Frecuencia óptima de trenes en la Línea A - Metro de Medellín

*Nota: Proyecto del curso modelación y simulación 4 - Profesor Henry Laniado Rodas

Julieth Stefanny Escobar Ramírez
Ingeniería Matemática
Universidad EAFIT
Medellín, Colombia
jsescobarr@eafit.edu.co

Sara Gallego Villada
Ingeniería Matemática
Universidad EAFIT
Medellín, Colombia
sgallegov@eafit.edu.co

Hernán Moreno Mora
Ingeniería Matemática
Universidad EAFIT
Medellín, Colombia
homorenom@eafit.edu.co

***Resumen*—En este proyecto se implementa el modelo del vendedor de periódicos al Metro de Medellín. Se busca conocer el número óptimo de trenes por día de la semana y hora de operación, para cumplir la demanda y reducir costos. Los datos usados en el proyecto fueron tomados de Datos abiertos - Metro de Medellín, donde están disponibles los datos de tráfico por hora y día, del año 2019 en adelante. Para este trabajo, vamos a utilizar los correspondientes al año 2022, Línea A, del 1 de enero hasta el 30 de septiembre.**

***Index Terms*—Metro de Medellín, simulación, optimización, problema del vendedor de periódicos.**

I. Introducción

El Metro de Medellín es el sistema de transporte público masivo tipo metro que sirve a la ciudad de su nombre, y a los municipios de su área metropolitana: Envigado, Sabaneta, Itagüí, Bello y La Estrella. El Metro cuenta con dos líneas, la línea A(Niquía - La Estrella) y la línea B(San Antonio - San Javier), varias líneas de Metrocable, una de Tranvía, y líneas de buses integradas al sistema.

El Metro de Medellin debe satisfacer una demanda de pasajeros en cada una de sus lineas, y dicha demanda es variable tanto durante las horas del día como durante los días de la semana. Por lo tanto, la oferta de trenes debe ser acorde con la demanda, con el fin de no tener trenes viajando con pocos pasajeros y de no tener suficientes trenes para atender la demanda en horas pico.

En el proyecto aplicamos el modelo del vendedor de periódicos para obtener la frecuencia óptima de trenes cada hora, puesto que, si los trenes viajan vacíos, hay un costo de operación que no se recupera, y si no se logra atender la demanda hay otro costo por pérdida de confianza por parte de los usuarios. Por lo tanto, analizamos el tráfico en la línea A del metro de Medellín en busca de la función de utilidad que permita determinar las frecuencias óptimas de operación, durante las horas del día y los días de la semana.

II. Objetivos

II-A. *Objetivo general*

Aplicar el módelo del vendedor de periódicos para encontrar el número óptimo de trenes en el Metro de Medellín para reducir costos y satisfacer la demanda.

II-A1. *Objetivos específicos:*

- Analizar los datos de la línea A del Metro de Medellín y establecer su distribución.
- Definir las variables del modelo para plantear la función de utilidad.
- Plantear la función de utilidad.
- Análisis de sensibilidad

III. Justificación

La cantidad de pasajeros que hacen uso de la Linea A del Metro de Medellín es variable durante las 20 horas diarias de operación (4 AM a 12 PM) los 7 días de la semana. Si las frecuencias no satisfacen la demanda, podría presentarse congestión en ciertas horas del día, con pasajeros descontentos que podrían dejar de usar el Sistema. Por el contrario, si las frecuencias son muy altas, el tren viajaría con pocos pasajeros generando un costo innecesario.

En nuestro trabajo buscamos generar un modelo que maximice la utilidad del sistema (y que minimice el costo de operación) mediante la determinación de la frecuencia óptima.

IV. MARCO TEÓRICO

El modelo del vendedor de periódicos es un problema fundamental en inventarios de control estocásticos; su objetivo es maximizar ganancias o minimizar costos esperados, determinando un inventario óptimo.

Es una análogía con el vendedor de periodicos, que debe hacer un pedido diario suficiente para satisfacer la demanda, pero sin que le sobren periódicos no vendidos al final del día. Si no satisface la demanda, se daña su reputación y podría perder clientes, y si le sobran periódicos no vendidos, tendrá un costo adicional.

Existen diversos estudios y aplicaciones de este modelo; en la sección $VI - B$ se explica la función de utilidad aplicada en nuestro proyecto. Khouja[1], plantea el siguiente acercamiento matemático general:

x Demanda, variable aleatoria
$f(x)$ función de densidad de probabilidad
$F(x)$ función de distribución acumulada.
P valor de venta por unidad.
C costo por unidad
V valor de salvamento por unidad.
S costo de penalización por escasez por unidad
C_o $C - V$, costo del exceso de unidades.
C_u $P - C + S$, costo unitario de los que escasean.
Q cantidad del pedido, una variable de decisión.

Donde las ganacias por periodo estan dadas por:

$$\pi = \begin{cases} (P-C)Q - S(x-Q), & si\ x \geq Q \\ Px + V(Q-x) - CQ, & si\ x < Q \end{cases} \tag{1}$$

Al simplificar y tomar el valor esperado de π se obtiene la siguiente ganancia esperada

$$E(\pi) = (P + S - C) \int_Q^{\infty} Q\, f(x)\, dx$$

$$-S \int_Q^{\infty} x f(x)\, dx + (P - V) \int_0^Q x f(x)\, dx$$

$$-(C - V) \int_0^Q Q f(x)\, dx \tag{2}$$

Para obtener el resultado que maximiza la función $F(Q)$

$$F(Q) = \frac{P + S - C}{P + S - V} \tag{3}$$

Con lo anterior se esperaria que se maximise $E(\pi)$ sin embargo esto puede no verse reflejado en la realidad, pues es más coherente alcanzar un objetivo de ganancias con las acciones de muchos directivos que con esta maximización.

IV-A. *Conceptos*

Los trenes del Metro de Medellín viajan con 6 vagones de 22.86 metros de largo por 3.2 metros de ancho cada uno. Si suponemos un máximo de 6 personas por metro cuadrado (incluidas las 148 que van sentadas) podemos estimar en 400 personas el cupo máximo por vagón para un total de 2400 personas por tren.

En horas pico, los trenes viajan hoy con una separación de 3 minutos, esto es, 20 trenes por hora en cada sentido, lo cual nos daría una capacidad máxima de 48000 personas por hora en cada sentido. En las 20 horas de operación, serían 960.000 personas por día en cada sentido.

Los datos muestran una utilización real de 597.200 personas por día en promedio, con un máximo de 739.009 (el día 228 del año, agosto 16, martes luego de un lunes festivo) y un mínimo de 53.959 (el día 170 del año, junio 19, domingo), lo cual indica una demanda máxima inferior a la capacidad máxima calculada.Los domingos, el máximo se reduce a la mitad, esto es, cerca de 350.000 personas.

Los promedios de pasajeros por hora, para los primeros 223 días de 2022 se muestran en la siguiente gráfica de barras (Figura 1):

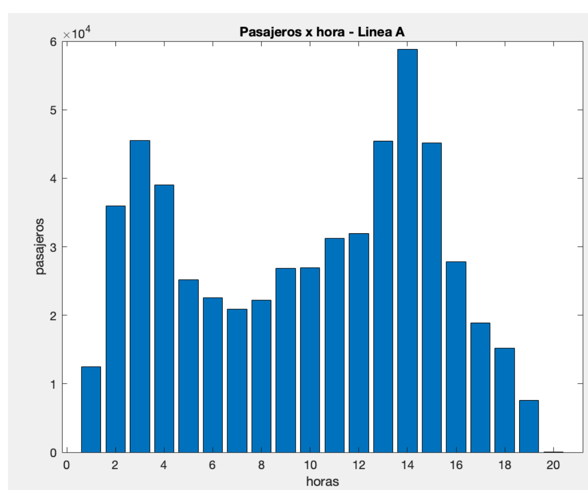

Figura 1. Los promedios de pasajeros por hora

Y estas son las medias por hora por día de la semana (el día 1 es sábado) (Figura 2)

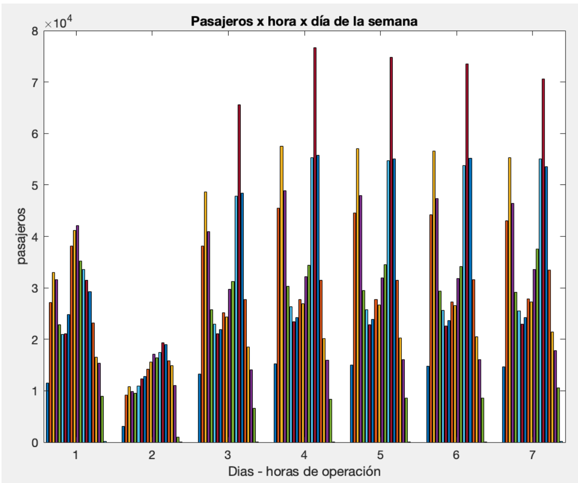

Figura 2. Medias por hora por día de la semana (el día 1 es sábado)

Con estos datos podemos calcular una matriz de medias (μ) y otra de desviación standard (σ). Cada una tendrá 7 filas (los 7 días de la semana) y 20 columnas (las 20 horas de operación por día), de tal forma que podremos obtener los valores correspondientes a cualquier día de la semana y cualquier hora de operación.

Como resultado, buscamos el que sea posible simular el tráfico por hora para una demanda esperada. Así, en caso de eventos especiales en la ciudad cuando se espera un incremento en el tráfico total, sería posible calcular anticipadamente la demanda por hora.

V. Metodología

Primero limpiamos el dataset obtenido de la página del metro Datos abiertos - Metro de Medellín para analizar los datos disponibles de la línea A.

Luego definimos las variables que contiene el módelo para el adecuado planteamiento de la función de utilidad. Posteriormente planteamos la función de utilidad. Simulamos el tráfico del metro en Matlab y hacemos la aplicación del módelo.

VI. Implementación

VI-A. *Datos disponibles*

En la página Datos abiertos - Metro de Medellín están disponibles los datos de tráfico por hora y día, del año 2019 en adelante. Para este trabajo, vamos a utilizar los correspondientes al año 2022 hasta el 30 de septiembre.

Los datos están disponibles en Excel, y mediante algunas transformaciones los convertimos en un archivo .txt que es aceptado por Matlab. El archivo es una matriz de 273 filas (días del año desde el 1 de enero 2022 hasta el 30 de septiembre 2022) y 20 columnas (horas de operación desde las 4 AM hasta las 12 PM). Un histograma de dicho archivo, donde se muestran los intervalos de pasajeros por hora en el eje x con un ancho de 3000, y el número de veces que dicho intervalo se presentó en los 273 días analizados en el eje y, es el siguiente:

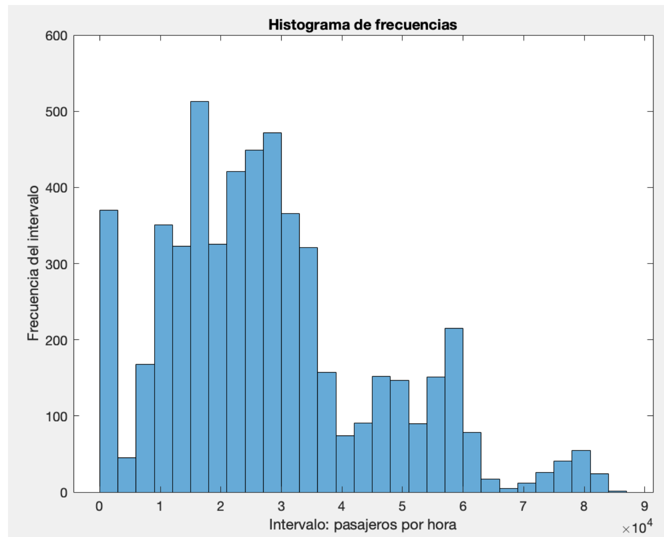

Figura 3. Intervalos de pasajeros por hora en 273 días

En un primer paso, generaramos datos simulados que se ajusten a la distribución de frecuencas de los datos crudos. Este modelo nos servirá en un futuro para simular una demanda específica en un día calquiera. Un ejemplo es el siguiente, con 100.000 usuarios (Figura 3):

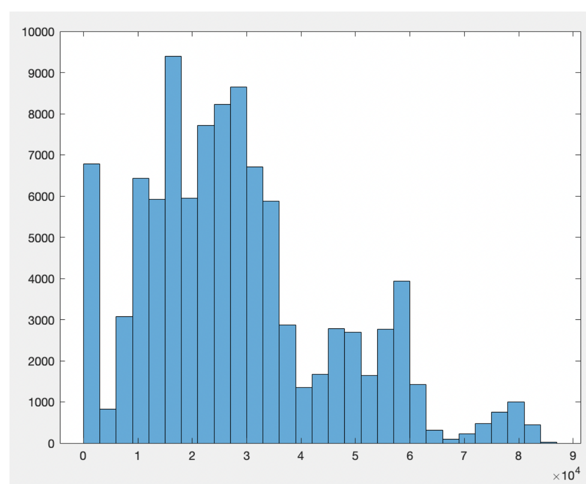

Figura 4. 100.000 usuarios

VI-B. Función de utilidad

Variables del modelo:

- p: Valor del tiquete: \$2500 en promedio
- n: capacidad de cada vagón (personas por vagón): 400 personas
- c: Costo de operación de un vagón durante una hora: estimado en \$1500, según los datos del presupuesto para el año 2022 [3]
- s: Valor de salvamento: mientras más peso en el tren, mayor desgaste de ruedas y rieles. Por lo tanto, si una persona no viaja, podemos suponer que contribuye a un menor desgaste. En la solución analítica, usaremos el 0.1 % del valor del tiquete (p), esto es, \$2.50
- k: Costo por demanda no satisfecha: trabajamos con un porcentaje del valor del tiquete. En la solución analítica, usaremos el 1 %del valor de p, esto es, \$25
- X: demanda en número de pasajeros por hora
- Q: Capacidad disponible por hora
- Función de utilidad:

$$U(Q) = pX - cQ + s(Q - X) \text{ si } x \leq Q$$

$$U(Q) = pQ - cQ - k(X - Q) \text{ si } x \geq Q \tag{4}$$

- Problema: maximizar la esperanza de la función de utilidad de acuerdo con la demanda esperada en cada hora de operación, y día de la semana.

Debemos encontrar la función de distribución de X ($X \sim F$) y la función de densidad $f(X)$. Para ello, debemos calcular:

$$E(U(Q)) = \int_{\infty}^{\infty} U(Q) f(x)dx \tag{5}$$

$= \int_0^Q (pX - cQ + (Q - X)s) f(x)dx + \int_Q^\infty (pQ - cQ + (X - Q)k) f(x)dx$

La primera integral es para $X \leq Q$ y la segunda para $X \geq Q$

Agrupando: $(p-s)\int_0^Q X f(X)dx + (s-c)Q \int_0^Q f(x)dx - k\int_Q^\infty X f(X)dx + (pQ - cQ + kQ)\int_Q^\infty f(x)dx$ (1)

Sabemos que la media

$\mu = \int_0^\infty X f(X)dx = \int_0^Q X f(X)dx + \int_Q^\infty X f(X)dx$

Por lo tanto: $\int_Q^\infty X f(X)dx = \mu - \int_0^Q X f(X)dx$

Pero: $\int_0^\infty f(X)dx = 1 = \int_0^Q f(X)dx + \int_Q \infty X f(X)dx$

Y por lo tanto: $\int_Q^\infty f(X)dx = 1 - \int_0^Q f(X)dx$

Reemplazando en (1): $E(U(Q)) = (p-s)\int_0^Q X f(X)dx + (s-c)Q\int_0^Q f(X)dx - k(\mu - \int_0^Q X f(x)dx) + (p - c + k)Q(1 - \int_0^Q f(X)dx$

Derivando con respecto a Q:
$\frac{\partial U}{\partial Q} = (p-s)Qf(Q) + (s-c)(Qf(Q) + \int_0^Q f(X)dx(1)) + kQf(Q) + (p-c+k) - (p-c+k)(Qf(Q) + \int_0^Q f(X)dx(1)$

$=(s-c)\int_0^Q f(X)dx - (p-c+k)\int_0^Q f(X)dx + (p-c+k) = 0$

Y por lo tanto:

$$\int \lbrack_0^Q f(X)dx(s-p-k) = c - p - k \tag{6}$$

De donde:

$$F(Q) = \frac{p-c+k}{p-s+k} \tag{7}$$

acotada entre 0 y 1.

Y por lo tanto podemos encontrar Q como

$$Q = F^{-1}(\frac{p-c+k}{p-s+k}) \tag{8}$$

VII. Simulación

No nos fue posible obtener los datos de costo de operación directamente del Metro, pero encontramos el presupuesto del año 2022 en internet Presupuesto Metro 2022, y con base en las cifras de costos de operación, estamos suponiendo un costo promedio de \$1500 por persona transportada.

Se simula el tráfico para una hora determinada, generando demanda aleatoria con distribución normal según los parámetros de μ y σ obtenidos del análisis de los datos históricos. Suponemos también una tarifa de \$2500 y un costo de operación por persona de \$1500. Usamos un valor de k (Costo por demanda insatisfecha) igual al 1 % de los ingresos dejados de percibir: $0{,}1(2500(X - Q))$.

Hacemos la simulación para valores de Q desde 1000 hasta 60000 con incrementos de 1000, y hacemos 500 simulaciones para cada valor de Q, y a cada una de ellas le calculamos la media. Tendremos al final 60 medias de los valores de la función de utilidad para una hora determinada, y graficamos dichas medias vs Q.

Al hacer la simulación para las 5 PM (hora 14 de operación) de un día sábado, obtenemos el siguiente gráfico de U vs Q (Figura 4).

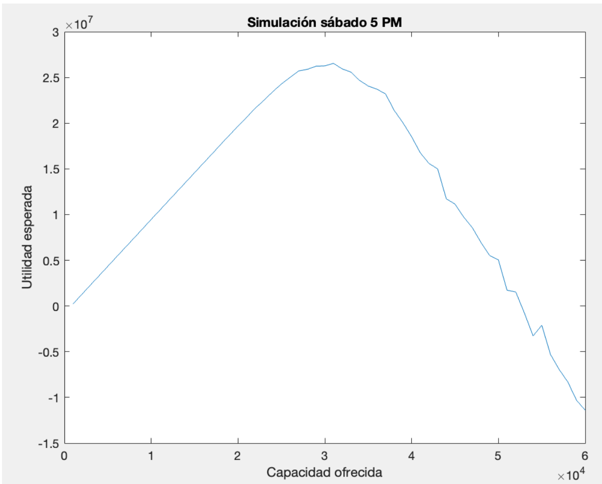

Figura 5. U vs Q

La interpretación es que la utilidad máxima se obtendría con una oferta entre 28000 y 32000 y que se perdería dinero con una oferta superior a 52000.
Analíticamente para este mismo caso, tenemos que $\mu = 31502$ y $\sigma = 5025$ (Obtenidas de las matrices de μ y σ calculadas anteriormente). Por lo tanto, usando la función inversa de Matlab (norminv), tenemos $norminv(0{,}4063, 31502, 5025) = 30311$, donde $\frac{2500-1000+25}{2500-2{,}50+25} = 0{,}4063$. Este número (30311) coincide con el máximo encontrado en la simulación.

La simulación para una hora de menor tráfico (11 AM), también del sábado, nos da lo siguiente (Figura 6):

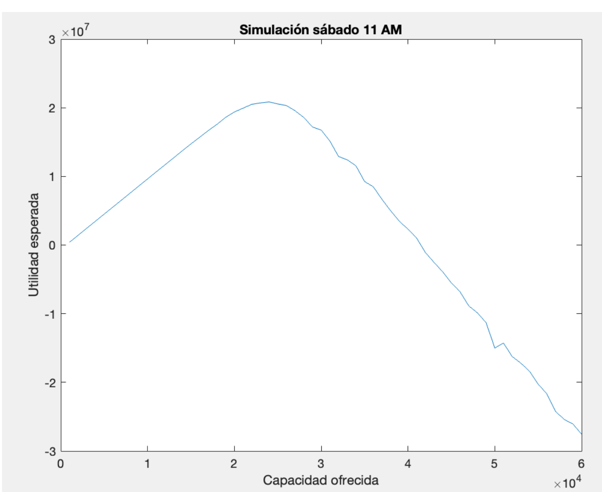

Figura 6. Simulación 11am

En este caso vemos que la utilidad máxima se logra con una oferta entre 23000 y 28000 pasajeros, y de allí en adelante es decreciente. Por encima de 40000 se genera una pérdida.

Analíticamente: $norminv(0{,}4063, 24824, 4045) = 23865$.

Calculamos ahora para las 6 PM de un día martes, la hora de máximo tráfico en la semana según los datos obtenidos:

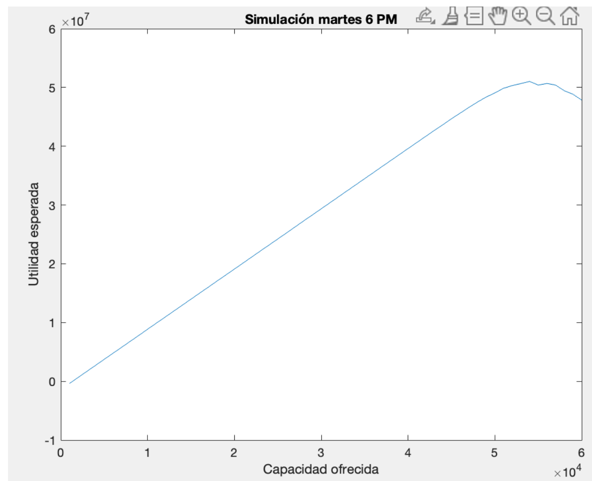

Figura 7. 6 PM de un día martes

Aquí podemos ver que la máxima utilidad se obtiene con una oferta superior a 51000 pasajeros por hora y que en ningún caso la función de utilidad es negativa para este día.
Para las 6pm también tenemos la solución análitica: donde anteriormente encontramos que la demanda sigue una distribución normal con parametros (55715,4952) por lo tanto aplicando la ecuación 7 y 8:

$$F(Q) = \frac{2500 - 1500 + 25}{2500 - 2{,}50 + 25}$$

$$F(Q) = 0{,}4063$$

$$Q = F^{-1}(0{,}4063)$$

Y de esto obtenemos:

$$Q = 54541$$

Lo que índica que el metro el martes a las 6pm debe ofrecer una capacidad máxima de 54541 personas por hora, lo que significa 1 tren cada 2.27 minutos.

Podemos obervar que este resultado es muy cercano al resultado de la simulación, lo que nos índica que la simulación ejecutada está teniendo buenos resultados.

Con el mismo procedimiento podemos calcular la oferta óptima (desde el punto de vista de la función de utilidad),

para cualquier hora del día y cualquier día de la semana, y con base en ella, teniendo en cuenta las cifras de capacidad de pasajeros por tren de 6 vagones indicada anteriormente, determinar la frecuencia óptima de los trenes para dicha hora. Las matrices de medias y desviaciones pueden recalcularse en cualquier momento con los datos históricos.

Como resultado se tendría la frecuencia óptima de trenes por hora del día y día de la semana. Además se podría calcular la frecuencia óptima para fechas especiales en las cuales se espera una demanda adicional o una demanda menor a la de los datos.

VIII. ANÁLISIS DE SENSIBILIDAD

VIII-A. *Sensibilidad a la tarifa*

Tomamos para el ejemplo el sábado a las 10 AM. Calculamos F(Q) para valores de p entre $500 y $5000, variando de a $100, y para cada F(Q) calculamos $F^{-1}(Q)$. Graficamos los precios vs. $F^{-1}(Q)$:

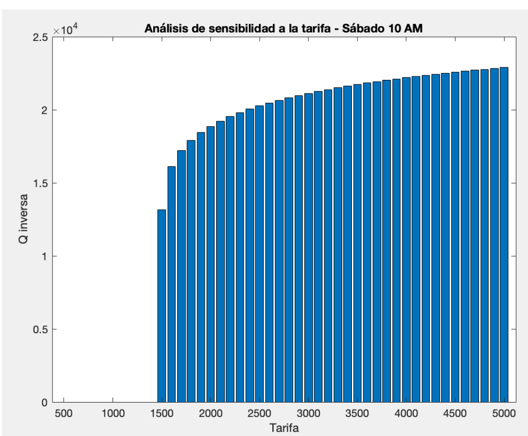

Figura 8. precios vs. $F^{-1}(Q)$

Podemos ver que para tarifas inferiores a $1500 (costo), no tenemos valores en $F^{-1}(Q)$ pues $F(Q) < 0$. De allí en adelante, empieza a incrementarse como era de esperarse.

VIII-B. *Sensibilidad al costo*

De forma similar, tomamos el mismo día sábado a las 10 AM y hacemos el cálculo con valores de c entre $500 y $5000 variando de a $100. Lo podemos observar en el siguiente gráfico (Figura 8):

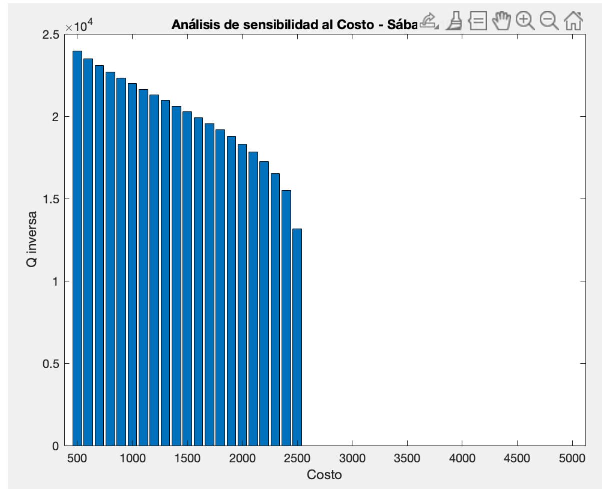

Figura 9. Valores de c entre $500 y $5000 variando de a $100

En este caso, la utilidad es decreciente hasta un costo igual a $2500. De allí en adelante es cero pues $F(Q) < 0$.

IX. TRABAJO FUTURO

Se espera que el modelo pueda extenderse al análisis de media varianza propuesto inicialmente en el estudio hecho por Kalymon[4], donde se emplea el modelo de Markowitz para evaluar el riesgo asociado al rendimiento de activos.
En este contexto, el parámetro α juega un rol fundamental en el análisis de carteras propuesto por Markowitz. El valor de α refleja la disposición del inversionista hacia el riesgo. Cuando α es igual a 0, se prioriza el rendimiento promedio sin considerar el riesgo. A medida que α aumenta, la importancia de reducir el riesgo se intensifica en busca de un equilibrio entre rendimiento y seguridad.
Empleando el parámetro α con $\alpha \geq 0$, si α es 0, se maximiza solamente el beneficio promedio. Sin embargo, a medida que α crece, indica que el tomador de decisiones está más dispuesto a sacrificar beneficio para mitigar el riesgo en la varianza. En otras palabras, no es posible mejorar el beneficio promedio sin asumir un mayor riesgo.
En resumen, el parámetro α juega un papel crucial al permitir que los inversores ajusten la relación entre riesgo y rendimiento en la construcción de carteras óptimas.

Bajo esta mirada, Wu[2] presenta el problema del vendedor de periodicos , que esta dado por:

$$max_{Q \geq 0}\{E[\pi(Q)] - \alpha Var[\pi(Q)]\} \tag{9}$$

Donde α denota la actitud de riesgo del decisor, $E[(Q)]$ la media del beneficio aleatorio y $Var[\pi(Q)]$ es la varianza de

la ganancia aleatoria dada por:

$$Var[\pi(Q)] = E[(\pi(Q)^2)] - (E[\pi(Q)])^2 \tag{10}$$

Con este acercamiento, teniendo un D como futura demanda estocástica durante los tiempos de venta, se puede obtener la cantidad óptima, aplicando un algoritmo de búsqueda unidimesional cuando no se puede obtener la solución óptima.
En futuras investigaciones se espera extender más esta propuesta aplicarla al proyecto.

X. Resultados y conclusiones

Este trabajo es un primer paso para desarrollar una aplicación más estructurada que permita optimizar no solo la operación de la Línea A sino de cualquier línea del Metro, Tranvía o Metrocable. Si se obtienen los datos de tráfico entre estaciones, sería posible determinar rutas expresas, esto es, que no paran en las estaciones de poca demanda a ciertas horas del día. Ello permitiría optimizar aún mas la operación del sistema.

Finalmente, si se logra un autorización del las directivas del Metro para acceder a las cifras de costos reales (aquí trabajamos sobre el presupuesto), se podrían lograr resultados más ajustados a la realidad de la operación del sistema.
Se espera continuar con la aplicación, así con la media varianza propuesta y datos precisos se pueden dar alcances más importantes en la investigación, para que el metro de Medellín tenga mejor desempeño y ganancias. El proceso no solo es aplicar el modelo del vendedor de periódicos al metro, sino que al mejorar este servicio los momento más críticos de transporte no solo disminuyan sino que mejoren significativamete.

En este caso obtuvimos resultados por simulación y de forma analítica. Sin embargo no se toma en cuenta la dinámica de sistemas, correlación entre los datos, costos imprevistos ni costos de reposición de equipos. Esperamos en un futuro poder continuar con esta investigación y así lograr resultados y análisis más amplios.

XI. Bibliografia

[1] Moutaz Khouja,The single-period (news-vendor) problem: literature review and suggestions for future research, Omega, Volume 27, Issue 5, 1999, Pages 537-553, ISSN 0305-0483, Moutaz Khouja

[2] Jun Wu, Jian Li, Shouyang Wang, T.C.E. Cheng, Mean–variance analysis of the newsvendor model with stockout cost, Omega, Volume 37, Issue 3, 2009, Pages 724-730, ISSN 0305-0483, https://doi.org/10.1016/j.omega.2008.02.005. Jun Wu y otros

[3] Presupuesto aprobado 2022 Presupuesto Metro de Medellin - 2022

[4] Kalymon, B. A. , 2009. Estimation Risk in the Portfolio Selection Model. Publicado en línea por Cambridge University Press. Kalymon