El valor optimo del indice de comportamiento

Irasema Pedroza Meza

2024-12-08

Tabla de contenidos

Prefacio

En este libro se desarrollara el proyecto de la clase de Aprendizaje Reforzado.

1 Introdución

En la literatura existen una variedad de modelos en los cuales consideran la vacunación como medida de prevención para el control de enfermedades [Jorge, y mas modelos]. Sin embargo, este medio de control llega a presentar algunas fallas como son: la falla de grado, falla en la toma y falla en la duración [Maclean]. En la modelación de enfermedades respiratorias con ecuaciones diferenciales, se suele considerar que las vacunas tienen una falla tipo de grado [Jorge y mas modelos]. Por otro lado, Pedroza, et. el, proponen un modelo de ecuaciones diferenciales donde consideran que la vacuna tiene dos tipos de fallas: la falla de grado y la falla en la toma.

En este ultimo trabajo, proponen un índice de comportamiento (ψ) el cual permite medir que tan riguroso pueden seguir las medidas de prevención una vez que son vacunados. En el modelo que proponen este índice solo afecta a los vacunados no inmunes

2 Formulación del Proceso de Decisión de Markov

Los estados de nuestro proceso de decisión de Markov representarán la proporción de la población en cada categoría del modelo propuesto por XXX:

- S_t: Fracción de susceptibles en el tiempo t.
- V_{+t}: Fracción de vacunados inmunes en el tiempo t.
- V_{-t}: Fracción de vacunados no inmunes en el tiempo t.
- I_t: Fracción de infectados en el tiempo t.
- R_t: Fracción de recuperados en el tiempo t.

El estado global del sistema en el tiempo t se presenta como

$$x_t = (S_t, V_{+t}, V_{-t}, I_t, R_t)$$

El escenario que consideraremos para cada $t \in 0, 1, \ldots, N$para el proceso de Markov:

- x_t: representa la dinámica de la enfermedad en el tiempo t.
- a_t: representa en qué escenario del índice de comportamiento se encuentra la población en el tiempo$t$.

Algunos supuestos que estaremos considerando para nuestro proceso son:

- Las personas cambian su comportamiento en el tiempo t de forma instantánea.
- Las únicas personas que pueden cambiar su comportamiento son los vacunados no inmunes.
- Supondremos que las personas cambian su comportamiento bajo una distribución uniforme $[0.5, 2]$.

Bajo los supuestos anteriormente mencionados, consideramos el siguiente Modelo de Control de Markov.

$$(\mathbf{X}, \{A(x) : x \in X\}, \mathbf{P}, \mathbf{C})$$

donde **X** es el espacio de los estados, $\{A(x) : x \in X\}$ es el espacio de las acciones admisibles, **P** es la ley de transicion de modelo y **C** es la funcion de costo.

En la siguiente figura se ilustra el cambio de un estado a otro por las acciones admisibles. Desarrollaremos la dinámica de la cadena de Markov en las siguientes secciones.

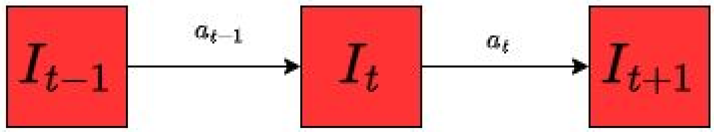

Figura 2.1: Diagrama del cambio de un estado a otro.

3 Dinámica del Modelo

En esta sección veremos cómo evoluciona el sistema atreves del tiempo t. La dinámica de la enfermedad esta representando por la siguiente cadena de Markov

- $S_{t+1} = \mu - \beta S_t I_t - (\mu + \phi)S_t + \omega V_{+t} + \xi R_t$
- $V_{+t+1} = \phi\_+(\sigma)S_t - (\mu + \omega)V_{+t}$
- $V_{-t+1} = \phi_-(\sigma)S_t - \psi(1-\sigma)\beta V_{-t} I_t - \mu V_{-t}$
- $I_{t+1} = \beta S_t I_t + \psi(1-\sigma)\beta V_{-t} I_t - (\mu + \gamma)I_t$
- $R_{t+1} = \gamma I_t - (\mu + \xi)R_t$

En la siguiente figura es una representacion de la cadena de Markov descrita anteriormente.

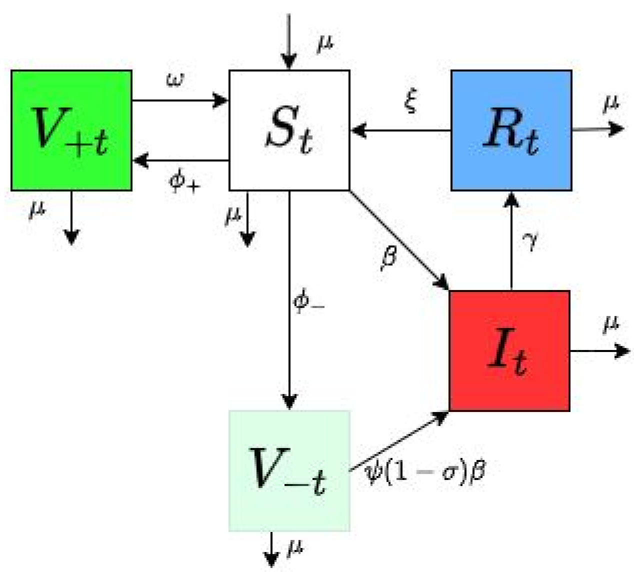

Figura 3.1: Diagrama de la cadena de Markov en el tiempo t

Las acciones representan los valores del índice de comportamiento (ψ). Supondremos que los valores del índice de comportamiento están restringidos al intervalo $[0.5, 2]$. Las acciones consideradas en nuestro modelo están dividas en tres principales acciones:

- Si $\psi \in [0.5, 0.9)$ entonces diremos que las personas se portan bien, es decir, que las personas siguen las medidas de prevención de forma estricta.
- Si $\psi \in [0.9, 1.2]$ entonces diremos que las personas se portan normal, es decir, que las personas siguen las medidas de prevención.
- Si $\psi \in (1.2, 2]$ entonces diremos que las personas se portan mal, es decir, que las personas no siguen las medidas de prevención.

Supondremos que las personas toman decisiones diarias mediante unas distribución $x_t \backsim Uni[0.5, 2]$ el periodo de observación será de un año. Como se ha mencionado con anterioridad el objetivo de este trabajo es encontrar el valor de ψ que permita tener la incidencia acumulada más pequeña al final del año. Entonces la probabilidad de tener i de incidencia acumulada y cambiar el comportamiento α en el tiempo t y pasar al estado j está dada por:

$$P_{ij}(\alpha) = P[x_{t+1} = j | x_t = i, a_t = \alpha]$$

4 Descripción y justificación del costo

En las secciones anteriores hemos mencionado que el objetivo es disminuir la incidencia acumulada en el tiempo $t = 365$. Por lo tanto, proponemos que la función de costo este dado por el costo de la enfermedad (C_I) y costo de vacunas con falla en la toma (C_V) sea

$$C_I(t) = c_i I_t$$

donde c_i es el costo unitario por persona infectada y I_t

$$C_{V-}(t) = c_v \phi\_V_-$$

c_v es el costo de generar vacunas con falla en la toma y ϕ_S es el número de susceptibles vacunados que no generan inmunidad. Ahora podemos calcular el impacto global que es:

$$C(t) = C_I(t) + C_{V-}(t)$$

Entonces el costo total esta dado por $C_{total} = \sum_{t=0}^{365} (0.5)^2 C(t)$

Las recompensas estarán determinadas mediante las acciones tomadas en el estado anterior. En la siguiente sección profundizaremos sobre las acciones y las recompensas que proponemos que se consideren esta clase de modelos.

5 Justificación de las acciones

En la sección **3** se observo como el conjunto principal de acciones que se consideran esta dada por intervalos. Consideramos que los valores de ψ va aumentando o disminuyendo en 0.1. Lo que nos lleva

- **Portarse bien:** serán los valores cuando $\psi = 0.5, 0.6, 0.7$ o $\psi = 0.8$
- **Portarse normal:** serán los valores cuando $\psi = 0.9, 1, 1.2$
- **Portarse mal:** serán los valores cuando $\psi = 1.3, 1.4, ...$ o $\psi = 2$

Cada una de estas acciones deben tener una recompensa y estara dada por la siguiente tabla

X_t	X_{t+1}	Recompesa
Portarse bien	Portarse bien	-0.5
Portarse bien	Portarse normal	1
Portarse bien	Portarse mal	2
Portarse normal	Portarse bien	0.5
Portarse normal	Portarse normal	1
Portarse normal	Portarse mal	-0.5
Portarse mal	Portarse bien	-1
Portarse mal	Portarse normal	-0.5
Portarse mal	Portarse mal	0.5