots s_{n-2}-ns_j^{n-1}+[(1-s_{j+1}^n) + s_{j+1}(1-s_{j+2}^n)+\cdots+$

$+s_{j+1}\cdots s_{n-3}(1-s_{n-2}^n)]-r_n(s_0,\cdots,s_{n-2})(1+s_{j+1}+s_{j+1}s_{j+2}+\cdots+s_{j+1}\cdots s_{n-2})$

$=0$ for $0 \le j \le n-4$,

$ns_{n-2} - ns_{n-3}^{n-1} + (1-s_{n-2}^n)-r_n(s_0,\cdots,s_{n-2})\cdot(1+s_{n-2})=0$, and

$n-n\,s_{n-2}^{n-1} - r_n(s_0,\cdots,s_{n-2}) = 0;$

(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 $(s_0,\cdots,s_{n-2})$ satisfying these $n-2$ equations, $r_n(s_0,\cdots,s_{n-2}) = 1 +(n-1)s_0^n$; 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 $(s_0,\cdots,s_{n-2})$ satisfying these equations, $r_n(s_0,\cdots,s_{n-2}) = 1 + (n-1)s_0^n = 1+\alpha.$

Let $B \subset \mathbb{R}^{n-1}$ be the region $B = \{(s_0,\cdots,s_{n-2});\ s_j \geq 0$ 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.). □

<u>Example 4.2.</u> 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 $E(X_1\vee\cdots\vee X_{10,000})<1.342 \sup\{EX_t:t\in T_{10,000}\}$.

<u>Corollary 4.3</u>. 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 $E(\sup X_i) \leq (1+(e-1)^{-1}) \sup\{EX_t:t\in T\}$.

<u>Proof</u>. Apply Proposition 3.6(b) to Theorem A. □

It is perhaps of some interest to indentify 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 $j=0,\cdots,n-2$, $\hat{s}_j$ 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 $j=1,\cdots,n-2$, 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 $j=0,\cdots,n-2$, $P(\hat{X}=V_{n-1}(\hat{X})) \in (\hat{p}_{n-1}-\varepsilon,\hat{p}_{n-1})$ 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 $\hat{p}_0,\cdots,\hat{p}_{n-2},\hat{p}_{n-1}-\varepsilon$, ε 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 $\varepsilon \searrow 0$. □

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, $(\hat{p}_0,\cdots,\hat{p}_9) \cong$ (0.711925, 0.070190, 0.047863, 0.037426, 0.030936, 0.026304, 0.022730, 0.019837, 0.017423, 0.015367)). For $\varepsilon>0$ small consider the random variables $\hat{X}=\hat{X}(n,\varepsilon)$ taking values 0, $\varepsilon\cdot v_1,\cdots,\varepsilon\cdot v_9$, and 1 with probabilities $\hat{p}_0,\cdots,\hat{p}_8$, $\hat{p}_9-\varepsilon$, and ε respectively, where $(v_1,\cdots,v_9) \cong$ (3.32872, 5.69852, 7.55198, 9.09031, 10.4247, 11.6234, 12.7317, 13.7818, 14.7974). For $\varepsilon>0$ 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 $\varepsilon>0$, let $Y_\varepsilon=X- V_n(X) + \varepsilon$. Then $E_n(Y_\varepsilon) = E_n(X) - V_n(X) + \varepsilon$, and $V_n(Y_\varepsilon) = \varepsilon$, so $1 + (E_n(X)-V_n(X))/\varepsilon = R_n(Y_\varepsilon) \nearrow \infty$ as $\varepsilon \searrow 0$.

§5. Proof of Theorem B and its extremal distributions.

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

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 $D_n(X) \le b_n$ 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 $0, V_1(X),\cdots,V_{n-1}(X)$, and 1 with probabilities $p_0,p_1,\cdots,p_n$ respectively. Let $s_j=p_0+p_1+\cdots+p_j$ 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 $D_n(X)$. $D_n(X)=D_n(s_0,s_1,\cdots,s_{n-1})$ where $D_n(s_0,\cdots,s_{n-1})$ is the continuous func defined on $\{(s_0,\cdots,s_{n-1}); 0 \le s_0 \le \cdots \le s_{n-1} \le 1\} \cup \{(s_0,\cdots,s_{n-1}); 0 < s_{n-1} < 1$ and $s_j > 0$ for $j=0,\cdots,n-2\}$ by

(11) $D_n(s_0,s_1,\cdots,s_{n-1}) = 0$ if $0=s_0=\cdots=s_i\leq\cdots\leq s_{n-1} = 1$,

$$\text{and } = \frac{(1-s_{n-1})\left\{\left(\sum_{j=1}^{n-1} s_{n-1}^{j}\right)s_0s_1\cdots s_{n-2}-s_0^n-s_0s_1^n-\cdots-s_0s_1\cdots s_{n-3}s_{n-2}^n\right\}}{(1-s_{n-1})(1+s_0+s_0s_1+\cdots+s_0s_1\cdots s_{n-3}) + s_0s_1\cdots s_{n-2}}$$

otherwise.

It remains only to show that

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

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

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

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

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

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

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

(14a) $\frac{\partial D_n}{\partial s_j}(s_0,\cdots,s_{n-1})=0$ for $j=0,\cdots,n-1$;

(b) $-ns_j^{n-1}-(n-1)s_0^n + (n-1)s_{j+1}^n = 0$ for $j=0,\cdots,n-2$,

$-ns_{n-1}^{n-1} - (n-1)s_0^n + (n-1)s_{n-1}^n + 1=0$, and at $(s_0,\cdots,s_{n-1})$ satisfying these n equations, $D_n(s_0,\cdots,s_{n-1}) = (n-1)s_0^n$; 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 $(s_0,\cdots,s_{n-1})$ satisfying these n equations, $D_n(s_0,\cdots,s_{n-1})=(n-1)s_0^n = \beta$.

Let C be the region $C=\{(s_0,\cdots,s_{n-1});\ 0 \le s_0 \le \cdots \le s_{n-1} \le 1\}$. Over this region C, $D_n \le 1$, 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), $0 \le D_n(s_0,\cdots,s_{n-1})=\beta \le 1$, 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 $D_n(s_0,\cdots,s_{n-1})=0$. Let $(s_0,\cdots,s_{n-1})$ be a boundary point of C satisfying $0 < s_0 \le \cdots \le s_j = s_{j+1} \le \cdots \le s_{n-1} < 1$. It can be shown from (13) that either $D_n(s_0,\cdots,s_{n-1}) \le 0$ or $(-1,1)\cdot\left(\frac{\partial D_n}{\partial s_j}, \frac{\partial D_n}{\partial s_{j+1}}\right) > 0$ if $0 \le j \le n-3$ and $(-1,1-s_{n-1})\cdot\left(\frac{\partial D_n}{\partial s_{n-2}}, \frac{\partial D_n}{\partial s_{n-1}}\right) > 0$ if $j=n-2$. From this observation one can find a point $(\bar{s}_0,\cdots,\bar{s}_{n-1})$ in the interior of C with $D_n(\bar{s}_0,\cdots,\bar{s}_{n-1}) > D_n(s_0,\cdots,s_{n-1})$. Thus the maximum for D_n in C is at $(\tilde{s}_0,\cdots,\tilde{s}_{n-1})$, and (12) follows.

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

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

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 $j = 0, \cdots, n-1$, $\tilde{s}_j$ 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 $j = 1, \cdots, n-1$, 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 $0, V_1(\tilde{Y}), \cdots, V_{n-1}(\tilde{Y})$, and 1 with probabilities $\tilde{p}_0, \cdots, \tilde{p}_n$ respectively. Then $D_n(\tilde{Y}) = b_n$.

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

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

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

§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, $\lim_{n\to\infty} E(X_1 \vee \cdots \vee X_n) = \lim_{n\to\infty} \sup\{EX_t : t \in T_n\}$ where $X_1, X_2, \cdots$ are independent random variables each with distribution that of X.

The parenthetical conclusion in Theorem B that $E(\min\{X_1, \cdots, X_n\}) \geq \inf\{EX_t : t \in T_n\} - b_n(b-a)$ is immediate by symmetry. In contrast, no corresponding universal constant exists for ratio comparisons of $E(\min\{X_1, \cdots, X_n\})$ 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. Y.S. Chow, H. Robbins, and D. Siegmund. Great Expectations: The Theory of Optimal Stopping, Houghton Mifflin, Boston (1971).

2. T.P. Hill and R.P. Kertz, "Ratio Comparisons of Supremum and Stop Rule Expectations," to appear.

3. T.P. Hill and R.P. Kertz, "Additive Comparisons of Stop Rule and Supremum Expectations of Uniformly Bounded Independent Random Variables," to appear.

4. U. Krengel and L. Sucheston, "On Semiamarts, Amarts, and Processes with Finite Value," Advances in Probability, Vol. 4 (1978), 197-266.