Funcionamiento de las Cadenas de Markov

Alejandra Velasco Zárate A01635453
José Antonio Juárez Pacheco A00572186
Jose Carlos Yamuni Contreras A01740285
Juan Manuel Hernández Solano A00572208
Mayra Sarahí De Luna Castillo A0163577

Octubre 2023

1 Cadena de Markov

Para analizar el comportamiento del mercado farmacéutico y conocer las probabilidades de que un producto sea desactivado o activado en una de las unidades de negocios de distribución masiva del grupo PiSA, Medicom, se hizo uso de las cadenas de Markov. Estas cadenas son modelos matemáticos que representan sistemas que cambian de estado de manera probabilística en intervalos discretos de tiempo, donde el estado futuro depende únicamente del estado presente, es decir, son una serie de eventos, en la cual la probabilidad de que ocurra un evento depende únicamente del evento inmediato anterior. A esto se le conoce como *pérdida de memoria*, donde solo recuerdan el último evento y esto condiciona las posibilidades de los eventos futuros. Esto será de mucha ayuda para predecir la compra o no compra de cada producto en cada perfil de venta (Hospitales, Farmacias y Distribuidores). Lo anterior se logró agrupando y filtrando la base de datos por *cliente_id* y *material_id* para realizar la matriz de transición, con el objetivo de obtener las probabilidades para n pasos, donde n son los meses. Los puntos a considerar en estas cadenas de Markov, es que se tienen dos estados planteados: $\mathscr{E} = \{0, 1\}$, donde 0 es que el producto fue comprado y 1 que el producto no fue comprado y los intervalos de tiempo se manejan en meses.
El diagrama de estado es el siguiente:

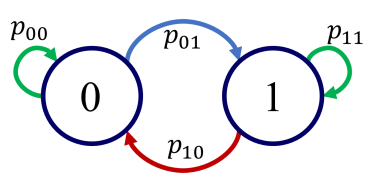

Figure 1: Diagrama de estados

Y la matriz de transición queda:

$$\mathcal{P} = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix}$$

donde p_{00} es la probabilidad de pasar al estado 0 dado que un paso anterior estaba en el estado 0, p_{01} es la probabilidad de pasar al estado 1 dado que un paso anterior estaba en el estado 0, p_{10} es la probabilidad de pasar al estado 0 dado que un paso anterior estaba en el estado 1 y p_{11} es la probabilidad de pasar al estado 1 dado que un paso anterior estaba en el estado 1. La suma de cada fila de la matriz de transición son siempre 1 (por ejemplo $p_{00} + p_{01} = 1$ y $p_{10} + p_{11} = 1$)

Una de las propiedades de las cadenas de Markov es que siempre convergen a probabilidades estables, esto permite encontrar el promedio a la larga o las probabilidades de estado estable para cada estado. Para lograr la convergencia de una probabilidad estable las cadenas de Markov tienen que cumplir con las siguientes condiciones:

1. **Irreducibilidad:** Desde cualquier estado es posible llegar a cualquier otro estado en un número finito de pasos, es decir, que todos los estados sean accesibles entre sí.
2. **Aperiodicidad:** La cadena debe ser aperiódica, lo que asegura que no esté atrapada en ciclos repetitivos y que con el tiempo cualquier estado pueda ser alcanzado.
3. **Recurrente:** Un estado recurrente es aquel desde el cual, una vez alcanzado, es seguro que volverá a visitarse en algún momento futuro. Para asegurar la convergencia, es importante que la cadena tenga al menos un estado recurrente positivo.

El cumplimiento de las condiciones anteriores da lugar a una cadena de Markov ergódica, esto significa que a medida que pasa el tiempo, la cadena convergerá hacia una distribución de probabilidad constante, independiente de las condiciones iniciales. Una vez sabiendo que se tiene una cadena érgodica, el siguiente paso es usar la descomposición espectral (la factorización de una matrizen términos de sus eigenvalores y eigenvectores) que se denota por la siguiente ecuación:

$$A = Q\Lambda Q^{-1}$$

donde Λ es una matriz diagonal de $(p \times p)$ cuyos elementos en la diagonal son los eigenvalores y Q es una matriz $p \times p$ cuya i-ésima columna es el eigenvector $\vec{v_i}$.

Para conseguir la descomposición, primero se calculan los eigenvalores y eigenvectores de la matriz haciendo uso de las siguientes fórmulas:

$$Eigenvalores : p(\lambda) = \det(P - \lambda I) = 0$$

$$Eigenvectores : (P - \lambda I)v_i = 0$$

Los eigenvalores se guardan en la variable Λ, y los eigenvectores se almacenan en la variable Q. Se calcula la matriz inversa Q^{-1}, y se realiza la matriz PP multiplicando las matrices Q, Λ y Q^{-1}, teniendo la matriz PP, la matriz diagonal Λ se eleva a la potencia n pasos. Y finalmente se calcula la matriz P_n con las probabilidades estables a n pasos, haciendo la multiplicación de Q, Λ^n y Q^{-1}.

Con las probabilidades constantes que arroja la matriz P_n se predicen y se realizan las recomendaciones de desactivar o activar los productos dependiendo de los porcentajes que se arroja en la matriz de 2x2. Si la probabilidad de compra es mayor al 70% se recomendará que se siga manteniendo la producción del producto.

Además de tener conocimiento si el producto se debe continuar produciendo o no, se añadió la recurrencia media, que se refiere al tiempo esperado o promedio que toma para que un proceso de Markov regrese a un estado particular después de haber salido de ese estado. Es decir, el tiempo que tardará un cliente en comprar (o no) un producto dado que ya lo había comprado (o no).

Lo anterior se calcula con los valores estables de las probabilidades de la cadena $\boldsymbol{\pi} = (\pi_0, \pi_1, \ldots)$. La probabilidad de que de la cadena llegue por primera vez al estado i partiendo del estado i, por ejemplo, que vuelva a comprar un producto dado que ya lo compro, esta dada por la fórmula

$$\mu_{ii} = \frac{1}{\pi_i}$$

Mientras que el tiempo medio de recurrencia de un estado recurrente j partiendo de un estado i, por ejemplo, que compre un producto dado que no lo había comprado está dado la siguiente fórmula

$$\mu_{ij} = 1 + \sum_{k \neq j} \pi_k \mu_{kj}$$

Realizando lo anterior, sabemos cuantos meses tiene que pasar para que se den los siguientes casos:

1. Que un cliente compre un producto dado que ya lo había comprado
2. Que un cliente no compre un producto dado que no lo había comprado
3. Que un cliente compre un producto dado que no lo había comprado
4. Que un cliente no compre un producto dado que lo había comprado

Esto puede ser de utilidad para saber y actuar sobre la capacidad de almacenamiento o de producción de los productos dependiendo de las acciones pasadas de los clientes