SELECCIÓN DE MODELOS BAJO EL ENFOQUE BAYESIANO: UNA APLICACIÓN AL ESTADO COGNITIVO DE LOS ADULTOS MAYORES EN EL ESTADO DE GUERRERO.

G.L. Díaz*,V. Sistachs Vega**, M. D. Covarrubias*, N. I. Hernández ***
[1]Unidad Académica de Matemática, Universidad Autónoma de Guerrero, México.
[2]Facultad de Matemática y Computación. Universidad de La Habana, Cuba.
Unidad Académica de Enfermería no.1, Universidad Autónoma de Guerrero.

ABSTRACT
The uncertainty problem is inherent to every statistical model and linked to it is the model selection topic. This paper presents a procedure for selection model in the presence of uncertainty called BMA (Bayesian Model Averaging) applied to logistic regression, proposed by Raferty (1995). For R implementation, we retook the ideas of Raftery, Painter and Volinsky(2005), such as the ones of Saminni and Parmeter(2011), said procedure is used to make the study analysis over the cognitive state of elders in Guerrero, Mexico. It is proposed the BMA as an alternative to take into account the models on this type of study.

KEYWORDS: BMA, model selection, logistic regression.

MSC: 62P15

RESUMEN
La problemática de la incertidumbre es inherente a todo modelo estadístico y vinculado a ella está el tema de selección de modelo. En este trabajo presentamos un procedimiento para la selección de modelos en presencia de incertidumbre llamado BMA (Bayesian Model Averaging) aplicado a regresión logística, propuesto por Raftery(1995). Para la implementación en R, se retoman las ideas de Raftery, Painter y Volinsky(2005), así como Saminni y Parmeter(2011), dicho procedimiento es utilizado para hacer el análisis del estudio sobre el estado cognitivo de los adultos mayores en Guerrero, México. Se propone el BMA como una alternativa para tomar en cuenta la incertidumbre de los modelos en este tipo de estudios.

1. INTRODUCCIÓN.

El tema de la incertidumbre inherente a todo modelo estadístico pocas veces es tratado explícitamente en las aplicaciones. Se puede decir que el procedimiento de la modelación consiste de dos fases, estimar un modelo y validar dicho modelo, después de haber pasado las pruebas con éxito se considera el modelo listo para su aplicación, pero sobre todo se asume como el modelo verdadero.

Muy relacionado con el tema de la incertidumbre está el problema de la selección de modelos, que según Gelfand y Dey (1994), tiene dos aspectos uno referido a si el modelo es adecuado y el otro, a ¿Cuál es el mejor modelo?, dentro de una colección bajo consideración.

Se define un modelo como una especificación de una distribución de cantidades observables (los datos) y no observables (los parámetros del modelo, observaciones perdidas, etc) y esta definición puede ser enfocada desde una perspectiva bayesiana. En el enfoque bayesiano los parámetros y los modelos son considerados aleatorios (f(y/Mi)) y expresan su incertidumbre en términos de distribución de probabilidad.

Entre los diferentes métodos bayesianos de selección de modelos están, los Factores de Bayes (FB), como un método para seleccionar entre dos posibles modelos y para el caso más general (más de 2 modelos) se utiliza el método BMA donde se habla de promediar los modelos (ver Claeskens, G. and Hjort, N. L.2008), también existen otros criterios como el AIC, BIC, etc.(ver Kadane, and Lazar 2004)

En el trabajo presentamos un método, el Bayesian Model Averagind que proporciona una vía formal para tomar en cuenta la incertidumbre en la selección de modelos. Ilustramos el método con una aplicación a un estudio de corte transversal investigando los factores de riesgo asociados al problema del estado cognitivo en adultos mayores. En el epígrafe 2 se muestra el uso del método de BMA para la selección de modelos bajo el paradigma bayesiano, así como una implementación para utilizarlo usando el software R. En el epígrafe 3 se presenta la aplicación de este método de selección en un estudio sobre el estado cognitivo de los adultos mayores en Guerrero, México y por último en el epígrafe 4 presentamos la discusión de los resultados.

2. MÉTODO BMA PARA LA SELECCIÓN DE MODELOS BAJO EL PARADIGMA BAYESIANO

La selección de variables ha sido reconocida como uno de los problemas más difundidos en la selección de modelos en aplicaciones estadísticas (George, 2000) y una gran cantidad de métodos han surgido durante los últimos 30 años, especialmente en el contexto de la regresión lineal (ver Miller, 1990, McQuarriet & Tsai, 1998, George 2000). Muchos investigadores se han enfocado en desarrollar diferentes criterios apropiados para la selección de modelos, tales como PRESS (Allen, 1971), Cp de Mallows (Mallows, 1973), Criterio de Akaike AIC (Akaike, 1973), Criterio de Información de Schwarz BIC (Schwarz, 1978), RIC de Foster y George (1994), Selección de modelos Bootstrap (Shao, 1996), aunque en la práctica se asume que hay disponible pocos modelos razonables. De cualquier forma los investigadores en realidad tienen que elegir uno o pocos mejores modelos de la enorme cantidad de potenciales modelos usando técnicas tales como Regresión Stepwise de Efroymson (1960) y sus diferentes variaciones, o por ejemplo, el algoritmo *leap-and-bounds* de Furnival y Wilson (1974).

Típicamente los investigadores usan ambos desarrollos, primero tratan de generar varios mejores modelos para diferentes números de variables y entonces seleccionar el modelo con mejor dimensión de acuerdo a uno de los criterios listados. Sin embargo, cualquier combinación de estos desarrollos para la selección de modelos no parece tener en cuenta la incertidumbre asociada con la selección de modelos y por lo tanto en la práctica se tiende a producir sesgos en las estimaciones y los procedimientos para la selección de variables son sospechosos (Lipkovich, 2002).

Los dos aspectos dos aspectos relacionados con el problema de la selección de modelos (la búsqueda de modelos y el criterio para la selección de modelos) son integrados con naturalidad en el modelo de promedios, el cual supera la deficiencia inherente de la selección de modelos determinista combinando (promediando) información de todos o un subconjunto de modelos cuando se hace estimación, inferencia o predicciones, en vez de usar sólo un modelo.

El BMA se está volviendo una herramienta de análisis de datos cada vez más popular que le permite a los investigadores tomar en cuenta la incertidumbre asociada con el proceso de la selección de modelos.

Muchas aplicaciones el BMA están relacionadas con el espacio de modelos confinados para alguna subclase especial, por ejemplo, hay aplicaciones del BMA para modelos gráficos (Madigan y Raftery, 1994), árboles de regresión (Chipman et al., 1998),regresión multivariada (Brown y Bannucci, 1998; Noble, 2000) y análisis de sobrevivencia (Volinsky, 1997) por mencionar algunos.

El enfoque bayesiano como se ha planteado ya, permite expresar la incertidumbre en términos de probabilidad, y bastan las reglas básicas del cálculo de probabilidades para poder hacer inferencias , el BMA (Bayesian Model Averaging) no es más que estadística bayesiana básica (ver Ando, T 2010). El BMA combina la predicción y estimación de parámetros obtenidos con diferentes modelos plausibles usando sus probabilidades a posterior como sus pesos.

De acuerdo con Madigan y Raftery (1994), si θ es la cantidad de interés, tal como un parámetro del modelo de regresión o una observación futura, entonces su distribución a posterior dados los datos D y un conjunto de K modelos es la mezcla de distribuciones a posterior (ver Leamer, 1972).

Así, como consecuencia de la regla o teorema de la probabilidad total, la probabilidad final BMA de θ viene dada por:

$$p(\theta|D) = \sum_{k=1}^{K} p\,(\theta|D, M_k) p(M_k|D) \qquad (1)$$

Siendo $p(\theta|D, M_k)$ distribución de probabilidad final de θ, dado el modelo M_k y los datos D, y $p(M_k|D)$la distribución de probabilidad final de M_k, tomado como el modelo verdadero, considerando que uno de los modelos propuestos es el verdadero.

La probabilidad final del modelo M_k, està dada por:

$$p(M_k|D) = \frac{p(D|M_k)p(M_k)}{\sum_{i-1}^{K} p(D|M_i)p(M_i)} \quad (2)$$

En esta expresión, $p(M_k|D)$ (2) es la integral de la función de verosimilitud del modelo M_k, resultado de integrar sobre los parámetros del modelo, es decir:

$$p(D|M_k) = \int p(D|\theta_k, M_k)p(\theta_k|M_k)\, d\theta_k \quad (3)$$

Siendo $\theta_k$el o los parámetros del modelo M_k y $p(D|\theta_k, M_k)$ la función de verosimilitud de θ_k para el modelo M_k , y $p(\theta_k|M_k)$ la probabilidad inicial de θ_k . Las probabilidades iniciales suelen considerarse iguales. Para calcular la integral en (3) utiliza la simple y precisa aproximación del BIC:

$$2\log p(D|M_k) \approx 2log p\left(D\middle|\hat{\theta}_k\right) - d_k \log(n) = -BIC_k \quad (4)$$

Donde $d_k = dim(\theta_k)$ es el número de parámetros independientes en M_k, y $\hat{\theta}_k$ es el estimador de máxima verosimilitud. Para regresión lineal, el BIC tiene la forma simple

$$BIC_k = nlog(1 - R_k^2) + p_k \log(n) \quad (5)$$

Donde R_k^2 es el valor de R^2 y p_k es el número de regresores para el k-ésimo modelo de regresión. Por (5), $BIC_k = 0$ para el modelo nulo sin variables regresoras.
Cuando nuestro interés se centra en los parámetros del modelo, digamos parámetros de regresión tal como β_1, (1) puede ser aplicado con $\Delta = \beta_1$. La media posterior del BMA de β_1 es justo un promedio de los pesos de las medias a posterior bajo cada uno de los modelos:

$$E[\beta_1|D] = \sum \tilde{\beta}_1^{(k)} p(M_k|D) \quad (6)$$

El cual se puede ver como un estimador puntual del modelo de promedios bayesianos. En (6), $\tilde{\beta}_1^{(k)}$ es la media posterior de β_1 bajo el modelo M_k y este se puede aproximarse por su correspondiente estimador de máxima verosimilitud $\hat{\beta}_1$ (Rafftery, 1995). Una expresión similar es posible para la desviación estándar a posterior, el cual puede verse como un error estándar del modelo de promedios bayesianos.
En la implementación del BMA existen dos dificultades: primero el cálculo de la integral en (3) y segundo promediar sobre todos los modelos cuando el número de modelos es grande como en (1) y (6). Para ello la integral de verosimilitud es aproximada por la aproximación del BIC (ec.4). La suma sobre todos los modelos es aproximada encontrando el mejor modelo usando el algoritmo *fast leaps and bounds* que fue introducido por Raftery(1995). Finalmente los modelos que son menos verosímiles a posterior que el mejor modelo son excluidos. Esta es una exhaustiva búsqueda para encontrar el modelo global óptimo.

2.1- Implementación del BMA en R

Para la implementación de la selección de modelos en el paquete estadístico R, se hace uso de la librería BMA, está permite aplicar la selección a modelos lineales, a modelos lineales generalizados y a modelos de sobreviviencia, además incluye funciones que permiten mostrar los resultados gráficamente (Raftery, Painter & Volinsky [2005], Amini y Parmeter, [2011]).
Como se mencionó en la sección anterior, el procedimiento BMA tiene dos dificultades:

1. Evaluar la integral para todos los modelos en (3), y
2. Promediar sobre todos los modelos, para obtener (1) y (6).

Para la selección de variables en modelos lineales generalizados, la función que debemos usar es **bic.glm** (para más detalles ver Raftery, Painter, y Volinky, [2005]; Amini y Parmeter, [2011]) y la integral es aproximada mediante el criterio de información Bayesiano BIC.

La suma sobre sobre todos los modelos posibles se aproxima mediante el algoritmo *leaps and bounds*. Este algoritmo fue propuesto por Furnival y Wilson (1974) para la selección de variables en regresión y ha sido aplicado en modelos linelaes, en modelos linelaes generalizados por Raftery (1995), y por último en modelos de sobrevivencia por Volinsky et al. (1997). Este algoritmo descarta los modelos con probabilidades finales menos verosímiles, encontrando el modelos globalmente óptimo.
Si el número de variables es muy grande, el algoritmo *leaps and bounds* puede hacerse notablemente lento. En estos casos, se puede acelerar el proceso de búsqueda modificando el valor por defecto de *maxCol*, (está establecido en 31 columnas). Si el número de variables es superior, entonces se procede por eliminación hacia atrás por etapas (backwards, stepwise) antes de aplicar *leaps and bounds*. Tratándose del caso de los modelos generalizados con probabilidad inicial conocida, se dispone también de la función *glib*, que aproxima la integral de la función de verosimilitud por el método de Laplace, Raftery (1996). Podemos ver el uso de esta función aplicada a un estudio de casos y controles en epidemiologia en Villefont, (2001).
El paquete BMA realiza el análisis asumiendo una distribución uniforme como modelo a priori y utiliza la aproximación del BIC (Bayesian Information Criterion) para construir las probabilidades a priori de los coeficientes de regresión (Raftery, Hoeting, Volinsky, Painter & Yeung, 2010). Además esta librería se construyó con base en el algoritmo de Raftery(1995).

3. ESTUDIO SOBRE EL ESTADO COGNITIVO DE LOS ADULTOS MAYORES EN GUERRERO, MÉXICO

El aumento en las expectativas de vida ha tenido implicaciones importantes para los sistemas de salud en el ámbito mundial. Las proyecciones señalan que entre 1980 y 2050, la expectativa de vida para las personas mayores de 60 años aumentará 77% [2]. Con ello incrementarán las enfermedades asociadas con la edad entre las que el deterioro cognoscitivo representa una condición que afecta de manera directa la calidad de la población adulta mayor y determinan un mayor uso de los servicios de salud (Banco Mundial, 1993).
El envejecimiento de la población implica una mayor demanda de servicios de salud. En este grupo de edad cada vez se presentan mayores tasas de morbilidad y necesidades de atención médica que en el resto de la población. Al mismo tiempo, los padecimientos de la población en edades avanzadas tienden a concentrarse en males crónico-degenerativos.
Las principales causas de muerte a nivel nacional de las personas de la tercera edad de ambos sexos en el año 2000 fueron las enfermedades cardiovasculares, neoplasias malignas, diabetes mellitus, enfermedades digestivas, respiratorias, del hígado y accidentes. En el año 2003 el Congreso del Estado aprobó la creación del Programa Pensión Guerrero, cuyo objetivo es apoyar económicamente a los adultos mayores de 65 años en los municipios de Acapulco de Juárez, Chilpancingo de los Bravo, Iguala de la Independencia, Taxco de Alarcón y José Azueta.
En 2004 se realizó un estudio entre la Secretaria de Desarrollo Social y la Escuela de Enfermería no. 1 de la Universidad Autónoma de Guerrero y uno de sus objetivos era evaluar el estado de saludos de los Adultos Mayores. Algunos de los indicadores obtenidos en ese estudio fueron: estado nutricional, seguridad social y accesibilidad a los servicios, vivienda, aspecto laboral, capacidad funcional, estado de salud, disfunciones físicas, estado cognitivo y estado anímico.
El estado de salud de los adultos mayores está asociado a distintos factores que influyen de manera sustancial en la calidad de vida que éstos puedan tener. Las variables que se analizaran en este estudio son las siguientes, ya que se consideran factores de riesgo para estar afectado en el Estado Cognitiva del adulto mayor:

Variables consideradas para el análisis:
1. (mpio) municipio (Acapulco(1), Chilpancingo(2), Iguala(3), Taxco(4), José Azueta(5))
2. (edad) Edad
3. (sexo) Sexo (Masculino (1) y Femenino (0))
4. (edo_civi) Estado Civil (Soltero (1), Unión libre (2), casado(3), Divorciado(4), Viudo(5))
5. (poblaci) Población (Urbana (1), Rural (0))
6. (esca_imc) Escala imc (Sobrepeso (1) y normal (0)
7. (edo_lab) Estado laboral (Trabaja (1), no trabaja (0))
8. (edo_sano) Estado de salud (Sano (1) y Enfermo (0))
9. (ABVD) Actividades básicas de la vida diaria (Dependiente(1), Independiente(0)), esta se obtuvo por medio de la escala de Kazt.

10. (AIVD) Actividades instrumentales de la vida diaria (Dependiente(1), Independiente(0)), esta se obtuvo por medio de la escala de Lawton Brody.
11. (EEC) escala estado cognitivo (Afectado(1), no afectado (0)), se obtuvo por medio de la prueba Pfeiffer.
12. (EEA) escala estado anímica (Afectado(1), no afectado (0)), se obtuvo por medio de Yesavage.

4. RESULTADOS Y DISCUSIÓN

El modelo que se utilizó fue el Bayesian Model Average (BMA) en regresión logística binaria para obtener un modelo que me permita calcular la probabilidad de que un adulto mayor se vea afectado del estado cognitivo, a partir de las variables que se consideraban factores de riesgo.
Lo primero que se estableció fue las distribuciones a priori de los parámetros vector β y σ^2 $q$ue se consideraron no informativas, es decir p(β,τ) ∝ τ-1 , donde $\tau = \frac{1}{\sigma^2}$
La distribución posterior del modelo con esas a priori seria

$$p(\beta,\tau|z) = N_p\left(\beta\middle|\hat{\beta}_w,\tau^{-1}(X^tWX)^{-1}\right)Ga\left(\tau\middle|\frac{n-p}{2},\frac{n-p}{2}\hat{\sigma}^2\right)$$

donde

$\hat{\beta}_w, = (X^tWX)^{-1}X^tWz,\ \hat{\sigma}^2 = \frac{1}{n}\left(z - X\hat{\beta}_w\right)^t\left(z - X\hat{\beta}_w\right)$

y W matriz diagonal donde $w_{ii} = \hat{\pi}_i(1-\hat{\pi}_i)$

Para resolver el problema se utilizó el paquete BMA que está en el lenguaje R (Raftery, et al.), y cuyas instrucciones para correr el BMA en este modelo aparecen en el siguiente cuadro.

Programa 1: BMA en regresión logística

```
*******ESTADO COGNITIVO DE LOS ADULTOS MAYORES*********
************************************************************

library(MASS)
library(splines)
library(survival)
library(leaps)
library(BMA)
datos<-read.table(base pension guerrero pocas variables.txt, header=T)
y<- datos$EEC
x<- data.frame(datos[,-11])
x$mpio<- as.factor(x$mpio)
x$sexo<- as.factor(x$sexo)
x$edo.civi<- as.factor(x$edo.civi)
x$poblaci<- as.factor(x$poblaci)
x$edo_lab<- as.factor(x$edo_lab)
x$es_sano<- as.factor(x$es_sano)
x$esca_imc<- as.factor(x$esca_imc)
x$ABVD<- as.factor(x$ABVD)
x$AIVD<- as.factor(x$AIVD)
#x$EEC<- as.factor(x$EEC)
x$EEA<- as.factor(x$EEA)

glm.out.FT<- bic.glm(x, y, glm.family=binomial)
summary(glm.out.FT)
plot(glm.out.FT,mfrow=c(3,3))
imageplot.bma(glm.out.FT)
```

El programa selecciono 18 modelos de los cuales en la **Tabla 1** se muestran sólo los 5 mejores que tiene una probabilidad a posteriori acumulada del 1.00, además en la tabla por columna con los nombre de la constante y las variables utilizadas en el problema, aparece otro bloque donde aparece *p!=0*, *EV y* SD, los cuales son, porcentaje las probabilidades finales de las variables para estar en el modelo ideal, EV que muestra los valores esperados BMA finales de los coeficientes y bajo las siglas SD las desviaciones estándar BMA finales ara cada coeficiente. En las siguientes columnas aparecen los coeficientes estimados de las variables que se incluyen en cada uno de los respectivos modelos. Al final se muestra el número de variables incluidas en los modelos, el R^2, el BIC y la probabilidad final del modelo.

	p!	**EV**	**SD**	**Modelo 1**	**Modelo 2**	**Modelo 3**	**Modelo 4**	**Modelo 5**
Intercepto	100	-1.6163	1.60196	-2.545e+00	-2.634e+00	2.458e-01	3.128e-01	-2.681e+00
Mpio								
2								
3								
4								
5								
edad	66	0.02580	0.02186	3.824e-02	4.029e-02	-	-	3.961e-02
sexo	100	-	-	-	-	-	-	-
1		-0.812	0.19509	-8.076e-01	-8.551e-01	-7.672e-01	-8.189e-01	-8.555e-01
Edo.civil	0	-	-	-	-	-	-	-
2	-	-	-	-	-	-	-	-
3	-	-	-	-	-	-	-	-
4	-	-	-	-	-	-	-	-
5	-	-	-	-	-	-	-	-
poblacion	100	-	-	-	-	-	-	-
1		-0.75012	0.17995	-7.487e-01	-7.597e-01	-7.410e-01	-7.555e-01	-7.462e-01
Esca_imc	2.5	-	-	-	-	-	-	-
1		0.005823	0.05241	-	-	-	-	2.306e-01
ABVD	65.7	-	-	-	-	-	-	-
1		0.518021	0.44374	7.716e-01		8.158e-01		7.791e-01
AIVD	100	-	-	-	-	-	-	-
1		0.825072	0.21492	7.344e-01	8.605e-01	8.661e-01	1.015e+00	7.553e-01
EEA	-	-	-	-	-	-	-	-
1	-	-	-	-	-	-	-	-
nvar				5	4	4	3	6
BIC	-	-	-	-2.999e+03	-2.998e+03	-2.998e+03	-2.996e+03	-2.993e+03
Post Prob	-	-	-	0.385	0.250	0.246	0.093	0.025

Tabla 1 Resultados de la corrida del paquete BMA en R, en el problema de la eclampsia**.**

En la tabla 1 se puede observar que hay tres variables que se incluyen en los 5 modelos, para el primer modelo las variables más importantes son actividades instrumentales de la vida diaria (preparar comida, manejar dinero, hacer compras, usar el teléfono, etc.), sexo y el tipo de población (urbana) ya que además de tener la probabilidad de inclusión más alta aparecen en todos los modelos, le sigue en importancia la edad y las actividades básicas de la vida diaria (caminar, bañarse, comer ponerse los zapatos, etc.) con 65% y 66% de probabilidad de inclusión en el modelo.
Por lo tanto el modelo quedaría de la siguiente manera

$$P(y=1) = \frac{exp[-2.54 + 0.0038(edad) - 0.080(sexo) - 0.080(poblaci) + 0.077(abvd) + 0.073(aivd)]}{1 - exp[-2.54 + 0.0038(edad) - 0.080(sexo) - 0.080(poblaci) + 0.077(abvd) + 0.073(aivd)]}$$

si tenemos una adulto mayor que sea mujer con edad de 70 años, que en zona urbana, que está afectada de su capacidad funcional (es decir afectada de sus AIVD y ABVD) la probabilidad de que se vea afectado de su estado cognitivo es de 11% de que se vea afectado de su estado cognitivo.
En la **Imagen1** se muestra las distribuciones finales BMA de los distintos coeficientes del ***modelo 1*** es el resultado de (1) tras hacer $\theta = \beta_i$.

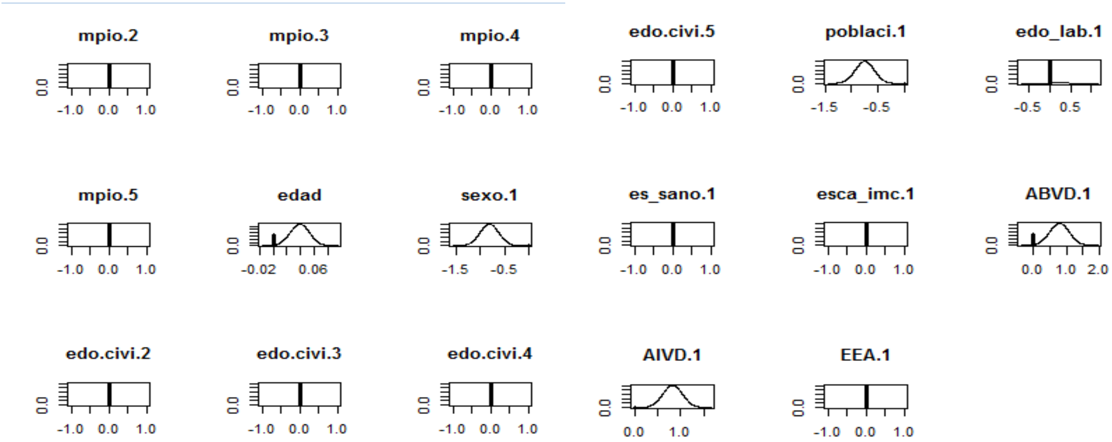

Imagen 1. Distribuciones finales del BMA

En la imagen 1 podemos ver la distribución final de las variables que se analizan y aquellas que quedaron incluidas en el modelo 1 son las que tiene forma de campana. Si observamos por ejemplo la función de densidad final del coeficiente de la variable ABVD, tenemos la siguiente imagen, ver imagen 2.

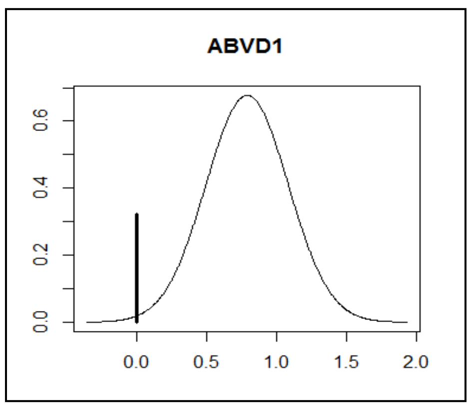

Imagen 2. Distribución final BMA de la variable ABVD

En la imagen 2 se puede apreciar que el máximo de la función corresponde a la probabilidad final de que la variable esté incluida en el modelo, es decir, la probabilidad de $\beta_4 = 0.657$, mientras que la barra vertical trazada en 0.00, representa la probabilidad de que dicha variable no se incluya en el modelo, este es, $\beta_4 = 0$, dicha probabilidad es 1-0.657=0.343, La función es el resultado de una mezcla de densidades normales y está escalada de manera que el máximo de la función se corresponda con la probabilidad final de que dicha variable este incluida en el modelo óptimo.
En la Imagen 3 se puede apreciar la inclusión de las variables (eje de las ordenadas) en los modelos obtenidos con el BMA (eje de las abscisas), con la particularidad de que la amplitud de las columnas representa de manera proporcional la probabilidad final del modelo y se colorea la parte correspondiente a la variable que se incluyó en dicho modelo, por ejemplo, en el modelo 1 se encuentran incluidas las variables edad (años), sexo, población actividades básicas y actividades instrumentales, los colores indican el signo del coeficiente en el modelo, siendo el azul el color que representa un valor positivo en los coeficientes y el color rojo representa un valor negativo en dicho coeficiente.

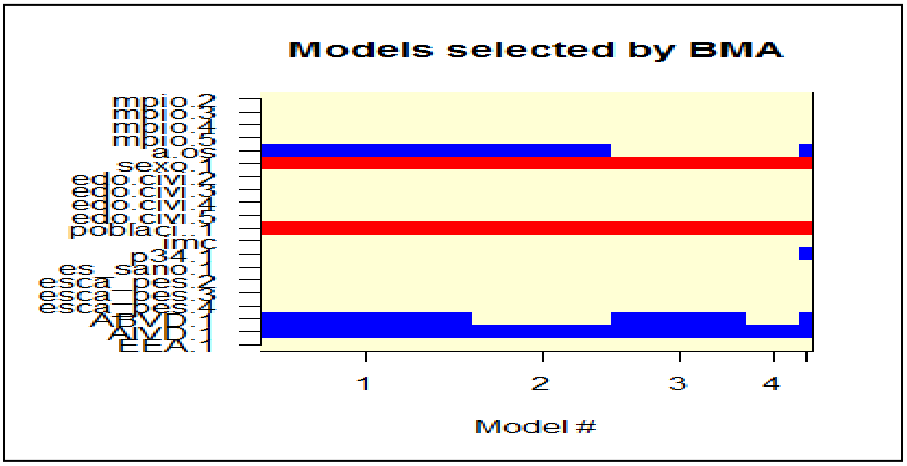

Imagen 3. Gráfico de inclusión de variables en el modelo BMA

5. CONCLUSIONES

El modelo seleccionado por BMA, expresa la variable Y (afectación del Estado cognitivo del adulto mayor) como función de la población a la que pertenece el adulto mayor, su edad, sexo, actividades básicas de la vida diaria y actividades instrumentales de la vida diaria, es importante resaltar que hay tres variables que parecen ser determinantes en relación con el Estado Cognitivo, estas son: AIVD, Sexo y Tipo de población. Cabe señalar que este es el primer trabajo que se realiza bajo este enfoque, sin embargo, existen tres trabajos en los cuales se estudia el estado funcional de los adultos mayores, en (Dorantes et al., 2007) se hace un análisis de regresión logística multifactorial, en (Díaz et al., 2012) se aplica un análisis de regresión logística y el otro utiliza un análisis de regresión multinomial (Díaz et al. 2011). Entre las variables que resultaron significativas en estos estudios, está la edad, género y el estado cognitivo que aparece como una covariable.
Es importante decir que el procedimiento presentado en este trabajo y el cual estamos proponiendo como una estrategia de análisis para este tipo de problemas (determinación de factores de riesgo) presenta notables ventajas sobre los análisis que tradicionalmente se realizan ya que además de su fácil implementación el paquete estadístico R proporciona resultados que se pueden observar gráficamente, pero sobre todo que en ellos se observan los mejores modelos y a su vez podemos ver cuales variables son las más importantes en cada uno de estos modelos y resulta ser una herramienta muy útil en investigaciones multidisciplinarias.
La utilización del paquete BMA en el entorno R constituye un aporte muy importante que favorecer la selección de modelos bajo el enfoque bayesiano y que en la actualidad está teniendo un desarrollo notable en las distintas áreas del conocimiento y más recientemente en aplicaciones en áreas de la Bioestadística, Epidemiología, y Salud en la cual está enmarcado el trabajo presentado.
En resumen, el procedimiento BMA (Bayesian Model Avering) constituye un herramienta que permite aplicar en enfoque Bayesiano al problema de la selección de modelos, tomando en cuenta la incertidumbre inherente a cualquier modelo estadístico y establece un establece un mecanismo coherente que, junto con la opinión de los expertos, permite elegir el modelo más apropiado.

RECEIVED NOVEMBER, 2013.
REVISED FEBRUARY, 2014

REFERENCIAS

[1] ANDO, T. (2010): **Bayesian Model Selection and Statistical Modeling**. Chapman & Hall/ CRC Press.
[2] AMINI S.M. and PARMETER C. F. (2011): Bayesian Model Averaging in R. **http://www.bus.miami.edu/_assets/files/faculty-and.../eco.../WP2011-9.pdf.**
[3] BANCO MUNDIAL (1994): **Informe sobre el desarrollo mundial 1993**. Invertir en salud. Oxford University Press, Washington:.

[4] CLAESKENS, G. AND HJORT, N. L. (2008): **Model Selection and Model Averaging**. Cambridge University Press, Cambridge.
[5] DÍAZ G. L., SISTACHS V. V., COVARRUBIAS M. D., ALARCÓN M. L., Y HERNÁNDEZ N. I. (2011): Capacidad funcional del adultos mayores de 65 años del programa pensión Guerrero: una aplicación del modelo de regresión multinomial, trabajo presentado el **4to Taller Latino Iberoamericano de Investigación de Operaciones,** Acapulco, México. sin publicar.
[6] DORANTES MENDOZA G., AVILA FUENTES J.A., Y GUTIERREZ ROBLEDO L.M. (2007): Factores asociados con la dependencia funcional en los adultos mayores: un análisis secundario del estudio nacional sobre salud y envejecimiento en México, 2001**. Rev. Panam. Salud Pública**. 22, 1-11.
[7] GELFAN, A.E. and DEY, D. K. (1994): Bayesian Model: asymptotics and exact calculations. **Journal of the Royal Statistical Society** B56, 510-514.
[8] LIPKOVICH, I.(2002): **Bayesian Model Averaging and variable selection in Multivariate Ecological Models**,. Dissertation, Blacksburg Virginia.
[9] KADANE, J. B. and LAZAR, N. A (2004): Methods and Criteria for Model Selection. **Journal of the American Statistical Association**, 99, 279-290.
[10] SÁNCHEZ, S, ET AL,(2001): Factores de riesgo preeclampsia en mujeres **HTTP://SISBIB.UNMSM.EDU.PE/BVREVISTAS/GINECOLOGIA/VOL_47N2/FACTORES_RIESG_PREECLAM.HTM** consultado en diciembre 2011.
[11] RAFTERY, A., C., PAINTER and VOLINSKY, I. (2005): **BMA: An R package for Bayesian Model Averaging,** R news, 5, 2-8.
[12] RAFTERY, A., HOETING, J., VOLINSKY, C., PAINTER, I. and YEUNG, K. Y. (2010): BMA: Bayesian Model Averaging. R package version 3.13.**URL: http://CRAN.Rproject.org/package=BMA**
[13] R DEVELOPMENT CORE TEAM (2010): R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria. **http://www.R-project.org,** consultado en enero 2012
[14] SALINAS N. S., DÍAZ G. L., COVARRUBIAS M. D., SISTACHS V. y V., HERNÁNDEZ N. I. (2012) Factores asociados a la Funcional de los adultos mayores en el Estado de Guerrero, **Trabajo Presentado En Segundo Encuentro Internacional de Medio Ambiente**. Sin publicar.