

**UNIVERSIDAD NACIONAL DE SAN ANTONIO**

**ABAD DEL CUSCO**

**FACULTAD DE CIENCIAS QUIMICAS, FISICAS Y MATEMATICAS**

**CARRERA PROFESIONAL DE MATEMATICA**



**“FACTORES ASOCIADOS DE RIESGO PARA QUE  
UNA PERSONA MUERA O SEA DIAGNOSTICADA  
CON EL VIRUS A (H1N1) MEDIANTE EL MODELO DE  
REGRESION LOGISTICA, EN EL DEPARTAMENTO  
DEL CUSCO, 2009”**

Tesis presentada por:

**Br: EFRAIN RODAS GUIZADO**

**Para Optar al Título Profesional de Licenciado en  
Matemática Mención Estadística.**

**ASESORA:**

***Mgt: Nelly M. Salazar Peña***

**CO-ASESOR:**

***Lic: Ray Milton Valderrama Álvarez***

Auspicia Concejo de Investigación de la “UNSAAC”

**CUSCO – PERU**

**2011**

## **PRESENTACION**

SEÑOR DECANO DE LA FACULTAD DE CIENCIAS QUIMICAS, FISICAS Y MATEMATICAS.

SEÑOR COORDINADOR DE LA CARRERA PROFESIONAL DE MATEMATICAS.

SEÑORES CATEDRATICOS MIEMBROS DEL JURADO.

Con el fin de optar el Título Profesional de Matemática Mención Estadística y en cumplimiento del Reglamento de Grados y Títulos de la Carrera Profesional de Matemáticas de la Facultad de Ciencias Químicas, Físicas y Matemáticas de la Universidad Nacional de San Antonio Abad del Cusco, presento a vuestra consideración el trabajo de tesis intitulada **“FACTORES ASOCIADOS DE RIESGO PARA QUE UNA PERSONA MUERA O SEA DIAGNOSTICADA CON VIRUS A (H1N1) MEDIANTE EL MODELO DE REGRESION LOGISTICA, EN EL DEPARTAMENTO DEL CUSCO, 2009”**.

En este trabajo de investigación se presenta la teoría y aplicación de los Modelos de Regresión Logística para conocer cuáles son los Factores de Riesgo que influyen para que una persona sea diagnosticada positiva o negativa con virus A (H1N1). La razón principal de este estudio es identificar los Factores más significativos de riesgo de la población en estudio.

**Atentamente:**

**Br. EFRAIN RODAS GUIZADO.**

## **AGRADECIMIENTOS**

### **QUIERO EXPRESAR MI AGRADECIMIENTO:**

A mis asesores de tesis Mgt. Nelly M. Salazar Peña y Lic. Ray Milton Valderrama Álvarez por su generosidad y predisposición al brindarme la oportunidad de recurrir a su capacidad y experiencia en un marco de confianza, afecto y amistad, fundamentales para la concreción de este trabajo.

También expreso mi agradecimiento a todos los docentes del departamento académico de Matemáticas y Estadística de la Universidad de San Antonio Abad del Cusco por sus enseñanzas en mi formación profesional.

**Efrain**

# INDICE

## CAPITULO I

### GENERALIDADES

1.1 DESCRIPCION DEL PROBLEMA.....	02
1.2 FORMULACIÓN DEL PROBLEMA.....	03
1.3 FORMULACIÓN DE OBJETIVOS.....	03
1.3.1 Objetivos general.....	03
1.3.2 Objetivos específicos.....	03
1.4 HIPOTESIS.....	04
1.5 JUSTIFICACION.....	04
1.6 PROCEDIMIENTO METODOLOGICO.....	04

## CAPITULO II

### MARCO TEORICO

2.1 ANTECEDENTES DEL PROBLEMA.....	05
2.2 FUNDAMENTOS TEORICOS.....	07
2.2.1 MODELOS LINEALES GENERALIZADOS.....	07
2.2.2 MODELO DE REGRESION CON RESPUESTA DICOTOMICA.....	09
2.2.3 INTRODUCCIÓN A LA REGRESION LOGISTICA.....	09
2.3 REGRESION LOGISTICA.....	10
2.3.1 DEFINICIÓN.....	10
2.3.2 FUNCIÓN DE VEROSIMILITUD.....	10
2.3.3 MODELO DE REGRESIÓN LOGÍSTICA BINARIA.....	11

2.3.4 FUNCIÓN DE VEROSIMILITUD.....	12
2.3.5 ESTIMACIÓN DE LOS PARÁMETROS.....	12
2.3.6 MODELO DE REGRESIÓN LOGÍSTICA SIMPLE.....	13
2.4 REGRESION LOGISTICA MULTIPLE.....	15
2.4.1 MODELO DE REGRESION LOGISTICA MULTIPLE.....	15
2.4.2 AJUSTE DEL MODELO DE REGRESION LOGISTICA MULTIPLE.....	16
2.4.3 METODO DE NEWTON – RAPHSON PARA ESTIMAR LOS PARAMETROS DEL MODELO DE REGRESION LOGISTICA.....	22
2.4.4 INTERPRETACION DE LOS COEFICIENTES DEL MODELO ESTIMADO.....	25
2.4.5 PRUEBA DE HIPOTESIS PARA LOS COEFICIENTES DEL MODELO DE REGRESION LOGISTICA.....	26
2.4.5.1 PRUEBA DE WALD.....	26
2.4.5.2 PRUEBA CHI-CUADRADO.....	28
2.4.5.3 ESTADISTICA CHI-CUADRADA DE PEARSON.....	29
2.4.5.4 DESVIANZA.....	29
2.4.5.5 TABLA DE CLASIFICACION.....	30
2.4.5.6 CONTRASTE BONDAD DE AJUSTE DE HOSMER- LEMESHOW.....	31
2.4.5.7 EL INCREMENTO DEL ESTADÍSTICO $-2\log L$ .....	32
2.5 DIAGNOSTICO DEL MODELO.....	33
2.5.1 EVALUACION DE $-2LL_0$ .....	33
2.5.2 RIESGO RELATIVO Y ODDS RATIO.....	34
2.5.3 RIESGO RELATIVO (RR).....	34

2.5.4 RAZON DE PREDOMINIO (OR).....	35
2.5.5 RESIDUOS DEL MODELO.....	36
2.5.6 MEDIDAS DE INFLUENCIA.....	37
2.6 VARIABLES EXPLICATIVAS CATEGORICAS – VARIABLES DUMMY.....	38
2.7 INTERPRETACIÓN DE LOS RESULTADOS.....	39
2.8 CONSIDERACIONES GENERALES.....	40
2.8.1 Definición de la patología.....	40
2.8.2 Etimología.....	40
2.8.3 Fisiopatología.....	42
2.8.4 Aspectos Epidemiológicos.....	43
2.8.5 Transmisión.....	44
2.9 Factores de riesgo asociados.....	45
2.9.1 Medio Ambiente.....	45
2.9.2 Estilos de vida.....	45
2.9.3 Grupos de Alto Riesgo.....	45
2.10 CONSIDERACIONES ESPECÍFICAS.....	46
2.10.1 Cuadro Clínico.....	46
2.10.1.1 Grupos de signos y síntomas relacionados con la patología.....	46
2.10.2 Diagnostico.....	46
2.10.2.1 Definiciones de Caso.....	46
2.10.2.2 Criterios de diagnostico.....	47
2.10.2.3 Diagnostico diferencial.....	48

<b>2.10.2.4</b> Signos de Alarma.....	49
<b>2.10.2.5</b> Criterios de Alta.....	49
<b>2.10.2.6</b> Pronóstico.....	50
<b>2.11</b> Fases de la Pandemia de Influenza según la Organización Mundial de Salud.....	52
<b>2.12</b> Obtención de la muestra para diagnóstico de Influenza A H1N1.....	53
<b>2.13</b> Información para la Comunidad.....	54

### **CAPITULO III**

#### **APLICACIÓN DEL ESTUDIO DEL VIRUS DE INFLUENZA A (H1N1)**

<b>3.1</b> DISEÑO.....	57
<b>3.2</b> ANALISIS DE LOS DATOS.....	59
<b>3.2.1</b> ANALISIS EXPLORATORIO PREVIO.....	59
<b>3.3</b> ANALISIS DE REGRESION LOGISTICA BINARIA CON CADA UNOS DE LOS FACTORES DE RIESGO.....	61
<b>3.4</b> Aplicando el Método Por pasos hacia atrás (Wald).....	66
<b>3.5</b> MEJOR MODELO DEL ESTUDIO.....	71
<b>CONCLUSIONES Y RECOMENDACIONES</b> .....	74
<b>CONCLUSIONES</b> .....	74
<b>RECOMENDACIONES</b> .....	75
<b>BIBLIOGRAFIA</b> .....	76
<b>ANEXOS</b> .....	78

ANEXO 1.....	79
ANEXO 2.....	80
ANEXO 3.....	86
ANEXO 4.....	94
ANEXO 5.....	98



## INTRODUCCION

La influenza es una enfermedad respiratoria aguda altamente transmisible de importancia global, que ha causado epidemias y pandemias por siglos.

En el siglo XX, en el mundo se han producido tres grandes pandemias, la primera, llamada “La Gripe Española” entre 1918 y 1919, durante la Primera Guerra Mundial, esta pandemia se caracterizó por ser de letalidad alta, se estima entre 40 y 100 millones de defunciones a nivel mundial. Al parecer, recién en marzo de 1919 llegó al Perú; sin embargo, no se cuentan con datos precisos sobre la mortalidad y letalidad causada en la población de nuestro país en esa época.

Esta pandemia fue estudiada por el Dr. Aníbal Corvetto Bisagno, quien estableció las relaciones de la influenza con la tuberculosis y concluyó que la gripe no siempre es un factor desfavorable para la tuberculosis. Desgraciadamente, el Dr. Corvetto adquirió la influenza y su salud se agravó seriamente.

La última pandemia se presentó en 1968, conocida como “La Gripe Asiática”, se extendió rápidamente por todo el mundo, la morbilidad fue muy alta; sin embargo, fue de menor letalidad que las anteriores. Desde entonces, la influenza se comporta como una enfermedad estacional que afecta a la población peruana durante el invierno. En los últimos años, se ha descrito brotes epidémicos de influenza en algunas comunidades indígenas amazónicas, en especial en aquellas de reciente contacto por parte de la Red Nacional de Epidemiología.

## CAPITULO I

### GENERALIDADES

#### PLANTEAMIENTO DE LA INVESTIGACION

##### 1.1. DESCRIPCION DEL PROBLEMA.

En este contexto desde que se declaró la alerta mundial por la aparición de un nuevo virus de Influenza A (H1N1), y la rápida propagación de éste por el mundo hasta alcanzar el grado VI de pandemia, declarada por la OMS. Se emitió a nivel del País con fecha 24 de Abril del 2009. La alerta epidemiológica Nacional, por parte de la Dirección General de Epidemiología-MINSA (D.G.E-MINSA), el 27 de Abril la DIRESA Cusco, emite la Alerta Epidemiológica Regional, el 28 de Abril se instala el Equipo de Alerta Respuesta en el Aeropuerto Internacional Velazco Astete de Cusco, el 29 de Abril, se Actualiza el Plan Regional de Contingencia y Respuesta frente a una posible pandemia de la Nueva Influenza A (H1N1) en Cusco, cuyo presupuesto estimado en 2'339,202.00 Nuevos Soles, fue aprobado por el gobierno regional Cusco, en reunión extraordinaria del Comité Regional de Defensa Civil, y la emisión de la Resolución Ejecutiva Regional declarando en emergencia la salud de la población del Cusco. El día 30 de Abril del presente año y el 02 de Mayo se instaló el Equipo de Alerta Respuesta en el principal Terminal Terrestre de Cusco, entre otras actividades destinadas a difundir, socializar y fortalecer la vigilancia epidemiológica de esta nueva Enfermedad a cargo de las diferentes Direcciones y equipos Técnicos de la DIRESA Cusco, bajo la responsabilidad Técnica de la Dirección Ejecutiva de Inteligencia Sanitaria DIRESA – Cusco.

Ante la amenaza de pandemia de la gripe porcina, que en realidad es una mezcla de los virus de la influenza aviar, porcina y humana, que tuvieron que combinarse en algún animal antes de saltar a los humanos, los científicos no saben qué conviene más: tratar de hacer una vacuna específica para este virus, con la probabilidad de que haya mutado para cuando esté lista, o seguir también con la vacuna convencional y usar ambas.

La ciencia se centra en tratar de identificar algún elemento común en los tres virus aviar, porcino y humano- para atacarlos a través de un solo agente. Esto tendría un espectro de protección mucho más amplio si logran aislar el factor común antes de que mute y antes de que llegue la próxima “temporada” gripal.

Hay mucha gente que piensa que se ha hecho demasiado escándalo con el brote de influenza “porcina” -ahora llamada “humana”-, pero ningún funcionario de ningún país está dispuesto a subestimar la gravedad de la pandemia, para que luego cobre muchas víctimas en la nación negligente.

Es cierto que como el internet permite que haya muchísima información, buena parte incomprendible, la alarma se incrementa, pero poca gente podrá criticar la decisión política y de salubridad pública de prevenir en vez de lamentar.

En una ciudad como Cusco, donde buena parte de la población vive hacinada y sin condiciones económicas ni culturales para practicar unas normas de higiene consideradas mínimas en los países ricos, y donde la desnutrición la hace más vulnerable, la vigilancia epidemiológica tiene que ser apoyada por todos los estamentos públicos y privados para evitar que haga estragos un enemigo tan ágil y potencialmente mortal como la influenza “humana”.

## **1.2. FORMULACIÓN DEL PROBLEMA.**

¿ Cuáles son los Factores asociados de riesgo para que una persona muera o sea diagnosticada con virus A (H1N1) mediante el modelo de Regresión Logística, en el Departamento del cusco, 2009?.

## **1.3. FORMULACIÓN DE OBJETIVOS.**

### **1.3.1. Objetivo general.**

Determinar cuáles son los Factores asociados de riesgo para que una persona muera o sea diagnosticada con virus A (H1N1) mediante el modelo de Regresión Logística, en el Departamento del cusco, 2009.

### **1.3.2. Objetivos específicos.**

- Elaborar un Marco Teórico adecuado del Modelo de Regresión Logística.
- Determinar la relación entre dos probabilidades, Razón de Predominio (OR),
- Aplicar el modelo de Regresión Logística a la nueva Influenza por virus A (H1N1).

#### **1.4. HIPOTESIS**

Los factores de riesgo más significativos con el virus A (H1N1) en el departamento del Cusco 2009, son Fiebre, Tos, Rinorrea, Sibilancia, Adenopatía, Cefalea, Mialgias.

#### **1.5. JUSTIFICACION**

El trabajo de investigación que lleva por título “Factores asociados de riesgo para que una persona muera o sea diagnosticada con virus A (H1N1) mediante el modelo de Regresión Logística, en el Departamento del cusco, 2009”. Se justifica por necesidad de brindar un estudio riguroso estadístico, que será de mucha utilidad para profesionales de Ciencias de la Salud, que requieran información estadística, será de utilidad necesaria para la toma de medidas de prevención en el tratamiento de pacientes afectados por virus A (H1N1).

Son estas cosas que me llevan a enfocarme en el estudio de Regresión Logística, y sus múltiples ventajas y aplicaciones, lo cual en este trabajo me avocaré en describir detalladamente los resultados para ver las mejoras en cuanto a los pacientes afectados por una enfermedad, se codificara como 1 la ocurrencia del evento de interés y como 0 la ausencia.

Finalmente, con el presente trabajo se quiere demostrar la validez de la hipótesis como también el logro de los objetivos que para este caso se tomo en cuenta.

#### **1.6. PROCEDIMIENTO METODOLOGICO**

Se trata de una investigación de tipo descriptivo, analítico, y explicativo, en el área de la Estadística.

Primeramente se va a describir el Modelo de Regresión Logística y su respectivo aplicativo en la Nueva Influenza por virus A (H1N1). Finalmente la aplicación se hará con los paquetes estadísticos como SPSS, Microsoft Excel 2007 y otros.

## CAPITULO II

### MARCO TEORICO

#### 2.1. ANTECEDENTES DEL PROBLEMA

En el siglo XX, en el mundo se han producido tres grandes pandemias, la primera, llamada “La Gripe Española” entre 1918 y 1919, durante la Primera Guerra Mundial, esta pandemia se caracterizó por ser de letalidad alta, se estima entre 40 y 100 millones de defunciones a nivel mundial. Al parecer recién en Marzo de 1919 llegó al Perú; sin embargo, no se cuentan con datos precisos sobre la mortalidad y letalidad causada en la población de nuestro país en esa época.

La siguiente pandemia fue la de 1957, que se inició en febrero al norte de China y rápidamente se diseminó a todo el mundo, llegó al Perú en la segunda quincena de Julio de ese mismo año. Los primeros casos se reportan en las ciudades fronterizas con Chile (Tarata, Locumba y Tacna), desde Tacna, Yunguyo y Desaguadero en el Departamento de Puno se diseminó de sur a norte hasta llegar a Lima. Las tasas de ataque observado fluctuaron entre 20 y 40% en la costa, mientras que en la sierra, fluctuaron alrededor de 70%, durante esta pandemia en el Perú, se notificaron 1081 defunciones (Letalidad de 0.8%).<sup>1</sup>

La última pandemia se presentó en 1968, conocida como “La Gripe Asiática”, se extendió rápidamente por todo el mundo, la mortalidad fue muy alta. Desde entonces, la Influenza se comporta como una enfermedad estacional que afecta a la población peruana durante el invierno.

En 1997, (OMS) Organización Mundial de la Salud da la alerta de un nuevo riesgo, al conocerse una de Influenza aviar en Hong Kong, donde el virus de Influenza aviar A (H5N1), cruzó la barrera de las especies infectando humanos, causando un brote con 18 casos y 6 defunciones.

El conocimiento de la existencia del Virus A (H5N1) (gripe aviar) aceleró el proceso para que los países elaboren un plan de preparación contra la pandemia de la Influenza bajo la guía de la OMS, de los cuales el país y particularmente la Región Cusco no puede ser ajena a ella.

---

<sup>1</sup> Plan de Implementación de la Vigilancia Clínico epidemiológica de la Nueva Influenza A H1N1 en el ámbito del Departamento: Cusco, Junio – 2009.

Asimismo el 29 de Abril del 2009, la OMS considerando la situación Epidemiológica reportada de casos presentados en diversos países ha considerado elevar el nivel de alerta correspondiente a la Fase 5, lo que implica que existe una inminencia de la presencia de casos en diversos países del mundo, de lo que por cierto la ciudad del Cusco, podría tener mayor probabilidad de contar o presentar casos de Influenza A (H1N1), dada su condición de ciudad cosmopolita por el turismo.

La técnica de la regresión logística se originó en la década de los 60 con el trabajo de Cornfield, Gordon y Smith<sup>1</sup> en 1967 Walter y Duncan la utilizan ya en la forma que la conocemos actualmente, o sea para estimar la probabilidad de ocurrencia de un proceso en función de ciertas variables. Su uso se incrementa desde principios de los 80 como consecuencia de los adelantos ocurridos en el campo de la computación.

En una búsqueda sistemática realizada en el MEDLINE encontramos que la regresión logística fue citada en 3 892 publicaciones durante el año 1999, en 285 de ellas conjuntamente con la diabetes mellitus (DM).

El modelo de Regresión Logística es un caso especial del modelo lineal generalizado como fue propuesto por Nelder y Wedderburn (1972) y ampliamente discutida en McCullagh y Nelder (1983).

El libro de Mcullagh y Nelder muestra la solución de las ecuaciones de verosimilitud en el modelo lineal generalizado y en regresión logística utilizando el método de Newton Rhapsón, esta solución puede ser obtenida por un método similar al cuadrado medio ponderado para el modelo de regresión ordinario. El método es llamado método iterativo del cuadrado medio ponderado, el cual también puede ser encontrado, por ejemplo en Andersen (1990) o Agreste (1996).

Residuales estandarizados y la distancia de Cook en regresión logística fueron discutidos en Pregibon (1981).

Análisis de regresión logística es también tratado en libros sobre análisis de datos categóricos y en muchos libros sobre análisis de regresión aplicada por ejemplo en Andersen (1990), Agresti(1996) y weisberg (1985).

Recientemente Hosmer y Lemeshow (1989) publicaron un libro especial en análisis de regresión logística. El modelo de regresión logística ha sido utilizado por muchos años ; pero no fue hasta que Truett, Corfield, y Kannel (1967) que aplicaron el modelo de regresión logística utilizando los datos de Framingham, el cual trata de un estudio del corazón, donde se pudo apreciar el poder y la aplicación de estos modelos.

En la universidad San Antonio Abad del Cusco, Edgar Centeno y Walter Chunga (1998) publicaron una tesis: el modelo de regresión logística y algunas aplicaciones. Que estudia el modelo de regresión logística para explicar y predecir fenómenos con variables de respuesta que tienen un comportamiento categórico, además analizan el Método de Máxima Verosimilitud para la estimación de los parámetros desconocidos y su respectiva aplicación en Biología en muestras con variable de respuesta categórica.<sup>2</sup>

## **2.2 . FUNDAMENTOS TEORICOS.**

### **2.2.1. MODELOS LINEALES GENERALIZADOS**

El modelo lineal generalizado (MLG) es una unificación de los modelos de regresión lineal y no lineal, que también permite incorporar distribuciones de respuesta no normales. En un modelo lineal generalizado la distribución de la variable respuesta solo necesita ser un miembro de la familia exponencial, que comprende las distribuciones normal, de Poisson, binomial, exponencial y gamma, entre sus miembros, además, el modelo lineal con error normal no es más que un caso especial del modelo lineal generalizado, por lo que en muchos aspectos se puede considerar que el modelo lineal generalizado es un método unificador de muchos aspectos del modelo y análisis empíricos de datos.

Se iniciara la presentación de esos modelos examinando el caso de la regresión logística, que es una situación en la que la variable de respuesta solo tiene dos resultados posibles, llamados, en forma genérica “éxito” y “fracaso” y se representan por 0 y 1. Obsérvese que la respuesta es cualitativa, en esencia, porque los calificativos “éxito” y “fracaso” son totalmente arbitrarios.

El modelo lineal general surge por la necesidad de expresar en forma Cuantitativa relaciones entre un conjunto de variables, en la que una de ellas se denomina variable respuesta o variable dependiente y las restantes son llamadas covariables, variables explicativas, o variables independientes.

Sea  $Y$  una variable aleatoria cuya función de distribución de probabilidad pertenece a una familia de distribuciones de probabilidades  $H$ , y es explicada por un conjunto

---

<sup>2</sup> Flores Manrique, Luz (Lima 2002). “Análisis estadístico de los factores de riesgo que influyen en la enfermedad Angina de Pecho”.

de variables  $X_1, X_2, \dots, X_n$  las cuales son fijadas antes de conocer Y. La esperanza condicional de Y es dada por:

$$E(Y / X_1, X_2, \dots, X_K) = \beta_0 + \beta_1 X_1 + \dots + \beta_K X_K \quad (1)$$

Luego, si se extrae una muestra aleatoria de tamaño  $n$   $\{(Y_i, X_{i1}, \dots, X_{in}) : i=1, 2, \dots, n\}$ , de una población en la cual la variable respuesta Y, y las variables Independientes  $X_1, X_2, \dots, X_n$  se relacionan linealmente, cada observación de la muestra puede ser expresada como:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \dots + \beta_K X_{iK} + \varepsilon_i \quad i=1, 2, \dots, n \quad (2)$$

En la ecuación (2), el término  $\varepsilon_i$  es una perturbación aleatoria no observable denominada error aleatorio, la cual tiene esperanza cero, varianza  $\sigma^2$  (constante); y dos errores cualesquiera  $\varepsilon_i, \varepsilon_j, \forall i \neq j$  son incorrelacionados entre sí. Utilizando notación matricial, podemos expresar (2) como:

$$Y = X\beta + \varepsilon \quad (3)$$

Donde,  $Y' = (Y_1, Y_2, \dots, Y_n)$  es un vector de variables aleatorias observables, denominado vector respuesta de orden  $n$ ; X es la matriz de variables independientes de orden  $n \times (k+1)$  y  $\beta$  es el vector de parámetros desconocidos de orden  $(k+1)$ . El vector de respuestas Y de la expresión (3) está formado por dos componentes, una sistemática y otra aleatoria. La primera componente constituida por la combinación lineal  $X\beta$ , predictor lineal, el cual es representado como:

$$\eta = X\beta \quad (4)$$

La segunda componente, formada por el vector aleatorio Y, con elementos independientes entre sí, caracterizada por una distribución  $h \in H$  con vector de esperanzas  $\mu$  y matriz de covarianza  $\sigma^2 I$ .

Por otro lado, calculando la esperanza de Y en (3) se tiene que:

$$E(Y) = X\beta = \mu$$



Una característica distintiva del modelo lineal general, es que la variable respuesta Y está medida en escala numérica, mientras que las covariables pueden ser numéricas o categóricas y además son independientes entre sí.

### **2.2.2. MODELO DE REGRESION CON RESPUESTA DICOTOMICA**

La regresión logística resulta útil para los casos en los que se desea predecir la presencia o ausencia de una característica o resultado según los valores de un conjunto de variables predictoras. Es similar a un modelo de regresión lineal pero está adaptado para modelos en las que la variable dependiente es dicotómica. Los coeficientes de regresión logística pueden utilizarse para estimar la razón de las ventajas (odds ratio) de cada variable independiente del modelo. La regresión logística se puede aplicar a un rango más amplio de situaciones de investigación que el análisis discriminante.

Para cada análisis, en el SPSS incluye en los casos: totales, seleccionados, validos. Para cada variable categórica: codificación de los parámetros. Para cada paso: variables introducidas o eliminadas, historia de iteraciones -2log de la verosimilitud, bondad de ajuste de Hosmer-Lemeshow, ji-cuadrado del modelo, ji-cuadrado de la mejora, tabla de clasificación correlaciones entre las variables. <sup>3</sup>

### **2.2.3. INTRODUCCIÓN A LA REGRESION LOGISTICA.**

El uso de procedimientos de regresión está extendido en Ciencias Económicas, Sociales, Experimentales y de la Salud. Con el método de regresión logística se pretende llegar a un modelo que explique una característica binaria (un “sí” o “no”, “éxito” frente a “fracaso”, “enfermo” frente a “sano”, etc.) y su probabilidad a partir de cierto número de variables observadas. Por ejemplo, un estudio epidemiológico puede recabar información sobre un número grande de sujetos, su sexo, edad, si fuma o no, cuánto ejercicio hace, ocupación, lugar de residencia, estado civil, etc. además de si han padecido o no infarto.

La regresión logística ayudaría a descartar cuáles de estas variables son realmente “explicativas” de la probabilidad de ser diagnosticado con el virus A (H1N1) y, para las que son explicativas, produciría un modelo matemático ajustado a nuestras observaciones que podría hacer predicciones. A veces el primer aspecto, es poder

---

<sup>3</sup> (Montgomery – Peck – Vining- 2004) “Introducción al Análisis de Regresión Lineal”.

descartar el efecto de una variable, es más importante incluso que el segundo. En Ciencias de la Salud, particularmente en Epidemiología, se plantean usualmente problemas de alta complejidad: de cada individuo son recogidas múltiples variables, las muestras son muy grandes. Para realizar el estudio estadístico es necesario el uso de bases de datos y paquetes estadísticos. La teoría, sin embargo, es relativamente sencilla y descansa sobre un principio universal: “el modelo que mejor ajusta las observaciones es aquél que hace más probable la muestra obtenida de entre todos los modelos posibles”. Este es el principio de “máxima verosimilitud”, y los métodos de ajuste a modelos logísticos, lineales, de Poisson u otros se basan en este principio.

## 2.3. REGRESION LOGISTICA

### 2.3.1. DEFINICIÓN

Sea  $Y$  una variable dependiente binaria que toma dos valores posibles etiquetados como 0 y 1.

Sean  $X_1, X_2, \dots, X_k$  un conjunto de variables independientes observadas con el fin de explicar y/o predecir el valor de  $Y$ .

El objetivo es determinar  $P[Y = 1 / X_1, X_2, \dots, X_k]$ , donde  $P$  indica probabilidad.

Por lo tanto:  $P[Y = 0 / X_1, X_2, \dots, X_k] = 1 - P[Y = 1 / X_1, X_2, \dots, X_k]$

Se construye un modelo de la forma:

$$P[Y = 1 / X_1, X_2, \dots, X_k] = P(X_1, X_2, \dots, X_k; \beta) \quad (5)$$

donde  $P(X_1, X_2, \dots, X_k; \beta) : R^k \times P[0,1]$  es una función que recibe el nombre de función de enlace (función de probabilidad) cuyo valor depende de un vector de parámetros  $\beta = (\beta_1, \dots, \beta_k)^T$ .

### 2.3.2. FUNCIÓN DE VEROSIMILITUD

Con el fin de estimar  $\beta$  y analizar el comportamiento del modelo considerado, observamos una muestra aleatoria simple de tamaño  $n$  dada por  $\{(x_i^T, y_i) ; i = 1, \dots, n\}$  donde  $x_i^T = (x_{i1}, \dots, x_{ik})$  es el valor de las variables

independientes e  $y_i = \{0,1\}$  es el valor observado de Y en el i-ésimo elemento de la muestra. <sup>4</sup>

$$Y / (x_1, x_2, \dots, x_k) \sim \text{Binomial} (1, P(Y = 1 / x_1, x_2, \dots, x_k; \beta))$$

Utilizando el hecho de que la variable dependiente toma sólo dos resultados (éxito y fracaso), cuando el número de éxitos en n repeticiones tiene una distribución binomial B(n,p).

La función de verosimilitud es:

$$L(\beta / (x_1^T, y_1), \dots, (x_n^T, y_n)) = \prod_{i=1}^n p_i^{y_i} (1 - p_i)^{1-y_i} \quad (6)$$

Donde

$$p_i = p(x_i^T, \beta) = p(x_{i1}, \dots, x_{ik}, \beta) ; i=1, \dots, n$$

### 2.3.3. MODELO DE REGRESIÓN LOGÍSTICA BINARIA

Sea

$$p(x_1, \dots, x_k; \beta) = G(\beta_0 x_0 + \dots + \beta_k x_k) \quad (7)$$

Donde  $G(x) = \frac{e^x}{1 + e^x}$  es la función de densidades acumuladas de la función

logística, el modelo normalmente conocido es:

$$\log \left( \frac{P(x_1, \dots, x_k; \beta)}{1 - P(x_1, \dots, x_k; \beta)} \right) = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k \quad (8)$$

Llamado modelo logit. Cuando la variable cualitativa toma el valor 1 en la expresión:

$$\frac{P[Y = 1 / X_1, \dots, X_k]}{P[Y = 0 / X_1, \dots, X_k]} = \frac{P(x_1, \dots, x_k; \beta)}{1 - P(x_1, \dots, x_k; \beta)} \quad (9)$$

se conoce con el nombre de factor de riesgo en el mundo de la medicina, donde la variable Y indica habitualmente la presencia de una determinada enfermedad, objeto de estudio y en ausencia toma el valor 0.

<sup>4</sup>Flores Manrique, Luz (Lima 2002). "Análisis estadístico de los factores de riesgo que influyen en la enfermedad Angina de Pecho".

### 2.3.4. FUNCIÓN DE VEROSIMILITUD

Teniendo en cuenta la forma matricial de:

$$P(X_1, \dots, X_k; \beta) = \frac{e^{x_i^T \beta}}{1 - e^{x_i^T \beta}} \quad (10)$$

Según (6) la función de verosimilitud viene dada por:

$$L(\beta / (x_1^T, y_1), \dots, (x_n^T, y_n)) = \prod_{i=1}^n \left[ \frac{e^{x_i^T \beta}}{1 + e^{x_i^T \beta}} \right]^{y_i} \left[ \frac{e^{x_i^T \beta}}{1 + e^{x_i^T \beta}} \right]^{1-y_i}$$

### 2.3.5. ESTIMACIÓN DE LOS PARÁMETROS

El vector de parámetros  $\beta$  se estima mediante el método de máxima verosimilitud que consiste en elegir el valor de  $\hat{\beta}$ , como estimador para  $(\beta)$  para el cual  $L(\beta)$  es máximo, se toma logaritmo a la función en (5) del siguiente modo:

$$l(\beta) = \log L(\beta) = \sum_{i=1}^n \log P(X_1, \dots, X_k; \beta) \quad (11)$$

se resuelve mediante la ecuación de verosimilitud

$$\frac{\partial l(\beta)}{\partial \beta} = \sum_{i=1}^n X_i (y_i - p_i) = 0 \quad (12)$$

Donde

$$p_i = p(X_i; \beta); i = 1, \dots, n$$

Mediante los métodos iterativos. Este método consiste en maximizar la función de verosimilitud de la muestra en función del parámetro  $\hat{\beta}$ . Este procedimiento es matemáticamente complejo a través del cálculo diferencial, pero lo que importa para el usuario es:

1º El proceso es iterativo, es decir se dan a los coeficientes unos valores arbitrarios (habitualmente, aunque no necesariamente, el valor 0). Algunos paquetes estadísticos (por ejemplo el PRESTA) preguntan por estos valores, otros (como el SPSS o el SAS) no y asumen 0. La solución final no depende de estos valores, pero sí el tiempo de cálculo y a veces puede ser necesario "jugar" con ellos.

2° A partir de estos valores iniciales y de los valores de la(s) variable(s) independiente(s) se calculan las matrices de varianzas y covarianzas.

3° Y a partir de la inversa de la matriz se calculan los nuevos estimadores, se comprueba si son la solución final se debe parar el proceso y en caso contrario se repite el proceso.<sup>5</sup>

### 2.3.6. MODELO DE REGRESIÓN LOGÍSTICA SIMPLE

Este modelo tiene la forma

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad \text{para } i = 1, 2, \dots, n \quad (13)$$

De esto se deduce que:

$$\text{Si } y = 1, \varepsilon_i = 1 - \beta_0 - \beta_1 x_i \quad (14)$$

$$\text{Si } y = 0, \varepsilon_i = -\beta_0 - \beta_1 x_i \quad (15)$$

Por lo tanto  $\varepsilon_i$ , no puede tener distribución normal debido a que toma valores discretos, el Modelo de Regresión Lineal Simple, no es aplicable para el caso de variable respuesta de tipo dicotómico.

En el Análisis de Regresión Lineal Simple, el punto inicial del proceso de estimación del modelo es un gráfico de dispersión de la variable respuesta versus la regresora, pero este gráfico resulta limitado cuando solo hay dos valores posibles para la variable respuesta, por tanto se debe usar otros gráficos esto resulta de la suavización de los valores de la variable respuesta con las variables regresoras.

$$\text{Sea } \pi(x) = E(y/x) \quad (16)$$

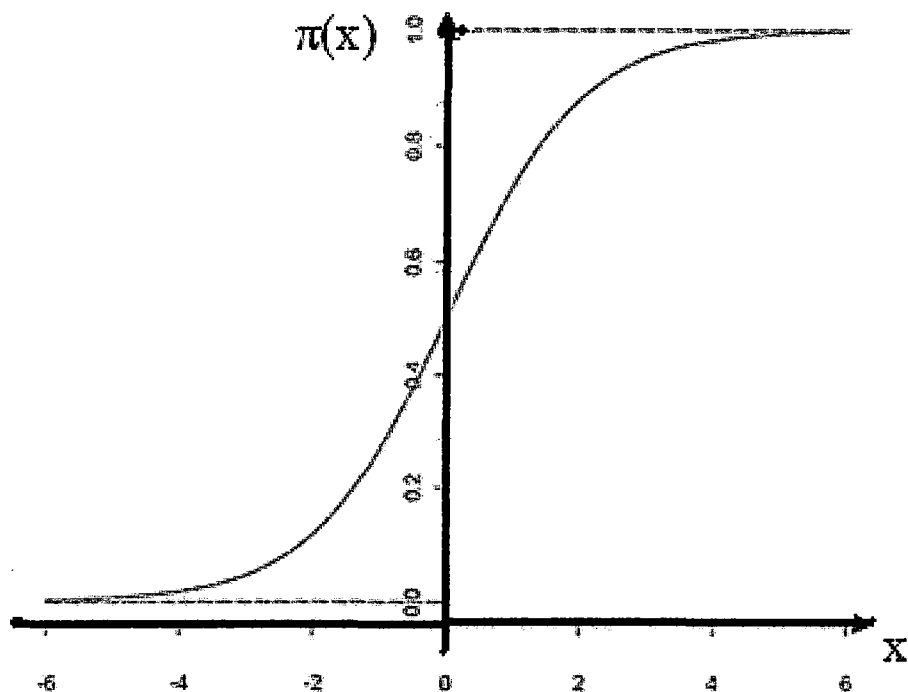
Que representa la media condicional de  $y = 1$  dado  $x$ , donde  $\pi(x)$  representa la probabilidad de que ocurra  $y = 1$ , ciertamente no se espera que (16) tenga relación lineal dentro del rango de la variable regresora.

¿Qué hay de relación entre  $\pi(x)$  y  $x$  para valores intermedios de  $x$ ? se espera una relación curvilínea. Para cualquier valor grande de  $x$ ,  $\pi(x)$  tomara valores cercanos a uno. El grafico que muestra el comportamiento de  $\pi(x)$  con  $x$  es:

---

<sup>5</sup> Flores Manrique, Luz (Lima 2002). "Análisis estadístico de los factores de riesgo que influyen en la enfermedad Angina de Pecho".

Figura N° 1: Función Logística



Curva en forma de S o sigmoide que tiene las propiedades requeridas para  $\pi(x)$  y que tiene las propiedades de una función de distribución de probabilidad acumulada, para esta probabilidad se usa la función de distribución acumulada de la distribución logística dado por: <sup>6</sup>

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} \quad (17)$$

(17) tiene un gráfico similar a la Figura N° 1 cuando  $\beta_0 < 0$  y  $\beta_1 > 0$ , además este modelo toma valores en el intervalo 0,1 .

Cuando,  $P y = 1 = 0.5$  el valor de x es:  $\frac{\beta_0}{\beta_1}$ , que brinda información muy útil.

Una transformación de  $\pi(x)$  que es lo central del estudio de la Regresión Logística es la transformación logit, transformación que se define en términos de  $\pi(x)$  como:

$$g x = \ln \frac{\pi(x)}{1 - \pi(x)} \quad (18)$$

<sup>6</sup> Salcedo Poma, Celia Mercedes (Lima 2002). "Estimación de ocurrencia de incidencias en declaraciones de pólizas de importación".

$$g(x) = \beta_0 + \beta_1 x$$

Lo importante de esta transformación es que tiene muchas propiedades semejantes al Modelo de Regresión Lineal simple, por ejemplo es lineal en sus parámetros, puede ser continua y puede tomar cualquier valor real dependiendo de  $x$ .

Para el Modelo de Regresión Lineal simple, la variable respuesta se expresa como:

$$y = E(y/x) + \varepsilon \quad (19)$$

Para la variable respuesta dicotómica lo expresamos como:

$$y = \pi(x) + \varepsilon \quad (20)$$

Veamos que ocurre con este modelo:

Si  $y = 1$ ,  $\varepsilon_i = 1 - \pi(x)$  y tiene probabilidad  $\pi(x)$

Si  $y = 0$ ,  $\varepsilon_i = -\pi(x)$  y tiene probabilidad  $1 - \pi(x)$

Entonces  $\varepsilon_i$  tiene distribución binomial con media cero y varianza  $\pi(x)[1 - \pi(x)]$  Por tanto la distribución condicional de la variable respuesta tiene distribución de probabilidad con media  $\pi(x)$ .

El lado izquierdo de (18) se llama también logaritmo de ODDS RATIO o razón de probabilidad de  $y = 1$  con  $y = 0$ , específicamente:

$$\text{ODDS RATIO} = \frac{\pi(x)}{1 - \pi(x)} \quad (21)$$

O también llamado razón de ventaja a favor de éxito.

## 2.4. REGRESION LOGISTICA MULTIPLE

En esta sección se generaliza el modelo de Regresión Logística Simple tratado en la sección anterior, es decir consideremos más de una variable regresora, en donde por lo menos una es de tipo cuantitativo.

### 2.4.1. MODELO DE REGRESION LOGISTICA MULTIPLE

Sea el vector de variables regresoras  $\vec{x}^T = (x_1, x_2, \dots, x_k)$  por el momento asumiremos que están medidas por lo menos bajo una escala intervalar. Sea la probabilidad condicional para que la variable respuesta sea igual a 1, denotado por:

$$p(y = 1 / \vec{x}) = \pi(\vec{x}) \quad (22)$$

El algoritmo del Modelo de Regresión Múltiple esta dado por:

$$g(\vec{x}_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}, \text{ para } i=1,2,\dots,n \quad (23)$$

Entonces el modelo de Regresión Logística Múltiple es:

$$\pi(\vec{x}) = \frac{e^{g(\vec{x})}}{1 + e^{g(\vec{x})}} \quad (24)$$

Al igual que en el caso de Regresión Lineal Múltiple si es necesario usar variables regresoras categóricas, por ejemplo si una tiene c niveles será necesario incorporar c-1 variables ficticias o llamadas dummy., así entonces el logit para un modelo con k variables regresoras y una variable categorica, por ejemplo la j-esima es:

$$g(\vec{x}) = \beta_0 + \beta_1 x_{i1} + \dots + \sum_{l=1}^{c-1} \beta_l D_{jl} + \beta_k x_{ik} \quad \text{para } i = 1,2,\dots,n$$

#### **2.4.2. AJUSTE DEL MODELO DE REGRESION LOGISTICA MULTIPLE**

El ajuste se efectúa a través del uso de los métodos de máxima verosimilitud, los mismos que se encuentran en los softwares estadísticos que permiten analizar datos mediante este método.

Asumimos que disponemos de una muestra n observaciones independientes  $(x_i, y_i)$ ,  $i=1,2,\dots, n$ ; donde  $y_i$  toma valores 0 ó 1, para estimar  $\vec{\beta}^T = (\beta_0, \beta_1, \dots, \beta_k)$  que es el vector de parámetros desconocidos.

Para el Modelo de Regresión Lineal Múltiple se usa el método de Mínimos Cuadrados para estimar  $\vec{\beta}^T$ , el cual minimiza la suma de cuadrados del error, pero cuando la variable respuesta es binaria aplicar este método no provee las mismas propiedades cuando es usado en variables respuestas continuas.

Por ello se usara el método de Máxima Verosimilitud, ya que obtendremos parámetros estimados que maximizan la probabilidad de obtener un conjunto de datos observados.

La función de verosimilitud expresa la probabilidad de los datos observados como una función de parámetros desconocidos. Los Estimadores de Máxima Verosimilitud de esos parámetros son aquellos que están en concordancia con los datos observados.



Consideremos el Método de Regresión Lineal Múltiple con mayor detalle, supongamos que se dispone de n objetos u observaciones donde para cada uno de ellos existe una respuesta que puede ser:

$$y_i = 0 \text{ o } y_i = 1$$

Sea  $\bar{y}^T = (y_1, y_2, \dots, y_n)$  donde  $y_i \sim B(1, \pi_i)$  y sea  $\bar{x}^T_i = (1, x_{i1}, \dots, x_{ik})$  la i-ésima observación para las k variables explicativas.

Así el Modelo de Regresión Logística está dada por la expresión:

$$g(\bar{x}_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}, \text{ para } i=1,2,\dots, n$$

$$P[y_i = 1 / \bar{x}_i] = \pi(\bar{x}_i) = \frac{e^{g(\bar{x}_i)}}{1 + e^{g(\bar{x}_i)}} \quad (25)$$

Equivalentemente

$$P[y_i = 1 / \bar{x}_i] = \frac{\text{Exp}(\beta_0 + \sum_{j=1}^k \beta_j x_{ij})}{1 + \text{Exp}(\beta_0 + \sum_{j=1}^k \beta_j x_{ij})} \quad (26)$$

y la probabilidad de que  $y_i$  es igual a cero es:

$$P[y_i = 0 / \bar{x}_i] = 1 - P[y_i = 1 / \bar{x}_i], \text{ y :}$$

$$P[y_i = 0 / \bar{x}_i] = \frac{1}{1 + \text{Exp}(\beta_0 + \sum_{j=1}^k \beta_j x_{ij})} \quad (27)$$

Para facilitar la notación usaremos la variable indicadora  $x_{0i} = 1, i=1,2,\dots, n$ .

Entonces (26) y (27) son respectivamente:

$$P[y_i = 1 / \bar{x}_i] = \pi(\bar{x}_i) = \frac{e^{\bar{\beta}\bar{x}_i^T}}{1 + e^{\bar{\beta}\bar{x}_i^T}} \quad (28)$$

$$P[y_i = 0 / \bar{x}_i] = 1 - \pi(\bar{x}_i) = \frac{1}{1 + e^{\bar{\beta}\bar{x}_i^T}} \quad (29)$$

Donde:  $\bar{x}_i = (x_{i0}, x_{i1}, \dots, x_{ik})$ , es el vector que contiene los valores de las variables explicativas  $\bar{\beta}^T = (\beta_0, \beta_1, \dots, \beta_k)$  es el vector de parámetros a ser estimado.

El i-ésimo logito es:

$$\lambda_i = \text{Ln} \left( \frac{\pi_i}{1 - \pi_i} \right) = \sum_{j=0}^k \beta_j x_{ij} \quad (30)$$

Como vemos, (30) es una función lineal simple del vector de observaciones  $\bar{x}_i$ , llamada transformación logística de la probabilidad  $\pi_i$ , o simplemente Logit o Logito de la ecuación, a la expresión (30) también se le llama Modelo Logístico Lineal.

A fin de obtener la estimación máximo verosímil para el vector  $\bar{\beta}$ , escribimos la función de densidad de probabilidad del vector  $\bar{y}$  el cual es proporcional a n funciones  $\beta(1, \pi_i)$ , esto es:

$$\begin{aligned}
 f(y_i; \pi_i) &= \prod_{i=1}^n \pi_i^{y_i} (1 - \pi_i)^{1-y_i} \\
 &= \prod_{i=1}^n \left( \frac{\pi_i}{1 - \pi_i} \right)^{y_i} (1 - \pi_i) \\
 &= \left\{ \prod_{i=1}^n (1 - \pi_i) \right\} \left\{ \prod_{i=1}^n \text{Exp} \left[ \text{Ln} \left( \frac{\pi_i}{1 - \pi_i} \right)^{y_i} \right] \right\} \\
 &= \left\{ \prod_{i=1}^n (1 - \pi_i) \right\} \text{Exp} \left[ \sum_{i=1}^n y_i \text{Ln} \left( \frac{\pi_i}{1 - \pi_i} \right) \right]
 \end{aligned} \tag{31}$$

Reemplazando (30) en (31), se obtiene:

$$\begin{aligned}
 f(y_i; \pi_i) &= \left\{ \prod_{i=1}^n (1 - \pi_i) \right\} \text{Exp} \left\{ \sum_{i=1}^n y_i \sum_{j=0}^k \beta_j x_{ij} \right\} \\
 &= \left\{ \prod_{i=1}^n (1 - \pi_i) \right\} \text{Exp} \left\{ \sum_{j=0}^k \left( \sum_{i=1}^n y_i x_{ij} \right) \beta_j \right\}
 \end{aligned} \tag{32}$$

El logaritmo natural de la función (32), llamado función soporte es:

$$l(\pi_i; y_i) = \sum_{j=0}^k \left( \sum_{i=1}^n y_i x_{ij} \right) \beta_j + \sum_{i=1}^n \ln(1 - \pi_i) \tag{33}$$

Pero (29):  $1 - \pi_i = [1 + \text{Exp}(\bar{\beta}^T \bar{x}_i)]^{-1}$ , luego

$$\text{Ln}(1 - \pi_i) = -\text{Ln}[1 + \text{Exp}(\bar{\beta}^T \bar{x}_i)]$$

$$\text{Ln}(1 - \pi_i) = -\text{Ln} \left[ 1 + \text{Exp} \left( \sum_{j=0}^k \beta_j x_{ij} \right) \right] \tag{34}$$

Reemplazando (34) en (33), se obtiene:

$$l(\pi_i; y_i) = \sum_{j=0}^k \left( \sum_{i=1}^n y_i x_{ij} \right) \beta_j - \sum_{i=1}^n \text{Ln} \left[ 1 + \text{Exp} \sum_{j=0}^k \beta_j x_{ij} \right] \tag{35}$$

Como vemos (35) es una función que ya no depende de  $\pi_i$ , sino de  $\beta_j$ , solamente, entonces lo denotamos como:

$$L(\vec{\beta}) = \sum_{j=0}^k \left( \sum_{i=1}^n y_i x_{ij} \right) \beta_j - \sum_{i=1}^n \text{Ln} \left[ 1 + \text{Exp} \left( \sum_{j=0}^k \beta_j x_{ij} \right) \right] \quad (36)$$

Es una función que depende exclusivamente del vector  $\vec{\beta}$ .

Definamos como:

$$t_j = \sum_{i=1}^n y_i x_{ij} \quad (37)$$

Entonces reemplazando (37) en (36) se tiene:

$$L(\vec{\beta}) = \sum_{j=0}^k \beta_j t_j - \sum_{i=1}^n \text{Ln} \left[ 1 + \text{Exp} \left( \sum_{j=0}^k \beta_j x_{ij} \right) \right] \quad (38)$$

Como (38) es una función exclusivamente del vector de parámetros  $\vec{\beta}$ , por el Teorema de Factorización, se tiene que  $t_j$ , para  $j=0,1,\dots,k$  son estadísticos suficientes para los parámetros  $\beta_j$ , para  $j=0,1,\dots,k$ .

La variable aleatoria  $t_j$ , dada en la expresión (38) es la suma de algunos de los términos de la matriz de diseño X, es decir se incluyen en la suma solamente los elementos que corresponden a una respuesta de tipo  $y=1$ .

Las ecuaciones de verosimilitud, se obtienen derivando (36) con respecto a los elementos de  $\vec{\beta}$  e igualando a cero:

$$\frac{\partial L}{\partial \beta_j} = \sum_{i=1}^n y_i x_{ij} - \sum_{i=1}^n x_{ij} \frac{\left[ \text{Exp} \left( \sum_{j=0}^k \beta_j x_{ij} \right) \right]}{\left[ 1 + \text{Exp} \left( \sum_{j=0}^k \beta_j x_{ij} \right) \right]} \quad (39)$$

Las ecuaciones de verosimilitud de (39) son:

$$\sum_{i=1}^n y_i x_{ij} - \sum_{i=1}^n x_{ij} \pi_i = 0 \quad j=0,1,2,\dots,k \quad (40)$$

Siendo  $x_{i0} = 1$ , equivalentemente a (40) es:

$$\sum_{i=1}^n x_{ij} (y_i - \pi_i) = 0 \quad j=0,1,2,\dots,k \quad (41)$$

Donde:

$$\pi_i = \frac{\text{Exp} \left( \sum_{j=0}^k \hat{\beta}_j x_{ij} \right)}{1 + \text{Exp} \left( \sum_{j=0}^k \hat{\beta}_j x_{ij} \right)}, \text{ para } i=1,2,\dots,n$$

Es el estimador máximo verosímil de  $\pi_i$  y se obtiene mediante  $\hat{\beta}_j$  y el vector  $\bar{x}_i$ .

La expresión (40) en su forma matricial es:

$$X^T (\bar{y} - \bar{\pi}) = X\bar{S} = \bar{0} \quad (42)$$

Estas ecuaciones son parecidas a las ecuaciones normales obtenidas para estimar el Modelo de Regresión Lineal Múltiple, pero no son lineales en  $\bar{\beta}$  lo cual hace que se use un método iterativo para determinar los valores del vector  $\bar{\beta}$ .

La obtención de  $\hat{\beta}_j$  mediante métodos iterativos; para  $j = 0,1,\dots, k$  se tratara en la siguiente sección, ahora obtendremos la varianza y covarianza de  $\bar{\beta}$ .

Sea  $X_{(n \times p)}$  la matriz de diseño, con  $p=k+1$ , con elementos.

$$X = \begin{bmatrix} 1 & x_{11} & \dots & x_{1k} \\ 1 & x_{21} & \dots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{nk} \end{bmatrix}$$

Las ecuaciones de verosimilitud en su forma matricial, de la expresión (42) es:

$$X^T \bar{y} = X^T \bar{\pi} \quad (43)$$

Donde  $\bar{\pi}^T = (\pi_1, \pi_2, \dots, \pi_n)$

$$\sum_{i=1}^n x_{ij} (y_i - \pi_i) = 0 \quad (44)$$

El método de estimación de las varianzas y covarianzas lo obtendremos de la matriz de la segunda derivada parcial de (44) y tiene la forma:

$$\frac{\partial^2 L}{\partial \beta_j^2} = - \sum_{i=1}^n x_{ij}^2 \pi_i (1 - \pi_i); \quad j=0,1,2,\dots,k \quad (45)$$

Reemplazando (45) para  $\pi_i$  se tiene:

$$\frac{\partial^2 L}{\partial \beta_j^2} = - \sum_{i=1}^n \frac{x_{ij}^2 \text{Exp} \left( \sum_{j=0}^k \beta_j x_{ij} \right)}{\left[ 1 + \text{Exp} \left( \sum_{j=0}^k \beta_j x_{ij} \right) \right]^2} \quad (46)$$

Para  $j=0,1,\dots, k$

$$\frac{\partial^2 L}{\partial \beta_j \partial \beta_l} = - \sum_{i=1}^n x_{ij} x_{il} \pi_i (1 - \pi_i) \quad (47)$$

Para  $j=0,1,2,\dots,k$  y  $l=0,1,2,\dots,k$

Reemplazando en (47) tenemos:

$$\frac{\partial^2 L}{\partial \beta_j \partial \beta_l} = - \sum_{i=1}^n x_{ij} x_{il} \frac{\text{Exp} \left( \sum_{j=0}^k \beta_j x_{ij} \right)}{\left[ 1 + \text{Exp} \left( \sum_{j=0}^k \beta_j x_{ij} \right) \right]^2} \quad (48)$$

Tanto (47) como (48) no son funciones de  $y_i$ , entonces la matriz de observación y la matriz de la segunda derivada esperada son idénticas.

Ahora bien la matriz que contiene el negativo de las ecuaciones (46) y (48) se denota como  $I(\vec{\beta})$ , llamada matriz de información; las varianzas y covarianzas de  $\hat{\beta}_j$  se obtienen tomando la inversa de esta matriz, esto es:

$$\text{Cov}(\vec{\beta}) = I^{-1}(\vec{\beta}) \quad (49)$$

Ver Cordeiro Gauss M.(1992).

Los estimadores de la varianza y covarianza, denotado por  $\hat{Cov}(\hat{\beta})$ , se obtiene evaluando  $\text{Cov}(\vec{\beta})$  en  $\hat{\beta}$ .

Entonces la matriz de información estimada, matricialmente tiene la forma:

$$\hat{I}(\hat{\beta}) = X^T V X \quad (50)$$

V es una matriz diagonal, de la forma:

$$V = \text{Diag} [\pi_i, (1 - \pi_i)]$$

de tamaño  $n \times n$ , además (50) es:

$$\hat{Cov}(\hat{\beta}) = (X^T V X)^{-1} \quad (51)$$

y es de tamaño  $(k+1)(k+1)$  escribiremos los elementos de la matriz (51).

$$\hat{Cov}(\hat{\beta}) = \begin{bmatrix} \sigma^2(\hat{\beta}_0) & \sigma(\hat{\beta}_0, \hat{\beta}_1) & \dots & \sigma(\hat{\beta}_0, \hat{\beta}_k) \\ \cdot & \sigma(\hat{\beta}_1) & \dots & \sigma(\hat{\beta}_1, \hat{\beta}_k) \\ \vdots & \vdots & \ddots & \vdots \\ \cdot & \cdot & \dots & \sigma^2(\hat{\beta}_k) \end{bmatrix}$$

Donde:

$\sigma^2(\hat{\beta}_j)$  es la varianza estimada de  $\hat{\beta}_j$ ,

$\sigma(\hat{\beta}_j, \hat{\beta}_i)$  es la covarianza estimada de  $\hat{\beta}_j$  y  $\hat{\beta}_i$ ,

$\sigma(\hat{\beta}_j)$  es el error estándar de  $\hat{\beta}_j$ ,

La matriz (51) será muy útil cuando se discuta el ajuste y la evaluación del Modelo de Regresión Logística.<sup>7</sup>

### **2.4.3. METODO DE NEWTON – RAPHSON PARA ESTIMAR LOS PARAMETROS DEL MODELO DE REGRESION LOGISTICA.**

Este es un método para resolver ecuaciones no lineales, como las obtenidas en (39) o equivalentemente en (40), y requieren una solución mediante métodos iterativos para hallar la estimación de los parámetros que es el máximo de la función (36).

Uno de los métodos más usados para resolver ecuaciones de este tipo, es el de Newton-Raphson, porque converge rápidamente. En la figura N° 2 se ilustra el método.

Tomado como estimación inicial  $x_1$ , se prolonga la tangente a la curva en este punto hasta interceptar con el eje de las abscisas al cual llamaremos punto  $x_2$ , entonces se toma  $x_2$  como la siguiente aproximación.

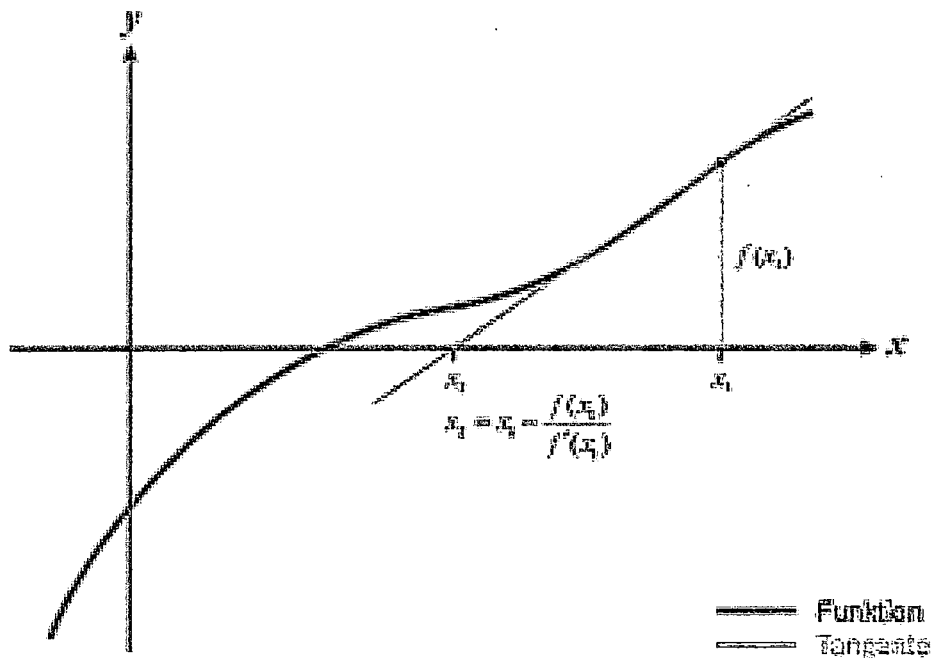
Este proceso continua hasta un valor de  $x_2$  haga que la función sea nula o suficientemente cercana a cero.

Para la estimación del vector  $\vec{\beta}$  se quiere hallar el máximo de una función; entonces usa la primera derivada, porque se anula en el punto máximo y la segunda derivada para calcular las tangentes. En nuestro caso es necesario hallar la segunda derivada para obtener la matriz de varianzas y covarianzas de los parámetros estimados.

---

<sup>7</sup> Salcedo Poma, Celia Mercedes (Lima 2002). "Estimación de ocurrencia de incidencias en declaraciones de pólizas de importación".

**FIGURA N° 2 Interpretación Geométrica del Método Iterativo de Newton Raphson.**



Entonces se usa el siguiente esquema iterativo:

$$\tilde{\beta}(t+1) = \tilde{\beta}(t) + [I(\tilde{\beta}(t))]^{-1} S\tilde{\beta}(t) \quad (52)$$

donde:  $S(\tilde{\beta})$  y  $I(\tilde{\beta})$  son las funciones de Score y de Información respectivamente.

La función de Score es un vector de tamaño  $k+1$ , donde el  $j$ -ésimo elemento de acuerdo a (39) es:

$$\frac{\partial L}{\partial \beta_j} = \sum_{i=1}^n (y_i - \pi_i^{(j)}) x_{ij} \quad (53)$$

La cual es similar a la expresión (46):

$$\sum_i x_{ij} (y_i - \pi_i) = 0 \quad j = 1, 2, \dots, k$$

La función de información es una matriz de tamaño  $(k+1)(k+1)$  donde el  $i,j$ -ésimo elemento  $(i,j)$  es:

$$\frac{\partial^2 L}{\partial \beta_j \partial \beta_i} = - \frac{\partial}{\partial \beta_i} \left[ \sum_{i=1}^n x_{ij} (y_i - \pi_i) \right]$$

$$\begin{aligned}
&= -\frac{\partial}{\partial \beta_i} \left[ \sum_{i=1}^n x_{ij} y_i - \sum_{i=1}^n x_{ij} \frac{e^{\tilde{\beta}^T \tilde{x}_i}}{1 + e^{\tilde{\beta}^T \tilde{x}_i}} \right] \\
&= \sum_{i=1}^n x_{ij} \left[ \frac{e^{\tilde{\beta}^T \tilde{x}_i} x_{ij} (1 + e^{\tilde{\beta}^T \tilde{x}_i}) - e^{\tilde{\beta}^T \tilde{x}_i} x_{ij} e^{\tilde{\beta}^T \tilde{x}_i}}{(1 + e^{\tilde{\beta}^T \tilde{x}_i})^2} \right] \\
&= \sum_{i=1}^n \frac{x_{ij} x_{ij} e^{\tilde{\beta}^T \tilde{x}_i}}{(1 + e^{\tilde{\beta}^T \tilde{x}_i})} \\
&= \sum_{i=1}^n x_{ij} x_{ij} \pi_i (1 - \pi_i) \quad j=0,1,\dots,k; i=0,1,\dots,k \quad (54)
\end{aligned}$$

donde  $\tilde{\pi}^{(t)}$ , es la t-ésima aproximación para  $\tilde{\pi}$ , obtenida de  $\tilde{\beta}^{(t)}$  mediante:

$$\pi_i^{(t)} = \frac{\text{Exp} \left( \sum_{j=0}^k \beta_j^{(t)} x_{ij} \right)}{\left[ 1 + \text{Exp} \left( \sum_{j=0}^k \beta_j^{(t)} x_{ij} \right) \right]} \quad (55)$$

entonces el próximo valor reemplazado en (52) es:

$$\tilde{\beta}^{(t+1)} = \tilde{\beta}^{(t)} - \{X^T V^{(t)} X\}^{-1} X^T (\tilde{y} - \tilde{\pi}^{(t)}) \quad (56)$$

La expresión (52) se usa para obtener  $\tilde{\pi}^{(t+1)}$  y así sucesivamente.

Después de dar un valor inicial  $\tilde{\beta}^{(0)}$ , se usa (52) para obtener  $\tilde{\pi}^{(0)}$  y para  $t > 0$  las iteraciones siguientes se efectuara usando (52) y (53).

En el límite,  $\tilde{\pi}^{(t)}$  y  $\tilde{\beta}^{(t)}$  converge a los EMV de  $\tilde{\pi}$  y converge en general en 5 o 6 iteraciones.

Existen software estadístico como el SAS y el SPSS con programas para estimar una regresión logística usando el método descrito. Una ventaja de este método es que en el paso final del proceso iterativo se obtienen la inversa de la función de información, que es asintóticamente la matriz de varianzas y covarianzas del vector  $\tilde{\beta}$  y permite efectuar inferencias sobre los parámetros basado en la teoría normal.



#### 2.4.4. INTERPRETACION DE LOS COEFICIENTES DEL MODELO ESTIMADO

Recordamos del modelo de regresión múltiple que el valor de un coeficiente significaba el cambio en unidades de la variable dependiente por cada unidad de la variable independiente a que se refiere el coeficiente, permaneciendo invariantes los valores del resto de variables independientes del modelo.

A nivel de coeficientes estimados exponencialmente la interpretación es muy similar y la diferencia estriba en que en este caso no se trata del cambio (incremento o disminución) de la probabilidad de la variable dependiente por cada unidad de cambio en las independientes, sino del incremento o disminución que se produce en el coeficiente entre  $P(Y=1) / P(Y=0)$ , expresado por :

$$\frac{P(Y = 1)}{P(Y = 0)} = e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k} \quad (57)$$

Más aun, están expresados en logaritmos, por lo que sería necesario transformarlos (tomando los valores del antilogaritmo) de tal forma que se evalúe mas fácilmente su efecto sobre la probabilidad. Los programas de computador lo hacen automáticamente calculando tanto el coeficiente real como el transformado. Utilizar este procedimiento no cambia en modo alguno la forma de interpretar el signo del coeficiente. Un coeficiente positivo aumenta la probabilidad, mientras que su valor negativo disminuye la probabilidad. Así pues si  $\beta$  es positivo, su transformación (antilog) será mayor a 1, y el odds ratio aumentara. Este aumento se produce cuando la probabilidad prevista de ocurrencia de un suceso aumenta y la probabilidad prevista de su no ocurrencia disminuye. Por lo tanto, el modelo tiene una elevada probabilidad de ocurrencia. De la misma forma, si  $\beta$  es negativo, el antilogaritmo es menor que 1 y el odds ratio disminuye. Un valor de cero equivale a un valor de 1, lo que no produce cambio en el odds.

## 2.4.5. PRUEBA DE HIPOTESIS PARA LOS COEFICIENTES DEL MODELO DE REGRESION LOGISTICA

Usualmente en la estimación del Modelo de Regresión Logística, como en el Modelo de Regresión Lineal Múltiple se efectúan pruebas con objetivos diferentes, siendo estos:

- Determinar si una variable explicativa tiene coeficientes igual a cero.
- Determinar si un conjunto de variables explicativas tienen coeficientes igual a cero.
- Determinar la calidad del ajuste global del modelo.

Veamos para cada objetivo, como se efectúa el análisis.

### 2.4.5.1. PRUEBA DE WALD

Wald (1943) estudio una prueba asintótica para estimaciones máximos verosímiles, y aseveró que los parámetros estimados en los modelos logísticos tienen una Distribución Normal para muestras grandes.

Esta prueba se usa para evaluar la significancia estadística de cada variable explicativa o regresora.

Sea  $\hat{\pi}^{(n)}$  que converge a los EMV de  $\tilde{\pi}$  y  $y_1, y_2, \dots, y_n$  variables respuesta binaria independiente cuyas probabilidades satisfacen.

$$\text{Logit}(\pi_i) = \tilde{x}_i^T \tilde{\beta}$$

$$\text{Donde } \pi_i = P[y_i = 1 / \tilde{x}_i]$$

Siendo  $\tilde{x}_i$  una observación que contiene los valores de las k variables explicativas con  $\tilde{x}_i^T = (1, x_{i1}, x_{i2}, \dots, x_{ik})$ .

Sin pérdida de generalidad seleccionaremos  $\beta_j$  como el parámetro de interés.

Supóngase que las hipótesis son.

$$\begin{aligned} H_0 : \beta_j &= \beta_{j0} \\ H_1 : \beta_j &\neq \beta_{j0} \end{aligned} \tag{58}$$

Sea  $\hat{\beta}_j$  un EMV de  $\beta_j$  y sea:

$I^{-1} = (X^T V X)^{-1}$  la inversa de la matriz de información muestral, entonces la estadística de Wald para docimar (57) es:

$$W = \frac{(\hat{\beta}_j - \beta_{j0})^2}{\sigma^2(\hat{\beta}_j)} \quad (59)$$

Donde  $\sigma(\hat{\beta}_j)$  es la estimación del error estándar de  $\hat{\beta}_j$ .

Bajo  $H_0$ ,  $W \sim \chi^2_{(1)}$  y para  $n$  suficientemente grande se tiene que:

$$z = \frac{\hat{\beta}_j - \beta_{j0}}{\sigma(\hat{\beta}_j)} \sim N\left(\left(\frac{\hat{\beta}_j - \beta_{j0}}{\sigma(\hat{\beta}_j)}\right), 1\right) \quad (60)$$

Por tanto:

$z^2 \sim \chi^2_{(\xi,1)}$ , es  $\chi^2$  con parámetro de no centralización:

$$\xi = \frac{(\beta_j - \beta_{j0})^2}{\sigma(\hat{\beta}_j)} \quad (61)$$

Pero la estadística  $W$ , tiene la propiedad que cuando el valor absoluto del coeficiente de Regresión es grande, el error estándar también lo es; esta situación hace que la estadística  $W$  sea pequeña y por lo tanto se puede rechazar  $\beta_j$  igual cero, cuando en realidad no debería rechazarse.

Por lo tanto, cuando se encuentra que un coeficiente es grande, es preferible no usar la estadística de Wald para efectuar docima individual. Sino se recomienda construir un método con y sin esa variable y basarse en la prueba de hipótesis de la diferencia entre los dos modelos.

Para las hipótesis estadísticas:

$$\begin{aligned} H_0 : \beta_j &= 0 \\ H_1 : \beta_j &\neq 0 \end{aligned} \quad (62)$$

La estadística (59) es:

$$W = \frac{(\hat{\beta}_j)^2}{\sigma^2(\hat{\beta}_j)} \quad (63)$$

Bajo  $H_0$  verdadero,  $W \sim \chi^2_{(1)}$  y para  $n$  suficientemente grande se tiene que:

$$z = \frac{\hat{\beta}_j}{\sigma(\hat{\beta}_j)} \sim N\left(\frac{\hat{\beta}_j}{\sigma(\hat{\beta}_j)}, 1\right) \quad (64)$$

Por tanto:

$$z^2 \sim X^2_{(\xi,1)}$$

$z^2$  se distribuye como una  $\chi^2_{(\xi, 1)}$  con parámetro de no centralización:

$$\xi = \left( \frac{(\beta_j)^2}{\sigma^2(\hat{\beta}_j)} \right) \quad (65)$$

Si la variable explicativa es categórica, los grados de libertad es igual al número de categorías o niveles de la variable menos uno.

**Decisión:** Si  $Wald > \chi^2_{\alpha, (1)gl}$  rechazamos  $H_0$  con un nivel de significancia de  $\alpha$  y concluimos que la variable independiente influye en la probabilidad de las características de la variable dependiente. Si la variable independiente es cualitativa los grados de libertad es igual al número de categorías menos 1.

#### 2.4.5.2. PRUEBA CHI-CUADRADO

Esta prueba sirve para lograr el segundo objetivo propuesto al iniciarse la sección 2.4.4 y sirve para docimar los coeficientes del modelo logístico.

Para elegir un modelo, se usa la prueba de razón de verosimilitud, para probar la hipótesis de que los coeficientes  $\beta_j$  correspondientes a las variables explicativas retiradas, digamos que q variables explicativas, del modelo son iguales a cero, siendo la hipótesis estadísticas:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_q = 0$$

$$H_1 : \beta_j \neq 0, \text{ para por lo menos un } j=1,2,\dots, q$$

Esta prueba se basa en la siguiente estadística:

$$\chi^2_{(q)} = -2[LnL_{p-q} - LnL_p] \quad (66)$$

Bajo la hipótesis de que los coeficientes de las variables retiradas son iguales a cero, la estadística (66) tiene una distribución asintótica  $\chi^2_{(q)gl}$ .

Los valores altos para esta estadística indican que una o más de las q variables retiradas tienen coeficientes de regresión distinto de cero.

La estadística  $\chi^2_{(q)}$  se usa también para probar si una variable explicativa determinada, por ejemplo  $x_k$ , muestra una asociación significativa (como factor de riesgo cuando se aplica a cosas de enfermedades) para con la variable respuesta en la presencia de las demás variables  $x_1, x_2, \dots, x_{k-1}$ .

### 2.4.5.3. ESTADÍSTICA CHI-CUADRADA DE PEARSON

Esta estadística sirve para evaluar el modelo ajustado en forma global. La estadística se basa en la comparación de los valores observados,  $y_i$ ; y sus respectivas probabilidades estimadas,  $\pi_i$ .

Las hipótesis estadísticas son:

$$H_0 : \beta_0 = \beta_1 = \dots = \beta_k = 0$$

$$H_1 : \beta_j \neq 0, \text{ para por lo menos un } j=1,2,\dots, k$$

Esta prueba se basa en la estadística Chi-cuadrada de Pearson, que está dada por:

$$\chi^2 = \sum_{i=1}^n \frac{(y_i - \pi_i)^2}{\pi_i(1 - \pi_i)} \quad (67)$$

$$\text{o equivalentemente } \chi^2 = \sum_{i=1}^n \frac{r_i^2}{v_{ii}} \quad (68)$$

donde:

$$r_i = (y_i - \pi_i)$$

$$v_{ii} = \text{Diag}(\hat{V}) = \pi_i(1 - \pi_i)$$

Bajo la hipótesis nula, de que el método se ajusta bien a los valores observados, la estadística (68) tiene distribución asintótica Chi-cuadrado

$$\chi^2_{(n-(k+1))}$$

Los valores altos de la estadística Chi-cuadrado de Pearson indican discrepancias con el modelo teórico.

La estadística (68) es inestable cuando  $\pi_i$  toma valores cercanos a cero o uno, por ello tomar en cuenta esta observación, cuando se realiza el análisis.

### 2.4.5.4. DESVIANZA

Otra forma de probar el ajuste global del modelo, es mediante la estadística llamada Desviianza, propuesta por Nelder y Wederbum

(1982), es análogo a la suma de cuadrados de los residuales del Modelo de Regresión Lineal Múltiple.<sup>8</sup>

Las hipótesis estadísticas son:

$$H_0 : \beta_1 = \dots = \beta_k = 0$$

$$H_1 : \beta_j \neq 0, \text{ para por lo menos un } j=1,2,\dots, k$$

Esta estadística se usa para evitar la inestabilidad de la estadística Chi-cuadrado de Pearson. La Desvianza esta dada por:

$$D_p = \sum_{i=1}^n d_i^2 \quad (69)$$

Donde:

$$d_i = \begin{cases} \sqrt{-2 \log \hat{p}_i} & \text{si } y_i = 1 \\ \sqrt{-2 \log(1 - \hat{p}_i)} & \text{si } y_i = 0 \end{cases} ; j=1,2,\dots,n$$

La Desvianza bajo la hipótesis nula, asintóticamente, es la misma que la distribución Chi-cuadrado de Pearson, es decir se distribuye  $\chi^2_{(n-(k+1))}$  y mide la discrepancia o el desvío entre el modelo bajo la investigación actual y el modelo saturado.

La estadística (69) para el modelo de regresión logística está dada por:

$$D = -2 \sum (y_i \log(\hat{\pi}_i) + (1 - y_i) \log(1 - \hat{\pi}_i)) \quad (70)$$

Cuando el modelo bajo investigación es verdadero se compara el valor D con el valor crítico  $\chi^2_{(n-p)}$  de una distribución  $\chi^2$  a un nivel de significación igual a  $\alpha$ , por tanto :

Si  $D > \chi^2_{(n-p)}$  el modelo se rechaza y

Si  $D \leq \chi^2_{(n-p)}$  el modelo no se rechaza.

Donde  $p=k+1$ .

#### 2.4.5.5. TABLA DE CLASIFICACION

También llamada Matriz de Confusión, es una forma sencilla de evaluar el ajuste del Modelo de Regresión Logística, no es tan objetiva pero se usa como indicador de bondad de ajuste.

---

<sup>8</sup> Salcedo Poma, Celia Mercedes (Lima 2002). "Estimación de ocurrencia de incidencias en declaraciones de pólizas de importación".

Es una tabla sencilla de 2x2 en el cual se muestra la distribución de los objetos que pertenecen a las categorías 1 y 2, es decir cuando  $y=0$  y cuando  $y=1$ , conjuntamente con la clasificación a cualquiera de las 2 categorías de acuerdo a la probabilidad estimada.

Para interpretar se hace mediante el porcentaje de objetos bien clasificados, esto es, aquellos que mediante la probabilidad estimada permanecen en su respectiva categoría. También se interpreta mediante el porcentaje de objetos mal clasificados, esto es, aquellos mediante la probabilidad estimada se asignan a categorías diferentes del cual fueron observados.

**TABLA DE CLASIFICACION**

GRUPO ACTUAL	GRUPO ESTIMADO		TOTAL MARGINAL
	0	1	
0	$n_{11}$	$n_{12}$	$n_{11} + n_{12}$
1	$n_{21}$	$n_{22}$	$n_{21} + n_{22}$
TOTAL MARGINAL	$n_{11} + n_{21}$	$n_{12} + n_{22}$	$n$

$\frac{n_{11} + n_{22}}{n} \times 100\%$  es el porcentaje de objetos bien clasificados mediante el Modelo de Regresión Logística estimado.

Por tanto, lo que se debe esperar es que este porcentaje sea lo más alto posible, a fin de concluir que el modelo obtenido clasifica bien a los objetos o individuos.

#### **2.4.5.6. CONTRASTE DE BONDAD DE AJUSTE DE HOSMER-LEMESHOW**

Este contraste evalúa la bondad de ajuste del modelo, es decir el grado en que la probabilidad predicha coincide con la observada, construyendo una tabla de contingencia a la que aplica un contraste. Para ello calcula los deciles de la probabilidades estimadas ( $p_i; i=1,2,\dots,n$ ),  $D_1, D_2, \dots, D_9$ , y divide los datos observados en 10 categorías dados por:

$$A_j = \{p_i \in [D_{j-1}, D_j] / i \in \{1, 2, \dots, n\}\} \quad ; \quad j=1, 2, \dots, 10$$

Donde  $D_0 = 0, D_{10} = 1$ .

Sean:

$n_j$  = Número de casos en  $A_j$  ;  $j=1, 2, \dots, 10$

$o_j$  = Número de  $y_j = 1$  en  $A_j$  ;  $j=1, 2, \dots, 10$

$$\bar{p}_j = \frac{1}{n_j} \sum_{i \in A_j} p_i \quad ; \quad j=1, 2, \dots, 10$$

El estadístico del contraste viene dado por:

$$T = \sum_{j=1}^{10} \frac{(o_j - n_j \bar{p}_j)^2}{n_j \bar{p}_j (1 - \bar{p}_j)} \quad (71)$$

y el p-valor del contraste es  $P[\chi_{\alpha}^2 \geq T_{obs}] \bullet z$ .

Las hipótesis a contrastar son:

$H_0$  : El modelo es adecuado

$H_1$  : El modelo no es adecuado

Estadístico de prueba es: si  $T \geq \chi_{\alpha, j-2}^2$  rechazamos  $H_0$  y concluimos que el modelo no es adecuado a un nivel de significancia  $\alpha$ .

#### 2.4.5.7. EL INCREMENTO DEL ESTADÍSTICO $-2\log L$

El estadístico  $-2\log L$  mide los cambios que se producen cuando se agrega o se quita una variable, donde  $L$  es la función de verosimilitud del modelo estudiado, puede oscilar entre 0 y 1, si el modelo se ajusta perfectamente a los datos tiene una verosimilitud igual a 1, de allí que  $-2\log L=0$ . Entonces diremos que el modelo se ajusta bien a los datos si tiene un valor pequeño de  $-2\log L$ , que es el logaritmo de la verosimilitud y se distribuye como una  $\chi^2$  (Chi-cuadrado), cuando el modelo incluye sólo la constante los grados de libertad es igual al número de casos menos uno  $(n-1)$ , y cuando se incluye la variable independiente sigue una distribución  $\chi^2$  con  $(n-k-1)$  grados de libertad, en el modelo de regresión logística simple es  $n-2$ , la diferencia entre estos dos valores de  $-2\log L$  se llama Devianza, esta



prueba cuando la variable  $x_i$  es significativa, se define como:

$D = -2 \log$  (verosimilitud del modelo sin la variable/verosimilitud del modelo con la variable)

$$D = -2 \sum_{i=1}^n \left[ Y_i \log \left( \frac{\hat{p}_i}{Y_i} \right) + (1 - Y_i) \log \left( \frac{1 - \hat{p}_i}{1 - Y_i} \right) \right]$$

Las hipótesis son:

$H_0$  : El modelo ajustado es significativo

$H_1$  : El modelo ajustado no es significativo

Estadístico de prueba es:

$D \sim \chi^2$  con  $(n-k-1)$  grados de libertad

### **Decisión**

Si  $D < X_{\alpha, (n-k-1)}^2$  no rechazamos  $H_0$ , es decir el modelo ajustado es significativo.

## **2.5. DIAGNOSTICO DEL MODELO**

Evalúan la bondad de ajuste caso por caso mediante el análisis de los resultados del modelo y de su influencia en la estimación del vector de parámetros del mismo.

Los residuos más utilizados son los siguientes: Residuos estandarizados, Residuos studentizados y Residuos desviación. Todos estos se distribuyen aproximadamente según  $N(0,1)$ , si el modelo ajustado es cierto.

### **2.5.1. EVALUACION DE $-2LL_0$**

El estadístico  $-2LL_0$  es muy importante en Regresión Logística.  $L$  Indica el logaritmo neperiano y  $L_0$  verosimilitud; puede oscilar entre 0 y 1, por lo tanto  $-2LL_0$  es el logaritmo neperiano de la verosimilitud. Las hipótesis a encontrar son las mismas que en los caso anteriores. La evaluación comienza conociendo el valor de este estadístico, que se distribuye como una ji-cuadrada. Para el modelo que contiene solo la constante los grados de libertad son iguales al número de casos menos 1; a continuación calculamos el valor cuando hemos incluido en el

modelo la variable independiente, los grados de libertad son iguales al número de casos menos el número de variables independientes, menos 1. En caso de la regresión logística simple, como hay una sola variable independiente, los grados de libertad son el número de casos menos 2. Calculamos la diferencia entre los dos valores del estadístico, la diferencia de un estadístico que se distribuye como un ji-cuadrado con grados de libertad igual a la diferencia de grados de libertad entre las dos situaciones, que es igual al número de variables independientes del modelo, uno en caso simple, el valor de la diferencia es un estadístico que se distribuye como una ji-cuadrado con 1 grado de libertad una posible formulación de hipótesis para la verosimilitud sería:

$$H_0 : L_0 = 0$$

$$H_1 : L_0 \neq 0$$

### **2.5.2. RIESGO RELATIVO Y ODDS RATIO**

El primer paso de un análisis estadístico sobre los coeficientes de un modelo de regresión, como ya se ha visto, es su significatividad. Si hay significancia estadística, en un segundo paso hay que estudiar la fuerza de asociación estadística.

Para estudiar la fuerza de asociación estadística existen muchos parámetros que pueden ser utilizados, nosotros tomaremos entre otros el riesgo relativo (RR) y el ODDS RATIO. También llamado razón de predominio (OR). Los dos pueden calcularse mediante derivaciones del modelo regresión logística, aunque el más utilizado en regresión logística es el OR por su facilidad de cálculo.

No siempre es correcto utilizar indistintamente RR y OR; esto depende de los diseños experimentales, en los estudios prospectivos, en los que la muestra se ha extraído aleatoriamente de la población, pueden utilizarse los dos.

En los estudios de casos y controles, y en otros casos en los que la muestra no es un reflejo estadísticamente representativo de la población, solo es correcto el uso del OR.

### **2.5.3. RIESGO RELATIVO O FACTOR DE PROTECCION (RR)**

Dentro del campo de las ciencias de la salud se considera factor de riesgo a una característica que aumente la probabilidad de padecer una determinada enfermedad. Factor de protección es una característica que disminuye la probabilidad de padecer una determinada enfermedad. El concepto de factor de

riesgo podemos generalizar a otros casos en los que la variable dependiente no sea una variable cuyas categorías sean padecer o no una determinada enfermedad. Los factores que aumentan la probabilidad de que ocurra un determinado suceso, sean estos de la naturaleza que sean, podemos denominarlos factores favorecedores y, a los que disminuyan la probabilidad, factores entorpecedores.

El RR es la razón entre la probabilidad de padecer la enfermedad en presencia del factor y la probabilidad de padecer la enfermedad en ausencia del factor.

El RR nos indica cuantas veces es más probable contraer la enfermedad en presencia del factor que en ausencia del factor. Si no existe ninguna relación entre la enfermedad y el factor, el valor de RR es 1, un número significativamente mayor que 1, indica que el factor que estamos evaluando es un factor de riesgo y un número menor que 1 indica que el factor es de protección.

No es correcto el cálculo de RR en todos los tipos de diseño; debe utilizarse cuando tengamos una muestra estadísticamente representativa de una población y la seguridad de que el factor que queramos evaluar es anterior al suceso en estudio.

La expresión para el riesgo relativo es:

$$RR = \frac{P(E / F)}{P(E / \bar{F})}$$

Donde:

F: significa presencia del factor.

$\bar{F}$  : en ausencia del factor.

#### **2.5.4. RAZON DE PREDOMINIO (OR)**

El predominio (PR), es una relación entre dos probabilidades. Hay un predominio en presencia del factor (PRF) y un predominio en ausencia del factor (PRN). El predominio en presencia del factor (PRF) es la razón entre la probabilidad de padecer la enfermedad en presencia del factor y la probabilidad de no padecer la enfermedad en presencia del factor.

El (PRN) es la razón entre la probabilidad de padecer la enfermedad en ausencia del factor y la probabilidad de no padecer la enfermedad en presencia del factor. La razón entre dos predominios, más conocida como (OR), es una medida de riesgo, su expresión es: <sup>9</sup>

$$OR = \frac{\text{odds (enfermedad / exp uestos )}}{\text{odds (enfermedad / no exp uestos )}}$$

$$OR = \frac{p(y = 1 / x = 1) / p(y = 0 / x = 1)}{p(y = 1 / x = 0) / p(y = 0 / x = 0)}$$

$$OR = \frac{(1 / 1 + e^{-(\alpha+\beta)}) / (1 - 1 + e^{-(\alpha+\beta)})}{(1 / 1 + e^{-(\alpha+\beta)}) / (1 - 1 + e^{-(\alpha+\beta)})}$$

$$OR = e^{\beta}$$

$OR = e^{\beta}$  : cuantifica la magnitud de la asociación entre la respuesta y el factor de interés.

### 2.5.5. RESIDUOS DEL MODELO

Los residuos más utilizados son los siguientes:

**Residuos estandarizados.**- Son el cociente entre los residuales y una estimación de la desviación estándar. <sup>10</sup>

$$z_i = \frac{y_i - \hat{p}_i}{\sqrt{\hat{p}_i(1 - \hat{p}_i)}} ; i=1,2,\dots,n \quad (72)$$

**Residuos studentizados.**- Son el cambio en el valor de la desviación del modelo si el caso es excluido.

$$st_i = \frac{y_i - \hat{p}_{(i)}}{\sqrt{\hat{p}_{(i)}(1 - \hat{p}_{(i)})}} ; i=1,2,\dots, n \quad (73)$$

donde  $\hat{p}_{(i)}$  es la estimación de  $p_i$  obtenida eliminando la observación  $i$  de la muestra.

---

<sup>10</sup> Desviación: es otra forma de probar el ajuste global del modelo, es mediante la estadística llamada Desviación, propuesta por Nelber y Wederbum (1982).

**Residuos Desvianza.-** Para cada observación la desvianza se calcula:

$$d_i = \begin{cases} \sqrt{-2 \log \hat{p}_i} & \text{si } y_i = 1 \\ \sqrt{-2 \log(1 - \hat{p}_i)} & \text{si } y_i = 0 \end{cases} ; j= 1,2,\dots, n \quad (74)$$

Todos estos residuos se distribuyen aproximadamente como una  $N(0,1)$ , si el modelo ajustado es correcto.

### 2.5.6. MEDIDAS DE INFLUENCIA

Cuantifican la influencia que cada observación ejerce sobre la estimación del vector de parámetros o sobre las predicciones hechas a partir del mismo, de modo que, cuanto más grande son, mayor es la influencia que ejerce una observación en la estimación del modelo.

#### Medida de Apalancamiento (Leverage)

Se utiliza para detectar observaciones que tienen un gran impacto en los valores predichos por el modelo.

Se calcula a partir de la matriz  $H=W^{-1/2} X (X^T W X)^{-1} X^T W^{-1/2}$  donde  $W = \text{diag} [p_i(1 - p_i)]$ . El apalancamiento para la observación  $i$ -ésima viene dado por el elemento  $i$ -ésimo de la diagonal principal de  $H$ ,  $h_{ii}$  y toma valores entre 0 y 1 con un valor medio de  $p/n$ .

Las dos medidas siguientes miden el impacto que tiene una observación en la estimación de  $a$ .

**Distancia de Cook.-** Mide la influencia en la estimación de  $a$ .

$$COOK_i = \frac{1}{p} (\hat{a} - \hat{a}_{(i)})^T X^T W X (\hat{a} - \hat{a}_{(i)}) \quad (75)$$

**DFBETA.-** Mide la influencia en la estimación de una componente de  $a$ ,  $a_i$

$$Dfbeta_{1i} = \frac{\hat{a}_i - \hat{a}_{1(i)}}{\text{std}(\hat{a}_i)} \quad (76)$$

Donde  $\hat{a}_i$ ,  $\hat{a}_{1(i)}$  denotan las estimaciones del modelo logístico de  $a$  y  $a_i$ , eliminando la  $i$ -ésima observación de la muestra y  $\text{std}(\hat{a}_i)$  el error estándar en la estimación de  $a_i$ .

## 2.6. VARIABLES EXPLICATIVAS CATEGORICAS – VARIABLES DUMMY

Si una de las variables explicativas es categórica, con  $c$  valores posibles, se crean  $c-1$  variables dicotómicas como variables explicativas también llamadas variables Dummy. Estas variables cuantifican el efecto de un valor de dichas variables con respecto a un valor de referencia.

Estas variables se usan cuando los datos se muestran como categorías, las categorías pueden ser:

**Nominales:** La variable simplemente indica diferentes categorías, las categorías no pueden ser ordenadas en un orden particular. Ejemplo; Sexo (hombre, mujer).

**Ordinales:** La variable además de estar agrupada en categorías puede ser ordenada. El que una categoría este en un orden superior que otra implica que su medida representa algo mayor que la otra. Ejemplo: clase social (baja, media, alta).

Cuando se tiene variables de este tipo se crean las variables llamadas dummy, si la variable tiene  $c$  categorías se usan  $c-1$  variables ficticias o dummy. La variable indica si un dato corresponde a una categoría o no veamos un ejemplo de cómo se hace esto:

Supongamos que tenemos una variable clase social, codificada 1: Baja, 2: Media, y 3: Alta, entonces creamos dos variables dummy:

Clase 1: 1 si el dato corresponde a la clase social baja, 0 si el dato no pertenece a la clase baja.

Clase 2: 1 si el dato corresponde a la clase social media, 0 si el dato no pertenece a la clase Media.

Como se puede ver estas nos permiten clasificar cualquier dato en una de las categorías existentes.

Supongamos que creamos las variables Dummy  $Ind_1$  e  $Ind_2$  para una variable  $X_1$  de tres categorías, como sigue:

$X_1$	$Ind_1$	$Ind_2$
1	1	0
2	0	1
3	0	0

En este caso la ecuación del modelo ajustado viene dada por:

$$\log \left( \frac{p(\text{Ind } 1, \text{Ind } 2; \beta)}{1 - p(\text{Ind } 1, \text{Ind } 2; \beta)} \right) = \beta_0 + \beta_1 \text{Ind } 1 + \beta_2 \text{Ind } 2 \quad (77)$$

Sea  $p_i = p[y = 1 / X_1 = i]$ ;  $i=1,2,3$ . Se tiene:

$$\frac{P_1}{1 - p_1} = e^{\beta_0 + \beta_1}, \frac{P_2}{1 - p_2} = e^{\beta_0 + \beta_2}, \frac{P_3}{1 - p_3} = e^{\beta_0}$$

Se sigue que:

$$\frac{P_1}{1 - p_1} / \frac{P_3}{1 - p_3} = e^{\beta_1}, \frac{P_2}{1 - p_2} / \frac{P_3}{1 - p_3} = e^{\beta_2}$$

Por lo tanto  $e^{\beta_i}$ ,  $i=1,2$  compara los odds ratio correspondientes a  $X_1=1, 2$ , frente al de la categoría de referencia  $X_1=3$ .

## 2.7. INTERPRETACIÓN DE LOS RESULTADOS

La interpretación de los resultados obtenidos se realiza a partir de la interpretación de los coeficientes del modelo. Para ello basta tener en cuenta que si el modelo ajustado es bueno, entonces se dice que el modelo es significativo, pero además se debe analizar el grado de asociación estadística que existen en sus parámetros, a partir de la ecuación (8) se tiene:

$$\log \left[ \frac{P(X_1, \dots, X_K; \beta)}{1 - P(X_1, \dots, X_K; \beta)} \right] = \beta_0 + \beta_1 X_1 + \dots + \beta_K X_K$$

donde el "odds ratio" que es el factor de riesgo está dado por la razón de esta expresión:

$$\frac{P(X_1, \dots, X_K; \beta)}{1 - P(X_1, \dots, X_K; \beta)} = e^{\beta_0 + \beta_1 X_1 + \dots + \beta_K X_K}$$

Entonces:

$$\frac{\frac{P(X_1 + 1, \dots, X_K; \beta)}{1 - P(X_1 + 1, \dots, X_K; \beta)}}{\frac{P(X_1 + 1, \dots, X_K; \beta)}{1 - P(X_1 + 1, \dots, X_K; \beta)}} = e^{\beta_1}$$

Por lo tanto,  $e^{\beta_1}$  es el factor de cambio en el "odds ratio" (OR) de riesgo si el valor de la variable  $X_1$  cambia en una unidad. Así, si  $\beta_1 > 0$  (ó  $\beta_1 < 0$ ) el factor será mayor que 1 y  $P(X_1, \dots, X_K; \beta)$  aumentará (disminuirá). Si  $\beta_1 = 0$  la variable  $X_1$  no ejerce ningún efecto sobre  $P_1$ .

$\beta_0$  es un ajuste de escala. Su mejor interpretación se obtiene calculando el valor de  $P(X_1, \dots, X_k; \beta)$  en los valores medios de  $X_1, \dots, X_k$  y usar como variables explicativas sus valores estandarizados.

En regresión logística la medida de asociación más empleada es el OR debido que el número  $e$  es la base de los logaritmos neperianos y elevados a un coeficiente de regresión logística del factor, si es mayor que 1 supone un aumento unitario, indica que el factor de riesgo es mayor.

Si el modelo de Regresión Logística es significativo y una de las variables independientes es dicotómica con valores de 0 y 1, el número  $e = 2.7182$  elevado al coeficiente de regresión logística es el OR, denominado factor de riesgo o protección que implica un aumento unitario de la variable independiente. En el caso de una variable cuantitativa,  $e$  elevado a  $\beta_1$  es el número de veces que aumenta la probabilidad de padecer una enfermedad por cada unidad de aumento de la variable independiente, o dicha de otra manera, cuántas veces es más probable que padezca la enfermedad una persona que presenta síntomas relacionadas a ella.

## **2.8. CONSIDERACIONES GENERALES**

### **2.8.1. Definición de la patología**

Enfermedad infecciosa aguda, causada por el nuevo virus de Influenza A (H1N1), que se transmite de humano a humano a través del contacto directo o indirecto con personas infectadas o sus secreciones respiratorias y que tiene una gran variabilidad clínica, desde casos asintomáticos hasta graves y fatales.

### **2.8.2. Etimología**

Esta enfermedad es causada por el virus de la influenza de tipo A sub tipo H1N1. Estos virus, pertenecientes a la familia Orthomyxoviridae al igual que los virus de tipo B y C, son partículas pleomorfas de apariencia esférica con una envoltura lipoproteica. Por debajo de la envoltura se encuentra una capa constituida por la proteína matriz (M1), que es el soporte estructural de la partícula. En el interior, existe un conjunto de ribo nucleoproteínas (vRNPs) que constituyen el genoma del virus.



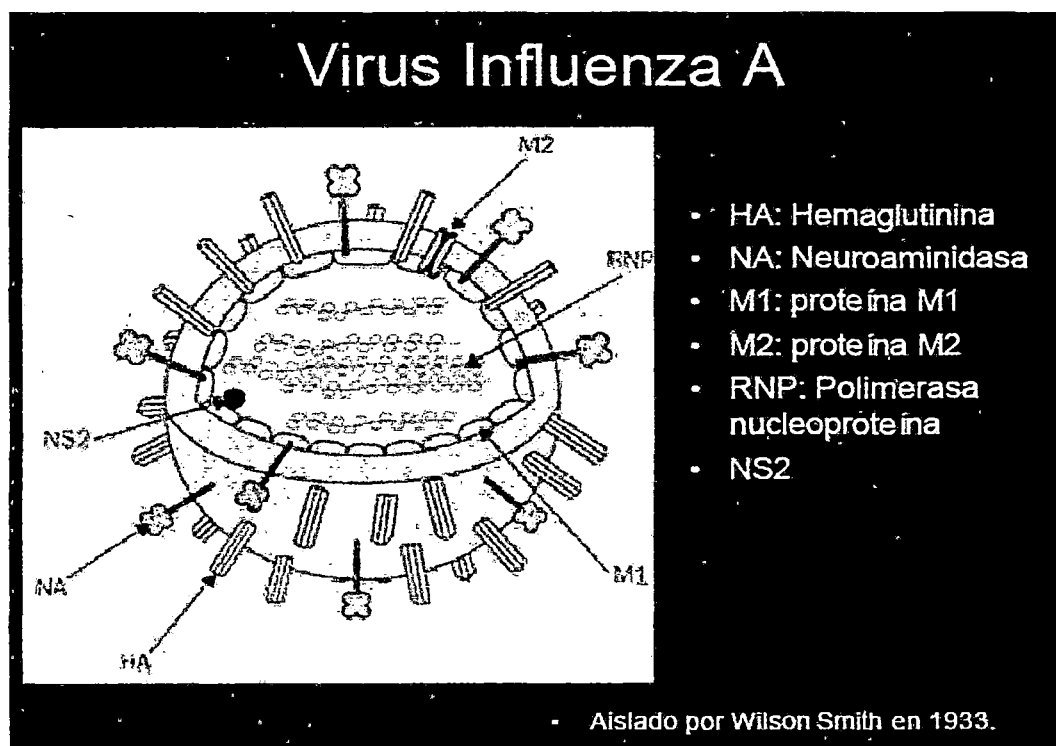
En la superficie del virus, se encuentran dos glicoproteínas, la hemaglutinina (HA) y la neuraminidasa (NA), cuya estructura determina los subtipos serológicos que podemos definir entre los virus gripales tipo A. Así existen 16 subtipos diferentes de HA (H1-H16) y 9 de NA (N1-N9).<sup>11</sup>

### INFLUENZA

#### Diferencia entre los Diferentes Subtipos

Características	Serotipo A	Serotipo B	Serotipo C
Huésped	Humanos, aves, cerdos, equinos	Solo humanos	Humanos, cerdos
Manifestación clínica	Grandes pandemias	Enfermedad severa no pandemia	Enfermedad leve

Figura N° 3



Guía técnica:

"Guía de práctica clínica para el diagnóstico y tratamiento de influenza A (H1N1)"; (MINSa 2009).

### 2.8.3. Fisiopatología

El virus de la influenza tipo A posee una alta patogenicidad debido a su capacidad de mutar en forma rápida y de incorporar en su estructura genes de virus que infectan a otras especies animales. La pandemia actual probablemente deriva de la recombinación de virus provenientes de animales (porcino, aves) y de humano. A la fecha, se desconoce con precisión como surge, pero si una especie es infectada simultáneamente por el virus de la influenza humana y el virus de la influenza porcina, existe la posibilidad que se produzca un intercambio genético que tiene como resultado un nuevo sub tipo viral. Por lo tanto, si la cepa (H1N1) de la influenza porcina se combina con una cepa de la influenza humana dentro de un mismo huésped, puede ocurrir una mutación que origine una nueva cepa capaz de replicarse y transmitirse entre humanos, con una mayor facilidad de transmisión y letalidad que la influenza común. Estos fenómenos de derivación genética han podido producirse con mayor facilidad en los cerdos, ya que estos animales son susceptibles de sufrir la infección por virus humanos y aviares en forma simultánea. Las personas que trabajan con cerdos pueden jugar un rol importante como “mezcladores” de cepas de virus de influenza, que lleven a la recombinación genética y al desarrollo de una nueva progenie con potencial pandémico.

Todos estos cambios podrían explicar la aparición de ciclos caracterizados por pandemias esporádicas, seguidos de periodos de endemia y epidemia, que suceden cuando la población que fue afectada inicialmente desarrolla nuevos anticuerpos contra el virus y se inmuniza.

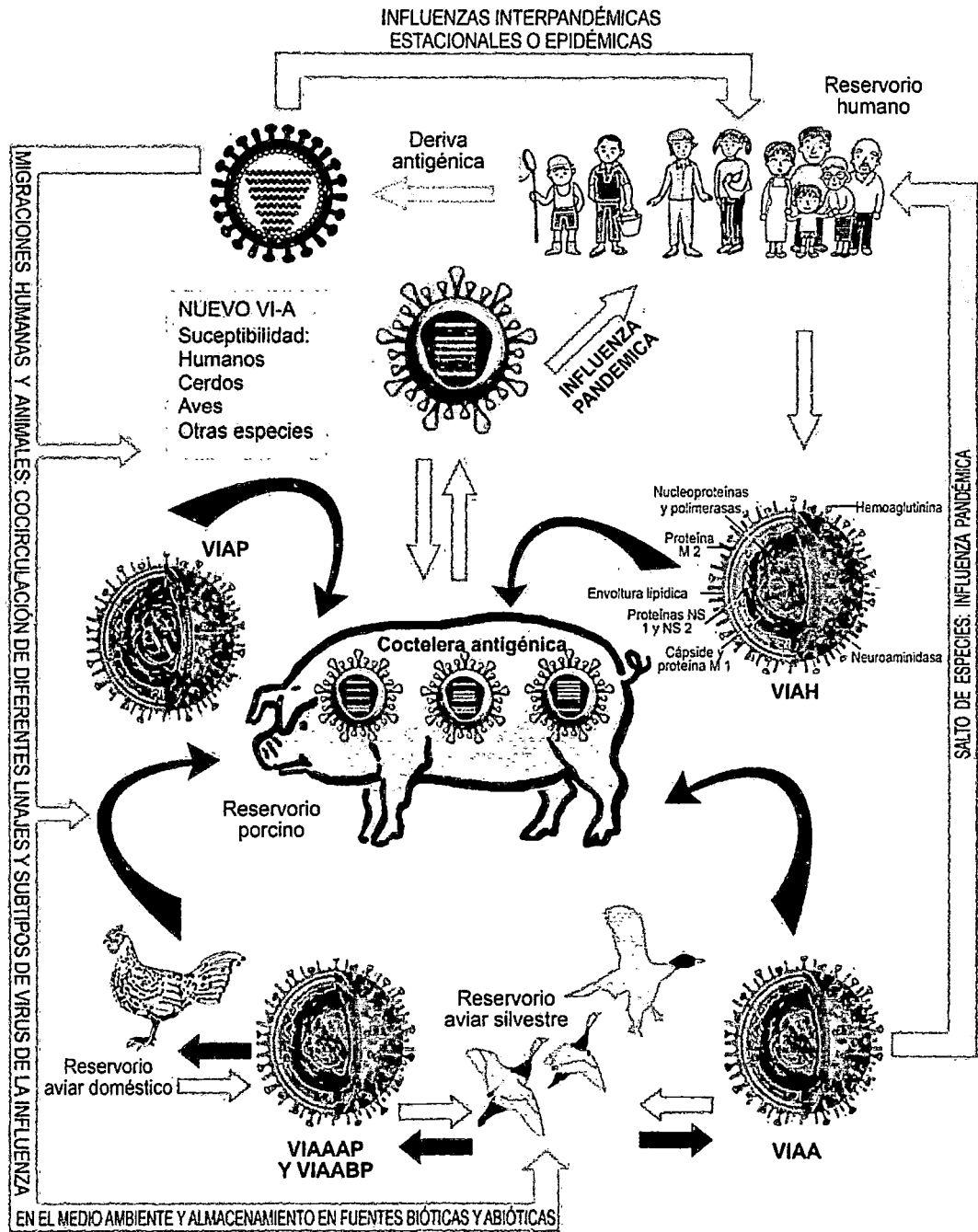
En la siguiente figura se observa de los VI-A del linaje humano, aviar y porcino, interaccionan en un ecosistema o medio ambiente común. En estas condiciones los VI-A siguen caminos de deriva antigénica menores, así como de reordenamientos o cambios antigénicos mayores. En el caso de los reordenamientos de los VI-A, el cerdo posee receptores para los VIAA y VIAH, por consiguiente puede infectarse simultáneamente con ambos linajes y el suyo mismo sirviendo de una coctelera antigénica. En ciertas oportunidades los VI-A pueden saltar directamente de una especie a otra como lo observado con el VIAA H5N1.<sup>12</sup>

---

Guía técnica:

“Guía de práctica clínica para el diagnóstico y tratamiento de influenza A (H1N1)”; (MINSA 2009).

Figura N° 4



#### 2.8.4. Aspectos Epidemiológicos

De lo ocurrido a lo largo de la historia, se ha observado que las pandemias de Influenza tienden a producirse en tres y cuatro veces cada siglo, como resultado de la aparición de un nuevo subtipo viral que se transmite entre humanos. En el

siglo xx, a la gran pandemia de Influenza de 1918-1919, que causo la muerte de aproximadamente 50 millones de personas en todo en mundo, surgieron las pandemias de 1957-1958 y 1968-1969.

Desde el 18 de marzo del 2009. El Gobierno de México comienza a reportar los primeros casos de una enfermedad semejante a la influenza humana ocasionados por el virus de la influenza A (H1N1). El día 21 de Abril el Centro para el Control de Enfermedades de los EE.UU. (CDC) reporta igualmente la presencia de 7 casos humanos confirmados de Influenza A (H1N1) en los EE.UU. (5 en California y 2 en Texas). Para el 25 del mismo mes. La Dirección General de la OMS realiza una declaración manifestando que la situación actual constituye una emergencia sanitaria de importancia internacional, aceptándose el comienzo de la actual pandemia de Influenza humana A (H1N1).

La OMS, en base a criterios acordados, ha establecido las fases de una pandemia que señalan las medidas de respuesta que los países deben adoptar ante una Potencial Pandemia de Influenza. Al 29 de Abril del 2009, el mundo se encuentra en la fase 5 de la Alerta Pandemica, ya que se han detectado casos de transmisión humano – humano en varios países del mundo.

#### **2.8.5. Transmisión**

La enfermedad se transmite por contacto directo, entre los contactos cercanos, incluyendo al personal de salud y familiares.

**Se define contacto cercano** aquel que sea ha cuidado, convivido, o tenido contacto directo con secreciones respiratorias o secreciones de un caso sospechoso, probable o confirmado de influenza A (H1N1).

Los mecanismos de transmisión son:

- Inhalación de gotitas infecciosas o núcleos de gotitas.
- Contacto directo con personas enfermas.
- Contacto indirecto con secreciones respiratorias y auto inoculación en el tracto respiratorio alto o mucosa conjuntival.

## **2.9. FACTORES DE RIESGO ASOCIADO**

### **2.9.1. Medio Ambiente**

De haber transmisión activa de virus de Influenza A (H1N1) en la comunidad, el riesgo será menor si se permanece mayor tiempo en casa, lo que resulta una medida importante para personas de edad avanzada o inmunosuprimidas. La transmisión no guarda relación con la época del año. Sin embargo para la influenza estacional esta es más frecuente en los meses de invierno.

### **2.9.2. Estilos de vida**

Los siguientes factores de riesgo deben causar sospecha de virus de influenza A (H1N1):

- Contacto cercano con casos confirmados de influenza A (H1N1) mientras el caso estuvo enfermo.
- Viajes reciente a áreas donde se han confirmado casos de influenza A (H1N1).

### **2.9.3. Grupos de Alto Riesgo:**

Una persona de alto riesgo de complicaciones para infección por el nuevo virus de influenza A (H1N1) está definida por los mismos criterios que para la influenza estacional.

- Niños menores de 5 años. El riesgo de complicaciones severas es más alto en niños menores de 2 años.
- Adultos de 60 años a más.
- Personas con las siguientes condiciones:
  - Enfermedades pulmonares crónicas (incluyendo el asma), cardiovasculares (excepto hipertensión), renales, hepáticas, hematológicas (incluyendo enfermedad de sickle cell), neurológicas, neuromusculares o metabólicas (incluyendo diabetes mellitus).
  - Inmunosupresión, incluyendo las ocasionadas por medicamentos o por HIV.
  - Mujeres embarazadas.
  - Personas menores de 19 años de edad quienes reciben terapia prolongada con ácido acetil salicílico.

- Residentes de casas de reposo u otros servicios de cuidados crónicos.

## **2.10. CONSIDERACIONES ESPECÍFICAS**

### **2.10.1. Cuadro Clínico**

#### **2.10.1.1. Grupos de signos y síntomas relacionados con la patología.**

Las manifestaciones de la Influenza A (H1N1) presentado en México (casos severos) y EE.UU. (casos leves) son similares a la influenza humana estacional, con enfermedades respiratorias aguda, incluyendo los siguientes signos y síntomas:

- Fiebre
- Tos seca (ocasionalmente con esputo)
- Dolor de garganta
- Dolores musculares
- Cefalea
- Disnea
- Rinorrea
- Escalofríos
- Diarrea y vomito (ocasional)

Las muertes producidas en México han sido por Neumonía y por falta respiratoria. No está claro si estas han sido por efecto del mismo virus o consiguiente proliferación bacteriana.

### **2.10.2. Diagnostico**

#### **2.10.2.1. Definiciones de Caso**

Las definiciones de caso pueden variar de acuerdo al desarrollo de la epidemia.

- **Caso sospechoso:**

Se considera a quien cumpla algunas de las siguientes condiciones:

- Personas con **fiebre mayor de 38°C** acompañada de al menos uno de los siguientes signos o síntomas:
  - Tos
  - Dolor de garganta
  - Rinorrea
- Que además:

- Estuvo en los 7 días previos al inicio de su enfermedad en una zona con casos confirmados de infección por virus de influenza A (H1N1).
- Tuvo contacto cercano con un caso confirmado de infección con virus de influenza A (H1N1).

- **Caso probable:**

- Un caso sospechoso con resultado de prueba de Influenza positiva para influenza A, pero no subtipificado por los reactivos comúnmente utilizados para detectar la infección por el virus de la influenza estacional.
- Persona que murió a causa de una infección respiratoria aguda explicada y con nexo epidemiológico con un caso probable o un caso confirmado.

- **Caso confirmado:**

- Persona con una prueba de laboratorio confirmatoria de infección con virus de la influenza A (H1N1) en un laboratorio de referencia, por una o más de las siguientes pruebas:
  - RT-PCR en tiempo real. (Reacción de la Cadena de Polimerasa en Transcripción Reversa).
  - Cultivo viral.

- **Caso descartado:**

Todo caso sospechoso o caso probable que luego de la investigación epidemiológica y de laboratorio tiene resultados negativos o la presencia de virus Influenza A (H1N1).<sup>13</sup>

#### **2.10.2.2. Criterios de diagnóstico:**

- **Epidemiológico:** Toda persona con enfermedad respiratoria leve o grave, de países o territorios en los cuales se ha detectado el virus de la Influenza A (H1N1).

---

Guía técnica:

“Guía de práctica clínica para el diagnóstico y tratamiento de influenza A (H1N1)”; (MINSa 2009).

- **Clínico:** Cuadro general o respiratorio agudo o grave, caracterizado generalmente por: fiebre alta acompañado de tos, dolor de garganta o Rinorrea.
- **Radiológico:** Presentación de infiltrados o consolidación, segmentación con broncograma aéreo.
- **Viroológico:** La confirmación laboratorial de la Influenza A (H1N1) requiere al menos uno de los siguientes:
  - a) Un análisis de RT PCR en tiempo real de la Influenza A (H1N1) positivo
  - b) Un cultivo viral positivo.

- **RT-PCR en tiempo real:**

Está recomendada para la confirmación del nuevo virus de influenza A (H1N1). La prueba de RT-PCR puede ser positiva para influenza A pero negativa para los subtipos estacionales H1 y H3. En ese caso se debe confirmar la presencia del nuevo virus de influenza A (H1N1) con un RT19 PCR específica. El tiempo aproximado para tener el resultado es de 24 hs, pero el protocolo de la OMS indica repetir las determinaciones negativas a fin de confirmar el diagnóstico por lo que se requiere un tiempo adicional.

- **Cultivo de virus:**

El aislamiento del nuevo virus de influenza A (H1N1) es diagnóstico de infección. Los resultados se pueden obtener entre 7 a 10 días por lo que no es útil para el manejo clínico. Un cultivo negativo no excluye infección por el nuevo virus. Para el diagnóstico del nuevo virus de influenza A (H1N1) se realiza el cultivo y secuenciación del genoma.

### 2.10.2.3. Diagnóstico diferencial

Para el diagnóstico diferencial pueden considerarse las siguientes entidades clínicas.

- Influenza estacional
- Neumonía bacteriana
- Neumonía viral (o por otros agentes)



- Bronquitis aguda
- Bronquiolitis
- Resfrió común
- Crisis asmática

#### **2.10.2.4. Signos de Alarma**

##### **ADULTOS**

- Dificultad para respirar o dolor en el pecho
- Cianosis
- Vomito o diarrea persistente
- Signos de deshidratación
- Trastorno del estado de conciencia
- Deterioro agudo de la función cardíaca
- Agravamiento de una enfermedad crónica.

##### **NIÑOS**

- Fiebre alta y dificultad para respirar
- Cianosis
- Aumento de frecuencia respiratoria:
  - Entre 2 y 11 meses : más de 50 respiraciones por minuto
  - Entre 1 y 5 años : más de 40 respiraciones por minuto
- Rechazo a la vía oral
- Signos de deshidratación ( Ej. llanto sin lagrimas)
- Irritación y/o convulsiones
- Trastorno del estado de conciencia

#### **2.10.2.5. Criterios de Alta**

##### **Criterios para alta de la Unidad de Cuidados Intensivos**

- Toleró entubación por 48 horas
- Respiración espontanea por 48 horas
- Hemodinamicamente estable.

#### **2.10.2.6. Pronóstico**

Hasta el momento la tasa de letalidad en los casos confirmados no ha sido mayor del 10% (experiencia mexicana), siendo mayormente afectados los grupos de riesgo.

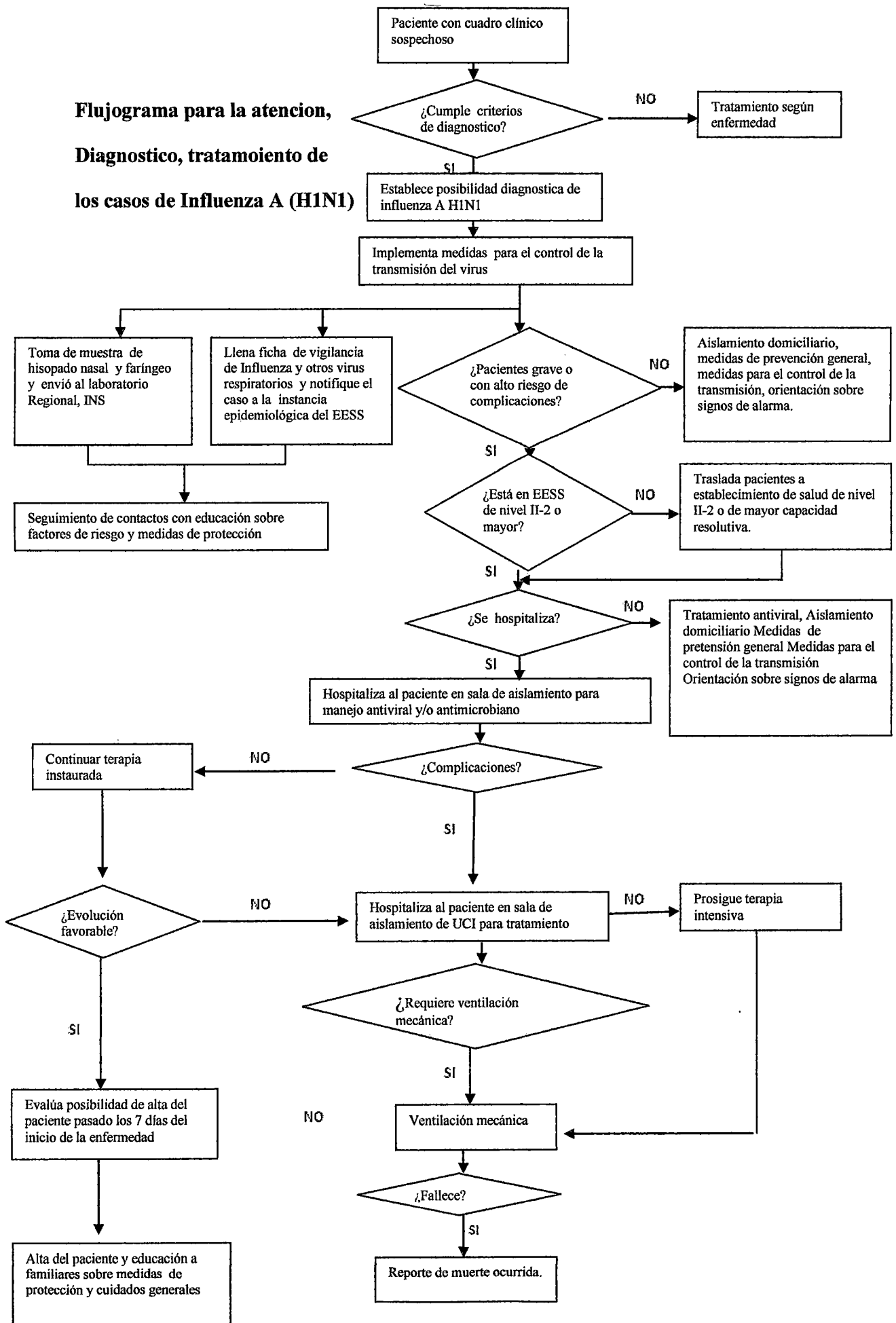
La muerte ocurrió en promedio de 9 a 10 días después del inicio de la enfermedad, la mayoría de los pacientes murieron a causa de la falta respiratoria progresiva, en el siguiente cuadro se muestra flujograma para la atención de casos de Influenza A (H1N1).<sup>14</sup>

---

Guía técnica:

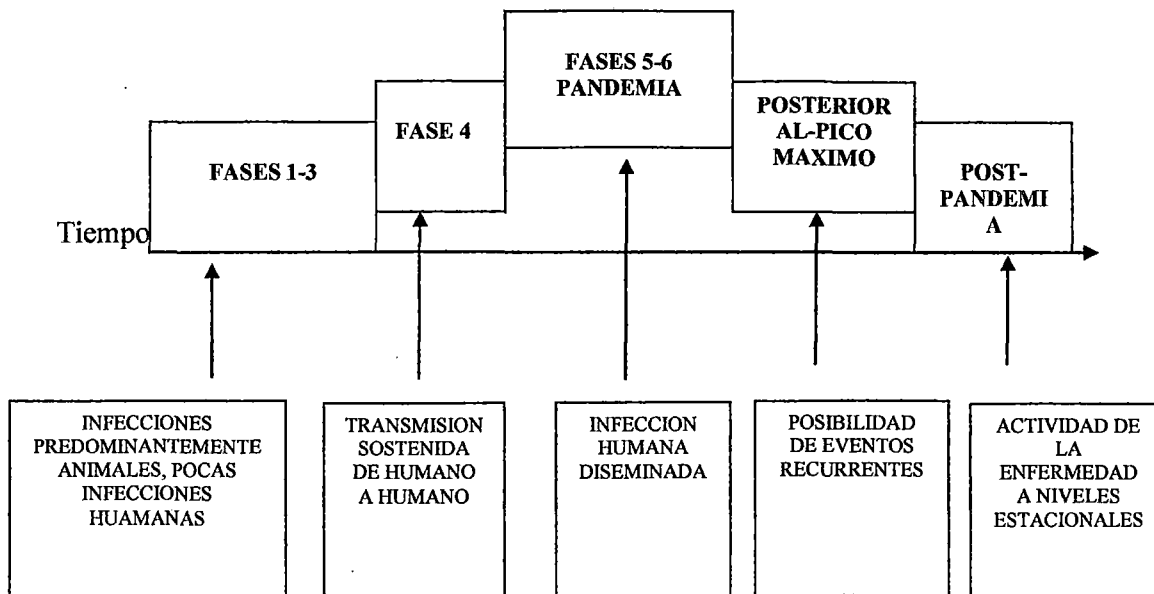
"Guía de práctica clínica para el diagnóstico y tratamiento de influenza A (H1N1)"; (MINSA 2009).

**Flujograma para la atención,  
Diagnostico, tratamiento de  
los casos de Influenza A (H1N1)**



**2.11. Fases de la Pandemia de Influenza según la Organización Mundial de la Salud.**

**FASES DE LA INFLUENZA PANDEMICA**



- **Fase 1:** no hay virus circulantes en animales que hayan afectado a seres humanos.
- **Fase 2:** hay virus circulantes ya sea en animales domésticos o salvajes que han causado infecciones humanas, considerándose una posible amenaza de pandemia.
- **Fase 3:** hay virus circulante en animales o recombinado humano-animal que han causado casos esporádicos o en pequeños grupos pero no se ha demostrado transmisión de persona a persona o si se produjo solo fue en casos de contactos íntimos. El riesgo de pandemia es bajo.
- **Fase 4:** hay virus circulantes en animales o recombinado humano-animal con transmisión demostrada de persona a persona y que ha causado brotes en comunidades, regiones o países. Esta fase implica un riesgo importante de pandemia aunque no se puede asegurar que la misma se produzca.
- **Fase 5:** se caracteriza por la propagación del virus de persona a persona al menos en dos países de una región de la OMS. Aunque la mayoría de los países

no estarán afectados en esta fase, la declaración de la fase 5 es un indicio claro de la inminencia de una pandemia.

• **Fase 6:** es la fase de pandemia e implica la aparición de un brote en al menos un tercer país de otra región OMS. La declaración de esta fase indica que esta en marcha una pandemia mundial.

**Post- Pandemia:** en el periodo posterior de la máxima actividad, la intensidad de la pandemia en la mayoría de los países con una vigilancia adecuada habrá disminuido por debajo de la observada en el momento álgido. En este periodo, la pandemia parece remitir, sin embargo, no pueden descartarse nuevas oleadas, y los países han de estar preparados para una segunda ola.

Las pandemias anteriores se han caracterizado por oleadas de actividad repetidas durante varios meses. Cuando el número de casos disminuye, se requiere una gran habilidad comunicadora para compaginar esa información con la advertencia de que pueda producirse otro ataque. Las olas pandémicas pueden sucederse a intervalos de meses, y cualquier señal de “relajación” puede resultar prematura.

En el periodo post-pandémico, los casos de gripe habrán vuelto a ser a comparables a los habituales de la gripe estacional. Cabe pensar en el virus pandémico se comportara como un virus estacional de tipo A. En esta fase es importante mantener la vigilancia y actualizar en consecuencia la preparación para una pandemia y los planes de respuesta. Puede requerirse una fase intensiva de recuperación y evaluación.

## **2.12. Obtención de la muestra para diagnóstico de Influenza A (H1N1)**

### **A. Muestra utilizada**

- Hisopado nasal y faríngeo combinado

### **B. Materiales:**

- Ficha clínica
- 02 Hisopos de dacron
- Medio de transporte viral
- Equipo de protección personal (respirador N95, guantes, lentes y mandilón)

### C. Suero

- 1.5 cc de suero de fase aguda y convaleciente (sueros pareados).  
Transportar en cadena de frío o guardar a -20°C.

### D. Procedimiento

- Obtención de la primera muestra por hisopado nasal (hasta los cornetes).
- Obtención de la muestra por hisopado faríngeo.
- Poner los dos hisopos en un tubo con medio de transporte.
- Colocar los tubos en una termo con bloquetas de hielo para su envío en menos de 12 horas.

#### Envío de muestra para el diagnóstico de Influenza A (H1N1)

<b>HOSPITALES (MINSA, ESSALUD, FFAA, FFPP, CLINICAS PARTICULARES)</b>	<b>DISA/DIRESA LABORATORIO REFERENCIAL</b>	<b>INSTITUTO NACIONAL DE SALUD</b>
- Identificación de caso sospechoso - Llenado de ficha de investigación - Obtención de muestras (en los 3 primeros días de la enfermedad).	- Investigación y seguimiento del caso - Envío inmediato de muestra al INS (2-8°C).	Procesamiento de muestra: - Inmunofluorescencia - PCR - Cultivo.

### 2.13. Información para la Comunidad

El virus de influenza o gripe se caracteriza por infectar no solo al hombre sino también a otros animales como por ejemplo aves y/o cerdos. Al mismo tiempo, este virus tiene la capacidad de mutar periódicamente o sea cambiar parte de sus características y esto puede ocurrir al intercambiar parte de su material genético entre distintos tipos de virus, hecho que puede ocurrir en animales como el cerdo. Al infectar al ser humano, estos virus “nuevos” tienen la capacidad de producir un mayor número de casos que frente a la gripe habitual, debido a que carecemos de defensas contra este virus de reciente

aparición. Si además de este hecho, el virus es capaz de transmitirse de persona a persona, se está frente a la posibilidad de una pandemia.

### **¿Qué es una pandemia de gripe?**

Una pandemia es la consecuencia de la aparición de un virus de influenza nuevo con diseminación global.

### **¿Qué se conoce de este nuevo virus?**

Si bien se lo ha denominado virus de la gripe porcina porque los estudios iniciales mostraban similitud con el virus que infecta a los cerdos en Estados Unidos, luego se pudo comprobar que consiste en una combinación de virus del cerdo, humano y de las aves.

### **¿Cómo se contagia este virus?**

Hasta el momento se conoce que el virus se transmite en forma similar al virus de la gripe habitual, a través de la tos y al estornudar. También cuando una persona toca una superficie contaminada con secreciones respiratorias y se lleva la mano a la boca o a la nariz.

### **¿Una persona puede infectarse comiendo carne de cerdo?**

El consumir carne de cerdo o derivados cocida o adecuadamente procesada no produce contagio.

### **¿Cuáles son los síntomas de esta enfermedad?**

Son similares a los de la gripe habitual: fiebre alta (superior a 38°C), dolores musculares, dolor de cabeza, tos seca y dolor de garganta.

### **¿Qué debe hacer una persona ante la sospecha de enfermedad?**

La persona que presenta los síntomas antes descriptos, No debe concurrir al colegio o al trabajo en esa situación. También se recomienda mantener los ambientes limpios y ventilados, evitar tomar contacto con otras personas y llamar médico a domicilio si lo tiene al alcance. <sup>15</sup>

---

Guía técnica:

"Guía de práctica clínica para el diagnóstico y tratamiento de influenza A (H1N1)"; (MINSAs 2009).

### **¿Cómo se puede prevenir la infección?**

Las medidas más adecuadas son:

- Cubrir la boca y la nariz con un pañuelo descartable al toser o estornudar
- Lavarse las manos con agua y jabón o con gel alcohólico en forma frecuente, en especial luego de toser o estornudar.
- Evitar tocarse la nariz, ojos o la boca luego de tocar otras superficies.
- Evitar contacto cercano con personas con síntomas de gripe.
- Se recomienda mantener los ambientes limpios y ventilados.

### **¿Sirve la vacuna antigripal que aplicamos habitualmente contra este virus?**

No está demostrado que la vacuna de la gripe que se está utilizando actualmente tenga impacto en la gravedad o prevenga la infección contra este tipo de virus.

### **¿Por qué entonces se debe vacunar con la vacuna habitual?**

Es muy importante que los grupos que tienen indicación de vacuna antigripal estacional reciban la misma, ya que el riesgo de complicaciones con la gripe habitual sigue siendo elevado en las personas pertenecientes a grupos de riesgo (mayores de 65 años, niños y adultos con enfermedades cardíacas y pulmonares crónicas, diabetes, disminución de las defensas, otras enfermedades crónicas renales, hematológicas, hepáticas, embarazadas, niños y adolescentes que reciben aspirina en forma crónica, todos los trabajadores de salud, convivientes de personas que pertenecen a los grupos de riesgo, personal que presta servicios esenciales).

### **¿Existe una vacuna contra este nuevo virus?**

Los virus de la nueva gripe A (H1N1) son antigénicamente muy diferentes de los virus (H1N1) de los seres humanos conocidos hasta ahora, por consiguiente las vacunas de la gripe estacional para las personas no proporcionan protección contra los virus de la nueva gripe A (H1N1).

### **¿Existe tratamiento contra este virus?**

Existe tratamiento antiviral que es efectivo si es administrado dentro de los dos días del comienzo de los síntomas. La indicación de este tratamiento debe ser efectuada por el médico que atiende al paciente de acuerdo a las normas existentes para el mismo. Se dispone de dos antivirales, oseltamivir y zanamivir que son útiles tanto para el tratamiento como así también para la prevención en los casos en que están indicados.



### CAPITULO III

#### APLICACIÓN DEL ESTUDIO DEL VIRUS DE INFLUENZA A (H1N1)

##### 3.1. DISEÑO

El objetivo del presente es identificar los factores de riesgo más significativos mediante el modelo de regresión logística que determina la probabilidad de que todos los pacientes sospechosos o casos probables que luego de la investigación epidemiológica y de laboratorio tienen resultados negativos o con la presencia de virus Influenza A (H1N1).

Se analiza la información correspondiente después de casos reportados de influenza A (H1N1) mediante el llenado de ficha de investigación clínico Epidemiológico.

Las variables que usaremos se definen en la siguiente tabla.

Tabla N° 3.1

##### Definición de Variables de estudio

NOMBRE	UNIDAD O CATEGORIA
EDAD	1 = "Edad menores a 21 años y mayores a 60 años" 0 = "Edad entre 21 y 60 años"
SEXO	1 = "Masculino" 0 = "Femenino"
CONTACTO CON PERSONAS CON SINTOMATOLOGIA EN ULTIMOS 7 DIAS	1 = "SI" 0 = "NO"
CONTACTO CON ANIMALES	1 = "SI" 0 = "NO"
VACUNACION ANTIGRI PAL	1 = "SI" 0 = "NO"
FIEBRE	1 = "SI" 0 = "NO"
TOS	1 = "SI" 0 = "NO"
DOLOR DE GARGANTA	1 = "SI" 0 = "NO"
RINORREA	1 = "SI" 0 = "NO"
EXPECTORACION	1 = "SI" 0 = "NO"

SIBILANCIAS	1 = "SI" 0 = "NO"
CONGESTION FARINGEA	1 = "SI" 0 = "NO"
OTALGIA	1 = "SI" 0 = "NO"
FOTOFOBIA	1 = "SI" 0 = "NO"
CONJESTION CONJUNTIVAL	1 = "SI" 0 = "NO"
VOMITOS	1 = "SI" 0 = "NO"
DOLOR ABDOMINAL	1 = "SI" 0 = "NO"
DIARREA	1 = "SI" 0 = "NO"
ADENOPATIAS	1 = "SI" 0 = "NO"
ASTENIA	1 = "SI" 0 = "NO"
CEFALEA	1 = "SI" 0 = "NO"
MIALGIAS	1 = "SI" 0 = "NO"
MALESTAR GENERAL	1 = "SI" 0 = "NO"
ERUPCION DERMICA	1 = "SI" 0 = "NO"
HOSPITALIZACION	1 = "SI" 0 = "NO"
FALLA MULTI ORGANICA	1 = "SI" 0 = "NO"
DIFICULTAD RESPIRATORIA	1 = "SI" 0 = "NO"
PERDIDA DE CONCIENCIA	1 = "SI" 0 = "NO"
COLAPSO CIRCULATORIO	1 = "SI" 0 = "NO"
MUERTE	1 = "SI"

	0 = "NO"
TRATAMIENTO ANTIBIOTICO	1 = "SI" 0 = "NO"
TRATAMIENTO ANTIVIRAL	1 = "SI" 0 = "NO"
RESULTADO DE LABORATORIO	1 = "Positivo Influenza A H1N1" 0 = "Negativo"

**Población:** Se toma a todos los casos sospechosos que fueron reportados en el Departamento del Cusco, siendo 924 personas que recurrieron a los diferentes establecimientos de salud.

- Personas con **fiebre mayor de 38°C** acompañada de al menos uno de los siguientes signos o síntomas:
  - Tos
  - Dolor de garganta
  - Rinorrea
- Que además:
  - Estuvo en los 7 días previos al inicio de su enfermedad en una zona con casos confirmados de infección por virus de influenza A (H1N1).
  - Tuvo contacto cercano con un caso confirmado de infección con virus de influenza A (H1N1).

**Unidad Estadística:** Serie de reportes de la nueva Influenza A (H1N1), mediante la ficha de Investigación Clínica – Epidemiológica, (Ver anexo 1).

### 3.2. ANALISIS DE LOS DATOS

#### 3.2.1. ANALISIS EXPLORATORIO PREVIO

El examen previo de los datos es un paso necesario, que permite obtener un conocimiento básico de los datos y las relaciones entre las variables. El conocimiento de las interrelaciones de variables puede ayudar enormemente a la especificación y refinamiento del modelo, así como proporcionar una perspectiva razonable para la interpretación de los resultados.

El análisis exploratorio es a menudo menospreciado o ignorado, sin embargo es una herramienta muy útil que nos permite también ver los problemas de los datos, tales como: datos ausentes, valores extremos o discordantes.

**Tabla N° 3.2**

**Casos reportados de influenza A (H1N1) al sistema de vigilancia Epidemiológica según país de procedencia.**

	RESULTADO		Total
	NEGATIVO	POSITIVO INFLUENZA A (H1N1)	
PAIS ARGENTINA	1	1	2
AUSTRALIA	0	2	2
BOLIVIA	0	1	1
BRASIL	0	2	2
CANADA	3	0	3
COLOMBIA	0	1	1
E.E.U.U	3	3	6
ESPAÑA	0	1	1
ISRAEL	1	4	5
ITALIA	1	0	1
JAPON	1	0	1
PERU	476	421	897
REINO UNIDO	1	0	1
TAIWAN	1	0	1
<b>Total</b>	<b>488</b>	<b>436</b>	<b>924</b>

Fuente: DIRESA CUSCO, 2009 (Elaboración propia)

En la tabla N° 3.2 se muestra el resumen de todos los pacientes sospechosos con la nueva influenza A (H1N1) que han sido reportados en el departamento del Cusco de los diferentes países del mundo, resultando positivos 436 de un total de 924 sospechosos.

**Tabla N° 3.3**

**Casos reportados de influenza A (H1N1) al sistema de vigilancia Epidemiológica según sexo y grupos etareos.**

	Edad según grupos etareos					Total
	MENORES A 5 AÑOS	DE 5 A 9 AÑOS	DE 10 A 19 AÑOS	DE 20 A 59 AÑOS	MAYORES A 60 AÑOS	
SEXO F	54	28	63	312	49	506
M	52	18	62	238	48	418
<b>Total</b>	<b>106</b>	<b>46</b>	<b>125</b>	<b>550</b>	<b>97</b>	<b>924</b>

Fuente: DIRESA CUSCO, 2009 (Elaboración propia)

En la tabla N° 3.3 se muestra los resultados según sexo clasificados en grupos etareos, fueron reportados 506 personas de sexo femenino y 418 personas de sexo masculino.

**Tabla N° 3.4**

**Casos reportados de influenza A (H1N1) al sistema de vigilancia Epidemiológica según grupos resultados de Laboratorio.**

		Frecuencia	Porcentaje
Válidos	NEGATIVO	488	52.8
	POSITIVO INFLUENZA A (H1N1)	436	47.2
	Total	924	100.0

Fuente: DIRESA CUSCO, 2009 (Elaboración propia)

La tabla N° 3.4 nos resume que el 52.8% fueron diagnosticados negativos por la nueva Influenza A H1N1 representa el total de todos los casos sospechosos y el 47.2% del total de los casos sospechosos fueron diagnosticados positivos con la nueva Influenza A H1N1.

### 3.3. ANALISIS DE REGRESION LOGISTICA BINARIA CON CADA UNO DE LOS FACTORES DE RIESGO.

Analizaremos cada uno de los Factores que influyen en que una persona sea diagnosticada con virus A (H1N1); la variable independiente si un persona tiene o no el virus cuyo resultado es determinado por Laboratorio como positivo o negativo.

De la tabla 3.4 observamos que casi el 50% fue positivo con Influenza A (H1N1). (Ver anexo 2).

**Tabla de Resumen N° 3.5**

VARIABLE	VALOR DE "P" ASOCIADO AL CONTRASTE	OR	IC AL 95% DE LA OR	
			INFERIOR	SUPERIOR
EDAD	0.651	1.062	0.818	1.380
SEXO	0.409	0.896	0.691	1.162
PAIS DE PROCEDENCIA	0.379	1.413	0.654	3.053
FIEBRE	0.000	3.957	2.674	5.855
TOS	0.000	4.485	2.641	7.619
DOLOR DE GARGANTA	0.010	1.825	1.300	2.562
RINORREA	0.000	2.704	1.947	3.755
EXPECTORACION	0.988	0.998	0.767	1.298
SIBILANCIA	0.000	0.531	0.382	0.737
CONGESTION FARINGEA	0.011	1.444	1.089	1.915
OTALGIA	0.105	1.295	0.947	1.771
FOTOFOBIA	0.894	1.023	0.733	1.427
CONJESTION CONJUNTIVAL	0.208	1.185	0.910	1.542
VOMITOS	0.508	1.104	0.824	1.477
DOLOR ABDOMINAL	0.472	1.108	0.838	1.465
DIARREA	0.324	0.848	0.611	1.177
ADENOPATIAS	0.102	0.701	0.458	1.073
ASTENIA	0.116	1.241	0.948	1.624
CEFALEA	0.000	2.283	1.670	3.122
MIALGIAS	0.001	1.636	1.239	2.161
MALESTAR GENERAL	0.068	1.351	0.978	1.868

ERUPCION DERMICA	0.475	1.214	0.713	2.067
HOSPITALIZACION	0.000	2.880	0.288	0.506
FALLA MULTIORGANICA	0.033	0.253	0.072	0.894
DIFICULTAD RESPIRATORIA	0.000	0.429	0.321	0.575
PERDIDA DE CONCIENCIA	0.010	0.271	0.101	0.730
COLAPSO CIRCULATORIO	0.060	0.338	0.109	1.045
MUERTE	0.060	17.352	2.282	131.910
TRATAMIENTO ANTIBIOTICO	0.941	1.010	0.780	1.303
TRATAMIENTO ANTIVIRAL	0.000	0.543	0.395	0.747

- En la variable Fiebre podemos decir que el valor de la OR es 3.957; lo que significa que una persona que presenta Fiebre tiene aproximadamente un riesgo de 4 veces más que una persona normal de ser diagnosticado con la nueva influenza A (H1N1), además es significativo  $P < 0.05$ , por lo que tenemos razones para tomar a la variable fiebre como un factor de riesgo para la Influenza A (H1N1).

- El resultado nos muestra que de la variable Tos tiene el resultado de OR que es 4.485; lo que significa que una persona que presenta Tos tiene aproximadamente un riesgo de 5 veces más que una persona normal de ser diagnosticado con la nueva influenza A (H1N1), además es significativo  $P < 0.05$ , por lo que tenemos razones para tomar a la variable Tos como un factor de riesgo para la Influenza A (H1N1).

- En la variable Dolor de Garganta podemos decir que el valor de la OR es 1.825; lo que significa que una persona que presenta Dolor de Garganta tiene aproximadamente un riesgo de 2 veces más que una persona normal de ser diagnosticado con la nueva influenza A (H1N1), además es significativo  $P < 0.05$ , por lo que tenemos razones para tomar a la variable Dolor de Garganta como un factor de riesgo para la Influenza A (H1N1).

- En la variable Rinorrea podemos decir que el valor de la OR es 2.704; lo que significa que una persona que presenta Dolor de Garganta tiene aproximadamente un riesgo de 3 veces más que una persona normal de ser diagnosticado con la nueva influenza A (H1N1), además es significativo  $P < 0.05$ , por lo que tenemos razones para tomar a la variable Rinorrea como un factor de riesgo para la Influenza A (H1N1).

- En la variable Expectorcación el valor de la OR es 0.998; lo que significa que una persona que presenta Expectorcación tiene aproximadamente un riesgo de 1 vez mayor

que una persona normal de ser diagnosticado con la nueva influenza A (H1N1), además no es significativo  $P>0.05$  por lo que tenemos razones para excluir a esta variable del modelo.

- En la variable Sibilancias podemos decir que el valor de la OR es 0.531; lo que significa que una persona que presenta Dolor de Sibilancias tiene aproximadamente un riesgo de 0.531 veces mayor que una persona normal de ser diagnosticado con la nueva influenza A (H1N1), además es significativo  $P<0.05$ , por lo que tenemos razones para tomar a la variable Sibilancias como un factor de riesgo para la Influenza (H1N1).

- En la variable Congestión faríngea podemos decir que el valor de la OR es 1.444; lo que significa que una persona que presenta Congestión faríngea tiene aproximadamente un riesgo de 1.5 veces mayor que una persona normal de ser diagnosticado con la nueva influenza A (H1N1), además es significativo  $P<0.05$ , por lo que tenemos razones para tomar a la variable Congestión faríngea como un factor de riesgo para la Influenza A (H1N1).

- En la variable Diarrea el valor de la OR es 0.848; lo que significa que una persona que presenta Diarrea tiene un riesgo de 0.848 veces mayor que una persona normal de ser diagnosticado con la nueva influenza A (H1N1), además no es significativo  $P>0.05$  por lo que tenemos razones suficientes para excluir a esta variable del modelo.

Seguidamente se construirá un modelo de regresión logística multivariante con las variables que son factores de riesgo para la Influenza A (H1N1); fiebre, tos, dolor de garganta, rinorrea, sibilancia, congestión faríngea, cefalea, mialgias, hospitalización, falla multiorganica, dificultad respiratoria, perdida de conciencia, tratamiento antiviral.

**Tabla N° 3.6**  
**Regresión Logística Múltiple**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a)								
FIEBRE	1.208	.219	30.489	1	.000	3.346	2.179	5.137
TOS	1.182	.298	15.741	1	.000	3.262	1.819	5.851
D_GARGANTA	-.066	.208	.101	1	.750	.936	.623	1.407
RINORREA	.479	.192	6.210	1	.013	1.614	1.108	2.352
SIBILANCIAS	-.418	.195	4.591	1	.032	.658	.449	.965
CONGESTION	.013	.171	.006	1	.939	1.013	.724	1.417
CEFALEA	.447	.202	4.878	1	.027	1.564	1.052	2.325
MIALGIAS	.137	.179	.587	1	.444	1.147	.807	1.631
HOSP	-.605	.208	8.409	1	.004	.546	.363	.822
FALLA_MULTILTI	-.500	.722	.481	1	.488	.606	.147	2.494
D_RESP	-.275	.211	1.687	1	.194	.760	.502	1.150
PER_CONC	-.455	.580	.615	1	.433	.634	.204	1.978
ANTIVIRALES	-.455	.185	6.041	1	.014	.634	.441	.912
Constante	-2.470	.355	48.339	1	.000	.085		

a Variable(s) introducida(s) en el paso 1: FIEBRE, TOS, D\_GARGANTA, RINORREA, SIBILANCIAS, CONGESTION, CEFALEA, MIALGIAS, HOSP, FALLA\_MULTILTI, D\_RESP, PER\_CONC, ANTIVIRALES.

La mayoría de las covariables incluidas en el modelo mantienen la significación estadística ( $p < 0,05$ ) en el contraste de hipótesis que las relaciona con la variable dependiente RESULTADO DE LABORATORIO, salvo las variables (dolor de garganta, congestión faríngea, mialgias, falla multiorganica dificultad respiratoria y pérdida de conciencia) que arroja un valor " $p > 0,05$ ". Un procedimiento automático hacia atrás habría excluido dichas variables del modelo final.

Sobre los valores de OR obtenidos podemos comprobar que son similares a los calculados en el análisis bivariante, aunque algo menores para las variables FIEBRE (pasa de 3.957 a 3.346), TOS (pasa de 4.485 a 3.262), DOLOR DE GARGANTA (pasa de 1.825 a 0.936), así sucesivamente y algo mayores para las variables SIBILANCIA (pasa de 0.531 a 0.658), CEFALEA (pasa de 2.283 a 1.564) PERDIDA DE CONCIENCIA (pasa de 0.271 a 0.634) y TRATAMIENTO ANTIVIRAL (pasa de 0.543 a 0.634).

Estos valores de OR están ajustados para cada variable y representan una estimación de su fuerza de asociación con RESULTADO DE LABORATORIO controladas todas las demás variables incluidas en el modelo.



Sin embargo su cálculo conlleva una mayor imprecisión en las estimaciones, como puede comprobarse viendo cómo los intervalos de confianza de las OR se han “separado” y los errores estándar de los coeficientes de regresión han aumentado.

Con estas trece variables, el modelo tiene una capacidad de clasificar correctamente al 52.8% de los casos analizados, como puede verse en la tabla siguiente.

**Tabla N° 3.7**  
**Tabla de clasificación(a,b)**

Observado		Pronosticado			
		RESULTADO		Porcentaje correcto	
		NEGATIVO	POSITIVO INFLUENZA A H1N1		
Paso 0	RESULTADO	NEGATIVO	488	0	100.0
		POSITIVO INFLUENZA A H1N1	436	0	.0
Porcentaje global					52.8

a En el modelo se incluye una constante.

b El valor de corte es .500

Y la proporción de la variabilidad de RESULTADO DE LABORATORIO que es explicado por este modelo no es demasiado buena (entre un 16.6% -ver R cuadrado de Cox y Snell- y un 22.2% -según el R cuadrado de Nagelkerke-), esto es, sigue existiendo un porcentaje importante de “influencia” sobre el hecho de tener un riesgo de ser diagnosticado con el virus de la nueva influenza A (H1N1) que no depende de las variables analizadas:

**Resumen de los modelos**

Paso	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	1109.996(a)	.166	.222

a La estimación ha finalizado en el número de iteración 5 porque las estimaciones de los parámetros han cambiado en menos de .001.

**UN MODELO DE PREDICCIÓN** como conclusión en este estudio, puede elaborarse la ecuación de regresión logística siguiente:

**Logit (p) = -0.247 + 1.280\*(FIEBRE) + 1.182\*(TOS) - 0.066\*(DOLOR DE GARGANTA) + 0.479\*(RINORREA) - 0.418\*(SIBILANCIA) + 0.013\*(CONGESTION FARINGEA) + 0.447\*(CEFALEA) + 0.137\*(MIALGIA) - 0.605\*(HOSPITALIZACION) - 0.500\*(FALLA MULTIORGANICA) - 2.75\*(DIFICULTAD RESPIRATORIA) - 0.455\*(PERDIDA DE CONCIENCIA) - 0.455\*(TRATAMIENTO ANTIVIRAL).**

Siendo  $\text{Logit}(p) = \ln(p / (1-p)) = \ln(\text{odds})$

Y donde  $P = P(Y = 1)$ , esto es, la probabilidad de que un individuo tenga la característica evaluada; en nuestro ejercicio que sea “Diagnosticada Positivo con el virus de la influenza A (H1N1)”, puesto que el valor “1” se asigna en la base de datos a los que resultan positivos con la nueva Influenza A (H1N1).

O alternativamente:

$$P(\text{POSITIVO}) = \frac{1}{1 + e^{(-0.247 + 1.280 * (\text{FIEBRE}) + 1.182 * (\text{TOS}) + \dots - 0.455 * (\text{TRATAMIENTO ANTIVIRAL}))}}$$

### 3.4. Aplicando el método por pasos hacia atrás (Wald)

Tabla N° 3.8

Pruebas omnibus sobre los coeficientes del modelo

		Chi-cuadrado	gl	Sig.
Paso 1	Paso	240.624	30	.000
	Bloque	240.624	30	.000
	Modelo	240.624	30	.000
Paso 2(a)	Paso	-3.211	1	.073
	Bloque	237.412	29	.000
	Modelo	237.412	29	.000
Paso 3(a)	Paso	-7.208	1	.007
	Bloque	230.204	28	.000
	Modelo	230.204	28	.000
Paso 4(a)	Paso	-.001	1	.976
	Bloque	230.203	27	.000
	Modelo	230.203	27	.000
Paso 5(a)	Paso	-.011	1	.917
	Bloque	230.192	26	.000
	Modelo	230.192	26	.000
Paso 6(a)	Paso	-.016	1	.898
	Bloque	230.176	25	.000
	Modelo	230.176	25	.000
Paso 7(a)	Paso	-.027	1	.869
	Bloque	230.149	24	.000
	Modelo	230.149	24	.000
Paso 8(a)	Paso	-.031	1	.861
	Bloque	230.118	23	.000
	Modelo	230.118	23	.000
Paso 9(a)	Paso	-.125	1	.724
	Bloque	229.994	22	.000
	Modelo	229.994	22	.000

Paso 10(a)	Paso	-.440	1	.507
	Bloque	229.554	21	.000
	Modelo	229.554	21	.000
Paso 11(a)	Paso	-.456	1	.500
	Bloque	229.098	20	.000
	Modelo	229.098	20	.000
Paso 12(a)	Paso	-.626	1	.429
	Bloque	228.472	19	.000
	Modelo	228.472	19	.000
Paso 13(a)	Paso	-1.014	1	.314
	Bloque	227.458	18	.000
	Modelo	227.458	18	.000
Paso 14(a)	Paso	-1.135	1	.287
	Bloque	226.323	17	.000
	Modelo	226.323	17	.000
Paso 15(a)	Paso	-1.441	1	.230
	Bloque	224.883	16	.000
	Modelo	224.883	16	.000
Paso 16(a)	Paso	-.852	1	.356
	Bloque	224.031	15	.000
	Modelo	224.031	15	.000
Paso 17(a)	Paso	-1.541	1	.214
	Bloque	222.490	14	.000
	Modelo	222.490	14	.000
Paso 18(a)	Paso	-2.070	1	.150
	Bloque	220.420	13	.000
	Modelo	220.420	13	.000
Paso 19(a)	Paso	-2.134	1	.144
	Bloque	218.286	12	.000
	Modelo	218.286	12	.000
Paso 20(a)	Paso	-2.337	1	.126
	Bloque	215.949	11	.000
	Modelo	215.949	11	.000
Paso 21(a)	Paso	-3.134	1	.077
	Bloque	212.815	10	.000
	Modelo	212.815	10	.000
Paso 22	Paso	10.689	1	.001
	Bloque	223.504	11	.000
	Modelo	223.504	11	.000

a Un valor de chi-cuadrado negativo indica que ha disminuido el valor de chi-cuadrado con respecto al paso anterior.

El proceso iterativo termina en el veintidós pasos, con un modelo que es significativamente mejor que el modelo con todas las variables (véase como han ido disminuyendo los grados de libertad del contraste Chi cuadrado de la Prueba Omnibus sobre el modelo, desde 30 en el primer paso (31 coeficientes, el de la constante y el de las 30 variables iniciales) hasta 12 en el 22º paso (se habrá quedado con 11 variables más la constante en la ecuación final de regresión). Seguidamente nos aporta información sobre los ajustes de los diferentes modelos que ha ido generando en cada paso, tanto a través del estadístico  $[-2 \log \text{verosimilitud}]$  como el R cuadrado.

**Tabla N° 3.9**

**Resumen de los modelos**

Paso	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	1037.385(a)	.229	.306
2	1040.596(a)	.227	.302
3	1047.804(b)	.221	.294
4	1047.805(b)	.221	.294
5	1047.816(b)	.221	.294
6	1047.832(b)	.221	.294
7	1047.859(b)	.220	.294
8	1047.890(b)	.220	.294
9	1048.014(b)	.220	.294
10	1048.454(b)	.220	.294
11	1048.910(b)	.220	.293
12	1049.536(b)	.219	.292
13	1050.550(b)	.218	.291
14	1051.685(b)	.217	.290
15	1053.125(b)	.216	.288
16	1053.977(b)	.215	.287
17	1055.518(b)	.214	.286
18	1057.588(b)	.212	.283
19	1059.722(b)	.210	.281
20	1062.059(b)	.208	.278
21	1065.193(b)	.206	.275
22	1054.504(a)	.215	.287

a La estimación ha finalizado en el número de iteración 20 por que se han alcanzado las iteraciones máximas. No se puede encontrar una solución definitiva.

b La estimación ha finalizado en el número de iteración 6 porque las estimaciones de los parámetros han cambiado en menos de .001.

Luego nos muestra un resumen de las tablas de clasificación que se han ido obteniendo con los valores predichos por el modelo en cada paso frente a los valores realmente observados. La mejoría es escasa en el % de predicción correcto:

**Tabla 3.10**  
**Tabla de clasificación global**

Observado	Pronosticado				
	RESULTADO		NEGATIVO	POSITIVO INFLUENZA A H1N1	Porcentaje correcto
Paso 1	RESULTADO	NEGATIVO	345	143	70.7
		POSITIVO INFLUENZA A H1N1	124	312	71.6
	Porcentaje global				71.1
Paso 2	RESULTADO	NEGATIVO	345	143	70.7
		POSITIVO INFLUENZA A H1N1	124	312	71.6
	Porcentaje global				71.1
Paso 3	RESULTADO	NEGATIVO	345	143	70.7
		POSITIVO INFLUENZA A H1N1	127	309	70.9
	Porcentaje global				70.8
Paso 4	RESULTADO	NEGATIVO	344	144	70.5
		POSITIVO INFLUENZA A H1N1	127	309	70.9
	Porcentaje global				70.7
Paso 5	RESULTADO	NEGATIVO	345	143	70.7
		POSITIVO INFLUENZA A H1N1	126	310	71.1
	Porcentaje global				70.9
Paso 6	RESULTADO	NEGATIVO	345	143	70.7
		POSITIVO INFLUENZA A H1N1	124	312	71.6
	Porcentaje global				71.1
Paso 7	RESULTADO	NEGATIVO	346	142	70.9
		POSITIVO INFLUENZA A H1N1	125	311	71.3
	Porcentaje global				71.1
Paso 8	RESULTADO	NEGATIVO	346	142	70.9
		POSITIVO INFLUENZA A H1N1	128	308	70.6
	Porcentaje global				70.8
Paso 9	RESULTADO	NEGATIVO	347	141	71.1
		POSITIVO INFLUENZA A H1N1	128	308	70.6
	Porcentaje global				70.9
Paso 10	RESULTADO	NEGATIVO	348	140	71.3
		POSITIVO INFLUENZA A H1N1	128	308	70.6
	Porcentaje global				71.0
Paso 11	RESULTADO	NEGATIVO	345	143	70.7
		POSITIVO INFLUENZA A H1N1	128	308	70.6
	Porcentaje global				70.7
Paso 12	RESULTADO	NEGATIVO	348	140	71.3
		POSITIVO INFLUENZA A H1N1	130	306	70.2
	Porcentaje global				70.8
Paso 13	RESULTADO	NEGATIVO	348	140	71.3
		POSITIVO INFLUENZA A H1N1	132	304	69.7
	Porcentaje global				70.6
Paso 15	RESULTADO	NEGATIVO	340	148	69.7
		POSITIVO INFLUENZA	126	310	71.1

		A H1N1			
	Porcentaje global				70.3
Paso 16	RESULTADO	NEGATIVO	341	147	69.9
		POSITIVO INFLUENZA A H1N1	126	310	71.1
	Porcentaje global				70.5
Paso 17	RESULTADO	NEGATIVO	341	147	69.9
		POSITIVO INFLUENZA A H1N1	129	307	70.4
	Porcentaje global				70.1
Paso 18	RESULTADO	NEGATIVO	343	145	70.3
		POSITIVO INFLUENZA A H1N1	132	304	69.7
	Porcentaje global				70.0
Paso 19	RESULTADO	NEGATIVO	340	148	69.7
		POSITIVO INFLUENZA A H1N1	127	309	70.9
	Porcentaje global				70.2
Paso 20	RESULTADO	NEGATIVO	346	142	70.9
		POSITIVO INFLUENZA A H1N1	134	302	69.3
	Porcentaje global				70.1
Paso 21	RESULTADO	NEGATIVO	344	144	70.5
		POSITIVO INFLUENZA A H1N1	133	303	69.5
	Porcentaje global				70.0
Paso 22	RESULTADO	NEGATIVO	346	142	70.9
		POSITIVO INFLUENZA A H1N1	133	303	69.5
	Porcentaje global				70.2

a El valor de corte es .500

Después de todo este análisis notamos que el modelo 14 tiene una tabla de mejor clasificación con un porcentaje global de 71.2% datos correctamente clasificados.

Y por último un resumen de los coeficientes y su significación estadística de las diferentes covariables que ha ido considerando en cada paso, extrayéndo en cada uno la menos signicativa, para volver a ajustar el modelo sin ella. Así puede comprobarse como tras el primer paso elimina las covariables FALLA MULTIORGANICA y COLAPSO CIRCULAR que son la que tienen una “p” asociada al contraste de Wald no significativo y de mayor valor (0,998).

Al final se queda con once variables predictoras, la mayoría de las variables son estadísticamente significativas en el contraste de su coeficiente de regresión (FIEBRE, RINORREA, SIBILANCIAS, ADENOPATIA, CEFALEA, HOPITALIZACION, DIFICULTAD RESPIRATORIA, COLAPSO CIRCILATORIO, MUERTE y TRATAMIENTO ANTIVIRAL), siendo este resultado diferente al obtenido por el procedimiento automático INTRODUCIR.

Así se puede ir construyendo modelos introduciendo y eliminando variables en cada paso, lo que le permitirá evaluar los cambios que se van produciendo en el los coeficientes y sus errores estándar, las modificaciones en el estadístico de ajuste [-2 log

verosimilitud] y en las tablas de clasificación de cada modelo, incluso optar por dejar covariables que, aun no siendo “estadísticamente significativas” en el contraste de su coeficiente, aportan información al modelo o permiten obtener estimaciones ajustadas de otras variables predictoras. Por supuesto también podrá introducir términos de interacción, imposibles en los procedimientos automáticos, como ya hemos comentado.

### 3.5. MEJOR MODELO DEL ESTUDIO.

Por lo tanto dado la finalidad del estudio y los resultados antes mencionadas seleccionamos el modelo 14 como el modelo mas adecuado para este estudio. (Ver anexo 3)

**Tabla N° 3.11**

**Regresión Logística Múltiple las mejores variables en la ecuación (c) en el paso 14**

	B	E.T.	Wald	Gf	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 14(a) Procedencia	.655	.452	2.098	1	.148	1.925	.793	4.671
FIEBRE	1.312	.228	33.033	1	.000	3.714	2.374	5.809
TOS	1.236	.309	16.014	1	.000	3.441	1.879	6.303
RINORREA	.509	.201	6.413	1	.011	1.663	1.122	2.466
SIBILANCIAS	-.446	.204	4.768	1	.029	.640	.429	.955
FOTOFOBIA	-.248	.205	1.474	1	.225	.780	.522	1.165
CONGES_CONJ UN	.199	.166	1.438	1	.230	1.221	.881	1.690
DIARREA	-.277	.190	2.126	1	.145	.758	.523	1.100
ADENOPATIAS	-.930	.246	14.279	1	.000	.395	.244	.639
CEFALEA	.481	.206	5.477	1	.019	1.618	1.081	2.422
MIALGIAS	.232	.185	1.571	1	.210	1.261	.877	1.813
HOSP	-.843	.220	14.753	1	.000	.430	.280	.662
D_RESP	-.340	.220	2.381	1	.123	.712	.462	1.096
PER_CONC	-1.100	.680	2.617	1	.106	.333	.088	1.262
MUERTE	4.936	1.269	15.128	1	.000	139.246	11.574	1675.192
ANTIBIOTICO	.274	.168	2.650	1	.104	1.315	.946	1.828
ANTIVIRALES	-.536	.199	7.274	1	.007	.585	.396	.864
Constante	-2.737	.367	55.659	1	.000	.065		

Como conclusión final en este estudio, puede elaborarse la ecuación de regresión logística siguiente:

$$\text{Logit (p)} = -2.737 + 0.655*(\text{PAIS DE PROCEDENCIA}) + 1.312*(\text{FIEBRE}) + 1.236*(\text{TOS}) + 0.509*(\text{RINORREA}) - 0.446*(\text{SIBILANCIA}) - 0.248*(\text{FOTOFOBIA}) + 0.199*(\text{CONGESTION CONJUNTIVAL}) - 0.277*(\text{DIARREA}) - 0.930*(\text{ADENOPATIAS}) + 0.481*(\text{CEFALEA}) + 0.232*(\text{MIALGIAS}) - 0.843*(\text{HOSPITALIZACION}) - 0.343*(\text{DIFICULTAD RESPIRATORIA}) - 1.100*(\text{PERDIDA DE CONCIENCIA}) + 4.936*(\text{MUERTE}) + 0.274*(\text{TRATAMIENTO ANTIBIOTICO}) - 0.536*(\text{TRATAMIENTO ANTIVIRALES}).$$

### 3.5.1. Interpretación de algunos de los coeficientes del modelo.

$\beta_0 = -2.737$  es la constante del modelo y significa que cuando todas las variables regresoras tengan un valor 0 la razón de probabilidad de que una persona sea diagnosticada con la Influenza A (H1N1) positivo o negativo tomara este valor. Para nuestro modelo este valor no tiene sentido porque no hay una persona negativa.

$\beta_1 = 0.665$  es el coeficiente de país de procedencia el cambio en unidades de la razón de probabilidades de ser diagnosticado positivo o negativo, cuando todas las variables regresoras excepto país de procedencia permanezca invariante. Como es un coeficiente positivo quiere decir que su transformación (antilog) será un valor mayor que 1 por lo que el riesgo aumentará. Por lo tanto con este modelo a valores altos de procedencia la probabilidad de ser diagnosticado positivo aumentará.

$\beta_2 = 1.312$  es el coeficiente de fiebre y significa el cambio en unidades de la razón de probabilidad de que una persona sea diagnosticada con la Influenza A (H1N1) positivo o negativo, cuando todas las variables regresoras excepto fiebre, permanezca invariantes. Como es un coeficiente positivo quiere decir que su transformación (antilog) será un valor mayor que 1, luego el riesgo aumentará. Por lo tanto con este modelo a valores altos de fiebre la probabilidad de ser diagnosticado positivo aumentara.

$\beta_{12} = -0.843$  es el coeficiente de hospitalización y significa el cambio en unidades de la razón de probabilidad de que una persona sea diagnosticada con la Influenza A (H1N1) positivo o negativo, cuando todas las variables regresoras excepto hospitalización, permanezcan invariantes. Como es un coeficiente negativo quiere decir que su transformación (antilog) será un valor menor que 1 por lo tanto el riesgo disminuirá, por lo tanto con este modelo a valores altos de ser hospitalizado la probabilidad de incidencia disminuirá.



$\beta_{17} = -0.536$  es el coeficiente de la variable tratamiento antiviral que significa si el paciente recibe tratamiento antiviral o no y significa el cambio en unidades de la razón de la probabilidad de que una persona sea diagnosticada con la Influenza A (H1N1) positivo o negativo, cuando todas las variables regresoras excepto la variable tratamiento antiviral permanezca invariante como las variables tratamiento antiviral son variables categóricas que solo toman los valores 0 y 1, la contribución al Odds del coeficiente será de 1 y la dummy toma el valor 0 y de  $e^{-0.536} = 0.5851$ .

## **RECOMENDACIONES**

1. Se le recomienda a la DIRESA Cusco la incorporación de profesionales estadísticos en sus diferentes áreas para su mejor análisis y a fin de mejorar la calidad de la información que brindan.
2. Se recomienda a la DIRESA Cusco, que aplique el modelo de Regresión Logística, y actualice en forma permanente estos porcentajes de riesgo.
3. Implementar la Biblioteca Especializada con libros relacionados a tópicos de Regresión Logística.
4. Es necesario buscar modelos como el del presente estudio para otros tipos de base de datos. Debe generarse un modelo diferente por cada base de datos puesto que ese modelo garantiza la precisión de la estimación.
5. Debido a la existencia y proliferación de softwares piratas, recomendamos a la Universidad la campaña de software estadístico con licencia para que los alumnos y tesisistas tengamos acceso a programas estadísticos originales que según sabemos son más completos que los programas piratas.

## CONCLUSIONES Y RECOMENDACIONES

### CONCLUSIONES

1. Se determinó el mejor modelo de la ecuación con los siguientes variables: “País de Procedencia, Fiebre, Tos, Rinorrea, Sibilancia, Fotofobia, Congestión Conjuntival, Diarrea, Adenopatías, Cefalea, Mialgias, Hospitalización, Dificultad Respiratoria, Pérdida de Conciencia, Muerte, Tratamiento Antibiótico, Tratamiento Antiviral”. Para que una persona sea diagnosticada con virus A (H1N1) mediante el modelo de Regresión Logística, con un porcentaje de clasificación correcta global de 71.2%.
2. Se utilizó el Método Por pasos hacia atrás (Wald), tomando en cuenta todas las variables y el resultado arrojó que las variables más significativas son: “Fiebre, Tos, Rinorrea, Sibilancias, Adenopatías, Cefalea, Hospitalización, Muerte, Tratamiento Antiviral”.
3. El ODDS Ratio para la variable Fiebre es 3.714, para la variable Tos es igual a 3.441, y Rinorrea es 1.663 (ver tabla N° 3.11), que son mayores a 1 por lo tanto son factores que incrementan la probabilidad de reportar como Positivo al paciente sospechoso.
4. Las variables como, sibilancias, Fotofobia presentan un ODDS Ratio igual a 0.640 y 0,780 menor que 1, por lo que son factores que disminuye la probabilidad de ser reportado Positivo pacientes sospechosos por la Influenza A (H1N1)
5. Por lo mencionado anteriormente el modelo de Regresión Logística es el más adecuado en la aplicación en Epidemiología.

## BIBLIOGRAFIA

- Álvarez Cáceres Rafael, (1995).** *Estadística Multivariante y no paramétrica con SPSS Aplicación a las ciencias de la salud*, Ediciones Díaz de Santos, S.A. Juan Bravo, 3A. 28006 Madrid España.
- Acuña Fernández Edgar, (2007).** *Análisis de Regresión*, Departamento de Matemáticas Universidad de Puerto Rico Reciento Universitario de Mayaguez.
- Cabero Moran Teresa y Martín Martín Quintín, (2008).** *Tratamiento estadístico de datos SPSS*. España, Thomson Editores Spain.
- Cornfield J, Gordon T, Smit Wn, (1961).** Quantal response curves for experimentally uncontrolled variables. Bulletin of the International Statistical Institute.
- Cordeiro, Gauss M. (1992)** Introducción a la teoría de la verosimilitud.
- Florez Manrique Luz, (2001).** *Tesis pre grado UNMSM*. Titulado “Análisis estadístico de los factores de riesgo que influyen en la enfermedad antigua de pecho”. Lima – Peru.
- Garson D. (2006).** *Logistic regression*. Disponible en:  
<http://www2.chass.ncsu.edu/garson/PA765/logistic.htm>. [Accesado el 15 de mayo de 2007].
- Hosmer & Lemeshow, (2000).** *Applied Logistic Regression* Wiley-Interscience.
- Kleinbaum G. David Mitchel Klein, (2010).** *Logistic Regression, Statistics for Biology and Health Third Edition USA*. Springer Science+Business Media, LLC.
- Montero Mercade Lidia (21 Setembre de 2.002)** Departament d'Estadística i Investigació Operativa Versió.
- Montgomery Douglas C. (2004)** *Introducción al Análisis de Regresión Lineal*. Compañía Editora Continental. México.
- McCullagh, Peter, Nelder, John, (1989).** *Generalized Linear Models*.
- Nelder and Wedderburn, (1972).** JRSS (A) paper is the primary source.
- Salinas M, Silva C. (2007).** *Modelos de regresión y correlación II. Regresión Lineal Múltiple*. Cienc Trab 23: 39 – 41.
- Salcedo Poma Celia Mercedes, (2002).** *Tesis pre grado UNMSM*. Titulado “Estimación de la ocurrencia de incidencias en declaraciones de pólizas de importación” Lima – Perú.
- Taucher E. 1997.** *Regresión Logística*. En: Bioestadística, Santiago: Editorial Universitaria, (25): 241-248.

**Truett J, Cornfield J, Kannel W. (1967).** Multivariate analysis of the risk of coronary heart disease in Framingham. *Biometrika*.

**Walter Sh, Duncan Db. (1967).** Estimación of the probability of as function of several independent variables. *Biometrica*.

**Guía Técnica:** “Guía de práctica clínica para el diagnóstico y tratamiento de influenza A (H1N1)”; (MINSa 2009).

[Http://sisbib.unmsm.edu.pe/bibvirtualdata/tesis/baisc/salcado\\_pc/enpdf/cap2.pdf](http://sisbib.unmsm.edu.pe/bibvirtualdata/tesis/baisc/salcado_pc/enpdf/cap2.pdf)-

<http://support.spss.com/Products.Ext/spss/docuemntation/spssforwindws/spanish/spss%20regresssion%2012.0.pdf>-

[http://sisbib.unmsm.edu.pe/bibvirtualdata/tesis/baisc/flores\\_ml/T\\_completo.PDF](http://sisbib.unmsm.edu.pe/bibvirtualdata/tesis/baisc/flores_ml/T_completo.PDF)

## **ANEXOS**

## ANEXO 1

Ficha de investigación clínico epidemiológico para tomar muestras a los pacientes sospechosos de la nueva Influenza A H1N1.

**Ministerio de Salud del Perú**  
 Ficha de investigación clínico epidemiológica  
 de Vigilancia de influenza y otros virus respiratorios (OVR), Infección Respiratoria Aguda Grave (IRAG),  
 IRAG inusitada, fallecimiento por IRAG

<b>ESTABLECIMIENTO NOTIFICANTE</b>			
Establecimiento de Salud: _____		DIRESA / DISA: _____	
Nombre del médico tratante: _____		Correo: _____	
Tipo de Vigilancia <input type="checkbox"/> Vigilancia de Influenza y OVR <input type="checkbox"/> IRAG <input type="checkbox"/> IRAG inusitada <input type="checkbox"/> Fallecido por IRAG			
<b>IDENTIFICACION DEL PACIENTE</b>			
Apellido paterno: _____		Apellido materno: _____	Nombres: _____
Fecha de nacimiento (dd/mm/aa): _____		Edad: _____	Años: _____ Meses: _____
Sexo: Masculino: <input type="checkbox"/> Femenino: <input type="checkbox"/>		DNI: _____	Ocupación: _____
Dirección y/o domicilio: _____		Teléfono: _____	
Departamento: _____		Provincia: _____	Distrito: _____
<b>ANTECEDENTES</b>			
Contacto con otras personas con sintomatología respiratoria en últimos 7 días <input type="checkbox"/> Sí <input type="checkbox"/> No			
Contacto con animales <input type="checkbox"/> Sí <input type="checkbox"/> No ¿Cuáles? _____			
País/lugar que visitó últimos 15 días (Lugar, fecha ingreso y salida): _____			
Vacunación contra influenza: <input type="checkbox"/> Sí <input type="checkbox"/> No Fecha de Vacunación (dd/mm/aa) _____			
<b>CLÍNICA</b>			
Fecha de inicio de síntomas: _____			
Fiebre:	<input type="checkbox"/> Sí <input type="checkbox"/> No	Medida con termómetro:	<input type="checkbox"/> Sí <input type="checkbox"/> No
Tos:	<input type="checkbox"/> Sí <input type="checkbox"/> No	Cefalea:	<input type="checkbox"/> Sí <input type="checkbox"/> No
Dolor garganta:	<input type="checkbox"/> Sí <input type="checkbox"/> No	Fotofobia:	<input type="checkbox"/> Sí <input type="checkbox"/> No
Rinorrea:	<input type="checkbox"/> Sí <input type="checkbox"/> No	Congestión conjuntival:	<input type="checkbox"/> Sí <input type="checkbox"/> No
Expectoración:	<input type="checkbox"/> Sí <input type="checkbox"/> No	Vómitos:	<input type="checkbox"/> Sí <input type="checkbox"/> No
Sibilancias:	<input type="checkbox"/> Sí <input type="checkbox"/> No	Dolor abdominal:	<input type="checkbox"/> Sí <input type="checkbox"/> No
Congestión faríngea:	<input type="checkbox"/> Sí <input type="checkbox"/> No	Diarrea:	<input type="checkbox"/> Sí <input type="checkbox"/> No
Cianosis:	<input type="checkbox"/> Sí <input type="checkbox"/> No	Taquipnea:	<input type="checkbox"/> Sí <input type="checkbox"/> No
Temperatura Máxima (°C):	_____	Adenopatías:	<input type="checkbox"/> Sí <input type="checkbox"/> No
Asma:	<input type="checkbox"/> Sí <input type="checkbox"/> No	Cefalea:	<input type="checkbox"/> Sí <input type="checkbox"/> No
Mialgias:	<input type="checkbox"/> Sí <input type="checkbox"/> No	Mal estar general:	<input type="checkbox"/> Sí <input type="checkbox"/> No
Erupción dérmica:	<input type="checkbox"/> Sí <input type="checkbox"/> No	Hipoxia:	<input type="checkbox"/> Sí <input type="checkbox"/> No
Otras manifestaciones: _____			
<b>Criterios de gravedad (compromiso sistémico)</b>			
Hospitalización:	<input type="checkbox"/> Sí <input type="checkbox"/> No	Dificultad respiratoria:	<input type="checkbox"/> Sí <input type="checkbox"/> No
Falla multi orgánica:	<input type="checkbox"/> Sí <input type="checkbox"/> No	Período de conciencia:	<input type="checkbox"/> Sí <input type="checkbox"/> No
Colapso circulatorio:	<input type="checkbox"/> Sí <input type="checkbox"/> No	Muerte:	<input type="checkbox"/> Sí <input type="checkbox"/> No
Diagnóstico presuntivo: _____			
<b>Laboratorio</b>			
Tipo de muestra	Fecha de obtención (dd/mm/aa)	Tipo de Prueba	Resultado
Hisopado nasal	_____	_____	_____
Hisopado faríngeo:	_____	_____	_____
Aspirado nasolaringeo	_____	_____	_____
Suero de fase aguda.	_____	_____	_____
Suero de fase convalescente:	_____	_____	_____
Otros: _____	_____	_____	_____
<b>Tratamiento</b>			
Antibióticos:	<input type="checkbox"/> Sí <input type="checkbox"/> No	¿Cuáles? _____	Antivirales: <input type="checkbox"/> Sí <input type="checkbox"/> No
		¿Cuáles? _____	
<b>Para casos hospitalizados por IRAG/IRAG inusitadas y fallecidos por IRAG</b>			
Fecha de Hospitalización: _____		Fecha de Alta: _____	
Co-mortalidad 1. _____		2. _____	
3. _____		Causa de la defunción: 1. _____	
Fecha de defunción: _____		2. _____	
Fecha de Censado de la ficha: _____		3. _____	
Fecha de notificación: _____			

Nombre y firma de la persona que llena la ficha: \_\_\_\_\_

## ANEXO 2

**SALIDA DEL SPSS: Modelo de Regresión Logística binaria con cada una de las variables, Tamaño de muestra 924 Aplicando el Método Introducir.**

**Variables en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) EDAD_1	.060	.134	.205	1	.651	1.062	.818	1.380
Constante	-.138	.087	2.545	1	.111	.871		

a Variable(s) introducida(s) en el paso 1: EDAD\_1.

**Variables en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) SEXO	-.109	.132	.682	1	.409	.896	.691	1.162
Constante	-.063	.089	.506	1	.477	.939		

a Variable(s) introducida(s) en el paso 1: SEXO.

**Variables en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) Procedencia	.346	.393	.775	1	.379	1.413	.654	3.053
Constante	-.123	.067	3.368	1	.066	.884		

a Variable(s) introducida(s) en el paso 1: Procedencia.

**Variables en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) FIEBRE	1.376	.200	47.338	1	.000	3.957	2.674	5.855
Constante	-1.264	.186	46.116	1	.000	.282		

a Variable(s) introducida(s) en el paso 1: FIEBRE.

**Variables en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) TOS	1.501	.270	30.835	1	.000	4.485	2.641	7.619
Constante	-1.479	.261	32.071	1	.000	.228		

a Variable(s) introducida(s) en el paso 1: TOS.



**Variables en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) D_GARGANTA	.602	.173	12.100	1	.001	1.825	1.300	2.562
Constante	-.602	.157	14.741	1	.000	.548		

a Variable(s) introducida(s) en el paso 1: D\_GARGANTA.

**Variables en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) RINORREA	.995	.168	35.246	1	.000	2.704	1.947	3.755
Constante	-.887	.150	35.134	1	.000	.412		

a Variable(s) introducida(s) en el paso 1: RINORREA.

**Variables en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) EXPECTORACION	-.002	.134	.000	1	.988	.998	.767	1.298
Constante	-.112	.103	1.169	1	.280	.894		

a Variable(s) introducida(s) en el paso 1: EXPECTORACION.

**Variables en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) SIBILANCIAS	-.633	.168	14.258	1	.000	.531	.382	.737
Constante	.016	.074	.049	1	.824	1.017		

a Variable(s) introducida(s) en el paso 1: SIBILANCIAS.

**Variables en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) CONGESTION	.368	.144	6.522	1	.011	1.444	1.089	1.915
Constante	-.368	.120	9.349	1	.002	.692		

a Variable(s) introducida(s) en el paso 1: CONGESTION.

**Variabes en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) OTALGIA	.259	.160	2.624	1	.105	1.295	.947	1.771
Constante	-.169	.075	5.134	1	.023	.844		

a Variable(s) introducida(s) en el paso 1: OTALGIA.

**Variabes en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) FOTOFOBIA	.023	.170	.018	1	.894	1.023	.733	1.427
Constante	-.117	.073	2.565	1	.109	.890		

a Variable(s) introducida(s) en el paso 1: FOTOFOBIA.

**Variabes en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) CONGES_CONJUN	.169	.135	1.585	1	.208	1.185	.910	1.542
Constante	-.180	.085	4.484	1	.034	.835		

a Variable(s) introducida(s) en el paso 1: CONGES\_CONJUN.

**Variabes en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) VOMITOS	.099	.149	.439	1	.508	1.104	.824	1.477
Constante	-.139	.077	3.258	1	.071	.870		

a Variable(s) introducida(s) en el paso 1: VOMITOS.

**Variabes en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) DOLOR_AD	.102	.142	.518	1	.472	1.108	.838	1.465
Constante	-.144	.079	3.311	1	.069	.865		

a Variable(s) introducida(s) en el paso 1: DOLOR\_AD.

**Variables en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) DIARREA	-.165	.167	.974	1	.324	.848	.611	1.177
Constante	-.081	.073	1.209	1	.272	.922		

a Variable(s) introducida(s) en el paso 1: DIARREA

**Variables en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) ADENOPATIAS	-.356	.217	2.681	1	.102	.701	.458	1.073
Constante	-.075	.070	1.164	1	.281	.928		

a Variable(s) introducida(s) en el paso 1: ADENOPATIAS.

**Variables en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) ASTENIA	.216	.137	2.468	1	.116	1.241	.948	1.624
Constante	-.250	.110	5.192	1	.023	.779		

a Variable(s) introducida(s) en el paso 1: ASTENIA.

**Variables en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) CEFALEA	.826	.160	26.731	1	.000	2.283	1.670	3.122
Constante	-.739	.140	27.700	1	.000	.478		

a Variable(s) introducida(s) en el paso 1: CEFALEA.

**Variables en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) MIALGIAS	.492	.142	12.015	1	.001	1.636	1.239	2.161
Constante	-.444	.117	14.385	1	.000	.642		

a Variable(s) introducida(s) en el paso 1: MIALGIAS.

**Variables en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) M_GENERA	.301	.165	3.324	1	.068	1.351	.978	1.868
Constante	-.353	.148	5.703	1	.017	.703		

a Variable(s) introducida(s) en el paso 1: M\_GENERAL.

**Variables en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) ERUPC_DERMI	.194	.271	.510	1	.475	1.214	.713	2.067
Constante	-.125	.068	3.363	1	.067	.883		

a Variable(s) introducida(s) en el paso 1: ERUPC\_DERMICA.

**Variables en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) HOSP	-.963	.144	44.761	1	.000	.382	.288	.506
Constante	.219	.082	7.072	1	.008	1.245		

a Variable(s) introducida(s) en el paso 1: HOSP.

**Variables en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) FALLA_MUL	-1.374	.644	4.551	1	.033	.253	.072	.894
Constante	-.093	.066	1.941	1	.164	.912		

a Variable(s) introducida(s) en el paso 1: FALLA\_MULT.

**Variables en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) D_RESP	-.845	.148	32.399	1	.000	.429	.321	.575
Constante	.142	.079	3.174	1	.075	1.152		

a Variable(s) introducida(s) en el paso 1: D\_RESP.

**Variables en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) PER_CON C	-1.304	.504	6.682	1	.010	.271	.101	.730
Constante	-.082	.067	1.522	1	.217	.921		

a Variable(s) introducida(s) en el paso 1: PER\_CONC.

**Variables en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) COLAPSO_CIR	-1.084	.576	3.545	1	.060	.338	.109	1.045
Constante	-.095	.066	2.037	1	.154	.909		

a Variable(s) introducida(s) en el paso 1: COLAPSO\_CIR.

**Variables en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) MUERTE	2.854	1.035	7.603	1	.006	17.352	2.282	131.910
Constante	-.146	.067	4.789	1	.029	.864		

a Variable(s) introducida(s) en el paso 1: MUERTE.

**Variables en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) ANTIBIOTICO	.010	.132	.006	1	.941	1.010	.780	1.308
Constante	-.118	.095	1.528	1	.216	.889		

a Variable(s) introducida(s) en el paso 1: ANTIBIOTICO.

**Variables en la ecuación**

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
							Inferior	Superior
Paso 1(a) ANTIVIRALES	-.610	.162	14.134	1	.000	.543	.395	.747
Constante	.022	.075	.090	1	.765	1.023		

a Variable(s) introducida(s) en el paso 1: ANTIVIRALES.  
**RESULTADOS DE REGRESION LOGISTICA MULTINOMIAL**

### ANEXO 3

**SALIDA DEL SPSS: Modelo de Regresión Logística todas las variables. Tamaño de muestra 924. Aplicando el Método por pasos hacia atrás (Wald).**

**Variables en la ecuación(c)**

		B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95.0% para EXP(B)	
		Inferior	Superior	Inferior	Superior	Inferior	Superior	Inferior	Superior
Paso 1(a)	EDAD_1	-.004	.161	.001	1	.979	.996	.726	1.366
	SEXO	-.059	.155	.144	1	.704	.943	.697	1.277
	Procedencia	.642	.457	1.976	1	.160	1.901	.776	4.656
	FIEBRE	1.280	.232	30.467	1	.000	3.596	2.283	5.665
	TOS	1.321	.319	17.121	1	.000	3.748	2.004	7.008
	D_GARGANTA	-.154	.218	.498	1	.480	.857	.559	1.314
	RINORREA	.531	.209	6.447	1	.011	1.701	1.129	2.563
	EXPECTORACION	-.145	.175	.681	1	.409	.865	.614	1.220
	SIBILANCIAS	-.409	.209	3.806	1	.051	.665	.441	1.002
	CONGESTION	.097	.183	.277	1	.598	1.101	.769	1.578
	OTALGIA	.028	.193	.021	1	.884	1.029	.705	1.501
	FOTOFOBIA	-.259	.212	1.490	1	.222	.772	.510	1.170
	CONGES_CONJUN	.155	.170	.835	1	.361	1.168	.837	1.629
	VOMITOS	.155	.185	.697	1	.404	1.167	.812	1.679
	DOLOR_AD	-.031	.184	.027	1	.868	.970	.676	1.392
	DIARREA	-.300	.208	2.074	1	.150	.741	.493	1.114
	ADENOPATIAS	-.908	.251	13.056	1	.000	.403	.246	.660
	ASTENIA	.156	.178	.764	1	.382	1.169	.824	1.657
	CEFALEA	.489	.221	4.894	1	.027	1.631	1.057	2.517
	MIALGIAS	.196	.195	1.008	1	.315	1.216	.830	1.781
	M_GENERAL	-.054	.222	.059	1	.808	.948	.613	1.465
	ERUPC_DERMICA	.112	.323	.120	1	.729	1.118	.594	2.107
	HOSP	-.891	.227	15.417	1	.000	.410	.263	.640
	FALLA_MULTIPLE	-18.319	6626.174	.000	1	.998	.000	.000	.
	D_RESP	-.293	.222	1.735	1	.188	.746	.483	1.154
	PER_CONC	-.694	.711	.953	1	.329	.500	.124	2.012
	COLAPSO_CIR	-18.693	6728.488	.000	1	.998	.000	.000	.
	MUERTE	40.447	9443.450	.000	1	.997	368181276511140500.000	.000	.
	ANTIBIOTICO	.286	.171	2.811	1	.094	1.331	.953	1.859
	ANTIVIRALES	-.520	.204	6.487	1	.011	.594	.398	.887
	Constante	-2.693	.395	46.362	1	.000	.068		
Paso 2(a)	EDAD_1	.000	.161	.000	1	.998	1.000	.729	1.371
	SEXO	-.051	.154	.110	1	.740	.950	.702	1.286
	Procedencia	.645	.458	1.987	1	.159	1.907	.777	4.677
	FIEBRE	1.280	.232	30.516	1	.000	3.597	2.284	5.666
	TOS	1.320	.320	17.042	1	.000	3.743	2.000	7.003
	D_GARGANTA	-.146	.218	.452	1	.502	.864	.564	1.324
	RINORREA	.543	.208	6.781	1	.009	1.721	1.144	2.589
	EXPECTORACION	-.157	.175	.802	1	.371	.855	.606	1.205
	SIBILANCIAS	-.417	.209	3.986	1	.046	.659	.438	.992
	CONGESTION	.104	.183	.325	1	.568	1.110	.775	1.590
	OTALGIA	.035	.193	.033	1	.855	1.036	.710	1.511
	FOTOFOBIA	-.258	.212	1.483	1	.223	.773	.510	1.170
	CONGES_CONJUN	.164	.170	.929	1	.335	1.178	.844	1.643

	VOMITOS	.159	.185	.735	1	.391	1.172	.815	1.686
	DOLOR_AD	-.030	.184	.026	1	.871	.970	.676	1.392
	DIARREA	-.311	.208	2.242	1	.134	.733	.487	1.101
	ADENOPATIAS	-.936	.250	13.959	1	.000	.392	.240	.641
	ASTENIA	.153	.178	.738	1	.390	1.165	.822	1.652
	CEFALEA	.488	.221	4.877	1	.027	1.629	1.056	2.513
	MIALGIAS	.207	.194	1.134	1	.287	1.230	.840	1.800
	M_GENERAL	-.061	.222	.076	1	.783	.941	.608	1.454
	ERUPC_DERMI CA	.115	.323	.127	1	.721	1.122	.595	2.116
	HOSP	-.872	.226	14.896	1	.000	.418	.268	.651
	D_RESP	-.307	.222	1.913	1	.167	.736	.476	1.137
	PER_CONC	-.748	.704	1.128	1	.288	.473	.119	1.882
	COLAPSO_CIR	-19.650	7687.61 3	.000	1	.998	.000	.000	.
	MUERTE	23.347	7687.61 3	.000	1	.998	137830 29892.7 39	.000	.
	ANTIBIOTICO	.286	.171	2.804	1	.094	1.330	.952	1.858
	ANTIVIRALES	-.546	.203	7.200	1	.007	.579	.389	.863
	Constante	-2.712	.395	47.035	1	.000	.066		
Paso 3(a)	EDAD_1	.005	.161	.001	1	.976	1.005	.733	1.377
	SEXO	-.051	.154	.108	1	.742	.951	.703	1.286
	Procedencia	.658	.460	2.044	1	.153	1.931	.783	4.761
	FIEBRE	1.288	.232	30.924	1	.000	3.624	2.302	5.706
	TOS	1.316	.320	16.958	1	.000	3.730	1.993	6.979
	D_GARGANTA	-.163	.218	.563	1	.453	.849	.555	1.301
	RINORREA	.546	.208	6.897	1	.009	1.726	1.149	2.594
	EXPECTORACI ON	-.185	.175	1.125	1	.289	.831	.590	1.170
	SIBILANCIAS	-.421	.208	4.098	1	.043	.656	.436	.987
	CONGESTION	.122	.183	.449	1	.503	1.130	.790	1.616
	OTALGIA	.034	.192	.031	1	.860	1.034	.709	1.508
	FOTOFOBIA	-.269	.211	1.629	1	.202	.764	.505	1.155
	CONGES_CONJ UN	.172	.169	1.026	1	.311	1.187	.852	1.655
	VOMITOS	.179	.186	.928	1	.335	1.196	.831	1.721
	DOLOR_AD	-.019	.184	.011	1	.916	.981	.684	1.407
	DIARREA	-.342	.207	2.732	1	.098	.710	.474	1.066
	ADENOPATIAS	-.929	.249	13.924	1	.000	.395	.242	.643
	ASTENIA	.135	.178	.575	1	.448	1.144	.807	1.622
	CEFALEA	.481	.221	4.746	1	.029	1.618	1.049	2.493
	MIALGIAS	.213	.194	1.212	1	.271	1.238	.847	1.810
	M_GENERAL	-.026	.221	.014	1	.907	.975	.632	1.502
	ERUPC_DERMI CA	.050	.317	.025	1	.875	1.051	.565	1.957
	HOSP	-.872	.226	14.850	1	.000	.418	.268	.652
	D_RESP	-.328	.222	2.183	1	.140	.720	.466	1.113
	PER_CONC	-1.199	.685	3.066	1	.080	.301	.079	1.154
	MUERTE	5.074	1.301	15.200	1	.000	159.827	12.469	2048.573
	ANTIBIOTICO	.279	.170	2.674	1	.102	1.321	.946	1.845
	ANTIVIRALES	-.538	.203	7.041	1	.008	.584	.393	.869
	Constante	-2.719	.396	47.184	1	.000	.066		
Paso 4(a)	SEXO	-.050	.154	.107	1	.743	.951	.703	1.286
	Procedencia	.657	.459	2.051	1	.152	1.929	.785	4.740
	FIEBRE	1.288	.231	31.134	1	.000	3.626	2.306	5.701
	TOS	1.317	.320	16.979	1	.000	3.731	1.994	6.979
	D_GARGANTA	-.164	.217	.567	1	.451	.849	.555	1.300
	RINORREA	.546	.208	6.917	1	.009	1.727	1.149	2.594
	EXPECTORACI ON	-.186	.173	1.159	1	.282	.830	.592	1.165
	SIBILANCIAS	-.421	.208	4.099	1	.043	.656	.437	.987
	CONGESTION	.123	.182	.451	1	.502	1.130	.791	1.616
	OTALGIA	.033	.191	.030	1	.862	1.034	.711	1.504
	FOTOFOBIA	-.270	.210	1.645	1	.200	.764	.506	1.153

	CONGES_CONJ UN	.172	.169	1.025	1	.311	1.187	.852	1.655
	VOMITOS	.179	.185	.940	1	.332	1.196	.833	1.719
	DOLOR_AD	-.019	.184	.011	1	.917	.981	.684	1.407
	DIARREA	-.342	.206	2.747	1	.097	.710	.474	1.064
	ADENOPATIAS	-.929	.248	13.979	1	.000	.395	.243	.643
	ASTENIA	.135	.178	.575	1	.448	1.144	.807	1.622
	CEFALEA	.481	.220	4.751	1	.029	1.617	1.050	2.491
	MIALGIAS	.213	.193	1.213	1	.271	1.237	.847	1.807
	M_GENERAL	-.026	.221	.014	1	.907	.975	.632	1.502
	ERUPC_DERMICA	.051	.316	.026	1	.873	1.052	.566	1.956
	HOSP	-.872	.226	14.862	1	.000	.418	.268	.651
	D_RESP	-.328	.222	2.182	1	.140	.720	.466	1.113
	PER_CONC	-1.199	.685	3.070	1	.080	.301	.079	1.153
	MUERTE	5.076	1.301	15.222	1	.000	160.064	12.501	2049.419
	ANTIBIOTICO	.279	.169	2.717	1	.099	1.322	.949	1.843
	ANTIVIRALES	-.538	.203	7.047	1	.008	.584	.393	.869
	Constante	-2.717	.391	48.375	1	.000	.066		
Paso 5(a)	SEXO	-.051	.154	.108	1	.742	.951	.703	1.285
	Procedencia	.659	.458	2.068	1	.150	1.933	.787	4.745
	FIEBRE	1.288	.231	31.125	1	.000	3.625	2.306	5.698
	TOS	1.316	.319	16.974	1	.000	3.729	1.994	6.975
	D_GARGANTA	-.162	.217	.561	1	.454	.850	.556	1.300
	RINORREA	.546	.208	6.917	1	.009	1.727	1.149	2.594
	EXPECTORACION	-.187	.173	1.170	1	.279	.830	.591	1.164
	SIBILANCIAS	-.421	.208	4.091	1	.043	.657	.437	.987
	CONGESTION	.123	.182	.457	1	.499	1.131	.791	1.617
	OTALGIA	.032	.191	.029	1	.865	1.033	.710	1.503
	FOTOFOBIA	-.270	.210	1.657	1	.198	.763	.505	1.152
	CONGES_CONJ UN	.171	.169	1.019	1	.313	1.186	.851	1.653
	VOMITOS	.174	.179	.952	1	.329	1.191	.839	1.690
	DIARREA	-.348	.199	3.068	1	.080	.706	.478	1.042
	ADENOPATIAS	-.930	.248	14.033	1	.000	.395	.243	.642
	ASTENIA	.134	.178	.569	1	.451	1.143	.807	1.620
	CEFALEA	.478	.219	4.767	1	.029	1.613	1.050	2.476
	MIALGIAS	.214	.193	1.226	1	.268	1.238	.848	1.808
	M_GENERAL	-.028	.220	.016	1	.898	.972	.632	1.495
	ERUPC_DERMICA	.051	.316	.026	1	.873	1.052	.566	1.956
	HOSP	-.871	.226	14.854	1	.000	.418	.269	.652
	D_RESP	-.328	.222	2.181	1	.140	.721	.466	1.113
	PER_CONC	-1.199	.685	3.065	1	.080	.302	.079	1.154
	MUERTE	5.077	1.300	15.246	1	.000	160.295	12.535	2049.746
	ANTIBIOTICO	.279	.169	2.714	1	.099	1.322	.948	1.843
	ANTIVIRALES	-.538	.203	7.042	1	.008	.584	.393	.869
	Constante	-2.716	.391	48.376	1	.000	.066		
Paso 6(a)	SEXO	-.051	.154	.112	1	.738	.950	.703	1.284
	Procedencia	.660	.458	2.070	1	.150	1.934	.788	4.749
	FIEBRE	1.288	.231	31.121	1	.000	3.624	2.305	5.697
	TOS	1.314	.319	16.967	1	.000	3.722	1.992	6.956
	D_GARGANTA	-.162	.217	.556	1	.456	.851	.556	1.301
	RINORREA	.547	.208	6.937	1	.008	1.728	1.150	2.595
	EXPECTORACION	-.189	.172	1.203	1	.273	.828	.591	1.160
	SIBILANCIAS	-.421	.208	4.113	1	.043	.656	.437	.986
	CONGESTION	.119	.180	.441	1	.507	1.127	.792	1.604
	OTALGIA	.032	.191	.029	1	.866	1.033	.710	1.503
	FOTOFOBIA	-.272	.210	1.674	1	.196	.762	.505	1.150
	CONGES_CONJ UN	.172	.169	1.035	1	.309	1.188	.853	1.654
	VOMITOS	.174	.179	.945	1	.331	1.190	.838	1.689
	DIARREA	-.348	.199	3.060	1	.080	.706	.479	1.043



	ADENOPATIAS	-.928	.248	14.018	1	.000	.395	.243	.643
	ASTENIA	.130	.175	.552	1	.457	1.139	.808	1.606
	CEFALEA	.472	.214	4.869	1	.027	1.603	1.054	2.438
	MIALGIAS	.210	.191	1.211	1	.271	1.234	.849	1.793
	ERUPC_DERMI CA	.052	.316	.027	1	.869	1.054	.567	1.958
	HOSP	-.871	.226	14.854	1	.000	.418	.269	.652
	D_RESP	-.329	.222	2.193	1	.139	.720	.466	1.112
	PÉR_CONC	-1.201	.684	3.080	1	.079	.301	.079	1.151
	MUERTE	5.079	1.302	15.215	1	.000	160.577	12.513	2060.602
	ANTIBIOTICO	.279	.169	2.716	1	.099	1.322	.949	1.843
	ANTIVIRALES	-.538	.202	7.065	1	.008	.584	.393	.868
	Constante	-2.724	.386	49.915	1	.000	.066		
Paso 7(a)	SEXO	-.052	.154	.113	1	.737	.950	.703	1.284
	Procedencia	.657	.458	2.055	1	.152	1.928	.786	4.732
	FIEBRE	1.289	.231	31.278	1	.000	3.631	2.311	5.705
	TOS	1.313	.319	16.952	1	.000	3.717	1.990	6.945
	D_GARGANTA	-.160	.217	.547	1	.460	.852	.557	1.303
	RINORREA	.545	.207	6.912	1	.009	1.725	1.149	2.590
	EXPECTORACI ON	-.188	.172	1.200	1	.273	.828	.591	1.160
	SIBILANCIAS	-.424	.207	4.191	1	.041	.654	.436	.982
	CONGESTION	.120	.180	.442	1	.506	1.127	.792	1.604
	OTALGIA	.033	.191	.031	1	.861	1.034	.711	1.504
	FOTOFOBIA	-.272	.210	1.682	1	.195	.762	.505	1.149
	CONGES_CONJ UN	.174	.169	1.068	1	.301	1.190	.855	1.656
	VOMITOS	.174	.179	.948	1	.330	1.190	.838	1.689
	DIARREA	-.348	.199	3.066	1	.080	.706	.478	1.042
	ADENOPATIAS	-.930	.248	14.063	1	.000	.395	.243	.642
	ASTENIA	.130	.175	.553	1	.457	1.139	.808	1.606
	CEFALEA	.472	.214	4.875	1	.027	1.603	1.054	2.438
	MIALGIAS	.211	.191	1.219	1	.270	1.235	.849	1.794
	HOSP	-.869	.226	14.831	1	.000	.419	.269	.653
	D_RESP	-.330	.222	2.206	1	.137	.719	.466	1.111
	PÉR_CONC	-1.198	.684	3.069	1	.080	.302	.079	1.153
	MUERTE	5.096	1.304	15.264	1	.000	163.320	12.671	2104.988
	ANTIBIOTICO	.280	.169	2.738	1	.098	1.323	.950	1.844
	ANTIVIRALES	-.538	.202	7.051	1	.008	.584	.393	.869
	Constante	-2.724	.386	49.912	1	.000	.066		
Paso 8(a)	SEXO	-.054	.153	.125	1	.724	.947	.702	1.279
	Procedencia	.652	.458	2.033	1	.154	1.920	.783	4.709
	FIEBRE	1.288	.230	31.257	1	.000	3.627	2.309	5.697
	TOS	1.311	.319	16.924	1	.000	3.711	1.987	6.930
	D_GARGANTA	-.159	.217	.540	1	.462	.853	.558	1.304
	RINORREA	.546	.207	6.935	1	.008	1.726	1.150	2.592
	EXPECTORACI ON	-.188	.172	1.198	1	.274	.828	.591	1.161
	SIBILANCIAS	-.422	.207	4.168	1	.041	.656	.437	.983
	CONGESTION	.120	.180	.445	1	.505	1.128	.792	1.605
	FOTOFOBIA	-.266	.207	1.654	1	.198	.767	.511	1.149
	CONGES_CONJ UN	.174	.169	1.071	1	.301	1.191	.856	1.656
	VOMITOS	.174	.179	.948	1	.330	1.190	.838	1.689
	DIARREA	-.345	.198	3.035	1	.081	.708	.481	1.044
	ADENOPATIAS	-.930	.248	14.069	1	.000	.395	.243	.641
	ASTENIA	.132	.175	.569	1	.451	1.141	.810	1.608
	CEFALEA	.474	.213	4.944	1	.026	1.607	1.058	2.442
	MIALGIAS	.213	.190	1.258	1	.262	1.238	.853	1.797
	HOSP	-.875	.223	15.351	1	.000	.417	.269	.646
	D_RESP	-.328	.222	2.189	1	.139	.720	.467	1.112
	PÉR_CONC	-1.199	.684	3.071	1	.080	.302	.079	1.153
	MUERTE	5.092	1.304	15.243	1	.000	162.641	12.623	2095.512
	ANTIBIOTICO	.282	.169	2.789	1	.095	1.326	.952	1.846
	ANTIVIRALES	-.538	.202	7.056	1	.008	.584	.393	.869

	Constante	-2.721	.385	49.939	1	.000	.066		
Paso 9(a)	Procedencia	.650	.457	2.018	1	.155	1.915	.781	4.695
	FIEBRE	1.287	.230	31.203	1	.000	3.623	2.306	5.692
	TOS	1.307	.318	16.853	1	.000	3.696	1.980	6.900
	D_GARGANTA	-.163	.216	.565	1	.452	.850	.556	1.299
	RINORREA	.550	.207	7.044	1	.008	1.733	1.155	2.600
	EXPECTORACION	-.189	.172	1.209	1	.272	.828	.591	1.160
	SIBILANCIAS	-.423	.207	4.180	1	.041	.655	.437	.983
	CONGESTION	.119	.180	.440	1	.507	1.127	.792	1.604
	FOTOFOBIA	-.261	.206	1.602	1	.206	.770	.514	1.154
	CONGES_CONJUN	.171	.168	1.028	1	.311	1.186	.853	1.649
	VOMITOS	.180	.178	1.018	1	.313	1.197	.844	1.696
	DIARREA	-.347	.198	3.069	1	.080	.707	.480	1.042
	ADENOPATIAS	-.932	.248	14.132	1	.000	.394	.242	.640
	ASTENIA	.133	.175	.574	1	.449	1.142	.810	1.609
	CEFALEA	.482	.212	5.136	1	.023	1.619	1.067	2.455
	MIALGIAS	.211	.190	1.236	1	.266	1.235	.851	1.794
	HOSP	-.872	.223	15.279	1	.000	.418	.270	.648
	D_RESP	-.329	.221	2.211	1	.137	.719	.466	1.110
	PÉR_CONC	-1.199	.684	3.076	1	.079	.302	.079	1.151
	MUERTE	5.111	1.306	15.325	1	.000	165.846	12.834	2143.136
	ANTIBIOTICO	.282	.169	2.779	1	.096	1.325	.952	1.845
	ANTIVIRALES	-.541	.202	7.166	1	.007	.582	.392	.865
	Constante	-2.744	.379	52.342	1	.000	.064		
Paso 10(a)	Procedencia	.658	.458	2.067	1	.151	1.932	.787	4.739
	FIEBRE	1.302	.229	32.239	1	.000	3.677	2.346	5.763
	TOS	1.317	.318	17.143	1	.000	3.733	2.001	6.964
	D_GARGANTA	-.145	.214	.455	1	.500	.865	.568	1.317
	RINORREA	.560	.206	7.354	1	.007	1.750	1.168	2.623
	EXPECTORACION	-.185	.172	1.155	1	.283	.831	.594	1.164
	SIBILANCIAS	-.414	.206	4.034	1	.045	.661	.441	.990
	FOTOFOBIA	-.259	.206	1.581	1	.209	.772	.515	1.156
	CONGES_CONJUN	.182	.167	1.182	1	.277	1.199	.864	1.665
	VOMITOS	.189	.177	1.137	1	.286	1.208	.853	1.711
	DIARREA	-.349	.198	3.111	1	.078	.706	.479	1.040
	ADENOPATIAS	-.921	.247	13.885	1	.000	.398	.245	.646
	ASTENIA	.138	.175	.622	1	.430	1.148	.815	1.617
	CEFALEA	.493	.212	5.428	1	.020	1.638	1.081	2.480
	MIALGIAS	.212	.190	1.246	1	.264	1.237	.852	1.795
	HOSP	-.874	.223	15.355	1	.000	.417	.269	.646
	D_RESP	-.328	.221	2.195	1	.138	.720	.467	1.112
	PÉR_CONC	-1.180	.682	2.994	1	.084	.307	.081	1.169
	MUERTE	5.058	1.300	15.150	1	.000	157.307	12.319	2008.733
	ANTIBIOTICO	.282	.169	2.782	1	.095	1.325	.952	1.845
	ANTIVIRALES	-.527	.201	6.885	1	.009	.590	.398	.875
	Constante	-2.732	.378	52.179	1	.000	.065		
Paso 11(a)	Procedencia	.640	.456	1.967	1	.161	1.896	.775	4.637
	FIEBRE	1.301	.229	32.266	1	.000	3.674	2.345	5.757
	TOS	1.293	.316	16.744	1	.000	3.644	1.962	6.771
	RINORREA	.538	.204	6.977	1	.008	1.712	1.149	2.552
	EXPECTORACION	-.190	.172	1.223	1	.269	.827	.591	1.158
	SIBILANCIAS	-.411	.206	3.974	1	.046	.663	.442	.993
	FOTOFOBIA	-.253	.206	1.511	1	.219	.776	.518	1.162
	CONGES_CONJUN	.181	.167	1.172	1	.279	1.198	.864	1.663
	VOMITOS	.186	.177	1.101	1	.294	1.204	.851	1.705
	DIARREA	-.348	.198	3.104	1	.078	.706	.479	1.040
	ADENOPATIAS	-.926	.247	14.048	1	.000	.396	.244	.643
	ASTENIA	.138	.175	.625	1	.429	1.148	.815	1.618

	CEFALEA	.471	.209	5.080	1	.024	1.602	1.063	2.413
	MIALGIAS	.201	.189	1.123	1	.289	1.222	.843	1.772
	HOSP	-.858	.222	15.006	1	.000	.424	.275	.654
	D_RESP	-.328	.221	2.194	1	.139	.720	.467	1.112
	PER_CONC	-1.162	.681	2.910	1	.088	.313	.082	1.189
	MUERTE	5.038	1.308	14.838	1	.000	154.150	11.877	2000.757
	ANTIBIOTICO	.285	.169	2.842	1	.092	1.329	.955	1.851
	ANTIVIRALES	-.536	.200	7.169	1	.007	.585	.395	.866
	Constante	-2.788	.370	56.670	1	.000	.062		
Paso 12(a)	Procedencia	.637	.455	1.965	1	.161	1.891	.776	4.611
	FIEBRE	1.300	.229	32.260	1	.000	3.669	2.343	5.746
	TOS	1.300	.316	16.930	1	.000	3.668	1.975	6.811
	RINORREA	.540	.204	7.032	1	.008	1.716	1.151	2.558
	EXPECTORACION	-.171	.170	1.013	1	.314	.843	.604	1.176
	SIBILANCIAS	-.416	.206	4.076	1	.043	.660	.440	.988
	FOTOFOBIA	-.241	.205	1.377	1	.241	.786	.525	1.175
	CONGES_CONJUN	.188	.167	1.276	1	.259	1.207	.871	1.675
	VOMITOS	.191	.177	1.161	1	.281	1.210	.855	1.712
	DIARREA	-.341	.197	2.989	1	.084	.711	.483	1.047
	ADENOPATIAS	-.917	.246	13.851	1	.000	.400	.247	.648
	CEFALEA	.496	.207	5.743	1	.017	1.642	1.094	2.463
	MIALGIAS	.230	.186	1.533	1	.216	1.259	.874	1.812
	HOSP	-.850	.221	14.772	1	.000	.427	.277	.659
	D_RESP	-.319	.221	2.088	1	.148	.727	.471	1.121
	PER_CONC	-1.163	.683	2.894	1	.089	.313	.082	1.193
	MUERTE	5.035	1.297	15.063	1	.000	153.749	12.091	1955.060
	ANTIBIOTICO	.283	.169	2.806	1	.094	1.327	.953	1.846
	ANTIVIRALES	-.523	.200	6.878	1	.009	.593	.401	.876
	Constante	-2.771	.369	56.319	1	.000	.063		
Paso 13(a)	Procedencia	.665	.453	2.156	1	.142	1.945	.800	4.724
	FIEBRE	1.303	.229	32.495	1	.000	3.681	2.352	5.763
	TOS	1.237	.309	15.986	1	.000	3.444	1.878	6.315
	RINORREA	.510	.201	6.426	1	.011	1.665	1.123	2.469
	SIBILANCIAS	-.444	.204	4.728	1	.030	.641	.430	.957
	FOTOFOBIA	-.256	.205	1.567	1	.211	.774	.518	1.156
	CONGES_CONJUN	.187	.167	1.261	1	.261	1.206	.870	1.672
	VOMITOS	.188	.177	1.131	1	.287	1.207	.853	1.708
	DIARREA	-.332	.197	2.849	1	.091	.717	.488	1.055
	ADENOPATIAS	-.925	.246	14.105	1	.000	.397	.245	.643
	CEFALEA	.482	.206	5.483	1	.019	1.620	1.082	2.426
	MIALGIAS	.222	.185	1.429	1	.232	1.248	.868	1.795
	HOSP	-.866	.221	15.407	1	.000	.421	.273	.648
	D_RESP	-.333	.220	2.285	1	.131	.717	.465	1.104
	PER_CONC	-1.144	.682	2.813	1	.094	.318	.084	1.213
	MUERTE	4.996	1.279	15.268	1	.000	147.790	12.060	1811.117
	ANTIBIOTICO	.270	.168	2.588	1	.108	1.311	.943	1.822
	ANTIVIRALES	-.528	.199	7.027	1	.008	.590	.399	.871
	Constante	-2.754	.368	55.962	1	.000	.064		
Paso 14(a)	Procedencia	.655	.452	2.098	1	.148	1.925	.793	4.671
	FIEBRE	1.312	.228	33.033	1	.000	3.714	2.374	5.809
	TOS	1.236	.309	16.014	1	.000	3.441	1.879	6.303
	RINORREA	.509	.201	6.413	1	.011	1.663	1.122	2.466
	SIBILANCIAS	-.446	.204	4.768	1	.029	.640	.429	.955
	FOTOFOBIA	-.248	.205	1.474	1	.225	.780	.522	1.165
	CONGES_CONJUN	.199	.166	1.438	1	.230	1.221	.881	1.690
	DIARREA	-.277	.190	2.126	1	.145	.758	.523	1.100
	ADENOPATIAS	-.930	.246	14.279	1	.000	.395	.244	.639
	CEFALEA	.481	.206	5.477	1	.019	1.618	1.081	2.422
	MIALGIAS	.232	.185	1.571	1	.210	1.261	.877	1.813

	HOSP	-.843	.220	14.753	1	.000	.430	.280	.662
	D_RESP	-.340	.220	2.381	1	.123	.712	.462	1.096
	PER_CONC	-1.100	.680	2.617	1	.106	.333	.088	1.262
	MUERTE	4.936	1.269	15.128	1	.000	139.246	11.574	1675.192
	ANTIBIOTICO	.274	.168	2.650	1	.104	1.315	.946	1.828
	ANTIVIRALES	-.536	.199	7.274	1	.007	.585	.396	.864
	Constante	-2.737	.367	55.659	1	.000	.065		
Paso 15(a)	Procedencia	.636	.452	1.979	1	.160	1.888	.779	4.579
	FIEBRE	1.332	.228	34.161	1	.000	3.789	2.424	5.922
	TOS	1.231	.308	15.965	1	.000	3.425	1.872	6.266
	RINORREA	.543	.199	7.433	1	.006	1.721	1.165	2.542
	SIBILANCIAS	-.430	.203	4.468	1	.035	.651	.437	.969
	FOTOFOBIA	-.181	.197	.852	1	.356	.834	.567	1.226
	DIARREA	-.276	.190	2.125	1	.145	.758	.523	1.100
	ADENOPATIAS	-.895	.243	13.523	1	.000	.408	.253	.658
	CEFALEA	.482	.206	5.484	1	.019	1.620	1.082	2.425
	MIALGIAS	.247	.185	1.777	1	.183	1.280	.891	1.839
	HOSP	-.832	.219	14.418	1	.000	.435	.283	.669
	D_RESP	-.332	.220	2.275	1	.131	.718	.467	1.104
	PER_CONC	-1.080	.679	2.532	1	.112	.339	.090	1.284
	MUERTE	4.939	1.297	14.496	1	.000	139.669	10.986	1775.616
	ANTIBIOTICO	.260	.168	2.410	1	.121	1.297	.934	1.801
	ANTIVIRALES	-.527	.198	7.048	1	.008	.590	.400	.871
	Constante	-2.726	.366	55.405	1	.000	.065		
Paso 16(a)	Procedencia	.646	.451	2.054	1	.152	1.908	.789	4.617
	FIEBRE	1.327	.228	33.923	1	.000	3.769	2.412	5.890
	TOS	1.243	.307	16.383	1	.000	3.467	1.899	6.329
	RINORREA	.518	.197	6.894	1	.009	1.678	1.140	2.469
	SIBILANCIAS	-.444	.203	4.793	1	.029	.641	.431	.955
	DIARREA	-.290	.189	2.359	1	.125	.748	.516	1.083
	ADENOPATIAS	-.908	.243	13.998	1	.000	.403	.251	.649
	CEFALEA	.473	.205	5.300	1	.021	1.605	1.073	2.400
	MIALGIAS	.229	.184	1.544	1	.214	1.257	.876	1.802
	HOSP	-.829	.219	14.383	1	.000	.436	.284	.670
	D_RESP	-.334	.219	2.314	1	.128	.716	.466	1.101
	PER_CONC	-1.066	.677	2.476	1	.116	.345	.091	1.299
	MUERTE	4.942	1.298	14.499	1	.000	139.987	11.001	1781.275
	ANTIBIOTICO	.264	.167	2.482	1	.115	1.302	.938	1.807
	ANTIVIRALES	-.524	.198	6.979	1	.008	.592	.401	.873
	Constante	-2.724	.366	55.450	1	.000	.066		
Paso 17(a)	Procedencia	.640	.448	2.039	1	.153	1.896	.788	4.563
	FIEBRE	1.338	.228	34.581	1	.000	3.811	2.440	5.953
	TOS	1.235	.307	16.229	1	.000	3.439	1.886	6.272
	RINORREA	.534	.196	7.391	1	.007	1.706	1.161	2.507
	SIBILANCIAS	-.429	.202	4.506	1	.034	.651	.438	.968
	DIARREA	-.281	.189	2.214	1	.137	.755	.522	1.093
	ADENOPATIAS	-.874	.241	13.128	1	.000	.417	.260	.670
	CEFALEA	.583	.185	9.875	1	.002	1.791	1.245	2.576
	HOSP	-.820	.218	14.114	1	.000	.441	.287	.676
	D_RESP	-.335	.219	2.331	1	.127	.716	.466	1.100
	PER_CONC	-1.109	.674	2.703	1	.100	.330	.088	1.238
	MUERTE	4.946	1.293	14.638	1	.000	140.588	11.158	1771.311
	ANTIBIOTICO	.258	.167	2.378	1	.123	1.294	.933	1.795
	ANTIVIRALES	-.512	.198	6.689	1	.010	.599	.406	.883
	Constante	-2.678	.363	54.550	1	.000	.069		
Paso 18(a)	FIEBRE	1.331	.227	34.417	1	.000	3.784	2.426	5.904
	TOS	1.201	.305	15.559	1	.000	3.325	1.830	6.040
	RINORREA	.542	.196	7.631	1	.006	1.720	1.171	2.527
	SIBILANCIAS	-.441	.202	4.771	1	.029	.643	.433	.956
	DIARREA	-.275	.188	2.125	1	.145	.760	.525	1.099
	ADENOPATIAS	-.890	.241	13.654	1	.000	.411	.256	.658

	CEFALEA	.560	.184	9.246	1	.002	1.751	1.220	2.512
	HOSP	-.805	.217	13.715	1	.000	.447	.292	.684
	D_RESP	-.350	.219	2.568	1	.109	.705	.459	1.081
	PER_CONC	-1.103	.676	2.661	1	.103	.332	.088	1.249
	MUERTE	4.915	1.287	14.585	1	.000	136.298	10.941	1697.964
	ANTIBIOTICO	.249	.167	2.226	1	.136	1.282	.925	1.778
	ANTIVIRALES	-.510	.198	6.628	1	.010	.601	.408	.885
	Constante	-2.603	.356	53.412	1	.000	.074		
Paso 19(a)	FIEBRE	1.320	.227	33.890	1	.000	3.744	2.400	5.839
	TOS	1.202	.305	15.532	1	.000	3.327	1.830	6.048
	RINORREA	.524	.195	7.178	1	.007	1.688	1.151	2.476
	SIBILANCIAS	-.437	.201	4.704	1	.030	.646	.435	.959
	ADENOPATIAS	-.881	.240	13.431	1	.000	.414	.259	.664
	CEFALEA	.561	.184	9.311	1	.002	1.753	1.222	2.513
	HOSP	-.789	.217	13.264	1	.000	.454	.297	.695
	D_RESP	-.375	.217	2.984	1	.084	.687	.449	1.052
	PER_CONC	-1.127	.679	2.753	1	.097	.324	.086	1.227
	MUERTE	4.944	1.296	14.549	1	.000	140.396	11.065	1781.477
	ANTIBIOTICO	.253	.166	2.324	1	.127	1.288	.930	1.785
	ANTIVIRALES	-.527	.197	7.138	1	.008	.590	.401	.869
	Constante	-2.634	.357	54.346	1	.000	.072		
Paso 20(a)	FIEBRE	1.349	.226	35.593	1	.000	3.855	2.475	6.006
	TOS	1.238	.304	16.614	1	.000	3.449	1.902	6.256
	RINORREA	.544	.195	7.782	1	.005	1.723	1.176	2.526
	SIBILANCIAS	-.423	.201	4.431	1	.035	.655	.442	.971
	ADENOPATIAS	-.860	.239	12.944	1	.000	.423	.265	.676
	CEFALEA	.559	.184	9.259	1	.002	1.750	1.220	2.509
	HOSP	-.732	.213	11.832	1	.001	.481	.317	.730
	D_RESP	-.372	.217	2.945	1	.086	.689	.450	1.054
	PER_CONC	-1.101	.682	2.607	1	.106	.333	.087	1.265
	MUERTE	5.019	1.312	14.626	1	.000	151.224	11.550	1979.982
	ANTIVIRALES	-.445	.190	5.501	1	.019	.641	.442	.930
	Constante	-2.620	.356	54.016	1	.000	.073		
Paso 21(a)	FIEBRE	1.357	.226	36.080	1	.000	3.884	2.495	6.048
	TOS	1.258	.303	17.246	1	.000	3.517	1.943	6.367
	RINORREA	.558	.194	8.244	1	.004	1.748	1.194	2.558
	SIBILANCIAS	-.413	.201	4.252	1	.039	.661	.446	.980
	ADENOPATIAS	-.866	.238	13.227	1	.000	.421	.264	.671
	CEFALEA	.559	.184	9.270	1	.002	1.749	1.220	2.507
	HOSP	-.775	.211	13.450	1	.000	.461	.305	.697
	D_RESP	-.388	.216	3.223	1	.073	.678	.444	1.036
	MUERTE	4.590	1.185	14.993	1	.000	98.449	9.645	1004.919
	ANTIVIRALES	-.420	.189	4.939	1	.026	.657	.454	.952
	Constante	-2.661	.356	55.727	1	.000	.070		
Paso 22(b)	FIEBRE	1.349	.226	35.709	1	.000	3.855	2.477	6.002
	TOS	1.269	.304	17.477	1	.000	3.557	1.962	6.447
	RINORREA	.550	.196	7.916	1	.005	1.733	1.182	2.543
	SIBILANCIAS	-.424	.201	4.435	1	.035	.654	.441	.971
	ADENOPATIAS	-.873	.240	13.190	1	.000	.418	.261	.669
	CEFALEA	.559	.184	9.201	1	.002	1.748	1.219	2.508
	HOSP	-.753	.211	12.687	1	.000	.471	.311	.713
	D_RESP	-.350	.217	2.600	1	.107	.705	.461	1.078
	MUERTE	-20.112	7723.33	.000	1	.998	.000	.000	
	ANTIBIOTICO		7						
	ANTIVIRALES	23.342	7723.33	.000	1	.998	137227	.000	
	Constante		7				48522.7		
							30		
		-.438	.190	5.317	1	.021	.645	.445	.936
	Constante	-2.650	.357	55.268	1	.000	.071		

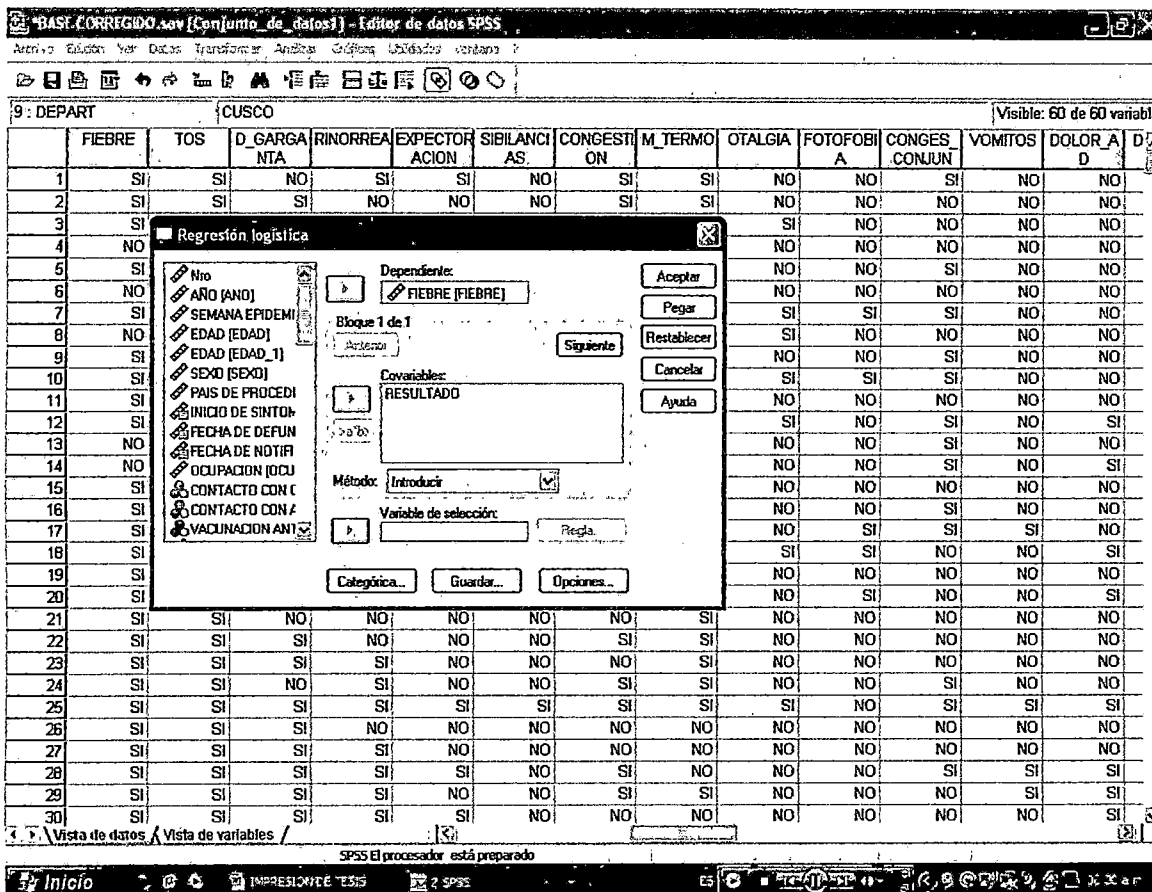
- a Variable(s) introducida(s) en el paso 1: EDAD\_1, SEXO, Procedencia, FIEBRE, TOS, D\_GARGANTA, RINORREA, EXPECTORACION, SIBILANCIAS, CONGESTION, OTALGIA, FOTOFOBIA, CONGES\_CONJUN, VOMITOS, DOLOR\_AD, DIARREA, ADENOPATIAS, ASTENIA, CEFALEA, MIALGIAS, M\_GENERAL, ERUPC\_DERMICA, HOSP, FALLA\_MULT, D\_RESP, PER\_CONC, COLAPSO\_CIR, MUERTE, ANTIBIOTICO, ANTIVIRALES.
- b Variable(s) introducida(s) en el paso 22: COLAPSO\_CIR.
- c Se ha detenido un procedimiento por pasos ya que al eliminar la variable menos significativa se obtuvo un modelo previamente ajustado.

### ANEXO 4

#### SALIDA DEL SPSS PASO A PASO: Modelo de Regresión Logística. Tamaño de muestra 924.

The screenshot shows the SPSS data editor for 'BASE CORREGIDO.sav'. The 'Análisis' menu is open, and 'Logística binaria...' is selected. The data table below shows the following variables and their values for 30 cases:

	TOS	D_GARGANTA	RINORREA	CONGESTION	M_TERMOM	OTALGIA	FOTOFOBIA	CONGESCONJUN	VOMITOS	DOLOR_AD	DIARREA	TE
1	SI	NO		SI	SI	NO	NO	SI	NO	NO	NO	TE
2	SI	SI		SI	SI	NO	NO	NO	NO	NO	NO	TE
3	SI	SI		SI	SI	NO	NO	NO	NO	NO	SI	TE
4	NO	SI		SI	SI	NO	NO	NO	NO	NO	NO	TE
5	SI	SI		SI	SI	NO	NO	SI	NO	NO	SI	TE
6	SI	SI		SI	SI	NO	NO	NO	NO	NO	SI	TE
7	SI	SI		SI	SI	NO	SI	SI	NO	NO	NO	TE
8	SI	SI		SI	SI	NO	NO	NO	NO	NO	SI	TE
9	SI	SI		SI	SI	NO	SI	SI	NO	NO	NO	TE
10	SI	SI		SI	SI	NO	SI	SI	NO	NO	NO	TE
11	SI	SI		SI	SI	NO	NO	NO	NO	NO	SI	TE
12	SI	SI		SI	SI	NO	NO	SI	NO	SI	NO	TE
13	NO	SI		SI	SI	NO	NO	SI	NO	NO	NO	TE
14	SI	SI		SI	SI	NO	NO	SI	NO	SI	SI	TE
15	SI	SI		SI	SI	NO	NO	NO	NO	NO	NO	TE
16	SI	SI		SI	SI	NO	NO	SI	NO	NO	NO	TE
17	SI	SI		SI	SI	NO	SI	SI	SI	NO	NO	TE
18	NO	SI		SI	SI	SI	SI	NO	NO	SI	NO	TE
19	SI	SI		SI	SI	NO	NO	NO	NO	NO	NO	TE
20	SI	SI		SI	SI	NO	SI	NO	NO	SI	SI	TE
21	SI	NO		SI	SI	NO	NO	NO	NO	NO	NO	TE
22	SI	SI		SI	SI	NO	NO	NO	NO	NO	NO	TE
23	SI	SI		SI	SI	NO	NO	NO	NO	NO	NO	TE
24	SI	NO		SI	SI	NO	NO	SI	NO	NO	NO	TE
25	SI	SI		SI	SI	SI	NO	SI	SI	SI	NO	TE
26	SI	SI		SI	SI	NO	NO	NO	NO	NO	NO	TE
27	SI	SI		SI	SI	NO	NO	NO	NO	NO	NO	TE
28	SI	SI		SI	SI	NO	NO	SI	SI	SI	NO	TE
29	SI	SI		SI	SI	NO	NO	NO	SI	SI	NO	TE
30	SI	SI		SI	SI	NO	NO	NO	NO	SI	NO	TE



## Regresión logística

[Conjunto\_de\_datos1] F:\RESULTADOS DE TESIS\BASE CORREGIDO.sav

Resumen del procesamiento de los casos

Casos no ponderados <sup>a</sup>		N	Porcentaje
Casos seleccionados	Incluidos en el análisis	924	100.0
	Casos perdidos	0	.0
	Total	924	100.0
Casos no seleccionados		0	.0
Total		924	100.0

a. Si está activada la ponderación, consulte la tabla de clasificación para ver el número total de casos.

Codificación de la variable dependiente

Valor original	Valor interno
NO	0
SI	1

## Bloque 0: Bloque inicial

Tabla de clasificación <sup>a,b</sup>

Observado			Pronosticado		
			FIEBRE		Porcentaje correcto
			NO	SI	
Paso 0	FIEBRE	NO	0	168	.0
		SI	0	756	100.0
Porcentaje global					81.8

a. En el modelo se incluye una constante.

b. El valor de corte es .500

Variables en la ecuación

	B	E.T.	Wald	gl	Sig.	Exp(B)	
Paso 0	Constante	1.504	.085	310.956	1	.000	4.500

Variables que no están en la ecuación

			Puntuación	gl	Sig.
Paso 0	Variables	RESULTADO	52.167	1	.000
	Estadísticos globales		52.167	1	.000

## Bloque 1: Método = Introducir

Pruebas omnibus sobre los coeficientes del modelo

		Chi-cuadrado	gl	Sig.
Paso 1	Paso	55.162	1	.000
	Bloque	55.162	1	.000
	Modelo	55.162	1	.000

Resumen de los modelos

Paso	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	821.048 <sup>a</sup>	.058	.095

a. La estimación ha finalizado en el número de iteración 5 porque las estimaciones de los parámetros han cambiado en menos de .001.



Tabla de clasificación <sup>a</sup>

Observado			Pronosticado		
			FIEBRE		Porcentaje correcto
			NO	SI	
Paso 1	FIEBRE	NO	0	168	.0
		SI	0	756	100.0
Porcentaje global					81.8

a. El valor de corte es .500

Variables en la ecuación

	B	E.T.	Wald	gl	Sig.	Exp(B)	
Paso 1	RESULTADO	1.376	.200	47.338	1	.000	3.957
1	Constante	1.003	.102	96.321	1	.000	2.725

a. Variable(s) introducida(s) en el paso 1: RESULTADO.

## ANEXO 5

### GLOSARIO DE TERMINOS

**Rinorrea:** Secreción nasal. La rinorrea es común, pero rara vez es grave. El drenaje de los senos paranasales inflamados o infectados puede ser espeso o decolorado. El exceso de moco puede bajar por la parte posterior de la garganta (goteo retranasal) u ocasionar una tos que generalmente empeora en la noche. La irritación en la garganta también puede resultar del excesivo drenaje de moco.

**Expectoración:** (Del latín *ex*, fuera, y *pectus*. Pecho). Fenómeno por el cual los productos formados en las vías respiratorias son expulsados fuera del pecho.

**Sibilancias:** Son un sonido silbante y agudo mientras la respiración que acontece cuando el aire fluye a través de las vías respiratorias estrechas. La sibilancia, también llamado estertor sibilante, es el sonido que hace el aire al pasar por las vías respiratorias congestionadas; se trata de un sonido agudo y silbante.

**Fotofobia:** Sensación desagradable y hasta dolorosa provocada por la luz más o menos intensa. La fotofobia es la intolerancia anormal a la luz. Es frecuente en personas con albinismo o puede ser debida por enfermedades relacionadas con el ojo o el sistema nervioso.

**Adenopatía:** Enfermedad de las glándulas en general y de los ganglios linfáticos en particular. Estos aumentan su tamaño como reflejo de defensa ante enfermedades como faringitis, catarro, tumoraciones y otros males de origen infeccioso. (Del griego *adén*, glándula, y *pathos*, enfermedad). Término genérico que se emplea para designar las inflamaciones crónicas de los ganglios linfáticos.

**Astenia:** La astenia (cansancio, debilidad generalizada) es frecuente en los pacientes con insuficiencia cardiaca. Se debe principalmente a que, como el corazón no bombea sangre en cantidades suficientes, los músculos no tienen todo el aporte de oxígeno y nutrientes que sería deseable.

**Cefalea:** El término cefalea (del latín cephalaea, hace referencia a los dolores y molestias localizadas en cualquier parte de la cabeza, en los diferentes tejidos de la cavidad craneana, en las estructuras que lo unen a la base del cráneo, los músculos.

**Mialgia:** Dolor originado en los músculos. Suele acompañar otros síntomas como decaimiento, fiebre, dolor de cabeza y en las enfermedades infecciosas. También suele asociarse a distintas enfermedades inmunológicas.

Las mialgias consisten en dolores musculares que pueden afectar a uno o varios músculos del cuerpo y pueden estar producidos por causas muy diversas. Estos dolores musculares pueden acompañarse en ocasiones de debilidad o pérdida de la fuerza y dolor a la palpación. También se asocia en ocasiones con calambres y contracturas de los músculos afectados.

**Otalgia:** Dolor de oídos. Se produce por infecciones y neoplasias del oído externo y medio, o como dolor referido de otros procesos patológicos. La causa más frecuente de otalgia es la otitis media aguda. Pérdida de sensación auditiva, uni o bilateral, que se caracteriza por su aparición en un corto período de tiempo, que puede variar desde breves instantes hasta unos pocos días.