



*Nombre del Alumno: Leticia Desiree Morales Aguilar*

*Nombre del tema: Conceptos de análisis*

*Segundo parcial*

*Nombre de la Materia: Salud pública.*

*Nombre del profesor: DR. Guillermo del Solar*

*Nombre de la Licenciatura: Medicina Humana*

## INTRODUCCIÓN

El análisis multivariable no es fácil de definir. En general, se refiere a aquellos métodos estadísticos que analizan simultáneamente diversas variables en cada individuo u objeto sobre el cual se investiga. Cualquier análisis simultáneo demás de dos variables, puede considerarse análisis multivariable. De hecho, muchas técnicas multivariable son la simple extensión de análisis univariados bivariados. Así, por ejemplo, la regresión simple (con una sola variable independiente), es una técnica multivariable cuando se extiende a varios regresores. Otras técnicas, sin embargo, como el análisis factorial o el análisis discriminante, están específicamente diseñadas para trabajar únicamente con estructuras multivariadas.

# Conceptos básicos de análisis multivariable

fenómenos de interés para la salud pública también tienen habitualmente múltiples causas.

- Variable dependiente (outcome, variable «y»): resultado o desenlace de interés.

• Variables independientes (predictors, variables «x»):

factores más fáciles de medir o recoger y que, según la hipótesis de estudio, pueden relacionarse con el resultado de interés. Las variables independientes pueden ser:

• Variable independiente principal: exposición (exposure) bajo estudio.

• Variables de confusión (confounding factors): otras variables independientes que se relacionan tanto con la variable dependiente como con la exposición y que pueden distorsionar la relación entre ambas.

Ejemplo sencillo de regresión logística:

Supongamos que se comparan 16 bebedores de alcohol ( $x = \text{alcohol} = 1$ ) frente a 44 abstemios ( $x = \text{alcohol} = 0$ ).

En los 16 bebedores ocurren 6 accidentes de tráfico ( $y = \text{AT} = 1$ ), pero en los abstemios ocurren sólo 4.

Variables con más de dos categorías: variables dummy

Cuando se desea introducir como independiente una variable cualitativa que tenga 3 o más categorías, se debe elegir primero cuál será la categoría de referencia (el suelo frente al que se compararán las demás) y crear una nueva variable para cada una de las demás categorías, es decir, se crearán tantas variables como categorías tenga la variable original menos una.

Variables independientes continuas:

Trabajar con quintiles haciendo 4 variables dummies suele ser un primer paso para valorar la relación entre una variable independiente continua (edad, kilómetros conducidos) y el acontecimiento de resultado (variable dependiente).

Predicciones y probabilidades absolutas:

La regresión logística permite hacer predicciones sobre probabilidades absolutas de presentar un acontecimiento. Si miramos la tabla 5-4 e imaginamos un varón ( $\text{sexo} = 1$ ) que tiene 26 años (está en el quintil inferior de la edad), que conduce 5.000km al año y que es bebedor ( $\text{alcohol} = 1$ ), su ecuación sería:

$$\ln(\text{odds de AT}) = a + b_1 \text{ alcohol} + b_2 \text{ sexo} + b_3 \text{ km1000} + b_4 \text{ edad}(\text{dummy}_i)$$

Modelo de regresión lineal múltiple

Las cosas se simplifican mucho cuando la variable dependiente es cuantitativa (peso, colesterol total, presión arterial). Entonces el modelo queda así:

$$y = a + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4 \dots + b_i x_i$$

Regresión de cox:

es una técnica muy difundida, que se parece mucho a la regresión logística, pero requiere tener en cuenta, además, los tiempos variables de seguimiento de cada participante en un estudio.

Tres usos principales del análisis multivariable:

Los usos principales de estos modelos de regresión multivariable son tres y se representan esquemáticamente en la figura.

- El uso más importante y más frecuente es el de controlar la confusión.
- El segundo uso es valorar la interacción o modificación del efecto.

El tercer uso es introducir términos cuadráticos o cúbicos que permitan que las relaciones sean en forma de «U» o sigan cualquier otro modelo distinto de la rígida línea recta

## CONCLUSIÓN

Los análisis estadísticos actuales intentan explicar un fenómeno teniendo en consideración varias variables simultáneamente y por eso se llaman análisis multivariados. Casi toda la estadística que se usa en salud pública consiste hoy en análisis multivariados.

Corresponden a diversos modelos de regresión. Lo más habitual es que se trate de valorar un desenlace o resultado relevante para la salud pública en función de una o más variables predictoras (covariables). Se tiene en cuenta, además, el posible papel de otras variables llamadas factores de confusión que sesgarían el análisis si no se controlasen adecuadamente. Los modelos multivariados también son útiles para valorar si los efectos de una variable independiente se modifican por otra (interacción) y para estudiar mejor la forma de las relaciones dosis-respuesta.

# Bibliografía

*Miguel, M. Conceptos de salud publica y estrategias preventivas.*