

Un modelo con información incompleta: aplicación al control
=====

de calidad
=====

F. Benito, IFOR-ETH Zürich

Resumen

En la aplicación de modelos de decisión markovianos se ignora a veces el estado concreto en que se encuentra el sistema, y se dispone únicamente de datos parciales en forma de una distribución de probabilidad en el espacio de estados.

Tales procesos markovianos con información incompleta se pueden transformar en procesos con información completa y discretizar el espacio de estados (informaciones) resultante. Así se consiguen calcular por aproximación las decisiones óptimas.

El método se ilustra con un modelo para el control de calidad con optimación de la política de producción y control.

§ 1 Introducción

En la producción en serie se suele admitir que la calidad del lote producido en un periodo depende del estado de la maquinaria que elabora el producto, y que ese estado puede variar de un periodo a otro debido a perturbaciones aleatorias. Si se considera un espectro discreto de posibles estados de la instalación y de calidades de los lotes producidos, se pueden representar esas variaciones por medio de una cadena de Markov.

Los test de calidad del producto son en muchos casos inexactos y aleatorios, de modo que la calidad actual en cada periodo no se

puede determinar con una certeza del 100%. Esta inseguridad no se refleja en los modelos usuales, y esto implica a menudo una inadmisibile simplificación de la realidad. En caso de varios métodos posibles de control de calidad, se plantea así mismo la pregunta: cuál se debe emplear en cada periodo, para conseguir un compromiso óptimo entre el costo del test y el valor de la información que proporciona ?. Precisemos con un ejemplo.

§ 2 Un ejemplo

En una fábrica de motores eléctricos pesados se producen dos unidades de un determinado modelo por mes. Cada motor puede alcanzar la calidad estándar o no. Un 'buen' ejemplar tiene un precio de venta de 10 Mio. y uno 'malo' sólo de 5 Mio. La materia prima cuesta 1 Mio. por motor. Hay dos modos de producir: simplemente producir (y la calidad del producto será no superior a la del periodo anterior) o, primero revisar la instalación y después producir (para mejorar la calidad con respecto al periodo anterior). Los costos por periodo son respectivamente 1 y 9 Mio. Se pueden someter a pruebas los motores producidos, lo que cuesta 0,2 Mio por motor.

Se busca hacer máxima la suma de las ganancias con un descuento del 1% mensual.

La producción mensual se puede caracterizar por el número de motores que no alcanzan los requerimientos estándar: 0, 1, 2 motores 'malos'. Supongamos que la producción evoluciona de un mes al siguiente de acuerdo con una cadena de Markov con tres estados, dos acciones y las probabilidades de tránsito:

<u>Producir</u>		Número de motores 'malos' en el periodo $t+1$		
		0	1	2
Número de motores 'malos' en el periodo t	0	0.64	0.32	0.04
	1	0.08	0.74	0.18
	2	0.01	0.18	0.81

Tabla 1 Probabilidades de tránsito en caso de producir

<u>Revisar</u>		Número de motores 'malos' en el periodo $t+1$		
		0	1	2
Número de motores 'malos' en el periodo t	0	0.64	0.32	0.04
	1	0.56	0.38	0.06
	2	0.49	0.42	0.09

Tabla 2 Probabilidades de tránsito con revisión previa

Las ganancias dependen, entre otros factores, de los gastos para probar los motores producidos. Hay por tanto que decidir también a este respecto.

Si un mes se decide probar un solo motor, se desconocerá el estado de la cadena de Markov correspondiente a la producción de ese mes. Según el resultado de la prueba se puede descartar uno de los estados extremos (ambos motores buenos o ambos malos), y en general se tienen las siguientes probabilidades condicionales acerca del resultado de la prueba

(2)		Resultado de la prueba	
		bueno	malo
Número de motores malos en la producción del mes	0	1.	0.
	1	0.5	0.5
	2	0.	1.

Tabla 3 Probabilidades de los resultados posibles, probando un solo motor

En caso de probar un solo motor tenemos un modelo de decisión markoviano con información incompleta: no se conoce el estado del sistema, sino que se observa indirectamente con un método determinado.

Si en cada periodo se puede elegir libremente uno de los tres métodos de observación (probar 0, 1, 2 motores) se trata de un modelo de decisión markoviano generalizado. Hay las siguientes seis acciones mixtas:

$$\left. \begin{array}{l} \text{producir} \\ \text{revisar} \end{array} \right\} \times \left\{ \begin{array}{l} \text{probar dos motores} \\ \text{probar un motor} \\ \text{no probar los motores} \end{array} \right.$$

Las ganancias inmediatas son: ganancia en un periodo = precios de venta de los motores producidos - gastos de producción y materias primas - importe de las pruebas.

Para los métodos de observación rigen las siguientes probabilidades condicionales:

(3)		Resultado de la prueba		
		0 malos	1 malo	2 malos
Estado: número de motores malos en la producción del mes	0	1.	0.	0.
	1	0.	1.	0.
	2	0.	0.	1.

Tabla 4 Probabilidades de los resultados posibles probando los dos motores producidos

(1)	Resultado de la prueba ?
Estado: número de motores malos en la producción del mes	0 1.
	1 1.
	2 1.

Tabla 5 El caso de no probar los motores equivale a obtener un resultado ficticio que no depende del estado del sistema

Si se decide probar sistemáticamente los dos motores producidos se obtiene como caso particular un modelo con información completa. Si no se prueban motores es un modelo sin información. La Fig. 1 representa el funcionamiento del sistema, incluida la prueba de motores.

Formalicemos estas consideraciones en un modelo abstracto (vid. (2)).

§ 3 Modelos de Decisión Markovianos con Información Incompleta

El MDM-II del ejemplo se caracteriza con los siguientes parámetros:

$$[Z, E, B, P, M^b, Q^b, \varphi_0, r, \beta]$$

Z = espacio de estados

= { 0, 1, 2 motores malos en la producción del mes } ; se designarán con $z = 1, 2, 3$

E = espacio de acciones

= { producir, revisar } ; estas acciones se designarán con $e = 1, 2$

B = espacio de los métodos de observación

= { no probar motores, probar 1 motor, probar 2 motores } ; se designarán con $b = 1, 2, 3$

P = ley de movimiento

$= \{p_{z',z}^e\}$; $p_{z',z}^e$ = probabilidad del tránsito $z' \rightarrow z$ cuando se elige la acción $e \in E$. Vid. Tablas 1 y 2 para $e = 1, 2$ respectivamente

M^b = conjunto de los resultados posibles de la observación con el método $b \in B$

En el ejemplo para $b = 2$ (probar un motor) es $M^2 = \{\text{bueno, malo}\}$; análogamente para $b=1$ ó 3

Q^b = ley de observación con el método $b \in B$

$= \{q^b(m^b | z)\}$; $q^b(m^b | z)$ = probabilidad de obtener el resultado $m^b \in M^b$ al observar con el método $b \in B$, si el estado

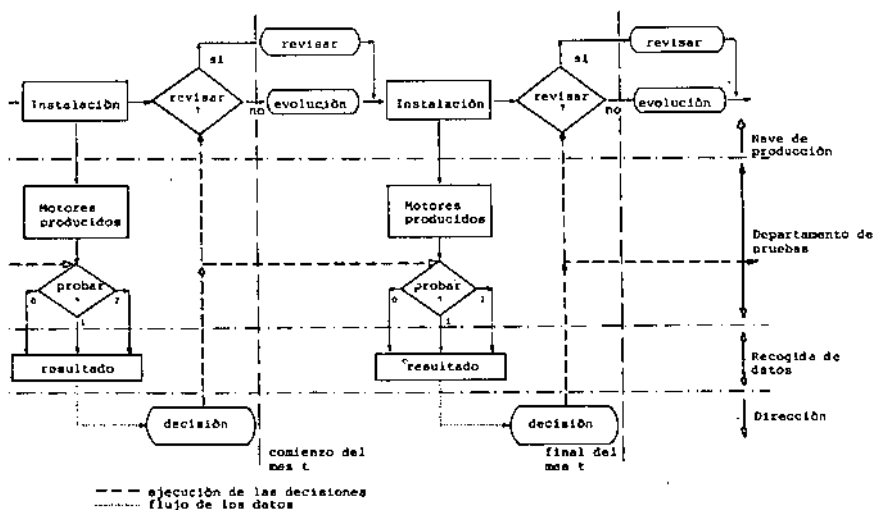


Fig. 1

es $z \in Z$ (vid. tablas 3, 4, 5)

φ_0 = informaci3n inicial

$= \{\varphi_0(1), \varphi_0(2), \varphi_0(3)\}$; distribuci3n de probabilidad en z :

$\varphi_0(z)$ = probabilidad de que al comienzo del primer periodo el estado sea $z \in Z$

r = ganancia inmediata

$= r(z, e, b)$ = precio de venta de la producci3n de calidad z , menos los gastos de producci3n y materia prima con la acci3n $e \in E$ y menos el importe de las pruebas $b \in B$ (vid. tabla 6)

$r(z, e, b)$:	$e = 1$ $b = 1$	$e = 1$ $b = 2$	$e = 1$ $b = 3$	$e = 2$ $b = 1$	$e = 2$ $b = 2$	$e = 2$ $b = 3$
$z = 1$	17.0	16.8	16.6	9.0	8.8	8.6
2	12.0	11.8	11.6	4.0	3.8	3.6
3	7.0	6.8	6.6	-1.0	-1.2	-1.4

Tabla 6

β = factor de descuento ($0 < \beta < 1$)

= valor actual de la cantidad de dinero unidad a recibir transcurrido un periodo (0.99)

La funci3n objetivo a optimar (maximar) en el MDM-II es:

$$E\left[\sum_{t=0}^{\infty} \beta^t \cdot r(z, e, b)\right] = \text{esperanza matemática de la ganancia total con descuento y horizonte ilimitado.}$$

Para un φ_0 dado, con la acci3n $e_0 \in E$ y el método de observaci3n $b_0 \in B$ (elegidos en el momento $t = 0$), obteniendo el resultado de la prueba $m_1 \in M$ (en el momento $t = 1$) la probabilidad de que el nuevo estado (calidad de la producci3n) sea $z \in Z$ es,

aplicando Bayes:

$$(3.1) \quad \varphi_1(z|\varphi_0, e_0, b_0, m_1^{b_0}) = \frac{q^{b_0(m_1^{b_0}|z)} \cdot \sum_{j=1}^3 p_{jz}^{e_0} \cdot \varphi_0(j)}{\sum_{\ell=1}^3 q^{b_0(m_1^{b_0}|\ell)} \cdot \sum_{j=1}^3 p_{j\ell}^{e_0} \cdot \varphi_0(j)}$$

(z = 1, 2, 3)

La suma de las componentes de φ_1 es la unidad. Dados e_0 y b_0 la nueva información φ_1 depende sólo del resultado de observación $m_1^{b_0}$. La probabilidad de obtener este resultado concreto viene medida por el denominador en (3.1); esta cantidad representa la probabilidad del tránsito $\varphi_0 \rightarrow \varphi_1$ en un modelo en que las informaciones juegan el papel de estados.

La fórmula (3.1) se puede generalizar para un periodo t cualquiera:

$$(3.2) \quad \varphi_{t+1}(z|\varphi_t, e_t, b_t, m_{t+1}^{b_t}) = \frac{q^{b_t(m_{t+1}^{b_t}|z)} \cdot \sum_{j=1}^3 p_{jz}^{e_t} \cdot \varphi_t(j)}{\sum_{\ell=1}^3 q^{b_t(m_{t+1}^{b_t}|\ell)} \cdot \sum_{j=1}^3 p_{j\ell}^{e_t} \cdot \varphi_t(j)}$$

(z = 1, 2, 3)

Con esta fórmula se puede conocer en cada periodo el estado del siguiente modelo de decisión markoviano con información completa, derivado del MDM-II:

$$\text{MDM-I}' : [\mathbb{E}, E, B, \bar{P}, (M^b), (Q^b), \varphi_0, r^\varphi, \beta]$$

\mathbb{E} = espacio de 'estados': espacio de las informaciones

$$= \{\varphi | \varphi(z) \geq 0, z \in Z, \sum_{z \in Z} \varphi(z) = 1\}$$

E = espacio de acciones: como en MDM-II

B = espacio de los métodos de observación: como en MDM-II

\bar{P} = ley de movimiento

$= \{\bar{p}_{\varphi', \varphi''}^e\} \cdot \bar{p}_{\varphi', \varphi''}^e$ = probabilidad de obtener la información $\varphi'' \in \mathcal{I}$ en el periodo $t+1$ partiendo de la información $\varphi' \in \mathcal{I}$ en el periodo t y habiendo elegido en ese periodo la acción $e \in E$. Los valores numéricos se calculan con la expresión del denominador en (3.2)

M^b = conjunto de los resultados posibles de la observación con el método $b \in B$: como en MDM-II. En el MDM-I' los resultados m^b sirven sólo para actualizar la información; por eso se ha escrito (M^b) entre paréntesis

Q^b = ley de observación con el método $b \in B$: como en MDM-II. Sirve sólo para actualizar la información y calcular la ley de movimiento \bar{P} , por lo que se ha escrito (Q^b)

φ_0 = 'estado' inicial: como en MDM-II

r^φ = ganancia inmediata

$$= r^\varphi(\varphi, e, b) = \sum_{z=1}^3 r(z, e, b) \cdot \varphi(z)$$

= esperanza matemática de la ganancia en un periodo con $e \in E$ y $b \in B$ dada la información (distribución) $\varphi \in \mathcal{I}$

β = factor de descuento: como en MDM-II

Como en el modelo derivado MDM-I' la ganancia en cada periodo es igual a la esperanza matemática de la ganancia en ese periodo para el MDM-II primitivo con la información disponible, podemos prescindir del MDM-II y optimar directamente en el MDM-I' la función objetivo:

$$E \left[\sum_{t=0}^{\infty} \beta^t \cdot r^\varphi(\varphi, e, b) \right]$$

Este MDM-I' cumple las condiciones de Blackwell (3), y posee por tanto una política óptima estacionaria (independiente del tiempo).

La existencia de tal política óptima no es de utilidad en la práctica ya que el espacio de 'estados' Φ tiene la potencia del continuo y no hay algoritmos prácticos para obtener dicha política. Como procedimiento utilizable se discretiza Φ , optimando en el modelo resultante.

§ 4 Discretización: MDM-I' \Rightarrow MDM-I*

$$\text{MDM-I*}: [\Phi^*, E, B, \tilde{P}, (M^b), (Q^b), \varphi_0, r^\varphi, \beta]$$

Los parámetros de este modelo son ya conocidos salvo Φ^* y \tilde{P} . Para un número entero fijo $d \in N$ ($d \geq 2$) se define

$$\Phi^* = \{\varphi^* | \varphi^*(z) = d_z \cdot \frac{1}{d}, d_z \in N, z \in Z, \sum_{z \in Z} d_z = d\}$$

Los elementos (en número finito) de Φ^* son los vértices de una retícula en el simplex Φ . Para nuestro ejemplo con $d = 3$:

$$\Phi^* = \{\varphi_1^*, \varphi_2^*, \dots, \varphi_{10}^*\} \quad (\text{vid. Fig. 2})$$

Cualquier elemento de Φ se puede expresar de modo único como combinación convexa de tres (tantos como elementos tiene Z) elementos determinados de Φ^* (los vértices del triangulito que lo contiene: vid. Fig. 3): $\varphi = \rho_6 \varphi_6^* + \rho_7 \varphi_7^* + \rho_9 \varphi_9^*$

con $\rho_6, \rho_7, \rho_9 \geq 0, \rho_6 + \rho_7 + \rho_9 = 1$.

Los coeficientes ρ_i dependen sólo de φ : $\rho_i(\varphi)$. Esta representación de los elementos de Φ se puede utilizar para calcular las probabilidades de tránsito entre elementos de Φ .

Como indica la Fig. 4 antes del tránsito se representa $\varphi \in \Phi$ como combinación de tres vértices, se obtiene el elemento imagen de cada vértice, y este de nuevo se representa como una combinación de elementos de Φ^* ; sea

$$(4.1) \quad \varphi = \rho_6(\varphi) \cdot \varphi_6^* + \rho_7(\varphi) \cdot \varphi_7^* + \rho_9(\varphi) \cdot \varphi_9^*$$

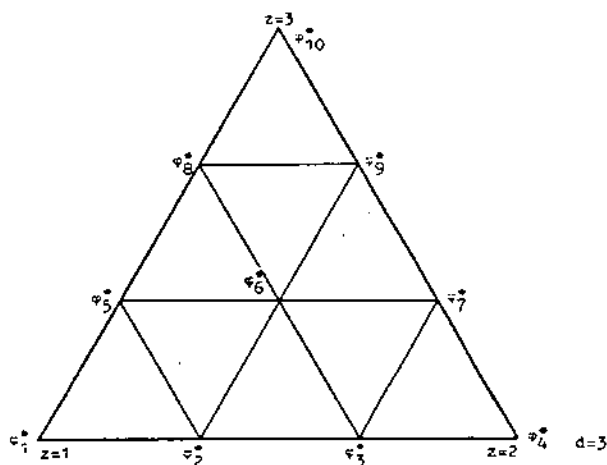


Fig. 2 Los puntos del triángulo representan los elementos de Φ y los vértices los de Φ^*

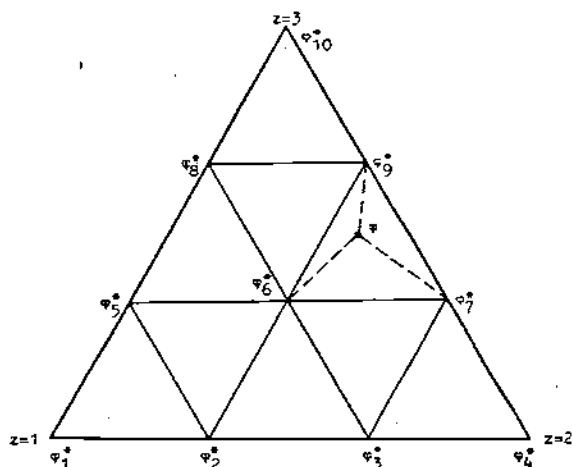


Fig. 3

El elemento imagen φ'_{m_e} de φ no se precisa calcular explícitamente ya que se cumple (vid. Fig. 4):

$$(4.2) \quad \bar{p}_{\varphi\varphi'}^e \cdot \varphi'_{m_e} = \rho_6(\varphi) \cdot \bar{p}_{\varphi_6^* \varphi_6^{**}}^e \cdot \varphi_6^{**} + \\ + \rho_7(\varphi) \cdot \bar{p}_{\varphi_7^* \varphi_7^{**}}^e \cdot \varphi_7^{**} + \rho_9(\varphi) \cdot \bar{p}_{\varphi_9^* \varphi_9^{**}}^e \cdot \varphi_9^{**}$$

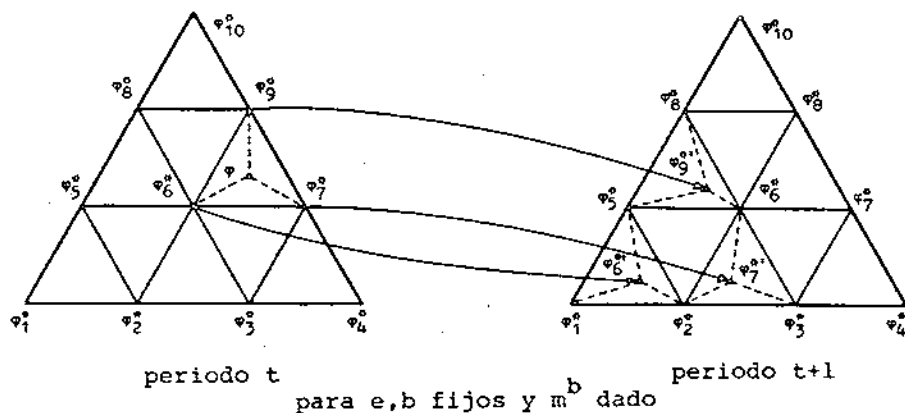


Fig. 4

A la derecha en la Fig. 4 sean las combinaciones:

$$\varphi_6^{*'} = p_1(\varphi_6^{*'}) \cdot \varphi_1^* + p_2(\varphi_6^{*'}) \cdot \varphi_2^* + p_5(\varphi_6^{*'}) \cdot \varphi_5^*$$

m^b_e m^b_e m^b_e m^b_e

$$\varphi_7^{*'} = p_2(\varphi_7^{*'}) \cdot \varphi_2^* + p_3(\varphi_7^{*'}) \cdot \varphi_3^* + p_6(\varphi_7^{*'}) \cdot \varphi_6^*$$

m^b_e m^b_e m^b_e m^b_e

$$\varphi_9^{*'} = p_5(\varphi_9^{*'}) \cdot \varphi_5^* + p_6(\varphi_9^{*'}) \cdot \varphi_6^* + p_8(\varphi_9^{*'}) \cdot \varphi_8^*$$

m^b_e m^b_e m^b_e m^b_e

Sustituyendo esto en (4.2) aparecen sólo elementos de Φ^* y probabilidades de tránsito entre elementos de Φ^* y sus imágenes (posiblemente no pertenecientes a Φ^*):

$$\tilde{P} = \{ \bar{p}_{\varphi^* \varphi^{*'}}^e \mid e \in E, b \in B, m^b \in M^b, \varphi^* \in \Phi^* \}$$

\tilde{P} es finito.

La ecuación que caracteriza los valores óptimos en el MDM-I' es:

$$(4.3) \quad V(\varphi) = \max_{\substack{e \in E \\ b \in B}} \{ r^e(\varphi, e, b) + \beta \sum_{\varphi' \in \Phi} \bar{p}_{\varphi \varphi'}^e \cdot V(\varphi') \}$$

Haciendo la aproximación:

$$V(\sum_j \rho_j(\varphi_{i1}^{*'}) \cdot \varphi_j^*) = \sum_j \rho_j(\varphi_{i1}^{*'}) \cdot \hat{V}(\varphi_j^*)$$

y teniendo en cuenta

$$\varphi = \sum_i \rho_i(\varphi) \cdot \varphi_i^*, \quad \rho_i(\varphi) \geq 0, \quad \sum_i \rho_i(\varphi) = 1$$

$$r^\varphi(\varphi, e, b) = \sum_i \rho_i(\varphi) \cdot r^\varphi(\varphi_i^*, e, b) \quad (\text{linear en } \varphi)$$

obtenemos de (4.3) con (4.2):

$$(4.4) \quad \hat{V}(\varphi) = \max_{\substack{e \in \bar{E} \\ b \in \bar{B}}} \left\{ \sum_i \rho_i(\varphi) \cdot r^\varphi(\varphi_i^*, e, b) + \beta \sum_{m^b \in M^b} \sum_i \rho_i(\varphi) \cdot \bar{p}_{\varphi_i^* \varphi_{i1}^{*'}}^e \cdot \sum_j \rho_j(\varphi_{i1}^{*'}) \cdot \hat{V}(\varphi_j^*) \right\}$$

La aproximación utilizada -una expresión lineal para la función convexa $V(\varphi)$ - origina un error que se puede acotar (vid. Benito (1))

$$(4.5) \quad 0 \leq \hat{V}(\varphi) - V(\varphi) \leq \beta \frac{1}{d} \max \{O_Z - U_Z\} \cdot \left[2 + \frac{n-1}{n(1-\beta)} \right] \xrightarrow{d \rightarrow \infty} 0$$

d = índice de la discretización (vid. definición de Φ^*)

n = número de elementos de Z

O_Z (resp. U_Z) = valor del estado $z \in Z$ en el primitivo MDM-II caso de maximizar (resp. minimar) suponiendo información completa

Para los elementos de \tilde{P} se cumple:

$$\bar{p}_{\varphi_i^* \varphi_{i1}^{*'}}^e = \sum_{m^b} q^b(m^b | z') \cdot \sum_{z \in Z} p_{zz'}^e \cdot \varphi_i^*(z)$$

Los $\varphi_{i1}^{*'} \cdot m^b$ se calculan según (3.2) e interesan sólo para calcular los coeficientes $\rho_j(\varphi_{i1}^{*'}) \cdot m^b$ en (4.4).

§ 5 Algoritmos

Este modo de proceder (§§ 3 y 4) se puede describir:

Algoritmo I: elaboración del modelo derivado y discretización

① Elijase un número entero d y construyase el conjunto Φ^* de vectores con n componentes:

$$\Phi^* = \{\varphi_i^* \mid \varphi_i^*(j) = d_j \cdot \frac{1}{d}, d_j \in N, j = 1, \dots, n, \sum_{j=1}^n d_j = d\}$$

Sea L el número de elementos de Φ^* ($i=1, \dots, L$)

② Para cada decisión mixta $(e, b) \in E \times B$, cada resultado de observación $m^b \in M^b$ y cada $\varphi_i^* \in \Phi^*$ calcúlense las imágenes

$$\varphi_{i, m^b e}^{*'}(j) = \frac{q^b(m^b | j) \cdot \sum_{z=1}^n \varphi_i^*(z) \cdot p_{z,j}^e}{\sum_{z'=1}^n q^b(m^b | z') \cdot \sum_{z=1}^n \varphi_i^*(z) \cdot p_{z,z'}^e}, \quad j = 1, \dots, n$$

y los valores

$$\bar{p}_{\varphi_i^{*'}, m^b e}^e = \sum_{z=1}^n q^b(m^b | z') \cdot \sum_{z=1}^n \varphi_i^*(z) \cdot p_{z,z'}^e$$

③ Para cada $\varphi_{i, m^b e}^{*'}$ determinense los n elementos φ_j^* de Φ^* vértices del n -poliedro que lo contiene y los coeficientes $\rho_j(\varphi_{i, m^b e}^{*'})$ de la combinación convexa

$$\varphi_{i, m^b e}^{*'} = \sum_j \rho_j(\varphi_{i, m^b e}^{*'}) \cdot \varphi_j^*$$

con $\rho_j(\varphi_{i, m^b e}^{*'}) \geq 0$, $\sum_j \rho_j(\varphi_{i, m^b e}^{*'}) = 1$, $j \in \{1, \dots, L\}$ y n suman-

dos como máximo

④ Cálculense para cada $\varphi_1^* \in \Phi^*$ y cada decisión mixta $(e,b) \in \text{ExB}$

$$r^\varphi(\varphi_1^*, e, b) = \sum_{z=1}^n \varphi_1^*(z) \cdot r(z, e, b).$$

La optimación en el MDM-I*, resolviendo la ecuación (4.4), se lleva a cabo con un algoritmo del tipo de Howard (4) o por programación lineal -vid. por ejemplo Mine (5)- utilizando como espacio de acciones el producto cartesiano ExB.

En el ejemplo la política óptima que se obtiene es

Política óptima

'estado' (información)	Producir o revisar	Número de motores a probar	Esperanza matemática de la ganancia total descontada: $\bar{V}(\varphi_1^*)$
φ_1^*	producir	0	1165.52
φ_2^*	producir	2	1160.89
φ_3^*	producir	2	1156.41
φ_4^*	producir	0	1152.02
φ_5^*	producir	1	1157.50
φ_6^*	producir	2	1152.98
φ_7^*	revisar	0	1149.17
φ_8^*	producir	1	1149.61
φ_9^*	revisar	0	1147.09
φ_{10}^*	revisar	2	1145.03

Tabla 7 $B = \{1, 2, 3\}$

§ 6 Uso de la política óptima del MDM-I*

La política óptima del MDM-I* se debe ampliar al MDM-I': hay que obtener una regla para las decisiones correspondientes también a los elementos de $\Phi - \Phi^*$.

Algoritmo II: ampliación de la política

① Sea $\varphi \in \Phi - \Phi^*$; determinense los n elementos $\varphi_i^* \in \Phi^*$, vértices del mínimo n -poliedro que lo contiene y los coeficientes $\rho_i(\varphi)$ de la combinación convexa

$$\varphi = \sum_i \rho_i(\varphi) \cdot \varphi_i^*$$

con $\rho_i(\varphi) \geq 0$, $\sum_i \rho_i(\varphi) = 1$, $i \in \{1, \dots, L\}$ (i toma como máximo n valores distintos)

② Cálculense los valores test:

$$\begin{aligned} TE(\bar{e}, \bar{b}) = & \sum \rho_i(\varphi) \cdot r^\varphi(\varphi_i^*, \bar{e}, \bar{b}) + \\ & + \beta \sum_{m \in M} \sum_b \sum_i \rho_i(\varphi) \cdot \bar{p}_{\varphi_i^*, \varphi_i^{*'}}^{\bar{e}} \cdot \sum_{m \in M} \sum_{\bar{e}} \rho_j(\varphi_i^{*'}) \cdot \bar{V}(\varphi_j^*), \end{aligned}$$

para todos los $(\bar{e}, \bar{b}) \in \text{ExB}$ donde los $\bar{p}_{\varphi_i^*, \varphi_i^{*'}}^{\bar{e}}$ y $\rho_j(\varphi_i^{*'})$ proceden del algoritmo I y $\bar{V}(\varphi_i^*)$ son los valores de los 'estados' del MDM-I* con la política óptima

③ Determinense los (\hat{e}, \hat{b}) tal que

$$TE(\hat{e}, \hat{b}) = \max_{\substack{\bar{e} \in E \\ \bar{b} \in B}} TE(\bar{e}, \bar{b})$$

Estas son decisiones mixtas (equivalentes entre si en caso de ser varias) a usar para φ , y se cumple $\hat{V}(\varphi) = TE(\hat{e}, \hat{b})$

Al aplicar la política en cada periodo se debe ir actualizando la información de acuerdo con (3.2).

Es de notar que a pesar del error en los valores de los estados (en general $V(\varphi) \neq \hat{V}(\varphi)$) la política obtenida puede coincidir con la óptima, cuya existencia garantiza Blackwell(3).

§ 7 Generalización a observaciones inseguras

En el ejemplo de la fabricación de motores hemos supuesto que las pruebas dan siempre un resultado correcto: un motor que se prueba y se califica como 'bueno' (resp. 'malo') lo es realmente.

Amplíemos ahora el ejemplo admitiendo una prueba rápida de los motores que a un buen motor lo califica como malo en el 10% de las veces, y reciprocamente un motor malo lo da el 20% de las veces como bueno. El importe de esta prueba rápida sea 0,1 Mio por motor.

Tenemos así en el MDM-II dos nuevos métodos de observación: someter a la prueba rápida uno o dos motores de la producción mensual (supongamos que por motivos técnicos no cabe utilizar la prueba rápida y la exacta en el mismo periodo).

Las leyes de observación aparecen en los cuadros siguientes.

(b=4) Hacer la prueba rápida de un motor Probabilidades de los resultados:			
		Resultado de la prueba rápida	
		$m^{b=1}$ bueno	$m^{b=2}$ malo
Estado: número de motores malos en la producción	0	0.90	0.10
	1	0.55	0.45
	2	0.20	0.80

Con el algoritmo I y la optimización en el MDM-I* se obtienen para el modelo con cinco métodos de observación los resultados de la tabla 8.

La posibilidad de someter ambos motores a la prueba rápida no se utiliza nunca en la política obtenida (optimando con los métodos de observación $b = 1-4$ se obtienen los mismos resultados de la tabla 8).

(b=5) Hacer la prueba rápida de dos motores Probabilidades de los resultados:				
		Resultados de las pruebas rápidas		
		$m^b=1$ ninguno malo	$m^b=2$ 1 malo	$m^b=3$ 2 malos
Estado: número de motores malos en la producción	0	0.81	0.18	0.01
	1	0.18	0.74	0.08
	2	0.04	0.32	0.64

Política óptima (cinco métodos de observación)				
'estado' (información)	Producir o revisar	Tipo de prueba	Número de mo- tores	Esperanza matemática de la ganancia total descontada
φ_1^*	producir	---	0	1166.45
φ_2^*	producir	rápida	1	1161.83
φ_3^*	producir	exacta	2	1157.31
φ_4^*	producir	---	0	1152.92
φ_5^*	producir	exacta	1	1158.42
φ_6^*	producir	exacta	2	1153.88
φ_7^*	revisar	---	0	1150.09
φ_8^*	producir	exacta	1	1150.53
φ_9^*	revisar	rápida	1	1148.01
φ_{10}^*	revisar	rápida	1	1145.94

Tabla 8 $B = \{1,2,3,4,5\}$

Las ganancias esperadas en la tabla 8 son superiores a las de la tabla 7. El aumento (un 0,08% en promedio) indica una cota superior de la inversión a hacer para introducir la prueba rápida de motores (valor marginal de los nuevos métodos de obser

Política óptima sin observación (no se prueban motores)		
'estado' (Información)	Producir o revisar	Esperanza matemática de la ganancia total con descuento
φ_1^*	producir	1138.54
φ_2^*	producir	1133.79
φ_3^*	producir	1129.40
φ_4^*	producir	1125.32
φ_5^*	producir	1128.85
φ_6^*	producir	1124.70
φ_7^*	revisar	1122.12
φ_8^*	revisar	1122.11
φ_9^*	revisar	1120.00
φ_{10}^*	revisar	1117.89

Tabla 9 $B = \{1\}$

Política óptima con observación directa (se prueban exactamente ambos motores)		
'estado' (Información)	Producir o revisar	Esperanza matemática de la ganancia total con descuento
φ_1^*	producir	1155.54
φ_2^*	producir	1151.08
φ_3^*	producir	1146.62
φ_4^*	producir	1142.16
φ_5^*	producir	1147.66
φ_6^*	producir	1143.20
φ_7^*	revisar	1139.30
φ_8^*	producir	1139.78
φ_9^*	revisar	1137.25
φ_{10}^*	revisar	1135.21

Tabla 10 $B = \{3\}$

vación).

Tienen un interés especial los casos sin observación (en el ejemplo equivale a admitir sólo el método de observación $b = 1$: no hacer pruebas de motores) y con observación directa (es decir, único método de observación admisible $b = 3$: prueba exacta de ambos motores: información completa).

Los resultados para estos casos aparecen en las tablas 9 y 10. La diferencia entre las ganancias esperadas da el valor marginal de la observación directa.

'estado' (información)	$B=\{3\}$	$B=\{1,2,3\}$	$B=\{1,2,3,4,5\}$
	Valor marginal de la observación directa	Valor marginal de los métodos de observación	Valor marginal de los métodos de observación
0			
φ_1^*	17.00	26.98	27.91
φ_2^*	17.29	27.10	28.04
φ_3^*	17.22	27.01	27.91
φ_4^*	16.84	26.70	27.60
φ_5^*	18.81	28.65	29.57
φ_6^*	18.50	28.28	29.18
φ_7^*	17.18	27.05	27.97
φ_8^*	17.67	27.50	28.42
φ_9^*	17.25	27.09	28.01
φ_{10}^*	17.32	27.14	28.05

Tabla 11 Esperanza matemática de las ganancias marginales acumuladas con descuento, frente al caso sin observación

Analogamente de las tablas 9 y 7 se desprende el valor marginal de los métodos de observación para el modelo del § 2 ($B=\{1,2,3\}$).

La tabla 11 resume estos valores. Multiplicando estas cantidades con $(1-\beta)$ se obtiene la ganancia marginal equivalente por periodo.

§ 8 Aplicaciones

Los MDM-II con varios métodos de observación encuentran aplicación en numerosos campos aparte del control de calidad: problemas de mantenimiento, concurrencia en mercados, procesos de aprendizaje, decodificación de canales de información, diagnóstico y tratamiento clínico, problemas de búsqueda de un objeto móvil, etc.

En la tabla 12 se indican algunos ejemplos.

Algunos campos de aplicación de MDM-II		
Campo	Estados del modelo	Resultados de la observación
Control de calidad	Calidad producida	Resultado de la muestra
Mantenimiento	Estado de la máquina	Resultado de la inspección
Aprendizaje	Nivel alcanzado	Respuesta a un test
Decodificación de canales de información	Estado del canal	Output con ruido
Tratamiento clínico	Estado fisiológico del paciente	Reacción a la medicación
Búsqueda de un objeto móvil	Ubicación actual del objeto	Resultado de los sondeos

Tabla 12

Referencias de aplicaciones concretas en estos campos se encuentran en Benito (1).

Bibliografia

- (1) Benito, F., Ein Modell der Qualitätskontrolle in der Serienproduktion bei kostspieligen Tests, Diss. ETH, Zürich, 1975
- (2) Benito, F., Markov-Modelle mit indirekter Beobachtung, Haupt Verlag, Bern, 1978
- (3) Blackwell, D., Discounted dynamic programming, Ann. Math. Stat, 36, 226-235, 1965
- (4) Howard, R.A., Dynamic Programming and Markov Processes, Wiley, New York, 1960
- (5) Mine, H., Osaki, S., Markovian Decision Processes, Elsevier, New York, 1970