

# 部分観測マルコフ決定過程と検診における費用 効果分析 (Partially observable Markov decision Processes and cost-effectiveness in health screening)

大阪電気通信大学・情報通信工学部・阪口昌彦

Masahiko SAKAGUCHI<sup>(1)</sup>

上智大学・経済学部・來島愛子

Aiko KURUSHIMA<sup>(2)</sup>

神奈川大学・理学部・堀口正之

Masayuki HORIGUCHI<sup>(3)</sup>

<sup>(1)</sup>Associate Professor of Department of Engineering Informatics, Faculty of Information and Communication Engineering, Osaka Electro-Communication University,

<sup>(2)</sup>Professor of Department of Economics, Faculty of Economics, Sophia University

<sup>(3)</sup>Professor of Department of Mathematics, Faculty of Science, Kanagawa University

## Abstract

本研究では、ヘルスケアプログラムとしてのスクリーニング(検診)プログラムの評価手法において、マルコフ決定過程による経済的評価の数理モデルの構築について考察する。具体的には、前立腺がん(prostate cancer)のスクリーニングにおける数理モデルからマルコフ決定過程を構成した場合の、状態空間、決定空間、推移法則、推移核およびベイズ更新と最適方程式について概観し、マルコフモデルによる費用効果分析の意義について議論する。また、費用効用解析としてのQALYs(質調整生存年)とマルコフモデル、費用効果分析との関係について述べる。

## 1 はじめに

QOL(生活の質, Quality of life)については, 医療や疾病に関わる具体的状況を要因とする評価方法のほかに, 仕事での満足度, 日常の生きがいなど, 個人の評価基準に大きく影響を受け定量化される評価尺度も用いられ, 多様な観点から総合的に「生活の質」を考えているといえる. 「生活の質」を保つ方法は, 個人によっても異なるだろうし, 満足度のように主観的な評価値が数値化されているものについては, 他者あるいは集団間の比較を行う場合に客観的に数値化されたデータとのペア(多次元データ)によって, 「生活の質」の数値化された総合的な評価基準が構築できると考える. 数値化によって, 個人の活動のあり方の改善や, 個々の状況での他の選択肢との比較も可能になると考えられる.

本研究では, マルコフ決定モデルとしての定式化を試み, 検診計画(screening planning)に対する考察, 提案を行うためのマルコフモデルの構成要素の状態空間, 決定空間, 推移法則, 推移核およびベイズ更新と最適方程式について概観し, マルコフモデルによる費用効果分析の意義について議論する.

## 2 Cost-effectiveness analysis

検診・診断・治療計画の指針策定研究として, 次の取り組みを見ていく.

Heijnsdijkらによる研究([5],[6])では, QALYs(Quality-Adjusted Life-Years)に関連する検診プログラム(Hearth Care Programmes)および診断・治療のアセスメントと有効性, 経済的評価をヨーロッパにおける前立腺がんに関するランダム化研究のデータを基礎として, 検診と陽性者への診断, 治療のそれぞれでのQALYsと検診・診断・治療措置等の各項目での効用推定値の積算によって, PSA (Prostate-specific antigen screening) の受診間隔や前立腺がんによる死亡数の減少, 過剰な診断の回避を目的としている.

例えば, [5]では, 生存年(Life-years)と質調整生存年(QALYs)の比較において, 数値シミュレーションによる1000人当たりの効用推定値の積算比較において, 55歳から70歳まで毎年のスクリーニングによって得られる生存年数とQALYsとの比(減少率)は, 4年間隔でのスクリーニングによる(1000人当たりの)52年の余命と41年のQALYsの比においても達成できることから, 長期間での経済的評価モデルにおいては, 過剰な診断の回避とPSAの費用効果の観点において, 4年間隔でのスクリーニングも有効であると考えられることができることを指摘している. ただし, ERSPC(European Randomized Study of Screening for Prostate cancer)におけるデータにもとづいた生存年数とQALYsの積算比で見ると, QOLには生存年数に対して23%のネガティブインパクトがあるとも報告しており, 前立腺がんの診断と治療の過程においては, 損失効用値のQOLへの数値の影響が少なくないことも示されている.

さらに, [6]においては, 検診・診断・治療のそれぞれにおける費用と健康状態, 推定された効用値とその対象期間に基づいて, 効用値の積算評価モデルとしてのマイクロシミュレーションによる分析および, 検診時間間隔, 受診・見受診群のコホート研究を行っている. 検診開始年齢, 検診期間間隔などについて, シナリオごとに得られるQALYsに対して, 期間中の総費用と比をとることによって, 得られる1年間のQALYsあたりでの費用効果分析を行っている. おおよそ, 2, 3年の間隔での検診の

受診, 治療後の65~72才のころでの検診参加の終了が費用効果分析における経済的評価が得られることを示している. しかしながら, 例えば, 治療診断の方針や, 得られているデータ値の地域特性の影響の有無などを考慮すれば, 費用効果分析における数理モデルの構築と, 評価関数による感度分析や, モデルを構成するパラメータ推定の手法が入手可能なデータセットからのものであることなどが求められ, また, 結果の分析においても多様な評価軸からの議論が必要であることは, 言うまでもない.

近年, 堀口([9])による部分観測マルコフ決定のもとでの乳がん検診プログラムのシナリオ評価の研究がある. 日本国内での, 罹患率, 検診結果の各種データ値, 診断後の生存率などを, 部分観測マルコフ決定過程における累積死亡リスクを評価関数として, 検診開始時期や検診受診期間を, 国内での推奨検診プログラムに準じたシナリオの優位性について検証をした. その研究で扱われたシナリオ評価では, 1年毎の推移確率行列として, 区分的(5年毎)に推移確率行列が変化する非定常マルコフ連鎖のもと, 検診群と非検診群の死亡リスクの比較を行っている. 推奨される年齢期間(40歳以降)での定期受診が, 想定した比較対象のシナリオにおいて一番死亡リスクを下げられるという結果が示されている.

経済的評価 (Economic Evaluation) を多段決定モデルにおいて行うときには, 想定される状態空間, 決定空間, 推移法則内での最適性が議論されているということに注意が必要である. 仮に, その数理モデルの最適化問題に最適解が存在し, 最適解が得られたとして, その最適解は, 数理モデル内での最適解であることを忘れてはならない. PSAの検診プログラムにおいても, 例えば[12]のモデルでは, 前立腺切除(RP, radical prostatectomy)を状態変数の一つに含み, 前立腺の治療過程における部分観測マルコフ決定過程の構築と分析を行っており, 数理モデルの構造は, 治療段階での医師の診療方針と患者の個別状況, 推奨される診療方針などによっても少なからず差異が生じることもある. また, [12]の研究においても言及されているが, 別の診断治療計画としてのPSA監視療法(active surveillance)などと, 部分観測マルコフ決定モデルの結果が直接比較可能であるとは言い難い. それぞれの診断治療計画でのリスク評価, 経済的評価は, かならずしも個人の不確実性に関して, 適切 (最適) な診断治療計画が示されるものではないが, 本研究での数理モデルとその評価手法を構築は, シミュレーションや感度分析をふくむモデル解析が可能になり, マイクロシミュレーションによる個別事例の発見, 分析につながるものと考えている.

### 3 Partially observed Markov decision model

この多段決定システムの状態空間 $S$ は, 次のように表すことができる.

$$\begin{aligned}
 S &= \{0(\text{no prostate cancer}), \\
 &\quad 1(T1 - G < 7), 2(T1 - G = 7), 3(T1 - G > 7) \\
 &\quad 4(T2 - G < 7), 5(T2 - G = 7), 6(T2 - G > 7) \\
 &\quad 7(T13+ - G < 7), 8(T3+ - G = 7), 9(T3+ - G > 7) \\
 &\quad 10(\text{death or worst imaginable death})\} \\
 &= \{0, 1, 2, \dots, 10\}
 \end{aligned}$$

ただし, T1, T2, T3+は癌の進行診断レベルを表し, Gの数値は, 生検によるグリソンスコア値を表す. また, 検診データは主に55歳以上を対象にして集計されることが多く, ここでは, 離散時刻  $T = \{55, 56, \dots, 75\}$  での検診時間間隔が1年以上の場合を扱うことにする. このようにおくことで, スクリーニング時間間隔を2年以上にして考察することができる. 費用としては, 例えばOECDのデータが利用可能で, 検診, 診療, 治療, 緩和ケア etc.の各種データを扱うことができる([6]). また, 決定空間を表す変数としては, 例えば, [12]のモデルでは, PSA検査の受診の有無, 生検の受診の有無, 前立腺の切除の有無の行動選択を考える. 堀口[9]のモデルのように, 検診プログラムの有効性を議論する場合には, PSA検診の有無と罹患率, PSA検診の精度による定期健診受診の有無による生存率(死亡リスク)の考察を行う場合には, PSA検診の受診後の精密検査等による診断治療を, それらが多段決定モデルでの吸収状態に到達する deteriorating systemとして定式化することが可能である(cf. [7]).

QOLとQALYsに関連する効用推定値の具体例として(cf. [5]), 例えば, 効用値を社会全般の共通認識における推計値として, 0から1の数値を考える. 死亡時を0, 完全な健康状態を1で表し, 事象に対する影響を受けた期間での区間幅のあるutility(効用)をとらえる. 例えば, スクリーニング(検診)によって, 1週間(1/52)の間, 効用値が0.99(区間値は0.99-1)と考えて, ひとりあたりのQALYsの損失値は

$$(1 - 0.99) \times \frac{1}{52} \doteq 0.00019$$

が平均余命期間(生存年, life years) に対して失われているとされる. 例えば, 生検を行う場合には, 効用推計値が0.9(0.87-0.94)であって, 損失値は

$$0.1 \times \frac{3}{52} \doteq 0.00058$$

となる.

マルコフ決定過程の構成要素として,

- (年齢による時刻を表す時間集合)  $T = \{55, 56, \dots, 75\}$
- core processの状態空間  $S = \{0, 1, 2, \dots, 10\} = X$ (set of states of Core Process)
- 決定空間  $A = \{0(\text{no screening}), 1(\text{screening}), 2(\text{clinical diagnosis without screening})\}$
- observation process の表す観測状態空間  $Y = \{0(\text{negative}), 1(T1 - G < 7), 2(T1 - G = 7), \dots, 9(T3+ - G > 7)\}$ (set of states of observation process)

前述したように, 決定空間については, 生検と切除も考慮するモデルも考えることができる. core processの状態空間をもとにした状態の推移の状況を図(1)に示す.

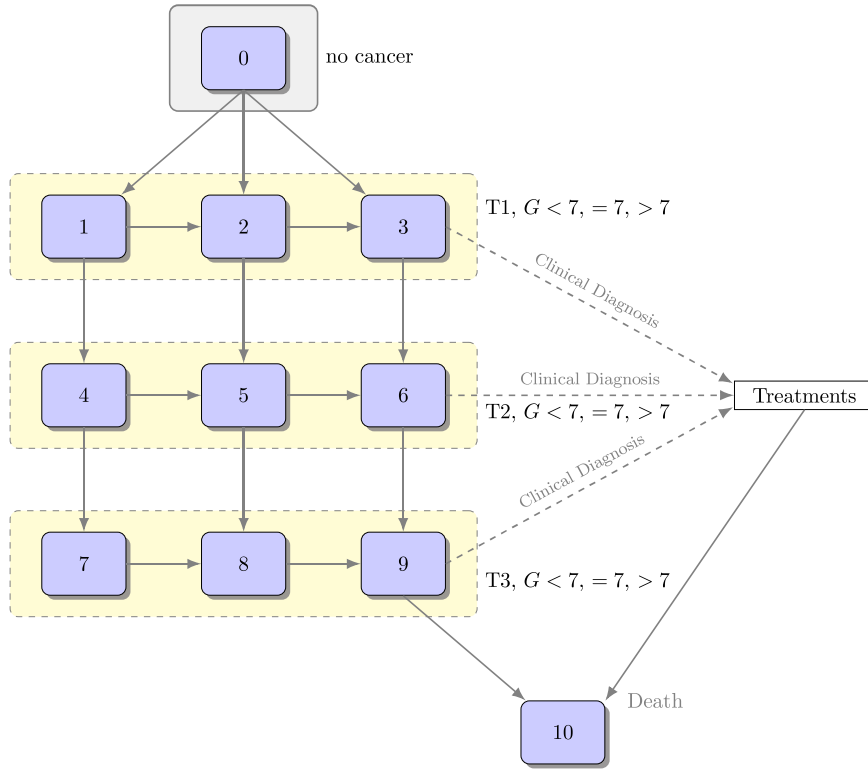


Figure 1: core processの各状態の推移状況

また、観測状態空間 $Y$ については、国内研究における前立腺がん検診のPSAと費用効果分析[4]にあるように、ベースラインPSAとされる2.0ng/mlとカットオフポイントの4.1ng/mlのほか、1.0ng/ml、3.0ng/mlの4つのPSA値による5つの状態分類から、検診期間間隔の変更や生検受診の選択を行うことが想定され、それに応じた観測状態空間 $Y$ の構成や感度分析も行うことが可能である。

さらに、推移確率行列 $P_{a_t}$ 、履歴集合 $H_n$ 、コスト関数 $C_a(x, y), x, y \in S$ 、観測状態(observation states  $y \in Y$ )に関する推移確率核 $q_a(y|i, j)$ を、通常の部分観測マルコフ決定モデル(Partially observable Markov decision model)として定義する。

このモデルでは、特に、

$$C_a(x, y) = [\underline{C}_a(x, y), \overline{C}_a(x, y)] \\ (= \{\underline{C}_a(x, y), \overline{C}_a(x, y)\} \text{ (low and high utility costs)})$$

を考慮することができる。

POMDPの構成として、

$$\mu_t(j) = \Phi_{a_{t-1}}(\mu_{t-1}, y_t)(j), j \in X, t \in T$$

について、policy  $\pi = (a_0, a_1, \dots, a_{T-1})$ のもと、 $\{\mu_t(\cdot)\}$ がマルコフ過程であり、次の定理が良く知られている。([10],[11])

**Theorem 1.** 任意の政策  $\pi = (a_0, a_1, \dots, a_{T-1})$  に対して, 確率分布の列  $\{\mu_i(\cdot)\}, i = 1, 2, \dots, T$  はマルコフ過程となる, すなわち, 任意の可測集合列  $\Gamma \in \mathcal{P}(X)$  に対して

$$Pr(\mu_i \in \Gamma | \mu_0, a_0, \mu_1, \dots, \mu_{t-1}, a_{t-1}) = Pr(\mu_i \in \Gamma | \mu_{i-1}, a_{i-1})$$

が成り立つ.

このとき, 次のDP方程式が成り立つ:

$$V_n(\mu) = \min_{a \in A} \left\{ C_a(\mu) + \sum_{y \in Y} V_{n-1}(\Phi_a(\mu, y)) Q_a(\mu; y) \right\}, n = 1, 2, \dots, N, \quad (1)$$

ただし,  $V_0$  は終端コスト関数である.

有限期間におけるマルコフ決定過程の最適化問題の結果から, 上記のDP方程式は, 最適性の原理によって,  $n = 1$  から  $N$  まで再帰的にDP方程式を解くことで最適解が得られる.

各期  $n = 1, 2, \dots, N$  に対して, DP方程式(1)のminimizerを  $a_{N-n}^*$  と表し,  $\pi^* = (a_0^*, a_1^*, \dots, a_{N-1}^*)$  とおく.

このとき, これらのDP方程式の解によって, 次の最適政策の存在が示される.

**Theorem 2.** DP方程式(1)のminimizerの最適解(最適政策列)は, 元問題のマルコフ決定過程の最適解と同値である, すなわち,  $V_{N, \pi^*}(\mu_0) = \sup_{\pi \in \Pi} V_{N, \pi}(\mu_0)$  が成り立つ.

POMDPsの数理モデルの詳細については, 例えば[3],[1],[8]を参照されたい.

## 4 cost effectiveness and cost utility analysis

マルコフモデルにおける評価関数の最小化, あるいは, 期待効用の最大化を考えるとときには, コスト関数および効用関数を適切に選べば, 有限期間問題においては動的計画法によって, 評価基準のもとでの最適化問題を再帰的に解くことが可能である. QOLおよびQALYsを価値関数評価とする費用効果分析もマルコフモデルにおいて解析可能である. 具体的には, 行動選択によって変化するQOLに関する期待効用値を, それぞれの行動選択によって生じる対象期間での積算値として評価することで, 行動選択の結果によるQALYs値の比較が可能で, 患者にとっての有益な情報提供のひとつとなる. 決定空間(行動空間)を適切に設定することで, 想定される選択と結果によって, どのようにQALYsの値が変化するか, 実情を反映したモデル構築が可能である. 評価手法は数理モデルを基に解析されるものであり, 医師や薬理学の専門家によるさらなるモデルの改良, 世代間での行動分析などにも適用できると考えられる. また, モデル化に関しては, 疫学上のデータ解析手法や, 構造モデルそのもののパラメータ値の適切性の議論も欠くことはできないと考える. 検診精度がほぼ同じもとでの費用効果分析であっても, 個々の患者の考え方, あるいは, ある一定の数値での検出力のもとで, 治療後の社会復帰を早く望む場合であるかどうかでも, 費用の最小化だけが最適な行動選択とはならないこともある. これらの議論は, 検診を含む診療計画における経済的評価全般に関する問題でもあり(cf. [2]), 多様なアプロー

チによる問題解決手法の構築とそれらの議論が、社会生活を豊かに送るための糧になりえると考えている。

QOL, QALYsの数値の利用に関しては、例えば、区間表現された効用推定値(utility estimates)を $S = \{0, 1, \dots, 10\}$ に対して利用(cf. [5],[6])があげられる。区間解析として、積算効用の上界値と下界値による行動選択の解釈も、マルコフモデルにおいて考察することができる。評価方法としては、費用効果(cost effectiveness)により推測される数値をもとにスクリーニングの各シナリオやスクリーニングなしの場合とのQALYsへの影響の比較も可能である。各シナリオ(finite policy  $\pi = (\pi_0, \pi_1, \dots, \pi_{N-1})$ )に対して、総コスト(net cost の増加分)をQALYsで割った比(平均コスト)で比較することで、費用効用分析(cost utility analysis)が可能である。感度解析としても、効用推定値の区間表現の下限值, 上限値を用いて行動選択分析をすることができる。得られる結果としては、費用効果分析, 費用効用分析のそれぞれで、例えばマイクロシミュレーションなどによる詳細な患者の動向分析が可能である。さらに言えば、多目的評価関数のもとでのパレート最適性の議論も可能であり、今後の研究課題の一つである。

## References

- [1] Bäuerle, N., Rieder, U.: Markov Decision Processes with Applications to Finance. Springer, Heidelberg (2011)
- [2] Drummond M.F. Sculpher M.J. Claxton K. et al. *Methods for the Economic Evaluation of Health Care Programmes*. Oxford University Press, Oxford, UK, 2015
- [3] Dynkin, E.B., Yushkevich, A.A.: Controlled Markov Processes Springer, New York (1979)
- [4] 後藤 励, 小林 恭, 光森 健二 マルコフモデルを用いた前立腺がん検診の費用効果分析 –最適な検診方法の設計– 医療経済研究, Vol.17, 2005, pp. 21–41
- [5] E.A.M. Heijnsdijk et. al., Quality-of-Life Effects of Prostate-Specific Antigen Screening, *The New England Journal of Medicine*, 2012, pp. 595–605.
- [6] E.A.M. Heijnsdijk et. al., Cost-effectiveness of Prostate Cancer Screening: A Simulation Study Based on ERSPC Data, *J Natl Cancer Instit*, 2015, pp. 1–9.
- [7] Hordijk, A. (1974). *Dynamic Programming and Markov Potential Theory*. Mathematical Centre Tracts, No. 51. Amsterdam: Mathematisch Centrum.
- [8] Hinderer, K., Rieder, U., Stieglitz, M.: Dynamic Optimization: Deterministic and Stochastic Models. Springer, Cham (2016)
- [9] M. Horiguchi, On an Approach to Evaluation of Health Care Programme by Markov Decision Model, In *Modern Trends in Controlled Stochastic Processes : Theory and Applications, V.III*, edited by: Alexey Piunovskiy, Yi Zhang, Springer, pp 341-354, 2021.

- [10] G. E. Monahan. A survey of partially observable Markov decision processes: Theory, models, and algorithms, *Manag. Sci.* 28(1982)1–16.
- [11] U. Rieder. Structural results for partially observed control models. *ZOR Methods and Models of Operations Research*, 1991; 35: 473–490.
- [12] J. Zhang and B. T. Denton, Partially observable Markov decision processes for prostate cancer screening, surveillance, and treatment: A budgeted sampling approximation method, In: *Decision Analytics and Optimization in Disease Prevention and Treatment*, John Wiley and sons, New York, 2018, pp. 201–222.

Masahiko SAKAGUCHI

Department of Engineering Informatics,  
Faculty of Information and Communication Engineering,  
Osaka Electro-Communication University,

Aiko KURUSHIMA

Department of Economics,  
Faculty of Economics,  
Sophia University

Masayuki Horiguchi

Department of Mathematics,  
Faculty of Science,  
Kanagawa University

大阪電気通信大学・情報通信工学部・阪口昌彦  
上智大学・経済学部・來島愛子  
神奈川大学・理学部 堀口正之