

Optimal Risk Probability for First Passage Models

| in Semi-Markov Decision Processes

Xianping Guo (Coauthor: Yonghui Huang)
(Zhongshan University, Guangzhou)

Peking University, July 2010

Outline

- 2 Motivation
- 2 Semi-Markov decision processes
- 2 Optimality problems
- 2 Optimality equation
- 2 Existence of optimal policies
- 2 Numerable examples

1. Motivation

Background: Reliability engineering, and risk analysis

Problem: $\sup_{\pi} P_i^{\pi}(\tau_B > \epsilon)$,

τ_B is an initial state

π is a policy

B is a given target set

τ_B is a first passage time to B

ϵ is a threshold value.

2. Semi-Markov Decision Processes

The model of SMDP:

$$fS; B; (A(i); i \in S); Q(t; j|i; a)g$$

where

$\in S$: the state space, a denumerable set;

$\in B$: a given target set, a subset of S ;

$\in A(i)$: finite set of actions available at $i \in S$;

$\in Q(t; j|i; a)$: semi-Markov kernel, $a \in A(i); i, j \in S$;

Notation:

π **Policy** π : A sequence $\pi = \{f_n; n = 0, 1, \dots, g\}$ of stochastic kernels f_n on the action space A given H_n satisfying

$$f_n(A(i_n) | (0; i_0; \dots; 0; a_0; \dots; t_{n-1}; i_{n-1}; \dots; a_{n-1}; t_n; i_n)) = 1$$

π **Stationary policy**: measurable $f, f(i; \cdot) \subseteq A(i)$ for all $(i; \cdot)$

π $P_{(i; \cdot)}^\pi$: Probability measure on $(S \times [0, 1]) \times (\prod_{i \in S} A(i))$

π $S_n; J_n; A_n$: n -th decision epoch, the state and action at the S_n , respectively.

Assumption A. There exist $\epsilon > 0$ and $\delta > 0$ such that

$$\sum_{j \in S} Q(\pm; j | i; a) \cdot \mathbf{1}_{|j-i| \leq \delta}; \text{ for all } (i; a) \in K:$$

Assumption A) $P_{(i; s)}^{\frac{1}{4}}(fS_1 = \mathbf{1} g) = 1$

Semi-Markov decision process $f(Z(t); A(t); t, 0 g):$

$$Z(t) = J_n; A(t) = A_n; \text{ for } S_n \leq t < S_{n+1}$$

The first passage time into B , is defined by

$$\tau_B := \inf \{ t \geq 0 \mid Z(t) \in B g; \text{ (with } \inf ; := \mathbf{1});$$

3. Optimality Problems

The risk probability:

$$F^{\frac{1}{4}}(i; s) := P_{(i; s)}^{\frac{1}{4}}(\mathcal{C}_B \cdot s)$$

The optimal value:

$$F_{\alpha}(i; s) := \inf_{\frac{1}{4}2\Pi} F^{\frac{1}{4}}(i; s);$$

Definition 1. A policy $\frac{1}{4}^{\alpha} 2 \mid$ is called optimal if

$$F^{\frac{1}{4}^{\alpha}}(i; s) = F_{\alpha}(i; s) \quad \forall (i; s) \in S \in R:$$

² Existence and computation of optimal policies ???

4. Optimality Equation

For $i \in B^c$; $a \in A(i)$, and $s \geq 0$, let

$$T^a u(i; s) := Q(s; Bji; a) + \sum_{j \in B^c} \int_0^s Q(dt; jji; a) u(j; s - t);$$

with $u \in F_{[0,1]}$ (the set of measurable functions $0 \leq u \leq 1$),

$$Q(s; Bji; a) := \sum_{j \in B} Q(s; jji; a); \quad T^a u(i; s) := 0 \text{ for } s < 0;$$

Then, define operators T and T^f :

$$Tu(i; s) := \min_{a \in A(i)} T^a u(i; s); \quad T^f u(i; s) := T^{f(i; s)} u(i; s);$$

for each stationary policy f .

Theorem 1. Let Under Assumption A, we have

(a) $F^f = \lim_{n \rightarrow \infty} u_n^f$, where $u_n^f := T^f u_{n-1}^f; u_{-1}^f := 1;$

(b) F^f satisfied the equation, $u = T^f u$, for all $f \in F;$

2 Theorem 1 gives an approximation of risk probability F^f .

For each $(i; s) \in B^c \in R_+$ and $\frac{1}{4} \in \mathbb{I}$, let

$$F_{i-1}^{\frac{1}{4}}(i; s) := 1;$$

$$F_n^{\frac{1}{4}}(i; s) := 1 - \sum_{m=0}^n P_{(i; s)}^{\frac{1}{4}}(S_m \cdot s < S_{m+1}; J_k \in B^c; 0 \cdot k \cdot m)$$

Theorem 2. Let $F_n^\alpha(i; s) := \inf_{\frac{1}{4}} F_n^{\frac{1}{4}}(i; s)$, then

(a) $F_{n+1}^\alpha = TF_n^\alpha$ for all $n \geq j - 1$, and $\lim_{n \rightarrow \infty} F_n^\alpha = F_\alpha$.

(b) F_α satisfies the **optimality equation**: $F_\alpha = TF_\alpha$.

(c) F_α is the maximal fixed point of T in $F_{[0;1]}$.

Remark 1.

\Rightarrow Theorem 2(a) gives a **value iteration algorithm** for computing the optimal value function F_α .

\Rightarrow Theorem 2(b) establishes the **optimality equation**.

5. Existence of Optimality Policies

To ensure the existence of optimal policies, we introduce the following condition.

Assumption B. For every $(i, s) \in B^c \times R$ and f ,

$$P_{(i,s)}^f(\zeta_B < 1) = 1:$$

To verify Assumption B, we have a fact below:

Theorem 3. If there exists a constant $\delta > 0$ such that

$$\sum_{j \in B} Q(1; j, i; a) \geq \delta \text{ for all } i \in B^c;$$

Theorem 4. Under Assumptions A and B, we have

- (a) F^f and F_α are the unique solution in $F_{[0;1]}$ to equations $u = T^f u$ and $u = Tu$, respectively;
- (b) any f , such that $F_\alpha = T^f F_\alpha$, is optimal;
- (c) there exists a stationary policy f^α satisfying the optimality equation: $F_\alpha = TF_\alpha = T^{f^\alpha} F_\alpha$; and such policy f^α is optimal.

Remark 2.

² Theorem 4(c) shows the existence of an optimal policy.

To give the existence of special optimal policies, let

$$A^\alpha(i; \delta) := \{f \in \mathcal{F} : \sum_{j \in A(i)} f_j F^\alpha(i; \delta) = T^\alpha F^\alpha(i; \delta)g\}$$

$$A^\alpha(i) := \bigcap_{\delta > 0} A^\alpha(i; \delta)$$

Theorem 5. If $\sup_{i \in B^c} \sup_{a \in A(i)} Q(t; B^c | i; a) < 1$ for some $t > 0$, and Assumptions A and B hold, then,

- (a) for any $g \in G := \{fg | g(i) \in A(i) \forall i \in S\}$, F^g is the unique solution in $F_{[0,1]}$ to the equation: $u = T^g u$;
- (b) there exists an optimal policy $f \in G$ if and only if $A^\alpha(i) \neq \emptyset$; for all $i \in B^c$.

5. Numerable examples

Example 5.1. Let $S = \{1, 2, 3\}$, $B = \{3\}$, where

state 1: the good state

state 2: the medium state

state 3: the failure state

Let $A(1) = \{a_{11}, a_{12}\}$; $A(2) = \{a_{21}, a_{22}\}$; $A(3) = \{a_{31}\}$.

The semi-Markov kernel is of the form:

$$Q(t; j | i; a) = H(t | i; a) p(j | i; a)$$

$H(t; j, i; a)$: the distribution functions of the sojourn time

$p(j; j, i; a)$: the transition probabilities.

$$H(t; j=1; a_{11}) := \begin{cases} 1 - e^{-0.25t}; & t \in [0; 25]; \\ 1; & t > 25; \end{cases}$$

$$H(t; j=2; a_{21}) := \begin{cases} 1 - e^{-0.20t}; & t \in [0; 20]; \\ 1; & t > 20; \end{cases}$$

$$H(t; j=3; a_{31}) := 1 - e^{-0.2t};$$

$$H(t; j=1; a_{12}) = 1 - e^{-0.08t};$$

$$H(t; j=2; a_{22}) = 1 - e^{-0.15t};$$

$$\begin{aligned}
p(1 \ j \ 1; a_{11}) &= 0; & p(2 \ j \ 1; a_{11}) &= \frac{9}{20}; & p(3 \ j \ 1; a_{11}) &= \frac{11}{20}; \\
p(1 \ j \ 1; a_{12}) &= 0; & p(2 \ j \ 1; a_{12}) &= \frac{1}{2}; & p(3 \ j \ 1; a_{12}) &= \frac{1}{2}; \\
p(1 \ j \ 2; a_{21}) &= \frac{1}{5}; & p(2 \ j \ 2; a_{21}) &= 0; & p(3 \ j \ 2; a_{21}) &= \frac{4}{5}; \\
p(1 \ j \ 2; a_{22}) &= \frac{1}{4}; & p(2 \ j \ 2; a_{22}) &= 0; & p(3 \ j \ 2; a_{22}) &= \frac{3}{4}; \\
p(3 \ j \ 3; a_{31}) &= 1;
\end{aligned}$$

Using the **value iteration algorithm** in Theorem 2, we obtain some computational results as in Figure 1 and Figure 2.

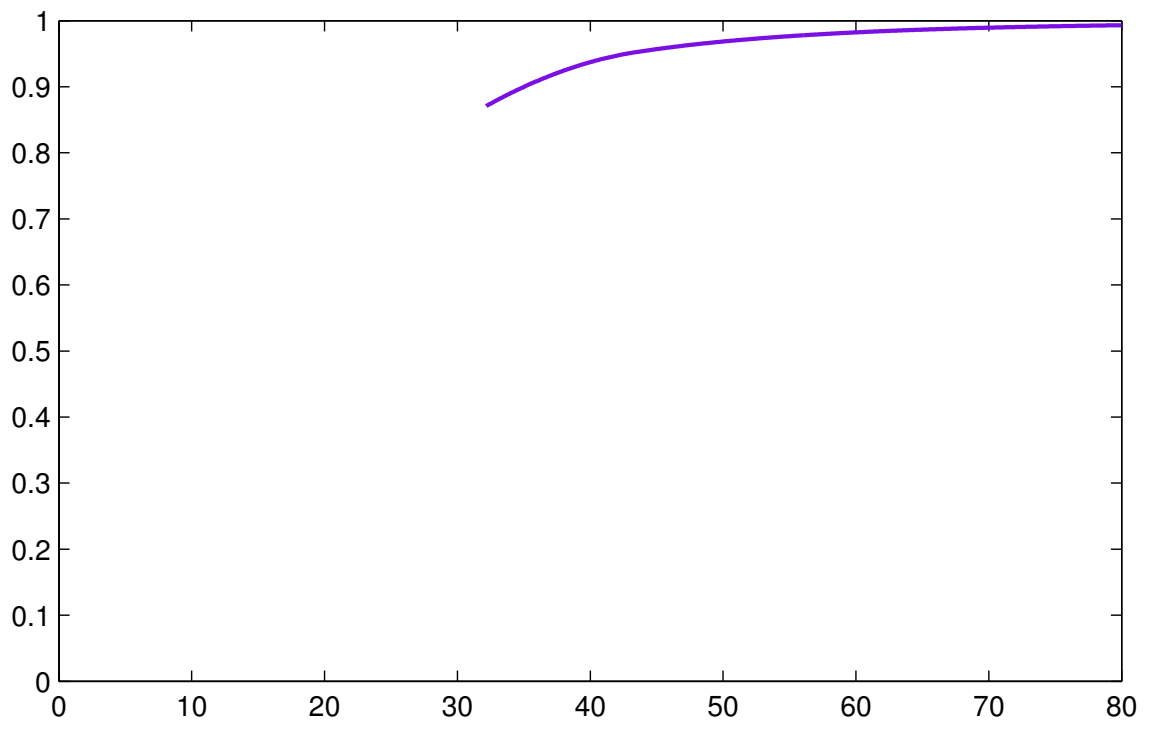


Figure 1. The functions $T^a F^a(i; s)$

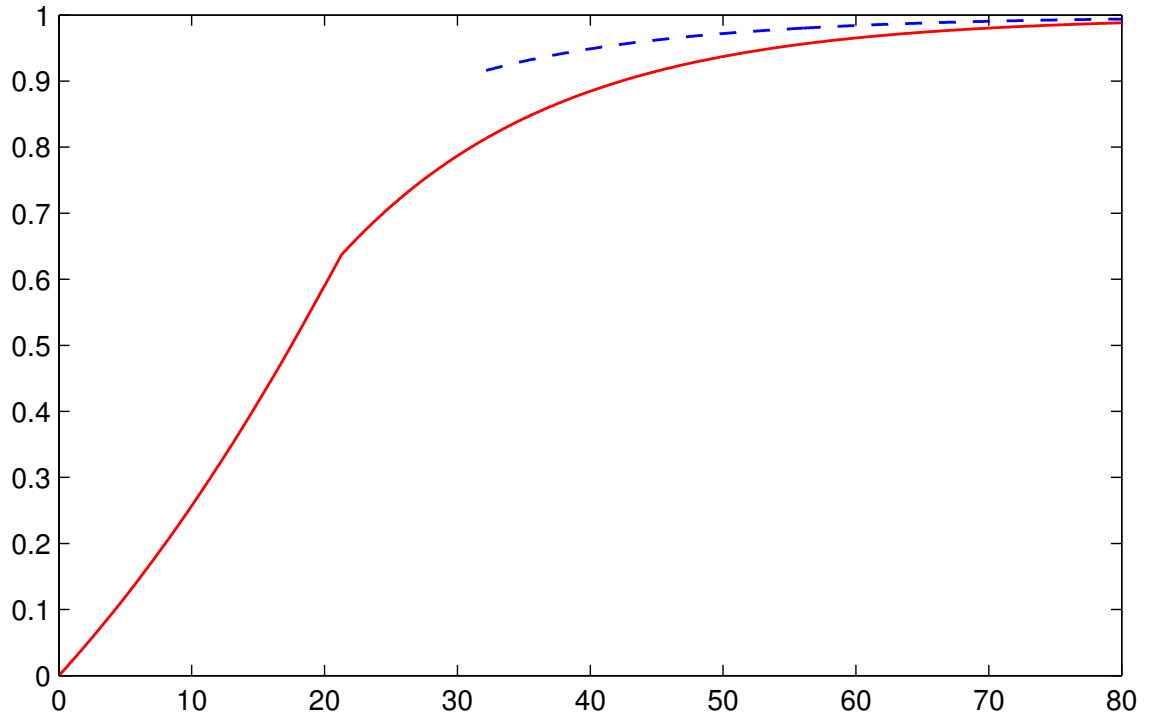


Figure 2. The value function $F^a(i; s)$

More clearly, we have

$$F^{\alpha}(1; \varsigma) = \begin{cases} T^{a_{11}} F^{\alpha}(1; \varsigma); & 0 \cdot \varsigma < 21:36; \\ T^{a_{11}} F^{\alpha}(1; \varsigma) = T^{a_{12}} F^{\alpha}(1; \varsigma); & \varsigma = 21:36; \\ T^{a_{12}} F^{\alpha}(1; \varsigma); & 21:36 < \varsigma < 29:3; \\ T^{a_{11}} F^{\alpha}(1; \varsigma) = T^{a_{12}} F^{\alpha}(1; \varsigma); & \varsigma = 29:3; \\ T^{a_{11}} F^{\alpha}(1; \varsigma) (= 0:7742); & \varsigma > 29:3; \end{cases}$$

$$F^{\alpha}(2; \varsigma) = \begin{cases} T^{a_{21}} F^{\alpha}(2; \varsigma); & 0 \cdot \varsigma < 18:54; \\ T^{a_{21}} F^{\alpha}(2; \varsigma) = T^{a_{22}} F^{\alpha}(2; \varsigma); & \varsigma = 18:54; \\ T^{a_{22}} F^{\alpha}(2; \varsigma); & 18:54 < \varsigma < 23:82; \\ T^{a_{21}} F^{\alpha}(2; \varsigma) = T^{a_{22}} F^{\alpha}(2; \varsigma); & \varsigma = 23:82; \\ T^{a_{21}} F^{\alpha}(2; \varsigma) (= 0:8542); & \varsigma > 23:82; \end{cases}$$

Define a policy f^α by

$$f^\alpha(1; s) = \begin{cases} a_{11}; & 0 \leq s \leq 21.36; \\ a_{12}; & 21.36 < s \leq 29.3; \\ a_{11}; & s > 29.3; \end{cases}$$

$$f^\alpha(2; s) = \begin{cases} a_{21}; & 0 \leq s \leq 18.54; \\ a_{22}; & 18.54 < s \leq 23.82; \\ a_{21}; & s > 23.82; \end{cases}$$

Then, we have

$${}^2 F^\alpha(i; s) = T^{f^\alpha} F^\alpha(i; s) \text{ for } i = 1, 2 \text{ and all } s \geq 0,$$

${}^2 f^\alpha$ is an optimal stationary policy.

$$A^{\alpha}(1; \varsigma) = \begin{cases} fa_{11}g; & 0 \cdot \varsigma < 21:36; \\ fa_{11}; a_{12}g; & \varsigma = 21:36; \\ fa_{12}g; & 21:36 < \varsigma < 29:3; \\ fa_{11}; a_{12}g; & \varsigma = 29:3; \\ fa_{11}g; & \varsigma > 29:3; \end{cases}$$

$$A^{\alpha}(2; \varsigma) = \begin{cases} fa_{21}g; & 0 \cdot \varsigma < 18:54; \\ fa_{21}; a_{22}g; & \varsigma = 18:54; \\ fa_{22}g; & 18:54 < \varsigma < 23:82; \\ fa_{21}; a_{22}g; & \varsigma = 23:82; \\ fa_{21}g; & \varsigma > 23:82; \end{cases}$$

Hence,

$$A^\alpha(1) = \bigcap_{s \geq 0} A^\alpha(1; s) = ; ; A^\alpha(2) = \bigcap_{s \geq 0} A^\alpha(2; s) = ; ;$$

which show there is no optimal policy in G .

Remark 3. This shows that the assumption in the previous literature is not satisfied for this example !!!

Example 5.2. Let $S = f1;2g$, $B = f2g$;

$$A(1) = fa_{11};a_{12}g; A(2) = fa_{21}g;$$

$Q(t;j j i; a)$ is given by

$$Q(t;j j 1; a_{11}) = \begin{cases} 1=2; & \text{if } t_s 1;j = 1;2; \\ 0; & \text{otherwise;} \end{cases}$$

$$Q(t;j j 1; a_{12}) = \begin{cases} 1; & \text{if } t_s 2;j = 2; \\ 0; & \text{otherwise;} \end{cases}$$

$$Q(t;j j 2; a_{21}) = \begin{cases} 1 j e^i t; & \text{if } t_s 0;j = 2; \\ 0; & \text{otherwise;} \end{cases}$$

Assumptions A and B holds in this example.

We now define a policy d as follows:

$$d(1; s) = \begin{cases} a_{12}; & 0 \leq s < 2; \\ a_{11}; & s \geq 2; \end{cases}$$

Then, by Theorem 1, we have $F^d(1; s) = \lim_{n \rightarrow \infty} F_n^d(1; s)$, which yields

$$F^d(1; s) = \begin{cases} 0; & 0 \leq s < 2; \\ 1; & s = 2; \\ 1=2; & 2 < s < 3; \end{cases}$$

Hence, $F^d(1; s)$ is not a distribution function of s .

Many Thanks !!!