

Clasificación de pacientes con presencia de Dengue, desde un enfoque bayesiano

Juan Diego Espinosa Hernandez*, Cristian Alberto Cortes Zarate*,

Prof: Johnatan Cardona Jimenez

Universidad Nacional de Colombia, sede Medellín

Resumen

The following study proposes a model that allows predicting whether or not a patient can be hospitalized, given that they already have dengue. Additionally, it takes into account the type of dengue that can be predicted for the patient, considering certain pathologies associated with known cases of dengue.

In order to achieve this, a logistic model with a Bayesian approach was used. This choice was made because our response variable comprises a categorical variable that is dichotomous, describing two types of dengue. Furthermore, categorical variables describing whether or not the patient has the associated symptoms were used as predictors. It should be clarified that both qualitative and quantitative variables can be added to the logistic model. The STAN programming language was used to perform this procedure since the posterior distribution cannot be found using conventional analytical methods. Therefore, it must be approximated through numerical methods. Finally, to compare models, the convergence of the model and the significance of the predictor variables are verified in advance.

1. Introducción

El siguiente trabajo propone un modelo que permita predecir si un paciente puede ser o no hospitalizado dado que ya tiene un dengue, además de esto teniendo en cuenta que tipo de dengue que podamos predecir que posee el paciente, y todo esto teniendo en cuenta ciertas patologías asociadas a los casos de dengue conocidos.

Para hacer esto posible se utilizó un modelo logístico con un enfoque bayesiano, y se llevó a cabo esta elección ya que nuestra variable respuesta comprende una variable categórica la cual es de tipo dicótoma y describe 2 tipos de dengue, además de esto se utilizaron como predictoras variables categóricas que describen si el paciente posee o no el síntoma asociado, cabe aclarar que al modelo logístico se le pueden añadir tanto variables cualitativas como cuantitativas. Para llevar a cabo este procedimiento se utilizó el lenguaje de programación STAN ya que la distribución posterior no se halla con métodos convencionales (analíticamente) por lo tanto debe aproximarse a través de métodos numéricos. Para finalizar para hacer una comparación de modelos antes se verifica la convergencia del modelo y significancia de las variables predictoras.

2. Glosario

- **Dolor abdominal:** Es el dolor que se siente en el área entre el pecho y la ingle, a menudo denominada región estomacal o vientre.
- **Somnolencia:** Las personas que están somnolientas pueden quedarse dormidas cuando no quieren o en momentos que pueden generar problemas de seguridad.
- **Hipotensión:** Se considera baja a la presión arterial cuya lectura es inferior a 90 milímetros de mercurio (mm Hg) para el número superior (sistólica) o 60 mm Hg para el número inferior (diastólica). Lo que se considera presión arterial baja para una persona puede estar bien para otra. La presión arterial baja puede no causar ningún síntoma evidente o puede causar mareos y desmayos. Algunas veces, la presión arterial baja puede poner en riesgo la vida.
- **Hepatomegalia:** Hepatomegalia es el agrandamiento del hígado por encima de su tamaño normal. Ciertas condiciones como una infección, parásitos, tumores, anemias, estados tóxicos, enfermedades de almacenamiento, insuficiencia cardíaca, enfermedad cardíaca congénita y trastornos metabólicos pueden hacer que el hígado se agrande.
- **Hemorragia mucosa:** Es la pérdida de sangre del tejido que recubre la nariz. El sangrado ocurre con más frecuencia en una fosa únicamente. Hipotermia: La hipotermia es una urgencia médica que ocurre cuando el cuerpo pierde calor más rápido de lo que lo produce, lo que provoca una peligrosa disminución de la temperatura corporal. La temperatura corporal normal es de alrededor de 98,6 °F (37 °C). La hipotermia se produce cuando la temperatura del cuerpo cae por debajo de 95 °F (35 °C).
- **Aumento de hematocritos:** Cuando una persona tiene niveles altos de hematocrito tiende a presentar estos síntomas: piel enrojecida, mareos, problemas de la vista, dolores de cabeza y agrandamiento del bazo
- **Acumulación de líquidos:** Es un aumento en el volumen del líquido intersticial, es decir, una acumulación excesiva de líquidos en los tejidos.

3. Contexto

El dengue es una enfermedad infecciosa transmitida por mosquitos del género *Aedes* que afecta a millones de personas cada año en el mundo. El dengue se produce por la infección de uno de los cuatro serotipos del dengue (DENV-1, DENV-2, DENV-3 y DENV-4), que pertenecen a la familia *Flaviviridae*. La infección por uno de estos virus confiere inmunidad específica y temporal contra ese tipo, pero no contra los otros tres. Por lo tanto, una persona puede sufrir hasta cuatro episodios de dengue en su vida. Los síntomas del dengue varían según el tipo de virus y el estado inmunológico del individuo. El dengue clásico se caracteriza por fiebre alta, dolor de cabeza, dolor muscular y articular, erupción cutánea y malestar general. El dengue con signos de alarma es una forma grave y potencialmente mortal que se presenta con sangrado de las encías, la nariz o la piel, disminución del número de plaquetas, hipotensión y shock. El dengue grave también puede manifestarse con daño hepático, renal o neurológico. El diagnóstico del dengue se basa en la clínica, la epidemiología y las pruebas

de laboratorio. Estas últimas incluyen la detección del virus o sus antígenos en la sangre, la serología para identificar los anticuerpos específicos y la biología molecular para determinar el genotipo viral. El tratamiento del dengue es principalmente sintomático y de soporte, con hidratación, analgésicos y antipiréticos. No existe una vacuna eficaz ni un tratamiento específico para el dengue. La prevención del dengue se centra en el control de los vectores y la educación sanitaria. El control de los vectores implica la eliminación o el tratamiento de los criaderos de mosquitos, el uso de repelentes e insecticidas y la protección personal con ropa adecuada y mosquiteros. La educación sanitaria busca concienciar a la población sobre los riesgos del dengue, los signos de alarma y la necesidad de acudir al médico ante cualquier síntoma. El dengue es una enfermedad que representa un importante problema de salud pública a nivel mundial. Se estima que cada año se producen entre 50 y 100 millones de casos de dengue clásico y entre 500 mil y un millón de casos de dengue hemorrágico. El conocimiento sobre el dengue, sus causas, tipos y consecuencias es fundamental para su prevención y control.

4. Motivación

Teniendo en cuenta todo lo anterior y para dar una atención más inmediata a los pacientes que requieren ser hospitalizados, se creó un modelo el cual su objetivo es: dadas ciertas patologías identificadas en un paciente que se sabe que tiene dengue, identificar el tipo de dengue y así determinar que tipo de prioridad y tratamiento se le debe de dar y así salvar muchas vidas en el proceso.

5. Descripción de la base de datos

La base de datos fue proporcionada por la página web **medata.gov.co** que es el portal de datos públicos de Medellín que trabaja bajo la premisa: “la información es de todos y para todos”. La estrategia MEData integra automáticamente la información estratégica de las dependencias de la Alcaldía de Medellín en una plataforma Big Data, con el fin de facilitar la obtención, gestión, manipulación, análisis, modelado, representación y entrega de datos para resolver problemas complejos de planificación y gestión. La base de datos trata acerca de un registro de pacientes atendidos en las Instituciones Prestadoras de Servicios de Salud con diagnóstico probable o confirmado de Dengue. La cual fue publicada el 24 de septiembre del año 2019, la información de forma cruda (sin hacerle ninguna depuración o cambios a la BD) consta de 54713 observaciones y 38 variables (columnas), entre las cuales se encuentran:

Cuadro 1: Variables de la BD

Nombre	Tipo
ID	number
semana	factor
edad_	number

Nombre	Tipo
uni_med__	factor
sexo__	factor
nombre_barrio	factor
comuna	factor
tipo_ss__	factor
cod_ase__	factor
fec_con__	factor
ini_sin__	factor
tip_cas__	factor
pac_hos__	factor
cod_dpto_r	factor
cod_mpio_r	factor
cod_dpto_o	factor
cod_mpio_o	factor
desplazami	factor
cod_mun_d	factor
clas_dengue	factor
fiebre	factor
cefalea	factor
dolrretroo	factor
malgias	factor
artralgia	factor
erupcionr	factor
dolor_abdo	factor
vomito	factor
somnolenci	factor
hipotensio	factor
hepatomeg	factor
hem_mucosa	factor
hipotermia	factor
aum_hemato	factor
caida_plaq	factor
acum_liquievento	factor
evento	factor
year__	factor

Cambios a la base de datos

De todo este conjunto de variables tomamos un total de 17 columnas:

Cuadro 2: Variables de interes

Nombre	Tipo
clas_dengue	factor
fiebre	factor
cefalea	factor
dolrretroo	factor
malgias	factor
artralgia	factor
erupcionr	factor
dolor_abdo	factor
vomito	factor
somnolenci	factor
hipotensio	factor
hepatomeg	factor
hem_mucosa	factor
hipotermia	factor
aum_hemato	factor
caida_plaq	factor
acum_liquievento	factor

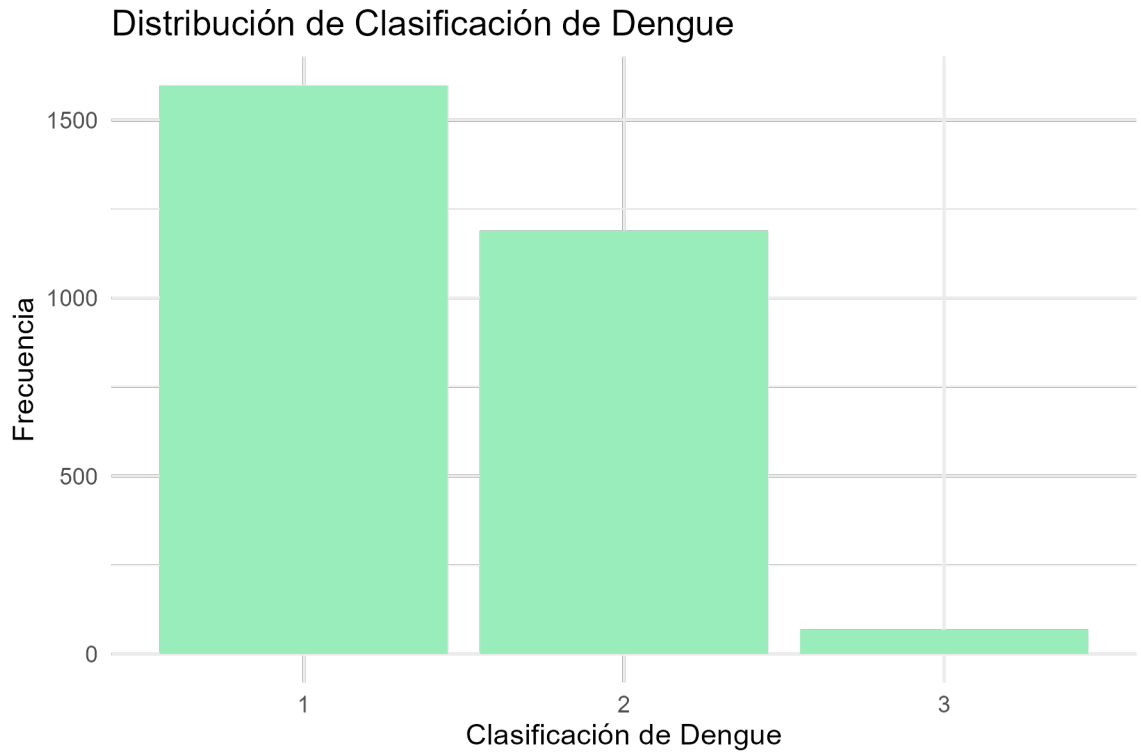
16 con patologias que posee los siguientes 3 niveles (variable categorica):

- **1:** Tiene el sintoma
- **2:** No tiene el sintoma
- **SD:** Sin informacion

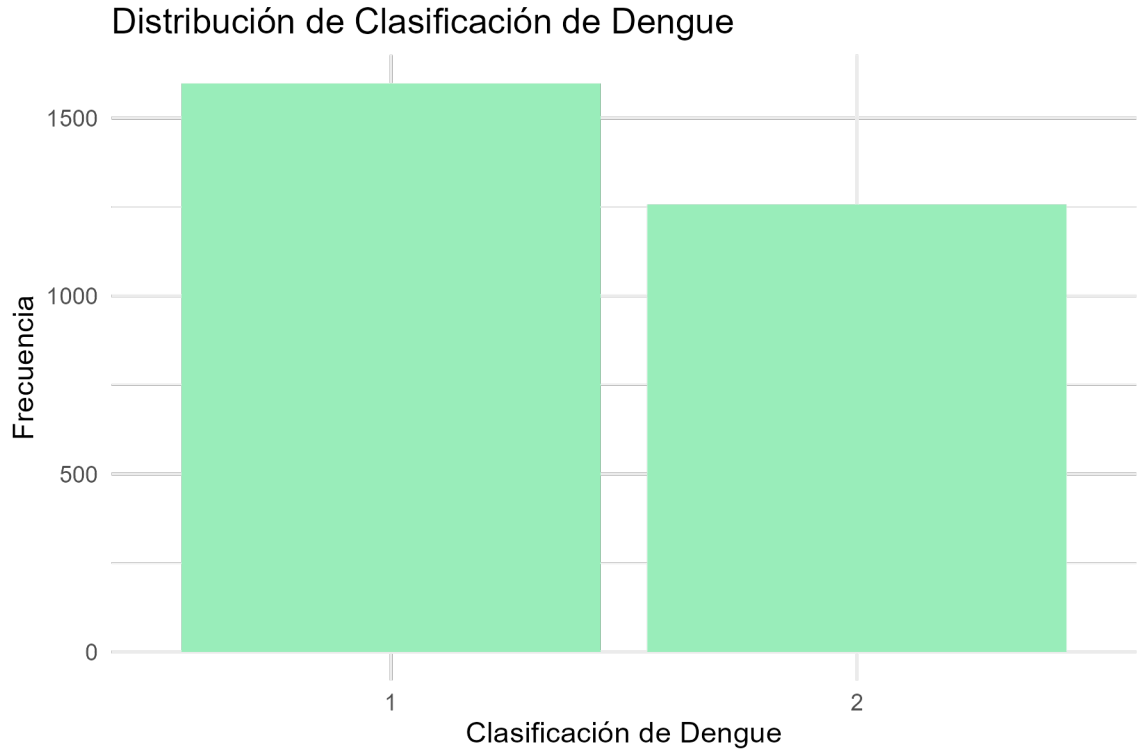
y por ultimo la clase de dengue, que posee 4 niveles:

- **1:** Dengue sin signos de alarma
- **2:** Dengue con signos de alarma
- **3:** Dengue grave
- **SD:** Sin informacion

Teniendo en cuenta estos niveles y que la falta de informacion en nuestra base de datos es tan alta, con un total de **50959** (los cuales estan distribuidos entre las distintas columnas que seleccionamos), lo cual es casi un **95 %** del el numero total de registros, asi que para no afectar la calidad de las estimaciones y confiabilidad de este modelo y ademas por temas computacionales se decidio eliminarlos, para quedar con un total de **2854** datos.



Por ultimo cabe mencionar que las observaciones correspondientes al dengue grave (nivel 3), las incluimos en el dengue con signos de alarma (nivel 2) por la poca cantidad de observaciones que habia en el dengue tipo 3 (68 observaciones) y ademas de esto para que todas nuestra variables fueran dicotomas.



Análisis de significancia de las variables predictoras con respecto a

La prueba de chi-cuadrado, también conocida como prueba de independencia o prueba de bondad de ajuste, es una prueba estadística utilizada para determinar si existe una asociación significativa entre dos variables categóricas en una población. En el contexto de un análisis de correspondencia, la prueba de chi-cuadrado puede ser utilizada para evaluar la independencia entre las variables en la tabla de contingencia.

La hipótesis nula de la prueba de chi-cuadrado establece que no hay asociación entre las variables categóricas en la población, mientras que la hipótesis alternativa sugiere que hay una asociación significativa. El estadístico de prueba chi-cuadrado se calcula como la suma de los residuos al cuadrado dividida por los valores esperados bajo la hipótesis nula.

$$\begin{cases} H_0 : \text{No existe una relacion significativa entre las variables} \\ H_a : \text{Existe una relacion significativa entre las variables} \end{cases}$$

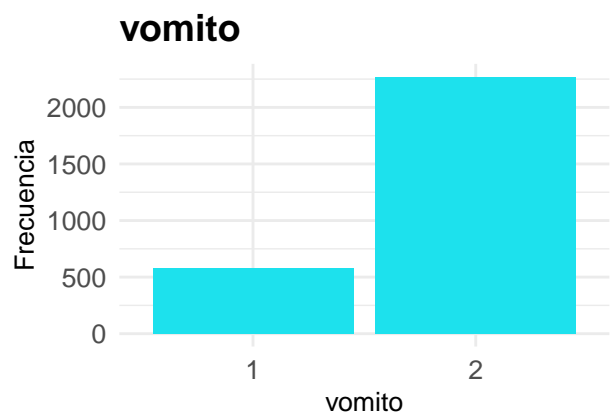
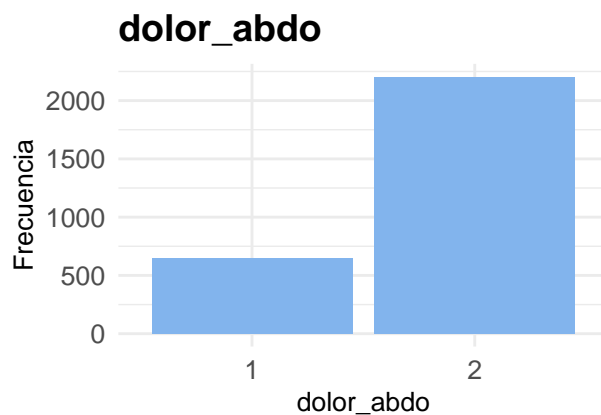
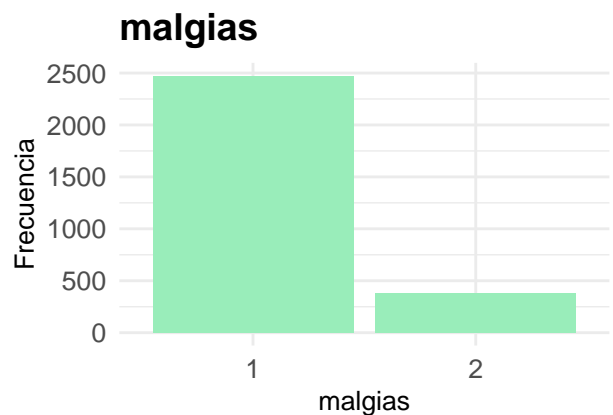
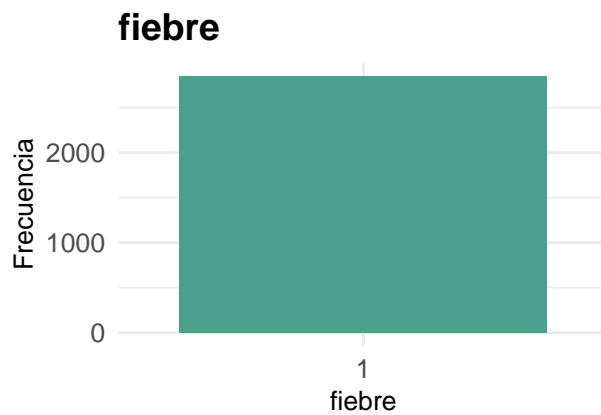
Cuadro 3: Significancia variables

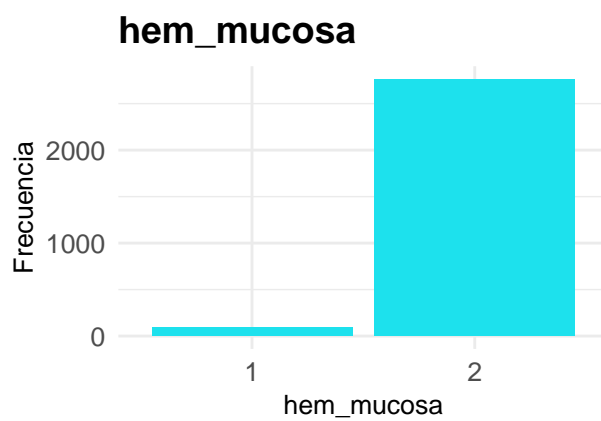
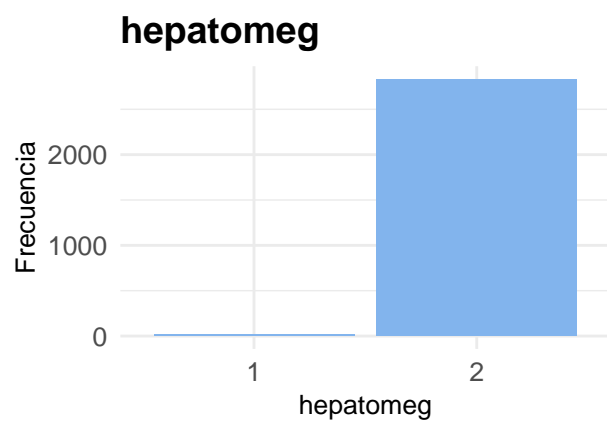
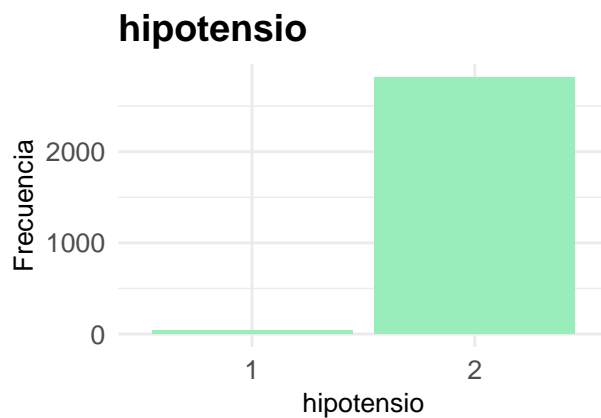
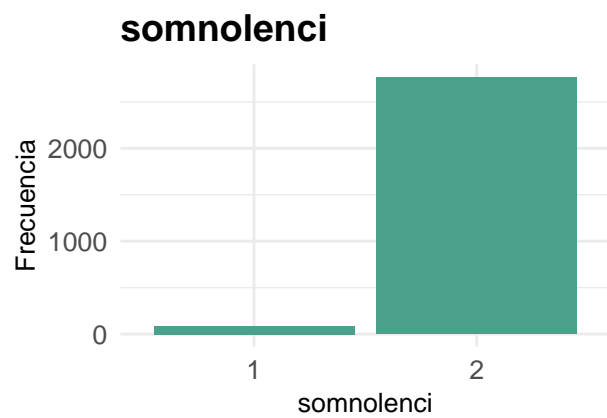
Variable	Valor_p
artralgia	0.9438053
dolrretroo	0.2241499
erupcionr	0.2012671
cefalea	0.0982721
malgias	0.0274714

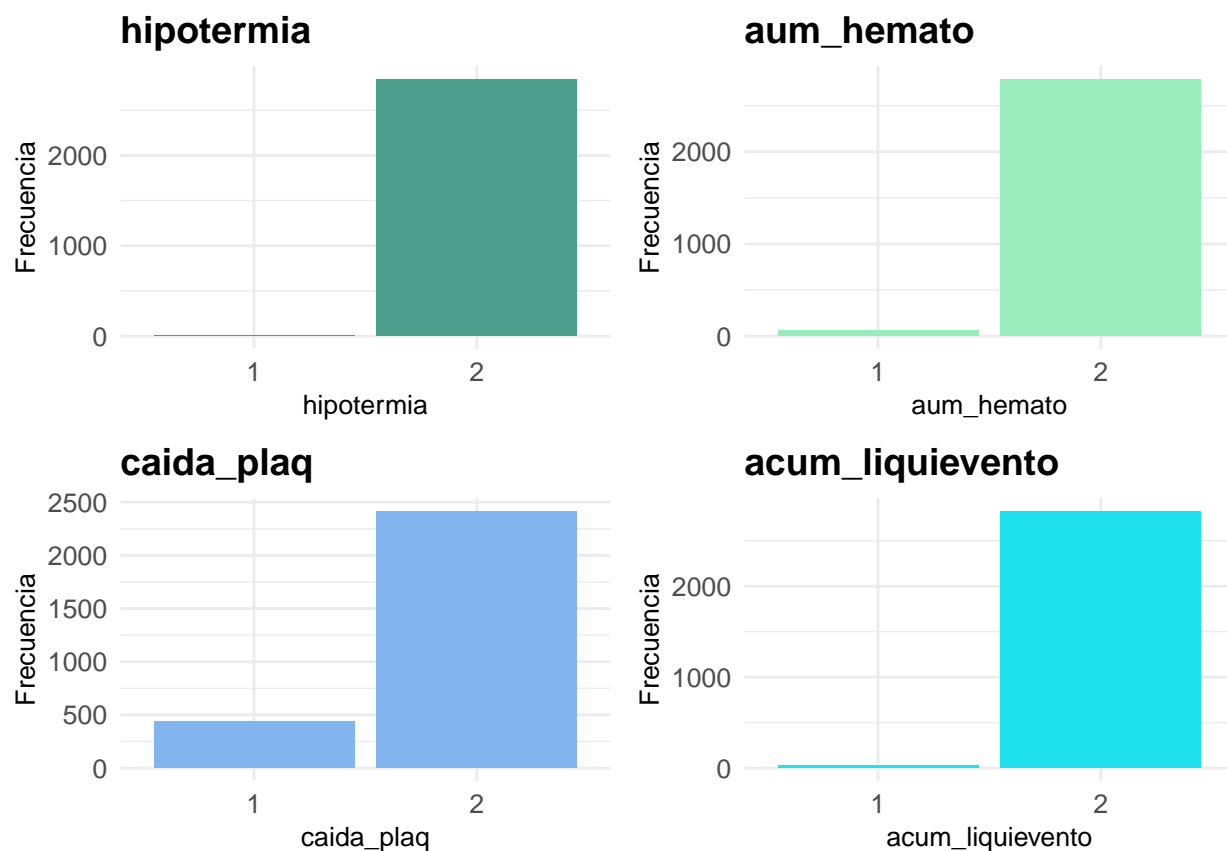
Variable	Valor_p
hipotermia	0.0045686
hepatomeg	0.0000007
acum_liquievento	0.0000000
fiebre	0.0000000
hipotensio	0.0000000
aum_hemato	0.0000000
somnolenci	0.0000000
hem_mucosa	0.0000000
caida_plaq	0.0000000
vomito	0.0000000
dolor_abdo	0.0000000

Tomando un $\alpha = 0.05$, teniendo suficiente evidencia estadística para decir que las variables: artralgias, dolor retro, erupción y cefalea no son significativas para explicar el comportamiento de la variable **clas_dengue**.

Estas son las variables resultantes:







Como se puede ver en estos graficos todos los pacientes tienen fiebre, por lo tanto no tendria sentido meterla en el modelo ya que no se mueve entre ambos mundos (dengue sin signos de alarma y dengue con signos de alarma) y eso le daria un sesgo al modelo

Cuadro 4: Vomito

.	1	2
1	0	1597
2	583	674

Cuadro 5: Dolor abdominal

.	1	2
1	0	1597
2	649	608

Cuadro 6: Hemorragia mucosa

.	1	2
1	0	1597
2	90	1167

Cuadro 7: Hipotermia

.	1	2
1	0	1597
2	8	1249

6. Metodología

Despues de analizar la base de datos y la información proporcionada, surge como pregunta de interes ¿Como optimizar la clasificación de pacientes que presentan dengue a partir de los sintomas que presentan?, esto permitiendo que su atención y uso de los recursos medicos disponibles sea más apropiada y aprovechada al maximo. Para poder responder a la interrogante y tomando en cuenta el análisis descriptivo de los datos, se plantean dos modelos donde uno corresponde a los sintomas comunes y otro a los sintomas poco comunes en pacientes con dengue (utilizando las variables que resultaron significativas en el test chi-cuadrado), para ambos la variable respuesta es **clas_dengue**, tanto la variable respuesta como las variables predictoras son de tipo binarias.

Dada la naturaleza del problema los modelos propuestos se definen como:

$$\begin{aligned} \mathbf{Y}_i &\sim \text{Bernoulli}(\theta_i), \quad y_i = 0, 1 \quad i = 1, \dots, n \\ \beta &\sim N(\mu, \sigma^2) \end{aligned}$$

De donde su función de enlace,

$$\begin{aligned} \text{Logit}(\theta_i) &= \ln\left(\frac{\theta_i}{1 - \theta_i}\right) = \beta^T \cdot \mathbf{X} \\ \theta_i &= \frac{e^{\beta^T \mathbf{X}}}{1 + e^{\beta^T \mathbf{X}}} \end{aligned}$$

6.1. Modelo 1: Sintomás comunes

El modelo 1 se describe como:

Natación	Variable	Función de la variable en ei modelo
Y	Clase de dengue	Variable de respuesta
X_1	Dolor abdominal	Predictora
X_2	Vomito	Predictora
X_3	Somnolencia	Predictora
X_4	Hipotensión	Predictora
X_5	Hepatomegalia	Predictora

Ejecutado en el lenguaje de programación stan se obtiene que:

	mean	semean	sd	2.5 %	25 %	50 %	75 %	97.5 %	neff	Rhat
beta[1]	162.23	1.97	60.59	57.63	118.57	157.46	200.97	291.64	946	1
beta[2]	-82.20	1.53	49.33	-193.85	-112.78	-75.08	-43.59	-11.35	1045	1
beta[3]	-56.97	1.09	38.10	-146.89	-78.73	-48.88	-27.07	-8.71	1223	1
beta[4]	-45.70	1.03	34.30	-131.48	-64.64	-37.33	-18.64	-5.12	1116	1
beta[5]	-2.96	1.53	47.79	-105.19	-30.36	-0.91	25.88	89.45	977	1
beta[6]	23.86	1.87	57.90	-92.95	-12.80	23.16	61.62	138.74	959	1

En general, estos resultados indican los valores estimados de los coeficientes del modelo junto con la incertidumbre asociada a cada estimación. Los intervalos de confianza del 95 % proporcionan un rango dentro del cual se espera que se encuentre el verdadero valor del coeficiente con un nivel de confianza del 95 %. Los valores bajos de la estadística Rhat y el número efectivo de muestras indican que las cadenas de MCMC han convergido de manera satisfactoria, lo que aumenta la confianza en los resultados obtenidos. Sin embargo cuando se observan los graficos de densidad posterior y los traceplot se aprecia que no alcanza a terminar de converger dado el poder de computo con el que se contaba para ejecutar el modelo.

6.2. Modelo 2: Sintomás menos comunes

El modelo 2 se describe como:

Natación	Variable	Función de la variable en ei modelo
Y	Clase de dengue	Variable de respuesta
X_1	Hemorragia en Mucosa	Predictora
X_2	Hipotermia	Predictora
X_3	Aumento de Hemato	Predictora
X_4	Acumulación de liquidos	Predictora

Ejecutado en el lenguaje de programación stan se obtiene que:

	mean	semean	sd	2.5 %	25 %	50 %	75 %	97.5 %	neff	Rhat
beta [1]	170.48	1.99	56.80	71.06	129.62	167.05	206.78	292.37	817	1.01
beta [2]	-43.88	1.05	33.47	-128.40	-61.15	-35.42	-18.19	-5.37	1013	1.01
beta [3]	-41.11	0.93	32.76	-124.62	-57.68	-32.72	-16.10	-3.76	1228	1.00
beta [4]	-45.69	1.20	34.08	-132.44	-63.04	-36.65	-19.30	-2.21	807	1.00
beta [5]	-40.76	0.94	31.45	-117.72	-58.39	-32.57	-15.92	-4.18	1123	1.00

7. Resultados

Continuando con la misma línea de reflexión, se escogieron tres criterios para comparar la calidad predictiva de los modelos:

7.1. DIC (Deviance Information Criterion)

El DIC (Deviance Information Criterion, por sus siglas en inglés) es una medida utilizada en la inferencia bayesiana y el modelado estadístico para evaluar la calidad de un modelo. Se basa en el principio de que un buen modelo debe ajustarse bien a los datos y, al mismo tiempo, ser parsimonioso.

Se expresa como:

$$DIC = -2 \log \left(f(y \mid \hat{\theta}_{\text{Bayes}}) + 2PDIC \right)$$

Donde $\vec{\theta}_{\text{Bayes}} = E[\theta \mid y]$ y $PDIC$ es el número efectivo de parámetros.

El valor del DIC se interpreta de manera relativa, es decir, se compara entre los modelos, los cuales están ajustados a los mismos datos. Un modelo con un valor de DIC más bajo se considera preferible, lo que indica un mejor equilibrio entre el ajuste y la complejidad en comparación con otros modelos.

Modelo	DIC
1 (Síntomas comunes)	601.2989
2 (Síntomas poco comunes)	986.0014

Dado que el modelo 1 presenta el menor DIC se considera que presenta un mejor ajuste a la hora de clasificar pacientes.

7.2. Curva ROC (Receiver Operating Characteristic)

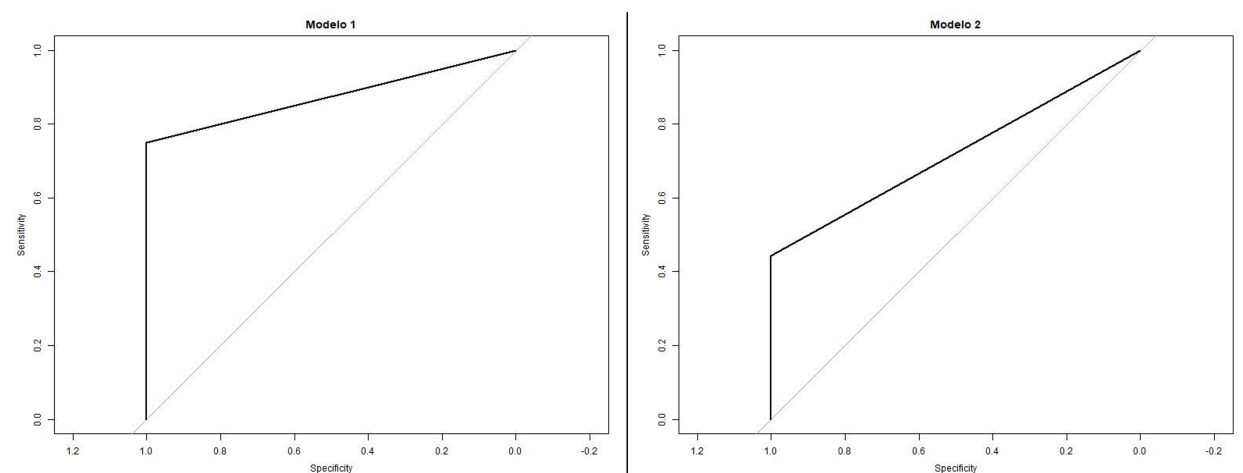
La curva ROC (Receiver Operating Characteristic) es una herramienta ampliamente utilizada para evaluar y visualizar el rendimiento de un clasificador en problemas de clasificación

binaria. Esta curva representa gráficamente la relación entre la sensibilidad (tasa de verdaderos positivos) y la especificidad (tasa de verdaderos negativos) en diferentes puntos de corte del clasificador.

En el eje X de la curva ROC se representa la tasa de falsos positivos, que es la proporción de instancias negativas incorrectamente clasificadas como positivas. En el eje Y se representa la tasa de verdaderos positivos, que es la proporción de instancias positivas correctamente clasificadas como positivas.

Cuanto más cercana esté la curva ROC al vértice superior izquierdo del gráfico, mejor será el rendimiento del clasificador, ya que se alcanzarán altas tasas de verdaderos positivos con bajos falsos positivos. El área bajo la curva ROC (AUC-ROC) es una medida comúnmente utilizada para resumir el rendimiento general del clasificador: un valor cercano a 1 indica un buen rendimiento, mientras que un valor cercano a 0.5 indica un rendimiento aleatorio.

Al realizar la curva ROC y calcular el área debajo de ella se obtuvo:



Modelo	AUC-ROC
1 (Síntomas comunes)	0.8747
2 (Síntomas poco comunes)	0.7214

El modelo 1 cuenta con un mayor AUC-ROC, lo que significa que presenta un mejor rendimiento a la hora de clasificar en la clase de dengue a los pacientes frente al modelo 2.

7.3. LOOCV (Leave-One-Out Cross-Validation)

El LOOCV (Leave-One-Out Cross-Validation) es una técnica de validación cruzada ampliamente utilizada en el campo del modelado estadístico. Consiste en dividir el conjunto de datos en n bloques, donde n es el número total de observaciones. En cada iteración, se ajusta el modelo utilizando $n-1$ bloques y se evalúa su rendimiento al predecir la observación excluida.

El LOOCV es especialmente útil cuando el tamaño del conjunto de datos es limitado. Proporciona una estimación precisa del rendimiento del modelo, ya que utiliza todos los datos

disponibles en cada iteración. Además, el LOOCV evita el sesgo de subestimación que puede ocurrir con otras técnicas de validación cruzada.

Modelo	<i>elpd_loo</i>	<i>p_loo</i>	looic
1 (Síntomas comunes)	-229.2	1.1	584.5
2 (Síntomas poco comunes)	-486.1	1.0	972.2

Dentro de las medidas de decisión de LOOCV, se tomaron en cuenta *elpd_loo*, *p_loo*, *looic*.

elpd_loo: ELPD calculado utilizando LOOCV. Es una medida de la calidad de ajuste y la capacidad predictiva del modelo. Cuanto mayor sea el valor, mejor será el modelo en términos de ajuste y predicción.

p_loo: Medida de precisión de la estimación de ELPD utilizando LOOCV. Cuanto menor sea el valor, mayor será la precisión de la estimación.

looic: Información de criterio de información bayesiano (BIC) calculada utilizando LOOCV. Es una medida de la calidad de ajuste y la complejidad del modelo. Valores más bajos indican un mejor ajuste y modelos más parsimoniosos

Como se observa en cada una de las tablas el modelo 1 es mejor que el modelo 2 en cada una de las medidas de decisión ya que presenta un mayor valor en el **elpd_loo**, un menor valor en el **p_loo** y un valor más bajo en el **looic**, aunque en esta última el modelo 2 tenga menos variables predictoras (lo cual se podría creer más parsimonioso), se obtiene un mejor resultado con el modelo 1.

7.4. Modelo Seleccionado

En resumen para los tres criterios de selección, el modelo con mejor ajuste y rendimiento predictivo es el conformado por las variables dolor abdominal, vómito, somnolencia, hipotensión, hepatomegalia, correspondientes al modelo 1 de síntomas comunes, cuyas estimaciones son:

Betas	Variables	2.5 %	50 %	97.5 %
$\beta[1]$	Intercepto	57.63	157.46	291.64
$\beta[2]$	Dolor abdominal	-193.85	-75.08	-11.35
$\beta[3]$	Vómito	-146.89	-48.88	0.76
$\beta[4]$	Somnolencia	-131.48	-37.33	-5.12
$\beta[5]$	Hipotensión	-105.19	-0.91	89.45
$\beta[6]$	Hepatomegalia	-92.95	23.16	138.74

Dado el formato de la estructura, es necesario realizar el cálculo de la exponencial de las estimaciones con el fin de facilitar su interpretación. Por lo tanto, podemos expresar que:

$$\begin{aligned}\log\left(\frac{\theta}{1-\theta}\right) &= \boldsymbol{\beta} \cdot \mathbf{X} \\ \log(\text{odds}) &= \boldsymbol{\beta} \cdot \mathbf{X} \\ \text{odds} &= e^{\boldsymbol{\beta} \cdot \mathbf{X}}\end{aligned}$$

	mean	2.5 %	97.5 %
beta[1]	2.854919×10^{70}	1.067557×10^{25}	4.546139×10^{126}
beta[2]	1.999832×10^{-36}	6.486564×10^{-85}	1.176949×10^{-5}
beta[3]	1.812355×10^{-25}	1.608731×10^{-64}	1.649283×10^{-4}
beta[4]	1.421485×10^{-20}	7.924311×10^{-58}	5.976023×10^{-3}
beta[5]	5.181892×10^{-2}	2.072829×10^{-46}	7.041114×10^{38}
beta[6]	$2.3028536472 \times 10^{10}$	4.288722×10^{-41}	1.794801×10^{60}

Para hacer este analisis hay que recordar que en nuestro modelo, como todas son variables categoricas se trata de tener o no tener el sintoma, teniendo en cuenta la tabla anterior, se puede decir que:

- El tener dolor abdominal da una probabilidad de un 99.99882 % de tener dengue tipo 1
- El tener vomito da una probabilidad de un 99.98351 % de tener dengue tipo 1
- El tener somnolencia da una probabilidad de un 99.4024 % de tener dengue tipo 1
- El tener hipotencion da una probabilidad de un 7.041114×10^{-40} % de tener dengue tipo 2, lo cual quiere decir que si se tiene este sintoma ya es un gran indicio de tener dengue tipo 2
- El tener hepatomegalia da una probabilidad de un 1.794801×10^{-62} % de tener dengue tipo 2, lo cual significa que este puede ser un sintoma que orienta a determinar que el paciente tiene este tipo de dengue

8. Conclusión

Como se demuestra en las investigaciones hechas por los expertos acerca de el dengue, ciertas patologias son indicativos fuertes de un tipo de dengue, lo cual da las indicaciones de que se posee un tipo de dengue en especifico dado que ya posee la enfermedad, y como vimos en la eleccion de modelos, los sintomas comunes ayudan mucho mas a predecir cual tipo de dengue tiene el paciente que los sintomas poco comunes ya que tal vez se puede llegar a asociar a otros problemas que tenga la persona. Por ultimo cabe resaltar que hay ciertas patologias que soy muy influyentes a la hora de determinar el tipo de dengue, lo cual se vio en el analisis del ODDS.

9. Anexos

En el siguiente link se redirecciona a un repositorio donde se encuentra todo el trabajo realizado y los codigos empleados para su desarrollo:

https://github.com/cristiancaz2311/Clasificacion_de_pacientes_con_Dengue-Bayesiana

o puede usar el siguiente código QR:



10. Referencias

- Alcaldía de Medellín. (n.d.). Alcaldía De Medellín. <https://www.medellin.gov.co/irj/portal/medellin?NavigationTarget=contenido/6991-MEDData-el-portal-de-datos-publicos-del-Municipio-de-Medellin>
- Ametameric. (n.d.). <https://pippin.gimp.org/ametameric/>
- Padilla, J. C. P., Rojas, D. P. R., & Sáenz-Gómez, R. S. (n.d.). Dengue en Colombia. Minsalud. [https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/INEC/INV/Dengue en Colombia.pdf](https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/INEC/INV/Dengue%20en%20Colombia.pdf)
- Dengue. (n.d.). Ministerio De Salud Y Protección Social. <https://n9.cl/ozwdm>
- Sosa, A. a. S., Ayala, F. A., Martínez De Cuellar, C. M. C., Borba, J. B., & Chamorro, G. C. (2012). Dengue guía del manejo clínico. Retrieved June 20, 2023, from https://iris.paho.org/bitstream/handle/10665.2/10101/9789996768422_esp.pdf?sequence=1&isAllowed=y
- Pando Ferrer, R., Ferrer, P. F., & Centro de Atención y Educación al Diabético de Bauta. (n.d.). Manifestaciones oftalmológicas del dengue, el Zika y el Chikungunya. Scielo. Retrieved June 20, 2023, from http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-21762021000400008

- Dengue. (s. f.). OPS/OMS | Organización Panamericana de la Salud. <https://n9.cl/nxtml>
- World Health Organization: WHO & World Health Organization: WHO. (2023). Dengue y dengue grave. [www.who.int. https://www.who.int/es/news-room/fact-sheets/detail/dengue-and-severe-dengue](https://www.who.int/es/news-room/fact-sheets/detail/dengue-and-severe-dengue)
- Síntomas y tratamiento del Dengue | CDC. (2021, 20 septiembre). Centers for Disease Control and Prevention. <https://www.cdc.gov/dengue/es/symptoms/index.html>
- Nrc, R. A. /. (2023, 17 mayo). Perú vive el peor brote de dengue del siglo. Aristegui Noticias. <https://aristeginoticias.com/1705/mundo/peru-vive-el-peor-brote-de-dengue-del-siglo/>
- Antulio, H. R. (s. f.). Factores de riesgos asociados a la infección por dengue en San Mateo, Anzoátegui, Venezuela. http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-21252011000300009