Regret-optimal policies in absorbing semi-Markov decision processes with multiple constraints

Yoshinobu Kadota[a], Masami Kurano[b] and Masami Yasuda[c]

[a]*Faculty of Education, Wakayama University , Wakayama 640-8510 Japan*
yoshi-k@math.edu.wakayama-u.ac.jp
[b]*Faculty of Education, Chiba University, Chiba 263-8522 Japan*
kurano@faculty.chiba-u.jp
[c]*Faculty of Science, Chiba University, Chiba 263-8522 Japan*
yasuda@math.s.chiba-u.ac.jp,

Abstract

We consider a constrained regret-optimization problem for semi-Markov decision processes. The expected regret-utility of the total reward is minimized subject to multiple expected regrete-utility constraints and the planning horizon is a reaching time to a given absorbing subset. By introducing a corresponding Lagrange function, a saddle-point theorem is proved. The existence of a constrained optimal policy is characterized by optimal action sets specified with a parametric utility.

Keywords: Semi-Markov decision process; Utility constraint; Lagrange technique; Saddle point; Optimal policy.

1 Introduction and notation

In decision making, it may be more approciate to evaluate each decision or policy under a regret-optimality criterion. Inour previous work[12], we had considered the general regret-constraint problem for absorbing semi-Markov decision processes(semi-MDP's), in which the expected utility of the total reward earned until the stopping time is minimized. Its regret-optimal policy is characterized by the corresponding optimality equation.

In this paper, we treat the constrained optimization problem for the same model as [12]. However, it often occurs, in a social life or in a business that that we should maxmimize the reward under several utility functions. For example, in the group decision making with different utility functions each player would like to maximize own specified utility function. In such a case, not only one type of expected utility but keeping other types higher than some given bound.

Here, we consider the constrained regret-optimization problem for semi-MDPs in which the expected regret-utility of the total reward earned until the reaching time to a given absorbing subset is minimized subject to multiple expected regret-utility constraints and the objectives is to show that the Lagrange approach to the utility-constraints case is made successfully. In fact, by introducing a corresponding Lagrange function, a saddle point theorem is obtained and the existence of a constrained optimal

policy had proved. Also a constrained optimal policy is characterized by optimal action sets specified tih a paremtric utility.

Similar to the previous work[12], we do not restrict the type of regret-utility functions and it is expected to enlarge the practical application of the optimization problem. Fro the utility discussions for MDPs and constrained MDPs, refer to [5, 6, 8, 11, 13] and their references. In remainder of this section, a constrained regret-utility optimizatio problem is formulated under the absorbing semi-MDPs model.

A semi-MDP is specified by the next five components:

(i) a countable state space: $S = \{0, 1, 2, \cdots\}$,

(ii) a finite action space: $A = \{1, 2, \cdots, m\}, m < \infty$,

(iii) a transition probability distribution: $\{(p_{ij}(a); j \in S) \,|\, i \in S, a \in A\}$,

(iv) a distribution function $\{F_{ij}(\cdot|a) \,|\, i, j \in S, a \in A\}$ of the time between transitions,

(v) an immediate reward r and a reward rate d which are functions from $S \times A$ to R_+, where $R_+ = [0, \infty)$.

When the system is in state $i \in S$ and action $a \in A$ is taken, then it moves to a new state $j \in S$ with the sojourn time τ, and the reward $r(i,a) + d(i,a)\tau$ is obtained, where the new state j and the sojourn time τ are distributed with $p_{i\cdot}(a)$ and $F_{ij}(\cdot|a)$ respectively. This process is repeated from the new state $j \in S$.

The sample space is the product space $\Omega = (S \times A \times R_+)^\infty$. Let X_n, Δ_n and τ_{n+1} be random quantities such that $X_n(\omega) = x_n$, $\Delta_n(\omega) = a_n$ and $\tau_{n+1}(\omega) = t_{n+1}$ for all $\omega = (x_0, a_0, t_1, x_1, a_1, t_2, \cdots) \in \Omega$ and $n = 0, 1, 2, \cdots$. Let $H_n = (X_0, \Delta_0, \tau_1, \cdots, X_n)$ be a history until time n. A policy $\pi = (\pi_0, \pi_1, \cdots)$ is a sequence of conditional probabilities $\pi_n = \pi_n(\cdot \mid H_n)$ such that $\pi_n(A \mid H_n) = 1$ for all histories $H_n \in (S \times A \times R_+)^n \times S$. The set of all policies is denoted by Π. A policy $\pi = (\pi_0, \pi_1, \cdots)$ is called stationary if there exists a function $f : S \to A$ such that $\pi_n(\{f(X_n)\} \mid H_n) = 1$ for all $n \geq 0$ and $H_n \in (S \times A \times R_+)^n \times S$. Such a policy is denoted by f^∞.

For any $\pi \in \Pi$, we assume that

(i) $Prob(X_{n+1} = j | X_0, \Delta_0, \tau_1, \cdots, X_n = i, \Delta_n = a) = p_{ij}(a)$

(ii) $Prob(\tau_{n+1} \leq t | X_0, \Delta_0, \tau_1, \cdots, X_n = i, \Delta_n = a, X_{n+1} = j) = F_{ij}(t|a)$

for all $n \geq 0$, $i, j \in S$ and $a \in A$.

For any Borel set D, we denote by $P(D)$ the set of all probility measures on D. From (i) and (ii), we can define the probability measure $P_\pi^\nu \in P(\Omega)$ with an initial distribution $\nu \in P(S)$ viz. $Prob(X_0 = i) = \nu(i), i \in S$ and $\pi \in \Pi$.

For any subset $J_0 \subset S$, called as absorbing set, let

$$N := \min\{n > 0 | X_n \in J_0\}, \text{where} \min \emptyset = \infty. \tag{1.1}$$

The present value $\{\widetilde{D}_\ell : \ell = 1, 2, \cdots\}$ and the total lapsed time $\{\widetilde{\tau}_\ell : \ell = 1, 2, \cdots\}$ of the process $\{X_n, \Delta_n, \tau_{n+1} : n = 0, 1, 2, \cdots\}$ until the ℓ-th time are defined respectively by

$$\begin{aligned} &\widetilde{D}_\ell := \textstyle\sum_{n=0}^{\ell-1} (r(X_n, \Delta_n) + \tau_{n+1} d(X_n, \Delta_n)) \quad \textit{and} \\ &\widetilde{\tau}_\ell := \textstyle\sum_{n=1}^{\ell} \tau_n, \quad (\ell \geq 1). \end{aligned} \tag{1.2}$$

Let $G, H_i (i = 1, 2, \cdots, k) : R_+ \times R_+ \to R$ be Borel-measurable function, which will be called regret-utility functions as describing the general evaluation of the differences between the target value and the present value.

For any given threshold vector $\alpha = (\alpha_1, \cdots, \alpha_k) \in R^k$ and constraint-target vector $h = (h_1, \cdots, h_k) \in R^k$, let

$$V(\nu, \alpha, h) := \left\{ \pi \in \Pi \,\middle|\, E_\pi^\nu[H_i(h_i \widetilde{\tau}_N, \widetilde{D_N})] \leq \alpha_i \text{ for all } i\,(1 \leq i \leq k) \right\} \tag{1.3}$$

Then, for a constant g^*, called a target value, our problem is given in the following.

Problem A:

$$\text{Minimize} \quad E_\pi^\nu[G(g^* \widetilde{\tau_N}, \widetilde{D_N})] \quad \text{subject to} \quad \pi \in V(\nu, \alpha, h).$$

The optimal solution $\pi^* \in V(\nu, \alpha, h)$ of Problem A, if it exists, is called a ν-constraint regret optimal policy or simply a constraint regret optimal policy.

By a slight modification of the proof of Theorem 3.2 and 3.3 in Borkar[3], we have the following assertion.

Lemma 1.1 *For any $\nu \in P(S)$, let $\varphi(\nu) := \left\{ P_\pi^\nu \in P(\Omega) \,\middle|\, \pi \in \Pi \right\}$. Then, $\varphi(\nu)$ is a convex set and compact in the weak topology w.r.t. a function of Ω.*

The motivation of considering Problem A has its origin to take a comparison between the target value and the present value, that is, G by $g^* \widetilde{\tau_N}$ and $\widetilde{D_N}$, and also $H_i (i = 1, 2, \ldots, k)$ by $h_i \widetilde{\tau}_N$ and $\widetilde{D_N}$. For example, under the condition of Markov chain corresponding a policy is possitive recurrent and ireducible, then an average criterion

$$\begin{aligned} & \sup\lim_T \frac{1}{T} E_\pi(\textstyle\sum_t \nu(X_t, \Delta) \,|\, i) \leq \delta \\ \Leftrightarrow \quad & E_\pi[\delta \tilde{\tau} - \textstyle\sum_t \nu(X_t, \Delta)] =: E_\pi[G[\delta \tilde{\tau}, \textstyle\sum_t \nu(X_t, \Delta)]] \geq 0 \end{aligned}$$

where we set $G(x, y) := x - y$ and the value δ is provided that it means a target value.

In Section 2, the saddle point statement for Problem A will be described and its result is applied to obtain the existence of a constraint optimal policy. In Section 3, characterization of a constraint optimal policy is given.

2 Saddle point theorem for constrained semi-MDP

Now we discuss the saddle point-theorem for Lagrangian associated with Problem A. Firstly, for any initial distribution $\nu \in P(S)$, Lagrangeian L^ν is defined as follows.

$$L^\nu(\pi, \lambda) := \sum_{i=1}^{k} \lambda_i (\alpha_i - E_\pi^\nu[H_i(h_i \widetilde{\tau_N}, \widetilde{D_N})]) - E_\pi^\nu[G(g^* \widetilde{\tau_N}, \widetilde{D_N})]) \tag{2.1}$$

for any $\pi \in \Pi$ and $\lambda = (\lambda_1, \cdots, \lambda_k) \in R_+^k$. Without any confusion, " $\lambda = (\lambda_1, \cdots, \lambda_k) \in R_+^k$ " will be written simply by "$\lambda \geq 0$".

The following statement on saddle points can be proved similarly to that of Luenberger [15] at Theorem 2 of pp.221. The proof is omitted.

Theorem 2.1 *Suppose that there exists $\pi* \in \Pi$ and $\lambda^* \geq 0$ such that L^ν with $\nu \in P(S)$ possesses a saddle point ar π^*, λ^*. That is,*

$$L^\nu(\pi, \lambda^*) \leq L^\nu(\pi^*, \lambda^*) \leq L^\nu(\pi^*, \lambda) \tag{2.2}$$

for each $\pi \in \Pi$ and $\lambda \geq 0$. Then, π^* solves Problem A and is a ν-constrained regret optimal policy.

This theorem motivates us to obtain a sufficient condition for theexistence of a saddle point associated with Lagrangian L^ν. We need the following assumption.

Assumption 2.1 (i) *There exists M_1 and M_2 such that*

$$0 \le r(i,a) \le M_1 < \infty, \quad 0 \le d(i,a) \le M_2 < \infty,$$

for all $i \in S$, $a \in A$.

(ii) *There exist $L > 0$, $B > 0$ such that $L \le \int_0^\infty t F_{ij}(dt|a) \le B$ for all i, j.*

(iii) *Regret-utility functions $G, H_i (i = 1,2,\cdots,k)$ are all lower semicontinuous.*

Assumption 2.2

$$K := \sup_{\pi \in \Pi} E_\pi^\nu(N) < \infty$$

Now we give sufficient condition for Assumption 2.2 to hold. Define $e(n), n = 1,2,\cdots$ by $e(n) = \sup_{i \in S} e_i(n)$, where $e_i(n) = \sup_{\pi \in \Pi} P_\pi^\nu(N > n)$. Then, it holds (cf. [12]) that $e(n+1) \le e(n)$ and $e(n+m) \le e(n)e(m)$ for all $m, n = 1,2,\cdots$.

Proposition 2.1 *Each of the following condition* (i) *or* (ii) *implies to satisfy Assumption 2.2.*

(i) $\sum_{n=1}^\infty e(n) < \infty$.

(ii) *There exists $0 < \eta_0 < 1$ and $n_0 > 1$ such that $e(n_0) < 1 - \eta_0$.*

Proof. Similar calculations as in [12], (ii) $\Rightarrow$ (i) $\Rightarrow$ Assumtion 2.2. □

Let, for each $\nu \in P(S)$ and $\pi \in \Pi$, define a class $\Phi(\nu)$:

$$F_\pi^\nu(x,y) := P_\pi^\nu(\widetilde{\tau_N} \le x, \widetilde{D_N} \le y) \tag{2.3}$$

$$\Phi(\nu) := \left\{ F_\pi^\nu(\cdot,\cdot) \,\middle|\, \pi \in \Pi \right\} \tag{2.4}$$

Here, with some abuse of notation, we define

$$L^\nu(F,\lambda) := \int_0^\infty \int_0^\infty g_\lambda(x,y) F(dx,dy) \tag{2.5}$$

for any $F \in \Phi(\nu)$ and $\lambda \ge 0$, where

$$g_\lambda(x,y) := \sum_{j=1}^k \lambda_j (\alpha_j - H_j(h_j x, y)) - G(g^* x, y) \tag{2.6}$$

Then, Lagrangian L^ν definded in (2.1) is obiously rewritten by $L^\nu(\pi,\lambda) = L^\nu(F,\lambda)$ with $F = F_\pi^\nu$. Thus, we have the following corollary.

Corollary 2.1 *Let $\pi^* \in \Pi$ and $\lambda^* \ge 0$. $L^\nu(\cdot,\cdot)$ with $\nu \in P(S)$ possesses a saddle point at π^*, λ^* if and only if the relation holds*

$$L^\nu(F,\lambda^*) \le L^\nu(F_{\pi^*}^\nu, \lambda^*) \le L^\nu(F_{\pi^*}^\nu, \lambda) \tag{2.7}$$

for all $F \in \Phi(\nu)$ and $\lambda > 0$. Then, π^* solves Problem A *and is a ν-constrained regret optimal policy.*

Lemma 2.1 *For any $\nu \in P(S)$, it holds that*

(i) $\Phi(\nu)$ *is convex and compact it the weak topology;*

(ii) $L^\nu(\cdot,\lambda)$ *is concave and upper semi-continuous for each $\lambda \geq 0$;*

(iii) $L^\nu(F,\cdot)$ *is convex and continuous for each $F \in \Phi(\nu)$.*

Proof. By Assumption 2.1 and 2.2, we observe that

$$0 \leq E_\pi^\nu[\widetilde{\tau_N}] \leq BK \quad \text{and} \quad 0 \leq E_\pi^\nu[\widetilde{D_N}] \leq (M_1 + M_2 B)K$$

for all $\pi \in \Pi$. Also, $\widetilde{\tau_N}$ and $\widetilde{D_N}$ are continuous, i.e., they continuos functions of Ω, so the claim (i) follows from Lemma1.1. By using the Assumption 2.2, the claim (ii) holds. For (iii), it is immediate from the definition of (2.5) and (2.6). □

From Lemma 2.1, Fan's minimax theorem(cf.[4]) could be applied to obtain the following lemma.

Lemma 2.2 *It holds, for any $\nu \in P(S)$,*

$$\inf_{\lambda \geq 0} \max_{F \in \Phi(\nu)} L^\nu(F,\lambda) = \max_{F \in \Phi(\nu)} \inf_{\lambda \geq 0} L^\nu(F,\lambda) \tag{2.8}$$

Henthforth, the common value in the both side of (2.8) will be denoted simply by L^*. In order to prove the existence of a saddle point with (2.7), a variant of the well-known Slater Condition is imposed.

Slater Condition: There exists a $\overline{\pi} \in \Pi$ such that

$$E_{\overline{\pi}}^\nu[H_i(h_i\widetilde{\tau_N}, \widetilde{D_N})] < \alpha_i \tag{2.9}$$

for all $i (1 \leq i \leq k)$.

Since $L^\nu(F_{\overline{\pi}}^\nu, \lambda) \longrightarrow \infty$ as $\|\lambda\| \to \infty$ under condition(2.9), the convex function $\max_{F \in \Phi(\nu)} L^\nu(F,\lambda)$ is bounded from below, so that there exists $\lambda^* \geq 0$ such that

$$\max_{F \in \Phi(\nu)} L^\nu(F,\lambda^*) \leq L^* \tag{2.10}$$

by (2.8). On the other hand, by Lemma2.2, there exists $F^* \in \Phi(\nu)$ with

$$L^\nu(F^*,\lambda) \geq L^* \tag{2.11}$$

for all $\lambda \geq 0$. Thus, applying Corollary 2.1, the following main theorem have obtained.

Theorem 2.2 *Under Slater condition(2.9), Lagrangian $L^\nu(\cdot,\cdot)$ with $\nu \in P(S)$ has a saddle point, i.e., there exists $\pi^* \in \Pi$ and $\lambda^* \geq 0$ satisfying (2.2).*

Also, from Theorem 2.1 and 2.2, the following corollary holds.

Corollary 2.2 *Under Slater condition(2.9), there exists a ν-constraint optimal policy for $\nu \in P(S)$.*

3 Characterization of optimal policy

Now we will derive some theoretical results, which are usful to seek a constraint optimal policy. Firstly, letting $\nu \in P(S)$ and for each $\lambda \geq 0$, a plocy $\pi \in \Pi$ is said to be g_λ-optimal if

$$E^{\nu}_{\pi^*}[g_\lambda(\widetilde{\tau_N}, \widetilde{D_N})] \geq E^{\nu}_{\pi}[g_\lambda(\widetilde{\tau_N}, \widetilde{D_N})]$$

for all $\pi \in \Pi$, where g_λ is defind in (2.6).

The following Lemma can be easily proved as [13].

Lemma 3.1 *Let $\overline{\pi} \in \Pi$ and $\overline{\lambda} = (\overline{\lambda}_1, \overline{\lambda}_2, \cdots, \overline{\lambda}_k) \geq 0$. Then, Lagrangian $L^*(\cdot,\cdot)$ given in (2.1) has a saddle point at $\overline{\pi}, \overline{\lambda}$ if and only if the following (i) – (iii) holds:*

(i) *$\overline{\pi}$ is $g_{\overline{\lambda}}$-optimal;*

(ii) *$\overline{\pi} \in V(\nu, \alpha, h)$;*

(iii) $\sum_{i=1}^{k} \overline{\lambda}_i \left(\alpha_i - E^{\nu}_{\overline{\pi}}[H_i(h_i \widetilde{\tau_N}, \widetilde{D_N})] \right) = 0.$

For any Borel set X, we denote by $B(X)$ the set of all bounded Borel measurable functions on X. We define an operator $U_\lambda(d)(c_0, c, x \,|\, i, a)$ for $d = (d_i; i \in S)$ with $d_i \in B(R^{k+2}_+)$ provided that $c_0 \in R$, $c = (c_1, c_2, \cdots, c_k) \in R^k$, $x \in R$ and $i \in S, a \in A$.

$$\begin{aligned} & U_\lambda(d)(c_0, c, x \,|\, i, a) \\ = \; & \sum_{j \in J} p_{ij}(a) \int_0^\infty d_j\left(c_0 + g^* t, c + ht, x + r(i,a) + d(i,a)t\right) F_{ij}(dt|a) \\ & + \sum_{j \in J_0} p_{ij}(a) \int_0^\infty g_\lambda\left(c_0 + g^* t, c + ht, r(i,a) + d(i,a)t\right) F_{ij}(dt|a) \end{aligned} \tag{3.1}$$

where $J = S \setminus J_0$ and

$$g_\lambda(c_0, c, x) = \sum_{j=1}^{k} \lambda_j (\alpha_j - H_j(c_j, x)) - G(c_0, x),$$

$c_0 \in R$, $c + ht = (c_1 + h_1 t, c_2 + h_2 t, \cdots, c_k + h_k t) \in R^k$, $x \in R$.

Now we define an optimal value function starting from the initial state $i \in S$ and with $(c_0, c, x) \in R^{k+2}_+$ by

$$\begin{aligned} & g^\lambda_i(c_0, c, x) \\ := \; & \inf_{\pi \in \Pi} E^{\{i\}}_\pi \left[\sum_{j=1}^{k} \lambda_j \left(\alpha_j - H_j(c_j + h_j \widetilde{\tau_N}, x + \widetilde{D_N}) \right) - G(c_0 + g^* \widetilde{\tau_N}, x + \widetilde{D_N}) \right] \end{aligned} \tag{3.2}$$

for $i \in S$. Then we can prove the following by the similar method of Theorem 2.1 in [12].

Lemma 3.2 *For $\lambda \geq 0$, the set of optimal value functions $g^\lambda = \{g^\lambda_i; i \in S\}$ is given as a unique solution of the optimality equation;*

$$g^\lambda_i(c_0, c, x) = \min_{a \in A} U_\lambda(g^\lambda)(c_0, c, x \,|\, i, a) \tag{3.3}$$

for all $i \in S$ and $(c_o, c, x) \in R^{k+2}_+$.

In order to determine an optimal policy, we define the set of λ-optimal actions, that is, $A^{\lambda}(c_0, c, x \mid i)$, by

$$A^{\lambda}(c_0, c, x \mid i) := \arg\min_{a \in A} U_{\lambda}(g^{\lambda})(c_0, c, x \mid i, a),$$

where $g^{\lambda} = (g_i^{\lambda}; i \in S)$ is a unique solution of (3.3). Then we have obtained the next theorem.

Theorem 3.1 *For any $\nu \in P(S)$, a policy $\pi \in V(\nu, \alpha, h)$ is a constrained optimal policy if and only if there exists $\lambda^* \geq 0$ such that*

(i) $P_{\pi^*}^{\nu}\left(\Delta_t \in A^{\lambda^*}(g^* \widetilde{\tau}_t, h\widetilde{\tau}_t, \widetilde{D_t} \mid X_t)\right) = 1 \quad$ *for all* $\quad t \geq 0$;

(ii) $\sum_{i=1}^{k} \lambda_i^* \left(\alpha_i - E_{\pi^*}^{\nu}\left[H_i(h_i \widetilde{\tau_N}, \widetilde{D_N}\right]\right) = 0$

Proof. Applying the results of Theorem 2.1 in [12], it can be shown that π^* is g_{λ^*}-optimal if and only if the above (i) holds. So this theorem could be proved by using Lemma 3.1. □

References

[1] Altman, E.; *Constrained Markov Decision Processes*, Chapman & Hall/CRC, 1999.

[2] Avriel, M.; *Nonlinear Programming, Analysis and Methods*, Prentice Hall, Inc., 1976.

[3] Borkar, V. S.; *Topics in Controlled Markov Chains*, Longman Scientific Technical, 1991.

[4] Borwein, J.M. and Zhuang, D.; On Fan's minimax theorem, *Math. Programming*, **34**, 232–244, 1986.

[5] Chung, K. J. and Sobel, M. J.; Discounted MDP's: Distribution functions and exponential utility maximization, *SIAM J. Control Optim.* **25**, 49-62, 1987.

[6] Denardo, E.V. and Rothblum, U.G.; Optimal stopping, exponential utility and linear programming, *Math. Prog.* **16**, 228–244, 1979.

[7] Fishburn, P.C.; Utility Theory for Decision Making, John Wiley & Sons, New York, 1970.

[8] Hinderer, K. and Waldmann, K.H. (2003): The critical discount factor for finite Markovian decision processes with an absorbing set. *Math. Mech. Oper. Res.* **57**: 1–19

[9] Howard, R.S. and Matheson, J.E.; Risk-sensitive Markov decision processes, *Manag. Sci.*, **8** (1972), 356–369.

[10] Kadota, Y., Kurano, M. and Yasuda, M.; Discounted Markov decision processes with general utility, In *Proceeding of APORS' 94*, 330–337, World Scientific, 1995.

[11] Kadota, Y., Kurano, M. and Yasuda, M.; On the general utility of discounted Markov decision processes, *Int. Trans. Opl Res.* **5**(1), 27–34, 1998.

[12] Kadota, Y., Kurano, M. and Yasuda, M.; Regret optimality in semi-Markov decision processes with an absorbing set, *Proceeding, The Sixth International COnference on Optimization: Techniques and Applications(ICOTA6)* Paper No.44, 1–14, 2004.

[13] Kadota, Y., Kurano, M. and Yasuda, M.; discounted Marokov decision processes with utility constraints, To appear in *Computers and Mathematics with Application*.

[14] Lippman, S.A.; Maximal average reward policies for Semi-Markov decision processes with arbitrary state and action space. *Ann. Math. Statist.*, **42** (1971), 1717–1726.

[15] Luenberger, D.; *Optimization by Vector Space Methods*, John Wiley, New York, 1969.

[16] Pratt, J.W.; Risk aversion in the small and in the large, *Econometrica*, **32** (1964), 122–136.

[17] Ross, S.M.; Applied Probability Models with Optimization Applications, Holden-Day, San Francisco, 1970.

[18] Sennot, L. I.; Constrained discounted Markov decision chains, *Probability in the Engineering and Information Sciences*, **5**, 463–475, 1991.