

CONTROL 3

1. En un secreto laboratorio de biotecnología están sufriendo sucesivos apagones de luz debido a problemas de desgaste del generador de energía (su única fuente de electricidad). Estos apagones han provocado pérdidas en ciertos cultivos de bacterias. El director del laboratorio ha decidido remediar esta situación, para lo cual se ha comunicado con el equipo de IO de la universidad local. Este equipo ha sido capaz de determinar que los tiempos de falla de los generadores siguen una distribución F de media $\mu < \infty$ y, además, han decidido proponer una política preventiva según la cual, un generador es reemplazado por otro nuevo en el instante en que falle o cuando cumpla T unidades de tiempo funcionando, lo que ocurra primero, donde $T > 0$ es un parámetro a determinar.

Suponga que cada vez que se debe reemplazar un generador (ya sea por falla o por que superó el tiempo T) se incurre en un costo K y, además, cuando el generador falla, se paga un costo adicional C , debido a la alteración de algunos procesos de producción. Cada vez que se instala un nuevo generador, éste comienza de inmediato a producir energía pero, durante un tiempo aleatorio S , con ley uniforme en $[0, T/2]$, se debe realizar una serie de ajustes que implican irregularidad en el voltaje. A causa de esto se producen pérdidas de producción, de valor proporcional (con coeficiente α) al tiempo de ajuste.

- Cuál es la distribución $F_T(x)$ que corresponde al tiempo entre sucesivos reemplazos de generadores? Cuál es la media μ_T ?
- En el largo plazo, cuántos reemplazos son hechos por unidad de tiempo?
- Cuál es el tiempo esperado entre fallas? En el largo plazo, cuántas fallas ocurren por unidad de tiempo?
- Obtenga una expresión para $C(T)$, el costo promedio en el largo plazo por unidad de tiempo, como función del parámetro T .
- Si $K = 2$, $C = 1$, $\alpha = 8$ y los tiempos de vida de los generadores son variables aleatorias uniformes en $[0, 1]$, obtenga el tiempo T^* que minimiza $C(T)$.
- Calcule el tiempo esperado hasta que se produzca la situación de crisis del laboratorio debido a que 3 generadores seguidos¹ fueron reemplazados por fallar, usando los valores de las constantes del apartado f) y el valor óptimo T^* .

Sol:

- Si X designa la va correspondiente a la vida de un generador, entonces el tiempo hasta el reemplazo es $X \wedge T := \min\{X, T\}$. Por lo tanto,

$$F_T(t) = 1 - P[\min\{X, T\} > t] = 1 - P[X > t, T > t] = 1 - P[X > t] \mathbf{1}_{\{T > t\}} = 1 - (1 - F(t)) \mathbf{1}_{\{T > t\}}.$$

Reorganizando los términos también se puede escribir

$$F_T(t) = \mathbf{1}_{\{T \leq t\}} + F(t) \mathbf{1}_{\{T > t\}}.$$

Es claro que F_T es discontinua, con un átomo en T y $\Delta F_T(T) = P[X < T] = F(T^-)$. La media se calcula usando la fórmula

$$\mu_T = \int_0^\infty (1 - F_T(t)) dt = \int_0^T (1 - F(t)) dt.$$

- Por el teorema elemental de renovación, el valor es μ_T^{-1} .
- Después de una falla comienza una serie de reemplazos de tipo “no falla”, en cantidad N , donde N es una va geométrica de parámetro² $p = F(T^-) = P[X < T]$, $q := 1 - p$, comienzan-

¹Modele esta situación como la aparición de un patrón en una serie de lanzamientos de una moneda.

²Aquí estamos suponiendo que si falla justo en T entonces no cuenta como falla.

do en 0. Esto porque los sucesivos reemplazos pueden verse como experimentos de Bernoulli independientes (cara=falla, sello= no falla). Entonces $P[N = n] = q^n p, n = 0, 1, \dots$ y $E[N] = q/p = P[X \geq T]/P[X < T]$. El tiempo que transcurre desde una falla hasta la proxima es igual a NT más el tiempo de vida del generador que se sabe va a fallar, dado por $X|X < T$, entonces la esperanza se calcula como

$$\mu_{ef} := T \frac{P[X \geq T]}{P[X < T]} + E[X|X < T] = T \frac{1 - F(T^-)}{F(T^-)} + E[X|X < T].$$

Por otra parte,

$$P[X > t|X < T] = P[t < X < T]/F(T^-) = \frac{F(T^-) - F(t)}{F(T^-)} = 1 - \frac{F(t)}{F(T^-)}, \quad t < T$$

y, por lo tanto,

$$E[X|X < T] = \int_0^\infty P[X > t|X < T] dt = \int_0^T \left(1 - \frac{F(t)}{F(T^-)}\right) dt = \frac{\mu_T - T(1 - F(T^-))}{F(T^-)}.$$

Combinando ambos resultados parciales, se llega a $\mu_{ef} = \mu_T / F(T^-)$.

Para calcular el número de fallas por unidad de tiempo, en el largo plazo, notamos que en un intervalo entre fallas, sólo contamos una falla, de manera que, del resultado para renovación con recompensa, obtenemos $1/\mu_{ef} = F(T^-)/\mu_T$.

Un razonamiento alternativo para calcular este último resultado es considerar la renovación con las va de tipo $X \wedge T$ y decir que recibimos una recompensa de valor 1, con probabilidad $F(T^-)$. Entonces, las recompensas son iid con esperanza $F(T^-)$ y de renovación con recompensa se obtiene $F(T^-)/\mu_T$. De aquí se deduce por cierto el valor de μ_{ef} .

- d) Veamos el costo a pagar en un intervalo de renovación, del tipo $X \wedge T$. Se paga seguro el costo K de un generador y se paga además C si ha fallado, lo cual suma $K + C\mathbf{1}_{\{X < T\}}$. Al lo anterior se agrega la perdida debida al tiempo S de ajuste, que se escribe αS . Sin embargo, podría ocurrir que el generador falle antes de que termine el período de ajuste, de manera que deberíamos cambiar³ αS por $\alpha(S \wedge X)$. El valor esperado del costo en un intervalo de renovación sería

$$E_{costo} = E[K + C\mathbf{1}_{\{X < T\}} + \alpha(S \wedge X)] = K + CF(T^-) + \alpha E(S \wedge X)$$

o bien

$$E'_{costo} = E[K + C\mathbf{1}_{\{X < T\}} + \alpha S] = K + CF(T^-) + \alpha T/2,$$

si ignoramos la posibilidad de que el generador falle antes de terminar el ajuste. Para calcular el costo por unidad de tiempo, en cualquiera de los casos, dividimos por el largo esperado del intervalo de renovación, dado por μ_T .

- e) Consideramos el caso $K = 2, C = 1, \alpha = 8$ y $X \rightarrow U[0, 1]$, de donde se obtiene $\mu_T = T(1 - T/2)$. De manera que

$$C(T) = \frac{2 + T + 8E(S \wedge X)}{T(1 - T/2)}$$

y

$$C'(T) = \frac{2 + 3T}{T(1 - T/2)}.$$

Notemos que (suponiendo S y X independientes)

$$E(S \wedge X) = \int_0^{T/2} \mu_s ds = \int_0^{T/2} s(1 - s/2) ds = \frac{T^2}{8} \left(1 - \frac{T}{6}\right).$$

³Aceptar cualquier versión como solución correcta.

Entonces

$$C(T) = \frac{2 + T + T^2(1 - T/6)}{T(1 - T/2)}.$$

Para encontrar T'^* vemos fácilmente que C' se minimiza en $T'^* = 2/3$. El caso de C es algo más complicado y se observa numéricamente que $T^* \sim 0,685$, aunque los valores $C'(T'^*)$ y $C(T^*)$ difieren bastante.

f) La crisis se produce cuando vemos el patrón 111, en lanzamientos iid de una moneda, con probabilidad de cara (1) dada por $p = F(T^-)$. El tiempo esperado hasta ver dicho patrón es $1/p + 1/p^2 + 1/p^3$ (ver solución en libro de Ross, pág. 126). Por ejemplo, si X es $U[0, 1]$ resulta $p = T \in [0, 1]$ y con $T = 1/2$ el tiempo esperado es 14; con $T = 3/4$ se obtiene 5,48 y con $T = 1/4$, el valor es 84.

2. W. Armijo ha trabajado por años en gestión de operaciones, habiendo logrado acumular una cuantiosa fortuna. Dado que Armijo se ha interesado siempre por la educación superior, decide invertir su capital en una universidad privada, a punto de irse a pique.

Armijo entiende que el buen manejo de la universidad requiere de un modelo estocástico para su gestión (inicialmente muy sencillo), cuyos elementos se describen a continuación. El estado de la universidad en el año n se caracteriza por las va X_n e Y_n , donde X_n indica la calidad e Y_n la acreditación. La calidad puede ser buena (1) o mala (0) mientras que la acreditación es lograda (1) o no lograda (0). Estas dos variables están estrechamente ligadas a los resultados de matrículas y por ende, a los ingresos que se obtienen por aranceles. Si $X_n = i, Y_n = j$ entonces el ingreso esperado por aranceles el año n es R_{ij} , con $i, j \in \{0, 1\}$. Estudios previos indican que $R_{00} < R_{01} < R_{10} < R_{11}$.

Cada año el propietario o rector de la universidad puede tomar una de tres decisiones respecto de las utilidades que haya obtenido por aranceles: (0) no hacer inversión alguna, lo cual significa traspasar dinero a sociedades relacionadas, etc. y tiene impacto negativo en la calidad y en la probabilidad de ser acreditada; (1) invertir apropiadamente, lo cual tiene un positivo impacto en la calidad de la universidad y en la probabilidad de ser acreditada y, finalmente, (2) no invertir pero gastar algo de dinero en sobornar a los funcionarios de la comisión de acreditación, lo cual tiene como efecto aumentar las probabilidades de obtener la acreditación pero, a la vez, expone a la universidad a fuertes sanciones en caso de ser descubierto el soborno. Suponga que la inversión anual es una cantidad I , no aleatoria, igual para cada año que se decida invertir. Por otra parte, el gasto en soborno es una cantidad S (muy inferior a I), no aleatoria, igual para cada año que se decida sobornar. Se sabe también que la universidad puede ser denunciada, con probabilidad p , en los medios y en tribunales, por un celoso guardián del sistema de acreditación, Sr. Patrick Glass. Si la denuncia ocurre pero la universidad no ha sobornado entonces se incurre en un costo C (publicidad, defensa, etc.) Pero, si ha habido soborno, la universidad debe pagar una fuerte multa M (muy superior a C) y se le niega la acreditación para el año correspondiente.

Cuando la universidad tiene buena calidad, nunca se recurre al soborno pero esta posibilidad si está abierta en caso contrario, cualquiera sea la acreditación previa que tenga.

Respecto a las probabilidades de transición del sistema se ha podido averiguar los siguiente:

- Si no invierte, la calidad no puede subir.
- Si no invierte y tiene mala calidad, entonces no será acreditada, cualquiera haya sido la acreditación del período previo.
- Si invierte, la calidad no puede bajar.
- Si invierte, la calidad es buena y está acreditada, entonces será acreditada con seguridad.
- Si soborna (lo que indica mala calidad) entonces la calidad no puede subir (puesto que no invierte) pero si podría ser acreditada, cualquiera sea la acreditación que tenga en el presente. Por supuesto, es más fácil ser acreditada si es que ya lo estaba.

Para las probabilidades, en general, se introducen los siguientes parámetros:

- α_i^a es la probabilidad de que teniendo calidad i y la acción sea a , en el siguiente período tenga calidad 1, $i \in \{0, 1\}, a \in \{0, 1, 2\}$.
- β_j^a es la probabilidad de que teniendo acreditación j y la acción sea a , en el siguiente período tenga acreditación 1, $j \in \{0, 1\}, a \in \{0, 1, 2\}$.

Con los 12 parámetros definidos (algunos de los cuales pueden ser 0 o 1) es posible calcular las probabilidades de transición. Por ejemplo, la probabilidad de pasar de buena calidad, no acreditada a buena calidad y acreditada, sin invertir es $\alpha_1^0\beta_0^0$. Más rigurosamente,

$$P[X_{n+1} = 1|X_n = 1, Y_n = 0, a = 0] = \alpha_1^0,$$

$$P[Y_{n+1} = 1|X_n = 1, Y_n = 0, a = 0] = \beta_0^0$$

y

$$P[X_{n+1} = 1, Y_{n+1} = 1|X_n = 1, Y_n = 0, a = 0] = \alpha_1^0\beta_0^0.$$

Lo anterior indica que sabiendo a y condicionalmente en X_n, Y_n , las va X_{n+1} e Y_{n+1} son independientes.

Finalmente, si el horizonte es N , se supone que el valor final de venta de la universidad es F_{ij} , dependiendo de la calidad y la acreditación que tenga. Se puede suponer que $F_{00} < F_{01} < F_{10} < F_{11}$.

- Formalizar la situación descrita más arriba como un Problema o Proceso de Decisión Markoviano, identificando claramente sus elementos: espacio de estados, acciones, recompensas, probabilidades de transición⁴ y función objetivo a optimizar.
- Escriba las ecuaciones de optimalidad de Bellman correspondientes al problema en a).
- Asigne valores sencillos a los parámetros, compatibles con la descripción hecha arriba, y realice una iteración con las ecuaciones de Bellman, indicando cual sería la decisión y el valor objetivo.

Sol:

- El proceso está descrito por una sucesión $(Z_n)_{n \geq 1}$ de pares de va con valores en $\{0, 1\}$. Es decir, $Z_n = (X_n, Y_n) \in E := \{0, 1\}^2$. Por otra parte, el espacio de decisiones es $A = \{0, 1, 2\}$, donde 0,1,2 tienen el significado descrito en el enunciado.

Sabiendo que las decisiones disponibles dependen del estado del sistema, consideramos los conjuntos A_{ij} , que corresponden a las decisiones posibles cuando el estado del sistema es (i, j) . Tenemos:

$$A_{00} = A, A_{01} = A, A_{10} = \{0, 1\}, A_{11} = \{0, 1\}.$$

Para las recompensas basta especificar la función $r(i, j, a)$ puesto que las cantidades descritas tales como ingresos, costos, multas, etc. son constantes en el tiempo. Tenemos:

0) $a = 0$:

- $r(0, 0, 0) = R_{00} - pC$,
- $r(0, 1, 0) = R_{01} - pC$,
- $r(1, 0, 0) = R_{10} - pC$,
- $r(1, 1, 0) = R_{11} - pC$.

1) $a = 1$:

- $r(0, 0, 1) = R_{00} - I - pC$,
- $r(0, 1, 1) = R_{01} - I - pC$,
- $r(1, 0, 1) = R_{10} - I - pC$,
- $r(1, 1, 1) = R_{11} - I - pC$.

2) $a = 2$:

- $r(0, 0, 2) = R_{00} - S - pM$,

⁴Las probabilidades 0 o 1 deben aparecer como tales.

$$\blacksquare r(0, 1, 2) = R_{01} - S - pM,$$

Las probabilidades de transición las presentamos en tres tablas, correspondientes a las decisiones $a = 0, 1, 2$. La tabla genérica está dada por: $(p((k, l)|(i, j), a)) =$

X		1	1	0	0	k
	Y	1	0	1	0	l
1	1	$\alpha_1^a \beta_1^a$	$\alpha_1^a(1 - \beta_1^a)$	$(1 - \alpha_1^a)\beta_1^a$	$(1 - \alpha_1^a)(1 - \beta_1^a)$	
1	0	$\alpha_1^a \beta_0^a$	$\alpha_1^a(1 - \beta_0^a)$	$(1 - \alpha_1^a)\beta_0^a$	$(1 - \alpha_1^a)(1 - \beta_0^a)$	
0	1	$\alpha_0^a \beta_1^a$	$\alpha_0^a(1 - \beta_1^a)$	$(1 - \alpha_0^a)\beta_1^a$	$(1 - \alpha_0^a)(1 - \beta_1^a)$	
0	0	$\alpha_0^a \beta_0^a$	$\alpha_0^a(1 - \beta_0^a)$	$(1 - \alpha_0^a)\beta_0^a$	$(1 - \alpha_0^a)(1 - \beta_0^a)$	
i	j					

aunque hay que tomar precauciones en el caso $a = 2$ tanto con el número de filas como con los valores porque está presente la posibilidad de que el soborno se descubra y no se consiga la acreditación.

Tomando en cuenta las restricciones descritas tenemos: cuando $a = 0$ (no hay inversión) no es posible aumentar la calidad, lo cual significa que $\alpha_0^0 = 0$. Por otra parte, si no invierte y tiene mala calidad, entonces no será acreditada, es decir $\beta_1^0 = \beta_0^0 = 0$. Llegamos entonces a

$$a = 0 \rightarrow$$

X		1	1	0	0	k
	Y	1	0	1	0	l
1	1	0	α_1^0	0	$1 - \alpha_1^0$	
1	0	0	α_1^0	0	$1 - \alpha_1^0$	
0	1	0	0	0	1	
0	0	0	0	0	1	
i	j					

Cuando $a = 1$ la calidad no puede disminuir, lo que equivale a $\alpha_1^1 = 1$. Por otra parte, se acredita con seguridad si la calidad es buena y está acreditada, es decir $\alpha_1^1 \beta_1^1 = 1$, de donde concluimos que $\beta_1^1 = 1$. Llegamos entonces a:

$$a = 1 \rightarrow$$

X		1	1	0	0	k
	Y	1	0	1	0	l
1	1	1	0	0	0	
1	0	β_0^1	$1 - \beta_0^1$	0	0	
0	1	α_0^1	0	$1 - \alpha_0^1$	0	
0	0	$\alpha_0^1 \beta_0^1$	$\alpha_0^1(1 - \beta_0^1)$	$(1 - \alpha_0^1)\beta_0^1$	$(1 - \alpha_0^1)(1 - \beta_0^1)$	
i	j					

Finalmente consideramos el caso $a = 2$. En primer término, las dos primeras filas de la matriz genérica no están porque no hay soborno si la calidad es 1. Cuando hay soborno, la calidad no puede subir, lo cual implica que $\alpha_0^2 = 0$. Por otra parte, para ser acreditado mediante soborno no puede haber inspección de Patrick Glass, lo cual ocurre con probabilidad $1 - p$ y es independiente del proceso (Z_n) . Resulta entonces

$$a = 2 \rightarrow$$

X		1	1	0	0	k
	Y	1	0	1	0	l
0	1	0	0	$\beta_1^2(1 - p)$	$1 - \beta_1^2(1 - p)$	
0	0	0	0	$\beta_0^2(1 - p)$	$1 - \beta_0^2(1 - p)$	
i	j					

La función objetivo a maximizar es la suma de las recompensas (con o sin tasa de descuento) hasta el horizonte N , es decir

$$\sum_{k=1}^{N-1} r(X_k, Y_k, a_k) + r_N(X_N, Y_N).$$

Notar que los a_k son decisiones aleatorias, dependientes del estado del sistema, con respecto a las cuales se optimiza.

b) Las ecuaciones de optimalidad de Bellman se escriben como

$$u_k(i, j) = \max_{a \in A_{ij}} \left\{ r(i, j, a) + \sum_{(i', j') \in E} u_{k+1}(i', j') p((i, j), (i', j'), a) \right\}, \quad 1 \leq k \leq N-1,$$

con $u_N(i, j) = F_{ij}$, $i, j \in E = \{0, 1\}$. Estas ecuaciones las especializamos según los 4 elementos del espacio de estados E . Comenzamos con $(i, j) = (0, 0)$ y leemos los valores en la tabla de recompensas: $r(0, 0, 0) = R_{00} - pC$, $r(0, 0, 1) = R_{00} - I - pC$, $r(0, 0, 2) = R_{00} - S - pM$. Por otra parte, de las probabilidades de transición obtenemos (solo listamos las positivas):

$$\begin{aligned} p((0, 0), (0, 0), 0) &= 1, \\ p((0, 0), (0, 0), 1) &= (1 - \alpha_0^1)(1 - \beta_0^1), p((0, 0), (0, 1), 1) = (1 - \alpha_0^1)\beta_0^1, \\ p((0, 0), (1, 0), 1) &= \alpha_0^1(1 - \beta_0^1), p((0, 0), (1, 1), 1) = \alpha_0^1\beta_0^1, \\ p((0, 0), (0, 0), 2) &= 1 - \beta_0^2(1 - p), p((0, 0), (0, 1), 2) = \beta_0^2(1 - p). \end{aligned}$$

Tenemos

$$\begin{aligned} u_k(0, 0) &= R_{00} + \max \{ -pC + u_{k+1}(0, 0), -I - pC + u_{k+1}(0, 0)(1 - \alpha_0^1)(1 - \beta_0^1) + \dots, \\ &\quad -S - pM + u_{k+1}(0, 0)(1 - \beta_0^2(1 - p)) + \dots \}. \\ &\quad \vdots \\ u_k(1, 1) &= R_{11} - pC + \max \{ u_{k+1}(1, 0)\alpha_1^0 + u_{k+1}(0, 0)(1 - \alpha_1^0), -I + u_{k+1}(1, 1) \}. \end{aligned}$$

c) Ahora damos algunos valores a los parámetros para realizar un par de iteraciones. Comenzamos con las probabilidades

- $p = 3/4$
- $\alpha_1^0 = 1/4, \alpha_0^1 = 3/4$
- $\beta_0^1 = 3/4, \beta_0^2 = 3/5, \beta_1^2 = 4/5$.

Luego las recompensas:

- $R_{00} = 10, R_{01} = 20, R_{10} = 30, R_{11} = 50$,
- $F_{00} = 100, F_{01} = 200, F_{10} = 300, F_{11} = 500$,
- $I = 25, S = 2, C = 4, M = 20$.

Reescribimos las ecuaciones de Bellman como sigue:

$$\begin{aligned} u_k(0, 0) &= \max \{ 7 + u_{k+1}(0, 0), \\ &\quad -18 + u_{k+1}(0, 0)\frac{1}{16} + u_{k+1}(0, 1)\frac{3}{16} + u_{k+1}(1, 0)\frac{3}{16} + u_{k+1}(1, 1)\frac{9}{16}, \\ &\quad -7 + u_{k+1}(0, 0)\frac{17}{20} + u_{k+1}(0, 1)\frac{3}{20} \}. \end{aligned}$$

$$\begin{aligned} u_k(0, 1) &= \max \{ 17 + u_{k+1}(0, 0), \\ &\quad -8 + u_{k+1}(0, 0)0 + u_{k+1}(0, 1)\frac{1}{4} + u_{k+1}(1, 0)0 + u_{k+1}(1, 1)\frac{3}{4}, \\ &\quad 3 + u_{k+1}(0, 0)\frac{4}{5} + u_{k+1}(0, 1)\frac{1}{5} \}. \end{aligned}$$

$$\begin{aligned} u_k(1, 0) &= \max \{ 27 + u_{k+1}(0, 0)\frac{3}{4} + u_{k+1}(1, 0)\frac{1}{4}, \\ &\quad 2 + u_{k+1}(0, 0)0 + u_{k+1}(0, 1)0 + u_{k+1}(1, 0)\frac{1}{4} + u_{k+1}(1, 1)\frac{3}{4} \}. \end{aligned}$$

$$u_k(1,1) = \text{máx}\{47 + u_{k+1}(0,0)\frac{3}{4} + u_{k+1}(1,0)\frac{1}{4}, \\ 22 + u_{k+1}(1,1)\}.$$

Para el caso $N - 1$ tenemos

$$u_{N-1}(0,0) = \text{máx}\{7 + 100, \\ -18 + 100\frac{1}{16} + 200\frac{3}{16} + 300\frac{3}{16} + 500\frac{9}{16}, \\ -7 + 100\frac{17}{20} + 200\frac{3}{20}\} \\ = \text{máx}\{107, 1453/4, 108\} = 1453/4 \Rightarrow a = 1.$$

$$u_{N-1}(0,1) = \text{máx}\{17 + 100, \\ -8 + 100 \cdot 0 + 200\frac{1}{4} + 300 \cdot 0 + 500\frac{3}{4}, \\ 3 + 100\frac{4}{5} + 200\frac{1}{5}\} \\ = \text{máx}\{117, 417, 123\} = 417 \Rightarrow a = 1.$$

$$u_{N-1}(1,0) = \text{máx}\{27 + 100\frac{3}{4} + 300\frac{1}{4}, \\ 2 + 300\frac{1}{4} + 500\frac{3}{4}\} \\ = \text{máx}\{177, 452\} = 452 \Rightarrow a = 1.$$

$$u_{N-1}(1,1) = \text{máx}\{47 + 100\frac{3}{4} + 300\frac{1}{4}, \\ 22 + 500\} \\ = \text{máx}\{197, 522\} = 522 \Rightarrow a = 1.$$

Observamos que, con los valores escogidos, en todos los casos conviene invertir.