

SOLUCIÓN CONTROL #3

1. Se ha observado que, gracias a gestiones realizadas por Armijo, los estudiantes de un conocido liceo emblemático de Santiago han adoptado un esquema democrático-aleatorizado para los períodos de toma y de clases en su establecimiento. Un período de clase dura un tiempo exponencial de parámetro λ_c mientras que un período de toma dura un tiempo exponencial de parámetro λ_t . Al terminar un período de clase o de toma, los estudiantes votan y deciden democráticamente si el período siguiente será de clases o de toma. Se sabe que las votaciones son independientes y que en cualquiera de ellas, la probabilidad de que gane la opción de clases es p_c y que gane la opción de toma es p_t , con $p_t = 1 - p_c$. Suponga que el liceo inicia actividades con un período de clases. Plantee un modelo estocástico que le permita analizar los siguientes temas.

- a) Ignorando fines de semana, días de fiesta o vacaciones, calcule la probabilidad de que en algún tiempo lejano del futuro, el liceo se encuentre en clase.

Sol: El modelo que se adapta a la situación descrita es el de renovación alternante. Designamos por (Z_n) las duraciones de los períodos contiguos de clases y por (Y_n) las duraciones de las tomas. Finalmente, las va $X_n = Z_n + Y_n$ representan los largos de los ciclos. Para obtener las distribuciones de las va Z_n, Y_n notamos que el primer período de clase se extiende mientras las votaciones lo indiquen así. Por lo tanto Z_1 es la suma de va exponenciales iid con parámetro λ_c pero el número de términos de es una va geométrica (independiente de las exponenciales) con parámetro p_t . Es decir,

$$Z_1 = \sum_{i=1}^N U_i,$$

donde los U_i son iid exp λ_c y N es va con $P[N = n] = p_c^{n-1} p_t, n = 1, 2, \dots$. Por supuesto, la misma situación aplica a las restantes va Z_2, Z_3, \dots que son iid. Ahora calculamos la distribución de Z_1 y para ello condicionamos en N :

$$P[Z_1 > z] = \sum_{n=1}^{\infty} P \left[\sum_{i=1}^N U_i > z \mid N = n \right] P[N = n].$$

Ahora recordamos que la suma de n va iid exponenciales de parámetro λ es una va gamma de parámetros λ, n . También recordamos que tiene densidad $\lambda(\lambda t)^{n-1} e^{-\lambda t} / (n-1)!$. Por lo tanto

$$P[X_1 > z] = \sum_{n=1}^{\infty} \int_z^{\infty} \frac{\lambda_c (\lambda_c t)^{n-1} e^{-\lambda_c t}}{(n-1)!} dt p_c^{n-1} p_t = \lambda_c p_t \int_z^{\infty} e^{-\lambda_c t} \sum_{n=1}^{\infty} \frac{(\lambda_c p_c t)^{n-1}}{(n-1)!} dt = e^{-\lambda_c p_t z}.$$

Es decir, Z_1 tiene ley exponencial de parámetro $\lambda_c p_t$. Un desarrollo totalmente análogo permite concluir que las va Y_i tienen distribución exponencial de parámetro $\lambda_t p_c$. Además, las va Z_i e Y_i son independientes. Así se completa la descripción del proceso de renovación alternante.

Para la probabilidad en el largo plazo de estar en clase usamos el resultado sobre probabilidad de encontrar el sistema en ON-OFF. Tenemos entonces que la probabilidad asintótica de estar en clase es

$$p_{ac} = \frac{E[Z_1]}{E[Z_1] + E[Y_1]} = \frac{(\lambda_c p_t)^{-1}}{(\lambda_c p_t)^{-1} + (\lambda_t p_c)^{-1}} = \frac{\lambda_t p_c}{\lambda_c p_t + \lambda_t p_c}.$$

- b) Se sabe que las tomas tienen costos y potenciales beneficios de acuerdo con el siguiente esquema: el costo por unidad de tiempo que el liceo está en toma es c_t ; el costo fijo para levantar una toma (una toma se levanta cuando se decide volver a clase) es c_l ; el beneficio asociado a un período de toma es una va B (iid entre períodos e independiente de las duraciones de los períodos y de los resultados de las votaciones) que toma el valor b_t con probabilidad q y 0 con probabilidad $1 - q$; el beneficio de un período de clase por unidad de tiempo es b_c . Calcule el beneficio neto por unidad de tiempo del liceo en el largo plazo.

Sol: Aquí podemos pensar en una renovación con recompensa. Cada ciclo, cuyo largo se distribuye como $X_1 = Z_1 + Y_1$, tiene asociada cierta recompensa que calculamos como sigue. El ciclo i tiene largo X_i , con valor esperado $(\lambda_c p_t)^{-1} + (\lambda_t p_c)^{-1}$, y recompensa (costo)

$$R_i = c_t Y_i + c_l - B_i,$$

donde B_i es el beneficio de la toma en el ciclo i . De acuerdo con un resultado del curso, la recompensa por unidad de tiempo, en el largo plazo, se calcula como el cociente entre el valor de la recompensa en un ciclo dividido por el largo esperado del ciclo. Es decir,

$$c = \frac{E[R_1]}{E[X_1]} = \frac{c_t(\lambda_t p_c)^{-1} + c_l - q b_t}{(\lambda_c p_t)^{-1} + (\lambda_t p_c)^{-1}}$$

- c) Suponga que el liceo comienza el año con clases. Calcule el tiempo total esperado de clase desde el comienzo del año hasta que se levanta la primera toma posterior al tiempo t .

Sol: Consideramos el proceso de renovación $N(t)$ asociado a las X_i (ciclo completo de clase y toma). Entonces, el tiempo total de clase hasta que se levanta la última toma antes de t es $T_c(t) = \sum_{i=1}^{N(t)+1} Z_i$. Notar que no estamos considerando el tiempo de clase exactamente hasta t . Para calcular el valor esperado de $T_c(t)$ usamos Wald, porque $N(t) + 1$ es un tdp para la sucesión $(Z_n, Y_n)_{n \geq 1}$. Si designamos por $m(t)$ la función de renovación de N tenemos

$$E[T_c(t)] = E[N(t) + 1]E[Z_1] = \frac{m(t) + 1}{\lambda_c p_t}.$$

2. Armijo ha ofrecido también sus servicios de asesoría al alcalde de otra comuna que enfrenta tomas de sus liceos (emblemáticos y no tanto). Se ha observado, durante 2011, que los períodos de clase y de toma se alternan y tienen duraciones aleatorias independientes con distribuciones respectivas F, G , con densidades f, g y valores esperados finitos μ_F, μ_G . Armijo sugiere que si una toma ha durado demasiado tiempo, entonces el alcalde debe usar la fuerza pública para desalojar el liceo tomado. Se sabe que una toma tiene un costo c por unidad de tiempo y si termina con desalojo, el costo se incrementa puntualmente en δ (no es proporcional a la duración), debido al deterioro de imagen la pública y de la convivencia. La estrategia planteada por Armijo consiste en desalojar un liceo tomado al cabo de T unidades de tiempo. Suponga que el liceo inicia actividades con un período de clases.

- a) Calcule la probabilidad de que en el largo plazo el alcalde llegue al liceo y lo encuentre en clases pero que éstas no duren más de x unidades de tiempo hasta la próxima toma.

Sol: Para modelar este problema consideramos un proceso de renovación alternante. Si la política de desalojo no se implementa entonces las va correspondientes a clases y tomas tienen distribuciones respectivas F, G . Digamos que un ciclo tiene un tiempo X de clase mas un tiempo Y de toma. Vamos a usar un método de solución similar al cálculo de la distribución asintótica de la edad (Ross, p.116). Asociamos una recompensa R al ciclo dado por X, Y . Si $X \leq x$ recibimos una recompensa X (o de manera equivalente, el estado del sistema es ON durante todo el período de clases y OFF cuando no hay clase). Si $X > x$ entonces la recompensa es x

(de manera equivalente, el sistema está en ON durante las últimas x unidades de tiempo de clase y OFF el tiempo restante del ciclo). Así la recompensa límite por unidad de tiempo, que corresponde a la proporción de tiempo que estamos en ON (estar en clase y que se termine antes de x unidades de tiempo) se calcula como el cociente

$$\frac{E[\min\{X_1, x\}]}{E[X_1] + E[Y_1]} = \frac{1}{\mu_F + \mu_G} \int_0^x (1 - F(t))dt.$$

En caso que la política de desalojo se haya implementado el razonamiento sigue siendo válido pero en la expresión de más arriba se debe cambiar μ_G por $\int_0^T (1 - G(t))dt$.

- b) Sabiendo que el beneficio por unidad de tiempo en un período de clase es b_c , muestre que el costo neto $\alpha(T)$, por unidad de tiempo, en el largo plazo, en que incurre el alcalde, como función de su variable de decisión T , se escribe como

$$\alpha(T) = \frac{c \int_0^T \bar{G}(x)dx + \delta \bar{G}(T) - b_c \mu_F}{\int_0^T \bar{G}(x)dx + \mu_F},$$

y muestre que α es creciente (no estricta) en valores T tales que $\alpha(T) \leq c - \delta h(T)$, donde $h(T) = g(T)/(1 - G(T))$ es la función de fallas de F (hazard).

Sol: Aquí se requiere escribir la recompensa asociada a un ciclo de clase-toma, teniendo en cuenta la política de desalojo. Las clases duran un tiempo aleatorio X con distribución F y la toma dura $\min\{Y, T\}$ (Y es la duración que tendría la toma sin desalojo). La recompensa (costo) será entonces

$$R = c \min\{Y, T\} + \delta \mathbf{1}_{\{Y > T\}} - b_c X.$$

Finalmente, la recompensa límite por unidad de tiempo se obtiene dividiendo la recompensa esperada en un ciclo por el largo esperado en un ciclo. Antes notamos que

$$E[\min\{Y, T\}] = \int_0^\infty P[\min\{Y, T\} > t]dt = \int_0^\infty P[Y > t]P[T > t]dt = \int_0^T (1 - G(t))dt.$$

Adoptando la notación $\bar{G} = 1 - G$ tenemos que $E[R] = c \int_0^T \bar{G}(t)dt + \delta \bar{G}(T) - b_c \mu_F$ y comprobamos que se obtiene la fórmula para $\alpha(T)$.

Para ver donde α es creciente calculamos la derivada α' y caracterizamos los valores de T para los cuales $\alpha'(T) \geq 0$. De la fórmula de α se obtiene:

$$\alpha'(T) \left(\int_0^T \bar{G}(x)dx + \mu_F \right) + \alpha(T) \bar{G}(T) = c \bar{G}(T) - \delta g(T).$$

Por lo tanto, $\alpha'(T) \geq 0$ si y solo si

$$(c - \alpha(T)) \bar{G}(T) - \delta g(T) \geq 0,$$

lo cual es equivalente a (dividiendo por $\bar{G}(T)$ y despejando α)

$$\alpha(T) \leq c - \delta h(T).$$

- c) Suponga que G es la distribución exponencial de parámetro 1. Escriba explícitamente $\alpha(T)$ en este caso y muestre que al alcalde le conviene desalojar inmediatamente una toma si $\delta < \delta^* := \mu_F(b_c + c)/(\mu_F + 1)$. Qué debería hacer el alcalde cuando $\delta > \delta^*$ y $\delta = \delta^*$?

Sol: Si G es la exponencial de parámetro 1, tenemos $\bar{G}(t) = e^{-t}$ y $h(t) = 1$. También, $\int_0^T \bar{G}(t)dt = 1 - e^{-T}$ y

$$\alpha(T) = \frac{c(1 - e^{-T}) + \delta e^{-T} - b_c \mu_F}{1 - e^{-T} + \mu_F}.$$

Al alcalde le conviene desalojar de inmediato una toma si el costo es creciente con T . Para que el costo sea creciente (en un valor de T dado) debe cumplirse $\alpha(T) < c - \delta$, lo cual equivale a

$$c(1 - e^{-T}) + \delta e^{-T} - b_c \mu_F < (c - \delta)(1 - e^{-T} + \mu_F).$$

Pero la desigualdad anterior se cumple si y solo si

$$c - b_c \mu_F < (c - \delta)(1 + \mu_F) \Leftrightarrow \delta < \frac{\mu_F(b_c + c)}{\mu_F + 1},$$

cualquiera sea $T \geq 0$. Es decir, cuando $\delta < \delta^*$ hay que desalojar en $T = 0$ (de inmediato). Si $\delta > \delta^*$ los costos son decrecientes y no conviene desalojar. Si $\delta = \delta^*$ los costos (por unidad de tiempo) son independientes de T y daría igual en qué momento desalojar.

- Para responder a las inquietudes de padres o apoderados de los estudiantes de liceos, el ministerio de educación ha instalado un call center a cargo de dos expertos en educación (Pedro y Juan) que atienden las llamadas a lo largo del día. Al comienzo de cada día el ministro inspecciona personalmente el call center, para saber si los expertos están operativos. El call center es redundante en el sentido que basta que uno de los expertos esté operativo para que el sistema funcione (aunque las llamadas pasarán mas tiempo en espera). El call center colapsa cuando ninguno de los dos expertos está en condiciones de trabajar. La tasa de falla de un experto depende de cuántos días lleva trabajando sin descanso y del estado de su compañero de trabajo (operativo o no). Un experto que lleva m días trabajando sin descanso, colapsará al día siguiente con probabilidad $r_1(m)$ si su colega está operativo y con probabilidad $r_2(m)$ si no lo está. Si al momento de la inspección un experto está colapsado, se le envía a descanso por T_0 días y vuelve como nuevo. También es posible enviar preventivamente a descansar a un experto que esté operativo, período que se extiende por T_1 días, con $1 \leq T_1 < T_0$. Cada día, al momento de la inspección, el ministro debe decidir si envía a descansar a uno o a ambos expertos o bien, los deja que sigan trabajando. El objetivo es minimizar el tiempo que el call center está caído, en un horizonte de N días.

Se pide formalizar la situación descrita más arriba como un Problema de Decisión Markoviano, identificando claramente sus elementos: espacio de estados, acciones, recompensas, probabilidades de transición y función objetivo a optimizar.

Indicación: Puede considerar como estado del call center al par de enteros (x_P, x_J) , con la siguiente interpretación: la primera coordenada se refiere a Pedro y la segunda a Juan. Si $x_P = i \geq 0$ entonces Pedro lleva i días trabajando sin descanso; si $x_P = -k < 0$ entonces a Pedro le quedan k días de descanso (y no está operativo). La misma interpretación vale para la coordenada de Juan x_J .

Sol: El estado del sistema lo representamos, tal como se sugiere en la indicación, por los pares del tipo (i, j) , donde i, j son enteros, posiblemente negativos. En principio los valores positivos no están acotados aunque es evidente que no pueden trabajar más allá de lo que duren sus propias existencias. Además, los valores negativos están acotados por $-T_0$ porque T_0 es el número máximo de días de descanso que se pueden tomar. Definimos

$$E = \{(i, j) \mid i, j \in \mathbf{Z}, -T_0 \leq i, j \leq M\},$$

donde M es el máximo de días que se puede trabajar sin descansar¹. El proceso estocástico asociado a estos estados lo designamos por $X_n = (U_n, V_n), n = 1, 2, \dots$

¹Si se omite M no importa

Las acciones que se pueden tomar en un día cualquiera son: enviar a descansar a Pedro y Juan si ambos están trabajando; enviar a descansar a Pedro si está trabajando; enviar a descansar a Juan si está trabajando; no enviar a ninguno. Más formalmente, debemos establecer conjuntos de acciones posibles $A_{i,j}$ que dependen del estado del sistema (i, j) . Consideremos las acciones P = Pedro se envía a descansar, J = Juan se envía a descansar y PJ = Pedro y Juan se envían a descansar, \emptyset = Nadie va a descansar. Entonces

$$A_{i,j} = \{P, J, PJ, \emptyset\}, \text{ si } i, j \geq 0;$$

$$A_{i,j} = \{P, \emptyset\}, \text{ si } i \geq 0, j < 0;$$

$$A_{i,j} = \{J, \emptyset\}, \text{ si } i < 0, j \geq 0;$$

$$A_{i,j} = \{\emptyset\}, \text{ si } i, j < 0.$$

Consideremos ahora las probabilidades de transición, que dependen del estado presente y de la acción que se tome. Sea $p((j, l)|(i, j), a), a \in A_{i,j}$ la probabilidad de que el sistema pase de (i, j) a (k, l) cuando la acción es a . Para formalizar las transiciones diremos que un trabajador que está colapsado al momento de la inspección inicia de inmediato su período de descanso. Veamos las posibles transiciones (con probabilidad positiva), comenzado con $i, j \geq 0$:

$$p((i+1, j+1)|(i, j), \emptyset) = (1 - r_1(i))(1 - r_1(j)),$$

$$p((-T_0, j+1)|(i, j), \emptyset) = r_1(i)(1 - r_1(j)),$$

$$p((i+1, -T_0)|(i, j), \emptyset) = (1 - r_1(i))r_1(j),$$

$$p((-T_0, -T_0)|(i, j), \emptyset) = r_1(i)r_1(j),$$

$$p((-T_1, j+1)|(i, j), P) = 1 - r_1(j), \text{ (ver nota}^2\text{)},$$

$$p((-T_1, -T_0)|(i, j), P) = r_1(j),$$

$$p((i+1, -T_1)|(i, j), J) = 1 - r_1(i),$$

$$p((-T_0, -T_1)|(i, j), J) = r_1(i),$$

$$p((-T_1, -T_1)|(i, j), PJ) = 1.$$

Supongamos ahora que $i < 0, j \geq 0$:

$$p((i+1, j+1)|(i, j), \emptyset) = 1 - r_2(j),$$

$$p((i+1, -T_0)|(i, j), \emptyset) = r_2(j),$$

$$p((i+1, -T_1)|(i, j), J) = 1.$$

Análogamente, si $i \geq 0, j < 0$:

$$p((i+1, j+1)|(i, j), \emptyset) = 1 - r_2(i),$$

$$p((-T_0, j+1)|(i, j), \emptyset) = r_2(i),$$

$$p((-T_1, j+1)|(i, j), P) = 1.$$

Finalmente si $i, j < 0$:

$$p((i+1, j+1)|(i, j), \emptyset) = 1.$$

²Cuando un experto se envía a descansar, estando operativo, suponemos que trabaja el día en que se tomó la decisión y el descanso comienza el día siguiente.

Con respecto a las recompensas supondremos que se recibe 1 unidad cada día que el call center está operativo. La recompensa de un día depende del estado y no de la decisión que tomemos ese día porque ésta se manifiesta al día siguiente. Tenemos entonces

$$r(i, j) = \mathbf{1}_{\{\max\{i, j\} \geq 0\}}.$$

La función objetivo a maximizar en un horizonte de N días es

$$E \left[\sum_{n=1}^N r(U_n, V_n) \right].$$