**Regresión logística (Regresión logística.sav)**

Tomando como punto de partida el creciente aumento en la tasa de suicidio adolescente, un grupo de investigadores japoneses decide investigar las variables psicológicas de riesgo y protección ante el suicidio en adolescentes de 3ro a 5to de secundaria de colegios de varios distritos de la Tokio. Los investigadores tenían particular interés en estudiar la probable conexión directa entre la depresión (medida con el Inventario de Depresión de Beck) y el haber tenido al menos algún intento de suicido.

Utilice la variable “Depresión” y determine si es que la hipótesis de los investigadores es estadísticamente probable. Al ser una variable dicotómica, los adolescentes sólo podían responder si habían intentado suicidarse o no. Para aquellos casos en los que la variable dependiente es dicotómica debes utilizar una Regresión con un modelo logístico. Estos son los pasos que debes seguir:

1. Para corroborar el funcionamiento global del modelo se utiliza la prueba ómnibus. Revisa la significación y evalúa la calidad del modelo.

LOGISTIC REGRESSION VARIABLES Suicidio

/METHOD=ENTER Depresión

/CRITERIA=PIN(.05) POUT(.10) ITERATE(20) CUT(.5).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Pruebas omnibus sobre los coeficientes del modelo** | | | | |
|  | | Chi cuadrado | gl | Sig. |
| Paso 1 | Paso | 8,858 | 1 | ,003 |
| Bloque | 8,858 | 1 | ,003 |
| Modelo | 8,858 | 1 | ,003 |

La significación es menor a .05 (p = .003). El modelo es, por tanto, significativo.

b) La regresión logística tiene una forma de medir ajuste similar a la de la regresión lineal. Sin embargo, en el caso de la regresión logística los estadísticos que miden ajuste son llamados Pseudo R2 y no se interpretan exactamente de la misma manera. Para este curso utilizaremos el R cuadrado de Nagelkerke. Utilizando los cuadros de “Ajuste del modelo” responde ¿Cuál es el ajuste del Modelo?

|  |  |  |  |
| --- | --- | --- | --- |
| **Resumen del modelo** | | | |
| Paso | -2 log de la verosimilitud | R cuadrado de Cox y Snell | R cuadrado de Nagelkerke |
| 1 | 220,287a | ,048 | ,067 |
| a. La estimación ha finalizado en el número de iteración 4 porque las estimaciones de los parámetros han cambiado en menos de ,001. | | | |

Asimismo, podemos concluir que hay una verosimilitud de 6,7% en este modelo en comparación con un modelo en el que no se postula ninguna VI como asociada a la tentativa de suicidio. El modelo es prácticamente irrelevante.

c) Recuerda que una de las funciones de la Regresión es predecir. Extrae la Tabla de clasificación (o Matriz de confusión) y determina qué tan adecuado es el pronóstico resultado del modelo ¿Qué se pronostica mejor? ¿La presencia o ausencia de tentativa de suicidio?

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Tabla de clasificacióna** | | | | | |
|  | Observado | | Pronosticado | | |
|  | Ha intentado suicidarse | | Porcentaje correcto |
|  | No | Si |
| Paso 1 | Ha intentado suicidarse | No | 108 | 12 | 90,0 |
| Si | 55 | 5 | 8,3 |
| Porcentaje global | |  |  | 62,8 |
| a. El valor de corte es ,500 | | | | | |

Se concluye que el modelo clasifica bien el 90% de los casos en los que no hubo intento de suicidio y sólo el 8,3% de casos donde sí hubo. El modelo no es del todo adecuado, al no predecir adecuadamente a los que sí han tenido al menos un intento de suicidio.

d) Para evaluar el “sentido” de cada una de las variables del modelo revisa el Beta y su significación en el cuadro de variables en la ecuación ¿Es significativa la variable identificada? ¿La relación es directa o inversa?

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Variables en la ecuación** | | | | | | | |
|  | | B | E.T. | Wald | gl | Sig. | Exp(B) |
| Paso 1a | Depresión | -,040 | ,014 | 8,504 | 1 | ,004 | ,961 |
| Constante | 1,279 | ,686 | 3,475 | 1 | ,062 | 3,593 |
| a. Variable(s) introducida(s) en el paso 1: Depresión. | | | | | | | |

Depresión tiene B1 = -,040. Es decir, a mayor depresión, menor probabilidad de haber intentado suicidarse.

e) Finalmente, para evaluar la importancia relativa de la(s) variables significativas utilizarás Exp(B). Este indicador se interpreta de la siguiente manera: Por cada punto adicional en la variable “Depresión” la nueva probabilidad de haber intentado suicidarse (Recuerda que la variable interés es dicotómica, donde 1=No y 2=Si) es igual a Exp(B) veces más que la anterior ¿Cuál es el puntaje Exp (B) en nuestro caso?

En nuestro caso, por cada punto adicional en la variable depresión, la nueva probabilidad de haber intentado suicidarse es de .961.