



Examen OPI


Solución sección B




Marco Tulio Pérez Ortega

Para resolver este problema es necesario responder las siguientes preguntas.

- ¿Qué variables influyen en la venta de paletas?
- ¿Qué días se pudieron vender más paletas?
- ¿Cuál es el promedio de ventas al día bajo ciertos parámetros?
- ¿Qué criterio de decisión debemos adoptar para surtir una máquina expendedora?



Parte I: ¿Qué variables influyen en la venta de paletas?




Variables que deben ser consideradas

Tenemos datos históricos, por lo que para cada máquina podemos saber y considerar como variables predictoras:


- Día de la semana.
- Mes.
- Día del mes.
- El día es feriado.

Variables que con las que no se cuenta pero pueden ser importantes y no difíciles de obtener son:

- Características del clima
 - Temperatura.
 - Fuerza de viento.
 - Humedad relativa.



Parte II: ¿Qué días se pudieron vender más paletas?




¿Qué días se pudieron vender más paletas?

Esta información puede ser obtenida mediante el número de paletas disponibles al final del día, si el número de paletas disponibles al final del día es cero entonces existe la posibilidad de haber vendido más paletas ese día.


A este tipo de días los llamaremos **d_f**.

A los días en los que sobraron paletas les llamaremos **d_i**.

De igual forma para las ventas denotaremos **v_f** las ventas de **d_f** y **v_i** a las ventas de **d_i**.



Parte II_a: ¿Cómo estimamos el número de paletas que pudieron ser vendidas?



Imputación mediante el vecino más cercano

1. Para un **d_f** buscamos un **d_i** con el mayor número de características en común (que comparte día de la semana, mes, es feriado, etc.)
2. Imputamos usando el día que tenga la venta máxima, de tal forma que si el **d_i** más cercano tiene una venta menor nos seguiremos quedando con la venta del **d_f**.

$$\text{ventas} = \max(v_i, v_f)$$

3. Por efectos de control debemos guardar el porcentaje histórico de **d_f** para cada máquina expendedora.



Parte III: ¿Cuál es el promedio de ventas al día
bajo ciertos parámetros?



Promedio de ventas al día bajo ciertos parámetros

Para responder esto, para cada máquina expendedora debemos :

- Seleccionar las variables predictoras.
 - Regresión lineal
 - Validación cruzada.
- Probar y seleccionar el modelo de predicción.
 - Validación cruzada sobre el error en un conjunto de prueba
- Predecir el promedio de ventas para los siguientes k días.
- Calcular un error estándar.



Parte IV: ¿Qué criterio de decisión debemos adoptar para surtir una máquina expendedora?



Criterio de decisión.

Sea c el número de paletas que hay en la máquina y x la variable aleatoria que representa el número de paletas vendidas ese día. Si $c < \text{capacidad de la máquina}$

- El criterio de decisión para cada máquina puede estar dado por:

Si $p(x > c) > 0.02$, Surtir máquina

Podemos considerar que x sigue o bien una distribución de Poisson o una distribución normal, ambas con la media estimada anteriormente.



Parte IV_a: ¿Cuántas paletas debemos surtir cada
a máquina expendedora?



Criterio de decisión.

Aquí debemos considerar distintos casos:

- Caso I: El número de ventas diario es cercano o igual a la capacidad de la máquina.
- Caso II: El número de ventas diario esperado es superior a 100 paletas.
- Caso III: El número de ventas diario esperado no es superior a 100 paletas.

Número de paletas por máquina expendedora

Para los primeros dos casos el número de paletas que se deben llevar es

$$\text{numero_paletas} = \max(\text{capacidad} - c, X)$$

Donde **X** está dado por la probabilidad:

$$p(x > X + c) < 0.02$$

Donde **c** es el número de paletas que hay en la máquina, **x** la variable aleatoria que representa el número de paletas vendidas ese día.

Número de paletas por máquina expendedora

Para el tercer caso debemos considerar que el número de paletas que debe haber en cada máquina expendedora al final del día debe ser menor a 100, de tal forma de que sea más conveniente no surtir esa máquina expendedora al siguiente día y al mismo tiempo la máquina no se quede sin paletas al siguiente día. Es decir buscamos **X** minimice:

$$p(X+c-x_1>100)+p(X + c - x_1 - x_2=0)$$

Donde **c** es el número de paletas que hay en la máquina, **x₁** y **x₂** las variables aleatorias que representan el número de paletas vendidas el día uno y el día dos respectivamente.

Finalmente el número de paletas está dado por:

$$\text{numero_paletas}=\max(\text{capacidad}-c, X)$$

Gracias por su atención!