**Grupo 23**

En primer lugar, para la construcción del modelo, decidimos realizar una cadena de Markov de tiempo continuo para conocer el comportamiento de la cantidad de personas que se encuentran en cada una de las estaciones en todo momento.

**Descripción del modelo**

Se realizó un modelo para cada estación.

Variables:

* X(t): Cantidad de personas en la estación Primera de Mayo en el tiempo t
* Y(t): Cantidad de personas en la estación NQS en el tiempo t
* Z(t): Cantidad de personas en la estación Caracas en el tiempo t

Temporalidad: Tiempo continuo.

Espacio de estados: Sx = {0, 1, … 57}; Sy = {0, 1, … 256}; Sz = {0, 1, … 213} ; Sw = {Sx ,Sy , Sz}

Factores de aleatoriedad:

* Demora tiempo llegada de trenes
* Tiempo desplazamiento de los pasajeros para subirse al tren
* Interacción estaciones

Matriz:

**Metodología:**

**1.Transformación de datos:**

Primero calculamos los tiempos entre llegadas entre usuarios a la estación. Para ello restamos los tiempos de llegadas entre 2 usuarios consecutivos y pasamos esta diferencia a minutos multiplicando por 24\*60.

**2. Prueba de Bondad:**

Se realizó una prueba de bondad para una distribución exponencial ya que se va a analizar el comportamiento de los tiempos entre llegadas. Para ello se utilizó el software estadístico R con las funciones “fitdist”, “gofstat” y los valores encontrados en minutos mencionados anteriormente. Como resultado, se obtuvo que los datos sí siguen un comportamiento exponencial, con un alfa del 1%.

A continuación se calcularon las tasas de llegada para cada estación. Esto se hizo promediando los datos de los tiempos entre llegada para cada estación y luego poniendo cada tiempo sobre 1. Para hallar las tasas de llegada de los trenes, se hizo la inversa de los tiempos de llegada dados en el enunciado. Estos valores se muestran a continuación.

tasa con la que llega una persona a una estación:

tasa con la que llega un tren a una estación:

**3. Análisis Transitorio; Valor Esperado y Varianza:**

Todo el procedimiento que se describe a continuación se hizo en R.

Con el objetivo de estimar y visualizar gráficamente el valor esperado y la varianza, se realizó un análisis transitorio calculando las probabilidades para los diferentes estados. Para encontrar el valor esperado “e” se elevó a la Q\*t por medio de la función expm, se multiplicó esto por alfa y por el estado correspondiente. Para definir este vector de estado inicial se asumió que todas las estaciones empiezan vacías. Esto se realizó para cada minuto de la primera hora de funcionamiento y se guardó en un vector. A su vez, para obtener se hizo el mismo procedimiento que para hallar , con la diferencia de que se multiplicó por cada estado al cuadrado. Finalmente se restaron y , para hallar la varianza, la cuál se guardó en un vector. Todos los valores esperados se graficaron en una misma gráfica, así como las varianzas con las funciones plot y lines.

*de tamaño*

**4. Análisis de Costo y Estado Estable:**

Posteriormente, para calcular los costos que se darán a largo plazo fue necesario un análisis de estado estable para cada una de las estaciones. Por lo tanto, para cada estación creamos una CMTC en R usando la función new, con Class igual a ctmc y la matriz Q. Luego calculamos el vector de probabilidades de estado estable utilizando las funciones steadyStates y Re (esta última para eliminar números imaginarios). Para sacar el costo fijo por hora, se transformó el costo fijo semanal teniendo en cuenta que se trabajan 7 días a la semana 7 horas cada día. Para sacar el valor esperado en estado estable hicimos una suma producto de la matriz de probabilidades en estado estable con los estados. Finalmente para sacar el costo total se multiplicó la suma producto anterior por el costo por hora del WIFI y se le sumó el costo fijo.

A partir del resultado obtenido, se probó su significancia mediante la construcción de un intervalo de confianza con un nivel de significancia del 5%.

; ; ;

|  |  |
| --- | --- |
| Límite inferior | $ 3.617.179 |
| Límite superior | $ 3.695.417 |

Por lo tanto, se concluye que, al caer el costo promedio en el intervalo calculado, el modelo planteado sí representa el sistema al predecir el comportamiento del costo de las tres estaciones en el largo plazo con una certeza del 95%.