

DETERMINACIÓN DE ESTADOS DE RIESGO EN PACIENTES CARDIOVASCULARES UTILIZANDO REGRESIÓN LOGÍSTICA

DE LUCA, María José; JAGOU, Nancy Elizabeth y LEÓN, María Natalia

Centro de Investigación, Facultad de Ciencias Económicas – UNaM
Campus Universitario - Ruta 12 km 7 1/2 - (3300) - POSADAS - Pcia. de Misiones

Un grupo de investigadores del área Matemática Aplicada, de la Facultad de Ciencias Económicas de la Universidad Nacional de Misiones, junto a médicos de la Fundación de Cardiología de Misiones (FUCAM), hemos iniciado una serie de trabajos interdisciplinarios a los fines de optimizar los procesos de tratamientos preventivos de cardiopatías en Misiones.

En esta dirección, una de las cuestiones consideradas fue la necesidad de mejorar la atención primaria de pacientes cardiovasculares unificando los criterios de asignación de riesgo por parte del personal de hospitales y sanatorios. Actualmente existen normas tabuladas que son utilizadas por los profesionales, que discriminan 5 categorías de niveles de riesgo¹.

Se pensó en generar un modelo que replique la clasificación utilizada y estudiar el poder de discriminación de dicha clasificación. Debido a las características de la variable de interés (ESTADO DE RIESGO) y de las variables explicativas (EDAD, PRESIÓN DIASTÓLICA, COLESTEROL, GLUCEMIA, TABAQUISMO Y ACTIVIDAD FÍSICA) fue necesario recurrir a un modelo que contemple más de dos categorías en la variable respuesta y que permita combinar variables continuas y discretas como variables explicativas, la **regresión logística politómica**.

Este modelo, con las mediciones iniciales de ciertos parámetros médicos influyentes, asigna el estado de riesgo más probable con respecto a las afecciones consideradas.

La expresión matemática que permite calcular la probabilidad de que un sujeto de la muestra pertenezca a la subpoblación j , es:

$$\pi_{j(x)} = \Pr(Y = j / \mathbf{x}) = \frac{\exp[g_j(\mathbf{x})]}{1 + \sum_{m=1}^{j-1} \exp[g_m(\mathbf{x})]} \quad j = 1, 2, \dots, J-1$$

$$\pi_{J(x)} = \Pr(Y = J / \mathbf{x}) = \frac{1}{1 + \sum_{m=1}^{j-1} \exp[g_m(\mathbf{x})]}$$

donde $g_j(\mathbf{x}) = \beta_{j0} + \sum_{k=1}^K \beta_{jk} x_k$, con $j = 1, 2, \dots, J-1$ y J la subpoblación de referencia.

La variable respuesta Y puede caer en una de las J clases posibles, codificada como $1, 2, \dots, J$, y está expresada en una escala nominal, mientras \mathbf{x} representa al vector de covariables de longitud K .

De los modelos analizados aquel que discrimina en tres estados de riesgo es el que posee un nivel de ajuste satisfactorio. Los estados definidos pueden interpretarse como de BAJO, MODERADO y ALTO RIESGO y las variables a considerar EDAD, PRESION DIASTOLICA, GLUCEMIA y la interacción entre SEXO y TAF².

Las logit que permiten estimar las probabilidades de pertenecer a uno de los tres niveles de riesgo, dado los valores de las variables predictoras para un sujeto son:

¹ Sin riesgo, bajo, moderado, alto y muy alto riesgo.

² Tabaquismo, actividad física y colesterol.

$$\hat{g}_1(x) = 39,523 - 0,141 \text{ EDAD} - 0,279 \text{ PDIAS} - 0,061 \text{ GLUCE} - 1,131 \text{ SEXO[F]} * \text{T AFC} - 1,742 \text{ SEXO[F]} * \text{T AFC}$$

$$\hat{g}_2(x) = 21,750 - 0,034 \text{ EDAD} - 0,157 \text{ PDIAS} - 0,035 \text{ GLUCE} - 0,626 \text{ SEXO[F]} * \text{T AFC} - 0,622 \text{ SEXO[F]} * \text{T AFC}$$

Usando estas estimaciones de las dos funciones logit, es posible crear un nuevo conjunto de datos que contenga las probabilidades estimadas de ser miembro de cualquiera de los tres grupos¹. El valor mayor de probabilidad se corresponde con la clasificación pronosticada para cada individuo.

En la siguiente tabla se muestra el nivel de ajuste entre los valores observados y los pronosticados por el modelo; los elevados porcentajes de clasificación correcta, hablan del ajuste satisfactorio del modelo.

Observed	Predicted			Percent Correct
	1	2	3	
1	162	41	0	79,8%
2	36	195	15	79,3%
3	0	35	88	71,5%

Análisis de resultados

Del estudio realizado se concluye que:

- las variables que inciden en forma directa en la determinación del estado de riesgo para enfermedades cardiovasculares son: EDAD, PRESION DIASTOLICA y GLUCEMIA.
 - la variable COLESTEROL no debe incluirse en el modelo como una variable continua sino como una variable dicotómica al igual que el TABAQUISMO y la ACTIVIDAD FÍSICA. Los valores 1 y 0 de estas variables indicarán presencia o ausencia del factor de riesgo, respectivamente,
 - la interacción que resulta significativa es SEXO × T AFC, donde T AFC (es la variable que permite acumular factores de riesgo cuando se da la presencia de COLESTEROL, TABAQUISMO y/o ACTIVIDAD FÍSICA). Esta interacción significativa indica un efecto diferente de estos factores de riesgo, según el SEXO.
 - las variables utilizadas por los criterios de clasificación recomendados por las organizaciones de salud, sólo permiten diferenciar con significancia estadística, tres niveles de riesgo.
- Es posible elaborar cuadros más simple y accesibles al personal que tiene a cargo la atención primaria de los pacientes.

¹ Bajo, moderado y alto riesgo.