

EJERCICIO 1

Aplicación de algoritmos de regresión

Profesora: Noemí Aedo Noa

I. Contexto del problema

El conflicto palestino-israelí ha generado desplazamientos forzados masivos, dejando a miles de personas en situación de vulnerabilidad. A pesar de numerosos informes cualitativos, existe una carencia de análisis cuantitativos robustos que permitan caracterizar las áreas más afectadas y las condiciones que aumentan la probabilidad de desplazamientos forzados.

La Cruz Roja Internacional ha decidido emprender un análisis de datos para abordar esta brecha de conocimiento, utilizando un modelo de regresión logística para identificar factores clave que contribuyen al desplazamiento forzado en territorio palestino.

Es importante destacar que, debido a las restricciones de acceso en ciertas áreas, los datos recopilados no abarcan la totalidad del territorio palestino, sino las regiones donde fue posible ingresar y recolectar información confiable. Por lo tanto, el análisis se centra en estas zonas y no pretende generalizar los resultados a lugares donde no fue posible recopilar datos. Los resultados del modelo servirán como base para futuros estudios y estrategias de intervención en la región.

El objetivo es clasificar si una región tiene alta probabilidad de haber experimentado desplazamiento forzado (1 = Sí, 0 = No) utilizando variables predictoras relacionadas con condiciones socioeconómicas, infraestructura y violencia.

II. Variables disponibles

- Desplazamiento (desplazamiento_forzado): Variable objetivo (1 = Sí, 0 = No).
- Ataques (numero_ataques): Número de incidentes violentos reportados en la región Palestina.
- Población desplazada (poblacion_desplazada): Número estimado de personas desplazadas.
- Acceso a servicios básicos (acceso_servicios): Proporción de la población con acceso a agua, electricidad y atención médica (0-100%).
- Nivel de destrucción (nivel_destruccion): Grado de destrucción de infraestructura en una escala del 1 al 5.
- Índice de pobreza (indice_pobreza): Porcentaje de la población bajo la línea de pobreza.
- Presencia militar (presencia_militar): Índice que mide la cantidad de tropas activas en la región (1-10).

III. Resultados

Calcule los estadísticos descriptivos de cada variables, los coeficientes del modelos y las métricas de rendimiento (sensibilidad, especificidad y precisión).

Tabla 1. Estadísticos descriptivos

Variable	Min	Max	Median	Mean
Desplazamiento forzado				
Número de ataques				
Población desplazada				
Acceso a servicios				
Nivel de destrucción				
Índice de pobreza				
Presencia militar				

Tabla 2. Coeficientes de regresión

Variable	Estimación del coeficiente	Error estándar	z-valor	p-valor
Intercepto				
Desplazamiento forzado				
Número de ataques				
Población desplazada				
Acceso a servicios				
Nivel de destrucción				
Índice de pobreza				
Presencia militar				

Tabla 3. Matriz de confusión del modelo elaborado

Precisión	Sensibilidad	Especificidad

IV. Preguntas

1. ¿Qué responsabilidad tienen los científicos de datos, antropólogos/arqueólogos y organismos internacionales en la forma en que estos modelos son usados en conflictos?. **Indiquen brevemente al menos dos responsabilidades para cada cargo (informáticos, antropólogos/arqueólogos y organismos internacionales).**
2. ¿Qué variables agregarían al modelo para robustecer los resultados?. **Indiquen al menos dos variables.**

3. Considerando los resultados obtenidos en la matriz de confusión, ¿de qué manera interpretarías los valores obtenidos para cada métrica (sensibilidad, especificidad y precisión) según el contexto del problema?

EJERCICIO 2

Aplicación de algoritmos de regresión

Profesora: Noemí Aedo Noa

I. Contexto del problema:

Durante la dictadura militar de Chile, miles de personas fueron detenidas y desaparecidas. En muchos de estos casos, se reportaron testimonios de torturas físicas, incluidas fracturas y otros tipos de traumas óseos como resultado de la violencia ejercida.

El objetivo de este análisis es construir un modelo de regresión logística que permita predecir la presencia de traumas óseos (variable dependiente) en función de diversas características relacionadas con el lugar y contexto de la desaparición.

Desarrollar un modelo predictivo utilizando datos históricos recopilados por organizaciones de derechos humanos, con el fin de identificar los factores que más influyen en la aparición de traumas óseos en personas detenidas y desaparecidas.

II. Variables disponibles:

- Trauma óseo (trauma_oseo): Variable objetivo (1 = Sí, 0 = No). Indica si la persona sufrió un trauma óseo durante su detención o desaparición.
- Zona geográfica (zona_geografica): Región de Chile donde se produjo la desaparición (1 = Norte, 2 = Centro, 3 = Sur).
- Violencia reportada (violencia_reportada): Nivel de violencia ejercida durante la detención (escala de 1 a 5, donde 1 es baja violencia y 5 es violencia extrema).
- Tiempo de detención (tiempo_detencion): Número de días que la persona estuvo detenida antes de ser liberada o desaparecer.
- Presencia de tortura (presencia_tortura): Si se reportó tortura física (1 = Sí, 0 = No).
- Acceso a atención médica (acceso_medico): Nivel de acceso a atención médica durante la detención (escala de 1 a 5, donde 1 es sin acceso y 5 es acceso total).
- Nivel de pobreza (indice_pobreza): Porcentaje estimado de pobreza en la zona donde se realizó la detención.
- Presencia militar (presencia_militar): Grado de presencia militar en la región (escala de 1 a 10, donde 1 es mínima presencia y 10 es máxima presencia).

III. Resultados

Calcule los estadísticos descriptivos de cada variables, los coeficientes del modelos y las métricas de rendimiento (sensibilidad, especificidad y precisión).

Tabla 1. Estadísticos descriptivos

Variable	Min	Max	Median	Mean
Zona geográfica				
Violencia reportada				
Tiempo de detención				
Presencia de tortura				
Acceso médico				
Índice de pobreza				
Presencia militar				

Tabla 2. Coeficientes de regresión

Variable	Estimación del coeficiente	Error estándar	z-valor	p-valor
Intercepto				
Zona geográfica				
Violencia reportada				
Tiempo de detención				
Presencia de tortura				
Acceso médico				
Índice de pobreza				
Presencia militar				

Tabla 3. Matriz de confusión del modelo elaborado

Precisión	Sensibilidad	Especificidad

IV. Preguntas

1. ¿Qué variables agregarían al modelo para robustecer los resultados?. Indiquen al menos dos variables.

2. Identifiquen qué variables estarían teniendo un impacto estadísticamente significativo sobre los traumas óseos