
Modelo logístico multinomial ordinal para la caracterización del dengue en el departamento del Caquetá entre los años 2012 hasta el 2019

Ordinal multinomial logistic model for the characterization of dengue in the department of Caquetá between the years 2012 to 2019

María Angélica Polanía Macías.^a
maria.polania@usantotomas.edu.co

Dagoberto Bermúdez Rubio.^b
dagobertobermudez@usantotomas.edu.co

Resumen

El dengue es una enfermedad vírica que afecta a la gran mayoría de países en zonas trópico y subtropicales, siendo transmitida por el mosquito *Aedes aegypti*. En el departamento del Caquetá, siendo una zona húmeda y con gran cantidad de zonas verdes, hace que sea su lugar de habitación. Debido a los altos números de personas afectadas por la picadura del mosquito que transmite el dengue, se quiere caracterizar por medio de un modelo logístico multinomial ordinal y exploración de datos con análisis multivariados, con la información brindada por la Secretaría Departamental de Salud del Caquetá, para así detectar cuáles son las variables que influyen más en la clasificación final del dengue (sin signos de alarma, con signos de alarma, dengue grave y mortalidad de dengue).

—
Palabras clave: (Dengue sin signos de alarma, Dengue con signos de alarma, Dengue grave, Mortalidad por dengue, Modelo logístico multinomial ordinal, V cramer).

Abstract

Dengue is a viral disease that affects the vast majority of countries in the tropics and subtropics, being transmitted by the *Aedes aegypti* mosquito. In the department of Caquetá, being a humid area with a large number of green areas, it makes it their place of residence. Due to the high numbers of people affected by the bite of the mosquito that transmits dengue, we want to characterize it by means of an ordinal multinomial logistic model and data exploration with multivariate analysis, with the information provided by the Caquetá Departmental Health Secretariat, in order to detect which are the variables that most influence the final classification of dengue (without warning signs, with warning signs, severe dengue and dengue mortality).

—
Keywords: (Dengue without warning signs, Dengue with warning signs, Severe Dengue, Mortality due to dengue, Ordinal multinomial logistic model, V cramer).

1. Introducción

Colombia por ser un país ubicado en el trópico y con regiones con altos porcentajes de humedad, es propicia para la picadura del zancudo hembra infectada por el mosquito *Aedes aegypti*, quien lleva en

^aEstudiante

^bDirector

ella la transmisión prevalente del dengue. La permanencia de estos mosquitos se debe a la ubicación geográfica, con mayor frecuencia en zonas urbanas con climas tropicales y subtropicales. La propagación del dengue en Colombia es generalmente por las variaciones climáticas, tal como el fenómeno del niño, en la gran mayoría de departamentos, donde actualmente convive el virus.

El incremento de los casos de dengue en el departamento de Caquetá, que aproximadamente en lo transcurrido en el año 2020 se registraron 1.857 casos y dos muertes, donde la humedad, los cambios climáticos, almacenamiento de aguas estancadas y residuales, mal uso de los humedales y de los residuos sólidos y una conciencia ecológica muy pobre en la gente, han contribuido a su crecimiento.

El departamento del Caquetá, está localizado en el sur del país y está conformado por 16 municipios: la capital, Florencia, Morelia, Bélen de los Andaquíes, San José del Fragua, Albania, Puerto Rico, San Vicente del Caguán, Cartagena del Chairá, El Doncello, Curillo, Solano, El Paujil, Valparaíso, Milán, Solita y La Montañita, donde se han reportado casos de dengue.

Debido a la problemática que se está viviendo en el departamento del Caquetá, se propone primero implementar un análisis multivariado, para la exploración de los datos. Después, a través de un modelo logístico multinomial de variable de escala ordinal, se caracteriza la clasificación del dengue. Es decir, se quiere observar cual es la probabilidad de que un paciente diagnosticado con dengue sea sin signos de alarma, con signos de alarma, dengue grave o mortalidad de dengue. Para ello, se dispone de datos suministrado por la Secretaría Departamental de Salud del Caquetá, teniendo información entre los años 2012 hasta el 2019.

2. Antecedentes

En Colombia, se realizó un estudio del comportamiento de la mortalidad implementando estadísticas descriptivas, tomando tasas de mortalidad del DANE. Chaparro-Nárvaez, León-Quevedo Catañeda-Orjuela en el 2016 [4], realizaron una investigación titulada “Comportamiento de la mortalidad por dengue en Colombia entre 1985 y 2012”, dando resultados generales por grupos de edad en casos de mortalidad en todo Colombia y dando los departamentos con alta mortalidad por dengue, quienes son: “río Magdalena, sectores de la Orinoquía, la Amazonía y la Costa Atlántica”.

También, se implementaron modelos bayesianos para el estudio del dengue en Colombia, como lo es el artículo de Montoya-González, Ortiz-Beltrán Santa-Guzmán en el año del 2017 [6], donde lo titulan “Modelo bayesiano para el estudio de la enfermedad del dengue en el departamento de Atlántico, Colombia, años 2010 a 2013”. El propósito del trabajo fue la implementación de datos de casos de dengue observando la relación que tiene con lo social, geográfico y económico en todo el departamento del Atlántico. Tomando el criterio de información de desviación (DIC), observan por cada año las variables que tienen relaciones afines con las características anteriormente mencionadas. Por último, los resultados obtenidos para esta investigación son que las variables afines con vivienda y el crecimiento de desorden en el casco urbano hacen que tengan mayores casos de dengue en los municipios del Atlántico.

A nivel internacional, utilizando métodos bayesianos, donde desean pronosticar el riesgo de la fiebre del dengue en Vietnam, Bett B, Grace D, Lee HS en el año 2019 [3], desarrollaron un modelo espacio-temporal, titulado “Análisis espacio-temporal de registros históricos (2001-2012) sobre la fiebre del dengue en Vietnam y desarrollo de un modelo estadístico para pronosticar el riesgo”, implementando un modelo bayesiano espacial jerárquico, agrupando por provincia y mes y dividiéndolo por conjuntos de capacitación y validación, agrupándolo por años. A la hora de la aplicación de los modelos espacio temporales, se basan en “los modelos Besag-York-Mollie (BYM) y autorregresivos (1), utilizando el criterio de información de desviación (DIC)”. Los resultados obtenidos fueron que las variables “temperatura mínima media mensual, las precipitaciones, el área bajo asentamiento urbano fueron significativas”

También, en Brasil realizaron un estudio basado en encuestas a hogares en la ciudad de Goiania entre los años 2001 y 2002. Siqueira-Junior, JB, Maciel, IJ Barcellos en el 2008 [12], estimaron un modelo aditivo generalizado (GAM), titulado “Análisis de puntos espaciales basado en encuestas de dengue a

nivel de hogares en el centro de Brasil”, donde tenían como objetivo observar el riesgo espacial en la región estudio. Los resultados dicen que las variaciones espaciales en el riesgo de picadura del zancudo *Aedes* tiene significancia al aumento de la edad y los bajos niveles de educación.

3. Objetivos

3.1. Objetivo general

Identificar los factores del grupo de clasificación del dengue (dengue sin signos de alarma, dengue con signos de alarma, dengue grave y mortalidad de dengue), por medio de la implementación de modelo de regresión logística multinomial ordinal para el departamento del Caquetá.

3.2. Objetivos específicos

3.2.1.

Organizar la información de la base de datos obtenida por medio de la Secretaría departamental de Salud del Caquetá

3.2.2.

Indicar la dinámica estadística que se va a utilizar para observar el comportamiento de las variables respecto a los grupos de clasificación del dengue.

3.2.3.

Analizar los resultados obtenidos, en busca de alternativas de solución o prevención de la población que se encuentra en algún grupo con signo de dengue.

3.2.4.

Mostrar los resultados estadísticos, con las autoridades de salud departamental, en la búsqueda de alternativas de mejoramiento que conlleve a la erradicación del dengue.

4. Marco teórico

El dengue es una epidemia que toma sus primeros registros de pacientes infectados en las regiones de Asia y Pacífico durante la segunda guerra mundial, dando inicio a la expansión del mosquito *Aedes aegypti*, esto es dado a la alta movilización de soldados y residentes para la protección y el bienestar de las personas por la situación que estaban viviendo en ese momento. Así mismo el mosquito iba dejando registros de pacientes infectados por el nuevo virus del dengue (DENV). La llegada de la enfermedad del dengue a América generó alarma a la Organización Panamericana de la Salud (OPS), dado que desde sus inicios y al transcurrir los años, han aumentado las personas infectadas y la morbilidad para esta enfermedad.

La mayor morbilidad presentada entre todos los arbovirus es del dengue por la picadura de mosquito *Aedes*, si bien, el arbovirus también se destaca con otros tipos de picaduras de artrópodos, tales como, insectos y arácnidos o también llamadas garrapatas. Según Messina, JP, Brady, OJ, Golding, N en

el 2019 [7], titulado “The current and future global distribution and population at risk of dengue” explican que el estimado de muertes son aproximadamente de 10.000 y 100 millones de personas infectadas presentando síntomas anualmente en áreas adecuadas para la transmisión del dengue.

El aumento de la temperatura al transcurrir los años, ha hecho que se intensifique las picaduras de mosquitos Aedes en las zonas donde actualmente conviven esta especie de artrópodos, dado que tendrían una mayor supervivencia y reproducción de esta especie. Algo que también preocupa es la colonización a países con bajo riesgo de transmisión de este mosquito o también libres de dengue.

Por otro lado en Colombia y en América Latina, Según la OPS publicado en el 2019 [9], titulado como “La OPS advierte sobre la compleja situación del dengue en América Latina y el Caribe” comentan que en el año 2019 se presentó la tasa más alta de infectados con más de tres millones de casos de dengue positivos y más de 1.300 personas con morbilidad, después de haber tenido una baja de casos confirmados por dengue entre los años 2017-2018. Pero el año 2019 no superó los casos entre los años 2015-2016, que hasta ahora son los años con mayores registros.

El dengue en sus comienzos no se consideraba una enfermedad con morbilidad en las personas, pero al transcurrir los años sus inicios de morbilidad empezaron con las complicaciones del dengue hemorrágico en niños de diferentes países con alta presencia del mosquito Aedes. Desde el año 1975, la Organización Mundial de la Salud (OMS), conformaron la primera forma de clasificar clínicamente el dengue. Se clasificó como:

Dengue clásico (DF): Es una enfermedad que se considera que uno de sus primeros hallazgos es fiebre alta y entre sus síntomas secundarios se encuentra: náuseas, vómitos, cefalea, artralgia, dolor retroorbitario, exantema, mialgia, manifestaciones hemorrágicas y leucopenia. Debido a la falta de especificidad de estos signos y síntomas clínicos, se requiere evidencia de laboratorio o epidemiológica de infección por el virus del dengue para un diagnóstico confirmado. (Srikiatkachorn A, Rothman AL, Gibbons RV, 2011) [13]

Dengue hemorrágico (DH): Consta de cuatro criterios clínicos: fiebre, tendencia hemorrágica (sangrado espontáneo o resultado positivo de la prueba del torniquete), trombocitopenia (recuento de plaquetas, 100000 células / mm³) y pérdida de plasma como se muestra por derrame pleural, ascitis, o 20% de hemoconcentración. Es importante señalar que la especificidad de este síndrome significa que un diagnóstico de dengue hemorrágico no requiere pruebas de laboratorio de infección por el virus del dengue. (Srikiatkachorn A, Rothman AL, Gibbons RV, 2011) [13]

En el año 2009, la OMS presenta un cambio en la forma de clasificar el dengue, debido a que los síntomas que se destinaron para ser dengue hemorrágico no se podía aplicar en la atención primaria en un centro médico. También, llegaron críticas en donde los síntomas que se especificaban no se podía confirmar que era dengue hemorrágico. Para dar solución a ello, la OMS decide investigar más a fondo las verdaderas sintomatologías que representen el dengue grave como lo es llamado hoy en día, donde según el Sistema de Vigilancia en Salud Pública (SIVIGILA) se clasifica de la siguiente manera:

- Dengue sin signos de alarma:
 - Fiebre persistente
 - Cefalea o también nombrado dolor de cabeza intenso.
 - Dolor retroocular o dolor intenso alrededor del ojo.
 - Mialgias o dolor persistente en los músculos del cuerpo.
 - Artralgias o dolor persistente en las articulaciones del cuerpo.
 - Erupción o rash o lesiones persistentes en la piel.
- Dengue con signos de alarma:

- Dolor abdominal intenso
- Vómito
- Diarrea
- Somnolencia o dificultad para conciliar el sueño.
- Hipotensión o presión arterial alta.
- Hepatomegalia o agrandamiento del hígado.
- Hemorragias importantes en mucosas o también sangrado de encías y de nariz.
- Hipotermia o disminución de la temperatura en el cuerpo.
- Aumento de hematocrito o también la disminución de plaquetas .
- Caída de plaquetas (< 100000)
- Acumulación de líquidos
- Dengue grave
 - Extravasación severo de plasma o derrame excesivo.
 - Hemorragia con compromiso hemodinámico o también dicho, sangrado vaginal, riñon y aparato digestivo alto y bajo.
 - Shock por dengue
 - Daño grave de órganos

4.1. Imputación por KNN

En la mayoría de casos el conjunto de datos que se va a estudiar, los valores faltantes están presentes; usualmente para que los resultados de cualquier estudio de investigación sin sesgos, se realiza o se utiliza un método de imputación de datos, donde dependiendo el método de imputar, va a completar la información y así mismo presentar resultados confiables.

K-Nearest-Neighbor es un algoritmo de Machine Learning, donde clasifica los valores buscando el valor más similar por su cercanía, siendo un algoritmo supervisado. Internamente lo que hace, es clasificar los datos que no tienen “etiqueta”, convirtiéndola a una clase de etiquetas debido a su cercanía o similitud de los datos. Para ello, utiliza la métrica de la distancia euclidiana a la hora de hacer la imputación:

$$d(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_i - q_i)^2} \quad (1)$$

Cabe resaltar que para el conjunto de datos que se está trabajando se utiliza el parámetro `n_neighbors` con tres vecinos más cercanos, según la distancia euclidiana.

4.2. Análisis de correspondencias múltiples

El análisis de correspondencias múltiples es una técnica que usualmente se utiliza para hacer descripción y exploración de los datos representados en tablas de contingencias, donde se recoge las frecuencias de dos o más variables categóricas. “La información de partida ahora es una matriz de dimensiones IJ , que representa las frecuencias absolutas observadas de dos variables cualitativas en n elementos. La primera

variable se representa por filas, y suponemos que toma I valores posibles, y la segunda se representa por columnas, y toma J valores posibles” (Peña, 2002) [10]

A diferencia del análisis de componentes principales (PCA) que su objetivo principal corresponde a obtener la medición de las variables numéricas y por el contrario, el análisis de correspondencias es representar en un espacio una matriz la medición de varias variables categóricas. (Montanero Fernández, 2015). [8]

El propósito del ACP y el ACM es de resumir la dimensionalidad de un conjunto de información generando variables denominadas factores o ejes, donde evidencian algunas diferencias de la información a estudio de acuerdo a las combinaciones que se presentan entre ellas. Dando como resultado la información de forma visual, tales como, tablas, gráficos o digramas, que nos va poder mostrar las distancias entre los individuos en el espacio. (Algañaraz, 2016). [2]

Primero, se busca la mejor proyección de la matriz F , donde se llamará frecuencias relativas a f_{ij}

$$\sum_{i=1}^I \sum_{j=1}^J f_{ij} = 1 \quad (2)$$

Para encontrar la medida de distancia que se debe utilizar , se toma R como la matriz de frecuencias relativas del total de filas.

$$R = D_f^{-1}F \quad (3)$$

Para encontrar las diferencias en la frecuencias relativas, se suman :

$$\frac{(r_{aj} - r_{bj})^2}{f_{.j}}, \text{ donde, } f_{.j} = \sum_{i=1}^I f_{ij} \quad (4)$$

Y así, se encuentra la distancia entre dos puntos, donde matricialmente se expresa:

$$D^2(r_a, r_b) = (r_a - r_b)' D_c^{-1} (r_a - r_b) \quad (5)$$

4.3. Correlación de Cramer V

La correlación es una medida que observa que tan correlacionadas estan entre variables o características de una información a estudio. Matemáticamente se expresa con la correlación de Pearson, donde tiene como finalidad sacar la correlación entre el intervalo de [-1,1]:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (6)$$

Ahora bien, en la fórmula anterior se representan para un conjunto de datos cuantitativos, pero, el conjunto de datos que se va a trabajar son categóricos y para ello se utiliza la correlación de Cramer V, que funciona como una medida basada en Ji cuadrado, y se expresa de forma matemática de la siguiente manera:

$$V = \sqrt{\frac{X^2}{n(\min[r, c] - 1)}} \quad (7)$$

4.4. Modelo lineal generalizado

Los modelos lineales generalizados representada con las siglas GLM, son una expansión de los modelos lineales, pero con una diferencia que se pueden representar con distribuciones pertenecientes a la familia exponencial, tales como, Poisson, Binomial, Multinomial, entre otros. La propuesta de estos modelos fueron hechas por Wedderburn y Nelder en el año 1972, buscando una forma de unificar otros tipos de modelos estadísticos. [5]

Para entender un poco más los modelos lineales generalizados, se representa el modelo de regresión múltiple, donde:

$$E[Y] = \beta_0 + \beta_1 X_1 + \dots + \beta_k X_k \quad (8)$$

Esto quiere decir que hay una relación lineal entre X y $E[Y]$, por lo tanto se relaciona una probabilidad de que ocurra un hecho con las demás variables explicativas. Un modelo lineal generalizado tiene tres componentes:

- Componente aleatorio: Identifica a la variable Y , nuestra variable respuesta; para ello se utiliza la distribución perteneciente a la familia exponencial uniparamétrica y se representa como su función de densidad:

$$f(y; \theta, \phi) = \exp\left[\frac{y\theta - b(\theta)}{a(\phi)}\right] + c(y, \phi) \quad (9)$$

- Componente sistemático: Representa a las variables explicativas o predictoras.
- Función enlace: Esta función relaciona la media $E[Y]$ y las variables explicativas. La función enlace transforma la media a un parámetro θ

$$g(E[Y]) = \theta \quad (10)$$

4.4.1. Modelo logístico multinomial ordinal

Se aplica un modelo logístico que se utiliza para variables respuestas ordinales llamados modelos logísticos acumulados. Según (Agresti, 2002) [1] Y la probabilidad acumulada de la variable respuesta, teniendo en cuenta que Y sea menor o igual que un valor j , definiéndose de la siguiente manera:

$$P(Y \leq j|x) = \pi_1(x) + \dots + \pi_j(x), j = 1, 2, \dots, J \quad (11)$$

[1]

Los logit acumulados se definen de la siguiente manera:

$$\text{logit}[P(Y \leq j|x)] = \log \frac{P(Y \leq j|x)}{1 - P(Y \leq j|x)} \quad (12)$$

[1]

$$= \log \frac{\pi_1(x) + \dots + \pi_j(x)}{\pi_j(x) + \dots + \pi_J(x)}, j = 1, \dots, J - 1 \quad (13)$$

[1]

Según (Agresti, 2002) [1], el modelo que se utiliza los logits acumulativos es el siguiente:

$$\text{logit}[P(Y \leq j|X)] = \alpha_j + \beta_1 X_1 + \dots + \beta_p X_p \quad (14)$$

[1]

Cada modelo logístico tiene su propio intercepto α_j siendo crecientes en j teniendo los mismo efectos en β . [1]

5. Resultados

5.1. Variables

En el presente estudio se toma información brindada por la Secretaría departamental de Salud del Caquetá, entregando tres bases de datos de dengue, dengue grave y mortalidad de dengue para todos los municipios del Caquetá, entre los años 2012 hasta el 2019. Seguidamente se profundiza la información que aborda todo lo relacionado con el dengue y sus altos casos de pacientes infetados en el departamento, para así tener más claridad a la hora de iniciar el estudio estadístico.

En las siguientes tablas se muestran las variables que componen la base de datos y la codificación de cada una de ellas tanto en R studio como en Python. Dado que la imputación de datos se hace en python.

Variable	Etiqueta	Codificación
Año	Año en el que ingreso el paciente diagnosticado con dengue.	2012 2013 2014 2015 2016 2017 2018 2019
Edad	Edad en la que ingreso el paciente diagnosticado con dengue.	Desde 1 hasta los 85 años.
sexo_	Sexo del paciente diagnosticado con dengue.	1. Femenino. 2. Masculino.
area_	Área de residencia del paciente diagnóstico con dengue.	1. Cabecera municipal 2. Centro poblado. 3. Rural disperso.
pac_hos_	Si el paciente diagnosticado con dengue fue hospitalizado o no.	1. Si_hosp 2. No_hosp
desplazami	Si el paciente diagnosticado con dengue se desplazo los últimos 15 días.	1. Si_desplaz 2. No_desplaz

Tabla 1: Variables

tip_cas	Clasificación inicial del paciente diagnosticado con dengue.	<ol style="list-style-type: none"> 1. Sospechoso. 2. Probable. 3. Confirmado por laboratorio. 4. Confirmado por clínica. 5. Confirmado por nexo epidemiológico.
Clasificación_final	Clasificación del dengue según el SIVIGILA y la OMS.	<ol style="list-style-type: none"> 1. Sin signos de alarma 2. Con signos de alarma 3. Dengue grave 4. Mortalidad por dengue
conducta	A donde fue remitido el paciente diagnosticado con dengue.	<ol style="list-style-type: none"> 1. No hospitalizado 2. Ambulatoria 3. Hospta. piso 4. Uni. cuidados intensivos 5. Observación 6. Remis. para hospitalización

Tabla 2: Variables

5.2. Imputación de los datos

Primero se observa el porcentaje de información de datos faltantes para cada una de las variables seleccionadas para este estudio, donde determinar que la variables sexo, área de residencia, desplazamiento del paciente en los últimos días y el tipo de conducta presentan información pérdida. Para ello es necesario utilizar un método de imputación, que en este caso se utilizó K- vecinos más cercanos para datos cualitativos.

```

año                0.000000
edad_              0.000000
sexo_              0.059844
area_              0.014961
tip_cas_          0.000000
pac_hos_          0.000000
desplazami        14.392579
    
```

Figura 1: Porcentajes de datos faltantes

```

Clasificación Final  0.044883
conducta             14.392579
    
```

Figura 2: Porcentajes de datos faltantes

5.3. Coeficiente Cramer V

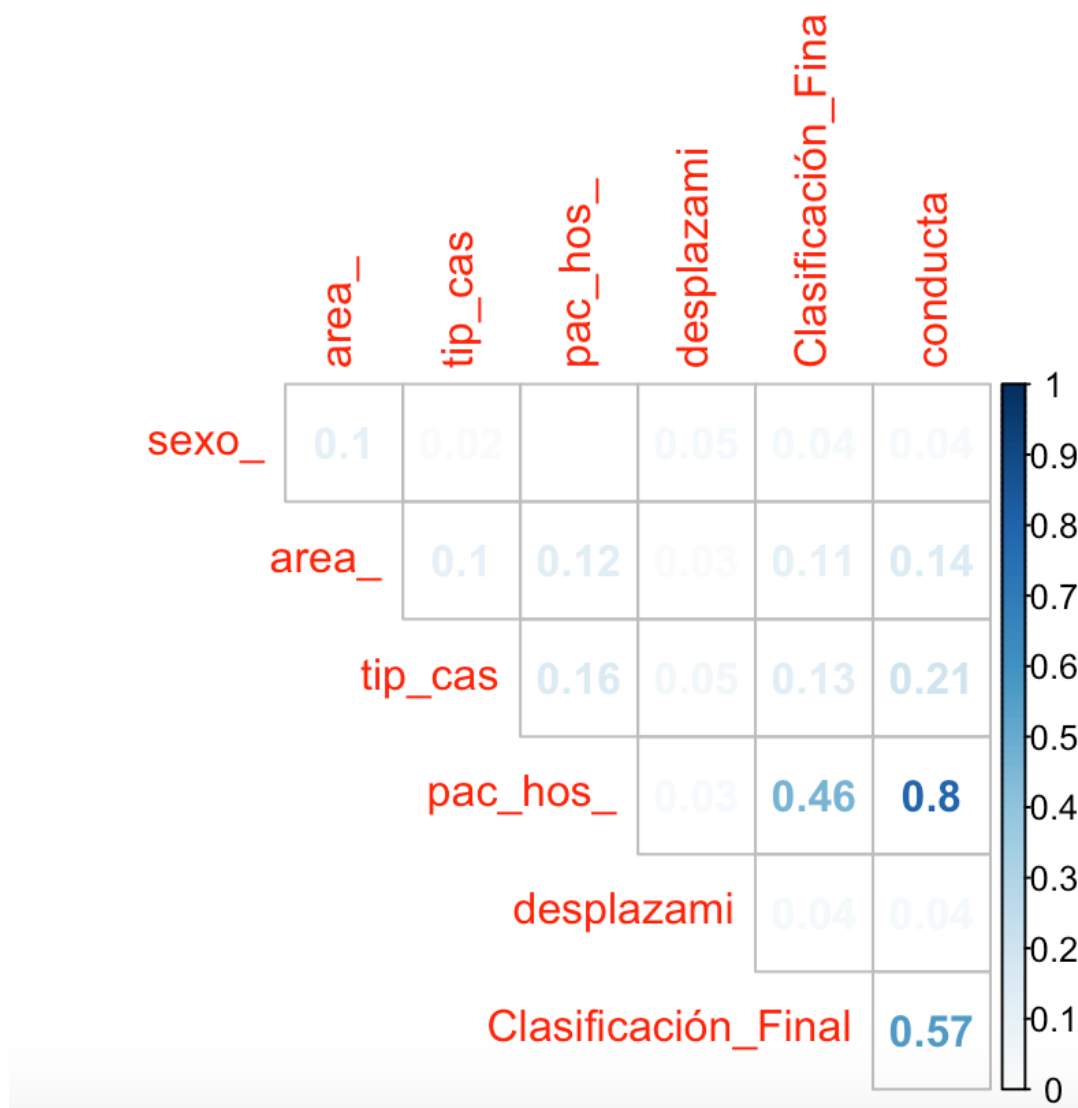


Figura 3: Mapa de calor de cramer

Se puede identificar en el mapa de calor a través de las variaciones de color las correlaciones entre variables. Una diferencia del análisis de correlaciones de Pearson, es que la correlación de cramer va entre un interval $[0,1]$. El coeficiente de cramer se utiliza para el análisis y visualización de datos categóricos. Habitualmente, los valores que están por encima de 0.30, se halla una relación entre variables; cuando dan valores de 0 o menores a 0.30 se dice que hay independecia entre variables y 1 es una relación perfecta entre variables.

En este caso, para este estudio se puede observar en el gráfico de calor, que en la variable clasificación final, que componen las categorías de la variable Y, donde reúne todas las sintomatolgías mencionadas anteriormente, se dice que tiene una correlación con si el paciente diagnosticado con dengue fue hospitalizado o no. Por lo general, es evidente que si un paciente es diagnosticado con dengue con síntomas de alarma o dengue grave se remite a hospitalización, como también si es dengue sin signos de alarma, en la mayoría de casos no se remite a hospitalización del paciente, debido a que los síntomas se pueden

controlar desde casa.

También, para la variable conducta, que en si, esta variable hace que el médico clasifique dependiendo el comportamiento del paciente, tiene una correlción casi perfecta con la variable si el paciente fue hospitalizado o no, debido a que hubo pacientes que fueron hospitalizados y dependiendo de los síntomas presentados en ese momento, el médico remitió al paciente a hospitalización en urgencias o en observación. De igual manera sucede con los pacientes que tuvieron conducta ambulatoria, ya que estas personas no fueron hospitalizados, pero si debieron acudir regularmente al centro médico.

Por último, se observa que clasificación final también tiene una correlación con la variable conducta, debido a que si los síntomas no generan alarmas a los médicos especialistas, se remite a una conducta ambulatoria, pero si son síntomas con signos de alarma , se es necesario que el paciente sea hospitalizado. También se debe tener en cuenta si los síntomas son clasificados dengue grave en la mayoría de casos el paciente es llevado la unidad de cuidados intensivos por la gravedad de los síntomas.

Clasificación final	Conducta				
	No hospit.	Ambulatorio	Hospit.	Uni. cuidados intensivos	Observa.
Sin signos de alar	0.000000	0.3500	0.1274	0.01256	0.0463
Con signos de alar	0.000000	0.04129	0.2748	0.0418	0.0537
Dengue grave	0.000000	0.00044	0.0014	0.0073	0.0004
Mortalidad	0.00074	0.000000	0.0000	0.0004	0.0001

Tabla 3: Proporción conducta-clasificación final

Se complementa el resultado de la correlación de cramer para observar el porcentaje de pacientes para cada una de las categorías de la variable conducta dada las categorías de la clasificación final.

5.4. Análisis de correspondencias múltiples

Se puede observar que en la dimensión 1 las variables que estan más correlacionadas son área de residencia, si el paciente fue hospitalizado o no. Y las variables que están mas correlacionadas con la dimesión 2 es la clasificación inicial.

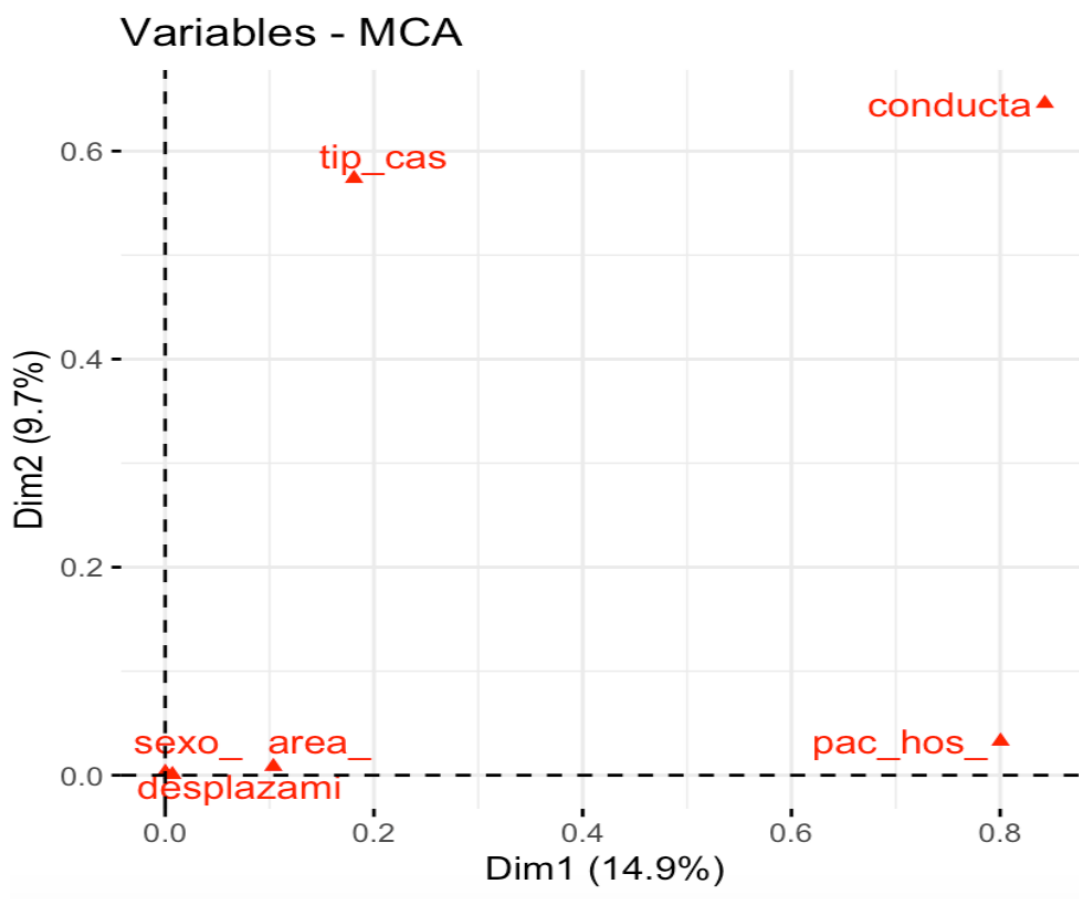


Figura 4: Correlaciones de las categorías

En la figura 5 se puede observar primero que toda la relación entre variables. Entre ellos se encuentra el primer grupo de categorías de si un paciente esta en unidad de cuidados intensivos y siendo un caso sospechoso de dengue inicialmente. En este caso, los pacientes que se clasifican inicialmente como un caso sospechoso de dengue han llegado a estar en cuidados intensivos, es decir, presentaron síntomas de dengue grave. En el segundo grupo, se dice que los pacientes residentes en el área centro poblado, es decir pacientes con viviendas con asentamientos rurales, tienden a ser diagnosticados con dengue sin signos de alarma debido a que son conducta ambulatoria, esto quiere decir, que no son hospitalizados. Y en el tercer grupo, se observa los pacientes residentes del área cabecera municipal, donde son los pacientes que mayor presentan dengue con signos de alarma y dengue grave, debido a las categorías que se relacionan, entre ellas está la confirmación del dengue tanto por laboratorio, como por clínica.

También, se observa la significancia en las dimensiones, entre ellas se encuentra la categoría sospechoso, donde es significativa para la dimensión 1 y 2, como también ambulatoria. Se encuentran las categorías Unidad de cuidados intensivos que solo afecta a la primera dimensión.

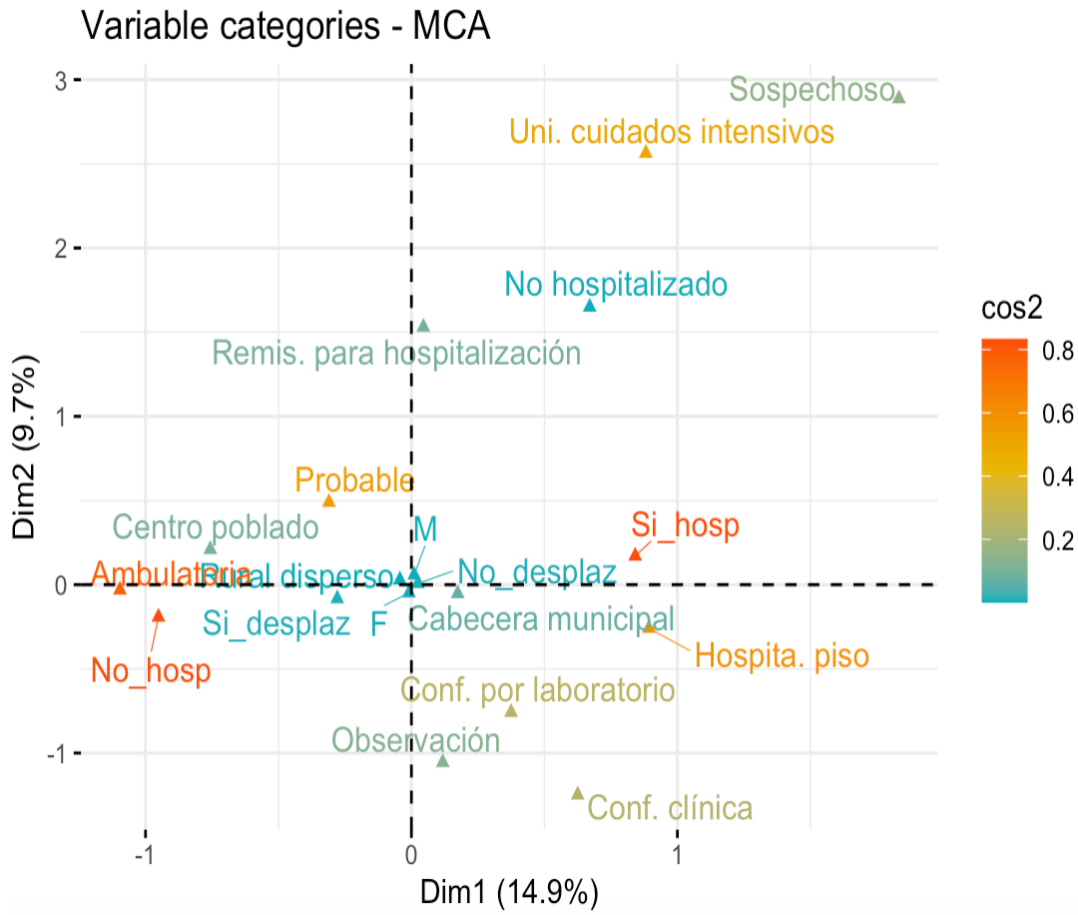


Figura 5: Importancia de las categorías

En este gráfico se observa la importancia o la contribución que hace las categorías, siendo la variable hospitalizado (si, no), conducta ambulatoria y hospitalizado por piso, las que más contibuyen.

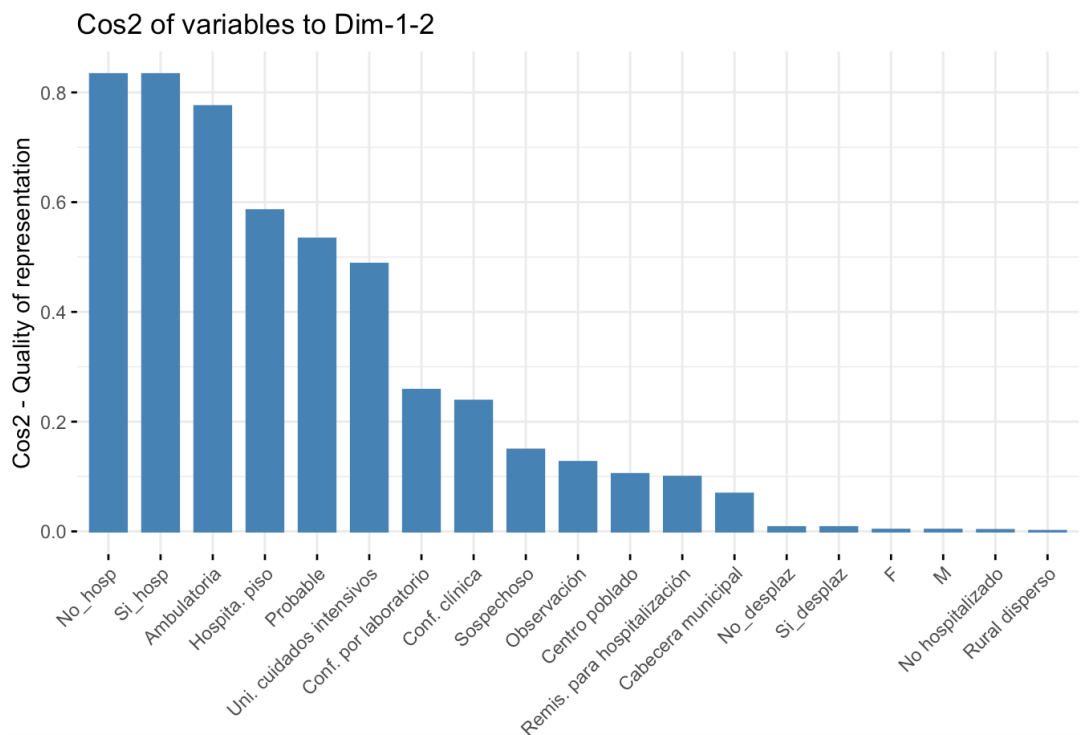


Figura 6: Contribución de las categorías

5.5. Modelo logístico multinomial ordinal

Se puede observar en el siguiente gráfico que se aplicó el modelo a través de los años y los resultados del modelo no cambian, esto quiere decir , que la información según los resultados siempre afecta a las mismas variables con respecto a la clasificación final.

Año	AIC	Residual Deviance
2018	7101.434	7069.434
2017	7101.434	7069.434
2016	7101.434	7069.434
2015	7101.434	7069.434
2014	7101.434	7069.434
2013	7101.434	7069.434
2012	7101.434	7069.434

Tabla 4: Resultados de a través de los años

Los resultados del modelo son los siguientes:

Coefficients:					
	Estimate	Std. Error	z value	Pr(> z)	
(Intercept):1	-4.522201	102.310117	-0.044	0.964744	
(Intercept):2	0.674087	102.310151	0.007	0.994743	
(Intercept):3	3.648364	102.310930	0.036	0.971554	
sexo_.L	-0.123642	0.043766	-2.825	0.004727	**
area_Centro poblado	0.175478	0.094832	1.850	0.064255	.
area_Rural disperso	0.507913	0.082017	6.193	5.91e-10	***
tip_cas.L	-1.373464	0.174890	-7.853	4.05e-15	***
tip_cas.Q	0.157438	0.133158	1.182	0.237071	
tip_cas.C	0.012833	0.074009	0.173	0.862335	
pac_hos_.L	0.225707	0.066398	3.399	0.000676	***
desplazami.L	-0.005749	0.079202	-0.073	0.942130	

Figura 7: Resultado modelo logístico multinomial ordinal año 2019

Como se puede observar, los resultados del modelo para el año 2019 se puede resaltar las variables con significancia con respecto a la variable interés que es clasificación final son, sexo, área rural disperso, la clasificación inicial con caso sospechoso y paciente hospitalizado. Esto quiere decir que, las mujeres del departamento del Caquetá han sido más propensas en tener dengue durante este año y el resto de años en estudio (2012-2018). En el departamento del Caquetá es muy común que los habitantes tengan casas hacia las afueras de los municipios, que normalmente son llamadas parcelas o fincas. A esto se le llama el área de rural disperso, donde en los resultados del modelo nos muestra que los pacientes más afectados por esta epidemia, son las personas que viven en este sector.

Por otra parte, se encuentra que la clasificación inicial del dengue, que es donde se observa los primeros síntomas del paciente y que probablemente el médico tome sus primeras hipótesis de lo que pueda tener el paciente. Entonces, según el modelo los pacientes en su primera clasificación, o sea antes de saber que es dengue, se clasifican como casos sospechosos o que tal vez por los síntomas iniciales se determine que es dengue. También, la mayoría de pacientes son hospitalizados con los primeros síntomas, esto se debe a que los pacientes llegan tanto con síntomas sin signos de alarma, como también con signos de alarma.

5.5.1. Gráficos de probabilidades por categoría con respecto a la variable edad.

A continuación se muestra unos gráficos donde muestran la probabilidad de una categoría por gráfico de la variable respuesta (Clasificación final) separa por sexo en primera instancia y tomando como variable x la única variable cuantitativa entre todas las variables que es edad. Y el segundo conjunto de gráficos se representa la misma probabilidad en Y pero esta vez separada por por área de residencia.

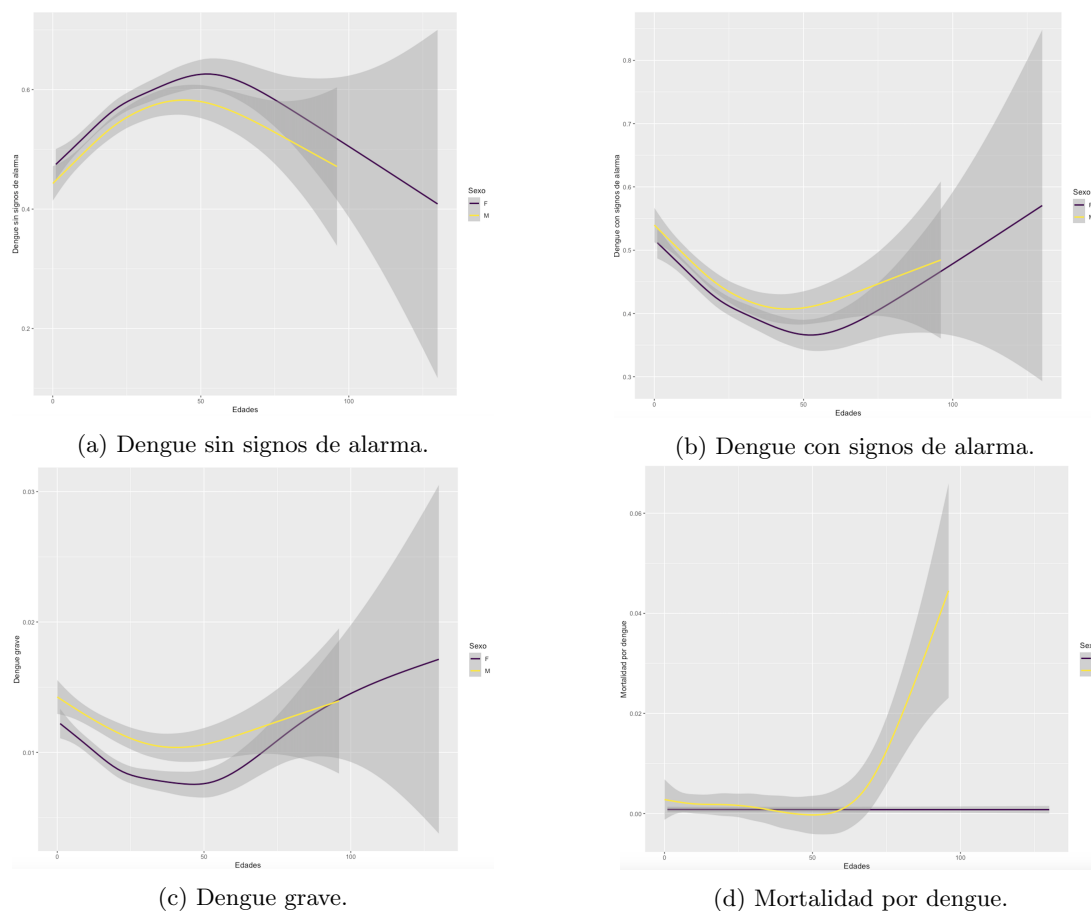


Figura 8: Gráfico de probabilidades separada por sexo

Se puede observar en este primer conjunto de gráficos que la mayor probabilidad está representada en la primera categoría que es dengue sin signos de alarma, donde hay más mujeres que hombres diagnosticadas con dengue sin signos de alarma, siendo las mujeres con 50 años las más afectadas por esta epidemia. También se muestra que hay mujeres mayores a 100 años afectadas por esta enfermedad. Para dengue con signos de alarma, se puede observar que la probabilidad es más baja, pero esta vez los hombres predominan esta categoría, donde se puede entender que hay más hombres con síntomas alarmantes y con una máxima de edad de alrededor de 95 años, pasando lo contrario con el sexo femenino que sigue habiendo mayores a 100 años. Es decir, la misma persona presentó otros síntomas en el mismo año.

También, en dengue grave se puede observar que la probabilidad es mucho más baja que las categorías anteriormente mencionadas. Se muestra también que hay más hombres con sintomatología grave que el sexo femenino y sigue la presencia de mujeres mayores a 100 años en esta categoría. Probablemente la misma persona se agrabó en el mismo año. Pero, en mortalidad por dengue, solamente murieron hombres con una probabilidad muy baja, esto quiere decir, que fueron muy pocos los hombres que fallecieron por dengue.

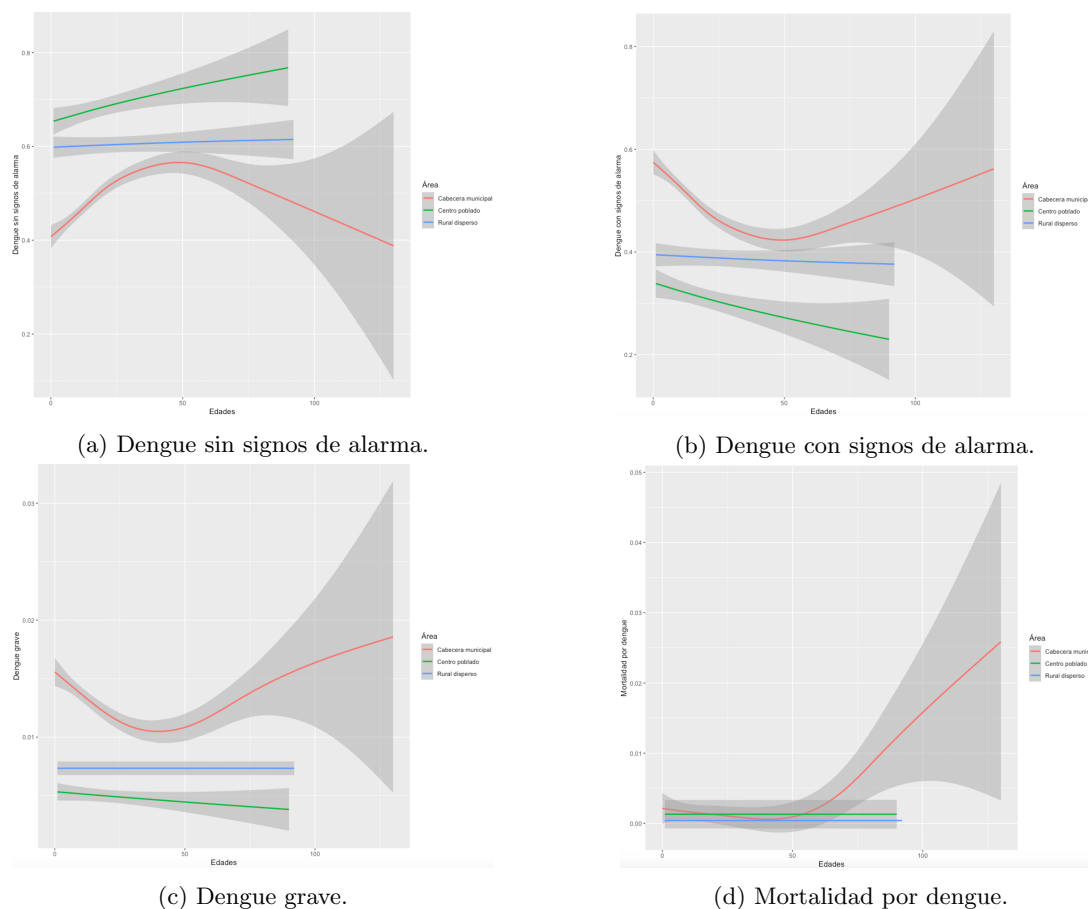


Figura 9: Gráfico de probabilidades separada por área

Para este caso que es por área de residencia, los pacientes diagnosticado con dengue sin signos de alarma, tiene mayor probabilidad las personas que son del área rural disperso y centro poblado de tener dengue con síntomas leves, con una edad menor a 100 años. También, se observa que las mujeres mayores a 100 años son del área cabecera municipal.

Por otra parte, el dengue con signos de alarma se encuentra con la mayor probabilidad los pacientes que viven en la cabecera municipal, al igual que la categoría de dengue grave. La mortalidad por dengue se encuentra localizado igualmente que las dos categorías mencionadas anteriormente, en la cabecera municipal. La prevalencia de esta epidemia dentro de la ciudad se debe a la acumulación de aguas, presencia de humedad en casas y edificios y humedales presentes dentro de la ciudad con aguas negras.

Como se puede observar, el gran porcentaje de pacientes se encuentran reunidas en los primeros síntomas del dengue, donde se categoriza como dengue sin signos de alarma y siendo remitidos a conducta ambulatoria, dado que el paciente no esta complicado de salud y solo es acudir regularmente a un centro médico.

Los hombres entre los intervalos de edad de mayores de 60 y menores a 100 años se concentran las muertes en el sexo masculino. Como se observa, la morbilidad en el departamento es bajo, pero aún así es preocupante, porque el dengue es una enfermedad que no debería tener morbilidad.

6. Conclusiones

En el análisis exploratorio multivariado de la base de datos, se concluye que la clasificación final del dengue (sin signos de alarma, con signos de alarma, dengue grave y mortalidad por dengue), está altamente correlacionado con la variable conducta.

Los pacientes que presentan dengue sin signos de alarma, el 35 % de la población, no presentan síntomas que compliquen su estado de salud, y se toma como una conducta ambulatoria; para pacientes con signos de alarma, el 27 % de la población fueron hospitalizados, sus síntomas exigen que deben ser tratados en el centro de salud, pacientes que presentaron síntomas con dengue grave, el 0.73 % de la población, por la gravedad son remitidos a la unidad de cuidados intensivos. La causa de mortalidad por dengue, en pacientes que llegaron a la UCI, es del 0.044 %, pacientes en observación el 0.014 % mientras que fallecieron personas que no fueron remitidas a hospitalización con un 0.074 %. Lo anterior confirma la complementación de las variables conducta y clasificación final en el momento de tomar decisiones para el seguimiento del paciente. La mortalidad por dengue se presenta en pacientes de sexo masculino y residentes en cabecera municipal. También se presentan pacientes de sexo femenino mayor a 100 años, quien fue diagnósticada con dengue sin signos de alarma y posteriormente pasa a ser dengue grave, la paciente mayor a 100 años se encontraba residiendo en el área de la cabecera municipal.

Se recomienda a la Secretaría de Salud, realizar campañas de prevención del dengue para reducir los índices de dengue sin y con signos de alarma.

Agradecerle a la Secretaría de Salud del Departamento del Caquetá por los datos brindados y realizar un segundo estudio sobre el dengue con la información que se tenga de los años 2020 y 2021, y se plantea la siguiente pregunta de investigación ¿Por qué los hombres tienen mayor probabilidad de ser infectado por el mosquito Aedes y morir?.

7. Bibliografía o Referencias

Referencias

- [1] Agresti, A. (2002) *Categorical Data Analysis. University of Florida. Gainesville, Florida. A John Wiley sons, INC., Publication*
- [2] Algañaraz, V. H. (2016) *El “Análisis De Correspondencias Múltiples” Como Herramienta Metodológica De Síntesis Teórica Y Empírica. Su Aporte Al Estudio Del Locus Universitario Privado Argentino (1955-1983). Relmecs, 6.*
- [3] Bett B, Grace D, Lee HS, et al. (2019) *Model Spation temporal analysis of historical records (2001-2012) on dengue fever in Vietnam and development of a statistical model for forecasting risk.*
- [4] Chaparro-Nárvaez, León-Quevedo Catañeda-Orjuela (2016) *Comportamiento de la mortalidad por dengue en Colombia entre 1985 y 2012.*, [urlsihttp://www.scielo.org.co/pdf/bio/v36s2/v36s2a14.pdf](http://www.scielo.org.co/pdf/bio/v36s2/v36s2a14.pdf)
- [5] Durbán María (2011). *Modelos Lineales Generalizados*, [urlhttp://halweb.uc3m.es/esp/Personal/personas/durban/esp/](http://halweb.uc3m.es/esp/Personal/personas/durban/esp/)
- [6] Montoya-González, Ortiz-Beltrán Santa-Guzmán (2017) *Modelo bayesiano para el estudio de la enfermedad del dengue en el departamento de Atlántico, Colombia, años 2010 a 2013*, [urlhttps:https://revistas.uptc.edu.co/index.php/perspectiva/article/view/7603/6677info](https://revistas.uptc.edu.co/index.php/perspectiva/article/view/7603/6677info)
- [7] Messina, J.P., Brady, O.J., Golding, N. (2019) *The current and future global distribution and population at risk of dengue. Nat Microbiol 4, 1508–1515*, [urlhttps://www.nature.com/articles/s41564-019-0476-8citeas](https://www.nature.com/articles/s41564-019-0476-8citeas)

- [8] Montanero, J. F. (2015) *Manual Abreviado De Estadística Multivariante*. Badajoz España: Universidad De Extremadura Departamento De Matematicas.
- [9] PAHO, Periodistas: Sebastián Oliel, Ashley Baldwin, Leticia Linn (2019) *PAHO warns of the complex situation of dengue in Latin America and the Caribbean*, url https://www.paho.org/hq/index.php?option=com_contentview=articleid=15365:paho-warns-of-the-complex-situation-of-dengue-in-latin-america-and-the-caribbeanItemid=135lang=en
- [10] Peña, D. (2002). *Manual Abrevi Análisis De Datos Multivariantes*. Madrid : University Carlos Iii De Madrid.
- [11] Stanforth, Moreno-Madriñan Ashby (2016) *Análisis exploratorio de las variables del nicho de la fiebre del dengue dentro de la cuenca del río Magdalena*, url <https://www.mdpi.com/2072-4292/8/9/770/htm>
- [12] Siqueira-Junior, JB, Maciel, IJ, Barcellos, C. et al.(2008) *Análisis de puntos espaciales basado en encuestas de dengue a nivel de hogares en el centro de Brasil*. *BMC Public Health* 8, 361, url <https://doi.org/10.1186/1471-2458-8-361>
- [13] Srikiatkhachorn, A., Rothman, AL, Gibbons, RV, Sittisombut, N., Malasit, P., Ennis, FA, Nimmanitya, S. y Kalayanarooj, S. (2011). *Dengue: la mejor forma de clasificarlo*. *Enfermedades infecciosas clínicas: una publicación oficial de la Sociedad de Enfermedades Infecciosas de América* , 53 (6), 563–567., url <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3202316/>