

Factores que influyen en los resultados de la prueba saber de undécimo grado presentadas en la Institución Educativa Agrícola la Holanda durante el periodo de 2017 a 2019

John Alexander Bello Vega

Universidad Nacional Abierta y a Distancia UNAD, Escuela de Ciencias de la Educación
ECEDU

Especialización en Educación Superior a Distancia

Asesor: Numar Álvarez Cardona

Granada, 23-03-2021

Tabla de contenido

Introducción	5
Planteamiento del problema	7
Descripción y características del Problema	7
Elaboración de la pregunta	7
Justificación	8
Objetivos	10
Objetivo general	10
Objetivos específicos	10
Marco referencial	11
Diseño metodológico	16
Enfoque	16
Caracterización de las variables	16
Demográficas	17
Académicas	17
Socio-económicas	17
Método	18
Instrumentos	19
Técnica	20
Modelos lineales (ML)	20
Modelos lineales generalizados (GLM)	21
Modelos Aditivos Generalizados (GAM)	22
Modelos Aditivos Generalizados para Localización, Escala y Forma (GAMLSS)	23
Análisis de los resultados	24
Discusión	38
Referencias	43

LISTA DE TABLAS

Tabla 1. Descripción de las variables explicativas	17
Tabla 2. Distribuciones	25
Tabla 3. Componente sistemático (demográfico)	28
Tabla 4. Componente sistemático (socio-económicos).....	30
Tabla 5. Componente sistemático (académico)	34
Tabla 6 Resumen variables demográficas influyentes	38
Tabla 7 Resumen variables socio-económicas influyentes	39
Tabla 8 Resumen variables socioeconómicas influyentes	40

LISTA DE FIGURAS

Prueba grafica lectura critica	26
Prueba de normalidad kolmogorov Smirnov	26
Ajuste de los datos	26
Prueba gráfica puntaje matemáticas.....	27
Prueba de normalidad kolmogorov Smirnov	27
Modelo óptimo componente sistemático (demográfico) lectura critica.....	28
Modelo óptimo componente sistemático (demográfico) matemáticas	29
Componente sistemático (socio-económicos) modelo optimo lectura critica	32
Componente sistemático (socio-económicos) modelo optimo matemáticas	33
Componente sistémico (académico) modelo optimo lectura critica	35
Componente sistémico (académico) modelo optimo matemáticas	36

Introducción

Este proyecto de investigación presenta los factores demográficos, académicos y socioeconómicos que intervienen en los puntajes de la prueba saber de undécimo grado de los estudiantes de la Institución Educativa Agrícola la Holanda, ubicada en la región del Ariari, Granada-Meta. Teniendo como fuente las bases de datos suministradas por el ICFES desde el año 2017 a 2019, la cual cuenta con información específica del estudiante al momento de inscribirse a la prueba, dicha base de datos, contaba inicialmente con cincuenta variables las cuales fueron depuradas, dado que no eran del interés para la investigación, o en su defecto presentaban un porcentaje alto de datos faltantes, quedando con un total de veinticuatro variables, de las cuales veintidós se clasificaron como variables explicativas (V.E) y dos como variables respuesta (V.R). Por otra parte, sólo se analizaron el puntaje en lectura crítica y matemáticas como V.R en función de las V.E, debido a que dichas materias son las únicas tenidas en cuenta para posicionar a la institución a nivel departamental y nacional, además de ser tenidas muy en cuenta a la hora de ingresar a una educación superior. En este orden de ideas se hace necesario determinar algunas de las características más reincidentes que resultan influyentes como son los antecedentes familiares, en particular los ingresos económicos y la educación que tengan o con que cuenten los padres que resultan estar asociadas de manera significativa y positiva con el desempeño académico de los estudiantes, Por otra parte, las características socio-económicas y académicas también están en estrecha relación demostrado tener una asociación con los puntajes, así como lo afirma el Banco Mundial. (2009). El rendimiento académico depende de infinitas variables, es por eso, que pretender buscar un modelo que explique en un cien por ciento cómo es afectado, se hace casi que imposible., aun así, con diferentes técnicas se logra una aproximación considerablemente significativa, en este sentido el presente trabajo busca

solucionar problemáticas inmediatas en donde la pobreza de un pueblo se mida por el nivel educativo que se tenga, por lo tanto, el acceso a la misma es un derecho que conlleva a beneficios u oportunidades que generan progreso prospectivo en la sociedad. Así se busque universalizar las posibilidades de acceso a la educación superior, si no se logra incidir sobre los factores que determinan los resultados del examen de estado, se continuará con la tendencia de la sobre selección de estudiantes con mayores capitales culturales y económicos, y, por tanto, no se lograrán mayores impactos en términos de equidad. (Correa, 2002, p. 135)

Las características familiares sobre el rendimiento académico empiezan a presentar un patrón de incidencia recurrente en cualquier investigación, así como también el plantel educativo resulta generar efectos considerables como lo afirma (Gaviria y Barrientos (2001)). En este sentido, se quiso responder a la siguiente pregunta de investigación: ¿Cuáles son los factores demográficos, socio-económicos y académicas que influyen en los resultados de la prueba saber de undécimo grado presentadas en la Institución Educativa Agrícola la Holanda durante periodo de 2017 a 2019? En el trabajo, se hace un análisis de modelamiento basado en los Modelos Aditivos Generalizados para Localización, Escala y Forma (GAMLSS). Finalmente, se muestran los resultados, discusión, conclusiones del trabajo y bibliografía con la que se realizó el estudio.

Planteamiento del problema

Descripción y características del Problema:

Los malos resultados en las pruebas Saber de undécimo grado es una realidad que viven muchos estudiantes y quizá es la mayor preocupación que aqueja a los padres de familia de la Institución Educativa Agrícola la Holanda ubicada en la región del Ariari-Granada Meta. Presentando una desmotivación en la comunidad educativa por la pérdida de oportunidades para seguir estudiando en Instituciones de Educación Superior (IES). Por lo tanto, se hace indispensable pensar en estrategias que garanticen el derecho a la educación al igual que la calidad educativa a fin de desarrollar las competencias curriculares que les permita vincularse productivamente a la vida universitaria como una forma de mejoramiento de la calidad de vida propia, y del contexto familiar regional.

Elaboración de la pregunta:

¿Cuáles son los factores demográficos, socio-económicos y académicas que influyen en los resultados de las pruebas saber de undécimo grado presentadas en la Institución Educativa Agrícola la Holanda durante periodo de 2017 a 2019?

Justificación.

El rendimiento escolar se constituye en uno de los objetos de estudio fundamentales a la hora de implementar acciones tendientes a analizar la práctica y políticas que regulan el proceso de enseñanza-aprendizaje. Sin embargo, tales análisis deben complementarse con las circunstancias relativas al entorno del estudiante pudiéndose solucionar desde un proyecto de investigación, el cual se establece como. Opción de grado que le permite al estudiante mostrar el resultado de un proceso de gestión del conocimiento a través de metodologías reconocidas por la comunidad académica. Los estudiantes que realizan un proyecto de investigación pueden ser incorporados en calidad de auxiliares a proyectos sistemáticos que se desarrollen en una línea de investigación y, eventualmente, participar en los semilleros de investigación. Igualmente, los trabajos de investigación deberán inscribirse en una de las líneas de investigación institucional o de escuela o de programa. (Hernández, y otros, (2019), p.28) Trabajo que sin lugar a duda generará conocimiento sobre la aplicación del conocimiento estadístico en el campo de la calidad de la educación.

El presente Proyecto de Investigación, pretende analizar los resultados obtenidos en las prueba saber de undécimo grado presentadas por los estudiantes de la Institución Educativa Agrícola la Holanda ubicada en la región del Ariari, con el fin de describir de manera concisa ¿cuáles son los factores demográficos, académicos y socio-económicos que influyen en los resultados de la pruebas Saber de undécimo grado presentadas entre los años 2017 a 2019?, el cual será llevado a cabo a partir del análisis secundario de bases de datos, con el fin de explorar y aprovechar la riqueza de la información que aporta para los docentes, estudiantes y directivos, obtenidas en las bases de datos del ICFES <https://www.icfesinteractivo.gov.co/> en sus evaluaciones semestrales. Estas bases de datos consolidan información que en muchos de los

casos no es aprovechada en procesos de mejora. Por otra parte, El posicionamiento de la institución educativa se ve afectada por los resultados de la prueba saber de undécimo grado, y los resultados que obtienen los estudiantes juegan un papel importante en dicho posicionamiento. Es así como el rendimiento escolar, se constituye en uno de los objetos de estudio fundamentales a la hora de realizar evaluaciones y análisis sobre la práctica y políticas que regulan el proceso de enseñanza-aprendizaje, al integrar un universo de circunstancias y variables, que no necesariamente dependen de los estudiantes y que intentan explicar dicha relación causa-efecto.

Objetivos.

Objetivo general

- Establecer aquellas variables de tipo demográfico, socio-económico y académico que intervienen en los resultados de la prueba saber de undécimo grado de la Institución Educativa Agrícola la Holanda

Objetivos específicos

- Describir los factores demográficos, socio-económicos y académicos según los resultados de la prueba Saber de undécimo grado de la Institución Educativa Agrícola la Holanda desde el año 2017 al año 2019.
- Identificar las variables demográficas, socio-económicas y académicas que influyen en el puntaje de la Prueba Saber de undécimo grado.
- Establecer acciones en las cuales la Institución Educativa Agrícola la Holanda puede intervenir para potenciar los puntajes en la prueba saber de undécimo grado.

Marco referencial

Los principales hallazgos de la OCDE (Organización para la Cooperación y el Desarrollo Económico) y el Banco Mundial, demuestran que los estudiantes están en desventaja respecto a otros países, debido a que están menos preparados, siendo además más jóvenes con estándares educativos realmente bajos a excepción de aquellos que poseen la ventaja de altos ingresos económicos, dicha revisión realizada en 2012, generando inequidad en el acceso y deserción (Mundial, 2012).

Algunas investigaciones resaltan a los padres y la función que cumplen en cuanto a la educación de sus hijos, en este sentido, el nivel educativo de ellos resulta ser determinante, llegando a la conclusión que, factores como el nivel educativo influyen en el rendimiento académico (Manrique & Carreño, 2014) en consecuencia el nivel educacional de los padres se considera un factor de alto impacto, así mismo, generalmente entre mayor sea el grado de estudio de los padres, mayor es el tiempo dedicado a los hijos, estando más atentos a su proceso educacional en cuanto a la calidad de la supervisión al mismo, otros trabajos que abordan este tema son el de (Hakkinen, y Woessmann (2003), llegando a la misma conclusión. En correspondencia a lo anterior, cabe mencionar la investigación que se realizó en 2017 por. Bello (2017) dado el impacto de los resultados en las pruebas saber para cualquier centro educativo generan cambios en los currículos, conllevando a reformularlos, además de las formas de evaluación y cómo la comunidad educativa prepara a los estudiantes para presentar la prueba. Por otra parte, se constituye como un referente de medición para los centros educativos en un ámbito local y nacional, posicionándolos a través de dicha prueba, de modo que conocer

aquellos factores que inciden en los resultados puede contribuir a mejorar la calidad académica y promover nuevos procesos de autoevaluación a fin de contribuir en una sociedad.

Investigaciones en las cuales analizan factores económicos y relacionados con el desempeño académico, teniendo en cuenta la formación y género de cada estudiante, dados los resultados en la prueba Saber Pro 2009, bajo los datos suministrados por el ICFES, demuestran que estos factores inciden en los resultados de la prueba, por lo tanto, se llega a la conclusión que el género y nivel de formación influyen en los resultados de las Pruebas Saber (Ramírez, 2014). Algunas de las investigaciones realizadas en Latinoamérica y las cuales se apoyan en el trabajo realizado por la estudiante de doctorado en Ciencias Sociales Carla Ramírez, se fundamentan en los trabajos de Garbanzo (2007). El cual analiza factores asociados al rendimiento académico en estudiantes de educación superior y su relación con la calidad de la educación pública. Por otra parte, Armenta, Pacheco & Pineda (2008) establecen una de las causas del bajo rendimiento académico de los estudiantes universitarios de la Facultad de Ciencias Humanas de la Universidad Autónoma de Baja California (UABC) llegando a la conclusión que existe una asociación con el rendimiento académico y la calidad de la educación pública. Por otra parte, Hurtado, T & García, G & López, C. (2015) sugieren que las pruebas suministradas por el ICFES son utilizadas en casi todas las instituciones de educación superior como un referente o indicador de selección para lograr ingresar a un determinado programa, en este sentido, el estudiante depende de su resultado ante la prueba de estado, la cual resulta ser de un alto interés para aquellos estudiantes que aspiran a formarse en una universidad, generalmente, para calcular el puntaje total de cada aspirante se adopta generalmente un esquema de ponderación donde los puntajes varían según la carrera, bajo la idea de que el puntaje de ingreso del estudiante está relacionado con factores académicos posteriores del estudiante, haciéndose esencial conocer

aquellos factores que inciden en una posible oportunidad al momento de ingresar a una educación superior. Trabajos de Conger & Long (2010) utilizan variables demográficas y educativas para explicar las diferencias de género en el desempeño y permanencia de los estudiantes universitarios de Texas, además del estudio de Türüt-Asık & Meltem (2007) el cual busca encontrar la existencia de diferencias por género significativas en el desempeño de los estudiantes de pregrado de la universidad pública más grande de Turquía, utilizan tres indicadores, puntaje de entrada a la universidad, desempeño en los cursos preparatorios de inglés y desempeño en el programa que estudian. Otros estudios a nivel internacional, realizados por Barahona, (2014) establecen que los factores asociados al rendimiento académico, no solo es una preocupación que atañe a los centros educativos nacionales, por el contrario, conocer dichos factores también es del interés para universidades como es el caso de Universidad de Atacama (UDA), teniendo como objetivo principal determinar los factores asociados al rendimiento académico, para el año lectivo 2010-2011, en estudiantes de primer año en carreras como Trabajo Social, Ingeniería, Derecho para un total de 258 alumnos, la metodología consistió en dividir el trabajo en dos partes, como primera medida, han utilizado el modelo de Regresión Múltiple con datos de corte transversal para determinar las variables predictores del rendimiento académico, en segunda instancia, se consideró significativo estimar las variables que inciden en la probabilidad de mejorar el rendimiento académico del estudiante haciendo uso del Modelo de Regresión Logística, últimamente el autor concluye bajo un análisis que las variables género, estudia y trabaja, en conformidad con la carrera, prueba verbal y matemática, resultaron ser estadísticamente significativas. Es decir, tendrían un efecto positivo sobre el rendimiento académico del estudiante, así mismo, un estudio llevado cabo por Rose et al. (2005) en la Universidad de Western de Australia se analizaron múltiples factores que podrían influir en el

rendimiento de los estudiantes, en dicha investigación se incluyeron variables tales como la habilidad académica del estudiante, el nivel socioeconómico, las características demográficas (lugar de procedencia), tipo de establecimiento educacional, público o privado, y puntaje prueba de acceso a la universidad. Pág. 27.

Jairo Garzón en su Investigación “Predicción del resultado en la prueba Saber Pro para Economía a partir de la información disponible en el proceso de admisión de la universidad Javeriana”, consideró el registro de acreditación de la universidad como una variable determinante, origen del estudiante y valor de la matrícula entre otras, intentando pronosticar el rendimiento académico en la prueba saber en relación con dichos factores, a fin de brindar información al estudiante en cuanto a la probabilidad de aumentar u obtener mejores resultados a partir del proceso de admisión., la investigación concluyó que los hombres tienen mayor probabilidad de obtener mejores resultados que las mujeres, así como también a menor edad mejores resultados. Algunas de las muchas investigaciones que respaldan este trabajo son las de Valens (2007) a través de un análisis multinivel que encuentra diferencias significativas entre los puntajes medios de las universidades y entre los obtenidos por los estudiantes, que son explicadas por sus características personales y por las características de la institución a la que pertenecen. Montenegro (2005) encontró una correlación positiva entre los puntajes por programa y el valor de la matrícula, el número de estudiantes evaluados y el pertenecer a jornada diurna, además de una correlación negativa si la institución es privada. Otras investigaciones, como las realizadas por Chica & Galvis, & Ramírez (2010) reconocen que, los centros educativos se posicionan según los puntajes obtenidos generalmente por las áreas de matemáticas y lenguaje, en este sentido, los resultados de las pruebas saber de undécimo grado permiten identificar los determinantes del rendimiento académico. Para ello, se utilizó el modelo Logit

Ordenado Generalizado, los resultados obtenidos demuestran que existe relevancia entre las variables socioeconómicas con en el desempeño para ambas áreas, Así también, las variables nivel de ingreso y nivel de escolaridad de los progenitores representan un impacto positivo y significativo en el resultado de las pruebas.

Diseño metodológico

Para este proyecto de investigación se analizarán los resultados obtenidos en la prueba saber de undécimo grado suministradas por las bases de datos del El Instituto Colombiano para el Fomento de la Educación Superior (ICFES) a los estudiantes de la Institución Educativa Agrícola la Holanda, se realizó una depuración de la base de datos para un total de 246 observaciones entre los años 2017 a 2019.

Enfoque

Para el análisis descriptivo se inició la exploración previa de los datos, concretamente con el puntaje obtenido por 246 estudiantes, realizando una selección de aquellos factores que pueden resultar ser influyentes a partir de los siguientes criterios: demográficos, socio-económicos y académicos. En este sentido, se procedió a realizar el análisis inferencial de las variables involucradas indicadas en la (tabla 1). Con enfoque mixto, como lo sugieren Johnson & Onwuegbuzie (2004) “el tipo de estudio donde el investigador mezcla o combina técnicas de investigación, métodos, enfoques, conceptos o lenguaje cuantitativo o cualitativo en un solo estudio” p. 17.

Caracterización de las variables

En estadística, se afirma que una variable es una característica que al ser medida en diferentes individuos es susceptible de adoptar diferentes valores (Camarero, 2013). Para ello hacemos uso del concepto de variable (todo aquello que cambia con el tiempo). Teniendo en cuenta lo anterior, podemos decir que una variable es un conjunto de valores que clasifica a la población objeto de análisis en distintos grupos (a través de distintas categorías clasificatorias);

para ello cada variable debe generarse a partir de un único y coherente criterio clasificatorio, es así que para este problema de investigación se ha determinado la siguiente clasificación.

Demográficas

Para esta categoría corresponden variables cuyas características de la población sean el sexo, edad, lugar de residencia, etc.

Académicas

En esta categoría se clasifican aquellas variables asociadas al rendimiento académico previo y al esfuerzo mostrado por el estudiante. Así, abarcan aquellos factores educativos que rodean directamente la intervención pedagógica o docente, como tiempo dedicado a internet, Máximo nivel educativo alcanzado por los padres entre otras.

Socio-económicas

A esta categoría pertenecen aquellas variables como estrato socioeconómico, forma de pago, estrato, ocupación de los padres entre otras.

A continuación, se exponen las variables que se consideraron en este proyecto de investigación.

Tabla 1. Descripción de las variables explicativas.

DEMOGRAFICAS	SOCIO-ECONOMICAS	ACADEMICAS
1. Género	1. Estrato	1. Tiempo dedicado a la lectura diaria
2. Situación de su hogar actual	2. El hogar cuenta con conexión a internet	2. Tiempo dedicado a internet.
3. Edad	3. El hogar cuenta con Servicio cerrado de televisión	3. Máximo nivel educativo alcanzado por el padre

4. Número de personas que conforman el hogar	4. Tiene computador en su hogar	4. Máximo nivel educativo alcanzado por la madre
	5. Número de horas que trabaja a la semana	5. Cuántos libros físicos o electrónicos hay en su hogar
	6. Número de dormitorios de la residencia	6. Puntaje en matemáticas
	7. Consumo de derivados lácteos	7. Puntaje lectura crítica
	8. Consumo de derivados cárnicos	
	9. Consumo de frutas, verduras y cereales	
	10. Tiene consola de videojuegos en casa.	
	11. Ocupación madre	
	12. Ocupación del padre	

Nota: La descripción de variables se presenta en tres factores: demográficos, socio-económicos y académicos., información tomada de los estudiantes a la hora de inscribirse a la prueba saber de undécimo grado suministrada en los diccionarios del ICFES para los años 2017 a 2019.

Fuente: Elaboración propia.

Método:

El método deductivo que en términos de sus raíces lingüísticas significa conducir o extraer está basado en el razonamiento, si nos remitimos a una definición enciclopédica, entendemos “que obra o procede por deducción” (Enciclopedia Salvat, 1997: 1139), y a su vez, deducción es definido como un “proceso según el cual a partir de ciertos enunciados (premisas)

se derivan otros (conclusiones)” (Enciclopedia Salvat, 1997: 1139). El término método concepto de origen griego que significa el camino hacia, es descrito como un conjunto de actividades que se formulan de acuerdo con una serie de pasos específicos, postulados con el fin de llegar a un resultado determinado (Hurtado León y Toro Garrido, 2007). Una metodología deductiva es aquella en la que se presenta una teoría, se formula una hipótesis de acuerdo con la teoría planteada, se lleva a cabo una observación partiendo de la hipótesis planteada, para finalmente confirmar mediante la observación realizada lo que en la hipótesis se propone. En este caso la deducción intrínseca del ser humano permite pasar de principios generales a hechos particulares. Lo anterior se traduce esencialmente en el análisis de los principios generales de un tema específico: una vez comprobado y verificado que determinado principio es válido, se procede a aplicarlo a contextos particulares (Bernal Torres, 2006).

Instrumentos:

Para este proyecto de investigación se tuvo en cuenta las bases de datos suministradas por el ICFES, obtenidas a través de <https://www.icfesinteractivo.gov.co/>, <https://www.icfes.gov.co/web/guest/investigadores-y-estudiantes-posgrado/acceso-a-bases-de-datos> . Los datos suministrados son de acceso libre en formato texto, posteriormente convertidas a formato Excel para su respectivo análisis, el instrumento cuenta con información referente a los resultados de las evaluaciones semestrales de la prueba saber de undécimo grado e información demográfica, socio-económica y académica al momento de inscribirse a la prueba, clasificada por año, institución, departamento, etc.

Técnica:

La técnica utilizada en este trabajo es a través de material bibliográfico o de campo obtenido por las bases de datos del ICFES, además se aplicará, Modelos Lineales (ML), Modelos Lineales Generalizados (GLM), Modelos aditivos generalizados (GAM), Modelos Aditivos Generalizados para Localización, Escala y Forma (GAMLSS), utilizando el **software R, versión**

4.0.3. R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Modelos lineales (ML)

El modelamiento consiste en la aplicación de una serie de procesos con el objeto de conseguir una explicación apropiada del comportamiento de una variable respuesta (datos) a partir de una o más variables explicativas (modelo), el modelo en general no puede ser perfecto, existe un error o residual. Díaz (como se citó en Bello, 2017, p. 8).

Las etapas del modelamiento estadístico se pueden resumir en:

Especificación: A través del análisis exploratorio de los datos se determinan los supuestos del componente aleatorio. Se establece la función del componente sistemático y se establece como los dos componentes son combinación en el modelo mediante la función enlace.

Selección: Se estiman los parámetros del componente sistemático y se valora la precisión de las estimaciones, calculando la discrepancia entre pares de modelos, lo anterior con el objetivo de seleccionar el modelo óptimo.

Evaluación: Se desea que el modelo represente la realidad lo más aproximada posible.

El más sencillo que podemos crear es un Modelo Lineal (LM), este logra cubrir un amplio aspecto de casos a pesar de sus estrictas restricciones como pueden ser o de hecho de que la variable respuesta debe adaptarse a una distribución normal o bien que la relación entre la

variable, o las variables explicativas, y la variable respuesta deba asumirse lineal. Pérez (como se citó en, Bello, 2017, p 11).

Los modelos lineales son un tipo de modelos de regresión univariante muestra que, entre otras, se dan dos asociaciones. La variable respuesta se considera que sigue una distribución normal y, además que la relación existente entre las variables explicativas y una variable respuesta tenga una forma lineal. Atendiendo la necesidad de que la variable respuesta de un modelo de regresión no siga una distribución normal aparece un Modelo Lineal Generalizado (GLM). Díaz (como se citó en, Bello, 2017).

Modelos lineales generalizados (GLM)

Para ajustar un modelo ML aparece un Modelo Lineal Generalizado (GLM). Este modelo sigue una relación lineal entre la variable respuesta y las covariables del modelo. No obstante, permite que la distribución asumida por la variable respuesta no sea necesariamente una distribución normal, pudiendo ser esta una Poisson o una Binomial, entre otras. Los Modelos GLM abarcan modelos de regresión ordinal con distribución no normal. Cuando se estudian varias mediciones simultáneamente, resulta útil conseguir una descripción parsimoniosa de los datos a través de un modelo matemático que explique, de alguna forma, las observaciones; a esto se le denomina Modelo óptimo, explicando la mayor variabilidad con el mínimo número de parámetros (principio de parsimonia). Tres componentes especifican a un modelo lineal generalizado. Díaz (como se citó en Bello, 2017,p. 9).

- Una variable respuesta
- Un conjunto de variables explicativas
- Una función de enlace

Modelos Aditivos Generalizados (GAM)

Según Pérez (como se citó en Bello, 2017, p. 11) un siguiente paso, antes de la creación de los Modelos Aditivos Generalizados (GAM), fue intentar construir relaciones no lineales entre la variable respuesta y las covariables del modelo utilizado, por ejemplo, los polinomios. Aun así, las suavizaciones de hecho puntualmente sobre los GLM no se aproximaban suficiente y dieron lugar a la de los modelos GAM. Estos últimos permiten la utilización de funciones suavizadoras para describir la relación entre la variable respuesta y las variables explicativas en un sentido no paramétrico. No obstante, y a pesar de las modificaciones de hecho sobre los GAM, todos estos modelos quedan cortos para aquellas distribuciones que tienen más de un parámetro (que son la gran mayoría) porque todos los modelos mencionados hasta ahora solo son capaces de modelar un parámetro de localización directamente a partir de las variables. En algunas ocasiones también se logra modelar un parámetro de escala, pero siempre a través de una relación con el parámetro de localización; por esta razón aparecieron los Modelos Aditivos de Localización, Escala y Forma (GAMLSS). Con la introducción de los GAM quedan flexibilizadas las dos condiciones más restrictivas de los modelos LM que eran la imposibilidad de considerar una familia distinta a la normal y de considerar una relación no lineal entre la variable respuesta y las distintas variables explicativas. Aun así, todo es mejorable y con la idea de introducir nuevas distribuciones adicionales a las de la familia exponencial y de poder modelar un parámetro de escala y los parámetros de forma (dependiendo de la distribución) directamente a partir de las variables explicativas, sin necesidad de utilizar la relación existente entre un parámetro de localización y los restantes parámetros, aparecieron los modelos GAMLSS.

Modelos Aditivos Generalizados para Localización, Escala y Forma (GAMLSS)

Según Pérez (como se citó en Bello, 2017, p. 12) los GAMLSS permiten llevar a cabo diversas mejoras, entre las más importantes están poder abandonar la familia exponencial para escoger la distribución de los datos entre una amplia gama de opciones, incluida algunas que presentan una fuerte asimetría y/o curtosis. Otra de las ventajas desde modelos es la posibilidad de modelar todos los parámetros de distribución directamente a partir de variables explicativas (Pérez, 2010). los GAMLSS pueden considerarse modelos de regresión semi paramétricos, son paramétricos en sentido que necesitan una distribución paramétrica para la variable respuesta y semi en sentido que lo modelado de los parámetros de la distribución en función de las variables explicativas puede implicar el uso de funciones suavizadoras no paramétricas. Según (Pérez 2010). En los modelos GAMLSS el supuesto que la variable respuesta Y pertenece a la familia exponencial y flexibilizada, permitiendo si trabajar con distribuciones con mayor asimetría o curtosis, tanto continuas como discretas. Además, la parte sistemática del modelo se amplía permitiendo modelar no solo la media (parámetro de localización) sino también la varianza (parámetro de escala), la asimetría o la curtosis (parámetro de forma). Incluye relaciones lineales como no lineales entre los parámetros de las distintas variables explicativas.

Análisis de los resultados

A continuación, se describe el comportamiento de las **variables respuesta** (**PUNT_LECTURA_CRITICA, PUNT_MATEMATICAS**) junto con la asignación de los parámetros según sea la distribución, para este trabajo sólo será analizado el parámetro μ interpretado como el valor esperado (promedio) en función de las **variables explicativas** descritas en el **componente sistemático**, el cual se clasificó en demográficas, socio-económicas y académicos., esto con el fin de determinar cuáles son los factores (con un nivel de significancia menor al 10%) que influyen en los resultados de la prueba saber de undécimo grado para la Institución Educativa Agrícola la Holanda, estableciendo los aspectos a mantener o mejorar en la institución, que brinden soporte para potenciar las competencias en los estudiantes., con acciones de mejoramiento en aquellos factores en los cuales pueda intervenir el centro educativo, o en su defecto darlos a conocer a la comunidad quien debe ser protagonista en la contribución de posibles acciones para aquellas variables influyentes que dependen únicamente del núcleo familiar. Posteriormente, se generará una discusión y conclusiones a partir de los resultados encontrados en cada uno de los modelos estadísticos.

Especificación del modelo

La construcción de los modelos se llevará a cabo a través de la especificación del modelo, selección (modelo más parsimonioso), posteriormente se evaluará teniendo en cuenta la bondad de ajuste y finalmente se interpretará los resultados.

La siguiente (tabla 2). Muestra la distribución para cada variable respuesta y sus correspondientes funciones de enlace.

Tabla 2. Distribuciones.

DISTRIBUCIÓN	VARIABLE RESPUESTA	PARÁMETROS			
		μ	δ	u	r
Normal (NO)	Puntaje lectura crítica	μ	δ		
Skew Power	Puntaje Matemáticas	μ	δ		r
Exponential Tipo 2 (SEP2)		μ	δ	u	r

Nota: Dado el comportamiento de las Variables Respuesta (VR). PUNT_LECTURA_CRITICA, PUNT_MATEMATICAS, y teniendo en cuenta la opción realAll se concluye que se distribuyen bajo una Normal (NO), Exponential tipo 2 (SEP2) respetivamente.

Fuente: Elaboración propia.

Para **PUNT_LECTURA_CRITICA**, el mejor ajuste que se presenta en los datos teniendo en cuenta el criterio AIC es la función Normal:

Imagen (1) Ajuste de los datos

```
> Ajustedist <- fitDist(P.español,type="realAll")
> Ajustedist$fit
      NO  SEP2   GA   GG  BCCGo  BCCG  SN2   PE  PE2  SN1  TF  TF2  exGAUS  GIG  GB2  BCPEo  BCPE  BCT
1768.616 1768.790 1769.547 1769.851 1770.034 1770.034 1770.367 1770.591 1770.591 1770.616 1770.616 1770.616 1771.547 1771.852 1772.020 1772.020 1772.034
      BCTo  SEP4  JSU  SHASH  SHASHo  SHASHo2  ST1  ST2  SEP3  SST  ST3  SEP1  GT  ST4  JSUo  ST5  EGB2  LOGNO2
1772.034 1772.052 1772.149 1772.217 1772.273 1772.273 1772.276 1772.276 1772.366 1772.367 1772.367 1772.589 1772.591 1772.615 1772.616 1772.616 1772.618 1773.076
      LOGNO  IG  LO  IGAMMA  WEI  WEI3  WEI2  NET  RG  GU  EXP  PARETO2  PARETO2o  GP
1773.076 1773.083 1774.421 1778.628 1781.361 1781.361 1781.361 1793.491 1794.489 1810.577 2453.626 2455.626 2455.628 2455.629
> |
```

Nota: Representa el ajuste de la variable respuesta PUNT_LECTURA_CRITICA a través del comando fitDist

Gráfica (1) Prueba gráfica.

```
> Ajustefamilia<-histDist(P.español,density=T,xlab="NORMAL",ylab="",main="PUNTAJE LECTURA CRÍTICA",nbins=30)
~
```

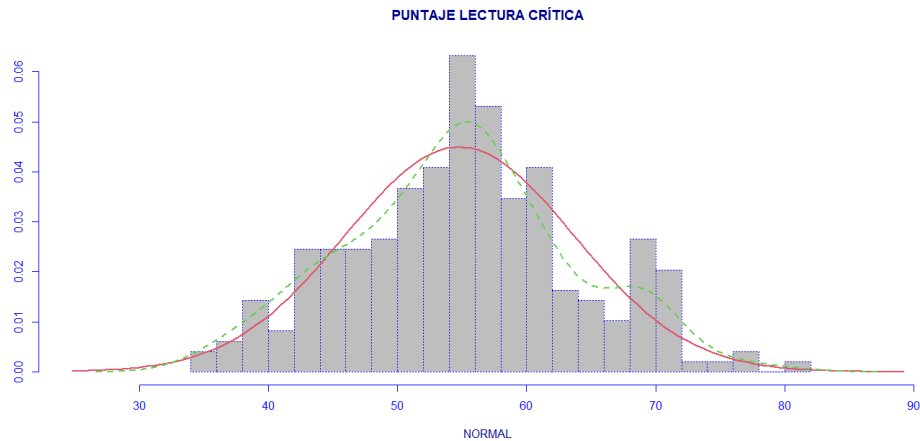


Ilustración 1 prueba grafica lectura critica

Nota: Representa el comportamiento de los datos (línea punteada) y su aproximación a la distribución Normal (línea roja).

Imagen (2) Prueba de normalidad kolmogorov Smirnov (mayor a 50 datos)

```
> lillie.test(P.español)
~
Lilliefors (Kolmogorov-Smirnov) normality test

data: P.español
D = 0.054391, p-value = 0.07604
```

Ilustración 2 Prueba de normalidad kolmogorov Smirnov

Nota: Dado que el p-valor es mayor que $\alpha = 0.05$ no se rechaza la hipótesis nula (H_0 = Se asume normalidad) Luego se concluye que la variable P. español es normal.

En canto a la variable respuesta **PUNT_MATEMATICAS**, el mejor ajuste que se presenta en los datos teniendo en cuenta el criterio AIC es la función SEP2:

Imagen (2) Ajuste de los datos

```
> ajustedist<- fitDist(P.matemáticas,type="realAll")
~
> ajustedist$fit
SEP2 SEP1 SEP3 SHASHo2 SHASHo BCCG BCCGo SHASH LO SEP4 BCPEo BCPE ST3 SST SN2 EGB2 JSU JSUo
1826.666 1826.813 1826.870 1827.159 1827.159 1827.160 1827.160 1827.470 1827.518 1827.649 1827.877 1827.877 1827.983 1827.983 1827.986 1828.218 1828.273 1828.273
ST2 BCTo BCT GB2 ST5 ST1 WEI WEI2 WEI3 PE2 PE ST4 TF TF2 GG NO GT exGAUS
1828.295 1828.347 1828.347 1828.366 1828.393 1828.414 1828.570 1828.570 1828.570 1828.854 1828.854 1829.015 1829.057 1829.057 1829.408 1829.507 1830.812 1831.507
SN1 NET GU GA GIG LOGNO2 LOGNO IG RG IGAMMA EXP GP PARETO2 PARETO2o
1831.507 1840.440 1849.798 1852.984 1854.984 1873.146 1873.146 1878.244 1902.808 1903.492 2435.060 2437.060 2437.061 2437.063
```

Ilustración 3 Ajuste de los datos

Nota: Representa el ajuste de la variable respuesta PUNT_MATEMATICAS a través del comando fitDist.

```
> Ajustefamilia<-histDist(P.matemáticas,density=T,xlab="SEP2",ylab="",main="PUNTAJE MATEMÁTICAS",nbins=30)
```

Gráfica (2) Prueba gráfica

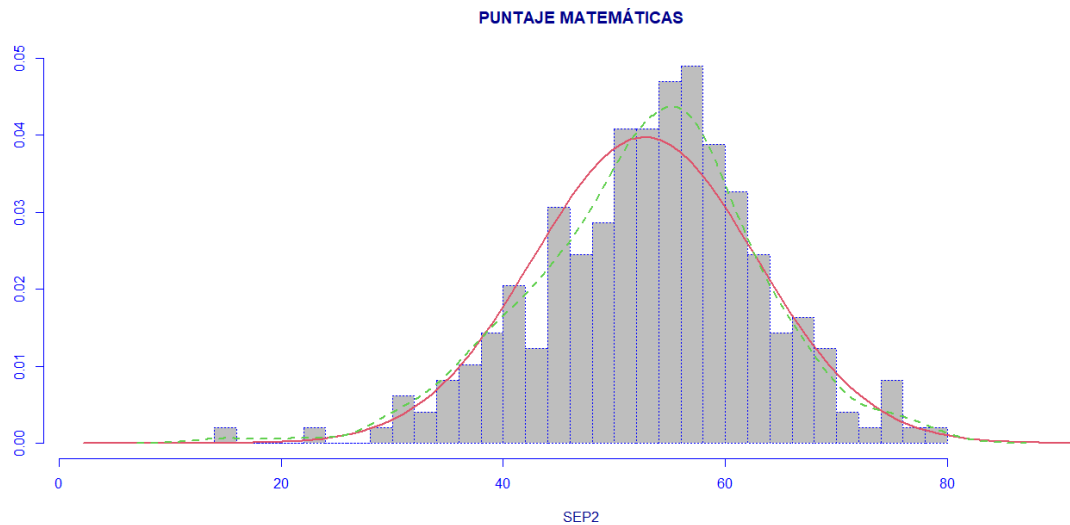


Ilustración 4 Prueba gráfica puntaje matemáticas

Nota: Representa el comportamiento de los datos (línea punteada) y su aproximación a la distribución Exponencial tipo 2 (SEP2) (línea roja).

Imagen (3) Prueba de normalidad kolmogorov Smirnov (mayor a 50 datos)

```
> lillie.test(P.matemáticas)

Lilliefors (Kolmogorov-Smirnov) normality test

data: P.matemáticas
D = 0.063886, p-value = 0.01696
```

Ilustración 5 Prueba de normalidad kolmogorov Smirnov

Nota: Dado que el p-valor es menor que $\alpha = 0.05$ se rechaza la hipótesis nula ($H_0 =$ Se asume normalidad) Luego se concluye que la variable P. matemáticas no es normal.

Componente sistemático

La siguiente (tabla 3). Muestra las variables explicativas a tener en cuenta para todos nuestros modelos Demográficos.

Tabla 3. Componente sistemático (demográfico)

Notación	Nombre
X1	Género
X2	Situación de su hogar actual
X3	Edad
X4	Número de personas que conforman el hogar

Nota: De acuerdo con la depuración que se realizó a la base de datos suministrada por el ICFES se determinan cuatro variables demográficas explicativas en función de cada una de las variables respuesta.

Fuente: Elaboración propia.

Modelo óptimo

A continuación, se presenta el modelo óptimo para **PUNT_LECTURA_CRITICA**, en la selección y mejor ajuste del modelo se utilizó el criterio StepGAICALL.A.

Imagen (4) Modelo óptimo

```
> summary(MODELO)
*****
Family: c("NO", "Normal")

Call: gamlss(formula = PUNT_LECTURA_CRITICA ~ EDAD.COD +
  ESTU_GENERO + FAMI_PERSONASHOGAR.COD + FAMI_SITUACIONECONOMICA,
  family = NO, data = a)

Fitting method: RS()

-----
Mu link function: identity
Mu Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  72.78232   12.64313   5.757 0.0000000682 ***
EDAD.COD     -0.98106    0.64185  -1.529   0.129
ESTU_GENERO[T.M]  0.03375    1.76489   0.019   0.985
FAMI_PERSONASHOGAR.COD[T.3 a 4]  0.24823    3.09503   0.080   0.936
FAMI_PERSONASHOGAR.COD[T.5 a 6] -0.05186    3.10918  -0.017   0.987
FAMI_PERSONASHOGAR.COD[T.7 a 8] -1.70263    4.22596  -0.403   0.688
FAMI_PERSONASHOGAR.COD[T.9 o más] 3.33157    7.17832   0.464   0.643
FAMI_SITUACIONECONOMICA[T.Mejor] -3.19136    1.96939  -1.620   0.108
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

-----
No. of observations in the fit: 128
Degrees of Freedom for the fit: 9
  Residual Deg. of Freedom: 119
                        at cycle: 2

Global Deviance: 933.7602
AIC: 951.7602
SBC: 977.4284
*****
```

Ilustración 6 Modelo óptimo componente sistemático demográfico lectura critica

Interpretación para μ (Resultados)

Nota: Se tiene que las variables demográficas (Edad, género, número de personas que conforman el hogar y situación económica) no explican el puntaje obtenido en lectura crítica. De acuerdo a los resultados se puede afirmar que un estudiante la Institución Educativa la Holanda obtiene en promedio 72.81 puntos o menos en lectura crítica sin tener en cuenta alguna de las variables de estudio.

Modelo óptimo

A continuación, se presenta el modelo óptimo para **PUNT_MATEMATICAS**, en la selección y mejor ajuste del modelo se utilizó el criterio StepGAICALL.A.

Imagen (5) Modelo óptimo

```
> summary(MODELO)
*****
Family:  c("SEP2", "Skew Exponential Power type 2")

Call:  gamlss(formula = PUNT_MATEMATICAS ~ EDAD.COD + ESTU_GENERO +
  FAMI_PERSONASHOGAR.COD + FAMI_SITUACIONECONOMICA,
  family = SEP2, data = a)

Fitting method: RS()

-----
Mu link function:  identity
Mu Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)      67.5078   13.3428   5.059 0.00000158 ***
EDAD.COD         -1.2008    0.6775  -1.772  0.0789 .
ESTU_GENERO[T.M]  4.0988    1.8665   2.196  0.0301 *
FAMI_PERSONASHOGAR.COD[T.3 a 4]  4.0867    3.2708   1.249  0.2140
FAMI_PERSONASHOGAR.COD[T.5 a 6]  5.6277    3.2781   1.717  0.0887 .
FAMI_PERSONASHOGAR.COD[T.7 a 8]  6.9749    4.4828   1.556  0.1224
FAMI_PERSONASHOGAR.COD[T.9 o más]  3.3496    7.5416   0.444  0.6578
FAMI_SITUACIONECONOMICA[T.Mejor] -0.8521    2.0867  -0.408  0.6838
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

-----
No. of observations in the fit: 128
Degrees of Freedom for the fit: 11
  Residual Deg. of Freedom: 117
                        at cycle: 9

Global Deviance:   946.6779
      AIC:         968.6779
      SBC:         1000.05
*****
```

Ilustración 7 Modelo óptimo componente sistemático demográfico matemáticas

Interpretación para μ (Resultados)

Nota: Un estudiante de la Institución Educativa Agrícola la Holanda obtiene en promedio 67.5 puntos en matemáticas sin tener en cuenta alguna de las variables de estudio. Ahora bien, con un nivel de significancia inferior al 10%, aquellos estudiantes que conviven entre 6 a 7 personas obtienen en promedio, 5.62 puntos más respecto a los estudiantes que conviven con 1 y 2 personas. Por otra parte, se puede afirmar, con un nivel de significancia inferior al 5%, que al comparar dos estudiantes donde el segundo difiere del primero en tan sólo su género, se tiene que aquellos que son hombres obtienen 4.9 puntos más que las mujeres. Sin embargo, si comparamos dos estudiantes que tan sólo difieren en que uno es mayor un año que el otro, se obtiene 1.2 puntos menos en matemáticas.

Componente sistemático

La siguiente (tabla 4) muestra las variables explicativas a tener en cuenta para todos nuestros modelos socio-económicos.

Tabla 4. Componente sistemático (socio-económicos)

Notación	Nombre
x1	Estrato socioeconómico de la residencia del estudiante según factura de energía
x2	El hogar cuenta con conexión a internet
x3	El hogar cuenta con Servicio cerrado de televisión
x4	Tiene computador en su hogar
x5	Número de horas que trabaja a la semana
x6	Número de dormitorios de la residencia
x7	¿Cuántas veces por semana se comen los siguientes alimentos en su hogar? Leche o derivados (queso, yogurt, etc.)
x8	¿Cuántas veces por semana se comen los siguientes alimentos en su hogar? Carne (pollo, pavo, res, cordero, cerdo, conejo, etc.), pescados o huevos
x9	¿Cuántas veces por semana se comen los siguientes alimentos en su hogar? Cereales (avena, granola), frutos secos (almendras, maní) o legumbres (frijoles, garbanzos, lentejas)
x10	Tiene consola de videojuegos en casa.
x11	Ocupación madre
x12	Ocupación padre

Nota: De acuerdo con la depuración que se realizó a la base de datos suministrada por el ICFES se determinan doce variables socio-económicas explicativas en función de cada una de las variables respuesta.

Fuente: Elaboración propia.

Modelo óptimo

A continuación, se presenta el modelo óptimo para **PUNT_LECTURA_CRITICA**, en la selección y mejor ajuste del modelo se utilizó el criterio StepGAICALL.A.

Imagen (6) Modelo óptimo

```
> summary(MODELO)
*****
Family: c("NO", "Normal")

Call: gam1ss(formula = PUNT_LECTURA_CRITICA ~ FAMI_TRABAJOLABORPADRE + FAMI_TRABAJOLABORMADRE + ESTU_HORASSEMANATRABAJA + FAMI_TIENECONSOLAVIDEOJUEGOS + FAMI_TIENEINTERNET + FAMI_TIENESERVICIOTV + FAMI_COMELECHEDERIVADOS, family = NO, data = a)

Fitting method: RS()

-----
Mu link function: identity
Mu Coefficients:

(Intercept)
FAMI_TRABAJOLABORPADRE [T.Es dueño de un negocio grande]
FAMI_TRABAJOLABORPADRE [T.Es dueño de un negocio pequeño]
FAMI_TRABAJOLABORPADRE [T.Es operario de máquinas o conduce vehículos]
FAMI_TRABAJOLABORPADRE [T.Es vendedor]
FAMI_TRABAJOLABORPADRE [T.Pensionado]
FAMI_TRABAJOLABORPADRE [T.Tiene un trabajo de tipo auxiliar administrativo]
FAMI_TRABAJOLABORPADRE [T.Trabaja como personal de limpieza]
FAMI_TRABAJOLABORPADRE [T.Trabaja como profesional]
FAMI_TRABAJOLABORPADRE [T.Trabaja en el hogar]
FAMI_TRABAJOLABORPADRE [T.Trabaja por cuenta propia]
FAMI_TRABAJOLABORMADRE [T.Es dueño de un negocio pequeño]
FAMI_TRABAJOLABORMADRE [T.Es vendedor]
FAMI_TRABAJOLABORMADRE [T.Tiene un trabajo de tipo auxiliar administrativo]
FAMI_TRABAJOLABORMADRE [T.Trabaja como personal de limpieza]
FAMI_TRABAJOLABORMADRE [T.Trabaja como profesional]
FAMI_TRABAJOLABORMADRE [T.Trabaja en el hogar]
FAMI_TRABAJOLABORMADRE [T.Trabaja por cuenta propia]
ESTU_HORASSEMANATRABAJA [T.Entre 11 y 20 horas]
ESTU_HORASSEMANATRABAJA [T.Menos de 10 horas]
FAMI_COMECARNEPESCADOHUEVO [T.3 a 5 veces por semana]
FAMI_COMECARNEPESCADOHUEVO [T.Nunca o rara vez comemos eso]
FAMI_COMECARNEPESCADOHUEVO [T.Todos o casi todos los días]
FAMI_COMECEREALFRUTOSLEGUMERE [T.3 a 5 veces por semana]
FAMI_COMECEREALFRUTOSLEGUMERE [T.Nunca o rara vez comemos eso]
FAMI_COMECEREALFRUTOSLEGUMERE [T.Todos o casi todos los días]
FAMI_CUARTOSHOGAR [T.Cuatro]
FAMI_CUARTOSHOGAR [T.Dos]
FAMI_CUARTOSHOGAR [T.Seis o mas]
FAMI_CUARTOSHOGAR [T.Tres]
FAMI_CUARTOSHOGAR [T.Uno]
FAMI ESTRATOVIVIENDA [T.Estrato 2]
FAMI ESTRATOVIVIENDA [T.Estrato 3]
FAMI ESTRATOVIVIENDA [T.Estrato 4]
FAMI ESTRATOVIVIENDA [T.Estrato 5]
FAMI ESTRATOVIVIENDA [T.Estrato 6]
FAMI_TIENECOMPUTADOR [T.Si]
FAMI_TIENECONSOLAVIDEOJUEGOS [T.Si]
FAMI_TIENEINTERNET [T.Si]
FAMI_TIENESERVICIOTV [T.Si]
FAMI_COMELECHEDERIVADOS [T.3 a 5 veces por semana]
FAMI_COMELECHEDERIVADOS [T.Nunca o rara vez comemos eso]
FAMI_COMELECHEDERIVADOS [T.Todos o casi todos los días]
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	46.29736	6.21297	7.452	0.00000000105 ***
FAMI_TRABAJOLABORPADRE [T.Es dueño de un negocio grande]	4.05999	9.86434	0.412	0.68237
FAMI_TRABAJOLABORPADRE [T.Es dueño de un negocio pequeño]	0.25828	4.32644	0.060	0.95263
FAMI_TRABAJOLABORPADRE [T.Es operario de máquinas o conduce vehículos]	1.49054	3.96742	0.376	0.70870
FAMI_TRABAJOLABORPADRE [T.Es vendedor]	10.36028	8.79921	1.177	0.24450
FAMI_TRABAJOLABORPADRE [T.Pensionado]	-1.03359	6.82751	-0.151	0.88027
FAMI_TRABAJOLABORPADRE [T.Tiene un trabajo de tipo auxiliar administrativo]	4.24643	5.89095	0.721	0.47430
FAMI_TRABAJOLABORPADRE [T.Trabaja como personal de limpieza]	0.82177	5.06623	0.162	0.87178
FAMI_TRABAJOLABORPADRE [T.Trabaja como profesional]	7.88529	3.65579	2.157	0.03575 *
FAMI_TRABAJOLABORPADRE [T.Trabaja en el hogar]	13.64439	4.64714	2.936	0.00498 **
FAMI_TRABAJOLABORPADRE [T.Trabaja por cuenta propia]	2.00202	3.29804	0.607	0.54652
FAMI_TRABAJOLABORMADRE [T.Es dueño de un negocio pequeño]	1.24320	4.54120	0.274	0.78538
FAMI_TRABAJOLABORMADRE [T.Es vendedor]	-0.00325	5.34391	-0.001	0.99952
FAMI_TRABAJOLABORMADRE [T.Tiene un trabajo de tipo auxiliar administrativo]	10.13527	6.95402	1.457	0.15112
FAMI_TRABAJOLABORMADRE [T.Trabaja como personal de limpieza]	-0.71191	5.12619	-0.139	0.89009
FAMI_TRABAJOLABORMADRE [T.Trabaja como profesional]	-5.80364	4.87070	-1.192	0.23896
FAMI_TRABAJOLABORMADRE [T.Trabaja en el hogar]	-2.03623	3.64203	-0.559	0.57852
FAMI_TRABAJOLABORMADRE [T.Trabaja por cuenta propia]	-6.05973	7.46667	-0.812	0.42081
ESTU_HORASSEMANATRABAJA [T.Entre 11 y 20 horas]	-8.68090	2.95889	-2.934	0.00501 **
ESTU_HORASSEMANATRABAJA [T.Menos de 10 horas]	1.83775	2.32576	0.790	0.43309
FAMI_COMECARNEPESCADOHUEVO [T.3 a 5 veces por semana]	4.16563	2.47048	1.686	0.09787 .
FAMI_COMECARNEPESCADOHUEVO [T.Nunca o rara vez comemos eso]	-5.55629	6.81128	-0.816	0.41844
FAMI_COMECARNEPESCADOHUEVO [T.Todos o casi todos los días]	-2.68514	3.01557	-0.890	0.37742
FAMI_COMECEREALFRUTOSLEGUMERE [T.3 a 5 veces por semana]	-5.99719	2.61595	-2.293	0.02603 *
FAMI_COMECEREALFRUTOSLEGUMERE [T.Nunca o rara vez comemos eso]	-7.92428	3.27375	-2.421	0.01910 *
FAMI_COMECEREALFRUTOSLEGUMERE [T.Todos o casi todos los días]	1.50364	2.95840	0.508	0.61346
FAMI_CUARTOSHOGAR [T.Cuatro]	4.14295	4.40532	0.940	0.35142
FAMI_CUARTOSHOGAR [T.Dos]	1.13642	4.29894	0.264	0.79258
FAMI_CUARTOSHOGAR [T.Seis o mas]	12.85910	9.33846	1.377	0.17453
FAMI_CUARTOSHOGAR [T.Tres]	5.34273	4.06790	1.313	0.19493
FAMI_CUARTOSHOGAR [T.Uno]	8.03088	6.38310	1.258	0.21407
FAMI ESTRATOVIVIENDA [T.Estrato 2]	7.32221	2.31391	3.164	0.00262 **
FAMI ESTRATOVIVIENDA [T.Estrato 3]	0.20990	4.60634	0.046	0.96383
FAMI ESTRATOVIVIENDA [T.Estrato 4]	-16.25998	10.21351	-1.592	0.11756
FAMI ESTRATOVIVIENDA [T.Estrato 5]	-10.07451	7.44795	-1.353	0.18213
FAMI ESTRATOVIVIENDA [T.Estrato 6]	-7.74893	9.74424	-0.795	0.43016
FAMI_TIENECOMPUTADOR [T.Si]	-0.75470	2.31233	-0.326	0.74547
FAMI_TIENECONSOLAVIDEOJUEGOS [T.Si]	1.85324	3.23577	0.573	0.56934
FAMI_TIENEINTERNET [T.Si]	0.61652	2.70482	0.228	0.82061
FAMI_TIENESERVICIOTV [T.Si]	2.30672	2.11926	1.088	0.28151
FAMI_COMELECHEDERIVADOS [T.3 a 5 veces por semana]	1.80079	2.90441	0.620	0.53801
FAMI_COMELECHEDERIVADOS [T.Nunca o rara vez comemos eso]	1.93702	2.74585	0.705	0.48375
FAMI_COMELECHEDERIVADOS [T.Todos o casi todos los días]	7.06947	3.44249	2.054	0.04516 *

```

-----
No. of observations in the fit: 95
Degrees of Freedom for the fit: 44
  Residual Deg. of Freedom: 51
                        at cycle: 2

```

```

Global Deviance:    648.2109
      AIC:          736.2109
      SBC:          848.5815

```

```

*****

```

Ilustración 8 componente sistemático (SOCIO-económicos) modelo optimo lectura critica

Interpretación para μ (Resultados)

Nota: Un estudiante de la Institución Educativa Agrícola la Holanda obtiene en promedio 46.29 puntos en Lectura crítica sin tener en cuenta alguna de las variables de estudio. Ahora bien, al comparar dos estudiantes donde el segundo difiere del primero en tan sólo la ocupación del padre se tiene que:

Con un nivel de significancia inferior al 5%, aquellos estudiantes que tienen padres con trabajo profesional o que trabajan en el hogar obtienen en promedio, 7.8 y 13.64 puntos más respectivamente, a comparación con estudiantes con padres bajo ocupaciones de agricultor, pesquero o jornalero. Por otra parte, se puede afirmar, con un nivel de significancia inferior al 1%, que al comparar dos estudiantes donde el segundo difiere del primero en tan sólo el estrato, se tiene que aquellos estudiantes de estrato 2 obtienen 7.32 puntos más que aquellos de estrato 1. Sin embargo, si comparamos dos estudiantes que tan sólo difieren en que uno trabaja y el otro no, se obtiene 8.68 puntos menos en lectura crítica. Así mismo, al comparar dos estudiantes con un nivel de significancia inferior al 5%, en donde el segundo difiere del primero en tan sólo el consumo de derivados lácteos, se tiene que aquellos estudiantes que consumen derivados lácteos casi todos los días obtienen 7.06 puntos más en lectura crítica que aquellos que consumen una o dos veces por semana.

Modelo óptimo

A continuación, se presenta el modelo óptimo para **PUNT_MATEMATICAS**, en la selección y mejor ajuste del modelo se utilizó el criterio StepGAICALL.A.

Imagen (7) Modelo óptimo

```
> summary(MODELO)
Warning in summary.gamss(MODELO) :
  summary: vcov has failed, option qr is used instead

*****
Family:  c("SEP2", "Skew Exponential Power type 2")

Call:  gamss(formula = PUNT_MATEMATICAS ~ ESTU_HORASSEMANTRABAJA + FAMI_COMECARNEPESCADOHUEVO + FAMI_CO
      FAMI_COMELECHEDERIVADOS + FAMI_CUARTOSHOGAR + FAMI ESTRATOVIVIENDA + FAMI_TIENECOMPUTADOR + FAMI_TIEN
      FAMI_TIENESERVICIOIV, family = SEP2, data = a)

Fitting method: RS()

-----
Mu link function: identity
Mu Coefficients:

(Intercept)          49.23159    6.70000    7.348  0.000000000154 ***
ESTU_HORASSEMANTRABAJA[T.Entre 11 y 20 horas] -6.89122    3.19084   -2.160  0.03552 *
ESTU_HORASSEMANTRABAJA[T.Menos de 10 horas]  4.46282    2.50808    1.779  0.08114 .
FAMI_COMECARNEPESCADOHUEVO[T.2 a 5 veces por semana]  7.53930    2.66414    2.820  0.00664 **
FAMI_COMECARNEPESCADOHUEVO[T.Nunca o rara vez comemos eso]  8.42139    7.34521    1.147  0.25692
FAMI_COMECARNEPESCADOHUEVO[T.Todos o casi todos los días]  0.17099    3.25196    0.053  0.95827
FAMI_COMECEREALFRITOSLEGUMERE[T.2 a 5 veces por semana] -8.49568    2.82101   -2.990  0.00428 **
FAMI_COMECEREALFRITOSLEGUMERE[T.Nunca o rara vez comemos eso] -5.66526    3.53028   -1.605  0.11473
FAMI_COMECEREALFRITOSLEGUMERE[T.Todos o casi todos los días]  0.09712    3.19021    0.030  0.97583
FAMI_COMELECHEDERIVADOS[T.2 a 5 veces por semana]  3.46705    3.13205    1.107  0.27351
FAMI_COMELECHEDERIVADOS[T.Nunca o rara vez comemos eso] -2.27522    2.96109   -0.768  0.44581
FAMI_COMELECHEDERIVADOS[T.Todos o casi todos los días]  5.03596    3.71234    1.357  0.18090
FAMI_CUARTOSHOGAR[T.Cuatro] -1.75183    4.75065   -0.369  0.71388
FAMI_CUARTOSHOGAR[T.Dos] -5.10622    4.63893   -1.101  0.27587
FAMI_CUARTOSHOGAR[T.Seis o mas]  14.75622   10.07050    1.465  0.14898
FAMI_CUARTOSHOGAR[T.Tres]  1.72428    4.38678    0.393  0.69589
FAMI_CUARTOSHOGAR[T.Uno] -4.85212    6.88346   -0.705  0.48408
FAMI ESTRATOVIVIENDA[T.Estrato 2]  6.67798    2.45530    2.676  0.00999 **
FAMI ESTRATOVIVIENDA[T.Estrato 3] -5.97234    4.96743   -1.202  0.23480
FAMI ESTRATOVIVIENDA[T.Estrato 4] -6.76215   11.01414   -0.614  0.54198
FAMI ESTRATOVIVIENDA[T.Estrato 5] -10.55807    8.02179   -1.315  0.19455
FAMI ESTRATOVIVIENDA[T.Estrato 6] -2.13054   10.50808   -0.203  0.84014
FAMI_TIENECOMPUTADOR[T.Si]  0.21280    2.49259    0.085  0.92242
FAMI_TIENECOMPUTADOR[T.No]  2.66827    3.48543    0.765  0.44759
FAMI_TIENEINTERNET[T.Si] -0.15788    2.51685   -0.054  0.95705
FAMI_TIENEINTERNET[T.No] -1.25324    2.28529   -0.548  0.58580
FAMI TRABAJOLABORPADRE[T.Es dueño de un negocio pequeño]  4.22304    4.89718    0.862  0.39254
FAMI TRABAJOLABORPADRE[T.Es vendedor]  0.66395    5.76283    0.115  0.90873
FAMI TRABAJOLABORPADRE[T.Tiene un trabajo de tipo auxiliar administrativo]  0.64401    7.49914    0.086  0.93190
FAMI TRABAJOLABORPADRE[T.Trabaja como personal de limpieza]  0.29225    5.52803    0.053  0.95805
FAMI TRABAJOLABORPADRE[T.Trabaja como profesional] -4.75390    5.25251   -0.912  0.36570
FAMI TRABAJOLABORPADRE[T.Trabaja en el hogar] -0.86702    3.52783   -0.221  0.82616
FAMI TRABAJOLABORPADRE[T.Trabaja por cuenta propia] -9.32740    8.05195   -1.158  0.25210
FAMI TRABAJOLABORPADRE[T.Es dueño de un negocio grande] -5.53790   10.62761   -0.521  0.60490
FAMI TRABAJOLABORPADRE[T.Es dueño de un negocio pequeño] -5.52838    4.66559   -1.185  0.24154
FAMI TRABAJOLABORPADRE[T.Es operario de máquinas o conduce vehículos]  1.26828    4.27843    0.296  0.76810
FAMI TRABAJOLABORPADRE[T.Es vendedor]  17.85210    9.48899    1.881  0.06564 .
FAMI TRABAJOLABORPADRE[T.Pensionado] -2.28698    7.36272   -0.311  0.75726
FAMI TRABAJOLABORPADRE[T.Tiene un trabajo de tipo auxiliar administrativo]  8.65627    6.35275    1.362  0.17899
FAMI TRABAJOLABORPADRE[T.Trabaja como personal de limpieza]  0.29678    5.46238    0.054  0.95689
FAMI TRABAJOLABORPADRE[T.Trabaja como profesional]  9.92662    3.94227    2.518  0.01499 *
FAMI TRABAJOLABORPADRE[T.Trabaja en el hogar]  8.18552    5.01142    1.632  0.10855
FAMI TRABAJOLABORPADRE[T.Trabaja por cuenta propia]  2.62866    3.55657    0.739  0.46222
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

-----

No. of observations in the fit: 95
Degrees of Freedom for the fit: 44
Residual Deg. of Freedom: 51
                        at cycle: 2

Global Deviance:      662.55
                   AIC:      750.55
                   SEC:      862.9206
*****
```

Ilustración 9 componente sistemático (socio-económicos) modelo optimo matemáticas

Interpretación para μ (Resultados)

Nota: Un estudiante de la Institución Educativa Agrícola la Holanda obtiene en promedio 49.23 puntos en matemáticas sin tener en cuenta alguna de las variables de estudio. Ahora bien, al comparar dos estudiantes donde el segundo difiere del primero en tan sólo la ocupación del padre se tiene que:

Con un nivel de significancia inferior al 1%, aquellos estudiantes que tienen padres con trabajo profesional obtienen en promedio, 9.92 puntos más, respecto a estudiantes con padres bajo ocupaciones de agricultor, pesquero o jornalero. Por otra parte, se puede afirmar, con un nivel de significancia inferior al 1%, que al comparar dos

estudiantes donde el segundo difiere del primero en tan sólo el estrato, se tiene que aquellos estudiantes de estrato 2 obtienen 6.67 puntos más que aquellos de estrato 1. Sin embargo, si comparamos dos estudiantes que tan solo difieren en que uno trabaja y el otro no, se obtiene 6.89 puntos menos en matemáticas. Así mismo, al comparar dos estudiantes con un nivel de significancia inferior al 5%, en donde el segundo difiere del primero en tan sólo el consumo de carene, se tiene que aquellos estudiantes que la consumen de 3 a 5 veces por semana obtienen 7.53 puntos más en matemáticas que aquellos que consumen una o dos veces por semana.

Componente sistemático

La siguiente (tabla 5) muestra las variables explicativas a tener en cuenta para todos nuestros modelos académicos.

Tabla 5. Componente sistemático (académico)

Notación	Nombre
x1	Tiempo dedicado a la lectura diaria
x2	Tiempo dedicado a internet.
x3	Máximo nivel educativo alcanzado por el padre
x4	Máximo nivel educativo alcanzado por la madre
x5	Cuántos libros físicos o electrónicos hay en su hogar excluyendo periódicos, revistas, directorios telefónicos y libros del colegio

Nota: De acuerdo con la depuración que se realizó a la base de datos suministrada por el ICFES se determinan cinco variables académicas explicativas en función de cada una de las variables respuesta.

Fuente: Elaboración propia

Modelo óptimo

A continuación, se presenta el modelo óptimo para **PUNT_LECTURA_CRITICA**, en la selección y mejor ajuste del modelo se utilizó el criterio StepGAICALL.A.

Imagen (8) Modelo óptimo

```
> summary(MODELO)
*****
Family: c("NO", "Normal")

Call: gamlss(formula = PUNT_LECTURA_CRITICA ~ ESTU_DEDICACIONLECTURADIARIA +
  FAMI_NUMLIBROS + ESTU_DEDICACIONINTERNET + FAMI_EDUCACIONMADRE +
  FAMI_EDUCACIONPADRE, family = NO, data = a)

Fitting method: RS()

-----
Mu link function: identity
Mu Coefficients:

```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	54.5239	4.9770	10.955	< 2e-16 ***
ESTU_DEDICACIONLECTURADIARIA[T.Entre 1 y 2 horas]	1.5780	3.2274	0.489	0.62642
ESTU_DEDICACIONLECTURADIARIA[T.Entre 30 y 60 minutos]	0.2291	2.1079	0.109	0.91375
ESTU_DEDICACIONLECTURADIARIA[T.No leo por entretenimiento]	0.8132	3.5111	0.232	0.81754
FAMI_NUMLIBROS[T.11 A 25 LIBROS]	3.2924	2.5202	1.306	0.19576
FAMI_NUMLIBROS[T.26 A 100 LIBROS]	2.9075	3.1280	0.929	0.35587
FAMI_NUMLIBROS[T.MÁS DE 100 LIBROS]	0.5412	5.4067	0.100	0.92056
ESTU_DEDICACIONINTERNET[T.Entre 1 y 3 horas]	4.5623	3.0099	1.516	0.13415
ESTU_DEDICACIONINTERNET[T.Entre 30 y 60 minutos]	2.0098	2.7926	0.720	0.47415
ESTU_DEDICACIONINTERNET[T.No Navega Internet]	1.4790	3.1497	0.470	0.64014
FAMI_EDUCACIONMADRE[T.Educación profesional incompleta]	-4.3875	6.9035	-0.636	0.52717
FAMI_EDUCACIONMADRE[T.Ninguno]	-9.1044	10.7575	-0.846	0.40029
FAMI_EDUCACIONMADRE[T.Primaria completa]	-2.1969	5.2894	-0.415	0.67918
FAMI_EDUCACIONMADRE[T.Primaria incompleta]	-1.6082	5.8283	-0.276	0.78343
FAMI_EDUCACIONMADRE[T.Secundaria (Bachillerato) completa]	-1.8237	5.1687	-0.353	0.72529
FAMI_EDUCACIONMADRE[T.Secundaria (Bachillerato) incompleta]	-3.4062	5.3052	-0.642	0.52296
FAMI_EDUCACIONMADRE[T.Técnica o tecnológica completa]	5.6604	5.3527	1.057	0.29398
FAMI_EDUCACIONMADRE[T.Técnica o tecnológica incompleta]	-0.2851	8.9447	-0.032	0.97467
FAMI_EDUCACIONPADRE[T.Ninguno]	-4.6466	6.8841	-0.675	0.50194
FAMI_EDUCACIONPADRE[T.Primaria completa]	-3.9443	6.0914	-0.648	0.51945
FAMI_EDUCACIONPADRE[T.Primaria incompleta]	-6.4293	5.5671	-1.155	0.25213
FAMI_EDUCACIONPADRE[T.Secundaria (Bachillerato) completa]	-1.7955	5.0539	-0.355	0.72347
FAMI_EDUCACIONPADRE[T.Secundaria (Bachillerato) incompleta]	-3.7767	5.7887	-0.652	0.51630
FAMI_EDUCACIONPADRE[T.Técnica o tecnológica completa]	-22.2994	8.0881	-2.757	0.00746 **
FAMI_EDUCACIONPADRE[T.Técnica o tecnológica incompleta]	3.0168	11.2451	0.268	0.78929

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
-----
No. of observations in the fit: 95
Degrees of Freedom for the fit: 26
  Residual Deg. of Freedom: 69
                        at cycle: 2

Global Deviance:      675.5618
      AIC:             727.5618
      SBC:             793.9626
*****
```

Ilustración 10 componente sistémico (académico) modelo óptimo lectura crítica

Interpretación para μ (Resultados)

Nota: Un estudiante de la Institución Educativa Agrícola la Holanda obtiene en promedio 54.52 puntos en Lectura crítica sin tener en cuenta alguna de las variables de estudio. Ahora bien, al comparar dos estudiantes donde el segundo difiere del primero en tan sólo el nivel educativo del padre se tiene que:

Con un nivel de significancia inferior al 1%, que aquellos que tienen madres con nivel educativo Técnico u tecnológico obtienen en promedio, 22.29 puntos menos respecto a los que tienen padres con educación profesional completa.

Modelo óptimo

A continuación, se presenta el modelo óptimo para **PUNT_MATEMATICAS**, en la selección y mejor ajuste del modelo se utilizó el criterio StepGAICALL.A.

Imagen (9) Modelo óptimo

```
> summary(MODELO)
Warning in summary.gamlss(MODELO) :
  summary: vcov has failed, option qr is used instead
*****
Family: c("SEF2", "Skew Exponential Power type 2")
Call: gamlss(formula = PUNT_MATEMATICAS ~ ESTU_DEDICACIONLECTURADIARIA + FAMI_NUMLIBROS + ESTU_DEDICACIONINTERNET + FAMI_EDUCACIONMADRE + FAMI_EDUCACIONPADRE, family = SEF2, data = a)
Fitting method: RS()
-----

Mu link function: identity
Mu Coefficients:

```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	50.7703	5.6166	9.039	2.53e-13 ***
ESTU_DEDICACIONLECTURADIARIA[T.Entre 1 y 2 horas]	-5.3123	3.6421	-1.459	0.1492
ESTU_DEDICACIONLECTURADIARIA[T.Entre 30 y 60 minutos]	-2.3959	2.3787	-1.007	0.3174
ESTU_DEDICACIONLECTURADIARIA[T.No leo por entretenimiento]	-8.8867	3.9623	-2.243	0.0281 *
FAMI_NUMLIBROS[T.11 A 25 LIBROS]	5.9083	2.8440	2.077	0.0415 *
FAMI_NUMLIBROS[T.26 A 100 LIBROS]	6.3952	3.5300	1.812	0.0744 .
FAMI_NUMLIBROS[T.MÁS DE 100 LIBROS]	4.5506	6.1015	0.746	0.4583
ESTU_DEDICACIONINTERNET[T.Entre 1 y 3 horas]	4.5498	3.3967	1.339	0.1848
ESTU_DEDICACIONINTERNET[T.Entre 30 y 60 minutos]	1.6071	3.1515	0.510	0.6117
ESTU_DEDICACIONINTERNET[T.No Navega Internet]	1.9023	3.5545	0.535	0.5942
FAMI_EDUCACIONMADRE[T.Educación profesional incompleta]	-7.4841	7.7906	-0.961	0.3401
FAMI_EDUCACIONMADRE[T.Ninguno]	-3.9329	12.1397	-0.324	0.7469
FAMI_EDUCACIONMADRE[T.Primaria completa]	-0.6561	5.9690	-0.110	0.9128
FAMI_EDUCACIONMADRE[T.Primaria incompleta]	0.4403	6.5772	0.067	0.9468
FAMI_EDUCACIONMADRE[T.Secundaria (Bachillerato) completa]	-3.0678	5.8328	-0.526	0.6006
FAMI_EDUCACIONMADRE[T.Secundaria (Bachillerato) incompleta]	-4.0272	5.9868	-0.673	0.5034
FAMI_EDUCACIONMADRE[T.Técnica o tecnológica completa]	1.1743	6.0405	0.194	0.8464
FAMI_EDUCACIONMADRE[T.Técnica o tecnológica incompleta]	-0.2065	10.0940	-0.020	0.9837
FAMI_EDUCACIONPADRE[T.Ninguno]	-4.8187	7.7686	-0.620	0.5371
FAMI_EDUCACIONPADRE[T.Primaria completa]	0.4831	6.8741	0.070	0.9442
FAMI_EDUCACIONPADRE[T.Primaria incompleta]	-4.4445	6.2825	-0.707	0.4817
FAMI_EDUCACIONPADRE[T.Secundaria (Bachillerato) completa]	-1.9176	5.7032	-0.336	0.7377
FAMI_EDUCACIONPADRE[T.Secundaria (Bachillerato) incompleta]	1.1690	6.5325	0.179	0.8585
FAMI_EDUCACIONPADRE[T.Técnica o tecnológica completa]	-16.5053	9.1273	-1.808	0.0749 .
FAMI_EDUCACIONPADRE[T.Técnica o tecnológica incompleta]	1.2391	12.6900	0.098	0.9225

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
-----

No. of observations in the fit: 95
Degrees of Freedom for the fit: 26
  Residual Deg. of Freedom: 69
                        at cycle: 2

Global Deviance:    698.53
                   AIC:    750.53
                   SBC:    816.9308
*****
```

Ilustración 11 componente sistémico (académico) modelo optimo matemáticas

Interpretación para μ (Resultados)

Nota: Un estudiante de la Institución Educativa Agrícola la Holanda obtiene en promedio 50.77 puntos en matemáticas sin tener en cuenta alguna de las variables de estudio. Ahora bien, al comparar dos estudiantes donde el segundo difiere del primero en tan sólo el nivel educativo del padre se tiene que: Con un nivel de significancia inferior al 10%, que aquellos que tienen padres con nivel educativo Técnico u tecnológica obtienen en promedio, 16.50 puntos menos respecto a los que tienen padres con educación profesional completa. Por otra parte, se puede

afirmar, con un nivel de significancia inferior al 5%, que al comparar dos estudiantes donde el segundo difiere del primero en tan sólo el número de libros que hay en la familia, se tiene que aquellos estudiantes con un número entre 11 y 25 textos obtienen 5.9 puntos más que aquellos que sólo poseen de 0 a 10 libros.

Discusión

El mejor ajuste que se presenta para puntaje en lectura crítica y matemáticas teniendo en cuenta el criterio AIC es la distribución Normal (NO) y Skew power exponential tipo 2 (SEP2) con un AIC de 1768,61 Y 1826,66 respectivamente.

La Tabla 6, 7 y 8 muestra los factores que influyen aun nivel de significancia del 1%, 5% o 10% para cada componente, siendo PLC, puntaje lectura crítica y PMAT, puntaje en matemáticas.

↑, indica que la variable influye a la hora de obtener puntajes altos. Por otra parte ↓ indica la posibilidad de obtener valores bajos, -, indica que la variable no hace parte, o no resultado significativa en el modelo óptimo.

Tabla 6 Resumen variables demográficas influyentes

VARIABLES DEMOGRÁFICAS (Explicativas)	GRADO DE INFLUENCIA (Variables respuesta)	
	PLC	PMAT
Género (masculino)	-	↑
Situación de su hogar actual	-	-
Edad	-	↓
Número de personas que conforman el hogar	-	↑

Nota: Se observa que ser hombre en comparación a las mujeres y convivir con más personas aumenta la posibilidad de obtener puntajes altos en matemáticas y lectura crítica. En cuanto a aquellos estudiantes con mayor edad disminuye la posibilidad de lograr puntajes altos. Por otra parte, ninguna de las variables demográficas resulto ser influyentes respecto al puntaje en lectura crítica, lo cual concuerda con lo planteado por Banco Mundial. (2009) pág. 36 y (Gaviria y Barrientos (2001)).

Fuente: Elaboración propia.

Tabla 7 Resumen variables socio-económicas influyentes

VARIABLES SOCIO- ECONÓMICAS (Explicativas)	GRADO DE INFLUENCIA (Variables respuesta)	
	PLC	PMAT
Estrato socioeconómico de la residencia del estudiante según factura de energía (estrato 2)	↑	↑
El hogar cuenta con conexión a internet	-	-
El hogar cuenta con Servicio cerrado de televisión	-	-
Tiene computador en su hogar	-	-
Número de horas que trabaja a la semana (Sí)	↓	↓
Número de dormitorios de la residencia		-
¿Cuántas veces por semana se comen los siguientes alimentos en su hogar? Leche o derivados (queso, yogurt, etc.) (casi todos los días)	↑	-
¿Cuántas veces por semana se comen los siguientes alimentos en su hogar? Carne (pollo, pavo, res, cordero, cerdo, conejo, etc.), pescados o huevos	-	↑
¿Cuántas veces por semana se comen los siguientes alimentos en su hogar? Cereales (avena, granola), frutos secos (almendras, maní) o legumbres (frijoles, garbanzos, lentejas)	-	-
Tiene consola de videojuegos en casa.	-	-
Ocupación madre	-	-
Ocupación padre (Trabajo profesional)	↑	↑

Nota: Se observa que consumir derivados lácteos casi todos los días aumenta la posibilidad de obtener puntajes altos en lectura crítica. En cuanto a la ocupación del padre como profesional obtiene mejores puntajes en lectura crítica y matemáticas, en comparación a una ocupación laboral de agricultor, pesquero o jornalero, lo cual

está de acuerdo con los planteamientos de (Coleman, 1966; Hanushek y Luque, 2003). Por otra parte, tener un estrato 2 aumenta las posibilidades de obtener buenos puntajes en lectura crítica y matemáticas, en comparación con estudiantes de estrato 1. Así mismo, al comparar dos estudiantes en tan sólo el consumo de carne, se tiene que aquellos estudiantes que la consumen de 3 a 5 veces por semana obtienen altos puntajes en matemáticas que aquellos que no la consumen. Así también un estudiante que trabaje obtiene meneos puntaje respecto a estudiantes que no lo hacen.

Fuente: Elaboración propia.

Tabla 8 Resumen variables socioeconómicas influyentes

VARIABLES ACADÉMICAS (Explicativas)	GRADO DE INFLUENCIA (Variables respuesta)	
	PLC	PMAT
Tiempo dedicado a la lectura diaria	-	-
Tiempo dedicado a internet.	-	-
Máximo nivel educativo alcanzado por el padre (Técnico o tecnológico)	↓	↓
Máximo nivel educativo alcanzado por la madre	-	-
Cuántos libros físicos o electrónicos hay en su hogar excluyendo periódicos, revistas, directorios telefónicos y libros del colegio (11 y 25 libros)	-	↑

Nota: Se observa que tener padre con nivel educativo técnico o tecnológico obtiene menores puntajes en lectura crítica y matemáticas, en comparación a tener niveles educativos profesionales, lo cual está de acuerdo con los planteamientos de (Ramírez, 2014) quien analizó algunas diferencias por nivel de formación evidenciando que en el nivel de formación técnico profesional, el nivel de estudios de la madre, específicamente el nivel técnico y tecnológico, tiene una asociación positiva y débil con el desempeño académico cuando se compara con estudiantes cuyas madres alcanzaron algún tipo de estudios de bachillerato. Por el contrario, para los estudiantes de programas profesionales, la educación de los padres resulta asociada débil y positivamente con su desempeño académico, particularmente en aquellos estudiantes cuyos padres alcanzaron algún nivel de posgrado en vez de algún tipo de estudios de bachillerato". Por otra parte, contar con un número de libros entre 11 y 25 en la familia aumenta la posibilidad de alcanzar mejores puntajes en matemáticas respecto a quienes sólo poseen una menor cantidad.

Fuente: Elaboración propia.

Conclusiones

Es importante mencionar que la evidencia estadística respalda los resultados encontrando para aquellas variables que influyen en los resultados de la prueba saber de undécimo grado a estudiantes de la Institución Educativa Agrícola la Holanda, ubicada en región del Ariari, Granada-Meta, por lo tanto, los resultados obtenidos sólo aplican a dicha población de estudio y a las características de la muestra seleccionada., en este sentido, es responsabilidad de la Institución Educativa desarrollar estrategias al interior, asimismo para autoevaluarse y tomar acciones de mejoramiento.

En cuanto a la variable respuesta lectura crítica, aquellas variables explicativas que resultaron significativamente influyentes son: el estrato, consumo de derivados lácteos, ocupación del padre (trabajo profesional). Obteniendo más puntos en dicha prueba. Por otra parte, el factor demográfico no explica el puntaje obtenido en lectura crítica.

Respecto a la variable respuesta puntaje matemáticas se determinó, que aquellas variables explicativas que afectan dicho puntaje son: género, número de personas que conforman el hogar, estrato, derivados lácteos, ocupación del padre (trabajo profesional), número de libros en la familia, obteniendo más puntos en dicha prueba.

En torno a los resultados encontrados, la Institución Educativa puede intervenir a través de estrategias a corto plazo, es así que, para mejorar dichos puntajes en la prueba saber de undécimo grado, se sugiere: capacitar al personal docente en cuanto a qué evalúa el ICFES y qué es realmente lo que está evaluando el docente, así como, informar y concientizar a los

estudiantes sobre lo pertinente e importancia de la prueba, la estructura de la misma y los beneficios que se pueden obtener de acuerdo a los resultados. Adicionalmente, es pertinente informar a la comunidad educativa (padres de familia) sobre los factores que afectan dichos puntajes a fin de concientizar a la comunidad sobre aquellas variables (número de libros en la familia, tiempo de lectura, aumento en el consumo de lácteos, verduras y cárnicos, entre otros.) en la cuales la institución no puede interponerse, a fin de intervenir dichas variables por medio de la comunidad.

Dado que en los aspectos demográficos la institución tampoco podría actuar, si se recomienda a la comunidad que desde el aspecto socio-económico y académico como lo es si el estudiante trabaja e incentivar la lectura se asuman estrategias internas, ya que estos factores influyen sobre la disminución de la probabilidad de obtener buenos resultados.

Considerando lo anterior se puede decir que el modelamiento es una herramienta que nos permitió explicar aquellas variables demográficas, académicas y socioeconómicas influyentes en la prueba saber de undécimo grado, a través de diversas técnicas estadísticas que van más allá de lo descriptivo.

Referencias

- Agudelo, A., Figueroa, Luz A. & Vásquez, L (2019, 8 de julio). *Relaciones causales de los factores que afectan el desempeño de los estudiantes en pruebas estandarizadas en Colombia*. Revista espacios, 40(23), 23-33
<http://www.revistaespacios.com/a19v40n23/19402323.html#:~:text=Categor%C3%ADa%20Educativa,como%20se%20analizar%C3%A1%20m%C3%A1s%20adelante.>
- Barahona, P. (2014). *Factores determinantes del rendimiento académico de los estudiantes de la Universidad de Atacama*. Estudios pedagógicos, 40(1), 25-39
<https://scielo.conicyt.cl/pdf/estped/v40n1/art02.pdf>
- Barrios, M., Frías M. (2016, 12 de junio). *Factores que influyen en el desarrollo y rendimiento escolar de los jóvenes de bachillerato*. Revista Colombiana de Psicología, 25(1), 63-82. <https://www.redalyc.org/pdf/804/804444652005.pdf>
- Carrascal, S., Sierra, I. (2014, 8 de febrero). *Influencia de los contextos de enseñanza en la calidad del aprendizaje universitario*. Revista Interuniversitaria de Formación del Profesorado, 29(2), 91-104.
<https://www.redalyc.org/pdf/274/27443659007.pdf>
- Castro, M., Ruiz, J. (2019, 10 de Octubre). *La educación secundaria y superior en Colombia vista desde las pruebas Saber*. Praxis & Saber, 10(24), 341-366.
<http://www.scielo.org.co/pdf/prasa/v10n24/2216-0159-prasa-10-24-341.pdf>
- Celis, M., Jiménez, O., Jaramillo, A. (2014, 21 de Abril) *¿Cuál es la brecha de la calidad educativa en Colombia en la educación media y en la superior?*, Universidad de Manizales.
<https://www.ices.gov.co/documents/20143/233733/Cual+es+la+brecha+de+calidad+educativa+en+colombia+en+educacion+media+y+superior.pdf>
- Cerón, S. (2020). *Análisis estadístico, multivariado en las pruebas saber Pro del programa de ingeniería en la producción acuícola de la universidad de Nariño 2016-2019*. [Tesis de especialización, Fundación universitaria los libertadores]. Repositorio Institucional Fundación Universitaria de los Libertadores].
https://repository.libertadores.edu.co/bitstream/handle/11371/3143/Cer%C3%B3n_Sandra_2020.pdf?sequence=4

- Charrys, A., Fuentes, A., Rodríguez, M., Prieto, N. (2016). *Factores Institucionales Asociados Al Logro Educativo De Las Pruebas Saber 11 De Ciencias Naturales En Colegios de la región del Caribe. Universidad del Norte*. [Tesis de maestría, Universidad del norte]. Repositorio universidad del Norte.
<http://manglar.uninorte.edu.co/bitstream/handle/10584/7514/ariel.pdf?sequence=1&isAllowed=y>
- Chávez, R., Redondo, J. (2007). *Variables y factores asociados al aprendizaje escolar. una discusión desde la investigación actual*. Estudios pedagógicos, 33(2), 155-175. <http://dx.doi.org/10.4067/S0718-07052007000200009>
- Chica, S., Galvis, D., Ramírez, A. (2010). *Determinantes del rendimiento académico en Colombia*. Revista universitaria EAFIT, 46(160), 48-72.
<https://repository.eafit.edu.co/bitstream/handle/10784/16801/document%20-%202020-07-06T150755.868.pdf?sequence=2&isAllowed=y>
- Duque, A., Ortiz, J. (2013, 20 de mayo). *Pruebas ICFES Saber 11 y su relación con el desempeño académico en estudiantes de primer semestre de psicología. Universidad Nacional de Colombia. Cuadernos Hispanoamericanos de Psicología*, 13(1), 26-36. <https://doi.org/10.18270/chps.v13i1.1355>
- Fernández, Y. O. (2014, 28 de Abril). *Variables académicas que influyen en el rendimiento académico de los estudiantes universitarios*. Investigación educativa, 15(27), 165-179. <https://core.ac.uk/download/pdf/304896164.pdf>
- Izquierdo Torres, N., Buitrago, D, (2020, 23 de Octubre). *¿Hay Desigualdad en la Elección de Carrera en Colombia?: Un Análisis de Algunos Factores Sociodemográficos?*. Social and Education History.
https://www.researchgate.net/publication/340217564_Hay_desigualdad_en_la_eleccion_de_carrera_en_Colombia_Un_analisis_de_algunos_factores_sociodemograficos
- Manrique, I., Carreño, C. (2014,21 de Abril). *Influencia de los padres en el rendimiento académico de sus hijos: una aproximación econométrica en el contexto de la*

- educación media colombiana*. Revista educación y desarrollo social.
<https://revistas.unimilitar.edu.co/index.php/reds/article/view/306/111>
- Marly Bahamón., L. R. (2014, 19 de Agosto). *Caracterización de la capacidad intelectual, factores sociodemográficos y académicos de estudiantes con alto y bajo desempeño en los exámenes Saber Pro - año 2012*. Universidad Simón Bolívar.
https://www.researchgate.net/publication/281559656_Caracterizacion_de_la_capacidad_intelectual_factores_sociodemograficos_y_academicos_de_estudiantes_con_alto_y_bajo_desempeno_en_los_examenes_Saber_Pro_-_ano_2012.
- Monroy, L. G. (2012). *Análisis estadístico de datos categóricos*. Bogotá: Universidad Nacional de Colombia.
- Morales, M. (2019). *Factores Determinantes En El Rendimiento De Los Estudiantes De La Región Pacífico-Colombiana, En Las Pruebas Saber-II*. [Tesis de pregrado, Universidad ICESI.] Repositorio Universidad ICESI.
https://repository.icesi.edu.co/biblioteca_digital/bitstream/10906/84736/1/TG02508.pdf
- Mundial, B. (2013, 23 de Enero). *La Educación Superior en Colombia*. Banco Mundial.
<https://www.oecd.org/education/skills-beyond-school/Evaluaciones%20de%20pol%C3%ADticas%20nacionales%20de%20Educaci%C3%B3n%20-%20La%20Educaci%C3%B3n%20superior%20en%20Colombia.pdf>
- Nacional, M. D. (2016, 26 de Abril). *La educación en Colombia*. Colombia.
https://www.mineducacion.gov.co/1759/articles-356787_recurso_1.pdf
- Pardo, J (2017). *Factores demográficos, académicos y socioeconómicos que influyen en los resultados del componente genérico de la prueba saber pro caso: licenciatura en matemáticas y estadística de la universidad pedagógica y tecnológica de Colombia, facultad seccional Duitama*. [Tesis de pregrado, Universidad Pedagógica y Tecnológica de Colombia]. Repositorio UPTC.
https://repositorio.uptc.edu.co/bitstream/001/2648/1/TGT_1262.pdf
- Pedraza, C., Triana, A. (2016, 18 de Octubre). *Factores asociados al desempeño en las pruebas saber de lenguaje y matemáticas*. Universidad Santo Tomás.

<http://soda.ustadistancia.edu.co/enlinea/congreso/congresoedu/4%20Curriculo%20y%20evaluacio%20B4n/4%204%20FACTORES%20ASOCIADOS%20AL%20DESEMPE%20EN%20LAS%20PRUEBAS%20SABER%20DE%20LENGUAJE%20Y%20MATEMATICAS.pdf>

Pérez, M. (2016), *Modelos aditivos generalizados para localización, escala y forma (gamlss)*, [Trabajo de maestría, Universidad de Vigo].

http://eio.usc.es/pub/mte/descargas/ProyectosFinMaster/Proyecto_1361.pdf

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Ramírez, C. (2015, 16 de abril). *Factores asociados al desempeño académico según nivel de formación pregrado y género de los estudiantes de educación superior*. Colombia Revista Colombiana de Educación, (66), 201-222.

<https://www.redalyc.org/pdf/4136/413635257009.pdf>

Ramírez, C. (2014, 17 de diciembre). *Factores socioeconomicos y educativos asociados con el desempeño académico, según nivel de formación y género de los estudiantes que presentaron la prueba saber pro en 2009*.

<https://www.icfes.gov.co/documents/20143/233983/Factores+socioeconomicos+y+educativos+asociados+a+desempeno+academico+segun+formacion+y+genero+Saber+Pro+2009.pdf>

Rodríguez, O., Cote, L & Malaver, J. (2019). *Factores asociados al desempeño en la prueba saber pro-2017: un estudio de los programas de psicología. Semillero de investigación en metodología y evaluación en psicología y educación / Universidad Nacional de Colombia*. Universidad Nacional de Colombia.

https://www.infopsicologica.com/documentos/2019/FACTORES_ASOCIADOS_DESEMPEÑO_PRUEBA_SABER_PRO.pdf

Romero, P., Arévalo, A. (2015). *Factores Que Afectan Los Resultados De Las Pruebas Saber 9° En El Área De Español, De La Institución Educativa Nuevo Horizonte, De La Ciudad De Girardot*. [Tesis de maestría, Universidad Del Tolima].

<http://repository.ut.edu.co/bitstream/001/1593/1/FACTORES%20QUE%20AFECTAN%20LOS%20RESULTADOS%20DE%20LAS%20PRUEBAS%20SABER>

%209%C2%B0%20EN%20EL%20%C3%81REA%20DE%20ESPA%C3%91OL
%2C%20DE%20LA%20INSTITUC.pdf

Rosário, P., Lourenco, A., Paiva, O., Rodríguez, A., Valle, A., Tuero, E. (2012, 12 de marzo). *Predicción del rendimiento en matemáticas: efecto de variables personales, socioeducativas y del contexto escolar*. *Psicothema*, 24(2), 289-295.
<http://www.psicothema.es/pdf/4013.pdf>

Timarán-Pereira, R., Caicedo-Zambrano, J., Hidalgo-Troya, A. (2019). *Árboles de decisiones para predecir factores asociados al desempeño académico de estudiantes de bachillerato en las pruebas saber 11°*. *Revista UPTC*, 9(2), 363-378.
https://revistas.uptc.edu.co/index.php/investigacion_duitama/article/view/9184/77

21