Ingeniería Investigación y Tecnología
volumen XXVI (número 2), abril-junio 2025 1-9
ISSN 2594-0732 FI-UNAM artículo arbitrado
Información del artículo: Recibido: 10 de enero de 2025, aceptado: 18 de febrero de 2025
Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) license
https://doi.org/10.22201/fi.25940732e.2025.26.2.011

Ingeniería Investigación y Tecnología

Método interactivo para resolver procesos de decisión secuenciales
Interactive method for solving sequential decision processes

Acosta-Flores José Jesús
Facultad de Ingeniería, UNAM
División de Ingeniería Mecánica e Industrial
Correo: jjaf@unam.mx
https://orcid.org/0000-0001-8101-5797

Resumen

El objeto de este artículo es presentar un método sencillo para el análisis de procesos de decisión secuenciales donde interactúa un decisor y un analista. Los problemas que se resuelven son secuenciales, probabilistas y con un objetivo, contemplando la cuantificación de la incertidumbre mediante procesos de Markov y la del comportamiento del decisor mediante las funciones de utilidad. El método de Ronald Howard, que emplea procesos de Markov y programación dinámica, se complementa en este artículo con la teoría de decisiones, señalando los pasos que se deben seguir para obtener la política óptima que establece para cada estado del sistema la alternativa que maximiza el logro del objetivo. Las limitaciones de este estudio son tres: válido solo para un decisor, un objetivo y procesos monodésmicos de Markov. La inclusión de la Teoría de decisiones, en forma estructurada, para que interactúen el decisor y al analista es totalmente original. Al considerar que el decisor puede tener un comportamiento de aversión, de propensión o de neutralidad al riesgo, se logra que las soluciones sugeridas sean las adecuadas para cada decisor, lo que conduce a ahorros considerables, ya que puede considerarse como si fuesen trajes a la medida. Finalmente, se ilustró el método con un ejemplo de la operación de un taxi entre tres alcaldías de la Ciudad de México.
Descriptores: Teoría de decisiones, procesos de Markov, programación dinámica, funciones utilidad, equivalentes bajo certeza.

Abstract

The purpose of this article is to present a simple method for the analysis of sequential decision processes where a decision-maker and an analyst interact. The problems that are solved are sequential, probabilistic and with an objective, contemplating the quantification of uncertainty through Markov processes and the behavior of the decision-maker through utility functions. Ronald Howard's method, which uses Markov processes and dynamic programming, is complemented in this article with decision theory, pointing out the steps that must be followed to obtain the optimal policy that establishes for each state of the system the alternative that maximizes the achievement of the objective. The limitations of this study are threefold: valid only for one decision-maker, one objective, and Markov monodesmic processes. The inclusion of Decision Theory, in a structured way, so that the decision-maker and the analyst interact is totally original. By considering that the decision-maker may have a behavior of aversion, proneness, or neutrality to risk, it is possible that the suggested solutions are appropriate for each decision-maker, which leads to considerable savings, since they can be considered as if they were tailor-made suits. Finally, the method was illustrated with an example of the operation of a taxi between three mayoralties in Mexico City.
Keywords: Decision Theory, Markov processes, dynamic programming, utility functions, equivalents under certainty.

Introducción

Los problemas de decisión pueden ser de una sola etapa o secuenciales, deterministas o probabilistas y con un objetivo o varios.

Los métodos para alcanzar la decisión óptima cuando los problemas son de una etapa, deterministas y un objetivo son: el valor presente neto, la relación beneficio costo, la relación efectividad costo, la tasa interna de retorno y el período de recuperación. En los tres primeros se utiliza una tasa de descuento, en el cuarto se calcula la tasa de rendimiento y en el quinto no se emplea ninguna tasa. Para varios objetivos se tienen los Métodos Multicriterio que son fáciles de utilizar, inclusive muchos de ellos están automatizados, lo cual ha hecho que su aplicación en problemas reales sea muy amplia (Hajkowiczs & Collins, 2007; Cinelli *et al*., 2014; Balali *et al.,* 2020; De Brito & Evers, 2016). Un método alternativo combina el análisis de dominancia con las permutas compensatorias de objetivos (Hammond *et al.,* 2000).

Los problemas secuenciales, probabilistas y un objetivo se han analizado con la Teoría de decisiones bayesianas. Dos de los pioneros son Raiffa H. (1968) y Schlaifer R. (1969) de la Universidad de Harvard, quienes consideran problemas con varias etapas para determinar la estrategia óptima que contemple las mejores decisiones dependiendo de los eventos que vayan ocurriendo.

Howard R. (1962) en su tesis doctoral en el Instituto Tecnológico de Massachusetts presenta un método que utiliza la programación dinámica para resolver procesos markovianos de decisión secuenciales. Presenta y resuelve problemas de operación de un taxi considerando tres estados (donde el estado de un sistema identifica sin ambigüedad su situación), de beisbol con 50 estados, y de reemplazo de maquinaria con 40 estados.

Por otra parte, Akbari & Sayarshad (2022) desarrollaron un modelo matemático basado en procesos de decisión de Markov (MDP) para coordinar la restauración de infraestructura vial y la logística de distribución. Este modelo, probado durante la recuperación del huracán Harvey, optimiza la asignación de recursos, reduciendo significativamente los tiempos de respuesta.

Otarola *et al.* (2024) proponen un marco que utiliza procesos de Markov para analizar el deterioro de sistemas críticos (como puentes y carreteras) sujetos a múltiples amenazas combinadas, como terremotos e inundaciones. Este marco permite modelar interacciones complejas y planificar estrategias de mitigación a lo largo del ciclo de vida de la infraestructura.

En Wang *et al.* (2025) se encuentra un modelo semi-Markoviano que introduce estrategias para la recuperación resiliente de sistemas de distribución, manejando incertidumbres inherentes a desastres naturales. Este enfoque se centra en mantener operaciones críticas mientras se prioriza la restauración. Desde Howard hasta Wang no han utilizado la Teoría de Decisiones.

Algunos métodos para analizar decisiones en problemas con incertidumbre, eventos discretos y una etapa son: criterio de Laplace, de max-min, de max-max y de lamentos. Estos métodos son sencillos y por ello se han aplicado mucho a problemas reales, no obstante que Raiffa (1968) los describe y muestra claramente que no es conveniente su uso. Posteriormente, De Neufville (1990) vuelve a describirlos agregando el criterio de índices ponderados, mostrando de nuevo por qué pueden conducir a decisiones erróneas.

En vez de usarlos es mejor determinar las funciones de probabilidad y usar el método de la función utilidad con atributos múltiples de Keeney (1976). El caso del Aeropuerto de la Ciudad de México es clásico: fue la primera aplicación importante de dicho enfoque y ha sido publicado extensivamente (Drake *et al.,* 1972; De Neufville & Marks, 1976; Keeney & Raiffa, 1976; De Neufville, 1990).

Acosta (2019 y 2024) presenta un algoritmo para analizar decisiones con objetivos múltiples y una etapa, que considera las funciones masa de probabilidad uniforme, normal, exponencial, Cauchy, Chi-cuadrada, Erlang, Gamma y Laplace.

Como se mencionó anteriormente Howard resuelve el problema de procesos secuenciales con varias etapas, pero no considera la Teoría de Decisiones para determinar la estructura de preferencias del decisor.

La finalidad del método sugerido en este artículo es presentar un procedimiento sencillo que guíe al analista y al decisor para que interactúen, considerando un proceso de Markov y las preferencias del decisor con el objetivo de obtener la mejor política para él. Esta política considera el análisis de las alternativas en cada estado considerando procesos de Markov, programación dinámica y la función utilidad del decisor.

Metodología

Para el análisis de problemas de decisión secuenciales se tiene como herramienta el Proceso de Markov. En este proceso dos conceptos fundamentales son los estados y las transiciones. Los estados definen la situación del sistema y las transiciones son los momentos en que el sistema cambia de un estado a otro o permanece en él.

Un proceso de Markov está totalmente definido cuando se conoce la matriz de probabilidades de transición $\boldsymbol{P} = [p_{ij}]$ donde pij es la probabilidad de pasar del estado i al estado j en una transición. La hipótesis de Markov establece que esta probabilidad depende sólo del estado i y no de la trayectoria para llegar a él.

El método que se presenta en este artículo es válido sólo cuando el proceso es monodésmico, y un proceso de Markov es monodésmico cuando tiene una sola cadena, es decir, cuando es posible pasar de cualquier estado a cualquier otro de la cadena (Howard, 1971).

Para resolver los procesos de Markov secuenciales Howard (1962) utilizó programación dinámica (Bellman, 1957) usando la función recursiva:

$$v_i(n) = q_i + \sum_{j=1}^{N} p_{ij} v_j (n-1) \quad (1)$$

Donde:

$v_i(n)$ = ganancias totales esperadas en las n transiciones siguientes si el sistema está ahora en el estado i
p_{ij} = probabilidad de pasar del estado i al estado j en una transición
r_{ij} = recompensa por pasar del estado i al estado j
N = total de estados
qi = $\sum_{j=1}^{N} p_{ij} r_{ij}$

La función recursiva (1) tiene como asíntota:

$$v_i (n) = ng + v_i \quad (2)$$

Donde g es la ganancia del proceso.

Después de las transiciones requeridas para llegar al límite se puede utilizar en vez de la función recursiva su asíntota, de manera que al sustituir (2) en (1) se obtiene:

$$ng + v_i = q_i + \sum_{j=1}^{N} p_{ij}[(n-1)g + v_j]$$

$$ng + v_i = q_i + n_g \sum_{j=1}^{N} p_{ij} - g\sum_{j=1}^{N} p_{ij} + \sum_{j=1}^{N} p_{ij} v_j$$

como $\sum_{j=1}^{N} p_{ij} = 1$ queda $g + v_i = q_i + \sum_{j=1}^{N} p_{ij} v_j$

Cuando se tienen varias alternativas k en cada estado, Howard (1962) utiliza un ciclo de iteraciones para obtener la política óptima.

Este ciclo está formado por la Operación de determinación de valores y por la Rutina para mejorar la política.

En la primera, se resuelve el sistema de ecuaciones $g + vi = qi + \sum_{j=1}^{N} p_{ij} v_j$ con $v_N = 0$

$i = 1, ..., N$, cuyas incógnitas son g y v_i

En la segunda, para cada estado, usando los resultados anteriores se encuentra la alternativa k que maximiza $q_i^k + \sum_{j=1}^{N} p_{ij}^{N} v_j$

El ciclo continúa hasta que la política nueva coincide con la anterior, siendo ésta la óptima.

Para considerar las preferencias del decisor, se obtiene la función utilidad que representa su comportamiento. Este comportamiento puede ser de aversión, de propensión o de neutralidad al riesgo.

Keeney (1992) menciona que en toda su carrera profesional ha observado que los decisores tienen un comportamiento de aversión o propensión constante al riesgo y por tanto su función utilidad es de tipo exponencial. En este artículo se utilizará esa experiencia contemplando únicamente funciones utilidad de tipo exponencial.

También, puesto que la utilidad de la suma de dos cantidades no es la suma de sus utilidades, para poder aplicar el método de Howard se utilizaron los equivalentes bajo certeza, EBC.

A continuación, se muestra el método que se sugiere, el cual consiste en utilizar el desarrollado por Howard complementándolo con el uso de la Teoría de Decisiones para considerar las preferencias del decisor mediante sus equivalentes bajo certeza.

Método interactivo para la solución de problemas de decisión secuenciales

Parte uno. Obtención de la función utilidad:

Paso 1-1: Defina el horizonte de planeación, el objetivo, su medida de efectividad y los estados del sistema.

Paso 1-2: Establezca si la medida de efectividad es monotónica creciente o decreciente.

Paso 1-3: Precise las probabilidades p_{ij}^{k} y las recompensas r_{ij}^{k} para cada alternativa k. Vea si el proceso de Markov es monodésmico. Si no es así pare, porque el método no es aplicable. Si es monodésmico continúe en el Paso 1-4.

Paso 1-4: Obtenga $a^* = \max \{r_{ij}^{k}\}$

Paso 1-5: Forme la lotería cuyas consecuencias son a* y 0, con probabilidades de 0.5 cada una. Pídale al decisor su equivalente bajo certeza (EBC) de esa lotería. Calcule el valor esperado (VE) de la lotería. Si la medida de efectividad es monotónica creciente y VE > EBC entonces se tiene un comportamiento de aversión al riesgo; si VE = EBC, el comportamiento es de neutralidad al riesgo; y si VE < EBC es de propensión al riesgo. Si la medida de efectividad es monotónica decreciente y VE > EBC entonces se tiene propensión al riesgo; si VE = EBC, neutralidad al riesgo; y si VE < EBC aversión al riesgo. Conociendo lo anterior determine, su u(x) en la Tabla 1.

Tabla 1. Función utilidad *u* (*x*), con *c* > 0

	Monotónica creciente	Monotónica decreciente
Aversión constante	$-e^{-cx}$	$-e^{cx}$
Neutralidad	x	-x
Propensión constante	$e_{ij}^{cr_{ij}^{k}}$	$e^{cr_{ij}^{k}}$

Paso 1-6: Para un comportamiento de aversión o propensión al riesgo formule la ecuación u(EBC) = 0.5u (a*) + 0.5u (0), la cual es una ecuación con una sola incógnita, el parámetro c. Resuelva dicha ecuación para tener c > 0 (puede usar el método Newton-Raphson).

Parte dos. Cálculo de los equivalentes bajo certeza:

Paso 2-1: Para cada alternativa k en el estado i, calcule el equivalente bajo certeza de su ganancia esperada inmediata usando la ecuación que se muestra en la Tabla 2.

Tabla 2. Equivalentes bajo certeza, EBC_i^k

	u(x) es monotónica creciente	u(x) es monotónica decreciente
Aversión constante	$\frac{-1}{c} Ln\left(\sum_{j=1}^{N} p_{ij}^{k} e^{-cr_{ij}^{k}}\right)$	$\frac{1}{c} Ln\left(\sum_{j=1}^{N} p_{ij}^{k} e^{cr_{ij}^{k}}\right)$
Neutralidad	$\sum_{j=1}^{N} p_{ij}^{k} r_{ij}^{k}$	$\sum_{j=1}^{N} p_{ij}^{k} r_{ij}^{k}$
Propensión constante	$\frac{1}{c} Ln\left(\sum_{j=1}^{N} p_{ij}^{k} e_{ii}^{cr_{ij}^{k}}\right)$	$\frac{-1}{c} Ln\left(\sum_{j=1}^{N} p_{ij}^{k} e_{i}^{-cr_{ij}^{k}}\right)$

Parte tres. Selección de la política óptima:

Paso 3-1: Haga $v_j = 0$ para j = 1, ... , N

Paso 3-2: Para cada estado i, encuentre la alternativa k'_i que maximiza

$$EBC_i^k + \sum_{j=1}^{N} p_{ij}^{k} v_j$$

Ahora k'_i es la decisión en el estado i, $EBC_i = EBC_i^{k'}$ y $P_{ij} = p_{ij}^{k'}$

La política es

$$d = \begin{bmatrix} k'_1 \\ k'_2 \\ \vdots \\ k'_N \end{bmatrix}$$

Paso 3-3: ¿La nueva política d coincide con la que se tenía? En caso afirmativo, esa política es la mejor opción y termina el método. Si no es así, continúe con el paso 3.4.

Paso 3-4: Resuelva el sistema de N ecuaciones con N + 1 incógnitas (g y las v_i) haciendo $V_N = 0$

$$g + v_i = EBC_i + \sum_{j=1}^{N} p_{ij} v_j \qquad i = 1, \ldots, N;$$

Regrese al Paso 3-2.

Finalmente, se ilustra el método aplicándolo a un ejemplo hipotético.

Ejemplo. Operación de un taxi

Considere el problema de un taxista que opera en tres alcaldías de la ciudad de México: Benito Juárez, Coyoacán y Tlalpan. En cualquiera de las tres alcaldías tiene tres alternativas:

1. Puede estar circulando, hasta que un pasajero solicite su servicio.
2. Puede conducir hacia la base más cercana y esperar en la línea hasta que le toque su turno para llevar pasaje.
3. Puede estacionarse y esperar una llamada de radio.

Parte uno. Obtención de la función utilidad:

Paso 1-1: Defina el horizonte de planeación, el objetivo, su medida de efectividad y los estados del sistema.

El horizonte de planeación es un año. El objetivo maximizar las ganancias del taxista y la medida de efectividad es el ingreso promedio en pesos promedio por viaje.

El sistema tiene tres estados que son:

Estado 1. El taxista se encuentra en la Alcaldía Benito Juárez.
Estado 2. El taxista está en la Alcaldía Coyoacán.
Estado 3. El taxista se encuentra en la Alcaldía Tlalpan.

Paso 1-2: Establezca si la medida de efectividad es monotónica creciente o decreciente.

La medida de efectividad es monotónica creciente porque el operador del taxi prefiere cantidades mayores de pesos promedio por viaje sobre cantidades menores.

Paso 1-3: Precise las probabilidades p_{ij}^{k} y las recompensas r_{ij}^{k} para cada alternativa k. Vea si el proceso de Markov es monodésmico. Si no es así, pare porque el método no es aplicable. Si es monodésmico continúe en el Paso 1-4.

La información solicitada se encuentra en la Tabla 3.

Como el proceso de Markov es monodésmico es posible continuar en el paso siguiente.

Paso 1-4: Obtenga a* = max $\left\{ r_{ij}^{k} \right\}$

$$a^{*} = 300$$

Paso 1-5: Forme la lotería cuyas consecuencias son a* y 0, con probabilidades de 0.5 cada una. Pídale al decisor su equivalente bajo certeza (EBC) de esa lotería. Calcule el valor esperado (VE) de la lotería. Si la medida de efectividad es monotónica creciente y VE > EBC entonces se tiene un comportamiento de aversión al riesgo; si VE = EBC, el comportamiento es de neutralidad al riesgo; y si VE < EBC es de propensión al riesgo. Si la medida de efectividad es monotónica decreciente y VE > EBC entonces se tiene propensión al riesgo; si VE = EBC, neutralidad al riesgo; y si VE < EBC aversión al riesgo. Conociendo lo anterior determine, su u(x) en la Tabla 4.

El taxista contestó que la mínima cantidad que estaría dispuesto a aceptar a cambio de la lotería que con probabilidad 0.5 puede obtener 0 o 300 pesos es 140 pesos, luego EBC = 140. El valor esperado de esa lotería es $VE = (0.5)(0) + (0.5)(300) = 150$.

Como la medida de efectividad es monotónica creciente y $VE > EBC$ entonces el decisor tiene un comportamiento de aversión al riesgo. Viendo la Tabla 1 la función utilidad que representa el comportamiento del taxista es $u\,(x) = -\,e^{-cx}$

Paso 1-6: Para un comportamiento de aversión o propensión al riesgo formule la ecuación u (EBC) = 0.5u(a*) + 0.5u (0), la cual es una ecuación con una sola incógnita, el parámetro c. Resuelva dicha ecuación para tener c > 0 (puede usar el método Newton-Raphson).

Luego la ecuación es: $-\,e^{-140c} = 0.5\,(-\,e^{-300c}) - 0.5$ de manera que, utilizando el algoritmo Newton-Raphson para determinar el valor de c:

$c_{n+1} = c_n - f(c)\,/\,f'(c)$; donde $f'(c)$ es la primera derivada de f con respecto a c, entonces:

$$f(c) = 0.5 + 0.5e^{-300c} - e^{-140c}$$

$$f'(c) = -\,150e^{-300c} + 140e^{-140c}$$

Se obtiene $c = 0.000891534$.

Tabla 3. Probabilidades y recompensas

Estado i	Alternativa k	Probabilidad p_{ij}^{k}			Recompensa r_{ij}^{k}		
		$j = 1$	$j = 2$	$j = 3$	$j = 1$	$j = 2$	$j = 3$
1	1	0.5	0.3	0.2	100	130	150
	2	0.3	0.5	0.2	80	110	130
	3	0.25	0.4	0.35	90	120	140
2	1	0.25	0.5	0.25	200	260	300
	2	0.3	0.4	0.3	110	140	160
	3	0.4	0.3	0.3	100	130	150
3	1	0.3	0.3	0.4	110	100	80
	2	0.1	0.4	0.5	100	90	70
	3	0.3	0.4	0.3	90	80	60

Tabla 4. Función utilidad $u\,(x)$, con $c > 0$

	Monotónica creciente	Monotónica decreciente
Aversión constante	$-e^{-cx}$	$-e^{cx}$
Neutralidad	x	-x
Propensión constante	e^{cx}	e^{-cx}

Parte dos. Cálculo de los equivalentes bajo certeza:

Paso 2-1: Para cada alternativa k en el estado i, calcule el equivalente bajo certeza de su ganancia esperada inmediata usando la ecuación que se muestra en la Tabla 5.

Usando:

$$EBC = \frac{-1}{c} Ln\left(\sum_{j=1}^{N} p_{ij}^{k} e^{-cr_{ij}^{k}}\right)$$

Se obtuvieron los valores que se muestran en la Tabla 6.

Parte tres. Selección de la política óptima:

Paso 3 -1: Haga $v_j = 0$ para $j = 1..., N$
$v_1 = v_2 = v_3 = 0$

Paso 3-2: Para cada estado i, encuentre la alternativa k'_j que maximiza:

$$EBC_i^k + \sum_{j=1}^{N} p_{ij}^{k} v_j$$

Examinando los EBC_i^k en la Tabla 6 se observa que la política consiste en escoger la alternativa 3 en el estado 1, la 1 en el estado 2 y la 1 en el estado 3, es decir:

$$\boldsymbol{d} = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$$

Las probabilidades de transición y equivalentes bajo certeza correspondientes a esta política son:

$$\boldsymbol{P} = \begin{bmatrix} 0.25 & 0.4 & 0.35 \\ 0.25 & 0.5 & 0.25 \\ 0.3 & 0.3 & 0.4 \end{bmatrix} \quad \text{y } \boldsymbol{EBC} = \begin{bmatrix} 119.3 \\ 254.4 \\ 94.93 \end{bmatrix}$$

Paso 3-3: ¿La nueva política **d** coincide con la que se tenía? en caso afirmativo, esa política es la mejor opción y termina el método. Si no es así, continúe con el paso 3.4. Como no hay política anterior, continuamos con el Paso siguiente.

Tabla 5. Equivalentes bajo certeza, EBC_i^k

	u(x) es monotónica creciente	u(x) es monotónica decreciente
Aversión constante	$\frac{-1}{c} Ln\left(\sum_{j=1}^{N} p_{ij}^{k} e^{-cr_{ij}^{k}}\right)$	$\frac{1}{c} Ln\left(\sum_{j=1}^{N} p_{ij}^{k} e^{cr_{ij}^{k}}\right)$
Neutralidad	$\sum_{j=1}^{N} p_{ij}^{k} r_{ij}^{k}$	$\sum_{j=1}^{N} p_{ij}^{k} r_{ij}^{k}$
Propensión constante	$\frac{1}{c} Ln\left(\sum_{j=1}^{N} p_{ij}^{k} e^{cr_{ij}^{k}}\right)$	$\frac{-1}{c} Ln\left(\sum_{j=1}^{N} p_{ij}^{k} e^{-cr_{ij}^{k}}\right)$

Tabla 6. EBC, equivalentes bajo certeza

Estado i	Alternativa k	Probabilidad p_{ij}^k			Recompensa r_{ij}^k			EBC_i^K
		$j = 1$	$j = 2$	$j = 3$	$j = 1$	$j = 2$	$j = 3$	
1	1	0.5	0.3	0.2	100	130	150	118.8
	2	0.3	0.5	0.2	80	110	130	104.9
	3	0.25	0.4	0.35	90	120	140	119.3
2	1	0.25	0.5	0.25	200	260	300	254.4
	2	0.3	0.4	0.3	110	140	160	136.8
	3	0.4	0.3	0.3	100	130	150	123.8
3	1	0.3	0.3	0.4	110	100	80	94.93
	2	0.1	0.4	0.5	100	90	70	80.94
	3	0.3	0.4	0.3	90	80	60	76.94

Paso 3-4: Resuelva el sistema de N ecuaciones con N + 1 incógnitas (g y las v_i) haciendo $v_N = 0$

$$g + vi = EBC_i + \sum_{j=1}^{N} p_{ij} v_j \quad i = 1,\ldots, N;$$

Las ecuaciones son:

$$g + v_1 = 119.3 + 0.25v_1 + 0.4v_2$$

$$g + v_2 = 254.4 + 0.25v_1 + 0.5v_2$$

$$g = 94.93 + 0.3\, v_1 + 0.3\, v_2$$

cuya solución es: $g = 166.5 \quad v_1 = 41.94 \quad v_2 = 196.7 \quad v_3 = 0$

Regresе al Paso 3-2.
Paso 3-2: Para cada estado i, encuentre la alternativa k'_i que maximiza:

$$EBC^k_i + \sum_{j=1}^{N} p^k_{ij} v_j$$

Estas cantidades se presentan en la Tabla 7.

Tabla 7. Para mejorar la política

Estado i	Alternativa K	$EBC^k_i + \sum_{j=1}^{N} p^k_{ij} v_j$
1	1	198.8
	2	215.8 ←
	3	208.5
2	1	363.3 ←
	2	228.1
	3	199.6
3	1	166.5
	2	163.8
	3	168.2 ←

Como puede verse en la Tabla 7, para i = 1 la cantidad en la columna de la extrema derecha se maximiza cuando $k = 2$. Para i = 2 se maximiza cuando $k = 1$. Para i = 3 se maximiza cuando $k = 3$.

De manera que la política nueva es:

$$\boldsymbol{d} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

Las probabilidades de transición y equivalentes bajo certeza correspondientes a esta política son:

$$\boldsymbol{P} = \begin{bmatrix} 0.3 & 0.5 & 0.2 \\ 0.25 & 0.5 & 0.25 \\ 0.3 & 0.4 & 0.3 \end{bmatrix} \quad \text{y } \boldsymbol{EBC} = \begin{bmatrix} 104.9 \\ 254.4 \\ 76.94 \end{bmatrix}$$

Paso 3-3: ¿La nueva política **d** coincide con la que se tenía? en caso afirmativo, esa política es la mejor opción y termina el método. Si no es así, continúe con el paso 3.4.

Como la política nueva no coincide con la anterior, continuamos en el Paso siguiente.

Paso 3-4: Resuelva el sistema de N ecuaciones con $N + 1$ incógnitas (g y las v_j) haciendo $v_N = 0$

$$g + v_i = EBC_i + \sum_{j=1}^{N} p_{ij} v_j \quad i = 1, \ldots, N;$$

Las ecuaciones son:

$$g + v_1 = 104.9 + 0.3v_1 + 0.5v_2$$

$$g + v_2 = 254.4 + 0.25v_1 + 0.5v_2$$

$$g = 76.94 + 0.3v_1 + 0.4v_2$$

cuya solución es: $g = 169 \quad v_1 = 47.41 \quad v_2 = 194.5 \quad v_3 = 0$

Regrese al Paso 3-2.

Paso 3-2: Para cada estado i, encuentre la alternativa k'_i que maximiza:

$$EBC^k_i + \sum_{j=1}^{N} p^k_{ij} v_j$$

Estas cantidades se presentan en la Tabla 8.

Tabla 8. Para mejorar la política

Estado i	Alternativa k	$EBC_i^k + \sum_{j=1}^{N} p_{ij}^k v_j$
1	1	200.9
	2	216.3 ←
	3	209
2	1	363.5 ←
	2	228.9
	3	201.1
3	1	167.5
	2	163.5
	3	169 ←

Como puede verse en la Tabla 8, para i = 1 la cantidad en la columna de la extrema derecha se maximiza cuando $k = 2$. Para i = 2 se maximiza cuando $k = 1$. Para i = 3 se maximiza cuando $k = 3$.

De manera que la política nueva es:

$$d = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

Las probabilidades de transición y equivalentes bajo certeza correspondientes a esta política son:

$$\boldsymbol{P} = \begin{bmatrix} 0.3 & 0.5 & 0.2 \\ 0.25 & 0.5 & 0.25 \\ 0.3 & 0.4 & 0.3 \end{bmatrix} \text{ y } \boldsymbol{EBC} = \begin{bmatrix} 104.9 \\ 254.4 \\ 76.94 \end{bmatrix}$$

Paso 3-3: ¿La nueva política **d** coincide con la que se tenía? en caso afirmativo, esa política es la mejor opción y termina el método. Si no es así, continúe con el paso 3.4.

Como la política nueva coincide con la anterior, paramos el método y la política óptima es:

$$\boldsymbol{d} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

Esta política consiste en lo siguiente:

Si está en la Alcaldía Benito Juárez debe conducir hacia la base más cercana y esperar formado hasta que le toque su turno para llevar pasaje.

Si está en la Alcaldía Coyoacán debe permanecer circulando hasta que soliciten su servicio.

Si está en la Alcaldía Tlalpan debe detenerse y esperar una llamada de radio.

Conclusiones

El método presentado en este artículo es interactivo y guía paso a paso al analista para que se facilite su interacción con el decisor.

Modela la incertidumbre mediante un proceso de Markov monodésmico y cuantifica las preferencias del decisor usando la función de utilidad que representa su estructura de preferencias, pudiendo ser de aversión constante, propensión constante o neutralidad al riesgo.

Complementa el método desarrollado por Howard (1962), quien usa procesos de Markov y Programación dinámica, usando la Teoría de Decisiones para que las decisiones recomendadas contemplen el comportamiento del decisor, lo cual conduce a mejores decisiones, especialmente cuando las cantidades involucradas son grandes, ya que las mejores políticas que establecen la mejor decisión en cada etapa y en cada estado del sistema para una compañía pequeña, que no puede afrontar mucho riesgo, son diferentes a las de compañías más grandes, con mayores recursos.

Las principales contribuciones en este artículo son: un método que guía paso a paso al analista para que pueda interactuar con el decisor; la utilización de la Teoría de Decisiones para determinar la función utilidad del decisor que represente su comportamiento; y el procedimiento para que el analista pueda determinar los valores de los equivalentes bajo certeza requeridos.

El método es aplicable cuando se tiene un solo decisor, varias etapas y la incertidumbre puede modelarse como un proceso de Markov monodésmico. Se ilustró su uso con un ejemplo hipotético sobre la operación de un taxi.

Referencias

Acosta-Flores, J. J.(2019). Algoritmo para analizar decisiones con objetivos múltiples bajo incertidumbre. *Ingeniería Investigación y Tecnología*, 20 (01), 1-8. https://doi.org/10.22201/fi.25940732e.2019.20n1.010

Acosta-Flores, J. J. (2024). Decisions with uncertainty using deterministic analysis. *Journal of Engineering Research*, 24(5), 1-14. https://journalspress.uk/index.php/LJER/article/view/895

Akbari, V., & Sayarshad, H. R. (2022). Integrated and coordinated relief logistics and road recovery planning problems. *Transportation Research*, Part D, 111: 103433. http://dx.doi.org/10.1016/j.trd.2022.103433

Balali, F., Nouri, J., Nasiri, A., & Zhao, T. (2020). *Data intensive industrial asset management*. Springer.

Bellman, R. (1957). *Dynamic programming*. Princeton University Press.

Cinelli, M., Coles, M., & Kirwan, K. (2014). Analysis of the potential of multi criteria decision analysis methods to conduct sustainability assessment. *Ecology Indicators*, 46, 138-148. https://doi.org/10.1016/j.ecolind.2014.06.011

De Brito, M. M., & Evers, M. (2016) Multi-criteria decision-making for flood risk management: A survey of the current state of the art. *Natural Hazards Earth Systems Science,* 16(4), 1019-1033. https://doi.org/10.5194/nhess-16-1019-2016

De Neufville, R. (1990). *Applied systems analysis: engineering planning and technology Management*. McGraw Hill Inc.

Drake, A. W., Keeney, R. L., & Morse, P. M. (1972). *Analysis of public systems*. The MIT press.

Hajkowicz, S., & Collins, K. (2007) A review of multiple criteria analysis for water resource planning and management. *Water Resources Management*, 21, 1553-1566. https://doi.org/10.1007/s11269-006-9112-5

Hammond, J. S., Keeney, R. L, & Raiffa, H. (2000) *Decisiones inteligentes. Guía práctica para tomar mejores decisiones.* Gestión 2000.

Howard, R. A. (1962) *Dynamic programming and markov processes*. London: The Technology Press of the MIT and John Wiley and Sons, Inc., New York.

Howard, R. A. (1971). *Dynamic probabilistic system.* John Wiley & Sons, Inc.

Keeney, R. L., & Raiffa, H. (1976). *Decisions with multiple objectives: Preferences and value tradeoffs*. John Wiley & Sons.

Keeney, R. L. (1992). *Value-Focused thinking. A path to creative decision making.* London, England: Harvard University Press, Cambridge, Massachusetts, USA.

Otálora, K., Iannacone, L., Gentile, R., & Galasso, C. (2024). Multi-hazard life-cycle consequence analysis of deteriorating engineering systems. *Structural Safety*, 111, 102515. https://doi.org/10.1016/j.strusafe.2024.102515&#8203

Raifffa, H. (1968). *Decision Analysis*. Addison Wesley.

Schlaifer, R. (1969). *Analysis of decisions under uncertainty*. Mcgraw-hill Inc.,USA.

Wang, C., Li, G., Wan, C., Wang, Z., Ju, P., & Lei, S. (2025). Uncertainty-Inflicted vent-Driven resilient recovery for distribution systems: A semi-markov decision process approach. *IEEE Transactions on Power Systems*, 40(1), 368-380. https://doi.org/10.1109/TPWRS.2024.3386851

Cómo citar:

Acosta-Flores, J. J. (2025). Método interactivo para resolver procesos de decisión secuenciales. *Ingeniería Investigación y Tecnología,* 26(02), 1-9. https://doi.org/10.22201fi.25940732e.2025.26.2.011