
Análisis estadístico de los factores de riesgo que inciden en la adquisición de tos ferina en recién nacidos de la ciudad de Bogotá

Statistical analysis of the risk factors to affect the acquisition of pertussis in newly born of the city of Bogotá

Edwin Ramiro Hernández Herrera^a
edwin.hernandez@usantotomas.edu.co

Diego Lemus Polania^b
dflemusp@gmail.com

Wilmer Pineda Ríos^c
wilmerpineda@usantotomas.edu.co

Resumen

A nivel mundial, se han presentado grandes avances tecnológicos que han posibilitado el aumento de estudios científicos, para el mejoramiento de la salud humana. De allí que muchos de estos estudios están enfocados al análisis de la eficiencia y efectividad de algunas vacunas, que ayudan a mitigar los síntomas o a prevenir enfermedades que pueden llegar a causar la muerte a un ser humano. Por tal razón, en la actualidad, la Secretaría Distrital de Salud se encuentra promoviendo el seguimiento a las madres gestantes para garantizar que transmitan anticuerpos a los recién nacidos, para evitar así la adquisición de la tos ferina. Por tal razón, la presente investigación propone la construcción de umbrales de clasificación, a través de la utilización métodos de simulación con el objetivo de determinar el comportamiento de los anticuerpos en las diferentes muestras tomadas en la ciudad de Bogotá, y así identificar los procesos de seroconversión, aumento, ausencia y/o transmisión de éstos. Adicionalmente, se plantean modelos de regresión gamma con el propósito de determinar los factores asociados a la producción de anticuerpos en la madre y la transmisión de éstos al recién nacido. De esta manera, el resultado final de la presente investigación consiste en el establecimiento de los perfiles de madres de la ciudad de Bogotá, que posee mayor probabilidad de transmisión de anticuerpos a sus hijos durante el período de gestación, con el fin de prevenir la adquisición de la tos ferina, siendo esta información relevante para el planteamiento de campañas preventivas y así tener mayor impacto en la población más vulnerable a esta enfermedad.

Palabras clave: Umbrales de clasificación, Tos ferina, Regresión Gamma.

Abstract

At the global level, great technological advances have made possible the increase of scientific studies, for the improvement of human health. Hence, many of these studies are focused on the analysis of the efficiency and effectiveness of some vaccines, which help to mitigate symptoms or prevent diseases. Specifically, the District Health Secretary is interested in determining the ability of pregnant mothers to develop antibodies against whooping cough and the transmissibility of these to the newborn in order to prevent it from catching the disease in its first trimester of life. In the present investigation, we

^aEstudiante de Estadística Universidad Santo Tomás Bogotá

^bConsultor Estadístico Independiente

^cDocente de Estadística Universidad Santo Tomás Bogotá

propose a method for the identification of the threshold of seroconversion in mothers, the threshold of transmissibility to the newborn and the threshold to identify the loss of antibodies in newborns. Finally, the factors associated with the production of antibodies in the mother and the transmission of these to the newborn were identified using gamma regression models.

Keywords: Classification thresholds, Pertusis, Gamma regression.

1. Introducción

El sector salud, en los últimos años ha presentado una gran evolución influencia por las nuevas tecnologías, que han permitido la adecuación y aplicación de nuevas técnicas científicas; especialmente en el campo de la epidemiología y cuyos resultados generan condiciones o factores que permiten el planteamiento de medicamentos, como por ejemplo el caso de las vacunas, que tienen como fin contrarrestar los efectos de una enfermedad o incluso prevenir la adquisición de la misma.

Es así, como la Secretaría Distrital de Salud (SDS) ha alertado a la comunidad en los últimos años sobre el aumento de los casos de adquisición de la tos ferina en recién nacidos y, a su vez, el Instituto Nacional de Salud (INS) (citado en Dario La Economía 2016) afirma que, para el año 2016 se reportaron 4131 casos probables de tos ferina, obteniéndose 369 casos confirmados en el 2016, de los cuales el 79.9% se presentaron en menores de un año. En consecuencia, con lo anterior, se han reportado diez muertes por esta causa, donde seis se han confirmado por laboratorio y cuatro continúan en estudio. Al ser la mortalidad infantil de gran relevancia para la SDS, esta institución estatal se encuentra interesada disminuir la tasa observada hasta el momento en el distrito sin importar su causa. En el caso de la tos ferina, la construcción de umbrales de seroconversión y la determinación de los factores asociados a la producción de anticuerpos en la madre y la transmisión de éstos al recién nacido toma fuerza; en ese sentido, establecer perfiles para madres gestantes de la ciudad de Bogotá resulta de gran importancia para aumentar la probabilidad de transmisión de anticuerpos a sus hijos durante el período de gestación, y así evitar que los recién nacidos se enfermen de tos ferina en su primer trimestre de vida. Para tal fin, se propone la utilización de un modelo de simulación para la determinación del umbral de seroconversión en las madres, umbral de transmisibilidad de anticuerpos al neonato y umbral para identificar la pérdida de títulos en recién nacidos. Por otro lado, para la identificación de los factores asociados con la adquisición de tos ferina se propone el uso de la regresión gamma; todo lo anterior enmarcado para la ciudad de Bogotá, dado que este tipo de estudio para población Latinoamérica no poseen desarrollo, ya que se focalizan en población estadounidense y europea.

En ese sentido, el presente estudio busca determinar el método más adecuado para la identificación del umbral de seroconversión en las madres (etapa 1), el umbral de transmisibilidad al neonato (etapa 2) y el umbral para identificar la pérdida de los anticuerpos en los recién nacidos (etapa 3). Con base en lo anterior, se realiza la modelación de una regresión Gamma $Y_i|\mu_i, \phi \sim \text{Gamma}(\mu_i, \phi)$, cuya variable objetivo es la cantidad de títulos y las variables significativas de dicha modelación son los factores que inciden en la adquisición de la tos ferina. Por otro lado, las variables significativas son traspolarlas en la modelación de una regresión logística $Y_i|\mu_i \sim \text{Bin}(\mu_i, n_i)$, cuya variable respuesta se crea a partir de la adjudicación de cortes para determinar si se presenta seroconversión, transmisión o ausencia de anticuerpos. Bajo dicha propuesta, se determina el mejor umbral a través del estadístico AUC y así lograr ratificar bajo la curva de ROC la selección del umbral.

Esta investigación, no sólo busca un impacto por el aumento actual de casos de recién nacidos con tos ferina, sino busca también la implementación de los umbrales y de los factores asociados en la producción de anticuerpos y su posterior transmisión al recién nacido para el sistema de salud colombiano. Lo anterior con el objetivo de promover campañas preventivas, así como el seguimiento estricto a la población más propensa a la adquisición de la tos ferina. Sumado a esto, los resultados de la presente investigación permitirán establecer datos para la población colombiana que, finalmente, a la fecha, no se tienen determinados, con lo que se puede detectar el riesgo en la adquisición de la enfermedad, y de

manera implícita controlar y/o disminuir la tasa de mortalidad en recién nacidos.

El documento tiene la siguiente estructura: en la sección 2 se plantea el contexto social de la investigación con respecto a la enfermedad de la tos ferina. En la sección 3 se presenta la teoría estadística de los modelos propuestos para la identificación de los factores asociados al comportamiento de los anticuerpos en la madre gestante y el neonato. En la sección 4 se encuentran los conceptos inmersos en la construcción de los umbrales de clasificación. En la sección 5 se enuncian los factores y escenarios considerados en el estudio de simulación y los resultados del mismo, seguidos por una aplicación empírica de los conceptos estadísticos considerados. En la última sección, se presentan las conclusiones de la investigación realizada y finalmente en el apéndice se puede evidenciar el código computacional utilizado en la investigación.

2. Tos ferina

Dentro del contexto de la presente investigación se encuentran las especificaciones en torno a la tos ferina, dado que es necesario conocer aspectos generales, epidemiológicos, clínicos, de diagnóstico y de tratamiento, los cuales serán abordados en esta sección.

2.1. Aspectos generales y clínicos

El Ministerio de Salud y Protección Social (2014) (MSPS) afirma que la tos ferina es una enfermedad aguda causada por la bacteria *Bordetella parapertusis*, la cual se encuentra dentro de la familia de cocobacilos, de la clase Gram-negativo no móvil capaz de producir múltiples productos activos, tanto antigénicos como biológicos. Resulta importante señalar que esta bacteria tiene como único huésped el ser humano, y su transmisión se da a través de las gotas que se producen al toser o estornudar, también se da al estornudar, adicional al presentar contacto con las secreciones respiratorias de una persona infectada. Esta enfermedad se cataloga como altamente contagiosa. (Ministerio de Salud y Protección Social 2014) La tos ferina, tal como lo plantea la SDS, posee una duración aproximada de ocho semanas y se desarrollan en tres importantes fases, a saber:

- **Fase catarral:** Se presentan los primeros síntomas, tales como malestar, anorexia y tos nocturna que a medida que evolucionan los días se convierte en diurna. Esta fase puede presentarse entre una o dos semanas.
- **Fase paraxística:** Se presenta tos en quinta, es decir tos con cinco a quince golpes en expiración que terminan en un silbido. Algunos síntomas adicionales que se pueden presentar es vómito, rara vez pérdida del conocimiento y convulsiones por la reducción de suministro de oxígeno al cerebro. Esta fase puede tener una duración entre dos o cuatro semanas.
- **Fase convaleciente:** Se da inicio entre la cuarta y sexta semana de la iniciación de la enfermedad, allí la tos se vuelve menos fuerte y el vómito desaparece. En esta etapa se puede presentar algunas complicaciones tales como encefalopatía con pérdida de conciencia, convulsiones, atelectasia pulmonar, neumonía, enfisema mediastinal, neumotórax, bronquiectasias, neumonía y finalmente la muerte.

Dado lo anterior (SDS 2018) indica que ha intensificado los esfuerzos en el proceso de vacunación, ya que se ha reportado que entre enero y marzo del presente año 28 casos, 50 % en menores de un año y la mayoría en las localidades de Suba, Usaquén y Ciudad Bolívar.

2.2. Aspectos Epidemiológicos

Tal como se mencionó anteriormente y lo reafirma (D.L Van Esso 2014), el agente etiológico más habitual de la tos ferina es la bacteria *Bordetella Pertusis*, destacando que dentro de este tipo de bacterias se han

identificado otras especies que son causantes del cuadro clínico indistinguible de la tos ferina. (D.L Van Esso 2014) expresa que este tipo de infección puede atacar a cualquier grupo de edad, pero los casos más graves se han detectado en neonatos y lactantes en los primeros meses de vida, Al ser una enfermedad altamente contagiosa, presenta mayor probabilidad de contagio en la fase catarral, en la cual se recomienda aislamiento mínimo de cinco días con suministro de antibiótico.

Este tipo de enfermedad no posee un comportamiento habitual al de otras enfermedades que son prevenibles a través de la vacunación y confieren inmunidad al paciente, por el contrario, con la vacunación se generan anticuerpos y proporciona una protección que va desapareciendo paulatinamente en el transcurrir del tiempo; de hecho, en algunos casos las personas, vacunadas pueden llegar a contraer la enfermedad después de recibir la vacuna. (D.L Van Esso 2014)

2.3. Diagnóstico

(D.L Van Esso 2014) Afirma que esta enfermedad es de fácil detección en lactantes y niños no vacunados, mientras su detección se complica en pacientes que el cuadro clínico no es caracterizado por la tos típica. El diagnóstico de laboratorio se realiza a través de una muestra de moco nasofaríngeo, dicha muestra puede adquirirse mediante aspirado, lavado nasal o frotis nasofaríngeo a través de escobillones de Dacrón o alginato cálcico. Una buena herramienta para la detección tardía de la enfermedad es la serología, pero a nivel clínico su utilización es poco usual. Sin embargo, en lactantes y niños pequeños se observa en el hemograma una leucocitosis a expensas de un aumento significativo de los linfocitos. (D.L Van Esso 2014))

2.4. Tratamiento

Para el control de esta enfermedad se deben aplicar dos tipos de tratamientos (SDS 2018): el sintomático y el etiológico. El primer tratamiento en mención tiene como fin disminuir la intensidad y la frecuencia de la tos, con la ayuda de la oxigenación, hidratación y la alimentación adecuada por parte del paciente. El segundo tratamiento tiene como objetivo eliminar la infección de la nasofaringe y así no permitir la transmisión. (D.L Van Esso 2014) afirma que la eficacia de este tipo de tratamiento depende de la semana en la cual se da inicio, dado que después de la tercera semana de evolución de la enfermedad el beneficio es casi nulo.

2.5. Prevención

Para contrarrestar la adquisición de la tos ferina, principalmente es mediante la vacunación, permitiendo así disminuir las tasas de mortalidad, aunque los resultados aún no están en ceros dado que en recién nacidos y lactantes por su corta edad, no pueden ser vacunados, o no han recibido el total de las dosis o son contagiados en su entorno familiar. En Colombia, la (SDS 2018) afirma que, para los niños entre los dos meses y seis años de edad, el cuadro de vacunación esta dado por cinco dosis: la prevalente a los 2, 4, y 6 meses y la DPT a los 18 meses y a los 5 años de edad. En Colombia, la (SDS 2018) afirma que, para los niños entre los dos meses y seis años de edad, el cuadro de vacunación esta dado por cinco dosis: la prevalente a los 2, 4, y 6 meses y la DPT a los 18 meses y a los 5 años de edad. Así mismo, la (SDS 2018) realiza algunas recomendaciones en caso de contraer la enfermedad, con el fin de prevenir su propagación:

- Tomar bastante líquido y comer frutas.
- Ingerir alimentos en porciones pequeñas y de manera frecuente.
- Si los menores reciben tratamiento en casa, se deben seguir las instrucciones del médico. Nunca les suministre medicamentos por su cuenta.

- Los niños y adultos que presenten síntomas de tos ferina deben permanecer en el hogar y evitar asistir a jardines infantiles, colegios o lugares de trabajo.
- Lavarse las manos de manera frecuente.

2.6. Modelo de regresión Gamma

Supóngase que se tienen $Y_1, Y_2, \dots, Y_i$ son variables aleatorias que poseen distribución Gamma con media μ y un coeficiente de variación $\phi^{-1/2}$ que se ajustan de la siguiente manera:

$$Y_i | \mu_i, \phi \sim \text{Gamma}(\mu_i, \phi), \quad (1)$$

cuya función de densidad se define como:

$$f(y_i | \mu_i, \phi) = \frac{1}{\Gamma(\phi)} \left(\frac{\phi y_i}{\mu_i} \right)^\phi \exp\left(-\frac{\phi y_i}{\mu_i}\right), \quad y_i > 0. \quad (2)$$

Rescribiendo la ecuación (2) en términos de la familia exponencial, se obtiene:

$$f(y_i | \mu_i, \phi) = \exp \phi \{(-y_i/\mu_i) - \ln(\mu_i)\} - \ln(\Gamma(\phi)) + \phi \ln(\phi y_i) - \ln(y_i) \quad (3)$$

Bajo la ecuación (3), se puede identificar el parámetro natural $\ln(-1/\mu_i)$ y el parámetro de dispersión ϕ , donde en términos del parámetro natural la esperanza y la varianza tenemos:

$$E(y_i | \mu_i, \phi) = b'(\theta_i) = \frac{-1}{\theta} = \mu_i \quad (4)$$

$$\text{Var}(y_i) = V_i = \theta^{-2} \quad (5)$$

Adicional a esto, al asumir medias diferentes se obtiene que el coeficiente de variación es el mismo, por tanto éste es igual a $CV(Y_i) = \theta^{1/2}$.

Dado que, dentro de los modelos lineales generalizados (MLG), (Agresti 2015) plantea la necesidad de tener tres componentes: el predictor lineal, el componente aleatorio y a función de enlace $g(\cdot)$. Para la regresión gamma, la función de enlace considerada fue la siguiente:

$$g(\mu_i) = \log(\mu_i) = \eta_i, \quad (6)$$

donde $\eta_i = X_i^T \beta$ representa el predictor lineal, y $X_i = (x_{i1}, x_{i2}, \dots, x_{ip})^T$ es el i -ésimo vector de p covariables para la i -ésimo valor observado de la variable de interés Y_i , y $\beta = (\beta_1, \beta_2, \dots, \beta_p)^T$ es el vector de parámetros desconocidos que serán estimados. Con respecto a la función de enlace como se observa en la ecuación (6) corresponde a un enlace logarítmico, pero es de aclarar que no es el único enlace que se podría considerar, ya existen otros enlaces como por ejemplo el enlace canónico expresado como $\mu_i = \eta_i^{-1}$.

2.7. Modelo de regresión logística

Supóngase que se tienen $Y_1, Y_2, \dots, Y_i$ son variables aleatorias con distribución bernoulli de media p que se ajustan de la siguiente manera:

$$Y_i | \mu_i \sim \text{Bernoulli}(\mu_i), \quad (7)$$

cuya función de densidad se define como:

$$\exp \left\{ y \log \left(\frac{\theta}{1-\theta} \right) + \log(1-\theta) \right\} \quad (8)$$

En términos de la familia exponencial, se obtiene como resultado:

$$f(Y_i|\mu_i) = \exp Y_i \ln \left(\frac{p}{1-p} \right) + n \ln(1-p) + \ln \binom{n}{Y_i} \quad (9)$$

Bajo la ecuación(9) se puede identificar el parámetro natural $\ln \left(\frac{p}{1-p} \right)$ y el parámetro de dispersión 1, en donde en términos del parámetro natural la esperanza y la varianza tenemos:

$$E(Y_i|\mu_i) = b'(\theta_i) = np = \mu_i \quad (10)$$

$$Var(Y_i) = V_k = np(1-p) \quad (11)$$

Dado que, el planteamiento realizado por (Agresti 2015)) donde se tiene que la función de enlace más utilizada para datos binarios, utilizada el enlace canónico o logit, el cual se define a continuación:

$$\text{logit}(i) = \log \left(\frac{p_i}{1-p_i} \right) \quad (12)$$

es decir,

$$\log \left(\frac{p_i}{1-p_i} \right) = \sum_{j=1}^p \beta_j X_{ij} \quad (13)$$

dado que $X_i = (x_{i1}, x_{i2}, \dots, x_{ip})^T$ es el i-ésimo vector de p covariables para la i-ésimo valor observado de la variable de interés Y_i , y $\beta = (\beta_1, \beta_2, \dots, \beta_p)^T$ es el vector de parámetros desconocidos que serán estimados.

2.8. Estimación de los parámetros del modelo

Teniendo en cuenta función de Score, (Gilberto A. 2013) expresa que el proceso iterativo de la estimación de los β 's se efectúa a través de la utilización del algoritmo de Fisher Scoring, el cual permite la solución del siguiente sistema de ecuaciones:

$$S(\hat{\beta}) = 0, \quad (14)$$

cuya solución está dada por:

$$\beta^{m+1} = (X^T W^{(m)} X)^{-1} X^T W^{(m)} z^m \quad m = 0, 1, \dots, n \quad (15)$$

donde,

$$z = \eta + W^{-1/2} V^{-1/2} (Y_i - \mu_i), \quad (16)$$

siendo que: —

$$\begin{aligned} \eta &= (\eta_1, \dots, \eta_i)^T \\ Y &= (Y_1, \dots, Y_i)^T \\ \mu &= (\mu_1, \dots, \mu_i)^T \\ V &= \text{diag} \mu_1, \dots, \mu_i \\ W &= \text{diag} w_1, \dots, w_i \quad \text{con} \quad w_i = (du_i/d\eta_i)^2 / \mu_i \end{aligned}$$

2.9. Umbrales de clasificación

En el proceso de identificación de los umbrales de seroconversión los modelos de regresión logística son ampliamente utilizados como sistemas de clasificación. El procedimiento empleado consiste en obtener una tabla de clasificación conocida como matriz de confusión ((Carlos 2017)). La mencionada matriz tiene la siguiente estructura:

		CONDICIÓN DE PREDICCIÓN	
		PREDICCIÓN POSITIVA	PREDICCIÓN NEGATIVA
CONDICIÓN VERDADERA	POBLACIÓN TOTAL		
	CONDICIÓN POSITIVA	VERDADERO POSITIVO	FALSO NEGATIVO
	CONDICIÓN NEGATIVA	FALSO POSITIVO	VERDADERO NEGATIVO

Figura 1: Matriz de confusión

Para la selección de los umbrales existen diferentes estadísticas, siendo el Accuracy el más usual ya que obtiene el umbral con mayor número de aciertos de clasificación según lo indica (Carlos 2017). Es de mencionar que los diferentes umbrales se logran observar a través de la curva de ROC, siendo esta la representación gráfica de la sensibilidad y el complemento de la especificidad, cuyos conceptos se definen más adelante.

El área bajo la curva ROC, denominada en la literatura relacionada como AUC, resulta de vital importancia para el presente estudio ya que su valor indica el grado de fiabilidad del clasificador de discernir adecuadamente entre las dos posibles estados (seroconversión y no seroconversión), es decir, indica la probabilidad de aciertos previstos para el modelo predictivo analizado. ((Carlos 2017))

Los valores referenciados por (Carlos 2017) para la interpretación del AUC se muestran a continuación:

- Valores entre 0.5 y 0.6 implican que el modelo ajustado no es el adecuado.
- Valores entre 0.6 y 0.75 implican que el modelo ajustado tiene una tasa de clasificación regular.
- Valores entre 0.75 y 0.9 implican que el modelo ajustado tiene una buena tasa de clasificación.
- Valores entre 0.9 y 0.97 implican que el modelo ajustado tiene una tasa de clasificación muy buena.
- Valores entre 0.97 y 1 implican que el modelo ajustado tiene una excelente tasa de clasificación.

Cabe destacar que este tipo de herramientas gráficas son muy utilizadas para seleccionar modelos de diagnóstico más óptimos, dado que, según (Carlos 2017) la curva ROC es independiente de la distribución de las clases de la población y se relaciona directamente con el costo/beneficio en la toma de decisiones diagnósticas.

2.10. Sensibilidad

Torres (2010) enuncia que la sensibilidad es el parámetro que permite calcular la probabilidad de obtener un verdadero positivo en el diagnóstico en el grupo que se encuentre bajo estudio, en este caso particular los pacientes que presentan seroconversión.

$$\text{sensibilidad} = \frac{VP}{CP} \quad (17)$$

dado que VP es la suma de verdaderos positivos y CP es condición positiva, que corresponde a la suma de verdaderos positivos y los falsos negativos.

2.11. Especificidad

La especificidad es el parámetro que permite determinar la probabilidad de obtener un verdadero negativo en el diagnóstico en el grupo de la condición negativa (Torres,2010), en este caso lo que no seroconvierten.

$$\text{especificidad} = \frac{VN}{CN} \quad (18)$$

donde que VN representan la suma de verdaderos negativos y CN la condición negativa, que corresponde a la suma de los falsos positivos y los verdaderos negativos.

3. Resultados y discusión

En las secciones anteriores se presentó la fundamentación teórica de la regresión gamma, regresión logística, con el fin de construir umbrales de clasificación para la identificación de procesos como seroconversión, ausencia y/o transmisión de anticuerpos en la madre gestante y en el neonato. Por tanto, a continuación, se expondrán los resultados de la aplicación propuesta.

3.1. Análisis descriptivo

Los datos empleados en la presente investigación fueron tomados de sistema de vigilancia epidemiológica de la SDS y está conformado por variables relacionadas con el periodo gestacional como son: número de embarazos, edad gestacional, peso gestacional, número de controles prenatales, entre otras, los datos comprende un periodo desde el año 2015 al 2017 de madres gestantes que participaron en el programa voluntariamente, aclarando que los datos no son públicos y se trabajaron bajo el permiso de dicha entidad. Dentro de la contextualización de los datos, es importante mencionar que también se trabajó la base de Nacidos vivos, de la misma entidad. Partiendo de lo anterior, a continuación, se listan las variables más relevantes para el desarrollo de la presente investigación:

- **Variable interés:** Número de anticuerpos desarrollados por cada una de las unidades de observación (madre gestante y neonato) en los diferentes momentos del estudio.
- **Covariables:** Número de embarazos, tipo de afiliación (contributivo, especial y subsidiado), número de consultas prenatales, peso según edad gestacional, tiempo intergenésico, diferencia día del parto y último día de vacunación contra la DTP, gestación, entre otras. De la base de datos nacidos vivos se seleccionaron variables como: tipo de parto (espontáneo, cesárea o instrumentado) multiplicidad parto, edad quinquenal de la madre y número de hijos vivos.

Con respecto a la variable objetivo se tiene cuatro periodos que se describen a continuación:

- **Periodo 1:** Se midió el número de títulos que presentó la madre gestante en el tercer mes de embarazo antes de recibir la vacuna DTPA. Se debe aclarar que las madres gestantes con títulos muy altos deben ser sacadas del estudio por una de dos posibles razones, fueron vacunadas previamente o ya adquirieron la enfermedad.

- **Periodo 2:** Se midió el número de títulos que presentó la madre gestante en el sexto mes de embarazo después de recibir la vacuna DTPA.
- **Periodo 3:** Se midió el número de títulos que presentó el neonato en el momento del nacimiento, muestra tomada del cordón umbilical.
- **Periodo 4:** Se midió el número de títulos que presentó el neonato a los tres meses de nacido.

Cabe resaltar la base tuvo un proceso de depuración y recodificación, el cual se realizó en el software estadístico **R**. A continuación, se presenta la densidad de la variable de interés en los cuatro periodos de tiempo mencionados se procedió a observar el comportamiento de la misma, para identificar el tipo de distribución como se muestra en la siguiente imagen:

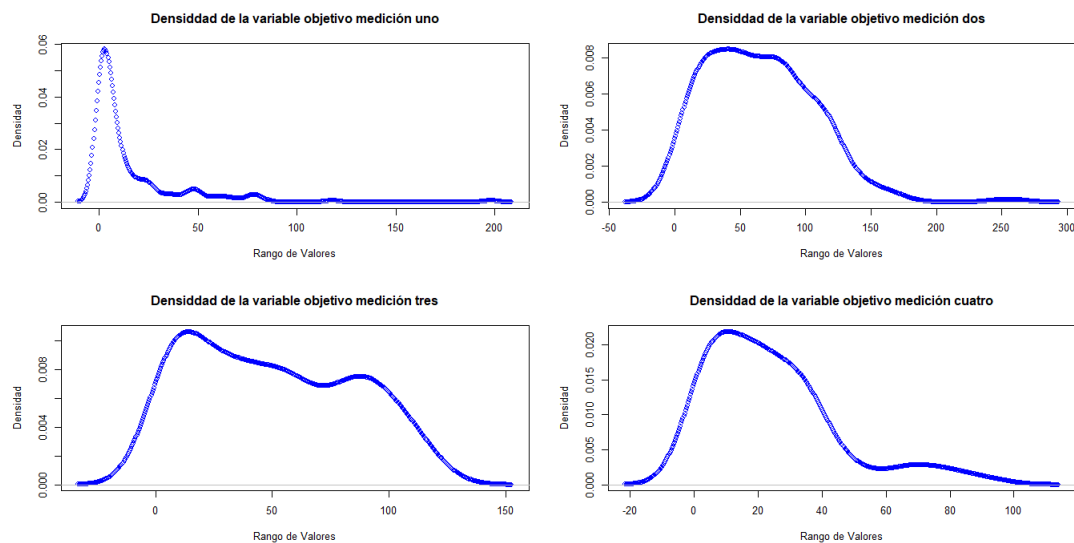


Figura 2: Gráfico de densidad número de títulos

Con base al gráfico 2 se observa que la variable número de títulos para las diferentes muestras consideradas en el estudio, presentando una distribución Gamma. De allí, que dado el contexto de la investigación se considerarán tres variables de interés para la modelación de una regresión gamma, las cuales se lista a continuación:

- **Diferencia entre periodo 2 con respecto al periodo 1:** Diferencia de la cantidad de títulos del periodo 2 con respecto al periodo uno. Se plantea dicha variable con el fin dar respuesta a la pregunta cuántos títulos debe desarrollar la madre gestante para que presente seroconversión.
- **Periodo 3:** Cantidad de títulos que posee el neonato al momento del nacimiento.
- **Diferencia entre periodo 3 y periodo 4:** Diferencia de la cantidad de títulos de las muestras del neonato. Tiene como fin dar respuesta a la pregunta ¿Cuántos títulos el neonato desarrolla durante sus primeros meses de vida?

Bajo el anterior planteamiento, es importante mencionar que, para obtener estabilidad de las variables de interés y los valores atípicos no afecten de manera significativa su comportamiento, se tomaran todos los valores inferiores a 120 títulos, con lo cual el gráfico de densidades se presenta a continuación:

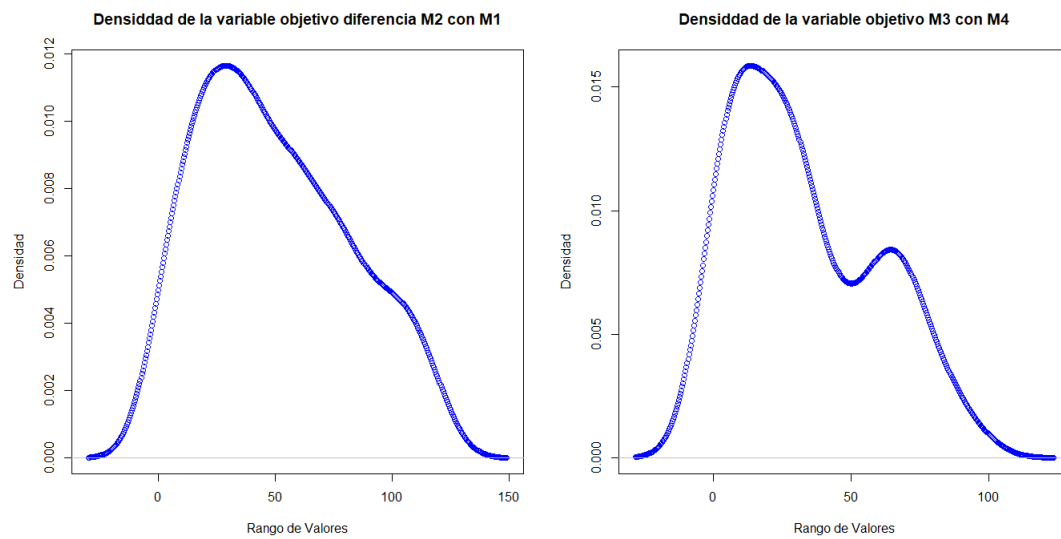


Figura 3: En el gráfico de densidad de la izquierda se evidencia la densidad entre la diferencia del periodo dos con respecto al periodo uno y el de la derecha corresponde a la densidad entre la diferencia del periodo tres con respecto al periodo cuatro.

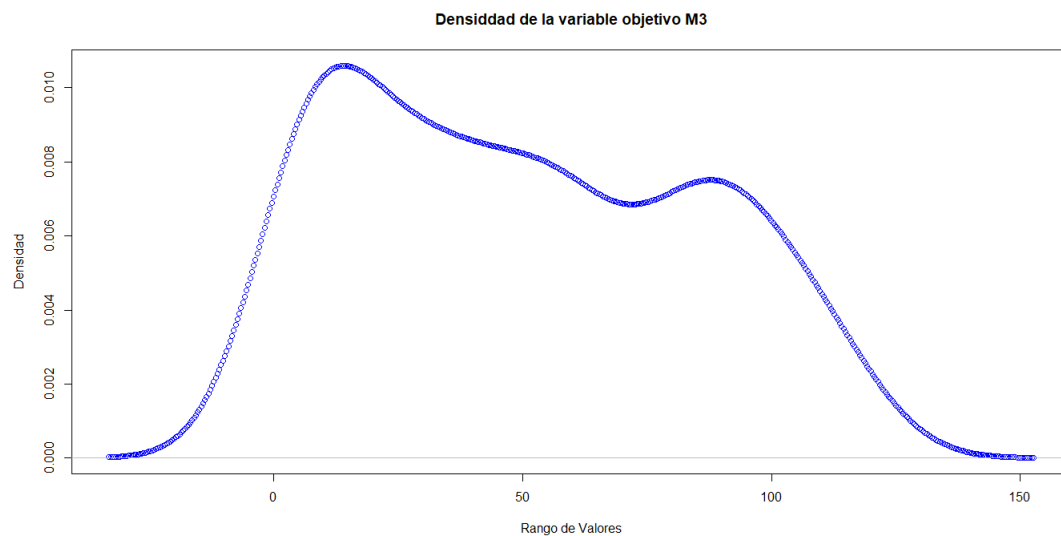


Figura 4: Gráfico de densidad número de títulos del periodo tres

3.2. Resultados de la estimación

Dados los resultados presentados en sección 3.1 se procede al planteamiento de los modelos Gamma, con el fin de identificar las variables significativas que inciden en el comportamiento de las variables de interés mencionadas en el apartado anterior.

3.2.1. Modelo de la diferencia de la cantidad de títulos entre el periodo 2 con respecto al periodo 1

Dado los planteamientos de la investigación en el presente apartado se exponen los resultados de la estimación del modelo gamma donde se podrá evidenciar los factores de riesgo con respecto a la adquisición de la tos ferina

En la tabla 1 se observa los resultados de la estimación de un modelo de regresión Gamma, donde se obtiene como factores de riesgo:

- Número de embarazos
- Tipo de afiliación
- Número de consultas prenatales
- Peso según edad gestacional

Tabla 1: Estimación parámetros modelo Dif. P2 y P1

	Estimate	Std. Error	t value	Pr(> t)
Intercept	4.00787	0.18228	21.987	< 2e-16
n_emb	-0.08244	0.03210	-2.568	0.01101
factor(Tipo_afiliacion)Especial	0.32684	0.19006	1.720	0.08716
factor(Tipo_afiliacion)Subsidiado	0.29157	0.11086	2.630	0.00925
factor(numconsul)6-10	0.11020	0.09608	1.147	0.25284
factor(numconsul)Cero	-0.32109	0.32223	-0.996	0.32033
factor(numconsul)más de 10	-0.74915	0.23724	-3.158	0.00186
factor(Peso_según.EG)Normal para EG	-0.19099	0.13526	-1.412	0.15961

Es de mencionar que la variable peso según edad gestacional, no fue significativa, pero dado el contexto de la investigación se observó la necesidad de tenerla en cuenta. Como parte de la validación del modelo en el figura 5 se observa un buen comportamiento de los residuales, a pesar de la presencia de colas pesadas, ya que el modelo logra capturar el comportamiento de los valores atípicos.

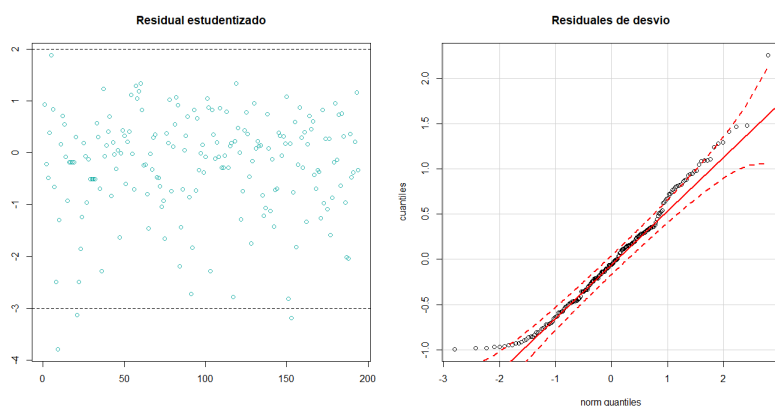


Figura 5: Residuales del modelo Dif. P2 y P1

En la figura 6 se observan los perfiles obtenidos bajo la estimación de este modelo, donde se evidencia que a medida que aumenta la cantidad de embarazos se espera que en promedio el número de títulos empiece

a disminuir. Adicional a la condición anterior, una mujer que tiene de 6 a 10 controles prenatales en su primer embarazo tiende a tener en promedio mayor de número de títulos que una mujer que tiene de 1 a 5 controles y de más de 10 controles prenatales. Al igual que al aumentar la cantidad de embarazos en promedio el número de títulos disminuye.

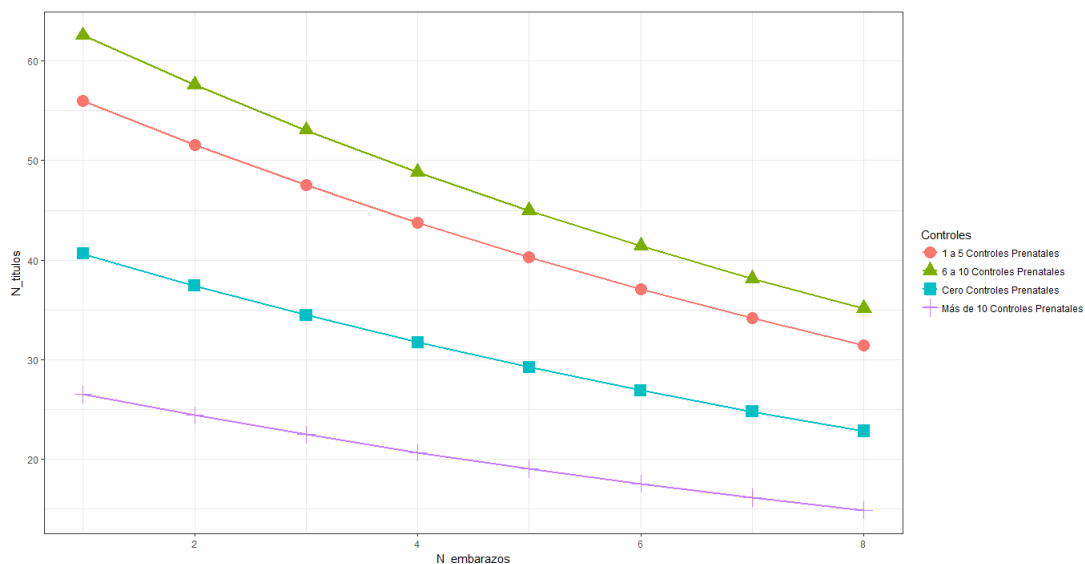


Figura 6: Perfiles modelo Dif. M2 y M1

Es de resaltar que la media de consultas prenatales se encuentra alrededor de 7 a 10 en una gestación normal. Por el contrario, si se realizan menos de cinco visitas los expertos consideran un embarazo no controlado, con lo que se evidencia que, si una mujer posee más de 10 controles, este se considera de alto riesgo, adicional al hecho que presenta en promedio disminución en la cantidad de títulos.

3.2.2. Modelo cantidad de títulos periodo 3

Es de resaltar que para este modelo solo se cuenta con el periodo 3, ya que es la muestra procedente del neonato, por tanto no se hace diferencia con el periodo 1 ni el periodo 2, dado que la unidad observacional cambia.

En la tabla 2 se observa los resultados de la estimación de un modelo de regresión Gamma, donde se obtiene como factores de riesgo:

- Tipo de parto
- Tiempo Inter-genésico
- Número de embarazos
- Diferencia de cantidad de títulos entre P2 y P1
- Diferencia día del parto y último día de vacunación contra la DTP

Para el caso de este modelo, se observa que todas las variables son significativas, y donde en el gráfico 7 se ve el comportamiento de los residuales del modelo, que al igual que el modelo anterior, presenta colas pesadas, pero su modelación logra captar dicho comportamiento, lo cual da luces de un buen ajuste.

Tabla 2: Estimación parámetros modelo de P3

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3.0881454	0.2373088	13.013	< 2e-16
factor(tipo_parto)Espontáneo	0.0823750	0.1145688	0.719	0.473170
factor(tipo_parto)Instrumentado	-1.4720041	0.4328803	-3.400	0.000846
factor(tiemp_inter2)Entre 5 a 6 años entre embarazos	-0.4672809	0.2471286	-1.891	0.060420
factor(tiemp_inter2)Entre 4 a 5 años entre embarazos	-0.3505289	0.2214455	-1.583	0.115380
factor(tiemp_inter2)Entre 2 a 3 años entre embarazos	-0.1434559	0.2082315	-0.689	0.491850
factor(tiemp_inter2)Entre 9 a 10 años entre embarazos	-0.8074128	0.4458305	-1.811	0.071977
factor(tiemp_inter2)Entre 8 a 9 años entre embarazos	-1.5715295	0.4099162	-3.834	0.000180
factor(tiemp_inter2)Entre 6 a 7 años entre embarazos	0.0048979	0.2779519	0.018	0.985962
factor(tiemp_inter2)Entre 7 a 8 años entre embarazos	-0.9982953	0.3334878	-2.993	0.003188
factor(tiemp_inter2)Entre 3 a 4 años entre embarazos	0.0628935	0.2171644	0.290	0.772481
factor(tiemp_inter2)Entre 1 a 2 años entre embarazos	-0.1928121	0.2241903	-0.860	0.391031
factor(tiemp_inter2)Menos de un año entre embarazos	-0.1275937	0.1949977	-0.654	0.513819
n_emb	0.1097841	0.0448424	2.448	0.015416
difuml	0.0042358	0.0008937	4.739	4.64e-06
Dif_tiempo2_inicio_cordon	0.0053785	0.0022513	2.389	0.018032

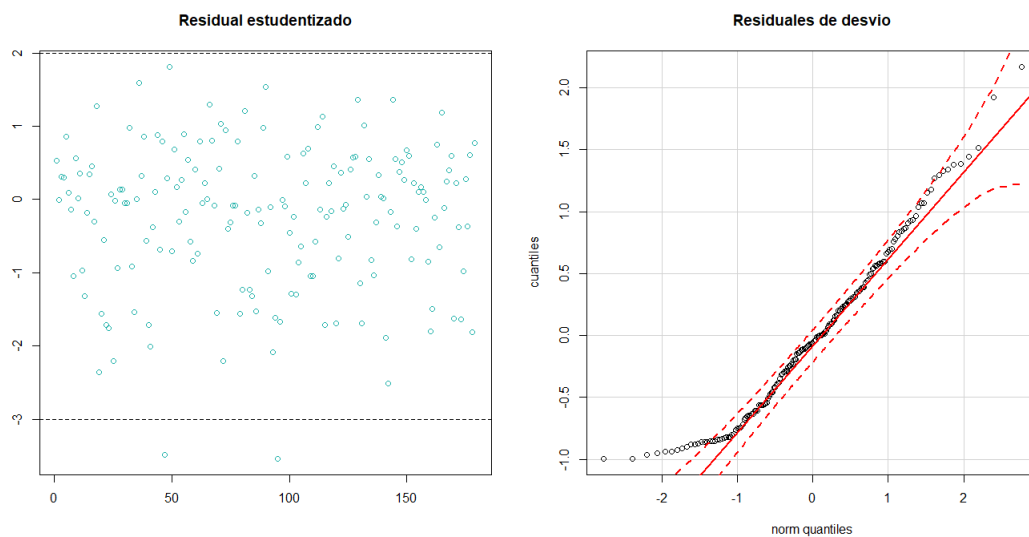


Figura 7: Residuales del modelo de M3

Por otro lado, en el gráfico 8 se visualizan los perfiles obtenidos bajo la estimación del modelo del periodo 3, donde se evidencia en promedio que el número de títulos que trasmite la mamá al neonato son pocos cuando el parto es instrumentado, sin importar la cantidad de embarazos que ha tenido la madre gestante como tampoco el tiempo intergenésico entre embarazos de más de dos años. Mientras que para el parto espontáneo y por cesárea a mayor número de embarazos la madre gestante presente la transmisión de anticuerpos, es decir una madre durante su primer embarazo, dando a luz a los 35 días posteriores de haberse vacunado en promedio trasmite más de 30 títulos al neonato.

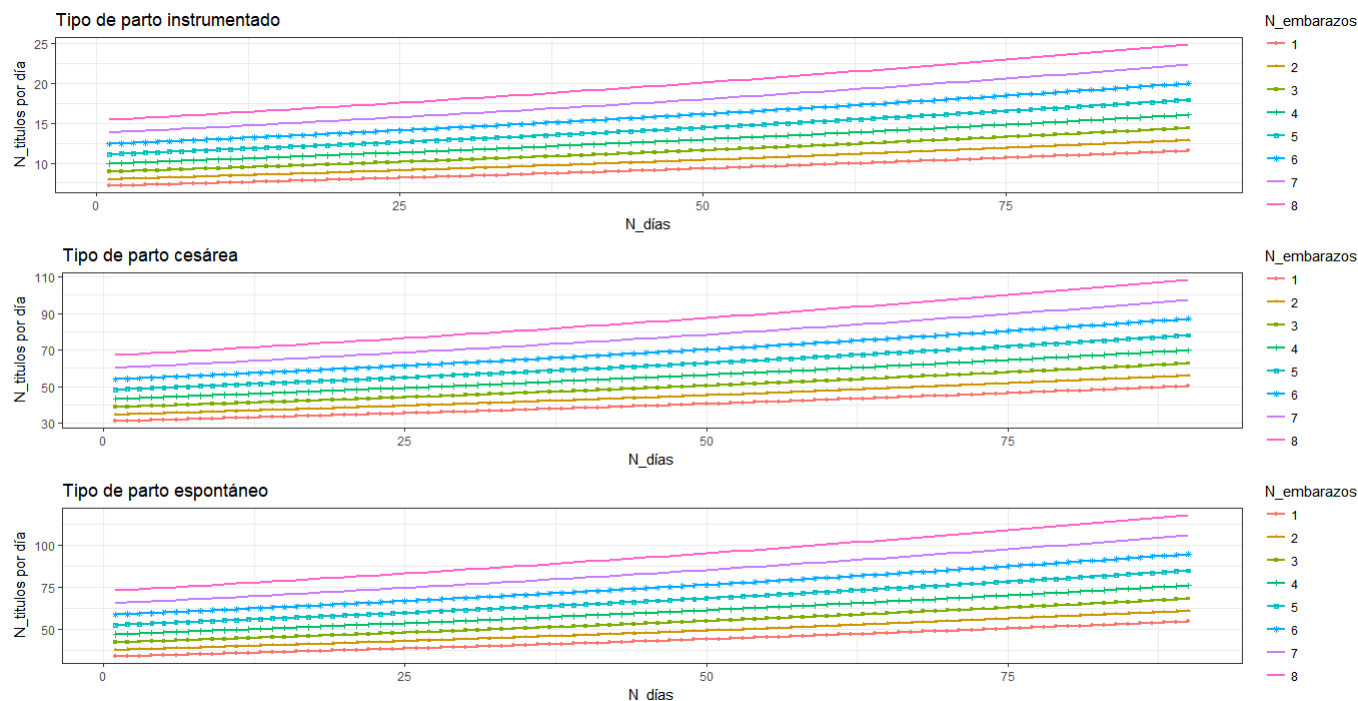


Figura 8: Perfiles modelo de P3

3.2.3. Modelo de la diferencia de la cantidad de títulos entre el periodo 3 y el periodo 4

Siguiendo con la estructura del presente documento se presentan a continuación los resultados de las estimaciones de los parámetros del modelo, haciendo énfasis que estas muestras son directamente del neonato.

En la tabla 3 se observa los resultados de la estimación de un modelo de regresión Gamma, donde se obtiene como factores de riesgo:

- Gestación
- Edad Quinquenal de la madre
- Número de consultas prenatales

Tabla 3: Estimación de los parámetros del modelo Dif. P3 y P4

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3.23347	0.23005	14.056	< 2e-16
factor(gestacion)Pretérmino	-0.88452	0.24908	-3.551	0.000635
factor(quinquel_madre)35 años o más	-0.69921	0.38482	-1.817	0.072828
factor(quinquel_madre)De 20 a 24 años	0.56363	0.26409	2.134	0.035772
factor(quinquel_madre)De 25 a 29 años	0.30964	0.27921	1.109	0.270645
factor(quinquel_madre)De 30 a 34 años	0.56420	0.30391	1.856	0.066931
factor(numconsul)6-10	-0.01495	0.16687	-0.090	0.928842
factor(numconsul)Cero	-3.19826	0.72943	-4.385	3.38e-05
factor(numconsul)más de 10	-0.41095	0.32866	-1.250	0.214673

Es de mencionar que para este caso todas las variables fueron significativas, y como parte de la validación del modelo en la figura 9 se muestra el comportamiento de los residuales del mismo, los cuales presentan buen ajuste a pesar la presencia de datos atípicos que se ven representado en las colas pesadas.

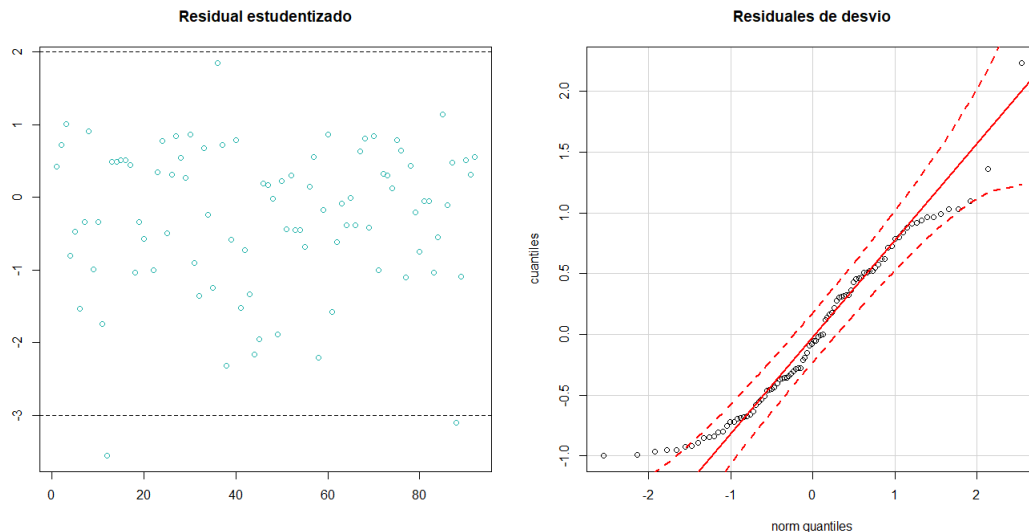


Figura 9: Residuales modelo de Dif. P3 y P4

Partiendo de las estimaciones del modelo, se tiene las siguientes interpretaciones:

- Si una paciente tiene una gestación a término, se encuentra entre la edad quinquenal de 15 a 19 años y ha asistido entre 1 a 5 consultas prenatales en promedio aumenta la cantidad de títulos en 3.2334 unidades, manteniendo las demás variables constantes.
- Si una paciente posee una gestación pretérmino en promedio disminuye la cantidad de títulos en 0.8845 unidades, manteniendo las demás variables constantes.
- Si una paciente se encuentra en la edad quincenal de 35 años o más en promedio disminuye la cantidad de anticuerpos en 0.69 unidades, manteniendo las demás variables constantes.
- Si la madre se encuentra entre la edad de 20 a 24 años en promedio aumenta la cantidad de anticuerpos en 0.5636 unidades, manteniendo las demás variables constantes.
- Si la madre se encuentra entre la edad de 25 a 29 años en promedio aumenta la cantidad de anticuerpos en 0.3096 unidades, manteniendo las demás variables constantes.
- Si la madre se encuentra entre la edad de 30 a 34 años en promedio aumenta la cantidad de títulos en 0.5642 unidades, manteniendo las demás variables constantes.
- Si la madre asiste entre 6 a 10 consultas prenatales en promedio disminuye la cantidad de títulos en 0.014 unidades, manteniendo las demás variables constantes.
- Si la madre no asiste a ninguna consulta prenatal en promedio disminuye la cantidad de anticuerpos en 3.1982 unidades, manteniendo las demás variables constantes.
- Si la madre asiste a más de 10 consultas prenatales en promedio disminuye la cantidad de anticuerpos en 0.4109 unidades, manteniendo las demás variables constantes.

3.3. Resultados umbrales de clasificación

Teniendo en cuenta los resultados propuestos en la sección 3.2, se identifican los factores influyentes en la cantidad de títulos, tanto para las madres como para el neonato. Es así como, a partir de dichos factores, se propone la modelación de una regresión logística, cuya variable respuesta está determinada en si hay seroconversión o ausencia de anticuerpos, entendiendo este proceso como la transmisión de títulos de la madre gestante al neonato.

Dado lo anterior, dicha variable es construida generando cortes, para así asignar las clases en mención, es decir, si es superior a un umbral determinado hay seroconversión y si es menor a dicho umbral se evidencia ausencia de anticuerpos. Bajo dicha premisa, se realizaron diversas simulaciones con diferentes umbrales y se construyó la curva ROC a través de la librería **plotROC** del software **R**, cuyo resultado final de selección del umbral estuvo dado por el mejor AUC.

3.4. Umbrales escenario modelo diferencia de P2 y P1

Según lo expuesto en el aparte anterior, en la figura 10 se observa el comportamiento de las curvas ROC con los diferentes cortes de la cantidad de títulos, empezando de 10 títulos hasta aproximadamente 80 títulos, en donde se logra inferir de manera general que el área bajo la curva de los umbrales inferiores a 30 títulos no es el máximo y no representa el mejor AUC, con lo cual se concluye que en este rango existe ausencia de títulos determinando que no se evidencia seroconversión. Por otro lado, para los umbrales de 38 títulos hasta 53 títulos presentan una buena cobertura bajo la curva, lo que es un indicio para determinar el umbral en donde se presenta seroconversión, es decir, se presenta transmisión de anticuerpos al hablar de la madre.

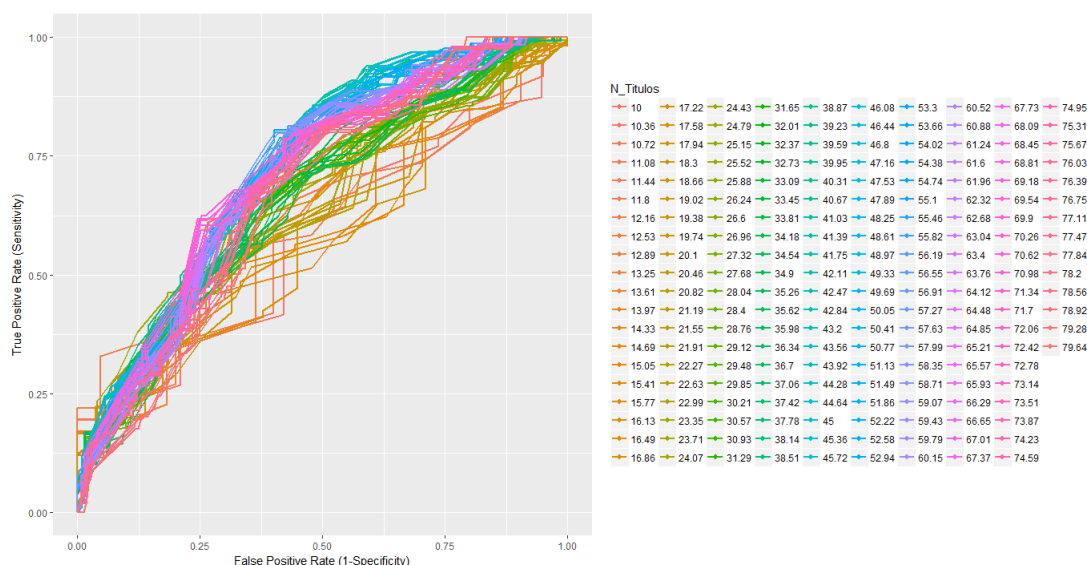


Figura 10: Curva ROC con totalidad de umbrales

De esta manera y, para evaluar de manera más precisa los cortes propuestos en las siguientes figuras, podrán evidenciar el comportamiento de los mismos, pero de forma agrupada.

En la figura 11, se evidencian las curvas para dos grupos de umbrales, los cuales son de 10-20 títulos y de 20-30 títulos, donde se evidencia un comportamiento no estable de la curva en cuanto a su forma, adicional al hecho que el mejor AUC se encuentra en un rango de 0,6475 y 0,6761 respectivamente, correspondientes a un ajuste malo. Por lo cual, se deduce que para este grupo de títulos se presenta ausencia de anticuerpos, por tanto, no hay presencia de seroconversión sino de ausencia de anticuerpos.

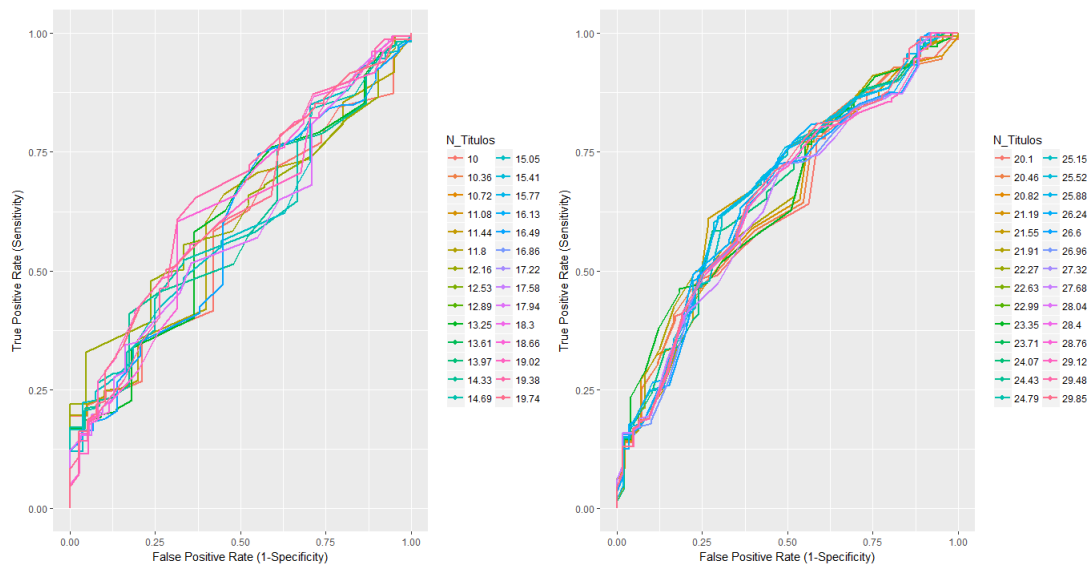


Figura 11: Curva ROC con umbrales de 10-20 títulos y 20-30 títulos

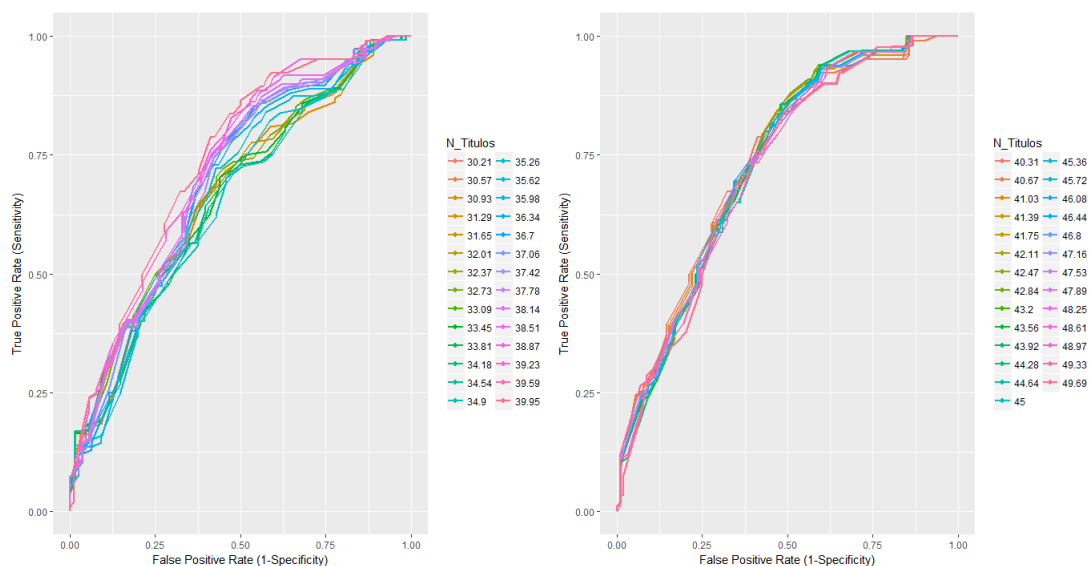


Figura 12: Curva ROC con umbrales de 30-40 títulos y 40-50 títulos

En la figura 12, se observan las curvas para los grupos de umbrales de 30-40 títulos y de 40-50 títulos, donde la curva en forma presenta mejor convexidad. Así mismo, se tiene que el mejor AUC es de 0.7379 y 0.7399, que corresponde en el primer caso regular ajuste y en el segundo caso a un buen ajuste, lo que da luces sobre el umbral a seleccionar; es decir el umbral en este caso sería de 41.03 títulos.

En la figura 13, los grupos de umbrales de 50-60 títulos y de 60-70 títulos presentan buena convexidad, y el mejor AUC en este caso es 0.7266 y 0.7234 respectivamente, lo que indica una disminución en los valores del AUC, por tanto perdiendo el objetivo de maximizar el área sobre la curva. Es así como, en el grupo de 70-80 títulos que se muestra en la figura 14, se tiene una situación similar en cuanto al valor del AUC que no representa la mejor área bajo la curva.

Dadas las anteriores descripciones, en la figura 15 se concluye que el umbral establecido en la cantidad de

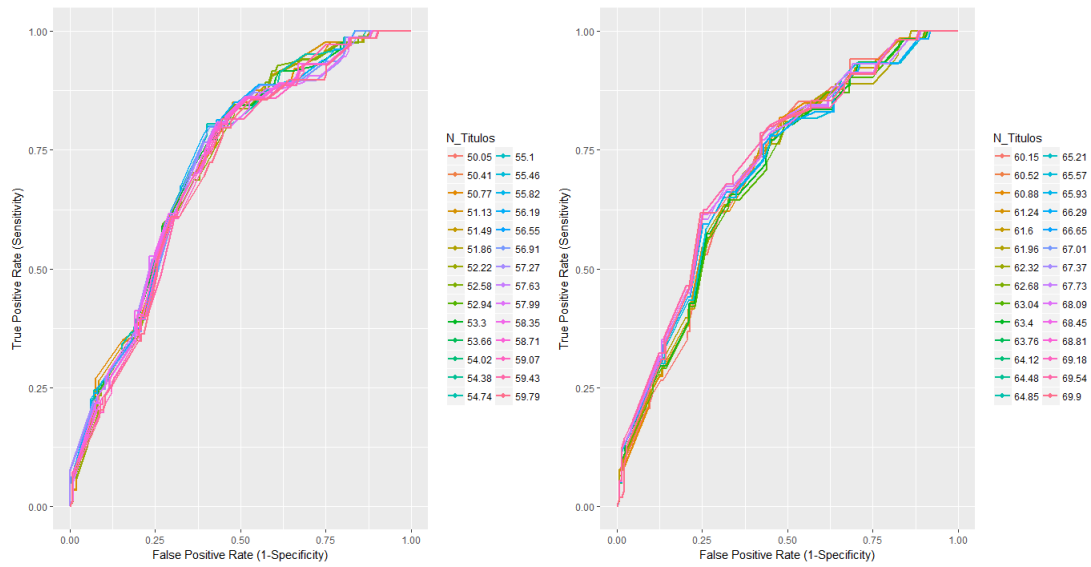


Figura 13: Curva ROC con umbrales de 50-60 títulos y 60-70 título

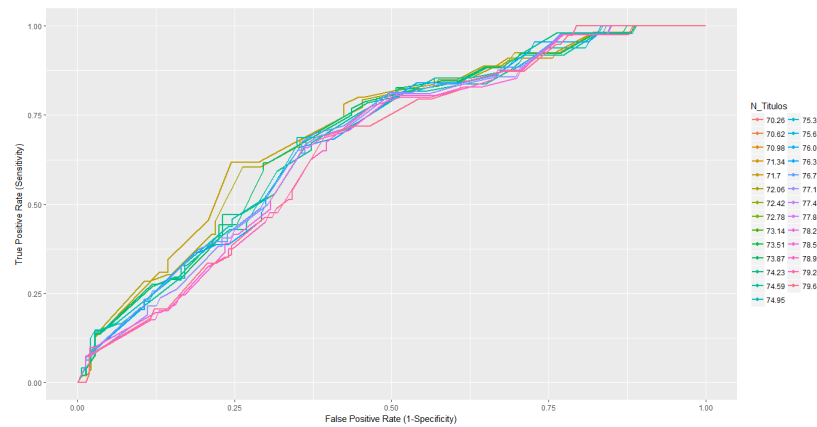


Figura 14: Curva ROC con umbrales de 70-80 títulos

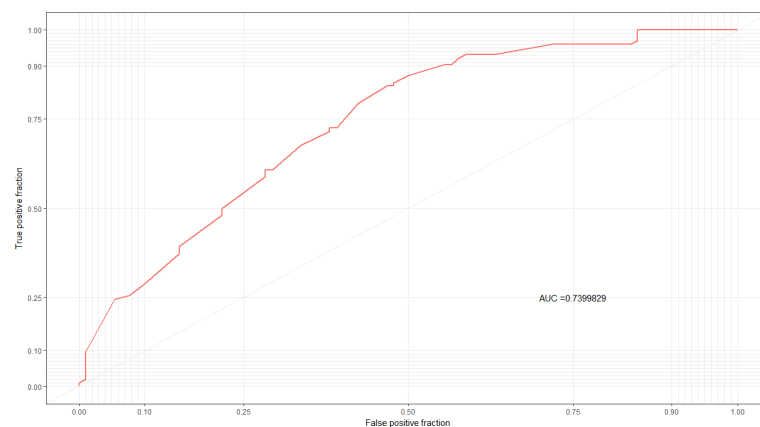


Figura 15: Curva ROC con el mejor umbral

títulos de la madre, para que se evidencie el proceso de seroconversión es de 41.03 títulos, obteniéndose la mejor área sobre la curva de 0.7399 correspondiente a un ajuste regular, destacando que al ser datos reales son buenos resultados.

3.5. Umbrales escenario modelo del P3

Para el caso del modelo que posee las muestras relacionadas con la madre y el neonato a través del cordón umbilical, se observa en la figura 16, que para los diferentes umbrales trabajados, el grupo que mejor presenta convexidad se encuentra entre 73 títulos y 80 títulos, por lo tanto se espera que el AUC de estos umbrales sean bajos, y su comportamiento no posibilite la elección del umbral para madre-neonato. A continuación se presentan las curvas dado el agrupa cimiento por cantidad de títulos.

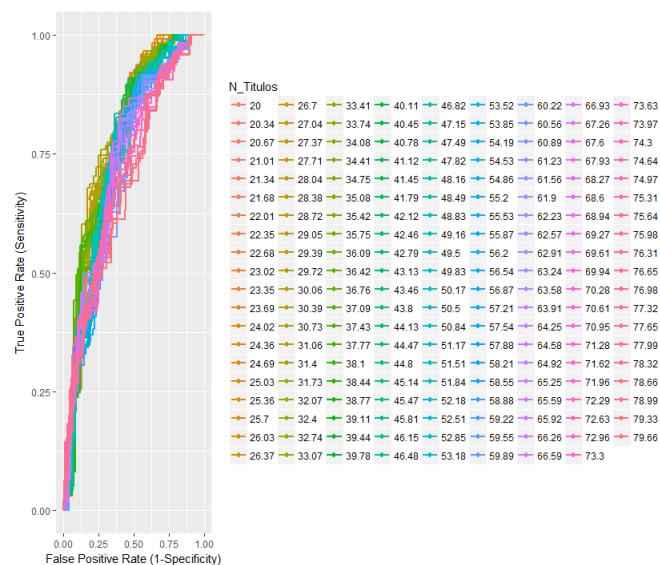


Figura 16: Curva ROC con totalidad de umbrales

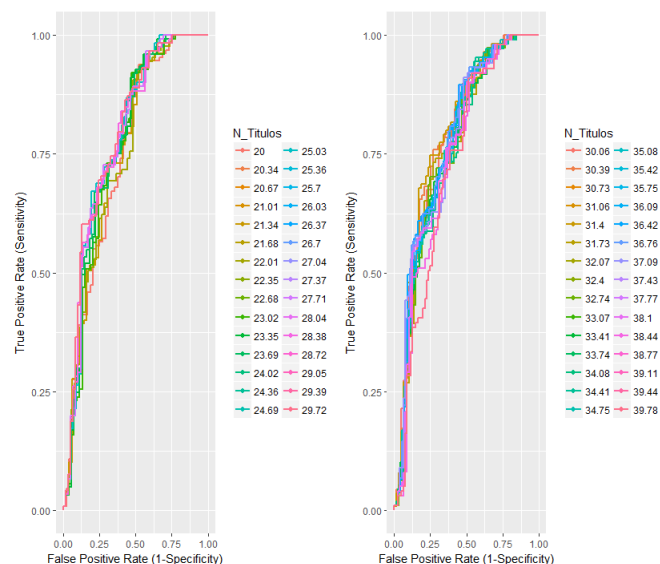


Figura 17: Curva ROC con umbrales de 20-30 títulos y 30-40 títulos

A partir de la figura 17, se observa las curvas generadas para los umbrales de 20-30 títulos y 30-40 títulos, donde el mejor AUC son de 0.7899 y 0.7953 respectivamente. Donde se logra ver que el segundo AUC corresponde al mejor y posee un buen ajuste del modelo, lo cual indica que la sensibilidad del mismo es adecuada para el comportamiento de los datos. Lo anterior, dejar ver que a partir de 30.73 títulos habría un proceso de transmisión de anticuerpos de la madre al neonato.

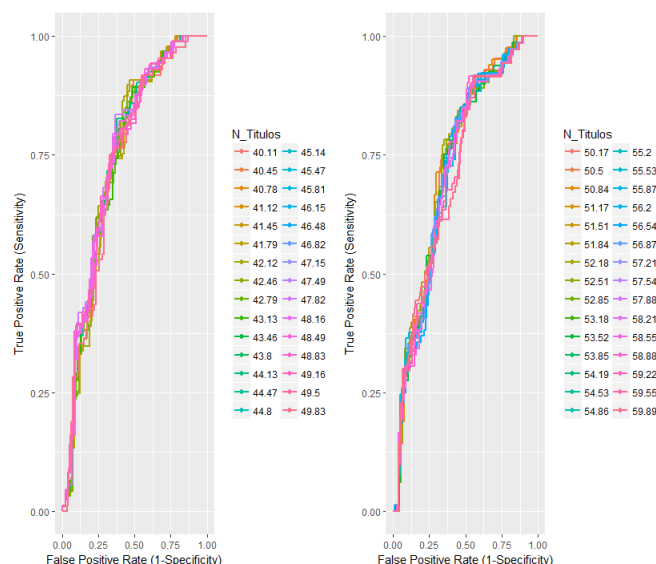


Figura 18: Curva ROC con umbrales de 40-50 títulos y 50-60 títulos

En la figura 18, se encuentran las curvas para los umbrales de 40-50 títulos y de 50-60 títulos, que de manera general presenta buena convexidad, adicional al hecho que el mejor AUC para cada una de las agrupaciones en mención son 0.7517 y 0.7359, lo que deja ver inestabilidad en los modelos dado que dicho estadístico empieza a disminuir, ya que no logra capturar de manera adecuado el comportamiento de los datos. Posterior a esto, se encuentran las curvas para los umbrales de 60-70 títulos y de 70-80 títulos las cuales se pueden visualizar en la figura 19, y cuyos AUC son 0.7307 y 0.7081 respectivamente. Al contrastar dichos resultados, se tiene que no son los que maximizan la sensibilidad del modelo.

Tomando en cuenta las acotaciones anteriores, en la figura 20 se observa que el umbral con el mejor AUC fue el correspondiente a 30.73 títulos, donde se concluye que para que la madre trasmita títulos al neonato, debe ascender 30.73 títulos ya que, de no presentarse esta situación, se debe realizar la vacunación a la mayor brevedad posible contra la tosferina, ya que es vulnerable para su adquisición.

3.6. Umbrales escenario model de diferencia P3 y P4

Con respecto a los umbrales correspondientes al neonato, hay que aclarar de antemano que la cantidad de la muestra disminuyó dado que no todas las madres estaban dispuestas a que se le tomarán las muestras al neonato. De este modo, en la figura 21 se observa claramente que el umbral objetivo se encuentra entre los 68-80 títulos aproximadamente, dado que es donde se generan las curvas con mayor área bajo curva. De allí, que en las siguientes gráficas se ratificará esta premisa siguiendo la estructura de presentación de los anteriores resultados.

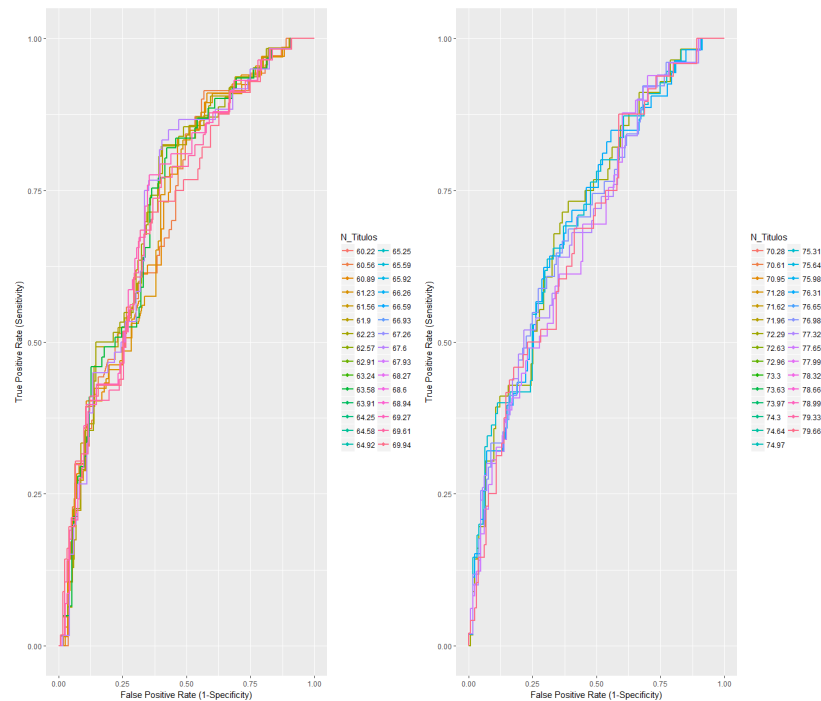


Figura 19: Curva ROC con umbrales de 60-70 títulos y 70-80 títulos

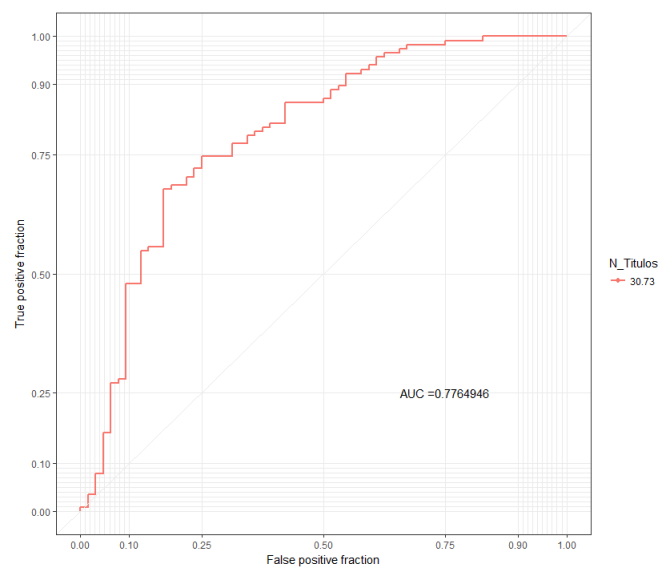
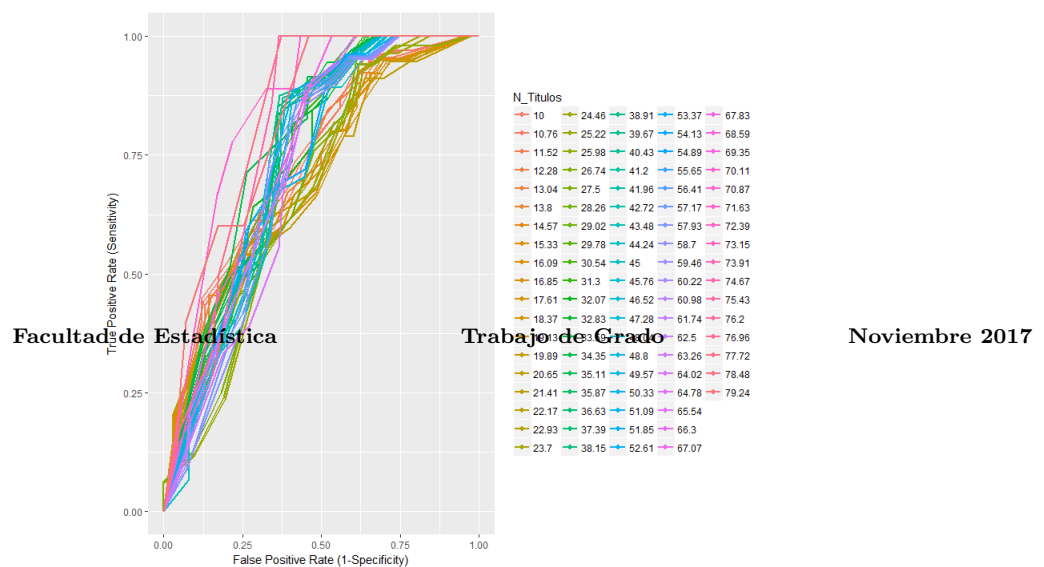


Figura 20: Curva ROC con el mejor umbral



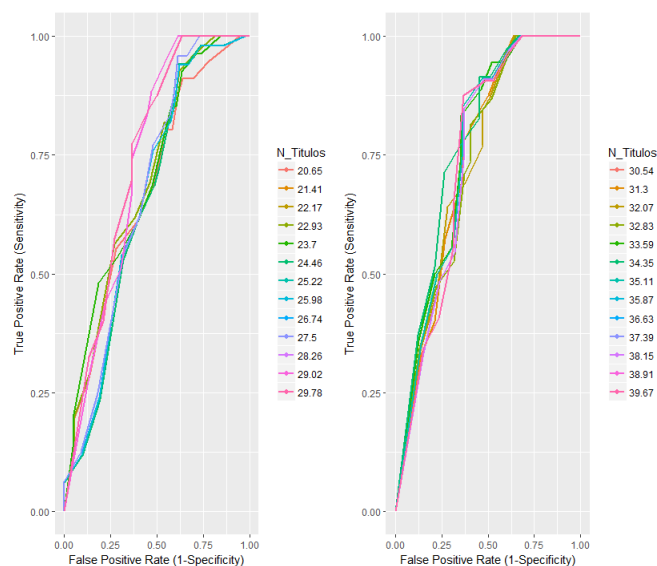


Figura 22: Curva ROC con umbrales de 20-30 títulos y 30-40 títulos

De este modo, se observa que en la figura 22, las curvas generadas para los umbrales entre 20-30 títulos y 30-40 títulos cuyos AUC son 0.7340 y 0.7423, concluyendo con estos estadísticos que los modelos poseen un mal ajuste, por lo tanto, se evidencia ausencia de anticuerpos. En figura 23 se observan las curvas generadas para las agrupaciones de 40-50 anticuerpos y 50-60 anticuerpos, cuyos mejores AUC corresponden a 0.7488 y 0.7488, lo que deja ver que buen ajuste del modelo en la primera agrupación en mención.

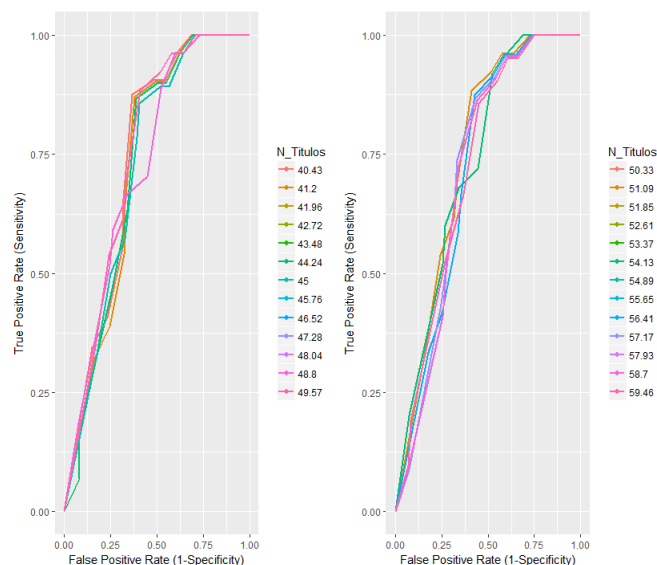


Figura 23: Curva ROC con umbrales de 40-50 títulos y 50-60 títulos

En la figura 24, se logra identificar buen cubrimiento bajo la curva para las agrupaciones de 60-70 títulos y 70-80 títulos, donde se encuentran AUC de 0.8453, lo que indica que el modelo se ajuste en ambas agrupaciones es bueno, y cuyo valor mínimo de títulos es 69.35 para identificar que hay un proceso la generación de defensas, de lo contrario como ya se mencionó anteriormente es necesaria la intervención médica mediante la vacuna de tos ferina, ya que el paciente se encuentra vulnerable y propenso a la

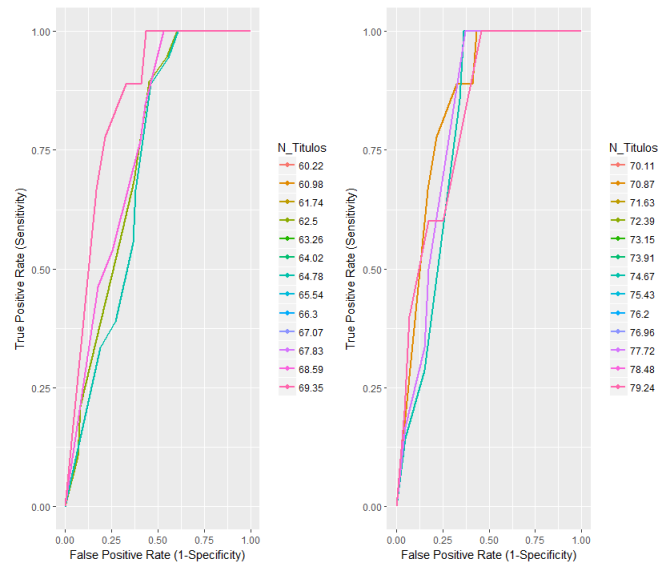


Figura 24: Curva ROC con umbrales de 60-70 títulos y 70-80 títulos

adquisición de dicha enfermedad.

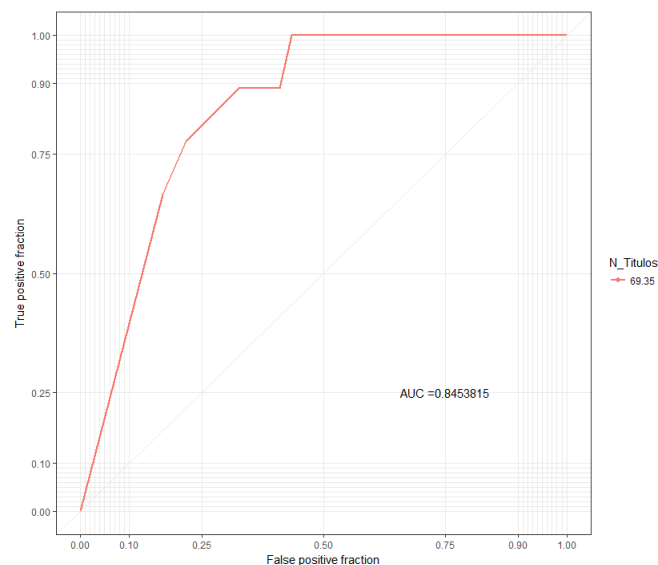


Figura 25: Curva ROC con el mejor umbral

Finalmente, en la figura 26 se encuentran las curvas ROC que poseen mayor AUC, para cada uno de los escenarios, es decir para el control de la madre gestante, relación entre la madre y el neonato, y por último el neonato, con lo que se puede evidenciar en el caso de las muestras de las madres gestantes tener una diferencia de la cantidad de anticuerpos igual o superior a 41.03 anticuerpos para tener un bajo riesgo de adquisición de la enfermedad, mientras que para la muestra donde se relaciona la madre y el neonato desde el cordón umbilical, se debe tener igual o superior a 30.73 títulos. Por último, el umbral perteneciente al neonato debe ser igual o superior a 69.35 títulos con el fin de no contraer la enfermedad.

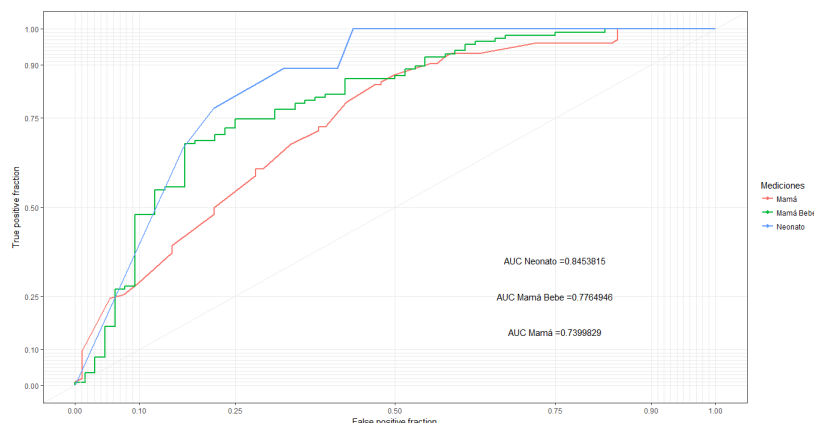


Figura 26: Curva ROC con los mejores umbrales

4. Conclusiones

- Los resultados obtenidos de las estimaciones del modelo gamma, permiten evidenciar que los factores de riesgo influyentes en la producción de anticuerpos de la madre gestante están relacionados con características físicas como el peso gestacional, la edad entre otros, así como características del periodo gestacional. Con lo cual se deja la puerta abierta a la estructuración de un programa de control y seguimiento para las madres gestantes y sus neonatos, con el fin de intervenir de manera temprana dichos pacientes y lograr la prevención de la adquisición de la tos ferina.
- Al determinar los valores de los umbrales en escenarios propios del proceso gestacional para la población colombiana específicamente en la ciudad de Bogotá, permite controlar e identificar a la población más vulnerable en la adquisición de la tos ferina. Resaltando que en los umbrales propios de la mamá y del bebe se evidencia una diferencia en la cantidad de títulos ya que para el primer escenario dicho umbral se encuentra alrededor de 40 y en el escenario del neonato alrededor de 70, esto debido a que los primeros meses de vida el neonato es más vulnerable a cualquier tipo de infección.
- En el proceso de manipulación de la información para la presente investigación se evidenciaron falencias en estándares de calidad de la data, así como una gran limitante en el acceso a dicha información, evidenciándose una brecha en el aspecto investigativo debido a que Colombia no cuenta con un banco de datos del sector salud y que este a su vez sea de fácil acceso para que los investigadores puedan acceder a éste y poder generar aportes en beneficio de la salud.

Referencias

- Agresti, A. (2015), *Foundations of Linear and Generalized Linear Models*, Wiley Series in Probability and Statistics.
- Carlos, B. (2017), Un paquete r para análisis masivo de modelos predictivos de regresión logística multivariante, y sus medidas de discriminación y de clasificación asociadas, Master's thesis, Universitat Oberta de Catalunya.
- de Salud, S. D. (2018), 'Distrito intensifica vacunación y recomendaciones contra la tosferina'.
[*http://bogota.gov.co/temas-de-ciudad/salud/distrito-intensifica-vacunacion-y-recomendaciones-contra-la-tosferina](http://bogota.gov.co/temas-de-ciudad/salud/distrito-intensifica-vacunacion-y-recomendaciones-contra-la-tosferina)

de Salud y Protección Social, M. (2014), *Guía de práctica clínica para la identificación y el manejo clínico de la tos ferina en menores de 18 años de edad. Sistema General de Seguridad Social en Salud.*, 43edn, Ministerio de Salud y Protección Social, Carrera 76, .

D.L Van Esso, A. (2014), ‘Actualización en tos ferina’, *Pediatr Integral* **2**(XVIII), 101–107.

Economía, D. L. (2016), ‘En Colombia hay 4.131 casos probables de tos ferina’.

*<http://diariolaeconomia.com/trabajo-y-salud/item/2541-en-colombia-hay-4-131-casos-probables-de-tos-ferina.html>

Gilberto A., P. (2013), *Modelos de Regresión con apoyo computacional*, Instituto de Matemática e Estadística.

R Core Team (2013), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.

*<http://www.R-project.org/>