



Universidad Nacional Mayor de San Marcos

Universidad del Perú. Decana de América

Facultad de Ciencias Matemáticas

Escuela Profesional de Estadística

Regresión logística: factores de riesgo que inciden al trabajo infantil - Encuesta Nacional de Hogares, Perú 2021

TRABAJO DE SUFICIENCIA PROFESIONAL

Para optar el Título Profesional de Licenciada en Estadística

AUTOR

Fiorella Ninet CASTILLO MAMANI

ASESOR

Dr. Helfer Joel MOLINA QUIÑONES

Lima, Perú

2022



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Castillo, F. (2022). *Regresión logística: factores de riesgo que inciden al trabajo infantil - Encuesta Nacional de Hogares, Perú 2021*. [Trabajo de suficiencia profesional de pregrado, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Matemáticas, Escuela Profesional de Estadística]. Repositorio institucional Cybertesis UNMSM.

Metadatos complementarios

Datos de autor	
Nombres y apellidos	Fiorella Ninet Castillo Mamani
Tipo de documento de identidad	DNI
Número de documento de identidad	47645613
URL de ORCID	https://orcid.org/0000-0002-4106-3249
Datos de asesor	
Nombres y apellidos	Helfer Joel Molina Quiñones
Tipo de documento de identidad	DNI
Número de documento de identidad	40014631
URL de ORCID	https://orcid.org/0000-0003-1307-7253
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	Zoraida Judith Huamán Gutierrez
Tipo de documento	DNI
Número de documento de identidad	09890094
Miembro del jurado 1	
Nombres y apellidos	Ofelia Roque Paredes
Tipo de documento	DNI
Número de documento de identidad	06243124
Datos de investigación	
Línea de investigación	A.3.2.6. Análisis de Datos y Modelamiento de Problemas de la Sociedad (Empresa, Instituciones, Poblaciones locales, regionales y nacionales)

Grupo de investigación	No aplica.
Agencia de financiamiento	Sin financiamiento
Ubicación geográfica de la investigación	<p>Edificio: - País: Perú Departamento: Lima Provincia: Lima Distrito: San Juan de Lurigancho Centro poblado: - Urbanización: Santa Maria Manzana y lote: Mz b3, Lte 08 Calle: - Latitud: -11.9649492 Longitud: -76.9690855</p>
Año o rango de años en que se realizó la investigación	Marzo 2022 - junio 2022
URL de disciplinas OCDE	Estadísticas, Probabilidad https://purl.org/pe-repo/ocde/ford#1.01.03



UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS

Universidad del Perú. Decana de América
FACULTAD DE CIENCIAS MATEMÁTICAS
ESCUELA PROFESIONAL DE ESTADÍSTICA

**ACTA DE SUSTENTACIÓN DEL TRABAJO DE SUFICIENCIA PROFESIONAL EN
LA MODALIDAD VIRTUAL PARA LA OBTENCIÓN DEL TÍTULO
PROFESIONAL DE LICENCIADA EN ESTADÍSTICA
(PROGRAMA DE TITULACIÓN PROFESIONAL 2022-I)**

En la UNMSM – Ciudad Universitaria – Facultad de Ciencias Matemáticas, siendo las 09:00 horas del sábado 23 de julio del 2022, se reunieron los docentes designados como Miembros del Jurado Evaluador (PROGRAMA DE TITULACIÓN PROFESIONAL 2022-I): Dra. Zoraida Judith Huamán Gutiérrez (PRESIDENTE), Dra. Ofelia Roque Paredes (MIEMBRO) y el Dr. Helfer Joel Molina Quiñones (MIEMBRO ASESOR), para la sustentación del Trabajo de Suficiencia Profesional titulado: **“REGRESIÓN LOGÍSTICA: FACTORES DE RIESGO QUE INCIDEN AL TRABAJO INFANTIL - ENCUESTA NACIONAL DE HOGARES, PERÚ 2021”**, presentado por la señorita **Bachiller Fiorella Ninet Castillo Mamani**, para optar el Título Profesional de Licenciada en Estadística.

Luego de la exposición del Trabajo de Suficiencia Profesional, la Presidente invitó a la expositora a dar respuesta a las preguntas formuladas.

Realizada la evaluación correspondiente por los Miembros del Jurado Evaluador, la expositora mereció la aprobación de **SOBRESALIENTE**, con un calificativo promedio de **DIECISIETE (17)**.

A continuación, los Miembros del Jurado Evaluador dan manifiesto que la participante **Bachiller Fiorella Ninet Castillo Mamani** en vista de haber aprobado la sustentación de su Trabajo de Suficiencia Profesional, será propuesta para que se le otorgue el Título Profesional de Licenciada en Estadística.

Siendo las 09:30 horas se levantó la sesión firmando para constancia la presente Acta.

Dra. Zoraida Judith Huamán Gutiérrez
PRESIDENTE

Dra. Ofelia Roque Paredes
MIEMBRO

Dr. Helfer Joel Molina Quiñones
MIEMBRO ASESOR



UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
Universidad del Perú. Decana de América
FACULTAD DE CIENCIAS MATEMÁTICAS
ESCUELA PROFESIONAL DE ESTADÍSTICA

INFORME DE EVALUACIÓN DE ORIGINALIDAD

El Director de la Escuela Profesional de Estadística, Dr. Roger Pedro Norabuena Figueroa, informa lo siguiente:

1. Operador del programa informático de similitudes: Dr. Roger Pedro Norabuena Figueroa
2. Documento evaluado: Trabajo de Suficiencia Profesional para optar el Título Profesional de Licenciado en Estadística, titulado: REGRESIÓN LOGÍSTICA: FACTORES DE RIESGO QUE INCIDEN AL TRABAJO INFANTIL – ENCUESTA NACIONAL DE HOGARES, PERÚ 2021
3. Autor de la tesis: FIORELLA NINET CASTILLO MAMANI
4. Fecha de recepción de la tesis: 30/11/2022
5. Fecha de aplicación del programa informático de similitudes: 30/11/2022
 - Software utilizado: Turnitin
6. Configuración del programa detector de similitudes:
 - Excluye textos entrecomillados
 - Excluye bibliografía
 - Excluye cadenas menores a 40 palabras
7. Porcentaje de similitudes según programa detector de similitudes: siete por ciento (7%)
8. Fuentes originales de las similitudes encontradas:
 - Fuentes de internet: 6 %
 - Publicaciones: 0 %
9. Calificación de originalidad:
 - Documento cumple criterios de originalidad, sin observaciones

Lima, 30 de noviembre del 2022



Firmado digitalmente por
NORABUENA FIGUEROA Roger
Pedro FAU 20148092282 soft
Motivo: Soy el autor del documento
Fecha: 01.12.2022 13:47:45 -05:00

Dr. Roger Pedro Norabuena Figueroa
Director

RESUMEN

El trabajo infantil es uno de los temas de mayor importancia que involucra la generación de nuevas políticas públicas enfocadas en su prevención y erradicación dentro del estado peruano, es por ello que en esta investigación abordaremos como principal estudio conocer los factores de riesgo que inciden al trabajo infantil. Considerando como fuente secundaria la Encuesta Nacional de Hogares 2021 la base de datos.

El estudio es descriptivo y transversal con un diseño observacional, no experimental. La selección de la muestra solo abarco niñas, niños y adolescentes entre 5 a 13 años de edad (19,546 casos). A partir de ello, se determinará los factores de riesgo utilizando el modelo de regresión logística múltiple. Donde, por medio de esta técnica Paramétrica se ha podido conocer los factores de riesgo que inciden al trabajo infantil en el 2021; en el cual se encontraron 12 variables que influyen al riesgo en que un menor se encuentre realizando algún tipo de trabajo ($X^2=46.88$; $p_valor=0.0000$). Dentro de estas se encuentra: sexo, área de residencia, región rural, macro región, la variable sabe leer y escribir, lengua materna, ultimo nivel de estudios que asistió el 2020, actualmente asiste algún centro de estudios, el tipo de centro de estudios, la relación de parentesco con el jefe del hogar, nivel de pobreza, N° miembros del hogar (variable agrupada).

Palabras clave: Trabajo infantil, Factores de riesgo, Regresión logística, odds ratio.

ABSTRACT

Child labor is one of the most important issues that involves the generation of various public policies focused on its prevention and eradication within the Peruvian state, which is why in this research we will address as the main study to know the risk factors that affect child labor. child labour. Considering the 2021 National Household Survey as a secondary source, the database.

The study is descriptive and cross-sectional with an observational, non-experimental design. The selection of the sample only included children and adolescents between 5 and 13 years of age (19,546 cases). From this, the risk factors will be determined using the multiple logistic regression model. Where, through this Parametric technique, it has been possible to know the risk factors that affect child labor in 2021; in which 12 variables were found that influence the risk that a minor is doing some type of work ($X^2=46.88$; $p_value=0.0000$). These include: sex, area of residence, rural region, macro region, the variable knows how to read and write, mother tongue, last level of studies attended in 2020, currently attends a study center, type of study center , the relationship of kinship with the head of the household, level of poverty, N° members of the household (grouped variable).

Keywords: Child labor, Risk factors, Logistic regression, odds ratio.

INDICE

I. INTRODUCCIÓN.....	8
II. INFORMACIÓN DEL LUGAR DONDE SE DESARROLLO LA ACTIVIDAD	11
2.1. Institución donde se desarrolló la actividad.....	11
2.2. Periodo de duración de la actividad.....	11
2.3. Finalidad y objetivos de la entidad	11
2.3.1. Finalidad	11
2.3.2. Objetivo.....	11
2.4. Razón social.....	11
2.5. Dirección postal	11
2.6. Correo electrónico del profesional a cargo	12
2.1. Organigrama del MIMP	13
III. DESCRIPCIÓN DE LA INVESTIGACIÓN.....	14
3.1. Organización de la investigación.....	14
3.2. Finalidad y objetivos de la investigación.....	15
3.2.1. Finalidad	15
3.2.2. Objetivo general.....	15
3.2.3. Objetivos específicos	15
3.3. Marco teórico.....	15
3.3.1. Antecedentes nacionales	15

3.3.2. Antecedentes internacionales.....	17
3.4. Bases teóricas.....	19
3.5. Metodología y procedimientos	24
3.5.1. Tipo y diseño estadístico.....	24
3.5.2. Población.....	25
3.5.3. Muestra	25
3.5.4. Variables de investigación	25
3.6. Resultados	27
3.6.1. Análisis descriptivo.....	27
3.6.2. Análisis del modelo de regresión logística múltiple.....	33
3.6.3. Interpretación de los parámetros.....	36
3.6.4. Evaluación de la capacidad predictiva del modelo	40
3.6.5. Evaluación de la bondad de ajuste	41
IV. CONCLUSIONES.....	41
V. RECOMENDACIONES.....	44
VI. BIBLIOGRAFIA	45
VII. ANEXOS.....	48

Índice de tablas

Tabla 1. Descripción de Variables de Estudio	25
Tabla 2. Menores de 5 a 13 años según variables sociodemográficas, ENAHO 2021	27
Tabla 3. Menores de 5 a 13 años según variables de ámbito educativo, ENAHO 2021 ..	28
Tabla 4. Frecuencia de NNA según las variables de ámbito del hogar, ENAHO 2	29
Tabla 5. Evaluación de la asociación de los factores demográficos en el TI.....	31
Tabla 6. Evaluación de la asociación de los factores de educación respecto al TI.....	31
Tabla 7. Evaluación de la asociación de los factores del hogar.....	33
Tabla 8. Modelo de regresión logística múltiple ajustado N°1.....	33
Tabla 9. Modelo de regresión logística múltiple ajustado N°2.....	34
Tabla 10. Selección del modelo	35
Tabla 11. Factores de riesgo que inciden al trabajo infantil	39
Tabla 12. Sensibilidad, especificidad y porcentaje de clasificación.....	40
Tabla 13. Tabla de comparación de las Predicciones estimadas del trabajo infantil	41
Tabla 14. Bondad de ajuste del modelo	41

Índice de figuras

Figura 1. Objetivos prioritarios de la PNMNNA.....	9
Figura 2. Indicadores estadísticos de Actividades económicas y trabajo infantil.....	10
Figura 3. Organigrama	13
Figura 4. Gráfico de Modelo Logístico.....	20
Figura 5. <i>Tabla de contingencia del cálculo de OR</i>	24
Figura 6. Porcentaje de NNA de 5 a 13 años en TI	30
Figura 7. Objetivos prioritarios e indicadores de la PNMNNA.....	49

I. INTRODUCCIÓN

En la actualidad, se han presentado diversos factores que han hecho que las niñas, niños y adolescentes (NNA) inicien actividades laborales a temprana edad. Según la Organización Internacional del Trabajo (OIT), “son 152 los millones de niños que trabajan incluso antes de aprender a leer o escribir”.

En nuestro país, los avances para erradicar la explotación del trabajo infantil se han visto truncadas y se han agravado con el inicio de la pandemia de COVID-19, su impacto en la sociedad origino el incremento de factores que ocasionan el trabajo infantil, entre estos se encuentra la pobreza, el desempleo de las personas adultas, la exclusión educativa, la caída de los ingresos familiares, la debilidad en la aplicación de las políticas y las percepciones culturales que lo naturalizan.

Hoy en día, en nuestro estado peruano hasta el momento se han admitido 4 planes de acción por la infancia y adolescencia (PNAIA), entre estos encontramos: 1^{er} Plan Nacional (1992 - 1995), 2^{do} Plan Nacional (1996 - 2000), el 1^{ro} contuvo 3 objetivos generales (Supervivencia, Desarrollo y Protección), 3^{er} plan abarco entre 2002 y 2010; este incorporo 4 objetivos estratégicos organizados según el ciclo de vida del NNA y finalmente tenemos al 4^{to} plan nacional (2012 – 2021). Con la intención de avanzar el fortalecimiento de las políticas y principios del Estado en materia de NNA, el 25 de junio del 2021 según el DECRETO SUPREMO N° 008-2021-MIMP se aprueba la Política Nacional Multisectorial para las Niñas, Niños y Adolescentes al 2030 (PNMNNA), el cual será el instrumento marco de políticas públicas en temas de niñez y adolescencia en nuestra nación, el cual lidera la acción del estado en sus tres niveles de régimen, a favor de asegurar el cumplimiento de los derechos de las NNA.

Cuya conducción está a cargo del MIMP, a través de la Dirección General de Niñas, Niños y Adolescentes, como ente rector del Sistema Nacional de Atención Integral al NNA.

En este marco, la PNMNNA cuenta con 5 Objetivos prioritarios - OP (Ver anexo 1):

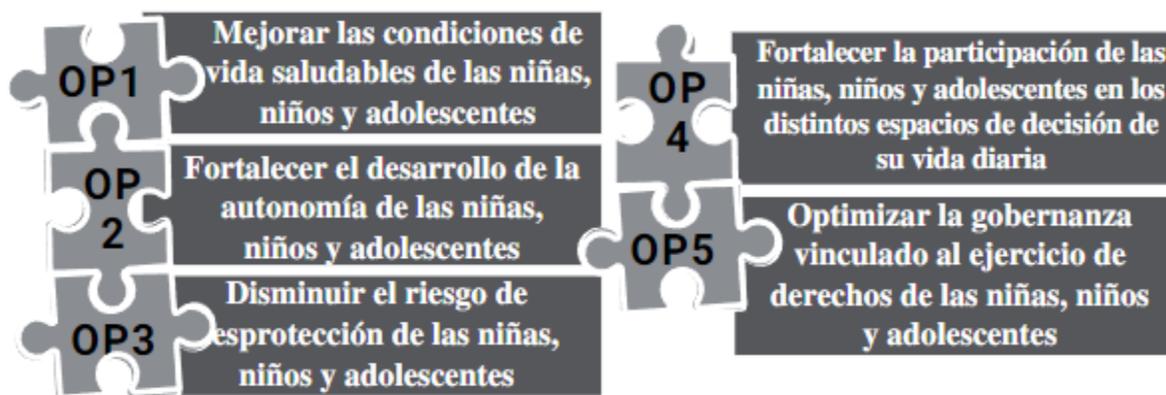


Figura 1. Objetivos prioritarios de la PNMNNA

Los objetivos prioritarios presentados en la gráfica 1, en total cuenta con 17 indicadores a nivel de resultados que medirán los cambios que se van obteniendo en beneficio para las NNA. Con el propósito de conocer los factores de riesgo que inciden al trabajo infantil en los menores de 5 a 13 años de edad nos centraremos en estudiar con mayor énfasis el OP3, el cual contempla 6 indicadores (Ver anexo 1). Para este estudio tomaremos como referencia el indicador del porcentaje de niños y niñas de 5 a 13 años de edad que trabajan por debajo de la edad mínima, ya que este indicador solo mide en que porcentaje los NNA se encuentran ocupados en alguna actividad económica mas no refleja la problemática global de trabajo infantil en los NNA. Está investigación promueve evaluar el impacto global que refleja el estudio del indicador de “Trabajo infantil en niños(as), adolescentes entre 5 a 13 años a nivel nacional”.

Según la Resolución Ministerial N° 114-2016-TR, el cual muestra el procedimiento de cómo obtener los indicadores de trabajo infantil, indica que, se considera el N° de niños, niñas de

5 a 11 años que realizan al menos 1 hora a la semana, una o más actividades económicas y aquellos adolescentes en trabajo intensivo en horas de 12 a 17 años.

En el estudio que abarcaremos, adaptaremos este indicador para reflejar el trabajo infantil en solo NNA de 5 a 13 años.

De esta manera se inicia el presente trabajo de suficiencia profesional, por la necesidad de reflejar a nivel global la problemática de los factores de riesgo que inciden al trabajo infantil en los niños, niñas y adolescentes de 5 a 13 años de edad.

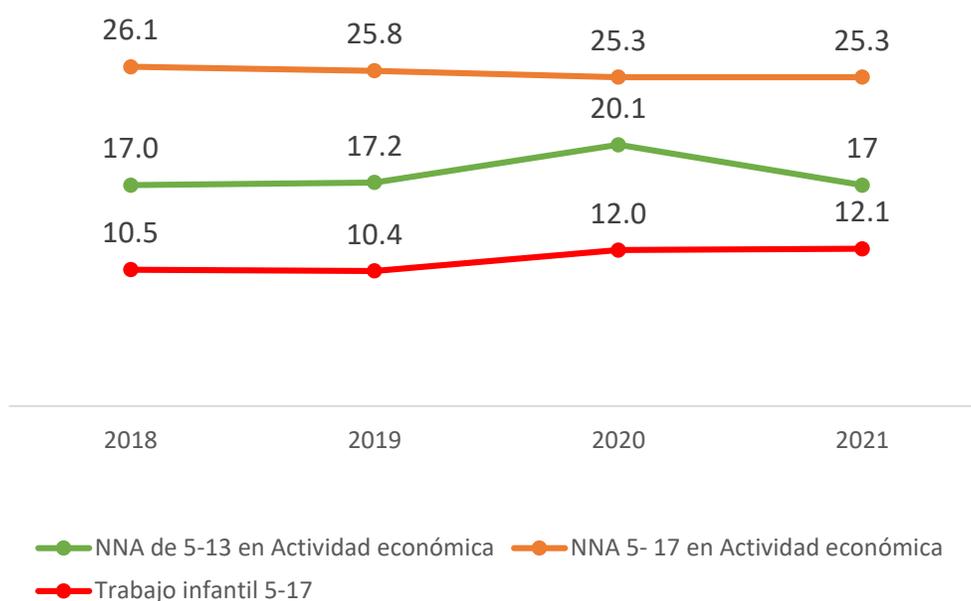


Figura 2. Indicadores estadísticos de Actividades económicas y trabajo infantil

Fuente. INEI-ENAH0

II. INFORMACIÓN DEL LUGAR DONDE SE DESARROLLA LA ACTIVIDAD

2.1. Institución donde se desarrolló la actividad

Dirección de Políticas de Niñas, Niños y Adolescentes de la Dirección General de Niñas, Niños y adolescentes del Ministerio de la Mujer y Poblaciones Vulnerables.

2.2. Periodo de duración de la actividad

Del 11 de marzo del 2021 hasta el 10 marzo del 2022

2.3. Finalidad y objetivos de la entidad

2.3.1. Finalidad

La DPNNA tiene como finalidad suscitar políticas sobre la infancia y adolescencia, además ve el cumplimiento y seguimiento de los objetivos del Plan Nacional de Acción por la Infancia y la Adolescencia 2012-2021, así como la PNMNNA al 2030.

2.3.2. Objetivo

Seguimiento y sistematización de la Estructuración Funcional Programática a nivel nacional y sectorial en el marco del PNAIA (Art. 3 de la Ley N° 28487). Así como también propone normas en materia de infancia y adolescencia PNAIA 2012-2021. Además, coordina, asesora y recopila información e implementa acciones multisectoriales en temas de niñez y adolescencia.

2.4. Razón social

Ministerio de la Mujer y Poblaciones Vulnerables

2.5. Dirección postal

Jr. Camaná 616 – Lima (7° piso)

2.6. Correo electrónico del profesional a cargo

ncastillo@mimp.gob.pe

2.1. Organigrama del MIMP

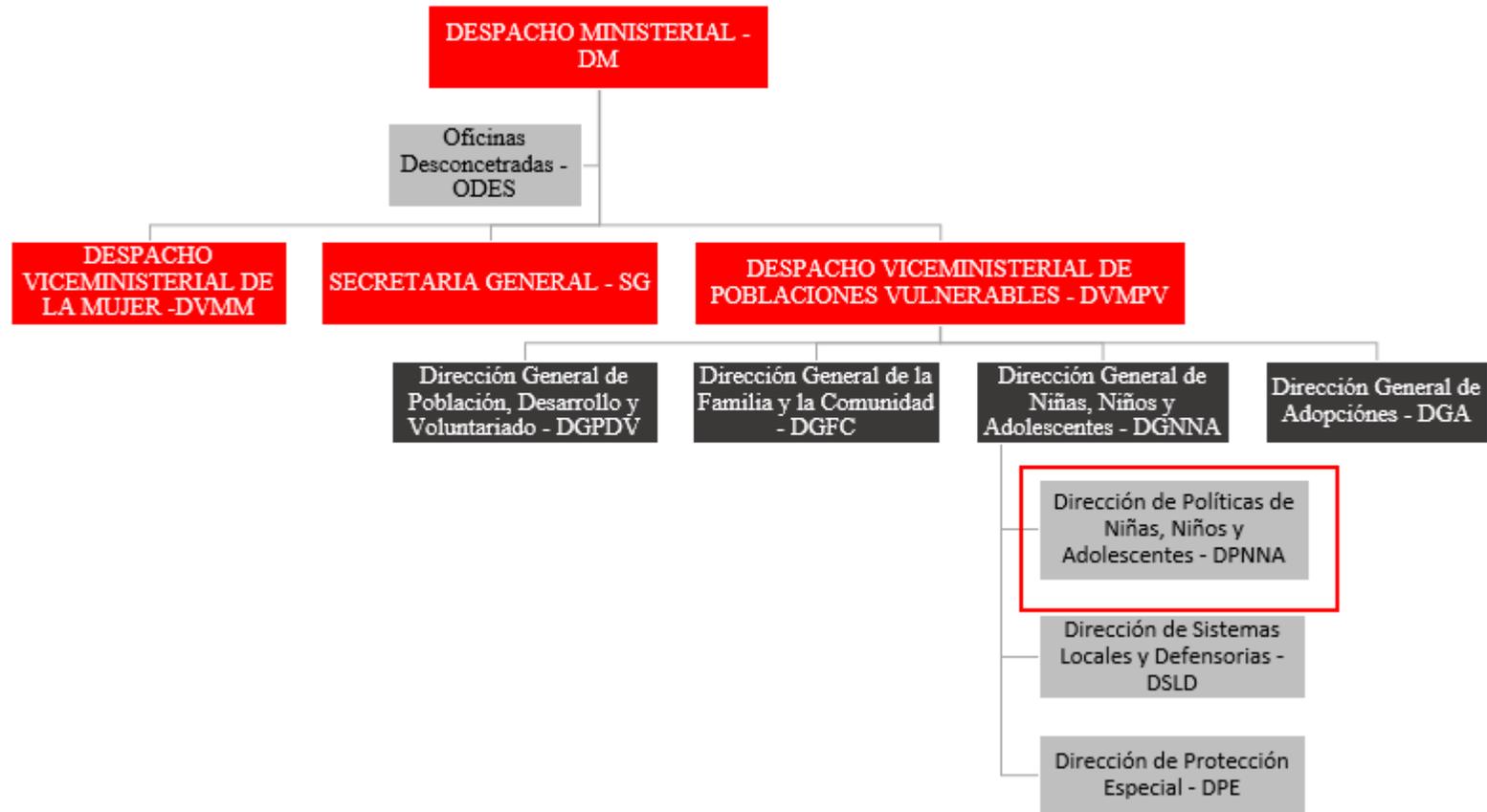


Figura 3. Organigrama

Nota. Organigrama nominal del MIMP. Referenciado de la página web.

(https://www.mimp.gob.pe/files/transparencia/organigrama/organigrama_nominal-04022021.pdf).

III. DESCRIPCIÓN DE LA INVESTIGACIÓN

3.1. Organización de la investigación

Determinar la problemática de estudio es el inicio a la elaboración de este trabajo, distinguir la existencia del trabajo infantil, particularmente en la población de 5 a 13 años y porque es importante determinar los factores de riesgo que inciden en este, será de importancia para futuros estudios y evaluación a la generación de nuevas políticas públicas.

Para tener un mejor entendimiento de los datos y las variables que utilizaremos nos basamos en diversas fuente bibliográficas relacionados al trabajo infantil, seguidamente iniciamos la preparación de los datos, el cual se extrajo de la página principal del INEI inei.inei.gov.pe/microdatos , el cual extrajimos nuestra base de datos en el programa STATA 16 de la Encuesta Nacional de Hogares 2021, posterior a ello iniciamos la extracción de información basándonos en los módulos de característica de vivienda y hogar, miembros del hogar, educación, empleo e ingresos y sumaria. seguidamente pasamos a unir los módulos de interés e iniciamos el proceso limpieza, construcción de nuevas variables*- con el fin de obtener la base de datos final (Para construir el indicador de trabajo infantil se consideró el factor de expansión **facpob07 “Correspondiente al módulo 200”** y **fac500a “correspondiente al módulo 500”**), según un estudio realizado por (Herrero & Bocco, 2007), hace mención que evaluar un estudio considerando las ponderaciones y otro no evaluándolo permiten de igual forma llegar a los mismos resultados estadísticos y no produce sesgos relevantes en las estimaciones permitiendo que los resultados sean mejores.

Finalmente se construyó el modelo de regresión logística múltiple con las variables que fueron significativas para el modelo y se procedió a evaluar la bondad de ajuste con la prueba de Hosmer Lemeshow.

3.2. Finalidad y objetivos de la investigación

3.2.1. Finalidad

Por medio de la técnica Paramétrica, haciendo uso del Modelo de Regresión Logística Múltiple se determinará los factores de riesgo que inciden al trabajo infantil en las niñas, niños y adolescentes de 5 a 13 años de edad.

3.2.2. Objetivo general

Determinar los factores de riesgo que inciden al trabajo infantil, tomando como base la Encuesta Nacional de Hogares, Perú 2021.

3.2.3. Objetivos específicos

- Obtener el porcentaje de NNA que se encuentran en trabajo infantil, para obtener los factores de riesgo que inciden en este.
- Construir un modelo de regresión logística que permita identificar los factores Sociodemográficos que inciden al riesgo para que un menor de edad entre 5 a 13 años se encuentre en situación de trabajo infantil.

3.3. Marco teórico

En este apartado conoceremos algunos antecedentes que influyeron en el desarrollo del presente estudio.

3.3.1. Antecedentes nacionales

Aguayo y otros (2019) en su artículo con el nombre “El trabajo infantil: La relación de los factores sociodemográficos, adaptación académica y la actividad laboral con la repitencia en niños y adolescentes de 10-17 años en los departamentos de Lima y Huánuco”. Se basa en conocer la relación de repitencia con los factores sociodemográficos (edad, lugar de residencia y

sexo), en **adolescentes** del grupo de 10-17 años de Lima y Huánuco. La investigación tuvo una muestra de 423 individuos, extraído de la base de datos de la encuesta nacional especializada del trabajo infantil (ETI)-INEI 2015. Del análisis descriptivo se obtuvo que no hay asociación en la repitencia con el factor sociodemográfico $p_valor > 0.05$, también se obtuvo que el modelo planteado según la prueba de Hosmer y Lemeshowes es adecuado ($X^2=3.546$; $p_valor=0.896$).

Libaque y otros (2017) en el Symposiun on Information Management and Big Data el cual tuvo la presentación del trabajo de investigación titulado “Predicting child labor in Peru: A comparison of logistic regression and neural networks techniques” busca realizar una comparación entre dos técnicas estadísticas “Regresión Logística versus Redes neuronales”, que le permitirá comparar el poder predictivo de los modelos tradicionales y los modernos en relación al trabajo infantil. En el modelo de regresión Logística el autor muestra las significancias de las variables a través de la prueba de chi-cuadrado ($\chi^2=2300.885$, $df=25$, $p=0.000$) y a través de la especificidad sensibilidad pudo obtener que su modelo predijo al 82,61% de los casos. Posteriormente paso a evaluar la prueba de Hosmer & Lemeshow. Respecto a la prueba de red Neuronal hizo uso de la sensibilidad y especificidad para evaluar la predictividad y midió la eficiencia del modelo con la prueba de MAPE=11.74%, finalmente realizó una comparación gráfica de ambos modelos mediante la curva ROC. Al concluir su estudio llegó a la conclusión que la sensibilidad de la técnica de la red neuronal fue superior a los valores obtenidos del modelo logit.

A pesar de estos resultados, el modelo logit no pudo superar a la Red Neuronal ya que esta logró capturar fuentes de variación que no son identificadas por la técnica logit. Una predicción precisa de este fenómeno.

Pinedo & Condori (2019) presentaron una tesis titulada “Incidencia de la situación socioeconómica en el trabajo infantil de la región del cusco en el año 2018” tuvo como objetivo identificar como la situación de trabajo en menores de 5 a 14 años (878 personas) en la region Cusco, para obtener ello usó el modelo econometrico conocido como regresión logit, obteniendo como resultado de la curva ROC un valor de 0.63 el cual se considero en este estudio como un valor aceptable y aplicando la prueba de Hosmer & Lemeshow, se obtuvo 0.1257 el cual tuvo como conclusión que el TI es ocasionado por las condiciones socioeconomicas de las familias del Cusco, esto involucra el nivel de ingreso del hogar, nivel que tuvo de estudios y la zona de residencia. El aporte de este antecedente es brindar información estadística para combatir el trabajo infantil en región Cuzco y permite enfocar el estudio a la contribución y mejora de ciertas politicas publicas encargadas a erradicar el trabajo infantil.

3.3.2. Antecedentes internacionales

Thiong'o (2020) en su trabajo de investigación que lleva por título “Modeling The Key Determinant of Child Labour In Kenya” tiene como objetivo hallar un modelo adaptable a la resolución de problemas relacionados con el Trabajo Infantil en niños de 5 a 14 años utilizando los datos de la encuesta de hogares KNBS de 2017. Para ello uso la siguiente metodología de regresión logística binaria de efectos mixtos para analizar los datos, el estudio encontró un mejor modelo con las técnicas de $AIC=4524.2$ Y $BIC=4604.5$, cuyo valor de chi-cuadrado no salió significativo $\chi^2 = 152.04$, $gl = 1$, $p = 0.0000$, el cual indica que el modelo es adecuado. Con lo cual se pudo concluir de las principales variables más importantes que por cada aumento en la edad de un niño, aumenta el trabajo infantil en un 11,9%, lo que es estadísticamente significativo $\alpha=0.0000\%$, así como el sexo de niño el cual indico que los hombres tienen más a realizar trabajo infantil 4.4%. En tanto, la característica del hogar, por cada año de aumento en el grado más alto

al que asiste el jefe del hogar, la probabilidad de que el niño del hogar esté involucrado en trabajo infantil disminuye en 2.5%.

Abdallah (2020) en su artículo científico que tiene por nombre “Using logistic regression to identify the factors affecting child labor in Red Sea State” busca aplicar la Regresión Logística para identificar los factores que afectan al TI en el Estado del mar, Para lograr su objetivo el estudio son niños entre 7-15 años. El tamaño de la muestra fue (133) niños. Luego de elegir las variables que caracterizaran su modelo. Los resultados mostraron -2 para el Valor de Log Likelihood para el modelo reducido, con una estadística de chi-cuadrado = 58,749 con (valor de $p = 0,0001$). Finalmente, los resultados en la tabla de clasificación fueron 82% de las predicciones correctas, haciendo uso de regresión logística binaria. Concluyendo que, resultados reveló que existe una relación entre niño mano de obra y las variables independientes (edad, N° de miembros de la familia, estado civil de los padres y deseo de los padres).

Villarruel y otros (2019) con el trabajo de titulación llamado “Factores que inciden en el trabajo infantil, Ecuador año 2017”, el cual analiza un conjunto de factores sociodemográficos (demográficos, sociales, culturales, económicos y académicos) que inciden al trabajo laboral en los NNA en Quito (Colombia), donde su principal objetivo fue obtener resultados descriptivos de las características sociodemográficas en Quito, donde realiza un estudio de tipo cuantitativo; además plantea un modelo económico de regresión logit, donde se evidencio que la probabilidad de que un menor trabaje para el 2017 se incrementa es cuando las características son de tipo sociales, tal es el caso como el sexo masculino, área rural, la pobreza por NBI y la edad, mientras que, entre las características culturales se encuentra la etnia indígena, montubia. Además, dentro de lo económico y demográfico se encontró lo siguiente: región sierra, amazonias y el empleo

del jefe de hogar es no reenumerado. La bondad de ajuste que se presenta (pseudo R-cuadrado) un 30.23%, siendo este de importancia para el estudio que presentaron los autores.

3.4. Bases teóricas

En esta sección extraeremos más información referente a bases teóricas del autor (Dobson AM & Barnett , 2018) para explicar el desarrollo de nuestro modelo de Regresión Logística Múltiple.

Modelo de Regresión Logística Simple

Los métodos de regresión se han convertido en un componente importante para el análisis de datos, el cual describe la relación entre una variable de respuesta y una o más variables predictivas. Pero en algunos casos se tiene que, la variable respuesta (variable dependiente) no es continua sino más bien categórica y toma sólo dos (binaria) o más (multinomial) posibles valores.

El Modelo de regresión logística establece que la variable respuesta Y es una variable dicotómica que toma dos respuestas, de tal manera que cada individuo de la muestra tiene que pertenecer a una y sólo una, de estas alternativas (clases o grupos).

$$Y = \begin{cases} 1 & p(Y = 1) = \pi \\ 0 & p(Y = 0) = 1 - \pi \end{cases}$$

Si la variable Y tiene distribución Bernoulli, las observaciones son independientes, entonces la variable aleatoria tiene distribución de Bernoulli con:

$$P(Y=1 / X=x) = E(Y/X=x) = \pi(x)$$

La varianza condicional de Y dado que $X=x$, es dada por:

$$V(Y/X=x) = \pi(x)(1-\pi(x))$$

La probabilidad condicional puede calcularse a partir de un modelo que tiene la forma de una curva sigmoidea, en particular esta curva sigmoidea puede ser la función logística.

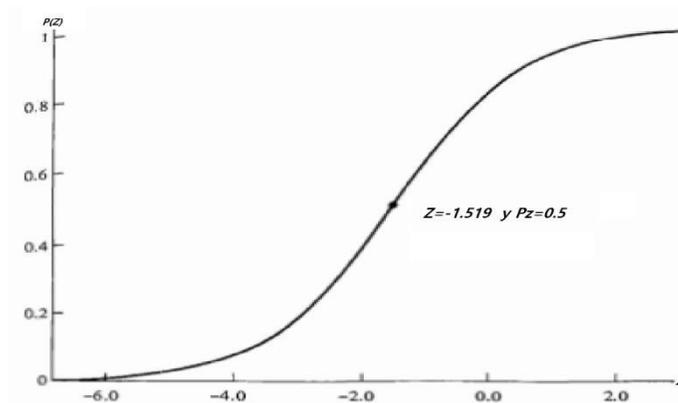


Figura 4. Gráfico de Modelo Logístico

La transformación utilizada para linealizar el modelo logístico corresponde al enlace canónico para una distribución binomial y también denominada enlace logit o transformación logit, dado por:

$$\theta = \ln\left(\frac{E(Y / X = x)}{1 - E(Y / X = x)}\right) = \ln\left(\frac{\pi(x)}{1 - \pi(x)}\right) \dots (1)$$

Aplicando la transformación de inversa a (1):

$$\exp\left\{\ln\left(\frac{\pi(x)}{1 - \pi(x)}\right)\right\} = e^{\theta}$$

$$\frac{\pi(x)}{1 - \pi(x)} = e^{\theta}$$

$$\frac{1 - \pi(x)}{\pi(x)} = \frac{1}{e^{\theta}}$$

$$\frac{1}{\pi(x)} - 1 = \frac{1}{e^\theta}$$

$$\frac{1}{\pi(x)} = 1 + \frac{1}{e^\theta} = \frac{e^\theta + 1}{e^\theta}$$

$$\pi = \pi(x) = E(y) = \frac{e^\theta}{1 + e^\theta}$$

Donde: $\theta = x\beta$; $E(y) = E(Y/X = x)$, se obtiene el modelo de regresión logística:

$$\pi(X) = \frac{e^{x\beta}}{1 + e^{x\beta}} \dots\dots\dots (2)$$

Modelo de Regresión Logística Múltiple

Podemos extender su aplicación de (1) para el caso en que se tenga varias variables explicativas, X_j donde $j=1,2,3,\dots,k$

$$\pi(X) = \frac{e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_K X_K}}{1 + e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_K X_K}} \dots\dots\dots (3)$$

Según (Dobson AM & Barnett, 2018) el modelo logístico nominal es usado cuando no tiene un orden natural en las categorías de respuesta. Tomamos una categoría referencial.

Supongamos que esta es la primera categoría. Entonces los logit para las otras categorías están definidas por:

$$\text{logit}(\pi_j) = \log\left(\frac{\pi_j}{\pi_1}\right) = X_j^T \beta_j, \text{ para } j=2,\dots,J \dots\dots\dots (4)$$

De la ecuación 4, (J-1) se utilizan simultáneamente para estimar los β_j . Una vez obtenidas las estimaciones de los parámetros β_j , los predictores lineales se pueden obtener de (3)

$$\hat{\pi}_j = \hat{\pi}_1 \exp(X_j^T \hat{\beta}_j) \quad \text{para } j = 2, \dots, J$$

Pero $\widehat{\pi}_1 + \widehat{\pi}_2 + \widehat{\pi}_3 + \dots + \widehat{\pi}_j = 1$, entonces

$$\widehat{\pi}_1 = \frac{1}{1 + \sum_{j=2}^J \widehat{\exp}(X_j^T b_j)}$$

y

$$\widehat{\pi}_j = \frac{\widehat{\exp}(X_j^T b_j)}{1 + \sum_{j=2}^J \widehat{\exp}(X_j^T b_j)}, \text{ para } j = 2, \dots, J$$

Las frecuencias esperadas, para cada patrón de covariable se obtiene multiplicando las probabilidades estimadas $\widehat{\pi}_j$ por la frecuencia total del patrón de covariable.

- **Estadísticas de Bondad de ajuste**

- **Estadística de Chi-cuadrado**

$$\chi^2 = \sum_{i=1}^N r_i^2, \text{ donde Residuales } = r_i = \frac{\text{Frecuencias observadas}_i - \text{Frecuencias esperadas}_i}{\sqrt{\text{Frecuencias esperadas}_i}}$$

- **Desviación:** Descrita con los valores máximos de log-verosimilitud

función para el modelo ajustado $l(b)$, y para el modelo maximal, $l(b_{max})$,

$$D = 2 [l(b_{max}) - l(b)]$$

Estadística de X^2 de razón de verosimilitud: Descrita en términos del valor máximo de la función de verosimilitud logarítmica para el modelo mínimo, $l(b_{min})$, y, $l(b)$.

$$M = 2 [l(b) - l(b_{min})]$$

- **Seudo R^2**

$$\text{Pseudo } R^2 = \frac{l(b_{min}) - l(b)}{l(b_{min})}$$

- **Criterio de informacion de Akaike**

$$AIC = -2l(\hat{\pi}; y) + 2p$$

Si el modelo se ajusta bien, tanto χ^2 y el D tienen asintóticamente, la distribución $\chi^2(N - P)$, donde P es el número de parámetros estimados. C tiene la distribución asintótica $\chi^2[P - (J - 1)]$ porque el modelo mínimo tendrá un parámetro para cada logit definido en (2). AIC se utiliza principalmente para comparaciones entre modelos que no están anidados. A menudo es más fácil interpretar los efectos de los factores explicativos en términos de razones de probabilidad que los parámetros β . Para simplificar, considere una variable de respuesta con J categorías y una variable explicativa dicotómica x que indica si un factor de "exposición" existe ($x = 1$) o no ($x = 0$). La razón de probabilidades de exposición para los valores j ($j = 2, \dots, J$) en combinación con la categoría de referencia $j=1$ sí;

$$OR_j = \frac{\pi_{jp}}{\pi_{ja}} \bigg/ \frac{\pi_{1p}}{\pi_{1a}}$$

donde π_{jp} y π_{ja} son las probabilidades de la variable respuesta j ($j = 1, \dots, J$), respectivamente. Para el modelo

$$\log\left(\frac{\pi_j}{\pi_1}\right) = \beta_{0j} + \beta_{1j}x, \quad j = 2, \dots, J$$

Donde las probabilidades logarítmicas son:

$$\log\left(\frac{\pi_{ja}}{\pi_{1a}}\right) = \beta_{0j} \text{ cuando } x = 0, \text{ indica que la exposición está ausente}$$

$$\log\left(\frac{\pi_{jp}}{\pi_{1p}}\right) = \beta_{0j} + \beta_{1j}x \text{ cuando } x = 1, \text{ indica que la exposicion esta presente}$$

Por lo tanto, el logaritmo de la razón de posibilidades se puede escribir como,

$$\log OR_j = \log \left(\frac{\pi_{jp}}{\pi_{1p}} \right) - \log \left(\frac{\pi_{ja}}{\pi_{1a}} \right)$$

$$\log OR_j = \beta_{1j}$$

- **ODDS ratio o razón de ventaja**

$$OR_j = e^{\beta_{1j}}$$

En la publicación de Pérez, (2015) sostiene que los OR (Odds Ratio) se pueden entender de una forma sencilla ya que si los valores son mayores a 1 representan el número de veces que se incrementa la posibilidad de presentar un fenómeno si está presente un factor asociado. Ahora si, el OR es menor a 1 representara un factor protector, y finalmente si el OR es igual a 1 indicaría ausencia de relación.

OR con IC<Li-Ls> estrecho ----->N° de sujetos estudiados fue suficiente

OR con IC<Li-Ls> amplio ----->N° de sujetos estudiados fue muy pequeño

TRABAJO INFANTIL		
Sexo	Trabaja (Y=1)	No trabaja (Y=0)
Hombre (X=1)	$P(Y=1/X=1) = \pi = \frac{e^{\beta_0 + \beta_1}}{1 + e^{\beta_0 + \beta_1}}$	$P(Y=0/X=1) = 1 - \pi = \frac{1}{1 + e^{\beta_0 + \beta_1}}$
Mujer (X=0)	$P(Y=1/X=0) = \pi = \frac{e^{\beta_0}}{1 + e^{\beta_0}}$	$P(Y=0/X=0) = 1 - \pi = \frac{e^{\beta_0}}{1 + e^{\beta_0}}$

Figura 5. *Tabla de contingencia del cálculo de OR*

3.5. Metodología y procedimientos

3.5.1. Tipo y diseño estadístico

Este estudio es de tipo aplicada por que busca encontrar los factores de riesgo que inciden al trabajo infantil con un enfoque cuantitativo, con un corte transversal. Además, cuenta con un

diseño Observacional No experimental, debido que se no se manipula deliberada las variables de interés, es decir, en el estudio no se modificó de forma intencional las variables extraídas de la Encuesta Nacional de Hogares (ENAHO) – 2021.

3.5.2. Población

Todos los menores de edad entre 5 a 13 años que residen habitualmente en la vivienda a nivel nacional en el Perú, 2021.

3.5.3. Muestra

La muestra está compuesta por 19,546 niñas, niños y adolescentes que residen habitualmente en la vivienda en el año 2021.

Para la extracción de la muestra se consideró el módulo 1 “Características de la Vivienda y del Hogar”, 2 “Características de los Miembros del Hogar”, 3 “Educación”, 5 “Empleo e Ingresos” y “sumarias” en formato .do (STATA 16), la base de datos de la ENAHO del INEI 2021 se extrajo de: <http://inei.inei.gob.pe/microdatos/>.

3.5.4. Variables de investigación

Tabla 1. Descripción de Variables de Estudio

- *Variable dependiente*

Variable	Código	Categoría	Tipo de Variable
Trabajo Infantil	0	No trabaja	Cualitativa
	1	Trabaja	(Dicotómica)

- *Variable independiente*

Variable	Código	Categoría	Tipo de Variable
Sexo del menor	1	Hombre	Cualitativa Nominal
	2	Mujer	
Área de residencia	1	Urbano	
	2	Rural	
Región Natural	1	Costa	
	2	Sierra	
	3	Selva	
Macro.región	1	Norte	
	2	Centro	
	3	Sur	
	4	Oriente	
	5	Lima Metropolitana	
¿Sabe leer y escribir?	0	No	
	1	Si	
Lengua Materna	1	Quechua	
	2	Castellano	
	3	Otras Lenguas	
¿Cuál.es.el.grado.o.año.de estudio.al.que.asistió.el.año pasado-2020?	1	Ed. Inicial/Ed. Básica	
	2	Primaria	
	3	Secundaria	
Actualmente, ¿ asiste.a.un centro.de.educación.básica o superior ?	1	Si	
	0	No	
¿Cuál es el centro de estudios en el que está matriculado?	1	Estatad	
	2	No estatal	
¿cuál.es.la.relación.de parentesco.con.el.jefe(a).del hogar?	1	hijo(a)/hijastro(a)	
	2	nieto(a)	
	3	otros	
Vivienda inadecuada	1	Vivienda Inadecuada	
	0	Vivienda adecuada	
Tipo de vivienda	1	Alquilada	
	2	Propia	
	3	Usufructuada	
Pobreza	1	Pobre extremo	
	2	Pobre no extremo	
	3	No pobre	
Total de personas en el hogar (Variable agrupada)	1	1-3 miembros	
	2	4-6 miembros	

Nota. Según INEI, Las macroregiones comprenden los siguientes departamentos: Norte: Tumbes, Piura, Lambayeque, Cajamarca, La Libertad. Centro: Lima (excluye provincia de Lima), Ancash, Junín, Cerro de Pasco, Huánuco, Huancavelica, Ayacucho, Ica. Sur: Arequipa, Moquegua, Tacna, Cusco, Madre de Dios, Apurímac, Puno. Oriente: Loreto, Ucayali, Amazonas, San Martín. Lima Metropolitana: Provincia Constitucional del Callao y provincia de Lima).

3.6. Resultados

3.6.1. Análisis descriptivo

- **Análisis univariado**

En esta sección, mostraremos de manera descriptiva la información de las variables de estudio, con el fin de tener conocimiento de la muestra la cual se está trabajando. En la tabla 1, se observa las frecuencias de las NNA de 5 a 13 años según el ámbito sociodemográfico, el cual muestra que, en el año 2021, el 50.8% de la población de 5 a 13 años de edad son hombres.

Según área de residencia, se tiene un mayor porcentaje en el área urbano (61.2%), respecto al área rural (38.9%). Mientras que, en la región natural de la costa se presentó 38.9% de menores de edad, lo mismo la macro región centro (16.7%).

Tabla 2. Menores de 5 a 13 años según variables sociodemográficas, ENAHO 2021

Variable	Característica	Nro. (N=19,546)	%
Sexo	Hombre	9,926	50.8
	Mujer	9,620	49.2
Área de residencia	Urbano	11,968	61.2
	Rural	7,578	38.8
Región natural	Costa	7,600	38.9
	Sierra	6,507	33.3
	Selva	5,439	27.8
Macroregión	Norte	4,190	21.4
	Centro	5,217	26.7
	Sr	3,642	18.6
	Oriente	4,055	20.8
	Lima Metropolitana	2,442	12.5

Según los indicadores de educación (Tabla 2) se obtuvo que el 52% de niños y adolescentes saben leer y escribir. Además, la gran mayoría de NNA tiene como primera lengua materna el castellano (90.8%). También se tiene que en el año 2020 el 65.3% de menores de edad a asistió en nivel primaria en comparación a los estudiantes que asistieron al nivel secundaria (7.5%). Además, se tiene que las instituciones estatales en el 2021 han tenido un mayor número de estudiantes matriculados entre las edades de 5 a 13 años de edad (74.8%).

Tabla 3. Menores de 5 a 13 años según variables de ámbito educativo, ENAHO 2021

Variable	Característica	Nro. (N=19,546)	%
¿Sabe leer y escribir ?	Si	10,169	52.0
	No	5,007	25.6
	Missing (.)	4,370	22.4
Lengua Materna	Quechua	1,091	5.6
	Castellano	17,745	90.8
	Otras lenguas	698	3.6
¿Cuál es el grado.o.año de estudios.al que.asistió.el año.pasado?	Missing (.)	12	0.1
	Inicial/Basica esp.	3,951	20.1
	Primaria	12,749	65.3
	Secundaria	1,457	7.5
¿Asiste actualmente aun.centro.de Estudios básica?	Missing (.)	1,389	7.1
	Si	15,103	90.2
¿Cuál es el centro de estudios en el que está matriculado?	No	1,639	9.8
	Estatal	14,620	74.8
	No estatal	2,121	10.8
	Missing (.)	2,805	14.4

Finalmente se describe las variables relacionadas a la característica del hogar, donde se obtuvo que la relación que tienen los niños/as y adolescentes (Residentes del hogar) con el jefe del hogar suelen ser hijos(as) o hijastros(as) (84.6%). Cabe resaltar que de este grupo de niños el 90.2% se encuentran en viviendas adecuadas. Referente a los niveles de pobreza se observa que hay un mayor grupo de NNA que se encuentran en situación de no pobres (66.2%).

Tabla 4. Frecuencia de NNA según las variables de ámbito del hogar, ENAHO 2

Variable	Característica	Nro. (N=19,546)	%
¿cuál es la relación de parentesco con el jefe(a) del hogar	hijo(a)/hijastro(a)	16,529	84.6
	nieto(a)	2,617	13.4
	otros	400	2.1
Necesidades básicas insatisfechas (vivienda inadecuada)	Vivienda adecuada	17,620	90.2
	Vivienda inadecuada	1,926	9.9
Pobreza	Pobre extremo	1,495	7.6
	Pobre no extremo	5,122	26.2
	No pobre	12,929	66.2

De la figura 5. Se tiene que, en el año 2021 a nivel nacional, el porcentaje de menores de edad entre 5 a 13 años que se encontraron en trabajo infantil (TI) fue 12.2%.

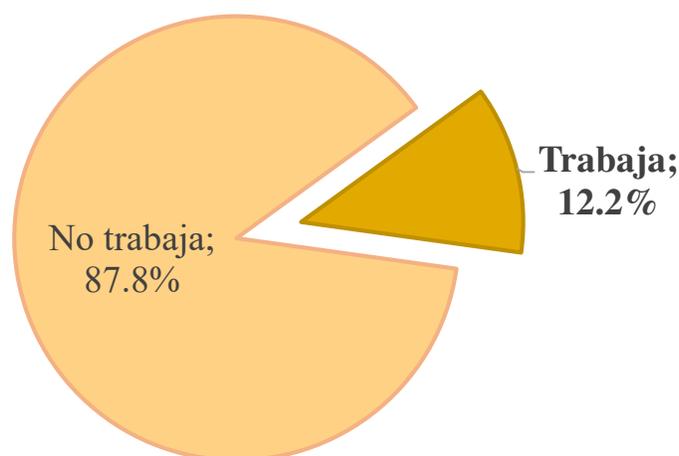


Figura 6. Porcentaje de NNA de 5 a 13 años en TI

- **Análisis bivariado**

A continuación, presentamos la relación existente entre las variables, de la tabla 5, se tiene como la prueba no paramétrica de chi cuadrado (χ^2) calcula la relación entre las variables del factor demográfico y el trabajo infantil. Los niños de edades entre 5 y 13 años que son del área rural tienen un mayor porcentaje en realizar algún trabajo seguido de los que se encuentran en la región natural sierra, al realizar la prueba se obtuvo que la probabilidad es menor al nivel de significancia ($X_{\text{área r.}}^2 = 3138.98, p < 0.05; X_{\text{región.}}^2 = 1641.59, p < 0.05$), con el que se rechaza la hipótesis-nula, y se concluye que tanto el área rural como la región natural se encuentran asociado con el trabajo infantil. Las otras variables del factor demográfico también tuvieron menor probabilidad ($p < 0.05$), con la cual rechazamos la hipótesis nula, se concluye que las variables restantes del factor demográfico también se encuentran asociados al trabajo infantil,

con excepción de la variable sexo del menor de edad ($X^2=0.0294$ $p_valor=0.8872$), lo que indicaría que aparentemente el sexo no tiene asociación al trabajo infantil.

Tabla 5. Evaluación de la asociación de los factores demográficos en el TI

Variable	Característica	Si trabaja	No trabaja	X^2	Sig.
		%	%		
Sexo	Hombre	6.2	44.7	0.0294	0.8872
	Mujer	6.0	43.1		
Área de residencia	Urbano	3.8	73.0	3138.98	0.0000
	Rural	8.4	14.8		
Región natural	Costa	1.8	49.9	1641.59	0.0000
	Sierra	7.8	24.3		
	Selva	2.6	13.7		
Macroregión	Norte	3.7	19.5	991.363	0.0000
	Centro	3.9	16.0		
	Sur	2.7	13.3		
	Oriente	1.6	10.0		
	Lima Metropolitana	0.3	28.9		

Mientras que, en la tabla 6 se observa la prueba de χ^2 entre las variables del factor educación y el trabajo infantil. Donde se tiene que, el 11.8% de menores de edad que realizaron trabajo infantil se encontraron matriculados en alguna institución del estado, mediante la prueba de χ^2 se encontró que todas las variables presentadas se asocian al trabajo infantil a excepción de la variable asistencia al centro de estudios ($X^2=3.2521$ $p_valor=0.2998$), lo que indicaría la ausencia de asociación respecto al trabajo infantil.

Tabla 6. Evaluación de la asociación de los factores de educación respecto al TI

Variable	Característica	Si trabaja	No trabaja	X^2	Sig.
		%	%		
¿Sabe leer y escribir ?	Si	10.9	56.8	80.56	0.0000
	No	3.4	28.8		
Lengua Materna	Quechua			852.98	0.0000
	Castellano	1.9	2.5		
	Otras lenguas	9.8	83.6		

		0.4	1.6		
¿Cuál es el grado o año de estudios al que asistió el año pasado?	Secundaria	0.5	7.8	297.77	0.0000
	Primaria	11.1	58.8		
	Educación inicial	1.5	20.4		
¿Asiste actualmente aun centro de Estudios básica?	Si	10.7	79.2	3.2521	0.2998
	No	1.4	8.7		
¿Cuál es el centro de estudios en el que está matriculado?	Estatal	11.8	72.0	318.57	0.0000
	No estatal	0.3	15.9		

De la tabla 7 se observa la prueba de χ^2 entre las variables del factor hogar y el trabajo infantil. De este se tiene, que todas las variables presentadas se asocian al trabajo infantil a excepción de la variable necesidades básicas insatisfechas ($X^2=3.6049$ p_valor=0.1541) y la variable miembros del hogar, lo que indicaría la ausencia de asociación respecto al trabajo infantil ($X^2=3.8965$ p_valor=0.3324).

Variable	Característica	Si trabaja		No trabaja	
		%	%	X ²	Sig.
¿cuál es la relación de parentesco con el jefe(a) del hogar	hijo(a)/hijastro(a)	11.1	72.7	100.22	0.0000
	nieto(a)	1.1	13.0		
	otros	0.1	2.1		
Necesidades básicas insatisfecha (vivienda inadecuada)	Vivienda adecuada	11.1	81.1	3.6049	0.1541
	Vivienda inadecuada	1.1	6.7		
vhogarocu	alquilada	0.5	10.1		

	Propia	10.0	61.1	190.4903	0.0000
	usufructuada	1.6	16.7		
Pobreza	Pobre extremo	1.7	5.1		
	Pobre no extremo	4.5	25.2	349.4749	0.0000
	No pobre	6.0	57.5		
Total de miembros del hogar	1-3	2.1	16.2		
	4-6	8.3	59.6	3.8965	0.3324
	>6	1.8	12.0		

Tabla 7. Evaluación de la asociación de los factores del hogar

3.6.2. Análisis del modelo de regresión logística múltiple

Se analizará el modelo ajustado incluyendo las 14 variables de estudio que desde un inicio planteamos, Cabe resaltar que estaremos considerando las estimaciones ponderadas, para eso utilizaremos en Stata 16 el comando svy: logit, el cual nos permitirá obtener el modelo ajustado, esto nos permitirá evaluar si al menos unas de las variables aportan significativamente al modelo.

De la tabla 8. Muestra como al menos una de la variable propuesta en el modelo es significativa ($X^2=70.31$; $P_valor=0.000$); se tiene que todas las variables ingresadas al modelo ajustado aportan significativamente al modelo planteado con excepción a la variable tipo de vivienda ($X^2=1.40$; $p_valor=0.161$).

Tabla 8. Modelo de regresión logística múltiple ajustado N°1

Trab_inf	B	S. E.	t	P > t 	95% IC	
sexo	.162983	.87385	-2.21	0.027	-.3077691	-.018197
Área residencia	1.831772	.1067845	17.15	0.000	1.622416	2.041128
Región Natural	.5819876	.0731922	7.95	0.000	.4384113	.7254039
Macroregión	.3593832	.8434474	-8.27	0.000	-.444484	-.2741228
¿Sabe. leer y escribir	.2467381	.1161291	-2.12	0.034	-.474406	-.0190541

Lengua materna	.8777941	.1294861	-6.78	0.000	-1.131657	-.6239312
Niveles de estudios	1.050281	.1220071	8.58	0.000	.8102971	1.290266
A. asiste. centroestudi	.478514	.1294945	3.70	0.000	.2246346	.7323933
Tipo Centro estudio	.9297824	.2541817	-3.66	0.000	-1.428116	-.4314486
Relación parentesco	.2862848	.100853	-2.84	0.005	-.4840112	-.0885584
Tipo de Vivienda	.3765847	.1393792	-2.70	0.007	-.6497635	-.1032459
Vivienda inadecuada	.1248673	.0891039	-1.40	0.161	-.0498245	.2995591
pobreza	-.1723005	.060127	-2.87	0.004	-.290182	-.0544191
Grupo.Miembros.hogar	-.1271284	.0645241	-1.97	0.049	-.2536226	-.0006182
_cons	-2.878794	.6336648	-4.54	0.000	-4.121121	-1.636468

($X^2=68.94$; $p_valor=0.000$).

A continuación, extraeremos la variable el cual no aporta significativamente al modelo planteado y luego por el método de Akaike y BIC realizaremos una comparación de cuál sería el mejor modelo.

De la tabla 9 se observa que nuestras 13 variables estudiadas son significativas, a partir de este análisis, de igual forma estaremos evaluando la parsimonia de los dos modelos estudiados, para ello haremos uso de los criterios de información.

Tabla 9. Modelo de regresión logística múltiple ajustado N°2

Trab_inf	B	S. E.	t	P > t	95% IC	
sexo	-.1635755	.0737944	-2.22	0.027	-.3082523	-.0188986
Área residencia	1.836896	.1066855	17.22	0.000	1.627735	2.046057
Región Natural	.5757992	.073627	7.82	0.000	.4314505	.720148

Macroregión	-.3588412	.0435583	-8.24	0.000	-.444239	.2734433
¿Sabe. leer y escribir	-.2442949	.1155568	-2.11	0.035	-.4708488	-.0177409
Lengua materna	-.8735918	.1296195	-6.74	0.000	-1.127716	-.6194675
Niveles de estudios	1.049546	.1221687	8.59	0.000	.8100296	1.289063
A. asiste. centroestudi	.4712575	.1289566	3.65	0.000	.2184328	.7240823
Tipo Centro estudio	-.928308	.2548246	-3.65	0.000	-1.426334	-.4302824
Relación parentesco	-.289487	.1008919	-2.87	0.004	.4872897	-.0916842
Vivienda inadecuada	-.3734112	.1398537	-2.69	0.007	-.6460318	-.1007906
pobreza	-.1753612	.0683857	-2.90	0.004	-.2937499	-.0569726
Grupo.Miembros.hogar	-.1327208	.0640353	-2.07	0.038	-.2582687	-.007181
_cons	-2.595657	.6581106	-3.99	0.000	-3.870225	-1.321088

($X^2=74.13$; $p_valor=0.000$).

De los criterios de información de akaike y el criterio de Criterio de Información Bayesiano (BIC) o criterio Schwarz (SIC), para obtener un adecuado modelo nos fijaremos en la parsimonia y en que en ambos coeficientes sean los mínimos.

De la tabla 10 se tiene que el modelo más parsimonioso que nos permitirá predecir correctamente será cuando tomen los valores siguientes: $aic=9313.6464$ y $bic=9418.394$.

Tabla 10. Selección del modelo

Criterios de Información	Modelo 1 (14 variables)	Modelo 2 (13 variables)
AIC	9313.6379	9313.6464
BIC	9425.8525	9418.394

3.6.3. Interpretación de los parámetros

Luego de ajustar el modelo de regresión logística múltiple se observa lo siguiente:

- El ser varón es 1.1719 veces más probable de que realice trabajo infantil, respecto a las mujeres (OR=1.172; IC<1.013-1.355>).
- Un niño, niña y adolescente que tenga área de residencia rural tiene 5.6040 veces más posibilidad que realice alguna actividad económica mayor o igual a una hora, frente aquellos que provienen de zonas urbanas (OR=1.172; IC<1.013-1.355>).
- Pertenecer a la región natural “sierra” tiene mayor riesgo de que un NNA se encuentre realizando trabajo infantil, respecto a los que son de la región costa (OR=3.0836; IC<2.323-4.09>).
- Las regiones que pertenecen a la Macroregión Norte (Tumbes, Piura, Lambayeque, Cajamarca, La libertad) presenta mayor riesgo de encontrar a menores de edad realizando trabajo infantil, respecto a los menores que habitan en la macroregión Centro (OR=1.5086; IC<1.186-1.918>).
- El en 2021 muchas de las actividades y trabajos realizados por algún niño fueron a través de un computador ya sea con o sin conexión a internet.
- Según INEI, en el cuarto trimestre del año **2021**, el 86,7% de las niñas, **niños** y adolescentes entre los 6 y 17 años de edad utilizaron **internet**.
- Hoy en día un niño que sabe leer y escribir tiene 1.2620 veces más posibilidades de realizar algún trabajo infantil, frente a un NNA que es analfabeto (OR=3.0836; IC<2.323-4.093>).
- La lengua materna quechua que aprendió el menor de edad tiene mayor repercusión en que un niño se encuentre realizando alguna actividad económica

mayor o igual a 1 hora, frente a aquellos que tienen como lengua materna el castellano (OR=2.1944; IC<1.634-2.946>).

- En el año 2020 la inmovilización social hizo que muchos menores de edad continúen estudiando desde casa, pero no todos gozaron de estos beneficios; por lo que los niños que tuvieron como último grado de estudios los niveles primarios tuvieron 2.896438 veces la posibilidad de realizar alguna actividad económica, con respecto a los menores que asistieron al nivel inicial o educación especial (OR=2.896438; IC<2.273791-3.689589>).
- Un niño que no asistió a algún centro de estudios en el 2021 tienen un mayor riesgo de encontrarse trabajando, respecto a los que si asistieron (OR=1.55198; IC<1.198-2.011>).
- Encontrarse estudiando en un centro de educación estatal es 2.37022 veces más probable de encontrar a un menor trabajando, frente a aquellos que proceden de instituciones no estatales (OR=2.37022; IC<1.433-3.922>).
- Un niño que tiene relación de parentesco con el jefe de hogar distinto a ser hijo(a)/hijastra(o) tiene 0.38246 veces más chance de que realice algún trabajo, frente aquellos que tienen un parentesco de “nietos”.
- Un menor de edad pobre no extremo tiene 1.30253 veces más chance de que realice trabajo infantil, respecto a los menores que pertenecen al grupo de pobre extremo (OR=1.30253; IC<0.736-1.217>).
- Tener más de 6 miembros en el hogar tiene mayor riesgo de que un menor perteneciente a ese hogar realice trabajo infantil, respecto a los hogares que

cuentan con menor número de miembros menor a 4 (OR=0.7531; IC<0.579-0.980>).

Tabla 11. Factores de riesgo que inciden al trabajo infantil

Factores	Características	B	Sig.	OR	95% de intervalo de confianza para OR	
					Inferior	Superior
	_constante	-4.324328	0.000	-5.1840	-5.184	-3.465
	Sexo					
	Hombre	.1586528	0.033	1.1719	1.013	1.355
	Área residencia					
	Rural	1.723474	0.000	5.6040	4.537	6.922
	Región Natural					
	Sierra	1.126084	0.000	3.0836	2.323	4.093
	Selva	.911826	0.000	2.4889	1.713	3.616
	Macro región					
	Norte	.411197	0.001	1.5086	1.186	1.918
	Sur	-.138615	0.278	0.8706	0.678	1.118
	Oriente	-.3435566	0.036	0.7092	0.514	0.977
	Lima metropo.	-1.326058	0.001	0.2655	0.118	0.597
	¿Sabe leer y escribir?					
	Si	.2334682	0.042	1.2620	1.008	1.582
	Lengua materna					
	Quechua	.7858976	0.000	2.1944	1.634	2.946
	Otra Lengua N.	-.1470732	0.463	0.8632	0.583	1.279
	Ultimo Nivel.de.estudios que.asistió.el.año.pasado					
	Primaria	.9571008	0.000	2.8964	2.273	3.690
	Actualmente asiste a algún centro de estudios					
	No	.4395305	0.001	1.55198	1.198	2.011
	Tipo Centro estudio					
	Estatat	.862984	0.001	2.37022	1.433	3.922
	Relación.parentesco.con el.feje.del.hogar					
	hijo(a)/hijastro(a)	.1177709	0.360	1.12499	0.874	1.445
	otros	-.9611333	0.007	0.38246	0.189	0.772
	Estado de la vivienda					
	vivienda adecuada	.1912396	0.153	1.21073	0.931	1.574
	pobreza					
	pobre no extremo	.2643054	0.040	1.30253	1.012	1.675
	no pobre	-.0550464	0.668	0.94644	0.736	1.217
	Miembros.hogar (por grupo)					
	4-6 personas	.0453709	0.659	1.0464	0.855	1.280
	>6 a personas	-.2835584	0.035	0.7531	0.579	0.980

($X^2=46.88$; $p_valor=0.0000$).

3.6.4. Evaluación de la capacidad predictiva del modelo

Para obtener un punto de corte conveniente donde se presenta la presencia de trabajo infantil en los NNA de 5 a 13 años, nos basaremos en la sensibilidad y especificidad que se tienen en tabla 12, donde hallaremos el punto que maximiza ambos indicadores.

Tabla 12. *Sensibilidad, especificidad y porcentaje de clasificación*

Punto de corte (p.c.)	Sensibilidad	Especificidad	Accuracy
0.50	46.53%	62.70%	60.73%
0.20	75.42%	56.20%	58.54%
0.10	88.23%	45.41%	50.63%

Como observamos en la tabla 12, el p.c. es 0.10. La sensibilidad obtenida fue de 88.23%. Es decir, el 50.63% de NNA en trabajo infantil fueron predichos correctamente por el modelo ajustado y que a pesar que no tiene un alto poder clasificatorio, nuestro modelo cuenta con un buen potencial predictivo.

Contando con una sensibilidad de 88.23 y una especificidad de 45.41 contamos con un alto poder predictivo, aunque aparentemente no se muestre esos resultados, se realizara una tabla de 2X2 en el cual se probara que el punto de corte es adecuado.

De la tabla 13 se tiene que del total de niños, niñas y adolescentes que trabajan del grupo edad de 5 a 13 años (12.2%) se ha predicho correctamente 312,572 niños, siendo un total de 5,510,027 niños en total que trabajan.

Lo que mostraría que nuestro modelo ha predicho correctamente a nuestra población de estudio.

Tabla 13. *Tabla de comparación de las Predicciones estimadas del trabajo infantil*

Trabajo Infantil	Trabajo Infantil Predicho		Total
	0	1	
	2,196,877	2,641,329	4838206
0	96.52%	81.67%	87.81%
	79,090	592,731	671,820
1	3.48%	18.33%	12.19%
Total	2,275,967	3,234,060	5,510,027
	100%	100%	100%

3.6.5. Evaluación de la bondad de ajuste

Como muestran los resultados de la prueba de Hosmer-Lemeshow, las estadísticas de la prueba de χ^2 fueron 7.99 con un valor de p_valor de 0.4343. Por lo tanto, el modelo de estudio se ajustaba bastante bien. Este hecho confirma que los datos se ajustaron bien por el modelo estimado.

Tabla 14. *Bondad de ajuste del modelo*

Número de observaciones	Número de Grupos	Hosmer-Lemeshow Chi-cuadrado χ^2	Sig.
19,546	10	14.63	0.0667

IV. CONCLUSIONES

El presente Trabajo de Suficiencia aborda la identificación de factores de riesgo que inciden al trabajo infantil, de todos los menores de 5 a 13 años edad que sean residentes

habituales y que hayan realizado al menos una hora a la semana a una o más actividades económicas para el grupo de 5 a 11 años y aquellos adolescentes en trabajo intensivo en horas de 12 a 13 años. Extraídos de la ENAHO 2021.

Para lograr este objetivo usamos el modelo de regresión logística múltiple y de allí se concluye que; al menos uno de los factores como el sexo, área de residencia, la región natural, macroregiones, el saber leer y escribir, lengua materna, ultimo nivel de estudios cursado en el 2020, actualmente asiste a algún centro de estudios, el tipo de institución , relación de parentesco con el jefe del hogar, el nivel de pobreza y el N° de miembros del hogar (por grupos), explica significativamente la variabilidad del trabajo infantil en menores de 5 a 13 años ($X^2= 46.88$; $p_valor=0.0000$), mientras que, se pudo rescatar que la variable “estado de la vivienda” no aporta significativamente al modelo planteado.

Finalmente, cuando se realizó el análisis de los factores de riesgos que inciden al trabajo infantil en el año 2021, se pudo apreciar que los factores sociodemográficos que resultaron ser significativos fueron los siguientes:

VARIABLES RELACIONADAS A LOS FACTORES SOCIODEMOGRÁFICOS:

- El sexo masculino del NNA
- El área de residencia Rural que procede el menor
- La región Natural de la sierra que se encuentra el niño/a
- El pertenecer a la macroregión del norte

VARIABLES RELACIONADAS A LOS FACTORES DE EDUCACIÓN:

- El ser alfabeto
- Tener como lengua materna el quechua

- Haber asistido en el último año (2020) al nivel primario
- Si “NO” asiste actualmente a algún centro educativo
- El tipo de institución educativa estatal

VARIABLES RELACIONADAS A FACTORES DE “CARACTERÍSTICAS DEL HOGAR”:

- El contar con otro tipo de relación con el jefe del hogar
- El nivel de pobreza de pobre no extremo
- El número de miembros del hogar mayor a 6 personas

V. RECOMENDACIONES

A partir de lo anterior, se sugiere: considerar los factores de riesgo que inciden al trabajo infantil a temprana edad ya que este permitirá evaluar diversas políticas públicas y generar programas que fortalezcan el bienestar de un niño, niña y adolescente.

Cabe resaltar que en los últimos años la evaluación del indicador del trabajo infantil en menores de 5 a 13 años de edad no ha abarcado muchas investigaciones, es por ello que se recomienda evaluarlo según la metodología de la resolución ministerial N° 114-2016-TR, ya que permitirá centrarnos en nuestro público objetivo el cual es menores en trabajo infantil de edad entre 5 a 13 años.

Si realizamos una evaluación entre las actividades económicas y el trabajo infantil que realiza un menor de 5 a 13 años de edad se tiene que: en el 2020 este indicador fue 20.1% y 14.24%, respectivamente y para el 2021 fue 17% y 12.2%, respectivamente. Con esta información podemos concluir que evaluar la actividad económica no reflejara con precisión la realidad del trabajo infantil. Además, conocer los factores de riesgo que inciden que un menor trabaje a temprana edad no implica que sean los mismos para un niño que realiza alguna actividad económica, es por eso que se aconseja estudiar el trabajo infantil para plasmar la situación actual que enfrenta muchos menores y abarcar solo a este grupo de niños para tener un mejor resultado en los estudios posteriores relacionados al trabajo infantil. Por ello se aconseja para posteriores estudios considerar la opción de realizar una regresión logística binaria de efectos mixtos respecto a las macro regiones para que se pueda medir más específicamente los efectos de los factores que inciden al trabajo infantil en los grupos de regiones.

VI. BIBLIOGRAFIA

- Abdallah, A. S. (2020). Using logistic regression to identify the factors affecting chil labor in Red Sea State. *IJAAS International journal of ADVANCED AND APPLIED SCIENCES*, 7, 12-19. <https://doi.org/https://doi.org/10.21833/ijaas.2020.10.002>
- Aguayo Castañón, P. J., Figueredo Arratea, J. H., Ito Calderón, A. D., Machaca Huamán , D. Á., Silva Ortecho, A. M., & Ucedo Silva, V. H. (2019). El trabajo infantil: La relación de los factores sociodemográficos, adaptación académica y la actividad laboral con la repitencia en niños y adolescentes de 10-17 años en los departamentos de Lima y Huánuco. *Revista Psicológica Herediana*, 1(12), 21-25. <https://doi.org/https://doi.org/10.20453/rph.v12i1.3632>
- Dobson AM, A. J., & Barnett , A. G. (2018). *An Introduction to Generalized Linear Models* (4 ed.). (J. Dobson, & A. G. Barnett Fassa, Trads.) Taylor & Francis Group, LLC. <https://doi.org/https://doi.org/10.1201/9781315182780>
- Herrero, V., & Bocco, M. (2007). Comparación de Ponderaciones en Regresiones Probit Simultáneas en un Modelo para la Estimación de la Participación Laboral. *Revista de Economía y Estadística*, 45(2), 95-124. <https://doi.org/https://doi.org/10.55444/2451.7321.2007.v45.n2.3841>
- INEI, MTP, & OTI. (2016). *Magnitud y Características del trabajo infantil en Perú:Informe de 2015--Análisis de la Encuesta Nacional de Hogares (ENAHO) y de Organización Internacional del trabajo*. https://www.ilo.org/ipec/Informationresources/WCMS_IPEC_PUB_28857/lang--es/index.htm

Libaque Saenz, C. F., Lazo, J., Lopez, K. G., & R. Bravo, E. (2017). Predicting child labor in

Peru: A comparison of logistic regression and neural networks techniques. En J. A.

Lossio Ventura, & H. Alatrística Salas (Ed.). Lima: CEUR Workshop Proceedings.

<http://ceur-ws.org/Vol-2029/paper5.pdf>

Liu, X. (2016). *Applied Ordinal Logistic Regression Using Stata: From Single-Level to*

Multilevel Modeling. California: SAGE Publications.

MINTRA. (s.f.). *MINTRA*. [http://www.mintra.gob.pe/archivos/file/SNIL/normas/2016-06-](http://www.mintra.gob.pe/archivos/file/SNIL/normas/2016-06-12_114-2016-TR_4724.pdf)

[12_114-2016-TR_4724.pdf](http://www.mintra.gob.pe/archivos/file/SNIL/normas/2016-06-12_114-2016-TR_4724.pdf)

Pérez Atanasio, J. M. (2015). Principios de bioestadística para comprender los resultados. *José*

Manuel Pérez Atanasio, 11(2), 91-95. <http://www.medigraphic.com/orthotips>

Pinedo Escalante, J. N., & Condori Yañez, V. R. (2019). INCIDENCIA DE LA SITUACIÓN

SOCIOECONÓMICA EN EL TRABAJO INFANTIL DE LA REGIÓN DEL CUSCO

EN EL AÑO 2018. *Tesis*. Universidad Andina del Cusco, Cusco, Cusco, Perú.

<https://hdl.handle.net/20.500.12557/3658>

Thiong'o, J. M. (November de 2020). *Modeling The Key Determinant of child Labour In Kenya*.

University of Nairobi Digital Repository:

[http://erepository.uonbi.ac.ke/bitstream/handle/11295/154374/Thiong'o_Modeling%20th](http://erepository.uonbi.ac.ke/bitstream/handle/11295/154374/Thiong'o_Modeling%20the%20Key%20Determinant%20of%20Child%20Labour%20in%20Kenya.pdf?sequence=1%22%20data-ved=%22ahUKEwj6p86dqPf3AhUNrZUCHTpGDPkQFnoECAcQAQ%22%20ping=%22/url?sa=t&source=web&a)

[e%20Key%20Determinant%20of%20Child%20Labour%20in%20Kenya.pdf?sequence=](http://erepository.uonbi.ac.ke/bitstream/handle/11295/154374/Thiong'o_Modeling%20the%20Key%20Determinant%20of%20Child%20Labour%20in%20Kenya.pdf?sequence=1%22%20data-ved=%22ahUKEwj6p86dqPf3AhUNrZUCHTpGDPkQFnoECAcQAQ%22%20ping=%22/url?sa=t&source=web&a)

[1%22%20data-](http://erepository.uonbi.ac.ke/bitstream/handle/11295/154374/Thiong'o_Modeling%20the%20Key%20Determinant%20of%20Child%20Labour%20in%20Kenya.pdf?sequence=1%22%20data-ved=%22ahUKEwj6p86dqPf3AhUNrZUCHTpGDPkQFnoECAcQAQ%22%20ping=%22/url?sa=t&source=web&a)

[ved=%22ahUKEwj6p86dqPf3AhUNrZUCHTpGDPkQFnoECAcQAQ%22%20ping=%](http://erepository.uonbi.ac.ke/bitstream/handle/11295/154374/Thiong'o_Modeling%20the%20Key%20Determinant%20of%20Child%20Labour%20in%20Kenya.pdf?sequence=1%22%20data-ved=%22ahUKEwj6p86dqPf3AhUNrZUCHTpGDPkQFnoECAcQAQ%22%20ping=%22/url?sa=t&source=web&a)

[22/url?sa=t&source=web&a](http://erepository.uonbi.ac.ke/bitstream/handle/11295/154374/Thiong'o_Modeling%20the%20Key%20Determinant%20of%20Child%20Labour%20in%20Kenya.pdf?sequence=1%22%20data-ved=%22ahUKEwj6p86dqPf3AhUNrZUCHTpGDPkQFnoECAcQAQ%22%20ping=%22/url?sa=t&source=web&a)

Villarruel Meythaler, R. E., Arguello Rojas, C. M., & Doicela Arguello, B. R. (s.f.). Factores que inciden en el trabajo infantil, Ecuador año 2017. *Titulación - Estadística*. Universidad Central del Ecuador, Quito. <http://www.dspace.uce.edu.ec/handle/25000/19273>

VII. ANEXOS

GLOSARIO

MIMP: Ministerio de la Mujer y Poblaciones Vulnerables

DGNNA: Dirección General de Niñas, Niños y Adolescentes

DPNNA: Dirección de Políticas de Niñas, Niños y Adolescentes

MINEDU: Ministerio de Educación

MINSA: Ministerio de Salud

INEI: Instituto Nacional de Estadística e Informática

ETI: Encuesta Nacional Especializada de Trabajo Infantil

ENAHO: Encuesta Nacional de hogares

PNMNNA: Política Nacional Multisectorial para las niñas, niños y Adolescentes

PNAIA: Plan Nacional de Acción por la Infancia y Adolescencia

NNA: Niñas, Niños y Adolescentes

TI: Trabajo Infantil

OIT: Organización Internacional del trabajo

Anexo 1. Objetivos prioritarios e indicadores de la PNMNNA

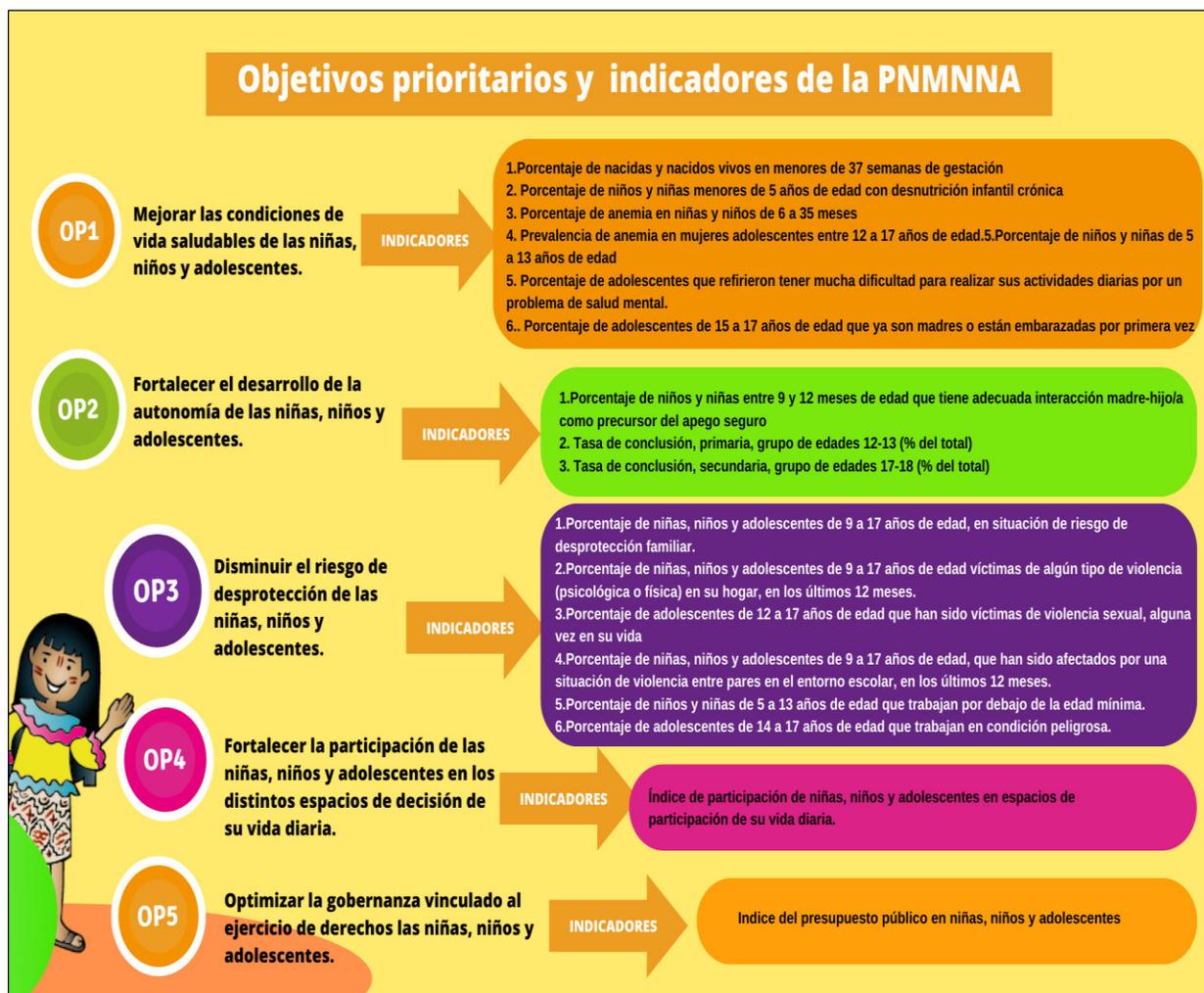


Figura 7. *Objetivos prioritarios e indicadores de la PNMNNA*

Fuente: Extraído de: <https://www.mimp.gob.pe/direcciones/dgna/contenidos/articulos.php?codigo=55>