Reprinted from IEEE TRANSACTIONS
ON *AUTOMATIC CONTROL*
Volume AC-12, Number 6, December, 1967
pp. 779-780

Control with Random Stopping

This correspondence relates control with random terminal time to reliability theory and to deterministic optimal control theory. The Hamilton–Jacobi–Bellman equation for random termination control stated here extends the linear quadratic case analyzed in Sivan[1] to general system dynamics, loss functions, and termination probabilities. Previous papers on this subject[2,3] discuss physical motivation. Random stopping control models are useful when estimates of plant parameters are improved, or the plant itself changes at random times. Queuing for a digital computer used in controlling several systems leads to the first case; component failures and catalyst deterioration exemplify the second.

If stopping time T is a positive random variable distributed as $F(\tau)=\Pr[T\leq\tau]$, the expectation of the integral portion of a cost functional of the control process

$$G = \int_0 g dt$$

where $g=g(t, \mathbf{x}, \mathbf{u})$, evaluated for any control $\mathbf{u}(t)$ and corresponding state vector $\mathbf{x}(t)$, is

$$E\{G\} = \int_0^\infty \int_0^T g dt dF(\tau).$$

The following lemma (an easy consequence of integration by parts) reduces this to a single integral.

Lemma

If T is a positive random variable distributed as $F(\tau)=\Pr[T\leq\tau]$, then

$$E\left\{ \int_0^T g(t, \mathbf{x}, \mathbf{u}) dt \right\} = \int_0^\infty g(t, \mathbf{x}, \mathbf{u})[1 - F(t)]dt. \quad (1)$$

The resulting infinite integral has the original integrand weighted by $1-F(t) =\Pr[T>t]$, the probability of nontermination by time t or "survival probability."

The lemma enables the optimization of the expected cost by standard procedures for any criterion, dynamic system, constraints, and probability of termination. Every procedure leads to equations which can be most simply stated in terms of $\nu(t)$, the conditional rate of failure function of reliability theory.

Definition

The conditional rate of failure function $\nu(t)$ has the property that, to terms of first order in dt,

$$\nu(t)dt = \Pr[t < T \leq t + dt \mid T > t] = \frac{dF(t)}{1 - F(t)}. \quad (2)$$

As (2) implies, $\nu(t)$ completely characterizes any continuous probability law. Density, distribution, and conditional rate of failure are all equivalent: any two can easily be expressed in terms of the third. See Klinger[2] or Papoulis,[4] pp. 109–111 and 246–248, for a general reference; Klinger[2] contains a table of $\nu(t)$ for several continuous probability laws.

Manuscript received July 24, 1967. This work was included in a Ph.D. dissertation, University of California, Berkeley, August 1966.

[1] R. Sivan, "The optimal control of linear systems with unknown final time," *IEEE Trans. Automatic Control* (*Short Papers*), vol. AC-11, pp. 529–530, July 1966.

[2] A. Klinger, "Continuous control with stochastic stopping time," RAND Corp., Santa Monica, Calif., RM-4993-PR, June 1966; also Preprints Joint Automatic Control Conf. (Philadelphia, Pa., June 1967), pp. 626–632.

[3] ——, "Identification and control of a randomly observed system," Ph.D. dissertation, University of California, Berkeley, August 1966.

[4] A. Papoulis, *Probability, Random Variables, and Stochastic Processes.* New York: McGraw-Hill, 1965.

The Hamilton–Jacobi–Bellman equation illustrates the role of conditional rate of failure. Consider a general system with dynamics

$$\dot{\mathbf{x}} = h(t, \mathbf{x}, u), \tag{3}$$

random terminal time, and minimum expected cost function $S(\mathbf{x}, t)$ given by

$$S(\mathbf{x}, t) = \min_{u} \left[E\left\{ \int_{t}^{T} g dt \,\middle|\, \mathbf{x}(t), \; t < T \sim \nu(t) \right\} \right]. \tag{4}$$

A standard argument involving the definition of expectation, principle of optimality, Taylor series expansion, and limits as $\Delta t \to 0$, yields

$$\frac{\partial S}{\partial t} - \nu(t) S + \min_{u} \left[\left\langle \frac{\partial S}{\partial \mathbf{x}}, h(t, \mathbf{x}, u) \right\rangle + g(t, \mathbf{x}, u) \right] = 0. \tag{5}$$

In (4) $\sim$ denotes "distributed as"; $\partial S/\partial \mathbf{x}$ in (5) is the row vector "gradient of S with respect to $\mathbf{x}$."

Note that the equivalent deterministic-terminal-time equation is (5) with $\nu(t) = 0$. Likewise, when $\nu(t) = 0$, the feedback Riccati equation of Klinger[2] [eq. (33); linear time-varying system, quadratic cost, stochastic stopping] corresponds to that obtained for control on a deterministic interval $[0, T]$ (see, e.g., Athans and Falb[5] p. 763).

These remarks and Theorems 1 and 2 suggest that control over a deterministic interval $[0, \overline{T}]$ with no terminal state cost is a limiting case, as variance approaches zero, of control for random stopping time T with expectation criterion, where $\overline{T} = E[T]$. Theorem 1 states the conditions for this to hold for the cost criteria; the proof is by Lebesgue's theorem on dominated convergence (see, e.g., Natanson,[6] p. 161).

Theorem 1

If T is a positive random variable with probability distribution $F_\sigma(t)$ of mean $\overline{T} < \infty$ and variance $\sigma^2 < \infty$, and $g(t) = g(t, \mathbf{x}(t), u(t))$ is non-negative and such that $\int_0^\infty g(t)/(1+t^2)\, dt < \infty$ for all $\mathbf{x}(t)$, $u(t)$ on $[0, \infty)$, then

$$\lim_{\sigma^2 \to 0} E\left\{ \int_0^T g(t, \mathbf{x}, u) dt \right\} = \int_0^{\overline{T}} g(t, \mathbf{x}, u) dt. \tag{6}$$

Hence the control optimal for the deterministic interval, u^o, where the superscript o denotes "optimal" is that for zero variance. The usefulness of the stochastic-stopping optimal control as an approximation to u^o for σ small is established by Theorem 2, which follows from a continuity result in differential equations (see, e.g., Birkhoff and Rota,[7] p. 107, corollary).

Theorem 2

If the distributions F_σ are continuous functions of variance σ^2, a unique optimal control u_σ^o exists for the criterion

$$E\left\{ \int_0^T g(t, \mathbf{x}, u) dt \,\middle|\, T \sim F_\sigma \right\} \tag{7}$$

for all $\sigma \geq 0$, and the functions Z_i of the canonical Hamiltonian equations for $z' = [\mathbf{x}', p']$, which follow from elimination of u via Pontryagin's maximum principle,

$$\dot{z} = Z_i(z, t, F_\sigma), \qquad t_i \leq t < t_{i+1}, \quad i = 1, 2, \cdots \tag{8}$$

in the limit satisfy Lipschitz conditions

$$|Z_i(z_1, t) - Z_i(z_2, t)| \leq L_i |z_1 - z_2|, \qquad Z_i(z, t) = \lim_{\sigma^2 \to 0} Z_i(z, t, F_\sigma) \tag{9}$$

for some finite constants L_i, $i = 1, 2, \cdots$, then

$$\lim_{\sigma^2 \to 0} u_\sigma^o = u^o. \tag{10}$$

Proofs of these theorems appear in Klinger.[3]

For some probability distributions, random stopping optimal controls can yield unbounded trajectories if the interruption is delayed until very large time values (an improbable event). For example, the time-invariant scalar linear quadratic exponential probability law case [$h(t, x, u) = ax + u$, $g(t, x, u) = x^2 + u^2$, $\nu(t) = \nu$ constant; see Klinger[2]] has a finite optimal trajectory as time approaches infinity *only if*

$$E[T] = \frac{1}{\nu} \geq \frac{a}{1 + a^2}.$$

If a is negative, the contrary inequality never holds; if a is positive and the contrary is true, the probability of the event $T \to \infty$ will be small: $E[T]$ will be bounded so the exponential distribution will assign almost all mass to the event T bounded. Thus instability of randomly stopped optimum trajectories should not occur in practice.

Allen Klinger
RAND Corporation
Santa Monica, Calif. 90406

[5] M. Athans and P. L. Falb, *Optimal Control.* New York: McGraw-Hill, 1966.
[6] I. P. Natanson, *Theory of Functions of a Real Variable,* vol. 1. New York: Ungar, 1961.
[7] G. Birkhoff and G. C. Rota, *Ordinary Differential Equations.* New York: Ginn, 1962.