

VALIDACIÓN Y CALIBRACIÓN DE INDICADORES DISCRIMINANTES PARA PACIENTES HOSPITALARIOS CON PROBLEMAS CARDIOVASCULARES

Abel Flores Laaz¹, Cesar Guerrero²

¹Ingeniero en Estadística e Informática, ICM-ESPOL, asflores@espol.edu.ec

²Mat., MSc. en Matemáticas ICM-ESPOL, cguerrer@espol.edu.ec

Resumen: El presente artículo muestra la utilización de técnicas estadísticas para adaptar el modelo Framingham a las características que presentan los pacientes con problemas cardiovasculares del Hospital Luis Vernaza de Guayaquil; fundamentalmente, consiste en estimar la probabilidad de riesgo o la probabilidad que una persona reingrese al Hospital por problemas cardiovasculares. Por ello, este estudio es en base a la información obtenida en el establecimiento médico (1994-2003), consistente en las variables que utiliza el modelo antes mencionado. A partir de dicha información, se llevó a cabo la validación y calibración del modelo que se utilizó en la presente investigación, donde la validación consiste en evaluar el modelo para los coeficientes proporcionados por Framingham y además se debe observar si los resultados se adaptan a las características de los casos de estudio; también se hizo un ajuste en base a la distribución de Supervivencia Weibull, para de esta forma obtener los valores de supervivencia base a 10 años. Debido que los valores obtenidos en la evaluación no se ajustan a los datos de estudio, se calibró los coeficientes de dicho modelo.

Summary: The present article shows the use of statistical techniques to adapt the model Framingham to the characteristics that the patients present with cardiovascular problems of the Hospital Luis Vernaza of Guayaquil; fundamentally, it consists on estimating the probability of risk or the probability that a person re-enter to the Hospital for cardiovascular problems. For it, this study is based on the information obtained in the medical establishment (1994-2003), consistent in the variables that it uses the pattern before mentioned. Starting from this information, it was carried out the validation and calibration of the model that it was used in the present investigation, where the validation consists on evaluating the pattern for the coefficients provided by Framingham and it should also be observed if the results adapt to the characteristics of the cases of study; an adjustment was also made based on the distribution of Survival Weibull, for this way to obtain the values of survival bases to 10 years. Due that the values obtained in the evaluation are not adjusted to the study data, it was gauged this model's coefficients.

Palabras Claves: Framingham, Weibull, Validación, Calibración, Cohorte.

1. INTRODUCCIÓN

El manejo de la información sobre los niveles elevados de colesterol, HDL-Colesterol, tensión arterial, diabético, fumador; de personas con problemas cardiovasculares, son fundamentales para que los doctores ofrezcan a los pacientes mejores medidas correctivas, y de esta manera poder disminuir la probabilidad que se presente otro problema cardiovascular.

La presente investigación trata determinar la probabilidad de re-ingreso para personas con problemas cardiovasculares, para ello, usa el estudio creado en USA-MASSACHUSETTS (Modelo de Framingham). La hipótesis que se plantea es: "El modelo Framingham puede ser utilizado para calcular la probabilidad de re-ingreso"; ya que posiblemente influya mucho las diferentes características de vida (USA-ECUADOR), y de la cual se derivan las siguientes interrogantes: ¿Qué género es más propenso a re-ingresar al hospital por problemas cardiovasculares?, ¿Influye la edad en tener mayor probabilidad de re-ingresar al Hospital?, ¿Los casos de estudio tienen una alta o aceptable probabilidad de no re-ingresar al Hospital Luis Vernaza?, ¿Los parámetros de Framingham (USA-MASSACHUSETTS) sobre-estimarán la probabilidad de re-ingresos de los casos de estudio?.

Estudios de este tipo han sido realizados en otros países, por lo cual, la presente investigación es la primera en nuestra ciudad y país, por lo tanto sería conveniente tomarla como un estudio piloto, dejando sentada las bases para un estudio mayor, como podría ser a nivel de país.

2. PROBLEMAS CARDIOVASCULARES

La enfermedad cardiovascular se refiere a las enfermedades del corazón y enfermedades del sistema de vasos sanguíneos (arterias, capilares, venas) de todo el organismo, en lugares tales como el cerebro, las piernas y los pulmones. La palabra cardiovascular se descompone en **Cardio** que se refiere al corazón, y **Vascular** al sistema de vasos sanguíneos. Bajo el término de enfermedades cardiovasculares se engloban un conjunto de procesos que afectan al corazón y al sistema vascular: enfermedades coronarias, valvulares, miocárdicas, congénitas, cerebrovasculares, vasculares periféricas y otras.

2.1. FACTORES DE RIESGO CARDIOVASCULAR

Los factores de riesgos que intervienen en el desarrollo de las enfermedades cardiovasculares están determinados por las siguientes variables: sexo, edad, herencia o antecedentes familiares, niveles elevados de colesterol total y LDL,

niveles de bajos de colesterol HDL, tabaquismo, hipertensión y diabetes.

Es fundamental conocer los niveles recomendables de colesterol, HDL colesterol y tensión arterial, los cuales serán presentados en las siguientes tablas:

TABLA I
Niveles de Colesterol y HDL-Colesterol

Categorías	Colesterol Total	Colesterol HDL
Recomendable	<200	>40
Cercano a los límites elevados (Riesgo bajo)	200-239	>35
Elevado (Riesgo alto)	>= 240	<35

Fuente: Hospital Luis Vernaza de Guayaquil, tesis 2005
Elaboración: Abel Flores

TABLA II
Niveles de Tensión Arterial

Categorías	Normal	Adulto
Presión Normal	< 120/80	< 130/85
Presión Alta	>= 140/90	>= 140/90

Fuente: Hospital Luis Vernaza de Guayaquil, tesis 2005
Elaboración: Abel Flores

3. TEORÍA DE RIESGO Y SUPERVIVENCIA

La función de probabilidad de la variable tiempo se puede especificar de varias formas: **Riesgo** ($h(t)$) es la probabilidad que se produzca un evento de forma aleatoria, basándonos en esta investigación, “*el riesgo es la probabilidad que el paciente vuelva a ingresar al Hospital por problemas cardiovasculares*”. **Supervivencia** ($S(t)$) “*es la probabilidad que el paciente no reingrese al Hospital por los problemas antes mencionado*”.

Cuando se habla de supervivencia o riesgo, es fundamental definir la unidad de observación, los sucesos aleatorios y el evento inicial. La unidad de observación en la presente investigación son las personas. El suceso aleatorio será el reingreso al hospital por problemas cardiovasculares y el evento inicial es la primera vez cuando ingresó por un problema cardiovascular.

Existen diversos métodos para poder estimar la función de probabilidad y supervivencia, las cuales pueden ser estimadas por los siguientes procedimientos no paramétricos:

Estimación de la Función de Probabilidad

$$\hat{f}(t) = \frac{\text{Número de Pacientes que fallecieron en el intervalo } t}{(\# \text{ total de pacientes})(\text{Amplitud del Intervalo})}$$

Estimación de la Función de Supervivencia

$$\hat{S}(t) = \frac{\# \text{ de pacientes con tiempos de supervivencia } \geq t}{\# \text{ total de pacientes}}$$

3.1. REGRESIÓN DE COX

La regresión de Cox es un método que permite estudiar datos de supervivencia, modelizando la función de azar de cada individuo como un producto de dos factores, dado que es interesante modelar no sólo la relación entre la tasa de supervivencia y el tiempo, sino también la posible relación con diferentes variables registradas para cada sujeto. Por lo tanto se trata de calcular la tasa de mortalidad como una función del tiempo y de las variables pronóstico.

$$h(t) = h_0(t) \cdot e^{(b_1x_1 + b_2x_2 + \dots + b_kx_k)}$$

El primer factor $h_0(t)$ depende del tiempo, y el segundo factor $e^{(b_1x_1 + b_2x_2 + \dots + b_kx_k)}$ dependen de las variables pronósticos que serán distintas para cada individuo. La función de supervivencia está dada: $S(t) = S_0(t) e^{(b_1x_1 + b_2x_2 + \dots + b_kx_k)}$, donde $S_0(t)$ es la función de supervivencia base. A esta función base se la puede estimar por métodos como el Kaplan-Meier, la distribución Weibull, etc.

3.2. DISTRIBUCIÓN DE WEIBULL

La función de densidad de la distribución Weibull está dada por:

$$f(x) = \begin{cases} \lambda\gamma \cdot x^{\gamma-1} e^{-\lambda x^\gamma} & \text{para } x > 0 \\ 0 & \text{en cualquier otra parte} \end{cases}$$

Esta función depende de dos parámetros λ y γ , evaluando la función para $\gamma=1$ la distribución Weibull se transforma en una distribución exponencial.

La supervivencia correspondiente a la función de riesgo de Weibull viene dada por la siguiente fórmula:

$$S(t) = \exp(-\lambda t^\gamma)$$

Nótese que calculando dos veces el logaritmo de la función de supervivencia, tendremos: $\ln[-\ln S(t)] = \ln \lambda + \gamma \ln t$, es decir las relaciones entre el logaritmo del logaritmo cambiando de signo de la supervivencia con el logaritmo de tiempo son lineales.

4. MODELO FRAMINGHAM

El modelo Framingham es utilizado para calcular el riesgo cardiovascular, el cual se basa en el modelo de Cox. El cálculo del riesgo Framingham utiliza las variables que intervienen en este modelo: **SEXO**, **EDAD**, el **COLESTEROL** sérico en mg/dl, fracción de colesterol ligado a lipoproteínas de alta densidad **HDL** en mg/dl, **PRESION SISTOLICA** en mmHg, **DIABETES** (No, Sí), **FUMADOR** (No, Sí), los cuales son factores de riesgo para problemas cardiovasculares.

Primero, se calcula el valor de la siguiente expresión, en base a los factores de riesgo mencionados anteriormente:

Para los hombres:

$$L_H = b_{E1} \cdot EDAD + b_C + b_H + b_T + b_D + b_F$$

Para las mujeres:

$$L_M = b_{E1} \cdot EDAD + b_{E2} \cdot EDAD^2 + b_C + b_H + b_T + b_D + b_F$$

Los **b** son los coeficientes correspondientes a cada variable, es por eso, que el modelo Framingham proporciona los valores de los coeficientes en base a las características de pacientes del país de origen (USA-MASSACHUSETTS). La tabla de coeficientes se presenta a continuación:

TABLA III
Coefficientes para el Modelo de Framingham

Coefficiente	Hombres	Mujeres
b_{E1} x Edad	0.04826	0.33766
b_{E2} x (Edad) ²	0	-0.00268
b_C Colesterol mg/dl		
< 160	-0.65945	-0.26138
160-199	0	0
200-239	0.17692	0.20771
240-279	0.50539	0.24385
≥ 280	0.65713	0.53513
b_H HDL- Col mg/dl		
< 35	0.49744	0.84312
35 - 44	0.24310	0.37796
45 - 49	0	0.19785
50 - 59	-0.05107	0
> 60	-0.48660	-0.42951
b_T Tensión arterial mmHg		
PAS < 120 PAD < 80	-0.00226	-0.53363
PAS < 130 PAD < 85	0	0
PAS < 140 PAD < 90	0.28320	-0.06773
PAS < 160 PAD < 100	0.52168	0.26288
PAS ≥ 160 PAD ≥ 100	0.61859	0.46573

b_D Diabetes		
NO	0	0
SI	0.42839	0.59626
b_F Fumador		
NO	0	0
SI	0.52337	0.29246

Fuente: Asociación de la Sociedad Española de Hipertensión
Elaboración: Luis M. Molinero

Una vez calculado el valor correspondiente de **L**, se le resta la cantidad **G** (función evaluada para los valores medios de las variables en el estudio) diferente para hombres o mujeres, donde: $G_{Hombres} = 3.0975$ y $G_{Mujeres} = 9.92545$

Exponenciamos ese valor calculado $B = \exp(L - G)$ y determinamos el valor de la expresión $1 - S^B$. **S** es la función de supervivencia base a 10 años, que es diferente para hombres y mujeres:

$$S_{Hombres} = 0.90015 \text{ y } S_{Mujeres} = 0.96246$$

5. UNIDADES DE ESTUDIO

Para llevar a cabo la investigación, se acudió al Hospital Luis Vernaza de la ciudad de Guayaquil. La Institución Médica nos facilitó una lista con los números correspondientes a fichas médicas de un periodo de 10 años (1994 - 2003). Se trató de obtener todos los datos y fichas de los pacientes con problemas cardiovasculares, cabe recalcar que existieron muchas fichas extraviadas, repetidas y algunas historias clínicas incompletas, lo cual complicó al momento de extraer toda la información necesaria para este tipo de estudio.

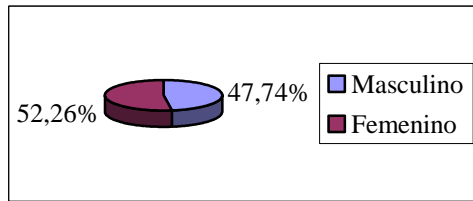
Se obtuvo 155 fichas médicas correspondientes a pacientes con re-ingreso, para de esta manera poder medir los días de reingresos por pacientes, este dato será fundamental para así encontrar la probabilidad de riesgo que estamos buscando.

Las variables utilizadas en este estudio de acuerdo a Framingham son: Edad, colesterol, HDL-colesterol, tensión arterial, diabetes y fumador.

5.1. ANÁLISIS DE VARIABLES

Los casos de estudio en la presente investigación, son pacientes con problemas cardiovasculares del Hospital Luis Vernaza de Guayaquil. De los pacientes que presentaron problemas cardiovasculares en el Hospital, tenemos que de un total de 155 pacientes: 74 eran hombres (47.74%) y 81 mujeres (52.26%).

GRÁFICO 1
Género de pacientes



Fuente: Validación y calibración del Modelo Framingham, tesis 2005.
Elaboración: Abel Flores

Siguiendo el modelo Framingham los valores medios de cada variable son:

TABLA IV
Valores Medios

Variables	Hombres	Mujeres
Edad	63,1081	60,6173
Colesterol	190,0676	209,9877
HDL- Colesterol	40,5270	42,3086
T.A.S.		
Sistólica	125,6486	131,5432
Diastólica	75,6757	77,9506
Diabetes		
No	0,6622	0,5926
SI	0,3378	0,4074
Fumador		
No	0,3919	0,8148
Si	0,6081	0,1852

Fuente: Validación y calibración del Modelo Framingham, tesis 2005.
Elaboración: Abel Flores.

Observando los resultados en la Tabla IV podemos concluir:

-El valor promedio de colesterol presentados por hombres y mujeres están cercanos a niveles de bajo riesgo (200-239 mg/dl).

-Los valores promedios HDL-Colesterol de los pacientes sometidos a estudio, no presentan mucha diferencia entre género.

-Los valores promedios de tensión se encuentran en niveles cercanos al riesgo, ya que sobrepasan el nivel recomendable (< 120/80 mmHg).

-De las personas sometidas a estudio, predominan las personas que no son diabéticas.

-En el caso de pacientes fumadores predominan las personas del sexo masculino, el 60.81% tienen el hábito de fumar, lo cual los hace más propensos a sufrir problemas cardiovasculares.

6. VALIDACIÓN Y CALIBRACIÓN DEL MODELO FRAMINGHAM

6.1. PASOS PARA LA VALIDACIÓN

1.- Se debe obtener los valores de (G_H, G_M) - Función evaluada para los valores medios de las variables de estudio.

Las funciones para el valor G están dadas por dos funciones:

$$G_H = bE_1 \cdot \hat{Edad} + b_C + b_H + b_T + b_D + b_F$$

$$G_M = b_{E1} \cdot \hat{Edad} + b_{E2} \cdot (\hat{Edad})^2 + b_c + b_H + b_T + b_D + b_F$$

Los coeficientes que se les asignarán a cada una de las variables, corresponden a los valores presentados en la Tabla III:

-Para hombres:

$$G_H = (0.04826 \cdot 63.1081) + 0 + 0.24310 + 0 + (0.3378 \cdot 0.42839) + (0.6081 \cdot 0.52337)$$

$$G_H = 3.7517$$

- Para Mujeres

$$G_M = (0.33766 \cdot 60.6173) + (-0.00268 \cdot (60.6173)^2) + 0.20771 + 0.37796 + 0 + (0.4074 \cdot 0.59626) + (0.1852 \cdot 0.29246)$$

$$G_M = 11.5032$$

2.- Obtener la probabilidad de supervivencia base a 10 años (S_{10})

Primero se estiman los valores de supervivencia condicionada (hombres y mujeres), para ello se construyó las siguientes tablas en base a las fórmulas expuesta en el punto 3.

TABLA V
Valores de la función de densidad y supervivencia para Hombres

Intervalo	Frec.	f(t)	S(t) Condicionado
(0-399]	50	0,00169	1
(399-798]	11	0,00037	0,32432
(798-1197]	5	0,00017	0,17568
(1197-1596]	3	0,00010	0,10811
(1596-1995]	1	0,00003	0,06757
(1995-2394]	2	0,00007	0,05405
(2394-2793]	1	0,00003	0,02703
(2793-3192]	1	0,00003	0,01351

Fuente: Validación y calibración del Modelo Framingham, tesis 2005.
Elaboración: Abel Flores.

TABLA VI
Valores de la función de densidad y supervivencia para Mujeres

intervalo	Frec.	f(t)	S(t) Condicionado
(0-340]	45	0,00163399	1
(340-680]	18	0,00065359	0,44444
(680-1020]	6	0,00021786	0,22222
(1020-1360]	5	0,00018155	0,14815
(1360-1700]	2	7,2622E-05	0,08642
(1700-2040]	3	0,00010893	0,06173
(2040-2380]	1	3,6311E-05	0,02469
(2380-2720]	1	3,6311E-05	0,01235

Fuente: Validación y calibración del Modelo Framingham, tesis 2005.
Elaboración: Abel Flores.

S(t) es condicionado porque los datos son en base a pacientes con re-ingreso, por este motivo existirán datos perdidos o censurados, puesto que existen personas que abandonaron el tratamiento, nunca más regresaron al Hospital, tal vez murieron o simplemente no sabemos que ha pasado con ellos. Por lo tanto, hacemos uso de técnicas matemáticas, para poder llevar a cabo el ajuste por medio de la función de Weibull.

Ante esto, se plantea la siguiente probabilidad:

$$P(S_{re-ingreso \leq T} > t) = P(S > t / S \leq T) = \frac{S(t) - c}{1 - c}$$

Donde **c** es una constante en la probabilidad ajustada condicionada, y reemplazando S(t) por la función de supervivencia de Weibull en la probabilidad ajustada condicionada, tendremos que:

$$P(S_{re-ingreso \leq T} > t) = \frac{e^{-\lambda \cdot t^\gamma} - c}{1 - c}$$

Basándose en esta probabilidad, se procede a hacer el ajuste a las tablas V y VI, por medio de regresión no lineal en MATLAB. Los resultados de los parámetros son:

-Hombres

$$C = 0.00002644 \quad R^2 = 0.9989$$

$$\lambda = 0.006508 \quad \gamma = 0.8672$$

- Mujeres

$$C = 0.005422 \quad R^2 = 0.9989$$

$$\lambda = 0.007537 \quad \gamma = 0.7829$$

Con esto, se reemplazó los parámetros en las funciones de supervivencia Weibull:

$$\text{S(t) hombre: } S(t) = e^{-0.006508t^{0.8672}}$$

$$S_{10} = 0.95$$

$$\text{S(t) mujeres: } S(t) = e^{-0.007537t^{0.7829}}$$

$$S_{10} = 0.96$$

Obtenidos los valores de G y S₁₀(t), podemos valorizar el modelo aplicando la formula para el cálculo de riesgo, la cual detallamos a continuación:

$$\left\{ \begin{array}{l} R = 1 - S^{\exp(L-G)}, \text{ donde:} \\ S = \text{función de Supervivencia base a 10 años} \\ L = (L_H \text{ o } L_M) \\ G = \text{(función evaluada para los valores} \\ \text{medios de las variables en el estudio)} \end{array} \right.$$

3.- Aplicación del Modelo

En la presente aplicación, se usarán los valores de G (G_H = 3.7517, G_M = 11.5032, S_H = 0.95 y S_M = 0.96) y los coeficientes (USA – Massachusetts).

Ejemplo:

-Validemos según este modelo el riesgo de un hombre de 58 años de edad, que es diabético, pero no fuma, tiene 150/90 mmHg de presión arterial, 208 mg/dl de colesterol total, 38 mg/dl de HDL.

Encontramos L_H de la siguiente forma:

$$L_H = b_{E1} \cdot EDAD + b_C + b_H + b_T + b_D + b_F$$

Por lo tanto, de la misma forma que hicimos en el cálculo del valor G, tenemos que:

$$L_H = (0.04826 \cdot 58) + 0.17692 + 0.24310 + 0.52168 + 0.42839 + 0$$

$$L_H = 4.16917$$

$$\text{Calculando } \exp(L_H - G_H), \text{ tenemos que:}$$

$$\exp(4.16917 - 3.7517) = 1.518163317$$

Luego, usando la función de supervivencia base a 10 años antes encontrada, el valor de

$$S(t) = e^{-0.006508t^{0.8672}} \text{ será:}$$

$$S_H = 0.95$$

Entonces la probabilidad que se produzca un incidente cardiovascular a 10 años será:

$$R = 1 - (0.95)^{1.518163317} = 0.0749168$$

$$\text{Riesgo} = 7.49 \%$$

Para efectos comparativos de resultados, en la tabla VIII se proporciona el riesgo medio a 10 años, de acuerdo al estudio Framingham:

TABLA VIII
Riesgo Medio a 10 años de acuerdo al estudio Framingham

Edad	Mujeres	Hombres
15-25	< 0.09 %	< 0.51 %
26-36	1.29 %	2.35 %
37-47	2.11 %	4.87 %
48-58	7.05 %	8.68 %
59-69	11.51 %	24.01 %
70-80	9.02 %	25.73 %
81-91	10.23 %	23.25 %

Fuente: Validación y calibración del Modelo Framingham, tesis 2005.
Elaboración: Abel Flores.

4.- Comparación de resultados

La Tabla VIII está basada en los coeficientes de USA-MASACHUSSETS, la cual debe ser comparada con la Tabla IX que contiene los valores de riesgo a 10 años según la tasa de re-ingreso (días) promedio ligados a los intervalos referentes a las edades de los casos de estudio, la cual, será presentada a continuación:

TABLA IX
Riesgo Medio a 10 años en base a las características de los casos de estudio.

Edad	Mujeres	Hombres
15-25	1.00 %	2.14 %
26-36	2.00 %	5.001 %
37-47	4.36 %	7.45 %
48-58	5.01 %	7.84 %
59-69	7.18 %	8.96 %
70-80	8.00 %	9.70 %
81-91	7.68 %	9.46 %

Fuente: Validación y calibración del Modelo Framingham, tesis 2005.
Elaboración: Abel Flores

Observando las dos tablas antes expuestas, podemos darnos cuenta que los resultados en base a los coeficientes del modelo FRAMINGHAM-USA no se ajustan a los valores de riesgo real en base a los casos de estudio; por este motivo, tenemos que encontrar los coeficientes para las características de los pacientes con problemas cardiovasculares del Hospital. Este proceso de hallar los coeficientes lo denominamos de calibración del modelo sometido a estudio que será explicado a continuación.

6.2. Calibración del Modelo

1.- Encontrar los coeficientes del Modelo.

En base al modelo Framingham, se parte de: $S(t) = S^B(1)$. El objetivo, es hacer lineal la función (1), para de esta forma hacer uso de las técnicas de regresión lineal múltiple y encontrar los coeficientes del modelo. Aplicando logaritmos en ambos lados en la función (1) quedaría:

$$\ln(-\ln(S(t))) = b_0 + \ln(B)$$

Dado que $B = \exp(L-G)$, reemplazando quedaría lo siguiente:

$$\ln(-\ln(S(t))) = b_0 + \ln(\exp(L-G))$$

$\ln(-\ln(S(t))) = b_0 + (L-G)$, y por género sería:

$$\begin{cases} b_0 + (L_H - G_H) \\ b_0 + (L_M - G_M) \end{cases}$$

– Consecuentemente, el modelo de regresión a utilizar para encontrar los coeficientes utilizando el criterio de estimación de mínimos cuadrados, sería:

- Hombres:

$$\begin{aligned} \ln(-\ln(S(t))) = & b_0 + b_{E1}(x_1 - \bar{x}_1) + b_C(x_2 - \bar{x}_2) + \\ & (2) \\ & + b_H(x_3 - \bar{x}_3) + b_T(x_4 - \bar{x}_4) + b_D(x_5 - \bar{x}_5) + b_F(x_6 - \bar{x}_6) \end{aligned}$$

- Mujeres:

$$\begin{aligned} \ln(-\ln(S(t))) = & b_0 + b_{E1}(x_1 - \bar{x}_1) + b_{E2}(x_2 - \bar{x}_2) + \\ & (3) \\ & + b_C(x_3 - \bar{x}_3) + b_H(x_4 - \bar{x}_4) + b_T(x_5 - \bar{x}_5) \\ & + b_D(x_6 - \bar{x}_6) + b_F(x_7 - \bar{x}_7) \end{aligned}$$

– Ahora faltaría encontrar el valor de $\ln(-\ln(S(t)))$, para de esta forma obtener los coeficientes $b_{E1}, b_{E2}, b_C, b_H, b_T, b_D, b_F$.

Para encontrar el valor de $\ln(-\ln(S(t)))$ usamos la función de supervivencia de la Weibull ($S(t) = e^{-\lambda t^\gamma}$) y le aplicamos logaritmo en ambos lados, con lo cual tendremos:

$$\ln(-\ln(S(t))) = \ln \lambda + \gamma \cdot \ln(t)$$

Entonces:

$$\ln(-\ln(S(t))) = \begin{cases} \text{hombres} & \ln(0.006508) + (0.8672) \cdot \ln(t) \\ \text{mujeres} & \ln(0.007537) + (0.7829) \cdot \ln(t) \end{cases}$$

Se ha encontrado los valores del lado izquierdo del modelo (2) y (3), para poder hallar los coeficientes.

2.- Recodificación de la variables

Se hace uso de las variables indicadoras “dummy”, para recodificar las variables, debido a que si observamos las variables: Colesterol, HDL Colesterol y tensión arterial, estas están

compuestas por 5 rangos en base a los niveles de las mismas: por ejemplo:

Rango (Niveles de Colesterol)	Rango (Niveles de HDL)	Rango (Niveles de T.A.S.)
<160	<35	PAS < 120 PAD <80
160-199	35-44	PAS < 130 PAD <85
200-239	45-49	PAS < 140 PAD <90
240-279	50-59	PAS < 160 PAD <100
≥ 280	≥ 60	PAS ≥ 160 PAD ≥ 100

Por consiguiente es necesario tener 5 coeficientes para cada variable antes mencionada, así obtener la tabla de estos en base a nuestros datos. Esta recodificación se la realiza utilizando las variables indicadoras para:

Generar tantas variables como el total de la categoría menos uno, obviamente esto se puede efectuar cuando una variable presente más de dos categorías. Cada nueva variable tomará valor 1 para un determinado nivel y 0 el resto, de tal forma que los individuos en una misma categoría tomarán valor 1 en una misma variable y 0 en el resto.

La categoría no considerada o categoría referencia, estará representada por el valor 0 en todas las nuevas variables. En nuestro caso los niveles normales en colesterol, HDL y T.A.S., son considerados como referencia para los demás. Mediante este esquema de codificación, los coeficientes de las nuevas variables reflejarán el efecto de las categorías representadas respecto al efecto de la categoría referencia. Por ejemplo las variables quedarían codificadas de la siguiente manera:

	x1	x2	x3	x4
bC1	1	0	0	0
bC2	0	0	0	0
bC3	0	1	0	0
bC4	0	0	1	0
bC5	0	0	0	1

En la tabla X se muestran los coeficientes obtenidos, utilizando los modelos de regresión lineal múltiple (2) y (3), aplicando las variables indicadoras.

3.- Resultados

TABLA X
Coefficientes para el modelo de Framingham en base a los datos sometidos a estudio

Coefficiente	Hombres	Mujeres
b_{E1} x Edad (bE1)	0.004144	0.0149
b_{E2} x (Edad)² (bE2)	0	-0.000102
b_C Colesterol mg/dl		
< 160 (bC1)	-0.273	-0.507
160-199 (bC2)	0	0
200-239 (bC3)	0.114	0.109
240-279 (bC4)	0.218	0.115
≥ 280 (bC5)	0.407	0.225
b_H HDL-Col mg/dl		
< 35 (bH1)	0.0367	0.600
35 – 44 (bH2)	0.01500	0.341
45 – 49 (bH3)	0	0.258
50 – 59 (bH4)	-0.0583	0
≥ 60 (bH5)	-0.291	-0.212
b_T Tensión arterial mmHg		
PAS < 120 PAD < 80 (bT1)	-0.717	-0.08040
PAS <130 PAD < 85 (bT2)	0	0
PAS <140 PAD < 90 (bT3)	0.05122	-0.0149
PAS < 160 PAD < 100 (bT4)	0.332	0.356
PAS ≥160 PAD ≥100 (bT5)	0.589	0.401
b_D Diabetes		
NO	0	0
SI	0.118	0.391
b_F Fumador		
NO	0	0
SI	0.09756	0.002838

Fuente: Validación y calibración del Modelo Framingham, tesis 2005.
Elaboración: Abel Flores.

4.- Evaluación en base a los coeficientes (casos de estudio).

El objetivo de esta evaluación, es efectuar los cálculos del modelo Framingham con los coeficientes obtenidos en la Tabla X.

Por este motivo las funciones para el valor G (hombres y mujeres) estarán dadas por las siguientes funciones:

- Para Hombres

$$G_H = 0.3757066$$

- Para Mujeres

$$G_M = 1.13822215$$

Se hace uso de los valores de supervivencia base a 10 años encontrados en la sección 6.1 (paso 2) que son:

$$S_H = 0.95 \quad \text{y} \quad S_M = 0.96$$

-Ejemplo:

-Validemos según este modelo el riesgo de un hombre de 58 años de edad, que es diabético, pero no fuma, tiene 150/90 mmHg de presión arterial, 208 mg/dl de colesterol total, 38 mg/dl de HDL

Encontramos L_H de la siguiente forma:

$$L_H = b_{E1} \cdot EDAD + b_C + b_H + b_T + b_D + b_F$$

Al igual que se hizo en el cálculo del valor G, tenemos que:

$$L_H = (0.004144 \cdot 58) + 0 + 0.01500 + 0.332 + 0.118 + 0$$

$$L_H = 0.705$$

Calculando $\exp(L_H - G_H)$, se tiene que:

$$\exp(0.705 - 0.3757066) = 1.389$$

La probabilidad de un evento cardiovascular a los 10 años será:

$$R = 1 - (0.95)^{1.389} = 0.063$$

$$\text{Riesgo} = 6.3 \%$$

7. CONCLUSIONES

1. De los casos estudiados con problemas cardiovasculares (155 registros), la proporción entre hombres (47.74%) y mujeres (52.26%) son aproximadamente iguales.
2. El valor promedio de colesterol presentados por los hombres (190.0676 mg/dl) y mujeres (209.9877 mg/dl). Los hombres están cercanos a (200-239 mg/dl) nivel de bajo riesgo, y en el caso de las mujeres están en el nivel citado.
3. Los valores promedio de HDL-Colesterol presentados por los hombres (40.5270 mg/dl) y mujeres (42.3086 mg/dl), no presentan mucha diferencia entre género.
4. Los valores promedio de tensión arterial presentados por los hombres (125.6486/75.6757 mmHg) y mujeres

(131.5432/77.9506 mmHg) se encuentran cercanos al riesgo ya que sobrepasan el nivel recomendable (< 120/80 mmHg).

5. En los casos de estudio con problemas cardiovasculares (74 hombres y 81 mujeres) predominan las personas que no son diabéticas, tanto hombres (66.22%) y mujeres (59.26%).
6. En el caso de los pacientes fumadores predominan las personas del sexo masculino, el 60.81% tienen el hábito de fumar, lo cual los hace más propensos a sufrir problemas cardiovasculares.
7. Los casos de estudio analizados son de pacientes con re-ingresos. por esto existen datos censurados. Para determinar la Supervivencia Real se utilizó un ajuste de Weibull condicionado, donde la probabilidad de supervivencia real no presenta mucha diferencia con la probabilidad de supervivencia condicionada ajustada.
8. Los valores del R^2 en el ajuste, para encontrar los parámetros de la función de supervivencia de la Weibull Condicionada son aceptables: hombres ($R^2 = 0.9989$) y mujeres ($R^2 = 0.9989$).
9. Se calibró el modelo Framingham debido a que el riesgo medio a 10 años en base a los coeficientes (USA-Massachusetts), sobrestima a los valores de riesgo real a 10 años en función a las características de los casos de estudio (mediante el ajuste Weibull).
10. Existe una gran variación entre los valores G según las características de (USA-Massachusetts, $G_H = 3.7517$ y $G_M = 11.5032$) y los casos de estudio ($G_H = 0.3757$, $G_M = 1.1382$), la cual, posiblemente se origina por los diferentes estilos de vida entre países (USA-ECUADOR).
11. La probabilidad de no reingresar al Hospital por problemas cardiovasculares (base a 10 años) es aceptable referente a los casos de estudio: hombres ($S_H = 0.95$) y mujeres ($S_M = 0.96$).
12. A mayor edad la probabilidad de re-ingresar al hospital por problemas cardiovasculares aumenta, dependiendo de los factores de riesgo del paciente.

13. Según los resultados se afirma que el hombre tiene mayor probabilidad de re-ingresar al hospital por problemas cardiovasculares que la mujer.
14. Se comprobó, que la probabilidad de re-ingresar al hospital (problemas cardiovasculares) según las características de los casos analizados, es baja con respecto a los parámetros de USA-MASSACHUSETTS.

REFERENCIAS BIBLIOGRÁFICAS

1. F. Louzada, J. Mazucheli y J. Achcar, "Análise de Sobrevivencia e Confiabilidade". (Instituto de Matemáticas y Ciencias Afines, IMCA, 2002), pp. 24-27
2. John E., I. Millar y M. Millar, "Estadística Matemática con Aplicaciones", (6ta. Edición, México, Pearson Educación, 2000), pp. 449-484
3. L. Zapata, "Factores predictorios de sobre vida en pacientes con melona, mediante el modelo regreseión de Cox" (Tesis, Instituto de Ciencias Matemáticas, Escuela Superior Politécnica del Litoral, 2004)
4. L. M. Molinero, "Modelos de riesgo cardiovascular – Estudio Framingham" (Asociación de la Sociedad Española de Hipertensión, 2003)
5. L. M. Molinero, "Variables Indicadoras Dummy" (Asociación de la Sociedad Española de Hipertensión, 2003)
6. L. M. Molinero, "Función de Supervivencia de la función Weibull" (Asociación de la Sociedad Española de Hipertensión, 2004)